Probability and Statistics I

STAT 3600 – Fall 2021

Le Chen
lzc0090@auburn.edu

Last updated on

July 4, 2021

Auburn University
Auburn AL

Chapter 3. Continuous Distributions

Chapter 3. Continuous Distributions

Definition 3.1-1 X is a *continuous* r.v. only if its range contains an interval (either finite or infinite) of real numbers.

Examples of continuous r.v. are the length of time it takes when waiting in line to buy frozen yogurt, the weight of a “1-pound” package of hot dogs, etc.

Definition 3.1-2 The function $f(x)$ is called the *probability density function (pdf)* of the *continuous* r.v. X, with space S if it satisfies the following conditions
(a) $f(x) \geq 0, \forall x \in S$.
(b) $\int_S f(x)dx = 1$.
(c) If $(a, b) \subset S$, then the probability of the event $\{a < X < b\}$ is

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

The corresponding distribution of probability is said to be of *continuous type*.

Definition 3.1-3 The *cumulative distribution function (cdf)* of a random variable X of the continuous type, defined in terms of the X, is given by

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt, \qquad -\infty < x < \infty.$$

Remark 3.1-1 By the fundamental theorem of calculus, we have, for x values for which the derivative $F'(x)$ exists,

$$F'(x) = f(x).$$

Example 3.1-1 Let X be a continuous r.v. with pdf

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

where k is a constant.
(a) Determine the value of k and sketch $f(x)$.
(b) Find and sketch the corresponding cdf $F(x)$.
(c) Find $\mathbb{P}\,(1/4 < X \leq 2)$.

Remark 3.1-2 Note that if X is continuous r.v.

$$\mathbb{P}(X = b) = 0, \qquad \text{for all real values of } b.$$

As a consequence,

$$\begin{aligned}
\mathbb{P}(a \le X \le b) &= \mathbb{P}(a < X < b) \\
&= \mathbb{P}(a \le X < b) \\
&= \mathbb{P}(a < X \le b) \\
&= F(b) - F(a).
\end{aligned}$$

Example 3.1-2 The pdf of a continuous r.v. X is given by

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

Find the corresponding cdf $F(x)$ and sketch $f(x)$ and $F(x)$.

Expected value of X, or the mean of X, is

$$\mu = \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

The variance of X is

$$\sigma^2 = Var(X) = \mathbb{E}[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx,$$

and the standard deviation of X is

$$\sigma = \sqrt{\mathrm{Var}(X)}.$$

The moment-generating function, if it exists, is

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx.$$

Example 3.1-3 Find the mean and variance of the r.v. X of Example 3.1-1.

Uniform distribution

Definition 3.1-4 A r.v. X is called a *uniform* r.v. over (a, b) if its pdf is given by

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

The corresponding cdf of uniform r.v. X is

$$F(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a < x < b \\ 1 & \text{if } x \geq b. \end{cases}$$

The mean and variance of the uniform r.v. X are

$$\mu = \mathbb{E}(X) = \frac{a+b}{2} \quad \text{and} \quad \sigma^2 = Var(X) = \frac{(b-a)^2}{12}.$$

The moment-generating function is

$$M(t) = \begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0. \end{cases}$$

Example 3.1-4 Let $f(x) = 1/2, -1 \leq x \leq 1$, and 0 otherwise be the pdf of X.
(a) Graph the p.d.f. and distribution function.
(b) Find the mean, variance, and mgf.

Example 3.1-5 If the moment-generating function of X is

$$M(t) = \frac{e^{5t} - e^{4t}}{t}, \quad t \neq 0 \qquad \text{and} \qquad M(0) = 1.$$

Find (a) $\mathbb{E}(X)$; (b) $Var(X)$; and (c) $\mathbb{P}\left(4.2 < X \leq 4.7\right)$.

Example 3.1-6 Let X have the pdf

$$f(x) = \begin{cases} xe^{-x} & \text{if } 0 \leq x < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Find the mgf, mean, and variance of r.v. X.

Definition 3.1-5 The *(100p)th percentile* is a number π_p such that the area under $f(x)$ to the left of π_p is p, that is,

$$p = \int_{-\infty}^{\pi_p} f(x)dx = F(\pi_p).$$

The 50th percentile is called the *median*. We let $m = \pi_{0.50}$. The 25th and 75th percentiles are called the *first* and and $q_3 = \pi_{0.75}$.

Example 3.1-7

Let X have the pdf

$$f(x) = \begin{cases} e^{-x-1} & -1 < x < \infty, \\ 0 & x \le -1. \end{cases}$$

(a) Find $\mathbb{P}(X \ge 1)$.

(b) Find mgf.

(c) Find the mean and variance.

(d) Find the first quartile, the second or median, and the third quartile.

Exercises from textbook:section 3.1: 1, 2, 3, 4, 6, 7, 8, 9, 10, 16, 18, 20.

