Soluciones a ejercicios seleccionados del Práctico 2.

El siguiente texto es un borrador y está sujeto a correcciones en clase.

Ejercicio 1

La función de distribución de una variable aleatoria X está dada por:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x < 1 \\ \frac{2}{3} & 1 \leq x < 2 \\ \frac{11}{12} & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

a) Dar el gráfico de F y hallar:

- i) $P\left(X > \frac{1}{2}\right)$
- ii) $P(X = 1)$
- iii) $P(X < 3)$,
- iv) $P(2 < X \leq 4)$.

b) ¿Es X una variable aleatoria discreta?

Solución: Para a) utilizamos la Proposición 1. Para i),

$$\begin{aligned} P\left(X > \frac{1}{2}\right) &= 1 - P\left(X \leq \frac{1}{2}\right) \\ &= 1 - F\left(\frac{1}{2}\right) = \frac{3}{4} \end{aligned}$$

Para ii), notar que

$$P(X = 1) = F(1) - \lim_{y \to 1^-} F(y) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$$

Para iii), notar que $P(X < 3) + P(X = 3) = F(3)$, de donde

$$P(X < 3) = F(3) - P(X = 3) = F(3) - [F(3) - \lim_{x \to 3^-} F(x)] = \lim_{x \to 3^-} F(x) = \frac{11}{12}$$

Finalmente

$$\begin{aligned} P(2 < X \leq 4) &= F(4) - F(2) \\ &= \frac{1}{12} \end{aligned}$$

Nota: recordar la importancia de graficar F y de señalar la relación entre las probabilidades calculadas y el gráfico, tal como se hizo en clase.

Ejercicio 2

Probar que para una variable aleatoria X con función de distribución F se cumplen las propiedades siguientes:

a) $\forall x \in \mathbb{R} : P(X > x) = 1 - F(x)$.

b) $\forall x, y \in \mathbb{R}$: si $x < y$ entonces $P(x < X \leq y) = F(y) - F(x)$.

Solución: Para probar a), fijemos $x \in \mathbb{R}$. Notar que si $A = \{\omega \in \Omega : X(\omega) > x\}$, entonces

$$P(X > x) = P(A) = 1 - P(A^c) = 1 - P(\{\omega \in \Omega : X(\omega) \leq x\}) = 1 - F(x)$$

donde hemos utilizado que $P(A) + P(A^c) = 1$ y la definición de función de distribución.

Para probar b), fijemos $x < y$ y definamos

$$A = \{\omega \in \Omega : X(\omega) \leq x\} \text{ y } B = \{\omega \in \Omega : x < X(\omega) \leq y\}.$$

Estos eventos son disjuntos con lo cual:

$$P(A \cup B) = P(A) + P(B).$$

Tengamos en cuenta ahora que $A \cup B = \{\omega \in \Omega : X(\omega) \leq y\}$ y, en consecuencia, se tiene que $P(A \cup B) = F(y)$. Así

$$F(y) = P(A \cup B) = P(A) + P(B) = P(x < X \leq y) + F(x),$$

equivalentemente

$$F(y) - F(x) = P(x < X \leq y).$$

Ejercicio 3

Una moneda cargada tiene probabilidad $p \in (0, 1)$ de que al lanzarla resulte "cara". Se lanza cuatro veces la moneda y se define la variable aleatoria discreta X: "número de caras en los cuatro lanzamientos".

a) Para la variable aleatoria X:

   i) Hallar y graficar la función de probabilidad de masa y la función de distribución.

   ii) Si se sabe que la prob abilidad de que aparezcan cuatro caras es 3^{-4}, hallar p.

b) Computar, utilizando la función de distribución las siguientes probabilidades:

   i) Que al menos en un lanzamiento aparezca cara.

   ii) $P(2 < X \leq 4)$.

iii) $P(X \geq 2)$.

c) Establecer relaciones entre las probabilidades y las gráficas del inciso a).

