Probability Theory

Chapter 6: Jointly Distributed Random Variables

Lecturer

Dr. Monjed H. Samuh

Applied Mathematics & Physics Department
Palestine Polytechnic University
(monjedsamuh@ppu.edu)

Term 191

Table of Contents

Joint Distribution Functions

- Sometimes we are interested in studying the relationship between two (or more) random variables X and Y.

- **Examples:**

  1. X person's height and Y person's weight.
  2. X person's total cholesterol and Y number of hours the person exercises per week.
  3. X price of an item and Y number of item sold.

Joint Distribution Functions

- For any two random variables X and Y, the joint cumulative probability distribution function (*jcdf*) of X and Y is given by

$$F(x, y) = P(X \leq x, Y \leq y), \quad -\infty < x, y < \infty.$$

- The *cdf* of X can be obtained from the *jcdf* of X and Y as follows:

$$\begin{aligned} F_x(x) &= P\left(X \leq x\right) \\ &= P\left(X \leq x, Y < \infty\right) \\ &= P\left(\lim_{y \to \infty} \{X \leq x, Y \leq y\}\right) \\ &= \lim_{y \to \infty} P\left(X \leq x, Y \leq y\right) \\ &= \lim_{y \to \infty} F(x, y) \\ &= F(x, \infty). \end{aligned}$$

- Similarly, the *cdf* of Y is given by $F_Y(y) = F(\infty, y)$.

- F_X and F_Y are sometimes referred to as the **marginal distributions** of X and Y.

Joint Distribution Functions

- $P(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + F(x, y)$.

  Proof:

$$
\begin{aligned}
P(X > x, Y > y) &= 1 - P\left(\{X > x, Y > y\}^c\right) \\
&= 1 - P\left(\{X > x\}^c \cup \{Y > y\}^c\right) \\
&= 1 - P\left(\{X \leq x\} \cup \{Y \leq y\}\right) \\
&= 1 - [P(X \leq x) + P(Y \leq y) - P(X \leq x, Y \leq y)] \\
&= 1 - F_X(x) - F_Y(y) + F(x, y).
\end{aligned}
$$

- $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1)$, whenever $x_1 < x_2$ and $y_1 < y_2$.

  HWK

Joint Distribution Functions

- In the case when X and Y are both discrete random variables, it is convenient to define the **joint probability mass function** of X and Y by

$$p(x, y) = P(X = x, Y = y)$$

- $\sum_{\forall x} \sum_{\forall y} p(x, y) = 1.$
- The **probability mass function** of X can be obtained from $p(x, y)$ by

$$p_X(x) = P(X = x) = \sum_{y:p(x,y)>0} p(x, y).$$

- Similarly,

$$p_Y(y) = P(Y = y) = \sum_{x:p(x,y)>0} p(x, y).$$

Joint Distribution Functions

Example

Suppose that 3 balls are randomly selected from an urn containing 3 red, 4 white, and 5 blue balls. If we let X and Y denote, respectively, the number of red and white balls chosen, then the joint probability mass function of X and Y,
$p(x, y) = P(X = x, Y = y)$, is given by

Joint Distribution Functions

Example

A local supermarket has three checkout counters. Two customers arrive at the counters at different times when the counters are serving no other customers. Each customer chooses a counter at random, independently of the other. Let Y_1 denote the number of customers who choose counter 1 and Y_2, the number who select counter 2. Find the joint probability function of Y_1 and Y_2.

Joint Distribution Functions

Example

Tossing a fair coin three times; let X denote the number of heads on the first two tosses, and Y denote the number of heads on the three tosses.

(1) Find the joint pdf of X and Y.

(2) Find the marginal pdf of X.

Ans. $\{(0, 2/8), (1, 4/8), (2, 2/8)\}$.

(3) Find the marginal pdf of Y.

Ans. $\{(0, 1/8), (1, 3/8), (2, 3/8), (3, 1/8)\}$.

Joint Distribution Functions

(4) Find $P(X > Y)$. **Ans.** 0.

(5) Find $P(X + Y = 2)$. **Ans.** $2/8$.

