Chapter 4

Continuous Random Variables and Their Probability Distributions

The Probability Distribution for a Continuous Random Variable

Review

Recall :

1. A random variable Y is continuous provided the set of possible values of Y is an interval (possibly infinite)
2. For continuous random variables probability is now interpreted as area.

$P(a \leq Y \leq b)$ = area under the curve associated with the probability function of Y.

The Probability Dist$^{\underline{n}}$ for a Continuous Random Variable

<u>Definition</u>: Let Y denote any random variable. The *cumulative distribution function* (*cdf*) of Y, denoted by F(y), is given by

$F(y) = P(Y \leq y)$ for $-\infty < y < \infty$.

Example

Y	1	2	3	4
P(Y=y)	.4	.3	.2	.1

Find F(y).

Example (con't)

Sol^n :

(i) What is F(y) for y < 0 or y < 1 ?

Since Y takes on value {1,2,3,4} it follows P(Y < y) = 0 for y < 0 (or y < 1).

(ii) P(Y ≤ 2.6) = P(Y ≤ 2) = .4 +.3 = .7

In general

$$F(y) = P(Y \le y) = \begin{cases} 0 & for \quad y < 1 \\ .4 & for \quad 1 \le y < 2 \\ .7 & for \quad 2 \le y < 3 \\ .9 & for \quad 3 \le y < 4 \\ 1 & for \quad 4 \le y \end{cases}$$

Properties of a Distribution Function

Note: Distribution function of discrete random variables are always step functions as F(y) increases only at a finite # of points.

Properties of a Distribution Function :

1. $F(-\infty) = \lim_{y \to -\infty} F(y) = 0.$
2. $F(\infty) = \lim_{y \to \infty} F(y) = 1.$
3. $F(y)$ is a nondecreasing function of y.
   (If y_1 and y_2 are any values such that $y_1 < y_2$, then $F(y_1) \le F(y_2)$.)

Definitions

Definition: Let Y denote a random variable with distribution function $F(y)$. Y is said to be *continuous* if the distribution function $F(y)$ is continuous, for $-\infty < y < \infty$.

→Y must take on values in an interval.

Remark: For continuous random variable Y,

$$P(Y = y) = 0.$$

Consequently,

$$P(\, a < Y < b \,) = P(\, a \leq Y \leq b \,)$$

Definition: Let F(y) be the distribution function for a continuous random variable Y. Then $f(y)$, given by

$$f(y) = \frac{dF(y)}{dy} = F'(y)$$

wherever the derivative exists, is called the *probability density function* (*pdf*) for the random variable Y.

Fundamental Theorem of Calculus

Recall the Fundamental Theorem of Calculus:

$$F(y) = \int_{-\infty}^{y} \frac{dF(t)}{dt} dt$$

$$= \int_{-\infty}^{y} f(t)dt \quad = \quad P(Y \le y)$$

→F(y) area under the curve f(y) between -∞ and y.

- Recalling that F(+∞) = 1 it follows that

$\int_{-\infty}^{\infty} f(t)dt = 1$ - total area under $f(y)$ is equal to 1

Properties of a Density Function

Theorem 4.2: If f(y) is a density function, then

1. $f(y) \geq 0$ for any value of y.

(Because $F(y_1) \leq F(y_2)$ whenever $y_1 < y_2$ it follows that $f(y) \geq 0$, for all y)

2. $\int_{-\infty}^{\infty} f(y)dy = 1.$

Example 1: Let Y be a continuous random variable with probability density function given by

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

Find F(y).

Theorem 4.3: If the random variable Y has density function $f(y)$ and $a \leq b$, then the probability that Y falls in the interval $[a,b]$ is

$$P(a \leq Y \leq b) = \int_a^b f(y)dy \quad .$$

Idea :

$$\begin{aligned} P(a < Y < b) &= P(a \leq Y < b) = P(a < Y \leq b) \\ &= P(a \leq Y \leq b) = P(Y \leq b) - P(Y \leq a) \\ &= F(b) - F(a) = \int_a^b f(y)dy \end{aligned}$$

Examples

Example1 : Given $f(y) = cy^2$, $0 \leq y \leq 2$, and $f(y) = 0$ elsewhere, find the value of c for which $f(y)$ is a valid density function.

Solution :

Example2 : Find $P(1 \leq Y \leq 2)$ for Example 1. Also find $P(1 < Y < 2)$.

Solution:

Example3 : Given $f(y) = ky(2-y)$, $0 < y < 2$, and $f(y) = 0$ elsewhere.

1. Find the value of k for which $f(y)$ is a valid density function.
2. Find $P(\frac{1}{2} \leq Y \leq 1)$.
3. Find the cumulative density function, F(Y).
4. Use F(Y) to find $P(Y \geq 1 \mid Y \leq 2)$.

