3 Distribuciones notables

3.1 Distribuciones discretas

3.1.1 La distribución de Bernoulli

Es el modelo discreto más sencillo: $0 \leq p \leq 1$. Hace referencia a situaciones en las que el resultado de un experimento sólo puede ser: se ha dado el suceso A o no se ha dado el suceso A. Por ejemplo, en el lanzamiento de una moneda sólo puede darse el suceso sale cara o su complementario no sale cara (sale cruz).

Por lo tanto, definimos la variable aleatoria X de la siguiente manera:

- $X = 1$ si se ha dado A.
- $X = 0$ si no se ha dado A, es decir, se ha dado el complementario A^c.

Si además, conocemos la probabilidad de que suceda A:

$$P[A] = p$$

y, por tanto,

$$P\left[A^c\right] = 1 - p$$

ya podemos definir la distribución de la variable aleatoria X. En estas condiciones diremos que X sigue una distribución de Bernoulli de parámetro p, que abreviaremos así $X \sim \text{Bernoulli}(p)$, y su función de masa de probabilidad se define así:

$$f(k) = P[X = k] = \left\{ \begin{array}{cc} p & \text{si } k = 1 (\text{ se ha dado } A) \\ 1 - p & \text{si } k = 0 \,(\text{ se ha dado } A^c) \end{array} \right\}$$

Gráficamente:

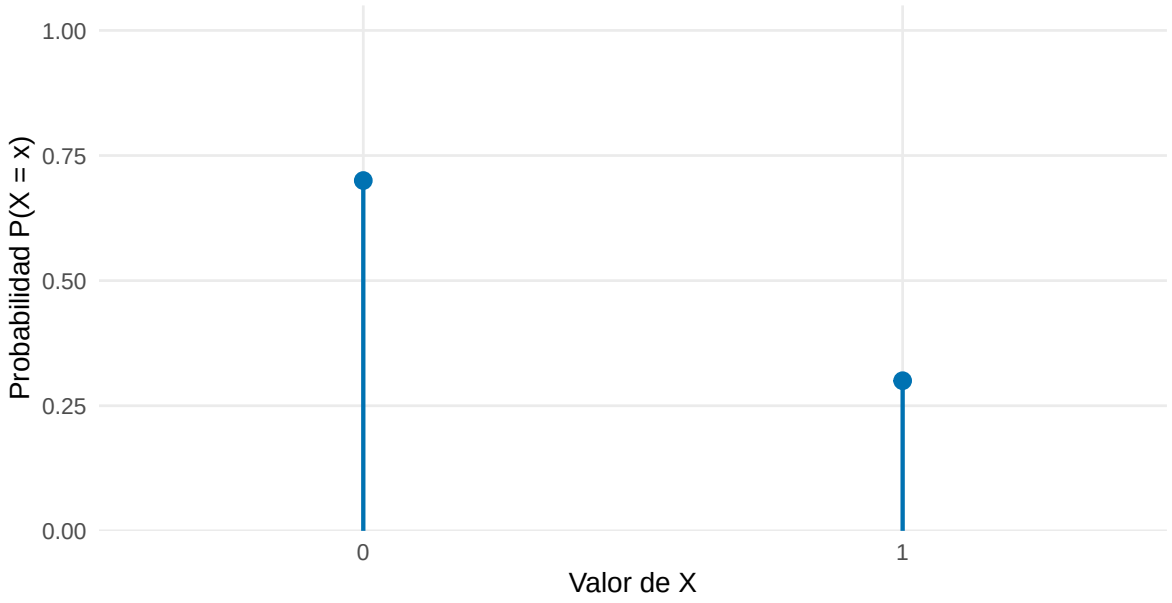


Figure 13: Masa de una Bernoulli con p = 0.3. Solo 0 y 1 tienen probabilidad positiva.

Mientras que la función de distribución será:

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0, \\ 1-p, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

Gráficamente:

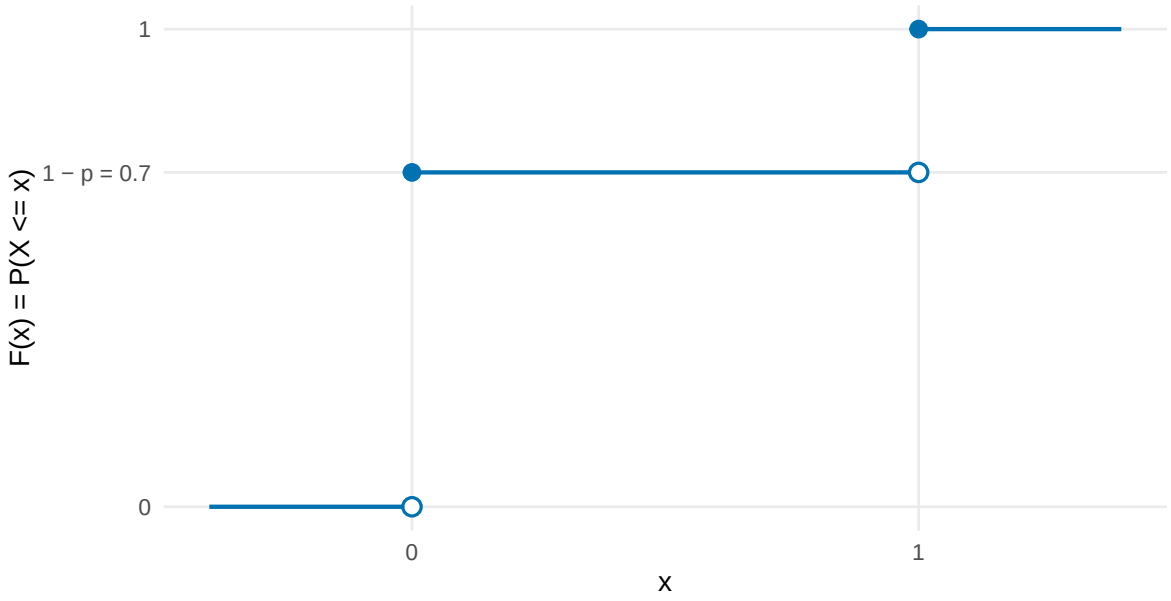


Figure 14: CDF de una Bernoulli con p = 0.3. Los puntos cerrados indican F(x); los abiertos, el límite por la izquierda.

Ejemplo 3.1.1.1 — Un resultado binario

Si X indica que una semilla germina y $p = 0.8$, entonces $P(X = 1) = 0.8$, $P(X = 0) = 0.2$ y $F(0) = 0.2$. La CDF acumula la masa situada a la izquierda, incluido el punto considerado.

3.1.1.1 Propiedades del modelo de Bernoulli

1) La esperanza vale $E(X) = p$.
2) La varianza vale $V(X) = p(1-p)$.

3.1.2 La distribución binomial

Al igual que el modelo de Bernoulli, hace referencia a experiencias con resultados dicotómicos (el resultado sólo puede ser A o $A^{\mathcal{C}}$). Sin embargo, en este modelo estamos interesados en la repetición de n veces una experiencia de este tipo, con independencia y una misma probabilidad p de éxito en todos los ensayos.

Tomemos el ejemplo del contaje del número de caras en el lanzamiento n veces de una moneda regular. Para concretar, vamos a suponer que disponemos de una moneda regular ($P[\text{cara}] = P[\text{cruz}] = 1/2$) que lanzamos cuatro veces. Es evidente que, en estas condiciones, la variable X: número de caras en cuatro lanzamientos independientes de una moneda regular es una variable aleatoria discreta que sólo puede tomar cinco posibles valores:

$$x = 0, 1, 2, 3, 4$$

Pasemos ahora a calcular la probabilidad de cada valor (en terminología estadística, vamos a calcular la función de masa de probabilidad de la variable X).

