Chapter 10
OTHER PROBABILITY DISTRIBUTIONS

IN THIS CHAPTER:

✔ *The Multinomial Distribution*
✔ *The Hypergeometric Distribution*
✔ *The Uniform Distribution*
✔ *The Cauchy Distribution*
✔ *The Gamma Distribution*
✔ *The Beta Distribution*
✔ *The Chi-Square Distribution*
✔ *Student's t Distribution*
✔ *The F Distribution*
✔ *Relationships Among Chi-Square, t, and F Distributions*

The Multinomial Distribution

Suppose that events $A_1, A_2, \ldots, A_k$ are mutually exclusive, and can occur with respective probabilities $p_1, p_2, \ldots, p_k$ where $p_1 + p_2 + \cdots + p_k + 1$. If $X_1, X_2, \ldots, X_k$ are the random variables, respectively, giving the number of times that $A_1, A_2, \ldots, A_k$ occur in a total of n trials, so that $X_1 + X_2 + \cdots X_k = n$, then

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) = \frac{n}{n_1! n_2! \cdots n_k!} p_1^{n_1} p_2^{n_k} \cdots p_k^{n_k} \qquad (1)$$

where $n_1 + n_2 + \cdots n_k = n$, is the joint probability function for the random variables $X_1, X_2, \ldots, X_k$.

This distribution, which is a generalization of the binomial distribution, is called the *multinomial distribution* since the equation above is the general term in the multinomial expansion of $(p_1 + p_2 + \cdots p_k)^n$.

The Hypergeometric Distribution

Suppose that a box contains b blue marbles and r red marbles. Let us perform n trials of an experiment in which a marble is chosen at random, its color observed, and then the marble is put back in the box. This type of experiment is often referred to as *sampling with replacement*. In such a case, if X is the random variable denoting the number of blue marbles chosen (successes) in n trials, then using the binomial distribution we see that the probability of exactly x successes is

$$P(X = x) = \binom{n}{x} \frac{b^x r^{n-x}}{(b+r)^n}, \qquad x = 0, 1, \ldots, n \qquad (2)$$

since $p = b/(b + r)$, $q = 1 - p = r/(b + r)$.

If we modify the above so that *sampling is without replacement*, i.e., the marbles are not replaced after being chosen, then

$$P(X = x) = \frac{\binom{b}{x}\binom{r}{n-x}}{\binom{b+r}{n}}, \quad x = \max(0,\, n - r), \ldots, \min(n, b) \qquad (3)$$

This is the *hypergeometric distribution.* The mean and variance for this distribution are

$$\mu = \frac{nb}{b+r}, \qquad \sigma^2 = \frac{nbr(b+r-n)}{(b+r)^2(b+r-1)} \qquad (4)$$

If we let the total number of blue and red marbles be N, while the proportions of blue and red marbles are p and $q = 1 - p$, respectively, then

$$p = \frac{b}{b+r} = \frac{b}{N}, \qquad q = \frac{r}{b+r} = \frac{r}{N} \quad \text{or} \quad b - Np, \quad r = Nq$$

This leads us to the following

$$P(X = x) = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} \qquad (5)$$

$$\mu = np, \qquad \sigma^2 = \frac{npq(N-n)}{N-1} \qquad (6)$$

Note that as $N \to \infty$ (or N is large when compared with n), these two formulas reduce to the following

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \tag{7}$$

$$\mu = np, \qquad \sigma^2 = npq \tag{8}$$

Notice that this is the same as the mean and variance for the binomial distribution. The results are just what we would expect, since for large N, sampling without replacement is practically identical to sampling with replacement.

Example 10.1 A box contains 6 blue marbles and 4 red marbles. An experiment is performed in which a marble is chosen at random and its color is observed, but the marble is not replaced. Find the probability that after 5 trials of the experiment, 3 blue marbles will have been chosen.

The number of different ways of selecting 3 blue marbles out of 6 marbles is $\binom{6}{3}$. The number of different ways of selecting the remaining 2 marbles out of the 4 red marbles is $\binom{4}{2}$. Therefore, the number of different samples containing 3 blue marbles and 2 red marbles is $\binom{6}{3}\binom{4}{2}$.

