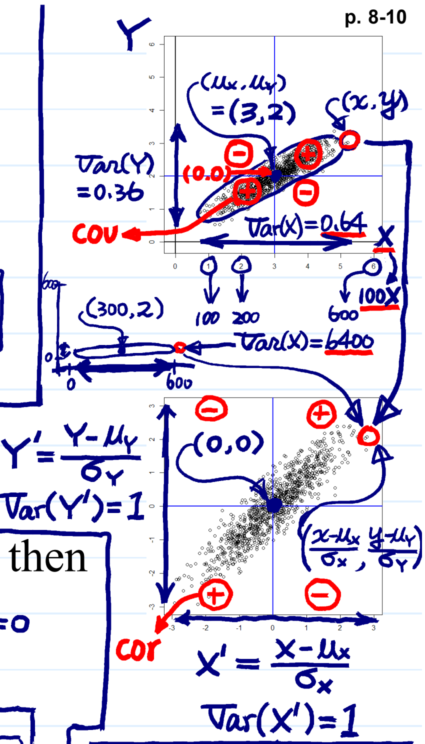
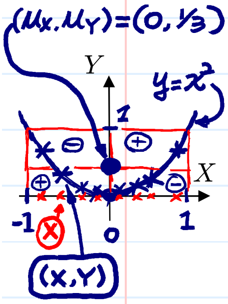
- Corollary. If X and Y are independent, then $Cov(X, Y)=0$, i.e., X and Y are uncorrelated. $\rho_{XY}=0$

Proof. When X, Y are independent,

$$E(XY)=E(X)E(Y)=\mu_X\mu_Y.$$

by Corollary in LNp.8-6





□ However, the converse statement is *not* necessarily true. → "uncorrelated" is a weaker condition than "indep.".

(e.g., let X~Uniform(−1, 1) and $Y=X^2$, then

$Cov(X, Y)=0$, ← $E(XY)=E(X^3)=\int_{-1}^{1} x^3 \frac{1}{2}dx=0$
$E(X)E(Y)=0\cdot(1/3)=0$

but X and Y are *not* independent).

Q: What's the dist. of $Y|X=x$?
$P(Y=y|X=x)=\begin{cases}1, & \text{if } y=x^2\\ 0, & \text{o.w.}\end{cases}$

Note. Y is a function of X. Zero correlation does not mean No relationship.

Q: Why is cor unit-free?

X', Y':
mean = 0
variance = 1

standardization (標準化, LNp.6-33)

■ Corollary.

$$\rho_{XY} = E\left[\underbrace{\left(\frac{X-\mu_X}{\sigma_X}\right)}_{X'}\underbrace{\left(\frac{Y-\mu_Y}{\sigma_Y}\right)}_{Y'}\right].$$

$E[(X-\mu_X)(Y-\mu_Y)]/\sigma_X\sigma_Y$

Proof. By definition.

➢ Example. If $(X_1, \ldots, X_m) \sim$ Multinomial$(n, m, p_1, \ldots, p_m)$, then

$$Cov(X_i, X_j) = -np_ip_j, \quad \text{for } 1 \le i \ne j \le m.$$

- Because $(X_1, X_2, X_3+\cdots+X_m) \sim$ Multinomial$(n, 3, p_1, p_2, p_3+\cdots+p_m)$, and

$$X_3+\cdots+X_m = n-X_1-X_2, \qquad p_3+\cdots+p_m = 1-p_1-p_2,$$

we have

$0 \le x_1+x_2 \le n$
$0 \le x_1 \le n$
$0 \le x_2 \le n$

$$E(X_1X_2) = \sum x_1x_2 \binom{n}{x_1, x_2, n-x_1-x_2} p_1^{x_1}p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}$$

$$= \sum \cancel{x_1}\cancel{x_2}\frac{n!}{\cancel{x_1}!\cancel{x_2}!(n-x_1-x_2)!}p_1^{x_1}p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}$$

$2 \le x_1+x_2 \le n$
$1 \le x_1 \le n-1$
$1 \le x_2 \le n-1$

(x_1-1) (x_2-1)

$$= n(n-1)p_1p_2\left[\sum \frac{(n-2)!}{(\underbrace{x_1-1}_{y_1})!(\underbrace{x_2-1}_{y_2})!(\underbrace{n-x_1-x_2}_{(n-2)-y_1-y_2})!} \times(p_1)^{x_1-1}(p_2)^{x_2-1}(1-p_1-p_2)^{n-x_1-x_2}\right]$$

$0 \le y_1+y_2 \le n-2$
$0 \le y_1 \le n-2$
$0 \le y_2 \le n-2$

joint pmf of Multinomial(n-2, 3, p_1, p_2, $1-p_1-p_2$)

$$= n(n-1)p_1p_2.$$

- WLOG, we can get $E(X_iX_j) = n(n-1)p_ip_j$, for $i \ne j$.

