Chapter 6

Hypothesis Testing and Confidence Intervals

Learning Objectives

- Test a hypothesis about a regression coefficient
- Form a confidence interval around a regression coefficient
- Show how the central limit theorem allows econometricians to ignore assumption CR4 in large samples
- Present results from a regression model

Hypotheses About β_1

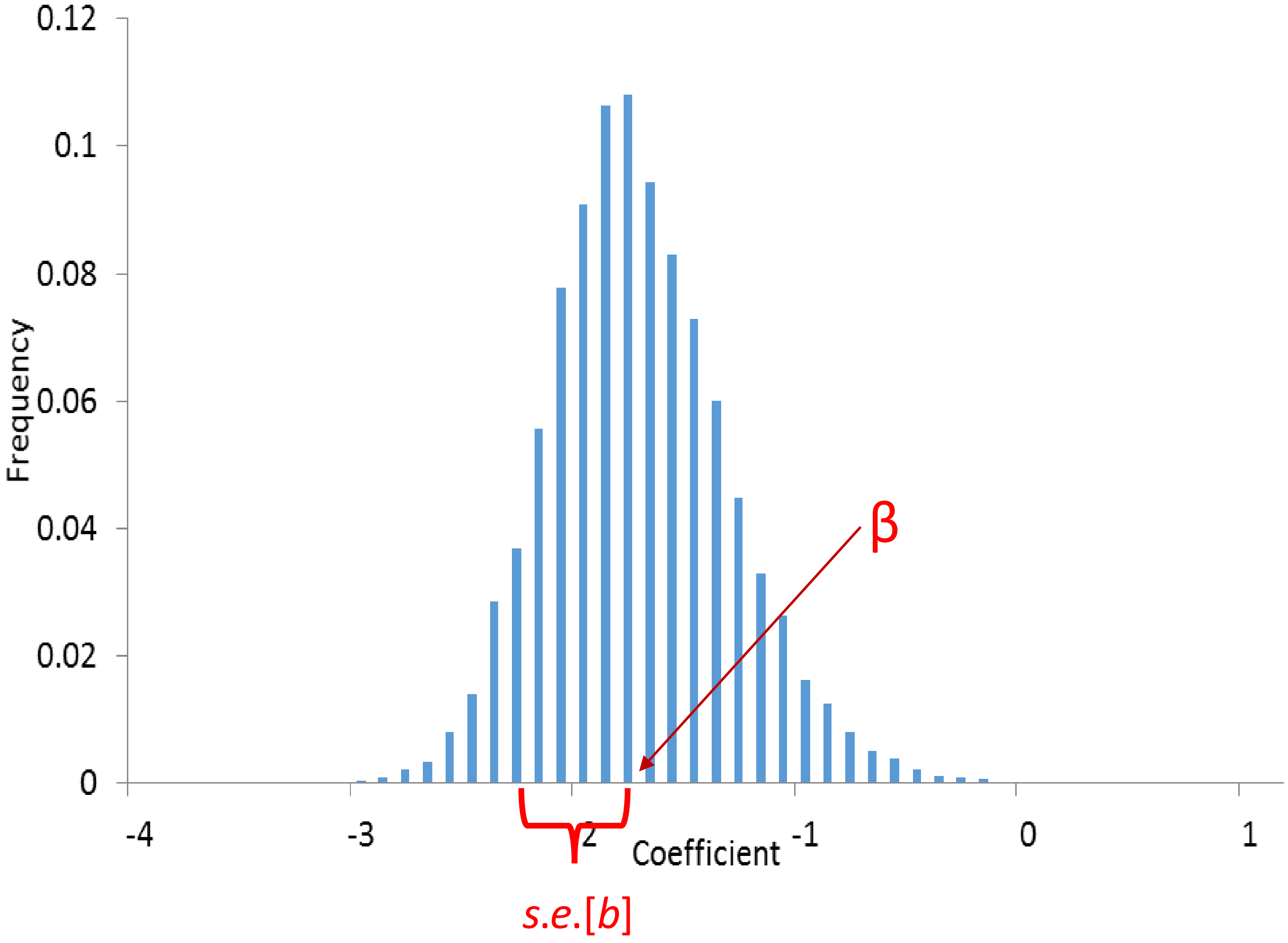
- We propose a value of β_1 and test whether that value is plausible based on the data we have
- Call the hypothesized value β_1^*
- Formal statement:

  Null hypothesis: H_0: $\beta_1 = \beta_1^*$

  Alternative hypothesis: H_1: $\beta_1 \neq \beta_1^*$

- Sometimes the alternative is one sided, e.g., H_1: $\beta_1 < \beta_1^*$
  - Use one sided alternative if only one side is plausible

Properties of *b*



The z-statistic

$$z = \frac{b_1 - \beta_1^*}{s.e.[b_1]}$$

For any hypothesis test:

(i) Take the difference between our estimate and the value it would assume under the null hypothesis, then

(ii) Standardize it by dividing by the standard error of the parameter

If z is a large positive or negative number, then we reject the null hypothesis. We conclude that the estimate is too far from the hypothesized value to have come from the same distribution.

