NILE TUITION ACADEMY
Excellence in Further Maths Tuition
niletuitionacademy.co.uk

A Level Further Maths

Topic 4: Statistics (Optional)

AQA Further Maths (7367)	Mark Scheme	100 marks

Instructions

- Answer **all** questions. You must show sufficient working - answers with no working may score no marks.
- You may use a calculator unless a question states otherwise.
- Give exact answers where required, otherwise to an appropriate degree of accuracy.
- There are 11 questions in this paper. The total mark is **100**.

1. A discrete random variable X has probability distribution given by P(X = x) = k x for x = 1, 2, 3, 4, where k is a constant.

(a) Find the value of k. **(3)**

✓ Σ P(X = x) = 1, so k(1 + 2 + 3 + 4) = 1 (1)
✓ 10k = 1 (1)
✓ k = 1/10 = 0.1 (1)

(b) Find E(X). **(3)**

✓ E(X) = Σ x P(X = x) = 0.1(1×1 + 2×2 + 3×3 + 4×4) (1)
✓ = 0.1(1 + 4 + 9 + 16) = 0.1 × 30 (1)
✓ = 3 (1)

2. The number of emails arriving at an office in a randomly chosen one-minute interval is modelled by a Poisson distribution with mean 3.

(a) Find the probability that exactly 2 emails arrive in a one-minute interval. Give your answer to 4 decimal places. **(2)**

✓ P(X = 2) = e^{-3} 3^{2} / 2! (1)
✓ = 0.2240 (1)

(b) Find the probability that more than 4 emails arrive in a one-minute interval. **(3)**

✓ P(X > 4) = 1 - P(X ≤ 4) (1)
✓ P(X ≤ 4) = 0.8153 (from tables / calculator for mean 3) (1)
✓ P(X > 4) = 1 - 0.8153 = 0.1847 (1)

3. A continuous random variable X has probability density function $f(x) = k x^2$ for $0 \le x \le 3$, and $f(x) = 0$ otherwise.

(a) Show that $k = 1/9$. **(3)**

✓ $\int_0^3 k x^2 \, dx = 1$ (1)
✓ $[k x^3/3]_0^3 = k(27/3) = 9k$ (1)
✓ $9k = 1$ so $k = 1/9$ as required (1)

(b) Find $P(1 \le X \le 2)$. **(3)**

✓ $P(1 \le X \le 2) = \int_1^2 (1/9) x^2 \, dx$ (1)
✓ $= (1/9)[x^3/3]_1^2 = (1/27)(8 - 1)$ (1)
✓ $= 7/27 = 0.259$ (3 s.f.) (1)

4. A discrete random variable Y has the probability distribution shown: $P(Y = 0) = 0.2$, $P(Y = 1) = 0.5$, $P(Y = 2) = 0.3$.

(a) Find $E(Y)$ and $E(Y^2)$. **(4)**

✓ $E(Y) = 0(0.2) + 1(0.5) + 2(0.3) = 1.1$ (1)
✓ correct method for $E(Y^2) = \Sigma y^2 P(Y = y)$ (1)
✓ $E(Y^2) = 0(0.2) + 1(0.5) + 4(0.3)$ (1)
✓ $= 0.5 + 1.2 = 1.7$ (1)

(b) Hence find $Var(Y)$. **(3)**

✓ $Var(Y) = E(Y^2) - [E(Y)]^2$ (1)
✓ $= 1.7 - 1.1^2 = 1.7 - 1.21$ (1)
✓ $= 0.49$ (1)

5. At a call centre, calls of type A arrive independently as a Poisson process at a mean rate of 2 per hour, and calls of type B arrive independently as a Poisson process at a mean rate of 5 per hour.

(a) State the distribution of the total number of calls (of either type) arriving in one hour. **(2)**

✓ the sum of independent Poisson variables is Poisson (1)
✓ total ~ Poisson with mean $2 + 5 = 7$ (1)

(b) Find the probability that exactly 6 calls of either type arrive in one hour. **(3)**

✓ $P(X = 6)$ where $X \sim \text{Poisson}(7)$ (1)
✓ $= e^{-7} 7^6 / 6!$ (1)
✓ $= 0.1490$ (4 d.p.) (1)

(c) State two conditions required for the number of type A calls to be well modelled by a Poisson distribution. **(3)**

✓ calls occur singly / one at a time (1)
✓ calls occur at a constant average rate (1)
✓ calls occur independently of one another (1)

6. The continuous random variable X has probability density function $f(x) = (3/8)\,x^2$ for $0 \le x \le 2$, and $f(x) = 0$ otherwise.

