

Probabilidad (Honores): Parcial 2

Jonathan Andrés Niño Cortés

23 de abril de 2015

1. EJERCICIO 1

- A) 1) Sea μ_1 una distribución en \mathbb{R} . Calcule el parámetro $c \in \mathbb{R}$ que falta.

$$\mu_1(dx) = c \frac{1}{x} \mathbb{I}\{x > 2\} dx$$

- 2) Sea μ_2 una distribución en \mathbb{R} . Calcule el parametro $c \in \mathbb{R}$ que falta.

$$\mu_2(dx) = 5 \frac{1}{x^c} \mathbb{I}\{x > 1\} dx$$

- 3) Sea μ_3 una distribución en \mathbb{R} . Calcule el parametro $c \in \mathbb{R}$ que falta.

$$\mu_3(dx) = 5 \frac{\ln(x)}{x^c} \mathbb{I}\{x > 1\} dx$$

Demostración. Para que una distribución μ este bien definida, es necesario que la integral

$$\int_{-\infty}^{\infty} \mu(dx) = 1$$

1)

$$\begin{aligned}
 \int_{-\infty}^{\infty} \mu_1(dx) &= \int_{-\infty}^{\infty} c \frac{1}{x} \mathbb{I}\{x > 2\} dx \\
 &= \int_2^{\infty} c \frac{1}{x} dx \\
 &= \lim_{t \rightarrow \infty} \int_2^t c \frac{1}{x} dx \\
 &= \lim_{t \rightarrow \infty} c \ln(x) \Big|_2^t \\
 &= \lim_{t \rightarrow \infty} c \ln(t) - c \ln(2) \\
 &= c\infty
 \end{aligned}$$

Luego, no hay ningún valor de c tal que esta integral de 1.

2) Por lo que vimos anteriormente $c \leq 1$ no va a funcionar, pues la integral impropia no va a converger. Entonces asumimos que $c > 1$.

$$\begin{aligned}
 \int_{-\infty}^{\infty} \mu_2(dx) &= \int_{-\infty}^{\infty} 5 \frac{1}{x^c} \mathbb{I}\{x > 1\} dx \\
 &= \int_1^{\infty} 5 \frac{1}{x^c} dx \\
 &= \lim_{t \rightarrow \infty} \int_1^t 5 \frac{1}{x^c} dx \\
 &= \lim_{t \rightarrow \infty} 5 \frac{1}{(1-c)x^{c-1}} \Big|_1^t \\
 &= \lim_{t \rightarrow \infty} 5 \left(\frac{1}{c-1} - \frac{1}{(c-1)t^{c-1}} \right) \\
 &= \frac{5}{c-1}
 \end{aligned}$$

Por lo tanto, $c = 6$ es el parámetro que le falta a la distribución probabilística.

3) Por lo que vimos anteriormente $c \leq 1$ no va a funcionar, pues la integral impropia no va

a converger. Entonces asumimos que $c > 1$.

$$\begin{aligned}
\int_{-\infty}^{\infty} \mu_3(dx) &= \int_{-\infty}^{\infty} 5 \frac{\ln(x)}{x^c} \mathbb{I}\{x > 1\} dx \\
&= \int_1^{\infty} 5 \frac{\ln(x)}{x^c} dx \\
&= \lim_{t \rightarrow \infty} \int_1^t 5 \frac{\ln(x)}{x^c} dx \\
&= \lim_{t \rightarrow \infty} 5 \frac{\ln(x)}{(1-c)x^{c-1}} \Big|_1^t - \int_1^t 5 \frac{1}{(1-c)x^c} \quad (\text{Integración por partes}) \\
&= \lim_{t \rightarrow \infty} 5 \frac{\ln(x)}{(1-c)x^{c-1}} \Big|_1^t - 5 \frac{1}{(1-c)^2 x^{c-1}} \Big|_1^t \\
&= \lim_{t \rightarrow \infty} 5 \frac{\ln(t)}{(1-c)t^{c-1}} - 5 \frac{\ln(1)}{(1-c)1^{c-1}} - 5 \frac{1}{(1-c)^2 t^{c-1}} + 5 \frac{1}{(1-c)^2 1^{c-1}} \\
&= \lim_{t \rightarrow \infty} \frac{5 \ln(t)}{(1-c)t^{c-1}} - \frac{5}{(1-c)^2 t^{c-1}} + \frac{5}{(1-c)^2} \\
&= \frac{5}{(1-c)^2} + \lim_{t \rightarrow \infty} \frac{5 \ln(t)}{(1-c)t^{c-1}} \\
&= \frac{5}{(1-c)^2} + \lim_{t \rightarrow \infty} \frac{5 \frac{1}{t}}{-(1-c)^2 t^{c-2}} \quad (\text{L'Hopital}) \\
&= \frac{5}{(1-c)^2} + \lim_{t \rightarrow \infty} \frac{5}{-(1-c)^2 t^{c-1}} \\
&= \frac{5}{(1-c)^2}
\end{aligned}$$

Entonces el parametro para completar la distribución es $c = \sqrt{5} + 1$.

□

B) Considere una familia enumerable de variables aleatorias $(X_i)_{i \in \mathbb{N}}$ con $X_i \sim \mathcal{U}_{[0,1]}$.

4) Calcule y dibuje la densidad de $X_1 + X_2$.

Demostración. Para calcular la densidad de la suma de variables independientes debemos calcular la convolución de las densidades de las variables que se van a sumar.

La función de densidad de una variable X_i es $f_i = \mathbb{I}\{0 \leq x \leq 1\}$.

Entonces la función de densidad g de $X_1 + X_2$ es

$$\begin{aligned}
g(x) &= \int_{\mathbb{R}} f_1(t) f_2(x-t) dt \\
g(x) &= \int_{\mathbb{R}} \mathbb{I}\{0 \leq t \leq 1\} \mathbb{I}\{0 \leq x-t \leq 1\} dt
\end{aligned}$$

Entonces debemos considerar dos casos diferentes. Si $x < 1$ o si $x \geq 1$. En el primer caso tenemos que $\mathbb{I}\{0 \leq t \leq 1\}\mathbb{I}\{0 \leq x - t \leq 1\} = 1$ si y solo si $0 < t < x$. Luego

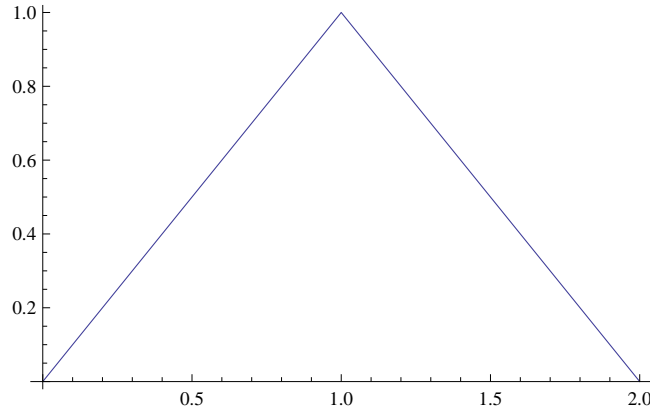
$$g(x) = \int_{\mathbb{R}} \mathbb{I}\{0 \leq t \leq 1\}\mathbb{I}\{0 \leq x - t \leq 1\}dt = \int_0^x dt = t \Big|_0^x = x$$

En el segundo caso entonces tenemos que $\mathbb{I}\{0 \leq x \leq 1\}(t)\mathbb{I}\{0 \leq x \leq 1\}(x - t) = 1$ si y solo si $x - 1 < t < 1$. Luego

$$g(x) = \int_{\mathbb{R}} \mathbb{I}\{0 \leq t \leq 1\}\mathbb{I}\{0 \leq x - t \leq 1\}dt = \int_x^1 -1^1 dt = t \Big|_{x-1}^1 = 2 - x$$

Entonces $g(x) = x\mathbb{I}\{0 \leq x \leq 1\} + (2 - x)\mathbb{I}\{1 \leq x \leq 2\}$

La gráfica de esta función es



□

5) Calcule y dibuje la densidad de $X_1 + X_2 + X_3$.

