Lecture 13

Longhai Li, October 21, 2021

Plan:

1. Correlation ρ ✓ § 2.4.
2. Indep. ✓ § 2.5

ρ Correlation Coefficient ; Covariance.

$$Cov(X, Y) = E((X - u_x) \cdot (Y - u_y))$$

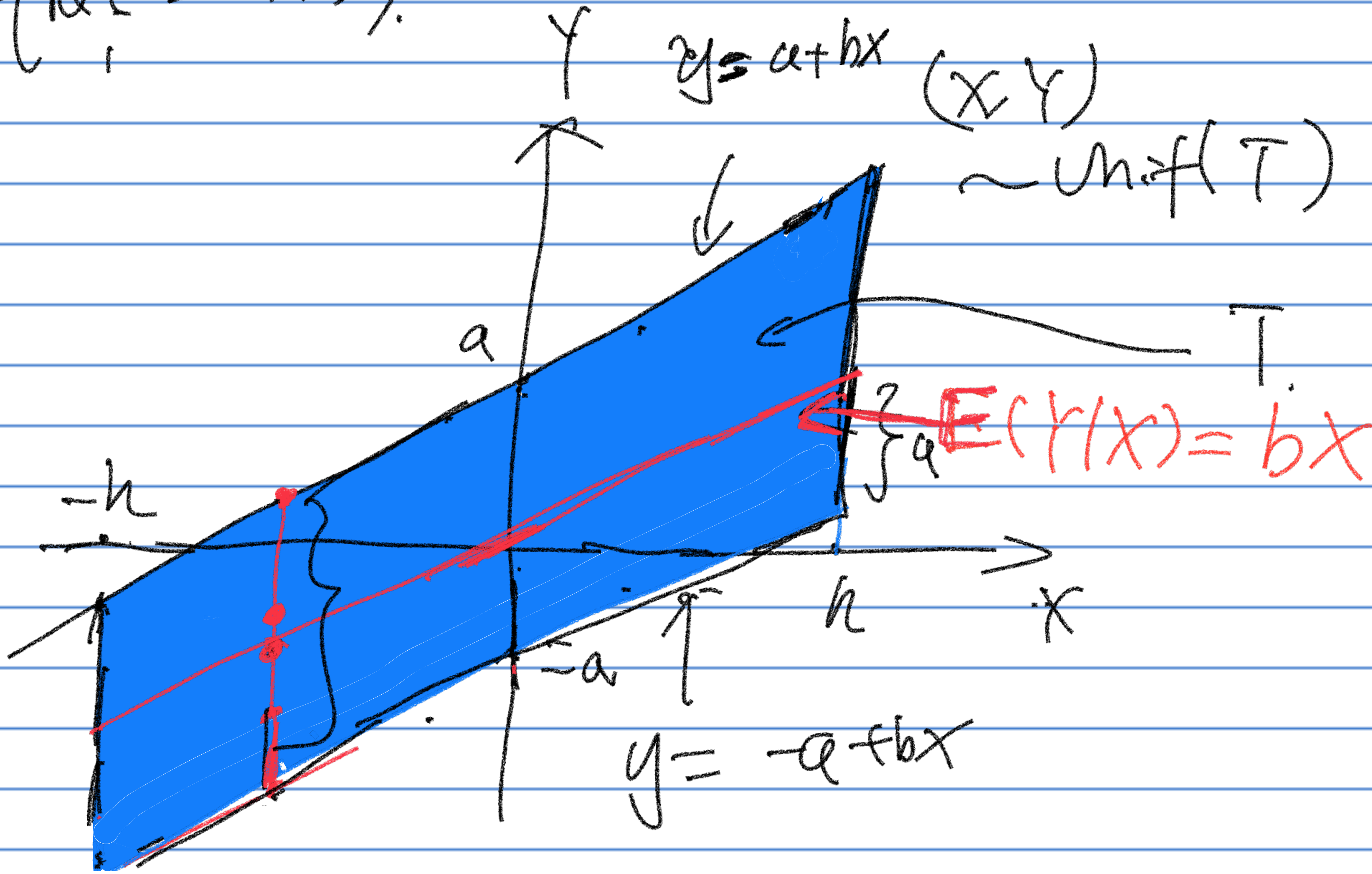
$$= E(XY) - E(X) \cdot E(Y)$$

$u_x = E(X)$. $\quad u_y = E(Y)$

$$Cov(X, X) = V(X)$$

$$\rho = \frac{Cov(X, Y)}{SD(X) \cdot SD(Y)} = \frac{Cov(X, Y)}{\sigma_X \cdot \sigma_Y}$$

