A PRIMER ON CONFIDENCE INTERVALS

Scott Alexander

Stat 570

Department of Mathematical Sciences

Montana State University

Bozeman, Montana 59715

1. OVERVIEW

In constructing meaningful confidence intervals it is convenient to first obtain a point estimate such as a maximum likelihood estimate or a method of moments estimate. Once a point estimate has been obtained, if possible, it is desirable to provide an interval based on the point estimate. Such point estimates are quite useful, however they leave something to be desired. The problem is that the probability that the estimator actually equals the value of the parameter being estimated is 0. The probability that a continuous random variable equals any one value is 0. Hence, it would be nice to include some measure of the possible error of the estimate. For instance, a point estimate might be given with an interval formed around the point estimate along with some indication as to how likely the true value of the parameter is to lie within the interval. Then instead of making inferences on point values of the parameter, it is possilbe to make inferences on interval values of the parameter. This form of estimation is known as interval estimation, which is to be the subject of this paper.

The paper will be divided into ~~six~~ main sections~~, the first being this overview section~~. Section 2 introduces and defines confidence intervals. Section 3 contains several examples of confidence intervals which are associated with the mean and variance of the normal distribution. In Section 4, two-sample problems are considered for the normal and binomial distributions. Several general methods of finding confidence intervals are given in Section 5, and Section 6 provides a summary.

2. CONFIDENCE INTERVALS

2.1 An Introduction to Confidence Intervals

One frequently sees estimates given in the form of the estimate plus or minus a certain amount. For instance, the Bureau of Labor Statistics may estimate the number of unemployed in a certain area to be 2.4 $\pm$ .3 million at a given time, feeling quite sure that the actual number is between 2.1 and 2.7 million. The average lifetimes of a certain kind of battery may be estimated to be 75 ± 3.29 hours with the idea that the average is very unlikely to be outside the range 71.71 to 78.29. Notice that these estimates are given in the form of intervals.

To better illustrate these ideas, consider a particular example. Suppose that a random sample (4.2, 6.4, 3.6, 8.6) of four observations is drawn from a normal population with an unknown mean μ and a known standard deviation 3. The maximum likelihood estimate of μ is the mean of the sample observations:

$$\bar{x} = 5.7.$$

I will now find upper and lower limits which are likely to contain the true unknown parameter value between them.

For samples of size 4 from the given distribution, the random variable

$$Z = \frac{\bar{X} - \mu}{3/2}$$

is normally distributed with mean 0 and variance 1, where $\bar{X}$ is the sample mean, and 3/2 is the standard deviation divided by the square root of the sample size: $\frac{\sigma}{\sqrt{n}}$. The quantity Z has a density function

$$f_Z(z) = \phi(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}},$$

which does not depend on the true value of the unknown parameter; so the probability that Z is between any two numbers can be computed without knowledge of the mean. For example,

$$P[-1.282 < Z < 1.282] = \int_{-1.282}^{1.282} \phi(z)\, dz = .90. \tag{2.1}$$

In (2.1) the inequality $-1.282 < Z$, or

$$-1.282 < \frac{\bar{X} - \mu}{3/2},$$

is equivalent to the inequality

$$\mu < \bar{X} + (3/2)(1.282) = \bar{X} + 1.923,$$

and the inequality

$$Z < 1.282$$

is equivalent to

$$\mu > \bar{X} - 1.923.$$

Thus, rewriting (2.1) in the form

$$P[\bar{X} - 1.923 < \mu < \bar{X} + 1.923] = .90,$$

and substituting 5.7 for $\bar{X}$ yields the interval

$$(3.777, 7.623).$$

In general, an interval with random endpoints is called a **random interval.** In particular, the interval $(\bar{X} - 1.923, \bar{X} + 1.923)$ is a random interval that contains the true value of μ with probability 0.90. That is, if samples of size 4 were repeatedly drawn from the normal population and if the random interval $(\bar{X} - 1.923, \bar{X} + 1.923)$ were computed for each sample, then the relative frequency of those intervals that contain the true unknown mean μ would approach 90 percent. The interval (3.777, 7.623) is referred to as a 90% **confidence interval** for μ. Because the estimated interval has known endpoints, it is not appropriate to say that it contains the true value of μ with probability 0.90. That is, the parameter μ, although unknown, is a constant, and this particular interval either does or does not contain μ. However, the fact that the associated random interval had probability 0.90, **prior** to estimation, might lead one to insist that they are "90% confident" that $3.777 < \mu < 7.623$. The probability, .90, is called the **confidence coefficient.**

Similarly, intervals with any level of confidence between 0 and 1 can be obtained. For example, using

$$P[-1.96 < Z < 1.96] = .95,$$

a 95 percent confidence interval for the true mean is obtained by converting the inequalities, as before, to get

$$P[\bar{X} - 2.94 < \mu < \bar{X} + 2.94] = .95$$

and then substituting 5.7 for $\bar{X}$ to get the interval (2.76, 8.64).

Notice that there are an infinite number of possible intervals with the same probability (with the same confidence coefficient). For example, because

$$P[-1.68 < Z < 2.70] = .95,$$

another 95 percent confidence interval for μ is given by the interval (1.65, 8.22). This interval could be thought of as being inferior to the one obtained before because its length, 6.57, is greater than the length, 5.88, of the interval (2.76, 8.64),and thus it gives less precise information about the location of μ.

The method of finding a confidence interval shown in the example above is a common method. The method involves finding, if possible, a function (the quantity Z above) only of the sample and the parameter to be estimated which has a distribution that does not depend on the parameter or any other parameters. Then any probability statement of the form $P[a < Z < b] = \gamma$ for known a and b, where Z is the function, will give rise to a probability statement about the parameter that can, possibly, be rewritten to give a confidence interval. This method, or technique, is fully described in Subsection 5.1 below. This technique is applicable in many important problems, but in others it is not because it is either impossible to find functions of the desired form or it is impossible to rewrite the derived probability statements. These problems can be dealt with by a more general technique to be described in Subsection 5.2.

2.2 Definition of Confidence Interval

In the last subsection a simple example was presented to give some intuition for confidence intervals. In this subsection confidence intervals are defined.

Definition 1: Confidence Interval

> Let $X_1, ..., X_n$ be a random sample from the pdf $f(x;\theta)$. Let L and U be two statistics satisfying $L < U$ such that
>
> $$P[L < \theta < U] = \gamma$$
>
> where $0 < \gamma < 1$ and γ does not depend on θ, then the interval $(l(\mathbf{x}), u(\mathbf{x}))$ is called a **$100\gamma\%$ confidence interval** for θ. The probability, γ, is called the confidence coefficient or confidence level, and the observed values $l(\mathbf{x})$ and $u(\mathbf{x})$ are called **lower** and **upper confidence limits,** respectively.

Notice that one or the other, but not both, of the two statistics L and U may be constant. That is, one of the two end points of the random interval (L, U) may be constant. Also, a distinction should be made between the random interval (L, U) and the observed interval $(l(\mathbf{x}), u(\mathbf{x}))$ as mentioned above. To emphasize this distinction, it is useful to call (L, U) an **interval estimator** and $(l(\mathbf{x}), u(\mathbf{x}))$ an **interval estimate** as is similar to the terminology used in point estimation.

Probably the most common interpretation of confidence intervals is based on the relative frequency property of probability. That is, under the assumptions that there exists a random sample of size n and that the sample came from a known distribution, then if the procedure is repeated many times, it is expected that about $100\gamma\%$ of the intervals constructed capture the true, but unknown, value of θ. Our confidence is in the method. Notice that the confidence coefficient reflects the long-term frequency interpretation of probability.

Often it is desirable to have either a lower or an upper confidence limit, but not both.

Definition 2: One-Sided Confidence Limits

Let $X_1, ..., X_n$ be a random sample from the pdf $f(x;\theta)$. Let L be a statistic such that

$$P[L < \theta] = \gamma$$

where $0 < \gamma < 1$ and γ does not depend on θ, then $l(\mathbf{x})$ is called a **one-sided lower 100γ% confidence limit** for θ. Similarly, let U be a statistic such that

$$P[\theta < U] = \gamma$$

where $0 < \gamma < 1$ and γ does not depend on θ, then $u(\mathbf{x})$ is called a **one-sided upper 100γ% confidence limit** for θ.

Example 1

$X_1, ..., X_n$ is a random sample of size n from the exponential distribution with pdf

$$f_X(x;\theta) = \frac{1}{\theta} e^{-x/\theta} I_{(0,\infty)}(x). \tag{2.2}$$

I will derive a one-sided lower 100γ% confidence limit for θ. It can be shown that $2\Sigma_{i=1}^{n} X_i/\theta = 2n\bar{X}/\theta \sim \chi^2(2n)$ where X_i $i = 1, ..., n$ are independent random variables with distribution (2.2) (Bain and Engelhardt 1989, pp. 215-216). Denote the γ^{th} percentile of a chi-square distribution with ν degrees of freedom by $\chi^2_\gamma(\nu)$. Then

$$\gamma = P[2n\bar{X}/\theta < \chi^2_\gamma(2n)]$$

$$= P[2n\bar{X}/\chi^2_\gamma(2n) < \theta].$$

If $\bar{x}$ is observed, then a one-sided lower 100γ% confidence limit is

$$l(\mathbf{x}) = 2n\bar{x}/\chi^2_\gamma(2n).$$

Similarly, for a one-sided upper $100\gamma\%$ confidence limit

$$\gamma = P[2n\bar{X}/\theta > \chi^2_{1-\gamma}(2n)]$$
$$= P[\theta < 2n\bar{X}/\chi^2_{1-\gamma}(2n)].$$

If $\bar{x}$ is observed, then a one-sided upper $100\gamma\%$ confidence limit is

$$u(\mathbf{x}) = 2n\bar{x}/\chi^2_{1-\gamma}(2n)$$

Suppose that it is desired to find a $100\gamma\%$ confidence interval for θ. If values of $\alpha_1 > 0$ and $\alpha_2 > 0$ are chosen such that $\alpha_1 + \alpha_2 = \alpha = 1-\gamma$, then it follows that

$$P[\chi^2_{\alpha_1}(2n) < 2n\bar{X}/\theta < \chi^2_{1-\alpha_2}(2n)] = 1 - \alpha_1 - \alpha_2,$$

and thus

$$P[2n\bar{X}/\chi^2_{1-\alpha_2}(2n) < \theta < 2n\bar{X}/\chi^2_{\alpha_1}(2n)] = \gamma.$$

Commonly, $\alpha_1 = \alpha_2$,. This equality is known as the **equal tailed** choice and implies that $\alpha_1 = \alpha_2 = \alpha/2$. The corresponding confidence interval has the form

$$(2n\bar{x}/\chi^2_{1-\alpha/2}(2n), 2n\bar{x}/\chi^2_{\alpha/2}(2n)).$$

For some problems, the equal tailed choice of α_1 and α_2 will provide a confidence interval of minimum expected length, but for others it will not. For example, the above corresponding confidence interval does not have this property because the chi-square distribution is not symmetric.

