Chapter 4: Expected Values (期望值)

- The Expected Value of a Random Variable(随机变量的期望)
- Variance and Standard Deviation(方差和标准差)
- Covariance and Correlation Coefficient(协方差和相关系数)
- Conditional Expectation (条件期望)

Chapter 4: Expected Values (期望值)

- The Expected Value of a Random Variable(随机变量的期望)
- Variance and Standard Deviation(方差和标准差)
- Covariance and Correlation Coefficient(协方差和相关系数)
- **Conditional Expectation (条件期望)**

Review

The properties of conditional **frequency** function

① $P\{X = x_i \mid Y = y_j\} \ge 0 \quad (i = 1, 2, \cdots)$

② $\sum_{i=1}^{\infty} P\{X = x_i \mid Y = y_j\} = 1$

The properties of conditional **density** function

① $f_{X|Y}(x \mid y) \ge 0$

② $\int_{-\infty}^{\infty} f_{X|Y}(u \mid y)du = 1$

It will be used to define conditional expectation.

When $X = x$, the conditional expectation definition of Y is

$$E(Y \mid X = x) = \sum_y y p_{Y|X}(y \mid x) \quad \text{(Discrete)}$$

$$E(Y \mid X = x) = \int y f_{Y|X}(y \mid x) dy \quad \text{(Continuous)}$$

Generally, the conditional expectation of function $h(Y)$ is

$$E[h(Y) \mid X = x] = \sum_y h(y) p_{Y|X}(y \mid x) \quad \text{(Discrete)}$$

$$E[h(Y) \mid X = x] = \int h(y) f_{Y|X}(y \mid x) dy \quad \text{(Continuous)}$$

Understand Conditional expectation $E(Y \mid X)$

If x is any value within the range of X, $E(Y|(X=x))$ always exists

This is the **function of X**, then **it is r.v.,** noted as $E(Y|X)$

In the same way we can define the expectation and variance for conditional expectation $E(Y|X)$ as follows:

$$E[E(Y \mid X)] \qquad\qquad D[E(Y \mid X)]$$

Theorem

$$\boxed{E(Y) = E[E(Y \mid X)]}$$

$$\boxed{D(Y) = D[E(Y \mid X)] + E[D(Y \mid X)]}$$

Theorem $\boxed{E(Y) = E[E(Y \mid X)]}$ **(Law of Total Expectation)**

Prove: In discrete case (continuous case is similar)

$$E[E(Y \mid X)] = \sum_x E(Y \mid X = x) p_X(x)$$

$$= \sum_x [\sum_y y p_{Y|X}(y \mid x)] p_X(x)$$

Law of total probability

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$$

$$= \sum_y y \sum_x p_{Y|X}(y \mid x) p_X(x)$$

$$= \sum_y y p_Y(y)$$

$$= E[Y].$$

Note: The expectation of a random variable Y can be calculated by first based on the condition of X to find E(Y|X), and then averaging this quantity with respect to X. i.e. weighting conditional expectations appropriately and summing or integrating.

Theorem $D(Y) = D[E(Y \mid X)] + E[D(Y \mid X)]$

Prove: In discrete case (continuous case is similar)

$$\boxed{D(X) = E(X - E(X))^2 = E(X^2) - [E(X)]^2}$$

$$D(Y) = E(Y^2) - [E(Y)]^2$$

By the Law of Total Expectation $\boxed{E(Y) = E[E(Y \mid X)]}$

$$D(Y) = E(E(Y^2|X)) - \{E[E(Y|X)]\}^2$$

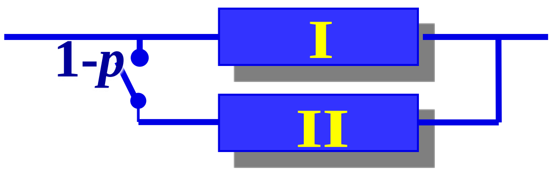
$$= \underbrace{E(E(Y^2|X)) - E\{[E(Y|X)]^2\}} + \underbrace{E\{[E(Y|X)]^2\} - \{E[E(Y|X)]\}^2}$$

$$D[E(Y|X)] = E\{[E(Y|X)]^2\} - \{E[E(Y|X)]\}^2$$

$$E[D(Y|X)] = E\{E(Y^2|X) - [E(Y|X)]^2\} = E[E(Y^2|X)] - E\{[E(Y|X)]^2\}$$

$$\boxed{D(Y) = D[E(Y \mid X)] + E[D(Y \mid X)]}$$

Example: Assume in a system shown in the diagram, the average life span of the component and its backup part all is μ. If the component breaks, the backup part will replace it automatically in the system. However, the probability of an error occurring during the replacement process is p. What is the average life span of the system?



$$\boxed{E(Y) = E[E(Y \mid X)] = \sum_x E(Y \mid X = x) p_X(x)}$$

Solve: Let T denote the life span of the system. E (T)

$$X = \begin{cases} 1, & \text{The replacement of backup part is successful} \\ 0, & \text{The replacement of backup part is failed} \end{cases}$$

Then $E(T \mid X = 1) = 2\mu, \qquad E(T \mid X = 0) = \mu,$

Therefore

$$\begin{aligned} E(T) &= E(T \mid X = 1) P\{X = 1\} + E(T \mid X = 0) P\{X = 0\} \\ &= 2\mu(1 - p) + \mu p = \mu(2 - p) \end{aligned}$$

Random Sum（随机和）

$$T = \sum_{i=1}^{N} X_i$$

What is the expectation and the variance of T?

