UNIVERSIDAD DE GRANADA

Departamento de Economía Aplicada

UN NUEVO MÉTODO PARA GENERAR DISTRIBUCIONES DE PROBABILIDAD. PROBLEMAS ASOCIADOS Y APLICACIONES.

MEMORIA presentada por José Callejón Céspedes para optar al grado de Doctor.

Fdo. José Callejón Céspedes

V°B° de los Directores de la Tesis:

Fdo. Rafael Herrerías Pleguezuelo
Catedrático de Universidad
Economía Aplicada

Fdo. Federico Palacios González.
Catedrático de Escuela Universitaria
Economía Aplicada

Granada, diciembre de 1994.

A mi esposa e hijos.

Agradecimiento

Por la ilusión que desde el primer momento me infundieron, el ánimo y el apoyo continuo y constante que de ellos he recibido, la paciencia que tienen ante las situaciones menos favorables y, sobre todo, por su amistad, de la que me honro, me resulta difícil expresar en unas pocas líneas mi más sincero agradecimiento, a los Profesores **D. Rafael Herrerías Pleguezuelo** y **D. Federico Palacios González,** directores de este trabajo.

Agradecimiento, porque sus conocimientos y sus esfuerzos los he tenido siempre que los he solicitado, porque me han dedicado tiempo de su trabajo y parte del tiempo que tenían previsto para su ocio.

No, no me olvido de ninguno de los Profesores que forman el grupo de Métodos Cuantitativos, de la Facultad de Ciencias Económicas y Empresariales, de la Universidad de Granada, por las muchas veces que me han escuchado y apoyado; de una forma u otra, todos ellos han colaborado para que este trabajo se haga una realidad.

CONTENIDO

3. DISTRIBUCIONES MULTIVARIANTES

4. APLICACIONES Y PROBLEMAS ASOCIADOS

INTRODUCCIÓN

En la introducción inicial queremos fijar las líneas generales de este trabajo, dejando para el resumen previo de cada capítulo el comentario de los objetivos y logros correspondientes.

La presente memoria se divide en cuatro capítulos, dedicados, los tres primeros a la generación de distribuciones de probabilidad univariantes, bivariantes y multivariantes, tanto continuas como discretas y el cuarto capítulo está dedicado a las aplicaciones y problemas estadísticos asociados. Al comienzo se establece el concepto amplio de función generadora y en concreto de generadora de una distribución de probabilidad. Se busca en todo caso un paralelismo entre ellos.

A continuación vamos a señalar las etapas por las que ha discurrido el trabajo que presentamos:

El primer capítulo está dedicado al estudio de las funciones generadoras de distribuciones univariantes de probabilidad, tanto continuas como discretas, cuyos antecedentes y logros pasamos a comentar:

Distribuciones univariantes continuas

Una forma de abordar el estudio global y conjunto de un gran número de distribuciones de probabilidades univariantes, se consigue con la formulación de sistemas de distribuciones que verifican una determinada ecuación funcional, bien diferencial para las distribuciones de tipo continuo, bien en diferencias finitas para las distribuciones de tipo discreto.

Los sistemas de distribuciones continuos univariantes más estudiados por sus aplicaciones son el de Pearson (1895), [35], y la familia exponencial, Loève [32], Zacks [48]. El primero de ellos responde a la ecuación diferencial lineal de primer orden y homogénea, Cansado [6], Elderton, [12] y Jonhson-Kotz, [26] :

$$\frac{f'(x)}{f(x)} = \frac{x - a}{b_0 + b_1 x + b_2 x^2} \qquad (1)$$

y el segundo surge de otra ecuación diferencial, derivada de la igualdad en la cota de Cramer - Rao cuando se alcanza la mínima varianza por el estimador T:

$$\frac{\partial \ln L(\vec{x},\theta)}{\partial \theta} = A(\theta)\ (T-\theta)$$

por lo que la función de densidad, [3], será del tipo:

$$f(x,\theta) = C(\theta)\ e^{B(\theta)T(x)}\ h(x) \qquad (2)$$

Evidentemente no todas las distribuciones, soluciones de (1) pueden expresarse como (2), y recíprocamente (2) no es solución de (1) en general, aunque existen distribuciones que satisfacen a (1) y a (2) simultáneamente, por ejemplo la familia $N(\mu,\sigma)$.

La relación entre las familias definidas por (1) y (2) puede verse, con todo detalle, en Herrerías (1975) [19]

Las distribuciones de probabilidad que generan (1) y (2) cubren la mayoría de las necesidades que sobre modelos probabilísticos continuos puedan presentarse. No obstante en la literatura especializada se encuentran otros sistemas que extienden y generalizan las distribuciones de (1) y (2); para nuestros propósitos posteriores citaremos dos ejemplos significativos: la extensión de L.K. Roy, (1971), [42] y la generalización de Herrerías, (1975), [19].

La primera se obtiene cambiando el segundo miembro de (1), que pasa a ser de la siguiente forma:

$$\frac{D\,f(x)}{f(x)} = \frac{f'(x)}{f(x)} = \frac{a_0 + a_1 x + a_2 x^2}{x\,(\,b_0 + b_1 x + b_2 x^2\,)} \tag{3}$$

y la segunda se obtiene cambiando el primer miembro, al utilizar el operador aleph, derivada generalizada, Aldanondo [1], en vez del operador derivada ordinaria

$$a_f^g[y(x)] = \frac{y'(x) - g(x)\,y(x)}{f(x)}$$

(Obsérvese que el operador derivada ordinaria D, es a_1^0).

Los sistemas obtenidos a partir de la generalización de Herrerías responden a la ecuación diferencial:

$$\frac{a_1^\varphi f(x)}{f(x)} = \frac{x - a}{b_0 + b_1 x + b_2 x^2}$$

es decir:

$$\frac{f'(x)}{f(x)} - \varphi(x) = \frac{x - a}{b_0 + b_1 x + b_2 x^2}$$

que se puede escribir

$$\frac{f'(x)}{f(x)} = \frac{x - a}{b_0 + b_1 x + b_2 x^2} + \varphi(x) \tag{4}$$

En [19] se hace notar que si $\varphi(x) = 1/x$, entonces (4) se transforma en (3), por lo que las nuevas distribuciones de (3), que no son de (1), son un caso particular de (4).

Esta expresión, (4), es la inspiradora de un nuevo concepto, el de función generadora para una distribución univariante continua, que será la herramienta básica de esta Memoria.

Manteniendo el primer miembro de la ecuación (4), se sustituye el segundo por una función g(x) a la que sólo se le exigen las condiciones necesarias y suficientes para que f(x) sea una función de densidad de una variable aleatoria, sobre un recinto determinado. Aparece de esta forma el concepto de función generadora de una distribución univariante continua.

Distribuciones univariantes discretas

Cuando se traslada la idea de función generadora al caso discreto, siguiendo el paralelismo ya establecido entre las familias de Pearson y de Ord, [34], univariantes, encontramos los precedentes en la extensión realizada por Herrerías (1976), [20], y la generalización realizada por Fajardo (1985), [13]. Este último estudio será nuestro punto de partida para las distribuciones discretas univariantes.

Lo más relevante que hemos obtenido en este primer capítulo ha sido:

* Establecer las condiciones que debe cumplir una función real de variable real, para que sea generadora de una distribución de probabilidad.
* Establecer los conceptos de función generadora de una distribución de probabilidad univariante tanto para el caso continuo como discreto.
* Describir el procedimiento de la mencionada generación.
* Dejar constancia del paralelismo que el concepto de función generadora establece entre los casos continuo y discreto.

El segundo capítulo está dedicado al estudio de los sistemas de funciones generadoras de una distribución bivariante de probabilidad. Se abordan también los casos continuo y discreto. En el último epígrafe de este capítulo se establecen condiciones necesarias y suficientes sobre las funciones que forman el sistema de generadoras para que dos variables aleatorias sean independientes.

Distribuciones bivariantes.

El caso continuo lo hemos desarrollado tomando como punto de partida la familia pearsoniana bivariante, definida por van Uven, (1947), [44]:

$$\begin{cases} \dfrac{\partial}{\partial x} \ln f(x,y) = \dfrac{L_1(x,y)}{Q_1(x,y)} \\ \\ \dfrac{\partial}{\partial y} \ln f(x,y) = \dfrac{L_2(x,y)}{Q_2(x,y)} \end{cases}$$

donde L_1 y L_2 son formas lineales y Q_1 y Q_2 son formas cuadráticas, y donde además se ha de verificar la igualdad de Schwartz para la integrabilidad del sistema

$$\frac{\partial^2}{\partial x \partial y} \ln f(x,y) = \frac{\partial}{\partial y}\frac{L_1(x,y)}{Q_1(x,y)} = \frac{\partial}{\partial x}\frac{L_2(x,y)}{Q_2(x,y)}$$

Fernández (1979), [16], extiende el sistema de van Uven, utilizando funciones cuadráticas y cúbicas para el numerador y el denominador, respectivamente, del segundo miembro de las ecuaciones que definen el sistema.

El sistema de distribuciones bivariantes discretas de tipo Pearson, análogo al estudiado por van Uven en el caso continuo, es estudiado por Herrerías - Cobos (1984) [21] y Herrerías - Calvete, (1985) [22] y (1986) [23], planteando el sistema de ecuaciones en diferencias finitas que lo define,

$$\begin{cases} \dfrac{\Delta_1 f_{r-1,s}}{f_{r-1,s}} = \dfrac{a_0 + a_1 r + a_2 s}{g_0 + g_1 r + g_2 s + g_3 r(r-1) + g_4 s(s-1) + g_5 rs} \\ \\ \dfrac{\Delta_2 f_{r,s-1}}{f_{r,s-1}} = \dfrac{b_0 + b_1 r + b_2 s}{h_0 + h_1 r + h_2 s + h_3 r(r-1) + h_4 s(s-1) + h_5 rs} \end{cases}$$

mostrando distribuciones que pertenecen a él, y encontrando una solución general basada en las relaciones de recurrencia de los momentos factoriales descendentes.

Aunque la generalización del sistema de Pearson discreto bivariante la realiza Fajardo (1985) [13], en la presente memoria se define el concepto de sistema de funciones generadoras, como un conjunto de dos funciones generadoras de distribuciones univariantes, una para cada distribución condicionada, que verifican las propiedades necesarias para la generación de la distribución conjunta.

En ambos casos, continuo y discreto, establecen las condiciones que debe cumplir un conjunto de dos funciones reales de variable real, para que sea un sistema de funciones generadoras de una distribución bivariante de probabilidad. Estas funciones, cada una por separado, son las funciones generadoras de las respectivas distribuciones condicionadas y, por tanto, se obtiene la distribución conjunta de probabilidad a partir de las funciones generadoras de las distribuciones condicionadas y sin el conocimiento previo de las distribuciones marginales.

Independencia de variables aleatorias.

El último epígrafe del capítulo de distribuciones bivariantes se dedica al estudio de independencia de dos variables aleatorias establecida sobre condiciones necesarias y suficientes en los sistemas de generadoras. Destacamos la facilidad de cálculo que conllevan los teoremas de independencia, formulados mediante las funciones generadoras, en vez de por las distribuciones marginales.

El tercer capítulo está dedicado a la generalización al caso multivariante de la obtención de distribuciones de probabilidad, a partir de los sistemas de funciones generadoras, generalizando también los teoremas de independencia entre subvectores aleatorios.

Distribuciones multivariantes. Independencia de variables aleatorias.

Basándonos en los trabajos previos de Steyn, (1960), [43], para el caso continuo, y de Fajardo, [13], para el caso discreto, la generación de distribuciones n-dimensionales se estudia de forma totalmente análoga a las bidimensionales.

Como novedad destaca la construcción de la distribución de un subvector condicionado a otro, a partir de un sistema de funciones generadoras que es un subsistema del sistema de funciones generadoras de la distribución conjunta.

La estructura de independencia entre las variables de un vector aleatorio constituye una de las principales aportaciones de los sistemas de funciones generadoras.

La independencia condicionada, definida por Dawid, (1979), [10], se estudia, de una forma totalmente similar, estableciendo condiciones necesarias y suficientes sobre la dependencia analítica de las funciones que constituyen los sistemas de funciones generadoras de cada uno de los subvectores.

Este capítulo de distribuciones multivariantes finaliza con el tratamiento que, a través de las funciones generadoras, se hace de los parámetros. El sistema de funciones generadoras permite la construcción de otra densidad conjunta donde el vector aleatorio está formado tanto por las variables como por los parámetros, lo que será de utilidad en el capítulo siguiente, dedicado mas de lleno a la utilidad de las funciones generadoras.

OTRAS APLICACIONES DE LAS FUNCIONES GENERADORAS.

El estudio de la independencia estocástica entre variables aleatorias es la primera de las aplicaciones de las funciones generadoras. En el cuarto capítulo se presentan otras aplicaciones que pertenecen al campo teórico de la Estadística Matemática, al campo económico y al campo financiero.

Veremos cómo la factorización de la función de densidad o de cuantía conjunta, surge de las ecuaciones diferenciales, subyacentes en los sistemas de generadoras y se pone de manifiesto el papel primordial que dichos factores juegan a la hora de interpretar las relaciones de dependencia e independencia entre las distintas partes del vector aleatorio.

Para obtener estos resultados se define una relación de equivalencia entre densidades de probabilidad mediante una relación diferencial, en caso continuo, o en diferencias en el caso discreto, que se asemeja a la metodología empleada en la definición y estudio de propiedades de los sistemas pearsonianos.

Propiedad de rectangularidad

Esta relación de equivalencia ha permitido demostrar una propiedad de rectangularidad para los sistemas de funciones generadoras, y como caso particular de ella se obtiene el principio de verosimilitud. El desarrollo de esta propiedad se ha realizado para distribuciones bivariantes continuas, discretas y n-variantes continuas y discretas, aunque sólo se incorporan a esta memoria las demostraciones que corresponden al caso bivariante, pues la generalización al caso n-variante es inmediata, tanto en los enunciados como en las demostraciones, no aportando novedades ni en la metodología ni en las conclusiones.

Las distribuciones mixtas de probabilidad se pueden obtener a partir del conjunto de dos funciones generadoras, una continua y la otra discreta. También los sistemas de funciones generadoras de este tipo de distribuciones verifican la propiedad de rectangularidad.

Familias conjugadas de distribuciones a priori.

La relación existente entre las generadoras de las funciones marginales y condicionadas, supone una relación entre las funciones generadoras de las distribuciones a priori y a posteriori y la derivada parcial del logaritmo de la verosimilitud con respecto al parámetro objeto de estudio,

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = g_\theta(\vec{x},\theta) - g(\theta)$$

que permite comprobar, de una forma sencilla, si una familia de distribuciones a priori es conjugada o no para una clase de verosimilitudes y, conocida la distribución a priori y la verosimilitud, conocer las funciones generadoras de la distribución a posteriori y por tanto la especificación completa de dicha distribución a posteriori.

Las funciones generadoras, la cantidad de información y el método de la máxima verosimilitud.

Bajo condiciones de regularidad, se establece una relación entre la matriz de información de Fisher y las varianzas de las funciones generadoras.

Así mismo las funciones generadoras permiten establecer una relación entre los métodos de estimación de máxima verosimilitud y de los momentos.

Generación de funciones que modelizan la curva de Lorenz.

La aportación en el campo económico viene dado por la utilización de la idea amplia de función generadora. Se estudian las funciones reales de variable real capaces de generar curvas que puedan modelizar una curva de Lorenz.

Casas, Herrerías y Núñez, (1990), [7], estudian dos formas de obtención de funciones que modelizan la curva de Lorenz, una mediante combinaciones lineales convexas de formas funcionales utilizadas en la estimación de dicha curva por otros autores y otra, utilizando la ecuación diferencial que genera la familia de distribuciones continuas univariantes de Pearson. Se estudian las formas potencial y exponencial, ambas satisfacen la ecuación diferencial de Pearson, y otra forma funcional que no satisface la mencionada ecuación. Inician el estudio de la relación entre la familia de Pearson y la modelización de la curva de Lorenz

Posteriormente, Lafuente (1994), [29], estudia ocho casos concretos de la función g(x) y obtiene las correspondientes soluciones de curva de Lorenz.

En la presente memoria se estudian las condiciones que una función g(x), real de variable real, ha de cumplir para que sea generadora de una curva de Lorenz, que son, evidentemente, distintas a las exigidas en el caso de la generación de distribuciones de probabilidad.

Este epígrafe concluye con el estudio de las condiciones que ha de cumplir una función de la forma

$$g(x) = \frac{x-a}{b_0 + b_1 x + b_2 x^2}$$

para que sea generadora de una curva de Lorenz.

Generación de funciones que modelizan una ley financiera.

Por lo que hace referencia a la aplicación al campo financiero, se utiliza la idea amplia de función generadora que permite establecer una relación directa entre ella y la ley financiera.

En 1987 Gil Peláez, [17], anticipa la idea de obtener una ley financiera a partir de una función real de variable real. Muy recientemente Cruz Rambaud, (1994), [9], ha estudiado los capitales financiero-aleatorios, estableciendo una relación entre las leyes financieras y las funciones de distribución de variables aleatorias.

La idea desarrollada en esta memoria consiste en generar la función de distribución de una variable aleatoria a partir de una función real de variable real y, obtenida la distribución, se establece una relación directa entre la ley financiera y la función generadora. Lo que permite, siguiendo la metodología de Cruz Rambaud, [9], disponer de leyes financieras diferentes a partir de las funciones generadoras.

Las condiciones que ha de cumplir esta función generadora son distintas según se pretenda obtener una ley de actualización o de capitalización, y por supuesto son diferentes de las exigidas para la generación de una distribución de probabilidad.

CAPÍTULO 1

GENERACIÓN DE DISTRIBUCIONES UNIVARIANTES DE PROBABILIDAD

En este primer capítulo se establecen los conceptos de función generadora de una distribución de probabilidad univariante tanto para el caso continuo como discreto.

En cada caso se acompañan de las condiciones que debe cumplir la función para que sea generadora de una distribución de probabilidad, se describe el procedimiento de dicha generación. Se acompañan varios ejemplos, mediante los cuales se pone de manifiesto que, para distribuciones usuales, es posible su reconocimiento a partir de la función generadora, así como el valor concreto de sus parámetros.

En las tablas que aparecen en este capítulo se relacionan las funciones generadoras de algunas distribuciones univariantes, en el primer caso continuas y en el segundo discretas, de probabilidad. (Johnson - Kotz, [25], [27]; Rényi, [39]).

Queda patente el paralelismo que el concepto de función generadora establece entre los casos continuo y discreto, y aunque definidas de forma diferente, juegan el mismo papel tanto en la generación como en las aplicaciones.

1.1 FUNCIÓN GENERADORA DE UNA DISTRIBUCIÓN CONTINUA UNIVARIANTE DE PROBABILIDAD.

1.1.1 Condiciones necesarias y suficientes para que una función sea generadora de una distribución continua univariante de probabilidad.

PROPOSICIÓN 1.1

Una función g, real de variable real, definida sobre el intervalo (a,b), es función generadora de una distribución continua de probabilidad sobre el intervalo (a,b) si y sólo si

$$\int_a^b e^{\int^x g(x)\,dx}\,dx < +\infty \quad (1) \qquad (1.1)$$

donde a puede llegar a ser $-\infty$ y b puede llegar a ser $+\infty$

En efecto:

Si la condición requerida se verifica, se puede definir la función de densidad de una v.a. X mediante la expresión:

$$f(x) = \begin{cases} K\,e^{\int^x g(x)\,dx} & \text{si } a < x < b \\ 0 & \text{en el resto} \end{cases} \qquad (1.2)$$

Obsérvese que es una verdadera función de densidad pues la función $e^{\int^x g(x)\,dx}$ es definida positiva y además la integral definida $\int_a^b e^{\int^x g(x)\,dx}\,dx$ es finita y distinta de cero, lo que permite calcular k con la condición de normalización.

---

(1) La expresión $\int^x g(x)\,dx$ representa, una cualquiera de las primitivas de g(x), y en particular la obtenida mediante el procedimiento de integración utilizado sin necesidad de añadir la constante aditiva.

La probabilidad así definida depende de la función g, del recinto sobre el que va a estar definida f y de las condiciones que se le impongan a la densidad resultante, tales como la continuidad o en su defecto el conocimiento de la magnitud de los saltos finitos. Para un mayor detalle se pueden observar los ejemplos que se analizan en el epígrafe 1.1.2.

Por otra parte, si $f(x)$ es la función de densidad de una variable aleatoria y $[a,b]$ es su dominio de definición, se puede calcular la correspondiente función generadora, que viene dada por:

$$g(x) = \frac{f'(x)}{f(x)} = \frac{d \ln f(x)}{dx} \tag{1.3}$$

para los puntos donde $f'(x)$ existe y $f(x)$ no es cero.
Será entonces:

$$\ln f(x) = \int^{x} g(x)\, dx + C$$

$$\int_a^b e^{\int^x g(x)\,dx}\, dx = \int_a^b e^{\ln f(x) - C}\, dx = e^{-C} \int_a^b f(x)\, dx = e^{-C} < +\infty$$

con lo que es posible encontrar para cada distribución continua univariante de probabilidad, la función $g(x)$, definida sobre los puntos donde $f(x)$ es estrictamente positiva y derivable, y que verifica (1.1)

1.1.2 Distribuciones generadas a partir de g(x). Ejemplos.

1. Sea $g(x) = 0 \quad \forall\, x \in (a,b)$

   Será entonces $f(x) = K \quad \forall\, x \in (a,b)$

   Integrando entre a y b, se obtiene el valor de $K = \dfrac{1}{b-a}$

   Se trata pues de la distribución uniforme en el intervalo de extremos reales a y b.

2. Sea $g(x) = -x \quad \forall\, x \in \mathbb{R}$

   Será entonces $f(x) = K\, e^{-\frac{1}{2}x^2} \quad \forall\, x \in \mathbb{R}$

Integrando entre $-\infty$ y $+\infty$ se obtiene el valor de k y la distribución resulta ser la N(0,1)

De igual forma, la distribución $N(\mu,\sigma)$ se genera a partir de $g(x) = -\frac{1}{\sigma^2}x + \frac{\mu}{\sigma^2}$

3. El siguiente ejemplo pone de manifiesto la necesidad de la continuidad de la función f(x) en los puntos donde no sea derivable; este es el caso de la **distribución triangular.** Sea la función

$$g(x) = \begin{cases} \dfrac{1}{x-a} & \text{si} \quad a < x < m \\ \dfrac{1}{b-x} & \text{si} \quad m < x < b \end{cases}$$

siendo a, m y b números reales, entonces:

$$f(x) = \begin{cases} k_1 e^{\int^x \frac{1}{(x-a)}dx} & \text{si} \quad a < x \le m \\ C_2 e^{\int^x \frac{1}{(b-x)}dx} & \text{si} \quad m \le x < b \end{cases}$$

y resolviendo se tiene:

$$f(x) = \begin{cases} k_1(x-a) & \text{si} \quad a < x \le m \\ k_2(b-x) & \text{si} \quad m \le x < b \end{cases}$$

que con las dos condiciones: una la continuidad de f en el punto de abcisa m y otra la condición de normalización,

$$\int_a^m k_1(x-a)\,dx + \int_m^b k_2(b-x)\,dx = 1$$

se calculan los valores de las constantes, resultando ser:

$$f(x) = \begin{cases} \dfrac{2}{(b-a)(m-a)}(x-a) & \text{si} \quad a < x \le m \\ \dfrac{2}{(b-a)(b-m)}(b-x) & \text{si} \quad m \le x < b \end{cases}$$

Se trata, como estaba previsto, de la distribución triangular definida en (a,b) y con moda m. (a < m < b).

Actuando de forma análoga sobre los puntos que definen la partición del dominio, se tratarán a las funciones f dadas por más de dos trozos.

Destaquemos que con la utilización de la función generadora, se sintetizan los métodos clásicos de generación de distribuciones de probabilidad ya que todas las funciones de la familia de Pearson univariante (1) o de su extensión, Roy (3), o de su generalización, Herrerías (4), pueden ser generadas mediante este procedimiento, pero además, con él, es posible obtener la función de densidad de una distribución que no pertenezca a las familias de tipo Pearson como se observa en el siguiente ejemplo.

4. Sea $g(x) = -\dfrac{\ln x + 1}{x}$, definida para todo x real positivo, se tienen entonces

$$f(x) = K\, e^{\int -\frac{\ln x + 1}{x}\,dx}$$

e integrando f(x) entre cero y más infinito, resulta:

$$f(x) = \frac{1}{\sqrt{2\pi}\, x}\, e^{-\frac{1}{2}(\ln x)^2}$$

definida para valores reales positivos.

Se ha generado la distribución Lognormal de parámetros cero y uno.

De igual forma se obtiene la lognormal de parámetros (μ, σ^2) sin más que hacer $g(x) = \dfrac{-1}{\sigma^2}\,\dfrac{\ln x + \sigma^2 - \mu}{x}$, definida sobre el mismo recinto.

5. Sea $g(x) = \frac{ax + b}{x}$, definida para todo x real positivo,

se tienen entonces $f(x) = K\, e^{ax}\, x^b$, e integrando f(x) entre

cero y más infinito, resulta: $f(x) = \frac{1}{\Gamma(b+1)\,(-1/a)^{b+1}}\, x^b\, e^{\frac{x}{1/a}}$

definida para valores reales positivos.

Con la condición de que $b > -1$ y $a < 0$, se ha generado la distribución gamma de parámetros $b+1$ y $-1/a$.

1.1.3 Familias de distribuciones y sus funciones generadoras. Ejemplos.

Dada una familia de funciones de densidad $\{ f(x/\theta) \}_{\theta \in \Theta}$, bajo las condiciones de derivabilidad de cada elemento de la familia, tiene asociada una familia de funciones generadoras $\{ g(x/\theta) \}_{\theta \in \Theta}$, siendo $g(x/\theta) = \frac{\partial}{\partial x} \ln f(x/\theta) \quad \forall\, \theta \in \Theta$

Así, la familia $N(\mu, \sigma^2)$ tiene como familia de generadoras

$$g(x/\mu, \sigma^2) = -\frac{x + \mu}{\sigma^2}$$

Y viceversa, si g(x) es un polinomio de grado uno, con coeficientes reales, y el coeficiente de x es negativo, entonces sobre el recinto $(-\infty, +\infty)$ se genera una distribución normal. Es decir, la familia de rectas

$$g(x/a,b) = ax + b \quad / \quad a < 0$$

constituye una familia de generadoras de la familia normal de media $\mu = -\frac{b}{a}$ y varianza $\sigma^2 = \frac{1}{a}$

Es inmediato comprobar que si g es de la forma

$$g(x/a,b) = \frac{a \ln x + b}{x} \quad \text{con } a \in \mathbb{R}^-,\ b \in \mathbb{R}$$

entonces, sobre el conjunto $\mathbb{R}^+$ se genera la distribución lognormal con $\mu = -\frac{1 + b}{a}$ y $\sigma^2 = -\frac{1}{a}$

La siguiente tabla muestra las funciones generadoras de algunas distribuciones univariantes continuas de probabilidad.

Distribución	Generadora, g(x) =
Uniforme (a,b)	0
Exponencial (λ)	$-\lambda$
Gamma (p,λ)	$\frac{p-1}{x} - \lambda$
Beta (p,q)	$\frac{p-1}{x} - \frac{q-1}{1-x}$
$N(\mu,\sigma^2)$	$-\frac{1}{\sigma^2}x + \frac{\mu}{\sigma^2}$
$\Lambda(\mu,\sigma^2)$	$-\frac{1}{\sigma^2}\,\frac{\ln x + \sigma^2 - \mu}{x}$
Pareto (x_0, a)	$-\frac{a+1}{x}$
Laplace (μ,λ)	$\begin{cases} \lambda^{-1} & x < \mu \\ -\lambda^{-1} & x > \mu \end{cases}$
Cauchy (a,μ)	$-\frac{2(x-\mu)}{a^2+(x-\mu)^2}$
Weibull (a,b,c)	$\frac{b-1}{x-c} - \frac{b(x-c)^{b-1}}{a}$
Weibull (α,β)	$\frac{\beta-1}{x} - \alpha\beta x^{\beta-1}$
Maxwell (σ^2)	$\frac{2}{x} - \sigma^2 x$
Rayleigh (a)	$\frac{1}{x} - \frac{2x}{a}$
$\chi^2(n)$	$\frac{n-2-x}{2x}$
t de Student (n)	$-\frac{(n+1)x}{n+x^2}$
f de Snedecor (n_1,n_2)	$\frac{1}{2}\left(\frac{n_1-2}{x} - \frac{(n_1+n_2)n_1}{n_2+n_1x}\right)$
z de Fisher (n_1,n_2)	$n_1 - \frac{(n_1+n_2)n_1e^{2x}}{n_2+n_1e^{2x}}$

1.2 FUNCIÓN GENERADORA DE UNA DISTRIBUCIÓN DISCRETA UNIVARIANTE. GENERACIÓN DE P(x) A PARTIR DE g(x).

En el caso discreto univariante, J.K. Ord (1967) [34], introdujo una ecuación en diferencias finitas, lineal de primer orden y homogénea, para generar funciones de cuantía de manera análoga al sistema de Pearson, que responde a la expresión:

$$\frac{\Delta p_{r-1}}{p_{r-1}} = \frac{a - r}{b_0 + b_1 r + b_2 r(r-1)} \qquad (1.4)$$

Este sistema fue ampliado por Herrerías (1976) [20], manteniendo el paralelismo con la extensión de Roy [42] en el caso continuo; la ecuación que define la extensión de Herrerías en el caso discreto es:

$$\frac{\Delta p_{r-1}}{p_{r-1}} = \frac{a_0 + a_1 r + a_2 r(r-1)}{b_0 + b_1 r + b_2 r(r-1) + b_3 (r-1)(r-2)} \qquad (1.5)$$

En ambos casos se mantiene el primer miembro de la ecuación en diferencias finitas, modificando sólo el segundo.

En la presente memoria consideraremos que el segundo miembro es una función a la que sólo se le exigirán las condiciones necesarias para que al resolver la ecuación en diferencias se obtenga una distribución de probabilidad.

Evidentemente, no se modifica el concepto de función generadora, con respecto al caso continuo, en el sentido de que se define igual al primer miembro de la ecuación funcional correspondiente, (1.4)

Queremos dejar constancia de que Fajardo (1985), [13], realiza la generalización del sistema de Pearson discreto univariante; parte de una ecuación en diferencias finitas y obtiene la solución general, caracterizando las funciones generatrices de probabilidad asociadas a las soluciones de la mencionada ecuación y estudia algunas familias de distribuciones discretas pertenecientes a este sistema.

El punto de vista que se toma en esta Memoria es diferente al de Fajardo [13], ya que estamos más interesados en estudiar el primer miembro de la ecuación en diferencias (1.4), que las distintas formas concretas que adopta el segundo miembro de dicha ecuación. Se pretende también una herramienta que permita el estudio de las propiedades de las variables discretas de forma análoga al de las continuas.

1.2.1 Condiciones para que una función sea generadora de una distribución de probabilidad univariante discreta.

1° Sea $\{x_n\}$ una sucesión de puntos de salto de una variable aleatoria X. Sea $\{p_n\}$ con $p_n > 0 \quad \forall\, n = 0,1,2,...$ la sucesión de cuantías de dichos valores, se define entonces

$$g(x_r) = \frac{\Delta p_{r-1}}{p_{r-1}} = \frac{p_r - p_{r-1}}{p_{r-1}} \qquad \forall\, r \in \mathbb{N},\ r \geq 1 \qquad (1.6)$$

Para una mejor notación, se utilizará la función

$$L(r) = \frac{p_r}{p_{r-1}} \qquad \forall\, r \geq 1$$

es decir

$$L(r) = g(x_r) + 1$$

2° De la definición anterior se puede deducir una relación de recurrencia, para r mayor que cero, sin más que despejar p_r:

$$p_r = p_{r-1}\, L(r) \quad \text{con} \quad r > 1$$

3° Se pueden ir deduciendo las siguientes relaciones entre los valores de los p_i y los de $L(i)$:

$$p_1 = p_0\, L(1)$$

$$p_2 = p_0 \prod_{i=1}^{2} L(i)$$

....

$$p_n = p_0 \prod_{i=1}^{n} L(i)$$

Puesto que $\Sigma_j \, p_j = 1$, resulta la relación:

$$p_0 + \sum_j p_0 \prod_{i=1}^{j} L(i) = 1 \qquad \text{con } j \geq 1$$

que, bajo las condiciones:

$$0 < k < L(i) \qquad \forall\, i \geq 1 \text{ [1]} \tag{1.7}$$

$$\sum_j \prod_{i=1}^{j} L(i) < +\infty \qquad \text{con } j \geq 1 \tag{1.8}$$

permite despejar p_0 :

$$p_0 = \frac{1}{1 + \sum_j \prod_{i=1}^{j} L(i)} \tag{1.9}$$

y por tanto será:

$$p_n = \frac{\prod_{i=1}^{n} L(i)}{1 + \sum_j \prod_{i=1}^{j} L(i)} \qquad \forall\, n > 0 \tag{1.10}$$

Así pues, cualquier función $g(x)$ tal que sobre una sucesión $\{x_n\}$ verifique las condiciones (1.7) y (1.8), define una distribución de probabilidades $\{p_n\}$, sobre los valores $\{x_n\}$ de la variable, mediante las relaciones (1.9) y (1.10).

Las condiciones necesarias para la generación de una distribución de probabilidad discreta se concretan en el lema y el teorema siguientes.

---

(1) El papel del valor k es asegurar que L(i) posee una cota inferior positiva.

LEMA 1.1

Sea $\{a_n\}$ una sucesión de números reales positivos para los que $\exists N \in \mathbb{N}$ / si $n > N \Rightarrow a_n < k < 1$ **, se verifica entonces que la suma** $\sum_{j=1}^{+\infty} \prod_{i=1}^{j} a_j$ **es finita.**

En efecto, se puede realizar la descomposición:

$$\sum_{j=1}^{+\infty} \prod_{i=1}^{j} a_i = \sum_{j=1}^{N} \prod_{i=1}^{j} a_i + \prod_{i=1}^{N} a_i \sum_{j=N+1}^{+\infty} \prod_{i=N+1}^{j} a_i$$

y solo faltará probar que la suma $\sum_{j=N+1}^{+\infty} \prod_{i=N+1}^{j} a_i$ es finita.

Sea $r = \operatorname{Supr}_{n>N} \{a_n\}$, se tendrá entonces que $r \leq k < 1$.

Por otra parte,

$$\sum_{j=N+1}^{+\infty} \prod_{i=N+1}^{j} a_i < \sum_{j=N+1}^{+\infty} \prod_{i=N+1}^{j} r = \sum_{j=N+1}^{+\infty} r^{j-N}$$

y puesto que $r < 1$, la suma será finita

$$\sum_{j=N+1}^{+\infty} \prod_{i=N+1}^{j} a_i < \sum_{j=N+1}^{+\infty} r^{j-N} = \frac{r}{1-r}$$

con lo que el lema queda demostrado.

TEOREMA 1.1

Dada una sucesión $\{x_n\}$ de números reales, monótona creciente, y una función g, real de variable real, la condición necesaria y suficiente para que dicha función genere una distribución de probabilidades sobre $\{x_n\}$ es que se verifiquen las dos condiciones siguientes:

a) $g(x_n) > -1 \quad \forall n$

b) $\exists N / \quad n > N \Rightarrow g(x_n) < k < 0$

En efecto, la condición es necesaria, pues si la función ha generado una distribución de probabilidad, de las relaciones:

$$L(i) = g(x_i) + 1$$

$$p_0 = \frac{1}{1 + \sum_j \prod_{i=1}^{j} L(i)}$$

$$p_n = \frac{\prod_{i=1}^{n} L(i)}{1 + \sum_j \prod_{i=1}^{j} L(i)} \qquad \forall\, n>0$$

y puesto que $\forall\, i \geq 0$ es $0 < p_i < 1$, se deduce que

$L(i)$ es positivo

$$\sum_{j=1}^{+\infty} \prod_{i=1}^{j} L(i) < +\infty$$

y por lo tanto:

1°) $g(x_i) > -1$ para todo i, esto es, se verifica a)

2°) $\sum_{j=1}^{+\infty} \prod_{i=1}^{j} (1 + g(x_i)) < +\infty$, lo que lleva consigo que a partir de un término en adelante $(1+g(x_i))$ ha de ser menor que uno y por tanto ocurre b)

Es condición suficiente, pues si se verifican las condiciones a) y b) se define la sucesión $\{a_n\} = \{1 + g(x_n)\}$, que por la condición a), todos los términos a_n son positivos y por b) todos son estrictamente menores que uno.

Se puede aplicar el lema 1.1 y resulta

$$\sum_{j=1}^{+\infty} \prod_{i=1}^{j} (1 + g(x_i)) < +\infty$$

y puesto que por definición es $L(i) = g(x_i) + 1$, se tiene

$$\sum_{j=1}^{+\infty} \prod_{i=1}^{j} L(i) < +\infty$$

lo que permite definir una distribución discreta de probabilidad sobre los términos a_n de la forma que se indica en las relaciones (1.9) y (1.10).

1.2.2 Ejemplos de distribuciones generadas a partir de g(x).

1. Sea $\{ x_n \} = \{ n \}$ y sea $g(x) = e^{-kx} - 1$, siendo k un número real positivo.

Evidentemente el producto es

$$\prod_{i=1}^{j} L(i) = \prod_{i=1}^{j} [1 + g(x_i)] = \prod_{i=1}^{j} e^{-ki} = e^{-k\sum_{i=1}^{j} i} = e^{-k\frac{j(j+1)}{2}}$$

por tanto

$$\sum_{j=1}^{+\infty} e^{-k\frac{j(j+1)}{2}} = \sum_{j=1}^{+\infty} [e^{-k}]^{\frac{j(j+1)}{2}} < \sum_{r=1}^{+\infty}[e^{-k}]^{r} = \frac{e^{-k}}{1-e^{-k}} < +\infty$$

y de acuerdo con las expresiones (1.9) y (1.10) obtiene el resultado:

$$p_0 = \frac{1}{1 + \sum_{j=1}^{+\infty} e^{-k\frac{j(j+1)}{2}}}$$

$$p_n = \frac{e^{-k\frac{n(n+1)}{2}}}{1 + \sum_{j=1}^{+\infty} e^{-k\frac{j(j+1)}{2}}}$$

con lo que se tiene generada una distribución de probabilidad, que no es del tipo Pearson - Ord.

2. Sea $\{ x_n \} = \{n-1\}$ y sea $g(x) = \frac{\lambda}{x} - 1$ con $\lambda > 0$

entonces el producto es

$$\prod_{i=1}^{j} L(i) = \prod_{i=1}^{j} (1 + g(x_i)) = \prod_{i=1}^{j} \left(\frac{\lambda}{i-1}\right) = \frac{\lambda^j}{j!}$$

y la suma

$$\sum_{j=1}^{+\infty} \frac{\lambda^j}{j!} = e^{\lambda} - 1$$

de donde se puede escribir:

$$p_0 = \frac{1}{1 + (e^{\lambda} - 1)} = e^{-\lambda}$$

$$p_n = \frac{\lambda^n}{n!} e^{-\lambda}$$

No cabía esperar otro resultado, sino la distribución de Poisson, pues de ella se había obtenido la función generadora.

3. Este ejemplo muestra una función que no permite generar distribuciones de probabilidad de una variable discreta con infinitos puntos de salto.

Obsérvese que si

$$g(x) = \frac{a - x}{b_0 + b_1 x + b_2 x(x-1)} \qquad (1.11)$$

siendo $b_2 \neq 0$, entonces

$$L(x) = g(x) + 1 = \frac{(b_0 + a) + (b_1 - 1)x + b_2 x(x-1)}{b_0 + b_1 x + b_2 x(x-1)}$$

y su límite, cuando x tiende a mas infinito, es uno; y como consecuencia

$$\lim_{j \to +\infty} \prod_{i=1}^{j} L(i) \neq 0$$

lo que indica que

$$\sum_{j=1}^{+\infty} \prod_{i=1}^{j} L(i) = +\infty$$

Así pues, ninguna función como la g(x) definida en (1.11) con $b_2 \neq 0$ podrá generar una distribución de probabilidades discreta con un número infinito de puntos de salto.

Si $b_2 = 0$ entonces

$$L(i) = g(x) + 1 = \frac{(b_0 + a) + (b_1 - 1)x}{b_0 + b_1 x}$$

y su límite, cuando x tiende a más infinito, es $\frac{b_1 - 1}{b_1}$, luego

$$\lim_{j \to +\infty} \prod_{i=1}^{j} L(i) = 0 \quad \text{si y solo si} \quad b_1 \geq 1.$$

Por tanto sólo pueden generar distribuciones con infinitos puntos de salto, entre las funciones del tipo (1.11), aquellas funciones que verifiquen necesariamente que $b_2 = 0$ y que $b_1 \geq 1$, lo que ocurre en el ejemplo 2 del párrafo precedente.

La siguiente tabla muestra las funciones generadoras de algunas distribuciones univariantes discretas de probabilidad.

Distribución	Generadora L(r) =
Binomial (n,p)	$\dfrac{r - (n+1)}{-\dfrac{1-p}{p}\, r}$
Poisson (λ)	$\dfrac{\lambda}{r}$
Geométrica (p)	$1-p$
Hipergeométrica (N_p, N_q, n)	$\dfrac{(N_p - r + 1)(n - r + 1)}{r(N_q - n + r)}$
Binomial negativa (r,p)	$\dfrac{(n + r - 1)(1-p)}{r}$
Serie logarítmica (θ)	$\dfrac{(r-1)\theta}{r}$

CAPÍTULO 2

GENERACIÓN DE DISTRIBUCIONES BIVARIANTES DE PROBABILIDAD

Establecido el concepto de función generadora de una distribución univariante de probabilidad, tanto en el caso continuo como discreto, corresponde ahora extender estas definiciones a las distribuciones bivariantes.

El presente capítulo se subdivide a su vez en tres partes:

En la primera se estudia la generación de una distribución bivariante continua, a partir de un sistema de dos funciones generadoras, siguiendo las directrices marcadas por van Uven (1947), [44], que define la familia pearsoniana bivariante como una extensión de la familia univariante.

Se establece el concepto de sistema de funciones generadoras de una distribución continua bivariante de probabilidad, que resulta estar formado por las funciones generadoras de cada una de las distribuciones condicionadas, lo que nos va a permitir obtener la función de densidad conjunta, $f(x,y)$, como solución de una ecuación diferencial, sin el concurso de las distribuciones marginales.

Bajo condiciones de regularidad, se establecen las relaciones entre las generadoras marginales y condicionadas, que son de gran importancia en el estudio de las familias conjugadas a través de las funciones generadoras y de la relación entre la cantidad de información de Fisher y las funciones generadoras.

En la segunda parte se realiza la generación de distribuciones de probabilidad en el caso bivariante discreto.

El procedimiento seguido es análogo al caso continuo, es decir a partir del sistema de funciones generadoras de las distribuciones condicionadas, conocer las condiciones necesarias para la obtención de la probabilidad conjunta.

La tercera parte estudia la independencia entre las variables y su relación con las funciones generadoras.
Se pone de manifiesto que la independencia estocástica se puede observar a partir de la dependencia analítica de las funciones generadoras de las variables.

2.1 SISTEMA DE FUNCIONES GENERADORAS DE UNA DISTRIBUCIÓN BIVARIANTE CONTINUA.

2.1.1 Funciones generadoras univariantes ligadas a una distribución bidimensional.

Dada la función de densidad conjunta, $f(x,y)$, el cálculo de probabilidades elemental nos permite obtener cuatro densidades unidimensionales asociadas con ella: las densidades marginales $f(x)$[1], $f(y)$ y las densidades condicionadas $f(x/y)$ y $f(y/x)$.

Como tales distribuciones unidimensionales, para cada una de ellas se puede obtener la función generadora correspondiente en la forma indicada en el capítulo 1,

$$g(x) = \frac{d \ln f(x)}{dx}$$

$$g(y) = \frac{d \ln f(y)}{dy}$$

(1) Siempre que no haya lugar a duda se va a denotar la densidad conjunta, las marginales y las condicionadas con el mismo signo, f. La misma consideración para las funciones generadoras.

$$g_x(x,y) = \frac{\partial}{\partial x} \ln f(x/y) \tag{2.1}$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(y/x) \tag{2.2}$$

las respectivas funciones

marginales y condicionadas. Las marginales tienen idénticas expresiones que en el capítulo anterior, mientras las condicionadas tienen la forma:

$$f(y/x) = k(x)\, e^{\int^y g_y(x,y)\, dy}$$

$$f(x/y) = k(y)\, e^{\int^x g_x(x,y)\, dx}$$

Con las mismas condiciones que en el caso univariante, acerca del valor finito de las integrales

$$\int_{-\infty}^{+\infty} e^{\int^y g_y(x,y)\, dy}\, dy \qquad \int_{-\infty}^{+\infty} e^{\int^x g_x(x,y)\, dx}\, dx$$

y el conocimiento de la continuidad en aquellos puntos donde las densidades condicionadas no sean derivables, es decir donde las funciones generadoras de las condicionadas no existan, se tiene entonces que:

$$f(y/x) = k(x)\, e^{\int^y g_y(x,y)\, dy}$$

siendo

$$k(x) = \left[\int_{-\infty}^{+\infty} e^{\int^y g_y(x,y)\, dy}\, dy \right]^{-1}$$

la constante de normalización.

$$f(x/y) = k(y)\, e^{\int^x g_x(x,y)\, dx}$$

siendo, igualmente

$$k(y) = \left[\int_{-\infty}^{+\infty} e^{\int^x g_x(x,y)\, dx}\, dx \right]$$

la constante de normalización.

Puesto que desde una densidad condicionada y su correspondiente marginal se puede obtener la densidad conjunta, está claro que las cuatro generadoras unidimensionales permiten la reconstrucción de la densidad conjunta, ahora bien, cabe preguntarse si este proceso de reconstrucción ha de pasar necesariamente por la obtención previa de las densidades unidimensionales, o si por el contrario, será factible la reconstrucción directa.

La siguiente proposición nos va a permitir la reconstrucción directa de la densidad conjunta a partir de las dos generadoras de las distribuciones condicionadas.

PROPOSICIÓN 2.1

Dada una densidad conjunta, f(x,y), las funciones generadoras de las densidades condicionadas se pueden expresar como:

$$g_x(x,y) = \frac{\partial}{\partial x} \ln f(x,y) \qquad \textbf{(2.3)}$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(x,y) \qquad \textbf{(2.4)}$$

La demostración de (2.3) resulta elemental, teniendo en cuenta la definición (2.1) y que $f(x/y) = \frac{f(x,y)}{f(y)}$, de donde se deduce que

$$\ln f(x/y) = \ln f(x,y) - \ln f(y) \qquad (2.5)$$

y derivando con respecto a x, se tiene la igualdad deseada. Análogamente se demuestra (2.4)

Obsérvese que existe una relación directa, [(2.3) y (2.4)], entre las generadoras condicionadas y la densidad conjunta, que nos va a permitir obtener f(x,y) como solución de una ecuación diferencial, sin el concurso de las generadoras marginales, como después se verá.

Esta obtención justifica la interpretación del par $\{ g_x(x,y), g_y(x,y) \}$ como sistema generador de la distribución conjunta de probabilidad.

2.1.2 Condiciones para la generación. Existencia de una diferencial total.

La importancia de la proposición 2.1 radica en que las generadoras condicionadas están directamente relacionadas con la función de densidad conjunta a través de la correspondiente derivada parcial del logaritmo neperiano.

PROPOSICIÓN 2.2

Supuestas ciertas las hipótesis sobre $g_x(x,y)$ y $g_y(x,y)$, necesarias para la generación de las respectivas densidades univariantes de probabilidad, la condición necesaria y suficiente para que el conjunto $\{ g_x(x,y), g_y(x,y)\}$ sea un sistema de generadoras de una densidad conjunta, $f(x,y)$, es que se verifique:

$$\frac{\partial}{\partial y} g_x(x,y) = \frac{\partial}{\partial x} g_y(x,y) \qquad \textbf{(2.6)}$$

En efecto, la condición necesaria es trivial, puesto que si existe una función de densidad, $f(x,y)$, de la cual resulta ser

$$g_x(x,y) = \frac{\partial}{\partial x} \ln f(x,y)$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(x,y)$$

entonces la condición (2.6) significa que

$$\frac{\partial^2}{\partial x\, \partial y} \ln f(x,y) = \frac{\partial^2}{\partial y\, \partial x} \ln f(x,y)$$

lo que se verifica siempre que $f(x,y)$ sea dos veces derivable.

En caso de que exista un conjunto de puntos en los que la función de densidad no sea derivable, se realiza análogo razonamiento que en el caso univariante.

Es condición suficiente pues si (2.6) es cierta, entonces $g_x(x,y)$ y $g_y(x,y)$ forman parte de una misma diferencial total, y por tanto existe $U(x,y)$ tal que

$$dU(x,y) = g_x(x,y)\,dx + g_y(x,y)\,dy$$

siendo $g_x(x,y) = \frac{\partial}{\partial x} U(x,y)$ y $g_y(x,y) = \frac{\partial}{\partial y} U(x,y)$

Evidentemente la densidad buscada es

$$f(x,y) = K\, e^{U(x,y)}$$

Obsérvese que las expresiones (2.3) y (2.4) constituyen un sistema similar al establecido por van Uven para la definición de la familia bivariante de Pearson.

Cabe destacar, como es de esperar, que la familia pearsoniana bivariante continua, es el caso particular en el que cada una de las generadoras condicionadas están constituidas por cocientes de una forma lineal entre otra cuadrática:

$$g_x(x,y) = \frac{L_1(x,y)}{Q_1(x,y)} \qquad g_y(x,y) = \frac{L_2(x,y)}{Q_2(x,y)}$$

donde como ya es conocido ha de ocurrir que

$$\frac{\partial}{\partial y}\frac{L_1(x,y)}{Q_1(x,y)} = \frac{\partial}{\partial x}\frac{L_2(x,y)}{Q_2(x,y)}$$

En definitiva, cualquier par de funciones $g_1, g_2 : R^2 \to R$, integrables sobre el recinto adecuado, para las que $\frac{\partial}{\partial y} g_1(x,y) = \frac{\partial}{\partial x} g_2(x,y)$, generan una distribución conjunta de probabilidad sobre un recinto finito.

