Pick Up the Warm Up
and work through as a group.

Warm Up
11.1 Day 2

5

Researchers were studying how playing a dancing video game impacts heart rate. They measured the heart rates (in beats per minute) of 15 subjects before they danced a song and again after they finished dancing the song. They want to use these results to estimate the average difference between before and after heart rates.

Which of these inference procedures is most appropriate?

Choose 1 answer:

(A) A z-interval for a proportion

(B) A two-sample t-interval for the difference of means

(C) A paired t-interval for the mean difference

(D) A two-sample z-interval for the difference of proportions

(E) A t-interval for slope

Warm up
11.1 Day 2

5

Researchers were studying how playing a dancing video game impacts heart rate. They measured the heart rates (in beats per minute) of 15 subjects before they danced a song and again after they finished dancing the song. They want to use these results to estimate the average difference between before and after heart rates.

Which of these inference procedures is most appropriate?

mean or proportion?

Choose 1 answer:

test or interval (to estimate)

one pop or two

paired?

(A) A z-interval for a proportion

(B) A two-sample t-interval for the difference of means

(C) A paired t-interval for the mean difference

(D) A two-sample z-interval for the difference of proportions

(E) A t-interval for slope

Warm up
11.1 Day 2

5

Researchers were studying how playing a dancing video game impacts heart rate. They measured the heart rates (in beats per minute) of 15 subjects before they danced a song and again after they finished dancing the song. They want to use these results to estimate the average difference between before and after heart rates.

Which of these inference procedures is most appropriate?

mean or proportion?

Choose 1 answer:

test or interval (to estimate)

one pop or two

paired?

(A) ~~A z-interval for a proportion~~

(B) A two-sample t-interval for the difference of means

(C) A paired t-interval for the mean difference

(D) ~~A two-sample z-interval for the difference of proportions~~

(E) A t-interval for slope

Warm up
11.1 Day 2

5

Researchers were studying how playing a dancing video game impacts heart rate. They measured the heart rates (in beats per minute) of 15 subjects before they danced a song and again after they finished dancing the song. They want to use these results to estimate the average difference between before and after heart rates.

Which of these inference procedures is most appropriate?

mean or proportion?

Choose 1 answer:

test or interval (to estimate)

(A) A ~~z-interval~~ for a proportion

(B) A two-sample t-interval for the ~~difference~~ of means

one pop or two

paired?

→ (C) A paired t-interval for the mean difference

(D) A two-sample ~~z-interval~~ for the difference of proportions

(E) A t-interval for ~~slope~~

6

A website streams movies and television shows to millions of users. Employees know that the average time a user spends per session on their website is 2 hours. The website changed its design, and they wanted to know if the average session length was longer than 2 hours. They randomly sampled 100 users and recorded their session lengths.

Which of these inference procedures is most appropriate?

Choose 1 answer:

(A) A paired t-test for the mean difference

(B) A two-sample t-test for the difference of means

(C) A t-test for a mean

(D) A z-test for a proportion

(E) A two-sample z-test for the difference of proportions

6

A website streams movies and television shows to millions of users. Employees know that the average time a user spends per session on their website is 2 hours. The website changed its design, and they wanted to know if the average session length was longer than 2 hours. They randomly sampled 100 users and recorded their session lengths.

Which of these inference procedures is most appropriate? mean or proportion?

Choose 1 answer:

test or interval (to estimate)

(A) A paired t-test for the mean difference

(B) A two-sample t-test for the difference of means

one pop or two
paired?

(C) A t-test for a mean

(D) A z-test for a proportion

(E) A two-sample z-test for the difference of proportions

6

A website streams movies and television shows to millions of users. Employees know that the average time a user spends per session on their website is 2 hours. The website changed its design, and they wanted to know if the average session length was longer than 2 hours. They randomly sampled 100 users and recorded their session lengths.

Which of these inference procedures is most appropriate? mean or proportion?

Choose 1 answer:

test or interval (to estimate)

(A) A paired t-test for the mean difference

(B) A two-sample t-test for the difference of means

one pop or two
paired?

