Lectures 18: Expectation and Variance of Continuous Random Variables

Anup Rao

May 13, 2019

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

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

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

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

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

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 1. We showed that the number of arrivals has the distribution

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

In the last lecture, we showed that this is a valid distribution on the number of arrivals, and gave some intuition for how to think of it as the limit of an infinite sequence of binomial distributions. What is the distribution of the number of arrivals in an interval of length t? It turns out that this is also a Poisson, but now the expectation is λt, so we have

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

as the probability of seeing k arrivals in any interval of length t.

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 $g(t) = p(X \leq t) = p(0 \leq X \leq t) = 1 - p(X > t)$. 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

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

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

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