1 Probabilidad. Definición y enunciados

1. Axiomas de probabilidad: Dados un experimento, y S un espacio muestral, se asigna a cada evento A una probabilidad $P(A)$ tal que:

    - $P(A) \in [0,1]$
    - $P(S) = 1$
    - Si $A_1, A_2, \cdots$ disjuntos, $P(\cup_{i=0}^{\infty} A_i) = \sum_{i=0}^{\infty} P(A_i)$

    (a)
    $$A \cup A^c = S$$
    y aplicando probabilidad a ambos lados,
    $$P(A \cup A^c) = P(S)$$
    $$P(A) + P(A^c) = 1$$
    pues A y A^c son disjuntos, y $P(S) = 1$.
    $$P(A) = 1 - P(A^c)$$

    (b)
    $$(A \setminus B) \cup B = A$$
    porque $B \subset A$. Aplicando probabilidad:
    $$P((A \setminus B) \cup B) = P(A)$$
    $$P(A \setminus B) + P(B) = P(A)$$
    pues B y $A \setminus B$ son disjuntos.
    $$P(A \setminus B) = P(A) - P(B)$$

2.

3.

4.

5.

2 Probabilidad condicional e independencia

6. Sea E = la persona está enferma, y V = la persona está vacunada. La probabilidad pedida es $P(\overline{V}|E)$.

$$P(\overline{V}|E) = \frac{P(\overline{V} \wedge E)}{P(E)} = \frac{0.13}{0.13 + 0.02} = \frac{13}{15}$$

7. Asumimos que la probabilidad de que un hijo cualquiera sea varón es de $\frac{1}{2}$. Sea V_1 = el primer hijo es varón, y V_2 = el segundo hijo es varón. La probabilidad pedida es

$$P(V_2|V_1) = \frac{P(V_1 \wedge V_2)}{P(V_1)} = \frac{1/4}{1/2} = \frac{1}{2}$$

8. Dadas las mismas suposiciones, la probabilidad pedida es:

$$P(V_1 \wedge V_2|V_1 \vee V_2) = \frac{P(V_1 \wedge V_2)}{P(V_1 \vee V_2)} = \frac{1/4}{1 - 1/4} = \frac{1}{3}$$

9. Notamos cada caso como una tupla de dos letras (M o V), y ponemos un asterisco antes de la letra del hijo que abrió la puerta.

   Los casos posibles (que corresponden al evento en que abre la puerta un varón) son: $\{(V * V), (*VV), (*VM), (M * V)\}$.

   De estos, los favorables son los dos primeros (que corresponden al evento en que hay dos hijos varones). Por lo tanto la probabilidad es $\frac{2}{4} = \frac{1}{2}$.

10. Sea $B_k = A_1 A_2 \cdots A_k$. Notar que $B_k = A_k B_{k-1}$ para $k > 1$.

    Los voy a demostrar por inducción en n. La propiedad se puede enunciar como

$$P(B_n) = P(A_1)P(A_2|B_1) \cdots P(A_n|B_{n-1})$$

    Caso base: $B_1 = A_1$. Esto ocurre por definición.

    Paso inductivo: vale la propiedad hasta B_{n-1} y quiero demostrarla para B_n.

$$P(A_n|B_{n-1}) = \frac{P(A_n B_{n-1})}{P(B_{n-1})} = \frac{P(B_n)}{P(B_{n-1})}$$
$$P(B_n) = P(B_{n-1})P(A_n|B_{n-1})$$

    Pero por HI,

$$P(B_{n-1}) = P(A_1)P(A_2|B_1) \cdots P(A_{n-1}|B_{n-2})$$

    Luego tenemos

$$P(B_n) = (P(A_1)P(A_2|B_1) \cdots P(A_{n-1}|B_{n-2}))P(A_n|B_{n-1})$$
$$P(B_n) = P(A_1)P(A_2|B_1) \cdots P(A_n|B_{n-1})$$

    Q.E.D.

11.

12.

13.

14.

15.

3 Convergencia en Distribución

26. Sea $S_n \sim \text{Bi}(n, \lambda/n)$.

$$P(S_n = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$\begin{aligned} lim_{n\to\infty} P(S_n = k) &= lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} lim_{n\to\infty} \frac{n!}{(n-k)!}\left(\frac{1}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} lim_{n\to\infty} \frac{n!}{(n-k)!n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

Veamos ambos factores del límite por separado:

$$lim_{n\to\infty} \frac{n!}{(n-k)!n^k} = lim_{n\to\infty} \prod_{i=n-k}^{n} \frac{i}{n} \quad > lim_{n\to\infty} \prod_{i=n-k}^{n} \frac{n-k}{n} = 1$$

Como la expresión está acotada por 1, tiende a 1.

$$\begin{aligned} lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} &= lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{\frac{n}{-\lambda}(n-k)\frac{-\lambda}{n}} \\ &= e^{-\lambda} \end{aligned}$$

Ponemos todo junto:

$$lim_{n\to\infty} P(S_n = k) = \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda} = \frac{\lambda^k e^{-\lambda}}{k!}$$

Que es la función de probabilidad puntual de la Poisson.

