Stats1 Chapter 7: Hypothesis Testing

One Tailed Tests

Doing a full one-tailed hypothesis test

We've done various bits of a hypothesis test, and haven't actually properly conducted one yet. Let's do an example!

John tosses a coin 8 times and it comes up heads 6 times. He claims the coin is **biased towards heads**. With a significance level of 5%, test his claim.

X is number of heads.
p is probability of heads.
$X \sim B(8, p)$

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

$H_0: p = 0.5$
$H_1: p > 0.5$

STEP 2: Write null and alternative hypotheses.

Assume H_0 is true, $X \sim B(8,0.5)$

$$P(X \geq 6) = 1 - P(X \leq 5)$$
$$= 1 - 0.8555$$
$$= 0.1445$$

STEP 3: Determine probability of observed test statistic (or 'more extreme'), assuming null hypothesis.
i.e. Determine probability we'd see this outcome just by chance.

14.45% > 5%, so insufficient evidence to reject H_0.
Coin is not biased.

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put in context of original problem.

C.D.F. Binomial table: $p = 0.5, n = 8$

x	$P(X \leq x)$
0	0.0039
1	0.0352
2	0.1445
3	0.3633
4	0.6367
5	0.8555
6	0.9648
7	0.9961

NEW TO A LEVEL 2017: The probability of 'the observed value or more extreme' is known as the p-value.

Alternative method using critical regions

We can also find the critical region and see if the test statistic lies within it.

John tosses a coin 8 times and it comes up heads 6 times. He claims the coin is **biased towards heads**. With a significance level of 5%, test his claim.

X is number of heads.
p is probability of heads.
$X \sim B(8, p)$

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

$H_0: p = 0.5$
$H_1: p > 0.5$

STEP 2: Write null and alternative hypotheses.

$P(X \geq 7) = 1 - 0.9648 = 0.0352$
Critical region is $X \geq 7$

STEP 3 (Alternative): Determine critical region.

6 is not in critical region, so do not reject H_0.
Coin is not biased.

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put in context of original problem.

C.D.F. Binomial table: $p = 0.5, n = 8$

x	$P(X \leq x)$
0	0.0039
1	0.0352
2	0.1445
3	0.3633
4	0.6367
5	0.8555
6	0.9648
7	0.9961

More on p-values

(Note that this is not covered in the Pearson textbook, but **is** in the specification)

Sheila wants to know if a coin is biased towards heads and throws it a large number of times, counting the number of heads. The p-value is less than 0.03. Conduct a hypothesis test at the 5% significance level.

Let p be the probability of heads.
$\boldsymbol{H_0: p = 0.5}$
$\boldsymbol{H_1: p > 0.5}$

$\boldsymbol{0.03 < 0.05}$ **so reject H_0.**
Sufficient evidence to suggest the coin is biased.

Note: Ordinarily we'd calculate the probability of seeing the observed number of heads 'or more extreme'. But this has already been done for us (i.e. the p-value), so we just need to compare this against the threshold.

Further Example

[Textbook] The standard treatment for a particular disease has a $\frac{2}{5}$ probability of success. A certain doctor has undertake research in this area and has produced a new drug which has been successful with 11 out of 20 patients. The doctor claims the new drug represents an improvement on the standard treatment.
Test, at the 5% significance level, the claim made by the doctor.

?

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

?

STEP 2: Write null and alternative hypotheses.

?

STEP 3: Determine probability of observed test statistic (or ‘more extreme’), assuming null hypothesis.

?

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put <u>in context of original problem</u>.

Further Example

[Textbook] The standard treatment for a particular disease has a $\frac{2}{5}$ probability of success. A certain doctor has undertake research in this area and has produced a new drug which has been successful with 11 out of 20 patients. The doctor claims the new drug represents an improvement on the standard treatment.
Test, at the 5% significance level, the claim made by the doctor.

X is number of patients for whom trial was successful.
p is probability of success in each patient.
$X \sim B(20, p)$

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

$H_0: p = 0.4$
$H_1: p > 0.4$

STEP 2: Write null and alternative hypotheses.

Assume H_0 is true, so $X \sim B(20,0.4)$
$P(X \geq 11) = 1 - P(X \leq 10) = 0.1275$

STEP 3: Determine probability of observed test statistic (or 'more extreme'), assuming null hypothesis.

12.75% > 5% so not enough evidence to reject H_0.
∴ New drug is no better than old one.

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put in context of original problem.

Test Your Understanding

Edexcel S2 Jan 2011 Q2

A student takes a multiple choice test. The test is made up of 10 questions each with 5 possible answers. The student gets 4 questions correct. Her teacher claims she was guessing the answers. Using a one tailed test, at the 5% level of significance, test whether or not there is evidence to reject the teacher's claim.

State your hypotheses clearly. **(6)**

?

Test Your Understanding

Edexcel S2 Jan 2011 Q2

A student takes a multiple choice test. The test is made up of 10 questions each with 5 possible answers. The student gets 4 questions correct. Her teacher claims she was guessing the answers. Using a one tailed test, at the 5% level of significance, test whether or not there is evidence to reject the teacher's claim.

