Estadística Descriptiva e Introducción a la Probabilidad. Relaciones de prácticas

ELÍAS MONGE SÁNCHEZ
DANIEL MORÁN SÁNCHEZ
JESÚS MUÑOZ VELASCO

Mayo 2023

Índice

5. Relación 5

5.1. Ejercicio 1:

1. Sea X una variable aleatoria con función masa de probabilidad P (X = i) = k_i; i = 1, ..., 20.
a) Determinar el valor de k, la función de distribución y las siguientes probabilidades:
$P(X = 4), P(X < 4), P(3 \leq X \leq 10), P(3 < X \leq 10), P(3 < X < 10)$.
b) Supongamos que un jugador gana 20 monedas si al observar esta variable obtiene un valor menor que 4, gana 24 monedas si obtiene el valor 4 y, en caso contrario, pierde una moneda. Calcular la ganancia esperada del jugador y decir si el juego le es favorable.

$\sum_{i=1}^{20} P(X = i) = \sum_{i=1}^{20} ki = 1 \xrightarrow{\sum_{i=1}^{20} i=210} 210k = 1 \rightarrow k = \frac{1}{210}$

$$F(x) = \begin{cases} 0 & \text{si } x < 1 \\ \sum_{i=0}^{E(x)} P(X = j) = \sum_{i=1}^{E(x)} \frac{1}{210} * i & \text{si } 1 \leq x < 20 \\ 1 & \text{si } x \geq 20 \end{cases}$$

Siendo E(x) la parte entera de x

a)

- $P(X = 4) = \frac{4}{210} = \frac{2}{105} = 0.019048$
- $P(X < 4) = P(X \leq 3) = \sum_{i=1}^{3} \frac{1}{210} * i = \frac{1+2+3}{210} = \frac{6}{210} = \frac{1}{35}$
- $P(3 \leq X \leq 10) = \sum_{i=3}^{10} \frac{1}{210} * i = \frac{3+4+5+6+7+8+9+10}{210} = \frac{26}{105} = 0.24761904761$
- $P(3 < X \leq 10) = P(4 \leq X \leq 10) = \sum_{i=4}^{10} \frac{1}{210} * i = \frac{4+5+6+7+8+9+10}{210} = \frac{7}{30} = 0.23333333333$
- $P(3 \leq X < 10) = P(4 \leq X \leq 9) = \sum_{i=4}^{9} \frac{1}{210} * i = \frac{4+5+6+7+8+9}{210} = \frac{39}{210} = 0.18571428571$

b)
Tomemos una nueva variable aleatoria Y, que será las monedas ganadas en función de X

$$y = g(x) = \begin{cases} \text{-1} & \text{si } 4 < x \leq 20 \\ 24 & \text{si } x = 4 \\ 20 & \text{si } 1 < x < 4 \end{cases}$$ y calculamos los valores de que Y en tome todos sus posibles valores en funcion de X

- $P(Y = -1) = P(X > 4) = 1 - P(X \leq 4) = 1 - P(X < 4) - P(X = 4) =$
  $1 - 0.02857142857 - 0.01904761904 = 0.95238095239$
- $P(Y = 24) = P(X = 4) = \frac{4}{210} = \frac{2}{105} = 0,01904761904$
- $P(Y = 20) = P(X < 4) = P(X \leq 3) = F_X(3) = \frac{1+2+3}{210} = \frac{1}{35} = 0.02857142857$

calculemos entonces la esperanza matemática:$(y_1 = -1, y_2 = 24, y_3 = 20)$

$$E[Y] = \sum_{i=1}^{n=3} y_i P(Y = y_i) = -1 * 0.95238095239 + \frac{24 * 2}{105} + \frac{20 * 1}{35} = 0.07619047618 \quad (1)$$

y como tenemos una esperanza de Y > 0, el juego es favorable al jugador.

5.2. Ejercicio 2:

Sea X el número de bolas blancas obtenidas al sacar dos de una urna con 10 bolas de las que 8 son blancas. Calcular:

a) Función masa de probabilidad y función de distribución.

La probabilidad de sacar x bolas blancas se calcula mediante la siguiente expresión:

$$P[X = x] = \frac{\binom{8}{x}\binom{2}{2-x}}{\binom{10}{2}} \qquad x = 0, 1, 2$$

Que es la función masa de probabilidad (en el tema 6 se ve que es una distribución hipergeométrica de parámetros $N = 10,\ N_1 = 8,\ n = 2$).

De la función anterior obtenemos:

$$P[X = 0] = \frac{\binom{8}{0}\binom{2}{2}}{\binom{10}{2}} = \frac{1}{45}$$
$$P[X = 1] = \frac{\binom{8}{1}\binom{2}{1}}{\binom{10}{2}} = \frac{16}{45}$$
$$P[X = 2] = \frac{\binom{8}{2}\binom{2}{0}}{\binom{10}{2}} = \frac{28}{45}$$

Y entonces la función de distribución $F_X(x) = P[X \leq x]$ queda:

$$F_X(x) = P[X \leq x] = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{45} & \text{si } 0 \leq x < 1 \\ \frac{17}{45} & \text{si } 1 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

b) Media, mediana y moda, dando la interpretación de cada una de estas medidas.

