Universidad de Granada
Facultad de Ciencias

APORTACIONES AL ESTUDIO DE LA DISTRIBUCION HIPERGEOMETRICA: CARACTERIZACION Y PROPIEDADES DE MONOTONIA

Héctor Manuel Ramos Romero

Tesis de Doctorado en Ciencias, Sección de Matemáticas juzgada el día 17 de Marzo de 1989 en la Facultad de Ciencias de la Universidad de Granada, y calificada con Apto Cum Laude por el siguiente tribunal:

Presidente: Dr. D. Ramón Gutierrez Jaimez, Catedrático de la Universidad de Granada.

Vocales: Dr. D. Luis Parras Guijosa, Catedrático de la Universidad de Málaga.

Dr. D. Juan Luis Romero Romero, Catedrático de la Universidad de Cádiz.

Dr. D. Antonino García Rendón, Catedrático de la Universidad de Cádiz.

Secretario: Dr. D. Elias Moreno Bas, Catedrático de la Universidad de Granada.

Director de la Tesis

Fdo. Jorge Ollero Hinojosa

Autor de la Tesis

Fdo. Héctor M. Ramos Romero

T-17-76

B. UNIVERSITARIA
GRANADA
Nº Documento 28943
Nº Copia 28958

FACULTAD DE CIENCIAS
BIBLIOTECA
GRANADA

APORTACIONES AL ESTUDIO DE LA DISTRIBUCION HIPERGEOMETRICA:

CARACTERIZACION Y PROPIEDADES DE MONOTONIA

UNIVERSIDAD DE GRANADA

FACULTAD DE CIENCIAS

Núm. 383

DON JUAN DE DIOS PEREZ JIMENEZ, PROFESOR TITULAR Y SECRETARIO DE LA FACULTAD DE CIENICAS - DE LA UNIVERSIDAD DE GRANADA,

C E R T I F I C O: Que la presente Tesis Doctoral corresponde a la presentada en esta Facultad por D. Hector Manuel Ramos Romero, que fué calificada con Apto "cum laude" el día 17 de Marzo de 1989 por el tribunal correspondiente que fué aprobado en Comisión de - Doctorado celebrada el día 22 de Fe-- brero de 1989, siendo el tema "Aportaciones al estudio de la distribución hipergeométrica: caracterizaición y - propiedades de monotonia".

Y para que conste y surta los efector correspondientes se extiende la presente certificación con el visto bueno del Iltmo. Sr. Decano en Granada a veintisiete de marzo de mil novecientos ochenta y nueve.

Vº. Bº.

EL DECANO,

FACULTAD DE CIENCIAS - UNIVERSIDAD DE GRANADA

UNIVERSIDAD DE GRANADA
Facultad de Ciencias
Fecha 30 MAR. 1989
SALIDA NUM. 383

APORTACIONES AL ESTUDIO DE LA DISTRIBUCION HIPERGEOMETRICA: CARACTERIZACION Y PROPIEDADES DE MONOTONIA

BIBLIOTECA UNIVERSITARIA
GRANADA
Nº Documento .....
Nº Copia .....

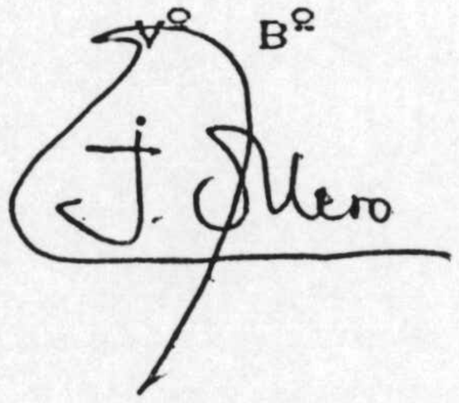
Memoria para optar al grado de Doctor en Ciencias, Sección de Matemáticas, que presenta el Licenciado

Héctor Manuel Ramos Romero.

Director de Tesis:

Prof. Dr. Jorge Ollero Hinojosa

V.º B.º



UNIVERSIDAD DE GRANADA. FACULTAD DE CIENCIAS

Doy las gracias

A mi amigo y maestro, Profesor Jorge Ollero Hinojosa, Director de esta memoria, por su magisterio, y por haberme iniciado en el oficio de investigar.

A Ana Mª Gómez Parra, por su inapreciable colaboración en el mecanografiado de originales.

A todos los que en algún momento se han interesado en este trabajo y, especialmente, al Profesor Antonino García Rendón.

Cádiz, Febrero de 1989

A Adela, y a nuestros hijos,
Aída, María y Héctor.

Número, acento, rima, nunca están
en el verso, sino siempre en el poeta.

(JUAN RAMON JIMENEZ)

INTRODUCCION GENERAL

INTRODUCCION GENERAL

Realizamos en esta memoria un estudio en profundidad de la distribución hipergeométrica, con dos vertientes claramente diferenciadas.

La primera parte de la memoria se desenvuelve según los esquemas habituales del estudio de una distribución de probabilidades; analizandose, en el capítulo I, la pertenencia de la hipergeométrica a familias de distribuciones, las funciones generatrices, momentos, principales propiedades, estimación de los parámetros, y aproximaciones por otras distribuciones, mientras que reservamos el capítulo II a teoremas de caracterización, por ser en este tema donde plasmamos parte de nuestra aportación personal.

La segunda parte de la memoria ofrece el aspecto más novedoso del tema: el referido a las ordenaciones de distribuciones en dispersión y aleatoriedad (capítulo III) y a las propiedades de monotonía que se derivan del análisis del comportamiento de la dispersión y de la aleatoriedad en la sucesión de distribuciones hipergeométricas $H_m(mN,p,n)$, que, como es bien sabido, converge hacia la binomial $B(n,p)$ (capítulo IV).

El calificativo de novedoso se justifica por el hecho de ser muy reciente la literatura que recoge este tema -la mayor parte de ella es posterior a 1980-, y por ser aquí donde se concreta el grueso de muestra aportación al estudio

de la hipergeométrica. En este sentido, el contenido del capítulo IV es, en su integridad, de nuestra exclusiva autoría.

Comentaremos, a continuación, con más detalles, el desarrollo de cada uno de los capítulos.

En el primer capítulo recopilamos los resultados más importantes existentes en la literatura acerca de la distribución hipergeométrica.

Partimos de tres libros básicos: el libro de Ord (1972) sobre familias de distribuciones, el manual de Johnson y Kotz (1969) sobre distribuciones discretas y el Patel, Kapadia y Owen (1976), sobre distribuciones en general. En estos dos últimos libros aparece recogida la mayor parte de dichos resultados, si bien, debido al carácter de manuales, vienen expresados de forma extremadamente concisa. Nuestra aportación en este capítulo ha consistido en completarlos, bien construyendo las demostraciones de los mismos, bien a través del estudio de los trabajos originales, comentandolos y llevando a cabo una síntesis de ellos, o bien, en algún caso, acabando de perfilar algún aspecto de la hipergeométrica, como puede ser la adscripción a determinadas familias de distribuciones que no hemos visto explícitamente recogida en la bibliografía revisada.

Queremos reseñar que el estudio está actualizado con

algunos resultados publicados en estos últimos años y, por tanto, no recopilados en los citados textos básicos.

Por otra parte, podremos observar que el grado de profundización y la intensidad con que se aborda cada uno de los aspectos no es uniforme, sino reflejo, en la mayoría de los casos, de la densidad bibliográfica correspondiente, aunque otras veces responde a nuestro personal interés, como ocurre con el trabajo de Uhlmann (1966). Este autor ofrece unos resultados, fruto de la comparación sistemática de la hipergeométrica con la binomial de igual media, que van a constituir uno de los pilares básicos sobre los que se van a sustentar nuestros resultados del capítulo IV.

Como parte de la informacion contenida en este primer capítulo es conocida por los lectores formados en nuestra disciplina, puede resultar preferible omitir inicialmente su lectura, llevando a cabo, a posteriori, la de aquellos apartados necesarios para seguir el desarrollo de los capítulos II y IV.

El estudio de la hipergeométrica en su vertiente más convencional se completa en el capítulo II, a cuyo contenido hemos considerado conveniente conferirle un tratamiento diferenciado, por ser aquí, como ya dijimos, donde aportamos nuestros primeros resultados personales, referidos a teoremas de caracterización de la familia hipergeométrica.

En primer lugar queremos señalar el hecho de que ninguno de los textos básicos ya mencionados recoge caracterización alguna para la distribución hipergeométrica, a diferencia de los que ocurre con otras distribuciones discretas.

La única caracterización encontrada es fruto de nuestra obligada revisión bibliográfica y se debe a Skibinsky (1970). Este autor fundamenta su resultado en el comportamiento de la distribución hipergeométrica H(N,X,n) ante la mixtura o composición, respecto del parámetro X (X=pN), con una distribución binomial B(N,p): la distribución resultante de la mixtura se corresponde con la binomial B(n,p).

Nosotros hemos estudiado en este capitulo la composición de la hipergeométrica H(N,X,n), respecto del segundo parámetro, con la hipergeométrica H(N',X',N) y comprobamos que la distribución resultante es, asimismo, hipergeométrica H(N',X',n).

Observamos, por tanto, que la distribución compuesta es, de nuevo, del mismo tipo que la supuesta para el parámetro X, lo que nos indujo a construir un teorema de caracterización siguiendo las pautas marcadas por Skibinsky, diferenciandonos de este autor en la distribución seguida por el parámetro.

Nos planteamos, a continuación, investigar si ese comportamiento singular de la distribución hipergeométrica se sigue manifestando cuando la distribución supuesta para el parámetro es una distribución cualquiera perteneciente al

esquema de urna de Polya al que, como sabemos, pertenecen las distribuciones binomial e hipergeométrica.

El resultado del análisis nos confirma la presunción y nos encontramos, por tanto, con la posibilidad de obtener un nuevo teorema de caracterización, del cual el de Skibinsky y el nuestro anterior son casos particulares.

La segunda parte de la memoria tiene como objetivo último el estudio de la sucesión de distribuciones hipergeométricas $H_m(mN,p,n)$; concretamente, lo que se refiere al comportamiento monótono de la dispersión y de la aleatoriedad en la citada sucesión.

Nuestro trabajo se sustenta en la posibilidad de ordenar parcialmente distribuciones en dispersión y en aleatoriedad, de manera que las medidas usuales de dichas características sean consistentes con el orden establecido, sin que sea necesario, por tanto, descender al estudio de medidas particulares.

En este sentido y ciñiendonos de momento a la ordenación en dispersión, se dirá que una variable aleatoria X es menor o igual en variabilidad que otra variable aleatoria Y, y se notará $X \underset{v}{\leq} Y$, si $E[g(X)] \leq E[g(Y)]$, para toda función g real y convexa.

Si tenemos en cuenta que cualquiera de las medidas usuales de dispersión puede ser expresada como valor esperado de función real convexa, esta definición nos asegura la

consistencia de las citadas medidas con la ordenación en variabilidad.

Aunque nuestro trabajo podía haberse desarrollado a partir de los resultados de un número muy limitado de autores, entre los que destacamos especialmente a Whitt (1985) y Hickey (1982-1983), no nos hemos querido restringir al estudio de los mismos, sino que, por el contrario, hemos dedicado el capítulo III a recopilar los principales resultados existentes en la literatura sobre este tema; labor que nos ha sido facilitada por ser, toda ella, de muy reciente publicación y, en su mayoría, fechada en esta última década. Naturalmente, en ese amplio contexto, la atención dedicada a los mencionados autores ha sido preferente.

En primer lugar, presentamos los diferentes órdenes parciales de distribuciones que han ido surgiendo ante la necesidad de formalizar la idea intuitiva de "ser más dispersa que", atendiendo a las diferentes vías -por otra parte, entrecruzadas-, por las que discurren y respetando, en la medida de lo posible, el orden cronológico.

Los caminos seguidos son, fundamentalmente, cuatro:

a) Ordenación de las distribuciones atendiendo a la dispersión respecto de un valor central.
b) Estudio de los cambios de signo de la función G-F, donde F

y G son, respectivamente, las funciones de distribución de las distribuciones comparadas.

c) Comparación de funciones adecuadas de cuantiles.

d) Estudio de la log-concavidad relativa.

Es precisamente la log-concavidad relativa la que nos va a proporcionar un acceso cómodo, a través de una relación interpuesta (la relación de variabilidad condicional uniforme, definida por Whitt), para poder establecer la comparación en variabilidad.

La ordenación en aleatoriedad se recoge en la siguiente sección de capítulo III y es desarrollada en exclusiva por Hickey a lo largo de diferentes trabajos. El fundamento de esta ordenación se encuentra en la teoría de mayoración de vectores.

El criterio de Hickey para comparar en aleatoriedad distribuciones discretas establece que, dadas dos distribuciones P y Q, con probabilidades puntuales $\{p_i\}$ y $\{q_i\}$, con i:1,2,... ,

$$P \underset{\alpha}{<} Q \quad \Leftrightarrow \quad \sum_{i=1}^{j} p_{[i]} > \sum_{i=1}^{j} q_{[i]} , \quad \forall j:1,2,\ldots ,$$

donde con $p_{[i]}$ y $q_{[i]}$ se notan las respectivas i-ésimas probabilidades puntuales, ordenadas éstas en forma decreciente.

Cualquiera de las medidas usuales de la aleatoriedad, como son la entropía de Shannon, la información de orden α de Renyi y la opuesta de la energía informacional de Onicescu, son consistentes con la ordenación establecida por Hickey.

Por fin, en el capítulo IV de la memoria, recogemos el grueso de nuestra aportación al estudio de la distribución hipergeométrica, aplicando resultados recopilados en el capítulo anterior.

(i) Comenzamos analizando el comportamiento monótono creciente de las medidas usuales de dispersión en la sucesión de hipergeométricas $H_m(mN,p,n)$, lo que pondrá de manifiesto que la manipulación de estas medidas puede presentar serios inconvenientes de índole operativo. Todos esas dificultades se soslayan demostrando que la familia H_m está ordenada en variabilidad.

Como paso previo para alcanzar este objetivo, demostraremos que

(ii) H_m converge de forma monótona creciente a la distribución B(n,p) en log-concavidad relativa.

(iii) Los elementos de la familia H_m no están ordenadas estocásticamente.

(iv) Como consecuencia de (ii) y (iii) podemos afirmar que la familia H_m está ordenada en variabilidad condicional uniforme y, por otra parte, al tener todas las distribuciones igual media, se concluye, finalmente, que están ordenadas en variabilidad.

(v) Seguidamente, y como una aplicación más del teorema anterior, ofrecemos un interesante resultado acerca del crecimiento monótono de la entropía Δ de las distribuciones pertenecientes a la familia H_m.

La validez del resultado no se circunscribe al caso de distribuciones hipergeométricas, sino que afecta a una clase más amplia de distribuciones: la clase $\mathcal{U}^*$ de las distribuciones fuertemente unimodales.

La siguiente sección está dedicada al estudio de la aleatoriedad.

(vi) En primer lugar, adecuaremos convenientemente los resultados de Uhlmann (1966), de manera que podamos comparar las funciones de distribución de dos hipergeométricas cualesquiera pertenecientes a la sucesión H_m , salvo en un intervalo I que contiene, a lo sumo, un número entero.

Siguiendo el criterio de Hickey y a través de la comparación de las funciones de distribución, ofrecemos un primer resultado parcial:

(vii) La aleatoriedad de las distribuciones de la sucesión de hipergeométricas $H_m(mN,p,n)$, con soporte común $\{0,1,...n\}$, crece monótonamente siempre que la media de las distribuciones sea un número entero.

Nuestro siguiente objetivo es generalizar este teorema, eliminando la restricción impuesta sobre la media.

Si nos fijamos en el criterio de Hickey, podemos observar que para j=1 se obtiene una condición necesaria para que $H_r \underset{\alpha}{<} H_s$: la probabilidad modal de la distribución H_r debe ser mayor que la correspondiente probabilidad en H_s. Esto pone de manifiesto la importancia del papel a desempeñar por la moda en nuestro estudio. Consecuentemente, dirigimos nuestros pasos en esa dirección:

(viii) Realizamos un amplio y detallado estudio sobre la localización de la moda M en la distribución hipergeométrica H(N,p,n), su posición relativa respecto del intervalo I y su comportamiento en la sucesión H_m .

(ix) Demostramos que siempre que las distribuciones $H_m(mN,p,n)$, con soporte común $\{0,1,..n\}$, no localicen la moda en los valores extremos del soporte, se cumple que la probabilidad modal es decreciente en la citada sucesión.

(x) Cuando la moda se localiza en los valores extremos del soporte, la sucesión de probabilidades modales será, por

el contrario, creciente. Sin embargo, esto no se continuará con un comportamiento opuesto de la aleatoriedad en la familia H_m, sino que, en este caso, sólo podremos dar resultados sobre medidas concretas.

La aplicación del criterio de Hickey se sigue con la comparación de la suma de las dos mayores probabilidades puntuales. En este sentido,

(xi) Obtendremos condiciones suficientes para poder afirmar, en cada caso, cuál de los dos enteros adyacentes a la moda tiene asignado mayor probabilidad, lo que nos permitirá localizar la segunda probabilidad puntual en orden decreciente.

(xii) Obtendremos, seguidamente, una condición suficiente que nos permite localizar las tres mayores probabilidades puntuales.

Todo este estudio previo será utilizado, finalmente, para generalizar nuestro primer resultado parcial sobre la aleatoriedad de las distribuciones de la familia $H_m(mN,p,n)$, con soporte común $\langle 0,1,...,n\rangle$:

(xiii) Siempre que la moda no se localice en los valores extremos del soporte, en cualquiera de las dos siguientes situaciones,

$$(1) \qquad p(n-1) + \frac{p(n-1)}{N} \leq M$$

$$(2) \qquad p(n-1) + \frac{p(n-1)}{N} > M \quad y \quad p(n+1) > M+q,$$

donde M es la moda de H(N,p,n) y q=1-p, se verifica que

$$H_r \underset{a}{<} H_s \ , \qquad \forall\ r<s \ .$$

Barcelona
L.e.H.G.P.

INDICE

PRIMERA PARTE

CAPITULO I
LA DISTRIBUCION HIPERGEOMETRICA

I.1 Introducción 2

I.2 Definición 4

I.3 Pertenencia a familias discretas de probabilidad .. 9
Sistema de Pearson discreto 9
Distribuciones fuertemente unimodales 11
Distribuciones de razón de verosimilitud monótona . 13
Distribuciones de razón de fallo creciente 15
Distribuciones " nuevo mejor que usado " 16

I.4 Funciones generatrices 17
Función generatriz de probabilidades 17
Función generatriz de momentos 18
Función característica 18

I.5 Momentos 18
Momentos factoriales 18
Momentos centrales 20
Momentos respecto del origen 23
Otros tipos de momentos 24
Desviación media 24
Momentos inversos factoriales 25
Momentos negativos respecto del origen 27

I.6 Principales propiedades 29
Simetría 31
Localización de la moda 32

I.7 Estimación de los parámetros 33
Estimación del parámetro X 34
Estimación del parámetro N 35
Estimación conjunta de N y X 39

I.8 Aproximaciones 40
Aproximaciones binomiales 40
Aproximación usual 41
Aproximación de Sandiford 44
Aproximación de Burr 47
Aproximación de Ord 47
Aproximación de Wise 49
Aproximación de Bennett 49
Aproximación de Molenaar 50
Comparación entre la distribución hipergeométrica y la binomial aproximante. Resultados de Uhlmann 51
Aproximaciones por la distribución de Poisson 58
Aproximación usual 59
Aproximación de Burr 60
Aproximación de Molenaar 60
Aproximaciones normales 61
Aproximación usual 62
Aproximación de Nicholson 63
Aproximación de Peizer 66
Aproximación normal a distribuciones discretas fuertemente unimodales 68
Aproximación por la distribución Beta 71

CAPITULO II
TEOREMAS DE CARACTERIZACION

II.1 Introducción 74
II.2 Mixtura de distribuciones 75
Definición 75
Mixturas de una distribución hipergeométrica 77
Mixtura con una distribución binomial 77
Mixtura con una distribución hipergeométrica .. 79
Mixtura con una distribución de Polya 81

II.3 Teoremas de caracterización 84
Caracterización de Skibinsky 85
Teorema de caracterización basado en la composición con una distribución hipergeométrica 88
Teorema de caracterización basado en la composición con una distribución de Polya 92

SEGUNDA PARTE

CAPITULO III

ORDENACIONES EN DISPERSION Y ALEATORIEDAD

III.1 Introducción 98
III.2 Ordenaciones en dispersión 100
Ordenación estocástica 102
Ordenaciones en dispersión respecto de un valor central 103
Ordenación de Birnbaum 103
Ordenación en variabilidad 104
Ordenación según el número de cambios de signo 106
Ordenaciones basadas en las funciones cuantiles ... 109
Ordenación en dispersión 110
Relación entre las ordenaciones en dispersión y variabilidad 116
Relación entre la ordenación en dispersión y la propiedad de unimodalidad fuerte 117
Relación entre la ordenación en dispersión y otras ordenaciones 119
Ordenación en dispersión y distribuciones de razón de fallo monótona 122
Variabilidad condicional uniforme 123
Relación con los cambios de signo 124
Relación con la ordenación en variabilidad 126

Log-concavidad relativa 129
Ordenación en log-concavidad relativa 130
Relación entre la log-concavidad relativa y la variabilidad condicional uniforme 131
III.3 Ordenaciones en aleatoriedad 134
Mayoración de vectores 134
Schur-convexidad 136
Medidas usuales de aleatoriedad 138
Ordenación en aleatoriedad 139
Algunos resultados generales sobre aleatoriedad ... 141
Ordenación en aleatoriedad en el caso continuo 144
Relación con los cambios de signo 146
Relación entre la ordenación en aleatoriedad y en dispersión 147

CAPITULO IV
PROPIEDADES DE MONOTONIA

IV.1 Introducción 151
IV.2 Monotonía de las medidas de dispersión 157
Estudio del coeficiente de asimetría 158
Monotonía de los momentos factoriales 160
Monotonía de los momentos ordinarios 160
Monotonía de los momentos centrales 161
Monotonía en dispersión 162
Estudio del orden estocástico 162
Monotonía en log-concavidad relativa 166
Monotonía en variabilidad condicional uniforme 168
Monotonía en variabilidad 169
Monotonía de la entropía 171
IV.3 Monotonía en aleatoriedad 176
Comparación de las funciones de distribución 179

Aleatoriedad en la familia hipergeométrica con valor esperado constante y entero 186
Aleatoriedad en la familia hipergeométrica con valor esperado constante 191
Observaciones previas 191
Posición relativa de la moda respecto del intervalo de indeterminación 197
Localización de la moda en la familia Hm(mN,p,n) 201
Aleatoriedad en la familia Hm(mN,p,n), cuando la moda se localiza en los valores extremos del soporte 204
Monotonía de la probabilidad modal 207
Localización de las dos mayores probabilidades puntuales 233
Localización de las tres mayores probabilidades puntuales 237
Aleatoriedad de las distribuciones $H_m(mN,p,n)$.. 239

Bibliografía 251

CAPITULO I

LA DISTRIBUCION HIPERGEOMETRICA

I.1 INTRODUCCION

Este primer capítulo de la memoria está dedicado a recopilar los resultados más importantes existentes en la literatura sobre la distribución hipergeométrica.

El lector interesado en esta distribución dispondrá aquí de una síntesis de dichos resultados y la correspondiente orientación bibliográfica y, en algunos casos, de la demostración acabada de los mismos.

Comenzamos el capítulo definiendo la distribución hipergeométrica y a continuación se estudia la pertenencia de esta distribución a familias discretas de probabilidad. Este aspecto es importante porque algunos resultados, como, por ejemplo, el valor de la desviación media, se pueden obtener como particularización de otros, más generales, cuya validez se extiende a todos los componentes de una familia de distribuciones.

Seguidamente se estudian la funciones generatrices, los momentos y las principales propiedades de la distribución hipergeométrica. Daremos, de forma explícita, el valor de los principales momentos y el de aquellos que permiten obtener los coeficientes de asimetría, curtosis y variación.

En la siguiente sección se tratarán tres supuestos

de estimación puntual de los parámetros: estimación del parámetro tamaño poblacional, estimación del parámetro tamaño de la clase, y estimación conjunta de ambos parámetros.

El grueso de este primer capítulo está dedicado al tema de las aproximaciones. Al contrario de lo que ocurre con el tema de las estimaciones, la bibliografía, en este caso, es amplia.

Le prestaremos una especial atención a las aproximaciones de la hipergeométrica por distribuciones binomiales, y dentro de este apartado queremos destacar el trabajo de Ulhmann (1966) -que será comentado ampliamente- en el que se ofrece unos interesantes resultados, fruto de la comparación sistemática entre la distribución hipergeométrica y su binomial aproximante usual.

Estos resultados tienen para nosotros un interés relevante, puesto que a partir de ellos abordaremos, en el último capítulo de esta memoria el estudio del comportamiento monótono, en aleatoriedad, de la familia hipergeométrica en su convergencia hacia la distribución binomial.

Además de las aproximaciones binomiales se recogerán aproximaciones por las distribuciones de Poisson y Normal, y una aproximación por la distribución Beta.

Queremos, por último, indicar que el conjunto de resultados sobre la distribución hipergeométrica se amplía con

el estudio de diversos casos de mixturas de esta distribución respecto del parámetro X (tamaño de la clase). Este estudio se incluye en el segundo capítulo de la memoria, en el que aportamos nuestras caracterizaciones de la distribución hipergeométrica, basadas en el comportamiento de esta distribución cuando es compuesta, respecto del parámetro X, con otras distribuciones pertenecientes al esquema de urna de Polya.

I.2. DEFINICION

La distribución hipergeométrica es una distribución de probabilidad discreta, definida en un intervalo de los números enteros no negativos. Depende de tres parámetros: N, X, n, números naturales, siendo $N \geq X$ y $N \geq n$.

El modelo del cual surge esta distribución es el siguiente: consideremos una población de tamaño N, de la cual X componentes presentan una característica determinada y realizamos n extracciones sin reemplazamiento. La variable aleatoria x, que da el número de elementos de la muestra que presentan dicha característica, decimos que sigue una distribución hipergeométrica H(N,X,n).

Si notamos por h(k; N,X,n) a la probabilidad de que la variable x tome el valor k, tendremos que la función de masa de probabilidad será:

$$h(k;\ N,X,n) = \frac{\left(\begin{array}{c} X \\ k \end{array}\right)\left(\begin{array}{c} N-X \\ n-k \end{array}\right)}{\left(\begin{array}{c} N \\ n \end{array}\right)} ,$$

con $\max(0,X+n-N) \leq k \leq \min(X,n)$.

Llamando p a la cantidad X/N, es decir, a la probabilidad de que un elemento de la población, elegido al azar, tenga la característica dada, y llamando q a 1-p, podremos también expresar:

$$h(k;\ N,X,n) = \frac{\left(\begin{array}{c} Np \\ k \end{array}\right)\left(\begin{array}{c} Nq \\ n-k \end{array}\right)}{\left(\begin{array}{c} N \\ n \end{array}\right)} .$$

La misma distribución es obtenida si se intercambian n y X. (Basta sustituir cada número combinatorio por su expresión factorial, para comprobarlo).

Sustituyendo los números combinatorios por su expresión factorial, reordenando, y simplificando, podemos escribir h(k; N,X,n) de la forma

$$\binom{n}{k} \frac{X(X-1)\ldots(X-k+1)(N-X)(N-X-1)\ldots(N-X-n+k+1)}{N(N+1)\ldots(N-n+1)} ,$$

y notando $A^{(r,c)}=A(A+c)(A+2c)\ldots(A+(r-1)c)$,

$$h(k;\ N,X,n) = \binom{n}{k} \frac{X^{(k,-1)}(N-X)^{(n-k,-1)}}{N^{(n,-1)}} \quad .$$

Con esta forma de la función de masa de probabilidad se pone de manifiesto que la distribución hipergeométrica es un caso particular de distribución de Polya.

En efecto, la función de masa de probabilidad de una distribución de Polya de parámetros N,X,n,c, es

$$p[k;\ N,X,n,c] = \frac{\binom{n}{k} X^{(k,c)}\ (N-X)^{(n-k,c)}}{N^{(n,c)}} \quad .$$

Es decir, la distribución hipergeométrica se obtiene cuando el parámetro c de la distribución de Polya es igual a menos uno.

Cuando c=0 se obtiene la distribución binomial, cuya función de masa de probabilidad es

$$p[k;\ n,p] = \binom{n}{k} p^k q^{n-k} \quad ,$$

con $p=\frac{X}{N}$ y $q=1-p$.

El nombre de la distribución que estamos estudiando lo recibe de la función hipergeométrica.

La función hipergeométrica viene definida por:

$$F(\alpha,\beta;\gamma;x) = 1 + \frac{\alpha\beta}{\gamma}\frac{x}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{x^2}{2!} + \ldots =$$

$$= \sum_{j=0}^{\infty} \frac{\alpha^{[j]}\beta^{[j]}}{\gamma^{[j]}}\frac{x^j}{j!} \quad \text{con } \gamma>0$$

siendo $a^{[j]} = a(a+1)(a+2)\ldots(a+j-1)$; $a^{[o]}=1$.

Esta serie es convergente para $|x|<1$. En el caso que α ó β sean enteros negativos, la función hipergeométrica tendrá sólo un número finito de términos distintos de cero.

Veamos ahora cuál es la relación que liga la distribución hipergeométrica con la función del mismo nombre.

Las probabilidades puntuales de la distribución hipergeométrica coinciden con los sucesivos términos del desarrollo de

$$\frac{(N-n)!(N-X)!}{N!(N-X-n)!}\,F(-n,-X;N-X-n+1;1) \ ,$$

es decir,

$$[1.1] \qquad h(k;\ N,X,n) = \frac{(N-n)!(N-X)!}{N!(N-X-n)!}\,F_k(-n,-X;N-X-n+1;1) \ ,$$

donde con el subíndice k queremos indicar el sumando del desarrollo de la función hipergeométrica que se obtiene cuando j=k.

En efecto:

$$\frac{(N-n)!(N-X)!}{N!(N-X-n)!}\ F_k(-n,-X;N-X-n+1;1) =$$

$$= \frac{(N-n)!(N-X)!}{N!(N-X-n)!}\ \frac{n(n-1)\ldots(n-k+1)X(X-1)\ldots(X-k+1)}{(N-X-n+1)\ldots(N-X-n+k)}\ \frac{1}{k!} =$$

$$= \frac{(N-n)!(N-X)!n!X!}{N!(N-X-n+k)!(n-k)!(X-k)!k!} = \frac{\left(\begin{array}{c}X\\k\end{array}\right)\left(\begin{array}{c}N-X\\n-k\end{array}\right)}{\left(\begin{array}{c}N\\n\end{array}\right)}\ .$$

Veamos ahora una forma alternativa de obtención de la hipergeométrica: como condicional de dos binomiales.

Sean X_1 y X_2 dos variables aleatorias binomiales independientes, con parámetros respectivos (N_1,p) y (N_2,p). Sabemos que X_1+X_2 es también binomial, de parámetros (N_1+N_2,p).

Consideremos la distribución condicional $X_1/(X_1+X_2)$.

Tendremos que:

$$p(X_1=k/(X_1+X_2=X) = \frac{\left(\begin{array}{c}N_1\\k\end{array}\right) p^k\ q^{N_1-k}\ \left(\begin{array}{c}N_2\\X-k\end{array}\right) p^{X-k}\ q^{N_2-X+k}}{\left(\begin{array}{c}N_1+N_2\\X\end{array}\right) p^X\ q^{N_1+N_2-X}} =$$

$$= \frac{\left(\begin{array}{c}N_1\\k\end{array}\right)\left(\begin{array}{c}N_2\\X-k\end{array}\right)}{\left(\begin{array}{c}N_1+N_2\\X\end{array}\right)} \quad \text{con } \max(0,X-N_2)\leq k\leq \min(N_1,X),$$

que como podemos ver corresponde a una distribución hipergeométrica $H(N_1+N_2,N_1,X)$ o bien, por la posibilidad de intercambiar los parámetros N_1 y X, a una $H(N_1+N_2,X,N_1)$.

I.3 PERTENENCIA A FAMILIAS DISCRETAS DE PROBABILIDAD

Vamos a tratar en este apartado sobre la pertenencia de la distribución hipergeométrica a las principales familias discretas de distribuciones de probabilidad.

SISTEMA DE PEARSON DISCRETO

La distribución hipergeométrica pertenece a la primera familia importante de distribuciones discretas; se trata del Sistema de Pearson discreto, al cual tambien pertenecen otras distribuciones conspicuas, como la binomial, binomial negativa, hipergeométrica negativa, y la de Poisson. El sistema de Pearson discreto está formado por aquellas distribuciones discretas que verifican la siguiente ecuación en diferencias finitas:

$$\Delta P_{j-1} = \frac{(a-j)P_{j-1}}{b_0+b_1 j+b_2 j(j-1)}$$

donde a, b_0, b_1, b_2, representan constantes; P_j, a la probabilidad de que la variable tome el valor j, y ΔP_{j-1} a la diferencia P_j-P_{j-1}.

Vamos a obtener el valor de las constantes antes

citadas, para la distribución hipergeométrica. Consideremos la relación de recurrencia que se verifica para las probabilidades puntuales de la hipergeométrica

$$[1.2] \qquad h_j = \frac{(X-j+1)(n-j+1)}{j(N-X-n+j)}\, h_{j-1} \ .$$

De aquí se tiene que

$$\Delta h_{j-1} = h_j - h_{j-1} = \left[\frac{(X-j+1)(n-j+1)}{j(N-X-n+j)} - 1\right] h_{j-1} \ .$$

Operando la expresión entre corchetes, se obtiene

$$\Delta h_{j-1} = \frac{\dfrac{(X+1)(n+1)}{N+2} - j}{\dfrac{N-X-n}{N+2}\, j + \dfrac{1}{N+2}\, j^2}\, h_{j-1} \ .$$

Si identificamos esta última expresión con la de ΔP_{j-1}, obtenemos que para la hipergeométrica,

$$a = \frac{(X+1)(n+1)}{N+2} \ ;$$

$$b_o = 0 \ ; \qquad b_2 = \frac{1}{N+2} \ ; \qquad b_1 = \frac{N-X-n+1}{N+2} \ .$$

Hacemos notar que si se verificase la ecuación reducida en diferencias finitas

$$\Delta P_{j-1} = \frac{(a-j)}{b_1 j}\, P_{j-1} \ ,$$

ello implicaría la pertenencia a otra de las familias importantes de distribuciones, la clase de distribuciones de series de potencias generalizadas. En la hipergeométrica esto no se cumple, ya que para que así fuese sería necesario que $b_2=1/(N+2)$ fuese cero.

Ord (1967) establece una clasificación para las diferentes distribuciones de esta familia, utilizando como criterio diferenciador los valores de

$$k = \frac{(b_1-b_2-1)^2}{4b_2(a+b_o)}$$

y del índice de dispersión I = cociente varianza/media.

La distribución hipergeométrica pertenece al Tipo I(a), con el criterio caracterizador de $I<1$ y $k>1$.

La clasificación completa, así como un estudio más detallado de esta familia de distribuciones puede verse en el libro de Ord (1972).

DISTRIBUCIONES FUERTEMENTE UNIMODALES

Una distribución discreta $\{p_n\}$, definida sobre los enteros, es unimodal si existe al menos un entero r, para el que se verifica:

$$p_{n-1} \leq p_n \quad \text{si} \quad n \leq r, \quad \text{y } p_{n+1} \leq p_n \quad \text{si} \quad n \geq r \ .$$

Una distribución discreta es fuertemente unimodal si la convolución con cualquier distribución unimodal es también unimodal.

La condición de ser fuertemente unimodal implica ser unimodal, ya que bastaría efectuar la convolución con la distribución que tiene toda la masa en el cero, y que evidentemente es unimodal, para obtener la propia distribución, que sería, por tanto, unimodal.

Según podemos ver en el trabajo de Keilson y Gerber (1971), una condición necesaria y suficiente para que una distribución discreta $\{p_n\}$ sea fuertemente unimodal es que

$$p_n^2 \geq p_{n+1} \, p_{n-1} \qquad \forall n \ .$$

Si acudimos de nuevo a la relación de recurrencia [1.2], entre las probabilidades puntuales, y obtenemos las expresiones correspondientes a h_{k+1} y h_{k-1} en función de h_k, tendremos que

$$h_{k+1}h_{k-1} = \frac{(X-k)(n-k)}{(k+1)(N-X-n+k+1)} \, \frac{k(N-X-n+k)}{(X-k+1)(n-k+1)} \, h_k^2 \ .$$

Para demostrar que $h_{k+1} \cdot h_{k-1}$ es menor o igual que h_k^2, bastará con comprobar que el factor multiplicador de h_k^2 es menor o igual que 1, para cualquier k:

$$\frac{X-k}{X-k+1}\,\frac{n-k}{n-k+1}\,\frac{k}{k+1}\,\frac{N-X-n+k}{N-X-n+k+1} \leq 1, \quad \forall k,$$

lo cual es claro, teniendo en cuenta que los cuatro denominadores son necesariamente enteros mayores o iguales que 1, ya que $\max(0,X+n-N) \leq k \leq \min(n,X)$.

De esta manera queda probada la pertenencia de la distribución hipergeométrica a la clase de distribuciones fuertemente unimodales y, en consecuencia, a la de las distribuciones unimodales.

Según el resultado de Ibragimov (1956), recogido en el artículo de Keilson y Gerber (1971), la condición necesaria y suficiente para el carácter fuertemente unimodal de una distribución continua es que la función de densidad sea log-cóncava. Precisamente, esta condición de log-concavidad la traducen Keilson y Gerber para el caso discreto por:

$$p_k^2 \geq p_{k-1}\, p_{K+1} \quad \forall k.$$

DISTRIBUCIONES DE RAZON DE VEROSIMILITUD MONOTONA

Una familia uniparamétrica $f_r(\omega)$, con ω variando en un intervalo, se dice que tiene razón de verosimilitud monótona respecto de un cierto estadístico $T=t(x_1,x_2,\ldots,x_n)$ si el cociente de verosimilitud

$$\frac{L(x_1, x_2, \ldots, x_n;\ \omega_1)}{L(x_1, x_2, \ldots, x_n;\ \omega_2)} \qquad \omega_1 < \omega_2$$

es una función monótona de T. [Mood, Graybill y Boes (1974)]

Veamos que la hipergeométrica es distribución de razón de verosimilitud monótona.

Para ello, consideremos $f_k(\omega)=h(k;N,\omega,n)$, es decir, suponemos n y N conocidos, y tomamos a X como parámetro. Por otra parte, tomaremos T=k.

La distribución hipergeométrica es de razón de verosimilitud monótona en k. En efecto, si calculamos el cociente,

$$\frac{h(k;\ N,X,n)}{h(k;\ N,X+1,n)} = \frac{(X-k+1)(N-X)}{(X+1)(N-X-n+k)} \quad .$$

Puede comprobarse fácilmente que este cociente es una función decreciente en k, por lo que

$$\frac{h(k;\ N,X_1,n)}{h(k;\ N,X_2,n)} \qquad \text{para} \quad X_1 < X_2,$$

será también decreciente en k, ya que podemos expresarlo como un producto de tales funciones:

$$\frac{h(k;\ N,X_1,n)}{h(k;\ N,X_2,n)} = \prod_i \frac{h(k;\ N,X_1+i-1,n)}{h(k;\ N,X_1+i,n)} \quad ,$$

con i variando desde cero hasta X_2-X_1+1.

Este resultado lo podemos ver en el libro de Ferguson (1967), p. 209.

