Random Variables and Distributions

Chapter 3 of Probability and Statistics

Yiping Cheng
ypcheng@bjtu.edu.cn
19-Nov-2019

School of Electronic and Information Engineering
Beijing Jiaotong University

Chapter Contents

Concepts of Random Variables and Distributions

Introducing Random Variables

The need for random variables comes from the fact that the sample space does not necessarily fit into the real line, which makes quantitative studies cumbersome.

Example

Tossing a Coin. A fair coin is tossed 10 times. In this experiment, the sample space S consists of the 2^{10} different sequences of 10 heads and/or tails that are possible. Let X stand for the real-valued function defined on S that counts the number of heads in each outcome.
For example, if s is the sequence HHTTTHTTTH, then $X(s) = 4$. For each possible sequence s consisting of 10 heads and/or tails, the value $X(s)$ equals the number of heads in the sequence. The possible values for the function X are $0, 1, \dots, 10$.

Definition of Random Variable

Definition: Let S be the sample space. A real-valued function that is defined on S is called a random variable.

$$S \xrightarrow{X} R$$

Examples

- In the previous example, the function X is a random variable. Another function, $Y = 10 - X$, which represents the number of tails, is also a random variable.
- **Measuring a Person's Height**. Consider an experiment in which a person is selected at random from some population and her height in inches is measured. This height is a random variable.

Distribution of a Random Variable

Remember that the sample space is associated with a probability measure. We will have the full distribution information of a random variable X, if given any subset C of the real line such that $X \in C$ is an event, we know its probability.

Definition [Distribution]: Let X be a random variable. The distribution of X is the collection of all probabilities of the form $P(X \in C)$ for all sets C of real numbers such that $X \in C$ is an event, where it is defined

$$X \in C = \{s \in S : X(s) \in C\}.$$

Remark

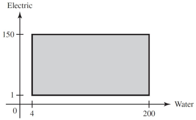
It is a consequence of the definition of the distribution of X that the distribution is itself a probability measure on the set of real numbers.

Distribution of a Random Variable – Examples

- **Tossing a Coin**. Consider again an experiment that a fair coin is tossed 10 times, and let X be the number of heads obtained. The possible values of X are $0, 1, 2, \ldots, 10$.

$$P(X = x) = \binom{10}{x} \frac{1}{2^{10}}, x = 0, 1, \ldots, 10$$

- **Demands for Utilities**. The shaded region in the figure below shows the sample space. Probability of an event $= \frac{\text{area of the event}}{\text{area of the sample space}}$.



Let X be the random variable that $X(s) = x$ when $s = (x, y)$. The distribution of X consists of all probabilities of the form $P(X \in C)$ for all sets C such that $X \in C$ is an event. In particular, if C is a subinterval of the interval $[4, 200]$, then

$$P(X \in C) = \frac{\text{length of interval C}}{200 - 4}.$$

Definition of the Distribution Function

The previous definition of distribution is awkward. It turns out that a distribution function contains all the information of distribution.
Definition [Distribution Function]: Let X be a random variable (abbreviated RV). The *(cumulative) distribution function* (abbreviated *c.d.f.*) of X is the function

$$F(x) = P(X \leq x) \text{ for } -\infty < x < \infty.$$

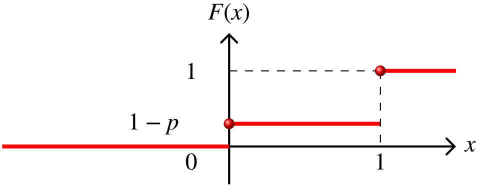
Remark

It should be emphasized that the distribution function given above is for EVERY random variable. Some may ask: What if for some x, the set $X \leq x$ is not an event? Actually, in a rigorous definition of random variable, it is required $X \leq x$ be an event for all x. Our previously given definition of random variable is, strictly speaking, not rigorous.

Distribution Function – Example

RV X takes only two values 0 and 1 with $P(X = 1) = p$. Find its c.d.f.
Solution: Since $P(X = 1) = p$ and $P(X = 0) = 1 - p$, we have

- for $x < 0$, $F(x) = P(\emptyset) = 0$,
- for $0 \le x < 1$, $F(x) = P(X \le x) = P(X = 0) = 1 - p$,
- for $x \ge 1$, $F(x) = P(X = 0 \text{ or } X = 1) = P(X = 0) + P(X = 1) = 1$.



In summary, $F(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$.

Continuity of Probability

Suppose $A_1 \supset A_2 \supset \cdots$. Then $\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$.

Proof.

Let $B_i = A_{i-1} - A_i$ for $i \geq 2$. It is obvious that the B_i's are mutually disjoint, and $A_1 - \bigcap_{n=1}^{\infty} A_n = \bigcup_{i=2}^{\infty} B_i$, and thus $P(A_1 - \bigcap_{n=1}^{\infty} A_n) = \sum_{i=2}^{\infty} P(B_i)$. This implies

$$P(A_1 - \bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \sum_{i=2}^{n} P(B_i) = \lim_{n\to\infty} P(\bigcup_{i=2}^{n} B_i) = \lim_{n\to\infty} P(A_1 - A_n),$$

$$P(A_1) - P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \left(P(A_1) - P(A_n)\right),$$

$$P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n).$$

□

Properties of the Distribution Function – 1,2

1. **Nondecreasing**: The function $F(x)$ is nondecreasing as x increases; that is, if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$.

Proof.

If $x_1 < x_2$, then the event $X \leq x_1$ is a subset of the event $X \leq x_2$. Hence, $P(X \leq x_1) \leq P(X \leq x_2)$, by property 4 of probability. □

2. **Limits at $\pm\infty$**: $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

Proof.

Let $A_n = \{\omega | X(\omega) \leq -n\}$. Then $A_1 \supset A_2 \supset \cdots$, and $\bigcap_{n=1}^{\infty} A_n = \emptyset$. By continuity of probability, $\lim_{n\to-\infty} F(n) = \lim_{n\to\infty} F(-n) = \lim_{n\to\infty} P(A_n) = 0$.
Let $B_n = \{\omega | X(\omega) > n\}$. Then $B_1 \supset B_2 \supset \cdots$, and $\bigcap_{n=1}^{\infty} B_n = \emptyset$. By continuity of probability, $\lim_{n\to\infty} 1 - F(n) = \lim_{n\to\infty} P(B_n) = 0$, thus $\lim_{n\to\infty} F(n) = 1$. □

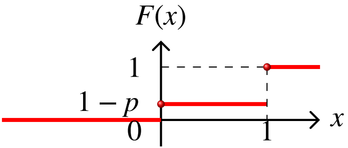
Properties of the Distribution Function – 3

The distribution function is not necessarily continuous, but it is guaranteed right-continuous.

