7 Confidence Intervals

Definition: Statistical inference provides methods for drawing conclusions about a population from sampled data.

In this chapter we discuss **point estimation** and **confidence intervals.**

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**, $\boldsymbol{p}$, and a **population mean**, $\boldsymbol{\mu}$.

In this chapter, $\boldsymbol{p}$ **and** $\boldsymbol{\mu}$ **are unknown,** and we assume that the observations are independent or *nearly* independent.

Point Estimation of both μ and p

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 (additional) 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.

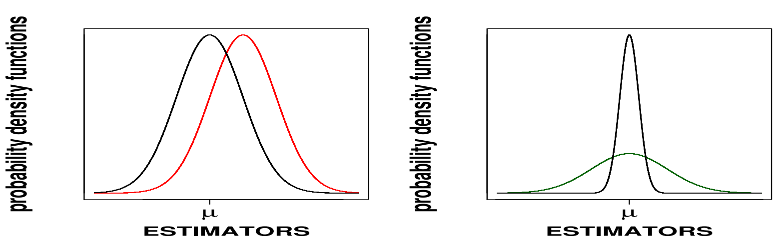
□

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.

□

7.3 Confidence Intervals for a Population Proportion, p

Let p = unknown population proportion.

Let $\hat{p}$ = sample proportion.

We make inferences on p using the point estimate $\hat{p}$.

We use large samples and apply the Central Limit Theorem; i.e., $\hat{p}$ is approximately normal for large n, for independent or nearly independent observations.

Confidence Interval on p

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

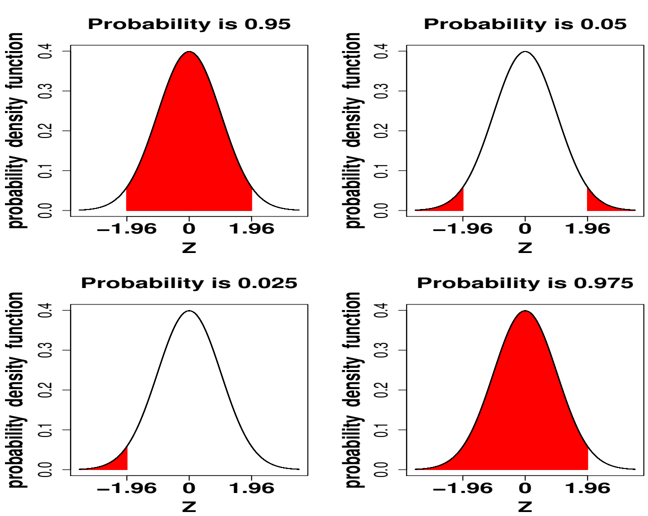
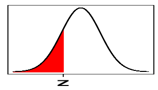
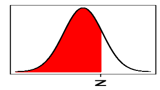
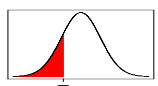
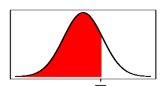


Table A.2 Cumulative Normal Distribution, pp. A-6 and A-7

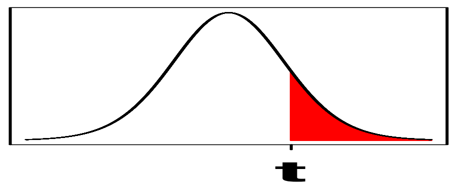
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table A.2 Cumulative Normal Distribution, pp. A-6 and A-7

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
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
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table A.3 Critical Values for the Student's t Distribution, p. A-8



Degrees of Freedom	Area in Right Tail									
	0.40	**0.25**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	0.254	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
z	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
	20%	50%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%
	Confidence Level									

□

Recall: $\hat{p} \overset{approx.}{\sim} N(\mu_{\hat{p}} = p, \sigma_{\hat{p}} = \sqrt{p(1-p)/n})$ if $np \geq 10$ and $n(1-p) \geq 10$.

