

MATH40005

Probability and Statistics

Revised Mock Examination 1 - Solutions and Marking Guide

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers
Difficulty calibration	A 32, B 20, C 12, D 16
Paper role	2026 structural reinforcement: one deliberately close paper, but with joint-density and position-counting variants rather than a line-by-line clone.

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

1(a) (3 marks)

Recognition signal	A finite experiment with equally likely ordered outcomes asks for all three components of a probability space.
First key step	Write the ordered sample space first; replacement means all 8^2 pairs are possible.
Historical basis	2025 Q1(a)(i); 2026 Q1(a)
Ability/method	finite probability-space construction; K01

Course-method solution. Take $\Omega = \{1, \dots, 8\}^2$, $F = \mathcal{P}(\Omega)$, and $P(\{(i, j)\}) = 1/64$. Hence $P(A) = |A|/64$ for every A subset Ω .

Marking points. 1 mark: ordered sample space; 1 mark: event space; 1 mark: probability measure

Common errors. Using unordered pairs; omitting F or P .

Final self-check. Check $|\Omega| = 64$ and the singleton probabilities sum to 1.

Difficulty allocation: A:3; estimated normal time 4 min.

1(b)(i) (5 marks)

Recognition signal	Fixed counts of several character types with repetitions and distinct symbols.
First key step	Separate position choices from character choices.
Historical basis	2025 Q2(a)(i)
Ability/method	multistage multiplication and repeated-element counting; K05

Course-method solution. Choose the two symbol positions, then five of the remaining positions for letters. Order the two distinct symbols. The count is $C(10, 2) C(8, 5) (5 \cdot 4) 52^5 10^3$.

Marking points. 2 marks: choose type positions; 1 mark: order the distinct symbols; 2 marks: character choices and product

Common errors. Using $C(5, 2)$ without ordering the two symbols; forgetting repetitions.

Final self-check. The multinomial form $10!/(5!3!2!)$ times $52^5 10^3$ times $C(5, 2)2!$ must agree.

Difficulty allocation: A:2, B:3; estimated normal time 7 min.

1(b)(ii) (4 marks)

Recognition signal	The random feature depends only on positions, not symbol identities.
First key step	Count adjacent unordered pairs among all unordered position pairs.
Historical basis	2025 Q2(a)(ii)
Ability/method	position-pattern probability; K05

Course-method solution. There are $C(10, 2)$ possible pairs of symbol positions and 9 adjacent pairs: $(1, 2), \dots, (9, 10)$. Therefore the probability is $9/C(10, 2) = 1/5$.

Marking points. 1 mark: denominator; 2 marks: adjacent-pair count; 1 mark: probability

Common errors. Dividing by $10!$ or by the total number of passwords.

Final self-check. The answer must lie between 0 and 1 and should be close to $2/10$.

Difficulty allocation: B:1, C:1, D:2; estimated normal time 6 min.

1(c) (5 marks)

Recognition signal	A shifted lattice distribution requires a floor function in the all-real CDF.
First key step	Re-index with $k = x - r$.
Historical basis	2026 Q1(c); 2020 Q2(c)(i)
Ability/method	support-aware finite/infinite geometric sums; K11

Course-method solution. Non-negativity is immediate. The sum is $p \sum_{k=0}^{\infty} (1-p)^k = 1$. For $x < r$, $F_X(x) = 0$. For $x \geq r$, $F_X(x) = \sum_{k=0}^{\lfloor x \rfloor - r} p(1-p)^k = 1 - (1-p)^{\lfloor x \rfloor - r + 1}$.

Marking points. 1 mark: non-negativity; 1 mark: total mass; 2 marks: finite geometric sum; 1 mark: all-real piecewise CDF

Common errors. Giving the CDF only at integer x ; missing $x < r$.

Final self-check. At $x=r$ the CDF should equal p , and it should tend to 1.

Difficulty allocation: A:2, C:3; estimated normal time 7 min.

1(d) (3 marks)

Recognition signal	A count of independent yes/no events is a sum of indicators.
First key step	Define success as failure on the specified day.
Historical basis	2026 Q2(c)(i)
Ability/method	indicator-to-binomial modelling; K09

Course-method solution. $Y_i \sim \text{Bernoulli}(1/d)$. Since the indicators are independent, $B \sim \text{Binomial}(n, 1/d)$.

Marking points. 1 mark: Bernoulli model; 1 mark: binomial model; 1 mark: parameters and independence

Common errors. Calling B Poisson without an approximation statement.

Final self-check. $E[B] = n/d$ is consistent with the binomial model.

Difficulty allocation: A:2, B:1; estimated normal time 3 min.

Total: 20 marks

Question 2

2(a) (3 marks)

Recognition signal	A joint density with an unknown constant on a triangular support.
First key step	Sketch/read the support and normalize.
Historical basis	2025 Q2(b)(i); 2023 Q3(b)(i)
Ability/method	joint-density normalization on non-rectangular support; K17,K22

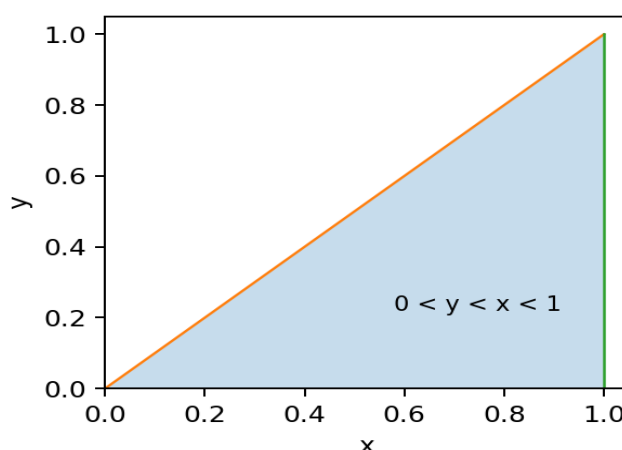
Course-method solution. Integrate over $0 < x < 1$ and $0 < y < x$: $\int_0^1 \int_0^x c(x+y) dy dx = c/2$. Hence $c=2$.

Marking points. 1 mark: correct region/bounds; 1 mark: integral; 1 mark: $c=2$ and non-negativity

Common errors. Integrating over the unit square.

Final self-check. Substitution $c=2$ must make the integral 1.

Difficulty allocation: A:2, B:1; estimated normal time 5 min.



2(b) (4 marks)

Recognition signal	Marginals require integrating along the support slice determined by the other variable.
First key step	For fixed x , y runs 0 to x ; for fixed y , x runs y to 1.
Historical basis	2025 Q2(b)(ii); 2023 Q3(b)(ii)
Ability/method	marginalization on triangular support; K18

Course-method solution. For $0 < x < 1$, $f_X(x) = \int_0^x 2(x+y) dy = 3x^2$. For $0 < y < 1$, $f_Y(y) = \int_y^1 2(x+y) dx = 1 + 2y - 3y^2$. Both are zero otherwise.

Marking points. 2 marks: f_X with support; 2 marks: f_Y with support

Common errors. Reversing the bounds or omitting supports.

Final self-check. Each marginal integrates to 1.

Difficulty allocation: A:2, B:2; estimated normal time 6 min.

2(c) (3 marks)

Recognition signal	The phrase “given $X=x$ ” in a continuous model signals a density ratio.
First key step	Compute f_X first, then divide.
Historical basis	2025 Q2(b)(iii); 2022 Q3(c)(iii)
Ability/method	conditional-density ratio and support; K18

Course-method solution. For $0 < x < 1$ and $0 < y < x$, $f_{\{Y|X=x\}}(y|x) = f_{XY}(x,y)/f_X(x) = 2(x+y)/(3x^2)$. It is defined for x with $f_X(x) > 0$, namely $0 < x < 1$.

Marking points. 1 mark: ratio formula; 1 mark: expression; 1 mark: support/domain

Common errors. Treating $P(X=x) > 0$; failing to state $0 < y < x$.

Final self-check. Integrating the conditional density from 0 to x gives 1.

Difficulty allocation: A:1, B:2; estimated normal time 5 min.

2(d) (2 marks)

Recognition signal	Independence must hold on the full support, not only where the joint density is positive.
First key step	Use support or check factorization.
Historical basis	2025 Q2(b)(iv); 2023 Q3(b)(iii)
Ability/method	support/factorization test for independence; K19

Course-method solution. No. The support is triangular, not a Cartesian product: for example $f_{XY}(1/4, 1/2) = 0$ while both marginals are positive. Equivalently, the conditional density depends on x .

Marking points. 1 mark: correct conclusion; 1 mark: decisive support/factorization argument

Common errors. Saying “not independent because x and y appear together” without proof.

Final self-check. Find one point where $f_X f_Y > 0$ but $f_{XY} = 0$.

Difficulty allocation: A:2; estimated normal time 3 min.

2(e) (5 marks)

Recognition signal	Covariance from a joint density requires three moments.
First key step	Write $E[XY] - E[X]E[Y]$ and compute each integral on the same support.
Historical basis	2024 Q2(b)(vi); 2022 Q2(b)
Ability/method	joint LOTUS moments and covariance; K25, K27

Course-method solution. Using LOTUS over $0 < y < x < 1$: $E[X] = 3/4$, $E[Y] = 5/12$, and $E[XY] = 1/3$. Hence $\text{Cov}(X, Y) = 1/3 - (3/4)(5/12) = 1/48$.

Marking points. 1 mark: $E[X]$; 1 mark: $E[Y]$; 2 marks: $E[XY]$; 1 mark: covariance

Common errors. Using marginal means but forgetting $E[XY]$; assuming independence.

Final self-check. The positive support ordering $Y < X$ makes a small positive covariance plausible.

Difficulty allocation: B:1, C:1, D:3; estimated normal time 8 min.

2(f) (3 marks)

Recognition signal	A ratio transformation is most directly handled by the event $Y \leq rX$.
First key step	Determine the range of R before integrating.
Historical basis	2021 Q3(c)(iii); 2023 Q3(b)(iv)
Ability/method	CDF method for a ratio on triangular support; K20, K22

Course-method solution. R lies in $(0, 1)$. For $0 < r < 1$, $F_R(r) = P(Y \leq rX) = \int_0^1 \int_0^{rx} 2(x+y) dy dx = (2r + r^2)/3$. Thus $F_R(r) = 0$ for $r \leq 0$ and 1 for $r \geq 1$.

Marking points. 1 mark: support/cases; 1 mark: correct integration region; 1 mark: evaluated CDF

Common errors. Using $y \leq r$ without multiplying by x ; missing endpoint cases.

Final self-check. At $r=0$ the CDF is 0 and at $r=1$ it is 1.

Difficulty allocation: D:3; estimated normal time 7 min.

Total: 20 marks

Question 3

3(a) (3 marks)

Recognition signal	This asks for the standard loss-minimization characterizations.
First key step	Recall the completed-square identity and the median property.
Historical basis	2024 Q4(a); 2026 Q3(a)
Ability/method	mean/median loss minimizers; K29

Course-method solution. The squared-loss minimizer is $a=E[X]$, and the minimum is $\text{Var}(X)$. Any median of X minimizes absolute loss.

Marking points. 1 mark each for mean, variance, median

Common errors. Confusing the two minimizers.

Final self-check. Squared loss has units squared and its minimum is a variance.

Difficulty allocation: A:3; estimated normal time 3 min.

3(b) (3 marks)

Recognition signal	MSE is the second moment of the estimation error.
First key step	Set $W=\theta_{\text{hat}}-\theta$.
Historical basis	2023 Q4(a); 2026 Q3(b)
Ability/method	bias-variance identity; K31

Course-method solution. Apply $E[W^2]=\text{Var}(W)+(E[W])^2$ to $W=\theta_{\text{hat}}-\theta$. Since $\text{Var}(\theta_{\text{hat}}-\theta)=\text{Var}(\theta_{\text{hat}})$ and $E[\theta_{\text{hat}}-\theta]=\text{Bias}_{\theta}(\theta_{\text{hat}})$, $\text{MSE}=\text{Var}(\theta_{\text{hat}})+\text{Bias}^2$.

