

Stats 100C – Linear Models

University of California, Los Angeles

Duc Vu

Fall 2021

This is stats 100C – Linear Models taught by Professor Christou. There is not an official textbook used for the course. Instead, handouts and reference materials are distributed and can be accessed through the class [website](#). You can find other math/stats lecture notes through my personal [blog](#). Let me know through my [email](#) if you notice something mathematically wrong/concerning. Thank you!

Contents

| | |
|--|-----------|
| 1 Lec 1: Sep 27, 2021 | 4 |
| 1.1 Simple Linear Regression Models | 4 |
| 1.2 Prediction Problem | 4 |
| 2 Lec 2: Sep 29, 2021 | 6 |
| 2.1 Linear Regression | 6 |
| 3 Lec 3: Oct 1, 2021 | 10 |
| 3.1 Gauss-Markov Theorem | 10 |
| 3.2 Estimation of Variance | 10 |
| 3.3 Distribution Theory | 12 |
| 4 Lec 4: Oct 4, 2021 | 13 |
| 4.1 Centered Model | 13 |
| 4.2 Distribution Theory Using the Centered Model | 13 |
| 5 Lec 5: Oct 6, 2021 | 16 |
| 5.1 Distribution Theory Using Non-Centered Model | 16 |
| 5.2 A Note on Gamma Distribution | 17 |
| 5.3 Coefficient of Determination | 17 |
| 6 Lec 6: Oct 8, 2021 | 19 |
| 6.1 Variance & Covariance Operations | 19 |
| 6.2 Inference | 20 |
| 6.3 Prediction Interval | 21 |
| 7 Lec 7: Oct 11, 2021 | 22 |
| 7.1 Hypothesis Testing | 22 |
| 8 Lec 8: Oct 13, 2021 | 25 |
| 8.1 Likelihood Ratio Test | 25 |
| 8.2 Power Analysis in Simple Regression | 26 |

| | |
|---|-----------|
| 9 Lec 9: Oct 15, 2021 | 28 |
| 9.1 Extra Sum of Squares Method | 28 |
| 9.2 Power Analysis Using Non-Central F Distribution | 30 |

List of Theorems

List of Definitions

| | | |
|-----|-----------------|----|
| 7.2 | F Distribution | 23 |
| 8.1 | Non-central t | 26 |

§1 | Lec 1: Sep 27, 2021

§1.1 Simple Linear Regression Models

Consider

$$Y_i = \mu + \varepsilon_i$$

with $\varepsilon_i \stackrel{\text{i.i.d}}{\sim} N(0, \sigma)$; specifically, $Y_1, \dots, Y_n \stackrel{\text{i.i.d}}{\sim} N(\mu, \sigma)$. We want to estimate μ and σ^2 using least squares or method of maximum likelihood (MML).

Method of Least Squares (OLS – Ordinary Least Squares):

$$\begin{aligned} \min Q &= \sum_{i=1}^n (Y_i - \mu)^2 \\ \frac{\partial Q}{\partial \mu} &= -2 \sum (Y_i - \mu) = 0 \\ \sum Y_i - n\hat{\mu} &= 0 \\ \implies \hat{\mu} &= \bar{Y} \end{aligned}$$

Method of Maximum Likelihood (MML):

$$\begin{aligned} f(y_i) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - \mu)^2} \\ &= (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}(y_i - \mu)^2} \\ L = f(y_1) \dots f(y_n) &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum (y_i - \mu)^2} \\ \ln L &= -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum (y_i - \mu)^2 \\ \frac{\partial \ln L}{\partial \mu} &= 0, \quad \frac{\partial \ln L}{\partial \sigma^2} = 0 \end{aligned}$$

Solve the above, we obtain the MLE of μ and σ^2

$$\hat{\mu} = \hat{y}, \quad \hat{\sigma}^2 = \frac{\sum (y_i - \hat{\mu})^2}{n} = \frac{\sum (y_i - \bar{y})^2}{n}$$

Notice that $\hat{\sigma}^2$ is biased and we adjust it to be unbiased as follows

$$S^2 = \frac{\sum (y_i - \bar{y})^2}{n-1}$$

§1.2 Prediction Problem

Given Y_1, \dots, Y_n , we want to predict a new Y , e.g., Y_0 . An educated guess here is

$$\hat{Y}_0 = \bar{Y}$$

1. Predictor assumption: $\hat{Y}_0 = \sum_{i=1}^n a_i Y_i$
2. We want \hat{Y}_0 to be unbiased, i.e., $E\hat{Y}_0 = \mu$

$$\begin{aligned} E \sum a_i Y_i &= \mu \\ \sum a_i EY_i &= \mu \\ \implies \sum a_i &= 1 \end{aligned}$$

3. Minimize the mean square error of prediction, i.e.,

$$E(Y_0 - \hat{Y}_0)^2 \quad \text{s.t.} \quad \sum a_i = 1$$

Notice that this is a constraint optimization problem, we use the method of Lagrange multiplier to obtain

$$\min Q = E(Y_0 - \hat{Y}_0)^2 - 2\lambda \left(\sum a_i - 1 \right)$$

Note: $EW^2 = \text{var}(W) + (EW)^2$

$$\begin{aligned} \min Q &= \text{var}(Y_0 - \hat{Y}_0) - 2\lambda \left[\sum a_i - 1 \right] \\ &= \text{var}(Y_0) + \text{var}(\hat{Y}_0) - 2 \text{cov}(Y_0, \hat{Y}_0) - 2\lambda \left[\sum a_i - 1 \right] \\ &= \sigma^2 + \sigma^2 \sum a_i^2 - 2\lambda \left[\sum a_i - 1 \right] \\ \frac{\partial Q}{\partial a_i} &= 2\sigma^2 a_i - 2\lambda = 0 \\ a_i &= \frac{\lambda}{\sigma^2} \end{aligned}$$

Notice that $a_1 = a_2 = \dots = a_n = \frac{\lambda}{\sigma^2}$. So

$$\sum a_i = \frac{n\lambda}{\sigma^2} = 1 \implies \lambda = \frac{\sigma^2}{n}$$

Thus, we can see that

$$a_i = \frac{1}{n}$$

and therefore since $\hat{Y}_0 = \sum a_i Y_i$, it follows that $\hat{Y}_0 = \bar{Y}$.

Prediction Interval:

$$Y_0 - \hat{Y}_0 \sim N\left(0, \sigma \sqrt{1 + \frac{1}{n}}\right)$$

Recall from 100B

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

So,

$$\frac{\frac{Y_0 - \hat{Y}_0 - 0}{\sigma \sqrt{1 + \frac{1}{n}}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2} / (n-1)}} = \frac{Y_0 - \hat{Y}_0}{S \sqrt{1 + \frac{1}{n}}} \sim t_{n-1}$$

We can now construct the prediction interval for Y_0 as follows

$$P\left(-t_{\frac{\alpha}{2}; n-1} \leq \frac{Y_0 - \hat{Y}_0}{S \sqrt{1 + \frac{1}{n}}} \leq t_{\frac{\alpha}{2}; n-1}\right) = 1 - \alpha$$

Finally, $Y_0 \in \hat{Y}_0 \pm t_{\frac{\alpha}{2}; n-1} S \sqrt{1 + \frac{1}{n}}$.

Remark 1.1. Compare this to the confidence interval for μ : $\mu \in \bar{Y} \pm t_{\frac{\alpha}{2}; n-1} \frac{S}{\sqrt{n}}$.