$Mo = 2 \Rightarrow$ La mayoría de veces se sacarán 2 bolas blancas
$Me = 2 \Rightarrow$ La probabilidad de sacar 2 o menos como de sacar 2 o más (2) es mayor o igual a 1/2

$$E[X] = \sum_{i=0}^{2} x_i P[X = x_i] = 0 \times \frac{1}{45} + 1 \times \frac{16}{45} + 2 \times \frac{28}{45} = 1.6$$

Al repetir el experimento varias veces, de media se sacarán aproximadamente 1.6 bolas blancas, por ejemplo si lo repetimos 10 veces lo esperado será sacar 16 bolas blancas en total.

c) Intervalo intercuartílico, especificando su interpretación.

$Q_1 = 1 \qquad Q_3 = 2 \Rightarrow R_I = 1$

El recorrido intercuartílico es 1 en comparación al recorrido que es 2. Esto parece indicar que la distribución es homogénea pero no es cierto, pues los cuartiles Q_1 y Q_3 están en la parte derecha de la distribución, lo que indica que la distribución está desplazada a la derecha.

5.3. Ejercicio 3:

El número de lanzamientos de una moneda hasta salir cara es una variable aleatoria con distribución $P(X = x) = 2^{-x};\ x = 1, 2, ...$

a) Probar que la función masa de probabilidad está bien definida.

b) Calcular la probabilidad de que el número de lanzamientos necesarios para salir cara esté entre 4 y 10.

c) Calcular los cuartiles y la moda de la distribución, interpretando los valores.

d) Calcular la función generatriz de momentos y, a partir de ella, el número medio de lanzamientos necesarios para salir cara y la desviación típica.

Solución:

a) Para que la función masa de probabilidad $P(X = x) = 2^{-x}$ esté bien definida basta con que se cumplan estas dos condiciones:

   - $0 \leq P(X = x) \leq 1\ \forall x \in \mathbb{N}$
   - $\sum_{x \in \mathbb{N}} P(X = x) = 1$

   Para la primera, como 2^{-x} es una función decreciente y positiva y $2^{-1} = \frac{1}{2} < 1$ podemos afirmar que $0 < 2^{-x} \leq \frac{1}{2}\ \forall x \in \mathbb{N}$

   Para la segunda tenemos que ver la siguiente serie converge a 1:

$$\sum_{x \in \mathbb{N}} 2^{-x} = \sum_{x \geq 0} 2^{-x} - 2^0 = \frac{1}{1 - \frac{1}{2}} - 1 = 2 - 1 = 1$$

   En efecto, por ser una serie geométrica, podemos comprobar que converge a 1.

b) La probabilidad de que se necesiten x lanzamientos para que salga cara es la que nos da la función masa de probabilidad $P(X = x) = 2^{-x}$. Por lo que la probabilidad que nos piden es:

$$P(4 \leq X \leq 10) = \sum_{n=4}^{10} P(X = n) = \sum_{n=4}^{10} 2^{-x} = \frac{1}{2^4} + \frac{1}{2^5} + \frac{1}{2^6} + \frac{1}{2^7} + \frac{1}{2^8} + \frac{1}{2^9} + \frac{1}{2^{10}} = 0.124$$

c) Los 3 cuartiles $Q_1,\ Q_2,\ Q_3$ dividen la función masa de probabilidad en 4 partes, de manera que:

$$P(X \leq Q_1) \geq \frac{1}{4} \Rightarrow Q_1 = 1$$
$$P(X \leq Q_2) \geq \frac{1}{2} \Rightarrow Q_2 = 1$$
$$P(X \leq Q_3) \geq \frac{3}{4} \Rightarrow Q_3 = 2$$

   Que los cuartiles estén tan a la izquierda de la distribución significa que la distribución es muy asimétrica y que la probabilidad de que salga cara en alguna de las primeras tiradas es alta en comparación con que tengamos que tirar la moneda muchas veces hasta que salga cara.
   La moda en variable discreta se define como:

$$P(X = Mo) = max\{P(X = x) : x \in E\}$$

   Siendo E $= \mathbb{N}$ en este caso. Como la función masa de probabilidad es decreciente, $Mo = 1$. Esto significa que si hacemos un número muy elevado de experimentos, en la mayoría de ellos solo habríamos tirado la moneda una vez.

d) La función generatriz de momentos se calcula como:

$$M_X(t) = E[e^{tX}] = \sum_{x\in\mathbb{N}} e^{tx} P(X = x) = \sum_{x\in\mathbb{N}} \frac{e^{tx}}{2^x} = \sum_{x\in\mathbb{N}} \left(\frac{e^t}{2}\right)^x = \frac{1}{1-\frac{e^t}{2}} - \left(\frac{e^t}{2}\right)^0 = \frac{1}{1-\frac{e^t}{2}} - 1 =$$

$$= \frac{\frac{e^t}{2}}{1-\frac{e^t}{2}} = \frac{e^t}{2-e^t} \ \forall t < ln2$$

