Lecture 16: Finish 9.2 Cat. Inference; Start 10.1 Inference for Quantitative Variables: Confidence Intervals

- Small P-values; Proportion Hypothesis Tests in Long Run
- Relating Proportion Test Results to Confidence Interval
- Inference for Means vs. Proportions
- Population Standard Deviation Known or Unknown
- Constructing CI for Mean (S.D. Known)
- Checking Normality
- Details of Confidence Interval for Mean

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-3)
  - Displaying and Summarizing (Lectures 3-8)
  - Probability (discussed in Lectures 9-14)
  - Statistical Inference
    - 1 categorical: conf. ints. (L14) , hyp.tests (began L15)
    - 1 quantitative
    - categorical and quantitative
    - 2 categorical
    - 2 quantitative

Example: *Cut-Offs for "Small" P-Value*

- **Background**: Bookstore chain will open new store in a city if there's evidence that its proportion of college grads is higher than 0.26, the national rate.
- **Question:** Choose cut-off (0.10, 0.05, 0.01):
  - if no other info is provided
  - if chain is enjoying considerable profits; owners are eager to pursue new ventures
  - if chain is in financial difficulties, can't afford losses if unsuccessful due to too few grads
- **Response:**
  - ______
  - ______
  - ______

Practice: 9.56 p.445

Definition

Statistically significant data: produce *P*-value small enough to reject H_0. *z* plays a role:

$$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{(\hat{p}-p_0)\sqrt{n}}{\sqrt{p_0(1-p_0)}}$$

Reject H_0 if *P*-value small; if |*z*| large; if…

- Sample proportion $\hat{p}$ far from p_0
- Sample size *n* large
- Standard deviation small (if p_0 is close to 0 or 1)

Role of Sample Size *n*

- **Large *n***: may reject H_0 even though observed proportion isn't very far from p_0, from a practical standpoint.

Very small *P*-value→strong evidence against Ho but *p* not necessarily very far from *po*.

- **Small *n***: may fail to reject H_0 even though it is false.

Failing to reject false Ho is 2nd type of error

Definition

- **Type I Error:** reject null hypothesis even though it is true (false positive)
  - Probability is cut-off α
- **Type II Error:** fail to reject null hypothesis even though it's false (false negative)

Hypothesis Test and Long-Run Behavior

Repeatedly carry out hypothesis tests of p=0.5, based on 20 coinflips, using cut-off 5%.

In the long run, 5% of the tests will reject $H_0: p=0.5$, even though it's true.

Hypothesis Test and Long-Run Behavior

20 coin flips

test Ho: p=.50 vs. Ha: p not equal .50
(reject if p-value<.05)

TTTTTHTHTTHHTHTHTTHH
proportion of heads 9/20=.45 — z=-.45, p-value=.655 → do not reject Ho

HTTHHTHHTTTHTHTTTHHT
proportion of heads 8/20=.40 — z=-.89, p-value=.371 → do not reject Ho

TTTHHHHHHHHTHTHHTHTT
proportion of heads 12/20=.60 — z=+.89, p-value=.371 → do not reject Ho

THHHHTHHHTHTHHHTHHHH
proportion of heads 15/20=.75 — z=+2.24, p-value=.025 → reject Ho

- repeated
- flips of 20
- coins

in the long run
95% of tests do not reject Ho
5% of tests reject Ho

TTHTTTTHTTTHTTHHTHHH
proportion of heads 8/20=.40 — z=-.89, p-value=.371 → do not reject Ho

