Part I: Linear Models

Chapter I.2

Multivariate Normal Distribution

Topics

To be discussed...

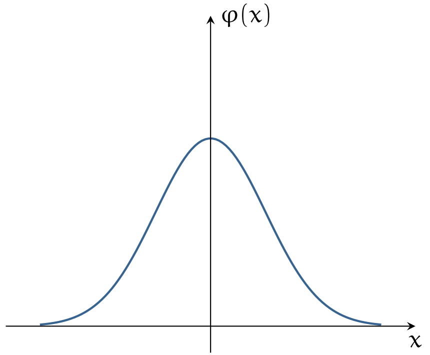
- definition of general multivariate normal distributions
- marginal & conditional distributions
- independence & correlation
- some illustrations (bivariate normal)
- linear transformations and applications

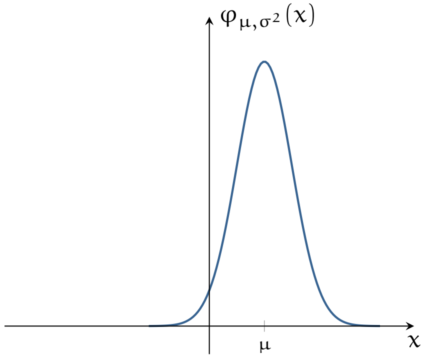
I.2.1 Univariate normal distribution $N(\mu, \sigma^2)$

- $X \sim N(\mu, \sigma^2)$ with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ has the PDF

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

- For $Z \sim N(0,1)$, we have $\boxed{X = \mu + \sigma Z}$; $\varphi = \varphi_{0,1}$.





I.2.2 Definition

Let $Z_1, \ldots, Z_k \overset{\text{iid}}{\sim} N(0,1)$, $\mathbf{Z} = (Z_1, \ldots, Z_k)'$, and $\boldsymbol{\mu} \in \mathbb{R}^p, A \in \mathbb{R}^{p \times k}$. Then:

- $\mathbf{Z}$ has a **k-dimensional standard normal distribution** (for short $\mathbf{Z} \sim N_k(0, I_k)$).
- $\mathbf{X} = \boldsymbol{\mu} + A\mathbf{Z}$ has a **p dimensional normal distribution with parameters μ and $\Sigma = AA'$** (for short $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$).

I.2.3 Remark

- In the situation of Definition I.2.2, we have
  - $E\mathbf{X} = \boldsymbol{\mu}$, Cov $(\mathbf{X}) = \Sigma$
  - Since rank(Σ) = rank(A) may be less than p, Σ may be a singular matrix.
- How can a multivariate normal distribution be generated with a **given $\boldsymbol{\mu} \in \mathbb{R}^p$ and $\Sigma \in \mathbb{R}^{p \times p}_{\geqslant 0}$**?

  ☞ How to choose A?

I.2.4 Corollary

Let $Z_1, \dots, Z_p \overset{\text{iid}}{\sim} N(0,1)$, $\mathbf{Z} = (Z_1, \dots, Z_p)'$, and $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}_{\geqslant 0}$. Then:

$$\mathbf{X} = \boldsymbol{\mu} + \Sigma^{1/2}\mathbf{Z} \sim N_p(\boldsymbol{\mu}, \Sigma).$$

I.2.5 Theorem

Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}_{>0}$ and determinant $\det(\Sigma)$. Then, $\mathbf{X}$ has the PDF

$$\boxed{f^{\mathbf{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^p \det(\Sigma)}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right), \quad \boldsymbol{x} = (x_1, \dots, x_p)' \in \mathbb{R}^p.}$$

I.2.6 Remark

A multivariate normal distribution $N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p$ and singular covariance-matrix $\Sigma \in \mathbb{R}^{p \times p}_{\geqslant 0}$ does not have a density function on $\mathbb{R}^p$!

Marginals & Conditionals

For a vector $\mathbf{x} \in \mathbb{R}^p$ and $\emptyset \neq K \subseteq \{1, \ldots, p\}$, let $\mathbf{x}_K = (x_i)_{i \in K}$.

