Ch.3 Inferences in regression analysis:

STAT 532

Assuming Normal error model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

β_0, β_1 parameters

X_i known const.

$\varepsilon_i \sim N(0, \sigma^2)$

3-1 Inferences concerning β_1

→ point estimator

$$\hat{\beta}_1 = b_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

sampling distribution of b_1 is normal, with

mean $E(b_1) = \beta_1$ and variance $\sigma^2(b_1) = \frac{\sigma^2}{\sum (X_i - \bar{X})^2}$

(sampling distribution of b_1 refers to different values of b_1 obtained with repeated sampling when the levels of X are held const.)

Proof:

$$b_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2} - \frac{\sum (X_i - \bar{X}) \bar{Y}}{\sum (X_i - \bar{X})^2}$$

$\sum (X_i - \bar{X}) = 0$

since $\sum (X_i - \bar{X}) = 0$ $\Rightarrow$ $\sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum (X_i - \bar{X}) Y_i$

then $b_1 = \frac{\sum (X_i - \bar{X}) Y_i}{\sum (X_i - \bar{X})^2}$

$\Rightarrow$ $b_1 = \sum k_i Y_i$ ← where $k_i = \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2}$

note that:

1. $\sum k_i = 0$
2. $\sum k_i X_i = 1$

$$\frac{\sum (X_i - \bar{X}) X_i}{\sum (X_i - \bar{X})^2} = \frac{\sum (X_i - \bar{X})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2} = 1$$

3. $\sum k_i^2 = \frac{1}{\sum (X_i - \bar{X})^2}$

Normality:

Since $b_1 = \sum k_i Y_i$ is a linear combination of Y_i

(Y_i ... independent normally distributed r.v.'s)

then b_1 is normally distributed

Mean: $E(b_1) = E(\sum k_i Y_i) = \sum k_i E(Y_i)$

$= \sum k_i (\beta_0 + \beta_1 X_i)$

$= \beta_0 \sum k_i + \beta_1 \sum k_i X_i = 0 + \beta_1 = \beta_1$ #

Variance: $\sigma^2(b_1) = \sigma^2(\sum k_i Y_i) = \sum k_i^2 \sigma^2(Y_i)$

$= \sigma^2 \cdot \sum k_i^2 = \frac{\sigma^2}{\sum (X_i - \bar{X})^2}$ #

Estimated variance:

$$s^2(b_1) = \frac{MSE}{\sum (X_i - \bar{X})^2}$$

unbiased point estimator of $\sigma^2(b_1)$, taking square root we obtain $s(b_1)$ ... point estimator for $\sigma(b_1)$.

* Sampling distribution of $\frac{b_1 - \beta_1}{s(b_1)}$

$b_1 \sim N(\beta_1, \sigma(b_1))$ then $\frac{b_1 - \beta_1}{\sigma(b_1)} \sim N(0,1)$

standardized statistic $\frac{b_1 - \beta_1}{s(b_1)}$ follows a t-distribution with $n-2$ degrees of freedom

proof:

$$\frac{b_1 - \beta_1}{s(b_1)} = \frac{b_1 - \beta_1}{\sigma(b_1)} \div \frac{s(b_1)}{\sigma(b_1)}$$

, $\frac{s^2(b_1)}{\sigma^2(b_1)} = \frac{MSE / \sum (X_i - \bar{X})^2}{\sigma^2 / \sum (X_i - \bar{X})^2} = \frac{MSE}{\sigma^2} = \frac{\frac{SSE}{n-2}}{\sigma^2}$

since $\frac{SSE}{\sigma^2} \sim \chi^2(n-2) \Rightarrow \frac{SSE}{\sigma^2(n-2)} \sim \frac{\chi^2(n-2)}{n-2}$

Recall χ^2-distribution

$$\chi^2(\nu) = Z_1^2 + Z_2^2 + \cdots + Z_\nu^2$$

where Z_i independent standard normally distributed r.v.'s

$$E[\chi^2(\nu)] = \nu$$

---

t-distribution

$$t(\nu) = \frac{Z}{\sqrt{\frac{\chi^2(\nu)}{\nu}}}$$ where Z and $\chi^2(\nu)$ are independent

