2: CONFIDENCE INTERVALS FOR THE MEAN; UNKNOWN VARIANCE

Now, we suppose that $X_1, \ldots, X_n$ are *iid* with unknown mean μ and *unknown* variance σ^2. Clearly, we will now have to estimate σ^2 from the available data. The most commonly-used estimator of σ^2 is the sample variance,

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

The reason for using the $n-1$ in the denominator is that this makes S_x^2 an **unbiased** estimator of σ^2. In other words, $E[S_x^2] = \sigma^2$. We will prove this later. A proof in the normal case follows from Section 4.8 of Hogg & Craig.

Note: Hogg and Craig use a denominator of n in their S^2. We, most textbooks, and most practitioners, however, use $n-1$. To minimize confusion, we will try for now to avoid using the symbol S^2.

• Question: What would happen if we used S_x in place of σ in the formula for the CI?

• Answer: It depends on whether the sample size is "large" or not.

Large-Sample Confidence Interval; Population Not Necessarily Normal

Theorem: The interval $\bar{X} \pm z_{\alpha/2} \frac{S_x}{\sqrt{n}}$ is an asymptotic level $1-\alpha$ CI for μ.

In other words, when the sample size is large, we can use S_x in place of the unknown σ, and the CI will still work.

Proof: It can be shown that S_x^2 converges in probability to σ^2. In other words,

$$\lim_{n \to \infty} Pr\,(|S_x^2 - \sigma^2| > \varepsilon) \to 0 \text{ for any } \varepsilon > 0.$$

As a result, the distribution of

$$\frac{\bar{X} - \mu}{S_x / \sqrt{n}}$$

converges to the standard normal distribution. Similarly to the proof from the previous handout, we get

$$Pr(\text{CI Contains } \mu) = Pr(-z_{\alpha/2} < \frac{\bar{X} - \mu}{S_x / \sqrt{n}} < z_{\alpha/2}) \to 1 - \alpha \ .$$

Small-Sample Confidence Interval;

Normal Population

- If the sample size is small (the usual guideline is $n \le 30$), and σ is unknown, then to assure the validity of the CI we will present here, we must assume that the population distribution is normal. This assumption is hard to check in small samples!

• The CI is $\bar{X} \pm t_{\alpha/2} \dfrac{S_x}{\sqrt{n}}$. ($t_{\alpha/2}$ is defined below.)

The Basics of t Distributions

When n is small, the quantity $t = \dfrac{\bar{X} - \mu}{S_x / \sqrt{n}}$ does not have a normal distribution, even when the population is normal.

Instead, t has a "Student's t distribution with $n-1$ degrees of freedom".

There is a different t distribution for each value of the degrees of freedom, ν.

The quantity $t_{\alpha/2}$ denotes the t-value such that the area to its right under the Student's t distribution (with $\nu = n - 1$) is $\alpha/2$. Note that we use $\nu = n - 1$, even though the sample size is n. Values of t_α are listed in Table 2, Page 599 of Jobson.

• Note that the last row of Table 2 is denoted by "∞". For practical purposes, any value of ν beyond 29 is usually considered "infinite". (Most tables stop at $\nu = 29$. Jobson's table is somewhat better, since he also has entries for $\nu =$ 30, 40, 50, 60, and 120.) In this case, the corresponding t distribution is essentially identical to the standard normal distribution. Here, it doesn't matter whether we use the CI $\bar{X} \pm t_{\alpha/2} \frac{S_x}{\sqrt{n}}$ or $\bar{X} \pm z_{\alpha/2} \frac{S_x}{\sqrt{n}}$ since they will be almost the same. Since t is asymptotically standard normal, the t_{α} values given in the "∞" row of Table 2 are identical to the z_{α} values defined earlier.

• On the other hand, if $\nu \leq 29$ the t distribution has "longer tails" (i.e., contains more outliers) than the normal distribution, and it is important to use the t–values of Table 2, assuming that σ is unknown. Here, the CI based on $t_{\alpha/2}$ will be wider than the (incorrect) one based on $z_{\alpha/2}$.

(Why does this happen, and why does it make sense?)

Eg 1: A random sample of 8 "Quarter Pounders" yields a mean weight of $\bar{x} = .2$ pounds, with a sample standard deviation of $s_x = .07$ pounds. Construct a 95% CI for the unknown population mean weight for all "Quarter Pounders".

Background: Definitions of χ^2 and t distributions

As in Section 1.3.3 of Jobson, we define the χ^2 distribution with ν degrees of freedom to be the distribution of the random variable $\chi^2_\nu = \sum_{i=1}^{\nu} Z_i^2$, where $Z_1, \cdots, Z_\nu$ are *iid* standard normal. The distribution is positive valued and is skewed to the right. The mean and variance are $E[\chi^2_\nu] = \nu$, $var[\chi^2_\nu] = 2\nu$. If $X_1, \ldots, X_n$ are *iid* $N(\mu, \sigma^2)$, then it can be shown that $(n-1)S_x^2/\sigma^2$ has a χ^2_{n-1} distribution.

Therefore, $S_x^2 \sim \sigma^2 \chi_{n-1}^2/(n-1)$, and we find that $E[S_x^2] = \sigma^2$, so that S_x^2 is unbiased for σ^2.

- The random variable

$$\frac{Z}{\sqrt{\chi_\nu^2/\nu}}$$

is said to have a t *distribution with* ν *degrees of freedom* if Z is standard normal and χ_ν^2 is independent of Z and has a χ_ν^2 distribution.

Establishing the Small-Sample CI

Theorem: If $X_1, \ldots, X_n$ are *iid* $N(\mu, \sigma^2)$, then the interval $\bar{X} \pm t_{\alpha/2} \frac{S_x}{\sqrt{n}}$ is a level $1-\alpha$ CI for μ.

Proof: It can be shown that $\bar{X}$ and S_x^2 are independent. (We will prove this later).

Define $Z = \sqrt{n}\,(\bar{X} - \mu)/\sigma$, which is standard normal.

Define $\chi^2_{n-1} = (n-1)S_x^2/\sigma^2$, which has a χ^2_{n-1} distribution. Define

$$t = \frac{Z}{\sqrt{\chi^2_{n-1}/(n-1)}} = \frac{\sqrt{n}\,(\bar{X} - \mu)}{S_x} = \frac{\bar{X} - \mu}{S_x/\sqrt{n}} \quad .$$

By its definition, t has a t distribution with $n-1$ degrees of freedom. Therefore, similarly to the earlier proofs,

$$Pr(\text{CI Contains } \mu) = Pr(-t_{\alpha/2} < \frac{\bar{X}-\mu}{S_x/\sqrt{n}} < t_{\alpha/2}) = 1 - \alpha \quad .$$