Chapter - 9

HYPOTHESES TEST CONCERNING PROPORTION (ATTRIBUTE)

Study Strategy and Learning Objectives

Please Remember The Following Reading Strategy and Learning Objectives:

Study Strategy:

1. First, read this section with the limited objective of simply trying to underst the following important key terms and concepts:
   population proportion p, sample proportion $\hat{p}$, sampling distribution of proportion, contingency table, independence of attributes, goodness-o fit-test, chi-square distribution, parametric tests, non parametric test.
2. Second, try to understand what they accomplish, why they are needed; develop the ability to calculate or select them.
3. Third, learn how to interpret them.
4. Fourth, read the section once again and try to understand the underlying theo

You will always enjoy much greater success if you understand what are doing, instead of blindly applying mechanical steps in order to obtain an ans that may or may not make any sense.

Learning Objectives

After careful study of this chapter you should be able to do the following:

1. Structure engineering decision-making problems as hypothesis tests proportion.
2. Structure comparative experiments involving two samples as hypothesis te
3. Test hypotheses and construct confidence intervals on the ratio of variances or standard deviations of two normal distributions.
4. Test hypotheses and construct confidence intervals on single proportion the difference in two population proportions.
5. Use the *P*-value approach for making decisions in hypotheses tests.
6. Compute power, type II error probability, and make sample size decisions one sample test and two-sample tests on proportions.
7. Explain and use the relationship between confidence intervals and hypothesis
8. Use the chi-square goodness-of-fit test, chi-square test for independence attributes to check distributional assumptions.
9. Use contingency table tests.

9.1 Testing a Hypothesis (or Claim) about a Proportion

In many engineering and management problems, we are concerned wit random variable that follows the binomial distribution. For example, conside production process that manufactures items that are classified as either acceptable defective. It is usually reasonable to model the occurrence of defectives with binomial distribution, where the binomial parameter p represents the proportion defective items produced.

Many of the methods used in sampling inspection, quality control reliability verification are based on tests of the null hypothesis that a proporti (*percentage* or *probability*) equals some specified constant.

The proportions can also represent probabilities or the decimal equivalents percents. The following are examples of the types of claims we will be able to test.

1. Fewer than 1/4 of all college graduates smoke.
2. Subjects taking the cholesterol-reducing drug Lipitor experience headaches at a rate that is greater than the 7% rate for people who do not take Lipitor.
3. The percentage of late-night television viewers who watch The Late Show with David Letterman is equal to 18%.
4. Based on early exit polls, the Maiost candidate for the presidency will win a majority (more than 50%) of the votes.

The required assumptions, notation, and test statistic are all given below. Basically, claims about a population proportion are usually tested by using a normal distribution as an approximation to the binomial distribution, as we did in previous sections. Instead of using the same exact methods we use a different but equivalent form of the test statistic shown below, and we don't include the correction for continuity (*because its effect tends to be very small with large samples*). If the given assumptions are not all satisfied, we may be able to use other methods not described in this section. In this section, all examples and exercises involve cases in which the assumptions are satisfied, so the sampling distribution of sample proportions can be approximated by the normal distribution.

Testing Hypotheses (Claims) About a Population Proportion p

Assumptions

1. The sample observations are a simple random sample. (Never forget the critical importance of sound sampling methods.)
2. The conditions for a binomial distribution are satisfied. (There is a fixed number of independent trials having constant probabilities, and each trial has two outcome categories of "success" and "failure.")
3. The conditions $np \geq 5$ and $nq \geq 5$ are both satisfied, so the binomial distribution of sample proportions can be approximated by a normal distribution with

$$\mu = np \text{ and } \sigma = \sqrt{npq}$$

Notation

n = sample size or number of trials, $\hat{p} = \frac{X}{n}$ (sample proportion)

p = population proportion (used in the null hypothesis).

$q = 1 - p$

Test Statistic for Testing a Claim About a Proportion

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Example 1: (*Finding the Test Statistic Z*): A survey of $n = 880$ randomly selected adult drivers showed that 56% (or $\hat{p} = 0.56$) of those respondents admitted to running red lights. Find the value of the test statistic for the claim that the majority of all adult drivers admit to running red lights.

Solution: We will see that there are assumptions that must be verified. For this example, assume that the required assumptions are satisfied and focus on finding the indicated test statistic. The null and alternative hypotheses are:

$$H_0: p = 0.5 \text{ and } H_1: p > 0.5.$$

Because we work under the assumption that the null hypothesis is true with

$p = 0.5$, we get the following test statistic:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

Interpretation: We know from previous chapters that a z score of 3.56 is exceptionally large. It appears that in addition to being "more than half," the sample result of 56% is *significantly* more than 50%. Here we show that the sample proportion of 0.56 (from 56%) does fall within the range of values considered to be significant because they are so far above 0.5 that they are not likely to occur by chance (assuming that the population proportion is $p = 0.5$)

9.2 Test of significance of a single proportion

If there is a single proportion p, we follow the following procedure for testing the significance of population proportions

Step 1. ***Set up hypotheses***

Null hypothesis: $H_0: p = p_0$ i.e., the population proportion has a specified value p_0.

Alternative hypothesis	$H_1: p \neq p_0$	(for *two-tailed test*)
or,	$H_1: p > p_0$	(for *right-tailed test*)
or,	$H_1: p < p_0$	(for *left-tailed test*)

Step 2. ***Level of significance*** **(α):**

Choose most commonly used α = 5% unless otherwise stated.

Step 3. ***Test statistic***: Under $H_0: p = p_0$, the test statistic

$$Z = \frac{\hat{p} - p}{S.E.(\hat{p})} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}} \sim N(0, 1) \quad [\because p = p_0, \hat{p} = \frac{X}{n}, q = 1 - p_0]$$

where, $\hat{p}$ = sample proportion $= \frac{X}{n}$; $S.E.(\hat{p}) = \sqrt{\frac{pq}{n}}$

X = number of successes; n = sample size

p = population proportion of success; $q = 1 - p$.

Step 4. ***Critical value***: Obtain the critical or tabulated value of test statistic Z at level of significance α.

Step 5. ***Decision***: If $|Z| > Z_\alpha$ for one tailed test or, $|Z| > Z_{\alpha/2}$ for two tailed test, reject H_0 and accept H_1. If $|Z| \leq Z_\alpha$ for one tailed test or $|Z| \leq Z_{\alpha/2}$ for two tailed test, accept H_0 and reject H_1.

Summary of Decision Rule for Z-test

Critical Regions for Testing $p = p_0$ (Large sample)	
Alternative hypothesis	*Reject H_0 if*
$p < p_0$	$Z < -z_\alpha$ i.e., $\lvert Z \rvert > z_\alpha$
$p > p_0$	$Z > z_\alpha$ i.e., $\lvert Z \rvert > z_\alpha$
$p \neq p_0$	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$ i.e., $\lvert Z \rvert > z_{\alpha/2}$

Remarks:

1. If the sampling is from a finite population of size N then standard error of sample proportion $\hat{p}$ is given by

$$S.E.(\hat{p}) = \sqrt{\frac{N - n}{N - 1} \frac{pq}{n}}$$

Then test statistic of sample proportion $\hat{p}$ is $Z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{N-n}{N-1}\dfrac{pq}{n}}} \sim N(0, 1)$

2. $(1 - \alpha)$ 100% confidence for estimating population proportion p is given by C.I. for $p = \hat{p} \pm z_{\alpha/2}\, S.E.(\hat{p})$

Example 2: (*Survey of Drivers - One sided test of proportion*): Of 880 randomly selected drivers, 56% admitted that they run red lights. An unknown reporter wrote this:" Nearly all Nepali drivers agree that running red lights is dangerous, but more than half admit they've done it, . . . , a survey found." Formulate and test an appropriate set of hypotheses.

Solution: The Traditional method

The original claim in symbolic form is $p > 0.5$(i.e., the majority or more than half of all Nepalese drivers run red lights.)

The opposite of the original claim is $p \le 0.5$.

Step 1. *Set up hypothesis*: Of the preceding two symbolic expressions, the expression $p > 0.5$ does not contain equality, so it becomes the alternative hypothesis. The null hypothesis is the statement that p equals the fixed value of 0.5. We can therefore express H_0 and H_1 as follows:

Null hypothesis: $H_0: p = 0.5$

Alternative hypothesis: $H_1: p > 0.5$

Step 2. *Level of significance* (α): In the absence of any special circumstances, we will select $\alpha = 0.05$ for the significance level.

(Because we are testing a claim about a population proportion p, the sample statistic $\hat{p}$ is relevant to this test, and the sampling distribution of sample proportions $\hat{p}$ is approximated by a normal distribution.)

Step 4. *Test statistic*: The test statistic is evaluated using $n = 880$ and $\hat{p} = 0.56$. In the null hypothesis we are assuming that $p = 0.5$, so $q = 1 - 0.5 = 0.5$. The test statistic is

$$z = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

This is a right tailed test, so the critical region is an area of $\alpha = 0.5$ in the right tail. Referring to Standard Normal Table A-3, we find that the critical value of $z_\alpha = 1.645$ is at the boundary of the critical region.

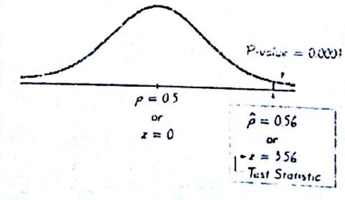
Step 5. *Decision*: Because the test statistic falls within the critical region, i.e., $|z| > z_\alpha$, we reject the null hypothesis.

Step 6. *Conclusion*: We conclude that there is sufficient sample evidence to support the claim that the majority of Nepalese drivers admit to running red lights.

The P-Value Method

For the hypothesis test described in the preceding example, the first three steps of the P-value method are the same as those shown in the above traditional method, so we now continue with Step 4.

Step 4. The test statistic is $z = 3.56$ as shown in the preceding traditional method. We now find the P-value (instead of the critical value) by using the following procedure, which is shown in Figure.



Right-tailed test:	P-value	=	area to right of test statistic Z
Left-tailed test:	P-value	=	area to left of test statistic Z
Two-tailed test:	P-value	=	twice the area of the extreme region bounded by the test statistic Z

Because the hypothesis test we are considering is right-tailed with a te statistic of $z = 3.56$, the P-value is the area to the right of $z = 3.5$. Referring to Standard Normal Table A-3, we see that for values of $z = 3.5$ and higher, we use 0.9999 for the cumulative area to the left of the te statistic. The area to the right of $z = 3.56$ is therefore $1 - 0.9999 = 0.000$ We now know that the P-value is 0.0001.

Step 5. Because the P-value of 0.0001 is less than or equal to the significance lev of $\alpha = 0.05$, we reject the null hypothesis

Step 6. As with the traditional method, we conclude that there is sufficient sampl evidence to support the claim that the majority of Nepalese admit t running red lights.

Example 3: (*Mendel's Genetics Experiments-Two sided test of proportion*): When Gregor Mendel conducted his famous hybridization experiments with pea one such experiment resulted in offspring consisting of 428 peas with green pods an 152 peas with yellow pods. According to Mendel's theory, ¼ of the offspring pea should have yellow pods. Use a 0.05 significance level with the P-value method t test the claim that the proportion of peas with yellow pods is equal to ¼.

Solution: After verifying that the assumptions are all satisfied, we begin with the P-value method. Note that $n = 428 + 152 = 580$, $\hat{p} = 152/580 = 0.262$, and for the purpose of the test, we assume that $p = 0.25$.

The original claim is that the proportion of peas with yellow pods is equal to ¼. We express this in symbolic form as $p = 0.25$.

The opposite of the original claim is $p \ne 0.25$.

Step 1. *Set up hypothesis*: Because $p \ne 0.25$ does not contain equality, it becomes H_1. We get

Null hypothesis: $H_0: p = 0.25$ (original claim)

Alternate Hypothesis $H_1: p \ne 0.25$

Step 2. *Level of significance* (α): The significance level is $\alpha = 0.05$.

Because the claim involves the proportion p, the statistic relevant to this test i the sample proportion $\hat{p}$, and the sampling distribution of sample proportion is approximated by the normal distribution (provided that the approximated at satisfied). (The requirements $np \ge 5$ and $nq \ge 5$ are both satisfied with $n = 58($ $p = 0.25$, and $q = 0.75$.)