Example 2

Consider a random sample of size n from the uniform distribution $X_i \sim \text{UNIF}(0,\theta)$, $\theta > 0$, and let $X_{n:n}$ be the largest order statistic. I will find a $100(1-\alpha)\%$ equal tailed confidence interval for θ. The pdf of X_i is

$$f_X(x;\theta) = \frac{1}{\theta} I_{(0,\theta)}(x),$$

so the CDF is

$$F_X(x;\theta) = \frac{x}{\theta} I_{(0,\theta)}(x).$$

The CDF of $X_{n:n}$ is

$$G_n(x_{n:n};\theta) = [F_X(x_{n:n};\theta)]^n = \left[\frac{x_{n:n}}{\theta}\right]^n I_{(0,\theta)}(x_{n:n}).$$

Let $Q = \frac{X_{n:n}}{\theta}$, then

$$\begin{aligned} F_Q(q) &= P[Q \le q] \\ &= P\left[\frac{X_{n:n}}{\theta} \le q\right] \\ &= P[X_{n:n} \le q\theta] \\ &= G_n(q\theta) \\ &= \left(\frac{q\theta}{\theta}\right)^n = q^n I_{(0,\theta)}(q) \end{aligned}$$

so

$$f_Q(q) = \frac{dF}{dq} = nq^{n-1} I_{(0,\theta)}(q),$$

This implies

$$Q = \frac{X_{n:n}}{\theta} \sim \text{BETA}(n,1).$$

Thus values of a and b can be found from the beta distribution such that

$$\begin{aligned} 1-\alpha &= P\left[a < \frac{X_{n:n}}{\theta} < b\right] \\ &= P[X_{n:n}/b < \theta < X_{n:n}/a] \end{aligned}$$

so that $(x_{n:n}/b, x_{n:n}/a)$ is a $100(1-\alpha)\%$ confidence interval for θ with confidence coefficient $1-\alpha$. If an equal tailed confidence interval is desired, values of a and b can be found by solving the following two equations

$$F_Q(a) = \alpha/2 \qquad \text{and} \qquad F_Q(b) = 1-\alpha/2.$$

Substituting in the CDF's, the solutions are

$$a^n = \alpha/2 \qquad \text{and} \qquad b^n = 1-\alpha/2$$

so

$$a = \sqrt[n]{\alpha/2} \qquad \text{and} \qquad b = \sqrt[n]{1-\alpha/2}.$$

Hence, substituting in for a and b the following $100(1-\alpha)\%$ confidence interval for θ is obtained

$$(x_{n:n} \,/\, \sqrt[n]{1-\alpha/2}, x_{n:n} \,/\, \sqrt[n]{\alpha/2}).$$

Also, the one-sided lower and upper $100(1-\alpha)\%$ confidence limits for θ can be found in a similar manner.

If

$$\frac{X_{n:n}}{\theta} \sim \text{BETA}(n, 1),$$

then

$$1 - \alpha = P[X_{n:n}/\theta < \sqrt[n]{1-\alpha}]$$

$$= P[X_{n:n} \,/\, \sqrt[n]{1-\alpha} < \theta].$$

If $x_{n:n}$ is observed, then a one-sided lower $100(1-\alpha)\%$ confidence limit is

$$l(\mathbf{x}) = \frac{x_{n:n}}{\sqrt[n]{1-\alpha}}.$$

Similarly, for a one-sided upper $100(1-\alpha)\%$ confidence limit

$$1 - \alpha = P[\frac{X_{n:n}}{\theta} > \sqrt[n]{\alpha}]$$

$$= P[\theta < X_{n:n} \,/\, \sqrt[n]{\alpha}],$$

If $x_{n:n}$ is observed, then a one-sided upper $100(1-\alpha)\%$ confidence limit is

$$u(\mathbf{x}) = \frac{x_{n:n}}{\sqrt[n]{\alpha}}.$$

3. CONFIDENCE INTERVALS FOR THE NORMAL DISTRIBUTION

Suppose $X_1, ..., X_n$ is a random sample of size n from the normal distribution with mean μ and variance σ^2, denoted by $N(\mu, \sigma^2)$. Often it is of interest to find limits within which a parameter should be expected to be found with certain probability. For example, it may be desired to know the lower (α_1) and upper (α_2) limits between which the average IQ (μ) of students at a certain college will fall, given a certain probability, say, of 95 percent. In symbols,

$$P[\alpha_1 < \mu < \alpha_2] = 0.95.$$

This section shows how such intervals (*confidence intervals*) are found when estimating the mean and the variance.

3.1 Confidence Intervals for the Mean

Let X be a random variable that is normally distributed with mean μ and variance σ^2. Two cases need to be considered depending on whether or not σ^2 is known. First consider the case when σ^2 is known. Following from Hogg and Craig (1970, p. 163), $\bar{X}$ has the distribution $N(\mu, \frac{\sigma^2}{n})$, and $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ has the distribution $N(0,1)$, whatever μ. Hence values of a and b (not depending on μ) can be found such that

$$P\left[a < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < b\right] = 1-\alpha.$$

This equation can be solved for μ to read

$$P[\bar{X} - b\sigma/\sqrt{n} < \mu < \bar{X} - a\sigma/\sqrt{n}] = 1-\alpha$$

so that $(\bar{X} - b\sigma/\sqrt{n}, \bar{X} - a\sigma/\sqrt{n})$ is a random interval having probability $1-\alpha$ of including the fixed, but unknown, value of μ. After the random experiment has been performed and the value of $\bar{x} = \sum_{i=1}^{n} x_i/n$ is observed, a $100(1-\alpha)\%$ confidence interval for μ is given by the interval $(\bar{x} - b\sigma/\sqrt{n}, \bar{x} - a\sigma/\sqrt{n})$ where a and b are found by using the normal distribution.

Because the normal distribution is unimodal and symmetric, the shortest such interval has $a = -b, b > 0$. Thus, if the γ^{th} percentile of a normal distribution is denoted by z_γ, then another form (equal tailed choice and shortest interval) of a $100(1-\alpha)\%$ confidence interval for μ is

$$(\bar{x} - z_{1-\alpha/2}\,\sigma/\sqrt{n}, \bar{x} + z_{1-\alpha/.} \quad \sqrt{n}).$$

One-sided confidence limits for μ are found in a s..ilar manner.

$$1 - \alpha = P\left[\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{1-}\right.$$

$$= P[\bar{X} - z_{1-\alpha}\,\sigma/\sqrt{n} <$$

If $\bar{x}$ is observed, then a one-sided lower $100(1-\alpha)$% ... nfidence limit is

$$l(\mathbf{x}) = \bar{x} - z_{1-\alpha}\,\sigma/\sqrt{}$$

Similarly, a one-sided upper $100(1-\alpha)$% confidenc ... it is

$$u(\mathbf{x}) = \bar{x} + z_{1-\alpha}\,\sigma/\sqrt{}$$

Example 3

Consider finding a 95% confidence interval for the av ... e IQ of the students attending a certain college. A random sample of 100 stud ... is selected, and the sample mean is found to be $\bar{x} = 110$. Assume that the IQ' ... distributed normally with a standard deviation, $\sigma = 20$ points.

From a normal table,

$$Z_{1-\alpha/2} = Z_{.975} = 1.$$

so

$$P\left[\bar{X} - \frac{1.96\sigma}{\sqrt{100}} < \mu < \bar{X} - \frac{1.9}{\sqrt{1}} \quad\right. = 0.95.$$

Thus, the specific 95% confidence interval is

$$\left(110 - 1.96 \times \frac{20}{\sqrt{100}}, 110 + 1.9 \quad \frac{20}{\sqrt{100}}\right)$$

or

$$(106.08, 113.92),$$

so (106.08, 113.92) is a 95% confidence interval for average IQ of the students attending a certain college.

Suppose that it is desired to find the one-sidec fidence limits for μ. From a normal table,

$$Z_{1-\alpha} = Z_{.95} = 1.64$$

so if $\bar{x}$ is observed, then a one-sided lower 100(1 – confidence limit is

$$l(\mathbf{x}) = \bar{x} - z_{1-\alpha}\,\sigma/\sqrt{n} = 110 - 1.645 \times \quad \sqrt{100} = 106.71$$

and a one-sided upper $100(1-\alpha)\%$ confidence limi

$$u(\mathbf{x}) = \bar{x} + z_{1-\alpha}\,\sigma/\sqrt{n} = 110 + 1.645 \times \quad \sqrt{100} = 113.29.$$

Notice that the confidence limits of the previou fidence intervals depend on σ. Thus, if σ^2 were not known, the end points of th idom interval would not be statistics, and the sample data would not yield an val with known end points. With a slightly different derivation it is possible to a confidence interval for μ, even if σ^2 is an unknown "nuisance parameter".

