CHAPTER 4
Continuous Variables and Their Probability Distributions

Motivation

- The Number of days must take one of the $n+1$ values $0, 1, 2, \dots, n$
- The daily rainfall at a specified geographical point
  - Each of the unaccountably infinite number of points in the interval $(0, 5)$ represents a distinct possible value of the amount of rainfall in a day
  - A random variable that can take on any value in an interval is called continuous
- The purpose of this chapter is to study probability distributions for continuous random variables

Motivation

- The probability distribution for a discrete random variable can always be given by assigning a nonnegative probability to each of the possible values the variable may assume
  - The sum of all the probabilities that we assign must be equal to 1
- Unfortunately, the probability distribution for a continuous random variable cannot be specified in the same way
  - It is mathematically impossible to assign nonzero probabilities to all the points on a line interval while satisfying the requirement that the probabilities of the distinct possible values sum to 1.
- Therefore, we must develop a different method to describe the probability distribution for a continuous random variable

Distribution Function

DEFINITION 4.1

Let Y denote any random variable. The *distribution function* of Y, denoted by $F(y)$, is such that $F(y) = P(Y \leq y)$ for $-\infty < y < \infty$.

- The nature of the distribution function associated with a random variable determines whether the variable is continuous or discrete

Distribution Function for Discrete Random Variable: Example

EXAMPLE 4.1

Suppose that Y has a binomial distribution with $n = 2$ and $p = \frac{1}{2}$. Find $F(y)$.

Distribution Function for Discrete Random Variable: Example

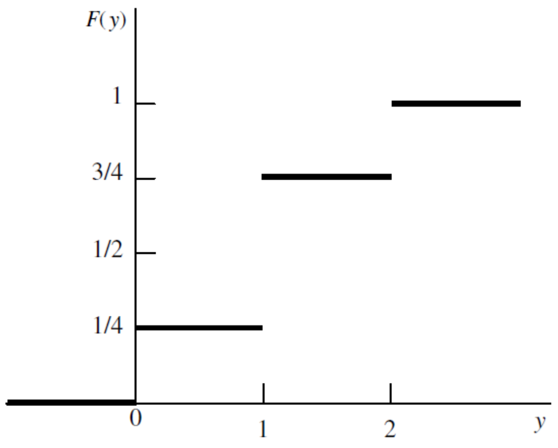
SOLUTION 4.1

The probability (mass) function for Y is given by

$$p(y) = \binom{2}{y}\left(\frac{1}{2}\right)^{y}\left(\frac{1}{2}\right)^{2-y}, \qquad y = 0,1,2.$$

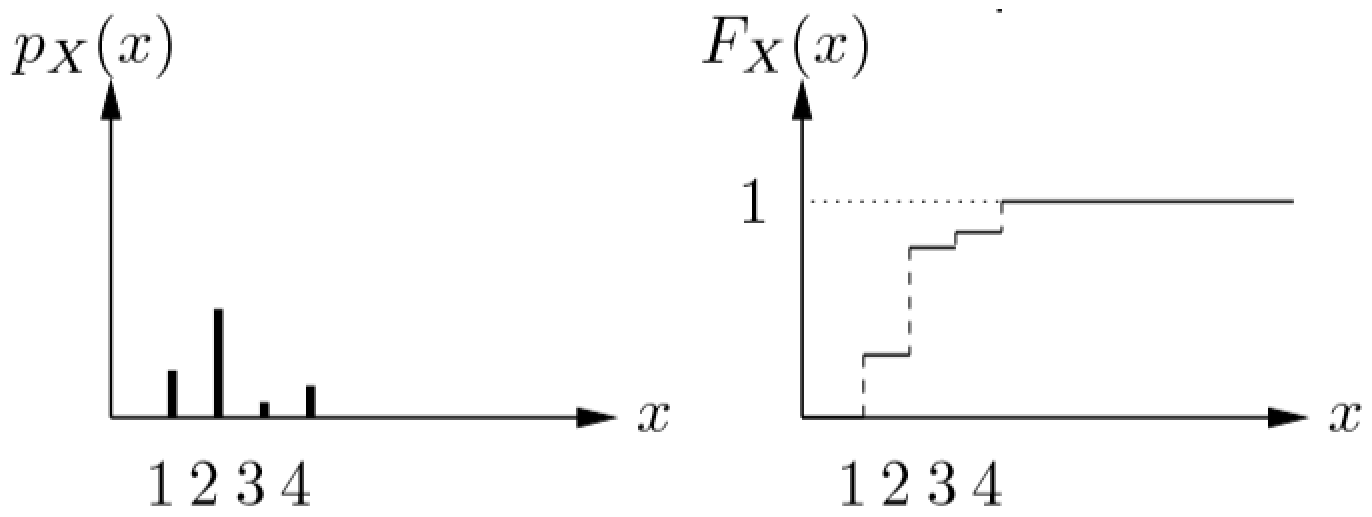
Since $p(0) = \frac{1}{4}, p(1) = \frac{1}{2}, p(2) = \frac{1}{4}$, the distribution function of Y is given by

$$F(y) = P(Y \le y) = \begin{cases} 0 & y < 0 \\ 1/4 & 0 \le y < 1 \\ 3/4 & 1 \le y < 2 \\ 1 & y \ge 2 \end{cases}$$



Distribution Function for Discrete Random Variable

- Distribution functions for discrete random variables are always **step functions** because the cumulative distribution function incases only at the finite or countable number of points with positive probability.



Properties of a Distribution Function

THEOREM 4.1 (Properties of a Distribution Function)

If $F(y)$ is a distribution function, then

1. $F(-\infty) \equiv \lim_{y \to -\infty} F(y) = 0$.
2. $F(\infty) \equiv \lim_{y \to \infty} F(y) = 1$.
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 $F(y_1) \leq F(y_2)$.]

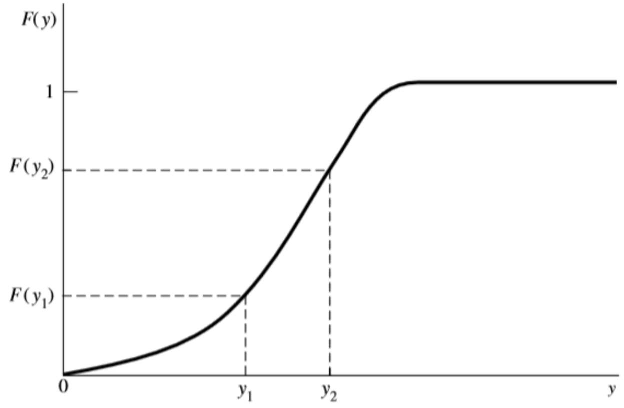
- Check that the distribution function discussed in EXAMPLE 4.1 has each of these properties.

Continuous Random Variable

DEFINITION 4.2

A random variable Y with distribution function $F(y)$ is said to be *continuous* if $F(y)$ is continuous, for $-\infty < y < \infty$.

- Suppose that, for all practical purposes, the amount of daily rainfall, Y, must be less that 6 inches .
- For every $0 \le y_1 < y_2 \le 6$, the interval (y_1, y_2) has a positive probability of including Y, no matter how close y_1 gets to y_2.
  - It follows that $F(Y)$ in this case should be a smooth, increasing function over some interval of real numbers

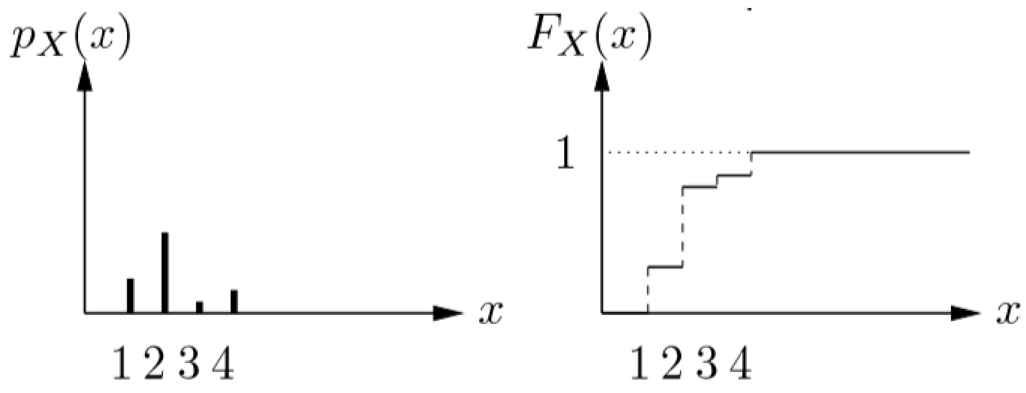


Continuous Random Variable

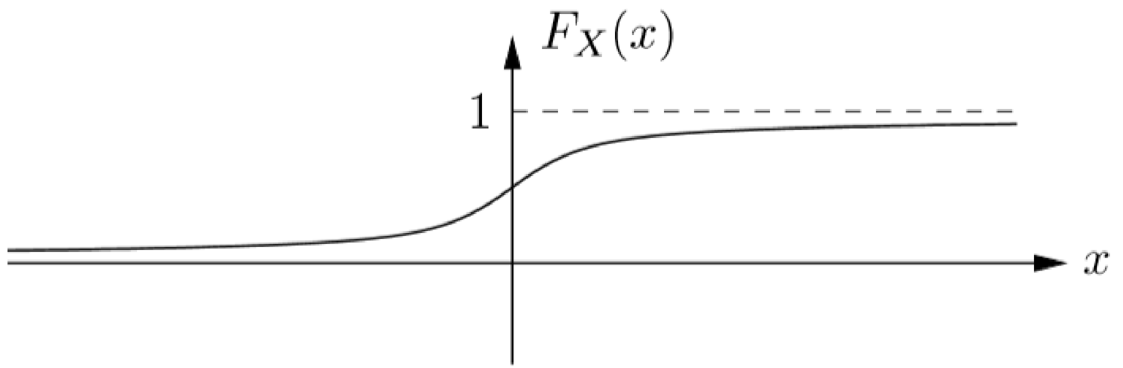
- If Y is a continuous random variable, then for any real number y,

