Part III

Estimation

Estimation

WHAT DO STATISTICIANS DO?

Statisticians help determine the sampling and data collection methods monitor the execution of the study and the processing of data, and advise on the strengths and limitations of the results. They must understand the nature of uncertainties and be able to draw conclusions in the context of particular statistical applications.

Surveys: Survey statisticians collect information from a carefully specified sample and extend the results to an entire population. Sample surveys might be used to:

- Determine which political candidate is more popular
- Discover what foods teenagers prefer for breakfast
- Estimate the number of children living in a given school district

Government Operations: Government statisticians conduct experiments to aid in the development of public policy and social programs. Such experiments include:

- Consumer prices
- Fluctuations in the economy
- Employment patterns
- Population trends

Scientific Research: Statistical sciences are used to enhance the validity of inferences in:

- Radiocarbon dating to estimate the risk of earthquakes
- Clinical trials to investigate the effectiveness of new treatments
- Field experiments to evaluate irrigation methods
- Measurements of water quality
- Psychological tests to study how we reach the everyday decisions in our lives

Business and Industry: Statisticians quantify unknowns in order to optimize resources. They:

- Predict the demand for products and services
- Check the quality of items manufactured in a facility
- Manage investment portfolios
- Forecast how much risk activities entail, and calculate fair and competitive insurance rates.

Source: http://www.amstat.org/Careers/index.cfm?fuseaction=whatisstatistics

General example:

1. What is the **average** life of Diehard batteries? $\mu = ?$
2. What is the **average** waiting time at a supermarket register? $\mu = ?$
3. What is the **average** length of time to finish a 2-year degree? $\mu = ?$
4. What **percentage** of college students graduate in 4 years? $P = ?$
5. What **percentage** of Dell customers will be shopping again from Dell? $P = ?$
6. What **percentage** of DMV applicant will be passing driving test for the first time? $P = ?$
7. What is the **average difference** in battery life between Diehard and Everlast brand? $\mu_D - \mu_E = ?$

Estimating One Population Mean μ

$$\mu = \overline{x} \pm E$$

$\overline{X}$ = **Point** estimate (Sample Mean)	E = Margin of error		
Decision making process on σ			
σ is population standard deviation	σ (known)	σ (unknown) $n > 30$	σ (unknown) $n \le 30$
Margin of Error	$E = z\frac{\sigma}{\sqrt{n}}$ (For **Z,** use Table page **3**	$E = z\frac{s}{\sqrt{n}}$ (For **Z,** use Table page **3)**	$E = t\frac{s}{\sqrt{n}}$ (For **t,** use Table page **4)**
Interval Estimate	$\mu = \overline{x} \pm E$	$\mu = \overline{x} \pm E$	$\mu = \overline{x} \pm E$
TI-83/84	*stat* → *tests* → *Option 7*	*stat* → *tests* → *Option 7*	*stat* → *tests* → *Option 8*

Width (difference between upper and lower bounds**)** $= 2E = UB - LB$ **so** $E = (UB - LB)/2$

Estimating Population Proportion P

$$P = \hat{p} \pm E$$

$\hat{p} = \frac{x}{n}$ **(Called p-hat is sample proportion and point estimate for population proportion)**	**E = Margin of error** $E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
TI-83 ***stat*** → ***test*** → ***Option A:***	

Estimating the *difference* between Two Populations

Mean $\mu_1 - \mu_2$	Proportion $P_1 - P_2$
$\mu_1 - \mu_2 = (\overline{x}_1 - \overline{x}_2) \pm E$	$P_1 - P_2 = (\hat{p}_1 - \hat{p}_2) \pm E$
Point **estimate** $= (\overline{x}_1 - \overline{x}_2)$	Point **estimate** $= (\hat{p}_1 - \hat{p}_2)$
$E = Z\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$E = Z\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
TI-83/84 ***stat*** → ***test*** → ***Option 9***	TI-83/84 ***stat*** → ***test*** → ***Option B:2***

Sample Size **Determining for the** Estimation **of Population**

Mean = μ	Proportion = P
$n = (Z\,S/E)^2$	$n = (Z/E)^2\,\hat{p}(1-\hat{p})$
If S **is** unknown **then estimate it by** $S = Range/4$	**If** $\hat{p}$ **is** unknown **then estimate it by** $\hat{p} = 0.5$