(C) A t-test for a mean

(D) A ~~z-test~~ for a proportion

(E) A two-sample ~~z-test~~ for the difference of proportions

6

A website streams movies and television shows to millions of users. Employees know that the average time a user spends per session on their website is 2 hours. The website changed its design, and they wanted to know if the average session length was longer than 2 hours. They randomly sampled 100 users and recorded their session lengths.

Which of these inference procedures is most appropriate? mean or proportion?

Choose 1 answer:

test or interval (to estimate)

(A) A paired t-test for the mean difference

(B) A two-sample t-test for the difference of means

one pop or two
paired?

(C) A t-test for a mean

(D) A ~~z-test~~ for a proportion

(E) A two-sample ~~z-test~~ for the difference of proportions

6

A website streams movies and television shows to millions of users. Employees know that the average time a user spends per session on their website is 2 hours. The website changed its design, and they wanted to know if the average session length was longer than 2 hours. They randomly sampled 100 users and recorded their session lengths.

Which of these inference procedures is most appropriate? mean or proportion?

Choose 1 answer:

test or interval (to estimate)

(A) A ~~paired t~~-test for the mean difference

(B) A ~~two-sample~~ t-test for the difference of means

one pop or two
paired?

(C) A t-test for a mean

(D) A ~~z-test~~ for a proportion

(E) A two-sample ~~z-test~~ for the difference of proportions

HW
Lottery

Today

Perform a full Chi-Square Test for Goodness of Fit

... and do a follow up analysis

Malcolm Gladwell "Outliers"

Tries to explain strange things

~ Pennsylvania town (unusually healthy)

~ Eastern European National Soccer Team with unusual birthdays

~ Similar in Canada

Pick Up Handout

- let's read about an observation in Canada

In his book Outliers, Malcolm Gladwell suggests that a hockey player's birth month has a big influence on his chance to make it to the highest levels of the game. Specifically, because January 1 is the cut-off date for youth leagues in Canada [where many National Hockey League (NHL) players come from], players born in January will be competing against players up to 12 months younger. The older players tend to be bigger, stronger, and more coordinated and hence get more playing time, more coaching, and have a better chance of being successful. To see if birth date is related to success (judged by whether a player makes it into the NHL), a random sample of 80 NHL players from a recent season was selected and their birthdays were recorded. The one-way table summarizes the data on birthdays for these 80 players.

Birthday	Jan–Mar	Apr–Jun	Jul–Sep	Oct–Dec
Number of players	32	20	16	12

Do these data provide convincing evidence that the birthdays of NHL players are not uniformly distributed across the four quarters of the year? If there is statistically significant evidence, perform a follow up analysis.

$\boxed{\text{STATE}}$

Use $\alpha = 0.05$

the year? If there is statistically significant evidence, perform a follow up analysis.

State Hypothesies: Significance Leve

H_0: The birthdays of all NHL players are uniformly distributed across the four quarters of the year.

H_a: The birthdays of all NHL players are not uniformly distributed across the four quarters of the year.

$n = 80$

Observed frequencies

Birthday	Jan–Mar	Apr–Jun	Jul–Sep	Oct–Dec
Number of players	32	20	16	12
	20	20	20	20

Expected frequencies to be Uniform

$80\left(\frac{1}{4}\right) = 20$

Plan Name of Procedure: Chi-square test for Goodness of fit

k conditions:

Random - data came from a random sample of NHL players. ✓

10% - Assuming that $80 < \frac{1}{10}$(of all NHL Players) ✓

Large Counts - All expected Counts $= 80(\frac{1}{4}) = 20 \geq 5$ ✓

Plan Name of Procedure: Chi-square test for Goodness of fit

k conditions:

Random - data came from a random sample of NHL players. ✓

10% - Assuming that $80 < \frac{1}{10}$(of all NHL Players) ✓

Large Counts - All expected Counts $= 80(\frac{1}{4}) = 20 \geq 5$ ✓

So we can generalize to all NHL Players

Note

Large Counts condition

- ensures that the probability distribution we use (chi-square distrib. in this case) to calculate P-Value is a good model.

