

Math 254 Assignment One

Name: _____

Due: In class on Monday July 25

Assignments must be completed on this paper. Marks may be deducted for not showing all your work.

1. [3 marks] A data set consisting of 320 measurements has a mean of 15 and a SD of $\sigma = 4$. What can you say about the number of measurements that are greater than 31?

$$\mu = 15 \quad \sigma = 4 \quad n = 320$$

$$31 = \mu + k\sigma \quad \text{Find } k.$$

$$31 = 15 + k(4)$$

$$\underline{k = 4}$$

$$\begin{aligned} \text{We know } &\geq \left(1 - \frac{1}{4^2}\right) 320 \\ &\geq 300 \text{ measurements} \end{aligned}$$

$$\begin{aligned} \text{lie in the interval } &[\mu - 4\sigma, \mu + 4\sigma] \\ &[-1, 31] \end{aligned}$$

So ≤ 20 measurements lie outside the interval.

≤ 20 measurements are greater than 31.

⊖ if you assumed data set is symmetric

2. [4 marks] a) Write the five-number summary for the data set below:

~~8, 12, 12, 13, 14, 15, 19, 22, 26, 26, 27, 27, 48~~

8, 12, 12, 13, 14, 15, 19, 22, 26, 26, 27, 27, 48

$$\text{Min} = 8$$

$$Q_1 = \frac{12+13}{2} = 12.5$$

$$\text{Median} = 19$$

$$Q_3 = \frac{26+27}{2} = 26.5$$

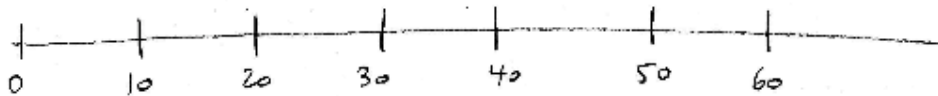
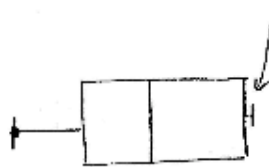
$$\text{Max} = 48$$

b) Draw a box plot for the data set above.

next largest
measurement
is 27

$$\text{IQR} = Q_3 - Q_1 = 14$$

$$1.5 \text{ IQR} = 21 \leftarrow \text{longest whisker allowed}$$



48 is an outlier

3. [3 marks] The following data gives systolic blood pressure (Y) versus a person's age in years (X).

Age	34	61	25	55	57	62	41	49	48	33
BP	120	175	105	180	140	145	140	160	155	120

a) Find the equation of the linear regression line. Round the coefficients to two decimal places.

$$y = bx + a \quad y \approx 1.54x + 72.32$$

b) Find the value of the correlation coefficient, rounded to two decimal places.

$$r \approx 0.81$$

c) Using your regression line from part a), what age has a predicted BP of 153.94?

$$153.94 = 1.54x + 72.32$$

$$x = 53$$

4. [3 marks] A red die, a blue die and a green die are rolled. Find the probability that the rolls sum to at most five.

$$n(S) = 6 \times 6 \times 6 = 216$$

Rolls summing to ≤ 5

111, 112, 121, 211, 221, 212, 122,
RBG, RBG, RBG, RBG, RBG, RBG, RBG

113, 131, 311
RBG, RBG, RBG

$$n(A) = 10$$

$$P(\text{rolls sum to } \leq 5) = \frac{10}{216} \approx 0.05$$

5. [4 marks] An employee PIN uses digits 1 through 9, and may be six or seven digits long. How many PIN's have at least one repeated digit?

$$\text{Total \# PIN'S} = 9^6 + 9^7$$

6 digits long OR 7 digits

$$\# \text{ with no repeats} = 9P6 + 9P7$$

6 digits OR 7 digits

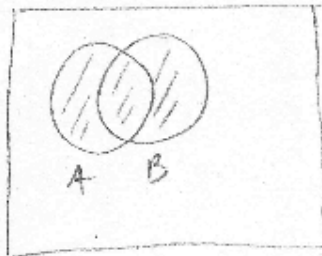
$$\# \text{ with } \geq 1 \text{ repeat} = (9^6 + 9^7) - (9P6 + 9P7) \quad [\text{As a \#: } 5,072,490]$$

6. [3 marks] A factory has four production lines. Line A produces 11% of all ball bearings; Lines B, C and D produce 31%, 22% and 36% respectively. Of the ball bearings produced on Line A, 0.3% are defective; for Lines B, C and D the percentages are 0.2%, 0.1% and 0.15% respectively. What is the probability that a defective ball bearing comes from Line B?

given that $d = \text{defective}$
it is defective
 \downarrow
 $P(B|d)$
 $= \frac{0.31 \times 0.002}{[0.11 \times 0.003 + 0.31 \times 0.002 + 0.22 \times 0.001 + 0.36 \times 0.0015]}$
 ≈ 0.363

7. [5 marks] You are given $P(A) = 0.4$, $P(B) = 0.58$ and $P(A \cup B) = 0.76$.

a) Find $P(A \cap B)$, $P(A|B)$ and $P(B|A)$



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$0.76 = 0.4 + 0.58 - P(A \cap B)$$

$$P(A \cap B) = 0.22$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \approx 0.38$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = 0.55$$

b) Are A and B independent? Justify your answer.
Notice $B \cap A$ is the same as $A \cap B$

No, $P(A|B) \neq P(A)$

[Alternatively: $P(B|A) \neq P(B)$
 $P(A \cap B) \neq P(A)P(B)$]

Only need to check 1 of the 3 rules.