Estimating Population or True Mean μ (σ is known and Large Samples, $n > 30$)

μ = Population or true mean $\bar{x} =$ **Sample mean** $\mu = \bar{x} \pm E$ **margin of error =** $E = z\frac{s}{\sqrt{n}}$

(for z- value use web-table sheet page 3)

Example: To estimate the average life of Diehard batteries. A sample of **64** batteries showed that **$\bar{x}$ = 50 months, s = 10 months**.

a) Construct a **95%** confidence interval for average life of all Diehard batteries.
b) Construct a **90%** confidence interval for average life of all Diehard batteries.
c) Construct a **99%** confidence interval for average life of all Diehard batteries.
d) Construct a **95%** confidence interval for average life of all Diehard batteries if **n = 100.**
e) Construct a **95%** confidence interval for average life of all Diehard batteries if **n = 36**.

μ = Population or true mean $\bar{x} =$ **Sample mean** $\mu = \bar{x} \pm E$ **margin of error =** $E = z\frac{s}{\sqrt{n}}$

Observation: compared with part a, write how the **error** (smaller/ larger) and **interval** get changed (narrower/ wider). Round answers in 2 decimal places. **All Answers on page 12**

	Z	n	$E = z\frac{s}{\sqrt{n}}$	$\mu = \bar{x} \pm E$	< μ <	**Observation** Compare all parts to part **(a),** draw conclusion about E , Interval
a	$z_{.95} =$ **1.96**	64	$1.96\frac{10}{\sqrt{64}} = 2.45$	50 ± 2.45	$47.55 < \mu < 52.45$ By 95% confidence, the average life of Diehard batteries is between 48 to 53 months	**Leave it blank**
b	$z_{.90} =$ 1.645	64				**E** gets **smaller**, ***narrower* Interval**
c	$z_{.99} =$ 2.58	64				**E** gets **larger**, **wider Interval**
d	$z_{.95} =$ 1.96	100				**E** gets **smaller**, ***narrower* Interval**
e	$z_{.95} =$ 1.96	36				**E** gets **larger**, **wider Interval**

Extra **P**ractice: Please do **p**roblems **1,6,10,11,14,15, 16, 18, 35** from **p**ractice **p**roblems of **part III**

Estimating Population or True Mean μ (σ is unknown, Small Samples $n \le 30$)

Note 1: When σ (population Standard deviation) is unknown and $n \le 30$, we use a different distribution called t-distribution,

μ = Population or true mean $\bar{x}$ = **Sample mean** $\mu = \bar{x} \pm E$ **margin of error =** $E = t\frac{s}{\sqrt{n}}$

(For t-value use P. 774 Triola or web-table sheet page 4)

Note 2: Finding the **t-value** based on a given confidence level and **df** = degree of freedom = n-1
Ex 1: a) **95%** confidence level, $n = 10$, $df = 10 - 1 = 9$, $\Rightarrow$ **t-value** will be $\Rightarrow$ $t = 2.262$
Ex 2: a) **99%** confidence level, $n = 13$, $df = 13 - 1 = 12$, $\Rightarrow$ **t-value** l Value will be $\Rightarrow$ $t = 3.055$

Example: To estimate the average life of Diehard batteries. A sample of **16** batteries showed that
$\bar{x} = 50$ **months**, $s = 10$ **months**.

a) Construct a **95%** confidence interval for average life of all Diehard batteries.
b) Construct a **90%** confidence interval for average life of all Diehard batteries.
c) Construct a **99%** confidence interval for average life of all Diehard batteries.
d) Construct a **95%** confidence interval for average life of all Diehard batteries if **n = 25.**
e) Construct a **95%** confidence interval for average life of all Diehard batteries if **n = 9**.