, $E[t(\nu)] = 0$

F-distribution

$$F(\nu_1, \nu_2) = \frac{\chi^2(\nu_1)}{\nu_1} \div \frac{\chi^2(\nu_2)}{\nu_2}$$ where $\chi^2(\nu_1)$ & $\chi^2(\nu_2)$ are independent

$$[t(\nu)]^2 = F(1, \nu)$$

---

continued proof:

hence, $$\frac{b_1 - \beta_1}{s(b_1)} \sim \frac{Z}{\sqrt{\frac{\chi^2(n-2)}{n-2}}}$$

$$\frac{b_1 - \beta_1}{s(b_1)} \sim t(n-2)$$ #

Confidence interval for β_1

$$P\left\{ t\left(\frac{\alpha}{2}; n-2\right) \leq \frac{b_1 - \beta_1}{s(b_1)} \leq t\left(1-\frac{\alpha}{2}; n-2\right) \right\} = 1-\alpha$$

since $t\left(\frac{\alpha}{2}; n-2\right) = -t\left(1-\frac{\alpha}{2}; n-2\right)$

then $(1-\alpha)100\%$ C.I. for β_1

$$\left[b_1 - t_{1-\frac{\alpha}{2}}\, s(b_1) \;,\; b_1 + t_{1-\frac{\alpha}{2}}\, s(b_1) \right]$$

Example: for Westwood company lot size application estimate 95% C.I. for β_1

$$s^2(b_1) = \frac{MSE}{\sum (X_i - \bar{X})^2} = \frac{7.5}{3,400} = 0.002206$$

$$s(b_1) = 0.04697$$

note:

$$\sum (X_i - \bar{X})^2 = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

$$\sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}$$

$\alpha = 0.05$ $\quad 1 - \frac{\alpha}{2} = 0.975$ $\quad t(0.975; 8) = 2.306$

$$2 - 2.306(0.046) \leq \beta_1 \leq 2 + 2.306(0.046)$$

$$1.89 \leq \beta_1 \leq 2.11$$

Tests concerning β_1

Since $\frac{b_1 - \beta_1}{s(b_1)} \sim t(n-2)$, test using t-distribution

Ex1 Two-sided test – Westwood application

Suppose a cost analyst is interested in testing whether or not there's a linear association between man-hours and lot size.

The two alternatives: $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

Test statistic $T = \frac{b_1}{s(b_1)}$, level of significance α

if $|T| \leq t_{1-\frac{\alpha}{2}} \Rightarrow$ conclude H_0

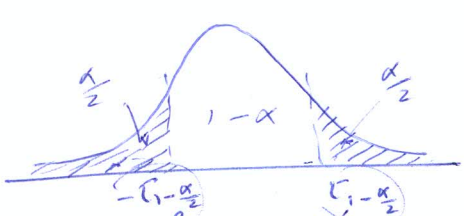
$|T| > t_{1-\frac{\alpha}{2}} \Rightarrow \sim H_1$

$\alpha = 0.05$, $b_1 = 2.$, $s(b_1) = 0.04697$

$t(0.975; 8) = 2.306$

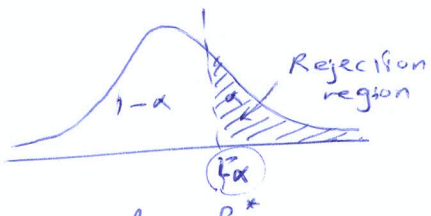
$|T| = \left|\frac{b_1}{s(b_1)}\right| = \left|\frac{2}{0.04697}\right| = 42.58 > 2.306$

accept H_1



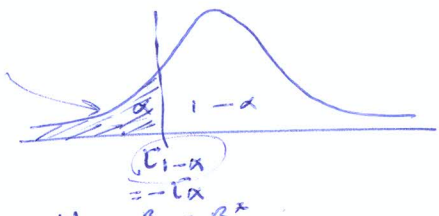
$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$



