# CHAPTER 4
# Continuous Random Variables and Probability Distributions

## ✠ Basic definitions and properties of continuous random variables

**Continuous random variable:** A random variable is continuous if its set of possible values is an entire interval of numbers.

**Probability distribution for continuous variables:** It is possible to construct a probability histogram (same as relative frequency histogram) for continuous variable. But by measuring the variable more and more finely, the resulting histogram approaches to a smooth curve. It is obvious that the total area under this curve is 1, also probability that the variable be between two points is the area under the curve between two points. It means probability distribution or **probability density function (pdf)** for a continuous random variable $X$ is a function $f(x)$, such that

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

$f(x)$ is a pdf if satisfies the following two conditions:

- $f(x) \geqslant 0$ for all $x$
- $\int_{-\infty}^{\infty} f(x)dx = 1 =$ area under the entire curve of $f(x)$

---

**Example:** Suppose that $X$ has following density function

$$f(x) = \begin{cases} 0.5x & 0 \leqslant x \leqslant 2 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate
**a.** $P(X \leqslant 1)$
**b.** $P(0.5 \leqslant X \leqslant 1.5)$
**c.** $P(1.5 \leqslant X)$

---

**Note:** For a continuous random variable $P(X = C) = 0$, then

$$P(a \leqslant X \leqslant b) = P(a < X \leqslant b) = P(a \leqslant X < b) = P(a < X < b)$$

---

**Example:** Consider the following function

$$f(x) = \begin{cases} 0.15e^{-0.15(x-0.5)} & x \geqslant 0.5 \\ 0 & \text{otherwise.} \end{cases}$$

**a.** Verify that $f(x)$ is a pdf
**b.** $P(X \leqslant 5)$

---

## ✠ Uniform Distribution

A continuous random variable $X$ has uniform distribution on interval $[a, b]$ if the pdf of $X$ is

$$f(x; a, b) = \begin{cases} \frac{1}{b-a} & a \leqslant X \leqslant b \\ 0 & \text{otherwise.} \end{cases}$$

## ✠ Cumulative Distribution Function

The cumulative distribution function (pdf) for a continuous random variable is

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(y)dy \quad \forall x$$

$F(x)$ is the area under the density curve to the left of $x$.

---

**Example:** Find cdf for uniform distribution on $[a, b]$ and then graph it.

---

Same as discrete random variable, the probabilities of intervals can be computed from $F(x)$ as

$$P(X > a) = 1 - F(a), \qquad P(a \leqslant X \leqslant b) = F(b) - F(a)$$

---

**Example:** Solve example 1 by using cdf.

---

- If $X$ is a continuous random variable with cdf $F(x)$ then
$$F'(x) = f(x)$$
- The median $\tilde{\mu}$ for a continuous random variable satisfies $0.5 = F(\tilde{\mu})$.
- The general, the (100p)th percentile of a distribution of a continuous random variable is defined by
$$p = F(\eta(\mu)) = \int_{-\infty}^{\eta(p)} f(y)dy \quad \eta(p) \text{ is } (100p)\text{th} \quad \text{percentile}$$

### ✠ Expected Value for Continuous Random Variables

The expected value or mean of a continuous random variable $X$ with $f(x)$ is

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

The expected value for a function of $X$ is

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

### ✠ Variance of Continuous Random Variables

The variance of a continuous random variable $X$ with pdf $f(x)$ and mean value $\mu$ is

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

and also standard deviation of $X$ is

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

The easier way for computing variance is the following formula

$$V(X) = E(X)^2 - (E(X))^2$$

---

**Example:** The pdf of $X$ is

$$f(x) = \begin{cases} 90x^8(1-x) & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

**a.** Obtain the cdf of $X$.

**b.** What is $P(X \leqslant 0.5)$?

**c.** What is $P(0.25 \leqslant X < 0.5)$?

**d.** What is $75^{th}$ percentile of the distribution?

**e.** What is the probability that $X$ is within 1 standard deviation of its mean value?

---

## ♠ The Normal Distribution

The normal distribution is the most important distribution in probability ad statistics, because it can be fit for a large number of random variables like weights, heights, ..... A continuous random variable $X$ has a **normal distribution with parameter** $\mu$ and $\sigma^2$, if the pdf of $X$ is