Demostración. Como ya calculamos la función de densidad de $X_1 + X_2$ la función de densidad de $X_1 + X_2 + X_3$ la podemos calcular como la convolución de g y f_3 . Defina esta convolución por h .

$$h(x) = \int_{\mathbb{R}} g(t)f_3(x - t)dt$$

$$h(x) = \int_{\mathbb{R}} [t\mathbb{I}\{0 \leq t \leq 1\} + (2 - t)\mathbb{I}\{1 \leq t \leq 2\}]\mathbb{I}\{0 \leq x - t \leq 1\}dt$$

$$h(x) = \int_{\mathbb{R}} t\mathbb{I}\{0 \leq t \leq 1\}\mathbb{I}\{0 \leq x - t \leq 1\}dt + \int_{\mathbb{R}} (2 - t)\mathbb{I}\{1 \leq t \leq 2\}\mathbb{I}\{0 \leq x - t \leq 1\}dt$$

De nuevo cada integral la debemos calcular por casos. Si $x < 1$ entonces como el caso anterior tenemos que la primera integral es diferente de 0 si $0 < t < x$, mientras que la

segunda integral siempre es 0. Luego

$$h(x) = \int_0^x t dx = \frac{t^2}{2} \Big|_0^x = \frac{x^2}{2}$$

Para el caso en que $1 \leq x < 2$ entonces la primera integral es diferente de 0 si $x-1 < t < 1$ y la segunda integral es diferente de 0 si $1 < t < x$.

Entonces

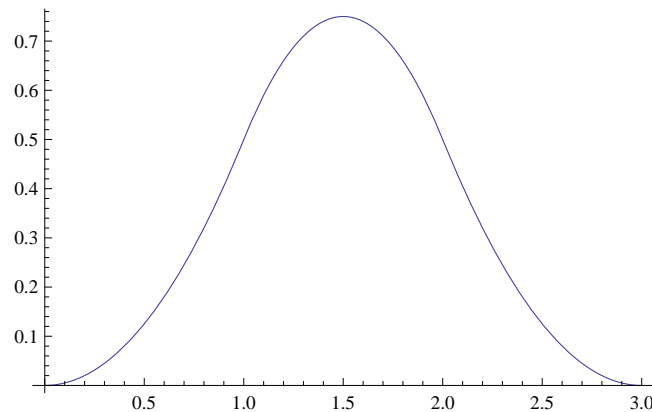
$$\begin{aligned} h(x) &= \int_{x-1}^1 t dt + \int_1^x (2-t) dt = \frac{t^2}{2} \Big|_{x-1}^1 + \left(2t - \frac{t^2}{2} \right) \Big|_1^x = \frac{1}{2} - \frac{(x-1)^2}{2} + 2x - \frac{x^2}{2} - \left(2 - \frac{1}{2} \right) \\ h(x) &= -x^2 + 3x - \frac{3}{2} \end{aligned}$$

Finalmente si $x \leq 2$ la primera integral siempre es 0 mientras que la segunda es diferente de 0 cuando $x-1 < t < 2$. Entonces

$$h(x) = \int_{x-1}^2 (2-t) dt = \left(2t - \frac{t^2}{2} \right) \Big|_{x-1}^2 = (4-2) - \left(2(x-1) - \frac{(x-1)^2}{2} \right) = \frac{x^2}{2} - 3x + \frac{9}{2}$$

Y entonces la función de distribución $\mathbb{I}\{0 \leq x \leq 1\} \frac{t^2}{2} + \mathbb{I}\{1 \leq x \leq 2\} (-x^2 + 3x - \frac{3}{2}) + \mathbb{I}\{2 \leq x \leq 3\} (\frac{x^2}{2} - 3x + \frac{9}{2})$.

La gráfica de esta función es



□

6) Calcule y dibuje la densidad de $X_1 + X_2 + X_3 + X_4$.

Demostración. De nuevo tenemos que calcular la densidad $j(x)$ como la convolución de h y f_4 .

$$\begin{aligned}
j(x) &= \int_{\mathbb{R}} h(x)f_4(x-t)dt \\
&= \int_{\mathbb{R}} \left(\mathbb{I}\{0 \leq t \leq 1\} \frac{t^2}{2} + \mathbb{I}\{1 \leq t \leq 2\} (-t^2 + 3t - \frac{3}{2}) + \mathbb{I}\{2 \leq t \leq 3\} (\frac{t^2}{2} - 3t + \frac{9}{2}) \right) \mathbb{I}\{0 \leq x-t \leq 1\} dt \\
&= \int_{\mathbb{R}} \mathbb{I}\{0 \leq t \leq 1\} \mathbb{I}\{0 \leq x-t \leq 1\} \frac{t^2}{2} dt + \int_{\mathbb{R}} \mathbb{I}\{1 \leq t \leq 2\} \mathbb{I}\{0 \leq x-t \leq 1\} (-t^2 + 3t - \frac{3}{2}) dt \\
&\quad + \int_{\mathbb{R}} \mathbb{I}\{2 \leq t \leq 3\} \mathbb{I}\{0 \leq x-t \leq 1\} (\frac{t^2}{2} - 3t + \frac{9}{2}) dt
\end{aligned}$$

Entonces de nuevo tenemos que considerar cuatro casos diferentes. Si $x < 1$ entonces la única integral que no es cero es la primera y justo no es cero cuando $0 < t < x$.

$$j(x) = \int_0^x \frac{t^2}{2} dt = \frac{t^3}{6} \Big|_0^x = \frac{x^3}{6}.$$

El segundo caso es cuando $1 \leq x \leq 2$. Entonces la primera integral es diferente de 0 cuando $x-1 < t < 1$ y la segunda integral es diferente de 0 cuando $1 < t < x$.

$$\begin{aligned}
j(x) &= \int_{x-1}^1 \frac{t^2}{2} dt + \int_1^x \left(-t^2 + 3t - \frac{3}{2} \right) dt \\
&= \frac{t^3}{6} \Big|_{x-1}^1 + \left(-\frac{t^3}{3} + \frac{3t^2}{2} - \frac{3t}{2} \right) \Big|_1^x \\
&= \frac{1}{6} - \frac{(x-1)^3}{6} + \left(-\frac{x^3}{3} + \frac{3x^2}{2} - \frac{3x}{2} \right) - \left(-\frac{1}{3} + \frac{3}{2} - \frac{3}{2} \right) \\
&= \frac{1}{6} - \frac{x^3}{6} + \frac{x^2}{2} - \frac{x}{2} + \frac{1}{6} - \frac{x^3}{3} + \frac{3x^2}{2} - \frac{3x}{2} + \frac{1}{3} \\
&= -\frac{x^3}{2} + 2x^2 - 2x + \frac{2}{3}
\end{aligned}$$

El tercer caso es cuando $2 \leq x \leq 3$ y en este caso la segunda integral es no cero cuando $x-1 < t < 2$ y la tercera integral es no cero cuando $2 < t < x$.

