Esperanza y covarianza de vectores aleatorios

- Esperanza de funciones de vectores aleatorios
- Covarianza y Correlación
- Media y Varianza Muestral
- Vector de Medias y Matriz de Covarianzas
- Sumas Aleatorias
- Mezclas de Distribuciones

Esperanza de funciones de vectores aleatorios

El **Teorema del Estadístico Inconsciente (TEI)** nos dice cómo calcular la esperanza de una función $g(X,Y)$ sin necesidad de encontrar primero la distribución de $g(X,Y)$.

Simulador: Esperanza de funciones en dados

Función g(X,Y): Calcular E[g(X,Y)]

TEI para vectores aleatorios:
Caso discreto: $\mathbb{E}[g(X,Y)] = \sum_x \sum_y g(x,y) f_{XY}(x,y)$
Caso continuo: $\mathbb{E}[g(X,Y)] = \iint_{\mathbb{R}^2} g(x,y) f_{XY}(x,y) dx dy$

Propiedad: Siempre se cumple que $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$. ¿Por qué?

La esperanza es lineal. Para $g(X,Y)=X+Y$, la integral/suma se separa en dos términos, cada uno dando la esperanza marginal correspondiente.

Ejercicio

Sean X,Y los resultados de lanzar dos dados justos de forma independiente (cada uno Uniforme Discreta en $\{1,\ldots,6\}$). Use el TEI para calcular $\mathbb{E}[\max(X,Y)]$ y verifique el resultado con el simulador de arriba.

Primero encontramos la f.m.p. de $\max(X,Y)$. Para $k=1,\ldots,6$: $$\mathbb{P}(\max(X,Y)\le k) = \mathbb{P}(X\le k)\mathbb{P}(Y\le k) = \left(\frac{k}{6}\right)^2 \implies \mathbb{P}(\max(X,Y)=k) = \frac{k^2-(k-1)^2}{36} = \frac{2k-1}{36}.$$ Aplicando el TEI (aquí, directamente sobre la distribución de $g(X,Y)=\max(X,Y)$, lo cual es válido porque ya conocemos su f.m.p.): $$\mathbb{E}[\max(X,Y)] = \sum_{k=1}^{6} k\cdot\frac{2k-1}{36} = \frac{1+6+15+28+45+66}{36} = \frac{161}{36} \approx 4.4722.$$ Compare este valor con el que reporta el simulador al seleccionar "g(X,Y) = max(X, Y)".

Covarianza y Correlación

La **covarianza** mide la asociación lineal entre dos variables: $$ \operatorname{cov}(X,Y) = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] $$ El **coeficiente de correlación** $\rho = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y}$ está normalizado entre -1 y 1.

Explora diferentes niveles de correlación

Correlación ρ: 0.8 0.80 Generar 200 puntos

Propiedades de la covarianza:

- $\operatorname{cov}(X,X) = \operatorname{var}(X)$
- $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$
- $\operatorname{cov}(aX, bY) = ab \operatorname{cov}(X,Y)$
- $\operatorname{cov}(X+Y, Z) = \operatorname{cov}(X,Z) + \operatorname{cov}(Y,Z)$
- Si $X \perp Y$, entonces $\operatorname{cov}(X,Y)=0$ (pero el recíproco no es cierto)

Contraejemplo: covarianza cero NO implica independencia

X\Y	-1	0	1	P(X=x)
-1	1/3	0	1/3	2/3
1	0	1/3	0	1/3
P(Y=y)	1/3	1/3	1/3	1

Verifica: En esta tabla, $\mathbb{E}[X]=-1/3$, $\mathbb{E}[Y]=0$, $\mathbb{E}[XY]=0$. ¿Son independientes X e Y?

No son independientes. Por ejemplo, $f(-1,-1)=1/3$ pero $f_X(-1)f_Y(-1) = (2/3)(1/3)=2/9$. ¡Covarianza cero pero dependientes!