Es evidente que la $P[X = 0]$ es igual a la probabilidad de que salgan cuatro cruces seguidas:

$$P[X = 0] = P[\text{cruz}, \text{cruz}, \text{cruz}, \text{cruz}] = \text{P}[\text{cruz}]^4 = (1/2)^4 = 0,0625$$

ya que la moneda es regular y, por tanto, $P[\text{cara}] = P[\text{cruz}] = 1/2$. La $P[X = 3]$ corresponde al suceso de que salgan tres caras (c en adelante) y una cruz (+ en adelante). Sin embargo, en este caso tenemos hasta cuatro posibles maneras de obtener dicho resultado, según el orden en que aparezcan las tres caras y la cruz:

+ccc	c + cc	cc + c	ccc+

También debería resultar evidente que la probabilidad de cada uno de estos sucesos es la misma:

$$P[+\text{ccc}] = P[\text{c} + \text{cc}] = P[\text{cc} + \text{c}] = P[\text{ccc}+] = (1/2)^4 = (1/2)^4 = 0,0625$$

de manera que, finalmente, la probabilidad de que salgan tres caras y una cruz es la suma de las probabilidades de los 4 casos anteriores:

$$P[X = 3] = 4(1/2)^4 = 0,25$$

Y así podríamos ir calculando el resto de casos. Podemos ver que, en este ejemplo, todos los casos tienen la misma probabilidad $(0,0625)$ y que el número total de casos posibles es 16. En términos de combinatoria dicho número se obtendría como variaciones con repetición de dos valores (cara o cruz) tomados de cuatro en cuatro (el número de lanzamientos de la moneda):

$$VR_2{}^4 = 2^4 = 16$$

En la siguiente tabla se muestran los dieciséis posibles resultados (c: cara; +: cruz):

k	Resultados
0	`++++`
1	`c+++`, `+c++`, `++c+`, `+++c`
2	`cc++`, `c+c+`, `c++c`, `+cc+`, `+c+c`, `++cc`
3	`ccc+`, `cc+c`, `c+cc`, `+ccc`
4	`cccc`

Si hacemos uso de nuestros conocimientos de combinatoria, comprobamos que el número de casos para cada posible valor $k(k = 0, 1, 2, 3, 4)$ puede calcularse como permutaciones con repetición de cuatro elementos tomados de k y $4 - k$:

$$RP_4^{k,4-k} = \frac{4!}{k!(4-k)!} = \binom{4}{k}$$

y obtenemos finalmente el número combinatorio 4 sobre k. En efecto, para el caso $k = 3$, tendríamos:

$$\binom{4}{3} = \frac{4!}{3!1!} = 4$$

que son los cuatro posibles casos que nos dan tres caras y una cruz. Finalmente, recordando que todos los casos tienen la misma probabilidad, se construye la siguiente tabla:

k: número de caras	Número de casos	$P(X = k)$
0	1	0,0625
1	4	0,2500
2	6	0,3750
3	4	0,2500
4	1	0,0625
Total	16	1

3.1.2.1 Los parámetros de la distribución binomial La tabla anterior es, justamente, la función de masa de probabilidad de nuestra variable X.

Función de masa de probabilidad de X	
k	$P[X = k]$
0	0,0625
1	0,2500
2	0,3750
3	0,2500
4	0,0625
En otro caso	0

Como hemos visto, para obtener los resultados anteriores, hemos tenido que definir dos valores:

1. n : el número de lanzamientos (repeticiones de la experiencia aleatoria en condiciones independientes), en nuestro caso $n = 4$.
2. p : la probabilidad de que salga cara ($P[c]$), en nuestro caso $p = 1/2$.

Se dice, por tanto, que la distribución binomial depende de dos parámetros: n y p. En nuestro ejemplo, diremos que X sigue una distribución binomial de parámetros $n = 4$ y $p = 1/2$. De forma abreviada:

$$X \sim B(n = 4; p = 1/2)$$

En el ejemplo que hemos visto, suponíamos que la moneda era regular y, por tanto,

$$P[c] = P[+] = 1/2$$

Si tenemos una moneda trucada con las siguientes probabilidades:

$$P[c] = 2/3 \quad \text{y} \quad P[+] = 1/3$$

diremos que en este caso la variable X : número de caras en cuatro lanzamientos independientes de nuestra moneda trucada sigue una distribución binomial de parámetros:

$$X \sim B(n = 4; p = 2/3)$$

El problema se nos complica levemente ya que ahora no todos los posibles resultados tienen la misma probabilidad. Veamos dos ejemplos:

- La probabilidad de obtener cuatro caras es:

$$P[cccc] = (2/3)^4 = 0,1975$$

- Si R_i es el resultado del lanzamiento i, la probabilidad de que el primero sea cara (c) y los otros tres sean cruces ($+$) es:

$$P\left[(R_1, R_2, R_3, R_4) = (c, +, +, +)\right] = (2/3)(1/3)^3 = 0,02469$$

Sin embargo, sí se cumplirá que la probabilidad de que todos los casos que resulten en el mismo número de caras y cruces tendrán la misma probabilidad. Por ejemplo, para los cuatro casos en los que el número total de caras es 1 y el de cruces 3:

$$P[c+++] = P[+c++] = P[++c+] = P[+++c] = (2/3)(1/3)^3 = 0,02469$$

Y, por tanto, la probabilidad de obtener una sola cara en el lanzamiento de nuestra moneda trucada será:

$$P[X = 1] = 4 \times (2/3)(1/3)^3 = 0,09877$$

O, generalizando, si $P[A] = p$ y $P\left[A^c\right] = 1 - p$ tenemos que

$$P[X = k] = c(n, k)p^k(1 - \mathrm{p})^{n-k} \quad \text{si } k = 0, 1, \ldots, n$$

donde $c(n, k)$ representa el número de posibles resultados en los que obtenemos k caras y $n - k$ cruces en n lanzamientos. Tal como hemos visto, dicho número se puede calcular como permutaciones con repetición de n unidades tomadas de k y $n - k$.

Todo lo anterior nos lleva a formular el modelo binomial a través de la siguiente función de masa de probabilidad:

$$f(k) = P[X = k] = \left\{ \begin{array}{ll} \binom{n}{k} p^k (1-p)^{n-k} & \text{si} \quad k = 0, \ldots, n \\ 0 & \text{en caso contrario} \end{array} \right\}$$

con lo que la función de distribución se calcularía:

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0, \\ \displaystyle\sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i} p^i (1-p)^{n-i}, & 0 \leq x < n, \\ 1, & x \geq n. \end{cases}$$

En R

```
x <- 0:4
masa <- dbinom(x, size = 4, prob = 0.5)
data.frame(x, masa, acumulada = pbinom(x, 4, 0.5))
```

```
##   x   masa acumulada
## 1 0 0.0625    0.0625
## 2 1 0.2500    0.3125
## 3 2 0.3750    0.6875
## 4 3 0.2500    0.9375
## 5 4 0.0625    1.0000
```

```r
sum(masa)  # 1
```

```
## [1] 1
```

```r
1 - pbinom(2, 4, 0.5)  # P(X >= 3), no 1 - F(3)
```

```
## [1] 0.3125
```

3.1.2.2 Propiedades del modelo binomial

1. La esperanza vale $E(X) = np$.
2. La varianza es $V(X) = np(1-p)$.
3. Es una generalización del modelo de Bernoulli. En efecto, la Binomial con $n = 1$ (una sola realización) coincide con la distribución de Bernoulli.
4. La suma de dos variables aleatorias binomiales independientes con igual parámetro p también sigue una distribución binomial:

$$X_1 \sim B\left(n = n_1; p = p_0\right) \quad \text{y} \quad X_2 \sim B\left(n = n_2; p = p_0\right)$$

Si definimos $Z = X_1 + X_2$ entonces,

$$Z \sim B\left(n = n_1 + n_2; p = p_0\right)$$

3.1.3 La distribución de Poisson

Se trata de un modelo discreto, pero en el que el conjunto de valores con probabilidad no nula no es finito, sino numerable. Se dice que una variable aleatoria X sigue la distribución de Poisson si su función de masa de probabilidad viene dada por:

$$f(k) = P[X = k] = \left\{ \begin{array}{ll} e^{-\lambda}\frac{\lambda^k}{k!} & \text{si } k = 0, 1, 2, \ldots \\ 0 & \text{en caso contrario} \end{array} \right\}$$

Como vemos, este modelo se caracteriza por un sólo parámetro λ, que debe ser positivo. Esta distribución suele utilizarse para contajes del tipo número de individuos por unidad de tiempo, de espacio, etc.