. **μ = Population or true mean** $\bar{x}$ = **Sample mean** $\mu = \bar{x} \pm E$ **margin of error =** $E = t\frac{s}{\sqrt{n}}$

Observation: compared with part a, write how the **error** and **interval** get changed (smaller/ larger/ narrower/ wider). **Round your answers in 2 decimal places.**

	n	df=n-1 t = ?	$E = t\frac{s}{\sqrt{n}}$	$\mu = \bar{x} \pm E$	< μ <	**Observation** Compare all parts to part **(a),** draw conclusion about E , Interval
a	16	df=16-1 df =15 $t_{.95} =$ **2.131**	$2.132\frac{10}{\sqrt{16}} = 5.33$	50 ± 5.33	$44.67 < \mu < 55.33$ By 95% confidence, the average life of Diehard batteries is between 45 to 55 months	**Leave it blank**
b	16	df =15 $t_{.90} =$				**E** gets smaller, $\Rightarrow$ **narrower Interval**
c	16	df =15 $t_{.99} =$				
d	25	df =24 $t_{.95} =$			$44.67 < \mu < 55.33$	
e	9					**E** gets **larger**, $\Rightarrow$ **wider Interval**

Extra **P**ractice: Please do **p**roblems **1,4,5,7,8,9,36,37** from **p**ractice **p**roblems of **part III**

Estimating Population Proportion

Example: To estimate the percentage of people who believe in UFO a researcher took a random sample of **200** people across America and found out that there were **40** who did believe in UFOs. Use proper computation and, **(for z- value use web-table sheet page 3)**

a) Construct a **95%** confidence interval for the percentage of people believing in UFO.

b) Construct a **90%** confidence interval for the percentage of people believing in UFO.

c) Construct a **99%** confidence interval for the percentage of people believing in UFO.

d) Construct a **95%** confidence interval for the percentage of people believing in UFO, if **x =80, n = 400**.

e) Construct a **95%** confidence interval for the percentage of people believing in UFO, if **x =20, n = 100**.

Write your answers in percentage and round your final answers in 2 decimal places.

P = **Population or true Proportion** $\hat{p} = \frac{x}{n} = \frac{40}{200} = 0.2 = 20\%$ $E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $P = \hat{p} \pm E$

Observation: compared with part a, write how the **error** and **interval** get changed (smaller/ larger/ narrower/ wider).

	Z	n	$E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$P = \hat{p} \pm E$	< **P** <	**Observation** Compare all parts to part **(a),** draw conclusion about E , Interval
a	$z_{.95} =$ 1.96	200	$1.96\sqrt{\frac{.2(1-.2)}{200}} = .0554$	$20\% \pm 5.54\%$	$14.46\% < P < 25.54\%$ By 95% confidence, the percentage of people believing in UFO is between 14 to 26%	**Leave it blank**
b	$z_{.90} =$ 1.645	200				**E** gets **smaller**, $\Rightarrow$ **narrower Interval**
c	$z_{.99} =$ 2.58	200				**E** gets **larger**, $\Rightarrow$ **wider Interval**
d	$z_{.95} =$ 1.96	400			$16.08\% < P < 23.92\%$	**E** gets **smaller**, $\Rightarrow$ **narrower Interval**
e		100				**E** gets **larger**, $\Rightarrow$ **wider Interval**

Extra **P**ractice: Please do **p**roblems **1,2,3,12,13,17,19,20,21** from **p**ractice **p**roblems of **part III**

Estimating Difference between Two Populations Means $(\mu_1 - \mu_2)$

Confidence interval between two population **means** $(\mu_1 - \mu_2)$

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm E \qquad E = z\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

(for z- value use web-table sheet page 3)

Example 1) Find a 99% confidence interval between μ_1 and μ_2.

	Diehard μ_1	**Everlast** μ_2
N	$n_1 = 44$	$n_2 = 36$
$\bar{x}$	$\bar{x}_1 = 51.8$	$\bar{x}_2 = 47.4$
S	$s_1 = 2.5$	$s_2 = 3.7$

Point estimate $= (\bar{x}_1 - \bar{x}_2) = (51.8 - 47.4) = 4.4$ $\qquad E = z\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.58\sqrt{\frac{2.5^2}{44} + \frac{3.7^2}{36}} = 1.86$

$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm E = 4.4 \pm 1.86$ $\qquad 2.54 < \mu_1 - \mu_2 < 6.26$

By 99% confidence, Diehard batteries on the average last between 2.5 to 6.26 months longer than Everlast.