27. Sea $U_n \sim \text{Uniforme}\{\frac{1}{n}, \cdots, \frac{n}{n}\}$.
Para cada par $n \in \mathbb{N}$, $k \in [0, 1]$, sea $x_{n,k} \in \mathbb{N}$ el mayor número tal que $\frac{x_{n,k}}{n} \leq k$.

Como tenemos que:

$$F_{U_n}(k) = \sum_{i=1}^{x_{n,k}} \frac{1}{n} = \frac{x_{n,k}}{n}$$

Lo que queremos demostrar es que $x_{n,k} \rightarrow_{n\rightarrow\infty} k$, para que U se vaya a una uniforme. Pero tenemos que dados k y n,

$$0 \leq \left(k - \frac{x_{n,k}}{n}\right) \leq \frac{1}{n}$$

Por lo que cuando $n \rightarrow \infty$, la expresión del medio se va a 0. Finalmente $\frac{x_{n,k}}{n} \rightarrow_{n\rightarrow\infty} k$, como queríamos demostrar.

28. Sean $Y_n \sim \text{Geom}(\frac{\lambda}{n})$.

    Vamos a demostrar la convergencia para $P\left(\frac{Y_n}{n} > y\right)$ porque la cuenta queda fácil así.

    Para las geométricas:

$$P\left(\frac{Y_n}{n} > y\right) = P(Y_n > n \cdot y) = \left(1 - \frac{\lambda}{n}\right)^{ny}$$

    Queremos saber:

$$\begin{aligned} lim_{n\rightarrow\infty}\left(1 - \frac{\lambda}{n}\right)^{ny} &= lim_{n\rightarrow\infty}\left(1 + \frac{-\lambda}{n}\right)^{\frac{n}{-\lambda}\frac{-\lambda}{n}ny} \\ &= lim_{n\rightarrow\infty}\left(1 + \frac{-\lambda}{n}\right)^{\frac{n}{-\lambda}(-\lambda y)} \\ &= e^{-\lambda y} \end{aligned}$$

    Con lo que $F_{\frac{Y}{n}}(y) = 1 - P\left(\frac{Y}{n} > y\right) = 1 - e^{-\lambda y}$, que es la F de una V.A. exponencial.

29. Sea $X_n = c + \frac{1}{n}$. Por un lado, $X_n > c$ siempre con lo que $F_{X_n}(c) = 0$. Por otro es claro que $X_n \rightarrow c$.

4 Vectores Aleatorios

30. PREGUNTAR

31. Sean Z_1 y Z_2 constantes de normalización de f_{X_1} y f_{X_2} respectivamente. $Z = Z_1 \cdot Z_2$.

$$P(X_1 = x_1 \wedge X_2 = x_2) = \frac{1}{Z} a^{x_1+x_2} = \frac{1}{Z_1} a^{x_1} \cdot \frac{1}{Z_2} a^{x_2}$$

Por otro lado,

$$\begin{aligned} P(X_1 = x_1) &= \sum_{x_2=1}^{\infty} \frac{1}{Z_1} a^{x_1} \cdot \frac{1}{Z_2} a^{x_2} \\ &= \frac{1}{Z_1} a^{x_1} \sum_{x_2=1}^{\infty} \frac{1}{Z_2} a^{x_2} \\ &= \frac{1}{Z_1} a^{x_1} \cdot 1 \end{aligned}$$

(vale 1 porque es la sumatoria sobre toda la densidad de X_2).

Análogamente, $P(X_2 = x_2) = \frac{1}{Z_2} a^{x_2}$. Como el producto de las densidades da la densidad conjunta, son independientes.

32.

$$P(X_1 \geq x_1 \wedge X_2 \geq x_2) = a^{x_1+x_2}$$
$$P(X_1 \geq x_1) = P(X_1 \geq x_1 \wedge X_2 \geq 0) = a^{x_1}$$

Análogamente $P(X_2 \geq x_2) = a^{x_2}$.

33.

($1 \Rightarrow 2$) Es fácil porque se define $g(x) = f_X(x)$ y $h(y) = f_Y(y)$.

($1 \Rightarrow 3$) Es lo mismo que el anterior, se define $G(x) = F_X(x)$ y $H(y) = F_Y(y)$.

($2 \Rightarrow 1$)

$$\begin{aligned} f_{X,Y}(x,y) &= g(x)h(y) \\ \int f_{X,Y}(x,y)dy &= \int g(x)h(y)dy \\ f_X(x) &= g(x) \int h(y)dy \end{aligned}$$

Es decir que $g(x) = C_x \cdot f_X(x)$ donde C_x es una constante.

Análogamente $h(y) = C_y \cdot f_Y(y)$ con C_y constante.

Entonces:

$$f_{X,Y}(x,y) = C_x C_y f_X(x) f_Y(y)$$

Y la constante vale 1 porque si integramos de los dos lados, $f_{X,Y}(x,y)$, $f_X(x)$ y $f_Y(y)$ son todas densidades e integran a 1.

Finalmente $f_{X,Y}(x,y) = f_X(x) f_Y(y)$, con lo que son independientes.

