A Level Statistics
AQA Past Exam Questions

TOPIC: Hypothesis Testing
One Way ANOVA

Candidates may use any calculator allowed by Pearson regulations. Calculators must not have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions ***on paper***
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Unless otherwise stated, statistical tests should be carried out at the 5% significance level.
- When a calculator is used, the answer should be given to three significant figures unless otherwise stated.

Information

- ***You may use the*** booklet ‘Statistical Formulae and Tables’
- There are **8** questions in this question paper. The total mark for this paper is ***60***
- The marks for **each** question are shown in brackets – use this as a guide as to how much time to spend on each question.

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.
- Check your answers if you have time at the end.

AQA_JUNE_2018_2

<table>
<tr><th>Solution</th><th>Mark</th><th>Total</th><th>Comment</th></tr>
<tr><td>$H_0 \quad \mu_{Shetland} = \mu_{Argyll} = \mu_{NCentral}$
H_1 at least 2 of the means differ 5% 1 tail
$T_{Shetland} = 64.23 \quad T_{Argyll} = 74.08 \quad T_{N\ Central} = 64.85$
$n_{Shetland} = 6 \quad n_{Argyll} = 5 \quad n_{N\ Central} = 7$
$T = 203.16 \quad \sum\sum x_{ij}^2 = 2404.94 \quad N = 18$</td><td>B1</td><td></td><td></td></tr>
<tr><td>Total SS
$2404.94 - \frac{203.16^2}{18} = \underline{111.94}$</td><td>M1PI</td><td></td><td>Total SS effort</td></tr>
<tr><td>Areas SS
$\frac{64.23^2}{6} + \frac{74.08^2}{5} + \frac{64.85^2}{7} - \frac{203.16^2}{18} = \underline{92.94}$</td><td>M1PI</td><td></td><td>Areas SS effort</td></tr>
<tr><td><table><tr><th></th><th>ss</th><th>df</th><th>ms</th></tr><tr><td>Between areas</td><td>92.94</td><td>2</td><td>46.47</td></tr><tr><td>Error</td><td>19.00</td><td>15</td><td>1.27</td></tr><tr><td>Total</td><td>111.94</td><td>17</td><td></td></tr></table></td><td>M1PI
B1PI
m1PI</td><td></td><td>dep error ss positive
error df = 15
method for ms ft dep B1 M1 previously</td></tr>
<tr><td>$F = \frac{46.47}{1.27} = 36.7$</td><td>m1PI</td><td></td><td>F test stat (awfw 30-40)</td></tr>
<tr><td>$F^{2}_{15} = 3.682$
$36.7 > 3.682$</td><td>B1</td><td></td><td>cv cao
or p = 0.00000167 <0.05</td></tr>
<tr><td>Reject H_0</td><td>A1 dep</td><td></td><td>Correct conclusion dep ts/cv correct</td></tr>
<tr><td>There is significant evidence of a difference in mean mercury concentration for at least two of the areas of Scotland.
<table><tr><th></th><th>Shetland</th><th>Argyll</th><th>N Central</th></tr><tr><td>mean</td><td>10.705</td><td>14.816</td><td>9.264</td></tr></table></td><td>E1PI</td><td></td><td>Correct conclusion in context – ‘at least two’ included
dep ts and cv correct and not too definite
PI if fully explained as below</td></tr>
<tr><td>Otters in the Argyll area of Scotland have significantly higher (bioaccumulations of) mercury, on average, than those otters in the N Central area or N Central lower than Argyll.</td><td>E1</td><td>11</td><td>Explanation in full of conclusion with some numerical justification attempt or reference to means (not referring to totals)</td></tr>
<tr><td>Total</td><td></td><td>11</td><td></td></tr>
</table>