If z is close to zero, then we cannot reject the null hypothesis. We conclude that it is a plausible value of the parameter.

How large should the *t* statistic be before we reject the null hypothesis?

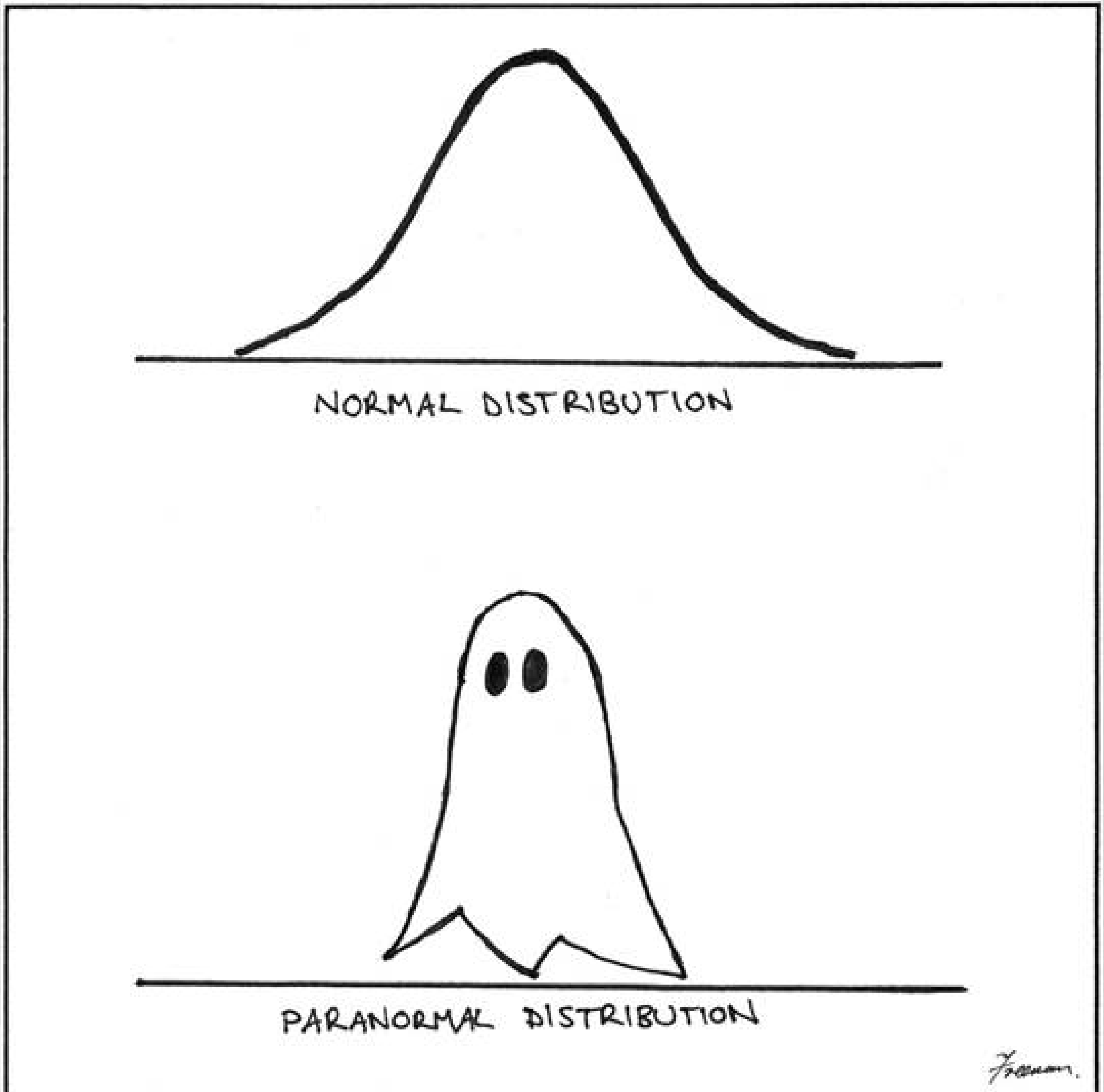
If CR1-CR3 hold (plus CR4, if the sample is small)

