- When there are more than a few data points, these calculations are tedious, so we use software (see R example on course web page).

11.4 Properties of the Least-Squares Estimators in SLR

- Our SLR model is $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, where $E(\varepsilon_i) = 0$.
- We now further assume $V(\varepsilon_i) = \sigma^2$ (the variance is <u>constant</u>, not depending on x).
- Thus $V(Y_i) =$ since $\beta_0 + \beta_1 x_i$ is ____________.

<u>Lemma</u>: $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i - \bar{x}) x_i$.

<u>Proof</u>:

<u>Theorem:</u> In the SLR model, $\hat{\beta}_0$ and $\hat{\beta}_1$ are <u>unbiased</u> estimators of β_0 and β_1.

<u>Proof:</u>

Theorem: $V(\hat{\beta}_0) = c_{00}\sigma^2$, where $c_{00} = \dfrac{\Sigma x_i^2}{n S_{xx}}$ and $V(\hat{\beta}_1) = c_{11}\sigma^2$, where $c_{11} = \dfrac{1}{S_{xx}}$.

Proof:

<u>Corollary</u>: $\text{cov}(\hat{\beta}_0, \hat{\beta}_1) = c_{01}\sigma^2$, where $c_{01} = \frac{-\overline{x}}{S_{xx}}$.

<u>Proof</u>:

- Note that $V(\hat{\beta}_0)$ and $V(\hat{\beta}_1)$ each involve $\sigma^2 = V(\varepsilon_i)$, which is usually unknown and must be estimated.
- Recall that in the single-sample situation, we used to estimate $E(Y_i)$ and used

to estimate $\sigma^2 = V(Y_i)$.

- In the regression setting, we use to estimate $E(Y_i)$ and so we use

to estimate σ^2.

- Why $(n-2)$ rather than $(n-1)$ in the denominator?
- There are 2 unknown quantities (β_0 and β_1) to be estimated when estimating $E(Y_i)$.

Theorem: MSE is an unbiased estimator of σ^2.

Proof:

Note: A convenient formula for hand calculation of SSE is

Example 1 again: Find $V(\hat{\beta}_0)$ and $V(\hat{\beta}_1)$, and estimate these by using MSE to estimate σ^2.

- Estimating $V(\hat{\beta}_0)$ and $V(\hat{\beta}_1)$ will be useful when we perform inference about β_0 and β_1.

The Normality Assumption

- To this point, we have assumed that $\varepsilon_1, \ldots, \varepsilon_n$ are independent with $E(\varepsilon_i) = 0$ and $V(\varepsilon_i) = \sigma^2$.
- We may also assume that the ε_i's are normally distributed (around the true regression line):

- In this case, since $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, then:

- Recall that $\hat{\beta}_0$ and $\hat{\beta}_1$ can each be written as a linear combination of the Y_i's, so they would each have a ____________ distribution.

Furthermore,

has a ________________________, similar to the result about in STAT 512.

- These distributional facts will allow us to make CIs and hypothesis tests in the regression setting.

11.5-11.6 Inferences About the β_i's

- In the SLR setting, we consider inference about

where a_0 and a_1 are constants.

Examples: If $a_0=1$, $a_1=0$, we have inference about

- If $a_0=0$, $a_1=1$, we have inference about
- If $a_0=1$, $a_1=x^*$, we have inference about

- Consider the estimator

- This is unbiased since

- Also,

- Plugging in results from Sec. 11.4,

- And if the Y_i's are normally distributed, then $\hat{\theta}$ is ____________________.
- Thus $Z = \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}}$ has a ____________________ distribution and could be used for inference about θ.
- But $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$ depends on σ^2 (typically unknown).

- If we use MSE in place of σ^2, then

has a ________________ with __________.

- So we can conduct __________ and CIs about any θ of interest.

General Hypothesis Test Procedure

H_0	H_a	Test Stat	RR

where

100(1−α)% CI for $\theta = a_0\beta_0 + a_1\beta_1$

Example: To test whether $E(Y)$ depends on x, we test

Perform this test with the Example 1 data, at $\alpha = .05$.

A 95% CI for

Interpretation:

Exercise: Let $a_0 = 1$, $a_1 = 0$ to get a 95% CI for β_0 (this is not very sensible in this example, however).

CI for an Expected Response

- Recall Example 1 data: When $x = 10.5$, we found a point estimate of $E(Y)$ to be:

- We can get a CI for $E(Y)$ by letting $a_0 =$, $a_1 =$

Example: 90% CI for expected race time for skiers with 10½-min time to exhaustion:

- With 90% confidence,

- Note: These CIs will be more precise when S_{xx} is ________, i.e., when the sample x-values are ______________.
- See course web page for examples in R.