Solución: Utilizando el espacio muestral y la función X, obtenemos la función de probabilidad de masa, $f_X(k) = P(X = k)$, del siguiente modo:

$$\begin{aligned}
P(X=0) &= P(\{(x,x,x,x)\}) = 1 \cdot q^4 \\
P(X=1) &= P(\{(c,x,x,x),\ldots(x,x,x,c)\}) = 4 \cdot p^3 q \\
P(X=2) &= \binom{4}{2} p^2 q^{4-2} = 6p^2q^2 \\
P(X=3) &= \binom{4}{3} p^3 q^{4-3} = 4p^3q^1 \\
P(X=4) &= \binom{4}{4} p^4 q^{4-4} = p^4
\end{aligned}$$

siendo $q = 1 - p$. Y, $f_X(k) = 0$ si $k \notin \{0,1,2,3,4\}$.

Luego F se define por partes, tal como lo hicimos en clase.

A partir de a)-ii) y observando que

$$P(X = 4) = p^4 = 3^{-4}$$

resulta que $p = 1/3$ y $q = 2/3$.

Computemos, las probabilidades para dos incisos de b).

i) La probabilidad de que al menos en un lanzamiento aparezca cara. Es decir

$$P(X \geq 1) = 1 - P(X = 0) = (2/3)^4 = 0.198$$

ii)
$$P(2 < X \leq 4) = P(X = 3) + P(X = 4) = 4(1/3)^3(2/3) + (1/3)^4 = 0.111.$$

Ejercicio 4

a) Se lanza un dado equilibrado y se define X la variable aleatoria discreta “resultado del lanzamiento”. Hallar $\mathrm{E}(X)$ y $\mathrm{VAR}(X)$.

b) Más generalmente, sea X una variable aleatoria discreta con rango $R_X = \{1, \ldots, M\}$ y tal que $P(X = i) = 1/M$ para cada $i \in R_X$. Hallar $\mathrm{E}(X)$.

Solución: Inciso a): si utilizamos que la función de probabilidad de masa es $f_X(x) = 1/6$ sobre el rango de X y la definición de esperanza, obtenemos

$$\mathrm{E}(X) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = 7/2.$$

Para la varianza

$$\text{VAR}(X) = \text{E}(X^2) - [\text{E}(X)]^= \frac{1}{6}91 - (7/2)^2 = 35/12$$

.

Para b), recordar la expresión de la suma de los primeros números naturales, con lo cual

$$\text{E}(X) = \frac{1}{M}(1 + 2 + 3 + \cdots + M) = \frac{1}{M}\frac{M(M+1)}{2} = \frac{M+1}{2}.$$

Intepretando $\text{E}(X)$ como centro de gravedad, ¿cuál es la diferencia si M es un número par o impar?.

Ejercicio 5

Consideremos un juego de preguntas en el que una persona recibe dos preguntas y debe decidir cuál responder primero. La Pregunta 1 se responderá correctamente con una probabilidad del 0.8 y la persona recibirá un premio de 1000, mientras que la Pregunta 2 se responderá correctamente con una probabilidad del 0.5 y la persona recibirá un premio de 2000. Si la primera pregunta intentada se responde incorrectamente, el juego termina, es decir, la persona no tiene permitido intentar la segunda pregunta. Si la primera pregunta se responde correctamente, la persona puede intentar la segunda pregunta. ¿Qué pregunta debe responderse primero para maximizar el valor esperado del dinero total del premio recibido?

Solución:

La respuesta no es evidente porque hay un desequilibrio. En efecto, intentar responder primero la Pregunta 2, la que lleva a un premio más valioso pero también con menos chances de acertar, conlleva el riesgo de no tener la oportunidad de intentar responder primero la Pregunta 1, con más chances de obtener algo.

El dinero total del premio recibido es una variable aleatoria X y calculemos el valor esperado $\text{E}(X)$ bajo los dos posibles órdenes de las preguntas.

Al espacio muestral hay que representarlo secuencialmente según si la Pregunta 1 es la primera respondida o la Pregunta 2 es la primera respondida. Ver Figura 1.

