Mock Exam 1

MATHEMATICS **9709**

Paper 5 Probability & Statistics 1 **1 hour 15 minutes**

MARK SCHEME

Maximum Mark: 50

Published

This mark scheme is published as an aid to teachers and candidates, to indicate the requirements of the examination. It shows the basis on which Examiners were instructed to award marks. It does not indicate the details of the discussions that took place at an Examiners' meeting before marking began, which would have considered the acceptability of alternative answers.

Mark schemes should be read in conjunction with the question paper and the Principal Examiner Report for Teachers.

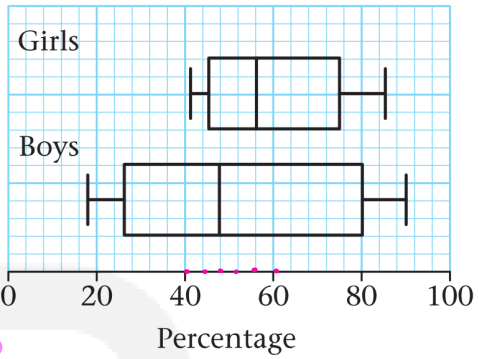
Cambridge International will not enter into discussions about these mark schemes.

Cambridge International is publishing the mark schemes for the October/November 2021 series for most Cambridge IGCSE™, Cambridge International A and AS Level components and some Cambridge O Level components.

By Shahab Yakoob

1. **STEM PLOT**

A group of students took a statistics test. The summary data for the percentage mark gained by boys and by girls is shown in the box plots opposite.



a Write down the percentage mark which 75% of the girls scored more than. ⇒ LQ = 45%.

b State the name given to this value.

c Compare and contrast the results of the boys and the girls.

✓ Girls performed better since the median % is higher [1+1+2]

✓ more dispersion in performance of Boys as IQR is more.

2. **MEAN + VARIANCE**

The test score, x, for a class of 14 students gives $\sum x = 919$ and $\sum x^2 = 60\,773$.

a) Calculate the mean and variance of the marks for this class.

$\bar{x} = \frac{\sum x}{n}$ $\quad sd^2 = \frac{\sum x^2}{n} - (\bar{x})^2$

$= \frac{919}{14}$ $\quad = \frac{60773}{14} - \left(\frac{919}{14}\right)^2$

$= 65.64$ $\quad = 32.318$

$= 32.32$

Another class of 15 students taking the same test scored a mean of 63.8, with a standard deviation of 5.58 marks.

b) Calculate $\sum x^2$ for the second class.

$\bar{x} = \frac{\sum x}{n}$ $\quad sd^2 = \frac{\sum x^2}{n} - (\bar{x})^2$

$63.8 = \frac{\sum x}{15}$ $\quad 5.58^2 = \frac{\sum x^2}{15} - (63.8)^2$

$\sum x = 957$ $\quad \sum x^2 = 61523.64$

$= 61500$

c) Calculate the mean and variance of the marks for all of the students in the two classes.

$\sum x_{overall} = 919 + 957$

$n_{overall} = 14 + 15$ [3+2+3]

$\bar{x}_{overall} = \frac{919 + 957}{14 + 15}$

$= 64.689$

$= 64.69$

$\sum x^2_{overall} = 60773 + 4537.486$

$= 122,296.646$

$sd^2_{overall} = \frac{122296.646}{29} - (64.689)^2$

$= 4217.13 - 4184.666$

$= 32.464$

$= 32.5$

Final Answers

(a) 65.6 (mean), 32.3 (variance) < 3 sf >

(b) $\sum x^2 = 61500$

(c) 64.7 (mean), 32.5 (variance)

3. **PROBABILITY**

The results of a survey of colours and types of cars are shown in the table.

	Saloon	Hatchback
Silver	65	59
Black	27	22
Other	16	19

124
49
35 ↓
208
108 100 →

One car is selected from the group at random.

a) Find the probability that the selected car is

i) a silver hatchback

ii) a hatchback

iii) a hatchback, given that it is silver.

b) Show that the type of car is not independent of its colour.