Si el recinto no es finito, para obtener una distribución de probabilidad, se habrá de cumplir además que la integral

$$\int_{-\infty}^{+\infty} e^{\int^{x} g_1(x,y)\,dx + \int^{y} g_2(x,y)\,dy}\,dx\,dy$$

sea finita.

2.1.3 Relación entre las funciones generadoras de las distribuciones marginales y las funciones generadoras de las distribuciones condicionadas de una distribución bidimensional.

A continuación se dan dos proposiciones que relacionan las funciones generadoras de las distribuciones marginales y las condicionadas de una misma distribución bidimensional; ambas serán utilizadas en el capítulo cuarto, donde se establece la relación entre la función cantidad de información de Fisher y las funciones generadoras.

Para la demostración de la proposición 2.4 será necesario suponer que de las condiciones de regularidad, [2], al menos se verifica que:

$$\frac{\partial}{\partial y}\int_{-\infty}^{+\infty} f(x/y)\,dx = \int_{-\infty}^{+\infty} \frac{\partial}{\partial y} f(x/y)\,dx \qquad \forall y$$

PROPOSICIÓN 2.3

Dada una función de densidad conjunta, f(x,y), las funciones generadoras de las distribuciones marginales se pueden expresar como:

$$g(y) = g_y(x,y) - \frac{\partial}{\partial y} \ln f(x/y) \tag{2.7}$$

$$g(x) = g_x(x,y) - \frac{\partial}{\partial x} \ln f(y/x) \tag{2.8}$$

La demostración de (2.7) es inmediata, derivando (2,5) con respecto a y. De forma similar se obtiene (2.8).

Esta proposición nos da una relación entre las funciones generadoras de las distribuciones marginales y las condicionadas de la misma distribución, en la que se observa claramente que las igualdades: $g(y) = g_y(x,y)$ y $g(x) = g_x(x,y)$ se verifican cuando las variables aleatorias X e Y son independientes.

PROPOSICIÓN 2.4

Dada una función de densidad conjunta $f(x,y)$, se verifica que, para cada y, $g(y)$ es igual al valor esperado de $g_y(x,y)$, bajo la distribución $f(x/y)$, es decir:

$$E_{f(x/y)}\ [\ g_y(x,y)\ ] = g(y) \tag{2.9}$$

$$\int_{-\infty}^{+\infty} g_y(x,y)\ f(x/y)\ dx = g(y) \tag{2.10}$$

para todo y, para el cual $f(x/y)$ está definida

En efecto, calculamos el valor esperado de la generadora con respecto a la verosimilitud:

$$E_{f(x/y)}\ [\ g_y(x,y)\ ] = \int_{-\infty}^{+\infty} g_y(x,y)\ f(x/y)\ dx =$$

y según la proposición 2.3 se puede escribir

$$= \int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial y} \ln f(x/y) + g(y) \right] f(x/y)\ dx =$$

$$= \int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial y} \ln f(x/y) \right] f(x/y)\ dx + g(y) \int_{-\infty}^{+\infty} f(x/y)\ dx$$

pero según Zacks, [48], lema 4.1.1, se verifica que

$$\int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial y} \ln f(x/y) \right] f(x/y)\ dx = 0 \quad \forall\, y$$

es decir, la primera de las dos integrales anteriores vale cero, mientras que la segunda, evidentemente es uno. Por tanto:

$$\int_{-\infty}^{+\infty} g_y(x,y)\ f(x/y)\ dx = g(y)$$

con lo que queda demostrada la proposición.

2.1.4 Función de densidad conjunta factorizada.

Resulta evidente que no tienen igual comportamiento los distintos factores de la densidad conjunta en las funciones generadoras condicionadas, es decir, los factores que sólo dependan de x, se perderán al tomar logaritmos y derivar con respecto a y; Igual situación para factores que sólo dependen de y, al tomar logaritmos y derivar con respecto a x.

Es interesante que la función de densidad conjunta de la distribución esté en forma factorizada, para tratar de conocer la información que no se pierde al calcular las generadoras condicionadas, y de cómo la información que se pierde en una de ellas se puede recuperar desde la otra.

Por tanto, en lo que sigue se supone que, tras obtener los posibles factores en x e y, siempre se puede expresar la función de densidad conjunta, en el caso mas general, como

$$f(x,y) = K\, h(x,y)\, l_1(x)\, l_2(y) \qquad (2.11)$$

siendo h(x,y) una función que no posee factores que dependan sólo de x o sólo de y, pudiendo llegar a valer uno, (cuando X e Y son variables aleatorias independientes), y siendo $l_1(x)$ y $l_2(y)$ factores en x e y respectivamente.

Resulta evidente que dada la función f(x,y) en la forma indicada en (2.11), las densidades marginales y condicionadas se expresan de la forma:

$$f(x) = K\, l_1(x) \int_{-\infty}^{+\infty} h(x,y)\, l_2(y)\, dy$$

$$f(y) = K\, l_2(y) \int_{-\infty}^{+\infty} h(x,y)\, l_1(x)\, dx$$

$$f(x/y) = \frac{K\, h(x,y)\, l_1(x)\, l_2(y)}{K\, l_2(y) \int_{-\infty}^{+\infty} h(x,y)\, l_1(x)\, dx} = \frac{h(x,y)\, l_1(x)}{\int_{-\infty}^{+\infty} h(x,y)\, l_1(x)\, dx}$$

$$f(y/x) = \frac{K\, h(x,y)\, l_1(x)\, l_2(y)}{K\, l_1(x) \int_{-\infty}^{+\infty} h(x,y)\, l_2(y)\, dy} = \frac{h(x,y)\, l_2(y)}{\int_{-\infty}^{+\infty} h(x,y)\, l_2(y)\, dy}$$

2.1.5 Obtención de las distribuciones marginales conocidas las distribuciones condicionadas.

En el epígrafe 2.2.2 se ha obtenido la distribución conjunta, salvo la constante de normalización, a partir de las generadoras de las distribuciones condicionadas.

Se pone ahora de manifiesto que, dadas las dos distribuciones condicionadas de una misma distribución conjunta de probabilidad, a través de un sencillo cálculo, se pueden determinar las distribuciones marginales, salvo las constantes de normalización, y por tanto la distribución conjunta.

PROPOSICIÓN 2.5

Sea (X,Y) una variable aleatoria continua bidimensional, de la que se conocen las distribuciones condicionadas, entonces se verifica que el cociente $\frac{f(x/y)}{f(y/x)}$ es igual al cociente de dos funciones, el numerador que sólo depende de x y el denominador que sólo depende de y: $\frac{f(x/y)}{f(y/x)} = \frac{f_1(x)}{f_2(y)}$

Las siguientes igualdades son evidentemente ciertas:

$$\frac{f(x/y)}{f(y/x)} = \frac{\dfrac{f(x,y)}{f_2(y)}}{\dfrac{f(x,y)}{f_1(x)}} = \frac{f_1(x)}{f_2(y)}$$

Con lo que, debido a la factorización, es inmediato separar la parte que sólo depende de x de la que sólo depende de y, salvo constantes que se obtienen al normalizar. Expresado según las fórmulas del epígrafe 2.2.4 es:

$$\frac{f(x/y)}{f(y/x)} = \frac{\dfrac{l_1(x)}{\int_{-\infty}^{+\infty} h(x,y)\ l_1(x)\ dx}}{\dfrac{l_2(y)}{\int_{-\infty}^{+\infty} h(x,y)\ l_2(y)\ dy}} = \frac{l_1(x) \int_{-\infty}^{+\infty} h(x,y)\ l_2(y)\ dy}{l_2(y) \int_{-\infty}^{+\infty} h(x,y)\ l_1(x)\ dx}$$

De la observación de la igualdad anterior, queda suficientemente claro que el numerador sólo depende de x y que el denominador sólo depende de y.

Como ejemplo se obtienen las densidades condicionadas y las marginales partiendo de un sistema de generadoras de la distribución normal bivariante:

Sean

$$g_y(x,y) = y - \mu_y - \rho \frac{\sigma_2}{\sigma_1}(x - \mu_x)$$

$$g_x(x,y) = x - \mu_x - \rho \frac{\sigma_1}{\sigma_2}(y - \mu_y)$$

(Obsérvese que $g_y(x,y) = 0$, $g_x(x,y) = 0$ son las rectas de regresión de y/x y x/y respectivamente. Es fácil comprobar que este hecho no se verifica en general. Queda para un estudio posterior la búsqueda de las distribuciones de probabilidad que verifican esta propiedad).

Las distribuciones condicionadas son:

$$f(y/x) \propto e^{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left(y - \mu_y - \rho \frac{\sigma_2}{\sigma_1}(x - \mu_x)\right)^2}$$

$$f(x/y) \propto e^{-\frac{1}{2(1-\rho^2)\sigma_1^2}\left(x - \mu_x - \rho \frac{\sigma_1}{\sigma_2}(y - \mu_y)\right)^2}$$

entonces el cociente es:

$$\frac{f(x/y)}{f(y/x)} \propto e^{-\frac{1}{2(1-\rho^2)}\left\{\left[\left(\frac{x-\mu_x}{\sigma_1}\right) - \rho\left(\frac{y-\mu_y}{\sigma_2}\right)\right]^2 - \left[\left(\frac{y-\mu_y}{\sigma_2}\right) - \rho\left(\frac{x-\mu_x}{\sigma_1}\right)\right]^2\right\}}$$

ahora bien:

$$\left[\left(\frac{x - \mu_x}{\sigma_1}\right) - \rho\left(\frac{y - \mu_y}{\sigma_2}\right)\right]^2 - \left[\left(\frac{y - \mu_y}{\sigma_2}\right) - \rho\left(\frac{x - \mu_x}{\sigma_1}\right)\right]^2 =$$

$$= (1 - \rho^2)\left[\left(\frac{x - \mu_x}{\sigma_1}\right)^2 - \left(\frac{y - \mu_y}{\sigma_2}\right)^2\right]$$

por tanto

$$\frac{f(x/y)}{f(y/x)} \propto e^{-\frac{1}{2(1-\rho^2)}(1-\rho^2)\left[\left(\frac{x-\mu_x}{\sigma_1}\right)^2 - \rho\left(\frac{y-\mu_y}{\sigma_2}\right)^2\right]} = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_1}\right)^2}}{e^{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_2}\right)^2}}$$

de donde, como era de esperar:

$$f(x) \propto e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_1}\right)^2} \qquad f(y) \propto e^{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_2}\right)^2}$$

2.1.6 Sistema de funciones generadoras de una función de densidad conjunta cuando se considera su forma factorizada.

PROPOSICIÓN 2.6

Dada la función de densidad conjunta en su forma factorizada, (2.11), las funciones generadoras de las distribuciones condicionadas son:

$$g_x(x,y) = \frac{\partial}{\partial x} \ln h(x,y) + \frac{\partial}{\partial x} \ln l_1(x) \qquad (2.12)$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln h(x,y) + \frac{\partial}{\partial y} \ln l_2(y) \qquad (2.13)$$

y las funciones generadoras de las distribuciones marginales son:

$$g(x) = \frac{\partial}{\partial x} \ln \int_{-\infty}^{+\infty} h(x,y)\, l_2(y)\, dy + \frac{\partial}{\partial x} \ln l_1(x)$$

$$g(y) = \frac{\partial}{\partial y} \ln \int_{-\infty}^{+\infty} h(x,y)\, l_1(x)\, dx + \frac{\partial}{\partial y} \ln l_2(y)$$

En efecto, de (2.11) se deduce que:

$$\ln f(x,y) = \ln k + \ln h(x,y) + \ln l_1(x) + \ln l_2(y)$$

y derivando parcialmente con respecto a x e y se obtienen las funciones generadoras de las distribuciones condicionadas.

Las funciones generadoras de las distribuciones marginales aparecen sin más que tomar logaritmos y derivar las expresiones de las funciones de densidad marginales, calculadas a partir de (2.11).

Se quiere hacer notar que la función $g_x(x,y)$ no depende del factor de la densidad conjunta, $l_2(y)$, y que la función $g_y(x,y)$ no depende del factor de la densidad conjunta, $l_1(x)$.

Este hecho es importante porque constituye la razón por la cual se verifica el Principio de Verosimilitud, como en el capítulo de aplicaciones veremos.

2.1.7 Reconstrucción de la función de densidad conjunta en forma factorizada.

En este epígrafe se expone una vía para la obtención de la densidad conjunta, a partir del sistema de generadoras.

Denotamos:

$$m_1(x,y) = \frac{\partial}{\partial x} \ln h(x,y) \tag{2.14}$$

$$n_1(x) = \frac{\partial}{\partial x} \ln l_1(x) \tag{2.15}$$

$$m_2(x,y) = \frac{\partial}{\partial y} \ln h(x,y) \tag{2.16}$$

$$n_2(y) = \frac{\partial}{\partial y} \ln l_2(y) \tag{2.17}$$

Las funciones generadoras condicionadas toman la forma:

$$g_x(x,y) = m_1(x,y) + n_1(x) \tag{2.18}$$

$$g_y(x,y) = m_2(x,y) + n_2(y) \tag{2.19}$$

Obsérvese que $m_2(x,y)$, función de x e y, puede ser constante o depender solo de x, pero no será una función únicamente de y, puesto que estaría englobada en la parte $n_2(y)$ esto significa que en $g_y(x,y)$ son inconfundibles los términos $m_2(x,y)$ y $n_2(y)$ y por tanto son perfectamente separables uno del otro.

El mismo razonamiento se puede hacer para las funciones $m_1(x,y)$ y $n_1(x)$.

Puesto que la densidad conjunta $f(x,y)$ se compone, salvo constante de normalización, de los tres factores $h(x,y)$, $l_1(x)$, $l_2(y)$ y todos ellos son recuperables desde los elementos $m_1(x,y)$, $m_2(x,y)$, $n_1(x)$, $n_2(y)$ por simple integración, es claro que los resultados obtenidos nos dan algunas alternativas para la reconstrucción de $f(x,y)$, que sin duda serán equivalentes al proceso desarrollado en el epígrafe 2.2.2

PROPOSICIÓN 2.7

La condición necesaria y suficiente para que dos funciones, expresadas en la forma (2.18) y (2.19), constituyan un sistema de generadoras de una distribución bidimensional es que

$$\frac{\partial}{\partial x} m_2(x,y) = \frac{\partial}{\partial y} m_1(x,y) \tag{2.20}$$

y que la integral $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^y m_2(x,y)\,dy + \int^x n_1(x)\,dx + \int^y n_2(y)\,dy}\,dy\,dx$ **sea finita.**

En efecto, la igualdad de Schwartz para la función $\ln f(x,y)$ equivale, bajo la forma factorizada, a la igualdad establecida para la función $h(x,y)$,

$$\frac{\partial^2}{\partial x\,\partial y} \ln h(x,y) = \frac{\partial^2}{\partial y\,\partial x} \ln h(x,y)$$

siendo la última de ellas, idéntica a la igualdad:

$$\frac{\partial}{\partial x} m_2(x,y) = \frac{\partial}{\partial y} m_1(x,y)$$

Evidentemente los límites de integración son, en todo caso, los extremos del dominio de definición de cada una de las variables.

Por otra parte, de las definiciones dadas en (2.14), ... (2.17), tenemos que

$$h(x,y) = e^{\int^y m_2(x,y)\,dy}$$

$$l_1(x) = e^{\int^x n_1(x)\,dx}$$

$$h(x,y) = e^{\int^x m_1(x,y)\,dx}$$

$$l_2(y) = e^{\int^y n_2(y)\,dy}$$

y como consecuencia se puede obtener f(x,y) a través de una de las dos expresiones siguientes:

a) $$f(x,y) = k \,.\, e^{\int^y m_2(x,y)\,dy + \int^x n_1(x)\,dx + \int^y n_2(y)\,dy} \qquad (2.21)$$

b) $$f(x,y) = k \,.\, e^{\int^x m_1(x,y)\,dx + \int^x n_1(x)\,dx + \int^y n_2(y)\,dy} \qquad (2.22)$$

siendo necesario que la integral

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^y m_2(x,y)\,dy + \int^x n_1(x)\,dx + \int^y n_2(y)\,dy}\, dy\, dx$$

sea finita, para calcular la constante k al imponer la condición de normalización.

Nótese que la integral

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^x m_1(x,y)\,dx + \int^x n_1(x)\,dx + \int^y n_2(y)\,dy}\, dy\, dx$$

también es finita, por la definición de las funciones $m_1(x,y)$, (2.5) y $m_2(x,y)$, (2.16).

2.1.8 Otras expresiones de la función de densidad conjunta.

De (2.21) y (2.22) se sigue que la función f(x,y) también se puede escribir:

$$f(x,y) = k \,.\, e^{\int^y g_y(x,y)\,dy + \int^x n_1(x)\,dx} \tag{2.23}$$

o bien

$$f(x,y) = k \,.\, e^{\int^x g_x(x,y)\,dx + \int^y n_2(y)\,dy} \tag{2.22}$$

Se observa que la expresión (2.23) está en concordancia con el hecho de que, como en su momento se destacó, la función $g_y(x,y)$ no depende del factor $l_1(x)$ de la densidad conjunta. Análogo razonamiento para (2.24).

Por otra parte, multiplicando (2.23) y (2.24) se llega a la expresión:

$$f^2(x,y) = k^2\, e^{\int^x n_1(x)\,dx}\, e^{\int^y n_2(y)\,dy}\, e^{\int^x g_x(x,y)\,dx + \int^y g_y(x,y)\,dy}$$

$$f^2(x,y) = k^2\, l_1(x)\, l_2(y)\, e^{\int^x g_x(x,y)\,dx + \int^y g_y(x,y)\,dy} \tag{2.25}$$

luego:

$$f(x,y) = k \left[l_1(x)\, l_2(y)\, e^{\int^x g_x(x,y)\,dx + \int^y g_y(x,y)\,dy} \right]^{\frac{1}{2}} \tag{2.26}$$

y también dividiendo (2.25) por (2.11) se tiene:

$$f(x,y) = k\, \frac{e^{\int^x g_x(x,y)\,dx + \int^y g_y(x,y)\,dy}}{h(x,y)}$$

$$f(x,y) = k\, \frac{e^{\int^x g_x(x,y)\,dx + \int^y g_y(x,y)\,dy}}{e^{\int^x m_1(x,y)\,dx}} \tag{2.27}$$

$$f(x,y) = k\, \frac{e^{\int^x g_x(x,y)\,dx + \int^y g_y(x,y)\,dy}}{e^{\int^y m_y(x,y)\,dy}} \tag{2.28}$$

2.1.9 Ejemplos de generación de distribuciones continuas bidimensionales.

En este epígrafe se desarrollan algunos ejemplos y se recogen en una tabla los sistemas de generadoras de algunas distribuciones continuas bivariantes de probabilidad, (Mardia, [33]).

1. En las distribuciones univariantes se propuso como primer ejemplo la distribución generada por la función $g(x) = 0$, que resultó ser la distribución uniforme en el intervalo de definición, (a,b)

Ahora se busca la distribución generada a partir del sistema formado por las funciones $\begin{cases} g_x(x,y) = 0 & a < x < b \\ g_y(x,y) = 0 & c < y < d \end{cases}$.

Será entonces: $f(x,y) = K = \dfrac{1}{(b-a)(d-c)}$, $a<x<b$; $c<y<d$

2. Se propuso también como ejemplo la distribución generada por la función $g(x) = -x$, ahora se busca la distribución generada a partir del sistema formado por las funciones

$$\begin{cases} g_x(x,y) = -x \\ g_y(x,y) = -y \end{cases}$$

Efectivamente, se verifica la igualdad entre derivadas parciales cruzadas, $\dfrac{\partial}{\partial y} g_x(x,y) = 0 = \dfrac{\partial}{\partial x} g_y(x,y)$.

Será entonces: $\begin{array}{ll} m_1(x,y) = 0 & n_1(x) = -x \\ m_2(x,y) = 0 & n_2(y) = -y \end{array}$

y por tanto: $f(x,y) \propto e^{-\int x dx - \int y dy}$

$$f(x,y) \propto e^{-\frac{1}{2}(x^2+y^2)}$$

Se trata de la distribución normal bivariante con $\rho = 0$

3. Dadas las funciones

$$\begin{cases} g_x(x,y) = \frac{1}{2}y - x \\ \\ g_y(x,y) = \frac{1}{2}x - y \end{cases}$$

entonces $\frac{\partial}{\partial y} g_x(x,y) = \frac{1}{2} = \frac{\partial}{\partial x} g_y(x,y)$.

Separando términos se tiene:

$$m_1(x,y) = \frac{1}{2}y \qquad n_1(x) = -x$$

$$m_2(x,y) = \frac{1}{2}x \qquad n_2(y) = -y$$

y por tanto: $f(x,y) \propto e^{\int \frac{1}{2} x dy - \int x dx - \int y dy}$

o bien $f(x,y) \propto e^{\int \frac{1}{2} y dx - \int x dx - \int y dy}$

obteniéndose $f(x,y) \propto e^{-\frac{1}{2}(-xy + x^2 + y^2)}$

El resultado es la distribución normal bivariante, cuyos parámetros se obtienen resolviendo el sistema de ecuaciones:

$$\begin{cases} 2\rho = (1-\rho^2)\,\sigma_1\,\sigma_2 \\ 1 = (1-\rho^2)\,\sigma_1^2 \\ 1 = (1-\rho^2)\,\sigma_2^2 \end{cases}$$

cuya solución es: $\rho = 0{,}5 \qquad \mu_1 = \mu_2 = 0 \qquad \sigma_1^2 = \sigma_2^2 = \frac{4}{3}$

Algunas distribuciones continuas bivariantes de probabilidad:

Distribución	generadoras $g_x(x,y) =$ $g_y(x,y) =$
$N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$	$-\dfrac{1}{(1-\rho^2)\,\sigma_x}\left(\dfrac{x-\mu_x}{\sigma_x} - \rho\,\dfrac{y-\mu_y}{\sigma_y}\right)$ $-\dfrac{1}{(1-\rho^2)\,\sigma_y}\left(\dfrac{y-\mu_y}{\sigma_y} - \rho\,\dfrac{x-\mu_x}{\sigma_x}\right)$
Uniforme (a) Morgenstern	$\dfrac{2a(2y-1)}{1+a(2x-1)(2y-1)}$ $\dfrac{2a(2x-1)}{1+a(2x-1)(2y-1)}$
Cauchy (c)	$-\dfrac{3}{2}\,\dfrac{2x}{c^2+x^2+y^2}$ $-\dfrac{3}{2}\,\dfrac{2y}{c^2+x^2+y^2}$
Beta (p_1, p_2, p_3)	$\dfrac{p_1-1}{x} - \dfrac{p_3-1}{1-x-y}$ $\dfrac{p_2-1}{y} - \dfrac{p_3-1}{1-x-y}$

Gamma (a, p, q) McKay	$\frac{p-1}{x} - \frac{q-1}{y-x}$ $\frac{q-1}{y-x} - a$
Exponencial (a) Gumbel	$\frac{a(1+ay)}{(1+ax)(1+ay)-a} - 1 - ay$ $\frac{a(1+ax)}{(1+ax)(1+ay)-a} - 1 - ax$
Pareto (p,a,b)	$-\frac{(p+2)b}{bx+ay-ab}$ $-\frac{(p+2)a}{bx+ay-ab}$
t de Student	$\frac{(f+2)(\rho y-x)}{(f-2)(1-\rho^2)+x^2+y^2-2\rho xy}$ $\frac{(f+2)(\rho x-y)}{(f-2)(1-\rho^2)+x^2+y^2-2\rho xy}$
F de Snedecor	$\frac{\frac{1}{2}\nu_1-1}{x} - \frac{\frac{1}{2}\nu\nu_1}{\nu_0+\nu_1x+\nu_2y}$ $\frac{\frac{1}{2}\nu_2-1}{y} - \frac{\frac{1}{2}\nu\nu_2}{\nu_0+\nu_1x+\nu_2y}$
Rhodes	$-1-\frac{pb}{ab-bx+ay} + \frac{p'b'}{a'b'+b'x+a'y}$ $-m+\frac{pa}{ab-bx+ay} + \frac{p'a'}{a'b'+b'x+a'y}$

2.1.10 La función g(x,y)

En el primer capítulo hemos visto que una distribución de probabilidad puede generarse a partir de una función real de variable real. En los apartados anteriores del presente capítulo se ha generado una distribución a partir de un sistema de dos funciones, consideradas cada una de ellas como funciones generadoras de las distribuciones condicionadas.

Cabe preguntarse ¿por qué no realizar la generación mediante una sola función, $g: \mathbb{R}^2 \to \mathbb{R}$, considerada como parcial segunda, cruzada de la densidad?. La respuesta a esta pregunta, aparece en la siguiente proposición; en ella se estudian las condiciones que han de cumplirse para que una función de este tipo genere una densidad bivariante de probabilidad.

Definimos la función g(x,y) como

$$g(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f(x,y) = \frac{\partial^2}{\partial y\, \partial x} \ln f(x,y) \quad ^{(1)}$$

Expresando f(x,y) en su forma factorizada,

$$f(x,y) = K\, h(x,y)\, l_1(x)\, l_2(y)\,,$$

es evidente que

$$g(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln h(x,y) \qquad (2.29)$$

Este resultado nos muestra, claramente, que no siempre es posible la reconstrucción de la densidad f(x,y), como veremos en la siguiente proposición, a partir sólo de g(x,y).

---

(1) La notación $g_{xy}(x,y)$ se simplificará, siempre que no haya lugar a duda, en g(x,y).

PROPOSICIÓN 2.8

Dada la distribución conjunta en su forma factorizada la condición necesaria y suficiente para que la función real de dos variables reales, $g(x,y) = \frac{\partial^2}{\partial x\,\partial y} \ln f(x,y)$ **, permita reconstruir por integración la distribución conjunta es que los factores** $l_1(x)$ **y** $l_2(y)$ **sean igual a uno.**

En efecto, es condición necesaria, pues si g(x,y) permite obtener f(x,y), esta será:

$$f(x,y) \propto e^{\int^x\int^y g(x,y)\,dxdy} = e^{\int^x\int^y \frac{\partial^2}{\partial x\,\partial y} \ln h(x,y)\,dx\,dy}$$

por tanto $f(x,y) = K\,h(x,y)$, de donde se deduce que $l_1(x) = \frac{1}{l_2(y)}$, y puesto que l_1 es función sólo de x, l_2 lo es sólo de y, y las constantes están englobadas en K, se deduce que $l_1(x) = l_2(y) = 1$

Es condición suficiente porque, si no hay factores individualizados, es decir si $l_1(x) = l_2(y) = 1$, entonces $f(x,y) = K\,h(x,y)$ y puesto que g(x,y) permite reconstruir, salvo constante, la función h(x,y), y por tanto la distribución conjunta se reconstruye a partir de g(x,y) cuando la integral

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^x\int^y g(x,y)\,dxdy}\,dx\,dy$$

definida sobre el dominio de las variables, es convergente.

La proposición resulta fácilmente comprensible, pues si $l_1(x) \neq 1$ o bien $l_2(y) \neq 1$ entonces la función g(x,y) no permite reconstruir la función de densidad conjunta, ya que la información que aportan los factores $l_1(x)$ y $l_2(y)$, se ha perdido al derivar, y no se recupera con la integración.

En el caso contrario, es decir cuando $l_1(x) = l_2(y) = 1$ las funciones generadoras de las distribuciones marginales, quedan de la forma

$$g_y(x,y) = \int^x g(x,y)\,dx \qquad (2.30)$$

$$g_x(x,y) = \int^y g(x,y)\,dy \qquad (2.31)$$

y por tanto la densidad conjunta será:

$$f(x,y) = K\,h(x,y) \propto e^{\int^x\int^y g(x,y)\,dx\,dy}$$

donde, evidentemente es

$$K = \left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^x\int^y g(x,y)\,dx\,dy}\,dx\,dy\right]^{-1}$$

Desde el punto de vista formal, se puede interpretar que existen factores en la densidad "responsables" de la relación entre ambas variables y otros factores exclusivos de cada una de ellas, pues se observa que:

1° Si no hay factor de interrelación, es decir si $h(x,y) = 1$, se trata de independencia.

2° Si son los factores $l_1(x)$ $l_2(y)$ los que valen uno, entonces $f(x,y) = K\,h(x,y)$, que es el caso desarrollado en la proposición 2.8

2.1.11 Relación entre la función g(x,y) y el sistema de funciones generadoras $\{g_x(x,y),\ g_y(x,y)\}$

a) Como ya se ha expuesto, en el caso más general, es decir aquel en el que existen factores de interrelación e individualizados, $g(x,y)$ no es función generadora, porque el sistema de generadoras condicionadas no es recuperable desde $g(x,y)$; en este caso las funciones generadoras de las distribuciones marginales, (2.30) y (2.31) quedan de la forma:

$$g_y(x,y) = \int^x g(x,y)\,dx + \frac{d}{dy} \ln l_2(y) \qquad (2.32)$$

$$g_x(x,y) = \int^y g(x,y)\,dy + \frac{d}{dx} \ln l_1(x) \qquad (2.33)$$

y estas últimas expresiones nos muestran que un sistema de generadoras equivalente al $g_x(x,y)$, $g_y(x,y)$ de condicionadas, es:

$$g(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln h(x,y)$$

$$n_1(x) = \frac{d}{dx} \ln l_1(x)$$

$$n_2(y) = \frac{d}{dy} \ln l_2(y)$$

donde necesariamente se ha de verificar:

$$\frac{\partial}{\partial y} \int^x g(x,y)\,dx = \frac{\partial}{\partial x} \int^y g(x,y)\,dy$$

b) Si hay independencia, evidentemente $g(x,y) = 0$ y el sistema generador queda reducido a:

$$n_1(x) = g_x(x,y) = g(x)$$

$$n_2(y) = g_y(x,y) = g(y)$$

c) Si no hay factores individualizados, estudiado en la proposición 2.8, es $n_1(x) = 0$; $n_2(y) = 0$, y por tanto:

$$g_y(x,y) = \int^x g(x,y)\,dx$$

$$g_x(x,y) = \int^y g(x,y)\,dy$$

2.2 FUNCIONES GENERADORAS DE UNA DISTRIBUCIÓN DISCRETA BIVARIANTE DE PROBABILIDAD.

2.2.1 Introducción.

El sistema de distribuciones bivariantes discretas de tipo Pearson, análogo al estudiado por van Uven en el caso continuo, es estudiado por Herrerías - Cobos (1984) [21] y Herrerías - Calvete (1985) [22], planteando el sistema de ecuaciones en diferencias finitas que lo define, mostrando distribuciones que pertenecen a él, y encontrando una solución general basada en las relaciones de recurrencia de los momentos factoriales descendentes.

La generalización del sistema de Pearson discreto bivariante la realiza Fajardo (1985), [13]; este autor considera el sistema formado por dos ecuaciones en diferencias parciales de primer orden:

$$\left.\begin{array}{l} G(r,s)\ p_{r+1,s} - M(r,s)\ p_{r,s} = 0 \\ H(r,s)\ p_{r,s+1} - N(r,s)\ p_{r,s} = 0 \end{array}\right\} \tag{2.34}$$

donde

$$M,\ N,: Z^{+}xZ^{+} \rightarrow \mathbb{R}\ ,\ G,\ H,: Z^{+}xZ^{+} \rightarrow \mathbb{R} - \{0\}$$

son funciones dadas y $p: Z^{+}xZ^{+} \rightarrow \mathbb{R}$ la función incógnita.

Demuestra que si se verifica:

$$\frac{M(r,s+1)}{G(r,s+1)}\cdot\frac{N(r,s)}{H(r,s)} = \frac{N(r+1,s)}{H(r+1,s)}\cdot\frac{M(r,s)}{G(r,s)} \tag{2.35}$$

para todo r y s tal que p_{rs} es distinto de cero, p es solución del sistema anterior.

Obtiene la solución general, que resulta ser:

$$
p_{r,s} = \begin{cases} p_{0,0} \prod_{i=0}^{r-1} \prod_{j=0}^{s-1} \dfrac{M(i,s)}{G(i,s)} \dfrac{N(0,j)}{H(0,j)} & r \geq 1 \ , \quad s \geq 1 \\ p_{0,0} \prod_{i=0}^{r-1} \dfrac{M(i,0)}{G(i,0)} & r \geq 1 \ , \quad s = 0 \\ p_{0,0} \prod_{j=0}^{s-1} \dfrac{N(0,j)}{H(0,j)} & r = 0 \ , \quad s \geq 1 \\ p_{0,0} & r = 0 \ , \quad s = 0 \end{cases}
$$

actuando $p_{0,0}$ como constante de normalización.

2.2.2 Funciones generadoras de la distribución discreta bivariante. Definiciones.

Los contenidos que aparecen a continuación difieren de los proporcionados por Fajardo en su aspecto formal, con objeto de darle un mismo tratamiento que en el caso continuo, desarrollado anteriormente, y destacar así el papel de las distribuciones condicionadas como generadoras de la distribución bivariante.

El procedimiento seguido es análogo al caso continuo, es decir a partir del sistema de generadoras condicionadas, conocer las condiciones necesarias para la obtención de la probabilidad conjunta. Este desarrollo supone darle entidad propia a los cocientes:

$$
\frac{p_{r,s}}{p_{r-1,s}} \quad y \quad \frac{p_{r,s}}{p_{r,s-1}}
$$

Dada una distribución de probabilidad discreta, conocidas las probabilidades $p_{r,s}$ $\forall\ r,s$, a través del cálculo de probabilidades se dispone de las distribuciones condicionadas; puesto que estas distribuciones son univariantes, se pueden definir sus generadoras $g_r(x_r, y_s)$, $g_s(x_r, y_s)$ de la forma:

$$g_r(x_r,y_s) = \frac{\Delta_r\, p(r-1/s)}{p(r-1/s)} = \frac{p(r/s) - p(r-1/s)}{p(r-1/s)} \qquad (2.36)$$

definida para r>0 y para todo s

$$g_s(x_r,y_s) = \frac{\Delta_s\, p(s-1/r)}{p(s-1/r)} = \frac{p(s/r) - p(s-1/r)}{p(s-1/r)} \qquad (2.37)$$

definida para s>0 y para todo r

De igual forma que en el caso continuo, la siguiente proposición nos va a permitir la reconstrucción directa de la distribución conjunta a partir de las dos generadoras de las distribuciones condicionadas, sin el concurso de las generadoras marginales, pues establece una relación entre las generadoras condicionadas y la probabilidad conjunta.

PROPOSICIÓN 2.9

Dada una distribución de probabilidad $p_{r,s}$, las funciones generadoras de las densidades condicionadas se pueden expresar como

$$g_r(x_r,y_s) = \frac{\Delta_r\, p_{r-1,s}}{p_{r-1,s}}$$

$$g_s(x_r,y_s) = \frac{\Delta_s\, p_{r,s-1}}{p_{r,s-1}}$$

en completa analogía al caso continuo.

En efecto, pues de (2.36) y (2.37) se deduce:

$$g_r(x_r,y_s) = \frac{\dfrac{p_{r,s}}{p_s} - \dfrac{p_{r-1,s}}{p_s}}{\dfrac{p_{r-1,s}}{p_s}} = \frac{p_{r,s} - p_{r-1,s}}{p_{r-1,s}} = \frac{\Delta_r\, p_{r-1,s}}{p_{r-1,s}} = \frac{p_{r,s}}{p_{r-1,s}} - 1$$

$$g_s(x_r.y_s) = \frac{\Delta_s\, p_{r,s-1}}{p_{r,s-1}} = \frac{p_{r,s} - p_{r,s-1}}{p_{r,s-1}} = \frac{p_{r,s}}{p_{r,s-1}} - 1$$

Obsérvese que se mantiene así la unicidad de criterios al pasar del univariante al bivariante y del continuo al discreto.

Sin embargo, y puesto que el protagonismo en la generación lo tienen los cocientes:

$$\frac{p_{r,s}}{p_{r-1,s}} \qquad y \qquad \frac{p_{r,s}}{p_{r,s-1}},$$

parece oportuno definir las funciones L, a las que también llamaremos generadoras condicionadas, de la forma:

$$L_r(r,s) = \frac{p_{r,s}}{p_{r-1,s}} \quad \forall\, r>0 \quad \forall s\geq 0 \tag{2.38}$$

$$L_s(r,s) = \frac{p_{r,s}}{p_{r,s-1}} \quad \forall\, s>0 \quad \forall r\geq 0 \tag{2.39}$$

Las relaciones entre las funciones g y L, como ya ocurriera en el caso univariante, son biunívocas:

$$L_r(r,s) = g_r(x_r,y_s) + 1$$

$$L_s(r,s) = g_s(x_r,y_s) + 1$$

2.2.3 Relaciones entre las funciones generadoras.

Las siguientes proposiciones establecen un método que, siendo trivialmente necesario, resultará también suficiente para la generación de la distribución conjunta.

PROPOSICIÓN 2.10

La probabilidad $p_{r,s}$ se puede expresar, a través de las funciones generadoras de una de las siguientes formas:

$$p_{r,s} = p_{0,0} \prod_{i=1}^{r} L_i(i,0) \prod_{j=1}^{s} L_j(r,j) \tag{2.40}$$

$$p_{r,s} = p_{0,0} \prod_{j=1}^{s} L_j(0,j) \prod_{i=1}^{r} L_i(i,s) \tag{2.41}$$

En efecto, las definiciones de las funciones generadoras, (2.38) y (2.39),

$$L_r(r,s) = \frac{p_{r,s}}{p_{r-1,s}} \quad \forall\, r>0 \quad \forall s\geq 0$$

$$L_s(r,s) = \frac{p_{r,s}}{p_{r,s-1}} \quad \forall\, s>0 \quad \forall r\geq 0$$

utilizando una relación de recurrencia, permiten escribir:

$$p_{r,s} = p_{0,s} \prod_{i=1}^{r} L_i(i,s) \tag{2.42}$$

$$p_{r,s} = p_{r,0} \prod_{j=1}^{s} L_j(r,j) \tag{2.43}$$

Ahora bien, haciendo s=0 en (2.42) se tiene:

$$p_{r,0} = p_{0,0} \prod_{i=1}^{r} L_i(i,0)$$

y sustituyendo en (2.43) se obtiene la expresión (2.40)

Análogamente para (2.41).

PROPOSICIÓN 2.11

Dada una distribución de probabilidad, p_{rs}, para todo r>0 y para todo s>0 se verifica la igualdad:

$$\prod_{i=1}^{r} L_i(i,0) \prod_{j=1}^{s} L_j(r,j) = \prod_{j=1}^{s} L_j(0,j) \prod_{i=1}^{r} L_i(i,s) \tag{2.44}$$

Resulta evidentemente, puesto que el valor de p_{rs} ha de coincidir para las igualdades (2.40) y (2.41), se igualan los segundos miembros y dividiendo después por p_{00} queda la igualdad.

Es decir, la igualdad (2.44) se verifica si y sólo si se trata de la misma distribución.

2.2.4 Generación de $p_{r,s}$. Sistema de funciones generadoras.

En las siguientes proposiciones se establecen las condiciones necesarias y suficientes para que un par de funciones sean generadoras de una distribución discreta bidimensional.

La siguiente proposición, 2.12, incluye un método para la construcción de la probabilidad buscada.

PROPOSICIÓN 2.12

La condición necesaria y suficiente para que el sistema de funciones $\{L_1(r,s), L_2(r,s)\}$ definidas de $\mathbb{N}^2$ en $\mathbb{R}^+$ sea generador de una distribución bivariante discreta de probabilidad, $p_{r,s}$, es que

$$\prod_{i=1}^{r} L_1(i,s) \cdot \prod_{j=1}^{s} L_2(0,j) = \prod_{j=1}^{s} L_2(r,j) \cdot \prod_{i=1}^{r} L_1(i,0) \qquad \textbf{(2.46)}$$

y que la suma

$$\sum_{r,s} \prod_{i=1}^{r} L_1(i,s) \prod_{j=1}^{s} L_2(r,j) + \sum_{r} \prod_{i=1}^{r} L_1(i,0) + \sum_{s} \prod_{j=1}^{s} L_2(0,j) \qquad \textbf{(2.47)}$$

sea finita, para todo r y s mayores o iguales que uno.

En efecto, es condición necesaria, ya que dada $p_{r,s}$ para todo r y para todo s, entonces existen dos funciones $L_r(r,s)$ y $L_s(r,s)$, definidas en (2.38) y (2.39) y las proposiciones 2.10 y 2.11 permiten asegurar que se verifica la igualdad (2.46), y que la suma es finita.

Es suficiente, pues dadas las funciones $\{L_1(r,s), L_2(s,r)\}$ se define:

$$K_{r0} = \prod_{i=1}^{r} L_1(i,0) \qquad \forall\, r \geq 1$$

$$K_{0s} = \prod_{j=1}^{s} L_2(0,j) \qquad \forall\, s \geq 1$$

mientras que para los valores de K_{rs} se podrían dar dos definiciones:

$$K_{rs} = \prod_{i=1}^{r} L_1(i,s)\ K_{0,s} = \prod_{i=1}^{r} L_1(i,s) \prod_{j=1}^{s} L_2(0,j) \qquad \forall\, r \geq 1,\ s \geq 1$$

$$K'_{rs} = \prod_{j=1}^{s} L_2(r,j)\ K_{r,0} = \prod_{j=1}^{s} L_2(r,j) \prod_{i=1}^{r} L_1(i,o) \qquad \forall\, r \geq 1,\ s \geq 1$$

Pero por la igualdad (2.46), se verifica que $K_{rs} = K'_{rs}$

lo que nos permite llamar

$$K_{00} = 1 + \left\{ \sum_{r,\,s} K_{r,s} + \sum_{r} K_{r0} + \sum_{s} K_{0s} \right\}$$

siempre que las sumas sean finitas,

y así definir

$$p_{rs} = \frac{K_{rs}}{K_{00}} \qquad \text{con } r>0 \text{ o } s>0$$

con lo que la constante de normalización es $p_{00} = \dfrac{1}{k_{00}}$

PROPOSICIÓN 2.13

Las condición (2.35), dada por Fajardo, [13], y la condición (2.46), dada en la proposición 2.12, sobre la necesidad y la suficiencia para la generación de una distribución discreta bivariante, son equivalentes.

En efecto si en (2.35), llamamos:

$$L_1(r,s) = \frac{M(r,s+1)}{G(r,s+1)}$$

$$L_2(r,s) = \frac{N(r+1,s)}{H(r+1,s)}$$

la igualdad se escribe:

$$L_1(r,s)\ L_2(r-1,s) = L_2(r,s)\ L_1(r,s-1)$$

que dejando fijo s y particularizando para distintos valores de r, queda

$$L_1(1,s)\ L_2(0,s) = L_2(1,s)\ L_1(1,s-1)$$

$$L_1(2,s)\ L_2(1,s) = L_2(2,s)\ L_1(2,s-1)$$

....

Multiplicando miembro a miembro y simplificando factores comunes resulta

$$L_2(0,s) \prod_{i=1}^{r} L_1(i,s) = L_2(r,s) \prod_{i=1}^{r} L_1(i,s-1)$$

Particularizando ahora para distintos valores de s

$$L_2(0,1) \prod_{i=1}^{r} L_1(i,1) = L_2(r,1) \prod_{i=1}^{r} L_1(i,0)$$

$$L_2(0,2) \prod_{i=1}^{r} L_1(i,2) = L_2(r,2) \prod_{i=1}^{r} L_1(i,1)$$

....

volviendo a multiplicar miembro a miembro y simplificando se obtiene

$$\prod_{j=1}^{s} L_2(0,j) \prod_{i=1}^{r} L_1(i,s) = \prod_{j=1}^{s} L_2(r,j) \prod_{i=1}^{r} L_1(i,0)$$

que es la condición (2.46).

Desarrollando los productos de (2.46), sustituyendo las funciones L por sus valores, queda una expresión de la que se desprende que la igualdad (2.35) es cierta para todo r entero positivo y para todo s también entero positivo.

Es necesario destacar que con el método de recurrencia lo que, en cierto modo, hacemos es reproducir para el caso discreto la metodología mediante la cual se obtiene la función potencial correspondiente a una diferencial total, [38], y que incluso la condición necesaria y suficiente para la generación (2.35) ó (2.46), es la correspondiente a la condición de Schwartz exigida en el caso continuo, como se pone de manifiesto en las siguientes proposiciones.

PROPOSICIÓN 2.14

La condición necesaria y suficiente para que el sistema de funciones $\{ L_1(r,s),\ L_2(r,s) \}$ **definidas de** $\mathbb{N}^2$ **en** $\mathbb{R}^+$ **sea generador de una distribución bivariante discreta de probabilidad,** $p_{r,s}$ **, es que**

$$L_1(r,s)\ p_{r-1\ s} = L_2(r,s)\ p_{r\ s-1} \tag{2.48}$$

para todo r y s mayores que uno.

En efecto, comprobaremos que las relaciones (2.46) y (2.48) son equivalentes.

a) Si la igualdad (2.46) es cierta, entonces multiplicando ambos miembros por p_{00}, obtenido por el procedimiento indicado, queda:

$$\prod_{i=1}^{r} L_1(i,s) \cdot \prod_{j=1}^{s} L_2(0,j)\ p_{00} = \prod_{j=1}^{s} l_2(r,j) \cdot \prod_{i=1}^{r} l_1(i,0)\ p_{00}$$

es decir:

$$\prod_{i=1}^{r} L_1(i,s)\ p_{0s} = \prod_{j=1}^{s} L_2(r,j)\ p_{r0}$$

y por tanto:

$$L_1(r,s) \left[\prod_{i=1}^{r-1} L_1(i,s)\ p_{0s} \right] = L_2(r,s) \left[\prod_{j=1}^{s-1} L_2(r,j)\ p_{r0} \right]$$

de donde se obtiene la igualdad:

$$L_1(r,s)\ p_{r-1\ s} = L_2(r,s)\ p_{r\ s-1}$$

b) Si para todo r,s>1 se verifica:

$$L_1(r,s)\ p_{r-1\ s} = L_2(r,s)\ p_{r\ s-1}$$

entonces

$$\left[\prod_{i=1}^{r-1} L_1(i,s) \cdot \prod_{j=1}^{s} L_2(0,j)\ p_{00} \right] L_1(r,s) =$$

$$= \left[\prod_{j=1}^{s-1} L_2(r,j) \cdot \prod_{i=1}^{r} L_1(i,0)\ p_{00} \right] L_2(r,s)$$

es decir:

$$\prod_{i=1}^{r} L_1(i,s) \cdot \prod_{j=1}^{s} L_2(0,j)\ p_{00} = \prod_{j=1}^{s} L_2(r,j) \cdot \prod_{i=1}^{r} L_1(i,0)\ p_{00}$$

que es tanto como afirmar que se verifica la igualdad (2.48).

El paralelismo entre la igualdad de derivadas segundas cruzadas, y la expresión (2.48), al cambiar derivadas por cocientes incrementales, Goldberg [18], se pone de manifiesto en la siguiente proposición:

PROPOSICIÓN 2.15 (SOBRE LAS DIFERENCIAS FINITAS).

Dada una función $z: \mathbb{R}^2 \to \mathbb{R}^+$, $z = z(x,y)$, **se verifica que:**

$$\frac{1}{z(x,\Delta y)} \cdot \Delta_x \left[\frac{\Delta_y\, z(x,y)}{z(x,y)}\right] = \frac{1}{z(\Delta x,y)} \cdot \Delta_y \left[\frac{\Delta_x\, z(x,y)}{z(x,y)}\right]$$

En efecto:

$$\frac{\Delta_y\, z(x,y)}{z(x,y)} = \frac{z(x,\Delta y) - z(x,y)}{z(x,y)}$$

$$\Delta_x \left[\frac{\Delta_y\, z(x,y)}{z(x,y)}\right] = \frac{z(\Delta x,\Delta y) - z(\Delta x,y)}{z(\Delta x,y)} - \frac{z(x,\Delta y) - z(x,y)}{z(x,y)}$$

$$\Delta_x \left[\frac{\Delta_y\, z(x,y)}{z(x,y)}\right] = \frac{z(\Delta x,\Delta y)\ z(x,y) - z(\Delta x,y)\ z(x,\Delta y)}{z(\Delta x,y)\ z(x,y)}$$

De igual forma se tiene:

$$\Delta_y \left[\frac{\Delta_x\, z(x,y)}{z(x,y)}\right] = \frac{z(\Delta x,\Delta y)\ z(x,y) - z(\Delta x,y)\ z(x,\Delta y)}{z(x,\Delta y)\ z(x,y)}$$

de donde resulta evidente que

$$\frac{1}{z(x,\Delta y)} \cdot \Delta_x \left[\frac{\Delta_y\, z(x,y)}{z(x,y)}\right] = \frac{1}{z(\Delta x,y)} \cdot \Delta_y \left[\frac{\Delta_x\, z(x,y)}{z(x,y)}\right] =$$

$$= \frac{\begin{vmatrix} z(x,y) & z(x,\Delta y) \\ z(\Delta x,y) & z(\Delta x,\Delta y) \end{vmatrix}}{z(x,y)\ z(\Delta x,y)\ z(x,\Delta y)}$$

2.2.5 La forma factorizada discreta.

De igual forma que para el caso continuo, en aquellas variables aleatorias discretas bidimensionales cuya función de cuantía admita un expresión analítica, la factorización es trivialmente cierta [1], pues se puede escribir p(r,s) en su aspecto mas general como:

$$p(r,s) = K\,h(r,s)\; l_1(r)\; l_2(s) \tag{2.50}$$

pudiéndose expresar las distribuciones marginales de la forma:

$$p(r,+) = K\, l_1(r) \sum_s h(r,s)\; l_2(s)$$

$$p(+,s) = K\, l_2(s) \sum_r h(r,s)\; l_1(r)$$

y por tanto las condicionadas son:

$$p(r/s) = \frac{h(r,s)\; l_1(r)}{\sum_r h(r,s)\; l_1(r)} \tag{2.51}$$

$$p(s/r) = \frac{h(r,s)\; l_2(s)}{\sum_s h(r,s)\; l_2(s)} \tag{2.52}$$

Las funciones generadoras de las distribuciones marginales quedan de la forma

$$L(r) = \frac{p(r,+)}{p(r-1,+)} = \frac{l_1(r)}{l_1(r-1)} \tag{2.53}$$

$$L(s) = \frac{p(+,s)}{p(+,s-1)} = \frac{l_2(s)}{l_2(s-1)} \tag{2.54}$$

y las funciones generadoras de las distribuciones condicionadas:

---

(1) La factorización de la función de cuantía siempre es posible, aún cuando la distribución esté dada por una tabla, considerando entonces que las distintas funciones están definidas punto a punto.

$$L_r(r,s) = \frac{p(r,s)}{p(r-1,s)} = \frac{h(r,s)\ l_1(r)}{h(r-1,s)\ l_1(r-1)} = \frac{h(r,s)}{h(r-1,s)} L(r) \tag{2.55}$$

$$L_s(r,s) = \frac{p(r,s)}{p(r,s-1)} = \frac{h(r,s)\ l_2(s)}{h(r,s-1)\ l_2(s-1)} = \frac{h(r,s)}{h(r,s-1)} L(s) \tag{2.56}$$

y llamando:

$$H_r(r,s) = \frac{h(r,s)}{h(r-1,s)}$$

$$H_s(r,s) = \frac{h(r,s)}{h(r,s-1)}$$

las expresiones (2.55) y (2.56) se transforman en:

$$L_r(r,s) = H_r(r,s)\ L(r)$$

$$L_s(r,s) = H_s(r,s)\ L(s)$$

PROPOSICIÓN 2.16

Toda la información sobre la distribución conjunta de una variable aleatoria discreta bidimensional, salvo la constante de normalización, está contenida en la pareja de distribuciones condicionadas, $p(r/s)$ y $p(s/r)$

En efecto, de forma análoga a como ocurre en las distribuciones continuas, el cociente $\frac{p(s/r)}{p(r/s)}$ permite obtener las distribuciones marginales, salvo la constante de normalización.

$$\frac{p(s/r)}{p(r/s)} = \frac{\dfrac{l_2(s)}{\sum_s h(r,s)\ l_2(s)}}{\dfrac{l_1(r)}{\sum_r h(r,s)\ l_1(r)}} = \frac{l_2(s) \sum_r h(r,s)\ l_1(r)}{l_1(r) \sum_s h(r,s)\ l_2(s)} =$$

$$= \frac{K\ l_2(s) \sum_r h(r,s)\ l_1(r)}{K\ l_1(r) \sum_s h(r,s)\ l_2(s)} = \frac{p(s)}{p(r)}$$

Para separar una marginal de otra basta llevar al numerador todos los elementos del cociente que dependan de s y al denominador los que dependan de r. Solo faltará encontrar K por la normalización, pudiéndose obtener por separado para cada una de las distribuciones marginales, aunque el resultado ha de ser el mismo.

