Ch.16 Statistical Inference Testing of Hypothesis

Define hypothesis

Any assumption about the population parameter which may or may not be true is known as hypothesis. There are two types' simple hypothesis and composite hypothesis

Simple hypothesis

A hypothesis which specify all the parameters of the distribution is called simple hypothesis

Example i) $H_0 : \mu = 20$ ii) $H_0 : \mu_1 = \mu_2$

Composite hypothesis

A hypothesis which does not specify all the parameters of the distribution is called composite hypothesis

Example i) $H_1 : \mu > 20$ ii) $H_1 : \mu_1 < \mu_2$

Define testing of hypothesis.

It is a procedure of accepting or rejecting a statement about a population parameter on the bases of the given sample information. It is also called tests of hypothesis.

Statistical hypothesis

An assumption about the unknown value of population parameter from which the sample is drawn and it may be true or false is called statistical hypothesis.

Define exact and inexact hypothesis.

Exact hypothesis

A hypothesis which takes one value of the parameter such as $H_0 : \mu = 20$ / A simple hypothesis must be an exact one while an exact hypothesis is not necessary a simple hypothesis.

Inexact hypothesis

A hypothesis which takes more than one value of the parameter such as $H_0 : \mu > 20$ .An inexact hypothesis is a composite hypothesis.

Differentiate between null hypothesis and alternative hypothesis?

Null hypothesis

Any assumption about the population parameter which is to be tested for the possible rejection under the assumption that is true is known as null hypothesis. It is denoted by H_0

Alternative hypothesis

Any other hypothesis which is different from null hypothesis is known as an alternative hypothesis. It is denoted by H_1

Differentiate between type-1 error and type-11 error?

Type-1 error

Reject H_0 when H_0 actually true then it is called type-1 error. It is also called level of significance. It is denoted by α .

Example:

Suppose a person is innocent, while judge declares him guilty. In this situation, the judge has committed type-1 error.

Type-11 error

Accept H_0 when H_0 actually false then it is called type-11 error. It is denoted by β .

Example

Suppose a person is guilty, while judge declares him innocent. In this situation, the judge has committed type-11 error.

Write down the comparison between α and β .

The comparison between α and β are given below.

i) The value of α commonly used 1% or 5%.

ii) β Depends both H_0 and H_1

iii) When α becomes smaller than β tends to become larger.

iv) When α becomes larger than β tends to become smaller

v) There is an inverse relationship between α and β

vi) We can both decreases α and β by increasing sample size.

$H_0 : \mu = 32$

$H_1 : \mu \neq 32$

Differentiate between one-Tailed test and two-tailed test?

One-Tailed test

When rejection region is taken one end of the distribution, then test is called one sided or one tailed test.

a) It is right sided test

$H_0: \mu \le 32$

$H_1: \mu > 32$

b) It is left sided test

$H_0: \mu \ge 32$

$H_1: \mu < 32$

Two-tailed test

When rejection region is taken on both ends of the distribution, then test is called two sided or two tailed test.

$H_0: \mu = 32$

$H_1: \mu \ne 32$

Differentiate between acceptance region and rejection region?

Ans: Acceptance region

The part of the sampling distribution of test-statistic which leads toward the rejection of the alternative hypothesis is known as assumption region. The assumption region of H_0 is denoted by $1-\alpha$. The non-shaded area of the distribution is called acceptance region.

Rejection region

The part of the sampling distribution of test-statistic which leads toward the rejection of the null hypothesis is known as rejection region. The acceptance region of H_1 is denoted by α. The shaded area of the distribution is called rejection region. It is also called critical region.

What is meant by test-statistic?

Ans: The formula which provides the base whether to accept or reject any assumption about the population parameter which may or may not be true is known as test-statistic.

Explain the terms level of significance..

Ans: The probability of reject H_0 when H_0 actually true then it is called type-1 error. It is also called level of significance. It is also called α risk. It is also called size of critical region. It is also called size of test. It is denoted by α.

Define level of confidence.

Ans: It is probability of accept H_0 when H_0 is true it is called level of confidence It is denoted by $1-\alpha$

i.e. $1-\alpha = P(Accepting\ H_0 / H_0\ is\ true)$

What is meant by critical value?

Ans: The value which separates the rejection and acceptance region is called the critical value of the test-statistic.

Define power of test.

Ans: It is probability of rejecting H_0 when H_0 is false it is power of test it is denoted by $1-\beta$

i.e. $1-\beta = P(rejecting\ H_0 / H_0\ is\ \ false)$

Define power of curve.

Ans: A graph of the probability of rejecting H_0 for all possible values of the population parameter not satisfying the null hypothesis is known as power of curve.

What type of signs used in any alternative hypothesis?

Ans: In which any one sign we used in alternative hypothesis

a) ≠ b) > c) <

What type of signs used in any null hypothesis?

Ans: In which any one sign we used in null hypothesis

a) = b) ≥ c) ≤

Write down the null and alternative hypothesis in each of the following cases.

i) The mean height of college students is 65 inches.

ii) A certain diet is supposed to increase the mean weight by no more than 5 kgs.

iii) The proportion of people favoring a certain candidate in fort-coming election is at least 60%.

Ans:

i) $H_0: \mu = 60$ And $H_1: \mu \ne 60$

ii) $H_0: \mu \le 60$ And $H_1: \mu > 60$

iii) $H_0: p = 60\% = 0.60$ And $H_1 p \ne 60\% = 0.60$

Formulate alternative hypothesis fro each of the following null hypothesis.

i) $H_0: \mu = 170$ And $H_1: ?$

ii) $H_0: \mu \le 30$ And $H_1: ?$

iii) $H_0 : \mu \geq 40$ And $H_1 : ?$

Ans: Null and alternative hypothesis

i) $H_0 : \mu = 170$ And $H_1 : \mu \neq 170$

ii) $H_0 : \mu \leq 30$ And $H_1 : \mu > 30$

iii) $H_0 : \mu \geq 40$ And $H_1 : \mu < 40$

Explain test of significance.

Ans: The method which make possible, by using sample observations either to accept or reject the null hypothesis at a level of significance that is not already given but decided according to the situation of problem, are called tests of significance.

Define $\alpha(alpha)$ and $\beta(beta)$.

Ans: $\alpha(alpha)$: indicates the probability of committing type-1 error

$\alpha = P(Rrjrct\ H_0 / H_0\ is\ true) = P(Type - 1errior)$

$\beta(beta)$: indicates the probability of committing type-11 error

$\beta = P(Accept\ H_0 / H_0\ is\ \ false) = P(Type - 11errior)$

Differentiate between hypothesis and statistical hypothesis

In general "the hypothesis" may be a statement, which may or may not be true about a phenomenon. But a statistical hypothesis is a statement about the population parameters, which may or may not be true. It means that in statistics we are only concerned with the population parameters; the quantities of interest.

Points to remember incase of hypothesis testing

a) Whether the sample size large or small?

b) Whether the population variability (variance/ error variance) is known or unknown?

c) Whether the shape of the population are normal or non normal?

i) Sample size is already given with the problem

ii) We know or assume the population to be normal

iii) We observe whether the population variance is given or not.

What are the major considerations for tests of significance?

The choice among various tests of significance depends on

i) The sample size

ii) Information about the population variability (variance)

a) If the variance of the population σ^2 is known "Z-test" is appropriate (irrespective of the normality of the population and the sample size).

b) If the variance of the population σ^2 is not known and sample size is large (n>30) "Z-test" is appropriate because the estimate of the population variance (s^2) from large sample is a satisfactory (unbiased) estimate of the population variance σ^2. (Since $E(s^2) = \sigma^2$)

c) If the variance of the population σ^2 is not known and sample size is small ($n \leq 30$) "t-test" is appropriate provided that the population is normal.

When "Z-test and t-test are recommended in hypothesis testing or interval estimation?

a) Z-test can be applied in the following conditions

i) If the variance of the population σ^2 is known "Z-test" is appropriate (irrespective of the normality of the population and the sample size).

ii) If the variance of the population σ^2 is not known and sample size is large (n>30) "Z-test" is appropriate because the estimate of the population variance (s^2) from large sample is a satisfactory (unbiased) estimate of the population variance σ^2.

b) t-test is applicable only, if the population variance is unknown and sample size is small $n \leq 30$ provided that the given population is normally distributed.

Which type of error does the level of significance refer?

A level of significance is the probability of rejecting the hypothesis when it is true also type one error committed, when we reject the null hypothesis, when it is true. So, levels of significance refer type one error.

Under what conditions can we not use the normal distribution but can use the "t: distribution to find confidence intervals for the unknown population mean?

When the sample size "n" is smaller than 30, and the population standard deviation "σ" is not known, we cannot use the normal distribution for determining confidence intervals for the unknown population mean but we can use the students "t" distribution.

What is the relationship between the "t" distribution and the standard normal distribution?

The "t" distribution is bell-shaped and is symmetrical about its mean zero. But if flatter or platy kurtic than the standard normal distribution so that more of its area falls within the tails. While there is only one standard normal distribution, there is different (t) distribution for each sample size "n" However, as "n" becomes larger, the (t) distribution approaches the standard normal distribution until, when $n \geq 30$, they are approximately equal.

