Varianza, condicionamiento y múltiples variables aleatorias discretas

Clase 5: *variance*, *conditioning*, *joint PMFs* e *independence*

6.431x — Probability: The Science of Uncertainty and Data

B.Sc. Cristian Lazo Quispe

BREIT (Brescia Institute of Technology) · Aporta — Grupo Breca

Programa SDS (Statistics and Data Science, MITx) · **BREIT** (Brescia Institute of Technology), iniciativa de Aporta —Grupo Breca— en alianza con MIT IDSS.
Basado en MIT 6.431x / MIT OCW RES.6-012 (CC BY-NC-SA 4.0).

clazoq@uni.pe · linkedin.com/in/cristian-lazo-quispe · github.com/CristianLazoQuispe

Dónde quedamos (Clase 4)

Variables aleatorias, en tres líneas

Una v.a. X es un número que depende del azar. Su **PMF** $p_X(x) = \mathbb{P}(X = x)$ reparte la masa; su **esperanza** $\mathbb{E}[X] = \sum_x x\, p_X(x)$ es el centro de masa. Regla del valor esperado: $\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$. Linealidad: $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$.

Lo que usaremos hoy

- $\mathbb{E}[X^2] \neq (\mathbb{E}[X])^2$ en general: la diferencia es la **varianza** de hoy.
- **Indicadores**: $\mathbb{E}[I_A] = \mathbb{P}(A)$ y $\mathbb{E}[X_1 + \cdots + X_n] = \sum \mathbb{E}[X_i]$ sin independencia (sombreros: $\mathbb{E}[X] = 1$).
- **Condicionar** (Clase 3): encoger el universo y renormalizar.

El salto de hoy

Pasamos de **una** v.a. a **dos**: la PMF conjunta, lo que significa que X e Y sean independientes, y por qué **la esperanza de una suma siempre se suma, pero la varianza solo a veces**. Con esto se cierra la Unit 4 y el material del **Midterm 1**.

¿Qué veremos hoy? (Unit 4 — Lectures 6 y 7)

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

Variance: promediar la distancia al cuadrado

Definición

Con $\mu = \mathbb{E}[X]$,

$$\mathrm{Var}(X) = \mathbb{E}\big[(X-\mu)^2\big] = \sum_x (x-\mu)^2 p_X(x) \ \geq 0$$

y la desviación estándar es $\sigma_X = \sqrt{\mathrm{Var}(X)}$. ¿Por qué al cuadrado? Porque $\mathbb{E}[X-\mu] = 0$ **siempre**: las desviaciones se cancelan.

La fórmula de cálculo (*the most important formula*)

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Juego de la Lecture 5 ($1, 2, 4$ con $0{,}2; 0{,}5; 0{,}3$): $\mathbb{E}[X] = 2{,}4$, $\mathbb{E}[X^2] = 7$, $\mathrm{Var}(X) = 7 - 5{,}76 = 1{,}24$, $\sigma \approx 1{,}11$.

Unidades y lectura

Si X está en soles, $\mathrm{Var}(X)$ está en soles^2; σ_X vuelve a soles: "la demora típica se aparta σ minutos del promedio".

Catálogo

v.a.	$\mathbb{E}[X]$	$\mathrm{Var}(X)$
Bernoulli(p)	p	$p(1-p)$
Uniforme $\{a..b\}$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+2)}{12}$
Bin(n, p)	np	$np(1-p)$
Geométrica(p)	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Bernoulli: $X^2 = X$, así que $\mathbb{E}[X^2] = p$ y Var $= p - p^2$. Máxima en $p = \frac{1}{2}$.

Problema: misma media, distinta varianza

Problema

X es uniforme en $\{1, 2, 3, 4, 5\}$; Y vale 1 o 5 con probabilidad $\frac{1}{2}$ cada uno; $Z = 3$ siempre. Las tres tienen media 3. Calcula sus varianzas y ordénalas.

Piénsalo antes de pasar a la solución.

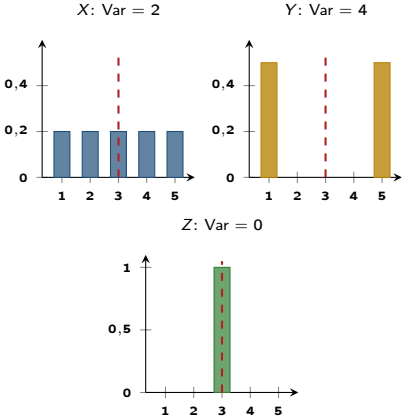
Solución del demo: misma media, distinta varianza

Problema

X es uniforme en $\{1,2,3,4,5\}$; Y vale 1 o 5 con probabilidad $\frac{1}{2}$ cada uno; $Z = 3$ siempre. Las tres tienen media 3. Calcula sus varianzas y ordénalas.

Solución

$\mathrm{Var}(X) = \frac{1}{5}[(1-3)^2+(2-3)^2+0+(4-3)^2+(5-3)^2] = \frac{10}{5} = 2.$
$\mathrm{Var}(Y) = \frac{1}{2}(1-3)^2 + \frac{1}{2}(5-3)^2 = 4.$ $\mathrm{Var}(Z) = 0.$
Y pone **toda** su masa a distancia 2 de la media; X pone parte a distancia 0 y 1. La varianza vale 0 solo si la v.a. es constante.
Error clásico: "tienen la misma media, así que se parecen". La media no ve la dispersión.



Línea roja: $\mu = 3$ en los tres casos. La varianza pondera cada $(x-\mu)^2$ por $p_X(x)$.

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

Cambio de escala y de origen

Teorema

$$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X), \qquad \sigma_{aX+b} = |a|\,\sigma_X.$$

Por qué, en dos pasos

- $Y = X + b$: $\mathbb{E}[Y] = \mu + b$ y $Y - \mathbb{E}[Y] = X - \mu$. **Las distancias a la media no cambian**: desplazar la masa no la dispersa.
- $Y = aX$: $(aX - a\mu)^2 = a^2(X-\mu)^2$. Las distancias se estiran por $|a|$; sus cuadrados, por a^2.

