

# STAT 315: Mean, variance, and covariance for discrete joint RV

Luc Rey-Bellet

University of Massachusetts Amherst

*luc@math.umass.edu*

November 4, 2025

## Expected value of function of joint random variables

If  $Y_1, Y_2$  are joint RV and  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a function then we can compute the expected value of  $g(Y_1, Y_2)$

### Expected value(discrete

For joint random discrete variables  $Y_1$  and  $Y_2$  with joint pdf  $p(y_1, y_2)$  and a function  $g(Y_1, Y_2)$  we have

$$E[g(Y_1, Y_2)] = \sum_{y_1, y_2} g(y_1, y_2) p(y_1, y_2) \quad \text{discrete RV}$$

### Expected value (continuous

For joint random continuous variables  $Y_1$  and  $Y_2$  with joint pdf  $f(y_1, y_2)$  and a function  $g(Y_1, Y_2)$  we have

$$E[g(Y_1, Y_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2) f(y_1, y_2) dy_1 dy_2 \quad \text{continuous RV}$$

# Linearity of expected value

## Linearity

- For any constant  $c$

$$E[c] = c$$

- For any function  $g(Y_1, Y_2)$  and any constant  $c$

$$E[c g(Y_1, Y_2)] = c E[g(Y_1, Y_2)]$$

- For any functions  $g(Y_1, Y_2)$  and  $h(Y_1, Y_2)$

$$E[g(Y_1, Y_2) + h(Y_1, Y_2)] = E[g(Y_1, Y_2)] + E[h(Y_1, Y_2)]$$

Same proof as for  $f(Y)$ !

# Independence and products

## Independence and products

If  $Y_1$  and  $Y_2$  are independent then for any functions  $g(Y_1)$  and  $h(Y_2)$

$$E[g(Y_1)h(Y_2)] = E[g(Y_1)]E[h(Y_2)]$$

For example independence implies that have

$$E[Y_1 Y_2] = E[Y_1]E[Y_2]$$

**Proof:** Independence means  $p(y_1, y_2) = p(y_1)p(y_2)$  and so

$$\begin{aligned} E[g(Y_1)h(Y_2)] &= \sum_{y_1, y_2} g(y_1)h(y_2)p(y_1)p(y_2) \\ &= \sum_{y_1} g(y_1)p(y_1) \sum_{y_2} h(y_2)p(y_2) \\ &= E[g(Y_1)]E[h(Y_2)] \end{aligned}$$

# Covariance

## Covariance of $Y_1$ and $Y_2$

If  $Y_1$  and  $Y_2$  are random variables with means  $\mu_1 = E[Y_1]$  and  $\mu_2 = E[Y_2]$  then the covariance of  $Y_1$  and  $Y_2$  is

$$\text{Cov}(Y_1, Y_2) = E[(Y_1 - \mu_1)(Y_2 - \mu_2)]$$

and the correlation coefficient  $\rho$  is

$$\rho = \rho(Y_1, Y_2) = \frac{\text{Cov}(Y_1, Y_2)}{\sigma_1 \sigma_2}$$

We say that  $Y_1$  and  $Y_2$  are

- positively correlated if  $\text{Cov}(Y_1, Y_2) > 0$
- negatively correlated if  $\text{Cov}(Y_1, Y_2) < 0$
- uncorrelated if  $\text{Cov}(Y_1, Y_2) = 0$

# Properties of covariance

- ① We have the formula

$$\text{Cov}(Y_1, Y_2) = E[Y_1 Y_2] - E[Y_1]E[Y_2]$$

- ②  $\text{Cov}(Y_1, Y_1) = V(Y_1)$  and so  $\rho(Y_1, Y_1) = 1$

- ③ We have Cauchy-Schwartz inequality

$$|E[Z_1 Z_2]| \leq \sqrt{E[Z_1^2]E[Z_2^2]}$$

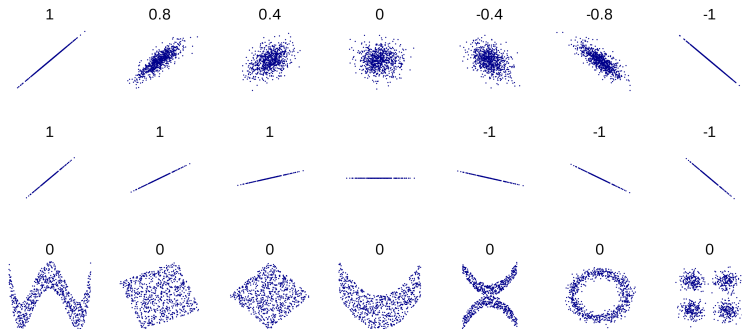
and as a consequence the correlation coefficient satisfies

$$-1 \leq \rho \leq 1$$

- ④ If  $Y_1$  and  $Y_2$  are independent then  $\text{Cov}(Y_1, Y_2) = 0$  and so  $Y_1$  and  $Y_2$  are uncorrelated.

