Course
MATH2361 Probability and Statistics
(Common for BSC BTECH, CSE, ECE, BBA)
@GITAM, deemed to be University

Unit-IV Curve fitting

UNIT-IV **Testing of Hypothesis and Large Sample Tests**

Dr Malikarjuna Reddy Doodipala
MSc, M.Phil, PGDCA, Ph D,
GITAM Deemed to be University
Best Teacher Awardee (State level)
Best Researcher Awardee (PLOS One J)
Philip M.Morse National tech.Awardee in OR
Out standing informative book award (Int.)

Testing of Hypothesis

Parametric Measures (Tests)

Large sampling Tests

Z-Test

- Mean(s),
- Standard Deviations
- Proportion(s))
- Correlation Coefficient

Small Sampling Tests

t-Test

- Mean(s)
- Paired means
- Correlation

Chi-square-test

- Variance
- Goodness
- Attributes

F-test

- Two-Variances
- ANOVA

Software

- MS Excel
- SPSS
- Graph pad prism
- Minitab
- R
- SAS
- Stata
- Micro soft Excel
- MATLAB
- Python etc.

Prerequisites

- One is supposed to have the knowledge of
  - Descriptive Statistics
  - Population, Sample, Parameter
  - Sampling Distribution
  - Estimate or Statistic
  - Estimator

Terms and notations

- Introduction to Hypothesis Testing
- Statistical Hypotheses
- Critical/Rejection Region
- Types of Errors
- Level of significance
- Power of the test
- Procedure of Hypothesis Testing Problem
- Sample Problems

Population, Sample, Parameter

- The field of statistical inference consists of those methods used to make decisions or draw conclusions about a population from the sample.
- A population consists of the totality of the observations (finite or infinite) which we are concerned. Each observation in a population is a value of some numeric variable X having some distribution (ex.Normal)
- These methods utilize the information contained in a sample from the population in drawing conclusions.
- A sample is a subset of a population.

Population, Sample, Parameter Cont'd….

- **Parameter**: Any statistical constant in population and it is unknown quantity it is necessary to estimate by statistic.
- **Sampling Distribution-** A sampling distribution is a distribution of all the possible values of a statistic for a given size sample selected from a population.
- There is a procedures for making inferences:

  - Statistical Hypothesis
- The purpose of hypothesis testing is to determine whether there is enough statistical evidence in favor of a certain belief about a parameter.

Estimator and Estimate

- **Statistic**: function of random samples or group of observations.
- Any sample statistic that is used to estimate a population parameter is called **estimator**.

  i.e. An estimator is a sample statistic to find parameter.
- An **estimate** is a specific observed value of a statistic. In other words every specific value of an estimator is an estimate.

Examples:

Unknown Parameter θ	Statistic $\hat{\Theta}$	Point Estimate $\hat{\theta}$
μ	$\bar{X} = \frac{\sum X_i}{n}$	$\bar{x}$
σ^2	$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$	s^2
p	$\hat{P} = \frac{X}{n}$	$\hat{p}$
$\mu_1 - \mu_2$	$\bar{X}_1 - \bar{X}_2 = \frac{\sum X_{1i}}{n_1} - \frac{\sum X_{2i}}{n_2}$	$\bar{x}_1 - \bar{x}_2$
$p_1 - p_2$	$\hat{P}_1 - \hat{P}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2}$	$\hat{p}_1 - \hat{p}_2$

For example : The S.E

- The standard error of a statistic is the SD of a sampling distribution
- If the S.E involves unknown parameters /quantities whose values can be estimated, substitution of these estimates into the standard error results in an estimated SE

$$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

Hypothesis Testing

- The purpose of hypothesis testing is to determine whether there is enough statistical evidence in favor of a certain belief about a parameter.
- **Examples**

1. Is there statistical evidence in a random sample of potential customers, that support the hypothesis that more than 75% of the potential customers will purchase a new products?