Joint Distribution Functions

Definition (Joint Probability Density Function)

We say that X and Y are **jointly continuous** if there exists a function $f(x, y) \geq 0$, defined for all real x and y ($\mathbb{R} \times \mathbb{R}$), having the property that, for every set C of pairs of real numbers

$$P((X, Y) \in C) = \iint_{(x,y)\in C} f(x, y)dxdy, \quad C \subset \mathbb{R}^2.$$

$f(x, y)$ is called the joint *pdf* of X and Y.

- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dxdy = 1.$
- If A and B are any sets of real numbers, then, by defining $C = \{(x, y) : x \in A, y \in B\}$, we have

$$P(X \in A, Y \in B) = \int_B \int_A f(x, y)dxdy = \int_A \int_B f(x, y)dydx.$$

Joint Distribution Functions

- Because

$$\begin{aligned} F(x,y) &= P\left(X \leq x, Y \leq y\right) \\ &= P\left(X \in (-\infty, x], Y \in (-\infty, y]\right) \\ &= \int_{-\infty}^{y} \int_{-\infty}^{x} f(u,v) du dv \end{aligned}$$

it follows, upon differentiation, that

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y),$$

wherever the partial derivatives are defined.

Joint Distribution Functions

Theorem (Marginal *pdf*'s)

If X and Y are jointly continuous, they are individually continuous, and their probability density functions can be obtained as follows:

- *For the random variable X:*

$$P(X \in A) = P\left(X \in A, Y \in (-\infty, \infty)\right)$$
$$= \int_A \int_{-\infty}^{\infty} f(x,y)dydx = \int_A f_X(x)dx,$$

where $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$ is thus the ***probability density function*** *of X.*

- *Similarly, the* ***probability density function*** *of Y is given by*

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

Joint Distribution Functions

Example

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} c & \text{if } 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases},$$

(1) Compute c. **Ans.** $c = 1$.

(2) Compute $P(0.2 < X < 0.5, 0 < Y < 0.7)$. **Ans.** 0.21.

Joint Distribution Functions

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} c & \text{if } 0 < x < 1, 0 < y < 1, y \leq x \\ 0 & \text{otherwise} \end{cases},$$

(1) Compute c.

Ans. $c = 2$.

(2) Compute $P(0 < X < 0.5, 0 < Y < 0.5)$.

Ans. 0.25.

Joint Distribution Functions

(3) Compute $P(0 < X < 0.5, 0 < Y < 0.5)$.

Ans. $c = 0.5$.

(4) Find the marginal *pdf* of X.

Ans. $2x,\ 0 < x < 1$.

(5) Find the marginal *pdf* of Y.

Ans. $2(1 - y),\ 0 < y < 1$.

Joint Distribution Functions

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2 & \text{if } 0 < x < 1, 0 < y < 1, 0 < x + y < 1 \\ 0 & \text{otherwise} \end{cases},$$

(1) Compute $P(X < 3/4, Y < 3/4)$. **Ans.** ....

Joint Distribution Functions

(2) Compute $P(X < 0.5, Y < 0.5)$.

Ans. 0.5.

(3) Compute $P(0 < X < 0.5)$.

Ans. 0.75.

(4) Find the marginal *pdf* of Y.

Ans. $Y \sim Beta(1, 2)$.

Joint Distribution Functions

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} kxy & \text{if } x > 0, y > 0, x + y \leq 1 \\ 0 & \text{otherwise} \end{cases},$$

(1) Compute k.

Ans. $k = 24$.

Joint Distribution Functions

(2) Find the marginal *pdf* of X.

Ans. $f_X(x) = 12x(1-x)^2, \;\; 0 \leq x \leq 1.$

(3) Find the marginal *pdf* of Y.

Ans. $f_Y(y) = 12y(1-y)^2, \;\; 0 \leq y \leq 1.$

(4) Find $E(X)$, $E(Y)$, $Var(X)$, and $Var(Y)$.

Joint Distribution Functions

Example

The joint density function of X and Y is given by

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

(1) Compute $P(X > 1, Y < 1)$.

Ans. $e^{-1}(1 - e^{-2})$.

Joint Distribution Functions

(2) $P(X < Y)$.

Ans. $1/3$.

(3) $P(X < a)$.

Ans. $1 - e^{-a}$.

Joint Distribution Functions

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} e^{-(x+y)} & \text{if } x > 0, y > 0, \\ 0 & \text{otherwise} \end{cases},$$

Find the density function of the random variable $Z = \frac{X}{Y}$.