3. **Continuity from the Right**: A c.d.f. is always continuous from the right; that is, $F(x) = F(x^+)$ at every point x.

Proof.

Let $x_1 > x_2 > \cdots$ and $\lim_{n\to\infty} x_n = x$. Then $F(x_n) = P(X \leq x_n) = P(X \leq x \bigcup x < X \leq x_n) = F(x) + P(x < X \leq x_n)$. By continuity of probability, we have $\lim\limits_{n\to\infty} P(x < X \leq x_n) = 0$, and hence $\lim\limits_{n\to\infty} F(x_n) = F(x)$. □



Determining Probabilities from the Distribution Function

1. For every value x, $P(X > x) = 1 - F(x)$.
2. For all values x_1 and x_2 such that $x_1 < x_2$,

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1).$$

3. For every value x, $P(X < x) = F(x^-)$.

Proof.

$P(X < x) = P(X \leq x - \frac{1}{n} \bigcup x - \frac{1}{n} < X < x) = F(x - \frac{1}{n}) + P(x - \frac{1}{n} < X < x)$.
Since $\lim_{n\to\infty} P(x - \frac{1}{n} < X < x) = 0$, we have
$P(X < x) = \lim_{n\to\infty} F(x - \frac{1}{n}) = F(x^-)$. □

4. For every value x, $P(X = x) = F(x) - F(x^-)$.

Classification of Univariate Distributions

1. Discrete Distributions
2. Continuous Distributions
3. Mixed Distributions

Definition of Discrete Distributions

Definition [Discrete Distribution/RV]: We say that an RV X has a *discrete distribution* or that X is a *discrete random variable* if there is

- a finite set of real numbers $\{x_1, x_2, \ldots, x_n\}$ such that $\sum_{i=1}^{n} P(X = x_i) = 1$, or
- a countable set of real numbers $\{x_1, x_2, \ldots\}$ such that $\sum_{i=1}^{\infty} P(X = x_i) = 1$.

Remark

- This definition is slightly different from the textbook version in that here we include the possibility that X may take other values; it is only that the probability of taking other values is zero.
- According to this definition, whether an RV is discrete or not depends only on its distribution, whereas the textbook definition does not have this property.

PF of Discrete Distributions

Definition [Probability Function/p.f.]: If an RV X has a discrete distribution, the *probability function* (abbreviated p.f.) of X is defined as the function f such that for every real number x, $f(x) = P(X = x)$.
Theorem: Let X be a discrete random variable with p.f. $f(x)$. If x is not in the set of real numbers mentioned in the definition, then $f(x) = 0$.
Theorem: If X has a discrete distribution, the probability of each subset C of the real line can be determined from the relation

$$P(X \in C) = \sum_{x_i \in C} f(x_i).$$

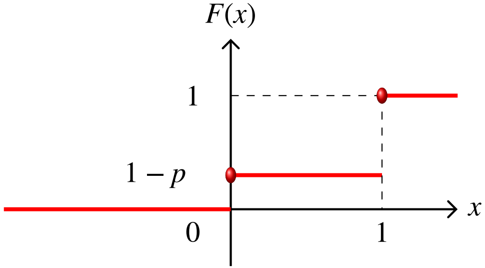
Observation: If X has a discrete distribution with p.f. $f(x)$. Then $F(x)$ has a staircase shape, having a jump of magnitude $f(x_i)$ at each x_i where $f(x_i) \neq 0$, and $F(x)$ will be constant between every pair of successive jumps.

Some Useful Discrete Distributions

1. Bernoulli Distributions
2. Binomial Distributions
3. Uniform Distributions on Integers
4. Hypergeometric Distributions (Chapter 5)
5. Poisson Distributions (Chapter 5)
6. Negative Binomial Distributions (Chapter 5)

Bernoulli Distributions

A random variable X that takes the value 1 with probability p and the value 0 with probability $q = 1 - p$ is said to have the Bernoulli distribution with parameter p.
Its distribution function is shown below:



Binomial Distributions

Consider a general experiment that consists of observing n independent repititions (trials) with only two possible results for each trial. For convenience, call the two possible results success and failure. Then the distribution of the number X of trials that result in success will be binomial with parameters n and p, where p is the probability of success on each trial.

The p.f. of X will be as follows:

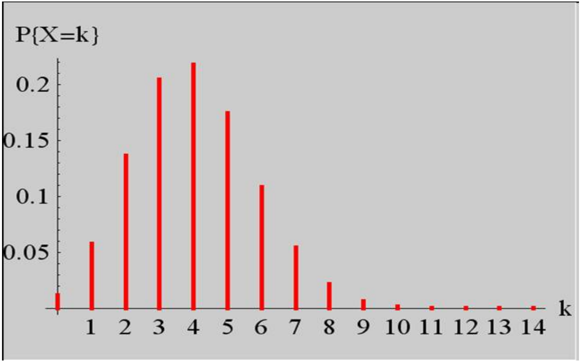
$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, .., n \\ 0 & \text{otherwise} \end{cases}$$

The number of heads in a sequence of 10 independent tosses of a fair coin, has the binomial distribution with parameters 10 and 1/2.

$$f(x) = P(X = x) = \binom{10}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{10-x}, x = 0, 1, \ldots, 10.$$

Graph of a PF of Bonomial Distribution

This graph shows the p.f. of the binomial distribution with $(n, p) = (20, 0.2)$.



Uniform Distribution on Integers

Let $a \leq b$ be integers. Suppose that the value of an RV X is equally likely to be each of the integers $a, ..., b$. Then we say that X has the uniform distribution on the integers $a, ..., b$.

$$f(x) = \begin{cases} \frac{1}{b-a+1} & x = a, ..., b \\ 0 & \text{otherwise} \end{cases}$$

Note: Random Variables Can Have the Same Distribution without Being the Same Random Variable.

