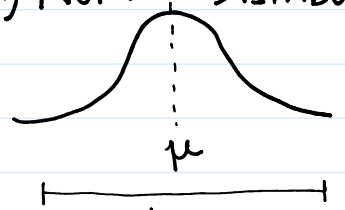
Hypothesis testing 15+16oct22

22 August 2022 08:04 PM

1) Normal Distribution. If r.v. $X \sim N(\mu, \sigma^2)$ (μ ↙ pop^n mean; σ^2 → pop^n variance)



2) Statistic → f^n of sample $(\bar{x}, s^2)$
parameter → f^n of population (μ, σ^2) [ALWAYS FIXED VALUES]

formulas: $\bar{x} = \frac{1}{n}\sum x_i$ $\quad s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$

μ → fixed const. $\quad \sigma^2$ = fixed const.

Population vs sample ✓ Theory of multiple samples and sampling distribution of the mean - lite version ✓

Problem to solve: I want to test if the average IQ score for all students enrolled in classes X-XII is 400 or not. {when pop^n variance is 300} → this info is sometimes given in Q.

Pop^n → all students of classes (X – XII)

sample → single Random sample of 'n' (=100) students is available.

Point estimate vs interval estimate

$\bar{x}$, s^2 → single points. Which are estimates of μ and σ^2.

Interval estimates give us a range in which pop^n parameter lies instead of single points. with certain degree of confidence.

$(\bar{x}-K \quad \bar{x} \quad \bar{x}+K)$ so $(\bar{x}-K, \bar{x}+K)$ is my interval estimate

Eg. I say: By our analysis we conclude that pop^n mean lies b/w (380 to 420) with 95% confidence. (400)

I'm 100% sure that $\mu \in (-\infty, \infty)$ least useful

usually taken → 95% .. .. $\mu \in (380 \text{ to } 420)$ better useful

90% .. .. $\mu \in (390 \text{ to } 410)$ better useful

Sampling distribution of the mean with test statistic

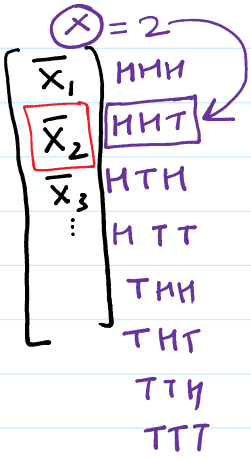
Pop^n: all students of class X - XII (size 'N')

we only have sample of size 'n'. → (sample 1)

→ (sample 2)

→ (sample 3)

⋮



Treat your $\bar{x}$ as a random variable, because $\bar{X}$ can take multiple values depending upon the sample chosen.

$\bar{X}$ is a random variable representing sample means which has taken specific value 402.5 in our particular case

INDEX	IQ score	index	sample1	index	sample2	index	sample3	index	sample4
1	315	894	362	9857	377	9613	328	1256	404
2	488	1687	500	5136	444	4764	436	3953	472
3	369	6965	362	1387	390	9450	381	9680	422
4	455	4916	300	3769	362	6854	448	6318	387
5	482	2183	402	6059	370	9384	436	3203	413
6	360	3775	318	8283	484	406	381	6380	318
7	489	4360	429	301	416	8468	348	3588	332
8	452	4902	441	6032	304	8508	302	8795	325
9	395	1788	425	3281	478	5707	479	5457	477
10	413	5072	311	7499	404	3741	367	4640	328
11	495	3882	311	47	429	3844	425	7372	488
12	493	5548	460	7998	480	6795	325	3588	332
13	494	9577	389	2585	312	2755	454	8856	334
14	469	5257	375	7091	455	7282	485	6186	342
15	420	1085	315	3506	478	3805	399	4700	363
16	384	5882	364	2333	388	6106	403	5603	462
17	405	2837	351	8210	342	6796	374	5286	384
18	439	3786	330	9682	489	6548	445	9896	329
19	317	5429	431	3444	438	8821	458	7747	345
20	476	7370	401	4792	306	712	376	8684	445
21	486								
22	409		378.85		407.3		402.5		385.1
23	392								
24	361								
25	367								
26	302								
27	487								
28	354								
29	349								
30	365								
31	367								
32	421								
33	443								
34	339								
35	306								
36	329								
37	439								
38	356								
39	489								

$$\mathbb{E}(\bar{X}) = \mu$$

Expected value of sample mean is population mean.

