

Chap. 4: Multivariate normal distribution

Rep: Random variables $Y_1, \dots, Y_n$

Random vector $\vec{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$

last week we considered means, variances and covariances. But not the joint distribution of all the random variables in such a vector.

Multivariate distribution:

Random vector $\vec{Y}$, with elements $Y_1, \dots, Y_n$, has joint density $f(\vec{y}) = f(y_1, \dots, y_n)$.

← continuous

This gives probabilities such as

$$\begin{aligned} P(\vec{Y} \leq \vec{y}) &= P(Y_1 \leq y_1, \dots, Y_n \leq y_n) \\ &= \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_n} f(y_1, \dots, y_n) dy_1 \dots dy_n \end{aligned}$$

"volume under the density"

Marginal densities:

$$f_1(y_1) = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{n-1} f(y_1, \dots, y_n) dy_2 \dots dy_n$$

$$f(y_2, \dots, y_n) = \int_{-\infty}^{\infty} f(y_1, \dots, y_n) dy_1$$

all variables except y_1

Conditional densities:

$$f(y_1 | y_2, \dots, y_n) = \frac{f(y_1, \dots, y_n)}{f(y_2, \dots, y_n)}$$

Independence:

If $Y_1, \dots, Y_n$ independent: $f(y_1, \dots, y_n) = \prod_{i=1}^n f(y_i)$ product of marginals
 $f(y_1 | y_2, \dots, y_n) = f(y_1)$

Example: Bivariate distribution

Y_1 : temperature in Trondheim, summer day

Y_2 : — " — Oslo, — " — (av. over 24 hours)

2 min: Sketch marginal densities

$f_1(y_1), f_2(y_2)$ and how Y_1 and Y_2 behave together (joint distn).

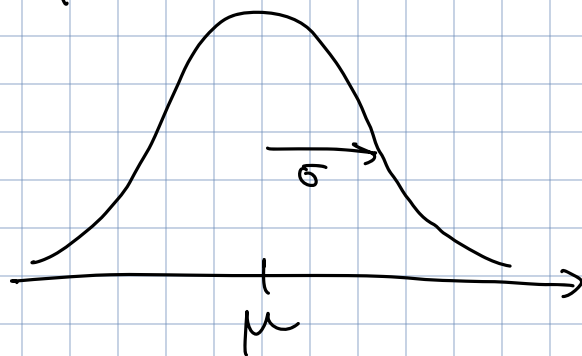
Slides

- 1) Estimated marginal densities, looks almost normally distributed $\mu_1 \approx 12.5$ $\mu_2 \approx 17$.
 $P(Y_1 \leq 10)$ = area under curve, left tail.
- 2) Joint observations, positive correlation.
- 3) Contour plot of joint distribution (3D)
 $P(Y_1 \leq 10, Y_2 \leq 15) \rightarrow$ volume under curve
- 4) Conditional distribution Oslo | Trondheim
 - Both mean and variance of distn in Oslo differs due to Trondheim temp.

4.1 Univariate Normal (+ rep.)

$$Y \sim N(\mu, \sigma^2)$$

$$f(y) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left(\frac{y-\mu}{\sigma} \right)^2}$$



$$Z \sim N(0, 1) \quad \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2}$$

$$\text{Transformations: } Y = \sigma Z + \mu, \quad Z = \frac{Y - \mu}{\sigma}$$

$f(y)$ derived from $\phi(z)$ and transformation formula

$$Z = w(y) = \frac{y - \mu}{\sigma}$$

$$f(y) = \phi(w(y)) \text{abs}\left(\frac{d}{dy} w(y)\right)$$

Rep: If $Y_1, \dots, Y_n$ independent $N(\mu, \sigma^2)$,

$$\sum Y_i \sim N(n\mu, n\sigma^2) \quad (*)$$

Proof: Use moment generating fn.

$$\begin{aligned} \text{Rep: For a random variable } X, \quad M_X(t) &= E(e^{tX}) \\ &= \int_{-\infty}^{\infty} e^{tx} f(x) dx. \end{aligned}$$

