## TOPIC: TWO MEANS – UNKNOWN, UNEQUAL VARIANCE

### Difference in Means: Hypothesis Tests

- In hypothesis tests given ***TWO*** samples, we test claims about the ____________ between the means.
  - Write $H_0$ as $\mu_1 =$ ______, i.e. $\mu_1 - \mu_2 =$____
  - Find the $P$-value using the ____________ of $n_1 - 1$ & $n_2 - 1$ for the degrees of freedom.

**EXAMPLE**

The table summarizes a study on the mean resting heart rate of males and females. Perform a hypothesis test using $\boldsymbol{\alpha = 0.05}$ to determine if there's a difference in between the two randomly sampled groups. Assume normal population distributions.

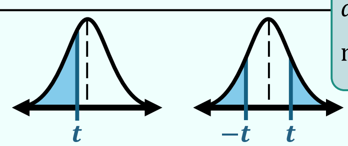
| Mean Resting Heart Rate | | | |
|---|---|---|---|
| **Group** | **Sample Size ($n$)** | **Sample Mean ($\bar{x}$)** | **Sample Std Dev ($s$)** |
| Males | 10 | 70.2 BPM | 5.8 BPM |
| Females | 11 | 81.4 BPM | 6.4 BPM |

Samples Random & Independent? ☐

$\sigma_1$ & $\sigma_2$ unknown & not equal? ☐

BOTH Samples Normal ***OR*** $n_1 > 30$, $n_2 > 30$ ☐

| | **1 Mean** | **2 Means** |
|---|---|---|
| **Hyp.** | $H_0: \mu = \#$<br>$H_a: \mu$ </>/≠ # | $H_0: \mu_1 = \mu_2$<br>$H_a: \mu_1$ **[ < \| > \| ≠ ]** $\mu_2$ |
| **Test Stat. (σ unknown)** | $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ | $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ |
| **P-Val.** | $df = n - 1$<br>Area "beyond" $t$ | $df = \min\begin{Bmatrix} n_1 - 1 \\ n_2 - 1 \end{Bmatrix}$ |
| **Conclude** | Because $P$-value **[ < \| > ]** $\alpha$, we **[ REJECT \| FAIL TO REJECT ]** $H_0$. There is **[ ENOUGH \| NOT ENOUGH ]** evidence to { restate $H_a$ } | |




**1)** $H_0$:

$H_a$:

**2)** **Males** **Females**

$n_1 =$ ___ , $n_2 =$ ___

$\bar{x}_1 =$ ___ , $\bar{x}_2 =$ ___

$s_1 =$ ___ , $s_2 =$ ___

$t =$

**3)** $df =$

$P$-value:

**4)** Because $P$-value **[ < | > ]** $\alpha$, we **[ REJECT | FAIL TO REJECT ]** $H_0$.
There is **[ ENOUGH | NOT ENOUGH ]** evidence that there is a difference in the mean resting heart rate between males & females.

**TOPIC: TWO MEANS – UNKNOWN, UNEQUAL VARIANCE**

## PRACTICE

Researchers are comparing the average number of hours worked per week by employees at two different companies. Below are the results from two independent random samples. Assuming population standard deviations are unknown and unequal, calculate the $t$-score for the difference in means, but do not find a $P$-value or state a conclusion.

**Company A:** $n_1 = 25$; $\bar{x}_1 = 22.4$ hours; $s_1 = 3.2$ hours

**Company B:** $n_2 = 16$; $\bar{x}_2 = 21.1$ hours; $s_2 = 2.9$ hours

**Recall**

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

## EXAMPLE

Use the information below to test the claim about the difference between two population means at the specified significance level. Assume samples are random and independent, normally distributed populations, and unknown and unequal population standard deviations.

Claim: $\mu_1 > \mu_2$ ; $\alpha = 0.10$

$n_1 = 32,\ \bar{x}_1 = 462,\ s_1 = 67$

$n_2 = 19,\ \bar{x}_2 = 431,\ s_2 = 85$

**TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES**

**Two Means Hypothesis Test ($\sigma$ Unknown) Using a TI-84**

◆ To perform a Hyp. Test for 2 pop. means with $\sigma$s unknown using a calculator, use the **`2-SampTTest`** function.

