Student's T Test

Deriving the Student's T Test from the Likelihood Ratio

Let $X_1, \dots, X_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(\mu_1, \sigma_1^2)$. Because the test is one-sided, I will consider the two alternatives:

$$\begin{aligned} H_0 &: \quad \mu_1 = \mu_0 = 5 \\ H_1 &: \quad \mu_1 > 5 \end{aligned}$$

μ_1 and σ_1^2 are both unknown, therefore

$$\Theta_0 = \{(\mu, \sigma^2) \in \mathbb{R} \times (0, +\infty), \mu = \mu_0\}$$

The likelihood ratio test is defined as:

$$T_n' = 2(\ell_n(\hat{\mu}_n^{MLE}, \widehat{\sigma^2}_n^{MLE}) - \ell_n(\hat{\mu}_n^0, \widehat{\sigma^2}_n^0)) \tag{1}$$

Where

$$(\hat{\mu}_n^0, \widehat{\sigma^2}_n^0) = \underset{(\mu, \sigma^2) \in \Theta_0}{\text{argmax}} \ell_n(\mu, \sigma^2)$$

We know that

$$\ell_n(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \tag{2}$$

We are trying to maximize $\ell_n(\mu, \sigma^2)$ under the constraint $\mu = \mu_0$. Let us define $g(\mu, \sigma^2) = \mu - \mu_0$. Our constraint therefore is $g(\mu, \sigma^2) = 0$. Using the Lagrange multiplier:

$$\begin{aligned} \mathcal{L}(\mu, \sigma^2, \lambda) &= \ell_n(\mu, \sigma^2) + \lambda g(\mu, \sigma^2) \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 + \lambda(\mu - \mu_0) \end{aligned}$$

We want to find the solution of $\nabla \mathcal{L}(\mu, \sigma^2, \lambda) = 0$.

$$\begin{gathered} \frac{\partial \mathcal{L}}{\partial \mu}(\mu, \sigma^2, \lambda) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu) + \lambda \\ \frac{\partial \mathcal{L}}{\partial \sigma^2}(\mu, \sigma^2, \lambda) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i - \mu)^2 \\ \frac{\partial \mathcal{L}}{\partial \lambda}(\mu, \sigma^2, \lambda) = \mu - \mu_0 \end{gathered}$$

We are only interested in finding μ and σ^2; using the last two equations, and equating them with 0, leads to:

$$\hat{\mu}_n^0 = \mu_0 \tag{3}$$

And for $\widehat{\sigma^2}_n^0$:

$$-\frac{n}{2\widehat{\sigma^2}^0_n} + \frac{1}{2\left(\widehat{\sigma^2}^0_n\right)^2}\sum_{i=1}^n (X_i - \mu_0)^2 = 0$$

$$-\frac{n}{2\left(\widehat{\sigma^2}^0_n\right)^2}\left(\widehat{\sigma^2}^0_n - \frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2\right) = 0$$

$$\widehat{\sigma^2}^0_n = \frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2 \tag{4}$$

We also know the forms of the maximum likelihood estimators:

$$\hat{\mu}_n^{MLE} = \bar{X}_n \tag{5}$$

$$\widehat{\sigma^2}^{MLE}_n = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2 \tag{6}$$

Injecting (5) and (6) into (2) gives us:

$$\ell_n(\hat{\mu}_n^{MLE}, \widehat{\sigma^2}^{MLE}_n) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\widehat{\sigma^2}^{MLE}_n) - \frac{1}{2}\frac{1}{\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2}\sum_{i=1}^n (X_i - \bar{X}_n)^2$$

$$\ell_n(\hat{\mu}_n^{MLE}, \widehat{\sigma^2}^{MLE}_n) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\widehat{\sigma^2}^{MLE}_n) - \frac{n}{2} \tag{7}$$

Similarly, by injecting (3) and (4) into (2):

$$\ell_n(\hat{\mu}_n^0, \widehat{\sigma^2}^0_n) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\widehat{\sigma^2}^0_n) - \frac{1}{2}\frac{1}{\frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2}\sum_{i=1}^n (X_i - \mu_0)^2$$

$$\ell_n(\hat{\mu}_n^0, \widehat{\sigma^2}^0_n) == -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\widehat{\sigma^2}^0) - \frac{n}{2} \tag{8}$$