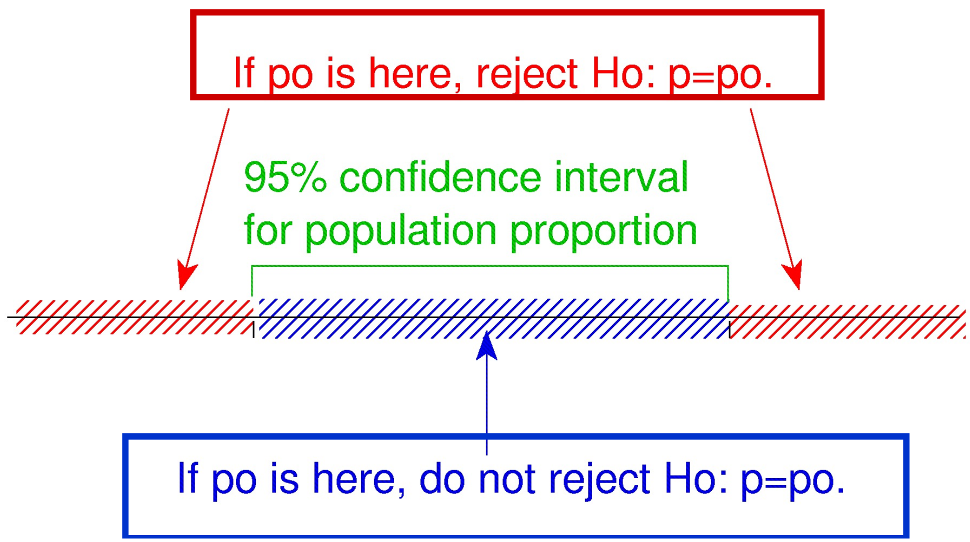
Confidence Interval and Hypothesis Test Results

- Confidence Interval: range of plausible values
- Hypothesis Test: decides if a value is plausible

Informally,

- If *po* is in confidence interval, don't reject Ho: $p=po$
- If *po* is outside confidence interval, reject Ho: $p=po$

Relationship between 95% confidence interval
and two-sided test with .05 as cut-off for p-value



Example: *Test Results, Based on C.I.*

- **Background**: A 95% confidence interval for proportion of all students choosing #7 "at random" from numbers 1 to 20 is (0.055, 0.095).
- **Question:** Would we expect a hypothesis test to reject the claim $p=0.05$ in favor of the claim $p>0.05$?
- **Response:**

Practice: 9.65b p.448

Example: *C.I. Results, Based on Test*

- **Background**: A hypothesis test did not reject H_0: p=0.5 in favor of the alternative H_a: p<0.5.
- **Question:** Do we expect 0.5 to be contained in a confidence interval for p?
- **Response:**

Practice: 9.38f p.440

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-3)
  - Displaying and Summarizing (Lectures 3-8)
  - Probability (discussed in Lectures 9-14)
  - Statistical Inference
    - 1 categorical (discussed in Lectures 14-16)
    - 1 quantitative: confidence intervals, hypothesis tests
    - categorical and quantitative
    - 2 categorical
    - 2 quantitative

Inference for Proportions or Means: Similarities

- 3 forms of inference (point estimate, CI, test)
- Point est. unbiased estimator for parameter if…?
- **Confidence Interval**: estimate ± margin of error
  = sample stat ± multiplier × s.d. of sample stat
  - Sample stat must be unbiased
  - Sample must be large enough so multiplier is correct

    Note: higher confidence→larger multiplier→wider interval
  - Pop at least 10*n* so s.d. is correct

    Note: larger sample→smaller s.d.→narrower interval

  Correct interpretation of interval; interval related to test.

Inference for Proportions or Means: Similarities

- **Hypothesis Test**: Does parameter = proposed value?
  - 3 forms of alternative (greater, less, not equal)
  - 4-steps follow 4 processes of statistics
    1. Data production: sample unbiased? n large? pop $\geq 10n$?
    2. Find sample statistic and standardize; is it "large"?
    3. Find P-value=prob of sample stat this extreme; is it "small"?
    4. Draw conclusions: reject null hypothesis if P-value is small
  - P-value for 2-sided alternative twice that for 1-sided
  - Cut-off level α (often 0.05) is probability of Type I Error (false positive)
  - Rejection: if sample stat far from proposed parameter, or n large, or spread small
  - Type II Error (false negative) also possible, especially for small n

Inference for Proportions or Means: Differences

- Different summaries for quantitative variables
  - Population mean μ
  - Sample mean $\bar{x}$
  - Population standard deviation σ
  - Sample standard deviation s

  (For proportions, s.d. could be calculated from n and p)
- Standardized statistic not always "z"
- No easy Rule of Thumb for what n is large enough to ensure normality; must examine shape of sample data.

