CHAPTER 11

Small Sampling Theory

SMALL SAMPLES

In previous chapters we often made use of the fact that for samples of size $N > 30$, called *large samples*, the sampling distributions of many statistics are approximately normal, the approximation becoming better with increasing N. For samples of size $N < 30$, called *small samples*, this approximation is not good and becomes worse with decreasing N, so that appropriate modifications must be made.

A study of sampling distributions of statistics for small samples is called *small sampling theory*. However, a more suitable name would be *exact sampling theory*, since the results obtained hold for large as well as for small samples. In this chapter we study three important distributions: Student's t distribution, the chi-square distribution, and the F distribution.

STUDENT'S t DISTRIBUTION

Let us define the statistic

$$t = \frac{\bar{X} - \mu}{s}\sqrt{N-1} = \frac{\bar{X} - \mu}{\hat{s}/\sqrt{N}} \qquad (1)$$

which is analogous to the z statistic given by

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}}.$$

If we consider samples of size N drawn from a normal (or approximately normal) population with mean μ and if for each sample we compute t, using the sample mean $\bar{X}$ and sample standard deviation s or $\hat{s}$, the sampling distribution for t can be obtained. This distribution (see Fig. 11-1) is given by

$$Y = \frac{Y_0}{\left(1 + \dfrac{t^2}{N-1}\right)^{N/2}} = \frac{Y_0}{\left(1 + \dfrac{t^2}{\nu}\right)^{(\nu+1)/2}} \qquad (2)$$

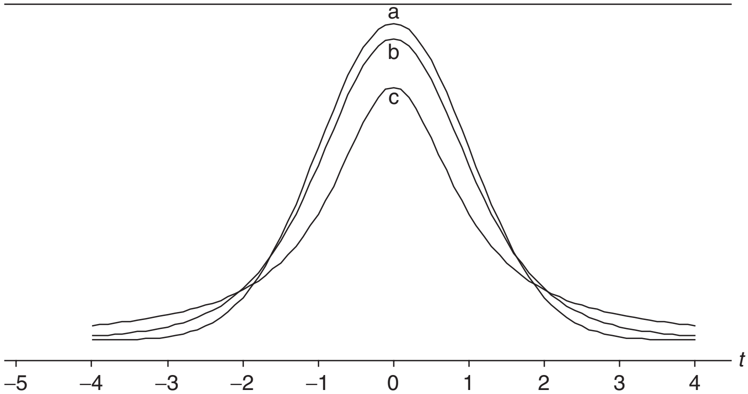

Fig. 11-1 (*a*) Standard normal, (*b*) Student t with $\nu = 5$, (*c*) Student t with $\nu = 1$.

where Y_0 is a constant depending on N such that the total area under the curve is 1, and where the constant $\nu = (N-1)$ is called the *number of degrees of freedom* (ν is the Greek letter *nu*).

Distribution (2) is called *Student's t distribution* after its discoverer, W. S. Gossett, who published his works under the pseudonym "Student" during the early part of the twentieth century.

For large values of ν or N (certainly $N \geq 30$) the curves (*2*) closely approximate the standardized normal curve

$$Y = \frac{1}{\sqrt{2\pi}} e^{-(1/2)t^2}$$

as shown in Fig. 11-1.

CONFIDENCE INTERVALS

As done with normal distributions in Chapter 9, we can define 95%, 99%, or other confidence intervals by using the table of the t distribution in Appendix III. In this manner we can estimate within specified limits of confidence the population mean μ.

For example, if $-t_{.975}$ and $t_{.975}$ are the values of t for which 2.5% of the area lies in each tail of the t distribution, then the 95% confidence interval for t is

$$-t_{.975} < \frac{\bar{X} - \mu}{s}\sqrt{N-1} < t_{.975} \tag{3}$$

from which we see that μ is estimated to lie in the interval

$$\bar{X} - t_{.975}\frac{s}{\sqrt{N-1}} < \mu < \bar{X} + t_{.975}\frac{s}{\sqrt{N-1}} \tag{4}$$

with 95% confidence (i.e. probability 0.95). Note that $t_{.975}$ represents the 97.5 percentile value, while $t_{.025} = -t_{.975}$ represents the 2.5 percentile value.

In general, we can represent confidence limits for population means by

$$\bar{X} \pm t_c \frac{s}{\sqrt{N-1}} \tag{5}$$

where the values $\pm t_c$, called *critical values* or *confidence coefficients*, depend on the level of confidence desired and on the sample size. They can be read from Appendix III.

The sample is assumed to be taken from a normal population. This assumption may be checked out using the Komogorov–Smirnov test for normality.

A comparison of equation (5) with the confidence limits ($\bar{X} \pm z_c\sigma/\sqrt{N}$) of Chapter 9, shows that for small samples we replace z_c (obtained from the normal distribution) with t_c (obtained from the t distribution) and that we replace σ with $\sqrt{N/(N-1)}s = \hat{s}$, which is the sample estimate of σ. As N increases, both methods tend toward agreement.

TESTS OF HYPOTHESES AND SIGNIFICANCE

Tests of hypotheses and significance, or decision rules (as discussed in Chapter 10), are easily extended to problems involving small samples, the only difference being that the *z score*, or *z statistic*, is replaced by a suitable *t score*, or *t statistic*.

1. **Means.** To test the hypothesis H_0 that a normal population has mean μ, we use the t score (or t statistic)

$$t = \frac{\bar{X}-\mu}{s}\sqrt{N-1} = \frac{\bar{X}-\mu}{\hat{s}}\sqrt{N} \tag{6}$$

where $\bar{X}$ is the mean of a sample of size N. This is analogous to using the z score

$$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{N}}$$

for large N, except that $\hat{s} = \sqrt{N/(N-1)}s$ is used in place of σ. The difference is that while z is normally distributed, t follows Student's distribution. As N increases, these tend toward agreement.

2. **Differences of Means.** Suppose that two random samples of sizes N_1 and N_2 are drawn from normal populations whose standard deviations are equal ($\sigma_1 = \sigma_2$). Suppose further that these two samples have means given by $\bar{X}_1$ and $\bar{X}_2$ and standard deviations given by s_1 and s_2, respectively. To test the hypothesis H_0 that the samples come from the same population (i.e., $\mu_1 = \mu_2$ as well as $\sigma_1 = \sigma_2$), we use the t score given by

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{1/N_1 + 1/N_2}} \qquad \text{where} \qquad \sigma = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}} \tag{7}$$

The distribution of t is Student's distribution with $\nu = N_1 + N_2 - 2$ degrees of freedom. The use of equation (7) is made plausible on placing $\sigma_1 = \sigma_2 = \sigma$ in the z score of equation (2) of Chapter 10 and then using as an estimate of σ^2 the weighted mean

$$\frac{(N_1-1)\hat{s}_1^2 + (N_2-1)\hat{s}_2^2}{(N_1-1)+(N_2-1)} = \frac{N_1 s_1^2 + N_2 s_2^2}{N_1+N_2-2}$$

where $\hat{s}_1^2$ and $\hat{s}_2^2$ are the unbiased estimates of σ_1^2 and σ_2^2.

THE CHI-SQUARE DISTRIBUTION

Let us define the statistic

$$\chi^2 = \frac{Ns^2}{\sigma^2} = \frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \cdots + (X_N-\bar{X})^2}{\sigma^2} \tag{8}$$

where χ is the Greek letter *chi* and χ^2 is read "chi-square."