Alternative, but very useful way to characterize a distn.

rules

$$M_{aX+b}(t) = E(e^{(aX+b)t}) = E(e^{aXt} e^{bt}) \\ = e^{bt} M_X(at)$$

$$M_{X_1+X_2}(t) = E(e^{(X_1+X_2)t}) \\ = E(e^{X_1t} e^{X_2t}) \\ \text{ind.} \Rightarrow E(e^{X_1t}) E(e^{X_2t}) \\ = M_{X_1}(t) M_{X_2}(t).$$

(*) Proof:

$$M_{Y_1+\dots+Y_n}(t) = \prod_{i=1}^n M_{Y_i}(t) = \prod_{i=1}^n e^{\mu t + \sigma^2 t^2/2} \\ = e^{\sum \mu t + \sigma^2 t^2/2} = e^{(n\mu)t + (n\sigma^2)t^2/2} \\ = M_U(t) \quad \text{where } U \sim N(n\mu, n\sigma^2).$$

4.2 multivariate normal

$$\vec{Y} \sim MVN(\vec{\mu}, \Sigma) \quad E(\vec{Y}) = \vec{\mu}, \text{cov}(\vec{Y}) = \Sigma$$

$$f(\vec{y}) = \frac{1}{(2\pi)^n |\Sigma|} e^{-\frac{1}{2}(\vec{y}-\vec{\mu})^T \Sigma^{-1} (\vec{y}-\vec{\mu})}$$

↑ generalized variance.

standardized distance from $\vec{\mu}$

Derivation of $f(\vec{y})$ (Sketch)

$Z_1, \dots, Z_n$ independent $N(0, 1) \rightarrow \vec{Z}$.

$$\begin{aligned} \phi(\vec{Z}) &= f(z_1, \dots, z_n) \stackrel{\text{ind.}}{=} \prod_{i=1}^n \phi(z_i) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z_i^2} = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2} \sum_{i=1}^n z_i^2} \\ &= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2} \vec{Z}^T \vec{Z}} \end{aligned}$$

joint density $\vec{Z} \sim \sqrt{\quad}$.

Define $\vec{Y} = \Sigma^{1/2} \vec{Z} + \vec{\mu}$ chap. 3.

We know: $E(\vec{Y}) = \vec{\mu}$, $\text{Cov}(\vec{Y}) = \Sigma$

Also: Y_i sum of independent Z -s $\rightarrow Y_i$ normal.

$f(\vec{y}) = ?$ Multivariate transformation formula

$$f(\vec{y}) = \phi(\vec{Z}) \overset{\text{pos def.}}{\text{abs}} \left(\det \left(\Sigma^{-1/2} \right) \right)$$

$\vec{Z} = \Sigma^{-1/2} (\vec{y} - \vec{\mu})$

$\uparrow$ Jacobian

$$= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2} (\vec{y} - \vec{\mu})^T (\Sigma^{-1/2})^T \Sigma^{-1/2} (\vec{y} - \vec{\mu})} |\Sigma^{-1/2}|$$

$\begin{bmatrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_1}{\partial y_2} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$

$\frac{\partial \vec{Z}}{\partial \vec{y}} = \Sigma^{-1/2}$
(similar to $\frac{1}{\sigma}$)

4.3 Moment gen. fun.

We want properties of $\vec{Y}$

Def: For a random vector $\vec{Y}$, the (joint) moment generating fn. is

$$\begin{aligned} M_{\vec{Y}}(\vec{t}) &= E(e^{\vec{t}^T \vec{Y}}) = E(e^{\sum_i t_i y_i}) \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{\sum_i t_i y_i} f(y_1, \dots, y_n) dy_1 \dots dy_n \end{aligned}$$

Properties (p. 92)

1) If $\vec{X}$ and $\vec{Y}$ have equal MGFs, $M_{\vec{X}}(\vec{t}) = M_{\vec{Y}}(\vec{t})$, then $\vec{X}$ and $\vec{Y}$ have the same distribution.

2) Consider random vector $\vec{Y} = \begin{pmatrix} \vec{Y}_1 \\ \vec{Y}_2 \end{pmatrix}$, with mgf $M_{\vec{Y}}(\vec{t})$.