Marking points. 1 mark: identity; 1 mark: variance step; 1 mark: bias step

Common errors. Expanding but dropping the cross term without explanation.

Final self-check. Both terms on the right are non-negative.

Difficulty allocation: A:2, B:1; estimated normal time 4 min.

3(c) (4 marks)

Recognition signal	Normal sample, unknown variance, small n: use Student t.
First key step	Take the square root of the sample variance before forming the standard error.
Historical basis	2020 Q6(a)(ii); 2024 Q4(c)(i)
Ability/method	unknown-variance t confidence interval; K36

Course-method solution. Use $(\bar{X}-\mu)/(S/\sqrt{n})\sim t_{14}$. Here $s=2.5$. The margin is $2.145(2.5/\sqrt{15})=1.385$, giving $(72.815, 75.585)$.

Marking points. 1 mark: t pivot and df; 1 mark: formula; 1 mark: margin; 1 mark: interval

Common errors. Using $z=1.96$ or $s^2/\sqrt{n}$.

Final self-check. The interval is centered at 74.2 and has symmetric endpoints.

Difficulty allocation: A:2, B:2; estimated normal time 6 min.

3(d) (5 marks)

Recognition signal	The same t pivot used for the interval becomes the test statistic.
First key step	Standardize the observed mean under H_0 .
Historical basis	2021 Q5(b)(i)-(iii); 2026 Q3(d)(ii)
Ability/method	one-sample t test and p-value; K39

Course-method solution. $T=(74.2-75)/(2.5/\sqrt{15})=-1.239$ with 14 df. The two-sided p-value is $2P(t_{14}\geq 1.239)=0.236$ (approximately). Since $p>0.05$, fail to reject H_0 ; this is not proof that $\mu=75$.

Marking points. 1 mark: hypotheses/test; 1 mark: statistic; 1 mark: df; 1 mark: p-value; 1 mark: conclusion

Common errors. One-sided p-value; saying accept H_0 .

Final self-check. The null value 75 lies inside the 95% interval, consistent with non-rejection.

Difficulty allocation: A:2, B:1, C:2; estimated normal time 8 min.

3(e) (5 marks)

Recognition signal	The question asks for both the random-variable property under H_0 and the observed-data interpretation.
First key step	State the tail event relative to the null distribution.
Historical basis	2021 Q5(d); 2024 Q5(a); 2026 Q3(e)
Ability/method	p-value definition, interpretation and null distribution; K41

Course-method solution. A p-value is the null probability of obtaining a test statistic at least as incompatible with H_0 as the observed value, using the pre-specified tail rule. Small values are evidence against H_0 , not the probability that H_0 is true. If the null distribution of a continuous test statistic is correctly specified (or the p-value uses a continuous null CDF), the p-value is $\text{Uniform}(0,1)$ under H_0 .

Marking points. 2 marks: definition; 1 mark: interpretation; 1 mark: uniformity conditions; 1 mark: limitation

Common errors. Calling it $P(H_0 \text{ true}|\text{data})$; omitting continuity/exactness.

Final self-check. Under H_0 , $P(p \leq \alpha) = \alpha$ for exact continuous p-values.

Difficulty allocation: A:2, B:1, C:2; estimated normal time 6 min.

Total: 20 marks

Question 4

4(a) (4 marks)

Recognition signal	Correlation of a linear combination is found from covariance bilinearity and variance expansion.
First key step	Convert correlation to covariance first.
Historical basis	2023 Q5(c); 2026 Q4(a)
Ability/method	covariance algebra for linear combinations; K27

Course-method solution. $\text{Cov}(X,Y)=(-1/3)(3)(2)=-2$. $\text{Cov}(X,Z)=\text{Var}(X)-2\text{Cov}(X,Y)=9+4=13$. $\text{Var}(Z)=9+4(4)-4\text{Cov}(X,Y)=33$. Hence $\text{Corr}(X,Z)=13/(3\sqrt{33})$.

Marking points. 1 mark: covariance; 1 mark: $\text{Cov}(X,Z)$; 1 mark: $\text{Var}(Z)$; 1 mark: correlation

Common errors. Using means; forgetting the cross term in $\text{Var}(Z)$.

Final self-check. The absolute value $13/(3\sqrt{33})$ is below 1.

Difficulty allocation: A:2, B:2; estimated normal time 6 min.

4(b) (5 marks)

Recognition signal	The target is uncorrelatedness, so compute one covariance directly.
First key step	Expand $\text{Cov}(X+Y,X-Y)$ term by term.
Historical basis	2024 Q5(c)
Ability/method	proof by covariance bilinearity and equal-variance cancellation; K27

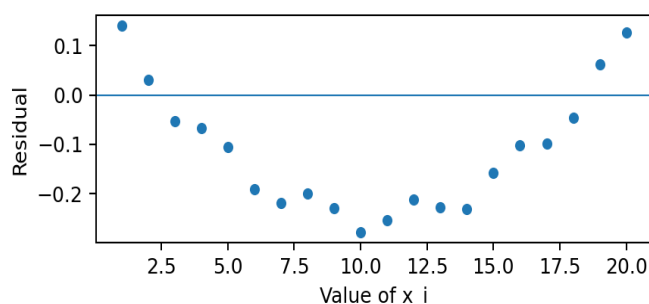
Course-method solution. By bilinearity and symmetry of covariance, $\text{Cov}(X+Y,X-Y)=\text{Var}(X)-\text{Cov}(X,Y)+\text{Cov}(Y,X)-\text{Var}(Y)=\text{Var}(X)-\text{Var}(Y)=0$. Therefore $X+Y$ and $X-Y$ are uncorrelated. Independence of X and Y is not required.

Marking points. 1 mark: expand covariance; 2 marks: use symmetry/cancellation; 1 mark: equal variances; 1 mark: conclusion

Common errors. Assuming X and Y are independent; confusing uncorrelatedness with independence.

Final self-check. The cross-covariance terms cancel and only the variance difference remains.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 7 min.



4(c) (4 marks)

Recognition signal	Residual structure overrides a superficially high R^2 .
First key step	Look for randomness around zero and constant spread.
Historical basis	2021 Q6(c); 2025 Q4(d); 2026 Q4(b)
Ability/method	residual diagnosis and transformation; K50

Course-method solution. No. The U-shaped residual pattern shows systematic curvature, so the linear mean structure is inadequate despite high R^2 . Try a transformed predictor such as x^2 or fit a polynomial/appropriate transformation, then recheck residuals.

Marking points. 1 mark: reject fit; 1 mark: identify pattern; 1 mark: R^2 limitation; 1 mark: specific remedy

Common errors. Saying the fit is good solely because R^2 is high.

Final self-check. After a remedy, residuals should show no systematic pattern.

Difficulty allocation: A:1, C:1, D:2; estimated normal time 5 min.

4(d) (5 marks)

Recognition signal	A support-dependent likelihood may be flat and non-unique.
First key step	Write the support inequalities before differentiating; no derivative is useful.
Historical basis	2026 Q4(f)
Ability/method	support-dependent non-unique MLE; K54

Course-method solution. The likelihood is 2^{-n} when $\theta \leq \min x_i$ and $\theta \geq \max x_i - 2$, and zero otherwise. Thus every θ in $[\max x_i - 2, \min x_i]$ is an MLE; the likelihood is flat on that interval.

Marking points. 1 mark: individual density/support; 1 mark: joint likelihood; 2 marks: convert all inequalities; 1 mark: full MLE set

Common errors. Forcing a unique endpoint; ignoring the zero-likelihood region.

Final self-check. The interval is nonempty exactly when the observed range is at most 2.

Difficulty allocation: C:1, D:4; estimated normal time 8 min.

4(e) (2 marks)

Recognition signal	A bootstrap resample is sampled with replacement and has the same size.
First key step	Check size and membership, not uniqueness.
Historical basis	2022 Q6(b)(i)-(ii); 2024 Q6(e)(i)-(ii); 2026 Q4(e)(i)-(ii)
Ability/method	bootstrap sample definition; K52

Course-method solution. (i) Yes: it has size 5 and every value comes from the original sample; repetition is allowed. (ii) No: it has only 4 observations, whereas a nonparametric bootstrap resample has the original size.

Marking points. 1 mark each classification with the decisive rule

Common errors. Rejecting repetitions; accepting a shorter sample.

Final self-check. A valid sample can repeat values but cannot introduce new values.

Difficulty allocation: A:1, B:1; estimated normal time 3 min.

Total: 20 marks

MATH40005

Probability and Statistics

Revised Mock Examination 2 - Solutions and Marking Guide

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers
Difficulty calibration	A 32, B 20, C 12, D 16
Paper role	2025 historical bridge: every high-mark 2025 ability family receives direct or method-level training, recalibrated to the 2026 2-hour paper.

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

1(a) (5 marks)

Recognition signal	Route categories form a partition and route-specific rates are conditional probabilities.
First key step	Write $P(L)=\sum_r P(L r)P(r)$.
Historical basis	2025 Q1(b)(i)
Ability/method	law of total probability with observational weights; K03

Course-method solution. By the law of total probability, $P(L)=0.04(0.5)+0.10(0.3)+0.05(0.2)=0.06$.

Marking points. 1 mark: notation/partition; 2 marks: LTP; 2 marks: calculation

Common errors. Averaging the late rates without route weights.

Final self-check. The weighted result 0.06 lies between 0.04 and 0.10.

Difficulty allocation: A:4, B:1; estimated normal time 6 min.

1(b) (3 marks)

Recognition signal	The conditioning direction is reversed from the provided late rate.
First key step	Use the total probability from part (a) as the denominator.
Historical basis	2025 Q1(b)(ii)
Ability/method	Bayes reversal after total probability; K08

Course-method solution. Bayes gives $P(A|L)=P(L|A)P(A)/P(L)=0.04(0.5)/0.06=1/3$.

Marking points. 1 mark: Bayes formula; 1 mark: numerator; 1 mark: result

Common errors. Reporting $P(L|A)=0.04$.

Final self-check. $P(A|L)=1/3$ is less than the prior $P(A)=1/2$ because A has the lowest late rate.

Difficulty allocation: A:2, B:1; estimated normal time 4 min.

1(c) (4 marks)

Recognition signal	An intervention changes conditional probabilities, so the denominator must be rebuilt.
First key step	Update the data-generating mechanism before reusing Bayes.
Historical basis	2025 Q1(b)(iii)
Ability/method	Bayes under a changed mechanism; K08,K03

Course-method solution. Now $P(L)=0.04(0.5)+1(0.3)+1(0.2)=0.52$. Therefore $P(A|L)=0.02/0.52=1/26$.

Marking points. 1 mark: update conditional rates; 1 mark: new total probability; 1 mark: Bayes; 1 mark: result

Common errors. Using the old $P(L)=0.06$.

Final self-check. The posterior should fall sharply because almost all B/C trains are late.

Difficulty allocation: B:1, C:1, D:2; estimated normal time 6 min.

1(d) (5 marks)

Recognition signal	Same combinatorial architecture as a multitype sequence.
First key step	Choose positions, then values.
Historical basis	2025 Q2(a)(i)
Ability/method	multitype password count; K05

Course-method solution. The count is $C(10,2)C(8,5)(5^*4)52^5 10^3$.

Marking points. 2 marks positions; 1 mark distinct-symbol ordering; 2 marks product

Common errors. Missing the order of symbols.

Final self-check. Check against $10!/(5!3!2!) C(5,2)2!52^5 10^3$.

Difficulty allocation: A:2, B:2, C:1; estimated normal time 6 min.

1(e) (3 marks)

Recognition signal	Plot selection depends on whether the goal is a relationship, robust distribution summary, or time order.
First key step	Classify the variable types and the analytical purpose.
Historical basis	2023 Q6(d)(i); 2023 Q6(d)(iii); 2023 Q6(d)(iv); 2026 Q3(c)(i)-(ii)
Ability/method	plot choice for relationship, robust summary and time series; K47

Course-method solution. (i) A scatterplot. (ii) A boxplot. (iii) A time-series/line plot in chronological order.

