

7 Statistical Inference: Confidence Intervals

7.1 What Are Point and Interval Estimates of Population Parameters?

In this chapter we discuss **point estimation** and **interval estimation**.

Definition: A **point estimate** is a *single number* (based on the data), used to estimate a population parameter.

Definition: An **interval estimate** is an *interval of numbers* (based on the data), used to estimate a population parameter.

We focus on estimating a **population proportion**, p , and a **population mean**, μ .

In this chapter, **p and μ are unknown.**

Point Estimation

What is a reasonable **point estimate** of μ ?

A desirable property of a point estimator is

unbiasedness; i.e., the mean of the point estimator is the population parameter.

For example, the mean of \bar{X} is μ , as discussed in chapter 6.

The tendency to overestimate μ is the same as the tendency to underestimate μ when using \bar{X} .

Example: Suppose that the survival period of terminally ill cancer patients beginning a new therapy is sampled for 10 patients.

Suppose the survival times in years for the 10 patients are 3.2, 5.6, 7.3, 1.3, 0.4, 2.6, 4.2, 6.4, 3.5, 3.9.

- (a) Estimate the **mean** survival time, μ , for the entire population of terminally ill cancer patients beginning this new therapy.
- (b) Estimate the **median** survival time for the entire population of terminally ill cancer patients beginning this new therapy.

□

Which estimator for center should one use, if the population is symmetric?

In general, if two different estimators are both **unbiased**, then the preferred one is the one with the smaller **variability** or **standard error**.

Recall: $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ and $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$, exactly for independent observations, and approximately for *nearly* independent observations.

Example: Discuss *bias* and *standard error* in the following *sampling distributions*, when estimating

μ .

Example: *Revisit cancer.* Suppose we wish to estimate the population proportion, p , of terminally ill cancer patients (beginning the new therapy) who will survive at least 6 more years.

□

Interval Estimation

Name the error associated with our point estimates \bar{X} and \hat{p} , when estimating μ and p , respectively?

A confidence interval on μ is $\bar{X} \pm$ (margin of error).

A confidence interval on p is $\hat{p} \pm$ (margin of error).

Example: A news organization reports a simple random **sample** (not a **census**) where the Democrat is defeating the Republican by a vote of 52% to 48% with a **margin of error** of 3%. Is it reasonable to conclude that the Democrat is winning, or is the election *too close to call*?

□

The **margin of error** for \hat{p} is typically

$$1.96 \times (\text{standard error}) = 1.96 \hat{\sigma}_{\hat{p}} = 1.96 \sqrt{\hat{p}(1 - \hat{p})/n}.$$

In chapter 4, the **margin of error** was crudely approximated to be $1/\sqrt{n}$.

Many more mathematical details follow in the remainder of this chapter.

7.2 How Can We Construct a Confidence Interval to Estimate a Population Proportion, p ?

Example: First, find the z -score such that $P(-z < Z < z) = 0.95$, where $Z \sim N(0, 1)$.

Standard normal table, pp. A1–A2										
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Standard normal table, pp. A1–A2										
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

t -table, p. A3							
		Confidence Level					
		80%	90%	95%	98%	99%	99.8%
		Right-Tail Probability					
df		$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
⋮		⋮	⋮	⋮	⋮	⋮	⋮
50		1.299	1.676	2.009	2.403	2.678	3.261
60		1.296	1.671	2.000	2.390	2.660	3.232
80		1.292	1.664	1.990	2.374	2.639	3.195
100		1.290	1.660	1.984	2.364	2.626	3.174
∞		1.282	1.645	1.960	2.326	2.576	3.090

□

Recall: $\hat{p} \stackrel{\text{approx.}}{\sim} N(\mu_{\hat{p}} = p, \sigma_{\hat{p}} = \sqrt{p(1-p)/n})$ if $np \geq 15$ and $n(1-p) \geq 15$.

Derivation of a 95% confidence

interval on p : *(You do NOT need to reproduce this derivation.)* For large enough sample sizes,

$$P(\mu_{\hat{p}} - 1.96\sigma_{\hat{p}} < \hat{p} < \mu_{\hat{p}} + 1.96\sigma_{\hat{p}}) \approx 0.95$$

$$P(p - 1.96\sqrt{p(1-p)/n} < \hat{p} < p + 1.96\sqrt{p(1-p)/n}) \approx 0.95$$

Solving for p ,

$$P(\hat{p} - 1.96\sqrt{p(1-p)/n} < p < \hat{p} + 1.96\sqrt{p(1-p)/n}) \approx 0.95$$

Since p is unknown, we write

$$P(\hat{p} - 1.96\sqrt{\hat{p}(1-\hat{p})/n} < p < \hat{p} + 1.96\sqrt{\hat{p}(1-\hat{p})/n}) \approx 0.95$$

A 95% confidence interval on p is $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$.

