4. Review of Basic Probability and Statistics

Outline:

4.1. Random Variables and Their Properties

4.2. Simulation Output Data and Stochastic Processes

4.3. Estimation of Means and Variances

4.4. Confidence Interval for the Mean

1

4.1. Random Variables and Their Properties

A random variable X is said to be *discrete* if it can take on at most a countable number of values, say, $x_1, x_2, \ldots$. The probability that X is equal to x_i is given by

$$p(x_i) = P(X = x_i) \text{ for } i = 1, 2, \ldots$$

and

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

2

where $p(x)$ is the *probability mass function*. The *distribution function* $F(x)$ is

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i)$$

for all $-\infty < x < \infty$.

3

Example 4.1: Consider the demand-size random variable of Section 1.5 of Law (2006) that takes on the values 1, 2, 3, 4, with probabilities 1/6, 1/3, 1/3, 1/6. The probability mass function and the distribution function are given in Figures 4.1 and 4.2.

4

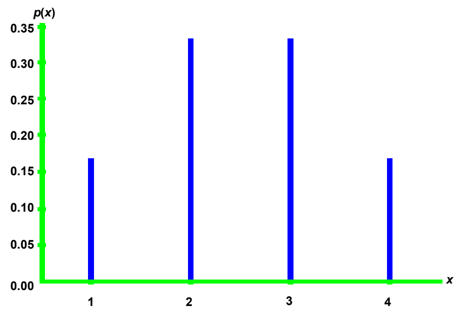

Figure 4.1. $p(x)$ for the demand-size random variable.

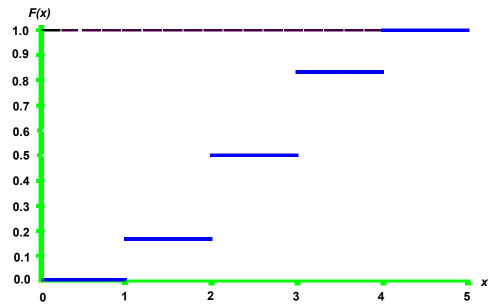

Figure 4.2. $F(x)$ for the demand-size random variable.

A random variable X is said to be ***continuous*** if there exists a nonnegative function $f(x)$, the ***probability density function***, such that for any set of real numbers B,

$$P(X \in B) = \int_B f(x)\,dx \text{ and } \int_{-\infty}^{\infty} f(x)\,dx = 1$$

(where "$\in$" means "contained in").

7

If x is a number and $\Delta x > 0$, then

$$P(X \in [x, x + \Delta x]) = \int_x^{x+\Delta x} f(y)\,dy$$

which is the left shaded area in Figure 4.3.

8

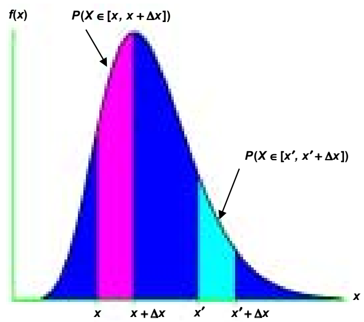

Figure 4.3. Interpretation of the probability density function *f*(*x*).

The distribution function *F*(*x*) for a continuous random variable *X* is

$$F(x) = P(X \le x) = P(X \in (-\infty, x])$$

$$= \int_{-\infty}^{x} f(y)\, dy \quad \text{for all } -\infty < x < \infty$$

10

Example 4.2: The probability density function and distribution function for an *exponential random variable* with mean β are defined as follows (see Figures 4.4 and 4.5):

$$f(x) = \frac{1}{\beta} e^{-x/\beta} \quad \text{for } x \ge 0$$

and

$$F(x) = 1 - e^{-x/\beta} \quad \text{for } x \ge 0$$

11

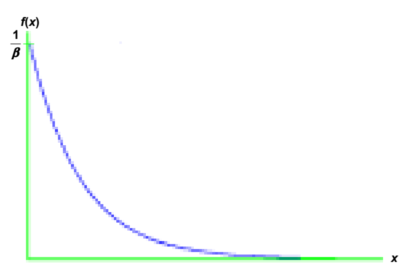

Figure 4.4. *f*(*x*) for an exponential random variable with mean β.