Do

Specific Formula

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Picture:

Work

Test Statistic

$$\chi^2 = \frac{(32-20)^2}{20} + \frac{(20-20)^2}{20} + \ldots$$

P-Value

$$= 7.2 + 0 + .8 + 3.2 =$$

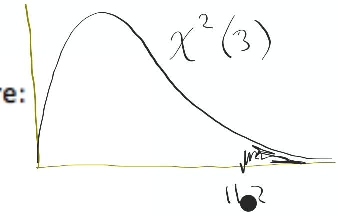
Do

$df = \text{categories} - 1 = 4 - 1 = 3$

Specific Formula

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Picture:



Work

$$\chi^2 = \frac{(32-20)^2}{20} + \frac{(20-20)^2}{20} + \ldots\ldots$$

$$= 7.2 + 0 + .8 + 3.2 =$$

Test Statistic $\chi^2 = 11.2$

P-Value

Table C → P-Value is between 0.01 and .02

or

lower upper df

$\chi^2\text{cdf}[11.2, 10000, 3]$

$= .011$

Conclude

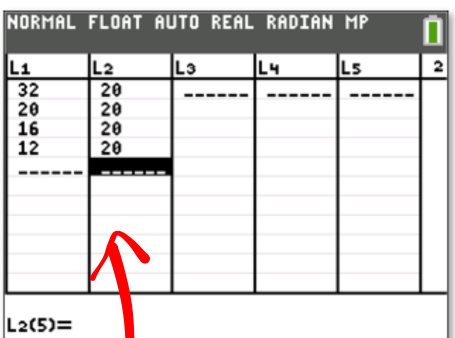
Conclude

Because the P-Value of 0.011 < α = 0.05, we reject H₀.
∴ We have convincing evidence that the birthdays of NHL players are not uniformly distributed across the four quarters of the year.

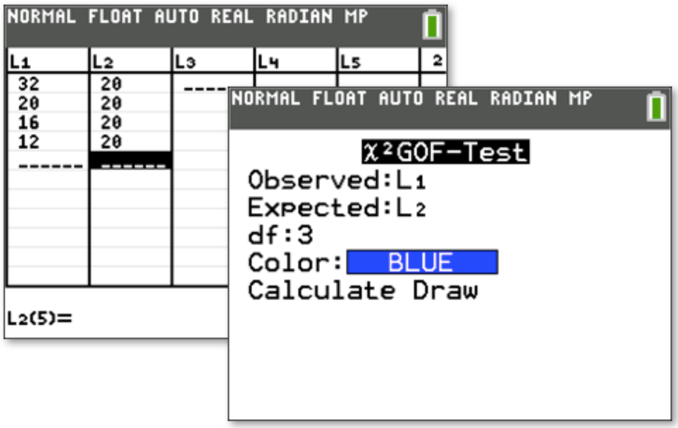
Never "accept H_0"