Derivation of a 95% confidence interval on p: *(You do NOT need to reproduce this derivation.)* For large enough sample sizes, and for independent or nearly independent observations,

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

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

Solving for p,

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

Since p is unknown, we write

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

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 the numbers of successes and failures both are at least 10. □

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

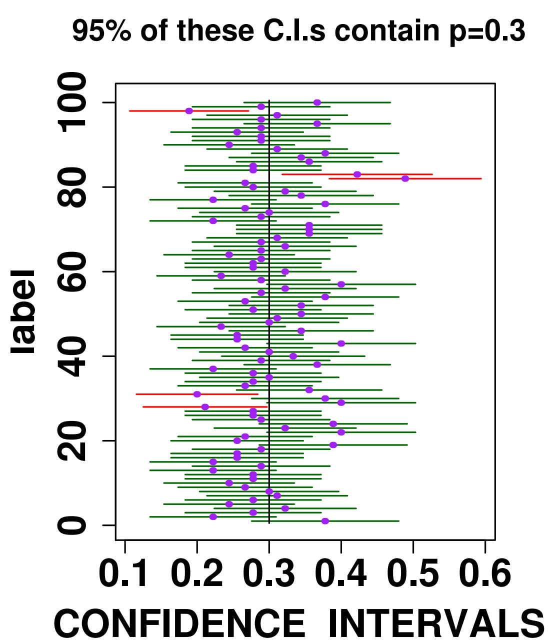
(a) (standard error) $= \sigma_{\hat{p}} = \sqrt{p(1-p)/n} \approx \sqrt{\hat{p}(1-\hat{p})/n}$.

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

(c) For independent or nearly independent observations, if the sample has at least 10 successes and at least 10 failures, then the 95% confidence interval on *unknown, fixed* p is $\hat{p} \pm$ (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.*

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

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 and with similar health conditions, 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) Now suppose that we want a **99%** confidence interval on p.

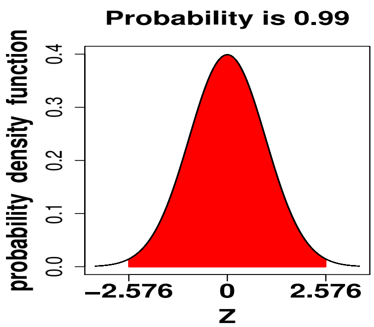
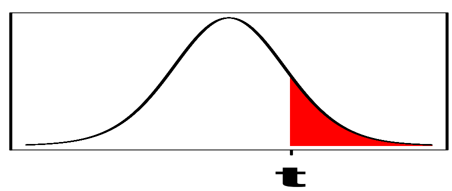
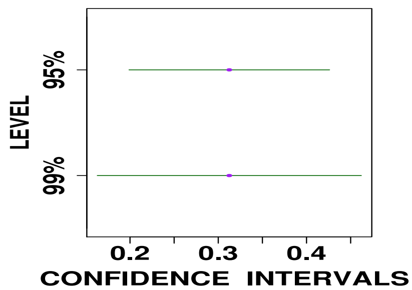


Table A.3 Critical Values for the Student's t Distribution, p. A-8



Degrees of Freedom	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Area in Right Tail									
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	0.254	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
z	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
	20%	50%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%
	Confidence Level									

(d) Which confidence interval is narrower?



(e) How can we increase the **level** of confidence without increasing the **width** of confidence interval?

□

Example: Suppose you borrow a lot of money from your parents to attend medical school. Your parents ask you, "How confident are you that you will graduate from medical school?"

□

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 people aged 15 to 49 from Botswana who are infected with HIV is between 23.9% and 26.3%.

I am almost 100% confident that the population proportion of people 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%?

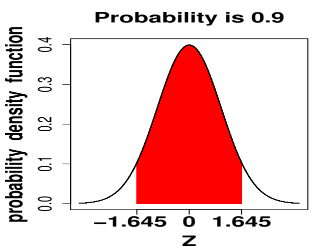
Sample Size Determination

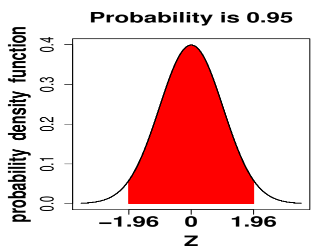
Recall: For at least 10 successes and at least 10 failures, a **confidence interval on *p***, the unknown population proportion, is

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

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.





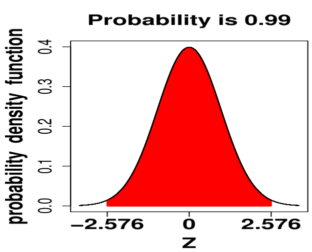
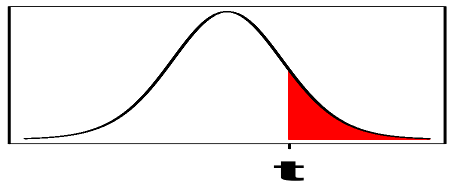


Table A.3 Critical Values for the Student's t Distribution, p. A-8



Degrees of Freedom	Area in Right Tail 0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	0.254	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
z	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
	20%	50%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%
	Confidence Level									

