Chris Piech
CS109

Handout #13
April 15th, 2016

Continuous Random Variables

So far, all random variables we have seen have been *discrete*. In all the cases we have seen in CS109 this meant that our RVs could only take on integer values. Now it's time for *continuous* random variables which can take on values in the real number domain. They usually represent measurements with arbitrary precision (eg height, weight, time).

Continuous Random Variables

Probability Density Functions

X is a Continuous Random Variable if there is a Probability Density Function (PDF) $f(x)$ for $-\infty \leq x \leq \infty$ such that:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

The following properties must also hold. These preserve the axiom that $P(a \leq X \leq b)$ is a probability:

$$0 \leq P(a \leq X \leq b) \leq 1$$
$$P(-\infty < X < \infty) = 1$$

A common misconception is to think of $f(x)$ as a probability. It is instead what we call a probability density. It represents probability/unit of X. Generally this is not particularly meaningful without either taking the interval over X or comparing it to another probability density. Of special note, the probability that a continuous random variable takes on a specific value (to infinite precision) is 0.

$$P(X = a) \int_a^a f(x)dx = 0$$

That is pretty different than in the discrete world where we often talked about the probability of a random variable taking on a particular value.

Cumulative Distribution Function

For a continuous random variable X the Cumulative Distribution Function, written $F(a)$ or as (CDF) is:

$$P(X = a) \int_a^a f(x)dx = 0$$

Example 1

Let X be a continuous random variable (CRV) with PDF:

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{when } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

In this function, C is a constant. What value is C? Since we know that the PDF must sum to 1:

$$\int_0^2 C(4x - 2x^2)dx = 1$$
$$C\left(2x^2 - \frac{2x^3}{3}\right)\Bigg|_0^2 = 1$$
$$C\left(\left(8 - \frac{16}{3}\right) - 0\right) = 1$$

And if you solve the equation for C you find that $C = 3/8$.

What is $P(X > 1)$

$$\int_1^{\infty} f(x)dx = \int_1^2 \frac{3}{8}(4x - 2x^2)dx = \frac{3}{8}\left(2x^2 - \frac{2x^3}{3}\right)\bigg|_1^2 = \frac{3}{8}\left[\left(8 - \frac{16}{3}\right) - \left(2 - \frac{2}{3}\right)\right] = \frac{1}{2}$$

Example 2

Let X be a random variable which represents the number of days of use before your disk crashes with PDF:

$$f(x) = \begin{cases} \lambda 3^{x/100} & \text{when } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

First, determine λ. Recall that $\int e^u du = e^u$:

$$\int \lambda 3^{x/100} dx = 1 \Rightarrow -100\lambda \int \frac{-1}{100} 3^{x/100} dx = 1$$

$$-100\lambda^{-x/100}\bigg|_0^{\infty} = 1 \Rightarrow 100\lambda = 1 \Rightarrow \lambda = \frac{1}{100}$$

What is the $P(X < 10)$?

$$F(10) = \int_0^{100} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}\bigg|_0^{10} = -e^{-1/10} + 1 \approx 0.095$$

Expectation and Variance

For continuous RV X:

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

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

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

For both continuous and discrete RVs:

$$E[aX + b] = aE[X] + b$$

$$Var(X) = E[(X - \mu)^2] = E[X^2] = (E[X]))^2$$

Uniform Random Variable

X is a Uniform Random Variable $X \sim Uni(\alpha, \beta)$ if:

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{when } \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases}$$

The key properties of this RV are:

$$P(a \leq x \leq b) = \int_a^b f(x) dx = \frac{b - a}{\beta - \alpha} \text{ (for } \alpha \leq a \leq b \leq \beta)$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx = \frac{x^2}{2(\beta - \alpha)}\bigg|_{\alpha}^{\beta} = \frac{\alpha + \beta}{2}$$

$$Var(X) = \frac{(\beta - \alpha)^2}{12}$$