Ejemplo 3.1.3.1 — Un recuento en un intervalo

Si el número de llamadas en un minuto se modela por una Poisson con $\lambda = 2$, $P(X = 0) = e^{-2} \approx 0.1353$. El parámetro indica el número medio en ese intervalo, no una probabilidad.

En R

```r
dpois(0, lambda = 2)  # P(X = 0)
```

```
## [1] 0.1353353
```

```r
ppois(2, lambda = 2, lower.tail = FALSE)  # P(X > 2)
```

```
## [1] 0.3233236
```

3.1.3.1 Propiedades del modelo de Poisson

1. Esperanza: $E(X) = \lambda$.
2. Varianza: $V(X) = \lambda$.

En esta distribución la esperanza y la varianza coinciden.

3. La suma de dos variables aleatorias independientes con distribución de Poisson resulta en una nueva variable aleatoria, también con distribución de Poisson, de parámetro igual a la suma de parámetros:

$$X_1 \sim P\left(\lambda = \lambda_1\right) \quad \text{y} \quad X_2 \sim P\left(\lambda = \lambda_2\right)$$

y definimos $Z = X_1 + X_2$, entonces,

$$Z \sim P\left(\lambda = \lambda_1 + \lambda_2\right)$$

Este resultado se extiende inmediatamente al caso de n variables aleatorias independientes con distribución de Poisson. En este caso, la variable suma de todas ellas sigue una distribución de Poisson de parámetro igual a la suma de los parámetros.

3.1.4 La distribución uniforme discreta

Tenemos esta distribución cuando el resultado de una experiencia aleatoria puede ser un conjunto finito de n posibles resultados, todos ellos igualmente probables.

Un ejemplo puede ser la variable X, puntuación en el lanzamiento de un dado regular. Esta variable toma seis valores posibles, todos con la misma probabilidad $p = 1/6$. La función de masa de probabilidad de esta variable será:

$$f(k) = P[X = k] = 1/6 \quad k = 1, 2, 3, 4, 5, 6$$

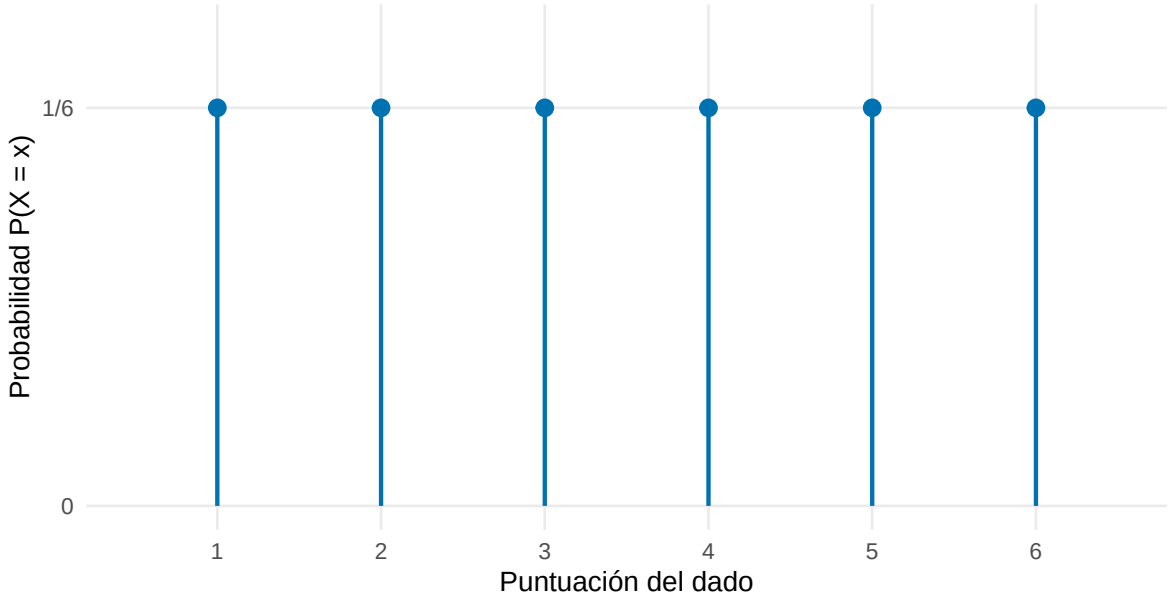


Figure 15: Masa de la puntuación de un dado regular: uniforme en los enteros de 1 a 6.

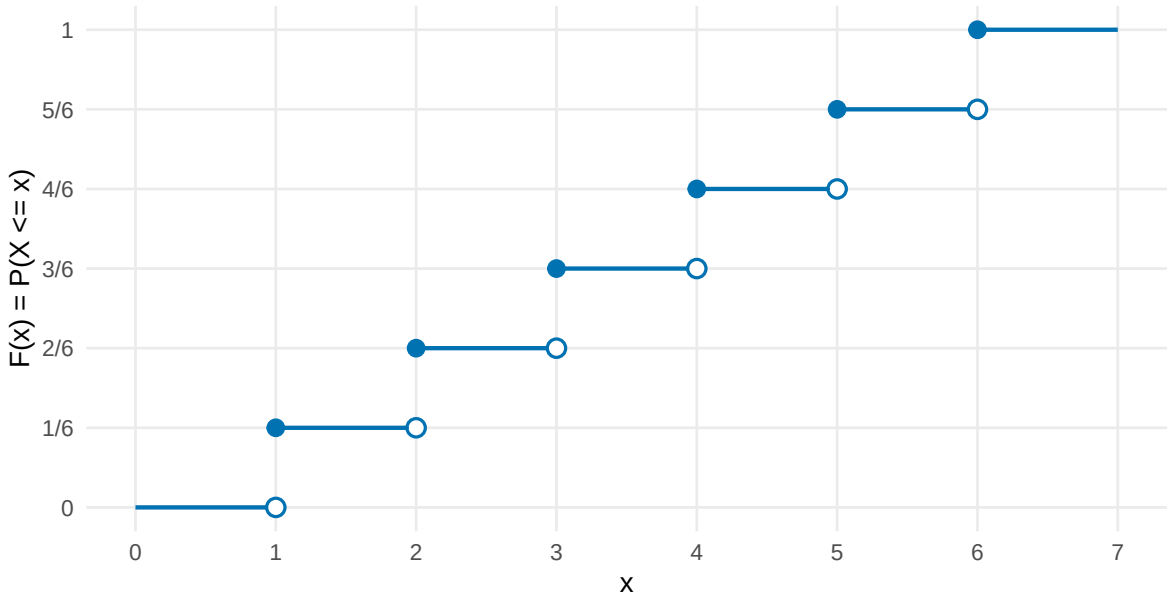


Figure 16: CDF de un dado regular. En cada entero se toma el valor del punto cerrado; no se interpolan los saltos.