If we consider samples of size N drawn from a normal population with standard deviation σ, and if for each sample we compute χ^2, a sampling distribution for χ^2 can be obtained. This distribution, called the *chi-square distribution*, is given by

$$Y = Y_0(\chi^2)^{(1/2)(\nu-2)}\, e^{-(1/2)\chi^2} = Y_0\chi^{\nu-2}\, e^{-(1/2)\chi^2} \tag{9}$$

where $\nu = N - 1$ is the *number of degrees of freedom*, and Y_0 is a constant depending on ν such that the total area under the curve is 1. The chi-square distributions corresponding to various values of ν are shown in Fig. 11-2. The maximum value of Y occurs at $\chi^2 = \nu - 2$ for $\nu \geq 2$.

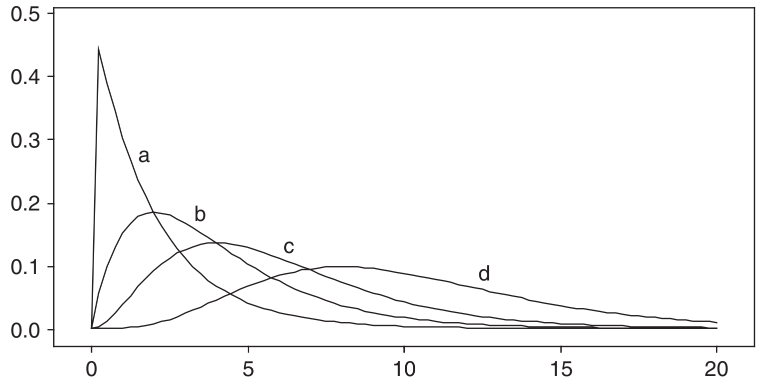

Fig. 11-2 Chi-square distributions with (a) 2, (b) 4, (c) 6, and (d) 10 degrees of freedom.

CONFIDENCE INTERVALS FOR σ

As done with the normal and t distribution, we can define 95%, 99%, or other confidence limits by using the table of the χ^2 distribution in Appendix IV. In this manner we can estimate within specified limits of confidence the population standard deviation σ in terms of a sample standard deviation s.

For example, if $\chi^2_{.025}$ and $X^2_{.975}$ are the values of χ^2 (called *critical values*) for which 2.5% of the area lies in each tail of the distribution, then the 95% confidence interval is

$$\chi^2_{.025} < \frac{Ns^2}{\sigma^2} < \chi^2_{.975} \tag{10}$$

from which we see that σ is estimated to lie in the interval

$$\frac{s\sqrt{N}}{\chi_{.975}} < \sigma < \frac{s\sqrt{N}}{\chi_{.025}} \tag{11}$$

with 95% confidence. Other confidence intervals can be found similarly. The values $\chi_{.025}$ and $\chi_{.095}$ represent, respectively, the 2.5 and 97.5 percentile values.

Appendix IV gives percentile values corresponding to the number of degrees of freedom ν. For large values of ν ($\nu \geq 30$), we can use the fact that $(\sqrt{2\chi^2} - \sqrt{2\nu - 1})$ is very nearly normally distributed with mean 0 and standard deviation 1; thus normal distribution tables can be used if $\nu \geq 30$. Then if χ^2_p and z_p are the pth percentiles of the chi-square and normal distributions, respectively, we have

$$\chi^2_p = \tfrac{1}{2}(z_p + \sqrt{2\nu - 1})^2 \tag{12}$$

In these cases, agreement is close to the results obtained in Chapters 8 and 9.

For further applications of the chi-square distribution, see Chapter 12.

DEGREES OF FREEDOM

In order to compute a statistic such as (*1*) or (*8*), it is necessary to use observations obtained from a sample as well as certain population parameters. If these parameters are unknown, they must be estimated from the sample.

The *number of degrees of freedom* of a statistic, generally denoted by ν, is defined as the number N of independent observations in the sample (i.e., the sample size) minus the number k of population parameters, which must be estimated from sample observations. In symbols, $\nu = N - k$.

In the case of statistic (*1*), the number of independent observations in the sample is N, from which we can compute $\bar{X}$ and s. However, since we must estimate μ, $k = 1$ and so $\nu = N - 1$.

In the case of statistic (*8*), the number of independent observations in the sample is N, from which we can compute s. However, since we must estimate σ, $k = 1$ and so $\nu = N - 1$.

THE *F* DISTRIBUTION

As we have seen, it is important in some applications to know the sampling distribution of the difference in means $(\bar{X}_1 - \bar{X}_2)$ of two samples. Similarly, we may need the sampling distribution of the difference in variances $(S_1^2 - S_2^2)$. It turns out, however, that this distribution is rather complicated. Because of this, we consider instead the statistic S_1^2/S_2^2, since a large or small ratio would indicate a large difference, while a ratio nearly equal to 1 would indicate a small difference. The sampling distribution in such a case can be found and is called the *F distribution*, named after R. A. Fisher.

More precisely, suppose that we have two samples, 1 and 2, of sizes N_1 and N_2, respectively, drawn from two normal (or nearly normal) populations having variances σ_1^2 and σ_2^2. Let us define the statistic

$$F = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} = \frac{N_1 S_1^2/(N_1 - 1)\sigma_1^2}{N_2 S_2^2/(N_2 - 1)\sigma_2^2} \tag{13}$$

where

$$\hat{S}_1^2 = \frac{N_1 S_1^2}{N_1 - 1} \qquad \hat{S}_2^2 = \frac{N_2 S_2^2}{N_2 - 1}. \tag{14}$$

Then the sampling distribution of F is called Fisher's F distribution, or briefly the F distribution, with $\nu_1 = N_1 - 1$ and $\nu_2 = N_2 - 1$ degrees of freedom. This distribution is given by

$$Y = \frac{CF^{(\nu_1/2)-1}}{(\nu_1 F + \nu_2)^{(\nu_1+\nu_2)/2}} \tag{15}$$

where C is a constant depending on ν_1 and ν_2 such that the total area under the curve is 1. The curve has a shape similar to that shown in Fig. 11-3, although this shape can vary considerably for different values of ν_1 and ν_2.

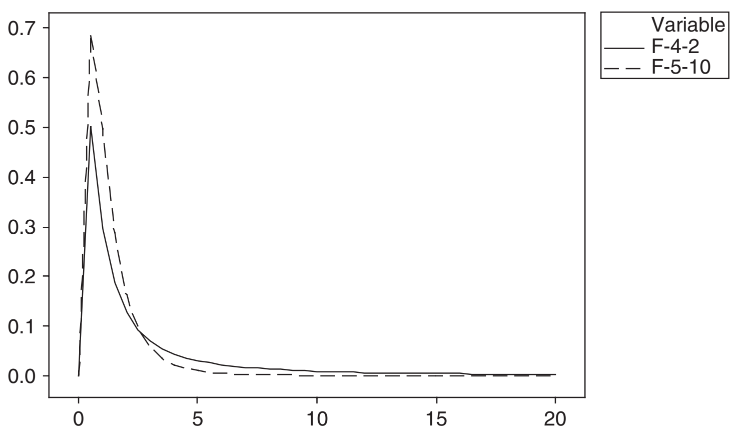

Fig. 11-3 The solid curve is the F distribution with 4 and 2 degrees of freedom and the dashed curve is the F distribution with 5 and 10 degrees of freedom.

Appendixes V and VI give percentile values of F for which the areas in the right-hand tail are 0.05 and 0.01, denoted by $F_{.95}$ and $F_{.99}$, respectively. Representing the 5% and 1% significance levels, these can be used to determine whether or not the variance S_1^2 is significantly larger than S_2^2. In practice, the sample with the larger variance is chosen as sample 1.