Vectors $\vec{Y}_1, \vec{Y}_2$ are independent if and only if $(\Leftrightarrow)$

$$M_{\vec{Y}}(\vec{t}) = M_{\vec{Y}_1}(\vec{t}_1) M_{\vec{Y}_2}(\vec{t}_2).$$

Thm. 4.3 If $\vec{Y} \sim \text{MVN}_n(\vec{\mu}, \Sigma)$ then

$$M_{\vec{Y}}(\vec{t}) = e^{\vec{t}^T \vec{\mu} + \vec{t}^T \Sigma \vec{t} / 2}$$

4.4 Properties

Linear functions

This is about sums,
before we only knew result
for sum of independent normal.

Thm 4.4 a.

$$\vec{Y} \sim \text{MVN}_n(\vec{\mu}, \Sigma)$$

positive definite
- invertible, full rank (n)
 $\vec{a}^T \Sigma \vec{a} > 0 \quad \vec{a} \neq \vec{0}$

$\vec{a}$ n-vector of constants

A $k \times n$ matrix, $k \leq n$, $\text{rank}(A) = k$.
full rank (k)

$$(i) \quad \vec{a}^T \vec{Y} = a_1 Y_1 + a_2 Y_2 + \dots + a_n Y_n$$

sum of dependent
normal vars.

$$\text{is } N(\vec{a}^T \vec{\mu}, \vec{a}^T \Sigma \vec{a})$$

follow from chapter 3.

$$(ii) \quad \underbrace{A\vec{Y}}_{k\text{-vector}} \sim \text{MVN}_k(\underbrace{A\vec{\mu}}_{k\text{-vector}}, \underbrace{A\Sigma A^T}_{k \times k \text{ pos. def. (full rank)}})$$

$k \times k$ pos. def.
(full rank)

Proof (i)

$$U = \vec{a}^T \vec{Y}, \text{ want } M_U(t) = e^{\mu_U(t) + (\sigma_U^2 t)^2 / 2}$$

$$M_U(t) = E(e^{Ut}) = E(e^{(\vec{a}^T \vec{Y})t}) = E(e^{(t\vec{a})^T \vec{Y}})$$

$$= M_{\vec{Y}}(t\vec{a}) = e^{(t\vec{a})^T \vec{\mu} + (t\vec{a})^T \Sigma (t\vec{a}) / 2}$$

we know
this one

$$= e^{t \underbrace{(\vec{a}^T \vec{\mu})}_{\mu_U} + t^2 \underbrace{(\vec{a}^T \Sigma \vec{a})}_{\sigma_U^2} / 2}$$

Corollary 1 $\underset{\substack{\uparrow \\ k \times n}}{A} \underset{\substack{\uparrow \\ k}}{\vec{Y}} + \vec{b} \sim \text{MVN}_k (A\vec{\mu} + \vec{b}, A\Sigma A^T)$

Thm 4.4 b) marginal dist^{ns} also normal

$$\vec{Y} = \begin{pmatrix} \vec{Y}_1 \\ \vec{Y}_2 \end{pmatrix} \quad \vec{\mu} = \begin{pmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{pmatrix} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

$$\vec{Y} \sim \text{MVN}(\vec{\mu}, \Sigma)$$

Then $\vec{Y}_1 \sim \text{MVN}(\vec{\mu}_1, \Sigma_{11})$.

Proof: $\vec{Y}_1 = A\vec{Y}$ with $A = \begin{bmatrix} I & 0 \end{bmatrix}$ $\swarrow \begin{smallmatrix} r \times r \end{smallmatrix}$ $\searrow \begin{smallmatrix} r \times (n-r) \end{smallmatrix}$

(thm 4.4 b) $A\vec{\mu} = \vec{\mu}_1 \quad A\Sigma A^T = \Sigma_{11}.$

Corollary 1: If $\vec{Y} \sim \text{MNN}(\mu, \Sigma)$,

$$Y_i \sim N(\mu_i, \sigma_{ii}^2).$$