Math 461 Spring 2024

Renming Song

University of Illinois Urbana-Champaign

March 29, 2024

Outline

Outline

HW8 is due today before the end of class.

Solution to HW8 will be on my homepage later this afternoon.

HW8 is due today before the end of class.

Solution to HW8 will be on my homepage later this afternoon.

Outline

1. General Info
2. 7.4 Covariance, variance of sums and correlations

Theorem

If X and Y are independent random variables, then for any functions ϕ and ψ on $\mathbb{R}$,

$$E[\phi(X)\psi(Y)] = E[\phi(X)]E[\psi(Y)].$$

Let's prove this in the absolute continuous case. Let f_X and f_Y be the density of X and Y respectively. Since X and Y are independent, the joint density of X and Y is $f_X(x)f_Y(y)$. Thus

$$E[\phi(X)\psi(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x)\psi(y)f_X(x)f_Y(y)dxdy$$

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

Theorem

If X and Y are independent random variables, then for any functions ϕ and ψ on $\mathbb{R}$,

$$E[\phi(X)\psi(Y)] = E[\phi(X)]E[\psi(Y)].$$

Let's prove this in the absolute continuous case. Let f_X and f_Y be the density of X and Y respectively. Since X and Y are independent, the joint density of X and Y is $f_X(x)f_Y(y)$. Thus

$$E[\phi(X)\psi(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x)\psi(y)f_X(x)f_Y(y)dxdy$$

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

Definition

The covariance $\text{Cov}(X, Y)$ of two random variables X and Y is defined by

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])].$$

One can easily check that

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

If X and Y are independent, then $\text{Cov}(X, Y) = 0$. But the converse is not true. When $\text{Cov}(X, Y) = 0$, we say that X and Y are uncorrelated. Independence implies uncorrelated, but not the other way around.

Definition

The covariance $\mathrm{Cov}(X, Y)$ of two random variables X and Y is defined by

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

One can easily check that

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

If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$. But the converse is not true. When $\mathrm{Cov}(X, Y) = 0$, we say that X and Y are uncorrelated. Independence implies uncorrelated, but not the other way around.

Definition

The covariance $\text{Cov}(X, Y)$ of two random variables X and Y is defined by

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])].$$

One can easily check that

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

If X and Y are independent, then $\text{Cov}(X, Y) = 0$. But the converse is not true. When $\text{Cov}(X, Y) = 0$, we say that X and Y are uncorrelated. Independence implies uncorrelated, but not the other way around.

Example 1

Suppose $P(X = 0) = P(X = 1) = P(X = -1) = \frac{1}{3}$ and

$$Y = \begin{cases} 0, & \text{if } X \neq 0, \\ 1, & \text{if } X = 0. \end{cases}$$

Then $XY = 0$, thus $E[XY] = 0$. We also have $E[X] = 0$ and $E[Y] = \frac{1}{3}$, thus $E[X]E[Y] = 0$. Hence $\text{Cov}(X, Y) = 0$. But X and Y are obviously not independent.

Example 1

Suppose $P(X=0)=P(X=1)=P(X=-1)=\frac{1}{3}$ and

$$Y=\begin{cases}0, & \text{if } X\neq 0,\\ 1, & \text{if } X=0.\end{cases}$$

Then $XY=0$, thus $E[XY]=0$. We also have $E[X]=0$ and $E[Y]=\frac{1}{3}$, thus $E[X]E[Y]=0$. Hence $\mathrm{Cov}(X,Y)=0$. But X and Y are obviously not independent.

Example 2

A box has 3 balls labeled 1, 2, 3. Two Balls are randomly selected without replacement. Let X be the number on the first ball and Y the number on the second. Find $\mathrm{Cov}(X, Y)$.

$$P(X=1, Y=2) = P(X=1, Y=3) = P(X=2, Y=1)$$
$$= P(X=2, Y=3) = P(X=3, Y=1) = P(X=3, Y=2) = \frac{1}{6}.$$

Thus

$$E[XY] = (2+3+2+3+6+6)\frac{1}{6} = \frac{11}{3}.$$

Since $P(X=1) = P(X=2) = P(X=3) = P(Y=1) = P(Y=2) = P(Y=3) = \frac{1}{3}$, $E[X] = E[Y] = 2$. Hence $\mathrm{Cov}(X, Y) = -\frac{1}{3}$.