Derivando y evaluando en el 0, podemos obtener el momento de orden 1, que es la esperanza:

$$M_X'(t) = \frac{e^t(2-e^t)+e^{2t}}{(2-e^t)^2} = \frac{2e^t}{(2-e^t)^2}$$

$$\underline{E[X]} = M_X'(0) = \frac{2}{1} = \underline{2}$$

Derivando otra vez y evaluando en el 0, podemos obtener el momento de orden 2, con el que podemos obtener la varianza y la desviación típica:

$$M_X''(t) = \frac{2e^t(2-e^t)^2+4e^{2t}(2-e^t)}{(2-e^t)^4} = \frac{2e^t(2-e^t)+4e^{2t}}{(2-e^t)^3} = \frac{4e^t+2e^{2t}}{(2-e^t)^3} = \frac{2e^t(2+e^t)}{(2-e^t)^3}$$

$$E[X^2] = M_X''(0) = \frac{2\times 3}{1} = 6$$

$$\sigma_X^2 = E[X^2] - E[X]^2 = 6 - 2^2 = 2 \Rightarrow \underline{\sigma_X = \sqrt{2}}$$

5.4. Ejercicio 4:

Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} k_1(x+1) & \text{si } 0 \leq x \leq 4 \\ k_2 x^2 & \text{si } 4 < x \leq 6 \end{cases} \quad \forall x \in [0,6]$$

Sabiendo que $P(0 \leq X \leq 4) = 2/3$, determinar k_1, k_2, y deducir su función de distribución.

Como f(x) es una función de densidad continua:

$$P(0 \leq X \leq 4) = 2/3 = \int_0^4 f(x)dx = \int_0^4 k_1(x+1)dx = k_1[\frac{x^2}{2} + x]_0^4 = k_1 * 12 \Rightarrow k_1 = \frac{2/3}{12} = \frac{1}{18}$$

Además, se debe cumplir que $\sum_{x \in \mathbb{N}} P(X = x) = 1 \Rightarrow P(4 < X \leq 6) = 1/3$

$$P(4 < X \leq 6) = 1/3 = \int_4^6 f(x)dx = \int_0^4 k_2(x^2)dx = k_2[\frac{x^3}{3}]_4^6 = k_2 * \frac{152}{3} \Rightarrow k_2 = \frac{1/3}{152/3} = \frac{1}{152}$$

La función de distribución $F_X(x)$ se calcula como:

$$F_X(x) = P[X \leq x] = \int_{-\infty}^x f(t)dt$$

En nuestro caso será:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{18}\int_0^x (t+1)dt & \text{si } 0 \leq x < 4 \\ \frac{2}{3} + \frac{1}{152}\int_4^x t^2 dt & \text{si } 4 \leq x < 6 \\ 1 & \text{si } x \geq 6 \end{cases}$$

Y resolviendo las integrales queda como:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x(x+2)}{36} & \text{si } 0 \leq x < 4 \\ \frac{x^3+240}{456} & \text{si } 4 \leq x < 6 \\ 1 & \text{si } x \geq 6 \end{cases}$$

5.5. Ejercicio 5:

La dimensión en centímetros de los tornillos que salen de cierta fábrica es una variable aleatoria, X, con función de densidad

$$f(x) = \frac{k}{x^2}, \quad 1 \leq x \leq 10$$

a) Determinar el valor de k, y obtener la función de distribución.

Para ser una función de densidad debe verificar $\int_{-\infty}^{+\infty} \frac{k}{x^2} dx = 1$

$$\int_{-\infty}^{+\infty} \frac{k}{x^2} dx = \cancel{\int_{-\infty}^{1} \frac{k}{x^2}} + \int_{1}^{10} \frac{k}{x^2} + \cancel{\int_{10}^{+\infty} \frac{k}{x^2}} = -k \int_{1}^{10} \frac{-1}{x^2} dx = -k(\frac{1}{10} - \frac{1}{1}) = \frac{9k}{10}$$

Por tanto

$$\frac{9k}{10} = 1 \Leftrightarrow k = \frac{10}{9}$$

La función de distribución será:

$$F_X(x) = \begin{cases} 0 & si \quad x \leq 1 \\ \int_1^x \frac{\frac{10}{9}}{x^2} dx = \frac{-10}{9}(\frac{1}{x} - 1) & si \quad 1 \leq x < 10 \\ 1 & si \quad x \geq 10 \end{cases}$$

b) Hallar la probabilidad de que la dimensión de un tornillo esté entre 2 y 5 cm.

Para hallar dicha probabilidad aplicaré que $F_X(B) = \int_B f(x)dx$ siendo $f(x)$ la función de densidad y B un intervalo contenido en Ω.

$$F_X(B) = \int_B f(x)dx = \int_2^5 \frac{\frac{10}{9}}{x^2} dx = \frac{-10}{9}(\frac{1}{5} - \frac{1}{2}) = \frac{1}{3}$$

Por tanto la probabilidad será de $\frac{1}{3}$.

c) Determinar la dimensión máxima del 50 % de los tornillos con menor dimensión y la dimensión mínima del 5 % con mayor dimensión.

La dimensión máxima del 50 % de los tornillos con menor dimensión va a coincidir con el valor de la mediana, es decir aquel $x_0 \in \Sigma / F_X(X) = 0,5$. Los cálculos quedan:

$$\frac{-10}{9}(\frac{1}{x} - 1) = \frac{1}{2} \ \Rightarrow \ \frac{1}{x} = \frac{-9}{20} + 1 \ \Rightarrow \ x = \frac{1}{\frac{-9}{20} + 1} = \frac{20}{11} = 1,81818$$

Y por tanto la dimensión máxima será de 1,81818 cm (aproximadamente).