Dos errores que valen puntos

- **El $+b$ no cuenta**: $\mathrm{Var}(X+3) = \mathrm{Var}(X)$, no $\mathrm{Var}(X)+3$.
- **El signo de a no cuenta**: $\mathrm{Var}(3-4X) = 16\,\mathrm{Var}(X)$, no $-4\,\mathrm{Var}(X)$.

Una varianza **nunca** es negativa: si te sale -8, el error es seguro.

De dónde sale $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mu^2$

$\mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 = \mathbb{E}[X^2] - \mu^2$.
Consecuencia: $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$ **siempre**.

Problema: Var$(3-4X)$, el $+3$ no cuenta

Problema

X es la demora (en minutos) de un bus, con $\mathrm{Var}(X) = 2$. Una app cobra una penalidad $C = 3 - 4X$ (en soles). Calcula $\mathrm{Var}(C)$ y σ_C. ¿Y si fuera $C' = 3 + 4X$?

Piénsalo antes de pasar a la solución.

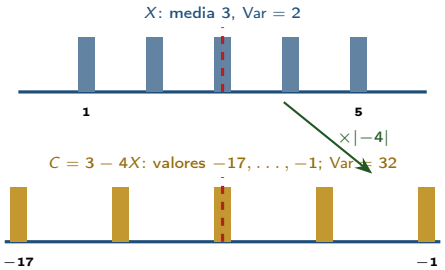
Solución del demo: $\mathrm{Var}(3-4X)$, el $+3$ no cuenta

Problema

X es la demora (en minutos) de un bus, con $\mathrm{Var}(X)=2$. Una app cobra una penalidad $C=3-4X$ (en soles). Calcula $\mathrm{Var}(C)$ y σ_C. ¿Y si fuera $C'=3+4X$?

Solución

$\mathrm{Var}(C)=(-4)^2\,\mathrm{Var}(X)=16\cdot 2=32$ soles^2; $\sigma_C=|-4|\sqrt{2}\approx 5{,}7$ soles. Para $C'=3+4X$: **lo mismo**, $\mathrm{Var}(C')=32$. El signo cambia la media ($\mathbb{E}[C]=3-4\mathbb{E}[X]$ frente a $3+4\mathbb{E}[X]$), no la dispersión. **Errores clásicos**: $-4\cdot 2=-8$ (negativa, imposible) o $16\cdot 2+3=35$ (el $+3$ adentro).



Las barras se separan 4 veces más (varianza $\times 16$) y el $+3$ solo las desliza.

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

PMF condicional: encoger el universo y renormalizar

Definiciones (si $\mathbb{P}(A) > 0$)

$$p_{X \mid A}(x) = \mathbb{P}(X = x \mid A) = \frac{\mathbb{P}(\{X = x\} \cap A)}{\mathbb{P}(A)}$$

$$\mathbb{E}[X \mid A] = \sum_x x\, p_{X \mid A}(x)$$

$$\mathrm{Var}(X \mid A) = \mathbb{E}[X^2 \mid A] - \big(\mathbb{E}[X \mid A]\big)^2$$

$p_{X \mid A}$ es una PMF de verdad: todo lo que vale para una PMF vale para ella.

Ejemplo de la Lecture 6

X uniforme en $\{1,2,3,4\}$, $A = \{X \geq 2\}$: $p_{X \mid A} = \frac{1}{3}$ en $\{2,3,4\}$, $\mathbb{E}[X \mid A] = 3$,
$\mathrm{Var}(X \mid A) = \frac{1}{3}[1 + 0 + 1] = \frac{2}{3}$ (antes: 2,5 y $\frac{5}{4}$).

Total expectation theorem

Si $A_1, \ldots, A_n$ es una partición,

$$\mathbb{E}[X] = \sum_i \mathbb{P}(A_i)\, \mathbb{E}[X \mid A_i].$$

Promedio **ponderado** de medias condicionales: cada rama aporta su media, pesada por su probabilidad.

Error clásico

Promediar las medias condicionales **sin pesos** ("$(4 + 12)/2$"). Las ramas no pesan igual.

Problema: un dado, dado que salió par

Problema

X es un dado justo. Te dicen que salió **par**. Calcula $p_{X \mid A}$, $\mathbb{E}[X \mid A]$ y $\mathrm{Var}(X \mid A)$. Compara con $\mathbb{E}[X] = 3{,}5$ y $\mathrm{Var}(X) = \frac{35}{12}$.

Piénsalo antes de pasar a la solución.

Solución del demo: un dado, dado que salió par

Problema

X es un dado justo. Te dicen que salió **par**. Calcula $p_{X \mid A}$, $\mathbb{E}[X \mid A]$ y $\mathrm{Var}(X \mid A)$. Compara con $\mathbb{E}[X] = 3{,}5$ y $\mathrm{Var}(X) = \frac{35}{12}$.

Solución

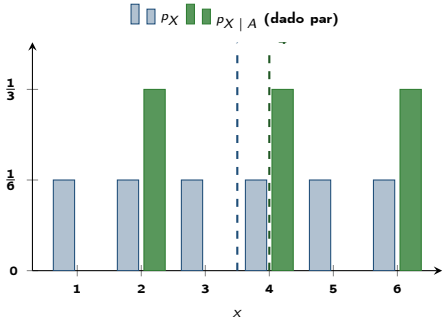
$A = \{2, 4, 6\}$, $\mathbb{P}(A) = \frac{1}{2}$. Para x par, $p_{X \mid A}(x) = \frac{1/6}{1/2} = \frac{1}{3}$; para x impar, 0.

$$\mathbb{E}[X \mid A] = \tfrac{1}{3}(2+4+6) = 4, \qquad \mathbb{E}[X^2 \mid A] = \tfrac{4+16+36}{3} = \tfrac{56}{3}$$

$$\mathrm{Var}(X \mid A) = \tfrac{56}{3} - 16 = \tfrac{8}{3} \approx 2{,}67.$$

La información mueve la media (de 3,5 a 4) y aquí reduce un poco la dispersión.