§2 | Lec 2: Sep 29, 2021

§2.1 Linear Regression

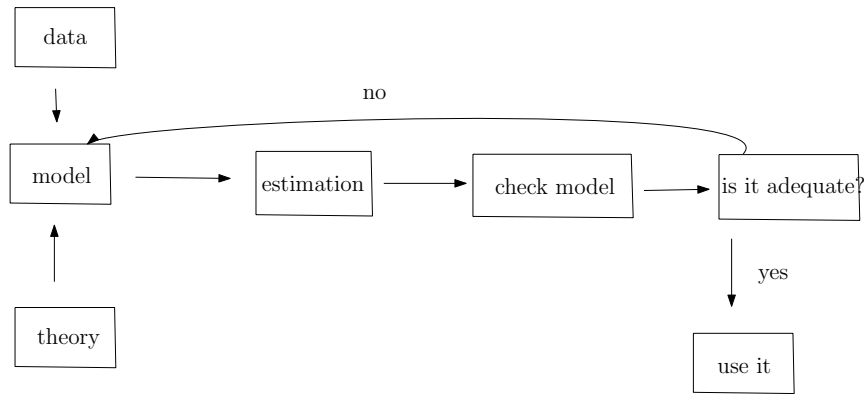
Consider a simple regression model

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

or $Y_i = \beta_1 X_i + \varepsilon_i$

Data:

| y | x |
|----------|----------|
| y_1 | x_1 |
| \vdots | \vdots |
| y_n | x_n |



where the parameters are

$$\begin{cases} \beta_0 : \text{intercept} \\ \beta_1 : \text{slope} \end{cases}$$

and X_1, \dots, X_n are predictors that are not random; $\varepsilon_1, \dots, \varepsilon_n$ are random error terms/disturbance/stochastic terms, and Y_1, \dots, Y_n are random response variable.

Assumption (Gauss-Markov Conditions):

$$E(\varepsilon_i) = 0, \quad \text{var}(\varepsilon_i) = \sigma^2$$

$\varepsilon_1, \dots, \varepsilon_n$ are independent. Using the Gauss-Markov conditions,

$$\begin{aligned} EY_i &= \beta_0 + \beta_1 X_i \\ \text{var}(Y_i) &= \sigma^2 \\ \min Q &= \sum \varepsilon_i^2 \\ \min Q &= \sum (Y_i - \beta_0 - \beta_1 X_i)^2 \\ \frac{\partial Q}{\partial \beta_0} &= -2 \sum (Y_i - \beta_0 - \beta_1 X_i) = 0 \\ \frac{\partial Q}{\partial \beta_1} &= -2 \sum (Y_i - \beta_0 - \beta_1 X_i) X_i = 0 \end{aligned}$$

So,

$$\begin{aligned} & \begin{cases} \sum y_i - n\beta_0 - \beta_1 \sum x_i = 0 \\ \sum x_i y_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2 = 0 \end{cases} \\ \Rightarrow & \begin{cases} n\beta_0 + \beta_1 \sum x_i = \sum y_i \\ \beta_0 \sum x_i + \beta_1 \sum x_i^2 = \sum x_i y_i \end{cases} \quad \text{-- normal equations} \end{aligned}$$

We can solve the above to get $\hat{\beta}_0, \hat{\beta}_1$.

$$\begin{aligned} & \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \\ & \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} \end{aligned}$$

Determinant of the matrix:

$$\begin{aligned} n \sum x_i^2 - \left(\sum x_i \right)^2 &= n \left[\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right] \\ &= n \sum (x_i - \bar{x})^2 \geq 0 \end{aligned}$$

If $x_1 = x_2 = \dots = x_n = \bar{x}$ then $\sum (x_i - \bar{x})^2 = 0$. From normal equations we get

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad \text{from (1)}$$

and plug (1) into (2) to obtain

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \frac{1}{n} (\sum x_i)(\sum y_i)}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

or

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

or

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} \quad (*)$$

or

$$\hat{\beta}_1 = \frac{\sum (y_i - \bar{y}) x_i}{\sum (x_i - \bar{x})^2}$$

or

$$\hat{\beta}_1 = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

Note: From (*), we have

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} \\ &= \frac{(x_1 - \bar{x}) y_1}{\sum (x_i - \bar{x})^2} + \dots + \frac{(x_n - \bar{x}) y_n}{\sum (x_i - \bar{x})^2} \\ &= k_1 y_1 + \dots + k_n y_n = \sum_{i=1}^n k_i y_i \end{aligned}$$

where $k_i = \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2}$. Notice that

$$\begin{aligned}\sum k_i &= 0 \\ \sum k_i^2 &= \frac{1}{\sum (x_i - \bar{x})^2} \\ \sum k_i x_i &= \frac{\sum (x_i - \bar{x}) x_i}{\sum (x_i - \bar{x})^2} = 1\end{aligned}$$

Properties of $\hat{\beta}_1$:

$$\begin{aligned}E\hat{\beta}_1 &= E \sum k_i y_i = \sum k_i E y_i \\ &= \sum k_i (\beta_0 + \beta_1 x_i) \\ &= \beta_0 \sum k_i + \beta_1 \sum k_i x_i \\ &= \beta_1 - \text{unbiased}\end{aligned}$$

For the variance,

$$\begin{aligned}\text{var}(\hat{\beta}_1) &= \text{var}\left(\sum k_i y_i\right) \\ &= \sum k_i^2 \text{var}(Y_i) \\ &= \frac{\sigma^2}{\sum (x_i - \bar{x})^2}\end{aligned}$$

Properties of $\hat{\beta}_0$:

$$\begin{aligned}\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \\ &= \sum \frac{y_i}{n} - \bar{x} \sum k_i y_i \\ &= \sum \left(\frac{1}{n} - \bar{x} k_i \right) y_i \\ &= \sum_{i=1}^n l_i y_i\end{aligned}$$

where $l_i = \frac{1}{n} - \bar{x} k_i$ and the properties of l_i are

$$\begin{aligned}\sum l_i &= 1 \\ \sum l_i^2 &= \sum \left(\frac{1}{n} - \bar{x} k_i \right)^2 = \sum \left(\frac{1}{n^2} + \bar{x}^2 k_i^2 - \frac{2}{n} \bar{x} k_i \right) \\ &= \frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} \\ \sum l_i x_i &= 0\end{aligned}$$

Now, we can easily show that $\hat{\beta}_0$ is unbiased

$$\begin{aligned}E\hat{\beta}_0 &= E \sum l_i y_i = \sum l_i E y_i \\ &= \sum l_i (\beta_0 + \beta_1 x_i) = \beta_0 \sum l_i + \beta_1 \sum l_i x_i \\ &= \beta_0\end{aligned}$$

Thus,

$$\text{var}(\hat{\beta}_0) = \text{var}\left(\sum l_i y_i\right) = \sigma^2 \sum l_i^2 = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}\right)$$

The fitted value is

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = \bar{y} + \hat{\beta}_1 (x_i - \bar{x})$$

and the residual is defined as

$$e_i = y_i - \hat{y}_i$$

with properties

$$\begin{aligned}\sum e_i &= 0 \\ \sum e_i x_i &= 0 \\ \sum e_i \hat{y}_i &= 0\end{aligned}$$

Estimation Using MML:

Assume $\varepsilon_1, \dots, \varepsilon_n \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma)$. Then $Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma)$. The log-likelihood function is

$$\ln L = -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum (y_i - \beta_0 - \beta_1 x_i)^2$$

So, we need to solve

$$\frac{\partial \ln L}{\partial \beta_0} = 0, \quad \frac{\partial \ln L}{\partial \beta_1} = 0$$

to get $\hat{\beta}_0, \hat{\beta}_1$ which are the same as least squares method.

$$\begin{aligned}\frac{\partial \ln L}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (y_i - \beta_0 - \beta_1 x_i)^2 = 0 \\ \hat{\sigma}^2 &= \frac{\sum e_i^2}{n}\end{aligned}$$

Then,

$$\sum (y_i - \bar{y})^2 = \sum \left(\underbrace{y_i - \hat{y}_i}_{e_i} + \hat{y}_i - \bar{y} \right)^2$$

in which we expand to get

$$\underbrace{\sum (y_i - \bar{y})^2}_{\text{SST}} = \underbrace{\sum e_i^2}_{\text{SSE}} + \underbrace{\sum (\hat{y}_i - \bar{y})^2}_{\text{SSR}}$$

in which

$$\begin{cases} \text{SST: sum of squares total} \\ \text{SSE: sum of squares error} \\ \text{SSR: sum of squares regression} \end{cases}$$

§3 | Lec 3: Oct 1, 2021

§3.1 Gauss-Markov Theorem

Recall

$$\hat{\beta}_1 = \sum k_i Y_i$$

where $k_i = \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2}$. Consider now

$$b_1 = \sum a_i Y_i$$

which is another unbiased estimator of β_1 . Then $E b_1 = \beta_1$ or $E \sum a_i Y_i = \beta_1$. So

$$\begin{aligned} \beta_1 &= \sum a_i E Y_i \\ &= \sum a_i (\beta_0 + \beta_1 X_i) \\ &= \beta_0 \sum a_i + \beta_1 \sum a_i X_i \end{aligned}$$