Example (2.4.3).



$$f(x,y) = \begin{cases} \dfrac{1}{4ah}, & -a+bx < y < a+bx;\ -h < x < h \\ 0, & \text{o.w.} \end{cases}$$

$$X \sim \text{Unif}((-h, h))$$

$$Y \mid X = x \sim \text{Unif}((-a+bx,\ a+bx))$$

$$E(Y \mid X) = \frac{2bX}{2} = bX$$

$$E(Y) = E(E(Y \mid X)) = b \cdot E(X)$$

$$= 0$$

$E(X) = 0$

$$E(XY) = E\big(E(XY \mid X)\big)$$

$$= E\big(X \cdot E(Y \mid X)\big)$$

$$= E(X \cdot bX)$$

$$= b\,E(X^2)$$

$$\mathrm{Cov}(X,Y) = E(XY) - \underbrace{E(X)}_{=0} \cdot \underbrace{E(Y)}_{=0}$$

$$= b\,E(X^2)$$

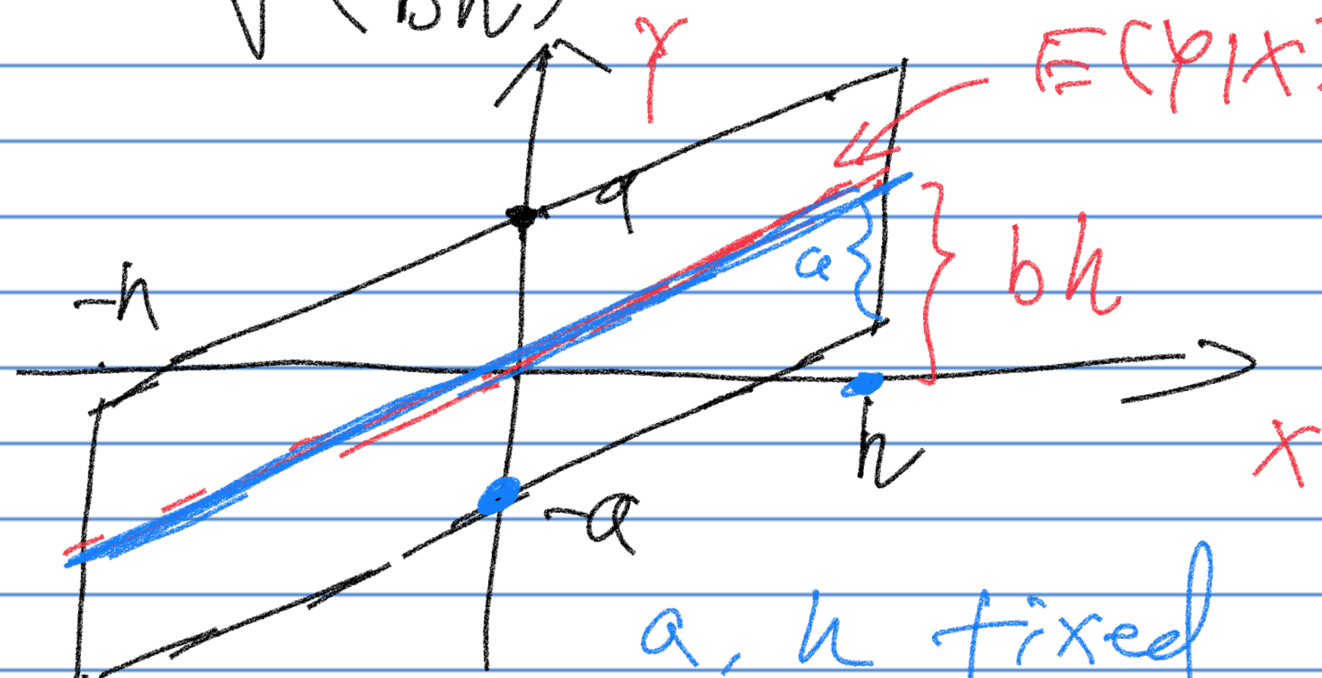
$$\mathrm{Cov}(X,Y) = b \cdot E(X^2) = b \cdot V(X)$$

