STAT 315: Joint PDF of continuous random variables

Luc Rey-Bellet

University of Massachusetts Amherst

luc@math.umass.edu

April 1, 2025

Bivariate or joint random variables

Suppose we have random experiment and make TWO measurements Y_1 and Y_2 or more....

Examples:

- We measure the height and weight of some individual in a population.
- We have an exponential random variable whose parameter β is itself random and obeys a certain distribution.
- We throw a dart at a random position (Y_1, Y_2) on a circular target.
- $\cdots$
- Sampling: We repeat an experiment n times and record the results $Y_1, Y_2, \cdots, Y_n$ of the experiments (the most important example!)

We need to describe the probability distribution of Y_1 and Y_2 together! This is called the joint (or bivariate) PDF $f(y_1, y_2)$ (continuous).

Joint PDF of continuous

Joint (or bivariate) PDF for continuous random variables

The joint continuous RV (Y_1, Y_2) have joint PDF $f(y_1, y_2)$ if

$$P(a_1 \leq Y_1 \leq b_1, a_2 \leq Y_2 \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(y_1, y_2) dy_1, dy_2$$

$$\text{with } 0 \leq f(y_1, y_2) \quad \text{and} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 dy_2 = 1$$

Example 1: The random variables (X, Y) have the joint PDF

$$f(x, y) = \begin{cases} 2e^{-2x} e^{-y} & \text{if } x \geq 0, y \geq 0 \\ 0 & \text{else} \end{cases}$$

Example 2: A gas station adds some gas to its tank every Monday morning. We write Y_1 (in $[0, 1]$) for the proportion of the tank being filled and Y_2 for the proportion of the tank which is sold to customers during the subsequent. Note that we must have $Y_2 \leq Y_1$. We propose the model

$$f(y_1, y_2) = \begin{cases} 3y_1 & \text{if } 0 \leq y_2 \leq y_1 \leq 1 \\ 0 & \text{else} \end{cases}$$

Example 3: Two friends, independently of each other, arrive at a random time between 12pm and 1pm at the blue wall. We can describe the time of their arrival by two random variable Y_1, Y_2 each with a uniform RV on $[0, 1]$ (measured in hours). The independence assumption leads to the model

$$f(y_1, y_2) = \begin{cases} 1 & \text{if } 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1 \\ 0 & \text{else} \end{cases}$$

Example 4: A (pretty bad) player throws a dart at a random point on a circular target of radius R. This can be described by a uniform distribution on a disk of radius R that is by the joint random variables (X_1, X_2) with pdf

$$f(x_1, x_2) = \begin{cases} \frac{1}{\pi R^2} & \text{if } x_1^2 + x_2^2 \leq R^2 \\ 0 & \text{else} \end{cases}$$

Example 5: In Bayesian statistics context one uses random variables whose parameters are themselves random variables. For example consider the joint PDF

$$f(x, y) = \begin{cases} y e^{-yx} e^{-y} & \text{if } x \geq 0, y \geq 0 \\ 0 & \text{else} \end{cases}$$

As we will see this describe a an exponential random variable whose scale parameters (i.e. with pdf $\lambda e^{-\lambda x}$) has a exponential distribution with parameter 1.

Marginal and conditional PDF

Marginal PDF of continuous random variables

If the joint continuous RV (Y_1, Y_2) has PDF $f(y_1, y_2)$ then the marginal PDFs of Y_1 and Y_2 are given by

$$f(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_2 \qquad f(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_1$$

Conditional PDF of continuous random variables

If the joint continuous RV (Y_1, Y_2) has PDF $p(y_1, y_2)$ then the conditional PDFs of Y_1 given $Y_2 = y_2$ is given by

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f(y_2)}$$

if $f(y_2) > 0$

Independence

Recall that the events A and B are independent if

$$P(A|B) = P(A) \text{ or } P(B|A) = P(B) \text{ or } P(A \cap B) = P(A)P(B)$$

Independence of continuous random variables

The continuous random variables Y_1 and Y_2 are independent if

$$f(y_1|y_2) = f(y_1) \text{ or } f(y_2|y_1) = f(y_2) \text{ or } f(y_1, y_2) = f(y_1)f(y_2)$$