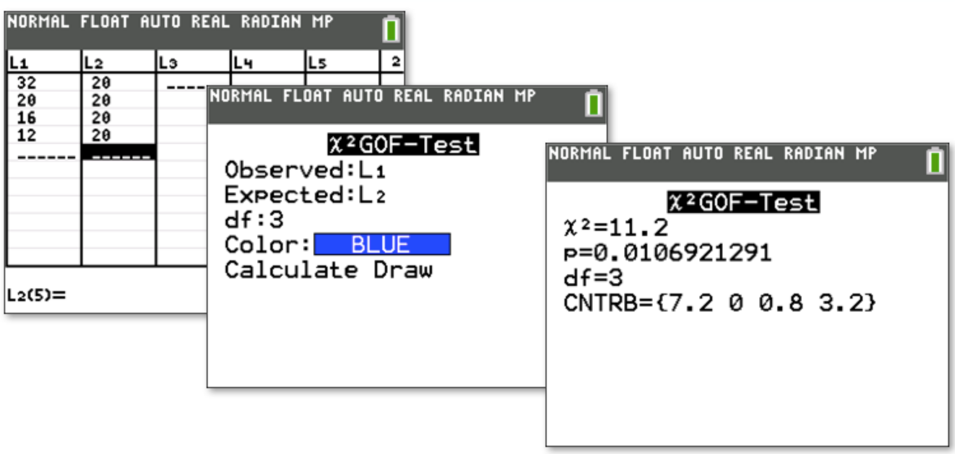
In a trial

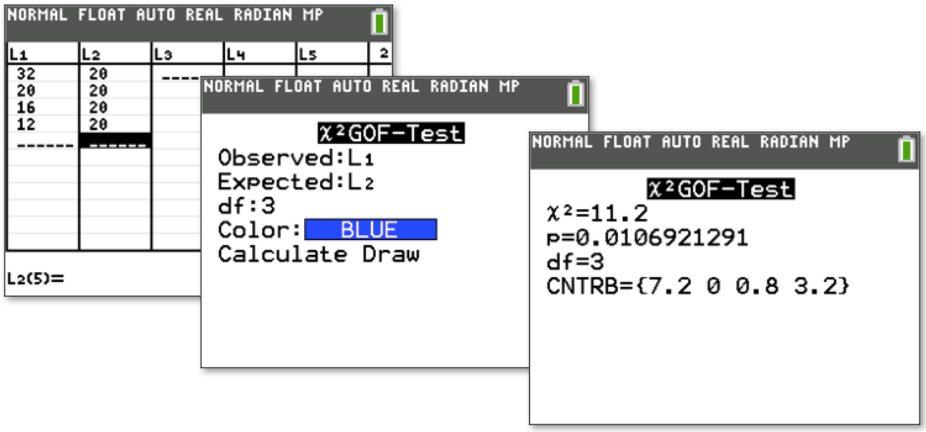
A verdict of "NOT GUILTY" doesn't mean the defendant is necessarily innocent.



enter counts


NORMAL FLOAT AUTO REAL RADIAN MP
L1 L2 L3 L4 L5 2
32 20
20 20
16 20
12 20
L2(5)=
NORMAL FLOAT AUTO REAL RADIAN MP
χ²GOF-Test
Observed:L1
Expected:L2
df:3
Color: BLUE
Calculate Draw


NORMAL FLOAT AUTO REAL RADIAN MP
L1 L2 L3 L4 L5 2
32 20
20 20
16 20
12 20
L2(5)=
NORMAL FLOAT AUTO REAL RADIAN MP
χ²GOF-Test
Observed:L1
Expected:L2
df:3
Color: BLUE
Calculate Draw
NORMAL FLOAT AUTO REAL RADIAN MP
χ²GOF-Test
χ²=11.2
p=0.0106921291
df=3
CNTRB={7.2 0 0.8 3.2}



Note: When you run the chi-square test for goodness of fit on the TI-84 calculator, a list of these individual components will be produced and stored in a list called CNTRB (for contribution).

Follow Up Analysis

If the sample data lead to a statistically significant result, we can conclude that our variable has a distribution different from the one stated.

If the sample data lead to a statistically significant result, we can conclude that our variable has a distribution different from the one stated.

To investigate *how* the distribution is different, start by identifying the categories that contribute the most to the chi-square statistic.

If the sample data lead to a statistically significant result, we can conclude that our variable has a distribution different from the one stated.

To investigate *how* the distribution is different, start by identifying the categories that contribute the most to the chi-square statistic.

Birthday	Observed	Expected	O − E	$(O-E)^2/E$
Jan–Mar	32	20	12	7.2
Apr–Jun	20	20	0	0.0
Jul–Sep	16	20	−4	0.8
Oct–Dec	12	20	−8	3.2

The two biggest contributions to the chi-square statistic came from Jan–Mar and Oct–Dec.

Birthday	Observed	Expected	O − E	$(O-E)^2/E$
Jan–Mar	32	20	12	7.2
Apr–Jun	20	20	0	0.0
Jul–Sep	16	20	−4	0.8
Oct–Dec	12	20	−8	3.2

The two biggest contributions to the chi-square statistic came from Jan–Mar and Oct–Dec.

