Statistics 2

CHAPTER 7
HYPOTHESIS TESTING

What is a hypothesis test? (pages 106 – 109)

- A **hypothesis** is a statement made about the value of a **population parameter** that we wish to test by collecting evidence in the form of a sample.

- In a statistical hypothesis test, the evidence comes from a sample which is summarized in the form of a statistic called the **test statistic.**

A hypothesis test consists of two hypotheses:

- The **null hypothesis**, denoted by H_0, is the hypothesis that we assume to be correct unless proved otherwise.
- The **alternative hypothesis**, denoted by H_1, tells us about the value of the population parameter if our assumption is shown to be wrong.

- A **hypothesis test** is a mathematical procedure to examine a value of the population parameter proposed by the null hypothesis H_0, compared to the alternative hypothesis H_1.

We perform a hypothesis test to decide whether or not we should reject the null hypothesis in favour of the alternative.

The decision whether or not to reject the null hypothesis for an observed value has to be based upon the idea that some values of X are unlikely under the null hypothesis and would be better explained by the alternative hypothesis.

Statisticians generally regard a probability of 5% as being unlikely and a probability of 1% as being very unlikely. This threshold probability is called the **level of significance**, a, the probability of rejecting H_0. We can perform the test at any significance level (usually 1%, 5% or 10%). This value is given in the exercise.

We can divide the x values into two regions.

1. The one that contains the values that **collectively** have a small chance of happening under the null hypothesis. We call this region the **critical region**.
2. The rest.

- The **critical region** is the range of values of the test statistic that would lead to you rejecting H_0.

- The value(s) on the boundary of the critical region are called **critical value(s).**

In S2, we will only examine hypothesis tests for the proportion, p, of a binomial distribution and hypothesis tests for the mean, λ, of a Poisson distribution.

- If the observed value of our test statistic (p or λ), lies in the critical region we have sufficient evidence to reject the null hypothesis H_0.
- If the observed value of our test statistic does not lie in the critical region, then we do not have sufficient evidence to reject the null hypothesis H_0.
  This does not mean that H_0 is true, rather that it remains a possibility and can be accepted at the moment, but later samples may cause it to be rejected.

There are two types of tests:

- A **one-tailed test** looks either for an increase in the value of a parameter or for a decrease in the value of a parameter, and has a single critical value.
- A **two-tailed test** looks for both an increase and a decrease in the value of a parameter, and has a twos critical values.

We can summarize hypothesis testing procedures as follows:

1. Identify the population parameter θ that you are going to test.
(In S2, this can be the proportion p for a binomial or the mean λ (or μ) for a Poisson)

2. Write down the null (H_0) and alternative (H_1) hypotheses. The alternative hypothesis will determine whether you want a one- or two-tailed test.

3. Specify the significance level α.

4. Use the following procedure to find out whether the observed value x of your test statistic falls in the critical region.

One-tailed tests

$H_0 : \theta = m \quad H_1 : \theta > m$ Reject H_0 if: $P(X \geq x) \leq a$

$H_0 : \theta = m \quad H_1 : \theta < m$ Reject H_0 if: $P(X \leq x) \leq a$

Two-tailed test

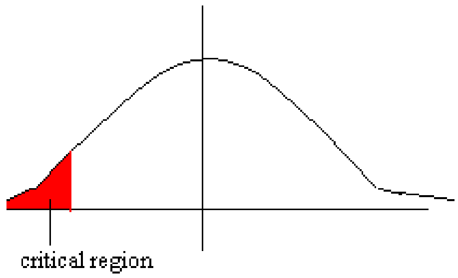
$H_0 : \theta = m \quad H_1 : \theta \neq m$ Reject H_0 if: $P(X \geq x) \leq \tfrac{1}{2}\alpha$ or $P(X \leq x) \leq \tfrac{1}{2}\alpha$

5. State your conclusion.
The following points should be addressed.
(a) Is the result **significant or not**?
(b) What are the implications in terms of the **context of the original problem**?

One-Tailed Test (pages 108 – 116)

We choose a critical region. In a one-tailed test, the critical region will have just one part (the shaded area below). If our sample value lies in this region, we reject the null hypothesis in favour of the alternative.

If we are looking for a definite decrease then, the critical region will be to the left.
If we are looking for a definite increase then, the critical region will be to the right.



<u>EXAMPLE 1:</u>

The random variable X follows a Poisson distribution with parameter 9. A sample observation gave a mean of 3. Carry out a hypothesis test on the mean of X, to test whether the mean has decreased. Use a 5% level of significance.

$$X \sim \text{Po}(9)$$

The hypotheses are:

$$H_0: \lambda = 9$$

$$H_1: \lambda < 9$$

We want to test if it is "reasonable" for the observed value of 3 to have come from a Poisson distribution with parameter 9. So what is the probability that a value as low as 3, has come from a Po(9)?

$$P(X \leq 3) = 0.0212$$

This has come from the Poisson tables under $\lambda = 9$

The probability 0.0212 is less than 0.05, so there is less than a 5% chance that the value has come from a Poisson distribution with mean 9.