Three Types of Inference Problem

Mean yearly earnings for sample of 446 students at a particular university was $3,776.

1. What is our best guess for the mean earnings of all students at that university?

   (Point Estimate)

2. What interval should contain mean earnings for all the students?

   (Confidence Interval)

3. Is this convincing evidence that mean earnings for all the students is less than $5,000?

   (Hypothesis Test)

Behavior of Sample Mean *(Review)*

For random sample of size n from population with mean μ, sample mean $\bar{X}$ has

- mean μ

→ $\bar{X}$ is *unbiased estimator* of μ

(sample must be random)

Example: *Checking if Estimator is Unbiased*

- **Background**: Anonymous on-line survey of intro stat students (various ages, majors) at a university produced sample mean earnings.
- **Questions:**
  - Is the sample representative of all students at that university? Does it represent *all* college students?
  - Were the values of the variable (earnings) recorded without bias?
- **Responses:**
  - Various ages, majors → ______________________________

    Socio-economic conditions depend on school → ____
  - Anonymous online survey → __________________________

Practice: 10.1b p.477

Example: *Point Estimate for* μ

- **Background**: In a representative sample of students at a university, mean earnings were $3,776.
- **Question:** What is our best guess for mean earnings of all students at that university?
- **Response:** $\bar{X}$ is an unbiased estimator for μ so ________ is our best guess.

Looking Ahead: *For point estimate we don't need to know s.d. For confidence intervals and hypothesis tests, to quantify how good our point estimate is, we must know sigma or estimate it with s. This makes an important difference in procedure. We also need n.*

Practice: 10.1a p.477

Three Types of Inference Problem *(Review)*

Mean yearly earnings for sample of 446 students at a particular university was $3,776.

1. What is our best guess for the mean earnings of all students at that university?

   (Point Estimate)

2. What interval should contain mean earnings for all the students?

   (Confidence Interval)

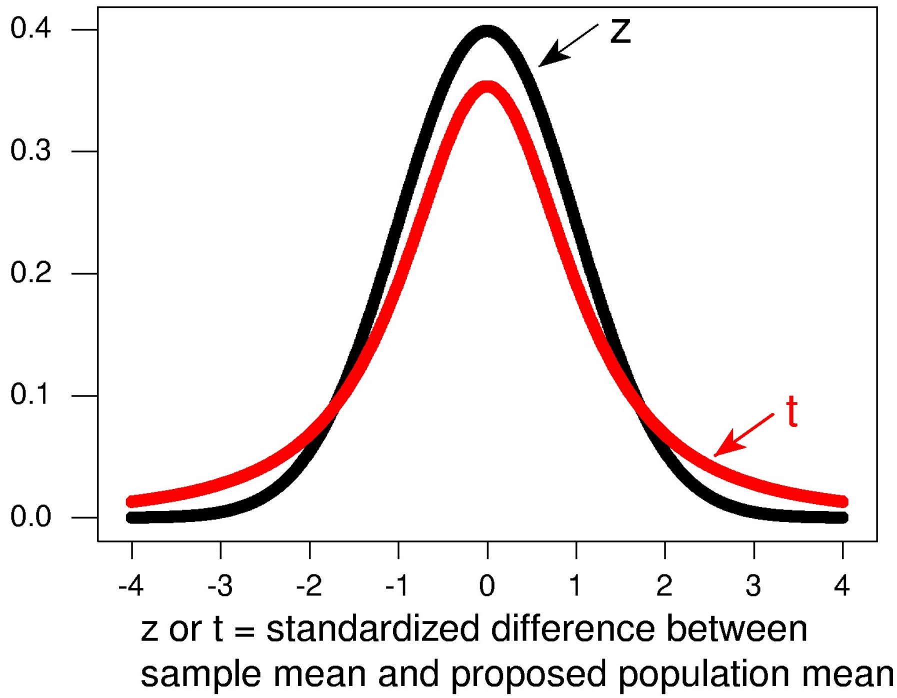
3. Is this convincing evidence that mean earnings for all the students is less than $5,000?

