

2 Variables Aleatorias y Distribuciones de Probabilidad

Problemas

1. El abdomen del cangrejo de mar común (*Carcinus moenas*) está integrado por siete segmentos dispuestos paralelamente. En los machos se suelen apreciar fusiones entre los segmentos 3, 4 y 5. Se considera la variable aleatoria $X = \text{"Número de segmentos fusionados"}$. Esta variable puede tomar los valores 0 (ninguna fusión), 1 (se fusionan los segmentos 3 y 4, ó el 4 y 5), y 2 (se fusionan los tres segmentos entre sí). A través de diversas consideraciones sobre la genética de esta población de cangrejos, se llega a la conclusión de que las probabilidades asociadas a esta variable son de la forma:

$$P(X = 0) = \frac{1}{a(b-a)} \quad P(X = 1) = \frac{1}{b-a} \quad P(X = 2) = \frac{1}{a}, \quad a > 1$$

- a) ¿Qué relación debe existir entre a y b para que esta distribución de probabilidad esté bien definida?.
- b) Calcula la esperanza y la varianza de esta variable aleatoria.
- c) Supongamos que $a=2$. Si se tienen 6 cangrejos, ¿cuál es la probabilidad de que ninguno presente fusiones en los segmentos señalados?.

Solución:

- a) Para que la distribución de probabilidad esté bien definida han de darse dos condiciones:

1) $0 \leq P(X = k) \leq 1, \quad k = 0, 1, 2$

2) $\sum_{k=0}^2 P(X = k) = 1$

Comenzando por la segunda de estas condiciones, debe ocurrir que:

$$\begin{aligned}
 P(X=0) + P(X=1) + P(X=2) &= 1 \\
 \frac{1}{a(b-a)} + \frac{1}{b-a} + \frac{1}{a} &= 1 \\
 \frac{1+a+b-a}{a(b-a)} &= 1 \\
 1+b &= ab-a^2 \\
 1+a^2 &= (a-1)b \\
 b &= \frac{1+a^2}{a-1}
 \end{aligned}$$

Llevando ahora este valor de b a la expresión de las probabilidades $P(X=k)$ para $k=0,1$, tenemos:

$$P(X=0) = \frac{1}{a(b-a)} = \frac{a-1}{a(a+1)}; \quad P(X=1) = \frac{a-1}{a+1}$$

donde hemos empleado que:

$$\frac{1}{b-a} = \frac{1}{\frac{1+a^2}{a-1} - a} = \frac{1}{\frac{1+a^2-a^2+a}{a-1}} = \frac{a-1}{a+1}$$

Podemos comprobar que efectivamente:

$$\begin{aligned}
 P(X=0) + P(X=1) + P(X=2) &= \frac{a-1}{a(a+1)} + \frac{a-1}{(a+1)} + \frac{1}{a} = \\
 &= \frac{a-1+a(a-1)+a+1}{a(a+1)} \\
 &= \frac{a-a+1+a^2-a+a+1}{a(a+1)} = 1
 \end{aligned}$$

Asimismo, como a es mayor que 1 según se indica en el enunciado, es trivial comprobar que las tres probabilidades $P(X=0)$, $P(X=1)$ y $P(X=2)$ están comprendidas entre 0 y 1.

- b) Por definición, la esperanza de una variable aleatoria discreta es $E[X] = \sum_k kP(X=k)$. En este caso:

$$\begin{aligned}
 E[X] &= 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) = \\
 &= \frac{a-1}{a+1} + 2 \cdot \frac{1}{a} = \frac{a(a-1)+2(a+1)}{a(a+1)} = \frac{a^2+a+2}{a(a+1)}
 \end{aligned}$$

Asimismo, la varianza es $\text{var}(X) = E[X^2] - (E[X])^2$, donde $E[X^2] = \sum_k k^2 P(X=k)$. En este caso:

$$\begin{aligned}
 E[X^2] &= 0^2 \cdot P(X=0) + 1^2 \cdot P(X=1) + 2^2 \cdot P(X=2) = \\
 &= \frac{a-1}{a+1} + 4 \cdot \frac{1}{a} = \frac{a(a-1)+4(a+1)}{a(a+1)} = \frac{a^2+3a+2}{a(a+1)} = \\
 &= \frac{(a+2)(a+1)}{a(a+1)} = \frac{a+2}{a}
 \end{aligned}$$

Por tanto:

$$\begin{aligned}
 \text{var}(X) &= E[X^2] - (E[X])^2 = \frac{a+2}{a} - \left(\frac{a^2+a+2}{a(a+1)} \right)^2 = \\
 &= \frac{a+2}{a} - \frac{a^4+a^2+2a+4+2a^3+4a^2}{a^2(a+1)^2} = \\
 &= \frac{a(a+1)^2(a+2) - a^4 - 2a^3 - 5a^2 - 2a - 4}{a^2(a+1)^2} = \\
 &= \frac{a^4 + 4a^3 + 5a^2 + 2a - a^4 - 2a^3 - 5a^2 - 2a - 4}{a^2(a+1)^2} = \\
 &= \frac{2a^3 - 4}{a^2(a+1)^2} = \frac{2(a^3 - 2)}{a^2(a+1)^2}
 \end{aligned}$$

- c) Si $a = 2$ entonces $P(X = 0) = \frac{a-1}{a(a+1)} = \frac{1}{6}$. Si suponemos que los 6 cangrejos son independientes (suposición razonable), y llamamos $X_i = \text{"Número de fusiones que presenta el cangrejo } i\text{"}$, la probabilidad de que ninguno presente fusiones es:

$$\begin{aligned}
 P(X_1 = 0 \cap X_2 = 0 \cap X_3 = 0 \cap X_4 = 0 \cap X_5 = 0 \cap X_6 = 0) &= \\
 = P(X_1 = 0) P(X_2 = 0) P(X_3 = 0) P(X_4 = 0) P(X_5 = 0) P(X_6 = 0) &= \left(\frac{1}{6} \right)^6
 \end{aligned}$$

2. El número de días que transcurren desde que cierto vertido industrial llega al mar hasta que puede considerarse que dicho vertido ha desaparecido del entorno del emisario por el cual fue expulsado, es una variable aleatoria cuya función de probabilidad se supone de la forma:

$$f(x) = \vartheta(0,7)^k, \quad k = 0, 1, 2, 3, \dots$$

siendo ϑ es una constante.