Independent Random Variables

Definition (Independence)

The random variables X and Y are said to be **independent** if,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B), \quad \forall A, B \subset \mathbb{R}.$$

In other words, X and Y are independent if the events $E_A = \{X \in A\}$ and $E_B = \{X \in B\}$ are independent.

- Let $A = (-\infty, x]$ and $B = (-\infty, y]$, then X and Y are independent iff

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y), \quad \forall x, y.$$

- In terms of F, X and Y are independent iff

$$F(x, y) = F_X(x)F_Y(y), \quad \forall x, y.$$

Independent Random Variables

- When X and Y are discrete random variables, the condition of independence is equivalent to
$$p(x, y) = p_X(x)p_Y(y), \quad \forall x, y.$$

- When X and Y are continuous random variables, the condition of independence is equivalent to
$$f(x, y) = f_X(x)f_Y(y), \quad \forall x, y.$$

- Thus, X and Y are independent if knowing the value of one does not change the distribution of the other.

- Random variables that are not independent are said to be **dependent**.

Independent Random Variables

Example

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} 6e^{-2x}e^{-3y} & \text{if } x > 0, y > 0 \\ 0 & \text{otherwise} \end{cases},$$

Are these random variables independent?

Independent Random Variables

Example

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} 24xy & \text{for } 0 < x < 1,\ 0 < y < 1,\ 0 < x + y < 1 \\ 0 & \text{otherwise} \end{cases},$$

Are X and Y independent?

Independent Random Variables

Definition

In general, the n random variables $X_1, X_2, \dots, X_n$ are said to be independent if, for all sets of real numbers $A_1, A_2, \dots, A_n$,

$$P\left(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n\right) = \prod_{i=1}^{n} P\left(X_i \in A_i\right)$$

Equivalently,

$$P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) = \prod_{i=1}^{n} P\left(X_i \leq x_i\right), \quad \forall x_1, x_2, \dots, x_n.$$

Independent Random Variables

Example

Let X, Y, Z be independent and uniformly distributed over $(0, 1)$. Compute $P(X \geq YZ)$.

Ans. $3/4$.

Sums of Independent Random Variables

Example

If X and Y are independent random variables, both uniformly distributed on $(0, 1)$, calculate the probability density of $X + Y$.

Solution:

- Let $Z = X + Y$, then
- the *pdf* of Z is given by

$$f_Z(z) = \begin{cases} z & \text{for } 0 < z < 1 \\ 2 - z & \text{for } 1 < z < 2 \\ 0 & \text{otherwise} \end{cases},$$

- The random variable Z is said to have a triangular distribution.

Sums of Independent Random Variables

Theorem (Closed under Convolutions)

If X and Y are independent gamma random variables with respective parameters (α_1, λ) and (α_2, λ), then $X + Y$ is a gamma random variable with parameters $(\alpha_1 + \alpha_2, \lambda)$.

Proof: *Next Chapter.*

Theorem

If $X_1, X_2, \ldots, X_n$ are independent gamma random variables with respective parameters (α_i, λ), then $\sum_{i=1}^{n} X_i$ is gamma with parameters $(\sum_{i=1}^{n} \alpha_i, \lambda)$.

Theorem

If $X_1, X_2, \ldots, X_n$, are independent random variables that are normally distributed with respective parameters μ_i, σ_i^2, $i = 1, \ldots, n$, then $\sum_{i=1}^{n} X_i$ is normally distributed with parameters $\sum_{i=1}^{n} \mu_i$ and $\sum_{i=1}^{n} \sigma_i^2$.

Proof: *Next Chapter.*

Sums of Independent Random Variables

Example

Let X_1 and X_2 be independent random variables with common exponential density with $\lambda = 1$, Find the pdf of $Y = X_1 + X_2$.

Ans. $Y \sim Ga(2, 1)$.

Sums of Independent Random Variables

Example

Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables, each having parameter λ, then $\sum_{i=1}^{n} X_i$ is gamma with parameters (n, λ).

