STA255: Statistical Theory

Chapter 4: Continuous Probability Distributions

Summer 2017

Continuous Random Variables

- A random variable Y is called continuous if it can take any value within a finite or infinite interval of the real line.
- Examples of intervals $(-\infty, \infty), (-\infty, a), (a, \infty), (a, b), (a, b], \ldots$
- Examples of experiments resulting in continuous variables:
    - Time: time required to assemble a computer, download time, failure time, lifetime of a computer system, ...
    - Physical measurements: weight, height, distance, velocity, temperature, connection speed,...

Probability Density Function

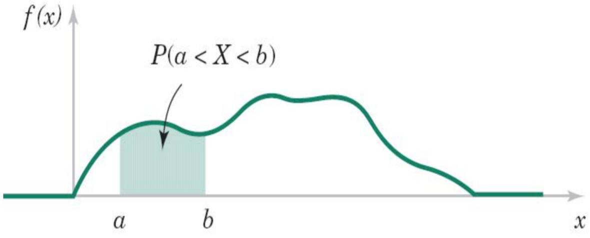
A probability density function (pdf), f(y), can be used to describe the probability distribution of a continuous random variable Y.

Definition (Probability Density Function)

A random variable Y is said to be continuous if there is a function f(y), called the probability density function (pdf), such that

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

Probability Density Function



Note

- $P(a \leq Y \leq b) = \int_a^b f(y)dy.$
- $P(Y = c) = \int_c^c f(y)dy = 0$ for all real number c
- $P(a \leq Y \leq b) = P(a < Y \leq b) = P(a \leq Y < b) = P(a < Y < b)$

Examples

1. Show that $f(y) = 5e^{-5y}, y \geq 0$, is a valid pdf.
   Solution:

2. Let $f(y) = y^2/18, y \in (-3, 2)$ and zero otherwise. Is f a valid pdf?
   Solution:

Definition (The Cumulative Distribution Function)

The cdf F(y) of a continuous random variable Y with pdf f(y) is defined for every number y by

$$F(y) = P(Y \leq y) = \int_{-\infty}^{y} f(t)dt.$$

Properties of F(y)

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 non-decreasing function of* y.
4. $P(a \leq Y \leq b) = F(b) - F(a).$
5. $P(Y \geq a) = P(Y > a) = 1 - P(Y \leq a) = 1 - F(a).$
6. *If* Y *is a continuous r.v. with cdf* $F(y)$ *then at every* y *at which* $F'(y)$ *exists:*

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

The Expected Value

Definition (The Expected Value (Mean))

If Y is a random variable with pdf $f(Y)$, then the expected value (the mean) of Y denoted by $E(Y)$ or μ is given by

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

The Expected Value of a Function

Let $g(Y)$ be a function of Y, then

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

The Variance of a Continuous Random Variable

The Variance of a Continuous Random Variable

The variance of Y, denoted by $V(Y)$ or σ^2 is given by

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

The Variance of a Function

Let $g(Y)$ be a function of Y, then

$$V(g(Y)) = \sigma^2_{h(y)} = E[(g(y) - \mu_{g(Y)})^2] = \int_{-\infty}^{\infty} (g(y) - \mu_{g(y)})^2 f(y)dy$$

Properties of Expectation and Variance

Properties of Expectation

- $E(c) = c$, c *is constant.*
- $E(cg(Y)) = cE(g(Y))$.
- $E(g_1(Y) + \ldots + g_m(Y)) = E(g_1(Y)) + \ldots + E(g_m(Y))$. *Here* $g_1(Y), \ldots, g_m(Y)$ *are* m *functions of* Y.

Properties of Varaince

- $V(Y) = E(Y^2) - (E(Y))^2$
- $V(aY + b) = \sigma^2_{aY+b} = a^2 V(Y) = a^2\sigma^2$
- $sd(aY + b) = \sigma_{aY+b} = |a|\sigma$

Example

Given

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

Find $E(Y)$ and $V(Y)$.

Solution:

Uniform Distribution $U(\theta_1, \theta_2)$

- The uniform random variable is used to model the behaviour of a continuous random variable whose values are uniformly or evenly distributed over a given interval.