- Determinar el valor de ϑ y calcular el número esperado de días que transcurren desde la emisión hasta la desaparición del vertido.
- Con probabilidad 0.95, ¿cuál es el número máximo de días que tarda en desaparecer el vertido?
- Sabiendo que tras 4 días aún no ha desaparecido el vertido, ¿cuál es la probabilidad de que desaparezca el quinto día?

Solución:

- a) Para que la función de probabilidad esté bien definida debe ocurrir que $\sum_k f(k) = 1$. En este caso, los valores de k van desde 0 hasta ∞ , por lo que:

$$\sum_{k=0}^{\infty} \vartheta (0,7)^k = 1 \Rightarrow \vartheta \sum_{k=0}^{\infty} (0,7)^k = 1 \Rightarrow \vartheta \frac{1}{1-0,7} = 1 \Rightarrow \vartheta = 0,3$$

Nota: Hemos utilizado aquí el valor de la suma de la *serie geométrica de razón ρ menor que la unidad* es $\sum_{k=0}^{\infty} \rho^k = \frac{1}{1-\rho}$. Este valor puede obtenerse del siguiente modo:

- 1) Llamamos $S_n = \sum_{k=0}^n \rho^k = 1 + \rho + \rho^2 + \rho^3 + \dots + \rho^{n-1} + \rho^n$
- 2) Calculamos $\rho S_n = \rho + \rho^2 + \rho^3 + \dots + \rho^{n-1} + \rho^n + \rho^{n+1}$
- 3) Restamos $S_n - \rho S_n = (1 + \rho + \rho^2 + \rho^3 + \dots + \rho^{n-1} + \rho^n) - (\rho + \rho^2 + \rho^3 + \dots + \rho^{n-1} + \rho^n + \rho^{n+1}) = 1 - \rho^{n+1}$, de donde se sigue que $(1 - \rho) S_n = 1 - \rho^{n+1}$.
- 4) Despejamos $S_n = \frac{1 - \rho^{n+1}}{1 - \rho}$.
- 5) Entonces, si $0 < \rho < 1$, se tiene que $\lim_{n \rightarrow \infty} \rho^n = 0$ y en tal caso $\sum_{k=0}^{\infty} \rho^k = S_{\infty} = \lim_{n \rightarrow \infty} S_n = \frac{1}{1 - \rho}$.

Para calcular la esperanza de esta variable hemos de hallar $E[X] = \sum_{k=0}^{\infty} k P(X = k) = \sum_{k=0}^{\infty} k \cdot 0,3 \cdot 0,7^k = 0,3 \sum_{k=0}^{\infty} k \cdot 0,7^k = 0,3 \cdot \frac{0,7}{(1-0,7)^2} = \frac{0,7}{0,3} = 2,3333$ días.

Nota: Ahora hemos utilizado el valor de la suma de la *serie aritmético-geométrica de razón ρ menor que la unidad* es $\sum_{k=1}^{\infty} k \rho^k = \frac{\rho}{(1-\rho)^2}$. Este valor puede obtenerse del siguiente modo:

- 1) Llamamos $S_n = \sum_{k=1}^n k \rho^k = \rho + 2\rho^2 + 3\rho^3 + \dots + (n-1)\rho^{n-1} + n\rho^n$
- 2) Calculamos $\rho S_n = \rho^2 + 2\rho^3 + \dots + (n-1)\rho^n + n\rho^{n+1}$
- 3) Restamos $S_n - \rho S_n = (\rho + 2\rho^2 + 3\rho^3 + \dots + (n-1)\rho^{n-1} + n\rho^n) - (\rho^2 + 2\rho^3 + \dots + (n-1)\rho^n + n\rho^{n+1}) = \rho + \rho^2 + \rho^3 + \dots + \rho^n - n\rho^{n+1}$, de donde se sigue que:

$$\begin{aligned} (1 - \rho) S_n &= -1 + (1 + \rho + \rho^2 + \rho^3 + \dots + \rho^n) - n\rho^{n+1} = \\ &= -1 + \frac{1 - \rho^{n+1}}{1 - \rho} - n\rho^{n+1} = \frac{\rho - \rho^{n+1} - n\rho^{n+1} + n\rho^{n+2}}{1 - \rho} = \\ &= \frac{\rho - (n+1)\rho^{n+1} + n\rho^{n+2}}{1 - \rho} = \frac{\rho(1 - (n+1)\rho^n + n\rho^{n+1})}{1 - \rho} \end{aligned}$$

- 4) Despejamos $S_n = \frac{\rho(1 - (n+1)\rho^n + n\rho^{n+1})}{(1-\rho)^2}$.
- 5) Entonces, si $0 < \rho < 1$, se tiene que $\lim_{n \rightarrow \infty} (n+1)\rho^n = 0$ y $\lim_{n \rightarrow \infty} n\rho^{n+1} = 0$, por lo que en tal caso $\sum_{k=1}^{\infty} k \rho^k = S_{\infty} = \lim_{n \rightarrow \infty} S_n = \frac{\rho}{(1-\rho)^2}$.