Birthday	Observed	Expected	O − E	$(O-E)^2/E$
Jan–Mar	32	20	12	7.2
Apr–Jun	20	20	0	0.0
Jul–Sep	16	20	−4	0.8
Oct–Dec	12	20	−8	3.2

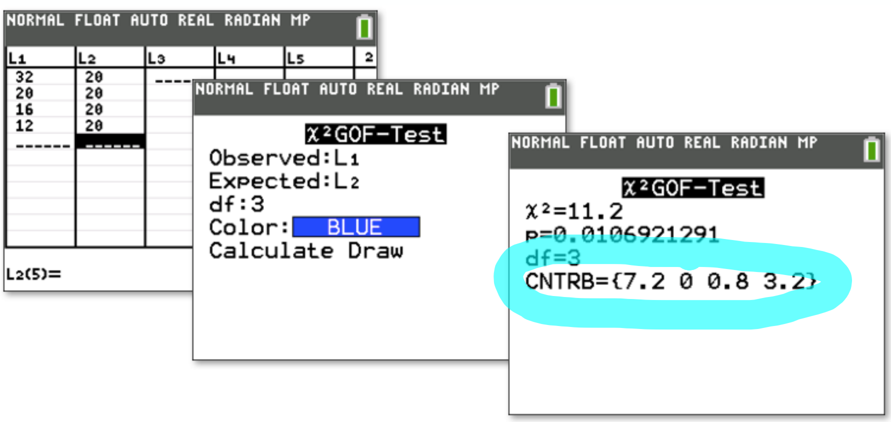
In October through December, 8 ***fewer*** players were born than expected.

The two biggest contributions to the chi-square statistic came from Jan–Mar and Oct–Dec.

Birthday	Observed	Expected	O − E	$(O-E)^2/E$
Jan–Mar	32	20	12	7.2
Apr–Jun	20	20	0	0.0
Jul–Sep	16	20	−4	0.8
Oct–Dec	12	20	−8	3.2

In October through December, 8 ***fewer*** players were born than expected.

In January through March, 12 ***more*** players were born than expected.



Note: When you run the chi-square test for goodness of fit on the TI-84 calculator, a list of these individual components will be produced and stored in a list called CNTRB (for contribution).

Lists CNTRB → L_3

AP® Exam Tip

You can use your calculator to carry out the mechanics of a significance test on the AP® Statistics exam.

But there's a risk involved. If you just give the calculator answer with no work, and one or more of your values is incorrect, you will likely get no credit for the "Do" step.

We recommend writing out the first few terms of the chi-square calculation followed by ". . .". This approach might help you earn partial credit if you enter a number incorrectly.

Be sure to name the procedure (chi-square test for goodness of fit) and to report the test statistic ($\chi^2 = 11.2$), degrees of freedom (df = 3), and *P*-value (0.011).

Car Colors
in Arizona

Car Colors in Arizona - Does the warm, sunny weather in Arizona affect a driver's choice of car color? Cass thinks that Arizona drivers might opt for a lighter color with the hope that it will reflect some of the heat from the sun. To see if the distribution of car colors in Oro Valley, near Tucson, is different from the distribution of car colors across North America, she selected a random sample of 300 cars in Oro Valley. The table shows the distribution of car color for Cass's sample in Oro Valley and the distribution of car color in North America, according to www.ppg.com.

Color	White	Black	Gray	Silver	Red	Blue	Green	Other	Total
Oro Valley sample	84	38	31	46	27	29	6	39	300
North America	23%	18%	16%	15%	10%	9%	2%	7%	100%

1. Do these data provide convincing evidence that the distribution of car color in Oro Valley differs from the North American distribution?

H_0: The distrib of car colors in Oro Valley is the same as the distrib of car colors across North America.

Ha: the distrib. of car colors in Oro Valley is not the same as the distribution of car colors across North America.

Use $\alpha = 0.05$

Observed

Color	White	Black	Gray	Silver	Red	Blue	Green	Other	Tot
Oro Valley sample	84	38	31	46	27	29	6	39	300
North America	23%	18%	16%	15%	10%	9%	2%	7%	100

Expected

$300(.23) = 69$

$300(.18) = 54$

etc

sample in Oro Valley and the distribution of car color in North America, according to www.ppg.com.

