Recap

$$\hat{\beta}_1 \sim N\left(\beta_1, \sigma^2/SSX\right)$$

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

$$e_i \sim N\left(0, \sigma^2\left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{SSX}\right)\right)$$

$$\hat{y}_i \sim N\left(\beta_0 + \beta_1 x_i, \underline{\quad ? \quad}\right)$$

Exercise:
calc. $var(\hat{y}_i)$

Properties:

1. $\sum_i e_i = 0$ $\qquad \sum_i (y_i - \hat{y}_i) = 0$

2. $\sum_i y_i = \sum_i \hat{y}_i$

3. $Q(\hat{\beta}_0, \hat{\beta}_1) = \sum_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = \sum_i e_i^2$
   is the minimum value of $Q(\beta_0, \beta_1)$.

4. $\sum_{i=1}^{n} e_i x_i = 0$

5. $\sum_{i=1}^{n} e_i \hat{y}_i = 0$

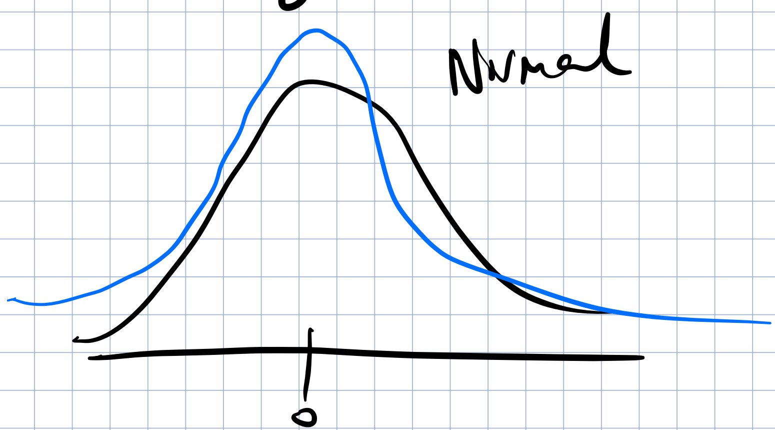
$$Y_i \sim N(\underline{\underline{\beta_0}} + \underline{\underline{\beta_1}} x_i, \sigma^2)$$

what if we want an estimator of this?

Remark:

$$\frac{\sum_i e_i^2}{\sigma^2} \sim \chi^2_{n-p}$$

p = # of betas in the model



Def: If I had $Z_1, \dots, Z_n \overset{iid}{\sim} N(0,1)$, then $V = Z_1^2 + Z_2^2 + \dots + Z_n^2 \sim \chi^2_n$

For SLR,

Notation:
$SSE = \sum_i e_i^2$

$$\boxed{\frac{\sum_i e_i^2}{\sigma^2} \sim \chi^2_{n-2}}$$

Pf later

I want to find an estimator of σ^2

I want it to be unbiased.

[If $V \sim \chi^2_{df}$, $E(V) = df$.]

$$E\left(\frac{\sum_i e_i^2}{\sigma^2}\right) = n-2$$

!!

$$\Rightarrow \frac{1}{\sigma^2} E\left(\sum_i e_i^2\right) = n-2$$

$$\Rightarrow \frac{1}{n-2} E\left(\sum_i e_i^2\right) = \sigma^2$$

$$= E\left(\frac{\sum_i e_i^2}{n-2}\right) = \sigma^2$$

take this as my estimator

Let $\hat{\sigma}^2 = \sum_i e_i^2 / n-2 \Rightarrow$

$$E(\hat{\sigma}^2) = E\left(\frac{\sum_i e_i^2}{n-2}\right) = \sigma^2,$$

therefore $\hat{\sigma}^2$ is unbiased!

The estimator $\hat{\sigma}^2$ is sometimes called the mean squared error (MSE).