- b) Si X es el número de días que tarda en desaparecer el vertido, hemos de calcular el valor M tal que $P(X \leq M) = 0,95$. Ahora bien:

$$P(X \leq M) = \sum_{k=0}^M P(X = k) = \sum_{k=0}^M 0,3 \cdot 0,7^k = 0,3 \frac{1 - 0,7^{M+1}}{1 - 0,7} = 1 - 0,7^{M+1}$$

Entonces:

$$P(X \leq M) = 0,95 \Leftrightarrow 1 - 0,7^{M+1} = 0,95 \Leftrightarrow 0,7^{M+1} = 0,05 \Leftrightarrow (M + 1) \ln(0,7) = \ln(0,05)$$

$$M = \frac{\ln(0,05)}{\ln(0,7)} - 1 = 7,3991$$

Si deseamos mostrar la solución con un número entero de días podríamos decir que $P(X \leq 8) \geq 0,95$.

- c) Se trata de una probabilidad condicionada:

$$\begin{aligned} P(X = 5 | X > 4) &= \frac{P(\{X = 5\} \cap \{X > 4\})}{P(X > 4)} = \frac{P(X = 5)}{1 - P(X \leq 3)} = \\ &= \frac{P(X = 5)}{1 - \sum_{k=0}^3 P(X = k)} = \frac{0,3 \cdot 0,7^5}{1 - 0,3 \sum_{k=0}^3 0,7^k} = \\ &= \frac{0,3 \cdot 0,7^5}{1 - 0,3 \frac{1 - 0,7^4}{1 - 0,7}} = \frac{0,3 \cdot 0,7^5}{0,7^4} = 0,21 \end{aligned}$$

3. La velocidad (en cm/seg) de una corriente marina en cierta zona es una variable aleatoria con función de densidad

$$f(x) = \frac{1}{100} e^{-\vartheta x}, \quad x \geq 0$$

- a) ¿Cuál es la velocidad media de la corriente?
 b) Se coloca una boya cuyo anclaje es capaz de resistir el empuje de una corriente de hasta 4 km/h. ¿Cuál es la probabilidad de que el anclaje se rompa?
 c) ¿Cuál es la probabilidad de que se rompa en un día en que la velocidad de la corriente ha superado ya los 3.5 km/h?

Solución:

Antes de realizar ningún cálculo necesitamos determinar el valor de la constante ϑ . Para ello imponemos la condición de que la probabilidad total (área total bajo la función de densidad)

debe ser 1. En este caso:

$$\begin{aligned}\int_0^{\infty} f(x) dx &= \int_0^{\infty} \frac{1}{100} e^{-\vartheta x} dx = \frac{1}{100} \int_0^{\infty} e^{-\vartheta x} dx = \frac{1}{100} \left[-\frac{1}{\vartheta} e^{-\vartheta x} \right]_0^{\infty} = \\ &= \frac{1}{100} \left(-\frac{1}{\vartheta} \right) [e^{-\infty} - e^0] = \frac{1}{100} \frac{1}{\vartheta}\end{aligned}$$

Entonces:

$$\int_0^{\infty} f(x) dx = 1 \Leftrightarrow \frac{1}{100} \frac{1}{\vartheta} = 1 \Leftrightarrow \vartheta = \frac{1}{100}$$

Por tanto, la función de densidad de la variable $X = \text{"Velocidad de la corriente"}$ es $f(x) = \frac{1}{100} e^{-\frac{1}{100}x}$, $x \geq 0$.

a) El valor medio, o esperado, de la velocidad de la corriente es:

$$E[X] = \int_0^{\infty} x f(x) dx = \int_0^{\infty} x \frac{1}{100} e^{-\frac{1}{100}x} dx$$

Esta integral puede resolverse fácilmente por partes; llamando $u = x$, $dv = \frac{1}{100} e^{-\frac{1}{100}x} dx$, se tiene que $du = dx$, $v = -e^{-\frac{1}{100}x}$ por lo que:

$$\int_0^{\infty} x \frac{1}{100} e^{-\frac{1}{100}x} dx = -x e^{-\frac{1}{100}x} + \int_0^{\infty} e^{-\frac{1}{100}x} dx = \left[-x e^{-\frac{1}{100}x} - 100 e^{-\frac{1}{100}x} \right]_0^{\infty} = 100$$

Así pues, la velocidad media de la corriente es de 100 cm/seg = 360.000 cm/hora = 3.6 km/hora.

b) Una velocidad de 4 km/hora equivale a $\frac{400,000 \text{ cm}}{3,600 \text{ seg}} = 111,11 \text{ cm/seg}$. La probabilidad de que el anclaje se rompa es entonces:

$$P(X \geq 111,11) = \int_{111,11}^{\infty} \frac{1}{100} e^{-\frac{1}{100}x} dx = \left[-e^{-\frac{1}{100}x} \right]_{111,11}^{\infty} = e^{-\frac{1}{100} \cdot 111,11} = 0,3292$$

c) Una velocidad de 3.5 km/hora es equivalente a $\frac{350,000 \text{ cm}}{3,600 \text{ seg}} = 97,22 \text{ cm/seg}$. La probabilidad de que el anclaje se rompa es ahora la probabilidad de que la velocidad supere la resistencia del anclaje (esto es, supere los 111.11 cm/seg.) sabiendo que ya supera los 97.22 cm/seg. La probabilidad pedida es, pues, una probabilidad condicionada:

$$\begin{aligned}P(X \geq 111,11 | X \geq 97,22) &= \frac{P(\{X \geq 111,11\} \cap \{X \geq 97,22\})}{P(X \geq 97,22)} = \\ &= \frac{P(X \geq 111,11)}{P(X \geq 97,22)} = \frac{e^{-\frac{1}{100} \cdot 111,11}}{e^{-\frac{1}{100} \cdot 97,22}} = \\ &= e^{-\frac{1}{100}(111,11-97,22)} = e^{-\frac{1}{100}13,89} = 0,87032\end{aligned}$$