Now consider the case in which σ^2 is unknown. ill be convenient to express the results in terms of the unbiased estimator of th iance,

$$S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})}{n-1}$$

in order to use some known distributional properti imely

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

and

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n -$$

In accordance with the definition of the Student's tribution, it is known that the random variable

$$T = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\frac{\sigma^2}{(n-1)}}}} = \frac{\bar{X}}{S_{/}}$$

has a t distribution with $n-1$ degrees of freedon ıatever the value of $\sigma^2 > 0$. Hence values of a and b can be found such that

$$1-\alpha = P\left[a < \frac{\bar{X}-\mu}{S/\sqrt{n}} \quad \right]$$

$$= P[\bar{X} - bS/\sqrt{n} < \mu < \bar{X} \cdot \quad /\sqrt{n}],$$

so that $(\bar{X} - bs/\sqrt{n}, \bar{X} - as/\sqrt{n})$ is a random i: ·al having probability $1-\alpha$ of including the fixed, but unknown, value of μ. A ·: the random experiment has been performed and the values of $\bar{x}$ and $s^2 = (n-1 \quad \sum_{i=1}^{n}(x_i-\bar{x})^2$ are observed, a $100(1-\alpha)\%$ confidence interval for μ is given by th· ·terval $(\bar{x}-bs/\sqrt{n}, \bar{x}-as/\sqrt{n})$ where a and b are found by using the Student's t (;ribution.

Because the graph of the pdf of the random · iable T is unimodal and symmetric about zero, the shortest such interval is ob· ıed with $a=-b, b>0$. Thus, if the γ^{th} percentile of the t-distribution with ν (:ees of freedom is denoted by $t_\gamma(\nu)$, then another form of a $100(1-\alpha)\%$ confide ؛ interval for μ is

$$(\bar{x} - t_{1-\alpha/2}(n-1)s/\sqrt{n}, \bar{x} + t_{1-\alpha} \quad n-1)s/\sqrt{n}).$$

One-sided confidence limits for μ are found ir similar manner.

$$1-\alpha = P\left[\frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{1-\alpha} \quad -1)\right]$$

$$= P[\bar{X} - t_{1-\alpha}(n-1)S/\, \quad < \mu].$$

If $\bar{x}$ and s are observed, then a one-sided lower 10· $-\alpha)\%$ confidence limit is

$$l(\mathbf{x}) = \bar{x} - t_{1-\alpha}(n-1) \quad \sqrt{n}.$$

Similarly, a one-sided upper $100(1-\alpha)\%$ confiden limit is

$$u(\mathbf{x}) = \bar{x} + t_{1-\alpha}(n-1 \quad \sqrt{n}.$$

Example 4

The following data are weights (in pounds) of students attending a certain college: 160, 170, 165, 175, 180. Assume that the data are observed values of a random sample from a normal distribution, $X_i \sim N(\mu, \sigma^2)$ Using the above sample find a 95% confidence interval for the average weight of students attending a certain college.

From a t table,

$$t_{1-\alpha/2}(n-1) = t_{.975}(4) = 2.78$$

so

$$P\left[\bar{X} - \frac{2.78S}{\sqrt{5}} < \mu < \bar{X} + \frac{2.78S}{\sqrt{5}}\right] = 0.95.$$

In this case $\bar{x} = 170$, and

$$s = \sqrt{\frac{\sum_{i=1}^{5}(x_i - 170)^2}{4}} \doteq 7.9$$

Thus, the specific 95% confidence interval is

$$\left(170 - 2.78 \times \frac{7.9}{\sqrt{5}}, 170 + 2.78 \times \frac{7.9}{\sqrt{5}}\right)$$

or

$$(160.2, 179.8),$$

so (160.2, 179.8) is a 95% confidence interval for the average weight of the students attending a certain college.

Suppose that it is desired to find the one-sided confidence limits for μ. From a t table,

$$t_{1-\alpha}(n-1) = t_{.95}(4) = 2.132,$$

so if $\bar{x}$ and s are observed, then a one-sided lower $100(1-\alpha)\%$ confidence limit is

$$l(\mathbf{x}) = \bar{x} - t_{1-\alpha}(n-1)s/\sqrt{n} = 170 - 2.132 \times 7.9/\sqrt{5} = 162.5,$$

and a one-sided upper $100(1-\alpha)\%$ confidence limit is

$$u(\mathbf{x}) = \bar{x} + t_{1-\alpha}(n-1)s/\sqrt{n} = 170 + 2.132 \times 7.9/\sqrt{5} = 177.5.$$

3.2 Confidence Intervals for the Variance

Because the normal distribution depends on σ^2, as well as on μ, information about σ^2 is needed to specify what normal distribution is being worked with. Although confidence intervals for σ^2 are not used as frequently as those for μ, they are still important, particularly if there is interest in the reliability and spread of the observations.

When constructing confidence intervals for σ^2, two cases must be considered depending on whether or not μ is assumed known. First consider the case when μ is a known number. The random variable $Z_i = (X_i - \mu)^2/\sigma^2 \sim \chi^2(1)$, and the random variable $Y = \sum_{i=1}^n Z_i \sim \chi^2(n)$ (Mood, Graybill and Boes 1974, p. 242). Hence from tables, values of a and b can be found such that

$$1-\alpha = P\left[a < \frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} < b\right]$$

$$= P\left[\frac{\sum_{i=1}^n (X_i-\mu)^2}{b} < \sigma^2 < \frac{\sum_{i=1}^n (X_i-\mu)^2}{a}\right]$$

so that $\left(\frac{\sum_{i=1}^n (X_i-\mu)^2}{b}, \frac{\sum_{i=1}^n (X_i-\mu)^2}{a}\right)$ is a random interval having probability $1-\alpha$ of including the fixed, but unknown, value of σ^2. After the random experiment has been performed and the values of x_i are observed, a $100(1-\alpha)\%$ confidence interval for σ^2 is given by the interval $\left(\frac{\sum_{i=1}^n (x_i-\mu)^2}{b}, \frac{\sum_{i=1}^n (x_i-\mu)^2}{a}\right)$ where a and b are found by using the chi-square distribution with n degrees of freedom.

Observe that there are no unique numbers $a < b$ such that $P[a < Y < b] = 1-\alpha$. As was mentioned in example 1, a common method is to form an equal tailed confidence interval, in which, $P[Y < a] = \alpha/2$ and $P[Y > b] = \alpha/2$. Thus another form (equal tailed choice, but not shortest interval) of a $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\left(\frac{\sum_{i=1}^n (x_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}, \frac{\sum_{i=1}^n (x_i-\mu)^2}{\chi^2_{\alpha/2}(n)}\right).$$

One-sided confidence limits for σ^2 are found in a similar manner.

$$1-\alpha = P\left[\frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} < \chi^2_{1-\alpha}(n)\right]$$

$$= P\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha}(n)} < \sigma^2\right].$$

If the values of x_i are observed, then a one-sided lower $100(1-\alpha)\%$ confidence limit is

$$l(\mathbf{x}) = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{1-\alpha}(n)}.$$

Similarly, a one-sided upper $100(1-\alpha)\%$ confidence limit is

$$u(\mathbf{x}) = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{\alpha}(n)}.$$

Example 5

The following data are IQ scores of students attending a certain college: 112, 128, 109, 127, 120, 118. Assume that the data are observed values of a random sample from a normal distribution with mean $\mu = 119$, $X_i \sim$ iid $N(119, \sigma^2)$. Using the above sample find a 98% confidence interval for the variance of the IQ scores of students attending a certain college.

From a χ^2 table,

$$\chi^2_{\alpha/2}(n) = \chi^2_{.01}(6) = 0.87$$

and

$$\chi^2_{1-\alpha/2}(n) = \chi^2_{.99}(6) = 16.81,$$

so

$$P\left[\frac{\sum_{i=1}^{6}(X_i-119)^2}{16.81} < \sigma^2 < \frac{\sum_{i=1}^{6}(X_i-119)^2}{0.87}\right] = 0.98.$$

In this case

$$\sum_{i=1}^{6}(x_i-119)^2 = 296$$

Thus, the specific 98% confidence interval is

$$\left(\frac{296}{16.81}, \frac{296}{0.87}\right)$$

or

$$(17.6, 340.2),$$

so (17.6, 340.2) is a 98% confidence interval for the variance of the IQ scores of the students attending a certain college.

Suppose that it is desired to find the one-sided confidence limits for σ^2. From a χ^2 table,

$$\chi^2_\alpha(n) = \chi^2_{.02}(6) = 1.13$$

and

$$\chi^2_{1-\alpha}(n) = \chi^2_{.98}(6) = 15.03$$

so if the values of x_i are observed, then a one-sided lower $100(1-\alpha)\%$ confidence limit is

$$l(\mathbf{x}) = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{1-\alpha}(n)} = \frac{\sum_{i=1}^{6}(x_i-119)^2}{15.03} = \frac{296}{15.03} = 19.69$$

and a one-sided upper $100(1-\alpha)\%$ confidence limit is

$$u(\mathbf{x}) = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{\alpha}(n)} = \frac{\sum_{i=1}^{6}(x_i-119)^2}{1.13} = \frac{296}{1.13} = 261.95.$$

Notice that the confidence limits of the previous confidence intervals depend on μ. Thus, if μ were not known, the end points of the random interval would not be statistics, and the sample data would not yield an interval with known end points. With a slightly different derivation it is possible to form a confidence interval for σ^2, even if μ is an unknown "nuisance parameter".

