Covariance & correlations

Lecture7b (STAT 24400 F24)

Definitions

For random variables (X, Y), the covariance is

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

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

(as long as σ_X, σ_Y are nonzero)

Cov corr basics

- Covariance and correlation are measures of dependence between two random variables having a joint distribution.
- The correlation is standardized covariance for random variables with variance.
- For any random variable X,

$$\text{Cov}(X, X) = \text{Var}(X)$$

and if $\sigma_X \neq 0$,

$$\text{Corr}(X, X) = \frac{\text{Cov}(X, X)}{\sigma_X \sigma_X} = \frac{\sigma_X^2}{\sigma_X \cdot \sigma_X} = 1.$$

Properties (range, perfect correlation)

1. **Correlation is always between 1 & -1.**

   $\text{Corr}(X, Y) = \pm 1$ if and only if $X = a + bY$ for some a, b with $b \neq 0$.

   For example,

   X = current temperature in °F, Y = current temperature in °C.
   X = weight gain in *lb*, Y = weight loss in *kg*.

Properties (linear transformation for variance)

2. **Linear transformations.**

   First, let's recall mean and variance under linear transformations:

$$\mathbb{E}(a+b\,X) = a + b\,\mathbb{E}(X), \quad \mathrm{Var}(a+b\,X) = b^2\,\mathrm{Var}(X)$$

e.g. for variance,

$$\mathrm{Var}(a+bX) = \mathbb{E}((a+bX-\mu_{a+bX})^2) = \mathbb{E}((a+bX-(a+b\mu_X))^2) = b^2\mathbb{E}((X-\mu_X)^2)$$

Similar derivation for covariance:

$$\begin{aligned}\mathrm{Cov}(a+b\,X, a'+b'\,Y) &= \mathbb{E}((a+b\,X-\mu_{a+bX})\cdot(a'+b'\,Y-\mu_{a'+b'Y}))\\ &= \mathbb{E}((a+bX-(a+b\mu_X))\cdot(a'+b'\,Y-(a'+b'\mu_Y)))\\ &= bb'\mathbb{E}((X-\mu_X)(Y-\mu_Y)) = bb'\mathrm{Cov}(X,Y)\end{aligned}$$

Properties (linear transformation for correlation)

For correlation: if b, b' are nonzero,

$$\begin{aligned}\mathrm{Corr}(a+bX, a'+b'Y) &= \frac{\mathrm{Cov}(a+bX, a'+b'Y)}{\sigma_{a+bX}\ \sigma_{a'+b'Y}}\\ &= \frac{bb'\mathrm{Cov}(X,Y)}{\sqrt{b^2\mathrm{Var}(X)}\ \sqrt{b'^2\mathrm{Var}(Y)}}\\ &= \mathrm{Corr}(X,Y)\cdot\mathrm{sign}(bb')\end{aligned}$$

where

$$sign(bb') = \begin{cases}1, & bb' > 0,\\ -1, & bb' < 0.\end{cases}$$

Properties (variance of sum of r.v.'s)

3. **Variance for sums of random variables.**

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathbb{E}[(X+Y-\mu_{X+Y})^2]\\ &= \mathbb{E}[(X+Y-\mu_X-\mu_Y)^2]\\ &= \mathbb{E}[(X-\mu_X)^2+(Y-\mu_Y)^2+2(X-\mu_X)(Y-\mu_Y)]\\ &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)\ .\end{aligned}$$

More generally,

$$\mathrm{Var}(X_1+\cdots+X_n) = \sum_i \mathrm{Var}(X_i) + 2\sum_{i<j}\mathrm{Cov}(X_i, X_j).$$

Properties (bilinearity)

4. **Covariance is bilinear.**

$$\mathrm{Cov}(X_1+X_2+\cdots+X_n, Y_1+Y_2+\cdots+Y_m) = \sum_{i=1}^n\sum_{j=1}^m \mathrm{Cov}(X_i, Y_j).$$

Proof:

$$\begin{aligned}\mathrm{Cov}\Big(\sum_i X_i, \sum_j Y_j\Big) &= \mathbb{E}\Big[\big(\sum_i X_i - \mu_{\sum_i X_i}\big)\cdot\big(\sum_j Y_j - \mu_{\sum_j Y_j}\big)\Big]\\ &= \mathbb{E}\Big[\big(\sum_i X_i - \sum_i \mu_{X_i}\big)\cdot\big(\sum_j Y_j - \sum_j \mu_{Y_j}\big)\Big]\\ &= \mathbb{E}\Big[\sum_i\sum_j (X_i-\mu_{X_i})\cdot(Y_j-\mu_{Y_j})\Big]\\ &= \sum_i\sum_j \mathbb{E}\Big[(X_i-\mu_{X_i})\cdot(Y_j-\mu_{Y_j})\Big]\\ &= \sum_i\sum_j \mathrm{Cov}(X_i, Y_j).\end{aligned}$$

def. of covariance

linearity of $\mathbb{E}(\cdot)$

linearity of $\mathbb{E}(\cdot)$

def. of covariance

Properties (under independence)

5. **Variance, covariance, and correlation under independence.**

Recall: $X \perp\!\!\!\perp Y \quad \Rightarrow \quad \mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$.