$H_0: \beta_1 = \beta_1^*$

$H_1: \beta_1 > \beta_1^*$



$H_0: \beta_1 = \beta_1^*$

$H_1: \beta_1 < \beta_1^*$

Ex2 Test whether or not β_1 is positive

$H_0: \beta_1 \leq 0$

$H_1: \beta_1 > 0$

for $\alpha = 0.05$ $\quad t(0.95; 8) = 1.860$

if $T \leq t_{1-\alpha}$ conclude H_0

$T > t_{1-\alpha} \quad \sim H_1$

$T = 42.58 > 1.86$ accept H_1

3.2 Inferences concerning β_0

Sampling distribution of b_0

$\hat{\beta}_0 = b_0 = \bar{Y} - b_1 \bar{X}$ (different values of b_0 obtained with repeated sampling)

Sampling distribution of b_0 is normal, with:

mean: $E(b_0) = \beta_0$ & variance: $\sigma^2(b_0) = \sigma^2 \frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2}$

$= \sigma^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right]$

normality of sampling distribution of b_0 follows because b_0 is a linear combination of Y_i

An estimator of $\sigma^2(b_0)$:

$$s^2(b_0) = MSE \frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2} = MSE \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right]$$

Sampling distribution of $\frac{b_0 - \beta_0}{s(b_0)} \sim t(n-2)$

Confidence interval of β_0

$(1-\alpha)100\%$ C.I. for β_0

$$\left(b_0 - t_{1-\frac{\alpha}{2}} s(b_0) \;,\; b_0 + t_{1-\frac{\alpha}{2}} s(b) \right)$$

Ex: Westwood company:

Find 90% C.I. for β_0 ?

$\alpha = 0.1 \quad 1 - \frac{\alpha}{2} = 0.95 \quad t_{(0.95;8)} = 1.86$

$$s^2(b_0) = MSE \frac{\sum X_i^2}{n(\sum (X_i - \bar{X})^2)} = (7.5) \frac{82{,}400}{10(3{,}400)} = 6.26471$$

$s(b_0) = 2.50294$

$10 + 1.86(2.50294) \leq \beta_0 \leq 10 + 1.86(2.50294)$

$5.34 \leq \beta_0 \leq 14.66$

3-3 Some considerations on making inferences concerning β_0 and β_1:

→ Estimators b_0 and b_1 have the property of asymptotic normality – their distributions approach normality – even if the distributions of Y are far from normal. For large sample t-value is replaced by z-value, in C.I. and decision rules.

→ Since model (3-1) assumes that X_i's are known constants, the confidence interval and risks of error are interpreted with repeated samples in which X observations are kept at the same levels.

→ Spacing of the X levels:

$$\sigma^2(b_1) = \frac{\sigma^2}{\sum(X_i - \bar{X})^2}, \quad \sigma^2(b_0) = \sigma^2 \frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2}$$

indicates that with fixed σ^2, n the variances of b_0, b_1 are affected by the spacing of the X-levels in the observed data.

→ Power of tests:

Consider the test $H_0: \beta_1 = \beta_{1.0}$

$H_1: \beta_1 \neq \beta_{1.0}$

Test statistic: $t^* = \dfrac{b_1 - \beta_{1.0}}{s(b_1)}$

for level of significance α:

if $|t^*| \leq t_{1-\frac{\alpha}{2}}$ conclude H_0

if $|t^*| > t_{1-\frac{\alpha}{2}}$ ~ H_1

The power of the test is: the probability that the decision rule will lead to conclude H_1 when H_1 in fact holds.

$$\text{Power} = P(|t^*| > t_{1-\frac{\alpha}{2}} \mid \delta)$$

$$\delta = \frac{|\beta_1 - \beta_{10}|}{\sigma(b_1)} \quad \text{measure of noncentrality}$$

Ex: West-wood company

find the power of the test when $\beta_1 = 0.25$

$$H_0: \beta_1 = \beta_{10} = 0$$

$$H_1: \beta_1 \neq \beta_{10} = 0$$

Assume $\sigma^2 = 10$, then $\sigma^2(b_1) = \dfrac{\sigma^2}{\sum (x_i - \bar{x})^2} = 0.00294$

$$\sigma(b_1) = 0.05423$$

then, $\delta = |0.25 - 0| / 0.05423 = 4.6$

since $\alpha = 0.05$ & $df = 8$ Reading the ordinate for $\delta = 4.6 \Rightarrow 97\%$, probability $= 0.97$ that we will conclude H_1 ($\beta_1 \neq 0$).