La dimensión mínima del 5 % con mayor dimensión hace referencia al percentil 95 y se calculará de igual manera que antes pero cambiando 0,5 por 0,95:

$$\frac{-10}{9}(\frac{1}{x} - 1) = \frac{95}{100} \ \Rightarrow \ \frac{1}{x} = \frac{-855}{1000} + 1 \ \Rightarrow \ x = \frac{1}{\frac{-855}{100} + 1} = \frac{200}{29} = 6,89655$$

Es decir, la dimensión mínima será de 6,89655 cm (aproximadamente).

d) Si Y denota la dimensión de los tornillos producidos en otra fábrica, con la misma media y desviación típica que X, dar un intervalo en el que tome valores la variable Y con una probabilidad mínima de 0.99.

(Hace falta la desigualdad de Chebyshev: no la hemos dado)

5.6. Ejercicio 6:

Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \frac{2x-1}{10} & \text{si } 1 < x \leq 2 \\ \\ 0.4 & \text{si } 4 < x \leq 6 \end{cases}$$

a) Calcular $P(1.5 < X \leq 2),\ P(2.5 < X \leq 3.5),\ P(4.5 \leq X < 5.5),\ P(1.2 < X \leq 5.2)$.

b) Dar la expresión general de los momentos no centrados y deducir el valor medio de X.

c) Calcular la función generatriz de momentos de X.

Solución:

a) Para calcular las respectivas probabilidades, se integra en los intervalos correspondientes de la función, teniendo en cuenta que la función vale 0 en los intervalos en los que no está definida explícitamente.

$$P(1.5 < X \leq 2) = \int_{1.5}^{2} f(x)dx = \int_{1.5}^{2} \frac{2x-1}{10}dx = \frac{x^2-x}{10}]_{1.5}^{2} = 0.125$$

$$P(2.5 < X \leq 3.5) = \int_{2.5}^{3.5} 0dx = 0$$

$$P(4.5 \leq X < 5.5) = \int_{4.5}^{5.5} f(x)dx = \int_{4.5}^{5.5} 0.4dx = 0.4(5.5-4.5) = 0.4$$

$$P(1.2 < X \leq 5.2) = \int_{1.2}^{2} f(x)dx + \int_{4}^{5.2} f(x)dx = \int_{1.2}^{2} \frac{2x-1}{10}dx + \int_{4}^{5.2} 0.4dx =$$

$$= \frac{x^2-x}{10}]_{1.2}^{2} + 0.4(5.2-4) = 0.176 + 0.48 = 0.656$$

b) La expresión general de los momentos no centrados m_k se calcula de la siguiente manera:

$$m_k = E[X^k] = \int_{-\infty}^{+\infty} x^k f(x)dx = \int_{1}^{2} x^k \frac{2x-1}{10}dx + \int_{4}^{6} 0.4x^k dx =$$

$$= \int_{1}^{2} \left(\frac{1}{5}x^{k+1} - \frac{1}{10}x^k\right) dx + \int_{4}^{6} 0.4x^k dx = \left(\frac{x^{k+2}}{5(k+2)} - \frac{x^{k+1}}{10(k+1)}\right)_1^2 + 0.4\left(\frac{x^{k+1}}{k+1}\right)_4^6$$

Desarrollando:

$$m_k = \frac{(k+1)(2^{k+3}-2) + (k+2)(2^{k+3}3^{k+1} - 2^{2k+4} - 2^{k+1} + 1)}{10(k+1)(k+2)}$$

Sustituyendo $k = 1$ se obtiene la esperanza de x:

$$E[X] = m_1 = \frac{2 \times (16-2) + 3 \times (16 \times 9 - 64 - 4 + 1)}{10 \times 2 \times 3} = \frac{259}{60} = 4.31667$$

c) La función generatriz de momentos la da la siguiente expresión:

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{+\infty} e^{tx} f(x)dx = \int_{1}^{2} e^{tx} \frac{2x-1}{10}dx + \int_{4}^{6} 0.4e^{tx}dx$$

Separando integrales queda:

$$M_X(t) = \frac{1}{5}\int_{1}^{2} xe^{tx}dx - \frac{1}{10}\int_{1}^{2} e^{tx}dx + 0.4\int_{4}^{6} e^{tx}dx$$

Donde las dos últimas son inmediatas, y la primera se resuelve por partes:

$$M_X(t) = \left(\frac{e^{tx}(tx-1)}{5t^2} - \frac{e^{tx}}{10t} \right)_1^2 + \left(\frac{0.4e^{tx}}{t} \right)_4^6$$

Y por último desarrollando:

$$M_X(t) = \frac{4te^{2t}(e^{4t} - e^{2t} + 1) - e^{2t}(t+2) - e^t(3t-2)}{10t^2}$$