Note: This is a large sample approximation, in that we insist that $n\hat{p} \geq 15$ and $n(1-\hat{p}) \geq 15$. \square

Recall, for a sample proportion, \hat{p} :

(a) (standard error) = $\sigma_{\hat{p}} = \sqrt{p(1-p)/n} \approx \sqrt{\hat{p}(1-\hat{p})/n}$.

(b) For 95% confidence, the (margin of error) = $z \times$
(standard error) $\approx 1.96\sqrt{\hat{p}(1-\hat{p})/n}$.

(c) For $n\hat{p} \geq 15$ and $n(1-\hat{p}) \geq 15$, the 95% confidence interval on *unknown, fixed* p is

$$\hat{p} \pm (\text{margin of error}) = \hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}.$$

Layman’s interpretation: We are 95% confident that the population proportion, p , lies in the confidence interval.

Mathematically rigorous interpretation: If we repeat the sampling procedure many times to construct many 95% confidence intervals on p , then approximately 95% of these 95% confidence intervals will contain the true value of p .

Example: *Estimating the success rate at the Charlottesville fertility clinic, called University of Virginia Assisted Reproductive Technology (ART)*

program, in year 2001.

64 women no older than 40 years-old attempted to get pregnant from using services at the UVa clinic in 2001.

Do these 64 women represent a simple random sample of women from the U.S.?

The population consists of all women no older than 40, from similar regions in 2001, who would seek clinical pregnancy services from this type of clinic.

Among those 64 women, 20 successfully gave live births (i.e., no miscarriages).

We want to estimate p , the population proportion of *similar* women who would give live births when using this clinic.

Hence, p is the population success rate of this clinic.

$X = 20$, the number of women who successfully gave live births.

(a) Determine the appropriate *point estimate* of p ,

the population success rate of this clinic.

- (b) Construct a **95%** confidence interval on p , the population success rate of this clinic.

Layman's interpretation: We are 95% confident that the population success rate of this clinic lies between 0.199 and 0.426.

Mathematically rigorous interpretation: If we repeat the sampling procedure many times to construct many 95% confidence intervals on p , the population success rate of this clinic, then approximately 95% of these 95% confidence intervals will contain p .

- (c) Which of the following are correct interpretations?

○ $P(0.199 < p < 0.426) \approx 0.95$

○ $P(0.199 < \hat{p} < 0.426) \approx 0.95$

- (d) Now suppose that we want a **99%** confidence interval on p .

t-table, p. A3							
		Confidence Level					
		80%	90%	95%	98%	99%	99.8%
		Right-Tail Probability					
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	
50	1.299	1.676	2.009	2.403	2.678	3.261	
60	1.296	1.671	2.000	2.390	2.660	3.232	
80	1.292	1.664	1.990	2.374	2.639	3.195	
100	1.290	1.660	1.984	2.364	2.626	3.174	
∞	1.282	1.645	1.960	2.326	2.576	3.090	

- (e) Which confidence interval is wider?
- (f) How can we increase the **level** of confidence without increasing the **width** of confidence interval?

□

Example: Do you prefer a **high** level (e.g., 99.9% level) of confidence or a **low** level (e.g., 50% level) of confidence in the following? You work for a bomb squad. A *red* wire and a *blue* wire are remaining. Cutting the *correct* wire results in *life*, but cutting the *wrong* wire results in *death*. Your partner says,

“Cut the *red* wire.” You respond, “How confident are you?”

□

Example: Do you prefer a **wide** confidence interval or a **narrow** confidence interval in the following?

The Joint United Nations Programme on HIV/AIDS (UNAIDS) is 95% confident that the population proportion of adults aged 15 to 49 from Botswana who are infected with HIV is between 23% and 32%.

I am almost 100% confident that the population proportion of adults aged 15 to 49 from Botswana who are infected with HIV is between 0.001% and 99.999%.

□

What is the optimal confidence level; e.g., 90%, 95%, or 99%?

