

CHAPTER FOUR

4. FUNCTIONS OF RANDOM VARIABLES

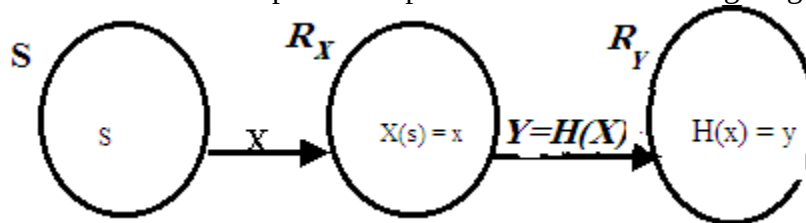
4.1. Equivalent events

4.2. Functions of discrete random variables and their distributions

4.3. Functions of continuous random variables and their distributions

Introduction

It is often the case that we know the probability distribution of the random variable X and we are interested in determining the probability distribution of some function of X . For example, given the distribution of X , we may be interested to determine the distribution of the mean of X . The type of problem considered in this chapter is as follows: Given the random variable X with known distribution and a function H , how do we find the probability distribution of a random variable $Y = H(X)$? The transformation process is presented in the following diagram.



Note that Y is a real valued function of the random variable X with domain R_X and range R_Y . Thus, the function Y itself is also a random variable.

More formally, let E be an experiment, S be a sample space associated with E and X be a random variable defined on S . Suppose that $Y = H(X)$ is a real valued function of X , then $Y = H(X)$ is also a random variable since for every $s \in S$, a value of Y is determined, say $y = H(X(s))$. As before, we call R_X the range space of X , the set of all possible values of the function X . Similarly, we define R_Y the range space of Y , the set of all possible values of the function Y .

Example

1. Let X be a random variable that represents the radius of a circle. Consider another variable Y that represents the area of a circle. Clearly, Y is a function of X , and their functional relationship is presented as $Y = \pi X^2$.
2. In standard statistical methods, the result of statistical hypothesis testing, estimation, or even statistical graphics does not involve a single random variable but, rather, functions of one or more random variables. As a result, statistical inference requires the distributions of these functions. For example, the use of averages of random variables is common. In addition, sums and more general linear combinations are important. We are often interested in the distribution of sums of squares of random variables, particularly in the use of analysis of variance techniques.

4.1. Equivalent events

Definition: Let C be an event associated with the range space of Y , R_Y , i.e. $C \subseteq R_Y$. Moreover, let $B \subseteq R_X$ be defined as $B = \{x \in R_X : H(x) \in C\}$. Then B and C are equivalent events. Meaning, each and every element in C is a function of a corresponding element in B .

Note

- ✓ Mathematically, if for all $H(x) \in C$, we can find event B such that $B = \{x \in R_X : H(x) \in C\}$ then $B \cong C$.
- ✓ When we speak of equivalent events (in the above sense), these events are associated with different sample spaces.

Example

1. Suppose that $H(x) = \pi x^2$. Then the events $B = \{X > 2\}$ and $C = \{Y > 4\pi\}$ are equivalent. Thus, for if $Y = \pi X^2$, then $\{X > 2\}$ occurs if and only if $\{Y > 4\pi\}$ occurs.
2. Consider an experiment of tossing a coin twice. Let X be a random variable indicating a count for heads. Let $B = \{1\}$ be an event with respect to R_X . Then find an event A such that $A \cong B$.

Definition: Suppose X is a random variable defined on the sample space S and R_X be its range space. Let $Y = H(X)$ be a random variable with range space R_Y . Then for any event $C \subseteq R_Y$, $P(C)$ is defined as: $P(C) = P\{x \in R_X : H(x) \in C\}$

In words : The probability of an event associated with the range space of Y is defined as the probability of the equivalent event (in terms of X).

In other words, we can evaluate probabilities involving events associated with the function of the random variable X if the p.d.f of X is known and if it is possible to determine the equivalent event. Note that if B and C are equivalent events, then $P(B) = P(C)$.