Example 2

A box has 3 balls labeled 1, 2, 3. Two Balls are randomly selected without replacement. Let X be the number on the first ball and Y the number on the second. Find $\mathrm{Cov}(X, Y)$.

$$P(X=1, Y=2) = P(X=1, Y=3) = P(X=2, Y=1)$$
$$= P(X=2, Y=3) = P(X=3, Y=1) = P(X=3, Y=2) = \frac{1}{6}.$$

Thus

$$E[XY] = (2+3+2+3+6+6)\frac{1}{6} = \frac{11}{3}.$$

Since $P(X=1) = P(X=2) = P(X=3) = P(Y=1) = P(Y=2) = P(Y=3) = \frac{1}{3}$, $E[X] = E[Y] = 2$. Hence $\mathrm{Cov}(X, Y) = -\frac{1}{3}$.

Example 3

The joint density of X and Y is given by

$$f(x,y) = \begin{cases} x+y, & 0<x<1, 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

Find $\mathrm{Cov}(X,Y)$.

$$E[XY] = \int_0^1 \int_0^1 xy(x+y)dxdy = \int_0^1 \int_0^1 (x^2y + xy^2)dxdy$$

$$= \int_0^1 \left(\frac{y}{3} + \frac{y^2}{2}\right)dy = \frac{1}{3}.$$

Example 3

The joint density of X and Y is given by

$$f(x,y) = \begin{cases} x+y, & 0<x<1, 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

Find $\text{Cov}(X, Y)$.

$$E[XY] = \int_0^1 \int_0^1 xy(x+y)dxdy = \int_0^1 \int_0^1 (x^2y + xy^2)dxdy$$
$$= \int_0^1 (\frac{y}{3} + \frac{y^2}{2})dy = \frac{1}{3}.$$

$$E[X] = \int_0^1 \int_0^1 x(x+y)dxdy = \int_0^1 \int_0^1 (x^2 + xy)dxdy$$
$$= \int_0^1 (\frac{1}{3} + \frac{y}{2})dy = \frac{7}{12}.$$

Similarly, $E[Y] = \frac{7}{12}$.

Thus

$$\text{Cov}(X, Y) = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}.$$

$$E[X] = \int_0^1 \int_0^1 x(x+y)dxdy = \int_0^1 \int_0^1 (x^2+xy)dxdy$$
$$= \int_0^1 (\frac{1}{3}+\frac{y}{2})dy = \frac{7}{12}.$$

Similarly, $E[Y] = \frac{7}{12}$.

Thus

$$\text{Cov}(X, Y) = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}.$$

Proposition

(i) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.

(ii) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

(iii) $\mathrm{Cov}(aX, Y) = a\mathrm{Cov}(X, Y)$.

(iv) $\mathrm{Cov}(\sum_{i=1}^{m} X_i, \sum_{j=1}^{n} Y_j) = \sum_{i=1}^{m} \sum_{j=1}^{n} \mathrm{Cov}(X_i, Y_j)$.

(i), (ii) and (iii) are obvious. Let's look at $\mathrm{Cov}(X_1 + X_2, Y)$.

$$\begin{aligned}\mathrm{Cov}(X_1 + X_2, Y) &= E[(X_1 + X_2)Y] - E[X_1 + X_2]E[Y] \\ &= E[X_1 Y] - E[X_1]E[Y] + E[X_2 Y] - E[X_2]E[Y] \\ &= \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y).\end{aligned}$$

The general case follows by induction.

Proposition

(i) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.

(ii) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

(iii) $\mathrm{Cov}(aX, Y) = a\mathrm{Cov}(X, Y)$.

(iv) $\mathrm{Cov}(\sum_{i=1}^m X_i, \sum_{j=1}^n Y_j) = \sum_{i=1}^m \sum_{j=1}^n \mathrm{Cov}(X_i, Y_j)$.