$$E(X) = 0, \quad V(X) = E(X^2)$$

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)\,V(Y)}} = \frac{b \cdot V(X)}{\sqrt{V(X)\,V(Y)}}$$

$$= b \cdot \frac{\sigma_x}{\sigma_y} = \frac{b \cdot h}{\sqrt{a^2 + b^2 h^2}}$$

$$\rho = \frac{1}{\sqrt{\left(\frac{a}{bh}\right)^2 + 1}}$$



$E(Y|X) = bX$

a, h fixed $b\uparrow$ $\rho\uparrow$

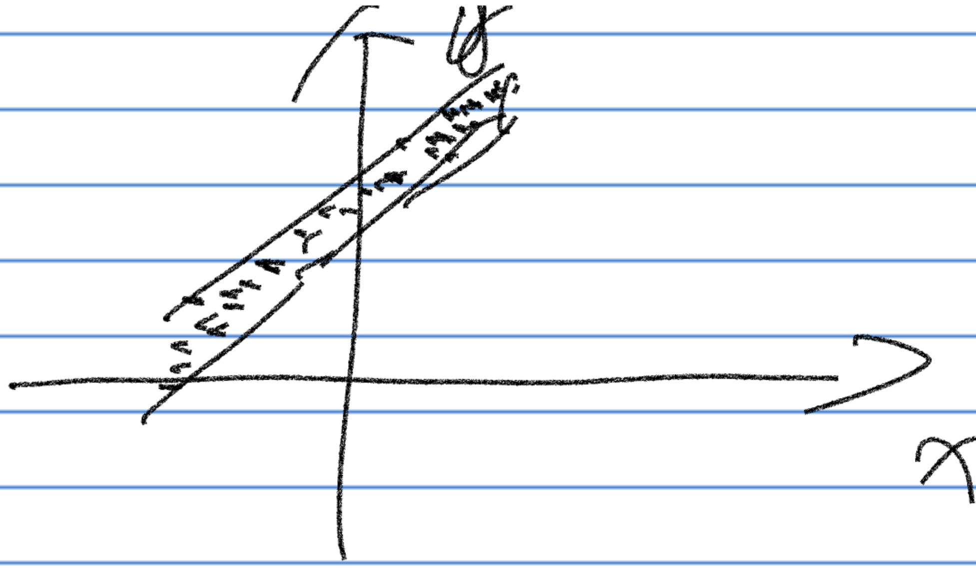
Remarks about ρ:

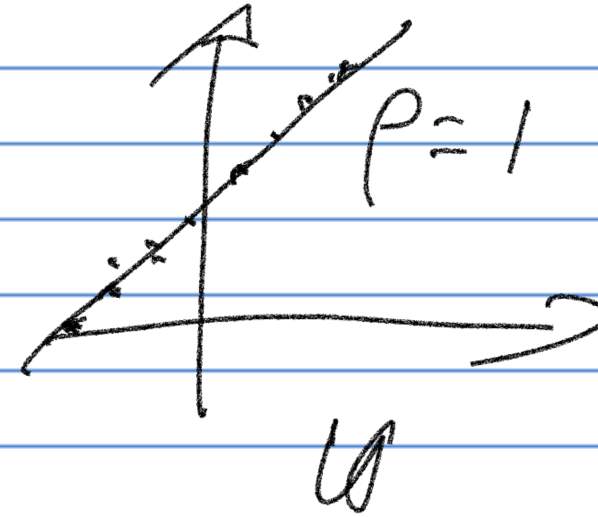
1. $-1 \le \rho \le 1$.