AQA_JUNE_2017_5

<table>
<tr><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>$H_O \quad \mu_{none} = \mu_{T1} = \mu_{T2}$
H_1 at least 2 (of the means) differ 1% 1 tail</td><td>B1</td><td></td><td>Suffices must be identified
Disallow ‘At least one differs’</td></tr>
<tr><td>
<table>
<tr><th>None</th><th>TR_1</th><th>TR_2</th></tr>
<tr><td>820</td><td>720</td><td>650</td></tr>
<tr><td>940</td><td>900</td><td>710</td></tr>
<tr><td>930</td><td>790</td><td>690</td></tr>
<tr><td>880</td><td>920</td><td>710</td></tr>
<tr><td>860</td><td>840</td><td>620</td></tr>
<tr><td>790</td><td>870</td><td>700</td></tr>
<tr><td>850</td><td>810</td><td>830</td></tr>
<tr><td></td><td>900</td><td></td></tr>
<tr><td></td><td>840</td><td></td></tr>
</table>
</td><td>M1
m1</td><td></td><td>Effort to separate categories of treatment PI
12 or more correctly placed PI</td></tr>
<tr><td>$T_{none} = 6070 \quad T_{TR1} = 7590 \quad T_{TR2} = 4910$
$n_{none} = 7 \quad n_{TR1} = 9 \quad n_{TR1} = 7$</td><td>m1</td><td></td><td>Total in each category effort</td></tr>
<tr><td>$T = 18570 \quad \sum\sum x_{ij}^2 = 15184700 \quad N = 23$</td><td>m1</td><td></td><td>T and $\sum\sum x_{ij}^2$ effort</td></tr>
<tr><td>Total SS
$15184700 - \frac{18570^2}{23} = \underline{191443.5}$</td><td>m1</td><td></td><td>ss Total</td></tr>
<tr><td>Treatments SS
$\frac{6070^2}{7} + \frac{7590^2}{9} + \frac{4910^2}{7} - \frac{18570^2}{23} = 115214.9$</td><td></td><td></td><td></td></tr>
<tr><td>
<table>
<tr><th></th><th>ss</th><th>df</th><th>ms</th></tr>
<tr><td>Between treats</td><td>115214.9</td><td>2</td><td>57607.5</td></tr>
<tr><td>Error</td><td>76228.6</td><td>20</td><td>3811.4</td></tr>
<tr><td>Total</td><td>191443.5</td><td>22</td><td></td></tr>
</table>
</td><td>m1

B1

m1</td><td></td><td>ss Error
(allow small slip – not if negative)

df Between Treats and Error - BOTH

ms Error and Between dep correct df</td></tr>
<tr><td>$F = \frac{57607.5}{3811.4} = 15.11 \qquad F_{20}^{2} = 5.849$</td><td>M1
A1</td><td></td><td><i>F</i> Between divided by Error
awfw 15.0 - 15.3 condone small arithmetic slips if <i>F</i> in range or <i>p</i> = 0.0001</td></tr>
<tr><td>$15.11 > \mathbf{5.849}$ Reject H_o</td><td>B1</td><td></td><td>cv=5.849 cao
or p = 0.0001 < 1%</td></tr>
<tr><td>The conclusion indicates that there is a significant difference between the mean level of immune cells ……
..for at least two of the treatments/treatment TR₂ clearly reduces the level of immune cells/ slows the progress of the disease more than treatment TR₁</td><td>E1dep

E1dep</td><td>14</td><td>Correct conclusion in context

mention of ‘at least two’ treatments
or
TR₂ identified as treatment that reduces more than TR₁</td></tr>
<tr><td>Total</td><td></td><td>14</td><td></td></tr>
</table>

AQA_JUNE_2016_1

<table>
<tr><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>H_O $\mu_{24+} = \mu_{12-24} = \mu_{2-12} = \mu_{less2}$
H_1 at least 2 (of the means) differ 5% 1 tail

$T_{24+} = 345.3$ $T_{12-24} = 328.7$ $T_{2-12} = 303.2$ $T_{less2} = 241.7$
$n_{24+} = 5$ $n_{12-24} = 5$ $n_{2-12} = 5$ $n_{-less2} = 4$

