, 26, 27, 28, 29, 30

Stat 231 - Statistics **Spring 2017**

Lecture 25, 26, 27, 28, 29, 30: June 26 - July 10, 2017

Lecturer: Suryapratim Banerjee *Notes By: Harsh Mistry*

25.1 Measurement Bias Testing

Objective : To test whether a scale is biased

Experiment : We take an object of a known weight (10) and measure of using the scale n times.

$$Y_i = \text{ith reading of the scale}$$

Let S = bias of our scale and R_i error on the ith measurement

$$Y_i = \text{true weight} + S + R_i$$

Conventions :

- If d is to the right of the median of the x^2
  $p - value = 2P(D \geq d)$
- If d is to the left of the median $p - value = 2P(D \leq d)$

25.2 Testing Using The Likelihood

Suppose we are not able to find D fro some problem. Assume n is large. We can then use the likelihood ratio test statistic as our discrepancy measure.

$$\triangle(\theta) = -2log\frac{L(\theta)}{L(\tilde{\theta}}$$

$$D = \mid \frac{\tilde{\theta} - \theta_o}{\sqrt{\frac{\theta(1-\theta)}{n}}} \mid$$

1. Find $L(\theta)$ and calculate $\hat{\theta}$
2. Calculate $\triangle(\theta_o$
3. Calculate the p-value where $d = \triangle$

25.3 Introduction To Regression Models

A linear regression model assumes that E(y) is a linear function of x

$$E(Y) = \alpha + \beta X$$

25.4 Simple Linear Regression Model

Assumptions

- Given x_i's, Y_i's have a Normal distribution.
- The mean of the Y_i's are a linear function of the x_i's
- The variance of the Y_I's are constant σ^2 (unknown)
  σ^2 is independent of x
- σ^2 is not a function of x and Y_i's are independent

25.4.1 Model

$$Y_i \sim G(\alpha + \beta x_i, \sigma)$$
$$\text{Alternate Form : } Y_i = \alpha + \beta x_i + R_i$$

25.4.2 Gauss-Markov Assumptions

- Y_i's independent
- Y_i's are Gaussian
- $E(Y_i = \alpha + \beta x_i$
- $Var(Y_i) = \sigma^2 \forall x$
- $E(Y_i)$ is not linear $\longrightarrow$ Non-linear regression models
- $V(Y_i) \neq \sigma^2 \longrightarrow$ Hetroscedastic Models
- More than one explanatory variable $\longrightarrow$ Multivariable Regression

25.4.3 Questions

- Interpretation of α, β, σ ?
  β is the increase in the mean of Y if x goes by 1 unit
  $\alpha = E(Y)$ when x = 0
- MLE for α, β, σ

- $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$ where $\hat{\beta} = \frac{S_{xy}}{S_{xx}}$
- $\hat{\sigma}^2 = \frac{1}{n}[X_{yy} - \hat{\beta}S_{xy}]$
- $S_{yy} = \sum(y_i - \bar{y})^2$
- $S_{xx} = \sum(x_i - \bar{x})^2$
- $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y})$

$$S^2 = \frac{1}{n-2}[S_{yy} - \beta S_{xy}$$

Let $a_i = \frac{1}{S_{xx}} \cdot (x_i - \bar{x})$

$$\hat{p} = \sum a_i y_i$$

25.4.4 Confidence interval for p

- $E(\tilde{p}) = \alpha\sigma + \beta = \beta$
- $V(\beta) = \frac{\sigma^2}{S_{xx}}$
- $$\frac{\tilde{\beta} - \beta}{\frac{S}{\sqrt{S_{xx}}}} \sim T_{n-2}$$
- Coverage Interval
$$\tilde{\beta} \pm t^* \frac{S}{\sqrt{S_{ss}}}$$
- D value for hypothesis testing
$$D = \mid \frac{\tilde{\beta} - \beta_0}{\frac{S}{\sqrt{S_{xx}}}} \mid$$