Which test based on Z distribution or normal distribution

i) Testing the hypothesis about the single population mean

a) When $n \geq 30$ *and* σ^2 *known*n b) When $n \geq 30$ *and* σ^2 *known*n

ii) Testing the hypothesis about the equality or difference of two population means

a) When $n_1, n_2 \geq 30$ *and* σ_1^2 *and* σ_2^2 *known and unequal*

b) When $n_1, n_2 \geq 30$ *and* σ_1^2 *and* σ_2^2 *known and equal*

c) When $n_1, n_2 \geq 30$ *and* σ_1^2 *and* σ_2^2 *unknown and unequal*

iii) Testing the hypothesis about the equality of two population mean when sample means are dependent or correlated

iv) Testing the hypothesis about the single population variance or standard deviation when $n \geq 30$

v) Testing the hypothesis about the equality of difference of two variances or SD when $n_1, n_2 \geq 30$

vi) Testing the hypothesis about the single proportion $n \geq 30$

vii) Testing the hypothesis about the equality or difference of two population proportion

a) When $n_1, n_2 \geq 30$ *and* P_1 *and* P_2 *unknown and unequal*

b) When $n_1, n_2 \geq 30$ *and* P_1 *and* P_2 *unknown and equal*

viii) Testing the hypothesis about the equality or difference of two population proportion when sample proportion are correlated

ix) Testing the hypothesis about the single population correlation coefficient

x) Testing the hypothesis about the equality or difference of two population correlation coefficients

Important points

Null hypothesis H_0	Alternative hypothesis H_1
Coin is unbiased Drug is ineffective in curing a particular disease The difference b/w two teaching methods are null or zero Not significantly different Insignificant At least $H_0 : \mu \geq \mu_0$ At most $H_0 : \mu \leq \mu_0$ Not more than $H_0 : \mu \leq \mu_0$ Not less than $H_0 : \mu \geq \mu_0$ No effect At least $H_0 : \mu = \mu_0$ Fair $H_0 : \mu = \mu_0$	Coin is biased Drug is effective The difference b/w two teaching methods are null or zero Significantly different Significant Better $H_1 : \mu > \mu_0$ Improved $H_1 : \mu > \mu_0$ Benifitted $H_1 : \mu > \mu_0$ Detericted $H_1 : \mu < \mu_0$ Above $H_1 : \mu > \mu_0$ Exceeds $H_1 : \mu > \mu_0$ More than $H_1 : \mu > \mu_0$ Taller $H_1 : \mu > \mu_0$ Increased $H_1 : \mu > \mu_0$ Biased $H_1 : \mu \neq \mu_0$ Fair $H_1 : \mu \neq \mu_0$ Raise $H_1 : \mu > \mu_0$ Higher $H_1 : \mu > \mu_0$ Inferior $H_1 : \mu < \mu_0$ Reduce $H_1 : \mu < \mu_0$

Q.16.5 (b): The proportion of families buying milk from company A in a certain city is believed to be p = 0.6.If a random samples of 10 families shows that 3 or less buy milk from company A, we shall reject the hypothesis that p = 0.6 in favour of the alternative p < 0.6.Evaluate α if p =0.6.Evaluate β for the alternatives p = 0.3, p = 0.4 and p = 0.5.

Solution:

The null and alternative hypothesis given as

$H_0 : P = 0.6$ and $H_1 : P < 0.6$

Let "X" denote the number of families buying milk from company "A".

Then the test-statistic is the binomial distribution with P=0.6 and n=10

The rejection region as given consists all the values from "X=0 to X=3"

Thus the probability of making type 1-error i.e α consists of $P(X \le 3)$

Hence

$$\alpha = P(X \le 3 \, when \, P = 0.6 \, and \, n = 10) = \sum_{X=0}^{3} b(x;10,0.6) = 0.0548 = from \, binomial \, probability \, tables$$

To compute β, the probability of type II-error, we need a specific alternative hypothesis

We are given $H_0 : P = 0.6$ and $H_1 : i) \; P = 0.3 \quad ii) \; P = 0.4 \quad iii) \; P = 0.5$

$$\beta = P(4 \le X \le 10 \; when \, H_1 : P = 0.3) = \sum_{X=4}^{10} b(x;10,0.3) = 1 - \sum_{X=4}^{10} b(x;10,0.3) = 1 - 0.6496 = 0.3504 \; from \, b.prb \, table$$

$$\beta = P(4 \le X \le 10 \; when \, H_1 : P = 0.4) = 1 - \sum_{X=4}^{10} b(x;10,0.4) = 1 - 0.3823 = 0.6177 \; from \, binomial \; prob \; table$$

$$\beta = P(4 \le X \le 10 \; when \, H_1 : P = 0.5) = 1 - \sum_{X=4}^{10} b(x;10,0.5) = 1 - 0.1719 = 0.8281 \; from \, binomial \; prob \; table$$

Q: 16.6(a): Define Type I and Type II-Errors in testing hypothesis. A normal distribution known to have a variance of 2.8.A one-tailed (increase) test is proposed of the form H_0:μ ≤ 14 versus H_1: μ > 14.Find the probability of making a type II error () with a sample size 2 if the significance level of the test is (i) 0.05, (ii) 0.01, when the true population mean is 16.5.

Solution: Given that $H_0 : \mu \le 14, H_1 : \mu > 14$,n=2, $\sigma^2 = 2.8$ and $\alpha = 0.05, 0.01$

First we find the vale of $\bar{x}$, the critical point which would lead to rejection of H_0 by the test statistic $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

i) The critical value of "Z" for one tailed test at $\alpha = 0.05$ from the normal curve area is 1.645. thus

$$1.645 = \frac{\bar{X} - 14}{\sqrt{2.8}/\sqrt{2}} \qquad hence \quad \bar{X} = 15.95$$

To compute β, the probability of type II-error, we use the $H_1 distribution \, N(16.5, \frac{2.8}{2})$ thus

$$\beta = (type \, II \; error \, / \, \mu = 16.5) = P(\bar{x} < 15.95) = P(Z < \frac{15.95 - 16.5}{\sqrt{\frac{2.8}{2}}}) = P(Z < -0.46) = 0.3228$$

ii) The critical value of "Z" for one tailed test at $\alpha = 0.05$ from the normal curve area is 1.645. Thus

$$1.645 = \frac{\bar{X} - 14}{\sqrt{2.8}/\sqrt{2}} \qquad hence \quad \bar{X} = 15.95$$

To compute β, the probability of type II-error, we use the $H_1 distribution \, N(16.5, \frac{2.8}{2})$ thus

$$\beta = (type \, II \; error \, / \, \mu = 16.5) = P(\bar{x} < 15.95) = P(Z < \frac{15.95 - 16.5}{\sqrt{\frac{2.8}{2}}}) = P(Z < -0.46) = 0.3228$$

iii) The critical value of "Z" for one tailed test at $\alpha = 0.01$ from the normal curve area is 2.33. Thus

$$2.33 = \frac{\bar{X} - 14}{\sqrt{2.8}/\sqrt{2}} \qquad hence \quad \bar{X} = 16.75$$

To compute β, the probability of type II-error, we use the $H_1 distribution\, N(16.5, \frac{2.8}{2})$ thus

$$\beta = (type\,II\ error / \mu = 16.5) = P(\bar{x} < 16.75) = P(Z < \frac{16.75 - 16.5}{\sqrt{\frac{2.8}{2}}}) = P(Z < 0.21) = 0.5832$$

Q.16.6(b): Given H_0: $\mu \geq 200$, H_1:$\mu < 200$, n = 100, $\alpha = 0.023$, and $\sigma = 25$.

(i) For what values of the sample mean $\bar{x}$ will H0 be accepted?

(ii) Compute β if μ is actually 191

(iii) What is the power of the test? What does it mean?

Solution:

i) To find the value of $\bar{x}$ the critical value of "Z" for one tailed test at $\alpha = 0.023$ from the normal curve area is -2.0. Thus

$$-2.0 = \frac{\bar{X} - 200}{25/\sqrt{100}} \qquad hence \quad \bar{X} = 195$$

ii) To compute β, the probability of type II-error, we use the $H_1 distribution\, N(191, \frac{625}{100})$ thus

$$\beta = (type\,II\ error / \mu = 191) = P(\bar{x} > 195) = P(Z > \frac{195 - 191}{\sqrt{\frac{625}{100}}}) = P(Z > 1.6) = 0.5 - 0.4452 = 0.0548$$

iii) Power of test $= 1 - \beta = 1 - 0.0548 = 0.9452$

16.7 c) Describe the general procedure for testing a hypothesis about a population parameter.

Solution: The procedure for testing a hypothesis about a population parameter involves the following six steps

i) State your problem and formulate an appropriate null and alternative hypothesis

ii) Choose the level of significance, α of the test, which is the probability of rejecting the null hypothesis if it is true.

iii) Choose an appropriate test-statistic, determine and sketch the sampling distribution of the test-statistic, assuming H_0 is true.

iv) Determine the rejection or critical region in such a way that the probability of rejecting the null hypothesis H_0, if it is true. It is depend on alternative hypothesis two decide one sided or two sided.

v) Compute the value of test-statistic on the basis of sample data on the basis we accept or reject H_0.

vi) Formulate the decision rule as follow

a) Reject the null hypothesis if our calculated value falls in critical region.

b) Otherwise accept H_0

Q: 16.8: Based on a sample of 25 observations from a normal population with $\sigma = 3$, the hypothesis H_0: $\mu = 67$ against H_1:$\mu > 67$ is tested at 5 % level of significance. Compute the probabilities of committing type-I errors, β and the powers of the test, when alternative hypothesis of 68.5, 68.0, 67.5 and 66 are used.

Solution:

Given $H_0 : \mu = 67$, $H_1 : \mu > 67$, $\alpha = 0.05$ and $\sigma = 3$. Since $H_1 : \mu > 67$, so we find using one sided test, the critical value for the decision rule as

$$C = \mu_0 + 1.645 \frac{\sigma}{\sqrt{n}} = 67.99$$

Since "4" values in H_1 are specified, so we associate the variables $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ and $\bar{x}_4$ with each of the four H_1-distribution. Using the H_1-distribution with $\mu = 68.5$ we calculate the value of β, the probability of type II error (say $\beta_{\mu=68.5}$) as

$$\beta_{\mu=68.5} = P(Type\ II\ error / \mu = 68.5)$$

$$\beta_{\mu=68.5} = P(\bar{x} < 67.99) = P(Z < \frac{67.99 - 68.5}{3/5}) = P(Z < -0.85) = P(-\infty < Z < 0) - P(-0.85 < Z < 0) = 0.1977$$

Using the H_1-distribution with $\mu = 68.0$ we calculate the value of β, the probability of type II error (say $\beta_{\mu=68.0}$) as

$$\beta_{\mu=68.0} = P(Type\ II\ error / \mu = 68.0)$$

$$\beta_{\mu=68.0} = P(\bar{x} < 67.99) = P(Z < \frac{67.99-68.0}{3/5}) = P(Z < -0.02) = P(-\infty < Z < 0) - P(-0.02 < Z < 0) = 0.4920$$

Using the H_1-distribution with $\mu = 67.5$ we calculate the value of β, the probability of type II error (say $\beta_{\mu=67.5}$) as

$$\beta_{\mu=67.5} = P(Type\ II\ error / \mu = 67.5)$$

$$\beta_{\mu=67.5} = P(\bar{x} < 67.99) = P(Z < \frac{67.99-67.5}{3/5}) = P(Z < 0.82) = P(0 < Z < 0.82) + P(0 < Z < \infty) = 0.7939$$

Using the H_1-distribution with $\mu = 66$ we calculate the value of β, the probability of type II error (say $\beta_{\mu=66}$) as

$$\beta_{\mu=66} = P(Type\ II\ error / \mu = 66)$$

$$\beta_{\mu=66} = P(\bar{x} < 67.99) = P(Z < \frac{67.99-66}{3/5}) = P(Z < 3.32) = P(0 < Z < 3.32) + P(0 < Z < \infty) = 0.9995$$

the power of the test for $\mu = \mu_1$ i.e. $P_w(\mu)$ is given by $1-\beta_{\mu=\mu_1}$

$$P_w(68.5) = 1-\beta_{\mu=68.5} = 1-0.1977 = 0.8023$$

$$P_w(68.0) = 1-\beta_{\mu=68.0} = 1-0.4920 = 0.5080$$

$$P_w(67.5) = 1-\beta_{\mu=67.5} = 1-0.7939 = 0.2061$$

$$P_w(66) = 1-\beta_{\mu=66} = 1-0.9995 = 0.0005$$

Q: 16.9: Given H_0: $\mu = \mu_0$ and H_1 : $\mu = \mu_1$ and α and β are probabilities of making type I and type-II errors respectively, show that for a one- sided hypothesis test, the required sample size n is given by the expression

$$n = \frac{\sigma^2 (Z_0 + Z_1)^2}{(\mu_1 - \mu_0)^2}$$

Also use this formula to find n when $\sigma = 12$, $\mu_0 = 28$, $\mu_1 = 32$, $\alpha = 0.05$ and $\beta = 0.01$.