Statistical software has added to the ability to find areas under the Student t distribution, the chi-square distribution, and the F distribution. It has also added to our ability to sketch the various distributions. We will illustrate this in the solved problems section in this chapter.

Solved Problems

STUDENT'S *t* DISTRIBUTION

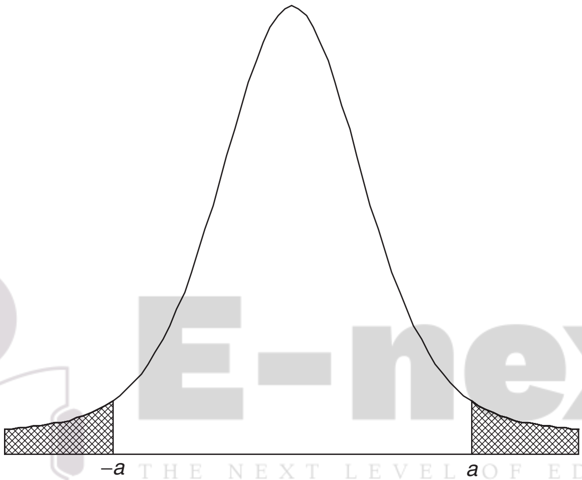

Fig. 11-4 Student's t distribution with 9 degrees of freedom.

11.1 The graph of Student's t distribution with nine degrees of freedom is shown in Fig. 11-4. Use Appendix III to find the values of a for which (*a*) the area to the right of a is 0.05, (*b*) the total shaded area is 0.05, (*c*) the total unshaded area is 0.99, (*d*) the shaded area on the left is 0.01, and (*e*) the area to the left of a is 0.90. Find (*a*) through (*e*) using EXCEL.

SOLUTION

(*a*) If the shaded area on the right is 0.05, the area to the left of a is $(1 - 0.05) = 0.95$ and a represents the 95th percentile, $t_{.95}$. Referring to Appendix III, proceed downward under the column headed υ until reaching entry 9, and then proceed right to the column headed $t_{.95}$; the result, 1.83, is the required value of t.

(*b*) If the total shaded area is 0.05, then the shaded area on the right is 0.025 by symmetry. Thus the area to the left of a is $(1 - 0.025) = 0.975$ and a represents the 97.5th percentile, $t_{.975}$. From Appendix III we find 2.26 to be the required value of t.

(*c*) 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 Appendix III we find that $t_{.995} = 3.25$.

(*d*) If the shaded area on the left is 0.01, then by symmetry the shaded area on the right is 0.01. From Appendix III, $t_{.99} = 2.82$. Thus the critical value of t for which the shaded area on the left is 0.01 equals -2.82.

(*e*) If the area to the left of a is 0.90, the a corresponds to the 90th percentile, $t_{.90}$, which from Appendix III equals 1.38.

Using EXCEL, the expression = TINV(0.1,9) gives 1.833113. EXCEL requires the sum of the areas in the two tails and the degrees of freedom. Similarly, =TINV(0.05,9) gives 2.262157, =TINV(0.01,9) gives 3.249836, =TINV(0.02,9) gives 2.821438, and =TINV(0.2,9) gives 1.383029.

11.2 Find the critical values of t for which the area of the right-hand tail of the t distribution is 0.05 if the number of degrees of freedom, ν, is equal to (*a*) 16, (*b*) 27, and (*c*) 200.

SOLUTION

Using Appendix III, we find in the column headed $t_{.95}$ the values (*a*) 1.75, corresponding to $\nu = 16$; (*b*) 1.70, corresponding to $\nu = 27$; and (*c*) 1.645, corresponding to $\nu = 200$. (The latter is the value that would be obtained by using the normal curve; in Appendix III it corresponds to the entry in the last row marked ∞, or infinity.)

11.3 The 95% confidence coefficients (two-tailed) for the normal distribution are given by ± 1.96. What are the corresponding coefficients for the t distribution if (*a*) $\nu = 9$, (*b*) $\nu = 20$, (*c*) $\nu = 30$, and (*d*) $\nu = 60$?

SOLUTION

For the 95% confidence coefficients (two-tailed), the total shaded area in Fig. 11-4 must be 0.05. Thus the shaded area in the right tail is 0.025 and the corresponding critical value of t is $t_{.975}$. Then the required confidence coefficients are $\pm t_{.975}$; for the given values of ν, these are (*a*) ± 2.26, (*b*) ± 2.09, (*c*) ± 2.04, and (*d*) ± 2.00.

11.4 A sample of 10 measurements of the diameter of a sphere gave a mean $\bar{X} = 438$ centimeters (cm) and a standard deviation $s = 0.06$ cm. Find the (*a*) 95% and (*b*) 99% confidence limits for the actual diameter.

SOLUTION

(*a*) The 95% confidence limits are given by $\bar{X} \pm t_{.975}(s/\sqrt{N-1})$.

Since $\nu = N - 1 = 10 - 1 = 9$, we find $t_{.975} = 2.26$ [see also Problem 11.3(*a*)]. Then, using $\bar{X} = 4.38$ and $s = 0.06$, the required 95% confidence limits are $4.38 \pm 2.26(0.06/\sqrt{10-1}) = 4.38 \pm 0.0452$ cm. Thus we can be 95% confident that the true mean lies between $(438 - 0.045) = 4.335$ cm and $(4.38 + 0.045) = 4.425$ cm.

(*b*) The 99% confidence limits are given by $\bar{X} \pm t_{.995}(s/\sqrt{N-1})$.

For $\nu = 9$, $t_{.995} = 3.25$. Then the 99% confidence limits are $4.38 \pm 3.25(0.06/\sqrt{10-1}) = 4.38 \pm 0.0650$ cm, and the 99% confidence interval is 4.315 to 4.445 cm.

11.5 The number of days absent from work last year due to job-related cases of carpal tunnel syndrome were recorded for 25 randomly selected workers. The results are given in Table 11.1. When the data are used to set a confidence interval on the mean of the population of all job-related cases of carpal tunnel syndrome, a basic assumption underlying the procedure is that the number of days absent are normally distributed for the population. Use the data to test the normality assumption and if you are willing to assume normality, then set a 95% confidence interval on μ.

Table 11.1

21	23	33	32	37
40	37	29	23	29
24	32	24	46	32
17	29	26	46	27
36	38	28	33	18

SOLUTION

The normal probability plot from MINITAB (Fig. 11-5) indicates that it would be reasonable to assume normality since the p-value exceeds 0.15. This p-value is used to test the null hypothesis that the data were selected from a normally distributed population. If the conventional level of significance, 0.05, is used then normality of the population distribution would be rejected only if the p-value is less than 0.05. Since the p-value associated with the Kolmogorov–Smirnov test for normality is reported as p-value > 0.15, we do not reject the assumption of normality.

The confidence interval is found using MINITAB as follows. The 95% confidence interval for the population mean extends from 27.21 to 33.59 days per year.

```
MTB > tinterval 95% confidence for data in cl
```

Confidence Intervals

```
Variable    N     Mean     StDev    SE Mean        95.0 % CI
days       25    30.40     7.72       1.54      (27.21, 33.59)
```

11.6 In the past, a machine has produced washers having a thickness of 0.050 inch (in). To determine whether the machine is in proper working order, a sample of 10 washers is chosen, for which the mean thickness is 0.053 in and the standard deviation is 0.003 in. Test the hypothesis that the machine is in proper working order, using significance levels of (a) 0.05 and (b) 0.01.