Sol^{n} :

1. Find the value of k for which $f(y)$ is a valid density function.

<u>Sol^n</u> :

2. $P(\frac{1}{2} \leq Y \leq 1) = ?$

Sol^n :

3. Find the cumulative density function, F(Y).

Sol^n :

4. $P(Y \geq 1 \mid Y \leq 2) =$

Expected Values for Continuous Random Variables

Expected Value

- As with discrete random variables we are also interested in the means, variances, and standard deviations of continuous random variables.

  Definition: The expected value of a continuous random variable Y is

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy$$

  provided that the integral exists.

  Analogy with the discrete case

$$E(Y) = \sum_{y} yp(y)$$

Expected Value

Theorem 4.4: Let g(Y) be a function of Y; then the expected value of g(Y) is given by

$$E(g(Y)) = \int_{-\infty}^{\infty} g(y) f(y) dy$$

Theorem 4.5: Let c be a constant, and let $g(Y), g_1(Y), \ldots, g_k(Y)$ be functions of a continuous random variable Y. Then the following results hold:

1. $E(c) = c$
2. $E[cg(Y)] = cE[g(Y)]$
3. $E[g_1(Y) + \ldots + g_k(Y) = E[g_1(Y)] + \ldots + E[g_k(Y)]$

Examples

<u>Example1</u> : Suppose Y is a continuous random variable with pdf

$$f(y) = \begin{cases} \dfrac{4}{\pi(1+y^2)} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find E(Y).

<u>Solution</u>:

Example2 : Suppose Y is a continuous random variable with pdf

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

Let $g(y)=e^{(3y/4)}$. Find E(g(Y))

Solution:

Variance

- Also as in the discrete case we define

$$V(Y) = E\left(\left(Y-\mu\right)^2\right) \quad , \quad \mu = E(Y)$$

$$= E(Y^2) - \left(E(Y)\right)^2$$

And

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f(y)dy \qquad \left(g(Y) = Y^2\right)$$

Examples

Example3: Suppose Y is a continuous random variable with pdf

$$f(y) = \begin{cases} \dfrac{4}{\pi(1+y^2)} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find V(Y).

Solution:

Example 4 : Daily total solar radiation for a specified location in Florida in October has probability density function given by

$$f(y)=\begin{cases}(3/32)(y-2)(6-y) & , \quad 2\le y\le 6\\ 0 & , \quad \text{otherwise}\end{cases}$$

with measurements in hundreds of calories. Find the expected daily solar radiation for October.

Solution:

- Let Y = # of daily solar radiation in October
- Want to find E(Y).

Example 5: Weekly CPU time used by an accounting firm has probability density function (measured in hours) given by

$$f(y) = \begin{cases} (3/64)y^2(4-y) & , \quad 0 \le y \le 4 \\ 0 & , \quad \text{otherwise} \end{cases}$$

(a) Find the expected value and variance of weekly CPU time.

(b) The CPU time costs the firm \$200 per hour. Find the expected value and variance of the weekly cost for CPU time.

(c) Would you expect the weekly cost to exceed \$600 very often?

(a) Find the expected value and variance of weekly CPU time.

Solution :

Let Y = weekly CPU time

(b) The CPU time costs the firm \$200 per hour. Find the expected value and variance of the weekly cost for CPU time.

Solution:

Let Y = weekly CPU time

C = weekly cost CPU time = $200Y$

(c) Would you expect the weekly cost to exceed \$600 very often?

$P(\, C > 600) = ?$

Solution:

Example 6: Given

$$F(y) = \begin{cases} 0, & y \le 0 \\ y/8, & 0 < y < 2 \\ y^2/16, & 2 \le y < 4 \\ 1, & y \ge 4 \end{cases}$$

Find the mean and variance of Y.

Solution:

Differentiate F(y) with respect to y, we have

$$f(y) = \begin{cases} 0, & y \le 0 \\ 1/8, & 0 < y < 2 \\ y/8, & 2 \le y < 4 \\ 0, & y \ge 4 \end{cases}$$

Solution (Con't)

Example 7 : Refer to Example 6.

(a) Find $P(1 \leq Y \leq 3)$.

(b) Find $P(Y \geq 1.5)$.

(c) Find $P(Y \geq 1 \mid Y \leq 3)$.

Solution:

(a) $P(1 \leq Y \leq 3)$

(b) $P(Y \geq 1.5)$

(c) $P(Y \geq 1 \mid Y \leq 3)$

The Uniform Probability Distribution

Uniform Probability Distribution

- Most basic pdf is the uniform distribution in which the pdf *f*(*y*) is constant.

  Definition: If $\theta_1 < \theta_2$, a random variable Y is said to have a continuous uniform probability distribution on the interval (θ_1, θ_2) if and only if the density function of Y is

$$f(y) = \begin{cases} \dfrac{1}{\theta_2 - \theta_1} & , \quad \theta_1 \le y \le \theta_2 \\ 0 & , \quad \text{elsewhere} \end{cases}$$

- θ_1 and θ_2 – the parameters of the uniform density

Examples

Example1 : Suppose $Y \sim$ uniform (θ_1, θ_2).

