

# Math 263

Exam I

February 5th 2015

The University of Arizona

KEY

Name: \_\_\_\_\_

Answers without adequate justification will not receive full credit, including multiple choice. Include units with your answer when appropriate, and box all answers unless an answer line is provided. By signing below I am agreeing to abide by the University of Arizona academic integrity policies and that all work done on this test is my own.

Signature: \_\_\_\_\_

### Tips for Success:

- Look through the entire test before starting to prioritize questions.
- If you get stuck on a question, move on and come back to it later.
- Do a quick reality check after each question: does my answer make sense? Did I include units? Did I show all my work?
- Read over the entire test at the end to make sure you didn't miss anything.
- For each question: take a deep breath, think slowly and deliberately at first, then work quickly once you see what to do.

### Special notes for 263 exams:

- If you use a calculator to compute statistical quantities such as a 5 number summary, linear regression, normal proportion etc, please indicate which function you used i.e. 1-Var Stats, LinReg(ax+b), normalcdf().
- Round all numerical answers to 2 decimal places unless otherwise specified.

1. Alison's IQ has a standardized score ( $z$ -score) of 1.5, as compared to other people her age.

(a) Which of the following is the correct interpretation of Alison's  $z$ -score?

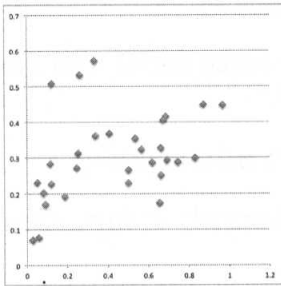
- Alison's IQ is 150
- Alison's IQ is 1.5 above the mean IQ for people of her age.
- Alison's IQ is 1.5 times the mean IQ of people her age.
- Alison's IQ is 1.5 standard deviations above the mean IQ of people her age.
- Only 1.5% of people Alison's age have an IQ higher than hers

(b) What percentage of people Alison's age have an IQ lower than her? Round your answer to 2 decimal places.

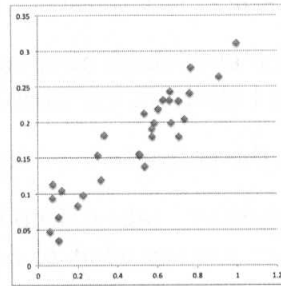
$z = 1.5$  so use  $z$ -table or Normal CDF ( )

To get  $F(1.5) \approx 0.9332$  or about 93.32%  
 ↑  
 CDF of  $N(0,1)$

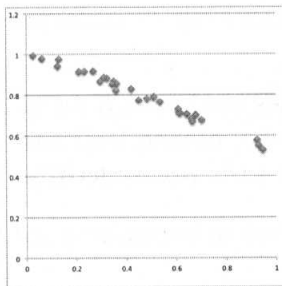
2. Choosing from the values  $r = 0.5, 0.9, -0.95, -0.01$ , label the following scatter plots with their approximate correlation coefficient.



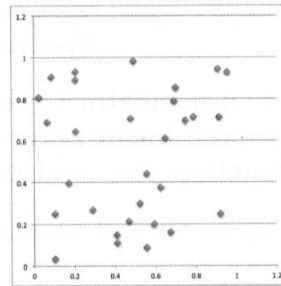
$r \approx 0.5$



$r \approx 0.9$



$r \approx -0.95$



$r \approx -0.01$

3. The mean home price in Seattle is currently \$475,000. Assume that home prices in Seattle are normally distributed with this mean and a standard deviation of \$100,000.

(a) *Approximately* what proportion of homes in Seattle are cheaper than \$275,000?

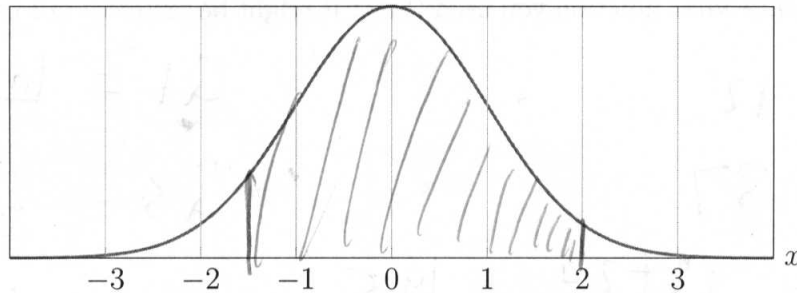
- 68%    95%    5%    2.5%    99.7%

"95% Rule: About 95% are between 275k & 675k"

(b) Find the z-scores for a home priced \$325,000 and a home priced \$675,000 (label them  $z_{325}$  and  $z_{675}$ ).

$$z_{325} = \frac{325 - 475}{100} = \boxed{-1.5} \quad z_{675} = \frac{675 - 475}{100} = \boxed{2}$$

(c) Shade the area on the standard normal distribution below corresponding to the proportion of houses worth between \$325,000 and \$675,000.



