HOMEWORK 1

STAT S 520

RAMPRASAD BOMMAGANTY (rbommaga)
21 JANUARY 2016

SOLUTIONS:

1. (a) Given that there are 59 students who developed polio out of the 201,000 students who were given the vaccine. The percentage calculation is as follows: (59/201000)*100 = 0.029%.

   Given that there are 142 students who developed polio out of the 201000 students who were given the placebo. The percentage calculation is as follows: (142/201000)*100 = 0.07%.

   (b) Given that there are 56 students who developed polio out of the 222,000 students who were given the vaccine. The percentage calculation is as follows: (56/222000)*100 = 0.025%.

   Given that there are 54 students who developed polio out of the 124,000 students who were not given the vaccine. The percentage calculation is as follows: (54/124000)*100 = 0.04%.

   (c) The polio percentage in the placebo group is higher than that of the percentage in the unvaccinated group. The reason being, the first experiment was a randomized controlled experiment whereas the study by NFIP did not involve any randomization. The participants in the second experiment are not distributed randomly, thereby leaving nothing to chance.

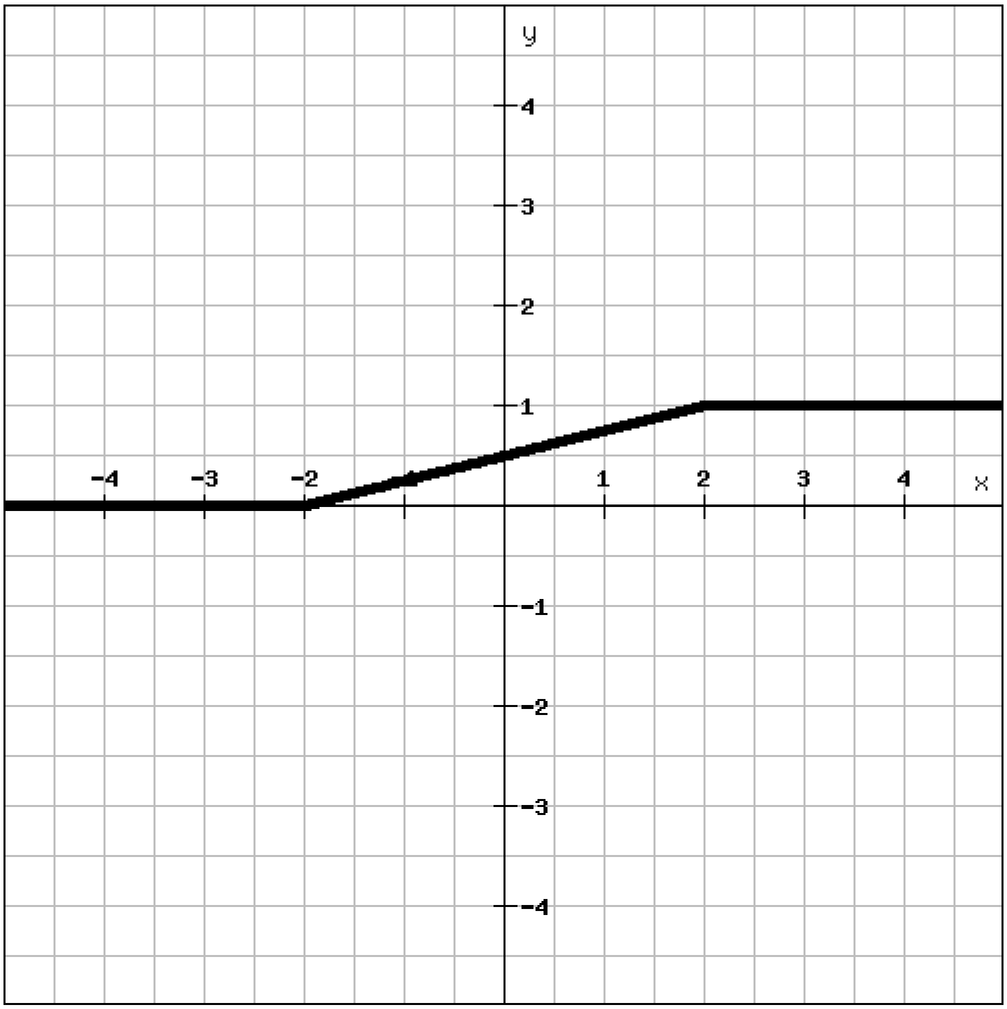
   The polio percentage in the NFIP unvaccinated group is strongly biased. Since the students were not picked randomly, the allocators might have consciously or sub-consciously allocated students belonging to the socially weaker sections to the vaccinated groups. The lack of randomization leads to the involvement of socio-economic status into the study which leads to incorrect and biased results.