Thus,

$$\begin{cases} \sum a_i = 0 \\ \sum a_i x_i = 1 \end{cases}$$

and we know that

$$\text{var}(b_1) = \text{var}\left(\sum_{i=1}^n a_i Y_i\right) = \sigma^2 \sum a_i^2$$

and

$$\text{var}(\hat{\beta}_1) = \sigma^2 \sum k_i^2 = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}$$

Now let $a_i = k_i + d_i$. Then,

$$\begin{aligned} \text{var}(b_1) &= \sigma^2 \sum (k_i + d_i)^2 \\ &= \sigma^2 \sum k_i^2 + \sigma^2 \sum d_i^2 + 2\sigma^2 \sum k_i d_i \end{aligned}$$

We need to show $\sum k_i d_i = 0$.

$$\begin{aligned} \sum k_i (a_i - k_i) &= \sum k_i a_i - \sum k_i^2 \\ &= \frac{\sum (x_i - \bar{x}) a_i}{\sum (x_i - \bar{x})^2} - \frac{1}{\sum (x_i - \bar{x})^2} \\ &= \frac{\sum x_i a_i}{\sum (x_i - \bar{x})^2} - \frac{\bar{x} \sum a_i}{\sum (x_i - \bar{x})^2} - \frac{1}{\sum (x_i - \bar{x})^2} \\ &= 0 \end{aligned}$$

So $\text{var}(b_1) \geq \text{var}(\hat{\beta}_1)$ and therefore $\hat{\beta}_1$ is the best linear unbiased estimator (BLUE).

§3.2 Estimation of Variance

Using MML

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n}$$

Is it unbiased?

$$E \hat{\sigma}^2 = \frac{\sum E e_i^2}{n} = \frac{\sum [\text{var}(e_i) + (E e_i)^2]}{n}$$

Note: $e_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$. So

$$Ee_i = E[Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i] = (\beta_0 + \beta_1 X_i) - (\beta_0 + \beta_1 X_i) = 0$$

Then,

$$E\hat{\sigma}^2 = \frac{\sum \text{var}(e_i)}{n}$$

Notice that

$$e_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$$

or

$$e_i = Y_i - \bar{Y} - \hat{\beta}_1(X_i - \bar{X})$$

where $\hat{Y}_i = \bar{Y} + \hat{\beta}_1(X_i - \bar{X})$. Substitute in and we get

$$\begin{aligned} \text{var}(e_i) &= \text{var}[Y_i - \bar{Y} - \hat{\beta}_1(X_i - \bar{X})] \\ &= \text{var}(Y_i) + \text{var}(\bar{Y}) + (X_i - \bar{X})^2 \text{var}(\hat{\beta}_1) - 2 \text{cov}(Y_i, \bar{Y}) - 2(X_i - \bar{X}) \text{cov}(Y_i, \hat{\beta}_1) \\ &\quad + 2(X_i - \bar{X}) \text{cov}(\bar{Y}, \hat{\beta}_1) \end{aligned}$$

Let's compute each term there.

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + \varepsilon_i \\ \text{var}(Y_i) &= \sigma^2 \\ \bar{Y} &= \beta_0 + \beta_1 \bar{X} + \frac{\sum \varepsilon_i}{n} \\ \text{var}(\bar{Y}) &= \frac{\sigma^2}{n} \\ \text{cov}(Y_i, \bar{Y}) &= \text{cov}\left(Y_i, \frac{Y_1 + \dots + Y_i + \dots + Y_n}{n}\right) \\ &= \frac{1}{n} \text{cov}(Y_i, Y_1) + \dots + \frac{1}{n} \text{cov}(Y_i, Y_i) + \dots + \frac{1}{n} \text{cov}(Y_i, Y_n) \\ &= \frac{\sigma^2}{n} \\ \text{cov}(Y_i, \hat{\beta}_1) &= \text{cov}\left(Y_i, \sum k_i Y_i\right) \\ &= \text{cov}(Y_i, k_1 Y_1) + \dots + \text{cov}(Y_i, k_i Y_i) + \dots + \text{cov}(Y_i, k_n Y_n) \\ &= k_1 \text{cov}(Y_i, Y_1) + \dots + k_i \text{cov}(Y_i, Y_i) + \dots + k_n \text{cov}(Y_i, Y_n) \\ &= \sigma^2 k_i = \sigma^2 \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2} \end{aligned}$$

Note: A property of covariance

$$\text{cov}(aY, bQ) = ab \text{cov}(Y, Q)$$

And for the last term,

$$\begin{aligned} \text{cov}(\bar{Y}, \hat{\beta}_1) &= \text{cov}\left(\frac{Y_1 + \dots + Y_n}{n}, k_1 Y_1 + \dots + k_n Y_n\right) \\ &= \text{cov}\left(\frac{Y_1}{n}, k_1 Y_1 + \dots + k_n Y_n\right) + \dots + \text{cov}\left(\frac{Y_n}{n}, k_1 Y_1 + \dots + k_n Y_n\right) \\ &= \frac{\sigma^2}{n} k_1 + \frac{\sigma^2}{n} k_2 + \dots + \frac{\sigma^2}{n} k_n \\ &= \frac{\sigma^2}{n} \sum k_i = 0 \end{aligned}$$

Now, we're ready to compute the variance

$$\begin{aligned}\text{var}(e_i) &= \sigma^2 + \frac{\sigma^2}{n} + \frac{\sigma^2(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} - \frac{2\sigma^2}{n} - \frac{2\sigma^2(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} \\ &= \sigma^2 \left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)\end{aligned}$$

Therefore,

$$\begin{aligned}E\hat{\sigma}^2 &= \frac{\sum \text{var}(e_i)}{n} = \sigma^2 \frac{\sum_{i=1}^n \left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)}{n} \\ &= \frac{(n-2)}{n} \sigma^2\end{aligned}$$

It follows that the unbiased estimator of σ^2 is

$$S_e^2 = \frac{n}{n-2} \sigma^2 = \frac{\sum e_i^2}{n-2}$$

§3.3 Distribution Theory

Let $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ and we assume $\varepsilon_1, \dots, \varepsilon_n \stackrel{\text{i.i.d}}{\sim} N(0, \sigma)$

$$\begin{aligned}\hat{\beta}_1 = \sum k_i Y_i &\implies \hat{\beta}_1 \sim N \left(\beta_1, \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}} \right) \\ \hat{\beta}_0 = \sum l_i Y_i &\implies \hat{\beta}_0 \sim N \left(\beta_0, \sigma \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}} \right)\end{aligned}$$

We will show $\frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$ in the next lecture.

§4 | Lec 4: Oct 4, 2021

§4.1 Centered Model

Consider the model: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i = 1, \dots, n$ and Gauss-Markov conditions hold, i.e.,

$$\begin{aligned} E[\varepsilon_i] &= 0 \\ \text{var}[\varepsilon_i] &= \sigma^2 \end{aligned}$$

for $i = 1, \dots, n$ and $\varepsilon_1, \dots, \varepsilon_n$ are independent (we assume $\varepsilon_1, \dots, \varepsilon_n \stackrel{\text{i.i.d}}{\sim} N(0, \sigma)$). This is non-centered model. Let's look at a centered model

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + \varepsilon_i \\ Y_i &= \beta_0 + \beta_1 \bar{X} + \beta_1 (X_i - \bar{X}) + \varepsilon_i \\ Y_i &= \gamma_0 + \beta_1 Z_i + \varepsilon_i \quad \text{-- centered model} \end{aligned}$$

where $\gamma_0 = \beta_0 + \beta_1 \bar{X}$ and $Z_i = X_i - \bar{X}$.

Note: $\sum z_i = \sum (x_i - \bar{x}) = 0$ and $\bar{z} = 0$. So,

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum (z_i - \bar{z}) y_i}{\sum (z_i - \bar{z})^2} = \frac{\sum z_i y_i}{\sum z_i^2} = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} \quad \text{-- same as non-centered model} \\ \hat{\gamma}_0 &= \bar{y} - \hat{\beta}_1 \bar{z} = \bar{y} \end{aligned}$$

Notice $\hat{y}_i = \bar{y} + \hat{\beta}_1 (x_i - \bar{x})$ which is the same as \hat{y}_i of the non-centered model.

