

## AP STATISTICS — UNIT 9 QUICK NOTES

**Model:**  $Y = \alpha + \beta x + \varepsilon$ ,  $\varepsilon \sim N(0, \sigma)$

**Sample fit:**  $\hat{y} = a + bx$ ,  $b = r \frac{s_y}{s_x}$ ,  $a = \bar{y} - b\bar{x}$

Residual SD:  $s = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}}$ ,  $SE_b = \frac{s}{\sqrt{\sum(x_i - \bar{x})^2}}$ ,  $df = n - 2$

---

### Conditions (LINER):

1. **L — Linearity:** Scatterplot/residual plot show no curvature.
  2. **I — Independence:** Observations independent; if sampling w/o replacement,  $n \leq 0.1N$ .
  3. **N — Normality:** Residuals roughly normal (histogram/NPP).
  4. **E — Equal variance:** Residual spread constant across  $x$ .
  5. **R — Randomness:** Data from random sample or random assignment.
- 

### SPDC Framework for Regression Inference:

**State** Identify the population slope  $\beta$  and state hypotheses or confidence level in context. Clearly define variables.

**Plan** Name the procedure ( $t$ -test or  $t$ -interval for slope) and check LINER conditions.

**Do** Perform the calculations — find  $b$ ,  $SE_b$ , test statistic  $t = \frac{b - \beta_0}{SE_b}$ , and  $p$ -value; or compute  $b \pm t^* SE_b$  for a CI.

**Conclude** Interpret the results in context, linking back to the slope and the original research question.

---

### Confidence Interval for $\beta$ :

$$\text{CI: } b \pm t_{n-2}^* \cdot SE_b$$

Interpretation: “We are  $C\%$  confident that the slope for the population is between . . .”

### Hypothesis Test for $\beta$ :

$$H_0 : \beta = 0 \quad H_a : \beta > 0, \beta < 0, \text{ or } \beta \neq 0$$

$$t = \frac{b - 0}{SE_b}, \quad p = 2 \cdot P(t_{n-2} > |t|)$$

Interpretation: Probability of getting a slope as extreme as  $b$  if  $H_0$  were true.

---

### TI-84 Instructions:

1. Enter  $x$  in L1,  $y$  in L2.
  2. Turn on diagnostics:  $2nd \rightarrow 0$  (CATALOG)  $\rightarrow$  DiagnosticOn.
  3. Run LinRegTTest:  $\text{STAT} \rightarrow \text{TESTS} \rightarrow \text{F:LinRegTTest}$ .
  4. Select  $L_1, L_2$ , alternative hypothesis ( $\neq, >, <$ ), and Calculate.
  5. Output:  $a$  (intercept),  $b$  (slope),  $s$ ,  $SE_b$ ,  $t$ ,  $p$ ,  $r$ ,  $r^2$ .
- 

### Quick Tips:

- Always check LINER before  $t$ -procedures.
- CI excluding 0  $\implies$  reject  $H_0$  at matching  $\alpha$ .
- Include context, variables, & units in interpretations.
- $R^2$ : percent of variation in  $y$  explained by  $x$ .

**Future Topics:** Multiple Linear Regression, Logistic Regression, GLMs, Machine Learning.

---