Assumptions: We have a random variable Y (the response variable) and fixed values x₁, x₂, … , xₙ of an explanatory variable X. We will assume that the random variable Y satisfies the following conditions (which are just rephrasings of the assumptions on pp. 675 – 676 of the textbook):

- **Linearity assumption:** There are constants β₀ and β₁ such that for each value x of X within some range of interest, E(Y|x) = β₀ + β₁x (The textbook uses μ_y instead of E(Y|x)).
- **Independence assumption:** The conditional distributions Y|x₁, Y|x₂, … , Y|xₙ are independent. (This will imply that the error random variables Y|x₁ − (β₀ + β₁x₁), Y|x₂ − (β₀ + β₁x₂) … , Y|xₙ − (β₀ + β₁xₙ) are independent. The book refers to these collectively as ε.)
- **Equal variance assumption:** All error variables Y|x − (β₀ + β₁x) (for x within the range of interest) have the same variance, which we will call σ².
- **Normality assumption:** Each conditional distribution Y|x₁, Y|x₂, … , Y|xₙ is normal. (This will imply that each error variable Y|xᵢ − (β₀ + β₁xᵢ) is normal.)

I. The formula for sₑ (p. 681): The reason this formula has n-2 in the denominator is similar to the reason that the formula for the ordinary sample standard deviation s has n-1 in the denominator: so its square will give an unbiased estimator of the population variance. (See the Chapter 18 handout “Why Does the Sample Variance Have n – 1 in the Denominator” for some details on that.)

Outline:
- The formula for sₑ² allows us to define a random variable Sₑ² in the regression context as follows:
  - The random process for Sₑ² is, “Randomly choose a sample y₁, y₂, … , yₙ in such a way that each yᵢ is a random observation from Y|xᵢ”.
  - The value of Sₑ² corresponding to this sample is sₑ² calculated using this sample.
  - Sₑ² is thus an estimator of σ².
- It can be proved (the proof is beyond the scope of this course) that E(Sₑ²) = σ², so Sₑ² is an unbiased estimator of σ².
- Note that this implies that if we used the formula with n-1, rather than n-2, in the denominator, we would get an estimator with expected value \( \frac{n-2}{n-1} \) σ², which means we would be consistently underestimating σ². This wouldn’t be too bad for large enough n, but could be a problem for small n. However, the estimator Sₑ² is also needed for deriving some of the other formulas and properties.

II. The formula for SE(b₁) (p. 682): (For more details, see notes Statistical Properties of Least Squares Estimators from a course in regression, available at http://www.ma.utexas.edu/users/mks/384Gfa08/384G08home.html)
• It is possible to prove that the least squares estimator obtained by using the formula for $b_1$ is an unbiased estimator of $\beta_1$. By abuse of notation: $E(b_1) = \beta_1$.

One proof depends on using the least squares equations to write $b_1$ as a certain linear combination of the sampled values $y_1, y_2, \ldots, y_n$. (This proof uses just the linearity assumption and the properties of expected values.)

• Applying the properties of variances to the same linear combination expression, and using the independence and constant variance (as well as linearity) assumptions leads to the formula $\text{Var}(b_1) = \frac{\sigma^2}{SXX}$, where $SXX = \sum (x_i - \overline{x})^2$

• Approximating $\sigma$ by $s_e$, noting that $SXX = (n-1)s_e^2$, and taking square roots then gives the formula for $\text{SE}(b_1)$ on p. 682

III. Why $\frac{b_1 - \beta_1}{\text{SE}(b_1)}$ has the t-distribution with n-2 degrees of freedom (p. 682): (For more details, see notes Inference for Simple Linear Regression from a course in regression, available at http://www.ma.utexas.edu/users/mks/384Gfa08/384G08home.html)

Recall from the handout Chi-Squared Distributions, t-Distributions, and Degrees of Freedom (Supplement to Chapter 23):

Definition: The $t$ distribution with k degrees of freedom is the distribution of a random variable which is of the form $\frac{Z}{\sqrt{U/k}}$ where

i. $Z \sim N(0,1)$
ii. $U \sim \chi^2(k)$, and
iii. $Z$ and $U$ are independent.

In that handout, this definition was used to show why (under the conditions for a one-sample t-test for a mean) $\frac{\overline{Y} - \mu}{s/\sqrt{n}}$ has a t-distribution. The reasoning showing that $\frac{b_1 - \beta_1}{\text{SE}(b_1)}$ has a t-distribution is similar. Here’s an outline:

• The fact (mentioned above) that $b_1$ is a certain linear combination of the sampled values $y_1, y_2, \ldots, y_n$ can be reframed to say that the estimator defined by $b_1$ is a linear combination of the random variables $Y|x_1, Y|x_2, \ldots, Y|x_n$.

• This plus the independence and normality assumptions implies that the estimator defined by $b_1$ (which by abuse of notation we will also call $b_1$) is normal.

• Since $E(b_1) = \beta_1$, standardizing $b_1$ says that $\frac{b_1 - \beta_1}{\text{SD}(b_1)} \sim N(0,1)$ (i.e., is standard normal. This will turn out to be the $Z$ in the definition of t-distribution.)

• From (II) above, $\text{SD}(b_1) = \sqrt{\frac{\sigma^2}{SXX}}$.

• Use algebra to re-express $\text{SE}(b_1)$ as follows:
\[ SE(b_1) = \sqrt{\frac{s_e^2}{SXX}} = \frac{\sigma^2}{\frac{\sigma^2}{s_e^2}} = \frac{SD(b_1)}{\sqrt{\frac{\sigma^2}{s_e^2}}} \]

- Now use this to re-express \( \frac{b_1 - \beta_1}{SE(b_1)} \) as

\[
\frac{b_1 - \beta_1}{SE(b_1)} = \frac{b_1 - \beta_1}{SD(b_1)} \sqrt{\frac{\sigma^2}{S_e^2}} = \frac{b_1 - \beta_1}{SD(b_1)} \sqrt{\frac{S_e^2}{\sigma^2}} \quad (*)
\]

- As remarked above, the numerator of the last expression in equation (*) is standard normal.
- There is a theorem (beyond the scope of this course) that says that (under the assumptions)
  a. \((n-2) \frac{s_e^2}{\sigma^2}\) has a \(\chi^2\) distribution with \(n-2\) degrees of freedom

  Notation: \((n-2) \frac{s_e^2}{\sigma^2} \sim \chi^2(n-2)\)

  b. \((n-2) \frac{s_e^2}{\sigma^2}\) is independent of \(b_1 - \beta_1\) (hence independent of the numerator in (*) )

- Putting this all together, we now see that (*) shows that \( \frac{b_1 - \beta_1}{SE(b_1)} = \frac{Z}{U/k} \), where
  - \(Z = \frac{b_1 - \beta_1}{SD(b_1)}\) is standard normal
  - \(k = n-2\)
  - \(U = (n-2) \frac{s_e^2}{\sigma^2}\) is \(\chi^2\) distribution with \(k\) degrees of freedom, and
  - \(U\) and \(Z\) are independent.
- This says that \( \frac{b_1 - \beta_1}{SE(b_1)} \) indeed has a t-distribution with \(n-2\) degrees of freedom.