Desigualdad de Cauchy-Schwarz: $$ (\mathbb{E}[XY])^2 \le \mathbb{E}[X^2]\mathbb{E}[Y^2] $$ De aquí se deduce que $|\operatorname{corr}(X,Y)| \le 1$.

Ejercicio

Una urna tiene tres bolas numeradas 1, 2, 3. Se extraen dos bolas sin reemplazo. Sea X el número de la primera bola extraída y Y el de la segunda. Determine $\operatorname{cov}(X,Y)$ y ρ_{XY}.

Hay 6 pares ordenados igualmente probables, cada uno con probabilidad 1/6: $(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)$. Por simetría, X y Y tienen marginal Uniforme Discreta en $\{1,2,3\}$, así que $\mathbb{E}[X]=\mathbb{E}[Y]=2$ y $\mathbb{E}[X^2]=\mathbb{E}[Y^2]=\frac{1+4+9}{3}=\frac{14}{3}$, de donde $\operatorname{var}(X)=\operatorname{var}(Y)=\frac{14}{3}-4=\frac{2}{3}$.

Para $\mathbb{E}[XY]$, sumamos xy sobre los 6 pares: $$ \mathbb{E}[XY] = \frac{1\cdot2+1\cdot3+2\cdot1+2\cdot3+3\cdot1+3\cdot2}{6} = \frac{22}{6} = \frac{11}{3}. $$ Entonces: $$ \operatorname{cov}(X,Y) = \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y] = \frac{11}{3}-4 = -\frac{1}{3}, \qquad \rho_{XY} = \frac{-1/3}{2/3} = -\frac{1}{2}. $$ La covarianza es negativa: si la primera bola extraída tiene un número alto, queda menos probabilidad de que la segunda también lo tenga (muestreo sin reemplazo).

Media y Varianza Muestral

Para una muestra $X_1, \ldots, X_n$ i.i.d. con media μ y varianza σ^2: $$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Simulador: Comportamiento de la media muestral

Distribución: **Tamaño muestral n:** 30 30

Generar muestra

Propiedades fundamentales:

- $\mathbb{E}[\bar{X}_n] = \mu$ (insesgado)
- $\operatorname{var}(\bar{X}_n) = \sigma^2/n$
- $\mathbb{E}[S^2] = \sigma^2$ (cuasivarianza insesgada)

La división por $n-1$ (en lugar de n) corrige el sesgo.

¿Por qué usamos n-1 en lugar de n para la varianza muestral?

Porque $\mathbb{E}[\frac{1}{n}\sum (X_i - \bar{X})^2] = \frac{n-1}{n}\sigma^2$. Al dividir por $n-1$ obtenemos un estimador insesgado: $\mathbb{E}[S^2] = \sigma^2$.

Ejercicio

Se registra la siguiente muestra de 5 observaciones: 4, 8, 6, 5, 7. Calcule $\bar{X}$ y S^2.

La media muestral es: $$\bar{X} = \frac{4+8+6+5+7}{5} = \frac{30}{5} = 6.$$ Las desviaciones respecto a la media son $-2, 2, 0, -1, 1$, y sus cuadrados suman $4+4+0+1+1=10$. Entonces: $$S^2 = \frac{1}{n-1}\sum_{i=1}^5 (X_i-\bar{X})^2 = \frac{10}{4} = 2.5.$$

Vector de Medias y Matriz de Covarianzas

> Para un vector aleatorio $\mathbf{X} = (X_1, \ldots, X_n)^T$: $$ \boldsymbol{\mu} = \mathbb{E}[\mathbf{X}] = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \operatorname{cov}(\mathbf{X}) = \mathbb{E}[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T] $$ $\boldsymbol{\Sigma}$ es simétrica y semidefinida positiva.

