2. Student's T Test

Deriving the Student's T Test from Likelihood Ratio

2.0/2得分 (计入成绩)

Let $X_1,\ldots,X_n \stackrel{iid}{\sim} X \sim \mathcal{N}(\mu_1,\sigma_1^2)$. Consider the null and alternative hypotheses

$$\begin{aligned} H_0 &: \mu_1 = 5 \\ H_1 &: \mu_1 \neq 5. \end{aligned}$$

Assume that μ_1 is not known, but σ_1^2 is known. The test statistic T_n' for the likelihood ratio test associated to the above hypothesis can be expressed in terms of n, $\overline{X}_n$, and σ_1^2.

What is T_n'?

(Enter **barX_n** for $\overline{X}_n$, and **sigma_1^2** for σ_1^2.)

$T_n' =$ 2*((25*n - 10*n* barX_n + n* (barX_n)^2)/(2*sigma_1^2))

Answer: n/(sigma_1^2)*(barX_n - 5)^2

STANDARD NOTATION

If σ_1^2 were unknown and we used the estimator $\widetilde{\sigma_1^2} = \frac{1}{n-1}\sum_i \left(X_i - \overline{X}_n\right)^2$ in **both log-likelihoods**, what would be the distribution of $\sqrt{T_n'}$?

- t_{n-1}
- t_n
- $|t_{n-1}|$ (selected)
- None of the above.

STANDARD NOTATION

Solution:

Recall that the MLE for a Gaussian statistical model is $(\overline{X}_n, \hat{\sigma}^2)$.

Therefore, by the definition of the likelihood-ratio test,

$$\begin{aligned} T_n' &= 2\left(\ell\left(X_1,\ldots,X_n;\overline{X}_n,\hat{\sigma}^2\right) - \ell\left(X_1,\ldots,X_n;5,\hat{\sigma}^2\right)\right) \\ &= 2\left(\frac{-1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 + \frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}\left(X_i - 5\right)^2\right) \\ &= \frac{1}{\hat{\sigma}^2}\left(\sum_{i=1}^{n}\left(-X_i^2 + 2X_i\overline{X}_n - \overline{X}_n^2 + X_i^2 - 10X_i + 25\right)\right) \\ &= \frac{1}{\hat{\sigma}^2}\left(2n\overline{X}_n^2 - n\overline{X}_n^2 - 10n\overline{X}_n + 25n\right) \end{aligned}$$

$$= \frac{n}{\hat{\sigma}^2}\left(\overline{X}_n^2 - 10\overline{X}_n + 25\right)$$
$$= \frac{n}{\sigma^2}\left(\overline{X}_n - 5\right)^2.$$

For the second question, observe that

$$\sqrt{T_n'} = \frac{\sqrt{n}}{\sigma}|\overline{X}_n - 5|.$$

Plugging in the estimator for σ_1^2, since

$$\frac{\sqrt{n}}{\widetilde{\sigma_1}}(\overline{X}_n - 5) \sim t_{n-1},$$

we conclude that the response "$|\mathbf{t_{n-1}}|$" is correct.

提交 你已经尝试了1次（总共可以尝试4次）

Answers are displayed within the problem

Introducing Another Sample

1/1得分 (计入成绩)

Let $Y_1, \ldots, Y_m \overset{iid}{\sim} Y \overset{iid}{\sim} N\left(\mu_2, \sigma_2^2\right)$ denote another sample, and assume that X's are independent of the Y's.

What is the distribution of $\overline{X}_n - \overline{Y}_m$?

- (selected) $N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$
- $N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} - \frac{\sigma_2^2}{m}\right)$
- $N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$
- None of the above.

Solution:

Since $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ are mutually independent, we know that $\overline{X}_n - \overline{Y}_m$ will have a normal distribution. It remains to compute the mean and variance.

$$\mathbb{E}\left[\overline{X}_n - \overline{Y}_m\right] = \mu_1 - \mu_2$$

by linearity of expectation. By independence, the variances are additive, so

$$\mathrm{Var}\left(\overline{X}_n - \overline{Y}_m\right) = \mathrm{Var}\left(\overline{X}_n\right) + \mathrm{Var}\left(\overline{Y}_m\right) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}.$$

Therefore the correct response is "$N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$".

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Test Statistic for a Two-Sample Test

1.0/1得分 (计入成绩)

Recall that $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$, $Y_1, \ldots, Y_m \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$, and the two samples are independent of one another. Consider the null and alternative hypotheses

$$\begin{aligned} H_0 &: \mu_1 \leq \mu_2 \\ H_1 &: \mu_1 > \mu_2. \end{aligned}$$

What is the test statistic T_n for the two-sample student's T test associated to H_0 and H_1?Express your answer in terms of $n, m, \hat{\sigma_1}^2, \hat{\sigma_2}^2, \overline{X}_n$, and $\overline{Y}_m$.

(Enter **barX_n** for $\overline{X}_n$, **barY_m** for $\overline{Y}_m$, **hat(sigma_1^2)** for $\widehat{\sigma_1^2}$, and **hat(sigma_2^2)** for $\widehat{\sigma_2^2}$.)

$T_n =$ (barX_n - barY_m)/sqrt(hat(sigma_1^2)/n+hat(sigma_2^2)/m)

Answer: (barX_n - barY_m)/sqrt(hat(sigma_1^2)/n + hat(sigma_2^2)/m)

STANDARD NOTATION

Solution:

Under the null hypothesis, we observe that

$$\mathbf{P}\left(\frac{\overline{X}_n - \overline{Y}_m - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{\sigma_1}^2}{n} + \frac{\hat{\sigma_2}^2}{m}}} > \tau\right) \leq \mathbf{P}\left(\frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{\hat{\sigma_1}^2}{n} + \frac{\hat{\sigma_2}^2}{m}}} > \tau\right)$$

Therefore, we define the test statistic to be

$$T_n = \frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{\hat{\sigma_1}^2}{n} + \frac{\hat{\sigma_2}^2}{m}}}.$$

提交 你已经尝试了1次（总共可以尝试4次）

Answers are displayed within the problem

Applying the Welch-Satterthwaite Formula

2/2得分 (计入成绩)

Suppose we observe $\overline{X}_n = 6.2, \overline{Y}_m = 6, \hat{\sigma_1}^2 = 0.1$, and $\hat{\sigma_2}^2 = 0.2$ with $n = 50$ and $m = 50$.

Using the Welch-Satterthwaite formula, what is the approximate number of degrees of freedom for the test statistic T_n?

88 **Answer:** 88

What is the p-value for this test?

(You may consult a table of values or use software for the student's T distribution.)

0.005 **Answer:** 0.0057

Solution:

According to the Welch-Satterthwaite formula, under H_0, T_n is approximately distributed as t_{88} because

$$\frac{\left(\frac{0.1}{50}+\frac{0.2}{50}\right)}{\sqrt{\frac{0.1^2}{50^2(50-1)}+\frac{0.2^2}{50^2(50-1)}}} \approx 88.2$$

Hence, the correct response to the first question is $\mathbf{88}$.

For the second question, we compute

$$T_n = \frac{6.2-6}{\sqrt{\frac{0.1}{50}+\frac{0.2}{50}}} \approx 2.582.$$

Consulting a table for the student's T distribution, we observe that $P\left(t_{88} > 2.582\right) \approx 0.0057$. Therefore, the correct answer to the second question is $\mathbf{0.0057}$.

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

讨论

主题： Unit 4 Hypothesis testing:Homework 7 / 2. Student's T Test

显示讨论