NTNU
Norwegian University of
Science and Technology

TMA4267 Linear Statistical Models V2014 (4)

Bivariate normal [1.5] (finally: independence) Maximum likelihood [1.6], Sums of squares [1.7]

Mette Langaas

To be lectured: January 14, 2014
wiki.math.ntnu.no/emner/tma4267/2014v/start/

The bivariate normal distribution

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

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

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

The bivariate normal (final part) [L4]

$(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$

random variable

Known: when X and Y are independent RVs

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

$$\rho = \frac{E((X-\mu_x)(Y-\mu_y))}{\sigma_x \sigma_y} = \frac{E(X \cdot Y) - \mu_x \mu_y}{\sigma_x \sigma_y} = 0$$

$$\int_{-\infty}^{\infty}\int x \cdot y \underbrace{f(x,y)}_{f_X(x) \cdot f_Y(y)} \, dx \, dy = \mu_x \cdot \mu_y$$

When we find that $\rho = 0$ we may in general <u>not</u> conclude that X and Y are independent.

Ex: $X \sim N(0,1)$, $Y = X^2$

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}(X,X^2) = E(X \cdot X^2) - \underbrace{E(X)}_{0} \cdot \overbrace{E(X^2)}^{1}$$

$$= E(X^3) - 0 = 0 - 0 = \underline{0}$$

$\mathrm{Var}(X) = E(X^2) - E(X)^2$

$N(0,1)$ symmetrish

X^3 odd $\quad \int_{-\infty}^{\infty} x^3 f(x)\,dx = 0$

But, when X and Y are bivariate normal then

$$\rho = 0 \iff X \text{ and } Y \text{ are independent}$$

We will prove using MGFs.

MGF

Univariate

$$M_X(t) = \mathrm{E}(\exp(tX))$$

Rules:

$$M_{aX}(t) = M_X(at)$$

$$M_{X+a}(t) = e^{at} M_X(t)$$

$$Y = \sum_{i=1}^{n} X_i \text{ then } M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

$$X_i \text{ i.i.d. } M_X(t) \text{ then } M_{\bar{X}}(t) = \prod_{i=1}^{n} M_X(t/n) = [M_X(t/n)]^n$$

MGF and independence

Moment generating function (MGF), univariate

$$M_X(t) = E(\exp(tX)) \quad \text{for } t \text{ real}$$

What is MGF of $X \sim N(\mu_x, \sigma_x)$?

$$M_X(t) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \int_{-\infty}^{\infty} e^{tx} \exp\left(-\frac{1}{2\sigma_x^2}(x-\mu_x)^2\right) dx$$

$$\left[\begin{array}{ll} u = \frac{x-\mu_x}{\sigma_x} & x = \sigma_x u + \mu_x \\ du = \frac{1}{\sigma_x} dx & \end{array}\right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(t\sigma_x u + \mu_x t - \frac{1}{2}u^2\right) du$$

completing the square

$$-\tfrac{1}{2}(u-\sigma_x t)^2$$

$$= -\tfrac{1}{2}(u^2 - 2\sigma_x u t + \sigma^2 t^2)$$

$$= e^{\mu_x t} \cdot e^{\frac{1}{2}\sigma_x^2 t^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(u-\sigma_x t)^2\right) du}_{1}$$

$$= \underline{\underline{\exp\left(\mu_x t + \tfrac{1}{2}\sigma_x^2 t^2\right)}}$$

MGF

Bivariate

$$M_{X,Y}(t_1, t_2) = \mathrm{E}(\exp(t_1 X + t_2 Y))$$

Rule:
X and Y is independent iff $M_{X,Y}(t_1, t_2) = M_X(t_1) \cdot M_Y(t_2)$.

Bivariate normal and independence

If X and Y are bivariate normal with correlation $\rho = 0$, then this means that X and Y are independent.
Proof: using the bivariate MGF.