$(3 \Rightarrow 2)$

$$\begin{aligned}
F_{X,Y}(x,y) &= G(x)H(y) \\
\frac{\delta F_{X,Y}(x,y)}{\delta x} &= G'(x)H(y) \\
\frac{\delta^2 F_{X,Y}(x,y)}{\delta x \delta y} &= G'(x)H'(y) \\
f_{X,Y}(x,y) &= G'(x)H'(y)
\end{aligned}$$

Entonces basta definir $g(x) = G'(x)$ y $h(y) = H'(y)$ para demostrar (2).

34. $x_1 \cdots x_n \cdots \geq 0$ independientes con media μ.

    Sea $N \in \{1 \cdots k\}$ (supongo que equiprobable) independiente de todos los x_i.

$$\begin{aligned}
E(\prod_{i=1}^{N} x_i) &= \sum_{n=1}^{k} \frac{1}{k} E(\prod_{i=1}^{n} x_i) \\
&= \frac{\sum_{n=1}^{k} \prod_{i=1}^{n} E(x_i)}{k} \\
&= \frac{\sum_{n=1}^{k} (\mu^n)}{k} \\
&= \frac{\mu^{k+1} - 1}{(\mu - 1)k}
\end{aligned}$$

    En el primer paso, hago suma sobre todos los casos para N, por la probabilidad de cada uno (que es $\frac{1}{k}$). Puedo sacar la productoria para afuera en el paso 2 por linealidad de la esperanza.

    El último paso es la resolución analítica de esa sumatoria.

35. $$P(R \leq a) = P(X^2 + Y^2 \leq a^2)$$

    Sean r, θ tales que $X = r \cdot sin(\theta)$ e $Y = r \cdot cos(\theta)$. Tenemos que $r^2 = X^2 + Y^2$.

$$P(R \leq a) = P(r^2 < a^2) = P(|r| < a)$$

suponiendo que $a > 0$.

Quiero hallar:

$$\int\int_D \frac{1}{2\pi} e^{\frac{-(x^2+y^2)}{2}} dy\ dx$$

para $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq a^2\}$.

Esto es lo mismo que integrar sobre:

$$\begin{aligned}
\int\int_D \frac{1}{2\pi} e^{\frac{-(x^2+y^2)}{2}} dy\ dx &= \int_0^a \int_0^{2\pi} \frac{1}{2\pi} e^{\frac{-r^2}{2}} \cdot r d\theta\ dr \\
&= \int_0^a \left(\frac{1}{2\pi} e^{\frac{-r^2}{2}} \cdot r \right) \cdot 2\pi\ dr \\
&= \int_0^a (e^{\frac{-r^2}{2}} \cdot r) dr \\
&= -\int_0^a ((-r) \cdot e^{\frac{-r^2}{2}}) dr \\
&= -(e^{\frac{-a^2}{2}} - e^{\frac{-0^2}{2}}) \\
&= 1 - e^{\frac{-a^2}{2}}
\end{aligned}$$

36.

$$F_{X+Y}(z) = P(X + Y \leq z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x} f_{X,Y}(x, y)\ dy\ dx$$

Sean $u = x$, $v = x + y$:

$$\begin{aligned}
F_{X+Y}(z) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{z} f_{X,Y}(u, v - u)\ dv\ du \\
&= \int_{-\infty}^{z} \left(\int_{-\infty}^{+\infty} f_{X,Y}(u, v - u)\ du \right)\ dv
\end{aligned}$$

Luego, lo de adentro del paréntesis es la densidad. Entonces:

$$f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_{X,Y}(u, z - u)\ du$$

37. Si reemplazamos la densidad por el producto de las densidades en la anterior:

$$f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(u) f_Y(z - u)\ du$$

38. Caso discreto:

$$\begin{aligned}
P(g(X) = x' \wedge h(Y) = y') &= \sum_{\{x:(g(x)=x')\}} \sum_{\{y:(h(y)=y')\}} P(X = x \wedge Y = y) \\
&= \sum_{\{x:(g(x)=x')\}} \sum_{\{y:(h(y)=y')\}} P(X = x)P(Y = y) \\
&= \left(\sum_{\{x:(g(x)=x')\}} P(X = x) \right) \left(\sum_{\{y:(h(y)=y')\}} P(Y = y) \right) \\
&= P(g(X) = x')P(h(Y) = y')
\end{aligned}$$

FALTA CASO CONTINUO PERO DEBE SER IGUAL, CON MENOR O IGUAL

39. Sale haciendo el gráfico (X, Y) y pintando el área que sirve, viendo que la proba de que no se crucen es como integrar un cuadrado de $\frac{5}{6}$ de lado.

40. Sea $Y = aX + b$. Supongamos primero que $a > 0$.

$$F_{aX+b}(y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

$$f_{aX+b}(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{a}$$

Reemplazando en la densidad de la normal:

$$\begin{aligned}
f_{aX+b}(y) &= f_X\left(\frac{y-b}{a}\right)\frac{1}{a} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}} \frac{1}{a} \\
&= \frac{1}{\sqrt{2\pi\sigma^2 a^2}} e^{-\frac{\left(\frac{y-b-a\mu}{a}\right)^2}{2\sigma^2}} \\
&= \frac{1}{\sqrt{2\pi(\sigma^2 a^2)}} e^{-\frac{y-(a\mu+b)}{2(\sigma^2 a^2)}}
\end{aligned}$$

Con lo que $Y \sim N(a\mu + b, a^2\sigma^2)$.