Therefore, $Cov(X_i, X_j) = E(X_iX_j) - E(X_i)E(X_j)$

$$= n(n-1)p_ip_j - (np_i)(np_j) = -np_ip_j.$$

- And, for $i \neq j$, **Note.** $X_1+X_2+X_3+\cdots+X_m=n$ → Why negative

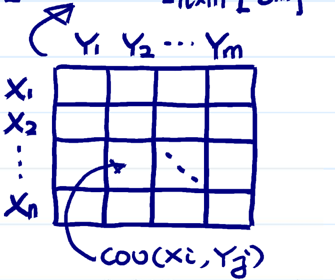
$$\underline{Cor(X_i,X_j)} = \frac{-\cancel{n}\sqrt{p_i}\sqrt{p_j}}{\sqrt{\cancel{np_i}(1-p_i)}\sqrt{\cancel{np_j}(1-p_j)}} = -\sqrt{\underbrace{\frac{p_i\,p_j}{(1-p_i)(1-p_j)}}}.$$

Q: When larger? When smaller?
Hint. $\frac{p}{1-p}$ ↑ as p ↑
Note. $0 \le p_i + p_j \le 1$

- Cov & Cor for Sums of Random Variables
  - Notation. In the following, let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be r.v.'s and $-\infty < a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_m < \infty$ are constants.
  - LNp.8-4 → Recall. $E(a_0+a_1X_1+\cdots+a_nX_n) = a_0+a_1E(X_1)+\cdots+a_nE(X_n)$.
  - Theorem (covariance of two sums).

$$Cov(\underline{a_0} + a_1X_1 + \cdots + a_nX_n, \underline{b_0} + b_1Y_1 + \cdots + b_mY_m)$$
$$= \sum_{i=1}^n \sum_{j=1}^m a_ib_jCov(X_i,Y_j) \rightarrow [a_1 \cdots a_n]\begin{bmatrix} cov(X_i,Y_j) \end{bmatrix}_{n\times m}\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

a_0, b_0 are gone.



Proof. Let $S = a_0+a_1X_1+\cdots+a_nX_n$, and $T = b_0+b_1Y_1+\cdots+b_mY_m$, then

$$S - \overbrace{E(S)}^{a_0+a_1\mu_{X_1}+\cdots+a_n\mu_{X_n}} = \sum_{i=1}^n a_i(X_i - \mu_{X_i}),$$
$$T - \overbrace{E(T)}^{b_0+b_1\mu_{Y_1}+\cdots+b_m\mu_{Y_m}} = \sum_{j=1}^m b_j(Y_j - \mu_{Y_j}),$$
$$[S - E(S)][T - E(T)] = \sum_{i=1}^n \sum_{j=1}^m a_ib_j(X_i - \mu_{X_i})(Y_j - \mu_{Y_j}).$$

Therefore, $Cov(S,T) = E\{[S - E(S)][T - E(T)]\}$

$$= \sum_{i=1}^n \sum_{j=1}^m a_ib_jE[(X_i - \mu_{X_i})(Y_j - \mu_{Y_j})]$$
$$= \sum_{i=1}^n \sum_{j=1}^m a_ib_jCov(X_i,Y_j).$$

mean of sum (LNp.8-4)

cf.

- Theorem (variance of sum). → $Cov(a_0+a_1X_1+\cdots+a_nX_n, a_0+a_1X_1+\cdots+a_nX_n)$

$$Var(\underline{a_0} + \overbrace{a_1X_1}^{Y_1=} + \cdots + \overbrace{a_nX_n}^{Y_n=})$$
$$= \sum_{i=1}^n \sum_{j=1}^n a_ia_jCov(X_i,X_j) \rightarrow [a_1,\cdots,a_n]\begin{bmatrix} cov(X_i,X_j) \end{bmatrix}_{n\times n}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$
$$= \underbrace{\sum_{i=1}^n a_i^2Var(X_i)}_{= Var(Y_i)}$$
$$+2\sum_{1\le i<j\le n} \underbrace{a_ia_jCov(X_i,X_j)}_{= Cov(Y_i,Y_j)}.$$

gone

covariance matrix of $X_1, \cdots, X_n$

$Var(X_i)$

symmetric matrix

Proof. $Cov(X_i, X_i)=Var(X_i)$ and $Cov(X_i, X_j)= Cov(X_j, X_i)$.

cf.