Observed

Expected

Color	White	Black	Gray	Silver	Red	Blue	Green	Other	Total
Oro Valley sample	84	38	31	46	27	29	6	39	300
North America	23%	18%	16%	15%	10%	9%	2%	7%	100%
	69	54	48	45	30	27	6	21	

1. Do these data provide convincing evidence that the distribution of car color in Oro Valley differs fro

STATE: Chi - Square test for goodness of fit

Random - Random sample of 300 cars

10% - $n = 300 < \frac{1}{10}$(all cars in Oro Valley)

Large Counts - Expected counts (69, 54, 48, 45, 30, 27, 6, 21) ≥ 5

PLAN:

STATE: Chi - Square test for goodness of fit

Random - Random sample of 300 cars

10% - $n = 300 < \frac{1}{10}$(all cars in Oro Valley)

Large Counts - Expected counts (69, 54, 48, 45, 30, 27, 6, 21) ≥ 5

PLAN:

PLAN

Chi - Square test for goodness of fit

Random - Random sample of 300 cars

10% - $n = 300 < \frac{1}{10}$(all cars in Oro Valley)

Large Counts - Expected counts (69, 54, 48, 45, 30, 27, 6, 21) ≥ 5

DO

$$\chi^2 = \frac{(84-69)^2}{69} + \frac{(88-54)^2}{54} + \ldots = 29.921 \qquad df = 8 - 1 = 7$$

P-value χ^2_{cdf}[29.921, 10000, 7] ≈ 0

DO:

$$\chi^2 = \frac{(84-69)^2}{69} + \frac{(88-54)^2}{54} + \ldots = 29.921 \qquad df = 8 - 1 = 7$$

P-value χ^2_{cdf}[29.921, 10000, 7] ≈ 0

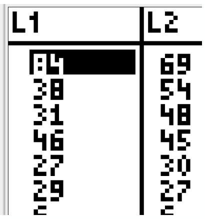
CONCLUDE: Because the P-value of approximately $0 < \alpha = 0.05$, we reject H_0

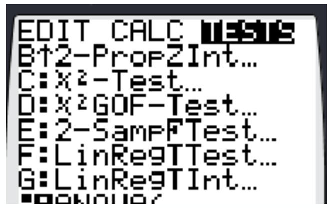
We have convincing evidence that the distrib. of car colors in Oro Valley is not the same as it is across North America.

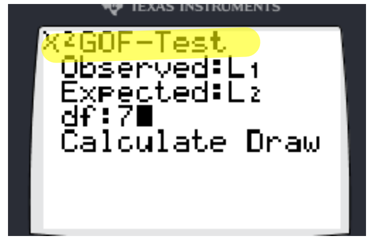
g

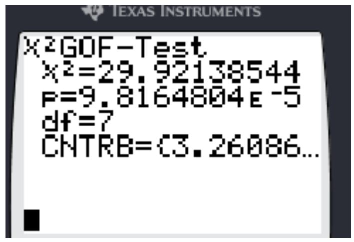
February 18, 2020

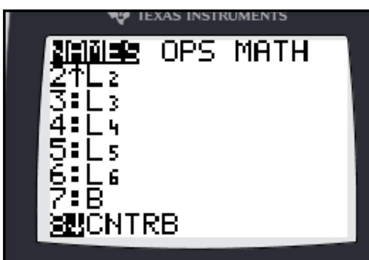
2. If there is convincing evidence of a difference in the distribution of car color, perform a follow-up analysis.

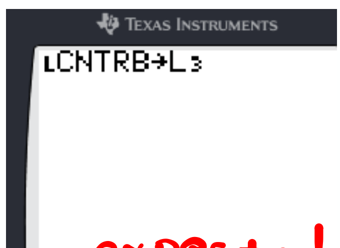






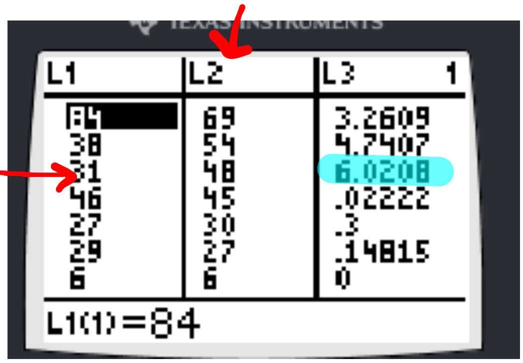






expected

gray



other

2. If there is convincing evidence of a difference in the distribution of car color, perform a follow-up analysis.

The two biggest contributions to the statistic came from gray and other colored cars. There were 17 fewer grays than expected and 10 more "other-colored" cars than expected.

See your Test

LCQ
11.1

finish LCQ (will not be accepted late)

11.1 .... 9, 13, 19-21

and study pp.717-721