Binomial hypothesis tests

S5.2 · Statistics · revision notes

Testing a claim about a population proportion p with the binomial distribution: writing hypotheses, finding a p-value or a critical region (one- or two-tailed), and writing a conclusion in context. A complete test is typically worth around 4 to 8 marks.

Key facts

- $H_0 : p = p_0$; $H_1 : p > p_0$, $p < p_0$ or $p \neq p_0$
- Upper tail: $P(X \geq x) = 1 - P(X \leq x - 1)$
- Critical region: the largest tail whose probability is no more than the significance level
- Actual significance level = probability of the critical region under H_0
- Under H_0: $X \sim B(n, p_0)$
- Lower tail: $P(X \leq x)$
- Two-tailed: at most half the significance level in each tail

Notes

Setting up

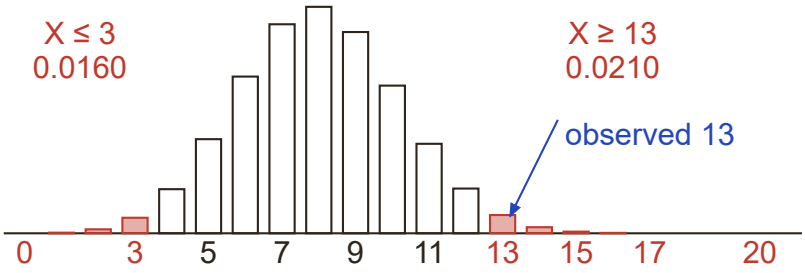
- Define p in context as the probability of success (the population proportion). $H_0 : p = p_0$, the claimed value.
- Choose H_1 from the wording: ‘increased’ or ‘more’ gives $p > p_0$; ‘decreased’ or ‘less’ gives $p < p_0$; ‘changed’ or ‘different’ gives $p \neq p_0$
- Under H_0, the number of successes in a sample of size n is $X \sim B(n, p_0)$

Method 1: p-value

- For $H_1 : p > p_0$, find $P(X \geq x) = 1 - P(X \leq x - 1)$, where x is the observed value; for $H_1 : p < p_0$, find $P(X \leq x)$
- In a two-tailed test, use the tail the observed value is in (above or below np_0, the number you expect under H_0) and compare with half the significance level. If the probability is smaller than the level you compare with, reject H_0

Method 2: critical region

- Upper tail: the smallest c with $P(X \geq c) \leq$ the significance level gives the critical region $X \geq c$. Lower tail: the largest c with $P(X \leq c) \leq$ the significance level gives $X \leq c$
- Use a table of cumulative probabilities on your calculator and quote the probabilities either side of each boundary.
- Two-tailed: each tail has probability at most half the significance level. The actual significance level is the sum of the two tail probabilities.



actual significance level = 0.0160 + 0.0210 = 0.0370

Two-tailed test, H_0: p = 0.4, n = 20: each tail at most 2.5%

- If a question asks for tails ‘as close as possible to’ a value instead, follow that wording.
- If the observed value is in the critical region, reject H_0

Conclusion

- Give the decision, then a conclusion in context that does not claim certainty, e.g. ‘there is evidence to suggest that the proportion of customers who use the app has increased’.

How to answer each type of question

M = method mark, A = answer mark, B = a mark on its own.

One-tailed test using a p-value

4 to 5 marks **Grade D**

1. Write the hypotheses, with p defined.
2. State the distribution under H_0
3. Find the probability of the observed value or a more extreme one.
4. Compare, decide and conclude in context.

Example. A company claims that 30% of its customers use its app. The manager believes that the proportion is higher. In a random sample of 25 customers, 12 use the app. Test the manager's belief at the 5% level of significance.

Model answer.

$H_0 : p = 0.3$, $H_1 : p > 0.3$, where p is the proportion of customers who use the app (B1)

Under H_0, $X \sim B(25, 0.3)$ (M1)

$P(X \geq 12) = 1 - P(X \leq 11)$

$= 1 - 0.9558 = 0.0442$ (A1)

$0.0442 < 0.05$, so reject H_0 (M1)

There is sufficient evidence to suggest that more than 30% of customers use the app (A1)

One-tailed critical region (lower tail)

3 to 4 marks **Grade C**

1. Use the distribution under H_0
2. Find the largest c with $P(X \leq c)$ no more than the significance level.
3. Quote the probabilities either side and write the region as $X \leq c$

Example. A treatment works for 70% of patients. A researcher suspects that a new batch is less effective. It is given to 20 patients, and X is the number for whom it works. Find the critical region for a test at the 5% level, and state the actual significance level.