Marking points. 1 mark each plot with purpose

Common errors. Using a histogram for a two-variable relationship; ignoring chronology in a stock series.

Final self-check. Each chosen plot directly displays the feature named in the question.

Difficulty allocation: A:2, B:1; estimated normal time 4 min.

Total: 20 marks

Question 2

2(a) (4 marks)

Recognition signal	A monotone transform of a uniform variable suggests the CDF method.
First key step	Solve $-\log U/\lambda \leq t$ carefully, noting the inequality direction.
Historical basis	2025 Q3(a)(i)
Ability/method	inverse-transform/CDF method; K20,K24

Course-method solution. For $t < 0$, $F_T(t) = 0$. For $t \geq 0$, $F_T(t) = P(U \geq e^{-\lambda t}) = 1 - e^{-\lambda t}$. Thus $T \sim \text{Exponential}(\lambda)$.

Marking points. 1 mark support; 2 marks CDF transformation; 1 mark identification

Common errors. Writing $P(U \leq e^{-\lambda t})$.

Final self-check. The density is $\lambda e^{-\lambda t}$ on $t > 0$.

Difficulty allocation: A:3, B:1; estimated normal time 5 min.

2(b) (4 marks)

Recognition signal	A maximum event becomes an intersection.
First key step	Write $\{M \leq m\} = \{X_1 \leq m, \dots, X_n \leq m\}$.
Historical basis	2025 Q3(a)(ii)
Ability/method	order-statistic CDF; K21

Course-method solution. For $0 < m < 1$, $F_M(m) = P(\text{all } X_i \leq m) = m^n$. It is 0 for $m \leq 0$ and 1 for $m \geq 1$. The density is $n m^{n-1}$ on $(0,1)$.

Marking points. 1 mark event rewrite; 1 mark independence; 1 mark piecewise CDF; 1 mark density

Common errors. Using $1 - (1-m)^n$, which is for a minimum-related event.

Final self-check. $F_M(1) = 1$ and the density integrates to 1.

Difficulty allocation: A:3, B:1; estimated normal time 5 min.

2(c) (4 marks)

Recognition signal	The exponential tilt of a normal density is handled by completing the square.
First key step	Combine exponents before integrating.
Historical basis	2025 Q3(b)(i)
Ability/method	MGF derivation by completing the square; K12,K24

Course-method solution. $M_X(t) = \int \exp(tx) (2\pi)^{-1/2} \exp(-x^2/2) dx$. Complete the square: $tx - x^2/2 = -(x-t)^2/2 + t^2/2$. The shifted normal density integrates to 1, so $M_X(t) = \exp(t^2/2)$, t real.

Marking points. 1 mark setup; 2 marks completing square; 1 mark result/domain

Common errors. Treating the integral as elementary without normalization.

Final self-check. $M_X(0) = 1$ and M_X is even.

Difficulty allocation: A:1, B:1, C:1, D:1; estimated normal time 7 min.

2(d) (4 marks)

Recognition signal	A sum of independent variables suggests multiplying MGFs.
First key step	Use the result from part (c).
Historical basis	2025 Q3(b)(ii)
Ability/method	MGF product and distribution identification; K12,K24

Course-method solution. Independence gives $M_S(t) = \prod \exp(t^2/2) = \exp(nt^2/2)$, the MGF of $N(0,n)$. Therefore $S \sim N(0,n)$.

Marking points. 1 mark independence/product; 1 mark MGF; 1 mark identification; 1 mark parameters

Common errors. Giving variance $\sqrt{n}$.

Final self-check. The mean is 0 and variance is n .

Difficulty allocation: A:1, B:1, C:1, D:1; estimated normal time 4 min.

2(e) (4 marks)

Recognition signal	Moment equality alone need not determine a distribution.
First key step	Add a condition ensuring the moment series defines the MGF near zero.
Historical basis	2025 Q3(c)(i)
Ability/method	moment determinacy via MGF existence; K34,K12

Course-method solution. A sufficient condition is that both MGFs exist on an open interval containing 0. Then their Taylor series have coefficients $E[U^k]/k!$ and $E[V^k]/k!$, so the MGFs agree near 0. Uniqueness of MGFs implies U and V have the same distribution.

Marking points. 1 mark condition; 2 marks equality of MGFs; 1 mark uniqueness conclusion

Common errors. Claiming moments always determine the distribution.

Final self-check. The argument must invoke a uniqueness theorem.

Difficulty allocation: B:1, C:1, D:2; estimated normal time 6 min.

Total: 20 marks

Question 3

3(a) (2 marks)

Recognition signal	Different means do not affect variance.
First key step	Apply variance to the fixed linear combination.
Historical basis	2025 Q4(a)(i)
Ability/method	variance of an independent sample mean; K26,K30

Course-method solution. By independence, $\text{Var}(\bar{X}) = n^{-2} \sum \text{Var}(X_i) = n^{-2} n \sigma^2 = \sigma^2/n$.

Marking points. 1 mark variance scaling; 1 mark independence sum

Common errors. Adding standard deviations.

Final self-check. For $n=1$ the formula reduces to σ^2 .

Difficulty allocation: A:2; estimated normal time 3 min.

3(b) (6 marks)

Recognition signal	The usual unbiasedness fails when the means differ.
First key step	Rewrite the centered sum before taking expectations.
Historical basis	2025 Q4(a)(ii)
Ability/method	expectation of sample variance under unequal means; K30,K31

Course-method solution. Use $\sum (X_i - \bar{X})^2 = \sum X_i^2 - n \bar{X}^2$. Now $E[\sum X_i^2] = \sum (\sigma^2 + \mu_i^2) = n \sigma^2 + \sum \mu_i^2$. Also $E[\bar{X}^2] = \sigma^2/n + \bar{\mu}^2$. Hence $E[(n-1)S^2] = (n-1)\sigma^2 + \sum (\mu_i - \bar{\mu})^2$, so $E[S^2] = \sigma^2 + [\sum (\mu_i - \bar{\mu})^2]/(n-1)$.

Marking points. 1 mark decomposition; 2 marks expectations; 2 marks algebra; 1 mark final expression

Common errors. Assuming $E[S^2] = \sigma^2$; omitting between-mean variation.

Final self-check. If all μ_i are equal, the extra term vanishes.

Difficulty allocation: A:1, B:2, C:1, D:2; estimated normal time 9 min.

3(c) (4 marks)

Recognition signal	Unknown variance and normality imply a t interval.
First key step	Use s , not s^2 .
Historical basis	2025 Q4(b)(i)
Ability/method	Student-t confidence interval; K36

Course-method solution. The interval is $\bar{x} \pm 1.796 s/\sqrt{12}$. Since $s = \sqrt{3}$, the margin is 0.898, so (180.102, 181.898).

Marking points. 1 mark pivot/df; 1 mark standard error; 1 mark margin; 1 mark interval

Common errors. Using 1.645 normal critical value.

Final self-check. Endpoints are symmetric around 181.

Difficulty allocation: A:3, B:1; estimated normal time 5 min.

3(d) (4 marks)

Recognition signal	Variance inference uses a chi-square pivot and reverses endpoints when inverted.
First key step	Write the probability statement before solving for σ^2 .
Historical basis	2025 Q4(b)(ii)
Ability/method	chi-square variance interval; K37

Course-method solution. Since $11S^2/\sigma^2 \sim \chi^2_{11}$, the interval is $(11 \cdot 3/21.920, 11 \cdot 3/3.816) = (1.505, 8.648)$.

Marking points. 1 mark pivot; 1 mark inversion/order; 1 mark substitution; 1 mark result

Common errors. Putting the larger chi-square quantile in the upper denominator.

Final self-check. The lower bound is smaller than s^2 and the upper bound larger.

Difficulty allocation: A:2, B:1, C:1; estimated normal time 6 min.

3(e) (4 marks)

Recognition signal	A Q-Q plot is judged by systematic departures from the reference line.
First key step	Inspect center and tails separately.
Historical basis	2025 Q4(d); 2023 Q6(d)(ii)
Ability/method	Q-Q diagnostic; K47

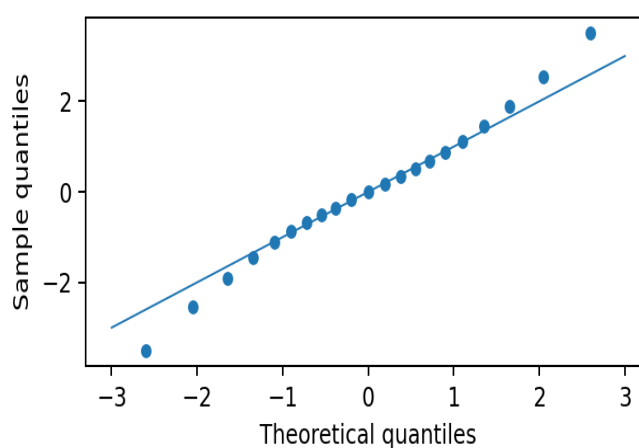
Course-method solution. The middle points are near the line but both tails deviate outward. This is evidence of heavier tails than a normal distribution, so the plot does not support a satisfactory normal model.

Marking points. 1 mark pattern; 1 mark tail interpretation; 1 mark conclusion; 1 mark qualified wording

Common errors. Requiring every point to be exactly on the line; ignoring tail curvature.

Final self-check. Heavy tails produce more extreme sample quantiles than the normal line.

Difficulty allocation: A:1, C:1, D:2; estimated normal time 4 min.



Total: 20 marks

Question 4

4(a) (5 marks)

Recognition signal	Both endpoints enter the support and the likelihood decreases with interval width.
First key step	Write constraints first, then shrink the interval to the sample range.
Historical basis	2025 Q6(a)
Ability/method	two-boundary support-dependent MLE; K54

Course-method solution. The likelihood is $(\theta_2 - \theta_1)^{-n}$ when $\theta_1 \leq x_{\min}$ and $\theta_2 \geq x_{\max}$, and zero otherwise. To maximize it, minimize the interval length subject to containing all data: $\theta_1_{\hat{}} = x_{\min}$ and $\theta_2_{\hat{}} = x_{\max}$.

Marking points. 1 mark likelihood/support; 2 marks optimization argument; 2 marks estimators

Common errors. Differentiating inside an unconstrained region; choosing sample mean.

Final self-check. The fitted support must contain every observation exactly with minimum width.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 7 min.

4(b) (4 marks)

Recognition signal	Writing estimators as linear combinations makes covariance algebra direct.
First key step	Use $\sum(x_i - \bar{x}) = 0$.
Historical basis	2025 Q6(b)(i)-(iii)
Ability/method	regression-estimator covariance; K49, K48

Course-method solution. Because $\sum(x_i - \bar{x}) = 0$, $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum(x_i - \bar{x})y_i$. Thus $\beta_{1\hat{}} = \sum a_i y_i$ with $a_i = (x_i - \bar{x})/S_{xx}$ and $\beta_{0\hat{}} = \bar{y} - \bar{x} \beta_{1\hat{}}$. $\text{Cov}(\bar{y}, \beta_{1\hat{}}) = 0$ and $\text{Var}(\beta_{1\hat{}}) = \sigma^2/S_{xx}$, so $\text{Cov}(\beta_{0\hat{}}, \beta_{1\hat{}}) = -\bar{x} \sigma^2/S_{xx}$.

Marking points. 1 mark identity; 1 mark linear-estimator form; 1 mark covariance pieces; 1 mark result

Common errors. Assuming the estimators are independent; dropping $\bar{x}$.

Final self-check. If $\bar{x} = 0$, the covariance is 0.

Difficulty allocation: A:1, B:1, C:1, D:1; estimated normal time 7 min.