$\bar{X}$ random variable → continuous r.v.

THEOREM: $\bar{X}$ follows a normal distribution with (mean μ and variance σ^2/n)

so ⇒ $\bar{X} \sim N(\mu, \sigma^2/n)$ where n is sample size

↳ Sometimes given in Q [Pop^n value]

unknown, that's why we will infer about them

$\bar{X}$: random variable

$\bar{x}$: known value from our sample to be used for inference

What is hypothesis?

What is hypothesis?

Give examples

A hypothesis is a potential explanation for something that happens or that you observe and think to be true. It can also be used to determine the relationship between two or more variables that you think might be related to each other.

Hypothesis: check if population mean is 400?

A statement about the population which we need to verify based upon our analysis of a sample available

Need of hypothesis? Pop NA

Hypotheses are usually written as if/then statements, such as *if someone eats a lot of sugar, then they will develop cavities in their teeth*. These statements identify specific variables (in this case, eating a large amount of sugar) and propose a result (in this case, teeth developing cavities)

Questions about population which we need answered.

Examples of a Hypothesis

- If I replace the battery in my car, then my car will get better gas mileage.
- If I eat more vegetables, then I will be healthier.
- If I add fertilizer to my garden, then my plants will grow faster.

pop" loses weigh faster?

STEP 1: construction of hypothesis.

Types of hypothesis : H0 and H1

Null Hypothesis: Base statement which is taken by default and which could be rejected by the analysis.
denoted by H_0

$H_0: \mu = 400$ [always about pop"]

Alternative Hypothesis: Counter statement to the null hypo.

$H_1: \mu \neq 400$ [always about pop"]

Types of hypothesis simple vs composite

Simple hypotheses uses $=, \neq$ sign
composite .. .. uses $>, <, \geq, \leq$ sign

Types of errors! Type 1 and 2

eg: If you're going for a COVID-19 test.

H_0: ✓
H_1: ✓
} only one can be true.

eg. If you're going for a COVID-19 test. H_1: ✓ one can be true.

→ You have COVID → test positive

→ You don't have COVID → test negative

→ means errors and errors are possible.

REALITY \ TEST RESULT	+	−
+	correct	TYPE 2 false negative
−	TYPE 1 false true	correct

(reject H_0 when H_0 is true → TYPE 1 err)
false positive

(accept H_0 when H_0 is false → TYPE 2 err)
false negative.

H_0: $\mu = 400$: H_1: $\mu \neq 400$ we need to find truth.

1) reality: [Pop^n μ is actualy 400] UNKNOWN TRUE STATEMENT
result: [test says to go with H_1] TYPE-1 error rejecting H_0 when it's true

2) reality: [Pop^n μ is actualy 200] UNKNOWN TRUE STATEMENT
result: [test say go with H_0] TYPE-2 error accept H_0 when it was actually false.

lays example

H_0: $\mu \geq 7$ ✓ accepte H_0
H_1: $\mu < 7$

Your $\bar{x}$ score for lays comes out as 6.85 You accept H_0 error
but it turned out as false actual $\mu = 4$
when you launched flavor
TYPE 2 error.

		ANSWER GOTTEN FROM SAMPLE ANALYSIS	
		accept H_0 ✓	reject H_0 ✓
POPULATION REALITY	Statement of H_0 is true	correct answer	TYPE-1 ERROR Prob. = α
	Statement of H_0 is false	TYPE-2 ERROR Prob. = β	correct answer

SYMBOLS: α → Probability of committing a type 1 error

β → Probability of committing a type 2 error

Because actual pop^n truth is unknown, in the mechanism where we choose whether to reject of accept H_0

we will make sure that α and β are minimum. Goal is to find whether to reject H_0 or not while keeping α and β at a min. level.

Test statistic and critical region

test statistic: a function of sample values which will help us decide whether to accept H_0 or reject H_0.