P. 1) Find a 90% confidence interval between the average life (in months) of "Diehard" and "Everlast" batteries.

	Diehard μ_1	**Everlast** μ_2
N	48	64
$\bar{x}$	52	48
S	10	8

Point estimate $= (\bar{x}_D - \bar{x}_E) = (52 - 48) = 4$ $\qquad E = z\sqrt{\frac{s_D^2}{n_D} + \frac{s_E^2}{n_E}} =$

$\mu_D - \mu_E = (\bar{x}_D - \bar{x}_E) \pm E = 4 \pm ?$

P. 2) Find a 99% confidence interval between the weights of Regular Coke and Regular Pepsi.

	Regular Coke μ_1	**Regular Pepsi** μ_2
N	$n_1 = 36$	$n_2 = 36$
$\bar{x}$	$\bar{x}_1 = 0.82410$	$\bar{x}_2 = 0.81682$
S	$s_1 = 0.007507$	$s_2 = 0.005701$

Extra **P**ractice: Please do **p**roblems **30,31, 33** from **p**ractice **p**roblems of part III

Estimating Difference between two Populations Proportions ($P_1 - P_2$)

Confidence interval between two population **proportions.** ($P_1 - P_2$)

$$P_1 - P_2 = (\hat{p}_1 - \hat{p}_2) \pm E \qquad E = Z\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

(for z- value use web-table sheet page 3)

Example. Find a 95% confidence interval between the proportions of female and male students who are passing stat class. $P_f - P_m = ?$

	Female	**Male**
n **= students**	$n_f =$ **100**	$n_m =$ **150**
x **= passing stat class**	$x_f =$ **73**	$x_m =$ **87**

Sample proportions for each one: $\hat{p}_f = \frac{73}{100} = .73$ $\qquad \hat{p}_m = \frac{87}{150} = .58$

Point estimate $(\hat{p}_f - \hat{p}_m) = .73 - .58 = .15 = 15\%$

$$E = Z\sqrt{\frac{\hat{p}_f(1-\hat{p}_f)}{n_f} + \frac{\hat{p}_m(1-\hat{p}_m)}{n_m}} = 1.96\sqrt{\frac{.73(1-.73)}{100} + \frac{.58(1-.58)}{150}} = 1.96\sqrt{.001971 + .001624} = .1175 = 11.75\%$$

$$P_f - P_m = (\hat{p}_f - \hat{p}_m) \pm E = 15\% \pm 11.75\% \qquad 3.25\% < P_f - P_m < 26.75\%$$

Estimated percentage of female students passing stat class is higher than male from 3.3 to 26.8%.

Practice 1. Find a 99% confidence interval between the proportions of women and men who lived more than 65 years in Alaska.

	Women	**Men**
n **= dead subject**	**240**	**400**
x **= lived longer than 65 years**	**144**	**192**

Sample proportions for each group. $\hat{p}_w = \frac{144}{240} =$ $\qquad \hat{p}_m = \frac{192}{400} =$

Point estimate $(\hat{p}_w - \hat{p}_m) =$

$$E = Z\sqrt{\frac{\hat{p}_w(1-\hat{p}_w)}{n_w} + \frac{\hat{p}_m(1-\hat{p}_m)}{n_m}} = 2.58\sqrt{\_\_\_\_ + \_\_\_\_}$$

Extra **P**ractice: Please do **p**roblems **32, 34** from **p**ractice **p**roblems of **part III**

Determining Sample Size for the Estimation of Mean (μ)

In practice, before estimating population mean, we need a sample, the question is how large of a sample? That depends on 3 factors, confidence level(z-value), margin of error(E), and prior information on standard deviation(s).

Use $n = \left(\frac{zs}{E}\right)^2$ when s is known from prior information **otherwise** use $s = \frac{Range}{4}$

(for z- value use web-table sheet page 3)

Example: How many Diehard batteries need to be sample to estimate by **95%** confidence level the mean life of its population? From a previous study, the standard deviation of batteries life was **6.2** months. The acceptable error for this estimation is to be within 2 months of the population mean?

$Z_{.95} = 1.96$ s = 6.2 months E = 2 months $n = \left(\frac{1.96 \bullet 6.2}{2}\right)^2 = 36.9 = 37$ batteries

Repeat the problems by changing

a) 95% confidence level to 99% confidence level
b) 95% confidence level to 90% confidence level
c) Error from 2 months to 1 month with 95 % confidence.
d) Error from 2 months to 4 months with 95 % confidence.

Observation: compared with part example, write how the sample size gets changed (smaller/ larger) and if possible by what proportion.