**EXAMPLE** Do a Hyp. Test to see if there's evidence $\mu_1 < \mu_2$ for $\alpha = 0.05$ using…

$(\boldsymbol{A})$ $\overline{x_1} = 23.75,\ s_{x_1} = 5.72,\ n_1 = 150$

$\overline{x_2} = 24.23,\ s_{x_2} = 4.91,\ n_2 = 175$

$H_0$:__________ $H_a$:__________ $P$-value: _______

Because $P$-value **[ < | > ]** $\alpha$, we **[ REJECT | FAIL TO REJECT ]** $H_0$, there is **[ ENOUGH | NOT ENOUGH ]** evidence to suggest…

---

$(\boldsymbol{B})$

| Sample Data | | | | | | |
|---|---|---|---|---|---|---|
| **Sample 1** | 13.2 | 12.4 | 15.6 | 12.98 | 19.9 | 20.11 |
| **Sample 2** | 25.3 | 13.85 | 19.1 | 22.68 | 16.7 | 24.59 |

$H_0$:__________ $H_a$:__________ $P$-value: _______

Because $P$-value **[ < | > ]** $\alpha$, we **[ REJECT | FAIL TO REJECT ]** $H_0$, there is **[ ENOUGH | NOT ENOUGH ]** evidence to suggest…

**HOW TO: Hyp. Test on TI-84 2 Means, $\sigma$ unknown**

**1)** If given data, enter in **`L1, L2`**

( STAT , **`1:EDIT…`**)

**2)** STAT > **TESTS**

**`4: 2-SampTTest`**

**3)** If given data: If given $\overline{x_1}, n_1, \ldots$:

| DATA | STATS |
|---|---|
| `List1: L1` | $\overline{x}$`1:` |
| `List2: L2` | `Sx1:` |
| `Freq1: 1` | `n1:` |
| `Freq2: 1` | $\overline{x}$`2:` |
| | `Sx2:` |
| | `n2:` |

`μ1: ≠μ2 <μ2 >μ2`

`Pooled: No Yes`

`Calculate Draw`

**TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES**

## PRACTICE

For $\overline{x_1} = 44$, $s_{x_1} = 2.3$, $n_1 = 50$, $\overline{x_2} = 40$, $s_{x_2} = 3.1$, & $n_2 = 75$, test if there is evidence that $\mu_1 > \mu_2$ for $\alpha = 0.01$ using a hypothesis test.

## PRACTICE

Perform a hypothesis test with $\alpha = 0.1$ to see if there's evidence that $\mu_1 \neq \mu_2$.

| Sample Data | | | | | | |
|---|---|---|---|---|---|---|
| Sample 1 | 11.2 | 12.3 | 13.6 | 16.32 | 12.3 | 17.81 |
| Sample 2 | 15.6 | 14.2 | 10.3 | 19.11 | 13.1 | 14.4 |

### HOW TO: Hyp. Test on TI-84 2 Means, $\sigma$ unknown

**1)** If given data, enter in **L1, L2**

(STAT, **1:EDIT…**)

**2)** STAT > **TESTS**

**4: 2-SampTTest**

**3)** If given data: If given $\overline{x_1}$, $n_1$,…:

| DATA | STATS |
|---|---|
| List1: L1 | $\overline{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\overline{x}$2: |
| | Sx2: |
| | n2: |

**μ1: ≠μ2 <μ2 >μ2**

**Pooled: No Yes**

**Calculate Draw**

**TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES**

**EXAMPLE**

A library is undergoing renovations and is unsure which section to increase, so they decide to take one sample each of children and adults and compare the average weekly reading time (in min.) for both groups. They get the following statistics:

Children: $\overline{x_1} = 125$, $s_{x_1} = 11.2$, $n_1 = 125$
Adults: $\overline{x_2} = 134$, $s_{x_2} = 12.3$, $n_2 = 135$

$(\boldsymbol{A})$ At the 0.1 significance level, test if there is enough evidence to suggest that the library's suspicion that adults read more per week is true.

$(\boldsymbol{B})$ The library will only increase the size of the adult section if there is evidence to suggest adults read for more time each week than children, otherwise, they will increase both sections. Based on the results of the test, what should they do?

**HOW TO: Hyp. Test on TI-84 2 Means, $\sigma$ unknown**

**1)** If given data, enter in **L1, L2**

(STAT, **1:EDIT…**)

**2)** STAT > **TESTS**

**4: 2-SampTTest**

**3)** If given data: | If given $\overline{x_1}$, $n_1$,…:

| DATA | STATS |
|---|---|
| List1: L1 | $\overline{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\overline{x}$2: |
| | Sx2: |
| | n2: |

**μ1: ≠μ2 <μ2 >μ2**

**Pooled: No Yes**

**Calculate Draw**

**TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES**

**EXAMPLE**

A tech retailer would like to make better stocking decisions and is curious if there is evidence that they sell more phone chargers than laptop chargers per day. They collect the following data to test the claim.

| Chargers Sold (Per Day) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Phone | 13 | 9 | 12 | 3 | 10 | 6 | 5 | 4 |
| Laptop | 11 | 5 | 14 | 7 | 8 | 4 | 8 | 7 |

(***A***) Perform a hypothesis test with a significance level of 0.01.

(***B***) Based on the results of the test above, should the retailer stock the same number of phone and laptop chargers? Why or why not?

