Chapter 21

What is a Confidence Interval?

Confidence interval for proportion p from a binomial distribution is

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Related to this, sampling distribution for $\hat{p}$ is approximately normal with a mean of p and a standard deviation of $\sqrt{\frac{p(1-p)}{n}}$. Confidence interval for mean μ is

$$\bar{x} \pm z^* \left(\frac{s}{\sqrt{n}}\right).$$

Related to this, sampling distribution for $\bar{x}$ is approximately normal with a mean of μ and a standard deviation of $\frac{\sigma}{n}$.

For both confidence intervals and sampling distributions, it is assumed a large random sample has been chosen.

Exercise 21.1 (What is a Confidence Interval?)

1. *Decayed teeth.* Of population of millions of American children who had their teeth checked at dentist in this past year, a random sample of size 75 is taken and it is found 42 of these children had decayed teeth.

   (a) point estimate of p is $\frac{42}{75}$ = **0.52** / **0.54** / **0.56** / **0.58**.

   (b) 95% confidence interval of p

   $$\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.56 \pm 2\sqrt{\frac{0.56(1-0.56)}{75}} \approx$$

   $\mathbf{0.56 \pm 0.09}$ / $\mathbf{0.56 \pm 0.11}$ / $\mathbf{0.56 \pm 0.13}$

2. *Credit card purchases.* Fifty-four (54) of 180 randomly selected credit card purchase slips (from thousands written) were made with Visa credit card at a department store.

(a) point estimate of p is $\frac{54}{180}$ = **0.2** / **0.3** / **0.4** / **0.5**.

(b) 95% confidence interval of p

$$\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3 \pm 2\sqrt{\frac{0.3(1-0.3)}{180}} \approx$$

0.333 ± 0.068 / **0.300 ± 0.068** / **0.3 ± 0.110**

(c) 95% confidence interval of p, 0.3 ± 0.068, could also be written as

i. from $0.300 - 0.068 = 0.232$ to $0.300 + 0.068 = 0.368$

ii. from $0.300 + 0.068 = 0.368$ to $0.300 - 0.068 = 0.232$

iii. from $0.300 - 0.070 = 0.230$ to $0.300 + 0.070 = 0.370$

(d) **True** / **False**. We are 95% *confident* population p of credit card purchases made with Visa in interval $(0.232, 0.368)$.

3. *Credit Card Purchase Slips Revisited.* Fifty-four (59) (not 54!) of 180 randomly selected purchases made with Visa.

(a) point estimate of p is $\frac{59}{180}$ ≈ **0.31** / **0.33** / **0.37**.

(b) point estimate $\hat{p} \approx 0.33$
same / **different from**
previous point estimate $\hat{p} = 0.30$

(c) 95% confidence interval of p

$$\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.33 \pm 2\sqrt{\frac{0.33(1-0.33)}{180}} \approx$$

0.330 ± 0.070 / **0.333 ± 0.070** / **0.33 ± 0.110**

(d) 95% confidence interval $0.330 \pm 0.070 = (0.260, 0.400)$
same as / **different from**
previous 95% confidence interval $0.300 \pm 0.068 = (0.232, 0.368)$

(e) Point estimate and confidence interval of *third* random sample most likely
same as / **different from**
previous point estimates and confidence intervals.

(f) Not all confidence intervals contain population proportion p. In fact, exactly **68%** / **95%** / **99.7%** of *all* possible CIs contain population proportion (and 5% do not), as shown below.

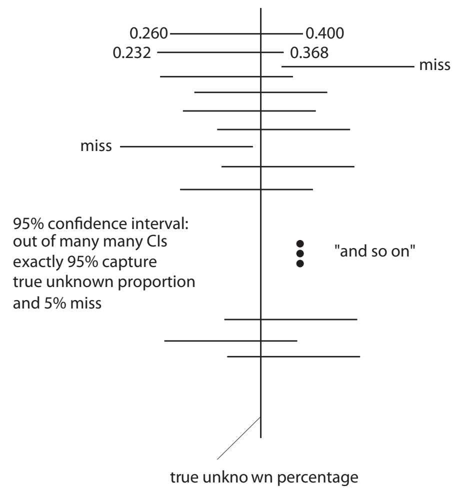

Figure 21.1 (Interpreting 95% confidence intervals)

4. *Tongue Studs.* Thirty-one of 305 randomly selected students had tongue studs.