Now the total number of different ways of selecting 5 marbles out of the 10 marbles (6 + 4) in the box is $\binom{10}{5}$. Therefore, the required probability is given by

$$\frac{\binom{6}{3}\binom{4}{2}}{\binom{10}{5}} = \frac{10}{21}$$

The Uniform Distribution

A random variable X is said to be *uniformly distributed* in $a \le x \le b$ if its density function is

$$f(x) = \begin{cases} 1/(b-a) & a \le x \le b \\ 0 & \text{otherwise} \end{cases} \qquad (9)$$

and the distribution is called a *uniform distribution*.

The distribution function is given by

$$F(x) = P(X \le x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x < b \\ 1 & x \ge b \end{cases} \qquad (10)$$

The mean and variance are, respectively

$$\mu = \frac{1}{2}(a+b), \qquad \sigma^2 = \frac{1}{12}(b-a)^2 \qquad (11)$$

The Cauchy Distribution

A random variable X is said to be *Cauchy distributed*, or to have the *Cauchy distribution*, if the density function of X is

$$f(x) = \frac{a}{\pi(x^2 + a^2)} \qquad a > 0, -\infty < x < \infty \tag{12}$$

The density function is symmetrical about $x = 0$ so that its median is zero. However, the mean and variance do not exist.

The Gamma Distribution

A random variable X is said to have the *gamma distribution*, or to be *gamma distributed*, if the density function is

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} & x > 0 \\ 0 & x \le 0 \end{cases} \qquad (\alpha, \beta > 0) \tag{13}$$

where $\Gamma(\alpha)$ is the *gamma function* (see Appendix A). The mean and variance are given by

$$\mu = \alpha\beta \qquad \sigma^2 = \alpha\beta^2 \tag{14}$$

The Beta Distribution

A random variable is said to have the *beta distribution*, or to be *beta distributed*, if the density function is

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} & 0 < x < 1 \\ 0 & \textit{otherwise} \end{cases} \qquad (\alpha, \beta > 0) \tag{15}$$

where $B(\alpha, \beta)$ is the *beta function* (see Appendix A). In view of the relation between the beta and gamma functions, the beta distribution can also be defined by the density function

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \textit{otherwise} \end{cases} \tag{16}$$

where α, β are positive. The mean and variance are

$$\mu = \frac{\alpha}{\alpha+\beta}, \quad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{17}$$

For $\alpha > 1$, $\beta > 1$ there is a unique mode at the value

$$x_{\text{mode}} = \frac{\alpha-1}{\alpha+\beta-2} \tag{18}$$

The Chi-Square Distribution

Let $X_1, X_2, \ldots, X_v$ be v independent normally distributed random variables with mean zero and variance one. Consider the random variable

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_v^2 \tag{19}$$

where χ^2 is called *chi square*. Then we can show that for all $x \geq 0$,

$$P(\chi^2 \le x) = \frac{1}{2^{v/2}\Gamma(v/2)} \int_0^x u^{(v/2)-1} e^{-u/2}\, du \qquad (20)$$

and $P(\chi^2 \le x) = 0$ for $x > 0$.

The distribution above is called the *chi-square distribution*, and v is called the *number of degrees of freedom*. The distribution defined above has corresponding density function given by

$$f(x) = \begin{cases} \dfrac{1}{2^{v/2}\Gamma(v/2)} x^{(v/2)-1} e^{-x/2} & x > 0 \\ 0 & x \le 0 \end{cases} \qquad (21)$$

It is seen that the chi-square distribution is a special case of the gamma distribution with $\alpha = v/2$ and $\beta = 2$. Therefore,

$$\mu = v, \qquad \sigma^2 = 2v \qquad (22)$$

For large v ($v \ge 30$), we can show that $\sqrt{2\chi^2} - \sqrt{2v-1}$ is very nearly normally distributed with mean 0 and variance one.

Three theorems that will be useful in later work are as follows:

Theorem 10-1: Let $X_1, X_2, \ldots, X_v$ be independent normally random variables with mean 0 and variance 1. Then $\chi^2 = X_1^2 + X_2^2 + \cdots + X_v^2$ is chi square distributed with v degrees of freedom.

Theorem 10-2: Let $U_1, U_2, \ldots, U_k$ be independent random variables that are chi square distributed with $v_1, v_2, \ldots, v_k$ degrees of freedom, respectively. Then their sum $W = U_1 + U_2 + \cdots U_k$ is chi square distributed with $v_1 + v_2 + \cdots v_k$ degrees of freedom.