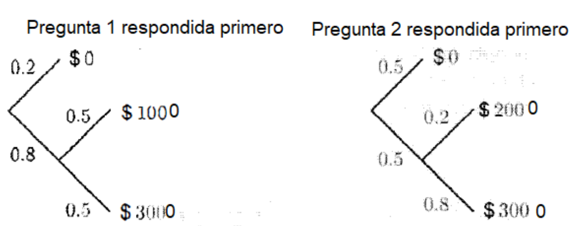

Figura 1: Descripción secuencial del espacio muestral

a) La Pregunta 1 es la primera espondida. Entonces

$$P(X = 0) = P(\text{ la pregunta 1 sea mal respondida}) = 0.2$$

$$\begin{aligned} P(X=1000) &= P(\text{ la pregunta 2 sea mal respondida dado que la 1 lo fue}) \\ &\cdot P(\text{ la pregunta 1 sea bien respondida }) = 0.5 \cdot 0.8 \end{aligned}$$

y

$$\begin{aligned} P(X=3000) &= P(\text{ la pregunta 2 sea bien respondida dado que la 1 lo fue}) \\ &\cdot P(\text{ la pregunta 1 sea bien respondida }) = 0.5 \cdot 0.8 \end{aligned}$$

Luego

$$\mathrm{E}(X) = 0 \cdot 0.2 + 1000 \cdot 0.5 \cdot 0.8 + 30000 \cdot 0.5 \cdot 0.8 = 1600 \text{ pesos}$$

b) La Pregunta 2 es la segunda respondida. Entonces

$$\begin{aligned} P(X=0) &= 0.5 \\ P(X=2000) &= 0.5 \cdot 0.2 \\ P(X=3000) &= 0.5 \cdot 0.8 \end{aligned}$$

y así

$$\mathrm{E}(X) = 2000 \cdot 0.5 \cdot 0.2 + 3000 \cdot 0.5 \cdot 0.8 = 1400 \text{ pesos.}$$

En consecuencia la estrategia más conveniente es responder primero la Pregunta 1.

Ejercicio 6

La temperatura de una ciudad se modela como una variable aleatoria con una media y una desviación estándar, ambas iguales a 10 grados Celsius. Un día se describe como "estándar" si la temperatura durante ese día se encuentra dentro de una desviación estándar de la media. ¿Cuál sería el rango de temperatura para un día estándar si la temperatura se expresara en grados Fahrenheit? (recordar que si X es la temperatura en grados Celsius, en grados Fahrenheit es $Y = 32 + 9X/5$).

Solución. Teniendo en cuenta que $Y = 32 + 9X/5$,

$$\mathrm{E}[Y] = 32 + 9\mathrm{E}[X]/5 = 32 + 18 = 50$$

y

$$\mathrm{VAR}(Y) = (9/5)^2 \mathrm{VAR}(X) = (9/5)^2 \cdot 100$$

Entonces el desvío estándar de Y es (9/5) 10=18. En consecuencia un día estándar en la escala Fahrenheit es uno para el cual la temperatura está en el rango [50-18,50+18]=[32, 68].

Ejercicio 7

Las Tablas 1 y 2 corresponden a funciones de probabilidad de masa de pares de vectores (X, Y) y (Z, T)

X/Y	0	2	4
0	1/9	1/9	1/9
1	1/9	1/9	1/9
2	1/9	1/9	1/9

Tabla 1: $f_{(X,Y)}$

Z/T	0	2	4
0	1/3	0	0
1	0	1/3	0
2	0	0	1/3

Tabla 2: $f_{(Z,T)}$

a) ¿Son X e Y variables aleatorias independientes? ¿Y Z y T son independientes?

b) ¿Qué valor asumen los coeficientes de correlación de ambos pares de variables?

Solución: Para la Tabla 1, es posible probar fácilmente que

$$f_X(x) = 1/3, \forall x = 0, 1, 2 \text{ y } f_Y(y) = 1/3, \forall y = 0, 2, 4$$

con lo cual

$$\forall x, y : f_{(X,Y)}(x, y) = f_X(x) f_Y(y).$$

En consecuencia, X e Y son independientes, con lo cual el coeficiente de correlación es cero.

En relación a la Tabla 2, es fácil mostrar que $E(T) = 2$, $\text{VAR}(T) = 8/3$, $\text{E}(Z) = 1$, $\text{VAR}(Z) = 2/3$, $\text{E}(ZT) = 10/3$, $\text{COV}(Z, T) = 4/3$

Luego el coeficiente es

$$\frac{4/3}{\sqrt{(2/3)8/3}} = 1$$

Pero también puede deducirse que es uno si se observa que

$$P(T = 2Z) = P(Z = 0, T = 0) + P(Z = 1, T = 2) + P(Z = 2, T = 4) = 3\frac{1}{3} = 1.$$

Ejercicio 8

Sean X e Y variables aleatorias discretas, a y b constantes. Probar que:

a) $\text{VAR}(aX + b) = a^2 \text{VAR}(X)$.

b) Si X e Y son variables aleatorias independientes entonces estń no correlacionadas y además $\text{VAR}(X + Y) = \text{VAR}(X) + \text{VAR}(Y)$.