Solution: the sample size "n" is given by $n = \frac{\sigma^2 (Z_0 + Z_1)^2}{(\mu_1 - \mu_0)^2}$

The value of Z under the μ_0-distribution at $\alpha = 0.05$ is Z_0=1.645 and The value of Z under the μ_1-distribution at $\beta = 0.01$ is Z_1=2.33

Putting the values in eq.(i)

$$n = \frac{\sigma^2 (Z_0 + Z_1)^2}{(\mu_1 - \mu_0)^2} = \frac{144(1.645+2.33)^2}{(32-28)^2} = 142$$

Hence the sample size is 142

Q: 16.10 The hypothesis H_0 : $\mu = 100$ is to be tested with $\alpha = 0.05$.The population standard deviation is known to be $\sigma = 10$.

a) Would a sample of size n = 100 result in a value of β less than 0.2 if .in fact, $\mu = 110$?

b) How large a sample would be required so that $\beta = 0.01$ if in fact , $\mu = 110$?

Solution: given that $H_0 : \mu = 100$, $\sigma = 10$, $\alpha = 0.05$ and n=100

a) Let us find the value of $\bar{x}$ (the critical point) which would lead to rejection of the hypothesis H_0 by the test statistic $Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$. The critical value of Z for two tailed test $\alpha = 0.05$ from the normal curve areas are "-1.96 and +1.96". Thus

$$\pm 1.96 = \frac{\bar{x}-100}{\frac{10}{\sqrt{100}}} \qquad After\ samplifying\ \bar{x} = 98.04\ and\ 101.96$$

That is, the hypothesis $H_0 : \mu = 100$ would be rejected if $\bar{x} < 98.04$ *or* $\bar{x} > 101.96$

To compute β, the probability of type II error, we use the distribution under the alternative hypothesis $H_1 : \mu = 110$

$$101.96 \text{ in standard units} = \frac{101.96-110}{\frac{10}{\sqrt{100}}} = -8.04$$

$$98.04 \text{ in standard units} = \frac{98.04-110}{\frac{10}{\sqrt{100}}} = -11.96$$

Thus β = Area under normal curve between Z=-8.04 and Z=-11.96 is <0.2

Hence β would result in less than 0.2 when in fact $H_1 : \mu = 110$

b) Solution: in this case, the sample size, "n" is given by the relation

$$n = \frac{\sigma^2(Z_0+Z_1)^2}{(\mu_1-\mu_0)^2} \qquad \text{(i)}$$

The value of Z under the μ_0-distribution at $\alpha = 0.05$ is Z_0=1.96 and The value of Z under the μ_1-distribution at $\alpha = 0.01$ is Z_1=2.33

Putting the values in eq.(i)

$$n = \frac{\sigma^2(Z_0+Z_1)^2}{(\mu_1-\mu_0)^2} = \frac{100(1.96+2.33)^2}{(110-100)^2} = 18.52$$

Hence the sample size must be at least 19

Q: 16.11: Suppose that H_0: μ = 200 miles and H_1 : μ > 200 miles. And α = 0.05 is required and β = 0.10 is acceptable when the true mean is 205 miles. Find the optimum sample size. It is estimated that σ = 15.What decision rule would you establish?

Solution: let μ_0 *and* μ_1 be the mean stated in the null and alternative hypothesis respectively and the values of normal vitiates of Z be Z_0 and Z_1

Then the sample size, n, is given by the relation

$$n = \frac{\sigma^2(Z_0+Z_1)^2}{(\mu_1-\mu_0)^2} \qquad \text{(i)}$$

The value of Z under the μ_0-distribution at $\alpha = 0.05$ is Z_0=1.645 and The value of Z under the μ_1-distribution at $\alpha = 0.10$ is Z_1=1.2

Putting the values in eq.(i)

$$n = \frac{\sigma^2(Z_0+Z_1)^2}{(\mu_1-\mu_0)^2} = \frac{225(1.645+1.28)^2}{(205-200)^2} = 77.04 = 77$$

Required sample size is 77

Describe the Procedure for testing of hypothesis about mean of normal population when population standard deviation is known when sample size small or large.

Procedure:

Setp-i: We set up our null and alternative hypothesis

a) $H_0 : \mu = \mu_0$ b) $H_0 : \mu \geq \mu_0$ c) $H_0 : \mu \leq \mu_0$

a) $H_1 : \mu \neq \mu_0$ b) $H_1 : \mu < \mu_0$ c) $H_1 : \mu > \mu_0$

Step-ii: Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and known σ.

Step-iii: Level of significance

$\alpha = (\textit{Comonly uesd } 5\% \textit{ or } 1\%)$

Step-iv: Test-statistic

a) $Z = \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} = \dfrac{\bar{X}-\mu_0}{S.E(\bar{X})}$ When sampling done with replacement

b) $Z = \dfrac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}\sqrt{\left(\frac{N-n}{N-1}\right)}} = \dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} = \dfrac{\bar{X}-\mu_0}{S.E(\bar{X})}$ Sampling done without replacement

If H_0 is true; it has z-distribution

Step-v: Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \mu \neq \mu_0$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : \mu < \mu_0$ We use one sided test

$Z_c < -Z_\alpha$

C) $H_1 : \mu > \mu_0$ We use one sided test

$Z_c > Z_\alpha$

Step-vi: Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

Step-vii: Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Q.16.12 (b): Past experience has shown that the scores of students who take a certain mathematics test are normally distributed with mean 75 and variance 36. The Mathematics Department members would like to know whether this year's group of 16 students is typical. They decide to test the hypothesis that this year's students are typical against the alternative that they are not typical. When the students take the test, the average score is 82. What conclusion should be drawn?

Solution:

i) We set up our null and alternative hypothesis

a) $H_0 : \mu = 75$

a) $H_1 : \mu \neq 75$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and known σ.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n = 16$ $\mu = 75$ $\overline{X} = 82$ $\sigma^2 = 36$ $\sigma = 6$

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{82-75}{\frac{6}{\sqrt{16}}} = 4.67$$

vii) Conclusion

Since our calculated value fall's in critical region then we rejects H_0.We conclude that the population mean not equal to 75 at 5% level of significance.

Q.16.13(b): A random sample of size 36 is taken from a normal population with a known variance $\sigma^2 = 25$.If mean of the sample is $\overline{x} = 42.6$, test the null hypothesis $\mu = 45$ against the alternative hypothesis $\mu < 45$ with $\alpha = 0.05$(α is the probability of committing type-I error)

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 45$

$H_1 : \mu < 45$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and known σ.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > Z_\alpha$

$Z_c > Z_{0.05}$

$Z_c > 1.645$

vi) Calculation

$n = 36 \quad \mu = 45 \qquad \overline{X} = 42.6 \qquad \sigma^2 = 25 \qquad \sigma = 5$

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{42.6 - 45}{\frac{5}{\sqrt{36}}} = -2.88$$

vii) Conclusion

Since our calculated value fall's in critical region then we rejects H_0.We conclude that the population mean is less than 45 at 5% level of significance.

Q: 16.14(a): The heights of college mail students are known to be normally distributed with a mean of 67.39 inches and σ = 1.30 inches. A random sample of 400 students showed a mean height of 67.47 inches. Using a 0.05 significance level, test the hypothesis $H_0 : \mu = 67.39$ against the alternative $H_1 : \mu > 67.39$.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 67.39$

$H_1 : \mu > 67.39$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and known σ.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > Z_\alpha$

$Z_c > Z_{0.05}$

$Z_c > 1.645$

vi) Calculation

$n = 400 \qquad \mu = 67.39 \qquad \overline{X} = 67.47 \qquad \sigma = 1.30$

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{67.47 - 67.39}{\frac{1.30}{\sqrt{400}}} = 1.23$$

vii) Conclusion

Since our calculated value fall's in critical region then we rejects H_0.We conclude that the population mean is grater than 67.39 at 5% level of significance.

Explain the general procedure for testing of hypothesis regarding the population mean when population standard deviation is unknown and the sample size is large.

Procedure:

Setp-i: We set up our null and alternative hypothesis

a) $H_0 : \mu = \mu_0$ b) $H_0 : \mu \geq \mu_0$ c) $H_0 : \mu \leq \mu_0$

a) $H_1 : \mu \neq \mu_0$ b) $H_1 : \mu < \mu_0$ c) $H_1 : \mu > \mu_0$

Step-ii: Assumption: A sample is drawn randomly and independently from a normal population with population mean μ when unknown σ .and large sample size $(n \geq 30)$

Step-iii: Level of significance

$\alpha = (Comonly \ \ uesd\ 5\% \ or \ 1\%)$

Step-iv: Test-statistic

a) $Z = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} = \dfrac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} = \dfrac{\overline{X} - \mu_0}{S.E(\overline{X})}$ When sampling done with replacement

b) $Z = \dfrac{\overline{X}-\mu_0}{\dfrac{S}{\sqrt{n}}\sqrt{\left(\dfrac{N-n}{N-1}\right)}} = \dfrac{\overline{X}-\mu_0}{\sigma_{\overline{X}}} = \dfrac{\overline{X}-\mu_0}{S.E(\overline{X})}$ Sampling done without replacement

If H_0 is true; it has z-distribution

Step-v: Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \mu \neq \mu_0$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : \mu < \mu_0$ We use one sided test

$Z_c < -Z_\alpha$

C) $H_1 : \mu > \mu_0$ We use one sided test

$Z_c > Z_\alpha$

Step-vi: Calculation

In this step we calculate the value of “Z” test statistic on the basis of sample data.