Now consider the case in which μ is not known. This case can be handled by recalling that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

Hence from tables values of a and b can be found such that

$$1-\alpha = P\left[a < \frac{(n-1)S^2}{\sigma^2} < b\right]$$

$$= P\left[\frac{(n-1)S^2}{b} < \sigma^2 < \frac{(n-1)S^2}{a}\right]$$

so that $\left(\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right)$ is a random interval having probability $1-\alpha$ of including the fixed, but unknown, value of σ^2. After the random experiment has been performed and the value of s^2 is observed, a $100(1-\alpha)\%$ confidence interval for σ^2 is given by the interval $\left(\frac{(n-1)s^2}{b}, \frac{(n-1)s^2}{a}\right)$ where a and b are found by using the chi-square distribution with $n-1$ degrees of freedom.

Once again, if a and b are replaced with the appropriate chi-square percentiles, the following equal tailed $100(1-\alpha)\%$ confidence interval for σ^2 is obtained when μ is unknown

$$\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}\right).$$

One-sided confidence limits for σ^2 are found in a similar manner.

$$1-\alpha = P\left[\frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\alpha}(n-1)\right]$$

$$= P\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)} < \sigma^2\right].$$

If the value of s^2 is observed, then a one-sided lower $100(1-\alpha)\%$ confidence limit is

$$l(\mathbf{x}) = \frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}.$$

Similarly, a one-sided upper $100(1-\alpha)\%$ confidence limit is

$$u(\mathbf{x}) = \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}.$$

Example 6

Suppose teachers are interested in the variability of the ACT performance of college bound seniors in Montana. They assume that the ACT test scores are normally distributed with mean μ and variance σ^2, and want a 98% confidence interval for σ^2. A random sample of 30 seniors and their ACT test scores gives $\bar{x} = 1150$ with $s^2 = 425$.

From a χ^2 table,

$$\chi^2_{\alpha/2}(n-1) = \chi^2_{.01}(29) = 14.26$$

and

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{.99}(29) = 49.59,$$

so

$$P\left[\frac{29S^2}{49.59} < \sigma^2 < \frac{29S^2}{14.26}\right] = 0.98.$$

In this case

$$s^2 = 425.$$

Thus, the specific 98% confidence interval is

$$\left(\frac{29 \times 425}{49.59}, \frac{29 \times 425}{14.26}\right)$$

or

$$(248.5, 864.3),$$

so (248.5, 864.3) is a 98% confidence interval for the variance of the ACT performance of college bound seniors in Montana.

Suppose that it is desired to find the one-sided confidence limits for σ^2. From a χ^2 table,

$$\chi^2_{\alpha}(n-1) = \chi^2_{.02}(29) = 15.57$$

and

$$\chi^2_{1-\alpha}(n-1) = \chi^2_{.98}(29) = 46.70$$

so if the value of s^2 is observed, then a one-sided lower 98% confidence limit is

$$l(\mathbf{x}) = \frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)} = \frac{29 \times s^2}{46.70} = \frac{29 \times 425}{46.70} = 263.92,$$

and a one-sided upper 98% confidence limit is

$$u(\mathbf{x}) = \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)} = \frac{29 \times s^2}{15.57} = \frac{29 \times 425}{15.57} = 791.59.$$

4. TWO-SAMPLE PROBLEMS

In this section, two-sample problems are examined. Usually two-sample problems fall into one of two different formats: either two different treatments are applied to two independent sets of similar subjects or the same treatment is applied to two different kinds of subjects. For example, testing whether rats raised by themselves (treatment X) react differently in a stress situation than rats raised with siblings (treatment Y) would be an example of the first type. On the other hand, examining two groups of students A(high income group) and B (low income group) to see whether there is a significant difference between their weekly allowances would be an example of the second type.

Inferences in two-sample problems usually reduce to a comparison of parameters such as means or probabilities. It might be assumed, for example, that the population of responses associated with, say, treatment X is normally distributed with mean μ_1 and standard deviation σ_1 while the Y distribution is normal with mean μ_2 and standard deviation σ_2. Then it may be of interest to know whether one population has a smaller mean than the other. For example, the weekly allowances of the students discussed above.

Sometimes, although less frequently, it becomes more relevant to compare the *variabilities* of two treatments. A food company, for example, trying to decide which of two types of machines to buy for filling cereal boxes would be concerned about the *average* weights of the boxes filled by each type, but they would also want to know something about the *variabilities* of the weights. Obviously, a machine that produced high proportions of "underfills" and "overfills" would be undesirable.

4.1 Two-Sample Normal Procedures

Assume that $X_1, ..., X_{n_1}$ is a random sample of size n_1 from $X \sim N(\mu_1, \sigma_1^2)$ and $Y_1, ..., Y_{n_2}$ is a random sample of size n_2 from $Y \sim N(\mu_2, \sigma_2^2)$. The sample sizes n_1 and n_2 need not be the same. Let $\bar{X}, S_1^2, \bar{Y}, S_2^2$ be the sample mean and sample variance for the respective samples from the two distributions. When sampling from two normal distributions, it is possible to find confidence intervals for the difference in the means and variances. These are the subjects of the next two subsections.

4.1.1 Procedure for means

Suppose that the random variable X has a normal distribution with unknown parameters μ_1 and σ_1^2. Assume that a modification can be made in conducting an experiment so that the mean of the distribution will be changed; say, increased. After the modification has been done, let the random variable now be denoted by Y, and suppose Y has a normal distribution with unknown parameters μ_2 and σ_2^2. It is hoped that μ_2 is greater than μ_1, or put another way, that $\mu_2 - \mu_1 > 0$. Thus it is desired to make statistical inferences by constructing a confidence interval for the difference $\mu_2 - \mu_1$ and observing whether the interval contains negative values or not.

When constructing confidence intervals for the difference of means, $\mu_2 - \mu_1$, two cases need to be considered depending on whether σ_1^2 and σ_2^2 are known or not. First consider the case when the distribution variances are possibly different, but known. Thus, $\bar{X} \sim \mathrm{N}\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$ and $\bar{Y} \sim \mathrm{N}\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$. In accordance with Hogg and Craig (1970, p. 158), the difference $\bar{Y} - \bar{X}$ is normally distributed with mean $\mu_2 - \mu_1$ and variance $\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}$. Thus the random variable

$$Z = \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}}$$

has the distribution N(0, 1), and it follows that

$$1 - \alpha = P[z_{\alpha/2} < Z < z_{1-\alpha/2}]$$

$$= P\left[(\bar{Y} - \bar{X}) - z_{1-\alpha/2}\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}} < \mu_2 - \mu_1 < (\bar{Y} - \bar{X}) - z_{\alpha/2}\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}\right]$$

so that $\left((\bar{Y} - \bar{X}) - z_{1-\alpha/2}\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}, (\bar{Y} - \bar{X}) - z_{\alpha/2}\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}\right)$ is a random interval having probability $1 - \alpha$ of including the fixed ,but unknown, difference of

$\mu_2 - \mu_1$. After the random experiment has been performed and the values of $\bar{y}$ and $\bar{x}$ are observed, a $100(1-\alpha)\%$ confidence interval for $\mu_2 - \mu_1$ is given by the interval $\left((\bar{y}-\bar{x}) - z_{1-\alpha/2}\sqrt{\frac{\sigma_2^2}{n_2}+\frac{\sigma_1^2}{n_1}}, (\bar{y}-\bar{x}) - z_{\alpha/2}\sqrt{\frac{\sigma_2^2}{n_2}+\frac{\sigma_1^2}{n_1}}\right)$.

Example 7

Two groups of pigs were fed different diets. A random sample of 9 pigs was selected from each group and their sample means were found to be $\bar{x} = 80$ lb., $\bar{y} = 90$ lb.

It is assumed that the weights are normally distributed and the standard deviations are $\sigma_1 = 9$ lb., $\sigma_2 = 18$ lb. Consider finding the 90% confidence interval for the difference of the means.

From a normal table,

$$Z_{1-\alpha/2} = Z_{.95} = 1.645$$

so

$$P\left[(\bar{Y}-\bar{X}) - 1.645\sqrt{\frac{18^2}{9}+\frac{9^2}{9}} < \mu_2 - \mu_1 < (\bar{Y}-\bar{X}) + 1.645\sqrt{\frac{18^2}{9}+\frac{9^2}{9}}\right] = 0.90.$$

Thus, the specific 90% confidence interval is

$$((80-90) - 1.645 \times 6.708, (80-90) + 1.645 \times 6.708)$$

or

$$(-21.035, 1.035),$$

so (-21.035, 1.035) is a 90% confidence interval for the difference of the means.

The second case to consider is when the distribution variances are possibly different and unknown. In this case, if the variances are different, an exact confidence interval can not be found. Hence, I will consider this case in two parts: First, when the variances are equal; and Second, when the variances are different.

Consider first when the distribution variances are equal. As was previously noted in this paper, it is known that

$$\frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2(n_1-1) \quad \text{and} \quad \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_2-1).$$

Hence,

$$V = \frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2).$$

Using the standard normal random variable, Z, from the first case now yields the random variable

$$T = \frac{Z}{\sqrt{V/(n_1 + n_2 - 2)}}$$

which has a t distribution with $n_1 + n_2 - 2$ degrees of freedom. The random variable T can be simplified by denoting what many text books call the "pooled" variance by

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

Then,

$$T = \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{S_p^2(\frac{1}{n_2} + \frac{1}{n_1})}}.$$

If $t_{\alpha/2} = t_{\alpha/2}(n_1 + n_2 - 2)$ and $t_{1-\alpha/2} = t_{1-\alpha/2}(n_1 + n_2 - 2)$, then

$$1 - \alpha = P[t_{\alpha/2} < T < t_{1-\alpha/2}]$$

$$= P\left[(\bar{Y} - \bar{X}) - t_{1-\alpha/2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_2 - \mu_1 < (\bar{Y} - \bar{X}) - t_{\alpha/2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$$

so that $\left((\bar{Y} - \bar{X}) - t_{1-\alpha/2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, (\bar{Y} - \bar{X}) - t_{\alpha/2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$ is a random interval having probability $1 - \alpha$ of including the fixed, but unknown, difference of $\mu_2 - \mu_1$. After the random experiment has been performed and the values of $s_1^2, s_2^2, \bar{y}$ and $\bar{x}$ are observed, a $100(1 - \alpha)\%$ confidence interval for μ is given by the interval $\left((\bar{y} - \bar{x}) - t_{1-\alpha/2}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, (\bar{y} - \bar{x}) - t_{\alpha/2}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$

Example 8

Two diets are being tested in a controlled laboratory experiment. Fifty overweight people have been randomly divided into two groups, say group X and group Y, of 15 and 35. Notice $n_1 \neq n_2$. A different diet is used for each group and each person's weight loss is recorded at the end of a 12 week period. If the weight losses are assumed to be normally distributed with equal variances and the results are

$\bar{X} = 9.8$ lbs., $s_1^2 = 19.7, \bar{Y} = 11.7$ lbs., $s_2^2 = 6.3$, consider finding the 95% confidence interval for the difference of the means.