4. La cantidad de arena (en kg.) depositada cada día en una playa debido a la meteorización de

las rocas del relieve próximo es una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} \frac{x^2 - 10000}{\vartheta} & -100 \leq x \leq 100 \\ 0 & |x| > 100 \end{cases}$$

- a) Determinar el valor de ϑ para que $f(x)$ sea función de densidad.
- b) ¿Cuál es la cantidad esperada de arena que se deposita cada día en la playa?
- c) ¿Cuál es la variabilidad de estos depósitos de arena?
- d) ¿Cuál es la probabilidad de que en un día cualquiera elegido al azar la playa gane arena?
¿y de que la pierda?
- e) ¿Cuál es la probabilidad de que en un día se ganen entre 20 y 50 kg de arena?
- f) ¿Cuál es la probabilidad de que en un día se ganen menos de 50 kg de arena si se sabe que ese día se han ganado más de 20 kg?

Solución:

- a) Nuevamente, para determinar ϑ debemos buscar el valor que hace que la probabilidad total sea 1. En este caso:

$$\begin{aligned} \int_{-100}^{100} f(x) dx &= \int_{-100}^{100} \frac{x^2 - 10000}{\vartheta} dx = \frac{1}{\vartheta} \left[\frac{x^3}{3} - 10000x \right]_{-100}^{100} = \\ &= \frac{1}{\vartheta} \left[\left(\frac{100^3}{3} - 10^6 \right) - \left(\frac{-100^3}{3} + 10^6 \right) \right] = \frac{1}{\vartheta} \left(\frac{2}{3} 10^6 - 2 \cdot 10^6 \right) \\ &= \frac{2 \cdot 10^6}{\vartheta} \left(-\frac{2}{3} \right) = -\frac{4 \cdot 10^6}{3\vartheta} \end{aligned}$$

Para que esta integral valga 1 ha de ocurrir que:

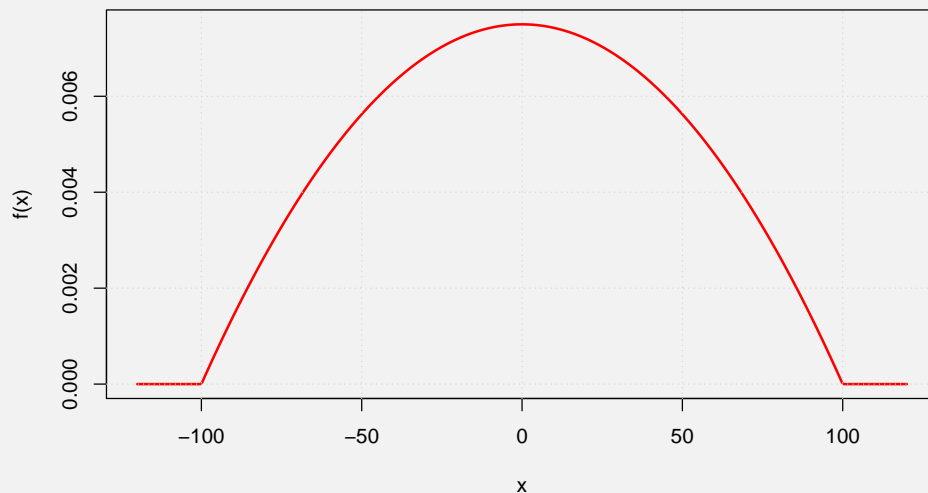
$$-\frac{4 \cdot 10^6}{3\vartheta} = 1 \Rightarrow \vartheta = -\frac{4}{3} \cdot 10^6$$

Por tanto, podemos reescribir la función de densidad de la variable X = "Cantidad de arena depositada en la playa diariamente" como:

$$f(x) = \frac{x^2 - 10000}{\vartheta} = -\frac{3}{4} \frac{x^2 - 10000}{10^6} = \frac{3(10000 - x^2)}{4 \cdot 10^6}, -100 \leq x \leq 100$$

El siguiente código R nos permite dibujar esta función:

```
> f=function(x){ifelse(abs(x)<=100,3*(10000-x^2)/4e6,0)}  
> x=seq(-120,120,length=500)  
> plot(x,f(x),type="l",col="red",lwd=2)  
>
```



Como puede apreciarse, esta función de densidad entre -100 y 100 es una parábola invertida.

- b) Si observamos la gráfica anterior, es evidente que la esperanza de esta variable es cero, ya que la función de densidad es simétrica respecto al origen, por lo que su centro de masas (esperanza) está en dicho punto. Desde un punto de vista empírico, ello significa que la playa está en equilibrio, y que en promedio recibe tanta arena como pierde.

