INFERENTIAL STATISTICS

INFERENCE ⟶ MAKING CONCLUSIONS

ON L.C. COURSE WE WILL BE ASKED TO MAKE CONCLUSIONS BASED ON SOME STATISTICS. OFTEN THIS IS VERY VAGUE AND YOU CAN JUST ANSWER USING COMMON SENSE OR A FEW STANDARD RESPONSES.

THERE ARE 2 "FORMAL" METHODS WHICH WE NEED TO LEARN.

- CONFIDENCE INTERVALS
- HYPOTHESIS TESTING

ALWAYS 95% CONFIDENT

THEY ARE VERY CLOSELY RELATED.

MARGIN OF ERROR

IF WE DO A SURVEY OF 1000 PEOPLE, WE CAN BE MUCH MORE "CONFIDENT" THAT OUR RESULTS ARE ACCURATE THAN IF WE SURVEY 10 PEOPLE.

WE SAY THAT THERE IS A BIG "MARGIN OF ERROR" IF WE HAVE A SMALL SAMPLE SIZE.

BIG SAMPLE SIZE = SMALL MARGIN OF ERROR

FORMULA : (THIS IS <u>REALLY</u> IMPORTANT TO REMEMBER THIS. IT'S NOT IN THE TABLES BOOK).

MARGIN OF ERROR = $\frac{1}{\sqrt{n}}$	n = SAMPLE SIZE

eg $n = 100$ M. OF E $= \frac{1}{\sqrt{100}} = 0.1 = 10\%$

$n = 1000$ M. OF E $= \frac{1}{\sqrt{1000}} = 0.032 = 3.2\%$.

USING MARGIN OF ERROR

LEAVING CERT
ALWAYS 95% CONFIDENCE

CONFIDENCE INTERVALS:

- WHEN WE CARRY OUT A SURVEY, WE ARE NOT ASKING EVERYBODY IN THE POPULATION.
- WE ARE USING OUR SURVEY (SAMPLE) TO "ESTIMATE" / MAKE CONCLUSIONS ABOUT THE WHOLE POPULATION.
- HOW CONFIDENT CAN WE BE? THAT DEPENDS ON THE SAMPLE SIZE / MARGIN OF ERROR.
- OUR SURVEY GIVES US A "BEST GUESS" WE CALL THIS $\hat{p}$ (P HAT)
- TO GET OUR INTERVAL WE GO UP + DOWN BY THE MARGIN OF ERROR.
- WE ARE THEN 95% SURE THAT THE TRUE POPULATION PROPORTION / STATISTIC IS WITHIN THIS INTERVAL / RANGE.
- SEE EXAMPLE ON NEXT PAGE

CONFIDENCE INTERVAL (EXAMPLE)

IN A SAMPLE OF 200 STUDENTS, 48 SAID THAT THEY SPEND AT LEAST ONE HOUR EACH DAY WATCHING T.V. CONSTRUCT A CONFIDENCE INTERVAL AT THE 95% LEVEL OF CONFIDENCE.

$\hat{p}$ (BEST GUESS) $= \frac{48}{200} = 0.24$

n (SAMPLE SIZE) $= 200$

Margin of Error $= \frac{1}{\sqrt{200}} = 0.07$

STEP 1

WRITE DOWN 3 THINGS

$\hat{p}$, n , M.OF.E.

$\hat{p}$ FROM QUESTION (WRITE AS A DECIMAL)

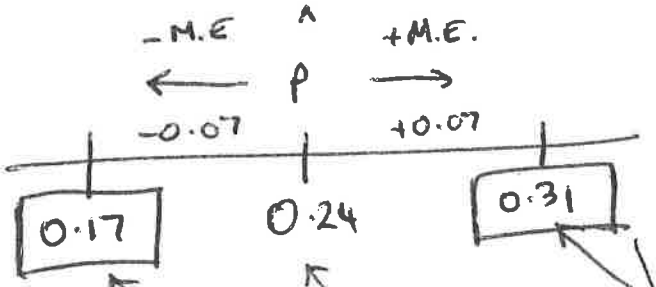
n FROM QUESTION

M.OF.E. USE FORMULA $\frac{1}{\sqrt{n}}$

REMEMBER THIS

STEP 2

DRAW CONFIDENCE INTERVAL PICTURE



STEP 3

FILL IN PICTURE

ALWAYS 95%

STEP 4

INTERPRET

WE ARE 95% SURE THAT THE PROPORTION OF STUDENTS THAT SPEND AT LEAST ONE HOUR WATCHING TV IS

BETWEEN 0.17 AND 0.31

HYPOTHESIS TESTING

OR STATEMENT WHICH HAS NOT BEEN PROVED

THIS IS ONE STEP FURTHER THAN CONFIDENCE INTERVALS. WE USE OUR CONFIDENCE INTERVAL TO ACCEPT / REJECT CLAIMS THAT COMPANIES / GOVERNMENTS etc MAKE:

eg PEPSI CLAIM THAT 75% OF PEOPLE PREFER PEPSI.

IN A SURVEY, 308 PEOPLE OUT OF 400 PREFERRED PEPSI.

TEST THEIR HYPOTHESIS AT THE 5% LEVEL OF CONFIDENCE.

"AT THE 5% LEVEL" IS ANOTHER WAY OF SAYING 95% CONFIDENCE

H_0: 75% PREFER PEPSI. ← WRITE DOWN THEIR HYPOTHESIS. WE CALL THIS "NULL HYPOTHESIS"

H_1: THE % WHO PREFER PEPSI IS NOT 75%. ← WRITE DOWN ALTERNATIVE

$\hat{p} = \frac{308}{400} = 0.77$

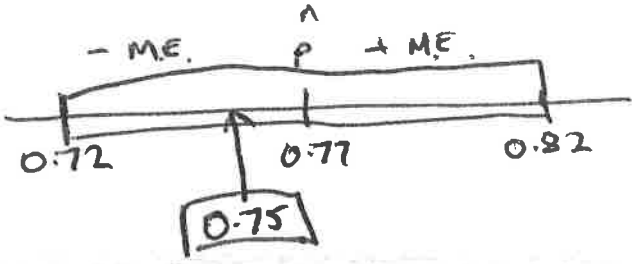
$n = 400$

$M.E. = \frac{1}{\sqrt{400}} = 0.05$

← PERFORM CONFIDENCE INTERVAL (SEE PREVIOUS PAGES)

eg 75% = 0.75

← MARK THE "CLAIM" ON THE DIAGRAM



IF THE CLAIM IS OUTSIDE THE CONFIDENCE INTERVAL, REJECT THE CLAIM. OTHERWISE DON'T REJECT.