Exercise 1. Give a real-world example of a joint distribution $P(x, y)$ where x is discrete and y is continuous. Do not use examples involving coins and dice. 1 point.

Answer 1: A real world example of a joint distribution P(x, y) where x is discrete and y is continuous would be the joint probability distribution of a soccer team having x amount of total players and y amount of fantasy score.

Exercise 2. What remains if I marginalize a joint distribution $P(v, w, x, y, z)$ over five variables with respect to variables w and y ? What remains if I marginalize the resulting distribution with respect to v ? 1 point

Answer 2: If a joint distribution P(v, w, x, y, z) is marginalized over five variable with respect to variable w and variable y then the remaining joint distribution would be P(v, x, z). If the resulting distribution is then marginalized with respect to v, the new joint distribution would be P(x, z).

Exercise 3. If variables x and y are independent and variables x and z are independent, does it follow that variables y and z are independent? 1 point

Answer 3: If variable x and y are independent and variable x and z are independent, it does not follow that variable y and z are necessary independent.

Exercise 4. Show that the following relation is true (1 points)

$$P(w, x, y, z) = P(x, y)P(z \mid w, x, y)P(w \mid x, y)$$

Answer 4:

P(w, x, y, z) = P(w, x, y | z) P(z)

Using bayes rules

P(z| w, x, y) = P(w, x, y | z) P(z)/P(w, x, y)

P(z | w, x, y)P(w, x, y) = P(w, x, y | z)P(z)

Hence

P(w, x, y, z) = P(z | w, x, y)P(w, x, y)

Since P(w, x, y) is the same as P(x, y, w)

[Adapted from: Machine Learning I, N. Bertschinger M. Kaschube V. Ramesh]

$P(w, x, y, z) = P(z \mid w, x, y)P(x, y, w)$

$P(w, x, y, z) = P(z \mid w, x, y)P(x, y \mid w)P(w)$

Applying bayes rules again

$P(w \mid x, y) = P(x, y \mid w)P(w) / P(x, y)$

$P(w \mid x, y)P(x, y) = P(x, y \mid w)P(w)$

Hence

$P(w, x, y, z) = P(z \mid w, x, y)P(w \mid x, y)P(x, y)$

Rearrange the term

$P(w, x, y, z) = P(x, y)P(z \mid w, x, y)P(w \mid x, y)$ True

Exercise 5. In my pocket there are two coins. Coin 1 is a fair coin, so the probability $P(h = 1 \mid c = 1)$ of getting heads is 0.5 and the likelihood $P(h = 0 \mid c = 1)$ of getting tails is also 0.5. Coin 2 is biased, so the probability $P(h = 1 \mid c = 2)$ of getting heads is 0.8 and the probability $P(h = 0 \mid c = 2)$ of getting tails is 0.2. I reach into my pocket and draw one of the coins at random. I assume there is an equal chance I might have picked either coin. Then I flip that coin and observe a head.

Think about the Bayesian framework and describe what is the prior, what is the likelihood in this case. 1 point

Use Bayes' rule to compute the posterior probability that I chose coin 2. 3 points

Answer 5:

The prior in this case is the probability of choosing coin 1 or coin 2.

The Likelihood in this case is the probability of getting a head given the chosen coin.

To compute the posterior probability we can use bayes rules.

We have

$P(h = 1 \mid c = 1) = 0.5$

$P(h = 0 \mid c = 1) = 0.5$

P(h = 1 | c = 2) = 0.8

P(h = 0 | c = 2) = 0.2

P(c = 1) = P(c = 2) = 0.5

We are trying to find P(c = 2| h = 1)

Using Bayes Rules

P(c = 2| h = 1) = P(h = 1 | c = 2) P(c = 2)/P(h=1)

We need to find P(h = 1). To do so we can use marginalization.

P(h = 1) = $\sum_c P\ (h = 1, c)$

P(h = 1) = P(h=1, c = 1) + P(h=1, c=2)

Using conditional probability

P(h = 1, c = 1) = P(h = 1 | c = 1)P(c = 1)

P(h = 1, c =2) = P(h = 1 | c = 2)P(c = 2)

Therefore P(h = 1) = P(h = 1 | c = 1)P(c = 1) + P(h = 1 | c = 2)P(c = 2)

Hence P(h = 1) = 0.5 * 0.5 + 0.8*0.5 = 0.25 + 0.4 = 0.65

Therefore

P(c =2 | h = 1) = 0.8*0.5/0.65 = 0.6154

Exercise 6. Consider a biased die where the probabilities of rolling sides {1,2,3,4,5,6} are {1/12,1/12,1/12,1/12,1/6,1/2}, respectively. What is the expected value of the outcome? If I roll the die twice, what is the expected value of the sum of the two rolls? 2 points

Answer 6:

If the dice is rolled once then

E[x] = $\sum_x xP(x)$

E[x] = 1*(1/12) + 2*(1/12) + 3*(1/12) + 4*(1/12)+ 5*(1/6) + 6*(1/2) = 4.67

If the dice is roll twice, the probability of the sum would look a little different.

Values can be from 2 to 12

For 2 it can be 1 + 1 in which P(2) will be (1/12)*(1/12) = 1/144

For 3 it can be 1 + 2 or 2+ 1 which P(3) will be (1/12)(1/12) + (1/12)(1/12) = 1/72

For 4 it can be 2+2 or 1+3 or 3+1 which P(4) will be (1/12)(1/12)+(1/12)(1/12)+(1/12)(1/12) = 1/48

For 5 it can be 1+4 or 4+1 or 2+3 or 3+2 which P(5) will be (1/12)(1/12) + (1/12)(1/12)+(1/12)(1/12)+(1/12)(1/12) = 1/36

For 6 it can be 1+5 or 5+1 or 2+4 or 4+2 or 3+3 which P(6) will be (1/12)(1/6) + (1/6)(1/12)+(1/12)(1/12)+(1/12)(1/12)+(1/12)(1/12) = 7/144

For 7 it can be 2+5 or 5+2 or 1+6 or 6+1 or 3+4 or 4+3 which P(7) will be (1/12)(1/6)+(1/6)(1/12)+(1/12)(1/2) + (1/2)(1/12) + (1/12)(1/12)+(1/12)(1/12) = 1/8

For 8 it can be 2+6 or 6+2 or 3+5 or 5+3 or 4+4 which P(8) will be (1/12)(1/2) + (1/2)(1/12)+(1/12)(1/6)+(1/6)(1/12)+(1/12)(1/12) = 17/144

For 9 it can be 3+6 or 6+3 or 5+4 or 4+5 which P(9) will be (1/12)(1/2) + (1/2)(1/12)+(1/6)(1/12)+(1/12)(1/6) = 1/9

For 10 it can be 4+6 or 6+4 or 5+5 which P(10) will be (1/12)(1/2)+(1/2)(1/12)+(1/6)(1/6) = 1/9

For 11 it can be 5+6 or 6+5 which P(11) will be (1/6)(1/2)+(1/2)(1/6) = 1/6

For 12 it can be 6+6 which P(12) will be (1/2)(1/2) = 1/4

Hence expected values of the sum of the two rolls is

E[y] = $\sum_y yP(y)$

E[y] = 2(1/144) + 3(1/72) + 4(1/48) + 5(1/36) + 6(7/144) + 7(1/8) + 8(17/144) + 9(1/9) + 10(1/9) + 11(1/6) + 12(1/4) = 9.33

[Adapted from: Machine Learning I, N. Bertschinger M. Kaschube V. Ramesh]