Criterion for independence

The random variables Y_1 and Y_2 are independent if and only if

$$f(y_1, y_2) = g(y_1)h(y_2) \quad -\infty < y_1, y_2 < \infty$$

for some function $g(x)$ and $h(y)$

Expected value of function of joint random variables

Expected value

For joint random variables Y_1 and Y_2 and a function $g(Y_1, Y_2)$ we have

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

with properties

Linearity of expectation

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

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:

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

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

and the correlation coefficient ρ is

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

We say that Y_1 and Y_2 are

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

Properties of covariance

1. We have the formula

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

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

3. We have Cauchy-Schwartz inequality

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

and as a consequence the correlation coefficient satisfies

$$-1 \le \rho \le 1$$

4. If Y_1 and Y_2 are independent then $\mathrm{Cov}(Y_1, Y_2) = 0$ and so Y_1 and Y_2 are uncorrelated.
But the converse is not always true

Linear combinations of random variables

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

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

- Covariance

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

Example

- Suppose $V[Y_1] = 2$ and $V[Y_2] = 8$. What are the possible values for the covariance $\mathrm{Cov}(Y_1, Y_2)$?
- What is $\mathrm{Cov}(1, Y)$?
- What is $\mathrm{Cov}(1 + Y, 3 - Y)$?
- Suppose $\rho(Y_1, Y_2) = .2$. What is $\rho(1 + Y_1, 3 - 2Y_2)$
- Suppose X and Y are independent with variance σ_X^2 and σ_Y^2. What is $V[X - Y]$
- In Example 2. $U = Y_1 - Y_2$ is the proportion of unsold gas. Compute the mean and the variance of U.

Illustration of covariances: diversifying your investment

• Think of X_1 and X_2 has the return on investment two investment strategies like investing in stocks, or in bonds, or in crypto, and so on.... If you invest 1 unit in startegy i then you will make a profit X_i.
• For simplicity think of all these strategies are equally good. That is we have $E[X_1] = E[X_2] = \mu$.
• Diversification strategy, pick $0 \le \alpha < 1$ and follow the strategy

$$X_\alpha = \alpha X_1 + (1-\alpha) X_2 \,.$$

We have

$$E[X_\alpha] = E[\alpha X_1 + (1-\alpha)X_2] = \alpha\mu + (1-\alpha)\mu = \mu$$

so they all give the same return, no matter what α is.
• To chose the less risky investment we need to analyze the variance.

- If X_1 and X_2 are independent we have

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

Differentiating to find the optimal α we have

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

For example if $\sigma_1^2 = 1$ and $\sigma_2^2 = 4$ then $\alpha^* = \frac{4}{5}$ and

$$V[X_{\alpha^*}] = (\frac{4}{5})^2 \times 1 + (\frac{1}{5})^2 \times 4 = \frac{4}{5} < 1$$

so the optimal variance is smaller than the smallest variance! In general

$$V[X_{\alpha^*}] = \frac{\sigma_4^2\sigma_1^2 + \sigma_2^2\sigma_1^4}{(\sigma_1^2+\sigma_2^2)^2} = \frac{\sigma_2^2\sigma_1^2}{(\sigma_1^2+\sigma_2^2)} < \sigma_1^2 \quad (\text{or } < \sigma_2^2)$$

- If X_1 and X_2 are negatively correlated then it is even better. We have

$$\mathrm{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2 < 0$$

and so

$$V[X_\alpha] = \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 + 2\alpha(1-\alpha)\rho\sigma_1\sigma_2 < \alpha^2\sigma^1 + (1-\alpha)^2\sigma_2^2$$

- For example if $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, and $\rho = -\frac{1}{2}$ then we have

$$V[X_\alpha] = \alpha^2 + 4(1-\alpha)^2 - 2\alpha(1-\alpha)$$

and differentiating gives $\alpha^* = \frac{5}{7}$ and

$$V[X_{\alpha^*}] = \frac{21}{49}$$

Mean and Variance of sample averages

Empirical or sample average

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

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

Then

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

and

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

Very important!!