Ejercicio 9

Suponga que sólo el 20 % de los automovilistas se detienen por completo en un cruce donde hay un semáforo con luz roja intermitente en todas las direcciones, cuando no haya otros automóviles visibles.

a) Hallar la probabilidad de que de 20 automovilistas, seleccionados al azar, que lleguen al cruce en estas condiciones:

   i) a lo sumo 5 se detengan por completo.
   ii) exactamente 5 se detengan por completo.
   iii) por lo menos 5 se detengan por completo.

b) ¿Cuántos, de los siguientes 20 automovilistas, se espera que se detengan por completo?

Solución: Sea X la variable "número de automovilistas que se detienen por completo en un total de 20 automovilistas". Luego

$$X \sim binomial(20, 1/5)$$

y

$$f_X(k) = \binom{20}{k}(1/5)^k(4/5)^{20-k},\ k = 0, \ldots, 20.$$

Para computar las probabilidades podemos utilizar R:

a) i)

$$P(X \leq 5) = \sum_{k=0}^{5}\binom{20}{k}(1/5)^k(4/5)^{20-k}.$$

Y en R

```
dbinom(0,20, 1/5)+ dbinom(1,20, 1/5)+ dbinom(2,20, 1/5)+
dbinom(3,20, 1/5)+dbinom(4,20, 1/5)+dbinom(5,20, 1/5)
[1] 0.8042078
```

O bien utilizamos la función de distribución para calcular $F_X(5)$:

```
> pbinom(5,20, 1/5)
[1] 0.8042078
```

ii) Aquí computamos $P(X = 5)$

```
> dbinom(5,20, 1/5)
[1] 0.1745595
```

iii) Queremos hallar

$$P(X \geq 5) = 1 - P(X < 5) = 1 - F_X(4)$$

```
> 1-pbinom(4,20, 1/5)
[1] 0.3703517
```

b) ¿Cuántos, de los siguientes 20 automovilistas, se espera que se detengan por completo? Nos preguntan cuál es la esperanza de la variable.

$$\mathrm{E}(X) = 20\frac{1}{5} = 4$$

Ejercicio 10

Un paquete de software consta de 12 programas, cinco de los cuales deben actualizarse. Si 4 programas son elegidos al azar para la prueba,

a) ¿Cuál es la probabilidad de que al menos dos de ellos deban actualizarse?

b) ¿Cuál es el número esperado de programas, de los cuatro elegidos, que deben ser actualizados?

Solución: Sea X la variable "número de programas que deben actualizarse". Luego

$$X \sim \text{hipergeométrica}(5, 4, 12)$$

con rango

$$R_X = \{k \in \mathbb{N} : \max\{0, n - N + D\} \leq k \leq \min\{n, D\}\} = \{0, 1, 2, 3, 4\} \tag{1}$$

y función de probabilidad de masa

$$f_X(k) = \frac{\binom{5}{k}\binom{7}{4-k}}{\binom{12}{4}}, k \in R_X.$$

a)

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - (P(X = 0) + P(X = 1)) = 1 - \sum_{k=0}^{1} \frac{\binom{5}{k}\binom{7}{4-k}}{\binom{12}{4}}$$

Y en R

```
>1 -(dhyper(0,5,7,4)+ dhyper(1,5,7,4))
[1] 0.5757576
```

O bien utilizamos la función de distribución para calcular $F_X(1)$:

```
> 1-phyper(1,5,7,4)
[1] 0.5757576
```

b) La esperanza esperanza de X es:

$$\mathrm{E}(X) = 4\ \frac{5}{12} = \frac{20}{12}$$

Ejercicio 11

Supongamos que el número de errores tipográficos sobre una única página de un libro tiene una distribución de Poisson de parámetro $\lambda = 1$. Calcular la probabilidad de que:

a) exista al menos un error en esa página.

b) exista a lo sumo 2 errores en esa página.

