Information Theory: Assignment #1

Due on Novermber 2017 at 3:10pm

Marek Smieja

Szymon Maszke

Problem 1

Given roll consisting of two dices and two events:

1. A even sum of roll
2. B even product of roll

calculate $P(A|B)$. Are those events independent?

Solution

Part A

We calculate probabilities for even sum of roll and even product of roll.

Table 1: Sum of rolls

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

For two undistinguishable dices probablity of their sum being even is equal to $P(A) = \frac{12}{36}$.

Table 2: Product of rolls

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

For two undistinguishable dices probablity of their product being even is equal to $P(B) = \frac{15}{36}$.

Part B

We calculate $P(A|B)$.

Given $P(A|B) = \frac{P(A,B)}{PB}$ and $P(A,B) = \frac{4}{36}$
the conditional probability is: $P(A|B) = \frac{\frac{4}{36}}{\frac{15}{36}}$ which equals $P(A|B) = \frac{4}{15}$

Part C

We test events independence based on condition: $A \perp B = P(A) * P(B)$
Given $P(A) = \frac{12}{36}$ and $P(B) = \frac{15}{36}$ we conclude that $P(A) * P(B) = \frac{15}{108}$.

Based on those results $\frac{15}{108} \neq \frac{4}{36}$ thus A and B are not independent events

Problem 2

Distribution of vector (X, Y) is given by the following table:

X, Y	1	2	3	4
1	0.1	0.2	0.1	0
2	0.3	0	0.2	0.1

Are X, Y independent? Find marginal distributions. Calculate distribution $X|Y = 1$ and it's mean.

Solution

Part A

We calculate marginal distributions based on equation:

$$P(X) = \sum_y P(X, Y) = \sum_y P(X|Y = y)P(Y = y) \tag{1}$$

Marginal distribution for X is therefore given by:

$$P(X = 1) = P(X = 1|Y = 1) + \cdots + P(X = 1|Y = 4) = 0.1 + 0.2 + 0.1 = 0.4 \tag{2}$$

$$P(X = 2) = \cdots = 0.3 + 0.2 + 0.1 = 0.6 \tag{3}$$

Therefore marginal distribution of X is given by vector $[0.4, 0.6]$
Analougous calculations for Y give us marginal distribution described by vector: $[0.4, 0.2, 0.3, 0.1]$

Part B

We test events independence based on condition: $X \perp Y = P(X) * P(Y)$

Given $P(X = 1, Y = 1) = 0.1$ let's calculate $P(X = 1) * P(Y = 1)$ which equals to $0.4 * 0.4 = 0.16$

We can see that $P(X = 1, Y = 1) \neq P(X = 1) * P(Y = 1)$ therefore those events are not linearly independent based on the condtion.

Part C

Distribution of X given Y = 1 is given by vector $[0.1, 0.3]$ and it's mean is $\frac{0.1+0.3}{2} = 0.2$

Problem 3

Defectiveness of certain product is 10%. You sample 3 pieces of forementioned. Derive distribution of random variable describing the number of defective products sampled.

Solution

In the following section defective sample will be indicated by D and non-defective by O

We have four possible outcomes of forementioned sampling:

- None of the products are defective
- One of the products is defective
- Two of the products are defective
- Every product is defective

For each of those events we have to calculate probability of event's occurence, therefore:

- $P(O, O, O) = 0.9 * 0.9 * 0.9 = 0.729$
- Probability of one defective product is given by:
  $P(D, O, O) + P(O, D, O) + P(O, O, D) = 0.1 * 0.9 * 0.9 * 3 = 0.243$
- Probability of two defective products is given by:
  $P(O, D, D) + P(D, O, D) + P(D, D, O) = 0.9 * 0.1 * 0.1 * 3 = 0.027$
- $P(D, D, D) = 0.1 * 0.1 * 0.1 = 0.001$

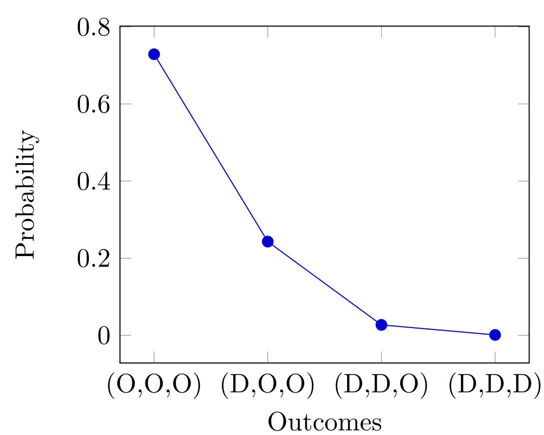
Summing the above:

$$P(O, O, O) + P(D, O, O) * 3 + P(O, D, D) * 3 + P(D, D, D) = 1 \tag{4}$$

proves correctness of the assumption about possible outcomes for this event.

Therefore distribution of product's defectiveness with 3 samples is given by vector:

$$[0.729, 0.243, 0.027, 0.001]$$



Problem 4

Two certain machines are available on the market; 60% produced by company A and 40% by company B. 2% of A's machines are defective and 1% of B's.
Calculate:

- Sampled machine is created by company A and is defective
- Probability of machine's defectiveness
- If machine is defective, what is the probability of it coming from A

Solution

Solutions to each problem provided in the same order:

- $P(A, D) = 0.6 * 0.02 = 0.012$
- We have to sum probabilities of defectiveness from different companies:
  $P(D) = P(A, D) + P(B, D) = 0.012 + 0.4 * 0.01 = 0.016$
- We have to use conditional probability and Bayes Rule, for this example taking the form of:
  $P(A|D) = \frac{P(A)*P(D|A)}{P(A)*P(D|A)+P(B)*P(D|B)}$, so:
  $P(A|D) = \frac{0.6*0.2}{0.6*0.2+0.4*0.1} = 0.75$

Problem 8

We have 100 workes whose seniority follows normal distribution $\mathcal{N}(10, 5)$.
Calculate workers whose seniority is:

- Shorter than 3 years $P(X < 3)$
- Between 3 and 5 years $P(3 < X < 5)$

Solution

First we have to standardize our distribution to obtain easily intepretable z values. For standarization we need the mean of distribution and standard deviation and follow the formula:

$$z = \frac{x - \mu}{\sigma} \tag{5}$$

- for 3: $z = \frac{3-10}{5} = -1.4$
- for 5: $z = \frac{5-10}{5} = -1$

Reading appropriate values from z-table gives us:

- Probability of worker with seniorty less than 3 years: 0.0808
  Next we multiply it by the count of workers and get:

$$w = \lfloor 0.0808 * 100 \rfloor = 8 \tag{6}$$

- Probability of worker with seniorty more than 3 years and less than 5: $0.15866 - 0.0808 = 0.07786$
  Next we multiply it by the count of workers and get:

$$w = \lfloor 0.07786 * 100 \rfloor = 7 \tag{7}$$

Conclusion: Approximately 8 workers work for less than 3 years and approximately 7 workers has seniority between 3 and 5 years.