PROBABILIDAD I

Segundo del Grado en Matemáticas

Hugo Marquerie

Profesor: Pablo Fernández Gallardo

Facultad de Ciencias - Universidad Autónoma de Madrid

Segundo cuatrimestre 2023 - 2024

February 1, 2024

1 Tema 1: Sucesos y probabilidades

1.1 Formalizando

Definición 1.1 (Espacio muestral). En un experimento aleatorio, el espacio muestral es el conjunto (no vacío) de sus posibles resultados y se denota por Ω. Puede ser:

1. Finito: $\Omega = \{\omega_1, \dots, \omega_N\}$
2. Infinito numerable: $\Omega = \{\omega_1, \omega_2, \dots\}$
3. Infinito no numerable, ej.: $\Omega = [0,1) \vee \Omega = \mathcal{P}([0,1))$

Definición 1.2 (Espacio de sucesos por Kolmogórov). Dado el espacio muestral Ω de un experimento aleatorio, $\mathcal{F} \subset \mathcal{P}(\Omega)$ es su espcio de sucesos

$$\iff (\mathcal{F} \neq \phi) \wedge \left(A \in \mathcal{F} \implies A^C \in \mathcal{F}\right) \wedge \left(A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{j=1}^{\infty} A_j \in \mathcal{F}\right)$$

Observación 1.1. De la definición se deduce:

• $\phi \in \mathcal{F} \wedge \Omega \in \mathcal{F}$ • $A_1, \dots, A_n \in \mathcal{F} \implies \bigcup_{j=1}^{n} A_j \in \mathcal{F}$ • $\forall A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$

Definición 1.3 (Función o medida de probabilidad). Dados espacio muestral (Ω) y de sucesos $(\mathcal{F})$ de un experimento aleatorio, la aplicación $P\colon \mathcal{F} \longrightarrow [0,1]$: es una medida de probabilidad

$$\iff (P(\Omega) = 1) \wedge \left[P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j) \iff A_i \cap A_j = \phi \text{ cuando } i \neq j\right]$$

Proposición 1.1. *De la definición se deduce:*

1. $P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j)$ *2.* $P\left(A^C\right) = 1 - P(A)$ *3.* $P(\phi) = 0$

4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ *5.* $A \subseteq B \implies P(A) \leq P(B)$

Demostración. (ejercicio) ■

Ejemplo 1.1. En un experimento aleatorio con espacio muestral finito, tomamos $\Omega = \{\omega_1, \dots, \omega_N\} \wedge \mathcal{F} = \mathcal{P} \to 2^N$. Asignamos $P(\{\omega_j\}) = p_j \wedge j = 1, \dots, N$ tales que $p_j \geq 0 \wedge \sum_{j=1}^{N} p_j = 1$. Entonces, $\forall A \in \mathcal{F} : P(A) = \sum_{\omega \in A} P(\omega)$

Caso particular: $\forall j \in \{1, \dots, N\} : p_j = \frac{1}{N} \implies \forall A \in \mathcal{F} : P(A) = \frac{|A|}{N} =$

$\dfrac{\text{"Casos favorables"}}{\text{"Total de casos"}}$

1. (Muy tonto) $\Omega \neq \phi$, tomas $A \subset \Omega : A \neq \phi, \Omega$.
   Dato $p \in (0,1)$. $\mathcal{F} = \{\phi, A, A^c, \Omega\}$ con $P(A) = p$.

2. (Bastante general) $\Omega = \{\omega_1, \dots, \omega_N\} \to |\Omega| = N$
   $\mathcal{F} = P(\Omega) \to |\mathcal{F}| = 2^N$. Dato: $p_1, \dots, p_N \geq 0 \implies \sum_{j=1}^{N} p_j = 1$. Asignamos $P(\{\omega_j\}) = p_j$ para cada $j = 1, \dots, N$.
   Defines, para $A \subset \Omega P(A) = \sum_{\omega \in A} P(\omega)$.
   Caso particular: $p_j = \dfrac{1}{N}, j = 1, \dots, N \implies P(A) = \dfrac{|A|}{N}$

3. Lanzas n veces la moneda.
   Dato: $p \in (0,1)$ $\Omega = \{111\cdots1, \dots, 000\cdots0\}, |\Omega| = 2^N$. $\mathcal{F} = P(\Omega)$.
   $P(\omega) = p^{\#\text{unos de } \omega}(1-p)^{\#\text{ceros de } \omega}$
   Comprobamos:
   $$\sum_{\omega \in \Omega} P(\omega) = \sum_{k=0}^{n} \left(\sum_{\substack{\omega \in \Omega \\ \#\text{ceros de } \omega = k}} P(\omega) \right) = \sum_{k=0}^{n} p^k (1-p)^{n-k} \left(|\{\omega \in \Omega : \# \text{ unos de } \omega \text{ en } k\}| \right)$$
   $$= \sum_{k=0}^{n} p^k (1-p)^{n-k} \binom{n}{k} = (p + 1 - p)^n = 1$$

4. Lanzamos moneda hasta que sale una cara. Dato $p \in (0,1)$.
   $\Omega = \{C, XC, XXC, \dots\} \wedge \mathcal{F} = P(\Omega)$.
   $\implies P(C) = p \wedge P(XC) = p(1-p) \wedge P(XXC) = p(1-p)^2 \wedge \dots$
   Comprobamos: $\sum_{j=1}^{\infty} p(1-p)^{j-1} = p \sum_{k=0}^{\infty} (1-p)^k = p \dfrac{1}{1-(1-p)} = 1$

1.2 Probabilidad condicionada y total, independencia y Bayes

Tienes $(\Omega, \mathcal{F}, P)$ y un suceso $A \in \mathcal{F} \to P(A)$. Llega "nueva información": ha ocurrido el suceso $B \in F \to$ ¿Debo reasignar la probabilidad de A?

