812 Section # 5

TA: Sarah Bouchat

3 October 2014

---

1 Probability Mass Functions and Cumulative Distributions

1.1 Definition & Concept Recap

Probability Mass Function: The pmf gives the probability that a random variable from that distribution takes on a given value.

$$P(Y = y) = p(y), \text{ sometimes written as } f(y)$$

The pmf is used to discuss *discrete* distributions. With continuous distributions, the probability that a random variable takes on any given value is 0, so we will use integrals/express this in terms of density. Note that all the values of the pmf must sum to 1.

Cumulative Distribution Functions: The cdf gives the probability that the random variable Y will be $\leq y$:

$$P(Y \leq y) = F(y)$$

The cdf is always a nondecreasing function.

Probability Density Functions: A probability density function (pdf) is a theoretical model for the frequency distribution of a population of measurements.

Remember that $\frac{\partial}{\partial y} CDF = PDF$. That is, $f(y) = \frac{\partial F(y)}{\partial y} = F'(y)$.

Theoretically, the **density** is the probability of being in the width of an interval. As these intervals go to 0 in width, we need to integrate.

Example: Imagine that Acme Enterprise makes machines that produce widgets. Acme wants to sell warranties that are appropriate so it collects data on when their machines break over time. That is, they ask how many of the machines break within 1 year? What about 2 years? 10 years? After many years of collecting data, Acme generates a relative frequency histogram. After a century in business, the relative frequency histogram starts to resemble a distribution, with a smooth curve. This is the probability density function for the life of their widget machines.

1.2 Distributions

These are the distributions we went over this week. What examples can you think of for each?

- Binomial
  - Note that each trial of a binomial is a Bernoulli. That is, Bernoullis are a special case of the binomial; the Binomial is a sum of Bernoulli distributions with success probability p. So, Binomial(2,p) $= Bernoulli(p) + Bernoulli(p)$.
  - Also remember that one of the fundamental characteristics of the Binomial is a fixed number of n identical trials.

- **Example:** I am polling American voters and want to know how many people I have to call before I talk to someone who will vote for a libertarian candidate. Can I model this with a binomial distribution? If so, why? If not, what distribution should I use and why?

- Geometric
- Poisson
  - Remember that Poisson processes are about **counts**.
  - In addition to the horses of the Prussian army, other classic examples of Poisson processes include queuing (arrival in waiting rooms, requests to web servers, calls to offices).
  - What is fundamental to assume about Poisson processes?
- Uniform
- Normal (Gaussian)

1.3 Exercises: PMFs and CDFs

Exercise 1 A cookie bakery sells two types of cookies: chocolate chip and oatmeal raisin. About 50% of customers prefer each type of cookie. Assume that customers may only purchase one cookie at a time (so each cookie sold is "independent" of consumer choices made before or after). Define the random variable X as the number of chocolate chip cookies among the next four cookies sold by the bakery.

a. Find the probability mass function of X
b. Find the cumulative distribution of X

Exercise 2 Consider again the experiment where we roll two regular six-sided dice. Let random variable Y be the sum of the two dice.

a. Find the probability mass function of Y
b. Find the cumulative distribution of Y

Exercise 3 [Walpole book] We flip a fair coin three times. Let W be a random variable defined as the number of heads in the three flips minus the number of tails.

a. Find the probability mass function of W
b. Find the cumulative distribution of W

Many thanks to Emily Sellars for past years' section materials!