Definition of Continuous Distributions

Definition [Continuous Distribution/RV]: We say that an RV X has a *continuous distribution* or that X is a *continuous random variable* if there exists a nonnegative function f defined on the real line, such that for all x,

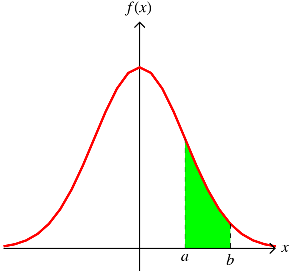
$$P(X \le x) = \int_{-\infty}^{x} f(t)dt.$$

Remark

- This definition is equivalent to the textbook version.
- According to this definition, the distribution function of a continuous distribution must be continuous. Actually, it is even stronger than that. The distribution function must be absolutely continuous.

PDF of Continuous Distributions

Definition [Probability Density Function/p.d.f.]: If an RV X has a continuous distribution, the function described in the definition of continuous distribution/RV is called the *probability density function* (abbreviated p.d.f.) of X.



Properties of PDF and Continuous Distribution

1. Obviously, according to definition,
   1. $f(x) \geq 0$,
   2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

2. The probability that X takes a value in the interval is the integral of f over the interval:

$$P(a < X \leq b) = \int_a^b f(x)dx.$$

3. Continuous distributions assign probability 0 to individual values. If X has a continuous distribution, then

$$\begin{aligned} P(X = a) &= 0 \text{ for all number } a \text{, and} \\ P(a < X \leq b) &= P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b). \end{aligned}$$

Relationship between PDF and CDF

- Let X have a continuous distribution. $F(x)$ is its c.d.f., and $f(x)$ is its p.d.f. Then F is continuous at every x,

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

- The values of each p.d.f. can be changed at a finite, or even countably infinite, set of points, without changing the value of the integral of the p.d.f. over any subset. Therefore for a continuous c.d.f. F, there are infinitely many p.d.f. f's.
- We shall give only one version of the p.d.f. and refer to this version as *the* p.d.f, of the distribution or the RV. This version is usually chosen to be as continuous as possible. At least, it must be chosen to ensure

$$f(x) = \frac{dF(x)}{dx} \text{ wherever } \frac{dF(x)}{dx} \text{ exists.}$$

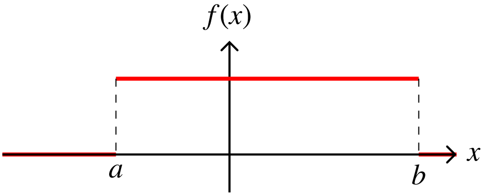
Some Useful Continuous Distributions

1. Uniform Distributions on Intervals
2. Normal Distributions (Chapter 5)
3. Gamma Distributions (Chapter 5)
4. Beta Distributions (Chapter 5)

Uniform Distributions on Intervals

Let a and b be two given real numbers such that $a < b$. Let X be a random variable such that it is known that $a \leq X \leq b$ and, for every subinterval of $[a, b]$, the probability that X will belong to that subinterval is proportional to the length of that subinterval. We then say that the random variable X has the uniform distribution on the interval $[a, b]$.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$



Note: Density Is Not Probability.

Other Examples of Continuous Distributions

Incompletely Specified p.d.f. Suppose that the p.d.f. of a certain random variable X has the following form:

$$f(x) = \begin{cases} cx & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases}$$

(1) Determine the value of c. (2) Find $P(1 \le X \le 2), P(X > 2)$.

Solution: (1) For every p.d.f., it must be true that $\int_{-\infty}^{\infty} f(x)dx = 1$, $\int_{-\infty}^{\infty} f(x)dx = \int_0^4 cxdx = 8c = 1$, Hence $c = \frac{1}{8}$.

(2) $P(1 \le X \le 2) = \int_1^2 \frac{1}{8}xdx = \frac{3}{16}$,
$P(X > 2) = \int_2^{\infty} f(x)dx = \int_2^4 \frac{1}{8}xdx = \frac{3}{4}$.

Other Examples of Continuous Distributions (Continued)

Unbounded RVs. The following function is a density function.

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0, \\ \frac{1}{(1+x)^2} & \text{for } x > 0. \end{cases}$$

Unbounded p.d.f.'s. The following function is a density function.

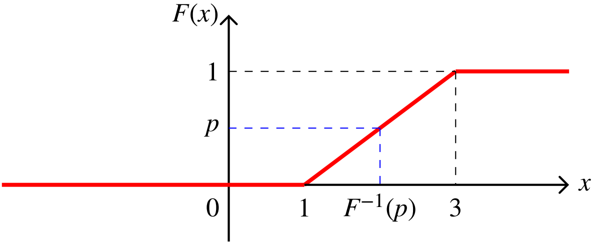
$$f(x) = \begin{cases} \frac{2}{3}x^{-1/3} & 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Quantile for Continuous Distributions

Given distribution function F, for each p strictly between 0 and 1, the p quantile of F is defined to be the smallest value x such that $F(x) \geq p$. When the distribution is continuous and one-to-one over the whole set of possible values of X, the inverse $F^{-1}(p)$ of F exists and equals the quantile function of F.

$$P(X \leq F^{-1}(p)) = p.$$

Example:



Mixed Distributions

- There are three extreme types of distributions: Discrete, Continuous, and Singular. The former two types of distributions have been given definitions here. It has been proved that every distribution can be expressed as a linear combination of three distributions which are discrete, continuous, and singular, respectively.