Step 6. *Test statistic*: The test statistic of $z = 0.67$ is found as follows:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.262 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{580}}} = 0.67$$

For this two-tailed test with the test statistic located to the right of the cente (because $z = 0.67$ is positive), the P-value is *twice* the area to the right of the tes statistic. Using Normal Table A-3, $z = 0.67$ has an area of 0.7486 to its left, so the area to the tight of $z = 0.67$ is $1 - 0.7486 = 0.2514$, which we double to get 0.5028.

Step 7. *Decision*: Because the *P*-value of 0.5028 is greater than the significance level of 0.05, we fail to reject the null hypothesis.

Interpretation: The methods of hypothesis testing never allow us to support a claim of equality, so we cannot conclude that the proportion of peas with yellow pods is equal to ¼. Here is the correct conclusion: There is not sufficient evidence to warrant rejection of the claim that ¼ of the offspring peas have yellow pods.

Example 4: (***One sided test of proportion***): A manufacturer of lenses is qualifying a new grinding machine and will qualify the machine if the percentage of polished lenses that contain surface defects does not exceed 2%. A random sample of 250 lenses contains six defective lenses. Formulate and test an appropriate set of hypotheses to determine if the machine can be qualified.

Solution: Here, population proportion, $p = p_0 = 0.02$, where p is the maximum probability to say that a machine can be qualified. If this is not true, then we can conclude that the machine cannot qualify. And

Sample size, $n = 250$, Sample proportion $\hat{p} = 6/250 = 0.024$

Therefore, standard error $= \sqrt{\frac{p(1-p)}{n}} = \frac{0.02(1-0.02)}{250} = 0.0089$

Step 1. ***Null hypothesis***: H_0: $p = p_0 = 0.02$. That is, we assume that the population from where the sample is drawn has the maximum proportion defective, 0.02. Or otherwise, we assume the percentage that the qualifying machine will contain maximum of polished lenses with surface defects is 2%.

Step 2. ***Alternative hypothesis***: H_1: $p > p_0$. That is the sample has come from a population whose proportion defective is more than 0.02. Or otherwise, we assume that the new grinding machine will not qualify if the percentages of polished lenses that contain surface defects exceed the maximum of 2%. Therefore, the proposed test is a right–tail test.

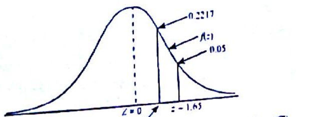
Step 3. ***Level of significance***: Since no specific level of significance (α) is proposed, it can be assumed as 5%. When $\alpha = 5\%$

or, $\alpha = 0.05$, we have for right–tail test $+z_\alpha = +z_{0.05} = +1.65$

Step 4. ***Test statistic***: Under H_0: $p = p_0$

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} = \frac{0.0254 - 0.02}{0.0089} = 0.607$$

Step 5. ***Decision***: Since calculated *Z*-value, *i.e.*, $z = 0.607$ is less than the tabulated *Z* value, i.e., $+z_{0.05} = +1.65$ the hypothesis is not rejected at 5% level of significance. Therefore, we conclude that the sample whose proportion is 0.0254 (6 defectives out of 250) has come from the population with proportion defective 0.02. Therefore, the machine can be said to be qualified.



Critical values for 5% (right tail) and the actual value (Z).
e.g. $P(Z \geq +1.65) = 0.05$.

P-value approach: We can also compare the level of significance α on the right - tail with the *p-value*, i.e., $P(Z > 0.607) = 1 - P(Z < 0.607) = 0.2719$ (from standard normal table A-3). Here, since 0.2719 is more than 0.05, we do not reject the null hypothesis at 5% level of significance.

Example 5: (***One sided test of proportion***): A semiconductor manufacturer produces controllers used in automobile engine applications. The customer does not prefer that the process fallout or fraction defective at a critical manufacturing greater than or equal to 0.05. The semiconductor manufacturer takes a random sample of 200 devices and finds that four of them are defective. Can the manufacturer demonstrate process capability for the customer? Test at 5% level of significance.

Solution: Here population proportion, $p = p_0 = 0.05$, where p is the proportion defective to say that the manufacturer does not conform to the required quality. If this is not true, then we can conclude that the manufacturer conforms to quality. Also, sample size, $n = 200$ and Sample proportion, $\hat{p} = \frac{4}{200} = 0.02$.

Therefore, standard error $= \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.05(1-0.05)}{250}} = 0.0154$

Step 1. ***Null hypothesis***: H_0: $p = p_0 = 0.05$. That is we assume that the population from where the sample is drawn has the proportion defective equal to 0.05. Or otherwise, we assume that at this level of proportion defective the manufacturer cannot demonstrate the capability of his product.

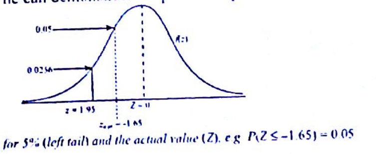
Step 2. ***Alternative hypothesis***: H_1: $p < p_0 = 0.05$. That is, the sample has come from a population whose proportion defective is less than 0.05. Or otherwise, we assume that the manufacturer can demonstrate process capability of its product. Therefore, the proposed test is a left-tail test.

Step 3. ***Level of significance***: Level of significance (α) is specified as 5%,. When $\alpha = 5\%$ or 0.05, we have for left-tail test $z_\alpha = z_{0.05} = -1.65$.

Step 4. ***Test statistic***: Under $H_0: p = p_0$

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} = \frac{0.02 - 0.05}{0.0154} = -1.95$$

Step 5. ***Decision***: Since the calculated *Z* value, i.e., $z = -1.95$ is less than the tabulated *Z* values, i.e., $+z_{0.05} = -1.65$, the hypothesis is rejected at 5%. Therefore, we conclude that the manufacturer is capable of meeting the requirement of the customer that the proportion defective should not exceed 0.05 and hence he can demonstrate the process capability for the customer.



Critical Values for 5% (left tail) and the actual value (Z). e.g. $P(Z \leq -1.65) = 0.05$

P-value approach:

We can also compare the level of significance α on the right tail with the p-value, i.e., $P(Z < -1.95) = 0.0256$ (*from standard normal table* A-3). Here, since 0.0256 is less than 0.05 we reject the null hypothesis at 5% level of significance.

Example 6: (*A stone test of proportion of transceivers:- one sided test*):
Transceivers provide wireless communication among electronic components of consumer products. Responding to a need for a fast, low–cost test of Bluetooth–capable transceivers, engineers developed a product test at the wafer level. In one set of trails with 60 devices selected from different wafer lots, 48 devices passed. Test the null hypothesis $p = 0.70$ against the alternative hypothesis $p > 0.70$ at the 0.95 level of significance.

Solution:

Step 1. *Null hypothesis* H_0: $p = 0.70$

Alternative hypothesis H_1: $p > 0.70$

Step 2. *Level of significance*: $\alpha = 0.05$

Step 3. *Test statistic*: Under H_0: $p = p_0$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{48 - 60(0.70)}{\sqrt{60(0.70)(0.30)}} = 1.69$$

Step 4. *Critical value*: $\alpha = 0.95$, $z_\alpha = z_{0.95} = 1.645$

Step 5. *Decision*: Since $z = 1.69$ is greater than 1.645, we reject the null hypothesis at level 0.05. In other words, there is sufficient evidence to conclude that the proportion of good transceivers that would be produced is greater than 0.70. The P–value, $P(Z > 1.69) = 0.455$, somewhat strengthens this conclusion.

Example 7: (*A one sided test concerning a population proportion p*): In a study designed to investigate whether certain detonator used with explosives in coal mining meet the requirement that at least 90% will ignite the explosive when charged. It is found that 174 out of 200 detonator function properly. Test the null hypothesis $p = 0.90$ against the alternative hypothesis $p < 0.90$ at the 0.05 level of significance. *[TU, BIE 2068 Jestha]*

Solution: Data given: x = number of success = 174, $n = 200$,

Step 1. *Null hypothesis*: H_0: $p = 0.90$ $(p = p_0)$

Alternative hypothesis: H_1: $p < 0.90$ (*left tailed test*)

Step 2. *Level of significance*: $\alpha = 0.05$

Step 3. *Critical value*: At $\alpha = 5\% = 0.05$, $z_\alpha = -1.645$

Criterion: Reject H_0 if $Z < -1.645$

where $z = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}}$

Step 4. *Test statistic*: $z = \frac{174 - 200(0.90)}{\sqrt{200(0.90)(0.10)}} = -1.41$

Step 5. *Decision*: Since $z > z_\alpha$, the null hypothesis H_0 cannot be rejected. In other words there is no sufficient evidence to say that the given kind of detonator fails to meet the required standard.

Example 8: (*Two sided test*): In a sample of 1000 people in Kathmandu district, 540 speak Nepali and rest speak Newari. Can we assume that both languages are equally popular in this district at 1% and 5% level of significance?

Solution : Date given: $n = 1000$ (samples size),

x = Number of people who speak Nepali = 540

Step 1. *Null hypothesis* H_0: $p = \frac{1}{2}$ (both are equally popular)

Alternative hypothesis H_1: $p \neq \frac{1}{2}$

Step 2. *Test statistic*: Under H_0, statistic

$$z = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{540 - 1000(\frac{1}{2})}{\sqrt{1000(\frac{1}{2})(1 - \frac{1}{2})}} = 2.53$$

Step 3. *Decision*: Significance or critical value of Z at 1% level of significance is $z_{0.005} = 2.58$ since the computed value of Z is less than critical v (*i.e.*, $z < z_{\alpha/2}$) so we accept H_0. Hence we conclude both languages are equ popular. Again, at $\alpha = 5\% = 0.05$ level $z_{\alpha/2} = z_{0.025} = 1.96$. Since $z > z_{\alpha/2}$ so, reject H_0. Hence at $\alpha = 5\%$ level of both languages are not equally popular.

Example 9: (*One sided test*): After a careful analysis, a company contemplating introduction of a new product has determined that it must capture a market shar 10% to break even. Anything greater than 10% will result in a profit for the comp In a survey, 400 potential customers are asked whether or not they would purchase product. If 52 people respond affirmatively, is this enough evidence to enable company to conclude that the product will produce a profit? (Use $\alpha = 0.05$)

Solution: Data given: Population proportion $p = 10\% = 0.10$,

Sample size $n = 400$

No. of successes $x = 52$, sample proportion $\hat{p} = \frac{x}{n} = 0.13$

Step 1. *Null hypothesis* H_0: $p = 0.10$

Alternative Hypothesis: $p > 0.10$. (*Right tailed test*)

Step 2. *Level of significance*: At $\alpha = 0.05$

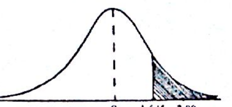
Step 3. *Test statistic*: Under H_0,

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{52 - 400 \times 0.10}{\sqrt{400 \times (0.10)(0.90)}} = 2$$

Step 4. At $\alpha = 0.05$, $z_\alpha = 1.645$

Step 5. *Decision*: As $z = 2 > 1.645$, reject H_0.

There is enough evidence to allow us to conclude that the product contribute a profit to the company.



Example 10: (*A one-sided test of the proportion of transceivers*): Transceiv provide wireless communication among electronic components of consumer produ Responding to a need for a fast, low-cost test of Bluetooth-capable transceiv engineers developed a product test at the wafer level. In one set of trials with 60 dev selected from different wafer lots, 48 devices passed. Test the null hypothesis p = (against the alternative hypothesis $p > 0.70$ at the 0.95 level of significance.

Solution:

Step 1. *Null hypothesis*: $p = 0.70$

Alternative hypothesis: $p > 0.70$.

Step 2. *Level of significance*: $\alpha = 0.05$

Step 3. *Criterion*: Reject the null hypothesis if $Z > 1.645$, where

$$Z = \frac{X - np_0}{\sqrt{n\,p_0(1 - p_0)}}$$

Sep 4. *Calculations*: Substituting $x = 48$, $n = 60$, and $p_0 = 0.70$ into the formula above, we get

$$z = \frac{48 - 60(0.70)}{\sqrt{60(0.07)\,(0.30)}} = 1.69$$

Step 5. *Decision*: Since $z = 1.69$ is greater than 1.645, we reject the null hypothesis at level 0.05. In other words, there is sufficient evidence to conclude that the proportion of good transceivers that would be produced is greater than 0.70. The P - value, $P(Z > 1.69) = 0.0455$, somewhat strengthens this conclusion.