$$z = \frac{b_1 - \beta_1}{s.e.[b_1]} \sim N(0,1)$$

Compare this to the *t*-statistic formula:

$$t = \frac{b_1 - \beta_1^*}{est\ s.e.[\mathrm{b}_1]}$$

- If the null hypothesis is true, the numerators of the two expressions are the same
  - If we knew the standard error, we would know the distribution of our *t*-statistic.
- Knowing the distribution means that we know which values are likely and which are unlikely.
- In particular, from the normal table, we would know that getting a *t*-statistic larger than 1.96 only happens with probability 2.5%.
- Getting a *t*-statistic less than -1.96 also only happens with probability 2.5%.


NORMAL DISTRIBUTION
PARANORMAL DISTRIBUTION

Example: California schools data

$$API_i = 951.87 - 2.11 FLE_i + e_i$$

Hypothesis: free-lunch eligibility (FLE) is uncorrelated with academic performance (API)

$$H_0 : \beta_1 = 0 \quad \text{vs} \quad H_1 : \beta_1 \neq 0$$

Put this into practice with our California schools data

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

(free-lunch eligibility, FLE, is uncorrelated with academic performance, API)

The ideal *t*-statistic would use the standard error from the population, i.e.,

Estimate from sample of 20 schools

$$z = \frac{-2.11 - 0}{0.41} = -5.15$$

Sd from whole population (usually don't know this)

The absolute value of this test statistic 5.15 > 1.96, so we reject the null hypothesis at 5% significance.

But We Don't Usually Know the Population Variance

$$t = \frac{-2.11 - 0}{0.36} = -5.86$$

Std error estimated from sample of 20 schools

- Exceeds the critical value (|-5.86| > 2.10), so we still reject the null hypothesis.
- But remember that we had to assume CR4 in order to perform this hypothesis test and construct the confidence interval.
- If there's any reason to think that the population errors are not normally distributed (picture a nice bell curve), this analysis will be very hard to defend to your colleagues.

How Big a Sample is Big Enough?

