Lectures 17 and 18: Continuous Random Variables

Anup Rao

February 14 and 16, 2018

We discuss continuous random variables.

So far, we have been considering random random variables that only take on discrete values. However, many real-world processes involve random variables that take *continuous* values. For example, in Buffon's needle experiment, we counted the number of times that the needle hit a line, which is a discrete random variable. However, we might have wanted to model the angle of the needle after it falls. That would be a continuous random variable. Another example has to do with the Poisson process—we could be interested in the time that the first request comes in, rather than just counting the number of server requests.

There is a basic issue with defining continuous random variables—the definition of what counts as a distribution no longer makes sense. Recall that we said that a distribution $p(x)$ must satisfy:

- For all x, $p(x) \geq 0$.
- $\sum_x p(x) = 1$.

The problem is the second condition. For example, if wanted to define a random variable that takes on a uniformly random value between 0 and 1, what should $p(x)$ be for $0 \leq x \leq 1$? Since the distribution is uniform, $p(x) = p(y)$ for all x, y in the interval. But then if we set $p(x) = \epsilon > 0$, we have

$$\sum_{0 \leq x \leq 1} p(x) = \sum_{0 \leq x \leq 1} \epsilon > 1.$$

So, we cannot set $p(x)$ to be positive to satisfy the definition.

Continuous random variables require a new definition for what a distribution is. Basically, we need to change sums to integrals.

There are two ways to define a distribution on real numbers. The first is using its *probability density function*, or pdf. The pdf is a function $f : \mathbb{R} \to \mathbb{R}$ with

- For all x, $f(x) \geq 0$.
- $\int_{-\infty}^{\infty} f(x) dx = 1$.

Given this description of the distribution, the probability that the random variable takes a value in a set S is just

$$p(X \in S) = \int_S f(x)\, dx.$$

Another way to describe the same distribution is using the *cumulative distribution function* or cdf. The cdf of a variable is a function $F : \mathbb{R} \to \mathbb{R}$ with $F(a) = p(X \le a)$. If the variable has pdf f, we must have

$$F(a) = p(X \le a) = \int_{-\infty}^{a} f(x)\, dx.$$

Sometimes, it is easier to compute the cdf of a distribution than to compute the pdf. If we have computed the cdf F, we can always recover the pdf f by setting $f(x) = \frac{dF(x)}{dx}$ to be its derivative.

Example: The uniform distribution

The uniform distribution on numbers between 0 and 1 has the pdf

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

The cdf is given by

$$F(a) = \int_{-\infty}^{a} f(x)\, dx = \begin{cases} 0 & \text{if } a \le 0, \\ a & \text{if } 0 \le a \le 1, \\ 1 & \text{if } 1 \le a. \end{cases}$$

There are a few important things to note here. First of all, the pdf of a continuous variable can actually take on values *larger* than 1. For example, the pdf of variable that is a uniformly random number in between 0 and $1/2$ is the function that is 2 in this interval, and 0 everywhere else. However, the cdf actually computes a probability, so it is always a number in between 0 and 1.

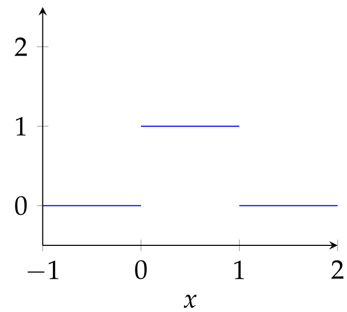
Events, expectation and events

ALL OF THE IDEAS WE HAVE DISCUSSED discrete random variables have analogues for continuous random variables as well. For example, if we have a continuous random variable X with pdf f and cdf g, then the probability of an event E is just

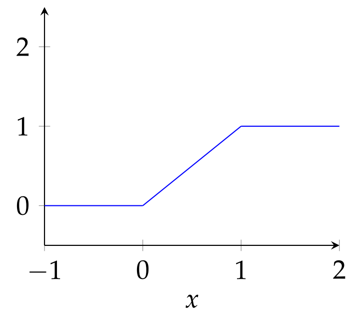
$$p(X \in E) = \int_{E} f(x)\, dx.$$

The expected value of the random variable is

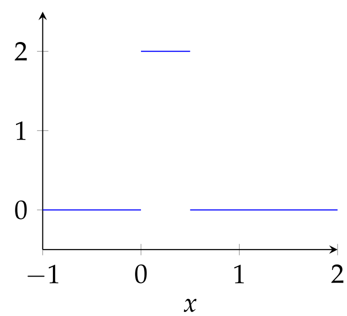
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx.$$



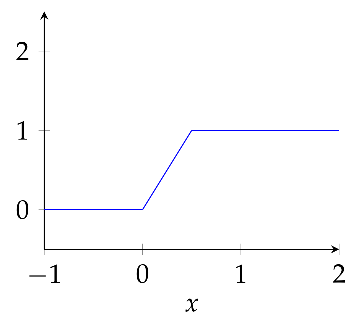
(a) pdf for a uniform number in between 0 and 1.



(b) cdf for a uniform number in between 0 and 1.



(c) pdf for a uniform number in between 0 and 1/2.



(d) cdf for a uniform number in between 0 and 1/2.

Figure 1: The pdf and cdf of a uniformly random number.