Para el caso negativo, la ecuación con la F se da vuelta, con lo que la f queda multiplicada por -1, pero se cancela al meter el $\frac{1}{a}$ en la constante de la raíz porque $-\sqrt{a^2} = a$.

Para ver la normalización, sea $Y = X - \mu$. $Y \sim N(0, \sigma^2)$. Luego sea $Z = \frac{Y}{\sigma^2}$. $Z \sim N(0, 1)$.

41. VER DESPUES

42. Primero hay que sacar la constante de normalización: se integra sobre todo el dominio y la constante es $\frac{1}{\text{Area}}$ para que la función integre a 1.

    Después es hacer la integral doble. Para el máximo es directamente $\int_{20}^{26}\int_{20}^{26} f_{X,Y}(x,y)dy\ dx$, y para el mínimo es $1-\int_{26}^{30}\int_{26}^{30} f_{X,Y}(x,y)dy\ dx$.

5 Esperanza

43. Sea $Y = g(X)$.

$$\begin{aligned} E(Y) &= \sum_y y \cdot p(Y = y) \\ &= \sum_y \sum_{\{x:g(x)=y\}} y \cdot p(X = x) \\ &= \sum_y \sum_{\{x:g(x)=y\}} g(x) \cdot p(X = x) \\ &= \sum_x g(x) \cdot p(X = x) \end{aligned}$$

Cuando reemplazo $y = g(x)$, es importante que en cada término es potencialmente un x distinto, pero como cada x es de la preimagen de g en y, las expresiones son equivalentes.

44. (a) Caso continuo:

$$E(X) = \int_0^{+\infty} x \cdot f_X(x)$$

Hago partes usando el reemplazo $du = f_X(x)dx$, $u = -(1 - F_X(x))$ (o sea, hago aparecer lo que quiero en la integral de partes). Quedan $v = x$, $dv = dx$.

$$\begin{aligned} E(X) &= \int_0^{+\infty} x \cdot f_X(x) \\ &= -x(1 - F_X(x))\Big|_0^{\infty} - \int_0^{+\infty} -(1 - F_X(x))dx \\ &= -x(1 - F_X(x))\Big|_0^{\infty} + \int_0^{+\infty} 1 - F_X(x)dx \end{aligned}$$

Queda ver que $x(1 - F_X(x))\Big|_0^{\infty} = 0$. Para $x = 0$ se ve reemplazando, falta ver qué pasa cuando $x \to +\infty$. Pero,

$$0 \leq x(1 - F_X(x)) = x \int_x^{+\infty} f_X(u)du \leq \int_x^{+\infty} u \cdot f_X(u)du$$

Como la esperanza está acotada, la parte de la derecha tiende a 0 cuando $x \to +\infty$.

(b) Caso discreto: Sean $\{x_0, x_1, \cdots, x_n, \cdots\}$ los elementos del rango de X, en orden creciente. Sea $x_0 = 0$ (si 0 no está en el rango, lo agrego con probabilidad asociada 0 porque si no está se me rompe todo con las geométricas).

Se cumple que $p(x_i) = F(x_i) - F(x_{i-1})$.

$$
\begin{aligned}
E(X) &= \sum_{i=1}^{+\infty} x_i \cdot p(x_i) \\
&= \sum_{i=1}^{+\infty} x_i \cdot (F(x_i) - F(x_{i-1})) \\
&= \sum_{i=1}^{+\infty} \sum_{j=1}^{i} (x_j - x_{j-1}) \cdot (F(x_i) - F(x_{i-1})) \\
&= \sum_{j=1}^{+\infty} \sum_{i=j}^{+\infty} (x_j - x_{j-1}) \cdot (F(x_i) - F(x_{i-1})) \\
&= \sum_{j=1}^{+\infty} (x_j - x_{j-1}) \cdot (1 - F(x_{j-1}))
\end{aligned}
$$

Lo último que quedó es la suma de cada rectangulito de la F (que en el caso discreto es constante salvo en los saltos que ocurren en los x_i). Luego equivale a la integral de la función[1].

45. (a) Caso Discreto: Supongamos que existe un $x_1 > 0$ tal que $p_X(x_1) > 0$.

$$0 = E(X) = \sum_{\{x \in Rn(X)\}} x \cdot p_X(x) \geq x_1 \cdot p_X(x_1) > 0$$

Luego, no existe tal x_1, y $p(X > 0) = 0 \Rightarrow p(X = 0) = 1$.