From a t table,

$$t_{1-\alpha/2}(n_1 + n_2 - 2) = t_{.975}(48) = 2.0106$$

so

$$P\left[(\bar{Y} - \bar{X}) - 2S_p\sqrt{\frac{1}{15} + \frac{1}{35}} < \mu_2 - \mu_1 < (\bar{Y} - \bar{X}) + 2S_p\sqrt{\frac{1}{15} + \frac{1}{35}}\right] = .95.$$

In this case,

$$s_p^2 = \frac{(15-1) \times 19.7 + (35-1) \times 6.3}{48} = 10.208.$$

Thus, the specific 95% confidence interval is

$$((11.7 - 9.8) - 2.0106 \times 10.208 \times 0.309, (11.7 - 9.8) + 2.0106 \times 10.208 \times 0.309)$$

or

$$(-4.442, 8.242),$$

so (-4.442, 8.242) is a 95% confidence interval for the difference of the means.

The second part, when the distribution variances are different, is a difficult problem and is known as the Behrens-Fisher problem. This problem is difficult because no random variable exists whose distribution does not depend on any unknown parameters. Because the variances are unknown it isn't possible to form a variable from the normal distribution, and because the variances are unequal it isn't possible to from a variable from the t distribution. According to Bain and Engelhardt (1989, p. 358), it is possible to form random variables which will lead to approximate confidence intervals but not exact. If it happens that the ratio of the variances is a known constant then the above T random variable can be used by replacing the ratios of the variances by the constant.

4.1.2 Procedure for variances

Suppose that the random variable X has a normal distribution with variance σ_1^2. Assume that although σ_1^2 is not known, it is found that the experimental values of X are widely spread, so that σ_1^2 must be quite large. It is believed that a certain modification in conducting the experiment may reduce the variance. After the modification has been done, let the random variable now be denoted by Y, and suppose Y has a normal distribution with variance σ_2^2. It is hoped that σ_2^2 is less than σ_1^2, or put another way, that $\sigma_2^2/\sigma_1^2 < 1$. Thus it is desired to make statistical inferences by constructing a confidence interval for the ratio σ_2^2/σ_1^2 and observing whether the interval contains values greater than one or not.

When constructing confidence intervals for the ratio of the variances, σ_2^2/σ_1^2, two cases need to be considered depending on whether the distribution means are known or not. First consider the case when the distribution means are known. Following from Mood et al. (1974, p. 242), the random variables

$$V = \sum_{i=1}^{n_1} \left(\frac{X_i - \mu_1}{\sigma_1} \right)^2 \sim \chi^2(n_1) \quad \text{and} \quad W = \sum_{j=1}^{n_2} \left(\frac{Y_j - \mu_2}{\sigma_2} \right)^2 \sim \chi^2(n_2)$$

are obtained. Thus from Mood et al. (1974, p. 247), it follows that

$$\frac{\frac{V}{n_1}}{\frac{W}{n_2}} \sim F(n_1, n_2)$$

where F(n_1, n_2) denotes the Snedecor's F Distribution with n_1 and n_2 degrees of freedom. Hence, if the $100\gamma^{th}$ percentile of the F distribution with n_1 and n_2 degrees of freedom is denoted by $f_\gamma(n_1, n_2)$. It follows that

$$1 - \alpha = P\left[f_{\alpha/2}(n_1, n_2) < \frac{n_2 V}{n_1 W} < f_{1-\alpha/2}(n_1, n_2) \right]$$

$$= P\left[\frac{f_{\alpha/2}(n_1, n_2) n_1 \sum_{j=1}^{n_2} (Y_j - \mu_2)^2}{n_2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{f_{1-\alpha/2}(n_1, n_2) n_1 \sum_{j=1}^{n_2} (Y_j - \mu_2)^2}{n_2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2} \right]$$

so that $\left(\frac{f_{\alpha/2}(n_1, n_2) n_1 \sum_{j=1}^{n_2} (Y_j - \mu_2)^2}{n_2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2}, \frac{f_{1-\alpha/2}(n_1, n_2) n_1 \sum_{j=1}^{n_2} (Y_j - \mu_2)^2}{n_2 \sum_{i=1}^{n_1} (X_i - \mu_1)^2} \right)$ is a random interval having probability $1 - \alpha$ of including the fixed ,but unknown, ratio σ_2^2/σ_1^2.

After the random experiment has been performed and the values of y_j and x_i are observed, a $100(1-\alpha)\%$ confidence interval for σ_2^2/σ_1^2 is given by the interval $\left(\frac{f_{\alpha/2}(n_1,n_2)n_1\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{n_2\sum_{i=1}^{n_1}(x_i-\mu_1)^2}, \frac{f_{1-\alpha/2}(n_1,n_2)n_1\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{n_2\sum_{i=1}^{n_1}(x_i-\mu_1)^2}\right)$.

Example 9

Suppose there are two machines A and B producing a certain item and it is wished to check whether the variances of the weights of the items are the same for both machines. Random samples of $n_1 = 4$ and $n_2 = 5$ are selected with the following results:

$$A : 15, 17, 16, 16 \quad (oz.)$$

$$B : 11, 12, 9, 11, 12 \ (oz.).$$

Assume that the data are from normal distributions with means $\mu_1 = 16$ and $\mu_2 = 11$. Using the above samples, consider finding a 95% confidence interval for the ratio of the variances of the weights of the items.

From an f table,

$$f_{\alpha/2}(n_1, n_2) = f_{.025}(4,5) = 0.107$$

and

$$f_{1-\alpha/2}(n_1, n_2) = f_{.975}(4,5) = 7.388,$$

so

$$P\left[\frac{0.107 \times 4 \times \sum_{j=1}^{5}(Y_j - 11)^2}{5 \times \sum_{i=1}^{4}(X_i - 16)^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{7.388 \times 4 \times \sum_{j=1}^{5}(Y_j - 11)^2}{5 \times \sum_{i=1}^{4}(X_i - 16)^2}\right] = 0.95.$$

In this case

$$\sum_{i=1}^{4}(X_i - 16)^2 = 2,$$

and

$$\sum_{j=1}^{5}(Y_j - 11)^2 = 6.$$

Thus, the specific 95% confidence interval is

$$(0.257, 17.731),$$

so (0.257, 17.731) is a 95% confidence interval for the ratio of the variances.

The second case to consider is when μ_1 and μ_2 are unknown. As was seen in the last subsection, the independent random variables $(n_1 - 1)S_1^2/\sigma_1^2$ and $(n_2 - 1)S_2^2/\sigma_2^2$ have chi-square distributions with $n_1 - 1$ and $n_2 - 1$ degrees of freedom, respectively. Then in accordance with the definition of the Snedecor's F distribution, it is known that the random variable

$$F = \frac{(n_1 - 1)S_1^2/[\sigma_1^2(n_1 - 1)]}{(n_2 - 1)S_2^2/[\sigma_2^2(n_2 - 1)]} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$$

has an F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom. Thus it follows that

$$1 - \alpha = P[f_{\alpha/2}(n_1 - 1, n_2 - 1) < F < f_{1-\alpha/2}(n_1 - 1, n_2 - 1)]$$

$$= P\left[f_{\alpha/2}(n_1 - 1, n_2 - 1)S_2^2/S_1^2 < \frac{\sigma_2^2}{\sigma_1^2} < f_{1-\alpha/2}(n_1 - 1, n_2 - 1)S_2^2/S_1^2\right]$$

so that $(f_{\alpha/2}(n_1 - 1, n_2 - 1)S_2^2/S_1^2, f_{1-\alpha/2}(n_1 - 1, n_2 - 1)S_2^2/S_1^2)$ is a random interval having probability $1 - \alpha$ of including the fixed ,but unknown, ratio $\frac{\sigma_2^2}{\sigma_1^2}$. After the random experiment has been performed and the values of s_2^2 and s_1^2 are observed, a $100(1 - \alpha)\%$ confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$ is given by the interval $(f_{\alpha/2}(n_1 - 1, n_2 - 1)s_2^2/s_1^2, f_{1-\alpha/2}(n_1 - 1, n_2 - 1)s_2^2/s_1^2)$.

Example 10

Suppose two different processes A and B are used to manufacture light bulbs. The life of the light bulbs of process A have a normal distribution with mean μ_a and standard deviation σ_a. Similarly, for B, it is μ_b and σ_b. Suppose that it is wished to check whether the variances of the light bulb lives are the same for both processes. Let random samples yield the following data:

$SampleA$	$SampleB$
$n_a = 17$	$n_b = 21$
$\bar{x}_a = 1200hr.$	$\bar{x}_b = 1300hr.$

$$s_a = 60hr. \qquad s_b = 50hr.$$

Using the above samples, consider finding a 98% confidence interval for the ratio of the variances. From an f table,

$$f_{\alpha/2}(n_a - 1, n_b - 1) = f_{.01}(16, 20) = 0.308$$

and

$$f_{1-\alpha/2}(n_a - 1, n_b - 1) = f_{.99}(16, 20) = 3.05,$$

so

$$P\left[(0.308)S_b^2/S_a^2 < \frac{\sigma_2^2}{\sigma_1^2} < (3.05)S_b^2/S_a^2\right] = 0.98.$$

Thus, the specific 98% confidence interval is

$$(0.308 \times (50)^2/(60)^2, 3.05 \times (50)^2/(60)^2)$$

or

$$(0.214, 2.118),$$

so (0.214, 2.118) is a 98% confidence interval for the ratio of the variances.