Step-vii: Conclusion

If our calculated value does not fall’s in critical region then we accept H_0 other wise we reject it.

Q.16.14 (b): The IQ’s of the college students are known to be normally distributed with a mean of 123.A random sample of 49 students showed an average IQ of $\bar{x}$ = 120.67 and S = 8.44.Test the hypothesis that μ = 123 against the alternative that it is less. Let α = 0.05.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 123$

$H_1 : \mu < 123$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and unknown σ but “n” is large

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z = \frac{\overline{X}-\mu}{S/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c < Z_\alpha$

$Z_c < Z_{0.05}$

$Z_c < -1.645$

vi) Calculation

$n = 49$ $\mu = 123$ $\overline{X} = 120.67$ $S = 8.44$

$$Z = \frac{\overline{X}-\mu}{S/\sqrt{n}} = \frac{120.67-123}{\dfrac{8.44}{\sqrt{49}}} = -1.93$$

vii) Conclusion

Since our calculated value fall’s in critical region then we rejects H_0.We conclude that the population mean is less than 123 at 5% level of significance.

Q: 16.15(a): A sample of size 40 from a non normal population yielded the sample mean $\bar{x}$ = 71 and $S^2 = 200$.Test H_0 : μ = 72 against H_1 : μ ≠ 72 using a 0.01 significance level.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 72$

$H_1 : \mu \neq 72$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and unknown σ and “n” is large.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.01}{2}}$

$|Z_c| \geq Z_{0.005}$

$|Z_c| \geq 2.57$

vi) Calculation

$n = 40$ $\mu = 72$ $\overline{X} = 71$ $S^2 = 200$ $S = 14.14$

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{71 - 72}{\frac{14.14}{\sqrt{40}}} = -0.45$$

vii) Conclusion

Since our calculated value does fall in critical region. Then we accept H_0 and we conclude that the population mean is 72 at 1% level of significance.

Q.16.15(b): Suppose that mean μ of a random variable X is unknown but the variance for X is known to be 144.Should we reject the null hypothesis H_0 : μ = 15 in favor of an alternative hypothesis H_0 : μ ≠ 15 at α = 0.05,if a random sample of 64 observations yielded a mean $\overline{x}$ = 12.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 15$

$H_1 : \mu \neq 15$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and known σ.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n = 64$ $\mu = 15$ $\overline{X} = 12$ $\sigma^2 = 144$ $\sigma = 12$

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{12 - 15}{\frac{12}{\sqrt{64}}} = -2.0$$

vii) Conclusion

Since our calculated value fall’s in critical region. Then we reject H_0 and we conclude that the population mean is not 15 at 5% level of significance.

Q: 16.16: It is claimed that an automobile is drawn on the average more than 20,000 kilometer per year. To test this claim, a random sample of 100 automobile owners is asked to keep a record of the kilometers they travel. Would you agree with the claim if the random sample showed an average of 23,500 kilometers and a standard deviation of 3900 kilometers? Use a 0.01 level of significance.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu \le 20000$

$H_1 : \mu > 20000$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and unknown σ but "n" is large.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > Z_\alpha$

$Z_c > Z_{0.01}$

$Z_c > 2.33$

vi) Calculation

$n = 100$ $\mu = 20000$ $\overline{X} = 23500$ $S = 3900$

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{23500 - 20000}{\frac{139}{\sqrt{100}}} = 8.97$$

vii) Conclusion

Since our calculated value fall's in critical region then we rejects H_0.We conclude that the population mean is grater than 20000 at 1% level of significance.

Q: 16.17(a): A sample of 900 members has a mean 2.4 inches. Could it be reasonably regarded as being a simple random sample from a large population whose mean is 2.9 inches and standard deviation 3.2 inches?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 2.9$

$H_1 : \mu \ne 2.9$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and unknown σ and "n" is large.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \ge Z_{\frac{\alpha}{2}}$

$|Z_c| \ge Z_{\frac{0.05}{2}}$

$|Z_c| \ge Z_{0.025}$

$|Z_c| \ge 1.96$

vi) Calculation

$n = 900$ $\mu = 2.9$ $\overline{X} = 2.4$ $S = 3.2$

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{2.4 - 2.9}{\frac{3.2}{\sqrt{900}}} = -4.69$$

vii) Conclusion

Since our calculated value fall's in critical region. Then we reject H_0 and we conclude that the population mean is not regarded 2.9 at 5% level of significance.

Q.16.17(b): A sample of size 400 has a mean 6.0”.Can it be regarded as a simple random sample from a large population with mean 6.2” and standard deviation 2.25”?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 6.2$

$H_1 : \mu \neq 6.2$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and unknown σ but “n” is large.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n = 400$ $\mu = 6.2$ $\bar{X} = 6.0$ $S = 2.25$

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{6.0 - 6.2}{\frac{2.25}{\sqrt{400}}} = -1.78$$

vii) Conclusion

Since our calculated value does not fall in critical region. Then we accept H_0 and we conclude that the population mean is regarded 6.2 at 5% level of significance.

Q: 16.18(a): A process is in control when the average amount of instant coffee that is picked in jar, is 6 oz. The standard deviation is 0.2 oz. A sample of 100 jars is selected at random and the sample average is found to be 6.1 oz. Is the process out of control?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu = 6$

$H_1 : \mu \neq 6$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and known σ.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq 1.96$

vi) Calculation

$n = 100$ $\mu = 6$ $\bar{X} = 6.1$ $\sigma = 0.2$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{6.1 - 6}{\frac{0.2}{\sqrt{100}}} = 5.0$$

vii) Conclusion

Since our calculated value fall in critical region then we rejects H_0.We conclude that the process is out of control at 5% level of significance.

Q.16.18 (b): Can you reject a claim that the average age of members of Parliament is at least 50, if a random sample of 36 members has a mean age of 48.7 with a standard deviation of 3.1 years. Assume all members ages are normally distributed, test at the 0.01 level.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu \geq 50$

$H_1 : \mu < 50$

ii) Assumption: A sample is drawn randomly and independently from a normal population with population mean μ and unknown σ but "n" is large.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \rightarrow N(0,1)$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c < -Z_\alpha$

$Z_c < -Z_{0.01}$

$Z_c < -2.33$

vi) Calculation

$n = 36$ $\quad$ $\mu = 50$ $\quad$ $\overline{X} = 48.7$ $\quad$ $S = 3.1$

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{48.7 - 50}{\frac{3.1}{\sqrt{36}}} = -2.52$$

vii) Conclusion

Since our calculated value fall in critical region then we rejects H_0.We conclude that we can reject the claim at 1% level of significance.

Describe the procedure for testing equality of means of two normal populations when population standard deviations are known ($\sigma_1 \neq \sigma_2$) and the sample sizes are large or small.

Procedure:

i) We set up our null and alternative hypothesis

a) $H_0 : \mu_1 - \mu_2 = \Delta$ $\quad$ b) $H_0 : \mu_1 - \mu_2 \leq \Delta$ $\quad$ c) $H_0 : \mu_1 - \mu_2 \geq \Delta$

a) $H_1 : \mu_1 - \mu_2 \neq \Delta$ $\quad$ b) $H_1 : \mu_1 - \mu_2 > \Delta$ $\quad$ c) $H_1 : \mu_1 - \mu_2 < \Delta$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2.known.

iii) Level of significance

$\alpha = (Comonly\ \ uesd\ 5\%\ or\ 1\%)$

iv) Test-statistic

a) $$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$ When sampling done with replacement

b) $$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}\left(\frac{N_1 - n_1}{N_1 - 1}\right) + \frac{\sigma_2^2}{n_2}\left(\frac{N_2 - n_2}{N_2 - 1}\right)}}$$ Sampling done without replacement

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \mu_1 - \mu_2 \neq \Delta$ $\quad$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : \mu_1 - \mu_2 > \Delta$ We use one sided test

$Z_c > Z_\alpha$

C) $H_1 : \mu_1 - \mu_2 < \Delta$ We use one sided test

$Z_c < -Z_\alpha$

vi) Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

vii) Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Q: 16.19(a): A sample of size 6 from a normal population with variance 24 gave $\bar{x}_1 = 15$.A sample of size 8 from a normal population with variance 80 gave $\bar{x}_2 = 13$.Test H_0: $\mu_1 - \mu_2 = 0$ against H_1: $\mu_1 - \mu_2 \neq 0$.Let $\alpha = 0.05$.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 known but unequal.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 6$ $\bar{X}_1 = 15$ $\sigma_1^2 = 24$ $n_2 = 8$ $\bar{X}_2 = 13$ $\sigma_2^2 = 80$

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$Z_c = \frac{(15-13)-(0)}{\sqrt{\frac{24}{6} + \frac{80}{8}}} = 0.53$$

vii) Conclusion

Since our calculated value does not fall in critical regions then we accept H_0.And we conclude that the difference of population means are insignificant at 5% level of significance.

Q: 16.19 (b): A random sample of size $n_1 = 25$, taken from a normal population with a standard deviation $\sigma_1 = 5.2$, has a mean $\bar{x}_1 = 81$.A second random sample of size $n_2 = 36$, taken from a different normal population with a standard deviation $\sigma_2 = 3.4$, has a mean $\bar{x}_2 = 76$.Test the hypothesis at the 0.06 level of significance that $\mu_1 = \mu_2$ against the alternative $\mu_1 \neq \mu_2$.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 known but unequal.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 25$ $\quad \bar{X}_1 = 81$ $\quad \sigma_1 = 5.2$ $\quad n_2 = 36$ $\quad \bar{X}_2 = 76$ $\quad \sigma_2 = 3.4$

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$Z_c = \frac{(81 - 76) - (0)}{\sqrt{\frac{(5.2)^2}{25} + \frac{(3.4)^2}{36}}} = 4.22$$

vii) Conclusion

Since our calculated value fall's in critical regions then we rejectt H_0. And we conclude that the difference of population means are significant at 5% level of significance.

Describe the procedure for testing equality of means of two normal populations when population standard deviations are known but equal ($\sigma_1 = \sigma_2$) and the sample sizes are large or large.