   (Hypothesis Test)

Inference About Mean Based on *z* or *t*

- σ known→ standardized $\bar{x}$ is *z*
- σ unknown→ standardized $\bar{x}$ is *t*

(may use *z* if σ unknown but *n* large)



Inference with t discussed after inference with z

Behavior of Sample Mean *(Review)*

For random sample of size *n* from population with mean μ, standard deviation σ, sample mean $\bar{X}$ has

- mean μ
- standard deviation $\frac{\sigma}{\sqrt{n}}$
- shape approximately normal for large enough *n*

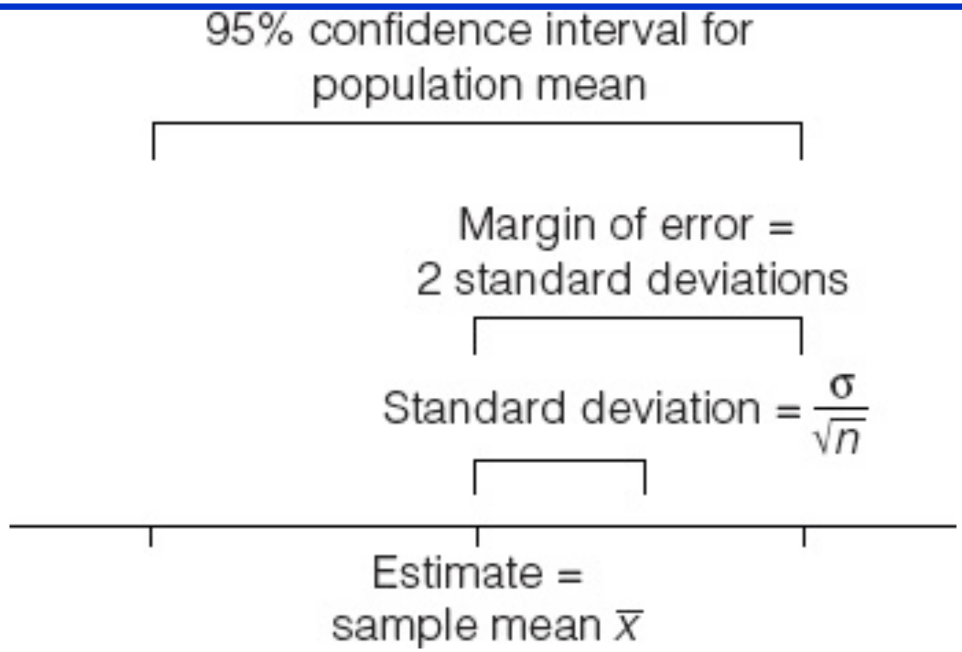
→Probability is 0.95 that $\bar{X}$ is within $2\frac{\sigma}{\sqrt{n}}$ of μ

Looking Ahead: Probability *results lead to* ***confidence*** *interval for* μ.

Confidence Interval for Population Mean

95% confidence interval for μ is $\bar{x} \pm 2\frac{\sigma}{\sqrt{n}}$

- Sample must be unbiased
- Population size must be at least 10*n*
- *n* must be large enough to justify multiplier 2 from normal distribution



Guidelines for Sample Mean Approx. Normal

Can assume shape of $\bar{X}$ for random samples of size *n* is approximately normal if

- Graph of sample data appears normal; or
- Graph of sample data fairly symmetric and *n* at least 15; or
- Graph of sample data moderately skewed and *n* at least 30; or
- Graph of sample data very skewed and *n* much larger than 30

A Closer Look: *Besides examining display, consider what shape we'd* ***expect*** *to see for the variable's distribution.*

