10 Two-Sample Tests

●● We may wish to compare two groups in an experiment.

Example: Which of two drugs is better?

●● We may wish to compare two populations in a survey.

Example: Compare the heights of females in Mexico vs. the United States, or the likelihood of developing cancer between smokers and nonsmokers.

□

When comparing two groups or two populations, we may use

(a) **Independent samples** (sections 10.1 and 10.3), OR

(b) **Dependent samples (matched pairs; two related populations) – IDEAL** (section 10.2).

Example: *Matched pairs.*

□

When **matched pairs** are not possible, use **independent samples**.

Example: *Independent samples.*

10.1 Comparing the Means of Two Independent Populations

Pooled-Variance *t* Test for the Difference Between Two Means

This topic assumes that $\sigma_1 = \sigma_2$, which is an unnecessary assumption, so we skip this topic.

Separate-Variance t Test for the Difference Between Two Means

This topic does NOT assume that $\sigma_1 = \sigma_2$, so we will address this topic.

In this section, we focus on **independent observations**, not **matched pairs**.

Construct independent t-test and independent t-confidence interval.

Population #1: Take independent or nearly independent observations from a population with mean μ_1 and positive finite standard deviation σ_1.

Let $\bar{X}_1$ be the sample mean and s_1 be the sample standard deviation, based on a sample of size n_1.

Population #2: Take independent or nearly independent observations from a population with mean μ_2 and positive finite standard deviation σ_2.

Let $\bar{X}_2$ be the sample mean and s_2 be the sample standard deviation, based on a sample of size n_2.

Assume that the two samples are independent of each other.

Question: Is $\mu_1 = \mu_2$, OR is $\mu_1 - \mu_2 = 0$?

Estimate: $(\mu_1 - \mu_2)$

What is the **point estimate** of $(\mu_1 - \mu_2)$?

What is the mean of $(\bar{X}_1 - \bar{X}_2)$?

It can be shown that since the samples are independent or nearly independent, then

$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$.

For the rest of this section, assume that all observations in the samples are **independent** or **nearly independent**, and both σ_1 and σ_2 are positive and finite.

If n_1 and n_2 are both large (usually $n_1 \geq 30$ and $n_2 \geq 30$, if none of the tails of the two distributions are too heavy), or if the two populations are approximately normal, then

$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$ **NOT PRACTICAL for inference**

is approximately standard normal, and

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$ **PRACTICAL**

is approximately t distributed, so a **confidence interval** on $(\mu_1 - \mu_2)$ is

$$\bar{X}_1 - \bar{X}_2 \pm t \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Degrees of freedom: When s_1 and s_2 are similar and n_1 and n_2 are close, then the **degrees of freedom** is close to $(\boldsymbol{n_1 + n_2 - 2})$. Otherwise, the **degrees of freedom** can be approximated conservatively by the **smaller** of $(\boldsymbol{n_1 - 1})$ and $(\boldsymbol{n_2 - 1})$.

Listed in your textbook is a very ugly but more accurate formula (10.4) for degrees of freedom, so we simply will use the above approximation.

How can we verify the normality assumption?

Again, the t procedures are **robust**.

Example: A study of zinc-deficient mothers was conducted to determine whether zinc supplementation during pregnancy results in babies with increased mean weight at birth. {*Data are available at* Goldenberg et al., *JAMA* 1995 (August 9); 274 (6): 463-468.}

Treatment #1 Zinc supplement group	Treatment #2 Placebo group
$n_1 = 294$	$n_2 = 286$
$\bar{X}_1 = 3214$ g	$\bar{X}_2 = 3088$ g
$s_1 = 669$ g	$s_2 = 728$ g

Is there sufficient evidence to support the claim that zinc supplementation results in increased mean birth weight, in comparison to a placebo? Test at level $\alpha = 0.05$.

(a) Do we need to assume that the two populations for birth weight are approximately normally distributed?

(b) Define your notation.

Let $\boldsymbol{\mu_1}$= *unknown* **population** mean birth weight in the **zinc**-supplemented group.

Let $\boldsymbol{\mu_2}$= *unknown* **population** mean birth weight in the **placebo** group.

(c) State the hypotheses.

