

Covariance of two Random Variables

The r^{th} & s^{th} product moment about means of rvs Y_1 & Y_2 , denoted by $\mu_{r,s}$ is expected value of $(Y_1 - \mu_1)^r (Y_2 - \mu_2)^s$ when Y_1 & Y_2 are discrete

$$\begin{aligned}\mu_{r,s} &= E[(Y_1 - \mu_1)^r (Y_2 - \mu_2)^s] \\ &= \sum_{Y_1} \sum_{Y_2} (y_1 - \mu_1)^r (y_2 - \mu_2)^s f(y_1, y_2)\end{aligned}$$

when Y_1 & Y_2 are continuous:

$$\mu_{r,s} = E[(Y_1 - \mu_1)^r (Y_2 - \mu_2)^s] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y_1 - \mu_1)^r (y_2 - \mu_2)^s f(y_1, y_2) dy_1 dy_2$$

$\mu_{1,1} \rightarrow$ Covariance of Y_1 & Y_2 , σ_{Y_1, Y_2} , $\text{cov}(Y_1, Y_2)$

If there is a high prob that large values of Y_1 will go large values of Y_2 & small values of Y_1 with small values of $Y_2 \rightarrow \text{cov}(Y_1, Y_2) > 0$

If there is a high prob that large values of Y_1 will go small values of Y_2 & vice versa
 $\text{cov}(Y_1, Y_2) < 0$

Covariance measures relationship, linear dependence or association between Y_1 & Y_2

$$\text{cov}(Y_1, Y_2) = \sigma_{Y_1, Y_2} = \mu_{1,1} - \mu_1 \mu_2 = E(Y_1, Y_2) - E(Y_1)E(Y_2)$$

If Y_1 & Y_2 are independent $E(Y_1, Y_2) = E(Y_1)E(Y_2)$
therefore $\text{cov}(Y_1, Y_2) = 0$

If $\text{cov}(Y_1, Y_2) = 0 \rightarrow$ does not imply Y_1 & Y_2 are independent

If Y_1 & Y_2 are completely dependent, $Y_1 = Y_2$,

$$\text{cov}(Y_1, Y_2) = \sigma_{Y_1, Y_2} = \sigma_1 \sigma_2$$

Correlation coefficient $\Rightarrow$ measure of dependence
 Y_1 & Y_2

$$\rho = \frac{\sigma_{Y_1, Y_2}}{\sigma_1 \sigma_2} = \frac{\text{cov}(Y_1, Y_2)}{\sigma_1 \sigma_2} \Rightarrow -1 \leq \rho \leq 1$$

$$\left. \begin{array}{l} Y_1 \uparrow \quad Y_2 \uparrow \\ Y_1 \downarrow \quad Y_2 \downarrow \end{array} \right\} \rho > 0 \Rightarrow 0 < \rho < 1$$

$$\left. \begin{array}{l} Y_1 \uparrow \quad Y_2 \downarrow \\ Y_1 \downarrow \quad Y_2 \uparrow \end{array} \right\} \rho < 0 \quad -1 < \rho < 0$$

$\rho \Rightarrow$ no units of measurement problem

$\rho \Rightarrow$ not imply causality

If $\rho = 0 \Rightarrow$ variables may or may not be independent

$$\underline{\text{Ex:}} \quad f(Y_1, Y_2) = \begin{cases} \frac{1}{4} (2Y_1 + Y_2), & 0 < Y_1 < 1 \quad 0 < Y_2 < 2 \\ 0 & \text{elsewhere} \end{cases}$$

find $\text{cov}(Y_1, Y_2)$

$$E(Y_1) = \frac{1}{4} \int_0^1 \int_0^2 Y_1 \cdot (2Y_1 + Y_2) dY_2 dY_1$$

$$= \frac{1}{4} \int_0^1 (4Y_1^2 + 2Y_1^2) \cdot dY_1 = \frac{1}{4} \left(\frac{4}{3} + 1 \right) = 5/12$$

$$E(Y_2) = \frac{1}{4} \int_0^2 \int_0^1 y_2 (2y_1 + y_2) \cdot dy_1 dy_2 = \frac{1}{4} \int_0^2 (y_2 + y_2^2) \cdot dy_2$$