4(c) (4 marks)

Recognition signal	Bootstrap approximates the sampling distribution by resampling the empirical distribution.
First key step	Resample observations, not residuals, for this simple statistic.
Historical basis	2025 Q6(c)
Ability/method	percentile bootstrap interval; K52

Course-method solution. Draw many independent resamples of size n with replacement from the observed values. Compute the median for each resample. Use the empirical 2.5th and 97.5th percentiles of the bootstrap medians as the percentile interval.

Marking points. 1 mark resample size/replacement; 1 mark replicate statistic; 1 mark repetition; 1 mark quantiles

Common errors. Sampling without replacement; using original-data quantiles rather than median replicates.

Final self-check. The interval endpoints are quantiles of the bootstrap statistic distribution.

Difficulty allocation: A:1, B:2, C:1; estimated normal time 5 min.

4(d) (7 marks)

Recognition signal	Known and unequal population variances signal a two-sample z test, not a pooled t test.
First key step	Write $\text{Var}(\bar{X} - \bar{Y}) = \sigma_1^2/n + \sigma_2^2/m$.
Historical basis	2025 Q5(b)
Ability/method	known-variance two-sample z test with p-value; K40

Course-method solution. Under H_0 , $Z = (\bar{X} - \bar{Y}) / \sqrt{10/15 + 4/12} = -3 / \sqrt{1} = -3$. The two-sided p-value is $2\Phi(-3) = 0.00270$. Since $0.00270 < 0.01$, reject H_0 ; the data provide evidence that the population mean heights differ. The assumptions are independent normal samples and the stated known variances.

Marking points. 1 mark hypotheses/assumptions; 2 marks standard error/statistic; 2 marks p-value; 1 mark comparison; 1 mark qualified conclusion

Common errors. Pooling the known variances; using a one-sided p-value; saying the means are proved unequal.

Final self-check. The standard error is exactly 1 and the observed z-score is -3.

Difficulty allocation: A:2, B:1, C:1, D:3; estimated normal time 10 min.

Total: 20 marks

MATH40005

Probability and Statistics

Revised Mock Examination 3 - Solutions and Marking Guide

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers
Difficulty calibration	A 32, B 20, C 12, D 16
Paper role	2023-2024 historical ability paper: collision/Bayes/measurability, random-count conditioning, curved supports, exact testing and compound Poisson.

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

1(a) (5 marks)

Recognition signal	"At least one collision" is easier via no collision.
First key step	Count injective assignments of weeks.
Historical basis	2023 Q1(a)
Ability/method	birthday collision complement; K07

Course-method solution. Use the complement. $P(\text{no shared week}) = (52)_2 / 52^2$, so the answer is $1 - (52)_2 / 52^2$.

Marking points. 1 mark assumptions; 2 marks complement count; 2 marks probability

Common errors. Adding pairwise collision probabilities.

Final self-check. The answer is close to 1 and lies in $[0,1]$.

Difficulty allocation: A:3, B:2; estimated normal time 6 min.

1(b) (5 marks)

Recognition signal	Exactly one pair requires excluding any further collision.
First key step	Build the configuration in disjoint stages.
Historical basis	2023 Q1(b)
Ability/method	exact single-collision count; K07, K05

Course-method solution. Choose the pair, choose their common day, then assign distinct days to the remaining $n-2$ people from the other 364 days: $C(n,2) \cdot 365 \cdot (364)_{n-2} / 365^n = C(n,2) 365! / [(366-n)! 365^n]$.

Marking points. 1 mark choose pair; 1 mark common day; 2 marks remaining injective assignment; 1 mark probability

Common errors. Using $C(n,2)/365$ only; allowing another pair.

Final self-check. For $n=2$ it would reduce to $1/365$, although the stated range begins at 3.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 8 min.

1(c) (5 marks)

Recognition signal	The symptom reverses the conditioning direction.
First key step	Compute the total rash probability first.
Historical basis	2023 Q1(c)
Ability/method	diagnostic Bayes calculation; K08, K03

Course-method solution. $P(R) = 0.01(0.95) + 0.90(0.05) = 0.0545$. Bayes gives $P(M|R) = 0.90(0.05) / 0.0545 = 90/109 \approx 0.826$.

Marking points. 1 mark notation; 2 marks total probability; 2 marks Bayes/result

Common errors. Using sensitivity alone as the posterior.

Final self-check. The posterior is much larger than prior 0.05 because rash is strongly associated with measles.

Difficulty allocation: A:3, B:2; estimated normal time 6 min.

1(d) (5 marks)

Recognition signal	A piecewise-defined random variable is checked through images and preimages.
First key step	Write the preimage of a singleton as a union on the two events.
Historical basis	2023 Q1(d)
Ability/method	measurability of a piecewise discrete variable; K15, K02

Course-method solution. The image of Z is contained in $\text{Im}(X) \cup \text{Im}(Y)$, hence countable. For any z , $Z^{-1}(\{z\}) = [X^{-1}(\{z\}) \cap A] \cup [Y^{-1}(\{z\}) \cap A^c]$, which is in $\mathcal{F}$. Thus Z is discrete/measurable.

Marking points. 1 mark countable image; 3 marks preimage identity/measurability; 1 mark conclusion

Common errors. Only showing countability; forgetting measurability.

Final self-check. Each component set is in $\mathcal{F}$ by closure under complements, intersections and unions.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 7 min.

Total: 20 marks

Question 2

2(a) (4 marks)

Recognition signal	Distributional marginals alone do not determine the sum without dependence information.
First key step	Try perfect dependence $Y=X$.
Historical basis	2023 Q3(a)
Ability/method	dependence counterexample for Poisson sums; K13,K04

Course-method solution. No. Take $\mu=\lambda$ and $Y=X$. Then $X+Y=2X$ takes only even values, whereas a non-degenerate Poisson variable assigns positive probability to odd values.

Marking points. 1 mark correct answer; 1 mark counterexample; 2 marks support justification

Common errors. Automatically using the independent-Poisson theorem.

Final self-check. Check the support of $2X$.

Difficulty allocation: A:2, C:1, D:1; estimated normal time 5 min.

2(b) (6 marks)

Recognition signal	A random number of variables signals total probability/expectation over N .
First key step	Condition on N first.
Historical basis	2023 Q2(b)
Ability/method	random-count conditioning and exponential-series sum; K23,K21,K13

Course-method solution. Condition on $N=n$. The conditional probability is $(1-x)^{n+1}$. Therefore $\sum_{n \geq 0} (1-x)^{n+1} e^{-\lambda} \lambda^n / n! = (1-x)e^{-\lambda} x$.

Marking points. 1 mark condition on N ; 2 marks conditional probability; 2 marks Poisson sum; 1 mark result

Common errors. Using N rather than $N+1$; forgetting independence.

Final self-check. At $x=0$ the result is 1; as x approaches 1 it tends to 0.

Difficulty allocation: A:2, B:1, C:1, D:2; estimated normal time 8 min.

2(c) (5 marks)

Recognition signal	The event is a half-plane intersection in a rotationally symmetric bivariate normal.
First key step	Either draw the wedge or integrate using $d\Phi = \phi dx$.
Historical basis	2023 Q2(c)
Ability/method	geometric/integral probability for independent normals; K22,K24

Course-method solution. Using the joint density, integrate $\int_0^\infty \phi(x)P(Y>-x)dx = \int_0^\infty \phi(x)\Phi(x)dx = [\Phi(x)^2/2]_0^\infty = 3/8$. A rotational-symmetry argument also gives a $3\pi/4$ wedge out of 2π .

Marking points. 1 mark region; 2 marks integral/symmetry; 1 mark evaluation; 1 mark result

Common errors. Treating $Y>-X$ as independent of X .

Final self-check. The wedge angle is 135 degrees, so $3/8$ is plausible.

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 7 min.

2(d) (5 marks)

Recognition signal	Thinning a Poisson count uses a conditional binomial allocation.
First key step	Write the joint pmf, not just marginal expectations.
Historical basis	historical Poisson thinning ability; 2024 Q3(a) related Poisson conditioning
Ability/method	Poisson splitting via joint pmf; K16,K13

Course-method solution. For $u, r \geq 0$, $P(U=u, R=r) = P(N=u+r)C(u+r, u)p^u(1-p)^r = e^{-\lambda}(\lambda p)^u / u! [\lambda(1-p)]^r / r!$. This factors into $\text{Poisson}(\lambda p)$ and $\text{Poisson}(\lambda(1-p))$ pmfs, proving both marginals and independence.

Marking points. 1 mark condition $N=u+r$; 2 marks algebra; 1 mark marginals; 1 mark factorization/independence

Common errors. Claiming independence from uncorrelatedness.

Final self-check. The two rates add to λ .

Difficulty allocation: A:2, B:2, D:1; estimated normal time 8 min.

Total: 20 marks

Question 3

3(a) (8 marks)

Recognition signal	The curved support must be rewritten in both integration orders.
First key step	Express it once as y from $\sqrt{x}$ to 1 and once as x from 0 to y^2 .
Historical basis	2023 Q3(b)(i)-(iii)
Ability/method	curved-support normalization and marginalization; K17,K18,K19,K22

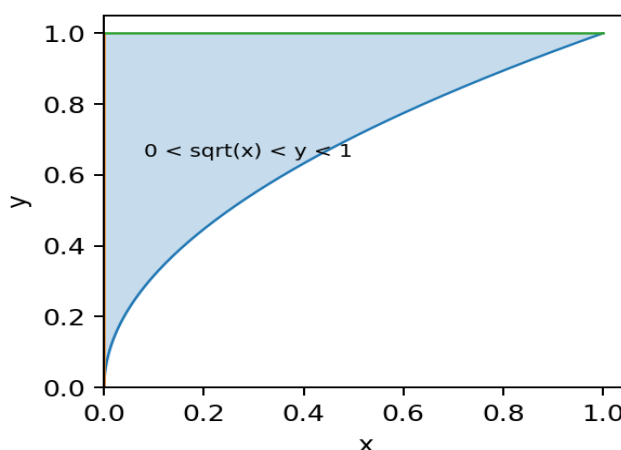
Course-method solution. The support is $0 < y < 1$, $0 < x < y^2$. Normalization gives $c \int_0^1 \int_0^{y^2} (x+y) dx dy = 1$, so $c = 20/7$. For $0 < x < 1$, $f_X(x) = c \int_{\sqrt{x}}^1 (x+y) dy = c(x/2 + 1/2 - x^{3/2})$. For $0 < y < 1$, $f_Y(y) = c \int_0^{y^2} (x+y) dx = c(y^4/2 + y^3)$. They are not independent because the support is not a rectangle.

Marking points. 2 marks normalization; 2 marks f_X ; 2 marks f_Y ; 2 marks independence/support

Common errors. Using triangular bounds $y < x$; factorizing only the formula and ignoring support.

Final self-check. Both marginals integrate to 1.

Difficulty allocation: A:5, B:2, D:1; estimated normal time 11 min.



3(b) (5 marks)

Recognition signal	A conditional CDF needs the conditional support and piecewise endpoint cases.
First key step	Divide by $f_X(x)$, then integrate from $\sqrt{x}$.
Historical basis	2023 Q3(b)(iv)
Ability/method	conditional CDF on curved support; K18

Course-method solution. For fixed x , Y lies in $(\sqrt{x}, 1)$ and $f_{Y|X=x}(v) = c(x+v)/f_X(x)$. Thus the CDF is 0 for $y \leq \sqrt{x}$, 1 for $y \geq 1$, and $[x(y-\sqrt{x}) + (y^2-x)/2] / [x(1-\sqrt{x}) + (1-x)/2]$ for $\sqrt{x} < y < 1$.

Marking points. 1 mark conditional density; 1 mark support cases; 2 marks integral; 1 mark expression

Common errors. Integrating from 0; not specifying x domain.

Final self-check. At $y=\sqrt{x}$ the middle formula is 0 and at $y=1$ it is 1.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 8 min.