**HOW TO: Hyp. Test on TI-84 2 Means, $\sigma$ unknown**

**1)** If given data, enter in **L1, L2**

(STAT, **1:EDIT…**)

**2)** STAT > **TESTS**

**4: 2-SampTTest**

**3)** If given data: If given $\overline{x}_1$, $n_1$,…:

| DATA | STATS |
|---|---|
| List1: L1 | $\overline{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\overline{x}$2: |
| | Sx2: |
| | n2: |

**μ1: ≠μ2 <μ2 >μ2**

**Pooled: No Yes**

**Calculate Draw**

**TOPIC: TWO MEANS – UNKNOWN, UNEQUAL VARIANCE**

**Difference in Means: Confidence Intervals**

◆ To make a Conf. Int. for $\mu_1 - \mu_2$, use point estimator $\bar{x}_1 - \bar{x}_2$ and margin of error:

New

$$E = t_{\alpha/2} \cdot \sqrt{\frac{\quad}{n_1} + \frac{\quad}{n_2}}$$

## EXAMPLE

The table summarizes a study on the mean resting heart rate of randomly sampled males and females. Make a 90% confidence interval for the difference in mean heart rate between males and females. Assume normal population distributions. What does this result suggest about the claim that there is no difference in mean resting heart rate between males & females?

| Mean Resting Heart Rate | | | |
|---|---|---|---|
| **Group** | **Sample Size ($n$)** | **Sample Mean ($\bar{x}$)** | **Sample Std Dev ($s$)** |
| Males | 10 | 70.2 BPM | 5.8 BPM |
| Females | 11 | 81.4 BPM | 6.4 BPM |

We are ______ % confident that the true difference in mean resting heart rate in males and females is between ________ & ________. Because this **[ DOES | DOES NOT ]** include 0, we **[ REJECT | FAIL TO REJECT ]** $H_0: \mu_1 = \mu_2$. There is **[ ENOUGH | NOT ENOUGH ]** evidence that there is a difference...

**HOW TO: Make a Confidence Interval for $p_1 - p_2$**

**1)** Verify samples are independent **AND** random ☐

Normal ***OR*** $n_1 > 30$, $n_2 > 30$ ☐

**2)** Find critical value: $t_{\alpha/2}$

**3)** Point estimate: $\bar{x}_1 - \bar{x}_2$

**4)** Margin of Error: $E$

**5)** Find upper & lower bounds

$\left((\bar{x}_1 - \bar{x}_2) - E, (\bar{x}_1 - \bar{x}_2) + E\right)$

**Recall**

$$df = \min\begin{Bmatrix} n_1 - 1 \\ n_2 - 1 \end{Bmatrix}$$

◆ If Conf. Int. *DOESN'T* include 0, we're confident of a *DIFFERENCE* between $\mu_1$ & $\mu_2$, so we ____________ $H_0$.

◆ If Conf. Int. *DOES* include 0, it's possible there is *NO DIFFERENCE* between $\mu_1$ & $\mu_2$, so we ____________ $H_0$.

## TOPIC: TWO MEANS – UNKNOWN, UNEQUAL VARIANCE

### PRACTICE

A researcher is comparing average number of hours slept per night by college students who work part-time versus those who don't. From survey data, they calculate $\bar{x}_1 = 6.82$ hours and $\bar{x}_2 = 6.57$ hours with a margin of error of 0.41. Should they reject or fail to reject the claim that there is no difference in hours slept between the two groups?

- **(A)** Reject
- **(B)** Fail to reject
- **(C)** There is not enough information to answer the question

## TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES

### Two Means Confidence Intervals ($\sigma$ Unknown) Using a TI-84

◆ To create a C.I. for the diff of 2 pop. means with $\sigma$s unknown using a calculator, use the `2-SampTInt` function.

**EXAMPLE** Create a C.I. for the difference between the 2 means for $\alpha = 0.10$ using…

**(A)** $\overline{x_1} = 23.05,\ s_{x_1} = 4.22,\ n_1 = 200$

$\overline{x_2} = 22.87,\ s_{x_2} = 4.87,\ n_2 = 150$

$C$-Level: ______ Confidence Interval: (______ , ______)

We are ___% sure the diff. btw. the 2 pop. means is btw. ___ & ___. The int. **[ DOES | DOESN'T ]** include ___, so there's **[ ENOUGH | NOT ENOUGH ]** evidence to suggest…

______________________________________________

**(B)**

| Sample Data | | | | | | |
|---|---|---|---|---|---|---|
| **Sample 1** | 12.2 | 12.8 | 18.7 | 16.98 | 19.1 | 20.89 |
| **Sample 2** | 24.3 | 14.57 | 18.1 | 24.88 | 19.7 | 25.52 |

$C$-Level: ______ Confidence Interval: (______ , ______)

