A **continuous random variable** can assume any value on some interval.

Recall that, in a relative frequency histogram, the area of a rectangle represents the relative frequency of the corresponding class and the height of a rectangle is called the density.

The probability of a continuous random variable is represented by the area under the curve, called the **probability density function** $f(x)$, which satisfies the following conditions:

$$f(x) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} f(x)dx = 1,$$

and the probability of the random variable being between a and b is

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

Note that for a continuous random variable X, $P(X = x) = 0$.

A continuous random variable X is a **uniform** random variable over the interval [a, b], if X can take any value in closed interval [a, b] and if the probability density function (pdf) of X is constant over this interval, that is,

$$f(x) = \frac{1}{b-a}, \quad \text{for } a \leq x \leq b.$$

For $a \leq c \leq d \leq b$, $P(c < X \leq d) = \frac{d-c}{b-a}$.

For a uniform random variable X

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

Example 1. The amount of time that a tourist waits to board a trolley car on Polk Street in San Francisco has a uniform distribution between 0 and 5 minutes.

a) Sketch the pdf of the amount of time.
b) Find the probability that a tourist waits more than 3 minutes.
c) Find the probability that a tourist waits less than 1 minute.
d) Find the expected value and standard deviation of the amount of time.

MSL Homework: 6.23, 6.24
HC Homework: 6.26 (on p. 239)

The continuous random variable with the following probability density function is called the **normal random variable** and its distribution is called the normal distribution.

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

Here, π=**3.14159**…, $\mu = E(X)$, and $\sigma^2 = Var(X)$.

The probability density function of a normal random variable is mound-shaped (or bell-shaped) and symmetric about its mean μ and has points of inflection at μ-σ and μ+σ.

The normal random variable with mean 0 and standard deviation 1 is called the **standard normal random variable**, and denoted by *Z*.

To find the probabilities of *Z* we will use the standard normal table, also called the *Z*-table, in Table E.2 of Appendix E on pp. 798-799 in the textbook. This cumulative probability table is consistent with MS Excel.

Various forms of probabilities in terms of cumulative probabilities:

$P(X > a) = 1 - P(X \le a)$ $\quad$ $P(X \ge a) = 1 - P(X < a)$

$P(a < X \le b) = P(X \le b) - P(X \le a)$ $\quad$ $P(a \le X \le b) = P(X \le b) - P(X < a)$

$P(a \le X < b) = P(X < b) - P(X < a)$ $\quad$ $P(a < X < b) = P(X < b) - P(X \le a)$

Example 2. Let *Z* denote the standard normal random variable. Find the following probabilities using the table in the handout.

$P(Z \le -0.25)$ $\quad$ $P(Z > -0.53)$ $\quad$ $P(0.38 < Z \le 1.42)$

$P(|Z| \le 1)$ $\quad$ $P(|Z| \ge 1.5)$ $\quad$ $P(Z = 0)$

In order to find $P(Z \le z)$ for a standard normal random variable *Z* (with mean 0 and variance 1), use **Formulas >f_x (Insert Function)> Statistical>NORM.S.DIST**. More specifically, suppose you want to find $P(Z \le 2.226)$. Then click **Formulas >f_x (Insert Function)> Statistical>NORM.S.DIST** and enter 2.226 for z.

Example 3. Let *Z* be the standard normal random variable. Find *z* satisfying the following:

$P(Z \le z) = 0.4052$ $\quad$ $P(Z \ge z) = 0.0322$

$P(-z \le Z \le z) = 0.4515$ $\quad$ $P(1 \le Z \le z) = 0.1219$

In order to find percentiles of a standard normal random variable *Z*, use **Formulas >f_x (Insert Function)> Statistical>NORM.S.INV**. More specifically, suppose you want to find the 90th percentile of Z. Then click **Formulas >f_x (Insert Function)> Statistical>NORM.S.INV** and enter 0.9 for Probability.