Para calcular esta esperanza de manera analítica hemos de resolver:

$$\begin{aligned} \int_{-100}^{100} x f(x) dx &= \int_{-100}^{100} x \frac{3(10000 - x^2)}{4 \cdot 10^6} dx = \frac{3}{4 \cdot 10^6} \left[10000 \frac{x^2}{2} - \frac{x^4}{4} \right]_{-100}^{100} = \\ &= \frac{3}{4 \cdot 10^6} \left[\left(10^4 \frac{100^2}{2} - \frac{100^4}{4} \right) - \left(10^4 \frac{(-100)^2}{2} - \frac{(-100)^4}{4} \right) \right] = 0 \end{aligned}$$

- c) Para calcular la variabilidad hemos de hallar la varianza:

$$\begin{aligned} var(X) &= E[X^2] - (E[X])^2 = E[X^2] = \int_{-100}^{100} x^2 \frac{3(10000 - x^2)}{4 \cdot 10^6} dx = \\ &= \frac{3}{4 \cdot 10^6} \left[10000 \frac{x^3}{3} - \frac{x^5}{5} \right]_{-100}^{100} = \\ &= \frac{3}{4 \cdot 10^6} \left[\left(10^4 \frac{100^3}{3} - \frac{100^5}{5} \right) - \left(10^4 \frac{(-100)^3}{3} - \frac{(-100)^5}{5} \right) \right] = \\ &= \frac{3}{4 \cdot 10^6} \left(\frac{2}{3} 10^{10} - \frac{2}{5} 10^{10} \right) = \frac{6 \cdot 10^{10}}{4 \cdot 10^6} \left(\frac{1}{3} - \frac{1}{5} \right) = \frac{3}{2} \cdot 10^4 \left(\frac{2}{15} \right) = 2000 \text{ kg}^2 \end{aligned}$$

Podemos expresar la variabilidad también en términos de la desviación típica:

$$\sigma = \sqrt{var(X)} = \sqrt{2000} = 44,721 \text{ kg}.$$

d) La probabilidad de que la playa gane arena es $P(X \geq 0)$ y la probabilidad de que pierda es $P(X \leq 0)$. De la simple observación de la gráfica de la función de densidad de esta variable, simétrica respecto al origen, se deduce que ambas probabilidades deben ser iguales a $1/2$. En cualquier caso, pueden calcularse analíticamente como:

$$\begin{aligned} P(X \geq 0) &= \int_0^{100} f(x) dx = \int_0^{100} \frac{3(10000 - x^2)}{4 \cdot 10^6} dx = \frac{3}{4 \cdot 10^6} \left[10000x - \frac{x^3}{3} \right]_0^{100} = \\ &= \frac{3}{4 \cdot 10^6} \left(10^6 - \frac{10^6}{3} \right) = \frac{3}{4} \left(1 - \frac{1}{3} \right) = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2} \end{aligned}$$

$$P(X \leq 0) = 1 - P(X \geq 0) = \frac{1}{2}$$

$$\begin{aligned} e) P(20 \leq X \leq 50) &= \int_{20}^{50} \frac{3(10000 - x^2)}{4 \cdot 10^6} dx = \frac{3}{4 \cdot 10^6} \left[10000x - \frac{x^3}{3} \right]_{20}^{50} = \\ &= \frac{3}{4 \cdot 10^6} \left[\left(5 \cdot 10^5 - \frac{50^3}{3} \right) - \left(2 \cdot 10^5 - \frac{20^3}{3} \right) \right] = \frac{3}{4 \cdot 10^3} \left(300 - \frac{125}{3} + \frac{8}{3} \right) = 0,19575 \end{aligned}$$

f) Se trata en este caso de una probabilidad condicionada:

$$\begin{aligned} P(X < 50 | X \geq 20) &= \frac{P(\{X < 50\} \cap \{X \geq 20\})}{P(X \geq 20)} = \frac{P(20 \leq X < 50)}{P(X \geq 20)} = \\ &= \frac{0,19575}{0,352} = 0,55611 \end{aligned}$$

donde:

$$\begin{aligned} P(X \geq 20) &= \int_{20}^{100} f(x) dx = \int_{20}^{100} \frac{3(10000 - x^2)}{4 \cdot 10^6} dx = \frac{3}{4 \cdot 10^6} \left[10000x - \frac{x^3}{3} \right]_{20}^{100} = \\ &= \frac{3}{4 \cdot 10^6} \left[\left(10^6 - \frac{10^6}{3} \right) - \left(2 \cdot 10^5 - \frac{20^3}{3} \right) \right] = \\ &= \frac{3}{4 \cdot 10^3} \left(800 - \frac{1000}{3} + \frac{8}{3} \right) = 0,352 \end{aligned}$$

5. La proporción X , en tanto por uno, en que cierto compuesto químico aparece presente en la composición de las algas de determinada especie es una variable aleatoria con función de densidad:

$$f(x) = (\vartheta + 1)x^\vartheta, 0 \leq x \leq 1$$

- Comprobar que $f(x)$ es siempre función de densidad, cualquiera que sea el valor de ϑ .
- ¿Cuál debe ser el valor de ϑ para que el valor esperado de X sea 0.7?
- ¿Cuál debe ser este valor para que $P(X \leq 0,7) = 0,95$?

Solución:

- a) Para comprobar que $f(x)$ es siempre función de densidad cualquiera que sea el valor de ϑ hay que verificar que $\int_0^1 f(x) dx = 1 \forall \vartheta$. En efecto:

$$\int_0^1 (\vartheta + 1) x^\vartheta dx = (\vartheta + 1) \int_0^1 x^\vartheta dx = (\vartheta + 1) \left[\frac{x^{\vartheta+1}}{\vartheta + 1} \right]_0^1 = (1^{\vartheta+1} - 0^{\vartheta+1}) = 1 \forall \vartheta$$

- b) La esperanza de esta variable aleatoria es:

$$\begin{aligned} E[X] &= \int_0^1 x f(x) dx = \int_0^1 x (\vartheta + 1) x^\vartheta dx = (\vartheta + 1) \int_0^1 x^{\vartheta+1} dx = \\ &= (\vartheta + 1) \left[\frac{x^{\vartheta+2}}{\vartheta + 2} \right]_0^1 = \frac{\vartheta + 1}{\vartheta + 2} \end{aligned}$$

Para que su valor sea 0.7:

$$\frac{\vartheta + 1}{\vartheta + 2} = 0,7 \Rightarrow \vartheta + 1 = 0,7(\vartheta + 2) \Rightarrow \vartheta - 0,7\vartheta = 1,4 - 1 \Rightarrow \vartheta = \frac{0,4}{0,3} = 1,3333$$