Sums of Independent Random Variables

Example

Let $Z \sim N(0,1)$, then find the *pdf* of Z^2.

$$f_{Z^2}(y) = \frac{1}{2\sqrt{y}}\left[f_Z(\sqrt{y}) + f_Z(-\sqrt{y})\right]$$
$$= \frac{\frac{1}{2}e^{-y/2}(y/2)^{1/2-1}}{\sqrt{\pi}}.$$

That is $Z^2 \sim Ga(\frac{1}{2}, \frac{1}{2}) = \chi^2_1$.

Sums of Independent Random Variables

Theorem

If $Z_1, Z_2, \ldots, Z_n$ are independent standard normal random variables, then $\sum_{i=1}^{n} Z_i^2$ is distributed as $\chi_n^2 = Ga(\frac{n}{2}, \frac{1}{2})$.

Theorem

If $X_1, X_2, \ldots, X_n$, are independent Poisson random variables with respective parameters λ_i, $i = 1, \ldots, n$, then $\sum_{i=1}^{n} X_i$ is distributed as Poisson with parameter $\sum_{i=1}^{n} \lambda_i$.

Proof: *Next Chapter.*

Theorem

If $X_1, X_2, \ldots, X_n$, are independent binomial random variables, each having parameters (m_i, p), $i = 1, \ldots, n$, then $\sum_{i=1}^{n} X_i$ is distributed as binomial with parameters $\left(\sum_{i=1}^{n} m_i, p\right)$.

Proof: *Next Chapter.*

Conditional Distributions: Discrete Case

- For any two events A and B, the conditional probability of A given B is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

- Let $A = \{X = x\}$ and $B = \{Y = x\}$, then we have the following definition.

Definition

- If X and Y are discrete random variables with joint pmf $p(x, y)$, then we define the **conditional probability mass function** of X given that $Y = y$, by

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)}.$$

- Similarly, the **conditional probability distribution function** of X given that $Y = y$ is defined by

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \sum_{x:X \leq x} p_{X|Y}(x|y).$$

Conditional Distributions: Discrete Case

- **Excercise**: Verify that $p_{X|Y}(x|y)$ is a pmf.

- For any events A and B,

$$P(X \in A|Y = y) = \sum_{x \in A} p_{X|Y}(x|y), \quad A \subset R_X,$$

  and

$$P(Y \in B|X = x) = \sum_{y \in B} p_{Y|X}(y|x), \quad B \subset R_Y.$$

- If X and Y are independent, then the conditional probability mass function of X given $Y = y$ is equal to the unconditioned probability mass function:

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)} = \frac{p_X(x)p_Y(y)}{p_Y(y)} = p_X(x).$$

Conditional Distributions: Discrete Case

Example

Suppose that $p(x, y)$, the joint probability mass function of X and Y, is given by

$$p(0,0) = 0.4,\ p(0,1) = 0.2,\ p(1,0) = 0.1,\ p(1,1) = 0.3.$$

1. Calculate the conditional probability mass function of X given that $Y = 1$.

   Ans. $\{(0, 2/5), (1, 3/5)\}$.

2. Compute $P(0 \leq X \leq 1 | Y = 1)$.

   Ans. 1.

Conditional Distributions: Discrete Case

Example

A balanced die with 3 faces, numbered 1,2,3 is rolled twice. Let X be the min of the two numbers obtained, and Y be the max of the two numbers obtained.

1. Obtain the joint pdf of (X, Y).

   Ans. $\{(1, 1, 1/9), (1, 2, 2/9), (1, 3, 2/9), (2, 2, 1/9), (2, 3, 2/9), (3, 3, 1/9)\}$.

2. Compute $P(X = x|Y = 2)$.

   Ans. $\{(1, 2/3), (2, 1/3)\}$.

3. Compute $P(Y = y|X = 1)$.

   Ans. $\{(1, 1/5), (2, 2/5), (3, 2/5)\}$.

Conditional Distributions: Discrete Case

Example

If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, calculate the conditional distribution of X given that $X + Y = n$.

Ans. $(X|X+Y) \sim Bin\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$.