5.7. Ejercicio 7:

Con objeto de establecer un plan de producción, una empresa ha estimado que la demanda de sus clientes, en miles de unidades del producto, se comporta semanalmente con arreglo a una ley de probabilidad dada por la función de densidad:

$$f(x) = \frac{3}{4}(2x - x^2), 0 \leq x \leq 2$$

a) ¿Qué cantidad deberá tener dispuesta a la venta al comienzo de cada semana para poder satisfacer plenamente la demanda con probabilidad 0.5?
b) Pasado cierto tiempo, se observa que la demanda ha crecido, estimándose que en ese momento se distribuye según la función de densidad:

$$f(y) = \frac{3}{4}(4y - y^2 - 3), 1 \leq y \leq 3$$

Los empresarios sospechan que este crecimiento no ha afectado a la dispersión de la demanda, ¿es cierta esta sospecha?
Damos por hecho que dicha función de probabilidad cumple las condiciones:

- $0 \leq P(X = x) \leq 1 \ \forall x \in \mathbb{N}$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

a) nos están pidiendo la Mediana de la distribución, por tanto buscamos

$$Me \in [0, 2] tq \int_{-\infty}^{Me} f(x)dx = 0.5 = \int_0^{Me} \frac{3}{4}(2x - x^2)dx = \frac{3}{4}(x^2 - \frac{x^3}{3}]_0^{Me}) = \frac{3}{4}(Me^2 - \frac{Me^3}{3}) \Rightarrow Me = 1$$

De lo que podemos concluir que necesitaríamos 1000 unidades para satisfacer dicha demanda.

b)

Para comprobar la suposición, utilizamos la desviación típica en ambas distribuciones, ya que tienen las mismas unidades y se pueden comparar de esa manera.

$$E[X] = \int_0^2 \frac{3}{4}x(2x - x^2)dx = \left[\frac{x^3}{2} - \frac{3}{16}x^4\right]_0^2 = 1$$

$$E[X^2] = \int_0^2 \frac{3}{4}x^2(2x - x^2)dx = \left[\frac{3}{8}x^4 - \frac{3}{20}x^5\right]_0^2 = 1.2$$

$$E[Y] = \int_1^3 \frac{3}{4}y(4y - y^2 - 3)dy = \left[y^3 - \frac{3}{16}y^4 - \frac{9}{8}y^2\right]_1^3 = 2$$

$$E[Y^2] = \int_1^3 \frac{3}{4}y^2(4y - y^2 - 3)dy = \left[\frac{3}{4}y^4 - \frac{3}{20}y^5 - \frac{3}{4}y^3\right]_1^3 = 4.2$$

$$\sigma_X = \sqrt{Var(X)} = \sqrt{E[X^2] - E[X]^2} = \sqrt{0.2} = \frac{\sqrt{5}}{5}$$

$$\sigma_Y = \sqrt{Var(Y)} = \sqrt{E[Y^2] - E[Y]^2} = \sqrt{0.2} = \frac{\sqrt{5}}{5}$$

Como $\sigma_X = \sigma_Y$, la suposición era cierta, no ha variado la dispersión.

5.8. Ejercicio 8:

Calcular las funciones masa de probabilidad de las variables Y = X +2 y Z = X^2, siendo X una variable aleatoria con distribución:

$$P(X=-2)=\frac{1}{5} \quad P(X=-1)=\frac{1}{10} \quad P(X=0)=\frac{1}{5} \quad P(X=1)=\frac{2}{5} \quad P(X=2)=\frac{1}{10}$$

¿Cómo afecta el cambio de X a Y en el coeficiente de variación?

En este caso, si E_X es la función masa de probabilidad, tengo que $Img(E_X)=\{\frac{1}{5},\frac{1}{10},\frac{2}{5}\}$
Por el teorema General del Cambio de Variable

$$P_Y(B)=P_X(g^{-1}(B)) \text{ donde g(x) es una función medible}$$

Por tanto, si en este caso $g(X)=Y=X+2$, entonces $g^{-1}(X)=Y-2$

$$P(Y=y)=P(X=y-2)$$

$$-2=y-2 \Rightarrow y=0 \Rightarrow P(Y=0)=P(X=-2)=\frac{1}{5}$$

$$-1=y-2 \Rightarrow y=1 \Rightarrow P(Y=1)=P(X=-1)=\frac{1}{10}$$

$$0=y-2 \Rightarrow y=2 \Rightarrow P(Y=2)=P(X=0)=\frac{1}{5}$$

$$1=y-2 \Rightarrow y=3 \Rightarrow P(Y=3)=P(X=1)=\frac{2}{5}$$

$$2=y-2 \Rightarrow y=4 \Rightarrow P(Y=4)=P(X=2)=\frac{1}{10}$$

En el caso $h(x)=Z=X^2$ tendría $E_X=\{-2,-1,0,1,2\} \xrightarrow{h(x)=x^2} h(E_X)=\{0,1,4\}$

$$P(Z=z)=P(X^2=z)$$

$$z^2=0 \Rightarrow z=0 \Rightarrow P(Z=0)=P(X=0)=\frac{1}{5}$$

$$z^2=1 \Rightarrow z=\pm 1 \Rightarrow P(Z=1)=P(X=1)+P(x=-1)=\frac{2}{5}+\frac{1}{10}=\frac{1}{2}$$

$$z^2=4 \Rightarrow z=\pm 2 \Rightarrow P(Z=4)=P(X=2)+P(x=-2)=\frac{1}{10}+\frac{1}{5}=\frac{3}{10}$$