2. 2.5% of people suffering from other symptoms in post covid period

Statistical Hypothesis

- Example from Class Age Data
- Null Hypothesis: B.Tech CSE Students are juniors (age 19). Mean age=19
- Alternative Hypothesis: Age doesn't matter and the mean age does not equal to19.
- We'll reject the hypothesis if mean is "too low" or "too high"

A statistical Hypothesis is a statement about parameters of one or populations

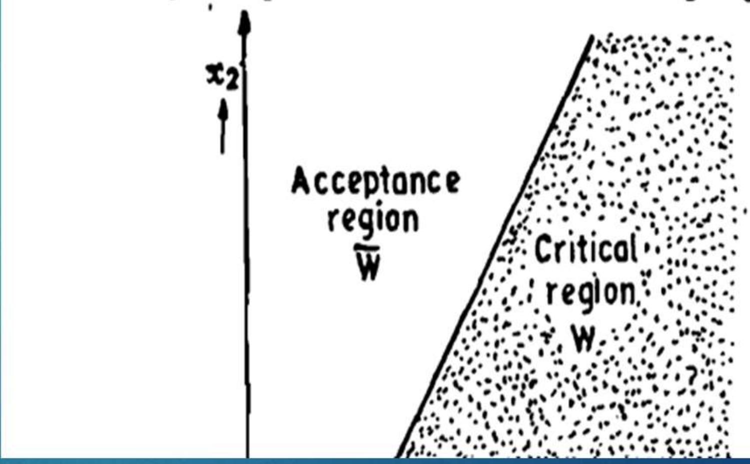
Hypothesis Testing

- **Null hypothesis –** Statement (any) regarding the value(s) of unknown parameter(s).
- Typically will imply no association between explanatory and response variables in our applications (will always contain an equality)
- **Alternative hypothesis** - Statement contradictory to the null hypothesis (will always contain an inequality)
- **Test statistic -** Quantity based on sample data and null hypothesis used to test between null and alternative hypotheses
- **level of significance** – a is known as level of significance or level (l.o.s) in other words it is rejection chance of statistical Null hypothesis for our testing problem

Critical & Acceptation regions

- The whole sample region (population) classified into two complimentary regions say W & Wc
- If the Xi fall in W region & Ho is rejected then the region is Critical
- If the Xi no fall in W region & Ho is accepted then the region is called acception region (Wc)

Suppose if the test is based on a sample of size 2, then the outcome set or the sample space is the first quadrant in a two-dimensional space and a test criterion will enable us to separate our outcome set into two complementary subsets, W and $\bar{W}$. If the sample point falls in the subset W, H_0 is rejected, otherwise H_0 is accepted. This is shown in the following diagram :



Types of Errors

- **Type I Error:**

  -Rejecting H_0 when it is actually true

  -Concluding a difference when one does not actually exist

- **Type II Error:**

  -Accepting H_0 when it is actually false (e.g. previous slide)

  -Concluding no difference when one does exist

	True State	
Test Result	**H_0 True**	**H_1 True**
Reject H_0	Type-I-Error	Correct Decision
Accept H_0	Correct Decision	Type-II-Error

Cont'd...

$$\alpha = P(\text{Type I Error}) = P(\text{Reject } \mathrm{H}_0 \text{ it is true})$$
$$\beta = P(\text{Type II Error}) = P(\text{Accept } \mathrm{H}_0 \text{ it is false})$$

Here Keep α, β reasonably small

In first case Alpha is known as level of significance or level (l.o.s) in other words it is rejection chance of statistical Null hypothesis for our testing problem.

Power of the test

- In general Probability of correct decision is known as power of the test and its denoted by 1-β

Where

$$\beta = P(\text{Type II Error}) =$$

$$P(\text{Accept } \mathrm{H}_0 \text{ it is false}) = \int_W L_1 dx$$

$$\alpha = P(\text{Type I Error}) = \int_{Wc} L_o dx$$

L_o, L_1 likely hood functions under Null and Alternate hypothesis respectively

Sample Problem:

- Let P be the probability that a coin will fall head in a toss and H0=p=1/2 against H1 p=3/4.

If coin is tossed five times and if H0 is rejected if more than 3 heads are obtained.

Find the probability of type I & II errors. Also find power of the test.

Given that

$\mathrm{H}_0: p = 1/2 \quad \mathrm{H}_0: p = 3/4$

$w = \{x: x > 3\}$

then $\bar{\mathrm{w}} = \{x: x \leq 3\}$ We know that X is binomial here.