Error clásico: responder 3,5 (ignorar A) o no dividir entre $\mathbb{P}(A)$.



Se borra la masa fuera de A y la que queda se reparte proporcionalmente: de $\frac{1}{6}$ a $\frac{1}{3}$.

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

La geométrica no tiene memoria

Memorylessness

X = lanzamientos hasta la primera cara, $p_X(k) = (1-p)^{k-1}p$. Dado $X > n$ (los primeros n fueron sello), $X - n$ vuelve a ser geométrica con el mismo p:

$$\mathbb{P}(X - n = k \mid X > n) = p_X(k)$$

$$\mathbb{P}(X > n + k \mid X > n) = \mathbb{P}(X > k) = (1-p)^k.$$

Por qué: arranque fresco (*fresh start*)

$$\mathbb{P}(X - 1 = 3 \mid X > 1) = \frac{\mathbb{P}(T_1 T_2 T_3 H_4)}{\mathbb{P}(T_1)}$$

$$= \frac{(1-p)^3 p}{1-p} = (1-p)^2 p = p_X(3)$$

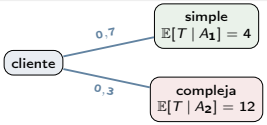
Los lanzamientos son independientes: la moneda no se acuerda de los fracasos. "Ya le toca salir" es la *falacia del jugador*.

Esperanza total en una agencia bancaria

70 % de clientes: operación simple, 4 min en promedio; 30 %: compleja, 12 min.

$$\mathbb{E}[T] = 0{,}7 \cdot 4 + 0{,}3 \cdot 12 = 2{,}8 + 3{,}6 = 6{,}4 \text{ min}$$

No $(4+12)/2 = 8$.



Problema: $\mathbb{E}[X] = 1/p$ en dos líneas, y "ya van tres sin seis"

Problema

(a) Usa la esperanza total y la falta de memoria para hallar la media de la geométrica sin sumar series.
(b) Lanzas un dado hasta el primer 6 y ya van 3 sin 6: ¿cuántos lanzamientos esperas en total? ¿Cuánto vale $\mathbb{P}(X = 5 \mid X > 3)$?

Piénsalo antes de pasar a la solución.

Solución del demo: $\mathbb{E}[X] = 1/p$ en dos líneas, y "ya van tres sin seis"

Problema

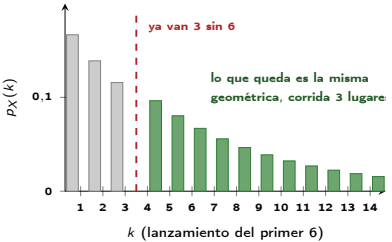
(a) Usa la esperanza total y la falta de memoria para hallar la media de la geométrica sin sumar series. (b) Lanzas un dado hasta el primer 6 y ya van 3 sin 6: ¿cuántos lanzamientos esperas en total? ¿Cuánto vale $\mathbb{P}(X = 5 \mid X > 3)$?

Solución

(a) Partición por el primer lanzamiento: cara (p) o sello ($1-p$).

$$\mathbb{E}[X] = p \cdot 1 + (1-p)\big(1 + \underbrace{\mathbb{E}[X-1 \mid X>1]}_{=\mathbb{E}[X] \text{ (sin memoria)}}\big)$$

Despejando: $\mathbb{E}[X] = 1 + (1-p)\mathbb{E}[X]$, luego $p\,\mathbb{E}[X] = 1$ y $\mathbb{E}[X] = \frac{1}{p}$. (b) Los que faltan son geométricos con media 6: $\mathbb{E}[X \mid X > 3] = 3 + 6 = 9$. Y
$\mathbb{P}(X = 5 \mid X > 3) = \mathbb{P}(X - 3 = 2) = \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{36}$.
Error clásico: $\mathbb{P}(X = 5) = (\frac{5}{6})^4 \frac{1}{6}$, que ignora lo que ya pasó.



Las barras verdes, renormalizadas, son exactamente $p_X(k-3)$.

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

Joint PMF: la tabla, las marginales y el relieve en 3D

Definición

$$p_{X,Y}(x,y) = \mathbb{P}(X = x \text{ y } Y = y)$$

y la masa total es 1: $\sum_x \sum_y p_{X,Y}(x,y) = 1$. **Marginales**: $p_X(x) = \sum_y p_{X,Y}(x,y)$ (sumar la columna) y $p_Y(y) = \sum_x p_{X,Y}(x,y)$ (sumar la fila).

Regla del valor esperado y linealidad

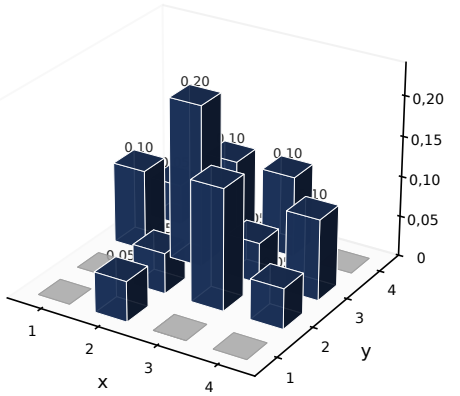
$$\mathbb{E}[g(X,Y)] = \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y)$$

Con $g = x + y$ la doble suma se separa: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ **sin ninguna hipótesis**.

Cuidado

La conjunta determina las marginales. **Las marginales no determinan la conjunta.**

la masa total es 1: suma de las 16 barras



Sobre cada par (x,y) se levanta una barra de altura $p_{X,Y}(x,y)$. Las 16 barras suman 1. Aplastar el relieve contra una pared da una marginal.