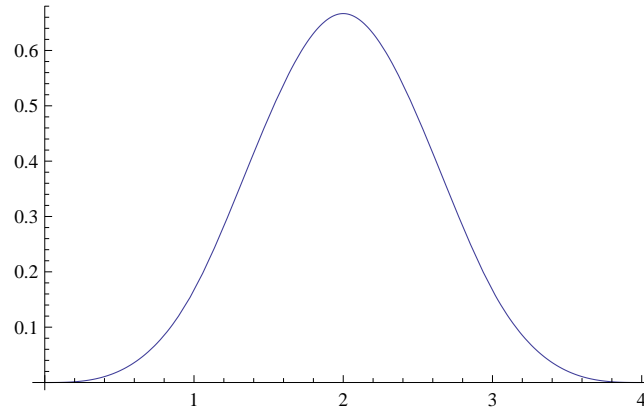
$$\begin{aligned}
j(x) &= \int_{x-1}^2 \left(-t^2 + 3t - \frac{3}{2} \right) dt + \int_2^x \left(\frac{t^2}{2} - 3t + \frac{9}{2} \right) \\
&= \left(-\frac{t^3}{3} + \frac{3t^2}{2} - \frac{3t}{2} \right) \Big|_{x-1}^2 + \left(\frac{t^3}{6} - \frac{3t^2}{2} + \frac{9t}{2} \right) \Big|_2^x \\
&= \left(-\frac{8}{3} + 6 - 3 \right) - \left(-\frac{(x-1)^3}{3} + \frac{3(x-1)^2}{2} - \frac{3(x-1)}{2} \right) + \left(\frac{x^3}{6} - \frac{3x^2}{2} + \frac{9x}{2} \right) - \left(\frac{8}{6} - 6 + 9 \right) \\
&= -\frac{8}{3} + 3 + \frac{x^3}{3} - x^2 + x - \frac{1}{3} - \frac{3x^2}{2} + 3x - \frac{3}{2} + \frac{3x}{2} - \frac{3}{2} + \frac{x^3}{6} - \frac{3x^2}{2} + \frac{9x}{2} - \frac{4}{3} - 3 \\
&= \frac{x^3}{2} - 4x^2 + 10x - \frac{22}{3}
\end{aligned}$$

Finalmente tenemos el caso cuando $x > 3$. En este caso la única integral que sobrevive es la ultima y no es cero cuando $x - 1$ a 3.

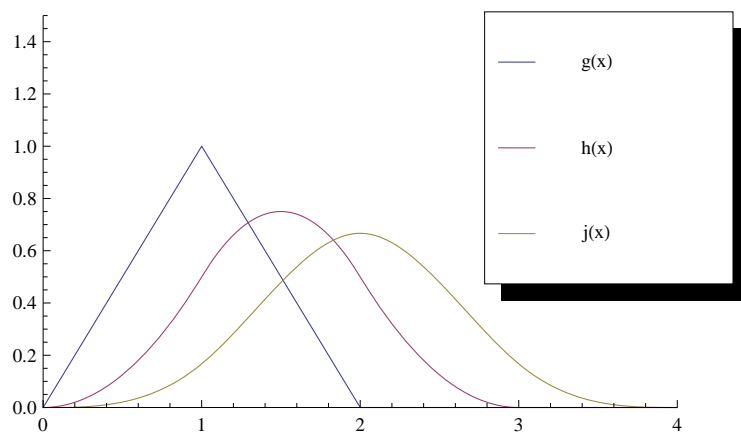
$$\begin{aligned}
j(x) &= \int_{x-1}^3 \left(\frac{t^2}{2} - 3t + \frac{9}{2} \right) dt \\
&= \left(\frac{t^3}{6} - \frac{3t^2}{2} + \frac{9t}{2} \right) \Big|_{x-1}^3 \\
&= \left(\frac{9}{2} - \frac{27}{2} + \frac{27}{2} \right) - \left(\frac{(x-1)^3}{6} - \frac{3(x-1)^2}{2} + \frac{9(x-1)}{2} \right) \\
&= \frac{9}{2} - \frac{x^3}{6} + \frac{x^2}{2} - \frac{x}{2} + \frac{1}{6} + \frac{3x^2}{2} - 3x + \frac{3}{2} - \frac{9x}{2} + \frac{9}{2} \\
&= -\frac{x^3}{6} + 2x^2 - 8x + \frac{32}{3}
\end{aligned}$$

Finalmente obtenemos que $j(x) = \mathbb{I}\{0 \leq x \leq 1\} \frac{x^3}{6} + \mathbb{I}\{1 \leq x \leq 2\} \left(-\frac{x^3}{2} + 2x^2 - 2x + \frac{2}{3} \right) + \mathbb{I}\{2 \leq x \leq 3\} \left(\frac{x^3}{2} - 4x^2 + 10x - \frac{22}{3} \right) + \mathbb{I}\{3 \leq x \leq 4\} \left(-\frac{x^3}{6} + 2x^2 - 8x + \frac{32}{3} \right)$

La gráfica de esta función es



Finalmente mostramos una gráfica que incluye las tres funciones juntas



□

- 7) Qué sospecha sobre la densidad de $\sum_{i=1}^n X_i$? Coleccione toda la información que usted sospecha.

Demostración. Denote por g_n la función de masa de $\sum_{i=1}^n X_i$. Una de las cosas que podemos demostrar es que las funciones g_n son continuas y alcanzan su máximo en $\frac{n}{2}$. También podemos observar que son simétricas con respecto a la recta $x = \frac{n}{2}$, es decir que $g_n(x) = g_n(n-x)$. Una de las cosas que sospecho es que los valores que toma esta función en punto enteros debe aproximarse a los valores que tomaría la distribución binomial $\mathcal{B}_{n, \frac{1}{2}}$ □

- C) Demuestre por cálculo directo para variables aleatorias $\perp(X, Y)$ las siguientes implicaciones:

- 8) $X \sim \mathcal{P}_\lambda, Y \sim \mathcal{P}_{\lambda'}$ entonces $X + Y \sim \mathcal{P}_{\lambda+\lambda'}$.

Demostración. Para nuestras variables discretas independientes X y Y la función de masa f de $X + Y$ esta dada por

$$f(x) = \sum_{t=0}^x f_1(t) f_2(x-t)$$

donde f_1 y f_2 son las funciones de masa de X y Y respectivamente.

Ahora las funciones de masa de la distribución de Poisson son $f_1(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ y $f_2(x) = \frac{e^{-\lambda'} (\lambda')^x}{x!}$.

Luego la función de masa de $X + Y$ esta dada por

$$\begin{aligned}
f(x) &= \sum_{t=0}^x f_1(t) f_2(x-t) \\
&= \sum_{t=0}^x \frac{e^{-\lambda} \lambda^t}{t!} \frac{e^{-\lambda'} (\lambda')^{x-t}}{(x-t)!} \\
&= e^{-(\lambda+\lambda')} \sum_{t=0}^x \frac{\lambda^t (\lambda')^{x-t}}{t! (x-t)!}
\end{aligned}$$

Recordemos que el coeficiente binomial esta definido como $\binom{x}{k} = \frac{x!}{k!(x-k)!}$. Luego tenemos que $\frac{1}{k!(x-k)!} = \frac{1}{x!} \binom{x}{k}$.