Case	Z	E	S	$n = \left(\frac{zs}{E}\right)^2$	**Observation** Compare all parts to example on this page and draw proper conclusion
a	$z_{.99} =$ 2.58				sample size gets larger
b	$z_{.90} =$ 1.645				
c	$z_{.95} =$ 1.96				sample size gets **larger** about 4 times
d	$z_{.95} =$ 1.96				

Extra **P**ractice: Please do **p**roblems **25,26,27,28** from **p**ractice **p**roblems of **part III**

Determining Sample Size for the Estimation of Proportion(P)

in practice, before estimating population proportion, we need a sample, the question is how large of a sample? That depends on 3 factors, confidence level(z-value), margin of error(E), and prior information on population proportion $\hat{p}$ *.*

Use $n = \left(\frac{Z}{E}\right)^2 \hat{p}(1-\hat{p})$ when $\hat{p}$ is known from prior information otherwise use $\hat{p} = 0.5$

(for z- value use web-table sheet page 3)

Example: How large should the sample size be if a researcher wants to estimate by **95%** confidence the percentage of people who will vote for the next democratic candidate? In the last election he got **76%** of the votes and the maximum error acceptable is **4%**

$Z_{.95} = 1.96$ $\hat{p} = .76$ E = 4%

$n = \left(\frac{1.96}{.04}\right)^2 .76(1-.76) = 437.9 = 438$ people needs to be sampled.

Repeat the problems by changing

a) **95%** confidence level to **99%** confidence level
b) **95%** confidence level to **90%** confidence level
c) Error from **.04 to .02** with 95% confidence level
d) Error from **.04 to .08** with 95% confidence level
e) $\hat{p}$ **is unknown** with 95% confidence level

Observation: compared with part a, write how the sample size gets changed (smaller/ larger) and if possible by what proportion.

Case	**Z**	**E**	$\hat{p}$	$n = \left(\frac{Z}{E}\right)^2 \hat{p}(1-\hat{p})$	**Observation** Compare all parts to example on this page and draw proper conclusion
a	$z_{.99} =$ 2.58	.04	0.76		sample size gets ***larger***
b	$z_{.90} =$ 1.645				
c	$z_{.95} =$ 1.96				
d	$z_{.95} =$ 1.96				sample size gets ***smaller*** by one fourth
e	$z_{.95} =$ 1.96				

Extra **P**ractice: Please do **p**roblems 22,23,24,29 from **p**ractice **p**roblems of **part III**

P4-b) E = 2.06 (90%) $\mu = 50 \pm 2.06$ $47.94 < \mu < 52.06$

P4-c) E = 3.23 (99%) $\mu = 50 \pm 3.23$ $46.77 < \mu < 53.23$

P4-d) E =1.96 (n=100) $\mu = 50 \pm 1.96$ $48.04 < \mu < 51.96$

P4-e) E = 3.27 (n=36) $\mu = 50 \pm 3.27$ $46.73 < \mu < 53.27$

P5-b) E = 4.38 (90%) $\mu = 50 \pm 4.38$ $45.62 < \mu < 54.38$

P5-c) E = 7.37 (99%) $\mu = 50 \pm 7.37$ $42.63 < \mu < 57.37$

P5-d) E = 4.13 (n = 25) $\mu = 50 \pm 4.13$ $45.87 < \mu < 54.13$

P5-e) E = 7.69 (n = 9) $\mu = 50 \pm 7.69$ $42.31 < \mu < 57.69$

P6-b) E =4.652 % P = 20 % ± 4.65 % 15.35% < P < 24.65%

P6-c) E = 7.297 % P = 20 % ± 7.30% 12.70% < P < 27.30%

P6-d) E = 3.92 % P = 20 % ± 3.92 % 16.08% < P < 23.92%

P6-e) E = 9.6 % P = 20 % ± 7.84 % 12.16% < P < 27.84%

P7) Diehard vs Everlast $E = 1.645\sqrt{\frac{10^2}{48} + \frac{8^2}{64}} = 2.89 \quad 1.11 < \mu_D - \mu_E < 6.89$

By 90% confidence, Diehard batteries on the average last between 1.11 to 6.89 months longer than Everlast.

