

Cheat sheet Práctica 4: Cálculo de Probabilidades

Juan Antonio Villatoro García – Grupo BioestadísticaR



**UNIVERSIDAD
DE GRANADA**

Todo el material para el conjunto de actividades de este curso ha sido elaborado y es propiedad intelectual del grupo BioestadísticaR formado por:

Antonio Martín Andrés

Juan de Dios Luna del Castillo,

Pedro Femia Marzo,

Miguel Ángel Montero Alonso,

Christian José Acal González,

Pedro María Carmona Sáez,

Juan Manuel Melchor Rodríguez,

José Luis Romero Béjar,

Manuela Expósito Ruíz,

Juan Antonio Villatoro García.

Todos los integrantes del grupo han participado en todas las actividades, en su elección, construcción, correcciones o en su edición final, no obstante, en cada una de ellas, aparecerán uno o más nombres correspondientes a las personas que han tenido la máxima responsabilidad de su elaboración junto al grupo de BioestadísticaR.

Todos los materiales están protegidos por la Licencia Creative Commons CC BY-NC-ND que permite "descargar las obras y compartirlas con otras personas, siempre que se reconozca su autoría, pero no se pueden cambiar de ninguna manera ni se pueden utilizar comercialmente".

Práctica 4: : Cálculo de probabilidades



Distribución Binomial

CONCEPTOS GENERALES

X: variable aleatoria con distribución binomial, $B(n,p)$

x: número de individuos que verifican una característica

p: probabilidad de que el individuo presente la característica (*prob* en las funciones).

n: número de repeticiones (*size* en las funciones).

FUNCIÓN	EXPLICACIÓN
dbinom(x, size, prob)	Calcula el valor de la función de probabilidad en el punto x de una variable con distribución $B(n,p)$.
pbinom(x, size, prob, lower.tail = TRUE)	Calcula la $P[X \leq x]$ de X que sigue una distribución $B(n,p)$. Si <i>lower.tail = FALSE</i> , calcula $P[X > x]$.
qbinom(a, size, prob, lower.tail = TRUE)	Calcula el percentil <i>a</i> -ésimo para una distribución Binomial $B(n,p)$, donde <i>a</i> toma un valor entre 0 y 1.
rbinom(r, size, prob)	Genera <i>r</i> valores aleatorios de una distribución binomial, $B(n,p)$

Ejemplos distribuciones

- $P[X=20]$ siendo X una distribución Binomial, $B(50,0.35)$

```
> dbinom(20,size=50,prob=0.35)
[1] 0.08750881
```

- $P[X \leq 5]$ siendo X una distribución de Poisson, $P(6)$

```
> ppois(5, lambda = 6)
[1] 0.4456796
```

- Valor de una normal $N(3,0.3)$ que deja por debajo al 85%

```
> a = 0.85
> qnorm(a, mean = 3, sd = 0.3)
[1] 3.31093
```

- Genera 15 valores aleatorios de una distribución normal $N(3,0.3)$

```
> rnorm(15, mean = 3, sd = 0.3)
[1] 2.807451 3.630845 3.323697 2.402412 3.159227
[6] 3.198425 3.134869 3.174063 3.056614 3.162409
[11] 3.273087 2.700590 3.214829 3.341973 3.248886
```

Distribución de Poisson

CONCEPTOS GENERALES

X: variable aleatoria con distribución de Poisson, $P(\lambda)$

λ : media de ocurrencias en el intervalo considerado (*lambda* en las funciones)

FUNCIÓN	EXPLICACIÓN
dpois(x, lambda)	Calcula el valor de la función de probabilidad en el punto x de una variable con distribución de Poisson, $P(\lambda)$.
ppois(x, lambda, lower.tail = TRUE)	Calcula la $P[X \leq x]$ de X que sigue una distribución $P(\lambda)$. Si <i>lower.tail = FALSE</i> , calcula $P[X > x]$.
qpois(a, lambda, lower.tail = TRUE)	Calcula el percentil <i>a</i> -ésimo para una distribución $P(\lambda)$, donde <i>a</i> toma un valor entre 0 y 1.
rpois(r, lambda)	Genera <i>r</i> valores aleatorios de una distribución de Poisson, $P(\lambda)$

Generación de números aleatorios

Función **sample(x, size, replace = FALSE)**

x = vector de la forma 1:n, siendo *n* el tamaño de la población
size = cantidad de números que queremos simular
replace = indica si la extracción se hace con reemplazamiento (TRUE) o no (FALSE)

Ejemplo: Obtener 20 números aleatorios sin reemplazamiento con números comprendidos entre el 1 y el 100.

```
> num.aleatorios=sample(1:100,20,replace=FALSE)
> num.aleatorios
[1] 90 29 9 91 67 23 43 31 50 14 95 10 30 25 47
[16] 1 15 34 36 39
```

Distribución Normal

CONCEPTOS GENERALES

X: variable aleatoria con distribución de Normal, $N(\mu,\sigma)$

μ : media de la distribución (*mean* en las funciones)

σ : desviación típica de la distribución (*sd* en las funciones)

FUNCIÓN	EXPLICACIÓN
dnorm(x, mean, sd)	Calcula el valor de la función de densidad en el punto x de una variable con distribución Normal $N(\mu,\sigma)$.
pnorm(x, mean, sd, lower.tail = TRUE)	Calcula la $P[X \leq x]$ de X que sigue una distribución Normal, $N(\mu,\sigma)$. Si <i>lower.tail = FALSE</i> , calcula $P[X > x]$.
qnorm(a, mean, sd, lower.tail = TRUE)	Calcula el percentil <i>a</i> -ésimo para una distribución Normal, $N(\mu,\sigma)$. (<i>a</i> toma un valor entre 0 y 1).
rnorm(r, mean, sd)	Genera <i>r</i> valores aleatorios de una distribución Normal, $N(\mu,\sigma)$

Extracción de una muestra aleatoria

Usamos primero la función **sample** para determinar las unidades de la muestras. Por ejemplo, si queremos seleccionar 20 individuos del data.frame *osteo*, generamos 20 números aleatorios entre 1 y 94 (números de individuos en el data.frame)

```
> num.aleatorios=sample(1:94,6,replace=FALSE)
> num.aleatorios
[1] 83 60 65 81 17 37
```

Posteriormente se seleccionan los individuos del data.frame

```
> osteo.muestra=osteo[num.aleatorios,]
> osteo.muestra
  num edad grupo_edad sexo peso talla imc tevo1
83 83 37 > 33 Mujer 49.5 150 22.00000 25
60 60 22 < 25 Hombre 61.7 160 24.10156 12
65 65 22 < 25 Mujer 49.0 158 19.62826 10
81 81 24 < 25 Hombre 59.7 170 20.65744 16
17 17 31 25 - 33 Hombre 62.0 168 21.96712 25
37 37 28 25 - 33 Hombre 47.0 153 20.07775 21
```