Bivariate Normal Distribution

p 81-84 $$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\right.$$

$$\left.\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\Bigg]\right\}.$$

$$-\infty < \mu_x, \mu_y < \infty$$
$$0 < \sigma_x^2, \sigma_y^2$$
$$-1 < \rho < +1$$

marginals $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$

Exercise: show

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left\{-\frac{1}{2}\frac{(x-\mu_x)^2}{\sigma_x^2}\right\}.$$

Exercise: show

$$f_{Y|X}(y|x) = \frac{1}{\sigma^2\sqrt{2\pi(1-\rho^2)}}\,\cdot$$

$$\cdot \exp\left\{\frac{1}{2}\frac{\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right]}{\sigma_y^2(1-\rho^2)}\right\}$$

and $E[Y|X=x] = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)$

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

$$\mathrm{Cov}(X, Y) = \rho\,\sigma_x \sigma_y.$$

$$\mathrm{Corr}(X, Y) = \rho$$

Note: X and Y are independent bivariate normal r.v.'s

if and only if

$\rho = 0$.

<u>Prediction</u>: A r.v. is observed and then based on the observed value, we predict a second r.v. Y

Let $g(X)$ denote the predictor function.

Choose $g(X)$ so it tends to be close to to Y.

<u>Criteria</u> minimum MSE, i.e., minimize

$$E[(Y-g(X))^2].$$

<u>Claim</u>: the best predictor is $g(X)=E[Y|X]$

i.e. $E[(Y-g(X))^2] \geq E[(Y-E[Y|X])^2]$.

<u>Pf</u>: condition on X and then take $E(\cdot)$

$$
\begin{aligned}
& E[(Y-g(X))^2|X] \\
&= E[(Y-E[Y|X]+E[Y|X]-g(X))|X]. \\
&= E[(Y-E[Y|X])^2|X] \\
&\qquad + E[(E[Y|X]-g(X))^2|X]. \\
&\qquad\quad + 2E[(Y-E[Y|X])(E[Y|X]-g(X))|X].
\end{aligned}
$$

note: $E[Y|X] - g(X)$ given X, can be treated as a constant.

Thus

$$E[(Y - E[Y|X])(E[Y|X] - g(X)) | X]$$

$$= (E[Y|X] - g(X))\, E[(Y - E[Y|X]) | X]$$

$$= (E[Y|X] - g(X))\,(E[Y|X] - \underbrace{E[E[Y|X]|X]}_{= E[Y|X]})$$

recall $E[E[Y|X]] = E[Y]$.

$$= 0$$

Therefore:

$$E[(Y - g(X))^2 | X] \geq E[(Y - E[Y|X])^2 | X].$$

finally taking $E[\cdot]$ on both sides

$$E[(Y - g(X))^2] \geq E[(Y - E[Y|X])^2].$$

<u>Exercise</u>: show $E[(X-a)^2]$ is minimized by $a = E[X]$.

It sometimes happens that the joint density of X and Y is not completely known, or $E[Y|X]$ is difficult to calculate. However,

$$\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \sigma_{xy}$$

are known, then we can at least determine the

"best linear predictor"

of Y on X.

Assume $g(X) = a + bX$

Find a and b that minimizes

$$E[(Y - (a+bX))^2]$$

Ans: $E[(Y-(a+bX))^2]$.

$$= E[Y^2 - 2aY - 2bXY + a^2 + 2abX + b^2X^2].$$

$$= E[Y^2] - 2a\,E[Y] - 2b\,E[XY]$$

$$+ a^2 + 2ab\,E[X] + b^2 E[X^2].$$

$$\frac{\partial}{\partial a} E[(Y-(a+bX))^2] = -2E[Y] + 2a + 2bE[X].$$

$$\frac{\partial}{\partial b} E[(Y-(a+bX))^2] = -2E[XY] + 2aE[X] + 2bE[X^2]$$

set each equation equal to zero,
solve for a & b.

<u>Exercise:</u> show

$$b = \frac{E[XY] - E[X]E[Y]}{E[X^2] - (E[X])^2}$$

$$= \frac{Cov(X,Y)}{Var(X)} = \rho \frac{\sigma_y}{\sigma_x}.$$

$$a = E[Y] - bE[X] = E[Y] - \rho \frac{\sigma_y}{\sigma_x} E[X].$$

and $\frac{\partial^2}{\partial a^2} > 0$, $\frac{\partial^2}{\partial b^2} > 0$, so minimum

Therefore, the best linear predictor of Y w.r.t. X is

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

$$= \underbrace{\left[\mu_y + \rho \frac{\sigma_y}{\sigma_x}\mu_x\right]}_{\text{y-intercept}} + \underbrace{\rho \frac{\sigma_y}{\sigma_x}}_{\text{slope}}\, x.$$

MSE

note: For (X, Y) bivariate Normal the best predictor is the best linear predictor $E[Y|X]$ //

The MSE of prediction

$$E[(Y - \mu_y - \rho \frac{\sigma_y}{\sigma_x}(X - \mu))^2].$$

$$= E[(Y - \mu_y)^2] + \rho^2 \frac{\sigma_y^2}{\sigma_x^2} E[(X - \mu_x)^2].$$

$$- 2\rho \frac{\sigma_y}{\sigma_x} E[(Y - \mu_y)(X - \mu_x)].$$

$$= \sigma_y^2 + \rho^2 \sigma_y^2 - 2\rho^2 \sigma_y^2$$

$$= \sigma_y^2 (1 - \rho^2).$$

note: For (X, Y) bivariate Normal The best linear predictor has MSE equal to ~~v.~~ $V(Y|X)$. //

Higher ρ lower prediction error