Hypothesis Testing & Confidence Intervals for SLR

Recall:

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2/SSX)$$

Consider:

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

$H_1: \beta_1 \neq 0$

$\alpha = 0.05$

Construct the z-stat:

$$Z = \frac{\hat{\beta}_1 - 0}{SD(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - 0}{\sigma / \sqrt{SSX}}$$

If $|Z| > Z^*_{1-\frac{\alpha}{2}} \Rightarrow$ reject H_0

The problem is that we don't know σ!! We can't do this.

We need to rely on our estimate of σ — the MSE.

If we replace σ by its estimate $\hat{\sigma}$, then the z-test becomes a t-test!

$\Rightarrow$

$$t = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}/\sqrt{SSX}}$$ is your new statistic.

This statistic follows a t_{n-2} dist under the null hypothesis.

<u>Decision Rule:</u>

If $|t| > t^*_{1-\frac{\alpha}{2}, n-2}$ then

<u>reject H_0</u>

We have found significant evidence to show that the true slope is not zero. Another way of saying this is that "X is a <u>significant</u> predictor of Y_i"

$$y_i = \beta_0 + \underset{\substack{\downarrow \\ 0 \text{ under } H_0}}{\beta_1} x_i + \varepsilon$$

$$= \beta_0 + \varepsilon_i$$

Similar tests exist for the intercept:

$H_0: \beta_0 = 0$
$H_1: \beta_0 \neq 0$

$$\Rightarrow t = \frac{\hat{\beta}_0 - 0}{\widehat{se}(\hat{\beta}_0)}$$

$$\Rightarrow t = \frac{\hat{\beta}_0 - 0}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{SSX}}}$$

Same Decision Rule:

If $|t| > t^{*}_{1-\frac{\alpha}{2}, n-2}$ then reject H_0.

Interpretation

Reject H_0:

We find evidence to believe that the mean response of an observation with $X=0$ is not zero.

$$SE(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$$

$\widehat{SE}(\hat{\theta}) \leftarrow$ plug in $\hat{\sigma}$ for σ

dist of $\hat{\beta}_1$

For ex:

$$SE(\hat{\beta}_1) = \sqrt{Var(\hat{\beta}_1)} = \sigma / \sqrt{SSX}$$

$$\widehat{SE}(\hat{\beta}_1) = \hat{\sigma} / \sqrt{SSX}.$$

"standard error" just means the standard deviation of some statistic.