4.2 Two-Sample Binomial Procedure

Many problems take the form of comparing two proportions. For example, there may be two processes A and B where A has P_1 and B has P_2 defective items. Suppose that it is of interest to know if $P_1 = P_2$. Then by constructing a confidence interval for $P_1 - P_2$ it is possible to check if zero is contained in the interval.

Let X_1 and X_2 be two independent random variables with binomial distributions BIN(n_1, P_1) and BIN(n_2, P_2), respectively. Thus, X_1 and X_2 can be thought of as denoting the numbers of successes observed in two independent sets of n_1 and n_2 Bernoulli trials, respectively. Because the mean and the variance of $X_1/n_1 - X_2/n_2$ are, respectively, $P_1 - P_2$ and $P_1(1 - P_1)/n_1 + P_2(1 - P_2)/n_2$, then it follows from the Central Limit Theorem that

$$\frac{(X_1/n_1 - X_2/n_2) - (P_1 - P_2)}{\sqrt{P_1(1-P_1)/n_1 + P_2(1-P_2)/n_2}} \xrightarrow{d} Z \sim N(0,1),$$

in which $\xrightarrow{d}$ denotes convergence in distribution as the sample size becomes large. Also it is known that $X_1/n_1(1 - X_1/n_1)/n_1 \xrightarrow{p} P_1(1-P_1)/n_1$ and $X_2/n_2(1 - X_2/n_2)/n_2 \xrightarrow{p} P_2(1-P_2)/n_2$, in which $\xrightarrow{p}$ denotes convergence in probablility as the sample size becomes large. Thus, dividing by

$$\left[\frac{X_1/n_1(1-X_1/n_1)/n_1 + X_2/n_2(1-X_2/n_2)/n_2}{P_1(1-P_1)/n_1 + P_2(1-P_2)/n_2}\right]^{1/2}$$

gives

$$\frac{(X_1/n_1 - X_2/n_2) - (P_1 - P_2)}{\sqrt{X_1/n_1(1-X_1/n_1)/n_1 + X_2/n_2(1-X_2/n_2)/n_2}} \xrightarrow{d} Z \sim N(0,1).$$

Let $Var_p = X_1/n_1(1 - X_1/n_1)/n_1 + X_2/n_2(1 - X_2/n_2)/n_2$,
then for large n_1 and n_2,

$$1 - \alpha \doteq P[z_{\alpha/2} < Z < z_{1-\alpha/2}]$$

$$= P\left[\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) - z_{1-\alpha/2}\sqrt{Var_p} < P_1 - P_2 < \left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) - z_{\alpha/2}\sqrt{Var_p}\right].$$

If the experimental values of X_1 and X_2 are, respectively, x_1 and x_2, then the interval $(x_1/n_1 - x_2/n_2) \pm z_{1-\alpha/2}\sqrt{x_1/n_1(1-x_1/n_1)/n_1 + x_2/n_2(1-x_2/n_2)/n_2}$ provides an approximate $100(1-\alpha)\%$ confidence interval for $P_1 - P_2$.

Example 11

Let two stochastically independent random variables Y_1 and Y_2, with binomial distributions that have parameters $n_1 = n_2 = 100, P_1$, and P_2, respectively, be observed to be equal to $y_1 = 50$ and $y_2 = 40$. Consider finding an approximate 90% confidence interval for $P_1 - P_2$.

From a normal table,

$$Z_{1-\alpha/2} = Z_{.95} = 1.645$$

so

$$P\left[(\frac{Y_1}{100}-\frac{Y_2}{100})-1.645\sqrt{Var_p}<P_1-P_2<(\frac{Y_1}{100}-\frac{Y_2}{100})+1.645\sqrt{Var_p}\right]=0.90.$$

In this case,

$$var_p = y_1/100(1-y_1/100)/100+y_2/100(1-y_2/100)/100$$

$$=50/100(1-50/100)/100+40/100(1-40/100)/100$$

$$=0.0049.$$

Thus, the specific approximate 90% confidence interval is

$$((50/100-40/100)-1.645\times 0.07,(50/100-40/100)+1.645\times 0.07)$$

or

$$(-0.015, 0.215),$$

so (-0.015, 0.215) is an approximate 90% confidence interval for the difference of the proportions.

5. METHODS OF FINDING CONFIDENCE INTERVALS

Through out this paper, confidence intervals have been constructed simply by using properties of CDF's. That is, first a random variable is constructed with a known distribution which doesn't depend on any unknown nuisance parameters. Then with that CDF, the following useful result is used, namely

$$P[a<x\leq b]=F(b)-F(a).$$

In most instances, $F(b)-F(a)$, is set to equal $1-\alpha$, and a and b are, respectively, the $\alpha/2$ and $1-\alpha/2$ quantiles of the distribution (the equal tailed choice). After a and b are determined by α, then the inequality $\{a<x\leq b\}$ in the probability statement can be rewritten, or inverted or "pivoted," to yield a probability statement for the parameter of interest. This technique for finding confidence intervals is called the *pivotal-quantity method* and will be more formally defined in the next subsection.

There is a problem with the technique described in the last paragraph. The problem being that it isn't always possible to find a random variable with a known distribution which doesn't depend on any unknown parameters. Hence, two other methods are presented in this section which circumvent this problem.

5.1 Pivotal-Quantity Method

As before, assume that $X_1, ..., X_n$ is a random sample of size n from the pdf $f(x;\theta)$ parametrized by θ. The goal is to find a confidence interval estimate of θ where other nuisance parameters may also be present.

Definition 3: Pivotal-Quantity

> A function $u(X_1, ..., X_n; \theta)$ whose distribution does not depend on any unknown parameters is known as a **pivot function** for θ. If $U = u(X_1, ..., X_n; \theta)$ is a random variable whose distribution does not depend on any unknown parameters, then U is called a **pivotal-quantity**.

Notice in the definition that the word "If" is used. This is because a pivotal-quantity doesn't always exist and a different method must be used. Examples of distributions which don't yield pivotal-quantities will be given in the next subsection.

Example 12

$X_1, ..., X_n$ is a random sample of size n from the exponential distribution with pdf

$$f_X(x;\eta) = e^{-(x-\eta)}\, I_{(\eta,\infty)}(x).$$

I will show that $Q = X_{1:n} - \eta$ is a pivotal-quantity and find its distribution.

The CDF of X_i is

$$F_X(x;\eta) = \int_{\eta}^{x} f_X(t;\eta)\, dt = \left[1 - e^{-(x-\eta)}\right] I_{(\eta,\infty)}(x).$$

The CDF of $X_{1:n}$ is

$$G_1(x_{1:n};\eta) = 1 - [1 - F_X(x_{1:n};\eta)]^n$$

$$= 1 - \left[1 - 1 + e^{-(x_{1:n} - \eta)}\right]^n$$

$$= \left[1 - e^{-n(x_{1:n} - \eta)}\right] I_{(\eta,\infty)}(x_{1:n}).$$

Thus,

$$F_Q(q) = P[Q \leq q]$$

$$= P[X_{1:n} - \eta \leq q]$$

$$= P[X_{1:n} \leq q + \eta]$$

$$= G_1(q + \eta)$$

$$= 1 - e^{-n(q+\eta-\eta)} = [1 - e^{-nq}]\, I_{(\eta,\infty)}(q + \eta)$$

$$= [1 - e^{-nq}]\, I_{(0,\infty)}(q)$$

so

$$f_Q(q) = \frac{dF}{dq} = ne^{-nq} I_{(0,\infty)}(q).$$

This implies

$$Q = X_{1:n} - \eta \sim \text{EXP}(1/n).$$

$Q = X_{1:n} - \eta$ is a random variable that is a function only of $X_1, ..., X_n$ and η, and its distribution ($\text{EXP}(1/n)$) does *not* depend on η or any other unknown parameters. Hence by the definition of pivotal-quantities, Q is a pivotal-quantity.

If $U = u(X_1, ..., X_n; \theta)$ is a pivotal-quantity, then for any fixed $0 < \gamma < 1$ there will exist values a and b depending of γ such that $P[a < U < b] = \gamma$. Now, if for each possible observed sample $(x_1, ..., x_n), a < u(x_1, ..., x_n; \theta) < b$ if and only if $\theta_l(x_1, ..., x_n) < \theta < \theta_u(x_1, ..., x_n)$ for sample statistics θ_l and θ_u (based on U, but not depending on θ), then the interval $(\theta_l(x_1, ..., x_n), \theta_u(x_1, ..., x_n))$ is a $100\gamma\%$ confidence interval for θ. Notice that different pairs of numbers a and b will produce different pairs of θ_l and θ_u. Hence, there are an infinite number of possible intervals with the same probability (see page 5 where this concept was noticed before).

Example 13

I will continue with example 12 and derive a $100\gamma\%$ equal-tailed confidence interval for η.

