CHAPTER III
CONTINUOUS RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

Department of Foundation Year
Institute of Technology of Cambodia

LIN Mongkolsery
sery@itc.edu.kh

2021-2022

Contents

Probability Distributions for Continuous Variables

Definition 1

Let X be a continuous rv. Then a **probability distribution** or **probability density function (pdf)** of X is a function $f(x)$ such that for any two numbers a, b with $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Property

For $f(x)$ to be a legitimate pdf, it must satisfy the following two conditions:

1. $f(x) \geq 0$ for all x
2. $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1.$

Probability Distributions for Continuous Variables

Example 2

Suppose that X is a continuous rv whose pdf is given by

$$f(x) = \begin{cases} C(4x - 2x^2), & 0 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine the value of C ?

(b) Find $P(X > 1)$.

Probability Distributions for Continuous Variables

Example 3

The current in a certain circuit as measured by an ammeter is a continuous random variable X with the following density function:

$$f(x) = \begin{cases} 0.075x + 0.2, & 3 \leq x \leq 5, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Graph the pdf and verify that the total area under the density curve is indeed 1.

(b) Calculate $P(X \leq 4)$. How does this probability compare to $P(X < 4)$?

(c) Calculate $P(3.5 \leq X \leq 4.5)$ and $P(4.5 < X)$.

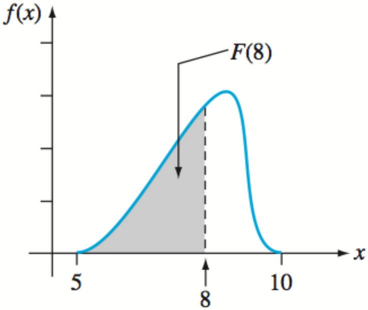
The Cumulative Distribution Function

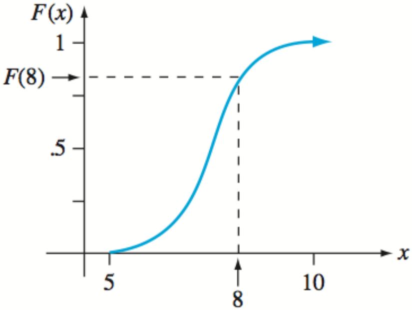
Definition 4

The **cumulative distribution function (cdf)** $F(x)$ for a continuous rv X is defined for every number x by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy.$$

The Cumulative Distribution Function





For each x, $F(x)$ is the area under the density curve to the left of x. This is illustrated in figure above, where $F(x)$ increases smoothly as x increases.

The Cumulative Distribution Function

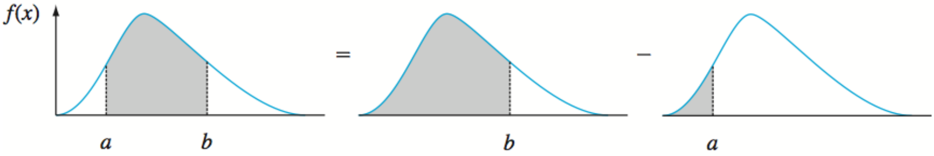
Theorem 1

Let X be a continuous rv with pdf $f(x)$ and cdf $F(x)$. Then for any number a,

$$P(X > a) = 1 - F(a)$$

and for any two numbers a and b with $a < b$,

$$P(a \leq x \leq b) = F(b) - F(a)$$



The Cumulative Distribution Function

Example 5

Suppose the pdf $f(x)$ of the magnitude X of a dynamic load on a bridge (in newtons) is given by

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x, & 0 \leq x \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

Find the cumulative distribution function.

Theorem 2

If X is a continuous rv with pdf $f(x)$ and cdf $F(x)$, then

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

provided that $F'(x)$ exists.