I.2.7 Theorem (parameters and marginals of a multivariate normal distribution)

Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}_{\geqslant 0}$ and $\emptyset \neq K \subseteq \{1, \ldots, p\}$ and $\Sigma_{K,K} = \text{Cov}(\mathbf{X}_K)$. Then:

1. $E\mathbf{X} = \boldsymbol{\mu}$
2. $\text{Cov}(\mathbf{X}) = \Sigma$
3. $\mathbf{X}_K \sim N(\boldsymbol{\mu}_K, \Sigma_{K,K})$ ('**marginals of normals are normal**')

I.2.8 Theorem (conditionals of a multivariate normal distribution)

Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}_{>0}$ and $\emptyset \neq K, L \subseteq \{1, \ldots, p\}$, $K \cap L = \emptyset$, $k = |K|$. Further, let $\Sigma_{K,L} = \text{Cov}(\mathbf{X}_K, \mathbf{X}_L)$ and $\Sigma_{KK|L} = \Sigma_{K,K} - \Sigma_{K,L}\Sigma_{L,L}^{-1}\Sigma_{K,L}'$. Then:

1. $\mathbf{X}_K \mid \mathbf{X}_L = \mathbf{x}_L \sim N_k(\boldsymbol{\mu}_K + \Sigma_{K,L}\Sigma_{L,L}^{-1}(\mathbf{x}_L - \boldsymbol{\mu}_L), \Sigma_{KK|L})$

   ('**conditionals of normals are normal**')
2. $E(\mathbf{X}_K \mid \mathbf{X}_L = \mathbf{x}_L) = \boldsymbol{\mu}_K + \Sigma_{K,L}\Sigma_{L,L}^{-1}(\mathbf{x}_L - \boldsymbol{\mu}_L)$

   The matrix $\Sigma_{K,L}\Sigma_{L,L}^{-1}$ is called **regression matrix**.
3. $\text{Cov}(\mathbf{X}_K \mid \mathbf{X}_L = \mathbf{x}_L) = \Sigma_{KK|L}$,

Independence & Correlation

I.2.9 Theorem (independence under multivariate normal distribution)

Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p\times p}_{\geqslant 0}$ and $\emptyset \neq K, L \subseteq \{1, \ldots, p\}$, $K \cap L = \emptyset$, $k = |K|$. Further, let $\Sigma_{K,L} = \operatorname{Cov}(\mathbf{X}_K, \mathbf{X}_L)$. Then:

1. $\mathbf{X}_K$ and $\mathbf{X}_L$ are independent if and only if $\Sigma_{K,L} = \mathbb{0}$
2. $\mathbf{X} = (X_1, \ldots, X_p)' \sim N_p(\mathbb{0}, I_p) \iff X_1, \ldots, X_p \overset{\text{iid}}{\sim} N(0, 1)$
3. $\mathbf{X} = (X_1, \ldots, X_p)' \sim N_p(\boldsymbol{\mu}, \Sigma)$ with a diagonal matrix $\Sigma = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_p^2)$

   $\iff X_1, \ldots, X_p$ are independent random variables with $X_j \sim N(\mu_j, \sigma_j^2)$, $1 \leqslant j \leqslant p$

Bivariate normal distribution

I.2.10 Example (Bivariate normal distribution)

A **bivariate normal distributed random vector** $\mathbf{X} = (X_1, X_2)'$ has the PDF (for $x_1, x_2 \in \mathbb{R}$)

$$f^{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} \right.\right. \\ \left.\left. - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right] \right) \quad \text{(C.1)}$$

with parameters $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$ and $\rho \in (-1, 1)$;

- for short: $(X_1, X_2)' \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$
- covariance matrix Σ as in Theorem I.2.7:

$$\boxed{\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \in \mathbb{R}^{2\times 2}}$$

with determinant $\det \Sigma = \sigma_1^2\sigma_2^2(1-\rho^2)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2)' \in \mathbb{R}^2$