Exponential distribution

Definition 3.2-1 A r.v. X is called an *exponential* r.v. with parameter $\lambda(> 0)$ if its pdf is given by

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & 0 \leq x < \infty, \\ 0 & 0 > x. \end{cases}$$

The corresponding distribution function is

$$F(x) = \begin{cases} 0 & \text{if } -\infty < x < 0, \\ 1 - e^{-x/\theta} & \text{if } 0 \leq x < \infty. \end{cases}$$

and the moment-generating function is

$$M(t) = \frac{1}{1 - \theta t}, \qquad \forall t < \frac{1}{\theta},$$

from which, we see that

$$\mu = M'(0) = \theta \quad \text{and} \quad \sigma^2 = M''(0) - [M'(0)]^2 = \theta^2.$$

Remark 3.2-1 It is useful to note that for an exponential random variable, X, we have

$$\mathbb{P}(X > x) = 1 - F(x) = 1 - (1 - e^{-x/\theta}) = e^{-x/\theta} \quad \text{when} \quad x > 0.$$

Hence, W, the waiting time until the first occurrence in a Poisson process($\lambda > 0$), has an exponential distribution with parameter $\theta = 1/\lambda$.

Example 3.2-1 It is known that the time (in hours) between consecutive traffic accidents can be described by the exponential r.v. X with parameter $\theta = 60$. Find (a) $\mathbb{P}(X \leq 60)$; (b) $\mathbb{P}(X > 120)$; (c) $\mathbb{P}(10 < X < 100)$; (d) the median time.

Example 3.2-2 Suppose that a certain type of electronic component has an exponential distribution with a mean life of 500 hours. If X denotes the life of this component (or the time to failure of this component), then

(a) What is the probability that this component will last at least 300 hours.

(b) Given that it has lasted at least 300 hours, what is the conditional probability that it will last at least another 600 hours.

The most interesting property of the exponential distribution is its ”memoryless” property,

$$\mathbb{P}(X > x + t | X > t) = \mathbb{P}(X > x).$$

By this we mean that if the lifetime of an item is exponentially distributed, then an item which has been in use for some hours is as good as a new item with regard to the amount of time remaining until the item fails.

The exponential distribution is the only continuous distribution which possesses this memoryless property.

Gamma distribution

Definition 3.2-2 The *gamma function* is defined by

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \qquad 0 < t.$$

Remark 3.2-2 For $t > 1$, by integration-by-parts, one finds the recursive relation:

$$\Gamma(t) = (t-1)\Gamma(t-1).$$

For example, $\Gamma(5) = 4\Gamma(4)$, $\Gamma(4) = 3\Gamma(3)$, $\Gamma(3) = 2\Gamma(2)$, $\Gamma(2) = 1\Gamma(1)$ and $\Gamma(1) = 1$ Thus, when n is a positive integer, we have

$$\Gamma(n) = (n-1)!$$

Here are some other special values:

$$\begin{aligned}
\Gamma\left(-\frac{3}{2}\right) &= \frac{4\sqrt{\pi}}{3} \\
\Gamma\left(-\frac{1}{2}\right) &= -2\sqrt{\pi} \\
\Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\
\Gamma\left(1\right) &= 0! = 1 \\
\Gamma\left(\frac{3}{2}\right) &= \frac{\sqrt{\pi}}{2} \\
\Gamma\left(2\right) &= 1! = 1 \\
\Gamma\left(\frac{5}{2}\right) &= \frac{3\sqrt{\pi}}{4}
\end{aligned}$$

and

$$\frac{1}{\Gamma(-3)} = \frac{1}{\Gamma(-2)} = \frac{1}{\Gamma(-1)} = \frac{1}{\Gamma(0)} = 0$$

Definition 3.2-3 A r.v. X has a *gamma distribution* if its pdf is defined by

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta}, \qquad 0 \le x < \infty.$$

Remark 3.2-3 W, the waiting time until the αth occurrence in a Poisson process($\lambda > 0$), has a gamma distribution with parameters α and $\theta = 1/\lambda$.

The moment generating function of X is

$$M(t) = \frac{1}{(1-\theta t)^\alpha}, \qquad t < 1/\theta$$

The mean and the variance are

$$\mu = \alpha\theta \quad \text{and} \quad \sigma^2 = \alpha\theta^2.$$

Example 3.2-3 Telephone calls enter a college switchboard at a mean rate of two-thirds of a call per minute according to a Poisson process. Let X denote the waiting time until the tenth call arrives.

(a) What is the pdf of X?

(b) What are the moment-generating function, mean, and variance of X.

Example 3.2-4 If X has a gamma distribution with $\theta = 4$ and $\alpha = 2$, find $\mathbb{P}(X < 5)$.