3(c) (7 marks)

Recognition signal	A monotone transform permits direct inversion; expectation can use LOTUS without the transformed density.
First key step	Find the transformed support first.
Historical basis	2023 Q3(c)
Ability/method	monotone CDF transform plus LOTUS; K20,K24

Course-method solution. W lies in $(1, 81)$. For $1 \leq w < 81$, $F_W(w) = P(X \leq w^{1/4}) = (w^{1/4} - 1)/2$. Hence $f_W(w) = w^{-3/4}/8$ on $(1, 81)$. By LOTUS, $E[W] = (1/2) \int_1^{81} x^{3/4} dx = (3^5 - 1)/10 = 24.2$.

Marking points. 2 marks piecewise CDF; 2 marks density/support; 2 marks expectation; 1 mark result

Common errors. Forgetting the factor $1/2$ or derivative $1/4$.

Final self-check. The density integrates to 1 and $E[W]$ lies between 1 and 81.

Difficulty allocation: A:3, B:2, C:1, D:1; estimated normal time 9 min.

Total: 20 marks

Question 4

4(a) (8 marks)

Recognition signal	A random sum asks for conditioning on the count.
First key step	Compute conditional mean and variance given N .
Historical basis	2024 Q3(b)(ii)-(iii); 2022 random-sum methods
Ability/method	compound-Poisson moments via total expectation/variance; K13,K23,K26

Course-method solution. Condition on N . $E[S|N]=Nm$, so $E[S]=\lambda m$. Also $\text{Var}(S|N)=Nv$. The law of total variance gives $\text{Var}(S)=E[Nv]+\text{Var}(Nm)=\lambda v+\lambda m^2=\lambda(v+m^2)$.

Marking points. 2 marks conditional mean; 2 marks expectation; 2 marks conditional variance; 2 marks total variance

Common errors. Using λv only; forgetting count variation.

Final self-check. If $Y_i=1$, $S=N$ and the variance formula gives λ .

Difficulty allocation: A:3, B:2, C:1, D:2; estimated normal time 10 min.

4(b) (5 marks)

Recognition signal	Fixed numbers selected without replacement imply a hypergeometric exact distribution.
First key step	Define the success count and use the tail at least as extreme.
Historical basis	2024 Q5(b)
Ability/method	exact hypergeometric tail test; K10,K44

Course-method solution. Under H_0 of random guessing, X , the number of true cola glasses selected, is Hypergeometric($N=12, K=6, n=6$). The one-sided p-value is $P(X \geq 5) = [C(6,5)C(6,1) + C(6,6)C(6,0)]/C(12,6) = 37/924 \approx 0.0400$. Reject H_0 at 0.05.

Marking points. 1 mark hypotheses/model; 2 marks tail probability; 1 mark value; 1 mark conclusion

Common errors. Using a binomial model; using only $P(X=5)$.

Final self-check. The p-value is just below 0.05, so the conclusion is sensitive but clear.

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 8 min.

4(c) (4 marks)

Recognition signal	Known population variance signals a normal critical value.
First key step	Compute $\sigma/\sqrt{n}$.
Historical basis	2020 Q6(a)(i); 2023 Q4(c)(i)
Ability/method	known-variance z confidence interval; K35

Course-method solution. Use $\bar{X} \sim N(\mu, 36/36)$. The standard error is 1, so the 95% interval is $102 \pm 1.96 = (100.04, 103.96)$.

Marking points. 1 mark distribution/pivot; 1 mark standard error; 1 mark formula; 1 mark interval

Common errors. Using t or sample variance.

Final self-check. The interval is centered at $\bar{x}$.

Difficulty allocation: A:1, B:1, C:2; estimated normal time 4 min.

4(d) (3 marks)

Recognition signal	The distinction is between a procedure's coverage and a posterior probability.
First key step	State what is random before data are observed.
Historical basis	2023 Q4(c)(ii); 2024 Q6(a)
Ability/method	frequentist confidence interpretation; K38

Course-method solution. The parameter is fixed and the interval is random before sampling. Over repeated samples, 95% of intervals produced by this procedure cover the true parameter. After observing the data, the realized interval either contains the parameter or it does not; it is not correct to assign 95% probability to the fixed parameter lying in this particular interval.

Marking points. 1 mark repeated-sampling coverage; 1 mark fixed parameter/random interval; 1 mark reject posterior interpretation

Common errors. Saying there is a 95% chance the fixed parameter is in the realized interval.

Final self-check. The confidence level belongs to the method over repetitions.

Difficulty allocation: A:1, B:1, C:1; estimated normal time 4 min.

Total: 20 marks

MATH40005

Probability and Statistics

Revised Mock Examination 4 - Solutions and Marking Guide

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers
Difficulty calibration	A 32, B 20, C 12, D 16
Paper role	2020-2022 core-method paper: definitions, stars-and-bars, double counting, geometric waits, estimator risk, inequalities, OLS, support likelihood and conjugacy.

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

1(a) (3 marks)

Recognition signal	A definition question requires all closure axioms.
First key step	List the three axioms.
Historical basis	2020 Q1(a)
Ability/method	sigma-algebra definition; K02

Course-method solution. F is a collection of subsets of Ω containing the empty set, closed under complements, and closed under countable unions.

Marking points. 1 mark each axiom

Common errors. Saying only “all events”; forgetting countable union.

Final self-check. The axioms imply Ω is in F and closure under countable intersections.

Difficulty allocation: A:3; estimated normal time 3 min.

1(b) (3 marks)

Recognition signal	A probability measure is more than a function assigning numbers.
First key step	State non-negativity, normalization and countable additivity.
Historical basis	2020 Q1(b)
Ability/method	probability-measure definition; K03

Course-method solution. $P:F \rightarrow [0,1]$ is non-negative, satisfies $P(\Omega)=1$, and is countably additive on pairwise disjoint events.

Marking points. 1 mark each axiom

Common errors. Using finite additivity only.

Final self-check. It follows that $P(\text{empty})=0$.

Difficulty allocation: A:3; estimated normal time 3 min.

1(c) (4 marks)

Recognition signal	Positive ordered parts correspond to bars in gaps, not unrestricted bins.
First key step	Place $n-1$ bars among $k-1$ gaps.
Historical basis	2020 Q1(e)
Ability/method	positive stars and bars; K06

Course-method solution. If $k < n$ there are none. If $k \geq n$, represent k stars and choose $n-1$ of the $k-1$ internal gaps for bars, giving $C(k-1, n-1)$.

Marking points. 1 mark impossible case; 2 marks stars-and-bars construction; 1 mark formula

Common errors. Using $C(k+n-1, n-1)$, which allows zeros.

Final self-check. For $k=n$ the formula gives 1.

Difficulty allocation: A:2, B:1, D:1; estimated normal time 5 min.

1(d) (4 marks)

Recognition signal	The factor $n-k$ suggests converting $C(n,k)(n-k)$ to $nC(n-1,k)$.
First key step	Start from the left and expand $(3+1)^{n-1}$.
Historical basis	2022 Q1(b)(i)
Ability/method	binomial identity proof; K05

Course-method solution. Use $nC(n-1,k) = C(n,k)(n-k)$. Then $n4^{n-1} = n(3+1)^{n-1} = \sum_{k=0}^{n-1} nC(n-1,k)3^k = \sum_{k=0}^{n-1} C(n,k)3^k 4^{n-k}$, with the $k=n$ term zero.

Marking points. 1 mark binomial theorem; 2 marks coefficient identity; 1 mark endpoint

Common errors. Differentiating incorrectly or losing the zero term.

Final self-check. Check $n=1$: both sides equal 1.

Difficulty allocation: A:1, B:2, C:1; estimated normal time 6 min.

1(e) (6 marks)

Recognition signal	A story proof requires two counts of exactly the same object.
First key step	Interpret $n-k$ as the choice of a special person among non-alcohol drinkers.
Historical basis	2022 Q1(b)(ii)
Ability/method	double counting/story proof; K05

Course-method solution. Count drink orders for n people with four choices (three alcoholic and water) together with a designated sober driver who drinks water. Left: choose the driver (n ways), then each other person has four choices, $n4^{\{n-1\}}$. Right: choose k people who take an alcoholic drink and choose one of 3 drinks for each, $C(n,k)3^k$; choose the driver among the remaining $n-k$ water drinkers. Sum over k .

Marking points. 2 marks common combinatorial object; 2 marks left count; 2 marks right count

Common errors. Giving two unrelated stories; not explaining the sum over k .

Final self-check. Both counts include a designated driver and all drink assignments.

Difficulty allocation: B:2, C:1, D:3; estimated normal time 9 min.

Total: 20 marks

Question 2

2(a) (4 marks)

Recognition signal	A discrete CDF is a right-continuous step function.
First key step	Accumulate masses in support order.
Historical basis	2020 Q2(b)
Ability/method	piecewise discrete CDF; K11

Course-method solution. $F_X(x)=0$ for $x<1$; $1/2$ for $1\leq x<3$; $5/6$ for $3\leq x<5$; and 1 for $x\geq 5$.

Marking points. 1 mark each interval

Common errors. Defining it only at 1,3,5.

Final self-check. Jumps equal the pmf and the final value is 1.

Difficulty allocation: A:4; estimated normal time 4 min.

2(b) (6 marks)

Recognition signal	A minimum is easiest through the event both variables exceed k.
First key step	Use survival functions.
Historical basis	2020 Q2(c)(ii)
Ability/method	minimum of independent geometric waits; K14,K04

Course-method solution. For integer $k\geq 0$, $P(\min(X,Y)>k)=P(X>k)P(Y>k)=[(1-p_1)(1-p_2)]^k$. Thus the minimum is $\text{Geom}(q)$ on $\{1,2,\dots\}$ with $q=1-(1-p_1)(1-p_2)$.

Marking points. 1 mark survival function; 1 mark independence; 2 marks algebra; 2 marks identify parameter/support

Common errors. Multiplying CDFs rather than survivals.

Final self-check. If p_1 or p_2 increases, q should increase.

Difficulty allocation: A:2, B:2, C:1, D:1; estimated normal time 7 min.

2(c) (6 marks)

Recognition signal	The target pattern can be decomposed into first reaching H, then reaching T while a trailing H remains available.
First key step	Use a state or two-stage waiting argument.
Historical basis	2020 Q2(d)
Ability/method	pattern waiting by states/geometric stages; K14,K23

Course-method solution. Wait for the first H: a $\text{Geom}(1/2)$ time with mean 2. After an H has occurred, wait for the first subsequent T; further H outcomes preserve a trailing H, so this is another $\text{Geom}(1/2)$ wait with mean 2. The total expected time is 4.

Marking points. 2 marks decomposition; 2 marks geometric means; 2 marks total/justification

Common errors. Treating non-overlapping pairs as independent trials.

Final self-check. A short Markov-state recursion gives the same answer 4.

Difficulty allocation: A:1, B:1, C:1, D:3; estimated normal time 8 min.

2(d) (4 marks)

Recognition signal	Partition conditioning expands each mass by total probability.
First key step	Start from the definition of expectation.
Historical basis	2020 Q3(d)(ii)
Ability/method	proof of total expectation; K23,K03

Course-method solution. $E[X]=\sum_x xP(X=x)=\sum_x x \sum_i P(X=x|B_i)P(B_i)$. Absolute convergence allows interchange: $=\sum_i P(B_i)\sum_x xP(X=x|B_i)=\sum_i E[X|B_i]P(B_i)$.

Marking points. 1 mark total probability; 1 mark interchange justification; 2 marks conclusion

Common errors. Interchanging infinite sums without the stated condition.

Final self-check. For a two-set partition it reduces to the familiar formula.

Difficulty allocation: A:1, B:2, C:1; estimated normal time 6 min.

Total: 20 marks

Question 3

3(a) (6 marks)

Recognition signal	A scaled sample sum requires separate expectation and variance calculations.
First key step	Compute E and Var before applying bias-variance.
Historical basis	2022 Q4(a)(i)-(ii); 2020 Q4(a)(iv)-(v)
Ability/method	bias/variance/MSE of an estimator; K31

Course-method solution. $E[T] = n\theta/(n+1)$, so $\text{Bias} = -\theta/(n+1)$. $\text{Var}(T) = n\sigma^2/(n+1)^2$. Hence $\text{MSE} = (\theta^2 + n\sigma^2)/(n+1)^2$.