MGF for bivariate case

$$M_{X,Y}(t_1,t_2) = E\left(\exp\left(t_1 X + t_2 Y\right)\right)$$

$$= \iint_{-\infty}^{\infty} \exp(t_1 x + t_2 y) \cdot \underbrace{f_X(x) \cdot f(y|x)}_{f(x,y)} \, dx \, dy$$

$N(\mu_x, \sigma_x^2)$ $\qquad$ $N(c_x, \sigma_y^2(1-\rho^2))$

$$= \int_{-\infty}^{\infty} \exp(t_1 x) \cdot f_X(x) \underbrace{\int_{-\infty}^{\infty} \exp(t_2 y) \cdot f(y|x) \, dy}_{\exp\left(c_x t_2 + \frac{1}{2} \sigma_y^2 (1-\rho^2) t_2^2\right)} \, dx$$

$$= \exp\left(\tfrac{1}{2}\sigma_y^2(1-\rho^2)t_2^2\right)\int_{-\infty}^{\infty}\exp(t_1 x + c_x t_2)\, f_X(x)\,dx$$

= Homework

$$= \exp\left\{\mu_x t_1 + \mu_y t_2 + \tfrac{1}{2}\left(\sigma_x^2 t_1^2 + 2\rho\sigma_x\sigma_y t_1 t_2 + \sigma_y^2 t_2^2\right)\right\}$$

$$= \overbrace{\exp\left(\mu_x t_1 + \tfrac{1}{2}\sigma_x^2 t_1^2\right)}^{N(\mu_x,\sigma_x^2)} \cdot \underbrace{\exp\left(\mu_y t_2 + \tfrac{1}{2}\sigma_y^2 t_2^2\right)}_{N(\mu_y,\sigma_y^2)} \cdot \exp\left(\rho\sigma_x\sigma_y t_1 t_2\right)$$

only when $\rho = 0$, $M_{X,Y}(t_1,t_2) = M_X(t_1)\cdot M_Y(t_2)$

MGF rule: X and Y are independent iff

$$M_{X,Y}(t_1, t_2) = M_X(t_1) \cdot M_Y(t_2)$$

We see this occurs if $\rho = 0$.

So, general result X and Y independent $\Rightarrow \rho = 0$ ok,

but also we have that $\rho = 0 \Rightarrow$ independence

for bivariate normal.

Mette.Langaas@math.ntnu.no, TMA4267V2014

Maximum likelihood

$$Y_i = \alpha + \beta x_i + \varepsilon_i, \qquad i = 1, \ldots, n,$$

and $\varepsilon_1, \ldots, \varepsilon_n$ are iid $N(0, \sigma^2)$.
See Exercise 1 - for parameter estimation using Maximum likelihood (ML), and notice that minimizing the sum of squared differences equals maximizing the normal likelihood.

Maximum likelihood and least squares [1.6]

$$Y|X=x \sim N\left(\mu_y + \rho \frac{\sigma_y}{\sigma_x}(x-\mu_x),\ \sigma_y^2(1-\rho^2)\right)$$

(X,Y) bivariate normal with parameters $(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$.

This may be written

$$(Y|X=x) = \underbrace{E(Y|X=x)}_{\mu_y + \rho \frac{\sigma_y}{\sigma_x}(x-\mu_x)} + \underbrace{\varepsilon}_{N(0,\sigma_y^2(1-\rho^2))}$$

Change in notation

$$Y = \alpha + \beta x + \underbrace{\varepsilon}_{N(0,\sigma^2)}$$

If we have a random sample $i = 1, \dots, n$, we write

$$Y_i = \alpha + \beta x_i + \varepsilon_i \quad \text{where}$$

distributed

$$\varepsilon_i \ \text{i.i.d.} \ N(0, \sigma^2)$$

independent identically

$\Downarrow$

Now we have a basic "simple linear regression model".

From TMA4245 we know how to estimate $(\alpha, \beta, \sigma^2)$ by least squares and maximum likelihood.

EI P3.

