Maths with Matt

Physics and maths resources

Edexcel A Level Further Mathematics

9FM0/4B - Further Statistics 2

Monday 23 June 2025

Questions and Fully Worked Model Solutions

This booklet places the **question and the solution together**. The mathematics is set out using proper fractions, powers, roots and exact notation, with extra written explanation so that a student can see *why* each step is being taken.

Solutions have been checked against the Pearson Edexcel Summer 2025 mark scheme. The aim is a clear **full-mark model response**, not merely a list of final answers.

Question wording is reproduced from the user-provided 2025 paper for study/reference. Pearson retains copyright in the original examination material.

Contents

Question 1 2

Question 2 4

Question 3 5

Question 4 6

Question 5 8

Question 6 10

Question 7 12

www.mathswithmatt.co.uk

Question 1

Question 1

A sample of 15 newborn babies gives $\sum x = 50.46$, $\sum y = 518.9$, $\sum x^2 = 171.828$, $\sum y^2 = 18004.47$ and $S_{xy} = 7.8284$, where x is birth weight (kg) and y is head circumference (cm). Find the requested regression and residual quantities.

Model solution

(a) The corrected sums of squares measure variation about the sample means, so use $S_{xx} = \sum x^2 - (\sum x)^2/n$, and similarly for y.

$$S_{xx} = 171.828 - \frac{50.46^2}{15} = 2.08056\ldots = 2.081, \qquad S_{yy} = 18004.47 - \frac{518.9^2}{15} = 53.98933\ldots = 53.989.$$

$$S_{xx} = 2.081 \text{ and } S_{yy} = 53.989 \text{ (3d.p.)}.$$

(3)

(b) For the regression of y on x, use $b = S_{xy}/S_{xx}$; then choose the intercept so that the line passes through $(\bar{x}, \bar{y})$.

$$b = \frac{7.8284}{2.08056\ldots} = 3.762\ldots, \qquad a = \bar{y} - b\bar{x} = \frac{518.9}{15} - 3.762\ldots\frac{50.46}{15} = 21.914\ldots$$

$$\hat{y} = 21.9 + 3.76x.$$

(3)

(c) The unexplained variation after fitting the regression line is the residual sum of squares.

$$\text{RSS} = S_{yy} - \frac{S_{xy}^2}{S_{xx}} = 53.98933\ldots - \frac{7.8284^2}{2.08056\ldots} = 24.5397\ldots$$

$$\text{RSS} = 24.5 \text{ (approximately)}.$$

(1)

(d) A residual is observed y minus predicted $\hat{y}$. First predict the head circumference for $x = 3.26$.

$$\hat{y} = 21.914\ldots + 3.762\ldots(3.26) = 34.202\ldots, \qquad e = 36.8 - 34.202\ldots = 2.598\ldots$$

$$\text{Residual} = 2.60\text{cm (approximately)}.$$

(2)

(e) Compare this one squared residual with the RSS for all 15 babies.

$$\frac{(2.598\ldots)^2}{24.5397\ldots} = 0.275\ldots$$

This one baby contributes about 27.5% of the total RSS, an unusually large share, s

(1)

www.mathswithmatt.co.uk

Question 2

Question 2

Two lecturers give marks to nine essays. Lecturer 1 gives $10, 16, 5, 17, 12, 7, 14, 11, 8$, and lecturer 2 gives $17, 22, 19, 24, 21, 15, 23, 18, 20$. Calculate Spearman's rank correlation coefficient and test, at 5%, whether the lecturers agree about rank order.

Model solution

(a) Spearman's coefficient is appropriate because agreement in rank order, rather than agreement in the numerical marks, is being assessed. Rank 1 is assigned to the smallest mark.

Student	A	B	C	D	E	F	G	H	I
R_1	6	2	9	1	4	8	3	5	7
R_2	8	3	6	1	4	9	2	7	5
$d = R_1 - R_2$	-2	-1	3	0	0	-1	1	-2	2

$$\sum d^2 = 4+1+9+0+0+1+1+4+4 = 24, \qquad r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(24)}{9(9^2-1)} = 0.8.$$

$$r_s = 0.800$$

(4)

(b) The claim is positive agreement, so use a one-tailed Spearman test.

$$H_0 : \rho_s = 0, \qquad H_1 : \rho_s > 0.$$

For $n = 9$, the 5% one-tailed critical value is 0.600. Since $0.800 > 0.600$, the statistic is in the critical region.