Example: *Revisiting Original Question*

- **Background**: Mean yearly earnings for 446 students at a particular university was \$3,776. Assume population standard deviation \$6,500.
- **Question:** Assuming sample is representative, what interval should contain population mean earnings?
- **Response:** 95% C.I. for μ is $\bar{x} \pm 2\frac{\sigma}{\sqrt{n}} =$

***A Closer Look:** 446 is large enough to offset right skewness.*

Practice: 10.9b-c p.478

Example: *C.I. as Range of Plausible Values*

- **Background**: Mean yearly earnings for 446 students at a particular university was \$3,776. Assume population standard deviation \$6,500. 95% confidence interval for μ is (3160, 4392).
- **Question:** Is \$5,000 a plausible value for population mean earnings?
- **Response:**

Looking Ahead: *This kind of decision is addressed more formally and precisely with a hypothesis test.*

Practice: 10.9d p.478

Example: *Role of Sample Size in C.I.*

- **Background**: Mean yearly earnings for 446 students at a particular university was \$3,776. Assume population standard deviation \$6,500.
  A 95% confidence interval for μ is $3,776 \pm 616$.
- **Question:** What would happen to the C.I. if *n* were one fourth the size (111 instead of 446)?
- **Response:** Divide *n* by 4 → ____________________

$$\bar{x} \pm 2\frac{\sigma}{\sqrt{n}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Practice: 10.9e p.478

Example: *Other Levels of Confidence*

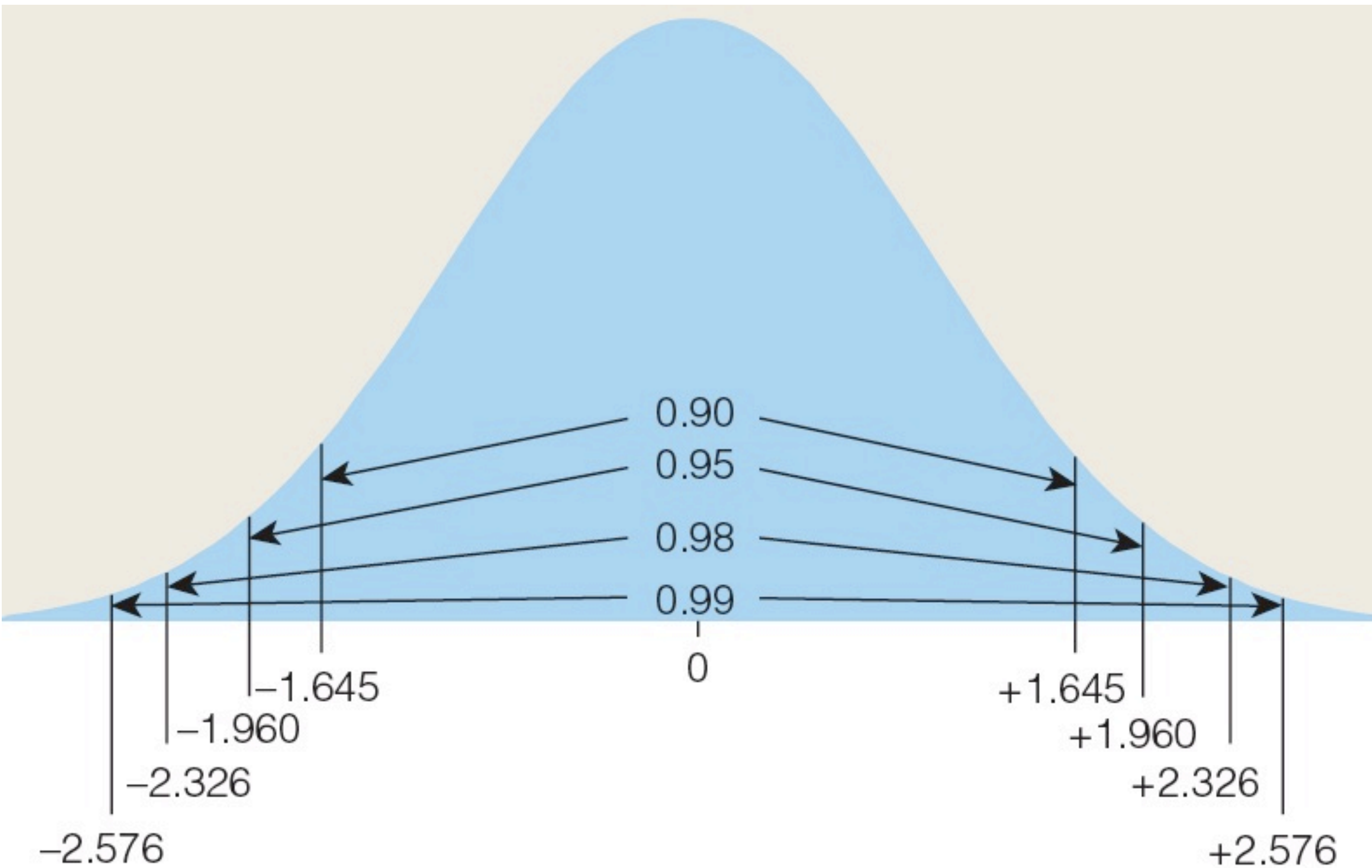
- **Background**: Mean yearly earnings for 446 students at a particular university was \$3,776. Assume population standard deviation \$6,500. A 95% confidence interval for μ

