STAT 441 DISCUSSION 2

1 Review

- Continuous random variable and discrete random variable: the former picks continuous values, the latter picks discrete values.

- Graphical displays of data: dotplot, stem-and-leaf plot, frequency histogram, relative frequency histogram, density histogram.

- Note that in a density histogram, areas of rectangles are relative frequencies, which sum to one.

- The range of the data is an interval from the minimum data point to maximum data point.

- Some useful statistics: sample mean, sample median, sample quartiles ($Q_1$, $Q_2$, $Q_3$), sample mode, sample range, sample variance, sample standard deviation. Interquartile range (IQR) = $Q_3 - Q_1$.

- Compute quartile $Q_1$, $Q_2$, $Q_3$.

- A useful approach to find percentile $Q$.
  (1) draw density histogram.
  (2) identify the class interval that includes the desired percentile $Q$, say this interval is $[a, b)$.
  (3) $Q$ divides the rectangle on $[a, b)$ into two parts, denote the area of left part by $A$, area of right part by $B$.
  (4) use formula $Q = \frac{Ab + Ba}{A + B}$ to find $Q$.

2 Problems

(a). A sample of $n = 90$ people are grouped according to age in the following frequency table.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 10)</td>
<td>9</td>
</tr>
<tr>
<td>[10, 25)</td>
<td>27</td>
</tr>
<tr>
<td>[25, 65)</td>
<td>54</td>
</tr>
<tr>
<td>Total</td>
<td>90</td>
</tr>
</tbody>
</table>

Answer the following questions:

(i). Calculate the group mean age and group standard deviation. Express in years and months.

(ii). Construct a relative frequency histogram.

(iii). Construct a density histogram.

(iv). What percentage of the sample falls between 15 and 35 years old?

(v). Calculate the group quartile ages $Q_1$, $Q_2$, $Q_3$. Express in terms of years and months.

(vi). Calculate the range and the interquartile range. Express in terms of years and months.
(b). Class I has \( n_1 = 19 \) students, class II has \( n_2 = 21 \) students. Scores of all the students from two classes are recorded. Suppose the two classes have the same mean scores, but class I has standard deviation \( s_1 = 5 \), class II has standard deviation \( s_2 = 10 \). If the two classes are pooled together, then what is the standard deviation of the pooled class?

(c). Suppose that \( x_1, \ldots, x_5 \) are 5 numbers.

• Find \( \bar{x} \) and \( s_x \)

• Let \( z_i = \frac{x_i - \bar{x}}{s_x} \) be the standardized numbers. Find \( \bar{z} \) and \( s_z \).