MSL Homework: 6.1, 6.3, 6.4

HC Homework: Do the following problems.

1. Let *Z* denote the standard normal random variable. Find the following probabilities using Table E.2 in the textbook.
   $P(Z \le -4.2)$ $\;$ $P(|Z| \le 2)$ $\;$ $P(|Z| > 1)$
2. Let *Z* be the standard normal random variable. Find *z* satisfying the following using Table E.2 in the textbook.

   $P(Z \le z) = 0.9554$ $\quad$ $P(Z \le z) = 0.3085$

$P(Z \geq z) = 0.1292$ $\qquad$ $P(-z \leq Z \leq z) = 0.6046$
$P(Z \geq z) = 0.7318$ $\qquad$ $P(-1 \leq Z \leq z) = 0.1219$

To find the probability of a normal random variable *X* with mean μ and standard deviation σ , we standardize *X* so that we can use the *Z*-table. Therefore to find $P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right)$, then look up the *Z*-table. And we have the following formula.

$$x_{(p)} = \mu_x + z_{(p)}\sigma_x,$$

where $x_{(p)}$ is the *p*-th percentile of *X* and $z_{(p)}$ is the *p*-th percentile of *Z*.

In order to find $P(X \leq x)$ for a normal random variable X with mean μ and variance σ^2, use **Formulas >f_x (Insert Function)> Statistical>NORM.DIST**. More specifically, suppose you want to find $P(X \leq 26)$ for a normal random variable X with mean 24 and variance 9. Then click **Formulas >f_x (Insert Function)> Statistical>NORM.DIST** and enter 26 for *x*, 24 for Mean, 3 for Standard_dev, and 1 for Cumulative. Therefore, when you use Excel, you do not need to standardize variables.

In order to find percentiles of a normal random variable *X* with mean μ and variance σ^2, use **Formulas >f_x (Insert Function)> Statistical>NORM.INV**. More specifically, suppose you want to find the 90th percentile of a normal random variable *X* with mean 24 and variance 9. Then click **Formulas >f_x (Insert Function)> Statistical>NORM.INV** and enter 0.9 for Probability, 24 for Mean, and 3 for Standard_dev. Therefore if you use Excel, you do not need to use the formula $x_{(p)} = \mu_x + z_{(p)}\sigma_x$ that requires the percentile of *Z*.

MSL Homework: 6.5, 6.6

Example 4. At a certain university, the SAT verbal part scores of 1st year students are normally distributed with mean 565 and standard deviation 75.

a) Find the proportion of 1st year students whose SAT verbal scores are between 500 and 650.
b) How high a verbal test score must be to be among the highest 5% test scores.
c) If 5 1st year students are randomly selected, what is the probability that there will be 3 students whose scores are between 500 and 650?

Example 5. Excess returns for stocks are determined by finding the difference between the return for a stock and the returns for firms in the market that have similar levels of risk. Suppose stocks listed on an exchange have a mean monthly excess return of 0.005 and a standard deviation of 0.004 and monthly excess returns are normally distributed.

a) Find the probability that a randomly selected stock resulting in a monthly excess return less than zero
b) How high a monthly excess return should be to be among the top 25% monthly excess returns.
c) If 10 stocks are randomly selected, what is the probability that there will be more than 5 stocks with positive monthly excess returns?

MSL Homework: 6.9, 6.11
HC Homework: 6.10

The empirical rule states that if data are mound shaped approximately, for example, 95% of them are within 2 standard deviation from the mean this is because

$$P(\mu - 2\sigma \le X \le \mu + 2\sigma) = P\left(\frac{\mu - 2\sigma - \mu}{\sigma} \le \frac{X - \mu}{\sigma} \le \frac{\mu + 2\sigma - \mu}{\sigma}\right) = P(-2 \le Z \le 2)$$
$$= 0.9544$$

HC Homework: Show that if data are mound shaped approximately 68% of them are within one standard deviation from the mean.