§4.2 Distribution Theory Using the Centered Model

Have

$$\begin{aligned} Y_i &\sim N(\gamma_0 + \beta_1 (X_i - \bar{X}), \sigma) \\ \hat{\beta}_1 &\sim \left(\beta_1, \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}} \right) \\ \hat{\gamma}_0 = \bar{Y} &\sim N\left(\gamma_0, \frac{\sigma}{\sqrt{n}}\right) \end{aligned}$$

Now, let's show that $\frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$. We have

$$\begin{aligned} \frac{Y_i - \gamma_0 - \beta_1 (X_i - \bar{X})}{\sigma} &\sim N(0, 1) \\ \frac{[Y_i - \gamma_0 - \beta_1 (X_i - \bar{X})]^2}{\sigma^2} &\sim \chi_1^2 \end{aligned}$$

It follows that

$$\frac{\sum_{i=1}^n [Y_i - \gamma_0 - \beta_1 (X_i - \bar{X})]^2}{\sigma^2} \sim \chi_n^2$$

Notice that $\frac{(n-2)S_e^2}{\sigma^2} = \frac{\sum \varepsilon_i^2}{\sigma^2}$. Let's manipulate this expression. First, let

$$L = \frac{\sum [Y_i - \gamma_0 - \beta_1 (X_i - \bar{X}) \pm \hat{\gamma}_0 \pm \hat{\beta}_1 (X_i - \bar{X})]^2}{\sigma^2}$$

Then,

$$\begin{aligned}
 L &= \frac{\sum \left[y_i - \hat{\gamma}_0 - \hat{\beta}_1(x_i - \bar{x}) + (\hat{\gamma}_0 - \gamma_0) + (\hat{\beta}_1 - \beta_1)(x_i - \bar{x}) \right]^2}{\sigma^2} \\
 &= \frac{\sum \left[e_i + (\hat{\gamma}_0 - \gamma_0) + (\hat{\beta}_1 - \beta_1)(x_i - \bar{x}) \right]^2}{\sigma^2} \\
 &= \frac{\sum e_i^2}{\sigma^2} + \frac{n(\hat{\gamma}_0 - \gamma_0)^2}{\sigma^2} + \frac{(\hat{\beta}_1 - \beta_1)^2 \sum (x_i - \bar{x})^2}{\sigma^2} + \frac{2(\hat{\gamma}_0 - \gamma_0) \sum e_i}{\sigma^2} \\
 &\quad + \frac{2(\hat{\beta}_1 - \beta_1) \sum e_i(x_i - \bar{x})}{\sigma^2} + \frac{2(\hat{\gamma}_0 - \gamma_0)(\hat{\beta}_1 - \beta_1) \sum (x_i - \bar{x})}{\sigma^2}
 \end{aligned}$$

So far,

$$\underbrace{\frac{\sum [y_i - \gamma_0 - \beta_1(x_i - \bar{x})]^2}{\sigma^2}}_{\chi_n^2} = \underbrace{\frac{(n-2)S_e^2}{\sigma^2}}_{?} + \underbrace{\frac{\hat{\gamma}_0 - \gamma_0}{\sigma/\sqrt{n}}}_{\chi_1^2} + \underbrace{\left[\frac{\hat{\beta}_1 - \beta_1}{\sigma/\sqrt{\sum (x_i - \bar{x})^2}} \right]^2}_{\chi_1^2}$$

$$Q = Q_1 + Q_2 + Q_3$$

Let's use moment generating function to find “?”. Notice that Q_1, Q_2, Q_3 are independent why?

$$M_Q(t) = M_{Q_1+Q_2+Q_3}$$

$$M_Q(t) = M_{Q_1}(t) \cdot M_{Q_2}(t) \cdot M_{Q_3}(t)$$

We have

$$Q \sim \chi_n^2 \implies M_Q(t) = (1 - 2t)^{-\frac{n}{2}}$$

$$Q_2 \sim \chi_1^2 \implies M_{Q_2}(t) = (1 - 2t)^{-\frac{1}{2}}$$

$$Q_3 \sim \chi_1^2 \implies M_{Q_3}(t) = (1 - 2t)^{-\frac{1}{2}}$$

$$\implies M_{Q_1}(t) = (1 - 2t)^{\frac{-n+2}{2}}$$

$$\implies Q_1 = \frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$$

Note: If $Y \sim \Gamma(\alpha, \beta)$ then

$$M_Y(t) = (1 - \beta t)^{-\alpha}$$

and

$$M_{cY}(t) = M_Y(ct)$$

Let's now find the distribution of s_e^2 .

$$S_e^2 = \frac{\sigma^2}{n-2} Q_1$$

$$M_{S_e^2}(t) = M_{\frac{\sigma^2}{n-2} Q_1}(t) = M_{Q_1}\left(\frac{\sigma^2}{n-2} t\right)$$

$$M_{S_e^2}(t) = \left(1 - \frac{2\sigma^2}{n-2} t\right)^{\frac{-n+2}{2}}$$

Therefore,

$$S_e^2 \sim \Gamma\left(\frac{n-2}{2}, \frac{2\sigma^2}{n-2}\right)$$

$$ES_e^2 = \sigma^2, \quad \text{var}(S_e^2) = \frac{2\sigma^4}{n-2}$$

Another way to show this result is to use the non-centered model

$$\frac{\sum \left(Y_i - \beta_0 - \beta_1 X_i \pm \hat{\beta}_0 \pm \hat{\beta}_1 X_i \right)^2}{\sigma^2}$$

§5 | Lec 5: Oct 6, 2021

§5.1 Distribution Theory Using Non-Centered Model

Recall that we want to show $\frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$ using the non-centered model $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ for $\varepsilon_1, \dots, \varepsilon_n \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma)$. Then, $Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma)$. Let

$$M = \frac{\sum (Y_i - \beta_0 - \beta_1 X_i \pm \hat{\beta}_0 \pm \hat{\beta}_1 X_i)^2}{\sigma^2} \sim \chi_n^2$$

Then,

$$\begin{aligned} M &= \frac{\sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i + (\hat{\beta}_0 - \beta_0) + (\hat{\beta}_1 - \beta_1)x_i)^2}{\sigma^2} \\ &= \frac{\sum e_i^2}{\sigma^2} + \frac{n(\hat{\beta}_0 - \beta_0)^2}{\sigma^2} + \frac{(\hat{\beta}_1 - \beta_1)^2 \sum x_i^2}{\sigma^2} + \frac{2(\hat{\beta}_0 - \beta_0) \sum e_i}{\sigma^2} + \frac{2(\hat{\beta}_1 - \beta_1) \sum e_i x_i}{\sigma^2} \\ &\quad + \frac{2(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1) \sum x_i}{\sigma^2} \\ &= \underbrace{\frac{\sum e_i^2}{\sigma^2}}_{\frac{(n-2)S_e^2}{\sigma^2}} + \underbrace{\frac{n(\hat{\beta}_0 - \beta_0)^2}{\sigma^2} + \frac{(\hat{\beta}_1 - \beta_1)^2 \sum x_i^2}{\sigma^2} + \frac{2(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1) \sum x_i}{\sigma^2}}_{?} \end{aligned} \quad (**)$$

Let $D = \hat{\beta}_0 + \hat{\beta}_1 \bar{X} = \bar{Y}$ and consider

$$\frac{(\hat{\beta}_1 - \beta_1)^2}{\text{var}(\hat{\beta}_1)} + \frac{(D - (\beta_0 + \beta_1 \bar{x}))^2}{\text{var}(D)} \quad (*)$$

Note: $\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}\right)$ and

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + \varepsilon_i \\ \bar{Y} &= \frac{\sum Y_i}{n} = \beta_0 + \beta_1 \bar{X} + \frac{\sum \varepsilon_i}{n} \end{aligned}$$