Ejemplo 1.2. Lanzas 10 veces la moneda (regular).
$A = \{\text{salen 6 caras}\} \implies P(A) = \dfrac{\binom{10}{6}}{2^10} \approx 20.51\%$ $B = \{\text{sale C en } 1^{\text{o}}\} \implies P(A)$ sube a $\dfrac{\binom{9}{5}}{2^9}$

Definición 1.4. Sea $(\Omega, \mathcal{F}, P)$, sucesos $A, B \in \mathcal{F}, P(B) > 0$, $P(A|B)$ es la probabilidad de

A condicionada a B

$$\iff P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Observación 1.2. En general, $P(A|B) \neq P(B|A)$

Proposición 1.2 (Calculamos $P(A|B)$ para cada $A \in \mathcal{F}$)**.** *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $B \in \mathcal{F}$ un suceso con $P(B) > 0$*

$\implies (\Omega, \mathcal{F}, Q_B)$*, donde* $Q_B \colon \mathcal{F} \longrightarrow [0,1] : Q_B(A) = P(A|B)$*, es un espacio de probabilidad*

Demostración.

$$\left(Q_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \in [0,1] \right) \wedge \left(Q_B(\Omega) = P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1 \right)$$

Sean $A_1, A_2, \dots \in \mathcal{F}$ disjuntos dos a dos.

$$\implies Q_B\left(\bigcup_{j=1}^{\infty} A_j\right) = P\left(\bigcup_{j=1}^{\infty} A_j \,\middle|\, B\right) = \frac{P\left(\bigcup_{j=1}^{\infty} A_j \cap B\right)}{P(B)}$$

$$= \frac{1}{P(B)} P\left(\bigcup_{j=1}^{\infty} (A_j \cap B)\right) = \frac{1}{P(B)} \sum_{j=1}^{\infty} P(A_j \cap B) = \sum_{j=1}^{\infty} Q_B(A_j)$$

$\blacksquare$

Definición 1.5 (Independencia). Sean $A, B \in \mathcal{F}$ dos sucesos con $P(A), P(B) \geq 0$ son independientes

$$\iff P(A|B) = P(A) \left(\wedge P(B|A) = P(B) \right) \text{ para entender}$$

$$\iff P(A \cap B) = P(A) \cdot P(B) \text{ la adecuada}$$

- A, B disjuntos $\implies$ no independientes.
- $A_1, \dots, A_N \in \mathcal{F}$ independientes $\iff \forall J \subset \mathbb{N}_N : P\left(\bigcap_{j \in I} A_j\right) = \prod_{j \in J} P(A_j)$

$$\iff P\left(\overline{A_1} \cap \dots \cap \overline{A_N}\right) = \prod_{i=1}^{N} P\left(\overline{A_i}\right) \text{ donde } \overline{A_i} = A_i, (A_i)^c$$

Ejercicio 1.2.1. Encontrar un espacio de probabilidad en la que haya un conjunto de sucesos independientes dos a dos pero no completamente independientes.

SOL: $\Omega = \{1,2,3,4\} \wedge A = \{1,2\} \wedge B = \{2,3\} \wedge C = \{1,3\}$

Proposición 1.3 (Regla de Bayes)**.** *Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $A, B \in \mathcal{F}$*

sucesos con $P(A), P(B) > 0$

$$\implies P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$$

Demostración.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \wedge P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$\implies P(A|B) \cdot P(B) = P(A \cap B) = P(B|A) \cdot P(A) \implies P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$$

■

1.2.1 Probabilidad Total

Tenemos una partición de $\Omega B_1, B_2, \cdots : (B_i \cap B_j = \phi \iff i \neq j) \wedge \left(\bigcup_j B_j = \Omega\right)$

$$\implies P\left(\bigcup_j (A \cap B_j)\right) = \sum_j (A \cap B) = \sum_j P(A|B_j) \cdot P(B_j)$$

Ejemplo 1.3. $U_1 = \{10b, 3n\} \wedge U_2 = \{5b, 5n\} \wedge U_3 = \{2b, 6n\}$

Procedimiento:

1. Sorteamos una urna $1/4 \to U_1 \wedge 1/4 \to U_2 \wedge 1/2 \to U_3$
2. Sacamos bola de la urna seleccionada.

¿Cuál es la probabilidad de que sea blanca?

$$P(b) = P(b|U_1)P(U_1) + P(b|U_2)P(U_2) + P(b|U_3)P(U_3)$$

Ejemplo 1.4 (Peso de la evidencia). $U_1 = \{80\%b, 20\%n\} \wedge U_2 = \{20\%b, 80\%n\}$

Procedimiento:

1. Sorteamos la urna con $1/2$ y $1/2$ de probabilidad.
2. Sacamos 10 bolas (con reemplazamiento).

Observamos la evidencia: $bb \dots nb$ ¿qué urna se usó?

$$P(U_1|5b5n) = P(5b5n|U_1)\frac{P(U_1)}{P(5b5n)} = \frac{P(5b5n|U_1)P(U_1)}{P(5b5n|U_1)P(U_1) + P(5b5n|U_2)P(U_2)}$$

$$P(5b5n|U_1) = \binom{10}{5} 0.8^5 0.2^5 = P(5b5n|U_2) \implies P(U_1|5b5n) = \frac{1}{2}$$

$$P(U_1|6b4n) = \frac{P(6b4n|U_1)P(U_1)}{P(6b4n|U_1)P(U_1) + P(6b4n|U_2)P(U_2)} = \frac{\binom{10}{6}0.8^6 0.2^4 \cdot 1/2}{\binom{10}{6}0.8^6 0.2^4 \cdot 1/2 + \binom{10}{6}0.8^4 0.2^6 \cdot 1/2} \approx 90\%$$

Ejemplo 1.5 (Falsos positivos/negativos). Hay una enfermedad $\to E \vee S$ y hay una prueba para detectar $\to + \vee -$. Datos: $P(+|E) = 95\% \wedge P(-|S) = 99\%$.

Te haces la prueba y sale $+$:

$$P(E|+) = P(+|E)\frac{P(E)}{P(+)} = \frac{P(+|E)P(E)}{P(+|E)P(E) + P(+|S)P(S)}$$

Conozco todas estas probabilidades excepto $p := P(E) \implies P(S) = 1 - p$.

Si definimos $f(p) := P(E|+)$

$$\implies \{f(0.5) = 98.95\% \wedge f(1/100) = 48.97\% \wedge f(1/1000) = 0.90\%\}$$

Es decir, si la incidencia es muy baja, no tiene sentido hacer pruebas masivamente porque la mayoría de positivos serán falsos.

