Trabajo Práctico Unidad 6
Probabilidad y Estadística

Alumnas:

Cipullo, Inés

Sullivan, Katherine

Universidad Nacional de Rosario

2021

Ejercicio 1

Sean las siguientes variables aleatorias:

- G_i: "resultado de la apuesta i"
- G_t: "suma de los resultados de las primeras 50 apuestas"

Tenemos que:

- $P(G_i = 0) = 0.9$
- $P(G_i = 1) = 0.1$
- $G_t = \sum_{i=1}^{50} G_i$

Considerando que el resultado de una apuesta es independiente del resultado del resto de apuestas y que $E(G_i)$ y $V(G_i)$ son finitas para todo $i = 1, ..., 50$ podemos aplicar el teorema central del límite y concluir que la variable G_t tiene una distribución aproximadamente normal de media μ y varainza σ^2, donde:

- $\mu = E(G_t) = 50 \cdot E(G_i) = 50 \cdot 0.1 = 5$
- $\sigma^2 = V(G_t) = 50 \cdot V(G_i) = 50 \cdot 0.09 = 4.5$

Por lo tanto, $G_t \sim N(5, 4.5)$

Sea S la cantidad de dinero con que la persona queda luego de las 50 apuestas. Queremos ver que $S \geq 0$.

Recordando que se apuesta una cantidad de dinero $c > 0$ que se pierde en caso de que $G_i = 0$ y que si $G_i = 1$ se ganan $5c$, tenemos que:

$$\begin{aligned} S &= 5 \cdot c \cdot G_t - c \cdot (50 - G_t) \\ &= 5 \cdot c \cdot G_t - (c \cdot 50 + c \cdot G_t) \\ &= 5 \cdot c \cdot G_t + c \cdot G_t - c \cdot 50 \\ &= 6 \cdot c \cdot G_t - c \cdot 50 \end{aligned}$$

Queremos que $S \geq 0 \implies 6c\ G_t - 50c \geq 0 \implies G_t \geq \frac{50}{6}$

Entonces, en síntesis lo que buscamos es $P(G_t \geq 8.33)$.

Para obtener este valor estandarizamos la variable:

$Z = \frac{G_t - \mu}{\sigma} = \frac{G_t - 5}{\sqrt{4.5}} = \frac{G_t - 5}{2.12}$

Y por lo tanto,

$P(G_t \geq 8.33) = P(Z \geq 1.57) = 1 - P(Z < 1.57) = 1 - 0.9418 = 0.0582$

Ejercicio 2

Sean las variables aleatorias:

- A: "longitud de bloques de tipo A"
- B: "longitud de bloques de tipo B"
- X: "suma de la longitud de 20 bloques de tipo A"
- Y: "suma de la longitud de 30 bloques de tipo B"
- R: "longitud del recipiente"

Contamos con los siguientes datos:

- $E(A) = 1.95$ y $\sqrt{V(A)} = 0.01 \implies V(A) = 0.0001$
- $E(B) = 0.83$ y $\sqrt{V(B)} = 0.02 \implies V(B) = 0.0004$
- $E(R) = 65$ y $\sqrt{V(R)} = 0.5 \implies V(R) = 0.7071$

Como A y B son variables aleatorias independientes con media y varianza finita, por el Teorema Central del Límite sabemos que las variables aleatorias X e Y tienen una distribución aproximadamente normal, donde $E(X) = 39$, $V(X) = 0.002$, $E(Y) = 24.9$ y $V(Y) = 0.012$.

Buscamos la probabilidad de que el ensamble formado por 20 bloques de tipo A y 30 bloques de tipo B entre en un recipiente. Podemos plantearlo como sigue:

$$P(X + Y \leq R) = P(X + Y - R \leq 0)$$

Como R tiene distribución normal y X e Y tienen distribución aproximadamente normal, por la propiedad reproductica de la distribucín normal:

$$X + Y + (-1) \cdot R \sim N(E(X) + E(Y) - E(R), \sqrt{V(X) + V(Y) - V(R)})$$

Definimos entonces una nueva variable aleatoria: $S = X + Y - R$, donde

$$S \sim N(-1.1, -0.6931)$$

En síntesis, queremos ver que $P(S \leq 0)$. Para obtener ese valor, debemos estandarizar la variable:

$Z = \frac{S-\mu}{\sigma^2} = \frac{S+1.1}{-0.6931}$

Buscamos en la tabla de distribución normal estándar

$$P(S \leq 0) = P(Z \leq -1.587) = 0.0559$$

Ejercicio 3

Sean las variables aleatorias:

- X: "número de defectos tipo D1 que presenta una pieza"
- Y: "número de defectos tipo D2 que presenta una pieza"

Contamos con los siguientes datos:

1. $E(X) = 0.3$ y $V(X) = 0.21$
2. $E(Y) = 0.8$ y $V(Y) = 0.56$
3. 20 % de piezas tienen 2 defectos tipo D2
4. 15 % de las piezas tienen 1 defecto tipo D1 y ninguno tipo D2
5. 50 % de las piezas que no tienen defectos tipo D1, tienen 1 defecto tipo D2

a)

X \Y	**0**	**1**	**2**	$p_X(x)$
0	0.25	0.35	0.1	0.7
1	0.15	0.05	0.1	0.3
$p_Y(y)$	0.4	0.4	0.2	1

Cálculos realizados:

- $p(1, 0) = 0.15$ (por lo especificado en el dato 4)
- $p_Y(2) = 0.2$ (por lo especificado en el dato 3)

- $E(X) = 0 \cdot p_X(0) + 1 \cdot p_X(1) = p_X(1) \implies p_X(1) = 0.3$
- $p_X(0) + p_X(1) = p_X(0) + 0.3 = 1 \implies p_X(0) = 0.7$
- $p(0,1) = p_X(0) \cdot 0.5 = 0.7 \cdot 0.5 = 0.35$ (por lo especificado en el dato 5)
- $E(Y) = 0 \cdot p_Y(0) + 1 \cdot p_Y(1) + 2 \cdot p_Y(2) = p_Y(1) + 2 \cdot 0.2 \implies p_Y(1) = 0.4$
- $p_Y(0) + p_Y(1) + p_Y(2) = p_Y(0) + 0.4 + 0.2 = 1 \implies p_Y(0) = 0.4$
- $p_Y(0) = p(0,0) + p(1,0) \implies 0.4 = p(0,0) + 0.15 \implies p(0,0) = 0.25$
- $p_X(0) = p(0,0) + p(0,1) + p(0,2) \implies 0.7 = 0.25 + 0.35 + p(0,2) \implies p(0,2) = 0.1$
- $p_Y(1) = p(0,1) + p(1,1) \implies 0.4 = 0.35 + p(1,1) \implies p(1,1) = 0.5$
- $p_Y(2) = p(0,2) + p(1,2) \implies 0.2 = 0.1 + p(1,2) \implies p(1,2) = 0.1$

b)

Las variables aleatorias X e Y son independientes si

$$p(x,y) = p_X(x) \cdot p_Y(y) \; \forall \; (x,y) \in R_{(X \times Y)}$$

Notar que en este caso $R_{(X \times Y)} = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2)\}$.

Pero tenemos que $p(0,0) = 0.25$ y $p_X(0) \cdot p_Y(0) = 0.7 \cdot 0.4 = 0.28$, por lo tanto X e Y no son independientes.

Calculamos entonces el Coeficiente de Correlación:

$$\begin{aligned} \rho_{xy} &= \frac{Cov(X,Y)}{\sqrt{V(X) \cdot V(Y)}} = \frac{E(X \cdot Y) - E(X) \cdot E(Y)}{\sqrt{V(X) \cdot V(Y)}} \\ &\overset{(*)}{=} \frac{0.25 - 0.3 \cdot 0.8}{\sqrt{0.21 \cdot 0.56}} = 0.0292 \end{aligned}$$

donde $(*)$ hace referencia al siguiente cálculo:

$$\begin{aligned} E(X \cdot Y) &= \sum_{(x,y) \in R_{(X \times Y)}} x \cdot y \cdot p(x,y) \\ &= p(1,1) + 2 \cdot p(1,2) = 0.05 + 2 \cdot 0.1 = 0.25 \end{aligned}$$

c)

La fórmula de probabilidad condicional es como sigue:

$$P(Y = y/X = x) = \frac{p(x,y)}{p_X(x)}$$

Queda entonces que

$$P(Y = 2/X = 0) = \frac{p(0,2)}{p_X(0)} = \frac{0.1}{0.7} = 0.1428$$

Este resultado refleja que dada una pieza que no tiene defectos tipo D1, la probabilidad de que tenga 2 defectos tipo D2 es 0.1428. Otra forma de verlo es pensar que el 14.28 % de las piezas que no tienen defectos D1, tienen 2 defectos D2.

d)

Buscamos la esperanza y la varianza del costo de reparación por pieza, siendo que un defecto tipo D1 tiene un costo de reparación de $3 y uno tipo D2 tiene un costo de reparación de $4.

Definimos una función sobre las variables aleatorias: $Z = H(X,Y)$, donde $H(x,y) = 3 \cdot x + 4 \cdot y$.

Luego, basta con calcular $E(Z)$ y $V(Z)$.

$$
\begin{aligned}
E(Z) &= \sum_{(x,y)\in R_{(X\times Y)}} H(x,y)\cdot p(x,y) \\
&= 4\cdot p(0,1) + 8\cdot p(0,2) + 3\cdot p(1,0) + 7\cdot p(1,1) + 11\cdot p(1,2) \\
&= 4\cdot 0.35 + 8\cdot 0.1 + 3\cdot 0.15 + 7\cdot 0.05 + 11\cdot 0.1 \\
&= 1.4 + 0.8 + 0.45 + 0.35 + 1.1 = 4.1
\end{aligned}
$$

$$
\begin{aligned}
V(Z) &= \sum_{(x,y)\in R_{(X\times Y)}} [H(x,y) - E(Z)]^2 \cdot p(x,y) \\
&= (-4.1)^2\cdot p(0,0) + (-0.1)^2\cdot p(0,1) + 3.9^2\cdot p(0,2) + \\
&\quad + (-1.1)^2\cdot p(1,0) + 2.9^2\cdot p(1,1) + 6.9^2\cdot p(1,2) \\
&= 16.81\cdot 0.25 + 0.01\cdot 0.35 + 15.21\cdot 0.1 + 1.21\cdot 0.15 + 8.41\cdot 0.05 + 47.61\cdot 0.1 \\
&= 4.2025 + 0.0035 + 1.521 + 0.1815 + 0.4205 + 4.761 = 11.09
\end{aligned}
$$