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

If this is not true and $P(Y = y) = p_0 > 0$, then $F(Y)$ would have a discontinuity (jump) of a size p_0 at the point y_0, violating the assumption that Y was continuous.



$P(Y = y) = p_0 > 0$



$P(Y = y) = 0$

probability density function

DEFINITION 4.3

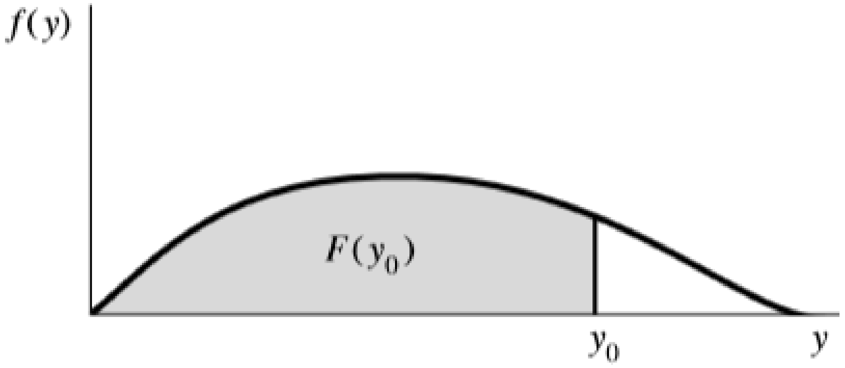
Let $F(y)$ be the distribution function for a continuous random variable Y. Then $f(y)$, given by

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

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

- If follows from Definitions 4.2 and 4.3 that $F(y_0)$ can be written as

$$F(y_0) = \int_{-\infty}^{y_0} f(t)\,dt$$



- The probability density function is a theoretical model for the frequency distribution (histogram) of a population of measurements.
  - For example, observations of the lengths of life of washers of a particular brand

probability density function

THEOREM 4.1 (Properties of a Distribution Function)

If $F(y)$ is a distribution function, then

1. $F(-\infty) \equiv \lim_{y \to -\infty} F(y) = 0.$
2. $F(\infty) \equiv \lim_{y \to \infty} F(y) = 1.$
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 $F(y_1) \leq F(y_2)$.]

continuous random variable Y

$$f(y) = \frac{dF(y)}{dy} = F'(y) \quad or \quad F(y) = \int_{-\infty}^{y} f(t)\,dt$$

THEOREM 4.2 (Properties of a Density Function)

If $f(y)$ is a density function for a continuous random variable, then

1. $f(y) \geq 0$ for all $y, -\infty < y < \infty$.
2. $\int_{-\infty}^{\infty} f(y)dy = 1.$

Distribution Function for Continuous Random Variable: Example

EXAMPLE 4.2

Suppose that

$$F(y) = \begin{cases} 0 & y < 0 \\ y & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

Find the probability density function for Y and graph it.

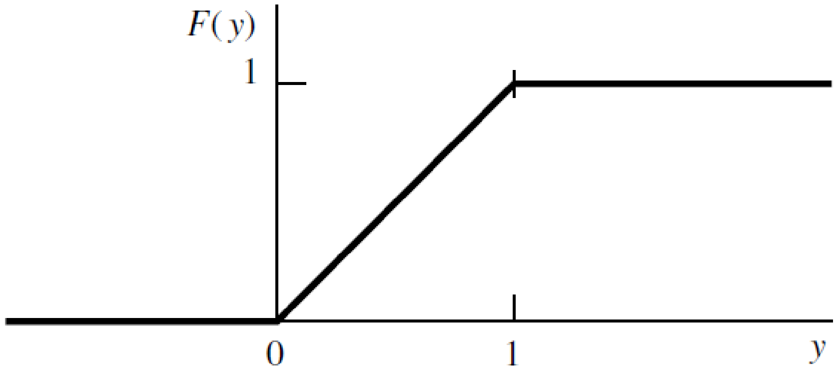
Distribution Function for Continuous Random Variable: Example

SOLUTION 4.2

Because the density function $f(y)$ is the derivative of the distribution function $F(y)$, when the derivative exists,

$$f(y) = \frac{dF(y)}{dy} = \begin{cases} \dfrac{d(0)}{dy} = 0 & y < 0 \\ \dfrac{d(y)}{dy} = 1 & 0 < y < 1 \\ \dfrac{d(1)}{dy} = 0 & y > 1 \end{cases}$$

Note that $f(y)$ is undefined at $y = 0$ and $y = 1$.



Distribution Function for Continuous Random Variable: Example

EXAMPLE 4.3

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

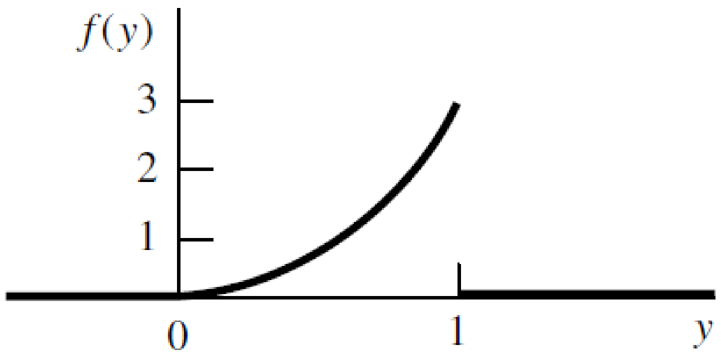
$$f(y) = \begin{cases} 3y^2 & 0 \leq y \leq 1 \\ 0 & elsewhere \end{cases}$$

Find $F(y)$. Graph both $f(y)$ and $F(y)$.

Distribution Function for Continuous Random Variable: Example

SOLUTION 4.3

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

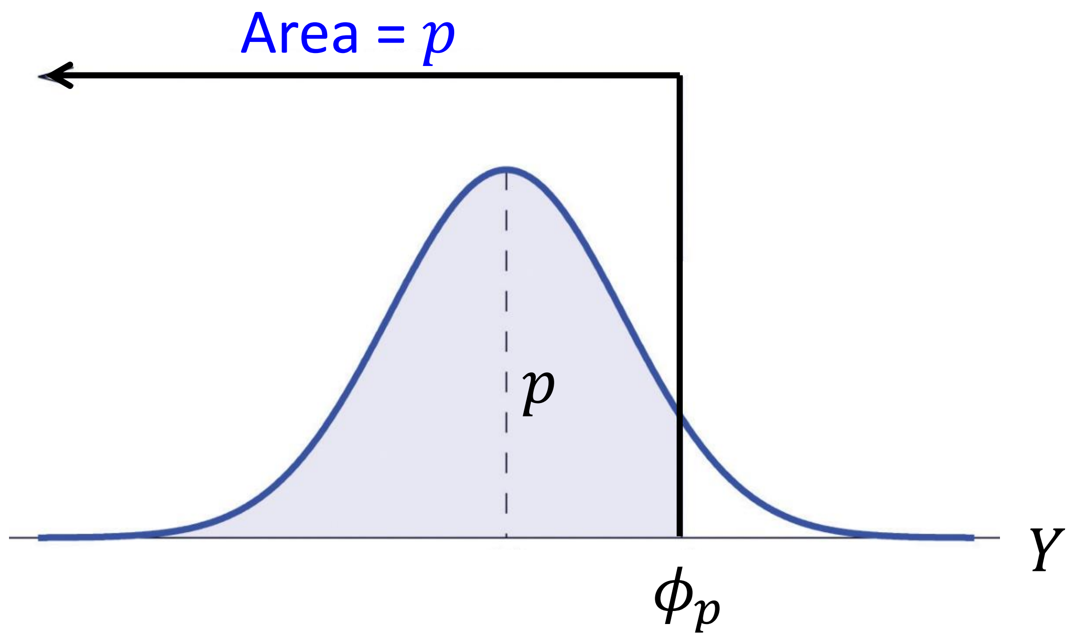


$$F(y) = \begin{cases} \int_{-\infty}^{y} 0dt = 0 & y < 0 \\ \int_{-\infty}^{0} 0dt + \int_{0}^{y} 3t^2dt = 0 + t^3]_0^y = y^3 & 0 \leq y \leq 1 \\ \int_{-\infty}^{0} 0dt + \int_{0}^{1} 3t^2dt + \int_{1}^{y} 0dt = 0 + t^3]_0^1 + 0 = 1 & 1 < y \end{cases}$$

Percentile

DEFINITION 4.4

Let Y denote any random variable. If $0 < p < 1$, the pth *quantile* of Y, denoted by ϕ_p, is the smallest value such that $P(Y \le \phi_p) = F(\phi_p) \ge p$. If Y is continuous, ϕ_p is the smallest value such that $F(\phi_p) = P(Y \le \phi_p) = p$. Some prefer to call ϕ_p the 100pth *percentile* of Y.