Therefore, our test is, using (1):

$$\begin{aligned}
T_n' &= 2(\ell_n(\hat{\mu}_n^{MLE}, \widehat{\sigma^2}^{MLE}_n) - \ell_n(\hat{\mu}_n^0, \widehat{\sigma^2}^0_n)) \\
&= 2\left(-\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\widehat{\sigma^2}^{MLE}_n) - \frac{n}{2} - \left(-\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\widehat{\sigma^2}^0) - \frac{n}{2}\right)\right) \\
&= n\left(-\ln(\widehat{\sigma^2}^{MLE}_n) + \ln(\widehat{\sigma^2}^0_n)\right)
\end{aligned}$$

$$T_n' = n\ln\left(\frac{\widehat{\sigma^2}^0_n}{\widehat{\sigma^2}^{MLE}_n}\right) \tag{9}$$

We can simplify further this expression, by using the definition of $\widehat{\sigma^2}^0_n$:

$$\begin{aligned}
\widehat{\sigma^2}_n^0 &= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n+\bar{X}_n-\mu_0)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left((X_i-\bar{X}_n)^2+2(X_i-\bar{X}_n)(\bar{X}_n-\mu_0)+(\bar{X}_n-\mu_0)^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2+\frac{2}{n}(\bar{X}_n-\mu_0)\left(\sum_{i=1}^{n}X_i-n\bar{X}_n\right)+\frac{n}{n}(\bar{X}_n-\mu_0)^2 \\
&= \widehat{\sigma^2}_n^{MLE}+\frac{2}{n}(\bar{X}_n-\mu_0)(n\bar{X}_n-n\bar{X}_n)+(\bar{X}_n-\mu_0)^2
\end{aligned}$$

$$\widehat{\sigma^2}_n^0 = \widehat{\sigma^2}_n^{MLE}+(\bar{X}_n-\mu_0)^2 \tag{10}$$

Which gives us the final expression of our test:

$$T'_n = n\ln\left(1+\frac{(\bar{X}_n-\mu_0)^2}{\widehat{\sigma^2}_n^{MLE}}\right) \tag{11}$$

What I believe to be the wrong answer

If, instead of finding the optimal $\widehat{\sigma^2}_n^0$ in Θ_0, we directly inject $\widehat{\sigma^2}_n^{MLE}$, we have instead the following expression:

$$\ell_n(\hat{\mu}_n^0,\widehat{\sigma^2}_n^{MLE}) = -\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln(\widehat{\sigma^2}_n^{MLE})-\frac{n}{2}\frac{\widehat{\sigma^2}_n^0}{\widehat{\sigma^2}_n^{MLE}}$$

Which leads to the alternative expression for T'_n:

$$\begin{aligned}
T'_n &= 2\left(-\frac{n}{2}\ln(\widehat{\sigma^2}_n^{MLE})-\frac{n}{2}+\frac{n}{2}\ln(\widehat{\sigma^2}_n^{MLE})+\frac{n}{2}\frac{\widehat{\sigma^2}_n^0}{\widehat{\sigma^2}_n^{MLE}}\right) \\
&= n\left(\frac{\widehat{\sigma^2}_n^0}{\widehat{\sigma^2}_n^{MLE}}-1\right) \\
&= n\frac{\widehat{\sigma^2}_n^{MLE}+(\bar{X}_n-\mu_0)^2-\widehat{\sigma^2}_n^{MLE}}{\widehat{\sigma^2}_n^{MLE}} \\
&= n\frac{(\bar{X}_n-\mu_0)^2}{\widehat{\sigma^2}_n^{MLE}}
\end{aligned}$$

And we find exactly what I think you are expecting; **but** this is, I think, an incorrect application of the likelihood ratio test, as we cannot consider that the maximizer $\widehat{\sigma^2}_n^0$ of ℓ_n restricted to Θ_0 is $\widehat{\sigma^2}_n^{MLE}$; this case corresponds to the scenario where σ_1^2 is **known** and equal to $\widehat{\sigma^2}_n^{MLE}$.

It is of course possible that my reasoning is wrong, but I don't see where I might have missed something.