Liklihood Ratio Derivation of T-Test

Ryan Lafferty

February 8, 2020

Theorem. Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$ where both parameters are unknown. Then likelihood ratio test for $\mu \leq \mu_0$ vs. $\mu > \mu_0$ is given by the following procedure: First we compute the test statistic $t = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s}$ and the critical value $t_{n-1,\alpha}$ given by the $(1-\alpha)$ quantile of the t distribution. If $t > t_{n-1,\alpha}$ we reject $H_0 : \mu \leq \mu_0$. Otherwise we fail to reject H_0.

Proof. Let $L(x;\theta) = L(x_1, ..., x_n; \mu, \sigma)$ be the likelihood function. We have

$$\begin{aligned}
\log L(x;\theta) &= \log \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \\
&= \log \left(\left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \right) \\
&= -\frac{n}{2} \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2
\end{aligned}$$

For given μ, the maximum of this quantity with respect to σ^2 is found by setting the derivative equal to 0 as follows.

$$\frac{\partial L(x;\theta)}{\partial \sigma^2} = -\frac{n}{2}\left(\frac{1}{\sigma^2}\right) + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

Multiplying both sides by $\frac{2\sigma^4}{n}$ gives $-\sigma^2 + \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ and rearranging to solve for σ^2 gives

$$\sigma^2 = \sigma^2_{\max} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2.$$

Note that maximizing the likelihood is the same thing as maximizing the log-likelihood since the log is a monotone increasing function. Now we have that

$$
\begin{aligned}
\sup_{\mu\in\mathbb{R},\sigma\in\mathbb{R}} \log L(x;\mu,\sigma^2) &= \sup_{\mu\in\mathbb{R}} \log L(x;\mu,\sigma^2_{\max}) \\
&= \max_{\mu\in\mathbb{R}} \left(-\frac{n}{2}\log 2\pi\sigma^2_{\max} - \frac{1}{2\sigma^2_{\max}} \sum_{i=1}^{n} (x_i-\mu)^2 \right) \\
&= \max_{\mu\in\mathbb{R}} \left(-\frac{n}{2}\log 2\pi\sigma^2_{\max} - \frac{n}{2} \right)
\end{aligned}
$$

First, note that $\frac{\partial\sigma^2_{\max}}{\partial\mu} = -\frac{2}{n}\sum_{i=1}^{n}(x_i-\mu) = -\frac{2}{n}\sum_{i=1}^{n} x_i + 2\mu = -2(\bar{x}-\mu)$. We next compute the partial with respect to μ of $\log L(x;\mu,\sigma^2_{\max})$ and set it to 0 as follows:

$$
\frac{\partial L(x;\mu,\sigma^2_{\max})}{\partial\mu} = n\frac{\bar{x}-\mu}{\sigma^2_{\max}} = 0
$$

From this we find that $\mu = \mu_{\max} = \bar{x}$. Therefore the denominator of the likelihood ratio test can be written as

$$
\left(\frac{1}{2\pi\sigma^2_{\max}}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2_{\max}}\sum_{i=1}^{n}(x_i-\bar{x})^2}.
$$

Here, $\sigma^2_{\max}$ is defined to be $\sum_{i=1}^{n}(x_i-\bar{x})^2$, so we can write the denominator as:

$$
\left(\frac{1}{2\pi\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{\frac{n}{2}} e^{-\frac{n}{2}}
$$

Next we need to find the numerator. We need to find a $\mu \le \mu_0$ that will maximize the quantity $-\frac{n}{2}\left(\log 2\pi\sum_{i=1}^{n}(x_i-\bar{x})^2\right) - \frac{n}{2}$ or equivalently, one that minimizes the quantity $\sigma^2_{\text{null}} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$. Now if $\bar{x} \le \mu_0$ then $\bar{x}$ will suffice for this case since we have already seen it will maximize the likelihood. On the other hand then I claim that if $\bar{x} \ge \mu_0$ then

$$
\sum_{i=1}^{n}(x_i-\mu)^2 \ge \sum_{i=1}^{n}(x_i-\mu_0)^2
$$

for all $\mu \le \mu_0$. The proof is as follows:

$$
\begin{aligned}
\sum_{i=1}^{n}(x_i-\mu)^2 &= \sum_{i=1}^{n}(x_i-\mu_0+\mu_0-\mu)^2 \\
&= \sum_{i=1}^{n}(x_i-\mu_0)^2 + \sum_{i=1}^{n}(\mu-\mu_0)^2 + 2(\mu-\mu_0)\sum_{i=1}^{n}(x_i-\mu_0) \\
&\ge \sum_{i=1}^{n}(x_i-\mu_0)^2 + 2n(\mu-\mu_0)(\bar{x}-\mu_0)
\end{aligned}
$$

which is clearly greater than $\sum_{i=1}^n (x_i - \mu_0)^2$ provided $\bar{x} \geq \mu_0$. So the numerator can be written as

$$\sup_{\mu \leq \mu_0, \sigma \in \mathbb{R}} L(x;\theta) = \begin{cases} \left(\frac{1}{2\pi\sigma^2_{\text{null}}}\right)^{\frac{n}{2}} e^{-\frac{n}{2}} & \text{if } \bar{x} \leq \mu_0 \\ \left(\frac{1}{2\pi\sigma^2_{\text{null}}}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2_{\text{null}}}\sum_{i=1}^n (x_i - \mu_0)^2} & \text{if } \bar{x} > \mu_0 \end{cases}$$

Now we have defined σ^2_{null} to be equal to $\begin{cases} \sum_{i=1}^n (x_i - \bar{x})^2 & \text{if } \bar{x} \leq \mu_0 \\ \sum_{i=1}^n (x_i - \mu_0)^2 & \text{if } \bar{x} > \mu_0 \end{cases}$