Ejemplo 1.6 (Sobre independencia). $(\Omega, \mathcal{F}, P) \wedge A_1, \ldots, A_n$ sucesos independientes tal que $\forall j \in \mathbb{N}_n : P(A_j) = \frac{1}{n}$. ¿Qué sabemos sobre $P\left(\bigcup_{j=1}^n A_j\right)$?

$$\text{En general, sabemos que } \frac{1}{n} \leq P\left(\bigcup_{j=1}^n A_j\right) \leq \sum_{j=1}^n P(A_j) \leq 1$$

$n = 2$: $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 3/4$

$n = 3$: $P(A \cup B \cup C) = \binom{3}{1}\frac{1}{3} - \binom{3}{2}\frac{1}{3^2} + \binom{3}{3}\frac{1}{3^3} = 19/27$

n general: $P\left(\bigcup_{j=1}^n A_j\right) = \sum_{j=1}^n P(A_j) - \sum_{1 \leq i < y \leq n} P(A_i \cap A_j) + \cdots$ (Inclusión exclusión)

$$\implies P\left(\bigcup_{j=1}^n A_j\right) = \binom{n}{1}\frac{1}{n} - \binom{n}{2}\frac{1}{n^2} + \binom{n}{3}\frac{1}{n^3} + \cdots = \sum_{j=1}^n \binom{n}{j}\frac{1}{n^j}(-1)^{j+1}$$

$$\implies P\left(\bigcup_{j=1}^n A_j\right) = 1 - \sum_{j=0}^n \binom{n}{j}\left(\frac{-1}{n}\right)^j = 1 - \left(1 - \frac{1}{n}\right)^n \xrightarrow{n\to\infty} 1 - \frac{1}{e}$$

1.2.2 Continuidad de la probabilidad: detalle técnico

Proposición 1.4. *Sea* $(\Omega, \mathcal{F}, P) \wedge A_1, \cdots : A_1 \subset A_2 \subset \cdots$ *una sucesión creciente de conjuntos*

$$\implies P\left(\bigcup_{j=1}^\infty A_j\right) = \lim_{n\to\infty} P(A_n)$$

Demostración. Se trata de describir $\bigcup_{j=1}^n A_j$ como la unión de conjuntos disjuntos.

$$P\left(\bigcup_{j=1}^n A_j\right) = P\left(A_1 \cup \bigcup_{j=1}^{n-1}(A_{j+1} \setminus A_j)\right) = P(A_1) + \sum_{j=1}^{n-1}(P(A_{j+1}) - P(A_j)) = P(A_n)$$

$$\implies P\left(\bigcup_{j=1}^{\infty} A_j\right) = P(A_1) + \sum_{j=1}^{\infty} \left(P(A_{j+1}) - P(A_j)\right) = \lim_{n\to\infty} P(A_n)$$

■

Proposición 1.5. *Si la sucesión $A_1, \dots$ es decreciente* $\implies P\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} P\left(A_n\right)$

Teorema 1.1 (Continuidad de la probabilidad)**.** *En* $(\Omega, \mathcal{F}, P) \wedge A_1, A_2, \dots$ *sucesión* $: \forall j \in \mathbb{N} : A_j \in \mathcal{F}$

$$\implies P\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} P\left(\bigcup_{j=1}^{n} A_j\right)$$

Demostración. ■

1.3 Variables aleatorias (discretas)

Definición 1.6. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidades, $X\colon \Omega \longrightarrow \mathbb{R}$ es una variable aleatoria discreta* (v.a.d.)

$$\iff (1)\ X(\Omega) \text{ es numerable*} \wedge\ (2)\ \forall x \in \mathbb{R} : \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}$$

En realidad, solo interesa (2) cuando $x = x_j$

Definición 1.7. Sea X una v.a.d. en $(\Omega, \mathcal{F}, P)$, p_X es su función de masa

$$\iff p_X\colon \mathbb{R} \longrightarrow [0,1] \wedge x \longmapsto p_X(x) = P(X = x) := P(\{\omega \in \Omega : X(\omega) = x\})$$

Vemos que

$$\sum_{j\geq 1} p_X(x_j) = \sum_{j\geq 1} P(X = x_j) = \sum_{j\geq 1} P(\{\omega \in \Omega : X(\omega) = x_j\}) = P\left(\bigcup_{j\geq 1}\{x_j\}\right) = P(\Omega) = 1$$

Lo relevante es la lista de posibles valores de X $\{x_1, x_2, \dots\}$ numerable y la lista (también numerable) de probabilidades $p_1, p_2, \dots$ donde $\forall j \geq 1 : p_j = P(x = x_j) \wedge p_j \geq 0 \wedge \sum_{j\geq 1} p_j = 1$

Teorema 1.2. *Sea* $S = \{x_1, x_2, \dots\}$ *un conjunto numerable y* $\Pi_1, \Pi_2, \dots, \Pi_j \geq 0 \wedge \sum_{j\geq 1} \Pi_j = 1$ *una lista*

$$\implies \exists (\Omega, \mathcal{F}, P) \wedge X \ \ v.a.d\ : \forall x \notin S : p_x(x) = 0 \wedge p_x(x_j) = \Pi_j$$

Demostración. Fijamos $\Omega = S$ y $\mathcal{F} = \mathcal{P}(S)$.

$$A \in \mathcal{F} \implies P(A) = \sum_{j : x_j \in A} \Pi_j \wedge X(x_j) = x_j$$

$\blacksquare$

Ejemplo 1.7 (Dados). 1. Uniforme en $\{1, \cdots, N\}, N \geq 2$.

$$S = \{1, \cdots, N\} \wedge \Pi_j = {}^1/N, \ldots, {}^1/N$$

2. X sigue una distribución de Bernoulli ($X \sim \operatorname{BER}(p)$) con parámetro p

$$\iff \begin{cases} p_X(x) = 0 \iff x \neq 0, 1 \\ p_X(1) = p \wedge p_X(0) = 1 - p \end{cases} \iff \begin{cases} 1, p \\ 0, 1 - p \end{cases} \quad \text{donde 1 es éxito y 0 fracaso}$$

3. X sigue una binomial de parámetro $n \geq 1 \wedge p \in (0,1)$ ($X \sim \operatorname{BIN}(n,p)$)

$$\iff S = \{0, 1, \ldots, n\} \wedge \forall j \in \{0, 1, \ldots, n\} : P(x = j) = \binom{n}{j} p^j (1-p)^{n-j}$$

Sirve para modelizar el número de caras que salen al lanzar n veces una moneda de probabilidad p.