So $\bar{Y} \sim N\left(\beta_0 + \beta_1 \bar{X}, \frac{\sigma}{\sqrt{n}}\right)$ and thus $\frac{D - (\beta_0 + \beta_1 \bar{X})}{\sigma/\sqrt{n}} \sim N(0, 1)$. It follows that each term in (*) follows chi-square distribution with 1 degree of freedom. Now, we have

$$\begin{aligned} (*) &= \frac{(\hat{\beta}_1 - \beta_1)^2}{\sigma^2} \sum (x_i - \bar{x})^2 + \frac{n(\hat{\beta}_0 - \beta_0)^2}{\sigma^2} + \frac{(\hat{\beta}_1 - \beta_1)^2}{\sigma^2} n \bar{x}^2 + \frac{2(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1)}{\sigma^2} \sum x_i \\ &= \frac{(\hat{\beta}_1 - \beta_1)^2 (\sum x_i^2 - n \bar{x}^2)}{\sigma^2} + \frac{n(\hat{\beta}_0 - \beta_0)^2}{\sigma^2} + \frac{(\hat{\beta}_1 - \beta_1)^2 n \bar{x}^2}{\sigma^2} + \frac{2(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1) \sum x_i}{\sigma^2} \end{aligned}$$

which is equivalent to the last three terms of (**). We just need to show that

$$\begin{aligned} \text{cov}(\bar{Y}, \hat{\beta}_1) &= 0 \\ \text{cov}(\bar{Y}, e_i) &= 0 \\ \text{cov}(\hat{\beta}_1, e_i) &= 0 \end{aligned}$$

Remark 5.1. Under normality, zero covariance implies independence.

§5.2 A Note on Gamma Distribution

Let $Q \sim \Gamma(\alpha, \beta)$. Then

$$\begin{aligned} EQ &= \alpha\beta \\ \text{var}(Q) &= \alpha\beta^2 \\ M_Q(t) &= (1 - \beta t)^{-\alpha} \\ EQ^k &= \frac{\Gamma(\alpha + k)\beta^k}{\Gamma(\alpha)} \end{aligned}$$

where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

is the Gamma function.

Property:

$$\begin{aligned} \Gamma(\alpha) &= (\alpha - 1)\Gamma(\alpha - 1) \\ \Gamma(\alpha + 1) &= \alpha\Gamma(\alpha) \end{aligned}$$

If α is an integer, then

$$\Gamma(\alpha) = (\alpha - 1)!$$

Recall that $S_e^2 \sim \Gamma\left(\frac{n-2}{2}, \frac{2\sigma^2}{n-2}\right)$

$$ES_e^2 = \sigma^2, \quad \text{var}(S_e^2) = \frac{2\sigma^4}{n-2}$$

Is S_e unbiased estimator of σ ?

$$\begin{aligned} ES_e &= E[S_e^2]^{\frac{1}{2}} \\ &= \frac{\Gamma\left(\frac{n-2}{2} + \frac{1}{2}\right) \left(\frac{2\sigma^2}{n-2}\right)^{\frac{1}{2}}}{\Gamma\left(\frac{n-2}{2}\right)} \\ &= \sigma \sqrt{\frac{2}{n-2}} \Gamma\left(\frac{n-1}{2}\right) / \Gamma\left(\frac{n-2}{2}\right) \\ &= \sigma A \end{aligned}$$

Thus, it's biased and we can adjust the result to be unbiased, i.e., $\frac{S_e}{A}$.

If $Y \sim \mathcal{X}_n^2$, then

$$M_Y(t) = (1 - 2t)^{-\frac{n}{2}}$$

which is $\Gamma\left(\frac{n}{2}, 2\right)$.

§5.3 Coefficient of Determination

Recall

$$\underbrace{\sum (y_i - \bar{y})^2}_{\text{SST}} = \underbrace{\sum e_i^2}_{\text{SSE}} + \underbrace{\sum (\hat{y}_i - \bar{y})^2}_{\text{SSR}}$$

where $\hat{Y}_i = \bar{y} + \hat{\beta}_1(x_i - \bar{x})$. We define R^2 as

$$R^2 = \frac{\text{SSR}}{\text{SST}} \quad \text{or} \quad R^2 = 1 - \frac{\text{SSE}}{\text{SST}}$$

and $0 \leq R^2 \leq 1$. We have

$$\begin{aligned}\text{var}(\hat{Y}_i) &= \text{var}\left(\bar{y} + \hat{\beta}_1(x_i - \bar{x})\right) \\ &= \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)\end{aligned}$$

Another way to show this is to express \hat{Y}_i as a linear combination of Y_1, \dots, Y_n .

$$\begin{aligned}\hat{Y}_i &= \bar{y} + \hat{\beta}_1(x_i - \bar{x}) \\ &= \frac{\sum y_j}{n} + (x_i - \bar{x}) \sum k_j y_j \\ &= \sum \left[\frac{1}{n} + (x_i - \bar{x})k_j \right] y_j \\ \text{var}(\hat{Y}_i) &= \sigma^2 \sum \left[\frac{1}{n} + (x_i - \bar{x})k_j \right]^2 \\ &= \sigma^2 \sum \left[\frac{1}{n^2} + (x_i - \bar{x})^2 k_j^2 + \frac{2}{n}(x_i - \bar{x})k_j \right] \\ &= \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right)\end{aligned}$$

Consider

$$e_i = y_i - \hat{y}_i = y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x}) = \sum a_l y_l - \frac{\sum y_l}{n} - (x_i - \bar{x}) \sum k_l y_l = \sum \left[a_l - \frac{1}{n} - (x_i - \bar{x})k_l \right] y_l$$

where

$$a_l = \begin{cases} 1, & \text{if } l = i \\ 0, & \text{otherwise} \end{cases}$$

§6 | Lec 6: Oct 8, 2021

§6.1 Variance & Covariance Operations

Have

$$\text{cov}\left(\sum a_i Y_i, \sum b_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \text{cov}(Y_i, Y_j) = \sum a_i b_i \text{cov}(Y_i, Y_i) = \sigma^2 \sum a_i b_i$$

because Y_1, \dots, Y_n are independent.

Example 6.1

Consider $\hat{\beta}_0$ and $\hat{\beta}_1$

$$\begin{aligned} \text{cov}(\hat{\beta}_0, \hat{\beta}_1) &= \text{cov}\left(\sum l_i Y_i, \sum k_i Y_j\right) \\ &= \sigma^2 \sum l_i k_i \\ &= \sigma^2 \sum \left[\left(\frac{1}{n} - k_i \bar{x}\right) k_i\right] \\ &= \sigma^2 \frac{1}{n} \sum k_i - \sigma^2 \bar{x} \sum k_i^2 \\ &= -\frac{\sigma^2 \bar{x}}{\sum (x_i - \bar{x})^2} \end{aligned}$$

Or

$$\begin{aligned} \text{cov}(\hat{\beta}_0, \hat{\beta}_1) &= \text{cov}(\bar{Y} - \hat{\beta}_1 \bar{X}, \hat{\beta}_1) \\ &= \text{cov}(\bar{Y}, \hat{\beta}_1) - \bar{X} \text{var}(\hat{\beta}_1) \\ &= \frac{-\bar{x} \sigma^2}{\sum (x_i - \bar{x})^2} \end{aligned}$$

Example 6.2

Consider \hat{Y}_i and \hat{Y}_j

$$\begin{aligned} \text{cov}(\hat{Y}_i, \hat{Y}_j) &= \text{cov}\left(\bar{y} + \hat{\beta}_1(x_i - \bar{x}), \bar{y} + \hat{\beta}_1(x_j - \bar{x})\right) \\ &= \frac{\sigma^2}{n} + 0 + 0 + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum (x_i - \bar{x})^2} \sigma^2 \\ &= \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum (x_i - \bar{x})^2}\right) \end{aligned}$$

When $i = j$,

$$\text{var}(\hat{Y}_i) = \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right)$$

Example 6.3 (Cont'd)

Notice that

$$\begin{aligned}
 \hat{Y}_i &= \bar{y} + \hat{\beta}_1(x_i - \bar{x}) = \frac{\sum y_l}{n} + (x_i - \bar{x}) \sum k_l y_l \\
 &= \sum \left[\frac{1}{n} + (x_i - \bar{x})k_l \right] y_l = \sum a_l y_l \\
 \hat{Y}_j &= \dots = \sum b_v y_v \\
 \text{cov}(\hat{Y}_i, \hat{Y}_j) &= \sigma^2 \sum a_l b_l \\
 &= \sigma^2 \sum \left[\frac{1}{n} + (x_i - \bar{x})k_l \right] \left[\frac{1}{n} + (x_j - \bar{x})k_l \right] \\
 &= \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum (x_i - \bar{x})^2} \right)
 \end{aligned}$$

