A Level Statistics
Practice Test 1: Hypothesis Testing

Instructions:
Answer all questions. Show your working clearly.
Calculators may be used unless stated otherwise.
Draw diagrams where appropriate to illustrate your solutions.
Time allowed: 3 hours

Section A: Fundamental Concepts [25 marks]

1. [**10 marks**] Define the basic concepts of hypothesis testing:

(a) Define a null hypothesis (H_0) and explain its purpose.

(b) Define an alternative hypothesis (H_1) and explain how it relates to the null hypothesis.

(c) Explain what is meant by a population parameter in the context of hypothesis testing.

(d) Give three examples of population parameters that might be tested.

(e) Explain the difference between one-tailed and two-tailed tests.

2. [**8 marks**] Define significance levels and p-values:

(a) Define the significance level () and explain its meaning.

(b) State three commonly used significance levels in statistical testing.

(c) Define a p-value and explain how it is interpreted.

(d) Explain the relationship between the p-value and the significance level in decision making.

3. [**7 marks**] Explain the types of errors in hypothesis testing:

(a) Define Type I error and state its probability.

(b) Define Type II error and explain when it occurs.

(c) Create a table showing the four possible outcomes of a hypothesis test.

(d) Explain why we cannot eliminate both types of errors simultaneously.

Section B: Setting Up Hypotheses [20 marks]

4. [**12 marks**] For each scenario, write appropriate null and alternative hypotheses:

(a) A manufacturer claims that the average weight of their chocolate bars is 50g. You suspect the bars are underweight.

(b) A new teaching method is being tested to see if it improves student performance compared to the current average score of 65%.

(c) A coin is suspected of being biased. The probability of heads should be 0.5 if fair.

(d) A medical researcher wants to test if a new drug has a different effect than the current drug, which has a 70% success rate.

(e) An environmental scientist claims that the mean temperature has increased from the historical average of 15°C.

(f) A quality control manager wants to test if the proportion of defective items has changed from the standard 2%.

5. [**8 marks**] Identify the type of test required for each scenario:

(a) Testing if the mean height of students has changed from 170cm.

(b) Testing if more than 60% of customers are satisfied with a new service.

(c) Comparing exam scores before and after a training program for the same students.

(d) Testing if the number of accidents follows a Poisson distribution.

(e) Comparing the mean salaries of two different professions.

(f) Testing if gender and voting preference are independent.

(g) Testing if a die is fair by examining the frequency of each outcome.

(h) Comparing the effectiveness of two different medications on independent groups.

Section C: Critical Regions and Test Statistics [25 marks]

6. [**10 marks**] Explain critical regions and critical values:

(a) Define a critical region and explain its purpose.

(b) Explain what critical values are and how they are determined.

(c) Describe how the critical region changes for one-tailed vs. two-tailed tests.

(d) Explain the relationship between significance level and the size of the critical region.

(e) Sketch diagrams showing critical regions for both one-tailed and two-tailed tests at $= 0.05$.

7. [**15 marks**] A factory claims that the mean weight of their products is 500g with standard deviation 20g. A sample of 25 products has a mean weight of 495g.

(a) State appropriate null and alternative hypotheses to test if the products are underweight.

(b) Calculate the test statistic for this scenario.

(c) Find the critical value for a one-tailed test at $= 0.05$.

(d) Determine the critical region for this test.

(e) Make a conclusion about the null hypothesis.

(f) Calculate the p-value for this test.

(g) Interpret the p-value in the context of the problem.

Section D: One-Sample t-Tests [30 marks]

8. [**12 marks**] Explain the one-sample t-test:

(a) State when a one-sample t-test is used instead of a z-test.

(b) Write the formula for the t-test statistic.

(c) Explain what degrees of freedom means and how to calculate it.

(d) State the assumptions required for a valid t-test.

(e) Explain how the t-distribution differs from the standard normal distribution.

9. [**18 marks**] A coffee shop claims that the average time to serve a customer is 3 minutes. A random sample of 16 customers shows the following service times (in minutes):

$$2.8, 3.2, 2.9, 3.5, 3.1, 2.7, 3.4, 3.0, 2.9, 3.3, 3.2, 2.8, 3.1, 3.6, 2.9, 3.0$$

(a) Calculate the sample mean and sample standard deviation.

(b) State appropriate hypotheses to test the shop's claim using $= 0.05$.

(c) Check that the assumptions for a t-test are satisfied.

(d) Calculate the t-test statistic.

(e) Find the critical value(s) for this test.