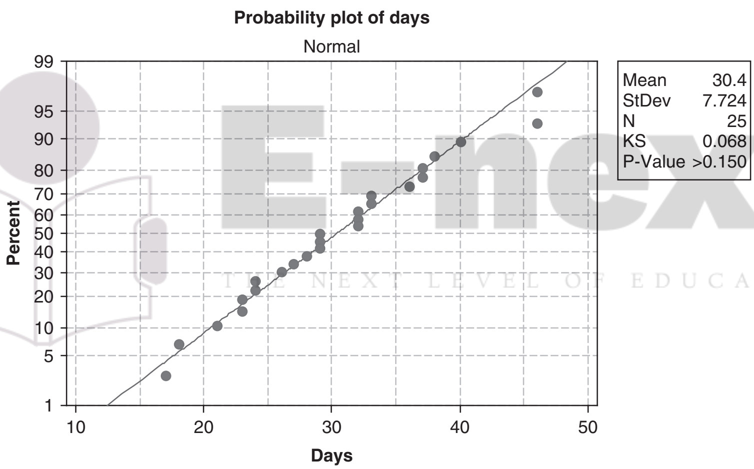

Fig. 11-5 Normal probability plot and Kolmogorov–Smirnov test of normality.

SOLUTION

We wish to decide between the hypotheses:

H_0 : $\mu = 0.050$, and the machine is in proper working order

H_1 : $\mu \neq 0.050$, and the machine is not in proper working order

Thus a two-tailed test is required. Under hypothesis H_0, we have

$$t = \frac{\bar{X} - \mu}{s}\sqrt{N-1} = \frac{0.053 - 0.050}{0.003}\sqrt{10-1} = 3.00$$

(a) For a two-tailed test at the 0.05 significance level, we adopt the decision rule:

Accept H_0 if t lies inside the interval $-t_{.975}$ to $t_{.975}$, which for $10 - 1 = 9$ degrees of freedom is the interval -2.26 to 2.26.

Reject H_0 otherwise

Since $t = 3.00$, we reject H_0 at the 0.05 level.

(*b*) For a two-tailed test at the 0.01 significance level, we adopt the decision rule:

Accept H_0 if t lies inside the interval $-t_{.995}$ to $t_{.995}$, which for $10 - 1 = 9$ degrees of freedom is the interval -3.25 to 3.25.

Reject H_0 otherwise

Since $t = 3.00$, we accept H_0 at the 0.01 level.

Because we can reject H_0 at the 0.05 level but not at the 0.01 level, we say that the sample result is *probably significant* (see the terminology at the end of Problem 10.5). It would thus be advisable to check the machine or at least to take another sample.

11.7 A mall manager conducts a test of the null hypothesis that $\mu = \$50$ versus the alternative hypothesis that $\mu \neq \$50$, where μ represents the mean amount spent by all shoppers making purchases at the mall. The data shown in Table 11.2 give the dollar amount spent for 28 shoppers. The test of hypothesis using the Student's t distribution assumes that the data used in the test are selected from a normally distributed population. This normality assumption may be checked out using anyone of several different *tests of normality*. MINITAB gives 3 different choices for a normality test. Test for normality at the conventional level of significance equal to $\alpha = 0.05$. If the normality assumption is not rejected, then proceed to test the hypothesis that $\mu = \$50$ versus the alternative hypothesis that $\mu \neq \$50$ at $\alpha = 0.05$.

Table 11.2

68	49	45	76	65	50
54	92	24	36	60	66
57	74	52	75	36	40
62	56	94	57	64	
72	65	59	45	33	

SOLUTION

The Anderson–Darling normality test from MINITAB gives a p-value $= 0.922$, the Ryan–Joyner normality test gives the p-value as greater than 0.10, and the Kolmogorov–Smirnov normality test gives the p-value as greater than 0.15. In all 3 cases, the null hypothesis that the data were selected from a normally distributed population would not be rejected at the conventional 5% level of significance. Recall that the null hypothesis is rejected only if the p-value is less than the preset level of significance. The MINITAB analysis for the test of the mean amount spent per customer is shown below. If the classical method of testing hypothesis is used then the null hypothesis is rejected if the computed value of the test statistic exceeds 2.05 in absolute value. The critical value, 2.05, is found by using Student's t distribution with 27 degrees of freedom. Since the computed value of the test statistic equals 18.50, we would reject the null hypothesis and conclude that the mean amount spent exceeds \$50. If the p-value approach is used to test the hypothesis, then since the computed p-value $= 0.0000$ is less than the level of significance (0.05), we also reject the null hypothesis.

Data Display

```
Amount
68        54    57    62    72    49    92    74    56
65        45    24    52    94    59    76    36    75
57        45    65    60    36    64    33    50    66
40

MTB > TTest 0.0 'Amount';
SUBC > Alternative 0.
```

T-Test of the Mean

```
Test of mu = 0.00 vs mu not = 0.00

Variable     N      Mean      StDev     SE Mean        T          P
Amount      28     58.07      16.61        3.14    18.50     0.0000
```

11.8 The intelligence quotients (IQs) of 16 students from one area of a city showed a mean of 107 and a standard deviation of 10, while the IQs of 14 students from another area of the city showed a mean of 112 and a standard deviation of 8. Is there a significant difference between the IQs of the two groups at significance levels of (*a*) 0.01 and (*b*) 0.05?

SOLUTION

If μ_1 and μ_2 denote the population mean IQs of students from the two areas, respectively, we have to decide between the hypotheses:

$H_0 : \mu_1 = \mu_2$, and there is essentially no difference between the groups.

$H_1 : \mu_1 \neq \mu_2$, and there is a significant difference between the groups.

Under hypothesis H_0,

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{1/N_1 + 1/N_2}} \qquad \text{where} \qquad \sigma = \sqrt{\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2}}$$

Thus $$\sigma = \sqrt{\frac{16(10)^2 + 14(8)^2}{16 + 14 - 2}} = 9.44 \qquad \text{and} \qquad t = \frac{112 - 107}{9.44\sqrt{1/16 + 1/14}} = 1.45$$

(*a*) Using a two-tailed test at the 0.01 significance level, we would reject H_0 if t were outside the range $-t_{.995}$ to $t_{.995}$, which for $(N_1 + N_2 - 2) = (16 + 14 - 2) = 28$ degrees of freedom is the range -2.76 to 2.76. Thus we cannot reject H_0 at the 0.01 significance level.

(*b*) Using a two-tailed test at the 0.05 significance level, we would reject H_0 if t were outside the range $-t_{.975}$ to $t_{.975}$, which for 28 degrees of freedom is the range -2.05 to 2.05. Thus we cannot reject H_0 at the 0.05 significance levels.

We conclude that there is no significant difference between the IQs of the two groups.

11.9 The costs (in thousands of dollars) for tuition, room, and board per year at 15 randomly selected public colleges and 10 randomly selected private colleges are shown in Table 11.3. Test the null hypothesis that the mean yearly cost at private colleges exceeds the mean yearly cost at public colleges by 10 thousand dollars versus the alternative hypothesis that the difference is not 10 thousand dollars. Use level of significance 0.05. Test the assumptions of normality and equal variances at level of significance 0.05 before performing the test concerning the means.

Table 11.3

Public Colleges			Private Colleges	
4.2	9.1	11.6	13.0	17.7
6.1	7.7	10.4	18.8	17.6
4.9	6.5	5.0	13.2	19.8
8.5	6.2	10.4	14.4	16.8
4.6	10.2	8.1	17.7	16.1

SOLUTION

The Anderson–Darling normality test from MINITAB for the public colleges data is shown in Fig. 11-6. Since the p-value (0.432) is not less than 0.05, the normality assumption is not rejected. A similar test for private colleges indicates that the normality assumption is valid for private colleges.

