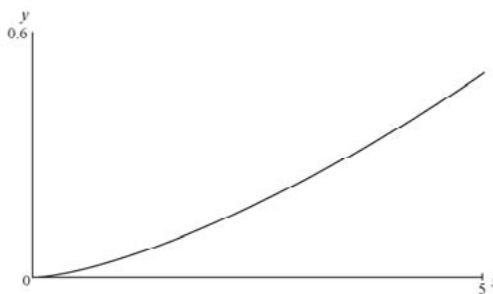


**Paper 4E: Further Statistics 2 Mark Scheme**

Question	Scheme	Marks	AOs
<b>1(a)</b>	$P(A > 3) = \frac{2}{5}$	B1	1.1b
	$\left(\frac{2}{5}\right)^3 = \frac{8}{125}$	M1 A1	1.1a 1.1b
		<b>(3)</b>	
<b>(b)</b>	$f(y) = \frac{3y^2}{125}$	M1	2.1
	$E(Y) = \int_0^5 \frac{3y^3}{125} dy$ $= \left[ \frac{3y^4}{500} \right]_0^5 \quad \left[ = \frac{15}{4} \right]$	M1	1.1b
	$Var(Y) = \int_0^5 \left( \frac{3y^4}{125} \right) dy - \left( \frac{15}{4} \right)^2$	M1	1.1b
	$= 0.9375^*$	A1*cs0	1.1b
		<b>(4)</b>	
<b>(c)</b>	Mode = 5	B1	1.2
	 <p>Or reason based on <math>\frac{df(y)}{dy} &gt; 0</math></p>	B1	2.4
		<b>(2)</b>	
<b>(d)</b>	From a sketch or mode > mean therefore it has negative skew	B1ft	2.4
		<b>(1)</b>	
<b>(e)</b>	$\frac{(2k)^3}{125} - \frac{k^3}{125} = 0.189$	M1	3.1a
	$\frac{7k^3}{125} = 0.189$	A1	1.1b
	$k = 1.5$	A1	1.1b
		<b>(3)</b>	
			<b>(13 marks)</b>

<b>Question 1 notes:</b>	
<b>(a)</b>	<p><b>B1:</b> <math>\frac{2}{5}</math> o.e. may be implied by a correct answer</p> <p><b>M1:</b> <math>\left(\text{"their"}\left(\frac{2}{5}\right)\right)^3</math> may be implied by a correct answer</p> <p><b>A1:</b> <math>\frac{8}{125}</math> o.e.</p>
<b>(b)</b>	<p><b>M1:</b> Realising that firstly need to find pdf <math>f(y)</math> and attempt to differentiate <math>F(y)</math></p> <p><b>M1:</b> Continuing the argument with an attempt to integrate <math>y \times \text{"their } f(y)\text{"}</math>  <math>y^n \rightarrow y^{n+1}</math></p> <p><b>M1:</b> Integrating <math>y^2 \times \text{"their } f(y)\text{"}</math> - [<math>\text{"their } E(Y)\text{"}</math>]<sup>2</sup> <math>y^n \rightarrow y^{n+1}</math></p> <p><b>A1*:</b> Complete correct solution no errors</p>
<b>(c)</b>	<p><b>B1:</b> 5 only</p> <p><b>B1:</b> Explain their reason by either an accurate sketch or <math>\frac{df(y)}{dy} &gt; 0</math> therefore an increasing function o.e.</p>
<b>(d)</b>	<p><b>B1ft:</b> Explaining the reason for their answer. Follow through their part(b) or mean from(d) and mode from(c). A correct sketch of <math>\text{"their } f(y)\text{"}</math> – may be seen anywhere in question or ft their mean and mode plus a correct conclusion</p> <p><b>NB:</b> Watch for gaming. A student who writes both negative skew with a reason and positive skew with a reason. Please send these to your Team Leader</p>
<b>(e)</b>	<p><b>M1:</b> Attempting to translate the problem into an equation using <math>2k</math> and <math>k</math>. Allow if the brackets are missing e.g. <math>\frac{2k^3}{125} - \frac{k^3}{125}</math>. No need for the 0.189</p> <p><b>A1:</b> A correct equation in any form</p> <p><b>A1:</b> A correct answer only</p>