Visualización de matriz de covarianzas 2D

Varianza de X_1 (σ_1^2): 1.0 **Varianza de X_2 (σ_2^2):** 1.0 **Covarianza σ_{12}:** 0.5

Propiedades:

- $\mathbb{E}[A\mathbf{X}] = A\boldsymbol{\mu}$
- $\operatorname{cov}(A\mathbf{X}) = A\boldsymbol{\Sigma}A^T$
- $\operatorname{vartot}(\mathbf{X}) = \operatorname{tr}(\boldsymbol{\Sigma}) = \sum_i \sigma_i^2 = \sum_i \lambda_i$

La **variación total** es la suma de las varianzas, igual a la suma de los valores propios.

¿Qué significa que la matriz de covarianzas sea diagonal?

Si $\boldsymbol{\Sigma}$ es diagonal, todas las covarianzas son cero: las variables están **no correlacionadas**. Pero cuidado: ¡no correlacionadas no implica independencia!

Ejercicio

Sea $\mathbf{X}=(X_1,X_2)^T$ con matriz de covarianzas $\boldsymbol{\Sigma}=\begin{pmatrix}2 & 1\\ 1 & 3\end{pmatrix}$. Use la fórmula $\operatorname{cov}(A\mathbf{X})=A\boldsymbol{\Sigma}A^T$ con $A=(1,1)$ para calcular $\operatorname{var}(X_1+X_2)$, y verifique el resultado con la fórmula usual de varianza de una suma.

Con $A=(1,1)$: $$ A\boldsymbol{\Sigma} = (1,1)\begin{pmatrix}2&1\\1&3\end{pmatrix} = (2+1,\ 1+3) = (3,4), $$ $$ A\boldsymbol{\Sigma}A^T = (3,4)\binom{1}{1} = 3+4 = 7. $$ Verificación con la fórmula usual: $\operatorname{var}(X_1+X_2)=\operatorname{var}(X_1)+\operatorname{var}(X_2)+2\operatorname{cov}(X_1,X_2) = 2+3+2(1)=7$. Ambos métodos coinciden, como debe ser: el enfoque matricial generaliza la fórmula escalar a cualquier combinación lineal del vector.

Sumas Aleatorias

Sea N una variable aleatoria entera no negativa, y $X_1, X_2, \ldots$ i.i.d. independientes de N. Definimos la suma aleatoria: $$ S_N = X_1 + X_2 + \cdots + X_N $$

Simulador: Elevador con personas

Número medio de personas (λ Poisson): 2.3 2.3 **Peso medio por persona (kg):** 66.25 66.25 Simular 1000 viajes

Resultados clave para sumas aleatorias: $$ \mathbb{E}[S_N] = \mathbb{E}[N] \cdot \mathbb{E}[X] $$ $$ \operatorname{var}(S_N) = \operatorname{var}(N) \cdot (\mathbb{E}[X])^2 + \mathbb{E}[N] \cdot \operatorname{var}(X) $$

¿Por qué aparece $\mathbb{E}[N] \cdot \operatorname{var}(X)$ y $\operatorname{var}(N) \cdot (\mathbb{E}[X])^2$?

Por la ley de varianza total: $\operatorname{var}(S_N) = \mathbb{E}[\operatorname{var}(S_N|N)] + \operatorname{var}(\mathbb{E}[S_N|N])$. El primer término da $\mathbb{E}[N]\sigma_X^2$ y el segundo $\sigma_N^2 \mu_X^2$.

Ejercicio

(Barrios y colaboradores, Cuaderno de Ejercicios CP2, sección 4, ej. 5) Sea $(X_i, i\in\mathbb{N})$ una sucesión de v.a.i.i.d. con f.d.p. común $f(x)=\lambda e^{-\lambda x}\mathbb{I}_{(0,\infty)}(x)$. Sea N una v.a. independiente de las X_i's, con distribución Uniforme en $\{1,2,3\}$. Considere la suma aleatoria $S=\sum_{i=1}^N X_i$. Encuentre su media y su varianza.