The ratio of numerator to denominator is then

$$\phi(x) = \begin{cases} 1 & \text{if } \bar{x} \leq \mu_0 \\ \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2}\right)^{\frac{n}{2}} & \text{if } \bar{x} \leq \mu_0 \end{cases}$$

We can rewrite $\sum_{i=1}^n (x_i - \bar{x})^2$ as

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})^2 &= \sum_{i=1}^n (x_i - \mu_0 + \mu_0 - \bar{x})^2 \\ &= \sum_{i=1}^n (x_i - \mu_0)^2 + n(\bar{x} - \mu_0)^2 - 2(\bar{x} - \mu_0)\sum_{i=1}^n (x_i - \mu_0) \\ &= \sum_{i=1}^n (x_i - \mu_0)^2 + n(\bar{x} - \mu_0)^2 - 2n(\bar{x} - \mu_0)^2 \\ &= \sum_{i=1}^n (x_i - \mu_0)^2 - n(\bar{x} - \mu_0)^2 \end{aligned}$$

and rearrange to get $\sum_{i=1}^n (x_i - \mu_0)^2 = n(\bar{x} - \mu_0)^2 + \sum_{i=1}^n (x_i - \bar{x})^2$ so that now

$$\phi(x) = \begin{cases} 1 & \text{if } \bar{x} \le \mu_0 \\ \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(\bar{x}-\mu_0)^2 + \sum_{i=1}^n (x_i - \bar{x})^2} \right)^{\frac{n}{2}} & \text{if } \bar{x} > \mu_0 \end{cases}$$

$$= \begin{cases} 1 & \text{if } \bar{x} \le \mu_0 \\ \left(\frac{1}{\left(\frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) + 1} \right)^{\frac{n}{2}} & \text{if } \bar{x} > \mu_0 \end{cases}$$

$$= \begin{cases} 1 & \text{if } \bar{x} \le \mu_0 \\ \left(\frac{1}{\left(\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sqrt{n-1}s} \right)^2 + 1} \right)^{\frac{n}{2}} & \text{if } \bar{x} > \mu_0 \end{cases}$$

$$= \begin{cases} 1 & \text{if } \bar{x} \le \mu_0 \\ \left(\frac{1}{\frac{t^2}{n-1} + 1} \right)^{\frac{n}{2}} & \text{if } \bar{x} > \mu_0, \end{cases}$$

where s is the sample variance and t is defined as in the statement of the theorem. Note that $\left(\frac{1}{\frac{t^2}{n-1} + 1} \right)^{\frac{n}{2}}$ is always less than or equal to 1 and is monotone decreasing in t. Thus, given any constant $c < 1$ there is some constant $C \in \mathbb{R}$ such that $\phi(x) < c$ if and only if $t > C$. (We know that $c < 1$ because if c were greater than 1 the test would be trivial). Therefore the likelihood ratio test procedure for this problem is exactly the procedure where the null hypothesis is rejected when $t > C$ for some constant C. We determine C by controlling Type I error as follows. Let α be the desired level of Type I error. We want to find C such that

$$P(t > C | \mu \le \mu_0) \le \alpha \text{ for all } \mu \le \mu_0$$

We will first show that, under the null hypothesis, the statistic t has a t distribution. Recall t is defined as the ratio of the two quantities $\bar{x} - \mu_0$ and $\frac{s}{\sqrt{n}}$, or equivalently, of the two quantities $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ and $\frac{s}{\sigma}$. We know that (when $\mu = \mu_0$) $\bar{x}$ is normal with mean μ and variance $\frac{\sigma^2}{n}$. Therefore, $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ is $N(0, 1)$. Furthermore, we know that $(n-1)\left(\frac{s}{\sigma}\right)^2$ is χ^2 with $n-1$ degrees of freedom, so $\frac{s}{\sigma}$ is the square root of a $\chi^2_{(n-1)}$ random variable divided by the square root of its degrees of freedom. Therefore this ratio exactly satisfies the definition of the t distribution. Hence if we let $t_{\alpha,n-1}$ be the $1-\alpha$ quantile for the t distribution with $n-1$ degrees of freedom, we have

$P(t > t_{\alpha,n-1}|\mu = \mu_0) = \alpha$. Now if $\mu = \mu_1 < \mu_0$ then

$$
\begin{aligned}
P(\frac{\sqrt{n}(\bar{x} - \mu_0)}{s} > t_{\alpha,n-1}|\mu = \mu_1) &= P(\frac{\sqrt{n}(\bar{x} - \mu_1)}{s} - \frac{\sqrt{n}(\mu_0 - \mu_1)}{s} > t_{\alpha,n-1}|\mu = \mu_1) \\
&\leq P(\frac{\sqrt{n}(\bar{x} - \mu_1)}{s} > t_{\alpha,n-1}|\mu = \mu_1) \\
&= \alpha
\end{aligned}
$$

Here the second line follows from the fact that $\frac{\sqrt{n}(\mu_0-\mu_1)}{s}$ is a non-negative random variable, so $\frac{\sqrt{n}(\bar{x}-\mu_1)}{s} - \frac{\sqrt{n}(\mu_0-\mu_1)}{s} > t_{\alpha,n-1} \Rightarrow \frac{\sqrt{n}(\bar{x}-\mu_1)}{s} > t_{\alpha,n-1}$. The third line follows from the fact that, following the same proof given on the last page, $\frac{\sqrt{n}(\bar{x}-\mu_1)}{s}$ has a t distribution whenever $\mu = \mu_1$. Therefore, for any $\mu \leq \mu_0$ we have

$$P(t > t_{\alpha,n-1}|\mu \leq \mu_0) \leq \alpha. \Box$$