

Statistics II Live Session

Aug 4, 2021

Hypothesis testing

See link from webex chat:

<https://drive.google.com/file/d/1r6b1fxh2zMPK-6AssCuq7BEHP7Tb33-3/>

PDF doc that contains the major formulae for tests.

Q1 (a) India's Covid vaccination coverage crosses 48 crore-mark, govt says

India's population > 18: 91 crores, Fraction vaccinated: $48/91 = 0.53$

Suppose you want to test the above claim. What will your null hypothesis be?

Q1 (b) India's Covid vaccination coverage crosses 48 crore-mark, govt says

India's population > 18: 91 crores, Fraction vaccinated: $48/91 = 0.53$

Suppose you suspect the above fraction is too high. What will your alternative hypothesis be?

Q1: India's Covid vaccination coverage crosses 48 crore-mark, govt says

India's population > 18: 91 crores, Fraction vaccinated: $48/91 = 0.53$

Google form: <https://forms.gle/FyJgfB7qimWAHLii8> (see webex chat)

Respond to question on vaccination status.

Fraction from survey will be shared.

Q1 (c) India's Covid vaccination coverage crosses 48 crore-mark, govt says

You need to design a test using the fraction from survey.

What are your assumptions about the data?

What is your condition for rejecting null?

What is the P-value? What is your conclusion?

Q1 (d) India's Covid vaccination coverage crosses 48 crore-mark, govt says

Suppose you desire a type-I error probability of 0.05 and power of 0.95 against the alternative that the fraction is 0.45. How many samples do you need? Use the normal approximation to the Gaussian distribution.

Q2 (a) Report in press: Average age of students in IITM's BSc programme is 24

Suppose you want to test the above claim. What will your null hypothesis be?

Q2 (b) Report in press: Average age of students in IITM's BSc programme is 24

Suppose you suspect the above is not correct. What will your alternative hypothesis be?

Q2 (c) Report in press: Average age of students in IITM's BSc programme is 24

Google form: <https://forms.gle/FyJgfB7qimWAHLii8> (see webex chat)

Respond to question on age.

Result from survey will be shared with you.

What will you compute from the data? What test will you run?

Q2 (d) Report in press: Average age of students in IITM's BSc programme is 24

What are your assumptions about the data? At a significance level of 0.05, what is your conclusion?

Q3 (a) Quiz II: Average marks in Stats II and Math II are equal

Assume that the standard deviation for marks of both subjects is 23.

Suppose you want to test the above claim. What will your null hypothesis be?

Q3 (b) Quiz II: Average marks in Stats II and Math II are equal

Assume that the standard deviation for marks of both subjects is 23.

Suppose you suspect Stats II has a higher average than Math II. What will your alternative hypothesis be?

Q3 (c) Quiz II: Average marks in Stats II and Math II are equal

Google form: <https://forms.gle/FyJgfB7qimWAHLii8> (see webex chat)

Please respond to questions on Stats II and Math II marks.

Result from survey will be shared with you.

What will you compute from the data? What test will you run?

Recall: The standard deviation for both subjects is 23.

Q3 (d) Quiz II: Average marks in Stats II and Math II are equal

Google form: <https://forms.gle/FyJgfB7qimWAHLii8> (see webex chat)

Please respond to questions on Stats II and Math II marks.

Result from survey will be shared with you.

What are your assumptions about the data? At a significance level of 0.05, what is your conclusion?

Recall: The standard deviation for both subjects is 23.

Q4 (a) Quiz II: Standard deviation of Stats II marks is 23

Suppose you want to test the above claim. What will your null hypothesis be?

Q4 (b) Quiz II: Standard deviation of Stats II marks is 23

Suppose you suspect the standard deviation is higher. What will your alternative hypothesis be?

Q4 (c) Quiz II: Standard deviation of Stats II marks is 23

Result from survey will be shared with you.

What will you compute from the data? What test will you run?

Q4 (d) Quiz II: Standard deviation of Stats II marks is 23

Result from survey will be shared with you.

What are your assumptions about the data? At a significance level of 0.05, what is your conclusion?

Q5 (a) Quiz II: Standard deviation of Stats/Math II marks are equal

Suppose you want to test the above claim. What will your null hypothesis be?

Q5 (b) Quiz II: Standard deviation of Stats/Math II marks are equal

Suppose you suspect Stats II has a lower standard deviation than Math II. What will your alternative hypothesis be?

Q5 (c) Quiz II: Standard deviation of Stats/Math II marks are equal

See results from earlier survey.

What will you compute from the data? What test will you run?

Q5 (d) Quiz II: Standard deviation of Stats/Math II marks are equal

See results from earlier survey.

What are your assumptions about the data? At a significance level of 0.05, what is your conclusion?

Q6 (a) Goodness of fit

Consider the following frequency table for olympic medals won by India from 1972 till 2016.

# medals	0	1	2	3	4	5	6	>6
fit	$(1-p)/2$	$(1-p)/2$	$p/6$	$p/6$	$p/6$	$p/6$	$p/6$	$p/6$
freq	4	5	1	1	0	0	1	0

Find an estimate of p using ML method.

Q6 (b) Goodness of fit

Consider the following frequency table for olympic medals won by India from 1972 till 2016.

# medals	0	1	2	3	4	5	6	>6
fit	$(1-p)/2$	$(1-p)/2$	$p/6$	$p/6$	$p/6$	$p/6$	$p/6$	$p/6$
freq	4	5	1	1	0	0	1	0
expected								

Is this a good fit using the chi-square test? What is the P-value?