Question	Scheme	Marks	AOs
<b>2(a)</b>	$H_0 : \rho = 0, H_1 : \rho > 0$	B1	2.5
	Critical value at 1% level is 0.8929	B1	1.1b
	$r_s < 0.8929$ so not significant evidence to reject $H_0$	M1	2.1
	The researcher's claim is not correct (at 1% level) <b>or</b> insufficient evidence for researcher's claim <b>or</b> there is insufficient evidence that water gets deeper further from inner bank <b>or</b> no (positive) correlation between depth of water and distance from inner bank	A1ft	2.2b
		<b>(4)</b>	
<b>(b)(i)</b>	The <b>ranks will remain the same</b> therefore there will be <b>no change</b> to the spearman's rank correlation coefficient	<b>B1</b>	2.4
<b>(ii)</b>	Spearman's rank correlation coefficient will <b>increase</b> since	<b>B1</b>	2.2a
	The <b>ranks are the same</b> for both distance and depth therefore <b><math>d = 0</math></b> however, <b><math>n</math> has increased</b> <b>or</b> the new position follows the pattern that large $b$ is associated with large $s$ and so $r_s$ will increase	<b>B1</b>	2.4
		<b>(3)</b>	
<b>(c)</b>	The mean of the tied ranks is given to each...	B1	2.4
	... then use PMCC	B1	2.4
		<b>(2)</b>	
<b>(9 marks)</b>			
<b>Notes:</b>			
<b>(a)</b>			
<b>B1:</b> Both hypotheses correct written using the notation $\rho$			
<b>B1:</b> awrt 0.893			
<b>M1:</b> Drawing a correct inference using their answer to part(a) and their CV			
<b>A1ft:</b> Drawing a correct inference in context using their answer to part(a) and their CV			
<b>(b)(i)</b>			
<b>B1:</b> Stating <b>no change</b> and an explanation including <b>ranks remain unchanged</b> o.e. and <b>no change o.e.</b>			
<b>(b)(ii)</b>			
<b>B1:</b> Interpreted the outcome of adding a point as <b>increased</b> oe			
<b>B1:</b> Explaining why. Need to mention the <b>ranks are the same for both</b> oe and <b><math>n</math> has increased</b> oe			
<b>(c)</b>			
<b>B1:</b> Explaining that the mean of the values for the tied ranks is given to both values			
<b>B1:</b> Explaining that the PMCC must be used			

Question	Scheme	Marks	AOs
<b>3(a)</b>	95% CI for $\mu$ uses $t$ value of <b>2.064</b>	B1	3.3
	$\frac{\hat{\sigma}}{\sqrt{25}} \times "2.064" = \frac{1}{2}(2.232 - 1.128)$ or $\frac{1}{2}(2.232 + 1.128) + "2.064" \times \frac{\hat{\sigma}}{\sqrt{25}} = 2.232$ (oe)	M1	2.1
	$\hat{\sigma} = \frac{2.76}{"2.064"}$ or 1.3372...	M1	1.1b
	$\hat{\sigma}^2 = 1.788... [=1.79 \text{ (3sf)}]$ *	A1*cso	1.1b
	<b>(4)</b>		
<b>(b)</b>	$12.401, < \frac{24 \times 1.79}{\sigma^2} <, 39.364$	B1 M1	1.1b 1.1a
	<b>1.09</b> < $\sigma^2$ < <b>3.46</b>	A1	1.1b
	<b>(3)</b>		
	<b>(7 marks)</b>		
<b>Notes:</b>			
<b>(a)</b>			
<b>B1:</b> Realising that the $t$ -distribution must be used as a model and finding the correct value awrt 2.06			
<b>M1:</b> Using the correct formula with a $t$ -value, $\frac{\hat{\sigma}}{\sqrt{25}} \times "t \text{ value}" = \frac{1}{2}(2.232 - 1.128)$ or $\frac{1}{2}(2.232 + 1.128) + "t \text{ value}" \times \frac{\hat{\sigma}}{\sqrt{25}} = 2.232$ or $\frac{1}{2}(2.232 + 1.128) - "t \text{ value}" \times \frac{\hat{\sigma}}{\sqrt{25}} = 1.128$			
<b>M1:</b> Rearranging one of these formula accurately to find a value of $\hat{\sigma}$			
<b>A1cso*:</b> A correct solution only using awrt 1.79			
<b>(b)</b>			
<b>B1:</b> awrt 12.4 or 39.4 May be implied by a correct confidence interval			
<b>M1:</b> $\frac{24 \times 1.79}{\sigma^2}$ May be implied by a correct confidence interval			
<b>A1:</b> awrt 1.09 and awrt 3.46			

