MATH036501

This question paper consists of 10 printed pages, each of which is identified by the reference **MATH0365**.

Graph paper is provided.
A formulae sheet is attached.
A normal table is attached.
Only approved basic scientific calculators may be used.

Examination for the Module MATH0365
(January 2006)

FOUNDATION PROBABILITY AND STATISTICS

Time allowed: **2 hours**

Attempt **ALL** questions in Section A and **TWO** questions from Section B.

Questions A1 to A10 require you to write down a single letter answer.
Questions A11 to A20 require you to write down a short explanation.

Sections A and B are each worth 50% of the examination marks.
Questions A11 to A20 are each worth 1.5 times the marks of questions A1 to A10.

SECTION A
Attempt all questions in Section A.
Questions A1 to A10 require you to write down a single letter answer.

A1. The data below refer to the closing price (in £) of 9 different shares traded on a stock exchange:

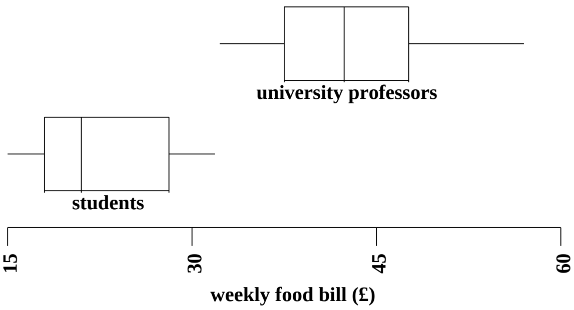
$$113, 118, 121, 126, 127, 131, 141, 147, 148$$

Which of the following are true?

(i) The lowest price is £118.
(ii) The median price is £127.
(iii) The range of the prices is £35.

A: all of these **B**: (iii) only **C**: (ii) (iii)

A2. The two box plots shown below summarise the weekly food bills of 13 students and 13 university professors:

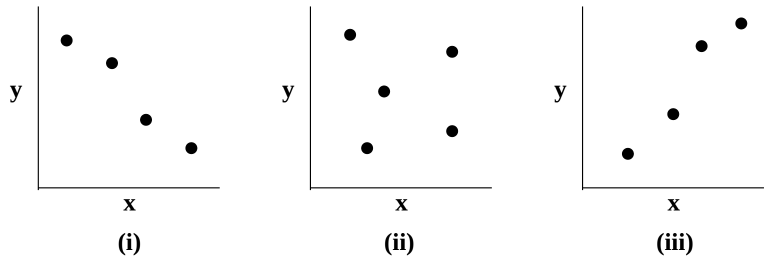


Which of the following are true?

(i) The median weekly food bill of students is greater than the median weekly food bill of university professors.
(ii) The interquartile range (IQR) of weekly food bills is similar for students and university professors.
(iii) students' weekly food bills show skewness.

A: (i) only **B**: (ii) (iii) **C**: all of these

A3. Which of the following scatter diagrams show an approximate positive linear relationship between the variables x and y?



A: (i) only **B**: (ii) only **C**: (iii) only

A4. A group of 30 business men spent between 5 and 25 minutes studying and investing in the stock market. The variables "hours studied", x, and "profit earned (in £)", y, follow a regression line $y = 8.51 + 5.01x$. Assuming the regression line is an appropriate summary of the data gathered, which of the following are true?

(i) The regression line can be used to predict a profit of $y = £83.66$ for a business man that studied the market for $x = 15$ minutes.

(ii) The regression line shows that a profit of $y = £2000$ can be achieved if a business man studies the market for approximately $x = 398$ minutes.

(iii) For the business men in the group, each extra minute studying the market was associated with an average of $£8.51$ extra profit.

A: (i) only **B**: (i) (ii) **C**: (i) (iii)

A5. If events A and B are such that $\mathrm{P}(A) = 0.3$, $\mathrm{P}(B) = 0.75$ and $\mathrm{P}(A|B) = 0.25$. Which of the following are true?

(i) Events A and B are mutually exclusive.

(ii) $\mathrm{P}(A') = 0.6$.

(iii) $\mathrm{P}(A \cap B) = 0.1875$.

A: (i) (iii) **B**: all of these **C**: (iii) only

A6. A small business employs eight permanent staff and four temporary staff. Two members of staff will be asked to work extra hours. Assuming each of the twelve members of staff are equally likely to be asked, what is the probability that both will be permanent members of staff?

