Topic Test: OxfordAQA AS Further Mathematics

Statistics

Name: ________________

Class: ________________

Date: ________________

Time: **71 minutes**

Marks: **59 marks**

Comments:

1 It is proposed to introduce, for all males at age 60, screening tests, A and B, for a certain disease.

Test B is administered only when the result of Test A is inconclusive.

It is known that 10% of 60-year-old men suffer from the disease.

For those 60-year-old men suffering from the disease:

- Test A is known to give a positive result, indicating a presence of the disease, in 90% of cases, a negative result in 2% of cases and a requirement for the administration of Test B in 8% of cases;
- Test B is known to give a positive result in 98% of cases and a negative result in 2% of cases.

For those 60-year-old men not suffering from the disease:

- Test A is known to give a positive result in 1% of cases, a negative result in 80% of cases and a requirement for the administration of Test B in 19% of cases;
- Test B is known to give a positive result in 1% of cases and a negative result in 99% of cases.

(a) Draw a tree diagram to represent the above information.

(4)

(b) (i) Hence, or otherwise, determine the probability that:

(A) a 60-year-old man, suffering from the disease, tests negative;

(B) a 60-year-old man, not suffering from the disease, tests positive.

(2)

(ii) A random sample of ten thousand 60-year-old men is given the screening tests. Calculate, to the nearest 10, the number who you would expect to be given an **incorrect** diagnosis.

(2)

(c) Determine the probability that:

(i) a 60-year-old man suffers from the disease given that the tests provide a positive result;

(ii) a 60-year-old man does not suffer from the disease given that the tests provide a negative result.

(5)

(Total 13 marks)

2 The continuous random variable X has a rectangular distribution defined by

$$f(x) = \begin{cases} k & -3k \leq x \leq k \\ 0 & \text{otherwise} \end{cases}$$

(a) (i) Sketch the graph of f.

(2)

(ii) Hence show that $k = \frac{1}{2}$.

(2)

(b) Find the **exact** numerical values for the mean and the standard deviation of X.

(3)

(c) (i) Find $P\left(X \geq -\frac{1}{4}\right)$.

(2)

(ii) Write down the value of $P\left(X \neq -\frac{1}{4}\right)$.

(1)

(Total 10 marks)

3 The random variable X denotes the number of trials necessary in order to obtain the first success.

(a) State **three** conditions which must apply in order that X may be modelled by a geometric distribution.

(3)

(b) The discrete random variable Y is such that $Y \sim \text{Geo}(p)$ and $P(Y = 1 \text{ or } 2) = 0.2775$.

(i) Determine the value of p.

(4)

(ii) Hence find the values of $E(Y)$ and $\text{Var}(Y)$.

(2)

(Total 9 marks)

4 The table shows the probability distribution for the number of weekday (Monday to Friday) morning newspapers, X, purchased by the Reed household per week.

x	0	1	2	3	4	5
P ($X = x$)	0.16	0.15	0.25	0.25	0.15	0.04

(a) Find values for E(X) and Var(X).

(3)

(b) The number of weekday (Monday to Friday) evening newspapers, Y, purchased by the same household per week is such that

E(Y) = 2.0, Var(Y) = 1.5 and Cov(X, Y) = –0.43

Find values for the mean and variance of:

(i) $S = X + Y$;

(ii) $D = X - Y$.

(5)

(c) The total cost per week, L, of the Reed household's weekday morning and evening newspapers may be assumed to be normally distributed with a mean of £2.31 and a standard deviation of £0.89.

The total cost per week, M, of the household's weekend (Saturday and Sunday) newspapers may be assumed to be independent of L and normally distributed with a mean of £2.04 and a standard deviation of £0.43.

Determine the probability that the total cost per week of the Reed household's newspapers is more than £5.

(5)
(Total 13 marks)

5 X is a discrete uniform distribution where x can take the value 1, 2 or 3

(a) Find P(X > 1)

(1)

(b) Derive the probability generating function, $G_X(t)$, of X.

(2)

(c) Y is a discrete random variable with probability generating function $G_Y(t)$ where

$$G_Y(t) = 0.6 + 0.4t$$

Given that X and Y are independent, find P($X + Y = 1$)

(3)
(Total 6 marks)

6 On a production line, components are checked for damage.

The random variable X represents the number of components that are checked before the first damaged component is found and can be modelled by a geometric distribution with $p = 0.045$

(a) (i) Find the probability that exactly eight products are checked before the first damaged product is found.