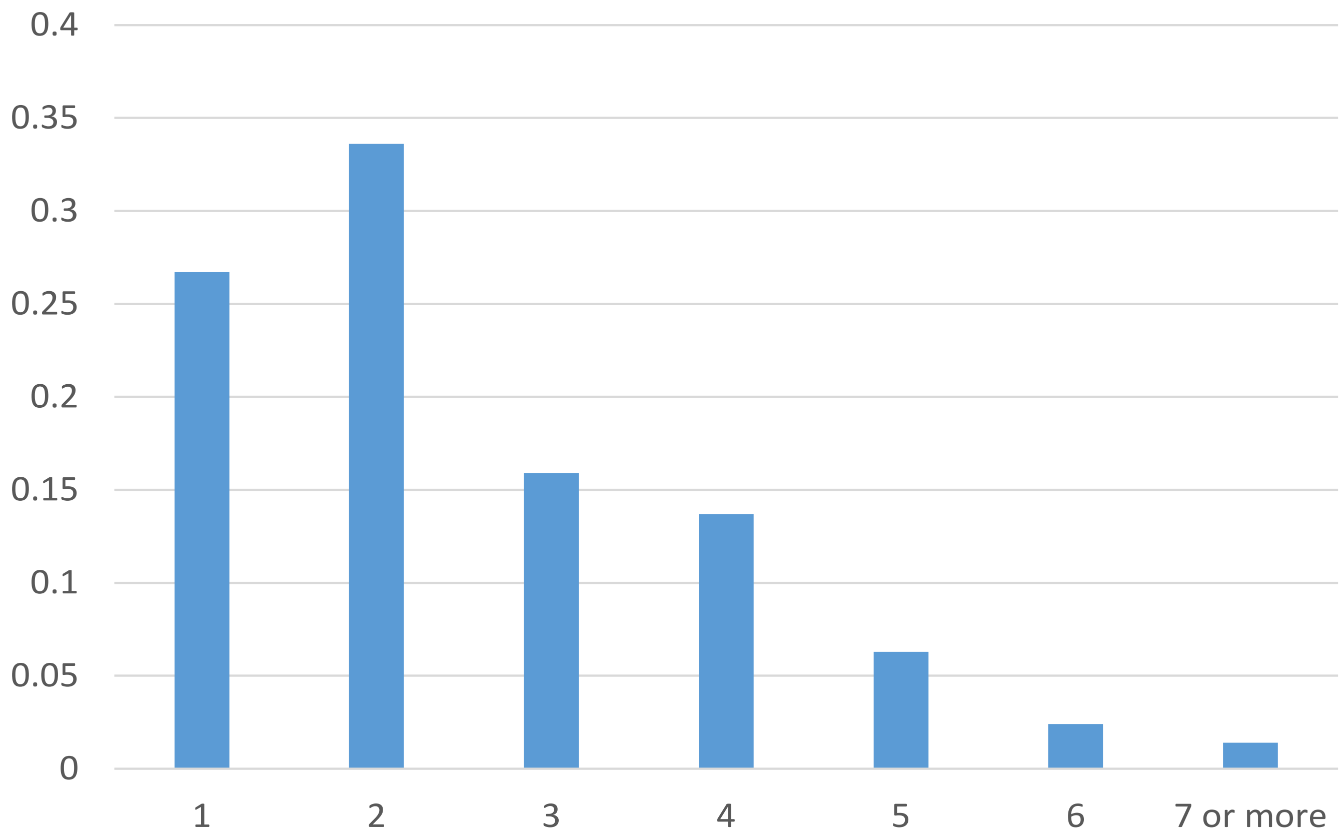


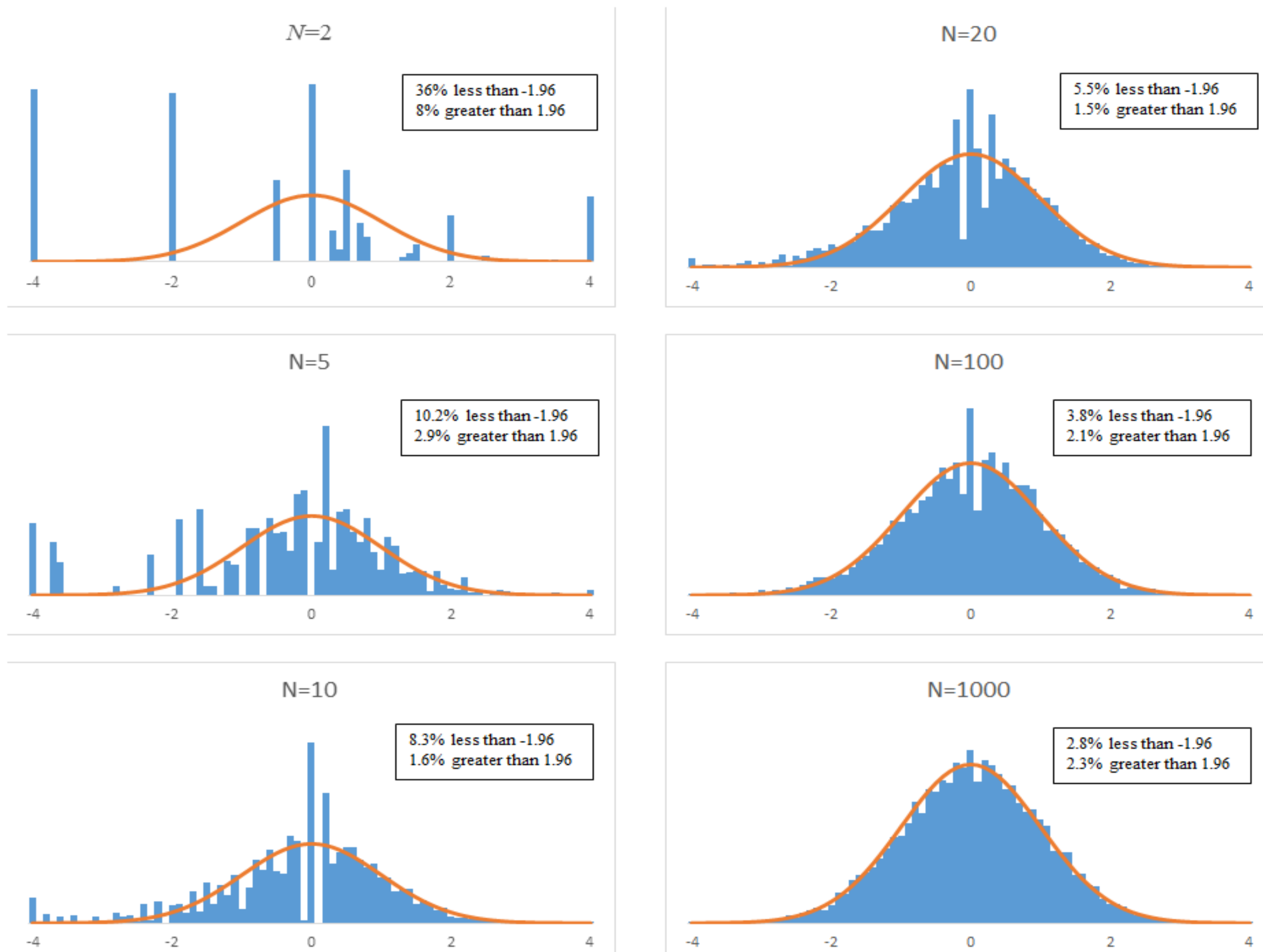
Figure 6.1. Number of People per Household; US Census 2010
Source: https://www.census.gov/hhes/families/data/households.html