Continuous probability distribution

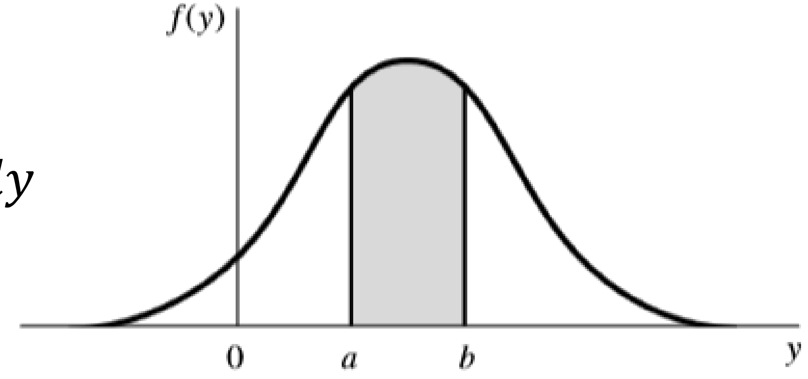
THEOREM 4.3

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

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

- From Chapter 1 we know that this probability corresponds to the area under the frequency distribution over the interval $a < y < b$.
- Because $f(y)$ is the theoretical counterpart of the frequency distribution, we would expect $P(a \le Y \le b)$ to be equal a corresponding area under the density function $f(y)$

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



Continuous probability distribution : Example

EXAMPLE 4.4

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.

Continuous probability distribution : Example

SOLUTION 4.4

$$F(\infty) = \int_{-\infty}^{\infty} f(y)dy = \int_{0}^{2} cy^2 dy = \frac{cy^3}{3}\bigg]_0^2 = \left(\frac{8}{3}\right)c = 1 \rightarrow \therefore c = \frac{3}{8}.$$

Expected Values of R.V.

DEFINITION 4.5

The expected value of a continuous random variable Y is

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

provided that the integral exists.

- The Expected value of a discrete random variable Y is

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

✓ The quantity $f(y)dy$ corresponds to $p(y)$ for the discrete case

✓ Integration is analogous to summation

Expected Values of a Function

THEOREM 4.4

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

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

provided that the integral exists.

- **Proof:**

  Similar to Theorem 3.2

Properties of Expectation

THEOREM 4.5

Let c be a constant and let $g(Y). g_1(Y), g_2(Y), \dots, g_k(Y)$ 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. $E[g_1(Y) + g_2(Y) + g_k(Y)] = E[g_1(Y)] + E[g_2(Y)] + \cdots + E[g_k(Y)]$.

- **Proof:**

  Similar to Theorem 3.3, 3.4, and 3.5

- Show that $V(Y) = E(Y^2) - \mu^2$

Expected Values of R.V. : Examples

EXAMPLE 4.6

In Example 4.4 we determined that $f(y) = (3/8)y^2$ for $0 \leq y \leq 2$, $f(y) = 0$ elsewhere, is a valid density function. If the random variable Y has this density function, find $\mu = E(Y)$ and $\sigma^2 = V(Y)$.

Expected Values of R.V. : Examples

SOLUTION 4.6

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy = \int_0^2 y\left(\frac{3}{8}\right)y^2dy = \left(\frac{3}{8}\right)\left(\frac{1}{4}\right)y^4\bigg]_0^2 = 1.5.$$

$$E(Y^2) = \int_{-\infty}^{\infty} y^2f(y)dy = \int_0^2 y^2\left(\frac{3}{8}\right)y^2dy = \left(\frac{3}{8}\right)\left(\frac{1}{5}\right)y^5\bigg]_0^2 = 2.4.$$

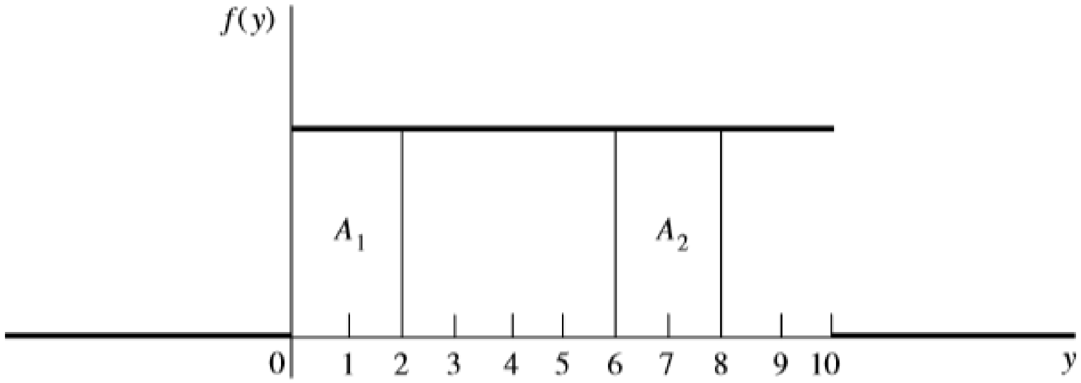
Thus, $\sigma^2 = V(Y) = E(Y^2) - [E(Y)]^2 = 2.4 - (1.5)^2 = 0.15.$

Motivation

- **Assume:**
  - ✓ A bus always arrives at a particular stop between 8:00 and 8:10 A.M.
  - ✓ The probability that the bus will arrive in any given subinterval of time is proportional only to the length of the subinterval
    - ➢ The bus is as likely to arrive between 8:00 and 8:02 as it is to arrive between 8:06 and 8:08.
- Let Y denote the length of time a person must wait for the bus if that person arrived at the bust stop at exactly 8:00
  - The relative frequency with which we observed a value of Y between 0 and 2 would be approximately the same as the relative frequency with which we observed a value of Y between 6 and 8.
  - Because the area under the curves represent probabilities for continuous random variables

$$A1 = A2,$$

$$P(0 \leq Y \leq 2) = P(6 \leq Y \leq 8)$$



The Uniform Probability Distribution

DEFINITION 4.6

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

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

Parameters

DEFINITION 4.7

The constants that determine the specific form of a density function are called *parameters* of the density function.

- The quantities θ_1 and θ_2 are parameters of the uniform density function
  - ✓ Meaningful numerical values associated with the theoretical density function
- Some continuous random variables in the physical, management, and biological sciences have approximately uniform probability distributions
  - ✓ If the number of events that occur in the time interval $(0, t)$ has a Poisson distribution,
    - ➢ Provided that exactly one such event has occurred in the interval $(0, t)$, the actual time of occurrence is distributed uniformly over this interval

The Uniform Probability Distribution : Examples

EXAMPLE 4.7

Arrivals of customers at a checkout counter follow a Poisson distribution. It is known that, during a given 30-minute period, one customer arrived at the counter. Find the probability that the customer arrived during the last 5 minutes of the 30-minute period.

