Chapter 11 & 2- Simple Linear Regression Model and Correlation

STAT 251

Lecture 36

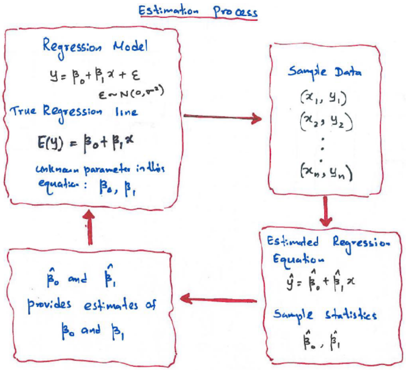
Inference in the Simple Linear Regression Model
Examples

Dr. Lasantha Premarathna

Chapter 11 & 2 - Learning Outcomes

- Scatter plot
- Covariance & Correlation
- Simple linear regression
- Least squares estimates in simple linear regression
- Interpret the parameters in a fitted linear model
- **Inference for the slope parameter - Confidence interval & hypothesis testing**

Simple Linear Regression



- $b_1 = \hat{\beta}_1 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \dfrac{\left(\sum_{i=1}^{n} x_i y_i\right) - n\bar{x}\bar{y}}{\left(\sum_{i=1}^{n} x_i^2\right) - n\bar{x}^2} = \dfrac{r s_y}{s_x}$
- $b_0 = \hat{\beta}_0 = \dfrac{\sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i}{n} = \bar{y} - \hat{\beta}_1 \bar{x}$

Inference About the Slope Parameter β_1

Hypothesis testing for β_1

1. In simple linear regression, we wish to test

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

- If no linear relationship exists between the two variables, we would expect the regression line to be horizontal, that is, to have a slope of zero.

- We want to see if there is a linear relationship
i.e. we want to see if the slope (β_1) is something other than zero.
Thus our alternative hypothesis become $H_1 : \beta_1 \neq 0$

Inference About the Slope Parameter β_1

2. Test statistic

$$t = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}} \sim t_{n-2}$$

test statistic has a $t-$distribution with $n-2$ df

- $s_{\hat{\beta}_1}$ is the estimated standard deviation of $\hat{\beta}_1$

  (i.e. $s_{\hat{\beta}_1}$ is the standard error of $\hat{\beta}_1$)

$$s_{\hat{\beta}_1} = \frac{s}{s_x \sqrt{n-1}}$$

where, $s^2 = \frac{SSE}{n-2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}$ (s^2 is the estimate for σ^2)

and $s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(\sum_{i=1}^{n} x_i^2) - n\bar{x}^2}{n-1}}$

Inference About the Slope Parameter β_1

Under the Null Hypothesis, $H_0 : \beta_1 = 0$

The test statistics is $t_{obs} = \dfrac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$

3. Find the critical value for the significance level α from the $t-$table with $n-2$ degrees of freedom.

4. Reject H_0 if $|t_{obs}| \geq t_{\frac{\alpha}{2},\,(n-2)}$

5. Conclusion

** you also can use $p-$value approach instead of critical value approach

Confidence Interval for the Slope β_1

A $(1-\alpha)100\%$ confidence Interval for β_1 is,

$$\Rightarrow \quad \text{point estimate} \ \pm \ \text{margin of error}$$

$$\Rightarrow \quad \hat{\beta}_1 \ \pm \ \left(t_{\frac{\alpha}{2},\,(n-2)} \ \times s_{\hat{\beta}_1} \right)$$

Example (Contd..)

The article "Characterization of Highway Runoff " for a particular location in BC gave following data and summaries

x = rainfall volume (m^3) and **y = runoff volume** (m^3)

x	5	12	14	17	23	30	40	47	55	67	72	81	96	112	127
y	4	10	13	15	15	25	27	46	38	46	53	70	82	99	100

$n = 15$ $\quad \sum_{i=1}^{n} x_i = 798$ $\quad \sum_{i=1}^{n} x_i^2 = 63{,}040$ $\quad \sum_{i=1}^{n} y_i = 643$ $\quad \sum_{i=1}^{n} y_i^2 = 41{,}999$ $\quad \sum_{i=1}^{n} x_i y_i = 51{,}232$

SST = 14,436 SSR= 14,079

(h) Calculate the point estimate of the standard deviation σ.

(i) Carry out a hypothesis test to decide whether there is a useful linear relationship between rainfall volume and runoff volume. Use $\alpha = 0.05$.

(j) Calculate 95% confidence interval for the true average change in runoff volume associated with a 1 m^3 increase in rainfall.

Example Solutions (Contd..)

(h) Calculate the point estimate of the standard deviation σ.

Point estimate of σ^2 is

$$\hat{\sigma}^2 = s^2 = \frac{SSE}{n-2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}$$

$$SST = SSR + SSE$$

$$14436 = 14079 + SSE$$

$$\Rightarrow \quad SSE = 357$$

$$\text{so,} \quad s^2 = \frac{SSE}{n-2} = \frac{357}{15-2} = 27.46$$

point estimate of σ is $s = \sqrt{27.46} = 5.24$

Example Solutions (Contd..)

(i) Carry out a hypothesis test to decide whether there is a useful linear relationship between rainfall volume and runoff volume. Use $\alpha = 0.05$.

Simple Linear Regression model: $Y = \beta_0 + \beta_1 X + \epsilon$

we test the hypotheses

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

test statistic

$$t = \frac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}} \sim t_{n-2}$$

$$\Rightarrow \quad s_{\hat{\beta}_1} = \frac{s}{s_x \sqrt{n-1}} = \frac{5.24}{38.35\sqrt{15-1}} = 0.0372$$

where $s = 5.34$ and $s_x = 38.35$ from part (h) and part (b)

Example Solutions (Contd..)

(i) then test statistic is

$$t_{obs} = \frac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}} \sim t_{n-2} = t_{13}$$

$$t_{obs} = \frac{0.826}{0.0372} = 22.2$$

$$\alpha = 0.05 \quad \Rightarrow \quad \frac{\alpha}{2} = 0.025 \quad \Rightarrow \quad t_{0.025,\,13} = 2.160$$

Since $|t_{obs}| = 22.2 > t_{0.025,\,13} = 2.160$

$\Rightarrow H_0$ is rejected at $\alpha = 0.05$ (observed test statistic value is in the rejection region. Since this $|t_{obs}|$ value is really large, we have strong evidence to reject H_0 at any usual α value)

<u>Conclusion:</u> We have strong evidence to conclude that there is a useful linear relationship between runoff volume and rainfall volume.

Example Solutions (Contd..)

(j) Calculate 95% confidence interval for the true average change in runoff volume associated with a 1 m^3 increase in rainfall.

True average change in rainfall volume associate with a 1 m^3 increase in rainfall volume is β_1

95% confidence Interval for β_1 is,

$$\hat{\beta}_1 \pm \left(t_{\frac{\alpha}{2},\,(n-2)} \times s_{\hat{\beta}_1} \right)$$

$$\begin{aligned}
\Rightarrow \quad & \hat{\beta}_1 \pm \left(t_{0.025,\,13} \times s_{\hat{\beta}_1} \right) \\
\Rightarrow \quad & 0.826 \pm 2.160\ (0.0372) \\
\Rightarrow \quad & 0.826 \pm 0.080 \\
\Rightarrow \quad & (0.746,\ 0.906)
\end{aligned}$$

$\Rightarrow$ 95% confidence that the true slope parameter (β_1 of the simple linear regression model is between 0.746 and 0.906.

End of Chapter 11

Please Complete the

Course/Instructor

Evaluation

Good Luck with your All Exams

All the very Best!