$\mathrm{p(x)} = \mathrm{n}_{c_x} p^x q^{n-x}$

$n = 5$

Sample Problem

- Since a coin is tossed five times and H0 is rejected if more than 3 heads are obtained.
- Now to find the probability of type I & II errors.
- We have

$$\alpha = P(\text{Type I Error})$$
$$= P(\text{Reject } \mathrm{H}_0 \text{ it is true})$$
$$= \mathrm{p}\{x : x > 3/\mathrm{H}_0 : p = 1/2\}$$
$$\sum_{4}^{5} p(x) = 3/16$$
$$\text{then } \bar{\mathrm{w}} = \{x : x \leq 3\}$$
$$\beta = P(\text{Type II Error})$$
$$= P(\text{Accept Ho it is false})$$
$$= \mathrm{p}\{x : x \leq 3/\mathrm{H1} : p = 3/4\}$$
$$= \sum_{0}^{3} p(x) = 47/128$$
$$1 - \beta = 1 - 47/128 = 81/128$$

2.Sample Problem:

- A discrete r.v X follows Poisson distribution with parameter λ,
- H0 is rejected if x > 3 to test the null hypothesis : λ=1 v/s alternative hypothesis: λ=2 find the probability of type I & II errors.

Hint:

$w = \{x : x > 3 / \mathrm{H}_0 : \lambda = 1\}$

$\bar{w} = \{x : x \leq 3 / \mathrm{H}_1 : \lambda = 2\}$

Ans : 0.004

0.94

0.06

3.Sample Problem

Let X follows N(μ, 4).1l unknown.
To test Ho : μ = -1'
against Hl : μ = 1.
Based on a sample of size 10 from this population.
we use the critical region $X_1 + 2X_2 + \dots + 10X_{10} \geq 0$
1.What IS its size?
2.What is the power of the test?

Hint:
$w = \{x : X_l + 2X_2 + \dots + 10X_{10} \geq 0\}$

3.Sample Problem: Solution

Solution. Critical Region $W = \{\mathbf{x} : x_1 + 2x_2 + \ldots + 10x_{10} \geq 0\}$.

Let $U = x_1 + 2x_2 + \ldots + 10x_{10}$

Since x_i's are *i.i.d.* $N(\mu, 4)$,

$$U \sim N\left[(1 + 2 + \ldots + 10)\,\mu,\ (1^2 + 2^2 + \ldots + 10^2)\,\sigma^2\right] = N(55\mu, 385\sigma^2)$$

$$\Rightarrow \quad U \sim N(55\mu, 385 \times 4) = N(55\mu, 1540) \qquad \ldots(*)$$

The size 'α' of the critical region is given by :

$$\alpha = P(\mathbf{x} \in W \mid H_0) = P(U \geq 0 \mid H_0) \qquad \ldots(**)$$

Under $H_0 : \mu = -1,\ U \sim N(-55, 1540)$

3.Sample Problem: Solution

$$\Rightarrow \qquad Z = \frac{U - E(U)}{\sigma_U} = \frac{U + 55}{\sqrt{1540}}$$

$$\therefore \text{ Under } H_0, \text{ when } U = 0,\ Z = \frac{55}{\sqrt{1540}} = \frac{55}{39{\cdot}2428} = 1{\cdot}4015$$

$$\begin{aligned} \therefore \quad \alpha &= P\,(Z \geq 1{\cdot}4015) && \text{[From (**)]} \\ &= 0{\cdot}5 - P(0 \leq Z \leq 1{\cdot}4015) \\ &= 0{\cdot}5 - 0{\cdot}4192 && \text{(From Normal Probability Tables)} \\ &= 0{\cdot}0808 \end{aligned}$$

Alternatively, $\alpha = 1 - P(Z \leq 1{\cdot}4015) = 1 - \Phi(1{\cdot}4015)$,
where $\Phi\,(\cdot)$ is the distribution function of standard normal variate.

