Hypothesis Testing Cheat Sheet

Edexcel FS1

A hypothesis test is used to investigate whether a population parameter differs from a certain value. In Stats and Mechanics Year 1, you learnt to carry out a hypothesis test concerning the parameter p for a binomial distribution. In this chapter, we will extend this further and learn to carry out hypothesis tests for the mean λ of a Poisson distribution and the parameter p for a geometric distribution.

Testing for the mean of a Poisson distribution

When carrying out a test for the mean of a Poisson distribution, you should use the following steps:

> If the test is one-tailed, H_1 will be $\lambda < c$ or $\lambda > c$ (depending on what you are told in the question. If the test is two-tailed, H_1 will be $\lambda \neq c$.

1) Define the test statistic.
2) State the null and alternative hypotheses. The null hypothesis will always be $\lambda = k$, where k is a constant determined based on the question. The alternative hypothesis will depend on whether the test is one-tailed or two-tailed.
3) Use the Poisson distribution with the mean from the null hypothesis to determine whether the observed value lies in the critical region.
4) Compare this probability to the significance level. If the probability is less than the significance level, then reject the null hypothesis. Otherwise, accept the null hypothesis.

> For example, if $H_1: \lambda < c$ then you need to find the probability that your test statistic is less than or equal to the observed value. Compare this probability to the significance level.

Example 1 (One-tailed test): The average number of flaws per 50m of cloth produced by a machine is found to be 2.3. After the machine is serviced, the number of flaws in the first 150m is found to be 2. Test, at the 5% level of significance, whether or not the average number of flaws has decreased.

1) Start by defining the test statistic. The mean was 2.3 for 50m of cloth, so for 150m of cloth the mean will be $2.3 \times 3 = 6.9$.	Let X be the number of flaws in 150m of cloth. Then $X \sim Po(6.9)$.
2) State the null and alternative hypotheses. We are testing whether the number of flaws has decreased, so this is a one-tail test.	$H_0: \lambda = 6.9$ $H_1: \lambda < 6.9$
3) We need to find whether the 2 flaws are significant enough to conclude that the mean has decreased. 2 is less than the mean of 6.9, so we are specifically testing the lower tail. This means we need to find $P(X \leq 2)$ rather than $P(X \geq 2)$.	$P(X \leq 2) = 0.032$
4) Compare this to the significance level (5%). State whether you are accepting or rejecting H_0, and give a conclusion in the context of the question.	$0.032 < 0.05$, so $X = 2$ does in fact lie in the critical region. $\therefore$ the result of the test is significant. There may be sufficient evidence to suggest that the average number of flaws has in fact decreased.

Example 2 (Two-tailed test): Breakdowns occur on a particular machine at a rate of 1.5 every week. A manager feels that the rate of breakdowns has changed and decides to monitor the machine. Over a 6-week period she finds that there are 13 breakdowns. Test at the 5% level of significance, whether or not the manager's suspicion is correct.

1) Start by defining the test statistic. The sample data given is for 6-weeks so our test statistic should model the number of breakdowns in 6 weeks.	Let X be the number of breakdowns in 6 weeks. Then $X \sim Po(9)$.
2) State the null and alternative hypotheses. We are testing whether the number of breakdowns has changed; there is no reference to an increase or decrease, so this is a two-tail test.	$H_0: \lambda = 9$ $H_1: \lambda \neq 9$
3) We need to find whether the 13 breakdowns in the 6-week period is significant. This is a two-tailed test so find both $P(X \geq 13)$ and $P(X \leq 13)$.	$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.876 = 0.124$
4) Because the test is two-tailed we must compare this to the half the significance level (2.5%). State whether you are accepting or rejecting H_0, and give a conclusion in the context of the question.	$0.124 > 0.025$, so $X = 13$ does not lie in the critical region. $\therefore$ the result of the test is insignificant. Accept H_0. There is insufficient evidence to suggest the rate of breakdowns has changed.

Finding critical regions for a Poisson distribution

Recall that the critical region is the set of values of your test statistic that would cause you to reject the null hypothesis. A critical value is a value on the boundary of the critical region. You need to be comfortable with finding critical region(s) for a one or two-tailed test for the mean of a Poisson distribution. It is often more convenient to use the tables for such questions, though using your calculator is perfectly fine too.