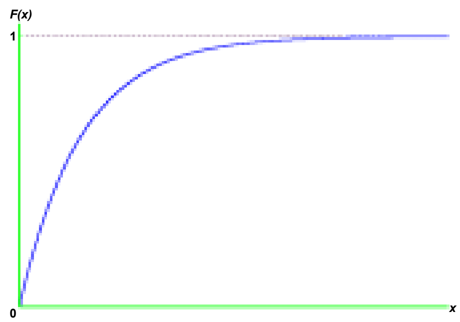

Figure 4.5. $F(x)$ for an exponential random variable with mean β.

The random variables X and Y are ***independent*** if knowing the value that one takes on tells us nothing about the distribution of the other.

The ***mean*** or ***expected value*** of the random variable X, denoted by μ or $E(X)$, is given by

$$\mu = \begin{cases} \sum_{i=1}^{\infty} x_i\, p(x_i) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x f(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

14

The mean is one measure of the central tendency of a random variable.

Problem 4.1: What are other measures?

Properties:

1. $E(cX) = cE(X)$, where c is a constant
2. $E(X + Y) = E(X) + E(Y)$ regardless of whether X and Y are independent

15

The ***variance*** of the random variable X, denoted by σ^2 or Var(X), is given by

$$\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2$$

The variance is a measure of the dispersion of a random variable about its mean (see Figure 4.6).

16

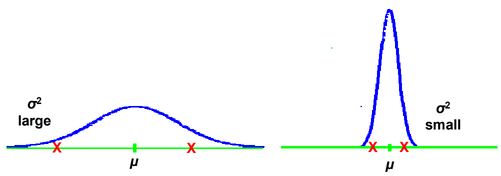

Figure 4.6. Density functions for continuous random variables with large and small variances.

Properties:

1. $Var(cX) = c^2 Var(X)$
2. $Var(X + Y) = Var(X) + Var(Y)$
   if X, Y are independent

18

The square root of the variance is called the ***standard deviation*** and is denoted by σ. It can be given the most definitive interpretation when X has a normal distribution (see Figure 4.7).

19

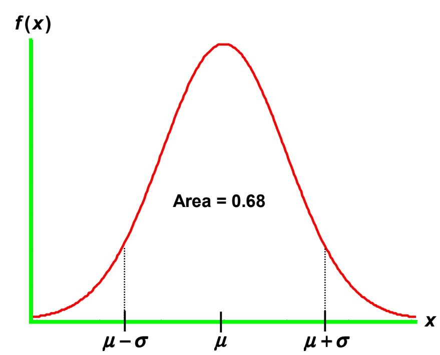

Figure 4.7. Density function for a N(μ, σ^2) distribution.

The ***covariance*** between the random variables X and Y, denoted by Cov(X, Y), is defined by

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$$
$$= E(XY) - E(X)E(Y)$$

The covariance is a measure of the dependence between X and Y. Note that Cov(X, X) = Var(X).

Definitions:

Cov(X, Y)	X and Y are
= 0	*uncorrelated*
> 0	*positively correlated*
< 0	*negatively correlated*

Independent random variables are also uncorrelated.

Note that, in general, we have

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)$$

If X and Y are independent, then

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$$

The ***correlation*** between the random variables X and Y, which is a measure of linear dependence (see next slide), is denoted by Cor(X, Y) and defined by

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}}$$

It can be shown that

$$-1 \le \text{Cor}(X, Y) \le 1$$

Suppose that $Y = aX + b$, where a and b are constants. Then

$$\text{Cor}(X,Y) = \begin{cases} 1 & \text{if } a > 0 \\ -1 & \text{if } a < 0 \end{cases}$$

25

4.2. Simulation Output Data and Stochastic Processes

A ***stochastic process*** is a collection of "similar" random variables ordered over time all defined relative to the same experiment. If the collection is X_1, X_2, ... , then we have a ***discrete-time*** stochastic process.