Para responder a la segunda cuestión deberé calcular los coeficientes de variación de Pearson para cada una de las variables:

$$CV_X=\frac{\sigma_X}{E[X]}=\frac{\sqrt{\sigma_X^2}}{E[X]}=\frac{\sqrt{E[(X-E[X])^2]}}{E[X]}$$

$$E[X]=(-2)\times\frac{1}{5}+(-1)\times\frac{1}{10}+0\times\frac{1}{5}+1\times\frac{2}{5}+2\times\frac{1}{10}=\frac{1}{10}$$

$$E[(X-E[X])^2]=E[X^2]-E[X]^2(-2)^2\times(\frac{1}{5})+(-1)^2\times(\frac{1}{10})+0^2\times(\frac{1}{5})+1^2\times(\frac{2}{5})+2^2\times(\frac{1}{10})-E[X]^2=$$

$$=\frac{17}{10}-(\frac{1}{10})^2=\frac{169}{100}=\frac{13^2}{10^2}$$

$$CV_X = \frac{\sqrt{\frac{13^2}{10^2}}}{\frac{1}{10}} = \frac{\frac{13}{\cancel{10}}}{\frac{1}{\cancel{10}}} = 13$$

$$CV_Y = \frac{\sigma_Y}{E[Y]} = \frac{\sqrt{\sigma_Y^2}}{E[Y]} = \frac{\sqrt{E[(Y - E[Y])^2]}}{E[Y]}$$

$$E[Y] = 0 \times \frac{1}{5} + 1 \times \frac{1}{10} + 2 \times \frac{1}{5} + 3 \times \frac{2}{5} + 4 \times \frac{1}{10} = \frac{21}{10}$$

$$E[(Y-E[Y])^2] = E[Y^2]-E[Y]^2 = 0^2\times\frac{1}{5}+1^2\times\frac{1}{10}+2^2\times\frac{1}{5}+3^2\times\frac{2}{5}+4^2\times\frac{1}{10}-E[Y]^2 = \frac{61}{10}-(\frac{21}{10})^2 = \frac{169}{100} = \frac{13^2}{10^2}$$

$$CV_Y = \frac{\sqrt{\frac{13^2}{10^2}}}{\frac{21}{10}} = \frac{\frac{13}{\cancel{10}}}{\frac{21}{\cancel{10}}} = \frac{13}{21}$$

$$\frac{CV_X}{CV_Y} = \frac{\cancel{13}}{\frac{\cancel{13}}{21}} = 21$$

Por tanto la desviación de la variable aleatoria Y es 21 veces menor que la desviación de X

Otra forma de resolver esta cuestión sería haciendo uso de las siguientes igualdades para transformaciones lineales:

$$E[aX + b] = aE[X] + b \qquad\qquad Var(aX + b) = a^2Var(X)$$

Entonces será:

$$E[Y] = E[X] + 2$$

$$Var(Y) = Var(X) \iff \sigma_Y = \sigma_X$$

Ahora calculamos la esperanza de X (de igual forma que antes) y de Y:

$$E[X] = (-2) \times \frac{1}{5} + (-1) \times \frac{1}{10} + 0 \times \frac{1}{5} + 1 \times \frac{2}{5} + 2 \times \frac{1}{10} = \frac{1}{10}$$

$$E[Y] = E[X] + 2 = \frac{1}{10} + 2 = \frac{21}{10} \quad \text{(mismo resultado que el calculado anteriormente)}$$

Y por último comparamos los coeficientes de variación:

$$C.V.(X) = \frac{\sigma_X}{|E[X]|} = 10\sigma_X$$

$$C.V.(Y) = \frac{\sigma_Y}{|E[Y]|} = \frac{\sigma_X}{\frac{21}{10}} = \frac{10}{21}\sigma_X$$

$$\frac{C.V.(X)}{C.V.(Y)} = 21 \qquad (\sigma_X \neq 0)$$

Llegamos a la misma conclusión que antes confirmando su validez.

5.9. Ejercicio 9:

Calcular las funciones de densidad de las variables $Y = 2X + 3$ y $Z = |X|$, siendo X una variable continua con función de densidad:

$$f_X(x) = \frac{1}{4}, \qquad -2 < x < 2$$

Solución:

$Y = 2X + 3 = h(X)$ estrictamente monótona, medible y derivable en $]-2,2[$

$h(]-2,2[) =]-1,7[$ $\qquad h^{-1}(y) = \frac{y-3}{2}\ \forall y \in]-1,7[$ $\qquad (h^{-1})'(y) = \frac{1}{2}\ \forall y \in]-1,7[$

Por el Teorema de Cambio de Variable de continua a continua:

$$f_Y(y) = f_X(h^{-1}(y)|(h^{-1})'(y)| = \frac{1}{4} \times |\frac{1}{2}| = \frac{1}{8}\ \forall y \in]-1,7[$$

Entonces la función de densidad queda:

$$f_Y(y) = \begin{cases} \frac{1}{8} & y \in]-1,7[\\ 0 & y \notin]-1,7[\end{cases}$$