A: $\frac{14}{33}$ **B**: $\frac{33}{14}$ **C**: $\frac{1}{11}$

A7. The discrete random variable X has expectation $\mathrm{E}(X) = 2$ and variance $\mathrm{Var}(X) = 3$. Define a random variable $Y = 4X + 10$. Which of the following are true?

(i) $\mathrm{Var}(Y) = 58$.
(ii) $\mathrm{E}(Y) = 18$.
(iii) $\mathrm{Var}(Y) = 48$.

A: (i) (ii) **B**: (ii) (iii) **C**: (ii) only

A8. The discrete random variable W follows a $\mathrm{B}(5, 0.25)$ distribution. Which of the following are true?

(i) $\mathrm{P}(W = 3) = 0.0879$.
(ii) $\mathrm{E}(W) = 1.25$.
(iii) $\mathrm{Var}(W) = 0.3125$.

A: (iii) only **B**: (i) (ii) **C**: all of these

A9. For $Z \sim \mathrm{N}(0, 1)$ which of the following are true?

(i) $\mathrm{P}(Z > z) = 1 - \mathrm{P}(Z < z)$.
(ii) $\mathrm{P}(Z < 1.43) = \Phi(1.43) = 0.9236$.
(iii) $\mathrm{P}(a < Z < b) = \Phi(a) - \Phi(b)$.

A: all of these **B**: (i) (iii) **C**: (i) (ii)

A10. The length of bolts produced by an engineering firm is modelled by a normal distribution with $\mu = 15$ mm and standard deviation $\sigma = 2.5$ mm. What is the probability that a bolt will be longer than 17 mm?

A:0.2119 **B**: 0.7881 **C**: 0.1056

Questions A11 to A20 require you to write down a short explanation.

A11. Explain what is meant by quantitative continuous data, and give an example.

A12. Explain what is meant by an outlier.

A13. Draw a scatter diagram that shows two variables, x and y, with a correlation coefficient, r, close to zero.

A14. Briefly explain the difference between correlation and causation, and give an example of two variables that are likely to be correlated but not causally related.

A15. For a statistical experiment, explain what is meant by an event, and give an example involving an ordinary six-sided die.

A16. Explain what is meant if events A and B are exhaustive, and give a probability statement that they must satisfy.

A17. Explain what is meant by the expectation of a discrete random variable, and given an example involving an ordinary six-sided die.

A18. Describe the situation where a discrete random variable can be modelled by a uniform distribution, and give an example.

A19. Explain the difference between a discrete random variable and a continuous random variable.

A20. Draw a sketch to explain how the probability density function, $\mathrm{f}(x)$, of a continuous random variable, X, can be used to calculate $\mathrm{P}(a < X < b)$.

CONTINUED...

SECTION B
Attempt TWO questions from Section B.

B1. (a) The following data refer to the waiting times (in minutes) of 17 adults at a local doctor's surgery:

$$39, 4, 59, 11, 54, 42, 5, 17, 15, 4, 8, 10, 8, 20, 17, 1, 43.$$

Construct a stem-and-leaf diagram to display the data.

(b) Calculate the median waiting time, the lower quartile, Q_1, and the upper quartile, Q_3. Construct a box plot to display the data.

(c) Calculate the quartile coefficients of skewness, and identify any outliers.

(d) What can you say about waiting times at the doctors surgery?

B2. (a) The manager of an insurance firm records the time (to the nearest minute) a random sample of 70 staff take for lunch. The data are summarised in the table below.

Time	Frequency f_i	Midpoint y_i	$y_i f_i$	$y_i^2 f_i$
5.5-15.5	4	10.5	42	441
15.5-25.5	15	20.5	307.5	6303.75
25.5-35.5	28	30.5	854	26047
35.5-55.5	20	40.5	810	32805
55.5-85.5	3	70.5	211.5	14910.75
	$n = \sum_{i=1}^{5} f_i = 70$		$\sum_{i=1}^{5} y_i f_i = 2225$	$\sum_{i=1}^{5} y_i^2 f_i = 80507.5$

(i) Produce a histogram to display the data.

(ii) Estimate the mean time a member of staff takes for lunch, $\bar{y}$, and the variance in the time a member of staff takes for lunch, s_y^2.

