

---

# Inference for a population mean

---

BPS chapter 18

# Objectives (BPS chapter 18)

## Inference about a Population Mean

- Conditions for inference
- The  $t$  distribution
- The one-sample  $t$  confidence interval
- Using technology
- Matched pairs  $t$  procedures
- Robustness of  $t$  procedures

# Conditions for inference about a mean

- ▣ We can regard our data as a **simple random sample** (SRS) from the population. This condition is very important.
- ▣ Observations from the population have a Normal distribution with mean  $\mu$  and standard deviation  $\sigma$ . In practice, it is enough that the distribution be symmetric and single-peaked unless the sample is very small. Both  $\mu$  and standard deviation  $\sigma$  are unknown.

## Sweetening colas



Cola manufacturers want to test how much the sweetness of a new cola drink is affected by storage. The sweetness loss due to storage was evaluated by 10 professional tasters (by comparing the sweetness before and after storage):

	Taster	Sweetness loss
□	1	2.0
□	2	0.4
□	3	0.7
□	4	2.0
□	5	-0.4
□	6	2.2
□	7	-1.3
□	8	1.2
□	9	1.1
□	10	2.3

Obviously, we want to test if storage results in a loss of sweetness, thus

$$H_0: \mu = 0 \text{ versus } H_a: \mu > 0$$

This looks familiar. However, here we do not know the population parameter  $\sigma$ .

- The population of all cola drinkers is too large.
- Since this is a new cola recipe, we have no population data.

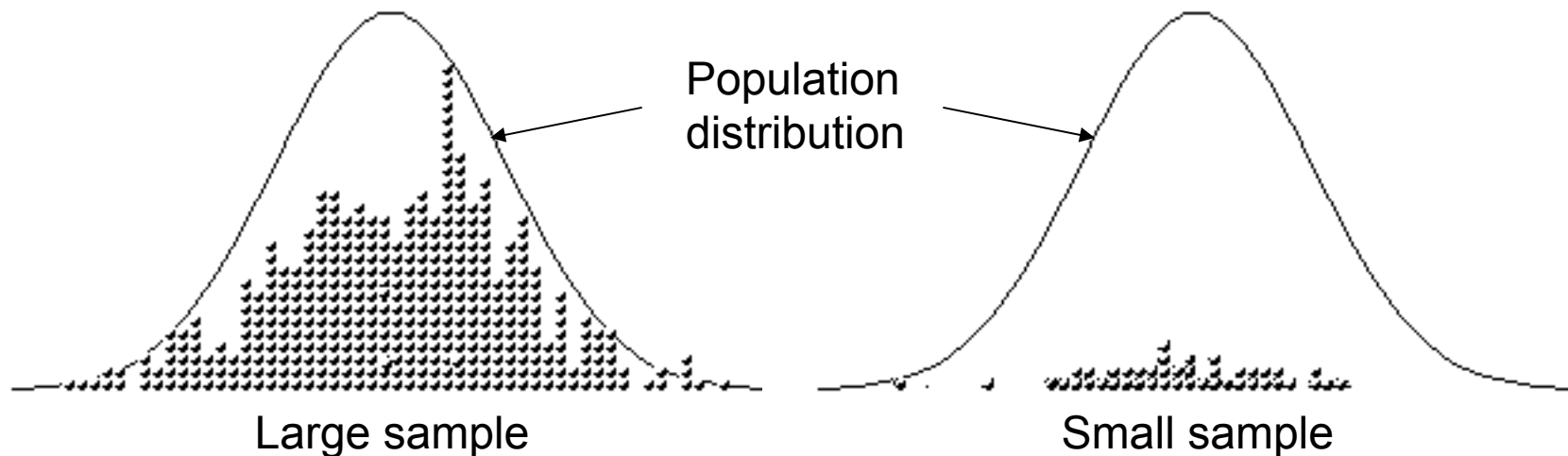
This situation is very common with real data.

# When $\sigma$ is unknown

The sample standard deviation  $s$  provides an estimate of the population standard deviation  $\sigma$ .

■ When the sample size is very large, the sample is likely to contain elements representative of the whole population. Then  $s$  is a very good estimate of  $\sigma$ .

■ But when the sample size is small, the sample contains only a few individuals. Then  $s$  is a more mediocre estimate of  $\sigma$ .



## Standard deviation $s$ — standard error of the mean $s/\sqrt{n}$

For a sample of size  $n$ ,  
the sample standard deviation  $s$  is:

$$s = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

$n - 1$  is the “degrees of freedom.”

The value  $s/\sqrt{n}$  is called the standard error of the mean **SEM**.

Scientists often present their sample results as the mean  $\pm$  SEM.



Example: A medical study examined the effect of a new medication on the seated systolic blood pressure. The results, presented as mean  $\pm$  SEM for 25 patients, are  $113.5 \pm 8.9$ . What is the standard deviation  $s$  of the sample data?

$$\begin{aligned} \text{SEM} = s/\sqrt{n} &\iff s = \text{SEM} \cdot \sqrt{n} \\ s &= 8.9 \cdot \sqrt{25} = 44.5 \end{aligned}$$