§6.2 Inference

Construct a confidence interval $1 - \alpha$ for β_1

$$P(L \leq \beta_1 \leq U) = 1 - \alpha$$

Know

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}\right)$$

and

$$\frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$$

Consider

$$\text{cov}(\hat{\beta}_1, e_i) = 0$$

Under normality, since their covariance is 0, $\hat{\beta}_1$ and S_e^2 are independent. Thus,

$$\frac{\frac{\hat{\beta}_1 - \beta_1}{\sigma / \sqrt{\sum (x_i - \bar{x})^2}}}{\sqrt{\frac{(n-2)S_e^2}{\sigma^2} / (n-2)}} = \frac{\hat{\beta}_1 - \beta_1}{S_e / \sqrt{\sum (x_i - \bar{x})^2}} \sim t_{n-2}$$

Pivot Method:

$$P\left(-t_{\frac{\alpha}{2}; n-2} \leq \frac{\hat{\beta}_1 - \beta_1}{S_e / \sqrt{\sum (x_i - \bar{x})^2}} \leq t_{\frac{\alpha}{2}; n-2}\right) = 1 - \alpha$$

and after some manipulation we get

$$P\left(\hat{\beta}_1 - t_{\frac{\alpha}{2}; n-2} \cdot \frac{S_e}{\sqrt{\sum (x_i - \bar{x})^2}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\frac{\alpha}{2}; n-2} \cdot \frac{S_e}{\sqrt{\sum (x_i - \bar{x})^2}}\right) = 1 - \alpha$$

We are $1 - \alpha$ confident that

$$\beta_1 \in \left[\hat{\beta}_1 \pm t_{\frac{\alpha}{2}; n-2} \cdot \frac{S_e}{\sqrt{\sum (x_i - \bar{x})^2}} \right]$$

For $\hat{\beta}_0$,

$$\hat{\beta}_0 \sim N \left(\beta_0, \sigma \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}} \right)$$

and we proceed similarly to obtain

$$\beta_0 \in \left[\hat{\beta}_0 \pm t_{\frac{\alpha}{2}; n-2} \cdot S_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}} \right]$$

Say if we want to construct a confidence interval for $\beta_0 - 2\beta_1$:

$$\begin{aligned} \text{var}(\hat{\beta}_0 - 2\hat{\beta}_1) &= \text{var}(\hat{\beta}_0) + 4 \text{var}(\hat{\beta}_1) - 4 \text{cov}(\hat{\beta}_0, \hat{\beta}_1) \\ &= \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} + \frac{4}{\sum (x_i - \bar{x})^2} + \frac{4\bar{x}}{\sum (x_i - \bar{x})^2} \right] \\ &= \sigma^2 \left[\frac{1}{n} + \frac{(\bar{x} + 2)^2}{\sum (x_i - \bar{x})^2} \right] \end{aligned}$$

So,

$$\hat{\beta}_0 - 2\hat{\beta}_1 \sim N \left(\beta_0 - 2\beta_1, \sigma \sqrt{\frac{1}{n} + \frac{(\bar{x} + 2)^2}{\sum (x_i - \bar{x})^2}} \right)$$

Thus, the C.I. is

$$\beta_0 - 2\beta_1 \in \left[\hat{\beta}_0 - 2\hat{\beta}_1 \pm t_{\frac{\alpha}{2}; n-2} \cdot S_e \sqrt{\frac{1}{n} + \frac{(\bar{x} + 2)^2}{\sum (x_i - \bar{x})^2}} \right]$$

§6.3 Prediction Interval

Prediction interval for Y_0 when $X = X_0$. Let's begin with error of prediction: $Y_0 - \hat{Y}_0$. We know

- $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
- $Y_0 = \beta_0 + \beta_1 X_0 + \varepsilon_0$
- $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 X_0$

So

$$\begin{aligned} E(Y_0 - \hat{Y}_0) &= 0 \\ \text{var}(Y_0 - \hat{Y}_0) &= \text{var}(Y_0) + \text{var}(\hat{Y}_0) - 2 \text{cov}(Y_0, \hat{Y}_0) \\ &= \sigma^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right) \end{aligned}$$

We apply the same procedure in the inference section

$$\left. \begin{aligned} Y_0 - \hat{Y}_0 &\sim N \left(0, \sigma \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \right) \\ \frac{(n-2)S_e^2}{\sigma^2} &\sim \chi_{n-2}^2 \end{aligned} \right\} \implies Y_0 \in \hat{Y}_0 \pm t_{\frac{\alpha}{2}; n-2} S_e \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

C.I. for EY_0 for a given $X = X_0$

$$\begin{aligned} \hat{Y}_0 &\sim N \left(\beta_0 + \beta_1 X_0, \sigma \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \right) \\ \frac{(n-2)S_e^2}{\sigma^2} &\sim \chi_{n-2}^2 \\ \implies EY_0 &\in \hat{Y}_0 \pm t_{\frac{\alpha}{2}; n-2} \cdot S_e \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \end{aligned}$$

§7 | Lec 7: Oct 11, 2021

§7.1 Hypothesis Testing

Consider the model:

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

Example 7.1

Hypothesis testing examples

$$\begin{aligned} H_0 : \beta_1 &= 0, & H_a : \beta_1 &\neq 0 \\ H_0 : \beta_1 &= 1, & H_a : \beta_1 &\neq 1 \\ H_0 : \beta_0 &= 0, & H_a : \beta_0 &\neq 0 \\ H_0 : \beta_0 + \beta_1 &= 0, & H_a : \beta_0 + \beta_1 &\neq 0 \\ H_0 : \left. \begin{array}{l} \beta_0 = \beta_0^* \\ \beta_1 = \beta_1^* \end{array} \right\}, & H_a : &\text{not true} \end{aligned}$$

Let's consider the following two-sided test

$$\begin{aligned} H_0 : \beta_1 &= 0 \\ H_a : \beta_1 &\neq 0 \end{aligned}$$

Recall under H_0 ,

$$\left. \begin{aligned} \hat{\beta}_1 &\sim N\left(0, \frac{\sigma}{\sqrt{\sum(x_i - \bar{x})^2}}\right) \\ \frac{(n-2)S_e^2}{\sigma^2} &\sim \chi_{n-2}^2 \end{aligned} \right\} \Rightarrow t = \frac{\hat{\beta}_1}{S_e / \sqrt{\sum(x_i - \bar{x})^2}} \sim t_{n-2}$$

We reject H_0 if $t > t_{\frac{\alpha}{2}; n-2}$ or $t < -t_{\frac{\alpha}{2}; n-2}$. Using a $1 - \alpha$ C.I.

$$\beta_1 \in \hat{\beta}_1 \pm t_{\frac{\alpha}{2}; n-2} \frac{S_e}{\sqrt{\sum(x_i - \bar{x})^2}}$$

For example, for $-2 \leq \beta_1 \leq 2$, we do not reject H_0 .

$$p\text{-value} = 2P(t > t^*)$$

We reject H_0 if $p\text{-value} < \alpha$.