Por lo tanto nuestra expresi3n resultante es

$$\begin{aligned}
f(x) &= e^{-(\lambda+\lambda')} \sum_{t=0}^x \frac{\lambda^t (\lambda')^{x-t}}{t! (x-t)!} \\
&= e^{-(\lambda+\lambda')} \sum_{t=0}^x \frac{1}{x!} \binom{x}{t} \lambda^t (\lambda')^{x-t} \\
&= \frac{e^{-(\lambda+\lambda')}}{x!} \sum_{t=0}^x \binom{x}{t} \lambda^t (\lambda')^{x-t}
\end{aligned}$$

Pero vemos que la sumatoria que queda es precisamente la f3rmula binomial de Newton para el polinomio $(\lambda + \lambda')^x$, luego la f3rmula final es

$$f(x) = \frac{e^{-(\lambda+\lambda')} (\lambda + \lambda')^x}{x!}$$

Que coincide precisamente con la funci3n de masa asociada a $\mathcal{P}_{\lambda+\lambda'}$. □

9) $X \sim \mathcal{B}_{n_1, p}$, $Y \sim \mathcal{B}_{n_2, p}$ entonces $X + Y \sim \mathcal{B}_{n_1+n_2, p}$.

Demostraci3n. En este caso las funciones de masa asociadas a X y Y son respectivamente $f_1(x) = \binom{n_1}{x} p^x (1-p)^{n_1-x}$ y $f_2(x) = \binom{n_2}{x} p^x (1-p)^{n_2-x}$. Entonces la funci3n de masa de $X + Y$ esta dada por

$$\begin{aligned}
f(x) &= \sum_{t=0}^x f_1(t) f_2(x-t) \\
&= \sum_{t=0}^x \binom{n_1}{t} p^t (1-p)^{n_1-t} \binom{n_2}{x-t} p^{x-t} (1-p)^{n_2-(x-t)} \\
&= \sum_{t=0}^x \binom{n_1}{t} \binom{n_2}{x-t} p^{t+x-t} (1-p)^{n_1-t+n_2-(x-t)} \\
&= \sum_{t=0}^x \binom{n_1}{t} \binom{n_2}{x-t} p^x (1-p)^{n_1+n_2-x} \\
&= p^x (1-p)^{n_1+n_2-x} \sum_{t=0}^x \binom{n_1}{t} \binom{n_2}{x-t}
\end{aligned}$$

Ahora concentrandonos en los coeficientes binomiales una de sus definiciones es que $\binom{n}{k}$ es el coeficiente que acompaña al k -ésimo término del polinomio $(x+1)^n$.

Entonces si tomamos $(x+1)^{n_1+n_2} = (x+1)^{n_1} (x+1)^{n_2}$ y por definición de la multiplicación de polinomios, el x -ésimo término estaría dado por $\sum_{t=0}^x a_t b_{x-t}$, es decir, $\binom{n_1+n_2}{x} =$

$$\sum_{t=0}^x \binom{n_1}{t} \binom{n_2}{x-t}.$$

Por lo tanto, llegamos a la expresió final

$$f(x) = p^x (1-p)^{n_1+n_2-x} \sum_{t=0}^x \binom{n_1}{t} \binom{n_2}{x-t} = \binom{n_1+n_2}{x} p^x (1-p)^{n_1+n_2-x}$$

que es la función de masa asociada a $\mathcal{B}_{n_1+n_2, p}$. □

- 10) $X \sim N_{m_1, \sigma_1^2}$, $Y \sim N_{m_2, \sigma_2^2}$ entonces $X + Y \sim N_{m_1+m_2, \sigma_1^2+\sigma_2^2}$.

Demostración. En este caso debemos utilizar las funciones de densidad de X y Y que

$$\text{son respectivamente } f_1(x) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2\sigma_1^2}} \text{ y } f_2(x) = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x-m_2)^2}{2\sigma_2^2}}.$$

Entonces la función de densidad asociada $X + Y$ esta dada por

$$\begin{aligned}
f(x) &= \int_{\mathbb{R}} \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(t-m_1)^2}{2\sigma_1^2}} \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x-t-m_2)^2}{2\sigma_2^2}} dt \\
&= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{(t-m_1)^2}{2\sigma_1^2} - \frac{(x-t-m_2)^2}{2\sigma_2^2}} dt \\
&= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{(t-m_1)^2}{2\sigma_1^2} - \frac{(x-t-m_2)^2}{2\sigma_2^2}} dt
\end{aligned}$$

Para simplificar los cálculos utilizamos la sustitución $z = t - m_1$.

$$f(x) = \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{z^2}{2\sigma_1^2} - \frac{(x-m_1-m_2-z)^2}{2\sigma_2^2}} dz$$

Y tome $a = x - m_1 - m_2$, pues todos estos son constantes en la integral.

$$\begin{aligned} f(x) &= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{z^2}{2\sigma_1^2} - \frac{(a-z)^2}{2\sigma_2^2}} dz \\ &= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{\sigma_2^2 z^2 + \sigma_1^2 (a-z)^2}{2\sigma_1^2 \sigma_2^2}} dz \\ &= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{\sigma_2^2 z^2 + \sigma_1^2 (z^2 - 2az + a^2)}{2\sigma_1^2 \sigma_2^2}} dz \\ &= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{-\frac{(\sigma_1^2 + \sigma_2^2)z^2 - 2\sigma_1^2 az + \sigma_1^2 a^2}{2\sigma_1^2 \sigma_2^2}} dz \end{aligned}$$

Ahora podemos completar cuadrados al polinomio $(\sigma_1^2 + \sigma_2^2)z^2 - 2\sigma_1^2 az + \sigma_1^2 a^2$

$$\begin{aligned} &(\sigma_1^2 + \sigma_2^2)z^2 - 2\sigma_1^2 az + \sigma_1^2 a^2 \\ &= (\sigma_1^2 + \sigma_2^2)z^2 - 2\sqrt{\sigma_1^2 + \sigma_2^2} \frac{\sigma_1^2 a}{\sqrt{\sigma_1^2 + \sigma_2^2}} z + \frac{\sigma_1^4 a^2}{\sigma_1^2 + \sigma_2^2} + \sigma_1^2 a^2 - \frac{\sigma_1^4 a^2}{\sigma_1^2 + \sigma_2^2} \\ &= \left(\sqrt{\sigma_1^2 + \sigma_2^2} z - \frac{\sigma_1^2 a}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right)^2 + \sigma_1^2 a^2 - \frac{\sigma_1^4 a^2}{\sigma_1^2 + \sigma_2^2} \end{aligned}$$

Entonces nuestra integral resultante es

$$\begin{aligned} f(z) &= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{\frac{(\sqrt{\sigma_1^2 + \sigma_2^2} z - \frac{\sigma_1^2 a}{\sqrt{\sigma_1^2 + \sigma_2^2}})^2 + \sigma_1^2 a^2 - \frac{\sigma_1^4 a^2}{\sigma_1^2 + \sigma_2^2}}{2\sigma_1^2 \sigma_2^2}} dz \\ &= \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{\mathbb{R}} e^{(\frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sqrt{2}\sigma_1 \sigma_2} z - \frac{\sigma_1^2 a}{\sqrt{2}\sigma_1 \sigma_2 \sqrt{\sigma_1^2 + \sigma_2^2}})^2 + \frac{a^2}{2\sigma_2^2} - \frac{\sigma_1^2 a^2}{2\sigma_2^2(\sigma_1^2 + \sigma_2^2)}} dz \end{aligned}$$