Uniform Distribution

A random variable Y is said to be uniformly distributed over the interval (θ_1, θ_2), denoted by $Y \sim U(\theta_1, \theta_2)$, if its density function is

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1} & \theta_1 \leq y \leq \theta_2 \\ 0 & elsewhere \end{cases}$$

Expectation and Variance of Uniform Distribution

If $Y \sim U(\theta_1, \theta_2)$, then:

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

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

Proof:

Expectation and Variance of Uniform Distribution

Example # 4.50

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

Normal Distribution $N(\mu, \sigma^2)$

- The normal distribution is one of the most commonly used probability distribution for applications.
- The Normal Distribution denoted by $N(\mu, \sigma^2)$ has two parameters associated with it; the mean μ and the standard deviation σ .

Normal Distribution

- *A random variable Y is said to be normally distributed if its density function is*

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

- *If $Y \sim N(\mu, \sigma^2)$,then:*

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

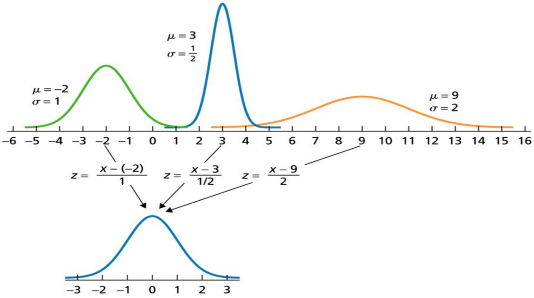
- *The standard normal distribution, $N(0, 1)$ is defined by*

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

Properties of Normal Distribution

- The normal distribution has a bell-shaped curve:
  - measurements concentrate near the mean μ.
  - symmetric distribution (Mean=Median).
  - the variance σ^2 summarizes the variability

    larger variance $\Rightarrow$ measurements are more variable.



Example: #4.68

The grade point averages (GPAs) of a large population of college students are approximately normally distributed with mean 2.4 and standard deviation 0.8. What fraction of the students will possess a GPA in excess of 3.0?

Solution:

Gamma Distribution

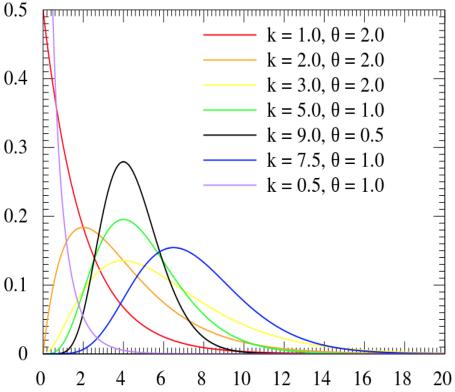
- A random variable Y is said to have a gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ if and only if

$$f(y) = \frac{1}{\beta^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{y/\beta}, \quad y \geq 0,$$

where $\Gamma(\alpha)$ is the gamma function: $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$.

- Properties of Gamma function:
- For $\alpha > 1, \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.
- If α is an integer, then $\Gamma(\alpha) = (\alpha - 1)!$
  Example: $\Gamma(4) = (4 - 1)! = 6$.
- $\Gamma(1/2) = \sqrt{\pi}$.
- α: the shape parameter β: the scale parameter.

Gamma Distribution



Source: wikipedia.org. In the figure, $k = \alpha$ and $\theta = \beta$.

Gamma Distribution

Gamma Distribution: $Y \sim Gamma(\alpha, \beta)$

The mean and the variance are:

$$\mu = \alpha\beta \text{ and } \sigma^2 = \alpha\beta^2$$

Chi-square Distribution

Chi-square Distribution

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

Expectation and Variance

If $Y \sim \chi^2_\nu$, *then*

$$E(Y) = \nu \text{ and } V(Y) = 2\nu$$

Exponential Distribution

Exponential Distribution

- *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*

$$f(y) = \frac{1}{\beta} e^{-y/\beta}, \quad y \geq 0.$$

- $Exp(\beta) = Gamma(1, \beta)$

Expectation and Variance

If $Y \sim Exp(\beta)$, then

$$E(Y) = \beta \text{ and } V(Y) = \beta^2$$

Example: #4.88

The magnitude of earthquakes recorded in a region of North America can be modelled as having an exponential distribution with mean 2.4, as measured on the Richter scale. Find the probability that an earthquake striking this region will

(a) exceed 3.0 on the Richter scale.

