COURSE CODE: AAS 3211/AHT 3216

COURSE TITLE: INTRODUCTION TO STATISTICS/ INTRODUCTORY STATISTICS

EXAM VENUE: STREAM: (BSc. Horticulture, BSc. Animal Science and BSc. Soil Science)

DATE: 3/12/18 EXAM SESSION: 9.00 – 11.00AM

TIME: 2HOURS

Instructions
1. Answer ALL questions in Section A (compulsory) and ANY TWO questions in Section B
2. Candidates are advised not to write on the question paper
3. Candidates must hand in their answer booklets to the invigilator while in the examination room
SECTION A [30 MARKS]

Answer ALL questions from this Section.

QUESTION ONE [SIX MARKS]
Define the following terms with examples:

a. Population [2 Marks]
b. Measure of variability [2 Marks]
c. Discrete Qualitative variable [2 Marks]

QUESTION TWO [SIX MARKS]
For the data:

<table>
<thead>
<tr>
<th>55</th>
<th>85</th>
<th>90</th>
<th>50</th>
<th>110</th>
<th>115</th>
<th>75</th>
<th>85</th>
<th>8</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>65</td>
<td>50</td>
<td>60</td>
<td>90</td>
<td>90</td>
<td>55</td>
<td>70</td>
<td>5</td>
<td>31</td>
</tr>
</tbody>
</table>

Find the:

a. Mode [1 Mark]
b. Mean [2 Marks]
c. Median [3 Marks]

QUESTION THREE [NINE MARKS]
For the following measurements: 13, 21, 9, 15, 13, 17, 21, 9, 19, 23, 11, 9, 21. Find the:

a. Median [2 Marks]
b. Lower and upper quartiles [4 Marks]
c. Mean Quartile deviation [3 Marks]
QUESTION FOUR [NINE MARKS]

a. i. Define variables [1 mark]

ii. Distinguish between Quantitative and Qualitative variables [2 marks]

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
<th>Age</th>
<th>Status</th>
<th>Diagnosis number</th>
<th>State</th>
<th>Waist</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>M</td>
<td>36</td>
<td>1</td>
<td>9696</td>
<td>N</td>
<td>34.5</td>
<td>2.45</td>
</tr>
<tr>
<td>B</td>
<td>M</td>
<td>37</td>
<td>0</td>
<td>1967</td>
<td>S</td>
<td>37.6</td>
<td>1.89</td>
</tr>
<tr>
<td>C</td>
<td>F</td>
<td>38</td>
<td>1</td>
<td>9644</td>
<td>E</td>
<td>36.7</td>
<td>2.06</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>35</td>
<td>1</td>
<td>9678</td>
<td>W</td>
<td>34.9</td>
<td>1.98</td>
</tr>
<tr>
<td>E</td>
<td>M</td>
<td>43</td>
<td>0</td>
<td>9675</td>
<td>S</td>
<td>41.3</td>
<td>2.05</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>40</td>
<td>0</td>
<td>9645</td>
<td>S</td>
<td>44.3</td>
<td>2.41</td>
</tr>
<tr>
<td>G</td>
<td>M</td>
<td>38</td>
<td>1</td>
<td>9608</td>
<td>E</td>
<td>35.7</td>
<td>2.8</td>
</tr>
<tr>
<td>H</td>
<td>F</td>
<td>41</td>
<td>1</td>
<td>9602</td>
<td>W</td>
<td>38.4</td>
<td>2.63</td>
</tr>
<tr>
<td>J</td>
<td>F</td>
<td>32</td>
<td>1</td>
<td>9598</td>
<td>N</td>
<td>40.4</td>
<td>2.34</td>
</tr>
<tr>
<td>K</td>
<td>M</td>
<td>38</td>
<td>0</td>
<td>9578</td>
<td>S</td>
<td>43.9</td>
<td>1.95</td>
</tr>
<tr>
<td>L</td>
<td>M</td>
<td>44</td>
<td>0</td>
<td>9554</td>
<td>E</td>
<td>41.8</td>
<td>2.35</td>
</tr>
</tbody>
</table>

b. Use the data above to identify:

i. Variables [2 marks]

ii. Qualitative variables [2 marks]

iii. Discrete quantitative variables [2 marks]

SECTION B [40 MARKS]

Answer ANY TWO questions from this Section.