(i), (ii) and (iii) are obvious. Let's look at $\mathrm{Cov}(X_1 + X_2, Y)$.

$$\begin{aligned}\mathrm{Cov}(X_1 + X_2, Y) &= E[(X_1 + X_2)Y] - E[X_1 + X_2]E[Y] \\ &= E[X_1 Y] - E[X_1]E[Y] + E[X_2 Y] - E[X_2]E[Y] \\ &= \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y).\end{aligned}$$

The general case follows by induction.

Combining (ii) and (iv) above, we get

$$\begin{aligned}\mathrm{Var}\Big(\sum_{i=1}^{n} X_i\Big) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \sum_{i<j} \mathrm{Cov}(X_i, X_j).\end{aligned}$$

In particular, if $X_1, \ldots, X_n$ are independent, then

$$\mathrm{Var}\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Combining (ii) and (iv) above, we get

$$\begin{aligned}\mathrm{Var}(\sum_{i=1}^{n} X_i) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).\end{aligned}$$

In particular, if $X_1, \ldots, X_n$ are independent, then

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Example 4

Suppose that X_1, X_2, X_3 are independent, $\mathrm{Var}(X_1) = \sigma_1^2$, $\mathrm{Var}(X_2) = \sigma_2^2$ and $\mathrm{Var}(X_3) = \sigma_3^2$. Find $\mathrm{Cov}(X_1 - X_2, X_2 + X_3)$.

$$\mathrm{Cov}(X_1 - X_2, X_2 + X_3) = \mathrm{Cov}(X_1, X_2 + X_3) - \mathrm{Cov}(X_2, X_2 + X_3)$$
$$= -\mathrm{Cov}(X_2, X_2 + X_3) = -\mathrm{Cov}(X_2, X_2) - \mathrm{Cov}(X_2, X_3)$$
$$= -\mathrm{Var}(X_2) = -\sigma_2^2.$$

Example 4

Suppose that X_1, X_2, X_3 are independent, $\mathrm{Var}(X_1) = \sigma_1^2$, $\mathrm{Var}(X_2) = \sigma_2^2$ and $\mathrm{Var}(X_3) = \sigma_3^2$. Find $\mathrm{Cov}(X_1 - X_2, X_2 + X_3)$.

$$\mathrm{Cov}(X_1 - X_2, X_2 + X_3) = \mathrm{Cov}(X_1, X_2 + X_3) - \mathrm{Cov}(X_2, X_2 + X_3)$$
$$= -\mathrm{Cov}(X_2, X_2 + X_3) = -\mathrm{Cov}(X_2, X_2) - \mathrm{Cov}(X_2, X_3)$$
$$= -\mathrm{Var}(X_2) = -\sigma_2^2.$$

Example 5

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with common mean μ and common variance σ^2.

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

is called the sample mean and

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

is called the sample variance. Find (a) $\mathrm{Var}(\overline{X})$; (b) $E[S^2]$.

$$\mathrm{Var}(\overline{X}) = \frac{1}{n^2}\mathrm{Var}(\sum_{i=1}^{n} X_i) = \frac{\sigma^2}{n}.$$

Example 5

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with common mean μ and common variance σ^2.

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

is called the sample mean and

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

is called the sample variance. Find (a) $\mathrm{Var}(\overline{X})$; (b) $E[S^2]$.

$$\mathrm{Var}(\overline{X}) = \frac{1}{n^2} \mathrm{Var}(\sum_{i=1}^{n} X_i) = \frac{\sigma^2}{n}.$$

For (b), we start with the following algebraic identity

$$
\begin{aligned}
(n-1)S^2 &= \sum_{i=1}^{n}(X_i - \mu + \mu - \overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i=1}^{n}(\overline{X} - \mu)^2 - 2(\overline{X} - \mu)\sum_{i=1}^{n}(X_i - \mu) \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 + n(\overline{X} - \mu)^2 - 2(\overline{X} - \mu)n(\overline{X} - \mu) \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 - n(\overline{X} - \mu)^2.
\end{aligned}
$$