Por último realice la sustitución $w = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sqrt{2}\sigma_1 \sigma_2} z - \frac{\sigma_1 a}{\sqrt{2}\sigma_2 \sqrt{\sigma_1^2 + \sigma_2^2}}$ y tome $b = \frac{a^2}{2\sigma_2^2} -$

$$\frac{\sigma_1^2 a^2}{2\sigma_2^2(\sigma_1^2 + \sigma_2^2)} = \frac{a^2}{2\sigma_2^2} \left(1 - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right) = \frac{a^2}{2\sigma_2^2} \frac{\sigma_1^2 + \sigma_2^2 - \sigma_1^2}{\sigma_1^2 + \sigma_2^2} = \frac{a^2}{2(\sigma_1^2 + \sigma_2^2)}.$$

Entonces la integral resultante es

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \int_{\mathbb{R}} e^{-w^2 - b} dw \\ &= \frac{e^{-b}}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \int_{\mathbb{R}} e^{-w^2} dw \end{aligned}$$

El valor de esta integral es igual a $\sqrt{\pi}$. Luego obtenemos la expresión.

$$\begin{aligned} f(x) &= \frac{e^{-b}}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \sqrt{\pi} \\ &= \frac{e^{-b}}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{a^2}{2(\sigma_1^2 + \sigma_2^2)}} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{(x - m_1 - m_2)^2}{2(\sigma_1^2 + \sigma_2^2)}} \end{aligned}$$

Que es precisamente la función de densidad asociada a $N_{m_1+m_2, \sigma_1^2+\sigma_2^2}$ □

2. EJERCICIO 2

En un típico texto alemán, la letra e es la más frecuente con un 12.5 % de todas las letras. Un software repasa un texto desconocido puramente aleatoriamente y reconoce con la certeza de 99 % la letra e correctamente. Con la probabilidad 0.1 % reconoce una e erróneamente.

1. Qué tan grande es la probabilidad que el software indique la letra e en un paso?

Demostración. Vamos a modelar el espacio de estados como $\Omega := \{E, \neg E\} \times \{D_e, \neg D_e\}$, donde E denote leer una e, y $\neg E$ su negación, mientras que D_e denota detectar una e, y $\neg D_e$ denota su negación.

Entonces vamos a definir los estados $A = \{(E, D_e), (E, \neg D_e)\}$ y $B = \{(E, D_e), (\neg E, D_e)\}$. donde A es leer una e y B es detectar una e.

El enunciado nos indica que $\mathbb{P}(A) = 0,125$ por lo que $\mathbb{P}(A^C) = 1 - \mathbb{P}(A) = 0,875$.

Además nos indica que $\mathbb{P}(B|A) = 0,99$ y que $\mathbb{P}(B|A^C) = 0,001$.

Entonces por la fórmula de probabilidad total, como A y A^C forman una partición del espacio Ω , tenemos que $\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^C)\mathbb{P}(B|A^C) = 0,125 * 0,99 + 0,875 * 0,001 = 0,1252375 = \frac{10019}{80000}$. □

2. Qué tan grande es la probabilidad que es efectivamente la letra e cuando detecta una e ?

Demostración. Esta probabilidad la podemos calcular usando la fórmula de Bayes. Tenemos que $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{0,99 * 0,125}{0,1252375} = \frac{0,12375}{0,1252375} = \frac{9900}{10019}$ \square

3. EJERCICIO 3: SIN MEMORIA IMPLICA LA DISTRIBUCIÓN EXPONENCIAL O GEOMÉTRICA

- 1) Demuestre que para toda las variables aleatorias con valores $[0, \infty)$ cuya distribución \mathbb{P}_X es absolutamente continua con respecto a la medida de Lebesgue con una densidad estrictamente positiva y que satisface

$$\forall t, s > 0, \quad 0 < \mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s) < 1$$

existe un $\lambda > 0$ tal que $\mathbb{P}_X = \text{EXP}(\lambda)$.

Demostración. En primer lugar vamos a desarrollar un poco más la expresión de la no memoria. Por Bayes tenemos que

$$\mathbb{P}(X > s) = \mathbb{P}((X > t + s | X > t)) = \frac{\mathbb{P}(X > t | X > t + s) \mathbb{P}(X > t + s)}{\mathbb{P}(X > t)}$$

Pero ahora tenemos que si $X > t + s$ entonces $X > t$, luego $\mathbb{P}(X > t | X > t + s) = 1$.

Luego tenemos que

$$\begin{aligned} \mathbb{P}(X > s) &= \frac{\mathbb{P}(X > t + s)}{\mathbb{P}(X > t)} \\ \mathbb{P}(X > t) \mathbb{P}(X > s) &= \mathbb{P}(X > t + s) \end{aligned}$$

Ahora recordemos que la función $F(x)$ asociada a la distribución esta definida como $F(x) = \mathbb{P}(X \leq x)$. Luego $\mathbb{P}(X > x) = 1 - \mathbb{P}(X \leq x) = 1 - F(x)$.

Entonces podemos escribir la expresión anterior como

$$(1 - F(t))(1 - F(s)) = 1 - F(s + t)$$

Ahora si tomamos logaritmos tenemos que

$$\begin{aligned} \ln((1 - F(t))(1 - F(s))) &= \ln(1 - F(s + t)) \\ \ln(1 - F(t)) + \ln(1 - F(s)) &= \ln(1 - F(s + t)) \end{aligned}$$

Entonces vemos que la función $h(x) = \ln(1 - F(x))$ es un automorfismo continuo de \mathbb{R} que preserva la suma.

Entonces podemos demostrar que $h(x) = \lambda x$ para algún $\lambda \in \mathbb{R}$. Tome $\lambda = h(1)$. Primero si $x \in \mathbb{N}$ podemos demostrar por inducción que $h(x) = xh(1)$. Ahora si tomamos $x \in \mathbb{Z}$ con $x \leq 0$ tenemos que $h(0) = h(1 - 1) = h(1) + h(-1) = 0$ por lo que $h(-1) = -h(1)$, luego por inducción podemos demostrar que $h(x) = xh(1)$.

Ahora para $x \in \mathbb{Q}$ consideremos sin pérdida de generalidad el caso cuando $x = \frac{1}{n}$ para $n \in \mathbb{N} \setminus \{0\}$. Entonces tenemos que $f(1) = \underbrace{f(\frac{1}{n}) + \cdots + f(\frac{1}{n})}_{n \text{ veces}}$. Luego por inducción podemos

probar que $f(1) = nf(\frac{1}{n})$ por lo que $f(\frac{1}{n}) = \frac{1}{n}f(1)$. Ya teniendo $\frac{1}{n}$ podemos calcular $f(\frac{m}{n})$ para cualquier $m \in \mathbb{Z}$ y probar como para los enteros que $f(\frac{m}{n}) = mf(\frac{1}{n}) = \frac{m}{n}f(1)$.

Finalmente podemos demostrar que $h(x) = h(1)x$ para cualquier $x \in X$ porque ya demostramos que $h|_{\mathbb{Q}}(x) = h(1)x$. Por topología sabemos que \mathbb{Q} es denso en \mathbb{R} y además tenemos un teorema que si f y g son dos funciones continuas de X , D es denso en X y las funciones son tales que $f(x) = g(x)$ para cualquier $x \in D$ entonces $f(x) = g(x)$ para cualquier $x \in X$. Este teorema nos permite concluir que $h(x) = h(1)x = \lambda x$.