DISTRIBUCIONES DE RAZON DE FALLO CRECIENTE

Una distribución discreta es de razón de fallo creciente si la razón de fallo r(j), definida por

$$r(j) = \frac{P[x=j]}{P[x \geq j]} \quad ,$$

es una función no decreciente.

En el libro de Barlow y Proschan (1975), pg. 76, se define:

Una función h(x), con $-\infty < x < \infty$, es PF_2 (frecuencia de Polya de orden 2), si

(i) $h(x) \geq 0$ para $-\infty < x \infty$, y

(ii) log h(x) es cóncava en $(-\infty, \infty)$.

Las funciones PF_2 son un caso particular de funciones de positividad total.

En el citado libro se demuestra, a continuación, que si una función de densidad es PF_2, entonces la correspondiente función de distribución verifica la condición de razón de fallo creciente.

En conclusión, puesto que ya hemos mencionado la equivalencia entre el carácter fuertemente unimodal de una distribución y la log-concavidad de su función de masa de probabilidad, podemos asegurar la pertenencia de la distribución hipergeométrica a la clase de distribuciones de razón de fallo creciente.

DISTRIBUCIONES "NUEVO MEJOR QUE USADO"

Una distribución con función de distribución F(x) es de la clase "nuevo mejor que usado" si

$$1-F(x+y) \leq (1-F(x))(1-F(y)) \qquad \forall\ x,y \geq 0.$$

Estas distribuciones son usadas en Teoría de la fiabilidad.

Barlow y Proschan (1975), pg. 159, demuestran que toda distribución de razón de fallo creciente es también de la clase "nuevo mejor que usado", por lo que la hipergeométrica queda asimismo adscrita a esta familia de distribuciones.

Para finalizar esta sección queremos mencionar la no pertenencia de la hipergeométrica a dos clases importantes de distribuciones: familia exponencial y distribuciones de serie de potencias generalizadas. Vease Patel, Kapadia y Owen (1976)

Propiedades, resultados y referencias de originales sobre familias de distribuciones pueden encontrarse en el mencionado libro de Patel, Kapadia y Owen (1976).

I.4. FUNCIONES GENERATRICES

FUNCION GENERATRIZ DE PROBABILIDADES

$$G(z) = \sum_k h_k \, z^k \ .$$

Vamos a obtenerla sustituyendo las probabilidades puntuales h_k, por su expresión [1.1]:

$$G(z) = \sum_k \frac{(N-n)!(N-X)!}{N!(N-X-n)!} \frac{(-n)^{[k]}(-X)^{[k]}}{(N-X-n+1)^{[k]}} \frac{1}{k!} z^k =$$

$$= \frac{(N-n)!(N-X)!}{N!(N-X-n)!} \sum_k \frac{(-n)^{[k]}(-X)^{[k]}}{(N-X-n+1)^{[k]}} \frac{z^k}{k!} =$$

$$= \frac{(N-n)!(N-X)!}{N!(N-X-n)!} F(-n,-X;N-X-n+1;z) \ .$$

La función generatriz de momentos y la función característica se obtienen sustituyendo z por e^t y e^{it}, respectivamente.

FUNCION GENERATRIZ DE MOMENTOS

$$M(t) = \frac{(N-n)!(N-X)!}{N!(N-X-n)!} F(-n,-X;N-X-n+1;e^{t}).$$

Otra forma de la función generatriz de momentos fue dada por Lessing (1973), pg. 115:

$$M(t) = \frac{(N-n)}{n!} \frac{\partial^{(n)}}{\partial_{y}^{(n)}} \left[(1+ye^{t})^{x}(1+y)^{N-x}\right]\Big]_{y=0} .$$

FUNCION CARACTERISTICA

$$\Phi(t) = \frac{(N-n)!(N-X)!}{N!(N-X-n)!} F(-n,-X;N-X-n+1;e^{it}) .$$

I.5. MOMENTOS

MOMENTOS FACTORIALES

El momento factorial de orden r es

$$\mu(r) = E\left[x^{(r)}\right], \text{ donde } x^{(r)}=x(x-1)\ldots(x-r+1) \ .$$

Para la hipergeométrica H(N,X,n), tendremos

$$\mu(r) = \sum_k k^{(r)} h_k = \sum_k k^{(r)} \frac{\left(\begin{array}{c} X \\ k \end{array}\right)\left(\begin{array}{c} N-X \\ n-k \end{array}\right)}{\left(\begin{array}{c} N \\ n \end{array}\right)} \ .$$

Naturalmente, el número de términos de la sumatoria, distintos de cero, depende del valor de los parámetros, ya que $\max(0,n+X-N)\leq k\leq \min(n,X)$, y, por otra parte, k debe ser mayor o igual que r, puesto que en caso contrario, $k^{(r)}=0$.

Utilizando la expresión factorial de los números combinatorios, y teniendo en cuenta que $a!=a^{(r)}(a-r)!$, podemos escribir:

$$\mu(r) = \sum_k k^{(r)} \left(\begin{array}{c} N-X \\ n-k \end{array}\right) \frac{x^{(r)}(x-r)!n^{(r)}(n-r)!(N-n)!}{k^{(r)}(k-r)!N^{(r)}(N-r)!(X-k)!} =$$

$$= \frac{X^{(r)}n^{(r)}}{N^{(r)}} \sum_k \frac{(X-r)!}{(k-r)!(X-k)!} \frac{(n-r)!(N-n)!}{(N-r)!} \binom{N-X}{n-k} =$$

$$= \frac{X^{(r)}n^{(r)}}{N^{(r)}} \sum_k \frac{\left(\begin{array}{c} X-r \\ k-r \end{array}\right)\left(\begin{array}{c} N-X \\ n-k \end{array}\right)}{\left(\begin{array}{c} N-r \\ n-r \end{array}\right)} \ .$$

Como, por el teorema de convolución de Vandermonde, la sumatoria es igual a uno, tendremos que

$$\mu(r) = \frac{X^{(r)} n^{(r)}}{N^{(r)}} \quad .$$

Como consecuencia inmediata podemos obtener los valores de la media y de la varianza para la hipergeométrica.

$$\text{Media } \mu = \mu(1) = \frac{X\ n}{N} = np \quad .$$

El momento factorial de orden 2,

$$\mu(2) = \frac{X(X-1)n(n-1)}{N(N-1)} = np\ \frac{(X-1)(n-1)}{N-1},$$

nos permite calcular la varianza:

$$\text{Varianza } \sigma^2 = \mu(2)+\mu(1)-\mu^2(1) = npq\ \frac{N-n}{N-1} \quad .$$

MOMENTOS CENTRALES

El momento central de orden r es

$$m_r = E[\,(x-\mu)^r\,] \quad .$$

Pearson (1924) considera la función hipergeométrica $F(-n,-X;N-X-n+1;e^y)$.

Llamando

$$\alpha=-n \ , \qquad \beta=-X \ , \qquad \gamma=N-X-n+1 \ ,$$

$$F_1(\alpha,\beta;\gamma;e^y) = \frac{(N-n)!\,(N-X)!}{N!\,(N-X-n)!}\, F(\alpha,\beta;\gamma;e^y)$$

es la función generatriz de momentos, y

$$e^{-npy}\ F_1(\alpha,\beta;\gamma;e^y),$$

sería la función generatriz de momentos centrales, por lo que

$$m_s = \frac{d^s}{dy^s}\,[e^{-npy}\ F_1(\alpha,\beta;\gamma;e^y)]_{y=o} \ .$$

A partir de aquí, y haciendo uso de que $F_1(\alpha,\beta;\gamma;x)$ satisface la ecuación diferencial de Forsyth

$$x(1-x)\ \frac{d^2F}{dx^2} + [\gamma-(\alpha+\beta+1)x]\ \frac{dF}{dx} - \alpha\beta F = 0 \ ,$$

tomando $x=e^y$, y trás un proceso operativo, Pearson llega a la siguiente relación en diferencias finitas, que se verifica entre los momentos centrales :

$$Nm_{r+1} = [(1+E)^r-E^r]\ [m_2-(X+n-2pn)m_1+pqn(N-n)m_o] \ ,$$

donde $m_o=1$ y $m_1=0$.

E representa el operador desplazamiento, cuyo comportamiento es:

$$Em_r = m_{r+1} \; ; \qquad E^s m_r = m_{r+s} \; .$$

Aplicando esta relación podemos calcular los momentos m_3 y m_4, obteniendo:

$$m_3 = \frac{pqn(N-n)}{(N-2)(N-1)} \; (1-2p)(N-2n);$$

$$m_4 = \frac{npq(N-n)}{(N-1)(N-2)(N-3)} \; A, \qquad \text{con}$$

$$A = N(N+1)-6n(N-n)+3pq(N^2(n-2)-Nn^2+6n(N-n)).$$

Como ya conocemos la media, la varianza y los momentos m_3 y m_4, podemos obtener los valores de los coeficientes de asimetría, curtosis y variación:

Coeficiente de asimetría:

$$\sqrt{\beta_1} = \frac{m_3}{m_2^{3/2}} = \frac{q-p}{\sqrt{npq\,\frac{N-n}{N-1}}} \; \frac{N-2n}{N-2} \; .$$

Coeficiente de curtosis:

$$\beta_2 = \frac{m_4}{m_2^2} = \frac{3(N-1)(N+6)}{(N-2)(N-3)} + \frac{(N-1)N(N+1)}{(N-n)(N-2)(N-3)}\left[1-\frac{6N}{N+1}\left[pq+\frac{n(N-n)}{N^2}\right]\right]\frac{1}{npq} \; .$$

Coeficiente de variación:

$$\frac{\sigma}{\mu} = \left[\frac{(N-X)(N-n)}{nX(N-1)}\right]^{1/2} \; .$$

MOMENTOS RESPECTO DEL ORIGEN

$$\mu_r = E(x^r) \quad .$$

Para el cálculo de los momentos respecto del origen puede utilizarse las relaciones que permiten obtenerlos a partir de los momentos centrales, o bien, a partir de los factoriales.

En el primer caso:

$$\mu_r = \sum_{j=0}^{r} \binom{r}{j} \mu_1^j \, m_{r-j} \quad .$$

En el segundo caso:

$$\mu_r = \sum_{j=0}^{r} \frac{\Delta^j O^r}{j!} \mu(j) \quad .$$

Con $\Delta^j O^r$ se quiere indicar $\Delta^j x^r / x=0$, siendo Δ el operador "diferencia hacia adelante",

$$\Delta f = f(x+1) - f(x).$$

Simbólicamente podríamos expresar: $\Delta = E-1$, siendo E es operador desplazamiento, que hemos visto con anterioridad.

Los coeficientes $\Delta^j O^r / j!$ reciben el nombre de números de Stirling de 2° tipo, y suelen notarse S_r^j.

OTROS TIPOS DE MOMENTOS

DESVIACION MEDIA

El primer momento central absoluto, o desviación media, está definido por

$$\nu_1 = E|x-\mu| \quad .$$

La desviación media para la hipergeométrica se obtiene a partir de la generalización que Kamat (1966) realizó de la propiedad de Johnson; propiedad ésta que permite el cálculo de la desviación media para algunas distribuciones del Sistema de Pearson discreto.

Para cualquier distribución perteneciente a esta familia, con la condición de que el primer valor que tome la variable sea el cero, y que la constante b_o sea asimismo cero -condiciones que sabemos cumple la hipergeométrica-, se verifica:

$$[1.3] \qquad (1-2b_2)\nu_1 = 2f_m[mb_1+(\mu-m)(1-2b_2)+m(m-1)b_2] \quad ,$$

siendo m la parte entera de la media y f_m la probabilidad de que la variable tome dicho valor.

Esta expresión es la particularización de una más

compleja, cuando se le impone las dos condiciones restrictivas arriba citadas.

Aplicando [1.3] y sustituyendo b_1 y b_2 por los valores que estas constantes toman para la hipergeométrica, y llamando, finalmente, $\delta(np)$ a (np-m) se obtiene:

$$\nu_1 = 2 \frac{\binom{X}{m}\binom{N-X}{n-m}}{\binom{N}{n}} N^{-1} [m(N-X-n+m)+n\delta(np)] .$$

Si np es entero, la desviación media tiene la expresión simplificada

$$\nu_1 = 2\sigma^2 h_m(1-N^{-1}) .$$

MOMENTOS INVERSOS FACTORIALES

Los momentos inversos factoriales, definidos como

$$E\left[\frac{1}{(x+r)^{(r)}}\right] \text{ó, equivalentemente, } E\left[\frac{1}{(x+1)^{[r]}}\right],$$

se pueden calcular por integracion sucesiva de la funcion generatriz de probabilidades, entre cero y uno, como límites de integración.

Así, el primer momento inverso factorial es

$$\mu_{-[1]} = \int_0^1 G(z)\,dz = \frac{(N-n)!(N-X)!}{N!(N-X-n)!} \int_0^1 F(-n,-X;N-X-n+1;z)\,dz\ .$$

Efectuando la integral se obtiene, para ésta, la siguientes expresion:

$$\frac{(N-X-n)}{(n+1)(X+1)} \sum_{k=1} F_k(-n-1,-X-1;N-X-n;1),$$

donde con F_k se indica el sumando que ocupa el lugar k+1 en el desarrollo de la función hipergeométrica F(-n-1,-X-1;N-X-n;1).

Teniendo en cuenta que

$$h(k;N,X,n) = \frac{(N-n)!(N-X)!}{N!(N-X-n)!} F_k(-n,-X;N-X-n+1;1),$$

se obtiene fácilmente la expresión:

$$\mu_{-[1]} = \frac{N+1}{(n+1)(X+1)} \sum_{k=1} h(k;N+1,X+1,n+1),$$

o lo que es igual,

$$\mu_{-[1]} = \frac{N+1}{(n+1)(X+1)} [1-h(O;N+1,X+1,n+1)].$$

El segundo momento inverso factorial es:

$$\mu_{-[2]} = \int_0^1 dz_1 \int_0^{z_1} G(z)\,dz.$$

Siguiendo el mismo razonamiento que para el primer momento inverso factorial, se obtiene la siguiente expresión:

$$\mu_{-[2]} = \frac{(N+1)^{[2]}}{(n+1)^{[2]}(X+1)^{[2]}} \left[1 - \sum_{k=0} h(k;N+2,X+2,n+2)\right] .$$

En general, para el momento inverso factorial de orden r, se obtiene:

$$\mu_{-[r]} = \frac{(N+1)^{[r]}}{(n+1)^{[r]}(X+1)^{[r]}} \left[1 - \sum_{k=0}^{r-1} h(k;N+r,X+r,n+r)\right] .$$

De aquí, puede obtenerse fácilmente la relación de recurrencia,

$$\mu_{-[r]} = \frac{N+r}{(n+r)(X+r)} \frac{1 - \sum_{k=0}^{r-1} h(k;N+r,X+r,n+r)}{1 - \sum_{k=0}^{r-2} h(k;N+r-1,X+r-1,n+r-1)} \mu_{-[r-1]}$$

entre los momentos inversos factoriales.

MOMENTOS NEGATIVOS RESPECTO DEL ORIGEN

Para poder considerar los momentos negativos respecto del origen, $E(x^{-r})$, es necesrio excluir de la variable el valor observado cero. La variable truncada que se obtiene se suele llamar hipergeométrica positiva.

Su función de probabilidad será:

$$h_k^+ = \frac{\binom{X}{k}\binom{N-X}{n-k}}{\binom{N}{n}-\binom{N-X}{n}} = \frac{\binom{x}{k}\binom{N-X}{n-k}}{\binom{N}{n}\left[1-\frac{\binom{N-X}{n}}{\binom{N}{n}}\right]} = \frac{h_k}{1-h_o} \ .$$

Para estos momentos no existe ninguna fórmula breve que proporcione su valor.

Uno de los autores que se ha dedicado a su estudio es Stephan (1945), quien los calcula a partir de un desarrollo en serie de x^{-a}.

Veamos cómo se haría el cálculo para E(1/x). Es facil probar, por induccion, que para x>0

$$\frac{1}{x} = \sum_{i=i}^{t} \frac{(i-1)!x!}{(x+i)!} + R_t(x) \ ,$$

siendo el resto $R_t(x) = \frac{t!(x-1)!}{(x+t)!}$.

Tendremos entonces que

$$E(1/x) = \sum_x \frac{1}{x} h_x^+ = \sum_{i=1}^{t} u_i + E(R_t(x)) \ ,$$

donde

$$u_i = \sum_x \frac{(i-1)!x!}{(x+i)!} h_x^+ = \sum_x \frac{(i-1)!x!}{(x+i)!} \frac{h_x}{1-h_o} \ ;$$

$$E(R_t(x)) = \sum_x \frac{t!(x-1)!}{(x+t)!} \frac{h_x}{1-h_o} \quad .$$

Fijemosnos en que

$$u_1 = \frac{\mu_{[-1]}}{1-h_o} \quad .$$

A partir de aquí, al igual que con los momentos inversos factoriales, puede establecerse una relación de recurrencia para los u_i.

E(1/x) puede entonces aproximarse empleando el número de términos que se desee del desarrollo en serie, el cual converge muy rápidamente..

Unas cotas para $E(R_t(x))$ vienen dadas en el artículo de Stephan, así como el desarrollo en serie de $E(x^{-a})$ que permite el cómputo de los momentos negativos de orden superior.

I.6. PRINCIPALES PROPIEDADES

Vamos a enunciar aquí una serie de propiedades y relaciones de recurrencia, de fácil demostración. Trataremos también, de la localización de la moda.

Notaremos $h_k=h(k;N,X,n)$, como hasta ahora, a la función de masa de probabilidad, y $H_k=H(k;N,X,n)$ a la función de distribución:

[1.4] a) $h(k+1;N,X,n) = \frac{(X-k)(n-k)}{(k+1)(N-X-n+k+1)} h(k;N,X,n)$

b) $h(k;N,X+1,n) = \frac{(X+1)(N-X-n+k)}{(X-k+1)(N-X)} h(k;N,X,n)$

c) $h(k;N+1,X,n) = \frac{(N+1-n)(N+1-X)}{(N+1-n-X+k)(N+1)} h(k;N,X,n)$

d) $h(k;N,X,n+1) = \frac{(N-X-n+k)(n+1)}{(n+1-k)(N-n)} h(k;N,X,n)$

[Johnson y Kotz (1969)]

e) $h(k;N,X,n) = h(k;N,n,X) = h(X-k;N,X,N-n) = h(n-k;N,N-X,n) =$

$= h(N-n-X+k;N,N-X,N-n)$

f) $H(k;N,X,n) = H(k;N,n,X) = 1-H(n-k-1;N,N-X,n) =$

$= 1-H(X-k-1;N,X,N-n) = H(N-n-X+k;N,N-X,N-n)$

[Lieberman y Owen (1961)]

Estas propiedades se demuestran fácilmente sustituyendo las probabilidades puntuales por su expresion factorial. Las igualdades relativas a la función de distribución se obtienen a partir de las correspondientes igualdades para la función de masa de probabilidad.

g) La hipergeométrica es simétrica si p=X/N=1/2, o, por poder intercambiarse los papeles de los parámetros X y n, si n/N=1/2.

Basta ver que

$$h_{np-k} = h_{np+k} \qquad \forall\ k:0,1,2,\ldots$$

En efecto, si p=1/2=X/N, entonces X=N-X. Por otra parte,

$$np-k = \frac{n}{2} - k = n - \frac{n}{2} - k = n - np - k, \quad y$$

$$np + k = \frac{n}{2} + k = n - \frac{n}{2} + k = n - np + k\ .$$

A partir de aquí tendremos las siguientes igualdades:

$$\binom{X}{np-k} = \binom{N-X}{n-np-k}; \qquad \binom{X}{np+k} = \binom{N-X}{n-np+k}$$

por lo que, ∀k,

$$h_{np-k}=\frac{\binom{X}{np-k}\binom{N-X}{n-np-k}}{\binom{N}{n}} = \frac{\binom{X}{np+k}\binom{N-X}{n-np+k}}{\binom{N}{n}}=h_{np+k}$$

quedando, por tanto, demostrada esta propiedad.

h) Localización de la moda.

Consideremos la relación de recurrencia [1.4]. A partir de ella tendremos que

$$h_{k+1} > h_k \Leftrightarrow \frac{(X-k)(n-k)}{(k+1)(N-X-n+k+1)} > 1 \ .$$

Operando en esta desigualdad se obtiene que

$$h_{k+1} > h_k \Leftrightarrow k < \frac{(n+1)(X+1)}{N+2} - 1 \ .$$

Igualmente, podemos obtener que

$$h_{k+1} = h_k \Leftrightarrow k = \frac{(n+1)(X+1)}{N+2} - 1 \ .$$

Concluimos, por tanto, que si $(n+1)(X+1)/(N+2)$ no es entero, la moda se localizará en el mayor entero menor que dicha cantidad; y en el caso que $(N+1)(X+1)/(N+2)$ sea entero, habrá dos valores modales:

$$\frac{(n+1)(X+1)}{N+2} \quad \text{y} \quad \frac{(n+1)(X+1)}{N+2} - 1 \ .$$

Por otra parte, la moda se localizará en la media cuando ésta sea entero, y en uno de los enteros adyacentes, cuando no lo sea.

Para comprobarlo basta demostrar que

$$0 < \frac{(n+1)(X+1)}{N+2} - \frac{nX}{N} < 1 \quad .$$

Demostrar la desigualdad de la izquierda equivale a demostrar que $2nX<N+Nn+NX$, lo cual es evidente teniendo en cuenta que n y X han de ser menores o iguales que N.

Demostrar la desigualdad de la derecha es equivalente a demostrar que $X+n-N<2n\frac{X}{N}+1$.

Si X+n-N es menor o igual que cero, se cumple, evidentemente; en caso contrario, el primer valor que toma la variable es precisamente X+n-N, por lo que, teniendo en cuenta que la media, nX/N, habrá de ser mayor o igual que dicho valor, la desigualdad quedaría demostrada.

I.7. ESTIMACION DE LOS PARAMETROS

Vamos a tratar en esta sección de tres supuestos de estimación puntual de los parámetros de la hipergeométrica : estimación del parámetro X, del parámetro N, y estimación conjunta de N y X.

ESTIMACION DEL PARAMETRO X

El estimar X, con N y n conocidos, es el caso más usual.

El estimador de máxima verosimilitud para X es el entero $\hat{X}$ que hace máxima la verosimilitud h(x;N,X,n), para el valor observado x.

Hagamos uso de la relación de recurrencia,

$$h(x;N,X+1,n) = \frac{(X+1)(N-X-n+x)}{(X-x+1)(N-X)}\, h(x;N,X,n).$$

Comparando el factor multiplicador de h(x;N,X,n) con la unidad, se puede obtener:

$$h(x;N,X+1,n) > h(x;N,X,n) \quad \Leftrightarrow \quad X < x(N+1)/n - 1$$

$$h(x;N,X+1,n) = h(x;N,X,n) \quad \Leftrightarrow \quad X = x(N+1)/n - 1$$

es decir, h(x;N,X,n) es creciente en X, alcanzando su máximo en el mayor entero que no excede a x(N+1)/n. Este será, por tanto, el estimador de máxima verosimilitud de X.

En el caso que x(N+1)/n sea entero, entonces,

$$\frac{x(N+1)}{n} - 1 \qquad \text{y} \qquad \frac{x(N+1)}{n}$$

son ambos estimadores de máxima verosimilitud de X.

Reseñamos que la varianza de x(N+1)/n es

$$\frac{(N+1)^2}{n} pq \frac{N-n}{N-1} \quad .$$

ESTIMACION DEL PARAMETRO N

Este caso se presenta ocasionalmente. Consideremos, por ejemplo, la estimación del número N de peces de un estanque. En primer lugar, son cogidos un número dado X de peces; son marcados y devueltos al estanque. Poco tiempo después, suficiente como para garantizar una dispersión aleatoria de los peces marcados, pero no tan largo como para que los cambios naturales afecten demasiado al tamaño de la población, se toma una muestra de tamaño n y se ve el número x de peces marcados que aparecen.

Veamos cuál sería el estimador de máxima verosimilitud para N.

Hagamos uso de la relación de recurrencia

$$h(x;N+1,X,n) = \frac{(N+1-n)(N+1-X)}{(N+1-n-X+x)(N+1)} h(x;N,X,n) \quad .$$

Comparando el factor multiplicador de h(x;N,X,n) con la unidad, llegaríamos a que

$$h(x;N+1,X,n) > h(x;N,X,n) \iff N < nX/x - 1$$

$$h(x;N+1,X,n) = h(x;N,X,n) \iff N = nX/x - 1$$

Por tanto, h(x;N,X,n) es creciente en N, alcanzando su máximo en el mayor entero que no excede a nX/x. Este máximo, $\hat{N}$, será el estimador de máxima verosimilitud de N.

En el caso que nX/x fuese entero,

$$\frac{nX}{x} - 1 \qquad y \qquad \frac{nX}{x}$$

serían ambos estimadores de máxima verosimilitud.

La varianza de nX/x es infinita si, como es habitual, $n \leq N-X$, es decir, si la variable toma el valor cero. No obstante, podemos hacernos alguna idea de la variabilidad, considerando la distribución hipergeométrica positiva.

Si x=0, cualquier valor arbitrario, por ejemplo, 2nX, (lo que equivale a considerar la media igual a 1/2), puede ser asignado como estimador.

Chapman (1951) aconseja el uso de otro estimador de N. Este autor, citado en el libro de Johnson y Kotz (1969), sugiere tomar

$$N' = \frac{(n+1)(X+1)}{x+1} - 1 \ .$$

Veamos que este estimador es aproximadamente

centrado. En efecto, si consideramos la expresión del primer momento inverso factorial,

$$E\left[\frac{1}{x+1}\right] = \frac{N+1}{(n+1)(X+1)}\left[1 - \frac{(N-n)!(N-X)!}{(N+1)!(N-n-X-1)!}\right],$$

el término entre corchetes del segundo miembro, es despreciable para valores de n/N y X/N suficientemente grandes, por lo que

$$E\left(\frac{1}{x+1}\right) \simeq \frac{N+1}{(n+1)(X+1)} \quad .$$

Si calculamos el valor esperado del estimador N',

$$E(N') = (n+1)(X+1)E\left(\frac{1}{x+1}\right) - 1 \simeq N \quad .$$

Utilizando el desarrollo en serie de factoriales

$$\frac{1}{x} = \sum_{i=1}^{t} \frac{(i-1)!x!}{(x+i)!} + R_t(x) \qquad \text{con} \qquad R_t(x) = \frac{t!(x-1)!}{(X+t)!}$$

que ya vimos en el cálculo de los momentos negativos respecto del origen, se obtiene que, para nX/N>10,

$$E(nX/x) \geq N\left[1 + \frac{N}{nX} + 2\left(\frac{N}{nX}\right)^2\right] \quad ,$$

por lo que el estimador nX/x tiene un pronunciado sesgo positivo, mientras que el estimador N' es, como ya hemos

visto, aproximadamente centrado.

Calcula, asimismo, la varianza de N', que es, aproximadamente,

$$N^2\left[(nX/N)^{-1} + 2(nX/N)^{-2} + 6(nX/N)^{-3}\right],$$

valor que toma también, aproximadamente, el error medio cuadrático, $E[(N'-N)^2]$, si n es lo suficientemente grande.

El error medio cuadrático de $\hat{N}$ (con $\hat{N}=2nX$, si $x=0$) es, aproximadamente,

$$N^2\left[3(nX/N)^{-1} + 6(nX/N)^{-2}\right] .$$

Esto hace que la comparación entre ambos estimadores sea favorable a N'.

Para n suficientemente grande, también podríamos tomar como estimador

$$N'' = \frac{(n+2)(X+2)}{x+2}$$

ya que

$$E(N'') \simeq N\left(1- \frac{N}{nX}\right)$$

$$\sigma^2(N'') \simeq N^2\left[(nX/N)^{-1} - (nX/N)^{-2} - (nX/N)^{-3}\right]$$

lo cual, indica que este estimador será al menos tan bueno como N'.

Johnson y Kotz (1969) indican que Pearson obtuvo, en 1989, la siguiente fórmula para N en términos de momentos:

$$N = \frac{m_5 m_3 - 10 m_4 m_2^2 + 6 m_3^3 + 2 m_2^4}{m_5 m_3 - 4 m_4 m_2^2 + 3 m_3^3 + 2 m_2^4}$$

sugiriendo que esta fórmula podría ser usada para estimar N; sin embargo, no ha sido nunca utilizada en este sentido y, probablemente, daría estimaciones muy inexactas.

Wasan (1970) demuestra que no existe estimador centrado para el parametro N de la hipergeometrica.

ESTIMACION CONJUNTA DE N Y X

En este caso las ecuaciones de máxima verosimilitud son muy complicadas de resolver.Sin embargo, puede ser interesante estudiar la eficiencia de los estimadores calculados por el método de los momentos.

Midiendo la eficiencia E por

$$E = \lim_{m \to \infty} \frac{\text{Varianza (Est. Max. Ver.)}}{\text{Varianza (Otro estimador)}} ,$$

cuando son tomadas m muestras, Ord (1972) obtiene:

(i) Si X ó N-X es pequeño y N no es grande, entonces E es pequeña.

(ii) Cuando $X \simeq N-X$ y $n/N < 1/2$, entonces,

$$E^{-1} \simeq 1 + \frac{16(n-2)}{3(N-n-1)(N-1)^4} ,$$

de manera que $E \geq 0.9$.

En este caso estaría aconsejado el uso de los estimadores de momentos.

(iii) Cuando N y X son grandes, el parametro p de la binomial límite está eficientemente estimado por la media.

I.8. APROXIMACIONES

APROXIMACIONES BINOMIALES

Fijemos, en primer lugar, la notación que vamos a utilizar:

Si ξ es una variable distribuida según una hipergeométrica H(N,X,n) y η se distribuye según una binomial B(n,p), notaremos

$$h(k;N,X,n) = h_k = prob(\xi=k)$$

$$H(k;N,X,n) = H_k = prob(\xi\leq k)$$

$$b(k;n,p) = b_k = prob(\eta=k)$$

$$B(k;n,p) = B_k = prob(\eta\leq k)$$

Cuando los parámetros N y X de la distribución hipergeométrica tienden a infinito, de modo que la proporción X/N se mantiene constante, igual a p, entonces, es bien conocido que las probabilidades puntuales h(k;N,X,n) convergen a b(k;n,p), es decir,

$$\lim_{\substack{N,X\to\infty \\ X/N=p}} \frac{\binom{X}{K}\binom{N-X}{n-K}}{\binom{N}{n}} = \binom{n}{K} p^K (1-p)^{n-k} \quad .$$

Esta convergencia permite aproximar las probabilidades hipergeométricas por las binomiales, lo que conlleva un claro ahorro de cálculo, por tener éstas una expresión más simplificada y por el manejo de dos parámetros, en lugar de los tres de la hipergeométrica.

Cuando las probabilidades hipergeométricas reales se sustituyen por las binomiales, se suele decir que se ha hecho una aproximación a población infinita.

Para valores de p cercanos a 0 ó 1, la convergencia no es muy rápida, pero en general se puede considerar suficientemente buena para los fines prácticos. Suele ser adecuado el uso de esta aproximación cuando $N \leq 50$ y $n \leq N/10$.

Teniendo en cuenta que en una hipergeométrica, los papeles de los parámetros X y n se pueden intercambiar, es claro que podemos aproximar $H(k;N,X,n)$ tanto por (I) $B(k;n,\frac{X}{N})$, como por (II) $B(k;X,\frac{n}{N})$.

Veamos otras dos aproximaciones alternativas. Si acudimos a la propiedad (e) de I.6, tendremos:

$$h(k;N,X,n) = h(n-k;N,N-X,n) \quad ,$$

y por poder intercambiar el segundo y tercer parámetro,

$$h(k;N,X,n) = h(n-k;N,n,N-X) \quad ;$$

por otra parte,

$$h(k;N,X,n) = h(X-k;N,X,N-n)$$

Las aproximaciones binomiales que corresponden a estas dos formas de la hipergeométrica serán:

(III) $b(n-k;N-X,\frac{n}{N})$ y (IV) $b(X-k;N-n,\frac{X}{N})$.

Para aproximar probabilidades acumuladas, acudimos a la propiedad (f) de I.6:

$$H(k;N,X,n) = 1 - H(n-k-1;N,N-X,n) = 1 - H(n-k-1;N,n,N-X) \ ,$$

y por otra parte,

$$H(k;N,X,n) = 1 - H(X-k-1;N,X,N-n)$$

Tendremos, por tanto, que las aproximaciones (III) y (IV), para las probabilidades acumuladas, serían:

$$\text{(III')} \quad 1 - B(n-k-1;N-X,\frac{n}{N})$$

$$\text{(IV')} \quad 1 - B(X-k-1;N-n,\frac{X}{N}).$$

Lieberman y Owen (1961), sugieren que la binomial que da una mejor aproximación es aquella con más pequeño parámetro "tamaño de la muestra", es decir, aconsejan el uso de aquella binomial cuyo primer parámetro coincida con mín(n,X,N-X,N-n).

Brunk, Holstein, y Williams (1968), hacen un estudio que permite comparar las aproximaciones binomiales (I) y (II), llegando a un resultado que está en la línea de la sugerencia de Lieberman y Owen.

En el artículo de Brunk, Holstein y Williams puede verse, también, una tabla donde se compara los errores máximos

cometidos cuando las citadas binomiales son usadas como aproximaciones de la distribución hipergeométrica, para varios valores de N, X y n.

Un nuevo camino de aproximación se obtiene igualando la media y la varianza de la binomial con las correspondientes media y varianza de la hipergeométrica. Esta nueva aproximación, debida a Sandiford (1960), es más exacta que la aproximación binomial usual estudiada anteriormente. Presenta también la ventaja de ser sesgada en la misma dirección y, aproximadamente, en la misma cuantía, que la distribución exacta.

Si llamamos B(r,p') a la binomial aproximante de Sandiford, al igualar la media y la varianza con las respectivas media y varianza de la hipergeométrica H(N,X,n), tendremos

$$rp' = np = n\frac{X}{N} \quad ;$$

$$rp'(1-p') = np(1-p)\frac{N-n}{N-1} = \frac{NX(N-X)(N-n)}{N^2(N-1)} \quad .$$

Los valores de r y p' habrán de calcularse resolviendo el sistema. Se le va a imponer a r la condición de ser número entero, restricción que se sugiere para facilitar el uso de tablas, pero no es estrictamente necesaria; una interpolación, para valores fraccionarios de r, permite un mejor rendimiento de la aproximación.

Despejando p' en la primera ecuación, y sustituyendo en la segunda, obtenemos el valor de r, que será el entero más próximo a

$$\frac{nX(N-1)}{N(N-1)-(N-X)(N-n)} \quad ,$$

y con el valor de r obtenido,

$$p' = \frac{np}{r} \quad .$$

En el límite, cuando N y X tienden a infinito, con X/N=p, las ecuaciones anteriores darían como solución r=n y p'=p, es decir, la nueva aproximación y la usual son iguales en el límite.

En las ecuaciones planteadas, si n>N-n, se toma N-n en lugar de n, es decir, como muestra se coge la menor de las dos particiones en las que el muestreo divide a la población. Asimismo, si X>N-X, se tomará N-X en lugar de X.

Debemos indicar, de pasada, que r es siempre menor o igual que mín(n,X).

En el caso que no se considerase la restricción de r entero, la varianza de la binomial de Sandiford coincide, por construcción, con la de la hipergeométrica. Comparemos ahora el tercer momento central, cuyos valores para la hipergeométrica, binomial usual y binomial de Sandiford son, respectivamente,

$$npq\ \frac{N-n}{N-1}\ (1-2p)\ \frac{N-2n}{N-2}\quad ,$$

$$npq(1-2p)\quad ,$$

$$npq\ \frac{N-n}{N-1}\left[(1-2p)\ \frac{N-n}{N-1}\ -\ \frac{n-1}{N-1}\right]\quad .$$

Puede verse que si n y X son ambos mayores o iguales que N/2, el tercer momento central de las dos aproximaciones son poco próximos al de la hipergeométrica; sin embargo, si n y X son ambos menores que N/2, el momento m_3 de la binomial de Sandiford es razonablemente próximo en valor al de la hipergeométrica, y en todos los casos, más próximo que el de la binomial usual.

En su artículo, Sandiford compara numéricamente ambas aproximaciones, para algunos valores de n relativamente pequeños en relación a N, y observa la mejor calidad de su aproximación. Como además, si n/N es grande se sabe que la aproximación binomial usual no es demasiado buena, concluye que la nueva aproximación es razonablemente ajustada.

Asimismo, aplica su aproximación a la función acumulativa de probabilidad, y para un ejemplo particular obtiene una aproximación casi tan precisa como la aproximación de Wise (que estudiaremos posteriormente).

Ord (1968) hace una crítica de la aproximación de

Sandiford, ya que para p=1/2, la hipergeométrica es simétrica, y en cambio, la binomial aproximante de Sandiford no lo sería. Otro aspecto negativo es el hecho de que puede cambiar el número de puntos con probabilidad distinta de cero. Por otra parte, si p y n/N son aproximadamente del mismo tamaño, y menores que 0.25, la simetría se preserva y la aproximación de Sandiford es buena.

Otras fórmulas para la aproximación de probabilidades puntuales fueron dadas por Burr (1952) y (1973).

Las probabilidades hipergeométricas h(k;N,X,n) se aproximan por:

$$\text{(I)} \qquad b(k;n,\frac{X}{N}) \left[1 + \frac{N}{2X(N-X)} \left[k - (k-nX/N)^2\right]\right]$$

$$\text{(II)} \qquad b(k;n,\frac{X}{N}) \left[1 + \frac{1}{2X} \left[k - (nX/N)^2\right]\right] \quad .$$

El error absoluto cometido es del orden de $\frac{1}{X^2}$.

Alternativamente, intercambiando los papeles de X y n,

$$\text{(I')} \qquad b(k;X,\frac{n}{N}) \left[1 + \frac{N}{2n(N-n)} \left[k - (k-nX/N)^2\right]\right]$$

$$\text{(II')} \qquad b(k;X,\frac{n}{N}) \left[1 + \frac{1}{2n} \left[k - (nX/N)^2\right]\right] \quad .$$

En la misma línea que Burr, Ord (1968) aproxima las

probabilidades puntuales hipergeométricas h(k;N,X,n) por la binomial usual modificada por un factor corrector:

$$b(k;n,\frac{X}{N})\ \left[1 + \frac{k(1-2X/N)+n(X/N)^2-(k-nX/N)^2}{2X(1-X/N)}\right] .$$

Para las probabilidades acumulativas, H(k;N,X,n) se aproxima por:

$$B(k;n,\frac{X}{N}) - \frac{n(n-1)X}{2N(N-X)} B(\frac{k}{n}) -2B(\frac{k-1}{n-1}) + B(\frac{k-2}{n-2}),$$

siendo

$$B(k/n) = \sum_{j=0}^{k} b(j;n,\frac{X}{N}).$$

El error absoluto cometido es del orden de $1/N^2$.

Ord presenta unas tablas comparativas de las aproximaciones: binomial usual, binomial de Sandiford, y la suya propia. Utilizando incluso los datos que Sandiford manejó para probar la validez de su aproximación respecto de la binomial usual, Ord muestra que su aproximación es mucho más precisa que la de Sandiford.

Hace también una comparación con la aproximación, ya citada anteriormente, de Wise, para probabilidades acumuladas, y la presenta como ligeramente menos precisa que esta última, lo que sin embargo, quedaría compensado por ser ésta de mayor complejidad.

Hemos pospuesto el estudio de la aproximación de Wise, cronológicamente anterior a las de Sandiford y Ord, por tratarse, en primer lugar, de una aproximación sólo para las probabilidades acumuladas y luego, por ser más precisa que las anteriormente citadas.

