The Principles of Interval Estimation and Hypothesis Testing

1. Introduction

In *Statistical Inference I* we described how to estimate the mean and variance of a population, and the properties of those estimation procedures. In *Statistical Inference II* we introduce two more aspects of statistical inference: **confidence intervals** and **hypothesis tests**. In contrast to a **point estimate** of the population mean β, like b = 17.158, a confidence interval estimate is a range of values which may contain the true population mean. A confidence interval estimate contains information not only about the location of the population mean but also about the precision with which we estimate it. A hypothesis test is a statistical procedure for using data to check the compatibility of a conjecture about a population with the information contained in a sample of data. Continuing the example from *Statistical Inference I*, suppose airplane designers have been basing seat designs based on the assumption that the average hip width of U.S. passengers is 16 inches. Is the information contained in the random sample of 50 hip measurements compatible with this conjecture, or not? These are the issues we consider in *Statistical Inference II*.

2. Interval Estimation for Mean of Normal Population When σ^2 is Known

Let Y be a random variable from a normal population. That is, assume $Y \sim N\left(\beta, \sigma^2\right)$. Assume that we have a random sample of size T from this population, $Y_1, Y_2, \ldots, Y_T$. The least squares estimator of the population mean is

$$b = \sum_{i=1}^{T} Y_i / T \tag{2.1}$$

This estimator has a normal distribution if the population is normal,

$$b \sim N\left(\beta, \sigma^2 / T\right) \tag{2.2}$$

For the present, let us assume that the population variance σ^2 is known. This assumption is not likely to be true, but making it allows us to introduce the notion of confidence intervals with few complications. In the next section we introduce methods for the case when σ^2 is unknown.

We can create a standard normal random variable from (2.2) by subtracting the mean and dividing by the standard deviation,

$$Z = \frac{b - \beta}{\sqrt{\sigma^2 / T}} = \frac{b - \beta}{\sigma / \sqrt{T}} \sim N(0,1) \tag{2.3}$$

The standard normal random variable Z has mean 0 and variance 1. That is, $Z \sim N(0,1)$. Let z_c be a "critical value" for the standard normal distribution, such that α = .05 of the probability is in the tails of the distribution, with $\alpha/2$ = .025 of the probability in each tail. From Table 1 at the end of *UE/2* the value of z_c = 1.96 when α = .05. This critical value is illustrated in Figure 1.

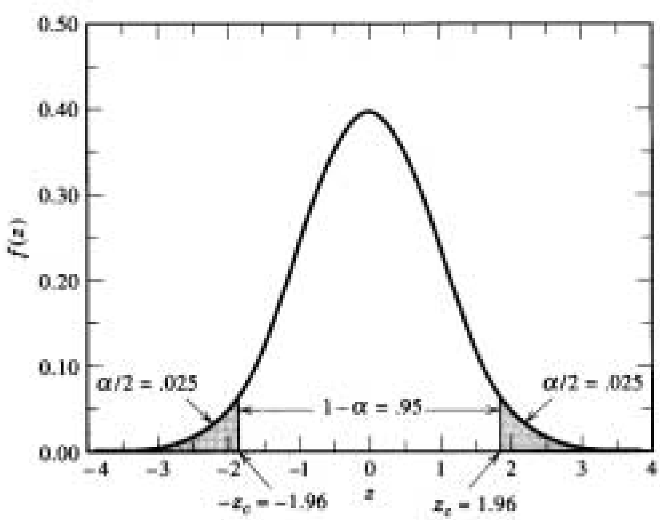

Figure 1 α = .05 critical values for the $N(0,1)$ distribution

Thus

$$P[Z \geq 1.96] = P[Z \leq -1.96] = 0.025 \tag{2.4}$$

and

$$P[-1.96 \leq Z \leq 1.96] = 1 - .05 = .95 \tag{2.5}$$

Substitute (2.3) into (2.5) to obtain

$$P\left[-1.96 \leq \frac{b-\beta}{\sigma/\sqrt{T}} \leq 1.96\right] = .95 \tag{2.6}$$

Multiplying through the inequality inside the brackets by $\sigma/\sqrt{T}$ yields

$$P\left[-1.96\sigma/\sqrt{T} \leq b-\beta \leq 1.96\sigma/\sqrt{T}\right] = .95 \tag{2.7}$$