- c) $P(X \leq 0,7) = \int_0^{0,7} (\vartheta + 1) x^\vartheta dx = [x^{\vartheta+1}]_0^{0,7} = 0,7^{\vartheta+1}$. Para que esta probabilidad sea 0.95:

$$0,7^{\vartheta+1} = 0,95 \Rightarrow (\vartheta + 1) \ln(0,7) = \ln(0,95) \Rightarrow \vartheta = \frac{\ln(0,95)}{\ln(0,7)} - 1 \Rightarrow \vartheta = -0,85619$$

6. En otra especie de algas, la concentración de este compuesto químico tiene por función de densidad la siguiente:

$$f(x) = \frac{2\vartheta_1 x + \vartheta_2}{\vartheta_1 + \vartheta_2}, \quad 0 \leq x \leq 1$$

- a) ¿Para qué valores de ϑ_1 y ϑ_2 es $f(x)$ una función de densidad correctamente definida?
- b) Demostrar que no existen valores de ϑ_1 y ϑ_2 para los que $E[X] = 0,7$ y $\text{var}(X) = 0,16$
- c) Demostrar que existen infinitos valores de ϑ_1 y ϑ_2 para los que $E[X] = 0,7$ y $\text{var}(X) = \frac{13}{300} = 0,043333$.
- d) Reparametrizar $f(x)$ de tal manera que los valores de sus parámetros sean únicos en el caso anterior. (Nota: Tener en cuenta que $f(x)$ es una recta entre 0 y 1)

Solución:

a) Para que $f(x)$ esté bien definida, debe verificar que $\int_0^1 f(x) dx = 1 \forall \vartheta$. Entonces:

$$\begin{aligned}\int_0^1 \frac{2\vartheta_1 x + \vartheta_2}{\vartheta_1 + \vartheta_2} dx &= \frac{1}{\vartheta_1 + \vartheta_2} \left[2\vartheta_1 \frac{x^2}{2} + \vartheta_2 x \right]_0^1 = \\ &= \frac{1}{\vartheta_1 + \vartheta_2} (\vartheta_1 + \vartheta_2) = 1 \quad \forall \vartheta_1, \vartheta_2, \text{ (Si } \vartheta_1 + \vartheta_2 \neq 0)\end{aligned}$$

Por tanto esta función está definida para cualesquiera valores de ϑ_1 y ϑ_2 , siempre y cuando $\vartheta_1 + \vartheta_2 \neq 0$.

b) Tenemos:

$$\begin{aligned}E[X] &= \int_0^1 x f(x) dx = \int_0^1 x \frac{2\vartheta_1 x + \vartheta_2}{\vartheta_1 + \vartheta_2} dx = \\ &= \frac{1}{\vartheta_1 + \vartheta_2} \left[2\vartheta_1 \frac{x^3}{3} + \vartheta_2 \frac{x^2}{2} \right]_0^1 = \frac{1}{\vartheta_1 + \vartheta_2} \left(\frac{2}{3}\vartheta_1 + \frac{1}{2}\vartheta_2 \right)\end{aligned}$$

$$\begin{aligned}E[X^2] &= \int_0^1 x^2 f(x) dx = \int_0^1 x^2 \frac{2\vartheta_1 x + \vartheta_2}{\vartheta_1 + \vartheta_2} dx = \\ &= \frac{1}{\vartheta_1 + \vartheta_2} \left[2\vartheta_1 \frac{x^4}{4} + \vartheta_2 \frac{x^3}{3} \right]_0^1 = \frac{1}{\vartheta_1 + \vartheta_2} \left(\frac{1}{2}\vartheta_1 + \frac{1}{3}\vartheta_2 \right)\end{aligned}$$

Sabemos que $\text{var}(X) = E[X^2] - (E[X])^2$, por lo que $E[X^2] = \text{var}(X) + (E[X])^2$. Como en este caso $\text{var}(X) = 0,16$ y $E[X] = 0,7$, se sigue que $E[X^2] = 0,16 + 0,49 = 0,65$. Sustituyendo estos valores en las expresiones anteriores:

$$\left. \begin{aligned} \frac{1}{\vartheta_1 + \vartheta_2} \left(\frac{2}{3}\vartheta_1 + \frac{1}{2}\vartheta_2 \right) &= 0,7 \\ \frac{1}{\vartheta_1 + \vartheta_2} \left(\frac{1}{2}\vartheta_1 + \frac{1}{3}\vartheta_2 \right) &= 0,65 \end{aligned} \right\} \Rightarrow \left. \begin{aligned} \left(\frac{2}{3}\vartheta_1 + \frac{1}{2}\vartheta_2 \right) &= 0,7(\vartheta_1 + \vartheta_2) \\ \left(\frac{1}{2}\vartheta_1 + \frac{1}{3}\vartheta_2 \right) &= 0,65(\vartheta_1 + \vartheta_2) \end{aligned} \right\}$$

de donde:

$$\left. \begin{aligned} \left(\frac{2}{3} - 0,7 \right) \vartheta_1 + \left(\frac{1}{2} - 0,7 \right) \vartheta_2 &= 0 \\ \left(\frac{1}{2} - 0,65 \right) \vartheta_1 + \left(\frac{1}{3} - 0,65 \right) \vartheta_2 &= 0 \end{aligned} \right\}$$