Esta operación es siempre factible porque el componente no factorizable se simplifica en el cociente.

2.2.6 Recurrencia en las funciones de cuantía condicionadas, bajo la factorización.

PROPOSICIÓN 2.17

Las distribuciones condicionadas se obtienen por separado, cada una de su correspondiente ecuación en diferencias, que constituyen el sistema Fajardo [13] de ecuaciones en diferencias parciales finitas.

En efecto: de las expresiones (2.55) y (2.56) se deducen las relaciones:

$$h(r,s)\ l_1(r) = h(r-1,s)\ l_1(r-1).L_r(r,s) \qquad (2.57)$$

$$h(r,s)\ l_2(s) = h(r,s-1)\ l_2(s-1).L_s(r,s) \qquad (2.58)$$

usando recurrencia sobre los segundos miembros, es decir sustituyendo se puede llegar a las expresiones:

$$h(r,s)\ l_1(r) = h(0,s)\ l_1(0) \prod_{i=1}^{r} L_i(i,s) \qquad r>0 \qquad (2.59)$$

$$h(r,s)\ l_2(s) = h(r,0)\ l_2(0) \prod_{j=1}^{s} L_j(r,j) \qquad s>0 \qquad (2.60)$$

si tenemos en cuenta (2.51) y (2.52), se puede escribir:

$$p(r/s) = K_1\ h(r,s)\ l_1(r) \qquad (2.61)$$

$$p(s/r) = K_2\ h(r,s)\ l_2(s) \qquad (2.62)$$

las cuatro últimas relaciones nos permiten expresar p(r/s) en función de los valores iniciales y usando recurrencia:

$$p(r/s) = \begin{cases} K_1\, h(0,s)\, l_1(0) \prod_{i=1}^{r} L_i(i,s) & r > 0 \\ K_1\, h(0,s)\, l_1(0) & r = 0 \end{cases} \qquad (2.63)$$

y análogamente para la otra condicionada:

$$p(s/r) = \begin{cases} K_2\, h(r,0)\, l_2(0) \prod_{j=1}^{s} L_j(r,j) & s > 0 \\ K_2\, h(r,0)\, l_2(0) & s = 0 \end{cases} \qquad (2.64)$$

y calculando las constantes por normalización resulta ser:

$$p(r/s) = \begin{cases} \dfrac{\prod_{i=1}^{r} L_i(i,s)}{1 + \sum_{r} \prod_{i=1}^{r} L_i(i,s)} & r > 0 \\[2ex] \dfrac{1}{1 + \sum_{r} \prod_{i=1}^{r} L_i(i,s)} & r = 0 \end{cases} \qquad (2.65)$$

$$p(s/r) = \begin{cases} \dfrac{\prod_{j=1}^{s} L_j(r,j)}{1 + \sum_{s} \prod_{j=1}^{s} L_j(r,j)} & s > 0 \\[2ex] \dfrac{1}{1 + \sum_{s} \prod_{j=1}^{s} L_j(r,j)} & s = 0 \end{cases} \qquad (2.66)$$

Reconstruidas las probabilidades condicionadas, la proposición 2.16, permite obtener las marginales de forma sencilla y, a través de ella, la distribución conjunta.

Obviamente, se está resolviendo el "sistema Fajardo", por esta otra vía, que muestra el protagonismo de las distribuciones condicionadas y demuestra que toda la información sobre la distribución conjunta se encuentra en ellas y que es aplicable al caso k-variante, como posteriormente veremos en el estudio de la independencia estocástica a través de la independencia analítica de las generadoras.

En el caso particular de que las variables sean independientes, puesto que la distribución conjunta ha de factorizarse en producto de marginales, entonces $h(r,s) = 1 \quad \forall\, r, s$ y las distribuciones marginales toman la forma:

$$p(r/s) = \frac{l_1(r)}{\sum_r l_1(r)}$$

$$p(s/r) = \frac{l_2(s)}{\sum_s l_2(r)}$$

con lo que las funciones L de las marginales coinciden con las condicionadas, como se esperaba:

$$L_r(r,s) = \frac{l_1(r)}{l_1(r-1)} = L(r)$$

$$L_s(r,s) = \frac{l_2(s)}{l_2(s-1)} = L(s)$$

De donde se siguen las siguientes expresiones de recurrencia:

$$l_1(r) = l_1(r-1)\, L(r)$$

$$l_2(s) = l_2(s-1)\, L(s)$$

Ejemplo. La distribución trinomial.

$$p(y/x) = \binom{n-x}{y}\left(\frac{p_2}{1-p_1}\right)^y\left(1-\frac{p_2}{1-p_1}\right)^{n-x-y} =$$

$$= \binom{n-x}{y}\frac{p_2^y\ (1-p_1-p_2)^{n-x-y}}{(1-p_1)^{n-x}}$$

$$p(x/y) = \binom{n-y}{x}\left(\frac{p_1}{1-p_2}\right)^x\left(1-\frac{p_1}{1-p_2}\right)^{n-x-y} =$$

$$= \binom{n-y}{x}\frac{p_1^x\ (1-p_1-p_2)^{n-x-y}}{(1-p_2)^{n-y}}$$

y por tanto el cociente es

$$\frac{p(x/y)}{p(y/x)} = \frac{\binom{n-y}{x} p_1^x(1-p_1)^{n-x}}{\binom{n-x}{y} p_2^y\ (1-p_2)^{n-y}} = \frac{\frac{(n-y)!}{x!}}{\frac{(n-x)!}{y!}}\ \frac{p_1^x\ (1-p_1)^{n-x}}{p_2^y\ (1-p_2)^{n-y}} =$$

$$= \frac{\frac{1}{x!\ (n-x)!}}{\frac{1}{y!\ (n-y)!}}\ \frac{p_1^x\ (1-p_1)^{n-x}}{p_2^y\ (1-p_2)^{n-y}} = \frac{\binom{n}{x} p_1^x(1-p_1)^{n-x}}{\binom{n}{y} p_2^y\ (1-p_2)^{n-y}}$$

de donde $p(x) = \binom{n}{x} p_1^x\ (1-p_1)^{n-x}$

$$p(y) = \binom{n}{y} p_2^y\ (1-p_2)^{n-y}$$

La tabla siguiente muestra los sitemas de funciones generadoras de algunas distribuciones bivariantes discretas de probabilidad

Distribución	Generadoras $\begin{cases} L_r(r,s) \\ L_s(r,s) \end{cases}$
Trinomial (n, p_1, p_2)	$\frac{n-r-s+1}{r} \frac{p_1}{1-p_1-p_2}$ $\frac{n-r-s+1}{s} \frac{p_2}{1-p_1-p_2}$
Hipergeométrica	$\frac{n-r-s+1}{r} \frac{N_p-r+1}{N-N_p-N_q-n+r+s}$ $\frac{n-r-s+1}{s} \frac{N_q-s+1}{N-N_p-N_q-n+r+s}$
Poisson	$\frac{\lambda_1-\lambda_3}{r} \frac{\sum_{i=0}^{\max(r,s)} \frac{r^{(i)}}{(\lambda_1-\lambda_3)^i} \frac{s^{(i)}}{(\lambda_2-\lambda_3)^i} \frac{\lambda_3^i}{i!}}{\sum_{i=0}^{\max(r,s)} \frac{(r-1)^{(i)}}{(\lambda_1-\lambda_3)^i} \frac{s^{(i)}}{(\lambda_2-\lambda_3)^i} \frac{\lambda_3^i}{i!}}$ $\frac{\lambda_2-\lambda_3}{s} \frac{\sum_{i=0}^{\max(r,s)} \frac{r^{(i)}}{(\lambda_1-\lambda_3)^i} \frac{s^{(i)}}{(\lambda_2-\lambda_3)^i} \frac{\lambda_3^i}{i!}}{\sum_{i=0}^{\max(r,s)} \frac{r^{(i)}}{(\lambda_1-\lambda_3)^i} \frac{(s-1)^{(i)}}{(\lambda_2-\lambda_3)^i} \frac{\lambda_3^i}{i!}}$
Binomial negativa	$\frac{(r+s+k-1)\, p_1}{r}$ $\frac{(r+s+k-1)\, p_2}{s}$
Logaritmica	$\frac{(r+s-1)\, p_1}{r}$ $\frac{(r+s-1)\, p_2}{s}$

2.2.7 Las funciones g(r,s) y L(r,s).

Siguiendo con el paralelismo entre distribuciones continuas y discretas, corresponde ahora definir las funciones g(r,s) y L(r,s). A la vista de la proposición 2.15, se define de una de las dos formas siguientes:

$$g(r,s) = \frac{\Delta_s\, g_r(r,s-1)}{p_{r,s-1}}$$

$$g(r,s) = \frac{\Delta_r\, g_s(r-1,s)}{p_{r-1,s}}$$

para r mayor que cero y para s mayor que cero.

Fácilmente se puede comprobar que coinciden.

$$g(r,s) = \frac{1}{p_{r,s-1}}\,\Delta_s\left(\frac{p_{r,s-1}}{p_{r-1,s-1}} - 1\right) = \frac{1}{p_{r,s-1}}\left(\frac{p_{r,s}}{p_{r-1,s}} - \frac{p_{r,s-1}}{p_{r-1,s-1}}\right)$$

$$g(r,s) = \frac{1}{p_{r-1,s}}\,\Delta_r\left(\frac{p_{r-1,s}}{p_{r-1,s-1}} - 1\right) = \frac{1}{p_{r-1,s}}\left(\frac{p_{r,s}}{p_{r,s-1}} - \frac{p_{r-1,s}}{p_{r-1,s-1}}\right)$$

La función g(r,s) también se puede expresar de la forma:

$$g(r,s) = \frac{\begin{vmatrix} p_{r-1,s-1} & p_{r-1,s} \\ p_{r,s-1} & p_{r,s} \end{vmatrix}}{p_{r-1,s-1}\, p_{r,s-1}\, p_{r-1,s}} =$$

$$= \frac{p_{r-1,s-1}\, p_{r,s} - p_{r,s-1}\, p_{r-1,s}}{p_{r-1,s-1}\, p_{r-1,s}\, p_{r,s-1}} = \frac{1}{p_{r-1,s-1}}\left(\frac{p_{r,s}\, p_{r-1,s-1}}{p_{r-1,s}\, p_{r,s-1}} - 1\right) \qquad (2.67)$$

Teniendo en cuenta que $L_r(r,s) = \frac{p_{r,s}}{p_{r-1,s}}$, parece lógico definir

$$L(r,s) = \frac{L_r(r,s)}{L_r(r,s-1)} \qquad (2.68)$$

que desarrollando es:

$$L(r,s) = \frac{L_r(r,s)}{L_r(r,s-1)} = \frac{\dfrac{p_{r,s}}{p_{r-1,s}}}{\dfrac{p_{r,s-1}}{p_{r-1,s-1}}} = \frac{p_{r,s}\, p_{r-1,s-1}}{p_{r-1,s}\, p_{r,s-1}}$$

Obsérvese que la definición

$$L(r,s) = \frac{L_s(r,s)}{L_s(r-1,s)}$$

es equivalente a (2.68)

Entre las funciones g(r,s) y L(r,s) existe la relación

$$g(r,s) = \frac{1}{p_{r-1,s-1}}\,(\,L(r,s) - 1) \qquad (2.69)$$

para todo r y s mayores que cero
que coincide con las establecidas para las funciones g y L en todos los casos anteriores.

2.3 INDEPENDENCIA DE VARIABLES ALEATORIAS

Son varias las condiciones de caracterización de la independencia de variables aleatorias recogidas en la literatura estadística, [28] [41], y para su aplicación tienen, en todas ellas, que calcularse las funciones marginales de distribución, o de densidad, o de cuantía, según sean del tipo continuo o discreto, o bien hay que determinar las funciones características o generatrices de momentos de las variables implicadas tanto marginal como conjuntamente, lo que conlleva cálculos de integrales o sumatorios que no tienen por qué ser rápidos ni simples.

Para establecer la condición de independencia de variables aleatorias se tendrá en cuenta el hecho de que siempre que la distribución condicionada de una variable dependa, en la forma funcional o en el intervalo de definición, de la otra variable, entonces las variables son dependientes; sin embargo cuando en una distribución condicionada no interviene otra variable, ni explícitamente en la función de densidad o de cuantía ni implícitamente en el intervalo de definición, las variables son independientes [45].

2.3.1 Condición de independencia para variables aleatorias continuas.

Supuesto que la región admisible de definición de las variables aleatorias continuas, sea tal que el dominio de definición de una variable aleatoria no dependa de el de la otra, se pueden enunciar el siguiente teorema y corolarios.

TEOREMA 2.1

Una condición necesaria y suficiente para que dos variables aleatorias continuas, con función de densidad conjunta $f(x,y)$, sean independientes es que:

$$g(x,y) = 0 \tag{2.70}$$

Para su demostración se comprobará que la condición anterior, (2.70), es equivalente a la factorización de la función de densidad conjunta por las funciones de densidad marginales, es decir:

$$f(x,y) = f_1(x) \,.\, f_2(y) \tag{2.71}$$

En efecto:

a) Si las variables aleatorias son independientes se verifica (2.71), y como consecuencia $\ln f(x,y) = \ln f_1(x) + \ln f_2(y)$ luego evidentemente se verifica que $g(x,y) = 0$

Recíprocamente, si $g(x,y) = 0$ entonces, utilizando la forma factorizada se tiene que

$$g(x,y) = \frac{\partial^2}{\partial y\, \partial x} \ln h(x,y) = 0$$

por tanto la función $\ln h(x,y)$ si no es constante, dependerá sólo de una de las dos variables.

Teniendo en cuenta que los factores que dependen sólo de x están englobados en $l_1(x)$, los que sólo dependen de y en $l_2(y)$ y los valores numéricos en la constante, se deduce que el único valor posible para $h(x,y)$ es uno.

Al sustituir ahora en la forma factorizada, la función $f(x,y)$ es igual a la constante por una función que sólo depende de x y por otra función que sólo depende de y, con lo que queda demostrado que las variables X e Y son independientes.

Se puede encontrar la solución de $f(x,y)$ en términos de las funciones generadoras.

b) Si $g(x,y) = 0$ entonces $g_x(x,y)$ es de la forma

$$g_x(x,y) = \frac{\partial}{\partial x} \ln f(x,y) = n_1(x) \tag{2.72}$$

pues $g(x/y)$ es una función que no depende de y, ya que al derivarla con respecto a y, esta derivada vale cero.

De igual forma

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(x,y) = n_2(y) \tag{2.73}$$

ya que $g(y/x)$ no depende de x, pues al derivar con respecto a x, vale cero.

De (2.72), integrando en x, se deduce que

$$\ln f(x,y) = \int^x g_x(x,y)\,dx + C_2(y) = \int^x n_1(x)\,dx + C_2(y) \tag{2.74}$$

derivando esta última expresión, (2.74), respecto a y, y teniendo en cuenta (2.73), resulta que:

$$\frac{\partial}{\partial y} \ln f(x,y) = n_2(y) = C_2'(y)$$

y por tanto

$$C_2(y) = \int^y n_2(y)\,dy + C \tag{2.75}$$

luego (2.74) se puede escribir como

$$\ln f(x,y) = \int^x n_1(x)\,dx + \int^y n_2(y)\,dy + C$$

de donde

$$f(x,y) = K\, e^{\int^x n_1(x)\,dx + \int^y n_2(y)\,dy} \tag{2.76}$$

siendo

$$k = \left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^x n_1(x)\,dx + \int^y n_2(y)\,dy}\,dx\,dy\right]^{-1} =$$

$$= \left[\int_{-\infty}^{+\infty} e^{\int^x n_1(x)\,dx}\,dx\right]^{-1} \cdot \left[\int_{-\infty}^{+\infty} e^{\int^y n_2(y)\,dy}\,dy\right]^{-1}$$

por tanto (2.76) queda:

$$f(x,y) = \left[\int_{-\infty}^{+\infty} e^{\int^x n_1(x)\,dx}\,dx\right]^{-1} e^{\int^x n_1(x)\,dx} \left[\int_{-\infty}^{+\infty} e^{\int^y n_2(y)\,dy}\,dy\right]^{-1} e^{\int^y n_2(y)\,dy} \tag{2.77}$$

y como de (2.77) se obtiene:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \left[\int_{-\infty}^{+\infty} e^{\int^x n_1(x)\,dx}\,dx\right]^{-1} e^{\int^x n_1(x)\,dx}$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \left[\int_{-\infty}^{+\infty} e^{\int^y n_2(y)\,dy}\,dy\right]^{-1} e^{\int^y n_2(y)\,dx},$$

se puede afirmar que $f(x,y) = f_1(x) \,.\, f_2(y)$

COROLARIO 2.1

En las mismas condiciones del teorema 2.1, es condición necesaria y suficiente para que las variables X e Y sean independientes, que la función $g_x(x,y)$ no dependa de y.(2.78)

Demostración:

Si las variables aleatorias X e Y son independientes, entonces se verifica que $\ln f(x,y) = \ln f_1(x) + \ln f_2(y)$ derivando con respecto a x se obtiene una función que no depende de y

Recíprocamente, si la función $g_x(x,y)$ no depende de y, entonces al hacer su derivada con respecto a y, $\frac{\partial}{\partial y} g_x(x,y)$, es cero y por el teorema 2.1, las variables X e Y son independientes.

COROLARIO 2.2

En las mismas condiciones del teorema, es condición necesaria y suficiente para que las variables X e Y sean independientes, que la función $g_y(x,y)$ no dependa de x.(2.79)

La demostración es idéntica a la del corolario 2.1

2.3.2 Ejemplos de determinación de la independencia para variables aleatorias continuas.

En los ejemplos que siguen se quiere mostrar la sencillez que ofrece este método para decidir sobre la independencia o no de dos variables aleatorias:

1) Sea (X,Y) la variable aleatoria bidimensional cuya función de densidad conjunta es:

$$f(x,y) = \frac{4}{5}(x+3y)\,e^{-(x+2y)} \qquad \forall\, x,y > 0$$

Entonces:

$$\ln f(x,y) = \ln \frac{4}{5} + \ln(x+3y) - (x+2y)$$

$$g_x(x,y) = \frac{\partial}{\partial x} \ln f(x,y) = \frac{1}{x+3y} - 1$$

y puesto que la función $g_x(x,y)$ sí depende de y, las variables aleatorias X e Y son dependientes.

2) Sea (X,Y) la variable aleatoria bidimensional cuya función de densidad conjunta es:

$$f(x,y) = \frac{1}{4} + \frac{xy}{4}(x^2 - y^2) \qquad \text{si } |x| \le 1 \text{ , } |y| \le 1$$

Entonces:

$$\ln f(x,y) = \ln(1 + x^3y - xy^3) - \ln 4$$

$$g_x(x,y) = \frac{\partial}{\partial x} \ln f(x,y) = \frac{3x^2y - y^3}{1 + x^3y - xy^3}$$

y puesto que la función $g_x(x,y)$ sí depende de y, las variables aleatorias X e Y son dependientes.

2.3.3 Condición de independencia para variables aleatorias discretas.

Supuesto que el dominio de definición de las variables aleatorias sea tal que $\forall x_r$, $\forall y_s$ la función de cuantía conjunta $P(x_r, y_s)$ sea estrictamente positiva, se pueden enunciar los lemas, teorema y corolarios siguientes:

LEMA 2.1

Sean X e Y dos variables aleatorias discretas, independientes, entonces $L_r(r,s) = \dfrac{p_{r,s}}{p_{r-1,s}}$ **no depende de s, para todo r.**

En efecto, puesto que las variables aleatorias X e Y son independientes resulta que $p(y_s/x_r) = p(y_s/x_t) \quad \forall r,s,t$, luego, en particular $p(y_s/x_r) = p(y_s/x_{r-1}) \quad \forall r,s$, de donde se deduce que el cociente

$$\frac{p_{r,s}}{p_{r-1,s}} = \frac{p_r}{p_{r-1}} \tag{2.80}$$

es independiente de s.

LEMA 2.2

Sean X e Y dos variables aleatorias discretas, independientes, entonces $L_s(r,s) = \dfrac{p_{r,s}}{p_{r,s-1}}$ **no depende de r, para todo s.**

La demostración se realiza igual forma que en el lema 2.1, obteniéndose

$$\frac{p_{r,s}}{p_{r,s-1}} = \frac{p_s}{p_{s-1}} \tag{2.81}$$

TEOREMA 2.2

Una condición necesaria y suficiente para que dos variables aleatorias discretas X_r, Y_s, de recorridos $0 \le r \le R \quad 0 \le s \le S$, **sean independientes es que**

$$\begin{cases} \Delta_s \dfrac{\Delta_r\, p_{r-1,s}}{p_{r-1,s}} = 0 & \forall r,\ s \\[2ex] \Delta_r \dfrac{\Delta_s\, p_{r,s-1}}{p_{r,s-1}} = 0 & \forall r,\ s \end{cases} \tag{2.82}$$

Demostración:

$$\Delta_s \frac{\Delta_r p_{r-1,s}}{p_{r-1,s'}} = \Delta_s \frac{p_{r,s} - p_{r-1,s}}{p_{r-1,s}} =$$

$$= \Delta_s \left[\frac{p_{r,s}}{p_{r-1,s}} - 1 \right] = \Delta_s \frac{p_{r,s}}{p_{r-1,s}} \qquad (2.83)$$

pero si las variables aleatorias son independientes, aplicando el lema 2.1, la expresión (2.83) es $\Delta_s \dfrac{p_r}{p_{r-1}} = 0$.

De forma análoga se comprueba, utilizando el lema 2.2, que si las variables son independientes entonces

$$\Delta_r \frac{\Delta_s p_{r,s-1}}{p_{r,s-1}} = 0$$

Recíprocamente si se verifica (2.82), las variables aleatorias discretas son independientes.

Efectivamente, en tal caso:

i) $\dfrac{\Delta_r p_{r-1,s}}{p_{r-1,s}}$ sólo depende de r $\Rightarrow$ $\dfrac{p_{r,s}}{p_{r-1,s}} = g_1(x_r)$

ii) $\dfrac{\Delta_s p_{r,s-1}}{p_{r,s-1}}$ sólo depende de s $\Rightarrow$ $\dfrac{p_{r,s}}{p_{r,s-1}} = g_2(y_s)$

con lo que se dispone del siguiente sistema de ecuaciones en diferencias parciales de primer orden:

$$p_{r,s} - g_1(x_r)\, p_{r-1,s} = 0 \quad \forall\, r \geq 1 \;\; y \;\; \forall\, s$$

$$p_{r,s} - g_2(y_s)\, p_{r,s-1} = 0 \quad \forall\, s \geq 1 \;\; y \;\; \forall\, r$$

y por recurrencia:

$$p_{r,s} = p_{0,s} \prod_{r=1}^{R} g_1(x_r) \qquad (2.84)$$

$$p_{r,s} = p_{r,0} \prod_{s=1}^{S} g_2(y_s) \qquad (2.85)$$

particularizando (2.84) para s = 0,

$$p_{r,0} = p_{0,0} \prod_{r=1}^{R} g_1(x_r)$$

y sustituyendo en (2.85), se obtiene para $p_{r,s}$ la expresión:

$$p_{r,s} = \begin{cases} p_{0,0} \prod\limits_{r=1}^{R} g_1(x_r) \prod\limits_{s=1}^{S} g_2(y_s) & r\geq 1 \;\; s\geq 1 \\ p_{0,0} \prod\limits_{r=1}^{R} g_1(x_r) & r\geq 1 \;\; s=0 \\ p_{0,0} \prod\limits_{s=1}^{S} g_2(y_s) & r=0 \;\; s\geq 1 \end{cases} \tag{2.86}$$

actuando $p_{0,0}$ como constante normalizadora.

Teniendo en cuenta que $\sum\limits_{r=0}^{R} \sum\limits_{s=0}^{S} p_{r,s} = 1$, resulta

$$p_{0,0}\left[1 + \sum_{r=1}^{R}\prod_{i=1}^{r} g_1(x_i) + \sum_{s=1}^{S}\prod_{j=1}^{s} g_2(y_j) + \right.$$

$$\left. + \left(\sum_{r=0}^{R}\prod_{i=1}^{r} g_1(x_i)\right)\left(\sum_{s=1}^{S}\prod_{j=1}^{s} g_2(y_j)\right)\right] = 1$$

por tanto

$$p_{0,0} = \frac{1}{1 + \sum\limits_{r=1}^{R} \prod\limits_{i=1}^{r} g_1(x_i)} \cdot \frac{1}{1 + \sum\limits_{s=1}^{S} \prod\limits_{j=1}^{s} g_2(y_j)}$$

$$p_{r,s} = \frac{\prod\limits_{r=1}^{R} g_1(x_r)}{1 + \sum\limits_{r=1}^{R} \prod\limits_{i=1}^{r} g_1(x_i)} \cdot \frac{\prod\limits_{s=1}^{S} g_2(y_s)}{1 + \sum\limits_{s=1}^{S} \prod\limits_{j=1}^{s} g_2(y_j)} \qquad r\geq 1 \quad s\geq 1$$

$$p_{r,0} = \frac{\prod\limits_{r=1}^{R} g_1(x_r)}{\left[1 + \sum\limits_{r=1}^{R} \prod\limits_{i=1}^{r} g_1(x_i)\right]\left[1 + \sum\limits_{s=1}^{S} \prod\limits_{j=1}^{s} g_2(y_j)\right]} \qquad r\geq 1$$

$$p_{0,s} = \frac{\prod_{s=1}^{S} g_2(y_s)}{\left[1 + \sum_{r=1}^{R} \prod_{i=1}^{r} g_1(x_i)\right]\left[1 + \sum_{s=1}^{S} \prod_{j=1}^{s} g_2(y_j)\right]} \qquad s \geq 1$$

Calculando además las marginales:

$$p_{r+} = \sum_{s=0}^{S} p_{r,s} = p_{r,0} + \sum_{s=1}^{S} p_{r,s}$$

$$p_{r+} = \frac{\prod_{r=1}^{R} g_1(x_r)}{1 + \sum_{r=1}^{R} \prod_{i=1}^{r} g_1(x_i)} \qquad r \geq 1$$

$$p_{+s} = \frac{\prod_{s=1}^{S} g_2(y_s)}{1 + \sum_{s=1}^{S} \prod_{j=1}^{s} g_2(y_j)} \qquad s \geq 1$$

$$p_{0+} = \sum_{s=0}^{S} p_{0,s} = p_{0,0} + \sum_{s=1}^{S} p_{0,s} = \frac{1}{1 + \sum_{r=1}^{R} \prod_{i=1}^{r} g_1(x_i)}$$

$$p_{+0} = \frac{1}{1 + \sum_{s=1}^{S} \prod_{j=1}^{s} g_2(y_j)}$$

Obsérvese que se verifica:

$p_{0+}\, p_{+0} = p_{00}$

$p_{0+}\, p_{+s} = p_{0s}$ con $s>0$

$p_{r+}\, p_{+0} = p_{r0}$ con $r>0$

$p_{r+}\, p_{+s} = p_{rs}$ con $r>0,\ s>0$

Por tanto las variables aleatorias son independientes.

COROLARIO 2.3

En las mismas condiciones del teorema 2.2, es condición necesaria y suficiente para que las variables X e Y sean independientes que se verifiquen:

i) La función $L_r(r,s)$ no dependa de s **(2.87)**

ii) La función $L_s(r,s)$ no dependa de r **(2.88)**

Demostración:

Si las variables aleatorias son independientes, entonces las condiciones (2.87) y (2.88) se verifican sin más que aplicar los lemas 2.1 y 2.2

Recíprocamente, si $\dfrac{p_{r,s}}{p_{r-1,s}}$ no depende de s, al aplicarle Δ_s su valor será cero, pero

$$\Delta_s \frac{p_{r,s}}{p_{r-1,s}} = \Delta_s \left[\frac{p_{r,s}}{p_{r-1,s}} - 1 \right] = \Delta_s \frac{\Delta_r p_{r-1,s}}{p_{r-1,s}} = 0 \qquad (2.89)$$

Si $\dfrac{p_{r,s}}{p_{r,s-1}}$ no depende de r, de igual modo, se obtiene

$$\Delta_r \frac{\Delta_s p_{r,s-1}}{p_{r,s-1}} = 0 \qquad (2.90)$$

Luego co. (2.89) y (2.90) se verifica la condición (2.82) y aplicando ahora el teorema 2.2, las variables X e Y son independientes.

COROLARIO 2.4

En las mismas condiciones del teorema 2.2, es condición necesaria y suficiente para que las variables X e Y sean independientes que se verifiquen:

i) La función $g_r(r,s)$ no dependa de s **(2.91)**

ii) La función $g_s(r,s)$ no dependa de r **(2.92)**

La demostración es consecuencia del corolario 2.3 y de la relación entre las funciones $g_r(r,s)$ y $L_r(r,s)$ por una parte y $g_s(r,s)$ y $L_s(r,s)$ por la otra.

TEOREMA 2.3

Una condición necesaria y suficiente para que dos variables aleatorias discretas X_r, Y_s, de recorridos $0 \le r \le R$ $0 \le s \le S$, **sean independientes es que** $g(r,s) = 0$

En efecto, si las variables son independientes entonces, según el corolario 2.4, la función $g(r/s)$ no depende de s y por tanto

$$g(r,s) = \frac{\Delta_s\, g_r(r,s-1)}{p_{r,s-1}} = \frac{1}{p_{r,s-1}}\, \Delta_s\, g_r(r,s-1) = 0$$

Recíprocamente, si $g(r,s) = 0$ entonces según (2.67)

$$g(r,s) = \frac{\begin{vmatrix} p_{r-1,s-1} & p_{r-1,s} \\ p_{r,s-1} & p_{r,s} \end{vmatrix}}{p_{r-1,s-1}\, p_{r,s-1}\, p_{r-1,s}} = 0$$

por tanto el determinante es cero

$$\begin{vmatrix} p_{r-1,s-1} & p_{r-1,s} \\ p_{r,s-1} & p_{r,s} \end{vmatrix} = 0$$

lo que significa que

$$p_{r-1,s-1}\, p_{r,s} - p_{r-1,s}\, p_{r,s-1} = 0$$

$$\frac{p_{r,s}}{p_{r-1,s}} = \frac{p_{r,s-1}}{p_{r-1,s-1}}$$

$$L_r(r,s) = L_r(r,s-1)$$

y puesto que ocurre para todo r y para todo s, todas las distribuciones condicionadas son iguales, coinciden con las marginales, y las variables son independientes.

COROLARIO 2.5

Una condición necesaria y suficiente para que dos variables aleatorias discretas X_r, Y_s, de recorridos $0 \leq r \leq R$ $0 \leq s \leq S$, **sean independientes es que**

$$L(r,s) = 1$$

Resulta evidente a la vista del teorema 2.3 y de la relación existente entre las funciones L(r,s) y g(r,s), (2.69).

CAPÍTULO 3

GENERACIÓN DE DISTRIBUCIONES MULTIVARIANTES DE PROBABILIDAD

El presente capítulo está subdividido en cuatro apartados: la generación de una distribución multivariante continua, la generación en el caso discreto, el estudio de la independencia entre subvectores y el tratamiento que damos a los parámetros en la generación de las distribuciones.

Los resultados del capítulo 2 se pueden extender de una forma sencilla al caso n-dimensional. Se mantienen los conceptos de funciones generadoras condicionadas, los sistemas y las condiciones que estos deben cumplir.

El tratamiento de las distribuciones multivariantes continuas se realiza siguiendo las directrices dadas por Steyn (1960), [43], al realizar la extensión del bivariante al multivariante a través de un sistema de igual número de ecuaciones que de variables componen el vector aleatorio.

Como novedad, con respecto al bidimensional, estudiamos la generación de un subvector condicionado a las restantes variables que forman el vector. Resulta importante para el desarrollo de la presente memoria, el hecho de que el sistema generador de la distribución condicionada sea un subsistema del generador de la densidad conjunta.

Se establecen las relaciones entre las generadoras marginales y condicionadas.

Para vectores aleatorios discretos se generaliza el caso bidimensional discreto y se siguen las directrices del multidimensional continuo.

Se pone de manifiesto que la independencia estocástica se puede observar a partir de la dependencia analítica de las funciones generadoras de las distintas variables.

En este caso multidimensional estudiamos la independencia de subvectores aleatorios, obtenidos de una partición del vector aleatorio, empezando por el caso de que la partición está constituida por sólo dos elementos hasta el caso de que cada elemento de la partición posea una sola variable aleatoria. En los teoremas correspondientes, 3.1 y 3.2, se demuestra, dos subvectores son estocásticamente independientes si y solo si las funciones del subsistema generador de un subvector no poseen variables comunes con las funciones que constituyen el subsistema generador del otro subvector.

Si en al menos una función de cada uno de los subsistemas correspondientes a dos subvectores existe una variable común que no pertenece a ninguno de los subvectores, estamos en el caso de independencia condicionada, en el sentido dado por A.P. Dawid [10].

Por último, si en al menos una función de cada uno de los subsistemas correspondientes a dos subvectores existe una variable común que pertenece a uno de los dos subvectores, estamos en el caso de dependencia estocástica entre los subvectores.

La coincidencia de enunciados, en el estudio de la independencia, entre los casos continuo y discreto, al referirse a las funciones generadoras, es debido a la propiedad, común, de cancelar las variables.

El último apartado de este capítulo muestra el tratamiento que a través de la generación puede darse a los parámetros de la función de verosimilitud.

3.1 GENERACIÓN DE UNA DISTRIBUCIÓN DE PROBABILIDAD CONTINUA n-DIMENSIONAL.

3.1.1 Extensión del caso bidimensional al n-dimensional.

Notación.

Buscando una mayor claridad, se introducen las siguientes notaciones:

* El conjunto $N = \{1,2,\ldots,n\}$, cuyos elementos son los números naturales comprendidos entre 1 y n, ambos inclusive.
* I es cualquier subconjunto de N; $I \subset N$
* $\wp(N)$ es el conjunto de las partes de N
* $\vec{x} \equiv (x_1, x_2, \ldots, x_n)$, vector del espacio n-dimensional.
* $\vec{x}_I$, subvector formado por todas las coordenadas del vector $\vec{x}$ que ocupan el lugar que corresponden a los elementos del conjunto I.
* $\vec{x}_{N-I}$, subvector formado por todas las coordenadas del vector $\vec{x}$ excepto las que ocupan el lugar que corresponden a los elementos del conjunto I.
* En particular, $\vec{x}_{N-\{i\}}$, subvector de $\vec{x}$, elemento de un espacio vectorial de dimensión n-1, formado por todas las coordenadas del vector $\vec{x}$ menos la iésima y que, siempre que no haya lugar a duda, se va denotar $\vec{x}_i$. Por tanto $\vec{x}_i \equiv (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.
* $\frac{\partial}{\partial I} \ln f(\vec{x})$ representa a la derivada parcial de orden el cardinal de I, con respecto una sola vez a cada una de las variables cuyo subíndice es un elemento de I.

Definición de función generadora.

Dada una densidad conjunta $f(\vec{x})$, conocemos las funciones univariantes condicionadas,

$$\{ f(x_i/\vec{x}_i) \}_{i=1,\ldots,n}$$

Para cada una de ellas existe la correspondiente función generadora unidimensional definida como

$$g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(x_i/\vec{x}_i)$$

La proposición 2.1 sigue siendo válida y por tanto la función generadora condicionada se puede escribir a partir de la densidad conjunta, es decir para todo i=1,...,n se verifica

$$g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(\vec{x})$$

Sistema de funciones generadoras.

Un conjunto de n funciones de $\mathbb{R}^n$ en $\mathbb{R}$, derivables

$$\{ g_1(\vec{x}), \ldots, g_n(\vec{x}) \}$$

constituye un sistema de funciones generadoras de una función de densidad conjunta de una distribución n-variante si y sólo si, se cumple la condición

$$\frac{\partial}{\partial x_j} g_i(\vec{x}) = \frac{\partial}{\partial x_i} g_j(\vec{x}) \qquad \forall\ i \neq j$$

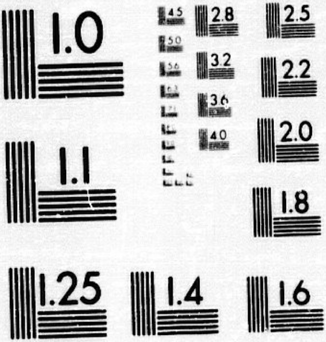
Conviene resaltar que, dado el sistema de generadoras $\{ g_i(\vec{x}) \}_{i=1,\ldots,n}$, todas y cada una de ellas son funciones reales de n variables reales, y la iésima componente es la función generadora de la distribución de la variable aleatoria X_i condicionada por las n-1 restantes.

La densidad conjunta es por tanto

$$f(\vec{x}) \propto K\, e^{U(\vec{x})}$$

donde $U(\vec{x})$ es la función potencial de la diferencial total:

$$dU(\vec{x}) = \sum_{i=1}^{n} g_i(\vec{x})\, dx_i$$



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

generadora unidimensional definida como

$$g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(x_i/\vec{x}_i)$$

La proposición 2.1 sigue siendo válida y por tanto la función generadora condicionada se puede escribir a partir de la densidad conjunta, es decir para todo i=1,...,n se verifica

$$g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(\vec{x})$$

Sistema de funciones generadoras.

Un conjunto de n funciones de $\mathbb{R}^n$ en $\mathbb{R}$, derivables

$$\{ g_1(\vec{x}), \ldots, g_n(\vec{x}) \}$$

constituye un sistema de funciones generadoras de una función de densidad conjunta de una distribución n-variante si y sólo si, se cumple la condición

$$\frac{\partial}{\partial x_j} g_i(\vec{x}) = \frac{\partial}{\partial x_i} g_j(\vec{x}) \qquad \forall\, i \neq j$$

Conviene resaltar que, dado el sistema de generadoras $\{ g_i(\vec{x}) \}_{i=1,\ldots,n}$, todas y cada una de ellas son funciones reales de n variables reales, y la iésima componente es la función generadora de la distribución de la variable aleatoria X_i condicionada por las n-1 restantes.

La densidad conjunta es por tanto

$$f(\vec{x}) \propto K\, e^{U(\vec{x})}$$

donde $U(\vec{x})$ es la función potencial de la diferencial total:

$$dU(\vec{x}) = \sum_{i=1}^{n} g_i(\vec{x})\, dx_i$$

3.1.2 Generación de un subvector condicionado.

En el estudio de las distribuciones n-variantes surgen nuevos elementos, que en el caso bidimensional no se dan. Se trata de las distribuciones multivariantes condicionadas de un subvector con respecto a otro, $f(\vec{x}_I/\vec{x}_{N-I})$.

Se hace necesario conocer el sistema de generadoras que le corresponde a estas distribuciones y su relación con el sistema de generadoras del vector completo.

Evidentemente no tienen sentido las situaciones en que I coincide bien con el conjunto vacío o bien con el conjunto N. De igual forma, si I es unitario, es el caso de la generación de una variable condicionada por todas las demás.

PROPOSICIÓN 3.1

El sistema de generadoras de la distribución multivariante dada por $f(\vec{x}_I/\vec{x}_{N-I})$ **es el subsistema de las generadoras de** $f(\vec{x})$ **constituido por los elementos correspondientes a las coordenadas de** $\vec{x}_I$

En efecto:

Sea $f(\vec{x})$ la densidad conjunta de las n variables aleatorias $X_1, \ldots, X_n$.

$f(\vec{x}_I/\vec{x}_{N-I})) = \dfrac{f(\vec{x})}{f(\vec{x}_{N-I})}$ es la función de densidad conjunta de una distribución multivariante y como tal tiene su sistema de generadoras, que notaremos:

$$\{ g_i(\vec{x}_I/ \vec{x}_{N-I}) \}_{i\in I}$$

que después de definidas en forma habitual y en virtud de la proposición 2.1 se puede escribir:

$$g_i(\vec{x}_I/ \vec{x}_{N-I}) = \frac{\partial}{\partial x_i} \ln f(\vec{x}_I/\vec{x}_{N-I})$$

para todo elemento del conjunto I.

De forma análoga, utilizando la mencionada proposición 2.1, resulta que

$$\frac{\partial}{\partial x_i} \ln f(\vec{x}_I/\vec{x}_{N-I}) = \frac{\partial}{\partial x_i} \ln f(\vec{x}) - \frac{\partial}{\partial x_i} \ln f(\vec{x}_{N-I})$$

para todo elemento de I.

Puesto que x_i es una coordenada del vector $\vec{x}_I$ y no del vector $\vec{x}_{N-I}$, se tiene que:

$$\frac{\partial}{\partial x_i} \ln f(\vec{x}_{N-I}) = 0$$

por lo que para $\forall\, i \in I$, la generadora se puede expresar:

$$g_i(\vec{x}_I/\vec{x}_{N-I}) = \frac{\partial}{\partial x_i} \ln f(\vec{x})$$

lo que nos permite escribir la siguiente identidad:

$$g_i(\vec{x}) \equiv g_i(\vec{x}_I/\vec{x}_{N-I}) \qquad \forall\ i \in I \tag{3.1}$$

La importancia de esta identidad radica en que dado un vector aleatorio multivariante: $\vec{X} = (X_1, \ldots, X_n)$, cuyo sistema de generadoras de la densidad es $\{ g_i(\vec{x}) \}_{i=1\ldots n}$, la distribución de un subvector $(X_{(1)}, \ldots, X_{(p)})$ de p componentes cualquiera, condicionado a las restantes n-p componentes, está generado por el subsistema:

$$\{ g_{(i)}(\vec{x}) \}_{i=1\ldots p}$$

3.1.3 Sistema de funciones generadoras de la distribución marginal de un subvector. Relación con las funciones generadoras de las distribuciones condicionadas.

De igual forma que en el capítulo anterior, se dan dos proposiciones que relacionan las funciones generadoras de las distribuciones marginales y las condicionadas de una misma distribución n-dimensional. Será también necesario suponer que se verifican las condiciones de regularidad ya mencionadas para el caso bidimensional.

PROPOSICIÓN 3.2

Dada una densidad conjunta, $f(\vec{x})$, las funciones generadoras de la densidad marginal de un subvector $\vec{X}_I$ y de una componente, se pueden expresar respectivamente como:

$$g_i(\vec{x}_I) = g_i(\vec{x}) - \frac{\partial}{\partial x_i} \ln f(\vec{x}_{N-I}/\vec{x}_I) \quad \forall\, i \in I \subset N \tag{3.2}$$

$$g(x_i) = g_i(\vec{x}) - \frac{\partial}{\partial x_i} \ln f(\vec{x}_i/x_i) \quad \forall\, i \in N \tag{3.3}$$

En efecto, si tenemos en cuenta que

$\forall\, I \subset N$ se verifica $f(\vec{x}_{N-I}/\vec{x}_I) = \dfrac{f(\vec{x})}{f(\vec{x}_I)}$,

sólo hay que tomar logaritmos neperianos, aplicar las definiciones de función generadora y ordenar oportunamente para obtener la expresión (3.2).

La expresión (3.3) es el caso particular de (3.2) en el que $I = \{i\}$.

Esta proposición nos da una relación entre las generadoras marginales y las condicionadas de la misma distribución, en la que se observa que las igualdades $g(x_i) = g_i(\vec{x})$ $\forall\, i \in N$ se verifican cuando las variables aleatorias X_i $\forall\, i \in N$ son independientes en su totalidad.

Dado un vector aleatorio $\vec{X}_N$ y un subvector $\vec{X}_I$, $I \subset N$, sabemos que el subsistema de generadoras de la densidad $f(\vec{x}_I/\vec{x}_{N-I})$ es $\{ g_i(\vec{x}) \}_{i \in I}$.

La siguiente proposición establece la forma del sistema de generadoras de la distribución marginal $f(\vec{x}_I)$:

PROPOSICIÓN 3.3

Si $\{ g_i(\vec{x}_I) \}_{i \in I}$ es el sistema de generadoras de la distribución marginal $f(\vec{x}_I)$, entonces, bajo condiciones de

regularidad, se verifica que:

$$g_i(\vec{x}_I) = E_{f(\vec{x}_{N-I}/\vec{x}_I)}[g_i(\vec{x})] \quad \forall\, i \in I \tag{3.4}$$

En efecto:

Sean χ_N, χ_I, χ_{N-I} los recintos de definición de las variables aleatorias $\vec{X}$, $\vec{X}_I$, $\vec{X}_{N-I}$ respectivamente.

Por definición

$$f(\vec{x}_I) = \int_{\chi_{N-I}} f(\vec{x})\, d\vec{x}_{N-I}$$

$$g_i(\vec{x}_I) = \frac{\partial}{\partial x_i} \ln f(\vec{x}_I) \qquad \forall\, i \in I$$

ahora bien:

$$\frac{\partial}{\partial x_i} \ln f(\vec{x}_I) = \frac{\frac{\partial}{\partial x_i} f(\vec{x}_I)}{f(\vec{x}_I)} = \frac{\frac{\partial}{\partial x_i} \int_{\chi_{N-I}} f(\vec{x})\, d\vec{x}_{N-I}}{f(\vec{x}_I)} \tag{3.5}$$

que bajo condiciones de regularidad, (3.5) se puede escribir como

$$\frac{\int_{\chi_{N-I}} \frac{\partial}{\partial x_i} f(\vec{x})\, d\vec{x}_{N-I}}{f(\vec{x}_I)} = \frac{\int_{\chi_{N-I}} \frac{\partial}{\partial x_i} \ln f(\vec{x}) \,.\, f(\vec{x})\, d\vec{x}_{N-I}}{f(\vec{x}_I)} \tag{3.6}$$

puesto que $f(\vec{x}_I)$ no depende de las variables respecto a las que se integra, (3.6) es igual

$$\int_{\chi_{N-I}} \left(\frac{\partial}{\partial x_i} \ln f(\vec{x}) \right) \left(\frac{f(\vec{x})}{f(\vec{x}_I)} \right) d\vec{x}_{N-I} =$$

$$= \int_{\chi_{N-I}} g_i(\vec{x})\, f(\vec{x}_{N-I}/\vec{x}_I)\, d\vec{x}_{N-I} \quad = \quad E_{f(\vec{x}_{N-I}/\vec{x}_I)}[g_i(\vec{x})]$$

En particular, si $I = \{i\}$ entonces

$$g(x_i) = E_{f(\vec{x}_i/x_i)}[g_i(\vec{x})] \tag{3.7}$$

se obtiene la misma expresión a la correspondiente al caso bidimensional, (2.9), que recordemos es:

$$g(x) = E_{f(y/x)}[\, g_x(x,y) \,]$$

3.1.4 Factorización de la función de densidad. Distribuciones marginales y distribuciones condicionadas [1].

Se realiza ahora el tratamiento de distribuciones de probabilidad n-variantes, de forma análoga a la desarrollada en el caso bidimensional, cuando la densidad conjunta está factorizada. Se prestará también atención a la generación de un subvector condicionado a otro subvector de la variable aleatoria n-variante.

Sea una variable aleatoria n-variante, $(X_1, X_2, ..., X_n)$; la función de densidad conjunta se escribirá como producto de todos los factores posibles, esto es:

$$f(\vec{x}) = \prod_{I \in \wp(N)} h(\vec{x}_I) \qquad (3.8)$$

En esta factorización se ha de tener en cuenta que:

a) Se sigue utilizando una sola letra, en este caso h, para designar a las distintas funciones que puedan aparecer.

b) Si $I = \emptyset$ entonces la función h no depende de ninguna variable: se trata de la constante de normalización.

c) Puede ocurrir que varios de estos factores sean igual a uno, por lo que no aparecen en la factorización.