En particular, si los valores son exactamente $1, 2, \ldots, n$, todos con igual probabilidad, su función de masa de probabilidad será:

$$f(k) = P[X = k] = 1/n \quad k = 1, 2, \ldots, n$$

3.1.4.1 Propiedades del modelo uniforme discreto

Sea n el número de valores equiprobables posibles:

3.1.4.2 Esperanza

$$E(X) = \frac{n+1}{2}$$

3.1.4.3 Varianza

$$V(X) = \frac{n^2 - 1}{12}$$

3.1.5 La distribución hipergeométrica

Este modelo presenta similitudes con el Binomial, pero sin la suposición de independencia de éste último. Veámoslo:

- Partimos de un conjunto formado por N individuos divididos en dos categorías mutuamente excluyentes: A y A^c; de manera que N_1 individuos pertenecen a la categoría A y N_2 individuos, a la categoría A^c. Por tanto, se cumple que

$$N = N_1 + N_2$$

- Si del conjunto anterior extraemos al azar n individuos sin reemplazamiento, con todas las muestras del mismo tamaño equiprobables ($n \leq N$), la variable X que representa el número k de individuos que pertenecen a la categoría A (de los n extraídos) tiene por función de masa de probabilidad:

$$f(k) = P[X = k] = \frac{\binom{N_1}{k}\binom{\mathrm{N}_2}{n-k}}{\binom{N}{n}}$$

si $\max\{0, \mathrm{n} - N_2\} \leq \mathrm{k} \leq \min\{N_1, n\}$

La dependencia se debe al hecho de que N es finito y las extracciones se efectúan sin reemplazamiento. Con reemplazamiento y extracciones independientes, la probabilidad de éxito permanece igual a N_1/N y el modelo es binomial, incluso si la población es finita.

3.1.5.1 Propiedades del modelo hipergeométrico

1. Esperanza: $E(X) = nN_1/N$.
2. Varianza: $\mathrm{Var}(X) = n\frac{N_1}{N}\frac{N_2}{N}\frac{N-n}{N-1}$, para $N > 1$

Ejemplo 3.1.5.1 — Sin reemplazamiento

En una urna con 10 bolas de cada uno de dos colores se extraen 12 sin reemplazamiento. El número X del primer color tiene media $12(10/20) = 6$; su varianza es $12(1/2)(1/2)(8/19) = 24/19$. El factor $(N-n)/(N-1)$ recoge la reducción de variabilidad respecto de la binomial.

En R

```
dhyper(6, m = 10, n = 10, k = 12)  # seis del primer color
```

```
## [1] 0.3500834
```

En `dhyper`, `m` y `n` son los tamaños de los dos grupos, y `k` el tamaño de la muestra.

3.1.6 La distribución geométrica o de Pascal

Definamos una experiencia aleatoria cuyo resultado sólo puede ser el suceso A o su complementario A^c, y que se repite secuencialmente hasta que aparece el suceso A por primera vez.

Definamos X como el número de fracasos (A^c) antes del primer éxito (A), en ensayos independientes con probabilidad constante p, donde $0 < p \leq 1$. Bajo estas condiciones, decimos que la variable X sigue una distribución geométrica o de Pascal de parámetro $p = P(A)$.

La función de masa de probabilidad puede deducirse fácilmente de la definición:

$$f(k) = P[X = k] = (1-p)^k p \quad k = 0, 1, 2, \ldots$$

Esta convención coincide con R. Otra convención cuenta el ensayo exitoso: $T = X + 1$ toma valores $1, 2, \ldots$ y tiene esperanza $1/p$.

Algunas puntualizaciones de la definición de X:

- Nótese que, en esta definición, se requieren independencia entre los ensayos y una probabilidad de éxito constante. La constancia de p por sí sola no implica independencia.
- Tal y como la hemos definido, X se refiere al número de lanzamientos hasta que se produce A, pero sin contabilizar el último caso en que se da A. Por dicha razón X puede tomar los valores $k = 0, 1, 2, \ldots$ con probabilidad no nula.

Un ejemplo de este modelo podría ser la experiencia consistente en lanzar sucesivamente un dado regular hasta que aparezca el número 6. Si definimos la variable aleatoria X como el número de lanzamientos sin un 6 antes del primero que lo muestre, queda claro que X sigue una distribución geométrica de parámetro $p = 1/6$.

En R

```r
dgeom(3, prob = 1/6)  # primer seis en el cuarto lanzamiento
```

```
## [1] 0.09645062
```

```r
pgeom(2, prob = 1/6)  # a lo sumo dos fracasos antes del seis
```

```
## [1] 0.4212963
```

3.1.6.1 Propiedades del modelo geométrico o de Pascal

1) Esperanza: $E(X) = (1-p)/p$
2) Varianza: $V(X) = (1-p)/p^2$

3.1.6.2 Preguntas

- ¿A qué suceso nos referimos cuando decimos $X = 0$? Respuesta.
  - Cuando decimos que $X = 0$ nos referimos al caso en que el 6 aparece en el primer lanzamiento. La probabilidad de que esto suceda, suponiendo un dado regular, es de 1/6 :

$$P[X = 0] = 1/6$$

- ¿Cuál es la probabilidad de que el primer 6 aparezca en el cuarto lanzamiento? Respuesta.
  - La probabilidad de que el primer 6 aparezca en el cuarto lanzamiento corresponde a:

$$P[X = 3] = (5/6)^3(1/6) = 0,0965$$

Fijémonos en que, si definimos A como el suceso sale un 6, la probabilidad anterior corresponde a la del suceso: $\{A^cA^cA^cA\}$ (en este orden).

3.1.7 La distribución binomial negativa

Puede definirse como una generalización del modelo geométrico o de Pascal. Así, dado un suceso A y su complementario A^c, cuando X representa el número de veces que se da A^c (ausencias, fallos, etc.) hasta que se produce r veces el suceso A, en una serie de repeticiones de la experiencia aleatoria en condiciones independientes, decimos que X sigue la distribución binomial negativa. Nótese que, cuando $r = 1$, tenemos exactamente el modelo geométrico.

En la interpretación de espera usamos $r = 1, 2, \ldots$ y $0 < p \leq 1$. Este modelo queda definido por dos parámetros p (la probabilidad de $A : p = P(A)$) y r (el número de veces que debe producirse A para que detengamos la experiencia).

La función de masa de probabilidad viene dada por:

$$f(k) = P[X = k] = \binom{k+r-1}{r-1} p^r q^k \quad k = 0, 1, 2, \ldots$$

donde q representa el complementario de $p : q = 1 - p$.

3.1.7.1 Propiedades del modelo binomial negativo

1. Esperanza: $E(X) = rq/p$
2. Varianza: $V(X) = rq/p^2$
3. Se cumplen las siguientes propiedades respecto a la función de masa de probabilidad:

$$f(0) = p^r \quad \text{y} \quad f(k+1) = \frac{(1-p)(k+r)}{k+1} f(k)$$

4. Este modelo se ajusta bien a contajes (números de individuos por unidad de superficie) cuando se produce una distribución contagiosa (los individuos tienden a agruparse).

5. La distribución binomial negativa puede definirse con mayor generalidad si tomamos r como un número real positivo cualquiera (no necesariamente entero). Pero, en dicho caso, se pierde el carácter intuitivo del modelo y se complican ligeramente los cálculos. Por dichas razones, se ha excluido dicha posibilidad en esta presentación.

3.1.8 Tabla resumen de las distribuciones discretas principales

En la tabla, $q = 1 - p$ y, para la hipergeométrica, $p = N_1/N$ y $N_2 = N - N_1$.

Distribución	Parámetros y soporte	Función de masa	Esperanza	Varianza
Bernoulli	$0 \leq p \leq 1$; $k = 0, 1$	$P(X = 0) = 1 - p$, $P(X = 1) = p$	p	$p(1-p)$
Binomial	$n \geq 1$ entero, $0 \leq p \leq 1$; $k = 0, \ldots, n$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$\lambda > 0$; $k = 0, 1, 2, \ldots$	$e^{-\lambda}\lambda^k/k!$	λ	λ
Uniforme discreta	$n \geq 1$ entero; $k = 1, \ldots, n$	$1/n$	$(n+1)/2$	$(n^2-1)/12$
Hipergeométrica	$N > 1$, $0 \leq n \leq N$; $\max(0, n - N_2) \leq k \leq \min(N_1, n)$, enteros	$\binom{N_1}{k}\binom{N_2}{n-k}/\binom{N}{n}$	np	$np(1-p)(N-n)/(N-1)$
Geométrica (fracasos)	$0 < p \leq 1$; $k = 0, 1, 2, \ldots$	$p(1-p)^k$	$(1-p)/p$	$(1-p)/p^2$
Binomial negativa (fracasos)	$r \geq 1$ entero, $0 < p \leq 1$; $k = 0, 1, 2, \ldots$	$\binom{k+r-1}{r-1} p^r (1-p)^k$	$r(1-p)/p$	$r(1-p)/p^2$

Las masas son cero fuera de los soportes indicados. En los casos límite $p = 0$ o $p = 1$ permitidos, se interpretan mediante las probabilidades del modelo, sin ambigüedad sobre 0^0.