Solve for n in

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

to obtain

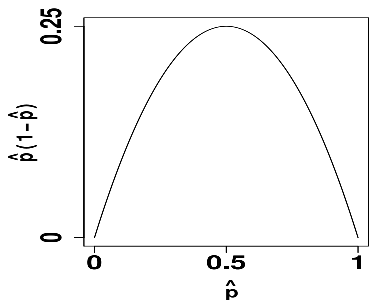
$$n = \hat{p}(1-\hat{p})(1.96/m)^2.$$

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

Two options:

(a) Use a preliminary *point estimate* $\hat{p}$, and then compute
$n = \hat{p}(1-\hat{p})(1.96/m)^2$, OR

(b) The maximum value of $n = \hat{p}(1-\hat{p})(1.96/m)^2$ occurs when $\hat{p} = 0.5$, so use
$n = 0.25(1.96/m)^2$ (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 $m = 0.03$.

7.1 Confidence Intervals for a Population

Mean μ, Standard Deviation σ Known

Scenario: μ is unknown, but σ is known.

What is needed in order for σ to be **known**?

We skip the textbook's formulas regarding Z, since such formulas require σ to be known when μ is unknown (and this is unrealistic).

7.2 Confidence Intervals for a Population Mean μ, Standard Deviation σ Unknown

Let μ = unknown population mean.

Let $\bar{X}$ = sample mean.

We make inferences on μ using the point estimate $\bar{X}$.

We use large samples and apply the Central Limit Theorem; i.e., $\bar{X}$ is approximately normal for large n, for independent or nearly independent observations. Alternatively, we start with an approximately normal population, in which case $\bar{X}$ is approximately normal for **any** n.

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; i.e., $N \geq 20n$. *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 positive finite σ), for **large** n (usually $n > 30$, if neither tail of the distribution is too heavy).

(d) (A special case) The sample mean, $\bar{X}$, is approximately normally distributed for Cases A and B (and positive finite σ, for **any** sample size n), if the **original population** is approximately **normally distributed**.

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

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

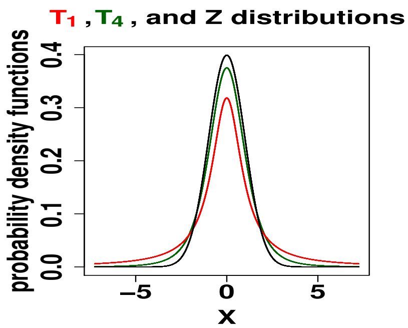
$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \overset{approx.}{\sim} t_{n-1}$$

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 "converges" 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.*

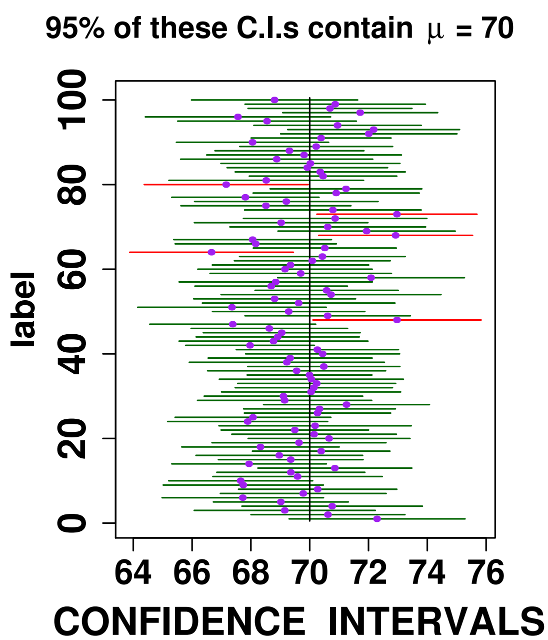


□

Confidence Interval on μ

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

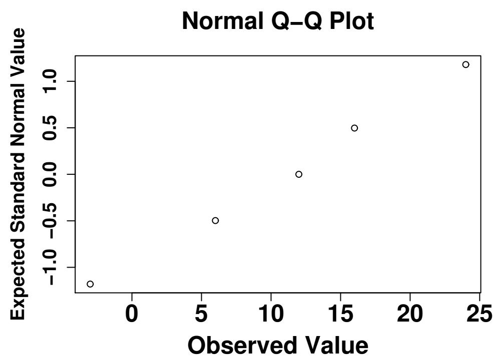
$\bar{X} \pm (\text{margin of error}) = \bar{X} \pm t_{n-1}(\text{standard error}) = \bar{X} \pm t_{n-1}s/\sqrt{n}$.



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.

