STAT 702/J702

September 12th, 2006

-Lecture 6-

Instructor: Brian Habing
Department of Statistics
Telephone: 803-777-3578
E-mail: habing@stat.sc.edu

Today

- Homework Solutions
- Binomial vs. Hypergeometric
- Discrete Random Variables

Ch. 1 #12) In a game of poker, five players are each dealt 5 cards from a 52-card deck. How many ways are there to deal the cards?

Ch. 1 #36) What is the coefficient of x^3y^4 in the expansion of $(x+y)^7$?

Ch. 1 #53) High-Risk, Medium-Risk, and Low-Risk have respective probabilities of 0.02, 0.01, and 0.0025 of filing claims. The proportion in each group are 0.1, 0.2, and 0.7 respectively. What proportion of claim filers were High-Risk?

Ch. 2 #27, but 500 not 100,000) Suppose a rare diseas has an incidence of 1 in 1,000. Assuming that members of the population are affected independently, find the probability of k cases in a population of 500 for k=0, 1, 2.

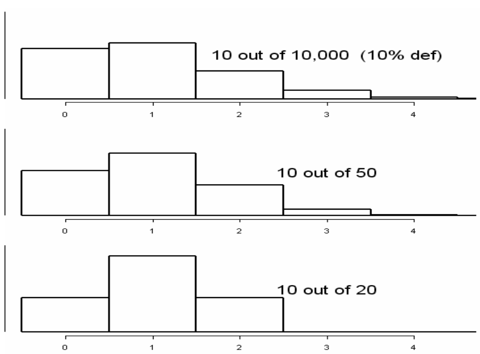
Binomial vs. Hypergeometric

When are the binomial and hypergeometric similar?

For a fixed sample size, when the population size goes to infinity the hypergeometric probability converges to the binomial probability.

What do you lose if you sample with replacement instead? (e.g. why not always use binomial?)

Imagine the case where the population size and sample size are nearly equal?



<u>Chapter 2 – Random Variables</u>

In many of the examples we saw before the event we were interested in was a number.

We converted
HHHHHHHHHTTTTTTTTT
to the number 10 for example.

A random variable is a function that assigns each sample point a numerical value.

Depending on whether the sample space is discrete or continuous the random variable can either be discrete or continuous.

We will begin by considering discrete random variables.

A discrete random variable X is defined by its <u>probability mass function</u> $p(x_i) = P(X=x_i)$

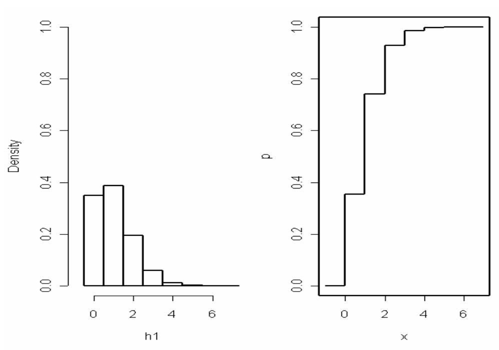
Such that $\Sigma_i \, p(x_i) = 1$

The p.m.f. can be graphed like a histogram.

The cumulative distribution function (cdf) is

$$F(x) = P(X \leq x)$$

This is a step function.



We can also construct numerical summaries of random variables.

The mean or expected value of a discrete random variable X is
$\mu_X = E(X) = \Sigma_i x_i \, p(x_i)$

The variance of a discrete random variable X is
$\sigma_X^2 = Var(X) = \Sigma_i (x_i - \mu_X)^2 \, p(x_i)$

Bernoulli Random Variable

The Bernoulli Random variable is the indicator of an event:

$X = I_A(\omega) =$ 1 if $\omega \in A$
0 otherwise

We can also write the Bernoulli p.m.f.

$p(x) = p^x(1-p)^{1-x}$ if $x = 0$ or 1
0 otherwise

This puts it in the same format regardless of the value of x, and also matches the form of the binomial.

We have already seen the Binomial Distribution

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x \in 0,\ldots n$$

And Hypergeometric Distribution

$$p(x) = \frac{\binom{r}{x}\binom{n-r}{m-x}}{\binom{n}{m}} \text{ for } x \in 0,\ldots \min(m,r)$$