# The $t$ distributions

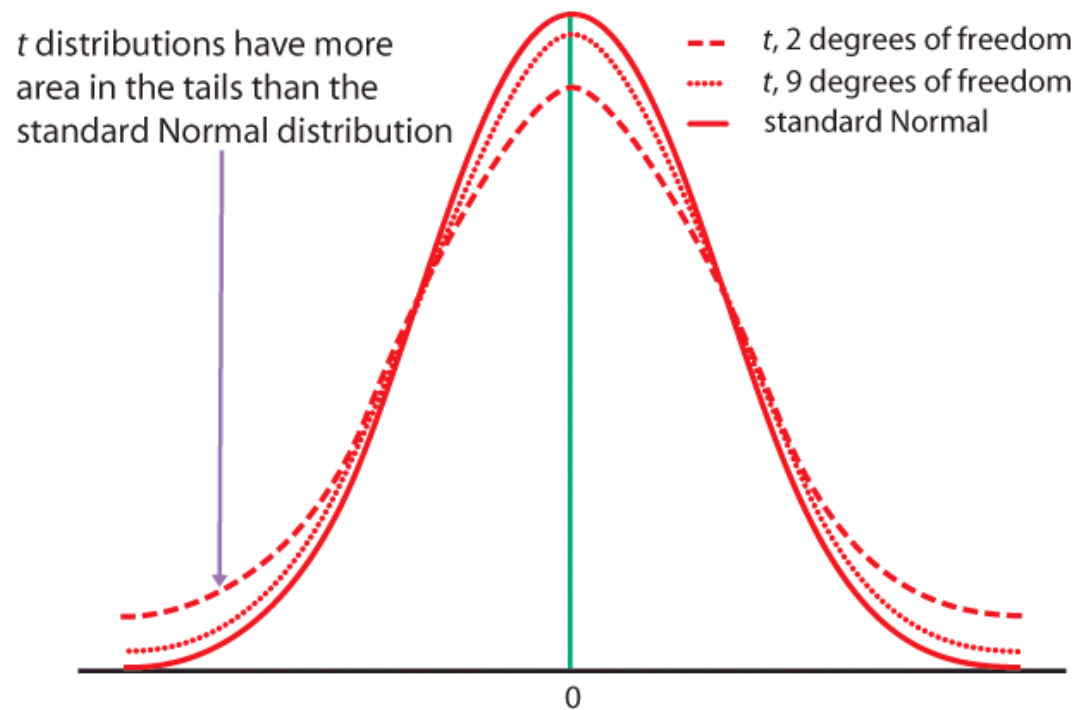
We test a null and alternative hypotheses with one sample of size  $n$  from a normal population  $N(\mu, \sigma)$ :

- When  $\sigma$  is known, the sampling distribution is normal  $N(\mu, \sigma/\sqrt{n})$ .
- When  $\sigma$  is estimated from the sample standard deviation  $s$ , then the sampling distribution follows a  **$t$  distribution  $t(\mu, s/\sqrt{n})$  with degrees of freedom  $n - 1$ .**

The value  $(s/\sqrt{n})$  is the standard error of the mean or SEM.

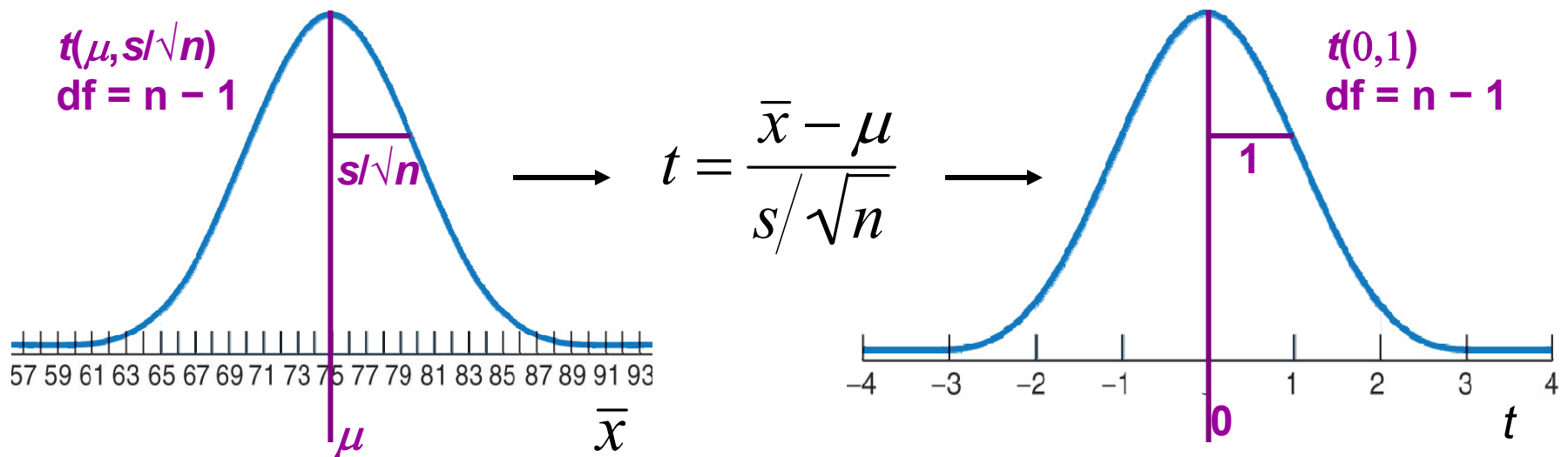
When  $n$  is very large,  $s$  is a very good estimate of  $\sigma$  and the corresponding  $t$  distributions are very close to the normal distribution.

The  $t$  distributions become wider for smaller sample sizes, reflecting the lack of precision in estimating  $\sigma$  from  $s$ .



## Standardizing the data before using Table C

As with the normal distribution, the first step is to standardize the data. Then we can use **Table C** to obtain the area under the curve.



*Here,  $\mu$  is the mean (center) of the sampling distribution, and the standard error of the mean  $s/\sqrt{n}$  is its standard deviation (width). You obtain  $s$ , the standard deviation of the sample, with your calculator.*

# Table C

When  $\sigma$  is unknown we use the sample standard deviation and a  $t$  distribution with “ $n - 1$ ” degrees of freedom (df).

Table C shows the z-values and  $t$ -values corresponding to landmark  $P$ -values/ confidence levels.

When  $\sigma$  is known, we use the normal distribution and the standardized z-value.