Example 11: A semiconductor firm produces logic devices. The contract with their customer calls for a fraction defective of no more than 0.05. They wish to test

H_0: $p = 0.05$,

H_1: $p_1 > 0.05$.

A random sample of 200 devices yields six defectives. The test statistic is

$$z = \frac{x - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{6 - 200(0.05)}{\sqrt{200\,(0.05)(0.95)}} = -1.30$$

Using $\alpha = 0.05$, we find that $z_{0.05} = 1.645$, and so we cannot reject the null hypothesis that $p = 0.05$.

9.3 Test of Hypotheses on Two Proportions

There are many real and important situations in which it is necessary to use sample data to compare two population proportions. In fact, a strong argument could be made that this section is one of the most important sections in the book because this is where we describe methods for dealing with two sample proportions. Although this section is based on proportions, we can deal with probabilities or we can deal with percentages by using the corresponding decimal equivalents. For example, we might want to determine whether there is a difference between the percentage of adverse reactions in a placebo group and the percentage of adverse reactions in a drug treatment group. We can convert the percentages to their corresponding decimal values and proceed to use the methods of this section.

When testing a hypothesis made about two population proportions or when constructing a confidence interval for the difference between two population proportions, we make the following assumptions and use the following notation.

Assumptions

1. We have proportions from two simple random samples that are *independent*, which means that the sample values selected from one population are not related to or somehow paired or matched with the sample values selected from the other population.
2. For both samples, the conditions $np \geq 5$ and $nq \geq 5$ are satisfied. That is, there are at least five successes and five failure in each of the two samples. (in many cases, we will test the claim that two populations have equal proportions so that $p_1 - p_2 = 0$. Because we assume that $p_1 - p_2 = 0$, it is not necessary to specify the particular value that p_1 and p_2 have in common. In such cases, the conditions $np \geq 5$ and $nq \geq 5$ can be checked by replacing p with the estimated pooled proportion S_P, which will be described later.)

Notation for Two Proportions

For population 1 we let

p_1 = *population* proportion

n_1 = size of the sample

x_1 = number of successes in the sample

$\hat{p}$ = (the *sample* proportion); $\hat{q}_1 = 1 - \hat{p}_1$

The corresponding meanings are attached to p_2, n_2, x_2, $\hat{p}_2$ and $\hat{q}_2$ which come from population 2.

Example 12: (*Finding the Numbers of Successes x_1 and x_2*): (The calculations for hypothesis tests and confidence intervals require that we have specific values for x_1, n_1, x_2, and n_2. Sometimes the available sample data include those specific numbers, but sometimes it is necessary to calculate the values of x_1 and x_2.)

Consider the statement that "when 734 men were treated with Viagra, 16% of them experienced headaches." and find x_1, x_2.

Solution: From that statement we can see that $n_1 = 734$ and $\hat{p}_1 = 0.16$, but the actual number of successes x_1 is not given. However, from $\hat{p} = x/n$, we know that $x_1 = n_1\hat{p}_1$ so that $x_1 = (734)(0.16) = 117.44$. But you cannot have 117.44 men who experienced headaches, because everyone either experiences a headache or not, and the number of successes x_1 must therefore be a whole number. We can round 117.44 to 117. We can now use $x_1 = 117$ in the calculations that require its value. It's really quite simple: 16% of 734 means 0.16×734, which results in 117.44, which we round to 117.

Hypotheses Tests

In previous section, we discussed tests of hypotheses made about a single population proportion. We will now consider tests of hypotheses made about two population proportions, but while testing claims that $p_1 = p_2$, we will use the following pooled (or combined) estimate of the value that p_1 and p_2 have in common. We can see from the form of the *pooled estimate* $\hat{p}$ that it basically combines the two different samples into one big sample.

Pooled Estimate of p_1 and p_2:

The pooled estimate of p_1 and p_2 is denoted by $\hat{p}$ and is given by $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$

We denote the complement of $\hat{p}$ by $\hat{q}$, so $\hat{q} = 1 - \hat{p}$.

Test Statistic for Two proportions (with H_0: $p_1 = p_2$)

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

where $p_1 - p_2 = 0$ (assumed in the null hypothesis)

$$\hat{p}_1 = \frac{x_1}{n_1} \text{ and } \hat{p}_2 = \frac{x_2}{n_2};\ \hat{p} = \frac{x_1 + x_2}{n_1 + n_2};\ \hat{q} = 1 - \hat{p}.$$

Testing for Constant Difference:

To test the null hypothesis that the difference between two population proportions is equal to a nonzero constant δ_0, use the test statistic

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \delta_0}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

As long as n_1 and n_2 are both large, the sampling distribution of the test statistic Z will be approximately the standard normal distribution.

Test Statistic for Two proportions (with $H_0: p_1 \neq p_2$)

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \delta_0}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

where $p_1 - p_2 = \delta_0$ (assumed in the null hypothesis)

$$\hat{p}_1 = \frac{x_1}{n_1} \text{ and } \hat{p}_2 = \frac{x_2}{n_2}; \quad \hat{q} = 1 - \hat{p}.$$

9.3.1 Large sample test significance of difference of two proportions

If there are independent populations with proportions p_1 and p_2, we follow the following procedure for testing the significance of difference between the two population proportions.

Step 1. *Set up hypothesis*

For equality of proportions $p_1 = p_2$ (i.e., $p_1 - p_2 = \delta_0 = 0$)

Null hypothesis: H_0: $p_1 = p_2$ i.e. two independent population proportions are same.

Alternative hypothesis: H_1: $p_1 \neq p_2$ (*Two tailed test*)

or, H_1: $p_1 > p_2$ (*Right tailed test*)

or, H_1: $p_1 < p_2$ (*Left tailed test*)

For inequality of proportions $p_1 \neq p_2$ (i.e., $p_1 - p_2 = \delta_0 \neq 0$)

Null hypothesis: H_0: $p_1 - p_2 = \delta_0$

Alternative hypothesis: H_1: $p_1 - p_2 \neq \delta_0$ (*Two tailed test*)

or, H_1: $p_1 - p_2 > \delta_0$ (*Right tailed test*)

or, H_1: $p_1 - p_2 < \delta_0$ (*Left tailed test*)

Step 2. ***Level of significance*** (α): Choose a appropriate level of significance. The most commonly used is $\alpha = 5\%$ unless otherwise stated.

Step 3. ***Test statistic:*** Under H_0: $p_1 = p_2$, the test statistic is

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad [\because H_0 : p_1 = p_2]$$

Under H_0: $p_1 - p_2 = \delta_0$, the test statistic is

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \delta_0}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

where $\hat{p}_1 = \frac{x_1}{n_1}$ = sample proportion of the first population.

$\hat{p}_2 = \frac{x_2}{n_2}$ = sample proportion of the second population

n_1 = sample size taken from first population

n_2 = sample size taken from second population

If $\hat{p}$ the common proportion is not known, we estimate $\hat{p}$ as follows

$$\hat{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

Step 4. ***Critical value:*** Obtain critical or tabulated value of test statistic Z at level of significance (α).

Step 5. ***Decision:*** If $|Z| > z_\alpha$ for one tailed test or, $|Z| > z_{\alpha/2}$ for two tailed test, then reject H_0 and accept H_1.

If $|Z| \leq z_\alpha$ for one tailed test or, $|Z| \leq z_{\alpha/2}$ for two tailed test, then accept H_0 and reject H_1.

Summary of Decision Rule for Z-test

Critical Regions for Testing $p = p_0$ (Large samples)	
Alternative hypothesis	*Reject H_0 if*
$p < p_0$	$Z < -z_\alpha$ i.e., $\lvert Z\rvert > z_\alpha$
$p > p_0$	$Z > z_\alpha$ i.e., $\lvert Z\rvert > z_\alpha$
$p \neq p_0$	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$ i.e., $\lvert Z\rvert > z_{\alpha/2}$

Remarks:

1. To test the null hypothesis that the difference between the two populatio proportions equals some constant. So, not necessarily 0, we can use the statist

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \delta_0}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

2. The $(1-\alpha)$ 100% Large sample confidence interval for estimating t difference of two proportions $(p_1 - p_2)$ is

$$C.I. \text{ for } (p_1 - p_2) = \left[(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}\right]$$

$$= \left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) \pm z_{\alpha/2}\sqrt{\frac{\frac{x_1}{n_1}\left(1 - \frac{x_1}{n_1}\right)}{n_1} + \frac{\frac{x_2}{n_2}\left(1 - \frac{x_2}{n_2}\right)}{n_2}}$$

Example 13: (*Racial Profiling*): For the sample data listed in following table, us 0.05 significance level to test the claim that the proportion of black drivers stopp by the police is greater than the proportion of white drivers who are stopped.

Black	White Drivers
$n_1 = 200$	$n_2 = 1400$
$x_1 = 24$	$x_2 = 147$
$\hat{p}_1 = \frac{x_1}{n_1} = \frac{24}{200} = 0.120$	$\hat{p}_2 = \frac{x_2}{n_2} = \frac{147}{1400} = 0.105$

Solution: (***We will now use the P-value method of hypothesis testing***)

The claim of a greater rate for black drivers can be represented by $p_1 > p_2$

If $p_1 > p_2$ is false, then $p_1 \leq p_2$.

Step 1. Because our claim of $p_1 > p_2$ does not contain equality, it becomes alternative hypothesis. Then null hypothesis is the statement equality, so we have

Null hypothesis $H_0: p_1 = p_2$

Alternative hypothesis $H_1: p_1 > p_2$ (original claim)

Level of significance: The significance level is $\alpha = 0.05$

Test statistic: We will use normal distribution (with the test statistic previously given) as an approximation to the binomial distribution. We have two independent samples, and the conditions $np \geq 5$ and $nq \geq 5$ are satisfied for each of the two samples. To check this, we note that in conducting this test, we assume that $p_1 = p_2$, where their common value is the pooled estimate $\hat{p}$ calculated as shown below with extra decimal places used to minimize rounding errors in later calculations.

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{24 + 147}{200 + 1400} = 0.106875$$

We $\hat{p} = 0.106875$, it follows that $\hat{q} = 1 - 0.106875 = 0.893125$.

We verify that $np \geq 5$ and $nq \geq 5$ for both samples as shown below, with p estimated by $\hat{p}$ and with q estimated by $\hat{q}$

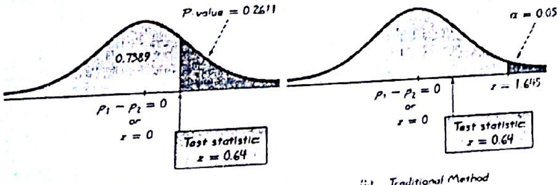
Sample 1	Sample 2
$n_1 p = (200)(0.106875) = 21.375 \geq 5$	$n_2 p = (1400)(0.106875) = 149.625 \geq 15$
$n_1 q = (200)(0.893125) = 178.625 \geq 5$	$n_2 q = (1400)(0.893125) = 1250.375 \geq 5$

Step 6. We can now find the value of the test statistic

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}\,\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 0.64$$

The P-value of 0.2611 is found as follows: This is a right-tailed test, so the *P*-value is the area to the right of the test statistic z = 0.64. (See Figure) Refer to Table A-3 and find that the area to the left of the test statistic z = 0.64 is 0.7389, so the *P*-value is 1 – 0.7389 = 0.2611. The test statistic and *P*-value are shown in Figure.

Step 7: Because the *P*-value of 0.2611 is greater than the significance level of $\alpha = 0.05$, we fail to reject the null hypothesis of $p_1 = p_2$.



(a) P-Value Method (b) Traditional Method

Interpretation: We must address the original claim that black drivers get stopped at a greater rate than white drivers. Because we fail to reject the null hypothesis, we conclude that there is not sufficient evidence to support the claim that the proportion of black drivers stopped by police is greater than that for white drivers. (See Figure for help in wording the final conclusion.) This does not mean that racial profiling has been disproved. It

means only that the evidence is not yet strong enough to conclude that the 12.0% rate for stopping black drivers is significantly greater than the 10.5% rate for stopping white drivers. The evidence might be strong enough with more data. In fact, data sets larger than those used in this example do suggest that racial profiling has been in effect.