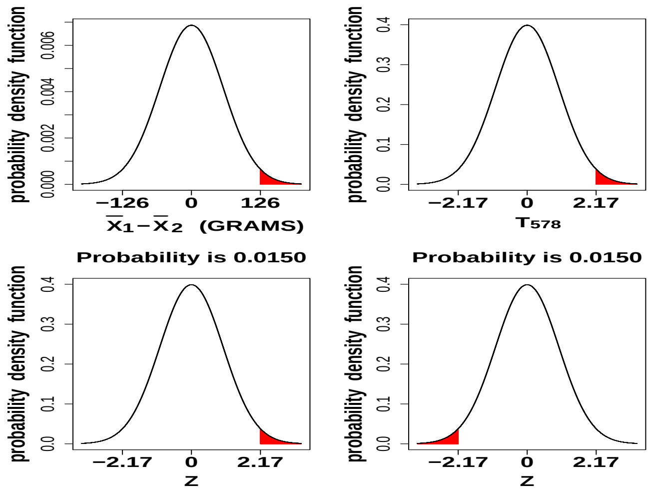
(d) Determine the value of the **standardized test statistic.**

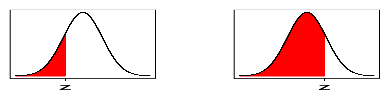
Let $\boldsymbol{\bar{X}_1}$= **sample** mean birth weight in the **zinc**-supplemented group.

Let $\boldsymbol{\bar{X}_2}$= **sample** mean birth weight in the **placebo** group.

(e) Determine the estimated number of degrees of freedom.

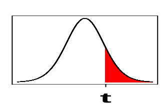
(f) Determine the P-value.





The Cumulative Standard Normal Distribution, pp. 914–915, Table E.2

	Cumulative Probabilities									
Z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮



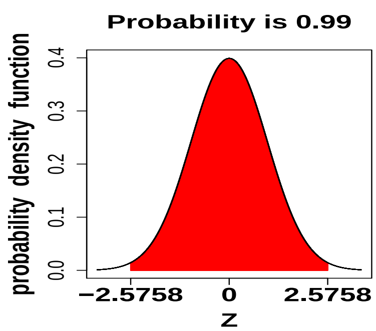
Critical Values of t, pp. 916–917, Table E.3

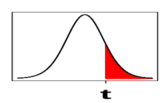
	Cumulative Probabilities					
	0.75	0.90	0.95	0.975	0.99	0.995
Degrees of Freedom	Upper-Tail Areas					
	0.25	.10	0.05	0.025	0.01	0.005
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	0.6770	1.2902	1.6604	1.9842	2.3646	2.6264
100	0.6770	1.2901	1.6602	1.9840	2.3642	2.6259
110	0.6767	1.2893	1.6588	1.9818	2.3607	2.6213
120	0.6765	1.2886	1.6577	1.9799	2.3578	2.6174
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758

(g) State the conclusion in statistical terms and in regular English.

We conclude that zinc supplementation during pregnancy among zinc-deficient mothers results in babies with increased mean weight at birth, in comparison to a placebo.

(h) Construct a **99%** confidence interval on $(\boldsymbol{\mu_1 - \mu_2})$.





Critical Values of t, pp. 916–917, Table E.3

	Cumulative Probabilities					
	0.75	0.90	0.95	0.975	0.99	0.995
Degrees of	Upper-Tail Areas					
Freedom	0.25	.10	0.05	0.025	0.01	0.005
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	0.6770	1.2902	1.6604	1.9842	2.3646	2.6264
100	0.6770	1.2901	1.6602	1.9840	2.3642	2.6259
110	0.6767	1.2893	1.6588	1.9818	2.3607	2.6213
120	0.6765	1.2886	1.6577	1.9799	2.3578	2.6174
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758

Layman's interpretation: We are 99% confident that the difference in population mean birth weights between zinc-users and placebo-users among zinc-deficient mothers lies between −23.7 grams and 275.7 grams.

Mathematically rigorous interpretation: If we repeat the sampling procedure many times to produce many 99% confidence intervals on $(\mu_1 - \mu_2)$, the difference in population mean birth weights between zinc-users and placebo-users among zinc-deficient mothers, then approximately 99% of these 99% confidence intervals will contain the true value of $(\mu_1 - \mu_2)$.

(i) Construct a **99%** confidence interval on $(\boldsymbol{\mu_2 - \mu_1})$.

□

10.2 Comparing the Means of Two Related Populations

Here, we pair the observations.

Construct paired-t test and paired-t confidence interval.

What are some examples of **paired observations**?

We assume the pairs of observations are independent or nearly independent, but we do **NOT** necessarily have independence **within** a pair.

Let $\boldsymbol{D}$ be the (observation in sample #1) − (observation in sample #2).

Again, we make inferences on the **difference between two means**, $(\mu_1 - \mu_2)$, or the **mean difference**, μ_D.

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

Assumptions:

(1) The observations are reasonably paired.

(2) The **differences** are independent or nearly independent (and σ_D is positive and finite).