3-4 Interval estimation of $E(Y_h)$

Let X_h be the level of X for which we are interested to estimate the mean response $E(Y_h)$,

$\hat{Y}_h = b_0 + b_1 X_h$, $\hat{Y}_h$ point estimator for $E(Y_h)$

→ sampling distribution of $\hat{Y}_h$ is normal, with

mean $E(\hat{Y}_h) = E(Y_h)$

variance $\sigma^2(\hat{Y}_h) = \sigma^2 \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]$

normality: $\hat{Y}_h$ is a linear combination of Y_i

mean: $E(\hat{Y}_h) = E(b_0 + b_1 X_h) = E(b_0) + X_h E(b_1)$

$= \beta_0 + \beta_1 X_h$

$\hat{Y}_h$ unbiased estimator of $E(Y_h)$

variance: (proof)

$\bar{Y} = \frac{1}{n} \sum Y_i$

$b_1 = \sum k_i Y_i$

since Y_i are independent random variables;

$\sigma(\bar{Y}, b_1) = \frac{1}{n} \sum k_i \sigma^2(Y_i) = \frac{\sigma^2}{n} \sum k_i$

since $\sum k_i = 0$, then $\sigma(\bar{Y}, b_1) = 0$

covariance betn. $\bar{Y}$ and b_1

note: we used theorem:

if Y_i are i/p r.v.'s then the covariance of two linear combinations of Y_i's $\sum a_i Y_i$ & $\sum c_i Y_i$

$= \sum a_i c_i \sigma^2(Y_i)$

To find the variance of $\hat{Y}_h$ we shall use the estimator in alternative form:

$$\hat{Y} = \bar{Y} + b_1(X - \bar{X})$$

then, $\sigma^2(\hat{Y}_h) = \sigma^2(\bar{Y} + b_1(X_h - \bar{X}))$

$$= \sigma^2(\bar{Y}) + (X_h - \bar{X})^2 \sigma^2(b_1)$$

$$= \frac{\sigma^2(Y_i)}{n} + (X_h - \bar{X})^2 \frac{\sigma^2}{\sum(X_i - \bar{X})^2}$$

$$= \sigma^2 \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum(X_i - \bar{X})^2} \right]$$

note: variation in $\hat{Y}_h$ from sample to sample will be greater when X_h is far from the mean $\bar{X}$.

Estimated variance of $\hat{Y}_h$:

$$s^2(\hat{Y}_h) = MSE \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum(X_i - \bar{X})^2} \right]$$

$s(\hat{Y}_h)$ ... square root of $s^2(\hat{Y}_h)$

→ sampling distribution of $[\hat{Y}_h - E(Y_h)] / s(\hat{Y}_h)$

$$\frac{\hat{Y}_h - E(Y_h)}{s(\hat{Y}_h)} \sim t(n-2)$$

→ Confidence interval for $E(Y_h)$

$(1-\alpha)100\%$ C.I. for $E(Y_h)$:

$$\left[\hat{Y}_h - t_{1-\frac{\alpha}{2}} s(\hat{Y}_h), \hat{Y}_h + t_{1-\frac{\alpha}{2}} s(\hat{Y}_h) \right]$$

Ex: Westwood company :

Find 90% C.I. for $E(Y_h)$ when lot size $X_h = 55$

$$\hat{Y}_h = 10.0 + 2.0(55) = 120$$

$$s^2(\hat{Y}_h) = MSE\left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\Sigma(X_i - \bar{X})^2}\right]$$

$$= 7.5\left[\frac{1}{10} + \frac{(55-50)^2}{3,400}\right] = 0.80515$$

$$s(\hat{Y}_h) = 0.89730$$

90% C.I. for $E(Y_h)$ $\quad \alpha = 0.1 \quad 1 - \frac{\alpha}{2} = 0.95$

$$120 - 1.86(0.8973) \leq E(Y_h) \leq 120 + 1.86(0.8973)$$

$$118.3 \leq E(Y_h) \leq 121.7$$

3-5 Prediction of new observations :

Now, we consider the prediction of a new response $Y_{h(new)}$, corresponding to a given level of X.