Example 14: *(Racial Profiling:- Two Sided Confidence Interval)*: Use the sample data given in Table of previous example to construct a 90% confidence interval estimate of the difference between the two population proportions. (The confidence level of 90% is comparable to the significance level of a $\alpha = 0.05$ used in the preceding right-tailed hypothesis test.)

Solution: With a 90% confidence level, $z_{\alpha/2} = 1.645$, a 90% confidence interval estimate of the difference between the two population proportions is

$$C.I.\text{ for }(p_1 - p_2) = \frac{x_1}{n_1} - \frac{x_2}{n_2} \pm z_{\alpha/2}\sqrt{\frac{\frac{x_1}{n_1}\left(1 - \frac{x_1}{n_1}\right)}{n_1} + \frac{\frac{x_2}{n_2}\left(1 - \frac{x_2}{n_2}\right)}{n_2}}$$

$$= 0.120 - 0.105 \pm 1.645\sqrt{\frac{(0.120)(1-0.120)}{200} + \frac{(0.105)(1-0.105)}{1400}}$$

or, $-0.025 < p_1 - p_2 < 0.055$

where $\hat{p}_1 = \frac{x_1}{n_1} = \frac{24}{200} = 0.120$ and $\hat{p}_2 = \frac{x_2}{n_2} = \frac{147}{1400} = 0.105$.

Interpretation: The confidence interval limits do contain 0, suggesting that there is not a significant difference between the two proportions. However, if the goal is to test for equality of the two population proportions, we should use *P*-value or traditional method of hypothesis testing; we should not base the decision on the confidence interval.

Example 15: Random samples of 400 male workers and 600 female workers were asked about their opinion of a project proposal on quality improvement. 200 male workers and 325 female workers were in favor of the proposal. Test the hypothesis that proportions of men and women in favor of proposal are same at 5% level of significance. *[TU BE, 2065 Chaitra]*

Solution: It is given that sample sizes $n_1 = 400$ and $n_2 = 600$ (large samples);

x = Number of male workers favoring proposal = 200 and

y = Number of female workers favoring proposal = 325.

Therefore, sample proportions are obtained as

$\hat{p}_1$ = The proportion of men favoring the proposal = 200/400 = 0.5

$\hat{p}_2$ = The proportion of women favoring the proposal = 325/600 = 0.541

It may be noted that these samples are drawn independently from the same organization, and hence we can pool to get an estimate of the overall population proportion, say $\hat{p}$. That is, the overall opinion favoring the proposal when both men and women are put together. Therefore, we have

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{x + y}{n_1 + n_2} = \frac{200 + 325}{400 + 600} = 0.525$$

Step 1. *Null hypothesis* $H_0: p_1 = p_2$. This means the proportion of men favoring the proposal and the proportion of women favoring the proposal are assumed equal.

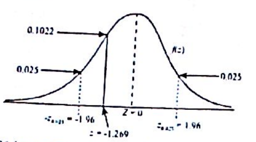
Alternative hypothesis: $H_1: p_1 \neq p_2$. This implies that the opinions of men and women workers on the proposal are not same. Therefore, it is a two tail test.

Step 2. *Level of significance* : Since level of significance (α) is specified as 5%, we have from standard normal table when α = 5% = 0.05, we have for left tail $-z_{\alpha/2} = -z_{0.025} = -1.96$ and for right tail $+z_{\alpha/2} = +z_{0.025} = +1.96$ i.e., for both tail $z_{\alpha/2} = 1.96$.

Step 3. *Test statistic*: Under $H_0: p_1 = p_2$

$$z = \frac{(\hat{p}-\hat{p})-(p_1-p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(0.5-0.541)-0}{\sqrt{(0.525)(1-(0.525))\left(\frac{1}{400}+\frac{1}{600}\right)}}$$

$$= \frac{-0.041}{0.0323} = -1.269$$



Critical values for 5% (two-tail), and the actual value (z) i.e., P(Z ≤ −1.269) = 0.1022

Step 4. *Decision*: Since the calculated Z value, *i.e.*, $z = -1.269$ is greater than the tabulated Z value, i.e. $z_{0.025} = -1.96$, at left tail, the hypothesis is not rejected at 5%. Therefore, we conclude that the proportions of opinions of men and women in favor of the project proposal are same.

P–value approach

We can also compare the level of significance α on the right tail with the p–value, i.e., $P(Z < -1.269) = 0.1022$ (from standard normal table A-3). Here, since 0.1022 is greater than 0.025, we do not reject the null hypothesis at 5% level of significance.

Example 16: From a lot of units produced by Machine A, a sample of 500 is drawn and tested for a quality characteristic. It is found that 16 units are not meeting the specification. Another sample of size 100 is drawn from the lot of similar units produced by Machine B and tested. In this case, only 3 units are found to be not meeting the specification. Test at 1% level of significance, whether there are in any significant difference of the proportions of defective units produced by the two machines.

Solution: It is given that sample size $n_1 = 500$ and $n_2 = 100$ (large samples); $x = 16$ and $y = 3$. Therefore, sample proportions are obtained as

$\hat{p}_1$ = The proportion of defectives produced by A = 16/500 = 0.032

$\hat{p}_2$ = The proportion of defectives produced by B = 3/100 = 0.030

It may be noted that these samples are from different populations (machines) so that we cannot pool the sample proportions and hence $\hat{p}_1$ and $\hat{p}_2$ are used as estimated for the unknown population proportions p_1 and p_2.

Step 1. *Null hypothesis*: $H_0: p_1 = p_2$. This means that the unknown populations related to the two machines are assumed equal.

Step 2. *Alternative hypothesis*: $H_1: p_1 \neq p_2$. This implies that the population proportions are not equal.

Step 3. *Level of significance*: Since level of significance (α) is specified as 1% we have from standard normal table A-3, when $\alpha = 1\% = 0.01$, we have fo left tail $-z_{\alpha/2} = -z_{0.005} = -2.58$ and for right tail $+z_{\alpha/2} = +z_{0.005} = +2.58$. Hence it is a two–tailed.

Step 4. *Test statistic*: Under $H_0: p_1 = p_2$, test statistic is

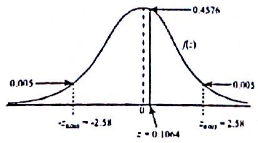
$$z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{(0.032-0.030)-0}{\sqrt{\frac{0.032(1-0.032)}{500}+\frac{0.03(1-0.03)}{100}}}$$

$$= \frac{0.002}{0.0188} = 0.1064$$

Step 5. *Decision*: Since the calculated Z–value, i.e., $z = 0.1064$ is less that tabulated Z–value i.e. $+z_{0.005} = 2.58$, at right tail, the hypothesis is no rejected at 1%. Therefore, we conclude that there is no significant differen between the two population proportions of two machines.

P–value approach:

We can also compare the level of significance α on the right tail with the p-value i.e., $P(Z > 0.1064) = 1 - P(Z < 0.1064) = 0.4576$ (from standard normal table A-3) Here, since 0.4576 is greater than 0.005, we do not reject the null hypothesis at 1% level of significance.



Critical values for 1% (two tail), and the actual value (z). eg. P(Z ≥ 0.1064) = 0.4576

Example 17(i): A study shows that 16 of 200 tractors produced on one assembly line required extensive adjustments before they could be shipped, while the same was true for 14 of 400 tractors produced on another assembly line. At the 0.01 level of significance, does this support the claim that the second production line does superior work?

Solution :

Step 1. *Null hypothesis*: $p_1 = p_2$

Alternative hypothesis: $p_1 > p_2$

Step 2. *Level of significance*: $\alpha = 0.01$

Step 3. *Criterion*: Reject the null hypothesis if $Z > 2.33$.

Step 4. *Test Statistic*: Substituting $x_1 = 16$, $n_1 = 200$, $x_2 = 14$, $n_2 = 400$, and

$\hat{p} = \frac{16+14}{200+400} = 0.05$ [∵ two samples are from same population]

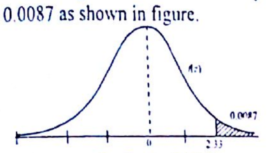
into the formula for Z, we get

$$z = \frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{\frac{16}{200}-\frac{14}{400}}{\sqrt{(0.05)(0.95)\left(\frac{1}{200}+\frac{1}{400}\right)}} = 2.38$$

Step 5. *Decision*: Since $z = 2.38$ exceeds 2.33, the null hypothesis must be rejected; we conclude that the true proportion of tractors requiring

Scanned by CamScanner

extensive adjustments is greater for first assembly line than for the second. The P-value is 0.0087 as shown in figure.



The P–value 0.0087

Example 17(ii): (*A large sample confidence interval for the difference of two proportions*). With reference to the preceding example, find the large sample 95% confidence interval for $p_1 - p_2$.

Solution: *Large sample confidence interval for the difference of two proportions*

$$\boxed{\frac{x_1}{n_1} - \frac{x_2}{n_2} \pm z_{\alpha/2}\sqrt{\frac{\frac{x_1}{n_1}\left(1-\frac{x_1}{n_1}\right)}{n_1} + \frac{\frac{x_2}{n_2}\left(1-\frac{x_2}{n_2}\right)}{n_2}}}$$

Since $\hat{p}_1 = \frac{x_1}{n_1} = \frac{16}{200} = 0.08$ and $\hat{p}_2 = \frac{x_2}{n_2} = \frac{14}{400} = 0.035$

$$C.I. \text{ for } (p_1 - p_2) = \frac{x_1}{n_1} - \frac{x_2}{n_2} \pm z_{\alpha/2}\sqrt{\frac{\frac{x_1}{n_1}\left(1-\frac{x_1}{n_1}\right)}{n_1} + \frac{\frac{x_2}{n_2}\left(1-\frac{x_2}{n_2}\right)}{n_2}}$$

$$= 0.08 - 0.035 \pm 1.96\sqrt{\frac{(0.08)(0.92)}{200} + \frac{(0.035)(0.0965)}{400}}$$

or, $\quad 0.03 < p_1 - p_2 < 0.087$

The first assembly line has a rate of extensive adjustment between 3 and 87 out of 1,000, higher than the rate for the second assembly line.

Example 18: The owner of a wholesale distributing firm would like to know the proportion of accounts receivable that are more than 60 days past due. The owner estimates that in the past the proportion has remained stable at 15%. A random sample of 200 current accounts receivable revealed that 44 were more than 60 days past due using 0.05 level of significance is there evidence that are more than 60 days past due has changed?

Solution: Data given $p = 15\% = 0.15\ (= p_0)$; $q = 1 - p = 1 - 0.15 = 0.85\ (= q_0)$

$x = 44$, $n = 200$

Step 1. *Null hypothesis*: H_0: 0.15. That is the proportion of accounts receivable that are more than 60 days past due has remained stable at 15%.

Alternative hypothesis: H_1: $p \neq 0.15$ (two tailed test). That is the proportion of accounts receivable more than 60 days past due has not remained stable at 15%.

Step 2. *Level of significance*: $\alpha = 5\%$

Step 3. *Test statistic*: $z = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{44 - 200 \times 0.15}{\sqrt{200 \times 0.15 \times 0.85}} = \frac{14}{5.05} = 2.77$

Step 4. *Critical value* : The tabulated value of Z at 5% level of significance is 1.96 i.e., $z_{\alpha/2} = 1.96$.

Step 5. *Decision*: Since $|z| > z_{\alpha/2}$, it is significant and H_0 is rejected. Hence H_1 is accepted which means that there is evidence that the proportion of accounts receivable that are more than 60 days past due has changed.

