STAT 8010 Statistical Methods I
Homework 5

Instructor: Whitney Huang (wkhuang@clemson.edu)

Due Date: April 23, 9:30am

Problem 1 Leaning Tower of Pisa

The following table provides the annual measurements of the lean (the difference between where a point on the tower would be if the tower were straight and where it actually is) from 1975 to 1987. The variable `Lean` are coded as tenths of a millimeter in excess of 2.9 meters, so that the 1975 lean was actually 2.9642 meters. The variable `Year` are coded as the last two digits of the year (e.g. 85 means the year of 1985). We would like to characterize lean over time by fitting a simple linear regression.

`Year`	75	76	77	78	79	80	81	82	83	84	85	86	87
`Lean`	642	644	656	667	673	688	696	698	713	717	725	742	757

(a) Identify the response variable (Y), the explanatory variable (X), and the sample size (n).

$$Y: \texttt{Lean}$$
$$X: \texttt{Year}$$
$$\texttt{Sample size:} \quad n = 13$$

(b) Use the fact that $\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) = 1696$, $\sum_{i=1}^{n}(X_i - \bar{X})^2 = 182$, and $\bar{Y} = 693.6923$ to compute the estimated slope $\hat{\beta}_1$ and intercept $\hat{\beta}_0$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{1696}{182} = 9.32$$
$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X} = 693.69 - 9.32 \times 81 = -61.12$$

(c) Compute the fitted value and the associated residual value in year 1983.

$$\hat{Y}_{x=83} = -61.12 + 9.3187 * 83 = 712.33$$
$$e_{x=83} = 713 - 712.33 = 0.67$$

The fitted value for lean at 1983 is 2.971233m and the residual is 6.7×10^{-5}m

(d) Compute $\hat{\sigma}$, the estimate of σ

$$\begin{aligned}\hat{\sigma} &= \sqrt{\frac{\sum_{i=1}^{13}(Y_i - \hat{Y}_i)^2}{13-2}} \\ &= \sqrt{\frac{192.29}{11}} = 4.18\end{aligned}$$

(e) Find the 95% confidence interval for β_1

The 95% CI for β_1 is $\hat{\beta}_1 \pm t_{0.05/2, df=11}\hat{\sigma}_{\hat{\beta}_1}$. From (b) we have $\hat{\beta}_1 = 9.32$. $t_{0.025,11} = 2.20$ can be found in t table (or type `qt(0.975, 11)` in R). Now we need to compute $\hat{\sigma}_{\hat{\beta}_1}$:

$$\hat{\sigma}_{\hat{\beta}_1} = \frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^{13}(X_i - \bar{X})^2}} = \frac{4.18}{\sqrt{182}} = 0.31$$

$\Rightarrow$ We have 95% CI for β_1:

$$(9.32 - 2.20 \times 0.31, 9.32 + 2.20 \times 0.31) = (8.64, 10.00)$$

(f) Test the following hypothesis: $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$ with $\alpha = 0.05$

Step 1: Compute the t test statistic $t_{obs} = \dfrac{\hat{\beta}_1 - 0}{\hat{\sigma}_{\hat{\beta}_1}} = \dfrac{9.32}{0.31} = 30.07$
Step 2: Compute the P-Value $= \mathrm{P}(|t_{11}| \geq |30.07|) = 6.5 \times 10^{-12}$
Step 3: Reject $H_0 : \beta_1 = 0$

$\Rightarrow$ Evidence suggests a positive linear relationship between `Lean` and `Year` at 0.05 level.

(g) Construct the 90% confidence interval for E[`Lean`] in year 1984

Similar to (e), we need to first compute $\hat{Y}_{x=84} = 721.65$, then find $t_{0.95,11} (= 1.80)$, and compute $\hat{\sigma}_{\hat{Y}_{x=84}} = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(84-81)^2}{182}} = 1.49$. Therefore we have the 90% CI for $\mathrm{E}[Y_{x=84}]$ is $(718.98, 724.32)$

(h) Calculate the ANOVA table and perform the F test with $\alpha = 0.05$.

Step 1: Find MSR. $\mathrm{SSR} = \sum_{i=1}^{13} (\hat{Y}_i - \bar{Y})^2 = 15804.48 \Rightarrow \mathrm{MSR} = \mathrm{SSR}/1 = 15804.48$
Step 2: Find MSE. From (d), we have MSE $= (\hat{\sigma})^2 = 17.48$
Step 3: Compute F test statistic $= \mathrm{MSR}/\mathrm{MSE} = 904.12$
Step 4: Compute the P-value $= 6.5 \times 10^{-12} < \alpha \Rightarrow$ RejectH_0

Problem 2 Residual Analysis

Use a statistical software to plot e_i's vs. x_i's to assess the model assumptions.

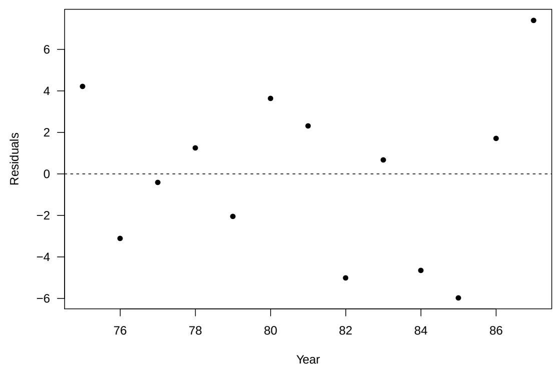

Figure 1: Residual plot

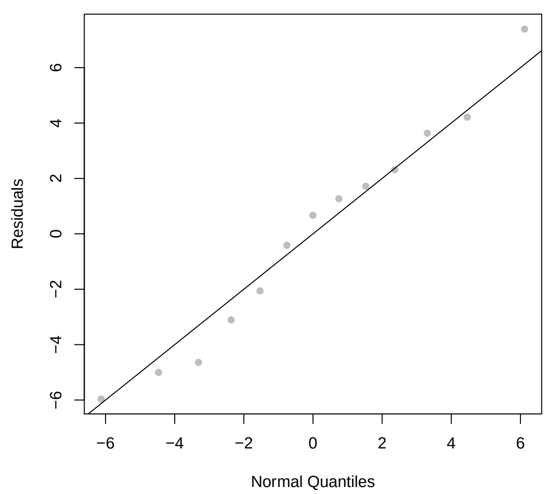

Figure 2: QQ plot for residuals