Confidence Intervals

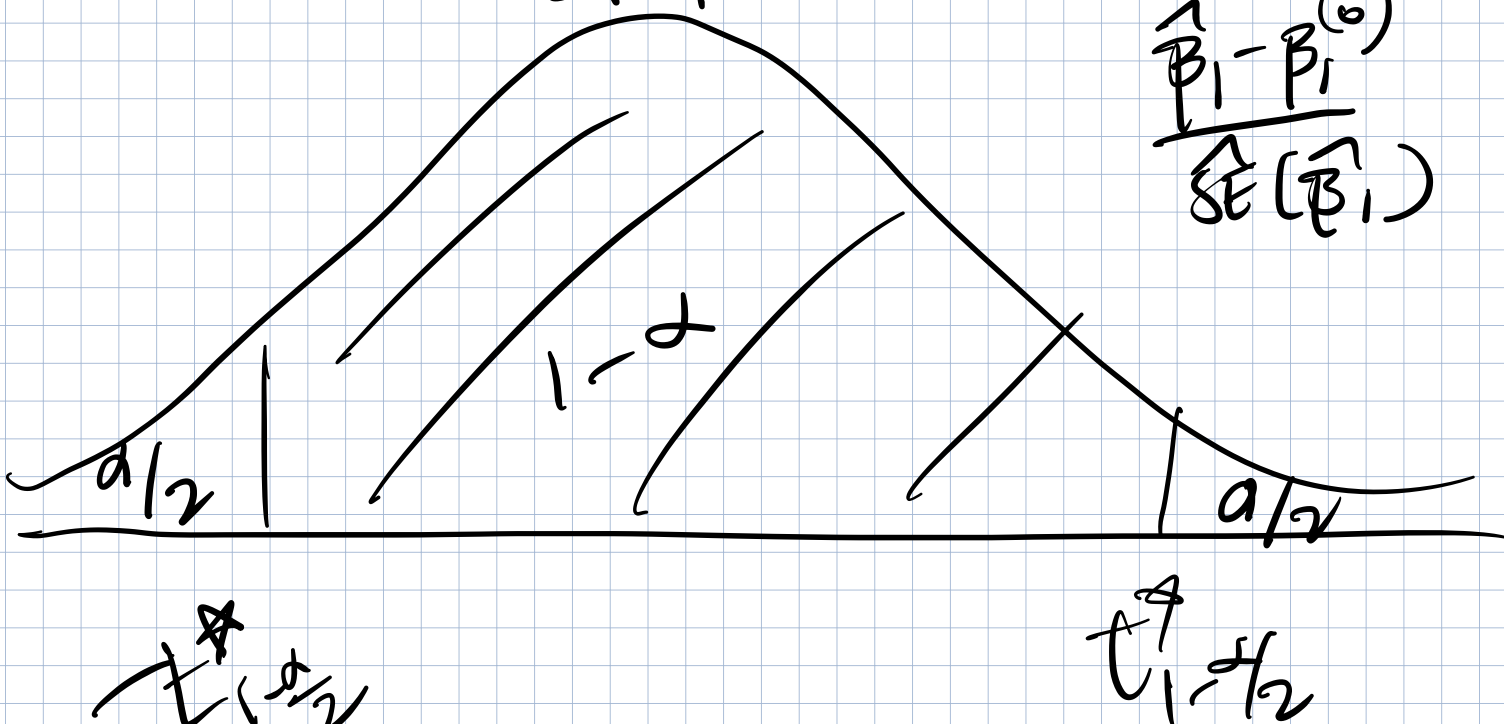
For the slope if we know

that $t = \dfrac{\hat{\beta}_1 - \beta_1}{\widehat{SE}(\hat{\beta}_1)} \sim t_{n-2}$ dist, ← null hypothesis value

then

$$P\left(-t^*_{1-\frac{\alpha}{2}, n-2} \le \frac{\hat{\beta}_1 - \beta_1}{\widehat{SE}(\hat{\beta}_1)} \le t^*_{1-\alpha/2, n-2}\right) = 1-\alpha$$

$$\frac{\hat{\beta}_1 - \beta_1^{(0)}}{\widehat{SE}(\hat{\beta}_1)}$$



— Solve for $\beta_1^{(0)}$

$\Rightarrow$

$$P\left(-t^*_{1-\frac{\alpha}{2},n-2}\,\widehat{SE}(\hat\beta_1) \le \hat\beta_1 - \beta_1^{(0)} \le t^*_{1-\frac{\alpha}{2},n-2}\,\widehat{SE}(\hat\beta_1)\right) = 1-\alpha$$

$$P\left(\hat\beta_1 + t^*_{1-\frac{\alpha}{2},n-2}\,\widehat{SE}(\hat\beta_1) \ge \beta_1^{(0)} \ge \hat\beta_1 - t^*_{1-\frac{\alpha}{2},n-2}\,\widehat{SE}(\hat\beta_1)\right) = 1-\alpha$$

$$\boxed{\hat\beta_1 \pm t^*_{1-\frac{\alpha}{2},n-2}\,\widehat{SE}(\hat\beta_1)}$$

$(1-\alpha)100\%$ CI for β_1

A similar construction gives us the CI for β_0:

$$\boxed{\hat{\beta}_0 \pm t^*_{1-\frac{\alpha}{2}, n-2} \, \widehat{SE}(\hat{\beta}_0)}$$

$(1-\alpha)100\%$ CI for β_0