(b) Caso Continuo: Sea $a > 0$ un valor tal que $f(a) > 0$. Existe un entorno de a en $\mathbb{R}$ (digamos $[a - \epsilon, a + \epsilon]$) donde la densidad es estrictamente positiva porque f es continua. Es decir que en $[a, a + \epsilon]$ también. Como toda la función es no negativa, se cumple:

$$\int_a^{+\infty} x \cdot f_X(x)dx \geq \int_a^{a+\epsilon} x \cdot f_X(x)dx > 0$$

Aplicando esto a la esperanza se tiene:

$$E(X) = \int_0^a x \cdot f_X(x)dx + \int_a^{+\infty} x \cdot f_X(x)dx \geq \int_a^{+\infty} x \cdot f_X(x)dx > 0$$

Luego no puede ser que exista un $a > 0$ con densidad positiva, entonces $P(X = 0) = 1$.

[1] Hay un dibujo que sirve para entender qué está sumando esta expresión en el apunte de Ferrari, al principio de la página 13.

46. Minimizar $E(X-c)^2$ es lo mismo que minimizar $E(X-c)^2 - E(X^2)$ porque es un número que no depende de c.

$$\begin{aligned} E(X-c)^2 - E(X^2) &= E((X-c)^2 - X^2) = E((X-c+X)(X-c-X)) \\ &= E((2X-c)c) = c \cdot E(2X-c) = c \cdot (2E(X)-c) \end{aligned}$$

que es una cuadrática que tiene sus ceros en $c=0$ y $c=2E(X)$. Por lo tanto el mínimo está en el vértice que es el promedio: $c=E(X)$.

47.

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} k \cdot \frac{e^{-\lambda}\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{(k-1)!} \\ &= e^{-\lambda} \cdot \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= e^{-\lambda} \cdot \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \cdot \lambda e^{\lambda} = \lambda \end{aligned}$$

48.

$$\begin{aligned} E(X) &= \sum_{i=1}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\ &= \sum_{i=1}^{n} i \frac{n!}{i! \cdot (n-i)!} p^i (1-p)^{n-i} \\ &= np \sum_{i=1}^{n} \frac{(n-1)!}{(i-1)! \cdot (n-i)!} p^{i-1} (1-p)^{n-i} \\ &= np \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{(n-1)-i} \\ &= np \end{aligned}$$

Porque el término de la sumatoria es la probabilidad puntual de una $Bi(n-1,p)$.

49. (a) $V(X) = E((x-\mu)^2) \geq 0$ porque es suma de producto de cosas no negativas.

(b) $\Rightarrow)$ $V(X) = E((X-\mu_X)^2) = V(0) = 0$.

$\Leftarrow$) Sea $Y = (X - \mu_X)^2$. Sabemos que $V(X) = E(Y) = 0$. Como es no negativa, por la propiedad 45 sabemos $P(Y = 0) = 1$. Finalmente, $P(X = \mu_X) = 1$.

(c) $V(X + b) = E((X + b - E(X + b))^2) = E((X + b - \mu_X - b)^2) = E((X - \mu_X)^2) = V(X)$

(d) $V(aX) = E((aX - E(aX))^2) = E(((aX - a\mu_X)^2) = E(a^2(X - \mu_X)^2) = a^2E((X - \mu_X)^2) = a^2V(X)$

50. Primero, el conjunto de las VA es un espacio vectorial real (o sea, las VA se portan bien con la suma y el producto por un escalar). Además, si definimos $< X, Y >= E(XY)$ vale que

- $< X, Y >=< Y, X >$
- $< X, X >= E(X^2) \geq 0$
- $E(X^2) = 0$ sii $X = 0$

así que es un producto interno del EV.

Entonces tenemos:

$$\begin{aligned}
\text{Cov}(X, Y)^2 &= E((X - \mu_X)(Y - \mu_Y))^2 \\
&=< X - \mu_X, Y - \mu_Y >^2 \\
&=< X - \mu_X, Y - \mu_Y >^2 \\
&\leq< X - \mu_X, X - \mu_X >< Y - \mu_Y, Y - \mu_Y > \\
&= E((X - \mu_X)^2)E((Y - \mu_Y)^2) \\
&= V(X)V(Y)
\end{aligned}$$

La desigualdad vale por CBS para espacios vectoriales.

51. (a)

$$\begin{aligned}
\text{Cov}(X, Y) &= E((X - \mu_X)(Y - \mu_Y)) \\
&= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X\mu_Y \\
&= E(XY) - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y \\
&= E(XY) - \mu_X\mu_Y
\end{aligned}$$

(b)

$$\text{Cov}(X, X) = E((X - \mu_X)^2) = V(X)$$

(c) Sean $U = X - \mu_X$ y $V = Y - \mu_Y$.

$$V(X) + V(Y) - |\text{Cov}(X, Y)| = E(U^2) + E(V^2) - |\text{Cov}(U, V)|$$

Si $\text{Cov}(X, Y) = E(UV) \geq 0$:

$$\begin{aligned}
V(X) + V(Y) - |\text{Cov}(X, Y)| &= E(U^2) + E(V^2) - E(UV) \\
&= E((U - V)^2) + E(UV) \\
&\geq E((U - V)^2) \geq 0
\end{aligned}$$

Si no,

$$\begin{aligned} V(X)+V(Y)-|\mathrm{Cov}(X,Y)| &= E(U^2)+E(V^2)+E(UV) \\ &= E((U+V)^2)-E(UV) \\ &\geq E((U+V)^2) \geq 0 \end{aligned}$$

(d) Es obvio porque el producto conmuta.