Taking expectation, we get

$$(n-1)E[S^2] = n\sigma^2 - n\mathrm{Var}(\overline{X}) = (n-1)\sigma^2,$$

For (b), we start with the following algebraic identity

$$
\begin{aligned}
(n-1)S^2 &= \sum_{i=1}^{n}(X_i - \mu + \mu - \overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i=1}^{n}(\overline{X} - \mu)^2 - 2(\overline{X} - \mu)\sum_{i=1}^{n}(X_i - \mu) \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 + n(\overline{X} - \mu)^2 - 2(\overline{X} - \mu)n(\overline{X} - \mu) \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 - n(\overline{X} - \mu)^2.
\end{aligned}
$$

Taking expectation, we get

$$(n-1)E[S^2] = n\sigma^2 - n\mathrm{Var}(\overline{X}) = (n-1)\sigma^2,$$

since $E[\overline{X}] = \mu$. Hence $E[S^2] = \sigma^2$.

Definition

The correlation coefficient $\rho(X, Y)$ of two random variables X and Y is defined by

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

We always have $|\rho(X, Y)| \leq 1$. $\rho(X, aX) = 1$ if $a > 0$, $\rho(X, aX) = -1$ if $a < 0$, and $\rho(X, Y) = 0$ if X and Y are independent. $|\rho(X, Y)| = 1$ if and only if $P(X = aY) = 1$ for some $a \neq 0$.

since $E[\overline{X}] = \mu$. Hence $E[S^2] = \sigma^2$.

Definition

The correlation coefficient $\rho(X, Y)$ of two random variables X and Y is defined by

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

We always have $|\rho(X, Y)| \leq 1$. $\rho(X, aX) = 1$ if $a > 0$, $\rho(X, aX) = -1$ if $a < 0$, and $\rho(X, Y) = 0$ if X and Y are independent. $|\rho(X, Y)| = 1$ if and only if $P(X = aY) = 1$ for some $a \neq 0$.

since $E[\overline{X}] = \mu$. Hence $E[S^2] = \sigma^2$.

Definition

The correlation coefficient $\rho(X, Y)$ of two random variables X and Y is defined by

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

We always have $|\rho(X, Y)| \leq 1$. $\rho(X, aX) = 1$ if $a > 0$, $\rho(X, aX) = -1$ if $a < 0$, and $\rho(X, Y) = 0$ if X and Y are independent. $|\rho(X, Y)| = 1$ if and only if $P(X = aY) = 1$ for some $a \neq 0$.

Now we are going to use the formula

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i\neq j} \mathrm{Cov}(X_i, X_j)$$
$$= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

to find the variance of some complicated random variables.

Example 6

Suppose S_n is a binomial random variable with parameters (n, p). S_n is the total number of successes in n indep trails each of which results in a success with probability p. For $i = 1, \ldots, n$, let $X_i = 1$ if the i-th trial results in a success and $X_i = 0$ otherwise. Then $X_1, \ldots, X_n$ indep Bernoulli random variables with parameter p and $S_n = \sum_{i=1}^{n} X_i$. Thus

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = np(1-p).$$

Now we are going to use the formula

$$\begin{aligned}\mathrm{Var}(\sum_{i=1}^{n} X_i) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i\neq j} \mathrm{Cov}(X_i, X_j)\\ &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).\end{aligned}$$

to find the variance of some complicated random variables.

Example 6

Suppose S_n is a binomial random variable with parameters (n, p). S_n is the total number of successes in n indep trails each of which results in a success with probability p. For $i = 1, \ldots, n$, let $X_i = 1$ if the i-th trial results in a success and $X_i = 0$ otherwise. Then $X_1, \ldots, X_n$ indep Bernoulli random variables with parameter p and $S_n = \sum_{i=1}^{n} X_i$. Thus

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = np(1-p).$$

Example 7

Let X be a negative binomial random variable with parameters (r, p). X is the number of trials needed in order to get r successes. Let Y_1 be the number of trials needed in order to get the first success; let Y_2 be the number of additional trials, after the first success, to get the second success, . . . , let Y_r be the number of additional trials, after the $(r-1)$-st success, to get the r-th success. Then $Y_1, \ldots, Y_r$ are independent geometric random variables with parameter p and $X = Y_1 + \cdots + Y_r$. Thus

$$\mathrm{Var}(X) = \sum_{i=1}^{r} \mathrm{Var}(Y_i) = \frac{r(1-p)}{p^2}.$$