But the converse is not always true

# Example of correlation coefficients



Correlation captures the linear dependence between RV (but not non-linear dependences) (third row)

The correlation reflects the noisiness and direction of a linear relationship (top row), but not the slope of that relationship (second row)

Image taken from [Wikipedia](#)

# Example: Discrete Random Variables

**Story:** A professor records how long students study ( $X$ ) and whether they pass an exam ( $Y$ ).

- $X \in \{0, 1, 2\}$ : number of hours studied
- $Y \in \{0, 1\}$ : pass (1) or fail (0)

**Joint Probability Distribution:**

$X$	$Y = 0$	$Y = 1$	$p_X(x)$
0	0.30	0.05	0.35
1	0.10	0.20	0.30
2	0.05	0.30	0.35
$p_Y(y)$	0.45	0.55	1

**Step 1: Expectations**

$$\mathbb{E}[X] = 0(0.35) + 1(0.30) + 2(0.35) = 1.0, \quad \mathbb{E}[Y] = 0(0.45) + 1(0.55) = 0.55.$$

**Step 2: Mixed Moment**

$$\mathbb{E}[XY] = 0(0.05) + 0(0.30) + 1(0.20) + 2(0.30) = 0.80.$$

**Step 3: Covariance**

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0.80 - (1.0)(0.55) = \boxed{0.25}.$$

**Interpretation:** A positive covariance (0.25) indicates that students who study more tend to have higher probability of passing. If study time and success were independent, covariance would be 0.



## Examples

- For the joint RV  $(X, Y)$  with PDF  $f(x, y) = 2e^{-2x}e^{-y}, x \geq 0, y \geq 0$  compute  $\text{Cov}(X, Y)$ .
- For gas tank problem with the joint RV  $(Y_1, Y_2)$  with joint PDF  $f(y_1, y_2) = 3y_1$  for

$$0 \leq y_2 \leq y_1 \leq 1$$

compute  $\text{Cov}(X, Y)$  and the correlation coefficient  $\rho(Y_1, Y_2)$ .

Compute the mean and the variance of  $Y_1 - Y_2$  (the quantity of unsold gas).

- Suppose  $Y_1$  and  $Y_2$  be the proportion of two chemical in a mixture so that we must have  $Y_1 + Y_2 \leq 1$ . We take the PDF to be

$$f(y_1, y_2) = 2 \quad \text{if } y_1 \geq 0, y_2 \geq 0, y_1 + y_2 \leq 1$$

Compute  $\text{Cov}(Y_1, Y_2)$  and the correlation coefficient  $\rho(Y_1, Y_2)$

## Correlation versus independence

- Take  $X$  uniform on  $[-1, 1]$  and  $Y = X^2$ . Are  $X$  and  $Y$  independent? Compute  $\text{Cov}(X, Y)$ .

- Consider the discrete RVs with joint PDF

$X \backslash Y$	-1	0	1	$P_X(x)$
-1	1/16	3/16	1/16	5/16
0	3/16	0	3/16	6/16
1	1/16	3/16	1/16	5/16
$P_Y(y)$	5/16	6/16	5/16	

Are  $X$  and  $Y$  independent?

Compute  $\text{Cov}(X, Y)$ .

# Linear combinations of random variables

## Linear combination

For random variables  $Y_1, Y_2$  and  $Z_1, Z_2$  and constants  $a_1, a_2$  and  $b_1, b_2$ .

