3.3 TOLERANCE LIMITS

The confidence intervals of Sections 3.1 and 3.2 provide interval estimates for unknown population parameters, such as the unknown probability p or the unknown quantile x_p, and a certain probability $1 - \alpha$ (confidence coefficient) that the unknown parameter is within the interval. Tolerance limits differ from confidence intervals in that tolerance limits provide an interval within which at least a proportion q of the population lies, with probability $1 - \alpha$ or more that the stated interval does indeed "contain" the proportion q of the population. A typical application would be in a situation where we are about to draw a random sample of size n, $X_1, X_2, \ldots, X_n$, and we want to know how large n should be so that we can be 95% sure that at least 90% of the population lies between $X^{(1)}$ and $X^{(n)}$, the largest and smallest observations in our sample. We may generalize somewhat and consider the question, "How large must the sample size n be so that at least a proportion q of the population is between $X^{(r)}$ and $X^{(n+1-m)}$ with probability $1 - \alpha$ or more?" The numbers q, r, m, and $1 - \alpha$ are known (or selected) beforehand, and only n needs to be determined.

Another typical situation is when a random sample of size n is available, and we want 95% confidence (or some other value of $1 - \alpha$) that the limits we choose will contain at least q of the population. What will the population proportion q be if we choose the two extreme values in the sample, $X^{(1)}$ and $X^{(n)}$, as our limits? Or should we choose the second most extreme values as our limits, $X^{(2)}$ and $X^{(n-1)}$? What proportion of the population will lie within those limits, with 95% confidence? In this version of the problem, q is the unknown quantity and is obtained after we know, or set, values for α, n, r, and m.

The preceding tolerance limits would be two-sided tolerance limits. One-sided tolerance limits are of the form, "At least a proportion q of the population is greater than $X^{(r)}$, with probability $1 - \alpha$," or "At least a proportion q of the population is less than $X^{(n+1-m)}$, with probability $1 - \alpha$." One-sided tolerance limits are identical with one-sided confidence intervals for quantiles, as will be shown in this section.

The population referred to here is either infinite or, if the population is finite, the sample is drawn with replacement so the X_is are independent. For finite populations where the sampling is without replacement and where the sample size n is small compared to the population size N, these methods are fairly accurate. More precise methods for finite populations may be found in Wilks (1962).

▶ Tolerance Limits

Data The data consist of a random sample $X_1, X_2, \ldots, X_n$ from a large population. Choose a confidence coefficient $1 - \alpha$ and a pair of positive integers r and m. Either we wish to determine the required sample size n after selecting a desired population proportion q (see Method A), or we wish to determine the population proportion q for a given sample size n (see Method B). We are trying to make the statement, "The probability is $1 - \alpha$ that the random interval from $X^{(r)}$ to $X^{(n+1-m)}$ inclusive contains a proportion q or more of the population." Note that we are using the convention $X^{(0)} = -\infty$ and $X^{(n+1)} = +\infty$, so that one-sided tolerance limits may be obtained by setting either r or m equal to zero.

Assumptions

1. The $X_1, X_2, \ldots, X_n$ constitute a random sample.
2. The measurement scale is at least ordinal.

Method A (to find n) If $r + m$ equals 1, that is, if either r or m equals zero as in a one-sided tolerance limit, read n directly from Table A5 for the appropriate values of α and q. If $r + m$ equals 2, read n directly from Table A6 for the appropriate values of α and q. If Tables A5 and A6 are not appropriate, use the approximation

$$\boxed{n \cong \frac{1}{4} x_{1-\alpha} \frac{1+q}{1-q} + \frac{1}{2}(r + m - 1)} \tag{1}$$

where $x_{1-\alpha}$ is the $(1 - \alpha)$ quantile of a chi-squared random variable with $2(r + m)$ degrees of freedom, obtained from Table A2.

Method B (to find q) For a given sample size n and selected values of α, r, and m, the approximate value of q, the proportion of the population, is given by

$$\boxed{q = \frac{4n - 2(r + m - 1) - x_{1-\alpha}}{4n - 2(r + m - 1) + x_{1-\alpha}}} \tag{2}$$

where $x_{1-\alpha}$ is the $(1 - \alpha)$ quantile of a chi-squared random variable with $2(r + m)$ degrees of freedom.

Tolerance Limit With a sample of size n, there is probability at least $1 - \alpha$ that at least q [or $(100)(q)\%$] of the population is between $X^{(r)}$ and $X^{(n+1-m)}$ inclusive. That is

$$\boxed{P(X^{(r)} \leq \text{at least a fraction } q \text{ of the population} \leq X^{(n+1-m)}) \geq 1 - \alpha} \tag{3}$$

For one-sided tolerance regions let either r or m equal zero, where $X^{(0)}$ and $X^{(n+1)}$ are considered to be $-\infty$ and $+\infty$ repectively, and proceed as described above.

EXAMPLE 1

3.3 Tolerance Limits

Method A: To find n, q is known

How large n should be, with $1-\alpha\%$ confident that $\geq q\%$ of the Population will be from $X^{(1)}$ & $X^{(n)}$ or

$$\boxed{X^{(r)} \leq \underset{\text{of the Pop.}}{\text{at least } q\%} \leq X^{(n+1-m)}}$$

Note: either r or $m=0$, it will be one sided Tolerence Limit.

Method B: To find q, n is known

Given known n, what proportion (q) of the Population (at least) are within a sample range. with $1-\alpha\%$ confident.

$$\boxed{\text{from } X^{(r)} \text{ and } X^{(n+1-m)}}$$

Note: either r or $m=0$ for one-sided tolerance Limit. $r=m=$ any #