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \infty < x < \infty,\ \sigma > 0,\ \infty < \mu < \infty$$

$\mu$ and $\sigma$ are parameters of normal distribution, and $X \sim N(\mu, \sigma^2)$ means random variable $X$ has normal distribution with parameters $\mu$ and $\sigma^2$.
Normal distribution is symmetric, then

$$\int_{-\infty}^{\mu} f(x)dx = \int_{\mu}^{\infty} f(x)dx = 0.5$$

To compute $P(a \leqslant X \leqslant b)$ when $X \sim N(\mu, \sigma^2)$ is

$$\int_a^b \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

For evaluating this expression, use standard normal ($\mu = 0,\ \sigma = 1$) which tabulated for different values of $a$ and $b$.
A random variable with $\mu = 0$ and $\sigma = 1$ is called a standard normal distribution and denoted by $Z$. The pdf of $Z$ is

$$f(z;0,1) = \frac{1}{\sqrt{2\pi}}\, e^{\frac{-x^2}{2}}$$

The cdf of $Z$ is $P(Z \leqslant z) = \int_{-\infty}^{z} f(y)dy$ which will denoted by $\Phi(z)$.

---

**Example:** Compute the following probabilities:

$P(0 \leqslant Z \leqslant 2.17),\ P(-2.5 \leqslant Z \leqslant 0),\ P(-2.5 \leqslant Z \leqslant 2.5),\ P(1.5 \leqslant Z),\ P(|Z| \leqslant 2.5)$

---

### $z_\alpha$ Notation

$z_\alpha$ denotes the value on $z$ axis for which $\alpha$ of the area under the $z$ curve lies to the right of $z_\alpha$. Thus $z_\alpha$ is the $100(1-\alpha)th$ percentile of the standard normal distribution.

---

**Examples:**

1. Determine the value of $c$ that makes the probability statement correct.
   $\Phi(c) = 0.9838$, $\qquad P(c \leqslant Z) = 0.121$, $\qquad P(c \leqslant |Z|) = 0.016$
2. Find the following percentiles for the standard normal distribution
   75th $\qquad$ 9th
3. Determine $z_\alpha$ for the following
   $\alpha = 0.0055$, $\qquad \alpha = 0.663$

---

**Normal Distribution:** If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{x-\mu}{\sigma}$ has a standard normal distribution, thus

$$P(a \leqslant X \leqslant b) = P(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$$

If the population of a variable is (approximately) normal, then
**1.** Roughly 68% of the values are within 1 $SD$ of the mean
**2.** Roughly 95% of the values are within 2 $SD_s$ of the mean
**3.** Roughly 99.7% of the values are within 3 $SD_s$ of the mean

In general the (100p)th percentile of any normal distribution is related to the (100p)th percentile of standard normal distribution as
(100p)th percentile for $N(\mu, \sigma) = \mu +$[(100p)th for standard normal ]$\times \sigma$

**The Normal Approximation to the Binomial Distribution**
When $X \sim b(x; n, p)$, mean and standard deviation are $\mu = np$ and $\sigma = \sqrt{npq}$, then if the probability histogram is not too skewed, $X$ has approximately normal distribution with same mean and standard deviation and for

finding probability based on normal distribution, need continuity correction, for example:

$$P(X \leqslant x) = (\text{area under the normal curve to the left of } x+0.5) = \Phi\left(\frac{x+0.5-np}{\sqrt{npq}}\right)$$

The condition for this approximation is both $np \geqslant 10$ and $nq \geqslant 10$.

---

**Example:** Suppose only 40% of all drivers in a certain state wear a seatbelt. A random sample of 500 drivers is selected, what is the probability that
**a.** Between 180 and 230 (inclusive) of the drivers in the sample wear a seatbelt?
**b.** Fewer than 170 of those in the sample wear a seatbelt? Fewer than 150?