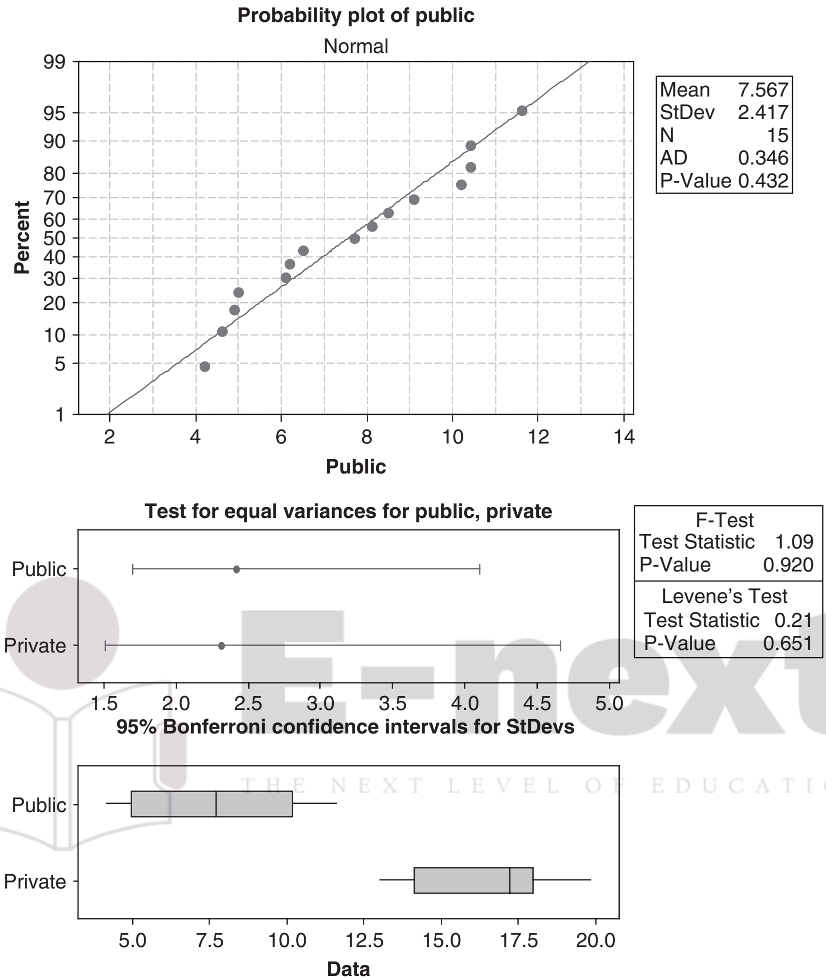

Fig. 11-6 Anderson–Darling test of normality and F-test of equal variances.

The F test shown in the lower part of Fig. 11-6 indicates that equal variances may be assumed. The pull-down menu **"Stat ⇒ Basic Statistics ⇒ 2-sample t"** gives the following output. The output indicates that we cannot reject that the cost at private colleges exceeds the cost at public colleges by 10,000 dollars.

Two-Sample T-Test and CI: Public, Private

```
Two-sample T for Public vs Private
            N      Mean     StDev    SE Mean
Public     15      7.57      2.42       0.62
Private    10     16.51      2.31       0.73

Difference = mu(Public) - mu(Private)
Estimate for difference: -8.9433
95% CI for difference: (-10.9499, -6.9367)
T-Test of difference = -10 (vs not =): T-Value = 1.09 P-Value = 0.287
DF = 23
Both use Pooled StDev = 2.3760
```

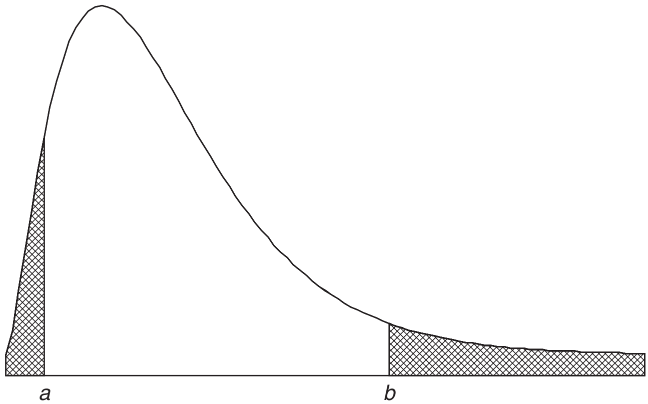

Fig. 11-7 Chi-square distribution with 5 degrees of freedom.

11.10 The graph of the chi-square distribution with 5 degrees of freedom is shown in Fig. 11-7. Using Appendix IV, find the critical values of χ^2 for which (*a*) the shaded area on the right is 0.05, (*b*) the total shaded area is 0.05, (*c*) the shaded area on the left is 0.10, and (*d*) the shaded area on the right is 0.01. Find the same answers using EXCEL.

SOLUTION

(*a*) If the shaded area on the right is 0.05, then the area to the left of b is $(1 - 0.05) = 0.95$ and b represents the 95th percentile, $\chi^2_{.95}$. Referring to Appendix IV, proceed downward under the column headed υ until reaching entry 5, and then proceed right to the column headed $\chi^2_{.95}$; the result, 11.1, is the required critical value of χ^2.

(*b*) Since the distribution is not symmetrical, there are many critical values for which the total shaded area is 0.05. For example, the right-hand shaded area could be 0.04 while the left-hand shaded area is 0.01. It is customary, however, unless otherwise specified, to choose the two areas to be equal. In this case, then, each area is 0.025. If the shaded area on the right is 0.025, the area to the left of b is $1 - 0.025 = 0.975$ and b represents the 97.5th percentile, $\chi^2_{.975}$, which from Appendix IV is 12.8. Similarly, if the shaded area on the left is 0.025, the area to the left of a is 0.025 and a represents the 2.5th percentile, $\chi^2_{.025}$, which equals 0.831. Thus, the critical values are 0.83 and 12.8.

(*c*) If the shaded area on the left is 0.10, a represents the 10th percentile, $\chi^2_{.10}$, which equals 1.61.

(*d*) If the shaded area on the right is 0.01, the area to the left of b is 0.99 and b represents the 99th percentile, $\chi^2_{.99}$, which equals 15.1.

The EXCEL answer to (*a*) is =CHIINV(0.05,5) or 11.0705. The first parameter in CHIINV is the area to the right of the point and the second is the degrees of freedom. The answer to (*b*) is =CHIINV(0.975,5) or 0.8312 and =CHIINV(0.025,5) or 12.8325. The answer to (*c*) is =CHIINV(0.9,5) or 1.6103. The answer to (*d*) is =CHIINV(0.01,5) or 15.0863.

11.11 Find the critical values of χ^2 for which the area of the right-hand tail of the χ^2 distribution is 0.05, if the number of degrees of freedom, ν, is equal to (*a*) 15, (*b*) 21, and (*c*) 50.

SOLUTION

Using Appendix IV, we find in the column headed $\chi^2_{.95}$ the values (*a*) 25.0, corresponding to $\nu = 15$; (*b*) 32.7, corresponding to $\nu = 21$; and (*c*) 67.5, corresponding to $\nu = 50$.

11.12 Find the median value of χ^2 corresponding to (*a*) 9, (*b*) 28, and (*c*) 40 degrees of freedom.

SOLUTION

Using Appendix IV, we find in the column headed $\chi^2_{.50}$ (since the median is the 50th percentile) the values (*a*) 8.34, corresponding to $\nu = 9$; (*b*) 27.3, corresponding to $\nu = 28$; and (*c*) 39.3, corresponding to $\nu = 40$.

