Quiz I

STAT 305, Section 7
Spring 2020

Instructions

- The quiz is scheduled for 80 minutes, from 02:10 to 03:30 PM. At 03:30 AM the exam will end.
- Total points for the exam is 88. Points for individual questions are given at the beginning of each problem. Show all your calculations clearly to get full credit. Put final answers in the box at the right (except for the diagrams!).
- A forumula sheet is attached to the end of the exam. Feel free to tear it off.
- Normal quantile table is attached to the end of the exam. Feel free to tear it off.
- You may use a calculator during this exam.
- Answer the questions in the space provided. If you run out of room, continue on the back of the page.
- If you have any questions about, or need clarification on the meaning of an item on this exam, please ask your instructor. No other form of external help is permitted attempting to receive help or provide help to others will be considered cheating.
- **Do not cheat on this exam.** Academic integrity demands an honest and fair testing environment. Cheating will not be tolerated and will result in an immediate score of 0 on the exam and an incident report will be submitted to the dean's office.

Name: ______________________________

Student ID: ______________________________

1. (2 points) Circle the **bold face** term that makes the following statement true:

   A measurement device that reports the measurements which are close to each other when repeatedly measuring the same thing is (**precise** or **accurate**).

2. (2 points) A number of issues concerning measurement must be addressed in the following order:

   (1) precision, validity, accuracy (2) accuracy, precision, validity
   (3) validity, accuracy, precision (4) validity, precision, accuracy

3. (2 points) For a complete(full) factorial study with three factors, each with 4 levels, the number of all possible combinations (i.e the least number of observation) is:

   (1) 12 (2) 64 (3) 81 (4) none of these

   $(\# \text{ of levels})^{\# \text{ of factors}} = 4^3 = 64$

4. (2 points) In a series of experiments to study the priority of a chemical product, the reactor temperature is set fixed at 550°C. The variable "reactor temperature" is a

   (1) response variable (2) controlled variable
   (3) blocking variable (4) experimental variable

5. A sample of size 6 was drawn from a population and the resulting observations are reported below.

110, 100, 105, 103, 105, 115

Using these observed values, report the following:

sorted data: 100, 103, 105, 105, 110, 115

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6$

(a) (2 points) the mean $\boxed{\bar{x} = 106.33}$

$$\bar{x} = \frac{1}{6}(x_1+x_2+x_3+x_4+x_5+x_6) = \frac{1}{6}(100+103+105+105+110+115)$$

$$= 106.33$$

(b) (3 points) the median $\boxed{Med. = 105}$

$Q(.25) \qquad Q(.75)$

$\frac{i-0.5}{n} = \frac{i-0.5}{6}$: $0.08 \quad \boxed{0.25} \quad 0.41 \quad 0.58 \quad \boxed{0.75} \quad 0.91$

$Q(.5) = ?$ $np + 0.5 = 6(0.5) + 0.5 = 3.5$, $[3.5] = 3 = i$

$Q(.5) = x_3 + (np - i + 0.5)(x_4 - x_3)$

$= 105 + (0.5)(105 - 105) = 105$

(c) (5 points) the variance $\boxed{s^2 = 28.66}$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{6-1}\Big((100-106.33)^2+(103-106.33)^2+(105-106.33)^2$$
$$+(105-106.33)^2+(110-106.33)^2+(115-106.33)^2\Big)$$

$$= \frac{143.33}{5} = 28.66$$

(d) (2 points) the standard deviation $\boxed{s = 5.35}$

$$s = \sqrt{s^2} = \sqrt{28.66} = 5.35$$

(e) (3 points) the value of $Q(.75)$ $\boxed{Q(.75) = 110}$

① Note: you could have found it just by looking at $\frac{i-0.5}{n}$ values. Then $Q(.75) = 110$.

② $n.p + 0.5 = 6(.75) + 0.5 = 5$, $[5] = 5 = i$

$Q(.75) = x_5$

$= 110$

(f) (4 points) the interquartile range

$IQR =$ 7

$IQR = Q(.75) - Q(.25)$

$Q(.25) = ?$

$np + 0.5 = 6(.25) + 0.5 = 2, \; [2] = 2 = i$

$Q(.25) = x_2 = 1.3$

$\Rightarrow IQR = 110 - 1.3 = 67$

(g) (5 points) give the coordinates (on a regular graph paper) of the upper right and lower left point that would appear on a normal plot of the data.