### Expected Value

$$E[a_1 Y_1 + a_2 Y_2] = a_1 E[Y_1] + a_2 E[Y_2]$$

### Variance

$$V(a_1 Y_1 + a_2 Y_2) = a_1^2 V(Y_1) + a_2^2 V(Y_2) + 2a_1 a_2 \text{Cov}(Y_1, Y_2)$$

### Covariance

$$\begin{aligned} \text{Cov}(a_1 Y_1 + a_2 Y_2, b_1 Z_1 + b_2 Z_2) = & a_1 b_1 \text{Cov}(Y_1, Z_1) + \\ & + a_1 b_2 \text{Cov}(Y_1, Z_2) + a_2 b_1 \text{Cov}(Y_2, Z_1) + a_2 b_2 \text{Cov}(Y_2, Z_2) \end{aligned}$$

## Practice: Variance, Covariance, and Correlation

- ➊  $X, Y$  are uncorrelated with  $\text{Var}(X) = 4$ ,  $\text{Var}(Y) = 9$ . Compute  $\text{Var}(2X - 3Y)$ .
- ➋  $X, Y$  have  $\text{Var}(X) = 1$ ,  $\text{Var}(Y) = 4$ ,  $\text{Cov}(X, Y) = 1$ . Compute:
  - (a)  $\text{Var}(X + Y)$
  - (b)  $\text{Var}(2X - Y)$
- ➌ Given  $\text{Var}(X + Y) = 25$ ,  $\text{Var}(X) = 9$ ,  $\text{Var}(Y) = 16$ , find  $\text{Cov}(X, Y)$ .
- ➍  $X_1, X_2, X_3$  are uncorrelated with  $\text{Var}(X_1) = 1$ ,  $\text{Var}(X_2) = 2$ ,  $\text{Var}(X_3) = 3$ . Compute  $\text{Var}(2X_1 - X_2 + 3X_3)$ .
- ➎  $X, Y, Z$  have  $\text{Var}(X) = \text{Var}(Y) = \text{Var}(Z) = 1$  and  $\text{Cov}(X, Y) = \text{Cov}(X, Z) = \text{Cov}(Y, Z) = \rho$ . Find  $\text{Var}(X + Y + Z)$  in terms of  $\rho$ .
- ➏  $X, Y$  have  $\text{Var}(X) = 9$ ,  $\text{Var}(Y) = 4$ ,  $\text{Cov}(X, Y) = -3$ . Compute  $\text{Var}(X + 2Y)$  and discuss the effect of negative covariance.
- ➐ (With correlation)  $E[X] = E[Y] = 0$ ,  $\text{Var}(X) = 4$ ,  $\text{Var}(Y) = 9$ ,  $\rho(X, Y) = 0.5$ . Compute:
  - (a)  $\text{Cov}(X, Y)$
  - (b)  $\text{Var}(2X - Y)$
  - (c)  $\rho(2X - Y, X)$

# Diversify to minimize risk

You are given 2 investments  $X_1$  and  $X_2$  which are independent, say  $X_1 = \text{crypto}$  and  $X_2 = \text{mongolian sheep}$ .

Assume that

$$E[X_1] = E[X_2] \quad \text{same average return}$$

$$V(X_1) = \sigma_1^2 \quad V(X_2) = \sigma_2^2 \quad \text{different risks}$$

: Goal is to minimize the risk of your investment: What should you do?

**Diversify:** Allocate proportion  $\alpha$  to  $X_1$  and  $(1 - \alpha)$  to  $X_2$  with  $0 \leq \alpha \leq 1$

$$X = \alpha X_1 + (1 - \alpha) X_2$$

$$V(X) = \alpha^2 V(X_1) + (1 - \alpha)^2 V(X_2) = \alpha^2 \sigma_1^2 + (1 - \alpha)^2 \sigma_2^2$$

Minimize with respect to  $\alpha$ :

$$0 = 2\alpha\sigma_1^2 - 2(1 - \alpha)\sigma_2^2 \implies \alpha = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

If  $\sigma_1^2 = 10$  (crypto) and  $\sigma_2^2 = 1$  (sheep) you should invest  $\frac{1}{11}$  in crypto and  $\frac{10}{11}$  in sheep.

# Mean and Variance of sample averages

## Empirical or sample average

Suppose  $Y_1, Y_2, \dots, Y_n$  are independent random variables with

$$E[Y_i] = \mu \quad V(Y_1) = \sigma^2$$

Then

$$E \left[ \frac{Y_1 + Y_2 + \dots + Y_n}{n} \right] = \mu$$

and

$$V \left( \frac{Y_1 + Y_2 + \dots + Y_n}{n} \right) = \frac{\sigma^2}{n}$$

Very important for later