(a) Find E(X). **(4)**

✓ $E(X) = \int_0^2 x\,f(x)\,dx = \int_0^2 (3/8)\,x^3\,dx$ (1)
✓ $= (3/8)[\,x^4/4\,]_0^2$ (1)
✓ $= (3/8)(16/4) = (3/8)(4)$ (1)
✓ $= 3/2 = 1.5$ (1)

(b) Find the median value of X, giving your answer to 3 significant figures. **(4)**

✓ median m satisfies $\int_0^m (3/8)\,x^2\,dx = 0.5$ (1)
✓ $(1/8)\,m^3 = 0.5$ (1)
✓ $m^3 = 4$ (1)
✓ $m = 4^{1/3} = 1.59$ (3 s.f.) (1)

7. The number of flaws per metre in a long roll of fabric is modelled by a Poisson distribution. Historically the mean number of flaws is 1.5 per metre. After a change to the manufacturing process, a manager wishes to test, at the 5% significance level, whether the mean number of flaws has decreased. In a randomly chosen one-metre length, 0 flaws are observed.

(a) State suitable null and alternative hypotheses for this test, defining any parameter you use. **(3)**

✓ let λ be the mean number of flaws per metre (1)
✓ $H_0: \lambda = 1.5$ (1)
✓ $H_1: \lambda < 1.5$ (one-tailed) (1)

(b) Carry out the test, showing your calculation of the relevant probability. **(4)**

✓ under H_0, $X \sim \text{Poisson}(1.5)$; find $P(X \le 0)$ (1)
✓ $P(X = 0) = e^{-1.5} = 0.2231$ (1)
✓ compare with 0.05: $0.2231 > 0.05$ (1)
✓ result is not significant, so do not reject H_0 (1)

(c) State the conclusion of the test in the context of the question. **(3)**

✓ there is insufficient evidence at the 5% level (1)
✓ to conclude that the mean number of flaws per metre has decreased (1)
✓ the observed value is consistent with a mean of 1.5 flaws per metre (1)

8. The mass, in grams, of a randomly chosen apple from a large orchard is modelled by a normal distribution with unknown mean μ and known standard deviation 8 g. A random sample of 16 apples has mean mass 152 g.

(a) Find a 95% confidence interval for μ. Give the endpoints to 1 decimal place. **(4)**

✓ confidence interval is x-bar $\pm z \times \sigma/\sqrt{n}$ with $z = 1.96$ (1)
✓ standard error $= 8/\sqrt{16} = 2$ (1)
✓ $152 \pm 1.96 \times 2 = 152 \pm 3.92$ (1)
✓ (148.1, 155.9) (1)

(b) Find a 99% confidence interval for μ, using $z = 2.576$. Give the endpoints to 1 decimal place. **(4)**

✓ 99% interval uses $z = 2.576$ (1)
✓ margin $= 2.576 \times 2 = 5.152$ (1)
✓ 152 ± 5.152 (1)
✓ (146.8, 157.2) (1)

(c) Explain why the 99% confidence interval is wider than the 95% confidence interval. **(3)**

✓ a higher confidence level requires a larger critical value z (1)
✓ $2.576 > 1.96$, so the margin of error is larger (1)
✓ a wider interval is needed to be more confident of containing μ (1)

9. A six-sided die is rolled 120 times. The observed frequencies of each score are: 1: 14, 2: 26, 3: 18, 4: 22, 5: 16, 6: 24. A chi-squared goodness-of-fit test is to be used, at the 5% significance level, to test whether the die is fair.

(a) State the null and alternative hypotheses, and find the expected frequency for each score under the null hypothesis. **(3)**

✓ H_0: the die is fair (each score equally likely); H_1: the die is not fair (1)
✓ under H_0 each probability = 1/6 (1)
✓ expected frequency $= 120 \times 1/6 = 20$ for each score (1)

(b) Calculate the value of the test statistic $\Sigma\, (O - E)^2 / E$. **(4)**

✓ compute $(O - E)^2/E$ for each: $(14-20)^2/20 = 1.8$, $(26-20)^2/20 = 1.8$ (1)
✓ $(18-20)^2/20 = 0.2$, $(22-20)^2/20 = 0.2$ (1)
✓ $(16-20)^2/20 = 0.8$, $(24-20)^2/20 = 0.8$ (1)
✓ $\Sigma = 1.8 + 1.8 + 0.2 + 0.2 + 0.8 + 0.8 = 5.6$ (1)