Problema: la tabla conjunta de la Lecture 6

Problema

Con la tabla (celdas en veinteavos), calcula $p_{X,Y}(1,3)$, $p_X(4)$, $p_Y(2)$, $\mathbb{P}(X = Y)$, $\mathbb{E}[X]$ y $\mathbb{E}[Y]$.

Piénsalo antes de pasar a la solución.

Solución del demo: la tabla conjunta de la Lecture 6

Problema

Con la tabla (celdas en veinteavos), calcula $p_{X,Y}(1,3)$, $p_X(4)$, $p_Y(2)$, $\mathbb{P}(X=Y)$, $\mathbb{E}[X]$ y $\mathbb{E}[Y]$.

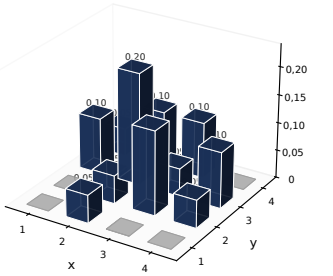
Solución

- $p_{X,Y}(1,3) = \frac{2}{20}$: una celda.
- $p_X(4) = 0+2+1+0 = \frac{3}{20}$: la columna $x=4$.
- $p_Y(2) = 0+1+3+1 = \frac{5}{20}$: la fila $y=2$.
- $\mathbb{P}(X=Y) = \frac{0+1+1+0}{20} = \frac{2}{20}$: la diagonal.
- $\mathbb{E}[X] = \frac{3+16+18+12}{20} = \frac{49}{20} = 2{,}45$;
  $\mathbb{E}[Y] = \frac{1+10+27+20}{20} = 2{,}9$.

Error clásico: confundir una celda ($p_{X,Y}$) con una suma de fila o columna (marginal).

$y\backslash x$	1	2	3	4	$p_Y(y)$
4	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	0	$\frac{5}{20}$
3	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{9}{20}$
2	0	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	$\frac{5}{20}$
1	0	$\frac{1}{20}$	0	0	$\frac{1}{20}$
$p_X(x)$	$\frac{3}{20}$	$\frac{8}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	1

la masa total es 1: suma de las 16 barras



Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

PMF condicional dada una v.a. e independencia

Conditional PMF (si $p_Y(y) > 0$)

$$p_{X \mid Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Receta: tomar la fila y y dividirla entre su suma.
Esperanza condicional:
$\mathbb{E}[X \mid Y = y] = \sum_x x\, p_{X \mid Y}(x \mid y)$, un número que depende de y.

Dos consecuencias

- Regla de multiplicación:
  $p_{X,Y} = p_Y\, p_{X \mid Y} = p_X\, p_{Y \mid X}$.
- Esperanza total:
  $\mathbb{E}[X] = \sum_y p_Y(y)\, \mathbb{E}[X \mid Y = y]$ (partición por el valor de Y).

Independence

$$p_{X,Y}(x,y) = p_X(x)\, p_Y(y)$$

para **todos** los pares (x, y). Equivale a $p_{X \mid Y}(x \mid y) = p_X(x)$: saber Y no cambia la PMF de X.

Cómo se decide

Para **afirmarla**: comprobar **todas** las celdas. Para **refutarla**: basta **una** que falle. Atajo: una celda en 0 con marginales positivas ya la refuta.

Con la tabla de la Lecture 6

$p_{X,Y}(1,1) = 0$ pero $p_X(1)p_Y(1) = \frac{3}{20} \cdot \frac{1}{20} \neq 0$: **no** son independientes.

Independencia, esperanzas y varianzas

Lo que vale **siempre**

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathrm{Var}(aX) = a^2\,\mathrm{Var}(X), \qquad \mathrm{Var}(X+b) = \mathrm{Var}(X).$$

Lo que vale **solo si X e Y son independientes**

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

y también $\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\,\mathbb{E}[h(Y)]$.

Varianza de la binomial

$X = I_1 + \cdots + I_n$ con indicadoras **independientes**, cada una Bernoulli(p): $\mathrm{Var}(X) = n\,p(1-p)$.

Por qué (medias cero)

$$\mathrm{Var}(X+Y) = \mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2] + 2\,\mathbb{E}[XY] + \mathbb{E}[Y^2].$$

El término cruzado $2\mathbb{E}[XY]$ se anula con independencia ($\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 0$). Sin ella, puede **sumar o restar**: es lo que la dependencia agrega a la dispersión.

La frase para memorizar

La esperanza de una suma **siempre** se suma. La varianza de una suma (y la esperanza de un producto) **solo con independencia**.

Problema: $\mathrm{Var}(X+Y)$ con $Y=X$, $Y=-X$ e independientes

Problema (Lecture 7)

$\mathrm{Var}(X)=\mathrm{Var}(Y)=1$. Calcula $\mathrm{Var}(X+Y)$ en tres casos: (i) $Y=X$; (ii) $Y=-X$; (iii) X e Y independientes. Y $\mathrm{Var}(X-3Y)$ si son independientes.

Piénsalo antes de pasar a la solución.

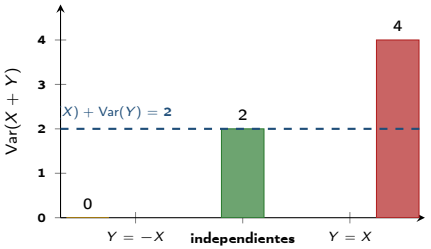
Solución del demo: $\mathrm{Var}(X+Y)$ con $Y=X$, $Y=-X$ e independientes

Problema (Lecture 7)

$\mathrm{Var}(X)=\mathrm{Var}(Y)=1$. Calcula $\mathrm{Var}(X+Y)$ en tres casos: (i) $Y=X$; (ii) $Y=-X$; (iii) X e Y independientes. Y $\mathrm{Var}(X-3Y)$ si son independientes.