Wise (1954) aproxima

$$H(k;N,X,n) \quad \text{por} \quad \sum_{j=0}^{k} \binom{n}{j} w^j (1-w)^{n-j} \ ,$$

donde $w = \frac{2X-k}{2N-n+1}$, es decir, se aproxima por la binomial de parámetros n y (2X-k)/(2N-n+1).

Como puede observarse, w es función de k, lo que obliga a considerar una binomial aproximante distinta para cada k. El precio de la mayor exactitud de esta aproximación es, por tanto, su mayor complejidad.

Del libro de Johnson y Kotz (1969) recogemos otra aproximación para las probabilidades acumuladas; fue construida por Bennett en 1965, en su tesis doctoral no publicada. Este autor aproxima, para $X \geq n$, las probabilidades acumuladas H(k;N,X,n) por

$$B(k;n,\frac{X}{N}) - \frac{1}{2} \frac{N(n-1)(X/N)(1-X/N)}{N-1} A,$$

con

$$A = \binom{n}{k+2} \left(\frac{X}{N}\right)^{k+2} \left(1 - \frac{X}{N}\right)^{n-k-2} - \binom{n}{k+1} \left(\frac{X}{N}\right)^{k+1} \left(1 - \frac{X}{N}\right)^{n-k-1} .$$

FACULTAD DE CIENCIAS BIBLIOTECA GRANADA

Por último, veamos la aproximación binomial más precisa. Se debe a Molenaar (1970) y es una modificación de la aproximación de Wise.

Recordamos que la binomial aproximante de Wise era B(n,w). La modificación de Molenaar consiste en tomar como segundo parámetro, w_1, siendo

$$w_1 = w + \frac{(n+1)(kw-(b-1)(1-w)) - k\,\frac{k+2}{w} + \frac{b^2-1}{1-w}}{6(2N-n+1)^2} \quad ;$$

donde b=n-k.

Una alternativa, perdiendo muy poca exactitud y simplificando sensiblemente, la propone el propio Molenaar. Consiste en tomar como segundo parámetro

$$w_2 = w + \frac{2n(nX/N - k - 0.5)}{3(2N-n+1)^2} \quad .$$

La aproximación de Molenaar, más compleja que cualquiera de las anteriores, es sorprendentemente precisa.

En Ramos (1985) hicimos un estudio numérico comparativo de cinco aproximaciones binomiales -aproximación usual, de Sandiford, de Ord, de Wise y de Molenaar-, que pone de manifiesto el grado de exactitud de esta última y, de otra parte, la sobrevaloración que hace Sandiford de su

aproximación.

Finalizamos este apartado comentando el trabajo de Uhlmann (1966). Este autor va a obtener unos interesantes resultados de la comparación sistemática entre la distribución hipergeométrica y su binomial aproximante.

Establezcamos, en primer lugar, la notación.

Notaremos

$$L_{n,c}(p) = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k};$$

y notaremos

$$L_{N,n,c}(p) = \begin{cases} 1 & \text{si } 0 \leq p \leq \frac{c}{N} \\ \sum_{k=0}^{c} \dfrac{\binom{pN}{k}\binom{N-pN}{n-k}}{\binom{N}{n}} & \text{si } \frac{c}{N} \leq p \leq 1 - \frac{n-c-1}{N} \\ 0 & \text{si } 1 - \frac{n-c-1}{N} \leq p \leq 1 \end{cases}$$

Como podemos observar, en las definiciones de $L_{n,c}(p)$ y $L_{N,n,c}(p)$ subyacen las funciones de probabilidad acumulada de la binomial B(n,p) y de la hipergeométrica H(N,p,n), respectivamente, consideradas como funciones de p, para un valor c, fijo, de la variable, con la salvedad de que no se hace ninguna restricción sobre los valores que puede

tomar p, sino que se considera que p puede tomar cualquier valor real.

Consideremos ahora la función auxiliar

$$L^{*}_{N,n,c}(p) = \sum_{k=0}^{c} \frac{\binom{pN}{k}\binom{N-pN}{n-k}}{\binom{N}{n}}$$

$L^{*}_{N,n,c}(p)$ es un polinomio en p, de grado n, que coincide con $L_{N,n,c}(p)$ en el intervalo $[c/N,1-(n-c-1)/N]$, para todo real p, y para cualesquiera números naturales N, n y c con $0 \le c < n \le N$.

Por otra parte, sabemos que

[1.4] $\quad L^{*}_{N,n,c}(p) = 1 \quad$ cuando $p = \frac{i}{N}$, con $i \le c$; $i \in \mathbb{Z}$

[1.5] $\quad L^{*}_{N,n,c}(p) = 0 \quad$ cuando $p = \frac{i}{N}$, con $i \ge N-n+c+1$; $i \in \mathbb{Z}$

La primera derivada, $\frac{d}{dp} L^{*}_{N,n,c}(p)$, como polinomio en p, de grado n-1, tiene n-1 raices distintas: $p'_1, p'_2, \ldots, p'_{n-1}$. Teniendo en cuenta [1.4] y [1.5], y considerando el carácter de función polinómica de $L^{*}_{N,n,c}(p)$, necesariamente

$$0 < p'_1 < \frac{1}{N} < p'_2 < \frac{2}{N} < \ldots < \frac{c-1}{N} < p'_c < \frac{c}{N} \ ,$$

y

$$\frac{N-n+c+1}{N} < p'_{c+1} < \frac{N-n+c+2}{N} < \ldots < \frac{N-1}{N} < p'_{n-1} < 1;$$

es decir, las n-1 raices se obtienen fuera del intervalo [c/N , 1-(n-c-1)/N], por lo que , como conclusión, en ese intervalo, $L^{*}_{N,n,c}(p)$ y $L_{N,n,c}(p)$ son funciones estrictamente decrecientes.

Apoyándose en la función auxiliar $L^{*}_{N,n,c}(p)$, Uhlmann (1966) compara las funciones $L_{N,n,c}(p)$ y $L_{N+1,n,c}(p)$, como primer paso en el estudio de la monotonía de la convergencia de $L_{N,n,c}(p)$ a $L_{n,c}(p)$, cuando $N\to\infty$, y obtiene el siguiente resultado:

1.1 TEOREMA:

(i) Para $c=0<n\leq N$,

$$L_{N,n,0}(p)-L_{N+1,n,0}(p) \text{ es } \begin{cases} 0 \text{ si } p=0 \text{ (ambas valen 1)} \\ <0 \text{ si } 0<p<1-\dfrac{n-1}{N+1} \\ 0 \text{ si } p\geq 1-\dfrac{n-1}{N+1} \text{ (ambas valen 0)} \end{cases}$$

(ii) Para $1\leq c=n-1<N$,

$$L_{N,n,0}(p) - L_{N+1,N,0}(p) \text{ es } \begin{cases} 0 \text{ si } 0 \leq p \leq \dfrac{n-1}{N+1} \text{ (ambas valen 1)} \\ >0 \text{ si } \dfrac{n-1}{N+1} < p < 1 \\ 0 \text{ si } p = 1 \text{ (ambas valen 0)} \end{cases}$$

[1.6] (iii) Para $0<c<n-1<N$ hay exactamente un número real p_o, con

$$\frac{c}{N} < p_o < 1 - \frac{n-c-1}{N},$$

tal que

$$L_{N,n,c}(p_o) = L_{N+1,n,c}(p_o).$$

(iv) Si n es un número impar mayor que 1, y $c=(n-1)/2$, entonces

$$L_{N,n,\frac{n-1}{2}}(p) - L_{N+1,n,\frac{n-1}{2}}(p) \text{ es}$$

$$\begin{cases} 0 & \text{si } 0 \le p \le \dfrac{n-1}{2(N+1)} \\ > 0 & \text{si } \dfrac{n-1}{2(N+1)} < p < \dfrac{1}{2} \\ = 0 & \text{si } p = 1/2 \\ < 0 & \text{si } \dfrac{1}{2} < p < 1 - \dfrac{n-1}{2(N+1)} \\ = 0 & \text{si } 1 - \dfrac{n-1}{2(N+1)} \le p \le 1 \end{cases}$$

El siguiente paso consiste en dar una mayor precisión al resultado [1.6], obteniendo el siguiente teorema:

1.2 TEOREMA:

Para $0<c<n-1<N$,

[1.7] $L_{N,n,c}(p) - L_{N+1,n,c}(p)$ es

$$\begin{cases} 0 & \text{si } 0 \le p \le \dfrac{c}{N+1} \\ > 0 & \text{si } \dfrac{c}{N+1} < p \le \dfrac{c}{n-1} - \dfrac{c}{(n-1)(N+1)} \\ < 0 & \text{si } \dfrac{c}{n-1} + \dfrac{n-c-1}{(n-1)(N+1)} \le p < 1 - \dfrac{n-c-1}{N+1} \\ 0 & \text{si } 1 - \dfrac{n-c-1}{N+1} \le p \le 1 \end{cases}$$

Con este teorema, el punto p_o de [1.6] queda localizado en el intervalo

$$\left(\frac{c}{n-1} - \frac{c}{(n-1)(N+1)} \, , \, \frac{c}{n-1} + \frac{n-c-1}{(n-1)(N+1)} \right).$$

A partir de este último teorema puede obtenerse fácilmente el siguiente, resultado de la comparación entre $L_{N,n,c}(p)$ y $L_{n,c}(p)$.

1.3 TEOREMA:

(i) $L_{N,n,o}(p) < L_{n,o}(p)$ si $0<p<1$;

$L_{N,n,n-1}(p) > L_{n,n-1}(p)$ si $0<p<1$;

verificandose la igualdad cuando p=0 y cuando p=1.

(ii) Para $0<c<n-1<N$,

$$L_{N,n,c}(p) - L_{n,c}(p) \text{ es } \begin{cases} = 0 & \text{si } p=0 \\ > 0 & \text{si } 0 < p \leq \frac{c}{n-1} - \frac{c}{(n-1)(N+1)} \\ < 0 & \text{si } \frac{c}{n-1} + \frac{n-c-1}{(n-1)(N+1)} \leq p < 1 \\ = 0 & \text{si } p = 1 \end{cases}$$

La demostración del apartado (i) es inmediata a partir de la definición de $L_{N,n,c}(p)$ y $L_{n,c}(p)$.

Para demostrar el apartado (ii) tengamos en cuenta que

$$\lim_{N\to\infty} L_{N,n,c}(p) = L_{n,c}(p),$$

por lo que podemos escribir

$$[1.8] \quad L_{N,n,c}(p) - L_{n,c}(p) = \sum_{i=0}^{\infty} [L_{N+i,n,c}(p) - L_{N+i+1,n,c}(p)]$$

A partir de [1.7], se habría obtenido que

$L_{N+i,n,c}(p) - L_{N+i+1,n,c}(p)$ es

$$\begin{cases} = 0 & \text{si } 0 \le p \le \dfrac{c}{N+i+1} \\ > 0 & \text{si } \dfrac{c}{N+i+1} < p \le \dfrac{c}{n-1}\,\dfrac{N+i}{N+i+1} \\ < 0 & \text{si } \dfrac{c}{n-1} + \dfrac{n-c-1}{(n-1)(N+i+1)} \le p < 1 - \dfrac{n-c-1}{N+i+1} \\ = 0 & \text{si } 1 - \dfrac{n-c-1}{N+i+1} \le p \le 1 \end{cases}$$

Si p=0, entonces, para todo i,

$$L_{N+i,n,c}(p) - L_{N+i+1,n,c}(p) = 0,$$

por lo que

$$L_{N,n,c}(p) - L_{n,c}(p) = 0.$$

Por otra parte, puesto que

$$\frac{N}{N+1} < \frac{N+i}{N+i+1} \qquad \forall i:1,2,.. \quad ,$$

obtendremos que si

$$0 < p \le \frac{c}{n-1}\,\frac{N}{N+1} \ ,$$

entonces

$$L_{N,n,c}(p) - L_{n,c}(p) > 0 \ ,$$

ya que, por pequeño que sea p, existirá un i suficientemente grande como para que p>c/(N+i+1), por lo que a partir de ese valor de i, los términos de la serie [1.8] son estrictamente positivos, siendo nulos los términos restantes.

Razonando en términos similares se completa la demostración del resultado.

Finalmente, para el caso en que n sea impar y c=(n-1)/2, se puede precisar más el resultado y se obtiene el siguiente teorema:

1.4 TEOREMA:

Si n es impar y c=(n-1)/2,

$$L_{N,n,\frac{n-1}{2}}(p) - L_{n,\frac{n-1}{2}}(p) \text{ es } \begin{cases} 0 & \text{si } p = 0 \\ >0 & \text{si } 0 < p < 1/2 \\ 0 & \text{si } p = 1/2 \\ <0 & \text{si } 1/2 < p < 1 \\ 0 & \text{si } p = 1 \end{cases}$$

APROXIMACIONES POR LA DISTRIBUCION DE POISSON

En primer lugar, establezcamos la notación a utilizar. Por $p(k;\lambda)$ representaremos la masa de probabilidad

en el punto k, de la distribución de Poisson de media λ.

Por tanto,

$$p(k;\lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!} \qquad k=0,1,2,\ldots \quad .$$

Es bien sabido que la binomial B(n,p) converge hacia la distribución de Poisson $P(\lambda)$, cuando el parámetro n crece de manera que el valor medio np se mantiene constante e igual a λ.

Esta convergencia permite aproximar probabilidades binomiales por probabilidades de Poisson. Consecuentemente, la convergencia de la hipergeométrica a la binomial, estudiada anteriormente, permite aproximar, en determinadas condiciones, las probabilidades hipergeométricas por las de Poisson.

Las condiciones que se debe cumplir para que la aproximación sea aconsejable, vienen obligadas por ambas convergencias. Es decir, X/N habrá de ser pequeño, y n grande. Naturalmente, estas condiciones son susceptibles de sustituirse por otras alternativas. Ya sabemos que se pueden intercambiar los papeles de los parámetros X y n en la hipergeométrica, por lo que podríamos tomar como condición igualmente satisfactoria para la aproximacion, el que n/N fuese pequeño y X grande. También podría tomarse N-X en lugar de X, cuando N-X sea pequeño; y N-n en lugar de n, cuando N-n sea grande, ya que los papeles desempeñados por las cantidades n y N-n, por un lado, y X y N-X, por otro, son equivalentes.

La aproximación usual de Poisson consiste en aproximar las probabilidades puntuales hipergeométricas por las correspondientes probabilidades de la distribución de Poisson que tiene por parámetro el valor medio de la hipergeométrica; es decir,

$$h(k;N,X,n) \simeq \frac{e^{-\frac{nX}{N}} (nX/N)^k}{k!} .$$

Mucho más precisa que esta aproximación es la que construye Burr (1973). En ella, las probabilidades puntuales de la distribución aproximante de Poisson van modificadas por un factor corrector que depende de los valores de los parámetros y del argumento de la función de masa de probabilidad hipergeométrica que se pretende aproximar.

$h(k;N,X,n)$ se aproxima por

$$p(k;nX/N) \left[1 + (\frac{1}{2X} + \frac{1}{2n}) [k - (k-nX/N)^2]\right].$$

Otra aproximación de Poisson fue dada por Molenaar (1970), tomando como Poisson aproximante aquella de parámetro

$$\lambda = nX/N + \frac{(n\ X/N\ -k)(2kN-nN+10kn)}{3N^2} .$$

Como podemos observar, λ es función de k, por lo que habrá una Poisson aproximante distinta para cada k.

APROXIMACIONES NORMALES

Emplearemos aquí las notaciones habitualmente establecidas para la distribución normal; es decir, para la normal unitaria N(0,1), la densidad de probabilidad en x vendrá dada por

$$\varphi(x) = (2\Pi)^{-1} e^{-\frac{x^2}{2}} ;$$

y la función de distribución en x será

$$\Phi(x) = \int_{-\infty}^{x} \varphi(u)\, du \quad .$$

Notaremos, por otra parte,

$$H_{ij} = \sum_{k=i}^{j} H(k;N,X,n) \quad .$$

La aproximación normal de distribuciones discretas tiene una historia que se remonta a los trabajos de De Moivre y Laplace, publicados en 1793 y 1812, respectivamente. El caso de aproximación para determinadas distribuciones, como la binomial y la hipergeométrica, es ampliamente conocido. Puede verse, a este respecto, lo tratado en Feller (1968).

La aproximación, usual -nominación que utilizaremos para distinguirla de otras aproximaciones normales que estudiaremos posteriormente-, consiste en aproximar la distribución hipergeométrica H(N,X,n), por la distribución normal $N(\mu_h,\sigma_h)$, siendo

$$\mu_h = \frac{nX}{n} \qquad y \qquad \sigma_h^2 = \frac{N-n}{N-1}\, pqn,$$

es decir, se aproxima por la distribución normal que tiene la misma media y varianza que la hipergeométrica que se pretende aproximar.

Naturalmente, el cálculo de las probabilidades se hará a través de la normal tipificada, N(0,1), cuya función de distribución se encuentra tabulada.

Como, por otra parte, se trata de aproximar una distribución discreta por una continua, se hará una corrección por continuidad.

En resumen: se aproximará

$$H_{ij} \quad \text{por} \quad \Phi(z_j + \frac{1}{2\sigma_h}) - \Phi(z_i - \frac{1}{2\sigma_h}),$$

donde

$$z_i = \frac{(i-\mu_h)}{\sigma_h}; \qquad z_j = \frac{(j-\mu_h)}{\sigma_h} \quad .$$

Las probabilidades puntuales hipergeométricas pueden aproximarse tomando i=j, ya que H_{ii}=h(i;N,X,n).

Las condiciones que aconsejan esta aproximación son que N sea grande y ninguna de las cantidades n, N-n, X y N-X sea demasiado pequeña.

Naturalmente, estos son resultados asintóticos, por lo que la aproximación, al tratarse realmente de valores finitos, conlleva un error. En el caso que estamos estudiando, el máximo error cometido es del orden de $1/\sigma_h$.

Para que la aproximación fuese útil, debería venir acompañada de una cota para el error cometido, preferentemente para el error relativo. Una cota así, como función de N, X, n, i y j, nos daría, en cualquier situación dada, si N, X y n son suficientemente grandes como para dar la exactitud deseada en la aproximacion.

En este sentido, Nicholson (1956) hace una generalización para la hipergeométrica, de un trabajo de Feller (1945) sobre la aproximación de la binomial a la normal, obteniendo unas cotas para la probabilidad acumulativa.

Define:

$$p_1 = \frac{X+1}{N+2}; \qquad q_1 = 1-p_1;$$

$$s = \frac{n+1}{N+2}; \qquad t = 1-s;$$

$$a = \frac{1}{6}\,(p_1-q_1)(t-s); \qquad \sigma = \sqrt{(N+1)tpq};$$

y para cada k,

$$x_k = [k + \frac{1}{2} - (n+1)p_1]/\sigma \quad .$$

Entonces, para

$$\sigma > 3; \qquad i \geq (n+1)p_1;$$

$$j + \frac{1}{2} \leq (n+1)p_1 + \frac{2}{3}\sigma^2;$$

$$n-j \geq 4; \qquad X-j \geq 4;$$

se verifica:

$$H_{ij} \leq \frac{N+1}{N+2} e^R \left[\Phi(\eta_{j+1}) - \Phi(\eta_i)\right] ,$$

donde

$$R = \frac{5(1-p_1q_1)(1-st)}{36\sigma^2} + \frac{2}{3(N+2)}$$

y

$$\eta_k = \frac{k-(n+1)p_1}{\sigma} + \frac{a}{\sigma}\left[\frac{k-(n+1)p_1}{\sigma}\right]^2 + \frac{2a}{\sigma} - \frac{1}{2\sigma^2} \quad ,$$

mientras que la desigualdad se verificaría en el otro sentido si

$$\eta_k = \frac{k-(n+1)p_1}{\sigma} + \frac{a}{\sigma}\left[\frac{k-(n+1)p_1}{\sigma}\right]^2 + \frac{2a}{\sigma} + \frac{M}{6\sigma} + \frac{1}{7\sigma} \quad ,$$

con $M = \frac{x_i^3}{\sigma} = (j+1/2-(n+1)p_1)^3\sigma^{-4}$.

Como un estimador puntual de H_{ij} sugiere la expresión

$$\frac{N+1}{N+2}\, e^R \left[\Phi(\eta_{j+1}) - \Phi(\eta_i)\right] ,$$

con η_k definido por:

$$\eta_k = \frac{k-(n+1)p_1}{\sigma} + \frac{a}{\sigma}\left[\frac{k-(n+1)p_1}{\sigma}\right]^2 + \frac{2a}{\sigma} .$$

Designando por $\hat{H}$ el valor estimado de H_{ij} , por L la cota inferior, y por U la cota superior, se obtiene la siguiente cota superior para el error relativo:

$$\frac{1}{L}\, máx(\hat{H}-L, U-\hat{H}) .$$

En una autocrítica de su método, Nicholson nos advierte que la aproximación obtenida es buena si la distribución es altamente sesgada ($|a|$ grande); sin embargo, para valores de a próximos a cero es muy poca la ventaja obtenida, a cambio de un proceso mucho más complicado que el procedimiento clásico.

Vamos a ver ahora otras dos aproximaciones normales a la hipergeométrica, recogidas de un reciente artículo de

Ling y Pratt (1984). Fueron obtenidas por Peizer, en 1968, en un trabajo no publicado, y son las mejores aproximaciones normales a la hipergeométrica. Surgen de una generalización de la aproximación normal de Peizer y Pratt para la distribución binomial, fuente de las mejores aproximaciones normales construidas para una serie de distribuciones mutuamente relacionadas: binomial negativa, Pascal, Beta, F de Snedecor, t de Student, Poisson, Gamma y Chi-cuadrado [Peizer y Pratt (1968) y Pratt (1968)].

Consideremos la hipergeométrica H(N,X,n), y vamos a hacer un cambio respecto de la notación que hemos utilizado hasta ahora. Llamaremos

$$a = k; \qquad b = n-k;$$

$$c = X-k; \qquad d = N-X-n+k;$$

$$s = N-X; \qquad m = N-n; \qquad r = X.$$

Por otra parte, llamaremos

$$A = a+0.5; \qquad B = b-0.5;$$

$$C = c-0.5; \qquad D = d+0.5.$$

Consideremos

$$L = A \ \mathrm{Ln} \frac{AN}{nr} + B \ \mathrm{Ln} \frac{BN}{ns} + c \ \mathrm{Ln} \frac{CN}{mr} + D \ \mathrm{Ln} \frac{DN}{ms} \quad .$$

La primera aproximación de Peizer aproxima las probabilidades acumuladas hipergeométricas H(k;N,X,n) por $\Phi(z_k)$, siendo

$$z_k = \frac{A'D'-B'C'}{|AD-BC|}\left[2L\,\frac{mnrsN'}{m'n'r's'N}\right]^{1/2},$$

donde

$$A' = A+ 1/6; \qquad B' = B+ 1/6 ;$$

$$C' = C+ 1/6; \qquad D' = D+ 1/6 ;$$

$$r' = r+ 1/6; \qquad s' = s+ 1/6 ;$$

$$n' = n+ 1/6; \qquad m' = m+ 1/6 ;$$

$$N' = N- 1/6.$$

La segunda aproximación de Peizer gana aún algo más de precisión tomando

$$z_k = \frac{A''D''-B''C''}{|AD-BC|}\left[2L\,\frac{mnrsN'}{m'n'r's'N}\right]^{1/2},$$

donde

$$A'' = A' + \frac{0.02}{A+0.05} + \frac{0.01}{n+1} + \frac{0.01}{r+1} \quad .$$

Análogamente se construye B", C" y D".

Un estudio más detallado de las dos aproximaciones de Peizer, así como un exhaustivo estudio empírico comparativo de diversas aproximaciones normales y binomiales a la

hipergeométrica, puede verse en el citado artículo de Ling y Pratt (1984).

Vamos a comentar, finalmente, un trabajo igualmente reciente de Quinn y Mac Gillivray (1986), en el que presentan condiciones suficientes para la convergencia hacia la Normal de sucesiones de variables aleatorias discretas fuertemente unimodales y con soporte consecutivo, es decir, con soporte de la forma $\{x\in\mathbb{Z};a\leq x\leq b\}$, donde a y b pueden ser, respectivamente, $-\infty$ y $+\infty$

Estas condiciones la aplican para obtener, de forma directa, la aproximación normal para la distribución hipergeométrica

1.5 TEOREMA:

Sea $\{X_n\}$ una sucesión de variables aleatorias discretas cuyas funciones de probabilidad $f_n(x)$ tienen soporte consecutivo y son fuertemente unimodales y sea

$$g_n(x) = \begin{cases} \dfrac{f_n(x)}{f_n(x-1)} & \text{si } a<x\leq b \\ \infty & \text{si } x\leq a \\ 0 & \text{si } x>b \end{cases}$$

Supongamos que existe una sucesión de enteros $\{\alpha_n\}$

tal que

(i) $g_n(x) \gtrless 1$ cuando $x \lessgtr \alpha_n$.

Sea $\beta_n=\{g_n(\alpha_n)-g_n(\alpha_n+1)\}^{-1/2}$ y supongamos que

(ii) $\beta_n \to \infty$, y

(iii) $\beta_n^2\{g_n(\alpha_n+j)-1\}=-j+k_n(j)$,

donde

$$\lim_{n\to\infty} \sup_{|j|\leq x\beta_n} |j^{-1}k_n(j)|=0 \qquad \forall x>0.$$

Entonces, para todo $x\in\mathbb{R}$,

$$\lim_{n\to\infty} \text{Prob}\,[(X_n-\alpha_n)/\beta_n\leq x] = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-\frac{1}{2}u^2}\,du \quad .$$

Los autores hacen hincapié en que, en la práctica, raras veces es necesario calcular el valor exacto de α_n y β_n , sino que suele ser suficiente cualquier aproximación de primer orden.

Este resultado se puede aplicar a la sucesión de hipergeométricas $\{H_n\}$, donde $H_n=H(n_r,n_s,n)$, con

n_r= parte entera de rn ; n_s= parte entera de sn ; $1<s\leq r-1$.

En este caso,

$$f_n(x)=\begin{cases} \dfrac{\binom{n_s}{x}\binom{n_r-n_s}{n-x}}{\binom{n_r}{n}} & ; \quad x:0,1,\ldots,n \\ 0 & \text{en otro caso} \end{cases}$$

Como vimos en [1.4](a) (sección I.6),

$$g_n(x+1) = \frac{f_n(x+1)}{f_n(x)} = \frac{(n_s-x)(n-x)}{(x+1)(n_r-n_s-n+x+1)} .$$

A partir de aquí se calcula los valores α_n tales que $g_n(x) \underset{(\leq)}{\geq} 1$ cuando $x \underset{(\geq)}{\leq} \alpha_n$, obteniendose que

$$\alpha_n = \frac{sn}{r} + 0(1) ,$$

donde, con la notación usual, $0(1)$ es tal que $\lim_{n->\infty} 0(1)=k$, con $0<|k|<\infty$.

Calculando $\beta_n = \{g_n(\alpha_n)-g_n(\alpha_n+1)\}^{-1/2}$, se obtiene

$$\beta_n = \left[\frac{sn(r-1)(r-s)}{r^3}\right]^{1/2} +0(1) .$$

Finalmente, se comprueba que la condición (iii) del teorema se verifica, por lo que $(X_n-\alpha_n) / \beta_n$ tiene una distribución que converge a la N(0,1).

APROXIMACION POR LA DISTRIBUCION BETA

Hemos visto, anteriormente, que existen unas condiciones de tamaño para los parámetros N, X y n que se deben cumplir para que la aproximación normal pueda considerarse satisfactoria. Bol'shev (1964) propone una aproximación, satisfactoria para $N \geq 25$, independiente del tamaño de X y de n, utilizando la distribución Beta.

Recordemos que la distribución Beta, de parámetros a y b, es una distribución continua cuya función de densidad es

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}; \qquad 0 \leq x \leq 1,\ a>0,\ b>0,$$

donde B(a,b) es la función Beta, definida

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \quad .$$

Si notamos por $I_p(a,b)$ la función de distribución en el punto p, es decir,

$$I_p(a,b) = \frac{\int_0^1 x^{a-1}(1-x)^{b-1}\,dx}{B(a,b)} \quad ,$$

la aproximación de Bol'shev consiste en aproximar

$$H(k;N,X,n) \quad \text{por} \quad I_{1-p}(n'-k+c,k-c+1),$$

donde

$$p = \frac{N(n+X-1)-2nX}{N(N-2)} \quad ;$$

$$c = \frac{n(n-1)X(X-1)}{(N-1)[(N-n)(N-X)+nX-N]} \quad ;$$

$$n' = \frac{(N-2)^2 nX(N-n)(N-X)}{(N-1)[(N-n)(N-X)+nX-N][N(n+X-1)-2nX]} \quad .$$

Para N tendiendo a infinito, esta aproximación es tan buena asintóticamente como cualquier aproximación conocida para la hipergeométrica, ya que los mismos teoremas límites se verifican para $H(k;N,X,n)$ que para I_{1-p}.

Además, $I_{1-p}(n'-k+c,k-c+1)$ verifica las mismas relaciones que vimos para $H(k;N,X,n)$ en I.6 (f).

En su artículo, Bol'shev compara su aproximación con la aproximación normal usual, en un caso en el que ésta funciona bien, y encuentra que al menos es tan satisfactoria como ella.

CAPITULO II

TEOREMAS DE CARACTERIZACION

II.1 INTRODUCCION

Este capítulo de la memoria tiene como objeto recoger parte de nuestra aportación al estudio de la distribución hipergeométrica; concretamente, lo referido a caracterizaciones de esta distribución, aspecto éste que hemos encontrado bastante desatendido en la literatura revisada -a diferencia de lo que ocurre con otras distribuciones discretas-, hasta el punto de sólo poder disponer de un teorema de caracterización de la hipergeométrica: el debido a Skibinsky (1970).

Este teorema se fundamenta en el comportamiento de la distribucion hipergeometrica ante la mixtura o composicion, respecto del parámetro X -tamaño de la clase-, con una distribución binomial.

Este comportamiento singular -la distribución compuesta resulta del mismo tipo que la supuesta para el parámetro X-, se manifiesta siempre que X siga una distribución perteneciente al esquema de urna de Polya.

La constatación de este hecho nos da pie, en primer lugar, para obtener un teorema de caracterización siguiendo una línea paralela al teorema de Skibinsky, diferenciandonos de este autor en la distribución seguida por el parámetro: binomial en su caso e hipergeométrica en el nuestro; y para, a

FACULTAD DE CIENCIAS BIBLIOTECA GRANADA

continuación, considerando X distribuido según una distribución de Polya, obtener un teorema de caracterización más general, del cual el de Skibinsky y el nuestro anterior son casos particulares.

II.2 MIXTURA DE DISTRIBUCIONES

DEFINICION

En el concepto mas amplio, si $F_j(x_1,x_2,\ldots,x_n)$, con $j=\ldots,-1,0,1,2,\ldots$, representa diferentes funciones de distribución, y a_j son tales que $\sum_{j=-\infty}^{+\infty} a_j=1$ y $a_j \geq 0$, $\forall j$, entonces,

$$F(x_1,x_2,\ldots,x_n) = \sum_{j=-\infty}^{+\infty} a_j\, F_j(x_1,x_2,\ldots,x_n),$$

es la función de distribución de lo que llamamos mixtura o composición de las distribuciones F_j.

Aquí vamos a tratar el caso particular en el que las funciones $F_j(x_1,x_2,\ldots,x_n)$ son de la forma $F(x;\theta)$ y los a_j

representan las probabilidades de la distribución del parámetro θ.

En este caso, la distribución obtenida por la mixtura tiene por función de distribución el valor esperado, respecto de la distribución del parámetro θ, de la función de distribución.

Como es lógico, nos vamos a centrar en la distribución hipergeométrica, y, en particular, vamos a estudiar casos de mixtura de esta distribución respecto del parámetro X (tamaño de la clase); es decir, consideraremos el parámetro X como variable aleatoria.

En cuanto a la distribución que ha de seguir el citado parámetro, nos vamos a interesar en aquellas que, posteriormente, van a servir de base para poder caracterizar la distribución hipergeométrica, objeto último de este capítulo.

Finalmente, la notación que utilizaremos para la mixtura respecto del parámetro X, será:

$$H(N,X,n) \underset{X}{\wedge} F \ ,$$

donde F representa la distribución que, en cada caso, siga el mencionado parámetro.

MIXTURAS DE UNA DISTRIBUCION HIPERGEOMETRICA

(i) $\quad H(N,X,n) \hat{X} B(N,p)$

La función de masa de probabilidad de la distribución resultante de esta mixtura será:

$$p(\xi=k) = \sum_{j} h(k;N,j,n)\ p(X=j) =$$

$$= \sum_{j=k}^{N-n+k} \frac{\binom{j}{k}\binom{N-j}{n-k}}{\binom{N}{n}} \binom{N}{j} p^j q^{N-j} \quad .$$

Sustituyendo los números combinatorios por su expresión factorial, simplificando, y reordenando de forma conveniente, obtenemos:

$$p(\xi=k) = \sum_{j=k}^{N-n+k} \binom{n}{k}\binom{N-n}{j-k} p^j q^{N-j} =$$

$$= \binom{n}{k} p^k q^{n-k} \sum_{j-k=0}^{N-n} \binom{N-n}{j-k} p^{j-k} q^{N-n-j+k} \quad ,$$

y, puesto que la última sumatoria es igual a la unidad, por ser la suma de las probabilidades de una binomial B(N-n,p), nos queda que

$$p(\xi=k) = \binom{n}{k} p^k q^{n-k} \quad .$$

En conclusión:

[2.1] $$H(N,X,n) \underset{X}{\wedge} B(N,p) = B(n,p)$$

Este resultado era previsible, si pensamos que la mixtura representaría el resultado de elegir aleatoriamente una muestra de tamaño N, de una población binomial, y a continuación, escoger, aleatoriamente y sin reemplazamiento, n elementos de dicha muestra. Como puede apreciarse, esto es equivalente a escoger de una vez n elementos de la población binomial.

Por otra parte, el hecho de que la distribución resultante de la mixtura sea de nuevo binomial, va a ser un criterio caracterizador de la hipergeométrica, como veremos en la siguiente sección de este capítulo.

El resultado anterior puede generalizarse suponiendo que el parámetro X sigue una distribución mixtura de binomiales, es decir,

$$p(X=j) = \sum_{i=1}^{m} \omega_i \binom{N}{j} p_i^j q_i^{N-j} \quad ,$$

$$\text{con } 0<p_i<1,\ 0<\omega_i<1;\ \sum_{i=1}^{m} \omega_i=1.$$

Realizando un estudio análogo al anterior, se obtiene:

$$H(N,X,n) \underset{X}{\wedge} \text{Mix } B(N,p_i) = \text{Mix } B(n,p_i).$$

(ii) $H(N,X,n) \underset{X}{\wedge} H(N',X',N)$

Supondremos, en este caso, que el parámetro X está distribuido según una hipergeométrica H(N',X',N).

La función de masa de probabilidad de la distribución resultante de la mixtura, vendrá dada por:

$$p(\xi=k) = \sum_j h(k;N,j,n)\ p(X=j) =$$

$$= \sum_j \frac{\binom{j}{k}\binom{N-j}{n-k}}{\binom{N}{n}} \frac{\binom{X'}{j}\binom{N'-X'}{N-j}}{\binom{N'}{n}},$$

donde máx(k,X'-N'+N) ≤ j ≤ mín(X',N-n+k).

Si sustituimos cada número combinatorio por su expresión factorial, simplificamos, y reordenamos de forma adecuada los factoriales resultantes, obtenemos que

$$p(\xi=k) = \sum_{j} \frac{\binom{X'}{k}\binom{N'-X'}{n-k}}{\binom{N'}{n}} \frac{\binom{N-n}{j-k}\binom{N'-N}{X'-j}}{\binom{N'-n}{X'-k}} \quad .$$

Extrayendo factores fuera de la sumatoria ,

$$p(\xi=k) = \frac{\binom{X'}{k}\binom{N'-X'}{n-k}}{\binom{N'}{n}} \sum_{j-k} \frac{\binom{N-n}{j-k}\binom{N'-N}{X'-j}}{\binom{N'-n}{X'-k}} \quad ,$$

donde $\max(0,X'+N-N'-k) \leq j-k \leq \min(X'-k,N-n)$.

La última sumatoria es igual a la unidad, por ser la suma de las probabilidades puntuales de una distribucion H(N'-n,N-n,X'-k); por tanto,

$$p(\xi=k) = \frac{\binom{X'}{k}\binom{N'-X'}{n-k}}{\binom{N'}{n}} \quad ,$$

que, como vemos, corresponde a la función de masa de probabilidad de una hipergeométrica H(N',X',n).

En conclusión:

[2.2] $$H(N,X,n) \underset{X}{\wedge} H(N',X',N) = H(N',X',n) \quad .$$

(iii) $H(N,X,n) \hat{X}$ Polya

Realizaremos, a continuación, el estudio de la mixtura con una distribución de Polya, mixtura sugerida en el libro de Johnson y Kotz (1969).

Recordemos que el esquema de urna de Polya es como sigue:

Supongamos una urna conteniendo M bolas blancas y N-M bolas negras. Se extrae una bola al azar, se anota su color, y se vuelve a introducir en la urna junto a c bolas más del mismo color. En estas condiciones, la operación se repite n veces. La variable aleatoria que indica el número de bolas blancas que se ha obtenido, decimos que sigue una distribución de Polya de parámetros N,M,n,c.

La función de masa de probabilidad viene dada por:

$$p_r = \frac{\binom{n}{r} M^{(r,c)} (N-M)^{(n-r,c)}}{N^{(n,c)}},$$

donde la notación $A^{(x,c)} = A(A+c)(A+2c)\ldots(A+(x-1)c)$.

Para c=-1 se obtiene la hipergeométrica, y para c=0 la distribución binomial, como casos particulares del esquema de urna de Polya.

Hemos visto que el comportamiento de la distribución hipergeométrica frente a una mixtura, respecto de X, con una binomial y con una hipergeométrica es el mismo: la distribución resultante de la mixtura sigue siendo binomial e hipergeométrica, respectivamente, cambiando el parámetro N por n.

Como vamos a demostrar a continuación, este comportamiento se va a seguir manteniendo si la mixtura se realiza con una distribución de Polya cualquiera; es decir, vamos a ver que para cualquier valor del parámetro c,

$$[2.3] \qquad H(N,X,n) \underset{X}{\wedge} Polya(N',M,N,c) = Polya(N',M,n,c) \ .$$

En efecto, la función de masa de probabilidad de la distribución resultante de la mixtura será:

$$p(\xi=k) = \sum_{j} h(k;N,j,n)\ p(X=j) =$$

$$= \sum_{j=k}^{N-n+k} \frac{\binom{j}{k}\binom{N-j}{n-k}}{\binom{N}{n}} \binom{N}{j} \frac{M^{(j,c)}\ (N'-M)^{(N-j,c)}}{N'^{(N,c)}} \ ,$$

Si sustituimos los números combinatorios por su expresión factorial, simplificamos, y reordenamos de forma adecuada, obtenemos:

$$p(\xi=k) = \sum_{j=k}^{N-n+k} \begin{pmatrix} n \\ k \end{pmatrix} \begin{pmatrix} N-n \\ j-k \end{pmatrix} \frac{M^{(j,c)}\ (N'-M)^{(N-j,c)}}{N'^{(N,c)}} .$$

Por otra parte, puede verse fácilmente que

$$M^{(j,c)} = M^{(k,c)}\ (M+kc)^{(j-k,c)}$$

$$N'^{(N,c)} = N'^{(n,c)}\ (N'+nc)^{(N-n,c)}$$

$$(N'-M)^{(N-j,c)} = (N'-M)^{(n-k,c)}\ (N'-M+nc-kc)^{(N-n-j+k,c)} .$$

Haciendo uso de esta descomposición, y sacando factores fuera de la sumatoria,

$$p(\xi=k) = \begin{pmatrix} n \\ k \end{pmatrix} \frac{M^{(k,c)}\ (N'-M)^{(n-k,c)}}{N'^{(n,c)}} .S \quad ,$$

donde

$$S = \sum_{j-k=o}^{N-n} \begin{pmatrix} N-n \\ j-k \end{pmatrix} \frac{(M+kc)^{(j-k,c)}\ (N'-M+nc-kc)^{(N-n-j+k,c)}}{(N'+nc)^{(N-n,c)}}$$

y, puesto que la sumatoria S es igual a la unidad, por ser la suma de las probabilidades puntuales de una distribución de Polya(N'+nc,M+kc,N-n,c), se obtiene, finalmente, que

$$p(\xi=k) = \begin{pmatrix} n \\ k \end{pmatrix} \frac{M^{(k,c)}\ (N'-M)^{(n-k,c)}}{N'^{(n,c)}} .$$

La distribución resultante de la mixtura, es, por tanto, una distribución de Polya(N',M,n,c).