Example 19: (*Testing for Constant Difference* i.e., $p_1 \neq p_2$ or $p_1 - p_2 = \delta_0 \neq 0$)
A cigarette manufacturing firm claims that its brand A of the cigarettes out sells its brand B by 8%. If it is found that 42 out of a sample 200 smokers prefer brand A and 18 out of another random sample of 100 smokers prefer brand B, test whether the stated 8% difference is a valid claim. Use 5% level of significance

[TU, BE, 2065 Kartik]

Solution : Following data are given:

$n_1 = 200$, $n_2 = 100$, $x_1 = 42$, $x_2 = 18$

The proportion of smokers preferring A, $\hat{p}_1 = 42/100 = 0.21$

The proportion of smokers preferring B, $\hat{p}_2 = 18/100 = 0.18$

Step 1. *Null hypothesis*: H_0: $p_1 - p_2 = 8\% = 0.08$

Alternative hypothesis: H_1: $p_1 - p_2 > 8\%$ (*Right tailed test*)

Step 2. *Level of significance*: $\alpha = 5\% = 0.05$. So, $z_\alpha = z_{0.05} = 2.58$

Step 3. *Test statistic*: Under H_0: $p_1 - p_2 = 0.08$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{(0.21 - 0.18) - 0.08}{\sqrt{\frac{(0.21)(0.79)}{200} + \frac{(0.18)(0.82)}{100}}} = -1.042$$

Step 4. *Decision*: Since $|z| < z_\alpha$ accept H_0 and reject H_1 Hence stated 8% difference is not a valid claim.

Example 20: (*Adverse Drug Reactions*).- The drug Viagra has become quite well known, and it has had a substantial economic impact on its producer. Pfizer Pharmaceuticals. In preliminary tests for adverse reactions, it was found that when 734 men were treated with Viagra, 16% of them experienced headaches. (There's some real irony there.) Among 725 men in a placebo group, 4% experienced headaches (based on data from Pfizer Pharmaceuticals).

(a) Using a 0.01 significance level, is there sufficient evidence to support the claim that among those men who take Viagra, headaches occur at a rate that is greater than the rate for those who do not take Viagra?

(b) Construct a 99% confidence interval estimate of the difference between the rate of headaches among Viagra users and the headache rate for those who are given a placebo. What does the confidence interval suggest about the two rates?

(c) Use a 0.05 significance level to test the claim that the headache rate of Viagra users is 10 percentage points more than the percentage for those who are given a placebo.

Solution: Similar to previous example.

9.4 Chi- Square (χ^2) Test

9.4.1 Introduction

The tests of significance like Z- test, t- test and F- test which are based on the assumption that the sample are drawn from normally distributed populations or approximately normally distributed populations are known as *parametric tests*.

χ^2- test (pronounced as *ki*) is a *non parametric test* because it depends only on the set of observed and expected frequencies and degrees of freedom. Since χ^2- test does not make any assumption about population parameters, it is also called a *distribution free test*. This test describes the magnitude of difference between observed frequencies and expected (theoretical) frequencies under certain assumptions.

It is defined as $\chi^2 = \sum \frac{(O - E)^2}{E}$

where O = observed frequencies

E = expected frequencies.

This test is good for nominal or ordinal scale of measurement. Nominal scale of measurement deals with the data which can be classified only into categories such as male and female, boy and girl, head and tail, juniors and seniors etc. whereas ordinal level of measurement assigns different ranks to above categorized data.

Definition: (*Chi-square Distribution*)

The square of standard normal variate is known as ***chi-square variate*** with 1 degree of freedom.

That is, if $Z = \frac{X - \mu}{\sigma} \sim N(0,1)$ then $Z^2 = \left(\frac{X - \mu}{\sigma}\right)^2$ is called a ***chi-square variate*** with 1 degree of freedom.

In general, if $X_1, X_2, \ldots, X_v$ are v independent normal variates with means $\mu_1, \mu_2, \ldots, \mu_v$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_v^2$ then

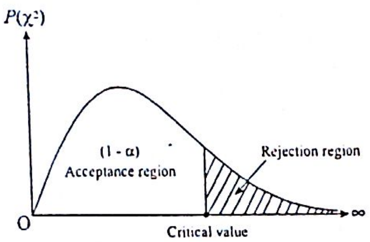
$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 = \left(\frac{X_1 - \mu_1}{\sigma_1}\right)^2 + \ldots + \left(\frac{X_n - \mu_v}{\sigma_n}\right)^2$$

Follows χ^2- distribution with v degrees of freedom.

The probability density function of chi-square distribution with v degree of freedom is given by

9.4.2 Properties of Chi-Square distribution

1. Since χ^2 is sum of squares, χ^2- distribution assumes non-negative values. That is, $0 \le \chi^2 \le \infty$.



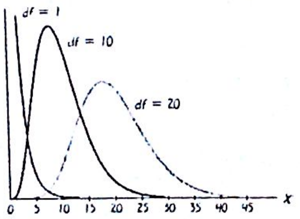


Figure: Chi-Square Distribution for 1, 10, and 20 Degrees of Freedom

χ^2 – distribution will be zero if difference of each pair is zero.

2. χ^2 – distribution is always positively skewed.
3. χ^2 – distribution for each number of degrees of freedom.
4. There is different χ^2 – distribution for each number of degrees of freedom.
5. It is based on one tailed test of the right hand side of the standard normal curve.
6. For χ^2 – distribution with v degrees of freedom (*d.f.*) mean = 2v, variance = 2v, mode = v − 2

9.4.3. Conditions for the validity of χ^2 - test

This test is used under the following assumptions:

1. Sample observations should be independent.
2. The observed frequency should be equal to the expected frequency.
3. The total frequency should be reasonably large, say, greater than 50.
4. No theoretical frequency should be less than 5.

If any theoretical frequency is less than 5, then for the applications of chi-square test it is pooled with the preceding or succeeding frequency so that the pooled frequency is more than 5 and adjusted the degree of freedom accordingly.

9.4.4 Applications of Chi-square distribution

The χ^2 – distribution has a large number of applications in statistics some of the application are;

1. To test 'goodness of fit'.
2. To test independence of attributes.
3. To test population variance.
4. To test the homogeneity.

Here we will discuss only 'goodness of fit' and independence of attributes.

9.5 χ^2 - test for Goodness-of-Fit

If we are given a set of observed frequencies obtained under some experiment and we are interested in knowing whether the experimental results support a particular theory or hypothesis, then test is said to be χ^2 - test for goodness of fit which describes the magnitude of the discrepancy between experimental values (observed values) and the theoretical values (expected values) under some theory or hypothesis.

The main objetive is to determine whether the distribution agrees with or "fits" some claimed distribution. If the observed values are close to the expected values under a hypothesis, the fit is said to be good. If, however, the difference between the two set of figures are found to be significant, the fit is not good.

Hence, χ^2-test for goodness of fit is used to test whether there is a significant difference between an observed frequency distribution and a theoretical (expected) frequency distribution. Because we test for how well an observed frequency distribution fits some specified theoritical distribution, this method is often called a *goodness-of-fit test*.

Assumptions

1. The data have been randomly selected.
2. The sample data consist of frequency counts for each of the different categories.
3. For each category, the expected frequency is at least 5. (The expected frequency for a category is the frequency that would occur if the data actually have the distribution that is being claimed. There is no requirement that the observed frequency for each category must be at least 5.)

Definition: A *goodness-of-fit test* is used to test the hypothesis that an observed frequeny distribution fits (or conforms to) some claimed distribution.

Notation:

Test statistic for *Goodness-of-Fit Test*:

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

where O = *Observed frequency* of an outcome.

E = *Expected frequency* (or *theoretical frequency*) of an outcome.

n = *Total number of trials.*

Critial values

1. Critical values are found in table by using $n-1$ degrees of freedom.
2. *Goodness-of-fit* hypothesis tests are always *right -tailed.*

9.5.1 The steps for testing χ^2-test for goodness-of-fit

The steps for testing χ^2-test for *goodncs-of-fit* are as follows:

Step 1. ***Null hypothesis:*** H_0 : There is no significant difference between observed (experimental) and the expected (theoretical) frequencies. In other words, the given data supports the theory (hypothesis).

Alternative hypothesis: H_1 : There is significant difference between observed and the expected frequencies. In other words, the given data does not support the theory.

Step 2. ***Test statistic:*** Under H_0 : The test statistic is

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

It follows χ^2 - distribution with $(n-1)$ degree of freedom.

Where,

O = *Observed frequency* of an outcome.

E = *Expected frequency* (or *theoretical frequency*) of an outcome which can be obtained under some theory or hypothesis.

Step 3. ***Critical value:*** Find the tabulated values (critical values) of χ^2 for $(n-1)$ $d.f$ at α level of significance (usually 5% or 1%)

Step 4. ***Decision:*** Make a decision by comparing the calculated value of χ^2 with tabulated values of χ^2.

If calculated $\chi^2 \le$ the tabulated χ^2, it is not significant and H_0 is accepted. Otherwise, it is rejected.

Remarks:

If $\chi^2 = 0$, we should be careful to question whether absolutely no difference exists between observed and expected frequencies.

Degree of Freedom ($d.f.$)

The number of values that can be chosen freely is called *degree of freedom*. It is denoted by ν (nu $d.f.$)

Thus, $d.f = n - 1$ (In general)

But, if we are needed to calculate the parameters of a theoretical distribution on the basis of given frequency distributed for obtaining the expectd (theoretical) frequencies of the distribution, then we subtract 1 $d.f.$ for each parameter estimated. If we are required to fit *Binomial* and *Poisson* distributions to the given set data. We lose 1 $d.f.$ for applying χ^2 - test, since for binomial distribution, we have estimated

only one parameter p, n being the given in the data and for the Poisson distribution we have estimated only one parameter λ.

In Binomial distribution, if the hypothetical value of p is given, then we do not lose any $d.f.$ for applying χ^2 - test.

However, if we required to fit the Normal distribution we lose 2 $d.f.$ for applying χ^2 because of calculating two parameters μ and σ^2 from the given data.

Thus, the degrees of freedom for the χ^2 - test of goodness of fit are given by

$d.f. = n - 1 - k_1 - k_2$, where

i) 1 $d.f.$ lost due to linear constraint $\Sigma O = \Sigma E = N$

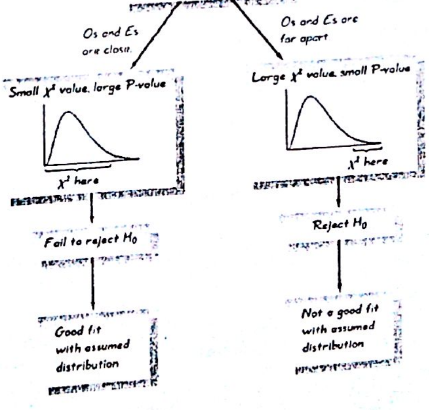
ii) k_1 is the number of parameter computed (if any) from the given data for estimating the expected frequency of the distribution.

iii) k_2 is the number of $d.f.$ lost due to polling expected frequencies which are less than 5.

9.5.2 Relationships Among the χ^2 Test Statistic, P-Value, and Goodness-of-Fit

The form of the χ^2 test statistic is such that close agreement between observed and expected values will lead to a small value of χ^2 and a large P-value. A large discrepancy between observed and expected values will lead to a large value of χ^2 and a small P-value. The hypothesis tests of this section are therefore always right-tailed, because the critical value and critical region are located at the extreme right of the distribution. These relationships are summarized and illustrated in Figure.

Once we know how to find the value of the test statistic and the critical value, we can test hypotheses by using the procedure introduced in previous Chapter.



Example 21: The number of automobile accidents per week in a certain community were as follows: 12, 8, 20, 2, 14, 10, 15, 6, 9, 4.

Are these frequencies in agreement with the belief that accident conditions were the same during the 10 week periods under considerations? [T.U. 2041 MBA/BIE 2068 Jestha]

Solution:

Step 1. *Null hypothesis H_0*: The given frequencies are consistent with the belief that accident conditions were the same during the 10 week period.

Alternative hypothesis H_1: The given frequencies are not consistent with the belief that accident conditions were the same during the 10 week period.

Step 2. *Test statistic*: Under H_0, the test statistic is $\chi^2 = \Sigma \frac{(O-E)^2}{E}$

where O = observed frequency

E = expected frequency = $\frac{\Sigma O}{n}$, since the number of accident expected each week would be the same.

Calculation of χ^2

O	$E = \frac{\Sigma O}{n}$	O – E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
12	10	2	4	0.4
8	10	–2	4	0.4
20	10	10	100	10
2	10	–8	64	6.4
14	10	4	16	1.6
10	10	0	0	0
15	10	5	25	2.5
6	10	–4	16	1.6
9	10	–1	1	0.1
4	10	–6	36	3.6
ΣO = 100	ΣE = 100			$\Sigma\frac{(O-E)^2}{E}$ = 26.6

Calculated $\chi^2 = \Sigma\frac{(O-E)^2}{E} = 26.6$. Degree of freedom $= n - 1 = 10 - 1 = 9$.