Subtracting b from each of the terms inside the brackets gives

$$P\left[-b-1.96\sigma/\sqrt{T} \leq -\beta \leq -b+1.96\sigma/\sqrt{T}\right] = .95 \tag{2.8}$$

Multiplying by −1 within the brackets reverses the direction of the inequalities giving

$$P\left[b-1.96\sigma/\sqrt{T} \leq \beta \leq b+1.96\sigma/\sqrt{T}\right] = .95 \tag{2.9}$$

In general,

$$P\left[b - z_c \frac{\sigma}{\sqrt{T}} \le \beta \le b + z_c \frac{\sigma}{\sqrt{T}}\right] = 1 - \alpha \tag{2.10}$$

where z_c is the appropriate critical value for a given value of tail probability α. In (2.10) we have defined the **interval estimator**

$$b \pm z_c \, \sigma / \sqrt{T} \tag{2.11}$$

Our choice of the phrase interval *estimator* is a careful one. The interval (2.11) defines a procedure that can be used for any sample of data. The interval endpoints are thus random variables. What (2.10) implies is that intervals constructed using (2.11), in repeated sampling from the population, have a $100(1-\alpha)\%$ chance of containing the population mean β.

(2.1) An Example Using Artificial Data

In order to use the interval estimation procedure defined in (2.11) we must have data from a normal population with a known variance. To illustrate the computation, and the meaning of interval estimation, we will create a sample of data using a computer simulation. Statistical software programs contain **random number generators**. These are routines that create values from a given probability distribution. Table 1 contains 30 values from a normal population with mean $\beta = 10$ and variance $\sigma^2 = 10$.

Table 1 30 values from N(10,10)		
11.939	11.407	13.809
10.706	12.157	7.443
6.644	10.829	8.855
13.187	12.368	9.461
8.433	10.052	2.439
9.210	5.036	5.527
7.961	14.799	9.921
14.921	10.478	11.814
6.223	13.859	13.403
10.123	12.355	10.819

Table 2 contains the least squares estimates and the lower and upper interval estimate values based on 10 samples like the one in Table 1.

Table 2 Results from 10 samples of data

Sample	*b*	*lower bound*	*upper bound*
1	10.206	9.074	11.338
2	9.828	8.696	10.959
3	11.194	10.062	12.326
4	8.822	7.690	9.953
5	10.434	9.302	11.566
6	8.855	7.723	9.986
7	10.511	9.380	11.643
8	9.212	8.080	10.343
9	10.464	9.333	11.596
10	10.142	9.010	11.273

Table 2 illustrates the sampling variation of the least squares estimator b. The sample means vary from sample to sample. In this simulation, or Monte Carlo, experiment we know the true population mean, $\beta = 10$, and the estimates b are centered at that value. The width of the interval estimates is $1.96\sigma\big/\sqrt{T}$. Note that while the point estimates b in Table 2 fall near the true value $\beta = 10$, not all of the interval estimates contain the true value. Intervals from samples 3, 4 & 6 do not contain the true value $\beta = 10$. However, in 10,000 simulated samples the average value of $b = 10.004$ and 0.9486% of intervals constructed using (2.11) contain the true parameter value $\beta = 10$.

These numbers reveal what is, and what is not, true about interval estimates.

- Any one interval estimate may or may not contain the true population parameter value.
- If *many* samples of size T are obtained, and intervals are constructed using the interval estimation procedure (2.11) with $(1-\alpha) = .95$, then 95% of them will contain the true parameter value.
- A 95% level of "confidence" represents the confidence (the probability that the interval estimator will provide an interval containing the true parameter value) we have in the procedure, not in any one interval estimate.
- Since 95% of intervals constructed using (2.11) will contain the true parameter $\beta = 10$, we will be surprised if an interval estimate based on one sample does not contain the true parameter. Indeed, the fact that 3 of the 10 intervals in Table 2 do not contain $\beta = 10$ is surprising, since out of 10 we would assume that only 1 95% interval estimate might not contain the true parameter. This just goes to show that what happens in any one sample,

or just a few samples, is not what statistical sampling properties tell us. Sampling properties tell us what happens in many repeated experimental trials.