2. Given function,

$$F(y) = \begin{cases} 0\ , y < -2 \\ y+2/4,\ -2 \leq y < 2 \end{cases}$$

1, $y \geq 2$

The graph is as follows:



(b) The formal mathematical expression for the piecewise function given is as follows:

$$F(y) = \begin{cases} 0.0\ , & y < 2 \\ 0.5, & 2 \leq y < 3 \\ 1.0, & y \geq 3 \end{cases}$$

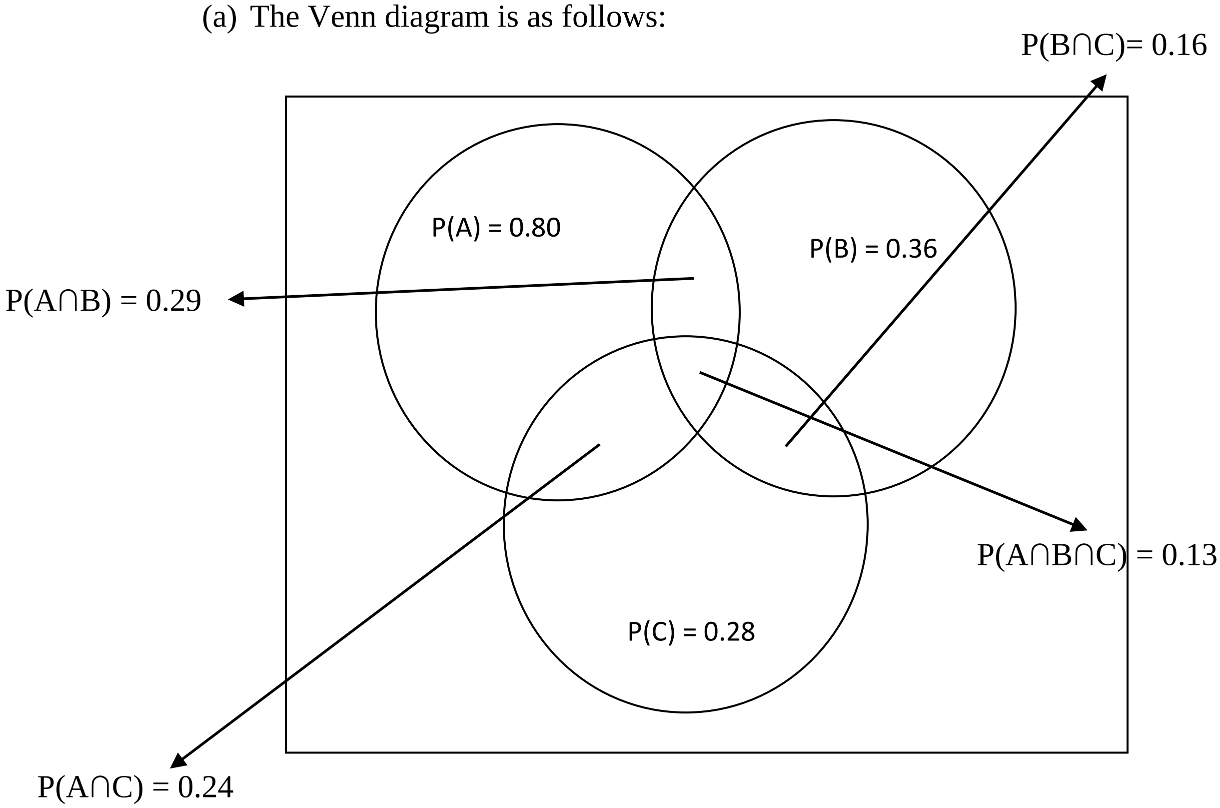
The graph is divided into three parts and is bounded at y = 2 and y = 3.

