Lecture 8: Discrete Expectation

Probability Theory and Applications

Fall 2005

September 27

Quiz Tuesday 10/4

Covers material covered through today with emphasis on materials covered in class, homeworks, book up through section 3.3.

Nothing on continuous random variables.

Can bring to class

- one page of notes
- simple calculator (no symbolic math)
- handout with distribution and tables

Problem

Suppose that a certain society values sons more than daughters. In this society, a couple will continue bearing children until they produce a son, at which point they will retire from the child-bearing business.

Would this family-planning scheme tend to produce more boys, or more girls?

Simulation

Let Y = size of famliy

Given Y, what is the number of girls per family.

What is number of boys per family?

Use coins to simulate 20 families, and calculate the number of girls per family.

Expectation of *Y*

Let *Y* be a discrete R.V with pdf *f(y),* then the expected value of *Y*, denoted *E(Y)* is

$$E(Y) = \sum_{y \in value\ set} y \cdot f(y)$$

Back to family planning

Family size follows a geometric distribution

Y = size of family

p = 0.5 probability of having a boy

$$f(y) = (1 - p)^{y-1} p \qquad y = 1, 2, 3, \ldots$$

$$E(Y) = \sum_{y=1}^{\infty} y f(y) = \sum_{y=1}^{\infty} y(1-p)^{y-1} p$$

$$= p \sum_{y=1}^{\infty} y(1-p)^{y-1}$$

$$= p \sum_{i=0}^{\infty} (i+1)(1-p)^{i}$$

$$= p \frac{1}{\left(1-\left(1-p\right)\right)^2} = \frac{p}{p^2} = \frac{1}{p}$$

Will there be more boys or girls

- Expected size of family

  $E(Y)=1/p=2$
- Expected number of girls per family

  $E(Y-1)=2-1=1$
- Expected number of boys per family

  1

On average equal number of boys and girls.

New family strategy

A couple will continue bearing children until they produce two daughters, at which point they will retire from the child-bearing business.

What is expected family size?

W=size of family

W~negative binomial(r=2,p=0.5)

Negative Binomial

Generalization of Geometric

X=# of Bernoulli trials to get r^th success

$$Given,\ p(success) = p$$

$$\underbrace{f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}\, p}_{Binomial(n=x-1,\,p)} \qquad x = r, r+1, \ldots$$

$$E(X) = \frac{r}{p}$$

New family strategy

A couple will continue bearing children until they produce two daughters, at which point they will retire from the child-bearing business.

What is expected family size?

$E(W)=r/p=2*2=4$

$E(Girls)=2$

$E(Boys)=E(W-2)=2$

Still about even number of boys and girls.

Example

- Car insurance pays $1000 if you get in an accident . Probability of getting in an accident is .005.

$Y = \text{amt. paid by insurace company}$

y	\$0	\$1000
$f(y)$	.995	.005

E(Y)=0(.995)+1000(.005)=$5

How much can you charge for insurance and make money? How risky is it?

Basic Properties of Expectation

a) Let *g(X)* be any function of R.V. *X*

$$E(g(X)) = \sum_{x_i \in value\ set} g(x_i) f(x_i)$$

Basic Properties of Expectation

b) If c is a constant,   E(c)=c

proof

$$E(c) = \sum_i cf(x_i) = c\sum_i f(x_i) = c*1 = C.$$

Basic Properties of Expectation

c) For constants *a* and *b*,

$$E(aX+b)=\sum_i (ax_i+b)f(x_i)$$

$$=a\sum_i x_i f(x_i)+b\sum_i f(x_i)$$

$$=a\overline{x}+b \qquad \overline{x}=E(X)$$

Theorem

Let $X_1, X_2, \ldots, X_n$ be random variables and

Let $E(X_i) = \overline{x}_i$.

Then

$$E(X_1 + X_2 + \ldots + X_n) = \overline{x}_1 + \overline{x}_2 + \ldots + \overline{x}_n$$

Variance of *Y*

Let *Y* be a discrete R.V with pdf *f(y),* then the variation of *Y*, denoted *E(Y)* is

$$\mathrm{var}(Y) = E[(Y - E(Y))^2]$$

$$= \sum_{y \in value\ set} \left(y - E(Y)\right)^2 f(y)$$

The standard deviation of Y is the square root of the variance.

Variance

- How much does the amount of payout vary year to year?

  variance = var(Y)=E[(Y-E(Y))²]

$[y-E(Y)]^2$	25	995^2
$f(y)$	$.995$	$.005$

$$\text{var}(Y) = 25 * .995 + 995^2 * .005 = 4975$$

$$\text{standard deviation } \sqrt{\text{var}(Y)} = 70.53$$

Insurance Example

Say I sell 100 insurance policies. Assume that accidents are independent. Let Y_i be the payout of the ith policy.

Find the variance of the payouts of the 100 policies?

Theorem

Let $X_1, X_2, \ldots, X_n$ be *independent* random variables and

Let $\mathrm{var}(X_i) = \sigma_i^2.$

Then

$$var(X_1 + X_2 + \ldots + X_n) = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2$$

$$\mathrm{var}(\sum_{i=1}^{100} Y_i) = 100 * 4975 = 497500$$

$$\text{standard deviation } = \sqrt{49750} = 705.34$$

Basic Properties of Variance

a) If c is a constant, var(c)=0.

proof

$$\mathrm{var}(c) = \sum_i (c - E(c))^2 f(x_i)$$

$$= \sum_i (c - c)^2 f(x_i) = 0$$

Basic Properties of Variance

b) For constants *a* and *b*,

$$\mathrm{var}(aX+b)=\sum_i [(ax_i+b)-E(aX+b)]^2 f(x_i)$$

$$=\sum_i [(ax_i+b-a\overline{x}-b]^2 f(x_i)$$

$$=a^2\sum_i [(x_i-\overline{x})]^2 f(x_i)$$

$$=a^2\,\mathrm{var}(X)$$

Basic Properties of Variance

C) Practical hint: $\mathrm{var}(X) = E(X^2) - E(X)^2$

$$\mathrm{var}(X) = \sum_i [(x_i - E(X)]^2 f(x_i)$$

$$= \sum_i [(x_i^2 - 2x_i E(X) + E(X)^2]^2 f(x_i)$$

$$= \sum_i (x_i^2) f(x_i) - 2E(X) \sum_i (x_i) f(x_i) + E(X)^2 \sum_i f(x_i)$$

$$= E(X^2) - 2E(X)E(X) + E(X)^2$$

$$= E(X^2) - E(X)^2$$