Theorem 10-3: Let V_1 and V_2 be independent random variables. Suppose that V_1 is chi square distributed with v_1 degrees of freedom while $V = V_1 = V_2$ is chi square distributed with v degrees of freedom, where $v > v_1$. Then V_2 is chi square distributed with $v - v_1$ degrees of freedom.

In connection with the chi-square distribution, the *t* distribution, the *F* distribution, and others, it is common in statistical work to use the *same symbol* for both the random variable and a value of the random variable. Therefore, percentile values of the chi-square distribution for v degrees of freedom are denoted by $\chi^2_{p,v}$, or briefly χ^2_p if v is understood, and not by $\chi^2_{p,v}$ or x_p. (See Appendix D.) This is an ambiguous notation, and the reader should use care with it, especially when changing variables in density functions.

Example 10.2. The graph of the chi-square distribution with 5 degrees of freedom is shown in Figure 10-1. Find the values for χ^2_1, χ^2_2 for which the shaded area on the right = 0.05 and the total shaded area = 0.05.

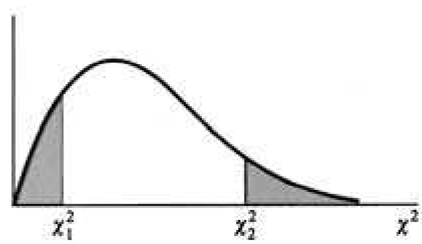

Figure 10-1

If the shaded area on the right is 0.05, then the area to the left of χ_2^2 is $(1 - 0.05) = 0.95$, and χ_2^2 represents the 95th percentile, $\chi_{0.95}^2$.

Referring to the table in Appendix D, proceed downward under the column headed v until entry 5 is reached. Then proceed right to the column headed $\chi_{0.95}^2$. The result, 11.1, is the required value of χ^2.

Secondly, since the distribution is not symmetric, there are many values for which the total shaded area = 0.05. For example, the right-handed shaded area could be 0.04 while the left-handed area is 0.01. It is customary, however, unless otherwise specified, to choose the two areas equal. In this case, then, each area = 0.025.

If the shaded area on the right is 0.025, the area to the left of χ_2^2 is $1 - 0.025 = 0.975$ and χ_2^2 represents the 97.5th percentile $\chi_{0.975}^2$, which from Appendix D is 12.8.

Similarly, if the shaded area on the left is 0.025, the area to the left of χ_1^2 is 0.025 and χ_1^2 represents the 2.5th percentile, $\chi_{0.025}^2$, which equals 0.831.

Therefore, the values are 0.831 and 12.8.

Student's *t* Distribution

If a random variable has the density function

$$f(t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{t^2}{v}\right)^{-(v+1)/2} \qquad -\infty < t < \infty \qquad (23)$$

it is said to have the *Student's t distribution*, briefly the *t distribution*, with v degrees of freedom. If v is large ($v \geq 30$), the graph of $f(t)$ closely approximates the normal curve, as indicated in Figure 10-2.

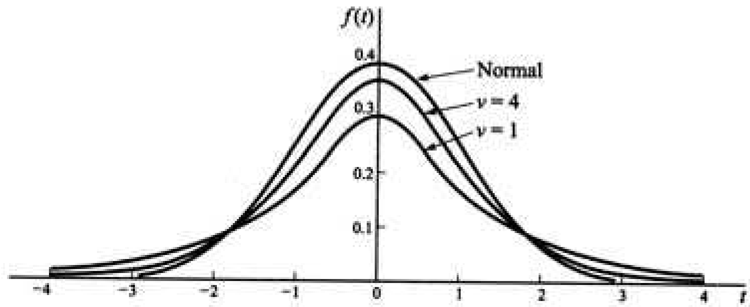

Figure 10-2

Percentile values of the t distribution for v degrees of freedom are denoted by $t_{p,v}$ or briefly t_p if v is understood. For a table giving such values, see Appendix C. Since the t distribution is symmetrical, $t_{1-p} = -t_p$; for example, $t_{0.5} = -t_{0.95}$.

For the t distribution we have

$$\mu = 0 \qquad \text{and} \qquad \sigma^2 = \frac{v}{v-2} \quad (v > 2) \qquad (24)$$

The following theorem is important in later work.

Theorem 10-4: Let Y and Z be independent random variables, where Y is normally distributed with mean 0 and variance 1 while Z is chi square distributed with v degrees of freedom. Then the random variable

$$T = \frac{Y}{\sqrt{Z/v}} \qquad (25)$$

has the t distribution with v degrees of freedom.