3. Given that,

   A = {quartz specimens are found}
   B = {tourmaline specimens are found}
   C ={aquamarine specimens are found }

   $P(A) = 0.80$, $P(B) = 0.36$, $P(C) = 0.28$, $P(A \cap B) = 0.29$, $P(A \cap C) = 0.24$, $P(B \cap C) = 0.16$ and $P(A \cap B \cap C) = 0.13$.

   (a) The Venn diagram is as follows:



(b) $P(A\cap B\cap C^c) = 0.29 – 0.13 = 0.16$
(c) $P(A\cap B^c\cap C^c) = 0.80 – 0.13 – (0.24-0.13) – (0.29-0.13) = 0.4$
(d) We can calculate the following:
$P(A\cap B^c\cap C^c) = 0.4$
$P(A^c\cap B\cap C^c) = 0.36 – (0.29+0.16) + 0.13 = 0.04$
$P(A^c\cap B^c\cap C) = 0.28 – (0.24+0.16) + 0.13 = 0.01$

$P(A^c\cap B^c\cap C^c) = 1 – (0.4+0.04+0.01) – (0.24+0.16+0.29- 2(0.13))$
$= 1- 0.45 – 0.43 = 0.12$
(e) $A^c \cap (B \cup C) = 0.28 – (0.24+0.16) + 0.13 + 0.36 – (0.29+0.16) + 0.13 + 0.16 - 0.13$
$= 0.28 – 0.4 + 0.36 – 0.45 + 0.13 + 0.16 = 0.08$

4. (a) Assume that there are four empty slots that have to be filled with the outcomes of each dice thrown. There will be a total of : $6 *6*6*6 = 6^4$ outcomes.

(b) For the first dice there are 6 possible outcomes, for the second there are (6-1) = 5 possible outcomes, for the third there are 4 possible outcomes, for the final one there are 3 possible outcomes. Therefore, the probability that each face shows a different number is: $6*5*4*3/ 6^4 = 5/18$.

(c) Given condition that the top four faces must sum to 5. The possible sequences are as follows: 2111, 1211, 1121, 1112 i.e., there are 4 of them.
The probability that the top four faces must sum to 5 are : $4/ 6^4 = 1/324$.

(d) The probability that at least one face shows an odd number = 1 – P(N) where event N denotes none of the odd faces appear.
$P(N) = 3 * 3* 3*3/ 6 * 6*6*6 = 1/16$
Probability that at least one face shows an odd number = 1 – (1/16) = 15/16.

(e) There are three odd numbers to choose from which can be done in 3 different ways. There are three even numbers to choose from which can be done in 3 different ways. The number of ways in which three odd and one even can be arranged in four empty slots is : 4!/3! Since the odd number chosen will be the same.
P(three top faces shows the same odd number and one other top face shows an even number) = $(3*3* 4!/3!)/ 6^4$.

5. (a) The total number of possible outcomes are 4^{10}.

P(five gimmels and five hehs) = $(10!/5!*5!)/4^{10}$
The ten obtained outcomes can be ordered in 10! Ways and since 5 of them are the same and order doesn't matter, we divide by 5!.

(b) The number of sequences where there are no nuns or shins are : 2^{10}, since there are 10 dice spins and each spin outcome can have 2 outcomes.
P(no nuns or shins) = $2^{10}/4^{10}$.

(c) Out of the possible two letter combinations, there are 6 ways to choose i.e., (nun,gimmel), (gimmel,heh), (heh,shin), (nun,shin), (nun,heh) and (gimmel,shin).

There are 2^{10} ordered sequences possible with the selected letter combinations. However, we remove the case of any one letter occupying all the positions, since it has been mentioned that two letters are present i.e., 2^{10} -2.

P(two letter are present and two letters are absent) = $(6* (2^{10}-2))/ 4^{10}$.

(d) Given that at least two letters are absent, it means that either two letters can be absent or three letters can be absent.

The case that two letters can be absent : $6*(2^{10}-2)$
The case that three letters can be absent : 4, since there are 4 different ways to choose among the 4 different letter combinations.

P(at least two letters are absent) = $(6*(2^{10}-2) + 4)/4^{10}$

DISCUSSION: This assignment was discussed with Sairam Rakshith Bhyravabhotla (bsairamr).

RESOURCES:

1. An Introduction to Statistical Inference and its Applications with R, Michael W. Trosset.

2. www.rechneronline.de (for graph).