QUESTION FIVE [20 MARKS]

The height of 40 students in a statistics class were measured to the nearest centimetre and recorded as below:
Use the following data for this question:

175 154 157 180 165 150 152 162
173 168 169 181 177 179 175 169
151 153 156 158 163 169 179 180
145 149 150 156 171 175 176 178
169 160 155 174 170 176 182 170

a. Use class intervals of 5 to group the data [7 marks]
b. Estimate the mean height of the students [6 marks]
c. Calculate the median height [7 marks]

QUESTION SIX [20 MARKS]
a. For the probability distribution

<table>
<thead>
<tr>
<th>y</th>
<th>0</th>
<th>1</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(y)</td>
<td>1/5</td>
<td>2/5</td>
<td>1/10</td>
<td></td>
</tr>
</tbody>
</table>

i. What is p (0) [1 mark]

ii. Find the expected value of y [2 marks]

iii. Find the standard deviation of y [3 marks]

b. The average personal yearly income of farmers in a county is $6 200, with standard deviation of $ 400. i. If a sample of 64 people is randomly chosen from the county, find the probability that the mean income for the sample exceeds $ 6 300. [8 marks]

c. An agricultural store has the following entries in one of the records.

<table>
<thead>
<tr>
<th>Unit</th>
<th>Yield</th>
<th>Treatment</th>
<th>Variety</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.8</td>
<td>A</td>
<td>1</td>
<td>Good</td>
</tr>
<tr>
<td>2</td>
<td>7.8</td>
<td>B</td>
<td>1</td>
<td>Poor</td>
</tr>
<tr>
<td>3</td>
<td>9.7</td>
<td>C</td>
<td>2</td>
<td>Fair</td>
</tr>
<tr>
<td>4</td>
<td>6.4</td>
<td>D</td>
<td>2</td>
<td>Poor</td>
</tr>
<tr>
<td>5</td>
<td>8.3</td>
<td>A</td>
<td>3</td>
<td>Fair</td>
</tr>
<tr>
<td>6</td>
<td>10.6</td>
<td>B</td>
<td>3</td>
<td>Good</td>
</tr>
<tr>
<td>7</td>
<td>15.5</td>
<td>C</td>
<td>4</td>
<td>Excellent</td>
</tr>
<tr>
<td>8</td>
<td>14.8</td>
<td>D</td>
<td>4</td>
<td>Excellent</td>
</tr>
</tbody>
</table>
Using the table answer the questions:

i. Identify the variables in the dataset [2 marks]

ii. Classify Treatment and Variety in terms of their measurement scale [4 marks]

QUESTION SEVEN [20 MARKS]

a. i. Explain what is meant by a binomial experiment [2 marks]

When a seed is planted, it will either germinate or not with equal chances. Three identical seeds are planted in a pot.

ii. Write down all possible results of the experiment [6 marks]

iii. Determine the probability distribution of X, the random variable that a seed germinates. [6 marks]

b. Find:

i. The mean number of seeds that germinated. [3 marks]

ii. The standard deviation of the number of germinated seeds. [3 marks]

QUESTION EIGHT [20 MARKS]

a. In a horticultural farm the heights of newly germinated plants, X, are normally distributed with a mean of 2.20 cm and a standard deviation of 0.42 cm. Find the probabilities that:

   i. P (x ≤ 1.8) [2 marks]
   ii. P (X > 3.0) [3 marks]
   iii. P (2.8 ≤ X ≤ 3.1) [3 marks]

b. In a goat farm the animals reared included Sanen, Alpine and Toggenbug. The animals found to be either infested by worms or not as shown in the table below.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Animal</th>
<th>Sanen</th>
<th>Alpine</th>
<th>Toggenbug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infested</td>
<td></td>
<td>18</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Not infested</td>
<td></td>
<td>52</td>
<td>70</td>
<td>58</td>
</tr>
</tbody>
</table>
Find:

i. total number of goats in the farm \[1 \text{ mark}\]

ii. If a goat is found to be infested what is the probability that it is alpine or sanen \[3 \text{ marks}\]

iii. probability that a goat not infested is Toggenbug \[2 \text{ marks}\]

c. Probability that a seed loquat germinating is 0.85. Three such seeds are tested and \(X\) = number of seeds that germinate is observed.

i. Construct the probability distribution of the random variable \(X\) \[3 \text{ marks}\]

ii. find \(P (1 < X \leq 3)\) \[3 \text{ marks}\]