

Hypothesis tests

A-Level Maths 9709 · Topic 6.5 revision notes

Everything from the 3 lessons of this topic on a few pages: the key idea of each lesson, the diagram to remember, the points to learn and the traps examiners set. Read it, cover it, test yourself.

1. Hypothesis tests with binomial and Poisson distributions

2. Hypothesis tests for a population mean

3. Type I and Type II errors

thelucidstem.com

1 Hypothesis tests with binomial and Poisson distributions

Syllabus 6.5

Assume H_0 is true. Find the probability of a result at least as extreme as the one observed (in the direction of H_1), or find the critical region. If that probability is below the significance level, the result lies in the critical region: reject H_0 . Always conclude in context and without certainty: "there is (in)sufficient evidence that ...".

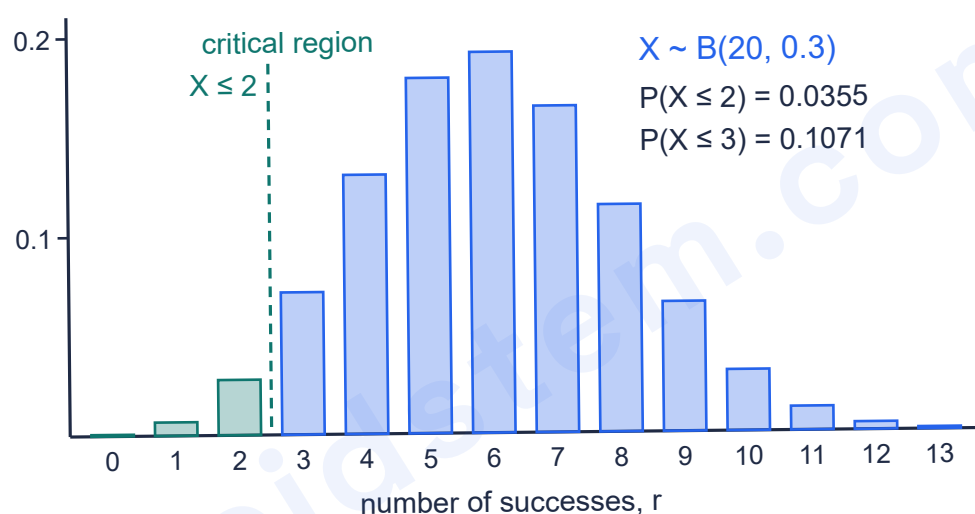


Fig. 1: $B(20, 0.3)$, drawn to scale. For a lower-tail test at the 5% level the critical region is $X \leq 2$, because $P(X \leq 2) = 0.0355 \leq 0.05$ but $P(X \leq 3) = 0.1071 > 0.05$; the actual significance level is 3.55%

Worked example: A one-tailed binomial test

A die is thrown 30 times and gives 9 sixes. Test, at the 5% significance level, whether the die is biased towards six.

1. $H_0 : p = \frac{1}{6}$; $H_1 : p > \frac{1}{6}$ (one-tailed). Under H_0 , $X \sim B(30, \frac{1}{6})$.
2. Probability of 9 or more sixes: $P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.9803 = 0.0197$.
3. $0.0197 < 0.05$, so the result is in the critical region: reject H_0 .

Answer: There is sufficient evidence at the 5% level that the die is biased towards six.

Hypotheses, distribution under H_0 , tail, compare, conclude.

- ✓ **Hypotheses:** $H_0 : p = p_0$ (or $\lambda = \lambda_0$); H_1 with $<$, $>$ or $\neq$ from the wording.
- ✓ **Distribution under H_0 :** $B(n, p_0)$ or $Po(\lambda_0)$, scaled to the observed period.
- ✓ **Tail probability:** observed value or more extreme, in the direction of H_1 .
- ✓ **Compare** with the significance level (half of it in each tail for a two-tailed test).
- ✓ **Conclude** in context, as evidence, never as proof.
- ✓ **Large n or λ :** normal approximation with continuity correction.

Watch out!

Compare the TAIL probability $P(X \geq 9)$ with 5%, never the single probability $P(X = 9)$. And never write " H_0 is true" or "this proves the die is biased": a test only gives evidence.