upper right: $\frac{6-0.5}{6} = 0.91 \longrightarrow Q(.9) = 115$

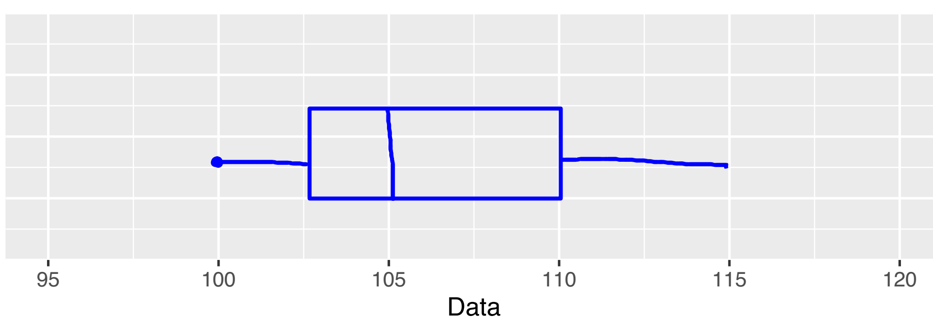
lower left on Normal graph $\Rightarrow$ Normal quantile of $\frac{1-0.5}{6} = .08$

by normal quantile table: $Q_N(.08) = -1.41$

upper right point = (0.91 , 115)

lower left point = (.08 , -1.41)

(h) (5 points) Using the axes below, create a box plot to summarize the data. Carefully label the axes



Whisker length: $1.5 * IQR = 1.5 * 7 = 10.5$

$Q(0.25) - 10.5 = 103 - 10.5 = 92.5$

$Q(0.75) + 10.5 = 110 + 10.5 = 120.5$

6. An environmental engineer is testing four methods for reducing the concentration of a certain lake pollutant found in Iowa lakes. To do this he first randomly selected 20 Iowa lakes from which he took water samples, then split each of the 20 samples into 4 portions, and randomly labeled the four portions 1, 2, 3, and 4. Finally, he attempted to reduce the concentration of each of the portions labeled 1 using the the first method, of each of the portions labeled 2 using the second method, of each of the portions labeled 3 using the third method, and of each of the portions labeled portion 4 using the fourth method. After the methods had been applied, he measured the change in concentration.

   (a) (3 points) Is this an experiment or an observational study? Explain.

   Experimental.
   The engineer is taking an active role in manipulation of the system under study by intentionally changing the cleaning meteds.

   (b) (2 points) What is the population under study?

   Iowa lakes

   (c) (2 points) What is the sample under study?

   20 selected Iowa lakes

   (d) Identify the following (if there was not one, simply put ”not used”).

      i. (2 points) Response variable(s):

      change in the concentration is the only response variable

      ii. (2 points) Experimental variable(s):

      The method used to clean the polution is the only experimental variable.

      iii. (2 points) What type of variable was (were) the experimental variable(s) in previous question (circle one):

      Quantitative **Qualitative**

iv. (2 points) Blocking variable(s):

The lakes from which the samples are taken are acting as a blocking variable. We are not interested in studying the effect of the lake on the response, but we can reasonably believe that the portions from the same lake's sample will be similar. So we are treating the lake the sample came from as a smaller, homogenous environment in our experiment. We also use all the methods on each lake's sample which is another indication that it is working as a block

(e) (2 points) Was replication used in this experiment? If so, where was it applied? If not, how could we have applied it?

No replication was used. While each cleaning method was used multiple times across the entire experiment, they were never used in the same block (meaning, for each of the 20 lakes we only used each method once). This means that we did not truly replicate.

7. A certain make of load bearing beam used in construction of large buildings is certified to withstand pressures of 2.5 tonnes per square inch. 30 beams were tested and the pressures at which they failed are collected in the table below.

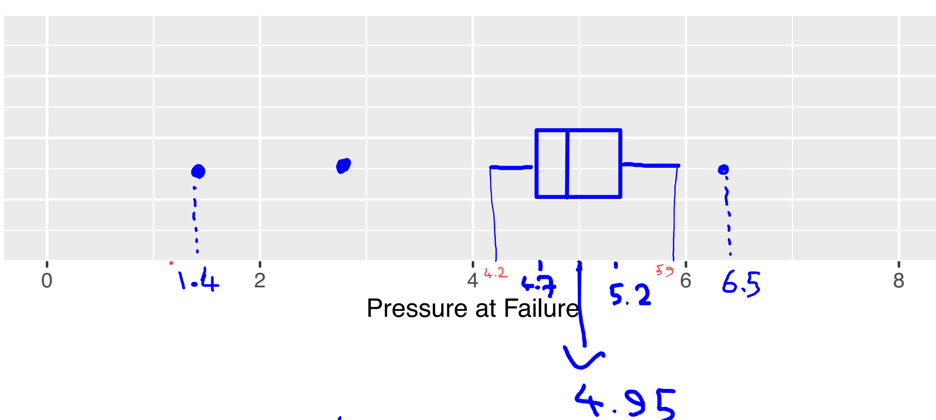
```
The decimal point is at the |

1 | 4
2 | 8
3 |
4 | 2555778899999
5 | 01122222234559
6 | 5
```

Note that here `1 | 4 represents 1.4`. In this case, the first quartile is $Q(.25) = 4.7$, the median is 4.95, and the third quartile is $Q(.75) = 5.2$.

