Math 302.102 Fall 2010
Summary of Lectures from October 20, 2010, through October 25, 2010

Recall that we can think of a random variable as the payout of a bet made on a chance experiment. Formally, a random variable is a function from the sample space S to the real numbers $\mathbb{R}$. If the chance experiment produces the outcome $\omega \in S$, then the value of the random variable is $X(\omega)$. When analyzing random variables, we do not always compute probabilities associated with the outcomes. Instead, we compute probabilities associated with the values of the random variable.

Two particularly important classes of random variables are the following. We say that X is a *continuous random variable* if there exists a non-negative function f such that

$$\mathbf{P}\{X \in B\} = \int_B f(x)\,\mathrm{d}x$$

for any $B \subseteq \mathbb{R}$. We call f the *density function of* X.

We say that X is a *discrete random variable* if X can take on at most countably many values. In this case, we can characterize X by the values

$$\mathbf{P}\{X = k\}$$

for $k = \dots, -2, -1, 0, 1, 2, \dots$. Some people might call the function $p(k) = \mathbf{P}\{X = k\}$ the *probability mass function* of X.

Remark. The important distinction between continuous and discrete random variables is that discrete random variables can assign strictly positive probability to individual points. For example, if X is binomial with $n = 5$ trials and $p = 0.3$ success probability, then

$$\mathbf{P}\{X = 1\} = \binom{5}{1}(0.3)^1(0.7)^4 = 0.36015.$$

However, if X is a continuous random variable, then

$$\mathbf{P}\{X = k\} = \int_k^k f(x)\,\mathrm{d}x = 0$$

for *any* k. In other words, there is no area under a density curve at a single point; the area of a rectangle of width 0 is 0. For example, if X is an exponential random variable with parameter $\lambda = 2$, then

$$\mathbf{P}\{X = 1\} = \int_1^1 2e^{-2x}\,\mathrm{d}x = 0.$$

It is often the case in applications that we are interested in the *average* payout of a bet. If X is a discrete random variable, then the *expected value of* X (also called the *expectation* of X or the *average value* of X) is simply a weighted average of the possible values of X; that is, the possible values of X weighted by their corresponding probabilities. Formally, if X is discrete, then the expected value of X is given by

$$\mathbb{E}(X) = \sum_{k=-\infty}^{\infty} k\mathbf{P}\{X = k\}.$$

However, if X is a continuous random variable, then the expected value of X is given by

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x.$$

Remark. It is perhaps worth noting that the expected value of X exists as a real number provided the sum/integral defining $\mathbb{E}(X)$ converges absolutely; that is, $\mathbb{E}(X)$ is well-defined provided $\mathbb{E}(|X|) < \infty$.

Example. If $X \sim \text{Exp}(\lambda)$ so that $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$, then

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_{0}^{\infty} \lambda x e^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}\int_{0}^{\infty} u e^{-u}\,\mathrm{d}u = \frac{1}{\lambda}$$

(skipping the details of the last step which uses integration-by-parts).

Example. If $X \sim \text{Gamma}(\alpha, \lambda)$ so that

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

for $x \geq 0$, then

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_{0}^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,\mathrm{d}x = \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_{0}^{\infty} x^{\alpha} e^{-\lambda x}\,\mathrm{d}x = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} = \frac{\alpha}{\lambda}$$

(using properties of the Gamma function).

The law of total probability for continuous random variables

Suppose that the continuous random variables X and Y are independent. If the density of X is $f_X(x)$ and the density of Y is $f_Y(y)$, then we can compute $\mathbf{P}\{X > Y\}$ by conditioning on the value of Y and using the continuous version of the law of total probability. That is,

$$\mathbf{P}\{X > Y\} = \int_{-\infty}^{\infty} \mathbf{P}\{X > Y \mid Y = y\} f_Y(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} \mathbf{P}\{X > y\} f_Y(y)\,\mathrm{d}y$$

where

$$\mathbf{P}\{X > y\} = \int_{y}^{\infty} f_X(x)\,\mathrm{d}x.$$

In other words,

$$\mathbf{P}\{X > Y\} = \int_{-\infty}^{\infty}\int_{y}^{\infty} f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Equivalently, we can condition on the value of X in which case we find

$$\begin{aligned}\mathbf{P}\{X > Y\} = \int_{-\infty}^{\infty} \mathbf{P}\{X > Y \mid X = x\} f_X(x)\,\mathrm{d}x &= \int_{-\infty}^{\infty} \mathbf{P}\{Y < x\} f_X(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{x} f_Y(y)\,\mathrm{d}y\right] f_X(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_Y(y) f_X(x)\,\mathrm{d}y\,\mathrm{d}x.\end{aligned}$$