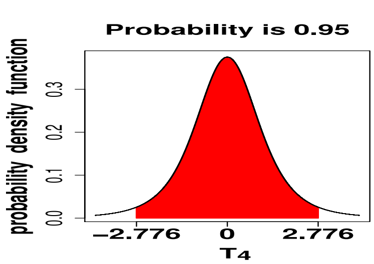
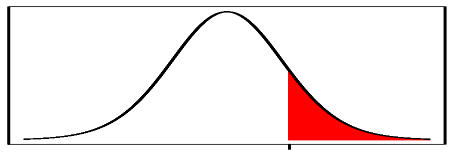


Table A.3 Critical Values for the Student's t Distribution, p. A-8



Degrees of Freedom	Area in Right Tail									
	0.40	**0.25**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
z	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
	20%	50%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%
	Confidence Level									

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

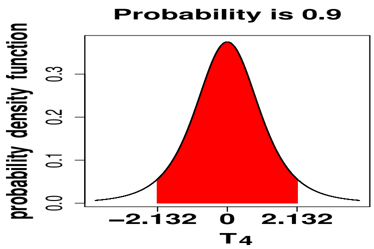
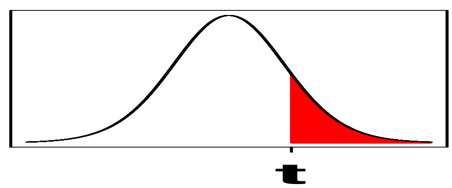


Table A.3 Critical Values for the Student's t Distribution, p. A-8

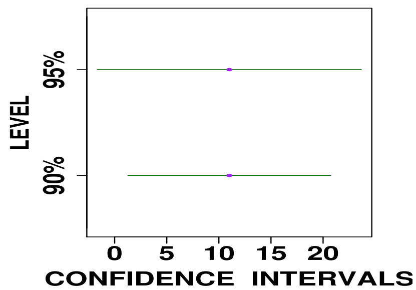


Degrees of Freedom	Area in Right Tail 0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
z	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
Confidence Level	20%	50%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%

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.

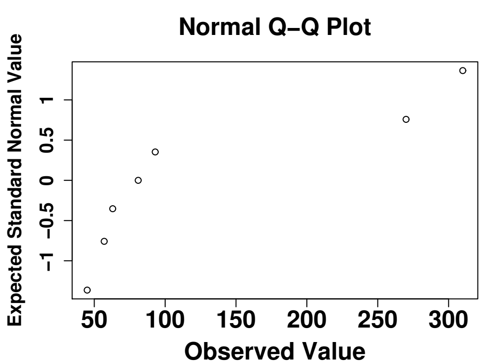
Mathematically rigorous interpretation: If we repeat the sampling procedure many times to construct many 90% confidence intervals on μ, the population mean weight loss of this exercise program, then approximately 90% of these 90% confidence intervals will contain the true value of μ.

(d) Which confidence interval is narrower?



□

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.



□

Sample Size Determination

Recall: For independent or nearly independent observations (and positive finite σ), if **the original population is approximately normal OR n is large**, then a

confidence interval on $\boldsymbol{\mu}$, 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

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

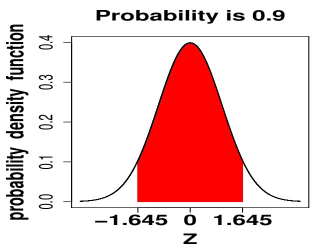
to obtain

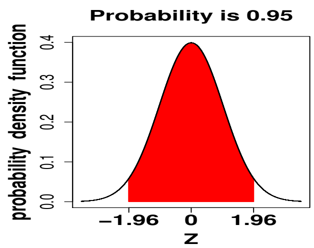
$$n = (t_{n-1}\ s/m)^2.$$

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.





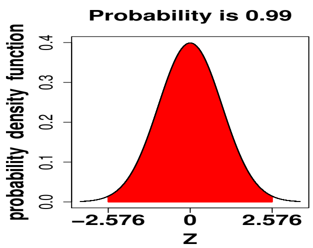
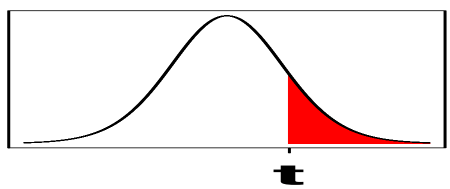


Table A.3 Critical Values for the Student's t Distribution, p. A-8



Degrees of Freedom	Area in Right Tail 0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	0.254	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
z	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
	20%	50%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%
	Confidence Level									

(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.