Reject H_0. There is significant evidence that the lecturers agree about the rank orde

(3)

Question 3

Question 3

Independent normal samples of cereal yields have summaries: without bacteria, $n = 6$, $\bar{x} = 1648$, $s = 143.8$; with bacteria, $n = 5$, $\bar{x} = 1910$, $s = 64.0$. Test equality of variances at 10%. A suitable test for increased mean yield has $p = 0.002$.

Model solution

(a) The samples are normal and independent, so an F-test compares their variances. Equality is two-tailed; at a 10% level, use 5% in each tail. Put the larger sample variance on top.

$$H_0 : \sigma^2_{\text{without}} = \sigma^2_{\text{with}}, \qquad H_1 : \sigma^2_{\text{without}} \neq \sigma^2_{\text{with}},$$

$$F = \frac{143.8^2}{64.0^2} = 5.0484\ldots$$

The upper critical value is $F_{5,4}(0.05) = 6.26$. Since $5.0484\ldots < 6.26$, this is not significant.

Do not reject H_0 : there is insufficient evidence of a difference in the yield variances.

(4)

(b)(i) The two-sample test for means being used requires the populations to have equal variances, so the result in part (a) permits that assumption (and hence pooling of variance) to be used.

The equal-variance condition required for the test of means is reasonable.

(1)

(b)(ii) The p-value 0.002 is far below conventional significance levels, so there is strong evidence for an increase in mean yield. The variance test found no evidence that variability changed.

Nitrogen-fixing bacteria increase mean cereal yield, without evidence of changing the variance.

(1)

Question 4

Question 4

Untreated apple weights are $N(131, 6^2)$ grams; the supermarket accepts weights from 125 g to 137 g. For 25 treated apples, $\sum x = 3250$ and $\sum x^2 = 422862$. Investigate the treatment and estimate the accepted proportion.

Model solution

(a) Standardise the two endpoints about the untreated mean. They are one standard deviation either side of 131 g.

$$P(125 < X < 137) = P(-1 < Z < 1) = 0.68268\ldots$$

$$0.683, \text{or} 68.3\%.$$

(1)

(b) The aim is a reduction in variability, so this is a lower-tailed chi-squared test for the variance. First calculate the unbiased sample variance.

$$H_0 : \sigma^2 = 6^2 = 36, \qquad H_1 : \sigma^2 < 36,$$

$$s^2 = \frac{422862 - 3250^2/25}{25 - 1} = 15.0833\ldots, \qquad \chi^2 = \frac{(25-1)s^2}{36} = 10.0556\ldots$$

With 24 degrees of freedom, the 5% lower-tail critical value is 13.848*. Since* $10.056 < 13.848$*, the statistic is in the lower critical region.*

$$\text{Reject} H_0. \text{There is evidence that treatment has reduced the variance of apple weigh}$$

(6)

(c) The population variance is unknown, so use a t*-interval with* 24 *degrees of freedom.*

$$\bar{x} \pm t_{24,0.975}\frac{s}{\sqrt{n}} = 130 \pm 2.064\frac{\sqrt{15.0833\ldots}}{\sqrt{25}} = 130 \pm 1.603\ldots$$

$$\text{A 95\% confidence interval is } (128.4, 131.6) \text{ g}.$$

(3)

(d)(i) The original mean, 131 g, lies in the confidence interval, so the data do not indicate that the mean has changed significantly. Part (b) found a significant reduction in variance.

Yes: the mean is consistent with 131 g and the variance has decreased.

