

# Two-sample Problems

The set-up: we wish to compare two population means  $\mu_1$  and  $\mu_2$ , so we take a sample from each.

Population	Sample size	Sample mean	Sample standard deviation
1	$n_1$	$\bar{x}_1$	$s_1$
2	$n_2$	$\bar{x}_2$	$s_2$

We look at the difference  $\bar{x}_1 - \bar{x}_2$

**Example:** What is the relationship between obesity and energy expended on daily activities?

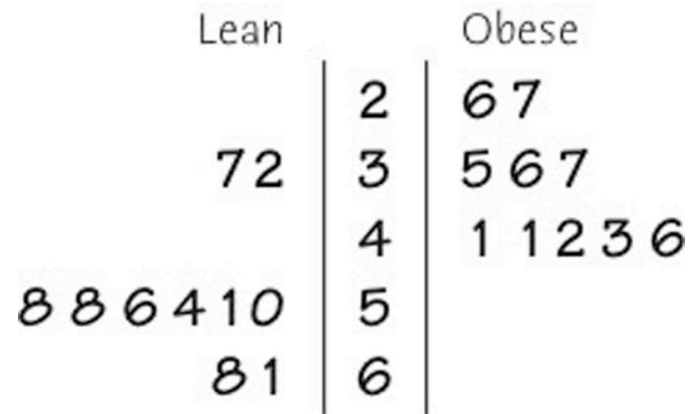
**TABLE 18.1** Time (minutes per day) spent in three different postures by lean and obese subjects

GROUP	SUBJECT	STAND/WALK	SIT	LIE
Lean	1	511.100	370.300	555.500
Lean	2	607.925	374.512	450.650
Lean	3	319.212	582.138	537.362
Lean	4	584.644	357.144	489.269
Lean	5	578.869	348.994	514.081
Lean	6	543.388	385.312	506.500
Lean	7	677.188	268.188	467.700
Lean	8	555.656	322.219	567.006
Lean	9	374.831	537.031	531.431
Lean	10	504.700	528.838	396.962
Obese	11	260.244	646.281	521.044
Obese	12	464.756	456.644	514.931
Obese	13	367.138	578.662	563.300
Obese	14	413.667	463.333	532.208
Obese	15	347.375	567.556	504.931
Obese	16	416.531	567.556	448.856
Obese	17	358.650	621.262	460.550
Obese	18	267.344	646.181	509.981
Obese	19	410.631	572.769	448.706
Obese	20	426.356	591.369	412.919

# STAND/WALK time:

	n	Mean $\bar{x}$	Std. dev. s
Group 1 (lean)	10	525.751	107.121
Group 2 (obese)	10	373.269	67.498

The difference is  $\bar{x}_1 - \bar{x}_2 = 525.751 - 373.269 = 152.482$  minutes.



**Two-sample t-Procedures:** The differences  $\bar{x}_1 - \bar{x}_2$  follow (approximately) a t-distribution. We can estimate the standard deviation by the standard error

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- Confidence interval for  $\mu_1 - \mu_2$  is  $(\bar{x}_1 - \bar{x}_2) \pm t^* SE$
- Significance Tests  $H_0: \mu_1 = \mu_2$ , use test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Use degrees of freedom  $df = \text{smaller of } n_1 - 1 \text{ and } n_2 - 1$

### Example: Daily activity and obesity.

	n	Mean $\bar{x}$	Std. dev. s
Group 1 (lean)	10	525.751	107.121
Group 2 (obese)	10	373.269	67.498

The difference is  $\bar{x}_1 - \bar{x}_2 = 525.751 - 373.269 = 152.482$  minutes.

**90% confidence interval** for  $\mu_1 - \mu_2$ :

df = 9  $\rightarrow$   $t^* = 1.833$

$$\begin{aligned}(\bar{x}_1 - \bar{x}_2) \pm t^* SE &= (525.751 - 373.269) \pm (1.833) \sqrt{\frac{107.121^2}{10} + \frac{67.498^2}{10}} \\ &= 152.482 \pm 73.390 = (79.09, 225.87)\end{aligned}$$

We are 90% confident the true difference in average daily minutes spent standing/walking is between 79.09 and 225.87.