3.Sample Problem: Solution

Power of the test is given by :

$$1-\beta = P(x \in W \mid H_1) = P(U \geq 0 \mid H_1)$$

Under $H_1 : \mu = 1,\quad U \sim N(55, 1540)$

$$\Rightarrow \quad Z = \frac{U - E(U)}{\sigma_U} = \frac{-55}{\sqrt{1540}} = -1{\cdot}40 \qquad \text{(when } U = 0\text{)}$$

$$\begin{aligned} \therefore \quad 1-\beta &= P(Z \geq -1{\cdot}40) \\ &= P(-1{\cdot}4 \leq Z \leq 0) + 0{\cdot}5 \\ &= P(0 \leq Z \leq 1{\cdot}4) + 0{\cdot}5 \qquad \text{(By symmetry)} \\ &= 0{\cdot}4192 + 0{\cdot}5 \\ &= 0{\cdot}9192 \end{aligned}$$

Alternatively,

$$1-\beta = 1 - P(Z \leq -1{\cdot}40) = 1 - \Phi(-1{\cdot}40).$$

Level of Significance and the Rejection Region:

One-Sided Tests:

$$H_0: \mu = 3$$

$$H_1: \mu < 3$$

$$H_0: \mu = 3$$

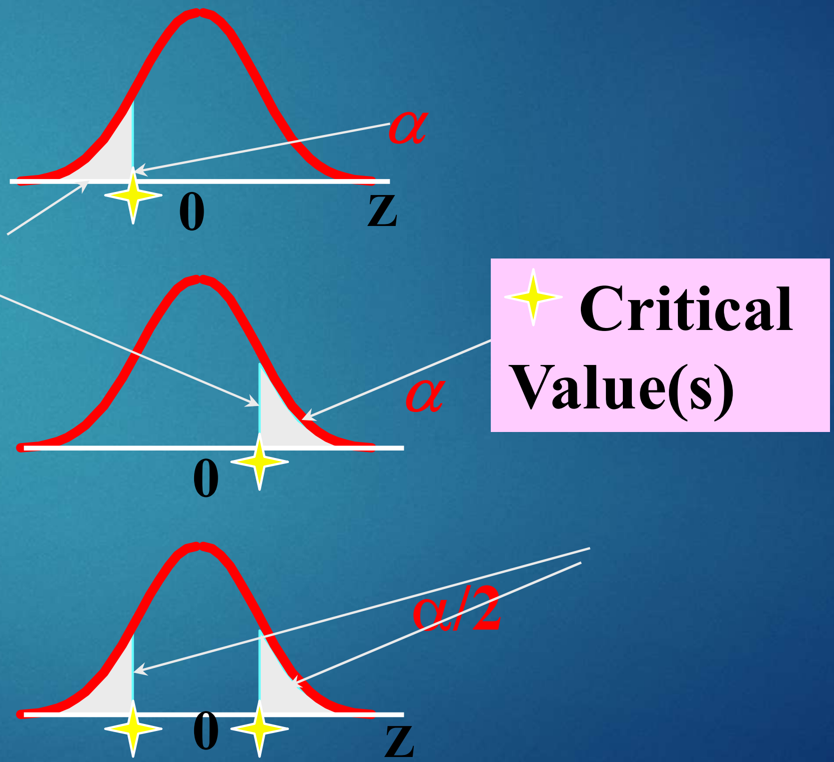
$$H_1: \mu > 3$$

Two-Sided Test:

$$H_0: \mu = 3$$

$$H_1: \mu \neq 3$$

Rejection Regions



Hypothesis Testing Procedure

$$\alpha$$

- **To solve the given hypothesis testing problem which consists the following five steps.**
- **Step 1.** set up the statistical null and alternate hypothesis.
- **Step 2.** List out the sample data which is available from the problem
- **Step3.** Use an appropriate test statistic formula.(by central limit theorem)
- **Step 4.** Using sample data, compute the test statistic and compare with its critical or table value.
- **Step 5.** give the conclusion at given l.o.s Alpha

Testing of Hypothesis (Traditional Approach)

- Choose a data sample from the population
- Assume whether the data is normally distributed or not
- Set up a null & alternative hypothesis that is

  $H_0: \mu =$ specified value

  $H_1: \mu \neq$ specified value, $\mu >$ specified value, $\mu <$ specified value
- Choose an alpha or significance level at 5% or 1%
- Use the test statistic(formula) Z-test, t-test etc.
- Decide the critical value : critical value is the value of test statistics which separates acceptance region from rejection region.
- Form a decision rule computation of test statistic value
- Conclusion or decision