Since $\alpha = 5\%$, the tabulated χ^2 at 5% level if significance for 10 *d.f.* is 16.5.
That is, $\chi_{1.05,10} = 16.9$

Step 3. *Conclusion*: Since the calculated χ^2 is greater than the tabulated χ^2, it is significant and H_0 is rejected and hence H_1 is accepted which means that accident conditions were not uniform over the 10 weeks.

Example 22: In a set of random numbers, the digits 0, 1, 2, . . . , 9 were found to have the following frequencies [T.U. 2050 MBA]

Digits:	0	1	2	3	4	5	6	7	8	9
Frequency	43	32	38	27	38	52	36	31	39	24

Solution:

Step 1. *Null hypothesis H_0*: The digits are not significantly different from those expected on the hypothesis of unit form distributions. In other words, hypothesis of uniform distribution of digits holds good hypothesis.

Step 2. *Alternative hypothesis H_1*: The digits are significantly different from those expected on the hypothesis of uniform distribution. In other words, hypothesis of uniform distribution of digits does not hold good.

Test statistic: Under H_0, the test statistic is $\chi^2 = \Sigma\frac{(O-E)^2}{E}$

where, O = observed frequency, E = expected frequency = $\frac{\Sigma O}{n}$

Calculation of χ^2

Digits	O	$E = \frac{\Sigma O}{n}$	O - E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
0	43	36	7	49	1.36
1	32	36	–4	16	0.44
2	38	36	2	4	0.11
3	27	36	–9	81	2.25
4	38	36	2	4	0.11
5	52	36	16	256	7.11
6	36	36	0	0	0
7	31	36	–5	25	0.69
8	39	36	3	9	0.25
9	24	36	–12	144	4
	ΣO = 360				$\Sigma\frac{(O-E)^2}{E}$ = 16.32

Calculated $\chi^2 = \Sigma\frac{(O-E)^2}{E} = 16.32$

Degree of freedom $= n - 1 = 10 - 1 = 9$
The tabulated value of χ^2 at 5% level of significance for 9 *d.f.* is 16.32
That is $\chi^2_{0.05,9} = 16.92$

Step 3. Conclusion: Since calculated value of χ^2 is less than the tabulated value of χ^2, it is not significant and H_0 is accepted which means that hypothesis of uniform distribution of digits holds good.

Example 23: Among 64 offspring's of a certain cross between guinea pigs, 34 were red, 10 were black and 20 were white. According to the genetic model, these numbers should be in the ratio 9 : 3 : 4. Are the data consistent with the model at 5 percent level? [T.U. 2045 MBA]

Solution:

Step 1. ***Null hypothesis H_0***: The data are consistent with the model. In other words, the data supports the theory.

Step 2. ***Alternative hypothesis H_1***: The data are not consistent with the model. In other words, the data does not support the theory.

Step 3. Test statistic: Under H_0, the test statistic is $\chi^2 = \Sigma\frac{(O-E)^2}{E}$

where O = observed frequency E = expected frequency which is calculated on the basis of given proportions 9 : 3 : 4

Calculations of χ^2

Pigs	*O*	*E*	O – E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
Red	34	$\frac{9}{16} \times 64 = 36$	–2	4	0.11
Black	10	$\frac{3}{16} \times 64 = 12$	–2	4	0.33
White	20	$\frac{4}{16} \times 64 = 16$	4	16	1
	ΣO = 64	ΣE = 64			$\Sigma\frac{(O-E)^2}{E}$ = 1.44

Calculated $\chi^2 = \Sigma\frac{(O-E)^2}{E} = 1.44$

Degree of freedom = $n - 1 = 3 - 1 = 2$

Tabulated $\chi^2_{0.05, 2} = 5.991$

Step 4. **Decisions:** Since calculate χ^2 is less than the tabulated χ^2, it is not significant and H_0 is accepted which means that the data are consistent with the genetic model.

9.6 χ^2- test for independence of attributes

In previous section, we considered categorical data summarized with frequency counts listed in a single row or column. Because the cells of the single row or column correspond to categories of a single variable (such as color), the tables in previous section are sometimes called ***one-way frequency tables***. In this section, we again consider categorical data summarized with frequency counts, but the cells correspond to two different variables. The tables we consider in this section are called ***contingency tables***, or ***two-way frequency tables***.

χ^2- test for independent of attributes (characteristics) is applied to test whether two or more attributes are associated or not i.e., whether the attributes are related or independent. For example, we may want to test whether there is any association between marriage and failure or association between gender and habits of smoking etc. For this, information can be summarized and presented in two ways classification table which is also called ***contingency table***.

Definitions

A ***contingency table*** (or ***two-way frequency table***) is a table in which frequencies correspond to two variables. (One variable is used to categorize rows, and a second variable is used to categorize columns.)

The frequency of occurrence of *successes* and *failures* presented in contingency table having r rows and c columns is called ***r × c contingency table.***

Following table which summarizes the fate of the passengers and crew when the Titanic sank on Monday, April 15, 1912, has two variables: a row variable, which indicates whether the person survived or died; and a column variable, which lists the demographic categories—men, women, boys, girls.

Titanic Mortality					
	Men	*Women*	*Boys*	*Girls*	*Total*
Survived	332	318	29	27	706
Died	13 60	104	35	18	1517
Total	1692	422	64	45	2223

Contingency tables are especially important because they are often used to analyze survey results. For example, we might ask subjects one question in which they identify their gender (male female), and we might ask another question in which they describe the frequency of their use of TV remote controls (often/ sometimes / never). The methods of this section can then be used to determine whether the use of TV remote controls is independent of gender. (We probably already know the answer to that one.) Applications of this type are very numerous, so the methods presented in this section are among those most often used.

Now we consider tests of independence, used to determine whether a contingency table's row variable is independent of its column variable.

Definition

A *test of independence tests* the null hypothesis that there is no association between the row variable and the column variable in a contingency table. (For the null hypothesis, we will use the statement that "the row and column variables are independent.")

It is very important to recognize that in this context, the word contingency refers to dependence, but this is only a statistical dependence, and it cannot be used to establish a direct cause-and-effect link between the two variables in question. For example, after analyzing the data in above table, we might conclude that whether a person survived the sinking of the Titanic is dependent on whether that person was a man, woman, boy, or girl, but that doesn't mean that the gender age category has some direct causative effect on surviving.

When testing the null hypothesis of independence between the row and column variables in a contingency table, the assumptions, test statistic, and critical values are described in the following box.

Assumptions:

1. The sample data are randomly selected.
2. The null hypothesis H_0 is the statement that the row and column variables are independent; the alternative hypothesis H_1 is the statement that the row and column variables are dependent.
3. For every cell in the contingency table, the expected frequency E is at least 5. (There is no requirement that every observed frequency must be at least 5. Also, there is no requirement that the population must have a normal distribution or any other specific distribution.)

Test Statistic for a Test of Independence : $\chi^2 = \sum\frac{(O-E)^2}{E}$

Critical values:

1. The critical values are found in Table by using

degrees of freedom = $(r-1)(c-1)$

where r is the number of rows and c is the number of columns.

2. In a test of independence with a contingency table, the critical region is located in the right tail only.

9.6.1 The procedure for testing the independence of two attributes presented in $r \times c$ contingency table

Step 1. Setting up hypotheses:

Null hypothesis H_0: two attributes (two categorical variables) are independent *i.e.*, there is no relationship (association) between them.

Alternative hypothesis H_1: two attributes are dependent *i.e.*, there is relationship association between them.

Step 2. *Test statistic*: Under H_0, the test statistic is

$$\chi^2 = \sum\frac{(O-E)^2}{E}$$

where O = observed frequency, E = expected frequency.

$E = \frac{RT \times CT}{N}$, RT = Row total, CT = Column total

N = Total sample size = grand Total

Special case of $r \times c$ contingency table: 2×2 contingency table

			RT
	a	b	a + b
	c	d	c + d
CT	a + c	b + d	N = a + b + c + d

The expected frequency for each cell can be obtained as:

$$E(a) = \frac{RT \times CT}{N} = \frac{(a+b)\,(a+c)}{N},\ E(C) = \frac{(c+d)\,(a+c)}{N}$$

$$E(b) = \frac{(a+b)\,(b+d)}{N},\ E(d) = \frac{(c+d)\,(b+d)}{N}$$

For 2 × 2 contingency table the value of χ^2 is directly calculated by using formula $\chi^2 = \frac{N(ad-bc)^2}{(a+b)\,(c+d)\,(a+c)\,(b+d)}$

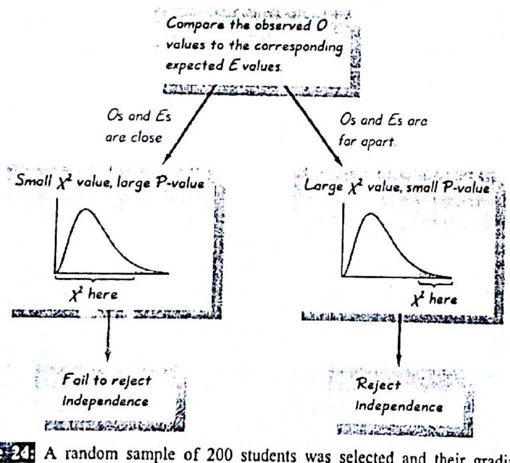
Step 3. ***Degree of freedom*** (*d.f.*): $\nu = (r-1) \times (c-1)$, where $n = r \times c$ = total number of frequencies.

Step 4. ***Critical value***: obtain the critical (tabulated value) of χ^2 for given level of significance for $(r-1)\,(c-1)$ *d.f.*

Step 5. ***Decision***: If calculated value of χ^2 is less than or is equal to tabulate value of χ^2, it is not significant and H_0 is accepted. Otherwise, it is rejected.

9.6.2 Relationship among Key Components in Test of Independence

As in previous Section, if observed and expected frequencies are close, the χ^2 test statistic will be small and the *P*-value will be large. If observed and expected frequencies are far apart, the χ^2 test statistic will be large and the *P*-value will be small. These relationships are summarized and illustrated in Figure



Example 24: A random sample of 200 students was selected and their grading ability in mathematics and interest in business administration were as given below. [TU 2051 / 2065]

Interest in Business Administration	Ability in Mathematics			
	Low	Average	High	Total
Low	60	15	15	90
Average	15	45	10	70
High	5	10	25	40
Total	80	70	50	200

Test whether there is any relationship between student interest in business administration and ability in mathematics.

Solution:

Step 1. ***Null hypothesis***: H_0: there exists no relationship between student interest in business administration and ability in mathematics.
Alternative hypothesis H_1: there exists relationship between student interest in business administration and ability in mathematics.

Step 2. ***Test statistic***: Under H_0, test statistic is $\chi^2 = \sum \frac{(O-E)^2}{E}$

where $E = \frac{RT \times CT}{N} = \frac{\text{Row total} \times \text{column total}}{\text{whole total}}$

Calculation of χ^2

O	E	O - E	$(O-E)^2$	$(O-E)^2/E$
60	(90 × 80) / 200 = 36	24	576	16
15	(90 × 70) / 200 = 31.5	−16.5	272.25	8.643
15	(90 × 50) / 200 = 22.5	−7.5	56.25	2.5
15	(90 × 80) / 200 = 28	−13	169	6.04
45	(70 × 70) / 200 = 24.5	20.5	420.25	17.153
10	(70 × 50) / 200 = 17.5	−7.5	56.25	3.214
5	(40 × 80) / 200 = 16	−11	121	7.5625
10	(40 × 70) / 200 = 14	−4	16	1.143
25	(50 × 40) / 200 = 10	15	225	22.5

∴ $\chi^2 = 84.7555 = 84.76$

Step 4. ***Level of significance*** (α): Take $\alpha = 5\% = 0.05$

Step 5. ***Degree of freedom*** $(d.f.) = (r-1) \times (c-1) = 2 \times 2 = 4$

Step 6. ***Critical value***: $\chi^2_\alpha(\nu) = \chi^2_{0.05}(4) = 9.488$

Step 7. ***Decision***: Since calculated value of $\chi^2 = 84.76$ > tabulated value of χ^2 9.88, so, H_0 is rejected. Hence we conclude that there exists relationship between student interest in business administration and ability in mathematics.

