NTNU
Norwegian University of
Science and Technology

TMA4267 Linear Statistical Models V2014 (3)

Bivariate normal [1.5]

Mette Langaas

To be lectured: January 13, 2014
wiki.math.ntnu.no/emner/tma4267/2014v/start/

The bivariate normal distribution

$$f(x, y) = ce^{-\frac{1}{2}Q(x,y)}$$

$$c = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$Q(x, y) = \frac{1}{(1-\rho^2)}[(\frac{x-\mu_X}{\sigma_X})^2 + (\frac{y-\mu_Y}{\sigma_Y})^2 - 2\rho(\frac{x-\mu_X}{\sigma_X})(\frac{y-\mu_Y}{\sigma_Y})]$$

Computational trick

$$Q(x,y)(1-\rho^2) = [(\frac{y-\mu_Y}{\sigma_Y}) - \rho(\frac{x-\mu_X}{\sigma_X})]^2 + (1-\rho^2)(\frac{x-\mu_X}{\sigma_X})$$

$$f(x,y) = \frac{1}{\sqrt{2\pi}\sigma_X}\exp[-\frac{1}{2}(\frac{x-\mu_X}{\sigma_X})^2]$$

$$\cdot \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_Y}\exp[-\frac{1}{2}(\frac{y-c_X}{\sqrt{1-\rho^2}\sigma_y})^2]$$

$$c_x = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$$

TMA 4267 Linear statistical models L3

The bivariate normal (cont.) [1.5]

Last time:

$$f(x,y) = \underbrace{f_x(x)}_{N(\mu_x,\sigma_x^2)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{1-\rho^2}} \frac{1}{\sigma_y} \cdot \exp\left(-\frac{1}{2} \frac{(y-c_x)^2}{\sigma_y^2(1-\rho^2)}\right)}_{N(c_x,\, \sigma_y^2(1-\rho^2))}$$

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

Remember: conditional distributions

$$f(y|x) = \frac{f(x,y)}{f_X(x)} \Leftrightarrow f(x,y) = f_X(x) \cdot f(y|x)$$

Thus, the conditional distribution of $Y|X=x$ is univariate normal

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

$$= \underbrace{\left(\mu_y - \rho\frac{\sigma_y}{\sigma_x}\mu_x\right)}_{\alpha} + \underbrace{\rho\frac{\sigma_y}{\sigma_x}}_{\beta} \cdot x$$

$E(Y|X=x)$ is a linear function in x.

Compare with a and b from least squares in a simple linear regression of response y as a function of covariate x

$$\hat{\rho} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

$$a = \bar{y} - b\bar{x} \quad \text{and} \quad b = \frac{S_{xy}}{S_{xx}} = \hat{\rho} \cdot \sqrt{\frac{S_{yy}}{S_{xx}}}$$

From $E(Y \mid X = x)$

$$\alpha = \mu_y - \beta \mu_x \qquad \beta = \rho \frac{\sigma_y}{\sigma_x}$$

Relation to Galton's equation:

$$E(Y \mid X=x) = \underset{\substack{\uparrow \\ E(Y)}}{\mu_y} - \rho \frac{\sigma_y}{\sigma_x} \underset{\substack{\uparrow \\ E(X)}}{\mu_x} + \rho \frac{\sigma_y}{\sigma_x} \cdot x$$

$$\frac{E(Y \mid X=x) - E(Y)}{\sigma_y} = \rho \, \frac{x - E(X)}{\sigma_x}$$

Galton said

$$Y - E(Y) = \rho\,(X - E(X))$$

and assumed that $\sigma_x = \sigma_y$

But, what is ρ in the bivariate normal?

Fact: If (X,Y) are bivariate normal $(\mu_x, \sigma_x^2, \mu_y, \sigma_y^2, \rho)$ then $\text{Corr}(X,Y) = \rho$.

$$\rho(X,Y) = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \cdot \sigma_Y}$$

BF p18

$$= \int_{-\infty}^{\infty}\int \left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) f(x,y)\,dx\,dy = \cdots = \rho$$

In the case that $\rho = \pm 1$, then X and Y are exactly linearly related, and we have a one-dimensional solution.

Interpretation:

$$\mathrm{Var}(Y \mid X=x) = (1-\rho^2)\sigma_y^2 \quad \text{and} \quad \mathrm{Var}(Y) = \sigma_y^2$$

the variability that can not be accounted for by x

$\Rightarrow$ $\rho^2\sigma_y^2$ is accounted for by the knowledge of x.

$$\frac{\mathrm{Var}(Y) - \mathrm{Var}(Y \mid X=x)}{\mathrm{Var}(Y)} = \frac{\sigma_y^2 - (1-\rho^2)\sigma_y^2}{\sigma_y^2} = \rho^2$$

proportion of variability in Y accounted for by knowledge of X.

Contours of $f(x,y)$

$$f(x,y) = c \cdot \exp\left(-\frac{1}{2} Q(x,y)\right)$$

When is $f(x,y)$ constant? When $Q(x,y) = d^2$

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

What kind of graphic figure is this?