Multiplicando ambas ecuaciones por 6 (para eliminar las fracciones):

$$\left. \begin{aligned} (4 - 4,2) \vartheta_1 + (3 - 4,2) \vartheta_2 &= 0 \\ (3 - 3,9) \vartheta_1 + (2 - 3,9) \vartheta_2 &= 0 \end{aligned} \right\} \Rightarrow \left. \begin{aligned} -0,2\vartheta_1 - 1,2\vartheta_2 &= 0 \\ -0,9\vartheta_2 - 1,9\vartheta_2 &= 0 \end{aligned} \right\}$$

Este sistema de ecuaciones es homogéneo; como el rango de su matriz de coeficientes es 2, su única solución es $\vartheta_1 = \vartheta_2 = 0$; pero en tal caso $f(x)$ no es una función de densidad, por lo que podemos concluir que no existen valores de los parámetros ϑ_1 y ϑ_2 que produzcan los valores de esperanza y varianza indicados.

c) En este caso $E[X^2] = \frac{13}{300} + \frac{49}{100} = \frac{160}{300} = \frac{16}{30}$, y procediendo como antes:

$$\left. \begin{aligned} \frac{1}{\vartheta_1 + \vartheta_2} \left(\frac{2}{3}\vartheta_1 + \frac{1}{2}\vartheta_2 \right) &= 0,7 \\ \frac{1}{\vartheta_1 + \vartheta_2} \left(\frac{1}{2}\vartheta_1 + \frac{1}{3}\vartheta_2 \right) &= \frac{16}{30} \end{aligned} \right\} \Rightarrow \left. \begin{aligned} \left(\frac{2}{3}\vartheta_1 + \frac{1}{2}\vartheta_2 \right) &= \frac{7}{10}(\vartheta_1 + \vartheta_2) \\ \left(\frac{1}{2}\vartheta_1 + \frac{1}{3}\vartheta_2 \right) &= \frac{16}{30}(\vartheta_1 + \vartheta_2) \end{aligned} \right\}$$

de donde:

$$\begin{cases} \left(\frac{2}{3} - \frac{7}{10}\right) \vartheta_1 + \left(\frac{1}{2} - \frac{7}{10}\right) \vartheta_2 = 0 \\ \left(\frac{1}{2} - \frac{16}{30}\right) \vartheta_1 + \left(\frac{1}{3} - \frac{16}{30}\right) \vartheta_2 = 0 \end{cases}$$

Multiplicando ambas ecuaciones por 30 (para eliminar las fracciones):

$$\begin{cases} (20 - 21) \vartheta_1 + (15 - 21) \vartheta_2 = 0 \\ (15 - 16) \vartheta_1 + (10 - 16) \vartheta_2 = 0 \end{cases} \Rightarrow \begin{cases} -\vartheta - 6\vartheta_2 = 0 \\ -\vartheta_1 - 6\vartheta_2 = 0 \end{cases}$$

Por tanto tenemos una única ecuación $\vartheta_1 + 6\vartheta_2 = 0$, que tiene infinitas soluciones: dado cualquier valor $a \in \mathbb{R} - \{0\}$, el par $\vartheta_2 = a, \vartheta_1 = -6a$ es solución.

d) Si observamos que $f(x)$ es en realidad la ecuación de una recta, podemos reparametrizar esta función de densidad como:

$$f(x) = \frac{2\vartheta_1 x + \vartheta_2}{\vartheta_1 + \vartheta_2} = \frac{2\vartheta_1}{\vartheta_1 + \vartheta_2} x + \frac{\vartheta_2}{\vartheta_1 + \vartheta_2} = \varphi_1 x + \varphi_2$$

y para que esté bien definida:

$$\int_0^1 f(x) dx = 1 \Rightarrow \int_0^1 (\varphi_1 x + \varphi_2) dx = 1 \Rightarrow \left[\varphi_1 \frac{x^2}{2} + \varphi_2 x \right]_0^1 = 1 \Rightarrow \frac{\varphi_1}{2} + \varphi_2 = 1$$

Por tanto:

$$\varphi_1 = 2(1 - \varphi_2)$$

y sustituyendo en la expresión anterior de $f(x)$:

$$f(x) = \varphi_1 x + \varphi_2 = 2(1 - \varphi_2)x + \varphi_2$$

Por tanto, en realidad $f(x)$ depende de un único parámetro; podemos prescindir del subíndice y escribir directamente:

$$f(x) = 2(1 - \varphi)x + \varphi, \quad 0 \leq x \leq 1$$

En particular, para los valores de $E[X]$ y $var(X)$ del apartado anterior se tenía que $\vartheta_1 + 6\vartheta_2 = 0$. Como $\varphi = \varphi_2 = \frac{\vartheta_2}{\vartheta_1 + \vartheta_2}$ resulta:

$$\varphi = \frac{\vartheta_2}{\vartheta_1 + \vartheta_2} = \frac{\vartheta_2}{-6\vartheta_2 + \vartheta_2} = \frac{\vartheta_2}{-5\vartheta_2} = -\frac{1}{5}$$

y la densidad en este caso se reduce a:

$$f(x) = 2\left(1 + \frac{1}{5}\right)x - \frac{1}{5} = \frac{2}{5}x + \frac{9}{5}$$

7. En otra especie de algas, la concentración de este compuesto químico tiene por función de

densidad la siguiente:

$$f(x) = \frac{3\vartheta_1 x^2 + 2\vartheta_2 x + 1}{\vartheta_1 + \vartheta_2 + 1}, \quad 0 \leq x \leq 1$$

a) ¿Para qué valores de ϑ_1 y ϑ_2 es $f(x)$ una función de densidad correctamente definida?

b) Determinar los valores de ϑ_1 y ϑ_2 sabiendo que $E[X] = 0,7$ y $\text{var}(X) = 0,06$