**TABLE C**  $t$  distribution critical values

degrees of freedom	Confidence level $C$												
	80%	85%	90%	95%	98%	99%	99.5%	99.8%	99.9%	99.95%	99.98%	99.99%	
1	1.000	1.578	1.963	2.078	2.314	2.771	3.083	3.182	3.362	3.501	3.659	3.851	4.009
2	0.816	1.061	1.386	1.533	1.699	2.074	2.306	2.353	2.479	2.576	2.689	2.818	2.938
3	0.765	0.978	1.280	1.415	1.577	1.962	2.157	2.202	2.317	2.398	2.499	2.627	2.746
4	0.741	0.941	1.190	1.327	1.487	1.860	2.045	2.089	2.203	2.283	2.383	2.511	2.629
5	0.727	0.920	1.156	1.295	1.455	1.824	2.009	2.053	2.167	2.246	2.346	2.474	2.592
6	0.718	0.908	1.134	1.274	1.434	1.801	1.986	2.029	2.143	2.222	2.322	2.450	2.568
7	0.711	0.900	1.119	1.259	1.419	1.789	1.974	2.017	2.131	2.210	2.310	2.438	2.556
8	0.706	0.899	1.108	1.248	1.408	1.779	1.964	1.997	2.111	2.190	2.290	2.418	2.536
9	0.703	0.893	1.100	1.240	1.400	1.772	1.957	1.989	2.103	2.182	2.282	2.410	2.528
10	0.700	0.879	1.093	1.233	1.393	1.766	1.951	1.983	2.097	2.176	2.276	2.404	2.522
11	0.697	0.876	1.088	1.228	1.388	1.760	1.945	1.977	2.089	2.168	2.268	2.396	2.514
12	0.695	0.873	1.083	1.223	1.383	1.755	1.940	1.971	2.083	2.162	2.262	2.390	2.508
13	0.694	0.870	1.079	1.219	1.379	1.751	1.935	1.966	2.077	2.156	2.256	2.384	2.502
14	0.692	0.868	1.076	1.216	1.376	1.748	1.932	1.963	2.075	2.154	2.254	2.382	2.500
15	0.691	0.866	1.074	1.214	1.374	1.746	1.930	1.961	2.073	2.152	2.252	2.380	2.498
16	0.690	0.865	1.071	1.211	1.371	1.744	1.928	1.959	2.071	2.150	2.250	2.378	2.496
17	0.689	0.863	1.069	1.209	1.369	1.742	1.926	1.957	2.069	2.148	2.248	2.376	2.494
18	0.688	0.862	1.067	1.207	1.367	1.740	1.924	1.955	2.067	2.146	2.246	2.374	2.492
19	0.688	0.861	1.066	1.206	1.366	1.739	1.923	1.954	2.065	2.144	2.244	2.372	2.490
20	0.687	0.860	1.064	1.204	1.364	1.738	1.922	1.953	2.064	2.143	2.243	2.371	2.489
21	0.686	0.859	1.063	1.203	1.363	1.737	1.921	1.952	2.063	2.142	2.242	2.370	2.488
22	0.686	0.858	1.061	1.201	1.361	1.736	1.920	1.951	2.062	2.141	2.241	2.369	2.487
23	0.685	0.858	1.060	1.200	1.360	1.735	1.919	1.950	2.061	2.140	2.240	2.368	2.486
24	0.685	0.857	1.059	1.198	1.358	1.734	1.918	1.949	2.060	2.139	2.239	2.367	2.485
25	0.684	0.856	1.058	1.197	1.357	1.733	1.917	1.948	2.059	2.138	2.238	2.366	2.484
26	0.684	0.856	1.058	1.196	1.356	1.732	1.916	1.947	2.058	2.137	2.237	2.365	2.483
27	0.684	0.855	1.057	1.195	1.355	1.731	1.915	1.946	2.057	2.136	2.236	2.364	2.482
28	0.683	0.855	1.056	1.194	1.354	1.730	1.914	1.945	2.056	2.135	2.235	2.363	2.481
29	0.683	0.854	1.055	1.193	1.353	1.729	1.913	1.944	2.055	2.134	2.234	2.362	2.480
30	0.683	0.854	1.055	1.192	1.352	1.728	1.912	1.943	2.054	2.133	2.233	2.361	2.479
40	0.681	0.851	1.050	1.303	1.654	2.021	2.123	2.423	2.704	2.971	3.307	3.851	
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496	
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.918	3.232	3.460	
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.198	3.416	
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390	
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.066	2.330	2.581	2.813	3.098	3.300	
$z^*$	0.674	0.841	1.036	1.282	1.646	1.960	2.064	2.326	2.576	2.807	3.091	3.291	
One-sided $P$	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005	
Two-sided $P$	.50	.40	.30	.20	.10	.05	.04	.02	.01	.005	.002	.001	

# Table A vs. Table C

Table A gives the area to the LEFT of hundreds of z-values.

It should only be used for normal distributions.

**TABLE A Standard normal probabilities**

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110

(...)

**TABLE C t distribution critical values**

degrees of freedom	Confidence level C											
	80%	80%	70%	50%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
1	1.000	1.378	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.80
3	0.785	0.978	1.250	1.638	2.353	3.182	3.482	4.841	6.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	4.998	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.908	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.898	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.853	2.282	2.398	2.821	3.280	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587

(...)

100	0.677	0.846	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
One-sided P	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
Two-sided P	.50	.40	.30	.20	.10	.05	.04	.02	.01	.005	.002	.001

Table C also gives the middle area under a t or normal distribution comprised between the negative and positive value of t or z.

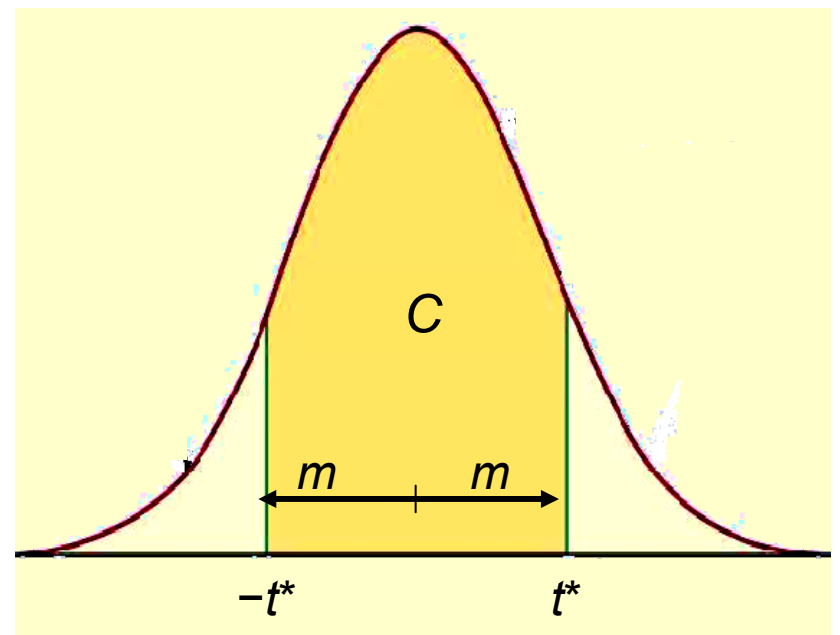
# Confidence intervals

Reminder: The **confidence interval** is a range of values with a **confidence level C** representing the probability that the interval contains the true population parameter.

