STAT 516: Multivariate Distributions

Joint Density Function and its Role

Prof. Michael Levine

April 25, 2020

Basics

- Similarly to the case of a single continuous random variables, we talk about the joint probability density function (pdf) - not about individual pdfs
- Notes of caution:
  1. The joint density function of all the variables does not equal the probability of a specific point in the multidimensional space
  2. The joint density function reflects a relative importance of a particular point
  3. For a general set A in the multidimensional space, the probability that a random vector $\mathbb{X}$ belongs to A is obtained by integrating the joint density function over the set A

Definition

- Let $\mathbb{X} = (X_1, \ldots, X_n)$ be an $n-$ dimensional random vector taking values in R^n for some n, $1 < n < \infty$
- We say that $f(x_1, \ldots, x_n)$ is the joint density or simply the density of $\mathbb{X}$ if, for all $a_1, \ldots, a_n, b_1, \ldots, b_n$ with $-\infty < a_i \leq b_i < \infty$

$$P(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \ldots, a_n \leq X_n \leq b_n)$$
$$= \int_{a_n}^{b_n} \cdots \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_n$$

PDF properties

- In order that a function $f : R^n \to R$ be a density function of some n-dimensional random vector, it is necessary and sufficient that
  1. $f(x_1, \ldots, x_n) \geq 0$ for any $(x_1, \ldots, x_n) \in R^n$
  2. $\int_{R^n} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1$

The joint cumulative density function (CDF)

- Let $\mathbb{X}$ be an $n-$ dimensional random vector with the density function $f(x_1, \ldots, x_n)$. The joint CDF, or simply the CDF, of $\mathbb{X}$ is

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_1} f(t_1, \ldots, t_n) dt_1 \ldots dt_n$$

- Both the CDF and the pdf completely specify the distribution of a continuous random vector
- The relation between the CDF and the pdf is

$$f(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F(x_1, x_2, \ldots, x_n)$$

Marginal densities

- Let $\mathbb{X} = (X_1, X_2, \ldots, X_n)$ be a continuous random vector with a joint density $f(x_1, x_2, \ldots, x_n)$. Let $1 \leq p \leq n$. Then, the marginal joint density of $(X_1, X_2, \ldots, X_p)$ is given by

$$f_{1,2,\ldots,p}(x_1, x_2, \ldots, x_p) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n) dx_{p+1} \ldots, dx_n$$

Independence

- Let $\mathbb{X} = (X_1, X_2, \ldots, X_n)$ be a continuous random vector with a joint density $f(x_1, x_2, \ldots, x_n)$. Then, $X_1, X_2, \ldots, X_n$ are independent if and only if

$$f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i)$$

- To prove one way, you have to integrate both sides of the factorization identity
- To prove the other way, you have to take the classical definition of independence

$$F(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_i(x_i)$$

and partially differentiate both sides successively with respect to $x_1, \ldots, x_n$

Example I

- **Bivariate Uniform** - $f(x, y) = 1$ in the $[0, 1] \times [0, 1]$ rectangle and 0 otherwise
- Note that f is always non-negative; moreover,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = \int_0^1 \int_0^1 f(x, y) dxdy = 1$$

- Thus, f is a valid bivariate density function
- Check that the marginal density

$$f(x) = \int_0^1 f(x, y)\, dy = 1$$

  for any $0 \le x \le 1$; thus, both $X \sim Unif[0, 1]$ and $Y \sim Unif[0, 1]$
- Furthermore, for any x, y $f(x, y) = f_1(x) f_2(y)$ and so X and Y are independent

Example II

- Uniform distribution on a triangle: $f(x, y) = 2$ for any $x, y \geq 0$ and $x + y \leq 1$
- In that case, the marginal density of X is

$$f_1(x) = \int_0^{1-x} 2\,dy = 2(1-x)$$

  for $0 \leq x \leq 1$
- Similarly, the marginal density of Y is $f_2(y) = 2(1-y)$ for $0 \leq y \leq 1$.
- Note that $P\left(X > \frac{1}{2} | Y > \frac{1}{2}\right) = 0$; at the same time,

$$P\left(X > \frac{1}{2}\right) = \int_{1/2}^{1} 2(1-x)\,dx = \frac{1}{4} \neq 0$$

- Thus, X and Y are not independent

Example III

- A joint density $f(x,y) = x\exp(-x(1+y))$ for $x, y \geq 0$
- It is obviously non-negative; note that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_{0}^{\infty}\int_{0}^{\infty} x\exp(-x(1+y))\,dxdy$$
$$= \int_{0}^{\infty} \frac{1}{(1+y)^2}\,dy = \int_{1}^{\infty} \frac{1}{y^2}\,dy = 1$$

- Verify that $f_1(x) = \exp(-x)$ for any $x \geq 0$ and $f_2(y) = \frac{1}{(1+y)^2}$ for any $y \geq 0$
- Thus, the factorization identity is not true and X and Y are not independent