ESE 520 Probability and Stochastic Processes

Lecture 5

"One-dimensional random variables and their probability distributions: concrete examples (continued)".

We start with some examples of important *continuous* random variables. (normally, they are referred to be *absolute continuous* random variables)

B1. Normal (Gaussian) probability distribution

For arbitrary numbers $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, we consider the function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, x \in \mathbb{R}. \tag{1}$$

It is clear that $f(x) \geq 0$ for all $x \in \mathbb{R}$. Moreover, it is not hard to see that the integral

$$I = \int_{-\infty}^{\infty} f(x)dx$$

exists and is finite. To find its value, we calculate

$$I^2 = (\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx)^2 =$$

- we make here a change of variables $z = \frac{x-\mu}{\sigma}$ -

$$(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz)^2 =$$

$$(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx)(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy) =$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)} dxdy =$$

- we switch here to polar coordinates using $x = r\cos\theta, y = r\sin\theta$ -

$$\frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} re^{-\frac{1}{2}r^2} drd\theta = 1.$$

It follows then that $I = 1$. In other words, the introduced function $f(x)$ must be a probability density function (pdf) of a continuous distribution on $\mathbb{R}$ which is called *a normal or Gaussian distribution.*

Definition 1. A continuous random variable X is said to have a normal or Gaussian probability distribution with parameters μ and σ^2 if it has the density of the form (1). One uses often then the notation $\mathcal{N}(\mu, \sigma^2)$ to express that fact.

B2. Uniform distribution on the interval (a, b)

Given two real numbers a and b so that $a < b$, we define

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & x \notin [a, b]. \end{cases} \tag{2}$$

Clearly, $f(x) \geq 0$ for all $x \in \mathbb{R}$ and $\int_{\mathbb{R}} f(x)dx = 1$. Therefore, $f(x)$ as defined generates a probability distribution called *uniform probability distribution on* (a, b).

Definition 2. A continuous random variable X is said to have a uniform probability distribution on the interval (a, b) if it has the density of the form (2). A common notation for the uniform distribution on (a, b) is $U(a, b)$.

B3. Gamma probability distribution

For arbitrary numbers $\alpha > 0$ and $\beta > 0$, define

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & x \leq 0, \end{cases} \tag{3}$$

where

$$\Gamma(a) := \int_0^\infty x^{a-1} e^{-x} dx$$

is so-called Gamma function. It is known that $\Gamma(a)$ is defined for all $a > 0$.

Some properties of the Gamma function (which can be verified by direct calculations) are:

$$\Gamma(1) = 1, \Gamma(a+1) = a\Gamma(a), \Gamma(1/2) = \sqrt{\pi}.$$

It can also be shown (you might consider it as an exercise!) that $\int_0^\infty f(x)dx = 1$ so that $f(x)$ defines a continuous probability distribution on $[0, \infty)$.

Definition 3. A continuous random variable X is said to have a *Gamma probability distribution* with parameters α and β if it has the density of the form (3). A usual notation for a Gamma distribution is $\Gamma(\alpha, \beta)$.

One important particular case of a Gamma distribution is when $\alpha = 1$. In this case the Gamma distribution is called an *exponential probability distribution.*

Definition 4. A continuous random variable X is said to have an exponential probability distribution if it has the density of the form

$$f(x) = \begin{cases} \frac{1}{\beta}e^{-x/\beta}, & x > 0 \\ 0, & x \leq 0. \end{cases} \tag{4}$$

The exponential probability distribution will arise later in the course when we will construct what is called a *Poisson process.*

Another important case of a Gamma probability distribution is a so-called χ^2 *probability distribution* arising in many applications, particularly in statistics.

Formally, it can be derived from a normal distribution through a simple transformation.

Assume that X has $\mathcal{N}(0,1)$ distribution and define $Y := X^2$.

We know that

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

How can we find the probability density of Y, the function $f_Y(y)$?

We proceed as follows: for any $y > 0$, we find

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) =$$

$$\int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_0^{\sqrt{y}} \frac{2}{\sqrt{2\pi}} e^{-x^2/2} dx =$$

- we make change of variables $u = x^2, du = 2xdx$ -

$$\int_0^y \frac{1}{\sqrt{2\pi u}} e^{-u/2} du.$$

Since $f_Y(y) = F'_Y(y)$, we obtain

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{-y/2}, & y > 0 \\ 0, & y \leq 0. \end{cases}$$

It is then not hard to see that the function $f_Y(y)$ obtained is nothing but the pdf of the Gamma probability distribution $\Gamma(\alpha = \frac{1}{2}, \beta = 2)$. This distribution is called the χ^2 probability distribution with one degree of freedom and is denoted by $\chi^2(1)$.