We have a set of data from a population with both  $\mu$  and  $\sigma$  unknown. We use  $\bar{x}$  to estimate  $\mu$ , and  $s$  to estimate  $\sigma$ , using a  $t$  distribution (df  $n - 1$ ).

## Practical use of $t$ : $t^*$

- ▢ C is the area under the  $t$  (df:  $n-1$ ) curve between  $-t^*$  and  $t^*$ .
- ▢ We find  $t^*$  in the line of Table C for df =  $n-1$  and confidence level C.
- ▢ The margin of error  $m$  is:  
$$m = t^* s / \sqrt{n}$$



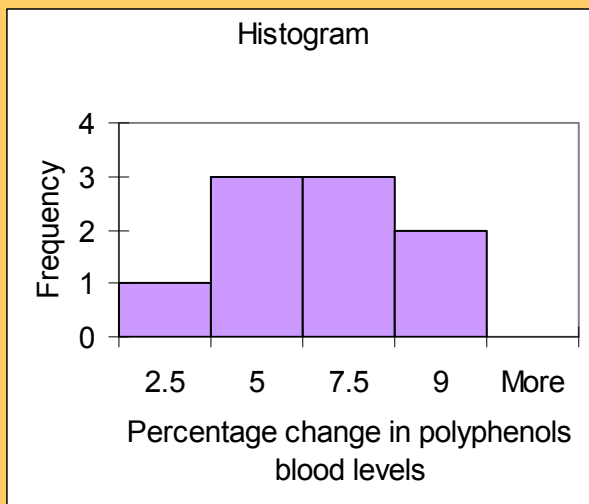
## Red wine, in moderation



Drinking red wine in moderation may protect against heart attacks. The polyphenols it contains act on blood cholesterol and thus are a likely cause.

To test the hypothesis that moderate red wine consumption increases the average blood level of polyphenols, a group of nine randomly selected healthy men were assigned to drink half a bottle of red wine daily for 2 weeks. Their blood polyphenol levels were assessed before and after the study and the percent change is presented here: **0.7 3.5 4 4.9 5.5 7 7.4 8.1 8.4**

Firstly: Are the data approximately normal?



0	7
1	
2	
3	5
4	09
5	5
6	
7	04
8	14

Normal?

When the sample size is small, histograms can be difficult to interpret.

There is a low value, but overall the data can be considered reasonably normal.

## Red wine, in moderation (*continued*)



What is the 95% confidence interval for the average percent change in blood polyphenols?

Sample average = 5.5;  $s = 2.517$ ;  $df = n - 1 = 8$

8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
(...)												
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level $C$											

The sampling distribution is a  $t$  distribution with  $n - 1$  degrees of freedom. For  $df = 8$  and  $C = 95\%$ ,  $t^* = 2.306$ .

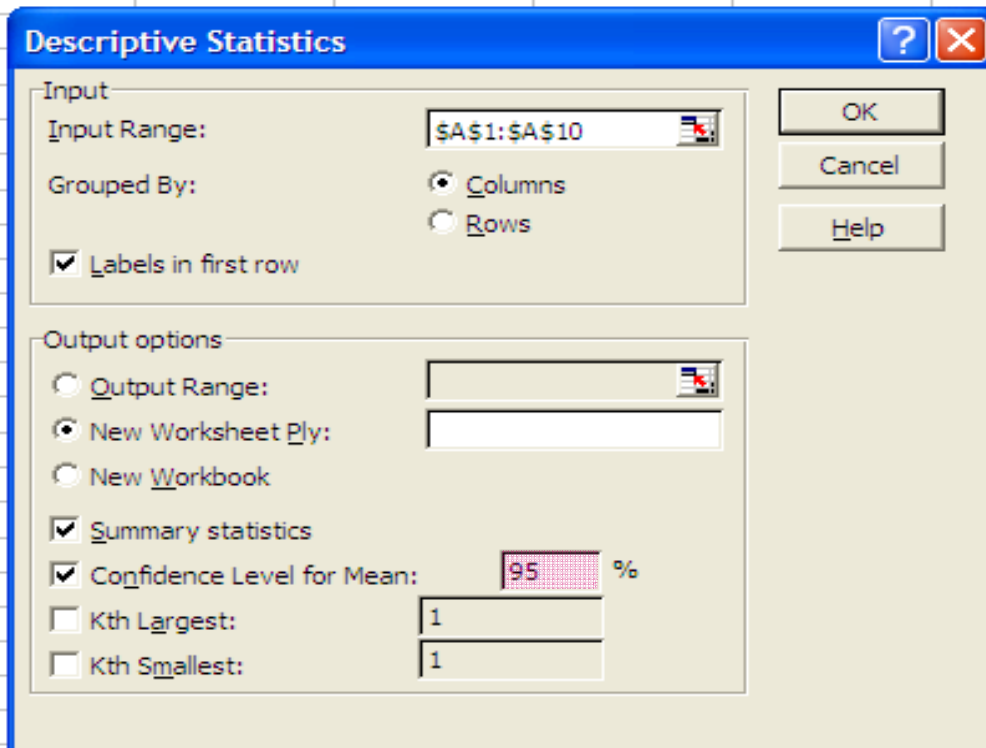
The margin of error  $m$  is :  $m = t^* s / \sqrt{n} = 2.306 * 2.517 / \sqrt{9} \approx 1.93$ .

The 95% confidence interval is therefore  $5.5 \pm 1.93$ .

**With 95% confidence, the average percent increase in polyphenol blood levels of healthy men drinking half a bottle of red wine daily is between 3.6% and 7.6%. Important: The confidence interval shows how large the increase is, but not if it can have an impact on men's health.**

# Excel

Menu: Tools/DataAnalysis: select “Descriptive statistics”



PercentChange	
Mean	5.5
Standard Error	0.838981 $s/\sqrt{n}$
Median	5.5
Mode	#N/A
Standard Deviation	2.516943
Sample Variance	6.335
Kurtosis	0.010884
Skewness	-0.7054
Range	7.7
Minimum	0.7
Maximum	8.4
Sum	49.5
Count	9
Confidence Level(95.0%)	1.934695 $m$

PercentChange	
Confidence Level(95.0%)	1.934695 $m$

**Warning !!! Do not use the function =CONFIDENCE(alpha, stdev, size).**  
 This assumes a normal sampling distribution and uses  $z^*$  instead of  $t^*$ !!!