26

If the collection is $\{X(t),\ t \geq 0\}$, then we have a ***continuous-time*** stochastic process.

27

Example 4.3:

For the single-server queueing system of Chapter 1, assume the following:

- The A_i's are independent and identically distributed (IID)
- The P_i's are IID
- The A_i's and P_i's are independent

28

Relative to the experiment of generating the A_i's and P_i's, one can define the discrete-time stochastic process of delays in queue D_1, D_2, ... as follows:

$D_1 = 0$

$D_{i+1} = \max\{D_i + P_i - A_{i+1}, 0\}$ for $i = 1, 2, \ldots$

29

Thus, the simulation maps the input random variables into the output process of interest.

Problem 4.2: Are D_i and D_{i+1} independent, positively correlated, or negatively correlated?

30

Other examples of stochastic processes:

- N_1, N_2, ... , where N_i = number of parts produced in the ith hour for a manufacturing system
- T_1, T_2, ... , where T_i = time in system of the ith part for a manufacturing system
- $\{Q(t), t \geq 0\}$, where $Q(t)$ = number of customers in queue at time t

31

- C_1, C_2, ... , where C_i = total cost in the ith month for an inventory system
- E_1, E_2, ... , where E_i = end-to-end delay of ith message to reach its destination in a communications network

32

Example 4.4: Consider the delay-in-queue process $D_1, D_2, \ldots$ for the $M/M/1$ queue with utilization factor ρ. Then the correlation function ρ_j between D_i and D_{i+j} is given in Figure 4.8.

33

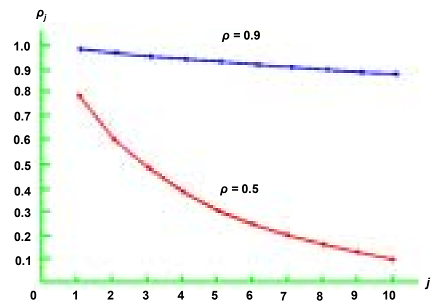

Figure 4.8. Correlation function ρ_j of the process $D_1, D_2, \ldots$ for the $M/M/1$ queue.

4.3. Estimation of Means and Variances

Let $X_1, X_2, \ldots, X_n$ be IID random variables with population mean and variance μ and σ^2, respectively.

35

Population parameter	Sample estimate	
μ	Sample mean $\bar{X}(n) = \dfrac{\sum_{i=1}^{n} X_i}{n}$	(1)
σ^2	Sample variance $S^2(n) = \dfrac{\sum_{i=1}^{n} [X_i - \bar{X}(n)]^2}{n-1}$	(3)
$\text{Var}[\bar{X}(n)] = \dfrac{\sigma^2}{n}$ (4)	$\hat{Var}[\bar{X}(n)] = \dfrac{S^2(n)}{n}$	(5)

Note that $\bar{X}(n)$ is an *unbiased estimator* of μ, i.e., $E[\bar{X}(n)] = E(X) = \mu$. (2)

Problem 4.3: Show that $\bar{X}(n)$ is an unbiased estimator of μ.

The difficulty with using $\bar{X}(n)$ as an estimator of μ without any additional information is that we have no way of assessing how close $\bar{X}(n)$ is to μ.

37

Because $\bar{X}(n)$ is a random variable with variance Var[$\bar{X}(n)$], on one experiment it may be close to μ while on another it may differ from μ by a large amount (see Figure 4.9).

The usual way to access the precision of $\bar{X}(n)$ as an estimator of μ is to construct a confidence interval for μ, which we discuss in the next section.