Conditional Distributions: Continuous Case

Definition

- If X and Y are continuous random variables with joint pdf $f(x, y)$, then we define the **conditional probability density function** of X given that $Y = y$, by

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

- If X and Y are jointly continuous, then, for any set A:

$$P(X \in A|Y = y) = \int_A f_{X|Y}(x|y)dx.$$

- In particular, by letting $A = (-\infty, a]$, we can define the **conditional cumulative distribution function** of X given that $Y = y$ by

$$F_{X|Y}(a|y) = P(X \leq a|Y = y) = \int_{-\infty}^{a} f_{X|Y}(x|y)dx.$$

- If X and Y are independent continuous random variables, then the conditional pdf of X given $Y = y$ is just the unconditional density of X.

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x).$$

Conditional Distributions: Continuous Case

Example

The joint density of X and Y is given by:

$$f(x,y) = \begin{cases} \frac{12}{5}x(2-x-y) & \text{for } 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases},$$

Compute the conditional density of X given that $Y = y$, where $0 < y < 1$.

Ans. $f_{X|Y}(x|y) = \frac{x(2-x-y)}{\frac{2}{3}-\frac{y}{2}}$.

- For example, when $y = \frac{1}{3}$: $f_{X|Y}\left(x\middle|\frac{1}{3}\right) = \frac{10}{3}x - 2x^2, \quad 0 < x < 1.$
- For checking: $\int_0^1 f_{X|Y}\left(x\middle|\frac{1}{3}\right) dx = 1.$

Conditional Distributions: Continuous Case

Example

The joint density of X and Y is given by:

$$f(x,y) = \begin{cases} \frac{x+y}{4} & \text{for } 0 < x < y < 2 \\ 0 & \text{otherwise} \end{cases},$$

(1) Find the conditional density of X given that $Y = y$.

Ans. $f_{X|Y}(x|y) = \frac{2(x+y)}{3y^2}$.

- Note that $\int_0^y f_{X|Y}(x|y) = 1$.

Conditional Distributions: Continuous Case

(2) Compute $f_{X|Y}(x|1.5)$.

Ans. $\frac{8}{27}(x + 1.5), \quad 0 < x < 1.5.$

(3) Compute $P(X < 1|Y = 1.5)$.

Ans. $\frac{16}{27}$.

Conditional Distributions: Continuous Case

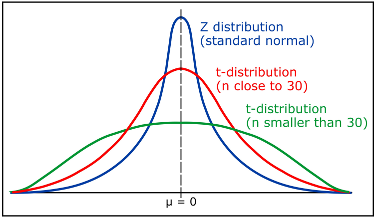
Definition (The t-Distribution)

If Z and Y are independent, with Z having a standard normal distribution and Y having a χ^2-distribution with n degrees of freedom, then the random variable T defined by

$$T = \frac{Z}{\sqrt{Y/n}}$$

is said to have a ***t*-distribution** with n degrees of freedom. Its density function is given by

$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}, \quad -\infty < t < \infty.$$



Conditional Distributions: Continuous Case

Definition (The Bivariate Normal Distribution)

Random variables X and Y are said to have a **bivariate normal distribution** if their joint *pdf* has the form

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]\right\},$$

where ρ is the correlation between X and Y (More about this next chapter).

Exercises:

- $X \sim N(\mu_X, \sigma_X^2)$.
- $Y \sim N(\mu_Y, \sigma_Y^2)$.
- $X|Y \sim N\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y),\ \sigma_X^2(1-\rho^2)\right)$.
- $Y|X \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X),\ \sigma_Y^2(1-\rho^2)\right)$.

Conditional Distributions: Continuous Case

Example

Let $(X, Y) \sim BN(160, 60, 225, 256, 0.7)$.

1. Compute $P(X < 175)$.

   Ans. $X \sim N(160, 225)$.

2. Compute $P(X < 175|y = 65)$.

   Ans. $\{X|(y = 65)\} \sim N(163.3, 114.8)$.

Order Statistics

Definition (Order Statistics)

Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed (*iid*) continuous random variables having a common density f and distribution function F. Define

- $X_{(1)} = \min(X_1, X_2, \ldots, X_n)$.
- $X_{(2)} = 2^{nd} \min(X_1, X_2, \ldots, X_n)$.

  $\vdots$
- $X_{(j)} = j^{th} \min(X_1, X_2, \ldots, X_n)$.

  $\vdots$
- $X_{(n)} = \max(X_1, X_2, \ldots, X_n)$.

The ordered values $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ are known as the **order statistics** corresponding to the random variables $X_1, X_2, \ldots, X_n$.

