Lecture 15

Andrei Antonenko

Feb 25, 2005

1 Multiple random variables

Assume the situation when we have 2 random variables X and Y. Their distribution can be given by the **joint PDF** $f_{X,Y}(x, y)$ such that

$$P((X, Y) \in B) = \iint_B f_{X,Y}(x, y)\, dx\, dy \tag{1}$$

In the special case,

$$P(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(x, y)\, dy\, dx \tag{2}$$

The normalization property for the PDF in the case of multiple random variables is the following:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1 \tag{3}$$

Now,

$$P(X \in A) = P(x \in A, -\infty < Y < \infty) = \int_A \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy\, dx. \tag{4}$$

Also, we have

$$P(X \in A) = \int_A f_X(x)\, dx. \tag{5}$$

Comparing last two equalities, we get the following formula for **marginal PDF:**

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy. \tag{6}$$

Obviously, similar formula is true for $f_Y(y)$:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx. \tag{7}$$

2 Two-dimensional uniform random variables

In the uniform case, when (X, Y) is distributed in some set S, the joint PDF should be equal to constant:

$$f_{X,Y}(x, y) = \begin{cases} \text{const}, & (x, y) \in S \\ 0, & \text{otherwise.} \end{cases}$$

The constant can be determined from the normalization property:

$$\iint_S \text{const}\ dx\, dx = 1,$$

and therefore

$$\text{const} = \frac{1}{\text{area}(S)}$$

Therefore, the formula for the PDF of the two-dimensional uniform distribution is

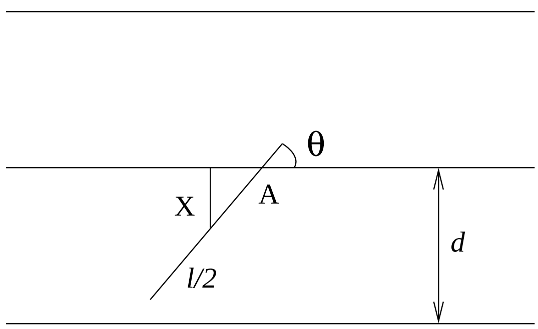
$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\text{area}(S)}, & (x, y) \in S \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

In the uniform case, we can see that the probability of $(X, Y) \in A$ is proportional to the area of A:

$$P((X, Y) \in A) = \frac{\text{area}(A)}{\text{area}(S)}. \tag{9}$$

Example 2.1 (Buffon's needle). Assume there are parallel lines drawn on the plane, such that the distance between two neighboring lines is d. Assume we drop the needle of the length $l < d$ on the plane. What is the probability that the needle will intersect the line?

We will characterize the position of the needle as shown on the next figure by the following parameters: X – the distance between the center of the needle and the closest line, and Θ – the acute angle between the needle and the line:



Let's notice, that X is a random variable uniformly distributed of the interval $[0, d/2]$, and Θ is a random variable, uniformly distributed on the interval $[0, \pi/2]$. The total area for X and *theta* is $(d/2) \cdot (\pi/2) = d\pi/4$, and therefore, the joint PDF of X and Θ is

$$f_{X,\Theta}(x, \theta) = \begin{cases} \frac{4}{\pi d}, & X \in [0, d/2], \Theta \in [0, \pi/2] \\ 0, & \text{otherwise} \end{cases}$$

The needle intersects the line in case the distance between the center of it and the point A is less than or equal to the half length of the needle $l/2$. From the right triangle, we see that the distance between center of the needle and A is equal to $X/\sin\Theta$. Therefore, we need to compute the following probability:

$$P\left(\frac{X}{\sin\Theta} < \frac{l}{2}\right) = P\left(X \le \frac{l\sin\Theta}{2}\right).$$

Using the joint PDF of X and Θ, we get:

$$\begin{aligned}
P\left(X \le \frac{l\sin\Theta}{2}\right) &= \int_0^{\pi/2}\int_0^{(l\sin\theta)/2} \frac{4}{\pi d}\,dx\,d\theta \\
&= \frac{4}{\pi d}\int_0^{\pi/2} \frac{l\sin\theta}{2}\,d\theta \\
&= \frac{2l}{\pi d}\int_0^{\pi/2} \sin\theta\,d\theta \\
&= \frac{2l}{\pi d}\left(-\cos\theta\right)_0^{\pi/2} \\
&= \frac{2l}{\pi d}.
\end{aligned}$$

3 Expectation

In case of two random variables X and Y, we can compute the expectation of any function $g(X,Y)$:

$$\mathbf{E}\left[g(X,Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy. \tag{10}$$

In the special case, when $g(X,Y) = aX + bY$, we have the following formula:

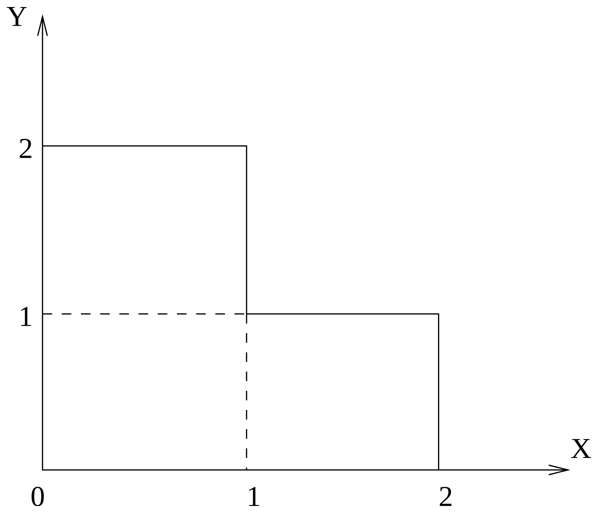
$$\mathbf{E}\left[aX + bY\right] = a\mathbf{E}\left[X\right] + b\mathbf{E}\left[Y\right]. \tag{11}$$

4 Conditioning on another random variable

Assume there are two random variable X and Y, and we know that $Y = y$. This information affects our beliefs about X. The conditional PDF of X given that $Y = y$ is given by the following formula:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}. \tag{12}$$

Example 4.1. Assume X and Y have uniform distribution in the area, shown on the next figure.