$$n! \sim n^n e^{-n} \sqrt{2\pi n} \implies \binom{n}{{}^n/2} \sim \frac{n^n e^{-n} \sqrt{2\pi n}}{({}^n/2)^{({}^n/2)} e^{-({}^n/2)} \sqrt{2\pi ({}^n/2)} ({}^n/2)^{({}^n/2)} e^{-({}^n/2)} \sqrt{2\pi ({}^n/2)}}$$

$$\implies \frac{n^n \sqrt{n}}{({}^n/2)^{({}^n/2)} \sqrt{2\pi ({}^n/2)} ({}^n/2)^{({}^n/2)} \sqrt{({}^n/2)}} = \frac{n^n \sqrt{n}}{({}^n/2)^n \sqrt{2\pi} ({}^n/2)} = \frac{n^n \sqrt{2}}{({}^n/2)^n \sqrt{\pi n}} = 2^n \sqrt{\frac{2}{\pi n}}$$

4. La variable X sigue una distribución geométrica de parámetro $p \in (0,1)$ ($X \sim \operatorname{GEOM}(p)$).

$$\iff S = \{1, 2, \ldots\} \wedge \forall j \geq 1 : P(x = j) = p(1-p)^{j-1}$$

Sirve para modelizar el número de lanzamientos hasta que sale un resultado C en cuestión.

Observación 1.3. Cuidado porque existen variables aleatorias que también se dicen de distribución geométrica en las que $S = \{0, 1, 2, \ldots\}$. Se habla de cuantas veces has obtenido el resultado complementario a C antes de que halla salido C.

5. La variable X sigue una distribución de Poisson con parámetro $\lambda > 0$ ($X \sim \operatorname{POISSON}(\lambda)$)

$$\iff S = \{0, 1, \ldots\} \wedge \forall j \geq 0 : P(x = j) = e^{-\lambda} \frac{\lambda^j}{j!}$$

Proposición 1.6. *Sea $X \sim \operatorname{BIN}(n,p)$ una v.a.d.*

$$\implies \textit{cuando } n \textit{ es grande, } \operatorname{BIN}(n,p) \sim \operatorname{POISSON}(np)$$

Demostración. Fijo $\lambda > 0 \wedge p = \dfrac{\lambda}{n}$.

$$\lim_{n\to\infty} \binom{n}{k} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n} \left(1 - \frac{\lambda}{n}\right)^{-k} = e^{-k} \frac{\lambda^k}{k!}$$

■

Ejemplo 1.8 (¿Hay más ejemplos?).

- Binomial negativa
- Hipergeométrica
- Dada cualquier serie convergente $\sum_{n=1}^{\infty} a_n = s$, se puede definir la variable aleatoria X tal que $S = \{1, 2, \dots\} \wedge P(x = k) = \dfrac{a_k}{s}$

1.3.1 Funciones transformadoras de variables aleatorias (discretas)

Sea X una v.a.d. y $g \colon \mathbb{R} \longrightarrow \mathbb{R}$ una función, definimos $Y := g(X)$.

$$\implies \{\omega \in \Omega : Y(\omega) = g(X(\omega)) = y\} \in \mathcal{F} \implies Y \text{ es una v.a.d}$$

Por otro lado,

$$\implies \forall y \in \mathbb{R} : P(Y = y) = P(g(X) = y) = P(X \in g^{-1}(y)) = \sum_{x \in g^{-1}(y)} P(X = x) = p_X(x)$$

Ejemplo 1.9 ($Y = x^2$).

$$p_Y(y) = P(Y = y) = P(X^2 = y) = \begin{cases} 0 \iff y < 0 \\ P(X = 0) = p_X(0) \iff y = 0 \\ P\left(X = \pm\sqrt{y}\right) = p_X\left(\sqrt{y}\right) + p_X\left(-\sqrt{y}\right) \iff y > 0 \end{cases}$$

1.3.2 Resúmenes: esperanza, varianza, momentos

Definición 1.8 (Esperanza). Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y X una v.a.d. con función de masa p_X, E es la esperanza de X (también llamada media o *expectatio*)

$$\iff E(X) = \sum_{x \in X(\Omega)} x \cdot P(X = x) = \sum_{j \geq 1} x_j \cdot p_X(x_j)$$

Pero ojo, solo si la serie es absolutamente convergente.

Si $x_1, \dots, x_N$ finito, la suma obviamente converge.

Si los x_j son positivos, la serie converge si y solo si es acotada. Si no lo es diverge a ∞.

Ejemplo 1.10.

- $x_1, \ldots, x_N \wedge p_1, \ldots, p_N \implies E(X) = \sum_{j=1}^{N} x_j \cdot p_j$
- $X \sim \mathrm{BER}\,(p) \implies E(X) = 1 \cdot p + 0 \cdot (1-p) = p$
- $X \sim \mathrm{UNIF}\,(1, \ldots, N) \implies E(X) = \frac{1}{N} \sum_{n=1}^{N} n = \frac{N+1}{2}$
- $X \sim \mathrm{BIN}\,(n, p) \implies E(X) = \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j} = np$

  Se obtiene derivando el binomio de Newton $(1+x)^n = \sum_{j=0}^{n} \binom{n}{j} x^j$

$$\implies \frac{\mathrm{d}}{\mathrm{d}x}(1+x)^n = \sum_{j=0}^{n} \binom{n}{j} j \cdot x^{j-1} \implies xn(1+x)^{n-1} = \sum_{j=0}^{n} \binom{n}{j} j x^j$$

  Si $x = \frac{p}{1-p} \implies \frac{p}{1-p} n \left(1 + \frac{p}{1-p}\right)^{n-1} = \sum_{j=0}^{n} \binom{n}{j} j \left(\frac{p}{1-p}\right)^j$

$$\implies \frac{p}{1-p} n \left(\frac{1}{1-p}\right)^{n-1} = \sum_{j=0}^{n} \binom{n}{j} j \left(\frac{p}{1-p}\right)^j$$

$$\implies np = (1-p)^n \sum_{j=0}^{n} \binom{n}{j} j \left(\frac{p}{1-p}\right)^j = \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j} = E(X)$$

- $X \sim \mathrm{GEOM}\,(p) \implies E(X) = \sum_{k=1}^{\infty} k \cdot p \cdot (1-p)^{k-1}$

$$\forall x : |x| < 1 : \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \implies \frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1} \implies E(X) = p \frac{1}{p^2} = \frac{1}{p}$$

- $X \sim \mathrm{POISSON}\,(\lambda) \implies E(X) = \sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} e^{-\lambda} = \lambda$

$$\implies E(X) = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda$$

Ejemplo 1.11. Sea X una v.a.d. con $\forall k \geq 0 : P(X = x) = \frac{6}{\pi^2} \frac{1}{k^2}$

$$\implies E(X) = \sum_{k=1}^{\infty} k \frac{6}{\pi^2} \frac{1}{k^2} = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k} \text{ que diverge a } \infty$$

Ejemplo 1.12. Sea X una v.a.d. que toma valores en $\{(-1)^{k+1}k : k \geq 1\} = \{1, -2, 3, -4, \dots\}$

$$P(X = (-1)^{k+1}k) = \frac{6}{\pi^2}\frac{1}{k^2} \implies E(X) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \text{ que sabemos que tiende a } \ln 2$$

Sin embargo, la serie no converge absolutamente, por tanto, mediante argumentos de reordenación, se puede argumentar que $E(X)$ toma cualquier valor real. Entonces $E(X)$ no tiene sentido.