Test $H_0 : \beta_1 = 0$ using the F statistics. Under H_0 ,

$$\begin{aligned} \hat{\beta}_1 &\sim N\left(0, \frac{\sigma}{\sqrt{\sum(x_i - \bar{x})^2}}\right) \\ \frac{\hat{\beta}_1 - 0}{\sigma / \sqrt{\sum(x_i - \bar{x})^2}} &\sim N(0, 1) \end{aligned}$$

Then,

$$\frac{\hat{\beta}_1^2 \sum(x_i - \bar{x})^2}{\sigma^2} \sim \chi_1^2$$

and we know

$$\frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$$

Therefore, we can form the F statistics

$$\frac{\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sigma_e^2} / 1}{\frac{(n-2)S_e^2}{\sigma^2} / (n-2)} = \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2} \sim F_{1, n-2}$$

Definition 7.2 (F Distribution) — Let $U \sim \mathcal{X}_n^2$ and $V \sim \mathcal{X}_m^2$ and U, V are independent. Then,

$$\frac{\frac{U}{n}}{\frac{V}{m}} \sim F_{n,m}$$

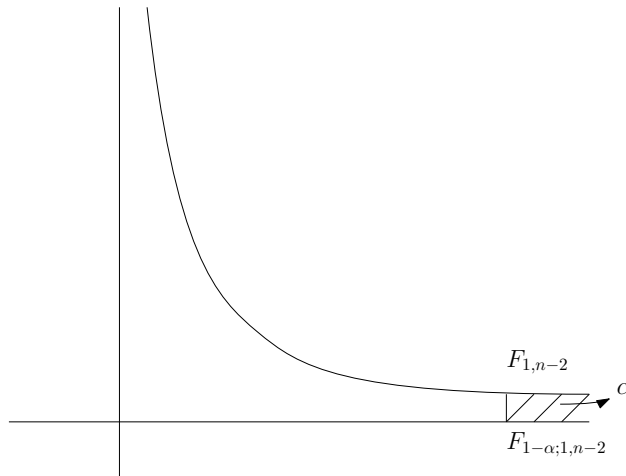
We can observe that $t_{n-2}^2 = F_{1, n-2}$. In general,

$$Z \sim N(0, 1)$$

$$U \sim \mathcal{X}_n^2$$

Z, U are independent

$$\frac{Z}{\sqrt{U/n}} \sim t_n \implies \frac{Z^2/1}{U/n} \sim F_{1, n}$$



Let's find the expected value of the F statistics.

- Denominator:

$$ES_e^2 = \sigma^2$$

- Numerator:

$$\begin{aligned} E\hat{\beta}_1^2 \sum (x_i - \bar{x})^2 &= \sum (x_i - \bar{x})^2 E\hat{\beta}_1^2 \\ &= \sum (x_i - \bar{x})^2 \left(\text{var}(\hat{\beta}_1) + (E\hat{\beta}_1)^2 \right) \\ &= \sum (x_i - \bar{x})^2 \left(\frac{\sigma^2}{\sum (x_i - \bar{x})^2} + \beta_1^2 \right) \\ &= \sigma^2 + \beta_1^2 \sum (x_i - \bar{x})^2 \end{aligned}$$

Under H_0 the ratio is approximately equal to 1. If H_0 is not true the ratio is greater than 1.

Now, for $\hat{\beta}_0$,

$$\left. \begin{aligned} \hat{\beta}_0 &\sim N\left(0, \sigma \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}\right) \\ \frac{(n-2)S_e^2}{\sigma^2} &\sim \chi_{n-2}^2 \end{aligned} \right\} \Rightarrow t = \frac{\hat{\beta}_0}{S_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}} \sim t_{n-2}$$

and consider $H_0 : \beta_1 = 1$ ($\beta_1 - 1 = 0$) and $H_a : \beta_1 \neq 1$ ($\beta_1 - 1 \neq 0$). Then under H_0 ,

$$\frac{\hat{\beta}_1 - 1}{\sigma / \sqrt{\sum (x_i - \bar{x})^2}} \sim N(0, 1)$$

Test Statistics:

$$\frac{\hat{\beta}_1 - 1}{S_e / \sqrt{\sum (x_i - \bar{x})^2}} \sim t_{n-2}$$

Using F statistics

$$\frac{(\hat{\beta}_1 - 1)^2 \sum (x_i - \bar{x})^2}{\sigma^2} \sim \chi_1^2$$

and thus

$$\frac{(\hat{\beta}_1 - 1)^2 \sum (x_i - \bar{x})^2}{S_e^2} \sim F_{1, n-2}$$

§8 | Lec 8: Oct 13, 2021

§8.1 Likelihood Ratio Test

Consider

$$\begin{aligned} Y_i &= \beta_1 X_i + \varepsilon_i \\ H_0 &: \beta_1 = 0 \\ H_a &: \beta_1 \neq 0 \end{aligned}$$

We know

$$\begin{aligned} \hat{\beta}_1 &\sim N\left(0, \frac{\sigma}{\sqrt{\sum x_i^2}}\right) \\ \frac{(n-1)S_e^2}{\sigma^2} &\sim \chi_{n-1}^2 \end{aligned}$$

So $t_{\text{test}}: \frac{\hat{\beta}_1}{S_e/\sqrt{\sum x_i^2}} \sim t_{n-1}$ and $F_{\text{test}}: \frac{\hat{\beta}_1^2 \sum x_i^2}{S_e^2} \sim F_{1, n-1}$.

Likelihood Ratio Test (LRT):

For testing: $H_0: \beta_1 = 0$

For the model: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

Show that this LRT is equivalent to the F statistic.

We reject H_0 if

$$\Lambda = \frac{L(\hat{w})}{L(\hat{\omega})} < k$$

where $L(\hat{w})$ is the maximized likelihood function under H_0 and $L(\hat{\omega})$ is maximized likelihood function under no restrictions. Under $H_0: \beta_1 = 0$, we have $Y_i = \beta_0 + \varepsilon_i$. The likelihood function is

$$\begin{aligned} L &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum (y_i - \beta_0)^2} \\ \ln L &= -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum (y_i - \beta_0)^2 \\ \hat{\beta}_0 &= \bar{y} \\ \hat{\sigma}_0^2 &= \frac{\sum (y_i - \bar{y})^2}{n} \end{aligned}$$

Under no restriction, the estimates are the MLEs of $\beta_0, \beta_1, \sigma^2$ which are $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\sigma}_1^2 = \frac{\sum \varepsilon_i^2}{n}$. Back to LRT, we have

$$\begin{aligned} \Lambda &= \frac{L(\hat{w})}{L(\hat{\omega})} \\ &= \frac{(2\pi\sigma_0^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_0^2} \sum (y_i - \bar{y})^2}}{(2\pi\sigma_1^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_1^2} \sum \varepsilon_i^2}} < k \end{aligned}$$

Note:

$$\begin{aligned} \sum (y_i - \bar{y})^2 &= n\sigma_0^2 \\ \sum \varepsilon_i^2 &= n\sigma_1^2 \end{aligned}$$

So,

$$\begin{aligned}\frac{(2\pi\hat{\sigma}_0^2)^{-\frac{n}{2}}e^{-\frac{n}{2}}}{(2\pi\hat{\sigma}_1^2)^{-\frac{n}{2}}e^{-\frac{n}{2}}} &< k \\ \frac{\hat{\sigma}_1^2}{\hat{\sigma}_0^2} &< k^{\frac{2}{n}} \\ \frac{\sum e_i^2/n}{\sum (y_i - \bar{y})^2/n} &< k^{\frac{2}{n}}\end{aligned}$$

Notice that

$$\begin{aligned}\sum (y_i - \bar{y})^2 &= \sum e_i^2 + \sum (\hat{y}_i - \bar{y})^2 \\ \sum (y_i - \bar{y})^2 &= \sum e_i^2 + \hat{\beta}_1^2 \sum (x_i - \bar{x})^2\end{aligned}$$

So,

$$\begin{aligned}\frac{\sum e_i^2}{\sum e_i^2 + \hat{\beta}_1^2 \sum (x_i - \bar{x})^2} &< k^{\frac{2}{n}} \\ \frac{1}{1 + \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sum e_i^2}} &< k^{\frac{2}{n}} \\ \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{(n-2)S_e^2} &> k^{-\frac{2}{n}} - 1 \\ \frac{\hat{\beta}_1 \sum (x_i - \bar{x})^2}{S_e^2} &> (n-2) \left(k^{-\frac{2}{n}} - 1 \right) = k'\end{aligned}$$

