MH2500 Probability and Introduction to Statistics

Handout 7 - Expected Values - III

Definition

The variance of a random variable measures its variability.

The **covariance** of two random variables measures their joint variability, or degree of association.

Definition

If X and Y are jointly distributed random variables with expectations μ_X and μ_Y, respectively, then the covariance of X and Y is

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

provided that the expectation exists.

Covariance is the average value of the product of the deviation of X from its mean and the deviation of Y from its mean.

Positive covariance means both X and Y tend to be are larger than their respective means or both X and Y tend to be smaller than their respective means.

Negative covariance means X tends to be are larger (smaller) than its mean while Y tends to be smaller (respectively larger) than its mean.

. . . .

$$Cov(X, Y) = E(XY) - E(X)E(Y).$$

Proof:

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

. . . .

If X and Y are independent, then $E(XY) = E(X)E(Y)$ and $Cov(X, Y) = 0$.

Example A

Let $H(x, y) = x^2y + y^2x - x^2y^2$, for $0 \le x, y \le 1$ be the bivariate cumulative distribution function of X and Y. Find $Cov(X, Y)$.

Proof: The marginal cdf of X is

$$F_X(x) = H(x, 1) = x^2 + x - x^2 = x.$$

Hence the marginal distribution of X is the uniform distribution on $[0, 1]$ and so $E(X) = 1/2$. By symmetry, $E(Y) = 1/2$.

The joint density is

$$h(x, y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}(x^2y + y^2x - x^2y^2) = 2x + 2y - 4xy.$$

Example A

Hence

$$E(XY) = \int \int \qquad dxdy$$

$$=$$

$$=$$

$$=$$

$$= \qquad .$$

Therefore,

$$Cov(X, Y) = \qquad = \qquad .$$

Example A

The covariance is negative and it can be seen from the graph that when X is less than its mean, Y tends to be larger than its mean and vice versa.

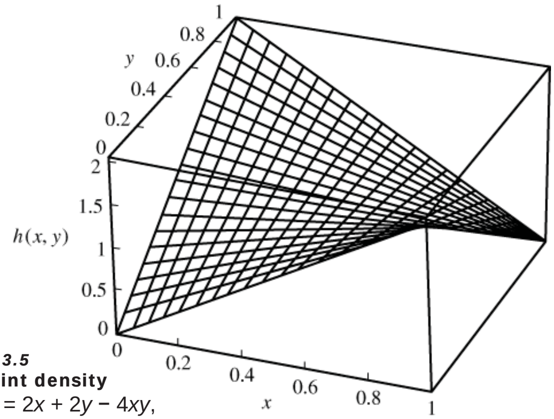

Figure 3.5
The joint density
$h(x, y) = 2x + 2y - 4xy$,
where $0 \le x \le 1$ **and** $0 \le y \le 1$,
which has uniform marginal densities.

Covariance of linear combinations

- Recall that $E(a + X) = a + E(X)$. Hence

$$\begin{aligned} Cov(a + X, Y) &= E\{ \qquad\qquad\qquad\qquad\qquad\qquad \} \\ &= E\{ \qquad\qquad\qquad\qquad\qquad\qquad \} \\ &= E\{ \qquad\qquad\qquad\qquad \} \\ &= Cov(X, Y). \end{aligned}$$

- Recall $E(aX) = aE(X)$. Hence

$$\begin{aligned} Cov(aX, bY) &= E\{ \qquad\qquad\qquad\qquad\qquad \} \\ &= E\{ \qquad\qquad\qquad\qquad\qquad \} \\ &= \\ &= abCov(X, Y). \end{aligned}$$

- Recall $E(X+Y) = E(X) + E(Y)$. Hence

$$
\begin{aligned}
& Cov(X, Y+Z) \\
& = E\{ \qquad\qquad\qquad\qquad\qquad\qquad \} \\
& = E\{ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \} \\
& = E\{ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \} \\
& = \\
& = Cov(X, Y) + Cov(X, Z).
\end{aligned}
$$

Likewise $\quad Cov(X+Y, Z) = Cov(X, Z) + Cov(Y, Z).$

- Combine the above gives

$$
\begin{aligned}
& Cov(aW + bX, cY + dZ) \\
& = \\
& = \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad .
\end{aligned}
$$

Theorem

In general, we have the following.

Theorem A

Suppose that $U = a + \sum_{i=1}^{n} b_i X_i$ and $V = c + \sum_{j=1}^{m} d_j Y_j$. Then

$$Cov(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_i d_j Cov(X_i, Y_j).$$

. . . .