Values of q_1 and q_2 need to be found so that

$$\gamma = 1 - \alpha = P[q_1 < X_{1:n} - \eta < q_2]$$

$$= P[X_{1:n} - q_2 < \eta < X_{1:n} - q_1].$$

Recall from example 12,

$$F_Q(q) = [1 - e^{-nq}]I_{(\eta,\infty)}(q).$$

If an equal-tailed confidence interval is desired, values of q_1 and q_2 can be found by solving the following two equations

$$F_Q(q_1) = \alpha/2 \quad \text{and} \quad F_Q(q_2) = 1 - \alpha/2.$$

Substituting in the CDF's, the solutions are

$$1 - e^{-nq_1} = \alpha/2 \quad \text{and} \quad 1 - e^{-nq_2} = 1 - \alpha/2$$

so

$$q_1 = \left(-\frac{1}{n}\right)\ln(1-\alpha/2) \quad \text{and} \quad q_2 = \left(-\frac{1}{n}\right)\ln(\alpha/2).$$

Hence, a $100\gamma\%$ equal-tailed confidence interval for η is given by

$$(x_{1:n} - q_2, x_{1:n} - q_1)$$

$$= \left(x_{1:n} + \left(\frac{1}{n}\right)\ln(\alpha/2), x_{1:n} + \left(\frac{1}{n}\right)\ln(1-\alpha/2)\right)$$

where $\gamma = 1 - \alpha$.

5.2 Statistical (General) Method

If a pivotal-quantity is not available, it may still be possible to find a confidence interval for a parameter θ. Specifically, let $X_1, ..., X_n$ be a random sample of size n from the pdf $f(x; \theta)$, and let $S = s(X_1, ..., X_n)$ be some statistic. Also, let $f_S(s; \theta)$ denote the pdf of S. The statistic S can be selected in many ways. For instance, S could be taken to be a sufficient statistic for θ, if one exists, or a point estimator such as the maximum-likelihood estimator of θ. Although this method works for S as a discrete random variable, I will assume that S is continuous to more clearly explain this technique. Now, for each possible value of θ, assume that two strictly monotone increasing functions, say $h_1(\theta)$ and $h_2(\theta)$, can be defined such that

$$\int_{-\infty}^{h_1(\theta)} f_S(s; \theta)\, ds = \alpha_1 \quad \text{and} \quad \int_{h_2(\theta)}^{\infty} f_S(s; \theta)\, ds = \alpha_2, \tag{5.1}$$

where α_1 and α_2 are constants satisfying $0 < \alpha_1, 0 < \alpha_2$, and $\alpha_1 + \alpha_2 < 1$.

If $S = s$ is an observed value of S, then values $\theta_L = \theta_L(s)$ and $\theta_U = \theta_U(s)$ are found by solving

$$h_2(\theta_L) = s \quad \text{and} \quad h_1(\theta_U) = s.$$

The above notation is used because θ_L and θ_U are both functions of s.

Thus, if θ_0 is the true value of θ, then $h_1(\theta_0) < s < h_2(\theta_0)$ if and only if $\theta_L = \theta_L(s) < \theta_0 < \theta_U = \theta_U(s)$ for any observed sample $x_1, ..., x_n$. But following from the definitions of $h_1(\theta)$ and $h_2(\theta)$,

$$P[h_1(\theta_0) < S < h_2(\theta_0)] = 1 - \alpha_1 - \alpha_2,$$

so

$$P[\theta_L(S) < \theta_0 < \theta_U(S)] = 1 - \alpha_1 - \alpha_2.$$

Hence, (θ_L, θ_U) is a $100(1 - \alpha_1 - \alpha_2)\%$ confidence interval for θ_0.

Observe that $h_1(\theta)$ and $h_2(\theta)$ are really not needed. For a given observed value s of the statistic $S, \theta_L = \theta_L(s)$ and $\theta_U = \theta_U(s)$ need to be found. θ_U can be found by solving for θ in the equation

$$\alpha_1 = \int_{-\infty}^{h_1(\theta)=s} f_S(s; \theta)\, ds. \tag{5.2}$$

θ_L can be found by solving for θ in the equation

$$\alpha_2 = \int_{h_2(\theta)=s}^{\infty} f_S(s;\theta)\,ds. \tag{5.3}$$

Example 14

Consider again a random sample from an exponential distribution with pdf

$$f_X(x;\theta) = \frac{1}{\theta} e^{-x/\theta} I_{(0,\infty)}(x).$$

I will find a confidence interval for θ by using the statistic $S = \sum_{i=1}^{n} X_i$. For given α_1 and α_2, $h_1(\theta)$ and $h_2(\theta)$ can be found. Because $2S/\theta \sim \chi^2(2n)$, it follows that

$$\alpha_1 = P[S \leq h_1(\theta)]$$

$$= P[2S/\theta \leq 2h_1(\theta)/\theta]$$

which implies

$$2h_1(\theta)/\theta = \chi^2_{\alpha_1}(2n)$$

and

$$h_1(\theta) = \theta\chi^2_{\alpha_1}(2n)/2.$$

Similarly,

$$\alpha_2 = P[S \geq h_2(\theta)]$$

$$= 1 - P[2S/\theta \leq 2h_2(\theta)/\theta]$$

which implies

$$2h_2(\theta)/\theta = \chi^2_{1-\alpha_2}(2n)$$

and

$$h_2(\theta) = \theta\chi^2_{1-\alpha_2}(2n)/2.$$

For observed $s = \sum_{i=1}^{n} x_i, \theta_L$ is such that $h_2(\theta_L) = s$; that is, $h_2(\theta_L) = \theta_L\chi^2_{1-\alpha_2}(2n)/2 = s$ or $\theta_L = 2s/\chi^2_{1-\alpha_2}(2n)$. Similarly, $\theta_U = 2s/\chi^2_{\alpha_1}(2n)$. So a $100(1-\alpha_1-\alpha_2)\%$ confidence interval for θ is given by $\left(\frac{2\sum_{i=1}^{n} x_i}{\chi^2_{1-\alpha_2}(2n)}, \frac{2\sum_{i=1}^{n} x_i}{\chi^2_{\alpha_1}(2n)}\right)$.

I mentioned at the beginning of this subsection that this method would work for discrete random variables as well as for continuous random variables. The only difference is that now the integrals in (5.1) to (5.3) need to be replaced by summations. Two popular discrete density functions are the Binomial and Poisson. One may be interested in confidence interval estimates of the parameters in each. The next two examples consider these two discrete density functions, respectively.

Example 15

$X_1, ..., X_n$ is a random sample of size n from the Bernoulli distribution with pdf

$$f_X(x;p) = p^x(1-p)^{1-x} I_{\{0,1\}}(x).$$

Consider finding a $100(1-\alpha)\%$ confidence interval for p.

It is known that $S = \sum_{i=1}^{n} X_i$ has a binomial distribution; that is, $P[S=s] = \binom{n}{s} p^s(1-p)^{n-s}$ for $s = 0, 1, ..., n$. I will not find explicit expressions for $h_1(p)$ and $h_2(p)$ in this example. Suppose $S = s_o$ is observed (necessarily an integer). Then p_U and p_L need to be solved for in each of the equations

$$\alpha_1 = \sum_{s=0}^{s_0} \binom{n}{s} p_U^s (1-p_U)^{n-s}$$

and

$$1-\alpha_2 = \sum_{s=0}^{s_0-1} \binom{n}{s} p_L^s (1-p_L)^{n-s}.$$

If $\alpha = \alpha_1 + \alpha_2$, a $100(1-\alpha)\%$ confidence interval for p is given by (p_L, p_U). To actually solve these equations, a computer program to evaluate the binomial distribution is useful. If $\alpha_1 = \alpha_2 = 0.05, n = 10$, and if $s = 2$ is observed then $p_L = 0.037$ and $p_U = 0.507$. It follows that (0.037, 0.507) is a conservative 90% confidence interval for p.

Example 16

$X_1, ..., X_n$ is a random sample of size n from the Poisson distribution with pdf

$$f_X(x;\mu) = \frac{e^{-\mu}\mu^x}{x!} I_{\{0,1,...\}}(x).$$

Consider finding a $100(1-\alpha)\%$ confidence interval for μ. It is known that $S = \sum_{i=1}^{n} X_i$ has a Poisson distribution with parameter $n\mu$. I will again not find explicit expressions for $h_1(\mu)$ and $h_2(\mu)$ in this example. The CDF of the Poisson distribution is related to the CDF of a chi-square distribution by $F_X(x;\mu) = 1 - H[2\mu; 2(x+1)]$ (Bain and Engelhardt 1989, p. 227) where $F_X(x;\mu)$ is the CDF of the Poisson. Thus, the confidence limits can easily be expressed in terms of chi-square percentiles. If I denote the CDF of a chi-square distribution with ν degrees-of- freedom by H($x;\nu$), then for an observed value $S = s$, μ_U and μ_L need to be solved for in each of the equations

$$\alpha_1 = 1 - H(2n\mu_U; 2s+2)$$

and

$$1 - \alpha_2 = 1 - H(2n\mu_L; 2s).$$

Thus,

$$2n\mu_U = \chi^2_{1-\alpha_1}(2s+2),$$

and thus

$$\mu_U = \chi^2_{1-\alpha_1}(2s+2)/2n.$$

Similarly,

$$2n\mu_L = \chi^2_{\alpha_2}(2s),$$

and

$$\mu_L = \chi^2_{\alpha_2}(2s)/2n.$$

If $\alpha = \alpha_1 + \alpha_2$, then a $100(1-\alpha)\%$ confidence interval for μ is given by $(\mu_L, \mu_U) = (\chi^2_{\alpha_2}(2\sum_{i=1}^{n} x_i)/2n, \chi^2_{1-\alpha_1}(2\sum_{i=1}^{n} x_i + 2)/2n)$.

5.3 Bootstrap Method

This subsection briefly describes the basis of the bootstrap and presents three closely related methods of using the bootstrap to find confidence intervals. Only the mechanics of these methods will be considered and not the theory behind them.

The bootstrap (Efron 1979, 1981, 1982) is widely viewed as a tool that can be used to find nonparametric confidence intervals in complex problems. The bootstrap is a general methodology for measuring the accuracy of an estimator. It is a computer-based method, which substitutes large amounts of computation in place of theoretical analysis. The bootstrap can be used to answer questions which are too complicated for traditional statistical analysis. In an era of declining computational costs, computer-intensive methods such as the bootstrap are becoming increasingly useful even for relatively simple problems.