- Here we give one example of discrete-continuous mixed distribution (Truncated Voltage example in the textbook):

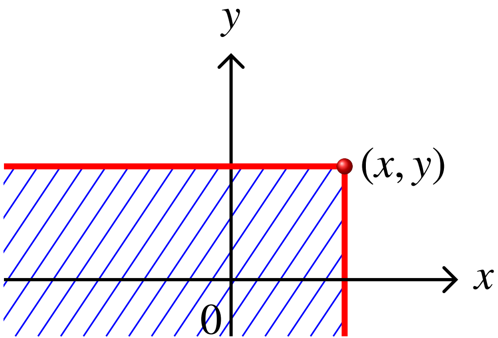
$$F(x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{x}{4}, & \text{if } 0 \le x < 3, \\ 1, & \text{if } x \ge 3. \end{cases}$$

Bivariate, Marginal, and Conditional Distributions

Distribution Function for Two RVs

Definition: The joint distribution function or joint cumulative distribution function (joint c.d.f.) of two random variables X and Y is defined as the function $F : \mathbb{R}^2 \mapsto [0, 1]$ such that for all x and y

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



Properties of Bivariate Distribution Functions 1–3

1. **Nondecreasing**: $F(x, y)$ is nondecreasing over both x and y.
2. **Limits at $\pm\infty$**: For all x, y,

$$\begin{aligned} F(-\infty, -\infty) &= F(-\infty, y) = F(x, -\infty) = 0 \\ F(\infty, \infty) &= 1, \\ F(x, \infty) &= F_X(x), \\ F(\infty, y) &= F_Y(y). \end{aligned}$$

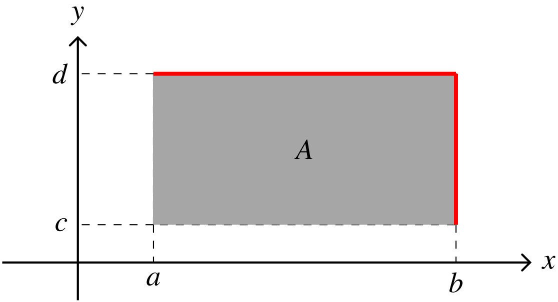
   where F_X and F_Y are the (univariate) distribution functions of X and Y, respectively.
3. **Continuity from the Right**:

$$F(x, y) = F(x^+, y^+).$$

Properties of Bivariate Distribution Functions 4

4. **Rectangular Inequality**: For any given numbers $a < b$, $c < d$,

$$F(b,d) - F(a,d) - F(b,c) + F(a,c) \geq 0.$$



Definition of Discrete Bivariate Distributions

Definition [Discrete Joint/Bivariate Distribution]: We say that two RVs X and Y have a *discrete joint distribution* if there is

- a finite set of real number pairs $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ such that $\sum_{i=1}^{n} P(X = x_i, Y = y_i) = 1$, or
- a countable set of real number pairs $\{(x_1, y_1), (x_2, y_2), \dots\}$ such that $\sum_{i=1}^{\infty} P(X = x_i, Y = y_i) = 1$.

Definition [Probability Function/p.f.]: If RVs X and Y have a discrete joint distribution, the *probability function* (abbreviated p.f.) of X and Y is defined as the function f such that for every point (x, y),

$$f(x, y) = P(X = x, Y = y).$$

Discrete Bivariate Distribution – Example

The joint p.f. f of X and Y is as specified in the following table:

Table: Joint p.f.

	$y=1$	$y=2$	$y=3$	$y=4$
$x=1$	0.1	0	0.1	0
$x=2$	0.3	0	0.1	0.2
$x=3$	0	0.2	0	0

Find $P(X \geq 2, Y \geq 2)$ and $P(X = 1)$.
Solution:

$$\begin{aligned} P(X \geq 2, Y \geq 2) &= f(2,2)+f(2,3)+f(2,4)+f(3,2) \\ &\quad +f(3,3)+f(3,4) \\ &= 0.5 \\ P(X=1) &= \sum_{y=1}^{4} f(1,y) = 0.2 \end{aligned}$$

Equivalence of Two Conditions

Question: What is the relationship between the two conditions:

A RVs X and Y have a discrete joint distribution.

B RV X has a discrete distribution and RV Y has a discrete distribution.

Answer: They are equivalent. Why?

Definition of Continuous Bivariate Distributions

Definition [Continuous Joint/Bivariate Distribution]: We say that two RVs X and Y have a *continuous joint distribution* if there exists a nonnegative function $f(x, y)$ defined over the entire xy-plane such that for all x, y,

$$P(X \le x, Y \le y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t, u) dt du.$$

The function f is called the joint probability density function (abbreviated joint p.d.f.) of X and Y.

Properties of Bivariate PDF and Continuous Distribution

1. Obviously, according to definition,
   1. $f(x, y) \geq 0$,
   2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.
2. The probability that (X, Y) takes a value in a subset C of the plane is the integral of f over C:

$$P((X, Y) \in C) = \int \int_C f(x, y) dx dy.$$

3. Continuous distributions assign probability 0 to any individual point even (even curved) line. So it does not matter whether the boundary points are included in the subset when computing the probability of (X, Y) in the subset.
4. The joint p.d.f. can be derived from the joint c.d.f. by using the relations

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

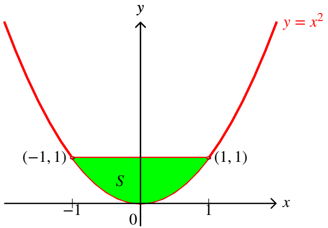
at every point (x, y) at which these second-order derivatives exist.

Continuous Bivariate Distributions – Example

Suppose that the joint p.d.f. of X and Y is specified as follows:

$$f(x,y) = \begin{cases} cx^2y & x^2 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(1) Determine the value of the constant c.
(2) Determine the value of $P(X \ge Y)$.



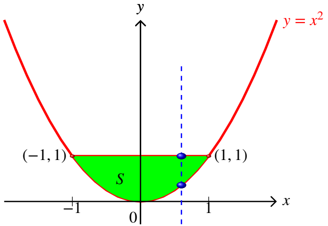
Continuous Bivariate Distributions – Example (Continued)

Solution: (1) The support S of (X, Y) is sketched. Since $f(x, y) = 0$ outside S, it follows that

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

$$= \int_{-1}^{1}\left[\int_{x^2}^{1} cx^2ydy\right]dx = \int_{-1}^{1}\left[cx^2\left(\frac{1}{2} - \frac{x^4}{2}\right)\right]dx = \frac{4}{21}c.$$

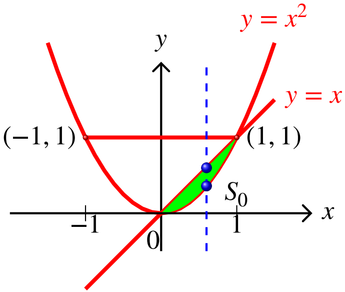
Thus, $c = \frac{21}{4}$. $f(x, y) = \begin{cases} \frac{21x^2y}{4} & x^2 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$.



