STAT 702/J702

September 9th, 2004

-Lecture 7-

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

Today

- Homework Solutions
- Expected Values in More Detail

Ch. 1 #18a) A lot of *n* items contains *k* defectives, and *m* are selected at random. How should *m* be chosen so that the probability of at least one defective is 0.90?

What is the value of *m* for *n* =1000 and *k* =10?

First note that this is a hypergeometric.

Also note that P[1 or more defectives]=0.9 is the same as P[0 defectives]=0.1.

So we get the formula…

$$P[0 \text{ defectives out of } m] = \frac{\binom{k}{0}\binom{n-k}{m}}{\binom{n}{m}} = 0.10$$

which is very messy. So we could expand it out…

Ch 1 #35a) Prove the following identity both algebraically and by interpreting its meaning combinatorically

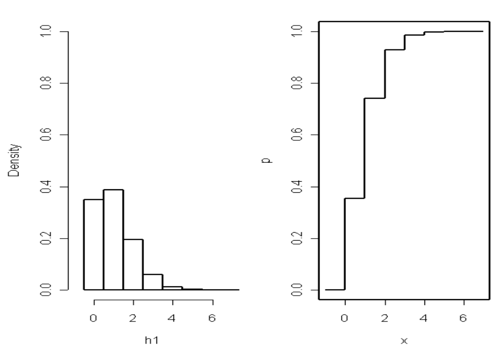
$$\binom{n}{r} = \binom{n}{n-r}$$

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

Chapter 2 – RVs (continued…)

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

The cumulative distribution function (cdf) is $F(x)=P(X\leq x)$



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
$\sigma_X^2 = Var(X) = \Sigma_i (x_i - \mu_X)^2 \, p(x_i)$

We have already seen the <u>Binomial Distribution</u>

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

And <u>Hypergeometric Distribution</u>

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

Notice that calculating the mean and variance of these distributions appears to be very unpleasant!

For example

$$E(X) = \sum_{x=0}^{n} x p(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

Expected Values for Discrete RVs (*from Chapter 4*)

Definition (pg. 111): If X is a discrete RV with p.m.f. p(x) then

$$E(X)= \Sigma_x x\, p(x)$$

when it exists.

Expected value of a function:

Theorem (Pg. 116): If X is a discrete RV with p.m.f. $p_X(x)$ then

$$E(g(X))= \Sigma_x g(x)\, p_X(x)$$

Proof: Let Y be the random variable where for any $\omega \in \Omega$, $Y(\omega)=g(X(\omega))$.

Let A_i be all the x's that correspond to y_i.

Note that this gives $p_Y(y_i)= \Sigma_{x\in A_i} p_X(x_i)$

And by definition

$$E(g(X))=E(Y)= \Sigma_i y_i\, p_Y(y_i)$$

Note that this gives $p_Y(y_i) = \Sigma_{x \in A_i} p_X(x)$
So… $E(g(X)) = E(Y) = \Sigma_i y_i\, p_Y(y_i)$

$= \Sigma_i y_i \{\Sigma_{x \in A_i} p_X(x)\}$

$= \Sigma_i \Sigma_{x \in A_i} y_i\, p_X(x)$

$= \Sigma_i \Sigma_{x \in A_i} g(x)\, p_X(x)$

$= \Sigma_x g(x)\, p_X(x)$ ∎

STAT 702/J702 B.Habing Univ. of S.C. 16

One place this is used is to get the formula for variance:

$\mathrm{Var}(X) \equiv E[(X-\mu_X)^2]$

$= \Sigma_i (x_i - \mu_X)^2\, p(x_i)$

STAT 702/J702 B.Habing Univ. of S.C. 17

The theorem also allows us to prove two results about a linear function of a random variable:

$g(X) = a + bX$

The constant a represents a shift and the multiplier b represents a change of scale.

STAT 702/J702 B.Habing Univ. of S.C. 18

$E(a + b X) = \Sigma_x (a+bx) p_X(x)$

$= \Sigma_x\{a\, p_X(x) + bx\, p_X(x)\}$

$= \Sigma_x a\, p_X(x) + \Sigma_x bx\, p_X(x)$

$= a\, \Sigma_x p_X(x) + b\, \Sigma_x x\, p_X(x)$

$= a + b\, E(x)$

STAT 702/J702 B.Habing Univ. of S.C. 19

$Var(a + b X) = E[((a + b X) - \mu_{a+bX})^2]$

$= E[(a + b X - (a + b\, \mu_X))^2]$

$= E[(b X - b\, \mu_X)^2]$

$= E[b^2 (X - \mu_X)^2]$

$= b^2 E[(X - \mu_X)^2]$

$= b^2 Var(X)$

STAT 702/J702 B.Habing Univ. of S.C. 20

Neither of these seem to help us with the finding the expected value and variance of the binomial though.

What could help us there is something that let us find the expectation and variance of a sum of independent random variables.

STAT 702/J702 B.Habing Univ. of S.C. 21

Theorem: (*special case of A on 119 and A on 131*)

Let $X_1, X_2, \ldots X_n$ be mutually independent random variables, then:

$\mu_{\Sigma X} = E(\Sigma_i X_i) = \Sigma_i E(X_i) = \Sigma \mu_X$

$\sigma_{\Sigma X}^2 = Var(\Sigma_i X_i) = \Sigma_i Var(X_i) = \Sigma\, \sigma_X^2$

Sketch of Proof: Consider the case of two random variables X and Y with p.m.f.s $p_X(x)$ and $p_Y(y)$ respectively.

$E(X+Y) = \Sigma_{x,y}(x+y)P(X=x, Y=y)$

$= \Sigma_{x,y}(x+y)P(X=x)P(Y=y)$

$= \Sigma_{x,y}(x+y)P(X=x)P(Y=y)$

$= \Sigma_{x,y}\, x\, P(X=x)P(Y=y) + \Sigma_{x,y}\, y\, P(X=x)P(Y=y)$

$= \Sigma_x\Sigma_y\, x\, P(X=x)P(Y=y) + \Sigma_x\Sigma_y\, y\, P(X=x)P(Y=y)$

$= \Sigma_x\Sigma_y\, x\, P(X=x)\, P(Y=y) + \Sigma_y\Sigma_x\, y\, P(Y=y)\, P(X=x)$

$= \Sigma_x \Sigma_y\, x\; P(X=x)\, P(Y=y)$
$+ \Sigma_y \Sigma_x\, y\; P(Y=y)\, P(X=x)$

$= \Sigma_x \{x\; P(X=x)\, \Sigma_y P(Y=y)\}$
$+ \Sigma_y \{y\, P(Y=y)\; \Sigma_x P(X=x)\}$

$= \Sigma_x\, x\; P(X=x) + \Sigma_y\, y\, P(Y=y)$

$= E(X) + E(Y)$

For a binomial X with sample size n and probability p,

$E(X) = np$

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