Suppose $X_1, ..., X_n$ are independent and identically distributed (iid) random variables from a population with unknown CDF F, and suppose the goal is to draw inferences about some parameter θ of the population. Let $\hat{\theta}$ be an estimator of θ and let $\hat{F}$ indicate the empirical probability distribution, the CDF that assigns mass $1/n$ to each X_i. The bootstrap approximates the sampling distribution of $\hat{\theta}$ under F by the sampling distribution of $\hat{\theta}$ under $\hat{F}$. This procedure is usually hard to carry out analytically, and it is often necessary to use the Monte Carlo algorithm as follows (Efron 1981, 1982):

1. Construct $\hat{F}$.
2. Draw a "bootstrap sample" from $\hat{F}$,

$$x_1^*, ..., x_n^* \overset{iid}{\sim} \hat{F},$$

and calculate the bootstrap estimate $\hat{\theta}^* = \hat{\theta}(x_1^*, ..., x_n^*)$.

3. Independently repeat step 2 B times (for some large B), obtaining $\hat{\theta}_b^*, b = 1, ..., B$. The CDF of the bootstrap distribution of $\hat{\theta}^*$ at y is approximated by $\hat{G}(y) = \#\{\hat{\theta}_b^* \leq y\}/B$, the number of bootstrap estimates less than or equal to y divided by the total number of replications.

I will present three different methods of using the bootstrap to set confidence intervals in order of increasing generality in respect to the characteristics of the sampling distribution such as bias, skewness, etc. All three methods use percentiles of $\hat{G}$ to define the confidence interval. They differ in which percentiles are used.

For a given α between 0 and .5, the simplest method is to define

$$\hat{\theta}_{LOW}(\alpha) = \hat{G}^{-1}(\alpha), \quad \hat{\theta}_{UP}(\alpha) = \hat{G}^{-1}(1-\alpha),$$

usually denoted simply $\hat{\theta}_{LOW}$, $\hat{\theta}_{UP}$. Efron's (1981, 1982) *percentile method* consists of taking

$$\theta \in [\hat{\theta}_{LOW}(\alpha), \hat{\theta}_{UP}(\alpha)] \tag{5.4}$$

as an approximate $100(1-2\alpha)\%$ central confidence interval for θ. Because $\alpha = \hat{G}(\hat{\theta}_{LOW})$ and $1-\alpha = \hat{G}(\hat{\theta}_{UP})$, the percentile method interval consists of the central $1-2\alpha$ proportion of the bootstrap distribution. Thus this method does not perform well for biased or skewed sampling distributions.

The bootstrap distribution for the sample median is median unbiased in the sense that $\hat{G}(x_{(m)}) = .50$ (splitting the probability at the sample median $x_{(m)}$, the middle order statistic). If $\hat{G}(\hat{\theta}) \neq .50$ then a bias-correction (Efron 1982) to the percentile method is called for. To be specific, define the bias-correction constant

$$z_0 = \Phi^{-1}(\hat{G}(\hat{\theta}))$$

where Φ is the standard normal CDF. Efron's (1981, 1982) *bias-corrected percentile method* (BC method) consists of taking

$$\theta \in [\hat{G}^{-1}(\Phi(2z_0 + z_\alpha)), \hat{G}^{-1}(\Phi(2z_0 + z_{1-\alpha}))] \tag{5.5}$$

as an approximate $100(1-2\alpha)\%$ central confidence interval for θ. Here z_α is the α^{th} percentile of the standard normal distribution.

Notice if $\hat{G}(\hat{\theta}) = .50$, that is if half of the bootstrap distribution of $\hat{\theta}^*$ is less than the observed value $\hat{\theta}$, then $z_0 = 0$ and (5.5) reduces to (5.4), the uncorrected percentile interval. However, even small differences of $\hat{G}(\hat{\theta})$ from .50 can make (5.5) much different from (5.4) as was shown in Efron and Tibsharani's (1986) law school example in section 2.

Even though the BC method works well for biased data, it does not perform well for skewed sampling distributions. For some skewed sample distributions the BC method only goes about half as far as it should toward achieving the asymmetry of the exact interval. Thus another constant is needed to help adjust for skewed distributions. To be specific, define the *"acceleration constant"*

$$a \doteq \frac{1}{6} \frac{\sum_{i=1}^{n} (U_i^0)^3}{\{\sum_{i=1}^{n} (U_i^0)^2\}^{3/2}},$$

where U_i^0 is the empirical influence function,

$$U_i^0 = \lim_{\epsilon \to 0} \frac{\hat{\theta}(\hat{F}[(1-\epsilon)\mathbf{p}^0 + \epsilon \boldsymbol{\delta}_i]) - \hat{\theta}(\hat{F}[\mathbf{p}^0])}{\epsilon}.$$

$\boldsymbol{\delta}_i$ is the ith coordinate vector, $\mathbf{p}^0$ is the vector $(1/n, 1/n, ..., 1/n)$, $\hat{F}[p_1, p_2, ..., p_n]$ is the weighted empirical distribution $\hat{F}[\mathbf{p}]$: probability p_i on $x_i i = 1, ..., n$, and $\hat{\theta}(\hat{F}[\mathbf{p}])$ is the bootstrap estimate obtained from using $\hat{F}[\mathbf{p}]$ in the Monte Carlo algorithm.

Efron's (1984) *bias-corrected percentile acceleration method* (BC_a method) consists of taking

$$\theta \in [\hat{G}^{-1}(\Phi(z[\alpha])), \hat{G}^{-1}(\Phi(z[1-\alpha]))], \tag{5.6}$$

where

$$z[\alpha] = z_0 + \frac{(z_0 + z_\alpha)}{1 - a(z_0 + z_\alpha)}, \tag{5.7}$$

and likewise for $z[1-\alpha]$, as an approximate $100(1-2\alpha)\%$ central confidence interval for θ.

If z_0 and a equal zero, then $z[\alpha] = z_\alpha$ and (5.6) becomes (5.4), the percentile method. In general z_0 and a do not equal zero, and formulas (5.6) and (5.7) make adjustments to the percentile method that are necessary to achieve higher order accuracy. Hence, this method appears to perform well for most all sampling distributions.

Suppose that a is set to equal 0 in (5.7), so $z[\alpha] = 2z_0 + z_\alpha$. Interval (5.6) with this definition of $z[\alpha]$ and $z[1-\alpha]$ becomes (5.5), the BC method. In other words, BC = BC_a, with $a = 0$. The constant z_0 is easier to obtain than the constant a which is why the BC method might be used.

To summarize this subsection, the progression from the percentile method to the BC_a method is based on a series of increasingly less restrictive assumptions about the sampling distribution such as no bias or no skewness. Each successive method requires more computation; first the bootstrap distribution $\hat{G}$, then the bias-correction constant z_0, and finally the constant a. However, all of these computations are algorithmic in character, and can be carried out in a somewhat automatic fashion.

6. SUMMARY

In summary, the purpose of this paper was to introduce the concept of an interval estimate or confidence interval. A point estimator, by itself, does not provide direct information about accuracy. An interval estimator gives one possible solution to this problem. The concept involves an interval the endpoints of which are statistics that include the true value of the parameter between them with a certain probability. This probability corresponds to the confidence level or significance level of the interval estimator. Ordinarily, the term confidence interval (or interval estimate) refers to the observed interval that is computed from data.

Three basic methods for constructing confidence intervals were discussed. The first, which is useful in certain applications where unknown nuisance parameters are present, involves the idea of a pivotal quantity. This amounts to finding a random variable that is a function of the observed random variables and the parameter of interest, but not of any other unknown parameters. It is also required that the distribution of the pivotal-quantity be free of any unknown parameters.

The second method, which is referred to as the general or statistical method, does not require the existence of a pivotal-quantity, but has the disadvantage that it cannot be used when a nuisance parameter is present. This method can be applied with any statistic whose distributions can be expressed in terms of the parameter. The percentiles are functions of the parameter, and the limits of the confidence interval are obtained by solving equations that involve certain percentiles and the observed value of the statistic.

The third method involves using the bootstrap. Actually three different methods were discussed; the percentile, the bias- corrected percentile, and the bias-corrected percentile acceleration. Each of these methods employ the bootstrap, however, they differ in that they are based on a series of increasingly less restrictive assumptions about the sampling distribution. Also each successive method requires more computation.

Interval estimates obtained by any of the three methods can be interpreted in terms of the relative frequency with which the true value of the parameter will be included in the interval, which corresponds to the probability that the interval estimator will contain the true value. Finally, I should mention that even though the methods discussed in this paper are perhaps the most common methods, they are not the only methods. For example, in some instances it may be known that a parameter in non-negative. Hence, it makes good sense to set the lower limit to be zero.

References

- Bain, Lee J. and Engelhardt, Max (1989), *Introduction to Probability and Mathematical Statistics*, Boston: PWS-KENT.
- Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.
- Efron, B. (1981), "Nonparametric Standard Errors and Confidence Intervals," *Canadian Journal of Statistics*, 9, 139-172.
- Efron, B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, National Science Foundation-Conference Board of the Mathematical Sciences Monograph 38, Philadelphia: Society for Industrial and Applied Mathematics.
- Efron, B. (1984), "Better bootstrap confidence intervals," Technical Report, Stanford University Department of Statistics.
- Efron, B. and Tibshirani, R. (1986), "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," *Statistical Science*, 1, 54-77.
- Hogg, Robert V. and Craig, Allen T. (1970), *Introduction to Mathematical Statistics*, New York: The Macmillan Co. Inc.
- Mood, Alexander M., Graybill, Franklin A. and Boes, Duane C. (1974), *Introduction to the Theory of Statistics*, New York: McGraw-Hill, Inc.