Lec. 34 04/02

7.3 Covariance

Def $\operatorname{Cov}(X,Y) = E\left[(X-E(X))(Y-E(Y))\right]$ ← covariance of X, Y

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\cdot \operatorname{Var}(Y)}}$$ ← correlation coefficient of X, Y.

Remarks: 1) $\operatorname{Cov}(X,Y)$
- >0 : positively correlated (bigger X tends to imply bigger Y)
- <0 : negatively correlated (bigger X tends to imply smaller Y)
- $=0$: uncorrelated (neither).

2) If X, Y is in ft $\Rightarrow \operatorname{Cov}(X,Y)$ is in ft^2.
$\rho_{X,Y}$ has no units, scale-independent.

3) $\operatorname{Cov}(X,X) = E[(X-E(X))^2] = \operatorname{Var}(X)$

$\Rightarrow \rho_{X,X} = 1$ This is the strongest possible positive correlation

4) $\operatorname{Cov}(X,-X) = -\operatorname{Var}(X)$

$\Rightarrow \rho_{X,-X} = -1$ This is the strongest possible negative correlation.

In fact, $\forall X, Y$:

$$-1 \le \rho_{X,Y} \le 1$$

(follows from Cauchy-Schwarz inequality)

Recall: $\operatorname{Var}(X) = E(X^2) - (E(X))^2$. Similarly for $\operatorname{Cov}(X,Y)$.

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

Proof $\mu_X := E(X)$, $\mu_Y := E(Y)$.

$$\operatorname{Cov}(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right] = E\left[XY - \mu_X Y - \mu_Y \cdot X + \mu_X\mu_Y\right]$$

$$= E(XY) - \mu_X \underbrace{E(Y)}_{=\mu_Y} - \mu_Y \underbrace{E(X)}_{=\mu_X} + \mu_X\mu_Y = E(XY) - \mu_X\mu_Y. \quad \text{QED.}$$

Ex (X = # bedrooms, Y = # iPads) – p. 65.

$p(x,y)$:

X \ Y	0	1	2
0	0.05	0.12	0.03
1	0.07	0.1	0.08
2	0.02	0.26	0.27

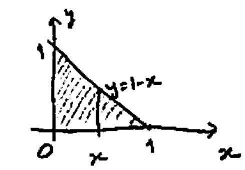
$E[X] = 1.35, \quad E[Y] = 1.24$

$$E[XY] = \sum_{x=0}^{2} \sum_{y=0}^{2} xy\, p(x,y)$$ (discrete version of Prop. on p.87)

$$= 1\cdot 1\cdot 0.1 + 1\cdot 2\cdot 0.08 + 2\cdot 1\cdot 0.26 + 2\cdot 2\cdot 0.27 = 1.86$$

$$Cov(X,Y) = 1.86 - 1.35\times 1.24 = 0.186$$ Positive correlation

Ex $(X,Y) \sim Unif(\Delta)$. $\quad Cov(X,Y) = ?\quad \rho_{X,Y} = ?$



$$f(x,y) = \begin{cases} 2, & (x,y)\in\Delta \\ 0, & \text{—} \end{cases}$$

$$E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)\,dx\,dy = 2\iint_{\Delta} x\,dx\,dy = 2\int_0^1 \left(\int_0^{1-x} x\,dy\right)dx$$

$$= 2\int_0^1 \underbrace{x(1-x)}_{x-x^2}\,dx = 2\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{1}{3}$$

$E[Y] = \frac{1}{3}$ by symmetry.

$$E[XY] = 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy = 2\iint_{\Delta} xy\,dx\,dy = 2\int_0^1 x\underbrace{\left(\int_0^{1-x} y\,dy\right)}_{\frac{y^2}{2}\Big|_0^{1-x} = \frac{(1-x)^2}{2}}dx$$

$$= \int_0^1 \underbrace{x(1-x)^2}_{x-2x^2+x^3}\,dx = 1 - \frac{2}{3} + \frac{1}{4} = \frac{1}{12}$$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{12} - \frac{1}{3}\cdot\frac{1}{3} = -\frac{1}{36}.$$ Negative correlation.

- $\rho_{X,Y} = ?$

$$E[X^2] = 2\iint_{\Delta} x^2\,dx\,dy = 2\int_0^1\left(\int_0^{1-x} x^2\,dy\right)dx = 2\int_0^1 \underbrace{x^2(1-x)}_{x^2-x^3} = 2\left(\frac{1}{3}-\frac{1}{4}\right) = \frac{1}{6}$$

$$Var(X) = E[X^2] - (E[X])^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.$$

$Var(Y) = \frac{1}{18}$ by symmetry.

$$\rho_{X,Y} = \frac{-1/36}{\sqrt{\frac{1}{18}\cdot\frac{1}{18}}} = -\frac{1}{2}.$$ Negative correlation, but weaker than when $Y = -X$.

Linear dependence:

Ex (Strongest correlation). Let $Y = kX + b$. $\quad Cov(X,Y) = ?\quad \rho_{X,Y} = ?$

$E[X] =: \mu_X, \qquad E[Y] = k\mu_X + b.$

- $cov(X,Y) = E[(X-\mu_X)(kX + \cancel{b} - k\mu_X - \cancel{b})] = k\cdot E[(X-\mu_X)^2] = k\cdot Var(X).$

$Var(X) =: \sigma_X, \qquad Var(Y) = k^2\sigma_X$

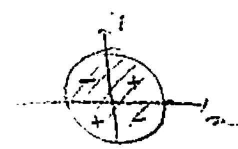
- $$\rho_{X,Y} = \frac{k\cdot Var(X)}{\sqrt{\sigma_X \cdot k^2\sigma_X}} = \frac{k}{|k|} = \begin{cases} 1 & \text{if } k>0 \\ -1 & \text{if } k<0 \end{cases}$$ ← strongest possible correlation.

Prop (Weakest corr.) If X, Y are independent then
$$Cov(X,Y) = 0,$$
ie. X, Y are uncorrelated.

Proof: $Cov(X,Y) = \underbrace{E[XY]}_{=E[X]\cdot E[Y]} - E[X]E[Y] = 0.$ QED

Ex (Uncorrelated does not always imply independent)

Let $(X,Y) \sim Unif(C)$, $\quad C = \{(x,y) : x^2+y^2 \le 1\}$



Then $E[X] = E[Y] = 0$, $\quad E[XY] = \iint_C xy\,dx\,dy = 0$ (Ex. compute, or argue by symmetry)

$\Rightarrow Cov(X,Y) = 0 - 0 = 0.$

However, X, Y are dependent (as we know).

Summary:
- Correlation $\rho_{X,Y}$ is always between $-1, 1$.
- $\rho_{X,Y} = \pm 1$ if there is a linear dependence between X, Y (one can show that this is the only case where $\rho_{X,Y} = \pm 1$)
- $\rho_{X,Y} > 0$ means larger values of X tend to imply larger values of Y
  < 0 — — — — — — smaller — — —
- $\rho_{X,Y} = 0$: "uncorrelated". Happens when X, Y are independent. But not always vice versa.