Marking points. 2 marks bias; 2 marks variance; 2 marks MSE

Common errors. Using $n^2\sigma^2$; omitting squared bias.

Final self-check. MSE is non-negative and decreases at order $1/n$ for fixed parameters.

Difficulty allocation: A:3, B:2, C:1; estimated normal time 7 min.

3(b) (4 marks)

Recognition signal	The event indicator converts a tail probability into an expectation bound.
First key step	Use $X \geq a \implies \mathbb{1}\{X \geq a\}$.
Historical basis	2020 Q4(b)
Ability/method	indicator proof of Markov inequality; K28

Course-method solution. For $a > 0$, $X \geq a \implies \mathbb{1}\{X \geq a\}$. Taking expectations gives $E[X] \geq aP(X \geq a)$, hence $P(X \geq a) \leq E[X]/a$.

Marking points. 1 mark pointwise inequality; 2 marks expectation; 1 mark result

Common errors. Assuming a density; Markov is distribution-free.

Final self-check. Equality can occur for a variable taking only 0 and a .

Difficulty allocation: A:2, B:2; estimated normal time 5 min.

3(c) (5 marks)

Recognition signal	Uniformity in p calls for $p(1-p) \leq 1/4$.
First key step	Bound the worst-case variance before solving for n .
Historical basis	2022 Q4(b); 2023 Q4(b)
Ability/method	distribution-free sample-size guarantee; K45, K28

Course-method solution. $\text{Var}(\hat{p}) = p(1-p)/n \leq 1/(4n)$. Chebyshev gives $P(|\hat{p} - p| \geq 0.02) \leq 1/[4n(0.02)^2]$. Require this ≤ 0.05 , so $n \geq 12500$.

Marking points. 1 mark variance; 1 mark uniform bound; 2 marks Chebyshev inequality; 1 mark n

Common errors. Using a normal approximation without permission.

Final self-check. Substitution $n=12500$ makes the upper error bound exactly 0.05.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 8 min.

3(d) (5 marks)

Recognition signal	Multiple tests require controlling the family-wise error, not comparing each to 0.01.
First key step	Divide α by the number of tests.
Historical basis	2020 Q4(c); 2022 Q5(c)
Ability/method	Bonferroni correction; K43

Course-method solution. The adjusted threshold is $0.01/100 = 0.0001$. Only 0.00002 and 0.000005 are below it.

Marking points. 2 marks threshold; 2 marks comparisons; 1 mark conclusion

Common errors. Using 0.01 for each test.

Final self-check. Exactly two values are below 10^{-4} .

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 5 min.

Total: 20 marks

Question 4

4(a) (6 marks)

Recognition signal	A quadratic RSS can be minimized by derivatives or completing the square.
First key step	Eliminate the intercept using the sample means.
Historical basis	2020 Q5(b); 2023 Q6(b)
Ability/method	OLS derivation; K48

Course-method solution. Minimize RSS. For fixed β_1 , the minimizing β_0 is $\bar{y} - \beta_1 \bar{x}$. Substitution gives $RSS = S_{yy} - 2\beta_1 S_{xy} + \beta_1^2 S_{xx}$, minimized at $\hat{\beta}_1 = S_{xy}/S_{xx}$. Thus $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.

Marking points. 2 marks intercept condition; 2 marks reduced RSS; 1 mark slope; 1 mark intercept

Common errors. Writing formulas without proving minimization.

Final self-check. The fitted line passes through $(\bar{x}, \bar{y})$.

Difficulty allocation: A:3, B:2, C:1; estimated normal time 9 min.

4(b) (5 marks)

Recognition signal	The parameter determines the support, so the maximum occurs at a boundary.
First key step	State the zero-likelihood region.
Historical basis	2020 Q5(d)
Ability/method	boundary MLE; K54

Course-method solution. $L(\theta) = \theta^n$ if $\theta \geq x_{\max}$ and all data are non-negative, otherwise 0. Since θ^n decreases, the smallest feasible θ maximizes it: $\hat{\theta} = x_{\max}$.

Marking points. 2 marks likelihood/support; 2 marks monotonicity/boundary; 1 mark estimator

Common errors. Differentiating $\log L = -n \log \theta$ and setting it to zero.

Final self-check. The estimate must be at least every observation.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 7 min.

4(c) (4 marks)

Recognition signal	Conjugacy is recognized by matching the gamma kernel.
First key step	Ignore constants independent of θ and collect powers/rates.
Historical basis	2020 Q6(c); 2022 Q6(e) related conjugacy method
Ability/method	Gamma conjugate update; K55

Course-method solution. The likelihood is $\theta^n \exp(-\theta \sum x_i)$. Multiplying by $\theta^{a-1} \exp(-b\theta)$ gives posterior kernel $\theta^{a+n-1} \exp[-(b + \sum x_i)\theta]$. Thus $\theta|x \sim \text{Gamma}(a+n, b + \sum x_i)$.

Marking points. 1 mark likelihood; 1 mark prior product; 1 mark kernel; 1 mark identified posterior

Common errors. Using scale instead of stated rate.

Final self-check. Posterior shape and rate both increase by data information.

Difficulty allocation: A:1, B:1, C:1, D:1; estimated normal time 6 min.

4(d) (5 marks)

Recognition signal	A covariance plus bounded Y can be converted into a variance bound.
First key step	Use Cauchy/correlation first, then the range variance bound.
Historical basis	2020 Q6(d)
Ability/method	non-trivial variance lower bound; K33, K27, K28

Course-method solution. $\text{Cov}(X, Y) = 4 - 6 = -2$. Since $\text{Corr}^2 \leq 1$, $\text{Var}(X) \geq \text{Cov}^2 / \text{Var}(Y) = 4 / \text{Var}(Y)$. For Y in $[2, 5]$, $\text{Var}(Y) \leq (5-2)^2/4 = 9/4$. Hence $\text{Var}(X) \geq 16/9$ and $\text{sd}(X) \geq 4/3$.

Marking points. 1 mark covariance; 1 mark correlation bound; 1 mark bounded-variance result; 2 marks combine

Common errors. Using $\text{Var}(Y) \geq 9/4$; the inequality direction is opposite.

Final self-check. Squaring the final bound gives $16/9$.

Difficulty allocation: A:2, C:1, D:2; estimated normal time 8 min.

Total: 20 marks

MATH40005

Probability and Statistics

Revised Mock Examination 5 - Solutions and Marking Guide

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers
Difficulty calibration	A 32, B 20, C 12, D 16
Paper role	Seven-year gap paper: measurability, preimage sigma-algebras, scale mixtures, mixed transformations, exact tests, quartiles, bounds and estimator construction.

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

1(a) (2 marks)

Recognition signal	Indicator identities reduce to membership equivalence.
First key step	Write the two “equals 1 iff” statements.
Historical basis	2023 Q2(a)(i)
Ability/method	indicator/preimage identity; K02

Course-method solution. Both sides equal 1 exactly when $f(u)$ belongs to A , equivalently u belongs to $f^{-1}(A)$, and equal 0 otherwise.

Marking points. 2 marks equivalence by cases

Common errors. Confusing image and preimage.

Final self-check. Test a point inside and outside the preimage.

Difficulty allocation: A:2; estimated normal time 3 min.

1(b) (6 marks)

Recognition signal	Preimages preserve set operations exactly.
First key step	Verify the three sigma-algebra axioms using preimage identities.
Historical basis	2023 Q2(a)(ii)
Ability/method	preimage sigma-algebra proof; K02

Course-method solution. Preimages satisfy $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(A^c) = [f^{-1}(A)]^c$, and $f^{-1}(\cup A_i) = \cup f^{-1}(A_i)$. Since G has the corresponding closure properties, the collection of preimages is a sigma-algebra.

Marking points. 1 mark empty set; 2 marks complements; 2 marks countable unions; 1 mark conclusion

Common errors. Trying to prove f is one-to-one; it is unnecessary.

Final self-check. No probability measure is needed.

Difficulty allocation: A:2, B:2, D:2; estimated normal time 8 min.

1(c) (4 marks)

Recognition signal	The standard parity construction separates pairwise from joint independence.
First key step	Use two independent bits and their xor.
Historical basis	2026 Q2(a)
Ability/method	pairwise-versus-mutual independence construction; K04

Course-method solution. Let U, V be independent Bernoulli($1/2$), and $W = U \text{ xor } V$. Each pair is uniform on $\{0,1\}^2$, hence independent. But W is determined by U, V , and $P(U=0, V=0, W=0) = 1/4$ differs from $(1/2)^3$.

Marking points. 1 mark construction; 2 marks pairwise checks; 1 mark failure of mutual independence

Common errors. Giving three mutually independent variables.

Final self-check. The triple has only four possible outcomes, not eight.

Difficulty allocation: A:2, B:1, C:1; estimated normal time 5 min.

1(d) (4 marks)

Recognition signal	Finite measurability is checked by singleton preimages.
First key step	Test the preimage of 0.
Historical basis	2020 Q2(a); 2022 Q1(a)(ii)
Ability/method	finite-space measurability check; K02

Course-method solution. No, because $X^{-1}(\{0\}) = \{1,3\}$ is not in $\mathcal{F}$. A valid example is $Y=0$ on $\{1,2\}$ and $Y=1$ on $\{3,4\}$; its singleton preimages belong to $\mathcal{F}$.

Marking points. 2 marks counterexample preimage; 2 marks valid example/justification

Common errors. Checking only the image values.

Final self-check. A measurable variable must be constant on each atom $\{1,2\}$ and $\{3,4\}$.

Difficulty allocation: A:2, B:1, C:1; estimated normal time 5 min.

1(e) (4 marks)

Recognition signal	Repeated finite trials produce a Cartesian product; a finite power set has size $2^{ \Omega }$.
First key step	Count outcomes before subsets.
Historical basis	2021 Q1(b)(i)-(iii)
Ability/method	product probability space and power set; K01,K05

Course-method solution. $\Omega = \{1, \dots, 6\}^7$, $|\Omega| = 6^7$, $\mathcal{F} = \mathcal{P}(\Omega)$, $P(A) = |A|/6^7$. The power set has cardinality 2^{6^7} .

Marking points. 1 mark sample space; 1 mark probability space; 2 marks power-set cardinality

Common errors. Answering 2^7 or 6^{2^7} .

Final self-check. The event-space cardinality is vastly larger than the sample-space cardinality.

Difficulty allocation: A:2, B:2; estimated normal time 5 min.

Total: 20 marks

Question 2

2(a) (7 marks)

Recognition signal	A random scale mixture signals conditional expectation.
First key step	Condition on the scale variable Y .
Historical basis	2021 Q2(d)
Ability/method	MGF of a scale mixture; K12,K23,K24

Course-method solution. Condition on $Y=y$: $E[e^{uZ}|Y=y]=M_X(u\sqrt{y})=\exp(u^2 y/2)$. Thus $M_Z(u)=M_Y(u^2/2)=[\beta/(\beta-u^2/2)]^\alpha$, requiring $u^2/2 < \beta$, i.e. $|u| < \sqrt{2\beta}$.

Marking points. 2 marks conditioning; 1 mark normal MGF; 2 marks gamma MGF; 1 mark domain; 1 mark result

Common errors. Replacing $E[\exp(u^2 Y/2)]$ by $\exp(u^2 E[Y]/2)$.

Final self-check. At $u=0$ the MGF is 1.

Difficulty allocation: A:1, B:1, C:2, D:3; estimated normal time 10 min.