The Uniform Probability Distribution : Examples

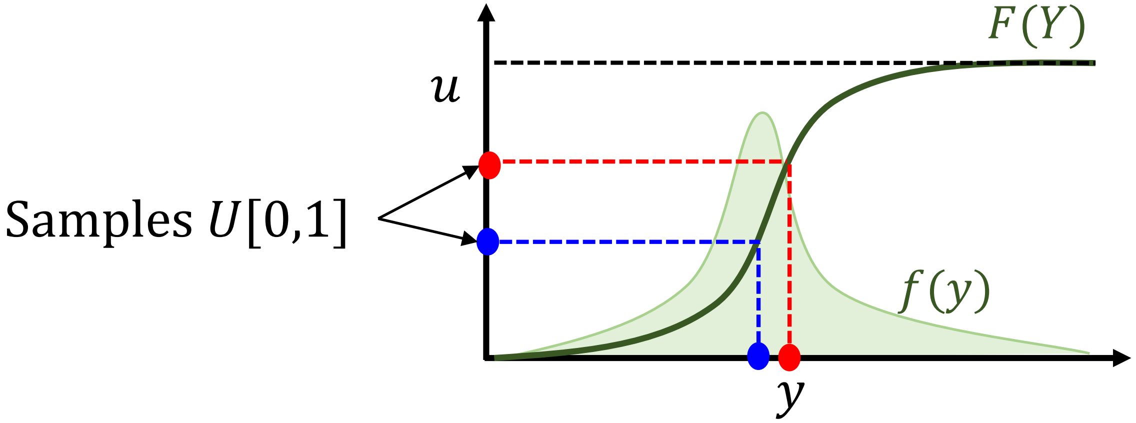
SOLUTION 4.7

The actual time of arrival follows a uniform distribution over the interval of $(0, 30)$. If Y denotes the arrival time, then

$$P(25 \leq Y \leq 30) = \int_{25}^{30} \frac{1}{30} dy = \frac{30-25}{30} = \frac{5}{30} = \frac{1}{6}.$$

Importance of Uniform Distribution

- The uniform distribution is very important for theoretical reasons
- Simulation studies are valuable techniques for validating models in statistics
- If we desire a set of observations on a random variable Y with distribution function $F(y)$,
  - ✓ We can obtain the desired results by transforming a set of observations on a uniform random variable
  - ✓ For this reason most computer systems contain a random number generator that generates observed values for a uniform random variable.



Inverse Transform Sampling

The Mean and Variance of a Uniform Distribution

THEOREM 4.6

If $\theta_1 < \theta_2$ and Y is a random variable uniformly distributed on the interval (θ_1, θ_2), then

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \text{ and } \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}.$$

Proof:

The Mean and Variance of a Uniform Distribution

THEOREM 4.6

If $\theta_1 < \theta_2$ and Y is a random variable uniformly distributed on the interval (θ_1, θ_2), then

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \text{ and } \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}.$$

Proof:

By Definition 4.5,

$$E(Y) = \int_{-\infty}^{\infty} y f(y) dy = \int_{\theta_1}^{\theta_2} y \left(\frac{1}{\theta_2 - \theta_1}\right) dy = \left(\frac{1}{\theta_2 - \theta_1}\right) \frac{y^2}{2}\bigg]_{\theta_1}^{\theta_2} = \frac{\theta_2^2 - \theta_1^2}{2(\theta_2 - \theta_1)} = \frac{\theta_2 + \theta_1}{2}.$$

Motivations

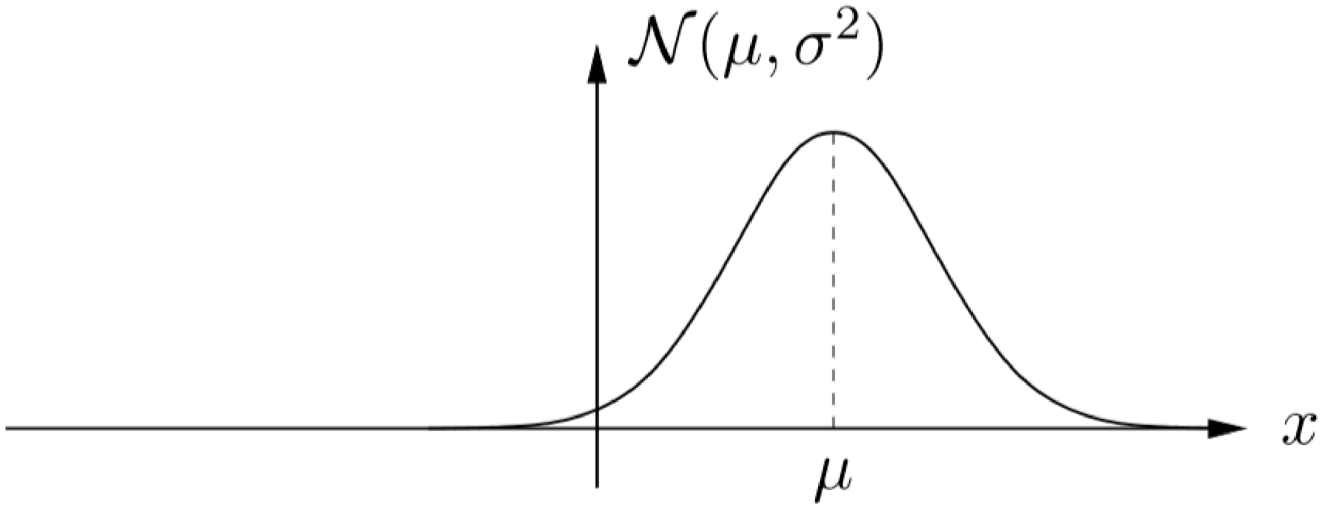
- The most widely used continuous probability distribution is the normal distribution
- The normal distribution looks like the bell shape
- This Chapter discusses the examples and exercises that illustrate some of the many random variables that have distributions that are closely approximated by a normal probability distribution
- Chapter 7 will present an argument that at least partially explains the common occurrence of normal distributions of data in nature

Normal Probability Distribution

DEFINITION 4.8

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

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



The Mean and Variance of Normal Probability Distribution

THEOREM 4.7

If Y is a normally distributed random variable with parameters μ and σ, then

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

Proof: Given in Section 4.9 where we derive the moment-generating function of a normally distributed random variable

- The parameter μ locates the center of the distribution
- The parameter σ measures its spread
- Area under the normal density function corresponding to $P(a \leq Y \leq b)$ require evaluation of the integral

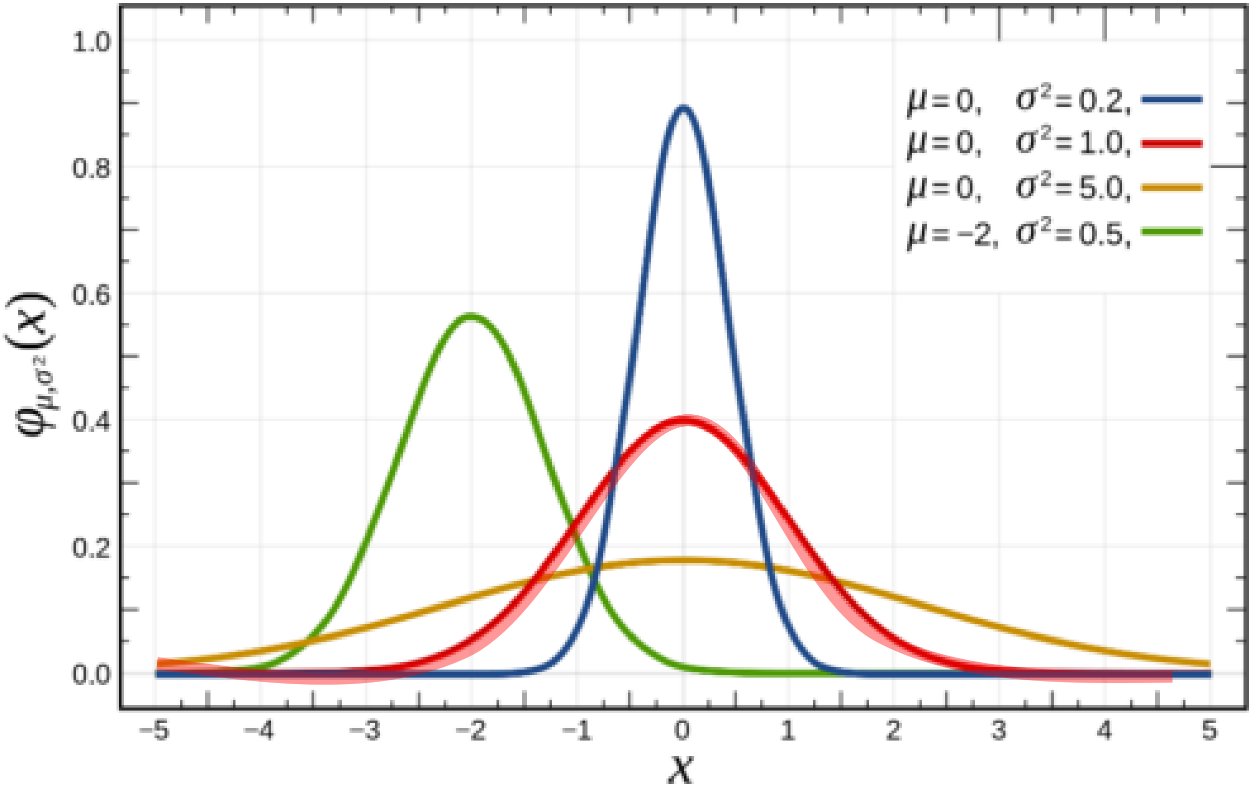
$$P(a \leq Y \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}\, dy$$

✓ The closed-form expression for this integral does not exist;

✓ Requires the use of numerical integration techniques

Standard Normal Random Variable

$$Y \sim N(\mu, \sigma^2) \xrightarrow{Z = \frac{Y - \mu}{\sigma}} Z \sim N(0, 1^2)$$



Normal Probability Distribution : Example

EXAMPLE 4.8

Let Z denote a normal random variable with mean 0 and standard deviation 1.

a. Find $P(Z > 2)$.
b. Find $P(-2 \leq Z \leq 2)$.
c. Find $P(0 \leq Z \leq 1.73)$.