- Corollary. If $X_1, \ldots, X_n$ are uncorrelated, then

$cor(X_i,X_j)=0, \forall i,j, i \neq j$

$$Var(\underbrace{a_0}_{gone} + \overbrace{a_1X_1}^{Y_1} + \cdots + \overbrace{a_nX_n}^{Y_n}) = \sum_{i=1}^n \overbrace{a_i^2Var(X_i)}^{Var(Y_i)}.$$

exchange of Var & Σ

$\frac{1}{n^2}\times n\sigma^2$

- Corollary. If $X_1, \ldots, X_n$ are uncorrelated and

$\frac{X_1+\cdots+X_n}{n} \rightarrow a_1=\cdots=a_n=\frac{1}{n}$

Variance exists

$$Var(X_1)=\cdots = Var(X_n) \equiv \sigma^2 < \infty,$$

Law of Large Number

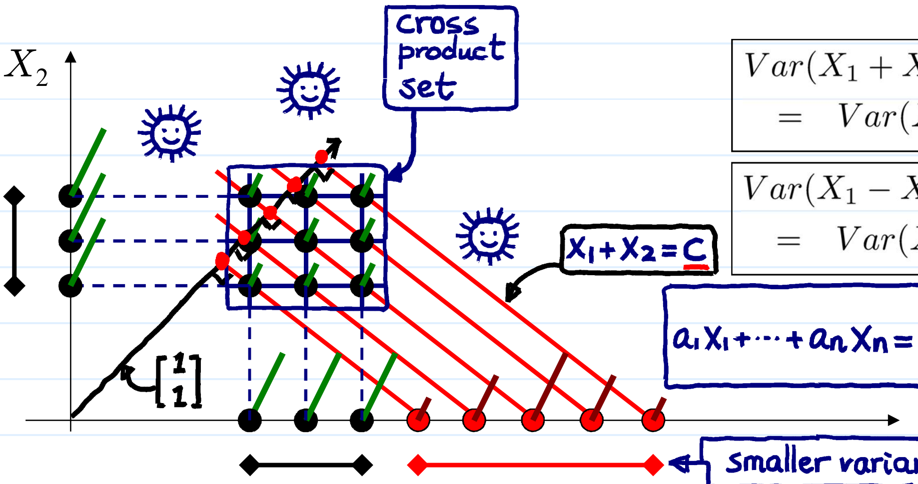
then $Var(\overline{X}_n) = \boxed{\sigma^2/n} \approx 0$ when n ↑ ∞

i.e. $\overline{X}_n \approx C_n$ when n large enough.
a constant → $C_n = ?$ $E(\overline{X}_n) = ?$

Independent

$Cov(X_1, X_2) = 0$

$\sigma_{X_1}^2 = \sigma_{X_2}^2$



$$Var(X_1 + X_2) = Var(X_1) + Var(X_2)$$

$$Var(X_1 - X_2) = Var(X_1) + Var(X_2)$$

$a_1X_1+\cdots+a_nX_n = [a_1,\cdots,a_n]\begin{bmatrix}X_1\\ \vdots\\ X_n\end{bmatrix}$

$\| \underset{\sim}{a}^T$ $\| \underset{\sim}{x}$

$\underset{\sim}{a}^T\underset{\sim}{x} = |\underset{\sim}{a}| \cdot \underline{\frac{\underset{\sim}{a}^T}{|\underset{\sim}{a}|}} \underline{\underset{\sim}{x}}$

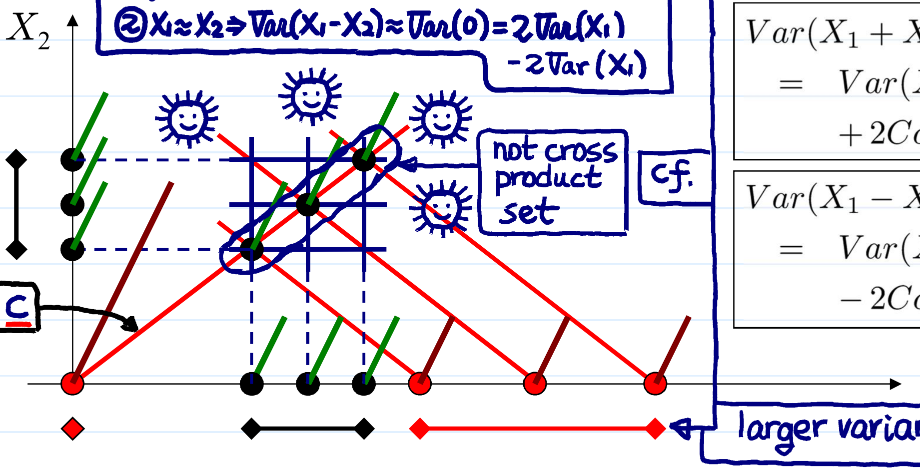
same marginals different joints

$X_1 = X_2$

$Cov(X_1, X_2) > 0$

$Cov(X_1, X_2) = \sigma_{X_1}^2 = \sigma_{X_2}^2$

Why? ① $X_1 \approx X_2 \Rightarrow Var(X_1+X_2) \approx Var(2X_1) = 4Var(X_1)$
② $X_1 \approx X_2 \Rightarrow Var(X_1-X_2) \approx Var(0) = 2Var(X_1) - 2Var(X_1)$



$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)$$

$$Var(X_1 - X_2) = Var(X_1) + Var(X_2) - 2Cov(X_1, X_2)$$