(a)(i) $\frac{59}{208}$

(ii) $\frac{100}{208}$

(iii) $\frac{59}{65+59}$ Silver or $\frac{59}{124}$

[1+1+2+3]

✓ P(Saloon) = $\frac{108}{208}$ P(HB) = $\frac{100}{208}$

✓ P(Silver) = $\frac{124}{208}$ P(Black) = $\frac{49}{208}$

✓ P(Saloon and Silver) = $\frac{65}{208}$

P(S and Silver) = P(Saloon) × P(Silver)

$\frac{65}{208} \stackrel{?}{=} \frac{108}{208} \cdot \frac{124}{208}$

$= \frac{13392}{43264}$

0.3125 ≠ 0.3095

Not indep

Shalini Yadav

The question can be approached in relation

to Saloon + Black

OR HB + Silver

OR HB + Black

4. DRV

Two fair cubical dice are thrown: one is red and one is blue. The random variable "M" represents the score on the red die minus the score on the blue die.

a Draw up the probability distribution table

b Write down E(M).

c Find Var(M).

[3+1+2]

RED

BLUE	1	2	3	4	5	6
1	0	1	2	3	4	5
2	−1	0	1	2	3	4
3	−2	−1	0	1	2	3
4	−3	−2	−1	0	1	2
5	−4	−3	−2	−1	0	1
6	−5	−4	−3	−2	−1	0

(a)

x^2	25	16	9	4	1	0	1	4	9	16	25
x	−5	−4	−3	−2	−1	0	1	2	3	4	5
p	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

(b) $E(M) = [(-5)(1) + (-4)(1) + (-3)(1) + (-2)(1) + (-1)(5) + 0(6) + 5(1) + (4)(1) + (3)(1) + (2)(1) + (1)(5)] \div 36$

$= 0$

(c) $E(M^2) = \dfrac{[25(1) + 16(2) + 9(3) + 4(4) + 1(5)](2) + 0(6)}{36}$

$= \frac{210}{36}$ OR $\frac{105}{18}$ OR $\frac{35}{6}$ OR $5\frac{5}{6}$ OR 5.83

$\text{Var}(M) = E(M^2) - (E(M))^2$

$= 5.83 - 0^2$

$= 5.83$

SY

TREE DIAG.

5. In a factory, machines A, B and C produce electronic components. Machine A produces 16% of the components, machine B produces 50% of the components and machine C produces the rest. Some of the components are defective. Machine A produces 4%, machine B 3% and machine C 7% defective components.

a Draw a tree diagram to represent this information.

Find the probability that a randomly selected component is

b produced by machine B and is defective,

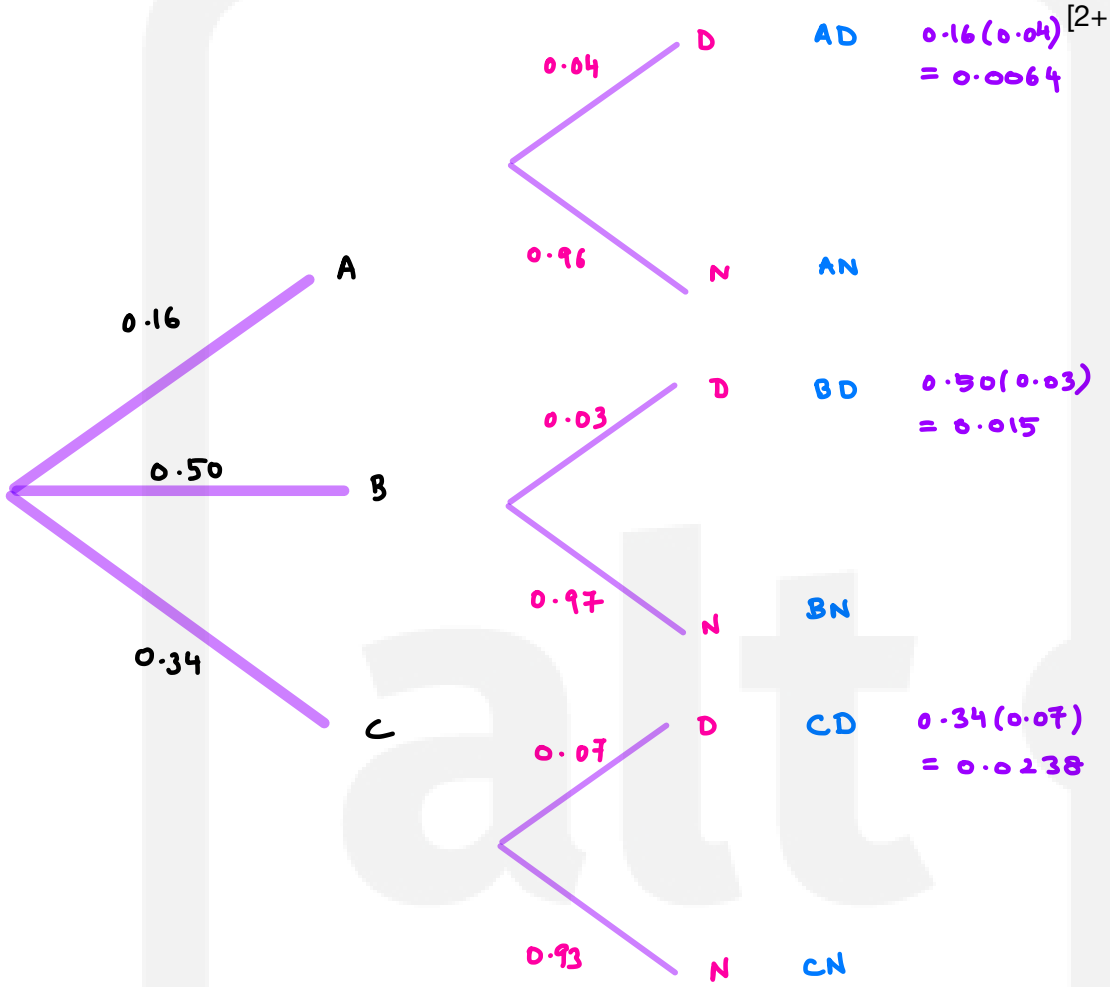
c defective.