T= 1218.9 $\sum\sum x_{ij}{}^2 = 78811.89$ N= 19</td><td>B1</td><td></td><td>Disallow if labels eg A,B,C,D used without identification
Allow 'population mean' for H_O
H_O $\mu_i = \mu_j$ for all i,j
H_1 $\mu_i \neq \mu_j$ for some i,j</td></tr>
<tr><td>Total SS
$78811.89 - \frac{1218.9^2}{19} = 616.3$</td><td>M1</td><td></td><td>Total SS effort</td></tr>
<tr><td>Times SS
$\frac{345.3^2}{5} + \frac{328.7^2}{5} + \frac{303.2^2}{5} + \frac{241.7^2}{4} - \frac{1218.9^2}{19} = 250.3$</td><td>M1</td><td></td><td>Times SS effort</td></tr>
<tr><td><table>
<tr><th></th><th>ss</th><th>df</th><th>ms</th></tr>
<tr><td>Between times</td><td>250.3</td><td>3</td><td>83.4</td></tr>
<tr><td>Error</td><td>**366.0**</td><td>**15**</td><td>24.4</td></tr>
<tr><td>Total</td><td>616.3</td><td>18</td><td></td></tr>
</table></td><td>M1 dep
B1
M1 dep</td><td></td><td>error ss dep SS above and (all SS) positive
error df PI
method for either ms ft PI</td></tr>
<tr><td>$F = \frac{83.4}{24.4} = \mathbf{3.42}$ $F^{3}_{15} = \mathbf{3.287}$
$3.287 < 3.42$ Reject H_o</td><td>A1 dep
B1</td><td></td><td>F test stat awrt 3.4
cv cao
Alt Allow p =0.0448 compared with 0.005 for A1 B1</td></tr>
<tr><td>There is significant evidence of a difference in mean coursework marks (for at least two of the handing in times).</td><td>A1 dep</td><td></td><td>Correct conclusion in context</td></tr>
<tr><td>Students handing in coursework more than 24 hours before the deadline gain higher marks, on average, than those handing in coursework less than 2 hours before the deadline</td><td>E1dep</td><td>**10**</td><td>Explanation, in full, of conclusion</td></tr>
<tr><td>Coursework **marks** are **normally distributed** with a **common variance.**</td><td>B1
E1</td><td>**2**</td><td>Normal and common variance
In context mentioning marks normally distributed or marks have common variance</td></tr>
<tr><td></td><td>**Total**</td><td>**12**</td><td></td></tr>
</table>

AQA_JUNE_2015_3a

<table>
<tr><th>Q3</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>(a)(i)</td><td>HO $\mu_A = \mu_B = \mu_C = \mu_D$
H_1 at least 2 of the means differ</td><td>B1</td><td></td><td>Hypotheses – must include 'at least two'</td></tr>
<tr><td></td><td>T = 701
$\sum x_i^2 = 33587$ N = 15</td><td></td><td></td><td></td></tr>
<tr><td></td><td>Between print providers ss
$\sum \frac{T_i^2}{n_i} = \frac{197^2}{4} + \frac{199^2}{4} + \frac{146^2}{4} + \frac{159^2}{3} = 33358.5$</td><td>M1</td><td></td><td>$\sum \frac{T_i^2}{n_i}$</td></tr>
<tr><td></td><td>$SS_{\text{print prov}} = 33358.5 - \frac{701^2}{15}$
$= 598.4$</td><td>M1</td><td></td><td>ss for printing providers</td></tr>
<tr><td></td><td>$SS_{\text{Total}} = 33587 - \frac{701^2}{15}$
$= 826.9$</td><td>M1</td><td></td><td>ss total</td></tr>
<tr><td></td><td><table><tr><th></th><th>SS</th><th>df</th><th>ms</th></tr><tr><td>Print prov</td><td>598.4</td><td>3</td><td>199.5</td></tr><tr><td>Error</td><td>228.5</td><td>11</td><td>20.8</td></tr><tr><td>Total</td><td>826.9</td><td>14</td><td></td></tr></table></td><td>B1
M1</td><td></td><td>df = 3 pp correct
ms ft if effort at ss seen</td></tr>
<tr><td></td><td>F ratio = $\frac{199.5}{20.8} = 9.60$ $F^{3}_{11} = 6.217 < 9.60$</td><td>m1
A1
B1</td><td>10</td><td>F ratio ms printing prov / error ms
F correct (9.5 – 9.7)
cv correct (or p = 0.0021 comp 0.01)</td></tr>
<tr><td></td><td>Reject H_O.</td><td>A1</td><td></td><td>Reject H_O</td></tr>
<tr><td>(ii)</td><td>The conclusion indicates that there is a significant difference between the mean scores for at least two of the providers. Printing provider C is clearly the worst and provider D is the best from consideration of means/medians. Choose D.</td><td>E1</td><td></td><td>At least two differ (may be in (i))or C differs from D</td></tr>
<tr><td></td><td><table><tr><th></th><th>A</th><th>B</th><th>C</th><th>D</th></tr><tr><td>Means</td><td>49.25</td><td>49.75</td><td>36.5</td><td>53</td></tr><tr><td>Meds</td><td>49.5</td><td>49</td><td>35</td><td>54</td></tr></table></td><td>E1</td><td>2</td><td>D best with reason/correct evidence seen(53/4)</td></tr>
<tr><td>(iii)</td><td>The scores are normally distributed with a common underlying variance</td><td>E1
E1</td><td>2</td><td>Must be in context – refer to scores.</td></tr>
</table>