Solución:

a) Para que $f(x)$ esté bien definida, debe verificar que $\int_0^1 f(x) dx = 1 \quad \forall \vartheta$. Entonces:

$$\begin{aligned} \int_0^1 \frac{3\vartheta_1 x^2 + 2\vartheta_2 x + 1}{\vartheta_1 + \vartheta_2 + 1} dx &= \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left[3\vartheta_1 \frac{x^3}{3} + 2\vartheta_2 \frac{x^2}{2} + x \right]_0^1 = \\ &= \frac{1}{\vartheta_1 + \vartheta_2 + 1} (\vartheta_1 + \vartheta_2 + 1) = 1 \quad \forall \vartheta_1, \vartheta_2, (\vartheta_1 + \vartheta_2 + 1 \neq 0) \end{aligned}$$

Por tanto esta función está definida para cualesquiera valores de ϑ_1 y ϑ_2 , siempre y cuando $\vartheta_1 + \vartheta_2 + 1 \neq 0$.

b) Tenemos:

$$\begin{aligned} E[X] &= \int_0^1 x f(x) dx = \int_0^1 x \frac{3\vartheta_1 x^2 + 2\vartheta_2 x + 1}{\vartheta_1 + \vartheta_2 + 1} dx = \\ &= \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left[3\vartheta_1 \frac{x^4}{4} + 2\vartheta_2 \frac{x^3}{3} + \frac{x^2}{2} \right]_0^1 = \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left(\frac{3}{4}\vartheta_1 + \frac{2}{3}\vartheta_2 + \frac{1}{2} \right) \end{aligned}$$

$$\begin{aligned} E[X^2] &= \int_0^1 x^2 f(x) dx = \int_0^1 x^2 \frac{3\vartheta_1 x^2 + 2\vartheta_2 x + 1}{\vartheta_1 + \vartheta_2 + 1} dx = \\ &= \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left[3\vartheta_1 \frac{x^5}{5} + 2\vartheta_2 \frac{x^4}{4} + \frac{x^3}{3} \right]_0^1 = \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left(\frac{3}{5}\vartheta_1 + \frac{1}{2}\vartheta_2 + \frac{1}{3} \right) \end{aligned}$$

$$\text{var}(X) = E[X^2] - (E[X])^2$$

c) Como $E[X] = 0,7$ y $\text{var}(X) = 0,06$, sustituyendo en las expresiones anteriores:

$$\left. \begin{aligned} \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left(\frac{3}{4}\vartheta_1 + \frac{2}{3}\vartheta_2 + \frac{1}{2} \right) &= 0,7 \\ \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left(\frac{3}{5}\vartheta_1 + \frac{1}{2}\vartheta_2 + \frac{1}{3} \right) - 0,7^2 &= 0,06 \end{aligned} \right\}$$

$$\left. \begin{aligned} \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left(\frac{3}{4}\vartheta_1 + \frac{2}{3}\vartheta_2 + \frac{1}{2} \right) &= 0,7 \\ \frac{1}{\vartheta_1 + \vartheta_2 + 1} \left(\frac{3}{5}\vartheta_1 + \frac{1}{2}\vartheta_2 + \frac{1}{3} \right) &= 0,55 \end{aligned} \right\}$$

$$\left. \begin{aligned} \left(\frac{3}{4}\vartheta_1 + \frac{2}{3}\vartheta_2 + \frac{1}{2} \right) &= 0,7 (\vartheta_1 + \vartheta_2 + 1) \\ \left(\frac{3}{5}\vartheta_1 + \frac{1}{2}\vartheta_2 + \frac{1}{3} \right) &= 0,55 (\vartheta_1 + \vartheta_2 + 1) \end{aligned} \right\}$$

Multiplicamos por 12 la primera ecuación y por 30 la segunda (para eliminar las fracciones):

$$\left. \begin{aligned} (9\vartheta_1 + 8\vartheta_2 + 6) &= 8,4 (\vartheta_1 + \vartheta_2 + 1) \\ (18\vartheta_1 + 15\vartheta_2 + 10) &= 16,5 (\vartheta_1 + \vartheta_2 + 1) \end{aligned} \right\}$$

$$\left. \begin{aligned} (9 - 8,4) \vartheta_1 + (8 - 8,4) \vartheta_2 &= 8,4 - 6 \\ (18 - 16,5) \vartheta_1 + (15 - 16,5) \vartheta_2 &= 16,5 - 10 \end{aligned} \right\}$$

$$\left. \begin{aligned} 0,6\vartheta_1 - 0,4\vartheta_2 &= 2,4 \\ 1,5\vartheta_1 - 1,5\vartheta_2 &= 6,5 \end{aligned} \right\}$$

Multiplicamos la primera ecuación por 5 y la segunda por -2:

$$\left. \begin{aligned} 3\vartheta_1 - 2\vartheta_2 &= 12 \\ -3\vartheta_1 + 3\vartheta_2 &= -13 \end{aligned} \right\}$$

Sumamos ambas ecuaciones:

$$\vartheta_2 = -1$$

De la primera ecuación:

$$3\vartheta_1 = 12 + 2\vartheta_2 = 12 + 2 \cdot (-1) = 10 \Rightarrow \vartheta_1 = \frac{10}{3}$$

El siguiente código R nos permite dibujar la función de densidad para estos valores de ϑ_1 y ϑ_2 :

$$f(x) = \frac{3\frac{10}{3}x + 2(-1)x + 1}{-1 + \frac{10}{3} + 1} = \frac{30x^2 - 6x + 3}{10} = 3x^2 - \frac{3}{5}x + \frac{3}{10}$$

```
> f=function(x){ifelse(x<0|x>1,0,3*x^2-0.6*x+3/10)}
> x=seq(-0.2,1.2,length=500)
> plot(x,f(x),type="l",col="red",lwd=2)
> abline(v=c(0,1),lty=2)
>
```

