

Continuous probability distributions, Part II  
Math 121 Calculus II  
Spring 2015

Continued from Part I.

Continuous random variables are studied by means of their cumulative distribution functions  $F(x)$  and their probability density functions  $f(s)$ . The probability density function  $f(x)$  is the derivative of the c.d.f.  $F(x)$ . Conversely, the c.d.f. is the integral of the density function.

$$f(x) = F'(x) \quad \text{and} \quad F(x) = \int_{-\infty}^x f(t) dt$$

**Expectation.** First we'll look at expectation of a discrete random variable, then of a continuous random variable.

Suppose that  $X$  is a random variable that only takes on a finite number of values, say  $P(X=3) = \frac{1}{2}$ ,  $P(X=4) = \frac{1}{3}$ , and  $P(X=5) = \frac{1}{6}$ . If you average the values that  $X$  takes on, weighted by the probabilities that  $X$  takes on those values, then you get what is called the *expected value*, *mathematical expectation*, or *mean* of  $X$ , denoted both  $E(X)$  and  $\mu_X$ . When there's only one random variable under discussion,  $\mu_X$  is usually abbreviated to just  $\mu$ .

For the example,  $E(X) = \frac{1}{2} 3 + \frac{1}{3} 4 + \frac{1}{6} 5 = \frac{22}{6} = 3 + \frac{2}{3}$ . It's not that  $X$  ever takes on this value, but  $3 + \frac{2}{3}$  is the expected average you would get if you averaged the outcomes of  $X$  if you could repeatedly perform an experiment with outcome  $X$ .

Now let's consider a continuous random variable  $X$  and come up with a definition of its expectation  $E(X)$ . In some cases it's easy to specify what the expectation should be. For instance, if  $X$  is uniformly distributed on the interval  $[a, b]$ , then an argument based on symmetry suggests  $E(X)$  should be  $\frac{1}{2}(a + b)$ , the middle of the interval  $[a, b]$ . But what if  $X$  is not symmetrically distributed? The exponential distribution has the density function

$$f(x) = \lambda e^{-\lambda x}$$

on  $[0, \infty)$ . What should its expectation be?

We would like to average the values that  $X$  takes on, weighted by the probabilities that  $X$  takes on those values. Over an infinitesimal interval  $[x, x + dx]$ , the value of  $X$  takes on the value (about)  $x$ . The probability that  $X$  lies in  $[x, x + dx]$  is the area under the curve  $f(x)$  over the interval  $[x, x + dx]$ , and since that's almost a rectangular area, that probability is  $f(x) dx$ . Thus, we want to add up  $x f(x) dx$  over all intervals  $[x, x + dx]$ . That leads to the definition of *expectation*  $E(X)$  for a continuous random variable  $X$

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

This expectation is also called the *mean* of  $X$  and denotes  $\mu_X$ .

The value of the integral  $\int_{-\infty}^{\infty} x f(x) dx$  can be interpreted as the  $x$ -coordinate of the center of gravity of the plane region between the  $x$ -axis and the curve  $y = f(x)$ . It is the point on the  $x$ -axis where that region will balance.

(Although in general the center of gravity is the quotient of two integrals

$$\left( \int x f(x) dx \right) / \left( \int f(x) dx \right)$$

for probability densities the denominator is 1.)

**Expectation for the exponential distribution.** When  $X$  is an exponential distribution with parameter  $\lambda$  we find its expectation to be

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} \lambda x e^{-\lambda x} dx.$$

Since the density function of this distribution is 0 for  $x < 0$ , we can restrict the integral to  $[0, \infty)$ . We'll need integration by parts to evaluate this integral. Let  $u = \lambda x$  and  $dv = e^{-\lambda x} dx$  so that  $du = \lambda dx$  and  $v = -\frac{1}{\lambda} e^{-\lambda x}$ . Then

$$\begin{aligned} \int \lambda x e^{-\lambda x} dx &= -x e^{-\lambda x} + \int e^{-\lambda x} dx \\ &= -x e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} \end{aligned}$$

Therefore,

$$E(X) = -x e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty}$$

Evaluating that expression at 0, we get  $-1/\lambda$ . Evaluating that expression as  $x \rightarrow \infty$ , we get 0. Therefore

$$E(X) = \frac{1}{\lambda}.$$

Thus, if the rate of events in the Poisson process is  $\lambda$  events per unit time, then the expected first event occurs at time  $1/\lambda$ .

**The normal distribution.** By far, the most important continuous random variable is the normal random variable (also called a Gaussian random variable). Actually, normal random variables form a whole family that includes the standard normal distribution  $Z$ . This  $Z$  has the probability density function

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

Note that the integral of this function is not an elementary function, that is, can't be expressed in terms of algebraic functions, trig functions, exponents and logs. Instead, its values are found from tables (which you can find in every statistics textbook) or approximated as needed by calculators or computers.

The rest of the normal distributions come from scaling and translating this the standard normal  $Z$ . If  $X = \sigma Z + \mu$ , where  $\sigma$  and  $\mu$  are constants, then  $X$  is called a *normal random*

variable with mean  $\mu$  and standard deviation  $\sigma$ . (The standard deviation of a random variable is a measure of spread of the random variable, second only in importance to the mean of the random variable.)

The density function for a normal distribution  $X \sim \sigma Z + \mu$  is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Note that  $\exp x$  is another notation used for  $e^x$  when the exponent is complicated. The expectation  $E(X)$  of this distribution is  $\mu$ , and its standard deviation is  $\sigma$ .

Normal distributions are very commonly found. They arise from “natural” variation. For instance if you measure the same thing many times you will get different measurements that tend to be normally distributed around the actual value. For an example from nature, different plants (or animals) growing in the same environment vary in their size and other measurable attributes, and these attributes tend to be normally distributed.

**The Central Limit Theorem.** This natural variation can be explained by summing together many small variations due to different factors. The mathematical explanation is found in the Central Limit Theorem or its many generalizations. It says that if you average many independent random variables  $X_1 + X_2 + \cdots + X_n$  all having the same mean  $\mu$  and standard deviation  $\sigma$ , then the average

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

is very close to a normal distribution with mean  $\mu$  and standard deviation  $\sigma/\sqrt{n}$ . More precisely, the limit as  $n \rightarrow \infty$  of  $(\bar{X} - \mu)\sqrt{n}$  is exactly a standard normal distribution.

Generalizations of the Central Limit Theorem weaken the requirements of independence and having the same means and standard deviations. That explains why when there are many factors that influence an attribute, the result is close to a normal distribution.

Many of the techniques of statistics rely on this central limit theorem to estimate from statistical measurements both the means and the standard deviations.

Math 121 Home Page at <http://math.clarku.edu/~ma121/>