Normal Probability Distribution : Example

SOLUTION 4.8

a. Refer to the Table 4, Appendix 3 of the textbook. The area opposite $z = 2.0$ is .0228, thus $P(Z > 2) = .0228$.

b. Because the density function is symmetric about the mean $\mu = 0$,

$$P(-2 \leq Z \leq 2) = 1 - 2(.0228) = .9544.$$

c. Because $P(Z > 0) = .5$, and $P(Z > 1.73) = .0418$ by the Table 4 in Appendix 3 of the textbook,

$$P(0 \leq Z \leq 1.73) = .5 - .0418 = .4582.$$

Normal Probability Distribution : Example

EXAMPLE 4.9

The achievement scores for a college entrance examination are normally distributed with mean 75 and standard deviation 10. What fraction of the scores lies between 80 and 90?

Normal Probability Distribution : Example

SOLUTION 4.9

Recall that z is the distance from the mean of a normal distribution expressed in units of standard deviation. Thus,

$$z = \frac{y - \mu}{\sigma}.$$

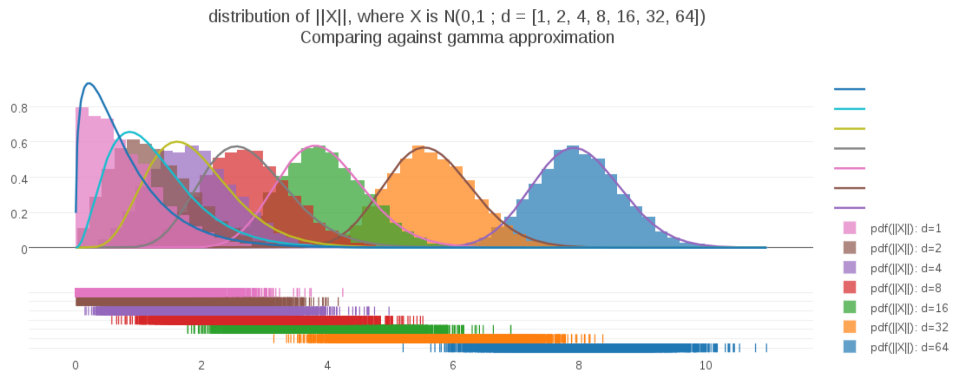
Then the desired fraction of the population is given by the area between

$$z_1 = \frac{80 - 75}{10} = .5 \ \ and \ z_2 = \frac{90 - 75}{10} = 1.5.$$

Therefore, the fraction is $.3085 - .0668 = .2417$.

Motivation

- Some random variables are always **nonnegative** and for various reasons yield distributions of data that are **skewed** (nonsymmetric) to the right.
  - ✓ The lengths of time between malfunctions for aircraft engines possess a skewed frequency distribution
  - ✓ The lengths of time between arrivals at a supermarket checkout queue
  - ✓ The lengths of time to complete a maintenance checkup for an automobile engine
- Most of the area under the density function is located near the origin, and the density function drops gradually as y increases
- The populations associated with these random variables frequently possess density functions that are adequately modeled by a gamma density function



The Gamma Probability Distribution

DEFINITION 4.9

A random variable Y is said to have a *gamma probability distribution with parameters* $\alpha > 0$ and $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & 0 \leq y < \infty \\ 0, & \text{elsewhere,} \end{cases}$$

where

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\,dy.$$

- The quantity $\Gamma(\alpha)$ is known as the *gamma function*
- If α is not an integer and $0 < c < d < \infty$, it is impossible to give a closed-form expression for

$$P(c \leq Y \leq d) = \int_c^d \frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}$$

The Properties of a Gamma Function

- The quantity $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$ is known as the *gamma function*

  ✓ $\Gamma(1) = 0$

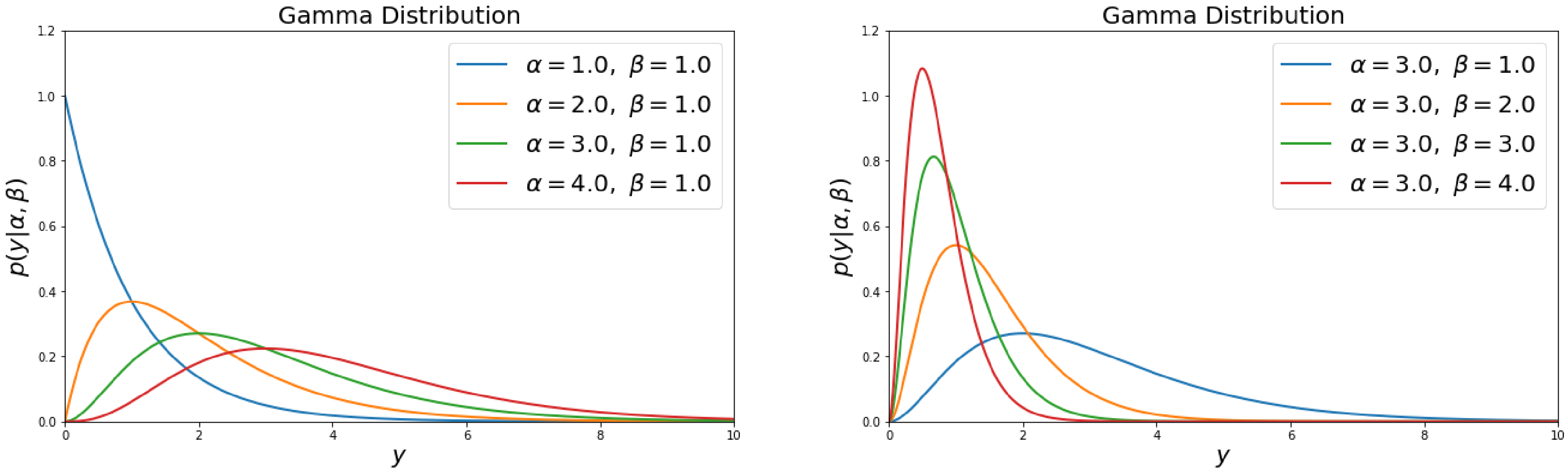
  ✓ $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for any $\alpha > 1$

  ✓ $\Gamma(n) = (n-1)!$ Provided that n is an integer

  ✓ $\int_0^\infty x^{\alpha-1} e^{-x/\beta} dx = \beta^\alpha \Gamma(\alpha)$ for $\beta > 0$

$$P(0 \le Y \le \infty) = \int_0^\infty \frac{y^{\alpha-1} e^{-y/\beta}}{\beta^\alpha \Gamma(\alpha)} dy = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-y/\beta} dy = \frac{1}{\beta^\alpha \Gamma(\alpha)} \beta^\alpha \Gamma(\alpha) = 1$$

Graphs of a Gamma Probability Distribution



https://nbviewer.jupyter.org/github/Jkparkaist/IE481/blob/master/Codes/L1_probabilityDistributions.ipynb

- α is called shaped parameters
  - Determine **the shape** of a gamma distribution
- β is scale parameter
  - Multiplying a gamma-distributed random variable by a positive constant produces a random variable that also has a gamma distribution with the same value of α but with an altered value of β

The Mean and Variance of a Gamma Distribution

THEOREM 4.8

If Y is a gamma distribution with α and β, then

$$\mu = E(Y) = \alpha\beta \quad \text{and} \quad \sigma^2 = V(Y) = \alpha\beta^2.$$

Proof:

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy = \int_0^{\infty} y\left(\frac{y^{\alpha-1}e^{-y/\beta}}{\beta^\alpha\Gamma(\alpha)}\right)dy$$

$$= \frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^{\infty} y^\alpha e^{-y/\beta}dy$$

Property: $\int_0^{\infty} y^{\alpha-1}e^{-y/\beta}dy = \beta^\alpha\Gamma(\alpha)$

$$= \frac{1}{\beta^\alpha\Gamma(\alpha)}[\beta^{\alpha+1}\Gamma(\alpha+1)] = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta$$

The Mean and Variance of a Gamma Distribution

THEOREM 4.8

If Y is a gamma distribution with α and β, then

$$\mu = E(Y) = \alpha\beta \quad \text{and} \quad \sigma^2 = V(Y) = \alpha\beta^2.$$

Proof:

$$E(Y^2) = \int_0^\infty y^2 \left(\frac{y^{\alpha-1} e^{-y/\beta}}{\beta^\alpha \Gamma(\alpha)} \right) dy$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty y^{\alpha+1} e^{-y/\beta} dy$$

Property: $\int_0^\infty y^{\alpha-1} e^{-y/\beta} dy = \beta^\alpha \Gamma(\alpha)$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} [\beta^{\alpha+2} \Gamma(\alpha+2)] = \alpha(\alpha+1)\beta^2$$

Property: $\Gamma(\alpha+2) = (\alpha+1)\Gamma(\alpha+1)$
$= (\alpha+1)\alpha\Gamma(\alpha)$

$$V(Y) = E[Y^2] - \big(E(Y)\big)^2 = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha\beta^2$$

Special Cases of a Gamma Probability Distribution

- When α is an integer
  - The distribution function of a gamma-distributed random variable can be expressed as a sum of certain Poisson probabilities **(Exercise 4.99) Derivation**

- When $\alpha = \nu$ and $\beta = 2$ (ν is an integer)
  - The distribution is called a *chi-square* (χ^2) *distribution with* ν *degrees of freedom*
  - The corresponding random variable is called a *chi-square* (χ^2) *random variable*

- When $\alpha = 1$
  - The distribution is called exponential distribution
  - The corresponding random variable is called an *exponential random variable*

Chi-Square Distribution

DEFINITION 4.10

Let ν be a positive integer. A random variable Y is said to have a *chi-square distribution with* ν *degrees of freedom* if and only if Y is a gamma-distributed random variable with parameters $\alpha = \nu/2$ and $\beta = 2$.