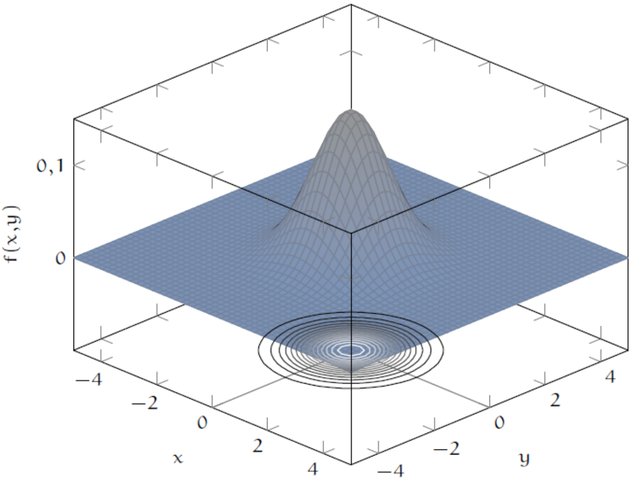


Figure: PDF of bivariate standard normal distribution $N_2(0, 0, 1, 1, 0)$

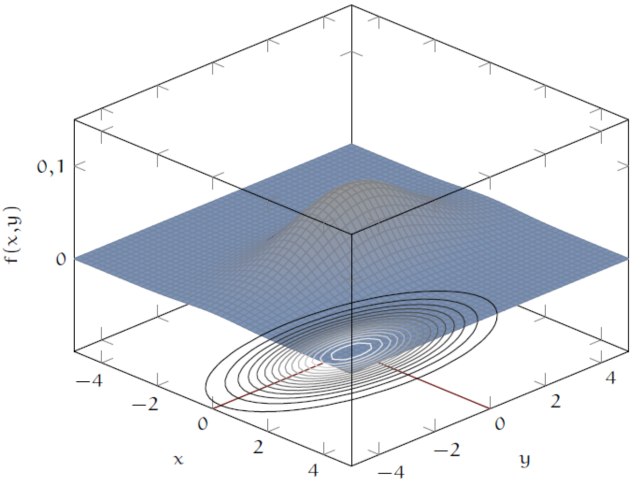


Figure: PDF of bivariate normal distribution $N_2(0, 0, 1, 4, 0)$

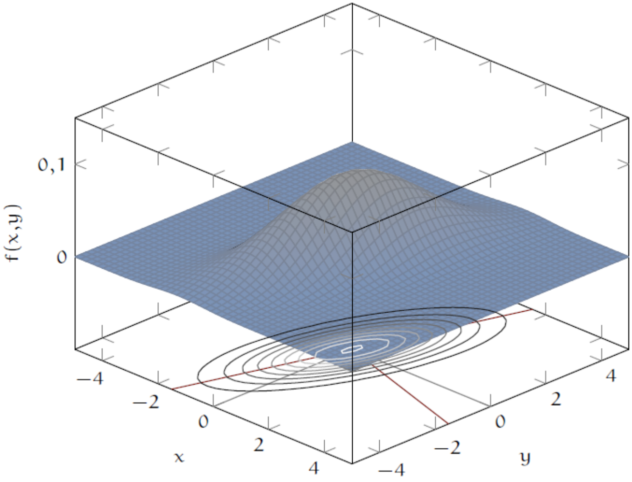


Figure: PDF of bivariate normal distribution $N_2(0, 0, 1, 4, \frac{1}{2})$

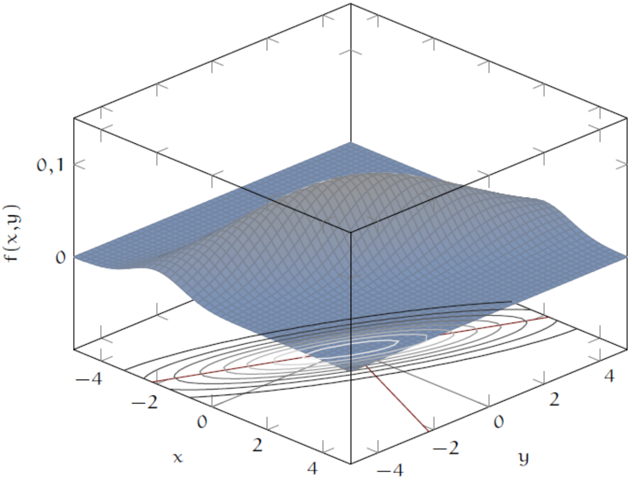


Figure: PDF of bivariate normal distribution $N_2(0, 0, 1, 4, \frac{4}{5})$

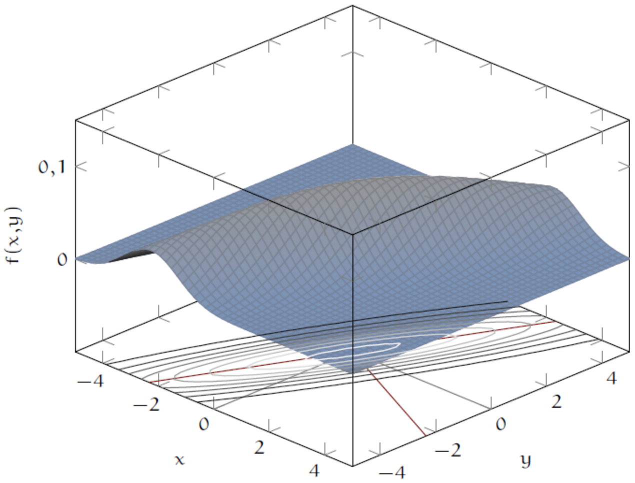
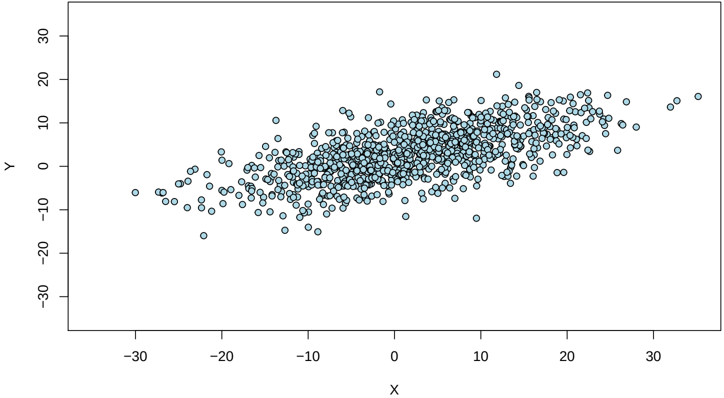
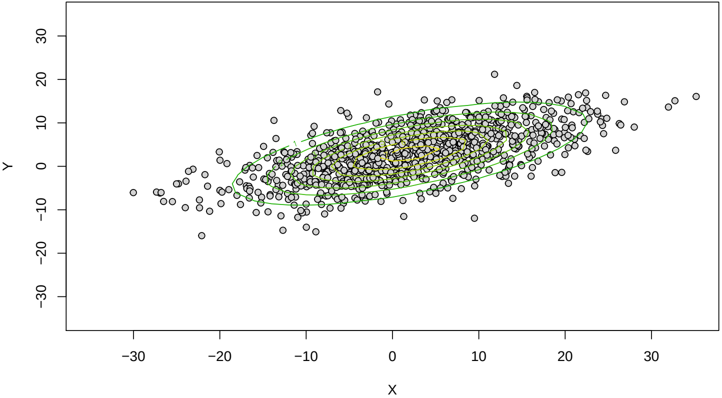


Figure: PDF of bivariate normal distribution $N_2(0, 0, 1, 4, \frac{9}{10})$

Simulated bivariate normal data



Simulated bivariate normal data with ellipses



I.2.11 Theorem

Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p\times p}_{\geqslant 0}$ and $\mathbf{a} \in \mathbb{R}^k, B \in \mathbb{R}^{k\times p}$, $1 \leqslant k \leqslant p$. Then:

$$\mathbf{Y} = \mathbf{a} + B\mathbf{X} \sim N_k(\mathbf{a} + B\boldsymbol{\mu}, B\Sigma B').$$

In particular, we get for $\Sigma \in \mathbb{R}^{p\times p}_{>0}$ and $\Sigma^{-1/2} = (\Sigma^{1/2})^{-1}$

$$\mathbf{Y} = \Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim N_p(\mathbf{0}, I_p).$$