(b) The manager suspects the time a member of staff takes for lunch might be linearly related to their salary. The following data are collected, where x denotes the staff member's salary (in units of £1000), and y denotes the time a member of staff takes for lunch (to the nearest minute).

x	30	25	20	20	15	50
y	15	20	25	22	30	80

$$S_{xx} = 783.33, \quad S_{yy} = 2890, \quad S_{xy} = 1220.$$

(i) Construct a scatter diagram to display the data.

(ii) Calculate the correlation coefficient, r.

(iii) What can you say about the relationship between salary and time taken for lunch?

B3. I invite you to toss three fair coins. All have sides labelled 'heads' and 'tails'. Once the coins have been tossed I record the number of 'heads' obtained and the number of 'tails' obtained.

(a) I define the events:

A: Exactly two heads are obtained.
B: At least two tails are obtained.
C: Exactly one tail is obtained.

(i) Calculate $\mathrm{P}(A)$, $\mathrm{P}(B)$ and $\mathrm{P}(C)$.

(ii) Calculate $\mathrm{P}(A \cap B)$ and $\mathrm{P}(A \cup B)$. Are events A and B mutually exclusive?

(iii) Calculate $\mathrm{P}(A \cap C)$ and $\mathrm{P}(A|C)$. Are events A and C statistically independent?

(b) Let the random variable X denote the number of tails obtained when the three coins are tossed.

(i) Determine the probability distribution of X.

(ii) Calculate $\mathrm{E}(X)$.

(iii) Calculate $\mathrm{E}(X^2)$ and $\mathrm{Var}(X)$.

(c) I will charge you £3 to toss the coins and pay you £$2X$ (i.e. twice the number of tails obtained). Let M denote your expected profit from tossing the coins. Write down the equation connecting M and X. Calculate your expected profit from tossing the coins, $\mathrm{E}(M)$, and the variance in your profit from tossing the coins, $\mathrm{Var}(M)$.

B4. (a) A theme park has 4 large roller coasters. From past experience the probability that a roller coaster will break down on any given day is 0.05. Use a suitable model to calculate the probability that exactly 2 of the 4 roller coasters will break down on a given day. What assumptions are you making?

(b) The theme park usually has one major accident every 10 years. Use a suitable model to calculate the probability that the theme park will have less than 3 major accidents over the next 15 years. What assumptions are you making?

(c) The queue for the theme park's most popular roller coaster contains 350 people. From past experience, the probability a person in the queue will faint whilst waiting is 0.02. Using a suitable approximation, determine the probability that exactly 10 of the people in the queue faint whilst waiting.

(d) The theme park's most popular roller coaster has a height restriction of 1.4 metres (i.e. people who are less than 1.4 metres tall are not able to ride the roller coaster). The theme park owners estimate the mean height of its customers to be $\mu = 1.7$ metres with a standard deviation of $\sigma = 0.24$ metres. Assuming customers' heights can be modelled by a normal distribution, calculate the probability that a randomly chosen customer will not be able to ride the roller coaster.

FORMULAE SHEET

Representation and summary of data

$$\text{median} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{if } n \text{ is odd,} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+2}{2}\right)}}{2} & \text{if } n \text{ is even.} \end{cases}$$

$$\text{mean} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

$$\bar{x} \text{ (grouped data)} = \frac{\sum_{i=1}^{m} x_i f_i}{n}, \quad n = \sum_{i=1}^{m} f_i.$$

$$\text{lower quartile} = Q_1 = x_{\left(\frac{n+3}{4}\right)}.$$

$$\text{upper quartile} = Q_3 = x_{\left(\frac{3n+1}{4}\right)}.$$

$$\text{range} = x_{(n)} - x_{(1)}.$$

$$\text{IQR} = Q_3 - Q_1.$$

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right].$$

$$s^2 \text{ (grouped data)} = \frac{1}{n-1}\left[\sum_{i=1}^{m} x_i^2 f_i - \frac{\left(\sum_{i=1}^{m} x_i f_i\right)^2}{n}\right], \quad n = \sum_{i=1}^{m} f_i.$$

$$s = \sqrt{s^2}.$$

$$\text{if } y_i = \frac{x_i - a}{b} \qquad \bar{x} = a + b\bar{y}, \quad s_x^2 = b^2 s_y^2.$$

$$\text{quartile coefficient of skewness} = \frac{Q_3 - (2 \times \text{median}) + Q_1}{Q_3 - Q_1}.$$

$$\text{outliers are outside the limits} \qquad \left[\frac{1}{2}(5Q_1 - 3Q_3), \frac{1}{2}(5Q_3 - 3Q_1)\right].$$

Correlation and regression

$$S_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}.$$

$$S_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}.$$

$$S_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n}.$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}.$$

In the regression line $y = a + bx$ $\qquad b = \frac{S_{xy}}{S_{xx}}, \quad a = \bar{y} - b\bar{x}.$

CONTINUED...