Write hypotheses about the PARAMETER (p or λ), not about the observed count, and scale λ to the period of the observation: 2 months at 3.2 per month means $H_0 : \lambda = 6.4$.

The continuity correction points AWAY from the tail: $X \geq 95$ becomes $Y > 94.5$, and $X \leq 8$ becomes $Y < 8.5$.

2 Hypothesis tests for a population mean

Syllabus 6.5

Under $H_0 : \mu = \mu_0$, $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$, so the test statistic is $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$. Compare z with the critical value (e.g. 1.645 one-tailed at 5%, ± 1.96 two-tailed at 5%). For a large sample with unknown σ , use the unbiased estimate s and the Central Limit Theorem.

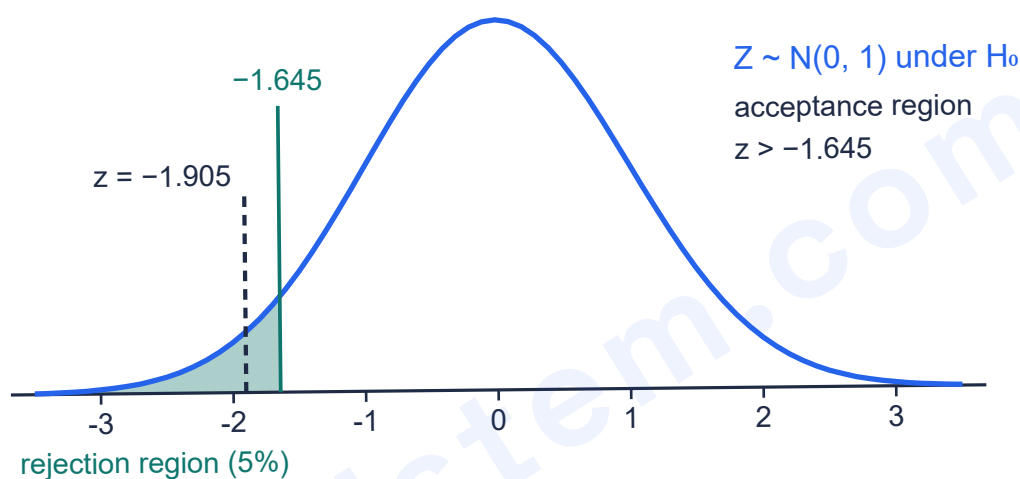


Fig. 1: the standard normal distribution under H_0 , drawn to scale. For a lower-tail test at 5%, the rejection region is $z < -1.645$ (shaded). The test statistic -1.905 from the bottling example lies inside it, so H_0 is rejected

Worked example: A one-tailed test

Volumes of drink in bottles are $N(\mu, 4^2)$ ml. The company claims $\mu = 500$. A random sample of 12 bottles has mean 497.8 ml. Test at the 5% level whether the mean is less than claimed.

1. $H_0 : \mu = 500$; $H_1 : \mu < 500$.
2. Under H_0 , $\bar{X} \sim N\left(500, \frac{16}{12}\right)$.
3. $z = \frac{497.8 - 500}{4/\sqrt{12}} = -1.905$.
4. Critical value -1.645 ; $-1.905 < -1.645$, so reject H_0 (Fig. 1). (Equivalently $\Phi(-1.905) = 0.0284 < 0.05$.)

Answer: There is sufficient evidence at the 5% level that the mean volume is less than 500 ml.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}, \text{ then}$$

compare.

- ✓ **Hypotheses:** $H_0 : \mu = \mu_0$; H_1 with $<$, $>$ or $\neq$.
- ✓ **Distribution:** $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$; say why it is normal.
- ✓ **Unknown σ , large n :** use the unbiased s^2 .
- ✓ **Compare:** z with the critical value, or the tail probability with the level.
- ✓ **Critical region for $\bar{x}$:** $\mu_0 \pm z \frac{\sigma}{\sqrt{n}}$.
- ✓ **Conclude** in context.

Watch out!

Divide by $\frac{\sigma}{\sqrt{n}}$, not by σ . With σ alone the bottling example gives $z = -0.55$ and the wrong decision.

In a two-tailed test at 5%, the critical value is 1.96, not 1.645. Using the one-tailed value makes you reject too easily.