Question	Scheme	Marks	AOs
<b>4(a)</b>	$H_0: \sigma_G^2 = \sigma_B^2, H_1: \sigma_G^2 \neq \sigma_B^2,$	B1	2.5
	$s_B^2 = \frac{1}{6}(56130 - 7 \times 88.9^2) = \frac{807.53}{6} = 134.6$	M1 A1	2.1 1.1b
	$s_G^2 = \frac{1}{7}(55746 - 8 \times 83.1^2) = \frac{501.12}{7} = 71.58$	A1	1.1b
	$\frac{s_B^2}{s_G^2} = 1.880\dots$	M1	3.4
	Critical value $F_{6,7} = 3.87$	B1	1.1b
	Not significant, variances can be treated as the same	A1 ft	2.2b
		(7)	
<b>(b)</b>	$H_0: \mu_B = \mu_G, H_1: \mu_B > \mu_G$	B1	2.5
	Pooled estimate of variance $s^2 = \frac{6 \times 134.6 + 7 \times 71.58}{13} = 100.6653\dots$	M1	3.1b
	Test statistic $t = \frac{88.9 - 83.1}{s \sqrt{\frac{1}{7} + \frac{1}{8}}} = \text{awrt } 1.12$	M1 A1	1.1b 1.1b
	Critical value $t_{13}(5\%) = 1.771$	B1	1.1b
	Insufficient evidence to support mother's claim	A1 ft	2.2b
		(6)	
	<b>(13 marks)</b>		
<b>Notes:</b>			
<b>(a)</b>			
<b>B1:</b> Both hypotheses correct using the notation $\sigma^2$ . Allow $\sigma$ rather than $\sigma^2$			
<b>M1:</b> Using a correct Method for either $s_B^2$ or $s_G^2$ May be implied by a correct value			
<b>A1:</b> awrt 135			
<b>A1:</b> awrt 71.6			
<b>M1:</b> Using the F-distribution as the model e.g. $\frac{s_B^2}{s_G^2}$			
<b>B1:</b> awrt 3.87			
<b>A1ft:</b> Drawing a correct inference following through their CV and value for $\frac{s_B^2}{s_G^2}$			
<b>(b)</b>			
<b>B1:</b> Both hypotheses correct using the notation $\mu$			
<b>M1:</b> For realising the need to find the pooled estimate for the test require from a correct interpretation of the question			
<b>M1:</b> Correct method for test statistic $t = \frac{88.9 - 83.1}{\text{"their } s" \sqrt{\frac{1}{7} + \frac{1}{8}}}$ May be implied by a correct			
awrt 1.12			
<b>A1:</b> awrt 1.12			
<b>B1:</b> awrt 1.77			
<b>A1ft:</b> Drawing a correct inference following through their CV and test statistic			

Question	Scheme	Marks	AOs
<b>5(a)</b>	Let $X = L - 4S$ then $E(X) = 19.6 - 4 \times 4.8$	M1	2.3
	$= 0.4$	A1	1.1b
	$\text{Var}(X) = \text{Var}(L) + 4^2 \text{Var}(S) = 0.6^2 + 16 \times 0.3^2$	M1	2.1
	$= 1.8$	A1	1.1b
	$P(X > 0) = [P(Z > \frac{0-0.4}{\sqrt{1.8}} = -0.298\dots\dots)]$	M1	2.1
	$= 0.617202\dots$ awrt <b>0.617</b>	A1	1.1b
	<b>(6)</b>		
<b>(b)</b>	$T = S_1 + S_2 + S_3 + S_4$ (May be implied by 0.36)	M1	3.3
	$T \sim N(19.2, 0.36)$ $E(T) = 19.2$	B1	1.1b
	$\text{Var}(T) = 0.36$ or $0.6^2$	A1	1.1b
		<b>(3)</b>	
<b>(c)</b>	Let $Y = L - T$ $E(Y) = E(L) - E(T) = [0.4]$	M1	3.3
	$\text{Var}(Y) = \text{Var}(L) + \text{Var}(T) = [0.72]$	M1	1.1b
	Require $P(-0.2 < Y < 0.2)$	M1	3.1a
	$= 0.16708\dots$ awrt <b>0.167</b>	A1	1.1b
		<b>(4)</b>	
<b>(13 marks)</b>			
<b>Notes:</b>			
<b>(a)</b>			
<b>M1:</b> Selecting and using an appropriate model i.e. $\pm(L - 4S)$ . May be implied by 0.4			
<b>A1:</b> 0.4 oe			
<b>M1:</b> For realising the need to use $\text{Var}(L) + 4^2\text{Var}(S)$ . Allow use of 0.6 for $\text{Var}(L)$ instead of $0.6^2$ and/or 0.3 for $\text{Var}(S)$ instead of $0.3^2$ may be implied by 1.8			
<b>A1:</b> 1.8 only			
<b>M1:</b> For realising $P(X > 0)$ is required and an attempt to find it e.g. $\frac{0-0.4}{\sqrt{\text{their Var}(X)}}$ but do not allow a negative $\text{Var}(X)$			
<b>A1:</b> awrt 0.617			
<b>(b)</b>			
<b>M1:</b> Selecting and using an appropriate model ie $S_1 + S_2 + S_3 + S_4$ : may be implied by 0.36			
<b>B1:</b> 19.2 only			
<b>A1:</b> 0.36			
<b>(c)</b>			
<b>M1:</b> Setting up and using the model $Y = L - T$ . May be implied by $E(Y) = E(L) - E(T)$			
<b>M1:</b> Using $\text{Var}(Y) = \text{Var}(L) + \text{Var}(T)$			
<b>M1:</b> Dealing with the modulus and realising they need to find $P(-0.2 < Y < 0.2)$			
<b>A1:</b> awrt 0.167			