Example1: Let X represent the radii of circles. Suppose that $H(x) = 2\pi x$.

$$B_1 = \{x > 2\} \text{ is equivalent to } C_1 = \{y > 4\pi\}$$

$$B_2 = \{x \leq 5\} \text{ is equivalent to } C_2 = \{y \leq 10\pi\}$$

$$\text{Hence, } P(B_1) = P(C_1) \text{ and } P(B_2) = P(C_2)$$

Example 2: Let X be a continuous random variable with p.d.f $f(x) = \begin{cases} e^{-x} & \text{if } x > 0 \\ 0 & \text{Otherwise} \end{cases}$

Suppose that a random variable Y is defined by $Y = \frac{X+1}{2}$. If $C = \{Y: Y \geq 7\}$, find $P(C)$.

Solution:

It can be seen that $R_X = \{X: X > 0\}$ and $R_Y = \{Y: Y > 1/2\}$, which implies that $C \subseteq R_Y$.

$$Y \geq 7 \Leftrightarrow \frac{X+1}{2} \geq 7 \Leftrightarrow X \geq 15$$

Thus, define B as $B = \{X: X \geq 15\}$ so that B and C are equivalent.

$$\text{Hence, } P(C) = P(B) = \int_{15}^{+\infty} e^{-x} dx = \frac{1}{e^{15}}$$

Suppose that the event C is defined as $C = \{Y > 5\}$. Determine the event B such that $B = \{x \in R_X: H(x) \in C\}$ where $Y = H(X) = 2X + 1$.

Exercise: Suppose that the length of time it takes Aaron to drive work everyday is a random variable X (measured in hours) with distribution function,

$$f(x) = \begin{cases} 8(x - \frac{1}{4}) & \text{if } \frac{1}{4} \leq x \leq \frac{3}{4} \\ 0 & \text{Otherwise} \end{cases}$$

If we define a random variable Y as $Y = 60X$, then find $P(Y < 45)$.

Remark: The other procedure for finding probabilities associated with Y is to find p.d.f of Y from the known p.d.f of X.

4.2. Functions of Discrete Random Variables and their Distributions

Case I: X is a discrete random variable:

If X is discrete random variable and $Y = H(X)$, then it follows immediately that Y is also a discrete random variable. That is, if the possible values of X are $x_1, x_2, \dots, x_n$ then the possible values of Y are $H(x_1), H(x_2), \dots, H(x_n)$.

Definition: If $x_1, x_2, \dots, x_n, \dots$ are the possible values of X, $P(x_i) = P(X = x_i)$, and H is a function such that to each value y there corresponds exactly one value x, then the probability distribution of Y is obtained as follows:

Possible values of Y: $y_i = H(x_i), i = 1, 2, \dots, n, \dots$

Probabilities of Y: $q(y_i) = P(Y = y_i) = P(x_i)$

Example 1: Let X assumes three values: -1, 0, and 1 with probabilities 1/3, 1/2, and 1/6 respectively. Then find the probability distributions for:

i) $Y = 3X + 1$

ii) $Y = X^3 + 5$

2. Suppose that X has the following probability distribution:

x:	-2	-1	0	1	2	3	4
P(X=x):	0.1	0.2	0.1	0.1	0.1	0.2	0.2

Find the probability distribution of $Y = 2X - 2$

Note

It may happen that several values of X lead to the same value of Y. In other words, the relation between X and Y may not necessarily be One-to-One, as the following examples illustrate.

Example

1. Recall the first example where X assumes three values: -1, 0, and 1. If $Y = X^2$, then find the possible values and their corresponding probabilities for Y.
2. Consider the second example given above, and suppose that $Y = X^2 + 2$. Then the possible values of Y are 2, 3, 6, 11 and 18 with probabilities 0.1, 0.3, 0.2, 0.2 and 0.2 respectively. For example, $Y = 3$ corresponds to not only one possible value in the range space of X but it refers to two possible values, namely $x = -1$ and $x = 1$. Accordingly $P(Y=3) = P(X=-1) + P(X=1) = 0.2 + 0.1$

Definition

1. Suppose that X is a discrete random variable with probability function $P(X=x)$. Then the probability distribution of $Y = H(X)$ is given by $P(Y=y) = P(H(X) = y) = \sum \{P(X = x_i) \}$ where $H(x_i) = y$
2. Suppose that X is a discrete random variable with probability function $f(x)$. Let $Y = H(X)$ define a one-to-one transformation between the values of X and Y so that the equation $y = H(x)$ can be uniquely solved for x in terms of y, say $x = w(y)$. Then the probability distribution of Y is $g(y) = f[w(y)]$.