→ Prediction interval when regression parameters are known :

Assume Westwood company example:

parameters of regression model are given as:

$$\beta_0 = 9.5 \quad , \quad \beta_1 = 2.1 \quad , \quad \sigma^2 = 10$$

$$E(Y) = \beta_0 + \beta_1 X$$

for $X_h = 40$, $E(Y_h) = 9.5 + 2.1(40) = 93.5$

probability distribution of Y at $X_h = 40$ is normal, has a mean $E(Y_h) = 93.5$ and std. dev. $\sigma = \sqrt{10} = 3.162$

Suppose we predict that $Y_{h(new)}$ for the next lot of $X_h = 40$, will be between :

$$E(Y_h) \pm 3\sigma = 93.5 \pm 3(3.162)$$

$$84.0 \leq Y_{h(new)} \leq 103.0$$

* The basic idea of a prediction interval is to choose a range in the distribution of Y wherein most of the observations will fall

When the regression parameters are known, the $1-\alpha$ prediction limits for $Y_{h(new)}$ are :

$$E(Y_h) \pm z_{1-\frac{\alpha}{2}} \sigma$$

→ Prediction interval for $Y_{h(new)}$ when regression parameters are unknown:

prediction limits for a new observation Y at at a given level X_h are obtained by:

$$\frac{\hat{Y}_h - Y}{s(Y_{h(new)})} \sim t(n-2)$$

we use $\hat{Y}_h$ rather than $E(Y_h)$ which is unknown

$(1-\alpha)$ prediction limits for a new observation:

$$\hat{Y}_h \pm t_{1-\frac{\alpha}{2}}\, s(Y_{h(new)})$$

$$\sigma^2(Y_{h(new)}) = \sigma^2(\hat{Y}_h - Y) = \sigma^2(\hat{Y}_h) + \sigma^2$$

$\sigma^2(\hat{Y}_h)$: variance of the sampling distribution of $\hat{Y}_h$

σ^2: variance of the distribution of Y

$$s^2(Y_{h(new)}) = s^2(\hat{Y}_h) + MSE$$

$$= MSE\left[1 + \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right]$$

Ex: Westwood company

Find 90% prediction interval for next observation at $X_h = 55$, where parameter values are unknown.

$$\hat{Y}_h = 120 \quad , \quad s^2(\hat{Y}_h) = 0.80515 \quad , \quad MSE = 7.5$$

$$s^2(Y_{h(new)}) = 0.80515 + 7.5 = 8.30515$$

$$s(Y_{h(new)}) = 2.88187$$

$$120 - 1.86(2.88187) \leq Y_{h(new)} \leq 120 + 1.86(2.88187)$$

$$114.6 \leq Y_{h(new)} \leq 125.4$$

→ Prediction of mean of m new observations for given X_h :

Mean value of Y to be predicted $\bar{Y}_{h(new)}$

$(1-\alpha)$ prediction limits :

$$\hat{Y}_h \pm t_{1-\frac{\alpha}{2}} \, s(\bar{Y}_{h(new)})$$

where : $$s^2(\bar{Y}_{h(new)}) = s^2(\hat{Y}_h) + \frac{MSE}{m}$$

$$= MSE \left[\frac{1}{m} + \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]$$

Ex: Westwood Company :

Find 90% prediction interval for $\bar{Y}_{h(new)}$ in 3 new production runs, each at $X_h = 55$

$$s^2(\bar{Y}_{h(new)}) = 0.80515 + \frac{7.5}{3} = 3.30515$$

$$s(\bar{Y}_{h(new)}) = 1.81801$$

$$120 - 1.86(1.81801) \leq \bar{Y}_{h(new)} \leq 120 + 1.86(1.81801)$$

$$116.6 \leq \bar{Y}_{h(new)} \leq 123.4$$

3-6 Considerations in applying regression analysis:

Uses of regression analysis:

1- To make inferences about regression parameters

2- To estimate the mean response for a given X.

3- To predict a new observation Y for a given X.

[1] Validity of the regression application depends upon whether basic conditions in the period ahead will be similar to those in the period in which regression analysis is based.

[2] In predicting new observation for Y, the independent variable X itself often has to be predicted.

[3] Inferences at levels of X which fall outside the range of observations.

[4] Statistical test that leads to a conclusion $\beta_1 \neq 0$ doesn't make a cause-and-effect relation betn. X and Y.