1. Find F(y).
2. Find E(Y), V(Y).

Solution:

(1) $F(y) = \int_{-\infty}^{y} f(t)dt$

(2) Find E(Y), V(Y)

Solution:

$$E(Y) = \int_{-\infty}^{\infty} y f(y) dy$$

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f(y) dy$$

Solution (Con't):

$$V(Y) = E(Y^2) - \left(E(Y)\right)^2$$

Example 2 :

Beginning at 12:00 midnight, a computer center is up for 1 hour and then down for 2 hours on a regular cycle. A person who is unaware of this schedule dials the center at a random time between 12:00 midnight and 5:00 A.M. What is the probability that the center is up when the person's call comes in?

Solution:

Let Y = the time that person's call comes in

Y ~ uniform (0,5)

midnight

$$f(y) = \frac{1}{\theta_2 - \theta_1} \quad , \quad \theta_1 \le y \le \theta_2$$

Recall

$$P(a \le Y \le b) = \int_a^b f(y)dy$$

Example 3 :

The cycle time for trucks hauling concrete to a highway construction site is uniformly distributed over the interval 50 to 70 minutes. What is the probability that the cycle time exceeds 65 minutes if it is known that the cycle time exceeds 55 minutes?

Solution:

Let Y = cycle time $Y \sim$ uniform $(50,70)$

Want $P(Y > 65 \mid Y > 55)$.

Alternative Solution:

Recall:

$$F(y) = P(Y \le y) \qquad f(y) = \frac{1}{70-50} = \frac{1}{20} \quad , \quad 50 \le y \le 70$$

$$F(y) = \int_{50}^{y} \frac{1}{70-50}\,dt = \int_{50}^{y} \frac{1}{20}\,dt = \frac{1}{20}\,t\Big|_{50}^{y} = \frac{y-50}{20} \qquad \text{for} \quad 50 \le y \le 70$$

$$F(y) = \begin{cases} 0 & , \quad y < 50 \\ \dfrac{y-50}{20} & , \quad 50 \le y \le 70 \\ 1 & , \quad y > 70 \end{cases}$$

F(65)

$$P(Y > 65 \mid Y > 55) = \frac{P(Y > 65)}{P(Y > 55)} = \frac{1 - P(Y \le 65)}{1 - P(Y \le 55)} = \frac{1 - \left(\dfrac{65-50}{20}\right)}{1 - \left(\dfrac{55-50}{20}\right)} = \frac{20-15}{20-5} = \frac{1}{3}$$

F(55)

Example 4: Refer to Example 3.

Find the mean and variance of the cycle times for the trucks.

Solution :

Let Y = cycle time

$$E(Y) = \frac{\theta_1 + \theta_2}{2}$$

$$V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}$$

The Normal Probability Distribution

Normal Distribution

Definition : A random variable Y is said to have a normal probability distribution if and only if , for $\sigma > 0$ and $-\infty < \mu < \infty$, the density function of Y is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/\left(2\sigma^2\right)} \quad , \quad -\infty < y < \infty$$

Normal Distribution

<u>Remarks</u> :

1. $P(a \le Y \le b) = \int_a^b f(y)dy \quad = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/\left(2\sigma^2\right)} dy$

   A closed-form expression for this integral does not exist; hence its evaluation requires the use of numerical integration techniques.

2. Always transform a normal random variable Y with parameters μ and σ to a standard normal random variable Z (μ=0 and σ=1) by using the relationship

$$Z = \frac{Y-\mu}{\sigma}$$

Normal Distribution

<u>Theorem 4.7</u> : If Y is a normally distributed random variable with parameters μ and σ, then

$$E(Y) = \mu \qquad \text{and} \qquad V(Y) = \sigma^2$$

Examples

Example 1 :

The grade point averages of a large population of college students are approximately normally distributed with mean 2.4 and standard deviation .8. What fraction of the students will possess a grade point average in excess of 3.0?

Solution:

Let Y = grade point average

$P(Y > 3.0) = ?$

Example 2 : Refer to Example 1.

Suppose that three students are randomly selected from the student body. What is the probability that all three will possess a grade point average in excess of 3.0?

Solution :

Let Y = # student possess a grade point average in excess of 3.0

$P(Y = 3) = ?$

Example 3:

Scores on an examination are assumed to be normally distributed with mean 78 and variance 36.

(a) What is the probability that a person taking the examination scores higher than 72?

Solution :

Let Y = scores on an examination

(b) Suppose that students scoring in the top 10% of this distribution are to receive an A grade. What is the minimum score a student must achieve to earn an A grade?

Solution:

We want to find y such that $P(Y > y) = .1$

(c) What must be the cutoff point for passing the examination if the examiner wants only the top 28.1% of all scores to be passing?

Solution:

We want to find y such that $P(Y > y) = .281$

(d) Approximately what proportion of students have scores 5 or more points above the score that cuts off the lowest 25%?

Solution:

Find y such that $P(Y < y) = .25$

(e) If it is known that a student's score exceeds 72, what is the probability that his or her score exceeds 84?

Solution:

$P(Y > 84 \mid Y > 72) = ?$