

Instructions: Solve the following problems regarding the requested order during your lecture.

1. Use c_i , r_j , and $n_{i,j}$ to compute the following probabilities using their relative frequencies :

1. joint probability $p(X = x_i, Y = y_i)$ using $n_{i,j}$:
2. marginal probability $p(X = x_i)$ using c_i :
3. marginal probability $p(X = x_i)$ using $n_{i,j}$:

Solution:

1. joint probability $p(X = x_i, Y = y_i)$ using $n_{i,j}$: By definition that is $p(X = x_i, Y = y_i) = \frac{n_{i,j}}{N}$
2. marginal probability $p(X = x_i)$ using c_i : By definition that is $p(X = x_i) = \frac{c_i}{N}$
3. marginal probability $p(X = x_i)$ using $n_{i,j}$: Using sum rule we compute the marginal, that is equivalent to express the number of items c_i in collum i by the sum over of all $n_{i,j}$ in that collum
$$(c_i = \sum_j n_{i,j}): p(X = x_i) = \frac{\sum_j n_{i,j}}{N}$$

2. Suppose that we have three colored boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities $p(r) = 0.2$, $p(b) = 0.2$, $p(g) = 0.6$, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting an apple? (Use Bayesian equation for that)

Solution: Express first the conditional probabilities for each box color.

$$\begin{aligned} p(x = apple|y = red) &= \frac{3}{10} \\ p(x = apple|y = green) &= \frac{3}{10} \\ p(x = apple|y = blue) &= \frac{1}{2} \end{aligned}$$

than use Bayesian equation to compute the probability of x given the prior to sample each box.

$$p(x = apple) = \sum_Y p(x = apple, Y) = \sum_Y p(x = apple|Y)p(Y)$$

with prior probabilities of $p = 0.2$ for red and blue, and $p = 0.6$ for green we get:

$$\begin{aligned} p(x = apple) &= p(x = apple|red)p(red) + p(x = apple|green)p(green) + p(x = apple|blue)p(blue) \\ &= \frac{3}{10} * 0.2 + \frac{3}{10} * 0.6 + \frac{1}{2} * 0.2 \\ &= \frac{6}{100} + \frac{18}{100} + \frac{10}{100} = \frac{34}{100} \end{aligned}$$

3. Write down the Bayesian equation. Label the individual parts with posterior, prior, and likelihood:

Solution:

$$p(X|Y) \sim p(X|Y)p(Y)$$

with:

posterior: $p(X)$

likelihood: $p(X|Y)$

prior: $p(Y)$

4. What is the difference between a continuous and discrete random variable.

Solution:

A Discrete Random Variable: A random variable X is called discrete if the range of X is finite or countably infinite.

A Continuous Random Variable: A random variable X is called continuous if the range of X is uncountably infinite.

5. How can we compute the probability that a continuous random variable takes a value in the range between a and b given that you know its probability density

Solution:

$$p(x \in [a, b]) = \int_a^b P(x)dx$$