Model answer.

$H_0 : p = 0.7$, $H_1 : p < 0.7$; $X \sim B(20, 0.7)$ (B1)

$P(X \leq 10) = 0.0480$ and $P(X \leq 11) = 0.1133$ (M1)

Critical region $X \leq 10$ (A1); actual significance level 0.0480 (B1)

Two-tailed critical regions and actual significance level

6 to 7 marks | **Grade B**

1. Write the hypotheses with $\neq$
2. Allow at most half the significance level in each tail.
3. Quote the probabilities either side of each boundary.
4. Add the two tail probabilities for the actual significance level.

Example. A gardener claims that 40% of seeds of a certain type germinate. A random sample of 20 seeds is planted to test whether the proportion is different from 40%.
(a) Using a 5% level of significance, find the critical region for the number of seeds that germinate.
(b) Find the actual significance level of the test.
(c) 13 of the seeds germinate. State the conclusion of the test.

Model answer.

(a) $H_0 : p = 0.4$, $H_1 : p \neq 0.4$ (B1); $X \sim B(20, 0.4)$

Lower tail: $P(X \leq 3) = 0.0160$, $P(X \leq 4) = 0.0510$ (M1), so $X \leq 3$ (A1)

Upper tail: $P(X \geq 13) = 1 - 0.9790 = 0.0210$, $P(X \geq 12) = 0.0565$, so $X \geq 13$ (A1)

(b) $0.0160 + 0.0210$ (M1) $= 0.0370$ (A1)

(c) 13 is in the critical region, so reject H_0: there is evidence to suggest that the proportion of seeds that germinate is not 40% (B1)

Smallest sample size for a significant result

3 marks | **Grade A***

1. Write the probability of the most extreme outcome in terms of n
2. Set it to be no more than the significance level.
3. Solve using logarithms and round up.

Example. A test of $H_0 : p = 0.2$ against $H_1 : p < 0.2$ is carried out at the 5% level using a sample of size n. Find the smallest value of n for which observing no successes would lead to rejecting H_0

Model answer.

$P(X = 0) = 0.8^n \leq 0.05$ (M1)

$n \geq \dfrac{\log 0.05}{\log 0.8} = 13.4$ (M1)

$n = 14$ (A1)

(check: $0.8^{13} = 0.0550$ and $0.8^{14} = 0.0440$)

Don't lose marks

- Using $P(X = x)$ instead of $P(X \geq x)$ or $P(X \leq x)$ for the p-value.

- Using $1 - P(X \leq x)$ for $P(X \geq x)$
- Writing the critical region as a probability instead of as values of X
- Using the full significance level in each tail of a two-tailed test.
- Conclusions that claim certainty, or that don't mention the context.

More tips

Memory tricks

- Use your calculator's table mode for binomial CD to list $P(X \leq x)$ for many values at once and spot the boundary.
- Write $P(X \geq c) = 1 - P(X \leq c - 1)$ every time: it avoids the most common slip.
- Two-tailed test: compare the observed value with np_0, the number you expect under H_0, to see which tail it is in, then test that tail only.

Exam technique

- For critical regions, show the probabilities either side of each boundary: they earn the method mark.
- Write critical regions in terms of X, e.g. $X \leq 3$ or $X \geq 13$
- Read the tail rule in the question: ‘at most 2.5%’ and ‘as close as possible to 2.5%’ can give different answers.

What each grade needs

What you need to be able to do, from the first marks up to the top grade.

Grade E — **Write hypotheses for a binomial test**
$H_0 : p = p_0$, with H_1 from the wording and p defined in context.

Grade D — **Carry out a one-tailed test with a p-value**
Find $P(X \geq x)$ or $P(X \leq x)$ under H_0 and compare it with the significance level.

Grade C — **Find a one-tailed critical region**
Find the boundary where the tail probability is no more than the significance level.

Grade B — **Find critical regions for a two-tailed test**
Allow at most half the significance level in each tail.

Grade B — **Find the actual significance level**
Find the probability of the critical region under H_0 (adding both tails for a two-tailed test).

Grade A — **Carry out a complete test in context**
Hypotheses, distribution, probability or critical region, decision and conclusion in context.

Grade A* — **Find a sample size that gives significance**
E.g. find the smallest n for which $X = 0$ is in the critical region.

Free notes, Full packs, practice papers and mock exams at chhetriacademy.com. Written for this site; not taken from past papers. Stuck? 1-to-1 tutoring with a free 30-minute trial lesson.