T and F Distributions

The purpose of both of these distributions is to allow for inferences about μ and σ in an unknown distribution. Both are quotients of known distributions.

Preliminaries

Sample Mean: Let $Y_1, \dots, Y_n$ be a random, independent sample from a distribution with mean μ and variance σ^2. Then,

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad \text{Sample Mean}$$

is a distribution with mean $\overline{\mu} = \mu$ and variance $\overline{\sigma}^2 = \frac{\sigma^2}{n}$. If the underlying distribution is a normal distribution, then $\frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ is a *standard* normal distribution.

Sample Variance: The *sample variance* is defined as

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2. \qquad \text{Sample Variance}$$

It is important to note that the sample variance is found for samples drawn from a distribution; for population standard deviation/variance, we use n instead of $n-1$ in the denominator.

When Y_i is a normal distribution, then $\frac{(n-1)S^2}{\sigma^2}$ is a χ^2 distribution with $n-1$ df — S^2 and $\overline{Y}$ are independent.

Definition of T Distribution

Let Z be a standard normal distribution, W be χ^2 with ν df, and Z and W be independent. Then,

$$T = \frac{Z}{\sqrt{W/\nu}}$$

has a T distribution with ν df.

Creating a T Distribution: Let Y_i be sampled from a normal distribution with mean μ and standard deviation σ.

Then, $Z = \frac{\overline{Y}-\mu}{\sigma/\sqrt{n}}$ is a standard normal distribution, and $W = \frac{(n-1)S^2}{\sigma^2}$ is χ^2 with $n-1$ df.

So,

$$\begin{aligned} T &= \frac{Z}{\sqrt{W/(n-1)}} \\ &= \frac{(\overline{Y}-\mu)\sqrt{n}}{\sigma}\sqrt{\frac{(n-1)\sigma^2}{S^2}} \\ &= \frac{(\overline{Y}-\mu)\sqrt{n}}{S} \end{aligned}$$

has a T distribution with $n-1$ df.

T Distribution: Let $Y_1, \dots, Y_6$ be samples from a normal distribution with unknown μ, σ. Estimate $P(|\overline{Y}-\mu| < (2S/\sqrt{n}))$.

Thus, we have

$$\begin{aligned} P\left(|\overline{Y}-\mu| \leq \frac{2S}{\sqrt{n}}\right) &= P\left(-2 \leq \frac{\sqrt{n}(\overline{Y}-\mu)}{S} \leq 2\right) \\ &= P(-2 \leq T \leq 2) \end{aligned}$$

Thus, for $n = 6$, we have that our random variable T has 5 df. By looking at a T distribution table, we can find that $P \approx 0.9$. We can also use R.