S1211Q Introduction to Statistics Lecture 22

Wei Wang

August 7, 2012

Inferential Problems concerning β_1

- Based on Least Squares method, we have the estimator of slope β_1, $\hat{\beta}_1$, but this doesn't answer some of the most important inferential problems.

Inferential Problems concerning β_1

- Based on Least Squares method, we have the estimator of slope β_1, $\hat{\beta}_1$, but this doesn't answer some of the most important inferential problems.
- Is $\hat{\beta}_1$ unbiased?

Inferential Problems concerning β_1

- Based on Least Squares method, we have the estimator of slope β_1, $\hat{\beta}_1$, but this doesn't answer some of the most important inferential problems.
- Is $\hat{\beta}_1$ unbiased?
- What's the (estimated) standard error?

Inferential Problems concerning β_1

- Based on Least Squares method, we have the estimator of slope β_1, $\hat{\beta}_1$, but this doesn't answer some of the most important inferential problems.
- Is $\hat{\beta}_1$ unbiased?
- What's the (estimated) standard error?
- How to get Confidence Interval of β_1?

Inferential Problems concerning β_1

- Based on Least Squares method, we have the estimator of slope β_1, $\hat{\beta}_1$, but this doesn't answer some of the most important inferential problems.
- Is $\hat{\beta}_1$ unbiased?
- What's the (estimated) standard error?
- How to get Confidence Interval of β_1?
- How to perform Hypothesis Test and get P-value about null hypothesis $H_0 : \beta_1 = 0$

Sampling Distribution of $\hat{\beta}_1$

- The least squares estimator $\hat{\beta}_1$ is an unbiased estimator, which mean that $E(\hat{\beta}_1) = \beta_1$.
- Also we have shown yesterday that the variance of this estimator is σ^2/S_{xx}. The estimated standard error is $s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}}$.
- In particular, under the assumption that the noise terms are normally distributed, the $\hat{\beta}_1$ is also normally distributed

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2/S_{xx})$$

Confidence Interval of β_1

- The way to build confidence interval for β_1 is the classical procedure, standardizing the estimator by subtracting its mean and then dividing by its estimated standard error.
- It turns out that the standardized variable

$$T = \frac{\hat{\beta}_1 - \beta_1}{S/\sqrt{S_{xx}}} = \frac{\hat{\beta}_1 - \beta_1}{S_{\hat{\beta}_1}}$$

  follows a t distribution with df $n - 2$.
- So a $100(1 - \alpha)\%$ CI for the slope β_1 is

$$\hat{\beta}_1 \pm t_{\alpha/2,n-2} \cdot s_{\hat{\beta}_1}$$

Hypothesis Testing

Null hypothesis: $H_0\colon \beta_1 = \beta_{10}$

Test statistic value: $t = \dfrac{\hat{\beta}_1 - \beta_{10}}{s_{\hat{\beta}_1}}$

Alternative Hypothesis	**Rejection Region for Level α Test**
$H_a\colon \beta_1 > \beta_{10}$	$t \geq t_{\alpha,n-2}$
$H_a\colon \beta_1 < \beta_{10}$	$t \leq -t_{\alpha,n-2}$
$H_a\colon \beta_1 \neq \beta_{10}$	either $t \geq t_{\alpha/2,n-2}$ or $t \leq -t_{\alpha/2,n-2}$

A P-value based on $n - 2$ df can be calculated just as was done previously for t tests in Chapters 8 and 9.

The **model utility test** is the test of $H_0\colon \beta_1 = 0$ versus $H_a\colon \beta_1 \neq 0$, in which case the test statistic value is the **t ratio** $t = \hat{\beta}_1/s_{\hat{\beta}_1}$.