It is of interest to note that the median values are very nearly equal to the number of degrees of freedom. In fact, for $\nu > 10$ the median values are equal to $(\nu - 0.7)$, as can be seen from the table.

11.13 The standard deviation of the heights of 16 male students chosen at random in a school of 1000 male students is 2.40 in. Find the (*a*) 95% and (*b*) 99% confidence limits of the standard deviation for all male students at the school.

SOLUTION

(*a*) The 95% confidence limits are given by $s\sqrt{N}/\chi_{.975}$ and $s\sqrt{N}/\chi_{.025}$.

For $\nu = 16 - 1 = 15$ degrees of freedom, $\chi^2_{.975} = 27.5$ (or $\chi_{.975} = 5.24$) and $\chi^2_{.025} = 6.26$ (or $\chi_{.025} = 2.50$). Then the 95% confidence limits are $2.40\sqrt{16}/5.24$ and $2.40\sqrt{16}/2.50$ (i.e., 1.83 and 3.84 in). Thus we can be 95% confident that the population standard deviation lies between 1.83 and 3.84 in.

(*b*) The 99% confidence limits are given by $s\sqrt{N}/\chi_{.995}$ and $s/\sqrt{N}/\chi_{.005}$.

For $\nu = 16 - 1 = 15$ degrees of freedom, $\chi^2_{.995} = 32.8$ (or $\chi_{.995} = 5.73$) and $\chi^2_{.005} = 4.60$ (or $\chi_{.005} = 2.14$). Then the 99% confidence limits are $2.40\sqrt{16}/5.73$ and $2.40\sqrt{16}/2.14$ (i.e., 1.68 and 4.49 in). Thus we can be 99% confident that the population standard deviation lies between 1.68 and 4.49 in.

11.14 Find $\chi^2_{.95}$ for (*a*) $\nu = 50$ and (*b*) $\nu = 100$ degrees of freedom.

SOLUTION

For ν greater than 30, we can use the fact that $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ is very closely normally distributed with mean 0 and standard deviation 1. Then if z_p is the z-score percentile of the standardized normal distribution, we can write, to a high degree of approximation,

$$\sqrt{2\chi^2_p} - \sqrt{2\nu - 1} = z_p \qquad \text{or} \qquad \sqrt{2\chi^2_p} = z_p + \sqrt{2\nu - 1}$$

from which $\chi^2_p = \frac{1}{2}(z_p + \sqrt{2\nu - 1})^2$.

(*a*) If $\nu = 50$, $\chi^2_{.95} = \frac{1}{2}(z_{.95} + \sqrt{2(50) - 1})^2 = \frac{1}{2}(1.64 + \sqrt{99})^2 = 67.2$, which agrees very well with the value of 67.5 given in Appendix IV.

(*b*) If $\nu = 100$, $\chi^2_{.95} = \frac{1}{2}(z_{.95} + \sqrt{2(100) - 1})^2 = \frac{1}{2}(1.64 + \sqrt{199})^2 = 124.0$ (actual value = 124.3).

11.15 The standard deviation of the lifetimes of a sample of 200 electric light bulbs is 100 hours (h). Find the (*a*) 95% and (*b*) 99% confidence limits for the standard deviation of all such electric light bulbs.

SOLUTION

(*a*) The 95% confidence limits are given by $s\sqrt{N}/\chi_{.975}$ and $s\sqrt{N}/\chi_{.025}$.

For $\nu = 200 - 1 = 199$ degrees of freedom, we find (as in Problem 11.14)

$$\chi^2_{.975} = \tfrac{1}{2}(z_{.975} + \sqrt{2(199) - 1})^2 = \tfrac{1}{2}(1.96 + 19.92)^2 = 239$$

$$\chi^2_{.025} = \tfrac{1}{2}(z_{.025} + \sqrt{2(199) - 1})^2 = \tfrac{1}{2}(-1.96 + 19.92)^2 = 161$$

from which $\chi_{.975} = 15.5$ and $X_{0.025} = 12.7$. Then the 95% confidence limits are $100\sqrt{200}/15.5 = 91.2$ h and $100\sqrt{200}/12.7 = 111.3$ h, respectively. Thus we can be 95% confident that the population standard deviation will lie between 91.2 and 111.3 h.

(*b*) The 99% confidence limits are given by $s\sqrt{N}/\chi_{.995}$ and $s\sqrt{N}/\chi_{.005}$.
For $\nu = 200 - 1 = 199$ degrees of freedom,

$$\chi^2_{.995} = \tfrac{1}{2}(z_{.995} + \sqrt{2(199) - 1})^2 = \tfrac{1}{2}(2.58 + 19.92)^2 = 253$$

$$\chi^2_{.005} = \tfrac{1}{2}(z_{.005} + \sqrt{2(199) - 1})^2 = \tfrac{1}{2}(-2.58 + 19.92)^2 = 150$$

from which $\chi_{.995} = 15.9$ and $\chi_{.005} = 12.2$. Then the 99% confidence limits are $100\sqrt{200}/15.9 = 88.9$ h and $100\sqrt{200}/12.2 = 115.9$ h, respectively. Thus we can be 99% confident that the population standard deviation will lie between 88.9 and 115.9 h.

11.16 A manufacturer of axles must maintain a mean diameter of 5.000 cm in the manufacturing process. In addition, in order to insure that the wheels fit on the axle properly, it is necessary that the standard deviation of the diameters equal 0.005 cm or less. A sample of 20 axles is obtained and the diameters are given in Table 11.4.

Table 11.4

4.996	4.998	5.002	4.999
5.010	4.997	5.003	4.998
5.006	5.004	5.000	4.993
5.002	4.996	5.005	4.992
5.007	5.003	5.000	5.000

The manufacturer wishes to test the null hypothesis that the population standard deviation is 0.005 cm versus the alternative hypothesis that the population standard deviation exceeds 0.005 cm. If the alternative hypothesis is supported, then the manufacturing process must be stopped and repairs to the machinery must be made. The test procedure assumes that the axle diameters are normally distributed. Test this assumption at the 0.05 level of significance. If you are willing to assume normality, then test the hypothesis concerning the population standard deviation at the 0.05 level of significance.

SOLUTION

The Shapiro–Wilk test of normality is given in Fig. 11-8. The large *p*-value (0.9966) would indicate that normality would not be rejected. This probability plot and Shapiro–Wilk analysis was produced by the statistical software package STATISTIX.

We have to decide between the hypothesis:

$$H_0: \sigma = 0.005 \text{ cm, and the observed value is due to chance.}$$

$$H_1: \sigma > 0.005 \text{ cm, and the variability is too large.}$$

The SAS analysis is as follows:

```
One Sample Chi-square Test for a Variance
Sample Statistics for diameter
N                Mean        Std. Dev.          Variance
-----------------------------------------------------------
20              5.0006         0.0046            215E-7


Hypothesis Test
Null hypothesis: Variance of diameter <= 0.000025
Alternative: Variance of diameter > 0.000025
Chi-square           Df          Prob
------------------------------------
16.358               19         0.6333
```

The large *p*-value (0.6333) would indicate that the null hypothesis would not be rejected.