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$$

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \underbrace{\hat{\alpha} - \beta x_i}_{\hat{Y}_i})^2$$

$$S^2 = \frac{1}{n-2} \text{ — " — }$$

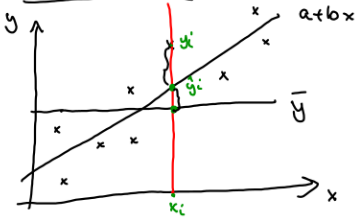
Sum of Squares

Let $\hat{y}_i = a + bx_i$.

$$\underbrace{\sum_{i=1}^{n}(y_i - \bar{y})^2}_{SS_T} = \underbrace{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}_{SS_R} + \underbrace{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}_{SS_E}$$

total=regression+error
Cross terms cancel due to the normal equations.

Sums of squares [1.7]



$$y_i = \bar{y} + (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

square and sum

$$\underbrace{\sum_{i=1}^{n}(y_i-\bar{y})^2}_{\substack{SS\text{ (book)} \\ [SST\text{ total}]}} = \underbrace{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}_{\substack{SSR \\ \text{regression}}} + \underbrace{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}_{\substack{SSE \\ \text{error}}}$$

since $2\sum_{i=1}^{n}(\hat{y}_i-\bar{y})(y_i-\hat{y}_i)=0$

due to normal equation. $\sum(y_i-a-bx_i)=0$

$$\sum(y_i-a-bx_i)\cdot x_i=0$$

$$\hat{y}_i=a+bx_i$$

Moreover

$$\frac{SSR}{SS}=\frac{\sum(\hat{y}_i-\bar{y})^2}{\sum(y_i-\bar{y})^2}=\cdots=r^2$$

proportion of variability explained by x.

empirical correlation coeff. squared

Mindmap

We draw a mindmap of Chapter 1, Bingham and Fry (2010).
You contribute with one word or fact - and we connect these.

$\Sigma = \begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(Y,X) & Var(Y) \end{bmatrix}$

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

MGF (tool)

variability σ_x, σ_y

S_{xx}, S_{yy}

SS, SSR, SSE

regression

$E(Y|x) = \alpha + \beta x$

Correlation ρ

r

$\frac{Cov(X,Y)}{\sigma_x \sigma_y}$

bivariate

Cond. prob

$\rho = 0 \Leftrightarrow X, Y$ indep.

for bivariate N.

$f(y|x) \sim N(c_x, \sigma_y^2(1-\rho^2))$

X, Y biv. N.

Contours

$f(x,y) = \text{const}$

elliptic $\leftarrow$ spectral decomposition

$A = P \Lambda P^T$

Mindmap Ch1

$$\frac{1}{2\pi}\frac{1}{\sqrt{\det(\Sigma)}}$$

$$f(x) = c \cdot \exp\{-\tfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\}$$

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

EXP2 matrix notation

bivariate normal

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

$X \sim N(\mu_x, \sigma_x^2)$
$Y \sim N(\mu_y, \sigma_y^2)$
$\rho = \mathrm{Cov}(X,Y)$

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

$X \longleftrightarrow Y$

$\rho = 0 \Leftrightarrow (X,Y)$ independent

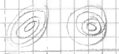
$Y|x \sim N(c_x, \sigma_y^2(1-\rho^2))$

$E(Y|X=x) = \alpha + \beta x$?

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

- if X,Y bivariate normal this is true
- else: may be a good approximation

contours: ellipses



Least squares & normal th.

$y = a + bx$
$\min \sum(y_i - a - bx_i)^2$

$Y = \alpha + \beta x + \varepsilon$

ε i.i.d $N(0,\sigma^2)$

$$\hat\alpha = \bar{Y} - \hat\beta\bar{x}$$

$$\hat\beta = \frac{S_{xY}}{S_{xx}} = \hat\rho\cdot\sqrt{\frac{S_{YY}}{S_{xx}}}$$

$$\frac{S_{xY}}{\sqrt{S_{xx}S_{YY}}}$$