(2)

(ii) Find the probability that more than five products are checked before the first damaged product is found.

(2)

(b) Derive the probability generating function, $G_X(t)$, of X in the form $\frac{at}{1-bt}$

where a and b are constants.

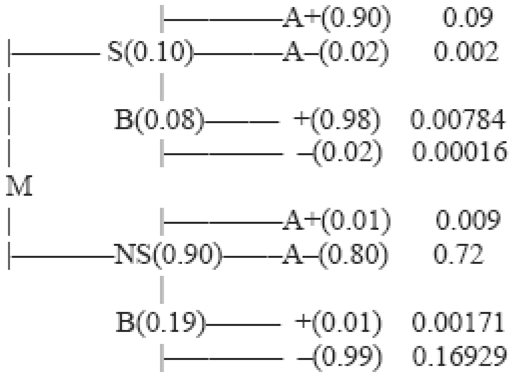
Fully justify your answer.

(4)

(Total 8 marks)

Mark schemes

1 (a)



S and NS with Ps or %s

B1

2 × (A+ and A–) with Ps or %s

B1

2 × (B) with Ps or %s

B1

2 × (B+ and B–) with Ps or %s

B1

Basic shape with labels, but without Ps or %s

(B2, 1)

Note:
The following BF and AF marks are
dependent on an essentially correctly-shaped tree diagram

4

(b) (i) **(A)** P(S and –) = 0.002 + 0.00016 = **0.00216**

F on (a); otherwise CAO

B1F

(B) P(NS and +) = 0.009 + 0.00171 = **0.01071**

F on (a); otherwise AWRT 0.0107

B1F

2

(ii) $E(N) = 10000 \times [(A) + (B)]$
Or equivalent

M1

$= 128.6$ to $128.7 \Rightarrow$ **130**
CAO

A1F

2

(c) (i) $P(S \mid +) = \dfrac{P(S \text{ and } +)}{P(+)} =$
Used

M1

$$\frac{0.09 + 0.00784}{0.09 + 0.00784 + 0.009 + 0.00171} = \frac{0.09784}{0.10855}$$

F on (a)
Otherwise correct

A1F

= 0.901 to 0.902
AWRT (0.90134)

A1

(ii) $P(NS \mid -) = \frac{P(NS \text{ and } -)}{P(-)} =$

Used; only if not scored in (i)

(M1)

$$\frac{0.72 + 0.16929}{0.002 + 0.00016 + 0.72 + 0.16929} = \frac{0.88929}{0.89145}$$

F on (a) and/or denominator (c)(i)
Otherwise correct

A1F

= **0.997 to 0.998**

AWFW (0.99758)

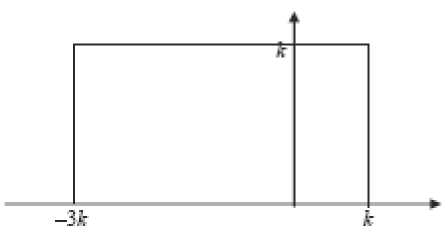
A1

Special cases:
Only numerators correct ⇒ (M1) B1 B1
Only denominators correct ⇒ (M1) B1 B1

5

[13]

2 (a) (i)



Horizontal line f(x) = k
From –3k to k

If $\frac{1}{2}$ then max. B1

B2,1

2

(ii) Area = $4k \times k = 1$

$k^2 = \frac{1}{4}$

***SC** If use $k = \frac{1}{2}$ to show that the Area = 1 then $\Rightarrow$ B1*

M1

$k = \frac{1}{2}$ ($k > 0$)
AG

A1

2

(b) $E(X) = \frac{1}{2}(-3k + k)$
$= -k$

$= -\frac{1}{2}$

CAO

B1

$\text{Var}(X) = \frac{1}{12}(k - -3k)^2 = \frac{16k^2}{12} = \frac{4k^2}{3}$

$= \frac{1}{3}$

CAO

M1

st. dev $(X) = \frac{1}{\sqrt{3}}$ or $\frac{\sqrt{3}}{3}$ or $\sqrt{\frac{1}{3}}$
*OE (**exact**)*

A1

3

(c) (i) $P\left(X \geq -\frac{1}{4}\right) = \frac{1}{2} \times \frac{3}{4}$

M1

$= \frac{3}{8}$ (0.375)