Procedure:

i) We set up our null and alternative hypothesis

a) $H_0 : \mu_1 - \mu_2 = \Delta$ $\quad$ b) $H_0 : \mu_1 - \mu_2 \leq \Delta$ $\quad$ c) $H_0 : \mu_1 - \mu_2 \geq \Delta$

a) $H_1 : \mu_1 - \mu_2 \neq \Delta$ $\quad$ b) $H_1 : \mu_1 - \mu_2 > \Delta$ $\quad$ c) $H_1 : \mu_1 - \mu_2 < \Delta$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2.known.

iii) Level of significance

$\alpha = (Comonly\ \ uesd\ 5\%\ or\ 1\%)$

iv) Test-statistic

a) $$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$ When sampling done with replacement

b) $$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1}\left(\frac{N_1 - n_1}{N_1 - 1}\right) + \frac{1}{n_2}\left(\frac{N_2 - n_2}{N_2 - 1}\right)}}$$ Sampling done without replacement

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \mu_1 - \mu_2 \neq \Delta$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : \mu_1 - \mu_2 > \Delta$ We use one sided test

$Z_c > Z_\alpha$

C) $H_1 : \mu_1 - \mu_2 < \Delta$ We use one sided test

$Z_c < -Z_\alpha$

vi) Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

vii) Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Q: 16.20: The two samples A and B detailed below were taken from normal populations of standard deviation 2.5.Decide whether the difference of sample means is significant at the 0.05 level of significance.

A	16	18	23	26	19	24	25	23	21	22
B	20	21	23	25	25	27	24	26	24	28

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 known but equal.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 10$ $\overline{X}_1 = \frac{\sum x}{n} = 21.7$ $\sigma_1 = \sigma_2 = \sigma = 2.5$ $n_2 = 10$ $\overline{X}_2 = 24.3$

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$Z_c = \frac{(21.7 - 24.3) - (0)}{2.5\sqrt{\frac{1}{10} + \frac{1}{10}}} = -2.33$$

vii) Conclusion

Since our calculated value fall's in critical regions then we reject H_0. And we conclude that the difference of population means are significant at 5% level of significance.

Describe the procedure for testing equality of means of two normal populations when population standard deviations are unknown and the sample sizes are large.

Procedure:

i) We set up our null and alternative hypothesis

a) $H_0 : \mu_1 - \mu_2 = \Delta$ b) $H_0 : \mu_1 - \mu_2 \le \Delta$ c) $H_0 : \mu_1 - \mu_2 \ge \Delta$

a) $H_1 : \mu_1 - \mu_2 \ne \Delta$ b) $H_1 : \mu_1 - \mu_2 > \Delta$ c) $H_1 : \mu_1 - \mu_2 < \Delta$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 unknown and large sample sizes.

iii) Level of significance

$\alpha = (Comonly \quad uesd\ 5\% \ or \ 1\%)$

iv) Test-statistic

a) $Z_c = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$ When sampling done with replacement

b) $Z_c = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1}\left(\dfrac{N_1 - n_1}{N_1 - 1}\right) + \dfrac{S_2^2}{n_2}\left(\dfrac{N_2 - n_2}{N_2 - 1}\right)}}$ Sampling done without replacement

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \mu_1 - \mu_2 \ne \Delta$ We used two sided test

$|Z_c| \ge Z_{\frac{\alpha}{2}}$

b) $H_1 : \mu_1 - \mu_2 > \Delta$ We use one sided test

$Z_c > Z_\alpha$

C) $H_1 : \mu_1 - \mu_2 < \Delta$ We use one sided test

$Z_c < -Z_\alpha$

vi) Calculation

In this step we calculate the value of “Z” test statistic on the basis of sample data.

vii) Conclusion

If our calculated value does not fall’s in critical region then we accept H_0 other wise we reject it.

Q.16.21: An examination was given to two classes of 40 and 50 students respectively.In the first class, mean grade was 74 with standard deviation of 8, while in the second class the mean grade was 78 with a standard deviation of 7.Is there a significant difference between mean grades (i) at 5% level of significance? (ii) at 1 % level of significance?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \ne \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 known but samples sizes are large.

iii) Level of significance

a) $\alpha = 5\% = 0.05$ b) $\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

a) $\alpha = 5\% = 0.05$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

b) $\alpha = 1\% = 0.01$

$|Z_c| \geq 2.58$

vi) Calculation

$n_1 = 40 \qquad \overline{X}_1 = 74 \qquad S_1 = 8 \qquad n_2 = 50 \qquad \overline{X}_2 = 78 \qquad S_2 = 7$

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$Z_c = \frac{(74-78)-(0)}{\sqrt{\frac{(8)^2}{40} + \frac{(7)^2}{50}}} = -2.49$$

vii) Conclusion

a) $\alpha = 5\% = 0.05$

Since our calculated value fall's in critical regions then we rejectt H_0. And we conclude that the difference of population means are significant at 5% level of significance.

b) $\alpha = 1\% = 0.01$

Since our calculated value does not fall in critical regions then we accept H_0. And we conclude that the difference of population means are insignificant at 1% level of significance.

Q.16.22: A manufacturer suspects a difference in the quality of the spare parts he receives from two suppliers. He obtains the following data on the service life of random samples of parts from two suppliers.

Supplier	Number in sample	Means	Standard Deviation
A	50	150	10
B	100	153	5

`Test whether the difference between the two sample means is statistically significant at the 1 % level of significance.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 unknown and large sample sizes.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.01}{2}}$

$|Z_c| \geq Z_{0.005}$

$|Z_c| \geq 2.58$

vi) Calculation

$n_1 = 50 \quad \bar{X}_1 = 150 \quad S_1 = 10 \qquad n_2 = 100 \quad \bar{X}_2 = 153 \quad S_2 = 5$

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$Z_c = \frac{(150 - 153) - (0)}{\sqrt{\frac{(10)^2}{50} + \frac{(5)^2}{100}}} = -2.0$$

vii) Conclusion

Since our calculated value does not fall in critical regions then we accept H_0. And we conclude that the difference of population means are significant at 1% level of significance.

Q.16.23 (a): A simple sample of heights of 6400 Englishmen has a mean of 67.85 inches and a standard deviation of 2.56 inches, while a simple sample of heights 1600 Australians has a mean of 68.55 inches and a standard deviation of 2.52 inches. Do the data indicate that Australians are on the average taller than Englishmen? Use level of significance 5%.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : \mu_1 \geq \mu_2$

$H_1 : \mu_1 < \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 unknown and sample sizes are large.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c < -Z_\alpha$

$Z_c < -1.645$

vi) Calculation

$n_1 = 6400 \quad \bar{X}_1 = 67.85 \quad S_1 = 2.56 \qquad n_2 = 1600 \quad \bar{X}_2 = 68.55 \quad S_2 = 2.52$

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$Z_c = \frac{(67.85 - 68.55) - (0)}{\sqrt{\frac{(2.56)^2}{6400} + \frac{(2.52)^2}{1600}}} = -9.91$$

vii) Conclusion

Since our calculated value fall's in critical regions then we rejectt H_0. And we conclude that the Australlian has the taller average than englishman significant at 5% level of significance.

Q.16.23(b): A potential buyer of lights bulbs bought 50 bulbs of each of 2 brands. Upon testing the bulbs, he found that brand A had a mean life of 1282 hours with a standard deviation of 80 hours, whereas brand B had a mean life of 1208 hours with a standard deviation of 94 hours. Can the buyer be quite certain that the two brands do differ in quality?

Solution:
i) We set up our null and alternative hypothesis

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 known but unequal.

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq Z_{\frac{0.05}{2}}$

$|Z_c| \geq Z_{0.025}$

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 50 \qquad \overline{X}_1 = 1282 \qquad S_1 = 80 \qquad\qquad n_2 = 50 \qquad \overline{X}_2 = 1208 \qquad S_2 = 94$

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$Z_c = \frac{(1282 - 1208) - (0)}{\sqrt{\frac{(80)^2}{50} + \frac{(94)^2}{50}}} = 4.24$$

vii) Conclusion

Since our calculated value fall's in critical regions then we rejectt H_0. And we conclude that the two brands differ in quality at 5% level of significance.

Q.16.24: A random sample of 80 light bulbs manufactured by company A had an average life time of 1258 hours with a standard deviation of 94 hours, while a random sample of 60 light bulbs manufactured by company B had an average lifetime of 1029 hours with a standard deviation of 68 hours. Because of the high cost of bulbs from company A, we are inclined to buy from company B unless the bulbs from company A will last over 200 hours longer on the average then those from company B. Run a test using α = 0.01 to determine from whom we should buy our bulbs.

Solution:
i) We set up our null and alternative hypothesis

$H_0: \mu_1 - \mu_2 \leq 200$

$H_1: \mu_1 - \mu_2 > 200$

ii) Assumption: The two samples of sizes n_1 *and* n_2 are randomly and independently drawn from two normal populations with population mean μ_1 *and* μ_2 when population variances are σ_1^2 *and* σ_2^2 unknown and sample sizes are large.

iii) Level of significance

$\alpha = 1\% = 0.01$

iv) Test-statistic

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > Z_\alpha$

$Z_c > 2.33$

vi) Calculation

$n_1 = 80 \quad \bar{X}_1 = 1258 \quad S_1 = 94 \qquad n_2 = 60 \quad \bar{X}_2 = 1029 \quad S_2 = 68$

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$Z_c = \frac{(1258 - 1029) - 200}{\sqrt{\frac{(94)^2}{80} + \frac{(68)^2}{60}}} = 2.12$$

vii) Conclusion

Since our calculated value does not fall in critical regions then we accept H_0. And we conclude that the average of company "A" land Company "B" are insignificant at 5% level of significance.

Q: 16.25(a): Explain how you test the hypothesis on proportions.

Explain the general procedure for testing of hypothesis regarding the population proportion "p" is unknown for a large sample.