- This distribution is certainly not normal

Distribution of the t-statistic Looks More Normal the Bigger the Sample

Average household size, 2010 US Census

$$t = \frac{b - 2.52}{s / \sqrt{N}}$$



Note: Blue bars show histogram of *t* statistics. Orange line is standard normal probability density function.

Figure 6.2. Central Limit Theorem Implies *t* statistic gets more Normal as *N* Increases

CA Schools Errors Look Close to Normal

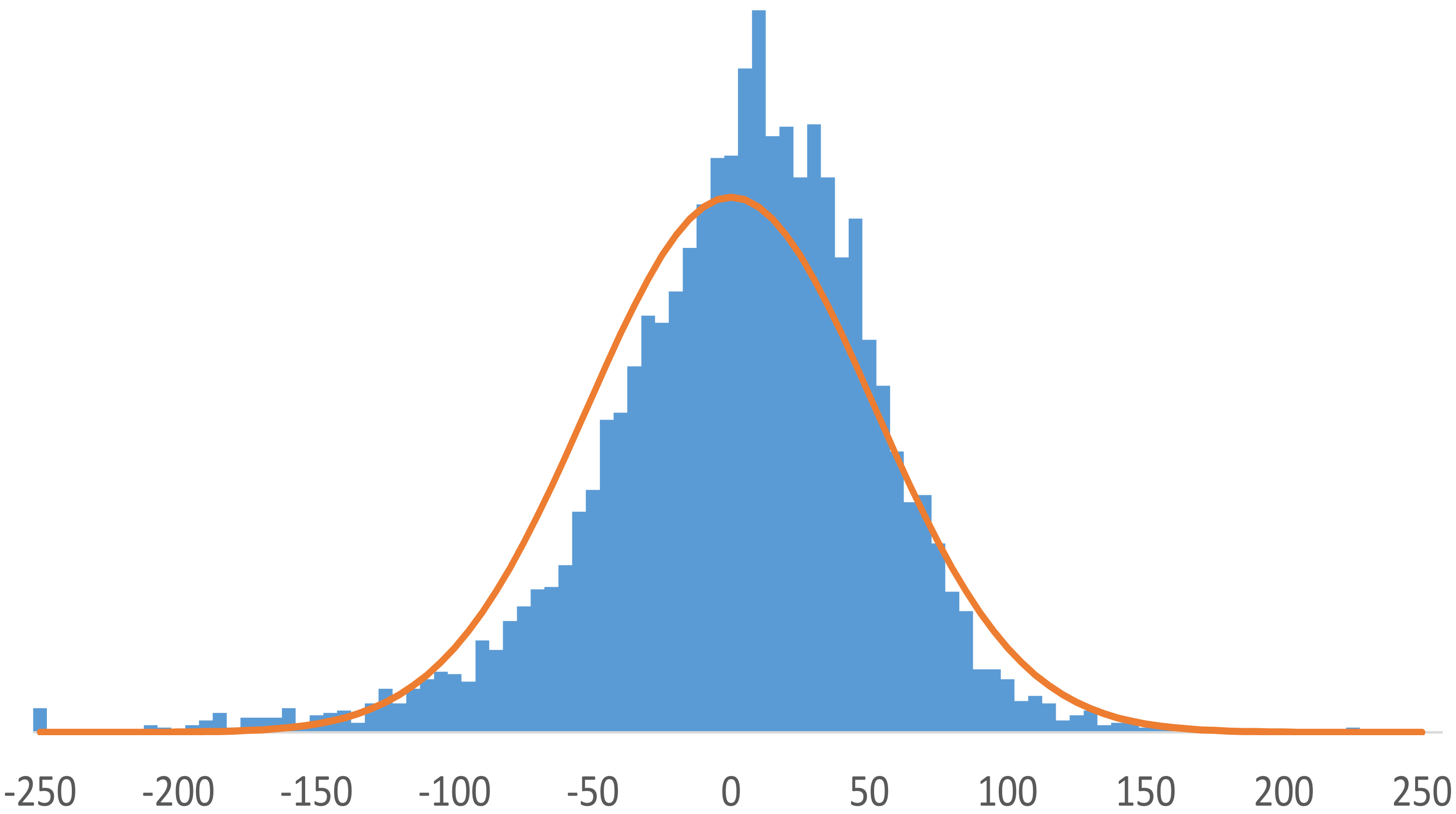


Figure 6.3. Histogram of Errors from Population Regression Using All 5765 CA Schools

Confidence Interval

- Range of values that we can reasonably expect the true population parameter to take on.
- It is the set of null hypotheses that you cannot reject.
- In a large sample, this means that a 95% confidence interval is all the β_1^* values for which

$$-1.96 \leq \frac{b_1 - \beta_1^*}{est\ s.e.[b_1]} \leq 1.96$$

- More common way to express the confidence interval: The 1-α% confidence interval is all the values in the range $\left[\underline{\beta}_1, \overline{\beta}_1\right]$, where

$$\underline{\beta}_1 = b_1 - c * \left(est\ s.e.[b_1]\right) \qquad \overline{\beta}_1 = b_1 + c * \left(est\ s.e.[b_1]\right)$$

where *c* is the critical value for a two-sided test.

Hypothesis Testing in Multiple Regression

$$t = \frac{b_k - \beta_k^*}{est\ s.e.[b_k]}$$

- $b_k = \frac{\sum_{i=1}^{N} v_{ki} y_i}{\sum_{i=1}^{N} v_{ki}^2}$

- $s.e.[b_k] = \frac{\sigma^2}{\sum_{i=1}^{N} v_{ki}^2}$

- $est\ s.e.(b_k) = \frac{s^2}{\sum_{i=1}^{N} v_{ki}^2}, \quad s^2 = \frac{SSE}{N-K-1} = \frac{\sum_{i=1}^{N} e_i^2}{N-K-1}$

Do We Have a Model?

- Compare the R^2 from two regressions, one with all the RHS variables in it, the other without any RHS variables (only an intercept, which will equal the mean of Y)