Al aplicar logaritmos neperianos:

$$\ln f(\vec{x}) = \sum_{I \in \wp(N)} \ln h(\vec{x}_I)$$

Y derivando con respecto a las variables cuyos subíndices pertenecen a I,

$$\frac{\partial}{\partial \vec{x}_I} \ln f(\vec{x}) = \sum_{J \in \wp(N),\ I \subset J} \frac{\partial}{\partial \vec{x}_I} \ln h(\vec{x}_J)$$

---

(1) Como en el caso bidimensional, y en tanto no haya lugar a confusión, se va a seguir notando tanto las distribuciones marginales como las condicionadas con la misma letra que la conjunta. Igual comentario para las generadoras.

Por lo que a las funciones generadoras se refiere, sólo hay que hacer I = {i}, para cada i desde 1 hasta n; su expresión es de la forma:

$$g_i(\vec{x}) = \sum_{J\in\wp(N),\ \{i\}\subset J} \frac{\partial}{\partial x_i} \ln h(\vec{x}_J)$$

Para un subvector la función marginal correspondiente tiene la forma:

$$f(\vec{X}_I/\vec{X}_{N-I}) = \frac{\prod_{J\in\wp(N),\ I\subset J} h(\vec{X}_J)}{\int_{X_I} \prod_{J\in\wp(N),\ I\subset J} h(\vec{X}_J)\, d\vec{X}_I} \tag{3.9}$$

y las funciones generadoras de la distribución son:

$$g_i(\vec{X}_I/\vec{X}_{N-I}) = \sum_{J\in\wp(N),\ I\subset J} \frac{\partial}{\partial x_i} \ln h(\vec{X}_J) \quad \forall i \in I$$

La factorización propuesta es la más general posible, si bien se pueden utilizar otras intermedias tales como asociar los factores de acuerdo con la partición elegida. (De este modo los factores resultantes pueden ser producto de otros).

Dado el subvector $\vec{x}_I$, para un I subconjunto de N, queda determinado el subvector complementario $\vec{x}_{N-I}$. En este caso la descomposición factorial se puede realizar como:

$$f(\vec{x}) = K\, h(\vec{x})\, l_1(\vec{x}_I)\, l_2(\vec{x}_{N-I})$$

donde en l_1 aparecen todos los factores que dependen de las variables que forman el subvector $\vec{x}_I$, en l_2 las variables que forman parte de $\vec{x}_{N-I}$, mientras que en h quedan relacionadas las variables de uno y otro subvector.

De esta forma:

$$f(\vec{x}_I) = K\, l_1(\vec{x}_I) \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} h(\vec{x})\, l_2(\vec{x}_{N-I})\, d\vec{x}_{N-I} \tag{3.10}$$

$$f(\vec{x}_I/\vec{x}_{N-I}) = \frac{h(\vec{x})\, l_1(\vec{x}_I)}{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} h(\vec{x})\, l_1(\vec{x}_I)\, d\vec{x}_I}$$

En este último caso, considerando I subconjunto unitario, es decir $\vec{x}_I$ de una sola variable, x_i, para cada i, la factorización se realiza de la siguiente forma:

$$f(\vec{x}) = K\, h(\vec{x})\ l_1(x_i)\ l_2(\vec{x}_i) \tag{3.11}$$

donde en l_1 están recogidos todos los factores que dependen sólo de la variable x_i, en l_2 los factores que no dependen de la variable x_i y por tanto $h(x_1,\ldots,x_n)$ es una función que liga la variable i con cualquier otra u otras distintas de ella.

Supuesta la densidad conjunta factorizada según (3.11), cualquier distribución marginal posee una densidad que tiene la expresión:

$$f(\vec{x}_i) = \int_{-\infty}^{+\infty} f(\vec{x})\, dx_i = K\, l_2(\vec{x}_i) \int_{-\infty}^{+\infty} h(\vec{x})\ l_1(x_i)\, dx_i \tag{3.12}$$

Las distribuciones condicionadas son de la forma:

$$f(x_i/\vec{x}_i) = \frac{h(\vec{x})\ l_1(x_i)}{\int_{-\infty}^{+\infty} h(\vec{x})\ l_1(x_i)\, dx_i} \tag{3.13}$$

puesto que el otro factor se simplifica.

Resulta evidente que

$$f(x_i/\vec{x}_i) \propto h(\vec{x})\ l_1(x_i)$$

y la constante de normalización es

$$K(\vec{x}_i) = \left[\int_{-\infty}^{+\infty} h(\vec{x})\ l_1(x_i)\, dx_i \right]^{-1}$$

3.1.5 Funciones generadoras, supuesta la factorización de la función de densidad conjunta.

La función generadora de la distribución de X_i condicionada al resto de las variables, $\vec{x}_i$, por definición es

$$g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(x_i/\vec{x}_i) =$$

$$= \frac{\partial}{\partial x_i} \ln h(\vec{x}) + \frac{\partial}{\partial x_i} \ln l_1(x_i) = m_i(\vec{x}) + n_i(x_i) \tag{3.14}$$

Se utiliza la misma notación que en el caso bidimensional:

$$\frac{\partial}{\partial x_i} \ln h(\vec{x}) = m_i(\vec{x})$$

$$\frac{\partial}{\partial x_i} \ln l_i(x_i) = n_i(x_i)$$

Razonando en sentido inverso, es evidente que:

$$f(x_i/\vec{x}_i) \propto e^{\int g_i(\vec{x})\, dx_i}$$

$$f(x_i/\vec{x}_i) \quad \alpha \quad e^{\int m_i(\vec{x})\, dx_i + \int n_i(x_i)\, dx_i}$$

y la constante de normalización adquiere ahora la siguiente forma:

$$K(\vec{x}_i) = \left[\int_{-\infty}^{+\infty} e^{\int^{x_i} g_i(\vec{x}) dx_i}\, dx_i \right]^{-1} =$$

$$= \left[\int_{-\infty}^{+\infty} e^{\int^{x_i} m_i(\vec{x}) dx_i + \int^{x_i} n_i(x_i) dx_i}\, dx_i \right]^{-1} =$$

$$= \left[\int_{-\infty}^{+\infty} h(\vec{x})\, l_i(x_i)\, dx_i \right]^{-1} \tag{3.15}$$

y teniendo en cuenta las expresiones (3.12) y (3.15) se puede escribir

$$K(\vec{x}_i) = \frac{K\, l_2(\vec{x}_i)}{f(\vec{x}_i)}$$

Expresión del sistema de funciones generadoras dada la función de densidad conjunta en forma factorizada.

Con la condición necesaria sobre la integrabilidad de las funciones n_i y m_j y la condición de igualdad de derivadas parciales cruzadas:

$$\frac{\partial}{\partial x_j} m_i(\vec{x}) = \frac{\partial}{\partial x_i} m_j(\vec{x}) \qquad \forall\ i \neq j$$

se puede realizar la integración y encontrar así la función primitiva.

En resumen, puesto que para todo $i = 1,2,...,n$ se verifica (3.13), de forma análoga al caso bidimensional, se puede afirmar que toda la información existente sobre $f(\vec{x})$ se encuentra en el conjunto de funciones

$$g_i(\vec{x}) = m_i(\vec{x}) + n_i(x_i) \qquad i=1,2,..,n \tag{3.16}$$

Si atendemos a la factorización en su caso más general, las generadoras condicionadas quedan de la forma:

$$g_i(\vec{x}) = \sum_{J\in\wp(N),\ \{i\}\subset J} \frac{\partial}{\partial x_i} \ln h(\vec{x}_J)$$

y si comparamos esta expresión con (3.16), es obvio que

$$m_i(\vec{x}) = \sum_{J\in\wp(N),\ \{i\}\subset J} \frac{\partial}{\partial x_i} \ln h(\vec{x}_J)$$

siendo J un conjunto de cardinal mayor o igual que dos.

Bajo estas condiciones se puede generar una distribución conjunta de probabilidad:

$$f(\vec{x}) \propto e^{\int^{x_j} m_j(\vec{x})\,dx_j + \sum_{i=1}^{n}\int^{x_i} n_i(x_i)\,dx_i} \tag{3.17}$$

para cualquier $j \in N$, definida en un recinto $D \subset \mathbb{R}^n$, sobre el que

$$\int \cdots \int_D e^{\int^{x_j} m_j(\vec{x})\,dx_j + \sum_{i=1}^{n}\int^{x_i} n_i(x_i)\,dx_i} dx_1 \ldots dx_n < +\infty \tag{3.18}$$

para cualquier j.

La constante de normalización es

$$K = \left[\int \cdots \int_D e^{\int^{x_j} m_j(\vec{x})\,dx_j + \sum_{i=1}^{n}\int^{x_i} n_i(x_i)\,dx_i} dx_1 \ldots dx_n \right]^{-1}$$

De la expresión (3.17) se obtiene la forma factorizada para la función de densidad conjunta, (3.11), sin más que utilizar las siguientes notaciones:

$$h(\vec{x}) = e^{\int^{x_j} m_j(\vec{x})\, dx_j} \qquad (3.19)$$

$$h_i(x_i) = e^{\int^{x_i} n_i(x_i)\, dx_i} \qquad (3.20)$$

resulta evidente que:

$$f(\vec{x}) \propto h(\vec{x})\; l_1(x_1)\ldots l_n(x_n)$$

Es decir, partiendo del sistema de generadoras se llega a expresar la densidad conjunta de la forma (3.11).

3.2 GENERACIÓN DE UNA DISTRIBUCIÓN DE PROBABILIDAD DISCRETA n-DIMENSIONAL.

3.2.1 Introducción.

El método a seguir para la generación de distribuciones discretas n-variantes de probabilidad queda diseñado tanto por la generación de las distribuciones multivariantes continuas, (epígrafe anterior de este capítulo), como la generación de las distribuciones bivariantes discretas (capítulo anterior), es decir, se definen las funciones generadoras de igual modo que en el caso discreto y se sigue la metodología multivariante.

Se pasa después a estudiar la generación de subvectores.

Notación.

A la notación utilizada para las multivariantes continuas se va a añadir la siguiente:

* Vector $\vec{i} = (0...0,1,0..0)$, vector de dimensión n en el que todas sus componentes son iguales a cero excepto la iésima que toma el valor uno.
* Por tanto el vector $r_j \vec{j} = (0,..,0,r_j,0..0)$ tiene todas sus coordenadas cero excepto la de lugar j que es igual a r_j.
* Como ejemplo de notación, en el caso de que $N = \{1,2,3,...8\}$ y $J = \{1,4,5,8\}$ entonces el vector $\vec{r} = \sum_{j \in J} r_j \vec{j}$ es:

$$\vec{r} = (r_1, 0, 0, r_4, r_5, 0, 0, r_8)$$

* Se define la función delta de R^2 en R:

$$\delta_j(r_i, t) = \begin{cases} 0 & i > j \\ t & i = j \\ r_i & i < j \end{cases}$$

siendo $\forall i \in N$, $\forall j \in J \subset N$

* Se define el vector

$\vec{g}_j(t) = (\ \delta_j(r_1,t),\ \delta_j(r_2,t),...,\delta_j(r_n,t)\ ) = (\delta_j(r_i,t))_{i \in N}$

3.2.2 Funciones generadoras de la distribución discreta multivariante.

Dada una distribución de probabilidad discreta, conocidas las probabilidades $p_{\bar{r}}$, se dispone de las distribuciones condicionadas; puesto que estas distribuciones son univariantes, se pueden definir las funciones:

$$g_i(x_{\bar{r}}) = \frac{\Delta_{r_i} p_{\bar{r}-\bar{i}}}{p_{\bar{r}-\bar{i}}} = \frac{p_{\bar{r}} - p_{\bar{r}-\bar{i}}}{p_{\bar{r}-\bar{i}}} \quad \forall r_i > 0 \qquad (3.21)$$

De igual forma que en el caso bivariante se definen las funciones generadoras L, de la forma

$$L_i(\bar{r}) = \frac{p_{\bar{r}}}{p_{\bar{r}-\bar{i}}} \quad \forall r_i > 0 \qquad (3.22)$$

Las relaciones entre las funciones g y L, como ya ocurriera en el caso univariante, son:

$$L_i(\bar{r}) = g_i(x_{\bar{r}}) + 1 \quad \forall i \in I$$

La siguiente proposición permite la reconstrucción de la distribución conjunta a partir de las generadoras de las distribuciones condicionadas, sin el concurso de las generadoras marginales, pues establece una relación entre las generadoras condicionadas y la probabilidad conjunta.

La proposición 2.10 en el caso multivariante queda enunciada como sigue:

PROPOSICIÓN 3.4

La probabilidad $p_{\bar{r}}$ se puede expresar a través de las funciones generadoras de la forma:

$$p_{\sum_{j\in J} r_j \bar{j}} = \prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\bar{g}_j(t)) \right\} p_{\bar{0}} \quad \forall J \in \wp(N) \text{ , } J \neq \emptyset \qquad \mathbf{(3.23)}$$

actuando $p_{\bar{0}}$ como constante normalizadora, cuyo valor es

$$p_{\bar{0}} = \frac{1}{1 + \sum_{J\subset N,\, j\neq\emptyset} p_{\sum_{j\in J} r_j \bar{j}}} =$$

$$= \frac{1}{1 + \sum_{J\subset N,\, j\neq 0} \prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\}}$$

La demostración se realiza de forma análoga al caso bivariante, a partir de las definiciones de las funciones generadoras,

$$L_i(\vec{r}) = \frac{p_{\vec{r}}}{p_{\vec{r}-\vec{1}}} \quad \forall\, r_i > 0$$

utilizando la relación de recurrencia,

$$p_{\vec{r}} = L_i(\vec{r})\, p_{\vec{r}-\vec{1}}$$

se obtiene la relación (3.23).

Esta relación, desarrollada en el caso n = 3, queda:

$$p_{r_1 00} = \prod_{t=1}^{r_1} L_1(t,0,0)\, p_{000}$$

$$p_{0 r_2 0} = \prod_{t=1}^{r_2} L_2(0,t,0)\, p_{000}$$

$$p_{00 r_3} = \prod_{t=1}^{r_3} L_3(0,0,t)\, p_{000}$$

$$p_{r_1 r_2 0} = \prod_{t=1}^{r_2} L_2(r_1,t,0)\, p_{r_1 00} = \prod_{t_2=1}^{r_2} L_2(r_1,t_2,0) \prod_{t_1=1}^{r_1} L_1(t_1,0,0)\, p_{000}$$

$$p_{0 r_2 r_3} = \prod_{t=1}^{r_3} L_3(0,r_2,t)\, p_{0 r_2 0} = \ldots$$

$$p_{r_1 0 r_3} = \prod_{t=1}^{r_3} L_3(r_1,0,t)\, p_{r_1 00} = \ldots$$

$$p_{r_1 r_2 r_3} = \prod_{t=1}^{r_3} L_3(r_1,r_2,t)\, p_{r_1 r_2 0} = \ldots$$

$$p_{r_1 r_2 r_3} = \prod_{t_3=1}^{r_3} L_3(r_1,r_2,t_3) \prod_{t_2=1}^{r_2} L_2(r_1,t_2,0) \prod_{t_1=1}^{r_1} L_1(t_1,0,0)\, p_{000}$$

PROPOSICIÓN 3.5

Dada una distribución de probabilidad, $p_{\vec{r}}$, $\forall \vec{r} \neq \vec{0}$, se verifican las igualdades:

$$\prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\} = \prod_{j\in (J)} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\}$$

siendo (J) una permutación cualquiera de los elementos del conjunto ordenado creciente J.

Resulta evidentemente, puesto que el valor de $p_{\vec{r}}$ ha de coincidir para las distintas posibilidades en las recurrencia.

3.2.3 Generación de $p_{\vec{r}}$. Sistema de funciones generadoras.

La siguiente proposición establece un método para la generación de la distribución conjunta.

PROPOSICIÓN 3.6

La condición necesaria y suficiente para que el sistema de funciones $\{ L_i((\vec{g}_i(t)) \}_{i=1,\dots,n}$ definidas de $\mathbb{N}^n$ en $\mathbb{R}^+$ sea generador de una distribución n-variante discreta de probabilidad, $p_{\vec{r}}$, es que

$$\prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\} = \prod_{j\in (J)} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\}$$

donde (J) sigue siendo cualquier permutación de J, y que la suma

$$\sum_{J\subset N, j\neq \emptyset} \prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\}$$

sea finita, para todo $\vec{r}$ distinto del vector nulo.

En efecto, es condición necesaria, ya que dada la probabilidad existen las funciones generadoras definidas en (3.22) y la proposición 3.2 permite asegurar que se verifica la igualdad y que la suma es finita.

Es suficiente, pues dadas las funciones $L_j(\vec{g}_j(t))$ se define:

$$K_{\sum_{j\in J} r_j \vec{j}} = \prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\} \qquad \forall\, J \in \wp(N)\ ,\ J \neq \emptyset$$

y siempre que la suma

$$\sum_{J\subset N,\, j\neq\emptyset} \prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\}$$

sea finita nos permite llamar

$$K_{\vec{0}} = 1 + \sum_{J\subset N,\, j\neq\emptyset} \prod_{j\in J} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\}$$

y así definir entonces:

$$p_{\vec{r}} = \frac{K_{\vec{r}}}{K_{\vec{0}}} \qquad \text{con } \vec{r} \neq 0$$

con lo que la constante de normalización es $p_{\vec{0}} = [k_{\vec{0}}]^{-1}$

3.3 ESTUDIO DE LA INDEPENDENCIA EN UNA DISTRIBUCIÓN CONTINUA MULTIVARIANTE.

3.3.1 Introducción.

Recordemos que en caso bivariante se definió una función

$$g(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f(x,y)$$

y que las variables X e Y son independientes cuando y sólo cuando $g(x,y)$ sea igual a cero, lo que equivale a que dado el sistema de funciones generadoras de la distribución conjunta, $g_y(x,y)$ es una función sólo de y, (no depende de x) y $g_x(x,y)$ lo es sólo de x, (no depende de y).

De forma análoga, la independencia estocástica de dos subvectores del vector aleatorio $\vec{X}$ puede deducirse a partir de las propiedades analíticas del sistema de generadoras.

En este epígrafe estudiamos primero la independencia estocástica entre dos subvectores $\vec{X}_I$ y $\vec{X}_{N-I}$ de $\vec{X}$, y, como caso particular, la independencia entre una variable X_i y las demás componentes de $\vec{X}$, $\vec{X}_i$.

A continuación se trata el caso en el que los subvectores $\vec{X}_I$ y $\vec{X}_J$ no son complementarios. Se distinguen dos situaciones: por una parte la independencia en el seno de la distribución marginal del vector $(\vec{X}_I,\vec{X}_J)$, y por otra parte la independencia en el seno de la distribución condicionada $(X_I,X_J/X_{N-I-J})$. (Dawid, [10]).
Se concluye con la independencia en su totalidad. (Feller, [14]; Fernández de Trocóniz, [15])

Como ya se hiciera para variables bidimensionales, se supone que la región admisible de definición de las variables aleatorias continuas, sea tal que el dominio de definición de una variable aleatoria no dependa del valor que tomen las otras. (rectangularidad del dominio).

En todos los casos la estructura de independencia se puede observar directamente en el sistema de generadoras de la densidad conjunta.

3.3.2 Independencia entre las variables de $\vec{X}_I$ y de $\vec{X}_{N-I}$

TEOREMA 3.1

Una condición necesaria y suficiente para que los subvectores complementarios $\vec{X}_I$ y $\vec{X}_{N-I}$ sean independientes es que

$$\frac{\partial}{\partial x_j} g_i(\vec{x}) = 0 \qquad \forall\, i \in I \;,\; \forall\, j \in N-I \qquad \textbf{(3.24)}$$

es decir:

$$\frac{\partial^2}{\partial x_i \, \partial x_j} \ln f(\vec{x}) = 0 \qquad \forall\, i \in I \;,\; \forall\, j \in N-I \qquad \textbf{(3.25)}$$

La demostración es similar a la realizada en el caso bivariante en el sentido de que la independencia quedará probada cuando se pueda factorizar la función de densidad conjunta como producto de funciones en las que las variables están separadas. (Fdez. de Trocóniz [15]).

Es condición necesaria pues si $\vec{X}_I$, $\vec{X}_{N-I}$ son independientes entonces $f(\vec{X}) = K\, f(\vec{X}_I)\, f(\vec{X}_{N-I})$

al tomar logaritmos es:

$$\ln f(\vec{X}) = \ln K + \ln f(\vec{X}_I) + \ln f(\vec{X}_{N-I})$$

y al derivar con respecto a una variable x_i del subvector $\vec{X}_I$ queda

$$g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(\vec{X}) = \frac{\partial}{\partial x_i} \ln f(\vec{X}_I)$$

de donde se deduce que

$$\frac{\partial}{\partial x_j} g_i(\vec{x}) = 0 \quad \forall\, i \in I \;,\; \forall\, j \in N-I$$

Resulta evidente entonces que las funciones que constituyen el subsistema generador de la variable $\vec{X}_I$ no dependen de las variables de $\vec{X}_{N-I}$

Pero además las funciones $g_j(\vec{x})$ $\forall\, j \in N-I$ no dependen de las variables del subvector $\vec{X}_I$, pues si $g_j(\vec{x})$ $j \in N-I$ dependiera de x_i entonces la derivada

$$\frac{\partial^2}{\partial x_j\, \partial x_i} \ln f(\vec{x}) \neq 0$$

y por la igualdad de Schwartz sería

$$\frac{\partial^2}{\partial x_j\, \partial x_i} \ln f(\vec{x}) = \frac{\partial^2}{\partial x_i\, \partial x_j} \ln f(\vec{x}) \neq 0$$

que está en contradicción con (3.25).

Es condición suficiente pues si se verifica (3.25) entonces las funciones que componen el subsistema generador de los elementos de $\vec{X}_I$ no dependen de las variables del subvector $\vec{X}_{N-I}$, y las funciones del subsistema generador de $\vec{X}_{N-I}$ tampoco dependen de los elementos de $\vec{X}_I$, es decir:

$\{ g_i(\vec{x}) \}_{i \in I}$ son funciones que sólo dependen de las variables x_i con $i \in I$, mientras que $\{ g_j(\vec{x}) \}_{j \in N-I}$ son funciones que sólo dependen de las variables x_j con $j \in N-I$.

Por otra parte, se sabe que

$$f(\vec{x}) = K\, e^{U(\vec{x})} \quad \text{donde} \quad dU(\vec{x}) = \sum_{l \in N} g_l(\vec{x})\, dx_l$$

y puesto que las funciones del primer sumatorio sólo dependen de las variables de $\vec{X}_I$, y las del segundo de las variables de $\vec{X}_{N-I}$, se puede escribir

$$dU(\vec{x}) = \sum_{i \in I} g_i(\vec{x}_I)\, dx_i + \sum_{j \in N-I} g_j(\vec{x}_{N-I})\, dx_j$$

que con la notación:

$$dU(\vec{X}_I) = \sum_{i\in I} g_i(\vec{X})\, dx_i \ ,$$

$$dU(\vec{X}_{N-I}) = \sum_{j\in N-I} g_j(\vec{X})\, dx_j$$

se tiene que:

$$dU(\vec{X}) = dU(\vec{X}_I) + dU(\vec{X}_{N-I})$$

de donde se deduce que la densidad conjunta se puede factorizar en la forma $f(\vec{X}) = K\, e^{U(\vec{X}_I)}\, e^{U(\vec{X}_{N-I})}$

En particular si $I = \{i\}$, la variable aleatoria X_i es independiente de cualquier otra del vector $\vec{X}$, si

$$\frac{\partial^2}{\partial x_i\, \partial x_j} \ln f(\vec{X}) = 0 \quad \forall\ j \neq i$$

3.3.3 Independencia entre las variables de $\vec{X}_I$ y de $\vec{X}_J$

En el caso de dos subvectores excluyentes, no complementarios, $I \cap J = \emptyset \quad K = N-I-J \neq \emptyset$, la dependencia o independencia entre las variables de uno y otro subvector se puede estudiar bajo la consideración de la densidad marginal $f(\vec{X}_I, \vec{X}_J)$, o bien bajo la densidad condicionada $f(\vec{X}_I, \vec{X}_J / \vec{X}_{N-I-J})$.

Entre las variables $\vec{X}_I$ y $\vec{X}_J$ existe independencia cuando la función de densidad marginal conjunta se puede factorizar de la forma:

$$f(\vec{X}_I, \vec{X}_J) = h(\vec{X}_I)\, h(\vec{X}_J)$$

donde las funciones h son definidas positivas, [15].

TEOREMA 3.2

Dada $f(\vec{X}_I, \vec{X}_J, \vec{X}_K)$, $\vec{X}_I$ $\vec{X}_J$ son independientes si y sólo si en el sistema de generadoras: $\{ g_i(\vec{X})_{i\in I}\ g_j(\vec{X})_{j\in J}\ g_k(\vec{X})_{k\in K} \}$ se verifica que no existe una variable x_l $l \in N$ que figure simultáneamente en una generadora, $g_{i_0}(\vec{X})$ $i_0 \in I$, y otra, $g_{j_0}(\vec{X})$ $j_0 \in J$

En efecto: la independencia entre los subvectores $\vec{X}_I$ y $\vec{X}_J$ implica que existen dos subvectores $\vec{X}_L$, $\vec{X}_{N-L}$, independientes tales que $\vec{X}_I \subset \vec{X}_L$ $\vec{X}_J \subset \vec{X}_{N-L}$. La densidad $f(\vec{x})$ se puede expresar:

$$f(\vec{x}) = h(\vec{X}_L)\, h(\vec{X}_{N-L})$$

y al calcular las generadoras no es posible encontrar una variable x_l común a los subsistemas correspondientes a X_I y X_J.

Si no existe una variable x_l $l \in N$ que figure simultáneamente en las generadoras $g_i(\vec{x})$ $i \in I$ y $g_j(\vec{x})$ $j \in J$, se construye una partición de K, llamando:

- $K_1 \subset K$ al conjunto de subíndices correspondientes a las variables de $\vec{X}_K$ que intervienen en $g_i(\vec{x})$ $i \in I$. $L = I + K_1$
- $K_2 \subset K$ al conjunto de subíndices correspondientes a las variables de $\vec{X}_K$ que intervienen en $g_j(\vec{x})$ $j \in J$. $M = J + K_2$
- $K_3 = K - K_1 - K_2$

Con la notación anterior, sea:

$$dU(\vec{x}) = \sum_{s \in N} g_s(\vec{x})\, dx_s$$

$$dU(\vec{x}_L) = \sum_{l \in L} g_l(\vec{x})\, dx_l$$

$$dU(\vec{x}_M) = \sum_{m \in M} g_m(\vec{x})\, dx_m$$

$$dU(\vec{x}_{K_3}) = \sum_{p \in K_3} g_p(\vec{x})\, dx_p$$

Teniendo en cuenta que los subconjuntos L, M y K_3 constituyen una partición de N, se puede escribir:

$$dU(\vec{x}) = dU(\vec{X}_L) + dU(\vec{X}_M) + dU(\vec{X}_{K_3})$$

y por tanto

$$f(\vec{x}) = K\, e^{U(\vec{X}_L)}\, e^{U(\vec{X}_M)}\, e^{U(\vec{X}_{K_3})}$$

como consecuencia, la densidad conjunta es igual al producto de tres funciones positivas:

$f(\vec{x}) = K\, h(\vec{X}_L)\; h(\vec{X}_M)\; h(\vec{X}_P)$

de donde se deduce que los vectores $\vec{X}_L$, $\vec{X}_M$, $\vec{X}_P$ son independientes dos a dos.

En particular, y puesto que $\vec{X}_I \subset \vec{X}_L$ y $\vec{X}_J \subset \vec{X}_M$, los subvectores aleatorios $\vec{X}_I$, $\vec{X}_J$ son independientes.

COROLARIO 3.1

Una condición necesaria y suficiente para que los subvectores $\vec{X}_I$, $\vec{X}_J$ y $\vec{X}_K$ con $I \cap J = \emptyset$ $K = N-I-J$ sean mutuamente independientes es que las funciones que constituyen el subsistema generador de la variable $\vec{x}_I$ sólo dependan de elementos de $\vec{x}_I$, y las funciones que constituyen el subsistema generador de la variable $\vec{x}_J$ sólo dependan de elementos de $\vec{x}_J$

Puesto que se verifica la igualdad de Schwartz, resulta evidente que las funciones que constituyen el subsistema generador de la variable $\vec{x}_K$ sólo dependen de elementos de $\vec{x}_K$, la demostración resulta evidente a partir del teorema anterior.

Este último corolario 3.1 se puede hacer extensivo al caso de una partición del vector $\vec{X}$ en subvectores $\vec{X}_{I_1}, \vec{X}_{I_2}, \ldots, \vec{X}_{I_s}$. Así los subvectores serán mutuamente independientes si y sólo si las funciones generadoras de cada subvector sólo dependen de las variables de dicho subvector.

Como caso particular de partición, las variables del vector aleatorio $\vec{X}$ son independientes en su totalidad si y sólo si cada función generadora depende sólo de la variable a la que corresponde, es decir:

$g_i(\vec{x}) = h_i(x_i) \quad \forall\, i \in N$

3.3.4 Independencia condicionada entre las variables de $\vec{X}_I$ y de $\vec{X}_J$

A.P. Dawid (1979), [10], establece el concepto de independencia condicionada:

Las variables $\vec{X}_I$ y $\vec{X}_J$ son independientes en el seno de la distribución condicionada al complemento, cuando la función de densidad de la distribución condicionada se puede factorizar de la forma:

$$f(\vec{x}_I,\vec{x}_J/\vec{x}_K) = h(\vec{x}_I,\vec{x}_K)\, h(\vec{x}_J,\vec{x}_K)$$

El teorema siguiente, 3.3, establece las condiciones necesarias y suficientes para este tipo de independencia, a través de los sistemas de funciones generadoras.

TEOREMA 3.3

Una condición necesaria y suficiente para que los vectores $\vec{X}_I$ y $\vec{X}_J$, con $I, J \in \wp(N)$, $I \cap J = \emptyset$, sean independientes en el seno de la distribución condicionada, $f(\vec{x}_I,\vec{x}_J/\vec{x}_{N-I-J})$, es que en el sistema de generadoras:

$$\{ g_i(\vec{x})_{i\in I} \,,\; g_j(\vec{x})_{j\in J} \,,\; g_k(\vec{x})_{k\in K} \}$$

las variables comunes en los subsistemas $g_i(\vec{x})$ $i\in I$ y $g_j(\vec{x})$ $j\in J$ pertenezcan al subvector $\vec{X}_K$

En efecto, si $\vec{X}_I$ $\vec{X}_J$ son independientes en el seno de la distribución condicionada, $f(\vec{x}_I,\vec{x}_J/\vec{x}_K) = h(\vec{x}_I,\vec{x}_K)h(\vec{x}_J,\vec{x}_K)$, y al calcular las generadoras

$$g_i(\vec{x}_{I+J}/\vec{x}_K) = g_i(\vec{x}) = \frac{\partial}{\partial x_i} \ln f(\vec{x}_I,\vec{x}_J/\vec{x}_K) \qquad \forall\, i\in I$$

no depende de las variables de $\vec{X}_J$.

$$g_j(\vec{x}_{I+J}/\vec{x}_K) = g_j(\vec{x}) = \frac{\partial}{\partial x_j} \ln f(\vec{x}_I,\vec{x}_J/\vec{x}_K) \qquad \forall\, j\in J$$

no depende de las variables de $\vec{X}_I$. Por tanto se verifica la condición.

Recíprocamente, si las únicas variables comunes en los subsistemas $g_i(\vec{x})$ $i \in I$ y $g_j(\vec{x})$ $j \in J$ pertenecen al subvector $\vec{X}_K$, de igual forma que en el teorema 3.2, se construye una partición de K.

Por otra parte, puesto que la densidad de la distribución condicionada $(\vec{X}_I, \vec{X}_J / \vec{X}_K)$ está generada a partir del sistema $\{ g_i(\vec{x})_{i \in I}\ ,\ g_j(\vec{x})_{j \in J} \}$, existen dos funciones $U(\vec{X}_L)$ y $U(\vec{X}_M)$ tales que $dU(\vec{x}_L) = \sum_{i \in I} g_i(\vec{x})\, dx_i$ $\quad dU(\vec{x}_M) = \sum_{j \in J} g_j(\vec{x})\, dx_j$,

lo que nos permite expresar $f(\vec{x}_I, \vec{X}_J / \vec{X}_K) = K\, e^{U(\vec{X}_L)}\, e^{U(\vec{X}_M)}$

Teniendo en cuenta que en general $U(\vec{X}_L)$ es una función del subvector $(\vec{x}_I, \vec{x}_K)$ y que $U(\vec{X}_M)$ lo es de $(\vec{x}_J, \vec{x}_K)$, se concluye que los subvectores $\vec{X}_I$ y $\vec{X}_J$ son independientes en el seno de la distribución condicionada, $f(\vec{x}_I, \vec{x}_J / \vec{x}_{N-I-J})$.

Siempre que consideremos la independencia entre los subvectores $\vec{X}_I$ y $\vec{X}_J$ como un caso particular de la independencia condicionada, se puede enunciar el siguiente corolario:

COROLARIO 3.2

Una condición necesaria y suficiente para que los subvectores $\vec{X}_I$ y $\vec{X}_J$ con $I \cap J = \emptyset$ sean independientes en el seno de la distribución condicionada $f(\vec{x}_I, \vec{x}_J / \vec{x}_{N-I-J})$ es que las funciones que constituyen el subsistema generador de la variable $\vec{x}_I$ no dependan de las variables de $\vec{x}_J$

La demostración resulta evidente. Obsérvese que en este caso las funciones $g_j(\vec{x})$ $\forall\, j \in J$ no dependen de las variables del subvector $\vec{X}_I$.

3.3.5 Ejemplos sobre independencia estocástica.

En los ejemplos que siguen se pone de manifiesto que la estructura de independencia se puede observar directamente en el sistema de generadoras.

1. La distribución de probabilidad de (X, Y, Z), dada por sus funciones generadoras, definidas en el abierto cero, uno:

$$g_x(x,y,z) = \frac{1}{x} + \frac{1}{x+z}$$

$$g_y(x,y,z) = \frac{1}{y}$$

$$g_z(x,y,z) = \frac{1}{x+z}$$

se observa cómo las variables X e Y son independientes, pues no existe una variable de la cual dependan analíticamente sus respectivas generadoras.

Lo mismo ocurre con las variables Y y Z.

Sin embargo entre X y Z no existe independencia, y tampoco independencia condicionada puesto que las generadoras comparten las variables x y z.

Este resultado se puede comprobar obteniendo la densidad conjunta

$$f(x,y,z) \propto xy(x+z) \quad 0 < x,\ y,\ z < 1$$

2. La distribución de probabilidad de (X_1, X_2, X_3, X_4), dada por sus funciones generadoras, definidas en el abierto cero, uno:

$$g_1(\vec{x}) = \frac{1}{x_1 + x_2} + \frac{1}{x_1 + x_3 + x_4} \tag{3.26}$$

$$g_2(\vec{x}) = \frac{1}{x_1 + x_2} \tag{3.27}$$

$$g_3(\vec{x}) = \frac{1}{x_3 + x_4} + \frac{1}{x_1 + x_3 + x_4} \tag{3.28}$$

$$g_4(\vec{x}) = \frac{1}{x_3 + x_4} + \frac{1}{x_1 + x_3 + x_4} \tag{3.29}$$

se observa que (3.27) sólo depende de x_1 y x_2 lo que nos podría hacer pensar que la variable aleatoria X_2 sólo depende de X_1, pero esto no es cierto, puesto que x_1 también está en (3.28) y (3.29), lo que establece una dependencia entre X_2 y (X_3, X_4).

Sin embargo sí existe independencia entre X_2 y (X_3, X_4) en $(X_2, X_3, X_4/X_1)$, puesto que la única variable común en (3.27) y {(3.28), (3.29)} es x_1.

Este resultado se puede comprobar en general, pues si una densidad conjunta que posee como únicos factores:

$$f(x_1,x_2,x_3,x_4) = K\, h(x_1,x_2)\, h(x_3,x_4)\, h(x_1,x_3,x_4)$$

la distribución de X_2 condicionada a las demás es

$$f(x_2/x_1,x_3,x_4) = \frac{h(x_1,x_2)}{\int_{-\infty}^{+\infty} h(x_1,x_2)\, dx_2}$$

y no resulta difícil probar que $f(x_2/x_1,x_3,x_4) = f(x_2/x_1)$, lo que nos indica que la variable aleatoria X_2 es estocásticamente independiente del vector (X_3,X_4).

Pero esta independencia se establece en relación a la distribución condicionada $(X_2,X_3,X_4/X_1)$, pues

$$f(x_2,x_3,x_4/x_1) = \frac{h(x_1,x_2)}{\int_{-\infty}^{+\infty} h(x_1,x_2)\, dx_2} \; \frac{h(x_3,x_4)\, h(x_1,x_3,x_4)}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x_3,x_4)\, h(x_1,x_3,x_4)\, dx_3 dx_4}$$

y sin embargo no es posible la factorización de $f(x_2,x_3,x_4)$ como producto de los factores $K\, h(x_2)\, h(x_3,x_4)$

El sistema de generadoras de esta distribución

$$g_1(\vec{x}) = \frac{\partial}{\partial x_1} \ln h(x_1,x_2) + \frac{\partial}{\partial x_1} \ln h(x_1,x_3,x_4)$$

$$g_2(\vec{x}) = \frac{\partial}{\partial x_2} \ln h(x_1,x_2)$$

$$g_3(\vec{x}) = \frac{\partial}{\partial x_3} \ln h(x_3,x_4) + \frac{\partial}{\partial x_3} \ln h(x_1,x_3,x_4)$$

$$g_4(\vec{x}) = \frac{\partial}{\partial x_4} \ln h(x_3,x_4) + \frac{\partial}{\partial x_4} \ln h(x_1,x_3,x_4)$$

muestra la independencia estocástica entre variables.

3.4 ESTUDIO DE LA INDEPENDENCIA EN UNA DISTRIBUCIÓN DISCRETA MULTIVARIANTE.

3.4.1 Introducción.

El tratamiento de la independencia en el caso discreto resulta totalmente análogo al caso continuo si para ello se utiliza la información de la dependencia analítica en el sistema de funciones generadoras, $L_i(\vec{r})$.

La estructura de independencia se puede observar directamente en el sistema de generadoras de la densidad conjunta.

Se supone que el dominio de definición de las variables aleatorias sea tal que la función de cuantía conjunta $p_{\vec{r}}$ sea estrictamente positiva.

3.4.2 Independencia entre las variables de $\vec{X}_I$ y de $\vec{X}_{N-I}$

TEOREMA 3.4

Una condición necesaria y suficiente para que los subvectores $\vec{X}_I$ y $\vec{X}_{N-I}$ sean independientes es que las funciones que constituyen el subsistema generador de la variable $\vec{x}_I$ no dependen de las variables de $\vec{X}_{N-I}$ y que las funciones que constituyen el subsistema generador de la variable $\vec{x}_{N-I}$ no dependen de las variables de $\vec{X}_I$, es decir

$$L_i(\vec{r}) = \frac{p_{\vec{r}}}{p_{\vec{r}-\vec{i}}} \quad \forall\, r_i>0 \;\; \forall\, i\in I \text{ **no depende de** } r_j \quad \forall\, j\in N-I$$

$$L_j(\vec{r}) = \frac{p_{\vec{r}}}{p_{\vec{r}-\vec{j}}} \quad \forall\, r_j>0 \;\; \forall\, j\in N-I \text{ **no depende de** } r_i \quad \forall\, i\in I$$

Es condición necesaria pues si los subvectores $\vec{X}_I$, $\vec{X}_{N-I}$ son independientes entonces cualquiera de las variables del primero es independiente de cualquier variable del segundo. En particular se verifica que

$$p(r_j/\vec{r}_j) = p(r_j/\vec{r}_j-\vec{i}) \qquad \text{con } i\in I,\ j\in N-I,$$

luego, $\dfrac{p_{\vec{r}}}{p_{\vec{r}_j}} = \dfrac{p_{\vec{r}-\vec{i}}}{p_{\vec{r}_j-\vec{i}}}$

y por tanto $L_i(\vec{r}) = \dfrac{p_{\vec{r}}}{p_{\vec{r}-\vec{i}}} = \dfrac{p_{\vec{r}_j}}{p_{\vec{r}_j-\vec{i}}}$ no depende de r_j.

De igual forma se comprueba que $L_j(\vec{r}) = \dfrac{p_{\vec{r}}}{p_{\vec{r}-\vec{j}}}$

no depende de r_i

Es condición suficiente; si las funciones que componen el subsistema generador de los elementos de $\vec{X}_I$ no dependen de las variables del subvector $\vec{X}_{N-I}$, y las funciones del subsistema generador de $\vec{X}_{N-I}$ tampoco dependen de los elementos de $\vec{X}_I$, es decir:

$$L_i(\vec{r}) = g_{1i}(\vec{r}_I) \qquad \forall\, r_i>0 \quad \forall\, i\in I$$

$$L_j(\vec{r}) = g_{2j}(\vec{r}_J) \qquad \forall\, r_j>0 \quad \forall\, j\in N-I$$

Si tomamos como convenio que $\prod_{t=1}^{r_1} L_j(\vec{g}_1(t))$ sea igual a uno cuando r_j valga cero, a partir de (3.23) se puede escribir que:

$$p_{\vec{r}} = \prod_{l\in N} \left\{ \prod_{t=1}^{r_l} L_l(\vec{g}_l(t)) \right\} p_{\vec{0}}$$

y puesto que las funciones generadoras tienen las variables separadas, el producto se puede descomponer en dos factores:

$$p_{\vec{r}} = \prod_{i\in I} \left\{ \prod_{t=1}^{r_i} L_i(\vec{g}_i(t)) \right\} \prod_{j\in N-I} \left\{ \prod_{t=1}^{r_j} L_j(\vec{g}_j(t)) \right\} p_{\vec{0}}$$

actuando $p_{\vec{0}}$ como constante de normalización.

De la factorización se deduce que los dos subvectores son estocásticamente independientes.

En particular si $I = \{i\}$, la variable aleatoria X_i es independiente de cualquier otra del vector $\vec{X}$, si $L_i(\vec{r})$ es una función que sólo depende de r_i y por tanto ninguna de las restantes variables del sistema de generadoras depende de r_i.

3.4.3 Independencia entre las variables $\vec{X}_I$ y $\vec{X}_J$

TEOREMA 3.5

Dada $p_{\vec{r}} = p(\vec{x}_I, \vec{x}_J, \vec{x}_K)$, **los subvectores** $\vec{X}_I$ $\vec{X}_J$ **son independientes si y sólo si en el sistema de generadoras:**

$$\{ L_i(\vec{r})_{i\in I},\ L_j(\vec{r})_{j\in J},\ L_k(\vec{r})_{k\in K} \}$$

se verifica que no existe una variable r_l $l\in N$ **que figure simultáneamente en una** $L_{i_0}(\vec{r})$ $i_0\in I$ **y otra** $L_{j_0}(\vec{r})$ $j_0\in J$ **generadora.**

En efecto, la independencia entre los subvectores $\vec{X}_I$ y $\vec{X}_J$ implica que existen dos subvectores $\vec{X}_L$, $\vec{X}_{N-L}$, independientes tales que $\vec{X}_I \subset \vec{X}_L$ $\vec{X}_J \subset \vec{X}_{N-L}$.

De este modo $p_{\vec{r}}$ se puede expresar como producto de marginales en las que no existen factores que dependan simultáneamente de variables de $\vec{r}_I$ y de $\vec{r}_J$, es decir

$$p_{\vec{r}} = p(\vec{x}_L +)\ p(\vec{x}_{N-L} +)$$

y al calcular las generadoras a través de los respectivos cocientes, se cancelan los factores invariantes y por tanto no es posible encontrar una variable r_l común a los subsistemas correspondientes a X_I y X_J.

Si no existe una variable r_l $l\in N$ que figure simultáneamente en las generadoras $L_i(\vec{r})$ $i\in I$ y $L_j(\vec{r})$ $j\in J$, entonces se construye una partición de K, llamando:

$K_1 \subset K$ al conjunto de subíndices correspondientes a las variables de $\vec{X}_K$ que intervienen en $L_i(\vec{r})$ $i\in I$

$K_2 \subset K$ al conjunto de subíndices correspondientes a las variables de $\vec{X}_K$ que intervienen en $L_j(\vec{r})$ $j \in J$

$K_3 = K - K_1 - K_2$

Se puede escribir así que:

$$p_{\vec{r}} = \prod_{l \in I + K_1} \left\{ \prod_{t=1}^{r_l} L_l(\vec{g}_l(t)) \right\} \prod_{m \in J + K_2} \left\{ \prod_{t=1}^{r_m} L_m(\vec{g}_m(t)) \right\} \prod_{p \in K_3} \left\{ \prod_{t=1}^{r_p} L_p(\vec{g}_p(t)) \right\} p_{\vec{0}}$$

(Se sigue utilizando $\prod_{t=1}^{r_l} L_j(\vec{g}_l(t) = 1$ cuando $r_l = 0$)

Los vectores $\vec{X}_L$, $\vec{X}_M$, $\vec{X}_P$ son independientes dos a dos, y puesto que $\vec{X}_I \subset \vec{X}_L$ y $\vec{X}_J \subset \vec{X}_M$, los subvectores aleatorios $\vec{X}_I$, $\vec{X}_J$ son independientes.

Dada una partición del vector $\vec{X}$ en subvectores $\vec{X}_{I_1}, \vec{X}_{I_2}, \ldots, \vec{X}_{I_s}$, los subvectores serán mutuamente independientes si y sólo si las funciones generadoras de cada subvector sólo dependen de las variables de dicho subvector.

Como caso particular de partición, las variables del vector aleatorio $\vec{X}$ son independientes en su totalidad si y sólo si cada función generadora depende sólo de la variable a la que corresponde, es decir: $L_i(\vec{r}) = h_i(r_i)$ $\forall\, i \in N$

3.4.4 Independencia condicionada entre las variables $\vec{X}_I$ y $\vec{X}_J$

TEOREMA 3.6

Una condición necesaria y suficiente para que los vectores $\vec{X}_I$ y $\vec{X}_J$, con $I, J \in \wp(N)$, $I \cap J = \emptyset$, sean independientes en el seno de la distribución condicionada, $p(\vec{r}_I, \vec{r}_J / \vec{r}_{N-I-J})$, es que en el sistema de generadoras:

$$\{ L_i(\vec{r})_{i \in I} \,,\; L_j(\vec{r})_{j \in J} \,,\; L_k(\vec{r})_{k \in K} \}$$

las variables comunes en los subsistemas $L_i(\vec{r})$ $i \in I$ **y** $L_j(\vec{r})$ $j \in J$ **pertenezcan al subvector** $\vec{X}_K$

En efecto, si $\vec{X}_I$ $\vec{X}_J$ son independientes en el seno de la distribución condicionada, $p(\vec{r}_I,\vec{r}_J/\vec{r}_K) = h(\vec{r}_I,\vec{r}_K)\,h(\vec{r}_J,\vec{r}_K)$, y al calcular las generadoras, $L_i(\vec{r}_{I+J}/\vec{r}_K) = L_i(\vec{r}) = \dfrac{p_{\vec{r}}}{p_{\vec{r}-\vec{i}}}$ no depende de las variables de $\vec{X}_J$ para todo i perteneciente a I, ya que se cancelan en los cocientes.
$L_j(\vec{r}_{I+J}/\vec{r}_K) = L_j(\vec{r}) = \dfrac{p_{\vec{r}}}{p_{\vec{r}-\vec{j}}}$, por la misma razón, no depende de las variables de $\vec{X}_I$ para todo j perteneciente a J.

Recíprocamente, si las únicas variables comunes en los subsistemas $L_i(\vec{r})$ $i \in I$ y $L_j(\vec{r})$ $j \in J$ pertenecen al subvector $\vec{X}_K$, entonces la densidad de la distribución condicionada $(\vec{X}_I,\vec{X}_J/\vec{X}_K)$, se puede expresar de la forma:

$$p(\vec{r}_I,\vec{r}_J/\vec{r}_K) = \prod_{l \in N-J} \left\{ \prod_{t=1}^{r_l} L_l(\vec{g}_l(t)) \right\} \prod_{m \in N-I} \left\{ \prod_{t=1}^{r_m} L_m(\vec{g}_m(t)) \right\} p_{\vec{0}}$$

es decir es el producto de las funciones $h(\vec{x}_I,\vec{x}_K)$ y $h(\vec{x}_J,\vec{x}_K)$, por tanto entre los subvectores $\vec{X}_I$ y $\vec{X}_J$ existe independencia condicionada.

De igual manera que para variables continuas, considerando la independencia como un caso particular de la independencia condicionada, para variables aleatorias discretas sigue teniendo validez el corolario 3.2.

3.5 TRATAMIENTO DE LOS PARÁMETROS EN LA GENERACIÓN DE LAS DISTRIBUCIONES.

3.5.1 Introducción.

Hasta ahora, en la expresión de la función de densidad no se han destacado los posibles parámetros que pudiera contener. Sin embargo, al hablar de una familia de densidades, todas ellas con la misma expresión funcional, pero con diferentes valores en sus parámetros, se hace necesario destacarlos en la notación.

Supongamos que la distribución del vector aleatorio $\vec{X} = (X_1, \ldots, X_n)$ contiene el vector $\vec{\theta} = (\theta_1, \ldots, \theta_k)$, perteneciente a un espacio paramétrico de dimensión k, entonces cualquier miembro de la familia se notará

$$f(x_1, \ldots x_n; \theta_1, \theta_2, \ldots, \theta_k) = f(\vec{x}; \vec{\theta})$$

Las funciones generadoras

$$g_i(\vec{x}; \vec{\theta}) = \frac{\partial}{\partial x_i} \ln f(x_i; \vec{\theta} / \vec{x}_i)$$

en función de la densidad conjunta, tienen la expresión:

$$g_i(\vec{x}; \vec{\theta}) = \frac{\partial}{\partial x_i} \ln f(\vec{x}; \vec{\theta})$$

La generación de la distribución se realiza de forma análoga al caso multivariante, teniendo en cuenta que si llamamos

$$dU(\vec{x}, \vec{\theta}) = \sum_{i=1}^{n} g_i(\vec{x}; \vec{\theta})\, dx_i$$

entonces

$$f(\vec{x}; \vec{\theta}) = K(\vec{\theta})\, e^{U(\vec{x}, \vec{\theta})}$$

siendo

$$K(\vec{\theta}) = \left[\int_{X_N} e^{U(\vec{x}, \vec{\theta})}\, d\vec{x} \right]^{-1}$$

o bien haciendo $h(\vec{x};\vec{\theta}) = e^{U(\vec{x},\vec{\theta})}$, queda

$$f(\vec{x};\vec{\theta}) = K(\vec{\theta})\ h(\vec{x},\vec{\theta})$$

es decir, la densidad conjunta se descompone en producto de dos funciones, en una de ellas se han separado todos los factores que dependen sólo de los parámetros, y la otra es una función de las variables y de los parámetros.

