

STAT-S520 Assignment 2

1)

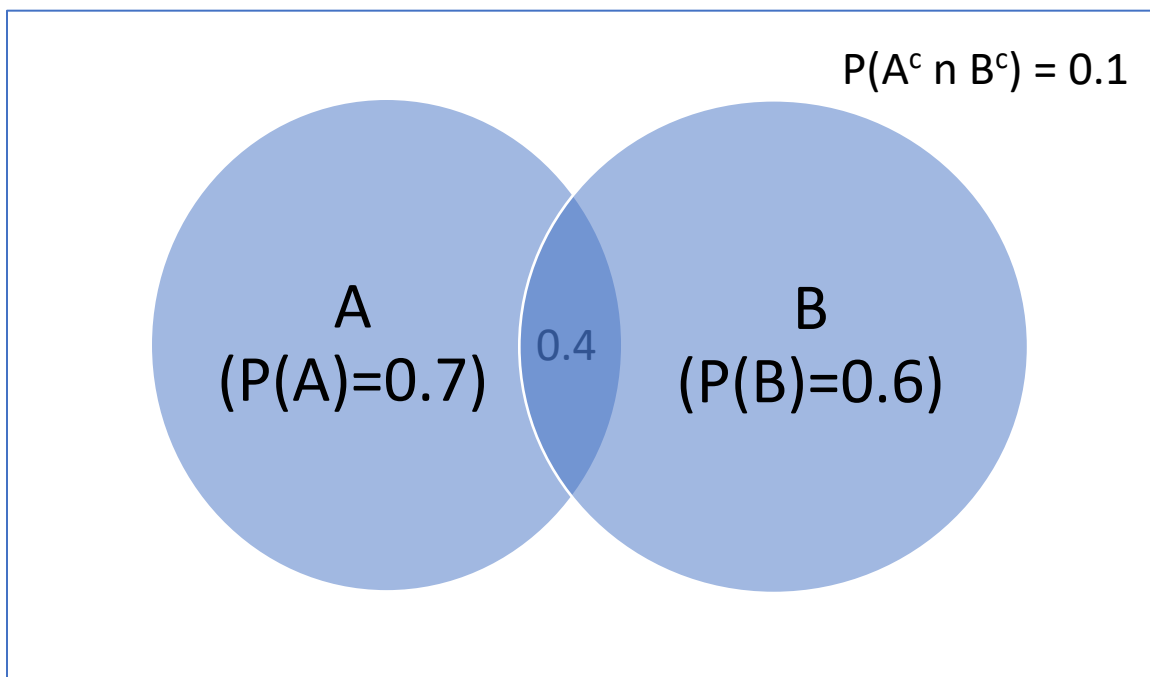
- a) Given, $P(A)=0.7$, $P(B)=0.6$, $P(A^c \cap B)=0.2$.

$$P(A^c \cap B) = P(B) - P(A \cap B), \text{ so } P(A \cap B) = 0.6 - 0.2 = 0.4.$$

$$\text{So, } P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.7 + 0.6 - 0.4 = 0.9$$

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - (0.9) = 0.1.$$

From the above probabilities, we can represent the events A and B in the following Venn-diagram:



- b) No, it's not possible for events A and B to be disjoint. If the events A and B were disjoint, then the probability of observing events A or B, $P(A \cup B) = P(A) + P(B) = 0.7 + 0.6 = 1.3 > 1$.
- c) The probability of $A \cup B^c$ can be written as,
- a. $P(A \cup B^c) = P(A) + P(B^c) - P(A \cap B^c) = 0.7 + 0.4 - 0.3 = 0.8$.
- d) If A and B were independent events $P(A \cap B) = P(A) * P(B)$. But, here, $P(A \cap B) = 0.4 \neq 0.7 * 0.6$. Hence, events A and B cannot be independent.
- e) $P(A/B) = P(A \cap B) / P(B) = 0.4 / 0.6 = 2/3 = 0.667$.