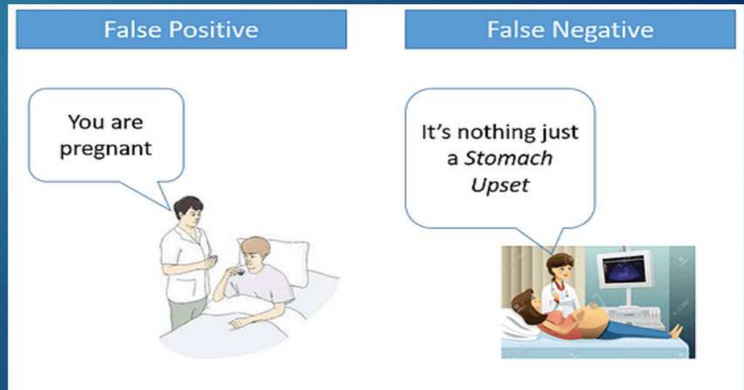
$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Role of P- value in testing of Hypothesis

- The P-value "likely" or "unlikely" by computing the probability – assuming that null hypothesis is true
- If P-value is small, say less than (or equal to) α, then it is "unlikely." it is "likely."
- If P-value is less than (or equal to) α, then the null hypothesis is rejected in favor of the alternative hypothesis.
- P-value is greater than α, then the null hypothesis is not rejected

		Fact (The Truth)	
Our Prediction (Model)		Ho is True	Ho is False
	Ho is True	Correct Decision (True Positive)	Type II Error (False Negative)
	Ho is False	Type I Error (False Positive)	Correct Decision (True Negative)



When test results fail (Image : quanthub.com)

Sampling

- It is a technique of drawing a sample from the population (business) to know the characteristics under study such as mean ,variance or s.d and proportion etc
- **Sample size:** Sample size is represented by lower case letter ‘n’
- If the sample size n is >30 large enough.
- If the sample size n is ≤ 30 small enough.

Sampling distribution

- A sampling distribution is a distribution of all the possible values of a statistic for a given size sample selected from a population.
- In the large sample case, the sampling distribution of sample statistic will be normally distributed, and we can use the z statistic.
- In the small sample case, the sampling distribution of sample statistic will be approximately normally distributed, and we can use the t, chi-square and F-statistic etc.

Sampling Distributions

- Sampling Distributions
  - Sampling Distribution of Sample Mean
  - Sampling Distribution of Sample Variance
  - Sampling Distribution of Sample Proportion

Sampling Tests

- **Large Sample Tests**

  - Sample size is > 30

- **Small Sample Tests**
- **-Sample size is <=30**

1. Z-test for Mean(TEST FOR MEAN)

Aim: This test is used to test the hypothetical population mean (or) random sample is drawn from the same population

Step 1. Set up the statistical null and alternate hypothesis.

Step 2. List out the sample data which is available.

Step 3. use the appropriate test statistic. Since, n is large by CLT statement Test for mean statistic