38

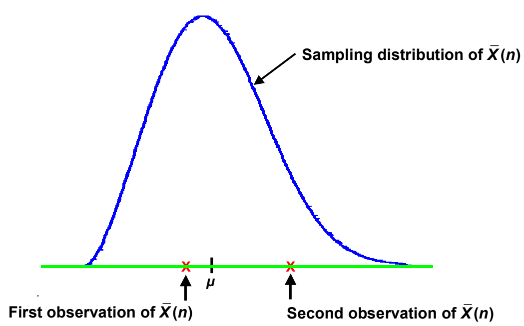

Figure 4.9. Two observations of the random variable $\bar{X}(n)$.

Example 4.5: Consider the bank with 5 tellers on p. 486-487 of Law. The following are the average delays in queue resulting from 10 independent replications of the simulation model:

1.53, 1.66, 1.24, …, 2.60

Since these observations are IID, they can be plugged into (1) through (5).

40

However, the delays in queue from one particular replication are not independent.

41

4.4. Confidence Interval for the Mean

Let $X_1, X_2, ..., X_n$ be IID random variables with mean μ. Then an (approximate) $100(1 - \alpha)$ percent ($0 < \alpha < 1$) *confidence interval* for μ is

$$\overline{X}(n) \pm t_{n-1,1-\alpha/2} \sqrt{S^2(n)/n} \qquad (6)$$

42

where $t_{n-1,\,1-\alpha/2}$ is the upper $1 - \alpha/2$ critical point for a t distribution with $n - 1$ df (see Figure 4.10).

43

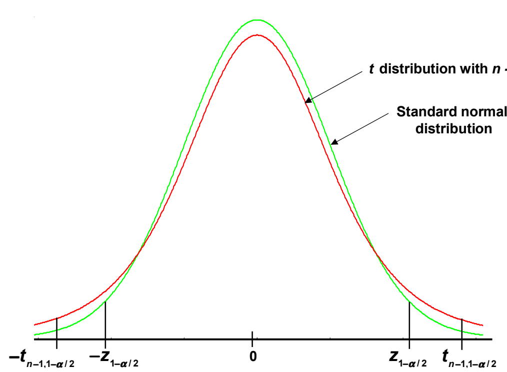

Figure 4.10. Standard normal distribution and *t* distribution with *n* - 1 df.

Notes:

- $t_{n-1, 1-\alpha/2} > z_{1-\alpha/2}$ for $n \geq 2$.
- $t_{n-1, 1-\alpha/2}$ decreases to $z_{1-\alpha/2}$ as n gets larger.
- $t_{n-1, 1-\alpha/2} \approx z_{1-\alpha/2}$ for $n = 50$
- As α gets smaller, the confidence interval half-length gets larger.

45

Interpretation of a confidence interval:

If one constructs a very large number of independent 100(1 - α) percent confidence intervals for μ each based on n observations, where n is sufficiently large, then the proportion of these confidence intervals that contain μ should be 1 - α (regardless of the distribution of X).

46

Alternatively, if X is $N(\mu,\sigma^2)$, then the coverage probability will be 1- α regardless of the value of n. If X is not $N(\mu,\sigma^2)$, then there will be a degradation in coverage for "small" n. The greater the skewness of the distribution of X, the greater the degradation (see pp. 256-257).

47

We used $t_{n-1, 1-\alpha/2}$ rather than $z_{1-\alpha/2}$ in (6) to help lessen the effect of skewness in the distribution of X and of "small" n.

48

Important characteristics of a confidence interval:

- **Confidence level (e.g., 90 percent)**
- **Half-length (see also p. 511)**

Problem 4.4: If we want to decrease the half-length by a factor of approximately 2 and *n* is "large" (e.g., 50), then to what value does *n* need to be increased?

49

Recommended reading

Chapter 4 in Law (2006)

50