   (a) point estimate of p is $\frac{31}{305} \approx$ **0.06** / **0.08** / **0.10**.

   (b) 95% confidence interval of p

   $$\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm 2\sqrt{\frac{0.1(1-0.1)}{305}} \approx$$

   0.1 ± 0.017 / **0.1 ± 0.034** / **0.1 ± 0.052**

   (c) 68% confidence interval of p

   $$\hat{p} \pm \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm \sqrt{\frac{0.1(1-0.1)}{305}} \approx$$

   0.1 ± 0.017 / **0.1 ± 0.034** / **0.1 ± 0.052**

   (d) 68% confidence interval is **smaller** / **longer** than 95% confidence interval

   (e) 99.7% confidence interval of p

   $$\hat{p} \pm 3\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm 3\sqrt{\frac{0.1(1-0.1)}{305}} \approx$$

   0.1 ± 0.017 / **0.1 ± 0.034** / **0.1 ± 0.052**

   (f) 99.7% confidence interval is **smaller** / **longer** than 95% confidence interval

   (g) Using Table 21.1 for 80% CI of p, $z^* =$ **1.28** / **1.64** / **1.96**, so

   $$\hat{p} \pm z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm 1.28\sqrt{\frac{0.1(1-0.1)}{305}} \approx$$

   0.1 ± 0.017 / **0.1 ± 0.022** / **0.1 ± 0.052**
   z^* is called the *critical value*

(h) Using Table 21.1 for 99% CI of p, $z^* =$ **2.28** / **2.58** / **2.96**, so

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm 2.58 \sqrt{\frac{0.1(1-0.1)}{305}} \approx$$

$\mathbf{0.1 \pm 0.017}$ / $\mathbf{0.1 \pm 0.022}$ / $\mathbf{0.1 \pm 0.044}$

5. *Properties, terminology: decayed teeth again.* Forty-two (42) of 75 randomly selected children had decayed teeth.

(a) Population proportion, p, of children with decayed teeth **known** / **unknown**

(b) Statistic $\hat{p}$ has mean $\boldsymbol{p}$ / $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ / $\sqrt{\frac{p(1-p)}{n}}$ estimated by $\hat{p} = \frac{42}{75} = 0.56$.

(c) Statistic $\hat{p}$ has SD of $\boldsymbol{p}$ / $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ / $\sqrt{\frac{p(1-p)}{n}}$ which is estimated by

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.56(1-0.56)}{75}} \approx 0.057.$$

(d) 95% confidence interval of p

$$\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.56 \pm 2\sqrt{\frac{0.56(1-0.56)}{75}} \approx 0.56 \pm 0.114$$

has margin of error $\sqrt{\frac{0.56(1-0.56)}{75}} \approx \mathbf{0.057}$ / $\mathbf{2}\sqrt{\frac{0.56(1-0.56)}{75}} \approx \mathbf{0.114}$

(e) Critical value, z^*, of 95% confidence interval of p is **1.28** / **1.64** / **1.96**

(f) Critical value, z^*, of 80% confidence interval of p is **1.28** / **1.64** / **1.96**

6. *Properties, terminology: credit card purchases again.* Fifty-four (54) of 180 randomly selected purchases made with Visa. Unknown to you, proportion of all purchases made with Visa is $p = 0.35$.