# The $t$ -test

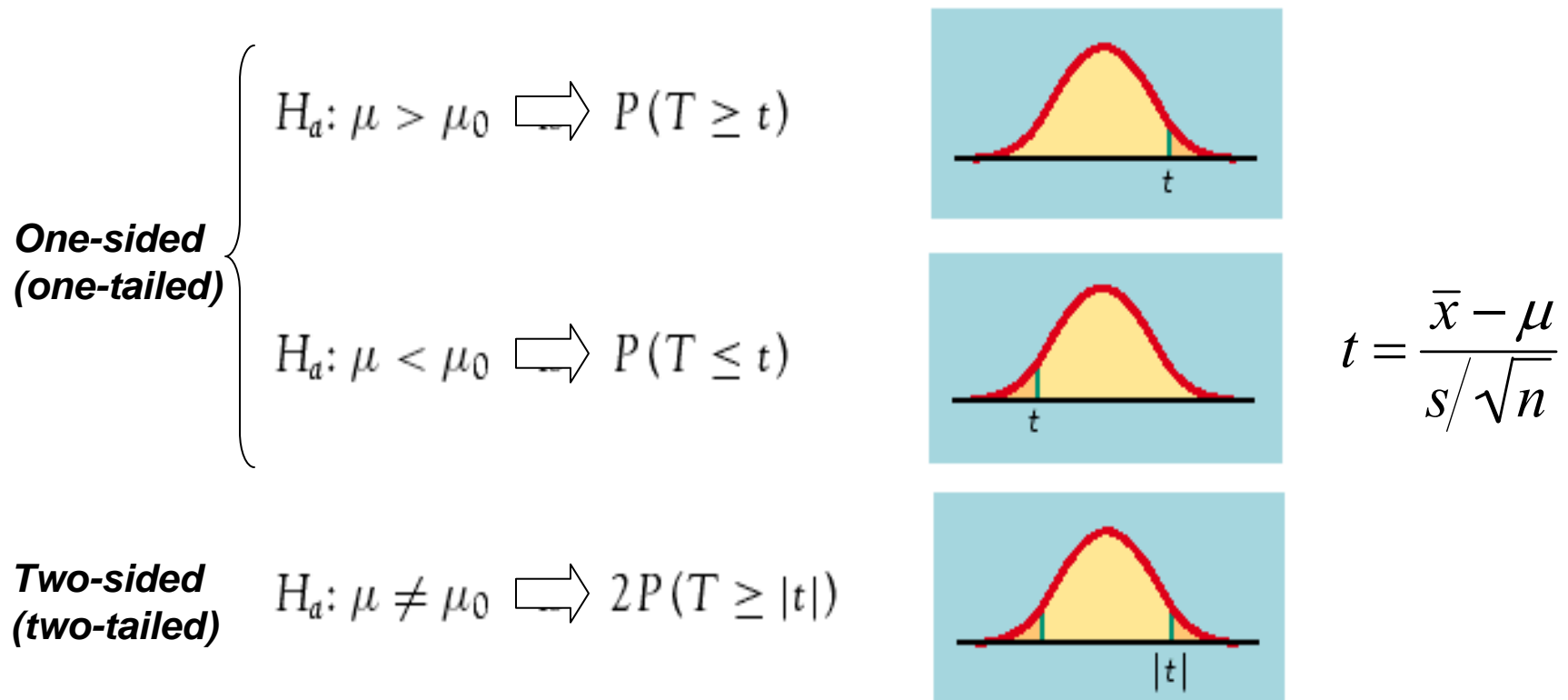
As in the previous chapter, a test of hypotheses requires a few steps:

1. Stating the null and alternative hypotheses ( $H_0$  versus  $H_a$ )
2. Deciding on a one-sided or two-sided test
3. Choosing a significance level  $\alpha$
4. Calculating  $t$  and its degrees of freedom
5. Finding the area under the curve *with Table C*
6. Stating the  $P$ -value and interpreting the result

## Review: test of significance

The **P-value** is the probability, if  $H_0$  is true, of randomly drawing a sample like the one obtained, or more extreme, in the direction of  $H_a$ .

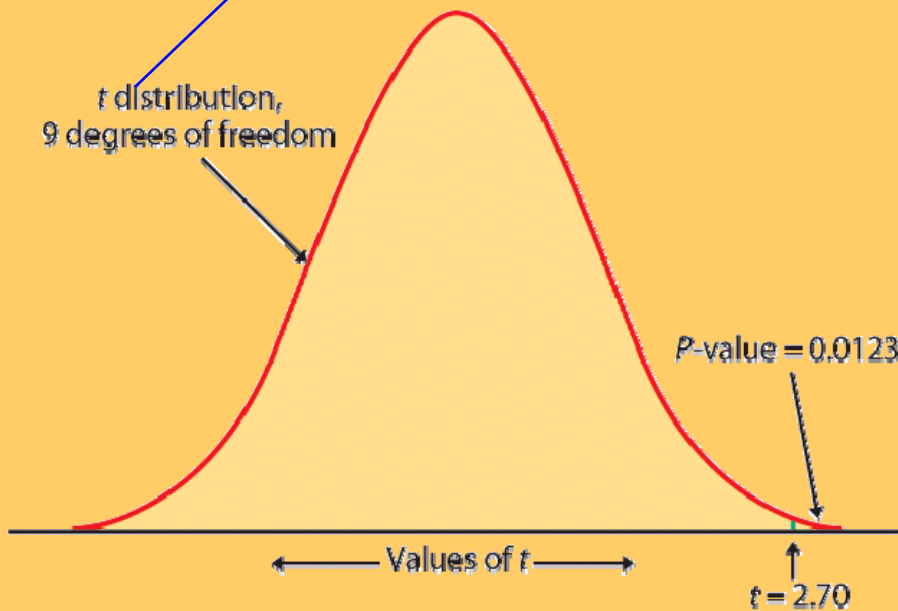
The P-value is calculated as the corresponding area under the curve, one-tailed or two-tailed depending on  $H_a$ :



# Table C How to:

1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781

1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
$z^*$	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
One-sided $P$	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
Two-sided $P$	.50	.40	.30	.20	.10	.05	.04	.02	.01	.005	.002	.001



The calculated value of  $t$  is 2.7.  
We find the two closest  $t$  values.

$$2.398 < t = 2.7 < 2.821$$

thus

$$0.02 > \text{upper tail } p > 0.01$$

For a one-sided  $H_a$ , this is the  $P$ -value (between 0.01 and 0.02);  
for a two-sided  $H_a$ , the  $P$ -value is doubled (between 0.02 and 0.04).