Entonces volviendo a la expresión anterior tenemos que $h(x) = \ln(1 - F(x)) = \lambda x$. Luego tomando exponencial a ambos lados tenemos que $1 - F(x) = e^{\lambda x}$ por lo que $F(x) = 1 - e^{\lambda x}$. Finalmente sabemos que la función de densidad se calcula como $f(x) = F'(x)$, entonces $f(x) = -\lambda e^{\lambda x}$. Como una de las hipótesis es que la distribución es absolutamente continua $f(x)$ debe ser siempre positivo por lo que $\lambda < 0$. Por lo que podemos escribir $f(x) = \lambda' e^{-\lambda' x}$ con $\lambda' = -\lambda > 0$, que es precisamente la función de densidad asociada a la distribución exponencial con parámetro λ' \square

- 2) Demuestre que para todas las variables aleatorias X con valores en \mathbb{N} cuya distribución \mathbb{P}_X satisface

$$\mathbb{P}(X > n + m | X > m) = \mathbb{P}(X > m) < 1 \quad \forall n, m \in \mathbb{N} \setminus \{0\}$$

Demostración. Por el mismo argumento que en el punto anterior llegamos a que

$$\mathbb{P}(X > n)\mathbb{P}(X > m) = \mathbb{P}(X > n + m)$$

Ahora podemos escribir $\mathbb{P}(X > x) = 1 - \mathbb{P}(X \leq x) = 1 - \sum_{i=1}^x \mathbb{P}(X = i)$, por lo que la expresión queda como

$$(1 - \sum_{i=1}^n \mathbb{P}(X = i))(1 - \sum_{i=1}^m \mathbb{P}(X = i)) = (1 - \sum_{i=1}^{n+m} \mathbb{P}(X = i)) \quad (3.1)$$

Ahora vamos a demostrar por inducción fuerte que

$$(1 - \sum_{i=1}^x \mathbb{P}(X = i)) = q^x \quad (3.2)$$

para algún $q \in \mathbb{R}$.

Para el caso base tenemos que $1 - \sum_{i=1}^1 \mathbb{P}(X = i) = 1 - \mathbb{P}(X = 1)$. Entonces si tomamos $q = 1 - \mathbb{P}(X = 1)$ vemos que se cumple la condición.

Ahora por la hipótesis de la no memoria tenemos que para cualquier $x > 1$, $1 - \sum_{i=1}^x \mathbb{P}(X = i) = (1 - \sum_{i=1}^n \mathbb{P}(X = i))(1 - \sum_{i=1}^m \mathbb{P}(X = i))$. para algunos $n, m \in \mathbb{N} \setminus \{0\}$. Entonces por la hipótesis de inducción tenemos que $1 - \sum_{i=1}^x \mathbb{P}(X = i) = q^n q^m = q^{n+m} = q^x$.

Esto nos permite calcular $\mathbb{P}(X = x)$ pues

$$\mathbb{P}(X = x) = 1 - \sum_{i=1}^{x-1} \mathbb{P}(X = i) - (1 - \sum_{i=1}^x \mathbb{P}(X = i)) = q^{x-1} - q^x = q^{x-1}(1 - q).$$

Finalmente tenemos que $\sum_{i=1}^{\infty} \mathbb{P}(X = i) = (1 - q) \sum_{i=1}^{\infty} q^{i-1}$ debe converger y debe ser mayor que 0. Por lo tanto, como esta es la serie geométrica tenemos que $0 \leq q < 1$. Por ultimo si tomamos $p = 1 - q$ tendríamos que $\mathbb{P}(X = x) = p(1 - p)^{x-1}$ con $0 < p \leq 1$ y por lo tanto tiene asociada la distribución geométrica con parámetro p . \square

4. EJERCICIO 4

1) Sea $X = (X_1, X_2) : \Omega \rightarrow \{0, 1, 2\}^2$ un vector aleatorio con marginales

$$\begin{aligned} \mathbb{P}_{X_1} &= \alpha_0 \delta_0 + \alpha_1 \delta_1 + \alpha_2 \delta_2 \\ \mathbb{P}_{X_2} &= \beta_0 \delta_0 + \beta_1 \delta_1 + \beta_2 \delta_2 \end{aligned}$$

Construya el sistema lineal de todas las ecuaciones y reduzcalo lo máximo posible con el algoritmo de Gauss. Cuántos grados de libertad existen para estas leyes de \mathbb{P}_X en el espacio producto $\{0, 1, 2\}^2$?

Demostración. Denotamos $\mathbb{P}_{ij} := \mathbb{P}(X_1 = i \cap X_2 = j)$. De estas ecuaciones marginales deducimos el siguiente sistema de ecuaciones.

$$\begin{aligned} \alpha_0 &= \mathbb{P}_{00} + \mathbb{P}_{01} + \mathbb{P}_{02} \\ \alpha_1 &= \mathbb{P}_{10} + \mathbb{P}_{11} + \mathbb{P}_{12} \\ \alpha_2 &= \mathbb{P}_{20} + \mathbb{P}_{21} + \mathbb{P}_{22} \\ \beta_0 &= \mathbb{P}_{00} + \mathbb{P}_{10} + \mathbb{P}_{20} \\ \beta_1 &= \mathbb{P}_{01} + \mathbb{P}_{11} + \mathbb{P}_{21} \\ \beta_2 &= \mathbb{P}_{02} + \mathbb{P}_{12} + \mathbb{P}_{22} \end{aligned}$$

Entonces la matriz asociada a este sistema es

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Y la matriz escalonada reducida resultante del algoritmo de Gauss (calculada mediante Wolfram Mathematica) es

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo que vemos que los grados de libertad para estas leyes son 5. □

2) Sea $X = (X_1, X_2, X_3) : \Omega \rightarrow \{0, 1\}^3$ un vector aleatorio con marginales

$$\begin{aligned} \mathbb{P}_{X_1} &= \alpha_0 \delta_0 + \alpha_1 \delta_1, \\ \mathbb{P}_{X_2} &= \beta_0 \delta_0 + \beta_1 \delta_1, \\ \mathbb{P}_{X_3} &= \gamma_0 \delta_0 + \gamma_1 \delta_1, \end{aligned}$$

Construya el sistema lineal de todas las ecuaciones y reduzcalo lo máximo posible con el algoritmo de Gauss. Cuántos grados de libertad existen para estas leyes de \mathbb{P}_X en el espacio producto $\{0, 1\}^3$?