La distribución multinomial describe un vector de recuentos y se reserva para la Unidad 4, «Distribuciones de probabilidad multidimensionales»; aquí nos limitamos a variables univariantes.

3.2 Distribuciones continuas

3.2.1 La distribución uniforme

La distribución uniforme es el modelo (absolutamente) continuo más simple. Corresponde al caso de una variable aleatoria que sólo puede tomar valores comprendidos entre dos extremos $a < b$, de manera que todos los intervalos de una misma longitud (dentro de (a, b)) tienen la misma probabilidad. También puede expresarse como el modelo probabilístico correspondiente a tomar un número al azar dentro de un intervalo (a, b).

De la anterior definición se desprende que la función de densidad debe tomar el mismo valor para todos los puntos dentro del intervalo (a, b) (y cero fuera del intervalo). Es decir,

$$f_X(x) = \left\{ \begin{array}{ll} \frac{1}{b-a} & \text{si } x \in (a,b) \\ 0 & \text{si } x \notin (a,b) \end{array} \right\}$$

Gráficamente:

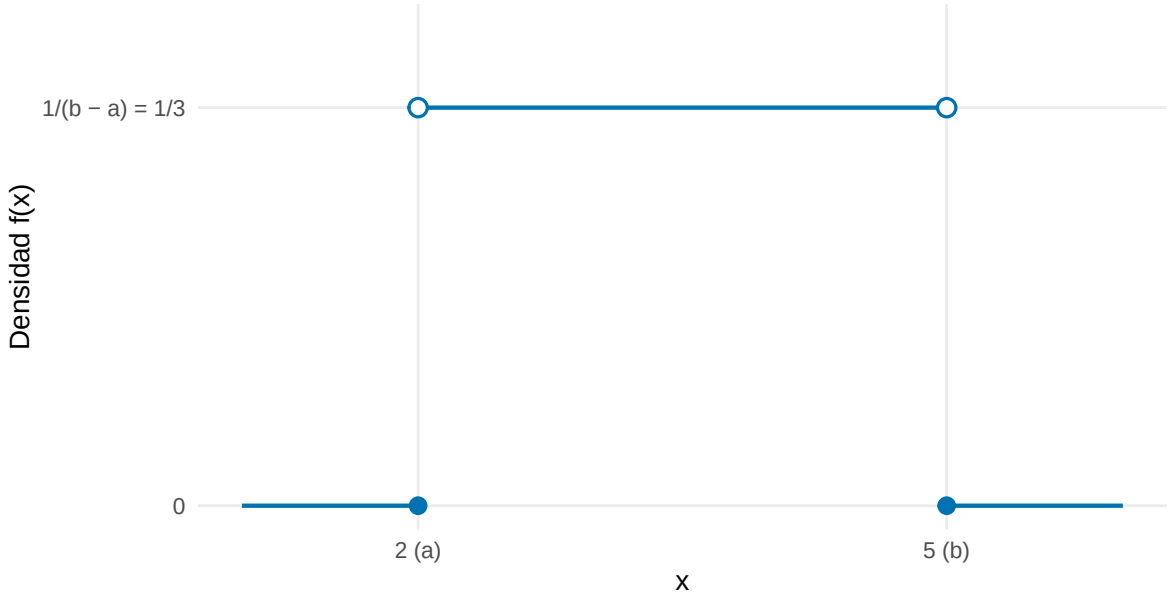


Figure 17: Densidad uniforme en (a, b), ilustrada con a = 2 y b = 5. Es cero fuera del intervalo; los valores aislados en los extremos no cambian las probabilidades.

La función de distribución se obtiene integrando la función de densidad y viene dada por:

$$F_X(x) = P(X \leq x) = \left\{ \begin{array}{ll} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } x \in (a,b) \\ 1 & \text{si } x \geq b \end{array} \right\}$$

Gráficamente:

Función de distribución del modelo uniforme

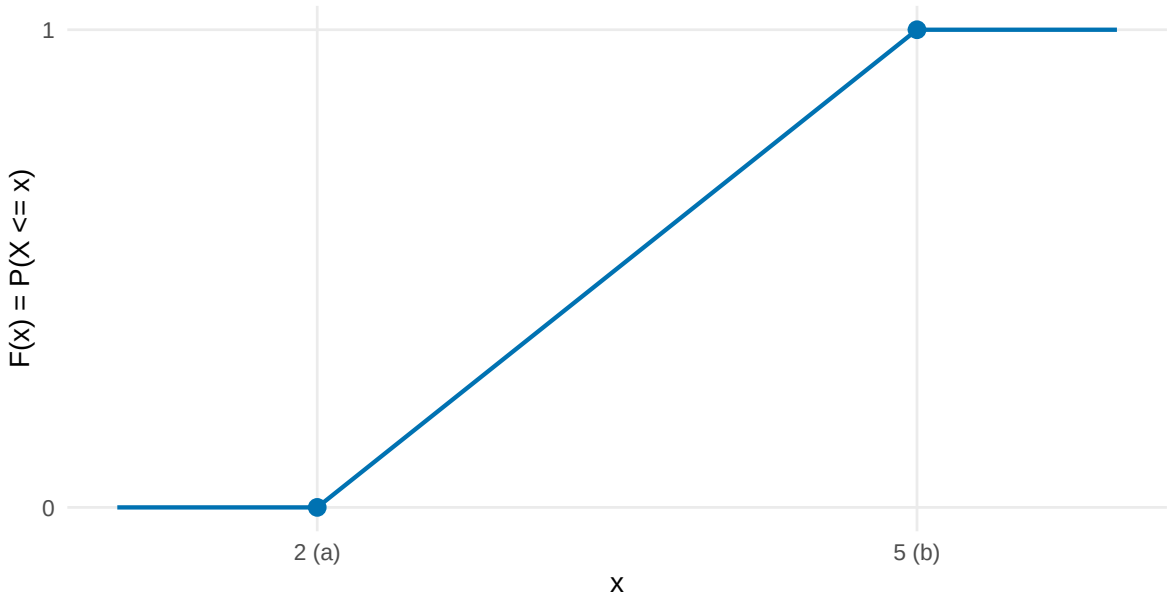


Figure 18: CDF uniforme con a = 2 y b = 5: continua, lineal en el intervalo, cero a la izquierda y uno a la derecha.

3.2.1.1 Propiedades del modelo uniforme

1. Su esperanza vale $(b + a)/2$
2. Su varianza es $(b - a)^2/12$

3.2.1.2 Una aplicación del modelo uniforme: el muestreo de Montecarlo En ciertos casos es útil simular el muestreo de una variable aleatoria con una distribución dada. El muestreo de Montecarlo es un procedimiento general para obtener muestras aleatorias de cualquier tipo de variable (discreta o continua) si su función de distribución es conocida o se puede calcular.

Supongamos que queremos generar una muestra procedente de una variable aleatoria X con función de distribución $F(x)$. El proceso comprende los siguientes pasos:

1. Obtener un valor aleatorio y entre cero y uno. Es decir, obtener una muestra de una distribución uniforme entre cero y uno. La mayoría de lenguajes de programación incorporan un generador de este tipo.
2. Buscar el cuantil correspondiente a y: $F^{-1}(y) = \inf\{x : F(x) \geq y\}$, para $0 < y < 1$.
3. Tomar $x = F^{-1}(y)$. Si la CDF es continua y estrictamente creciente en su recorrido, coincide con la inversa ordinaria. En una distribución discreta, la igualdad $y = F(x)$ puede no tener solución, pero el cuantil está definido.
4. Si queremos obtener una muestra con n individuos debemos repetir los pasos anteriores n veces.