$$= 3776 \pm 2\frac{6500}{\sqrt{446}} = 3776 \pm 616 = (3160, 4392)$$

- **Question:** How would we construct intervals at 90% or 99% confidence?
- **Response:**

Other Levels of Confidence

"Inside" probabilities correspond to various multipliers.



Other Levels of Confidence *(Review)*

Confidence level 95% uses multiplier 2. Other levels use other multipliers, based on normal curve.

More precise multiplier for 95% is 1.96 instead of 2.

Level	Multiplier
90%	1.645
95%	1.960
98%	2.326
99%	2.576

Example: *Other Levels of Confidence*

- **Background**: Mean yearly earnings for 446 students at a particular university was \$3,776. Assume population standard deviation \$6,500.
- **Question:** What are 90% and 99% confidence intervals for population mean earnings?
- **Response:** Interval is $3776 \pm \text{multiplier}\frac{6500}{\sqrt{446}}$
  - **90% C.I.** ______________________ $= (3270, 4282)$
  - **99% C.I.** ______________________ $= (2983, 4569)$

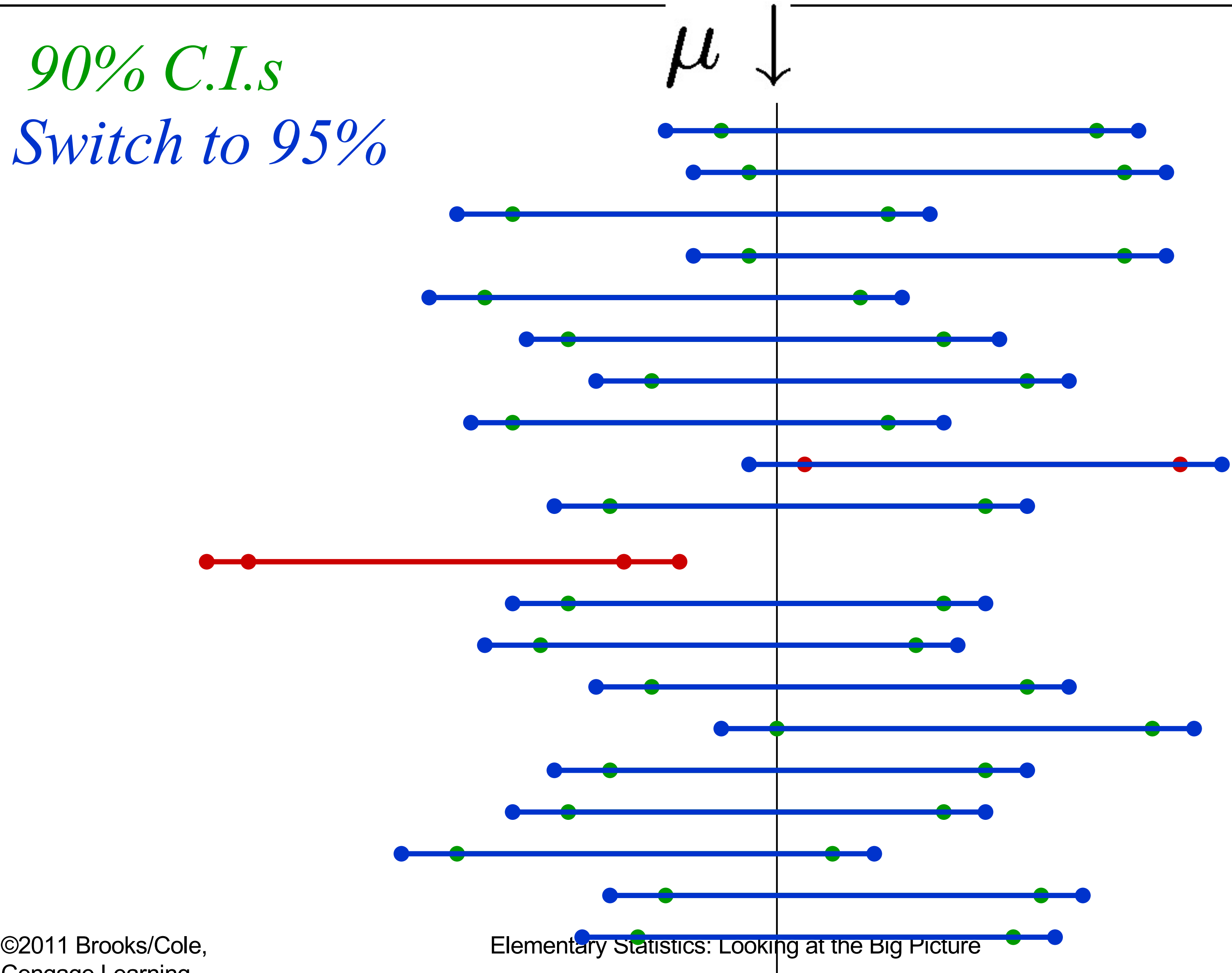
Tradeoff: higher level of confidence→_____ precise interval

Wider Intervals ←→ More Confidence

Consider illustration of many 90% confidence intervals in the long run: 18 in 20 should contain population parameter.

If they were widened to 95% intervals (multiply s.d. by 2 instead of 1.645), then they'd have a higher probability (19 in 20) of capturing population parameter.

Wider Intervals ←→ More Confidence



