Dilip Nikhil Francies , 2001274638

1. For each of the following pairs of events, explain why A and B are dependent or independent

e) Consider the population of Hollywood feature films produced during the 20th century. A movie is selected at random from this population. Let A denote the event that the movie was filmed in color and let B denote the event that the movie is a western.

Answer :
P(A) = Probability that the movie was filmed in color
P(B) = Probability that the movie belongs to a western genre.
These two events are dependent events. The number of colored movies that were made in the 20th century increased in the early 1900s due to advancement in visual technology, as a result the western movies that were shot in black and white before 1900s were now shifted to color.
Hence, the probability of choosing a colored movie now has altered the probability of choosing a western genre movie, hence the two distributions are different, the events are dependent.

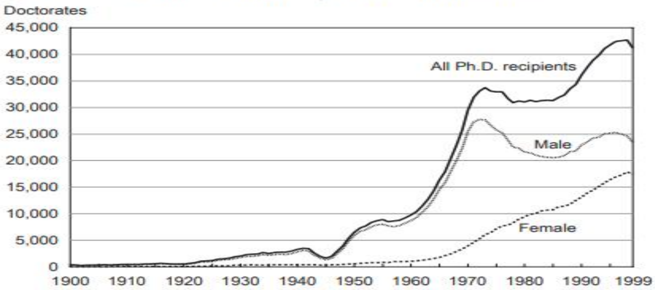
F) Consider the population of U.S. college freshmen, from which a student is randomly selected. Let A denote the event that the student attends the College of William & Mary, and let B denote the event that the student graduated from high school in Virginia.

Answer :
P(A) = student attends the college of William and Mary.
P(B) = student graduated from high school in Virginia.
The two events are dependent.
According to the class profile of College of William and Mary, [link] 60% of the students in their college from a total of [1643] for the class of 2026 are residents of Virginia. Because the distribution is skewed, and it is more than likely that a random person I choose from William and Mary would have graduated from high school in Virginia, hence the two events are dependent.

F) Consider the population of all persons (living or dead) who have earned a Ph.D. from an American university, from which one is randomly selected. Let A denote the event that the person's Ph.D. was earned before 1950 and let B denote the event that the person is female.

Answer :
P(A) = Probability of a person's PhD was earned before 1950.
P(B) = The probability that the chosen person is female.
According to the data and trends seen in the below distribution of total PhD's earned, it is quite straightforward that the number of males earned more PhD than females before 1950. Hence the probability of a PhD chosen before 1950 would more likely be more of a male's PhD than a females one. **Hence these two are dependent events.**

FIGURE 3-2. Doctorates awarded, by sex of recipient: 1900–99



SOURCES: NSF/NIH/USED/NEH/USDA/NASA, Survey of Earned Doctorates and Doctorate Records File (1920–99); U.S. Office of Education annual and biennial reports (1900–19).

2 lightning:

(a) Given that the probability of ~~xxx the~~ ~~citizen~~ a lightning striking a male out of the entire population of citizens who were struck by a lightning, we have the probability,

$$P(A|B) = \underline{\underline{0.85}}.$$

The probability of a lightning striking other than a male is

$$P(A^c/B) = \underline{\underline{0.15}}.$$

2(b) $P(A)$ = Event that a person is a male.
$P(B)$ = probability that a lightning strikes a person.

$P(A|B) = 0.85$
$P(A) = ?$

Because we are considering the entire population of U.S residents, the probability that an event of choosing a male would be close to 0.5

hence, $P(A) = \underline{\underline{0.5}}.$

For Two events to be independent, we know that

$$P(A|B) = P(A)$$
$$0.85 \neq 0.5$$

***2] C**.*
It is a curious fact that approximately 85% of all U.S. residents who are struck by lightning are men. Consider the population of U.S. residents, from which a person is randomly selected. Let A denote the event that the person is male and let B denote the event that the person will be struck by lightning.

(c) Suggest reasons why P(A|B) is so much larger than P(Ac|B). It is tempting to joke that men don't know enough to come in out of the rain! Why might there be some truth to this possibility; i.e., why might men be more reluctant to take precautions than women? Can you suggest other explanations?