Continuous Bivariate Distributions – Example (Continued 2)

(2)

$$
\begin{aligned}
P(X \geq Y) &= \int\int_{S_0} f(x,y)dxdy = \int_0^1 \left[\int_{x^2}^{x} \frac{21x^2y}{4} dy\right] dx \\
&= \int_0^1 \left[\frac{21}{8}(x^4 - x^6)\right] dx = \frac{3}{20}.
\end{aligned}
$$



Non-Equivalence of Two Conditions

Question: What is the relationship between the two conditions:

A RVs X and Y have a continuous joint distribution.

B RV X has a continuous distribution and RV Y has a continuous distribution.

Answer: They are NOT equivalent.

Marginal Distributions

Often, we start with a joint distribution of two random variables and then want to find the distribution of just one of them. The latter distribution is called *marginal distribution*, relative to the joint distribution.

Theorem: If X and Y have a discrete joint distribution for which the joint p.f. is f, then the (marginal) p.f. f_X of X is

$$f_X(x) = \sum_{\text{All } y} f(x, y).$$

Similarly, the (marginal) p.f. f_Y of Y is

$$f_Y(y) = \sum_{\text{All } x} f(x, y).$$

Marginal Distributions (Continued)

Theorem: If X and Y have a continuous joint distribution with joint p.d.f. f, then the (marginal) p.d.f. $f_X(x)$ of X is

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy, \quad -\infty < x < \infty.$$

Similarly, the (marginal) p.d.f. $f_Y(y)$ of Y is

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

Proof.

$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(t,y)dtdy = \int_{-\infty}^{x} dt \int_{-\infty}^{\infty} f(t,y)dy$. Therefore by definition of p.d.f, $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$. Similarly, $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$. □

Marginal Distributions - Example

The joint p.f. $f(x,y)$ of X and Y is as specified in the following table:

	$y=1$	$y=2$	$y=3$	$y=4$
$x=1$	0.1	0	0.1	0
$x=2$	0.3	0	0.1	0.2
$x=3$	0	0.2	0	0

Find the marginal distribution $f_X(x)$.
Solution: The p.f. f_X of X is

$$\begin{aligned} f_X(1) &= 0.1+0+0.1+0=0.2 \\ f_X(2) &= 0.3+0+0.1+0.2=0.6 \\ f_X(3) &= 0+0.2+0+0=0.2 \end{aligned}$$

Table: p.f. f_X of X

x	1	2	3
$f_X(x)$	0.2	0.6	0.2

Independence of Two Random Variables

Definition: It is said that two random variables X and Y are independent, if for all x, y,

$$F(x, y) = F_X(x)F_Y(y)$$

where F is the joint c.d.f. of X and Y, and F_X and F_Y are the c.d.f.s of X and Y, respectively.

Remark

This definition is to say that for all x, y,

$$P(X \leq x \cap Y \leq y) = P(X \leq x)P(Y \leq y).$$

It is able to ensure that, for every two sets A and B of real numbers such that $X \in A$ and $Y \in B$ are events,

$$P(X \in A \cap Y \in B) = P(X \in A)P(Y \in B).$$

Criteria of Independence of Two RVs

It is more usual p.f.s or p.d.f.s are used for specifying a distribution. So we need criteria of independence using p.f.s or p.d.f.s.

Discrete: Let X and Y have a discrete joint distribution with joint p.f. f. Let f_X and f_Y be the p.f.s of X and Y, respectively. Then X and Y are independent if and only if for all real numbers x and y,

$$f(x,y) = f_X(x)f_Y(y).$$

Continuous: Let X and Y have a continuous joint distribution. Then X and Y are independent if and only if the joint p.d.f. f of X and Y, the p.d.f. f_X of X, and the p.d.f. f_Y of Y, can be chosen so that for all real numbers x and y,

$$f(x,y) = f_X(x)f_Y(y).$$

Checking Independence - Example 1

Determine whether X and Y are independent or not.

Table: Joint p.f. with marginals

	$y=1$	$y=2$	$y=3$	$y=4$	Marginal
$x=1$	0.1	0	0.1	0	0.2
$x=2$	0.3	0	0.1	0.2	0.6
$x=3$	0	0.2	0	0	0.2
Marginal	0.4	0.2	0.2	0.2	1

$$0 = f(1,2) \neq f_X(1) \times f_Y(2) = 0.2 \times 0.2$$

Thus, X and Y are not independent.

Checking Independence - Example 2

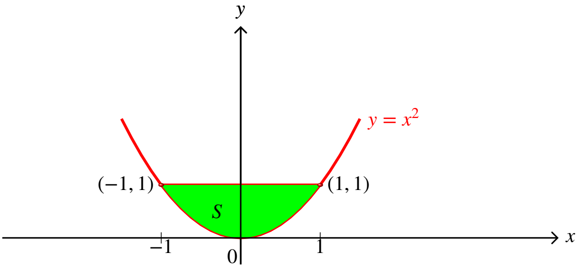
Suppose that the joint p.d.f. of X and Y is specified as

$$f(x,y) = \begin{cases} \frac{21}{4}x^2y & x^2 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

(1) Determine the marginal p.d.f. of X.

(2) Determine the marginal p.d.f. of Y.

(3) Check if X and Y are independent or not.

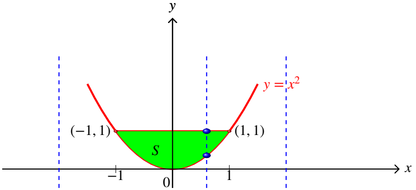


Checking Independence - Example 2 (Continued)

Solution: (1)

- If $x < -1$ or $x > 1$, $f_X(x) = 0$
- If $-1 \leq x \leq 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{x^2}^{1} \frac{21}{4}x^2 y dy = \frac{21}{8}x^2(1-x^4).$$



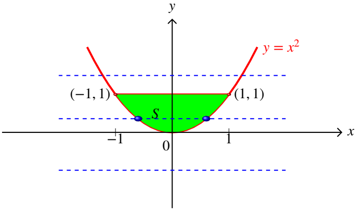
In summary, $f_X(x) = \begin{cases} \frac{21}{8}x^2(1-x^4) & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$.