$$= \frac{1}{4} \left(2 + \frac{8}{3} \right) = \frac{14}{12}$$

$$E(Y_1, Y_2) = \frac{1}{4} \int_0^1 \int_0^2 (2y_1^2 y_2 + y_1 y_2^2) dy_2 dy_1 = \frac{1}{4} \int_0^1 (4y_1^2 + \frac{8}{3} y_1) dy_1$$

$$= \frac{1}{4} \left(\frac{4}{3} + \frac{4}{3} \right) = \frac{2}{3}$$

$$\text{cov}(Y_1, Y_2) = E(Y_1, Y_2) - E(Y_1)E(Y_2) = \frac{2}{3} - \left(\frac{7}{12}\right)\left(\frac{14}{12}\right) = -\frac{1}{12}$$

Ex: $f(y_1, y_2) = \begin{cases} \frac{1}{8} (y_1 + y_2) & , 0 < y_1 < 2, 0 < y_2 < 2 \\ 0 & \text{elsewhere} \end{cases}$

Find $\text{cov}(Y_1, Y_2)$

$$E(Y_1) = \frac{1}{8} \int_0^1 \int_0^2 (y_1^2 + y_1 y_2) dy_2 dy_1 = \frac{1}{8} \int_0^1 (2y_1^2 + 2y_1) dy_1 = \frac{1}{8} \left(\frac{2}{3} + 1 \right) = \frac{5}{12}$$

$$E(Y_2) = \frac{1}{8} \int_0^2 \int_0^1 (y_1 y_2 + y_2^2) \cdot dy_1 dy_2 = \frac{1}{8} \int_0^2 \left(\frac{1}{2} y_2 + y_2^2 \right) dy_2 = \frac{1}{8} \left(1 + \frac{8}{3} \right) = \frac{11}{12}$$

$$E(Y_1, Y_2) = \frac{1}{8} \int_0^1 \int_0^2 (y_1^2 y_2 + y_1 y_2^2) dy_2 dy_1 = \frac{1}{8} \int_0^1 (2y_1^2 + \frac{8}{3} y_1) dy_1$$

$$\frac{1}{8} \left(\frac{2}{3} + \frac{4}{3} \right) = \frac{2}{8}$$

$$\text{cov}(Y_1, Y_2) = \frac{2}{8} - \left(\frac{5}{12}\right)\left(\frac{11}{12}\right) = -\frac{1}{8}$$

Ex: $f(y_1, y_2) = \begin{cases} \frac{1}{216} (2y_1 + y_2) & 2 \leq y_1 \leq 6 \text{ and } 0 \leq y_2 \leq 5 \\ \text{elsewhere} \end{cases}$

find $\text{cov}(Y_1, Y_2)$ & ρ

$$E(Y_1) = \frac{1}{216} \int_2^6 \int_0^5 y_1 (2y_1 + y_2) dy_2 dy_1 = \frac{268}{63}$$

$$E(Y_2) = \frac{1}{216} \int_2^6 \int_0^5 y_2 (2y_1 + y_2) dy_1 dy_2 = \frac{170}{63}$$

$$E(Y_1 Y_2) = \frac{1}{216} \int_2^6 \int_0^5 y_1 y_2 (2y_1 + y_2) dy_2 dy_1 = \frac{80}{7}$$

$$E(Y_1^2) = \frac{1}{216} \int_2^6 \int_0^5 y_1^2 (2y_1 + y_2) dy_2 dy_1 = \frac{1220}{63}$$

$$E(Y_2^2) = \frac{1}{216} \int_2^6 \int_0^5 y_2^2 (2y_1 + y_2) dy_1 dy_2 = \frac{1175}{126}$$

$$\text{Var}(Y_1) = E(Y_1^2) - [E(Y_1)]^2 = \frac{1220}{63} - \left(\frac{268}{63}\right)^2 = \frac{5036}{3969}$$

$$\text{Var}(Y_2) = E(Y_2^2) - [E(Y_2)]^2 = \frac{1175}{126} - \left(\frac{170}{63}\right)^2 = \frac{16225}{7938}$$

$$\begin{aligned} \text{Cov}(Y_1, Y_2) &= E(Y_1 Y_2) - E(Y_1) \cdot E(Y_2) \\ &= \frac{80}{7} - \frac{268}{63} \cdot \frac{170}{63} = \frac{-200}{3969} \end{aligned}$$

$$\rho = \frac{\text{Cov}(Y_1, Y_2)}{\sigma_1 \sigma_2} = \frac{-200/3969}{\sqrt{(5036/3969)(16225/7938)}} = -0.03127$$

Expected Value & Variance of Linear Functions of RV's

Let $y_1, y_2, \dots, y_n$ & $x_1, x_2, \dots, x_m$ be rv's with

$$E(y_i) = \mu_i \quad E(x_j) = \mu_j$$

$$i = 1, 2, \dots, n \quad j = 1, 2, \dots, m$$

Define

$$U_1 = \sum_{i=1}^n a_i y_i, \quad U_2 = \sum_{j=1}^m b_j x_j$$

for constant $a_1, a_2, \dots, a_n$ & $b_1, b_2, \dots, b_m$. Then,

$$a) E(U_1) = \sum_{i=1}^n a_i \mu_i$$

$$b) V(U_1) = \sum_{i=1}^n a_i^2 V(y_i) + 2 \sum_i \sum_j a_i a_j \text{Cov}(y_i, y_j) \text{ where double}$$

sums over all pairs (i, j) with $i < j$

$$c) \text{Cov}(U_1, U_2) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(y_i, x_j)$$

If $y_1, y_2, \dots, y_n$ are independent then

$$V(U_1) = \sum_{i=1}^n a_i^2 V(y_i)$$

Ex: Let y_1, y_2, y_3 be independent rv's with $E(y_1) = 4$

$$E(y_2) = 9, \quad E(y_3) = 3.$$

$$\text{Var}(y_1) = 3, \quad \text{Var}(y_2) = 7, \quad \text{Var}(y_3) = 5$$

Define $U_1 = 2y_1 - 3y_2 + 4y_3$ & $U_2 = y_1 + 2y_2 - y_3$

Find mean & variance of U_1 & U_2 & $\text{Cov}(U_1, U_2)$

$$E(U_1) = 2E(y_1) - 3E(y_2) + 4E(y_3) = 2(4) - 3(9) + 4(3) = -7$$

$$\text{Var}(U_1) = 2^2 \text{Var}(y_1) + 3^2 \text{Var}(y_2) + 4^2 \text{Var}(y_3)$$

$$= 4(3) + 9(7) + 16(5) = 155$$

$$E(U_2) = E(Y_1) + 2E(Y_2) - E(Y_3) = 4 + 2(9) - 3 = 19$$

$$\text{Var}(U_2) = \text{Var}(Y_1) + 2^2 \text{Var}(Y_2) + \text{Var}(Y_3) = 3 + 4(7) + 5 = 36$$

$$\begin{aligned} \text{Cov}(U_1, U_2) &= (2)(1) \text{Var}(Y_1) - (3)(2) (\text{Var}(Y_2)) + 4(-1) \text{Var}(Y_3) \\ &= 6 - 42 - 20 = -56 \end{aligned}$$

eski s inay sorusu

$$\text{Ex: } f(y_1, y_2) = \begin{cases} \frac{1}{3} (y_1 + y_2) & 0 < y_1 < 1 \quad 0 < y_2 < 2 \\ 0 & \text{elsewhere} \end{cases}$$

find variance $W = 3Y_1 + 4Y_2 - 5$

$$E(Y_1) = \frac{1}{3} \int_0^1 \int_0^2 y_1^2 + y_1 y_2 \, dy_2 \, dy_1 = 8/9$$

$$E(Y_1^2) = \frac{1}{3} \int_0^1 \int_0^2 y_1^3 + y_1^2 y_2 \, dy_2 \, dy_1 = 7/18$$

$$V(Y_1) = \frac{7}{18} - \left(\frac{8}{9}\right)^2 = \frac{13}{162}$$

$$E(Y_2) = \frac{1}{3} \int_0^2 \int_0^1 y_1 y_2 + y_2^2 \, dy_1 \, dy_2 = 11/9$$

$$E(Y_2^2) = \frac{1}{3} \int_0^2 \int_0^1 y_1 y_2^2 + y_2^3 \, dy_1 \, dy_2 = 16/9$$

$$V(Y_2) = \frac{16}{9} - \left(\frac{11}{9}\right)^2 = \frac{23}{81}$$

$$E(Y_1, Y_2) = \frac{1}{3} \int_0^1 \int_0^2 y_1^2 y_2 + y_1 y_2^2 \, dy_2 \, dy_1 = \frac{2}{3}$$

$$\text{Cov}(Y_1, Y_2) = E(Y_1, Y_2) - E(Y_1)E(Y_2)$$

$$= \frac{2}{3} - \left(\frac{8}{9}\right)\left(\frac{11}{9}\right) = -\frac{1}{81}$$

$$\text{Var}(W) = 3^2 \text{Var}(Y_1) + 4^2 \text{Var}(Y_2) + 2(3)(4) \text{Cov}(Y_1, Y_2)$$

$$= 9\left(\frac{13}{162}\right) + 16\left(\frac{23}{81}\right) + 24\left(-\frac{1}{81}\right) = \frac{805}{162}$$

Conditional Expectations

If Y_1 is a discrete r.v. $f(y_1, y_2)$ is value of cont. prob. distr. of Y_1 given $Y_2 = y_2$ conditional expectation of $g(Y_1)$ is

$$E(g(Y_1) | Y_2 = y_2) = \sum_{y_1} g(y_1) \cdot p(y_1 | y_2)$$

If Y_1 is a cont. r.v. $f(y_1 | y_2)$ is value of cond. prob. density of Y_1 given $Y_2 = y_2$ conditional expectation of $g(Y_1)$ is $E(g(Y_1) | Y_2 = y_2) = \int_{-\infty}^{\infty} g(y_1) f(y_1 | y_2) \cdot dy_1$

Let $g(Y_1) = Y_1$, conditional mean of Y_1 given $Y_2 = y_2 \Rightarrow \mu_{Y_1|Y_2} = E(Y_1 | Y_2) \Rightarrow E(Y_1) = E[E(Y_1 | Y_2)]$

$$\begin{aligned} \sigma_{Y_1|Y_2}^2 &= E[(Y_1 - \mu_{Y_1|Y_2})^2 | Y_2] = E(Y_1^2 | Y_2) - \mu_{Y_1|Y_2}^2 = \\ &= E[Y_1^2 | Y_2] - [E(Y_1 | Y_2)]^2 \\ E(Y_1) &= E[E(Y_1 | Y_2)] \end{aligned}$$

Ex:

Find conditional expectation of Y_2 given $Y_1 = 2$ given $f(y_1, y_2) = \begin{cases} \frac{1}{42}(2y_1 + y_2) & 0 < y_1 < 2 \text{ and } 0 < y_2 < 3 \\ 0 & \text{elsewhere} \end{cases}$

$$f(y_2 | 2) = \frac{f(y_1, y_2)}{f(y_1)} = \frac{(2y_1 + y_2)/42}{22/42} = \frac{4 + y_2}{22}$$

$$E[Y_2 | Y_1 = 2] = \sum_{y_2} y_2 \left(\frac{4 + y_2}{22} \right) = 0\left(\frac{4}{22}\right) + 1\left(\frac{5}{22}\right) + 2\left(\frac{6}{22}\right) + 3\left(\frac{7}{22}\right) = \frac{19}{11}$$

Ex: Joint Pdf = $f(y_1, y_2)$ is

$$f(y_1, y_2) = \begin{cases} 8y_1 y_2 & 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

Find

a) cond expect of y_2 given y_1

b) " " " " " "

c) " variance of y_2 given y_1

$$f(y_1) = \int_0^{y_1} 8y_1 y_2 dy_2 = 4y_1^3 \text{ for } 0 \leq y_1 \leq 1$$

$$h(y_2) = \int_{y_1=y_2}^1 8y_1 y_2 dy_1 = 4y_2(1-y_2^2) \quad 0 \leq y_2 \leq 2$$

$$f(y_1 | y_2) = \frac{f(y_1, y_2)}{h(y_2)} = \frac{8y_1 y_2}{4y_2(1-y_2^2)} = \frac{2y_1}{1-y_2^2} \text{ for } y_2 \leq y_1 \leq 1$$

$$f(y_2 | y_1) = \frac{f(y_1, y_2)}{g(y_1)} = \frac{8y_1 y_2}{4y_1^3} = \frac{8y_1 y_2}{4y_1^3} = \frac{2y_2}{y_1^2} \quad 0 \leq y_2 \leq y_1$$

