Handout[1] 6

Oct 5, 2022

[Topics]:
Expectation
Variance
Covariance

1 Expectation of Discrete Random Variable

The **expectation** of a random variable X, denoted as $\mathbb{E}(X)$, is the probability-weighted value. **Expectation** is also called **mean** or **expected value**.

Definition 1.1 Suppose X is a discrete random variable with all possible values $x_1, x_2, x_3, \cdots, x_n, \cdots$ (finite or countably infinite), with the probability mass function pmf $f(x)$, then the expectation of X is

$$\mathbb{E}(X) = \sum_i x_i f(x_i).$$

Example 1.1 Let X=1 with probability p and $X = 0$ with probability $(1-p)$, then the expectation of X is

$$\mathbb{E}(X) = 1 \times p + 0 \times (1-p) = p.$$

Example 1.2 For a fair die throw, the expectation is

$$\begin{aligned}\mathbb{E}(x) &= \sum_i x_i f(x_i) = \sum_{i=1}^{6} x_i \times \frac{1}{6} \\ &= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} \\ &= \frac{7}{2}.\end{aligned}$$

Convergence of Expectation : Note that we need a default condition that at least one of the following sums is finite:

$$\sum_{x_i \ positive} x_i f(x_i), \qquad \sum_{x_i \ negative} x_i f(x_i)$$

If both are finite, then $\mathbb{E}(x) = \sum_i x_i f(x_i)$ is well defined.
If $\sum_{x_i \ positive} x_i f(x_i)$ is finite while $\sum_{x_i \ negative} x_i f(x_i)$ is infinite, then we define $\mathbb{E}(X)$ to be $-\infty$.
If $\sum_{x_i \ positive} x_i f(x_i)$ is infinite while $\sum_{x_i \ negative} x_i f(x_i)$ is finite, then we define $\mathbb{E}(X)$ to be $+\infty$.
If both are infinite, then $\mathbb{E}(x) = \sum_i x_i f(x_i)$ is not defined.

Example 1.3 Suppose X has support 2^k, $k = 1, 2, 3, \cdots$, with pmd $f(2^k) = 2^{-k}$, then its expectation is

$$\begin{aligned}\mathbb{E}(X) &= \sum_i x_i f(x_i) \\ &= \sum_{k=1}^{\infty} 2^k \cdot 2^{-k} \\ &= \sum_{k=1}^{\infty} 1 \\ &= +\infty.\end{aligned}$$

[1] This handout is made by Hongkai Wang for Econ 329 Economic Statistics. It is adapted from Professor Wiseman's lectures, however, all errors are mine.

2 Expectation of Continuous Random Variable

Definition 2.1 If X is continuously distributed with pdf $f(x)$, then its **expectation** is defined as

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x \cdot f(x)dx.$$

Example 2.1 Suppose $f(x) = 1$ on $0 < x < 1$ and $f(x) = 0$ otherwise, then the expectation of X is

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{+\infty} x \cdot f(x)dx \\ &= \int_{0}^{1} x \cdot 1dx \\ &= \frac{1}{2}x^2\Big|_0^1 \\ &= \frac{1}{2}.\end{aligned}$$

Convergence of Expectation : Similarly as for discrete random variable, in order to well define $\mathbb{E}(X)$, at least one of the following integrals is finite:

$$\int_0^{+\infty} xf(x)dx, \qquad \int_{-\infty}^{0} xf(x)dx$$

If both integrals are infinite, then $\mathbb{E}(X)$ is not defined, or say it does not exist.