Example 3.2-5 In a medical experiment, a rat has been exposed to some radiation. The experimenters believe that the rat's survival time X (in weeks) has the pdf

$$f(x) = \frac{3x^2}{120^3} e^{-(x/120)^3} \qquad 0 < x < \infty.$$

(a) What is the probability that the rat survives at least 100 weeks? (Ans. $e^{-(125/216)}$)
(b) Find the expected value of the survival time. (Ans: $120\Gamma(4/3)$)

Chi-square distribution

Definition 3.2-4 A r.v. X has a *chi-square distribution*, denoted as $X \sim \chi^2(r)$, if its pdf is defined by

$$f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-r/2}, \qquad 0 \leq x < \infty.$$

Remark 3.2-4 The chi-square distributions is a gamma distribution with $\theta = 2$ and $\alpha = r/2$.

The mean and the variance of this chi-square distribution are, respectively,

$$\mu = \alpha\theta = \left(\frac{r}{2}\right) 2 = r \quad \text{and} \quad \sigma^2 = \alpha\theta^2 = \left(\frac{r}{2}\right) 2^2 = 2r$$

and the moment-generating function is

$$M(t) = (1 - 2t)^{-r/2} \qquad \forall t < 1/2.$$

Example 3.2-6 If $X \sim \chi^2(23)$, find the following:
(a) $\mathbb{P}(14.85 < X < 32.01)$. (Use Table IV)
(b) Constants a and b such that $\mathbb{P}(a < X < b) = 0.95$ and $\mathbb{P}(X < a) = 0.025$.
(c) The mean and the variance of X.
(d) $\chi^2_{0.05}(23)$ and $\chi^2_{0.95}(23)$.

Exercises from textbook: Section 3.2: 1, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15, 17, 18, 21, 22

Definition 3.3-1 A r.v. X follows the *normal distribution* with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, denoted as $X \sim N(\mu, \sigma^2)$, if its pdf is defined by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty.$$

The moment-generating function of X is

$$M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

Consequently,

$$\mathbb{E}(X) = M'(0) = \mu,$$

$$\mathrm{Var}(X) = M''(0) - [M'(0)]^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

That is, the parameters μ and σ^2 in the pdf of X are the mean and the variance of X.

Standard normal distribution

If $Z \sim N(0, 1)$, we say that Z has a standard normal distribution.

The distribution function of Z is

$$\Phi(z) = \mathbb{P}(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw.$$

Values of $\Phi(z)$ for $z \geq 0$ is given in Table Va in the appendix of the textbook.

Because of the symmetry of the standard normal pdf it is true that

$$\Phi(-z) = 1 - \Phi(z), \qquad \forall z \in \mathbb{R}.$$

Again, because of symmetry of the standard normal pdf, when $z > 0$,

$$\Phi(-z) = \mathbb{P}(Z \leq -z) = \mathbb{P}(Z > z)$$

can be read directly from Table Vb.

Example 3.3-1 If $Z \sim N(0, 1)$, find
(a) $\mathbb{P}(0.53 < Z \leq 2.06)$.
(b) $\mathbb{P}(-0.79 \leq Z < 1.52)$.
(c) $\mathbb{P}(Z > -1.77)$.
(d) $\mathbb{P}(Z > 2.89)$.

In statistical applications, we are often interested in finding a number z_α such that

$$\mathbb{P}(Z \geq z_\alpha) = \alpha,$$

where Z is $N(0,1)$ and α is usually less that 0.5. That is, z_α is the $100(1-\alpha)th$ percentile (sometimes called the upper 100α percent point) for standard normal distribution.

Example 3.3-2 Find the values of (a) $z_{0.01}$; (b) $-z_{0.005}$; (c) $z_{0.0475}$.

Theorem 3.3-1 If $X \sim N\left(\mu, \sigma^2\right)$, then $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$.

Theorem 3.3-1 can be used to find probabilities relating to $X \sim N\left(\mu, \sigma^2\right)$, as follows:

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

since $(X-\mu)/\sigma$ is $N(0,1)$.

Example 3.3-3 If the moment-generating function of X is $M(t) = \exp(166t + 200t^2)$, find (a) The mean of X; (b) The variance of X; (c) $\mathbb{P}(170 < X < 200)$; (d) $\mathbb{P}(148 < X < 172)$; (e) $\mathbb{P}(|X - 166| < 40)$.

Theorem 3.3-2 If $X \sim N(\mu, \sigma^2)$, then $V = (X - \mu)^2/\sigma^2 = Z^2 \sim \chi^2(1)$.

Example 3.3-4 If Z is $N(0, 1)$, find values of c such that (a) $\mathbb{P}(Z \geq c) = 0.025$; (b) $\mathbb{P}(|Z| \leq c) = 0.9$.

Exercises from textbook: Section 3.3: 1, 3, 5cd, 6, 7, 8, 9, 11, 12, 13, 17.