- The actual significance level is the probability of incorrectly rejecting H_0.

> In other words, this is the probability that the test statistic falls into the critical region, assuming H_0 is correct.

Example 3: Millie manufactures printed material. She knows that defects occur randomly in the manufacturing process at a rate of 1 every 7 metres. Once a week the machinery is cleaned and reset. Millie then takes a random sample of 35 metres of material from the next batch produced to test if there has been any change in the rate of defects.
a) Stating your hypotheses clearly and using a 10% level of significance, find the critical region for this test. You should choose your critical region so that the probability of rejection is less than 0.05 in each tail.
b) State the actual significance level of this test.

a) Start by defining the test statistic. The sample data given is for 35 metres so our test statistic should model the number of defects in 35 metres.	Let X be the number of defects in 35m. Then $X \sim Po(5)$.
State the null and alternative hypotheses. We are testing whether the rate of defects has changed; there is no reference to an increase or decrease, so this is a two-tail test.	$H_0: \lambda = 5$ $H_1: \lambda \neq 5$
To find the critical region, we look at each tail separately. The significance level is 10%, so we use a 5% significance level in each tail.	For the lower tail, we are looking for the greatest value of k such that $P(X \leq k) < 0.05$. Using tables: $\Rightarrow P(X \leq 1) = 0.0404 < 0.05$ $\Rightarrow P(X \leq 2) = 0.1247 > 0.05$ $X = 1$ is the largest value in the lower critical region so $X \leq 1$ is the lower critical region.
For the upper tail:	For the upper tail, we are looking for the smallest value of k such that $P(X \geq k) < 0.05$. Using tables: $\Rightarrow P(X \geq 9) = 1 - P(X \leq 8) = 0.0681 > 0.05$ $\Rightarrow P(X \geq 10) = 1 - P(X \leq 9) = 0.0318 < 0.05$ $X = 10$ is the smallest value in the upper critical region so $X \geq 10$ is the upper critical region.
Put both parts of the critical region together.	So, the complete critical region is given by $X \leq 1 \cup X \geq 10$
b) The actual significance level is simply the probability of incorrectly rejecting H_0, which is given by `$P(X \leq 1) + P(X \geq 10)$.	$0.0404 + 0.0318 = 0.0722$

Using a Poisson approximation

Some questions will require you to approximate a binomial distribution by a Poisson distribution in order to carry out the test. This can be done if the probability of success in the binomial test is very low and the number trials is big (usually >=50).

Example 4: During an influenza epidemic, 4% of the population of a large city were affected on a given day. The manager of a factory that employs 250 people found that 17 of the employees at his factory were absent, claiming to be suffering from influenza. Using a Poisson approximation to the binomial distribution and a 5% level of significance, test whether or not the proportion of employees suffering from influenza at his factory was larger than that of the whole city.

1) Start by defining the test statistic. The sample size is 250 and the known probability is 0.04.	Let X be the number of employees suffering from influenza. Then $X \sim B(250, 0.04)$.
2) State the null and alternative hypotheses. We are testing whether the proportion has increased. This is a one-tail test referring to the population parameter of the binomial distribution.	$H_0: p = 0.04$ $H_1: p > 0.04$
3) We apply the Poisson approximation.	$mean = np = 250 \times 0.04 = 10$ $\therefore X \approx \sim Po(10)$
Now we need to find whether the figure of 17 absences is significant. This is an upper tail test so we find $P(X \geq 17)$.	$P(X \geq 17) = 1 - P(X \leq 16) = 0.027$
4) Compare this to the significance level (5%). State whether you are accepting or rejecting H_0, and give a conclusion in the context of the question.	$0.027 < 0.05$, so $X = 17$ does in fact lie in the critical region. $\therefore$ the result of the test is significant. There is sufficient evidence to suggest that the proportion of employees suffering at his factory was larger than that of the whole city.

Testing for the mean of a Geometric distribution

You will only need to carry out a one-tail test for the geometric distribution. You should use the following steps:

> If the test is one-tailed, H_1 will be $p < c$ or $p > c$. If the test is two-tailed, H_1 will be $p \neq c$.