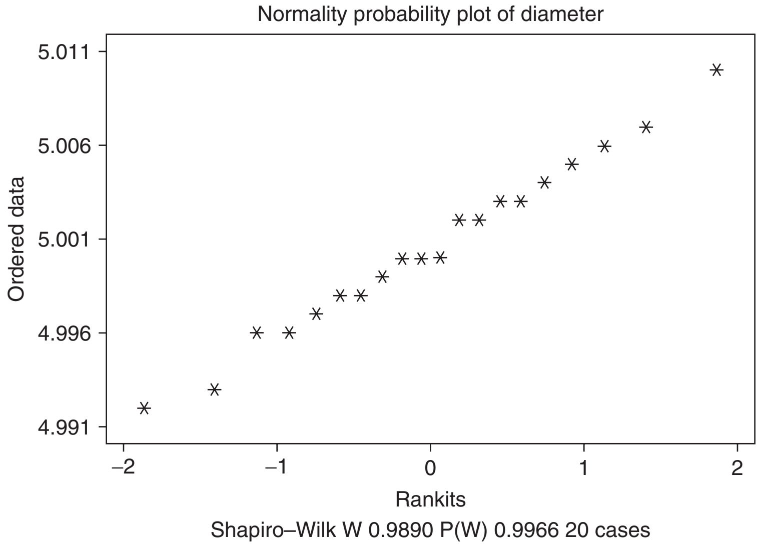

Fig. 11-8 Shapiro–Wilk test of normality from STATISTIX.

11.17 In the past, the standard deviation of weights of certain 40.0-ounce packages filled by a machine was 0.25 ounces (oz). A random sample of 20 packages showed a standard deviation of 0.32 oz. Is the apparent increase in variability significant at the (*a*) 0.05 and (*b*) 0.01 levels?

SOLUTION

We have to decide between the hypotheses:

$H_0 : \sigma = 0.25$ oz, and the observed result is due to chance.

$H_1 : a > 0.25$ oz, and the variability has increased.

The value of χ^2 for the sample is

$$\chi^2 = \frac{Ns^2}{\sigma^2} = \frac{20(0.32)^2}{(0.25)^2} = 32.8$$

(*a*) Using a one-tailed test, we would reject H_0 at the 0.05 significance level if the sample value of χ^2 were greater than $\chi^2_{.95}$, which equals 30.1 for $\nu = 20 - 1 = 19$ degrees of freedom. Thus we would reject H_0 at the 0.05 significance level.

(*b*) Using a one-tailed test, we would reject H_0 at the 0.01 significance level if the sample value of χ^2 were greater than $\chi^2_{.99}$, which equals 36.2 for 19 degrees of freedom. Thus we would not reject H_0 at the 0.01 significance level.

We conclude that the variability has probably increased. An examination of the machine should be made.

THE *F* DISTRIBUTION

11.18 Two samples of sizes 9 and 12 are drawn from two normally distributed populations having variances 16 and 25, respectively. If the sample variances are 20 and 8, determine whether the first sample has a significantly larger variance than the second sample at significance levels of (*a*) 0.05, (*b*) 0.01, and (*c*) Use EXCEL to show that the area to the right of 4.03 is between 0.01 and 0.05.

SOLUTION

For the two samples, 1 and 2, we have $N_1 = 9$, $N_2 = 12$, $\sigma_1^2 = 16$, $\sigma_2^2 = 25$, $S_1^2 = 20$, and $S_2^2 = 8$. Thus

$$F = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} = \frac{N_1 S_1^2/(N_1 - 1)\sigma_1^2}{N_2 S_2^2/(N_2 - 1)\sigma_2^2} = \frac{(9)(20)/(9-1)(16)}{(12)(8)/(12-1)(25)} = 4.03$$

(*a*) The degrees of freedom for the numerator and denominator of F are $\nu_1 = N_1 - 1 = 9 - 1 = 8$ and $\nu_2 = N_2 - 1 = 12 - 1 = 11$. Then from Appendix V we find that $F_{.95} = 2.95$. Since the calculated $F = 4.03$ is greater than 2.95, we conclude that the variance for sample 1 is significantly larger than that for sample 2 at the 0.05 significance level.

(*b*) For $\nu_1 = 8$ and $v_2 = 11$, we find from Appendix VI that $F_{.01} = 4.74$. In this case the calculated $F = 4.03$ is less than 4.74. Thus we cannot conclude that the sample 1 variance is larger than the sample 2 variance at the 0.01 significance level.

(*c*) The area to the right of 4.03 is given by =FDIST(4.03,8,11) or 0.018.

11.19 Two samples of sizes 8 and 10 are drawn from two normally distributed populations having variances 20 and 36, respectively. Find the probability that the variance of the first sample is more than twice the variance of the second sample.

Use EXCEL to find the exact probability that F with 7 and 9 degrees of freedom exceeds 3.70.

SOLUTION

We have $N_1 = 8$, $N_2 = 10$, $\sigma_1^2 = 20$, and $\sigma_2^2 = 36$. Thus

$$F = \frac{8S_1^2/(7)(20)}{10S_2^2/(9)(36)} = 1.85\frac{S_1^2}{S_2^2}$$

The number of degrees of freedom for the numerator and denominator are $\nu_1 = N_1 - 1 = 8 - 1 = 7$ and $\nu_2 = N_2 - 1 = 10 - 1 = 9$. Now if S_1^2 is more than twice S_2^2, then

$$F = 1.85\frac{S_1^2}{S_2^2} > (1.85)(2) = 3.70$$

Looking up 3.70 in Appendixes V and VI, we find that the probability is less than 0.05 but greater than 0.01. For exact values, we need a more extensive tabulation of the F distribution.

The EXCEL answer is =FDIST(3.7,7,9) or 0.036 is the probability that F with 7 and 9 degrees of freedom exceeds 3.70.

Supplementary Problems

STUDENT'S *t* DISTRIBUTION

11.20 For a Student's distribution with 15 degrees of freedom, find the value of t_1 such that (*a*) the area to the right of t_1 is 0.01, (*b*) the area to the left of t_1 is 0.95, (*c*) the area to the right of t_1 is 0.10, (*d*) the combined area to the right of t_1 and to the left of $-t_1$ is 0.01, and (*e*) the area between $-t_1$ and t_1 is 0.95.

11.21 Find the critical values of t for which the area of the right-hand tail of the t distribution is 0.01 if the number of degrees of freedom, v, is equal to (*a*) 4, (*b*) 12, (*c*) 25, (*d*) 60, and (*e*) 150 using Appendix III. Give the solutions to (*a*) through (*e*) using EXCEL.

11.22 Find the values of t_1 for Student's distribution that satisfy each of the following conditions:

(*a*) The area between $-t_1$ and t_1 is 0.90 and $\nu = 25$.

(*b*) The area to the left of $-t_1$ is 0.025 and $\nu = 20$.

(*c*) The combined area to the right of t_1 and left of $-t_1$ is 0.01 and $\nu = 5$.

(*d*) The area to the right of t_1 is 0.55 and $\nu = 16$.

11.23 If a variable U has a Student's distribution with $\nu = 10$, find the constant C such that (*a*) $\Pr\{U > C\} = 0.05$, (*b*) $\Pr\{-C \leq U \leq C\} = 0.98$, (*c*) $\Pr\{U \leq C\} = 0.20$, and (*d*) $\Pr\{U \geq C\} = 0.90$.

11.24 The 99% confidence coefficients (two-tailed) for the normal distribution are given by ± 2.58. What are the corresponding coefficients for the t distribution if (*a*) $\nu = 4$, (*b*) $\nu = 12$, (*c*) $\nu = 25$, (*d*) $\nu = 30$, and (*e*) $\nu = 40$?

11.25 A sample of 12 measurements of the breaking strength of cotton threads gave a mean of 7.38 grams (g) and a standard deviation of 1.24 g. Find the (*a*) 95%, (*b*) 99% confidence limits for the actual breaking strength, and (*c*) the MINITAB solutions using the summary statistics.

11.26 Work Problem 11.25 by assuming that the methods of large sampling theory are applicable, and compare the results obtained.