2) In a class of 35 students- 11 students are undergrads and 24 students are graduate students. Of the undergraduates- 4 are female and 7 are male. Of the grad students 5 are female and 19 are male. The given information can be summarized in the table below-

	Male	Female	Total
Undergraduates	7	4	11
Graduates	19	5	24
Total	26	9	35

- a) The probability of selecting a student who is an undergraduate given that the student is a male can be written as follows

$$P(\text{Undergraduate}/\text{Male}) = P(\text{Undergraduate and Male})/P(\text{Male}).$$

$$P(\text{Undergraduate and Male}) = 7/35.$$

$$P(\text{Male}) = (7+19)/35 = 26/35.$$

$$\text{So, } P(\text{Undergraduate}/\text{Male}) = 7/26.$$

- b) The probability of selecting a graduate student in first pick= $24/35$, while the probability of selecting a graduate student in the first attempt and undergraduate student in the second attempt without replacement can be written as -

$$P(\text{First student} = \text{Graduate, Second student} = \text{Undergraduate}) = 24/35 * 11/34.$$

Since we know that the first student is a graduate student, the conditional probability can be written as-

$$P((\text{First student} = \text{Graduate, Second student} = \text{Undergraduate})/\text{First student} = \text{Graduate}) = 11/34.$$

- 3) e) In a population of Hollywood feature films produced during the 20th century, the event A of selecting a movie that was filmed in colored is **dependent** on event B that denotes the movie was western. Since the color and type (western) are two attributes of any movie in the sample, if we know only one of the events between A and B, then the likelihood of observing the unknown event is affected (conditional probability).

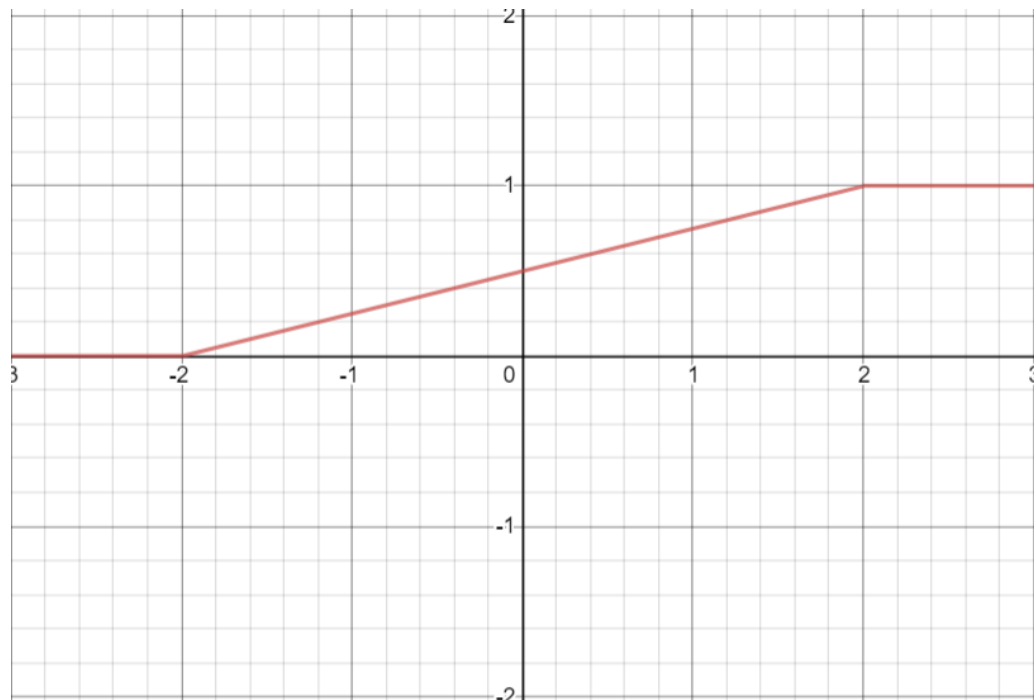
f) In a population of US college freshmen, the event A of selecting a student who attends the College of William & Mary, is **dependent** on event B that the student graduated from a High school in Virginia. Since both the educational institutions are in the same state, there's a high chance that a student graduating from high school attends a college in the same state. So, if we randomly select a student from the population of US college freshmen and know only one of the attributes between events A and B, that's if we know whether he/she graduated from a high

school in Virginia or that he/she attends the College of William & Mary, then the likelihood of observing the unknown event between A and B is affected (conditional probability).

g) From a population of all the persons (living or dead) who have earned a PhD from an University in the United States, if we randomly select a person, then event A that the person's Ph.D. was earned before 1950 and event B that the person is female are **dependent**. For example, we know that the number of women earning a PhD before 1950 is significantly less than the number of women earning PhD since 1950. So, if we know that a person in the sample has earned a PhD before 1950, then there's a high chance that the person is not female. Therefore, there's a clear dependency between events A and B.

4)

a)



Plotted the graph using, "Desmos", which is an online graph plotting tool.