AQA_JUNE_2013_3

Q	Solution	Marks	Total	Comments
3(a)(i)	*Low level* *Medium level* *High level* $T_{\text{low}} = 85.8$ $T_{\text{med}} = 108.6$ $T_{\text{high}} = 85.6$ $n_{\text{low}} = 5$ $n_{\text{med}} = 6$ $n_{\text{high}} = 5$ $T = 280$ $\sum\sum x_{ij}^2 = 4910.2$ $N = 16$ $\sum \frac{T_i^2}{n_i} = \frac{85.8^2}{5} + \frac{108.6^2}{6} + \frac{85.6^2}{5}$ $= 4903.46$			
	$\text{SS}_{\text{treats}} = 4903.46 - \frac{280^2}{16}$ $= \mathbf{3.46}$	M1		SS for treatments
	$\text{SS}_{\text{Total}} = 4910.2 - \frac{280^2}{16}$ $= \mathbf{10.2}$	M1		SS for total
	$\begin{array}{\|l\|l\|l\|l\|} \hline & \text{SS} & \text{df} & \text{ms} \\ \hline \text{Treats} & 3.46 & 2 & \mathbf{1.73} \\ \hline \text{Error} & \mathbf{6.74} & 13 & \mathbf{0.52} \\ \hline \text{Total} & 10.2 & 15 & \\ \hline \end{array}$	M1dep M1dep		Error SS ft (not –ve) Either ms correct method (SS/df)
	$\text{F} = \frac{1.73}{0.52} = \mathbf{3.33}$	M1dep		Method for F (ft) 'their ms treats/ms error'
		A1		3.1–3.5
	$\text{F}\,^{2}_{13} = \mathbf{3.806}$	B1 B1		df correct 2,13 cv correct (or $p = 0.068 > 0.05$ B2)
	H_0 $\mu_{low} = \mu_{med} = \mu_{high}$ H_1 at least 2 of the means differ oe One mean sig different from others	B1		Hypotheses
	3.806 > 3.33 **Accept $\mathbf{H_0}$.** There is no significant evidence of a difference in mean breaking strength for the 3 thread treatment levels.	A1	10	Conclusion correct
(a)(ii)	Since there is **no significant difference** detected between mean **breaking strength** for the **three** thread **treatments/levels, the company should not be advised** to use any one particular treatment level.	E1		No difference in strengths for treatments
		E1	2	Could not advise company to use a specific level of treatment or choose cheapest/easiest to obtain
(b)	The **Kruskal–Wallis** test as this is distribution free so **does not depend on assumption that breaking strengths are normally distributed.**	B1		Kruskal–Wallis
		E1dep	2	Does not require underlying normal distribution/distribution free.
	Total		**14**	