$Z = |X| = h(X)$ medible, no derivable, no estrictamente monótona

$$h(x) = \begin{cases} x & \text{si } -2 < x < 0 \\ -x & \text{si } 0 \leq x < 2 \end{cases} \quad \forall z \in]-2,2[$$

$$h^{-1}(z) = \begin{cases} h_1^{-1}(z) = z \\ h_2^{-1}(z) = -z \end{cases} \quad \forall z \in [0,2[$$

$$(h^{-1})'(z) = \begin{cases} (h_1^{-1})'(z) = 1 \\ (h_2^{-1})'(z) = -1 \end{cases} \quad \forall z \in [0,2[$$

Se requiere aplicar la generalización del Teorema de Cambio de Variable continua a continua, ya que cada valor de $z \in [0,2[$ tiene 2 preimágenes:

$$f_Z(z) = f_X(h_1^{-1}(z))|(h_1^{-1})'(z)| + f_X(h_2^{-1}(z))|(h_2^{-1})'(z)| = \frac{1}{4} \times |1| + \frac{1}{4} \times |(-1)| = \frac{1}{2}\ \forall z \in [0,2[$$

Entonces la función de densidad queda:

$$f_Z(z) = \begin{cases} \frac{1}{2} & z \in [0,2[\\ 0 & z \notin [0,2[\end{cases}$$

5.10. Ejercicio 10:

Si X es una variable aleatoria con función de densidad $f(x) = \frac{e^{-|x|}}{2} si - \infty < x < \infty$, hallar su función de distribución y las probabilidades de cada uno de los siguientes sucesos:
$a)|X| \leq 2$.
b) $|X| \leq 2$ o $X \geq 0$.
c) $|X| \leq 2$ y $X \leq -1$.
d) $X^3 - X^2 - X - 2 \leq 0$.
e) X es irracional.
Primero separamos el valor absoluto:

$$f(x) = \begin{cases} \frac{e^x}{2} & x \in]-\infty, 0[\\ \\ \frac{e^{-x}}{2} & x \in [0, \infty[\end{cases}$$

Veamos que dicha función de densidad cumple las condiciones:

- $0 \leq P(X = x) \leq 1 \ \forall x \in \mathbb{N}$

  $$f'(x) = \begin{cases} \frac{e^x}{2} & x \in]-\infty, 0[\\ \\ \frac{-e^{-x}}{2} & x \in [0, \infty[\end{cases}$$

  es fácil ver que $f'(x) > 0 \ \forall x \in]-\infty, 0[$ y $f'(x) < 0 \ \forall x \in [0, \infty[$, por tanto f crece desde -infinito hasta el 0 y después decrece hasta el infinito. Además, como f(x) $>$ 0 $\forall x \in \mathbb{R}$ podemos ver que f(0) = 1/2 en ambas ecuaciones por lo que sí, está acotada entre 0 y 1/2.

- $\int_{-\infty}^{+\infty} f(x)dx = 1$

  $$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} \frac{e^x}{2}dx + \int_{0}^{+\infty} \frac{e^{-x}}{2}dx = \frac{e^x}{2}]_{-\infty}^{0} + \frac{-e^x}{2}]_{0}^{\infty} = \frac{1}{2} + \frac{1}{2} = 1 \qquad (2)$$

  Por tanto podemos concluir que cumple ambas condiciones, y es una función de densidad.

Hallemos ahora la función de distribución:

$$F(X) = \begin{cases} \frac{e^x}{2} & x \in]-\infty, 0[\\ \\ 1 - \frac{e^{-x}}{2} & 0 \in [x, \infty[\end{cases}$$

Hallemos las probabilidades de cada uno de los siguientes sucesos:

- a) $P(|X| \leq 2) = P(-2 \leq X \leq 2) = F(2) - F(-2) = 1 - \frac{e^{-2}}{2} - \frac{e^{-2}}{2} = 1 - e^{-2}$
- b) $P(|X| \leq 2 \text{ o } X \geq 0) = P(X \geq -2) = 1 - P(X \leq 2) = 1 - \frac{e^{-2}}{2} = 1 - \frac{1}{2e^2} = \frac{2e^2-1}{2e^2}$
- c) $P(|X| \leq 2 \text{ y } X \leq -1) = P(-2 \leq X \leq -1) = \int_{-2}^{-1} \frac{e^x}{2} = \frac{e^{-1}}{2} - \frac{e^{-2}}{2} = \frac{e-1}{2e^2}$
- d) $P(X^3 - X^2 - X - 2 \leq 0)$
  $X^3 - X^2 - X - 2 \leq 0 \Leftrightarrow X \leq 2$
  $P(X \leq 2) = 1 - \frac{e^{-2}}{2} = 1 - \frac{1}{2e^2} = \frac{2e^2-1}{2e^2}$
- e) $P(X \in (R - Q)\ ) = 1$ (por que sabemos que el conjunto de los números racionales es numerable, o sea, es insignificante en comparación al de los irracionales, entonces la integral se hace igualmente.