"Accept H_0 " overstates the case. Prefer "do not reject H_0 ; there is insufficient evidence that ...". Failing to find evidence of a change is not evidence that there was no change.

3 Type I and Type II errors

Syllabus 6.5

Type I error: rejecting H_0 when it is true; $P(\text{Type I}) = P(\text{result in critical region} \mid H_0 \text{ true})$, the actual significance level. Type II error: not rejecting H_0 when it is false; $P(\text{Type II}) = P(\text{result in acceptance region} \mid \text{a given true value})$.

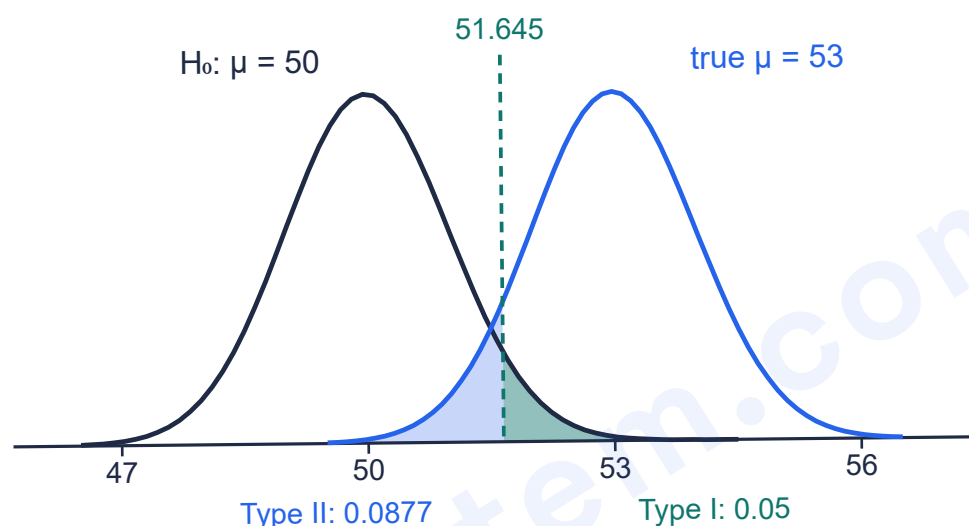


Fig. 1: the sample mean $X \sim N(\mu, 1)$ for a test of $H_0: \mu = 50$ against $H_1: \mu > 50$ at 5%, drawn to scale. Teal: Type I error, area 0.05 to the right of 51.645 under H_0 . Blue: Type II error when μ is really 53, area 0.0877 to the left of 51.645

Worked example: Errors in context

A test of $H_0: p = 0.3$ against $H_1: p < 0.3$ is used to decide whether a seed treatment reduces the failure rate p . Describe a Type I error and a Type II error in context.

1. Type I: rejecting H_0 when it is true.
2. In context: concluding that the treatment reduces the failure rate when in fact it does not.
3. Type II: not rejecting H_0 when it is false.
4. In context: concluding there is no evidence of a reduction when the treatment really does reduce the failure rate.

Answer: Type I: a false claim that the treatment works. Type II: failing to detect that it does work.

Critical region first, then the right distribution.

- ✓ **Type I:** reject a true H_0 . Probability = critical region under H_0 .
- ✓ **Type II:** keep a false H_0 . Probability = acceptance region under the stated true value.
- ✓ **Discrete:** $P(\text{Type I})$ is the actual significance level, at most the nominal one.
- ✓ **Normal:** $P(\text{Type I})$ equals the significance level; find the critical $\bar{x}$ using μ_0 .
- ✓ **Context:** describe each error in the words of the question.
- ✓ **Reducing both:** a larger sample lowers $P(\text{Type II})$ at a fixed significance level.

Watch out!

Do not describe the errors with "accept" and "reject" mixed up. Write the definition, then translate it into the context: Cambridge often gives one mark for each.

For a discrete test, $P(\text{Type I})$ is NOT the significance level (5%). It is the probability of the actual critical region, here 0.0355 or 0.0174.

Find the critical value using H_0 's mean, but the Type II probability using the TRUE mean.

Standardising the critical value with 50 again just gives back 0.95.