Solución

(i) $X+Y=2X$: $\mathrm{Var}(2X)=4\,\mathrm{Var}(X)=4$. (ii) $X+Y=0$: Var $=0$. (iii) $\mathrm{Var}(X)+\mathrm{Var}(Y)=2$.
Independientes:
$\mathrm{Var}(X-3Y)=\mathrm{Var}(X)+(-3)^2\,\mathrm{Var}(Y)=1+9=10$ (no $1-3=-2$).
Error clásico: responder 2 en (i) o (ii). "Sumar varianzas" **necesita** independencia; la dependencia positiva amplifica y la negativa cancela.



Mismas varianzas marginales; la varianza de la suma va de 0 a 4 según la dependencia.

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

Para pensar (1): la tabla conjunta, condicionada

Pregunta 1

Con la tabla conjunta de la Lecture 6 (celdas en veinteavos):

1. Halla $p_{X \mid Y}(x \mid 2)$ para $x = 1, 2, 3, 4$ y $\mathbb{E}[X \mid Y = 2]$.
2. ¿Son X e Y independientes? Justifica con **una** celda.
3. Condiciona en $C = \{X \leq 2,\ Y \geq 3\}$. ¿Son independientes *dentro* de C?

$y \backslash x$	1	2	3	4	$p_Y(y)$
4	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	0	$\frac{5}{20}$
3	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{9}{20}$
2	0	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	$\frac{5}{20}$
1	0	$\frac{1}{20}$	0	0	$\frac{1}{20}$
$p_X(x)$	$\frac{3}{20}$	$\frac{8}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	1

Solución 1 (1/4): la fila $y = 2$, renormalizada

Paso 1: aislar la fila

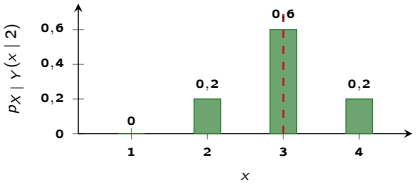
Condicionar en $Y = 2$ es quedarse con la fila $y = 2$ de la tabla: $(0,\ 1,\ 3,\ 1)$ en veinteavos. Su suma es la marginal: $p_Y(2) = \frac{0+1+3+1}{20} = \frac{5}{20}$.

Paso 2: dividir entre la suma

$$p_{X\,|\,Y}(x\,|\,2) = \frac{p_{X,Y}(x,2)}{p_Y(2)} = \frac{(0,1,3,1)/20}{5/20}$$

$$= \left(0,\ \tfrac{1}{5},\ \tfrac{3}{5},\ \tfrac{1}{5}\right) \quad \text{para } x = 1,2,3,4.$$

Comprobación: $0 + \frac{1}{5} + \frac{3}{5} + \frac{1}{5} = 1$. Es una PMF en x. **Error clásico**: reportar la fila sin dividir (la conjunta), que no suma 1.



La fila $(0, 1, 3, 1)$ conserva su **forma**; solo cambia la escala.

Paso 3: la media condicional

$$\mathbb{E}[X \mid Y = 2] = 1{\cdot}0 + 2{\cdot}\tfrac{1}{5} + 3{\cdot}\tfrac{3}{5} + 4{\cdot}\tfrac{1}{5} = \tfrac{2+9+4}{5} = 3$$

Compara con $\mathbb{E}[X] = 2{,}45$: saber que $Y = 2$ **sube** la media de X.

Solución 1 (2/4): ¿independientes? Basta una celda que falle

Qué habría que comprobar

$p_{X,Y}(x,y) = p_X(x)\,p_Y(y)$ en **las 16 celdas**. Para *refutarla* basta **una** que falle.

La celda $(1,1)$

$$p_{X,Y}(1,1) = 0, \qquad \text{pero}$$

$$p_X(1)\,p_Y(1) = \tfrac{3}{20}\cdot\tfrac{1}{20} = \tfrac{3}{400} \neq 0.$$

No son independientes.

Atajo

Celda en **0** con marginales **positivas** $\Rightarrow$ no hay independencia. Aquí sirven las cinco celdas con 0.

La misma conclusión, en lenguaje de condicionales

$$p_{X\mid Y}(1\mid 1) = \frac{p_{X,Y}(1,1)}{p_Y(1)} = \frac{0}{1/20} = 0$$

mientras que $p_X(1) = \frac{3}{20}$. Saber que $Y = 1$ **cambia** la PMF de X: de hecho la obliga a valer 2. Eso es exactamente dependencia.

$y\backslash x$	1	2	3	4	$p_Y(y)$
4	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	0	$\frac{5}{20}$
3	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{9}{20}$
2	0	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	$\frac{5}{20}$
1	**0**	$\frac{1}{20}$	0	0	$\frac{1}{20}$
$p_X(x)$	$\frac{3}{20}$	$\frac{8}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	1

La celda $(1,1)$ es 0 aunque $p_X(1) > 0$ y $p_Y(1) > 0$.

Solución 1 (3/4): las medias condicionales lo confirman

Si fueran independientes...

... todas las filas tendrían la **misma forma** ($p_{X \mid Y}(\cdot \mid y) = p_X$) y **todas las medias condicionales valdrían** $\mathbb{E}[X] = 2{,}45$.

Fila por fila (dividir entre $p_Y(y)$)

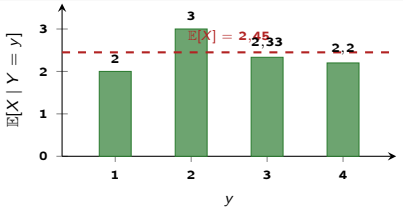
- $y = 1$: fila $(0, 1, 0, 0)$, $\mathbb{E}[X \mid Y = 1] = 2$.
- $y = 2$: fila $(0, 1, 3, 1)/5$, $\mathbb{E}[X \mid Y = 2] = 3$ (paso anterior).
- $y = 3$: fila $(2, 4, 1, 2)/9$, $\mathbb{E}[X \mid Y = 3] = \frac{2+8+3+8}{9} = \frac{21}{9} \approx 2{,}33$.
- $y = 4$: fila $(1, 2, 2, 0)/5$, $\mathbb{E}[X \mid Y = 4] = \frac{1+4+6}{5} = \frac{11}{5} = 2{,}2$.