Checking Independence - Example 2 (Continued 2)

(2)

- If $y < 0$ or $y > 1$, $f_Y(y) = 0$
- If $0 \leq y \leq 1$,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2 y dx = \frac{7}{2}y^{5/2}.$$



In summary, $f_Y(y) = \begin{cases} \frac{7}{2}y^{5/2} & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}.$

Checking Independence - Example 2 (Continued 3)

$$(3)\, f(x,y) = \begin{cases} \frac{21}{4}x^2y & x^2 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

$$f_X(x) = \begin{cases} \frac{21}{8}x^2(1-x^4) & -1 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

$$f_Y(y) = \begin{cases} \frac{7}{2}y^{5/2} & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

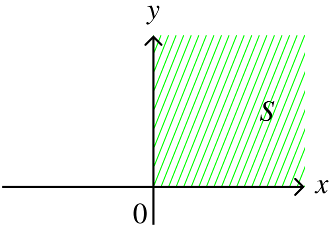
$\frac{21}{4}x^2y \neq \frac{21}{8}x^2(1-x^4) \times \frac{7}{2}y^{5/2}$, $f(x,y) \neq f_1(x) \times f_2(y)$. Therefore X and Y are not independent.

Checking Independence - Example 3

Suppose that the joint p.d.f. of X and Y is as follows:

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

(1) Determine the marginal p.d.f. of X.
(2) Determine the marginal p.d.f. of Y.
(3) Check if X and Y are independent or not.

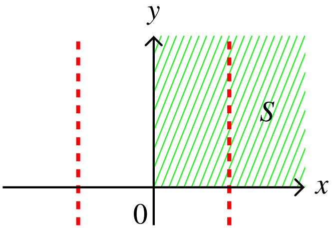


Checking Independence - Example 3 (Continued)

Solution: (1)

- If $x < 0, f_X(x) = 0$
- If $x \geq 0,$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^{\infty} 2e^{-x-2y}dy = e^{-x}.$$



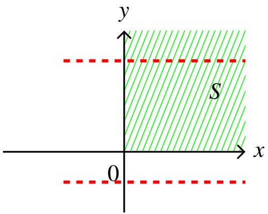
In summary, $f_X(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}.$

Checking Independence - Example 3 (Continued 2)

(2)

- If $y < 0, f_Y(y) = 0$
- If $y \geq 0$,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^{\infty} 2e^{-x-2y}dx = 2e^{-2y}.$$



In summary, $f_Y(y) = \begin{cases} 2e^{-2y} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}.$

Checking Independence - Example 3 (Continued 3)

(3) $f(x,y) = f_X(x) \times f_Y(y)$, X and Y are independent.

Note: If X and Y are independent, then $h(X)$ and $g(Y)$ are independent no matter what the functions h and g are.

Discrete Conditional Distributions

Definition: Let X and Y have a discrete joint distribution with joint p.f. f. For each y such that $f_Y(y) > 0$, define conditional p.f. of X given $Y = y$ as

$$\begin{aligned} f_{X|Y}(x|y) &= P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{f(x,y)}{f_Y(y)} = \frac{f(x,y)}{\sum_{\text{All } x} f(x,y)}. \end{aligned}$$

According to this definition, for each x such that $f_X(x) > 0$, the conditional p.f. of Y given $X = x$ is

$$\begin{aligned} f_{Y|X}(y|x) &= P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} \\ &= \frac{f(x,y)}{f_X(x)} = \frac{f(x,y)}{\sum_{\text{All } y} f(x,y)}. \end{aligned}$$

Discrete Conditional Distributions – Example

Table: Joint p.f. with marginals

	$y=1$	$y=2$	$y=3$	$y=4$	Marginal
$x=1$	0.1	0	0.1	0	0.2
$x=2$	0.3	0	0.1	0.2	0.6
$x=3$	0	0.2	0	0	0.2
Marginal	0.4	0.2	0.2	0.2	1

The joint p.f of (X, Y) is given. Determine
(1) the conditional p.f. of X given that $Y=1$;
(2) the conditional p.f. of Y given that $X=2$.

Solution: (1) From the table, $f_Y(1) = 0.4$. The general equation for the conditional probabilities $f_{X|Y}(x|1)$ is

$$f_{X|Y}(x|1) = \frac{f(x,1)}{f_Y(1)} = \frac{f(x,1)}{0.4}.$$

We then have

$$f_{X|Y}(1|1) = \frac{1}{4}, \quad f_{X|Y}(2|1) = \frac{3}{4}, \quad f_{X|Y}(3|1) = 0.$$

Table: Conditional probabilities $f_{X|Y}(x|1)$

	$x=1$	$x=2$	$x=3$
$f_{X\|Y}(x\|1)$	$\frac{1}{4}$	$\frac{3}{4}$	0

Exercise: Find $f_{X|Y}(x,2)$, $f_{X|Y}(x,3)$, $f_{X|Y}(x,4)$.

(2) From the table, $f_X(2) = 0.6$. The general equation for the conditional probabilities $f_{Y|X}(y|2)$ is

$$f_{Y|X}(y|2) = \frac{f(2,y)}{f_X(2)} = \frac{f(2,y)}{0.6}.$$

We then have

$$f_{Y|X}(1|2) = \frac{1}{2}, f_{Y|X}(2|2) = 0, f_{Y|X}(3|2) = \frac{1}{6}, f_{Y|X}(4|2) = \frac{1}{3}.$$

Table: Conditional probabilities $f_{Y|X}(y|2)$

	$y = 1$	$y = 2$	$y = 3$	$y = 4$
$f_{Y\|X}(y\|2)$	$\frac{1}{2}$	0	$\frac{1}{6}$	$\frac{1}{3}$

Exercise: Find $f_{Y|X}(y|1), f_{Y|X}(y|3)$.