More generally, if one chooses so-called independent random variables X_1, X_2,...,X_n each of which has the distribution $\mathcal{N}(0,1)$ and one defines a new random variable

$$Y := X_1^2 + X_2^2 + ... + X_n^2,$$

then it can be shown (we'll do it later) that Y will have $\Gamma(\alpha = \frac{n}{2}, \beta = 2$ distribution called the χ^2 distribution with n degrees of freedom, also denoted as $\chi^2(n)$.

B4. Rayleigh probability distribution

Consider the function

$$f(r) = \begin{cases} re^{-r^2/2}, & r > 0 \\ 0, & r \leq 0. \end{cases}$$

It is clear that $f(r) \geq 0$ and $\int_{\mathbb{R}} f(r)dr = 1$ so that this function determines a continuous probability distribution of $[0, \infty)$ called the *Rayleigh probability distribution*. Any random variable X having the pdf of this form is then said to have a Rayleigh probability distribution.

One way to come up with this distribution is to consider a point in $\mathbb{R}^2$ randomly chosen so that its coordinates (X_1, X_2) are such that X_1 and X_2 follow the distribution $\mathcal{N}(0,1)$ and are independently chosen. If we define by

$$R = \sqrt{X_1^2 + X_2^2},$$

the distance from that point to the origin $(0,0)$, then the random variable R will follow the Rayleigh probability distribution (to be shown later in the course).

Remark 1. After we discussed in detail discrete probability distributions and continuous probability distributions on $\mathbb{R}$, one could think that any particular probability distribution on $\mathbb{R}$ belongs entirely to one of those classes or it is a finite sum of discrete and continuous distributions. It turns out that it is not exactly the case as there exists another class of probability distributions which does not belong to any of the mentioned classes but has the properties of both of them. Such distributions are called *"singular probability distributions"*.

Since any probability distribution on $\mathbb{R}$ is characterized uniquely by its cdf, a singular probability distributions has then a "singular cdf". A classical example of a singular cdf is due to G. Cantor and it is known as "Cantor example". We will not consider singular distributions in this course.

Example. (Cantor's example) You can omit that example as it is not a required part of this course. If interested, you can read it.

We will construct a probability distribution on $\Omega = [0,1]$ which is called a singular distribution. We will do so by constructing a corresponding cdf for that distribution.

Define a function F_1 as

$$F_1(x) = \begin{cases} \frac{1}{2}, & 1/3 < x < 2/3 \\ 0, & x = 0 \\ 1, & x = 1 \end{cases}$$

and we define F_1 on the remaining intervals $(0, 1/3)$ and $(2/3, 1)$ by linear interpolation.

As the result of this construction, F_1 will be a cdf on $[0, 1]$.

Next, using F_1, we construct F_2 in a similar way as

$$F_2(x) = \begin{cases} \frac{1}{2}, & 1/3 < x < 2/3 \\ 1/4, & 1/9 < x < 2/9 \\ 3/4, & 7/9 < x < 8/9 \\ 0, & x = 0 \\ 1, & x = 1 \end{cases}$$

and we use a linear interpolation to define $F_2(x)$ on the remaining intervals.

By continuing in a similar way, we construct a sequence of functions

$$F_1(x), F_2(x), ...$$

where all $F_i(x)$ are continuous and increasing functions.

Thus,

$$\lim_{n\to\infty} F_n(x) = F(x), x \in [0, 1]$$

will exist and is a continuous function that must be a cdf.

The constructed function $F(x)$ is called *the Cantor function.*

Now, to see the properties of $F(x)$, we recall from calculus:

A point x is called a point of increase of $F(x)$ if, for all $\epsilon > 0$, it holds

$$F(x + \epsilon) - F(x) > 0.$$

Let $\mathcal{N} := \{x \in [0, 1] |$ x is a pont of increase of $F\}$.

Then

$$\lambda(\mathcal{N}) = 0,$$

where λ is the Lebesgue measure on $[0, 1]$.

Indeed, from the construction of $F(x)$ it follows that the total length of all intervals from $[0,1]$ where $F(x)$ is constant is

$$1/3 + 2/9 + 4/27 + ... = \frac{1}{3}\sum_{n=0}^{\infty}(\frac{2}{3})^n = 1$$

since the last series is a geometric one.

At the same time, if μ is the corresponding measure for the cdf $F(x)$, then it holds

$$\mu(\mathcal{N}) = 1$$

but

$$\lambda(\mathcal{N}) = 0.$$

In other words, the Lebesgue measure (the total length) of all points of growth of the function $F(x)$ is zero whereas the probability measure μ on $[0,1]$ is concentrated only on the points from $\mathcal{N}$.

Such cdf $F(x)$ and its corresponding probability measure μ are then called singular since they have the properties of both probability distribution classes: the discrete one and the continuous one.