Probability

$$
\begin{aligned}
\mathrm{P}(A') &= 1-\mathrm{P}(A).\\
\mathrm{P}(A\cup B) &= \mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B).\\
\mathrm{P}(A|B) &= \frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}.\\
n! &= n\times(n-1)\times(n-2)\times\cdots\times 2\times 1,\\
0! &= 1.\\
{}^nP_r &= \frac{n!}{(n-r)!}.\\
\binom{n}{r} &= \frac{n!}{(n-r)!r!}.
\end{aligned}
$$

A box contains N balls. The balls are of k different types. There are N_1 balls of type 1, N_2 balls of type 2 etc. (with $\sum_{i=1}^k N_i = N$). The probability that the sample contains exactly n_1 balls of type 1, n_2 balls of type 2 etc. (with $\sum_{i=1}^k n_i = n$) is:

$$\frac{\binom{N_1}{n_1}\times\binom{N_2}{n_2}\times\cdots\times\binom{N_k}{n_k}}{\binom{N}{n}}.$$

Discrete random variables

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_x x\mathrm{P}(X=x).\\
\mathrm{E}(X^2) &= \sum_x x^2\mathrm{P}(X=x).\\
\mathrm{Var}(X) &= \mathrm{E}(X^2)-[\mathrm{E}(X)]^2.\\
\mathrm{E}(aX+b) &= a\mathrm{E}(X)+b.\\
\mathrm{Var}(aX+b) &= a^2\mathrm{Var}(X).
\end{aligned}
$$

discrete uniform probability function $\quad \mathrm{f}(x)=\frac{1}{n}, \quad (x=1,\ldots,n).$

if X is discrete uniform then $\quad \mathrm{E}(X)=\frac{1}{2}(n+1), \quad \mathrm{Var}(X)=\frac{1}{12}(n^2-1).$

if $X\sim\mathrm{B}(n,p)$ $\quad \mathrm{f}(x)=\binom{n}{x}p^x(1-p)^{n-x}, \quad (x=0,1,2,\ldots,n).$

if $X\sim\mathrm{B}(n,p)$ $\quad \mathrm{E}(X)=np, \quad \mathrm{Var}(X)=np(1-p).$

if $X\sim\mathrm{Po}(\lambda)$ $\quad \mathrm{f}(x)=\frac{\mathrm{e}^{-\lambda}\lambda^x}{x!}, \quad (x=0,1,2,\ldots).$

if $X\sim\mathrm{Po}(\lambda)$ $\quad \mathrm{E}(X)=\lambda, \quad \mathrm{Var}(X)=\lambda.$

if $X\sim\mathrm{B}(n,p)$ with $n>50$ and $p<0.1$ $\quad X$ is approximately $\mathrm{Po}(\lambda=np)$.

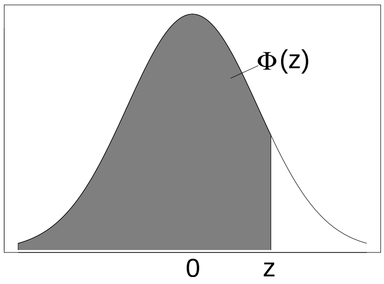
The normal distribution

if $X\sim\mathrm{N}(\mu,\sigma^2)$ $\quad \mathrm{E}(X)=\mu, \quad \mathrm{Var}(X)=\sigma^2.$

if $X\sim\mathrm{N}(\mu,\sigma^2)$ $\quad Z=\frac{X-\mu}{\sigma}\sim\mathrm{N}(0,1).$

TABLE OF $\Phi(z)$ FOR THE STANDARD NORMAL DISTRIBUTION

For $Z \sim \mathrm{N}(0,1)$, the table shows $\Phi(z) = \mathrm{P}(Z < z)$ where $z \geq 0$.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
+0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
+0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
+0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
+0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
+0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
+0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
+0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
+0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8133
+0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
+1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
+1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
+1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
+1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
+1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
+1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
+1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
+1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
+1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
+1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
+2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
+2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
+2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
+2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
+2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
+2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
+2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
+2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
+2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
+2.9	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
+3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
+3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
+3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
+3.3	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
+3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

END