Answer : The reason why P(A/B) is larger than that of P(Ac\B) is because, firstly majority of the men usually tend to work outdoors on fields, roads, high altitude stations etc that do come with a certain level of risk. So irrespective of weather conditions, one has to get the work done. Moreover, men who are in the agriculture field tend to take shelter under a tree during heavy rainfall and thunderstorm which increases the probability of getting struck by a lightning even more. Women on the other side, a fair portion of them tend to work indoors, blue collar jobs, teaching etc which is significantly safer than being on the field or repairing a broken piece of tech on a skyscraper.
Even though, precautions are taken to a certain extent to get a particular task done, nobody can expect to get hit by a lightning. Take for instance fishing, when the weather abruptly turns bad, depending on how far a fisher man is in the sea from the shore, he is at risk of getting struck. Considering the fact that 90.3% of workforce in fishing is men and other points mentioned, it is only fair that the value P(A/B) > P(Ac/B).

(3) Probability that a randomly selected woman from this population has breast cancer.

Answer: Contingency table.

	Positive	Negative
Cancer [0.428%]	82.6%	17.4%
No cancer [99.572%]	9.6%	90.4%

3(a) Probability of randomly selected woman has breast cancer & +ve.

$$= P(B \& +ve) = 0.428\% \times 82.6\%$$

$$= \underline{\underline{0.3535\%}} \text{ or } \underline{\underline{0.003535}}$$

3(b). Probability that a random woman tests positive.

$$P(+ve) = P(B \text{ and } +ve) + P(\text{No } B \text{ and } +ve)$$

$$= 0.003535 + \frac{99.572}{100} \times \frac{9.6}{100}$$

$$= 0.003535 + 0.99572 + 0.096$$

$$= \underline{\underline{0.09912}}$$

Scanned with OKEN Scanner

3(c). Probability that woman ~~test~~ ~~positive~~ has cancer given that she tests +ve.

$P(A)$ = Probability that woman has cancer
$P(B)$ = Probability that she tests +ve.

$P(A|B) = ?$

from Bayes theorem we have

$$\Rightarrow P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$

$$= \frac{\frac{0.428}{100} \times \frac{82.6}{100}}{0.09912} = \underline{\underline{0.0356}}$$

3(d) The probability of atleast one with Breast cancer, provided she was tested positive,

$= 1 - (P(\text{no one with Breast cancer})) - (1)$

Now, P(~~no~~ one woman with breast cancer)

$\Rightarrow$ 1 - P(woman has cancer & tests positive).

$= 1 - 0.0356$

Now, for 10 women, the probability is

$= \underline{\underline{(1 - 0.0356)^{10}}} = 0.6959.$

Substituting in the above equation (1)

$1 - 0.6959 = \underline{\underline{0.3041}}$

(4) let X be a random variable with the following cumulative distribution function.

$$F(y) = \begin{cases} 0 & y < 0 \\ y/2 & 0 \le y < 1 \\ \frac{y+1}{4} & 1 \le y < 3 \\ 1 & y \ge 3 \end{cases}$$

4(a) what $P(X \le 2)$?

$P(X \le 2) = F(2) = \frac{2+1}{4} = 3/4$ $\left[\frac{y+1}{4}\right]$

4(b) $P(X > 2) = 1 - P(X \le 2)$

$= 1 - [F(2)]$

$= 1 - 3/4 = 1/4$

4(c) $P(0.5 < X \le 2.5)$?

$= P(X \le 2.5) - P(X \le 0.5)$

$= F(2.5) - F(0.5)$

$= \frac{2.5+1}{4} - \frac{0.5}{2}$

$= \frac{3.5}{4} - 0.25 \Rightarrow 0.625$

4(d) $P(X=1) = ?$

$$P(X=1) = P(X \le 1) - P(X < 1).$$

$$= \cancel{F(1)} \frac{1+1}{4} - \frac{1}{2}.$$

$$= \frac{2}{4} - \frac{1}{2} = 0$$

4(e) Let q be the number such that $F(q) = 0.6?$

$\Rightarrow$ solve the C.D.F for q.

$\Rightarrow 0.6 = \frac{y+1}{4}$. [Substitute q in y]

$$\frac{q+1}{4} = 0.6.$$

$$q+1 = 2.4 \Rightarrow 1.4$$

(5) Weighted die is tossed.

$$F(x) = \begin{cases} \frac{7-x}{20} & x \in \{1,2,3,4,5\} \\ 0 & x \in \{6\}. \end{cases}$$

5 (a) P.M.F of X is the function defined by:

$$f(x) = \begin{cases} \frac{7-x}{20} & x \in \{1,2,3,4,5\} \\ 0 & x \in \{6\} \end{cases}$$

$$f(1) = P(X=1) = \frac{7-1}{20} = \frac{6}{20}$$

$$f(2) = P(X=2) = \frac{7-2}{20} = 5/20$$

$$f(3) = P(X=3) = \frac{7-3}{20} = 4/20$$

$$f(4) = P(X=4) = \frac{7-4}{20} = 3/20$$

$$f(5) = P(X=5) = \frac{7-5}{20} = 2/20$$

$$f(6) = P(X=6) = 0$$

5(b) For Cumulative Distribution function we add values of probability for its respective X values.