Cuatro medias distintas: la fila **sí** depende de y.

Chequeo con la esperanza total

$\mathbb{E}[X] = \sum_y p_Y(y)\,\mathbb{E}[X \mid Y = y]$ tiene que devolver $\frac{49}{20}$:

$$\frac{1}{20}\cdot 2 + \frac{5}{20}\cdot 3 + \frac{9}{20}\cdot\frac{21}{9} + \frac{5}{20}\cdot\frac{11}{5} = \frac{2+15+21+11}{20} = \frac{49}{20}. \checkmark$$



Con independencia, las cuatro barras estarían a la altura de la línea roja.

Solución 1 (4/4): dentro de $C = \{X \le 2,\ Y \ge 3\}$ sí son independientes

Paso 1: quedarse con el bloque y renormalizar

Las cuatro celdas de C tienen masa $2, 4, 1, 2$ (veinteavos); $\mathbb{P}(C) = \frac{9}{20}$. Dividiendo cada celda entre $\frac{9}{20}$:

$$(1,3) \to \tfrac{2}{9}, \quad (2,3) \to \tfrac{4}{9}, \quad (1,4) \to \tfrac{1}{9}, \quad (2,4) \to \tfrac{2}{9}.$$

Paso 2: marginales condicionales

Columnas: $p_{X \mid C}(1) = \frac{2}{9} + \frac{1}{9} = \frac{1}{3}$,
$p_{X \mid C}(2) = \frac{4}{9} + \frac{2}{9} = \frac{2}{3}$.
Filas: $p_{Y \mid C}(3) = \frac{2}{9} + \frac{4}{9} = \frac{2}{3}$,
$p_{Y \mid C}(4) = \frac{1}{9} + \frac{2}{9} = \frac{1}{3}$.

$y \backslash x$	1	2	$p_{Y \mid C}$
4	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
3	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{2}{3}$
$p_{X \mid C}$	$\frac{1}{3}$	$\frac{2}{3}$	1

Paso 3: comprobar las **cuatro** celdas

$\frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$ ✓ $\quad \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$ ✓ $\quad \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$ ✓
$\frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$ ✓
Dentro de C, X e Y **son independientes** (*conditionally independent*).

Moraleja

La independencia **depende del universo**. Dentro de C las filas $(2,4)$ y $(1,2)$ son **proporcionales**: eso es ser independientes.

Para pensar (2): mismas marginales, ¿misma conjunta?

Pregunta 2

Dos monedas justas, $X, Y \in \{0,1\}$ ($1 =$ cara). Se proponen **tres** tablas conjuntas:

(i) $\frac{1}{4}$ en cada una de las cuatro celdas;

(ii) $\frac{1}{2}$ en $(0,0)$ y $\frac{1}{2}$ en $(1,1)$;

(iii) $\frac{1}{2}$ en $(0,1)$ y $\frac{1}{2}$ en $(1,0)$.

1. Verifica que las tres tienen las **mismas** marginales.
2. Calcula $\mathbb{E}[X+Y]$ en los tres casos.
3. Calcula $\mathbb{E}[XY]$ y $\mathrm{Var}(X+Y)$ en los tres casos. ¿Qué cambió y qué no?

Solución 2 (1/4): las marginales son idénticas

Sumar filas y columnas en cada tabla

(i) $y\backslash x$	0	1	p_Y
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
p_X	$\frac{1}{2}$	$\frac{1}{2}$	1

(ii) $y\backslash x$	0	1	p_Y
1	0	$\frac{1}{2}$	$\frac{1}{2}$
0	$\frac{1}{2}$	0	$\frac{1}{2}$
p_X	$\frac{1}{2}$	$\frac{1}{2}$	1

(iii) $y\backslash x$	0	1	p_Y
1	$\frac{1}{2}$	0	$\frac{1}{2}$
0	0	$\frac{1}{2}$	$\frac{1}{2}$
p_X	$\frac{1}{2}$	$\frac{1}{2}$	1

Conclusión del paso 1

En las tres, $p_X = p_Y = (\frac{1}{2}, \frac{1}{2})$: dos monedas justas. Por separado, X e Y son **indistinguibles** en los tres casos: $\mathbb{E}[X] = \mathbb{E}[Y] = \frac{1}{2}$ y $\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Lo que las tablas dicen que no se ve en las marginales

- (i) Las celdas son el producto de las marginales ($\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$): **independientes**. Dos monedas distintas.
- (ii) $Y = X$ siempre: la **misma** moneda vista dos veces.
- (iii) $Y = 1 - X$ siempre: la segunda es la **opuesta** de la primera.

Moraleja parcial

Conocer X e Y "por separado" no es conocer el par. Las marginales **no** determinan la conjunta.

Solución 2 (2/4): la esperanza de la suma no ve la dependencia

Por linealidad, sin mirar la tabla

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = \tfrac{1}{2} + \tfrac{1}{2} = 1$$

en los tres casos. La linealidad **solo usa las marginales**; por eso da lo mismo en (i), (ii) y (iii).

Comprobación celda por celda (regla del valor esperado)

- (i): $(0+0)\tfrac{1}{4} + (0+1)\tfrac{1}{4} + (1+0)\tfrac{1}{4} + (1+1)\tfrac{1}{4} = \tfrac{0+1+1+2}{4} = 1$.
- (ii): $(0+0)\tfrac{1}{2} + (1+1)\tfrac{1}{2} = 1$.
- (iii): $(0+1)\tfrac{1}{2} + (1+0)\tfrac{1}{2} = 1$.

Pero la PMF de $S = X + Y$ sí cambia

$p_S(s)$	$s=0$	$s=1$	$s=2$
(i) independientes	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
(ii) $Y = X$	$\frac{1}{2}$	0	$\frac{1}{2}$
(iii) $Y = 1 - X$	0	1	0

Lectura

Tres PMFs distintas con el **mismo centro de masa** 1: una en el medio, una en los extremos, una concentrada. Ya sabemos cómo distinguirlas: por la **varianza**.