The Mean and Variance of Chi-Square Distribution

THEOREM 4.9

If Y is a chi-square random variable with ν degrees of freedom, then

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

Proof: Trivial by Theorem 4.8. ($\alpha = \nu/2, \beta = 2$)

$$\mu = E(Y) = \alpha\beta = (\nu/2)2 = \nu$$

$$\sigma^2 = V(Y) = \alpha\beta^2 = (\nu/2)2^2 = 2\nu$$

Chi-Square Random Variable: Examples

EXERCISE 6.46

Suppose that Y has a gamma distribution with $\alpha = \frac{n}{2}$ for some positive integer n and β equal to some specified value. Use the method of moment-generating functions to show that $W = \frac{2Y}{\beta}$ has a χ^2 distribution with n degrees of freedom.

Chi-Square Random Variable: Examples

SOLUTION 6.46

$$m_W(t) = E[e^{tW}] = E\left[e^{t(2Y/\beta)}\right] = m_Y(2t/\beta) = \left(1 - \beta * (2t/\beta)\right)^{-(n/2)} = (1 - 2t)^{-n/2}$$

Since this is the moment-generating function of χ^2 distribution with n degrees of freedom, it follows that W is a χ^2 random variable with n degrees of freedom.

Exponential Distribution

DEFINITION 4.11

A random variable Y is said to have an *exponential distribution with parameter* $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \frac{1}{\beta} e^{-y/\beta}, & 0 \le y < \infty \\ 0, & elsewhere. \end{cases}$$

- The gamma density function in which $\alpha = 1$ is called the exponential density function

$$p(y) = \frac{y^{\alpha-1} e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)} = \frac{y^{1-1} e^{-y/\beta}}{\beta^{1}\Gamma(1)} = \frac{1}{\beta} e^{-y/\beta}$$

- The exponential density function is often useful for modeling the length of life of electronic components

The Mean and Variance of Exponential Distribution

THEOREM 4.10

If Y is an exponential random variable with parameter β, then

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

Proof: Trivial by Theorem 4.8. ($\alpha = 1$)

Exponential Distribution : Examples

EXAMPLE 4.10 (Memoryless property)

Suppose that Y has an exponential probability density function. Show that, if $a > 0$ and $b > 0$,

$$P(Y > a + b | Y > a) = P(Y > b).$$

Exponential Distribution : Examples

SOLUTION 4.10 (Memoryless property)

From the definition of conditional probability,

$$P(Y > a + b|Y > a) = \frac{P(Y > a + b, Y > a)}{P(Y > a)} = \frac{P(Y > a + b)}{P(Y > a)}$$

Now,

$$P(Y > a + b) = \int_{a+b}^{\infty} \frac{1}{\beta} e^{-y/\beta} dy = -e^{-y/\beta}\Big]_{a+b}^{\infty} = e^{-(a+b)/\beta}.$$

Similarly,

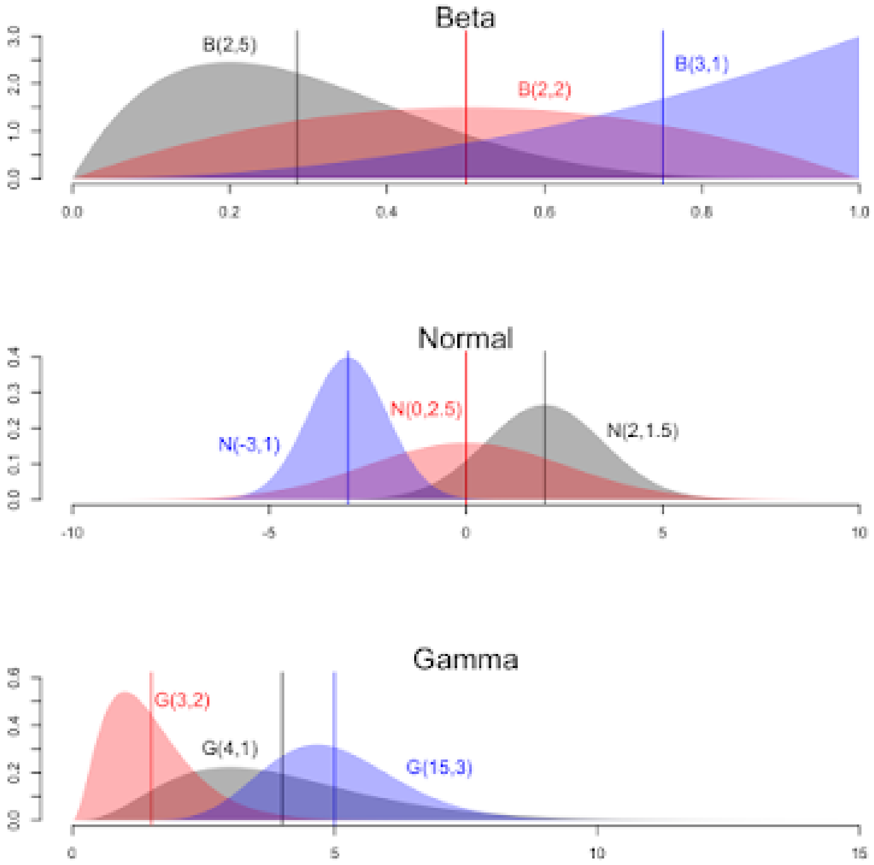
$$P(Y > a) = \int_{a}^{\infty} \frac{1}{\beta} e^{-y/\beta} dy = e^{-a/\beta},$$

and

$$P(Y > a + b|Y > a) = \frac{e^{-(a+b)/\beta}}{e^{-a/\beta}} = e^{-b/\beta} = P(Y > b).$$

Motivations

- We need a random variable that can describe proportions, such as
  - The proportion of impurities in a chemical product
  - The proportion of time that a machine is under repair



Beta Probability Distribution

DEFINITION 4.12

A random variable Y is said to have a *beta probability distribution with parameters* $\alpha > 0$ and $\beta > 0$ if and only if the density function of Y is

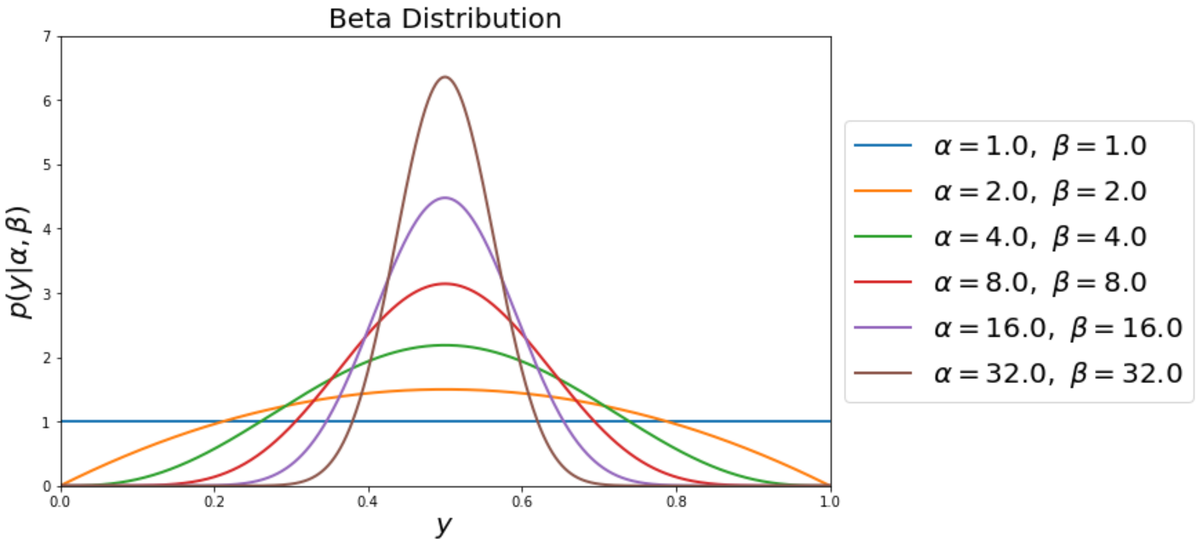
$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}, & 0 \leq y \leq 1 \\ 0, & \text{elsewhere,} \end{cases}$$

