6. A Two-Sample Test on Standardized Test Scores

(a)

3/3 points (graded)

The National Assessment of Educational Progress tested a simple random sample of 1000 thirteen year old students in both 2004 and 2008 and recorded each student's score. The average and standard deviation in 2004 were 257 and 39, respectively. In 2008, the average and standard deviation were 260 and 38, respectively.

Your goal as a statistician is to assess whether or not there were statistically significant changes in the average test scores of students from 2004 to 2008. To do so, you make the following modeling assumptions regarding the test scores:

- $X_1, \ldots, X_{1000}$ represent the scores in 2004.
- $X_1, \ldots, X_{1000}$ are iid Gaussians with standard deviation 39.
- $\mathbb{E}[X_1] = \mu_1$, which is an unknown parameter.
- $Y_1, \ldots, Y_{1000}$ represent the scores in 2008.
- $Y_1, \ldots, Y_{1000}$ are iid Gaussians with standard deviation 38.
- $\mathbb{E}[Y_1] = \mu_2$, which is an unknown parameter.
- $X_1, \ldots, X_n$ are independent of $Y_1, \ldots, Y_n$.

You define your hypothesis test in terms of the null $H_0 : \mu_1 = \mu_2$ (signifying that there were not significant changes in test scores) and $H_1 : \mu_1 \neq \mu_2$. You design the test

$$\psi = \mathbf{1}\left(\sqrt{n}\left|\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}}\right| \geq q_{\eta/2}\right).$$

where q_η represents the $1 - \eta$ quantile of a standard Gaussian.

Hint: Under $H_0 : \mu_1 = \mu_2$, the test statistic is distributed as a standard Gaussian:

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}} \sim N(0, 1)$$

You are encouraged to check this.(Compute the mean and variance and recall that the sum of iid Gaussians is again Gaussian.)

What is the largest possible value of η so that ψ has level 10%?

$\eta =$ 0.1 **Answer:** 0.1

If ψ is designed to have level 10%, would you **reject** or **fail to reject** the null hypothesis given the 2008 data?

- (selected) Reject
- Fail to reject

What is the p-value for this data set?

0.0818

Answer: 0.0815

Solution:

Recall the definition of quantiles a standard Gaussian: q_η is the number such that

$$\eta = P\left(Z \geq q_\eta\right)$$

where $Z \sim N\left(0,1\right)$. Observe that under the null hypothesis,

$$\sqrt{n}\left|\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2+39^2}}\right| \sim N\left(0,1\right).$$

By symmetry,

$$P\left(|Z| \geq q_{\eta/2}\right) = 2P\left(Z \geq q_{\eta/2}\right).$$

Thus our goal is to choose the smallest η such that $P\left(|Z| \geq q_{\eta/2}\right) \leq 0.1\%$. We get

$$2P\left(Z \geq q_{\eta/2}\right) = 0.1 \Rightarrow \eta = 0.1.$$

To determine if we should reject or accept the null based on the 2008 data, we need to compute $q_{0.1/2} = q_{0.05}$. Using computational software or a table, we find that $q_{0.05} \approx 1.64$. Now we evaluate our test statistic on the 2008 data:

$$\sqrt{n}\left|\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2+39^2}}\right| = \sqrt{1000}\left|\frac{260-257}{\sqrt{38^2+39^2}}\right| \approx 1.7422$$

Hence, $\psi = 1$, and we would **reject** the hypothesis that there were no changes in test scores between 2004 and 2008.

To compute the p-value for this data set, we let $Z \sim N\left(0,1\right)$ and compute using a table

$$P\left(|Z| > 1.7422\right) = 2P\left(Z > 1.7422\right) \approx 0.0815$$

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

显示讨论

主题：Unit 2 Foundation of Inference:Homework 3: Introduction to Hypothesis Testing / 6. A Two-Sample Test on Standardized Test Scores

认证证书是什么？