AQA_JUNE_2012_5a

Q	Solution	Marks	Total	Comments
5(a)(i)	$A(20/29)$ $B(30/49)$ $C(50+)$ $T_A = 10.67$ $T_B = 16.03$ $T_C = 16.39$ $n_A = 5$ $n_B = 6$ $n_C = 6$ $T = 43.09$ $\sum\sum x_{ij}^2 = 111.138 \quad N = 17$ $\sum \frac{T_i^2}{n_i} = \frac{10.67^2}{5} + \frac{16.03^2}{6} + \frac{16.39^2}{6}$ $= 110.37$			
	$SS_{Ages} = 110.37 - \frac{43.09^2}{17}$ $= 1.148$	M1		SS for ages
	$SS_{Total} = 111.138 - \frac{43.09^2}{17}$ $= 1.917(5)$	M1		SS for total (can be implied in table)
	$\begin{array}{\|l\|c\|c\|c\|} \hline & \text{SS} & \text{df} & \text{MS} \\ \hline \text{Ages} & 1.148 & 2 & 0.574 \\ \hline \text{Error} & 0.769(5) & 14 & 0.055 \\ \hline \text{Total} & 1.9175 & 16 & \\ \hline \end{array}$	M1 m1		Error SS ft (not –ve) Method for MS (dep error ss/df)
	$F = \frac{0.574}{0.055} = 10.44$	m1 A1		Method for F (ft) 10.2–10.6 (or p =)
	$F^{2}_{14} = 6.515 < 10.44$	B1 B1		df correct 2, 14 cv correct CAO allow $p = 0.00167$
	H_0 $\mu_A = \mu_B = \mu_C$ H_1 at least 2 of the pop means differ	B1		hypotheses – subscripts identified OE
	Reject H_0	A1	10	
(ii)	There is significant evidence of a difference in mean satisfaction scores for the 3 age groups so at least 2 groups differ. Ages 20/29 sig less satisfied than those aged 50+	E1	1	In context
(iii)	The normal populations of satisfaction scores have a common variance	E1	1	For either normally distributed satisfaction scores or populations of satisfaction scores have a common variance

AQA_JUNE_2011_3

<table>
<tr><th>Q</th><th>Solution</th><th>Marks</th><th>Total</th><th>Comments</th></tr>
<tr><td>3(a)</td><td>$T_A = 2857$ $T_B = 2490$ $T_C = 3190$
$n_A = 6$ $n_B = 5$ $n_C = 7$
T = 8537
$\sum\sum x_{ij}^2 = 4\,067\,243$ N = 18
$\sum \frac{T_i^2}{n_i} = \frac{2857^2}{6} + \frac{2490^2}{5} + \frac{3190^2}{7}$
= 4054156.7</td><td></td><td></td><td></td></tr>
<tr><td></td><td>SS Methods $= 4054156.7 - \frac{8537^2}{18}$
= 5247.3</td><td>M1</td><td></td><td>SS for methods</td></tr>
<tr><td></td><td>SS Total $= 4\,067\,243 - \frac{8537^2}{18}$
= 18333.6</td><td>M1</td><td></td><td>SS for total</td></tr>
<tr><td></td><td><table><tr><th></th><th>SS</th><th>df</th><th>ms</th></tr><tr><td>Methods</td><td>5247.3</td><td>2</td><td>2623.6</td></tr><tr><td>Error</td><td>13086.3</td><td>15</td><td>872.4</td></tr><tr><td>Total</td><td>18333.6</td><td>17</td><td></td></tr></table></td><td>m1
m1</td><td></td><td>Error SS ft (not –ve)
Method for MS – both correct ft incorrect df</td></tr>
<tr><td></td><td>$F = \frac{2623.6}{872.4} = 3.01$</td><td>m1
A1</td><td></td><td>Method for F ft
2.8 – 3.2 3.01/in range with no method seen allow 6 marks (or p = 0.080)</td></tr>
<tr><td></td><td>H_0 $\mu_A = \mu_B = \mu_C$
H_1 at least 2 of the means differ</td><td></td><td></td><td></td></tr>
<tr><td></td><td>$F^{2}_{15} = 3.682 > 3.01$</td><td>B1
B1</td><td></td><td>df correct
cv correct</td></tr>
<tr><td></td><td>Accept H_0. There is no significant evidence of a difference in(mean) reading achievement scores for the 3 methods. Allow no difference in teaching methods.</td><td>A1</td><td>9</td><td>correct ts/cv and conclusion in context</td></tr>
<tr><td>(b)</td><td>Assumptions:
Reading scores are normally distributed for each method</td><td>E1</td><td></td><td>Normal mentioned</td></tr>
<tr><td></td><td>The normal populations of reading scores have a common variance</td><td>E1</td><td>2</td><td>Explanations in some sort of context (scores appears) in one of the comments here
Disallow ‘random’, ‘no interaction’</td></tr>
<tr><td></td><td>Total</td><td></td><td>11</td><td></td></tr>
</table>