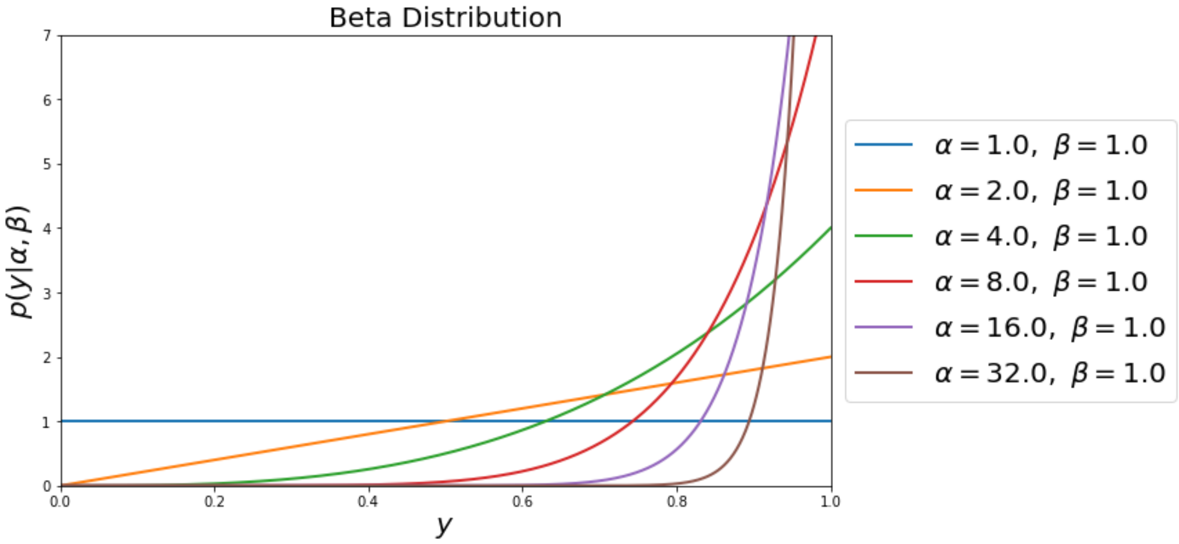
where

$$B(\alpha,\beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

- The beta density function is a two-parameter density function defined over the closed interval $0 \leq y \leq 1$
- Defining y over the interval $0 \leq y \leq 1$ does not restrict the use of the beta distribution
    - If $c \leq y \leq d$, then $y^* = (y-c)/(d-c)$ defines a new variable such that $0 \leq y^* \leq 1$

The Graphs of a Beta Density Function





https://nbviewer.jupyter.org/github/Jkparkaist/IE481/blob/master/Codes/L1_probabilityDistributions.ipynb

- The graphs of beta density functions are different for various values of the two parameters α and β.

The Mean and Variance of Beta Probability Distribution

THEOREM 4.11

If Y is a beta-distributed random variable with parameters $\alpha > 0$ and $\beta > 0$, then

$$\mu = E(Y) = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \sigma^2 = V(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Proof:

The Mean and Variance of Beta Probability Distribution

THEOREM 4.11

If Y is a beta-distributed random variable with parameters $\alpha > 0$ and $\beta > 0$, then

$$\mu = E(Y) = \frac{\alpha}{\alpha+\beta} \quad \text{and} \quad \sigma^2 = V(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Proof:

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy = \int_0^1 y\left(\frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}\right)dy$$

$$= \frac{1}{B(\alpha,\beta)}\int_0^1 y^{\alpha}(1-y)^{\beta-1}dy$$

$$= \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)}$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\alpha}{\alpha+\beta}.$$

The derivation of variance will be on the next session.

Beta Probability Distribution : Examples

EXAMPLE 4.11

A gasoline wholesale distributor has bulk storage tanks that hold fixed supplies and are filled every Monday. Of interest to the wholesaler is the proportion of this supply that is sold during the week. Over many weeks of observation, the distributor found that this proportion could be modeled by a beta distribution with $\alpha = 4$ and $\beta = 2$. Find the probability that the wholesaler will sell at least 90% of her stock in a given week.

Beta Probability Distribution : Examples

SOLUTION 4.11

If Y denotes the proportion sold during the week, then

$$f(y) = \begin{cases} \dfrac{\Gamma(4+2)}{\Gamma(4)\Gamma(2)} y^3(1-y) & 0 \le y \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$P(Y > .9) = \int_{.9}^{\infty} f(y)dy = \int_{.9}^{1} 20(y^3 - y^4)dy = 20\left\{\frac{y^4}{4}\right]_{.9}^{1} - \left.\frac{y^5}{5}\right]_{.9}^{1}\right\} = 20(.004) = .08.$$

Comments

- Density functions are theoretical models for populations of real data that occur in random phenomena
  - How do we know which model to use?
  - How much does it matter if we use the wrong density as our model for reality?

- The purpose of a probabilistic model : a framework to see population
  - We are unlikely ever to select a density function that provides a perfect representation of nature
  - The purpose a probabilistic model is to provide the mechanism for making inferences about a population based on information contained in a sample
  - A good model is one that yields good inferences about the population of interest

- Selecting a proper model
  - Theoretical consideration:
    - ✓ Poisson distribution can be assumed based on the random behavior of events in time
    - ✓ Knowing this, we can model the time between random event using Exponential distribution

  - Form a frequency histogram for data drawn from the population and to chose a density function that would visually appear to give a similar frequency curve.

Comments

- Not all model selection is completely subjective
  - Statistical procedures are available to test a hypothesis that a population frequency distribution is of a particular type
  - We can also calculate a measure of goodness of fit for several distribution and select the best
- Studies of man common inferential methods have been made to determine the magnitude of the errors of inference introduced by incorrect population models
  - Many statistical methods of inference are insensitive to assumptions about the form of the underlying population frequency distribution
- The uniform, normal, gamma, beta distribution offer an assortment of density functions that fit many population frequency distributions.

kth Moment About the Origin and Central Moment

DEFINITION 4.13

If Y is a continuous random variable, then the kth *moment about the **origin*** is the given by

$$\mu_k' = E\left(Y^k\right), \qquad k = 1,2,\dots.$$

The kth *moment about the **mean***, or the kth *central moment*, is given by

$$\mu_k = E[(Y-\mu)^k], \qquad k = 1,2,\dots.$$

kth Moment About the Origin and Central Moment : Examples

EXAMPLE 4.12

Find μ_k' for the uniform random variable with $\theta_1 = 0$ and $\theta_2 = \theta$.

kth Moment About the Origin and Central Moment : Examples

SOLUTION 4.12

$$\mu_k' = E(Y^k) = \int_{-\infty}^{\infty} y^k f(y)dy = \int_0^{\theta} y^k \left(\frac{1}{\theta}\right) dy = \frac{y^{k+1}}{\theta(k+1)}\bigg]_0^{\theta} = \frac{\theta^k}{k+1}.$$

Moment Generating Function

DEFINITION 4.14

If Y is a continuous random variable, then the *moment-generating function of* Y is given by

$$m(t) = E(e^{tY}).$$

The moment-generating function is said to exist if there exists a constant $b > 0$ such that $m(t)$ is finite for $|t| \leq b$.

Finding Using a Moment Generating Function

$$m(t) = E(e^{tY}) = \int_{-\infty}^{\infty} e^{ty} f(y)dy = \int_{-\infty}^{\infty} \left(1 + ty + \frac{t^2y^2}{2!} + \frac{t^3y^3}{3!} + \cdots\right) f(y)dy$$

$$= \int_{-\infty}^{\infty} f(y)dy + \int_{-\infty}^{\infty} yf(y)dy + \frac{t^2}{2!}\int_{-\infty}^{\infty} y^2 f(y)dy + \frac{t^2}{3!}\int_{-\infty}^{\infty} y^3 f(y)dy + \cdots$$

$$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^2}{3!}\mu_3' + \cdots$$

- This take the same form for both discrete and continuous random variables

$$\left.\frac{d^k m(t)}{dt^k}\right]_{t=0} = \mu_k'$$

Moment Generating Function : Examples

EXAMPLE 4.13

Find the moment-generating function for a gamma-distributed random variable.