Example 25: Two groups of 100 people each were taken for testing the use of vaccine 15 persons contracted the disease out of the inoculated persons, while contracted the disease in the other group. Test the efficiency of the vaccine using value. At 5% level for one degree of freedom, the value of $\chi^2 = 3.84$.

Solution: The above data can be arranged in the form 2 × 2 contingency table as follows.

	Attack	Not attack	Total (RT)
Inoculated	15	10−15 =85	100
Not inoculated	25	100−25=75	100
Total (CT)	40	160	N = 200

Step 1. ***Null hypothesis*** H_0: The vaccine is not effective in curing the disease.
Alternative hypothesis H_1: The vaccine is effective in curing the disease

Step 2. ***Test statistic***: Under H_0 the test statistic is

$\chi^2 = \sum \frac{(O-E)^2}{E}$, where $E = \frac{RT \times CT}{N}$

Calculation of χ^2

O	E	$O-E$	$(O-E)^2$	$(O-E)^2/E$
15	(100×40) /200 = 20	−5	25	1.25
85	(100×160) /200 = 80	5	25	0.3125
25	(100×40) /200 = 20	5	25	1.25
75	(100×160) /200 = 80	−5	25	0.3125
200				3.125

Calculated $\chi^2 = \sum \frac{(O-E)^2}{E} = 3.125$

Step 3. ***Degree of freedom*** $(d.f.) = (r-1) \times (c-1) = (2-1)(2-1) = 1$

Step 4. ***Level of significance*** $(\alpha) = 5\%$

Step 5. ***Critical value*** Tabulated value $\chi^2_{0.05}(1) = 3.84$

Step 6. *Decision*: Since calculated χ^2 < tabulated χ^2, it is not significant and H_0 is accepted which means that the vaccine is not effective.

Example 26: *(The chi square test of independence).* To determine whether there really is a relationship between an employee's performance in the company's training program and his or her ultimate success in the job. the company takes a sample of 400 cases from its very extensive files and obtains he results shown in the following table:

		Performance in training program			
		Below average	*Average*	*Above average*	*Total*
Success in job (employer's rating)	*Poor*	23	60	29	112
	Average	28	79	60	167
	very good	9	49	63	121
	Total	60	188	152	400

Use the 0.01 level of significance to test the null hypothesis that performance in the training program and success in the job are independent

Solution:

Step 1. ***Null hypothesis***: Performance in training program and success in job are independent.

Alternative hypothesis: Performance in training program and success in job are dependent.

Step 2. ***Level of significance***: $\alpha = 0.01$

Step 3. ***Criterion***: reject the null hypothesis if $\chi^2 > 13.277$, the value of $\chi^2_{0.01}$ for $(3-1)(3-1) = 4$ degrees of freedom, where χ^2 is given by the formula

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Step 4. ***Calculation***: Calculating first the expected cell frequencies for the first two cells of the first two rows, we get

$$E(23) = e_{11} = \frac{(112)(60)}{400} = 16.80,\ E(60) = e_{12} = \frac{(112)(188)}{400} = 52.64$$

$$E(28) = e_{21} = \frac{(167)(60)}{400} = 25.05,\ E(79) = e_{22} = \frac{(167)(188)}{400} = 78.49$$

Then, by subtraction, we find that the expected frequencies for the third cell of the first two rows are 42.56 and 63.46, and those for the third row are 18.15, 56.87 and 45.98. Thus,

$$\chi^2 = \frac{(23-16.80)^2}{16.80} + \frac{(60-52.64)^2}{52.64} + \frac{(29-42.56)^2}{42.56}$$
$$+ \frac{(28-25.05)^2}{25.05} + \frac{(79-78.49)^2}{78.49} + \frac{(60-63.46)^2}{63.46}$$
$$+ \frac{(9-18.15)^2}{18.15} + \frac{(49-56.87)^2}{56.87} + \frac{(63-45.98)^2}{45.98} = 20.179$$

Step 5. *Decision*: Since $\chi^2 = 20.179$ exceeds 13.277, the null hypothesis must be rejected; we conclude that there is a dependence between an employee's performance in the training program and his or her success in the job.

Note: We pursue this example further in order to determine the form of the dependence.

Example 27: *(Testing the equality of three proportions using the χ^2 statistic)* Samples of three kinds of materials, subjected to extreme temperature changes, produced the results shown in the following table:

	Material A	*Material B*	*Material C*	*Total*
Crumbled	41	27	22	90
Remained intact	79	53	78	210
Total	120	80	100	300

Use the 0.05 level of significance to test whether, under the stated conditions, the probability of crumbling is the same for the three kinds of materials.

Solution:

Step 1. *Null hypothesis*: $p_1 = p_2 = p_3$

Alternative hypothesis: p_1, p_2 and p_3 are not all equal.

Step 2. *Level of significance*: $\alpha = 0.05$

Step 3. *Criterion*: Reject the null hypothesis if $\chi^2 > 5.991$, the value of $\chi^2_{0.05}$ for $3 - 1 = 2$ degrees of freedom, where χ^2 is given by the formula

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Step 4. *Calculations*: The expected frequencies for the first two cells of the first row are $e_{11} = \frac{90 \times 120}{300} = 36$ and $e_{12} = \frac{90 \times 80}{300} = 24$

and, as it can be shown that the sum of expected frequencies for any row or column equals that corresponding observed frequencies, we find by subtraction that $e_{13} = 90 - (36 + 24) = 30$, and that the expected frequencies for the second row are $e_{21} = 120 - 36 = 84$, $e_{22} = 80 - 24 = 56$, and $e_{23} = 100 - 30 = 70$. Then substituting these values together with the observed frequencies into the formula for χ^2, we get

$$\chi^2 = \frac{(41-36)^2}{36} + \frac{(27-24)^2}{24} + \frac{(22-30)^2}{30} + \frac{(79-84)^2}{84} + \frac{(53-56)^2}{56} + \frac{(78-70)^2}{70}$$
$$= 4.575$$

Step 5. *Decision*: Since $\chi^2 = 4.575$ does not exceed 5.991, the null hypothesis cannot be rejected; in other words, the data do not refute the hypothesis that, under the stated conditions, the probability of crumbling is the same for the three kinds of material.

Example 28: *(Exploring the form of dependence).*

With reference to the preceding example, find the individual contributions to the chi square statistic.

Solution: We display the contingency table, but this time we conclude the expected frequencies just below the observed frequencies.

Performance in training program

Success in job (employer's rating)	Below average	Average	Above average	Total
Poor	23 16.80	60 52.64	29 42.56	112
Average	28 25.05	79 78.49	60 63.46	167
Very good	9 18.15	49 56.87	63 45.98	121
Total	60	188	152	400

Also, we write the χ^2 statistic as the sum of the contributions.

$$\chi^2 = 2.288 + 1.029 + 4.320 + 0.347 + 0.003 + 0.189 + 4.613 + 1.089 + 6.300 = 20.179$$

From the two displays, it is clear that here is a positive dependence between performance in training and job success. For the three individual cells with the largest contributions to χ^2, the ***above average-very good*** cell frequency is high, whereas the ***above average-poor*** and ***below average- very good*** cell frequencies are low.

Example 29: (*Titanic Sinking*): Refer to the Titanic mortality data in Table. We will treat the 2223 people aboard the Titanic as a sample. We could take the position that the Titanic data constitute a population and therefore should not be treated as a sample, so that methods of inferential statistics do not apply. Let's stipulate that the data are sample data randomly selected from the population of all theoretical people who would find themselves in the same conditions. Realistically, no other people will actually find themselves in the same conditions, but we will make that assumption for the purposes of this discussion and analysis. We can then determine whether the observed differences have statistical significance. Using a 0.05 significance level, test the claim that when the Titanic sank, whether someone survived or died is independent of whether the person is a man, woman, boy, or girl.

Table: Titanic Mortality

	Men	Women	Boys	Girls	Total
Survived	332	318	29	27	706
Died	1360	104	35	18	1517
Total	1692	422	64	45	2223

Solution:

Step 1: The null hypothesis H_0 and alternative hypothesis H_1 are as follows:

H_0: Whether a person survived is independent of whether the person is a man, woman, boy, or girl.

H_1: Surviving the *Titanic* sinking and being a man, woman, boy, or girl are dependent.

Step 2: The significance level is $\alpha = 0.05$. Because the data are in the form of a contingency table, we use the χ^2 distribution with this test statistic:

$$\chi^2 = \sum\frac{(O-E)^2}{E} = \frac{(332-537.360)^2}{537.360} + \frac{(318-134.022)^2}{134.022} + \frac{(29-20.326)^2}{20.326} + \frac{(27-14.291)^2}{14.291} + \frac{(1360-1154.640)^2}{1154.640} + \frac{(104-287.978)^2}{287.978} + \frac{(35-43.674)^2}{43.674} + \frac{(18-30.709)^2}{30.709}$$

$$= 78.481 + 252.555 + 3.702 + 11.302 + 36.525 + 117.536 + 1.723 + 5.260$$
$$= 507.084$$

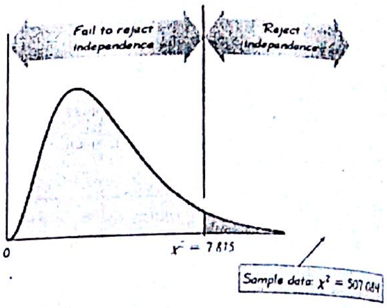


Figure: Test of Independence for the Titanic Mortality Data.

The critical value is $\chi^2 = 7.815$ and it is found from Table A-5 by noting that $\alpha = 0.05$ in the right tail and the number of degrees of freedom is given by $(r-1)(c-1) = (2-1)(4-1) = 3$. The test statistic and critical value are shown in Figure. Because the test statistic falls within the critical region, we reject the null hypothesis that whether a person survived is independent of whether the person is a man, woman, boy, or girl. It appears that whether a person survived the Titanic and whether that person is a man, woman, boy, or girl are dependent variables.

Exercise 9

Theoretical Question

Proportion

1. Write down the steps for testing hypothesis of population proportion for large sample size. **[TU BE 2068 Bhadra/2068 Magh(Back)]**
2. Describe the procedure of test of significance between two population proportions **[TU, BE, 2066 Magh/ 2064 Poush]**

Numerical problems

1. A sample of 600 persons selected at random from a large city gives the result that males are 53%. Is there reason to doubt the hypothesis that males and females are equal number in the city?
   [Ans : Z =] **[TU, BE, 2065 Chaitra (Re)**
2. Microsoft estimated last year that 35% of potential software buyers the new planning to wait to purchase the new operating system, window panes, until an upgrade has been released. After an advertising campaign to reassure to public Microsoft surveyed 3,000 people and found 950 who were still skeptical people has been decreased ? **[TU, BE, 2063 Ashadh]**
3. A manufacture claims that at least 95% of the equipments which he supplied to a factory conformed to specification. An examination of a sample of 200 pieces of equipment revealed that 18 were faulty. Test the claim at 5% and 1% level of significance.

4. An airline claims that only 6% of all lost luggage's is never found. If, in a random sample, 17 out of 200 pieces of lost luggage are not found, test the null hypothesis $P = 0.06$ against the alternative hypothesis $P > 0.06$ at the 0.05 level of significance.
[Ans : $Z = 1.489$; Accept H_0]

5. Suppose that 4 out of 13 undergraduate engineering's students state that they will go on to graduated school. Test the dean's claim that 60% of the undergraduate students will go on to graduated school, using the alternative hypothesis $P > 0.60$ and the level of significance $\alpha = 0.05$.
[Hint: Use binomial table to determine the probability of getting "at most 4 successes in 13 trials" when $p = 0.60$]

6. Random sample 400 men and 600 women were asked whether they would like certain TV program, 200 men and 325 women were in favour of the proposal. It is claimed that the proportion of men and women in favour of the proposal are same. Can we support the claim at 5% level of significance? .
[TU, BE 2065 chaitra] [Reject H_0]

χ^2 Test

Theoretical Question

1. Write application of chi-square distribution.

Numerical problems

Goodness of fit

1. The following table lists the frequency distribution of cars sold at an auto dealership during the past 10 months.

Months	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*
Cars sold	23	17	15	10	14	12	13	15	26	25

Using the 5% significance level, will you conclude that the number of cars sold at this dealership is the same for each month?
[Ans: $\chi^2_{0.05,9} = 16.92$, H_0 is Rejected]

2. The following figures show the distributor of digits in number chosen at random from a telephone directory.

Digits	0	1	2	3	4	5	6	7	8	9
Frequency	925	875	965	900	935	950	875	800	875	900

Test the hypothesis that the digits are distributed randomly through a telephone directory.
[Ans: $\chi^2_{0.05,9} = 16.92$, H_0 is Rejected]

Independence of attributes

1. From the following data can you conclude that there is association between the purchase of brand and geographical region?