Continuous Conditional Distributions

Definition: Let X and Y have a continuous joint distribution with joint p.d.f. f. For each y such that $f_Y(y) > 0$, define conditional p.d.f. of X given $Y = y$ as

$$
\begin{aligned}
f_{X|Y}(x|y) &= \lim_{\Delta x \to 0, \Delta y \to 0} \frac{P(X \in [x, x+\Delta x] | Y \in [y, y+\Delta y])}{\Delta x} \\
&= \lim_{\Delta x \to 0, \Delta y \to 0} \frac{P(X \in [x, x+\Delta x], Y \in [y, y+\Delta y])}{P(Y \in [y, y+\Delta y])\Delta x} \\
&= \frac{f(x,y)}{f_Y(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx}.
\end{aligned}
$$

According to this definition, for each x such that $f_X(x) > 0$, the conditional p.d.f. of Y given $X = x$ is

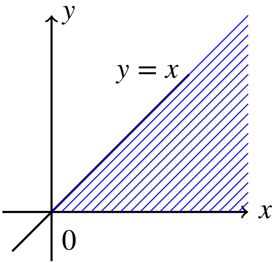
$$
f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dy}.
$$

Continuous Conditional Distributions – Example

Suppose that X and Y have a joint continuous distribution with joint p.d.f.

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

Find (1) $f_{X|Y}(x|y)$, (2) $f_{Y|X}(y|x)$, (3) $P(X < 2|Y = 1)$.

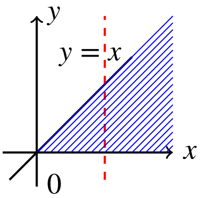


Conditional Distributions – Example 2 (Continued)

Solution: First, determine the marginal p.d.f. $f_X(x)$ of X:

- If $x < 0, f_X(x) = 0$
- If $x \geq 0$,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^x e^{-x}dy = xe^{-x}.$$



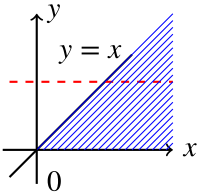
In summary, $f_X(x) = \begin{cases} xe^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}.$

Conditional Distributions – Example 2 (Continued 2)

Then, determine the marginal p.d.f. $f_Y(y)$ of Y:

- If $y < 0, f_Y(y) = 0$
- If $y \geq 0$,

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



In summary, $f_Y(y) = \begin{cases} e^{-y} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}.$

Conditional Distributions – Example 2 (Continued 3)

1. If $y \geq 0$, then $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} e^{y-x}, & 0 \leq y \leq x \\ 0, & \text{otherwise.} \end{cases}$
2. If $x \geq 0$, then $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \frac{1}{x}, & 0 < x, y \leq x \\ 0, & \text{otherwise.} \end{cases}$
3. $P(X < 2|Y = 1) = \int_{-\infty}^{2} f_{X|Y}(x|1)dx = \int_{1}^{2} e^{1-x}dx = 1 - e^{-1}.$

Remark

- The term $P(X < 2|Y = 1)$ is illegal in strict probability theory sense, since $P(Y = 1) = 0$. It should be understood as the limit $\lim_{\epsilon \to 0} P(X < 2||Y - 1| < \epsilon)$.
- $P(a < X < b|Y = y) = \int_a^b f_{X|Y}(x|y)dx$. Conditional Distributions Behave Just Like Distributions.

Multivariate Distributions

All the concepts for two random variables extend naturally to more than two random variables.

Functions of One or More Random Variables

1. Functions of One Random Variable
2. Functions of Two or More Random Variables

Functions of One Random Variable

Suppose X is an RV whose distribution is known. $Y = r(X)$ where r is a known function. We want to find the distribution of Y.

- RV X with a Discrete Distribution
- RV X with a Continuous Distribution

RV X with a Discrete Distribution

The process is best illustrated with an example. Let X have a discrete distribution with p.f. $f(x)$ shown below. Find the p.f. $g(y)$ of $Y = |X - 5|$.

x	1	2	3	4	5	6	7	8	9
$f(x)$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Solution: The possible values of Y are 0, 1, 2, 3, and 4.

$$\begin{aligned}
g(0) = P(Y = 0) = P(|X - 5| = 0) = P(X = 5) = 1/9, \\
g(1) = P(Y = 1) = P(|X - 5| = 1) = P(X = 4 \cup X = 6) = 2/9, \\
g(2) = P(Y = 2) = P(|X - 5| = 2) = P(X = 3 \cup X = 7) = 2/9, \\
g(3) = P(Y = 3) = P(|X - 5| = 3) = P(X = 2 \cup X = 8) = 2/9, \\
g(4) = P(Y = 4) = P(|X - 5| = 4) = P(X = 1 \cup X = 9) = 2/9.
\end{aligned}$$

y	0	1	2	3	4
$g(y)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$

Note: $g(y) = P(Y = y) = P(r(X) = y) = \sum_{x:\, r(x)=y} f(x)$.

RV X with a Continuous Distribution - Method 1

This method is based on basic theory. Suppose that the p.d.f. of X is $f(x)$ and that another random variable is defined as $Y = r(X)$. The c.d.f. $G(y)$ of Y can be derived as

$$G(y) = P(Y \leq y) = P(r(X) \leq y) = \int_{x:\, r(x)\leq y} f(x)dx.$$

If the random variable Y also has a continuous distribution, its p.d.f. $g(y)$ can be obtained from the relation

$$g(y) = \frac{dG(y)}{dy}.$$

This relation is satisfied at every point y where G is differentiable.

RV X with a Continuous Distribution - Method 1 Example

Let X have uniform distribution on interval $[-1, 1]$, so

$$f(x) = \begin{cases} \frac{1}{2} & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Determine the p.d.f. of the random variable $Y = X^2$.
Solution: Since $Y = X^2, -1 \leq X \leq 1$, then Y must be in the interval $[0, 1]$.
(1) If $y < 0$, $G(y) = P(Y \leq y) = P(\emptyset) = 0$.
(2) If $y > 1$, $G(y) = P(Y \leq y) = P(S) = 1$.
(3) If $0 \leq y \leq 1$, $G(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx = \sqrt{y}$.

Therefore the p.d.f. $g(y)$ of Y is $g(y) = \begin{cases} \frac{1}{2y^{1/2}} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}.$

RV X with a Continuous Distribution - Method 2

This method is applicable only when r is one-to-one and differentiable but it obtains the p.d.f. of Y directly.

Theorem: Let X be an RV for which the p.d.f. is $f(x)$ and for which $P(a < X < b) = 1$. Let $Y = r(X)$, and suppose that $r(x)$ is differentiable and one-to-one for $a < x < b$. Let (α, β) be the image of the interval (a, b) under the function r. Let $s(y)$ be the inverse function of $r(x)$, Then the p.d.f. $g(y)$ of Y is

$$g(y) = \begin{cases} f[s(y)] \times \left|\frac{ds(y)}{dy}\right| & \alpha < y < \beta \\ 0 & \text{otherwise} \end{cases}$$

RV X with a Continuous Distribution - Method 2 Example

Suppose that X has the p.d.f.

$$f(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Determine the p.d.f. of the random variable $Y = X^2$.