N is r.v. and has limited values with expectation and variance;
$X_i (i = 1, \dots, N)$ has the same mean value $E(X)$ and variance D(X);
X_i and N are independent.

Insurance companies received N clams at a certain time. Using r.v. X_i to denote the cost of each claim.

Real cases

There are N customers in a shopping mall. The consumption amount of i_{th} customer is $r.v. X_i$

The amount of tasks of a service line is N. The time that is needed for i_{th} service is $r.v. X_i$.

$X_i(i = 1, \dots, N)$ has the same mean value $E(X)$ and variance D(X); X_i and N are independent.

$$T = \sum_{i=1}^{N} X_i$$

Solve:

For $E(T \mid N = n) = nE(X)$, $E(T \mid N) = NE(X)$,

$D(T \mid N) = ND(X)$,

Then

$$E(T) = E[E(T \mid N)] = E[NE(X)] = E(N)E(X).$$

$$\boxed{E(T) = E(N)E(X)}$$

- Thus the average cost for N claims is the average value of the random claim number N multiplying the average cost for one claim.
- Thus the average consumption amount for N customers is the average value of the random customer number N multiplying the average amount for one customer.
- Thus the average time for N services is the average value of the random number N multiplying the average time for one service.

$X_i (i = 1, \ldots, N)$ has the same mean value $E(X)$ and variance D(X); X_i and N are independent.

Solve: In discrete case (continuous case is similar)

If X and Y are independent $D(X+Y) = D(X) + D(Y)$

$$E(T\mid N) = NE(X),\ D(T\mid N) = ND(X),$$

$$\boxed{D(Y) = D[E(Y\mid X)] + E[D(Y\mid X)]}$$

So
$$D[E(T\mid N)] = D[NE(X)] = [E(X)]^2 D(N),$$

$$E[D(T\mid N)] = E(N)D(X),$$

Therefore
$$D(T) = D[E(T\mid N)] + E[D(T\mid N)]$$

$$\boxed{D(T) = [E(X)]^2 D(N) + E(N)D(X).}$$

And
$$D(T\mid N = n) = D(\sum_{i=1}^{n} X_i) = nD(X),$$

Thus the uncertainty of total time T is from the randomness of N and X.
[If N=n is fixed, then $D(T) = nD(X)$]

Exercise: The number of insurance claims N in a certain time period has the expected value being equal to 900 and its standard deviation is 30. As would be the case if the number N were a Poisson random variable with expected value 900. Suppose that the average claim value X is \$1000 and the standard deviation is \$500. What is the expected value of the total claim T? What is the variance of T? $E(T) = ?$ $D(T) = ?$

$E(N) = ?$ $D(N) = ?$ $E(X) = ?$ $D(X) = ?$

$$\boxed{E(T) = E[E(T \mid N)] = E[NE(X)] = E(N)E(X).}$$

$$\boxed{D(T) = [E(X)]^2 D(N) + E(N)D(X)}$$

The **Expected value** of the total claims T:

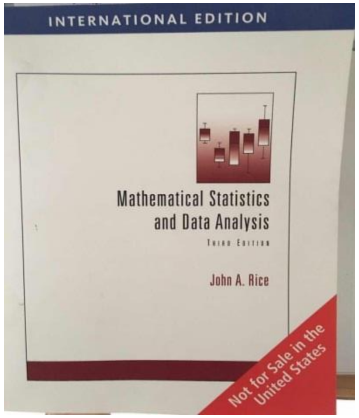
$$E(T) = E(X) \times E(N) = 1000 \times 900 = 900{,}000$$

The **Variance** of the total claims T:

$$D(T) = [E(X)]^2 D(N) + E(N)D(X).$$

$$= 1000^2 \times 30^2 + 900 \times 500^2 = 1.125 \times 10^9$$

Homework



P172：67、77

Supplementary Questions

1. Suppose X and Y are independent, prove that $E(X|Y=y)=E(X)$.

2. Suppose the joint density function of (X,Y) is

$$f(x,y)=\begin{cases} ke^{-(x+y)}, & 0\le y\le x \\ 0, & otherwise \end{cases}$$

(1). Compute $Cov(X,Y)$ and ρ_{XY}
(2). Compute $E(X|Y=y),\ E(Y|X=x)$
(3). Find the density function of $E(X|Y)$ and $E(Y|X)$.