Q. 1.

(a) Let P be the probability that a randomly chosen male has at least a Bachelor’s Degree.

Therefore, any degree which is equal to a Bachelor’s degree or above it can be considered for this case.

Hence, from the table we can take the sum of Probability of Bachelor’s Degree and Graduate or Professional Degree under male column which is

0.16 + 0.09 = 0.25

Now, we know that P(A) = Number Outcomes / Total number of Possible outcomes, where A is the desired event.

Here, A is that a randomly chosen male has at least a Bachelor’s Degree.

P(A) = 0.25 / 1.00 = 0.25 or 25%

Hence, the probability that a randomly chosen male has at least a Bachelor’s Degree is 0.25 or 25%.

(b) Let P be the probability that a randomly chosen female has at least a Bachelor’s Degree.

Therefore, any degree which is equal to a Bachelor’s degree or above it can be considered for this case.

Hence, from the table we can take the sum of Probability of Bachelor’s Degree and Graduate or Professional Degree under female column which is

0.17 + 0.09 = 0.26

Now, we know that P(A) = Number Outcomes / Total number of Possible outcomes, where A is the desired event.

Here, A is that a randomly chosen female has at least a Bachelor’s Degree.

P(A) = 0.26 / 1.00 = 0.26 or 26%

Hence, the probability that a randomly chosen male has at least a Bachelor’s Degree is 0.26 or 26%.

(c) Since the highest attained degree of the married couple is an independent event, we can directly apply the formula of P(A and B) = P(A) \* P(B). This law can be only applied if A and B are independent event.

Therefore,

P(Male and Female both having at least a Bachelor’s Degree) = P(Male having at least a Bachelor’s Degree) \* P(Female having at least a Bachelor’s Degree)

P(Male and Female both having at least a Bachelor’s Degree) = 0.25 \* 0.26

P(Male and Female both having at least a Bachelor’s Degree) = 0.065

Therefore, the probability that a man and woman getting married, and both have at least a bachelors degree is 0.065 or 6.5%.

(d) I assumed in part (c) where I assumed that both the events are independent. I assumed this because it is very obvious that the highest degree of a male does not affect the highest degree of female and vice versa. Hence, it is reasonable to assume that both the events are independent.

Q. 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1,1 | 1,2 | 1,3 | 1,4 | 1,5 | 1,6 |
| 2,1 | 2,2 | 2,3 | 2,4 | 2,5 | 2,6 |
| 3,1 | 3,2 | 3,3 | 3,4 | 3,5 | 3,6 |
| 4,1 | 4,2 | 4,3 | 4,4 | 4,5 | 4,6 |
| 5,1 | 5,2 | 5,3 | 5,4 | 5,5 | 5,6 |
| 6,1 | 6,2 | 6,3 | 6,4 | 6,5 | 6,6 |

In this question, the sample space of rolling two dices can simple be seen as a 6 x 6 matrix with pair of all possible outcomes. In total, there can be 36 possible outcomes, if two dice are rolled simultaneously. The sample space can be seen in the above table clearly.

(a) Here, the probability of events where the sum is not 5 can take longer time. Hence, we utilize the complement property which states that P(A) = 1 – P(A’). Here P(A) is the probability of event A happening and P(A’) is probability of event A not happening.

In this question, P(A) = P(Sum of the dice is 5) and P(A’) = P(Sum of the dice is not 5)

There are 4 outcomes where the sum of the dice is 5 which are (4,1), (3,2), (2,3), (1,4).

Therefore, P(Sum of the dice is 5) = 4/36 = 1/9 = 0.11 = 11%

Using the complement law,

P(Sum of dice is not 5) = 1 – P(Sum of dice is 5)

P(Sum of dice is not 5) = 1 – 0.11

P(Sum of dice is not 5) = 0.89 or 89%

Hence, the probability that sum of dice is not 5 is 0.89 or 89%.

(b) Here, the sum on of both the rolled dice should be at least 7. So, the possible sum can be 7, 8, 9, 10, 11, and 12.

Let P(A) = P(Sum is at least 7)

Possible outcomes where the sum is at least 7 can be observed in 21 cases which are (6,1), (5,2), (4,3), (3,4), (2,5), (1,6), (6,2), (5,3), (4,4), (3,5), (2,6), (6,3), (5,4), (4,5), (3,6), (6,4), (5,5), (4,6), (6,5), (5,6), (6,6).

So therefore, P(Sum is at least 7) = 21/36 = 7/12 = 0.58 = 58%

Hence, probability that the sum of dice is at least 7 is 0.58 or 58%.

(c) Here, let us assume that P(A) = P(Sum more than 8)

The possible outcomes where sum is greater than 8 can be cases where the sum of the dice is 9, 10, 11, or 12.

So, the possible favorable conditions are (6,3), (5,4), (4,5), (3,6), (6,4), (5,5), (4,6), (6,5), (5,6), (6,6).