Solución 2 (3/4): $\mathbb{E}[XY]$ y $\mathrm{Var}(X+Y)$ sí ven la dependencia

$\mathbb{E}[XY]$: solo la celda $(1,1)$ aporta

$XY = 1$ únicamente cuando $X = Y = 1$, así que $\mathbb{E}[XY] = p_{X,Y}(1,1)$:

$$\text{(i) } \tfrac{1}{4}, \qquad \text{(ii) } \tfrac{1}{2}, \qquad \text{(iii) } 0.$$

Compara con $\mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$: coincide **solo** en (i), el caso independiente.

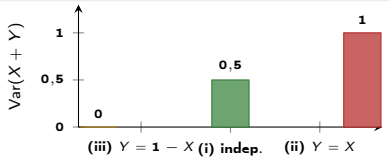
Lectura

En (ii) el producto vale 1 con el doble de frecuencia que "si fueran independientes"; en (iii) nunca vale 1. $\mathbb{E}[XY]$ **sí** mira la tabla completa.

$\mathrm{Var}(X+Y)$ con la PMF de $S = X+Y$ (slide anterior)

$\mathbb{E}[S] = 1$ en los tres casos, así que $\mathrm{Var}(S) = \mathbb{E}[S^2] - 1$:

- (i) $\mathbb{E}[S^2] = 0\cdot\frac{1}{4} + 1\cdot\frac{1}{2} + 4\cdot\frac{1}{4} = \frac{3}{2}$, $\mathrm{Var}(S) = \frac{1}{2} = \mathrm{Var}(X) + \mathrm{Var}(Y)$.
- (ii) $\mathbb{E}[S^2] = 0\cdot\frac{1}{2} + 4\cdot\frac{1}{2} = 2$, $\mathrm{Var}(S) = 1 = \mathrm{Var}(2X) = 4\cdot\frac{1}{4}$.
- (iii) $\mathbb{E}[S^2] = 1\cdot 1 = 1$, $\mathrm{Var}(S) = 0$: S es constante.



Solución 2 (4/4): el culpable es el término cruzado

La identidad general

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)+2\big(\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]\big)$$

Con $\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{1}{4}$ y $\mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{4}$:

- (i) $\frac{1}{4}+\frac{1}{4}+2\big(\frac{1}{4}-\frac{1}{4}\big) = \frac{1}{2}$: término cruzado **nulo**.
- (ii) $\frac{1}{2}+2\big(\frac{1}{2}-\frac{1}{4}\big) = 1$: la dependencia **positiva** amplifica.
- (iii) $\frac{1}{2}+2\big(0-\frac{1}{4}\big) = 0$: la dependencia **negativa** cancela.

Ese paréntesis se llamará *covariance* en la Clase 8.

Lo que cambió y lo que no

No cambió: las marginales, $\mathbb{E}[X]$, $\mathbb{E}[Y]$ y $\mathbb{E}[X+Y]$.

Cambió: la PMF de la suma, $\mathbb{E}[XY]$ y $\mathrm{Var}(X+Y)$.

Sumar varianzas solo fue válido en (i).

Moraleja de la pregunta 2

Las marginales dicen cómo es cada variable **sola**. La conjunta dice, además, cómo se **mueven juntas**. La linealidad de $\mathbb{E}$ no necesita lo segundo; la varianza de la suma, sí.

Para pensar (3): los sombreros, ahora con varianza

Pregunta 3

n personas dejan su sombrero en una caja y cada una saca uno al azar (todas las asignaciones son igualmente probables). $X =$ número de personas que recuperan **su** sombrero. En la Clase 4 vimos que $\mathbb{E}[X] = 1$ para todo n.

1. Escribe $X = I_1 + \cdots + I_n$ y calcula $\mathbb{E}[I_i^2]$ y $\mathbb{E}[I_i I_j]$ para $i \neq j$.
2. Calcula $\mathbb{E}[X^2]$ y $\mathrm{Var}(X)$.
3. ¿Qué daría "sumar las varianzas de las I_i"? ¿Por qué no es la respuesta?

Solución 3 (1/3): indicadoras y lo que ya sabíamos

Las indicadoras

$I_i = 1$ si la persona i recupera su sombrero, 0 si no. Entonces $X = I_1 + \cdots + I_n$.

$$\mathbb{P}(I_i = 1) = \frac{1}{n}, \qquad \mathbb{E}[I_i] = \frac{1}{n}$$

(hay n sombreros posibles para la persona i y uno es el suyo). **Clase 4:** $\mathbb{E}[X] = \sum_i \mathbb{E}[I_i] = n \cdot \frac{1}{n} = 1$, sin independencia.

Los cuadrados

Como I_i solo vale 0 o 1, $I_i^2 = I_i$ y

$$\mathbb{E}[I_i^2] = \mathbb{E}[I_i] = \frac{1}{n}.$$

Los productos cruzados

$I_i I_j = 1$ solo si **ambas** personas aciertan. Por la regla de multiplicación:

$$\mathbb{E}[I_i I_j] = \mathbb{P}(I_i = 1)\,\mathbb{P}(I_j = 1 \mid I_i = 1) = \frac{1}{n} \cdot \frac{1}{n-1}.$$

Si la persona i ya tiene el suyo, a j le quedan $n-1$ sombreros y uno es el suyo.

Ojo

$\frac{1}{n(n-1)} > \frac{1}{n^2} = \mathbb{E}[I_i]\mathbb{E}[I_j]$: las indicadoras **no** son independientes. Que una persona acierte hace un poco *más* probable que otra acierte (hay menos sombreros ajenos circulando).