Sustituyendo $k(\vec{\theta})$ por su valor, se puede escribir:

$$f(\vec{x};\vec{\theta}) = \frac{h(\vec{x},\vec{\theta})}{\int_X h(\vec{x},\vec{\theta})\, d\vec{x}} \qquad (3.30)$$

3.5.2 Distribución n+k variante.

Siempre que la integral $\int_X\int_\Theta h(\vec{x},\vec{\theta})\, d\vec{x}\, d\vec{\theta}$ sea finita, definimos:

$$f_1(\vec{x},\vec{\theta}) = \left[\int_X\int_\Theta h(\vec{x},\vec{\theta})\, d\vec{x}\, d\vec{\theta} \right]^{-1} h(\vec{x},\vec{\theta})$$ (1)

que resulta ser la función de densidad de una variable aleatoria n+k variante, esto es $f_1(\vec{x},\vec{\theta}) = K\, h(\vec{x},\vec{\theta})$, siendo

$$K = \left[\int_X\int_\Theta h(\vec{x},\vec{\theta})\, d\vec{x}\, d\vec{\theta} \right]^{-1}$$

y calculando las distribuciones condicionadas, resulta

$$f_1(\vec{x}/\vec{\theta}) = \frac{h(\vec{x},\vec{\theta})}{\int_X h(\vec{x},\vec{\theta})\, d\vec{x}} \qquad (3.31)$$

$$f_1(\vec{\theta}/\vec{x}) = \frac{h(\vec{x},\vec{\theta})}{\int_\Theta h(\vec{x},\vec{\theta})\, d\vec{\theta}} \qquad (3.32)$$

---

(1) Queremos hacer notar que las funciones $f(\vec{x};\vec{\theta})$ y $f_1(\vec{x},\vec{\theta})$, evidentemente son distintas, pues mientras la primera está definida de $\mathbb{R}^n$ en $\mathbb{R}$, la segunda lo hace de $\mathbb{R}^{n+k}$ en $\mathbb{R}$.

Obsérvese que las expresiones, (3.30) y (3.31), de las funciones $f(\vec{x};\vec{\theta})$ y $f_1(\vec{x}/\vec{\theta})$ correspondientes a la densidad de la variable aleatoria $\vec{X}$ con parámetros $\vec{\theta}$, y a la densidad de la variable aleatoria $\vec{X}$ condicionada por la variable aleatoria $\vec{\theta}$, respectivamente, son formalmente idénticas.

Se puede enunciar que formalmente no existe diferencia entre las funciones $f(\vec{x};\vec{\theta})$ y $f(\vec{x}/\vec{\theta})$, hecho que nos permite considerar a la familia de densidades conjuntas de n variables aleatorias, para cada juego de valores de los parámetros, como una distribución de las variables condicionada por los parámetros, considerados como variables.

3.5.3 Funciones generadoras de una distribución n-variante con parámetros. Generación de la distribución $f_1(\vec{x}/\vec{\theta})$

Los parámetros, $\vec{\theta}$, pueden considerarse como valores fijos, (arbitrarios), de las variables aleatorias, y por tanto la notación $\{g_i(\vec{x};\vec{\theta})\}_{i=1,\dots,n}$ es interpretable como funciones generadoras condicionadas que, siguiendo la definición, vienen dadas por las expresiones:

$$g_i(\vec{x};\vec{\theta}) = \frac{\partial}{\partial x_i} \ln h(\vec{x},\vec{\theta}) + \frac{\partial}{\partial x_i} \ln l_1(x_i,\vec{\theta})$$

que con la notación habitual:

$$m_i(\vec{x},\vec{\theta}) = \frac{\partial}{\partial x_i} \ln h(\vec{x},\vec{\theta})$$

$$n_i(x_i,\vec{\theta}) = \frac{\partial}{\partial x_i} \ln l_1(x_i,\vec{\theta})$$

es

$$g_i(\vec{x};\vec{\theta}) = m_i(\vec{x},\vec{\theta}) + n_i(x_i,\vec{\theta}) \tag{3.33}$$

Si se cumple la integrabilidad de las correspondientes funciones y la condición de Schwartz,

$$\frac{\partial}{\partial x_i} m_j(\vec{x},\vec{\theta}) = \frac{\partial}{\partial x_j} m_i(\vec{x},\vec{\theta}) \qquad \forall\, i \neq j$$

se puede generar la distribución condicionada $f_1(\vec{x}/\vec{\theta})$, pues

$$f(\vec{x};\vec{\theta}) = K(\vec{\theta})\; e^{\int m_j(\vec{x};\vec{\theta})\,dx_j + \sum_{i=1}^{k} \int n_i(x_i;\vec{\theta})\,dx_i}$$

para un j dado; la constante $K(\vec{\theta})$ se calculará mediante la condición de normalización.

Es claro que para construir la función $f(\vec{x},\vec{\theta})$, como distribución n+k variante, al considerar $\vec{\theta}$ como variables, es necesario conocer el sistema generador completo:

$$\{ g_i(\vec{x},\vec{\theta}) \}_{i=1,\ldots,n}$$

$$\{ g_j(\vec{x},\vec{\theta}) \}_{j=n+1,\ldots,n+k}$$

para el que, evidentemente se ha de seguir verificando la igualdad de Schwartz entre las derivadas segundas.

Obsérvese que para la generación de $f(\vec{x}_I/\vec{x}_{N-I})$, formalmente es equivalente interpretar $\vec{x}_{N-I}$ como valores fijos de las variables que establecen la distribución condicionada o como parámetros de la función.

Una diferencia entre los enfoques clásico y bayesiano de la estadística radica, posiblemente, en aceptar o no como plausible la hipótesis de que se puede obtener una muestra para la variable $\vec{x}_I$, manteniendo constantes los valores de $\vec{x}_{N-I}$ y siempre que estas pueden considerarse aproximadamente constantes durante el tiempo en que se utilicen las conclusiones obtenidas de la muestra.

3.5.4 Ejemplos de la generación de una distribución bidimensional con parámetros.

La normal bivariante.

Se trata de obtener la densidad conjunta de una variable aleatoria bidimensional en la que intervienen parámetros.

Sean las funciones:

$$m_1(x,y) = \frac{\partial}{\partial x} U(x,y) = \frac{\rho}{(1 - \rho^2)} \frac{y - \mu_y}{\sigma_x \sigma_y}$$

$$m_2(x,y) = \frac{\partial}{\partial y} U(x,y) = \frac{\rho}{(1 - \rho^2)} \frac{x - \mu_x}{\sigma_x \sigma_y}$$

$$n_1(x) = \frac{-1}{1 - \rho^2} \left(\frac{x - \mu_x}{\sigma_x^2}\right)$$

$$n_2(y) = \frac{-1}{1 - \rho^2} \left(\frac{y - \mu_y}{\sigma_y^2}\right)$$

Se puede escribir:

$$U(x,y) = \frac{\rho}{(1 - \rho^2)\, \sigma_x \sigma_y} u(x,y)$$

en cuyo caso ha de ser:

$$\frac{\partial}{\partial x} u(x,y) = y - \mu_y \qquad (3.34)$$

$$\frac{\partial}{\partial y} u(x,y) = x - \mu_x \qquad (3.35)$$

integrando (3.34) con respecto a x,

$$u(x,y) = \int (y - \mu_y)\, dx + C(y) = (y - \mu_y)\, x + C(y) \qquad (3.36)$$

y realizando la derivada parcial de u(x,y), según (3.36), con respecto a y, ha de ser:

$$\frac{\partial}{\partial y} u(x,y) = x + \frac{d\, C(y)}{dy} = x - \mu_x$$

por tanto:

$$\frac{d\,C(y)}{dy} = -\mu_x$$

e integrando ahora en y:

$$C(y) = -y\mu_x + K$$

y como consecuencia:

$$u(x,y) = (y-\mu_y)x - y\mu_x + K = xy - x\mu_y - y\mu_x + k \qquad (3.37)$$

También puede obtenerse integrando (3.35) con respecto a y, siguiendo un desarrollo paralelo, se llega lógicamente al mismo valor

$$u(x,y) = (x-\mu_x)y - x\mu_y + K = xy - x\mu_y - y\mu_x + k \qquad (3.38)$$

De una forma u otra

$$\ln h(x,y) = U(x,y) = \frac{\rho}{(1-\rho^2)\,\sigma_x\,\sigma_y}\left[(y-\mu_y)x - y\mu_x\right]$$

Se puede calcular también:

$$\ln l_1(x) = \int \frac{-1}{1-\rho^2}\left(\frac{x-\mu_x}{\sigma_x^2}\right)dx = \frac{-1}{2(1-\rho^2)}\left(\frac{x-\mu_x}{\sigma_x}\right)^2$$

análogamente para $l_2(y)$:

$$\ln l_2(y) = \frac{-1}{2(1-\rho^2)}\left(\frac{y-\mu_y}{\sigma_y}\right)^2$$

así pues:

$$f(x,y) \propto e^{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{(y-\mu_y)x - y\mu_x}{\sigma_x\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]}$$

Comparando esta solución con la usual, se observa la ausencia del término $\mu_x\mu_y$ en la función $u(x,y)$. Es lógico que así ocurra, pues el factor que corresponde al mencionado término, $e^{\frac{\mu_x\mu_y}{(1-\rho^2)\sigma_x\sigma_y}}$ forma parte de la constante de normalización, $K(\mu_x, \mu_y, \rho, \sigma_x, \sigma_y)$, y no de las funciones h, l_1, l_2, ya que no depende de las variables x e y.

Obsérvese que también es solución: $u(x,y) = (y-\mu_y)(x-\mu_x) + K'$ sin mas que hacer: $K = \mu_x\mu_y + K'$

La distribución beta bivariante.

La densidad de la distribución beta bidimensional de parámetros $p_1 + 1$, $p_2 + 1$, $p_3 + 1$, se obtiene a partir de las funciones:

$$m_1(x,y) = \frac{\partial}{\partial x} U(x,y) = -\frac{p_3}{1-x-y} \qquad (3.39)$$

$$m_2(x,y) = \frac{\partial}{\partial y} U(x,y) = -\frac{p_3}{1-x-y} \qquad (3.40)$$

$$n_1(x) = \frac{p_1}{x} \qquad\qquad n_2(y) = \frac{p_2}{y}$$

siendo $0 < x < 1$; $0 < y < 1$; $x + y < 1$.

Integrando (3.39) con respecto a x,

$$U(x,y) = p_3 \ln(1-x-y) + C(y) \qquad (3.41)$$

derivando (3.41) con respecto a y,

$$\frac{\partial}{\partial y} U(x,y) = -\frac{p_3}{1-x-y} + \frac{d\,c(y)}{dy}$$

y teniendo en cuenta (3.40) se deduce que $\frac{d\,c(y)}{dy} = 0$, luego

$C(y) = k$ y por tanto

$$\ln h(x,y) = U(x,y) = \ln(1-x-y)^{p_3} + k$$

También puede obtenerse U(x,y) integrando (3.40) con respecto a y, siguiendo un desarrollo paralelo, se llega lógicamente al mismo valor

Por otra parte:

$$\ln l_1(x) = \int \frac{p_1}{x}\,dx = \ln x^{p_1} + k'$$

$$\ln l_2(y) = \int \frac{p_2}{y}\,dy = \ln y^{p_2} + k''$$

de esta forma resulta:

$$f(x,y) \propto x^{p_1}\, y^{p_2}\, (1-x-y)^{p_3}$$

CAPÍTULO 4

APLICACIONES Y PROBLEMAS ASOCIADOS

En los capítulos anteriores se han definido los sistemas de funciones generadoras de distribuciones de probabilidad multivariantes, tanto continuas como discretas y se ha realizado una aplicación concreta, el estudio de la independencia estocástica.

En este capítulo se recogen, a modo ilustrativo, otras aplicaciones, que pueden dar una idea de la riqueza de posibilidades que tiene la utilización de las funciones generadoras. Estas aplicaciones pertenecen al campo teórico de la Estadística Matemática, al campo económico y al campo financiero.

1) En la primera de las aplicaciones veremos cómo la factorización, utilizada en capítulos anteriores, surge de una forma natural desde las ecuaciones diferenciales, subyacentes en los sistemas de generadoras y se pone de manifiesto el papel primordial que dichos factores juegan a la hora de interpretar las relaciones de dependencia e independencia entre las distintas partes del vector aleatorio.

Para obtener estos resultados se define una relación de equivalencia entre densidades de probabilidad mediante una relación diferencial, en caso continuo, o en diferencias en el caso discreto, que se asemeja a la metodología empleada en la definición y estudio de propiedades de los sistemas pearsonianos. Demostramos una propiedad sobre la rectangularidad de los sistemas de funciones generadoras.

2) Se estudian las distribuciones de probabilidad de tipo mixto y se demuestra que los sistemas de funciones generadoras verifican las mismas propiedades, que los sistemas de los tipos continuo y discreto, estudiados anteriormente.

3) En la siguiente aplicación comprobamos, de una forma sencilla, si una familia de distribuciones a priori es conjugada o no para una clase de verosimilitudes y, conocidas las distribuciones a priori y las verosimilitudes, conocemos las funciones generadoras de las distribuciones a posteriori y por tanto la distribución y sus parámetros.

4) Bajo condiciones de regularidad, se establece una relación entre la matriz de información de Fisher y las varianzas de las funciones generadoras. Así mismo las funciones generadoras permiten establecer una relación entre los métodos de la máxima verosimilitud y de los momentos.

5) La aportación en el campo económico viene dado por la utilización de la idea amplia de función generadora. Se estudian las funciones reales de variable real capaces de generar curvas que puedan modelizar una curva de Lorenz.

Se estudian las condiciones necesarias y suficientes para que una función g(x), real de variable real, sea generadora de una curva de Lorenz. Estas condiciones se aplican a ejemplos concretos, llegando a establecer las condiciones sobre los coeficientes de una función que verifica la ecuación de Pearson, para que sea también generadora de una curva de Lorenz.

6) Por lo que hace referencia a la aplicación al campo financiero, se utiliza la idea amplia de función generadora que permite establecer una relación directa entre ella y la ley financiera. Esto se consigue generando la función de distribución de una variable aleatoria a partir de una función real de variable real.

4.1 RECTANGULARIDAD DE LOS SISTEMAS DE GENERADORAS.

En el primero de los apartados demostramos una propiedad sobre la rectangularidad de los sistemas de generadoras, relacionada con el principio de verosimilitud, para variables bidimensionales de tipo continuo.

A continuación se realiza el mismo estudio para sistemas de generadoras de variables bidimensionales de tipo discreto.

Idéntico estudio se ha realizado e idénticos resultados se han obtenido para los casos multivariante continuo y multivariante discreto; no se incluyen en esta memoria las demostraciones de los teoremas para no cansar al lector.

De la relación de equivalencia que se establece entre distribuciones de probabilidad mediante una relación diferencial, en caso continuo, o en diferencias en el caso discreto, destaca la importancia del elemento canónico, constituido por el factor común de todas las funciones de una misma clase, de cuya forma se deduce la estructura de dependencia entre variables aleatorias.

Cada elemento del conjunto cociente define una estructura de dependencia distinta, y en el caso particular de que el elemento canónico sea igual a uno, se trata de la clase de equivalencia a la que pertenecen todas las distribuciones cuyas variables aleatorias son independientes en su totalidad.

Se muestra cómo los sistemas de funciones generadores de las densidades de probabilidad son una herramienta eficaz para el estudio del conjunto cociente de las relaciones definidas, surgiendo de forma natural y sencilla las propiedades.

Como consecuencia de ellas se puede asegurar que sobre las propiedades analíticas, meramente estructurales, del sistema de funciones generadoras de la distribución, radica lo que se conoce como rectangularidad de los sistemas de funciones generadoras; en un caso particular de esta rectangularidad se sustenta el principio de verosimilitud.

4.1.1 Distribuciones continuas bidimensionales. Relación entre funciones de densidad conjunta.

*** Definición:**

Dos funciones de densidad $f_1(x,y)$, $f_2(x,y)$ **, definidas sobre** $D \subset \mathbb{R}^2$**, son equivalentes si y sólo si las respectivas funciones**

$$g_1(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_1(x,y)$$

$$g_2(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_2(x,y)$$

coinciden, es decir:

$$f_1(x,y) \ R \ f_2(x,y) \quad \Leftrightarrow \quad \frac{\partial^2 \ln f_1(x,y)}{\partial x\, \partial y} = \frac{\partial^2 \ln f_2(x,y)}{\partial x\, \partial y} \tag{4.1}$$

Evidentemente, la relación así definida es una relación de equivalencia.

*** Clases de equivalencia. Representante canónico. Conjunto cociente.**

Las funciones de densidad de una clase de equivalencia $\{ f_\lambda(x,y)\}_{\lambda\in\Lambda}$, tienen en común

$$g(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_\lambda(x,y) \qquad \forall\, \lambda \in \Lambda$$

lo que significa que:

$$\ln f_\lambda(x,y) = \int^x \int^y g(x,y)\, dx\, dy + V_\lambda(x) + W_\lambda(y) + C_\lambda \qquad \forall\, \lambda \in \Lambda$$

y los sistemas de funciones generadoras correspondientes

$\{ g_{\lambda,x}(x,y), g_{\lambda,y}(x,y) \}_{\lambda\in\Lambda}$ vienen dados por:

$$g_{\lambda,x}(x,y) = \frac{\partial}{\partial x} \ln f_\lambda(x,y) = \int^y g(x,y)\, dy + V'_\lambda(x) \tag{4.2}$$

$$g_{\lambda,y}(x,y) = \frac{\partial}{\partial y} \ln f_\lambda(x,y) = \int^x g(x,y)\, dx + W'_\lambda(y) \tag{4.3}$$

Esto sugiere elegir como representante canónico de la clase, $f_{\lambda_0}(x,y)$ aquel elemento para el cual

$V_{\lambda_0}(x) = W_{\lambda_0}(y) = 0 \quad \forall x, \ \forall y$ es decir:

$$f_{\lambda_0}(x,y) = K\, e^{\int^x\int^y g(x,y)\,dx\,dy} \tag{4.4}$$

siendo la constante de normalización:

$$K = \left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\int^x\int^y g(x,y)\,dx\,dy}\,dx\,dy \right]^{-1}$$

La expresión (4.4) permite interpretar $g(x,y)$ como elemento generador del representante canónico de la clase y, en cierta forma, como generador de la misma.

*** Sistemas de funciones generadoras, asociados a las funciones de densidad de una clase de equivalencia.**

PROPOSICIÓN 4.1

Dada una clase de equivalencia $C = \{ f_\lambda(x,y) \}_{\lambda\in\Lambda}$, **las funciones generadoras de la distribuciones condicionadas de la densidad canónica se obtienen al integrar**

$$g_{c,y}(x,y) = \int^x g(x,y)\,dx \tag{4.5}$$

$$g_{c,x}(x,y) = \int^y g(x,y)\,dy \tag{4.6}$$

Es consecuencia inmediata de la definición de densidad canónica y de las expresiones (4.2) y (4.3).

*** Condiciones necesarias y suficientes para la equivalencia.**

TEOREMA 4.1

Dos funciones de densidad $f_1(x,y)$, $f_2(x,y)$ **son equivalentes, si y sólo si la diferencia entre las generadoras condicionadas respectivas no depende de la variable que actúa de**

condicionante, es decir:

$$g_{2,y}(x,y) - g_{1,y}(x,y) = C_2(y) \quad \text{no depende de } x \qquad \textbf{(4.7)}$$

$$g_{2,x}(x,y) - g_{1,x}(x,y) = C_1(x) \quad \text{no depende de } y \qquad \textbf{(4.8)}$$

Demostración.

a) Es condición necesaria, pues si $f_2(x,y)\ R\ f_1(x,y)$ entonces se verifican las igualdades entre derivadas segundas cruzadas que permiten obtener las funciones generadoras de cada una de las funciones de densidad:

$$g_{1,y}(x,y) = \frac{\partial}{\partial y} \ln f_1(x,y) = \int g(x,y)\,dx = \int^x g(x,y)\,dx + C_{12}(y)$$

$$g_{2,y}(x,y) = \frac{\partial}{\partial y} \ln f_2(x,y) = \int g(x,y)\,dx = \int^x g(x,y)\,dx + C_{22}(y)$$

De igual forma se obtienen:

$$g_{1,x}(x,y) = \int^y g(x,y)\,dy + C_{11}(x)$$

$$g_{2,x}(x,y) = \int^y g(x,y)\,dy + C_{12}(x)$$

restando adecuadamente, resulta evidente que se verifican las relaciones (4.7) y (4.8).

b) Para probar la suficiencia basta derivar la expresión (4.7) con respecto a x o bien la expresión (4.8) con respecto a y, pues en ambos casos se deduce la equivalencia de las funciones de densidad f_1 y f_2

En particular si f_c es la densidad canónica de la clase de equivalencia a la que pertenece f, entre las funciones generadoras respectivas se verifica:

$$g_x(x,y) = g_{c,x}(x,y) + C_1(x)$$

$$g_y(x,y) = g_{c,y}(x,y) + C_2(y)$$

TEOREMA 4.2

Las funciones de densidad $f_1(x,y)$ y $f_2(x,y)$ son equivalentes si y sólo si existen dos funciones reales de variable real, positivas, $l_1(x)$, $l_2(y)$ tales que

$$f_2(x,y) = K.f_1(x,y).l_1(x).l_2(y) \qquad \mathbf{(4.9)}$$

donde K es un número real.

En efecto, si las funciones de densidad son equivalentes entonces

$$\frac{\partial^2 \ln f_2(x,y)}{\partial x\, \partial y} = \frac{\partial^2 \ln f_1(x,y)}{\partial x\, \partial y}$$

luego

$$\frac{\partial^2}{\partial x\, \partial y}\left(\ln \frac{f_1(x,y)}{f_2(x,y)}\right) = 0$$

y según el teorema 2.1, el cociente $\dfrac{f_1(x,y)}{f_2(x,y)}$ se puede escribir

igual a un producto de dos funciones positivas, una de ellas depende sólo de x y la otra sólo de y; queda así demostrada la necesidad.

Recíprocamente, si existen dos funciones $l_1(x)$ y $l_2(y)$, definidas positivas, tales que

$$f_2(x,y) = K.f_1(x,y).l_1(x).l_2(y)$$

entonces al tomar logaritmos y derivar, con respecto a x, y, es evidente que se obtiene la igualdad:

$$\frac{\partial^2 \ln f_2(x,y)}{\partial x\, \partial y} = \frac{\partial^2 \ln f_1(x,y)}{\partial x\, \partial y}$$

luego las funciones $f_1(x,y)$ y $f_2(x,y)$ son equivalentes.

TEOREMA 4.3

Dos funciones de densidad son equivalentes si y sólo si la forma factorizada de ambas tiene común el factor h(x,y)

En efecto, es condición necesaria, pues si $f_1(x,y)$ y

$f_2(x,y)$ son equivalentes, entonces existen $l_1(x)$ y $l_2(y)$, definidas positivas, tales que $f_2(x,y) = K\, f_1(x,y)\, l_1(x)\, l_2(y)$,

luego $\dfrac{f_2(x,y)}{f_1(x,y)} = K\, l_1(x)\, l_2(y)$, por tanto el factor $h_2(x,y)$ de $f_2(x,y)$ y el factor $h_1(x,y)$ de $f_1(x,y)$ son ambos la unidad (único valor constante posible) o son la misma función.

También es condición suficiente pues suponiendo que

$$f_1(x,y) = K_1\, h(x,y)\, l_{11}(x)\, l_{21}(y)$$

$$f_2(x,y) = K_2\, h(x,y)\, l_{12}(x)\, l_{22}(y)$$

entonces:

$$\frac{f_2(x,y)}{f_1(x,y)} = \frac{K_2\, h(x,y)\, l_{12}(x)\, l_{22}(y)}{K_1\, h(x,y)\, l_{11}(x)\, l_{21}(y)} = \frac{k_2}{K_1}\, \frac{l_{12}(x)}{l_{11}(x)}\, \frac{l_{22}(y)}{l_{21}(y)}$$

y denotando:

$$K = \frac{k_2}{K_1}, \quad l_1(x) = \frac{l_{12}(x)}{l_{11}(x)}, \quad l_2(y) = \frac{l_{22}(y)}{l_{21}(y)}$$

es claro que $f_2(x,y) = K\, f_1(x,y)\, l_1(x)\, l_2(y)$, de donde se deduce que las dos funciones de densidad están relacionadas.

Desde el punto de vista de la factorización, el elemento canónico de una clase de funciones de densidad equivalentes es la función de densidad

$$f(x,y) = K\, h(x,y)\, l_1(x)\, l_2(y)$$

tal que

$$l_1(x) = 1; \quad l_2(y) = 1,$$

es decir a aquella función de densidad que no admite factores que dependan analíticamente de una sola variable.

Las funciones generadoras canónicas tienen la forma:

$$g_{c,x}(x,y) = \frac{\partial}{\partial x} \ln h(x,y)$$

$$g_{c,y}(x,y) = \frac{\partial}{\partial y} \ln h(x,y)$$

Identificamos cada elemento del conjunto cociente con la función no factorizable h(x,y) que posee.

Como caso particular, e importante, la función $h(x,y) = 1$ se identifica con la clase de funciones de densidad con variables aleatorias independientes.

Las dos proposiciones siguientes establecen unas relaciones entre las funciones generadoras de las funciones de densidad de una clase y las funciones generadoras canónicas de dicha clase.

PROPOSICIÓN 4.2

Sea $f(x,y)$ **una densidad de la clase cuyo elemento canónico es** $f_c(x,y)$**, sea** $f(x,y) = K\, f_c(x,y)\, l_1(x)\, l_2(y)$**, entonces entre las generadoras se verifican las relaciones:**

$$g_x(x,y) = g_{c,x}(x,y) + \frac{d \ln l_1(x)}{dx}$$

$$g_y(x,y) = g_{c,y}(x,y) + \frac{d \ln l_2(y)}{dy}$$

(Expresiones que están en concordancia con (4.2) y (4.3)).

La demostración es inmediata, sin más que derivar la forma factorizada con respecto a una u otra variable.

Puesto que los elementos que definen las funciones $g_x(x,y)$ y $g_y(x,y)$ son perfectamente separables en sumandos, que dependen de las dos variables o de una sola, las generadoras del elemento canónico se leen directamente del sistema de cualquier elemento de la clase.

PROPOSICIÓN 4.3

Dada una clase definida por h(x,y). Sea $\{g(x/y),\ g(y/x)\}$ **el sistema de funciones generadoras de un elemento de dicha clase, entonces se verifica:**

$$\frac{\partial}{\partial y}\int^{x} g_x(x,y)\,dx = g_{c,y}(x,y) \qquad \textbf{(4.10)}$$

$$\frac{\partial}{\partial x}\int^{y} g_y(x,y)\,dy = g_{c,x}(x,y) \qquad \textbf{(4.11)}$$

En efecto, como ya es sabido:

$$f(x,y) = K\,h(x,y)\;l_1(x)\;l_2(y)$$

$$g_x(x,y) = \frac{\partial}{\partial x}\ln h(x,y) + \frac{d}{dx}\ln l_1(x)$$

$$g_y(x,y) = \frac{\partial}{\partial y}\ln h(x,y) + \frac{d}{dy}\ln l_2(y)$$

Haciendo la integración en x, salvo constantes,

$$\int^{x} g_x(x,y)\;dx = \ln h(x,y) + \ln l_1(x)$$

al derivar con respecto a y

$$\frac{\partial}{\partial y}\int^{x} g_x(x,y)\;dx = \frac{\partial}{\partial y}\ln h(x,y) = g_{c,y}(x,y)$$

De forma análoga se demuestra (4.11)

4.1.2 Forma rectangular del sistema de funciones generadoras de las distribuciones bivariantes continuas.

El siguiente teorema demuestra que, *con la primera función generadora de una función de densidad y la segunda función generadora de cualquier otra función de densidad equivalente a la primera, se obtiene un sistema de funciones generadoras de una tercera función de densidad que pertenece a dicha clase de equivalencia.*

En un caso particular de esta propiedad analítica, meramente estructural, de los sistemas de funciones generadoras de una clase de equivalencia, radica lo que se conoce como principio de verosimilitud. (Un enunciado del mismo se puede ver en Berger, J.O. (1988), [4], pág. 28).

TEOREMA 4.4

Sea $C = \{ f_\lambda(x,y) \}_{\lambda\in\Lambda}$ **una clase de equivalencia de funciones de densidad.**

Sea $G = \left\{ g_{\lambda,x}(x,y),\ g_{\lambda,y}(x,y) \ \middle/ \ g_{\lambda,x}(x,y) = \frac{\partial}{\partial x} \ln f_\lambda(x,y), \right.$

$$\left. g_{\lambda,y}(x,y) = \frac{\partial}{\partial y} \ln f_\lambda(x,y) \ ; \quad \forall f_\lambda(x,y) \in C \right\}$$

la clase de sistemas de funciones generadoras de los elementos de C.

Sea G_1 **la proyección 1 de** G

Sea G_2 **la proyección 2 de** G

Entonces se verifica que $G \equiv G_1 \times G_2$

es decir, la clase constituida por los sistema de generadoras de todas las densidades pertenecientes a $\{ f_\lambda(x,y) \}_{\lambda\in\Lambda}$ **es el producto cartesiano** $\{ g_{\lambda,x}(x,y) \}_{\lambda\in\Lambda} \times \{ g_{\lambda,y}(x,y) \}_{\lambda\in\Lambda}$

En efecto:

Por definición es evidente que $G \subset G_1 \times G_2$, así que el teorema quedará demostrado cuando se compruebe que $G \supset G_1 \times G_2$

Dado un par cualquiera,

$$(g_{1,x}(x,y),\ g_{2,y}(x,y)) \in G_1 \times G_2$$

es evidente que $g_{1,x}(x,y)$ forma parte del sistema de funciones generadoras de una cierta $f_1(x,y)$ de la clase C, luego

$$g_{1,x}(x,y) = \frac{\partial}{\partial x} \ln f_1(x,y)$$

y por tanto

$$\frac{\partial}{\partial y} g_{1,x}(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_1(x,y)$$

Análogamente $g_{2,y}(x,y)$ forma parte del sistema de funciones generadoras de $f_2(x,y)$, densidad de la misma clase de equivalencia C, luego

$$\frac{\partial}{\partial x}\, g_{2,y}(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_2(x,y)$$

Puesto que $f_1(x,y)$ y $f_2(x,y)$ son funciones de la misma clase de equivalencia,

$$\frac{\partial^2}{\partial x\, \partial y} \ln f_1(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_2(x,y)$$

y teniendo en cuenta que:

$$\frac{\partial}{\partial y}\, g_{1,x}(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_1(x,y)$$

$$\frac{\partial}{\partial x}\, g_{2,y}(x,y) = \frac{\partial^2}{\partial x\, \partial y} \ln f_2(x,y)$$

se verifica la igualdad

$$\frac{\partial}{\partial y}\, g_{1,x}(x,y) = \frac{\partial}{\partial x}\, g_{2,y}(x,y)$$

que es la condición necesaria y suficiente para que el par ($g_{1,x}(x,y)$, $g_{2,y}(x,y)$) sea un sistema de funciones generadoras de una densidad conjunta de probabilidad (Proposición 2.2)

Para concluir la demostración sólo falta probar que esta función de densidad, $f_3(x,y)$ pertenece la clase de equivalencia C. Y así es, puesto que:

$$\frac{\partial^2}{\partial x\, \partial y} \ln f_3(x,y) = \frac{\partial}{\partial x}\, g_{2,y}(x,y) = \frac{\partial}{\partial y}\, g_{1,x}(x,y) =$$

$$= \frac{\partial^2}{\partial x\, \partial y} \ln f_2(x,y) = \frac{\partial^2}{\partial x\, \partial y}\, f_1(x,y)$$

Consecuencias de la forma rectangular del sistema de funciones generadoras

Las consecuencias de los teoremas anteriores se pueden resumir en los siguientes apartados:

1° Dada una verosimilitud $f(x/y)$, se obtiene $g_x(x,y)$, y con ella se conoce $g(x,y)$ y por tanto $h(x,y)$, quedando así determinada una clase de equivalencia, $\mathscr{E}$ de densidades de probabilidad $f(x,y)$.

2° Todas las funciones generadoras $g_y(x,y)$ admisibles para la verosimilitud dada, constituyen, según el teorema 4.4, la proyección 2 de dicha clase.

3° Cualquier otra verosimilitud $f_1(x/y)$ correspondiente a una función de densidad de $\mathscr{E}$, (de la que se obtiene $g_{1_x}(x,y)$), admite el mismo "juego" de distribuciones a posteriori.

4° Por el teorema 4.1, la diferencia entre las dos funciones generadoras de la misma clase sólo depende de x:

$$g_{1,x}(x,y) - g_x(x,y) = C(x)$$

al integrar con respecto a x, y puesto que

$$f_1(x/y) = K_1(y)\ e^{\int^x g_{1,x}(x,y)\,dx} \qquad f(x/y) = K(y)\ e^{\int^x g_x(x,y)\,dx}$$

se deduce

$$f_1(x/y) = K\ f(x/y)\ l_1(x)\ l_2(y)$$

5° Como conclusión se puede enunciar que la condición necesaria y suficiente para que dos verosimilitudes $f_1(x/y)$ y $f_2(x/y)$ admitan la misma clase de distribuciones a posteriori es que se verifique la condición

$$f_1(x/y) = f_2(x/y)\ l(x)\ m(y)$$

6° El principio de verosimilitud considera el caso particular en el que $m(y) = 1$.

4.1.3 Distribuciones continuas n-dimensionales. Relación entre funciones de densidad.

La relación de equivalencia definida entre funciones de densidad de distribuciones bidimensionales se generaliza al caso n-dimensional, si bien viene asociada a una partición no trivial del vector aleatorio. Se obtiene la propiedad de rectangularidad de las funciones generadoras, como generalización de la estudiada en el caso bidimensional.

Definición.

Sea $I_1,\ldots,I_r$ con $1 < r \leq n$ **una partición del conjunto N, decimos que las densidades de probabilidad** $f_1(\vec{x})$ **y** $f_2(\vec{x})$ **están relacionadas si y sólo si, se verifica**

$$\frac{\partial^2}{\partial x_i\,\partial x_j}\ln f_1(\vec{x}) = \frac{\partial^2}{\partial x_i\,\partial x_j}\ln f_2(\vec{x}) \quad \forall\, i \in I_l, \quad \forall\, j \in I_s, \quad \forall\, l \neq s \tag{4.12}$$

Obsérvese que si consideramos la matriz de las derivadas de segundo orden del neperiano de las funciones de densidad,

$$\left(g_{i,j}(\vec{x})\right)_{i=1,2,\ldots,n\,;\ j=1,2,\ldots,n} \qquad \text{donde} \qquad g_{i,j}(\vec{x}) = \frac{\partial^2}{\partial x_i\,\partial x_j}\ln f(\vec{x}),$$

la relación se puede expresar en forma matricial: **"dos funciones de densidad son equivalentes, respecto a una partición, si y sólo si los menores de las matrices respectivas, definidos por los subvectores de la partición, coinciden.**

Puesto que no aportan ni una metodología ni unas conclusiones nuevas, y con objeto de hacer una más fluida su lectura, se omiten de esta memoria las demostraciones de los teoremas que llevan a la rectangularidad de los sistemas de funciones generadoras.

Condiciones necesarias y suficientes para la equivalencia.

TEOREMA 4.5

Dos funciones de densidad $f_1(\vec{x})$, $f_2(\vec{x})$ son equivalentes, según la partición $\rho = \{ I_1, \ldots, I_r \}$, si y sólo si la diferencia entre las generadoras de la misma variable x_i , $i \in I_s$ sólo depende de las variable del subvector $\vec{X}_{I_s}$, es decir:

$$g_{2,x_i}(\vec{x}) - g_{1,x_i}(\vec{x}) = C(\vec{X}_{I_s}) \qquad \mathbf{(4.13)}$$

COROLARIO 4.1

Las funciones de densidad $f_1(\vec{x})$ y $f_2(\vec{x})$ son equivalentes según la partición $\rho = \{ I_1, \ldots, I_r \}$, si y sólo si existen funciones reales de variables reales, positivas, $C_s(\vec{x}_{I_s})$ tales que

$$f_2(\vec{x}) = K f_1(\vec{x}) \prod_{s=1}^{r} C_s(\vec{x}_{I_s}) \qquad \mathbf{(4.16)}$$

donde K es un número real.

COROLARIO 4.2

Dos funciones de densidad $f_1(\vec{x})$, $f_2(\vec{x})$ son equivalentes según la partición $\rho = \{ I_1, \ldots, I_r \}$, si y sólo si tienen en común todos los factores que ligan elementos de dos subvectores diferentes de la partición, es decir:

$$f_1(\vec{x}) = h(\vec{x}) \prod_{j=1}^{r} h_1(\vec{x}_{I_j}) \qquad \mathbf{(4.17)}$$

$$f_2(\vec{x}) = h(\vec{x}) \prod_{j=1}^{r} h_2(\vec{x}_{I_j}) \qquad (4.18)$$

Tal y como ocurre en el caso bidimensional, el sistema de funciones generadoras de la función de densidad canónica de la clase a la que pertenecen $f_1(\vec{x})$ y $f_2(\vec{x})$ viene dado por

$$\{ g_{c,x_i}(\vec{x}) \}_{i \in N} \quad \text{con } g_{c,x_i}(\vec{x}) = \frac{\partial}{\partial x_i} \ln h(\vec{x}) \qquad (4.19)$$

La clase de equivalencia en la que $h(\vec{x}) = 1$, igual que en el caso bidimensional, está formada por todas las densidades que hacen a los subvectores $\vec{X}_{I_1},\ldots,\vec{X}_{I_r}$ totalmente independientes. En este caso, todos los menores de la matriz $(g_{i,j}(\vec{x}))_{i=1,2,\ldots,n;\ j=1,2,\ldots,n}$, asociados a la partición tienen los elementos de la diagonal principal uno y el resto de ellos cero.

4.1.4 Forma rectangular del sistema de funciones generadoras de las distribuciones continuas multivariantes.

El siguiente teorema demuestra que, *dada una partición* $\{I_1,\ldots,I_r\}$, *si se eligen* $f_1, \ldots, f_r$, *funciones de densidad en una misma clase de equivalencia y de cada una de ellas se selecciona el subsistema asociado a un elemento de la partición, (diferente al escogido en las demás), el sistema completo así formado genera una densidad que también está en dicha clase de equivalencia.*

TEOREMA 4.6

Dada una partición $I_1, \ldots, I_r$, de N, si llamamos proyección I_l $(1 \le l \le r)$ de la clase de sistemas de generadoras $G = \{ (g_{\lambda,x_i}(\vec{x}))_{i\in N} \}_{\lambda\in\Lambda}$ a la clase $G = \{ (g_{\lambda,x_i}(\vec{x}))_{i\in I_l} \}_{\lambda\in\Lambda}$, entonces se verifica que G es el producto cartesiano de

$$G = \{ (g_{\lambda,x_i}(\vec{x}))_{i\in I_1} \}_{\lambda\in\Lambda} \times\cdots\times \{ (g_{\lambda,x_i}(\vec{x}))_{i\in I_r} \}_{\lambda\in\Lambda}$$

seguido de la ordenación de las funciones generadoras en el orden creciente de los subíndices de las variables.

COROLARIO 4.3

Para cualquier partición obtenida reuniendo en uno sólo al menos dos elementos de la partición $I_1, \ldots, I_r$, los subsistemas ligados a esta nueva partición siguen cumpliendo el teorema de la rectangularidad. En particular, cualquier partición binaria, obtenida de ese modo, cumple la rectangularidad.

Las consecuencias de la forma rectangular del sistema de funciones generadoras son la generalización de las enunciadas para el caso bivariante.

4.1.5 Relación entre distribuciones de probabilidad discretas bidimensionales.

En este epígrafe se realiza un desarrollo paralelo al caso bivariante continuo. La definición de una relación de equivalencia entre distribuciones de probabilidad, a través de las funciones generadoras, nos llevan, también, a las mismas conclusiones

Definición:

Dos distribuciones de probabilidad $p_1(x_r,y_s)$, $p_2(x_r,y_s)$ [1], **son equivalentes si y sólo si las respectivas funciones** [2] $L_1(r,s)$ $L_2(r,s)$ **coinciden, es decir si y sólo si**

$$\frac{p_1(r,s)\ p_1(r-1,s-1)}{p_1(r-1,s)\ p_1(r,s-1)} = \frac{p_2(r,s)\ p_2(r-1,s-1)}{p_2(r-1,s)\ p_2(r,s-1)} \quad \forall r,s > 0 \tag{4.23}$$

Evidentemente la relación así definida es una relación de equivalencia.

La relación (4.23) también se puede establecer en términos de las funciones $g(r,s)$:

$$p_1(r,s)\ R\ p_2(r,s) \leftrightarrow p_1(r-1,s-1)\ g_1(r,s) = p_2(r-1,s-1) g_2(r,s)$$

para todo r y s mayores que cero.

De igual manera que en variables continuas, los teoremas que siguen nos muestran condiciones necesarias y suficientes para que dos distribuciones de probabilidad sean equivalentes.

TEOREMA 4.7

Dos distribuciones de probabilidad $p_1(r,s)$, $p_2(r,s)$ **son equivalentes, si y sólo si el cociente entre las generadoras condicionadas respectivas no depende de la variable que actúa de condicionante, es decir:**

---

(1) En adelante simplificamos la notación: $p(x_r,y_s) \equiv p(r,s)$

(2) Funciones definidas en el epígrafe 2.2.7 de esta memoria.

$$\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)} = C_2(r) \quad \text{no depende de } s \qquad \textbf{(4.24)}$$

$$\frac{L_{1,s}(r,s)}{L_{2,s}(r,s)} = C_2(s) \quad \text{no depende de } r \qquad \textbf{(4.25)}$$

Obsérvese la similitud de las expresiones anteriores con las correspondientes a las distribuciones continuas:

$$g_{2,y}(x,y) - g_{1,y}(x,y) = C_2(y) \quad \text{no depende de } x \qquad (4.7)$$

$$g_{2,x}(x,y) - g_{1,x}(x,y) = C_1(x) \quad \text{no depende de } y \qquad (4.8)$$

Demostración.

a) Es condición necesaria, pues si $p_2(r,s)$ R $p_1(r,s)$ entonces $\frac{L_1(r,s)}{L_2(r,s)} = 1$, y según la definición de la función L,

$$\frac{L_{1,r}(r,s)}{L_{1,r}(r,s-1)} \cdot \frac{L_{2,r}(r,s-1)}{L_{2,r}(r,s)} = 1$$

por lo que

$$\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)} = \frac{L_{1,r}(r,s-1)}{L_{2,r}(r,s-1)}$$

el cociente, $\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)}$ permanece invariante al cambiar s por s-1, luego es una función que sólo depende de r.

Análogamente se demuestra (4.25)

b) Para probar la suficiencia se utiliza el razonamiento seguido en la necesidad en sentido inverso, pues si el cociente $\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)}$

no depende de s, en particular para s y s-1 coincide

$$\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)} = \frac{L_{1,r}(r,s-1)}{L_{2,r}(r,s-1)}$$

luego

$$\frac{L_{1,r}(r,s)}{L_{1,r}(r,s-1)} = \frac{L_{2,r}(r,s)}{L_{2,r}(r,s-1)}$$

es decir las funciones L_1 y L_2 coinciden.

TEOREMA 4.8

Las distribuciones de probabilidad $p_1(r,s)$ y $p_2(r,s)$ son equivalentes si y sólo si, existen dos funciones reales de variable real positivas, $l_1(r)$ y $l_2(s)$, tales que

$$p_2(r,s) = K\, p_1(r,s)\, l_1(r)\, l_2(s) \qquad \textbf{(4.26)}$$

donde K es un número real.

En efecto, si las distribuciones son equivalentes entonces aplicando el teorema 4.7, $\dfrac{L_{2,r}(r,s)}{L_{1,r}(r,s)} = C(r)$, esto es

$$\frac{p_2(r,s)/p_2(r-1,s)}{p_1(r,s)/p_1(r-1,s)} = C(r)$$

por lo que

$$\frac{p_2(r,s)}{p_1(r,s)} = C(r)\,\frac{p_2(r-1,s)}{p_1(r-1,s)}$$

reiterando el procedimiento:

$$\frac{p_2(r,s)}{p_1(r,s)} = \prod_{i=1}^{r} C(i)\,\frac{p_2(0,s)}{p_1(0,s)}$$

y denotando $\quad l_1(r) = \prod_{i=1}^{r} C(i), \qquad l_2(s) = \dfrac{p_2(0,s)}{p_1(0,s)}$

es evidente que se verifica (4.26)

Recíprocamente, si (4.26) es cierta, calculamos el cociente

$$\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)} = \frac{p_1(r,s)/p_1(r-1,s)}{p_2(r,s)/p_2(r-1,s)} = \frac{p_1(r,s)\;p_2(r-1,s)}{p_2(r,s)\;p_1(r-1,s)}$$

sustituyendo p_2 por su valor en función de p_1, y queda:

$$\frac{L_{1,r}(r,s)}{L_{2,r}(r,s)} = \frac{p_1(r,s)\;K\,p_1(r-1,s)\;l_1(r-1)\;l_2(s)}{K\,p_1(r,s)\;l_1(r)\;l_2(s)\;p_1(r-1,s)} = \frac{l_1(r-1)}{l_1(r)}$$

por lo que el cociente $\dfrac{L_{1,r}(r,s)}{L_{2,r}(r,s)}$ sólo depende de r, y según el teorema 4.7, las distribuciones son equivalentes.

TEOREMA 4.9

Dos funciones de cuantía son equivalentes si y sólo si, se pueden expresar de la forma:

$$p_1(r,s) = K_1\, h(r,s)\, l_{11}(r)\, l_{21}(s) \qquad \textbf{(4.27)}$$

$$p_2(r,s) = K_2\, h(r,s)\, l_{12}(r)\, l_{22}(s) \qquad \textbf{(4.28)}$$

En efecto, es condición necesaria, ya que si $p_1(r,s)$ y $p_2(r,s)$ son equivalentes, entonces, por el teorema 4.8, existen $l_1(r)$ y $l_2(s)$, definidas positivas, tales que

$$\frac{p_2(r,s)}{p_1(r,s)} = K\, l_1(r)\, l_2(s)$$

por tanto los posibles factores que dependan de las dos variables han de ser comunes en las dos distribuciones, pues de otro modo el cociente no se simplificaría.

También es condición suficiente puesto que si dividimos (4.28) entre (4.27) se obtiene (4.26), y por tanto son equivalentes.

COROLARIO 4.4

Si dos distribuciones de probabilidad $p_1(r,s)$, $p_2(r,s)$ son equivalentes entonces entre sus respectivas distribuciones condicionadas se verifica:

$$p_1(r/s) = p_2(r/s)\, K_1(r)\, K_2(s) \qquad (4.29)$$

En efecto, pues si las dos distribuciones de probabilidad son equivalentes, por el teorema 4.9 existen funciones que permiten expresarlas de la forma (4.27) y (4.28), por tanto

$$p_1(r/s) = \frac{h(r,s)\, l_{11}(r)}{\sum_r h(r,s)\, l_{11}(r)} = \frac{h(r,s)\, l_{11}(r)}{V_1(s)} \qquad (4.30)$$

$$p_2(r/s) = \frac{h(r,s)\, l_{12}(r)}{\sum_r h(r,s)\, l_{12}(r)} = \frac{h(r,s)\, l_{12})}{V_2(s)} \qquad (4.31)$$

al dividir (4.30) entre (4.31) $\qquad \dfrac{p_1(r/s)}{p_2(r/s)} = \dfrac{l_{11}(r)\, V_2(s)}{l_{12}(r)\, V_1(s)}$

y por tanto se verifica (4.29)

Desde el punto de vista de la factorización, el elemento canónico de una clase de distribuciones equivalentes es la distribución expresada

$$p(r,s) = K\, h(r,s)\; l_1(r)\; l_2(s) \quad \text{tal que } l_1(r) = 1; \quad l_2(s) = 1 \tag{4.32}$$

es decir a aquella función de cuantía que no admite factores función sólo de r ni sólo de s. Tal como ocurría en el caso de distribuciones de probabilidad de tipo continuo.

Identificamos cada elemento del conjunto cociente con la función no factorizable h(r,s).

Las generadoras canónicas tienen la forma:

$$L_{c,r}(r,s) = \frac{h(r,s)}{h(r-1,s)}$$

$$L_{c,s}(r,s) = \frac{h(r,s)}{h(r,s-1)}$$

Obsérvese que

a) Para cualquier distribución de la clase definida por h(r,s) se verifica que

$$L(r,s) = \frac{h(r,s)\; h(r-1,s-1)}{h(r-1,s)\; h(r,s-1)} \tag{4.33}$$

b) Cuando las dos variables son independientes, h(r,s) = 1 para todo r y s mayores que cero, y por tanto todas las distribuciones cuyas variables aleatorias sean estocásticamente independientes, pertenecen a la misma clase de equivalencia, aquella en la que L(r,s) = 1

c) Toda la clase se puede generar desde el elemento canónico, mediante el producto de dos factores positivos $l_1(r)$ y $l_2(r)$. Véase (4.32).

(Idénticos resultados a los obtenidos anteriormente para las distribuciones de probabilidad continuas).

Relación entre las funciones generadoras de elementos de una clase y las funciones generadoras del elemento canónico de dicha clase.