7.3 How Can We Construct a Confidence Interval to Estimate a

Population Mean, μ ?

The t distribution

Case *A*: Sample **with** replacement. *Hence, observations are independent.*

Case *B*: Sample **without** replacement, but the population size is quite large compared to n . *Hence, observations are nearly independent.*

(a) $\mu_{\bar{X}} = \mu$ *always.*

(b) $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ (called the **standard error** of \bar{X}), exactly for Case *A* and approximately for Case *B*.

(c) (A version of the Central Limit Theorem) The sample mean, \bar{X} , is approximately normally distributed for Cases *A* and *B* (and finite σ), for **large** n (usually $n \geq 30$, if neither tail of the distribution is too heavy).

(d) (A special case) The sample mean, \bar{X} , is approx-

imately normally distributed for Cases *A* and *B* (and finite σ , for **any** sample size n), if the **original population** is approximately **normally distributed**.

Therefore, for independent or nearly independent observations (and finite σ), if **the original population is approximately normal OR n is large**, then

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \underset{\text{approx.}}{\sim} N(0, 1), \text{ and}$$
$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \underset{\text{approx.}}{\sim} t_{n-1}$$

Thus, T has a t distribution with $(n - 1)$ degrees of freedom.

The t distribution is symmetric about zero, has no units, and has heavier tails than the standard normal distribution.

As the degrees of freedom gets large, then s “gets close

to" σ , so the t distribution starts to converge to the standard normal distribution.

Example: Below are the *probability density functions* of a t distribution with *one* degree of freedom, a t distribution with *four* degrees of freedom, and the *standard normal distribution*.

□

Confidence interval on μ

For independent or nearly independent observations (and finite σ), if **the original population is approximately normal OR n is large**, then a confidence interval on μ is

$$\begin{aligned}\bar{X} \pm (\text{margin of error}) &= \bar{X} \pm t_{n-1}(\text{standard error}) = \\ &= \bar{X} \pm t_{n-1}s/\sqrt{n}.\end{aligned}$$

Example: A sample of individuals participating in a rigorous exercise program results in the following weight losses in pounds: $\{16, 6, 24, -3, 12\}$.

The population consists of all *similar* individuals who would be willing to participate in this rigorous exercise program, if offered the opportunity.

- (a) Are the assumptions for constructing a confidence interval satisfied?
- (b) Construct a **95%** confidence interval on the population mean weight loss.

t-table, p. A3

		Confidence Level					
		80%	90%	95%	98%	99%	99.8%
		Right-Tail Probability					
<i>df</i>		$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
1		3.078	6.314	12.706	31.821	63.657	318.309
2		1.886	2.920	4.303	6.965	9.925	22.327
3		1.638	2.353	3.182	4.541	5.841	10.215
4		1.533	2.132	2.776	3.747	4.604	7.173
5		1.476	2.015	2.571	3.365	4.032	5.893
⋮		⋮	⋮	⋮	⋮	⋮	⋮

Layman’s interpretation: We are 95% confident that the population mean weight loss, μ , of this exercise program is between -1.66 pounds and 23.66 pounds.

Mathematically rigorous interpretation: If we repeat the sampling procedure many times to construct many 95% confidence intervals on μ , the population mean weight loss of this exercise program, then approximately 95% of these 95% confidence intervals will contain the true value of μ .

(c) Which of the following are correct interpretations?

⊙ $P(-1.66 \text{ pounds} < \mu < 23.66 \text{ pounds}) \approx 0.95$

⊙ $P(-1.66 \text{ pounds} < \bar{X} < 23.66 \text{ pounds}) \approx 0.95$

⊙ 95% of the population of weight losses lies between -1.66 pounds and 23.66 pounds.

⊙ $P(-1.66 \text{ pounds} < X < 23.66 \text{ pounds}) \approx 0.95$

(d) Construct a **90%** confidence interval on the population mean weight loss.

<i>t</i> -table, p. A3							
		Confidence Level					
		80%	90%	95%	98%	99%	99.8%
		Right-Tail Probability					
<i>df</i>	<i>t</i> _{.100}	<i>t</i> _{.050}	<i>t</i> _{.025}	<i>t</i> _{.010}	<i>t</i> _{.005}	<i>t</i> _{.001}	
1	3.078	6.314	12.706	31.821	63.657	318.309	
2	1.886	2.920	4.303	6.965	9.925	22.327	
3	1.638	2.353	3.182	4.541	5.841	10.215	
4	1.533	2.132	2.776	3.747	4.604	7.173	
5	1.476	2.015	2.571	3.365	4.032	5.893	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	

Layman’s interpretation: We are 90% confident that the population mean weight loss, μ , of this exercise program is between 1.28 pounds and 20.72 pounds.