# Excel

## TDIST(x, degrees\_freedom, tails)

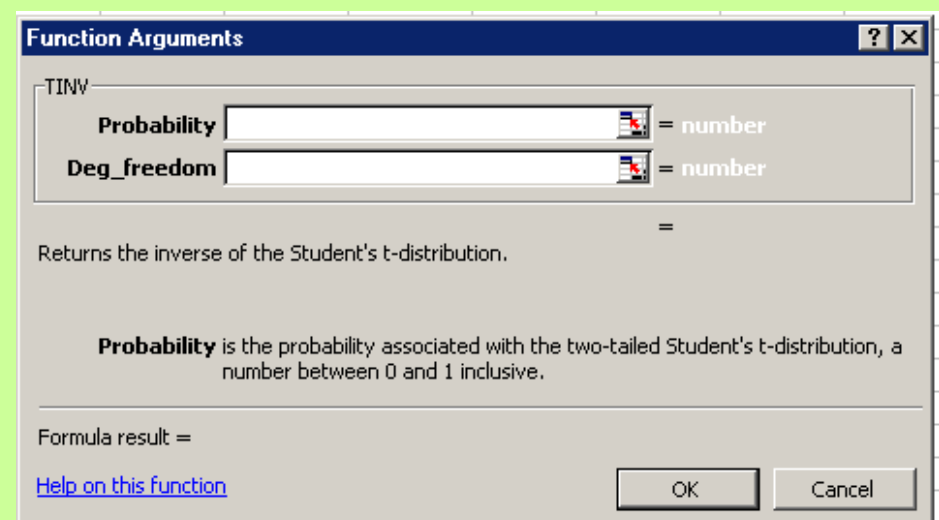
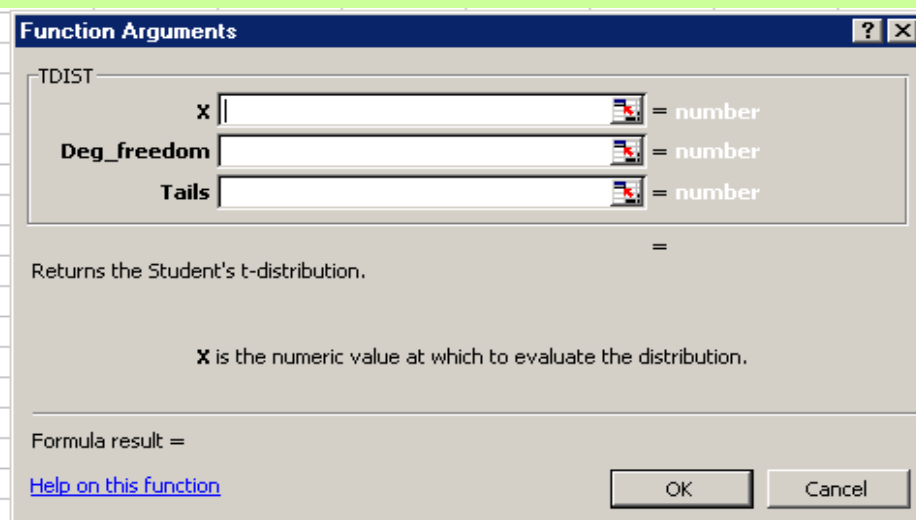
TDIST =  $p(X > x)$ , where  $X$  is a random variable that follows the  $t$  distribution ( $x$  positive). Use this function in place of a table of critical values for the  $t$  distribution or to obtain the  $P$ -value for a calculated, positive  $t$ -value.

- ❑  $X$  is the *standardized* numeric value at which to evaluate the distribution (“ $t$ ”).
- ❑ *Degrees\_freedom* is an integer indicating the number of degrees of freedom.
- ❑ *Tails* specifies the number of distribution tails to return. If tails = 1, TDIST returns the one-tailed  $P$ -value. If tails = 2, TDIST returns the two-tailed  $P$ -value.

## TINV(probability, degrees\_freedom)

Returns the  $t$ -value of the Student's  $t$ -distribution as a function of the probability and the degrees of freedom (for example,  $t^*$ ).

- ❑ *Probability* is the probability associated with the two-tailed Student's  $t$  distribution.
- ❑ *Degrees\_freedom* is the number of degrees of freedom characterizing the distribution.





## Sweetening colas (*continued*)

Is there evidence that storage results in sweetness loss for the new cola recipe at the 0.05 level of significance ( $\alpha = 5\%$ )?

$H_0: \mu = 0$  versus  $H_a: \mu > 0$  (one-sided test)

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1.02 - 0}{1.196/\sqrt{10}} = 2.70$$

$$df = n - 1 = 9$$

- ▣ the critical value  $t_\alpha = 1.833$   
 $t > t_\alpha$  thus the result is significant.
- ▣  $2.398 < t = 2.70 < 2.821$ , thus  $0.02 > p > 0.01$   
 $p < \alpha$ , thus the result is significant.