Teorema 1.3. *Sea X una v.a.d. que toma los valores x_j y probabilidades p_j con $j \geq 1$. Sea $g\colon \mathbb{R} \longrightarrow \mathbb{R}$ una función.*

$$\implies E(g(X)) = \sum_{j\geq 1} g(x_j)p_j$$

Demostración. POR REVISAR

$$\sum_{j\geq 1} g(x_j)p_j = \sum_{j\geq 1} g(x_j)P(X = x_j) = \sum_{j\geq 1} g(x_j) \sum_{\substack{\omega\in\Omega \\ X(\omega)=x_j}} P(\{\omega\}) = \sum_{\omega\in\Omega} g(X(\omega))P(\{\omega\}) = E(g(X))$$

■

Observación 1.4.

1. Si X es tal que $P(X = a) = 1 \implies E(X) = a$
2. $X \sim \text{UNIF}\left(\{-1, 0, 1\}\right) \wedge Y = X^2 \implies E(X) = 0 \wedge E(Y) = 2/3$
3. $E(aX + b) = aE(X) + b$ porque
$$\sum_{j\geq 1}(ax_j + b)p_j = a\sum_{j\geq 1} x_jp_j + b\sum_{j\geq 1} p_j = aE(X) + b$$
4. En general $E(g(X)) \neq g(E(X))$

Ejemplo 1.13 ($X \sim \text{POISSON}(\lambda) \implies E(X) = \lambda$). Si $Y = g(X) = e^X$

$$\implies E(Y) = E(e^Y) = \sum_{j=0}^{\infty} e^j \frac{\lambda^j}{j!}e^{-\lambda} = e^{-\lambda}\sum_{j=0}^{\infty} \frac{(e\lambda)^j}{j!} = e^{-\lambda}e^{e\lambda} = e^{\lambda(e-1)} \neq e^{\lambda}$$

Definición 1.9. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y X una v.a.d. con función de masa p_X, $V(X)$ es la varianza de X

$$\iff V(X) = E\left[(X - E(X))^2\right]$$

Si X toma valores $x_1, x_2, \dots$ con probabilidades $p_1, p_2, \dots$ y denominamos $\mu := E(X)$

$$\implies V(X) = \sum_{j\geq 1}(x_j - \mu)^2 \cdot p_j = \sum_{j\geq 1}(x_j - \mu)^2 \cdot p_j = \sum_{j\geq 1}(x_j^2 - 2x_j\mu + \mu^2) \cdot p_j$$

$$\implies V(X) = \sum_{j\geq 1} x_j^2 p_j - 2\sum_{j\geq 1} x_j \mu p_j + \sum_{j\geq 1} \mu^2 p_j$$

Observación 1.5.

1. Es medidad de dispersión de X alrededor de μ.
2. $V(X) \geq 0$
3. $V(X) = 0 \implies P(X = E(X)) = 1$
4. $V(aX+b) = E[(aX+b-E(aX+b))^2] = E[(aX-aE(X))^2] = a^2V(X)$
5. Las unidades de $V(X)$ son las de X^2
$$\implies \text{definimos la desviación típica de } X \text{ como } \sigma(X) := \sqrt{V(X)}$$
6. ¿Por qué no $E(|X-E(X)|)$?
Porque el valor absoluto no es diferenciable y no se puede trabajar con él.

Ejemplo 1.14.

1. $X \sim \text{BER}\,(p) \implies E(X) = p \wedge V(X) = p - p^2 = p(1-p)$
2. $X \sim \text{UNIF}\,(\{1,\dots,N\}) \implies E(X) = \dfrac{N+1}{2} \wedge V(X) = \dfrac{N^2-1}{12}$
$$\implies V(X) = \frac{1}{N}\sum_{j=1}^{N} j^2 = \frac{1}{N}\cdot\frac{N(N+1)(2N+1)}{6} = \frac{(N+1)(N-1)}{12}$$
3. $X \sim \text{BIN}\,(n,p) \implies E(X) = np \wedge V(X) = np(1-p)$

   Demostración. ■
4. $X \sim \text{GEOM}\,(p) \implies E(X) = \dfrac{1}{p} \wedge V(X) = \dfrac{1-p}{p^2}$

   Demostración. ■
5. $X \sim \text{POISSON}\,(\lambda) \implies E(X) = \lambda \wedge V(X) = \lambda$

   Demostración.
$$E(X^2) = \sum_{k=0}^{\infty} k^2 e^{-k}\frac{\lambda^k}{k!}$$
■

Definición 1.10 (Momentos de X). Sea X una v.a.d. con función de masa p_X, μ_k es el k-ésimo momento de $X \iff \mu_k = E\left((X-E(X))^k\right)$

Observación 1.6.

1. $\mu_1 = 0$
2. $\mu_2 = V(X)$
3. μ_3 es la asimetría de X
4. μ_4 es la curtosis de X

Teorema 1.4 (Desigualdad de Markov)**.** *Sea X una v.a.d. "no negativa" ($P(X < 0) = 0$) y con $E(X)$ finita.*

$$\implies \forall t > 0 : P(X \geq t) \leq \frac{E(X)}{t}$$

Demostración. **Notación:** $(\Omega, \mathcal{F}, P), A \subset \mathcal{F}\ \mathbb{1}_A \longrightarrow \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{cases}$

Fijamos $t > 0$ y definimos $Y_t = t \cdot \mathbb{1}_{\{x \geq t\}} = \begin{cases} t, P(x \geq t) \\ 0, 1 - P(x \geq t) \end{cases}$

Observación 1.7. $\forall \omega : Y_t(\omega) \leq X(\omega) \implies E(Y_t) = t \cdot P(X \geq t) \leq E(X)$

