Applied Statistics

for Computer Science BSc, Exam

Probability Theory and Mathematical Statistics

for Computer Science Engineering BSc, Term grade

István Fazekas
University of Debrecen

2020/21 fall

This work was supported by the construction EFOP-3.4.3-16-2016-00021. The project was supported by the European Union, co-financed by the European Social Fund.

Main topics

1. Probability theory
2. Statistics

Mathematical tools: combinatorics, calculus

Computer tool: Matlab

Book:

Yates, Goodman:

Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers

Lecture 9

Joint distributions

The joint distribution function

Definition. Let X and Y be random variables. Then the function

$$\boxed{F(x, y) = P(X < x, Y < y)}, \quad x, y \in \mathbb{R}, \tag{1}$$

is called the joint cumulative distribution function (joint CDF) of X and Y.

The joint distribution function...

Theorem. $F : \mathbb{R}^2 \to \mathbb{R}$ is a joint CDF if and only if

a) F is monotone increasing in both variables;

b) F is left continuous in both variables;

c) $\lim_{x \to -\infty} F(x, y) = \lim_{y \to -\infty} F(x, y) = 0$
and $\lim_{x \to \infty, y \to \infty} F(x, y) = 1$;

d) $F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \geq 0$
for all (a_1, a_2), (b_1, b_2) with $a_1 < b_1$, $a_2 < b_2$.

The joint distribution function...

Proof. Let F be the joint CDF of X and Y.
a), b) and c) are similar to the univariate case.
The proof of d) is

$$F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2 =$$

$$= P\{(X, Y) \in [a_1, b_1) \times [a_2, b_2)\} \geq 0$$

Conversely, if F satisfies a), b), c) and d), then the function $P([a_1, b_1) \times [a_2, b_2)) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2)$ defines a probability on $\mathbb{R}^2$. One can see that the joint CDF of the random variables $X(x, y) = x$ and $Y(x, y) = y$ is the function F.

Marginal distribution functions

Let the joint CDF of X and Y be $F(x, y)$.
Then the CDF of X is $F_X(x) = \lim_{y \to \infty} F(x, y)$
and the CDF of Y is $F_Y(y) = \lim_{x \to \infty} F(x, y)$.
F_X and F_Y are called the marginal CDF's of $F(x, y)$.
Proof.

$$F_X(x) = P(X < x) = P(X < x, Y \in \mathbb{R}) =$$

$$= \lim_{y_n \to \infty} P(X < x, Y < y_n) = \lim_{y_n \to \infty} F(x, y_n)$$

because of the continuity of the probability.

Example for joint distribution function (exponential)

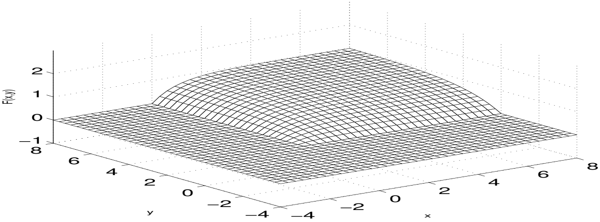

Figure: $F(x, y)$ when $\lambda = 1$, $\mu = 2$

Let $F(x, y) = 1 - e^{-\lambda x} - e^{-\mu y} + e^{-(\lambda x + \mu y)}$, if $x, y > 0$,
and $F(x, y) = 0$ else.
Then $F(x, y)$ is a joint CDF.

The joint density function

The joint distribution of X and Y is called absolutely continuous if there exists a two variable function f so that the joint CDF F is

$$\boxed{F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, dv\, du}\ , x, y \in \mathbb{R}$$

for all $x, y \in \mathbb{R}$.
f is called the joint probability density function (joint PDF) of X and Y.
Remark. If f is the joint PDF of X, Y, then

$$\boxed{P((X,Y) \in B) = \iint_B f(x,y)\, dx\, dy}$$

for any two-dimensional Borel set B.

The joint density function...

Theorem. $f : \mathbb{R}^2 \to \mathbb{R}$ is a joint PDF if and only if

a) it is measurable,

b) it is non-negative,

c) $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$

Marginal density functions

Let the joint PDF of X and Y be f.
Then both X and Y have PDF and they are

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy, \;\; x \in \mathbb{R}; \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx, \;\; y \in \mathbb{R}.$$

f_X and f_Y are called the marginal PDF's.

The joint PDF and the expectation

If X and Y are random variables and $g : \mathbb{R}^2 \to \mathbb{R}$ is a measurable function, then
$g(X, Y)$ is a random variable.
If $f(x, y)$ is the joint PDF of X and Y, then

$$\boxed{\mathbb{E}g(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy}$$

Independence of random variables

X and Y are called independent if their joint CDF is the product of the two marginal CDF's, that is

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y), \quad x, y \in \mathbb{R}. \tag{2}$$

Remark. (2) is equivalent to

$$P(X \in B_1, Y \in B_2) = P(X \in B_1)P(Y \in B_2) \tag{3}$$

for any Borel sets B_1 andB_2.
(3) shows that in the case of independence the events connected to X are independent of the events connected to Y.