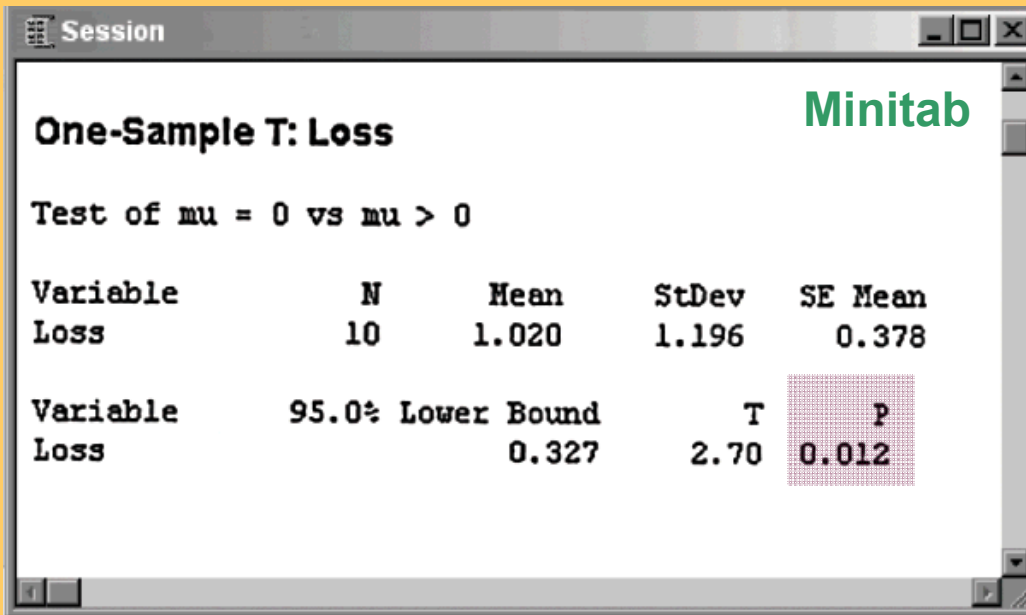
Taster	Sweetness loss
1	2.0
2	0.4
3	0.7
4	2.0
5	-0.4
6	2.2
7	-1.3
8	1.2
9	1.1
10	2.3
<hr/>	
Average	1.02
Standard deviation	1.196

9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
One-sided $P$	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
Two-sided $P$	.50	.40	.30	.20	.10	.05	.04	.02	.01	.005	.002	.001

The  $t$ -test has a significant  $p$ -value. We reject  $H_0$ .

**There is a significant loss of sweetness, on average, following storage.**

## Sweetening colas (*continued*)



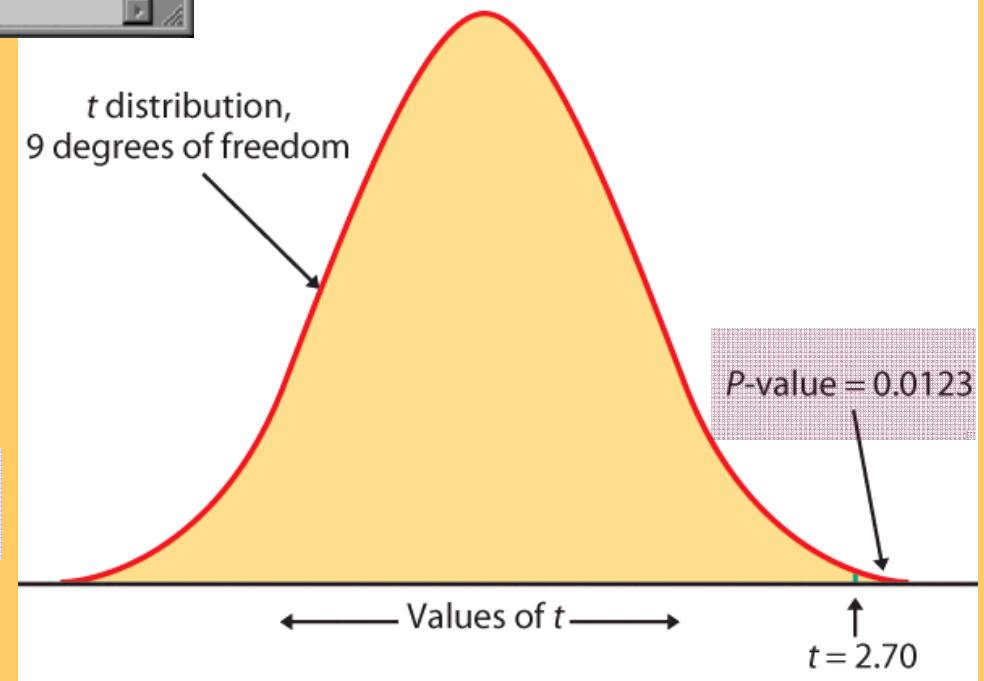
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1.02 - 0}{1.196/\sqrt{10}} = 2.70$$

$$df = n - 1 = 9$$

In Excel, you can obtain the precise *P*-value once you have calculated *t*:

Use the function `dist(t, df, tails)`

`"=tdist(2.7, 9, 1),"` which gives 0.01226



## Red wine, in moderation (*continued*)

Does moderate red wine consumption increase the average blood level of polyphenols in healthy men?

$H_0: \mu = 0$  versus  $H_a: \mu > 0$   
(one-sided test)

Sample average = 5.5;  $s = 2.517$ ;

$$t = (5.5 - 0) / (2.517 / \sqrt{9}) \approx 6.556$$



Test statistic would be off the chart to the right

8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
One-sided $P$	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
Two-sided $P$	.50	.40	.30	.20	.10	.05	.04	.02	.01	.005	.002	.001

From Table C,  $df = 8$ :  $t > 5.041$  and therefore  $p > 0.0005$ .

The  $P$ -value is very small (well below 1%), and thus the result is very significant.

**Moderate red wine consumption significantly increases the average polyphenol blood levels of healthy men. Important: This test does not say how large the increase is, or what the impact on men's health is.**

# Matched pairs *t* procedures

Sometimes we want to compare treatments or conditions at the individual level. These situations produce two samples that are not independent — they are related to each other. The members of one sample are identical to, or matched (paired) with, the members of the other sample.

- ❑ Example: Pre-test and post-test studies look at data collected on the same sample elements before and after some experiment is performed.
- ❑ Example: Twin studies often try to sort out the influence of genetic factors by comparing a variable between sets of twins.
- ❑ Example: Using people matched for age, sex, and education in social studies allows us to cancel out the effect of these potential lurking variables.