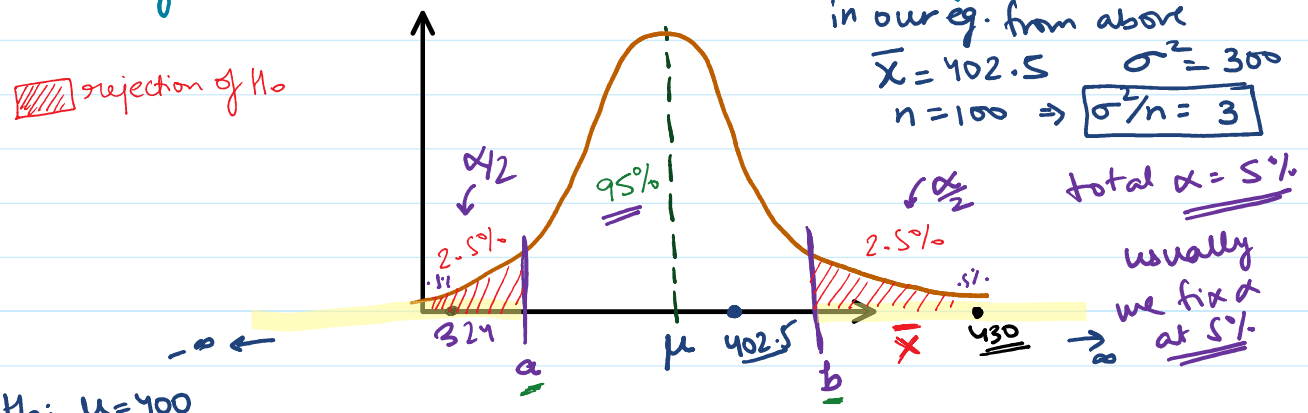
eg. sample mean.

using the theorem above: $\bar{X} \sim N(\mu, \sigma^2/n)$ [3]

in our eg. from above

$\bar{x} = 402.5$ $\sigma^2 = 300$

$n = 100 \Rightarrow \sigma^2/n = 3$



$H_0: \mu = 400$

$H_1: \mu \neq 400$

I need to decide a & b such that if $\bar{x} < a$ then reject H_0 & if $\bar{x} > b$ then reject H_0.

we will accept H_0 if our $\bar{x}$ lies b/w a, b since 402.5 lies b/w a & b we accept H_0 at 95% level of confidence.

Level of significance - allowing for some error

critical region is from $-\infty$ to a & b to ∞ where if $\bar{x}$ comes in critical region we reject H_0.

Inferential Statistics Vs Hypothesis Testing

Let's examine the fundamental distinction between hypothesis testing and inferential statistics.

Inferential statistics: When you don't have a starting point number, inferential statistics are used to find a population parameter (typically the population mean). To determine the sample mean, you begin by sampling. Then, you estimate the population mean from the sample mean using the confidence interval.

Hypothesis testing: To validate your assertion (or hypothesis) regarding the population parameter, use hypothesis testing. By conducting a hypothesis test, you can decide whether there is sufficient data to conclude that the population parameter hypothesis is true or false.

Steps to solve any problem:

1. Formulate the problem using H_0 and H_A $H_A = H_1$
2. Fetch sample, calculate statistic (under H_0)
3. Calculate tabulated and critical value
4. Decide on the basis of type 1 / 2 error or p value
5. Conclude on the basis of H_0

STEP - 1

Null Hypothesis

A null hypothesis is the default position that assumes that variables have no relation to each other.

It is denoted as H_0.

Null hypothesis (H_0): The status quo

Continuing with previous example:

Hypothesis: *"If an office provides snacks, employees will take fewer off-site breaks"*

Null Hypothesis(H_0): *The number of off-site breaks employees take is not related to the availability of food.*

Simple vs composite
1 tailed / 2 tailed

Alternate Hypothesis - Example

Find the null hypothesis' alternate result that contradicts it. Both the null hypothesis and the original hypothesis are different from the alternative hypothesis.

eg. Is this vaccine effective? test using hypothesis

Deck p12-19 onwards example

Errors in Hypothesis Testing

Alpha and beta types

We either fail to reject the null hypothesis in hypothesis testing or do so. We do not, however, always choose wisely. Every time a null hypothesis is rejected or not rejected, some error is involved.