Interpretation of XX% Confidence Interval

- We are XX% confident that the interval contains the unknown parameter.
- XX% intervals' long-run probability of capturing the unknown parameter is XX%.

Example: *Interpreting Confidence Interval*

- **Background**: A 95% confidence interval for mean U.S. household size μ is (2.166, 2.714).
- **Question:** Which of the following are true?
  - Probability is 95% that μ is in the interval (2.166, 2.714).
  - 95% of household sizes are in the interval (2.166, 2.714).
  - Probability is 95% that $\bar{x}$ is in the interval (2.166, 2.714).
  - We're 95% confident that $\bar{x}$ is in interval (2.166, 2.714).
  - We're 95% confident that μ is in interval (2.166, 2.714).
  - The probability is 95% that our sample produces an interval which contains μ.
- **Response:** ______________________________

Practice: 10.10 p.479

Lecture Summary
(More Hypothesis Tests for Proportions)

- Cut-off for "small" *P*-value
- Statistical significance; role of *n*, Type I or II Error
- Hypothesis tests in long-run
- Relating tests and confidence intervals

Lecture Summary
(Inference for Means: Confidence Interval)

- Inference for means vs. proportions
  - Similarities (many)
  - Differences: population s.d. may be unknown
- Constructing CI for mean with z (pop. s.d. known)
- Checking assumption of normality
- Role of sample size
- Other levels of confidence
- Interpreting the confidence interval