From previous maths courses:

$x^2+y^2=1$: circle with centre in $(0,0)$ and axes of halflength 1 in x and y direction

$\left(\frac{x}{\sigma_x}\right)^2+\left(\frac{y}{\sigma_y}\right)^2=1$: ellipse in $(0,0)$ with axes of halflength σ_x in x-direction and σ_y in y- — " —.

$\left(\frac{x-\mu_x}{\sigma_x}\right)^2+\left(\frac{y-\mu_y}{\sigma_y}\right)^2=1$: same as above, but now centre in (μ_x,μ_y)

But we have an additional cross term.
Long version E1 F2

Solution: use diagonalization (spectral decomp).

- write $Q(x,y)$ in vector-matrix notation

$$x = \begin{bmatrix} x \\ y \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \quad \Sigma = \begin{bmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \rho \cdot \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix} \leftarrow \text{covariance matrix}$$

$$Q(x) = \underbrace{(x-\mu)^T}_{z^T} \underbrace{\Sigma^{-1}}_{A} \underbrace{(x-\mu)}_{z}$$

- use eigenvector-eigenvalue decomp. $A = P \Lambda P^T$

- use $\underbrace{z^T P}_{w} \underbrace{\Lambda}_{\begin{bmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{bmatrix}} \underbrace{P^T z} = \frac{1}{\lambda_1} w_1^2 + \frac{1}{\lambda_2} w_2^2$

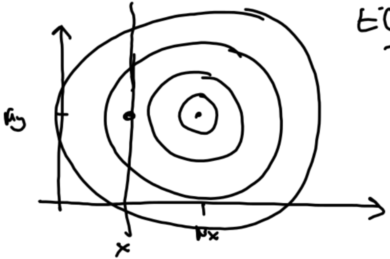
to see that we have an ellipse centered in $\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$ and have axes in the direction of the eigenvectors of Σ, with halflengths $\sqrt{\lambda_i} \cdot d$, when λ_1, λ_2 are the eigenvalues of Σ. $\Rightarrow$ E1.P2

Contours and $E(Y|X=x)$

1) $\sigma_x = \sigma_y$ and $\rho = 0$: $Q(x,y) = \left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2$

equal

$\rightarrow$ circles

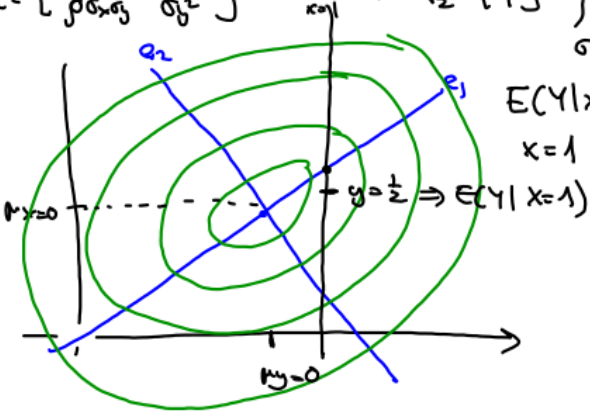


$$E(Y|x) = \mu_y + \underset{\substack{\uparrow \\ 0}}{\rho} \frac{\sigma_x}{\sigma_y}(x-\mu_x)$$

2) $\sigma_x = \sigma_y$ and $\rho = 0.5$ E1P2

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \rho \cdot \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix} \rightarrow e_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad e_2 = \frac{1}{\sqrt{2}}$$

$\sigma_x = \sigma_y$



$E(Y \mid x) = 0.5$

$x = 1$

Findings

- The marginal distributions $f_X(x)$ and $f_Y(y)$ are both normal, and with parameters (μ_X, σ_X^2) and (μ_Y, σ_Y^2), respectively.
- The conditional distribution $f(y|x)$ is also normal with $\mathrm{E}(Y|X = x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$ and $\mathrm{Var}(Y|X = x) = (1 - \rho^2)\sigma_Y^2$.
- $\mathrm{Corr}(X, Y) = \rho$.
- ρ^2 can be explained to be the proportion of variability of Y accounted for by knowledge of X.
- The contours of the bivariate normal are ellipses centered in (μ_X, μ_Y), see Exercise 1, Problem 2.

4 figures

You have 4 figures - they are red, lillac, black and white. Look at the 4 3D figures. Link the four characteristics below with the figure color.
Write the answer on a piece of paper in the figure box. Hand in your answer in the box at the entrance.

1. $\sigma_X = 1$, $\sigma_Y = 1$, $\rho = 0$.
2. $\sigma_X = 1$, $\sigma_Y = 2$, $\rho = 0$.
3. $\sigma_X = 1$, $\sigma_Y = 1$, $\rho = 0.5$.
4. $\sigma_X = 1$, $\sigma_Y = 2$, $\rho = 0.3$.