Therefore, the number of favorable outcomes are 10.

P(Sum more than 8) = 10/36 = 5/18 = 0.28 = 28%

We know the complement property which states that P(A) = 1 – P(A’). Here P(A) is the probability of event A happening and P(A’) is probability of event A not happening.

Therefore, P(Sum more than 8) = 1 – P(Sum not more than 8)

0.28 = 1 – P(Sum not more than 8)

P(Sum not more than 8) = 1 – 0.28

P(Sum not more than 8) = 0.72 = 72%

Therefore, probability that the sum of the dice will not be more than 8 is 0.72 or 72%.

Q. 3.

(a) First, let me define what is a mutually exclusive event. A mutually exclusive means two events cannot occur simultaneously. For example, if a statement is true, then it cannot be false at the same time. In a similar way, when there is a person with no health coverage, the same person cannot have health coverage at the same time. This explanation clearly supports the fact that being excellent health and have health coverage are mutually exclusive.

(b) Let P(A) = P(randomly chosen individual has excellent health)

From the given data,

P(randomly chosen individual has excellent health) = 0.2329

Therefore, the probability that a randomly chosen individual has excellent health is 0.2329 or 23.29%.

(c) The question asks to find the probability that a randomly chosen individual has excellent health given that he has health insurance. This is a case of classic conditional probability problem.

Let us assume P(A) = P(randomly chosen individual has excellent health) and P(B) = P(randomly chosen individual has healthcare coverage).

P(A) = P(randomly chosen individual has excellent health) = 0.2329

P(B) = P(randomly chosen individual has healthcare coverage) = 0.8738

The formula for conditional probability is P(A|B) = P(A ∩ B) / P(B)

P(A ∩ B) = 0.2099

P(B) = 0.8738

Therefore, P(A|B) = 0.2099/0.8738

P(A|B) = 0.2402 = 24.02%

Therefore, probability that a randomly chosen individual has excellent health given that he has insurance is 0.2402 or 24.02%.

(d) ) The question asks to find the probability that a randomly chosen individual has excellent health given that he does not have health insurance. This is again a case of classic conditional probability problem.

Let us assume P(A) = P(randomly chosen individual has excellent health) and P(B) = P(randomly chosen individual does not have healthcare coverage).

P(A) = P(randomly chosen individual has excellent health) = 0.2329

P(B) = P(randomly chosen individual does not have healthcare coverage) = 0.1262

The formula for conditional probability is P(A|B) = P(A ∩ B) / P(B)

P(A ∩ B) = 0.0230

P(B) = 0.1262

Therefore, P(A|B) = 0.0230/0.1262

P(A|B) = 0.1822 = 18.22%

Therefore, probability that a randomly chosen individual has excellent health given that he does not have insurance is 0.1822 or 18.22%.

(e) Two events are said to be independent if the outcome of one event does not change the outcome of another event. We can further test whether two events are independent or not mathematically by using the formula P(A ∩ B) = P(A) . P(B) where A and B are the two events.

In case of the given problem, let us assume that P(A) = P(having excellent health) and P(B) = P(having healthcare coverage)

From the given table,

P(A) = P(having excellent health) = 0.2329

P(B) = P(having healthcare coverage) = 0.8738

P(A ∩ B) = 0.2099

Now to check the independency, we will use the above formula P(A ∩ B) = P(A) . P(B)

Substituting the above values in the above equation, we get,

0.2099 = 0.2329 . 0.8738

0.2099 ≠ 0.2035

Hence, mathematically it can be established that the events having excellent health and having healthcare coverage are not independent.

Q. 4.

It is given that P(A) = 0.3 and P(B) = 0.7

(a) We already know that P(A and B) = P(A) \* P(B). But this formula is applicable if and only if event A and B are independent. In question (a) it is not specified whether events A and B are independent or not. Hence, we cannot compute P(A and B) even if P(A) and P(B) are given.

(b)

i. Since it is given the events A and B are independent. We can find P(A and B) by the formula P(A and B) = P(A) \* P(B). It is given that P(A) = 0.3 and P(B) = 0.7. So, putting that values in the given equation, we get

P(A and B) = 0.3 \* 0.7.

P(A and B) = 0.21 = 21%

Therefore, P(A and B) is 0.21 or 21%.

ii. We can find the P(A or B) by the formula P(A or B) = P(A) + P(B) – P(A and B). It is given that P(A) = 0.3 and P(B) = 0.7. So, putting that values in the given equation, we get

P(A or B) = 0.3 + 0.7 – 0.21

P(A or B) = 0.79 = 79%

Therefore, P(A or B) is 0.79 or 79%.

iii. We can find P(A|B) by the formula P(A ∩ B) / P(B). Here, P(A ∩ B) is also denoted as P(A and B) which was already calculated in i.. Thus, by this relation, P(A ∩ B) = P(A and B) = 0.21. P(B) is already given as 0.7.