Moment Generating Function : Examples

SOLUTION 4.13

$$m(t) = E(e^{tY}) = \int_0^\infty e^{ty}\left[\frac{y^{\alpha-1}e^{-y/\beta}}{\beta^\alpha\Gamma(\alpha)}\right]dy$$

$$= \frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty y^{\alpha-1}\exp\left[-y\left(\frac{1}{\beta}-t\right)\right]dy$$

$$= \frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty y^{\alpha-1}\exp\left[-\frac{y}{\beta/(1-\beta t)}\right]dy$$

$$= \frac{1}{\beta^\alpha\Gamma(\alpha)}\left[\left(\frac{\beta}{1-\beta t}\right)^\alpha\Gamma(\alpha)\right]$$

Property: $\int_0^\infty y^{\alpha-1}e^{-y/\beta}dy = \beta^\alpha\Gamma(\alpha)$

$$= \frac{1}{(1-\beta t)^\alpha}, \qquad \text{for} \quad t < \frac{1}{\beta}.$$

Moment Generating Function : Examples

EXAMPLE 4.14

Expand the moment-generating function of Example 4.13 into a power series in t and thereby obtain μ_k'.

Moment Generating Function : Examples

SOLUTION 4.14

From Example 4.13, $m(t) = (1 - \beta t)^{-\alpha}$,

$$m(t) = (1 - \beta t)^{-\alpha} = 1 + (-\alpha)(1)^{-\alpha-1}(-\beta t) + \frac{(-\alpha)(-\alpha - 1)(1)^{-\alpha-2}(-\beta t)^2}{2!} + \cdots$$

$$= 1 + t(\alpha\beta) + \frac{t^2[\alpha(\alpha + 1)\beta^2]}{2!} + \frac{t^3[(\alpha(\alpha + 1)(\alpha + 2)\beta^3]}{3!} + \cdots$$

Because μ_k' is the coefficient of $t^k/k!$,

$$\mu_1' = \mu = \alpha\beta$$
$$\mu_2' = \alpha(\alpha + 1)\beta^2$$
$$\mu_3' = \alpha(\alpha + 1)(\alpha + 2)\beta^3$$

or in general, $\mu_k' = \prod_{i=1}^{k}(\alpha + i - 1)\,\beta^k, k \in \mathbb{N}$

Moment Generating Function

THEOREM 4.12

Let Y be a random variable with density function $f(y)$ and $g(Y)$ be a function of Y. Then the moment-generating function for $g(Y)$ *is*

$$E\left[e^{tg(Y)}\right] = \int_{-\infty}^{\infty} e^{tg(y)} f(y) dy .$$

Proof: Trivial by Definition 4.14 and Theorem 4.4.

- Finding the moments of a function of a random variable is frequently facilitated by using its moment-generating function.

Finding Moments using Moment Generating Function : Examples

EXAMPLE 4.16

Let $g(Y) = Y - \mu$, where Y is a normally distributed random variable with mean μ and variance σ^2. Find the moment-generating function for $g(Y)$.

Finding Moments using Moment Generating Function : Examples

SOLUTION 4.16

The moment-generating function of $g(Y)$ is given by

$$m(t) = E\left[e^{tg(Y)}\right] = E\left[e^{t(Y-\mu)}\right] = \int_{-\infty}^{\infty} e^{t(y-\mu)} \left[\frac{\exp[-(y-\mu)^2/2\sigma^2]}{\sigma\sqrt{2\pi}}\right] dy$$

To integrate, let $u = y - \mu$. Then $du = dy$ and

$$\begin{aligned}
m(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tu} e^{-u^2/(2\sigma^2)} du \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(u^2 - 2\sigma^2 tu)\right] du \\
&= e^{t^2\sigma^2/2} \int_{-\infty}^{\infty} \frac{\exp[-(1/2\sigma^2)(u^2 - 2\sigma^2 tu + \sigma^4 t^2)]}{\sigma\sqrt{2\pi}} du \\
&= e^{t^2\sigma^2/2} \int_{-\infty}^{\infty} \frac{\exp[-(u-\sigma^2 t)^2/2\sigma^2]}{\sigma\sqrt{2\pi}} du = e^{t^2\sigma^2/2},
\end{aligned}$$

$$p(u) = \frac{\exp[-(u-\sigma^2 t)^2/2\sigma^2]}{\sigma\sqrt{2\pi}} = N(\sigma^2 t,\, \sigma^2)$$

The Roles of Moments

- Moments can be used as numerical descriptive measures to describe the data that we obtain in an experiment
- Moments can be used in a theoretical sense to prove that a random variable possesses a particular probability distribution
  - If two random variables Y and Z possess identical moment-generating functions, then Y and Z possess identical probability distributions.

Chi-Square Random Variable: Examples

EXERCISE 6.46

Suppose that Y has a gamma distribution with $\alpha = \frac{n}{2}$ for some positive integer n and β equal to some specified value. Use the method of moment-generating functions to show that $W = \frac{2Y}{\beta}$ has a χ^2 distribution with n degrees of freedom.

Chi-Square Random Variable: Examples

SOLUTION 6.46

$$m_W(t) = E[e^{tW}] = E\left[e^{t(2Y/\beta)}\right] = m_Y(2t/\beta) = \left(1 - \beta * (2t/\beta)\right)^{-(n/2)} = (1 - 2t)^{-n/2}$$

Since this is the moment-generating function of χ^2 distribution with n degrees of freedom, it follows that W is a χ^2 random variable with n degrees of freedom.

Motivations

- Even if the exact distributions are unknown for random variables of interest, knowledge of the associated means and standard deviations permits us to deduce meaningful bounds for the probabilities of events that are often of interest

Tchebysheff's Theorem

THEOREM 4.13 (Tchebysheff's Theorem)

Let Y be a random variable with finite mean μ and variance σ^2. Then, for any $k > 0$,

$$P(|Y-\mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad or \quad P(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof:

Tchebysheff's Theorem

THEOREM 4.13 (Tchebysheff's Theorem)

Let Y be a random variable with finite mean μ and variance σ^2. Then, for any $k > 0$,

$$P(|Y-\mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad or \quad P(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof:

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

$$= \int_{-\infty}^{\mu-k\sigma} (y-\mu)^2 f(y)dy + \int_{\mu-k\sigma}^{\mu+k\sigma} (y-\mu)^2 f(y)dy + \int_{\mu+k\sigma}^{\infty} (y-\mu)^2 f(y)dy$$

$$\geq \int_{-\infty}^{\mu-k\sigma} (y-\mu)^2 f(y)dy + \int_{\mu+k\sigma}^{\infty} (y-\mu)^2 f(y)dy \qquad \because \int_{\mu-k\sigma}^{\mu+k\sigma} (y-\mu)^2 f(y)dy \geq 0$$

$$\geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(y)dy + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(y)dy \qquad \because (y-\mu)^2 \geq k^2\sigma^2$$

$$\geq k^2\sigma^2 \left[\int_{-\infty}^{\mu-k\sigma} f(y)dy + \int_{\mu+k\sigma}^{\infty} f(y)dy\right]$$

Dividing by $k^2\sigma^2 \rightarrow \; P(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

Tchebysheff’s Theorem : Examples

EXAMPLE 4.17

Suppose that experience has shown that the length of time Y (in minutes) required to conduct a periodic maintenance check on a dictating machine follows a gamma distribution with $\alpha = 3.1$ and $\beta = 2$. A new maintenance worker takes 22.5 minutes to check the machine. Does this length of time to perform a maintenance check disagree with prior experience?

Tchebysheff's Theorem : Examples

SOLUTION 4.17

The mean and variance for the length of maintenance check times (based on prior experience) are

$$\mu = \alpha\beta = (3.1)(2) = 6.2 \quad and \quad \sigma^2 = \alpha\beta^2 = (3.1)(2^2) = 12.4.$$

It follows that $\sigma = \sqrt{12} = 3.52$. Notice that $y = 22.5$ minutes exceeds the mean $\mu = 6.2$ minutes by 16.3 minutes, or $k = \frac{16.3}{3.52} = 4.63$ standard deviations. Then from Tchebysheff's theorem,

$$P(|Y - 6.2| \geq 16.3) = P(|Y - \mu| \geq 4.63\sigma) \leq \frac{1}{(4.63)^2} = .0466.$$

This probability is based on the assumption that the distribution of maintenance times has not changed from prior experience. Then, observing that $P(Y \geq 22.5)$ is small, we must conclude either that our new maintenance worker has generated by chance a lengthy maintenance time that occurs with low probability or that the new worker is somewhat slower than preceding ones. Considering the low probability for $P(Y \geq 22.5)$, we favor the latter view.

ETC

The Cumulative Distribution Function for the Beta Random Variable

- The cumulative distribution function for the beta random variable is commonly called the incomplete beta function:

$$F(y) = \int_0^y \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}\,dt = I_{\mathrm{y}}(\alpha,\beta)$$

- $F(y)$ is related to the binomial probability function. Integration by parts can be used to show that for $0 < y < 1$, and α and β

$$F(y) = \int_0^y \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}\,dt = \sum_{i=\alpha}^{n} \binom{n}{i} y^i (1-y)^{n-i}$$

where $n = \alpha + \beta - 1$

✓ The sum on the right-hand side of this expression is just the sum of probabilities associated with a binomial random variable with $n = \alpha + \beta - 1$ and $p = y$