I.2.12 Remark

- The transformation
$$\mathbf{Y} = \Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$$
considered in Theorem I.2.11 is called **Mahalanobis transformation**.
- Considering the Euclidean norm $\|\mathbf{Y}\|$ of $\mathbf{Y}$, we get
$$\|\mathbf{Y}\|^2 = \mathbf{Y}'\mathbf{Y} = (\Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}))'\Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{X} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{X} - \boldsymbol{\mu}) = \|\mathbf{X} - \boldsymbol{\mu}\|^2_\Sigma, \text{ say.}$$
$\|\mathbf{X}\|^2_\Sigma$ is called **Mahalanobis norm** of $\mathbf{X}$. For random vectors $\mathbf{X}$ and $\mathbf{Y}$ of the same dimension p, $\|\mathbf{X} - \mathbf{Y}\|^2_\Sigma$ is called **Mahalanobis distance** of $\mathbf{X}$ and $\mathbf{Y}$.

I.2.13 Corollary

1. Let $\mathbf{X} = (X_1, \ldots, X_p)' \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\Sigma = (\sigma_{ij})_{i,j}$ and $\overline{X} = \frac{1}{p}\sum\limits_{j=1}^{p} X_j$. Then:

$$\overline{X} = \frac{1}{p}\mathbb{1}_p'\mathbf{X} \sim N\left(\frac{1}{p}\sum_{j=1}^{p}\mu_j, \frac{1}{p^2}\mathbb{1}_p'\Sigma\mathbb{1}_p\right) = N\left(\frac{1}{p}\sum_{j=1}^{p}\mu_j, \frac{1}{p^2}\sum_{i,j}\sigma_{ij}\right).$$

2. If $X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$, $\sigma^2 > 0$, then
   - $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$
   - $\boxed{\sqrt{n}\,\frac{\overline{X} - \mu}{\sigma} \sim N(0, 1)}$

I.2.14 Theorem

Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}_{\geqslant 0}$. Then, for $A \in \mathbb{R}^{k \times p}, B \in \mathbb{R}^{r \times p}$ with $k, r \in \mathbb{N}$, we get:

- $A\mathbf{X}$ and $B\mathbf{X}$ are independent if and only if $A\Sigma B' = 0$.

I.2.15 Theorem

Let $p \geqslant 2$, $\mathbf{Z} = (Z_1, \ldots, Z_p)' \sim N_p(0, \sigma^2 I_p)$, $\overline{Z} = \frac{1}{p}\sum_{j=1}^p Z_j$, and $E_p = I_p - \frac{1}{p}\mathbb{1}_{p \times p}$. Then:

- $\overline{Z}$ and $\mathbf{Z} - \overline{Z}\mathbb{1}_p = E_p\mathbf{Z}$ are independent.
- $\overline{Z}$ and $S_Z = \frac{1}{p-1}\sum_{j=1}^p (Z_j - \overline{Z})^2$ are independent.

I.2.16 Lemma

Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(0,1)$. Then:

1. $X_1^2 \sim \chi^2(1)$
2. $\sum\limits_{i=1}^{n} X_i^2 \sim \chi^2(n)$

I.2.17 Corollary

Let $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ with $\mu \in \mathbb{R}$, $\sigma > 0$ and

$$\widehat{\sigma}_\mu^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \mu)^2, \quad \widehat{\sigma}^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_j - \overline{X})^2.$$

mu is known, so we can include all the random variables(n) s.t. the degrees of freedom is n

estimate of unbiased corrected standard deviation

we measure the sample mean and get the squared sum of the difference between this sample mean and each variable. In this way, we loose one random variable, so it results in n-1

Then:

1. $\frac{n}{\sigma^2} \widehat{\sigma}_\mu^2 \sim \chi^2(n)$.
2. $\overline{X}$ and $\widehat{\sigma}^2$ are independent.
3. $\frac{n-1}{\sigma^2} \widehat{\sigma}^2 \sim \chi^2(n-1)$ (if $n \geqslant 2$)