Given that a randomly selected component is defective,

d find the probability that it was produced by machine B.

[2+1+2+2]

(a)



(b) BD → 0.015

(c) AD + BD + CD

0.0064 + 0.015 + 0.0238 = 0.0452

(d) Given → Defective

$$\frac{BD}{AD + BD + CD} = \frac{0.015}{0.0064 + 0.015 + 0.0238} = \frac{0.015}{0.0452} \text{ OR } \frac{75}{226}$$

OR 0.332

Shahrat Yakoob

PROBABILITY

6. In a college there are 100 students taking A level French, German or Spanish. Of these students, 64 are female and the rest are male. There are 50 French students of whom 40 are female and 30 German students of whom 10 are female.

Find the probability that a randomly chosen student

a is taking Spanish,

b is male, given that the student is taking Spanish.

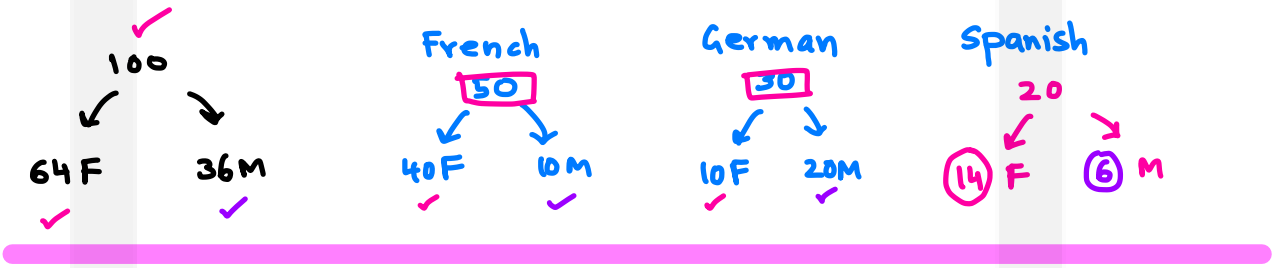
College records indicate that 70% of the French students, 80% of the German students and 60% of the Spanish students have applied for University.

A student is chosen at random.

c Find the probability that this student has applied for University.

d Given that the student had applied to University, find the probability that the student is studying French.

[1+2+2+2]



(a) $P(\text{Spanish}) = \frac{20}{100}$ or $\frac{1}{5}$

(b) $\frac{6}{14+6} = \frac{6}{20}$ or $\frac{3}{10}$

FU 0.5 (0.7) = 0.35

GU 0.3 (0.8) = 0.24

SU 0.2 (0.6) = 0.12

(c) FU + GU + SU

= 0.35 + 0.24 + 0.12

= 0.71

(d) $\frac{FU}{FU + GU + SU} = \frac{0.35}{0.35 + 0.24 + 0.12} = \frac{0.35}{0.71}$ or $\frac{35}{71}$ or 0.493

SY

BINOMIAL DIST.

7. On a production line, on average 6% of the bottles of lemonade are not filled properly.

a) If five bottles are examined, find the probability that exactly one of them is not filled properly.

b) If 2000 bottles are examined, use a normal approximation to find the probability that less than 100 of them are not filled properly.