The Cumulative Distribution Function

Definition 6

Let p be a number between 0 and 1. The $(100p)$**th percentile** of the distribution of a continuous rv X, denoted by $\eta(p)$, is defined by

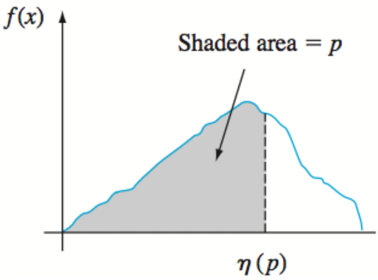
$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)dy$$

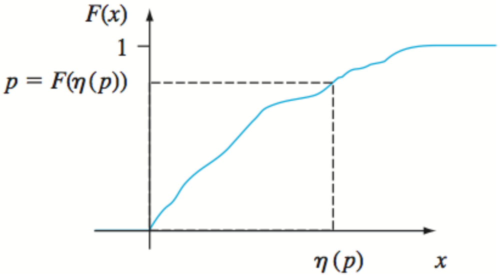
Definition 7

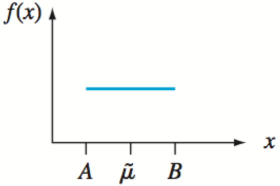
The **median** of a continuous distribution, denoted by $\widetilde{\mu}$, is the 50th percentile, so $\widetilde{\mu}$ satisfies $0.5 = F(\widetilde{\mu})$. That is,

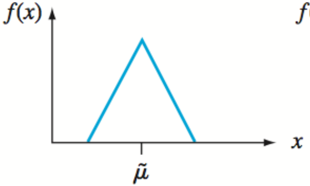
$$\frac{1}{2} = F(\widetilde{\mu}) = \int_{-\infty}^{\widetilde{\mu}} f(y)dy.$$

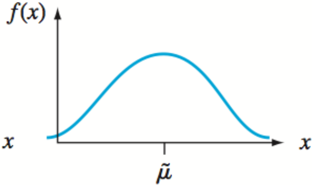
The Cumulative Distribution Function











Expected Values

Definition 8

Let X be a continuous rv with pdf $f(x)$.

1. The **expected** or **mean value** of X is

$$\mu_X = E(X) = \int_{-\infty}^{\infty} xf(x)dx.$$

2. The **variance** of X is

$$V(X) = E[(X-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

3. The **standard deviation** of X is

$$\sigma_X = \sqrt{V(X)}$$

Expected Values

Definition 9

The **moment-generating function (mgf)** of a continuous rv X, if it exists, is

$$M(t) = E\left(e^{tX}\right) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx$$

Theorem 3

If X is a continuous rv with pdf $f(x)$ and $h(X)$ is any function of X, then

$$E[h(X)] = \int_{-\infty}^{\infty} h(x).f(x) dx$$

Expected Values

Theorem 4

Let X is a continuous rv. Then

1. $E(aX + b) = aE(X) + b$
2. $V(x) = E(X^2) - E^2(X)$
3. $V(aX + b) = a^2 V(X)$
4. $\sigma_{aX+b} = |a|\sigma_X$
5. $M^{(n)}(t) = E[X^n e^{tX}]$
6. $E[X] = M'(0)$
7. $V(X) = M''(0) - [M'(0)]^2$

Example 10

The weekly demand for propane gas (in 1000s of gallons) from a particular facility is an rv X with pdf

$$f(x) = \begin{cases} 2\left(1 - \dfrac{1}{x^2}\right), & 1 \le x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Compute the cdf of X.

(b) Obtain an expression for the $(100p)$th percentile. What is the value of $\widetilde{\mu}$?

(c) Compute $E(X)$ and $V(X)$.

(d) If 1.5 thousand gallons are in stock at the beginning of the week and no new supply is due in during the week, how much of the 1.5 thousand gallons is expected to be left at the end of the week?

The Uniform Distribution

Definition 11

A continuous rv X is said to have a **uniform distribution** on the interval $[a, b]$ if the pdf of X is

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise} \end{cases}$$

In this case, we write $X \sim \mathrm{U}(a, b)$.

Theorem 5

If $X \sim \mathrm{U}(a, b)$, then

$$E(X) = \frac{a+b}{2}, \qquad V(X) = \frac{(b-a)^2}{12}, \qquad M(t) = \begin{cases} \dfrac{e^{tb} - e^{ta}}{t(b-a)}, & t \neq 0 \\ 1, & t = 0 \end{cases}$$

The Uniform Distribution

Example 12

Buses arrive at a specified stop at 15-minute intervals starting at 7 A.M. That is, they arrive at $7, 7:15, 7:30, 7:45$, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and $7:30$, find the probability that he waits

(a) less than 5 minutes for a bus;

(b) more than 10 minutes for a bus.