1) Define the test statistic.
2) State the null and alternative hypotheses. The null hypothesis will always be $p = k$, where k is a constant. The alternative hypothesis will be $p < k$ or $p > k$, depending on what you are told in the question.
3) Use the Geometric distribution with the value of p from the null hypothesis to determine whether the observed value lies in the critical region.
4) Compare this probability to the significance level. If the probability is less than the significance level, then reject the null hypothesis. Otherwise, accept the null hypothesis.

Important: unlike hypothesis testing with other distributions, when you are testing for an increase in p for a geometric distribution, you need to consider the lower tail of the distribution. Likewise, when you are testing for a decrease in p, you need to consider the upper tail of the distribution.

The following results are useful when carrying out hypothesis tests for a Geometric distribution. If $X \sim Geo(p)$, then:

- $P(X = x) = p(1-p)^{x-1}$
- $P(X \leq x) = 1 - (1-p)^x$
- $P(X \geq x) = (1-p)^{x-1}$

> You came across these in Chapter 3.

Example 5: Marie claims that she scores a penalty on 30% of her attempts. One of her rivals claims that she is overstating her ability. In an attempt to prove her case, Marie takes consecutive shots until she scores her first penalty. She scores her first penalty on her 10th shot. Test her rival's claim, using a 5% level of significance and clearly stating your null and alternative hypotheses.

1) Start by defining the test statistic. We assume the value of p stated (0.3) is correct.	Let X be the number of attempts until a penalty is scored. Then $X \sim Geo(0.3)$.
2) State the null and alternative hypotheses. We are testing whether the probability is overstated, so the alternative hypothesis is $p < 0.3$.	$H_0: p = 0.3$ $H_1: p < 0.3$
3) We need to find whether 1 goal in 10 shots is significant enough to conclude that p is lower than 0.3. So we find $P(X \geq 10)$ to determine whether $X = 10$ is in the critical region. Be careful: we don't instead find $P(X \leq 10)$ is because lower values of X correlate to bigger values of p, which is not what we are testing for.	$P(X \geq 10) = (1-0.3)^{10-1} = 0.7^9 = 0.040$
4) Compare this to the significance level (5%). State whether you are accepting or rejecting H_0, and give a conclusion in the context of the question.	$0.040 < 0.05$, so $X = 10$ does in fact lie in the critical region. $\therefore$ the result of the test is significant. There is sufficient evidence to suggest that the 30% figure is overstated.

Finding critical regions for a geometric distribution

Finding the critical region for a geometric distribution is done in a very similar way to the Poisson distribution. If we are testing for the lower tail, then we need to find the greatest value of k such that $P(X \leq k) < 0.05$. If we are testing for the upper tail instead, then we need to find the smallest value of k such that $P(X \geq k) < 0.05$.

Example 6: It is known that 15% of products produced by a machine are defective. Products are tested, one at a time, until the first defective one is encountered. The machine is serviced, and it is hoped that this has reduced the proportion of defective products.
a) Find the critical region for a hypothesis test that the proportion of defective products has reduced. Use a 5% level of significance.
b) Find the actual significance level.

a) Start by defining the test statistic.	Let X be the number of products tested until a defective one is found. Then $X \sim Geo(0.15)$.
To find the critical region, we set $P(X \geq k) < 0.05$ and solve for k. We are testing for a decrease in p so consider the upper tail of the distribution.	$P(X \geq k) = (1-0.15)^{k-1} = 0.85^{k-1} < 0.05$ $0.85^{k-1} < 0.05$
Take logs of both sides.	$\ln(0.85^{k-1}) < \ln 0.05$ $(k-1)\ln(0.85) < \ln 0.05$
Divide by ln 0.85 before rearranging for k. Since $\ln 0.85 < 0$, the inequality flips.	$k - 1 > \frac{\ln 0.05}{\ln 0.85}$ $k > 19.433 \ldots$
State the critical region.	So critical region is $X \geq 20$.
b) The actual significance level is simply the probability of incorrectly rejecting H_0, which is given by $P(X \geq 20)$.	$P(X \geq 20) = (1-0.15)^{20-1} = 0.0456\ (3\ s.f.)$