Stat 451
11-1-16
1

$$E\left(\left[(X-\mu_X)+v(Y-\mu_Y)\right]^2\right) \geq 0$$

$$\parallel$$

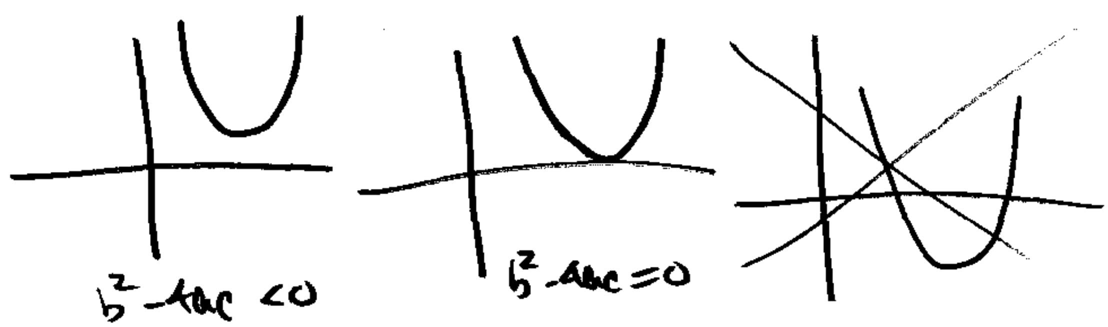
$$f(v)$$

$$f(v) = E\left[(X-\mu_X)^2 + v^2(Y-\mu_Y)^2 + 2v(X-\mu_X)(Y-\mu_Y)\right]$$
$$= E(X-\mu_X)^2 + v^2 E(Y-\mu_Y)^2 + 2v\,E\left((X-\mu_X)(Y-\mu_Y)\right)$$
$$= \sigma_X^2 + v^2\sigma_Y^2 + 2v\,\sigma_{XY}$$
$$= \sigma_Y^2 v^2 + 2\sigma_{XY} v + \sigma_X^2$$

This is a parabola in v, opening up.

2

But $f(v) \geq 0$ for all v



$$b^2-4ac = 4\sigma_{XY}^2 - 4\sigma_Y^2\sigma_X^2 \leq 0$$
$$\sigma_{XY}^2 \leq \sigma_X^2\sigma_Y^2$$
$$\frac{\sigma_{XY}^2}{\sigma_X^2\sigma_Y^2} \leq 1 \qquad \rho^2 \leq 1 \qquad -1 \leq \rho \leq 1$$

③

$b^2 - 4ac = 0 \iff f(v) = 0$ for some v

$\iff \rho^2 = 1$

But $f(v) = E\left[\left((X - \mu_x) + v(Y - \mu_Y)\right)^2\right]$

$$\therefore \left[(X - \mu_x) + v(Y - \mu_Y)\right]^2 = 0$$

$$(X - \mu_x) + v(Y - \mu_Y) = 0$$

$$X = -v(Y - \mu_Y) + \mu_x$$

$\rho = \pm 1$ if and only if one of the random variables is a perfect linear function of the other.

④

Defn: The multinomial distribution

Run n independent trials.

There are r possible outcomes in each trial, with probabilities $p_1, p_2, \ldots, p_r$

$$p_1 + p_2 + \ldots + p_r = 1$$

X_1 = # of type 1 outcomes in n trials

X_2 = # " " 2 " " " "

$\vdots$

X_r = # " " r " " " "

$$X_1 + X_2 + \ldots + X_r = n$$

(5)

$$P(X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_r = x_r)$$

$$= \frac{n!}{x_1!\, x_2! \cdots x_r!} p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r}$$

The random variable

X_i has a binomial distribution

with parameters n and p_i

So $E[X_i] = np_i$ and $V[X_i] = np_i q_i$

(where $q_i = 1 - p_i$)

(6)

The random variable $X_i + X_j$

has a binomial distribution with

parameter n and $p_i + p_j$

So $V[X_i + X_j] = n(p_i + p_j)(1 - p_i - p_j)$

$\parallel$

$V[X_i] + V[X_j] + 2\,\mathrm{Cov}(X_i, X_j)$

⑦

$$n(p_i+p_j)(1-p_i-p_j) = np_iq_i + np_jq_j + 2\,\mathrm{Cov}(X_i,X_j)$$

$$n\left[\cancel{p_i} - \cancel{p_i^2} - p_ip_j + \cancel{p_j} - p_ip_j - \cancel{p_j^2}\right]$$

$$= n\left[\cancel{p_i} - \cancel{p_i^2} + \cancel{p_j} - \cancel{p_j^2}\right] + 2\,\mathrm{Cov}(X_i,X_j)$$

$$-2np_ip_j = 2\,\mathrm{Cov}(X_i,X_j)$$

$$\therefore\ \mathrm{Cov}(X_i,X_j) = -np_ip_j$$

⑧

$$\mathrm{Corr}(X_i,X_j) = \frac{-\cancel{n}p_ip_j}{\sqrt{\cancel{n}p_iq_i\,\cancel{n}p_jq_j}}$$

$$= -\frac{p_ip_j}{\sqrt{p_iq_ip_jq_j}} = -\sqrt{\frac{p_ip_j}{q_iq_j}}$$