Question	Scheme	Marks	AOs																																												
<b>6(a)</b>	$\left[ b = \frac{S_{xm}}{S_{xx}} = -0.0277576 \right]$	M1	3.3																																												
	$[a = \bar{m} - b\bar{x} = 1.278 + 0.0277576 \times 8.5 = 1.5139]$																																														
	$m = 1.5139 - 0.02775...x$	A1	1.1b																																												
		<b>(2)</b>																																													
<b>(b)</b>	$RSS = 0.12756 - \frac{(-2.29)^2}{82.5}$	M1	1.1b																																												
	$= 0.06399^*$	A1*	1.1b																																												
		<b>(2)</b>																																													
<b>(c)</b>	<table border="1"> <thead> <tr> <th><math>x</math></th> <th><math>m</math></th> <th><math>m = a + bx</math></th> <th><math>\varepsilon</math></th> </tr> </thead> <tbody> <tr><td>4</td><td>1.50</td><td>1.4029</td><td>+0.0971</td></tr> <tr><td>5</td><td>1.20</td><td>1.3752</td><td>-0.1752</td></tr> <tr><td>6</td><td>1.40</td><td>1.3474</td><td>+0.0526</td></tr> <tr><td>7</td><td>1.40</td><td>1.3196</td><td>+0.0804</td></tr> <tr><td>8</td><td>1.23</td><td>1.2919</td><td>-0.0619</td></tr> <tr><td>9</td><td>1.30</td><td>1.2641</td><td>+0.0359</td></tr> <tr><td>10</td><td>1.20</td><td>1.2364</td><td>-0.0364</td></tr> <tr><td>11</td><td>1.15</td><td>1.2086</td><td>-0.0586</td></tr> <tr><td>12</td><td>1.25</td><td>1.1808</td><td>+0.0692</td></tr> <tr><td>13</td><td>1.15</td><td>1.1531</td><td>-0.0031</td></tr> </tbody> </table>	$x$	$m$	$m = a + bx$	$\varepsilon$	4	1.50	1.4029	+0.0971	5	1.20	1.3752	-0.1752	6	1.40	1.3474	+0.0526	7	1.40	1.3196	+0.0804	8	1.23	1.2919	-0.0619	9	1.30	1.2641	+0.0359	10	1.20	1.2364	-0.0364	11	1.15	1.2086	-0.0586	12	1.25	1.1808	+0.0692	13	1.15	1.1531	-0.0031	M1	3.4
	$x$	$m$	$m = a + bx$	$\varepsilon$																																											
	4	1.50	1.4029	+0.0971																																											
	5	1.20	1.3752	-0.1752																																											
	6	1.40	1.3474	+0.0526																																											
	7	1.40	1.3196	+0.0804																																											
	8	1.23	1.2919	-0.0619																																											
	9	1.30	1.2641	+0.0359																																											
	10	1.20	1.2364	-0.0364																																											
	11	1.15	1.2086	-0.0586																																											
	12	1.25	1.1808	+0.0692																																											
	13	1.15	1.1531	-0.0031																																											
	A1	1.1b																																													
	<b>(2)</b>																																														
<b>(d)</b>	The point (5, 1.2) is an outlier	B1ft	2.2b																																												
		<b>(1)</b>																																													
<b>(e)(i)</b>	It is a valid piece of data so should be used <b>or</b> It does not follow the pattern according to the residuals so may contain an error making the result invalid so should be removed	B1	2.4																																												
<b>(ii)</b>	$a = \bar{m} - b\bar{x} = 1.28667 + 0.03765 \times 8.88889 = 1.6213$	M1	3.3																																												
	$m = 1.6213 - 0.03765x$	A1	1.1b																																												
<b>(iii)</b>	$m = 1.6213 - 0.03765 \times 15$																																														
	$= 1.056$ or awrt 1.06	B1ft	3.4																																												
<b>(iv)</b>	The model is only reliable if the values are limited to those in the given range so probably not reliable	B1	3.5b																																												
		<b>(5)</b>																																													
<b>(12 marks)</b>																																															

