9.5 The t-test

Tests for samples from a Normal population are one of the most common hypothesis procedures. Here, we focus on the case when both the mean μ and variance σ^2 of the population are unknown.

We could start by looking at the likelihood ratio to find an MLR statistic. But unlike the case when variance is known, this is a non-starter. Think about the hypotheses:

$$H_0 : \mu \leq \mu_0 \quad H_1 : \mu > \mu_0$$

What is parameter space? We actually omitted the parameter σ^2 from our hypotheses. So we can't actually use MLR to get an efficient test. What alternatives do we have?

Recall that if $\mathbf{X}$ is a sample from $N(\mu, \sigma^2)$, we can form a one-sided γ-level confidence interval for μ as

$$(A, \infty) \quad \text{where } A = \bar{X} - F_{n-1}^{-1}(\gamma)\frac{S}{\sqrt{n}}$$

Since this should give the range of plausible values for μ, we might reject H_0 in favor of H_1 if $\mu_0 < A$. This procedure indeed gives a test of size $\alpha_0 = 1 - \gamma$. Explicitly, we reject H_0 if

$$\mu_0 < \bar{X} - F_{n-1}^{-1}(\gamma)\frac{S}{\sqrt{n}} \iff \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} > F_{n-1}^{-1}(\gamma)$$

This is called the **t-test for a population mean**.

Let $T = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$, which has $t(n-1)$ when $\mu = \mu_0$. Thus

$$P(T > F_{n-1}^{-1}(\gamma) | \mu = \mu_0) = P(F_{n-1}(T) > \gamma | \mu = \mu_0) = 1 - \gamma$$

by the Universality of the Uniform. Moreover, if $\mu < \mu_0$, then

$$T = \frac{X - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{X - \mu}{\frac{S}{\sqrt{n}}} - \frac{\mu_0 - \mu}{\frac{S}{\sqrt{n}}} = T^* - W \qquad T^* \sim t(n-1), W > 0$$

So

$$P(T > F_{n-1}^{-1}(\gamma)|\mu) = P(T^* > F_{n-1}^{-1}(\gamma) + W|\mu) < P(T^* > F_{n-1}^{-1}(\gamma)|\mu) = 1 - \gamma$$

Hence, this procedure is indeed a size $\alpha_0 = 1 - \gamma$ test.

Based on the preceding arguments, we can show that the power function $\pi(\mu, \sigma^2|\delta)$ for this procedure δ has the following properties:

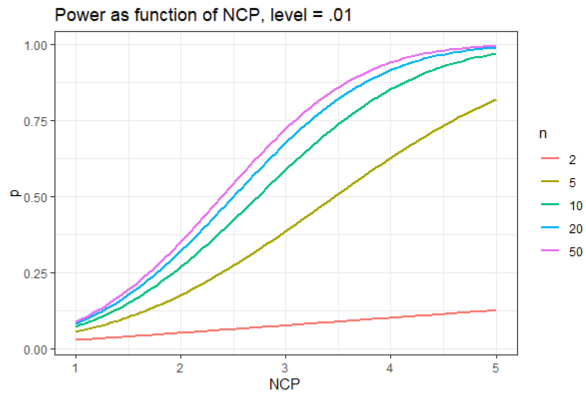
1. $\pi(\mu, \sigma^2|\delta) = \alpha_0$ when $\mu = \mu_0$.
2. $\pi(\mu, \sigma^2|\delta) < \alpha_0$ when $\mu < \mu_0$.
3. $\pi(\mu, \sigma^2|\delta) > \alpha_0$ when $\mu > \mu_0$.
4. $\pi(\mu, \sigma^2|\delta) \to 0$ when $\mu \to -\infty$.
5. $\pi(\mu, \sigma^2|\delta) \to 1$ when $\mu \to \infty$.

However, finding a particular formula for $\pi(\mu, \sigma^2|\delta)$ is more difficult. Note for example when $\mu < \mu_0$,

$$\pi(\mu, \sigma^2|\delta) = P(T > F_{n-1}^{-1}(\gamma)|\mu) = P(T^* > F_{n-1}^{-1}(\gamma) + W|\mu)$$

which involves random variables T^* and W on both sides of the inequality. That is, when $\mu \neq \mu_0$, the variable T is **not** t-distributed. Instead, it is said to have the **noncentral t-distribution** with noncentrality parameter $\text{ncp} = \frac{\mu - \mu_0}{\sigma/\sqrt{n}}$.

Note that R can return the PDF, CDF, QF, and random variates for non-central t by adding an `ncp =...` argument inside `dt`, `pt`, `qt`, `rt`.



Nevertheless, we can still calculate p-values of an observation using facts (1) and (2) above. For a test of

$$H_0 : \mu \leq \mu_0 \quad H_1 : \mu > \mu_0$$

suppose we observe the statistic $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$. Then we would reject H_0 at level α_0 iff

$$t \geq F_{n-1}^{-1}(1-\alpha_0) \iff F_{n-1}(t) \geq 1-\alpha_0 \iff \alpha_0 \geq 1 - F_{n-1}(t)$$

Hence, the p-value for t is $1 - F_{n-1}(t)$.

Two-sided Hypotheses

Suppose instead we consider the two-sided alternative hypothesis:

$$H_0 = \mu = \mu_0 \quad H_1 \mu \neq \mu_0$$

We can derive a α_0 size test from $\gamma = 1 - \alpha_0$ confidence interval, just as we did for the 1-sided test.

In this case, the t-test takes the form. Reject H_0 if

$$|T| \geq F_{n-1}^{-1}(1-\alpha_0/2) \quad \text{where} \quad t = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$$

Ex 1: A batch of stout beer is best when it has an original gravity (OG) close to 1.071. Suppose the OG of beer is $N(\mu, \sigma^2)$. We sample 5 OG measurements from a batch of beer and find $\bar{x} = 1.0686$ and $s = 0.0064$. We perform a 2-sided hypothesis test of $H_0 : \mu = 1.071$ vs. $H_1 : \mu \neq 1.071$. Our t-statistic is then

$$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{1.0686 - 1.071}{0.0064\sqrt{5}} = -0.84$$

At the $\alpha_0 = 0.05$ level, $F_{n-1}(1-\alpha_0/2) =$ `qt(.975, 4)` $= 2.78$. Since $|t| < 2.78$, we do not reject H_0. Moreover, our statistic had p-value of

$$p-\text{value } = 2P(T \geq |t|) = \texttt{2*(1-pt(.84, 4))} = 0.448$$

With all that said, if $\mu = 1.0686$ and $s = 0.0064$, then nc $= \frac{1.0686-1.071}{0.0064/\sqrt{5}}$ and so

$$\begin{aligned} P(|T| > 2.78) =& 1 - P(T < 2.78) + P(T < -2.78) \\ =& \texttt{1-pt(2.78, ncp = nc, df = 4)+pt(-2.78, ncp = nc, df = 4)} \\ =& 0.1 \end{aligned}$$

which shows that this test may not have particularly high power if the true parameter values are close to those actually observed.