Mathematically rigorous interpretation: If we repeat the sampling procedure many times to construct many 90% confidence intervals on μ , the population mean weight loss of this exercise program, then approximately 90% of these 90% confidence intervals will contain the true value of μ .

(e) Which confidence interval is wider?

□

Example: In a simple random sample from a large population, the following observations were taken: $\{45, 310, 93, 63, 81, 270, 57\}$. Construct a **95%** confidence interval on the population mean.

□

7.4 How Do We Choose the Sample Size for a Study?

Population proportion, p

Recall: For $n\hat{p} \geq 15$ and $n(1 - \hat{p}) \geq 15$, a

confidence interval on p , the unknown population proportion, is

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

The **margin of error** on \hat{p} is $m = z\sqrt{\hat{p}(1 - \hat{p})/n}$, which is **half the width** of the confidence interval.

Suppose we want to construct a 95% confidence interval on p , where the **margin of error**, m , is selected prior to drawing the sample.

<i>t</i> -table, p. A3						
Confidence Level						
	80%	90%	95%	98%	99%	99.8%
Right-Tail Probability						
<i>df</i>	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.090

What sample size, n , is needed?

Solve for n in

$$m = 1.96 \sqrt{\hat{p}(1 - \hat{p})/n}$$

to obtain

$$n = \hat{p}(1 - \hat{p})(1.96/m)^2.$$

What is the drawback when using the above formula for n ?

Two options:

(a) Use a preliminary *point estimate* \hat{p} , and then compute

$$n = \hat{p}(1 - \hat{p})(1.96/m)^2, \quad \text{OR}$$

(b) The maximum value of $n = \hat{p}(1 - \hat{p})(1.96/m)^2$ occurs when $\hat{p} = 0.5$, so use

$$n = 0.25(1.96/m)^2 \quad (\text{conservative sample size}).$$

Example: *Revisit the Charlottesville fertility clinic.* A sample of 64 women resulted in 20 live births. However, the population success rate, p , of this clinic is unknown. What sample size n is needed to obtain a **95%** confidence interval on p with **margin of error** approximately equal to **0.06**, using:

(a) 0.3125, as the initial *point estimate* of p ?

(b) no initial *point estimate* of p ?

Example: *Revisit the Charlottesville fertility*

clinic, again! What sample size n is needed to obtain a **90%** confidence interval on p with **margin of error** approximately equal to **0.06**, using:

- (a) 0.3125, as the initial *point estimate* of p ?
- (b) no initial *point estimate* of p ?
- (c) Repeat part (b) using $m = 0.03$.

Population mean, μ

Recall: For independent or nearly independent observations (and finite σ), if **the original population is approximately normal OR n is large**, then a **confidence interval on μ** , the unknown population mean, is

$$\bar{X} \pm t_{n-1} s / \sqrt{n}.$$

The **margin of error** on \bar{X} is $m = t_{n-1} s / \sqrt{n}$, which is **half the width** of the confidence interval.

Suppose we want to construct a 95% confidence interval on μ , where the **margin of error**, m , is selected

prior to drawing the sample.

What sample size, n , is needed?

Solve for n in

$$m = t_{n-1} s / \sqrt{n}$$

to obtain

$$n = (t_{n-1} s / m)^2.$$

For large n , what is t_{n-1} (approximately)?

What is the drawback when using the above formula for n ?

Example: Based on a sample of 41 personal incomes, $\bar{X} = \$43,000$ and $s = \$30,000$. Let μ be the unknown population mean income.

- (a) What sample size n is needed to obtain a **95%** confidence interval on μ with **margin of error** approximately equal to **\$2,000**?
- (b) What sample size n is needed to obtain a **95%**

confidence interval on μ with **margin of error** approximately equal to **\$1,000**?

- (c) What sample size n is needed to obtain a **99%** confidence interval on μ with **margin of error** approximately equal to **\$1,000**?

Remark: Decreasing the **margin of error** by half results in quadrupling the required sample size, for a fixed level of confidence.

Remark: Increasing the level of confidence for fixed m requires a larger sample size.