Así que $\forall t > 0 : P(X \geq t) \leq \dfrac{E(X)}{t}$ ∎

Teorema 1.5 (Desigualdad de Chebyshev)**.** *En $(\Omega, \mathcal{F}, P)$, sea X una v.a.d. con $E(X)$ y $V(X)$ finitas.*

$$\implies \forall \lambda > 0 : P(|X - E(X)| \geq \lambda \cdot \sigma(X)) \leq \frac{1}{\lambda^2}$$

$$\iff P\left(|X - E(X)| \geq \alpha\right) \leq \frac{V(X)}{\alpha^2}$$

Demostración. Definimos $Y = |X - E(X)|^2$ y aplicamos la desigualdad de Markov.

$$\implies \forall t > 0 : P(Y \geq t) \leq \frac{E(Y)}{t} \implies \forall t > 0 : P(|X - E(X)|^2 \geq t) \leq \frac{E(Y)}{t}$$

Como $E(Y) = E(|X - E(X)|^2) = V(X)$ por la def de varianza,

$$\implies \forall t > 0 : P\left(|X - E(X)| \geq \sqrt{t}\right) \leq \frac{V(X)}{t}$$

Definimos $\alpha := \sqrt{t} \implies \forall \alpha > 0 : P\left(|X - E(X)| \geq \alpha\right) \leq \dfrac{V(X)}{\alpha^2}$

y para la desigualdad equivalente definimos $\lambda := \dfrac{\alpha}{\sigma(X)}$

$$\implies \forall \lambda > 0 : P\left(|X - E(X)| \geq \lambda \cdot \sigma(X)\right) \leq \frac{1}{\lambda^2}$$

$\blacksquare$

1.3.3 Esperanza condicionada

Definición 1.11 (Esperanza condicionada). Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad $B \in \mathcal{F}$ un suceso tal que $P(B) > 0$ y X una v.a.d. con esperanza $E(X)$, $E(X|B)$ es la esperanza de X condicionada a B

$$\iff E(X|B) = \sum_{x \in X(\Omega)} x \cdot P(X = x|B) = \sum_{x \in X(\Omega)} x \cdot \frac{P(\{X = x\} \cap B)}{P(B)}$$

(Siempre que la serie sea absolutamente convergente).

Teorema 1.6 (Esperanza total)**.** *En $(\Omega, \mathcal{F}, P)$, sea X una v.a.d. y $\{B_1, B_2, \dots\}$ una partición de Ω*

$$\implies E(X) = \sum_{i \geq 1} E(X|B_i) \cdot P(B_i)$$

(Siempre que la serie sea absolutamente convergente).

Demostración. POR REVISAR

$$\sum_{i \geq 1} E(X|B_i) \cdot P(B_i) = \sum_{i \geq 1} \sum_{x \in X(\Omega)} x \cdot P(X = x|B_i) \cdot P(B_i) = \sum_{x \in X(\Omega)} x \cdot \sum_{i \geq 1} P(X = x|B_i) \cdot P(B_i)$$

$$\sum_{i \geq 1} P(X = x|B_i) \cdot P(B_i) = P(X = x)$$

$\blacksquare$

Ejemplo 1.15. Lanzamos una moneda con probabilidad p de cara y $1-p$ de cruz y definimos X como la longitud de la racha inicial, i.e. el número de caras/cruces consecutivas.

$$\implies E(X) = E(X|C) \cdot p + E(X|\times) \cdot (1-p)$$

$$\implies E(X) = \left(\sum_{j \geq 1} j \cdot P(X = j|C) \right) p + \left(\sum_{j \geq 1} j \cdot P(X = j|\times) \right) (1-p)$$

$$\implies E(X) = p \sum_{k=1}^{\infty} k p^{k-1} (1-p) + (1-p) \sum_{k=1}^{\infty} k (1-p)^{k-1} p$$

$$\implies E(X) = \frac{1}{1-p} \cdot p + \frac{1}{p} \cdot (1-p) = \frac{1}{p(1-p)} - 2$$

También se puede abordar el problema pensando en las variables geométricas.

1.4 Varias variables aleatorias

En $(\Omega, \mathcal{F}, P)$, sea $\mathcal{X} = X_1, X_2, \ldots, X_n$ una colección de variables aleatorias discretas.

En el caso $n = 2$, tenemos X e Y variables aleatorias discretas, se genera una tabla con las probabilidades conjuntas:

$$p_{X,Y}(x, y) := P(\omega \in \Omega : X(\omega) = x \wedge Y(\omega) = y) = P(X = x \wedge Y = y)$$

Definición 1.12 (Función de masa conjunta). En $(\Omega, \mathcal{F}, P)$, sean X e Y v.a.d., $p_{X,Y}$ es su función de masa conjunta

$$\iff p_{X,Y} \colon \mathbb{R}^2 \longrightarrow [0, 1] \wedge \forall (x, y) \in X(\Omega) \times Y(\Omega) : p_{X,Y}(x, y) = P(X = x \wedge Y = y)$$

tal que $\displaystyle\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} p_{X,Y}(x, y) = 1$

Ya tenemos X e Y en $(\Omega, \mathcal{F}, P)$ con $P_{X,Y}$

1. ¿Qué sabemos de X e Y por separado?
2. Esperanzas: nos interesa calcular $E(X), E(Y), E(X + Y), E(X \cdot Y)$
3. Independencia

Definición 1.13 (Funciones marginales). En $(\Omega, \mathcal{F}, P)$, sean X e Y v.a.d., p_X, p_Y son sus funciones de masa marginales

$$\iff p_X(x) = \sum_{y \in Y(\Omega)} P_{X,Y}(x, y) \wedge p_Y(y) = \sum_{x \in X(\Omega)} P_{X,Y}(x, y)$$

Teorema 1.7. *En $(\Omega, \mathcal{F}, P)$, sean X e Y v.a.d. y sea $g \colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ una función*

$$\implies E(g(X, Y)) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} g(x, y) \cdot P_{X,Y}(x, y)$$

Si converge absolutamente.