Demostración. Para estas leyes el sistema de ecuaciones asociado es

$$\begin{aligned} \alpha_0 &= \mathbb{P}_{000} + \mathbb{P}_{001} + \mathbb{P}_{010} + \mathbb{P}_{011} \\ \alpha_1 &= \mathbb{P}_{100} + \mathbb{P}_{101} + \mathbb{P}_{110} + \mathbb{P}_{111} \\ \beta_0 &= \mathbb{P}_{000} + \mathbb{P}_{001} + \mathbb{P}_{100} + \mathbb{P}_{101} \\ \beta_1 &= \mathbb{P}_{010} + \mathbb{P}_{011} + \mathbb{P}_{110} + \mathbb{P}_{111} \\ \gamma_0 &= \mathbb{P}_{000} + \mathbb{P}_{010} + \mathbb{P}_{100} + \mathbb{P}_{110} \\ \gamma_1 &= \mathbb{P}_{001} + \mathbb{P}_{011} + \mathbb{P}_{101} + \mathbb{P}_{111} \end{aligned}$$

La matriz asociada a este sistema es

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Y la matriz reducida es

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & -1 & -1 & -2 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo que vemos que los grados de libertad de estas leyes son 4. □

3) Sea $X = (X_1, X_2, X_3) : \Omega \rightarrow \{0, 1\}^3$ un vector aleatorio con marginales

$$\mathbb{P}_{(X_1, X_2)} = \mathbb{P}_{(X_1, X_3)} = \mathbb{P}_{(X_2, X_3)} = \beta_{00}\delta_{(0,0)} + \beta_{01}\delta_{(0,1)} + \beta_{10}\delta_{(1,0)} + \beta_{11}\delta_{(1,1)}$$

Construya el sistema lineal de todas las ecuaciones y reduzcalo lo máximo posible con el algoritmo de Gauss. Cuántos grados de libertad existen para estas leyes de \mathbb{P}_X en el espacio producto $\{0, 1\}^3$?

Demostración. Para estas leyes el sistema de ecuaciones es

$$\begin{aligned} \beta_{00} &= \mathbb{P}_{000} + \mathbb{P}_{001} \\ \beta_{00} &= \mathbb{P}_{000} + \mathbb{P}_{010} \\ \beta_{00} &= \mathbb{P}_{000} + \mathbb{P}_{100} \\ \beta_{01} &= \mathbb{P}_{010} + \mathbb{P}_{011} \\ \beta_{01} &= \mathbb{P}_{001} + \mathbb{P}_{011} \\ \beta_{01} &= \mathbb{P}_{001} + \mathbb{P}_{101} \\ \beta_{10} &= \mathbb{P}_{100} + \mathbb{P}_{101} \\ \beta_{10} &= \mathbb{P}_{100} + \mathbb{P}_{110} \\ \beta_{10} &= \mathbb{P}_{010} + \mathbb{P}_{110} \\ \beta_{11} &= \mathbb{P}_{110} + \mathbb{P}_{111} \\ \beta_{11} &= \mathbb{P}_{101} + \mathbb{P}_{111} \\ \beta_{11} &= \mathbb{P}_{011} + \mathbb{P}_{111} \end{aligned}$$

La matriz asociada a este sistema es

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Y la matriz reducida es

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo que concluimos que los grados de libertad son 7. □

5. EJERCICIO 5: EL DILEMA DE LOS PRESOS

En una prisión hay tres presos condenados a muerte. Antonio, Bárbara y Coco. A través de un sorteo justo en el cual todos los tres tienen el mismo chance uno de los presos es perdonado, pero sin que los presos conozcan el resultado.

Antonio, que por tanto tiene la probabilidad de supervivencia de un tercio pregunta al guardián, que conoce el resultado: Dime uno de mis compañeros que va a morir. El guardián responde verazmente: Barbara. Ahora Antonio especula: Como o bien se muere Coco o bien me muero yo, tengo la probabilidad de supervivencia de un 50%.

Usted puede asumir que la probabilidad que el guardián diga Barbara o Coco es la misma, si sabe que Antonio está perdonado.

- 1) Modele este experimento con ayuda de variables aleatorias. Calcule la probabilidad de supervivencia de Antonio, si este sabe que Barbara morirá.

Demostración. Vamos a modelar primero el sorteo mediante la variable aleatoria $X : \Omega_1 \rightarrow \mathbb{R}$ donde $\Omega_1 = \{A, B, C\}$ y la distribución asociada es la uniforme discreta, es decir, que $\{X = j\}$ indica que el prisionero j es perdonado y $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \frac{1}{3}$. En segundo lugar vamos a modelar la respuesta del guardián mediante la variable aleatoria $Y : \Omega_2 \rightarrow \Omega_2$ donde $\Omega_2 = \{B, C\}$ y donde Y tiene asociada la distribución de Bernoulli para un valor p que aún necesitamos calcular. Por lo tanto el sistema total esta modelado por el vector $(X, Y) : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}^2$. Finalmente las variables son tales que $X(A) = 0$, $X(B) = 1$, $X(C) = 2$, $Y(C) = 0$ y $Y(B) = 1$.

Ahora en primer lugar calculamos $\mathbb{P}(Y = 1)$, esta esta dada por la fórmula de probabilidad total como $\mathbb{P}(Y = 0) = \mathbb{P}(X = 0)\mathbb{P}(Y = 0|X = 0) + \mathbb{P}(X = 1)\mathbb{P}(Y = 0|X = 1) + \mathbb{P}(X = 2)\mathbb{P}(Y = 0|X = 2)$.

Ahora en el enunciado mencionan que $\mathbb{P}(Y = 0|X = 0) = \frac{1}{2}$. Pero además podemos deducir que $\mathbb{P}(Y = 0|X = 1) = 0$, puesto que el guardián no puede decir que Bárbara va a morir si ella es perdonada.

Por ultimo tenemos que $\mathbb{P}(Y = 0|X = 2) = 1$ puesto que en este caso como Coco es perdonado y no puede mencionar nada sobre Antonio la única opción para el guardián es decir que Bárbara morirá.

Por lo tanto $\mathbb{P}(Y = 0) = \frac{1}{3}\frac{1}{2} + 0 + \frac{1}{3} = \frac{1}{2}$.

Ahora la probabilidad de que Antonio viva dado que el guardián dijo Bárbara la podemos calcular por Bayes como

$$\mathbb{P}(X = 0|Y = 0) = \frac{\mathbb{P}(Y = 0|X = 0)\mathbb{P}(X = 0)}{\mathbb{P}(Y = 0)} = \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

□

- 2) Usted está de acuerdo con el razonamiento de Antonio? En caso de que sí, por qué? En caso de que no, por qué no?

Demostración. Como vimos en el punto anterior, Antonio esta equivocado pues su probabilidad de supervivencia dado que conoce que Bárbara morirá sigue siendo de $\frac{1}{3}$.

Ahora obsérvese que la probabilidad de que Coco viva dado que Antonio sabe que Bárbara va a morir sube al doble pues por Bayes tenemos que

$$\mathbb{P}(X = 2|Y = 0) = \frac{\mathbb{P}(Y = 0|X = 2)\mathbb{P}(X = 2)}{\mathbb{P}(Y = 0)} = \frac{1\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

Entonces lo que Antonio debería concluir es que su probabilidad de supervivencia sigue siendo la misma pero la probabilidad de que Coco sobreviva es mayor. Luego si pudiera debería intentar sustituirse por Coco. □

6. EJERCICIO 6

Sea $n \geq 1$ fijo. Consideremos el n -ésimo tiro de una moneda justa y los eventos

$$A_k \triangleq \text{Cara en el } k\text{-ésimo tiro, } k = 1, \dots, n,$$

$$B \triangleq \text{El número de la aparición de cara es par.}$$

- 1) Construya un espacio de probabilidad adecuado.

Demostración. Podemos modelar cada uno de los k -ésimos tiros de la moneda con una variable X_k de Bernoulli tal que $X_k = 1$ si el k -ésimo tiro es cara y $X_k = 0$ de lo contrario. Como la moneda es justa tenemos que $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}$.