Example 1: Suppose the probability function of a discrete random variable X is

x	-1	0	1	2
$P(X = x)$	$\frac{7}{30}$	$\frac{8}{30}$	$\frac{4}{30}$	$\frac{11}{30}$

Find the probability function of Y if

- a) $Y=X^2$ b) $Y=2X+5$

Example 2: Let X be a discrete random variable having the possible values $1, 2, 3, \dots, n, \dots$ and suppose that $P(X = x) = \frac{1}{2^x}$. Find the probability distribution of Y if

$$Y = \begin{cases} 1 & \text{if } X \text{ is even} \\ -1 & \text{if } X \text{ is odd} \end{cases}$$

Case II: If X is a continuous random variable.

X may assume all real values while Y is defined to be +1 if $X > 0$ and -1 if $X < 0$. In order to obtain the probability distribution of Y, simply determine the equivalent event (in the range

space R_x) corresponding to the different values of Y . In the above case, $P(Y = 1) = P(X > 0)$ and $P(Y = -1) = P(X \leq 0)$.

In general case, if $\{Y = y_i\}$ is equivalent to an event, say A , in the range space of X , then

$$g(y_i) = P(Y = y_i) = \int_A f(x)dx$$

4.3. Functions of Continuous Random Variables and their Distributions

In this section, our interest is to determine the Probability Density Function of a function of a continuous random variable. Let X be a continuous random variable and Y be another random variable which is a function of X . That is $Y=H(X)$. Moreover, assume that the random variable Y is a continuous random variable. To find or determine the p.d.f of $Y=H(X)$, we follow the steps below.

- Obtain the Cumulative Distribution Function (cdf) of Y , denoted by $G(y)$, where $G(y) = P(Y \leq y)$ in terms of the cdf or p.d.f of X by finding an event A in R_x which is equivalent to the event $Y \leq y$.
- Differentiate $G(y)$ with respect to y in order to obtain $g(y)$, which is going to be the p.d.f of Y .
- Determine those values of y in the range space of Y for which $g(y) > 0$.

Example 1: Suppose that X has p.d.f given by:

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } 0 < x < 2 \\ 0 & \text{Otherwise} \end{cases}$$

Let $H(X)=2X+4$. Find the p.d.f of $Y=H(X)$

Exercise: Let a random variable X has a density function $f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{Otherwise} \end{cases}$. Let $H(X)=e^x$. Then find the p.d.f of $Y=H(X)$.

Example: suppose that X has the pdf

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

If $Y = 3X + 1$, then find the pdf of Y

Solution:

The other method of obtaining $g(y)$.

$$G(y) = P(Y \leq y) = P\left(X \leq \frac{y-1}{3}\right) = F\left(\frac{y-1}{3}\right),$$

Where F is the cdf of X ; that is

$$F(x) = P(X \leq x).$$

In order to calculate the derivative of G , $G'(y)$, we use the chain rule for differentiation as follows:

$$\begin{aligned} \frac{d}{dy} G(y) &= \\ \frac{d}{dy} G(y) &= \frac{d}{du} G(y) \frac{du}{dy} \text{ ; where } u = \frac{y-1}{3} \end{aligned}$$

Hence

$$G'(y) = F'(u) * \frac{1}{3} = f(u) * \frac{1}{3} = 2 * \left(\frac{y-1}{3}\right) * \frac{1}{3} = \frac{2}{9}(y-1)$$

Theorem: Let X be a continuous random variable with pdf f , where $f(x) > 0$ for $a < x < b$. suppose that $y = H(x)$ is a strictly monotone (increasing or decreasing) function of x . Assume that this function is differentiable (and hence continuous) for all x . Then the random variable Y defined as $Y=H(X)$ has a pdf g given by,

$$g(y) = f(H^{-1}(y)) * \left| \frac{d}{dy} H^{-1}(y) \right| \text{ where } H^{-1}(y) = x$$

Example

- Let X be a random variable with pdf: $f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

Find the pdf of $Y = e^{-x}$

- Suppose that the pdf of a random variable X is given by: $f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

Find the pdf of $Y = \ln(X)$

Theorem: Let X be a continuous random variable with pdf f and $Y = X^2$. Then the random variable Y has the pdf given by: $g(y) = \frac{1}{2\sqrt{y}} [f(\sqrt{y}) + f(-\sqrt{y})]$

Example

Let the pdf of X is given as $f(x) = 1/2$ if $-1 < x < 1$, and 0 else where. If $Y = 4 - X^2$, then find the pdf of Y .