---

In these cases, we use the paired data to test the difference in the two population means. The variable studied becomes  $X = x_1 - x_2$ , and

$$H_0: \mu_{\text{difference}} = 0; H_a: \mu_{\text{difference}} > 0 \text{ (or } < 0, \text{ or } \neq 0)$$

**Conceptually, this does not differ from tests on one population.**

## Sweetening colas (revisited)



The sweetness loss due to storage was evaluated by 10 professional tasters (comparing the sweetness before and after storage):

	Taster	Sweetness loss
□	1	2.0
□	2	0.4
□	3	0.7
□	4	2.0
□	5	-0.4
□	6	2.2
□	7	-1.3
□	8	1.2
□	9	1.1
□	10	2.3

We want to test if storage results in a loss of sweetness, thus

$$H_0: \mu = 0 \text{ versus } H_a: \mu > 0$$

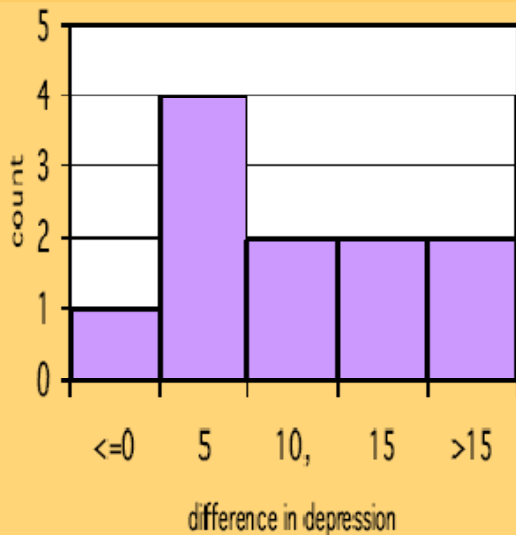
Although the text did not mention it explicitly, this is a pre-/post-test design, and the variable is the difference in cola sweetness before and after storage.

A matched pairs test of significance is indeed just like a one-sample test.

# Does lack of caffeine increase depression?

Individuals diagnosed as caffeine-dependent are deprived of all caffeine-rich foods and assigned to receive daily pills. At some time, the pills contain caffeine and at another time they contain a placebo. Depression was assessed.

- There are two data points for each subject, but we will only look at the difference.
- The sample distribution appears appropriate for a *t*-test.



11 "difference" data points

Subject	Depression with Caffeine	Depression with Placebo	Placebo - Caffeine
1	5	16	11
2	5	23	18
3	4	5	1
4	3	7	4
5	8	14	6
6	5	24	19
7	0	6	6
8	0	3	3
9	2	15	13
10	11	12	1
11	1	0	-1

## Does lack of caffeine increase depression?

For each individual in the sample, we have calculated a difference in depression score (placebo minus caffeine).

There were 11 “difference” points, thus  $df = n - 1 = 10$ .

We calculate that  $\bar{x} = 7.36$ ;  $s = 6.92$

$$H_0: \mu_{\text{difference}} = 0 ; H_a: \mu_{\text{difference}} > 0$$

$$t = \frac{\bar{x} - 0}{s/\sqrt{n}} = \frac{7.36}{6.92/\sqrt{11}} = 3.53$$

Subject	Depression with Caffeine	Depression with Placebo	Placebo - Caffeine
1	5	16	11
2	5	23	18
3	4	5	1
4	3	7	4
5	8	14	6
6	5	24	19
7	0	6	6
8	0	3	3
9	2	15	13
10	11	12	1
11	1	0	-1

For  $df = 10$ ,  $3.169 < t = 3.53 < 3.581$ , therefore  $0.005 > p > 0.0025$ .

**Caffeine deprivation causes a significant increase in depression.**

# Robustness

The  $t$  procedures are exactly correct when the population is distributed exactly normally. However, most real data are not exactly normal.

The  $t$  procedures are **robust** to small deviations from normality. This means that the results will not be affected too much. Factors that do strongly matter are:

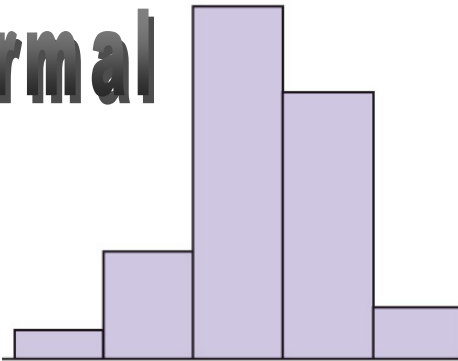
- ❑ **Random sampling.** The sample **must** be an SRS from the population.
- ❑ **Outliers and skewness.** They strongly influence the mean and therefore the  $t$  procedures. However, their impact diminishes as the sample size gets larger because of the Central Limit Theorem.

Specifically:

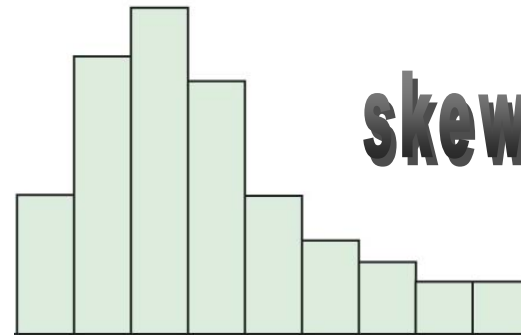
- ❑ When  $n < 15$ , the data must be close to normal and without outliers.
- ❑ When  $15 > n > 40$ , mild skewness is acceptable, but not outliers.
- ❑ When  $n > 40$ , the  $t$  statistic will be valid even with strong skewness.

# Reminder: Looking at histograms for normality

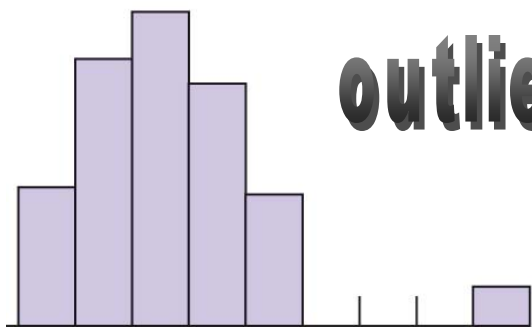
**normal**



**skewed**



**outlier**



**bimodal**