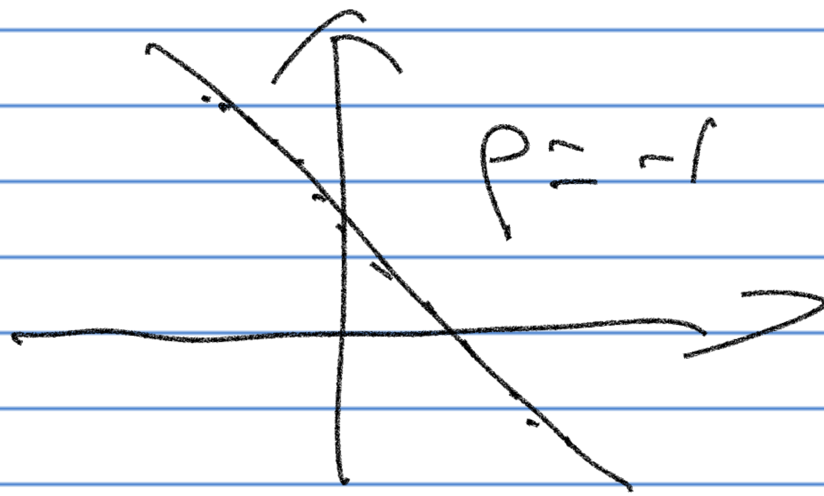
$$|Cov(X,Y)| \le SD(X) \cdot SD(Y)$$

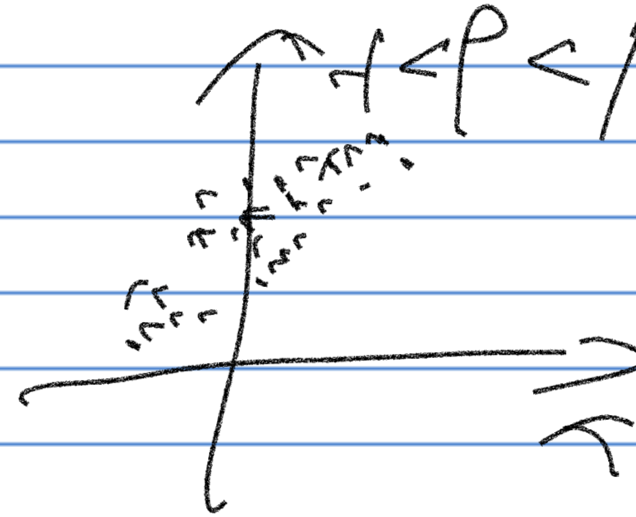
C-S Inequality.

2. $\rho = \pm 1$, perfect linear relationship.


y
x


ρ = 1
y


ρ = -1


-1 < ρ < 1
x

Thm:

If $E(Y|X)$ is a linear function

then $E(Y|X) = \mu_y + \rho \dfrac{\sigma_y}{\sigma_x}(X - \mu_x)$

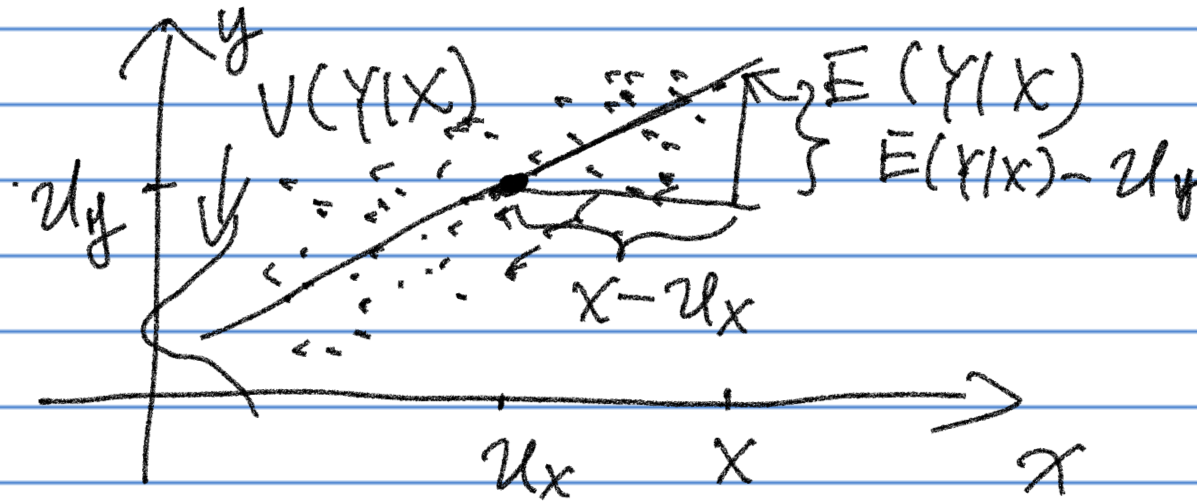
$$\frac{E(Y|X) - \mu_y}{\sigma_y} = \rho \cdot \frac{X - \mu_x}{\sigma_x}$$

$$E(V(Y|X)) = \sigma_y^2 (1 - \rho^2)$$

$$V(E(Y|X)) = \sigma_y^2 \rho^2$$

$$\rho^2 = \frac{V(E(Y|X))}{V(Y)} = \frac{\sigma_y^2 \rho^2}{\sigma_y^2}$$

ρ^2 is the proportion of Variance of Y that is explained by X.



$$\rho = \frac{[E(Y|X) - u_y] / \sigma_y}{(X - u_x) / \sigma_x}$$

$$\frac{\rho \cdot \sigma_y}{\sigma_x} = \frac{E(Y|X) - u_y}{X - u_x} = \beta$$

$$\sigma_Y^2 = V(Y) = \underbrace{\mathbb{E}(V(Y|X))}_{\text{Varian within groups}} + \underbrace{V(\mathbb{E}(Y|X))}_{\text{Variance btw groups.}}$$

$$\rho^2 = \frac{V(\mathbb{E}(Y|X))}{V(Y)}$$

Pf: Assume $E(Y|X) = a + bX$

We will link a, b to $\rho, \mu_x, \mu_y, \sigma_x, \sigma_y$

$$E(Y) = E(E(Y|X)) = a + b\,E(X)$$

$$(\mu_y = a + b\,\mu_x\,,$$

$$E(XY) = E(X\,E(Y|X))$$

$$= E(X\cdot(a + bX))$$

$$= a\,\mu_x + b\,E(X^2)$$

$$E(XY) = a\,\mu_x + b\cdot(\sigma_x^2 + \mu_x^2)$$

b.c. $E(X^2) - [E(X)]^2 = \sigma_x^2$

$$\rho = \frac{E(XY) - EX \cdot EY}{\sigma_x \cdot \sigma_y}$$

$$= \frac{a\,\mu_x + b\cdot(\sigma_x^2 + \mu_x^2) - \mu_x \mu_y}{\sigma_x\ \sigma_y}$$

solving a, b given $\rho, \mu_x, \mu_y, \sigma_x, \sigma_y$:

$$\begin{cases} a = \mu_y - \rho \cdot \frac{\sigma_y}{\sigma_x} \mu_x \\ b = \rho \cdot \frac{\sigma_y}{\sigma_x} \end{cases}$$

- $E(Y|X) = a + bX$

$$= \mu_y + \rho \frac{\sigma_y}{\sigma_x} (X - \mu_x)$$

- $V(E(Y|X)) = \rho^2 \cdot \frac{\sigma_y^2}{\cancel{\sigma_x^2}} \cdot \cancel{V(X)} = \rho^2 \sigma_y^2$

Corrolary:

If $E(Y|X) = a$, where a is a constant

then $\rho_{X,Y} = 0$.

pf: $b = 0$. $b = \rho \cdot \frac{\sigma_y}{\sigma_x}$. $\rho = 0$.

