Name: *Caleb McWhorter — Solutions* **MAT 221**
Fall 2014 **Problem Set 3**

Problem 1: Fill in the blank:

(a) The five number summary consists of the

min, Q_1, $median$, Q_3, and max.

(b) Bar charts are for $categorical$ variables while box plots and histograms are for $quantitative$ variables.

(c) You look for outliers by calculating $1.5 \cdot IQR$ then looking for numbers less than $median - 1.5 \cdot IQR$ or numbers greater than $median - 1.5 \cdot IQR$.

(d) If the mean is less than the median, then the distribution of data is $left$ skewed.

(e) The z-score measures how many $standard\ deviations$ a data point is from the $mean$.

(f) Scatterplots show the relationship between what kind of variables? $quantitative$

Problem 2: The scores on a university examination is normally distributed with mean 62 and standard deviation 11.

(a) What proportion of the students scored at least 80?

$$z_{80} = \frac{80-62}{11} = \frac{18}{11} = 1.63 \rightsquigarrow 0.9484$$

Therefore, $P(\geq 80) = 1 - P(< 80) = 1 - 0.9484 = 0.0516$.

(b) What proportion of the students scored between 70 and 80?

$$z_{80} = \frac{80-62}{11} = \frac{18}{11} = 1.63 \rightsquigarrow 0.9484$$

$$z_{70} = \frac{70-62}{11} = \frac{8}{11} = 0.73 \rightsquigarrow 0.7673$$

Therefore, the proportion of students scoring between 70 and 80 is $0.9484 - 0.7673 = 0.1811$.

(c) If the top 5% of students are awarded a merit certificate, what is the lowest mark that a student can have and still be awarded a merit certificate?

top 5% $\rightsquigarrow$ *bottom 95%* $\rightsquigarrow 1.645 = z$. *Then*

$$1.645 = z = \frac{x-62}{11}$$

Then $18.095 = x - 62$ *so that* $x = 80.095$. *Therefore, a student must score at least* 80.095 *to be in the top 5% of students taking the exam.*

Problem 3: There were 8 students in a class. The average grade (out of 100) of each student and her/his score on the final exam (out of 100) were recorded. The record is given below:

Observation #	1	2	3	4	5	6	7	8	Mean	StDev.
Quiz (x)	85	78	99	87	79	71	88	99	85.75	9.867
Final exam (y)	80	72	98	85	82	65	92	90	83	10.784

The correlation for a linear regression for this data was $r = 0.9049$.

(a) Find the equation of the least square regression line.

$b_1 = r\frac{s_y}{s_x} = 0.9049 \cdot \frac{10.784}{9.867} = 0.99$ *and* $b_1 = \overline{y} - b_1\overline{x} = 83 - 0.99 \cdot 85.75 = -1.89$. *Therefore,* $y = 0.99x - 1.89$.

(b) Use the regression line to predict the final exam score of a student whose average quiz grade is 82.

$y = 0.99(82) - 1.89 = 79.29$

(c) Calculate the residual for observation #6.

$y = 0.99(71) - 1.89 = 68.40$. *We have residual* $=$ *observed* $-$ *predicted* $= 65 - 68.40 = -3.40$.

Problem 4: Consider the following numbers:

$$4 \quad 17 \quad 18 \quad 19 \quad 23 \quad 25 \quad 27 \quad 27 \quad 29 \quad 32 \quad 35 \quad 40 \quad 42 \quad 44 \quad 46 \quad 61$$

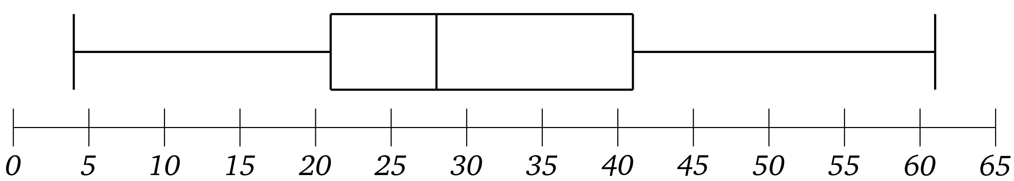
(a) Compute the five number summary for the data above.

Min	Q_1	*Median*	Q_3	*Max*
4	*21*	*28*	*41*	*61*

(b) Are there any outliers? Justify your answer.

$1.5 \cdot IQR = 1.5 \cdot (41 - 21) = 1.5 \cdot 20 = 30$. *Then* $Q_1 - 1.5 \cdot IQR = 21 - 30 = -9$ *and* $Q_3 + 1.5 \cdot IQR = 41 + 30 = 71$. *Therefore, there are no outliers.*

(c) Sketch a box plot for the above data set.



(d) Compute the mean and standard deviation for the data above.

$\overline{x} = 30.56$ *and* $s = 13.87$