## Example: Daily activity and obesity.

	n	Mean $\bar{x}$	Std. dev. s
Group 1 (lean)	10	525.751	107.121
Group 2 (obese)	10	373.269	67.498

The difference is  $\bar{x}_1 - \bar{x}_2 = 525.751 - 373.269 = 152.482$  minutes.

**Is this evidence that lean people are more active?**

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 > \mu_2$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{525.751 - 373.269}{\sqrt{\frac{107.121^2}{10} + \frac{67.498^2}{10}}} = 3.808$$

Table C (df = 9) gives P between 0.001 and 0.0025.  
This is evidence that lean people are more active.

**Example:** Do brand name batteries last longer than generic?

To test, measure lifespan of batteries in portable CD player.

---

	Sample size	Sample mean	Sample standard deviation
Brand name	6	$\bar{x}_1 = 206.0$	$s_1 = 10.3$
Generic	6	$\bar{x}_2 = 187.4$	$s_2 = 14.6$

---

**95% confidence interval:**  $df = 5 \rightarrow t^* = 2.571$

$$\begin{aligned}(\bar{x}_1 - \bar{x}_2) \pm t^* SE &= (206.0 - 187.4) \pm (2.571) \sqrt{\frac{10.3^2}{6} + \frac{14.6^2}{6}} \\ &= 18.6 \pm 18.75 = (-0.15, 37.35)\end{aligned}$$

We are 95% confident that the mean difference in lifespan between brand name and generic batteries is between  $-0.15$  and  $37.35$  minutes.

	Sample size	Sample mean	Sample standard deviation
Brand name	6	$\bar{x}_1 = 206.0$	$s_1 = 10.3$
Generic	6	$\bar{x}_2 = 187.4$	$s_2 = 14.6$

Are these samples evidence at the  $\alpha = 0.05$  level that brand name batteries last longer?

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 > \mu_2$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{206.0 - 187.4}{\sqrt{\frac{10.3^2}{6} + \frac{14.6^2}{6}}} = 2.55$$

df = 5 Table C gives P between 0.025 and 0.05

This is evidence at the  $\alpha = 0.05$  level that name brand batteries last longer.

**Example:** Do people offer different prices when buying from friends?

To test this, sample of people offered bids for a used camera.

buying from friend	275	300	260	300	255	275	290	300
buying from stranger	260	250	175	130	200	225	240	

	Sample size	Sample mean	Sample standard deviation
friend	8	$\bar{x}_1 = 281.88$	$s_1 = 18.31$
stranger	7	$\bar{x}_2 = 211.43$	$s_2 = 46.43$

Is this evidence that people make different sized offers when buying from friends?

	Sample size	Sample mean	Sample standard deviation
friend	8	$\bar{x}_1 = 281.88$	$s_1 = 18.31$
stranger	7	$\bar{x}_2 = 211.43$	$s_2 = 46.43$

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{281.88 - 211.43}{\sqrt{\frac{18.31^2}{8} + \frac{46.43^2}{7}}} = 3.77$$

Table C (using  $df = 6$ , Two-sided P) gives P-value between 0.005 and 0.0025.

This is strong evidence that people make different sized offers to friends.

- We use  $df = \text{smaller of } n_1 - 1 \text{ and } n_2 - 1$ ; results using this are approximate and give conservative estimates for confidence intervals and P-values.

There is a more precise (and ugly) formula for  $df$ , which we will not use. (See p. 486 if you're curious.)

- 2-Sample tests with the TI 83/84:  
confidence intervals **STAT** → **TESTS** → **0:2-SampTInt**

significance tests **STAT** → **TESTS** → **4:2-SampTTest**

(Always set Pooled: **No**)

(TI 83/84 answers use the exact formula for  $df$ , so they will differ slightly from calculations that use  $df = \text{smaller of } n_1 - 1 \text{ and } n_2 - 1$ .)

## Using the 2-Sample $t$ Procedures:

- $n_1 + n_2$  *less than 15*: Use if data appear close to Normal. If data is skewed, or outliers, do not use.
- $n_1 + n_2$  *at least 15*: The  $t$  procedures can be used except if outliers or strong skewness.
- $n_1 + n_2$  *large (roughly  $n_1 + n_2 \geq 40$ )*: The  $t$  procedures can be used even for skewed distributions.