$$a) E(y_2 | y_1 = y_1) = \int_{-\infty}^{\infty} y_2 f(y_2 | y_1) dy_2 = \int_0^{y_1} y_2 \left(\frac{2y_2}{y_1^2} \right) dy_2$$

$$b) E(y_1 | y_2 = y_2) = \int_{-\infty}^{\infty} y_1 f(y_1 | y_2) dy_1 = \int_{y_2}^1 y_1 \left(\frac{2y_1}{1-y_2^2} \right) dy_1$$

$$= \frac{2(1-y_2^3)}{3(1-y_2^2)}$$

$$c) E[(y_1 - \mu_{y_2}) | y_1 = y_1] = \int_{-\infty}^{\infty} (y_2 - \mu_{y_2})^2 f(y_2 | y_1) dy_2$$

$$= \int_0^{y_1} \left(y_2 - \frac{2y_1}{3} \right)^2 \left(\frac{2y_2}{y_1^2} \right) dy_2$$

Multinomial Prob. Distribution

Where # of possible outcomes per trial is more than two
Generalization of binomial experiment

Properties

- 1) Experiment consists of N independent trials
- 2) outcome of each trial falls into one of k classes or cells
- 3) Prob. that outcome of a single trial falls into cell i , is p_i where $i = 1, 2, 3, \dots, k$
it remains same from trial to trial, $p_1 + p_2 + \dots + p_k = 1$
- 4) trials are independent
- 5) R.V.'s $y_1, y_2, y_3, \dots, y_k$ equal # of trial from which outcome fall into cell i
 $y_1 + y_2 + \dots + y_k = n$

Assume that $p_1, p_2, \dots, p_k$ are such that $\sum_{i=1}^k p_i = 1$ & $p_i > 0$,
 $i = 1, 2, \dots, k$

R.V.'s $y_1, y_2, \dots, y_k$ are said to have multinomial distr. with parameters

$n, p_1, p_2, \dots, p_k$ if joint prob. function of $y_1, y_2, \dots, y_k$ is
$$p(y_1, y_2, \dots, y_k) = \frac{n!}{y_1! y_2! \dots y_k!} p_1^{y_1} p_2^{y_2} \dots p_k^{y_k} \quad \text{where for each}$$

$i, y_i = 0, 1, 2, \dots, n, i=1, y_i = n$

If $y_1, y_2, \dots, y_k$ have a multinomial distr. with parameters
 $n, p_1, p_2, \dots, p_k$

- 1- $E(y_i) = np_i$, $V(y_i) = np_i q_i$ where $q_i = 1 - p_i$
- 2- $\text{cov}(y_s, y_t) = -np_s p_t$, where $s \neq t$

Ex:

A large lot of manufactured items contains 7.10 with exactly one defect, 8.7. with more than one defect & remainder with no defects

10 items were randomly selected from this lot for sale

Y_1 denotes number of items with exactly one defect

Y_2 " " " more than one defect

$Y_3 \rightarrow$ no defect

Repair costs are $Y_1 + 3Y_2$

find mean & variance of repair costs

$$E(Y_1) = 10 \cdot (0.1) = 1$$

$$E(Y_2) = 10 \cdot (0.05) = 0.5$$

$$V(Y_1) = 10 \cdot (0.1) (0.9) = 0.9$$

$$V(Y_2) = 10 \cdot (0.05) (0.95) = 0.475$$

$$E(Y_1 + 3Y_2) = E(Y_1) + 3E(Y_2) = 1 + 3(0.5) = 2.5$$

$$\begin{aligned} V(Y_1 + 3Y_2) &= V(Y_1) + 3^2 V(Y_2) + 2 \cdot 1 \cdot 3 \cdot \text{Cov}(Y_1, Y_2) \\ &= 0.9 + 9(0.475) + 6(-0.05) = 4.875 \end{aligned}$$

Bivariate Normal Distribution

Two rvs X & Y are said to have bivariate normal distr. if $aX + bY$ has a normal distr. for all $a, b \in \mathbb{R}$

Then two rvs X & Y are said to have standard bivariate normal distribution if their joint pdf is

$$f(x, y) = \frac{1}{2\pi\sqrt{1-p^2}} \exp\left(-\frac{1}{2(1-p^2)}(x^2 - 2pxy + y^2)\right)$$

where $-1 \leq p \leq 1$