The Normal Distribution

Definition 13

A continuous rv X is said to have a **normal distribution** with parameters μ and σ (or μ and σ^2), where $-\infty < \mu < \infty$ and $0 < \sigma$, if the pdf of X is

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

In this case, we write $X \sim \mathrm{N}(\mu, \sigma^2)$.

Theorem 6

If $X \sim \mathrm{N}(\mu, \sigma^2)$, then

$$E(X) = \mu, \qquad V(X) = \sigma^2, \qquad M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

The Standard Normal Distribution

Definition 14

The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution**. A random variable having a standard normal distribution is called a **standard normal random variable** and will be denoted by Z. The pdf of $Z \sim \mathrm{N}(0, 1)$ is

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

The graph of $f(z)$ is called the standard normal (or z) curve. Its inflection points are at 1 and -1. The cdf of Z is

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} f(y)dy.$$

Nonstandard Normal Distributions

Theorem 7

Let $X \sim \mathrm{N}(\mu, \sigma^2)$. If $Z = \dfrac{X - \mu}{\sigma}$, then $Z \sim \mathrm{N}(0, 1)$. Thus

$$\begin{aligned} P(a \le X \le b) &= P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \\ P(X \le a) &= \Phi\left(\frac{a-\mu}{\sigma}\right) \\ P(X \ge b) &= 1 - \Phi\left(\frac{b-\mu}{\sigma}\right) \end{aligned}$$

Nonstandard Normal Distributions

Example 15

Determine the value of the constant c that makes the probability statement correct:

(a) $\Phi(c) = .9838$

(b) $P(c \leq Z \leq 1) = .291$

(c) $P(c \leq Z) = .121$

(d) $P(-c \leq Z \leq c) = .668$

(e) $P(c \leq |Z|) = .016$

Nonstandard Normal Distributions

Example 16

Scores on an examination are assumed to be normally distributed with mean 65 and variance 36.

1. What is the probability that a person taking the examination scores higher than 70?
2. 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?
3. What must be the cutoff point for passing the examination if the examiner wants only the top 75% of all scores to be passing?
4. Approximately what proportion of students have scores 5 or more points above the score that cuts off the lowest 15%?
5. If it is known that a student's score exceeds 72, what is the probability that his or her score exceeds 84?

Approximating the Binomial Distribution

Theorem 8

Let X be a binomial rv based on n trials with success probability p. Then if the binomial probability histogram is not too skewed, X has approximately a normal distribution with $\mu = np$ and $\sigma = \sqrt{npq}$. In particular, for $x =$a possible value of X,

$$P(X \leq x) = B(x; n, p) \approx \left(\begin{array}{c} \text{area of the normal curve} \\ \text{to the left of } x + 0.5 \end{array} \right)$$

$$= \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right)$$

In practice, the approximation is adequate provided that both $np \geq 10$ and $nq \geq 10$, since there is then enough symmetry in the underlying binomial distribution. $B(x; n, p)$ is the cdf of $\mathrm{Bin}(n, p)$.

Approximating the Binomial Distribution

Example 17

Suppose only 75% of all drivers in a certain state regularly wear a seat belt. A random sample of 500 drivers is selected. What is the probability that

(a) Between 360 and 400 (inclusive) of the drivers in the sample regularly wear a seat belt?

(b) Fewer than 400 of those in the sample regularly wear a seat belt?

The Exponential Distribution

Definition 18

A continuous rv X is said to have an **exponential distribution** ($X \sim \mathrm{Exp}(\lambda)$) with parameter $\lambda (\lambda > 0)$ if the pdf of X is

$$f(x) = \begin{cases} \dfrac{1}{\lambda} e^{-\frac{x}{\lambda}}, & x \geq 0, \\ 0, & \text{otherwise} \end{cases}$$

Theorem 9

If $X \sim \mathrm{Exp}(\lambda)$, then

1. $F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\frac{x}{\lambda}}, & x \geq 0. \end{cases}$
2. $E(X) = \lambda, \qquad V(X) = \lambda^2, \qquad M(t) = \dfrac{1}{1 - \lambda t} \quad , t < \dfrac{1}{\lambda}.$

The Exponential Distribution

Example 19

Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = 10$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait

(a) more than 10 minutes;

(b) between 10 and 20 minutes.