Solución: Sea X la variable "número de errores tipográficos en una única página". Luego

$$X \sim \text{Poisson}(1)$$

y

$$f_X(k) = \frac{e^{-1}\ 1^k}{k!}, \text{con } k = 0, 1, ...$$

a)
$$P(X \geq 1) = 1 - F(0) = 1 - \frac{e^{-1}1^0}{0!} = 1 - \frac{1}{e} = 0.6321206$$

Y en R

```
>1 - dpois(0, 1)
[1] 0.6321206
```

O bien utilizamos la función de distribución para calcular $F_X(0)$:

```
> 1 - ppois(0, 1)
[1] 0.6321206
```

b)
$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \sum_{k=0}^{2} \frac{e^{-1}\ 1^k}{k!} = 0.919$$

Y en R

```
> dpois(0,1) + dpois(1,1) + dpois(2,1)
[1] 0.9196986
```

O con la función de distribución para calcular $F_X(2)$:

```
> ppois(2, 1)
[1] 0.9196986
```

Ejercicio 12

El número de visitas realizadas en un día entre semana en una determinada página web se decide modelizar por una variable de Poisson de media 8.

a) ¿Cuál es la probabilidad de que en un día se reciban más de 4 visitas? Y ¿entre siete y diez visitas (ambos valores incluidos)?

b) ¿Qué valor asume el número medio del número de visitas?.

Solución: Sea X la variable "número de visitas realizadas en un día entre semana en una página web".

Aquí

$$X \sim Poisson(8)$$

y

$$f_X(k) = \frac{e^{-8}\ 8^k}{k!}, \text{con } k = 0, 1, ...$$

a)

$$P(X > 4) = 1 - P(X \leq 4) = 1 - F(4) = 1 - \sum_{k=0}^{4} \frac{e^{-8}\ 8^k}{k!} = 0.9003676$$

Y en R

```
> 1 - (dpois(0, 8) + dpois(1, 8) + dpois(2, 8) + dpois(3, 8) + dpois(4, 8))
[1] 0.9003676
```

O bien con la función de distribución para calcular $F_X(4)$:

```
> 1 - ppois(4, 8)
[1] 0.9003676
```

b)

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \sum_{k=0}^{2} \frac{e^{-1}\ 1^k}{k!} = 0.919$$

Y en R

```
> dpois(0,1) + dpois(1,1) + dpois(2,1)
[1] 0.9196986
```

O con la función de distribución para calcular $F_X(2)$:

```
> ppois(2, 1)
[1] 0.9196986
```

Para responder la segunda parte de la pregunta, planteamos

$$\begin{aligned} P(7 \leq X \leq 10) &= P(7 < X \leq 10) + P(X = 7) \\ &= F(10) - F(7) + f(7) \\ &= 0.3629 \end{aligned}$$

ó también se puede pensar como

$$P(6 < X \leq 10) = F(10) - F(6) = 0.3629$$

Y en R

```
> ppois(10,8) - ppois(7,8) + dpois(7,8) #primera resolución
[1] 0.5025115
> ppois(10,8) - ppois(6,8) #segunda resolución
[1] 0.5025115
```

Ejercicio 13

Una persona usuaria de una computadora intenta recuperar su contraseña. La persona sabe que puede ser una de 4 posibles contraseñas de las que dispone. Prueba sus contraseñas hasta que encuentra la correcta. Sea X : "número de contraseñas incorrectas probadas hasta encontrar la correcta". Hallar $E(X)$ y $\text{VAR}(X)$.

Solución: Sea X la variable "número de contraseñas incorrectas probadas hasta encontrar la correcta"

$$X \sim \text{geométrica}(p)$$

con $p = \frac{1}{4}$ y $k = 1, 2, ...$

Luego la esperanza y varianza para esta variable viene dada por

$$\mathrm{E}(X) = \frac{1}{p} = 4$$

y

$$\text{VAR}(X) = \frac{q}{p^2} = \frac{\frac{3}{4}}{(\frac{1}{4})^2} = 12.$$

Nota. Aquí estamos haciendo un supuesto que es poco real y es que las pruebas son independientes unas de otras.