Solution:

Example: #4.88

(b) fall between 2.0 and 3.0 on the Richter scale.

Solution:

Beta Distribution

- It is useful to model phenomena constrained to a finite interval of possible values.

Beta Distribution

The beta probability distribution with parameters α and β is

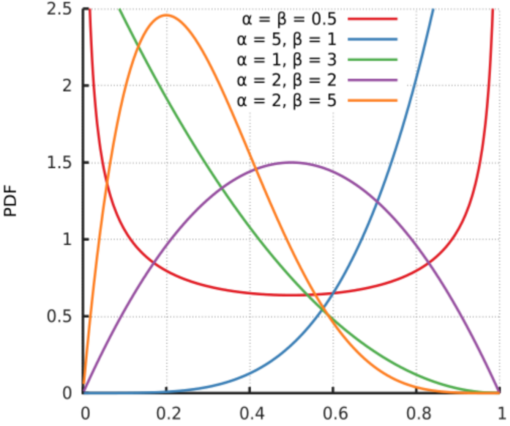
$$f(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}, \quad 0 < y < 1.$$

- Note: since $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1} dy = 1$, we have

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

- The uniform is a special case of the beta distribution with $\alpha = \beta = 1$.

Beta Distribution



Beta Distribution

Beta Distribution: $Y \sim \text{beta}(\alpha, \beta)$

The mean and the variance are:

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

Example

A gasoline wholesale distributor uses bulk storage tanks to hold a fixed supply. The tanks are filled every Monday. Of interest to the wholesaler is the proportion of the supply sold during the week. Over many weeks, this proportion has been observed to match fairly well a beta distribution with $\alpha = 4$ and $\beta = 2$.

(a) Find the expected value and the variance of this proportion.

Solution:

Example

(b) Find the probability that the wholesale will sell at least 90% of the stock in a given week?

Solution:

The Moment Generating Function

Moments

- *The kth moment of a random variable Y taken about the origin is defined to be* $E(y^k), k = 1, 2, \ldots$
- $E(Y)$ *is called the first moment about the origin.*
- $E(Y^2)$ *is called the second moment about the origin.*
- $E(Y^k)$ *is called the k moment about the origin.*
- *Definition: The kth moment of a random variable Y taken about its mean is defined to be* $E[(Y - \mu)^k],\ k = 1, 2, \ldots$
- $E[(Y - \mu)^k]$ *is called the k-th central moment about the mean.*

The Moment Generating Function:

Definition (Moment Generating Function)

The moment generating function (mgf) of a random variable Y , denoted by m(t), is defined to be

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

Theorem

$$E(Y^k) = m^{(k)}(0)$$

where $m^{(k)}(0)$ is the kth derivative of m(t) when $t = 0$.

The Moment Generating Function:

Theorem

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

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

Example

If $Y \sim \mathit{Gamma}(\alpha, \beta)$, find the moment generating function and use it to find the mean and variance of distribution.

Solution:

Example:

Example #4.137

If Y is a random variable with mgf $m(t)$ and U is given by $U = aY + b$, show that the mgf of U is $e^{tb}m(at)$. If Y has mean μ and variance σ^2, use the mgf of U to derive the mean and variance of U.

Solution:

Example:

Example:

Example #4.16

Let Y follow

$$f(y) = \begin{cases} c(2-y) & 0 \leq y \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

(a) Find the constant c and $F(y)$

(b) Graph $f(y)$ and $F(y)$.

(c) Find $P(1 \leq Y \leq 2)$.

Solution:

Example

Example

Example

Example: #4.96

Suppose that a random variable Y has a probability density function given by

$$f(y) = \begin{cases} ky^3 e^{-y/2} & y > 0 \\ 0 & elsewhere \end{cases}$$

(a) Find the value of k that makes $f(y)$ a density function.

Solution:

Example: #4.96

(b) Does Y have a χ^2 distribution? If so, how many degrees of freedom?
Solution:

(c) What are the mean and standard deviation of Y?
Solution: