Ch. 7 Properties of Expectations

Recall the definitions of expected value of a random variable X ...

- X is discrete:

$$E[X] = \sum_x x\, p(x)$$

where $p(x)$ is the probability mass function of X

also

$$E[g(X)] = \sum_x g(x) p(x)$$

expected value of a function of X.

(see proposition 4.1, p. 122 textbook)

- X is continuous:

$$E[X] = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

where $f(x)$ is the probability density function of X

also

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x)\, dx$$

expected value of a function of X

(see Proposition 2.1, p. 181 textbook)

What about jointly distributed random variables X, Y?

7.2 Expectation of Sums of Random Variables

Here we have the following results generalizing the previous ones ~~be~~ just listed.

Proposition 2.1

discrete case: If X and Y have a joint pmf $p(x,y)$ then

$$E\left[g(X,Y)\right] = \sum_{y}\sum_{x} g(x,y)\,p(x,y)$$

continuous case: If X and Y have a joint pdf $f(x,y)$ then

$$E\left[g(X,Y)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,f(x,y)\,dx\,dy$$

(Proof, see textbook, p. 281)

EXAMPLE (important special case)

Suppose X and Y are jointly continuous with joint pdf $f(x,y)$.

What is $E[X+Y]$? (i.e. $g(X,Y) = X+Y$)

$$E[X+Y] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x+y)\,f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,f(x,y)\,dx\,dy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\,f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{+\infty} x \underbrace{\left[\int_{-\infty}^{+\infty} f(x,y)\,dy\right]}_{= f_X(x) \text{ marginal pdf of } X} dx + \int_{-\infty}^{+\infty} y \underbrace{\left[\int_{-\infty}^{+\infty} f(x,y)\,dx\right]}_{= f_Y(y) \text{ marginal pdf of } Y} dy$$

So

$$E[X+Y] = \int_{-\infty}^{+\infty} x f_X(x)\,dx + \int_{-\infty}^{+\infty} y f_Y(y)\,dy$$

$$= E[X] + E[Y]$$

(assuming $E[X]$ and $E[Y]$ are finite) (see p. 282 textbook)

see also notes p. 145 Corollary 9.2

More generally..., for random variables $X_1, X_2, \ldots, X_n$

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n] \quad (*)$$

extend previous argument using induction.

This result (*) holds for ~~[illegible]~~ jointly continuous $X_1, \ldots, X_n$ as well as for jointly discrete $X_1, \ldots X_n$.

EXAMPLE (2c, p. 283 textbook)

Suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables each having ~~[illegible]~~ cumulative distribution function, F, and expected value $E[X_i] = \mu$ for $i = 1, 2, \ldots, n$.

Define the sample mean as

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

(i.e. $\bar{X}$ is a function of random variables $X_1, X_2, \ldots, X_n$)

Compute $E[\bar{X}]$.

$$E[\bar{X}] = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right)$$

} by our result on previous page

$$= \frac{1}{n}\sum_{i=1}^{n} E[X_i]$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu$$

$$= \frac{1}{n}(n \cdot \mu) = \mu$$

The expected value of the sample mean is the mean of the distribution.

EXAMPLE (Expectation of Binomial Random Variable)

Let X be a binomial R.V. with parameters n and p

recall, this means

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

$$= P\{X = i\}$$

see section 4.6 p.127

i = # of successes
p = probability of a success
n = total # of trials

Let

$$X = X_1 + X_2 + \ldots + X_n$$

where

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ trial is a success} \\ 0 & \text{if } i^{th} \text{ trial is a failure} \end{cases}$$

Bernoulli R.V.

think of this as a clever way to count up # of successes.

Then

Note that ~~each~~ the X_i's ~~is~~ are independent (trials ~~have~~ has success or failure independent of ~~previous~~ other trials outcomes)

Also $E[X_i] = 1 \cdot p + 0 \cdot (1-p) = p$

Therefore,

$$E[X] = E[X_1 + \ldots + X_n]$$
$$= E[X_1] + \ldots + E[X_n]$$ independence of X_i
$$= p + \ldots + p$$
$$= np$$ ← we had found this result previously

(notes p. 120 123 10.18.18)

same example in notes p. 146-147 10.23.18

EXAMPLE (Expectation of a Hypergeometric Random Variable)

Let X be a hypergeometric R.V. with parameters N, n, m. Recall

$$P\{X=i\} = p(i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \quad \text{for } i=0,1,\ldots,n$$

see textbook p.151 (Section 4.8.3)

- select n balls from an urn with
  m white balls
  $N-m$ black balls

X = # of white balls selected

without replacement!

We found $E[X] = \frac{mn}{N}$

Let $X = X_1 + X_2 + \ldots + X_m$

where

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ white ball is selected} \\ 0 & \text{otherwise} \end{cases}$$

again a clever way to add up how many white balls are selected.

Note $P\{X_i = 1\} = \frac{n}{N} \equiv p$

n chances out of N for white ball i to be selected = proportion of all balls selected.

$$= \frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$$

then $P\{X_i = 0\} = 1-p$

As in the previous example, X_i are each Bernoulli R.V.

So $E[X_i] = 1\cdot p + 0\cdot(1-p) = p$.

Then

$$E[X] = E[X_1 + X_2 + \dots + X_m]$$
$$= E[X_1] + E[X_2] + \dots + E[X_m]$$
$$= p + p + \dots + p$$
$$= mp$$
$$= m\left(\frac{n}{N}\right)$$

← See also notes p. 147, 140, 107 ...

Recall, we had found that the expected value for X was the same for drawing balls with and w/o replacement $\left[\begin{matrix} m \text{ white balls} \\ N-m \text{ black balls}\end{matrix} \quad \text{select } n \text{ balls}\right]$

with replacement	without replacement
$P\{X=i\} = \binom{n}{i} p^i (1-p)^{n-i}$ $\quad i=0 \dots n$	$P\{X=i\} = \dfrac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$ $\quad i=0,\dots n$
$p = \dfrac{m}{N}$	
$E[X] = np = \dfrac{nm}{N}$	$E[X] = \dfrac{mn}{N}$

same expected value

7.4 Covariance, Variance of Sums, and Correlations

Recall: For a single R.V.

Discrete: $E[X] = \sum_x x\,p(x)$

$$Var(X) = E\left[(X-\mu)^2\right] = E[X^2] - (E[X])^2$$

$$\mu = E[X]$$

Continuous

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx \quad = \mu$$

$$Var(X) = E\left[(X-\mu)^2\right] = E[X^2] - (E[X])^2$$

Def: Covariance between X and Y

The covariance between X and Y, denoted by $Cov(X,Y)$, is

$$Cov(X,Y) = E\left[(X-E[X])(Y-E[Y])\right]$$

Note:

- $Cov(X,Y) = E\left[XY - E[X]Y - E[Y]X + E[X]E[Y]\right]$

$$= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y]$$

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

- If X and Y are independent (say for continuous R.V. case)

$$E[XY] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f(x,y)\,dx\,dy$$

but $f(x,y) = f_X(x)\, f_Y(y)$ if X and Y are indep.

so

$$E[XY] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f_X(x)\, f_Y(y)\,dx\,dy$$

$$= \int_{-\infty}^{+\infty} x f_X(x)\,dx \cdot \int_{-\infty}^{+\infty} y f_Y(y)\,dy$$

$$= E[X] \cdot E[Y].$$

- So, if X and Y are independent R.V.

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]\cdot E[Y]$$

$$= E[X]\cdot E[Y] - E[X]\cdot E[Y]$$

$$\boxed{\mathrm{Cov}(X,Y) = 0},$$

$\leftarrow$ true for continuous X, Y and for discrete X, Y.