(f) Determine the p-value for the test.

(g) State your conclusion in the context of the problem.

(h) Calculate a 95% confidence interval for the true mean service time.

(i) Explain how the confidence interval supports your hypothesis test conclusion.

Section E: Chi-Squared Goodness of Fit [25 marks]

10. [**8 marks**] Explain the chi-squared goodness of fit test:

(a) State the purpose of a goodness of fit test.

(b) Write the formula for the chi-squared test statistic.

(c) Explain what ”observed” and ”expected” frequencies represent.

(d) State the assumptions required for a valid chi-squared test.

11. [**17 marks**] A researcher claims that exam grades follow the distribution: A (20%), B (30%), C (35%), D (10%), F (5%). In a sample of 200 students, the observed frequencies are: A (45), B (55), C (75), D (18), F (7).

(a) State the null and alternative hypotheses for this test.

(b) Calculate the expected frequencies for each grade.

(c) Verify that the assumptions for the chi-squared test are met.

(d) Calculate the chi-squared test statistic.

(e) Determine the degrees of freedom for this test.

(f) Find the critical value at $= 0.05$.

(g) Calculate the p-value for the test.

(h) State your conclusion about whether the data fits the claimed distribution.

(i) Identify which grade categories contribute most to the test statistic.

Section F: Two-Sample t-Tests [35 marks]

12. [**12 marks**] Explain two-sample t-tests:

(a) Distinguish between independent and paired samples.

(b) State when to use an independent samples t-test.

(c) Write the formula for the pooled variance in an independent samples t-test.

(d) State when to use a paired samples t-test.

(e) Write the test statistic formula for a paired t-test.

(f) State the assumptions for both types of two-sample t-tests.

13. [**12 marks**] Two groups of students use different study methods. Group A ($n = 20$) has mean score 78 with standard deviation 8. Group B ($n = 18$) has mean score 82 with standard deviation 7.

(a) State hypotheses to test if there's a significant difference between the methods.

(b) Calculate the pooled standard deviation.

(c) Calculate the standard error of the difference in means.

(d) Calculate the t-test statistic.

(e) Determine the degrees of freedom.

(f) Find the critical value for $= 0.05$ (two-tailed).

(g) State your conclusion about the difference in study methods.

14. [**11 marks**] A fitness program claims to reduce weight. Eight participants' weights (kg) before and after the program are:

Participant	1	2	3	4	5	6	7	8
Before	78	85	72	90	68	82	76	88
After	75	83	71	87	66	80	74	85

(a) Calculate the differences (Before - After) for each participant.

(b) Calculate the mean and standard deviation of the differences.

(c) State appropriate hypotheses for testing the program's effectiveness.

(d) Calculate the t-test statistic for the paired test.

(e) Find the p-value and state your conclusion at $= 0.05$.

(f) Calculate a 95% confidence interval for the mean weight loss.

Section G: Chi-Squared Test for Independence [30 marks]

15. [**10 marks**] Explain the chi-squared test for independence:

(a) State the purpose of testing for independence between two variables.

(b) Explain what a contingency table represents.

(c) Write the formula for expected frequencies in a contingency table.

(d) State the null and alternative hypotheses for independence testing.

(e) Explain how degrees of freedom are calculated for this test.

16. [**20 marks**] A study examines the relationship between gender and preferred subject. The data collected is:

	Mathematics	Science	English	Total
Male	45	35	20	100
Female	30	40	30	100
Total	75	75	50	200

(a) State the hypotheses for testing independence between gender and subject preference.

(b) Calculate all expected frequencies for the contingency table.

(c) Verify that the assumptions for the chi-squared test are satisfied.

(d) Calculate the chi-squared test statistic.

(e) Determine the degrees of freedom.

(f) Find the critical value at = 0.01.

(g) Calculate the p-value for the test.

(h) State your conclusion about the independence of gender and subject preference.

(i) Identify which cells contribute most to the test statistic.

(j) Calculate and interpret the standardized residuals for each cell.

Section H: Multiple Testing Scenarios [25 marks]

17. [**12 marks**] A pharmaceutical company tests a new drug. They claim it's effective for 80% of patients. In a trial of 50 patients, 35 showed improvement.

(a) State hypotheses to test the company's claim.

(b) Identify the appropriate test and check its assumptions.

(c) Calculate the test statistic.

(d) Find the p-value for a two-tailed test.

(e) State your conclusion at = 0.05.

(f) Discuss whether the sample size is adequate for this test.