Actual	Decision: H_0 true (Fail to reject)	Decision: H_0 false (Rejecting H_0)
H_0 true	TRUE NEGATIVE Correct decision: **Confidence level** (prob 1 - α)	FALSE POSITIVE Type I Error: **Significance level/Size (α)** (prob α)
H_0 false	FALSE NEGATIVE Type II Error: Fail to reject (prob β)	TRUE POSITIVE Correct decision: **Power** (prob 1 - β)

STEP-2

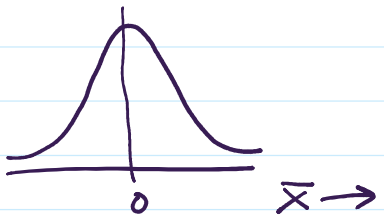
- calculating the sample statistic (say sample mean) {under H_0}
- see the sampling distribution of your statistic

$H_0 : \mu = 400$

$H_0 : \mu = 400$

$\bar{x} = 402.5$ {we got from our sample}

then $\bar{x}$ is the test statistic for 'μ'



also

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

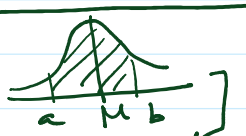
so $\bar{X} - \mu \sim N(0, \sigma^2/n)$ ∵ we shifted each value back by μ so center is 0 now

also $\dfrac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$ [subtract μ & ÷ by std. dev]

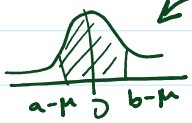
$$\Rightarrow \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$ then to calculate a, b

---

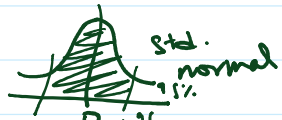
$$P[a < \bar{X} < b] = 0.95$$



$$\Rightarrow P[a-\mu < \bar{X}-\mu < b-\mu] = 0.95$$



$$\Rightarrow P\left[\frac{a-\mu}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{b-\mu}{\sigma/\sqrt{n}}\right] = 0.95$$



$$\Rightarrow P\left[l < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < r\right] = 0.95$$

$$l = \bar{x} - 1.96\frac{\sigma}{\sqrt{n}} \quad \text{and} \quad r = \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$$

If $H_0 : \mu = 400$ If this 400 lies b/w $\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}$ and $\bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$ then accept H_0

$$402.5 - 1.96\sqrt{\frac{300}{100}} \quad , \quad 402.5 + 1.96\sqrt{\frac{300}{100}}$$

$$402.5 - 3.39 \qquad 402.5 + 3.39$$

$$= (399.11) \qquad (405.89)$$

$= (399.11)$ (405.89)

with 95% confidence you can say that μ lies b/w 399.11 to 405.89

here $H_0: \mu = 400$

using this range based upon sample values what is your decision?!

ACCEPT H_0 ✓✓

(399.11, 405.89) is called a 95% confidence interval.

NORMAL DISTRIBUTION CLASS, CALCULATING PROB.

Steps to solve any problem:

1. Formulate the problem using H_0 and H_A ✓
2. Fetch sample, calculate statistic (under H_0)
3. Calculate tabulated and critical value
4. Decide on the basis of type 1 / 2 error or p value
5. Conclude on the basis of H_0

template to follow for each question

set $\alpha = 5\%$ ← given in Q

CONCEPTS OVER PRACTISE START

STEP 1

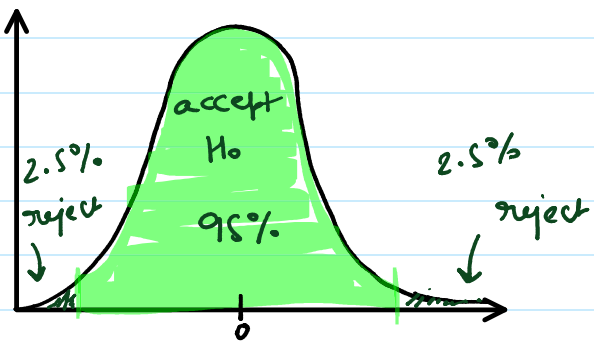
$H_0: \mu = 400$

$H_A: \mu \neq 400$

STEP 2

sample statistic: $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

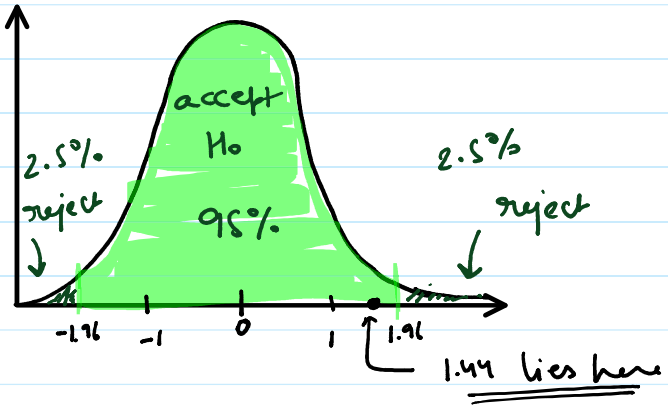
$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$ [2 steps to get std. normal]