	Region		
	Central	*Eastern*	*Western*
Purchase brand	40	55	45
Do not purchase brand	60	45	55

Use 5% level of significance. [$\chi^2 = 4.687$, H_o is accepted] *[TU 2061]*

2. Do the following data provide evidence of the effectiveness of inoculation in preventing tuberculosis: *[TU 2046]*

	Attacked	*Not Attacked*	*Total*
Inoculated	20	300	320
Not inoculated	80	600	680
Total	100	900	1000

[Ans: $\chi^2 = 7.36$ It is significant and H_o is rejected]

3. A sample of 500 workers of a factory according to gender and nature of work is as follows:

Nature of work	*Gender*		*Total*
	Male	*Female*	
Technical	200	100	300
Non-technical	50	150	200
Total	250	250	500

Test at 5% level of significance whether exists an relationship between gender and nature of work. [Ans: $\chi^2 = 83.33$, H_o is rejected] *[TU 2057]*

4. Four hundred employees of a factory are classified according to their level and decisions. Do you agree with the statement that decisions very according to level of employee? [TU 2057]

Decision	*Senior officers*	*Officer*	*Junior officer*	*Total*
Quick	60	80	70	210
Slow	40	60	90	190
Total	100	140	160	400

[Ans: χ^2 = 8.377; Not significant. accept H_o]

5. Test of the fidelity and selectivity of 190 digital radio receivers produced the results shown in the following table:

Selectivity		*Fidelity*		
		Low	*Average*	*High*
	Low	6	12	32
	Average	33	61	18
	High	13	15	0

Use the 0.01 level of significance to test whether there is relationship (dependence) between fidelity and selectivity. [Ans: $\chi^2 = 54.328$, reject H_0]

Additional Exercise

THE SAMPLING DISTRIBUTION OF THE SAMPLE PROPORTION

Confidence Intervals + Hypothesis Testing

Confidence Intervals

1. In a survey carried out in a large city, 150 households out of a random sample of 250 owned at least one pet. Find approximate 95% confidence limits for the proportion of households within the city who own at least one pet.

2. In a market research survey, 30 people out of a random sample of 100 people said that they used a particular brand of washing powder. Find 98% confidence limits for the proportion of people in the area who use this brand of washing powder.

3. A special edition commemorative coin is tossed 500 times, each toss being independent of each other, and on 300 occasions the coin lands face up. Find a 95% confidence interval for the probability that the coin lands face up on any single toss. Explain briefly how your confidence interval suggests that the coin is biased.

4. The manager of a large supermarket wishes to judge the effect of a new layout on the customers. On the day that the layout was introduced the first 200 customers in the store were asked whether or not they approved of the new layout. Comment on the manner in which the sample was chosen, and suggest a way of obtaining a more suitable sample. Out of a suitably chosen sample of 200 customers, 148 approved of the new layout. Calculate an approximate 95% confidence interval for the population percentage of customers who approve of the new layout.

5. In a random sample of 40 adult males, 35 said that they contributed to a personal pension plan.
   i) Calculate an approximate 95% confidence interval for the population proportion, p, of adult males who contribute to a personal pension plan.
   ii) Estimate the size of sample required to produce a 95% confidence interval of width 0·1 for p.
6. Of a random sample of 400 adults, 250 said that they would vote for the *Manic* party at the next election.
   i) Calculate an approximate 98% confidence interval for the population proportion, p, of adults who would vote for the *Manic* party at the next election.
   ii) Estimate the size of sample required to produce a 98% confidence interval of width 0·15 for p.
   iii) It later emerged that the 400 adults who took part in the survey were contacted by telephone. Explain briefly why this could lead to bias within the survey.
7. In a postal survey of 500 households, 330 said that they thought they were being overcharged for the public services within their area.
   i) Calculate an approximate 99% confidence interval for the population proportion, p, of households who thought they were being overcharged for public services within their area.
   ii) ii) Estimate the size of sample required to estimate the value of p to within 99% confidence limits of ± 0 025.

Hypothesis testing: large samples.

8. It is thought that the proportion of defective items produced by a particular machine is 0·1. A random sample of 100 items is inspected and found to contain 16 defective items. Does this provide evidence, at the 5% level, that the machine is producing more defective items than expected?
9. A coin is tossed 200 times and 115 heads obtained. Is there evidence, at the 1% level, that the coin is biased towards heads?
10. In an investigation into ownership of mobile phones, 100 randomly chosen people were interviewed and 15 found to own a mobile phone. Using the evidence of this sample, test, at the 10% level of significance, the hypothesis that the proportion of people owning a mobile phone is 20% against the alternative hypothesis that the proportion is less than 20%. {Hint: continuity correction; represent 15 by 15·5 to give the null hypothesis the greatest chance of being accepted!}
11. At a previous election, 30% of the electorate voted for the MHT party. Prior to the next election a telephone survey was conducted to attempt to predict the forthcoming result. 330 out of 1000 people telephoned said that they intended to vote for the MHT party.
   (i) The number of people who intend to vote for the MHT party is modeled by a binomial distribution. Use a normal approximation to this binomial distribution to carry out a hypothesis test at the 5% significance level to test the claim that support for the MHT party has changed. Give suitable null and alternative hypotheses and state your conclusion clearly.
   (ii) State briefly why this survey might not give an accurate prediction of the forthcoming result.
12. In order to test whether a particular coin is biased towards heads, it is tossed 150 times and the number of heads obtained, X, counted.
   (i) State suitable null and alternative hypotheses, involving a probability, for significance test at the 5% level.
   (ii) For the test described in (i), find the probability of making a Type *II* error if the probability of obtaining a head on the coin is in fact 0·55.
13. Previous experience suggests that support for the *Radical* party is 40% amongst the local electorate. A random sample of 200 voters is selected to try and determine if this support has diminished at all.
   (i) State suitable null and alternative hypotheses, involving a probability, for

significance test at the 2% level.
   (ii) For the test described in (i), find the probability of making a Type *II* error if the actual proportion of voters supporting the *Radical* party is 0.32.
14. The process of manufacturing a certain kind of dinner plate results in a proportion 0.15 of faulty plates. An alteration is made to the process which is intended to reduce the proportion of faulty plates.
   (i) State suitable null and alternative hypotheses for a statistical test of the effectiveness of the alteration.
   (ii) In order to carry out the test, the quality control department counts the number of faulty plates in a random sample of 2000 plates. If 260 or fewer faulty plates are found then it will be accepted that the alteration does result in a reduction in the proportion of faulty plates. Calculate the significance level of this test using a suitable normal approximation.
   (iii) Calculate the probability of making a Type *II* error in the above test, given that the alteration results in a decrease in the proportion of faulty plates to 0·13.
15. A die is suspected of bias towards showing more sixes than would be expected of an ordinary die. A test is devised to test the null hypothesis $p = 1/6$ where p is the probability of the die showing a six, against the alternative $p > 1/6$. The test involves throwing the die 120 times and rejecting the null hypothesis if the number of sixes obtained is m or more. Use a normal approximation to the binomial distribution to find the value of m for which the probability of making a Type I error is 0·01.

Hypothesis testing: small samples.

16. A coin is tossed 6 times and 5 heads are obtained. Test at the 5% level whether the coin is biased towards heads.
17. A die is thrown 9 times and it shows a six on 3 occasions. Is the die biased in favor of showing a six? Test at the 1% level.
18. The probability that a certain type of seed germinates is 0.4. The seeds undergo new kind of treatment, and when a packet of 10 seeds is tested, 8 germinate. Is there evidence, at the 5% level, of a change in the germination rate? {2 – t tailed test!}
19. In a test of 10 true–false questions, a student gets 8 correct. The student claims sh was not guessing.
   (i) State suitable null and alternative hypotheses, involving a probability, f significance test of this student's claim.
   (ii) Carry out the test, at the 5% level, stating your conclusion clearly.
   (iii) For the test described in (i), find the smallest level of the test which wou result in the null hypothesis being rejected.
20. A die is suspected of bias towards showing more sixes than would be expected of an ordinary die. In order to test this, it is decided to throw the die 12 times. The nu hypothesis $p = 1/6$ where p is the probability of the die showing a six, will b rejected in favor of the alternative hypothesis $p > 1/6$ if the number of sixe obtained is 4 or more. Calculate, to 3 decimal places, the probability of making
   (i) a Type I error, (ii) a Type II error if, in fact, $p = 1/2$
21. An enthusiastic gardener claimed that she could never work in the garden at th weekend because "It always rains on Saturday and Sunday when I'm at home an it's always fine on weekdays when I'm not!". She noted the weather for the ne month and recorded that, out of 10 wet days, 5 were either a Saturday or a Sunda The gardener's claim may be modeled by regarding her observation as a sing sample from a Bin(10, p) distribution. Given that one would expect two out o every seven wet days to be either a Saturday or a Sunday, the null hypothesis, p 2/7 may be tested against the alternative hypothesis, $p > 2/7$
   Carry out a hypothesis test to test her claim at the 5% significance level.

(ANSWERS) Where appropriate, answers are given to 3 significant figures.

1. $0.539 < p < 0.661$. 2) $0.193 < p < 0.407$. 3) $0.557 < p < 0.643$.
4. Sample could yield biased results as people might influence each other in bein interviewed together. Also the sample will likely contain friends and family etc. A

more appropriate sample could be obtained by interviewing, say, every 20th customer. $67.9\% < p < 80.1\%$.

5. (i) $0.772 < p < 0.977$, (ii) n needs to be at least 169
6. (i) $0.569 < p < 0.681$, (ii) n needs to be at least 226, (iii) the sample is biased against those who do not own a telephone.
7. (i) $0.605 < p < 0.715$, (ii) n needs to be at least 2383.
   {The answers to questions (8) - (15) assume that continuity corrections are made. }
8. Z test = 1.833. Reject H_0. there is sufficient evidence to suggest that the machine is producing too many defective etc.
9. Z test = 2·051. Accept H_0. there is insufficient evidence to suggest that the coin is biased towards heads etc.
10. Z test = −1.125. Accept H_0, there is insufficient evidence to suggest that the proportion is < 20% etc.
11. (i) H_0: $p = 0.3$, H_1: $p \neq 0.3$, Z test = 2.036. Reject H_0 etc. (ii) The telephone survey is biased as not all of the electorate will have a telephone.
12. (i) H_0: $p = 1/2$, H_1: $p > 1/2$, (ii) {Accept H_0 if $X \leq 85$ heads} P(Type II error) = 0·6586.
13. X = "the number of voters from the sample of 200 who support the *Radical* party".
   (i) H_0: $p = 0.4$, H_1: $p < 0.4$, (ii) {Accept H_0 if $X \geq 66$}, (ii) P(Type II error) = 0.4102.
14. (i) H_0: $p = 0.15$, H_1: $p < 0.15$, (ii) P(Type I error) = 0.0067 ⇒significance level = 0.67%, (iii) P(Type II error) = 0·4868.
15. m = 30 sixes.
16. Accept H_0. there is insufficient evidence to suggest that the coin is biased towards heads at the 5% level.
17. Accept H_0, there is insufficient evidence to suggest that the die is biased towards showing a six at the 1% level.
18. Reject H_0. there is sufficient evidence, at the 5% level, to suggest that the germination rate has changed.
19. (i) H_0: $p = 1/2$, H_1: $p > 1/2$, (ii) accept H_0. there is insufficient evidence, at the 5% level, to suggest that the student was doing anything other than guessing. (iii) 5.46875%.
20. (i) P(Type *I* error) = 0.125, (ii) P(Type *II* error) = 0.073.
21. Accept H_0. there is insufficient evidence to support the gardener's claim at the 10% level.