Demostración. ■

Observación 1.8. Si $g \colon \mathbb{R} \longrightarrow \mathbb{R}$

$$\implies E(g(x)) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} g(x) \cdot P_{X,Y}(x, y) = \sum_{x \in X(\Omega)} \left(g(x) \sum_{y \in Y(\Omega)} P_{X,Y}(x, y) \right)$$

$$\implies E(g(X)) = \sum_{x \in X(\Omega)} g(x) p_X(x)$$

De manera análoga, $E(g(Y)) = \sum_{y \in Y(\Omega)} g(y) p_Y(y)$

Ejemplo 1.16 **($E(aX+bY) = aE(X) + bE(Y)$).**

$$E(aX+bY) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (ax+by) \cdot P_{X,Y}(x,y)$$

$$\implies E(aX+bY) = a \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} x \cdot P_{X,Y}(x,y) + b \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} y \cdot P_{X,Y}(x,y)$$

Definición 1.14 (Independencia de v.a.d.). En $(\Omega, \mathcal{F}, P)$, sean X e Y v.a.d., X e Y son independientes

$$\iff \forall (x,y) \in X(\Omega) \times Y(\Omega) : X = x \text{ y } Y = y \text{ son sucesos independientes.}$$

$$\iff \forall (x,y) \in X(\Omega) \times Y(\Omega) : P(X = x \wedge Y = y) = P(X = x) \cdot P(Y = y)$$

Teorema 1.8. *En $(\Omega, \mathcal{F}, P)$, sean X e Y v.a.d., X e Y son independientes*

$$\iff \exists\, g, h \colon \mathbb{R} \longrightarrow \mathbb{R} : P_{X,Y}(x,y) = g(x) \cdot h(y)$$

Ejemplo 1.17.

$$P_{X,Y}(x,y) = e^{-(\lambda+\mu)} \frac{\lambda^x \mu^y}{x!y!} \text{ con } x, y \in \mathbb{Z} \text{ y } \lambda, \mu > 0$$

$\implies$ Se puede interpretar como $X \sim \text{POISSON}(\lambda)$ e $Y \sim \text{POISSON}(\mu)$ independientes.

Observación 1.9. Si X e Y son independientes $\implies E(X \cdot Y) = E(X) \cdot E(Y)$.
Sin embargo, la implicación recíproca no es cierta. **(Motivo de excomunión)**

Definición 1.15 (Covarianza). En $(\Omega, \mathcal{F}, P)$, sean X e Y v.a.d., $\operatorname{cov}(X,Y)$ es la covarianza de X e Y

$$\iff \operatorname{cov}(X,Y) = E\left[(X - E(X)) \cdot (Y - E(Y))\right] = E(X \cdot Y) - E(X) \cdot E(Y)$$

29/02/2024

Observación 1.10.

1. Cálculo de la covarianza

$$\operatorname{cov}(X,Y) = \sum_{i,j} x_i y_j P(X = x_i \wedge Y = y_j) - \left(\sum_i x_i P(X = x_i)\right)\left(\sum_j y_j P(Y = y_j)\right)$$

2. Signo de la covarianza (y coeficiente de correlación)

$$\operatorname{cov}(X,Y) = \sum_{i,j}(x_i - E(X))(y_j - E(Y))P(X = x_i \wedge Y = y_j)$$

   Entonces, si la covarianza es positiva, X e Y tienden a crecer juntas. Si es negativa, tienden a decrecer juntas. Si es 0, no hay relación lineal entre X e Y.

3. Cálculo fundamental

$$\begin{aligned} V(X+Y) &= E((X+Y)^2) - (E(X+Y))^2 \\ &= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - (E(X))^2 - 2E(Y)E(X) - (E(Y))^2 \\ &\implies V(X+Y) = V(X) + V(Y) + 2\operatorname{cov}(X,Y) \end{aligned}$$

   Pero cuidado: $V(X-Y) = V(X) + V(Y) - 2\operatorname{cov}(X,Y)$

   De forma más general:

$$\begin{aligned} V(aX+bY) &= V(aX) + V(bY) + 2\operatorname{cov}(aX,bY) \\ &= a^2V(X) + b^2V(Y) + 2ab\operatorname{cov}(aX,bY) \end{aligned}$$

$$\implies \boxed{V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\operatorname{cov}(aX,bY)}$$

4. Si X e Y son independientes $\implies \operatorname{cov}(X,Y) = 0$

Definición 1.16 (coeficiente de correlación).

$$\rho := \frac{\operatorname{cov}(X,Y)}{\sqrt{V(X)\cdot V(Y)}}$$

$$\implies |\rho(X,Y)| \leq 1$$

Proposición 1.7 (Desigualdad de Cauchy-Schwarz)**.**

$$E(X\cdot Y)^2 \leq E(X^2)\cdot E(Y^2)$$

La igualdad se da cuando una variable es transformación lineal de la otra, i.e. $Y = aX + b$.

Demostración*.* Definimos $W = sX + Y, W^2 \geq 0$ con probabilidad 1.

$$\begin{aligned} 0 \leq E\left(W^2\right) &= E\left((sX+Y)^2\right) = E\left(s^2X^2 + Y^2 + 2sXY\right) \\ &= s^2E\left(X^2\right) + E\left(Y^2\right) + 2sE(XY) \\ &= E(X^2)\cdot s^2 + 2E(XY)\cdot s + E(Y^2) \end{aligned}$$

$$\implies 4E(XY)^2 - 4E\left(X^2\right)E\left(Y^2\right) \leq 0 \implies E(XY)^2 \leq E\left(X^2\right)E\left(Y^2\right)$$

$\blacksquare$

Por tanto,

$$\operatorname{cov}(X,Y)^2 = E((X-E(X))(Y-E(Y)))^2 \leq V(X)\cdot V(Y)\operatorname{cov}(X,Y)$$

04/03/2024

05/03/2024

Dato: $(a_n)_{n=0}^{\infty}\ x \longmapsto \displaystyle\sum_{n=0}^{\infty} a_n x^n$

Existe $R : \frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n} : |x| < R \implies$ converge $\wedge$ $|x| > R \implies$ no converge.

Ejemplo 1.18.

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x \wedge \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, |x| < 1 \wedge \sum_{n=0}^{m} \binom{m}{n} x^n = (1+x)^m$$

1.4.1 Funciones generatrices

$$(a_n)_{n=0}^{\infty} \longleftrightarrow f(x) = \sum_{n=0}^{\infty} a_n x^n \implies \begin{cases} x\cdot f(x) \longleftrightarrow (0, a_0, \dots) \\ x\cdot f'(x) \longleftrightarrow (0a_0, 1a_1, 2a_2, \dots) \end{cases}$$

$$f(x)\cdot g(x) = \sum_{n=0}^{\infty} a_n x^n \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) x^n$$

1.4.2 Funciones generatrices de probabilidad

Definición 1.17. Sea X una variable aleatoria discreta que toma valores en $\{0,1,2,\dots\}$ donde $\forall j \geq 0 : p_j = P(X=j)$, $G_X(s)$ es su función generatriz de probabilidad

$$\iff G_X(s) = \sum_{n=0}^{\infty} p_n s^n$$

Ejemplo 1.19.