11.27 Five measurements of the reaction time of an individual to certain stimuli were recorded as 0.28, 0.30, 0.27, 0.33, and 0.31 seconds. Find the (*a*) 95% and (*b*) 99% confidence limits for the actual reaction time.

11.28 The mean lifetime of electric light bulbs produced by a company has in the past been 1120 h with a standard deviation of 125 h. A sample of eight electric light bulbs recently chosen from a supply of newly produced bulbs showed a mean lifetime of 1070 h. Test the hypothesis that the mean lifetime of the bulbs has not changed, using significance levels of (*a*) 0.05 and (*b*) 0.01.

11.29 In Problem 11.28, test the hypothesis $\mu = 1120$ h against the alternative hypothesis $\mu < 1120$ h, using significance levels of (*a*) 0.05 and (*b*) 0.01.

11.30 The specifications for the production of a certain alloy call for 23.2% copper. A sample of 10 analyses of the product showed a mean copper content of 23.5% and a standard deviation of 0.24%. Can we conclude at (*a*) 0.01 and (*b*) 0.05 significance levels that the product meets the required specifications?

11.31 In Problem 11.30, test the hypothesis that the mean copper content is higher than in the required specifications, using significance levels of (*a*) 0.01 and (*b*) 0.05.

11.32 An efficiency expert claims that by introducing a new type of machinery into a production process, he can substantially decrease the time required for production. Because of the expense involved in maintenance of the machines, management feels that unless the production time can be decreased by at least 8.0%, it cannot afford to introduce the process. Six resulting experiments show that the time for production is decreased by 8.4% with a standard deviation of 0.32%. Using significance levels of (*a*) 0.01 and (*b*) 0.05, test the hypothesis that the process should be introduced.

11.33 Using brand A gasoline, the mean number of miles per gallon traveled by five similar automobiles under identical conditions was 22.6 with a standard deviation of 0.48. Using brand B, the mean number was 21.4 with a standard deviation of 0.54. Using a significance level of 0.05, investigate whether brand A is really better than brand B in providing more mileage to the gallon.

11.34 Two types of chemical solutions, A and B, were tested for their pH (degree of acidity of the solution). Analysis of six samples of A showed a mean pH of 7.52 with a standard deviation of 0.024. Analysis of five samples of B showed a mean pH of 7.49 with a standard deviation of 0.032. Using the 0.05 significance level, determine whether the two types of solutions have different pH values.

11.35 On an examination in psychology, 12 students in one class had a mean grade of 78 with a standard deviation of 6, while 15 students in another class had a mean grade of 74 with a standard deviation of 8. Using a significance level of 0.05, determine whether the first group is superior to the second group.

THE CHI-SQUARE DISTRIBUTION

11.36 For a chi-square distribution with 12 degrees of freedom, use Appendix IV to find the value of χ_c^2 such that (*a*) the area to the right of χ_c^2 is 0.05, (*b*) the area to the left of χ_c^2 is 0.99, (*c*) the area to the right of χ_c^2 is 0.025, and (*d*) find (*a*) through (*c*) using EXCEL.

11.37 Find the critical values of χ^2 for which the area of the right-hand tail of the χ^2 distribution is 0.05 if the number of degrees of freedom, ν, is equal to (*a*) 8, (*b*) 19, (*c*) 28, and (*d*) 40.

11.38 Work Problem 11.37 if the area of the right-hand tail is 0.01.

11.39 (*a*) Find χ_1^2 and χ_2^2 such that the area under the χ^2 distribution corresponding to $\nu = 20$ between χ_1^2 and χ_2^2 is 0.95, assuming equal areas to the right of χ_2^2 and left of χ_1^2.

(*b*) Show that if the assumption of equal areas in part (*a*) is not made, the values χ_1^2 and χ_2^2 are not unique.

11.40 If the variable U is chi-square distributed with $\nu = 7$, find χ_1^2 and χ_2^2 such that (*a*) $\Pr\{U > \chi_2^2\} = 0.025$, (*b*) $\Pr\{U < \chi_1^2\} = 0.50$, and (*c*) $\Pr\{\chi_1^2 \le U \le \chi_2^2\} = 0.90$.

11.41 The standard deviation of the lifetimes of 10 electric light bulbs manufactured by a company is 120 h. Find the (*a*) 95% and (*b*) 99% confidence limits for the standard deviation of all bulbs manufactured by the company.

11.42 Work Problem 11.41 if 25 electric light bulbs show the same standard deviation of 120 h.

11.43 Find (*a*)$\chi_{.05}^2$ and (*b*) $\chi_{.95}^2$ for $\nu = 150$ using $\chi_p^2 = \frac{1}{2}(z_p + \sqrt{2\upsilon - 1})^2$ and (*c*) compare the answers when EXCEL is used.

11.44 Find (*a*) $\chi_{.025}^2$ and (*b*) $\chi_{.975}^2$ for $\nu = 250$ using $\chi_p^2 = \frac{1}{2}(z_p + \sqrt{2\upsilon - 1})^2$ and (*c*) compare the answers when EXCEL is used.

11.45 Show that for large values of ν, a good approximation to χ^2 is given by $(\upsilon + z_p\sqrt{2\nu})$, where z_p is the pth percentile of the standard normal distribution.

11.46 Work Problem 11.39 by using the χ^2 distributions if a sample of 100 electric bulbs shows the same standard deviation of 120 h. Compare the results with those obtained by the methods of Chapter 9.

11.47 What is the 95% confidence interval of Problem 11.44 that has the least width?

11.48 The standard deviation of the breaking strengths of certain cables produced by a company is given as 240 pounds (lb). After a change was introduced in the process of manufacture of these cables, the breaking strengths of a sample of eight cables showed a standard deviation of 300 lb. Investigate the significance of the apparent increase in variability, using significance levels of (*a*) 0.05 and (*b*) 0.01.

11.49 The standard deviation of the annual temperatures of a city over a period of 100 years was 16° Fahrenheit. Using the mean temperature on the 15th day of each month during the last 15 years, a standard deviation of annual temperatures was computed as 10° Fahrenheit. Test the hypothesis that the temperatures in the city have become less variable than in the past, using significance levels of (*a*) 0.05 and (*b*) 0.01.

THE *F* DISTRIBUTION

11.50 Find the values of F in parts *a*, *b*, *c*, and *d* using Appendix V and VI.
(*a*) $F_{0.95}$ with $V_1 = 8$ and $V_2 = 10$.
(*b*) $F_{0.99}$ with $V_1 = 24$ and $V_2 = 11$.
(*c*) $F_{0.85}$ with $N_1 = 16$ and $N_2 = 25$.
(*d*) $F_{0.90}$ with $N_1 = 21$ and $N_2 = 23$.

11.51 Solve Problem 11.50 using EXCEL.

11.52 Two samples of sizes 10 and 15 are drawn from two normally distributed populations having variances 40 and 60, respectively. If the sample variances are 90 and 50, determine whether the sample 1 variance is significantly greater than the sample 2 variance at significance levels of (*a*) 0.05 and (*b*) 0.01.

11.53 Two companies, *A* and *B*, manufacture electric light bulbs. The lifetimes for the *A* and *B* bulbs are very nearly normally distributed, with standard deviations of 20 h and 27 h, respectively. If we select 16 bulbs from company *A* and 20 bulbs from company *B* and determine the standard deviations of their lifetimes to be 15 h and 40 h, respectively, can we conclude at significance levels of (*a*) 0.05 and (*b*) 0.01 that the variability of the *A* bulbs is significantly less than that of the *B* bulbs?