3.2.1.3 Generación de una muestra procedente de una distribución binomial Supongamos que queremos simular el experimento de contar el número de caras obtenidas en 5 lanzamientos de una moneda trucada con probabilidad de cara igual a 0,75. Es decir, queremos obtener una muestra de una distribución binomial con $n = 5$ y $p = 0,75$.

Siguiendo los pasos anteriores deberemos obtener un número al azar entre 0 y 1 (un valor procedente de una distribución uniforme entre 0 y 1) y si este valor es menor o igual a 0,75 diremos que ha salido cara y, si es

superior a 0,75, cruz. Al sumar los cinco indicadores de cara obtenemos una realización binomial. También podemos generar directamente una binomial con su función cuantil:

En R

```
set.seed(303)
sum(runif(5) <= 0.75)  # agregar cinco Bernoulli
```

```
## [1] 4
```

```
qbinom(runif(1), size = 5, prob = 0.75)  # inversión directa
```

```
## [1] 5
```

Son dos realizaciones independientes del mismo modelo, no tienen por qué coincidir.

3.2.2 La distribución exponencial

Este modelo suele utilizarse para variables que describen el tiempo hasta que se produce un determinado suceso.

Su función de densidad es de la forma:

$$f(x) = \left\{ \begin{array}{ll} \frac{1}{\alpha}\exp\left(-\frac{x}{\alpha}\right) & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x \leq 0 \end{array} \right\}$$

Como vemos, este modelo depende de un único parámetro α que debe ser positivo: $\alpha > 0$. Aquí α es la escala y también la esperanza; la tasa es $1/\alpha$.

La función de distribución se obtiene integrando la de densidad y es de la forma:

$$F(x) = \left\{ \begin{array}{ll} 1 - \exp\left(-\frac{x}{\alpha}\right) & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x \leq 0 \end{array} \right\}$$

Ejemplo 3.2.2.1 — Espera y escala

Si una bombilla tiene vida exponencial de media $\alpha = 2$ años, $P(X > 3) = e^{-3/2} \approx 0.2231$. La tasa es $1/2$ por año; no debemos pasar la media 2 como argumento `rate` en R.

En R

```
pexp(3, rate = 1/2, lower.tail = FALSE)
```

```
## [1] 0.2231302
```

3.2.2.1 Propiedades del modelo exponencial

1. Su esperanza es α.
2. Su varianza es α^2.
3. La *carencia de memoria* significa $P(X > s + t \mid X > s) = P(X > t)$, para $s, t \geq 0$. Por ejemplo, si la variable X mide el tiempo de vida y sigue una distribución exponencial, significará que la probabilidad de que siga con vida dentro de 20 años es la misma para un individuo que a fecha de hoy tiene 25 años que para otro que tenga 60 años.
4. En un proceso de Poisson homogéneo de tasa λ (con incrementos independientes y estacionarios), el tiempo entre dos sucesos consecutivos sigue una distribución exponencial de parámetro $\alpha = 1/\lambda$.

3.2.3 La distribución normal

Se trata, sin duda, del modelo continuo más importante en estadística, tanto por su aplicación directa, veremos que muchas variables de interés general pueden describirse por dicho modelo, como por sus propiedades, que han permitido el desarrollo de numerosas técnicas de inferencia estadística. En realidad, el nombre de normal proviene del hecho de que durante un tiempo se creyó, por parte de médicos y biólogos, que todas las variables naturales de interés seguían este modelo.

Su función de densidad viene dada por la fórmula:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad \text{donde } -\infty < x < +\infty$$

que, como vemos, depende de dos parámetros μ (que puede ser cualquier valor real) y σ (que ha de ser positiva). Usaremos sistemáticamente $X \sim N(\mu, \sigma^2)$: el segundo parámetro es la varianza y σ es la desviación típica. En R, el argumento `sd` recibe σ, no σ^2. Por ejemplo, si nos referimos a una distribución normal con $\mu = 0$ y $\sigma = 1$ lo abreviaremos $N(0, 1)$.

Ejemplo 3.2.3.1 — Varianza y desviación típica

Si $X \sim N(100, 15^2)$, la desviación típica es 15. Para calcular $P(X \leq 115)$ tipificamos: $(115 - 100)/15 = 1$, de modo que la probabilidad es $\Phi(1) \approx 0.8413$.

En R

```
pnorm(115, mean = 100, sd = 15)
```

```
## [1] 0.8413447
```

```
qnorm(0.975, mean = 100, sd = 15)
```

```
## [1] 129.3995
```

La función de densidad del modelo normal tiene forma de campana, que habitualmente se denomina campana de Gauss. De hecho, este modelo se conoce también con el nombre de distribución gaussiana.

3.2.3.1 Propiedades del modelo normal

1. Su esperanza es μ.
2. Su varianza es σ^2y, por tanto, su desviación típica es σ.
3. Es simétrica respecto a su media μ, por la forma de su densidad.
4. Media, moda y mediana coinciden (μ).
5. Cualquier transformación lineal de una variable con distribución normal seguirá también el modelo normal. Si $X \sim N(\mu, \sigma^2)$ y definimos $Y = aX + b (\text{con } a \neq 0)$, entonces $Y \sim N(a\mu + b, a^2\sigma^2)$. Es decir, la esperanza de Y será $a\mu + b$ y su desviación típica, $|a|\sigma$.
6. Cualquier combinación lineal de variables normales independientes sigue también una distribución normal. Es decir, dadas n variables aleatorias independientes con distribución $X_i \sim N\left(\mu_i, \sigma_i^2\right)$ para $i = 1, 2, \ldots, n$ la combinación lineal: $Y = a_n X_n + a_{n-1} X_{n-1} + \ldots + a_1 X_1 + \mathrm{a}_0$ sigue también el modelo normal:

$$Y \sim N\left(a_0 + \sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

3.2.4 La función de distribución del modelo normal

La función de distribución del modelo normal se debería calcular, como en el resto de distribuciones continuas, integrando la función de densidad:

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} \mathrm{dt}$$

Pero nos encontramos con el problema de que no hay una primitiva expresable con funciones elementales; la CDF se calcula numéricamente. Sin embargo, si somos incapaces de calcular la función distribución no podremos efectuar ningún cálculo con este modelo. ¿Cómo solucionamos el problema?

Una primera solución podría consistir en aproximar la integral a través de técnicas de cálculo numérico. Sin embargo, dado que el conjunto de valores que pueden tomar los parámetros μ y σ son infinitos, deberíamos repetir el proceso para cada valor diferente de algún parámetro. Afortunadamente, podemos ahorrarnos el esfuerzo aprovechando la propiedad de que cualquier transformación lineal de una variable normal sigue también el modelo normal. Por tanto, replantearemos cualquier problema en términos de una normal concreta, que suele ser la N$(0, 1)$, de la siguiente manera:

Si $X \sim N(\mu, \sigma^2)$ y entonces definimos $Z = (\mathrm{X} - \mu)/\sigma$ se cumplirá que $Z \sim N(0, 1)$

y, por tanto:

$$F_X(x) = P[X \leq x] = P\left[\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right] = P\left[Z \leq \frac{x-\mu}{\sigma}\right] = F_Z\left(\frac{x-\mu}{\sigma}\right)$$

A la distribución $N(0, 1)$, es decir, la que tiene por media cero y por desviación típica uno, se le denomina normal reducida o tipificada. En cambio, al proceso de transformación del cálculo de la función de distribución de una normal cualquiera a través de la normal tipificada, se le denomina tipificación.

Debemos remarcar que el proceso de tipificación no resuelve el problema de la ausencia de una primitiva elemental. Sin embargo, sí es posible, mediante técnicas de cálculo numérico, obtener la integral numérica correspondiente y elaborar unas tablas que podemos consultar. Naturalmente, la tipificación permite que con una sola tabla, la de la $N(0, 1)$, tengamos suficiente.