~~F(x) = for x~~

$$P(1 \le X < 2) = \frac{6}{20}$$

$$P(2 \le X < 3) = P(X=2) + P(X=1)$$
$$= \frac{5}{20} + \frac{6}{20} = \frac{11}{20}$$

$$P(3 \le X < 4) = \frac{11}{20} + \frac{4}{20} = \frac{15}{20}$$

$$P(4 \le X < 5) = \frac{15}{20} + \frac{3}{20} = \frac{18}{20}$$

$$P(5 \le X < 6) = \frac{18}{20} + \frac{2}{20} = 1$$

$$P(X \ge 6) = 0$$

$$\text{C.D.F} = F(y) = \begin{cases} 6/20 & 1 \le y < 2 \\ 11/20 & 2 \le y < 3 \\ 15/20 & 3 \le y < 4 \\ 18/20 & 4 \le y < 5 \\ 1 & 5 \le y < 6 \\ 0 & y \ge 6 \end{cases}$$

URBAN EDGE COLORFUL YOU

5 ⓒ To find the expected value of x, multiply its probability.

$$Ex = \left(1 \times \frac{6}{20}\right) + 2\left(\frac{5}{20}\right) + 3 \times \frac{4}{20} + 4 \times \frac{3}{20} + 5 \times \frac{2}{20}$$

$$= \frac{6}{20} + \frac{1}{2} + \frac{6}{10} + \frac{6}{10} + \frac{1}{2}$$

$$= \underline{\underline{2.5}}$$

5(d) Variance of X

x	$f(x)$	$x(f(x))$	x^2	$x^2 f(x)$
1	0.3	0.3	1	0.3
2	0.25	0.5	4	1
3	0.2	0.6	9	1.8
4	0.15	0.6	16	2.4
5	0.1	0.5	25	2.5
6	0	0	36	0
		2.5		8

~~Var =~~

$$\text{Var } X = \sum_x (x-\mu)^2 \cdot f(x)$$

$$\Rightarrow (1-2.5)^2 0.3 + (2-2.5)^2 0.25 + (3-2.5)^2 0.2 + (4-2.5)^2 0.15 + (5-2.5)^2 \times 0.1$$

$$\Rightarrow 2.25 \times 0.3 + 0.0625 + 0.05 + 0.3375 + 0.625$$

$$\Rightarrow 0.675 + 0.0625 + 0.05 + 0.3375 + 0.625$$

$$\Rightarrow \underline{\underline{1.75}}.$$

5 (e) Standard deviation $= \sqrt{\text{variance}}$

$= \underline{\underline{1.322}}$.

6. Urn contains 10 tickets:

(a) P.M.F of X:

$P(X=1) = 4/10 = 2/5$ or 0.4.

$P(X=2) = 1/10 = 0.1$

$P(X=5) = 2/10 = 0.2$.

$P(X=10) = 3/10 = 0.3$.

$$f(x) = \begin{cases} 0.4 & x=1 \\ 0.1 & x=2 \\ 0.2 & x=5 \\ 0.3 & x=10 \end{cases}$$

6(b) Cumulative distribution function:

$P(1 \le X < 2) = 0.4$.

$P(2 \le X < 5) = 0.5$

$P(5 \le X \ 10) = 0.7$

$P(X \le 10) = 1$.

$$F(y) = \begin{cases} 0.4 & 1 \le y < 2 \\ 0.5 & 2 \le y < 5 \\ 0.7 & 5 \le y < 10 \\ 1 & 10 \le y \end{cases}$$

6(c) Expected value of X

$EX = 1 \times 0.4 + 2 \times 0.1 + 5 \times 0.2 + 10 \times 0.3$

$= 0.4 + 0.2 + 1 + 3$.

$= \underline{\underline{4.6}}$.

6(d) Variance $= \sum_x (x-\mu)^2 f(x)$

$= (1-4.6)^2 \times 0.4 + (2-4.6)^2 \times 0.1 + (5-4.6)^2 \times 0.2 + (10-4.6)^2 \times 0.3$

$= 5.184 + 0.676 + 0.032 + 8.748$

$\sigma^2 = 14.64.$

6(e) Standard deviation $= \sqrt{\sigma^2}$

$= 3.826$