2(b) (6 marks)

Recognition signal	A product with a positive discrete multiplier is a finite mixture of scaled densities.
First key step	Condition on Y and use the CDF method.
Historical basis	2022 Q2(c)(i)
Ability/method	continuous-discrete mixture transform; K15,K20,K23

Course-method solution. Conditioning on Y , $F_Z(z)=\sum_{y=1}^{10} F_X(z/y)P(Y=y)$. Differentiating gives $f_Z(z)=\sum_{y=1}^{10} [1/y]f_X(z/y)P(Y=y)$. It is non-negative and integrates to $\sum_y P(Y=y)=1$ after the change $v=z/y$.

Marking points. 1 mark condition/CDF; 2 marks differentiation; 1 mark non-negativity; 2 marks integral check

Common errors. Omitting the Jacobian $1/y$.

Final self-check. Each component density integrates to its mixture weight.

Difficulty allocation: A:2, B:2, D:2; estimated normal time 9 min.

2(c) (7 marks)

Recognition signal	An ordered simplex has nested bounds.
First key step	Normalize first, then reuse the same bounds for the moment.
Historical basis	2020 Q3(c)
Ability/method	three-dimensional normalization and LOTUS; K17,K22,K20

Course-method solution. The ordered region has volume $1/6$, so $c=6$. Direct LOTUS gives $E[XYZ]=6 \int_0^1 \int_0^z \int_0^y xyz \, dx \, dy \, dz = 1/8$.

Marking points. 2 marks normalization/constant; 3 marks triple integral; 2 marks result

Common errors. Assuming X,Y,Z independent.

Final self-check. The order-statistic symmetry gives $c=3!=6$.

Difficulty allocation: A:2, B:1, C:1, D:3; estimated normal time 10 min.

Total: 20 marks

Question 3

3(a) (5 marks)

Recognition signal	Fixed-size classification without replacement gives a hypergeometric null.
First key step	Use the upper tail including the observed result.
Historical basis	2024 Q5(b) variant
Ability/method	exact hypergeometric test; K10,K44

Course-method solution. Under H_0 , $X \sim \text{Hypergeometric}(10,5,5)$. The p-value $P(X \geq 4) = [C(5,4)C(5,1) + C(5,5)C(5,0)]/C(10,5) = 26/252 = 13/126 \approx 0.103$. Fail to reject H_0 .

Marking points. 1 mark model/hypotheses; 2 marks tail; 1 mark value; 1 mark conclusion

Common errors. Using only $P(X=4)$.

Final self-check. The p-value exceeds 0.05.

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 7 min.

3(b) (4 marks)

Recognition signal	Tukey outliers are based on quartiles and 1.5 IQR fences.
First key step	Sort before splitting the sample.
Historical basis	2023 Q4(d)
Ability/method	quartiles and Tukey outliers; K46

Course-method solution. Sorted: 6.1,7.2,8.3,9.4,10.5,11.6,12.7,14.3,15.9,24.0. Lower half median $q_1=8.3$; upper half median $q_3=14.3$. IQR=6.0, fences $[-0.7,23.3]$, so 24.0 is the only outlier.

Marking points. 1 mark sorting/quartiles; 1 mark IQR; 1 mark fences; 1 mark outlier

Common errors. Including both middle observations in halves for even n .

Final self-check. The identified value must lie outside a computed fence.

Difficulty allocation: A:2, B:1, C:1; estimated normal time 6 min.

3(c) (3 marks)

Recognition signal	A reversal after aggregation is the signature of Simpson's paradox.
First key step	Compare subgroup weights and baseline risks.
Historical basis	2021 Q6(d); 2023 Q5(d)
Ability/method	Simpson reversal/confounding; K51

Course-method solution. This is Simpson's paradox. It occurs when treatment allocation is unbalanced across groups with different baseline risks; aggregation changes the weights and can reverse the within-group comparison. The risk group is a confounder.

Marking points. 1 mark name; 1 mark unequal weighting; 1 mark contextual explanation

Common errors. Calling it random error without explaining the weighting mechanism.

Final self-check. Stratified comparisons preserve the within-group direction.

Difficulty allocation: A:1, B:1, C:1; estimated normal time 4 min.

3(d) (5 marks)

Recognition signal	No distribution is specified, so use the variance bound rather than normal theory.
First key step	Translate confidence and width separately.
Historical basis	2021 Q4(b)
Ability/method	Chebyshev sample-size design; K45,K28

Course-method solution. Since only the bound $\sigma < 10$ is known, report the interval $\bar{X} \pm 10/\sqrt{0.01n} = \bar{X} \pm 100/\sqrt{n}$. Chebyshev gives coverage at least 0.99. Its total width is $200/\sqrt{n}$. The strict requirement $200/\sqrt{n} < 4$ is equivalent to $n > 2500$. Hence the minimum integer is $n = 2501$.

Marking points. 1 mark use known bound 10; 1 mark half-width $100/\sqrt{n}$; 1 mark total width; 1 mark strict inequality $n > 2500$; 1 mark minimum integer $n = 2501$.

Common errors. Using the unknown actual sigma in a reported interval; accepting $n = 2500$, for which the bound-based width is exactly 4; using z critical values.

Final self-check. $200/\sqrt{2501} < 4$, whereas $200/\sqrt{2500} = 4$.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 8 min.

3(e) (3 marks)

Recognition signal	The two errors condition on different truths.
First key step	State the decision and the true state.
Historical basis	2022 Q5(a); 2025 Q5(c)
Ability/method	testing error trade-off; K42

Course-method solution. Type I error is rejecting a true H_0 , with probability α for a level- α test. Type II error is failing to reject a false H_0 . With other design features fixed, increasing α enlarges the rejection region, increasing Type I error and generally decreasing Type II error/increasing power.

Marking points. 1 mark each definition; 1 mark trade-off

Common errors. Saying both decrease when α increases.

Final self-check. A larger rejection region makes rejection more frequent under both null and alternative.

Difficulty allocation: A:1, B:1, C:1; estimated normal time 4 min.

Total: 20 marks

Question 4

4(a) (6 marks)

Recognition signal	Known-variance lognormal location estimation becomes least squares on log data.
First key step	Discard theta-free terms and differentiate.
Historical basis	2022 Q6(a)
Ability/method	regular log-likelihood MLE; K53

Course-method solution. The log-likelihood in θ is constant $-[1/(2\sigma^2)]\sum(\log x_i - \theta)^2$. Differentiating gives $\sum(\log x_i - \theta) = 0$, so $\hat{\theta} = n^{-1}\sum \log x_i$. The second derivative is $-n/\sigma^2 < 0$.

Marking points. 2 marks log-likelihood; 2 marks score equation; 1 mark maximum check; 1 mark estimator

Common errors. Using the arithmetic mean of x_i .

Final self-check. The estimate is the sample mean of transformed observations.

Difficulty allocation: A:2, B:1, C:1, D:2; estimated normal time 8 min.

4(b) (4 marks)

Recognition signal	The intercept normal equation forces mean residual zero.
First key step	Differentiate RSS with respect to β_0 .
Historical basis	2022 Q6(d)
Ability/method	OLS mean-point property; K48

Course-method solution. The normal equation for the intercept is $\sum(y_i - \beta_0 - \beta_1 x_i) = 0$, so $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$. Hence at $x = \bar{x}$, $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} = \bar{y}$.

Marking points. 2 marks normal equation; 1 mark intercept relation; 1 mark point conclusion

Common errors. Using the result without justification.

Final self-check. The fitted residuals sum to zero.

Difficulty allocation: A:3, B:1; estimated normal time 5 min.

4(c) (4 marks)

Recognition signal	Moment equality requires a determinacy condition.
First key step	Use MGF existence near zero.
Historical basis	2025 Q3(c)(i)
Ability/method	moment determinacy; K34, K12

Course-method solution. If both MGFs exist on an open interval around 0, their Taylor expansions have the same moment coefficients, so the MGFs agree there. MGF uniqueness yields equal distributions.

Marking points. 1 mark condition; 2 marks MGF equality; 1 mark uniqueness

Common errors. Claiming the conclusion without a condition.

Final self-check. MGFs that agree near zero uniquely determine distributions.

Difficulty allocation: A:1, B:1, C:1, D:1; estimated normal time 5 min.

4(d) (3 marks)

Recognition signal	The bootstrap distribution is the distribution of replicate medians.
First key step	Resample with replacement at the original size.
Historical basis	2025 Q6(c)
Ability/method	bootstrap percentile interval; K52

Course-method solution. Generate many size- n resamples with replacement, compute each median, and take the empirical 2.5th and 97.5th percentiles.

Marking points. 1 mark resampling; 1 mark statistic; 1 mark percentiles

Common errors. Taking percentiles of raw data.

Final self-check. The endpoints come from bootstrap medians.

Difficulty allocation: A:2, B:1; estimated normal time 4 min.

4(e) (3 marks)

Recognition signal	Correct the upward bias of $\bar{X}^2$ using an unbiased variance estimate.
First key step	Compute $E[\bar{X}^2]$.
Historical basis	2021 Q4(a)(iii)
Ability/method	constructing an unbiased estimator; K32

Course-method solution. Since $E[\bar{X}^2] = \mu^2 + \sigma^2/n$ and $E[S^2] = \sigma^2$, the estimator $\bar{X}^2 - S^2/n$ is unbiased for μ^2 .

Marking points. 1 mark moment identity; 1 mark estimator; 1 mark unbiasedness

Common errors. Using $\bar{X}^2$ alone.

Final self-check. Taking expectations returns exactly μ^2 .

Difficulty allocation: A:2, B:1; estimated normal time 5 min.

Total: 20 marks

MATH40005

Probability and Statistics

Revised Mock Examination 6 - Solutions and Marking Guide

Time allowed	2 hours
Questions	4
Total marks	80
Format	Closed book; justify all answers
Difficulty calibration	A 32, B 20, C 12, D 16
Paper role	Seven-year mixed transfer pressure paper: combines distinct historical families, changes information order and context, and preserves the 2026 total difficulty.

- Answer all four questions.
- Start each question on a new page in an actual timed attempt.
- State assumptions and named results used.
- Statistical tables or stated critical values may be used where supplied.

Question 1

1(a) (6 marks)

Recognition signal	The factor $n-k$ signals a distinguished member among a complement set.
First key step	Identify the binomial coefficient identity $nC(n-1,k)=(n-k)C(n,k)$.
Historical basis	2022 Q1(b) method, altered parameter
Ability/method	algebraic and narrative double counting; K05

Course-method solution. Algebra: $n(4+1)^{n-1} = \sum nC(n-1,k)4^k = \sum C(n,k)(n-k)4^k$. Story: n people choose among four non-water drinks and water, with one designated driver who drinks water. Count by choosing driver first, or by choosing k non-water drinkers and then the driver among $n-k$ water drinkers.

Marking points. 3 marks algebra; 3 marks common-object story

Common errors. Giving only algebra or only a story.

Final self-check. Check $n=1$.

Difficulty allocation: A:2, B:2, C:1, D:1; estimated normal time 9 min.

1(b) (6 marks)

Recognition signal	Comparing posteriors under changed prevalence tests understanding of prior/base-rate effects.
First key step	Write Bayes separately for each prevalence.
Historical basis	2022 Q3(a); 2024 Q1(c)
Ability/method	Bayes/base-rate transfer; K08

Course-method solution. At prevalence 0.01: $0.95(0.01)/[0.95(0.01)+0.05(0.99)]=0.1610$. At 0.10: $0.095/[0.095+0.045]=0.6786$. The likelihood characteristics are unchanged, but the prior odds increase; Bayes updates prior odds by the same likelihood ratio.

Marking points. 2 marks first posterior; 2 marks second; 2 marks explanation

Common errors. Holding the denominator fixed.

Final self-check. Posterior odds equal prior odds times 19.

Difficulty allocation: A:2, B:2, C:1, D:1; estimated normal time 8 min.

1(c) (5 marks)

Recognition signal	Exact one-pair counting requires excluding further collisions.
First key step	Choose the pair first.
Historical basis	2023 Q1(b)
Ability/method	exact birthday collision; K07

Course-method solution. The probability is $C(12,2)*365*(364)_{-10}/365^{12}$.