A1

2

(ii) $P\left(X \neq -\frac{1}{4}\right) = 1$

B1

1

[10]

3 (a) Independent trials
Two outcomes OE
Constant probability of success
Unlimited number of trials

Any three

E1 × 3

3

(b) (i) $p + p(1 - p) = 0.2775$

$1 - (1 - p)^2 = 0.2775$

M1

$p^2 - 2p + 0.2775 = 0$

$(1 - p)^2 = 0.2775$

m1

$p = 0.15$ $(0 < p < 1)$

$(1 - p) = 0.85$

m1

$p = 0.15$

A1

4

(ii) $E(Y) = \frac{1}{0.15} = 6.67$

ft on 0 < p < 1

B1F

$\text{Var}(Y) = \frac{0.85}{0.15^2} = 37.8$

B1

2

[9]

4 (a) $E(X) =$ **<u>2.2</u>**

CAO

B1

$\text{Var}(X) = E(X^2) - 2.2^2 =$

Used; or equivalent

M1

$6.8 - 4.84 = 1.96$

CAO

A1

3

(b) (i) $E(S) = E(X) + 2.0 = 4.2$

F on (a)

B1F

$\text{Var}(S) = \text{Var}(X) + 1.5 + 2 \times (-0.43)$

Used for S or D

M1

$= 2.6$

F on (a)

A1F

(ii) $E(D) = E(X) - 2.0 = 0.2$

F on (a)

B1F

$Var(D) = Var(X) + 1.5 - 2 \times (-0.43)$
$= 4.32$

F on (a)

A1F

5

(c) $L \sim N(2.31, 0.89^2)$ $M \sim N(2.04, 0.43^2)$

$T = L + M \sim N(\mathbf{4.35}, \mathbf{0.977})$

Both CAO; SD = 0.98843

B1B1

$P(T > 5) = P\left(Z < \frac{5 - 4.35}{\sqrt{0.977}}\right)$

Standardising 5 or 5.01 using C's mean & SD

M1

$= P(Z > 0.66) = 1 - P(Z < 0.66)$

Area change

m1

0.25 to 0.26

AWFW (0.25540)

A1

5

[13]

5

	Answer	Mark	Comments
(a)	$\frac{2}{3}$	B1	

(b)	$G_X(t) = \sum_{n=1}^{3} t^n \left(\frac{1}{3}\right)$	M1	Applies formula for $G_X(t)$
	$= \frac{1}{3}(t + t^2 + t^3)$ oe	A1	Ignore subsequent incorrect working

Alternative method 1		
(c) $G_{X+Y}(t) = \frac{1}{3}(t + t^2 + t^3)(0.6 + 0.4t)$	M1	Multiplies their $G_X(t)$ and $G_Y(t)$
$\frac{1}{3}t \times 0.6$	M1	Multiplies out the required terms to find the coefficient of t Implied by correct answer
= 0.2	A1	Accept 1/5 oe

Alternative method 2		
Y Bernoulli, p = 0.4 or Y~B(1, 0.4)	M1	Identifies distribution of Y
P(X + Y = 1) means X = 1 and Y = 0 or P(X + Y = 1) = P(X = 1)P(Y = 0)	M1	Identifies possible combinations of X and Y
$= \frac{1}{3}(1 - 0.4) = 0.2$	A1	Accept 1/5 oe
		Total 6 marks

6

	Answer	Mark	Comments
(a)(i)	$P(X = 8) = (1 - 0.045)^7 \times 0.045$	M1	Correct expression Can be implied by correct answer
	= 0.0326	A1	AWRT (From calculator 0.03260138983)

(ii)	$P(X > 5) = (1 - 0.045)^5$	M1	Correct expression Can be implied by correct answer
	= 0.794	A1	AWRT (From calculator 0.7943590686)

(b)	$G_X(t) = E(t^x) = \sum_{x=1}^{\infty} t^x \times p(1-p)^{x-1}$	M1	Applies probability generating function formula unsimplified or simplified
	$= pt(1 + (p-1)t + (p-1)^2t^2 + \cdots)$	M1	Simplifies expression Constants may be unsimplified or simplified
	$= pt\left(\frac{1}{1-(1-p)t}\right)$	M1	Applies sum to infinity formula to evaluate the summation Constants may be unsimplified or simplified
	$= \frac{0.045t}{1-0.955t}$	A1	CAO
			Total 8 marks