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

where $\frac{\sigma}{\sqrt{n}} = \mathrm{SE}$

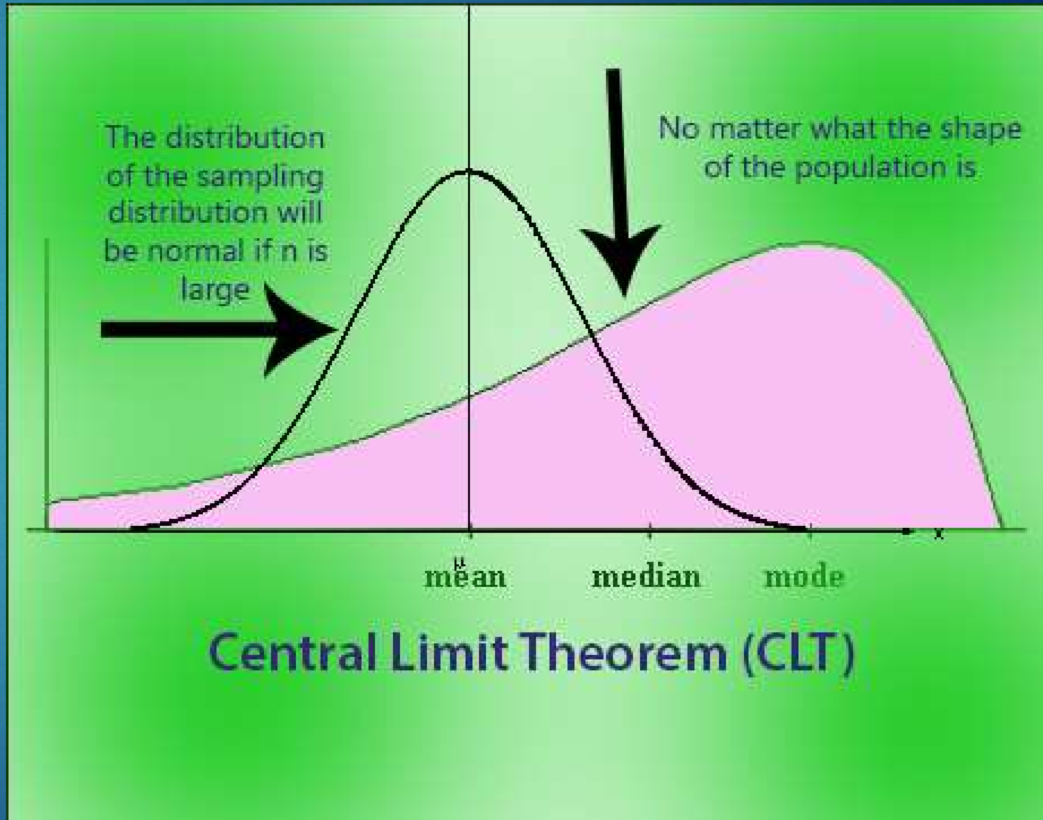
Step 4. Compute the test statistic

Step 5. Give the conclusion at given l.o.s. α

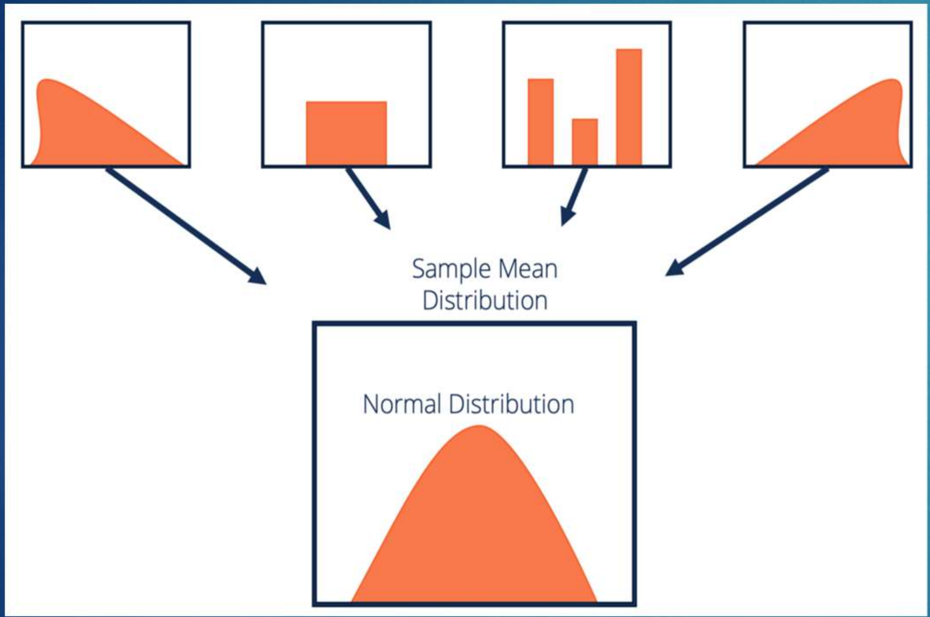
i.e If cal value <= table value accept Null hypothesis. Otherwise, we reject it.

Central limit Theorem (brief)

- The **central limit theorem** states that if you have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population with replacement , then the distribution of the sample means will be approximately normally distributed



Central limit Theorem (brief)



What exactly does the CLT say? — Theory

$$S_n = X_1 + \cdots + X_n \qquad Z_n = \frac{S_n - n\mu}{\sqrt{n}\,\sigma} \qquad Z \sim N(0,1)$$

Central Limit Theorem: For every z: $\lim_{n\to\infty} \mathbf{P}(Z_n \le z) = \mathbf{P}(Z \le z)$

- CDF of Z_n converges to normal CDF
- results for convergence of PDFs or PMFs (with more assumptions)
- results without assuming that the X_i are identically distributed
- results under "weak dependence"
- proof: uses "transforms": $\mathbf{E}[e^{sZ_n}] \to \mathbf{E}[e^{sZ}]$, for all s

Confidence Interval for a Mean

- **If you have a" large” sample….**
- A confidence interval for a population mean is:

$$\bar{x} \pm Z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$$

- where the average, standard deviation, and n depend on the sample, and Z depends on the confidence level.