Let we draw the samples randomly and independently from a binomial population having binomial proportion (P). Then according to central limit theorem, the sampling distribution of $\hat{P}$ is approximately normally distribution with mean (P) and standard deviation $\sqrt{\frac{pq}{n}}$ for sufficiently large sample size.

$$Z_C = \frac{\hat{P} - P_0}{\sqrt{\frac{pq}{n}}}$$

When $\hat{P} = \frac{x}{n}$ then

$$Z_C = \frac{\frac{x}{n} - P_0}{\sqrt{\frac{pq}{n}}} = \frac{x - nP_0}{\sqrt{npq}}$$

Procedure:

1) Set up our null and alternative hypothesis

a) $H_0 : P = P_0$ b) $H_0 : P \geq P_0$ c) $H_0 : P \leq P_0$

a) $H_1 : P \neq P_0$ b) $H_1 : P < P_0$ c) $H_1 : P > P_0$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with population proportion

iii) Level of significance

$\alpha = (Comonly \ \ uesd\ 5\% \ or\ 1\%)$

iv) Test-statistic

a) $Z_C = \dfrac{\hat{P} - P_0}{\sqrt{\frac{\hat{p}\hat{q}}{n}}}$ When "P" is unknown

b) $Z_C = \dfrac{\hat{P} - P_0}{\sqrt{\frac{pq}{n}}}$ When "P" is known

Or

a) $Z_C = \dfrac{X - nP_0}{\sqrt{np_0 q_0}}$ Without continuity correction

b) $Z_C = \dfrac{(X \pm \frac{1}{2}) - nP_0}{\sqrt{np_0 q_0}}$ With continuity correction

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : P \neq P_0$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : P < P_0$ We use one sided test

$Z_c < -Z_\alpha$

C) $H_1 : P > P_0$ We use one sided test

$Z_c > Z_\alpha$

vi) Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP_0}{\sqrt{np_0 q_0}}$$

$Z_C = \dfrac{(X + \frac{1}{2}) - nP_0}{\sqrt{np_0 q_0}}$ When $X < nP_o$ we add ½

$Z_C = \dfrac{(X - \frac{1}{2}) - nP_0}{\sqrt{np_0 q_0}}$ When $X > nP_o$ we subtract ½

vii) Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Q: 16.25(b): A basketball player has hit on 60% of his shots from the floor. If on the next 100 shots he makes 70 baskets, would you say that his shooting has improved? Use a 0.05 level of significance.

Solution:

i) Set up our null and alternative hypothesis

$H_0 : P \leq 0.60$

$H_1 : P > 0.60$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$Z_C = \dfrac{\hat{P} - P}{\sqrt{\frac{pq}{n}}}$ Without continuity correction

$Z_C = \dfrac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}}$ With continuity correction

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > Z_\alpha$

$Z_c > Z_{0.05}$

$Z_c > 1.645$

vi) Calculation

$n=100 \qquad X=70 \qquad P=60\%=0.60 \qquad P=1-q=40\% \qquad \hat{p}=\frac{X}{n}=\frac{5}{20}=0.25$

$$Z_C=\frac{\hat{P}-P}{\sqrt{\frac{pq}{n}}}=\frac{0.7-0.60}{\sqrt{\frac{(0.60)(0.40)}{100}}}=2.04$$

$$Z_C=\frac{(X\pm\frac{1}{2})-nP_0}{\sqrt{npq}}=\frac{(70-\frac{1}{2})-100(0.60)}{\sqrt{100(0.60\times 0.40)}}=1.939 \qquad \text{Where } X>nP$$

vii) Conclusion

Since our calculated value fall's in critical region with and without continuity condition then we reject H_0 and we conclude that the shooting has improved at 5% level of significance.

Q: 16.26(a): A coin is tossed 900 times and heads appear 490 times. Does this result support the hypothesis that the coin is unbiased?

Solution:

i) Set up our null and alternative hypothesis

$$H_0: P=\frac{1}{2}=0.5$$

$$H_1: P\neq\frac{1}{2}\neq 0.5$$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha=0.05$

iv) Test-statistic

$$Z_C=\frac{\hat{P}-P}{\sqrt{\frac{pq}{n}}} \qquad \text{Without continuity correction}$$

$$Z_C=\frac{(X\pm\frac{1}{2})-nP}{\sqrt{npq}} \qquad \text{With continuity correction}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c|\geq 1.96$

vi) Calculation

$n=900 \qquad X=490 \qquad P=0.5 \qquad q=1-P=0.5 \qquad \hat{p}=\frac{X}{n}=\frac{490}{900}=0.544$

$$Z_C=\frac{\hat{P}-P}{\sqrt{\frac{pq}{n}}}=\frac{0.544-0.50}{\sqrt{\frac{(0.50)(0.50)}{900}}}=2.64$$

$$Z_C=\frac{(X\pm\frac{1}{2})-nP}{\sqrt{npq}}=\frac{(490-\frac{1}{2})-900(0.5)}{\sqrt{900(0.5*0.5)}}=2.63 \qquad \text{Where } X>nP$$

vii) Conclusion

Since our calculated value fall's in critical region with and without continuity condition then we reject H_0 and we conclude that the coin is biased at 5% level of significance.

Q.16.26 (b): The sex distribution of 98 births reported in a newspaper was 52 boys and 46 girls.Is this consistent with an equal sex division in the population?

Solution:

i) Set up our null and alternative hypothesis

$$H_0: P=\frac{1}{2}=0.5$$

$$H_1: P\neq\frac{1}{2}\neq 0.5$$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$Z_C = \dfrac{\hat{P} - P}{\sqrt{\dfrac{pq}{n}}}$ Without continuity correction

$Z_C = \dfrac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}}$ With continuity correction

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq 1.96$

vi) Calculation

$n = 98$ $X(boys) = 52$ $P = 0.5$ $q = 1 - P = 0.5$ $\hat{p} = \dfrac{X}{n} = \dfrac{52}{98} = 0.531$

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.531 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{98}}} = 0.614$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(52 - 0.5) - 98(0.5)}{\sqrt{98(0.5 * 0.5)}} = 0.505$$ Where $X > nP$

vii) Conclusion

Since our calculated value does not fall in critical region in case of with and without continuity condition then we accept H_0 and we conclude that the equal sex division at 5% level of significance.

Q: 16.27: In a poll of 10,000 voters selected at random from all the voters in a certain district, it is found that 5,180 voters in favor of particular candidate. Test the null hypothesis that the proportion of all the voters in the district, who favor the candidate is equal to or less than 50% Against the alternative that it is greater than 50%.Use a 0.05 level of significance.

Solution:

i) Set up our null and alternative hypothesis

$H_0 : P \leq 0.5$

$H_1 : P > 0.5$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$Z_C = \dfrac{\hat{P} - P}{\sqrt{\dfrac{pq}{n}}}$ Without continuity correction

$Z_C = \dfrac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}}$ With continuity correction

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > 1.645$

vi) Calculation

$n = 10000 \qquad X = 5180 \qquad P = 0.5 \qquad q = 1 - P = 0.5 \qquad \hat{p} = \frac{X}{n} = \frac{5180}{10000} = 0.518$

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.518 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{10000}}} = 3.6$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(5180 - \frac{1}{2}) - 10000(0.5)}{\sqrt{10000(0.5 * 0.5)}} = 3.59$$

vii) Conclusion

Since our calculated value fall's in critical region in both cases then we reject H_0 and we conclude that the proportion is greater than 50% at 5% level of significance.

Q: 16.28(a): In the inspection of a product, it is found that in a random sample of 200 units, 12 are defective. Is this consistent with an average of 5 percent set as a standard?

Solution:

i) Set up our null and alternative hypothesis

$H_0 : P = 5\% = 0.05$

$H_1 : P \neq 0.05$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} \qquad \text{Without continuity correction}$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} \qquad \text{With continuity correction}$$

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq 1.96$

vi) Calculation

$n = 200 \qquad X = 12 \qquad P = 0.05 \qquad q = 1 - P = 0.95 \qquad \hat{p} = \frac{X}{n} = \frac{12}{200} = 0.06$

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.06 - 0.050}{\sqrt{\frac{(0.05)(0.95)}{200}}} = 0.65$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(12 - 0.5) - 200(0.05)}{\sqrt{200(0.05 * 0.95)}} = 0.487 \qquad \text{Where } X > nP$$

vii) Conclusion

Since our calculated value does not fall in critical region in both cases then we accept H_0 and we conclude that the proportion of 5% is standard at 5% level of significance.

Q.16.28 (b): A sample of size 78 from a binomial population gave 35 successes. Test the null hypothesis that the true proportion of successes is 0.55 against the alternative that it is less. Let α = 0.05.

Solution:

i) Set up our null and alternative hypothesis

$H_0 : P = 0.55$

$H_1 : P < 0.55$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}}$$ Without continuity correction

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}}$$ With continuity correction

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > Z_\alpha$

$Z_c > Z_{0.05}$

$Z_c > 1.645$

vi) Calculation

$n = 78$ $\quad X = 35$ $\quad P = 0.55$ $\quad q = 1 - P = 0.45$ $\quad \hat{p} = \frac{X}{n} = \frac{35}{78} = 0.45$

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.45 - 0.55}{\sqrt{\frac{(0.55)(0.45)}{78}}} = -1.78 = -1.80$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(35 + 0.5) - 78(0.55)}{\sqrt{78(0.55 * 0.45)}} = -1.68$$ Where $X < nP$

vii) Conclusion

Since our calculated value fall's in critical region in both cases then we reject H_0 and we conclude that the proportion is less than 0.55 at 5% level of significance.

Q: 16.29(a): The manufacturer of a patent medicine claimed that it was 90% effective in relieving an allergy for a period of 8 hours. In a sample of 200 people who had the allergy, the medicine provided relief for 160 people. Determine whether the manufacturer's claim is legitimate at the α = 0.01.

Solution:

i) Set up our null and alternative hypothesis

$H_0 : P \geq 90\%$

$H_1 : P < 90\%$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.01$

iv) Test-statistic

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}}$$ Without continuity correction

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}}$$ With continuity correction

If H_0 is true; it has z-distribution

v) Critical region

$Z_c < -2.33$

vi) Calculation

$n = 200$ $\quad X = 160$ $\quad P = 90\% = 0.90$ $\quad q = 1 - P = 0.10$ $\quad \hat{p} = \frac{X}{n} = \frac{160}{200} = 0.80$

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.8 - 0.90}{\sqrt{\frac{(0.90)(0.10)}{200}}} = -4.71$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(160 + 0.5) - 200(0.9)}{\sqrt{200(0.9 * 0.1)}} = -4.60 \qquad \text{Where } X < nP$$

vii) Conclusion

Since our calculated value fall's in critical region in both cases then we reject H_0 and we conclude that the claim is not legitimate at 5% level of significance.

Q.16.29 (b): It is claimed that 90% of men cannot tell the difference between two different brands of cheese, but of the members of a random sample of 500 men, 72 could distinguish between them. Is the claim justified?