(c) Using 5 degrees of freedom and a critical value of 11.07, complete the test and state your conclusion in context. **(4)**

✓ degrees of freedom $= 6 - 1 = 5$ (1)
✓ compare 5.6 with the critical value 11.07 (1)
✓ $5.6 < 11.07$, so do not reject H_0 (1)
✓ there is insufficient evidence at the 5% level that the die is unfair (1)

(d) State the distribution that the test statistic approximately follows under the null hypothesis. **(2)**

✓ the chi-squared distribution (1)
✓ with 5 degrees of freedom (1)

10. A survey records whether 200 people prefer tea or coffee, classified by age group. The observed frequencies are: under 30 - tea 30, coffee 50; 30 or over - tea 60, coffee 60. A chi-squared test for independence is carried out at the 5% level to test whether drink preference is independent of age group.

(a) State the null and alternative hypotheses for the test. **(3)**

✓ H_0: drink preference and age group are independent (1)
✓ H_1: drink preference and age group are not independent (1)
✓ i.e. there is an association between drink preference and age group (1)

(b) Show that the expected frequency for people aged under 30 who prefer tea is 36. **(4)**

✓ row total (under 30) = 80; column total (tea) = 90; grand total = 200 (1)
✓ expected = (row total × column total) / grand total (1)
✓ = (80 × 90) / 200 (1)
✓ = 36 as required (1)

(c) The value of the test statistic is found to be 2.78. The test has 1 degree of freedom. Using a critical value of 3.841, state the conclusion of the test in context. **(3)**

✓ degrees of freedom = (2 - 1)(2 - 1) = 1, consistent with the given value (1)
✓ 2.78 < 3.841, so do not reject H_0 (1)
✓ there is insufficient evidence of an association between drink preference and age group (1)

(d) State why a continuity correction is sometimes applied to a 2 by 2 contingency table, and give the corrected form of one term of the test statistic. **(3)**

✓ with 1 degree of freedom the chi-squared approximation can be improved (Yates's correction) (1)
✓ each term becomes $(|O - E| - 0.5)^2 / E$ (1)
✓ this reduces the size of the statistic, lowering the chance of a Type I error / over-rejection (1)

11. The continuous random variable X has probability density function $f(x) = (1/4)(3 - x)$ for $1 \le x \le 3$, and $f(x) = 0$ otherwise. (You may assume f is a valid probability density function.)

(a) Find the cumulative distribution function F(x) for $1 \le x \le 3$. **(4)**

✓ $F(x) = \int_1^x (1/4)(3 - t)\,dt$ (1)
✓ $= (1/4)[\,3t - t^2/2\,]_1^x$ (1)
✓ $= (1/4)[\,(3x - x^2/2) - (3 - 1/2)\,]$ (1)
✓ $F(x) = (1/4)(3x - x^2/2 - 5/2) = (6x - x^2 - 5)/8$ (1)

(b) Find E(X). **(4)**

✓ $E(X) = \int_1^3 x\,(1/4)(3 - x)\,dx = (1/4)\int_1^3 (3x - x^2)\,dx$ (1)
✓ $= (1/4)[\,(3/2)x^2 - x^3/3\,]_1^3$ (1)
✓ at x = 3: (13.5 - 9) = 4.5; at x = 1: (1.5 - 1/3) = 7/6 (1)
✓ E(X) = (1/4)(4.5 - 7/6) = (1/4)(10/3) = 5/6 = 0.833 (3 s.f.) (1)

(c) Find $E(X^2)$ and hence Var(X). **(5)**

✓ $E(X^2) = (1/4)\int_1^3 (3x^2 - x^3)\,dx$ (1)
✓ $= (1/4)[\,x^3 - x^4/4\,]_1^3 = (1/4)[\,(27 - 81/4) - (1 - 1/4)\,]$ (1)
✓ = (1/4)[27/4 - 3/4] = (1/4)(24/4) = (1/4)(6) = 3/2 (1)
✓ $Var(X) = E(X^2) - [E(X)]^2 = 3/2 - (5/6)^2$ (1)
✓ = 3/2 - 25/36 = 54/36 - 25/36 = 29/36 = 0.806 (3 s.f.) (1)