- b) $P(X \leq 1) = F(1) = (1+2)/4 = 3/4$
- c) $P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - 3/4 = 1/4$.
- d) $P(X \geq 1) = 1 - F(1) = 1/4$, same as above, because the random variable X is continuous.
- e) $P(-1.5 < X < 0.5) = P(X < 0.5) - P(X \leq -1.5) = F(0.5) - F(-1.5) = 5/8 - 1/8 = 1/2$.
- f) $P(|X| > 1) = P(X < -1) + P(X > 1) = F(-1) + (1 - F(1)) = 1/4 + (1 - 3/4) = 1/4 + 1/4 = 1/2$.

5) The ten tickets in the urn are labelled as follows:

$\{1, 1, 1, 1, 2, 5, 5, 10, 10, 10\}$.

From the above distribution, we can consider the label on the tickets as a discrete random variable X which can take values from the set $\{1, 2, 5, 10\}$.

The probability mass function of the random variable X is the probability of observing a sample from its distribution. So, for X, the probability mass function (PMF) can be written as-

$$p(X) = 4/10 = 0.4, \text{ if } X=1,$$

$$p(X) = 1/10 = 0.1, \text{ if } X=2,$$

$$p(X) = 2/10 = 0.2, \text{ if } X=5,$$

$$p(X) = 3/10 = 0.3, \text{ if } X=10.$$

$$p(X) = 0, \text{ otherwise}$$

- b) The cumulative distribution function of X, $F(y) = P(X \leq y)$. So, using the above probabilities, we can write the CDF of X as below-

$$F(X) = \begin{cases} 0, & X < 1 \\ 0.4, & 1 \leq X < 2 \\ 0.5, & 2 \leq X < 5 \\ 0.7, & 5 \leq X < 10 \\ 1, & 10 \leq X \end{cases}$$

- c) The expected value of X, $E(X) = \sum X_i p(X_i) = 1(0.4) + 2(0.1) + 5(0.2) + 10(0.3) = 0.4 + 0.2 + 1 + 3 = 4.6$.

- d) The variance of X can be written as $\text{Var}(X) = E[X^2] - (E[X])^2$,

$$X^2 = [1, 4, 25, 100]$$

$$a. E[X^2] = (0.4) * 1 + (0.1) * 4 + (0.2) * 25 + (0.3) * 100 = 0.4 + 0.4 + 5 + 30 = 35.8.$$

$$\text{Var}(X) = 35.8 - (4.6)^2 = 14.64.$$

- e) The standard deviation of X = $(\text{Var}(X))^{0.5} = (14.64)^{0.5} = 3.83$.

- 6) The probability that a receiver correctly guesses the illuminated light at the sender's end is $1/5 = 0.2$, and the probability of getting it wrong is, $(1 - 0.2) = 0.8$.

- a) If X is the random Variable that represents the number of symbols that the receiver guessed correctly in a series of 25 trials, then the expected value of X can be written as-

$$E[X] = {}^{25}C_0 * (0.8)^{25} * (0.2)^0 + {}^{25}C_1 * (0.8)^{24} * (0.2)^1 + {}^{25}C_2 * (0.8)^{23} * (0.2)^2 + \dots + {}^{25}C_{25} * (0.8)^0 * (0.2)^{25}$$

Since the random variable X follows a binomial distribution, the expected value of X can be written as,

$E[X] = n * p$, where n is the number of trials, and p is the probability of guessing the answer right.

$$\text{So, } E[X] = 25 * 0.2 = 5.$$

b) The probability mass function of X can be written as,

- i. $p(X=r) = nCr * p^r * (1-p)^{(n-r)}$, where n = Number of Trials, r = Number of correct guesses, p = probability of guessing the light correctly in a trial.

Since guessing correctly in more than 7 trials is indicative of Extrasensory perception (ESP), the probability of having ESP can be written as, $P(X>7)$.

$$P(X>7) = 1 - P(X \leq 7) = 1 - ({}^{25}C_0 * (0.8)^{25} * (0.2)^0 + {}^{25}C_1 * (0.8)^{24} * (0.2)^1 + {}^{25}C_2 * (0.8)^{23} * (0.2)^2 + \dots + {}^{25}C_7 * (0.8)^{18} * (0.2)^7)$$

$$\text{So, } P(X>7) = 0.109$$

Also Using R,

$$P(X>7) = \text{R-code } (-> 1 - \text{pbinom}(7, 25, 0.2)) = 0.109122.$$

- c) The probability that a receiver has ESP is 0.109, so the probability that at least one person in a group of 20 receivers has ESP can be written as,

$$1 - (0.109)^{20} = 0.9008 \sim 0.901.$$