

## Probability & Statistics Overview

1. You have a fair 3-sided die with 3 possible outcomes  $\{1\}, \{2\}$  and  $\{3\}$ .

- a) If you roll the die once, what is the probability of obtaining a 3?  $\frac{1}{3}$
- b) If you roll the die once, what is the probability of obtaining an odd number?  $\frac{2}{3}$
- c) If you roll the die twice, what is the probability of obtaining  $[1,1]$ ?  $\frac{1}{9}$
- d) If you roll the die twice, what is the probability of obtaining  $[1,3]$ ?  $\frac{1}{9}$

2. Given the following table of blood groups and gender

Blood Group	Males	Females	Total
O	22	18	40
A	10	25	35
B	10	3	13
AB	8	4	12
Total	50	50	100

- a) What is  $P(B \cap \text{Male})$ ?  $\frac{10}{100} = 0.1$
- b)  $P(\text{Male})$   $\frac{50}{100} = 0.5$
- c)  $P(\text{Female} | A)$   $\frac{25}{35} = 0.71$

3. The outcome of an American Roulette in the casino is the following:

1,22,17,27,35,11,3,18,0,36,14,17,12,21,25,34,7,5



Assuming that you know for sure that the roulette is fair, is it a good idea to bet for a black number next or it irrelevant?

NO, It is irrelevant

4. You toss a normal 6-sided dice twice, what is the probability of obtaining the sequence [1,3]?

$$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

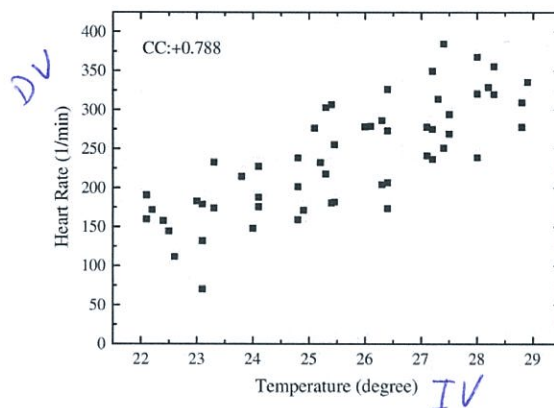
5. Let say you have a population of people: [Mary, Daniel, Josh, Susan, Michael, Peter, Andreas, Martin, Craig]. Draw a sample of size 3 out of this population.

[Mary, Martin, Andreas]

6. Classify the following variables describing a human as quantitative or categorical:

1. ethnicity *Categorical*
2. height *Quantitative*
3. IQ *Quantitative*
4. gender *Categorical*

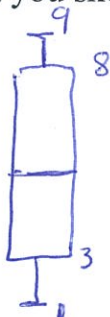
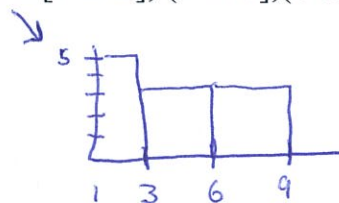
7. In the following relationship between variables Temperature and Heart Rate, which one is the independent variable and which one is the dependent variable?



8. For a random variable X we draw a sample S such that S = [5, 6, 1, 3, 8, 9, 2, 3, 3, 4, 9]

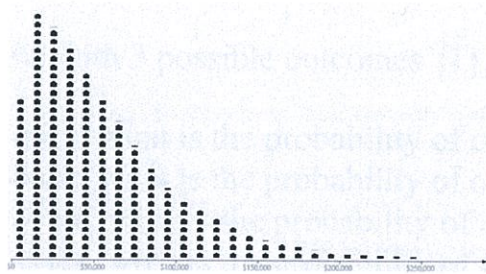
- a) Find the mean of S *4.82*
- b) What is the median of S? *4*
- c) What is the mode of S? *3*
- d) Manually sketch a histogram of the data using bins: [1 to 3], (3 to 6], (6 to 9]
- e) What is the range of the data? *[1-9]*
- f) What is the variance of S? *7.24*
- g) What is the standard deviation of S? *2.69*
- h) Can you sketch a rough box plot of S

*1, 2, 3, 3, 3, 4, 5, 6, 8, 9, 9*



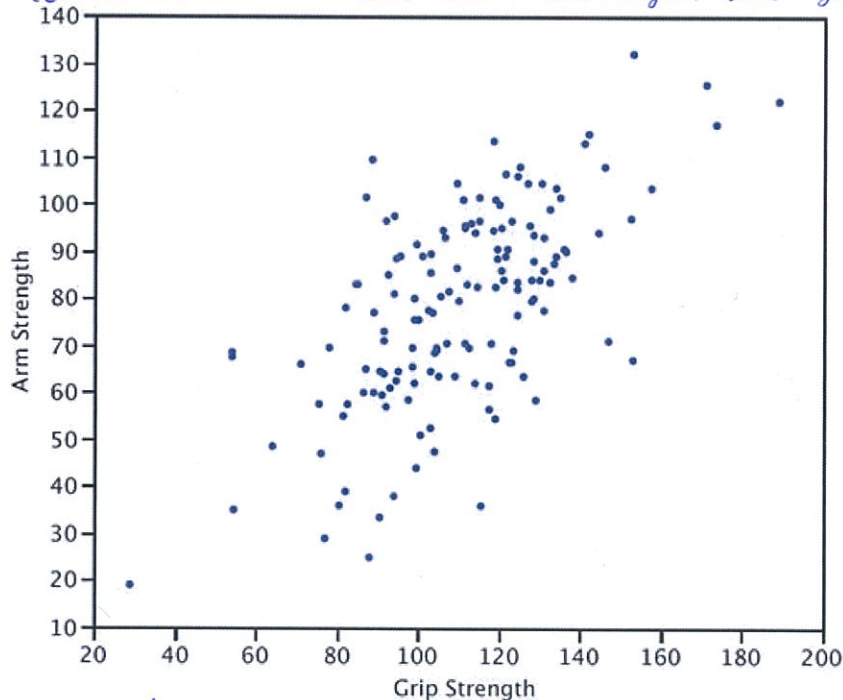
2, 2015

9. In the following distribution of household income, which is higher the median or the mean? *The Mean*



10. In your own words, explain what does grip strength suggests about arm strength according to the following scatter plot? Bonus question: can you infer from this graph that grip strength **causes** someone to have a lot of arm strength?

*Positive correlation between arm strength and grip strength*



*NO, correlation DOES not imply causation*

11. Identify which of the following distributions is: uniform, bimodal, positively skew, normal, negatively skew