The Gamma Function

Definition 20

The **gamma function** and the **incomplete gamma function** are defined respectively by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt, \quad \alpha > 0.$$

$$\Gamma(x, \alpha) = \int_0^x t^{\alpha-1} e^{-t} dt, \quad \alpha > 0, x > 0.$$

Theorem 10

1. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1), \quad \alpha > 1.$
2. $\Gamma(n) = (n-1)!, \quad n \in \mathbb{N}.$
3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

The Gamma Distribution

Definition 21

A continuous rv X is said to have a **gamma distribution** with parameters $\alpha > 0$ and $\beta > 0$ if the pdf of X is

$$f(x) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

In this case, we write $X \sim \mathrm{Gam}(\alpha, \beta)$. If $\beta = 1$, we call **standard gamma distribution**.

The Gamma Distribution

Theorem 11

If $X \sim \mathrm{Gam}(\alpha, \beta)$, then

$$E(X) = \alpha\beta, \qquad V(X) = \alpha\beta^2, \qquad M(t) = \frac{1}{(1-\beta t)^\alpha}, \quad t < \frac{1}{\beta}.$$

Theorem 12

Let $X \sim \mathrm{Gam}(\alpha, \beta)$. Then for any $x > 0$, the cdf of X is given by

$$P(X \leq x) = F(x) = \Gamma\left(\frac{x}{\beta}, \alpha\right)$$

The Chi-squared Distribution

Definition 22

Let ν be a positive integer. Then a random variable X is said to have a **chi-squared distribution** with parameter ν, we write $X \sim \chi^2(\nu)$, if the pdf of X is the gamma density with $\alpha = \dfrac{\nu}{2}$ and $\beta = 2$.

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

The parameter ν is called the **number of degree of freedom** of X.

The Chi-squared Distribution

Theorem 13

If $X \sim \chi^2(\nu)$, then

$$E(X) = \nu, \qquad V(X) = 2\nu, \qquad M(t) = (1-2t)^{-\frac{\nu}{2}}, \quad t < \frac{1}{2}.$$

Theorem 14

If the random variable $X \sim \mathrm{N}(\mu, \sigma^2), \sigma^2 > 0$, then the random variable $V = \dfrac{(X-\mu)^2}{\sigma^2} = Z^2 \sim \chi^2(1)$.

Beta Distribution

Definition 23

The **beta function** is defined by

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \alpha, \beta > 0$$

where Γ is the gamma function.

Definition 24

The continuous rv X has a **beta distribution**($X \sim \mathrm{Bet}(\alpha, \beta)$) with parameters $\alpha > 0$ and $\beta > 0$ if its density function is given by

$$f(x) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Beta Distribution

Theorem 15

If $X \sim \mathrm{Bet}(\alpha, \beta)$, then

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

Lognormal Distribution

Definition 25

The continuous rv X has a **lognormal distribution** if the random variable $Y = \ln(X)$ has a normal distribution with mean μ and standard deviation σ. The resulting density function of X is

$$f(x; \mu, \sigma) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2\sigma^2}[\ln(x) - \mu]^2}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

We write $X \sim \mathrm{Log}(\mu, \sigma)$.

Theorem 16

If $X \sim \mathrm{Log}(\mu, \sigma)$, then

$$E(X) = e^{\mu + \sigma^2/2} \quad \text{and} \quad V(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

Weibull Distribution

Definition 26

The continuous rv X has a **Weibull distribution**, with parameters α and β, if its density function is given by

$$f(x;\alpha,\beta)=\begin{cases}\alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}}, & x>0\\ 0, & \text{otherwise.}\end{cases},\quad \alpha>0,\beta>0.$$

We write $X\sim \mathrm{Wei}(\alpha,\beta)$.

Theorem 17

If $X\sim \mathrm{Wei}(\alpha,\beta)$, then

$$E(X)=\alpha^{-1/\beta}\Gamma\left(1+\frac{1}{\beta}\right),V(X)=\alpha^{-2/\beta}\left\{\Gamma\left(1+\frac{2}{\beta}\right)-\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right\}$$