Entonces el espacio de probabilidad Ω va a ser el producto de los espacios de cada una de las variables, es decir, $\Omega = \{0, 1\}^n$.

Por otra parte podemos considerar la variable aleatoria Y asociada al número de veces que aparece cara en n tiros de la moneda. Por un lado tenemos que $Y = X_1 + \dots + X_n$, pero sabemos de antemano que la distribución de Y es $\mathcal{B}_{n, \frac{1}{2}}$ \square

- 2) Calcule $\mathbb{P}(A_i)$ para $1 \leq i \leq n$.

Demostración. Por la manera como modelamos el experimento tenemos que $\mathbb{P}(A_i) = \mathbb{P}(X_i = 1) = \frac{1}{2}$. \square

- 3) Calcule $\mathbb{P}(B)$.

Demostración. La probabilidad de B esta dada por la suma de las probabilidades de los eventos que el número de monedas sea par. Si $n = 2k + 1$ esta probabilidad es

$$\mathbb{P}(B) = \sum_{i=0}^k \mathbb{P}(Y = 2i) = \sum_{i=0}^k \binom{n}{2i} \frac{1}{2^{2i}} \frac{1}{2^{n-2i}} = \sum_{i=0}^k \binom{n}{2i} \frac{1}{2^n} = \frac{1}{2^n} \sum_{i=0}^k \binom{n}{2i}.$$

Finalmente podemos probar que $\sum_{i=0}^k \binom{n}{2i} = 2^{n-1}$. Esto puede ser hecho mediante un argumento de combinatoria pues esto equivale a contar el número de secuencias de 1's y 0's que tienen un número par de 1's.

Para el caso base tenemos que si $n = 1$ las únicas cadenas posibles son 0 y 1 y claramente se cumple que el número de cadenas con un número par de 1's es $2^{n-1} = 1$.

Ahora supongamos que el enunciado se cumple para n y tomemos las cadenas de tamaño $n + 1$. Estas cadenas se forman recursivamente agregando un 0 o un 1 al final de las cadenas de tamaño n . Entonces si primero tomamos las cadenas que terminan en 0, vemos que como el 0 no cambia la paridad de la suma de los dígitos, este número es igual al número de cadenas con un número par de 1's, es decir, por hipótesis de inducción, 2^{n-1} s. Por otro lado, para las que terminan en 1 este último dígito si cambia la paridad, luego es igual al número de cadenas de tamaño n con un número impar de 1's. Pero esto es igual a $2^n - 2^{n-1} = 2^{n-1}(2 - 1) = 2^{n-1}$. Por lo tanto el total de cadenas con un número par de 1's es $2^{n-1} + 2^{n-1} = 2^{n-1} \cdot 2 = 2^n$.

Por lo tanto concluimos que $\mathbb{P}(B) = \frac{1}{2^n} 2^n = \frac{1}{2}$. \square

- 4) Calcule $\mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n)$ en dependencia del número n .

Demostración. Bueno suponiendo que en los n lanzamientos se obtuvo cara, tenemos que la ocurrencia de B depende de si n es par o impar. Si n es par entonces tenemos que B se da siempre por lo que $\mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n)$ si n es impar entonces $\mathbb{P}(B)$ nunca ocurre por lo que $\mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n) = 0$. Luego $\mathbb{P}(B) = \frac{(-1)^n + 1}{2}$ \square

- 5) Es correcto que

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n \cap B) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n) \mathbb{P}(B)?$$

Demostración. Por la fórmula de Bayes tenemos que

$$\mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n) = \frac{\mathbb{P}(A_1 \cap \cdots \cap A_n \cap B)}{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)}.$$

En un punto posterior demostramos que la familia $\{A_1, \cdots, A_n\}$ es independiente por lo que $\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)$. Luego

$$\begin{aligned} \mathbb{P}(A_1 \cap \cdots \cap A_n \cap B) &= \mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n) \\ &= \mathbb{P}(A_1) \mathbb{P}(A_2) \cdots \mathbb{P}(A_n) \mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n) \end{aligned}$$

Y en los puntos anteriores ya calculamos valores para las probabilidades de estos eventos.

$$\begin{aligned} \mathbb{P}(A_1 \cap \cdots \cap A_n \cap B) &= \mathbb{P}(A_1) \mathbb{P}(A_2) \cdots \mathbb{P}(A_n) \mathbb{P}(B|A_1 \cap A_2 \cap \cdots \cap A_n) \\ &= \frac{1}{2} \frac{1}{2} \cdots \frac{1}{2} \frac{(-1)^n + 1}{2} = \frac{(-1)^n + 1}{2^{n+1}} \end{aligned}$$

Entonces concluimos que la afirmación no es correcta puesto que el resultado que predice es $\mathbb{P}(A_1) \cdots \mathbb{P}(A_n) \mathbb{P}(B) = \frac{1}{2^{n+1}}$ que es general diferente a lo mostrado anteriormente. \square

- 6) La familia $\{A_1, \cdots, A_n, B\}$ es una familia independiente?

Demostración. Si esta familia fuera independiente entonces la afirmación anterior hubiera sido correcta. Como ese no fue el caso concluimos que la familia no es independiente. \square

- 7) La familia $\{A_1, \cdots, A_n\}$ es una familia independiente?

Demostración. Por la naturaleza del experimento tenemos que cada lanzamiento es independiente de cualquier otro lanzamiento que se halla realizado. Este debe ser el caso si la variable Y efectivamente tiene la distribución binomial. Luego la familia $\{A_1, \cdots, A_n\}$ es independiente. \square

- 8) La familia $\{A_{i_1}, \cdots, A_{i_{n-1}}, B\}$ para cualquier $\{i_1 \cdots i_{n-1}\} \subset \{1, \cdots, n\}$ es una familia independiente?

Demostración. De nuevo la pregunta es si $\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_{n-1}} \cap B) = \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_{n-1}}) \mathbb{P}(B)$. Entonces consideremos $\mathbb{P}(B|A_{i_1} \cap \dots \cap A_{i_{n-1}})$. En primer lugar, esto podría depender de la paridad de n . Sin embargo, si n es par, entonces B se daría si el A_i que no se esta teniendo en cuenta no se da. Por otra parte si n es impar, entonces B depende de sí el A_i que no se esta teniendo en cuenta efectivamente se da. Entonces en ambos casos la probabilidad es igual a $\frac{1}{2}$.

De nuevo por la definición de probabilidad condicional tenemos que $\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_{n-1}} \cap B) = \mathbb{P}(B|A_{i_1} \cap \dots \cap A_{i_{n-1}}) \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_{n-1}})$. Y como la familia de A_i 's es independiente tenemos que esto es igual a $\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_{n-1}} \cap B) = \mathbb{P}(B|A_{i_1} \cap \dots \cap A_{i_{n-1}}) \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_{n-1}}) = \underbrace{\frac{1}{2} \dots \frac{1}{2}}_{n-1 \text{ veces}} \frac{1}{2} = \frac{1}{2^n}$

Entonces vemos que en este caso si coincide con lo predicho por la fórmula inicial. Por lo tanto si es independiente. Luego en nuestro modelo podemos reemplazar cualquiera de las variables X_i por Y y el modelo no cambiaria. \square

9) Qué hemos aprendido por este ejercicio?

Demostración. Si bien el evento B depende de todos los estados A_1, \dots, A_n , cuando dejamos de considerar uno cualquiera de los eventos entonces se vuelve independiente. \square