$R^2 = \dfrac{SSR}{SST}$ for ≥ 3 variables

ρ^2 for bivariate variables.

R^2 is a more general concept of ρ^2

Independence

X & Y are indep. if

$$f(x,y) = f_X(x) \cdot f_Y(y) \quad \text{or}$$

$$f(x|y) = f_X(x) \quad \text{or}$$

$$f(y|x) = f_Y(x)$$

Thm: If X & Y are indep.

then $\rho_{X,Y} = 0$. $Cov(X,Y) = 0$.

pf: $f(x,y) = f_X(x) \cdot f_Y(y)$

$$E(X \cdot Y) = \iint x \cdot y \cdot f_X(x) \cdot f_Y(y)\,dx\,dy$$

$$= \int y \cdot f_Y(y)\,dy \cdot \int x \cdot f_X(x)\,dx$$

$$= E(Y) \cdot E(X)$$

$$\mathrm{Cov}(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

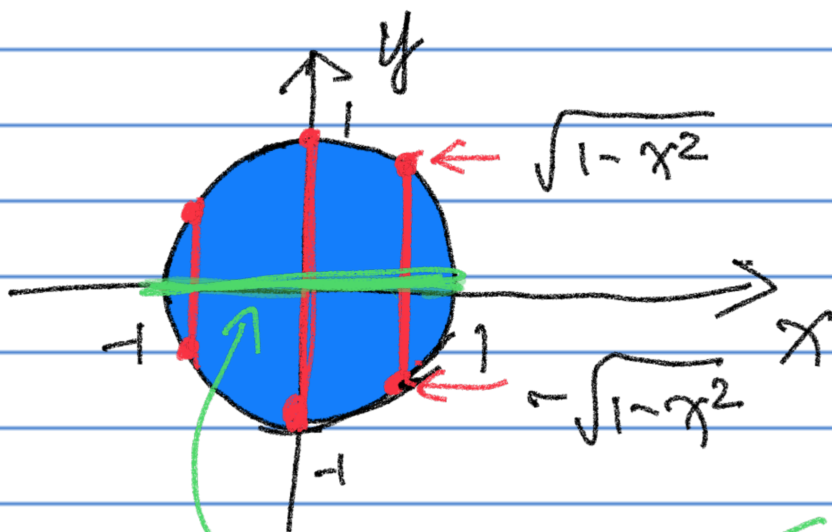
$$= 0$$

$$\rho_{X,Y} = 0$$

Example:

$$(X, Y) \sim \mathrm{Unif}(S^2)$$

$$S^2 = \{(x, y) \mid x^2 + y^2 \le 1\}$$



$$Y|X \sim \mathrm{Unif}((-\sqrt{1-X^2}, \sqrt{1-X^2}])$$

$$f(y|x) \neq f_X(x)$$

X and Y are not indep!

$$\underline{E(Y|X)} = 0 = 0 + 0 \cdot X$$

$$E(XY) = E(X \cdot \underbrace{E(Y|X)}_{=0})$$

$$= E(X \cdot 0) = 0$$

$$Cov(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

$$= 0 - 0 \cdot 0 = 0$$

Alternatively $b = \rho \cdot \frac{\sigma_y}{\sigma_x} = 0 \Rightarrow \rho = 0$

$$\rho = 0$$

Thm: If $X \perp Y$

then $E(g(X) \cdot h(Y)) = E(g(X)) E(h(Y))$.

Thm: If $X \perp Y$

then $F(x, y) = F_X(x) \cdot F_Y(y)$.

Thm: If $X \perp Y$

then $P(X \in A, Y \in B)$

$= P(X \in A) \cdot P(Y \in B)$.

Thm: If $X \perp Y$

then $M_{X,Y}(t_1, t_2)$

$$= E\left(e^{t_1 X + t_2 Y}\right)$$

$$= E\left(e^{t_1 X} \cdot e^{t_2 Y}\right)$$

$$= E\left(e^{t_1 X}\right) \cdot E\left(e^{t_2 Y}\right)$$

$$= M_X(t_1) \cdot M_Y(t_2)$$