Chris Piech
CS109

Handout #18
April 20th, 2016

Exponentials and Joint Distributions

Exponential Random Variable

An Exponential Random Variable $X \sim Exp(\lambda)$ represents the time until an event occurs. It is parametrized by $\lambda > 0$, the rate at which the event occurs. This is the same λ as in the Poisson distribution.

Properties

The Probability Density Function (PDF) for an Exponential is:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}$$

The expectation is $E[X] = \frac{1}{\lambda}$ and the variance is $Var(X) = \frac{1}{\lambda^2}$

There is a closed form for the Cumulative distribution function (CDF):

$$F(x) = 1 - e^{-\lambda x} \text{ where } x \geq 0$$

Example 1

Let X be a random variable that represents the number of minutes until a visitor leaves your website. You have calculated that on average a visitor leaves your site after 5 minutes and you decide that an Exponential function is appropriate to model how long until a person leaves your site. What is the $P(X > 10)$?

We can compute $\lambda = \frac{1}{5}$ either using the definition of $E[X]$ or by thinking of how much of a person leaves every minute (one fifth of a person). Thus $X \sim Exp(1/5)$.

$$\begin{aligned} P(X > 10) &= 1 - F(10) \\ &= 1 - (1 - e^{-\lambda 10}) \\ &= e^{-2} \approx 0.1353 \end{aligned}$$

Example 2

Let X be the # hours of use until your laptop dies. On average laptops die after 5000 hours of use. If you use your laptop for 7300 hours during your undergraduate career (assuming usage = 5 hours/day and four years of university), what is the probability that your laptop lasts all four years?

We can compute $\lambda = \frac{1}{5000}$ either using the definition of $E[X]$. Thus $X \sim Exp(1/5000)$.

$$\begin{aligned} P(X > 7300) &= 1 - F(7300) \\ &= 1 - (1 - e^{-7300/5000}) \\ &= e^{-1.46} \approx 0.2322 \end{aligned}$$

Joint Distributions

Often you will work on problems where there are several random variables (often interacting with one another). We are going to start to formally look at how those interactions play out.

For now we will think of joint probabilities with two events X and Y.

Discrete Case

In the discrete case a joint probability mass function tells you the probability of any combination of events $X = a$ and $Y = b$:

$$p_{X,Y}(a,b) = P(X = a, Y = b)$$

This function tells you the probability of all combinations of events (the "," means "and"). If you want to back calculate the probability of an event only for one variable you can calculate a "marginal" from the joint probability mass function:

$$p_X(a) = P(X = a) = \sum_y P_{X,Y}(a,y)$$
$$p_Y(b) = P(Y = b) = \sum_x P_{X,Y}(x,b)$$

In the continuous case a joint probability density function tells you the relative probability of any combination of events $X = a$ and $Y = y$.

In the discrete case, we can define the function $p_{X,Y}$ non-parametrically. Instead of using a formula for p we simply state the probability of each possible outcome.

Continuous Case

Random variables X and Y are Jointly Continuous if there exists a Probability Density Function (PDF) $f_{X,Y}$ such that:

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x,y)dy\, dx$$

Using the PDF we can compute marginal probability densities:

$$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y)dy$$
$$f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b)dx$$

Lemmas

Here are two useful lemmas. Let $F(a,b)$ be the Cumulative Density Function (CDF):

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F(a_2,b_2) - F(a_1,b_2) + F(a_1,b_1) - F(a_2,b_1)$$

And did you know that if Y is a non-negative random variable the following hold (for discrete and continuous random variables respectively):

$$E[Y] = \sum_{i=1}^{n} P(Y \geq i)$$
$$E[Y] = \int_{0}^{\infty} P(Y \geq i)di$$

Example 3

A disk surface is a circle of radius R. A single point imperfection is uniformly distributed on the disk with joint PDF:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi R^2} & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{else} \end{cases}$$

Let D be the distance from the origin: $D = \sqrt{X^2 + Y^2}$. What is $E[D]$? Hint: use the lemmas