We are ___% sure the diff. btw. the 2 pop. means is btw. ___ & ___. The int. **[ DOES | DOESN'T ]** include ___, so there's **[ ENOUGH | NOT ENOUGH ]** evidence to suggest…

**HOW TO: C.I. on TI-84**
**2 Means, $\sigma$ unknown**

**1)** If given data, enter in **L1, L2**

( STAT , **1:EDIT…**)

**2)** STAT > **TESTS**

**0: 2-SampTInt**

**3)** If given data: If given $\overline{x_1}, n_1, \ldots$:

| DATA | STATS |
|---|---|
| List1: L1 | $\overline{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\overline{x}$2: |
| | Sx2: |
| | n2: |

C-level:

Pooled: **No** Yes

**Calculate** Draw

**TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES**

### PRACTICE

For $\overline{x_1} = 29$, $s_{x_1} = 3.3$, $n_1 = 40$, $\overline{x_2} = 26$, $s_{x_2} = 3.8$, & $n_2 = 75$, create a confidence interval for the difference of the two means to test if there's evidence that $\mu_1 > \mu_2$ for $\alpha = 0.05$.

### PRACTICE

Create a confidence interval with $\alpha = 0.1$ for the difference between the 2 population means to see if there's evidence that $\mu_1 \neq \mu_2$.

| Sample Data | | | | | | |
|---|---|---|---|---|---|---|
| Sample 1 | 23.4 | 24.6 | 30.1 | 25.6 | 18.9 | 26.4 |
| Sample 2 | 31.2 | 17.5 | 22.1 | 21.8 | 25.8 | 27.5 |

**HOW TO: C.I. on TI-84**
**2 Means, $\sigma$ unknown**

**1)** If given data, enter in **L1, L2**

( STAT , **1:EDIT…**)

**2)** STAT > **TESTS**

**0: 2-SampTInt**

**3)** If given data: | If given $\overline{x_1}$, $n_1$,…:

| DATA | STATS |
|---|---|
| List1: L1 | $\overline{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\overline{x}$2: |
| | Sx2: |
| | n2: |

**C-level:**

**Pooled: No Yes**

**Calculate**

## TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES

### EXAMPLE

A restaurant chain is hiring new waitstaff. To discover which of the two locations is more in need of extra hands, they collect data on a sample of wait times at each location and get the following statistics:

Location A: $\overline{x_1} = 15.4$, $s_{x_1} = 3.2$, $n_1 = 35$
Location B: $\overline{x_2} = 18.6$, $s_{x_2} = 4.3$, $n_2 = 35$

$(\boldsymbol{A})$ For $\alpha = 0.1$, create a confidence interval to see if there is enough evidence to suggest that the chain's suspicion that Location B needs more assistance is true.

$(\boldsymbol{B})$ If there is enough evidence to suggest wait times are higher for Location B, then new employees will be stationed there. Otherwise, new employees will be evenly distributed among the two locations. Based on the confidence interval from part (A), where should new employees be stationed?

### HOW TO: C.I. on TI-84
### 2 Means, $\sigma$ unknown

**1)** If given data, enter in **L1, L2**

(STAT, **1:EDIT…**)

**2)** STAT > **TESTS**

**0: 2-SampTInt**

**3)** If given data: | If given $\overline{x_1}$, $n_1$,…:

| DATA | STATS |
|---|---|
| List1: L1 | $\overline{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\overline{x}$2: |
| | Sx2: |
| | n2: |

**C-level:**

**Pooled: No Yes**

**Calculate**

## TOPIC: TWO MEANS - UNKNOWN, UNEQUAL VARIANCES

### EXAMPLE

A manager of two hair salons is interested in seeing if their locations have different customer satisfaction ratings, so they collect samples of customer ratings (out of 10 stars) from Salon A and Salon B.

| Customer Satisfaction Ratings | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| A | 8 | 9 | 6 | 3 | 7 | 6 | 8 | 4 |
| B | 7 | 5 | 9 | 7 | 10 | 4 | 2 | 7 |

(*A*) Create a 90% confidence interval for the difference of the 2 means.

(*B*) Based on the results of the test above, can the manager conclude one location has a higher satisfaction rating? Why or why not?

### HOW TO: C.I. on TI-84
### 2 Means, $\sigma$ unknown

**1)** If given data, enter in **L1, L2**

(STAT, **1:EDIT…**)

**2)** STAT > **TESTS**

**0: 2-SampTInt**

**3)** If given data: | If given $\bar{x}_1$, $n_1$,…:

| DATA | STATS |
|---|---|
| List1: L1 | $\bar{x}$1: |
| List2: L2 | Sx1: |
| Freq1: 1 | n1: |
| Freq2: 1 | $\bar{x}$2: |
| | Sx2: |
| | n2: |

**C-level:**

**Pooled: No Yes**

**Calculate**