Solution:

i) We state our null alternative hypothesis

$H_0 : P = 0.90$

$H_1 : P < 0.90$

ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} \qquad \text{Without continuity correction}$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} \qquad \text{With continuity correction}$$

If H_0 is true; it has z-distribution

v) Critical region

$Z_c < -Z_\alpha$

$Z_c < -Z_{0.05}$

$Z_c < -1.645$

vi) Calculation

$n = 500$ $\qquad X'(could\ distiguish\ between them) = 72$

$X'(could\ not\ distiguish\ between them) = 500 - 72 = 428$ $\qquad p = 90\% = 0.90$

$\hat{p} = \frac{X}{n} = \frac{428}{500} = 0.856$ $\qquad q = 1 - p = 1 - 0.90 = 0.10$

$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.856 - 0.90}{\sqrt{\frac{(0.90)(0.10)}{500}}} = -3.28$$

$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(428 + 0.5) - 500(0.90)}{\sqrt{500(0.90 * 0.10)}} = -3.21 \qquad \text{Where } X < nP$$

vii) Conclusion

Since our calculated value fall's in critical region in case of with and without continuity correction then we reject H_0 and we conclude that the head occur less than 90% at 5% level of significance.

Q: 16.30: An electrical company claimed that at least 95% of the parts which they supplied on a government contract conformed to specifications. A sample of 400 parts was tested, and 45 did not meet specifications. Can we accept the company's claim at a 0.05 level of significance?

Solution:
i) We state our null and alternative hypothesis
$H_0 : P \geq 0.95$
$H_1 : P < 0.95$
ii) Assumption: A sample is drawn randomly and independently from a binomial normal population sufficiently large sample size with known p.
iii) Level of significance
$\alpha = 0.05$
iv) Test-statistic
$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} \qquad \text{Without continuity correction}$$
$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} \qquad \text{With continuity correction}$$
If H_0 is true; it has z-distribution
v) Critical region
$Z_c < -Z_\alpha$
$Z_c < -Z_{0.05}$
$Z_c < -1.645$
vi) Calculation
$n = 400$ $\quad X'(did\ not\ specification) = 45$ $\quad X(specification) = 400 - 45 = 355$
$p = 50\% = 0.50$ $q = 1 - p = 1 - 0.50 = 0.50$ $\quad \hat{p} = \frac{X}{n} = \frac{355}{400} = 0.8875$
$$Z_C = \frac{\hat{P} - P}{\sqrt{\frac{pq}{n}}} = \frac{0.8875 - 0.95}{\sqrt{\frac{(0.95)(0.05)}{400}}} = -5.73$$
$$Z_C = \frac{(X \pm \frac{1}{2}) - nP}{\sqrt{npq}} = \frac{(355 + 0.5) - 400(0.95)}{\sqrt{400(0.95 * 0.05)}} = -5.62 \qquad \text{Where } X < nP$$

vii) Conclusion
Since our calculated value fall's in critical region in case of with and without continuity correction then we reject H_0 and we conclude that the head occur less than 90% at 5% level of significance.

Difference between two population proportion

Suppose we draw two random samples of size n_1 and n_2 from two binomial population with unknown proportion of success P_1 and P_2 respectively. Sample proportion of two samples are $\hat{P}_1 = \frac{X_1}{n_1}$ and $\hat{P}_2 = \frac{X_2}{n_2}$
Then sampling distribution of the differences between two sample proportions $\hat{P}_1 - \hat{P}_2$ is approximately normal with mean $P_1 - P_2$ and standard deviation $\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$ for approximately large sample sizes
$Z_c = \frac{\hat{P}_1 - \hat{P}_2 - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$ Approximately standard normal variate. When P_1 *and* P_2 are not known then for large sample sizes then replaced with sample proportion $Z_c = \frac{\hat{P}_1 - \hat{P}_2 - (P_1 - P_2)}{\sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}}$

Testing of hypothesis between two population proportion

Procedure

i) We state our null and alternative hypothesis

a) $H_0 : P_1 - P_2 = \Delta$ b) $H_0 : P_1 - P_2 \geq \Delta$ c) $H_0 : P_1 - P_2 \leq \Delta$

a) $H_1 : P_1 - P_2 \neq \Delta$ b) $H_1 : P_1 - P_2 < \Delta$ c) $H_1 : P_1 - P_2 > \Delta$

ii) Assumption: Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be unequal. Also $n_1, n_2 > 30$

iii) Level of significance

$\alpha = (Comonly \quad uesd\ 5\%\ or\ 1\%)$

iv) Test-statistic

a) $Z_C = \dfrac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}}$ When "$P_1 - P_2 \neq 0$ but known"

c) $Z_C = \dfrac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\dfrac{\hat{p}_1 \hat{q}_1}{n_1} + \dfrac{\hat{p}_2 \hat{q}_2}{n_2}}}$ When "$P_1 - P_2 \neq 0$" but unknown

b) $Z_C = \dfrac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\dfrac{1}{n_1} + \dfrac{1}{n_2})}}$ When "$P_1 - P_2 = 0$" but unknown

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : P_1 - P_2 \neq \Delta$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : P_1 - P_2 < \Delta$ We use one sided test

$Z_c < -Z_\alpha$

C) $H_1 : P_1 - P_2 > \Delta$ We use one sided test

$Z_c > Z_\alpha$

vi) Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

$\hat{P}_c = \dfrac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2} = \dfrac{X_1 + X_2}{n_1 + n_2}$ $\hat{q}_c = 1 - \hat{P}_c$

$\hat{P}_1 = \dfrac{X_1}{n_1}$ And $\hat{P}_2 = \dfrac{X_2}{n_2}$

$\hat{q}_1 = 1 - \hat{P}_1$ And $\hat{q}_2 = 1 - \hat{P}_2$

vii) Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Q: 16.31(a): A random sample of 150 light bulbs manufactured by a firm X showed 12 defective bulbs while a random sample of 100 light bulbs manufactured by another firm Y showed 4 defective bulbs. Is there a significant difference between the proportions of two firms?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : P_1 = P_2$

$H_1 : P_1 \neq P_2$

ii) Assumption: Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be unequal. Also $n_1, n_2 > 30$

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}}$$ When “$P_1 - P_2 = 0$”

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 150$ $X_1 = 12$ $n_2 = 100$ $X_2 = 4$

$\hat{P}_1 = \frac{12}{150} = 0.08$ $\hat{P}_2 = \frac{4}{100} = 0.04$

$\hat{q}_1 = 1 - \hat{P}_1 = 1 - 0.08 = 0.92$ $\hat{q}_2 = 1 - \hat{P}_2 = 1 - 0.04 = 0.96$

$\hat{P}_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{12 + 4}{150 + 100} = 0.064$ $\hat{q}_c = 1 - \hat{P}_c = 1 - 0.064 = 0.936$

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.08 - 0.04) - 0}{\sqrt{(0.064)(0.936)\left(\frac{1}{150} + \frac{1}{100}\right)}} = \frac{0.04}{0.0316} = 1.27$$

vii) Conclusion

Since our calculated value does not fall’s in critical region then we accept H_0 and we conclude that there is no difference between the proportions of two firms at 5% level of significance.

Q.16.31(b): Random samples of 500 men and 500 women are selected to determine whether the proportions of men and women favouring a political candidate are different. Perform a hypothesis test at 5 % level if, in the sample 225 men and 275 women favor the candidate. What is implied by the test result?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : P_1 = P_2$

$H_1 : P_1 \neq P_2$

ii) Assumption: Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be equal. Also $n_1, n_2 > 30$

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}}$$ When “$P_1 - P_2 = 0$”

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > 1.645$

vi) Calculation

$n_1 = 500$ $X_1 = 225$ $n_2 = 500$ $X_2 = 275$

$\hat{P}_1 = \frac{225}{500} = 0.45$ $\hat{P}_2 = \frac{275}{500} = 0.55$

$\hat{q}_1 = 1 - \hat{P}_1 = 1 - 0.45 = 0.55$ $\hat{q}_2 = 1 - \hat{P}_2 = 1 - 0.55 = 0.45$

$\hat{P}_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{225 + 275}{500 + 500} = 0.50$ $\hat{q}_c = 1 - \hat{P}_c = 1 - 0.50 = 0.50$

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.45 - 0.55) - 0}{\sqrt{(0.5)(0.5)\left(\frac{1}{500} + \frac{1}{500}\right)}} = \frac{-0.10}{0.0316} = -3.16$$

vii) Conclusion

Since our calculated value does not fall's in critical region then we accept H_0 and we conclude that the machine has not improved at 5% level of significance.

Q: 16.32(a): A machine puts out 16 imperfect articles in a sample of 500. After machine is overhauled, it puts 3 imperfect articles in a batch of 100. Has the machine been improved?

Solution:

i) We set up our null and alternative hypothesis

$H_0 : P_1 = P_2$

$H_1 : P_1 > P_2$

ii) Assumption: Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be unequal. Also $n_1, n_2 > 30$

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}}$$ When "$P_1 - P_2 = 0$"

If H_0 is true; it has z-distribution

v) Critical region

$Z_c > 1.645$

vi) Calculation

$n_1 = 500$ $X_1 = 16$ $n_2 = 100$ $X_2 = 3$

$\hat{P}_1 = \frac{16}{500} = 0.032$ $\hat{P}_2 = \frac{3}{100} = 0.03$

$\hat{q}_1 = 1 - \hat{P}_1 = 1 - 0.032 = 0.968$ $\hat{q}_2 = 1 - \hat{P}_2 = 1 - 0.03 = 0.97$

$\hat{P}_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{16+3}{500+100} = 0.032$ $\hat{q}_c = 1 - \hat{P}_c = 1 - 0.032 = 0.968$

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.032 - 0.03) - 0}{\sqrt{(0.032)(0.968)\left(\frac{1}{500} + \frac{1}{100}\right)}} = \frac{0.002}{0.01928} = 0.104$$

vii) Conclusion

Since our calculated value does not fall's in critical region then we accept H_0 and we conclude that the machine has not improved at 5% level of significance.

