# Covariance + Correlation.

One more example.!

$X \sim \text{Uniform}(1,2)$, $Y \mid X = x \sim \text{Exponential}(x)$

$\text{Cov}(X,Y)$, $\rho(X,Y)$ ?

---

$$\text{Cov}(X,Y) = \mathbb{E}(X-\mu_X)(Y-\mu_Y)$$

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

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \text{Cov}(U,V)$$

$$U = \frac{X-\mu_X}{\sigma_X}, \quad V = \frac{Y-\mu_Y}{\sigma_Y}$$

↳ standardizations.

---

$$\text{Cov}(X,Y) = \mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y.$$

$$\mathbb{E}X = \frac{1+2}{2} = \frac{3}{2}$$

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

$$= \mathbb{E}[1/X]$$

mean of Exponential($X$)

$$= \int_1^2 1/x \, dx = \ln 2$$

$$\mathbb{E}\{XY\} = \mathbb{E}\{\mathbb{E}\{XY|X\}\}$$

$$= \mathbb{E}[X \, \overset{= 1/x}{\mathbb{E}\{Y|X\}}]$$

$$= \mathbb{E}\left\{\frac{X}{X}\right\} = \mathbb{E}\,1 = 1$$

$$\mathrm{Cov}(X,Y) = 1 - \ln 2\,(3/2)$$

$$\mathrm{Var}(X) = \frac{(2-1)^2}{12} = 1/12$$

$$\mathrm{Var}(Y) = \mathrm{Var}(\mathbb{E}\{Y|X\}) + \mathbb{E}[\mathrm{Var}(Y|X)]$$

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

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

$$= 2\mathbb{E}\left[\tfrac{1}{X^2}\right] - \left(\mathbb{E}\left[\tfrac{1}{X}\right]\right)^2$$

$$= 2\int_1^2 \tfrac{1}{x^2}\, dx - \Big(\underbrace{\int_1^2 \tfrac{1}{x}\, dx}_{\ln 2}\Big)^2$$

$$= 2\left(1 - \tfrac{1}{2}\right) - (\ln 2)^2$$

$$= 1 - (\ln 2)^2$$

$$\rho(X, Y) = \frac{1 - \tfrac{3}{2}\ln 2}{\sqrt{\tfrac{1}{12}\left(1 - (\ln 2)^2\right)}} \approx -0.19.$$

---

Multiple RVs (more than 2).

n - RV's

$$X_1, X_2, \ldots, X_n$$

<u>Joint Distribution.</u> CDF.

$$F_{X_1, X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n)$$

Discrete (joint PMF)

$$P_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n).$$

$$= P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

Jointly continuous case (PDF).

$$P((X_1, X_2, \dots, X_n) \in R) = \int \dots \int_R f_{X_1 X_2 \dots X_n}(x_1, \dots, x_n) dV$$

Independence

$X_1, X_2, \dots, X_n$ are independent.

if

$$F_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) \dots F_{X_n}(x_n)$$

Discrete

$$P_{X_1 X_2 \dots X_n}(x_1, \dots, x_n) = P_{X_1}(x_1) \dots P_{X_n}(x_n)$$

Continuous

$$f_{X_1, X_2 \ldots X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n)$$

Also if $X_1, X_2, \ldots X_n$ are independent.

$$\mathbb{E}[X_1 X_2 \ldots X_n] = [\mathbb{E} X_1][\mathbb{E} X_2] \ldots [\mathbb{E} X_n].$$

---

Sums of RVs.

$$Y = X_1 + X_2 + \ldots + X_n$$

Often the case when taking measurements

$$\mathbb{E} Y = \mathbb{E}[X_1 + X_2 + \ldots + X_n].$$

$$= \mathbb{E} X_1 + \mathbb{E} X_2 + \ldots + \mathbb{E} X_n$$

$$\mathrm{Var}(Y) = \mathrm{Cov}\left( \sum_{i=1}^{n} X_i , \sum_{i=1}^{n} X_i \right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{\substack{i=1 \\ j=1}}^{n} \mathrm{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

**Ex** Flip coin $n$-time count consecutive TT pairs.

$$X_i = \begin{cases} 1 & i^{th} \text{ pair is TT} \\ 0 & \text{otherwise.} \end{cases}$$

H T T T H H H

$$Y = \sum_{i=1}^{n-1} X_i = \# \text{ of TT pairs}$$

$$\mathbb{E}Y = \sum_{i=1}^{n-1} \mathbb{E}X_i = \frac{(n-1)}{4}$$

$$\mathrm{Var}(Y) = \sum_{i=1}^{n-1} \underbrace{\mathrm{Var}(X_i)}_{= pq = \frac{1}{4}\left(\frac{3}{4}\right)} + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

$\mathrm{Cov}(X_i, X_j) = 0 \quad |i-j| > 1$ "no overlap".

Overlap $\Rightarrow \operatorname{Cov}(X_i, X_{i+1})$

$$= \underbrace{\mathbb{E}[X_i X_{i+1}]}_{= 1/8} - \underbrace{\mathbb{E}X_i}_{= 1/4}\, \underbrace{\mathbb{E}X_{i+1}}_{= 1/4}$$

$$X_i X_{i+1} = \begin{cases} 1 & \text{if } i^{th}, i+1^{th} \text{ pair are TT} \\ 0 & \text{otherwise} \end{cases}$$

$$\sim \text{Bernoulli}(\underbrace{1/8}_{1/2 \cdot 1/2 \cdot 1/2})$$

$$\operatorname{Cov}(X_i, X_{i+1}) = 1/8 - 1/16 = 1/16.$$

$$\operatorname{Var}(Y) = (n-1)(1/4)(3/4) + 2(n-2)\, 1/16$$

$$= \frac{5n-9}{16}.$$