Este resultado era, de nuevo, previsible, si nos fijamos en que la operación mixtura es el resultado de escoger aleatoriamente un subconjunto de n elementos, de una muestra de tamaño N, extraida según un esquema de urna de Polya; y esto es equivalente a escoger directamente una muestra de tamaño n, extraida según dicho esquema.

II.3 TEOREMAS DE CARACTERIZACION

En esta Sección recogemos nuestra aportación al estudio de la caracterización de la familia hipergeométrica, concretada en dos teoremas, basados ámbos en el comportamiento de la citada distribución frente a la composición, respecto del parámetro tamaño de la clase, con otras distribuciones pertenecientes al esquema de urna de Polya.

Este planteamiento ya fue usado por Skibinsky (1970), quien tomando como punto de partida el comportamiento de la hipergeométrica frente a la mixtura con una binomial (vease [2.1]), obtiene el siguiente teorema de caracterización:

2.1 TEOREMA

Si $h_0, h_1, \ldots, h_N$ son funciones que toman valores en $\{0,1,\ldots,n\}$, y por $b(i;r,\theta)$ notamos la probabilidad de que la variable binomial de parámetros r y θ tome el valor i, entonces,

$$h_j(i) = \frac{\binom{j}{i}\binom{N-j}{n-i}}{\binom{N}{n}}, \qquad \begin{array}{l}\text{con } i=0,1,\ldots,n \\ \quad\; j=0,1,\ldots,N \\ \quad\; n\leq N\end{array}$$

si y sólo si

a) los h_j son independientes de θ

b) $\sum_{j=0}^{N} h_j(i)\, b(j;N,\theta) = b(i;n,\theta)$.

Demostración:

La demostración de este teorema, en uno de los sentidos, ha sido ya vista en el apartado (i) de la Sección anterior. Veamosla en el otro sentido.

Supongamos que

$$b(i;n,\theta) = \sum_{j=0}^{N} h_j(i)\, b(j;N,\theta),$$

es decir,

$$\binom{n}{i}\theta^{i}(1-\theta)^{n-i} = \sum_{j=o}^{N} h_j(i)\binom{N}{j}\theta^{j}(1-\theta)^{N-j} =$$

$$= \sum_{j=o}^{N} h_j(i)\binom{N}{j}\theta^{j}\theta^{j-i}(1-\theta)^{n-i}(1-\theta)^{N-n-j+i} =$$

$$= \binom{n}{i}\theta^{i}(1-\theta)^{n-i} \sum_{j=o}^{N} h_j(i)\frac{\binom{N}{j}\theta^{j-i}(1-\theta)^{N-n-j+i}}{\binom{n}{i}} .$$

De aquí, se tiene que la última sumatoria es necesariamente igual a la unidad. Vamos a recurrir ahora, a otra sumatoria que, para cualquier valor de θ, es siempre igual a la unidad, por ser la suma de las probabilidades puntuales de la variable binomial B(N-n,θ). Ella es:

$$\sum_{j=i}^{N-n+i} \binom{N-n}{j-i}\theta^{j-i}(1-\theta)^{N-n-(j-i)} .$$

Tendremos, por tanto que

$$\sum_{j=o}^{N} h_j(i)\frac{\binom{N}{j}\theta^{j-i}(1-\theta)^{N-n-(j-i)}}{\binom{n}{i}} = 1 =$$

$$= \sum_{j=i}^{N-n+i} \binom{N-n}{j-i}\theta^{j-i}(1-\theta)^{N-n-(j-i)} \qquad \forall\ 0<\theta<1 .$$

Si identificamos los coeficientes de las

potencias de θ en ambas sumatorias, teniendo en cuenta que $h_j(i)$ es independiente de θ, obtendremos:

$$h_j(i) = 0 \quad \text{si} \quad j<i \quad \acute{o} \quad j>N-n+i$$

$$h_j(i) = \frac{\binom{N-j}{j-i}\binom{n}{i}}{\binom{N}{j}} \qquad \text{si } i\leq j\leq N-n+i$$

y para que el teorema quede demostrado basta hacer uso, por último, de la igualdad

$$\frac{\binom{j}{i}\binom{N-j}{n-i}}{\binom{N}{n}} = \frac{\binom{n}{i}\binom{N-n}{j-i}}{\binom{N}{j}} .$$

Con esto queda demostrado que la familia $\{h_j\}$ con $j=0,1,\ldots,N$ es la familia hipergeométrica de parámetros N y n (tamaño de población y tamaño de muestra, respectivamente).

El primero de nuestros teoremas de caracterización parte de un planteamiento paralelo al de Skibinsky (1970), sin más que cambiar la distribución seguida por el parámetro X.

Consideremos ahora que éste sigue una distribución hipergeométrica H(N',X',N); en este caso, sabemos por [2.2]

que

$$H(N,X,n) \hat{X} H(N',X',N) = H(N',X',n) \ .$$

Veamos que esta propiedad es caracterizadora de la familia hipergeomérica.

2.2. TEOREMA

Sea $\{F_x\}$, X:0,1,...,N, una familia de funciones de distribución que asignan probabilidades no nulas sólo a los enteros $\{0,1,...,n\}$. Condición necesaria y suficiente para que esta familia sea la formada por las hipergeométricas H(N,X,n), es que se verifique:

(1) Su composición cuando X sigue una hipergeométrica H(N',X',N) es la hipergeométrica H(N',X',n).

(2) Las distribuciones F_x no dependen del valor N'.

Demostración:

Condición necesaria:

La condición (2) se cumple trivialmente. En cuanto a la condición (1) está demostrada con el resultado [2.2].

Condición suficiente:

Sea F_x una familia de distribuciones que no depende de N' y que cumple que

$$[2.4] \qquad F_x \hat{X} H(N',X',N) = H(N',X',n) \ .$$

Veamos que F_x sólo puede ser H(N,X,n).

La condición [2.4], aplicada a cada uno de los valores en los que está definida la distribución, exige que

$$[2.5] \qquad \sum_{x=o}^{N} f_x(k)\ h(X;N',X',N) = h(k;N',X',n).$$

Si consideramos k fijo, [2.5] representa una ecuación con N+1 incógnitas $f_o, f_1, \ldots, f_N$, y, puesto que sabemos que F no depende de N', podemos, variando N' desde N' hasta N'+N, plantear un sistema de N+1 ecuaciones con N+1 incógnitas. Este sistema de ecuaciones admite, al menos la solución

$$f_x(k) = \frac{\binom{X}{k}\binom{N-X}{n-k}}{\binom{N}{n}} \ ,$$

como consecuencia inmediata de la condición necesaria del teorema. Si probamos que el sistema anterior tiene solución única, habremos demostrado el teorema. Al

sustituir en [2.5] las probabilidades hipergeométricas por su expresión correspondiente, se obtiene

$$\sum_{x=o}^{N} f_x(k) \frac{\binom{X'}{X}\binom{N'-X'}{N-X}}{\binom{N'}{N}} = \frac{\binom{X'}{k}\binom{N'-X'}{n-k}}{\binom{N'}{n}},$$

con $N':N',N'+1,\ldots,N'+N,$

que también podemos escribir como

$$[2.6] \qquad \sum_{x=o}^{N} f_x(k) \binom{X'}{X}\binom{N'-X'}{N-X} = \frac{\binom{N'}{N}\binom{X'}{k}\binom{N'-X'}{n-k}}{\binom{N'}{n}},$$

con $N':N',N+1,\ldots,N'+N.$

El determinante de la matriz de los coeficientes del sistema [2.6] es de la forma

$$\begin{vmatrix}
\binom{X'}{0}\binom{N'-X'}{N} & \binom{X'}{1}\binom{N'-X'}{N-1} & \cdots & \binom{X'}{N}\binom{N'-X'}{0} \\
\binom{X'}{0}\binom{N'+1-X'}{N} & \binom{X'}{1}\binom{N'+1-X'}{N-1} & \cdots & \binom{X'}{N}\binom{N'+1-X'}{0} \\
\vdots & \vdots & \cdots & \vdots \\
\binom{X'}{0}\binom{N'+N-X'}{N} & \binom{X'}{1}\binom{N'+N-X'}{N-1} & \cdots & \binom{X'}{N}\binom{N'+N-X'}{0}
\end{vmatrix}$$

Extrayendo factores fuera del determinante, y llamando, por comodidad, M a N'-X', se obtiene

$$[2.7] \qquad \binom{X'}{0}\binom{X'}{1}\cdots\binom{X'}{N} \begin{vmatrix} \binom{M}{N} & \binom{M}{N-1} & \binom{M}{N-2} & \cdots & \binom{M}{0} \\ \binom{M+1}{N} & \binom{M+1}{N-1} & \binom{M+1}{N-2} & \cdots & \binom{M+1}{0} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \binom{M+N}{N} & \binom{M+N}{N-1} & \binom{M+N}{N-2} & \cdots & \binom{M+N}{0} \end{vmatrix}$$

Puesto que el factor multiplicador de [2.7] es distinto de cero, nos queda probar que el determinante también es distinto de cero.

Para ello nos basamos en la conocida propiedad de los números combinatorios:

$$\binom{a+1}{b} = \binom{a}{b-1} + \binom{a}{b}.$$

Aplicandola y restando sucesivamente a cada fila del determinante la fila anterior, se obtiene un determinante equivalente, de la forma

$$\begin{vmatrix} \binom{M}{N} & \binom{M}{N-1} & \binom{M}{N-2} & \cdots & \binom{M}{1} & 1 \\ \binom{M}{N-1} & \binom{M}{N-2} & \binom{M}{N-3} & \cdots & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \binom{M}{1} & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{vmatrix}$$

que, evidentemente, es distinto de cero.

Tanto la caracterización de Skibinsky como nuestro anterior teorema de caracterización, se basan en la particularidad de que la distribución compuesta de una hipergeométrica, respecto del parámetro X, resulta del mismo tipo que la supuesta para el parámetro. Teniendo en cuenta que tanto binomial como hipergeométrica son distribuciones pertenecientes al esquema de urna de Polya, pensamos que debería ser posible extender ambos resultados, englobandolos en un único teorema de caracterización de la distribución hipergeométrica, basado en la composición con distribuciones de Polya.

2.3 TEOREMA

Sea $\langle F_x \rangle$, X:0,1,...,N, una familia de funciones de distribución que asignan probabilidades no nulas sólo a los enteros $\langle 0,1,...,n \rangle$. Condición necesaria y suficiente para que esta familia sea la constituida por las hipergeométricas H(N,X,n), es que se verifique:

(1) Su composición cuando X sigue una Polya $\mathcal{P}(N',X',N,c)$ es la Polya $\mathcal{P}(N',X',n,c)$.

(2) Las distribuciones F_x no dependen del valor de N'.

Demostración:

Condición necesaria:

La condición (2) se cumple trivialmente. Por otra parte, la condición (1) quedó probada con el resultado [2.3].

Condición suficiente:

Sea F_x una familia de distribuciones que no depende de N' y que cumple que

$$[2.8] \qquad F_x \hat{X} \mathcal{P}(N',X',N,c) = \mathcal{P}(N',X',n,c).$$

Veamos, entonces, que F_x sólo puede ser H(N,X,n).

La condición [2.8], aplicada a cada uno de los valores en los que esta definida la distribución, exige que

$$[2.9] \qquad \sum_{x=o}^{N} f_x(k)\ p(x;N',X',N,c) = p(k;N',X',n,c)\ .$$

Ahora bien, esta expresión, para un valor k fijo, representa una ecuación con N+1 incógnitas $f_o, f_1, \ldots, f_N$.

Teniendo en cuenta la no dependencia de F respecto de N', podemos, variando N' desde N' hasta N'+N, plantear un sistema de N+1 ecuaciones con N+1 incógnitas.

Dicho sistema de ecuaciones admite, al menos, la solución

$$f_x(k) = \frac{\binom{X}{k}\binom{N-X}{n-k}}{\binom{N}{n}},$$

como consecuencia inmediata de la condicion necesaria del teorema. Si demostramos que el sistema anterior tiene solución única, el teorema quedaría demostrado.

Sustituyendo en la ecuación [2.9] las probabilidades de las distribuciones de Polya por su valor, y pasando los denominadores al segundo miembro, tendremos

$$\sum_{X=o}^{N} f_x(k)\binom{N}{X} X'^{(X)}(N'-X')^{(N-X)} =$$

[2.10]

$$= \frac{\binom{n}{k} X^{(k)} (N'-X')^{(n-k)} N'^{(N)}}{N'^{(n)}}.$$

Estamos prescindiendo, en la notación ahora utilizada, de explicitar la constante c (cuarto parámetro de la distribución de Polya), de manera que

$$A^{(\alpha)} = A(A+c)(A+2c)\ldots(A+(\alpha-1)c)$$
$$A^{(0)} = 1$$

Llamando M a N'-X' y variando N' desde N' hasta N'+N, la ecuación [2.10] da lugar a un sistema de ecuaciones lineales con un determinante de la matriz de los coeficientes de la forma

$$\binom{N}{0}\cdots\binom{N}{N}X'^{(0)}\cdots X'^{(N)}\begin{vmatrix} M^{(N)} & M^{(N-1)} & \cdots & M^{(0)} \\ (M+1)^{(N)} & (M+1)^{(N-1)} & \cdots & (M+1)^{(0)} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ (M+N)^{(N)} & (M+N)^{(N-1)} & \cdots & (M+N)^{(0)} \end{vmatrix}$$

Para demostrar que el sistema [2.10] tiene solución única basta ver que el determinante anterior es distinto de cero, puesto que el factor multiplicador del mismo siempre lo es. Este determinante podemos escribirlo, con ayuda de E, el operador desplazamiento, en la forma

$$\begin{vmatrix} M^{(N)} & M^{(N-1)} & M^{(N-2)} & \cdots & M^{(0)} \\ EM^{(N)} & EM^{(N-1)} & EM^{(N-2)} & \cdots & EM^{(0)} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ E^{N}M^{(N)} & E^{N}M^{(N-1)} & E^{N}M^{(N-2)} & \cdots & E^{N}M^{(0)} \end{vmatrix}$$

Restando a cada fila la fila anterior y reiterando este proceso un total de N veces, de forma que en el proceso i restemos a la fila j la fila j-1, para j:N+1,N,N-1...,i+1, se obtiene

$$\begin{vmatrix} M^{(N)} & M^{(N-1)} & \cdots & M^{(1)} & M^{(0)} \\ (E-1)M^{(N)} & (E-1)M^{(N-1)} & \cdots & (E-1)M^{(1)} & (E-1)M^{(0)} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ (E-1)^{N}M^{(N)} & (E-1)^{N}M^{(N-1)} & \cdots & (E-1)^{N}M^{(1)} & (E-1)^{N}M^{(0)} \end{vmatrix}$$

donde el binomio (E-1) representa al operador diferencia. Ahora bien, puesto que $M^{(s)}$ es un polinomio en M de grado s, se verifican las siguientes igualdades

$$(E-1)^{j}M^{(k)} = 0 \qquad \text{si } j>k$$

y

$$(E-1)^{k}M^{(k)} = k! \ ,$$

por lo que todos los elementos del determinante por debajo de la diagonal secundaria son nulos, mientras que los de esa diagonal son distintos de cero. Por tanto, como queriamos demostrar, el sistema tiene solución única.

CAPITULO III

ORDENACIONES EN DISPERSION Y ALEATORIEDAD

III.1 INTRODUCCION

Este capítulo está dedicado a presentar los fundamentos que nos permitan abordar el objetivo principal de esta memoria, consistente en el estudio de propiedades de monotonía de la distribución hipergeométrica; centrándonos fundamentalmente en el comportamiento monótono de las medidas de dispersión y de aleatoriedad de la familia hipergeométrica, en su convergencia hacia la distribución binomial, objetivo éste que desarrollamos en el cuarto capítulo.

Para realizar un análisis sistemático de las propiedades de distribuciones de probabilidad se ha usado habitualmente alguno de los siguientes tres caminos:

a) Las propiedades de las distribuciones son medidas cuantitativamente. A veces, tales medidas son puramente heurísticas, pero una necesidad natural es que sean consistentes con algún orden parcial correspondiente a la misma propiedad. Ejemplos de lo anterior son las medidas clásicas de posición, dispersión, sesgo, curtosis, etc.

b) Podemos clasificar las distribuciones según los atributos en consideración; así, por ejemplo, pueden ser clasificadas las distribuciones univariantes en simétricas, con sesgo hacia la derecha y con sesgo hacia la izquierda, o

las distribuciones de tiempo de vida, en distribuciones de razón de fallo creciente, de razón de fallo decreciente, etc.

A menudo tal clasificación se basa en una ordenación parcial.

c) Dada una pareja de variables aleatorias X e Y con funciones de distribución F_X y F_Y, respectivamente, y considerada una determinada propiedad en estudio, podemos decir que $F_X \prec F_Y$, o bien, alternativamente, $X \prec Y$, si F_Y posee la propiedad en cuestión en igual o mayor grado que F_X.

Exigiremos que la relación " $\prec$ " sea reflexiva, es decir,

$$F_X \prec F_X \quad ,$$

y transitiva, es decir:

$$\text{si } F_X \prec F_Y \text{ y } F_Y \prec F_Z \text{, entonces } F_X \prec F_Z \quad .$$

Sin embargo, no es necesario que todas las parejas de distribuciones sean comparables. Habremos, por tanto, establecido en el conjunto de las distribuciones un orden parcial, - con más precisión, un preorden parcial, al no haberse exigido la propiedad antisimetrica - respecto de una determinada propiedad.

Este tercer camino será el utilizado para ordenar distribuciones en dispersión y en aleatoriedad, entendiendo que una determinada relación ordena distribuciones en

dispersión cuando existen medidas de dispersión que son consistentes con el citado orden, y análogamente para el caso de ordenaciones en aleatoriedad.

La idea de ordenar distribuciones respecto de una propiedad considerada, es más bien reciente. Las primeras incursiones en este campo las podemos encontrar en Mann y Whitney (1947), quienes introducen la noción de ordenación estocástica y de ordenación por posición. Al año siguiente, Birnbaum (1948) propone una ordenación por dispersión.

No obstante, como iremos viendo, la mayor densidad bibliográfica sobre este tema es posterior a 1980.

En el desarrollo de este capítulo nos vamos a centrar en las ordenaciones en dispersión y en aleatoriedad, y le dedicaremos una especial atención a los trabajos de Whitt (1985) y de Hickey (1982-1983-1986), por ser en los que, en mayor medida, vamos a fundamentar el último capítulo de la memoria.

III.2 ORDENACIONES EN DISPERSION

Es frecuente encontrarse con una pareja de variables aleatorias X e Y, con funciones de distribución F y G, respectivamente, en la que es intuitivamente claro que una de

ellas es más dispersa que la otra.

A partir de aquí surge la necesidad de formalizar la intuición, y han ido apareciendo diversas ordenaciones parciales de las distribuciones atendiendo a la dispersión.

Estas ordenaciones surgen por diferentes motivos: en problemas de inferencia no paramétrica; en Teoría de colas; para obtener desigualdades probabilísticas; en el estudio de las colas de una distribución, etc. Para definir tales ordenaciones se ha seguido, fundamentalmente, dos caminos:

Una primera forma de abordar el problema consiste en ordenar F y G según la dispersión respecto de algún valor central μ, que habitualmente es la media o la mediana. Estas ordenaciones comparan, estocásticamente, expresiones tales como $|X - \mu_x|$ y $|Y - \mu_y|$, o bien otras funciones convexas de $(X - \mu_x)$ y $(Y - \mu_y)$, donde μ_x y μ_y son los puntos centrales apropiados de F y G, respectivamente.

El segundo camino consiste en que F y G son ordenadas comparando funciones adecuadas de parejas de cuantiles. Por ejemplo, si F^{-1} y G^{-1} son las funciones cuantiles (inversas) de F y G, respectivamente, se comparan las siguientes parejas de expresiones:

$$F^{-1}(\beta) - F^{-1}(\alpha) \quad \text{con} \quad G^{-1}(\beta) - G^{-1}(\alpha), \qquad 0 < \alpha < \beta < 1,$$

o bien

$$F^{-1}(\beta) \,/\, F^{-1}(\alpha) \quad \text{con} \quad G^{-1}(\beta) \,/\, G^{-1}(\alpha), \qquad 0 < \alpha < \beta < 1.$$

Habría aún una tercera vía para ordenar F y G según

algunas propiedades de $G^{-1}F$, utilizadas comúnmente en teoría de la fiabilidad. En este caso, aunque el objetivo pueda no ser la ordenación en dispersión, también podría ser utilizado en este sentido.

ORDENACION ESTOCASTICA

La noción de ordenación estocástica para variables aleatorias es un concepto relativamente familiar y usado. De hecho, aparecerá en el desarrollo de este capítulo en relación con diferentes ordenaciones, bien de forma explícita, en la definición de la variabilidad condicional uniforme, o bien implícitamente, a través de conceptos equivalentes, en otras ordenaciones.

Se dice que X es estocásticamente menor o igual que Y, y se nota $X \underset{st}{\leq} Y$, si

$$P[X>t] \leq P[Y>t], \qquad -\infty<t<\infty \quad .$$

Si F y G son las funciones de distribución respectivas de X e Y , podemos también escribir

$$[3.1] \qquad F \underset{st}{\leq} G \iff F(t) \geq G(t), \qquad -\infty < t < \infty \quad .$$

Otras condiciones equivalentes para la ordenación

estocástica son las siguientes:

[3.2] a) $X \underset{st}{\leq} Y$.

[3.3] b) $E\ \Phi(X) \leq E\ \Phi(Y)$ para toda función Φ no decreciente tal que los valores esperados existen.

c) $\Phi(X) \underset{st}{\leq} \Phi(Y)$ para toda función Φ no decreciente.

La demostración de estas equivalencias pueden verse en el capítulo 17 del libro de Marshall y Olkin (1979).

ORDENACIONES EN DISPERSION RESPECTO DE UN VALOR CENTRAL

Birnbaum (1948) dice que una variable aleatoria X es más *puntiaguda* (menos variable) respecto de μ_x que otra variable aleatoria Y respecto de μ_y, donde μ_x y μ_y son sus respectivos valores centrales , si

[3.4] $P\left[\ |Y - \mu_y| \geq t\ \right] \geq P\left[\ |X - \mu_x| \geq t\ \right] \quad \forall\ t \geq 0$.

Birnbaum restringe su estudio a variables aleatorias con densidades simétricas respecto de μ_x y μ_y, respectivamente, y deduce algunas propiedades de esta ordenación. Es evidente, a partir de [3.4], que si las

probabilidades en las colas de Y son mayores que en las de X, en los valores centrales la densidad de X debe elevarse respecto de la de Y, lo cual justifica el nombre dado por Birnbaum a la relación.

Por otra parte, la expresión [3.4] se puede escribir en términos de ordenación estocástica como

$$[3.5] \qquad | X - \mu_x | \underset{st}{\leq} | Y - \mu_y | \quad .$$

Si ahora hacemos uso de la equivalencia entre [3.2] y [3.3], la expresión [3.5] también puede escribirse como

$$[3.5] \qquad E \left[g \left(| Y - \mu_y | \right) \right] \geq E \left[g \left(| X - \mu_x | \right) \right] ,$$

para toda función g no decreciente tal que los valores esperados en [3.5] existan.

En el caso en que $\mu_x = E [X]$ y $\mu_y = E [Y]$, se puede obtener fácilmente, a partir de [3.5], que la desviación media y la varianza de Y son mayores que las respectivas desviación media y varianza de X, por lo que la ordenación de Birnbaum satisface, en principio, características deseables para una ordenación en dispersión.

Otra ordenación parecida es la ordenación dilatación - término recogido de Shaked (1985) -, u ordenación en variabilidad, según nominación usada en Whitt (1985).

Diremos que una variable aleatoria X es menor o igual en variabilidad que otra variable aleatoria Y, y lo notaremos por $X \underset{v}{\leq} Y$, si

[3.6] $$E\,[h(X)] \leq E\,[h(Y)],$$

para toda función real convexa h, tal que los valores esperados estén definidos.

Alternativamente, podemos decir que $F \underset{v}{\leq} G$, siendo F y G las funciones de distribución de X e Y, respectivamente.

Obsérvese que si los valores esperados E[X] y E[Y] existen, entonces la condición [3.6] implica que ambos valores esperados coinciden. En este caso, podemos ver que [3.6] formaliza la idea de que Y es más dispersa que X. En efecto, cualquiera de las medidas conocidas del grado de dispersión, como son los momentos ordinarios y los momentos centrales - de orden par en ambos casos-, de la variable X, será menor que la correspondiente de Y, ya que los citados momentos se pueden expresar como valores esperados de funciones reales y convexas.

Para poder ordenar en variabilidad distribuciones con distintos valores esperados se puede considerar el orden

[3.7] $$(X - \mu_x) \underset{v}{\leq} (Y - \mu_y) \quad \text{si} \quad E\left[h(X - \mu_x)\right] \leq E\left[h(Y - \mu_y)\right],$$

para toda función real convexa h, para la cual los valores

esperados de [3.7] existen, siendo μ_x y μ_y los respectivos valores esperados de X e Y .

Alternativamente, para el caso de desigual media, se puede considerar :

a) Si $E[x] \leq E[Y]$, entonces $X \underset{<}{\leq} Y$ si $E\left[h(X)\right] \leq E\left[h(Y)\right]$, para toda función real, convexa, y no decreciente.

b) Si $E[X] \geq E[Y]$, entonces $X \underset{<}{\leq} Y$ si $E\left[h(X)\right] \leq E \left[h(Y)\right]$, para toda función real, convexa, y no creciente.

Consideremos ahora variables aleatorias X e Y con funciones de distribución respectivas F y G, absolutamente continuas, y con funciones de densidad f y g, respectivamente.

Otro camino usado para comparar la variabilidad de las variables aleatorias X e Y es a través de los cambios de signo de las funciones g - f o G - F . Este camino ha sido estudiado por Shaked (1980).

Vamos a definir, en primer lugar, el número de cambios de signo de una función. Sea a(x) una función definida en un subconjunto I de la recta real. El número de cambios cambios de signo de la función a en I se define por

[3.8] $$S(a) = \sup S\ [\ a(x_1),\ a(x_2),\ \ldots\ ,\ a(x_n)\ ]\ ,$$

donde $S\,[\,a(x_1),\,a(x_2),\ldots,a(x_n)\,]$ es el número de cambios de signo en la secuencia anterior, sin contar los términos ceros, y considerando que el supremo de [3.8] se extiende sobre todos los conjuntos de puntos $x_1 < x_2 < \ldots < x_m$, con $x_i \in I$ y con $m < \infty$.

Una condición que, razonablemente, se le debe exigir a $g - f$ para que g sea, en algún sentido, más variable que f es que

[3.9] $\quad S\,(\,g - f\,) = 2$, con secuencia de signos +, −, + .

La condición [3.9] implica que

[3.10] $\quad S\,(\,G - F\,) = 1$, con secuencia de signos +, − .

Si $E\,[X] \geq E\,[Y]$, entonces [3.10] implica que

$$\int_{-\infty}^{t} F(x)\,dx \;\leq\; \int_{-\infty}^{t} G(x)\,dx \quad , \quad \forall\, t \quad ,$$

o, equivalentemente, se podría afirmar que la condición

[3.11] $$E\left[\,h(X)\,\right] \;\leq\; E\left[\,h(Y)\,\right] \quad ,$$

se verifica para toda función real h, convexa y no creciente, para la que los valores esperados de [3.11] estén definidos.

Similarmente, si $E[X] \leq E[Y]$, entonces [3.10] implica que

$$\int_t^{\infty} \left[1 - F(x) \right] dx \leq \int_t^{\infty} \left[1 - G(x) \right] dx \quad , \quad \forall\, t \quad ,$$

o, equivalentemente, la condición [3.11] se verifica para toda función real h , convexa y no decreciente, para la que los valores esperados estén definidos.

Un caso de especial interés ocurre cuando E[X]=E[Y], ya que, en este caso, la condición [3.11] se verificaría para toda función h, real y convexa .

En el caso de igual media tendremos, por tanto, que

[3.12] $S(g - f)=2$, con secuencia +,-,+ , implica que $F \underset{v}{\leq} G$

Cualquiera de las condiciones [3.9] o [3.10], pueden ser utilizadas para definir un orden parcial en el conjunto de las distribuciones, que estaría incluido en la ordenación en variabililidad. No se trataría de ordenaciones equivalentes puesto que se puede construir ejemplos en los que $X \underset{v}{\leq} Y$ y, sin embargo, $S(g-f) > 2$.

En un trabajo posterior, Shaked (1982) realiza un amplio y detallado estudio de la ordenación según el número de cambios de signo, conectándola con la ordenación en dispersión que se establece a partir de la comparación de funciones adecuadas de parejas de cuantiles - ordenación que vamos a ver seguidamente -, y conectándola, asimismo, con la primitiva ordenación de Birnbaum (1948). Retomaremos, por tanto, más adelante, el trabajo de Shaked.

ORDENACIONES BASADAS EN LAS FUNCIONES CUANTILES

En términos generales, podemos decir que en la sección anterior las ordenaciones tienen en cuenta la dispersión relativa a un punto fijo que, habitualmente, es la media.

Ahora vamos a considerar ordenaciones basadas en la dispersión o *extensión* de la variable aleatoria a lo largo de su soporte.

Sean X e Y variables aleatorias, y F y G sus respectivas funciones de distribución. Saunders (1978) y Saunders y Moran (1978) consideraron parejas de distribuciones que satisfacen

$$[3.13] \qquad F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \qquad \text{para} \quad 0<\alpha<\beta<1$$

donde $F^{-1}(u) = \inf \{ x / F(x) \geq u \}$.

Cuando $F(0)=0$ y $G(0)=0$, consideraron también la desigualdad

$$\frac{F^{-1}(\beta)}{F^{-1}(\alpha)} \geq \frac{G^{-1}(\beta)}{G^{-1}(\alpha)} \quad , \text{ para cualesquiera } \quad 0<\alpha<\beta<1$$

La relación [3.13] permite a los citados autores

ordenar la familia de la distribución Gamma y la familia de cocientes de variables Gamma.

Fueron Lewis y Thompson (1981) quienes introdujeron el concepto de ordenación en dispersión, a partir de la relación [3.13], estableciendo un orden parcial en el espacio de las funciones de distribución de probabilidades.

En este sentido, se dirá que

$$F \underset{disp}{\leq} G \quad \Leftrightarrow \quad F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha),$$

para cualesquiera $0<\alpha<\beta<1$.

Alternativamente, podríamos notar $X \underset{disp}{\leq} Y$.

La expresión [3.13] indica que la diferencia entre cualesquiera dos percentiles de G, no es menor que la que existe entre los correspondientes percentiles de F.

Para distribuciones simétricas, la condición [3.13] podría escribirse

$$F^{-1}(\alpha) - F^{-1}(1/2) \underset{(\geq)}{\leq} G^{-1}(\alpha) - G^{-1}(1/2) \quad \text{siempre que } \alpha \underset{(\geq)}{\leq} 1/2$$

Es fácil ver que, para toda variable aleatoria X, y para cualquier constante $a > 1$,

$$a\,X \underset{disp}{\geq} X \quad .$$

Shaked (1982), como ya anunciamos, hace un detallado estudio de esta ordenación, relacionándola con el número de cambios de signo y obteniendo una serie de resultados interesantes.

Para evitar supuestas complicaciones en las demostraciones, impone unas restricciones de partida, que consisten en considerar funciones de distribución estrictamente crecientes y continuas en su soporte, exigiéndole a éste ser un intervalo. Sin embargo, la validez de los resultados obtenidos es general, como ponen de manifiesto Lynch, Mimmack y Proschan (1983), quienes extienden los resultados para cualquier funcion de distribucion.

Saunders (1984) abunda en el tema, y ofrece demostraciones alternativas a algunos de los resultados de Shaked (1982) y de Lynch, Mimmack y Proschan (1983).

Entre los resultados de Shaked (1982) destacaremos los siguientes:

3.1 TEOREMA:

$$F \underset{disp}{\leq} G \quad \Leftrightarrow \quad S(F_c - G) \leq 1 \,, \quad \forall c \in \mathbb{R} \,, \quad \text{y, en el caso de igualdad, la secuencia de signos es } -, + \,.$$

FACULTAD DE CIENCIAS BIBLIOTECA GRANADA

(Hacemos notar que con F_c se está indicando una traslación de F).

A partir de esta caracterización, Shaked hace ver que si $F \underset{disp}{\leq} G$, y si F y G son simétricas respecto del cero, entonces F es *más puntiaguda* respecto de cero que G, estableciendo, de esta manera, una conexión del orden en dispersión con el orden definido por Birnbaum (1948).

En el caso en que los soportes de F y G sean, en ambos casos, el intervalo $[0,\infty)$, Shaked obtiene el siguiente

3.2 TEOREMA:

$F \underset{disp}{\leq} G$ si y sólo si

[3.14] 1) $F(x) - G(x) \geq 0$, $\forall\, x \in (0,\infty)$, y

2) $S(F_c - G) \leq 1$, para cualquier $c>0$, y en caso de igualdad, la secuencia de signos es $-, +$.

Observemos que la condición [3.14] equivale a exigir la ordenación estocástica para las distribuciones F y G.

3.3 LEMA:

Sean F y G dos distribuciones absolutamente continuas, y sean f y g sus correspondientes

densidades. Si, dada c, $S(f_c - g) \leq 2$ y, en caso de igualdad, la secuencia de signos -, +, - ,se cumple entonces que $S(F_c - G) \leq 1$ y, en caso de igualdad, la secuencia de signos es -, + .

A partir de este lema y de los dos teoremas anteriores, Shaked (1982) demuestra los siguientes resultados:

3.4 TEOREMA:

Sean F y G dos distribuciones absolutamente continuas, cuyos soportes son intervalos, y sean f y g sus correspondientes densidades. Si $S(f_c - g) \leq 2$ para cualquier constante real c, y en caso de igualdad, la secuencia de signos es -, +, - , entonces $F \underset{disp}{\leq} G$.

3.5 TEOREMA:

Sean F y G dos distribuciones absolutamente continuas con soporte $[0,\infty)$, y sean f y g sus correspondientes densidades. Si $S(f_c - g) \leq 2$, para cualquier constante c positiva, con secuencia de signos -, +, - en caso de igualdad, y si $F \underset{st}{\leq} G$, entonces $F \underset{disp}{\leq} G$.

3.6 TEOREMA:

Si F y G son diferenciables en sus soportes, con densidades f y g , respectivamente, entonces $F \underset{disp}{\leq} G$ si y sólo si

$$g\left[G^{-1}(u)\right] \leq f\left[F^{-1}(u)\right] , \quad \forall\, u \in (0,1)$$

3.7 TEOREMA:

Sea I_F e I_G los soportes respectivos de F y G Entonces $F \underset{disp}{\leq} G$ si y sólo si existe una función

$$\phi : I_F \dashrightarrow I_G ,$$

tal que

[3.15] $$Y \overset{d}{=} \phi(X) ,$$

donde la notación $\overset{d}{=}$ indica igualdad en distribución, y

[3.16] si $x < x'$, entonces $\phi(x') - \phi(x) \geq x' - x$.

Aquí, $\phi = G^{-1}F$. Cuando ϕ es diferenciable, la condición [3.16] es equivalente a

$$\phi\,'(x) \geq 1$$

Puede observarse que las condiciones [3.15] y [3.16] son coherentes con la idea de que Y es más variable que X.

Finalmente, Shaked estudia también el caso en que X e Y son variables aleatorias positivas con probabilidad 1, y obtiene el siguiente resultado:

3.8 TEOREMA:

$$\log X \underset{disp}{\leq} \log Y \iff G^{-1}(\beta)/G^{-1}(\alpha) \geq F^{-1}(\beta)/F^{-1}(\alpha)\ ,$$

para cualesquiera $0 < \alpha < \beta < 1$.

De esta manera, se puede ordenar en dispersión parejas de logaritmos de variables aleatorias.

Queremos insistir en el hecho de que los resultados obtenidos por Shaked (1982) son generalizables a funciones de distribución en general, como ponen de manifiesto Lynch, Mimmack y Proschan (1983), y en una versión mejorada, Saunders (1984).

Posteriormente, el propio Shaked (1985) hace una recopilación de sus resultados, acomodando la formulación de los enunciados a la validez general de los mismos.

RELACION ENTRE LAS ORDENACIONES EN DISPERSION Y VARIABILIDAD

Recogemos de Whitt (1985) unos resultados que conectan las ordenaciones en dispersión y en variabilidad.

Sean X e Y variables aleatorias con funciones de distribución F y G , respectivamente. Si

$$F \underset{disp}{\leq} G \quad , \quad y \quad EX + c \leq EY \quad ,$$

siendo c una constante, entonces

$$[3.17] \qquad E\, h(X+c) \leq E\, h(Y) \quad ,$$

para toda función h , real, convexa, y no decreciente.

Si $EX + c \geq EY$, la relación [3.17] se verificaría para toda función real, convexa, y no creciente.

En otras palabras, si $X \underset{disp}{\leq} Y$, entonces $X \underset{v}{\leq} Y+c$ para alguna constante c, o bien, equivalentemente,

$$X \underset{disp}{\leq} Y \quad \Rightarrow \quad (X - EX) \underset{v}{\leq} (Y - EY) \quad .$$

En el otro sentido, es fácil construir ejemplos tales que $(X - EX) \underset{v}{\leq} (Y - EY)$, y sin embargo, $S(G - F) > 1$, por lo que no se verificaría que $X \underset{disp}{\leq} Y$.

Con este resultado se completa el dado en [3.12] .

RELACION ENTRE LA ORDENACION EN DISPERSION Y LA PROPIEDAD DE UNIMODALIDAD FUERTE

Diremos que una distribución H es dispersiva si

$$H * F \underset{disp}{\leq} H * G \quad , \quad \text{siempre que} \quad F \underset{disp}{\leq} G \ .$$

Con la notación "*" indicamos convolución.

La siguiente caracterización para las distribuciones dispersivas fue dada por Lewis y Thompson (1981):

3.9 TEOREMA:

Sea H una distribución no degenerada. Entonces, H es dispersiva si y sólo si H es absolutamente continua con densidad log-cóncava.

Queremos hacer notar que la log-concavidad de la densidad de H equivale a decir -Ibragimov (1956)- que H es fuertemente unimodal.

Recordemos que una distribución es fuertemente unimodal si su convolución con cualquier distribución unimodal es unimodal.

Lynch, Mimmack y Proschan (1983) ofrecen una demostración más sencilla del anterior teorema de

caracterización de las distribuciones dispersivas, usando la positividad total, camino éste cuyo uso parece aconsejable en problemas de cambios de signo.