Order Statistics

- For two random variables, X_1 and X_2, we have:

  1. $P(X_1 \leq X_2)=$ ______.
  2. $P(X_{(1)} \leq X_{(2)})=$ ______.

- For three random variables, X_1, X_2, and X_3, we have:

  1. $P(X_1 \leq X_2 \leq X_3)=$ ______.
  2. $P(X_{(1)} \leq X_{(2)} \leq X_{(3)})=$ ______.

$\vdots$

- For n random variables, $X_1, X_2, \ldots, X_n$, we have:

  1. $P(X_1 \leq X_2 \leq \cdots \leq X_n)=$ ______.
  2. $P(X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)})=$ ______.

Order Statistics

Example

Let $X_1, X_2, \ldots, X_n$ be *iid* $f(x)$. Let $Y = \max(X_1, X_2, \ldots, X_n) = X_{(n)}$. Find the density function of Y.

Ans. $f_Y(y) = nF^{n-1}(y)f(y)$.

Order Statistics

Example

Let $X_1, X_2, \ldots, X_n$ be *iid* $f(x)$. Let $Y = \min(X_1, X_2, \ldots, X_n) = X_{(1)}$. Find the density function of Y.

Ans. $f_Y(y) = n(1 - F(y))^{n-1} f(y)$.

Order Statistics

Example

Let $X_1, X_2, \ldots, X_n$ be *iid* $U(0,1)$. Let $Y_1 = X_{(1)}$ and $Y_n = X_{(n)}$.

1. Find the density function of Y_1.

   Ans. $Y_1 \sim Be(1, n)$.

2. Find the density function of Y_n.

   Ans. $Y_n \sim Be(n, 1)$.

3. Compute $E(Y_n)$.

   Ans. $E(Y_n) = \frac{n}{n+1}$.

Order Statistics

- The *jpdf* of $Y_1 = X_{(1)}, Y_2 = X_{(2)}, \ldots, Y_n = X_{(n)}$ is given by

$$f_{Y_1,Y_2,\ldots,Y_n}(y_1, y_2, \ldots, y_n) = n! f(y_1) f(y_2) \cdots f(y_n), \quad y_1 < y_2 < \cdots < y_n.$$

Proof: Let us consider the case $n = 2$ and find the joint pdf for Y_1 and Y_2.

Order Statistics

- The density function of the j^{th}-order statistic $Y_j = X_{(j)}$ is given by

$$f_{Y_j}(y) = \binom{n}{j-1,\ n-j,\ 1} [F(y)]^{j-1} [1 - F(y)]^{n-j} f(y).$$

Order Statistics

- The joint pdf of the order statistics $Y_i = X_{(i)}$ and $Y_j = X_{(j)}$ when $i < j$ is

$$f_{Y_i, Y_j}(y_i, y_j) = \binom{n}{i-1,\ j-i-1,\ n-j} [F(y_i)]^{i-1} \left[F(y_j) - F(y_i)\right]^{j-i-1} \left[1 - F(y_j)\right]^{n-j} f(y_i)f(y_j), \quad y_i < y_j.$$

Example (Distribution of the range of a random sample)

Suppose that n iid random variables $X_1, X_2, \ldots, X_n$ are observed. The random variable R defined by $R = X_{(n)} - X_{(1)}$ is called the range of the observed random variables. If the random variables X_i have distribution function F and density function f, then find the distribution of R.

Joint Probability Distribution of Functions of Random Variables

Theorem (Transformation Technique (univariate case))

Let X be a continuous random variable having pdf $f_X(x)$. Suppose that $g(x)$ is a strictly monotonic (increasing or decreasing), differentiable (and thus continuous) function of x. Then the random variable Y defined by $Y = g(X)$ has a pdf given by

$$f_Y(y) = \left|\frac{d}{dy} g^{-1}(y)\right| f_X\left(g^{-1}(y)\right),$$

where $g^{-1}(y)$ is defined to equal that value of x such that $g(x) = y$.

Example

Let X have the pdf given by $f_X(x) = 2x$, $0 < x < 1$. Find the pdf of $Y = -4X + 3$.