Therefore, by substituting above values in P(A|B) = P(A ∩ B) / P(B), we get

P(A|B) = 0.21 / 0.7

P(A|B) = 0.3 = 30%

Therefore, P(A|B) is 0.3 or 30%.

(c) As we already know, for events A and B to become independent, P(A and B) = P(A ∩ B) = P(A) . P(B)

Hence, if we look at this test, then there can only be one unique value of P(A and B) for which A and B are independent which is 0.21. For any other value, events A and B will not be independent. Hence, in the given condition where P(A and B) is given as 0.1, events A and B will not be independent.

(d) It is given in the question that P(A and B) = 0.1. We already know that P(A|B) = P(A ∩ B) / P(B) and P(B) = 0.7.

Therefore, by substituting the known values in the equation P(A|B) = P(A ∩ B) / P(B), we get

P(A|B) = 0.1 / 0.7

P(A|B) = 0.1428 = 14.28%

Therefore, P(A|B) is 0.1428 or 14.28%.

Q. 5.

(a) Z > -1.13

Finding probability using the R programming,

1 – pnorm(-1.13, mean = 0, sd = 1)

0.8707619

Histogram

Description automatically generated

(b) Z < 0.18

Finding probability using the R programming,

1 – pnorm(0.18, mean = 0, sd = 1)

0.5714237

Chart, histogram

Description automatically generated

(c) Z > 8

Finding probability using the R programming,

1 – pnorm(8, mean = 0, sd = 0)

6.661338e-16

Chart, histogram

Description automatically generated

(d) |Z| < 5

Finding probability using the R programming,

p1 <- pnorm(-0.5, mean = 0, sd = 1)

p2 <- pnorm(0.5, mean = 0, sd = 1)

p2 - p1

0.3829249

Histogram

Description automatically generated

Q. 6.

From the given question, we can see that,

X verbal = 160

X quant = 157

Mean verbal = 151

Sd verbal = 7

Mean quant = 153

Sd quant = 7.67

(a) The shorthand for normal distribution of verbal score is,

N(μ = 151, σ = 7)

The shorthand for normal distribution of quantitative score is,

N(μ = 153, σ = 7.67)

(b) We know that Z = (X - μ)/σ where Z is the Z score, X is the individual value, μ is the mean, and σ is the standard deviation. Substituting the given values in the above equation, we get

For Sofia’s verbal Z-score,

Z verbal = (X verbal - Mean verbal)/ Sd verbal

Z verbal = (160 – 151)/7

Z verbal = 9 / 7

Z verbal = 1.2857

For Sofia’s Quantitative Z-score,

Z quant = (X quant - Mean quant)/ Sd quant

Z quant = (157 – 153)/7.67

Z quant = 4 / 7.67

Z quant = 0.5215

Chart, line chart, histogram

Description automatically generated

(c) The Z-scores give the idea of the distance of the individual score from their mean. In this case the Z quant gives the distance or the number of standard deviation that the individual score is away from the mean and similar explanation can be given for Z verbal.

(d) Since the Z verbal is greater than Z quant, this indicates that Sophia’s score is farther towards the positive direction for Verbal as compared to Quantitative. This means, there are more people who scored less than Sophia’s verbal score as compared to Quantitative score. Hence, we can say that Sophia scored better in Verbal than in Quantitative if compared to others.

(e) We can use the Z-score lookup table to find her percentile. But I chose to use R programming and the pnorm() function that helps to get the percentile of a normal distribution. It takes the Z-score as input and gives the percentile as output.

For percentile of Sophia’s verbal score,

The command that I used is pnorm(Z verbal) = pnorm(1.2857) = 0.9007261

P verbal =0. 9007261 ≈ 90th percentile

Similarly, for percentile of Sophia’s quantitative score,

The command that I used is pnorm(Z quant) = pnorm(0.5215) = 0.6989907

P quant =0. 6989907 ≈ 70th percentile

(f) Percentage of test takes who did better than Sophia in Verbal is,

100 – P verbal = 100 – 90 = 10%

Therefore, 10% test takers did better than Sophia in verbal.

Similarly, Percentage of test takes who did better than Sophia in Quantitative is,

100 – P quant = 100 – 70 = 30%

Therefore, 30% test takers did better than Sophia in quantitative.

(g) Simply comparing the raw scores from the two sections could lead to an incorrect conclusion as to which student did better because the scales used to calculate verbal and quantitative raw scores are different and hence, we cannot just take the raw score. For example, a student may have a very good score and still do worse than most of the class. This can hinder our conclusion, and this is the reason why raw scores are really not a good measure of scoring performances.

(h) Yes, the answer will change. It is because since the distribution will not be normal, it will be some other type of distribution which will have a different curve and thus the area under the curve and the percentiles will also change. Thus, changing the answers to part (e) and (f).