Hoy en día, cada vez se utilizan menos tablas como la mencionada anteriormente, ya que los ordenadores, junto con los abundantes programas estadísticos existentes nos resuelven este problema. Sin embargo, la ausencia de una primitiva elemental persiste y, aunque nosotros no seamos conscientes, los programas informáticos realizan el proceso de tipificación para simplificar el problema.

3.2.5 La distribución gamma

Este modelo es una generalización del modelo exponencial ya que, en ocasiones, se utiliza para modelar variables que describen el tiempo hasta el p-ésimo suceso de un proceso de Poisson homogéneo cuando p es entero. La forma p puede ser cualquier real positivo en la definición matemática.

Su función de densidad es de la forma:

$$f(x) = \left\{ \begin{array}{lll} \frac{1}{\alpha^p \Gamma(p)} e^{-\frac{x}{\alpha}} x^{p-1} & \text{si} & x > 0 \\ 0 & \text{si} & x \leq 0 \end{array} \right\}$$

Como vemos, este modelo depende de dos parámetros positivos: escala α y forma p. En R se indican como `scale = alpha` y `shape = p`; la tasa equivalente es `rate = 1/alpha`.

La función $\Gamma(p)$ es la denominada función Gamma de Euler que representa la siguiente integral:

$$\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx$$

que verifica $\Gamma(p+1) = p\Gamma(p)$, con lo que, si p es un número entero positivo, $\Gamma(p+1) = p!$. Es decir, si evaluamos la función Gamma en números enteros positivos se obtiene el factorial del valor anterior en el que se ha realizado la evaluación. Por ejemplo:

$$\Gamma(5) = 4! = 4 \times 3 \times 2 = 24$$

3.2.5.1 Propiedades de la distribución gamma

1. Su esperanza es $p\alpha$.
2. Su varianza es $p\alpha^2$
3. La distribución gamma $(\alpha, p = 1)$ es una distribución exponencial de parámetro α. Es decir, el modelo exponencial es un caso particular de la Gamma con $p = 1$.
4. Dadas dos variables aleatorias **independientes** con distribución gamma y escala α común

$$X \sim G(\alpha, p_1) \text{ y } Y \sim G(\alpha, p_2)$$

se cumplirá que la suma también sigue una distribución gamma

$$X + Y \sim G(\alpha, p_1 + p_2)$$

Una consecuencia inmediata de esta propiedad es que, si tenemos k variables aleatorias con distribución exponencial de parámetro α (común) e independientes, la suma de todas ellas seguirá una distribución $G(\alpha, k)$.

Ejemplo 3.2.5.1 — Suma de esperas

La suma de cuatro vidas independientes exponenciales de media 2 años es $G(\alpha = 2, p = 4)$. Su esperanza es 8 años y su varianza 16 años cuadrados. La escala se conserva y las formas se suman.

En R

```
pgamma(8, shape = 4, scale = 2)  # probabilidad de no superar 8 años
```

```
## [1] 0.5665299
```

3.2.6 La distribución de Cauchy

Se trata de un modelo continuo cuya función de densidad es:

$$f(x) = \frac{1}{\pi(1+x^2)} \quad \text{para} \quad -\infty < x < \infty$$

Cuya integral nos proporciona la función de distribución:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)} dt = \frac{1}{\pi}[\arctan(t)]_{t=-\infty}^{t=x} = \frac{1}{2} + \frac{\arctan(x)}{\pi}$$

3.2.6.1 Propiedades de la distribución de Cauchy La Cauchy estándar es simétrica respecto de cero, que es su mediana y su moda, pero no tiene esperanza, y su varianza no está definida. La simetría no garantiza la existencia de la esperanza:

$$\int_0^b \frac{x}{\pi(1+x^2)}\, dx = \frac{\log(1+b^2)}{2\pi} \longrightarrow +\infty.$$

La parte negativa también tiene integral infinita en valor absoluto; no se pueden cancelar ambas divergencias. El valor principal simétrico cero no es la esperanza. Sus momentos enteros de orden positivo no son finitos.

3.2.7 La distribución de Weibull

Se trata de un modelo continuo asociado a variables del tipo tiempo de vida, tiempo hasta que un mecanismo falla, etc. La función de densidad de este modelo viene dada por:

$$f(x) = \left\{ \begin{array}{ll} \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{array} \right\}$$

que, como vemos, depende de dos parámetros: $\alpha > 0$ y $\beta > 0$, donde α es un parámetro de escala y β es un parámetro de forma (lo que proporciona una gran flexibilidad a este modelo).

La función de distribución se obtiene por la integración de la función de densidad y vale:

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-(x/\alpha)^{\beta}}, & x \geq 0. \end{cases}$$

Para calcular su CDF en R usamos `pweibull(x, shape = beta, scale = alpha)`. El valor de una densidad en el punto aislado cero puede elegirse sin alterar las probabilidades.

3.2.7.1 Propiedades de la distribución Weibull

1. Si tomamos $\beta = 1$ tenemos una distribución exponencial.
2. Su esperanza vale:

$$E(X) = \alpha\Gamma\left(\frac{1}{\beta} + 1\right)$$

3. Su varianza vale:

$$V(X) = \alpha^2 \left\{ \Gamma\left(\frac{2}{\beta} + 1\right) - \left[\Gamma\left(\frac{1}{\beta} + 1\right)\right]^2 \right\}$$

donde $\Gamma(x)$ representa la función Gamma de Euler definida anteriormente.

3.2.8 Tabla resumen de las principales distribuciones continuas

Las densidades son cero fuera del soporte indicado. En la normal, σ es la desviación típica; en exponencial y Gamma, α es la escala. Para abreviar la densidad Weibull escribimos $z = x/\alpha$; en la normal, ϕ es la densidad de $N(0, 1)$.

Distribución	Parámetros y soporte	Densidad
Uniforme	$a < b$; $a < x < b$	$1/(b-a)$
Exponencial	$\alpha > 0$; $x > 0$	$e^{-x/\alpha}/\alpha$
Normal	$\mu \in \mathbb{R}$, $\sigma > 0$; $x \in \mathbb{R}$	$\phi((x-\mu)/\sigma)/\sigma$
Gamma	$\alpha, p > 0$; $x > 0$	$\frac{x^{p-1}e^{-x/\alpha}}{\alpha^p \Gamma(p)}$
Cauchy estándar	$x \in \mathbb{R}$	$1/(\pi(1+x^2))$
Weibull	$\alpha, \beta > 0$; $x > 0$	$(\beta/\alpha)z^{\beta-1}e^{-z^{\beta}}$

Distribución	Esperanza	Varianza
Uniforme	$(a+b)/2$	$(b-a)^2/12$
Exponencial	α	α^2
Normal	μ	σ^2
Gamma	$p\alpha$	$p\alpha^2$
Cauchy estándar	No existe	No definida
Weibull	$\alpha\Gamma(1+1/\beta)$	$\alpha^2[\Gamma(1+2/\beta)$ $-\Gamma(1+1/\beta)^2]$

3.3 Distribuciones con R (y Python)

El lenguaje estadístico R es muy potente en cuanto al cálculo con distribuciones de probabilidad.

Dado que el trabajo con distribuciones de probabilidad usando R está muy estandarizado y explicado en múltiples fuentes no repetiremos aquí estas explicaciones.

Como complemento a los cálculos anteriores, os referimos a las siguientes fuentes de información que podéis utilizar para aprender como hacer los cálculos con R y también una aplicación que os permite visualizar casi cualquier distribución conocida.

R Tutorials

Explicación detallada y de nivel básico del manejo de las principales distribuciones con R

https://www.r-tutor.com/elementary-statistics/probability-distributions

The distribution Zoo

Permite visualizar de forma interactiva distintas distribuciones y proporciona información diversa sobre sus propiedades e incluso su aplicación.

https://ben18785.shinyapps.io/distribution-zoo/

Distribution explorer

Más completo que los anteriores. No se basa en R sino en python.

https://distribution-explorer.github.io/index.html

3.4 La familia exponencial de distribuciones

En el estudio de las propiedades de los estimadores, vemos que algunas distribuciones se comportan mejor que otras. Muchas veces, este buen comportamiento refleja una estructura común que proviene de pertenecer a una misma familia de distribuciones llamada familia exponencial.