[2+4]

(a) $X \sim B(5, 0.06)$

$P(X=1) = {}^5C_1 (0.94)^4 (0.06)^1$

$= 0.2342$

(b) $X \sim B(2000, 0.06)$

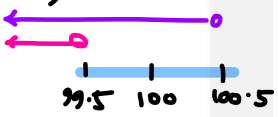
$n = 2000 \quad p = 0.06 \quad q = 0.94$

$np = 120 \quad nq = 1880 \quad npq = 112.8$

since $np > 5$ and $nq > 5$

$\Rightarrow X \sim N(120, 112.8)$

$P(X < 100)$ CC



$P(X < 99.5)$

$P\left(z < \frac{99.5 - 120}{\sqrt{112.8}}\right)$

$P(z < -1.930)$

$P(z > 1.930)$

$1 - \Phi(1.930)$

$1 - 0.9732$

0.0268

OR 2.68%

Shahab Yakoob

NORMAL DIST.

8. The heights of seedlings are normally distributed. Given that 10% of the seedlings are taller than 15 cm and 5% are shorter than 4 cm, find the mean and standard deviation of the heights.

$X \sim N(\mu, \sigma^2)$

[5]

$P(X > 15) = 0.10$

$P(X < 15) = 0.90$

$P\left(Z < \frac{15-\mu}{\sigma}\right) = 0.90$

$P(Z < 1.282) = 0.90$

$\frac{15-\mu}{\sigma} = 1.282$

$\frac{15-\mu}{1.282} = \sigma$ — [1]

$P(X < 4) = 0.05$

$P(X > 4) = 0.95$

$P\left(Z > \frac{4-\mu}{\sigma}\right) = 0.95$

$P(Z > -1.645) = 0.95$

$\frac{4-\mu}{\sigma} = -1.645$

$\frac{4-\mu}{-1.645} = \sigma$ — [2]

Solve Simultaneously

$\mu = 10.2$

$\sigma = 3.76$

SY

The End!

PROBABILITY & STATISTICS

Summary statistics

For ungrouped data:

$$\bar{x} = \frac{\Sigma x}{n}, \qquad \text{standard deviation} = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$$

For grouped data:

$$\bar{x} = \frac{\Sigma xf}{\Sigma f}, \qquad \text{standard deviation} = \sqrt{\frac{\Sigma(x-\bar{x})^2 f}{\Sigma f}} = \sqrt{\frac{\Sigma x^2 f}{\Sigma f} - \bar{x}^2}$$

Discrete random variables

$$\mathrm{E}(X) = \Sigma xp, \qquad \mathrm{Var}(X) = \Sigma x^2 p - \{\mathrm{E}(X)\}^2$$

For the binomial distribution $\mathrm{B}(n, p)$:

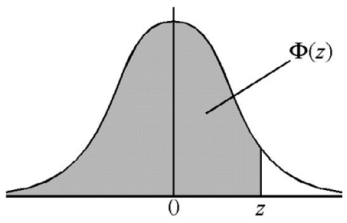
$$p_r = \binom{n}{r} p^r (1-p)^{n-r}, \qquad \mu = np, \qquad \sigma^2 = np(1-p)$$

THE NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1, then, for each value of z, the table gives the value of $\Phi(z)$, where

$$\Phi(z) = P(Z \leqslant z).$$

For negative values of z, use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
											ADD								
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621	2	5	7	9	12	14	16	19	21
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830	2	4	6	8	10	12	14	16	18
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015	2	4	6	7	9	11	13	15	17
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177	2	3	5	6	8	10	11	13	14
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319	1	3	4	6	7	8	10	11	13
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441	1	2	4	5	6	7	8	10	11
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545	1	2	3	4	5	6	7	8	9
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633	1	2	3	4	4	5	6	7	8
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706	1	1	2	3	4	4	5	6	6
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767	1	1	2	2	3	4	4	5	5
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817	0	1	1	2	2	3	3	4	4
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857	0	1	1	2	2	2	3	3	4
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890	0	1	1	1	2	2	2	3	3
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916	0	1	1	1	1	2	2	2	2
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936	0	0	1	1	1	1	1	2	2
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952	0	0	0	1	1	1	1	1	1
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964	0	0	0	0	1	1	1	1	1
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974	0	0	0	0	0	1	1	1	1
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981	0	0	0	0	0	0	0	1	1
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986	0	0	0	0	0	0	0	0	0

Critical values for the normal distribution

If Z has a normal distribution with mean 0 and variance 1, then, for each value of p, the table gives the value of z such that

$$P(Z \leqslant z) = p.$$

p	0.75	0.90	0.95	0.975	0.99	0.995	0.9975	0.999	0.9995
z	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

BLANK PAGE

BLANK PAGE