Math 207
Introduction to Statistics

Regression Review Exercises

1. Here is some data:

x	y
0	1
1	3
2	2

a) Sketch the data and the line $y = x + 1$ then calculate the RMS error for the line

b) Sketch the data and the line $y = \frac{3}{2}x + 1$ then calculate the RMS error for the line

c) Sketch the data and the line $y = \frac{2}{3}x + 1$ then calculate the RMS error for the line

d) For the given data, $\text{sd}_x = \text{sd}_y = \sqrt{2/3}$ and $r = 1/2$. Calculate the sd line and find its RMS error.

e) Calculate the regression line and find its RMS error.

2. For a given data set we find that mean$(x) = 10$, mean$(y) = 0$, $sd_x = 1$, $sd_y = 2$ and $r = 1/5$.

a) Predict y if $x = 11$.

b) If $x = 11$, about what range of y values contains about 68% of the data?

c) If $x = 11$, about what range of y values contains about 95% of the data?

d) If $x = 8$, about what range of y values contains about 68% of the data?

e) If $x = 8$, about what range of y values contains about 95% of the data?

f) Predict x if $y = 2$.

g) If $y = 2$, about what range of x values contains about 68% of the data?

h) If $y = 2$, about what range of x values contains about 95% of the data?

i) If $y = -1$, about what range of x values contains about 68% of the data?

j) If $y = -1$, about what range of x values contains about 95% of the data?

3. Calculate the given area under the standard normal curve.

a) to the left of $z = 1.5$

b) to the right of $z = 2.5$

c) between -1.5 and 2.5

4. For what value of z is each statement true?

a) area to the left of z is 35%

b) area to the right of z is 10%

c) area between z and $-z$ is 10%

5. Suppose a measurement has a normal distribution with mean = 10 and sd = 4. Calculate the percent of measurements that satisfy the following.

a) at most 16

b) at least 0

c) between 4 and 20