Aceptando que F tenga una densidad de Lebesgue positiva sobre $\mathbb{R}$, Droste y Wefelmeyer (1985) demuestran la equivalencia de las siguientes tres proposiciones:

[3.18] (i) F es fuertemente unimodal

[3.19] (ii) $F * G \underset{disp}{\geq} F \quad , \quad \forall \; G \in \mathbb{D}$

[3.20] (iii) $F * G \underset{disp}{\geq} F \quad , \quad \forall \; G \in \mathbb{D}_o$

donde con $\mathbb{D}$ notamos el conjunto de las distribuciones sobre $\mathbb{R}$, y con $\mathbb{D}_o$ al conjunto de las distribuciones con masa igual a $1/2$ en el cero, y masa igual a $1/2$ en cualquier otro punto $a \in \mathbb{R}$.

El cuadro de equivalencias lo completa Klaassen (1985), demostrando que la condición

[3.21] $$F * G_1 \underset{disp}{\geq} F * G_2 \quad ,$$

para todo par de distribuciones G_1 y G_2 tales que $G_1 \underset{disp}{\geq} G_2$, es equivalente a las condiciones [3.18], [3.19], y [3.20].

Teniendo en cuenta la equivalencia entre [3.18], [3.19] y [3.20], y que Lewis y Thompson (1981) probaron que [3.18] es equivalente a [3.21], y puesto que las implicaciones

[3.21] ⇒ [3.19] ⇒ [3.20]

son triviales, a Klaassen (1985) , para completar el cuadro de equivalencias ,le basta con probar la implicación

[3.20] ⇒ [3.18]

Finalmente, queremos reseñar que en un reciente trabajo, Capéràa (1986) pone de manifiesto que la caracterización dada por Droste y Wefelmeyer (1985) para las distribuciones fuertemente unimodales es una consecuencia inmediata de un teorema de Pratt (1964).

RELACION ENTRE LA ORDENACION EN DISPERSION Y OTRAS ORDENACIONES

En este apartado vamos a relacionar la ordenación en dispersión con dos ordenaciones bien conocidas en teoría de la fiabilidad : la ordenación por colas y la ordenación "estrella".

Dadas dos distribuciones F y G , decimos que F es menor o igual que G en la ordenación por colas, y lo notamos $F \underset{c}{\leq} G$, si

[3.22] $G^{-1}F(x) - x$ es no decreciente en x ,

donde $G^{-1}(u) = \inf \{ x / G(x) \geq u \}$.

Este orden parcial ha sido usado por diversos autores, entre ellos Docksum (1969), en problemas de inferencia, para obtener cotas de eficiencia de test.

Deshpande y Kochar (1983) ponen de manifiesto la equivalencia entre la anterior ordenación y el orden en dispersión. En efecto, recordemos que una de las caracterizaciones del orden en dispersión era

$$[3.23] \quad F \underset{disp}{\leq} G \iff F^{-1}(\beta)-F^{-1}(\alpha) \leq G^{-1}(\beta)-G^{-1}(\alpha) \quad , \; 0<\alpha<\beta<1.$$

Si tomamos en [3.23] $\alpha=F(x)$ y $\beta=F(y)$, con $x \leq y$, es fácil ver que se obtiene la condición [3.22] .

Nosotros, por nuestra parte, queremos llamar la atención sobre el hecho de que una de las caracterizaciones dadas por Shaked (1982) para el orden en dispersión , concretamente la expresada por las condiciones [3.15] y [3.16], consiste, en esencia, en identificar el citado orden en dispersión con la ordenación por colas ; bien es cierto que este último orden no es citado de forma expresa por el autor.

Por otra parte, para distribuciones que cumplan que $F(0)=G(0)=0$, otra ordenación parcial definida por Shaked (1982) es la caracterizada por la condición

$$[3.24] \quad G^{-1}(\beta)/\, G^{-1}(\alpha) \geq F^{-1}(\beta)/\, F^{-1}(\alpha) \quad , \quad \text{para} \quad 0<\alpha<\beta<1 \; ;$$

condición que equivalía a decir que $\log X \underset{disp}{\leq} \log Y$.

Si en [3.24] hacemos $\alpha = F(x)$ y $\beta = F(y)$, con $x \leq y$, vemos que la condición [3.24] se verifica si y sólo si

[3.25] $\dfrac{G^{-1} F(x)}{x}$ es no decreciente en x.

La condición [3.25] es usada en la literatura para definir la ordenación "estrella", $F \underset{*}{\leq} G$ [Barlow y Proschan (1975)]. Es decir, para variables aleatorias X e Y positivas con probabilidad uno, con funciones de distribución respectivas F y G , tendremos que

$$\log X \underset{disp}{\leq} \log Y \iff F \underset{*}{\leq} G .$$

Finalmente, Deshpande y Kochar (1983) demuestran que si F y G son absolutamente continuas, con $F(0) = G(0) = 0$, y las correspondientes densidades son tales que $f(0) \geq g(0) > 0$, entonces

[3.26] $$F \underset{*}{\leq} G \Rightarrow F \underset{disp}{\leq} G .$$

Sathe (1984) demuestra que la única condición que es necesario exigir para que se cumpla [3.26] es que

$$\lim_{x \to 0+} \frac{G^{-1} F(x)}{x} \geq 1 .$$

ORDENACION EN DISPERSION Y DISTRIBUCIONES DE RAZON DE FALLO MONOTONA

Queremos, en este breve apartado, recoger los resultados de Bartoszewics (1985), que relacionan la ordenación en dispersión con las distribuciones de razón de fallo monótono.

Si las distribuciones F y G tienen por densidades a f y g , las respectivas funciones de razón de fallo vendrán definidas por

$r_F(t) = f(t)/[1 - F(t)]$ y $r_G(t) = g(t)/[1 - G(t)]$, $t \geq 0$.

Bartoszewics (1985) demuestra los siguientes teoremas:

3.10 TEOREMA:

Sean F y G distribuciones absolutamente continuas, con soporte $[0,\infty)$, tales que $F(0)=G(0)=0$.

Si $r_G(t) \leq r_F(t)$, $\forall$ $t \geq 0$, y F ó G son distribuciones de razón de fallo decreciente, entonces $F \underset{disp}{\leq} G$.

3.11 TEOREMA:

Sean F y G distribuciones absolutamente continuas con soporte $[0,\infty)$, tales que $F(0)=G(0)=0$.

Si $F \underset{disp}{\leq} G$ y F ó G son distribuciones de razón de fallo creciente, entonces $r_G(t) \leq r_F(t)$, $\forall$ $t \geq 0$

VARIABILIDAD CONDICIONAL UNIFORME

Vamos a desarrollar aquí un amplio comentario del reciente e interesante trabajo de Whitt (1985).

Este autor define una nueva relacion entre distribuciones, como es la ordenación en variabilidad condicional uniforme (OVCU). Relacionando la OVCU con las ordenaciones por cambios de signo, demuestra que la ordenación en variabilidad condicional uniforme implica otras ordenaciones conocidas.

Finalmente, a través de la log-concavidad relativa de una distribución respecto de otra, obtiene una condición suficiente para la OVCU. El índice de log-concavidad relativa va a ser el instrumento aconsejable para probar la ordenación en variabilidad condicional uniforme.

Consideraremos variables aleatorias X e Y cuyas respectivas funciones de distribución F y G, son absolutamente

continuas, y con funciones de densidad o funciones de masa de probabilidad f y g , respectivamente. (La función de masa de probabilidad se obtiene, como es usual, cuando se trata de continuidad absoluta respecto de una medida de conteo sobre los enteros).

Se dice que f es menor o igual en variabilidad condicional uniforme que g , y se notará por $X \underset{uv}{\leq} Y$, o bien, alternativamente, por $f \underset{uv}{\leq} g$, si se verifican las siguientes tres condiciones:

(i) soporte(f) $\subseteq$ soporte(g) .

(ii) La razón $\frac{f(t)}{g(t)}$ es unimodal, entendiendo por moda un supremo.

(iii) f y g no están ordenadas estocásticamente.

El siguiente teorema relaciona la variabilidad condicional uniforme con los cambios de signo:

3.12 TEOREMA:

Admitiendo que soporte(f) $\subseteq$ soporte(g) , $f \underset{uv}{\leq} g$ si y sólo si $S(cf - g) \leq 2$ para toda constante positiva, verificándose la igualdad cuando c=1 , y en este caso, la secuencia de signos es -, +, - .

Demostración:

Condición necesaria: supongamos que $f \underset{uv}{\leq} g$. Al ser f/g unimodal en el sentido de supremo, también lo será cf/g , para $c>0$. Por tanto, $S(cf/g - 1) \leq 2$, o lo que es igual, $S(cf - g) \leq 2$.

Para $c=1$, $S(f - g) = 2$ ya que si $S(f - g) < 2$, f y g estarían ordenadas estocásticamente. Se ve fácilmente que, en este caso, la secuencia de signos es $-, +, -$.

Condición suficiente: supongamos ahora que $f \underset{uv}{\leq} g$ no fuese cierto. Si f y g estuviesen ordenadas estocásticamente, entonces $S(f -g) < 2$. Supongamos, entonces, que la condicion que falla es la unimodalidad de f/g y que, sin embargo, $S(f - g) = 2$ con secuencia de signos $-, +, -$, y comprobaremos que se llega a una contradicción.

En efecto, si f/g no es unimodal, habrá tres puntos $t_1 < t_2 < t_3$, tales que

$$\frac{f}{g}(t_1) > \frac{f}{g}(t_2) < \frac{f}{g}(t_3) \quad .$$

Sea c una constante tal que

$$\frac{f}{g}(t_2) < c^{-1} < \min \left[\frac{f}{g}(t_1), \frac{f}{g}(t_2) \right] .$$

Entonces,

$$S(f/g - c^{-1}) > 2 \ , \quad \text{o bien} \quad S(f/g - c^{-1}) = 2,$$

y en el caso de igualdad la secuencia de signos sería +, −, + , y sin embargo, $f/g - c^{-1}$ tiene el mismo número de cambios de signo que $cf - g$, como se puede comprobar fácilmente.

Podemos añadir, por nuestra parte, como corolario al anterior teorema, que - teniendo en cuenta el resultado [3.12] -, si las variables aleatorias tienen igual media,

[3.27] $$f \underset{uv}{\leq} g \quad \Rightarrow \quad F \underset{v}{\leq} G \ .$$

Es conveniente resaltar el hecho de que la ordenación en variabilidad condicional uniforme no es, a pesar de la denominación, relación de orden, puesto que no se cumple la propiedad transitiva. En Whitt (1985) puede verse un ejemplo en el que no se verifica la transitividad.

Por otra parte, ya hemos visto anteriormente que el orden en dispersión, $\underset{disp}{\leq}$, es más fuerte que el orden en variabilidad, $\underset{v}{\leq}$. Sin embargo, la ordenación en variabilidad condicional uniforme y la ordenación en dispersión no son comparables, como se pone de manifiesto con los contraejemplos adecuados, en el citado artículo.

Supongamos ahora que nos fijamos en las distribuciones condicionadas que se obtienen cuando X e Y

toman valores en un conjunto A, donde A es tal que

$$P[X \in A] > 0 \qquad \text{y} \qquad P[Y \in A] > 0$$

Lo natural es que el conjunto A sea un intervalo (continuo o discreto), pero no es necesario. Así, por ejemplo, si f y g tienen soportes sobre los enteros, A podría tratarse del conjunto de los enteros pares.

Las distribuciones condicionadas son absolutamente continuas, con densidades o funciones de masa de probabilidad

$$f_A(t) = \frac{f(t)}{P[X \in A]} \quad , \quad g_A(t) = \frac{g(t)}{P[Y \in A]} \quad , \text{ con } t \in A$$

Llamaremos X_A e Y_A a las correspondientes variables aleatorias.

Es conocido que la unimodalidad se preserva bajo multiplicación escalar de las densidades y bajo la restricción a un subconjunto; por tanto, la unimodalidad de $f(t)/g(t)$ se puede extender a $f_A(t)/g_A(t)$.

Naturalmente, entra dentro de lo posible que $\frac{f_A(t)}{g_A(t)}$ sea monótona, lo cual correspondería al orden de razón monótona de verosimilitud, que notaremos por $\underset{mr}{\leq}$.

El siguiente teorema resume las implicaciones que el orden $\underset{uv}{\leq}$ establece para las comparaciones estocásticas de las distribuciones condicionadas.

3.13 TEOREMA:

Si $X \underset{uv}{\leq} Y$, entonces, para cada subconjunto A del soporte de Y, se verifica uno y sólo uno de los siguientes casos:

(i) f/g es constante en A, por lo que $X_A \overset{d}{=} Y_A$;

(ii) f/g es creciente en A, por lo que $X_A \underset{mr}{\geq} Y_A$;

(iii) f/g es decreciente en A, por lo que $X_A \underset{mr}{\leq} Y_A$;

(iv) f/g es unimodal en A, pero no monótona, y $S(g_A-f_A)=1$ con secuencia de signos +, - ; en este caso, $X_A \underset{st}{\geq} Y_A$;

(v) f/g es unimodal en A , pero no monótona, con $S(g_A-f_A)=1$, con secuencia de signos -, + ; en este caso, $X_A \underset{st}{\leq} Y_A$;

(vi) f/g es unimodal en A, pero no monótona, con $S(g_A-f_A)=2$,con secuencia de signos +, -, + ; en este caso, $X_A \underset{uv}{\leq} Y_A$.

Naturalmente, los casos (i), (ii) y (iii), sólo pueden darse si A cae enteramente a un lado de la moda de f/g. Por otra parte, el caso (vi) es el único en el que se preserva la ordenación en variabilidad condicional uniforme.

3.14 COROLARIO:

Supongamos que $X \underset{uv}{\leq} Y$.

(i) $E[X/X\in A] \leq E[Y/Y\in A] \Rightarrow E[h(X)/X\in A] \leq E[h(Y)/Y\in A]$,
para toda función convexa h, no decreciente.

(ii) $E[X/X\in A] \geq E[Y/Y\in A] \Rightarrow E[h(X)/X\in A] \leq E[h(Y)/Y\in A]$,
para toda función convexa h, no creciente.

(iii) $E[X/X\in A] = E[Y/Y\in A] \Rightarrow E[h(X)/X\in A] \leq E[h(Y)/Y\in A]$,
para toda función convexa h .

Hacemos notar que tanto el teorema como el corolario siguen siendo válidos si condicionamos sobre $X\in A$ e $Y\in B$, con $A\subset B$. En este caso, el soporte de X_A estaría incluido en el de Y_B .

LOG-CONCAVIDAD RELATIVA

Puede verse en el capítulo V del libro de Keilson (1979) que una condición suficiente para que f/g sea unimodal es que f/g sea log-cóncava.

Vamos a considerar, en lo que sigue, que los soportes de f y g son intervalos de $\mathbb{R}$, o conjuntos de

enteros consecutivos en el caso de funciones de masa de probabilidad.

Una función de densidad es log-cóncava si

$$f[\lambda x + (1-\lambda)y] \geq f(x)^{\lambda} f(y)^{1-\lambda} \quad \text{con} \quad 0 \leq \lambda \leq 1 ,$$

para todo x e y pertenecientes al dominio ; o de forma equivalente, si

$$f(y_2 - x_2)\, f(y_1 - x_1) \geq f(y_2 - x_1)\, f(y_1 - x_2)$$

para $x_1 < x_2$ e $y_1 < y_2$.

Una función de masa de probabilidad f es log-cóncava si

$r_f(k) \geq 1$ para todo k perteneciente al soporte de f,

donde $r_f(k)$ es el índice de log-concavidad, definido por

[3.28] $$r_f(k) = \frac{f^2(k)}{f(k-1)f(k+1)}$$

Como ya sabemos, el concepto de log-concavidad es equivalente al de unimodalidad fuerte.

Diremos que una función de densidad, o función de masa de probabilidad, f es log-cóncava relativa a otra g , y se notará

$f \underset{lc}{\leq} g$, si f/g es log-cóncava.

En el caso de densidades, f/g es log-cóncava si el soporte de f está contenido en el de g , y

$$r_f(x_1, x_2, y_1, y_2) \geq r_g(x_1, x_2, y_1, y_2)$$

para todo $x_1 < x_2$ y para todo $y_1 < y_2$, donde

$$r_f(x_1, x_2, y_1, y_2) = \frac{f(y_2 - x_2)\, f(y_1 - x_1)}{f(y_2 - x_1)\, f(y_1 - x_2)} .$$

Para el caso de funciones de masa de probabilidad, f es log-cóncava relativa a g, si

(i) soporte(f) $\subseteq$ soporte(g) y

(ii) $r_f(k) \geq r_g(k)$, $\forall k \in$ soporte(f)

con $r_f(k)$ y $r_g(k)$ definidos según [3.28].

Es de destacar que la log-concavidad relativa nos va a proporcionar un camino cómodo para el estudio de la variabilidad condicional uniforme, y a través de ésta, como ya hemos indicado en el corolario [3.27], nos permitirá abordar la ordenación en variabilidad.

A este respecto, el resultado clave es el siguiente:

[3.29] Si $f \underset{lc}{\leq} g$ y $S(f - g) = 2$, entonces $f \underset{uv}{\leq} g$.

En efecto, si $f \underset{lc}{\leq} g$ entonces

(i) f/g es unimodal

(ii) soporte(f) $\subseteq$ soporte(g) ,

y si $S(f - g) = 2$ entonces f y g no están ordenadas estocásticamente; cumpliendose , por tanto, la tercera condición exigida para que $f \underset{uv}{\leq} g$.

Por otra parte, y a diferencia de lo que ocurre con la variabilidad condicional uniforme, la log-concavidad relativa es transitiva.

Para finalizar, vamos a enunciar aquí una serie de propiedades adicionales recogidas del mencionado trabajo de Whitt(1985).

(i) Una densidad en $[0,\infty)$ es log-cóncava si y sólo si es log-cóncava relativa respecto de la densidad exponencial.

(ii) Una función de masa de probabilidad sobre los enteros no negativos es log-cóncava si y sólo si es log-cóncava relativa respecto de la distribución geométrica.

(iii) En el caso de intervalos [a,b], o conjuntos de enteros consecutivos, la distribución que servirá de referencia será la uniforme, continua y discreta, respectivamente.

(iv) La log-concavidad se preserva bajo convolución, y la log-convexidad bajo mixturas.

(v) En el caso de la variabilidad condicional uniforme, no se cumple que $X \underset{uv}{\leq} aX$ para todo $a > 1$.

(vi) Sean f, g y f_n , g_n con $n \geq 1$, funciones de masa de probabilidad con soporte sobre los enteros no negativos. Entonces

1.- $f_n(k) \dashrightarrow f(k)$ cuando $n \to \infty$, para todo k, si y sólo si $f_n(0) \dashrightarrow f(0)$, $f_n(1) \dashrightarrow f(1)$, y $r_{f_n}(k) \dashrightarrow r_f(k)$, para todo $k \geq 1$.

2.- Si $f_n \underset{uv}{\leq} g_n$ para todo n, $f_n(k) \dashrightarrow f(k)$, y $g_n(k) \dashrightarrow g(k)$, cuando $n \to \infty$, para todo k, entonces $f \underset{uv}{\leq} g$.

(vii) Si $f \underset{uv}{\leq} g$, entonces $f \underset{uv}{\leq} [\, pf + (1-p)g \,]$, con $0 < p < 1$.

(viii) $\underset{uv}{\leq}$ y $\underset{lc}{\leq}$ no se preservan bajo convolución.

III.3 ORDENACION EN ALEATORIEDAD

La posibilidad de ordenar parcialmente las distribuciones en aleatoriedad es desarrollada por Hickey.

Hickey (1982 y 1983) introduce la ordenación en aleatoriedad para distribuciones discretas, fundamentándose en la Teoría de mayoración de vectores recogida de Marshall y Olkin (1979).

Seguidamente, Hickey (1984) extiende parte de sus resultados al caso continuo; y, finalmente, en su trabajo más reciente, Hickey (1986) relaciona la ordenación en aleatoriedad con la ordenación en dispersión.

Definamos, en primer lugar, el concepto de mayoración de vectores.

Sean dos vectores , $\vec{x}$ e $\vec{y}$, de un espacio de dimensión n , de componentes respectivas x_i e y_i , con i tomando valores desde 1 hasta n, y sean $x_{[i]}$ e $y_{[i]}$, con i = 1,...,n , sus respectivas componentes ordenadas decrecientemente, es decir,

$$x_{[1]} \geq x_{[2]} \geq ..... \geq x_{[n]} ,$$

y análogamente para el vector $\vec{y}$.

Se dice que el vector $\vec{x}$ está mayorado por el

vector $\vec{y}$, y se nota $\vec{x} \underset{m}{\leq} \vec{y}$, si

$$\text{(i)} \quad \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

[3.30] $$\text{(ii)} \quad \sum_{i=1}^{j} x_{[i]} \leq \sum_{i=1}^{j} y_{[i]} , \qquad j = 1,2,\ldots, n-1.$$

Si algunas de las desigualdades de [3.30] es estricta, se dice que la mayoración también lo es, y se notará por $\vec{x} \underset{m}{<} \vec{y}$.

Esta terminología fue introducida por Hardy, Littlewood y Polya (1951); nosotros la recogemos del libro de Marshall y Olkin (1979), pg. 7.

Definiciones equivalentes son:

(i) $\vec{x} \underset{m}{\leq} \vec{y}$

(ii) $\vec{x} = \vec{y}\, P$, siendo P una matriz doblemente estocástica.

[3.31] (iii) $\sum_i \phi(x_i) \leq \sum_i \phi(y_i)$,para toda función ϕ , continua y convexa.

En la definición de mayoración de vectores subyace la idea intuitiva de que las componentes del vector mayorado son menos dispersas o, equivalentemente, presentan mayor similitud que las componentes del vector que mayora.

Las funciones que preservan el orden de mayoración fueron estudiadas sistemáticamente por Schur y, en su honor, a tales funciones se las denomina funciones Schur-convexas.

Una función $\phi : \mathbb{R}^n \dashrightarrow \mathbb{R}$ es Schur-convexa si

$$x \underset{m}{\leq} y \Rightarrow \phi(x) \leq \phi(y) .$$

Si además

$$x \underset{m}{<} y \Rightarrow \phi(x) < \phi(y) ,$$

diremos que ϕ es estrictamente Schur-convexa.

Una función ϕ es Schur-cóncava si $-\phi$ es Schur-convexa.

3.15 PROPOSICION:

Si ϕ es simétrica y convexa, entonces ϕ es Schur-convexa.

Entenderemos que ϕ es simétrica cuando $\phi(x)=\phi(x\Pi)$, siendo $x\Pi$ el vector resultante de cualquier permutación de las componentes del vector x .

En el libro de Marshall y Olkin (1979), pg.7, se indica que esta importante proposición fue, en esencia,

probada por Schur, si bien este autor restringió el dominio de ϕ a $\mathbb{R}^n_{++}$, donde $\mathbb{R}^n_{++} = \{ (x_1, x_2, \ldots, x_n) / x_i > 0 , \forall i \}$.

3.16 PROPOSICION:

Si ϕ es simétrica y estrictamente convexa sobre conjuntos de la forma $\{ z / \Sigma z_i = c \}$, entonces ϕ es estrictamente Schur-convexa.

Esta proposición está igualmente demostrada en el capítulo tercero del mencionado libro de Marshall y Olkin , capítulo que está dedicado al estudio, en profundidad, de las funciones Schur-convexas.

Las funciones Schur-convexas, por preservar el orden de mayoración, son las funciones adecuadas para ser utilizadas como medidas de la desigualdad, ya que la idea intuitiva de que las componentes de un vector presentan mayor similitud que las de otro vector, subyace, como dijimos, en la definición del orden de mayoración de vectores.

La idea de desigualdad o, alternativamente, de similitud, surge en muy variados contextos; uno de ellos, es el campo de la Teoría de la Información. En este contexto, toma forma en los conceptos, todos ellos intercambiables, de aleatoriedad, incertidumbre, y falta de información.

La Teoría matemática de la Información desarrollada al margen de la Teoría de mayoración - veanse los libros de Rényi (1976) y Gil (1981) -, nos ofrece la posibilidad de comparar distribuciones de probabilidad, especialmente en el caso discreto, utilizando para ello determinadas medidas de aleatoriedad.

Sea ξ una variable aleatoria discreta que toma valores x_i, con $i=1,\ldots,n$, con probabilidades $p(x_i)>0$, para todo i, siendo $\sum_{i=1}^{n} p(x_i) = 1$.

Entre las medidas más usuales de la aleatoriedad están:

a) Entropia de Shannon, definida por

$$H(\xi) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)$$

b) Energía informacional de Onicescu:

$$\aleph(\xi) = \sum_{i=1}^{n} p(x_i)^2$$

c) Información de Rényi de orden α :

$$I_\alpha(\xi) = \frac{1}{1-\alpha} \log \sum_{i=1}^{n} p(x_i)^\alpha \quad , \text{ con } \alpha > 1 .$$

Hickey (1982 y 1983) se da cuenta - y en ello radica su mayor mérito - que la Teoría de mayoración proporciona el

soporte idóneo para comparar en aleatoriedad distribuciones discretas de probabilidad, al margen de medidas concretas, de forma que cualquier medida razonable de la aleatoriedad, entre ellas las tres citadas, sea consistente con la ordenación establecida. En este sentido puede darse la siguiente definición:

3.17 DEFINICION:

Dadas dos distribuciones discretas, P y Q , con probabilidades puntuales $\{p_i\}$ y $\{q_i\}$, y con i=1,2,..., diremos que P es al menos tan aleatoria como Q , si los vectores $\vec{p}$ y $\vec{q}$, cuyas componentes respectivas son $\{p_i\}$ y $\{q_i\}$ y las componentes nulas necesarias para que ambos vectores tengan la misma dimensión, verifican que $\vec{p} \underset{m}{\leq} \vec{q}$.

Si $\vec{p} \underset{m}{<} \vec{q}$, se dirá que P es más aleatoria que Q.

Hickey (1982) da la siguiente definición:

3.18 DEFINICION:

Una función continua ϕ , que toma valores reales sobre la clase de las distribuciones discretas, $\mathbb{P}$, es una medida de la aleatoriedad siempre que sea simétrica y cóncava, siendo la concavidad estricta sobre la subclase

de distribuciones que tienen sólo un número finito de probabilidades positivas.

Como vemos, Hickey lo que hace es exigir a las medidas de aleatoriedad que sean Schur-cóncavas (estrictamente Schur-cóncavas en el caso de que las distribuciones tengan sólo un número finito de probabilidades positivas), o lo que es igual, que sus opuestas sean consistentes con el orden de mayoración establecido entre los vectores de probabilidad de las distribuciones.

De acuerdo con la anterior definición, la entropía de Shannon y las medidas de información de Rényi serían medidas de aleatoriedad, mientras que en el caso de la energía informacional de Onicescu, lo seria su opuesta. Se demuestra a partir de la proposición C.1, pg.44, del libro de Marshall y Olkin (1979).

Por otra parte, es fácil ver que una distribución degenerada (es decir, cuando la probabilidad está concentrada en un sólo punto) es menos aleatoria que cualquier distribución no degenerada, y también que la distribución uniforme que toma n valores es más aleatoria que cualquier otra distribución que tome al menos n valores con probabilidad mayor que cero.

El siguiente teorema ofrece una condición suficiente para que ϕ tome valores finitos en todo $\mathbb{P}$.

3.19 TEOREMA

Sea u_n , con $n=1,2,....$, la distribución uniforme que toma n valores, y sea ϕ una medida de aleatoriedad.

Si la sucesión creciente $\{\phi(u_n)\}$ converge a $k<\infty$, entonces

$$\phi (P) \leq k \quad , \quad \text{para todo} \quad P \in \mathbb{P} .$$

ALGUNOS RESULTADOS GENERALES SOBRE ALEATORIEDAD

Exponemos aquí varios resultados sobre aleatoriedad, recogidos, todos ellos, de Hickey (1983).

(i) Sea ϕ una medida de aleatoriedad, y sean X e Y variables aleatorias discretas, distribuidas conjuntamente. Se verifica

[3.32] $$\phi (X) \geq \phi (X / Y) .$$

Si X e Y son independientes, entonces en [3.32] se verifica la igualdad. Para espacios muestrales finitos, o bien para espacios muestrales infinitos, con la condición de que la medida ϕ tome sólo valores finitos, entonces se verifica también el recíproco; es decir,

$$\phi\,(X) = \phi\,(X\,/\,Y) \quad \Rightarrow \quad X \text{ e } Y \text{ son independientes.}$$

En el caso finito dimensional, este resultado es una simple consecuencia de la concavidad estricta de ϕ . En el caso infinito se sigue de la desigualdad de Jensen.

(ii) La entropía de Shannon pertenece a la clase de medidas de aleatoriedad que tienen la forma

[3.33] $$\phi\,(P) = \sum_{i=1}^{\infty} g(p_i) \quad ,$$

siendo g una función continua y estrictamente cóncava y siendo $g(0)=0$.

Se sabe que la entropia admite una interpretacion como valor esperado : es el valor esperado de la variable aleatoria $-\log\,(p_i)$. Una interpretación similar puede darse para cualquier medida de la forma [3.33].

En efecto, si g es estrictamente cóncava, y $g(0)=0$, tendremos que

[3.34] $$g(ax) \geq a\,g(x) \text{ , con } x > 0 \text{ , } \quad \forall\, a \in [0,1] \text{ .}$$

Para comprobarlo basta con tener en cuenta que si g es cóncava, entonces

$$g(ax+by) \geq a\,g(x) + b\,g(y) \text{ , } \quad \forall\, a,b \geq 0 \text{ tales que } a+b=1.$$

En el caso particular que y=0, se obtendrá [3.34] y de ahí se sigue que

[3.35] $$\frac{g(ax)}{ax} \geq \frac{g(x)}{x} \quad , \quad \text{para} \quad x > 0 \ .$$

Tomando $u(x) = g(x)/x$, podemos escribir

$$\phi(P) = \sum_{i=1}^{\infty} g(p_i) = \sum_{i=1}^{\infty} p_i \, u(p_i) \ .$$

donde u es una función decreciente puesto que se verifica [3.35]. En conclusión, cualquier medida de la forma [3.33] puede ser interpretada como el valor esperado de la variable aleatoria que toma los valores $u(p_i)$, siendo u una función decreciente.

(iii) Para cualquier variable aleatoria discreta X se verifica que X es al menos tan aleatoria como h(X), para cualquier función real h . Precisando más: en el caso que h sea inyectiva, como es obvio, la aleatoriedad será la misma ya que los vectores de probabilidad tendrán las mismas componentes, pero en el caso de que no se cumpla la inyectividad, X será más aleatoria que h(X), como puede concluirse a partir de la definición de mayoración de vectores.

(iv) Sean X e Y dos variables aleatorias discretas. La distribución de X+Y es más aleatoria que la de X y que la de

Y, salvo que una de ellas, por ejemplo Y, sea degenerada, en cuyo caso X+Y y X tienen la misma aleatoriedad.

(v) Sea P_θ una familia de distribuciones discretas con un parámetro θ que toma valores en $\Theta \subset \mathbb{R}$. Diremos que θ es un parámetro de incertidumbre si P_{θ_1} es más aleatoria que P_{θ_2} para todo $\theta_1 < \theta_2$ o para todo $\theta_1 > \theta_2$.

Para la distribución binomial B(n,p) se demuestra que, para un valor fijo de p, el parámetro n es un parámetro de incertidumbre. Los parámetros de las distribuciones de Poisson y geométrica son , asimismo, de incertidumbre.

ORDENACION EN ALEATORIEDAD EN EL CASO CONTINUO

En Hickey (1984) el autor extiende, en la medida de lo posible, sus resultados al caso continuo.

Consideraremos, en adelante, variables aleatorias X e Y , con funciones de distribución absolutamente continuas y con funciones de densidad f y g , respectivamente.

Las dos siguientes condiciones son equivalentes :

[3.36] $$\int_{-\infty}^{\infty} u[g(x)]\,dx \geq \int_{-\infty}^{\infty} u[f(x)]\,dx ,$$

para toda función u, continua y cóncava.

[3.37] $$\int_{-\infty}^{t} g^{*}(x)\,dx \;\leq\; \int_{-\infty}^{t} f^{*}(x)\,dx \quad , \quad \forall\, t \ ,$$

donde f^{*} y g^{*} son las reordenaciones decrecientes de f y g, respectivamente, es decir,

$$f^{*}(x) = \sup \{c:\ m(c) > x\}, \text{ para } x > 0,$$

con $m(c) = \mu\{\ x:\ f(x) > c\ \}$, y donde μ indica la medida de Lebesgue.

Hickey nos indica que la demostración de la equivalencia puede verse en Hardy, Littlewood y Polya (1929).

Si estas condiciones se verifican se dirá que g está mayorada por f, y se notara $g \underset{m}{\leq} f$.

Podemos observar que las condiciones [3.36] y [3.37] serían, para el caso continuo, las condiciones análogas a [3.31] y [3.30], respectivamente.

Como en el caso discreto, si $g \underset{m}{\leq} f$ podemos considerar que existe una mayor similitud en los valores tomados por g.

3.20 DEFINICION:

Si $g \underset{m}{\leq} f$ diremos que g es al menos tan aleatoria como f y lo notaremos $f \underset{a}{\leq} g$ .Si $g \underset{m}{\leq} f$, pero f no está mayorada por g, se dirá que g es más aleatoria que f, y

se notará por $f \underset{a}{<} g$, o, indistintamente, por $g \underset{m}{<} f$.

Hacemos notar, a este respecto, que para distintas funciones de densidad es posible que, simultáneamente, $g \underset{m}{\leq} f$ y $f \underset{m}{\leq} g$. Por ejemplo, si $f(x)=g(x-a)$, para alguna constante a, entonces, a partir de [3.37], se puede concluir que $g \underset{m}{\leq} f$ y $f \underset{m}{\leq} g$, simultáneamente.

Al igual que en el caso discreto, el grado de aleatoriedad de una función de densidad podrá ser medido por una función continua ϕ que sea Schur-cóncava, es decir,

$$\phi[g(x)] \geq \phi[f(x)] \quad \text{siempre que} \quad g \underset{m}{\leq} f$$

El siguiente resultado de Hickey (1984) conecta la aleatoriedad con los cambios de signo.

3.21 TEOREMA:

Sean f y g funciones de densidad definidas para $x \geq 0$, continuas y no crecientes en x. Si $f(x)-g(x)$ cambia de signo sólo una vez cuando x va de cero a infinito, y $f(0)$ es mayor que $g(0)$, entonces g es más aleatoria que f.

Se demuestra fácilmente viendo que se verifica la condición [3.37].

3.22 COROLARIO:

Consideremos f y g funciones de densidad continuas y simétricas respecto de a y b, respectivamente. Si f(x) y g(x) son no crecientes en x>a y x>b, respectivamente, y f(a) es mayor que g(b), y además f(x-a)-g(x-b) tiene dos cambios de signo, entonces g es más aleatoria que f.

Con sus correspondientes modificaciones, este teorema también se verifica en el caso discreto.

Finalmente, al igual que en el caso discreto, si X e Y son independientes, se verifica que X+Y es mas aleatoria que X y más aleatoria que Y.

RELACION ENTRE LA ORDENACION EN ALEATORIEDAD Y EN DISPERSION

En su más reciente trabajo, Hickey (1986) relaciona los conceptos de ordenación en aleatoriedad y de ordenación en dispersión y en variabilidad.

Todos los resultados vienen dados para el caso continuo, aunque el autor indica que algunos son, asimismo, ciertos en el caso discreto.

Los dos siguientes teoremas relacionan las ordenaciones en aleatoriedad, dispersión y variabilidad, en cuanto que presentan un cierto comportamiento común.

3.23 TEOREMA:

Tanto si $f \underset{disp}{\leq} g$, como si $g \underset{a}{\leq} f$, se verifica que

$$\mu\langle x: f(x)>0\rangle \leq \mu\langle x: g(x)>0\rangle \ .$$

3.24 TEOREMA:

Si $f \underset{v}{\leq} g$, y los soportes de f y g están constituidos por un sólo intervalo, entonces

$$\mu\langle x: f(x)>0\rangle \leq \mu\langle y: g(y)>0\rangle$$

El resultado fundamental de Hickey (1986) consiste en probar que, bajo ciertas condiciones, la ordenación en dispersión es más fuerte que la ordenación en aleatoriedad.

Llamemos $\mathcal{U}$ a la familia de densidades unimodales, esto es, $g \in \mathcal{U}$ si $\langle x: g(x) \geq c\rangle$ es un intervalo para todo $c>0$.

3.25 TEOREMA:

Sean f y g funciones de densidad con $g \in \mathcal{U}$. Entonces,

$$f \underset{disp}{\leq} g \quad \Rightarrow \quad f \underset{a}{\leq} g$$

El recíproco no es cierto ya que, como hemos visto anteriormente, la convolución de f y g está mayorada por f; por el contrario, Lewis y Thompson (1981) observaron que la relación $f \underset{disp}{\leq} f*g$ no se cumple necesariamente, por lo que, en general,

$$f \underset{a}{\leq} g \quad \text{no implica que} \quad f \underset{disp}{\leq} g \ .$$

Shaked (1982) probó que la ordenación en dispersión implicaba ordenación en variabilidad; por el contrario, Hickey (1986) hace ver que entre la ordenación en aleatoriedad y en variabilidad existe poca relación. Incluso si se impone la restricción de que ambas distribuciones sean unimodales, ni la ordenación en aleatoriedad implica ordenación en variabilidad, ni recíprocamente.

CAPITULO IV

PROPIEDADES DE MONOTONIA

IV1 INTRODUCCION

En este capítulo vamos a aplicar los resultados recopilados en el capítulo anterior, al estudio de la sucesión de hipergeométricas $H_m(mN,p,n)$ que, como sabemos, converge a la binomial $B(n,p)$. Este estudio se va a centrar fundamentalmente en dos aspectos: el comportamiento de la dispersión y el de la aleatoriedad en la citada sucesión, que desarrollaremos en las secciones segunda y tercera, respectivamente.

Comenzaremos analizando el comportamiento monótono creciente de las medidas usuales de dispersión, a la vez que ponemos de manifiesto que la manipulación de estas medidas en el caso de la distribución hipergeométrica puede presentar serios inconvenientes de índole operativo. Todos esas dificultades se soslayan demostrando - primero de nuestros objetivos- que la familia H_m está ordenada en variabilidad. Al afirmar que $h_m \underset{v}{\leq} h_{m+1}$, $\forall m:1,2,\ldots$, estamos asegurando que si X e Y son variables aleatorias que se distribuyen según las hipergeométricas h_m y h_{m+1} ,respectivamente, entonces

$$E[g(X)] \leq E[g(Y)],$$

para toda función g real y convexa, lo que es tanto como decir que la dispersión de las distribuciones de la familia H_m crece monótonamente, ya que cualquiera de las medidas usuales de

dispersión pueden ser expresadas como valores esperados de funciones reales convexas.

Como paso previo para alcanzar este objetivo, demostraremos que

(i) H_m converge de forma monótona creciente a la distribución B(n,p) en log-concavidad relativa, es decir,

$$h_m \underset{lc}{\leq} h_{m+1} \ , \quad \forall m=1,2,\ldots \ .$$

(ii) Los elementos de la familia H_m no están ordenadas estocásticamente.

Como consecuencia de (i) y (ii) podemos afirmar que la familia H_m está ordenada en variabilidad condicional uniforme y, por otra parte, al tener todas las distribuciones igual media, se concluye, finalmente, que están ordenadas en variabilidad.

La ordenación en log-concavidad relativa de la sucesión H_m , nos va a permitir demostrar, seguidamente, la monotonía de la entropía Δ de las distribuciones de la citada familia. El interés de este resultado crece al tener en cuenta que la validez del mismo no se circunscribe al caso de distribuciones hipergeométricas, sino que se extiende a toda la clase $\mathbb{U}^*$ de distribuciones fuertemente unimodales.