(a) Population proportion, $p = 0.35$ is **critical value** / **mean** / **parameter** / **statistic** / **bias**

(b) Sample proportion, $\hat{p} = \frac{54}{180} = 0.3$ is **critical value** / **mean** / **parameter** / **statistic** / **bias**

(c) Sample proportion $\hat{p}$ has mean $\boldsymbol{p} = \mathbf{0.35}$ / $\hat{\boldsymbol{p}} = \frac{\mathbf{54}}{\mathbf{180}} = \mathbf{0.3}$

(d) Sample proportion $\hat{p}$ has *approximate* standard deviation (SD) of

i. $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.3(1-0.3)}{180}} \approx 0.034$

ii. $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.35(1-0.35)}{180}} \approx 0.036$

(e) If simple random sample (SRS) increased from $n = 180$ to $n = 4 \times 180 = 720$,

   (i) both mean and SD $\frac{1}{2}$ as large.
   (ii) same mean, but standard deviation $\frac{1}{2}$ as large.
   (iii) mean $\frac{1}{2}$ as large, same standard deviation.
   (iv) same mean, but standard deviation $\frac{1}{4}$ as large.
   (v) mean $\frac{1}{4}$ as large, same standard deviation.

   Hint: SD $\sqrt{\frac{0.35(1-0.35)}{180}} \approx 0.036$ becomes $\sqrt{\frac{0.35(1-0.35)}{720}} \approx 0.018$

7. *CI for mean: diet.* From a random sample 64 middle aged men on a diet, mean fat intake is 37 grams and SD in fat intake is 32 grams.

(a) point estimate of μ is $\bar{x} =$ **31** / **37** / **64**.

(b) 95% confidence interval of population mean fat intake μ

$$\bar{x} \pm 2\left(\frac{s}{\sqrt{n}}\right) = 37 \pm 2\frac{32}{\sqrt{64}} =$$

37 ± 7.84 / **37 ± 8** / **37 ± 32**

(c) We are 95% confident actual mean fat intake is in interval

   i. (29, 45)
   ii. (27.75, 46.25)

(d) Using Table 21.1 for 95% CI of μ, $z^* =$ **1.96** / **2** / **2.58**, so

$$\bar{x} \pm z^*\left(\frac{s}{\sqrt{n}}\right) = 37 \pm 1.96\frac{32}{\sqrt{64}} \approx$$

37 ± 7.84 / **37 ± 9.065** / **37 ± 32**

(e) **True** / **False**. If we took many SRSs of middle aged men, 95% of all 95% CIs calculated would include population mean, μ.

(f) Match columns.

Column I	Column II
(a) population	(A) mean fat intake of 64, $\bar{x}$
(b) sample	(B) fat intakes of 64
(c) statistic	(C) mean fat intake of all middle aged men, μ
(d) parameter	(D) fat intakes of all middle aged men

terms	(a)	(b)	(c)	(d)
weight control example				

8. *CI for mean: salaries.* From a SRS of 40 employees, mean salary is \$26,000 with SD in salary of \$15,000.

   (a) point estimate of μ is $\bar{x} =$ **40** / **\$15,000** / **\$26,000**.

   (b) 95% confidence interval of population (company) mean salary μ

$$\bar{x} \pm 2\left(\frac{s}{\sqrt{n}}\right) = \$26,000 \pm 2\left(\frac{\$15,000}{\sqrt{40}}\right) \approx$$

   \$26,000 ± \$2,371.71 / **\$26,000 ± \$4,743.42** / **\$26,000 ± \$8,003.21**

   (c) Using Table 21.1 for 70% CI of μ, $z^* =$ **0.84** / **1.04** / **1.28**, so

$$\bar{x} \pm z^*\left(\frac{s}{\sqrt{n}}\right) = \$26,000 \pm 1.04\left(\frac{\$15,000}{\sqrt{40}}\right) \approx$$

   \$26,000 ± \$2,466.58 / **\$26,000 ± \$3,743.42** / **\$26,000 ± \$4,003.21**

   (d) Match columns.

terms	widget example
(a) population	(A) mean salary of all employees, μ
(b) sample	(B) salaries of 40 employees
(c) statistic	(C) salaries of all employees
(d) parameter	(D) mean salary of 40 employees, $\bar{x}$

terms	(a)	(b)	(c)	(d)
widget example				