Example 10.3. The graph of Student's *t distribution* with 9 degrees of freedom is shown in Figure 10-3. Find the value of t_1 for which the shaded area on the right = 0.05 and the total unshaded area = 0.99.

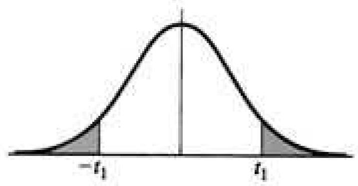

Figure 10-3

If the shaded area on the right is 0.05, then the area to the left of t_1 is (1 – 0.05) = 0.095, and t_1 represents the 95th percentile, $t_{0.95}$. Referring to the table in Appendix C, proceed downward under the column headed v until entry 9 is reached. Then proceed right to the column headed $t_{0.95}$. The result 1.83 is the required value of t.

Next, if the total unshaded area is 0.99, then the total shaded area is $(1 - 0.99) = 0.01$, and the shaded area to the right is $0.01 / 2 = 0.005$. From the table we find $t_{0.995} = 3.25$.

The *F* Distribution

A random variable is said to have the *F* distribution (named after R. A. Fisher) with v_1 and v_2 *degrees of freedom* if its density function is given by

$$f(u) = \begin{cases} \dfrac{\Gamma\left(\dfrac{v_1+v_2}{2}\right)}{\Gamma\left(\dfrac{v_1}{2}\right)\Gamma\left(\dfrac{v_2}{2}\right)} v_1^{v_1/2} v_2^{v_2/2} u^{(v_1/2)-1}(v_2+v_1u)^{-(v_1+v_2)/2} & u > 0 \\ 0 & u \le 0 \end{cases} \tag{26}$$

Percentile values of the F distribution for v_1, v_2 degrees of freedom are denoted by $F_{p,v1,v2}$, or briefly F_p if v_1 and v_2 are understood.

For a table giving such values in the case where p = 0.95 and p = 0.99, see Appendix E.

The mean and variance are given, respectively, by

$$\mu = \frac{v_2}{v_2 - 2} \quad (v_2 > 2) \qquad \text{and} \qquad \sigma^2 = \frac{2v_2^2(v_1 + v_2 + 2)}{v_1(v_2 - 4)(v_2 - 2)^2} \tag{27}$$

The distribution has a unique mode at the value

$$u_{\text{mode}} = \left(\frac{v_1 - 2}{v_1}\right)\left(\frac{v_2}{v_2 + 2}\right) \quad (v_1 > 2) \tag{28}$$

The following theorems are important in later work.

Theorem 11-5: Let V_1 and V_2 be independent random variables that are chi square distributed with v_1 and v_2 degrees of freedom, respectively. Then the random variable

$$V = \frac{V_1 / v_1}{V_2 / v_2} \tag{29}$$

has the F distribution with v_1 and v_2 degrees of freedom.

Theorem 10-6:

$$F_{1-p.v_2,v_1} = \frac{1}{F_{p,v_1,v_2}} \tag{30}$$

Remember

While specially used with small samples, Student's *t* distribution, the chi-square distribution, and the *F* distribution are all valid for large sample sizes as well.

Relationships Among Chi-Square, *t*, and *F* Distributions

Theorem 10-7: $$F_{1-p,1,v} = t^2_{1-(p/2),v} \tag{31}$$

Theorem 10-8: $$F_{p,v,\infty} = \frac{\chi^2_{p,v}}{v} \tag{32}$$

Example 10.4. Verify Theorem 10-7 by showing that $F_{0.95} = t^2_{0.975}$.

Compare the entries in the first column of the $F_{0.95}$ table in Appendix E with those in the *t* distribution under $t_{0.975}$. We see that

$$161 = (12.71)^2,\quad 18.5 = (4.30)^2,\quad 10.1 = (3.18)^2,\quad 7.71 = (2.78)^2,\text{ etc.,}$$

which provides the required verification.

Example 10.5. Verify Theorem 10-8 for $p = 0.99$.

Compare the entries in the last row of the $F_{0.99}$ table in Appendix E (corresponding to $v_2 = \infty$) with the entries under $\chi^2_{0.99}$ in Appendix D. Then we see that

$$6.63 = \frac{6.63}{1}, \quad 4.61 = \frac{9.21}{2}, \quad 3.78 = \frac{11.3}{3}, \quad 3.32 = \frac{13.3}{4}, \text{ etc.,}$$

which provides the required verification.