PROPOSICIÓN 4.4

Si $L_{c,r}(r,s)$ y $L_{c,s}(r,s)$ **son las generadoras del representante canónico de una clase de equivalencia, entonces un elemento pertenece a dicha clase si y sólo si sus generadoras admiten la expresión:**

$$L_s(r,s) = L_{c,s}(r,s)\ z_2(s)$$

$$L_r(r,s) = L_{c,r}(r,s)\ z_1(r)$$

Es una consecuencia inmediata del teorema 4.7

Puesto que los elementos que definen las funciones $L_r(r,s)$ y $L_s(r,s)$ son separables en factores que dependen de las dos variables o de una sóla, las generadoras del elemento canónico se leen directamente del sistema de cualquier elemento de la clase, después de factorizar.

PROPOSICIÓN 4.5

Para generar la función de cuantía, representante canónico de una clase de equivalencia, basta conocer sólo una de estas dos funciones: $L_{c,r}(r,s)$ ó $L_{c,s}(r,s)$

En efecto, el elemento canónico es $p_c(r,s) = K_c\, h(r,s)$, donde $h(r,s)$ no admite factores, que sólo dependan de r o que sólo dependan de s; por tanto el cociente $\dfrac{h(r,s)}{h(r-1,s)}$ no es simplificable, lo que significa que en $L_{c,r}(r,s)$ no hay ningún elemento cancelado y siempre se puede escribir de manera que en su numerador se reconozca $h(r,s)$. Localizada dicha función, está calculada $p_c(r,s)$ salvo la constante que se determinará mediante normalización.

COROLARIO 4.5

Dada $L_r(r,s)$ **queda determinada la clase de equivalencia de todas las distribuciones conjuntas de la que forma parte.**

En efecto, en virtud de la proposición 4.4

$$L_r(r,s) = L_{c,r}(r,s)\ z_1(r)$$

el factor $z_1(r)$ es perfectamente reconocible y por tanto separable, quedando determinada entonces $L_{c,r}(r,s)$

Conocida $L_{c,r}(r,s)$ por la proposición 4.5, se construye $p_c(r,s) = K\,h(r,s)$ y la clase de equivalencia está determinada a través de su representante canónico.

COROLARIO 4.6

Dada $L_r(r,s)$ **considerada como una verosimilitud, según el corolario 4.5, se conoce la clase de equivalencia y las distribuciones a posteriori admisibles son de la forma:**

$$L_s(r,s) = L_{c,s}(r,s)\ z_2(s) \qquad \forall\ z_2(s) > 0$$

y como consecuencia a cualquier otra verosimilitud de la clase le corresponde el mismo conjunto de distribuciones a posteriori admisibles.

4.1.6 Forma rectangular del sistema de funciones generadoras.

Se trata de la misma propiedad ya demostrada en el caso de distribuciones bidimensionales continuas.

TEOREMA 4.10

Sea $C = \{ p_\lambda(r,s) \}_{\lambda\in\Lambda}$ **una clase de equivalencia de distribuciones**

Sea $G = \{ L_{\lambda,1}(r,s)\ ,\ L_{\lambda 2}(r,s)\ /\ L_{\lambda,1}(r,s) = \dfrac{p(r,s)}{p(r-1,s)}$

$$L_{\lambda,2}(r,s) = \frac{p(r,s)}{p(r,s-1)}; \qquad \forall\ p_\lambda(r,s) \in C \}$$

la clase de sistemas de generadoras de C**,**

Sea G_1 la proyección 1 de G

Sea G_2 la proyección 2 de G

Entonces se verifica que

$$G \equiv G_1 \times G_2$$

es decir, la clase constituida por los sistema de generadoras de todas las distribuciones pertenecientes a $\{ p_\lambda(r,s) \}_{\lambda\in\Lambda}$ es el producto cartesiano $\{ L_{\lambda,1}(r,s) \}_{\lambda\in\Lambda} \times \{ L_{\lambda,2}(r,s) \}_{\lambda\in\Lambda}$

En efecto:

Dado un par cualquiera,

$$(L_{1,1}(r,s),\ L_{2,2}(r,s)) \in G_1 \times G_2 \quad ^{(1)}$$

$L_{1,1}(r,s)$ forma parte del sistema de generadoras de una distribución $p_1(r,s)$ de la clase C.

Sea $L_1(r,s)$ la función L asociada a esta distribución; para todo r y s mayores que cero se verifica:

$$L_1(r,s) = \frac{L_{1,1}(r,s)}{L_{1,1}(r,s-1)} \tag{4.34}$$

Calculando el producto

$$P_1 = \prod_{i=1}^{r} \left(\prod_{j=1}^{s} L_1(i,j) \right)$$

queda

$$P_1 = \prod_{i=1}^{r} \left(\frac{L_{1,1}(i,1)}{L_{1,1}(i,0)} \frac{L_{1,1}(i,2)}{L_{1,1}(i,1)} \cdots \frac{L_{1,1}(i,s)}{L_{1,1}(i,s-1)} \right) =$$

$$= \prod_{i=1}^{r} \left(\frac{L_{1,1}(i,s)}{L_{1,1}(i,0)} \right) = \frac{\prod_{i=1}^{r} L_{1,1}(i,s)}{\prod_{i=1}^{r} L_{1,1}(i,0)} \tag{4.35}$$

---

(1) Es conveniente recordar que el primer subíndice se refiere a la distribución a la que pertenece y el segundo a la variable respecto de la cual se establece la diferencia.

Análogamente $L_{2,2}(r,s)$ forma parte del sistema de generadoras de $p_2(r,s)$, distribución de la misma clase de equivalencia C, y como consecuencia, calculando el producto correspondiente

$$P_2 = \prod_{j=1}^{s}\left(\prod_{i=1}^{r} L_2(i,j)\right)$$

su valor es

$$P_2 = \frac{\prod_{j=1}^{s} L_{2,2}(r,j)}{\prod_{j=1}^{s} L_{2,2}(0,j)} \tag{4.36}$$

Puesto que las dos distribuciones pertenecen a la misma clase, para todo r y s mayores que cero, ha de ser

$$L_1(r,s) = L_2(r,s)$$

y los producto P_1 y P_2 también lo son. Igualando (4.35) y (4.36) se obtiene la condición necesaria y suficiente para que el par $(L_{1,1}(r,s), L_{2,2}(r,s)) \in G_1 \times G_2$ sea un sistema de generadoras de una distribución discreta bivariante de probabilidad (Proposición 2.12).

Para concluir la demostración sólo falta probar que esta distribución $p_3(r,s)$ pertenece a la clase de equivalencia C.

Es evidente, que $p_3(r,s)$ pertenece a la clase de equivalencia C, pues $L_{1,1}(r,s)$ es la primera componente del sistema de generadoras de $p_3(r,s)$, luego

$$L_3(r,s) = \frac{L_{1,1}(r,s)}{L_{1,1}(r,s-1)} \tag{4.37}$$

de las expresiones (4.34) y (4.37) se deduce que $L_3(r,s) = L_1(r,s)$, es decir $p_1(r,s)$ y $p_3(r,s)$ pertenecen a la misma clase de equivalencia.

Consecuencias de la forma rectangular del sistema de generadoras.

Exactamente las mismas que en el caso de distribuciones de probabilidad de tipo continuo,

1° Dada una verosimilitud se obtiene $L_r(r,s)$, con ella se conoce $L(r,s)$ y por tanto $h(r,s)$, quedando así determinada una clase de equivalencia C de distribuciones de probabilidad $p(r,s)$.

2° Todas las $L_s(r,s)$ admisibles para la verosimilitud dada, constituyen, según el teorema 4.10, la proyección 2 de dicha clase.

3° Cualquier otra verosimilitud $L_{1,r}(r,s)$ correspondiente a una distribución de C admite el mismo "juego" de distribuciones a posteriori.

4° Por el corolario 4.4 la relación entre las dos verosimilitudes que pertenecen a la misma clase es

$$p_1(r/s) = p_2(r/s)\ K_1(r)\ K_2(s)$$

5° Como conclusión se puede enunciar que la condición necesaria y suficiente para que dos verosimilitudes $p_1(r/s)$ y $p_2(r/s)$ admitan la misma clase de distribuciones a posteriori es que se verifique la condición

$$p_1(r/s) = p_2(r/s)\ K_1(r)\ K_2(s)$$

4.1.7 Relación entre distribuciones discretas n-dimensionales.

Igual que se hizo en la generalización del caso continuo, en este epígrafe nos limitamos a establecer las proposiciones correspondientes.

Definición.

Sea $I_1,...,I_r$ con $1 < r \le n$ **una partición del conjunto N, decimos que las distribuciones de probabilidad** $p_1(\vec{r})$ **y** $p_2(\vec{r})$ **están relacionadas si y sólo si, se verifica que**

$$L^1_{i,j}(\vec{r}) = L^2_{i,j}(\vec{r}) \quad \forall\, i \in I_l, \quad \forall\, j \in I_s, \quad \forall\, l \neq s \tag{4.38}$$

TEOREMA 4.11

Dos funciones de cuantía $p_1(\vec{r})$, $p_2(\vec{r})$ **son equivalentes, si y sólo si, el cociente entre las funciones generadoras de la misma variable** r_i , $i \in I_s$ **sólo depende de las variables del subvector** $\vec{r}_{I_s}$**, es decir:**

$$\frac{L^2_i(\vec{r})}{L^1_i(\vec{r})} = C(\vec{r}_{I_s}) \tag{4.39}$$

TEOREMA 4.12

Las distribuciones de probabilidad $p_1(\vec{r})$ **y** $p_2(\vec{r})$ **son equivalentes si y sólo si existen funciones reales de variables reales, positivas,** $C_s(\vec{r}_{I_s})$ **tales que**

$$p_2(\vec{r}) = K\, p_1(\vec{r}) \prod_{s=1}^{r} C_s(\vec{r}_{I_s}) \tag{4.40}$$

donde K es un número real.

TEOREMA 4.13

Dos distribuciones de probabilidad $p_1(\vec{r})$, $p_2(\vec{r})$ son equivalentes según la partición $\wp = \{ I_1, \ldots, I_r \}$, si y sólo si, tienen en común todos los factores que ligan elementos de dos subvectores diferentes de la partición, es decir:

$$p_1(\vec{r}) = h(\vec{r}) \prod_{j=1}^{r} h_1(\vec{r}_{I_j}) \qquad \textbf{(4.41)}$$

$$p_2(\vec{r}) = h(\vec{r}) \prod_{j=1}^{r} h_2(\vec{r}_{I_j}) \qquad (4.42)$$

4.1.8 Forma rectangular del sistema de funciones generadoras.

TEOREMA 4.14

Dada una partición I_1, ..., I_r, de N, si llamamos proyección I_l $(1 \le l \le r)$ de la clase de sistemas de funciones generadoras $G = \left\{ (L_i^{\lambda}(\vec{r}))_{i\in N} \right\}_{\lambda\in\Lambda}$ a la clase $G = \left\{ (g_{\lambda, x_i}(\vec{x}))_{i\in I_1} \right\}_{\lambda\in\Lambda}$, entonces se verifica que G es el producto cartesiano de

$$G = \left\{ (L_i^{\lambda}(\vec{r}))_{i\in I_1} \right\}_{\lambda\in\Lambda} x \ldots x \left\{ (L_i^{\lambda}(\vec{r}))_{i\in I_r} \right\}_{\lambda\in\Lambda}$$

seguido de la ordenación de las funciones generadoras en el orden creciente de los subíndices de las variables.

Las consecuencias de la forma rectangular del sistema de funciones generadoras coinciden con las ya expuestas en los casos anteriores.

4.2 DISTRIBUCIONES MIXTAS DE PROBABILIDAD.

4.2.1 Introducción.

Resulta relativamente frecuente el estudio de distribuciones en las que una variable es discreta y la otra continua. Por ejemplo, X es una variable aleatoria discreta que sigue una binomial $B(n,p)$, donde, a su vez, el parámetro p sigue una distribución beta de parámetros conocidos, [4].

El estudio de esta distribución conjunta lleva al tipo mixto de distribuciones. (Wilks [46], pág. 47).

Sea Y una variable aleatoria continua, cuya densidad es $f(y)$. Sea X una variable aleatoria discreta, cuya función de cuantía, condicionada por Y, viene dada por $p(x/y)$.
Consideramos la función

$$F(x,y) = P[X \leq x \ , \ Y \leq y] = \sum_{r=0}^{x} \int_{-\infty}^{y} p(r/s)\, f(s)\, ds$$

que es definida positiva, monótona creciente y verifica:

$$\sum_{x} \int_{-\infty}^{+\infty} p(x/y)\, f(y)\, dy = \int_{-\infty}^{+\infty} f(y) \ \left(\sum_{x} p(x/y) \right) dy = 1$$

así disponemos de una distribución mixta de probabilidad.

En lo que sigue se considera que

$$f(x,y) = p(x/y)\, f(y) \qquad (4.47)$$

donde $p(x/y)$ es la función de cuantía de la variable discreta X, condicionada a la variable continua Y, $f(y)$ es la función de densidad de la variable Y.

Para la distribución conjunta (X,Y), así definida las distribuciones marginales son

$$p(x) = \int_{X_y} p(x/y)\, f(y)\, dy$$

$$f(y)$$

y las condicionadas

$p(x/y)$ (verosimilitud, discreta)

$$f(y/x) = \frac{p(x/y)\ f(y)}{\int_{\chi_y} p(x/y)\ f(y)\ dy}$$ (posteriori, continua)

4.2.2 Funciones generadoras. Propiedades.

Se definen las funciones generadoras de la forma habitual:

$$L_x(x,y) = \frac{p(x/y)}{p(x-1/y)} \tag{4.48}$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(y/x) \tag{4.49}$$

Manteniendo la notación y los criterios anteriores, la generadora de la distribución marginal de Y, es

$$g(y) = \frac{\partial}{\partial y} \ln f(y)$$

PROPOSICIÓN 4.6

También para el tipo mixto de distribuciones se verifica:

$$L_x(x,y) = \frac{f(x,y)}{f(x-1,y)} \tag{4.50}$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(x,y) \tag{4.51}$$

En efecto, en (4.49) se sustituye $f(y/x)$ por su valor

$$g_y(x,y) = \frac{\partial}{\partial y} \ln p(x/y) + \frac{\partial}{\partial y} \ln f(y) - \frac{\partial}{\partial y} \ln \int_{\chi_y} p(x/y)\ f(y)\ dy$$

y puesto que la integral sólo depende de x, entonces

$$g_y(x,y) = \frac{\partial}{\partial y} \ln p(x/y) + \frac{\partial}{\partial y} \ln f(y) = \frac{\partial}{\partial y} \ln f(x,y)$$

es decir (4.51) es cierto.

Sustituyendo $p(x/y)$ por $\frac{f(x,y)}{f(y)}$ en (4.48)

$$L_x(x,y) = \frac{p(x/y)}{p(x-1/y)} = \frac{\dfrac{f(x,y)}{f(y)}}{\dfrac{f(x-1/y)}{f(y)}} = \frac{f(x,y)}{f(x-1,y)}$$

queda demostrado (4.50)

PROPOSICIÓN 4.7

Dada la distribución $f(x,y) = p(x/y)\, f(y)$ **con las consideraciones de (4.47), se verifica que**

$$E_{p(x/y)}[\, g_y(x,y)\,] = g(y) \tag{4.52}$$

En efecto, si en (4.47) tomamos logaritmos y derivamos con respecto a y queda:

$$\frac{\partial}{\partial y} \ln f(x,y) = \frac{\partial}{\partial y} \ln p(x/y) + \frac{\partial}{\partial y} \ln f(y)$$

es decir

$$g_y(x,y) = \frac{\partial}{\partial y} \ln p(x/y) + g(y) \tag{4.53}$$

calculando el valor esperado de cada uno de los términos, con respecto a $p(x/y)$

$$E_{p(x/y)}[g_y(x,y)] = E_{p(x/y)}\left[\frac{\partial}{\partial y} \ln p(x/y)\right] + E_{p(x/y)}[g(y)] \tag{4.54}$$

si tenemos en cuenta que (Zacks, [48]),

$$E_{p(x/y)}\left[\frac{\partial}{\partial y} \ln p(x/y)\right] = 0$$

y puesto que $g(y)$ no depende de x, se verifica

$$E_{p(x/y)}[g(y)] = g(y)$$

entonces de (4.54) se deduce la certeza de (4.52)

PROPOSICIÓN 4.8

Dada la distribución de probabilidad $f(x,y) = p(x/y)\, f(y)$ **con las consideraciones de (4.47), se verifica que**

$$g_y(x,y) = g_y(x-1,y) + \frac{\partial}{\partial y} \ln L_x(x,y) \tag{4.55}$$

Resulta evidente al tomar logaritmos en (4.48) y derivar con respecto a y.

Queremos hacer notar que este mismo resultado se obtiene si, utilizando las definiciones (4.48) y (4.49), calculamos la diferencia entre las dos generadoras.

$$g_y(x,y) - g_y(x-1,y) = \frac{\partial}{\partial y} \ln p(x/y) - \frac{\partial}{\partial y} \ln p(x-1/y) =$$

$$= \frac{\partial}{\partial y} \ln \frac{p(x/y)}{p(x-1/y)} = \frac{\partial}{\partial y} \ln L_x(x,y)$$

PROPOSICIÓN 4.9

Entre las funciones generadoras de una distribución mixta, definida en (4.47) se verifica la relación

$$g(y) = g_y(0,y) - \frac{\partial}{\partial y} \ln p(0/y) \qquad (4.56)$$

Su demostración es evidente, pues es la igualdad (4.53) en el caso particular de que $x = 0$.

Se ha destacado la relación (4.56) porque se utiliza para definir la función generadora de la distribución marginal de y, necesaria en el proceso de reconstrucción de $f(x,y)$ a partir del sistema de funciones generadoras, como más adelante se detalla.

Cabe destacar también que la expresión (4.56) se puede obtener a partir de (4.55):
Utilizando recurrencia sobre x,

$$g_y(x,y) = g_y(0,y) + \sum_{i=1}^{x} \frac{\partial}{\partial y} \ln L_x(i,y)$$

$$g_y(x,y) = g_y(0,y) + \frac{\partial}{\partial y} \ln \prod_{i=1}^{x} L_x(i,y)$$

Pero

$$\prod_{i=1}^{x} L_x(i,y) = \frac{p(1/y)}{p(0/y)} \frac{p(2/y)}{p(1/y)} \cdots \frac{p(x/y)}{p(x-1/y)} = \frac{p(x/y)}{p(0/y)}$$

luego

$$g_y(x,y) = g_y(0,y) + \frac{\partial}{\partial y} \ln p(x/y) - \frac{\partial}{\partial y} \ln p(0/y)$$

Al calcular ahora el valor esperado referido a $p(x/y)$, puesto que $E_{p(x/y)}[\,g_y(x,y)\,] = g(y)$, $E_{p(x/y)}\left[\frac{\partial}{\partial y}\ln p(x/y)\right] = 0$ y que $g_y(0,y)$ y $\frac{\partial}{\partial y}\ln p(0/y)$ no dependen de x, se tiene la expresión deseada.

4.2.3 Sistema de funciones generadoras.

El siguiente teorema establece las condiciones necesarias y suficientes para que un conjunto de dos funciones, cada una de ellas generadora de la respectiva distribución unidimensional, sea un sistema de funciones generadoras de la distribución de tipo mixto descrita en (4.47)

TEOREMA 4.15

La condición necesaria y suficiente para que el conjunto de funciones $\{L(x,y), g(x,y)\}$, donde $L(x,y)$ es función generadora de una distribución discreta, $p(x/y)$, y $g(x,y)$ es función generadora de una distribución continua $f(y/x)$, sea un sistema de funciones generadoras de una distribución de probabilidad de tipo mixto, (4.47) es que se verifique

$$g(x,y) = g(x-1,y) + \frac{\partial}{\partial y}\ln L(x,y) \qquad (4.57)$$

En efecto:

De la proposición 4.8 se desprende que (4.57) es condición necesaria.

Para probar la suficiencia se realiza el proceso constructivo de la distribución de probabilidad y después se prueba que las funciones generadoras de la misma, calculadas a partir de las definiciones (4.48) y (4.49), coinciden con las funciones $L(x,y)$ y $g(x,y)$.

1° El primer paso consiste en generar la distribución $p(x/y)$ a partir de la función $L(x,y)$. (Véase la generación de una distribución discreta en el capítulo 1).

$$p(x/y) = \frac{\prod_{i=1}^{x} L(i,y)}{1 + \sum_{j} \prod_{i=1}^{j} L(i,y)} \qquad \forall\, x>0$$

$$p(0/y) = \frac{1}{1 + \sum_{j} \prod_{i=1}^{j} L(i,y)}$$

se puede escribir

$$p(x/y) = p(0/y) \prod_{i=1}^{x} L(i,y) \qquad \forall\, x>0$$

2° Para construir la distribución de probabilidad $f(y)$, motivados por la igualdad (4.56), se define su función generadora por (4.56) como

$$g(y) = g_y(0,y) - \frac{\partial}{\partial y} \ln p(0/y)$$

y en este caso

$$f(y) = K\, e^{\int^y g_y(0,y)\, dy - \int^y \frac{\partial}{\partial y} \ln p(0/y)\, dy}$$

$$f(y) = \frac{K\, e^{\int^y g_y(0,y)\, dy}}{p(0/y)}$$

3° Calculamos la función $f(x,y)$,

$$f(x,y) = p(x/y)\, f(y) = K \prod_{i=1}^{x} L(i,y)\; e^{\int g(0,y)\, dy} \qquad (4.58)$$

4° Calculamos ahora las funciones generadoras a partir de (4.58), y resultan ser:

$$L_x(x,y) = \frac{f(x,y)}{f(x-1,y)} = \frac{\prod_{i=1}^{x} L(i,y)}{\prod_{i=1}^{x-1} L(i,y)} = L(x,y)$$

$$g_y(x,y) = \frac{\partial}{\partial y} \ln f(x,y) = \sum_{i=1}^{x} \frac{\partial}{\partial y} \ln L(i,y) + g(0,y)$$

puesto que se ha de verificar (4.57), se sustituye $L(i,y)$ por $g(i,y) - g(i-1,y)$ y entonces

$$g_y(x,y) = \sum_{i=1}^{x} [g(i,y) - g(i-1,y)] + g(0,y)$$

$$g_y(x,y) = [g(x,y) - g(0,y)] + g(0,y) = g(x,y)$$

se observa que cada una de ellas coincide con las funciones de las que inicialmente hemos partido.

4.2.4 Ejemplos

1. Dadas las funciones:

$$L(x,y) = \frac{y}{x}$$

$$g(x,y) = -3 + \frac{x}{y}$$

definidas para $x = 0, 1, 2, \ldots \;;\quad 0 < y < +\infty$

Se verifica la condición (4.57),

$$g_y(x,y) = g_y(x-1,y) + \frac{\partial}{\partial y} \ln L_x(x,y)$$

pues

$$g(x-1,y) = -3 + \frac{x-1}{y}$$

$$\frac{\partial}{\partial y} \ln L(x,y) = \frac{1}{y}$$

Se genera, a partir de la función L, la distribución $p(x/y)$ (resuelto en el epígrafe 1.3.2), cuya solución es la distribución de Poisson.

$$p(x/y) = e^{-y}\,\frac{y^x}{x!}$$

y por tanto

$$p(0/y) = e^{-y}$$

$$\frac{\partial}{\partial y}\ln p(0/y) = -1$$

Se calcula ahora g(y)

$$g(y) = g(0,y) - \frac{\partial}{\partial y}\ln p(0/y) = -3 - (-1) = -2$$

luego f(y) es la exponencial negativa

$$f(y) = K\,e^{\int^{y}(-2)\,dy}$$

$$f(y) = 2\,e^{-2y}$$

siendo la distribución conjunta de la forma

$$f(x,y) = p(x/y)f(y) = e^{-y}\frac{y^x}{x!}\,2\,e^{-2y} = 2\,e^{-3y}\,\frac{y^x}{x!}$$

2. Dadas las funciones:

$$L(x,y) = \frac{(n - x + 1)\,y}{x\,(1 - y)}$$

$$g(x,y) = \frac{x + p - 1}{y} + \frac{x + 1 - n - q}{1 - y}$$

definidas para $x = 0,1,2,\ldots,n\,;\quad 0 \le y \le 1\,;\quad p,\, q > 0$

Se verifica la condición (4.57), pues

$$g(x,y) = \frac{x + p - 2}{y} + \frac{x - n - q}{1 - y}$$

$$\frac{\partial}{\partial y}\ln L(x,y) = \frac{1}{y} + \frac{1}{1 - y}$$

Se genera, a partir de la función L, la distribución p(x/y), cuya solución es la distribución binomial

$$\prod_{i=1}^{j} L(i,y) = \frac{n(n-1)\ldots(n-j+1)}{1.2.3\ldots j}\,\frac{y^j}{(1-y)^j} = \frac{n!}{j!\,(n-j)!}\,\frac{y^j}{(1-y)^j}$$

$$\sum_j \prod_{i=1}^{j} L(i,y) = \binom{n}{1}\frac{y}{1-y} + \binom{n}{2}\frac{y^2}{(1-y)^2} + \ldots + \binom{n}{n}\frac{y^n}{(1-y)^n}$$

entonces

$$(1-y)^n \sum_j \prod_{i=1}^{j} L(i,y) = \binom{n}{1} y\,(1-y)^{n-1} + \ldots + \binom{n}{n-1} y^{n-1}(1-y) + \binom{n}{n} y^n$$

$$(1-y)^n \sum_j \prod_{i=1}^{j} L(i,y) = [y + (1-y)]^n - \binom{n}{0}(1-y)^n = 1 - (1-y)^n$$

$$\sum_j \prod_{i=1}^{j} L(i,y) = (1-y)^{-n} - 1$$

luego

$$p(0/y) = \frac{1}{1 + (1-y)^{-n} - 1} = (1-y)^n$$

y por tanto $\quad \frac{\partial}{\partial y} \ln p(0/y) = \frac{-n}{1-y}$

siendo además

$$p(x/y) = p(0/y) \prod_{i=1}^{x} L(i,y) = (1-y)^n \binom{n}{x} \frac{y^x}{(1-y)^x} = \binom{n}{x} y^x (1-y)^{n-x}$$

Se calcula ahora $g(y)$

$$g(y) = g(0,y) - \frac{\partial}{\partial y} \ln p(0/y) = \frac{p-1}{y} + \frac{1-n-q}{1-y} + \frac{n}{1-y} = \frac{p-1}{y} + \frac{1-q}{1-y}$$

luego $f(y)$ es la distribución beta de parámetros p y q

$$f(y) = K\, e^{\int^y \left(\frac{p-1}{y} + \frac{1-q}{1-y}\right) dy}$$

$$f(y) = K\, e^{\ln y^{p-1} + \ln (1-y)^{q-1}} = K\, y^{p-1} (1-y)^{q-1}$$

siendo la distribución conjunta de la forma

$$f(x,y) = p(x/y)\, f(y) = \frac{1}{B(p,q)} \binom{n}{x} y^x (1-y)^{n-x} y^{p-1} (1-y)^{q-1}$$

$$f(x,y) = \frac{1}{B(p,q)} \binom{n}{x} y^{x+p-1} (1-y)^{n-x+q-1}$$

que corresponde a la beta-binomial

4.2.5 Rectangularidad de los sitemas de funciones generadoras de las distribuciones de tipo mixto

Definición:

Dos funciones $f_1(x,y)$, $f_2(x,y)$ **, son equivalentes si y sólo si**

$$\frac{\partial}{\partial y} \ln L_x^1(x,y) = \frac{\partial}{\partial y} \ln L_x^2(x,y) \tag{4.59}$$

Esta misma relación se puede definir utilizando las funciones generadoras continuas:

$$f_1(x,y)\ R\ f_2(x,y) \Leftrightarrow g_y^1(x,y) - g_y^1(x-1,y) = g_y^2(x,y) - g_y^2(x-1,y)$$

Obsérvese que

$$\frac{\partial}{\partial y} \ln L_x^1(x,y) = \frac{\partial}{\partial y} \ln \frac{f_1(x,y)}{f_1(x-1,y)} =$$

$$\frac{\partial}{\partial y} \ln f_1(x,y) - \frac{\partial}{\partial y} \ln f_1(x-1,y) = g_y^1(x,y) - g_y^1(x-1,y)$$

análogamente para la función $L_x^2(x,y)$, por lo que las dos definiciones coinciden.

TEOREMA 4.16

Dos funciones $f_1(x,y)$, $f_2(x,y)$ **son equivalentes, si y sólo si**

$$\frac{L_x^1(x,y)}{L_x^2(x,y)} = C_1(x) \tag{4.60}$$

Demostración.

Es condición necesaria, pues si $f_2(x,y)\ R\ f_1(x,y)$ entonces

$$\frac{\partial}{\partial y} \ln L_x^1(x,y) = \frac{\partial}{\partial y} \ln L_x^2(x,y)$$

$$\frac{\partial}{\partial y} \ln \frac{L_x^1(x,y)}{L_x^2(x,y)} = 0$$

por tanto se verifica (4.60)

b) Para probar la suficiencia basta tomar logaritmos y después derivar, con respecto a x, la expresión (4.60); se deduce la igualdad en las derivadas citadas y por tanto la equivalencia de las funciones f_1 y f_2

TEOREMA 4.17

Dos densidades $f_1(x,y)$, $f_2(x,y)$ son equivalentes, si y sólo si

$$g_y^1(x,y) - g_y^2(x,y) = C_2(y) \qquad \textbf{(4.61)}$$

La demostración es inmediata a partir del teorema 4.16 y la igualdad (4.47).

TEOREMA 4.18

$f_1(x,y)$ y $f_2(x,y)$ son equivalentes si y sólo si existen dos funciones reales de variable real, positivas, $l_1(x)$, $l_2(y)$ tales que

$$f_2(x,y) = K.f_1(x,y).l_1(x).l_2(y) \qquad \textbf{(4.62)}$$

donde K es un número real.

En efecto, si $f_1(x,y)$ y $f_2(x,y)$ son equivalentes entonces aplicando el teorema 4.16, $\dfrac{L_x^2(x,y)}{L_x^1(x,y)} = C(x)$,es decir

$$\frac{f_2(x,y)}{f_1(x,y)} = C(x)\ \frac{f_2(x-1,y)}{f_1(x-1,y)}$$

reiterando el procedimiento:

$$\frac{f_2(x,y)}{f_1(x,y)} = \prod_{i=1}^{x} C(i)\ \frac{f_2(0,y)}{f_1(0,y)}$$

y denotando $l_1(x) = \prod_{i=1}^{x} C(i)$, $l_2(y) = \dfrac{f_2(0,y)}{f_1(0,y)}$

es evidente que se verifica (4.62)

Recíprocamente, si (4.62) es cierta, calculamos el cociente

$$\frac{L_x^1(x,y)}{L_x^2(x,y)} = \frac{f_1(x,y)\ f_2(x-1,y)}{f_2(x,y)\ f_1(x-1,y)} =$$

sustituyendo f_2 por su valor en función de f_1, y queda:

$$\frac{f_1(x,y)\ K\ f_1(x-1,y)\ l_1(x-1)\ l_2(y)}{K\ f_1(x,y)\ l_1(x)\ l_2(y)\ f_1(x-1,y)} = \frac{l_1(x-1)}{l_1(x)}$$

por lo que el cociente $\frac{L_x^1(x,y)}{L_x^2(x,y)}$ sólo depende de x, y según el teorema 4.16, las distribuciones de probabilidad son equivalentes.

TEOREMA 4.19

Dos distribuciones de probabilidad son equivalentes si y sólo si, se pueden expresar de la forma:

$$f_1(x,y) = K_1\ h(x,y)\ l_{11}(x)\ l_{21}(y) \qquad \textbf{(4.63)}$$

$$f_2(x,y) = K_2\ h(x,y)\ l_{12}(x)\ l_{22}(y) \qquad \textbf{(4.64)}$$

La demostración coincide con la del teorema similar del caso discreto bidimensional.

TEOREMA 4.20

Dos funciones de densidad son equivalentes si y sólo si, la forma factorizada de ambas tiene común el factor h(x,y)
Es válida la demostración realizada en el caso bidimensional continuo.

Desde el punto de vista de la factorización, el elemento canónico de una clase de densidades equivalentes es la función $f(x,y) = K\ h(x,y)\ l_1(x)\ l_2(y)$ tal que $l_1(x) = 1$; $l_2(y) = 1$, es decir aquella función que no admite factores que dependan analíticamente de una sola variable.

TEOREMA 4.21

Sea $C = \{ f_\lambda(x,y) \}_{\lambda \in \Lambda}$ **una clase de equivalencia de funciones de densidad.**

Sea $G = \left\{ L_x^\lambda(x,y) \, , \, g_y^\lambda(x,y) \;\Big/\; L_x^\lambda(x,y) = \dfrac{f_\lambda(x,y)}{f_\lambda(x-1,y)} , \right.$

$$\left. g_y^\lambda(x,y) = \frac{\partial}{\partial y} \ln f_\lambda(x,y) \; ; \qquad \forall\, f_\lambda(x,y) \in C \right\}$$

la clase de sistemas de funciones generadoras de las funciones de densidad de C.

Sea G_1 **la proyección 1 de** G

Sea G_2 **la proyección 2 de** G

Entonces se verifica que $G \equiv G_1 \times G_2$

es decir, la clase constituida por los sistema de generadoras de todas las densidades pertenecientes a $\{ f_\lambda(x,y) \}_{\lambda \in \Lambda}$ **es el producto cartesiano** $\{ L_x^\lambda(x,y) \}_{\lambda \in \Lambda} \times \{ g_y^\lambda(x,y) \}_{\lambda \in \Lambda}$

En efecto, dado un par cualquiera,

$$(L_x^1(x,y) \, , \, g_y^2(x,y) \,) \in G_1 \times G_2$$

es evidente que $L_x^1(x,y)$ forma parte del sistema de generadoras de una cierta $f_1(x,y)$ de la clase C.

Análogamente $g_y^2(x,y)$ forma parte del sistema de generadoras de $f_2(x,y)$, luego se ha cumplir (4.55), es decir

$$g_y^2(x,y) - g_y^2(x-1,y) = \frac{\partial}{\partial y} \ln L_x^2(x,y) \tag{4.65}$$

Puesto que $f_1(x,y)$ y $f_2(x,y)$ son funciones de densidad de la misma clase de equivalencia, se cumple (4.59), esto es

$$\frac{\partial}{\partial y} \ln L_x^1(x,y) = \frac{\partial}{\partial y} \ln L_x^2(x,y)$$

entonces, sustituyendo en (4.65) queda

$$g_y^2(x,y) - g_y^2(x-1,y) = \frac{\partial}{\partial y} \ln L_x^1(x,y)$$

que es la condición que debe cumplir el par $(L_x^1(x,y),\ g_y^2(x,y)) \in G_1 \times G_2$ para que sea sistema de generadoras de una distribución $f_3(x,y)$.

Para concluir la demostración sólo falta probar que $f_3(x,y)$ pertenece la clase de equivalencia C. Y así es, puesto que la función generadora $L_x(x,y)$ la tiene en común con f_1, y la relación de equivalencia se establece sobre estas funciones generadoras.

Consecuencias de la forma rectangular del sistema de generadoras

Las consecuencias de la rectangularidad coinciden con las expuestas en todos los demás casos.

Como conclusión se puede enunciar que la condición necesaria y suficiente para que dos verosimilitudes $f_1(x/y)$ y $f_2(x/y)$, (X discreta, Y continua), admitan la misma clase de distribuciones a posteriori es que se verifique la condición

$$f_1(x/y) = f_2(x/y)\ l(x)\ m(y)$$

El principio de verosimilitud considera el caso particular en el que $m(y) = 1$.

4.3 LAS FUNCIONES GENERADORAS Y LAS FAMILIAS CONJUGADAS DE DISTRIBUCIONES A PRIORI.

La definición sobre familias conjugadas, adoptada en esta memoria es la dada por Berger, ([4], pág. 130):

"Denotamos por $\mathcal{F}$ la clase de funciones de densidad $f(x/\theta)$ **(indexadas por θ). Una clase $\wp$ de distribuciones a priori se dice Familia Conjugada para $\mathcal{F}$ si la distribución a posteriori** $\pi(\theta/x)$ **pertenece a la clase $\wp$,** $\forall f \in \mathcal{F}$ **y** $\pi \in \wp$**".**

En lo que sigue cuando utilicemos un elemento de la familia de verosimilitudes, con parámetro desconocido, le llamaremos verosimilitud.

La proposición 2.3 establece una relación entre las funciones generadoras de las distribuciones marginales y condicionadas, de la forma:

$$\frac{\partial}{\partial y} \ln f(x/y) = g_y(x,y) - g(y)$$

La proposición 3.2 generaliza esta relación para distribuciones multivariantes:

$$\frac{\partial}{\partial x_i} \ln f(\vec{x}_i/x_i) = g_i(\vec{x}) - g(x_i) \qquad \forall\, i \in N$$

A partir de esta última igualdad, considerando una variable n-dimensional $\vec{x}$ y un parámetro θ se puede escribir una relación entre las funciones generadoras de las distribuciones a priori y a posteriori y la derivada parcial del logaritmo de la verosimilitud con respecto al parámetro objeto de estudio:

$$\frac{\partial}{\partial \theta} \ln f(\vec{x}/\theta) = g_\theta(\vec{x},\theta) - g(\theta) \tag{4.66}$$

La relación (4.66) permite, comprobar de una forma sencilla si una familia de distribuciones a priori es conjugada o no para una clase de verosimilitudes y, conocidas las distribuciones a priori y las verosimilitudes, conocer las funciones generadoras de las distribuciones a posteriori.

En la tabla que aparece más adelante se da una relación de familias conjugadas, suficientemente conocida, (Berger, [4];

Wolf-Rüdiger, [47]), pero que en esta memoria es tratada desde el punto de vista de las funciones generadoras.

A continuación se desarrollan algunos ejemplos de los que aparecen en la tabla.

4.3.1 Familias de distribuciones unidimensionales.

1. La familia $N(\mu,\sigma^2)$, es conjugada para la verosimilitud de una muestra, procedente de una población $N(\theta,\tau^2)$ de parámetro θ desconocido, siendo los restantes parámetros τ^2, μ, σ^2 conocidos

Puesto que θ sigue una distribución a priori $\theta \sim N(\mu,\sigma^2)$, entonces

$$f(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(\theta-\mu)^2}$$

y la función generadora es:

$$g(\theta) = -\frac{1}{\sigma^2}\theta + \frac{\mu}{\sigma^2}$$

Por otra parte, dada

$$f(x/\theta) = \frac{1}{\sqrt{2\pi}\,\tau} e^{-\frac{1}{2\tau^2}(x-\theta)^2}$$

se tiene

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = \frac{x}{\tau^2} - \frac{\theta}{\tau^2}$$

Para una muestra aleatoria simple, de dicha variable aleatoria,

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \frac{\sum_{i=1}^{n} x_i}{\tau^2} - \sum_{i=1}^{n} \frac{\theta}{\tau^2} = \frac{n\overline{x}}{\tau^2} - \frac{n\theta}{\tau^2}$$

Aplicando (4.66), la función generadora de la distribución a posteriori es:

$$g(\theta/\vec{x}) = \frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) + g(\theta) = -\left(\frac{n}{\tau^2} + \frac{1}{\sigma^2}\right)\theta + \left(\frac{n\overline{x}}{\tau^2} + \frac{\mu}{\sigma^2}\right)$$

que corresponde a una distribución normal, donde el coeficiente de θ es el opuesto del inverso de la varianza, y el término independiente el cociente entre la media y la varianza.
Luego la distribución a posteriori es

$$N\left(\frac{n\bar{x}\sigma^2 + \mu\tau^2}{n\sigma^2 + \tau^2}, \frac{\tau^2\sigma^2}{n\sigma^2 + \tau^2}\right)$$

2. La familia $Be(\alpha,\beta)$, es conjugada para la verosimilitud de una muestra procedente de una población $BN(r,\theta)$ de parámetro θ desconocido.

En efecto, si θ sigue una distribución a priori $\theta \sim Be(\alpha,\beta)$, entonces

$$f(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

y la función generadora es:

$$g(\theta) = \frac{\alpha-1}{\theta} - \frac{\beta-1}{1-\theta}$$

Dada

$$f(x/r,\theta) = \binom{r+x-1}{x}\theta^r(1-\theta)^x$$

con θ desconocida y r conocida, se tiene

$$\frac{\partial}{\partial\theta}\ln f(x/r,\theta) = \frac{r}{\theta} - \frac{x}{1-\theta}$$

para una muestra aleatoria simple, procedente de esta distribución es

$$\frac{\partial}{\partial\theta}\ln f(\vec{x}/r,\theta) = \frac{nr}{\theta} - \frac{n\bar{x}}{1-\theta}$$

Aplicando (4.66), la función generadora de la distribución a posteriori resulta ser:

$$g(\theta/r,\vec{x}) = \frac{\partial}{\partial\theta}\ln f(\vec{x}/r,\theta) + g(\theta) = \frac{nr+\alpha-1}{\theta} - \frac{n\bar{x}+\beta-1}{1-\theta}$$

que corresponde a una distribución beta, donde el numerador de la primera fracción más uno es el primer parámetro y el

numerador de la segunda fracción más uno es el segundo parámetro.
Por tanto la distribución a posteriori es

$$Be(nr + \alpha, n\bar{x} + \beta)$$

3. Es conocido que para la verosimilitud de una muestra procedente de una variable aleatoria con función de densidad

$$f(x/\theta) = e^{A(x)\theta + B(x) - \psi(\theta)}$$

una familia conjugada es

$$f(\theta) = K(\mu, \lambda)\, e^{\theta\mu - \lambda\psi(\theta)}$$

(Lee, [30]; Robert, [40]).

En efecto, la función generadora de la distribución a priori es:

$$g(\theta) = \mu - \lambda\psi'(\theta)$$

Teniendo en cuenta que

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = A(x) - \psi'(\theta)$$

entonces, para una muestra aleatoria simple es

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \sum_{i=1}^{n} A(x_i) - n\psi'(\theta)$$

aplicando (4.66), la función generadora de la distribución a posteriori es:

$$g(\theta/\mu, \lambda, \vec{x}) = \left(\mu + \sum_{i=1}^{n} A(x_i)\right) - (n+\lambda)\,\psi'(\theta)$$

que corresponde a una distribución del mismo tipo que la priori, donde los nuevos parámetros son $\left(\mu + \sum_{i=1}^{n} A(x_i),\ n + \lambda\right)$.

En particular si $A(x) = x$, la función generadora de la distribución es

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \sum_{i=1}^{n} A(x_i) - n\psi'(\theta)$$

y sus parámetros

$$(\mu + n\bar{x},\ n + \lambda)$$

DISTRIBUCIÓN A PRIORI. Generadoras $g(\theta)$ (1)	VEROSIMILITUDES. Parcial del logaritmo. $\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta)$ (2)	GENERADORAS DE DISTRIBUC. A POSTERIORI. $g(\theta/\vec{x}) = (1) + (2)$	DISTRIBUC. A POSTERIORI Parámetros
Beta (p,q) $\frac{p-1}{\theta} - \frac{q-1}{1-\theta}$	Binomial (n,θ) $\frac{x}{\theta} - \frac{n-x}{1-\theta}$	$\frac{x+p-1}{\theta} - \frac{n+q-x-1}{1-\theta}$	Beta $(p+x,\ n+q-x)$
Beta (p,q) $\frac{p-1}{\theta} - \frac{q-1}{1-\theta}$	Bin. negat. (r,θ) $\frac{nr}{\theta} - \frac{n\overline{x}}{1-\theta}$	$\frac{nr+p-1}{\theta} - \frac{n\overline{x}+q-1}{1-\theta}$	Beta $(nr+p,\ n\overline{x}+q)$
Beta (p,q) $\frac{p-1}{\theta} - \frac{q-1}{1-\theta}$	Geométrica (θ) $\frac{n}{\theta} - \frac{n\overline{x}}{1-\theta}$	$\frac{n+p-1}{\theta} - \frac{n\overline{x}+q-1}{1-\theta}$	Beta $(n+p, n\overline{x}+q)$
Gamma (p,a) $\frac{p-1}{\theta} - a$	Poisson (θ) $-n + \frac{n\overline{x}}{\theta}$	$\frac{n\overline{x}+p-1}{\theta} - (a+n)$	Gamma $(p+n\overline{x},\ a+n)$
Gamma (p,a) $\frac{p-1}{\theta} - a$	Exponencial (θ) $\frac{n}{\theta} - n\overline{x}$	$\frac{p+n-1}{\theta} - (a+n\overline{x})$	Gamma $(p+n,\ a+n\overline{x})$
Gamma (p,a) $\frac{p-1}{\theta} - a$	Gamma (ν,θ) $\frac{n\nu}{\theta} - n\overline{x}$	$\frac{n\nu+p-1}{\theta} - (a+n\overline{x})$	Gamma $(n\nu+p, a+n\overline{x})$

Gamma (p,a) $\frac{p-1}{\theta}-a$	Uniforme en $(0,\theta)$ $-\frac{n}{\theta}$	$\frac{p-n-1}{\theta}-a$	si $p-n>0$ Gamma $(p-n,\ a)$
Gamma (p,a) $\frac{p-1}{\theta}-a$	Normal (μ,σ^2) siendo $\theta=\frac{1}{\sigma^2}$ $\frac{n/2}{\theta}-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2$	$\frac{\frac{n}{2}+p-1}{\theta}-\left(a+\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$	Gamma $\left(\frac{n}{2}+p,\ a+\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$
Gamma (p,a) $\frac{p-1}{\theta}-a$	Pareto (x_0,θ) $\frac{n}{\theta}+\ln x_0^n-\sum_{i=1}^{n}\ln x_i$	$\frac{p+n-1}{\theta}-\left(a-\ln x_0^n+\sum_{i=1}^{n}\ln x_i\right)$	Si $a-\ln x_0^n+\sum_{i=1}^{n}\ln x_i>0$ Gamma $(p+n,\ a-\ln x_0^n+\sum_{i=1}^{n}\ln x_i)$
Gamma (p,a) $\frac{p-1}{\theta}-a$	Pareto (θ,a) $\frac{na}{\theta}$	$\frac{na+p-1}{\theta}-a$	Gamma $(p+na,\ a)$
Gamma (p,a) $\frac{p-1}{\theta}-a$	Maxwell (θ) $\frac{n}{\theta}-\frac{1}{2}\sum x_i^2$	$\frac{p+n-1}{\theta}-\left(a+\frac{1}{2}\sum_{i=1}^{n}x_i^2\right)$	Gamma $\left(p+n,\ a+\frac{1}{2}\sum x_i^2\right)$
Normal (μ,σ^2) $-\frac{1}{\sigma^2}\theta+\frac{\mu}{\sigma^2}$	Normal (θ,τ^2) $\frac{n\bar{x}}{\tau^2}-\frac{n}{\tau^2}\theta$	$\frac{n\bar{x}}{\tau^2}+\frac{\mu}{\sigma^2}-\left(\frac{n}{\tau^2}+\frac{1}{\sigma^2}\right)\theta$	Normal $\left(\frac{n\bar{x}\sigma^2+\mu\tau^2}{n\sigma^2+\tau^2},\ \frac{\tau^2\sigma^2}{n\sigma^2+\tau^2}\right)$
Pareto (ω,a) $-\frac{a+1}{\theta}$	Uniforme en $(0,\theta)$ $-\frac{n}{\theta}$	$-\frac{a+n+1}{\theta}$	Pareto $(\omega',\ a+n)$ $\omega'=\text{máx}(\omega,x_1..x_n)$

Lognormal (μ, σ^2) $-\frac{1}{\theta} - \frac{\frac{1}{\sigma^2}(\ln\theta-\mu)}{\theta}$	Uniforme en $(0, \theta)$ $-\frac{n}{\theta}$	$-\frac{1}{\theta} - \frac{\frac{1}{\sigma^2}(\ln\theta-\mu+\sigma^2 n)}{\theta}$	Lognormal $(\mu - n\sigma^2, \sigma^2)$
Maxwell (σ^2) $\frac{2}{\theta}-\sigma^2\theta$	$f(x/\theta) = e^{-A(x)\theta^2}$ $-2\sum_{i=1}^{n} A(x_i)\,\theta$	$\frac{2}{\theta}-\left(\sigma^2+2\sum_{i=1}^{n} A(x_i)\right)\theta$	Maxwell $\sigma^2+2\sum_{i=1}^{n} A(x_i)$
.....	.....	.....	.....

Obsérvese el caso de la familia gamma conjugada para una familia de verosimilitudes normal de media conocida y varianza desconocida. El parámetro que sigue una distribución gamma no es la varianza, sino su inverso, es decir la precisión.

Cabe preguntarse si para una familia de verosimilitudes normal de media conocida y varianza desconocida, la familia gamma es conjugada con ella.

Sea la familia de distribuciones $N(\mu, \sigma^2)$, donde μ es conocido y σ^2, desconocido, sigue una gamma de parámetros p,a.

En estas condiciones la familia gamma no es conjugada para la familia de verosimilitudes normal.

Sea

$$f(x/\theta) = \frac{1}{\sqrt{2\pi\theta}}\, e^{-\frac{1}{2\theta}(x-\mu)^2}$$

entonces

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = \frac{-1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2}$$

y por tanto

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\mu,\theta) = \frac{-n}{2\,\theta} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\,\theta^2}$$

Si la distribución a priori es Gamma (p,a), su función generadora es

$$g(\theta) = \frac{p-1}{\theta} - a$$

y la función generadora de la distribución a posteriori

$$g(\theta/\vec{x}) = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\,\theta^2} + \frac{p-1-\frac{n}{2}}{\theta} - a$$

y esta función generadora no corresponde a una distribución gamma.

4.3.2 Familias de distribuciones bivariantes.

De igual modo que en las univariantes, los sistemas de funciones generadoras de las correspondientes distribuciones y las respectivas parciales con respecto a los parámetros de las familias de verosimilitudes, nos pueden facilitar, el reconocimiento rápido de las familias conjugadas. A modo de ejemplo podemos comprobar que

La familia de distribuciones de Dirichlet es conjugada para la verosimilitud trinomial.