[5] Confidence coeff.'s for prediction limits apply for a single level of X for a given sample.

3-7 Case when X is random:

Normal error model assumes that X values are known constants. In many situations, it may be preferable to consider both X and Y random variables. All previous results for model (3-1) still apply if:

(1) The conditional distributions of Y_i given X_i are normal and independent

(2) X_i's are independent r.v.'s with probability distribution $g(X_i)$ doesn't involve $\beta_0, \beta_1, \sigma^2$.

3-8 Analysis of Variance Approach to regression analysis:

→ Partitioning of total sum of squares:

The measure of total variation of Y_i, SSTO, sum of squared deviations:

$$SSTO = \sum (Y_i - \bar{Y})^2$$

Using regression approach, deviation of the Y observations around regression line $Y_i - \hat{Y}_i$

The measure of variation with the regression model:

$$SSE = \sum (Y_i - \hat{Y}_i)^2 \quad \text{(error sum of squares)}$$

Difference between these two sums of squares:

$$SSR = \sum (\hat{Y}_i - \bar{Y})^2 \quad \text{(regression sum of squares)}$$

$$SSR = SSTO - SSE$$

Formal development of partitioning:

$$Y_i - \bar{Y} = \hat{Y}_i - \bar{Y} + Y_i - \hat{Y}_i$$

total deviation | deviation of fitted regression value around $\bar{Y}$ | deviation around regression line

$$\underbrace{\sum (Y_i - \bar{Y})^2}_{SSTO} = \underbrace{\sum (\hat{Y}_i - \bar{Y})^2}_{SSR} + \underbrace{\sum (Y_i - \hat{Y}_i)^2}_{SSE}$$

proof:

$$\sum (Y_i - \bar{Y})^2 = \sum [\hat{Y}_i - \bar{Y} + Y_i - \hat{Y}_i]^2$$

$$= \sum [(\hat{Y}_i - \bar{Y})^2 + (Y_i - \hat{Y}_i)^2 + 2(\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i)]$$

$$= \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y}_i)^2 + 2\sum (\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i)$$

$$\sum (\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i) = \sum \hat{Y}_i (Y_i - \hat{Y}_i) - \bar{Y} \sum (Y_i - \hat{Y}_i)$$

(each of the two terms on the right $\to 0$)

$$SSTO = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = \sum Y_i^2 - n\bar{Y}^2$$

$$SSR = b_1 \left[\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} \right]$$

$$= b_1 \left[\sum (X_i - \bar{X})(Y_i - \bar{Y}) \right]$$

or $$SSR = b_1^2 \sum (X_i - \bar{X})^2$$

Degrees of freedom

$SSE \sim (n-2)$ & $SSR \sim 1$

$SSTO \sim n-1$

* Mean squares:

regression mean square: $MSR = \frac{SSR}{1} = SSR$

error mean square $MSE = \frac{SSE}{n-2}$

Analysis of variance table:

ANOVA table for simple regression:

Source of variation	SS	df	MS	E(MS)
Regression	$SSR = \sum(\hat{Y}_i - \bar{Y})^2$	1	$MSR = \frac{SSR}{1}$	$\sigma^2 + \beta_1^2 \sum(X_i - \bar{X})^2$
Error	$SSE = \sum(Y_i - \hat{Y}_i)^2$	$n-2$	$MSE = \frac{SSE}{n-2}$	σ^2
Total	$SSTO = \sum(Y_i - \bar{Y})^2$	$n-1$		

$$SSTO = \sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2$$

$SSTO = \sum Y_i^2$ Total uncorrected sum of squares

correction of mean $SS = n\bar{Y}^2$

Modified ANOVA table for simple regression:

Source of variation	SS	df	MS
Regression	$SSR = \sum(\hat{Y}_i - \bar{Y})^2$	1	$MSR = \frac{SSR}{1}$
Error	$SSE = \sum(Y_i - \hat{Y}_i)^2$	$n-2$	$MSE = \frac{SSE}{n-2}$
Total	$SSTO = \sum(Y_i - \bar{Y})^2$	$n-1$	
Correction for mean	$SS = n\bar{Y}^2$	1	
Total uncorrected	$SSTOU = \sum Y_i^2$	n	