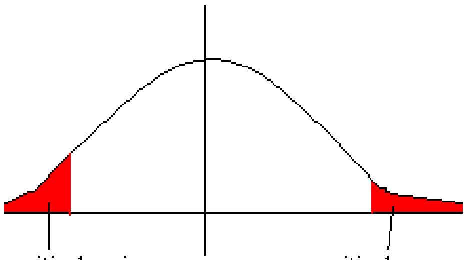
Therefore, there is enough evidence to reject the null hypothesis in favour of the alternative at the 5% level.

REMARK:

The probability is greater than 0.01. So, we would not reject the null hypothesis in favour of the alternative at the 1% level.

Two-Tailed Test (pages 108 – 116)

In a two-tailed test, we are looking for either an increase or a decrease. So, for example, H_0 might be that the mean is equal to 9 (as before). This time, however, H_1 would be that the mean is not equal to 9. In this case, therefore, the critical region has two parts:



<u>EXAMPLE 2:</u>

Adam claims that a specific 10 p coin is not fair. He tossed the coin 10 times and he got 7 heads. Test Adam's belief at a 10% significance level, on the basis of this sample.

- Tossing a coin has two possible outcomes: Heads or Tails.
- Assuming that the coin is fair, there is a 50% chance of getting heads and a 50% chance of getting tails. ($p = 0.5$)
- There is also a fixed number of trials (10 tosses).
- Each toss of a coin is independent of the previous one.

Therefore, let X = the number of heads obtained when tossing a coin 10 times.

$$X \sim B\,(10, 0.5)$$

We want to test the parameter p of this Binomial distribution at the 10% level.

If the coin is fair, $p = 0.5$. This is the null hypothesis. We want to test if p is other than 0.5. So, the alternative hypothesis will be $p \neq 0.5$. We have a two-tailed test.

$$H_0:\ p = 0.5$$
$$H_1:\ p \neq 0.5$$

Now, because the test is 2-tailed, the critical region has two parts. Half of the critical region is to the right and half is to the left. So the critical region contains both the top 5% of the distribution and the bottom 5% of the distribution (since we are testing at the 10% level).

If H_0 is true, $X \sim B(10, 0.5)$.

If the null hypothesis is true, what is the probability that X is 7 or above?

$P(X \geq 7) = 1 - P(X < 7) = 1 - P(X \leq 6) =$

$= 1 - 0.8281 = 0.1719$

Use the Binomial tables under $n = 10$ and $p = 0.5$

The probability 0.1719 is not less than 0.05. The probability that X is at least 7 is not less than 0.05 (5%).

So, at the 10% level, there is not significant evidence to reject the null hypothesis.

∗ Adam's belief is wrong.
There is not enough evidence that the coin is biased.

It is very important to make a comment relating your conclusion to the context of the question.

REMARK:

There are cases when we need to use an approximation (Poisson or Normal) in order to perform the calculations easily. See Chapter 5 and Section 2.5 of the book.

Critical Values (pages 116 – 117)

<u>EXAMPLE 3:</u>

In a particular city it was found, over a period of time that X, the number of cases of a certain medical condition reported in a month, has a Poisson distribution with mean 3.5. A new investigation was carried on in the month of September to test whether or not the number of occurrences has changed.

(a) State the hypotheses that should be used.

(b) Find a two-tail critical region, where each tail is as close as possible to 5%, for this test.

(c) Determine the actual significance level for this test.

(a) Let X = the number of cases of a certain medical condition reported in a month and follows a Poisson distribution,

$$X \sim Po\,(3.5)$$

The two hypotheses are:

$$H_0\colon \lambda = 3.5$$

$$H_1\colon \lambda \neq 3.5$$

(b)

This is a two-tailed test at a 10% significance level, i.e. 5% on each side.
So, let c_1 and c_2 be the two critical values.
To reject the null hypothesis the following probabilities have to hold.

$P(X \le c_1) \le 0.05$

$P(X \ge c_2) \le 0.05$
$1 - P(X \le c_2 - 1) \le 0.05$
$P(X \le c_2 - 1) \ge 0.95$

From the Poisson tables under $\lambda = 3.5$, we find that:

$P(X = 0) = 0.0302$ ✓
$P(X \le 1) = 0.1359$

We look for the two probabilities which are as close as possible to 0.05.

$P(X \le 6) = 0.9347$ ✓
$P(X \le 7) = 0.9733$

We look for the two probabilities which are as close as possible to 0.95.

So, $c_1 = 0$

$c_2 - 1 = 6$
$c_2 = 7$

Therefore, Critical Region: $x = 0$, $x \ge 7$.

(c) Actual Significance Level = $P(reject\ \ H_0) =$

$= P(X = 0) + P(X \ge 7) = P(X = 0) + 1 - P(X \le 6) =$
$= 0.0302 + 1 – 0.9347 = 0.0955 = 9.55\%$