En efecto, si el vector de los parámetros (θ_1,θ_2) de la $B(n,\theta_1,\theta_2)$, sigue una distribución $Be(p_1,p_2,p_3)$, el sistema de funciones generadoras de la distribución a priori es

$$g_1(\theta_1,\theta_2) = \frac{p_1-1}{\theta_1} - \frac{p_3-1}{1-\theta_1-\theta_2}$$

$$g_2(\theta_1,\theta_2) = \frac{p_2-1}{\theta_2} - \frac{p_3-1}{1-\theta_1-\theta_2}$$

Las correspondientes parciales de la verosimilitud son:

$$\frac{\partial}{\partial\theta_1} \ln f(x,y/\theta_1,\theta_2) = \frac{x}{\theta_1} - \frac{n-x-y}{1-\theta_1-\theta_2}$$

$$\frac{\partial}{\partial\theta_2}\ln f(x,y/\theta_1,\theta_2) = \frac{y}{\theta_2} - \frac{n-x-y}{1-\theta_1-\theta_2}$$

y por tanto el sistema de funciones generadoras de la distribución a posteriori es

$$g_1(\theta_1,\theta_2,x,y) = \frac{x+p_1-1}{\theta_1} - \frac{n-x-y+p_3-1}{1-\theta_1-\theta_2}$$

$$g_2(\theta_1,\theta_2,x,y) = \frac{y+p_2-1}{\theta_2} - \frac{n-x-y+p_3-1}{1-\theta_1-\theta_2}$$

que como podemos observar se trata de la beta bivariante de parámetros $(x+p_1,\ y+p_2,\ n+p_3-x-y)$

La tabla recoge las distribuciones bivariantes: beta - binomial y beta - binomial negativa.

DISTRIBUCIÓN A PRIORI. Generadoras $g_\theta(\theta,\omega)$ (1) $g_\omega(\theta,\omega)$ (2)	VEROSIMILITUDES. Parcial del logaritmo. $\frac{\partial}{\partial\theta}\ln f(\bar{x}/\theta,\omega)$ (3) $\frac{\partial}{\partial\omega}\ln f(\bar{x}/\theta,\omega)$ (4)	GENERADORAS DE DISTRIBUC. A POSTERIORI. $g_\theta(\theta,\omega/\bar{x}) = (1)+(3)$ $g_\omega(\theta,\omega/\bar{x}) = (2)+(4)$	DISTRIBUC. A POSTERIORI Parámetros
Beta bivariante (p_1,p_2,p_3) $\frac{p_1-1}{\theta} - \frac{p_3-1}{1-\theta-\omega}$ $\frac{p_2-1}{\omega} - \frac{p_3-1}{1-\theta-\omega}$	Binomial bivar. (n,θ,ω) $\frac{x}{\theta} - \frac{n-x-y}{1-\theta-\omega}$ $\frac{y}{\omega} - \frac{n-x-y}{1-\theta-\omega}$	$\frac{x+p_1-1}{\theta} - \frac{n+p_3-x-y-1}{1-\theta-\omega}$ $\frac{x+p_1-1}{\theta_1} - \frac{n+p_3-x-y-1}{1-\theta_1-\theta_2}$	Beta bivariante p_1+x p_2+y $n+p_3-x-y$
Beta bivariante (p_1,p_2,p_3) $\frac{p_1-1}{\theta} - \frac{p_3-1}{1-\theta-\omega}$ $\frac{p_2-1}{\omega} - \frac{p_3-1}{1-\theta-\omega}$	Bin. negativa (k,θ,ω) $\frac{x}{\theta} - \frac{k}{1-\theta-\omega}$ $\frac{y}{\omega} - \frac{k}{1-\theta-\omega}$	$\frac{x+p_1-1}{\theta} - \frac{p_3+k-1}{1-\theta-\omega}$ $\frac{y+p_2-1}{\omega} - \frac{p_3+k-1}{1-\theta-\omega}$	Beta bivariante p_1+x p_2+y p_3+k

4.3.3 Familias de distribuciones multivariantes.

Tiene un desarrollo análogo a los casos ya estudiados. Podemos comprobar que **la familia de distribuciones normal multivariante es conjugada para la verosimilitud de una muestra, procedente de una población normal k-variante.**

En efecto,
La función de densidad de X es:

$$f(X/\theta,\Sigma) = \frac{1}{(2\pi\,|\Sigma|)^{k/2}}\, e^{-\frac{1}{2}(X-\theta)'\Sigma^{-1}(X-\theta)}$$

entonces:

$$\frac{\partial}{\partial\theta}\ln f(X/\theta,\Sigma) = \Sigma^{-1}(X-\theta)$$

y para la verosimilitud de una muestra aleatoria simple de n elementos:

$$\frac{\partial}{\partial\theta}\ln f(\bar{X}/\theta,\Sigma) = n\,\Sigma^{-1}(\bar{X}-\theta)$$

Por otra parte, el sistema de funciones generadoras de la distribución a priori es

$$g(\theta) = -\,A^{-1}\theta + A^{-1}\mu$$

y por tanto el sistema de funciones generadoras de la distribución a posteriori es

$$g(\theta/\bar{X}) = -\,(n\Sigma^{-1} + A^{-1})\,\theta + (n\Sigma^{-1}\bar{x} + A^{-1}\mu)$$

que como podemos observar se trata de la normal multivariante de parámetros

$$\mathrm{Cov}(\theta) = (n\,\Sigma^{-1} + A^{-1})^{-1}$$

$$\mathrm{Media}(\theta) = (n\,\Sigma^{-1} + A^{-1})^{-1}\,(n\Sigma^{-1}\bar{x} + A^{-1}\mu)$$

Queda resumido en la siguiente tabla:

DISTRIBUCIÓN A PRIORI. Generadoras $g(\theta)$ (1)	Normal (μ, A) $-A^{-1}\theta + A^{-1}\mu$
VEROSIMILITUDES. Parcial del logaritmo. $\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta)$ (2)	Normal (θ, Σ) $n\Sigma^{-1}(\overline{X}-\theta)$
GENERADORAS DE DISTRIBUC. A POSTERIORI. $g(\theta/\vec{x}) = (1) + (2)$	$-(n\Sigma^{-1} + A^{-1})\theta + (n\Sigma^{-1}\overline{X} + A^{-1}$
DISTRIBUC. A POSTERIORI Parámetros	Normal $(n\Sigma^{-1} + A^{-1})^{-1}(n\Sigma^{-1}\overline{X} + A^{-1}\mu)$ $(n\Sigma^{-1} + A^{-1})^{-1}$

4.3.4 Verosimilitudes bivariantes y distribuciones a priori univariantes.

Comprobamos que la familia de distribuciones univariantes gamma es conjugada para una familia de verosimilitudes gamma bivariante. Es decir cuando el parámetro «a» de una verosimilitud es desconocido, pero se sabe que sigue una ley gamma univariante.

DISTRIBUCIÓN A PRIORI. Generadoras $g(\theta)$ (1)	VEROSIMILITUDES. Parcial del logaritmo. $\frac{\partial}{\partial\theta} \ln f(\vec{x},\vec{y}/\theta)$ (2)	GENERADORAS DE DISTRIBUC. A POSTERIORI. $g(\theta/\vec{x}) = (1) + (2)$	DISTRIBUC. A POSTERIORI Parámetros
Gamma (p_0, a) $\frac{p_0-1}{\theta} - a$	Gamma (Mckay) (p,q,θ) $\frac{n(p+q)}{\theta} - n\overline{y}$	$\frac{np+nq+p_0-1}{\theta} - (a+n\overline{y})$	Gamma $np+nq+p_0$ $a+n\overline{y}$

4.3.5 Condiciones para que una verosimilitud admita una familia conjugada concreta.

La relación (4.66) también permite conocer las condiciones que debe cumplir la parcial del logaritmo de una verosimilitud para que una familia de distribuciones a priori sea conjugada para dicha verosimilitud.

A continuación aparecen cinco proposiciones: para la familias a priori normal, gamma, beta, de función generadora polinómica y en general para cualquier familia.

La familia normal.

PROPOSICIÓN 4.10

La condición necesaria y suficiente para que una verosimilitud $f(\vec{x}/\theta)$ $(\vec{x} \equiv (x_1, x_2, \ldots, x_n)$ muestra $)$ **admita una familia de distribuciones a priori conjugada que sea normal es que la parcial con respecto al parámetro, del logaritmo de la verosimilitud sea lineal en el parámetro:**

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = C(\vec{x})\,\theta + D(\vec{x}) \,, \qquad \textbf{siendo} \quad C(\vec{x}) < 0.$$

En efecto, si θ tiene una distribución a priori $N(\mu, \sigma^2)$ entonces

$$g(\theta) = \frac{-1}{\sigma^2}\,\theta + \frac{\mu}{\sigma^2}$$

Si la distribución a posteriori, $f(\theta/\vec{x})$, es normal, entonces

$$g(\theta/\vec{x}) = -A(\vec{x})\,\theta + B(\vec{x}) \qquad \text{con } A(\vec{x}) > 0$$

y por tanto

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \left(-A(\vec{x}) + \frac{1}{\sigma^2}\right)\theta + \left(B(\vec{x}) - \frac{\mu}{\sigma^2}\right)$$

es lineal en θ.

Recíprocamente:

$$\text{Si } \theta \to N(\mu, \sigma^2)\,, \text{ entonces } \quad g(\theta) = \frac{-1}{\sigma^2}\,\theta + \frac{\mu}{\sigma^2}$$

Sea $\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta)$ lineal en θ; es decir

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = C(\vec{x})\,\theta + D(\vec{x})$$

en este caso la función generadora de la distribución a posteriori es:

$$g(\theta/\vec{x}) = \frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) + g(\theta) = \left(C(\vec{x}) - \frac{1}{\sigma^2}\right)\theta + \left(D(\vec{x}) + \frac{\mu}{\sigma^2}\right)$$

Si el coeficiente de θ es negativo, es decir si

$$C(\vec{x}) < \frac{1}{\sigma^2} \qquad \forall\, \vec{x},\ \forall\, \sigma^2$$

y por tanto ha de ser $C(\vec{x}) < 0 \quad \forall\, \vec{x}$,
entonces se genera una distribución normal de esperaranza

$E(\theta/\vec{x}) = \dfrac{D(\vec{x})\,\sigma^2 + \mu}{1 - C(\vec{x})\,\sigma^2}$ y varianza $Var(\theta/\vec{x}) = \dfrac{\sigma^2}{1 - C(\vec{x})\,\sigma^2}$

La familia gamma.

PROPOSICIÓN 4.11

La condición necesaria y suficiente para que una verosimilitud $f(\vec{x}/\theta)$ **admita una familia de distribuciones a priori conjugada que sea gamma es que:**

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \frac{C(\vec{x})}{\theta} - D(\vec{x}) \qquad \textbf{(4.67)}$$

siendo $C(\vec{x}) > 0,\ D(\vec{x}) > 0 \quad \forall\vec{x},\ \forall a,\ \forall p.$

En efecto, si la distribución a priori es gamma de parámetros conocidos p, a entonces

$$g(\theta) = \frac{p-1}{\theta} - a$$

Si la distribución a posteriori es gamma, entonces

$$g(\theta/\vec{x}) = \frac{A(\vec{x}) - 1}{\theta} - B(\vec{x}) \qquad \text{con } A(\vec{x}) > 0\ ,\ B(\vec{x}) > 0 \qquad \forall\, \vec{x}$$

y por tanto

$$\frac{\partial}{\partial\theta}\ln f(x/\theta) = \frac{A(\vec{x}) - p}{\theta} - (B(\vec{x}) - a)$$

que es de la forma (4.67).
Recíprocamente, si la distribución a priori es Ga(p,a), entonces

$$g(\theta) = \frac{p-1}{\theta} - a$$

Sea

$$\frac{\partial}{\partial\theta}\ln f(\vec{x}/\theta) = \frac{C(\vec{x})}{\theta} - D(\vec{x})$$

en este caso la función generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \frac{\partial}{\partial\theta}\ln f(\vec{x}/\theta) + g(\theta) = \frac{C(x)}{\theta} - D(x) + \frac{p-1}{\theta} - a$$

$$g(\theta/\vec{x}) = \frac{C(\vec{x}) + p - 1}{\theta} - (D(\vec{x}) + a)$$

que con las condiciones:

$$C(\vec{x}) + p > 0 \qquad \forall\vec{x} \;\; \forall p$$

$$D(\vec{x}) + a > 0 \qquad \forall\vec{x} \;\; \forall a$$

y por tanto

$$C(\vec{x}) > 0 \qquad \forall\vec{x}$$

$$D(\vec{x} > 0 \qquad \forall\vec{x}$$

permite asegurar que se trata de una distribución gamma de parámetros $C(\vec{x}) + p$, $D(\vec{x}) + a$

En particular la verosimilitud de una muestra procedente de una población con distribución de Poisson de parámetro θ desconocido, es:

$$\frac{\partial}{\partial\theta}\ln f(\vec{x}/\theta) = \frac{n\overline{x}}{\theta} - n$$

$$C(\vec{x}) = n\overline{x} > 0$$

$$D(\vec{x}) = n > 0$$

y por tanto la familia gamma es conjugada para esta verosimilitud.

De igual forma se puede particularizar esta proposición para los demás casos que aparecen en la tabla.

La familia beta.

PROPOSICIÓN 4.12

La condición necesaria y suficiente para que una verosimilitud $f(\vec{x}/\theta)$ **admita una familia de distribuciones a priori conjugada que sea beta, sobre un intervalo [a, b], es que**

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \frac{C(\vec{x})}{\theta - a} - \frac{D(\vec{x})}{b - \theta} \qquad \textbf{(4.68)}$$

siendo $C(\vec{x}) > 0,\ D(\vec{x}) > 0 \quad \forall\, \vec{x}$

En efecto, si la distribución a priori es beta de parámetros conocidos p, q entonces

$$g(\theta) = \frac{p - 1}{\theta - a} - \frac{q - 1}{b - \theta}$$

Si la distribución a posteriori es beta, entonces

$$g(\theta/\vec{x}) = \frac{A(\vec{x}) - 1}{\theta - a} - \frac{B(\vec{x}) - 1}{b - \theta} \qquad \text{con } A(\vec{x}) > 0\ ,\ B(\vec{x})\ \ 0 \qquad \forall\, \vec{x}$$

y por tanto

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = \frac{A(\vec{x}) - p}{\theta - a} - \frac{B(\vec{x}) - q}{b - \theta}$$

que es de la forma (4.68).

Recíprocamente, si la distribución a priori es beta de parámetros conocidos p, q,

$$g(\theta) = \frac{p - 1}{\theta - a} - \frac{q - 1}{b - \theta}$$

sea

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \frac{C(\vec{x})}{\theta - a} - \frac{D(\vec{x})}{b - \theta}$$

entonces la función generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \frac{C(x) + p - 1}{\theta - a} - \frac{D(x) + q - 1}{b - \theta}$$

que con las condiciones

$$C(\vec{x}) + p > 0 \qquad \forall \vec{x},\ \forall p$$

$$D(\vec{x}) + q > 0 \qquad \forall \vec{x},\ \forall q$$

y por tanto

$$C(\vec{x}) > 0 \quad \forall \vec{x}$$
$$D(\vec{x}) > 0 \quad \forall \vec{x}$$

se trata de una distribución beta de parámetros $C(\vec{x}) + p$, $D(\vec{x}) + q$ en el intervalo [a,b].

En particular, si la verosimilitud es $B(n,\theta)$, entonces

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \frac{x}{\theta} - \frac{n-x}{1-\theta}$$

$$C(\vec{x}) = x > 0$$
$$D(\vec{x}) = n - x > 0$$

y por tanto la familia beta es conjugada para esta verosimilitud.

De igual forma se puede particularizar esta proposición para los demás casos que aparecen en la tabla.

Ejemplo no recogido en la tabla.

Dada la verosimilitud

$$f(\vec{x}/\theta) = \theta^{\Sigma x_i} (1 - \theta)^{n-\Sigma x_i} , \quad \text{con} \quad \theta \in (0,1)$$

entonces

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \frac{\Sigma x_i}{\theta} - \frac{n - \Sigma x_i}{1 - \theta}$$

Si la función generadora de la distribución a priori es la correspondiente a una beta de parámetros p y q:

$$g(\theta) = \frac{p - 1}{\theta} - \frac{q - 1}{1 - \theta}$$

entonces la generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \frac{\Sigma x_i + p - 1}{\theta} + \frac{n - \Sigma x_i + q - 1}{1 - \theta}$$

que con las condiciones

$$\Sigma x_i + p > 0 , \quad n - \Sigma x_i + q > 0 \quad \forall \vec{x}$$

corresponde a la distribución beta de parámetros

$$\left(\Sigma x_i + p , \quad n - \Sigma x_i + q , 0 , 1 \right)$$

La familia de distribuciones con función generadora polinómica.

PROPOSICIÓN 4.13

La condición necesaria y suficiente para que una verosimilitud $f(\vec{x}/\theta)$ **admita una familia de distribuciones a priori conjugada cuya generadora sea una función polinómica, de grado menor o igual que n, es que:**

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \sum_{i=0}^{n} A_i(\vec{x})\, \theta^i \qquad (4.69)$$

En efecto, sea $$g(\theta) = \sum_{i=0}^{n} a_i\, \theta^i$$

la función generadora de la distribución a priori. Si la función generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \sum_{i=0}^{n} a_i(\vec{x})\, \theta^i$$

entonces $$\frac{\partial}{\partial\theta} \ln f(x/\theta) = \sum_{i=0}^{n} [\, a_i(\vec{x}) - a_i\,]\, \theta^i$$

que es de la forma (4.69).

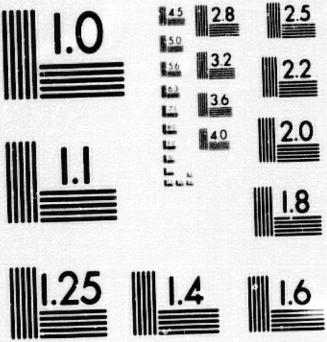
Recíprocamente, si $$g(\theta) = \sum_{i=0}^{n} a_i\, \theta^i$$

y si $$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \sum_{i=0}^{n} A_i(\vec{x})\, \theta^i$$

entonces la función generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \sum_{i=0}^{n} [\, A_i(\vec{x}) + a_i\,]\, \theta^i$$

polinómica de grado menor o igual que n, (si $A_n(\vec{x}) = -a_n$ el grado de la función generadora es menor que n).



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

La familia de distribuciones con función generadora polinómica.

PROPOSICIÓN 4.13

La condición necesaria y suficiente para que una verosimilitud $f(\vec{x}/\theta)$ **admita una familia de distribuciones a priori conjugada cuya generadora sea una función polinómica, de grado menor o igual que n, es que:**

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \sum_{i=0}^{n} A_i(\vec{x})\, \theta^i \qquad (4.69)$$

En efecto, sea $\quad g(\theta) = \sum_{i=0}^{n} a_i\, \theta^i$

la función generadora de la distribución a priori. Si la función generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \sum_{i=0}^{n} a_i(\vec{x})\, \theta^i$$

entonces $\quad \frac{\partial}{\partial\theta} \ln f(x/\theta) = \sum_{i=0}^{n} [\, a_i(\vec{x}) - a_i\,]\, \theta^i$

que es de la forma (4.69).

Recíprocamente, si $\quad g(\theta) = \sum_{i=0}^{n} a_i\, \theta^i$

y si $\quad \frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = \sum_{i=0}^{n} A_i(\vec{x})\, \theta^i$

entonces la función generadora de la distribución a posteriori es

$$g(\theta/\vec{x}) = \sum_{i=0}^{n} [\, A_i(\vec{x}) + a_i\,]\, \theta^i$$

polinómica de grado menor o igual que n, (si $A_n(\vec{x}) = -\, a_n$ el grado de la función generadora es menor que n).

La conjugación en general.

PROPOSICIÓN 4.14

Una familia de distribuciones, dada por sus generadoras,

$$F \equiv \{g(\theta/\vec{\eta})\}_{\vec{\eta}\in\Omega\subset\mathbb{R}^k}$$

es conjugada para la familia de verosimilitudes $f(\vec{x}/\theta)$ **si y sólo si, existe una función**

$$\vec{\varphi} : \mathbb{R}^n \times \Omega \dashrightarrow \Omega$$

$$(\vec{x},\vec{\eta}) \dashrightarrow \vec{\varphi}(\vec{x},\vec{\eta})$$

tal que

$$\frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) = g(\theta/\vec{\varphi}(\vec{x},\vec{\eta})) - g(\theta/\vec{\eta}) \qquad \textbf{(4.70)}$$

En efecto, si la igualdad (4.70) es cierta, entonces $g(\theta/\vec{\eta})$ es la función generadora de la distribución a priori y $g(\theta/\vec{\varphi}(\vec{x},\vec{\eta}))$ de la distribución a posteriori que pertenece a la familia, por tanto la familia es conjugada para la verosimilitud considerada.

Recíprocamente, si la familia es conjugada para la verosimilitud, entonces dada $\vec{x}$ y dada una distribución a priori de la familia, $g(\theta/\eta_1)$ $\forall\, \eta_1 \in \Omega$, existe otra única distribución de la familia que hace el papel de distribución a posteriori que representamos por $g(\theta/\eta_2)$.

Ahora bien si el par

$$(\, g(\theta/\eta_1) \ ,\ g(\theta/\eta_2)\,)$$

son priori-posteriori, para una $\vec{x}$ dada entonces

$$g(\theta/\eta_2) = \frac{\partial}{\partial\theta} \ln f(\vec{x}/\theta) + g(\theta/\eta_1) \qquad \forall\, \theta$$

por tanto η_2 es una función que depende de $\vec{x}$ y de η_1,

$$\vec{\varphi} : \mathbb{R}^n \times \Omega \to \Omega$$

$$(\vec{x},\eta_1) \to \eta_2 = \vec{\varphi}(\vec{x},\eta_1)$$

que evidentemente, por construcción verifica (4.70).

EJEMPLO.

Es conocido que la familia $N(\mu,\eta^2)$ es conjugada para la familia de verosimilitudes $N(\theta,1)$.

Se verifica

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = x - \theta$$

Dada una distribución a priori

$$g(\theta/\mu_1,\eta_1^2) = -\frac{\theta - \mu_1}{\eta_1^2}$$

se sabe que la distribución a posteriori es otra

$$g(\theta/\mu_2,\eta_2^2) = -\frac{\theta - \mu_2}{\eta_2^2}$$

y puesto que

$$g(\theta/\mu_2,\eta_2^2) = \frac{\partial}{\partial\theta} \ln f(x/\theta) + g(\theta/\mu_1,\eta_1^2)$$

sustituyendo:

$$-\frac{\theta - \mu_2}{\eta_2^2} = x - \theta - \frac{\theta - \mu_1}{\eta_1^2}$$

$$-\frac{\theta - \mu_2}{\eta_2^2} = \frac{-(\eta_1^2 + 1)\theta + \mu_1 + x\eta_1^2}{\eta_1^2}$$

$$-\frac{1}{\eta_2^2}\theta + \frac{\mu_2}{\eta_2^2} = -\frac{1}{\frac{\eta_1^2}{\eta_1^2+1}}\theta + \frac{\mu_1 + x\eta_1^2}{\eta_1^2}$$

luego

$$\eta_2^2 = \frac{\eta_1^2}{1 + \eta_1^2}$$

$$\mu_2 = \frac{\mu_1 + x\eta_1^2}{\eta_1^2}\eta_2^2 = \frac{\mu_1 + x\eta_1^2}{1 + \eta_1^2}$$

4.4 SISTEMAS DE FUNCIONES GENERADORAS, MATRIZ DE INFORMACIÓN Y MÉTODO DE MÁXIMA VEROSIMILITUD, BAJO CONDICIONES DE REGULARIDAD.

Entre las aplicaciones que se pueden encontrar en la Estadística Matemática, ciertos conceptos habituales como son la cantidad de Información de Fisher, la matriz de Información, el Método de la Máxima Verosimilitud, entre otros, pueden ser interpretados desde el punto de vista de las funciones generadoras. En todo el desarrollo se supone que se verifican determinadas condiciones de regularidad, que a continuación se detallan.

Dada una función de densidad condicionada, $f(x/y)$, diremos que se cumplen las condiciones de regularidad si y sólo si:

$$\frac{\partial}{\partial y}\int_{-\infty}^{+\infty} f(x/y)\,dx = \int_{-\infty}^{+\infty} \frac{\partial}{\partial y} f(x/y)\,dx \qquad \forall y$$

$$\frac{\partial^2}{\partial y^2}\int_{-\infty}^{+\infty} f(x/y)\,dx = \int_{-\infty}^{+\infty} \frac{\partial^2}{\partial y^2} f(x/y)\,dx \qquad \forall y$$

Estas condiciones son consecuencia inmediata de las establecidas en Apostol (1972), [2]. Se han acomodado a las notaciones seguidas en la presente memoria.

4.4.1 La cantidad de información como varianza de la función generadora $g_y(x,y)$.

De Groot ([11], pág. 403), define la cantidad de información de Fisher como:

$$I(y) = \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial y} \ln f(x/y)\right)^2 f(x/y)\,dx$$

demuestra que también se puede expresar como

$$I(y) = -\int_{-\infty}^{+\infty} \left(\frac{\partial^2}{\partial y^2} \ln f(x/y)\right) f(x/y)\,dx$$

y que coincide con la varianza

$$I(y) = Var_{f(x/y)}\left(\frac{\partial}{\partial y} \ln f(x/y)\right)$$

El siguiente teorema expresa la cantidad de información en términos de las funciones generadoras.

TEOREMA 4.22

Dada la verosimilitud f(x/y) y para todo y, la función cantidad de información de Fisher coincide con la varianza de la función generadora $g_y(x,y)$ **es decir:**

$$I(y) = Var(g_y(x,y))$$

En efecto:

$$I(y) = \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial y} \ln f(x/y)\right)^2 f(x/y)\, dx =$$

aplicando la proposición 2.3 se escribe en términos de g

$$= \int_{-\infty}^{+\infty} (g_y(x,y) - g(y))^2 f(x/y)\, dx$$

y puesto que por la proposición 2.4, el valor esperado, con respecto a la distribución $f(x/y)$, de $g_y(x,y)$ coincide con $g(y)$, se puede afirmar que;

$$\int_{-\infty}^{+\infty} (g_y(x,y) - g(y))^2 f(x/y)\, dx = Var[g_y(x,y)]$$

El hecho de que la cantidad de información de Fisher coincida con la varianza de la generadora $g_y(x,y)$, nos indica que, para cada y, la información será mayor cuanto mayor dispersión exista entre las funciones generadoras, $g_y(x,y)$.

Obsérvese que en el caso de independencia entre las variables X e Y, las funciones generadoras $g_y(x,y)$ y $g(y)$ coinciden, la Var ($g_y(x,y)$) es cero y por tanto la información que se tiene es nula.

EJEMPLO

Sea $$f(x/y) = \frac{1}{y} e^{-\frac{x}{y}} \qquad x>0\ , \quad \forall\, y>0$$

entonces $$\ln f(x/y) = -\ln y - \frac{x}{y}$$

$$\frac{\partial}{\partial y} \ln f(x/y) = -\frac{1}{y} + \frac{x}{y^2}$$

Por tanto, de acuerdo con la proposición 2.3, se puede adoptar como funciones generadoras canónicas el par:

$$g(y) = \frac{1}{y}$$

$$g_y(x,y) = \frac{x}{y^2}$$

Cualquier par de generadoras admisibles toma la forma:

$$g(y) = \frac{1}{y} + C(y)$$

$$g_y(x,y) = \frac{x}{y^2} + C(y)$$

En cualquier caso:

$$g_y(x,y) - g(y) = \frac{x}{y^2} - \frac{1}{y} = \frac{\partial}{\partial y} \ln f(x/y)$$

es un invariante, es decir, no cambia de forma mientras no lo hace $f(x/y)$. Será pues:

$$E[g_y(x,y) - g(y)] = \int_0^{+\infty} \frac{x}{y^2} \frac{1}{y} e^{-\frac{x}{y}} dx - \int_0^{+\infty} \frac{1}{y} \frac{1}{y} e^{-\frac{x}{y}} dx =$$

$$= \frac{1}{y^2} \int_0^{+\infty} x \frac{1}{y} e^{-\frac{x}{y}} dx - \frac{1}{y} \int_0^{+\infty} \frac{1}{y} e^{-\frac{x}{y}} dx = \frac{1}{y^2} y - \frac{1}{y} 1 = 0$$

Calculando ahora la parcial segunda con respecto a y

$$\frac{\partial^2}{\partial y^2} \ln f(x/y) = \frac{\partial}{\partial y}\left(-\frac{1}{y} + \frac{x}{y^2}\right) = \frac{1}{y^2} - \frac{2x}{y^3}$$

Por tanto la función de información es

$$I(y) = -\int_0^{+\infty}\left(\frac{\partial^2}{\partial y^2}\ln f(x/y)\right) f(x/y)\,dx =$$

$$= \int_0^{+\infty}\left(\frac{2x}{y^3} - \frac{1}{y^2}\right)^2 \frac{1}{y}\, e^{-\frac{x}{y}}\,dx =$$

$$= \frac{2}{y^3}\int_0^{+\infty}\frac{x}{y}\, e^{-\frac{x}{y}}\,dx - \frac{1}{y^2}\int_0^{+\infty}\frac{1}{y}\, e^{-\frac{x}{y}}\,dx = \frac{2y}{y^3} - \frac{1}{y^2} = \frac{1}{y^2}$$

También I(y) se puede obtener como

$$\int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial y}\ln f(x/y)\right)^2 f(x/y)\,dx = \int_0^{+\infty}\left(-\frac{1}{y} + \frac{x}{y^2}\right)^2 \frac{1}{y}\, e^{-\frac{x}{y}}\,dx =$$

$$\frac{1}{y^4}\int_0^{+\infty}(x-y)^2\,\frac{1}{y}\, e^{-\frac{x}{y}}\,dx = \frac{1}{y^4}\, y^2 = \frac{1}{y^2}$$

Por otra parte, para todo y:

$$\mathrm{Var}(g(y/x)) = \int_0^{+\infty}\left[g(y/x) - \int_0^{+\infty} g(y/x)\, f(x/y)\,dx\right]^2 f(x/y)\,dx =$$

$$= \int_0^{+\infty}[g(y/x) - g(y)]^2 f(x/y)\,dx = \int_0^{+\infty}\left(\frac{x}{y^2} - \frac{1}{y}\right)^2 \frac{1}{y} e^{-\frac{x}{y}}\,dx = \frac{1}{y^2}$$

Estimación de la varianza a través de las observaciones muestrales.

En la igualdad $E[g_y(x,y)] = g(y)$ se sustituye el valor esperado por su estimador, media muestral, $\frac{1}{n}\, g_y(x,y)$, entonces la ecuación $\frac{1}{n}\, g_y(x,y) = g(y)$ se identifica con el método de la máxima verosimilitud.

La distribución asintótica del estimador máximo-verosímil θ es normal de media el parámetro y varianza $\frac{1}{n\,I(y)}$.

Puesto que $I(y)$ es una varianza, cabe estimarla por su varianza muestral

$$\hat{I}(y) = \frac{1}{n} \sum_{i=1}^{n} (g_y(x_i, y))^2 - \left(\frac{1}{n} \sum_{i=1}^{n} g_y(x_i, y) \right)^2$$

Este método, como se estudia a continuación, no se aleja de "la filosofía del método de la máxima verosimilitud", y permite estimar la cantidad de información conociendo el sistema de funciones generadoras.

4.4.2 La cantidad de información, matriz de covarianzas de las funciones generadoras $g_y(x,y)$.

En el caso de que la distribución de probabilidad de la variable aleatoria $(\vec{X}/\vec{Y}) = (X_1, \ldots, X_n/Y_1, \ldots, Y_k)$, se define la matriz de información de Fisher

$$\left(I_{i,j}(\vec{y}) \right)_{i=1,\ldots,k \;\; j=1,\ldots,k}$$

siendo

$$I_{i,j}(\vec{y}) = \int_{\chi_x} \left(\frac{\partial}{\partial y_i} \ln f(\vec{x}/\vec{y}) \right) \left(\frac{\partial}{\partial y_j} \ln f(\vec{x}/\vec{y}) \right) f(\vec{x}/\vec{y})\, d\vec{x}$$

y χ_x el recinto de definición del vector aleatorio $\vec{X}$, (Ibragimov, [24]), que coincide con la covarianza:

$$I_{i,j}(\vec{y}) = \mathrm{Cov}\left(\frac{\partial}{\partial y_i} \ln f(\vec{x}/\vec{y}),\ \frac{\partial}{\partial y_j} \ln f(\vec{x}/\vec{y}) \right)$$

TEOREMA 4.23

Dada la verosimilitud $f(\vec{x}/\vec{y})$ **y para todo** $\vec{y}$ **, la matriz de información de Fisher coincide con la matriz de covarianzas de las funciones generadoras** $g_{y_i}(\vec{x},\vec{y})$ con $i=1,\ldots,k$, **es decir:**

$$I_{ij}(\vec{y}) = \mathrm{Cov}[\, g_{y_i}(\vec{x},\vec{y}) \,,\, g_{y_j}(\vec{x},\vec{y}) \,] \quad \forall\, i,j=1,\ldots,k$$

En efecto:

$$I_{i,j}(\vec{y}) = \int_{\chi_x} \left(\frac{\partial}{\partial y_i} \ln f(\vec{x}/\vec{y}) \right) \left(\frac{\partial}{\partial y_j} \ln f(\vec{x}/\vec{y}) \right) f(\vec{x}/\vec{y})\, d\vec{x} =$$

que aplicando la igualdad (3.2), de la proposición 3.2, se escribe en términos de las generadoras

$$= \int_{\chi_x} [\, g_{y_i}(\vec{x},\vec{y}) - g_{y_i}(\vec{y}) \,]\, [\, g_{y_j}(\vec{x},\vec{y}) - g_{y_j}(\vec{y}) \,]\, f(\vec{x}/\vec{y})\, d\vec{x} =$$

y puesto que la igualdad (3.4) de la proposición 3.3, dice

$$E_{f(\vec{x}/\vec{y})}[\, g_{y_i}(\vec{x},\vec{y}) \,] = g_{y_i}(\vec{y}) \quad \forall\, i=1,2,\ldots,k$$

se puede afirmar que

$$I_{ij}(\vec{y}) = \mathrm{Cov}[\, g_{y_i}(\vec{x},\vec{y}) \,,\, g_{y_j}(\vec{x},\vec{y}) \,] \quad \forall\, i,j=1,\ldots,k$$

Sobre la utilización de este resultado en la inferencia se pueden hacer idénticos comentarios que en el epígrafe 4.4.1

TEOREMA 4.23

Dada la verosimilitud $f(\vec{x}/\vec{y})$ **y para todo** $\vec{y}$ **, la matriz de información de Fisher coincide con la matriz de covarianzas de las funciones generadoras** $g_{y_i}(\vec{x},\vec{y})$ con $i=1,...,k$, **es decir:**

$$I_{ij}(\vec{y}) = \mathrm{Cov}[\, g_{y_i}(\vec{x},\vec{y}) \,,\, g_{y_j}(\vec{x},\vec{y})\,] \qquad \forall\, i,j=1,..,k$$

En efecto:

$$I_{i,j}(\vec{y}) = \int_{\mathcal{X}_x} \left(\frac{\partial}{\partial y_i} \ln f(\vec{x}/\vec{y})\right)\left(\frac{\partial}{\partial y_j} \ln f(\vec{x}/\vec{y})\right) f(\vec{x}/\vec{y})\, d\vec{x} =$$

que aplicando la igualdad (3.2), de la proposición 3.2, se escribe en términos de las generadoras

$$= \int_{\mathcal{X}_x} [\, g_{y_i}(\vec{x},\vec{y}) - g_{y_i}(\vec{y})\,]\,[\, g_{y_j}(\vec{x},\vec{y}) - g_{y_j}(\vec{y})\,]\, f(\vec{x}/\vec{y})\, d\vec{x} =$$

y puesto que la igualdad (3.4) de la proposición 3.3, dice

$$E_{f(\vec{x}/\vec{y})}[\, g_{y_i}(\vec{x},\vec{y})\,] = g_{y_i}(\vec{y}) \qquad \forall\, i=1,2,...,k$$

se puede afirmar que

$$I_{ij}(\vec{y}) = \mathrm{Cov}[\, g_{y_i}(\vec{x},\vec{y}) \,,\, g_{y_j}(\vec{x},\vec{y})\,] \qquad \forall\, i,j=1,..,k$$

Sobre la utilización de este resultado en la inferencia se pueden hacer idénticos comentarios que en el epígrafe 4.4.1

4.5 SISTEMAS DE FUNCIONES GENERADORAS EN LOS MÉTODOS DE ESTIMACIÓN MÁXIMO VEROSÍMIL Y DE LOS MOMENTOS.

En todos y cada uno de los estudios sobre la familia univariante continua de Pearson, en sus extensiones, Roy [42], y en sus generalizaciones Herrerías [19], así como en la familia univariante discreta Ord [34] y su extensión Herrerías [20], se establece una relación entre los parámetros de las ecuaciones que las definen y los momentos muestrales, estimándolos utilizando el método de los momentos.

Este método coincide, bajo determinadas condiciones, con el método de máxima verosimilitud, (Kendall & Stuart, [28]; Callejón - Santos, [5]).

En este epígrafe demostramos que el método de la máxima verosimilitud es equivalente a un *método de momentos generalizado*, en el sentido de que los estimadores son las soluciones del sistema de ecuaciones que se obtiene al igualar en cada ecuación el valor esperado de una determinada función, con la media muestral correspondiente.

4.5.1 Comparación de los dos métodos de estimación.

Utilizando las funciones generadoras, en el lema, teorema y corolario siguientes se hace una comparación de estos métodos de estimación puntual.

LEMA 4.1

Dada una función generadora: $g_x(x,\theta_0,\ldots,\theta_r)$ **, se verifica**

$$\frac{\partial}{\partial\theta_i} \ln K(\theta_0,\ldots,\theta_r) = -E\left[\frac{\partial}{\partial\theta_i}\int^x g_x(x,\theta_0,\ldots,\theta_r)\,dx\right]$$

siendo $K(\theta_0,\ldots,\theta_r)$ **la constante de normalización.**

En efecto, llamando $\vec{\theta} = (\theta_0,..,\theta_r)$, por definición es

$$K(\vec{\theta}) = \left(\int_{-\infty}^{+\infty} e^{\int^x g_x(x,\vec{\theta})\,dx}\, dx \right)^{-1}$$

luego $$\ln K(\vec{\theta}) = -\ln \int_{-\infty}^{+\infty} e^{\int^x g_x(x,\vec{\theta})\,dx}\, dx$$

y en consecuencia,

$$\frac{\partial}{\partial\theta_i} \ln K(\vec{\theta}) = - \frac{\int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial\theta_i} \int^x g_x(x,\vec{\theta})\, dx \right) e^{\int^x g_x(x,\vec{\theta})\,dx}\, dx}{\int_{-\infty}^{+\infty} e^{\int^x g_x(x,\vec{\theta})\,dx}\, dx}$$

teniendo en cuenta que la función de densidad generada por $g_x(x,\vec{\theta})$ es

$$f(x/\vec{\theta}) = K(\vec{\theta})\, e^{\int^x g_x(x,\vec{\theta})\,dx}$$

es claro que

$$\frac{\partial}{\partial\theta_i} \ln K(\vec{\theta}) = - \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial\theta_i} \int^x g_x(x,\vec{\theta})\, dx \right) f(x/\vec{\theta})\, dx =$$

$$= - E\left[\frac{\partial}{\partial\theta_i} \int^x g_x(x,\vec{\theta})\, dx \right]$$

Notaciones:

Para simplificar la notación, llamaremos

$$d_i(x;\vec{\theta}) = \frac{\partial}{\partial\theta_i} \int^x g_x(x,\vec{\theta})\, dx \qquad (4.71)$$

Dada una muestra aleatoria simple, de tamaño n, $X_1,...,X_n$, notaremos

$$\overline{d}_i(\vec{\theta}) = \frac{1}{n} \sum_{j=1}^{n} d_i(x_j;\vec{\theta}) \qquad (4.72)$$

TEOREMA 4.24

Si existe un estimador $\hat{\theta}$ de máxima verosimilitud, este es una de las posibles soluciones del sistema de ecuaciones:

$$E_x\,[d_i(x;\vec{\theta})] = \bar{d}_i(\vec{\theta}) \quad \text{con} \quad i=0,1,\ldots,r$$

En efecto, el elemento de verosimilitud es

$$f(x_1,\ldots,x_n/\vec{\theta}) = K^n(\vec{\theta})\, e^{\sum_{j=1}^{n}\int^{x_j} g_{x_j}(x_j,\vec{\theta})\,dx_j}$$

entonces

$$\ln f(x_1,\ldots,x_n/\vec{\theta}) = n \ln K(\vec{\theta}) + \sum_{j=1}^{n}\int^{x_j} g_{x_j}(x_j,\vec{\theta})\,dx_j$$

por tanto, al derivar con respecto a cada uno de los r+1 parámetros se puede escribir el sistema:

$$\frac{\partial}{\partial\theta_i}\ln f(x_1,\ldots,x_n/\vec{\theta}) = n\frac{\partial}{\partial\theta_i}\ln K(\vec{\theta}) + \sum_{j=1}^{n}\frac{\partial}{\partial\theta_i}\int^{x_j} g_{x_j}(x_j,\vec{\theta})\,dx_j$$

y como consecuencia, en virtud del lema 4.1 y utilizando las notaciones (4.71) y (4.72), se obtiene

$$\frac{\partial}{\partial\theta_i}\ln f(x_1,\ldots,x_n/\vec{\theta}) = -\,n\,E[d_i(x;\vec{\theta})] + n\,\bar{d}_i(\vec{\theta})$$

de donde se deduce que los dos sistemas de ecuaciones siguientes son idénticos:

Primer sistema

$$\frac{\partial}{\partial\theta_i}\ln f(x_1,\ldots,x_n/\vec{\theta}) = 0 \qquad i=0,1,\ldots,r \qquad (4.73)$$

Segundo sistema

$$E[d_i(x;\vec{\theta})] = \bar{d}_i(\vec{\theta}) \qquad i=0,1,\ldots,r \qquad (4.74)$$

Se estiman unas medias poblacionales mediante sus respectivas medias muestrales, y se despejan los valores de los parámetros correspondientes. Se trata pues, de un método de momentos generalizado.

COROLARIO 4.7

En el caso particular de que la función generadora sea polinómica, $g_x(x,\vec{\theta}) = \sum_{i=0}^{r} \theta_i x^i,$ [1] **entonces el sistema (4.74) se puede escribir:**

$$M_i(\theta) = m_i \qquad i=1,\dots,r+1$$

siendo

$$M_i(\theta) = E[x^i]$$

$$m_i = \frac{1}{n}\sum_{j=1}^{n} x_j^i$$

En efecto, calcularemos por separado cada uno de los miembros del segundo sistema.

Para calcular el primer miembro integramos en x

$$\int^x g_x(x,\vec{\theta})\,dx = \sum_{i=0}^{r} \int^x \theta_i x^i\,dx = \sum_{i=0}^{r} \theta_i \frac{x^{i+1}}{i+1}$$

derivamos ahora con respecto a θ_i

$$d_i(x;\vec{\theta}) = \frac{\partial}{\partial\theta_i}\int g_x(x,\vec{\theta})\,dx = \frac{x^{i+1}}{i+1}$$

y su valor esperado es

$$E[d_i(x;\vec{\theta})] = \frac{M_{i+1}(\vec{\theta})}{i+1}$$

El segundo miembro queda

$$\bar{d}_i(\vec{\theta}) = \frac{1}{n}\sum_{j=1}^{n} d_i(x_j;\vec{\theta}) = \frac{1}{n}\sum_{j=1}^{n}\frac{x_j^{i+1}}{i+1} = \frac{m_{i+1}}{i+1}$$

---

(1) Obsérvese que si la función generadora es una función polinómica, entonces la densidad viene dada por

$$f(x;\vec{\theta}) = K(\vec{\theta})\, e^{\sum_{i=0}^{r} \frac{\theta_i}{i+1} x^{i+1}}$$

y por tanto el sistema (4.74) es

$$M_{i+1}(\theta) = m_{i+1} \qquad i=0,\ldots,r$$

o lo que es equivalente

$$M_i(\theta) = m_i \qquad i=1,\ldots,r+1$$

Supuestas las condiciones necesarias de existencia de los sistemas y de unicidad de las soluciones, para las distribuciones en las que el segundo miembro de la ecuación diferencial correspondiente a la definición de la familia univariante de Pearson o cualquiera de sus extensiones sea una función polinómica, los métodos de estimación de los parámetros de máxima verosimilitud y de los momentos coinciden.

Es importante que el resultado de este corolario coincida con el enunciado dado por Kendall & Stuart [28], puesto que se han utilizado métodos distintos, siendo aquel un caso particular del que en esta memoria se ha desarrollado.

4.5.2 El método de la máxima verosimilitud y su interpretación desde el sistema de funciones generadoras.

Supongamos $f(x/\theta)$, una distribución condicionada. El estimador de máxima verosimilitud sobre una observación X, es la solución de la ecuación

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = 0 \qquad (4.75)$$

Considerando la distribución conjunta $f(x,\theta)$ (1), de la que $f(x/\theta)$ es una condicionada, por la proposición 2.3 se verifica

$$g_\theta(x,\theta) = g(\theta) + \frac{\partial}{\partial\theta} \ln f(x/\theta) \qquad (4.76)$$

la ecuación (4.75) es:

$$\frac{\partial}{\partial\theta} \ln f(x/\theta) = -g(\theta) + g_\theta(x,\theta) = 0$$

(1) Véase la interpretación de los parámetros, (capítulo 3).

es decir el estimador máximo verosímil es el valor del parámetro donde se produce la intersección de la función generadora a priori y la función generadora a posteriori, cualquiera que sea la distribución a priori que propongamos, ya que la relación (4.76) es válida para cualquier distribución.

Todo lo expresado anteriormente sigue siendo válido para una muestra de tamaño n

$$g_\theta(x_1..x_n, \theta) = g(\theta) + \frac{\partial}{\partial\theta} \ln f(x_1,...,x_n/\theta)$$

por tanto la expresión

$$\frac{\partial}{\partial\theta} \ln f(x_1,...,x_n/\theta) = -g(\theta) + g_\theta(x_1...x_n, \theta) = 0$$

proporciona el estimador de máxima verosimilitud.

En el caso de varios parámetros, se parte de la verosimilitud $f(\vec{x}/\vec{\theta})$ y la distribución a priori $f(\vec{\theta})$, la distribución conjunta es: $f(\vec{x},\vec{\theta}) = f(\vec{x}/\vec{\theta})\, f(\vec{\theta})$
por lo que según la proposición 3.2 se verifica

$$g_{\theta_j}(\vec{x},\vec{\theta}) = g_{\theta_j}(\vec{\theta}) + \frac{\partial}{\partial\theta_j} \ln f(\vec{x}/\vec{\theta})$$

y el sistema de ecuaciones que plantea el método de la máxima verosimilitud es:

$$\frac{\partial}{\partial\theta_j} \ln f(\vec{x}/\vec{\theta}) = 0 \qquad (j=1,...,k)$$

Escrito en términos de funciones generadoras:

$$g_{\theta_j}(\vec{x},\vec{\theta}) - g_{\theta_j}(\vec{\theta}) = 0 \qquad (j=1,...,k) \qquad (4.77)$$

Si la distribución a priori está totalmente factorizada (variables aleatorias independientes) entonces la relación (4.77) queda:

$$g_{\theta_j}(\vec{x},\vec{\theta}) - g(\theta_j) = 0 \qquad (j=1,...,k)$$

4.6 LA FUNCIÓN GENERADORA EN LAS CURVAS DE LORENZ.

Una aplicación directa de la función generadora, introducida en el primer capítulo de esta Memoria, que entra dentro del campo económico, es la de utilizarla como modelizadora de las curvas de Lorenz

A partir de la distribución de probabilidad de una variable aleatoria finita no negativa y de tipo continuo, X, la curva de Lorenz correspondiente viene dada por

$$L(p) = \frac{\int_0^{F^{-1}(p)} t\, dF(t)}{E(x)}, \qquad 0 \leq p \leq 1 \tag{4.78}$$

siendo F(x) la función de distribución y E(x) el valor esperado de dicha variable, (Casas - Núñez, [8]).

Casas, Herrerías y Núñez, [7], presentan dos formas de obtención de funciones que modelizan la curva de Lorenz, una mediante combinaciones lineales convexas de formas funcionales utilizadas en la estimación de dicha curva por otros autores y otra, utilizando la ecuación diferencial que genera la familia de distribuciones continuas univariantes de Pearson. Se estudian la formas potencial y exponencial, ambas satisfacen la ecuación diferencial de Pearson, y otra forma funcional que no satisface la mencionada ecuación.

Puesto que la ecuación diferencial que define la familia de Pearson

$$\frac{f'(x)}{f(x)} = \frac{x - a}{b_0 + b_1 x + b_2 x^2} \tag{4.79}$$

dio la idea de generalizar el segundo miembro, obteniendo de esta forma el concepto de función generadora de una distribución continua univariante de probabilidad, cabe suponer que una función

$$g(x) = \frac{f'(x)}{f(x)} \tag{4.80}$$

que cumpla determinados requisitos, dará lugar a modelizaciones de la curva de Lorenz, tanto si g(x) es racional, del tipo exigido en (4.79), como si es de cualquier otro tipo.

Casas-Herrerías y Núñez [7] inician el estudio de la relación entre la familia de Pearson (4.79), ecuación diferencial (4.80), y la modelización de la curva de Lorenz

Lafuente en su tesis doctoral, [29], estudia ocho casos concretos de la función g(x) que, agrupados según su forma más general, se citan a continuación. Todos ellos verifican las condiciones para la generación de la curva de Lorenz. Lafuente obtiene las correspondientes soluciones utilizando la ecuación diferencial (4.80). No todas las funciones utilizadas verifican la ecuación que define la familia de Pearson.

