Markov - L1 - Review

Class preliminaries.

We will start with a review. I will assume you are familiar with the laws of probability, independence, conditional probability, expectation, etc. We will review discrete and continuous random variables and the relationships between them.

Discrete Random Variables

Suppose we have a random variable X that can take values $x_1, x_2, \ldots$ then its probability mass function (PMF) is

$$p(x) = P(X = x)$$

Recall it should satisfy $\sum_{x_i} p(x_i) = 1$

Let's look at some examples:

Bernoulli Random Variable (or "Bernoulli trials")

Here X can take only two values, say 0 and 1:

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases} \qquad E[X] = p \qquad Var[X] = p(1-p)$$

Geometric

Do independent Bernoulli trials until you get a 1. Let X = # of trials. Then its PMF is

$$p(k) = \underbrace{(1-p)^{k-1}}_{k-1 \text{ } 0\text{'s}} \underbrace{p}_{\text{one } 1} \qquad E[X] = \frac{1}{p} \qquad Var[X] = \frac{1-p}{p^2}$$

Binomial

Do n Bernoulli trials, let X = number of 1's. Then the PMF is

$$p(k) = \underbrace{\binom{n}{k}}_{\text{choose position of } k \text{ ones among } n \text{ trials}} \underbrace{p^k}_{k \text{ ones}} \underbrace{(1-p)^{n-k}}_{n-k \text{ zeros}}$$

$E[X] = np$

$Var[X] = np(1-p)$

Poisson

If $n \to \infty$, $p \to 0$, and $np \to \lambda$ (i.e., many trials, a small probability of success) then the binomial can be approximated by a Poisson distribution with parameter λ:

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad E[X] = \lambda \qquad Var[X] = \lambda$$

Now let's move to continuous distributions.

Continuous Random Variables

Suppose X is a random variable that can take values in $\mathbb{R}$. Then its <u>probability density function</u> (PDF) satisfies

$$P(X \in A) = \int_A f(x)\,dx$$ for any measurable set (think intervals)

Here are some of the most important continuous random variables.

Gaussian

A Gaussian (or "normal") random variable has PDF

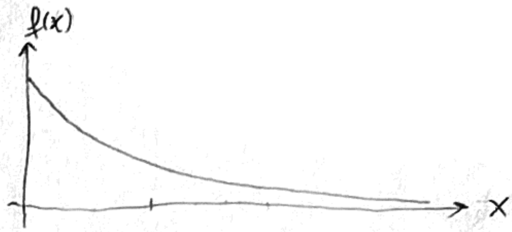
$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad E[X] = \mu \qquad Var[X] = \sigma^2$$

The central limit theorem says that a sum of N independant, identically distributed random variables converges to a Gaussian random variable as $N \to \infty$ (after proper rescaling).

In particular, since a Binomial can be thought as a sum of Bernoulli random variables, a Binomial can be approximated by a Gaussian for large n.

Exponential

An exponential distribution has the PDF



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

$$E[X] = \frac{1}{\lambda}$$

$$Var[X] = \frac{1}{\lambda^2}$$

Let's spend some time with the exponential distribution. When we study Poisson processes later (processes where events occur at "random" times) we'll see the time of the first event arrival is exponentially distributed. Let's find the probability that we have to wait more than x:

$$P(X \geq x) = \int_x^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_x^\infty = e^{-\lambda x}$$

Note that the cumulative distribution function
$$\begin{aligned} F(x) &= P(X \leq x) \\ &= 1 - P(X > x) \\ &= 1 - e^{-\lambda x} \end{aligned}$$

The exponential distribution has the "memoryless property":

$$P(X > T_2 | X > T_1) = P(X > T_2 - T_1) \qquad (T_2 > T_1)$$

To show this, recall conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad \text{so}$$

$$P(X > T_2 \mid X > T_1) = \frac{P(X > T_2 \text{ AND } X > T_1)}{P(X > T_1)} = \frac{P(X > T_2)}{P(X > T_1)}$$

$$= \frac{e^{-\lambda T_2}}{e^{-\lambda T_1}} = e^{-\lambda(T_2 - T_1)} = P(X > T_2 - T_1)$$