Stat 312: Lecture 21
Linear Regression II.

Moo K. Chung
mchung@stat.wisc.edu

Dec 7, 2004

Concepts

1. *Maximum likelihood estimation.* Given a linear model
$$Y_j = \beta_0 + \beta_1 x_j + \epsilon_j$$
with $\epsilon_j \sim N(0, \sigma^2)$, we will estimate β_0, β_1 using the maximum likelihood estimation. Note that
$$Y_j \sim N(\beta_0 + \beta_1 x_j, \sigma^2).$$
The density function for Y_j are
$$f(y_j) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big[\frac{(y_j - \beta_0 - \beta_j x_j)^2}{2\sigma^2}\Big].$$
The loglikelihood is
$$L(\beta_0, \beta_1) = const - \frac{1}{2\sigma^2}\sum_{j=1}^n (y_j - \beta_0 - \beta_1 x_j)^2.$$
We maximize the loglikelihood which is equivalent to minimizing the sum of the residuals
$$\frac{1}{2\sigma^2}\sum_{j=1}^n (y_j - \beta_0 - \beta_1 x_j)^2.$$
Hence MLE is LSE in linear regression.

2. The least squares estimation for β_1 are given by
$$\hat\beta_1 = \frac{S_{xy}}{S_{xx}}$$
where $S_{xy} = n(\overline{xy} - \bar x \bar y)$. It can be shown that
$$\hat\beta_1 = \frac{S_{xy}}{S_{xx}} = \sum_{j=1}^n c_j Y_j,$$
where
$$c_j = (x_j - \bar x)/S_{xx}.$$
So LSE is a linear estimation. From $\sum_{j=1}^n c_j = 0$, $\sum_{j=1}^n c_j x_j = 1$ and $\sum_{j=1}^n c_j^2 = S_{xx}^{-1}$ we can show that
$$\mathbb{E}\hat\beta_1 = \beta_1$$
showing unbiasness. Further, $\mathbb{V}\hat\beta_1 = \sigma^2/S_{xx}$. Since $\hat\beta_1$ is a linear combination of normals, $\hat\beta_1 \sim N(\beta_1, \sigma^2/S_{xx})$.

3. We are interested in hypothesis testing
$$H_0 : \beta_1 = 0 \text{ vs. } H_1 : \beta_1 \neq 0.$$
Inference on the slope parameter β_1 is based on test statistic
$$T = \frac{\hat\beta_1 - \beta_1}{S_{\hat\beta_1}} \sim t_{n-2},$$
where $S_{\hat\beta_1} = \hat\sigma/\sqrt{S_{xx}}$ and $\hat\sigma = SSE/(n-2)$. It can be shown that $SSE = S_{yy} - S_{xy}^2/S_{xx}$, we get $S_{\hat\beta_1} = \frac{1}{\sqrt{n-2}}\sqrt{\frac{S_{yy}}{S_{xx}} - (\frac{S_{xy}}{S_{xx}})^2}$. Then we reject H_0 if $|t| > t_{\alpha/2,n-2}$ at $100(1-\alpha)\%$ significance. We don't usually compute the test statistic by hand. Use R-package.

Example. We continue Lecture 19 example.

```
>summary(lm(y~x))

Call: lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-10.908  -6.312   1.758   4.354  10.836

Coefficients:
   Estimate Std. Error t value Pr(>|t|)

(Intercept) 29.48 13.23 2.22 0.06
x            0.55  0.17 3.12 0.01 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*'

Residual standard error: 7.647 on 8 degrees
of freedom Multiple R-Squared: 0.5519,
Adjusted R-squared:0.4959 F-statistic:9.854
1 and 8 DF, p-value: 0.01383
```