1) $g(x) = ax + b\ln x + c + \frac{d}{x}, \quad a, c \geq 0; \; b \leq 0; \; d > 0$

definida $\forall x > 0$
y sus casos particulares

$$g(x) = \frac{d}{x}, \quad d > 0$$

$$g(x) = c + \frac{1}{x}, \quad c > 0$$

$$g(x) = \ln A + \frac{1}{x}, \quad A > 1$$

$$g(x) = 2x + \frac{d}{x}, \quad d \geq 0$$

2) $g(x) = \frac{c}{x} + \frac{a(1-x)^{a-1}}{b(1-(1-x)^a)}, \quad \begin{matrix} 0 \leq a \leq 1 \\ 0 < b \leq 1 \\ c \geq 0 \end{matrix} \quad$ definida $\forall x > 0$

(a y c no son cero a la vez),
y sus casos particulares

$$g(x) = \frac{a(1-x)^{a-1}}{b(1-(1-x)^a)}, \qquad \begin{matrix} 0 \le a \le 1 \\ 0 \le b \le 1 \end{matrix}$$

$$g(x) = \frac{c}{x} + \frac{a(1-x)^{a-1}}{1-(1-x)^a}, \qquad \begin{matrix} 0 \le a \le 1 \\ c \ge 0 \end{matrix}$$

3) $$g(x) = \frac{k\, e^{kx}}{e^{kx} - 1}, \qquad k > 0 \quad \text{definida } \forall x > 0$$

Si una modelización de la curva de Lorenz, $f(x)$ verifica (4.80) entonces ha de ser de la forma

$$f(x) = C\, e^{\int^x g(x)\, dx}$$

sujeta a las condiciones:

a) $\lim_{x \to 0} f(x) = 0$

b) $f(1) = 1$ que permitirá determinar la constante C

c) f creciente en $[0,1]$ es decir $f'(x) > 0 \quad \forall x \in [0,1]$

d) f cóncava en $[0,1]$ es decir $f''(x) > 0 \quad \forall x \in [0,1]$

La siguiente proposición estudia las condiciones necesarias y suficientes que ha de cumplir la función g para que la solución obtenida a de la ecuación (4.80) verifique las condiciones de contorno especificadas: a), b), c) y d).

4.6.1 Condiciones que debe cumplir una función g(x) para que sea función generadora de una curva de Lorenz.

PROPOSICIÓN 4.15

Sea g(x) una función real de variable real.

Llamamos $$G(x) = \int^x g(x)\, dx$$

Para que $$f(x) = e^{G(x) - G(1)} \qquad \textbf{(4.81)}$$

sea una curva de Lorenz es condición necesaria y suficiente que se verifiquen las relaciones:

1) $\lim_{x \to 0} G(x) = -\infty$

2) $g(x) > 0 \quad \forall\, x \in [0,1]$

3) $(g(x))^2 + g'(x) \geq 0 \quad \forall\, x \in [0,1]$

En efecto, si $f(x)$ es una función que modeliza la curva de Lorenz, entonces es posible definir

$$g(x) = \frac{d}{dx} \ln f(x) = G'(x)$$

de donde se deduce que:

a) $f(0) = 0$ implica que $\lim_{x \to 0} G(x) = -\infty$

b) Por definición $f'(x) > 0$; $f'(x) = f(x)\, g(x)$ y puesto que $f(x)$ es positiva, ha de ser $g(x) > 0$

c) Por definición $f''(x) \geq 0$; $f''(x) = f(x)\,[(g(x))^2 + g'(x)]$ y puesto que $f(x)$ es positiva, ha de ser $(g(x))^2 + g'(x) \geq 0$

Recíprocamente, dada una función $g(x)$ que verifica las propiedades 1) 2) y 3), por la definición (4.81), $f(x)$ es una función definida positiva para $x>0$, y por otra parte

a') Si $\lim_{x \to 0} G(x) = -\infty$ entonces $\lim_{x \to 0} f(x) = 0$

b') Si $g(x) > 0 \quad \forall\, x \in [0,1]$ la derivada de $f(x)$,

$$f'(x) = f(x)\, G'(x) = f(x)\, g(x)$$

puesto que $f(x)$ es positiva, $f'(x)$ también es positiva; luego la función f en el intervalo $[0,1]$ es creciente.

c') Si $(g(x))^2 + g'(x) \geq 0 \quad \forall\, x \in [0,1]$ entonces la segunda derivada de $f(x)$

$$f''(x) = f(x)\,[(g(x))^2 + g'(x)]$$

también es positiva, y por tanto $f(x)$ en el intervalo $[0,1]$ es cóncava

Obsérvese que a'), b') y c') , junto a la necesaria existencia de $G(1)$, pues sin ello no sería posible la definición (4.81) se obtiene una función que modeliza la curva de Lorenz. (la definición (4.81) asegura que la función pasa por el punto $(1,1)$.

En los ejemplos que siguen se estudian funciones que permiten obtener la curva de Lorenz y otras para las que no es posible tal construcción.

EJEMPLOS

1) En primer lugar se estudia una función que sí verifica las condiciones necesarias y suficientes para la modelización de la curva de Lorenz:

Sea la función

$$g(x) = a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \frac{a_3}{x^3} + \ldots + \frac{a_n}{x^n}$$

$$\text{con } a_0 > 0 \,,\; a_1 > \frac{n}{2} \,,\; a_2, a_3, \ldots a_n > 0$$

entonces

$$G(x) = a_0 x + a_1 \ln x - \frac{a_2}{x} - \frac{a_3}{2x^2} - \ldots - \frac{a_n}{(n-1)\,x^{n-1}}$$

$$G(1) = a_0 - a_2 - \frac{a_3}{2} - \ldots - \frac{a_n}{n-1}$$

se verifica que

$$\lim_{x \to 0} G(x) = -\infty$$

$$g(x) \geq 0 \qquad \forall\, x \in [0,1]$$

$$(g(x))^2 + g'(x) = a_0^2 + \frac{2a_0a_1}{x} + \frac{a_1(a_1-1) + 2a_0a_2}{x^2} +$$

$$+ \frac{2a_0a_3 + 2a_2(a_1-1)}{x^3} + \frac{a_2^2 + 2a_0a_4 + a_3(2a_1-3)}{x^4} +$$

$$+ \frac{2a_0a_5 + 2a_2a_3 + 2a_4(a_1-2)}{x^5} + \ldots + \frac{a_n^2}{x^n}$$

y puesto que con $a_0 > 0 \,,\; a_1 > \frac{n}{2} \,,\; a_2, a_3, \ldots a_n > 0$

la suma $(g(x))^2 + g'(x)$ es positiva $\forall\, x \in [0,1]$

y por tanto se puede generar una curva de Lorenz, cuya forma funcional es:

$$f(x) = e^{G(x) - G(1)} =$$

$$x^{a_1}\, e^{a_0(x-1) + a_2\left(1-\frac{1}{x}\right) + \frac{a_3}{2}\left(1-\frac{1}{x^2}\right) + \dots + \frac{a_n}{n-1}\left(1-\frac{1}{x^{n-1}}\right)}$$

2) Por el contrario, si $g(x)$ es una función polinómica, a partir de ella no se puede generar una curva de Lorenz, pues es ese caso su integral $G(x)$ también es un polinomio, lo que hace imposible que $\lim\limits_{x \to 0} G(x) = -\infty$.

En particular, $g(x) = 0$ no es generadora de una curva de Lorenz.

4.6.2 Condiciones para que una función que verifica la ecuación de Pearson sea función generadora de una curva de Lorenz.

Casas, Herrerías y Núñez, [7], manifiestan el deseo de conocer si todas las curvas que satisfacen la ecuación diferencial que define la familia de Pearson cumplen la condición necesaria de las curvas de Lorenz.

A continuación se hace un estudio completo de las condiciones que se han de cumplir para que una función definida de la forma

$$g(x) = \frac{x - a}{b_0 + b_1 x + b_2 x^2}$$

permita construir una curva de Lorenz.

I) Caso en el que $b_1 = b_2 = 0\,;\quad b_0 \neq 0$

PROPOSICIÓN 4.16

La función $g(x) = \dfrac{x - a}{b_0}$ $\quad (b_0 \neq 0)$ **no permite generar la curva de Lorenz.**

En efecto pues $G(x) = \dfrac{1}{b_0}\left(\dfrac{x^2}{2} - ax\right)$ cuyo límite, cuando x tiende a cero, no es menos infinito.

II) Caso en el que $b_2 = 0$; $b_1 \neq 0$

PROPOSICIÓN 4.17

La condición necesaria y suficiente para que la función

$g(x) = \dfrac{x - a}{b_0 + b_1 x}$ $(b_1 \neq 0)$ **permita generar una curva de Lorenz**

es que

a sea negativo

b_0 sea cero

$0 < b_1 \leq -a$

En efecto, es necesaria pues si la función

$$f(x) = e^{G(x) - G(1)}$$

modeliza una curva de Lorenz, siendo

$$G(x) = \int^x g(x)\,dx = \frac{x}{b_1} - \frac{b_1 a + b_0}{b_1^2} \ln (b_0 + b_1 x)$$

entonces el límite de G(x) cuando x tiende a cero ha de ser igual a menos infinito

$$\lim_{x \to 0} G(x) = - \frac{b_1 a + b_0}{b_1^2} \ln (b_0) = -\infty$$

de donde se deduce que $b_0 = 0$; $b_1 > 0$; $\dfrac{a}{b_1} < 0$ y por tanto

$a < 0$

por lo que la función queda de la forma:

$$g(x) = \frac{x - a}{b_1 x}$$

siendo

$$G(x) = \frac{x}{b_1} - \frac{a}{b_1} \ln (b_1 x)$$

y por tanto, G(1) existe,

$$G(1) = \frac{1}{b_1} - \frac{a}{b_1} \ln (b_1)$$

De las condiciones ya obtenidas para a, b_0 y b_1 se deduce que g(x) es positiva para todo x del intervalo cerrado [0,1].

Puesto que la fracción

$$(g(x))^2 + g'(x) = \frac{(x-a)^2 + ab_1}{b_1^2 x^2}$$

ha de ser positiva o cero, se ha de cumplir

$$(x-a)^2 + ab_1 \geq 0 \qquad \forall\, x \in [0,1]$$

que al ser a negativo, se deduce que

$$a^2 + ab_1 \geq 0$$

$$a + b_1 \leq 0$$

$$b_1 \leq -a$$

La suficiencia se prueba a partir de las condiciones de los coeficientes siguiendo los mismos pasos en orden inverso, y la función que modeliza la curva de Lorenz es:

$$f(x) = e^{G(x)-G(1)} = x^{-\frac{a}{b_1}}\, e^{\frac{(x-1)}{b_1}}$$

III) Caso en el que $b_2 \neq 0$

El estudio de este caso se realiza atendiendo a las posibles raíces del denominador: dos reales distintas, una real doble o dos complejas.

PROPOSICIÓN 4.18

La condición necesaria y suficiente para que la función

$$g(x) = \frac{x-a}{b_0 + b_1 x + b_2 x^2} \qquad (b_2 \neq 0)$$

en la que el denominador posee dos raíces reales, α, β, **distintas, permita generar una curva de Lorenz es que los coeficientes cumplan:**

$b_0 = 0$

$b_1 > 0$

$a \leq -b_1$

$0 < b_2 \leq 1 + \dfrac{a(a+b_1)}{1-2a}$

En efecto, es condición necesaria, pues dada la función $g(x) = \dfrac{x - a}{b_0 + b_1 x + b_2 x^2}$, cuyo denominador posee dos raíces reales, α, β, distintas, y supongamos que la función

$$f(x) = e^{G(x) - G(1)} \tag{4.82}$$

modeliza una curva de Lorenz,
siendo

$$G(x) = \int^x g(x)\,dx = \frac{1}{b_2}\left(\frac{\alpha - a}{\alpha - \beta}\ln(x-\alpha) + \frac{\beta - a}{\beta - \alpha}\ln(x-\beta)\right)$$

entonces el límite de $G(x)$ cuando x tiende a cero ha de ser igual a menos infinito:

$$\lim_{x \to 0} G(x) = \frac{1}{b_2}\left(\frac{\alpha - a}{\alpha - \beta}\ln(-\alpha) + \frac{\beta - a}{\beta - \alpha}\ln(-\beta)\right) = -\infty$$

de donde se deduce que una de las dos soluciones es cero y la otra negativa, es decir $\alpha < 0$; $\beta = 0$ o bien $\alpha = 0$; $\beta < 0$, (no se estudia la posibilidad $\alpha = \beta = 0$, pues estamos en el caso de dos soluciones reales distintas).

Sin pérdida de generalidad supongamos que $\beta = 0$; en este caso se ha cumplir además que $\dfrac{a}{b_2 \alpha} > 0$.

Por tanto, hasta ahora se tiene:

$$b_0 = 0, \quad b_1 \neq 0, \quad b_2 \neq 0, \quad \alpha < 0, \quad \frac{a}{b_2} < 0$$

La función puede escribirse como:

$$g(x) = \frac{x - a}{b_1 x + b_2 x^2} \quad (b_1 \neq 0\ ; \quad b_2 \neq 0)$$

entonces las soluciones del denominador son $\beta = 0$, $\alpha = -\dfrac{b_1}{b_2}$ y

puesto que α ha de ser negativo se deduce que b_1 y b_2 tienen el mismo signo.

La función G(x) se puede escribir:

$$G(x) = \frac{b_1 + ab_2}{b_1 b_2} \ln (b_2 x + b_1) - \frac{a}{b_1} \ln x$$

y dado que $0 \le x \le 1$ y que b_1 y b_2 tienen el mismo signo, para que G(x) exista, b_1 y b_2 deben ser positivos y por tanto a negativo.

Las dos condiciones de contorno hacen que a sea negativa que b_0 valga cero y que b_1 y b_2, sean positivos.

Puesto que la función f(x), dada en (4.82) modeliza una curva de Lorenz, su derivada ha de ser positiva, por tanto se verifica

$$g(x) = \frac{x - a}{b_1 x + b_2 x^2} > 0$$

que, junto a las condiciones que deben cumplir los coeficientes, b_1 y b_2 sólo se deduce que $a < x$, cosa que también ya es conocida.

Sólo queda por estudiar cómo condiciona a los coeficientes la concavidad de la curva de Lorenz. Puesto que

$$(g(x))^2 + g'(x) \ge 0 \quad \forall\, x \in [0,1]$$

para todo x del intervalo cerrado [0,1] se ha de cumplir que

$$\frac{(1-b_2)x^2 - 2a(1-b_2)x + a(a+b_1)}{(b_1 x + b_2 x^2)^2} \ge 0$$

y por tanto

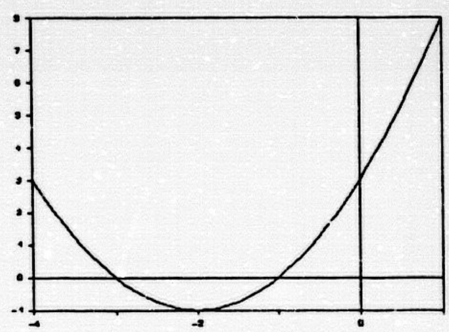
$$(1-b_2)x^2 - 2a(1-b_2)x + a(a+b_1) \ge 0 \quad \forall\, x \in [0,1] \tag{4.83}$$

Este trinomio de segundo grado se estudia bajo distintos supuestos para el coeficiente b_2, del que ya sabemos ha de ser positivo:

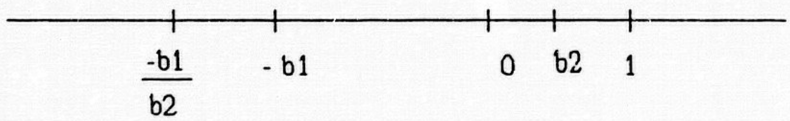
1) Si $0 < b_2 < 1$, la función $y = (1-b_2)x^2 - 2a(1-b_2)x + a(a+b_1)$, posee un mínimo en el punto $(a, a(b_1+ab_2))$. La inecuación (4.83) se verifica cuando:

a) El mínimo está en el segundo cuadrante o sobre el propio eje de las x, es decir $a(b_1+ab_2) \geq 0$ y puesto que a es negativo $ab_2 + b_1 \leq 0$, luego $a < -\frac{b_1}{b_2}$

b) El mínimo está en el tercer cuadrante, pero la parábola corta al eje y en un valor mayor o igual que cero, es decir: $a(ab_2 + b_1) < 0$ a la vez que $a(a+b_1) \geq 0$ y de ambas condiciones se deduce que $-\frac{b_1}{b_2} < a \leq b_1$

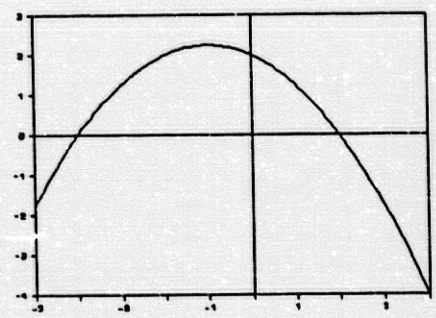


Obsérvese que $-\frac{b_1}{b_2} < -b_1$ puesto que $0 < b_2 < 1$, por tanto se puede concluir que para $0 < b_2 < 1$ la inecuación se verifica si y sólo si $a \leq -b_1$.



2) Si $b_2 > 1$, la función $y = (1-b_2)x^2 - 2a(1-b_2)x + a(a+b_1)$, posee un máximo en el punto $(a, a(b_1+ab_2))$. La inecuación (4.83) se verifica cuando:

a) El máximo está por encima del eje de las x, $a(b_1+ab_2) \geq 0$, es decir $a < -\frac{b_1}{b_2}$, y además una de las dos raíces es mayor o igual que uno, (la otra evidentemente es negativa).



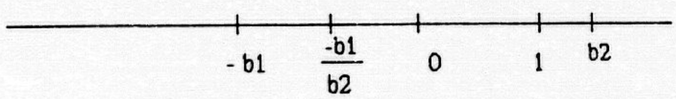
Para que una raíz sea mayor o igual que uno, cuando x=1, la parábola ha de ser positiva o cero, esto es:

$$1-b_2-2a+2ab_2+a^2+ab_1 \geq 0$$

$$b_2 \leq 1 + \frac{a(a+b_1)}{1-2a}$$

y para que el conjunto de posibles valores de b_2 no sea vacío ha de ocurrir que $\frac{a(a+b_1)}{1-2a} > 0$, de donde se deduce que $a < -b_1$

Obsérvese que $-b_1 < -\frac{b_1}{b_2}$ puesto que $b_2 > 1$, por tanto se puede concluir que para $b_2 > 1$ la inecuación se verifica si y sólo si $a < -b_1$.



3) Si $b_2 = 1$ entonces $a(a + b_1) \geq 0$. Puesto que a es negativa, entonces $a \leq -b_1$

Recíprocamente el mismo camino seguido en la necesidad se puede utilizar para demostrar la suficiencia, pues si los coeficientes verifican las condiciones exigidas, se verifica la inecuación (4.83), por lo que f es cóncava, g(x) es positiva, luego f es creciente, existe G(1) y el límite de G(x) cuando x tiende a cero es menos infinito, de donde se deduce que la función definida en (4.82) modeliza una curva de Lorenz.

PROPOSICIÓN 4.19

La condición necesaria y suficiente para que la función

$$g(x) = \frac{x - a}{b_0 + b_1 x + b_2 x^2} \quad (b_2 \neq 0),$$

cuyo denominador posee una raíz real doble α, permita generar una curva de Lorenz es que

$b_0 = 0$

$b_1 = 0$

$a < 0$

$0 < b_2 \leq 1 + \frac{a^2}{1-2a}$

En efecto, es condición necesaria, pues dada la función $g(x) = \frac{x - a}{b_0 + b_1 x + b_2 x^2}$, cuyo denominador posee una raíz real doble, α, y supongamos que la función

$$f(x) = e^{G(x) - G(1)} \tag{4.84}$$

modeliza una curva de Lorenz, siendo

$$G(x) = \int^x g(x)\, dx = \frac{1}{b_2}\left(\ln\,(x-\alpha) + \frac{a-\alpha}{x-\alpha}\right)$$

entonces el límite de G(x) cuando x tiende a cero ha de ser igual a menos infinito:

$$\lim_{x \to 0} G(x) = \frac{1}{b_2}\left(\ln\,(-\alpha) + \frac{a-\alpha}{-\alpha}\right) = -\infty$$

y para que esto ocurra ha de ser $\alpha = 0$, $(b_0 = b_1 = 0)$ y $b_2 > 0$. Entonces:

$$g(x) = \frac{x - a}{b_2 x^2}$$

$$G(x) = \frac{1}{b_2}\left(\ln\,(x) + \frac{a}{x}\right)$$

Puesto que la función f modeliza la curva de Lorenz, ha de ser

$$g(x) = \frac{x - a}{b_2 x^2} > 0$$

y teniendo en cuenta que $b_2 > 0$ ha de ocurrir que $x - a > 0$ y por tanto $a < 0$

Sólo queda por estudiar cómo condiciona a los coeficientes la concavidad de la curva de Lorenz:

$$(g(x))^2 + g'(x) \geq 0 \quad \forall x \in [0,1]$$

Es decir, para todo x del intervalo [0,1], se verifica

$$\frac{(1-b_2)x^2 - 2a(1-b_2)x + a^2}{(b_2x^2)^2} \geq 0$$

y por tanto

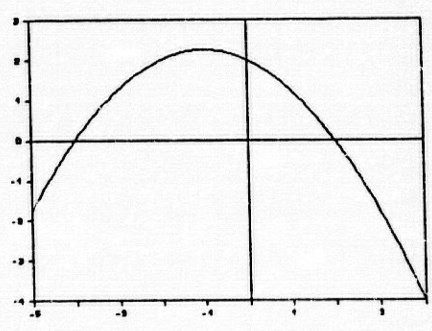
$$(1-b_2)x^2 - 2a(1-b_2)x + a^2 \geq 0 \quad \forall x \in [0,1]$$

(4.85)

De la misma forma que en la proposición anterior, este trinomio de segundo grado se estudia bajo distintos supuestos para el coeficiente b_2, del que ya sabemos ha de ser positivo:

1) Si $0 < b_2 < 1$, la función $y = (1-b_2)x^2 - 2a(1-b_2)x + a^2$, posee un mínimo en el punto (a, a^2b_2). La inecuación (4.85) se verifica siempre pues la ordenada del mínimo es positiva.

2) Si $b_2 > 1$, la función $y = (1-b_2)x^2 - 2a(1-b_2)x + a^2$, posee un máximo en el punto $(a, a^2b_2))$, que pertenece al segundo cuadrante. La inecuación (4.85) se verifica cuando una de las dos raíces es mayor o igual que uno, (la otra evidentemente es negativa).



De igual forma que antes, cuando x=1, la parábola ha de ser positiva o cero, esto es:

$$1-b_2-2a+2ab_2+a^2 \geq 0$$

$$b_2 \leq 1 + \frac{a^2}{1-2a}$$

Obsérvese que el conjunto de posibles valores de b_2 no es vacío pues $\frac{a^2}{1-2a} > 0$, ya que $a < 0$

3) Si $b_2 = 1$ entonces queda $a^2 \geq 0$, que evidentemente siempre se verifica.

La demostración de la suficiencia se puede hacer siguiendo los mismos pasos pero en orden inverso.

Obsérvese que esta proposición es el caso particular de la anterior en el que $b_1 = 0$

PROPOSICIÓN 4.20

Una función de la forma

$$g(x) = \frac{x - a}{b_0 + b_1 x + b_2 x^2} \qquad (b_2 \neq 0) \qquad \textbf{(4.86)}$$

en la que el denominador no posee raíces reales, no permite generar una curva de Lorenz.

En efecto, pues dada la función $g(x) = \dfrac{x - a}{b_0 + b_1 x + b_2 x^2}$, cuyo denominador no posee raíces reales, sean α, β dos números reales tales que

$$b_0 + b_1 x + b_2 x^2 = b_2 ((x - \alpha)^2 + \beta^2) \qquad \alpha, \beta \in \mathbb{R},$$

calculando la función G(x), resulta:

$$G(x) = \int^x g(x)\, dx = \frac{1}{2b_2} \ln ((x-\alpha)^2 + \beta^2) + \frac{\alpha - a}{\beta b_2} \operatorname{arctg}\left(\frac{x-\alpha}{\beta}\right)$$

el límite de G(x) cuando x tiende a cero ha de ser igual a menos infinito, lo que supone que $\alpha = \beta = 0$. Entonces estamos en el caso anterior. Por lo tanto a partir de una función de la forma (4.86) no se puede construir un modelo de curva de Lorenz.

4.7 LEYES FINANCIERAS ASOCIADAS A VARIABLES ALEATORIAS.

4.7.1 Introducción.

Gil Peláez, [17], anticipa la idea de obtener una ley financiera a partir de una función real de variable real, pero no realiza su estudio pues considera que se necesitan exigencias sobre integrabilidad y continuidad de la función de partida.

Recientemente Cruz Rambaud, en su tesis doctoral, [9], ha estudiado los capitales financiero-aleatorios, estableciendo una relación entre las leyes financieras y las funciones de distribución de variables aleatorias.

La idea desarrollada en esta memoria consiste en generar la función de distribución de una variable aleatoria a partir de una función real de variable real.

Obtenida la distribución, utilizando la relación establecida por Cruz, se establece una relación directa entre la ley financiera y la función generadora.

Como veremos más adelante, algunas de las condiciones que ha de cumplir esta función generadora son distintas según se pretenda obtener una ley de actualización o de capitalización.

Cruz obtiene las expresiones para generar leyes financieras aleatorias, en sentido amplio, a partir de la función de distribución del siguiente modo:

Estudia los capitales financiero-aleatorios (C_0,τ) donde el vencimiento τ es una variable aleatoria, que toma valores en un subconjunto del conjunto de los números enteros, con función de distribución $t \rightarrow F(t)$

Sustituye esta variable por ξ, las proyecciones de $(C_0, t-h)$, $h \in \{nZ\}$ en t, según una ley financiera, distinguiendo entre los dos supuestos:

1) Considera el tanto de muerte o terminación en el intervalo $[t, t+n]$ de la variable τ :

$$m(t,n) = -\frac{\Delta_n p(t)}{p(t)} = -\frac{p(t+n) - p(t)}{p(t)} \qquad (4.87)$$

siendo $p(t) = P(\tau > t)$ la función de pervivencia en t del capital (C_0, τ). Supone que esta tasa de terminación coincide con el rédito de contracapitalización de la ley en ese mismo período, $[t, t+n]$, y establece para la ley financiera de actualización o descuento la expresión

$$A(t,n) = Id_R \frac{p(t)}{p(t+n)} \qquad (4.88)$$

Puesto que se trata de un sistema multiplicativo, para el período $[t, t + h.n]$ establece la igualdad

$$A(t,h) = Id_R \frac{p(t)}{p(t+h)} = Id_R \frac{1 - F(t)}{1 - F(t+h)} \qquad (4.89)$$

2) El segundo supuesto considera el tanto de crecimiento en el intervalo $[t, t+n]$ de la variable τ:

$$c(t,n) = \frac{F(t+n) - F(t)}{F(t)} \qquad (4.90)$$

y supone que esta tasa coincide con el rédito de capitalización de la ley en $[t, t+n]$, y establece para la ley financiera de capitalización la expresión

$$L(t,n) = Id_R \frac{F(t+n)}{F(t)} \qquad (4.91)$$

y, para el período $[t, t + h.n]$, es

$$L(t,h) = Id_R \frac{F(t+h)}{F(t)} \qquad (4.92)$$

4.7.2 Función generadora de una ley financiera de descuento.

Para las leyes financieras de descuento se entenderá la función generadora de la supervivencia como

$$g(x) = \frac{d}{dx} \ln(1 - F(x)) = \frac{-F'(x)}{1 - F(x)} \qquad (4.93)$$

siendo F(x) la función de distribución de una variable aleatoria, definida en el intervalo [a,b], (a puede tender a menos infinito y b a más infinito).

La función de distribución, obtenida de (4.93) se puede expresar de la forma:

$$F(x) = 1 - e^{k} e^{G(x)}$$

donde $G(x) = \int^{x} g(x)\,dx$, la primitiva de g(x), cuya constante aditiva vale cero y k es un número real.

Las condiciones que ha de cumplir F para que sea función de distribución se trasladan a la función g, y resulta

PROPOSICIÓN 4.21

Una función g real de variable real es generadora de una ley financiera de descuento si y sólo si

1) **g(x) es integrable y continua salvo en un conjunto numerable de puntos, donde puede presentar saltos finitos.**
2) $g(x) \leq 0$
3) $\lim\limits_{x \to a} G(x) = -k \quad (k \in \mathbb{R})$
4) $\lim\limits_{x \to b} G(x) = -\infty$

En efecto:

1) La continuidad exigida supone que la función de distribución es continua a la derecha. Es evidente la exigencia de la integrabilidad de la función generadora.

2) Se ha de verificar $g(x) \le 0$, lo que equivale a decir que $F(x)$ es monótona creciente pues

$$F'(x) = - e^k e^{G(x)} G'(x) = - e^k e^{G(x)} g(x) \ge 0$$

3) $\lim\limits_{x \to a} G(x) = - k \quad (k \in \mathbb{R})$, pues equivale a $\lim\limits_{x \to a} F(x) = 0$, lo que nos permite calcular k: $k = - \lim\limits_{x \to a} G(x)$

4) $\lim\limits_{x \to b} G(x) = - \infty$, pues se verifica si y sólo si $\lim\limits_{x \to b} F(x) = 1$

Determinado ya el valor del número real k, la función de distribución viene dada por

$$F(x) = 1 - e^{G(x) - \lim\limits_{x \to a} G(x)}$$

Es decir, la función de distribución queda determinada, de forma única, a partir de una función $g(x)$ que verifique las condiciones anteriores, y para un valor concreto de la constante de integración.

También una función $g(x)$, real de variable real, que verifique las propiedades anteriores genera una ley financiera, pues la expresión (4.89) es

$$L(t,h) = Id_R \frac{1 - F(t)}{1 - F(t+h)} = Id_R \frac{e^k e^{G(t)}}{e^k e^{G(t+h)}}$$

$$L(t,h) = Id_R \; e^{G(t) - G(t+h)} = Id_R \; e^{-\int_t^{t+h} g(x)\,dx}$$

Obsérvese que la ley no depende de la constante de integración.

Por lo tanto, la función $F(x) = 1 - e^k e^{G(x)}$, si $k \ne - \lim\limits_{x \to a} G(x)$ no es de función de distribución de una cierta variable aleatoria y sin embargo da lugar a la misma ley financiera que $F(x) = 1 - e^{G(x) - \lim_{x \to a} G(x)}$, es decir la ley financiera viene asociada, de forma biunívoca a la función g(x) y no a la función de distribución de una variable aleatoria.

Ejemplos

1°) Sea $$g(x) = \begin{cases} -\lambda & \text{si } x > 0 \quad (\lambda > 0) \\ 0 & \text{si } x < 0 \end{cases}$$

entonces $$G(x) = \begin{cases} -\lambda x & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Se verifican las condiciones necesarias y suficientes para generar una distribución, pues g(x) es definida negativa, continua e integrable y además

$$\lim_{x \to 0} G(x) = 0$$

$$\lim_{x \to +\infty} G(x) = -\infty$$

y por tanto genera una ley expresada:

$$A(t,h) = Id_R \; e^{G(t) - G(t+h)} = Id_R \; e^{\lambda h}$$

2°) Sea $$g(x) = \begin{cases} -\dfrac{1}{b-x} & \text{si } a < x < b \\ 0 & \text{en el resto} \end{cases}$$

entonces $$G(x) = \begin{cases} \ln(b-x) & \text{si } a < x < b \\ 0 & \text{en el resto} \end{cases}$$

Se verifican las condiciones necesarias y suficientes para generar una distribución, pues g(x) es definida negativa, continua e integrable y además

$$\lim_{x \to a} G(x) = k \in \mathbb{R}$$

$$\lim_{x \to b} G(x) = -\infty$$

por tanto se genera ley financiera:

$$A(t,h) = Id_R \; e^{G(t) - G(t+h)} = Id_R \; \frac{b-t}{b-h-t}$$

3°) Sea

$$g(x) = -\lambda + \frac{d}{dx}\ln\left(\sum_{j=0}^{r-1}\frac{(\lambda x)^j}{j!}\right), \quad x\geq 0,\ \lambda>0,\ r\in\mathbb{N}$$

entonces $$G(x) = -\lambda x + \ln\left(\sum_{j=0}^{r-1}\frac{(\lambda x)^j}{j!}\right)$$

Se verifican las condiciones necesarias y suficientes para generar una distribución, pues

$$\frac{d}{dx}\ln\left(\sum_{j=0}^{r-1}\frac{(\lambda x)^j}{j!}\right) = \frac{\lambda\sum_{j=1}^{r-1}\frac{(\lambda x)^{j-1}}{(j-1)!}}{\sum_{j=0}^{r-1}\frac{(\lambda x)^j}{j!}} = \frac{\lambda\sum_{j=0}^{r-2}\frac{(\lambda x)^j}{j!}}{\sum_{j=0}^{r-1}\frac{(\lambda x)^j}{j!}} < \lambda$$

y por tanto g(x) es definida negativa, es continua e integrable y además

$$\lim_{x\to 0} G(x) = \lim_{x\to 0}\left(-\lambda x + \ln\left(\sum_{j=0}^{r-1}\frac{(\lambda t)^j}{j!}\right)\right) = 0$$

$$\lim_{x\to +\infty} G(x) = -\infty$$

calculando $$G(t) - G(t+h) = \lambda h + \ln\frac{\sum_{j=0}^{r-1}\frac{(\lambda t)^j}{j!}}{\sum_{j=0}^{r-1}\frac{(\lambda(t+h))^j}{j!}}$$

se obtiene la ley

$$A(t,h) = Id_R\, e^{\lambda h}\,\frac{\sum_{j=0}^{r-1}\frac{(\lambda t)^j}{j!}}{\sum_{j=0}^{r-1}\frac{(\lambda(t+h))^j}{j!}}$$

4°) Sea $g(x) = \dfrac{-e^{x-\alpha}}{1+e^{x-\alpha}} = -\dfrac{d}{dx}\ln(1+e^{x-\alpha}) \quad \forall x\in\mathbb{R}$

$$G(x) = -\ln(1+e^{x-\alpha})$$

se verifican las condiciones necesarias para que se pueda generar una ley financiera, cuya expresión es

$$A(t,h) = Id_R\ e^{G(t)-G(t+h)} = Id_R\,\frac{1+e^{t+h-\alpha}}{1+e^{t-\alpha}}$$

4.7.3 Función generadora de una ley financiera de capitalización.

Para las leyes financieras de capitalización la función generadora viene dada por

$$g(x) = \frac{d}{dx} \ln F(x) = \frac{F'(x)}{F(x)} \qquad (4.94)$$

siendo, igualmente, F(x) la función de distribución de una variable aleatoria, definida en el intervalo [a,b].

La función de distribución, obtenida de (4.93) se puede expresar de la forma:

$$F(x) = e^k \, e^{G(x)}$$

donde $G(x) = \int^x g(x)\, dx$, la primitiva de g(x), cuya constante aditiva vale cero y k es un número real.

Las condiciones que ha de cumplir F para que sea función de distribución se trasladan a la función g, y resulta

PROPOSICIÓN 4.22

Una función g real de variable real es generadora de una ley financiera de capitalización si y sólo si,

1) **g(x) es integrable y continua salvo en un conjunto numerable de puntos, donde puede presentar saltos finitos.**

2) $g(x) \geq 0$

3) $\lim_{x \to a} G(x) = -\infty$

4) $\lim_{x \to b} G(x) = -k, \quad k \in \mathbf{R}$

En efecto:

1) Es evidente la necesidad de integrabilidad de la función generadora, así como de la continuidad salvo en un conjunto numerable de puntos, donde puede presentar saltos finitos, para que la distribución sea continua a la derecha.

2) Se ha de verificar $g(x) \geq 0$, lo que equivale a decir que $F(x)$ es monótona creciente pues

$$F'(x) = e^k\, e^{G(x)}\, G'(x) = e^k\, e^{G(x)}\, g(x) \geq 0$$

3) $\lim_{x \to a} G(x) = -\infty$, pues se verifica si y sólo si $\lim_{x \to a} F(x) = 0$.

4) $\lim_{x \to b} G(x) = -k \in \mathbf{R}$, que equivale a $\lim_{x \to b} F(x) = 1$, lo que nos permite calcular k: $k = -\lim_{x \to b} G(x)$

Determinado ya el valor del número real k, la función de distribución viene dada por

$$F(x) = e^{G(x) - \lim_{x \to b} G(x)}$$

Si una función $g(x)$ que verifica las propiedades anteriores, genera una distribución. La expresión (4.92) de la ley financiera queda

$$L(t,h) = Id_R\, \frac{F(t+h)}{F(t)} = \frac{e^k\, e^{G(t+h)}}{e^k\, e^{G(t)}}$$

$$L(t,h) = Id_R\, e^{G(t+h) - G(t)} = Id_R\, e^{\int_t^{t+h} g(x)\, dx}$$

Las mismas observaciones que en el caso de la ley de descuento, es decir que la ley no depende de la constante de integración, es decir es única para cada función $g(x)$, y sin embargo existen funciones que sin ser de distribución dan lugar a la misma ley financiera.

Ejemplos

1°) Sea $$g(x) = \begin{cases} \dfrac{\lambda}{e^{\lambda x} - 1} & \text{si } x > 0 \quad (\lambda > 0) \\ 0 & \text{si } x < 0 \end{cases}$$

entonces $$G(x) = \begin{cases} \ln\,(1 - e^{-\lambda x}) & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Se verifican las condiciones necesarias y suficientes para generar una distribución, pues g(x) es definida positiva, continua e integrable y además

$$\lim_{x\to 0} G(x) = -\infty; \qquad \lim_{x\to+\infty} G(x) = 0$$

y por tanto genera una ley expresada:

$$L(t,h) = Id_R \; e^{G(t+h)-G(t)} = Id_R \frac{1-e^{-\lambda(t+h)}}{1-e^{-\lambda t}}$$

2°) Sea
$$g(x) = \begin{cases} \dfrac{1}{x-a} & \text{si } a < x < b \\ 0 & \text{en el resto} \end{cases}$$

entonces
$$G(x) = \begin{cases} \ln(x-a) & \text{si } a < x < b \\ 0 & \text{en el resto} \end{cases}$$

Se verifican las condiciones necesarias y suficientes para generar una distribución, pues g(x) es definida positiva, continua e integrable y además

$$\lim_{x\to a} G(x) = -\infty$$

$$\lim_{x\to b} G(x) = \ln(b-a)$$

por tanto se genera una ley financiera:

$$L(t,h) = Id_R \; e^{G(t+h)-G(t)} = Id_R \frac{t+h-a}{t-a}$$

$$L(t,h) = Id_R \left(1 + \frac{h}{t-a}\right)$$

3°) Sea $g(x) = \dfrac{1}{1+e^{x-\alpha}} = -\dfrac{d}{dx} \ln\left(1+e^{-(x-\alpha)}\right) \quad \forall x \in \mathbb{R}$

$$G(x) = -\ln\left(1+e^{-(x-\alpha)}\right)$$

se verifican las condiciones necesarias para que se pueda generar una ley financiera, cuya expresión es

$$L(t,h) = Id_R \; e^{G(t+h)-G(t)} = Id_R \frac{1+e^{-(t-\alpha)}}{1+e^{-(t+h-\alpha)}}$$

$$L(t,h) = Id_R \, e^{h} \frac{1+e^{t-\alpha}}{1+e^{t+h-\alpha}}$$

EPÍLOGO

INVESTIGACIONES FUTURAS SOBRE LAS FUNCIONES GENERADORAS.

La idea de función generadora surge cuando, utilizando la metodología propia de las familias pearsonianas, se pretende obtener la distribución de probabilidad a partir de una determinada función real de variable real, cumpla o no la ecuación que define la familia de Pearson. Los distintos pasos, del continuo al discreto, del univariante al bivariante y al multivariante, han sido los propios de la citada metodología, pero, en todo caso, las únicas exigencias sobre las funciones generadoras vienen dadas porque sea posible la generación de una distribución de probabilidad. De este modo queda establecido el concepto de sistema de funciones generadoras, unificando en su tratamiento distintas distribuciones, a la vez que se pone de relieve el paralelismo entre los casos continuo y discreto.

Cuando comparábamos los distintos estudios existentes en la literatura sobre las familias pearsonianas y el concepto de función generadora, decíamos que "*se pretende también una herramienta que permita el estudio de las propiedades de las variables ...* ", y como tal hemos usado las funciones generadoras en este trabajo. Hemos podido comprobar el buen funcionamiento de las mismas en determinados temas concretos: recordemos la facilidad de cálculo que conllevan los teoremas de independencia; la relación que, de una forma sencilla, permite comprobar si una familia de distribuciones a priori es conjugada, o no lo es, para una clase de verosimilitudes; la propiedad de rectangularidad; la cantidad de información como matriz de covarianzas, etc.

Por otra parte el concepto amplio de función generadora nos ha permitido aumentar el campo de aplicación, generando no ya distribuciones de probabilidad, sino también funciones que modelizan la curva de Lorenz, o leyes financieras.

La diversidad de sus aplicaciones y la sencillez de algunos resultados nos animan a pensar que su utilidad puede ir mucho más lejos, no sólo en el campo de la Estadística Teórica, sino también en estudios de modelización de fenómenos financieros, económicos, etc.

A la vista de los resultados hasta ahora obtenidos, en principio, parece oportuno estudiar, bajo el prisma de las funciones generadoras, los siguientes temas:

1° Como hemos visto, en el cuerpo de este trabajo, las funciones generadoras de la distribución normal bivariante y las funciones de regresión están relacionadas. Esto induce a pensar que en la caracterización de distribuciones, las funciones generadoras podrían jugar un papel clarificador y sistematizador interesante. Es una cuestión que sólo se ha tocado tangencialmente y queda abierta para su estudio futuro.

2° De forma análoga, en algunos aspectos que hemos estudiado de la Estadística Matemática, a modo de aplicación de las funciones generadoras, se ha observado un fenómeno de simplificación en procesos demostrativos y, en algunas ocasiones, un papel clarificador de los propios conceptos; estos resultados animan a investigar más profundamente el papel de las funciones generadoras en este campo y llevar su aplicación hasta los problemas sobre los que actualmente se está trabajando.

3° Puesto que las funciones generadoras han surgido de la metodología pearsoniana, sería interesante estudiar todos los problemas de inferencia de esta metodología desde un punto de vista más general.

4° También los resultados obtenidos en el campo del análisis bayesiano parecen augurar un buen futuro a la utilización de las funciones generadoras como herramienta en el estudio y análisis de alguno de los problemas actuales.

5° Dado que en el caso discreto la función generadora tiene el aspecto de una tasa de variación [1], quizá una forma de profundizar en su tratamiento podría ser la incorporación de los conocimientos de Matemática Financiera, y estudiar la interrelación de una forma más profunda de lo que en este trabajo hemos hecho.

(1) Véase por ejemplo **Martín Pliego, F.J.** (1994). Introducción a la Estadística Económica y Empresarial. AC.

BIBLIOGRAFÍA

[1] **Aldanondo, I.** (1969). Ecuaciones diferenciales lineales de coeficientes constantes en derivadas parciales en el operador "aleph". Publicaciones del Seminario de Matemáticas García de Galdeano, Vol. 10, pág. 47 - 70. Universidad de Zaragoza.

[2] **Apostol, T.M.** (1960). Análisis Matemático. Reverté.

[3] **Barra, J.R.** (1971) Notions fondamentales de statistique mathématique. Dunod-Université.

[4] **Berger, J.O.** (1988). Statistical Decision Theory and Bayesian Analysis. Springer-Verlag. New York.

[5] **Callejón, J. y Santos, M.** (1994). Comparación de dos métodos de estimación: Estimación máximo verosímil y método de los momentos. VIII Reunión Anual de ASEPELT-España, volumen I. pág. 245 - 249. Universitat de las Illes Balears.

[6] **Cansado, E.** (1950). Exposición sistemática de las distribuciones de Pearson. Trab. de Estadística. Vol 1. Cuad. III, pág. 279 - 287.

[7] **Casas, J.M., Herrerías, R. y Núñez, J.** (1990). Familias de formas funcionales para estimar la curva de Lorenz. IV Reunión Anual ASEPELT-ESPAÑA. Murcia.

[8] **Casas, J.M. y Núñez, J.** (1991). Sobre la medición de la desigualdad y conceptos afines. Actas de la V Reunión Anual ASEPELT-ESPAÑA. Las Palmas.

[9] **Cruz Rambaud, S.** (1994). Nuevo enfoque de las leyes financieras a través de la teoría algebraica de autómatas. Tesis Doctoral. U.N.E.D.

[10] **Dawid, A.P.** (1979). Conditional Independence in Statistical Theory. J.R. Statist. Soc. B, 41, N°. 1, pág. 1 - 31.

[11] **Degroot, M.H.** (1988). Probabilidad y Estadística. Addison - Wesley Iberoamericana. México.

[12] **Elderton, W.P. y Johnson N.L.**(1969), System of frequency curves. Cambridge University Press.

[13] **Fajardo, M.A.** (1985). Generalizaciones de los sistemas pearsonianos discretos. Universidad de Extremadura.

[14] **Feller, W.** (1968). An Introduction to Probability. Theory and its applications. Vol I. Wiley.

[15] **Fernández de Trocóniz, A.** (1980).Introducción a las teorías de las Probabilidades. Estadística clásica y Estadística bayesiana. Edición del autor.

[16] **Fernández, F.** (1979). Una extensión del sistema de Pearson bivariante o sistema de van Uven. Publicaciones de la Facultad de Ciencias de Granada.

[17] **Gil Peláez, L.** (1987). Matemática de las Operaciones financieras. Ed. AC. Madrid.

[18] **Goldberg, S.** (1964). Introducción a las ecuaciones en diferencias finitas. Marcombo.

[19] **Herrerías, R.**(1975). Sobre las estructuras estadísticas de Pearson y exponenciales, problemas asociados. Publicaciones de la Facultad de Ciencias de Granada.

[20] **Herrerías, R.** (1976). Extensión del sistema de distribuciones discretas de Pearson. Cuadernos de Estadística Matemática. Serie A. n° 3, pág. 30-36. Facultad de Ciencias. Granada.

[21] **Herrerías, R. y Cobos, J.** (1984). Solución general para un tipo de sistemas de distribuciones de probabilidad bivariantes discretas. Cuadernos Aragoneses de Economía N° 8, pág. 133 - 137

[22] **Herrerías, R. y Calvete, H.** (1985). Un sistema de distribuciones discretas bivariantes. Estadística Española, 109, 15-28.

[23] **Herrerías, R. y Calvete, H.** (1986). Estudio del sistema de distribuciones de probabilidad bivariantes discretas del tipo Pearson-Ord. Cuadernos Aragoneses de Economía, N° 10, pág. 81-86.

[24] **Ibragimov, I.A. y Has'minskii, R.Z.**(1981). Applications of Mathematics. Statistical Estimation. Asymptotic Theory. Springer-Verlag. New York.

[25] **Johnson, N.L. y Kotz, S** (1970). Distributios in Statistic: Continuous univariate distributions-1. John Wiley & Sons.

[26] **Johnson, N.L. y Kotz, S** (1972). Distributios in Statistic: Continuous multivariate distributions. Wiley.

[27] **Johnson, N.L. y Kotz, S** (1972). Distributios in Statistic: Discrete distributions. Wiley & Sons.

[28] **Kendall, M.G. y Stuart, A.** (1967, 1969). The avanced theory of statistics. Vol. I and II. Griffin. London.

[29] **Lafuente Lechuga, M.** (1994). Medidas de cuantificación de la Desigualdad: La Desigualdad de la Renta en España según la E.P.F. 1990-91. Tesis Doctoral. Universidad de Murcia.

[30] **Lee, P.** (1989). Bayesian Statistics: an Introduction. Oxford University Press.

[31] **Lehmann, E.** (1959). Testinng statistical hypotheses. Wiley.

[32] **Loève, M.** (1963). Probability Theory. Van Nostraud Company. New York.

[33] **Mardia, K.V.** (1970). Families of bivariate distributions. Griffin. London

[34] **Ord,J.K.** (1967). On a system of discrete distributions. Biometrika, 54, pág. 649-656.

[35] **Pearson, k.** (1895). Memoir on skew variation in homogeneous material. Phil. Trans. Roy. Soc., A. nº 186, pág. 343 - 414.

[36] **Pearson, K.** (1923). Notes on skew frequency surfaces. Biometrika. Vol. 15. pág. 222-230.

[37] **Pearson, K.** (1923). On non-skew frequency surfaces. Biometrika. Vol. 15. pág. 231-244.

[38] **Puig Adam, P.** (1973). Curso Teórico Práctico de Cálculo Integral aplicado a la Física y Técnica. Editorial Biblioteca Matemática.

[39] **Rényi, A.** (1976). Cálculo de Probabilidades. Reverté.

[40] **Robert, C.** (1992). L'Analyse Statistique Bayesienne. Ed. Economica. Paris.

[41] **Rohatgi, V.K.** (1976). An Introduction to Probability Theory and Mathematical Statistics. John Wiley & Sons. New York.

[42] **Roy, L.K.** (1971) "An extension of the Pearson systems of frecuency curves". Trabajos de Estadística e I.O. Vol XXII. Cuad. 1 y 2, pág. 113-123.

[43] **Steyn, H.S.** (1960). On regression properties of discrete sistems of probability functions. Proceeding of the Royal Academy of Sciences, Amsterdan. Vol 63, pág. 302 - 311.

[44] **Uven, M.J. van.** (1947 - 48). Extensions of Pearson's probability distributions to two variable. Proceeding of the Royal Academy of Sciences. Amsterdam, Vol. 50, pág. 1063-1070 y 1252-1264; Vol.51, pág. 41-52 y 191-196.

[45] **Wadsworth, G.P. Bryan, J.G.** (1974). Applications of Probability and Randon Variables. Mc Graw Hill.

[46] **Wilks, S.S.** (1962). Mathematical Satatistics. John Wiley & Sons. New York.

[47] **Wolf-Rüdiger, K.** (1986). Statistical decision theory and insurance ratemaking. Methods of Operations Research, 57. XI symposium on operations research. pág. 559 - 581.

[48] **Zacks, S.** (1971). The theory of statistical inference. Wiley.