

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

**Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education**

MATHEMATICS

4734

Probability & Statistics 3

MARK SCHEME

Specimen Paper

MAXIMUM MARK	72
---------------------	-----------

This mark scheme consists of 4 printed pages.

<p>1 Model for call-outs is Poisson Mean is $\frac{1}{5}(6+2)$ = 1.6 Probability is $1 - 0.9212$ = 0.0788</p>	<p>B1 M1 A1 M1 A1</p>	<p>For any implication of Poisson For summing two relevant parameters For correct mean of 1.6 For relevant use of tables For correct answer</p>										
<p>2 Assume $F = E + M_1 + M_2 + \dots + M_{50}$, where the masses of the 50 matches in a box are independent the mass of the empty box is independent of the masses of the matches $20.0 = 12.5 + 50\mu$ Hence mean mass of a match is 0.15 grams $0.4^2 = 0.2^2 + 50\sigma^2$ Hence standard deviation is 0.049 grams</p>	<p>B1 B1 M1 A1 M1 A1 A1</p>	<p>(The relation itself may be implied) For one relevant valid assumption For another relevant valid assumption For attempting $E(F)$ in terms of μ For correct value 0.15 For attempting $\text{Var}(F)$ as a sum For correct equation For correct value 0.049</p>										
<p>3 (i) $\bar{x} = 25.0055$ $s^2 = \frac{1}{79} \left(0.2287 - \frac{0.44^2}{80} \right)$ = 0.00286... Interval is $25.0055 \pm 2.576 \sqrt{\frac{0.00286}{80}}$ Hence $24.99(0) < \mu < 25.02(1)$</p> <hr/> <p>(ii) The sample size of 80 is sufficient large for the Central Limit Theorem to apply, so it is not necessary to assume a normal distribution</p>	<p>B1 M1 A1 M1 B1 A1</p> <hr/> <p>M1 A1</p>	<p>For correct sample mean, or equivalent; the 25 may be taken into account later For correct unsimplified expression For correct unbiased estimate For calculation of the form $\bar{x} \pm z\sqrt{(s^2/n)}$ For relevant use of $z = 2.576$ For correct interval, stated to an appropriate degree of accuracy</p> <hr/> <p>For mention of sample size and CLT For the correct conclusion and reason</p>										
<p>4 (i) $f_e = 100 \times \int_5^{10} 0.1e^{-0.1x} dx$ = $100[-e^{-0.1x}]_5^{10}$ = $100(e^{-0.5} - e^{-1}) = 23.87$</p> <hr/> <p>(ii) Combining: <table style="display: inline-table; vertical-align: middle;"><tr><td>f_o</td><td>49</td><td>22</td><td>20</td><td>9</td></tr><tr><td>f_e</td><td>39.35</td><td>23.87</td><td>23.25</td><td>13.53</td></tr></table> Test statistic is $\frac{9.65^2}{39.35} + \frac{1.87^2}{23.87} + \frac{3.25^2}{23.25} + \frac{4.53^2}{13.53}$ = 4.484 This is less than 6.251 Hence there is a satisfactory fit</p>	f_o	49	22	20	9	f_e	39.35	23.87	23.25	13.53	<p>M1 A1 M1 M1 A1</p> <hr/> <p>M1 M1 A1 M1 A1\checkmark</p>	<p>For attempting to integrate $f(x)$ For correct indefinite integral For multiplying by total frequency For use of correct limits For obtaining given answer correctly</p> <hr/> <p>For combining the last two classes For correct calculation process For correct value 4.48 For comparison with the correct critical value For correct conclusion, in terms of the fit</p>
f_o	49	22	20	9								
f_e	39.35	23.87	23.25	13.53								