Solution: $y = r(x) = x^2$ is one-to-one for $0 < x < 1$ and has its inverse $s(y) = y^{1/2}, 0 < y < 1$.

$$g(y) = \begin{cases} 3[y^{1/2}]^2|\frac{1}{2}y^{-1/2}| = \frac{3}{2}y^{1/2} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Functions of Two or More Random Variables

Suppose $X_1, X_2, \ldots, X_n$ are RVs whose joint distribution is known. $(Y_1, Y_2, \ldots, Y_m) = r(X_1, X_2, \ldots, X_n)$ where r is a known function. We want to find the distribution of $Y_1, Y_2, \ldots, Y_m$.

- RVs $X_1, X_2, \ldots, X_n$ with a Discrete Joint Distribution
- RVs $X_1, X_2, \ldots, X_n$ with a Continuous Joint Distribution

RVs $X_1, X_2, \ldots, X_n$ with a Discrete Joint Distribution

The process of finding the distribution of $Y_1, Y_2, \ldots, Y_m$ is conceptually easy. We will not illustrate this because it is relatively unimportant. However, under this umbrella there is an important theorem.

Theorem [Bernoulli and Binomial Distributions]: Assume that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) random variables having the Bernoulli distribution with parameter p. Let $Y = X_1 + \ldots + X_n$. Then Y has the binomial distribution with parameters n and p.

RVs $X_1, \dots, X_n$ with a Continuous Joint Distribution

This is a relatively difficult area and hence we only consider the following problems.

- $Y = X_1 + X_2$ with independent X_1, X_2
- $Y = \max(X_1, \dots, X_n)$ with independent $X_1, \dots, X_n$
- $Y = \min(X_1, \dots, X_n)$ with independent $X_1, \dots, X_n$

Finding PDF of $Y = X_1 + X_2$

Theorem: Let X_1 and X_2 be independent continuous random variables and let $Y = X_1 + X_2$. Then the p.d.f. of Y is

$$g(y) = \int_{-\infty}^{\infty} f_1(y - z) f_2(z) dz$$

where f_1 and f_2 are the p.d.f.s of X_1 and X_2, respectively.

Proof.

Let G be the c.d.f. of Y. From the assumption we know that X_1 and X_2 have continuous joint distribution with joint p.d.f. $f_{X_1,X_2}(x_1, x_2) = f_1(x_1)f_2(x_2)$. It then follows that
$G(y) = P(Y \leq y) = P(X_1 + X_2 \leq y) = \int\int_{x_1+x_2\leq y} f_1(x_1)f_2(x_2)dx_1dx_2 = \int\int_{t\leq y} f_1(t - x_2)f_2(x_2)dtdx_2 = \int_{-\infty}^{y} dt \int_{-\infty}^{\infty} f_1(t - x_2)f_2(x_2)dx_2$. By definition of p.d.f., it follows that the p.d.f. g of Y is

$$g(y) = \int_{-\infty}^{\infty} f_1(y - x_2)f_2(x_2)dx_2 = \int_{-\infty}^{\infty} f_1(y - z)f_2(z)dz.$$

□

Finding PDF of $Y = X_1 + X_2$ – Example

Suppose that X_1 and X_2 are independent random variables with common distribution having p.d.f.

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

Find the p.d.f. of $Y = X_1 + X_2$.
Solution: If $y \le 0$, then for all z, either $z \le 0$ or $y - z \le 0$, which means either $f(z) = 0$ or $f(y - z) = 0$, which further implies $f(z)f(y - z) = 0$ for all z, thus $g(y) = \int_{-\infty}^{\infty} 0dz = 0$.
If $y > 0$, $g(y) = \int_{-\infty}^{\infty} f(y - z)f(z)dz$

$$= \int_0^y 2e^{-2(y-z)}2e^{-2z}dz = 4ye^{-2y}.$$

Thus, $g(y) = \begin{cases} 4ye^{-2y} & y > 0 \\ 0 & \text{otherwise} \end{cases}.$

Finding PDF of $Y = \max(X_1, \dots, X_n)$

Theorem: Suppose that $X_1, \dots, X_n$ are mutually independent and have c.d.f.s $F_1(x), \dots, F_n(x)$, respectively. Let

$$Y = \max(X_1, \dots, X_n).$$

Then the c.d.f. and p.d.f of Y are

$$\begin{aligned} G(y) &= F_1(y) \cdots F_n(y), \\ g(y) &= \frac{dG(y)}{dy}. \end{aligned}$$

Moreover, if $F_1 = \dots = F_n = F$ and $f(x) = \frac{dF(x)}{dx}$, then

$$\begin{aligned} G(y) &= F^n(y), \\ g(y) &= nF^{n-1}(y)f(y). \end{aligned}$$

Finding PDF of $Y = \min(X_1, \ldots, X_n)$

Theorem: Suppose that $X_1, \ldots, X_n$ are mutually independent and have c.d.f.s $F_1(x), \ldots, F_n(x)$, respectively. Let

$$Y = \min(X_1, \ldots, X_n).$$

Then the c.d.f. and p.d.f of Y are

$$\begin{aligned} G(y) &= 1 - [1 - F_1(y)] \cdots [1 - F_n(y)], \\ g(y) &= \frac{dG(y)}{dy}. \end{aligned}$$

Moreover, if $F_1 = \ldots = F_n = F$ and $f(x) = \frac{dF(x)}{dx}$, then

$$\begin{aligned} G(y) &= 1 - [1 - F(y)]^n, \\ g(y) &= n[1 - F(y)]^{n-1} f(y). \end{aligned}$$

Homework for Chapter 3

1. Part 1:
   - P100: 2, 6
   - P106: 4, 5, 8
   - P116: 2, 5, 15
   - P129: 3, 4, 5
2. Part 2:
   - P140: 2, 4, 8
   - P151: 2, 4
   - P174: 1, 2, 8
   - P187: 6, 8
   - P203: 16, 18