Ans. $f_Y(y) = (3 - y)/8, \; -1 < y < 3.$

Joint Probability Distribution of Functions of Random Variables

- Let (X_1, X_2) be jointly continuous random variables with joint pdf $f_{X_1,X_2}(x_1, x_2)$. It is sometimes necessary to obtain the joint distribution of the random variables Y_1 and Y_2, which arise as functions of X_1 and X_2.

- Specifically, let $Y_1 = g_1(X_1, X_2)$ and $Y_2 = g_2(X_1, X_2)$ for some functions g_1 and g_2. Assume that the functions g_1 and g_2 satisfy the following conditions:

  1. The equations $y_1 = g_1(x_1, x_2)$ and $y_2 = g_2(x_1, x_2)$ can be uniquely solved for x_1 and x_2 in terms of y_1 and y_2, with solutions given by, say, $x_1 = h_1(y_1, y_2)$, $x_2 = h_2(y_1, y_2)$.

  2. The functions g_1 and g_2 have continuous partial derivatives at all points (x_1, x_2) and are such that the 2×2 determinant

$$J(x_1, x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \frac{\partial g_1}{\partial x_1}\frac{\partial g_2}{\partial x_2} - \frac{\partial g_1}{\partial x_2}\frac{\partial g_2}{\partial x_1} \neq 0$$

  at all point (x_1, x_2).

- Under these two conditions, it can be shown that the random variables Y_1 and Y_2 are jointly continuous with joint density function given by

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2)\left|J(x_1, x_2)\right|^{-1},$$

where $x_1 = h_1(y_1, y_2)$, and $x_2 = h_2(y_1, y_2)$.

Example

Let X_1 and X_2 be jointly continuous random variables with probability density function f_{X_1,X_2}. Let $Y_1 = X_1 + X_2$, $Y_2 = X_1 - X_2$. Find the joint density function of Y_1 and Y_2 in terms of f_{X_1,X_2}.

Ans. $f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2} f_{X_1,X_2}\left(\frac{1}{2}(y_1 + y_2), \frac{1}{2}(y_1 - y_2)\right)$.

Joint Probability Distribution of Functions of Random Variables

- If X_1 and X_2 are *iid* $U(0,1)$, then

$$f_{Y_1,Y_2}(y_1,y_2)=\begin{cases}\frac{1}{2} & \text{for } 0<y_1+y_2<2,\ 0<y_1-y_2<2\\ 0 & \text{otherwise}\end{cases},$$

- If X_1 and X_2 are independent exponential random variables with respective parameters λ_1 and λ_2, then

$$f_{Y_1,Y_2}(y_1,y_2)=\begin{cases}\frac{\lambda_1\lambda_2}{2}\exp\left\{-\lambda_1\left(\frac{y_1+y_2}{2}\right)-\lambda_2\left(\frac{y_1-y_2}{2}\right)\right\} & \text{for } 0<y_1+y_2,\ 0<y_1-y_2\\ 0 & \text{otherwise}\end{cases},$$

Joint Probability Distribution of Functions of Random Variables

- If X_1 and X_2 are *iid* $N(0,1)$, then

$$\begin{aligned} f_{Y_1,Y_2}(y_1,y_2) &= \frac{1}{4\pi} e^{-\left[\left(\frac{y_1+y_2}{8}\right)^2+\left(\frac{y_1-y_2}{8}\right)^2\right]} \\ &= \frac{1}{\sqrt{4\pi}} e^{-\frac{y_1^2}{4}} \frac{1}{\sqrt{4\pi}} e^{-\frac{y_2^2}{4}}, \quad -\infty < y_1, y_2 < \infty. \end{aligned}$$

  - $X_1 + X_2$ and $X_1 - X_2$ are normal with mean 0 and variance 2.
  - $X_1 + X_2$ and $X_1 - X_2$ are independent.

Theorem

If X_1 and X_2 are independent random variables having a common distribution function F, then $X_1 + X_2$ will be independent of $X_1 - X_2$ if and only if F is a normal distribution function.

Example

If X and Y are independent gamma random variables with parameters (α, λ) and (β, λ), respectively, compute the joint density of $U = X + Y$ and $V = \frac{X}{X+Y}$.