<p>5 (i) $P(X < a) = P(-a < X < a)$ $= \int_{-a}^0 (1+x) dx + \int_0^a (1-x) dx$ $= \left[x + \frac{1}{2}x^2 \right]_{-a}^0 + \left[x - \frac{1}{2}x^2 \right]_0^a = 2a - a^2$</p>	<p>M1 A1 A1</p>	<p>For consideration of two areas, or equiv For integrals or equivalent trapezia 3 For showing the given answer correctly</p>				
<p>(ii) $P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = 2\sqrt{y} - y$ Hence the pgf of Y is $\frac{d}{dy}(2\sqrt{y} - y) = \frac{1}{\sqrt{y}} - 1$</p>	<p>M1 A1 M1 A1</p>	<p>For expression of $P(X^2 \leq y)$ in terms of y For correct expression $2\sqrt{y} - y$ For differentiation of previous expression 4 For showing the given answer correctly</p>				
<p>(iii) $E(Y) = \int_0^1 y^{\frac{1}{2}} - y dy = \left[\frac{2}{3}y^{\frac{3}{2}} - \frac{1}{2}y^{\frac{1}{2}} \right]_0^1 = \frac{1}{6}$ $E(X^2) = \int_{-1}^0 (x^2 + x^3) dx + \int_0^1 (x^2 - x^3) dx$ $= \left[\frac{1}{3}x^3 + \frac{1}{4}x^4 \right]_{-1}^0 + \left[\frac{1}{3}x^3 - \frac{1}{4}x^4 \right]_0^1 = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$</p>	<p>M1 A1 M1 A1</p>	<p>For the correct integral in terms of y For correct answer $\frac{1}{6}$ For the correct integrals in terms of x 4 For the correct answer correctly obtained</p>				
<p>(iv) $E(\sqrt{Y}) = \int_0^1 y^{\frac{1}{2}} g(y) dy = \int_0^1 (1 - y^{\frac{1}{2}}) dy$ $= \left[y - \frac{2}{3}y^{\frac{3}{2}} \right]_0^1 = \frac{1}{3}$</p>	<p>M1 A1</p>	<p>For forming the correct integral 2 For the correct answer $\frac{1}{3}$</p>				
13						
<p>6 (i) H_0: shoppers' views and age are independent, H_1: shoppers' views and age are not independent Exp frequencies under H_0 are <table style="display: inline-table; vertical-align: middle;"><tr><td>163.56</td><td>184.44</td></tr><tr><td>306.44</td><td>345.56</td></tr></table> Test statistic is $\frac{22.94^2}{163.56} + \frac{22.94^2}{184.44} + \frac{22.94^2}{306.44} + \frac{22.94^2}{345.56}$ $= 9.31\dots$ This is greater than the critical 0.5% value of 7.879 Hence there is very strong evidence to reject H_0 and conclude that views about changing to metric units are not independent of age</p>	163.56	184.44	306.44	345.56	<p>B1 M1 A1 M1 A1 M1 A1√</p>	<p>For stating both hypotheses For correct method for expected frequencies For all four correct For correct calculation process, inc Yates For correct value of the test statistic For a relevant (1 df) comparison 7 For correctly justifying the given answer (the final two marks remain available if Yates' correction is omitted)</p>
163.56	184.44					
306.44	345.56					
<p>(ii) $H_0: p_1 = p_2, H_1: p_1 \neq p_2$ Under H_0 the sample value of the common proportion is $\frac{187+161}{1000} = 0.348$ Test statistic is $\frac{\frac{187}{470} - \frac{161}{530}}{\sqrt{0.348 \times 0.652 \times \left(\frac{1}{470} + \frac{1}{530} \right)}}$ $= 3.118$ This is greater than the 0.2% (two-tail) critical value of 3.090 Hence this test supports the conclusion of part (i)</p>	<p>B1 B1 M1 A1 A1 M1 A1√</p>	<p>For both hypotheses stated For correct value of estimated p For num $p_1 - p_2$ and denom using attempted s.d. based on a common estimate of p For completely correct expression For correct value of the test statistic For a relevant comparison using the normal distribution For any relevant comparison or comment 7</p>				
14						

<p>7 (i) (a) $H_0: \mu_d = 0, H_1: \mu_d \neq 0$ $\bar{d} = 4.1667$</p> $s^2 = \frac{486}{11} - \frac{50^2}{11 \times 12} = 25.2424$ <p>Test statistic is $\frac{4.1667 - 0}{\sqrt{(25.2424/12)}}$ $= 2.873$</p> <p>This is greater than the critical value 2.718 Hence there is enough evidence to reject H_0 and conclude that there is a difference between the times for the two methods</p>	<p>B1 B1 M1 A1 M1 A1 M1 A1√</p>	<p>For both hypotheses stated For correct mean difference (subtraction can be either way round) For calculation of unbiased variance estimate For correct value 25.24... For correct standardising process For correct value of test statistic For a relevant comparison using t tables 8 For correctly stated conclusion in context</p>
<p>(b) Population of differences is normal</p>	<p>B1</p>	<p>1 For correct statement</p>
<p>(c) Interval is $4.1667 \pm 2.201 \sqrt{\frac{25.2424}{12}}$</p> <p>Hence $0.97 < \mu_d < 7.36$</p>	<p>M1 B1 A1</p>	<p>For calculation of the form $\bar{d} \pm t \sqrt{(s^2/n)}$ For relevant use of $t = 2.201$ 3 For correct interval</p>
<p>(ii) (a) Variation in the speed of individual workers is not eliminated, and may be large compared with the difference between the methods that is being tested</p>	<p>B1</p>	<p>1 For any relevant valid statement</p>
<p>(b) Both samples are from normal populations The population variances are equal</p>	<p>B1 B1</p>	<p>For a correct statement about normality 2 For a correct statement about the variances</p>
<p>15</p>		