- Form the Wald statistic:

$$W = \frac{R^2_{alt} - R^2_{null}}{(1 - R^2_{alt}) / (N - K_{alt} - 1)}$$

- K_{alt} in the denominator is the number of X variables in the main model

- The more the RHS variables "explain" the variation in Y, the bigger this test statistic will be

- Under the null hypothesis, the Wald statistic has a Chi-square distribution with q degrees of freedom, where q is the number of RHS variables omitted under the null hypothesis

Example: Do We Have a Model to Predict API in CA Schools?

$$W = \frac{0.86 - 0}{(1 - 0.86)/(20 - 2 - 1)} = 104.43$$

Because 104.43 > 5.99, we reject the null hypothesis at 5% significance

Degrees of Freedom	Significance level		
	0.10	0.05	0.01
1	2.71	3.84	6.64
2	4.61	5.99	9.21
3	6.25	7.82	11.35
4	7.78	9.49	13.28
5	9.24	11.07	15.09

How to Present Regression Results 1

$$Y_i = 777.17 - 0.51X_{1i} + 2.34X_{2i} + e_i$$
$$(37.92) \quad (0.40) \quad (0.47)$$

$$\text{Sample size} = 20$$
$$R^2 = 0.84$$

How to Present Regression Results 2

Variable	Estimated Coefficient	Standard Error	t-Statistic
Free-lunch eligibility	–0.51	0.40	–1.28
Parents' education	2.34	0.47	5.01
Constant	777.17	37.92	20.50
Sample size	20		
R^2	0.84		

What We Learned

- We reject the null hypothesis of zero relationship between free lunch eligibility (FLE) and academic performance.
  - Our result is the same whether we drop CR4 and invoke the central limit theorem (valid in large samples) or whether we impose CR4 (necessary in small samples).
- Confidence intervals are narrow when the sum of squared errors is small, the sample is large, or there's a lot of variation in X.
- How to present results from a regression model.