The variance is

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2,$$

where here $\mu = \mathbb{E}[X]$ is the expectation of X. The standard deviation is

$$\sigma(X) = \sqrt{\mathrm{Var}[X]}.$$

Markov's inequality and Chebyshev's inequality remain unchanged for continuous random variables:

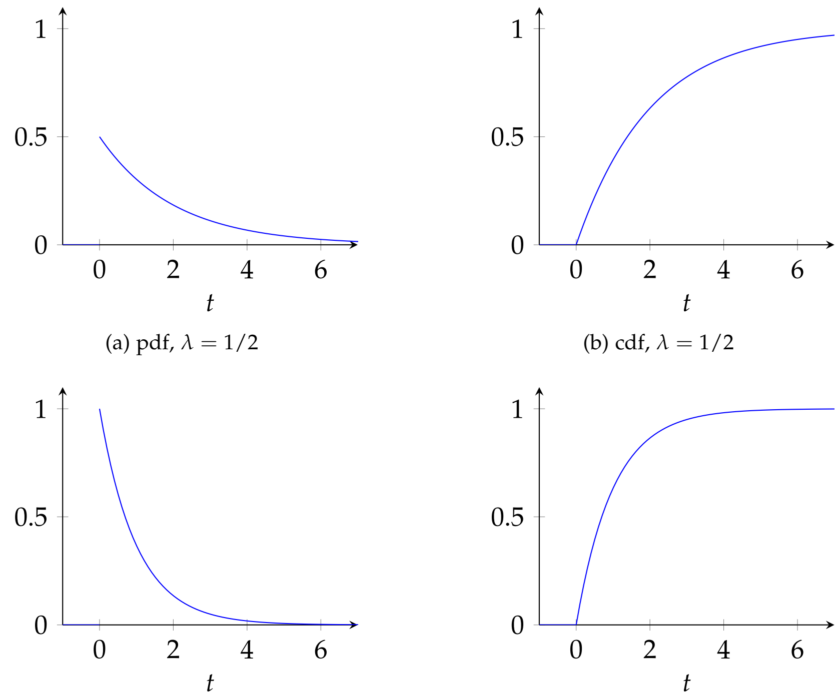
Fact 1 (Markov's inequality)**.** *If X is a non-negative continuous random variable, then*

$$p(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}.$$

Fact 2 (Chebyshev's inequality)**.**

$$p(|X - \mathbb{E}[X]| \geq \alpha) \leq \frac{\mathrm{Var}[X]}{\alpha^2}.$$

Just like for discrete variables, linearity of expectation holds, and $\mathrm{Var}[aX+b] = a^2\mathrm{Var}[X]$. Similarly, if $Y = h(X)$, then $\mathbb{E}[Y] = \int_{-\infty}^{\infty} h(x)f(x)\,dx$.



(a) pdf, $\lambda = 1/2$

(b) cdf, $\lambda = 1/2$

(c) pdf, $\lambda = 1$

(d) cdf, $\lambda = 1$

Figure 2: The pdf and cdf of the arrival time in the Poisson process, which are distributed according to the exponential distribution with parameter λ.

Example: The uniform distribution

Suppose X is uniformly distributed between 0 and 1, with pdf f and cdf g as described above. Then the expected value of X is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{1} dx = (1/2)x\Big|_0^1 = 1/2.$$

The variance is

$$\begin{aligned}\mathrm{Var}[X] &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \\ &= \int_0^1 (x-1/2)^2\,dx \\ &= \frac{1}{3}\cdot(x-1/2)^3\Big|_0^1 \\ &= \frac{(1/8)+(1/8)}{3} = \frac{1}{12}.\end{aligned}$$

The probability that $X \leq 0.8$ is just $g(0.8) = 0.8$.

Example: The exponential distribution

In the last lecture, we defined the Poisson process, which is a probabilistic process that is useful for modeling the arrivals of requests at a server, or the cars going through an intersection at a given time. We talked about a discrete random variable associated with the process, namely the number of arrivals in an interval of time of length τ. We showed that the number of arrivals has the distribution

$$p_\lambda(k) = e^{-\lambda\tau} \cdot \frac{(\lambda\tau)^k}{k!}.$$

Let X denote the *first arrival* in the Poisson process after time 0. What are the pdf and cdf of X? The cdf is easier to compute, so let us start with that. We need to compute $F(t) = p(X \leq t) = p(0 \leq X \leq t) = 1 - p(X > t)$. But we see that $p(X > t)$ is the same as the probability that there are 0 arrivals in the interval $[0, t]$. So we have

The exponential distribution is basically the continuous analogue of the geometric distribution.

$$F(t) = 1 - p(X > t) = 1 - e^{-\lambda t}.$$

To recover the pdf of the first arrival time, we can just differentiate $F(t)$ to get

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

To calculate the expected waiting time, we have

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} t f(t)\, dt \\ &= \int_0^\infty \lambda t e^{-\lambda t}\, dt \\ &= \int_0^\infty (-t) \cdot (-\lambda e^{-\lambda t})\, dt \\ &= -t \cdot e^{-\lambda t}\Big|_0^\infty - \int_0^\infty -e^{-\lambda t}\, dt \\ &= -t \cdot e^{-\lambda t}\Big|_0^\infty - e^{-\lambda t}/\lambda\Big|_0^\infty \\ &= 1/\lambda. \end{aligned}$$

integration by parts

Intuitively, if you expect to see λ arrivals in a unit time interval, you should wait λ time for the first arrival.

The variance turns out to be $1/\lambda^2$.