Q.16.32 (b): A manufacturer of house dresses sent out advertising by mail. He sent samples of material to each of 2 groups of 1,000 women, for one group, he enclosed a white return envelope and for the other group, a blue envelope. His received orders from 10% and 13% respectively. Do the data indicate that the Colour of the envelope has an effect on the sales? Use 5% level of significance.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : P_1 = P_2$

$H_1 : P_1 < P_2$

ii) Assumption: Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be equal. Also $n_1, n_2 > 30$

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c \hat{q}_c (\frac{1}{n_1} + \frac{1}{n_2})}}$$ When "$P_1 - P_2 = 0$"

If H_0 is true; it has z-distribution

v) Critical region

$Z_c < -1.645$

vi) Calculation

$n_1 = 1000$ $\quad$ $n_2 = 1000$

$\hat{P}_1 = 10\% = 0.10$ $\quad$ $\hat{P}_2 = 13\% = 0.13$

$\hat{q}_1 = 1 - \hat{P}_1 = 1 - 0.10 = 0.90$ $\quad$ $\hat{q}_2 = 1 - \hat{P}_2 = 1 - 0.13 = 0.87$

$\hat{P}_c = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1 + n_2} = \frac{1000(0.10) + 1000(0.13)}{1000 + 1000} = 0.115$ $\quad$ $\hat{q}_c = 1 - \hat{P}_c = 1 - 0.115 = 0.885$

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c\hat{q}_c(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.10 - 0.13) - 0}{\sqrt{(0.155)(0.885)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} = \frac{0.030}{0.01426} = -2.10$$

vii) Conclusion

Since our calculated value fall in critical region then we reject H_0 and we conclude that the blue envelop help the sale at 5% level of significance.

Q: 16.33: A civil service examination is given to a group of a 200 candidates. On the basis of their total scores, the 200 candidates are divided into two groups, the upper 30% and the remaining 70%. Consider the first question on the examination. Among the first group, 40 had the correct answer, whereas among the second group, 80 had the correct answer. On the basis of these results, can one candidate that the first question is no good at discriminating ability of the type being examined here.

Solution:

i) We set up our null and alternative hypothesis

$H_0 : P_1 = P_2$

$H_1 : P_1 \neq P_2$

ii) Assumption Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be equal. Also $n_1, n_2 > 30$

iii) Level of significance

$\alpha = 0.05$

iv) Test-statistic

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c\hat{q}_c(\frac{1}{n_1} + \frac{1}{n_2})}}$$ When "$P_1 - P_2 = 0$"

If H_0 is true; it has z-distribution

v) Critical region

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 30\%\ 0f\ \ 200 = 60$ $\quad$ $X_1 = 40$ $\quad$ $n_2 = 70\%\ of\ 200 = 140$ $\quad$ $X_2 = 80$

$\hat{P}_1 = \frac{40}{60} = 0.67$ $\quad$ $\hat{P}_2 = \frac{80}{140} = 0.57$

$\hat{q}_1 = 1 - \hat{P}_1 = 1 - 0.67 = 0.33$ $\quad$ $\hat{q}_2 = 1 - \hat{P}_2 = 1 - 0.57 = 0.43$

$\hat{P}_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{40 + 80}{60 + 140} = 0.6$ $\quad$ $\hat{q}_c = 1 - \hat{P}_c = 1 - 0.60 = 0.40$

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{p}_c\hat{q}_c(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.67 - 0.57) - 0}{\sqrt{(0.6)(0.4)\left(\frac{1}{60} + \frac{1}{140}\right)}} = \frac{0.10}{0.0756} = 1.322$$

vii) Conclusion

Since our calculated value does not fall's in critical region then we accept H_0 and we conclude that there is no difference between the proportions of two groups at 5% level of significance.

Example 16.17: A candidate for mayor in a large city believes that he appeals to at least 10 per cent more of the women voters than the men voters. He hires the services of a poll taking

organization, and they find that 62 of 100 women interviewed support the candidate, and 69 0f 150 men support him. At the 0.05 significance level, is the hypothesis accepted or rejected?
Solution:
i) We state our null and alternative hypothesis
$H_0 : P_1(women) - P_2(men) \leq 0.10$
$H_1 : P_1(women) - P_2(men) > 0.10$
ii) Assumption: Samples are random, independent and drawn from binomial distribution. Since the sample sizes is sufficiently large, so it follows normal distribution population proportions P_1, P_2 are unknown and assumed to be unequal. Also $n_1, n_2 > 30$
iii) Level of significance
$\alpha = 5\% = 0.05$
iv) Test-statistic

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

When " $P_1 - P_2 \neq 0$ " but unknown

If H_0 is true; it has z-distribution
v) Critical region
It is naturally depend on alternative hypothesis
$H_1 : P_1 - P_2 > \Delta$ We use one sided test
$Z_c > Z_\alpha$
$Z_c > 1.645$
vi) Calculation
$n_1 = 100$ $X_1 = 62$ $n_2 = 150$ $X_2 = 69$

$\hat{P}_1 = \frac{X_1}{n_1} = \frac{62}{100} = 0.62$ And $\hat{P}_2 = \frac{X_2}{n_2} = \frac{69}{150} = 0.46$

$\hat{q}_1 = 1 - \hat{P}_1 = 0.38$ And $\hat{q}_2 = 1 - \hat{P}_2 = 0.54$

$$Z_C = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}} = \frac{0.62 - 0.46 - 0.10}{\sqrt{\frac{(0.62 * 0.38)}{100} + \frac{(0.46 * 0.54)}{150}}} = 0.95$$

vii) Conclusion
since our calculated value does not fall's in critical region then we accept H_0 at 5% level of significance.

Testing of single population standard deviation or variance

Procedure
i) We state our null and alternative hypothesis
a) $H_0 : \sigma = \sigma_0$ b) $H_0 : \sigma \geq \sigma_0$ c) $H_0 : \sigma \leq \sigma_0$
$H_1 : \sigma \neq \sigma_0$ $H_1 : \sigma < \sigma_0$ $H_1 : \sigma > \sigma_0$
ii) Assumption: Sample is random and independent drawn from normal population large sample size
iii) Level of significance
$\alpha = (Comonly \quad uesd \; 5\% \; or \; 1\%)$
iv) Test-statistic

$$Z_C = \frac{S - \sigma_0}{\sigma_0 / \sqrt{2n}}$$

If H_0 is true; it has z-distribution
v) Critical region
It is naturally depend on alternative hypothesis
a) $H_1 : \sigma \neq \sigma_0$ We used two sided test
$|Z_c| \geq Z_{\frac{\alpha}{2}}$
b) $H_1 : \sigma < \sigma_0$ We use one sided test

$Z_c < -Z_\alpha$

C) $H_1 : \sigma > \sigma_0$ We use one sided test

$Z_c > Z_\alpha$

vi) Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

$$Z_C = \frac{S - \sigma_0}{\sigma_0 / \sqrt{2n}}$$

vii) Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Example!6.18: A random sample of size 100 from a normal population showed a standard deviation 8.9. Test at the 0.05 level of significance the hypothesis that $\sigma = 7.5$ against the alternative hypothesis $\sigma \neq 7.5$.

Solution:

i) We state our null and alternative hypothesis

a) $H_0 : \sigma = 7.5$

$H_1 : \sigma \neq 7.5$

ii) Assumption: Sample is random and independent drawn from normal population large sample size

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_C = \frac{S - \sigma}{\sigma / \sqrt{2n}}$$

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \sigma \neq 7.5$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq 1.96$

vi) Calculation

$n = 100$ $S = 8.9$ $\sigma = 7.5$

$$Z_C = \frac{S - \sigma}{\sigma / \sqrt{2n}} = \frac{8.9 - 7.5}{7.5 / \sqrt{2(100)}} = 2.64$$

vii) Conclusion

Since our calculated value fall in critical region then we reject H_0 and we conclude that $H_1 : \sigma \neq 7.5$ at 5% level of significance.

The difference between two population standard deviation for large sample sizes (n_1,n_2>120) or variances

Procedure

i) We state our null and alternative hypothesis

a) $H_0 : \sigma_1 = \sigma_2$ b) $H_0 : \sigma_1 \geq \sigma_2$ c) $H_0 : \sigma_1 \leq \sigma_2$

$H_1 : \sigma_1 \neq \sigma_2$ $H_1 : \sigma_1 < \sigma_2$ $H_1 : \sigma_1 > \sigma_2$

ii) Assumption: Sample is random and independent drawn from normal population large sample sizes $n_1, n_2 \geq 120$

iii) Level of significance

$\alpha = (Comonly\ \ uesd\ 5\%\ or\ 1\%)$

iv) Test-statistic

$$Z_C = \frac{S_1 - S_2 - (\sigma_1 - \sigma_2)}{\sqrt{\frac{\sigma_1^2}{2n_1} + \frac{\sigma_2^2}{2n_2}}}$$ When (σ_1, σ_2) are known

$$Z_C = \frac{S_1 - S_2}{\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}}}$$ When (σ_1, σ_2) are unknown

If H_0 is true; it has z-distribution

v) Critical region

It is naturally depend on alternative hypothesis

a) $H_1 : \sigma_1 \neq \sigma_2$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

b) $H_1 : \sigma_1 < \sigma_2$ We use one sided test

$Z_c < -Z_\alpha$

C) $H_1 : \sigma_1 > \sigma_2$ We use one sided test

$Z_c > Z_\alpha$

vi) Calculation

In this step we calculate the value of "Z" test statistic on the basis of sample data.

$$Z_C = \frac{S_1 - S_2 - (\sigma_1 - \sigma_2)}{\sqrt{\frac{\sigma_1^2}{2n_1} + \frac{\sigma_2^2}{2n_2}}}$$

vii) Conclusion

If our calculated value does not fall's in critical region then we accept H_0 other wise we reject it.

Q: 16.34: The standard deviation of a simple sample of 1,000 members is 5.9 years and that of an independent sample of 900 members is 6.1 years. Show that the samples can be reasonably regarded as drawn from equally variable normal populations.

Solution:

i) We state our null and alternative hypothesis

a) $H_0 : \sigma_1 = \sigma_2$

$H_1 : \sigma_1 \neq \sigma_2$

ii) Assumption: Sample is random and independent drawn from normal population large sample sizes $n_1, n_2 \geq 120$

iii) Level of significance

$\alpha = 5\% = 0.05$

iv) Test-statistic

$$Z_C = \frac{S_1 - S_2}{\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}}}$$ When (σ_1, σ_2) are unknown

If H_0 is true; it has z-distribution

v) Critical region

a) $H_1 : \sigma_1 \neq \sigma_2$ We used two sided test

$|Z_c| \geq Z_{\frac{\alpha}{2}}$

$|Z_c| \geq 1.96$

vi) Calculation

$n_1 = 1000$ $n_2 = 900$ $S_1 = 5.9$ $S_2 = 6.1$

$$Z_C = \frac{S_1 - S_2}{\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}}} = \frac{5.9 - 6.1}{\sqrt{\frac{(5.9)^2}{2(1000)} + \frac{(6.1)^2}{2(900)}}} = -1.02$$

vii) Conclusion

Since our calculated value does not fall's in critical region then we accept H_0 at 5% level of significance.