18. [**13 marks**] A researcher studies three teaching methods by measuring improvement scores: Method A: 12, 15, 18, 14, 16, 13, 17 Method B: 18, 22, 20, 19, 21, 23, 20 Method C: 16, 19, 17, 20, 18, 21, 19

(a) Calculate the mean and standard deviation for each method.

(b) State hypotheses to test if all three methods are equally effective.

(c) Explain why a one-way ANOVA would be more appropriate than multiple t-tests.

(d) Perform pairwise comparisons between methods using two-sample t-tests.

(e) Calculate the test statistics for all three comparisons.

(f) Discuss the issue of multiple comparisons and Type I error inflation.

(g) State conclusions about which methods differ significantly at $= 0.05$.

Section I: Power and Effect Size [20 marks]

19. **[10 marks]** Explain statistical power and effect size:

(a) Define statistical power and state its relationship to Type II error.

(b) List four factors that affect the power of a statistical test.

(c) Explain what effect size measures and why it's important.

(d) Define Cohen's d for measuring effect size in t-tests.

(e) Explain the trade-off between significance level and power.

20. **[10 marks]** A researcher wants to detect a difference of 5 points in mean test scores ($= 12$) between two groups.

(a) Calculate Cohen's d for this effect size.

(b) Classify this effect size as small, medium, or large.

(c) Explain how increasing the sample size would affect the power.

(d) Describe how changing from 0.05 to 0.01 would affect power.

(e) Suggest ways to increase power without changing the significance level.

Section J: Comprehensive Applications [30 marks]

21. **[15 marks]** A quality control manager monitors production lines. Historical data shows 3% defect rate. Recent inspection of 500 items found 20 defective.

(a) Test if the defect rate has changed using $= 0.05$.

(b) Calculate the exact p-value for this test.

(c) Use the normal approximation to verify your result.

(d) Calculate a 95% confidence interval for the true defect rate.

(e) Interpret the relationship between the confidence interval and hypothesis test.

(f) Determine the sample size needed to detect a change to 4% defect rate with 80% power.

22. **[15 marks]** A marketing team studies customer satisfaction across three store locations:

	Satisfied	Neutral	Dissatisfied
Store A	45	15	10
Store B	38	22	15
Store C	52	12	8

Additionally, satisfaction scores (1-10 scale) from a random sample: Store A: 8.2, 7.5, 8.8, 7.9, 8.1, 7.6, 8.3, 7.8 Store B: 7.1, 6.8, 7.5, 6.9, 7.2, 6.7, 7.4, 7.0

(a) Test if satisfaction levels are independent of store location using the categorical data.

(b) Calculate standardized residuals to identify which stores/categories deviate most.

(c) Test if the mean satisfaction scores differ between Stores A and B.

(d) Calculate Cohen's d for the difference between Stores A and B.

(e) If Store C has a target mean satisfaction score of 8.0, test if it meets this target.

(f) Discuss the advantages and limitations of each approach to analyzing satisfaction.

Answer Space

Use this space for your working and answers.

Formulae and Key Concepts

Test Statistics:

One-sample t-test: $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ with df = n-1

Two-sample t-test: $t = \frac{\bar{x_1}-\bar{x_2}}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ with df $= n_1 + n_2 - 2$

Paired t-test: $t = \frac{\bar{d}}{s_d/\sqrt{n}}$ with df = n-1

One-sample z-test: $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$

Chi-Squared Tests:

Test statistic: $\chi^2 = \sum \frac{(O_i-E_i)^2}{E_i}$

Goodness of fit df = categories - 1 - parameters estimated

Independence test df = (rows - 1)(columns - 1)

Expected frequency: $E_{ij} = \frac{R_i \times C_j}{n}$

Pooled Standard Deviation:

$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$

Effect Size:

Cohen's d: $d = \frac{|\bar{x_1}-\bar{x_2}|}{s_p}$

Small: d = 0.2, Medium: d = 0.5, Large: d = 0.8

Error Types:

Type I Error: Reject H_0 when H_0 is true, P(Type I) =

Type II Error: Fail to reject H_0 when H_0 is false, P(Type II) =

Power = 1 -

Critical Values (= 0.05):

z: ± 1.96 (two-tailed), 1.645 (one-tailed)

t varies with degrees of freedom

χ^2 varies with degrees of freedom

Confidence Intervals:

Mean: $\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

Proportion: $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

END OF TEST

Total marks: 305

For more resources and practice materials, visit:
stepupmaths.co.uk