Independence and density function

Let the joint distribution of X and Y be absolutely continuous. Then X and Y are independent if and only if their joint PDF is the product of the two marginal PDF's, that is

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y), \quad x,y \in \mathbb{R}. \tag{4}$$

The covariance

$$\boxed{\operatorname{cov}(X, Y) = \mathbb{E}\Big((X - \mathbb{E}X)(Y - \mathbb{E}Y)\Big)}$$

To calculate the covariance use

$$\operatorname{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$$

and

$$\mathbb{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\, dx\, dy$$

The covariance measures the strength of dependence of X and Y.
The properties of the covariance are the same as in the discrete case.

Remark. If X and Y are independent and both of them have finite expectation, then

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

and $\operatorname{cov}(X, Y) = 0$

Random vectors

Let $X_1, X_2, \ldots, X_n$ be random variables. Then

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

is called a random vector. Its expectation vector is

$$\mathbb{E}\boldsymbol{X} = \begin{pmatrix} \mathbb{E}X_1 \\ \mathbb{E}X_2 \\ \vdots \\ \mathbb{E}X_n \end{pmatrix}$$

Random vectors...

Its covariance matrix is

$$\mathrm{Var}(\boldsymbol{X}) = \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{cov}(X_1, X_2) & \cdots & \mathrm{cov}(X_1, X_n) \\ \mathrm{cov}(X_2, X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(X_n, X_1) & \mathrm{cov}(X_n, X_2) & \cdots & \mathrm{Var}(X_n) \end{pmatrix}.$$

Multivariate normal distribution...

We call $\boldsymbol{X}$ non-degenerate n-dimensional normal random variable if its PDF is

$$\boxed{f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}(\det D)^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^\top D^{-1}(\boldsymbol{x}-\boldsymbol{m})\right\}} \tag{5}$$

Here $\boldsymbol{x} \in \mathbb{R}^n$, and $\boldsymbol{m}$ is the expectation vector of $\boldsymbol{X}$ and the positive definite symmetric matrix D is its covariance matrix. Notation $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{m}, D)$

The two-dimensional normal distribution

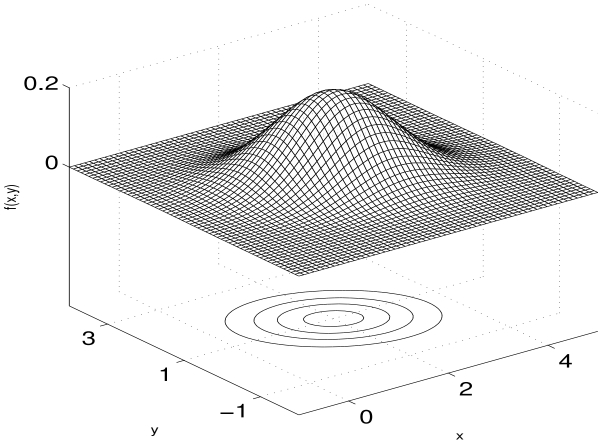

Figure: The two-dimensional normal PDF

The two-dimensional normal distribution

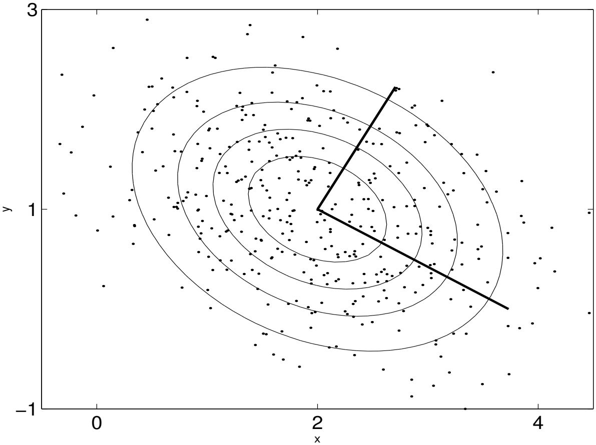

Figure: Concentration ellipses of a two-dimensional normal distribution and a sample from the same population

The three-dimensional normal distribution

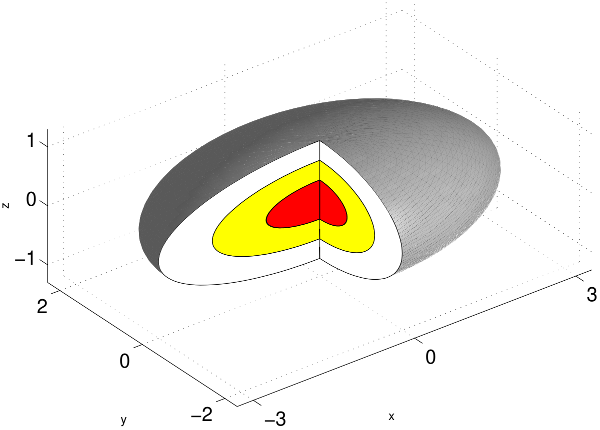

Figure: Concentration ellipsoids of a three-dimensional normal distribution