Appendix I

The t Distribution

The t distribution is a symmetric distribution that is peaked at the centre, and approaches the horizontal axis the farther from the centre of the distribution the t value is. For each degree of freedom there is a different t distribution. For small degrees of freedom, the t distribution is considerably more varied than is the normal distribution, but as the degrees of freedom increase, the t distribution approaches the normal distribution. When there are more than 30 degrees of freedom, the normal rather than the t distribution is used.

The t distribution is often required when the sample size is small. Small random samples produce statistics which can be quite variable from sample to sample. When sample sizes are larger, many statistics are normally distributed, but such is not the case for smaller sample sizes. In particular, the mean of a random sample from a normally distributed population has a t distribution. For some other statistics that do not have an exact t distribution, the t distribution may provide quite a close approximation to the sampling distribution, when the sample size is small.

The t table which follows gives t values for the standardized t distribution for various confidence and significance levels. These t values are given for 1 through 30 degrees of freedom, and for $40, 50, \ldots, 100$ degrees of freedom. This t table was generated by the SHAZAM program, and is not as accurate as t tables provided in many textbooks. The t values in the upper right of the table will differ somewhat from those in other, more accurate, tables. But most of the t values in the table are accurate to within 0.001.

The t distribution is shown in Figure I.1 (p. 910). The t values along the horizontal axis are standardized t values. That is, for each degree of

freedom, this t distribution has a mean of 0 and a standard deviation of 1. The vertical axis represents the probability of occurrence for each t value. In percentages, the area in the middle of the distribution is equal to the confidence level shown in the table.4 The proportion of the area in each tail of the distribution is equal to value shown in the column '1 Tailed.' The values in the body of the table are the $t = t1$ values. Since the distribution is symmetric, only the t values on the positive side of the mean are shown in the table.

In the table, only selected areas under the t distribution are shown. When using the table to determine the t value for an interval estimate, select one of the confidence levels shown, and determine the number of degrees of freedom associated with the statistic. Then go to the appropriate combination to determine the t value. For example, at 95% confidence, and 10 degrees of freedom, $t = 2.228$. With respect to Figure **??**, this means that 95% of the area would be in the centre of the distribution, between the shaded areas, and for 10 degrees of freedom, the shaded areas would begin at $t = -2.228$ on the left, and $t = 2.228$ on the right. For a two tailed test of significance at the 0.05 level of significance, the critical values would be -2.228 and +2.228.

The areas in one tail of the distribution are those given in the '1 Tailed' column. For a one tailed significance test in the positive direction, with 15 degrees of freedom, there is 0.01 of the area in the tail of the t distribution that lies to the right of $t = 2.603$.