University of Trieste

International School for Advanced Studies

The Abdus Salam International Centre for Theoretical Physics

Probabilistic Machine Learning

Homeworks 1

Author:
Marco Sciorilli

March 17, 2021

Exercise 1

Let X be a continuous random variable with PDF

$$f(x) = \begin{cases} \frac{5}{32}x^4 & \text{for } 0 < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

and let $Y = X^2$.

1. Find the PDF of Y.

First of all, let's find the PDF of Y by substitution $Y = X^2$, following the rule

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}(g^{-1}(y))\right| \tag{1}$$

As $g^{-1}(y) = \sqrt{y}$ and $\left|\frac{d}{dy}(g^{-1}(y))\right| = \frac{1}{2}y^{-\frac{1}{2}}$, we get

$$f(y) = \begin{cases} \frac{5}{32}y^2 \cdot \frac{1}{2}y^{-\frac{1}{2}} & \text{for } 0 < \sqrt{y} \leq 2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{5}{64}y^{\frac{3}{2}} & \text{for } 0 < y \leq 4 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

2. Find the CDF of Y.

The CDF is defined as

$$F_Y(y) = P(Y \leq y) = \int_{-\infty}^{y} f_Y(t)dt \tag{3}$$

So in our case, as $f(y)$ is equal to 0 outside the interval $0 < y \leq 4$

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \int_{0}^{y} f_Y(t)dt = \int_{0}^{y} \frac{5}{64}t^{\frac{3}{2}}dt = \left[\frac{5}{64}\cdot\frac{2}{5}t^{\frac{5}{2}}\right]\Big|_0^y = \frac{y^{\frac{5}{2}}}{32} \text{ for } 0 < y \leq 4 \tag{4}$$

So F_Y is

$$F_Y(y) = \begin{cases} \frac{1}{32}y^{\frac{5}{2}} & \text{for } 0 < y \leq 4 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

3. Find $\mathbb{E}[Y]$.

$\mathbb{E}[Y]$ for a continuous variable is defined as

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} yf(y)dy = \int_{0}^{4} \frac{5}{64}y^{\frac{5}{2}}dy = \left[\frac{5}{64}\cdot\frac{2}{7}t^{\frac{7}{2}}\right]\Big|_0^4 = 4 \cdot \frac{5}{7} \tag{6}$$

Exercise 2

Suppose that the joint PDF of X and Y is

$$f(x,y) = \begin{cases} \frac{15}{4}x^2 & \text{for } 0 \leq y \leq 1 - x^2 \text{ and } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the marginal PDFs of X and Y.

The marginal PDFS are defines as

$$\begin{aligned} f_X(x) &= \int_{\chi_y} f_{X,Y}(x,y)dy = \int_0^{1-x^2} \frac{15}{4}x^2 dy \\ f_Y(y) &= \int_{\chi_x} f_{X,Y}(x,y)dx = \int_{-1}^{1} \frac{15}{4}x^2 dx \end{aligned}$$

So we get

$$\begin{aligned} f_X(x) &= \frac{15}{4}x^2 \cdot \Big[y\Big]\Big|_0^{1-x^2} == \frac{15}{4}x^2 \cdot (1 - x^2) \\ f_Y(y) &= \Big[\frac{15}{12}x^3\Big]\Big|_{-1}^{1} = \frac{15}{6} \end{aligned}$$

Exercise 3

Let X and Y be continuous random variables with joint PDF

$$f(x,y) = \begin{cases} 6e^{-(2x+3y)} & \text{for } x, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

1. Are X and Y independent?

Lets calculate their joint distribution

$$\begin{aligned} f_X(x) &= \int_{\chi_y} f_{X,Y}(x,y)dy = \int_0^{\infty} 6e^{-(2x+3y)}dy = 6e^{-2x}\Big[-\frac{1}{3}e^{-3y}\Big]\Big|_0^{\infty} = 2e^{-2x} \\ f_Y(y) &= \int_{\chi_x} f_{X,Y}(x,y)dx = \int_0^{\infty} 6e^{-(2x+3y)}dx = 6e^{-3y}\Big[-\frac{1}{2}e^{-2x}\Big]\Big|_0^{\infty} = 3e^{-3y} \end{aligned}$$

And we notice that

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \tag{7}$$

So the two variable are independent.

2. Find $\mathbb{E}[Y|X > 2]$.
As the two variables are independent

$$\mathbb{E}[Y|X > 2] = \mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{0}^{\infty} 3ye^{-3y} dy = \frac{1}{3} \tag{8}$$

3. Find $P(X > Y)$.

By definition, for continuous variables

$$P(A) = P(x \in A) = \int_A P(x) dx \tag{9}$$

In our case $A = (y, \infty)$, so

$$P(x > y) = \int_{y}^{\infty} 2e^{-2x} dx = \Big[-e^{-2x} \Big] \Big|_y^{\infty} = e^{-2y} \tag{10}$$

Exercise 4

Let X and Y be two continuous random variables with joint PDF

$$\begin{cases} x + \frac{3}{2}y^2 & \text{for } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the MAP and the ML estimates of X given $Y = y$.

The log-likelihood expression of the PDF is

$$l(x) = \log(L(x)) = \sum_{i=1}^{n} \log P(Y_i|x) = \log\bigg(x + \frac{3}{2}y^2 \bigg) \tag{11}$$

But in the interval considered, the likelihood doesn't as a maximum and is always increasing. So the ML estimation of x is $x = 1$.

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} = \frac{x + \frac{3}{2}y^2 dy}{\int x + \frac{3}{2}y^2 dx} = \frac{2x + 3y^2}{1 + 3y^2} \tag{12}$$

So, as the MAP estimate is defined as

$$\hat{\theta}_{MAP} = argmax_{\theta} \Big[p(\theta) \prod_i p(x_i|\theta) \Big] \tag{13}$$

In our case, as before, $p(y)p(x|y)$ doesn't have a maximum in $[0, 1]$, and is always increasing, so the MAP estimate is $y = 1$.

Exercise 5

Find the VC-dimension of the set of the hyperplanes in a d-dimensional space.

Hint: consider the problem of binary classification in $\mathbb{R}^d$.

Given m points in a d-dimensional space, they can be shattered by a set of hyperplanes if and only if the position vectors of the remaining points after each division is linearly independent (otherwise the remaining points would lay on the same hyperplane, making the division impossible through an hyperplane, as for example three aligned points on an plane is not separable using lines).
Therefore the VC dimension of the set of hyperplanes in a d-dimensional space is d+1. Picking a point as the origin, we can choose d other points such that their position vectors are linearly independent. However for $m > d+2$, as $d+1$ vectors can't be linearly independent in a d-dimensional space, after the first division we would surely have linearly dependent position vectors, resulting in some points which could not be separated using hyperplanes.