Critical Value of Z (table)

Probability(Alpha or l.o.s)	Critical value	α
One tail	z_{α}	**Two Tail**
0.05	**1.65**	**0.10 (10%)**
0.025	**1.96**	**0.05 (5%)**
0.01	**2.33**	**0.02 (2%)**
0.005	**2.58**	**0.01 (1%)**

From a class of college 55 students are selected at random in a particular geographic region, the mean and standard deviation of scores on a reading test are 96 points, and 12 points, respectively. Test the hypothesis that scores of in a particular school who received a mean score of 100 whole, or are their scores surprisingly low at 5% los?

- Sol: we are given that $\bar{x}$= 96 and $\sigma = 12$

Problem

Solution

- Null hypothesis H0: $\mu = 100$,
  alternateH1: hypothesis $\mu < 100$
- We begin by calculating the standard error of the mean: $SE = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{55}} = \frac{12}{7.42} = 1.62$
- Next we calculate the z-test value, $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$
- which is the distance from the sample mean to the population mean in units of the standard error:
- Cal value $z = \frac{M - \mu}{SE} = \frac{96 - 100}{1.62} = -2.47$

Cont'd..

Now compare this with its critical or table value at 5% level (From Z- table. it is 1.96)

Here Zcal (2.47) value > Ztab(1.96) value

We reject our hypothesis.

We conclude that the scores of in a particular school who received a mean score of 100 whole, or

their scores surprisingly low at 5% los

Example: Interval estimators (For mean)

A Random sample of 59 students spent an average of Rs.273.20 on textbooks. Sample standard with deviation was 94.40. find CI for Mean at 5% LoS.

$$\mu = \overline{x} \pm Z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}} \right) =$$

$$273.20 \pm 1.96 \left(\frac{94.4}{\sqrt{59}} \right) = 273.20 \pm 24.09$$

=(249.11 , 297.29)

We can be 95% confident that the average amount spent by all students was between 249.11 and 297.29.

2.Test for Difference of Means Z Test (Variances Known)

- **Aim: This test is used to test the hypothetical population means of two populations (or) random sample is drawn from the same population**
- **Assumptions**
  - **Samples are randomly and independently drawn from normal distributions(Populations)**
  - **Population variances are known**
  - **Statistical Hypothesis:**

$$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 - \mu_2 \neq 0$$

$$\text{or} \quad H_0: \mu_1 - \mu_2 = 0 \quad H_1: \mu_1 \neq \mu_2$$

- **Test Statistic**

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Z-test for Difference of Means

A random sample of 40 mothers gave birth to children with an average weight of 2.8 kg, at the time of birth in hospital A in an year with s.d of 0.6kg and another random sample of 60 mothers gave birth to children with an average weight of 3.2 kg, at the time of birth in hospital B in same year with s.d of 0.8kg.

Test whether the hypothesis there is any significance difference b/w the Weights of the two groups of mothers at 5%los

Ho:there is no significance difference b/w the Weights of the two groups of mothers at 5%los
H1:there is significance difference b/w the Weights of the two groups of mothers at 5%los

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

$$Z = \frac{2.8 - 3.2}{\sqrt{\frac{(0.6)^2}{40} + \frac{(0.8)^2}{60}}} = \frac{0.4}{0.14} = 2.86$$

Z cal = 2.86

at $\alpha = 0.05$ Z tab = 1.96

Now compare cal Value with tab Value

Zcal > Ztab

Sampling Tests for Attributes

- Test for Proportion(P):

  -This is used to test the hypothetical population proportion

- Difference of Proportions(P1-P2)

  - This is used to test the difference of hypothetical population proportions of two distinct populations

- Sample Proportion(p):

$$p = \frac{\text{No of individuals have an attribute(x)}}{\text{Sample size(n)}}$$

$$i.e \quad p = \frac{x}{n} \quad \text{and } q = 1 - p$$

$$Z = \frac{p - P}{\sqrt{p(1-p)/n}}$$

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

***P* -** is population proportion of successes

X- is the number of successes in sample

n- is the sample size

p = ***x*/*n***, is observed(sample) proportion of successes

When $np \geq 5$ and $n(1-p) \geq 5$, use ***Z*** test

Drug has been shown to provide relief for 70% of all patients. A competing new drug provides relief for 231 of 300 randomly selected patients. Is New Drug more effective than old drug at α=0.01? Or is there any difference between old and new drug.