Assume $Y = 1.5$. In this case the conditional distribution of X is uniform on the interval from 0 to 1. Now, assume $Y = 0.5$. In this case, the conditional distribution of X is uniform on the interval 0 to 2:

$$f_{X|Y}(x|1.5) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases} \qquad f_{X|Y}(x|0.5) = \begin{cases} 1/2, & 0 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

Example 4.2. Consider the circular uniform distribution, when X and Y are distributed uniformly inside the circle with radius r. In this case the joint PDF is

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi r^2}, & x^2 + y^2 \le r^2 \\ 0, & \text{otherwise} \end{cases}$$

Let's find the marginal PDF of Y:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \\ &= \frac{1}{\pi r^2} \int_{x^2+y^2 \le r^2} dx \\ &= \frac{1}{\pi r^2} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} dx \\ &= \frac{2\sqrt{r^2-y^2}}{\pi r^2}, \quad -r \le Y \le r. \end{aligned}$$

Therefore, we see that the marginal PDF of Y is not uniform.

Now, let's compute the conditional PDF of X.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1}{2\sqrt{r^2-y^2}}, \quad -\sqrt{r^2-y^2} \le x \le \sqrt{r^2-y^2}.$$

Since the conditional PDF of X does not depend on x, we can deduce that as the value of Y is fixed, X is distributed uniformly from $-\sqrt{r^2-y^2}$ to $\sqrt{r^2-y^2}$.

5 Continuous Bayes' rule

From the definition of the conditional PDF we have the following equality:

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x) = f_Y(y) f_{X|Y}(x|y).$$

Expressing $f_{X|Y}(x|y)$, we get:

$$f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{f_Y(y)}. \tag{13}$$

Furthermore, we have:

$$\begin{aligned} f_Y(y) &= \int f_{X,Y}(t,y)\,dt \\ &= \int f_X(t) f_{Y|X}(y|t)\,dt. \end{aligned}$$

Therefore, substituting into the previous formula, we get the **continuous Bayes' rule:**

$$f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{\int f_X(t) f_{Y|X}(y|t)\,dt}. \tag{14}$$

Example 5.1. Suppose the lifetime of the light bulb Y is distributed exponentially with parameter λ. Assume that λ is also a random variable, uniformly distributed on $[0, 1/2]$. We test a light bulb, and observe that the total working time of it is equal to y. What can we now say about the distribution of λ?

Let's use the continuous Bayes' rule. We know:

$$f_\Lambda(\lambda) = \begin{cases} 2, & 0 \le \lambda \le 1/2 \\ 0, & \text{otherwise} \end{cases}$$

Moreover, the conditional distribution of Y given $\Lambda = \lambda$ is exponential:

$$f_{Y|\Lambda}(y|\lambda) = \lambda e^{-\lambda y}.$$

Therefore, using the continuous Bayes' rule, we have:

$$f_{\Lambda|Y}(\lambda|y) = \frac{2\lambda e^{-\lambda y}}{\displaystyle\int_0^{1/2} 2te^{-ty}\,dt}.$$

This formula gives us a new distribution for Λ after we got the information the the light bulb was working for y hours.

Example 5.2. Here we will consider mixed case, when one of the variables is discrete, and another one is continuous. Assume the signal S is transmitted over the phone line. The possible values of the signal are $S = 1$ with probability p, and $S = -1$ with probability $1 - p$. Assume the phone line is noisy, i.e. instead of S, the received signal is $Y = S + N$, where N is a normal random variable, corresponding to noise, with mean equal to 0 and variance σ^2:

$$N \sim N(0, \sigma^2).$$

Assume, the value received is equal to y. What is the probability that the transmitted signal S is equal to 1?

We know the following information:

$$P(S = 1) = p, \qquad P(S = -1) = 1 - p.$$

Moreover, given that the signal transmitted was $S = 1$, we see that $Y = S + N = 1 + N$ is a normal random variable with mean 1 and variance σ^2:

$$Y|\{S = 1\} \sim N(1, \sigma^2),$$

and thus

$$f_{Y|S}(y|1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-1)^2/2\sigma^2}.$$

By the similar reasoning, if the transmitted signal is $S = -1$, we have $Y = S + N = N - 1$, and therefore,

$$Y|\{S = -1\} \sim N(-1, \sigma^2),$$

and thus

$$f_{Y|S}(y|-1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y+1)^2/2\sigma^2}.$$

By Bayes' rule we have:

$$\begin{aligned} P(S = 1|Y = y) &= \frac{p_S(1) f_{Y|S}(y|1)}{p_S(1) f_{Y|S}(y|1) + p_S(-1) f_{Y|S}(y|-1)} \\ &= \frac{p \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-1)^2/2\sigma^2}}{p \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-1)^2/2\sigma^2} + (1-p) \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-(y+1)^2/2\sigma^2}} \end{aligned}$$