State your hypotheses clearly. **(6)**

$H_0 : p = 0.2 \quad H_1 : p > 0.2$		B1
Under H_0, $X \sim \text{Bin}(10, 0.2)$		B1
$P(X \geq 4) = 1 - P(X \leq 3)$	OR $P(X \leq 4) = 0.9672$	M1
$= 1 - 0.8791$	$P(X \geq 5) = 0.0328$	
$= 0.1209$	CR $X \geq 5$	A1
0.1209>0.05. Insufficient evidence to reject H_0 so teacher's claim is supported.		M1A1ft
		[6]

Note the mark for stating distribution of X under null hypothesis.

Note two-mark conclusion.

Exercise 7.3

Pearson Applied Year 1/AS

Pages 47-48

Homework Exercise

1 A single observation, x, is taken from a binomial distribution $B(10, p)$ and a value of 5 is obtained. Use this observation to test H_0: $p = 0.25$ against H_1: $p > 0.25$ using a 5% significance level.

2 A random variable has distribution $X \sim B(10, p)$. A single observation of $x = 1$ is taken from this distribution. Test, at the 5% significance level, H_0: $p = 0.4$ against H_1: $p < 0.4$.

3 A single observation, x, is taken from a binomial distribution $B(20, p)$ and a value of 10 is obtained. Use this observation to test H_0: $p = 0.3$ against H_1: $p > 0.3$ using a 5% significance level.

4 A random variable has distribution $X \sim B(20, p)$. A single observation of $x = 3$ is taken from this distribution. Test, at the 1% significance level, H_0: $p = 0.45$ against H_1: $p < 0.45$.

5 A single observation, x, is taken from a binomial distribution $B(20, p)$ and a value of 2 is obtained. Use this observation to test H_0: $p = 0.28$ against H_1: $p < 0.28$ using a 5% significance level.

6 A random variable has distribution $X \sim B(8, p)$. A single observation of $x = 7$ is taken from this distribution. Test, at the 5% significance level, H_0: $p = 0.32$ against H_1: $p > 0.32$.

7 A dice used in playing a board game is suspected of not giving the number 6 often enough. During a particular game it was rolled 12 times and only one 6 appeared. Does this represent significant evidence, at the 5% level of significance, that the probability of a 6 on this dice is less than $\frac{1}{6}$?

Homework Exercise

8 The success rate of the standard treatment for patients suffering from a particular skin disease is claimed to be 68%.

a In a sample of n patients, X is the number for which the treatment is successful. Write down a suitable distribution to model X. Give reasons for your choice of model.

A random sample of 10 patients receives the standard treatment and in only 3 cases was the treatment successful. It is thought that the standard treatment was not as effective as it is claimed.

b Test the claim at the 5% level of significance.

9 A plant germination method is successful on average 4 times out of every 10. A horticulturist develops a new technique which she believes will improve the number of plants that successfully germinate. She takes a random sample of 20 seeds and attempts to germinate them.

a Using a 5% level of significance, find the critical region for a test to check her belief. **(4 marks)**

b Of her sample of 20 plants, the horticulturalist finds that 14 have germinated. Comment on this observation in light of the critical region. **(2 marks)**

Problem-solving

In this question you are told to find the critical region in part **a**. You will save time by using your critical region to answer part **b**.

10 A polling organisation claims that the support for a particular candidate is 35%. It is revealed that the candidate will pledge to support local charities if elected. The polling organisation think that the level of support will go up as a result. It takes a new poll of 50 voters.

a Describe the test statistic and state suitable null and alternative hypotheses. **(2 marks)**

b Using a 5% level of significance, find the critical region for a test to check the belief. **(4 marks)**

c In the new poll, 28 people are found to support the candidate. Comment on this observation in light of the critical region. **(2 marks)**

Homework Answers

1 $0.0781 > 0.05$
There is insufficient evidence to reject H_0.

2 $0.0464 < 0.05$
There is sufficient evidence to reject H_0 so $p < 0.04$.

3 $0.0480 < 0.05$
There is sufficient evidence to reject H_0 so $p > 0.30$.

4 $0.0049 < 0.01$
There is sufficient evidence to reject H_0 so $p < 0.45$.

5 $0.0526 > 0.05$
There is insufficient evidence to reject H_0 so there is no reason to doubt $p = 0.28$.

6 $0.0020 < 0.05$
There is sufficient evidence to reject H_0 so $p > 0.32$.

7 $0.3813 > 0.05$
There is insufficient evidence to reject H_0 (not significant).
There is no evidence that the probability is less than $\frac{1}{6}$.
There is no evidence that the dice is biased.

8 **a** Distribution $B(n, 0.68)$.
Fixed number of trials.
Outcomes of trials are independent.
There are two outcomes success and failure.
The probability of success is constant.

b $P(X \leqslant 3) = 0.0155 < 0.05$. There is sufficient evidence to reject the null hypothesis so $p < 0.68$.
The treatment is not as effective as claimed.

9 **a** Critical region is $X \geqslant 13$

b 14 lies in the critical region, so we can reject the null hypothesis. There is evidence that the new technique has improved the number of plants that germinate.

10 **a** The number of people who support the candidate.
H_0: $p = 0.35$, H_1: $p > 0.35$

b Critical region is $X \geqslant 24$

c 28 lies in the critical region, so we can reject the null hypothesis. There is evidence that the candidate's level of popularity has increased.