P7) Regular Coke vs Regular Pepsi $0.00324 < \mu_1 - \mu_2 < 0.01133$

P8) $\hat{p}_w = \frac{144}{240} = .60$ $\hat{p}_m = \frac{192}{400} = .48$ **Point estimate** $\left(\hat{p}_w - \hat{p}_m\right) = (.60 - .48) = .12 = 12\%$

$$E = Z\sqrt{\frac{\hat{p}_w(1-\hat{p}_w)}{n_w} + \frac{\hat{p}_m(1-\hat{p}_m)}{n_m}} = 2.58\sqrt{\frac{.6(1-.6)}{240} + \frac{.48(1-.48)}{400}} = .104 = 10.4\%$$

$$\left(P_w - P_m\right) = \left(\hat{p}_w - \hat{p}_m\right) \pm E = 12\% \pm 10.4 \qquad 1.6\% < P_w - P_m < 22.4\%$$

P9-a) 99% E =2 n = 64 **P9-b)** 90% E = 2 n = 27

P9-c) 95% E = 1 n = 148 **P9-d)** 95% E = 4 n = 10

P8-a) 99% E = .04 n = 759 **P10-b**) 90% E = .04 n = 310

P10-c) 95% E = .02 n = 1752 **P10-d**) 95% E = .08 n = 110

P10-e) 95% E = .04 n = 601

Page 4 solutions

	Z	n	$E = z\frac{s}{\sqrt{n}}$	$\mu = \overline{x} \pm E$	< μ <	**Observation** Compare all parts to part **(a),** draw conclusion about E , Interval
b	$z_{.90} =$ 1.645	64	$1.645\frac{10}{\sqrt{64}} = 2.06$	50 ± 2.06	47.94 < μ < 52.06 By 90% confidence, the average life of **all** Diehard batteries is between 48 to 52 months	**E** gets **smaller**, ***narrower* Interval**
c	$z_{.99} =$ 2.58	64	$2.58\frac{10}{\sqrt{64}} = 3.22$	50 ± 3.22	46.78 < μ < 53.22 By 99% confidence, the average life of **all** Diehard batteries is between 47 to 53 months	**E** gets **larger**, **wider Interval**
d	$z_{.95} =$ 1.96	100	$1.96\frac{10}{\sqrt{100}} = 1.96$	50 ± 1.96	48.04 < μ < 51.96 By 95% confidence, the average life of **all** Diehard batteries is between 48 to 52 months	**E** gets **smaller**, ***narrower* Interval**
e	$z_{.95} =$ 1.96	36	$1.96\frac{10}{\sqrt{36}} = 3.27$	50 ± 3.27		**E** gets **larger**, **wider Interval**

Page 5 solutions

	n	**df=n-1 t = ?**	$E = t\frac{s}{\sqrt{n}}$	$\mu = \overline{x} \pm E$	< μ <	**Observation** Compare all parts to part **(a),** draw conclusion about E , Interval
b	16	df =15 $t_{.90} =$	$1.753\frac{10}{\sqrt{16}} = 4.38$	50 ± 4.38	$45.67 < \mu < 54.38$	
c	16	df =15 $t_{.99} =$	$2.947\frac{10}{\sqrt{16}} = 7.37$	50 ± 7.37	$42.63 < \mu < 57.37$	**E** gets **larger**, $\Rightarrow$ **wider Interval**
d	25	df =24 $t_{.95} =$	$2.064\frac{10}{\sqrt{25}} = 4.13$	50 ± 4.13	$45.67 < \mu < 54.13$	
e	9		$2.306\frac{10}{\sqrt{9}} = 7.69$	50 ± 7.69	$42.31 < \mu < 57.69$	**E** gets **larger**, $\Rightarrow$ **wider Interval**

Page 6 solutions

	Z	n	$E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$P = \hat{p} \pm E$	< **P** <	**Observation** Compare all parts to part **(a),** draw conclusion about E , Interval
b	$z_{.90} =$ 1.645	200	$1.645\sqrt{\frac{.2(1-.2)}{200}} = .0465$	$20\% \pm 4.65\%$	$15.35\% < P < 24.65\%$	**E** gets **smaller**, ⇒ **narrower Interval**
c	$z_{.99} =$ 2.58	200	$2.58\sqrt{\frac{.2(1-.2)}{200}} = .0730$	$20\% \pm 7.30\%$	$12.7\% < P < 27.3\%$	**E** gets **larger**, ⇒ **wider Interval**
d	$z_{.95} =$ 1.96	400	$1.96\sqrt{\frac{.2(1-.2)}{400}} = .0392$	$20\% \pm 3.92\%$	$16.08\% < P < 23.92\%$	**E** gets **smaller**, ⇒ **narrower Interval**
e		100	$1.96\sqrt{\frac{.2(1-.2)}{100}} = .0784$	$20\% \pm 7.84\%$	$12.16\% < P < 27.84\%$	**E** gets **larger**, ⇒ **wider Interval**