Example 2.2 Suppose $f(x) = x^{-2}$ for $x > 1$, and 0 otherwise, then its expectation is

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{+\infty} x \cdot f(x)dx \\ &= \int_{1}^{+\infty} x \cdot x^{-2}dx \\ &= \int_{1}^{+\infty} \frac{1}{x}dx \\ &= log(x)\Big|_1^{+\infty} \\ &= +\infty.\end{aligned}$$

3 Expectation of Function of Random Variable

Theorem 3.1 Suppose X is a random varibale, and $Y = g(X)$, where $y = g(x)$ is a function of x.
(1) If X is discrete random variable with pmf $f(x)$, then

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_i g(x_i)f(x_i).$$

(1) If X is continuous random variable with pdf $f(x)$, then

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x).$$

Example 3.1 Suppose X is a discrete random variable with support $\{-1, 0, 1\}$, and the pmf is given by

$$f(-1) = P(X = -1) = 0.2, \quad f(0) = P(X = 0) = 0.5, \quad f(1) = P(X = 1) = 0.3.$$

Find $\mathbb{E}(X^2)$.
Solution : $\mathbb{E}(X^2) = (-1)^2 \cdot f(-1) + 0^2 \cdot f(0) + 1^2 \cdot f(1) = 0.5$.

Example 3.2 : Suppose $f(x) = 1$ on $0 \leq x \leq 1$ and 0 otherwise. Find $\mathbb{E}(X^2)$.
Solution : $\mathbb{E}(X^2) = \int_0^1 x^2 \cdot 1dx = \frac{1}{3}x^3\Big|_0^1 = \frac{1}{3}$.

Note that in general, $\mathbb{E}(g(X)) \neq g(\mathbb{E}(X))$. As in **example 3.1**, $\mathbb{E}(X) = 0.1$, hence $g(\mathbb{E}(X)) = 0.1^2 = 0.01$, while $\mathbb{E}(g(X)) = \mathbb{E}(X^2) = 0.5$.

Theorem 3.2 Linearity of Expectation : For any constant a and b,

$$\mathbb{E}(a + bX) = a + b\mathbb{E}(X).$$

Proof : If X is continuous, then for constant a and b,

$$\begin{aligned}
\mathbb{E}(a + bX) &= \int_{-\infty}^{+\infty} (a + bx)f(x)dx \\
&= \int_{-\infty}^{+\infty} af(x)dx + \int_{-\infty}^{+\infty} bxf(x)dx \\
&= a \underbrace{\int_{-\infty}^{+\infty} f(x)dx}_{=1 \ by \ property \ of \ pdf} + b \underbrace{\int_{-\infty}^{+\infty} xf(x)dx}_{=\mathbb{E}(X) \ by \ definition} \\
&= a + b\mathbb{E}(X).
\end{aligned}$$

If X is discrete, then

$$\begin{aligned}
\mathbb{E}(a + bX) &= \sum_i (a + bx_i)f(x_i) \\
&= \sum_i af(x_i) + \sum_i bx_i f(x_i) \\
&= a \underbrace{\sum_i f(x_i)}_{=1} + b \underbrace{\sum_i x_i f(x_i)}_{=\mathbb{E}(X)} \\
&= a + b\mathbb{E}(X).
\end{aligned}$$

Example 3.3 Suppose X is random variable with $\mathbb{E}(X) = 5$, suppose $Y = 3 + 7X$, find $\mathbb{E}(Y)$.
Solution : $\mathbb{E}(Y) = \mathbb{E}(3 + 7X) = 3 + 7\mathbb{E}(X) = 38$.

Important Property : Let X and Y be two random variables, then
(1) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ whenever X and Y are independent or not;
(2) $\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ only when X and Y are independent.

We will take a deeper look at this in the chapter of joint random variables.

4 Median

Definition 4.1 : Let X be a random variable, then its median is a number m such that

$$\mathbb{P}(X \leq m) \geq \tfrac{1}{2}, \quad \text{and} \quad \mathbb{P}(X \geq m) \geq \tfrac{1}{2}.$$

Example 4.1 Let X be a discrete random variable with pmf given by:

$$f(1) = 0.1, \quad f(2) = 0.2, \quad f(3) = 0.3, \quad f(4) = 0.4.$$

Then its median is 3. As $\mathbb{P}(X \leq 3) = P(X = 0or1or2) = f(1) + f(2) + f(3) = 0.6 \geq \frac{1}{2}$ and $\mathbb{P}(X \geq 3) = P(X = 3or4) = f(3) + f(4) = 0.7 \geq \frac{1}{2}$.