El teorema asegura que si X e Y son variables aleatorias discretas, con funciones de masa de probabilidad f y g, respectivamente, y $g \in \mathbb{U}^*$,entonces

$$f \underset{v}{\leq} g \quad \Rightarrow \quad \Delta f \leq \Delta g \quad ,$$

verificandose la igualdad sólo si $X \overset{d}{=} Y$.

La siguiente sección está dedicada al estudio de la aleatoriedad de las distribuciones de la sucesión H_m , al margen de medidas concretas -la entropía es una de ellas-, de la aleatoriedad. Para este estudio seguiremos el método proporcionada por Hickey (1982) para comparar distribuciones discretas en aleatoriedad.

En primer lugar, adecuaremos convenientemente los resultados de Uhlmann (1966), de forma que nos permitan comparar las funciones de distribución de dos hipergeométricas cualesquiera pertenecientes a la sucesión H_m ,salvo para valores pertenecientes a un intervalo I, que tiene la forma

$$I = \left(p(n-1) - \frac{q(n-1)}{N} \ , \ p(n-1) + \frac{p(n-1)}{N} \right)$$

y que denominaremos intervalo de indeterminación.

Utilizando estos resultados y siguiendo el criterio de Hickey, ofrecemos un primer resultado parcial:

La aleatoriedad de las distribuciones de la sucesión

de hipergeométricas $H_m(mN,p,n)$, con soporte común $\langle 0,1,\ldots n\rangle$, crece monótonamente siempre que la media de las distribuciones sea un número entero.

Nuestro siguiente objetivo es generalizar este teorema, eliminando la restricción impuesta sobre la media.

El criterio de Hickey (1982) para la comparación en aleatoriedad establece que

$$H_r \underset{a}{<} H_s \quad \Leftrightarrow \quad \sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]} \quad \forall\ j:1,2,\ldots n-1 \ .$$

donde con $h_{r[i]}$ y $h_{s[i]}$ se nota la i-ésima probabilidad puntual de las distribuciones H_r y H_s ,respectivamente, ordenadas dichas probabilidades en orden decreciente.

Observemos que si j=1, se obtiene que una condición necesaria para que $H_r \underset{a}{<} H_s$ es que la probabilidad modal de la distribución H_r sea mayor que la correspondiente probabilidad en H_s . Esto pone de manifiesto la importancia del papel desempeñado por la moda, de manera que nuestros primeros pasos se conducen en esa dirección. En este sentido,

(i) Estudiamos la localización de la moda en la distribución hipergeométrica, así como su posición relativa respecto del intervalo de indeterminación y su comportamiento en la sucesión H_m .

La posición relativa del valor modal de la distribución H_m respecto del intervalo de indeterminación, da lugar a una variada casuística, siendo distintas, en cada una de los casos, las condiciones previas requeridas para poder completar la generalización del teorema sobre la aleatoriedad.

(ii) Demostramos que siempre que las distribuciones $H_m(mN,p,n)$, con soporte común $\langle 0,1,..n\rangle$, no localicen la moda en los valores extremos del soporte, se cumple que la probabilidad modal es decreciente en la citada sucesión.

El caso en que la moda se localice en cero o en n, lo estudiamos separadamente. En esta tesitura, el comportamiento de las distribuciones de la sucesión H_m es singular, en cuanto que la correspondiente sucesión de probabilidades modales es creciente. No obstante, esto no se continua con un comportamiento opuesto de la aleatoriedad sino que, como pondremos de manifiesto, no es posible comparar de manera absoluta la aleatoriedad: sólo se podrán dar resultados sobre medidas concretas.

Siguiendo con el criterio de Hickey, para j=2, habra de compararse la suma de las dos mayores probabilidades puntuales, lo cual nos obliga a determinar la localización de la segunda probabilidad puntual, en orden decreciente.

En este sentido, obtenemos dos teoremas en los que

ofrecemos condiciones suficientes para poder afirmar, en cada caso, cuál de los dos enteros adyacentes a la moda tiene asignada mayor probabilidad.

Seguidamente, y en la misma línea, demostramos que si la media es mayor o igual que la moda M, entonces h(M+1) es mayor que h(M-2).

Con este resultado conseguiremos localizar las tres mayores probabilidades puntuales de una distribución hipergeométrica.

Todo este estudio previo será utilizado para generalizar nuestro primer resultado parcial sobre la aleatoriedad de las distribuciones de la familia $H_m(mN,p,n)$, con soporte común $\langle 0,1,\ldots,n\rangle$, y obtener que, siempre que la moda no se localice en los valores extremos del soporte, en cualquiera de las dos siguientes situaciones,

$$\text{(i)} \quad p(n-1) + \frac{p(n-1)}{N} \leq M$$

$$\text{(ii)} \quad p(n-1) + \frac{p(n-1)}{N} > M \quad \text{y} \quad p(n+1) > M+q,$$

donde M es la moda de H(N,p,n) y q=1-p, se verifica que

$$H_r \underset{a}{<} H_s \,, \quad \forall\, r<s \quad .$$

IV.2 MONOTONIA DE LAS MEDIDAS DE DISPERSION

Consideremos la sucesión de hipergeométricas $H_m(mN,p,n)$ que, como es bien sabido, converge hacia la binomial B(n,p).

Vamos a estudiar, en primer lugar, el comportamiento monótono del coeficiente de asimetría y el de algunos momentos en esa sucesión, para, a continuación, obtener un resultado más general que permita asegurar la monotonía creciente de las medidas usuales de dispersión.

Por conveniencia operativa, en el estudio que sigue, consideraremos la familia hipergeometrica H(N,p,n), donde los parámetros p y n permanecerán constantes y haremos crecer de forma continua el valor del parámetro N, obviando, de esta manera, el hecho de que el valor de dicho parámetro, así como el del parámetro X (X=pN), puedan no ser enteros y por tanto no compatibles con el modelo hipergeométrico, si bien, a veces, y por el mismo motivo, nos centraremos en la sucesión $H_m(mN,p,n)$, restrigiendo, de este modo, el campo de variabilidad del primer parámetro.

En lo que sigue consideraremos, asimismo, que $1<n<N-1$, excluyendo los casos triviales en los que al ser n=1 o n=N-1, la función de masa de probabilidad de todos los términos de la sucesión de hipergeométricas coincide con la de la distribución binomial límite.

ESTUDIO DEL COEFICIENTE DE ASIMETRIA

Como vimos en el primer capítulo de la memoria, el coeficiente de asimetría $\sqrt{\beta_1}$ para la distribución hipergeométrica H(N,p,n) viene dado por

$$\sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} \frac{\sqrt{N-1}}{\sqrt{N-n}} \frac{N-2n}{N-2} \quad ,$$

donde, como es habitual, q=1-p.

Desglosaremos el estudio en tres casos:

(i) p=1/2

En este caso, $\sqrt{\beta_1}=0$, por lo que las distribuciones de la familia $H_m(mN,p,n)$ son todas simétricas.

(ii) p<1/2

Si p<q, entonces $\sqrt{\beta_1}>0$, salvo quizás para el primer término de la sucesión $H_m(mN,p,n)$, para el cual puede ocurrir que N-2n sea negativo o cero. A partir del segundo término ($m\geq 2$), necesariamente, mN-2n es una cantidad estrictamente positiva por lo que la asimetría es positiva.

Veamos que las distribuciones de la sucesión poseen asimetría creciente.

Bastará para ello con estudiar el comportamiento de

$$u = \frac{\sqrt{N-1}}{\sqrt{N-n}} \frac{N-2n}{N-2}$$

donde $u>0$. Tomando logaritmos y derivando respecto de N, se obtiene

$$\frac{\partial Lu}{\partial N} = \frac{(n-1)[4(N-1)(N-n)-(N-2n)(N-2)]}{2(N-2n)(N-2)(N-1)(N-n)} \quad .$$

Si hemos excluido el caso en el que N-2n pueda ser menor o igual que cero, el cociente anterior es siempre positivo, por lo que, en conclusión, la asimetría es creciente.

(iii) $p>1/2$

Este caso es simétrico del anterior, de manera que la asimetría de las distribuciones de la sucesión de hipergeométricas es negativa, salvo quizás para el primer término de la sucesión, para el que $\sqrt{\beta_1}$ puede ser positivo o cero. En cualquier caso, el coeficiente de asimetría decrecerá monótonamente al aumentar el tamaño de la población.

Finalmente, es claro que al tomar

$$\lim_{N\to\infty} \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}}$$

se obtiene el valor del coeficiente de asimetría de la distribución binomial límite B(n,p).

MONOTONIA DE LOS MOMENTOS FACTORIALES

El momento factorial de orden r de la distribución hipergeométrica H(N,p,n) viene dado por

$$\mu(r) = \frac{p(pN-1)\ldots(pN-r+1)n(n-1)\ldots(n-r+1)}{(N-1)\ldots(N-r+1)} .$$

Consideremos esta expresión como función de N, para $1<r<\min(n,pN)$.

Los factores $\frac{pN-j}{N-j}$, con $1\leq j\leq r-1$, son todos ellos positivos y crecientes en N, como puede comprobarse fácilmente, por lo que el momento factorial $\mu(r)$ es creciente en la sucesión $H_m(mN,p,n)$.

Para r=1, el primer momento factorial es la media, que permanece constante en la mencionada sucesión.

MONOTONIA DE LOS MOMENTOS ORDINARIOS

Vimos en el primer capítulo de la memoria que el momento ordinario de orden r podía obtenerse a partir de los momentos factoriales:

$$\mu_r = \sum_{j=1}^{r} S_r^j \, \mu(j),$$

donde los coeficientes S_r^j reciben el nombre de números de Stirling de 2º tipo y son todos ellos positivos.

Como para $j>1$ acabamos de ver que $\mu(j)$ es creciente en N (y para $j=1$ es constante), tendremos que los momentos ordinarios de orden mayor que uno, como sumatoria de términos crecientes en N, son crecientes en N.

MONOTONIA DE LOS MOMENTOS CENTRALES

Puede estudiarse fácilmente el comportamiento monótono creciente del momento central de segundo orden (varianza),

$$m_2 = npq \frac{N-n}{N-1} \quad ,$$

sin más que ver que $\frac{N-n}{N-1}$ es creciente en N.

Podríamos, igualmente, demostrar el crecimiento del momento central de orden cuatro, expresándolo en función de los factoriales, si bien la demostración resulta ya más complicada. La demostración, siguiendo el mismo camino, para momentos centrales de orden par superior a cuatro puede devenir tarea imposible.

Hemos querido llegar a este punto para realzar la importancia, en el estudio que realizaremos a continuación, de la log-concavidad relativa como instrumento que permitirá demostrar de una manera cómoda, obviando la dificultad operativa que entraña el uso de la expresión de los momentos de la distribución hipergeométrica, el comportamiento monótono creciente de cualquier medida razonable de dispersión, entre las que se encuentran los momentos ordinarios y centrales de orden par y los momentos centrales absolutos, en la sucesión $H_m(mN,p,n)$.

MONOTONIA EN DISPERSION

ESTUDIO DEL ORDEN ESTOCASTICO

4.1. LEMA:

Si h_{m_1} y h_{m_2} son las funciones de masa de probabilidad de dos distribuciones cualesquiera pertenecientes a la familia $H_m(mN,p,n)$, con $m_1<m_2$, entonces soporte(h_{m_1}) $\subseteq$ soporte(h_{m_2}), verificandose la igualdad sólo si los valores cero y n pertenecen a los soportes de ambas distribuciones.

Demostración:

Consideremos la distribución hipergeométrica H(mN,p,n). El soporte de esta distribución es un conjunto de enteros consecutivos $\{a,a+1,\ldots,b\}$, donde

$$a = \max(0,n-mqN) \quad ; \quad b = \min(mpN,n) \ .$$

Nos interesa estudiar el comportamiento del soporte cuando a y b no coinciden, respectivamente, con 0 y n, es decir, cuando el soporte depende de m.

Si a=n-mqN, el valor de a disminuye al aumentar m; y si b=mpN, el valor de b aumenta al aumentar m, por lo que podemos concluir que

$$m_1 < m_2 \ \Rightarrow \ \text{soporte}(h_{m_1}) \subseteq \text{soporte}(h_{m_2}) .$$

4.2 LEMA:

Dadas dos distribuciones H_{m_1} y H_{m_2}, con $m_1<m_2$, pertenecientes a la familia $H_m(mN,p,n)$, se cumple:

(i) $H_{m_1}(0) \leq H_{m_2}(0)$

(ii) $H_{m_1}(n-1) \geq H_{m_2}(n-1)$,

verificandose la igualdad en (i) y (ii) sólo si los valores de 0 y n, respectivamente, no pertenecen a los soportes de ambas distribuciones.

Demostración:

La probabilidad puntual en el cero para una distribución cualquiera de la familia es de la forma

$$h_m(0) = \frac{\left[\begin{array}{c} pmN \\ 0 \end{array}\right]\left[\begin{array}{c} qmN \\ n \end{array}\right]}{\left[\begin{array}{c} mN \\ n \end{array}\right]} = q\,\frac{(qmN-1)\ldots(qmN-n+1)}{(mN-1)\ldots(mN-n+1)} \quad .$$

Cada factor (qmN-j)/(mN-j), con j:1,2,...,n-1, es positivo y creciente en m, como puede comprobarse fácilmente. Por tanto, salvo que el cero no pertenezca al soporte de h_{m_1} y h_{m_2}, en cuyo caso $H_{m_1}(0)=H_{m_2}(0)=0$, se verificará

$$H_{m_1}(0)<H_{m_2}(0).$$

Análogamente, puede comprobarse el crecimiento de las probabilidades puntuales en el valor n; salvo que n no pertenezca al soporte de las distribuciones, en cuyo caso $H_{m_1}(n-1) = H_{m_2}(n-1) = 1$, en el resto de los casos se verifica, para la probabilidad acumulada en el punto n-1, que

$$H_{m_1}(n-1)>H_{m_2}(n-1).$$

4.3 TEOREMA:

Las distribuciones de la familia $H_m(mN,p,n)$ no están ordenadas estocásticamente.

Demostración:

Consideremos dos distribuciones cualesquiera H_{m_1} y H_{m_2}, con $m_1 < m_2$, pertenecientes a la familia.

Si los valores 0 y n pertenecen al soporte de ambas distribuciones, el teorema es consecuencia inmediata del lema 4.2, ya que

$$H_{m_1}(0) < H_{m_2}(0) \quad ; \qquad H_{m_1}(n-1) > H_{m_2}(n-1) \quad ;$$

y, por tanto, no existe ordenación estocástica.

Si no es asi, debemos considerar tres casos:

(i) soporte(h_{m_1}) = $\{a,a+1,\ldots,n\}$, con $a>0$.

En este caso, por el lema 4.1,

soporte(h_{m_2}) = $\{a',\ldots,n\}$, con $a'<a$,

por lo que $H_{m_1}(a')=0 < H_{m_2}(a')$; y por el lema 4.2,

$$H_{m_1}(n-1) > H_{m_2}(n-1).$$

(ii) soporte(h_{m_1}) = $\{0,1,\ldots,b)$, con $b<n$.

En este caso, por el lema 4.1,

$$\text{soporte}(h_{m_2}) = \{0,1,\ldots b'\}, \text{ con } b<b',$$

por lo que $H_{m_1}(b)=1 > H_{m_2}(b)$; y por el lema 4.2,

$$H_{m_1}(0) < H_{m_2}(0).$$

(iii) $\text{soporte}(h_{m_1}) = \{a,\ldots,b\}$, con $0<a$ y $b<n$.

Razonando como en los dos casos anteriores tendremos que

$$H_{m_1}(a')=0 < H_{m_2}(a') \; ; \qquad H_{m_1}(b)=1 > H_{m_2}(b) \quad .$$

Como puede verse, en ningún caso existirá ordenación estocástica.

MONOTONIA EN LOG-CONCAVIDAD RELATIVA

4.4 TEOREMA:

La familia de hipergeométricas $H_m(mN,p,n)$ converge de forma monótona creciente a la distribución binomial B(n,p) en log-concavidad relativa.

Demostración:

Vimos en la sección III.2 de esta memoria que una función de masa de probabilidad f es log-cóncava

relativa a otra g, y se nota $f \underset{lc}{\leq} g$, si

(i) soporte(f) $\subseteq$ soporte (g), y
(ii) $r_f(k) \geq r_g(k)$, $\forall\ k \in$ soporte(f),

donde $r_f(k)$ es el índice de log-concavidad, definido por

$$r_f(k) = \frac{f^2(k)}{f(k-1)f(k+1)} .$$

En el caso de una distribución hipergeométrica perteneciente a la familia $H_m(mN,p,n)$ tenemos que

$$r_{h_m}(k) = \frac{(k+1)(mpN-k+1)(mqN-n+k+1)(n-k+1)}{k(mpN-k)(mqk-n+k)(n-k)} .$$

Puesto que $(k+1)(n-k+1)/[k(n-k)]$ es mayor que cero, para estudiar la monotonía respecto de m de $r_{h_m}(k)$ basta con estudiar el comportamiento de

$$u = \frac{(mpN-k+1)(mqN-n+k+1)}{(mpN-k)(mqN-n+k)} .$$

Tomando logaritmos y derivando respecto de m,

$$\frac{\partial L[u]}{\partial m} = pN\left(\frac{1}{mpN-k+1} - \frac{1}{mpN-k}\right) + qN\left(\frac{1}{mqN-n+k+1} - \frac{1}{mqN-n+k}\right).$$

Como es fácil ver,

$$\frac{\partial L[u]}{\partial m} < 0 ,$$

por lo que $r_{h_m}(k)$ es estrictamente decreciente en m.

Por otra parte, el lema 4.1 nos asegura que

$$\text{soporte}(h_m) \subseteq \text{soporte}\ (h_{m+1}),\ \forall m:1,2,\ldots$$

Hemos demostrado, en conclusión, que

$$h_m \underset{lc}{\leq} h_{m+1}\ ,\quad \forall m:1,2,\ldots$$

es decir, la sucesión de distribuciones $h_m(mN,p,n)$ converge de forma monótona creciente en log-concavidad relativa a la distribución binomial B(n,p).

MONOTONIA EN VARIABILIDAD CONDICIONAL UNIFORME

A partir del teorema 4.4 se obtiene :

4.5 COROLARIO:

$$h_m \underset{uv}{\leq} h_{m+1}\ ,\quad \forall m:1,2,\ldots$$

Para demostrar que los elementos de la familia $h_m(mN,p,n)$ están ordenados en variabilidad condicional uniforme, basta tener en cuenta que al estarlo en log-concavidad relativa se cumplen dos de las tres condiciones exigidas. La condición restante es que las distribuciones de la familia no estén ordenadas estocásticamente, lo cual fue probado con el teorema 4.3.

4.6 COROLARIO:

$S(c\ h_m - h_{m+1}) \leq 2$, $\forall m:1,2,\ldots$, para toda constante positiva c, verificandose la igualdad cuando c=1, y en este caso, la secuencia de signos es -, +, -.

Este corolario es consecuencia inmediata del anterior y del teorema 3.12.

MONOTONIA EN VARIABILIDAD

Al tener todas las distribuciones de la familia $H_m(mN,p,n)$ la misma media, entonces, como consecuencia directa del corolario anterior y de la condición [3.12], se cumple :

4.7 COROLARIO:

$$h_m \underset{v}{\leq} h_{m+1} , \quad \forall m:1,2,\ldots$$

Este último corolario nos asegura que si X e Y son variables aleatorias distribuidas, respectivamente, según H_m y H_{m+1}, entonces, para toda función real convexa g,

$$E[g(X)] \leq E[g(Y)] ,$$

FACULTAD DE CIENCIAS BIBLIOTECA GRANADA

Como sabemos, cualquiera de las medidas usuales de dispersión: momentos centrales y ordinarios de orden par, y momentos absolutos, se pueden expresar como valores esperados de funciones reales convexas.

Como habíamos anticipado, el uso de la relación de log-concavidad relativa ha permitido demostrar, de una forma cómoda, el comportamiento monónoto creciente en dispersión de los términos de la sucesión de hipergeométricas $H_m(mN,p,n)$, con carácter general, es decir, sin descender al caso particular de medidas concretas de dispersión, cuya manipulación, en el caso de distribuciones hipergeométricas, presentaba la dificultad que ya pusimos de manifiesto.

En resumen:

4.8. COROLARIO:

La familia de distribuciones hipergeométricas $H_m(mN,p,n)$ es monótona creciente en dispersión.

Queremos hacer notar que en el caso de una variable aleatoria hipergeométrica, al tratarse de una variable discreta no negativa, el corolario 4.7 nos permite, asimismo, asegurar el comportamiento monónoto creciente en la sucesión $H_m(mN,p,n)$ de cualquier momento ordinario de orden mayor que

uno, -no sólo de los pares-, y de los momentos factoriales, también de orden mayor que uno, por tratarse, en este caso, de valores esperados de funciones convexas, como se puede comprobar fácilmente.

MONOTONIA DE LA ENTROPIA

Como una aplicación más del corolario 4.7, vamos a demostrar, a continuación, la monotonía de la entropía de las distribuciones pertenecientes a la familia $H_m(mN,p,n)$.

Aunque la entropia, como vimos en la seccion III.3, es una medida de aleatoriedad, y el estudio de la monotonía en aleatoriedad será el que abordemos en la siguiente sección de este capítulo, nos parece conveniente adelantar aquí el resultado sobre la entropía; en primer lugar, por tratarse de una aplicación de la convergencia monótona de la citada familia en log-concavidad relativa, y después, porque la validez del resultado así obtenido, no va a estar limitada por ningún tipo de restricción previa, lo que no ocurrirá, desafortunadamente, en los resultados que obtendremos sobre monotonía en aleatoriedad. Además, el resultado trascenderá la distribución hipergeométrica y podrá ser planteado como un teorema más general.

4.9 TEOREMA:

La entropía de las distribuciones de la familia de hipergeométricas $H_m(mN,p,n)$ es estrictamente monótona creciente.

Demostración:

Para una distribución hipergeométrica con función de masa de probabilidad h(x), la entropía viene dada por la expresión

$$\Delta = - \sum_i h(xi)\ L[h(x_i)],$$

lo que permite ser expresada en terminos de valor esperado :

$$\Delta = E[-L[h(x_i)]].$$

Sabemos que por pertenecer la distribución hipergeométrica a la clase de distribuciones fuertemente unimodales, su función de masa de probabilidad es log-cóncava, y, por tanto, $-L[h(x)]$ es una función convexa.

Sin embargo, no vamos a poder aplicar directamente el corolario 4.7 del teorema de convergencia monótona en log-concavidad relativa de la sucesión $H_m(mN,p,n)$ para demostrar la monotonía de la entropía, puesto que aunque $-L[h_m(x)]$ sea una función convexa, es una función

distinta para cada elemento de la familia.

La demostración del teorema va a requerir utilizar el Lema de Gibbs, cuyo enunciado es el siguiente:

Dados dos sistemas de números $r_1, r_2, \ldots, r_n$, y $s_1, s_2, \ldots s_m$, positivos, tales que $m \geq n$ y

$$\sum_{k=1}^{n} r_k \geq \sum_{k=1}^{n} s_k \quad ,$$

siempre se verifica que

$$-\sum_{k=1}^{n} r_k \, L(r_k) \leq -\sum_{k=1}^{n} r_k \, L(s_k),$$

cumpliendose la igualdad únicamente si $r_k = s_k$, $\forall k$.

La demostración del lema, para el caso n=m, puede verse en el libro de Gil (1981), pg. 22.

Consideremos ahora la siguiente función:

$$g(x) = -L[\, h(x/m_2)\,],$$

donde $h(x/m_2)$ representa las probabilidades puntuales de la distribución hipergeométrica $H(m_2 N, p, n)$.

Como ya hemos indicado anteriormente, $h(x/m_2)$ es log-cóncava, por lo que g(x) es una función convexa.

Aplicando ahora el corolario 4.7, tendremos que si $m_1 < m_2$, entonces $E[\, g(x)/m_1] \leq E[\, g(x)/m_2]$, es decir,

$$-\sum_{k=o}^{n} h(k/m_1) \; L[\, h(k/m_2)] \leq -\sum_{k=o}^{n} h(k/m_2) \; L[\, h(k/m_2)] \; ,$$

donde la expresión del segundo miembro es la entropía de la distribución H_{m_2}, y la notaremos por Δ_2.

Por otra parte, aplicando el lema de Gibbs, tendremos que

$$-\sum_{k=o}^{n} h(k/m_1)\ L[h(k/m_1)] < -\sum_{k=o}^{n} h(k/m_1)\ L[h(k/m_2)]\ ,$$

y, puesto que el primer miembro de la desigualdad es la entropía Δ_1 de la distribución H_{m_1}, tendremos finalmente que $\Delta_1<\Delta_2$, quedando, por tanto, demostrado el crecimiento estrictamente monótono de la entropía en la familia de hipergeométricas $H_m(mN,p,n)$.

Aunque en la demostración hemos estado considerando, implícitamente, que el soporte de h_{m_1} y h_{m_2} es, en ambos casos, $\langle\ 0,1,..,n\ \rangle$, la validez general del resultado está garantizada por ser soporte(h_{m_1})$\subseteq$ soporte(h_{m_2}).

4.10 COROLARIO:

La distribución binomial B(n,p) tiene mayor entropía que la distribución hipergeométrica H(N,p,n).

En el desarrollo de la demostración del teorema

anterior, podemos observar que su validez es posible extenderla, en tanto en cuanto no hemos utilizado en ningún momento propiedades particulares o exclusivas de la distibución hipergeométrica, sino que hemos utilizado unas condiciones más generales como son la log-concavidad y la log-concavidad relativa.

El teorema anterior es, por tanto, susceptible de ser generalizado:

Sea $\mathbb{U}^*$ la clase de funciones de probabilidad fuertemente unimodales, o, equivalentemente, log-cóncavas, y por Δ_f notaremos la entropía de la variable aleatoria con función de probabilidad f.

4.11 TEOREMA:

Sean X e Y variables aleatorias discretas con funciones de masa de probabilidad f y g, respectivamente, y sea $g \in \mathbb{U}^*$. Entonces,

$$f \underset{v}{\leq} g \quad \Rightarrow \quad \Delta_f \leq \Delta_g \quad ,$$

verificandose la igualdad sólo si $X \overset{d}{=} Y$.

IV.3 MONOTONIA EN ALEATORIEDAD

Esta Sección va a estar dedicada al estudio del crecimiento monótono de la aleatoriedad de las distribuciones de la familia de hipergeométricas $H_m(mN,p,n)$.

Para este estudio se seguirá el método proporcionado por Hickey (1982), para comparar distribuciones discretas en aleatoriedad, al margen del uso de medidas concretas, de manera que cualquiera de las medidas usuales de la aleatoriedad, como la entropía de Shannon, la energía informacional de Onicescu, o la información de Renyi de orden α, son consistentes con la ordenación establecida. El trabajo de Hickey fue desarrollado y comentado ampliamente en el tercer capítulo de esta memoria.

Recordemos que el soporte idóneo para comparar en aleatoriedad distribuciones discretas, lo suministraba la teoría de mayoración de vectores. Apoyándose en ella, Hickey (1982) da la siguiente definición:

> Dadas dos distribuciones discretas P y Q, con probabilidades puntuales $\{p_i\}$ y $\{q_i\}$, con i=1,2,..., diremos que P es al menos tan aleatoria como Q, y se notará $Q \underset{a}{\leq} P$, si los vectores $\vec{p}$ y $\vec{q}$, cuyas componentes respectivas son $\{p_i\}$ y $\{q_i\}$ y las componentes nulas necesarias para que ambos vectores tengan la misma dimensión, verifica que $\vec{p} \underset{m}{\leq} \vec{q}$.

En el caso que $\vec{p} \underset{m}{<} \vec{q}$, se dirá que P es más aleatoria que Q.

Por otra parte, $\vec{p} \underset{m}{\leq} \vec{q}$ ($\vec{p}$ está mayorado por $\vec{q}$) si

$$[4.1] \qquad (i) \quad \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i \qquad n = \text{dimensión de } \vec{p} \text{ y } \vec{q}$$

$$[4.2] \qquad (ii) \quad \sum_{i=1}^{j} p_{[i]} \leq \sum_{i=1}^{j} q_{[i]}, \quad \forall j:1,2,\ldots,n-1 \quad ,$$

donde la notación $p_{[i]}$ y $q_{[i]}$ representa la i-ésima probabilidad puntual, cuando las probabilidades puntuales han sido ordenadas decrecientemente; de manera que $\sum_{i=1}^{j} p_{[i]}$ indica la suma de las j mayores probabilidades puntuales para la distribución P.

Si alguna de las desigualdades de [4.2] fuese estricta, se dirá que la mayoración es estricta y se notará $\vec{p} \underset{m}{<} \vec{q}$.

Cuando estemos considerando vectores de probabilidad, la condición [4.1] se verifica obviamente. Debemos centrar, por tanto, nuestra atención en la condición [4.2].

Hickey define una medida de aleatoriedad como una función continua ϕ, simétrica y cóncava, que toma valores reales sobre la clase de las distribuciones discretas, siendo

la concavidad, estricta sobre la subclase de distribuciones que tienen sólo un número finito de probabilidades positivas.

Lo que este autor exige con las condiciones de simetría y concavidad, de acuerdo con la proposición 3.15, es la schur-concavidad de la función ϕ (schur-concavidad estricta en el caso que las distribuciones tengan sólo un número finito de probabilidades positivas), o lo que es lo mismo, que sus opuestas sean consistentes con el orden de mayoración establecido entre los vectores de probabilidad de las distribuciones; es decir,

$$\text{si} \quad \vec{p} \underset{m}{\leq} \vec{q} \quad \Rightarrow \quad \phi(\vec{q}) \leq \phi(\vec{p}).$$

De acuerdo con esta definicion, la entropia de Shannon, las medidas de información de Renyi y la opuesta de la energía informacional de Onicescu, son medidas de aleatoriedad.

El primer resultado que vamos a obtener va a consistir en demostrar el crecimiento monótono de la aleatoriedad de las distribuciones de la sucesión de hipergeométricas $H_m(mN,p,n)$, en el caso particular de que la media μ, común a todas las distribuciones de la sucesión, sea un número entero.

En la demostración de este teorema, así como en el estudio del caso más general -cuando no se imponga restricción sobre el valor de la media-, van a jugar un papel importante

los resultados de Uhlmann (1966), fruto de una comparación sistemática entre la hipergeométrica y la binomial de igual media, que va a permitir comparar las funciones de distribución de las citadas distribuciones, así como las de dos distribuciones hipergeométricas cualesquiera que pertenezcan a la familia $H_m(mN,p,n)$.

Un amplio comentario del estudio de Uhlmann (1966) está recogido en la sección VIII del capítulo I

COMPARACION DE LAS FUNCIONES DE DISTRIBUCION

Uhlmann obtiene que para $0<c<n-1<N$,

$$L_{N,n,c}(p) - L_{N+1,n,c}(p) \quad \text{es}$$

$$[4.3] \quad \begin{cases} 0 & \text{si} \quad 0 \le p \le \dfrac{c}{N+1} \\ >0 & \text{si} \quad \dfrac{c}{N+1} < p \le \dfrac{c}{n-1} - \dfrac{c}{(n-1)(N+1)} \\ <0 & \text{si} \quad \dfrac{c}{n-1} + \dfrac{n-c-1}{(n-1)(N+1)} \le p < 1 - \dfrac{n-c-1}{N+1} \\ 0 & \text{si} \quad 1 - \dfrac{n-c-1}{N+1} \le p \le 1 \end{cases}$$

donde

$$L_{N,n,c}(p) = \begin{cases} 1 & \text{si} \quad 0 \le p \le \frac{c}{N} \\ \sum_{k=0}^{c} \frac{\binom{pN}{k}\binom{qN}{n-k}}{\binom{N}{n}} & \text{si} \quad \frac{c}{N} \le p \le 1 - \frac{n-c-1}{N} \\ 0 & \text{si} \quad 1 - \frac{n-c-1}{N} \le p \le 1 \end{cases}$$

con q=1-p, como es habitual.

Como podemos ver, para un valor p fijo, que sea compatible con el parámetro N, $L_{N,n,c}(p)$ coincide con el valor de la función de distribución de la hipergeométrica H(N,p,n), en el punto c.

Podemos, por tanto, reinterpretar el resultado [4.3], expresándolo en función de c, y obtendremos que para $p \in (0,1)$ y $0 < c < n-1 < N$,

$L_{N,p,n}(c) - L_{N+1,p,n}(c)$ es

[4.4]
$$\begin{cases} 0 & \text{si} \quad c \le n-q(N+1)-1 \\ < 0 & \text{si} \quad n-q(N+1)-1 < c \le \frac{p(n-1)(N+1)-n+1}{N} \\ > 0 & \text{si} \quad p(n-1)\frac{N+1}{N} \le c < p(N+1) \\ 0 & \text{si} \quad c \ge p\,(N+1) \end{cases}$$

Consideremos ahora la sucesión de distribuciones hipergeométricas $H_m(mN,p,n)$ y admitamos que todas las

distribuciones tienen soporte común de la forma

$$\text{soporte}(h_m) = \{0,1,\ldots,n\}, \quad \forall m.$$

Para la obtención de nuestro siguiente teorema esta restricción es innecesaria; la imponemos por una mayor simplicidad y atendiendo a que en nuestros resultados posteriores sí nos vamos a ver obligados a exigirla. En cualquier caso, podemos observar que esta restricción sólo afecta a los primeros términos de la sucesión, ya que si tenemos en cuenta la forma del soporte de una distribución hipergeométrica, para un m suficientemente grande los valores 0 y n pertenecen al soporte.

En estas condiciones sabemos, por el lema 4.2, que, para dos distribuciones cualesquiera H_r y H_s con $r<s$, pertenecientes a la citada familia, se cumple:

$$\text{(i)} \quad H_r(0) < H_s(0)$$

$$\text{(ii)} \quad H_r(n-1) > H_s(n-1) \qquad [4.5]$$

4.12 TEOREMA:

Sean h_r y H_s dos distribuciones cualesquiera pertenecientes a la sucesión de hipergeométricas $H_m(mN,p,n)$, con $p\in(0,1)$, $n>1$, y con soporte común $\{0,1,\ldots,n\}$.

Si $r<s$, se cumple que para $0 \leq c \leq n-1$,

$$H_r(c)-H_s(c) \quad \text{es} \quad \begin{cases} < 0 & \text{si} \quad c \leq \dfrac{p(n-1)(N_r+1)-n+1}{N_r} \\ > 0 & \text{si} \quad c \geq \dfrac{p(n-1)(N_r+1)}{N_r} \end{cases}$$

donde $H_r(c)$ y $H_s(c)$ representan el valor de la probabilidad acumulada en el punto c para las distribuciones H_r y H_s , respectivamente, y donde N_r representa el parámetro tamaño poblacional en la distribución H_r , es decir, $N_r=rN$.

Demostración:

Podemos escribir, utilizando la notación de Uhlmann que

$$[4.6] \quad L_{N_r,p,n}(c)-L_{N_s,p,n}(c) = \sum_{i=o}^{u} L_{N_r+i,p,n}(c)-L_{N_r+i+1,p,n}(c)$$

con $u = N_s-N_r+1$.

A partir de [4.4], para cada uno de los términos de la sumatoria anterior, se tendría que si $0<c<n-1$,

$$L_{N_r+i,p,n}(c)-L_{N_r+i+1,p,n}(c) \text{ es}$$

$$[4.7]\quad \begin{cases} 0 & \text{si} \quad c \leq n-q(N_r+i+1)-1 \\ < 0 & \text{si} \quad n-q(N_r+i+1)-1 < c \leq \dfrac{p(n-1)(N_r+i+1)-n+1}{N_r+i} \\ > 0 & \text{si} \quad \dfrac{p(n-1)(N_r+i+1)}{N_r+i} \leq c < p(N_r+i+1) \\ 0 & \text{si} \quad c \geq p(N_r+i+1) \end{cases}$$

Teniendo en cuenta que el soporte de una distribución hipergeométrica H(N,p,n) es de la forma

$$\{c\in\mathbb{Z}/\text{máx}(0,n-qN)\leq c\leq \text{mín}(n,pN)\} \ ,$$

si admitimos, por hipótesis, que el soporte de las distribuciones comparadas es $\{0,1,\ldots,n\}$, tendremos que

(i) $\quad n \leq pN_r$,

por lo que el valor $p(N_r+i+1)$ queda fuera del soporte;

(ii) $\quad 0 \geq n-qN_r$,

por lo que el valor $n-q(N_r+i+1)-1$ es menor que cero y, por tanto, queda fuera del soporte.

Como además, por [4.5], quedan comparadas las funciones de distribución de H_r y H_s en los puntos 0 y

n-1, podemos simplificar la expresión [4.7] y escribir:

Si $0 \leq c \leq n-1$,

$$L_{N_r+i,p,n}(c) - L_{N_r+i+1,p,n}(c) \quad \text{es}$$

$$[4.8] \qquad \begin{cases} < 0 & \text{si} \quad c \leq \dfrac{p(n-1)(N_r+i+1)-n+1}{N_r+i} \\ > 0 & \text{si} \quad c \geq \dfrac{p(n-1)(N_r+i+1)}{N_r+i} \end{cases}$$

Si tenemos en cuenta ahora que

$$\frac{p(n-1)(N_r+1)-n+1}{N_r} \leq \frac{p(n-1)(N_r+i+1)-n+1}{N_r+i} \quad , \qquad \forall i: 0,1,\ldots,$$

obtenemos que si $c \leq [p(n-1)(N_r+1)-n+1]/N_r$, cada término de la sumatoria [4.6] es negativo, por lo que, retomando la notación de función de distribución,

$$H_r(c) - H_s(c) < 0 \quad ;$$

por otra parte,

$$\frac{p(n-1)(N_r+1)}{N_r} \geq \frac{p(n-1)(N_r+i+1)}{N_r+i}, \qquad \forall i: 0,1,\ldots,$$

luego, razonando igualmente, si $c \geq [p(n-1)(N_r+1)]/N_r$, cada término de la sumatoria [4.6] es positivo, y, por tanto,

$$H_r(c) - H_s(c) > 0,$$

quedando demostrado el teorema.

4.13 COROLARIO:

Teniendo en cuenta que la sucesión $H_m(mN,p,n)$ converge a la distribución binomial B(n,p), si $0 \leq c \leq n-1$, $n>1$ y $0<p<1$,

$$H_r(c)-B(c) \text{ es } \begin{cases} < 0 & \text{si} \quad c \leq \dfrac{p(n-1)(N_r+1)-n+1}{N_r} \\ > 0 & \text{si} \quad c \geq \dfrac{p(n-1)(N_r+1)}{N_r} \end{cases}$$

donde $H_r(c)$ y B(c) son los valores de la probabilidad acumulada en el punto c de las distribuciones $H_r(rN,p,n)$ y B(n,p), respectivamente, y $N_r=rN$.

Otra forma de expresar los resultados del teorema 4.12 es la siguiente: bajo las hipótesis del mencionado teorema se cumple que si $r<s$,

$$\begin{aligned} H_r(c) &< H_s(c) \quad \text{si} \quad c \leq p(n-1) - \delta \\ H_r(c) &> H_s(c) \quad \text{si} \quad c \geq p(n-1) + \varepsilon \end{aligned} \qquad [4.9]$$

donde $\varepsilon = \dfrac{p(n-1)}{N_r}$ y $\delta = \dfrac{q(n-1)}{N_r}$.

Observemos que $\varepsilon+\delta = \dfrac{n-1}{N_r} < 1$.