(1)

(d)(ii) Estimate the treated distribution by $N(131, 3.884^2)$*, using the target mean and the sample standard deviation.*

$$P(125 < X < 137) = P\left(\frac{-6}{3.884} < Z < \frac{6}{3.884}\right) = P(-1.545 < Z < 1.545) = 0.876\ldots$$

$$\text{Estimated proportion} = 0.876, \text{ or } 87.6\%.$$

(2)

www.mathswithmatt.co.uk

Question 5

Question 5

$X \sim U[1,4]$. A sample of ten observations has maximum M, with $F_M(y) = ((y-1)/3)^{10}$ for $1 \leqslant y \leqslant 4$. Find the stated probabilities, the density and mode of M, and $E(M)$.

Model solution

(a)(i) A uniform probability is interval length divided by the full width, which is $4-1=3$.

$$P(1.8 < X < 3.2) = \frac{3.2-1.8}{4-1} = \frac{7}{15}.$$

$$\frac{7}{15}.$$

(1)

(a)(ii) Given $X > 1.8$, *restrict attention to the interval* $(1.8, 4)$. *The favourable part is* $(3.2, 4)$.

$$P(X > 3.2 \mid X > 1.8) = \frac{P(X > 3.2)}{P(X > 1.8)} = \frac{0.8/3}{2.2/3} = \frac{4}{11}.$$

$$\frac{4}{11}.$$

(2)

(b) Use the CDF of the maximum: $P(M > 3.75) = 1 - F_M(3.75)$.

$$1 - \left(\frac{3.75-1}{3}\right)^{10} = 0.581096\ldots$$

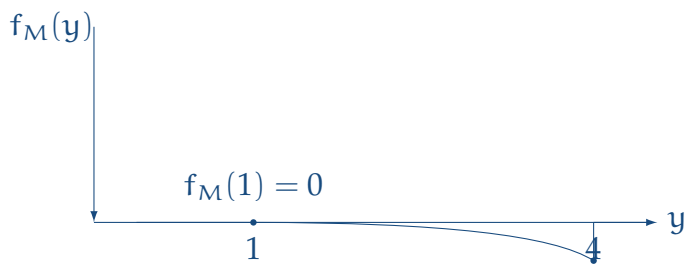
$$0.581 (\text{approximately}).$$

(1)

(c)(i) Differentiate the CDF on its support to obtain the density.

$$f_M(y) = \frac{d}{dy}\left(\frac{y-1}{3}\right)^{10} = \frac{10}{3}\left(\frac{y-1}{3}\right)^{9}, \qquad 1 \leqslant y \leqslant 4.$$

The required sketch has an increasing curve from $(1, 0)$ *to* $(4, 10/3)$, *with* $f_M(y) = 0$ *outside* $[1, 4]$.



$$f_M(y) = \frac{10}{3}\left(\frac{y-1}{3}\right)^9 \text{ for } 1 \leqslant y \leqslant 4, \text{ drawn as the stated increasing curve.}$$

(2)

(c)(ii) Since the density increases throughout the support, its greatest value is at the right endpoint.

$$\text{mode}(M) = 4.$$

(1)

(d) Use $E(M) = \int y f_M(y)\,dy$. *The substitution* $u = (y-1)/3$ *makes the power simple:* $y = 3u + 1$, $dy = 3\,du$.

$$E(M) = \int_1^4 y \frac{10}{3}\left(\frac{y-1}{3}\right)^9 dy = 10\int_0^1 (3u+1)u^9 du$$

$$= 10\left[\frac{3u^{11}}{11} + \frac{u^{10}}{10}\right]_0^1 = \frac{30}{11} + 1 = \frac{41}{11}.$$

$$E(M) = \tfrac{41}{11}.$$

(4)

Question 6

Question 6

Delivery rates (minutes per km) have summaries for A: $n_A = 61$, $\bar{x}_A = 3.7$, $s_A^2 = 1.64$; and for B: $n_B = 51$, $\sum x_B = 142.8$, $\sum x_B^2 = 461.34$. A paired sample of ten customers gives A rates $5.4, 3.6, 2.5, 4.7, 2.3, 3.2, 3.4, 4.1, 5.9, 4.5$ and B rates $4.9, 3.9, 2.3, 4.5, 2.6, 2.8, 3.1, 3.8, 5.2, 4.3$.