Solution:

- The joint density of X and Y is given by

$$f_{X,Y}(x,y) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda(x+y)} x^{\alpha-1} y^{\beta-1}, \quad x, y > 0.$$

- If $g_1(x,y) = x + y$ and $g_2(x,y) = \frac{x}{x+y}$, then

$$J(x,y) = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & \frac{-x}{(x+y)^2} \end{vmatrix} = \frac{-1}{x+y}$$

Joint Probability Distribution of Functions of Random Variables

Example

Solution (Cont'd):

- As the equations $u = x + y$, $v = \frac{x}{x+y}$ have their solutions $x = uv$ and $y = u(1 - v)$, we see that

$$\begin{aligned} f_{U,V}(u, v) &= uf_{X,Y}(uv, u(1 - v)) \\ &= \frac{\lambda e^{-\lambda u}(\lambda u)^{\alpha+\beta-1}}{\Gamma(\alpha + \beta)} \frac{\Gamma(\alpha + \beta) v^{\alpha-1}(1 - v)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}, \quad u > 0, 0 < v < 1. \end{aligned}$$

- $X + Y$ and $\frac{X}{(X+Y)}$ are independent.

- $X + Y$ having a gamma distribution with parameters $(\alpha + \beta, \lambda)$ and $\frac{X}{(X+Y)}$ having a beta distribution with parameters (α, β).

Example

Let X_1 and X_2 be independent with $f_{X_1,X_2}(x_1, x_2) = e^{-x_1}e^{-x_2}$, $x_1, x_2 > 0$. Let $Y = X_1 + X_2$. Obtain the pdf of Y.

Ans. $f_Y(y) = ye^{-y}, \; y > 0.$

Example

Let X_1 and X_2 be iid $N(0,1)$. Obtain the pdf of $Y = \frac{X_2}{X_1}$.

Ans. Cauchy with $\theta = 0$; $f_Y(y) = \frac{1}{\pi(1+y^2)}, \; -\infty < y < \infty$.

Example

Let X_1, X_2, and X_3 be independent standard normal random variables. If $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_1 - X_2$, and $Y_3 = X_1 - X_3$, compute the joint density function of $Y_1,\ Y_2,\ Y_3$.

Solution:

- The joint density of X_1, X_2 and X_3 is given by

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}\sum_{i=1}^{3} x_i}, \quad -\infty < x_1, x_2, x_3 < \infty.$$

- If $g_1(x_1, x_2, x_3) = x_1 + x_2 + x_3$, $g_2(x_1, x_2, x_3) = x_1 - x_2$, and $g_3(x_1, x_2, x_3) = x_1 - x_3$, then

$$J(x_1, x_2, x_3) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{vmatrix} = 3.$$

Example

Solution (Cont'd):

- As the equations $y_1 = x_1 + x_2 + x_3$, $y_2 = x_1 - x_2$, and $y_3 = x_1 - x_3$ have their solutions $x_1 = \frac{y_1+y_2+y_3}{3}$, $x_2 = \frac{y_1-2y_2+y_3}{3}$, and $x_3 = \frac{y_1+y_2-2y_3}{3}$ we see that

$$\begin{aligned} f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) &= \frac{1}{3} f_{X_1,X_2,X_3}\left(\frac{y_1+y_2+y_3}{3}, \frac{y_1-2y_2+y_3}{3}, \frac{y_1+y_2-2y_3}{3}\right) \\ &= \frac{1}{3(2\pi)^{3/2}} e^{-\frac{1}{2}Q(y_1,y_2,y_3)}, \quad -\infty < y_1, y_2, y_3 < \infty, \end{aligned}$$

where

$$\begin{aligned} Q(y_1,y_2,y_3) &= \left(\frac{y_1+y_2+y_3}{3}\right)^2 + \left(\frac{y_1-2y_2+y_3}{3}\right)^2 + \left(\frac{y_1+y_2-2y_3}{3}\right)^2 \\ &= \frac{1}{3}y_1^2 + \frac{2}{3}y_2^2 + \frac{2}{3}y_3^2 - \frac{2}{3}y_2 y_3. \end{aligned}$$

- See Example 7e Page 284.

Problems and Exercises

PROBLEMS

PAGE 290:

2, 6, 7, 8, 15, 17, 19, 20, 21, 27, 40, 41, 44, 55, 58, 59

THEORETICAL EXERCISES

PAGE 294:

5, 7, 9, 28, 35