(e)

$$\begin{aligned} \mathrm{Cov}(aX+bY,Z) &= E((aX-a\mu_X+bY-b\mu_Y)(Z-\mu_Z)) \\ &= E((aX-a\mu_X)(Z-\mu_Z)+(bY-b\mu_Y)(Z-\mu_Z)) \\ &= E((aX-a\mu_X)(Z-\mu_Z))+E((bY-b\mu_Y)(Z-\mu_Z)) \\ &= a\cdot E((X-\mu_X)(Z-\mu_Z))+b\cdot E((Y-\mu_Y)(Z-\mu_Z)) \\ &= a\cdot \mathrm{Cov}(X,Z)+b\cdot \mathrm{Cov}(Y,Z) \end{aligned}$$

52. (a) Sean $U = X-\mu_X$, $V = Y-\mu_Y$.

$$V(X+Y) = E(U+V)^2 = E(U^2+V^2+2UV) = V(U)+V(V)+2\mathrm{Cov}(U,V)$$

$$V(X+Y) = V(X)+V(Y)+2\mathrm{Cov}(X,Y)$$

(b)

$$\mathrm{Cov}(X,Y) = E((X-\mu_X)(Y-\mu_Y)) = E(XY)-E(X)E(Y) = 0$$

53. BUSCAR EJEMPLO DESPUES

54. BUSCAR EJEMPLO DESPUES

55.

$$\begin{aligned} \rho(aX+b,cY+d) &= \frac{Cov(aX+b,cY+d)}{\sqrt{V(aX+b)V(cY+d)}} \\ &= \frac{ac\cdot Cov(X,Y)}{\sqrt{a^2c^2V(X)V(Y)}} \\ &= \frac{ac}{|ac|}\cdot\rho(X,Y) \\ &= \mathrm{sg}(ac)\cdot\rho(X,Y) \end{aligned}$$

56. Sean $U = X-\mu_X$, $V = Y-\mu_Y$. Sabemos que

$$\rho(X,Y) = \frac{\mathrm{Cov}(U,V)}{\sqrt{\mathrm{Var}(U)\mathrm{Var}(V)}} = \frac{< U,V >}{\sqrt{< U,U >< V,V >}}$$

$$\rho(X,Y)^2 = \frac{< U,V >^2}{< U,U >< V,V >}$$

El cociente es ≤ 1 por CS. Además sabemos que ocurre la igualdad cuando:

$$\begin{aligned}
V &= \alpha U \\
(Y - \mu_Y) &= \alpha(X - \mu_X) \\
Y - \alpha X &= \mu_Y - \alpha \cdot \mu_X = \beta \\
Y &= \alpha X + \beta
\end{aligned}$$

57.

$$E(aX) = \sum_{x \in \mathrm{Rg}(X)} aX \cdot P(X = x) = a \sum_{x \in \mathrm{Rg}(X)} X \cdot P(X = x) = a \cdot E(X)$$

$$\begin{aligned}
E(X + Y) &= \sum_{x \in \mathrm{Rg}(X)} \sum_{y \in \mathrm{Rg}(Y)} (x + y) P(X = x \wedge Y = y) \\
&= \sum_{x \in \mathrm{Rg}(X)} \sum_{y \in \mathrm{Rg}(Y)} x \cdot P(X = x \wedge Y = y) + \sum_{x \in \mathrm{Rg}(X)} \sum_{y \in \mathrm{Rg}(Y)} y \cdot P(X = x \wedge Y = y) \\
&= \sum_{x \in \mathrm{Rg}(X)} x \sum_{y \in \mathrm{Rg}(Y)} P(X = x \wedge Y = y) + \sum_{y \in \mathrm{Rg}(Y)} y \sum_{x \in \mathrm{Rg}(X)} P(X = x \wedge Y = y) \\
&= \sum_{x \in \mathrm{Rg}(X)} x \cdot P(X = x) + \sum y \in \mathrm{Rg}(Y) y \cdot P(Y = y) \\
&= E(X) + E(Y)
\end{aligned}$$

La propiedad vale para la sumatoria de n términos por inducción.

58.

$$E(aX) = \int_{-\infty}^{+\infty} aX \cdot f_X(x) = a \int_{-\infty}^{+\infty} X \cdot f_X(x) = a \cdot E(X)$$

$$\begin{aligned}
E(X + Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) f_{X,Y}(x, y) dy \ dx \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot f_{X,Y}(x, y) dy \ dx + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \cdot f_{X,Y}(x, y) dx \ dy \\
&= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy \ dx + \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx \ dy \\
&= \int_{-\infty}^{+\infty} x \cdot f_X(x) \ dx + \int_{-\infty}^{+\infty} y \cdot f_Y(y) \ dy \\
&= \int_{-\infty}^{+\infty} x \cdot f_X(x) \ dx + \int_{-\infty}^{+\infty} y \cdot f_Y(y) \ dy \\
&= E(X) + E(Y)
\end{aligned}$$

La propiedad vale para la sumatoria de n términos por inducción.