Esta forma del teorema pone de manifiesto que como

máximo existe un sólo valor de c para el que la comparación de las funciones de distribución queda sin determinar; esto ocurre cuando $c \in (p(n-1)-\delta, p(n-1)+\varepsilon)$.

ALEATORIEDAD EN LA FAMILIA HIPERGEOMETRICA CON VALOR ESPERADO CONSTANTE Y ENTERO

El teorema que vamos a demostrar a continuación, permite asegurar la monotonía de la aleatoriedad, que crece estrictamente, de las distribuciones de la familia de hipergeométricas $H_m(mN,p,n)$ en un caso particular: cuando la media es un numero entero.

Veamos, en primer lugar, el siguiente lema:

4.14 LEMA:

Si μ es la media de las distribuciones $H_m(mN,p,n)$, es decir, $\mu=pn$, se tiene que

$$\mu-1 < p(n-1) - \frac{q(n-1)}{N_m} < p(n-1) + \frac{p(n-1)}{N_m} < \mu \quad ,$$

siendo $N_m=mN$.

En efecto:

$$\text{a)} \quad \left.\begin{array}{l} p \in (0,1) \\ p+q = 1 \\ \dfrac{n-1}{N_m} < 1 \end{array}\right\} \Rightarrow \; p + q\,\frac{n-1}{N_m} < 1;$$

restando a ambos miembros de la desigualdad la cantidad pn,

$$p(1-n) + q\,\frac{n-1}{N_m} < 1 - pn \quad ,$$

de donde, finalmente,

$$\mu - 1 < p(n-1) - \frac{q(n-1)}{N_m} \quad .$$

b) Sabemos que $p>p(n-1)/N_m$; sumando a ambos miembros de la desigualdad la cantidad p(n-1) se obtiene

$$\mu = p_n > p(n-1) + \frac{p(n-1)}{N_m} \quad .$$

4.15 TEOREMA:

Consideremos la sucesión de distribuciones hipergeométricas $H_m(mN,p,n)$, con $n>1$, $p\in(0,1)$, y con soporte común $\{0,1,\ldots,n\}$. Si la media de las distribuciones, $\mu = pn$, es un número entero, entonces, dadas dos distribuciones cualesquiera H_r y H_s, con $r<s$,

se verifica que H_r es menos aleatoria que H_s; es decir,

$$r<s \quad \Rightarrow \quad H_r \underset{a}{<} H_s \quad .$$

Demostración:

Compararemos la aleatoriedad de ambas distribuciones utilizando el criterio de Hickey (1982):

$$H_r \underset{a}{<} H_s \quad \Leftrightarrow \quad \sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]}, \quad \forall j:1,\ldots,n-1,$$

donde $h_{r[i]}$ indica la i-ésima probabilidad puntual de la distribución H_r, ordenadas dichas probabilidades en orden decreciente, de forma que $\sum_{i=1}^{j} h_{r[i]}$ es la suma de las j mayores probabilidades puntuales para la citada distribución.

Vimos en la Sección I.6 de esta memoria que si la media μ de una distribución hipergeométrica es un número entero, la moda M se localiza en dicho número entero.

Puesto que, por el lema 4.14, el intervalo $(p(n-1)-q(n-1)/N_r,\ p(n-1)+p(n-1)/N_r)$ está comprendido entre los enteros $\mu-1$ y μ, es decir, entre M-1 y M, el teorema 4.12 nos va a permitir comparar las funciones de distribución de H_r y H_s, en todo punto K perteneciente al soporte de ambas distribuciones.

Si r<s , tendremos que

[4.10] $$H_r(M-k) \leq H_s(M-k), \ \forall k:1,\ldots,M+1,$$

verificandose la igualdad sólo si k=M+1;

[4.11] $$H_r(M+t) \geq H_s(M+t), \ \forall t:0,\ldots,n-M,$$

verificandose la igualdad sólo si M+t=n.

Si tenemos en cuenta ahora el carácter unimodal de la distribución hipergeométrica, sabemos que las probabilidades puntuales crecen a la izquierda de la moda, que en este caso coincide con la media, y decrecen a la derecha. Vamos a poder, por tanto, expresar la suma de las j mayores probabilidades puntuales como la suma de j probabilidades puntuales consecutivas:

$$\sum_{i=1}^{j} h_{s[i]} = \sum_{i=u}^{v} h_s(i) = H_s(v) - H_s(u-1),$$

con u≤M≤v, y finalmente, como v=M+t y u-1=M-k, para algún t y k números naturales tales que 0≤t≤n-M y 1≤k≤M+1,

$$\sum_{i=1}^{j} h_{s[i]} = H_s(M+t) - H_s(M-k).$$

Análogamente, se obtendría que

$$\sum_{i=1}^{j} h_{r[i]} = H_r(M+t') - H_r(M-k'),$$

para t' y k' números naturales con las mismas condiciones que t y k.

Por [4.10] y [4.11] tenemos que

[4.12] $$\sum_{i=1}^{j} h_{s[i]} = H_s(M+t) - H_s(M-k) < H_r(M+t) - H_r(M-k).$$

Hacemos notar que la desigualdad es estricta porque las igualdades en [4.10] y [4.11] no pueden, en nuestro caso, verificarse simultáneamente, ya que para que así fuese, t+k=n+1, mientras que al estar la sumatoria de [4.12] extendida hasta un j que, como máximo, alcanza el valor n-1, tenemos que $t+k = j+1 \leq n$.

Como, por otra parte, $H_r(M+t)-H_r(M-k)$ es la suma de j probabilidades puntuales de la distribución H_r, la citada diferencia será menor o igual que la suma de las j mayores probabilidades puntuales de H_r, $\sum_{i=1}^{j} h_{r[i]}$, luego

$$\sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]} \quad , \forall j:1,\ldots,n-1,$$

quedando demostrado que

$$H_r \underset{a}{<} H_s \ .$$

Un estudio, realizado en términos similares, para la aleatoriedad de distribuciones binomiales con valor esperado entero, puede verse en Ollero (1985).

4.16 COROLARIO:

Puesto que la sucesión de hipergeométricas $H_m(mN,p,n)$ converge hacia la distribución B(n,p), bajo la hipótesis del teorema anterior, se verifica que

$$H_r \underset{a}{<} B \quad , \qquad \forall r:1,2,\ldots$$

ALEATORIEDAD EN LA FAMILIA HIPERGEOMETRICA CON VALOR ESPERADO CONSTANTE

OBSERVACIONES PREVIAS

Nuestro próximo objetivo es generalizar el teorema 4.15 y demostrar el crecimiento monótono de la aleatoriedad de las distribuciones de la familia $H_m(mN,p,n)$, sin exigir que la media sea un número entero. Si eliminamos esta restricción nos encontramos con que el intervalo

[4.13] $$\left(p(n-1) - \frac{q(n-1)}{N},\ p(n-1) + \frac{p(n-1)}{N} \right)$$

puede contener a un número entero, punto en el cual no podríamos comparar los valores de las funciones de probabilidad acumulada de las distribuciones de la citada familia de hipergeométricas, interrumpiendose, por este motivo, el razonamiento usado en la demostración del teorema 4.15. Incluso, aunque el anterior intervalo no contenga a ningún número entero, si la moda de las distribuciones comparadas es menor que el extremo inferior del intervalo, nos encontramos con la imposibilidad de comparar, por el procedimiento utilizado en el teorema 4.15, las probabilidades puntuales en el valor modal, lo que, de acuerdo con el criterio de Hickey, impediría comparar en aleatoriedad las distribuciones en cuestión.

En la demostración del teorema 4.15 generalizado seguiremos utilizando los resultados del teorema 4.12 para comparar las probabilidades acumuladas, pero además ahora vamos a necesitar una serie de resultados previos que serán objeto de estudio a continuación, y que podemos resumirlos en:

a) Estudio detallado del comportamiento de la moda de las distribuciones de la familia $H_m(mN,p,n)$, lo cual incluye un análisis de la posición relativa del valor modal respecto del intervalo [4.13] al que denominaremos intervalo de indeterminación.

La posición relativa del valor modal respecto del intervalo de indeterminación da lugar a una variada casuística

que nos obliga a un análisis de cada caso, ya que en cada uno de ellos son distintas las condiciones previas que se necesita demostrar para poder completar la generalización del teorema. Hay, no obstante, un resultado fundamental que utilizaremos en todos los casos: la demostración del decrecimiento de las probabilidades puntuales en la moda de las distribuciones de la familia $H_m(mN,p,n)$.

b) En alguno de los casos necesitaremos comparar las probabilidades puntuales en los enteros adyacentes a la moda de una distribución hipergeométrica, para, de este modo, poder determinar la posición, respecto del intervalo de indeterminación, de los dos valores con mayor probabilidad puntual.

c) En uno de los casos se necesitará comparar las probabilidades puntuales en los puntos M+1 y M-2, siendo M el valor modal, y además habría que demostrar el decrecimiento de la suma de las probabilidades puntuales en M y M-1, de las distribuciones de la familia que estamos estudiando.

A la vista de lo anterior, podemos comprender que el demostrar la monotonía en aleatoriedad de las distribuciones de la sucesión de hipergeométricas $H_m(mN,p,n)$ no va resultar tarea cómoda.

Queremos llamar la atención en este punto sobre un resultado de Hickey (1986), recogido en el capítulo III de la

memoria, que simplifica enormemente el estudio de la comparación en aleatoriedad, pero que, desafortunadamente para nosotros, no vamos a poder utilizar.

El teorema 3.17 de Hickey (1986) indica que si f y g son funciones de densidad, con g perteneciente a la familia de densidades unimodales, entonces

$$f \underset{disp}{\leq} g \quad \Rightarrow \quad f \underset{a}{\leq} g$$

es decir, la ordenación en dispersión, bajo ciertas condiciones, es más fuerte que la ordenación en aleatoriedad.

En primer lugar, la demostración de Hickey considera únicamente el caso de distribuciones continuas, si bien deja la puerta abierta para la validez del teorema en el caso discreto.

Nosotros queremos poner de manifiesto aquí que, en cualquier caso,no podríamos utilizar el resultado de Hickey, ya que las distribuciones hipergeométricas $H_m(mN,p,n)$ no se encuentran ordenadas en dispersión.

Recordemos, a este respecto, que una de las caracterizaciones de Shaked (1982), (Teorema 3.1), para la ordenación en dispersión es

$$F \underset{disp}{\leq} G \Leftrightarrow S(F_c - G) \leq 1, \quad \forall c \in \mathbb{R},$$

y, en el caso de igualdad, la secuencia de signos es -, +, siendo F_c una traslación de F.

Consideremos las distribuciones hipergeométricas H(20,0.25,3) y H(40,0.25,3) pertenecientes a la sucesión H_m(20m,0.25,3), y llamemos F y G a sus respectivas funciones de distribución. Tenemos que

$$F(t) = \begin{cases} 0 & \text{si} \quad t<0 \\ 0.39912 & \text{si} \quad 0\leq t<1 \\ 0.85965 & \text{si} \quad 1\leq t<2 \\ 0.99123 & \text{si} \quad 2\leq t<3 \\ 1 & \text{si} \quad t\geq 3 \end{cases}$$

$$G(t) = \begin{cases} 0 & \text{si} \quad t<0 \\ 0.41093 & \text{si} \quad 0\leq t<1 \\ 0.85121 & \text{si} \quad 1\leq t<2 \\ 0.98785 & \text{si} \quad 2\leq t<3 \\ 1 & \text{si} \quad t\geq 3 \end{cases}$$

Consideremos ahora c=0.5 y obtengamos $F_{0.5}$(t):

$$F_{0.5}(t) = \begin{cases} 0 & \text{si} \quad t<0.5 \\ 0.39912 & \text{si} \quad 0.5\leq t<1.5 \\ 0.85965 & \text{si} \quad 1.5\leq t<2.5 \\ 0.99123 & \text{si} \quad 2.5\leq t<3.5 \\ 1 & \text{si} \quad t\geq 3.5 \end{cases}$$

Podemos observar fácilmente que el número de cambios de signo, $S(F_{0.5} - G)=4$, lo cual nos sirve como contraejemplo

para demostrar que la familia de distribuciones hipergeométricas $H_m(mN,p,n)$ no se encuentra ordenada en dispersión.

Finalmente queremos indicar que, puesto que nuestro objetivo es comparar en aleatoriedad dos distribuciones hipergeométricas, atendiendo al criterio de Hickey, podemos considerar en lo que sigue, sin pérdida alguna de generalidad, que el parámetro p de las distribuciones es menor o igual que 1/2.

En efecto, si H_r y H_s son distribuciones pertenecientes a la familia $H_m(mN,p,n)$,

$$H_r \underset{\alpha}{<} H_s \Leftrightarrow \sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]} , \quad \forall j:1,\ldots,n-1$$

Si $p>1/2$, bastaría demostrar, por simetría, que $H_r^* \underset{\alpha}{<} H_s^*$, donde con H_r^* y H_s^* notamos, respectivamente, las distribuciones hipergeométricas $H_r^*(rN,1-p,n)$ y $H_r^*(sN,1-p,n)$ ya que, como vimos en la sección I.6 de esta memoria, se cumple, para las probabilidades puntuales de la hipergeométrica, que

$$h(c;N,p,n) = h(n-c;N,1-p,n),$$

por lo que, para todo i,

$$h_{r[i]} = h^*_{r[i]} ; \qquad h_{s[i]} = h^*_{s[i]} ;$$

y, por tanto,

$$H_r \underset{a}{<} H_s \Leftrightarrow H^*_r \underset{a}{<} H^*_s \ .$$

POSICION RELATIVA DE LA MODA RESPECTO DEL INTERVALO DE INDETERMINACION

Vimos en la Sección I.6 de la memoria que la moda M de una distribución hipergeométrica H(N,p,n) es el mayor entero menor o igual que (n+1)(pN+1)/(N+2), y, en el caso de que esta última cantidad fuese número entero, la distribución tendría dos valores modales:

$$\frac{(n+1)(pN+1)}{N+2} \qquad \text{y} \qquad \frac{(n+1)(pN+1)}{N+2} - 1.$$

Habitualmente, notaremos por M^* al valor cuya parte entera determina la moda, es decir,

$$M^* = \frac{(n+1)(pN+1)}{N+2} \ ;$$

$$M = [M^*].$$

También podemos expresar

$$M^* = p(n+1) + \frac{(n+1)(1-2p)}{N+2} \quad ,$$

como se comprueba fácilmente, resultando el manejo de esta expresión más útil a nuestros propósitos.

Veamos ahora cuál es el máximo valor que puede tomar la moda, cuando $p \leq 1/2$. Calculemos, en primer lugar, la siguiente diferencia:

$$M^* - \frac{n+1}{2} = \frac{(n+1)N}{2(N+2)} (2p-1) \quad .$$

A partir de aquí, tendremos que

$$p \leq 1/2 \Rightarrow M^* \leq \frac{n+1}{2} \quad ,$$

alcanzandose la igualdad sólo si $p=1/2$. En conclusión :

a) $p < 1/2 \Rightarrow M \leq n/2$

[4.14] b) $p=1/2$ y n par $\Rightarrow M=n/2$

c) $p=1/2$ y n impar $\Rightarrow$ la distribución es bimodal con $M_1 = \frac{n+1}{2}$ y $M_2 = \frac{n-1}{2}$.

Estos resultados serán utilizados más adelante.

4.17 LEMA:

Si $p \leq 1/2$, se verifica que

$$0 < M^* - \left[p(n-1) + \frac{p(n-1)}{N}\right] < 1.$$

Demostración:

La demostración de la desigualdad de la izquierda es trivial a partir de la definición de M^*. Puesto que si $p \leq 1/2$, M^* es mayor o igual que $p(n+1)$, es suficiente con comprobar que

$$p(n+1) > p(n-1) + \frac{p(n-1)}{N}.$$

Para demostrar la desigualdad de la derecha, partamos de que

[4.15] $$\frac{n+1}{N+2}(1-2p) \leq 1-2p,$$

lo cual es cierto por ser $p \leq 1/2$ y $(n+1)/(N+2) < 1$.

Restando al primer miembro de [4.15] la cantidad positiva $p(n-1)/N$, tendremos

$$\frac{n+1}{N+2}(1-2p) - \frac{p(n-1)}{N} < 1-2p,$$

y, finalmente, sumando y restando al primer miembro la cantidad pn y reordenando convenientemente la expresión, se obtiene

$$p(n+1) + \frac{(n+1)(1-2p)}{N+2} - \left[p(n-1) + \frac{p(n-1)}{N}\right] < 1,$$

donde los dos primeros sumandos constituyen la expresión de M*.

4.18 LEMA:

Si $p \leq 1/2$, la moda de cualquiera de las distribuciones hipergeometricas pertenecientes a la familia $H_m(mN,p,n)$ se localiza en la parte entera de $p(n-1)+\frac{p(n-1)}{N}$, o en la parte entera de $p(n-1)+\frac{p(n-1)}{N}+1$, localizandose en ambos valores en caso de bimodalidad.

Demostración:

Este lema es consecuencia inmediata del lema 4.17 y de la expresión del valor modal de la hipergeométrica como parte entera de M*.

Podemos hacer la siguiente interpretación del resultado anterior:

FACULTAD DE CIENCIAS
BIBLIOTECA
GRANADA

(i) Si el intervalo de indeterminación,

$$\left(p(n-1) - \frac{q(n-1)}{N},\ p(n-1) + \frac{p(n-1)}{N}\right)$$

no contiene a ningún número entero, la moda de las distribuciones de la familia $H_m(mN,p,n)$, con $0<p\leq 1/2$, se localiza en uno de los enteros adyacentes al intervalo, y en ambos en caso de bimodalidad.

(ii) Si el citado intervalo contiene a un número entero, la moda se localizará en dicho número entero o en el siguiente, y en ambos en caso de bimodalidad.

Estos resultados sobre la moda se aplicarán más adelante, cuando dispongamos de todos los recursos necesarios para poder abordar el teorema general sobre la aleatoriedad de la familia de hipergeométricas que estamos estudiando.

LOCALIZACION DE LA MODA EN LA FAMILIA $H_m(mN,p,n)$

Consideremos una distribución hipergeométrica cualquiera H_m, con la consabida restricción de $0<p\leq 1/2$. Como sabemos, su moda M_m se localiza en la parte entera de M_m^*,

siendo

[4.16] $$M_m^* = p(n+1) + \frac{(n+1)(1-2p)}{mN+2} \ .$$

4.19 LEMA:

M_m^* es una función decreciente en m, con

$$\lim_{m\to\infty} M_n^* = p(n+1).$$

Demostración:

La demostración es inmediata; basta tener en cuenta que, al ser $0<p\leq 1/2$, el segundo sumando de [4.16] es positivo y decrece conforme aumenta m.

4.20 LEMA:

Si p(n+1) es entero, la moda de todas las distribuciones pertenecientes a la familia $H_m(mN,p,n)$, con $0<p\leq 1/2$, se localiza en p(n+1).

Demostración:

Este resultado es, asimismo, inmediato; basta observar que si $0<p\leq 1/2$,

$$0 < \frac{(n+1)(1-2p)}{mN+2} < 1,$$

y, en este caso, $[M_m^*]=p(n+1)$, $\forall m$.

4.21 TEOREMA:

Si notamos por K a la parte entera de p(n+1), la moda de las distribuciones pertenecientes a la sucesión de hipergeométricas $H_m(mN,p,n)$, con $0<p\leq 1/2$, se localiza:

en K, si $m \geq \frac{n+1-2(K+1)}{N[K+1-p(n+1)]}$;

en K+1, si $m \leq \frac{n+1-2(K+1)}{N[K+1-p(n+1)]}$.

Demostración:

Teniendo en cuenta el lema 4.19, la moda se localizará en K a partir de un valor de m suficientemente grande: siempre que

[4.17] $$p(n+1) + \frac{(n+1)(1-2p)}{mN+2} \leq K+1.$$

Despejando el valor de m en la desigualdad [4.17],

se obtiene

$$m \geq \frac{n+1-2(K+1)}{N[K+1-p(n+1)]} ,$$

condición necesaria y suficiente para que la moda de la distribución H_m se localice en la parte entera de p(n+1).

En el caso de igualdad, la distribución sería bimodal, con valores modales K y K+1.

ALEATORIEDAD EN LA FAMILIA $H_m(mN, p, n)$, CUANDO LA MODA SE LOCALIZA EN LOS VALORES EXTREMOS DEL SOPORTE

En el lema 4.2 vimos que si $r<s$, $H_r(0) \leq H_s(0)$, verificandose la igualdad sólo si el valor cero no pertenece al soporte de ambas distribuciones; por tanto, si H_r y H_s localizan su moda en el valor cero, $H_r(0)<H_s(0)$, o lo que es lo mismo, por ser el cero el primer valor con probabilidad no nula,

$$h_r(0) < h_s(0),$$

por lo que las probabilidades puntuales en el valor modal es estrictamente creciente en la sucesión H_m .

Naturalmente, estamos implícitamente considerando que $0<p\leq 1/2$, para que la moda pueda localizarse en el cero. En el caso simétrico, $1/2\leq p<1$, la moda podría localizarse en el valor n, y, en este caso, se obtiene, análogamente, que

$$h_r(n) < h_s(n).$$

Este resultado refleja un comportamiento singular de la sucesión de hipergeométricas cuando la moda se localiza en cero o en n, ya que para cualquier otro valor modal, como vamos a demostrar más adelante, la sucesión de las probabilidades modales es estrictamente decreciente.

En el caso particular de la moda localizada en el cero, no podemos afirmar, de acuerdo con el criterio de Hickey, que $H_r <_{\alpha} H_s$. Sin embargo, tampoco puede afirmarse que $H_s <_{\alpha} H_r$, como podemos probar fácilmente con un contraejemplo. En efecto, consideremos las distribuciones hipergeométricas H_1 y H_2, pertenecientes a la familia $H_m(30m,0.1,3)$, que localizan la moda en el cero. Las probabilidades puntuales de ambas distribuciones son, respectivamente,

$h_1(0) = 0.72044$	$h_2(0) = 0.72484$
$h_1(1) = 0.25936$	$h_2(1) = 0.25091$
$h_1(2) = 0.01995$	$h_2(2) = 0.02367$
$h_1(3) = 0.00025$	$h_2(3) = 0.00058$

Si calculamos la entropía de ambas distribuciones, -medida de la aleatoriedad definida por $\Delta=-\sum_{i=1}^{n} h(x_i)L[h(x_i)]$- y la notamos, respectivamente, Δ_1 y Δ_2, se obtiene

$$\Delta_1 = 0.6664118$$

$$\Delta_2 = 0.6731123$$

por lo que al ser $\Delta_1<\Delta_2$, no es cierto que $H_2 <_{\alpha} H_1$.

Queremos hacer notar aquí que el resultado $\Delta_1<\Delta_2$ estaba, en cualquier caso, garantizado por nuestro teorema 4.9.

Tenemos, en conclusion, que cuando la moda se localiza en cero o en n, no es posible comparar de manera absoluta la aleatoriedad de la sucesión; en tal caso sólo podemos dar resultados sobre medidas concretas. Si seguimos considerando el ejemplo anterior y tomamos ahora otra medida de aleatoriedad, como es la opuesta de la energía informacional de Onicescu, definida ésta por $\aleph=\sum_{i=1}^{n}[h(x_i)]^2$, obtenemos que

$$\aleph_1 = 0.5866994$$

$$\aleph_2 = 0.5889094$$

por lo que al ser $-\aleph_2<-\aleph_1$, esta medida tiene un comportamiento

contrario al de la entropía.

Finalmente, se podría estudiar hasta qué orden α estaría garantizado el crecimiento de la entropía generalizada o medida de información de Rényi, que, como vimos en la sección III.3, es otra de las medidas usuales de aleatoriedad.

MONOTONIA DE LA PROBABILIDAD MODAL

Consideremos la familia de hipergeométricas $H_m(mN,p,n)$, con $0<p\leq 1/2$ y con soporte $(h_m)=\langle 0,1,\ldots,n\rangle$, $\forall m$.

La primera de las restricciones, por razones de simetría, como ya pusimos de manifiesto, no va a suponer ninguna pérdida de generalidad, mientras que la segunda de ellas nos obliga, en el peor de los casos, a asegurar la validez de los resultados que obtengamos, sólo a partir de un cierto término de la sucesión en adelante: para m lo suficientemente grande como para que $mpN \geq n$.

Si llamamos K a la parte entera de $p(n+1)$, vimos en el teorema 4.21 que la moda de las distribuciones de la familia H_m se localiza en K+1 si

$$m \leq \frac{n+1-2(K+1)}{N[K+1-p(n+1)]} ,$$

y en K si m es mayor o igual que dicho valor. (En el caso de igualdad la distribución sería bimodal). Como es obvio, puede ocurrir que todas las distribuciones de la sucesión localicen la moda en K.

El estudio de la monotonía de la probabilidad modal lo vamos a realizar en dos pasos: comenzaremos restringiendonos a las distribuciones de la sucesión que localicen la moda en el mismo valor M, para seguidamente generalizar el resultado. En cualquier caso, excluiremos los valores modales extremos 0 y n, que presentan un comportamiento singular, caso que ha sido estudiado en el apartado anterior.

La demostración del teorema que asegura el decrecimiento de las probabilidades modales es larga y laboriosa. Buscando una mayor claridad en el desarrollo de la misma, la hemos descompuesto en una serie de lemas previos, interconectados, que desembocan finalmente en el teorema.

Por otra parte, por la citada conexión de los lemas, hemos prescindido de explicitar en las hipótesis aquello que es común a todos ellos, como el significado de los parámetros intervinientes y el campo de valores de los mismos, de todo lo cual, por otra parte, el lector tiene conocimiento previo.

4.22 LEMA:

Si M es la moda de una distribución hipergeométrica H(N,p,n), con $0<p\leq 1/2$,

$$M-1 \leq p(n-1),$$

verificandose la igualdad sólo cuando p=1/2 y n es impar.

Demostración:

Sabemos que el valor modal M se localiza en la parte entera de M*, donde

$$M^* = p(n+1) + \frac{(n+1)(1-2p)}{N+2} .$$

De aquí,

$$M^*-p(n-1) = 2p + \frac{(n+1)(1-2p)}{N+2} .$$

El segundo miembro de esta última igualdad, considerado como función de p, es creciente, como puede comprobarse fácilmente; el máximo valor lo alcanzará cuando p=1/2, obteniendose, en este caso,

[4.18] $$M^*-p(n-1) = 1.$$

Tenemos, por tanto, que

$$M^*-p(n-1) \leq 1,$$

lo cual implica que

$$M-p(n-1) \leq 1,$$

quedando demostrada la primera parte del lema.

Por otra parte, teniendo en cuenta [4.18],

$$M-p(n-1)=1 \Leftrightarrow \begin{cases} p=1/2 \\ \text{y} \\ M=M^*=\frac{n+1}{2} \end{cases} \Leftrightarrow \begin{cases} p=1/2 \\ \text{y} \\ n \text{ es impar} \end{cases}$$

4.23 LEMA:

Sea M la moda de una distribución hipergeométrica H(N,p,n), con $0<p\leq 1/2$ y $M\geq 1$. Consideremos los vectores

$$X(\frac{1}{p},\frac{2}{p},\ldots,\frac{M-1}{p}), \quad Y(0,\frac{1}{q},\frac{2}{q},\ldots,\frac{n-M-1}{q}), \quad Z(1,2,\ldots,n-1).$$

Se cumple:

(i) Todas las componentes del vector X y las primeras componentes del vector Y se encuentran mayoradas, cada una de ellas por una componente distinta, por las j_M-1 primeras componentes del vector Z, siendo j_M el menor entero mayor o igual que M/p .

Si M=1, el vector X no estaría definido; en este caso, el proceso de mayoración anteriormente descrito se cumple entre las componentes de Y y Z.

(ii) Si $j_M > n-1$, o lo que es lo mismo, si $p(n-1) < M$, entonces cada componente de X y de Y está mayorada, cada una de ellas por una componente distinta de Z.

(iii) Si $j_M \leq n-1$, las últimas $n-j_M$ componentes de Z están mayoradas por las correspondientes $n-j_M$ últimas componentes de Y.

Demostración:

(i) Llamemos j_i al menor entero mayor o igual que i/p. Las componentes del vector Z pueden ser expresadas de la siguiente forma:

$$Z(1,\ldots,j_1,\ldots,j_2,\ldots,j_{M-1},\ldots,(j_M),\ldots,n-1).$$

La existencia de la componente j_{M-1} está garantizada

por el lema 4.22, el cual nos asegura que

$$\frac{M-1}{p} \leq n-1,$$

por lo que, en virtud de la definición de j_{M-1},

$$\frac{M-1}{p} \leq j_{M-1} \leq n-1.$$

No podemos asegurar lo mismo para la componente j_M, que puede no existir; j_M será mayor que n-1 cuando $p(n-1)<M$.

Como consecuencia inmediata de la definición de los j_i, tendremos que la componente i del vector X esta mayorada por la componente j_i del vector Z, para i:1,...,M-1, de forma que cada componente de X está mayorada por una componente distinta de Z.

Las restantes componentes de Z las podemos expresar como sigue:

[4.19] j_i+s, con $s:1,2,\ldots,j_{i+1}-j_i-1$
$i=0,1,\ldots,M-1$
$j_o=0$

a las que habría que añadir, en el caso que $j_M \leq n-1$, las últimas componentes, que serían de la forma

[4.20] j_M+s, con $s:0,1,\ldots,n-1-j_M$

Por otra parte, las componentes de Y pueden ser expresadas de la forma

[4.21] $\frac{j_i+s-(i+1)}{q}$, con $s:1,2,\ldots,j_{i+1}-j_i-1$ $i=0,1,\ldots,M-1$

a las que habría que añadir, en el caso que $j_M \leq n-1$, las últimas componentes, que serían de la forma

[4.22] $\frac{j_M+s-M}{q}$, con $s:0,1,\ldots,n-1-j_M$

Vamos ahora a comprobar que cada componente de Y, de la forma [4.21], está mayorada por la correspondiente componente de Z, de la forma [4.19].

En efecto: de la definición de los j_i tenemos que

$$j_{i+1}-1 < \frac{i+1}{p} ,$$

y de aquí,

$$j_i+[j_{i+1}-j_i-1] < \frac{i+1}{p} ;$$

por tanto,

[4.23] $$j_i+s < \frac{i+1}{p}\ , \qquad \forall s:1,\ldots,j_{i+1}-j_i-1.$$

Si en [4.23] sustituimos p por 1-q y reordenamos convenientemente los términos, se obtiene

$$j_i+s > \frac{j_i+s-(i+1)}{q} \qquad \forall s:1,\ldots,j_{i+1}-j_i-1,$$

quedando demostrada la proposición, y con ello, el primer apartado del lema.

(ii) Este apartado es una consecuencia inmediata del apartado (i) para el caso que $j_M > n-1$.

(iii) Para demostrar este apartado tenemos que ver que cada componente de Z, de la forma [4.20], está mayorada por la correspondiente componente de Y, de la forma [4.22], es decir, hay que probar que

[4.24] $$j_M+s < \frac{j_M+s-M}{q}\ , \qquad \forall s:0,1,\ldots,n-1-j_M\ .$$

En efecto, si en [4.24] sustituimos q por 1-p, y reordenamos convenientemente los términos, se obtiene la condición equivalente

$$j_M+s > \frac{M}{p}\ ,$$

condición que se verifica $\forall s \geq 0$, por ser j_M el menor

entero mayor o igual que M/p.

4.24 LEMA:

$$n - j_M \leq j_1 - 1$$

Demostración:

En virtud del teorema 4.21, sabemos que

$$p(n+1) \leq M+1,$$

donde la igualdad se alcanza sólo si p(n+1) es entero y p=1/2, dandose, en este caso, una situación de bimodalidad. Dividiendo en la desigualdad anterior los dos miembros por p, y reordenando términos, se obtiene

[4.25] $$n - \frac{M}{p} \leq \frac{1}{p} - 1.$$

Por otra parte, por ser j_i el menor entero mayor o igual que i/p, tenemos que

[4.26] $$j_1 \geq \frac{1}{p}$$

[4.27] $$j_M \geq \frac{M}{p}$$

Utilizando conjuntamente las desigualdades [4.25], [4.26] y [4.27], se obtiene

$$n-j_M \leq n - \frac{M}{p} \leq \frac{1}{p} - 1 \leq j_1 - 1,$$

por lo que

$$n-j_M \leq j_1-1.$$

4.25 LEMA:

$$n-j_M < \frac{n}{M+1}$$

Demostración:

Por la localización del valor modal, sabemos que

[4.28] $$pn < M+1 .$$

Si multiplicamos los dos miembros de la desigualdad anterior por M ($M \geq 1$) y sumamos a ambos miembros la cantidad pn, se obtiene

$$(pn-M)(M+1) < pn,$$

y de aquí,

$$n - \frac{M}{p} < \frac{n}{M+1} .$$

Finalmente, como $j_M \geq \frac{M}{p}$,

[4.29] $$n - j_M < \frac{n}{M+1} .$$

4.26 LEMA:

[4.30] $$\left(i - \frac{i-1}{q}\right)(n-i)\left(\frac{n-M-i}{q}\right) - \left(\frac{n-M-i}{q} - n+i\right)i\left(\frac{i-1}{q}\right) > 0$$

$\forall i: 1, 2, \ldots, n - j_M$.

Demostración:

Analizaremos, en primer lugar, cada uno de los factores de [4.30].

a) Como consecuencia inmediata de los valores que toma i, tenemos:

$$i > 0$$

$$n - i > 0$$

$$\frac{i-1}{q} \geq 0$$

b) A partir de las desigualdades [4.28] y [4.29] se obtiene

$$i \leq n-j_M < \frac{n}{M+1} < \frac{1}{p} .$$

De aquí,

$$i < \frac{1}{p} ,$$

y sustituyendo p por 1-q en esta última desigualdad y reordenando términos se obtiene

$$i - \frac{i-1}{q} > 0 .$$

c) Sabemos, por nuestras hipótesis de partida que excluyen los valores modales o y n, que $M \leq n-1$.

Si multiplicamos los dos miembros de la desigualdad anterior por M y le sumamos a cada uno de ellos la cantidad n, y, finalmente, reordenamos términos, se obtiene

$$\frac{n}{M+1} \leq n-M .$$

Teniendo en cuenta ahora que $i < \frac{n}{M+1}$, tendremos que $i < n-M$, de donde

$$\frac{n-M-i}{q} > 0 .$$

Una vez analizado el signo de los factores de [4.30], tenemos que el minuendo es siempre positivo,

mientras que el signo del sustraendo será el del factor

$$\frac{n-M-i}{q} \; -n+i.$$

Por ello, para probar el lema será suficiente con demostrar

[4.31] (i) $$i - \frac{i-1}{q} > \frac{n-M-i}{q} - n+i$$

[4.32] (ii) $$n-i > i$$

[4.33] (iii) $$\frac{n-M-i}{q} > \frac{i-1}{q}$$

Operando en [4.31] se obtiene la desigualdad equivalente $pn<M+1$, que sabemos que es cierta.

La desigualdad [4.32] es asimismo fácil de comprobar, puesto que al ser $M \geq 1$,

$$i < \frac{n}{M+1} \leq \frac{n}{2} .$$

Para demostrar [4.33] partiremos de que $1 \leq M \leq n-1$:

$$n \geq M+1 \Rightarrow n(1-M) \leq 1-M^2 \Rightarrow n-nM \leq 1-M^2$$

$$\Rightarrow 2n+M-nM \leq n+M+1-M^2$$

$$\Rightarrow 2n \leq (M+1)(n-M+1)$$

$$\Rightarrow \frac{n}{M+1} \leq \frac{n-M+1}{2}$$

y, puesto que $i<\frac{n}{M+1}$, se obtiene

$$i < \frac{n-M+1}{2} \Rightarrow n-M-i > i-1 \Rightarrow \frac{n-M-i}{q} > \frac{i-1}{q} .$$

4.27 LEMA:

Si $p(n-1)>M$ e $i:1,2,\ldots,n-j_M$, entonces

$$-2i(n-i)+2\left(\frac{i-1}{q}\right)\left(\frac{n-M-i}{q}\right)$$

es creciente en i.

Demostración:

Consideremos la función

$$f(i) = -2i(n-i) + 2(\frac{i-1}{q})(\frac{n-M-i}{q}),$$

y su derivada,

$$f'(i) = -2n +4i + \frac{2}{q^2} (n-M-2i+1),$$

que debemos probar que es positiva.

Partamos de que $p(n-1)>M$ y $M\geq 1$, por lo que $pn>1$ y, de aquí, $pnq>q$.

Podemos afirmar, por tanto, que

$$p(n-1)+pnq > M+q.$$

Sumando a ambos miembros de la desigualdad la cantidad p,

$$pn+pnq > M+1;$$

multiplicando ambos miembros por M-1 y reordenando convenientemente se obtiene

$$\frac{M-1}{M+1}\, n(1-q^2) \geq M-1,$$

o, equivalentemente,

[4.34] $$\left(n - \frac{2n}{M+1}\right)(1-q^2) \geq M-1.$$

Puesto que por el lema 4.25

$$i \leq n-j_M < \frac{n}{M+1},$$

[4.34] implica que

$$(n-2i)(1-q^2) > M-1.$$

Multiplicando por 2 ambos miembro de esta última

expresión, y reordenando términos, se obtiene, finalmente,

$$\frac{2}{q^2}\,(n-2i-M+1) \rangle 2n-4i,$$

quedando probado, de esta forma, que $f'(i)\rangle 0$ y, consecuentemente, el lema.

4.28 LEMA:

Si $p(n-1) \rangle M$,

[4.35] $$\frac{M+1-pn}{q}\,N^2 + \left[2\,\frac{i-1}{q}\,\frac{n-M-i}{q} - 2i(n-i)\right]N \geq 0$$

para $i:1,2,\ldots,n-j_M$.

Demostración:

Por el teorema 4.21 sobre la localización de la moda, sabemos que $pn\langle M+1$, por lo que el coeficiente de N^2 en [4.35] es siempre positivo e independiente de i. Por el lema 4.27, sabemos que el coeficiente de N en la citada expresión es decreciente en i; por lo tanto, para demostrar el lema será suficiente probar que [4.35] se verifica para el menor valor de i, es decir, para i=1.

Debemos demostrar, pues, que

$$\frac{M+1-pn}{q} N^2 - 2(n-1)N \geq 0$$

o, equivalentemente, que

[4.36] $$(M+1)N \geq pnN + 2q(n-1)$$

Por la hipótesis del lema, tenemos que $pn \geq M$, y, puesto que $q \geq 1/2$,

$$pn \geq M \Rightarrow pn \geq M + \frac{1}{2} - q \Leftrightarrow (1-q)n \geq M + \frac{1}{2} - q$$

$$\Leftrightarrow n - M - \frac{1}{2} \geq qn - q \Leftrightarrow 2n - 2M - 1 \geq 2qn - 2q$$

$$\Leftrightarrow pnN + 2n - 2M - 1 \geq pnN + 2qn - 2q$$

[4.37] $$\Rightarrow pnN + n + pN - 2M - 1 \geq pnN + 2qn - 2q,$$

puesto que, por nuestra hipótesis de partida, sabemos que $pN \geq n$.

Finalmente, agrupando términos en [4.37], podemos escribir

[4.38] $$(n+1)(pN+1) - 2(M+1) \geq pnN + 2q(n-1) .$$

Por otra parte, recordemos que al ser M valor modal, se cumple que

$$\frac{(n+1)(pN+1)}{N+2} \leq M+1,$$

y de aquí se obtiene fácilmente que

[4.39] $$(M+1)N \geq (n+1)(pN+1)-2(M+1).$$

De [4.38] y [4.39] resulta

$$(M+1)N \geq pnN+2q(n-1),$$

quedando probado [4.36] y, consecuentemente, el lema.