Note that the median may not be unique.
Example 4.2 Let X be a discrete random variable with pmf given by:

$$f(1) = 0.1, \quad f(2) = 0.4, \quad f(3) = 0.3, \quad f(4) = 0.2.$$

Since $\mathbb{P}(X \leq 2) = \frac{1}{2} \geq \frac{1}{2}$ and $\mathbb{P}(X \geq 3) = \frac{1}{2} \geq \frac{1}{2}$, any number $m \in [2, 3]$ is a median.

Median of Continuous Random Variable For a continuous random varibale X, its median is the 0.5 quantile $F^{-1}(0.5)$.

Example 4.3 Suppose X is continuous random varibale with pdf given by

$$f(x) = \begin{cases} 4x^3 & if \quad 0 < x < 1 \\ 0 & otherwise \end{cases}$$

Then its cdf is

$$F(x) = \begin{cases} 0 & if \quad x \leq 0 \\ x^4 & if \quad 0 < x < 1 \\ 1 & if \quad x \geq 1 \end{cases}$$

Now solving for $F(m) = \frac{1}{2}$ we get $m = (\frac{1}{2})^{\frac{1}{4}}$. i,e. $(\frac{1}{2})^{\frac{1}{4}}$ is the median of X.

Example 4.4 Suppose X is continuous random varibale with pdf given by

$$f(x) = \begin{cases} \dfrac{1}{2} & if \quad 0 \leq x \leq 1 \\ 1 & if \quad \dfrac{5}{2} \leq x \leq 3 \\ 0 & otherwise \end{cases}$$

Then its cdf is

$$F(x) = \begin{cases} 0 & if \quad x \leq 0 \\ \dfrac{1}{2}x & if \quad 0 < x < 1 \\ \dfrac{1}{2} & if \quad 1 \leq x \leq \dfrac{5}{2} \\ x - 2 & if \quad \dfrac{5}{2} << 3 \\ 1 & if \quad x \geq 3 \end{cases}$$

Then the median of X is any number $m \in [1, \frac{5}{2}]$.

Theorem 4.1 If X is symmetricly distributed, then the median of X and the expectation of X are identical.

5 Variance

Definition 5.1 Variance : For a random variable X, the variance of X, denoted by $Var(X)$, is defined by

$$Var(X) = \mathbb{E}((X - \mathbb{E}(X))^2).$$

If X is discrete with expectation $\mathbb{E}(X)$, then its variance is $Var(X) = \sum_i (x_i - \mathbb{E}(X))^2 f(x_i)$.
If X is continuous with expectation $\mathbb{E}(X)$, then its variance is $Var(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 \cdot f(x) dx$.

Definition 5.1 Standard Deviation : The standard deviation of a random varibale X, denoted as δ_X, is the square root of its variance

$$\delta_X = \sqrt{Var(X)}.$$

Example 5.1 Suppose $X \sim Bernoulli(p)$, then $\mathbb{E}(X) = 1\times p + 0\times(1-p) = p$, hence $Var(X) = (1-p)^2\times p + (0-p)^2\times(1-p) = p(1-p)$.

Example 5.2 Suppose $X \sim Uniform[0,1]$, then $\mathbb{E}(X) = 0.5$, hence $Var(X) = \int_0^1 (x-0.5)^2 \times 1 \times dx = \frac{1}{3}(x-0.5)^3\Big|_0^1 = \frac{1}{12}$.

Theorem 5.1 $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
Proof :

$$\begin{aligned} Var(X) &= \mathbb{E}((X - E(X))^2) \\ &= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + (\mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \end{aligned}$$

Properties of Variance

Property 5.1 For any random variable X, $Var(X) \geq 0$.

Property 5.2 If X is constant, then $Var(X) = 0$.
Proof : For X to be constant, say $X = C$, then $\mathbb{E}(X) = C$. Hence $Var(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}((C - C)^2) = 0$.

Property 5.3 If X is a random varaible and a is a constant, then $Var(aX) = a^2 Var(X)$.
Proof : By Theorem 5.1,

$$\begin{aligned} Var(aX) &= \mathbb{E}((aX)^2) - (\mathbb{E}(aX))^2 \\ &= \mathbb{E}(a^2X^2) - (a\mathbb{E}(X))^2 \\ &= a^2\mathbb{E}(X^2) - a^2(\mathbb{E}(X))^2 \\ &= a^2(\mathbb{E}(X^2) - (\mathbb{E}(X))^2) \\ &= a^2 Var(X). \end{aligned}$$

Property 5.4 If X is a random varaible and a is a constant, then $Var(a + X) = Var(X)$.
Proof : By Theorem 5.1,

$$\begin{aligned} Var(a + X) &= \mathbb{E}((a + X)^2) - (\mathbb{E}(a + X))^2 \\ &= \mathbb{E}(a^2 + X^2 + 2aX) - (a + \mathbb{E}(X))^2 \\ &= a^2 + \mathbb{E}(X^2) + 2a\mathbb{E}(X) - a^2 - (\mathbb{E}(X))^2 - 2a\mathbb{E}(X) \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= Var(X). \end{aligned}$$

Property 5.5 Let X and Y be two random varibales, then

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y).$$

If X and Y are independent, then $Cov(X, Y) = 0$, hence $Var(X + Y) = Var(X) + Var(Y)$.

6 Covariance and Correlation

Definition 6.1 Covariance : For two random variabels X and Y, the covariance of X and Y, denoted as $Cov(X,Y)$, is defined by

$$Cov(X,Y) = \mathbb{E}\{(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\}.$$

It is well defined if such expectation exists. Otherwise, we say the covariance is undefined.

If $Cov(X,Y) = 0$, we say X and Y are uncorrelated.
If $Cov(X,Y) > 0$, we say X and Y are positively correlated.
If $Cov(X,Y) < 0$, we say X and Y are negatively correlated.

Theorem 6.1 For random variable X and Y such that $Var(X) < \infty$ and $Var(Y) < \infty$,

$$Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Proof : By definition,

$$\begin{aligned} Cov(X,Y) &= \mathbb{E}\{(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\} \\ &= \mathbb{E}\{XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)\} \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y). \end{aligned}$$

Definition 6.2 Correlation : Let X and Y be two random variables with variance $Var(X)$ and $Var(Y)$, then the correlation of X and Y, denoted as $\rho(X,Y)$, is defined as

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

Note that an important property for correlation is $-1 \le \rho \le 1$.

Properties of Covariance
Property 6.1 If X and Y are independent random varaibles with $0 < Var(X) < \infty$ and $0 < Var(Y) < \infty$, then $Cov(X,Y) = \rho(X,Y) = 0$.
Proof : Since X and Y are independent, we have $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Hence $Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) = 0$, which also leads to $\rho(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = 0$.

Note that the inverse is not true generally: $Cov(X,Y) = 0$ or $\rho(X,Y) = 0$ does not imply X and Y are independent.

Property 6.2 $Cov(X,Y) = Cov(Y,X)$.

Property 6.3 $Cov(X,X) = Var(X)$.

Property 6.4 If X is a random varaible and c is a constant, then $Cov(X,c) = 0$.
Proof :

$$\begin{aligned} Cov(X,c) &= \mathbb{E}(Xc) - \mathbb{E}(X)\mathbb{E}(c) \\ &= c\mathbb{E}(X) - \mathbb{E}(X) \cdot c \\ &= 0. \end{aligned}$$

Property 6.5 If X and Y are random variables and a is a constant, then $Cov(aX,Y) = aCov(X,Y)$ and $Cov(X,aY) = aCov(X,Y)$ **Proof :**

$$\begin{aligned} Cov(aX,Y) &= \mathbb{E}(aXY) - \mathbb{E}(aX)\mathbb{E}(Y) \\ &= a\mathbb{E}(XY) - a\mathbb{E}(X)\mathbb{E}(Y) \\ &= a(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)) \\ &= aCov(X,Y). \end{aligned}$$

7 Examples

7.1 Bernoulli Random Variable: $P(X = 1) = p$, $P(X = 0) = 1 - p$.
The expectation is:

$$\begin{aligned}\mathbb{E}(X) &= 1 \times p + 0 \times (1 - p) \\ &= p.\end{aligned}$$

And

$$\begin{aligned}\mathbb{E}(X^2) &= 1^2 \times p + 0^2 \times (1 - p) \\ &= p.\end{aligned}$$

The variance is

$$\begin{aligned}Var(X) &= (1 - p)^2 \times p + (0 - p)^2 \times (1 - p) \\ &= p - p^2.\end{aligned}$$

Or using Theorem 5.1

$$\begin{aligned}Var(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= p - p^2.\end{aligned}$$

7.2 Binomial Random Variables:

$$f(x) = P(X = x) = \tbinom{n}{x} p^k (1 - p)^{n-x},\ x = 0, 1, 2, 3, \cdots, \text{ and } 0 < p < 1$$

The expectation is

$$\begin{aligned}\mathbb{E}(X) &= \sum_{x=0}^{\infty} x \binom{n}{x} p^x (1 - p)^{n-x} \\ &= \sum_{x=1}^{\infty} n \binom{n-1}{x-1} p^x (1 - p)^{n-x} \\ &= np \sum_{x=1}^{\infty} \binom{n-1}{x-1} p^{x-1} (1 - p)^{n-x} \\ &= np \sum_{x=1}^{\infty} \binom{n-1}{x-1} p^{x-1} (1 - p)^{n-x} \\ &= np \cdot \underbrace{\sum_{\tilde{x}=0}^{\infty} \binom{\tilde{n}}{\tilde{x}} p^{\tilde{x}} (1 - p)^{\tilde{n}-\tilde{x}}}_{sum\ of\ pmf\ of\ \tilde{X} \sim Binomial(\tilde{n}, p)}, \quad where\ \tilde{n} = n - 1 \quad and\ \tilde{x} = x - 1 \\ &= np.\end{aligned}$$

And $\mathbb{E}(X^2) = np + n(n - 1)p^2$. (Derivation not required)
Then

$$\begin{aligned}Var(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= np + n(n - 1)p^2 - (np)^2 \\ &= np(1 - p).\end{aligned}$$

7.3 Uniform Random Variable: $X \sim Unifrom[a, b]$, then $f(x) = \frac{1}{b-a}$ for $x \in [a, b]$.
The expectation is

$$\mathbb{E}(X) = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{2(b-a)} x^2 \Big|_a^b = \frac{a+b}{2}.$$

And

$$\mathbb{E}(X^2) = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{3(b-a)} x^3 \Big|_a^b = \frac{a^2+b^2+ab}{3}.$$

So

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{a^2+b^2+ab}{3} - (\frac{a+b}{2})^2 = \frac{(b-a)^2}{12}.$$

7.4 Exponential Random Variable: $X \sim Exponential(\lambda)$.
For $X \sim Exponential(\lambda)$, we have $\mathbb{E}(X) = \frac{1}{\lambda}$ and $\mathbb{E}(X^2) = \frac{2}{\lambda^2}$, what is $Var(X)$?

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{2}{\lambda^2} - (\frac{1}{\lambda})^2 = \frac{1}{\lambda^2}.$$

7.5 Suppose X and Y are two independent random variables, and $E(X) = 0$, $E(Y) = 3$, $E(X^2) = 1$, $E(Y^2) = 7$. What is the covariance of X and Y?
Solution : Since X and Y are independent, we have $Cov(X, Y) = 0$.

Good Luck on Your Midterm!