59. Sea $Z = X - Y$. Entonces $P(Z \geq 0) = 1$.

$$E(Z) = \int_{-\infty}^{+\infty} z \cdot f_Z(z)\ dz \geq \int_{-\infty}^{+\infty} 0 \cdot f_Z(z)\ dz = \int_{-\infty}^{+\infty} 0\ dz = 0$$

$$E(Z) = E(X) + E(Y) \Rightarrow E(X) + E(Y) \geq 0 \Rightarrow E(X) \geq E(Y)$$

60.

$$E(X) = \sum_{x \in \mathrm{Rg}(X)} x \cdot P(X = x) = c \cdot P(X = c) = c \cdot 1 = c$$

6 Esperanza Condicional

61.

X \ Y	2	3	T
0	0	1/8	1/8
1	3/8	0	3/8
2	3/8	0	3/8
3	0	1/8	1/8
T	3/4	1/4	1

No son independientes.

$$E(X|Y=2) = \frac{1\cdot\frac{3}{8} + 2\cdot\frac{3}{8}}{\frac{3}{8}+\frac{3}{8}} = 3/2$$

62.

$$E(X|Y=1) = \frac{2\cdot 1/6 + 4\cdot 1/6 + 6\cdot 1/6}{1/2} = 4$$

$$E(Y|X\geq 4) = \frac{1\cdot 1/6 + 0\cdot 1/6 + 1\cdot 1/6}{1/2} = 2/3$$

63. Sean $N \sim P(\lambda)$ y $X_i \sim \mathrm{Bi}(p)$.

$$\begin{aligned}
E\left(\sum_{i=1}^{N} X_i\right) &= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{N} X_i \middle| N=k\right)\cdot P(N=k) \\
&= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{k} X_i\right)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty}\left(\sum_{i=1}^{k}\mu_X\right)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} kp\frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-\lambda}\lambda\sum_{k=1}^{\infty} p\frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda}\lambda p\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} \\
&= e^{-\lambda}\lambda p e^{\lambda} \\
&= \lambda p
\end{aligned}$$

64.

$$\begin{aligned}
f_{X,Y}(x,y) &= \lambda^2 e^{-\lambda y}\mathbb{I}_{(0<x<y)} \\
f_X(x) &= \int_{-\infty}^{+\infty} \lambda^2 e^{-\lambda y}\mathbb{I}_{(0<x<y)}\ dy \\
&= \int_x +\infty \lambda^2 e^{-\lambda y}\mathbb{I}_{(0<x)}\ dy \\
&= (-\lambda e^{-\lambda y})\Big|_x^{+\infty}\mathbb{I}_{(0<x)} \\
&= (\lambda e^{-\lambda x})\mathbb{I}_{(0<x)}
\end{aligned}$$

Entonces $X \sim E(\lambda)$. Por otro lado:

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{+\infty} \lambda^2 e^{-\lambda y}\mathbb{I}_{(0<x<y)}\ dx \\
&= \int_0^y \lambda^2 e^{-\lambda y}\mathbb{I}_{(0<y)}\ dx \\
&= \lambda^2 e^{-\lambda y}\cdot x\Big|_0^y \mathbb{I}_{(0<y)} \\
&= \lambda^2 e^{-\lambda y}\cdot y\mathbb{I}_{(0<y)} \\
&= \frac{\lambda^2\cdot y^{2-1}e^{-\lambda y}}{(2-1)!}\mathbb{I}_{(0<y)}
\end{aligned}$$

Entonces $Y \sim \Gamma(2,\lambda)$, por lo que $Z = Y - X \sim E(\lambda)$.

65.

$$\begin{aligned}
E\left(\sum_{i=1}^{N} X_i\right) &= \sum_{j=1}^{\infty} P(N=j)\cdot E\left(\sum_{i=1}^{N} X_i \Big| N=j\right) \\
&= \sum_{j=1}^{\infty} P(N=j)\cdot E\left(\sum_{i=1}^{j} X_i\right) \\
&= \sum_{j=1}^{\infty} P(N=j)\cdot \left(\sum_{i=1}^{j} E(X_i)\right) \\
&= \sum_{j=1}^{\infty} P(N=j)\cdot j\cdot E(X_1) \\
&= E(X_1)\sum_{j=1}^{\infty} j\cdot P(N=j) \\
&= E(X_1)E(N)
\end{aligned}$$

66. Sea N la cantidad de túneles recorridos hasta salir. $N \sim G(1/3)$. Sea T el tiempo total que se tarda en salir, y T_i el tiempo que lleva recorrer cada

túnel salvo el último.

$$E(T_i) = 1/2 \cdot 2h + 1/2 \cdot 3h = 2,5h$$

Ahora usamos Wald para resolver la esperanza condicional:

$$\begin{aligned}
E(T) &= E\left(1h + \sum_{n=1}^{N-1} T_i\right) = 1h + E\left(\sum_{n=1}^{N-1} T_i\right) \\
&= 1h + E(N-1)E(T_i) \\
&= 1h + (3h - 1h) \cdot 2,5h = 6h
\end{aligned}$$

7 Vectores Aleatorios

73. Sea $\mathbb{I}_{[\epsilon,+\infty)}(X)$ la indicadora de $X \geq \epsilon$.

$$\mathbb{I}_{[\epsilon,+\infty)}(X) \cdot \epsilon \leq X$$

Porque si $X < \epsilon$ el lado izquierdo da 0, y si no es obvio que $X \geq \epsilon$.