Marking points. 1 mark pair; 1 mark common day; 2 marks distinct remaining days; 1 mark denominator

Common errors. Using the complement of no collision, which includes multiple collisions.

Final self-check. The numerator has 12 labeled people throughout.

Difficulty allocation: A:1, B:1, C:1, D:2; estimated normal time 7 min.

1(d) (3 marks)

Recognition signal	Independent finite stages form a Cartesian product.
First key step	Use ordered pairs.
Historical basis	2025 Q1(a)(i); 2021 Q1(b)
Ability/method	finite product probability space; K01

Course-method solution. $\Omega = \{1,2,3,4\} \times \{H,T\}$, $F = P(\Omega)$, and each singleton has probability $1/8$, so $P(A) = |A|/8$.

Marking points. 1 mark each component

Common errors. Omitting the event space.

Final self-check. There are 8 equally likely outcomes.

Difficulty allocation: A:3; estimated normal time 3 min.

Total: 20 marks

Question 2

2(a) (8 marks)

Recognition signal	A conditional expectation from a joint density needs normalization, marginalization and a conditional integral.
First key step	Compute f_X before forming the conditional density.
Historical basis	2022 Q3(c)(i)-(iii)
Ability/method	conditional expectation on triangular support; K17,K18,K23

Course-method solution. Normalization gives $c=24/7$. For $0 < x < 1$, $f_X(x) = c x [3/2 - x - x^2/2]$. The conditional density is $(y+1)/[3/2 - x - x^2/2]$ for $x < y < 1$. Hence $E[Y|X=x] = [5/6 - x^3/3 - x^2/2]/[3/2 - x - x^2/2]$, $0 < x < 1$.

Marking points. 2 marks c ; 2 marks marginal; 1 mark conditional density; 2 marks integral; 1 mark domain

Common errors. Using y from 0 to 1 rather than x to 1.

Final self-check. The conditional density integrates to 1.

Difficulty allocation: A:4, B:2, C:1, D:1; estimated normal time 11 min.

2(b) (7 marks)

Recognition signal	Random sum $\rightarrow$ condition on N .
First key step	Use both total expectation and total variance.
Historical basis	2024 Q3(b)(ii)-(iii)
Ability/method	compound-Poisson moments; K13,K23

Course-method solution. $E[S|N] = Nm$ and $\text{Var}(S|N) = Nv$. Therefore $E[S] = \lambda m$ and $\text{Var}(S) = E[Nv] + \text{Var}(Nm) = \lambda(v + m^2)$.

Marking points. 2 marks conditional moments; 2 marks mean; 3 marks total variance

Common errors. Omitting $\text{Var}(E[S|N])$.

Final self-check. For deterministic $Y_i = 1$, recover $\text{Poi}(\lambda)$.

Difficulty allocation: A:2, B:2, C:1, D:2; estimated normal time 9 min.

2(c) (5 marks)

Recognition signal	Combine an order-statistic CDF with a decreasing transformation.
First key step	Handle the inequality reversal under $-\log$.
Historical basis	2025 Q3(a); historical transformation chain
Ability/method	order statistic followed by monotone transform; K21,K20,K24

Course-method solution. For $t \geq 0$, $F_T(t) = P(M \geq e^{-t}) = 1 - P(M < e^{-t}) = 1 - e^{-nt}$. Thus $T \sim \text{Exponential}(n)$. For $t < 0$ the CDF is 0.

Marking points. 1 mark support; 2 marks transformation; 1 mark CDF; 1 mark identification

Common errors. Writing $e^{-t/n}$ as the survival rather than 1 minus it.

Final self-check. Mean should be $1/n$: larger n makes M closer to 1 and T smaller.

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 7 min.

Total: 20 marks

Question 3

3(a) (6 marks)

Recognition signal	Unequal means create extra between-mean variation.
First key step	Use the sum-of-squares identity.
Historical basis	2025 Q4(a)(i)-(ii)
Ability/method	non-identical-normal sample variance expectation; K30

Course-method solution. Using $\sum (X_i - \bar{X})^2 = \sum X_i^2 - n\bar{X}^2$ gives $E[S^2] = \sigma^2 + \sum (\mu_i - \bar{\mu})^2 / (n-1)$.

Marking points. 2 marks decomposition; 2 marks expectations; 2 marks result

Common errors. Applying the i.i.d. unbiasedness theorem.

Final self-check. Equal means remove the extra term.

Difficulty allocation: A:1, B:2, C:1, D:2; estimated normal time 8 min.

3(b) (7 marks)

Recognition signal	The tested linear combination is a sum, but independent sample means still add variances.
First key step	Standardize $\bar{X} + \bar{Y} - 360$ using the pooled variance.
Historical basis	2023 Q5(b)
Ability/method	pooled two-sample t test for a linear combination; K40

Course-method solution. The pooled variance is 4. Under H_0 , $T = (\bar{X} + \bar{Y} - 360) / [s_p \sqrt{1/n + 1/m}] = (357.5 - 360) / [2 \sqrt{1/16 + 1/9}] = -3.00$ with 23 df. The two-sided p-value is about 0.0064, so reject H_0 .

Marking points. 1 mark hypotheses/assumptions; 1 mark pooled variance; 2 marks statistic/df; 2 marks p-value; 1 mark conclusion

Common errors. Using $\bar{X} - \bar{Y}$; using df $n + m$.

Final self-check. The absolute statistic 3 is clearly beyond a typical 5% two-sided threshold.

Difficulty allocation: A:2, B:2, C:1, D:2; estimated normal time 10 min.

3(c) (4 marks)

Recognition signal	A high-confidence variance interval is wide and asymmetric.
First key step	Invert the chi-square inequalities carefully.
Historical basis	2026 Q3(d)(i)
Ability/method	chi-square variance interval; K37

Course-method solution. The interval is $(11 \cdot 3 / 26.757, 11 \cdot 3 / 2.603) = (1.233, 12.678)$.

Marking points. 1 mark pivot; 1 mark correct quantiles/order; 1 mark substitution; 1 mark result

Common errors. Swapping endpoints.

Final self-check. The interval contains $s^2 = 3$.

Difficulty allocation: A:2, B:1, C:1; estimated normal time 5 min.

3(d) (3 marks)

Recognition signal	Continuity permits the probability integral transform without randomisation.
First key step	Write p as a function of the null CDF.
Historical basis	2024 Q5(a)(i); 2026 Q3(e)
Ability/method	p-value definition and probability-integral-transform proof; K41

Course-method solution. For observed t , the upper-tail p-value is $p = 1 - F_0(t)$, the null probability of a result at least as large. Under H_0 , $F_0(T) \sim \text{Uniform}(0,1)$ by the probability integral transform; therefore $1 - F_0(T)$ is also $\text{Uniform}(0,1)$. Thus $P_{H_0}(p \leq \alpha) = \alpha$.

Marking points. 1 mark definition; 1 mark probability integral transform; 1 mark uniform conclusion

Common errors. Calling p the probability H_0 is true; stating uniformity without proof.

Final self-check. The result implies $P(p \leq \alpha) = \alpha$ under H_0 .

Difficulty allocation: A:2, B:1; estimated normal time 5 min.

Total: 20 marks

Question 4

4(a) (5 marks)

Recognition signal	A one-parameter linear-in-beta model is minimized by a normal equation.
First key step	Differentiate with respect to beta.
Historical basis	2024 Q6(c)
Ability/method	least squares for a transformed predictor; K48

Course-method solution. $RSS(\beta) = \sum (y_i - \beta x_i)^2$. Differentiate: $-2 \sum x_i^2 (y_i - \beta x_i) = 0$, giving $\hat{\beta} = \frac{\sum x_i^2 y_i}{\sum x_i^4}$ (provided the denominator is positive). The second derivative is $2 \sum x_i^4 > 0$.

Marking points. 2 marks RSS/derivative; 2 marks estimator; 1 mark minimum condition

Common errors. Using $\sum x_i y_i / \sum x_i^2$.

Final self-check. Units and the second derivative check the formula.

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 7 min.

4(b) (3 marks)

Recognition signal	Residuals should have approximately constant spread around zero.
First key step	Compare the vertical spread at small and large x.
Historical basis	2020 Q6(b) residual-diagnostic method; 2026 Q4(b)
Ability/method	heteroscedastic residual diagnosis and variance-stabilising remedy; K50

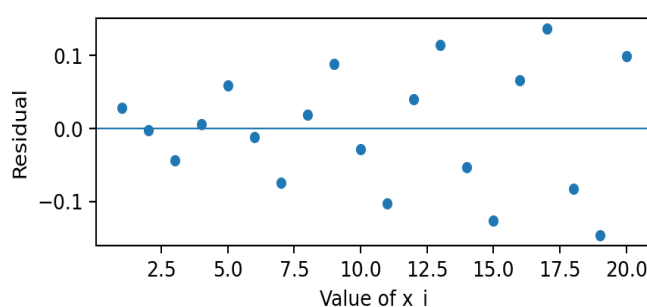
Course-method solution. The residual spread increases with x, giving a fan shape. This indicates non-constant error variance (heteroscedasticity), so the constant-variance regression assumption is doubtful. A variance-stabilising response transformation or an appropriately weighted regression may improve the fit; residuals must then be checked again.

Marking points. 1 mark fan/increasing spread; 1 mark heteroscedasticity conclusion; 1 mark justified remedy

Common errors. Diagnosing nonlinear mean curvature when the central trend remains near zero; relying only on R^2 .

Final self-check. After a successful remedy, the residual spread should be roughly constant across x.

Difficulty allocation: A:2, D:1; estimated normal time 4 min.



4(c) (5 marks)

Recognition signal	The likelihood is flat on the feasible support interval.
First key step	Translate all data constraints simultaneously.
Historical basis	2026 Q4(f)
Ability/method	non-unique support MLE; K54

Course-method solution. The likelihood is 1 when $\max x_{i-1} \leq \theta \leq \min x_i$ and zero otherwise. Hence all θ in $[x_{\max}-1, x_{\min}]$ are MLEs, provided the observed range is at most 1.

Marking points. 1 mark likelihood; 2 marks support inequalities; 1 mark MLE set; 1 mark feasibility

Common errors. Choosing only $x_{\min}$.

Final self-check. The interval length equals $1-(x_{\max}-x_{\min})$.

Difficulty allocation: A:2, B:1, C:1, D:1; estimated normal time 7 min.

4(d) (4 marks)

Recognition signal	Geometric likelihood is conjugate to a beta prior.
First key step	Collect exponents of θ and $1-\theta$.
Historical basis	2021 Q6(b)
Ability/method	Beta-geometric conjugacy; K55

Course-method solution. The likelihood is $\theta^n(1-\theta)^{\sum x_i}$. Multiplying by $\theta^{a-1}(1-\theta)^{b-1}$ gives $\text{Beta}(a+n, b+\sum x_i)$.

Marking points. 1 mark likelihood; 1 mark kernel; 1 mark parameters; 1 mark identification

Common errors. Using $b+\sum x_i$ rather than $b+\sum x_{i-n}$.

Final self-check. Both posterior parameters remain positive for valid data.

Difficulty allocation: A:1, B:1, C:1, D:1; estimated normal time 6 min.

4(e) (3 marks)

Recognition signal	Time trends can create spurious correlation.
First key step	Separate descriptive association from causal inference.
Historical basis	2024 Q5(d); 2026 Q4(d)
Ability/method	spurious correlation interpretation; K51

Course-method solution. The sample shows strong association, but neither causation nor a substantive relationship follows. Shared time trends, confounding or coincidence can produce a spurious correlation; causal evidence requires more than r .

Marking points. 1 mark association; 1 mark no causation; 1 mark explanation

Common errors. Saying correlation proves no relation; it does describe the sample.

Final self-check. Detrending or design information would be needed for stronger claims.

Difficulty allocation: A:2, B:1; estimated normal time 4 min.

Total: 20 marks