Solution:

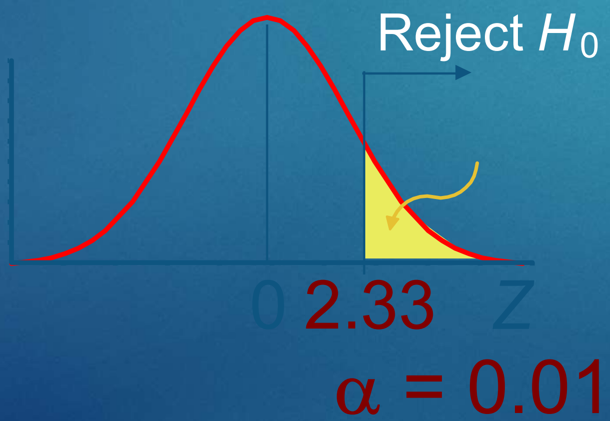
H_0: p =.70

H_1: p > .70

P_s = 231/300=.77

n = 300

Critical Value (s):



Test Statistic:

$$Z = \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{.77 - .70}{\sqrt{\frac{.70(1-.70)}{300}}} = 2.65$$

Decision: $Z_{cal} > Z_{tab}$

Reject H_0 at α = .01

Conclusion:

There is evidence that new drug is more effective than old drug.

Significance test for $p_1 - p_2$

The test statistic

- The test statistic has a standard form:

$$z = \frac{(p_1 - p_2)}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$P = p_1 + p_2/n_1 + n_2, Q = 1 - P$$

The null hypothesis is

$H_{o:}$ There is no difference between the population proportions.

This means that any difference we observe is due to random chance.

i.e $H_{o:}\ p_1 - p_2 = 0$

$H_{1:}\ p_1 - p_2$ is not 0

•(Choose Alpha here.) a=5%

- Where P is the overall weighted average or combined population proportion.
- This means we are assuming equal proportions in the two populations.

Conclusion:

- Use the sample data compute Z- Test Value
- Now compare This Value With its table value of z-test Value at given
- If Z cal-Value<Z tab-value at given l.o.s accept Null Hypothesis other wise it is rejected

Sample Problem

- In a busy city a survey was conducted to construct a fly over near by their residences : “Do you favor or not to the proposal? on two independent populations male and female With the following data ‘.’”
- Population favor to the proposal Total(n)

Population	favor to the proposal	Total(n)
male	122	345
Female	268	1900

- Test the hypothesis that the male and female population are equally favor to the same proposal at 5%level.

Significance test for $p_1 - p_2$: Example

Assumptions: sample size large enough that the sampling distribution of P_1 - P_2 is approximately normal, independent groups

1. **Hypothesis: H_o: $P_1 - P_2 = 0$**
2. **Test statistic:**

$$z = \frac{(p_1 - p_2)}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$P = p_1 + p_2/n_1 + n_2, Q = 1 - P$$

z = (122/345 – 268/1900) /

SQRT[(390/2245)*(1 - 390/2245)*(1/345 + 1/1900)]

= 9.59

3. **Critical value : Z-value: at 0.05 it is 1.96 from table thus Z-cal > z-tab**
4. **conclusion:**

reject H_0: Male and female population have difference in opinion to construct flyover near by their residences.

Comparisons of two independent population proportions: Confidence Interval

- confidence interval:

$$C.I = (P_1 - P_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

$$p_1 = \frac{x}{n_1}, p_1 = \frac{y}{n_2}$$

- Notice that there is <u>no</u> overall weighted average or

  Combined population proportion as there is in a significance test for proportions.

Any Questions? Suggestions?

Thank you

Feedback to

drmallikreddyd@gmail.com