We reject H_0 if

$$\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2} > k'$$

Recall we stated that we reject H_0 if $\Lambda = \frac{L(\hat{w})}{L(\hat{\omega})} < k$. Let's find k . First, we need α (type I error). Before that, we know

$$\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2} \sim F_{1, n-2}$$

So,

$$P\left(F_{1, n-2} > k' \mid H_0 \text{ is true}\right) = \alpha$$

§8.2 Power Analysis in Simple Regression

Using the non-central t distribution

Definition 8.1 (Non-central t) — Let $Z \sim N(\delta, 1)$ and $U \sim \mathcal{X}_n^2$ and Z and U are independent. Then,

$$\frac{Z}{\sqrt{U/n}} \sim t_n \text{ (NCP} = \delta\text{)}$$

Back to the t ratio. If H_0 is true,

$$\frac{\frac{\hat{\beta}_1}{\sigma/\sqrt{\sum (x_i - \bar{x})^2}}}{\sqrt{\frac{(n-2)S_e^2}{\sigma^2}/(n-2)}}$$

follows central t_{n-2} in which the numerator follows standard normal distribution. If H_0 is not true, then the numerator follows $N\left(\frac{\beta_1\sqrt{\sum(x_i-\bar{x})^2}}{\sigma}, 1\right)$. Thus, the ratio follows t_{n-2} (NCP = $\frac{\beta_1\sqrt{\sum(x_i-\bar{x})^2}}{\sigma}$). Finally, the power is

$$1 - \beta = P(t_{n-2}(\text{NCP}) > t_{\frac{\alpha}{2}; n-2}) + P(t_{n-2}(\text{NCP}) < -t_{\frac{\alpha}{2}; n-2})$$

§9 | Lec 9: Oct 15, 2021

§9.1 Extra Sum of Squares Method

So far, we have learnt several ways for hypothesis testing for, e.g.,

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + \varepsilon_i \\ H_0 &: \beta_1 = 0 \\ H_a &: \beta_1 \neq 0 \end{aligned}$$

which are

1. t statistics
2. F statistics
3. Likelihood ratio test
4. Extra sum of square principle (reduced and full model)

$$\frac{(SSE_R - SSE_F)/(df_R - df_F)}{SSE_F/df_F} \sim F_{1,n-2}$$

$$\left. \begin{aligned} SSE_F &= \sum e_i^2 \\ df_F &= n - 2 \end{aligned} \right\}$$

Under $H_0: \beta_1 = 0$ we have a reduced model

$$Y_i = \beta_0 + \varepsilon_i \implies \hat{\beta}_0 = \bar{y}$$

Therefore $SSE_R = \sum (y_i - \bar{y})^2$ and $df_R = n - 1$. Thus,

$$\frac{(\sum (y_i - \bar{y})^2 - \sum e_i^2) / (n - 1 - (n - 2))}{\sum e_i^2 / (n - 2)}$$

Note:

$$\underbrace{\sum (y_i - \bar{y})^2}_{\text{SST}} = \underbrace{\sum e_i^2}_{\text{SSE}} + \underbrace{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}_{\text{SSR}}$$

So,

$$\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2} \sim F_{1,n-2}$$

$$\left(\frac{\hat{\beta}_1}{S_e / \sqrt{\sum (x_i - \bar{x})^2}} \right) \sim t_{n-2}^2$$

Example 9.1

Use the extra sum of squares method to test

$$H_0 : \beta_1 = 1$$

$$H_a : \beta_1 \neq 1$$

Reduced model: $Y_i = \beta_0 + x_i + \varepsilon_i$

$$Y_i - x_i = \beta_0 + \varepsilon_i$$

$$\hat{\beta}_0 = \bar{y} - \bar{x}$$

$$\begin{aligned} SSE_R &= \sum (y_i - x_i - (\bar{y} - \bar{x}))^2 \\ &= \sum (y_i - \bar{y} - (x_i - \bar{x}))^2 \\ &= \sum (y_i - \bar{y})^2 + \sum (x_i - \bar{x})^2 - 2 \sum (x_i - \bar{x})(y_i - \bar{y}) \end{aligned} \quad (*)$$

Note:

$$\begin{aligned} \sum (y_i - \bar{y})^2 &= \sum e_i^2 + \hat{\beta}_1^2 \sum (x_i - \bar{x})^2 \\ \hat{\beta}_1 &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \\ \Rightarrow \sum (x_i - \bar{x})(y_i - \bar{y}) &= \hat{\beta}_1 \sum (x_i - \bar{x})^2 \end{aligned}$$

So, we have

$$\begin{aligned} (*) &= \sum e_i^2 + \hat{\beta}_1^2 \sum (x_i - \bar{x})^2 + \sum (x_i - \bar{x})^2 - 2\hat{\beta}_1 \sum (x_i - \bar{x})^2 \\ SSE_R &= \sum e_i^2 + (\hat{\beta}_1 - 1)^2 \sum (x_i - \bar{x})^2 \end{aligned}$$

Test statistics:

$$\begin{aligned} &\frac{(SSE_R - SSE_F)/(df_R - df_F)}{SSE_F/df_F} \\ &= \frac{\left(\sum e_i^2 + (\hat{\beta}_1 - 1)^2 \sum (x_i - \bar{x})^2 - \sum e_i^2 \right) / (n - 1 - (n - 2))}{\sum e_i^2 / (n - 2)} \\ &= \frac{(\hat{\beta}_1 - 1)^2 \sum (x_i - \bar{x})^2}{S_e^2} \sim F_{1, n-2} \end{aligned}$$

Proof. Under H_0 ,

$$\begin{cases} \hat{\beta}_1 \sim N\left(1, \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}\right) \\ \frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2 \end{cases}$$

So,

$$\begin{aligned} &\frac{\left[\frac{(\hat{\beta}_1 - 1)}{\sigma / \sqrt{\sum (x_i - \bar{x})^2}} \right]^2 / 1}{\frac{(n-2)S_e^2}{\sigma^2} / (n - 2)} \\ &= \frac{(\hat{\beta}_1 - 1)^2 \sum (x_i - \bar{x})^2}{S_e^2} \sim F_{1, n-2} \end{aligned} \quad \square$$

§9.2 Power Analysis Using Non-Central F Distribution

Definition 9.2 — 1. $Y \sim N(\mu, 1)$ then $Y^2 \sim \chi_1^2$ ($\theta = \mu^2$)

2. Suppose $Y \sim N(\mu, \sigma)$

$$\frac{Y}{\sigma} \sim N\left(\frac{\mu}{\sigma}, 1\right)$$

$$\frac{Y^2}{\sigma^2} \sim \chi_1^2 \left(\theta = \frac{\mu^2}{\sigma^2}\right)$$

MGF of $Y \sim \chi_1^2$ (NCP = θ). Then

$$M_Y(t) = (1 - 2t)^{-\frac{1}{2}} e^{\theta \frac{t}{1-2t}}$$

If $\theta = 0 \implies M_Y(t) = (1 - 2t)^{-\frac{1}{2}}$.

Consider now

$$Y_1, Y_2, \dots, Y_n \stackrel{\text{i.i.d.}}{\sim} N(\mu, \sigma)$$

Find distribution of $Q = \frac{Y_1^2}{\sigma^2} + \dots + \frac{Y_n^2}{\sigma^2}$.

$$M_Q(t) = \left[(1 - 2t)^{-\frac{1}{2}} e^{\frac{\mu^2}{\sigma^2} \frac{t}{1-2t}} \right]^n$$

$$= (1 - 2t)^{-\frac{n}{2}} e^{\frac{n\mu^2}{\sigma^2} \frac{t}{1-2t}}$$

$$Q = \frac{\sum Y_i^2}{\sigma^2} \sim \chi_n^2 \left(\theta = \frac{n\mu^2}{\sigma^2} \right)$$

Non-Central F Distribution: Let $U \sim \chi_n^2$ (NCP = θ) and $V \sim \chi_m^2$. If U, V are independent, then

$$\frac{U/n}{V/m} \sim F_{n,m} \text{ (NCP = } \theta \text{)}$$

Back to simple regression:

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}\right)$$

$$\frac{\hat{\beta}_1}{\sigma / \sqrt{\sum (x_i - \bar{x})^2}} \sim N\left(\frac{\beta_1}{\sigma / \sqrt{\sum (x_i - \bar{x})^2}}, 1\right)$$

$$\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sigma^2} \sim \chi_1^2 \left(\theta = \frac{\beta_1^2 \sum (x_i - \bar{x})^2}{\sigma^2} \right)$$

$$\frac{(n-2)S_e^2}{\sigma^2} \sim \chi_{n-2}^2$$

$$\frac{\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sigma^2} / 1}{\frac{(n-2)S_e^2}{\sigma^2} / (n-2)} \sim F_{1,n-2} \left(\theta = \frac{\beta_1^2 \sum (x_i - \bar{x})^2}{\sigma^2} \right)$$

Thus,

$$1 - \beta = P(F_{1,n-2}(\theta) > F_{1-\alpha;1,n-2})$$