

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

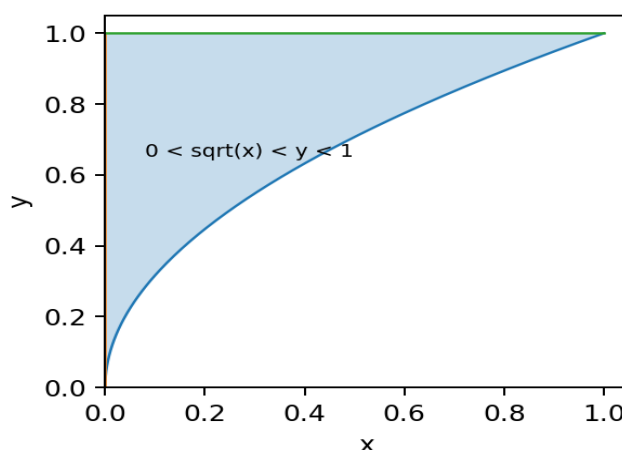
Course-method solution. The support is $0 < y < 1$, $0 < x < y^2$. Normalization gives $c \int_0^1 \int_0^{y^2} (x+y) dx dy = 1$, so $c = 20/7$. For $0 < x < 1$, $f_X(x) = c \int_{\sqrt{x}}^1 (x+y) dy = c(x/2 + 1/2 - x^{3/2})$. For $0 < y < 1$, $f_Y(y) = c \int_0^{y^2} (x+y) dx = c(y^4/2 + y^3)$. They 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