(d) Find the proportion of houses in Seattle that are worth between \$325,000 and \$675,000.

$$F(2) - F(-1.5) \approx \cancel{0.9772 - 0.0668} \\ = \boxed{0.9104} \text{ or about } \underline{91.04\%}$$

(e) What is the price of a home at the 90th percentile? Hint: the proportion of homes cheaper will be 0.9.

Want z-score such that  $F(z) = 0.9$   
 Use table or  $\text{InvNorm}(0.9, 475, 100)$   
 To get  $\sim \boxed{\$603,155}$  ( $z \approx 1.28$ )

4. The Environmental Protection Agency records data on the fuel economy of different makes of cars. Some of the variables they collect are listed below.

(a) Identify the following variables as *quantitative* or *categorical*:

(i) Manufacturer (Ford, Toyota, etc) *Categorical*

(ii) Weight of the car (in pounds) *Quantitative*

(iii) Size of the car (compact, full-size, pickup, etc) *Categorical*

Data on the gas mileage of 20 randomly selected cars are listed below. The data is sorted for convenience.

12	13	15	16	16	18	18	18	19	19
20	20	22	23	24	24	26	27	31	37

(b) Find the 5 number summary for this data set. If you use a calculator, make sure to indicate what function you used (hint: it might be faster to *not* use your calculator).

$$\text{Min} = 12$$

$$\text{max} = 37$$

$$\text{med} = \frac{19+20}{2} = 19.5$$

$$Q1 = \frac{16+18}{2} = 17$$

$$Q3 = \frac{24+24}{2} = 24$$

(c) Using the  $1.5 \cdot \text{IQR}$  rule, does this data have any outliers?

$$\text{IQR} = 24 - 17 = 7 \quad \text{so} \quad 1.5 \cdot \text{IQR} = 10.5$$

$$Q1 - 1.5 \cdot \text{IQR} = 6.5$$

$$Q3 + 1.5 \cdot \text{IQR} = 34.5$$

So only  $\boxed{37}$  is a potential outlier

5. There are two major tests of readiness for college, the ACT and the SAT. ACT scores are reported on a scale from 1 to 36. The distribution of ACT scores for more than 1 million students in a recent high school graduating class was roughly normal with mean  $\mu = 20.8$  and standard deviation  $\sigma = 4.8$ . SAT scores are reported on a scale from 400 to 1600<sup>1</sup>. The SAT scores for 1.4 million students in the same graduating class were roughly normal with mean  $\mu = 1026$  and standard deviation  $\sigma = 209$ .

Suppose a student scores a 1310 on the SAT. Assuming both tests are measuring the same thing, what score on the ACT is equivalent to this score? Hint: standardized scores.

$$Z_{\text{SAT}} = \frac{1310 - 1026}{209} \approx 1.36$$

so

$$Z_{\text{ACT}} = 1.36 \quad \text{if}$$

$$\frac{X_{\text{SAT}} - 20.8}{4.8} = 1.36$$

$$\Rightarrow X_{\text{SAT}} = 1.36 \cdot 4.8 + 20.8 \approx \boxed{27.33}$$

6. A study found a correlation of  $r = -0.61$  between the gender of a worker and his or her income. For each of the following statements, circle T or F to indicate if the statement is true or false.

- (a) (T/F) The study shows that women earn less than men on average
- (b) (T/F) An arithmetic mistake was made. Correlation must be positive.
- (c) (T/F) This measurement makes no sense -  $r$  can only be measured between two *quantitative* variables.

<sup>1</sup>At least they were when this data was collected - the SAT is now out of 2400.



- (c) For the lean body mass of 36.1 kilos, what is the *predicted* value given by the regression line? *Show work.*

$$LBM = 0.036 MR + (-0.47)$$

From data table, MR for 36.1 is 995

So  $LBM = 0.036(995) - 0.47 = 35.35$  (c) 35.35 kilos

- (d) What is the *residual* for the lean body mass of 36.1? *Show work.*

$$\begin{aligned} \text{Resid} &= \text{ACTUAL} - \text{Predicted} \\ &= 36.1 - 35.35 \\ &= 0.75 \text{ kilos} \end{aligned}$$

(d) \_\_\_\_\_

- (e) What fraction of the variation in lean body mass is explained by variation in metabolic rate? Round your answer to two decimal places. *Show work.*

This is the "definition" of  $R^2$ .

From calculator,  $R^2 = 0.8822$

(e) \_\_\_\_\_