Now for $g(x) = x - \mu_X$, $h(y) = y - \mu_Y$,

$\text{Cov}(X, Y) = \mathbb{E}[(X-\mu_X)\cdot(Y-\mu_Y)] = \mathbb{E}(X-\mu_X)\cdot\mathbb{E}(Y-\mu_Y) = 0 \cdot 0 = 0$

Then $\quad X \perp\!\!\!\perp Y \quad \Rightarrow \quad \text{Corr}(X, Y) = 0$

And, under $X \perp\!\!\!\perp Y$,

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) = \text{Var}(X) + \text{Var}(Y)\ .$$

More generally, if $X_1, \ldots, X_n$ are mutually independent,

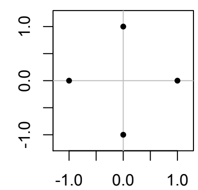
$$\text{Var}(X_1 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n)$$

Properties (correlations vs. independence)

Note: $X \perp\!\!\!\perp Y \Rightarrow \text{Corr}(X, Y) = 0$ but $\text{Corr}(X, Y) = 0 \not\Rightarrow X \perp\!\!\!\perp Y$.

Example:

(X, Y) is drawn uniformly at random from the four points $(0,1)$, $(0,-1)$, $(1,0)$, $(-1,0)$



- $\mathbb{E}(X) = \frac{1}{2}\cdot 0 + \frac{1}{4}\cdot 1 + \frac{1}{4}\cdot -1 = 0$
- Similarly $\mathbb{E}(Y) = 0$
- $\mathbb{E}(XY) = 0$, since $XY = 0$ almost surely
- $\Rightarrow \text{Cov}(X, Y) = \mathbb{E}((X-\mu_X)(Y-\mu_Y)) = \mathbb{E}(XY) = 0$

So, $\text{Corr}(X, Y) = 0$, but $X \not\!\perp\!\!\!\perp Y$

Remark: Only *some* types of dependence are captured by covariance/correlation!

Properties (shortcut formula)

6. **Shortcut for calculating covariance.**

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

Proof:

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]\\ &= \mathbb{E}(XY) - \mathbb{E}(X)\cdot\mu_Y - \mu_X\cdot\mathbb{E}(Y) + \mu_X\mu_Y\\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).\end{aligned}$$

(Note: Compare to $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, proved earlier)

Example (bivariate normal)

Example

If (X, Y) has a bivariate normal distribution with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$, then $\text{Corr}(X, Y) = \rho$.

The proof can be done using the joint density of (X, Y):

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right\}$$

Optional exercise: Show for the case $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 1$,

$$\Rightarrow \quad \mathbb{E}(XY) = \iint_{\mathbb{R}^2} \frac{xy}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)}dx\,dy = \rho,$$

then $\text{Corr}(X, Y) = \textit{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY) = \rho$

Example (binomial)

Example

Let $X \sim \text{Binomial}(n, p)$.

We can write

$$X = X_1 + \cdots + X_n$$

where $X_i = \mathbb{1}_{\text{success on } i\text{th trial}}$.

Then since $X_1, \ldots, X_n$ are mutually independent,

$$\text{Var}(X) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) = n \cdot p(1-p) \,.$$

Example (correlation in a card hand)

Example

Draw a hand of 10 cards.

$X = \#$ Kings, $Y = \#$ Aces.

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

Question to consider before we do the calculation:

Should the covariance be positive or negative?

Example (correlation in card hand)

We can write

$$X = X_1 + \cdots + X_{10}, \qquad Y = Y_1 + \cdots + Y_{10}$$

where

$$X_i = \mathbb{1}_{i\text{th card is King}} \,, \qquad Y_i = \mathbb{1}_{i\text{th card is Ace}} \,.$$

$$\mathbb{E}(X) = \sum_{i=1}^{10} \mathbb{E}(X_i) = \sum_{i=1}^{10} \frac{4}{52} = \frac{40}{52}, \qquad \mathbb{E}(Y) = \frac{40}{52}$$

$$\mathbb{E}(XY) = \sum_{i=1}^{10}\sum_{j=1}^{10} \underbrace{\mathbb{E}(X_iY_j)}_{\substack{=0,\text{ if } i=j,\text{ or} \\ =\frac{4}{52}\cdot\frac{4}{51},\text{ if } i\neq j}} = \sum_{i,j=1,\cdots,10; i\neq j} \mathbb{E}(X_iY_j) = 90 \cdot \frac{4}{52} \cdot \frac{4}{51}$$

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 90 \cdot \frac{4}{52} \cdot \frac{4}{51} - \frac{40}{52} \cdot \frac{40}{52} = -0.0487.$$

Recap and look ahead

We learned

- Joint distribution (order statistics)
- covariance and correlation

Next:

- Conditional expectation

  $\mathbb{E}(X \mid Y = y) = \sum_x x \cdot p_{X|Y}(x \mid y)$ (discrete case)
  $\mathbb{E}(X \mid Y = y) = \int_x x \cdot f_{X|Y}(x \mid y)$ (continuous case)
  $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X))$ (Tower law)

- Conditional variance

  $\text{Var}(X \mid Y = y) = \mathbb{E}[(X - \mathbb{E}(X \mid Y = y))^2 \mid Y = y]$
  $\text{Var}(Y) = \mathbb{E}(\text{Var}(Y \mid X)) + \text{Var}(\mathbb{E}(Y \mid X))$ (Law of total variance)