(3) **n is large** (usually $n \geq 30$, if neither tail of the distribution *of the* **differences** is too heavy), or the **differences** are **approximately normal.**

Then, the standardized test statistic is $(\bar{D} - \mu_D)/(s_D/\sqrt{n}) \overset{approx.}{\sim} t_{n-1}$.

Confidence interval on μ_D is $\bar{D} \pm t_{n-1}\ s_D/\sqrt{n}$.

Example: *Hypothetical data.* Test at level $\alpha = 0.05$ whether the **population** mean (systolic reading of) blood pressure is reduced by more than 10 when using a placebo. The data consist of the following *before* and *after* blood pressure readings of five patients: $\{(190, 180), (220, 205), (242, 214), (175, 156), (201, 177)\}$.

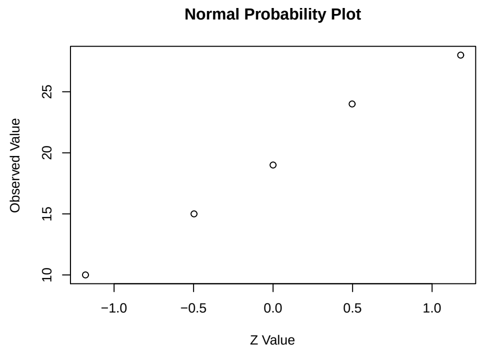
(a) Define your notation.

Let $\boldsymbol{D}$ be the difference in blood pressure, *before* minus *after*.

Let $\boldsymbol{\mu_D}$ be the *unknown* **population** mean difference in blood pressure.

(b) State the hypotheses.

(c) Check the assumptions.



(d) Determine the value of the **standardized test statistic.**

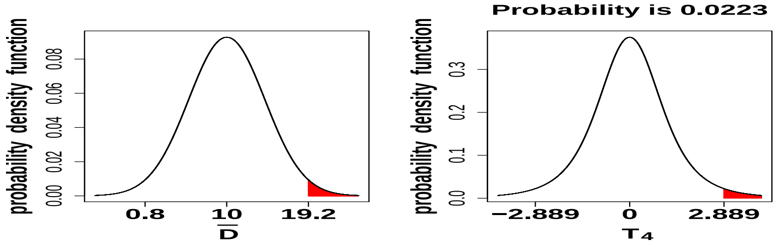
Let $\bar{D}$ be the **sample** mean difference in blood pressure.

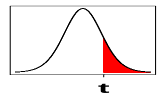
Let s_D be the **sample** standard deviation of the difference in blood pressure.

Goal: Construct a one-sample t test on μ_D.

(e) How many **degrees of freedom** are associated with this test?

(f) Determine the P-value.





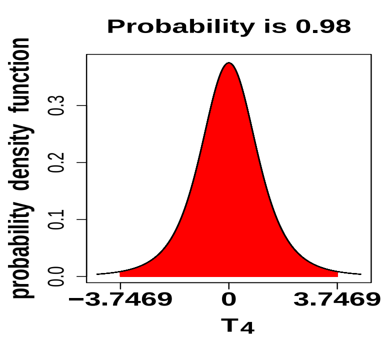
Critical Values of t, pp. 916–917, Table E.3

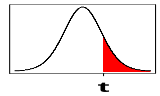
	Cumulative Probabilities					
	0.75	0.90	0.95	0.975	0.99	0.995
Degrees of	Upper-Tail Areas					
Freedom	0.25	.10	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3138	12.7062	31.8205	63.6567
2	0.8165	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
⋮	⋮	⋮	⋮	⋮	⋮	⋮

(g) State the conclusion in statistical terms and in regular English.

We conclude that the **population** mean (systolic reading of) blood pressure is reduced by more than 10 when using a placebo.

(h) Construct a **98%** confidence interval on μ_D.





Critical Values of t, pp. 916–917, Table E.3

	Cumulative Probabilities					
	0.75	0.90	0.95	0.975	0.99	0.995
Degrees of	Upper-Tail Areas					
Freedom	0.25	.10	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3138	12.7062	31.8205	63.6567
2	0.8165	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Layman's interpretation: We are 98% confident that μ_D, the population mean reduction in (systolic reading of) blood pressure due to the placebo effect, is between 7.27 and 31.13, when using a placebo.

Mathematically rigorous interpretation: If we repeat the sampling procedure many times to produce many 98% confidence intervals on μ_D, the population mean reduction in (systolic reading of) blood pressure due to the placebo effect, then approximately 98% of these 98% confidence intervals will contain the true value of μ_D.

□