Para $N\sim\text{Uniforme}\{1,2,3\}$: $\mathbb{E}[N]=2$, $\mathbb{E}[N^2]=\frac{1+4+9}{3}=\frac{14}{3}$, así que $\operatorname{var}(N)=\frac{14}{3}-4=\frac{2}{3}$. Para cada $X_i\sim\text{Exponencial}(\lambda)$: $\mathbb{E}[X]=1/\lambda$, $\operatorname{var}(X)=1/\lambda^2$. Aplicando las fórmulas de sumas aleatorias: $$ \mathbb{E}[S] = \mathbb{E}[N]\cdot\mathbb{E}[X] = \frac{2}{\lambda}, $$ $$ \operatorname{var}(S) = \operatorname{var}(N)(\mathbb{E}[X])^2 + \mathbb{E}[N]\operatorname{var}(X) = \frac{2}{3}\cdot\frac{1}{\lambda^2} +

2\cdot\frac{1}{\lambda^2} = \frac{8/3}{\lambda^2}. \]

Mezclas de Distribuciones

Una **mezcla** de distribuciones tiene densidad: $$ f_X(x) = \sum_{i=1}^k p_i f_i(x), \quad \text{con } p_i > 0, \sum p_i = 1 $$ Es como elegir al azar de cuál distribución proviene la observación.

Mezcla de dos normales (distribución contaminada)

Probabilidad de la normal estándar (p): 0.7 0.70 **Desviación de la contaminante (σ):** 3 3.0

Propiedades de la mezcla: $$ \mathbb{E}[X] = \sum_{i=1}^k p_i \mu_i $$ $$ \operatorname{var}(X) = \sum_{i=1}^k p_i \sigma_i^2 + \sum_{i=1}^k p_i (\mu_i - \bar{\mu})^2 $$ (varianza dentro de los grupos + varianza entre las medias)

Observa: Cuando σ es grande, ¿cómo cambia la forma de la densidad?

Aparece una "cola" más pesada. Las mezclas de normales son útiles para modelar datos atípicos o poblaciones heterogéneas (ej: pesos de personas de diferentes edades).

Ejemplo adicional: Poisson con media aleatoria Gamma produce una distribución Binomial Negativa (modelo para número de accidentes).

Ejercicio

Un tiempo de espera X sigue una mezcla de dos exponenciales: con probabilidad $p=0.6$ proviene de Exponencial($\lambda_1=1$), y con probabilidad 0.4 de Exponencial($\lambda_2=3$). Calcule $\mathbb{E}[X]$ y $\operatorname{var}(X)$.

Las medias de cada componente son $\mu_1=1/\lambda_1=1$ y $\mu_2=1/\lambda_2=1/3$. Usando la fórmula de la media de una mezcla: $$ \mathbb{E}[X] = 0.6(1) + 0.4\left(\frac{1}{3}\right) = 0.7333. $$ Para la varianza usamos $\operatorname{var}(X)=\sum p_i\sigma_i^2+\sum p_i(\mu_i-\bar\mu)^2$, con $\sigma_1^2=1$, $\sigma_2^2=1/9\approx0.1111$ (varianza de cada exponencial es $1/\lambda^2$) y $\bar\mu=\mathbb{E}[X]=0.7333$: $$ \text{dentro de grupos: } 0.6(1)+0.4(0.1111) = 0.6444, $$ $$ \text{entre medias: } 0.6(1-0.7333)^2+0.4(0.3333-0.7333)^2 = 0.6(0.2667)^2+0.4(-

0.4)^2 = 0.1067, \] \[\operatorname{var}(X) = 0.6444+0.1067 = 0.7511. \] Se puede verificar directamente con \(\operatorname{var}(X)=\mathbb{E}[X^2]-(\mathbb{E}[X])^2\), usando \(\mathbb{E}[X^2]=\sum p_i\cdot\frac{2}{\lambda_i^2}=0.6(2)+0.4(2/9)=1.2889\): \(1.2889-0.7333^2=0.7511\). ✓