4.29 LEMA:

Para $i:1,2,\ldots,n-j_M$ y $p(n-1)>M$,

[4.40] $$\left(i-\frac{i-1}{q}\right)\left(N-\frac{n-M-i}{q}\right)(N-n+i)-\left(\frac{n-M-i}{q}-n+i\right)(N-i)\left(N-\frac{i-1}{q}\right) > 0$$

Demostración:

El primer miembro de [4.40] es una función cuadrática en N, que podemos expresar como

$$\frac{M+1-pn}{q}N^2 + \left[2\,\frac{i-1}{q}\,\frac{n-M-i}{q} - 2i(n-i)\right]N +$$

$$+ \left(i - \frac{i-1}{q}\right)(n-i)\,\frac{n-M-i}{q} - \left(\frac{n-M-i}{q} - n+i\right)i\,\frac{i-1}{q} ,$$

expresión que sabemos que es positiva, como consecuencia inmediata de los lemas 4.26 y 4.28.

4.30 TEOREMA:

Sea la sucesión de hipergeométricas $H_m(mN,p,n)$, con $0<p\leq 1/2$ y con soporte $(h_m)=\langle 0,1,\ldots,n\rangle$, $\forall m$.

Se cumple que $h_m(M;mN,p,n)$ es decreciente en m, siempre que las distribuciones que se generan al variar m localicen su moda en M, donde $1\leq M\leq n-1$.

Demostración:

Por una mayor economía operativa, en lugar de demostrar el decrecimiento en m de $h_m(M;mN,p,n)$, probaremos el decrecimiento en N , de $h(M;N,p,n)$, manteniendo inalterables el resto de las hipótesis del teorema.

Las probabilidades puntuales en el valor modal, $h(M;N,p,n)$, serán notadas, abreviadamente, por $h(M)$.

Recordemos que

$$h(M) = \frac{\binom{pN}{M}\binom{qN}{n-M}}{\binom{N}{n}} = \binom{n}{M}\frac{pN!}{(pN-M)!}\frac{qN!}{(qN-n+M)!}\frac{(N-n)!}{N!}$$

Como $\left[\begin{smallmatrix} n \\ M \end{smallmatrix} \right]$ es independiente de N, bastará comprobar el decrecimiento de la expresión

$$\frac{pN!}{(pN-M)!}\ \frac{qN!}{(qN-n+M)!}\ \frac{(N-n)!}{N!}\ ,$$

cuyo logaritmo es

$$[4.41] \qquad \sum_{j=o}^{M-1} L(pN-j) + \sum_{j=o}^{n-M-1} L(qN-j) - \sum_{j=o}^{n-1} L(N-j).$$

Si consideramos una variación continua del parámetro N en el intervalo de valores compatibles con nuestra hipótesis, una condición suficiente para el decrecimiento de h(M) resultará de exigir que la primera derivada de [4.41], respecto de N,

$$\sum_{j=o}^{M-1} \frac{p}{pN-j} + \sum_{j=o}^{n-M-1} \frac{q}{qN-j} - \sum_{j=o}^{n-1} \frac{1}{N-j},$$

sea negativa.

Esta última expresión puede ser también escrita como

$$[4.42] \qquad \sum_{j=1}^{M-1} \frac{1}{N-j/p} + \sum_{j=o}^{n-M-1} \frac{1}{N-j/q} - \sum_{j=1}^{n-1} \frac{1}{N-j}\ ,$$

siempre que M>1, y en el caso particular de que M=1, la

expresión resultante sería

$$\sum_{j=0}^{n-2} \frac{1}{N-j/q} - \sum_{j=1}^{n-1} \frac{1}{N-j} \cdot$$

Hacemos notar que el caso M=n-1 no podrá darse al estar impuesta la restricción $p \leq 1/2$.

Consideremos ahora los vectores

$$X\left(\frac{1}{p}, \frac{2}{p}, \ldots, \frac{M-1}{p} \right), \; Y\left(0, \frac{1}{q}, \ldots, \frac{n-M-1}{q} \right), \; Z(1,2,\ldots,n-1)$$

Sabemos, por el apartado segundo del lema 4.23, que si $p(n-1)<M$, entonces cada componente de X y de Y está mayorada, cada una de ellas por una componente distinta de Z. Por tanto, si $p(n-1)<M$, en [4.42] se verifica que cada sumando de las dos primeras sumatorias está mayorado por un sumando distinto de la última sumatoria; consecuentemente, la expresión [4.42] es negativa y el teorema estaría demostrado.

Si $p(n-1) \geq M$, por el primer apartado del lema 4.23, sabemos que todas las componentes del vector X y las primeras componentes del vector Y se encuentran mayoradas, cada una de ellas por una componente distinta, por la j_M-1 primeras componentes del vector Z, donde j_M

es el menor entero mayor o igual que M/p, y, en virtud del tercer apartado del mismo lema, las últimas $n-j_M$ componentes de Z están mayoradas por las correspondientes $n-j_M$ últimas componentes de Y.

Todo ello, aplicado a la expresión [4.42], nos permite asegurar que, salvo los últimos $n-j_M$ sumandos de la segunda sumatoria, los sumandos restantes de las sumatorias del minuendo están mayorados, cada uno de ellos por un sumando distinto de la sumatoria sustraendo.

Para probar que la expresión [4.42] es negativa, será suficiente con probar que la aportación positiva de esos $n-j_M$ últimos sumandos queda ampliamente compensada por la aportacion negativa del resto de los sumandos.

Concretando, una condición suficiente para que [4.42] sea menor que cero es que se verifique

$$[4.43] \qquad \sum_{j=0}^{j_1-2} \frac{1}{N-j/q} + \sum_{j=j_M-M}^{n-M-1} \frac{1}{N-j/q} < \sum_{j=1}^{j_1-1} \frac{1}{N-j} + \sum_{j=j_M}^{n-1} \frac{1}{N-j} .$$

Por el lema 4.24, sabemos que $n-j_M \leq j_1-1$; por tanto, las primeras sumatorias de cada miembro de [4.43] pueden ser descompuestas como sigue:

$$\sum_{j=0}^{j_1-2} \frac{1}{N-j/q} = \sum_{j=0}^{n-j_M-1} \frac{1}{N-j/q} + \sum_{j=n-j_M}^{j_1-2} \frac{1}{N-j/q}$$

$$\sum_{j=1}^{j_1-1} \frac{1}{N-j} = \sum_{j=1}^{n-j_M} \frac{1}{N-j} + \sum_{j=n-j_M+1}^{j_1-1} \frac{1}{N-j} \ .$$

Por el lema 4.23, se cumple que

$$\sum_{j=n-j_M}^{j_1-2} \frac{1}{N-j/q} < \sum_{j=n-j_M+1}^{j_1-1} \frac{1}{N-j} \ .$$

En virtud de lo anterior, para probar que la condición [4.43] es cierta, será suficiente con probar que

[4.44] $$\frac{1}{N-\frac{i-1}{q}} + \frac{1}{N-\frac{n-M-i}{q}} < \frac{1}{N-i} + \frac{1}{N-(n-i)} \ ,$$

$$\text{con } i:1,2,\ldots,n-j_M \ .$$

Teniendo en cuenta que todos los denominadores son positivos, [4.44] puede ser escrita como

[4.45] $$\left(i-\frac{i-1}{q}\right)\left(N-\frac{n-M-i}{q}\right)(N-n+i)-\left(\frac{n-M-i}{q}-n+i\right)(N-i)\left(N-\frac{i-1}{q}\right) > 0$$

$$\text{con } i:1,2,\ldots,n-j_M,$$

donde, por hipótesis, $p(n-1)>M$.

La expresión [4.45] se cumple ciertamente, como vimos en el lema 4.29, por lo que, finalmente, el teorema queda demostrado.

Por último, vamos a generalizar el resultado del teorema 4.30, eliminando la restricción de que las distribuciones de la sucesión H_m tengan que localizar la moda en un mismo valor M.

4.31 TEOREMA:

Sea la sucesión de hipergeométricas $H_m(mN,p,n)$, con $0<p\leq 1/2$ y con soporte $(h_m)=\langle 0,1,\ldots,n\rangle$, $\forall m$, y sea M_m la moda de la distribución H_m.

Si $1\leq M_m \leq n-1$, $\forall m$, se cumple que $h_m(M_m;mN,p,n)$ es decreciente en m.

Demostración:

Llamemos

$$m^* = \frac{n+1-2(K+1)}{N[K+1-p(n+1)]} ,$$

donde K es la parte entera de p(n+1).

Por el teorema 4.21, sabemos que

$$m\leq m^* \Rightarrow M_m=K+1$$

$$m\geq m^* \Rightarrow M_m=K ,$$

donde se aprecia que el caso de bimodalidad se corresponde con $m=m^*$.

Si obviamos el hecho de que m^* puede no ser entero, lo que obliga a que la definición de las probabilidades hipergeométricas trascienda el modelo de urnas y no exija valores enteros para los parámetros tamaño poblacional y tamaño de la clase, lo que, por otra parte, no perturba el resultado sobre la localización de la moda, podemos afirmar:

(i) $h_m(K+1;mN,p,n)$ es decreciente en m, $\forall m \leq m^*$.

(ii) $h_{m*}(K+1;m^*N,p,n)=h_{m*}(K;m^*N,p,n)$.

(iii) $h_m(K;mN,p,n)$ es decreciente en m, $\forall m \geq m^*$.

Las afirmaciones (i) y (iii) son consecuencia directa de la aplicación del teorema 4.30, y (ii) es consecuencia de la bimodalidad.

Con esto, queda demostrado el decrecimiento de la probabilidad modal en la familia H_m.

Como ya anunciamos, la restricción impuesta sobre el parámetro p, puede ser eliminada por razones de simetría ya que si la distribución H(N,p,n) localiza su moda en M, la

distribución H(N,q,n) lo hará en n-M, y

$$h(M;N,p,n) = h(n-M;N,q,n).$$

A todos los efectos, la generalización del último teorema está asegurada y podrá ser reformulado como sigue:

4.32 TEOREMA:

Sea la sucesión de hipergeométricas $H_m(mN,p,n)$, con soporte $(h_m)=\langle 0,1,\ldots,n\rangle$, $\forall m$, y sea M_m la moda de la distribución H_m .

Si $1\leq M_m \leq n-1$, $\forall m$, se cumple que $h_m(M_m;mN,p,n)$ es decreciente en m.

4.33 COROLARIO:

Puesto que la sucesión $H_m(mN,p,n)$ converge hacia la binomial B(n,p), bajo las hipótesis del teorema anterior, se cumple que

$$h_m(M_m) \rightarrow b(M), \quad \forall m:1,2,\ldots$$

siendo M la moda de B(n,p).

LOCALIZACION DE LAS DOS MAYORES PROBABILIDADES PUNTUALES

La localización de las dos mayores probabilidades puntuales la determinaremos a través de la comparación de las probabilidades en los enteros adyacentes a la moda, ofreciendo, en los dos siguientes teoremas, condiciones suficientes para poder afirmar, en cada caso, a cuál de dichos valores le corresponde mayor probabilidad.

Como es natural, seguimos manteniendo todas nuestras hipótesis previas, y, entre ellas, la exigencia de ser $M \geq 1$ y $pN \geq n$.

4.34. TEOREMA:

Si $p \leq 1/2$ y $pn \leq M$, entonces

$$h(M-1) \geq h(M+1) ,$$

verificandose la igualdad sólo si $p=1/2$ y n es par.

Demostración:

a.1) Si $p=1/2$, y n es par, sabemos que la distribución es simétrica y con un único valor modal en $n/2$, por lo que, evidentemente, $h(M-1)=h(M+1)$.

a.2) Si $p=1/2$ y n es impar, la distribución es bimodal, localizando su moda en $M=\frac{n+1}{2}$ y $M-1=\frac{n-1}{2}$ (vease [4.14]), por lo que resulta claro que $h(M-1)>h(M+1)$.

b) Si $p<1/2$, vamos a partir de la equivalencia de las dos siguientes condiciones:

$$h(M-1) > h(M+1)$$

$$[4.46] \quad M(M+1)(qN-n+M)(qN-n+M+1) > (n-M)(n-M+1)(pN-M)(pN-M+1)$$

como se puede comprobar sustituyendo las probabilidades hipergeométricas por sus expresiones correspondientes.

Si $pn\leq M$, entonces

$$[4.47] \quad M = pn+d, \quad \text{con } d\geq 0$$

Sustituyendo, en [4.46], M por su expresión [4.47], se obtiene

$$[4.48] \quad (pn+d)(pn+d+1)(q[N-n]+d)(q[N-n]+d+1) > (qn-d)(qn-d+1)(p[N-n]-d)(p[N-n]-d+1).$$

Debemos, por tanto, probar que bajo las hipotesis del teorema, [4.48] es cierta. Para ello, teniendo en cuenta que todos los factores de la citada expresión son positivos, será suficiente probar

(i) $(pn+d)(q[N-n]+d) \geq (qn-d)(p[N-n]-d)$

(ii) $(pn+d+1)(q[N-n]+d+1) > (qn-d+1)(p[N-n]-d+1)$.

En efecto,

(i) $(pn+d)(q[N-n]+d) \geq pn\ q(N-n) = qn\ p(N-n) \geq$
$\geq (qn-d)(p[N-n]-d)$.

(ii) $p < 1/2 \Rightarrow N-n > n \Leftrightarrow (q-p)(N-n) > (q-p)n$
$\Leftrightarrow pn+q(N-n) > qn+p(N-n)$
$\Leftrightarrow pn+q(N-n)+pnq(N-n)+1 > qn+p(N-n)+pnq(N-n)+1$
$\Leftrightarrow (pn+1)(q[N-n]+1) > (qn+1)(p[N-n]+1)$.

De aquí,

$(pn+d+1)(q[N-n]+d+1) \geq (pn+1)(q[N-n]+1) >$
$> (qn+1)(p[N-n]+1) \geq (qn-d+1)(p[N-n]-d+1)$,

quedando demostrado el teorema.

4.35 TEOREMA:

Si $p \leq 1/2$ y $p(n+1) > M+q$, entonces

$$h(M+1) > h(M-1)$$

Demostración:

Por [4.46] sabemos que la condición a demostrar es

$$[4.49] \qquad (n-M)(n-M+1)(pN-M)(pN-M+1) > M(M+1)(qN-n+M)(qN-n+M+1)$$

Puesto que

$$p(n+1) > M+q \quad \Rightarrow \quad pn > M+q-p \ ,$$

podemos escribir

$$[4.50] \qquad M = pn-d, \quad \text{con } d>q-p \ .$$

Sustituyendo, en [4.49], M por su expresión [4.50], se obtiene

$$[4.51] \qquad \begin{aligned} &(qn+d)(qn+d+1)(p[N-n]+d)(p[N-n]+d+1) > \\ &> (pn-d)(pn-d+1)(q[N-n]-d)(q[N-n]-d+1). \end{aligned}$$

Debemos, por tanto, probar que, bajo las hipótesis del teorema, se cumple [4.51]. Para ello, teniendo en cuenta que todos los factores de esta última expresión son positivos, será suficiente probar

$$\text{(i)} \qquad (qn+d)(p[N-n]+d) > (pn-d)(q[N-n]-d)$$

$$\text{(ii)} \qquad (qn+d+1)(p[N-n]+d+1) > (pn-d+1)(q[N-n]-d+1).$$

En efecto,

(i) $(qn+d)(p[N-n]+d) > qn\ p[N-n] = pn\ q[N-n] >$

$> (pn-d)(q[N-n]-d).$

(ii) Partamos de que $d>q-p\geq 0$

$$d+p > q \Leftrightarrow (d+p)[N-n] > q[N-n] \Rightarrow$$

$$\Rightarrow (d+p)[N-n]+(p+q)nd+(q-p)n+4d > q[N-n].$$

Sumando a cada miembro de esta última desigualdad, la cantidad $qnp[N-n]+d^2+1$, y reordenando convenientemente los términos, se obtiene

$$(qn+d+1)(p[N-n]+d+1) > (pn-d+1)(q[N-n]-d+1),$$

quedando demostrada la segunda de las condiciones, y, consecuentemente, el teorema.

LOCALIZACION DE LAS TRES MAYORES PROBABILIDADES PUNTUALES

La determinaremos a través de la comparación de las probabilidades $h(M+1)$ y $h(M-2)$.

4.36 TEOREMA:

$$pn \geq M \Rightarrow h(M+1) > h(M-2)$$

Demostración:

Si $M<2$, $h(M-2)=0$ (recordemos que, por nuestras hipótesis previas, el soporte de las distribuciones es de la forma $\{0,1,\ldots,n\}$), por lo que el teorema se verifica evidentemente.

Consideremos $M \geq 2$. Puesto que M es valor modal, se cumple que $h(M)/h(M-1) \geq 1$; por ello, para demostrar el teorema, será suficiente probar que

$$\frac{h(M+1)}{h(M-2)} > \frac{h(M)}{h(M-1)} ,$$

o, lo que es lo mismo, que

$$\frac{h(M-1)}{h(M-2)} > \frac{h(M)}{h(M+1)} .$$

Sustituyendo las probabilidades hipergeométricas por sus expresiones correspondientes, se obtiene la condición equivalente

$$(n-M)(n-M+2)(pN-M)(pN-M+2) > \quad [4.52]$$

$$> (M-1)(M+1)(qN-n+M-1)(qN-n+M+1)$$

Si $pn \geq M$, entonces

$$M = pn-d, \quad \text{con } 0 \leq d < 1 \quad . \quad [4.53]$$

Sustituyendo, en [4.52], M por su expresión [4.53], se obtiene

$$(qn+d)(qn+d+2)(p[N-n]+d)(p[N-n]+d+2) \rangle \quad [4.54]$$

$$\rangle (pn-d-1)(pn-d+1)(q[N-n]-d-1)(q[N-n]-d+1).$$

Hemos obtenido, por tanto, una condición suficiente que vamos a demostrar; para ello tendremos en cuenta que todos los factores de [4.54] son positivos.

Se cumple la siguiente secuencia de desigualdades:

$$(qn+d)(qn+d+2)(p[N-n]+d)(p[N-n]+d+2) \rangle$$

$$\rangle qn\ qn\ p[N-n]\ p[N-n] = p^2n^2\ q^2[N-n]^2 \rangle$$

$$\rangle (p^2n^2-1)(q^2[N-n]^2-1) = (pn-1)(pn+1)(q[N-n]-1)(q[N-n]+1) \rangle$$

$$\rangle (pn-1-d)(pn+1-d)(q[N-n]-1-d)(q[N-n]+1-d),$$

quedando, de esta forma, probada la desigualdad [4.54], y, consecuentemente, el teorema.

ALEATORIEDAD DE LAS DISTRIBUCIONES Hm(mN, p, n)

Vamos a abordar, a continuación, la generalización del teorema 4.15 que aseguraba la monotonía creciente de la

aleatoriedad de las distribuciones pertenecientes a la sucesión de hipergeométricas Hm(mN,p,n), en el caso particular de que la media fuese un número entero; eliminaremos, por tanto, la restricción impuesta sobre el valor medio de las distribuciones.

Remitimos al lector al apartado de observaciones previas a la generalización del teorema 4.15, para recordar que la supresión de la citada restricción da lugar a una variada casuística que debemos, como paso previo, analizar, y en la que juega un papel fundamental el intervalo de indeterminación

$$I = \left[p(n-1) - \frac{q(n-1)}{N},\ p(n-1) + \frac{p(n-1)}{N}\right],$$

y su posición relativa respecto del valor modal M.

Teniendo en cuenta el lema 4.18, los casos que se pueden presentar los vamos a agrupar en :

(i) $M-1 \leq p(n-1) - \frac{q(n-1)}{N} < p(n-1) + \frac{p(n-1)}{N} \leq M$

(ii) $M-1 \in I$

(iii) $M < p(n-1) + \frac{p(n-1)}{N} < M+1$

Antes de pasar al análisis de cada uno de los casos, hacemos observar que, en todos ellos, se pretende comparar en

aleatoriedad dos distribuciones H_r y H_s, con $r<s$, pertenecientes a la familia $H_m(mN,p,n)$. Si llamamos M_r y M_s a las modas respectivas de H_r y H_s, sabemos, por el teorema 4.21, que, o bien ambas coinciden, o bien $M_r=M_s+1$.

Vamos a llamar, en lo que sigue, M a M_s.

CASO (i)

Si

$$M-1 \leq p(n-1) - \frac{q(n-1)}{N} < p(n-1) + \frac{p(n-1)}{N} \leq M ,$$

el razonamiento seguido en el teorema 4.15, para demostrar que

$$H_r \underset{a}{<} H_s , \quad \forall r<s ,$$

es válido en su totalidad.

CASO (ii)

Por el teorema 4.32, tenemos que

[4.55] $$h_r(M_r) > h_s(M_s),$$

o lo que es lo mismo,

$$\sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]} , \qquad \text{para } j=1.$$

Tomemos $M=M_s$ y consideremos que $M-1 \in I$. Veamos que, entonces, $pn<M$.

En efecto:

$$M-1 \in I \Rightarrow p(n-1) - \frac{q(n-1)}{N} < M-1 \Rightarrow pn+q - \frac{q(n-1)}{N} < M,$$

y basta tener en cuenta que $\frac{n-1}{N} <1$ para poder afirmar que $pn<M$.

Estamos, por tanto, en condiciones de aplicar el teorema 4.34, que nos asegura que

$$h_s(M-1) \geq h_s(M+1).$$

Como consecuencia,

$$\sum_{i=1}^{2} h_{s[i]} = h_s(M)+h_s(M-1) = H_s(M)-H_s(M-2)$$

De esta forma vemos que, para $j\geq 2$,

$$\sum_{i=1}^{j} h_{s[i]} = H_s(M+t)-H_s(M-K), \quad \text{con } t\geq 0;\ K\geq 2,$$

y, por tanto, estamos en condiciones de aplicar los resultados

del teorema 4.12 y de seguir, en su totalidad, el razonamiento empleado en el teorema 4.15, para demostrar que

$$H_r \underset{a}{<} H_s \ , \qquad \forall r<s \ ,$$

ya que la única dificultad se planteaba en M-1, donde, por pertenecer M-1 al intervalo de indeterminación, no podíamos haber comparado $H_r(M-1)$ con $H_s(M-1)$; y esta dificultad, que hubiese afectado al caso j=1, fue salvada gracias al teorema 4.32 que aseguraba el decrecimiento de la probabilidad modal en la sucesión H_m .

CASO (iii)

Si

$$M < p(n-1) + \frac{p(n-1)}{N} < M+1,$$

tendremos que M<pn<M+1, donde la segunda de las desigualdades viene asegurada por el hecho de que al ser M la moda, entonces p(n+1)<M+1. Por tanto,

$$pn = M+d, \qquad \text{con } 0<d<1.$$

Vamos a dividir el caso (iii) en dos subcasos:

(iiia) y (iiib) según que $p+d>q$ o no, respectivamente.

(iiia) Al igual que en el caso (ii), el teorema 4.32 nos asegura que

$$h_r(M_r) > h_s(M_s),$$

y, por tanto,

$$\sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]}, \quad \text{para } j=1.$$

Por otra parte, si $p+d>q$, el teorema 4.35 nos permite afirmar que

$$h_s(M+1) > h_s(M-1),$$

lo cual es tanto como decir que, para $j\geq 2$,

$$\sum_{i=1}^{j} h_{s[i]} = H_s(M+t)-H_s(M-K), \quad \text{con } t\geq 1 \text{ y } K\geq 1,$$

y, por tanto, al igual que en el caso (ii), estamos en condiciones de seguir en su totalidad el razonamiento utilizado en el teorema 4.15 para llegar a demostrar que

$$H_r \underset{a}{<} H_s, \quad \forall r<s .$$

(iiib) En este caso puede ocurrir que $h_s(M-1) > h_s(M+1)$, lo cual nos impedirá asegurar que

$$\sum_{i=1}^{2} h_{r[i]} > \sum_{1=1}^{2} h_{s[i]} .$$

de acuerdo con el criterio utilizado hasta ahora. En efecto,

$$\sum_{i=1}^{2} h_{s[i]} = h_s(M-1)+h_s(M) = H_s(M)-H_s(M-2),$$

y, si bien $H_s(M-2) > H_r(M-2)$ -por el teorema 4.12-, sin embargo, si $M \in I$, no podemos comparar $H_s(M)$ con $H_r(M)$, y si $M \notin I$, tendríamos que $H_s(M) > H_r(M)$, con lo cual, en cualquiera de los casos, no podriamos afirmar nada respecto de la comparacion de las dos mayores probabilidades puntuales de H_r y H_s y, por tanto, no podríamos aplicar el criterio de Hickey para comparar en aleatoriedad.

Este dificultad va a quedar limitada a la comparación de las dos mayores probabilidades puntuales ya que, por el teorema 4.36, tenemos que

$$h_s(M+1) > h_s(M-2),$$

de forma que, para $j \geq 3$, sí podemos afirmar que

$$\sum_{i=1}^{j} h_{r[i]} > \sum_{i=1}^{j} h_{s[i]} .$$

La dificultad ya señalada podría solventarse si se demuestra el decrecimiento de la suma de las dos mayores probabilidades puntuales en la sucesión H_m , lo que podría entenderse como una generalización del teorema 4.32, estudio éste que no vamos a abordar en esta memoria.

En conclusión, si excluimos el caso (iiib), podemos asegurar el crecimiento monótono de la aleatoriedad en la familia H_m, lo que resumimos con el siguiente teorema.

4.37 TEOREMA:

Sea la sucesión de distribuciones hipergeométricas $H_m(mN,p,n)$, con $0<p\leq 1/2$ y con soporte$(h_m)=\{0,1,\ldots,n\}$, $\forall m$, y sea $1\leq M_m\leq n-1$, $\forall m$, donde M_m es la moda de H_m .

Si H_r y H_s , con $r<s$, son dos distribuciones cualesquiera pertenecientes a la citada sucesión, en cualquiera de los dos siguientes casos,

[4.56] (i) $$p(n-1) + \frac{p(n-1)}{N} \leq M_s$$

[4.57] (ii) $$p(n-1) + \frac{p(n-1)}{N} > M_s \quad y \quad p(n+1)>M_s+q \ ,$$

se verifica que

$$H_r \underset{a}{<} H_s \ .$$

Si quisieramos reformular este teorema, de manera que las condiciones de validez del mismo no vengan dadas en función de la moda de uno de los términos que se están comparando, sino que sean comunes para cualquier par de términos comparados, necesitamos demostrar el siguiente lema:

4.38 LEMA:

$$M \geq 2 \quad \Rightarrow \quad p(n+1) > (M-1)+q$$

Demostración:

Si M es la moda, sabemos que

$$p(n+1) + \frac{(n+1)(1-2p)}{N+2} \geq M \ ,$$

luego

$$p(n+1)-M+1 \geq 1 - \frac{(n+1)(1-2p)}{N+2} \ .$$

Para demostrar el lema será suficiente probar que

$$1 - \frac{(n+1)(1-2p)}{N+2} > q \ ,$$

o, lo que es lo mismo,

$$p > \frac{(n+1)(1-2p)}{N+2} \ .$$

En efecto:

$$M \geq 2 \quad \Rightarrow \quad pn > 1 \quad \Rightarrow \quad 2pn+2p > 1-2p \quad \Leftrightarrow$$

$$\Leftrightarrow \quad n+2p > n+1-2pn-2p \quad \Leftrightarrow \quad n+2p > (n+1)(1-2p) \quad \Rightarrow$$

$$\Rightarrow \quad pN+2p > (n+1)(1-2p),$$

puesto que estamos trabajando con la hipótesis previa $pN \geq n$; y, finalmente, se obtiene

$$p > \frac{(n+1)(1-2p)}{N+2} \ ,$$

quedando probado el lema.

Llamemos ahora M a la moda del primer miembro de la sucesión, es decir, a la moda de H(N,p,n).

En primer lugar, si

$$p(n-1) + \frac{p(n-1)}{N} > M \quad y \quad p(n+1) > M+q \ ,$$

necesariamente, todos los términos de la sucesión localizan la moda en M, y la condición [4.57] se verifica, $\forall s$.

Por otra parte, si

$$p(n-1) + \frac{p(n-1)}{N} < M,$$

podría ocurrir que a partir de un cierto término de la sucesión, la moda pase a localizarse en M-1, pero, en este caso, por el lema 4.38, podemos asegurar que

$$p(n+1) > (M-1)+q ,$$

de forma que seguiremos estando en una de las dos situaciones donde la validez del teorema está garantizada.

Todo lo anterior nos permite reformular el teorema como sigue:

4.39 TEOREMA:

Sea la sucesión de distribuciones hipergeométricas $H_m(mN,p,n)$, con $0<p\leq 1/2$ y con soporte$(h_m)=\{0,1,\ldots,n\}$, $\forall m$, y sea $1\leq M_m\leq n-1$, $\forall m$.

En cualquiera de las dos situaciones siguientes,

[4.58] (i) $p(n-1) + \frac{p(n-1)}{N} \leq M$

[4.59] (ii) $p(n-1) + \frac{p(n-1)}{N} > M$ y $p(n+1) > M+q$,

donde M es la moda de H(N,p,n), se verifica que

$$H_r \underset{a}{<} H_s \ , \qquad \forall r<s \ .$$

Finalmente, como ya pusimos de manifiesto en el apartado dedicado a observaciones previas a la generalización del teorema 4.15, la restricción impuesta sobre el parámetro p obedece únicamente a razones de orden operativo.

En el caso de la sucesión $H^{*}_{m}(mN,p^{*},n)$, con $p^{*}>1/2$, bastaría considerar $p=1-p^{*}$ y ver si, para este valor de p, nos encontramos en alguna de las situaciones [4.58] o [4.59], que permiten asegurar que

$$H_r \underset{a}{<} H_s \ , \qquad \forall r<s \ .$$

en cuyo caso,

$$H^{*}_{r} \underset{a}{<} H^{*}_{s} \ , \qquad \forall r<s \ .$$

4.40 COROLARIO:

Puesto que la sucesión $H_m(mN,p,n)$ converge hacia la binomial B(n,p), bajo las condiciones del teorema anterior, se cumple que

$$H_r \underset{a}{<} B \ , \qquad \forall r \ .$$

BIBLIOGRAFIA

BARLOW, R.E. y PROSCHAN, F. (1975): " Statistical theory of reliability and life testing ". *Holt, Rinehart and Winston, Inc.*, Nueva York.

BARTOSZEWICZ, J. (1985): " Dispersive ordering and monotone failure rate distributions ". *Adv. Appl. Prob.*, 17, pp. 472-474.

BIRNBAUM, Z.W. (1948): " On random variables with comparable peakedness ". *Ann. Math. Statist.*, 19, pp. 76-81.

BOL'SHEV, L.N. (1964): " Distributions related to the hypergeometric distribution ". *Theory of probability and its applications*, 9, pp. 619-624.

BRUNK, H.D., HOLSTEIN, J.E. y WILLIAMS, F. (1968): " A comparison of binomial approximations to the hypergeometric distribution ". *American Statistician*, 22, pp.24-26.

BURR, I. (1952): " Formulae for approximating the hypergeometric and the binomial by the Poisson distribution ". (Abstract) *A.M.S.*, 23, p.145.

BURR, I. (1973): "Some approximate relations between terms of hypergeometric, binomial and Poisson distributions ". *Commu. Statis.*, 1, pp. 297-301.

CAPERAA, P. (1986): " Caracterisation des lois fortement unimodales ". *Adv. Appl. Prob.*, 18, pp. 277-278.

CHAPMAN, D.G. (1951): " Some properties of the hypergeometric distribution with applications to zoological sample censuces ". *Univ. of California Publications in Statistics*, 1, pp. 131-159.

DESHPANDE, J.V. y KOCHAR, S.C. (1983): " Dispersive ordering is the same as tail-ordering ". *Adv. Appl. Prob.* 15, pp. 686-687.

DOKSUM, K. (1969): " Star-shaped transformations and the power of rank test ". *Ann. Math. Statist.*, 40, pp. 1167-1176.

DROSTE, W. y WEFELMEYER (1985): " A note on strong unimodality and dispersivity ". *J. Appl. Prob.*, 22, pp. 235-239.

FELLER, W. (1945): " On the Normal approximation to the binomial distribution ". *Ann. Math. Stat.*, *vol.*16, pp. 319-329.

FELLER, W. (1968): " Probability theory and its applications". *Vol.* 1, *3ª Ed.*, *John Wiley and Sons.* Nueva York.

FERGUSON, T.S. (1967): "Mathematical Statistics". *Academic Press, Inc.*, Nueva York.

GIL, P. (1981): " Teoría matemática de la información ". *Ed. ICE.* Madrid.

HARDY, G.H., LITTLEWOOD, J.E. y POLYA, G. (1929): " Some inequalities satisfied by convex functions ". *Messenger of Math.*, 58, pp. 45-52.

HARDY, G.H., LITTLEWOOD, J.E. y POLYA, G. (1951): " Inequalities ". *Cambridge University Press.* Londres y Nueva York.

HICKEY, R.J. (1982): " A note on the measurement of randomness ".*J.Appl. Prob.*, 19, pp. 229-232.

HICKEY, R.J. (1983): " Majorisation, randomness and some discrete distributions ". *J. Appl. Prob.* , 20, pp. 897-902.

HICKEY, R.J. (1984): " Continuous majorisation and randomness" *J. Appl. Prob.*, 21, pp. 924-929.

HICKEY, R.J. (1986): " Concepts of dispersion in distributions: a comparative note ". *J. Appl. Prob.*, 23, pp. 914-921.

IBRAGIMOV, I.A. (1956): " On the composition of the unimodal distribution ". *Theory of probability and its applications, Vol.* 1, pp. 255-260.

JOHNSON, N.L. y KOTZ, Z. (1969): " Discrete distributions ". John Wiley and Sons. Nueva York.

KAMAT, A.R. (1966): " A generalization of Johnson's property of the mean deviation for a class of discrete distributions ". *Biometrika* , 53, pp. 285-287.

KEILSON, J. (1979): " Markov chain models - Rarity and exponentiality ". *Ed. Springer-Verlag.* Nueva York.

KEILSON, J. y GERBER, H. (1971): " Some results for discrete unimodality ". *J.A.S.A.*, 66, pp. 386-389.

KLAASSEN, C.A.J. (1985): "Strong unimodality". *Adv.Appl.Prob.*, 17, pp. 905-907.

LESSING, R. (1973) : "An alternative expression for the hypergeometric moment generating function ". *Amer. Statist.*, 27, p. 115.

LEWIS, T. y THOMPSON, J.W. (1981): " Dispersive distributions, and the connection between dispersivity and strong unimodality ". *J. Appl. Prob.*, 18, pp. 76-90.

LIEBERMAN, G.J. y OWEN, D.B. (1961): " Tables of the hypergeometric probability distribution ". Stanford: *Stanford University Press.*

LING, R.F. y PRATT, J.W. (1984): " The accuracy of Peizer approximations to the hypergeometric distribution, with comparisons to some other approximations ". *J.A.S.A.*, 79, pp. 49-60.

LYNCH, J. MIMMACK, G. y PROSCHAN, F. (1983): " Dispersive ordering results ". *Adv. Appl. Prob.*, 15, pp. 889-891.

MANN, H.B. y WHITNEY, D.R. (1947): " Stochastic ordering ". *Ann. Math. Statist.*, 18, pp. 50-60.

MARSHALL, A.W. y OLKIN, I. (1979): " Inequalities: theory of majorization and its applications ". *Academic Press.* Nueva York.

MOLENAAR, W. (1970): " Approximations to the Poisson, binomial, and hypergeometric distributional functions ". *Mathematical Centre Tracts*, 31. Mathematisch Centrem, Amsterdam.

MOOD, A.M., GRAYBILL, F.A. y BOES, D.C. (1974) : " Introduction to the theory of statistics ". *McGraw-Hill Book Co.* Nueva York.

NICHOLSON, W.L. (1956): " On the normal approximation to the hypergeometric distribution ". *Ann. Math. Statist.*, 27, pp. 471-483.

OLLERO, J. (1985): "Aleatoriedad de distribuciones binomiales con valor esperado entero ". *VII Congreso de GMEL, VOL II*, pp. 233-236.

ORD, J.K. (1967): " On a system of discrete distributions ". *Biometrika*, 54, pp. 649-656.

ORD, J.K. (1968): " Approximations to distribution functions which are hypergeometric series ". *Biometrika*, 55, pp. 243-248.

ORD, J. K. (1972): " Families of frequency distributions ". *Griffin.* Londres.

PATEL, J.K., KAPADIA, C.H. y OWEN, D.B. (1976): " Handbook of statistical distributions ". *Marcel Dekker Inc.* N. York.

PEARSON, K. (1924): " On the moments of the hypergeometrical series ". *Biometrika*, 16, pp. 157-162.

PEIZER, D.B. y PRATT, J.W. (1968): " A Normal approximation for Binomial, F, Beta, and other common, related tail probabilities, I ". *J.A.S.A.*, 63, pp. 1416-1456.

PRATT, J.W. (1964): " Risk aversion in the small and in the large ". *Econometrica*, 32, pp. 122-136.

PRATT, J.W. (1968): " A Normal approximation for Binomial, F, Beta, and other common, related tail probabilities, II ". *J.A.S.A.*, 63, pp. 1457-1483.

QUINN, B.G. y MacGILLIVRAY (1986): " Normal approximations to discrete unimodal distributions ". *J.Appl. Prob.*, 23, pp. 1013-1018.

RAMOS, H. (1985): " La distribución hipergeométrica ". Tesina. *Universidad de Granada. Facultad de Ciencias.*

RENYI, A. (1976): " Cálculo de probabilidades ". *Ed. Reverté.* Barcelona.

SAUNDERS, D.J. (1984): "Dispersive ordering of distributions". *Adv. Appl. Prob.*, 16, pp. 693-694.

SAUNDERS, I.W. (1978): " Locating bright spots in a point process ". *Adv. Appl. Prob.*, 10, pp. 587-612.

SAUNDERS, I.W. y MORAN, P.A.P. (1978): " On the quantiles of the gamma and F distributions ". *J. Appl. Prob.*, 15, pp 426-432.

SANDIFORD, P.J. (1960). " A new binomial approximation for use in sampling from finite populations ". *J.A.S.A.*, 55, pp. 718-722.

SATHE, Y.S. (1984): " A comment on star-ordering and tail-ordering ". *Adv. Appl. Prob.*, 16, p.692.

SHAKED, M. (1980): " On mixtures from exponential families ". *J.R. Statist. Soc., B* 42, pp. 192-198.

SHAKED, M. (1982): " Dispersive ordering of distributions ". *J. Appl. Prob.*, 19, pp. 310-320.

SHAKED, M. (1985): " Ordering distributions by dispersion ". *Encyclopedia of Statistical Sciences*, 5, pp. 485-490. *Ed. Kotz and Johnson.*

SKIBINSKY, M. (1970): " A characterization of hypergeometric distributions ". *J.A.S.A.*, 65, pp. 926-929.

STEPHAN, F.F. (1945): " The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate ". *Ann. of Math. Statist.*, 16, pp. 50-61.

UHLMANN, W. (1966): " Vergleich der hypergeometrischen mit der Binomial- Verteilung ". *Metrika*, 10, pp. 145-158.

WASAN, M.T. (1970): " Parametric estimation ". *McGraw-Hill Book. Co.* Nueva York.

WHITT, W. (1985): " Uniform conditional variability ordering of probability distributions ". *J. Appl. Prob.*, 22, pp. 619-633.

WISE, M.E. (1954): " A quickly convergent expansion for cumulative hypergeometric probabilities, direct and inverse ". *Biometrika*, 41, pp. 317-329.

BIBLIOTECA UNIVERSITARIA DE GRANADA

800028958