Si tomamos esperanza a ambos lados, siendo que la esperanza de la indicadora es $P(X \geq \epsilon)$,

$$P(X \geq \epsilon) \cdot \epsilon \leq E(X) \Rightarrow P(X \geq \epsilon) \leq \frac{E(X)}{\epsilon}$$

Si tomamos la VA $(X - E(X))^2$ (sabemos que es ≥ 0), y reescribimos la constante como ϵ^2,

$$P((X - E(X))^2 \geq \epsilon^2) \leq \frac{E((X-\mu)^2)}{\epsilon^2}$$
$$P(|X - E(X)| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}$$

Que es la fórmula de Chebychev.

74.

8 Funciones Generadoras de Momentos

78. Quiero ver que

$$M_{X_1+X_2+\cdots+X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

Como las X_i son independientes, tenemos que

$$M_{X_1+X_2+\cdots+X_n}(t) = E(e^{t\sum_{i=1}^{n} X_i}) = E(\prod_{i=1}^{n} e^{tX_i}) = \prod_{i=1}^{n} E(e^{tX_i}) = \prod_{i=1}^{n} M_{X_i}(t)$$

79.

$$M_{aX+b}(t) = E(e^{atX+bt}) = E(e^{atX}) \cdot E(e^{bt}) = E(e^{atX}) \cdot e^{bt} = M_X(at) \cdot e^{bt}$$

$E(e^{bt}) = e^{bt}$ porque e^{bt} es una constante.

80. Quiero calcular $M_X(t) = E(e^{tX})$ para $X \sim N(0,1)$.

$$E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tX} \cdot f_X(x) dX = \int_{-\infty}^{+\infty} e^{tX} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2}} dX$$
$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{tX-\frac{X^2}{2}} dX$$

Ahora multiplico y divido por $e^{-\frac{t^2}{2}}$ para completar a una normal:

$$= e^{\frac{t^2}{2}} \cdot \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{tX-\frac{X^2}{2}} \cdot e^{-\frac{t^2}{2}} dX$$
$$= e^{\frac{t^2}{2}} \cdot \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(X^2-2tX+t^2)} dX = e^{\frac{t^2}{2}} \cdot \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(X-t)^2}{2}} dX$$

Notamos que lo que está adentro de la integral es $f_X(x)$ si $X \sim N(t,1)$. Por lo tanto integra a 1 :) y la FGM de la normal es:

$$M_X(t) = e^{\frac{t^2}{2}}$$

81. Sea $X \sim P(\lambda)$.

$$M_X(t) = E(e^{tX}) = \sum_{k=0}^{\infty} e^{tk} \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{e^{tk}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t\lambda)^k}{k!}$$

Ahora, sabemos que $\sum_{i=0}^{n} \frac{x^i}{i!} = e^x$ (se deriva de la f de la Poisson, de hecho).

$$M_X(t) = e^{-\lambda} e^{e^t \lambda} = e^{e^t \lambda - \lambda} = e^{\lambda(e^t - 1)}$$

Ahora vamos a usar esto para calcular la FGM de una suma de Poissons. Si da la FGM de una Poisson, para cierto λ, demostramos que la suma es $P(\lambda)$ por unicidad de la FGM.

Sean $X_i \sim P(\lambda_i)$ VA independientes. Como son independientes (por item 78):

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t) = \prod_{i=1}^n e^{\lambda_i (e^t - 1)} = e^{\sum_{i=1}^n \lambda_i (e^t - 1)} = e^{(e^t - 1) \sum_{i=1}^n \lambda_i}$$

Por lo tanto, $\sum_{i=1}^n X_i \sim P(\sum_{i=1}^n \lambda_i)$.

9 Coupon collector

131. Voy a asumir que empiezo a contar desde 0 y no desde 1 porque me hace feliz, y los índices quedan más lindos.

    Primero, T_i es el número de intentos hasta el primer éxito (encontrar un elemento nuevo), con una probabilidad de éxito de:

    $$\frac{\text{elementos nuevos}}{\text{elementos totales}} = \frac{n-i}{n}$$

    Además, las T_i son independientes.

    En particular, T_0 (el tiempo hasta el primer elemento) tiene una probabilidad de éxito de 1.

    Esta definición de T_i quiere decir que $T_i \sim G(\frac{n-i}{n})$. Por esta razón, $E(T_i) = \frac{n}{n-i}$.

    $$E(T) = \sum_{i=0}^{n-1} E(T_i)$$

    $$E(T) = \sum_{i=0}^{n-1} \frac{n}{n-i}$$

    $$E(T) = n \cdot \sum_{i=0}^{n-1} \frac{1}{n-i}$$

    $$E(T) = n \cdot \sum_{i=1}^{n} \frac{1}{i}$$

    FALTA TERMINAR