STEP 3

You will have some $\bar{x}$ from sample.
convert that into Z-score: $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ eg. $\frac{402.5-400}{\sqrt{300/100}}$

$= 1.44$



since 1.44 lies in the acceptance region we accept H_0

STEP 4:

Critical values: Points which are boundary of acceptance region:

here critical values: ± 1.96

since calculate z (1.44 here) lies b/w -1.96 to 1.96

It lies in acceptance region.

STEP-5

Making a decision

Examples:

1. If the average commute time is at least thirty minutes, then $H_0 \geq 30$ and $H_1 < 30$ indicate the test is lower-tailed because the critical region will be on the left side of the distribution.

1. The test is upper-tailed if the average commute time is no more than 30 minutes, in which case $H_0 \leq 30$ and $H_1 > 30$, indicating that the critical region will be on the right side of the distribution.

1. Given that the critical region will be on both sides of the distribution and the average commute time is 30 minutes, the test is two-tailed if $H_0=30$ and $H_1 \neq 30$.

The **Critical Value Method** or **p-Value Method** is utilised to determine the critical values for the critical region.

Critical Value Method ✓

The critical value method requires the following steps to be followed to reach a decision:

1. Create the hypothesis, identify **H_0 and H_1**, and confirm the test that needs to be performed.

Critical Value Method ✓

The critical value method requires the following steps to be followed to reach a decision:

1. Create the hypothesis, identify **H_0 and H_1**, and confirm the test that needs to be performed.
1. Using the value of as a starting point, calculate the value of z-Critical Value (Zc) (Significance Level).
1. Based on the test, determine the critical values (UCV and LCV) from Zc.
1. The sample mean(x) for the critical values is the foundation for the decision.

↙ given in Q

usually we use 95% confidence limits ∵ $\alpha = 5\%$

If $\alpha = 1\%$ we will 99% confidence limits

critical values for 99% will be

p-Value in Hypothesis Testing

The probability that the null hypothesis will not be rejected can be summed up as the p-Value. It is for clarity purposes only and is not the official definition.

The formal definition is: *"The p-value is the probability of obtaining test results at least as extreme as the results observed, under the assumption that the null hypothesis is correct".*

The steps we must take to choose the null hypothesis using the p-value method are as follows:

1. Create the hypothesis, identify H_0 and H_1, and confirm the test that needs to be performed.
2. Determine Z's value for the sample mean.
3. Utilizing the z-table, determine the p-value for the specified z-score.
4. Based on the p-value for the specified value of α, make a decision (significance)

Using the z-score, determine the p-value.

[accept H_0 if $p > \alpha$ / reject H_0 if $p < \alpha$] to remember

p-value : P(getting a more extreme sample)

$P(Z >$ test statistic$)$

Demo p28onwards

Demo p33 onwards

} read from deck has another example

STEP - 5

- decision and conclusion

Practice Question -1

According to the manufacturer, the average lifespan of a product is 36 months. An auditor determines the product's average life to be 34.5 months after selecting 49 of the product samples. The four-month population standard deviation. Utilizing the critical value method, test the manufacturer's claim at a 3% significance level.

Practice Question -1 (Solution)

Step-1: Formulate the Hypothesis

H_0:μ = 36 months and

H_1: $\mu \neq$ 36 months

Step-2: Z-critical value

For a 3% significance level, you would have two critical regions on both sides with a total area of 0.03. So, the area of the critical region on the right side would be 0.015, which means that the area till UCV (cumulative probability of that point) would be 1 - 0.015 = 0.985. The z-score for 0.9850 in the z-table is 2.17.

Step-3: The critical values can be calculated from $\mu \pm Zc \times (\sigma/\sqrt{N})$ as $36 \pm 2.17(4/\sqrt{49}) = 36 \pm 1.24$ which comes out to be 37.24 and 34.76.

Step-4: The critical values for this test are 37.24 and 34.76. The sample mean in this case is 34.5 months, which is less than lower critical values. So this implies that the sample mean lies in the critical region, and you can reject the null hypothesis.