(a) (10 points) Complete the following frequency table:

Value Range	Frequency	Relative Frequency	Cumulative Relative Frequency
0.00 - 2.00	1	$1/30 = 0.0\overline{33}$	$0.0\overline{33}$
2.01 - 4.00	1	$1/30 = 0.0\overline{33}$	$0.0\overline{66}$
4.01 - 6.00	27	$27/30 = 0.9$	$0.9\overline{66}$
6.01 - 8.00	1	$1/30 = 0.0\overline{33}$	1

(b) (10 points) Using the axes below, create a box plot to summarize the data. Carefully label the axes.



$IQR = Q(.75) - Q(.25)$

$= 5.2 - 4.7 = 0.5$

$1.5 * IQR = 0.75$

$Q(0.25) - 0.75 = 3.95$

$Q(0.75) + 0.75 = 5.95$

(c) (4 points) Are there any unusually low or high observations? If so, what pressures caused those beams to fail?

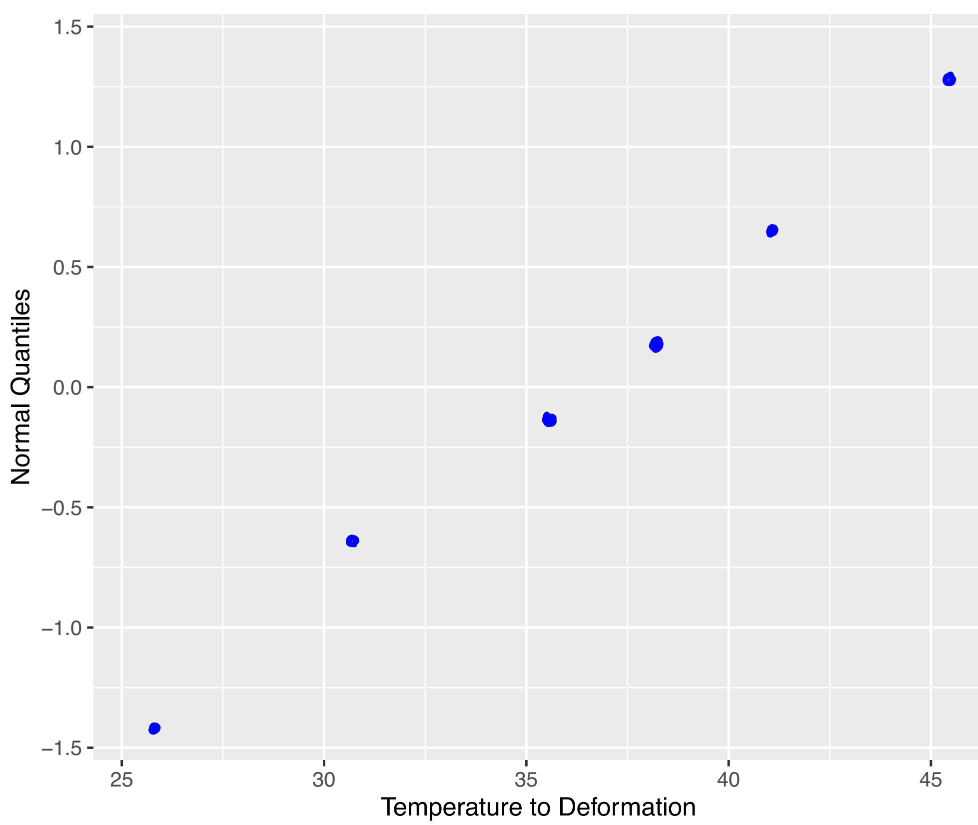
There are 3 (Potential) outliers in the data, at pressure level 1.4 & 2.8 & 6.5

(d) (10 points) The company also measured the heat at which the beams begin to deform (in 10,000 degrees Celsius). The values are collected below:

31.3, 41.8, 25.3, 35.3, 45.3, 37.3

First complete the table for the quantiles of the data ($Q(p)$) and then create a theoretical Q-Q plot using the following quantiles from the normal distribution as the theoretical quantiles.

i	1	2	3	4	5	6
$p = \frac{i-.5}{6}$	0.08	0.25	0.42	0.58	0.75	0.92
$Q_N(p)$	-1.41	-0.67	-0.2	0.2	0.67	1.41
$Q(p)$	25.3	31.3	35.3	37.3	41.8	45.3



What does this graph tell us about the temperature at which the beams deform?

The distribution (shape) of the the temp. at which the beams deform is Bell-shape. i.e. the data has a Normal distribution.