---

## The Gamma Distribution

Normal distribution is bell shape and symmetric, but there is many random variables that have a skewed situation. For these kind of variables first define the gamma function. For $\alpha > 0$, the gamma function $\Gamma(\alpha)$ is

$$\Gamma(\alpha) = \int_o^\infty x^{\alpha-1} e^{-x}$$

This function has following properties:

1. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1) \quad \alpha > 1$
2. $\Gamma(\alpha) = (\alpha - 1)! \quad \forall\alpha$ positive integer
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

In general, continuous random variable $X$ has gamma distribution of the pdf of $X$ is

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

$\alpha$ and $\beta$ are the parameters of distribution and $\alpha > 0, \ \beta > 0$.
For standard gamma distribution $\beta = 1$, then pdf for standard gamma ran-

dom variable is

$$x;\alpha) = \begin{cases} \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} & x \geqslant 0 \\ 0 & \text{otherwise.} \end{cases}$$

The mean and variance for a gamma random variable are

$$E(X) = \mu = \alpha\beta \qquad V(X) = \sigma^2 = \alpha\beta^2$$

**Computing probabilities for Gamma distribution**

For computing probabilities for a gamma random variable, same as normal distribution, probability can find for standard gamma by using table. Divided $x$ by $\beta$ change any kind of gamma distribution to standard gamma, then

$$P(X \leqslant x) = F(x;\alpha,\beta) = F(\frac{x}{\beta};\alpha)$$

---

**Examples:**

1.

- Evaluate the following

$$\Gamma(6) \qquad \Gamma(5/2) \qquad F(5;4)$$

- Let $X$ have a standard gamma distribution with $\alpha = 7$. Evaluate

$$P(X \leqslant 5) \qquad P(3 < X < 8) \qquad P(X < 4 \text{ or } X > 6)$$

2. Suppose the time taken by a homeowner to mow his lawn is an random variable $X$ having a gamma distribution with parameters $\alpha = 2$ and $\beta = \frac{1}{2}$. What is the probability that is takes:

**a.** At most 1 hour to mow the lawn?

**b.** At least 2 hours to mow the lawn?

**c.** Between 0.5 and 1.5 hours to mow the lawn?

---

♠ The Exponential Distribution

$X$ has an exponential distribution if pdf of $X$ is

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0 \\ 0 & \text{otherwise.} \end{cases}$$

Exponential distribution is a special case of gamma distribution with $\alpha = 1$ and $\beta = \frac{1}{\lambda}$, then

$$\mu = \alpha\beta = \frac{1}{\lambda} \qquad \sigma^2 = \alpha\beta^2 = \frac{1}{\lambda^2}$$

The cdf of exponential random variable is

$$F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geqslant 0 \\ 0 & \text{otherwise.} \end{cases}$$

The exponential distribution has two important applications:

- This distribution is a good model for the distribution of times between the occurrence for two successive events. As before the number of event in a time interval follows Poisson distribution, it can be shown that if the time interval is $t$ and average number of events in a unit of time is $\alpha$, then $\lambda = \alpha t$, then the distribution of time between two successive events is exponential with parameter $\lambda = \alpha$

- "Memoryless property"

$$P(X \geqslant t + t_0 | X \geqslant t_0) = \frac{P[(X \geqslant t + t_0) \cap (X \geqslant t_0)]}{P(X \geqslant t_0)}$$
$$= \frac{P(X \geqslant t + t_0)}{P(X \geqslant t_0)} = \frac{1 - F(t + t_0; \lambda)}{1 - F(t_0; \lambda)} = e^{-\lambda t}$$

This property is useful for distribution of component lifetime and it means the distribution of additional lifetime is exactly the same as original distribution of lifetime , in other words, the distribution of the remaining lifetime is independent of current age.

---

**Example:**

The exponential distribution with mean value 6 MPa is used as a model for the distribution of stress range in certain bridge connections. Find
a) Probability that stress range is at most 10 MPa.
b) Probability that stress range is between 5 and 10 MPa.

---

**Suggested Exercises for Chapter 4:** 3, 5, 9, 11, 13, 19, 21, 33,35, 37, 41, 45, 47, 53, 55, 59, 63, 65, 67, 99, 101, 105, 107