Solución 3 (2/3): $\mathbb{E}[X^2]$ con sus términos cruzados

Elevar la suma al cuadrado

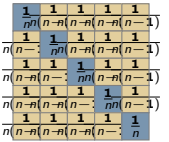
$$X^2 = \Big(\sum_i I_i\Big)^2 = \underbrace{\sum_i I_i^2}_{n \text{ términos}} + \underbrace{\sum_{i \neq j} I_i I_j}_{n(n-1) \text{ términos}}$$

Hay n cuadrados y $n(n-1)$ productos con $i \neq j$ (cada par ordenado).

Tomar esperanza (linealidad: vale siempre)

$$\mathbb{E}[X^2] = n \cdot \tfrac{1}{n} + n(n-1) \cdot \tfrac{1}{n(n-1)} = 1 + 1 = 2$$

$\mathbb{E}[I_i I_j]$, con $n = 5$



diagonal: $n \cdot \frac{1}{n} = 1$

fuera: $n(n-1) \cdot \frac{1}{n(n-1)} = 1$

La varianza

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 2 - 1 = 1$$

para todo $n \geq 2$.

Solución 3 (3/3): por qué no se pueden sumar las varianzas

Lo que daría "sumar varianzas"

Cada I_i es Bernoulli$(\frac{1}{n})$: $\mathrm{Var}(I_i) = \frac{1}{n}\left(1 - \frac{1}{n}\right)$.
Sumando las n: $\sum_i \mathrm{Var}(I_i) = 1 - \frac{1}{n}$ (p. ej. 0,98 con $n = 50$). Es la varianza de una Bin$(n, \frac{1}{n})$: n intentos **independientes**. Pero los aciertos no lo son.

La diferencia está en los términos cruzados

$$\mathrm{Var}(X) = \sum_i \mathrm{Var}(I_i) + \sum_{i \neq j} \left(\mathbb{E}[I_i I_j] - \mathbb{E}[I_i]\mathbb{E}[I_j]\right)$$

$$= \left(1 - \tfrac{1}{n}\right) + n(n-1)\left(\tfrac{1}{n(n-1)} - \tfrac{1}{n^2}\right) = 1 - \tfrac{1}{n} + \tfrac{1}{n} = 1.$$

El caso $n = 2$ lo muestra sin fórmulas

Dos personas: o ambas aciertan o ninguna.
$I_1 = 1$ **obliga** a $I_2 = 1$. Dependencia total.
$X \in \{0, 2\}$ con prob. $\frac{1}{2}$: $\mathbb{E}[X] = 1$,
$\mathrm{Var}(X) = \frac{1}{2}(1) + \frac{1}{2}(1) = 1$. ✓

Resumen de la pregunta 3

- $\mathbb{E}[X] = 1$ y $\mathrm{Var}(X) = 1$ para cualquier n: ni la media ni la dispersión crecen con el grupo.
- La esperanza se obtuvo por linealidad (sin independencia); la varianza **necesitó** los $\mathbb{E}[I_i I_j]$, porque sumar varianzas exige independencia.
- Media 1 y varianza 1 es la huella de una Poisson(1): lo veremos más adelante.

Contenido

1. Varianza: cuánto se dispersa la masa
2. Propiedades de la varianza
3. Condicionar una v.a. en un evento
4. Esperanza total y la geométrica sin memoria
5. PMF conjunta y marginales
6. Condicionar en una v.a. e independencia
7. Para pensar: tres preguntas
8. Cierre

Kahoot de repaso (Clase 5)

Jugamos ahora

10 preguntas sobre **varianza, condicionamiento, PMF conjunta e independencia**: $\mathrm{Var}(aX + b)$, misma media distinta varianza, dado par, marginales y condicionales de una tabla, esperanza total, geométrica sin memoria, $\mathrm{Var}(X + Y)$ con dependencia, mismas marginales y los sombreros con varianza. Las 3 primeras son de calentamiento (30 s); el resto, de 1 a 1,5 min. Unos 10 minutos.

- Cada pregunta tiene una **trampa**: el error clásico está entre las opciones.
- Antes de responder una varianza: ¿me salió **negativa** o con un $+b$ adentro?
- Al condicionar: ¿**dividí** entre $\mathbb{P}(A)$ o $p_Y(y)$?
- Antes de sumar varianzas o factorizar $\mathbb{E}[XY]$: ¿comprobé la **independencia**?

Resumen de la Clase 5

Lo que hay que saber hacer

1. Calcular $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ y aplicar $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$ (el $+b$ y el signo de a no cuentan; nunca negativa).
2. Reconocer las varianzas del catálogo: $p(1-p)$, $\frac{(b-a)(b-a+2)}{12}$, $np(1-p)$, $\frac{1-p}{p^2}$.
3. **Condicionar** una v.a.: tomar la fila y **renormalizar**; usar la **esperanza total** como promedio ponderado de medias condicionales.
4. Usar la **falta de memoria** de la geométrica ($\mathbb{E}[X] = 1/p$ en dos líneas).
5. Leer una **tabla conjunta**: celdas, marginales (sumar filas/columnas), condicionales (dividir), $\mathbb{E}[g(X, Y)]$.
6. Decidir **independencia** ($p_{X,Y} = p_X p_Y$ en todas las celdas) y saber que $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ y $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ **solo** valen con ella.
7. Calcular varianzas de conteos con indicadoras **y sus términos cruzados**.

Siguiente clase (Clase 6)

Unit 5: variables aleatorias **continuas**, densidades (PDF), la normal y el condicionamiento continuo. Todo lo de hoy, cambiando sumas por integrales.

¡Gracias!

Clase 5 — Varianza, condicionamiento y múltiples v.a. discretas (6.431x, Unit 4)

clazoq@uni.pe · linkedin.com/in/cristian-lazo-quispe · github.com/CristianLazoQuispe

BREIT (Brescia Institute of Technology) · Aporta — Grupo Breca · en alianza con MIT IDSS

Basado en MIT 6.431x / MIT OCW RES.6-012 (Bertsekas & Tsitsiklis), CC BY-NC-SA 4.0.