Recall $Var(X) = \sigma_2 = E[(X-m_1)^2]$, where $m_1 = E(X)$.

3.6. Covariance

Ex 1. Let $E(X) = \mu_1$, $E(Y) = \mu_2$.

a) Confirm

$$\boxed{Var(X+Y) = Var(X) + Var(Y) + 2E[(X-\mu_1)(Y-\mu_2)]} \quad (1)$$

Answer. $E(X+Y) = \mu_1 + \mu_2$

$X+Y-E(X+Y) = X+Y-\mu_1-\mu_2 = (X-\mu_1) + (Y-\mu_2)$

$Var(X+Y) = E[((X-\mu_1) + (Y-\mu_2))^2] = E[(X-\mu_1)^2] + E[(Y-\mu_2)^2]$

$+ 2E(X-\mu_1)(Y-\mu_2) = Var(X) + Var(Y) + 2E[(X-\mu_1)(Y-\mu_2)]$.

b) Find $Var(X+Y)$, assuming X, Y are independent.

Answer. $\boxed{Var(X+Y) = Var(X) + Var(Y)}$, because

$$E[(X-\mu_1)(Y-\mu_2)] \underset{X,Y \text{ are indep.}}{=} E(X-\mu_1)\,E(Y-\mu_2) = 0$$

Def. Covariance of X, Y is the number

$Cov(X,Y) = E[(X-\mu_1)(Y-\mu_2)]$, where $\mu_1 = E(X)$, $\mu_2 = E(Y)$

Note 1. $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$

2. $Cov(X,Y) = 0$ if X, Y are independent. In that case,

$Var(X+Y) = Var(X) + Var(Y)$

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

4. $\mathrm{Cov}\left(\sum_{i=1}^n a_i X_i, Y\right) = \sum_{i=1}^n a_i \mathrm{Cov}(X_i, Y)$

Why 4.? Let $E(Y) = \mu$, $E(X_i) = \mu_i$. Then $E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i \mu_i$,

$$\sum_{i=1}^n a_i X_i - E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i (X_i - \mu_i)$$

$$\mathrm{Cov}\left(\sum_{i=1}^n a_i X_i, Y\right) = E\left[\sum_{i=1}^n a_i (X_i - \mu_i)(Y-\mu)\right] = \sum_{i=1}^n a_i \mathrm{Cov}(X_i, Y).$$

5. (Consequence of 4)

$$\mathrm{Cov}\left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j\right) = \sum_{i,j} a_i b_j \mathrm{Cov}(X_i, Y_j)$$

Variance covariance expansion

$$\mathrm{Var}\left(\sum_{i=1}^n a_i X_i\right) = \mathrm{Cov}\left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^n a_j X_j\right)$$

$$= \sum_{i,j=1}^n a_i a_j \mathrm{Cov}(X_i, X_j) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X_i) + \sum_{i\neq j} a_i a_j \mathrm{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^n a_i^2 + 2\sum_{i<j} a_i a_j \mathrm{Cov}(X_i, X_j).$$

In particular, $\mathrm{Var}(aX+bY) = a^2 \mathrm{Var}(X) + b^2 \mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y)$.

$\mathrm{Var}(aX+bY) = a^2 \mathrm{Var}(X) + b^2 \mathrm{Var}(Y)$. if $X, Y$ are indep.

Def. $X, Y$ are uncorrelated if $\mathrm{Cov}(X,Y) = 0$.

If $X, Y$ are independent, then they are uncorrelated.

Meaning of $\mathrm{Cov}(X,Y) = E[(X-\mu_1)(Y-\mu_2)]$ ($\mu_1 = E(X)$, $\mu_2 = E(Y)$):

$\mathrm{Cov}(X,Y) > 0$ if $X-\mu_1$ and $Y-\mu_2$ are predominantly of the same sign.

$\mathrm{Cov}(X,Y) < 0$ if " ——— " of different sign

$\mathrm{Cov}(X,Y)=0$ if $X,Y$ uncorrelated

Def. Let $\sigma_1=\sqrt{\mathrm{Var}(X)}$, $\sigma_2\ \sqrt{\mathrm{Var}(Y)}$.

Correlation (correlation coefficient) of $X,Y$ is the number

$$\rho=\rho_{X,Y}=\frac{\mathrm{Cov}(X,Y)}{\sigma_1\sigma_2}$$

Basic properties 1. $\mathrm{Cov}(X,Y)=\rho\sigma_1\sigma_2$

2. Always $-1\le\rho\le 1$, equivalently, $|\rho|\le 1$

Ex 1. Let $\mathrm{Var}(X)=\sigma_1^2$, $\mathrm{Var}(Y)=\sigma_2^2$ with correlation $\rho=\rho(X,Y)$. Show that

a) $Y=\rho\frac{\sigma_2}{\sigma_1}X+Z$ with $\mathrm{Cov}(X,Z)=0$.

Such a representation is unique: if $Y=aX+\tilde{Z}$ with $\mathrm{Cov}(X,\tilde{Z})=0$, then $\tilde{Z}=Z$, $a=\rho\frac{\sigma_2}{\sigma_1}$.

b) $\mathrm{Var}(Z)=\sigma_2^2(1-\rho^2)$

Answer a) Since $Z=Y-\rho\frac{\sigma_2}{\sigma_1}X$,

$$\mathrm{Cov}(X,Z)=\mathrm{Cov}\left(X,Y-\rho\frac{\sigma_2}{\sigma_1}X\right)=\mathrm{Cov}(X,Y)-\rho\frac{\sigma_2}{\sigma_1}\mathrm{Var}(X)=$$

$$=\rho\sigma_1\sigma_2-\rho\frac{\sigma_2}{\sigma_1}\sigma_1^2=0.$$

b) $\mathrm{Var}(Z)=\mathrm{Var}(Y)+\cancel{\rho^2\frac{\sigma_2^2}{\sigma_1^2}\sigma_1^2}-2\rho\frac{\sigma_2}{\cancel{\sigma_1}}\rho\cancel{\sigma_1}\sigma_2=\sigma_2^2(1-\rho^2)$.

Comments on Ex 1:

1. $\mathrm{Var}(Z)=\sigma_2^2(1-\rho^2)\ge 0\implies\rho^2\le 1$, equivalently, $|\rho|\le 1$.

2. "Extreme coloration" is when $\rho^2=1$, $\rho=\pm 1$.

In this case, $\mathrm{Var}(Z) = \sigma_2^2(1-\rho^2) = 0 \Rightarrow Z = c$ (constant), and

$Y = \pm \frac{\sigma_2}{\sigma_1} X + c$ : $Y$ is linear function of $X$.

3. The roles of $X$ and $Y$ can be reversed:

$X = \rho \frac{\sigma_1}{\sigma_2} Y + V$ with $\mathrm{Cov}(V, Y) = 0$, and $\mathrm{Var}(V) = \sigma_1^2(1-\rho^2)$.