Page 7

	Diehard μ_1	**Everlast** μ_2
N	48	64
$\bar{x}$	52	48
S	10	8

Point estimate $= (\bar{x}_D - \bar{x}_E) = (52-48) = 4$

$$E = z\sqrt{\frac{s_D^2}{n_D} + \frac{s_E^2}{n_E}} = 1.645\sqrt{\frac{10^2}{48} + \frac{8^2}{64}} = 2.89$$

$$\mu_D - \mu_E = (\bar{x}_D - \bar{x}_E) \pm E = 4 \pm 2.89$$

$$1.11 < \mu_D - \mu_E < 6.89$$

Page 8

Sample proportions for each group. $\hat{p}_w = \frac{144}{240} = 0.6 = 60\%$ $\hat{p}_m = \frac{192}{400} = .48 = 48\%$

Point estimate $\left(\hat{p}_w - \hat{p}_m\right) = .6 - .48 = 60\% - 48\% = 12\%$

$$E = Z\sqrt{\frac{\hat{p}_w(1-\hat{p}_w)}{n_w} + \frac{\hat{p}_m(1-\hat{p}_m)}{n_m}} = 2.58\sqrt{\frac{.6(1-.6)}{240} + \frac{.48(1-.48)}{400}} = 0.104 = 10.4\%$$

$$\left(P_w - P_m\right) = \left(\hat{p}_w - \hat{p}_m\right) \pm E = 12\% \pm 10.2\% \qquad 1.8\% < P_w - P_m < 22.2\%$$

We are 99% confident that proportions of women who lived more than 65 years in Alaska exceeds men between 1.8 to 22.2%.

Page 9 solutions

Case	Z	E	S	$n = \left(\frac{zs}{E}\right)^2$	**Observation** Compare all parts to example on this page and draw proper conclusion
a	$z_{.99} =$ 2.58		6.2	$n = \left(\frac{2.58 \bullet 6.2}{2}\right)^2 = 64$	sample size gets larger
b	$z_{.90} =$ 1.645		6.2	$n = \left(\frac{1.645 \bullet 6.2}{2}\right)^2 = 27$	sample size gets smaller
c	$z_{.95} =$ 1.96		6.2	$n = \left(\frac{1.96 \bullet 6.2}{4}\right)^2 = 148$	sample size gets **larger** about 4 times
d	$z_{.95} =$ 1.96		6.2	$n = \left(\frac{1.96 \bullet 6.2}{8}\right)^2 = 10$	sample size gets smaller about one fourth

---

Page 10 solutions

Case	Z	E	$\hat{p}$	$n = \left(\frac{Z}{E}\right)^2 \hat{p}(1-\hat{p})$	**Observation** Compare all parts to example on this page and draw proper conclusion
a	$z_{.99} =$ 2.58	.04	0.76	$n = \left(\frac{2.58}{.04}\right)^2 .76(1-.76) = 758.83 = 759$	sample size gets ***larger***
b	$z_{.90} =$ 1.645	.04	0.76	$n = \left(\frac{1.645}{.04}\right)^2 .76(1-.76) = 308.48 = 309$	sample size gets ***smaller***
c	$z_{.95} =$ 1.96	.02	.76	$n = \left(\frac{1.96}{.02}\right)^2 .76(1-.76) = 1751.77 = 1752$	sample size gets 4 times ***larger***
d	$z_{.95} =$ 1.96	.08	.76	$n = \left(\frac{1.96}{.08}\right)^2 .76(1-.76) = 109.49 = 110$	sample size gets ***smaller*** by one fourth
e	$z_{.95} =$ 1.96	.04	.50	$n = \left(\frac{1.96}{.04}\right)^2 .50(1-.50) = 600.25 = 601$	sample size gets ***larger***