5.11. Ejercicio 11:

Sea X una variable aleatoria con función de densidad

$$f(x) = 1, \;\; 0 \leq x \leq 1$$

Encontrar la distribución de las variables:

a) $Y = \frac{X}{1+X}$

Es claro que Y es una variable aleatoria de tipo continuo ya que para cada valor de Y existe un único X. Por tanto su función de densidad se obtiene aplicando la fórmula de Cambio de Variable a cada antiimagen y sumando las funciones resultantes:

$$y = g(x) = \frac{x}{1+x} \;\Rightarrow\; x = \frac{y}{1-y} \;\Rightarrow\; g^{-1}(y) = \frac{y}{1-y}$$

$$(g)'(x) = \frac{(1+x)-x}{(1+x)^2} = \frac{1}{(1+x)^2} > 0 \;\Rightarrow\; g \text{ es estrictamente creciente}$$

$$g(0) = 0 \qquad g(1) = \frac{1}{2} \;\Rightarrow\; Im(g(x)) = [0, \frac{1}{2}]$$

$$(g^{-1})'(x) = \frac{(1-y)+y}{(1-y)^2} = \frac{1}{(1-y)^2} > 0 \qquad \forall y \in [0, \frac{1}{2}]$$

Por el teorema de cambio de Variable de Continua a Continua tengo:

$$f_Y(Y) = f_X(g^{-1}(y))|\frac{dg^{-1(y)}}{dy}| = f_X(\frac{x}{1-x})|\frac{1}{(1-y)^2}| = \frac{1}{(1-y)^2} \qquad \forall y \in [0, \frac{1}{2}]$$

Una vez tengo calculada la función de densidad puedo calcular la función de distribución

$$F_Y(y) = \begin{cases} 0 & si \quad y < 0 \\ \int_{-\infty}^{y} f_Y(t)\; dt & si \quad 0 \leq y < \frac{1}{2} \\ 1 & si \quad y >= \frac{1}{2} \end{cases}$$

$$\int_{-\infty}^{y} f_Y(t)\; dt = \cancel{\int_{-\infty}^{0}} 0\; dt + \int_{0}^{y} \frac{1}{(1-t)^2}\; dt = (\frac{1}{1-y} - \frac{1}{1-0}) = \frac{1-(1-y)}{1-y} = \frac{y}{1-y}$$

Y finalmente queda

$$F_Y(y) = \begin{cases} 0 & si \quad y < 0 \\ \frac{y}{1-y} & si \quad 0 \leq y < \frac{1}{2} \\ 1 & si \quad y >= \frac{1}{2} \end{cases}$$

b) $Z = \begin{cases} -1 & si \quad X < 3/4 \\ 0 & si \quad X = 3/4 \\ 1 & si \quad X > 3/4 \end{cases}$

La variable Z es claramente discreta y aplicando el Teorema de Cambio de Variable de continua a discreta tengo que

$$\forall z \in E_Z, \;\; P\{Z = z\} = \int_{g^{-1}(y)} f_X(x) dx$$

En este caso tengo que

$$P(Z = 0) = P(X = \frac{3}{4}) = \int_{\frac{3}{4}}^{\frac{3}{4}} 1\, dx = 0$$

$$P(Z=-1) = P(X<\frac{3}{4}) = \int_{-\infty}^{\frac{3}{4}} 1\,dx = \cancel{\int_{-\infty}^{0} 1\,dx} + \int_{0}^{\frac{3}{4}} 1\,dx = \frac{3}{4}$$

$$P(Z=1) = P(X>\frac{3}{4}) = \int_{\frac{3}{4}}^{+\infty} 1\,dx = \int_{\frac{3}{4}}^{1} 1\,dx + \cancel{\int_{1}^{+\infty} 1\,dx} = 1-\frac{3}{4} = \frac{1}{4}$$

Por tanto me queda que la función masa de probabilidad es

$$P_Z(z) = \begin{cases} \frac{3}{4} & si \quad z=-1 \\ 0 & si \quad z=0 \\ \frac{1}{4} & si \quad z=1 \end{cases}$$

Lo que me lleva a la función de distribución:

$$F_Z(z) = \begin{cases} 0 & si \quad z<-1 \\ \frac{3}{4} & si \quad -1\leq x<1 \\ 1 & si \quad z\geq 1 \end{cases}$$

5.12. Ejercicio 12:

Sea X una variable aleatoria simétrica con respecto al punto 2, y con coeficiente de variación 1. ¿Qué puede decirse acerca de las siguientes probabilidades?:

- $P(-8 < X < 12)$
- $P(-6 < X < 10)$

Solución:

Sabiendo que es simétrica con respecto al punto 2, se puede concluir que su esperanza vale 2.
Como se define el coeficiente de variación de Pearson como el cociente entre la desviación típica y la esperanza, al valer este 1, podemos concluir que la desviación típica vale 2, y por tanto, la varianza será 4. Como la varianza existe y es finita, se puede aplicar la desigualdad de Chebyshev:

$$P(|X - E[X]| < k) \geq 1 - \frac{Var(X)}{k^2}$$

En nuestro caso:

$$P(|X - 2| < k) \geq 1 - \frac{4}{k^2} \iff P(2 - k < X < 2 + k) \geq 1 - \frac{4}{k^2}$$

Tomando $k = 10$ y $k = 8$ se acotan las probabilidades del enunciado:

$$P(-8 < X < 12) \geq 1 - \frac{4}{100} = 0.96$$

$$P(-6 < X < 10) \geq 1 - \frac{4}{64} = 0.9375$$

Se puede concluir que dada una distribución con las características del enunciado, las probabilidades centrales serán muy altas.