Model solution

(a) The unbiased mean estimate is the sample mean, and the unbiased variance divides the corrected sum of squares by $n - 1$.

$$\bar{x}_B = \frac{142.8}{51} = 2.8, \qquad s_B^2 = \frac{461.34 - 142.8^2/51}{50} = 1.2264\ldots$$

$$\bar{x}_B = 2.8 \text{minutesperkmands}_B^2 = 1.23 (\textit{minutes per km})^2.$$

(3)

(b) A pooled two-sample t*-test requires independent samples from normal populations with equal population variances.*

Assume normally distributed delivery rates with equal variances.

(1)

(c) This is a two-tailed pooled two-sample t*-test because the question asks whether the two mean rates differ.*

$$H_0 : \mu_A = \mu_B, \qquad H_1 : \mu_A \neq \mu_B,$$

$$s_p^2 = \frac{60(1.64) + 50(1.2264\ldots)}{61 + 51 - 2} = 1.4536\ldots,$$

$$t = \frac{3.7 - 2.8}{\sqrt{1.4536\ldots}\sqrt{1/61 + 1/51}} = 3.934\ldots$$

There are $61 + 51 - 2 = 110$ *degrees of freedom. The 5% two-tailed critical values are* ± 1.982. *Since* $|3.934| > 1.982$, *reject* H_0.

There is significant evidence of a difference in mean delivery rates.

(7)

(d) The two observations for each customer form a pair, so analyse differences to remove customer-to-customer variation. Let $D = A - B$.

$$H_0 : \mu_D = 0, \qquad H_1 : \mu_D \neq 0,$$

$$d = (0.5, -0.3, 0.2, 0.2, -0.3, 0.4, 0.3, 0.3, 0.7, 0.2),$$

$$\bar{d} = 0.22, \text{qquads}_D = 0.31552\ldots, \qquad t = \frac{0.22 - 0}{0.31552\ldots/\sqrt{10}} = 2.205\ldots$$

With 9 degrees of freedom, the 5% two-tailed critical values are ± 2.262. Since $|2.205| < 2.262$, the result is not significant.

Do not reject H_0 : there is insufficient evidence of a difference in delivery rates for comparable custo

(6)

www.mathswithmatt.co.uk

Question 7

Question 7

$X_1, \ldots, X_5$ is a random sample from $N(\mu, \sigma^2)$. Nina uses $\overline{X}$ and Jonah uses $Y = (X_1 + X_5)/2$ to estimate μ. Their 95% confidence intervals are $(21.823, 38.177)$ and $(27.029, 37.371)$.

Model solution

(a) Let $D = \overline{X} - Y$. Both estimators are unbiased, so $E(D) = 0$. Write D in terms of independent sample observations before finding its variance.

$$D = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5} - \frac{X_1 + X_5}{2} = \frac{-3X_1 + 2X_2 + 2X_3 + 2X_4 - 3X_5}{10}.$$

$$\mathrm{Var}(D) = \frac{(9 + 4 + 4 + 4 + 9)\sigma^2}{100} = \frac{3}{10}\sigma^2.$$

Thus $D \sim N(0, 3\sigma^2/10)$, so standardisation gives

$$P(\overline{X} - Y > \sigma) = P\left(Z > \frac{\sigma}{\sigma\sqrt{3/10}}\right) = P\left(Z > \sqrt{\frac{10}{3}}\right) = 0.03394\ldots$$

$$P(\overline{X} - Y > \sigma) = 0.034 (\text{approximately}).$$

(7)

(b) $\overline{X}$ uses all five observations, whereas Y uses only two, so $\overline{X}$ has the smaller standard error and therefore the narrower confidence interval.

$$\text{Nina calculated } (27.029, 37.371), \text{ because it is the narrower interval.}$$

(1)

(c) Use Nina's interval. Its half-width is $1.96\sigma/\sqrt{5}$, because $\overline{X}$ has variance $\sigma^2/5$.

$$\frac{37.371 - 27.029}{2} = 1.96\frac{\sigma}{\sqrt{5}}, \qquad \sigma = \frac{10.342\sqrt{5}}{2(1.96)} = 5.899\ldots$$

$$\sigma = 5.9 (\text{approximately}).$$

(3)