1. $X \sim \operatorname{BER}(p) \implies G_X(s) = (1-p) + ps$
2. $X \sim \operatorname{BIN}(n,p) \implies G_X(s) = \displaystyle\sum_{j=1}^{n} \binom{n}{j} (1-p)^{n-j} (ps)^j = (1-p+ps)^n$

3. $X \sim \text{GEOM}\,(p) \implies G_X(s) = \sum_{j=1}^{\infty}(1-p)^{j-1}ps^j = \dfrac{ps}{1-(1-p)s}$

4. $X \sim \text{POISSON}\,(\lambda) \implies G_X(s) = \sum_{j=0}^{\infty} e^{-\lambda}\dfrac{\lambda^j}{j!}s^j = e^{\lambda(s-1)}$

¿Para qué?

1. Cálculo de momentos con $(p_n)_{n=0}^{\infty}$

$$G_X(s) = \sum_{n=0}^{\infty} p_n s^n \implies G'_X(s) = \sum_{n=1}^{\infty} np_n s^{n-1} \implies G'_X(1) = \sum_{n=1}^{\infty} np_n = E(X)$$

Si seguimos derivando, obetenemos

$$\implies G''_X(s) = \sum_{n=2}^{\infty} n(n-1)p_n s^{n-2} = \sum_{n=2}^{\infty} n^2 p_n s^{n-2} - \sum_{n=2}^{\infty} np_n s^{n-2}$$

$$\implies G''_X(1) = \sum_{n=2}^{\infty} n^2 p_n - \sum_{n=2}^{\infty} np_n = E(X^2) - E(X)$$

$$\implies V(X) = E(X^2) - E(X)^2 = G''_X(1) + G'_X(1)\,(1 - G'_X(1))$$

Ejemplo 1.20.

(a) $X \sim \text{BER}\,(p) \implies G_X(s) = (1-p) + ps$

$$\implies G'_X(s) = p = E(X) \wedge G''_X(s) = 0 \implies V(X) = p(1-p)$$

(b) $X \sim \text{BIN}\,(n,p) \implies G_X(s) = (1-p+ps)^n$

$$\implies G'_X(s) = n(1-p+ps)^{n-1}p \implies G'_X(1) = np = E(X)$$

$$\implies G''_X(s) = n(n-1)(\cdots)^{n-2}p^2 \implies G''_X(1) = n(n-1)p^2$$

$$\implies V(X) = n(n-1)p^2 + np(1-np) = np(1-p)$$

2. Suma de independientes

Teorema 1.9. *Sean X, Y dos v.a.d. independientes con valores en $\{0,1,2,\dots\}$ y con funciones generatrices de probabilidad $G_X(s), G_Y(s)$ respectivamente*

$$\implies G_{X+Y}(s) = G_X(s)\cdot G_Y(s)$$

***Demostración*.**

$$G_{X+Y}(s) = \sum_{n=0}^{\infty} P(X+Y=n)s^n = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} P(X=k \wedge Y=n-k)\right)s^n$$

Por otro lado,

$$\left.\begin{aligned} G_X(s)\cdot G_Y(s) &= \left(\sum_{n=0}^{\infty} P(X=n)s^n\right)\left(\sum_{n=0}^{\infty} P(Y=n)s^n\right) \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} P(X=k)P(Y=n-k)\right)s^n \end{aligned}\right. \implies G_{X+Y}(s) = G_X(s)\cdot G_Y(s)$$

■

Otra manera:

$$G_X(s) = E(s^X) \wedge G_{X+Y}(s) = E\left(s^{X+Y}\right) = E\left(s^X\cdots^Y\right) = E(s^X)\cdot E(s^Y) = G_X(s)\cdot G_Y(s)$$

■

Corolario 1.1. Sean $X_1, X_2, \ldots, X_n$ v.a.d. independientes con valores en $\{0,1,2,\ldots\}$ y con funciones generatrices de probabilidad $G_{X_1}(s), G_{X_2}(s), \ldots, G_{X_n}(s)$ respectivamente

$$\implies G_{X_1+X_2+\cdots+X_n}(s) = \prod_{i=1}^{n} G_{X_i}(s)$$

Si las X_i son "idénticas"

$$\implies G_{X_1+X_2+\cdots+X_n}(s) = \left(G_{X_1}(s)\right)^n$$

06/03/2024

Teorema 1.10 (Unicidad)**.** *Sean X, Y dos v.a.d. con valores en $\{0,1,2,\ldots\}$ y con funciones generatrices de probabilidad $G_X(s), G_Y(s)$ respectivamente*

$$\implies G_X(s) = G_Y(s) \iff \forall n \geq 0 : P(X=n) = P(Y=n)$$

Ejemplo 1.21. Sean $X \sim \text{POISSON}\,(\lambda) \wedge Y \sim \text{POISSON}\,(\mu)$ independientes con $\lambda, \mu > 0$. Definimos $Z = X + Y$.

$$\implies \forall x \geq 0 : P(Z=k) = P(X+Y=k) = \sum_{j=0}^{k} P(X=j \wedge Y=k-j) = \cdots$$

Pero, a través de funciones generatrices obetenemos:

$$G_X(s) = e^{\lambda(s-1)} \wedge G_Y(s) = e^{\mu(s-1)} \implies G_Z(s) = e^{(\lambda+\mu)(s-1)} \implies Z \sim \text{POISSON}\,(\lambda+\mu)$$

Ejemplo 1.22.

1. Sean $I_1, I_2, \ldots, I_n$ v.a.d. independientes con $\forall k \in \mathbb{N}_n : I_k \sim \text{BER}\,(p)$ y definimos $Z = I_1 + I_2 + \cdots + I_n$.

$$\implies G_Z(s) = \left[(1-p)+ps\right]^n \implies Z \sim \text{BIN}\,(n,p)$$

2. Sean $X \sim \text{BIN}\,(n, p) \,\wedge\, Y \sim \text{POISSON}\,(\lambda)$ independientes y definimos $Z = X + Y$.

$$\implies G_Z(s) = ((1-p)+ps)^n \cdot e^{\lambda(s-1)}$$