Definición (caso uniparamétrico con soporte fijo): Sea f_θ una familia de probabilidades que depende de un parámetro unidimensional $\{f_\theta(x), \theta \in \Theta \subseteq \mathbb{R}\}$ tal que el soporte $S(\theta) = \{x \mid f_\theta(x) > 0\}$ no depende de θ. Si existen funciones de los parámetros $Q(\theta)$ y $C(\theta)$ y funciones del valor observado, $T(x)$ y $h(x)$, tales que la función de densidad puede escribirse como:

$$f_\theta(x) = C(\theta)h(x)\exp\{Q(\theta)\cdot T(x)\}$$

Para cada θ, $C(\theta) > 0$ incorpora la normalización y $h(x) \geq 0$. La expresión representa una densidad o una masa según el modelo. Diremos que $f_\theta(x)$ *pertenece a la familia exponencial de distribuciones.*

La familia exponencial no representa un nuevo tipo de distribuciones, sino la constatación de que muchas distribuciones comunes, que pueden reformularse para ajustarse a la expresión anterior, pertenecen a esta familia.

Veamos algunos ejemplos de que esto es efectivamente así.

3.4.1 Ejemplos de distribuciones de esta familia

3.4.1.1 Distribución de Poisson

Ejemplo 3.4.1.1.1 — Poisson en forma exponencial

La ley de Poisson pertenece a la familia exponencial uniparamétrica.

Efectivamente,

$$f_\lambda(x) = e^{-\lambda}\frac{\lambda^x}{x!} = \exp\{-\lambda + x\log\lambda - \log(x!)\}$$

y si hacemos

$$Q(\lambda) = \log(\lambda) \quad T(x) = x \quad C(\lambda) = e^{-\lambda} \quad h(x) = 1/x!, \quad x = 0, 1, 2, \ldots, \quad \lambda > 0$$

se hace evidente que f_λ pertenece a la familia exponencial.

3.4.1.2 Distribución normal uniparamétrica La ley normal depende de dos parámetros μ y σ. Fijado uno de ellos, nos queda una distribución que depende de un solo parámetro, y de aquí la denominación "normal uniparamétrica".

Si, con el subíndice "0", indicamos el parámetro fijado, tenemos:

$$f_\sigma = \left\{N\left(\mu_0, \sigma^2\right), \sigma > 0\right\} \text{ normal uniparamétrica, de parámetro } \sigma^2,$$
$$f_\mu = \left\{N\left(\mu, \sigma_0^2\right), \mu \in \mathbb{R}\right\} \text{ normal uniparamétrica, de parámetro } \mu.$$

Si queremos considerar ambos parámetros a la vez, debemos extender la definición al caso de parámetros k-dimensionales. En estos materiales no trataremos esta extensión.

3.4.1.2.1 Caso 1: Fijando la media μ_0

Ejemplo 3.4.1.2.1.1 — Media fija

Consideramos la distribución normal $N(\mu_0, \sigma^2)$, donde fijamos $\mu = \mu_0$ y σ^2 es el parámetro libre.

La función de densidad de probabilidad es

$$f_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu_0)^2}{2\sigma^2}\right\}$$

Vamos a reescribir esta función en forma de la familia exponencial. Primero, reorganizamos los términos de la densidad:

$$f_\sigma(x) = \frac{1}{\sqrt{2\pi}}\cdot\sigma^{-1}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu_0)^2\right\}$$

Ahora identificamos las funciones que se corresponden con la forma de la familia exponencial $f_\theta(x) = C(\theta)h(x)\exp\{Q(\theta)T(x)\}$:

- $Q(\sigma) = -\frac{1}{2\sigma^2}$
- $T(x) = (x-\mu_0)^2$
- $C(\sigma) = \frac{1}{\sqrt{2\pi}\sigma}$

- $h(x) = 1$

Esto confirma que la distribución normal, con μ_0 fijo, pertenece a la familia exponencial.

3.4.1.2.2 Caso 2: Fijando la varianza σ_0^2

Ejemplo 3.4.1.2.2.1 — Varianza fija

Ahora consideramos la distribución $N(\mu, \sigma_0^2)$, donde la varianza está fijada y el parámetro libre es μ.

La función de densidad es

$$f_\mu(x) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma_0^2}\right\}$$

Vamos a reescribir esta función de la misma manera:

$$f_\mu(x) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-\frac{1}{2\sigma_0^2}(x^2 - 2\mu x + \mu^2)\right\}$$

Identificamos las funciones correspondientes:

- $Q(\mu) = \frac{\mu}{\sigma_0^2}$
- $T(x) = x$
- $C(\mu) = \exp(-\mu^2/(2\sigma_0^2))/(\sqrt{2\pi}\sigma_0)$
- $h(x) = \exp(-x^2/(2\sigma_0^2))$

Esto prueba que la distribución normal con σ_0 fijo pertenece a la familia exponencial.

3.4.2 Distribución binomial

Ejemplo 3.4.2.1 — Binomial en forma exponencial

La distribución binomial es un ejemplo interesante, puesto que, a priori, no parece tener la estructura propia de la distribución exponencial, cosa que sí pasa con la distribución de Poisson o con las normales uniparamétricas que acabamos de ver.

Sin embargo, tras aplicar algunas transformaciones se puede ver como, también esta distribución pertenece a la familia exponencial.

Con n fijo y $0 < p < 1$, la función de masa de la binomial es

$$f(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$$

Reescribimos esta función en términos exponenciales:

$$f(x; n, p) = \binom{n}{x} \exp\{x \log(p) + (n-x)\log(1-p)\}$$

Agrupamos los términos dependientes de x:

$$f(x; n, p) = \binom{n}{x} \exp\left\{x \log\left(\frac{p}{1-p}\right) + n \log(1-p)\right\}$$

Identificamos las funciones correspondientes a la familia exponencial:

- $Q(p) = \log\left(\frac{p}{1-p}\right)$
- $T(x) = x$
- $C(p) = (1-p)^n$
- $h(x) = \binom{n}{x}$, para $x = 0, \ldots, n$

Por lo tanto, la distribución binomial pertenece a la familia exponencial.

3.4.3 Importancia y utilidad de la familia exponencial

Muchas de las distribuciones usadas para modelar gran cantidad de situaciones prácticas pertenecen a esta familia.

Esto significa que es posible estudiar sus propiedades en conjunto. Las propiedades comunes deben demostrarse para toda la familia bajo sus hipótesis; no basta con comprobarlas en uno de sus miembros.

A continuación, se describen algunas aplicaciones de trabajar con esta familia:

3.4.4 Los modelos lineales generalizados (GLMs)

Una de las aplicaciones más importantes de la familia exponencial es su uso en los **Modelos Lineales Generalizados (GLMs)**.

Estos modelos nos permiten extender la regresión lineal clásica a diferentes tipos de datos, como los resultados binarios (por ejemplo, éxito o fracaso), mediante la *regresión logística*, recuentos de eventos (como el número de llamadas recibidas en una hora) mediante la *regresión de Poisson*, y muchos otros.

Gracias a la estructura de la familia exponencial, podemos conectar la media de la variable que estamos modelando con las variables explicativas de forma flexible, lo que hace posible aplicar GLMs en una amplia variedad de situaciones.

3.4.5 Estimación en la familia exponencial

La estructura exponencial permite obtener resultados generales sobre estimación, siempre bajo las condiciones del teorema utilizado. Pertenecer a esta familia no garantiza por sí solo que cualquier estimador sea preciso, insesgado o adecuado para los datos.

La elección del modelo debe justificarse por el tipo de variable, el mecanismo de muestreo y los datos; no únicamente por su pertenencia a esta familia.