<b>Question 6 notes:</b>	
<b>6(a)</b>	
<b>M1:</b>	Realising the need to use $b = \frac{S_{xm}}{S_{xx}}$ and $a = \bar{m} - b\bar{x}$
<b>A1:</b>	$m = \text{awrt } 1.51) - (\text{awrt } 0.0278) x$ . Award M1A1 for correct equation
<b>(b)</b>	
<b>M1:</b>	Using $S_{mm} - \frac{(S_{xm})^2}{S_{xx}}$
<b>A1*:</b>	awrt 0.064
<b>(c)</b>	
<b>M1:</b>	Using the model in part (a) i.e. $m - ("1.5139" - "0.02775"x)$ implied by a correct value
<b>A1:</b>	All correct. Award M1A1 for a list of correct residuals
<b>(d)</b>	
<b>B1:</b>	Inferring from the residuals that the outlier is (5, 1.2) ft their residuals.
<b>(e)(i)</b>	
<b>B1:</b>	Explaining why the outlier should be removed or not.
<b>(ii)</b>	
<b>M1:</b>	Removing the outlier and refining the model by finding a new regression line.
<b>A1:</b>	$m = (\text{awrt } 1.62) - (\text{awrt } 0.0377)x$
<b>(iii)</b>	
<b>B1ft:</b>	using their model in e(i) with $x = 15$ . awrt 1.06 or ft their e(ii)
<b>(iv)</b>	
<b>B1:</b>	Realising the limitations of the model by stating it is <u>not reliable</u> and giving the reason why i.e. extrapolation/out of range o.e.



Question	Scheme	Marks	AOs
<b>7(a)</b>	$S_{xx} = \sum (10s)^2 - \frac{(\sum 10s)^2}{10}$	M1	2.1
	$2658.9 = 100 \sum (s)^2 - \frac{100(\sum s)^2}{10}$	M1	1.1b
	$2658.9 = 100 S_{ss}$		
	$S_{ss} = 26.589 *$	A1*cso	1.1b
		<b>(3)</b>	
<b>(b)</b>	$64 = \sum_{i=1}^{10} 10(d_i - 9)$	M1	3.1a
	$64 = 10 \sum_{i=1}^{10} d_i - 900$		
	$\sum_{i=1}^{10} d_i = 96.4$	A1	1.1b
	$S_{dd} = 1081.74 - \frac{("96.4")^2}{10}$	M1	1.1b
	$= 152.444$		
	$r = 0.935$	A1ft	1.1b
		<b>(4)</b>	
<b>(c)</b>	Linear correlation is significant but scatter diagram suggests a non-linear relationship between the level of serum magnesium, and the level of the disease protein	B1	3.5a
		<b>(1)</b>	
<b>(8 marks)</b>			
<b>Notes:</b>			
<b>(a)</b>			
<b>M1:</b> Attempting to use $S_{xx} = \sum x^2 - \frac{(\sum x)^2}{10}$ with $x = 10s$			
<b>M1:</b> Substituting in 2658.9 and dealing with the 10 correctly			
<b>A1*:</b> cso A complete solution with no errors leading to 26.589 only			
<b>(b)</b>			
<b>M1:</b> Realising that either $64 = \sum_{i=1}^{10} 10(d_i - 9)$ or $64 = 10 \sum_{i=1}^{10} d_i - 900$ o.e. must be used. May be implied by seeing 96.4			
<b>A1:</b> 96.4 only			
<b>M1:</b> Attempting to use $S_{dd} = \sum d^2 - \frac{(\sum d)^2}{10}$ may be implied by 0.935			
<b>A1ft:</b> awrt 0.935 ft "their 96.4"			
<b>(c)</b>			
<b>B1:</b> A correct comment comparing their value of $r$ and the scatter diagram in context			