Here is an application of Theorem A. Since $Var(X) = Cov(X, X)$,

$$\begin{aligned} Var(X+Y) &= Cov(X+Y, X+Y) \\ &= Cov(X,X) + Cov(X,Y) + Cov(Y,X) + Cov(Y,Y) \\ &= Var(X) + Var(Y) + 2Cov(X,Y). \end{aligned}$$

Corollary

Corollary A

$$Var\left(a+\sum_{i=1}^{n} b_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j Cov(X_i, X_j).$$

Corollary B

Suppose the X_i's are independent. Then

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

Example B

Suppose $X_1, X_2, \ldots, X_n$ are independent Bernoulli random variables with $P(X_i = 1) = p$. Let $Y = X_1 + X_2 + \cdots + X_n$. Find $Var(Y)$.

Earlier, we saw that for a Bernoulli random variable,

$$Var(X_i) = p(1-p).$$

Hence

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

Correlation coefficient

Definition

If X and Y are jointly distributed random variables and the variance and covariance of both X and Y exist and the variances are nonzero, then the correlation of X and Y, denoted by ρ, is

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

- Both numerator and denominator have the same units and so correlation has no units.
- Suppose both X and Y are subject to linear transformations, both Variance and Covariance changes but ρ does not change.
- Since ρ does not depend on the units of measurement, it is usually a more useful measure of association.

Example D

Referring back to Example A where $h(x, y) = 2x + 2y - 4xy$ and X and Y are marginally uniform. Find $Var(X)$, $Var(Y)$ and ρ.

. . . .

$$E(X^2) = \int_0^1 \int_0^1 x^2(2x + 2y - 4xy)\ dydx$$
$$= \int_0^1 2x^3 + x^2 - 2x^3\ dx\ = \int_0^1 x^2\ dx = \frac{1}{3}.$$

Recall from Example A that $E(X) = \frac{1}{2}$. Hence

$$Var(X) = \frac{1}{3} - \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{12}.$$

Similarly, $Var(Y) = \frac{1}{12}$. Recall from Example A that $Cov(X, Y) = \frac{2}{9}$. Hence

$$\rho = \qquad\qquad\qquad .$$

Notations

Denote by $\sigma_{XY} = Cov(X, Y)$. Then

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y},$$

i.e.,

$$\sigma_{XY} = \rho \sigma_X \sigma_Y.$$

Theorem B

Theorem B

We always have

$$-1 \leq \rho \leq 1.$$

Furthermore, $\rho = \pm 1$ if and only if $P(Y = a + bX) = 1$ for some constants a and b.

Proof:

$$\begin{aligned}
0 &\leq Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) \\
&= Var\left(\frac{X}{\sigma_X}\right) + Var\left(\frac{Y}{\sigma_Y}\right) + 2Cov\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) \\
&= \frac{Var(X)}{\sigma_X^2} + \frac{Var(Y)}{\sigma_Y^2} + 2\frac{Cov(X, Y)}{\sigma_X \sigma_Y} \\
&= 1 + 1 + 2\rho.
\end{aligned}$$

Hence $\rho \geq -1$.

A similar argument shows that

$$0 \leq Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2 - 2\rho,$$

which shows that $\rho \leq 1$.

Suppose $\rho = 1$. Then

$$Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0,$$

which implies that

$$P\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = c\right) = 1$$

for some constant c. This is equivalent to $P(Y = a + bX) = 1$ for some constants a and b. Similar argument for $\rho = -1$.

Example F –Bivariate Normal Distribution

Let X and Y follow a bivariate normal distribution. Show that the covariance of X and Y is $\rho\sigma_X\sigma_Y$.

. . . .

The covariance is

$$Cov(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x,y)\ dxdy.$$

By the change of variable $u = (x-\mu_X)/\sigma_X$ and $v = (y-\mu_Y)/\sigma_Y$, we find that

$$Cov(X, Y) = \frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}uv\exp\left[-\frac{1}{2(1-\rho^2)}(u^2+v^2-2\rho uv)\right]dudv.$$

By completing the square,

$$u^2+v^2-2\rho uv = (u-\rho v)^2 + v^2(1-\rho^2),$$

Example F con't

we find that the expression becomes

$$\frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} v e^{-v^2/2} \left(\int_{-\infty}^{\infty} u \exp\left[-\frac{1}{2(1-\rho^2)} (u-\rho v)^2 \right] du \right) dv.$$

The inner integral is the mean of a random variable with distribution $N[\rho v, (1-\rho^2)]$, i.e.,

$$\frac{1}{\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} u \exp\left[-\frac{1}{2(1-\rho^2)} (u-\rho v)^2 \right] du = \rho v$$

Hence

$$Cov(X, Y) = \frac{\rho \sigma_X \sigma_Y}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v^2 e^{-v^2/2} dv = \rho \sigma_X \sigma_Y.$$

(Again we used the fact that $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v^2 e^{-v^2/2} dv = 1$, which follows from Gamma distribution.)