

Exercise 1: Brief Introduction to Statistics
(to be returned on Nov 3, 2015, 8:15 in HS 101.00.026,
or before in building 102, 1st floor, 'Anbau')

Prof. Dr. Moritz Diehl, Robin Verschueren, Fabian Girrbaach and Jesus Lago

In this exercise we investigate some important facts from mathematical statistics in theory and in numerical experiments.

For the last exercise, create a MATLAB script called `main.m` with your code, possibly calling other functions/scripts. From running this script, all the necessary results and plots should be clearly visible. Compress all the files/functions/scripts necessary to run your code in a `.zip` file and send it to `jesuslagogarcia@gmail.com`. Please state your name and the names of your team members in the e-mail.

Exercise Tasks

1. Given the following x-y data set, compute the vector mean and the covariance matrix. (1 point)

x	2.72	4.53	2.75	1.94	4.60	4.23	2.77	1.49	2.56
y	0.84	1.28	0.74	1.44	1.39	-0.25	0.05	0.26	0.49

2. Given a random variable X with a normal propability density $p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$ with μ the mean and σ^2 the variance of X , write down the probability density function $p_Y(y)$ of the random variable $Y = 5X + 3$, using μ and σ . (1 point)
3. The covariance matrix of a vector valued random variable X in \mathbb{R}^n with mean $\mathbb{E}\{X\} = \mu_X$ is defined by

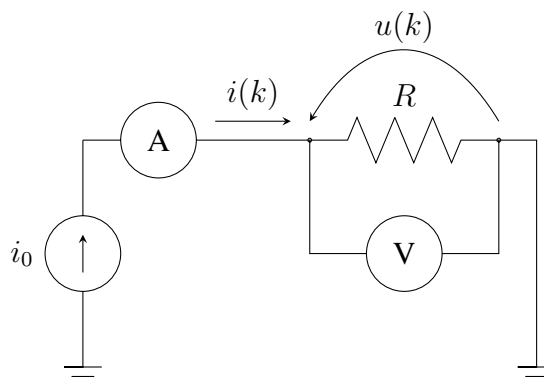
$$\text{cov}(X) := \mathbb{E}\{(X - \mu_X) \cdot (X - \mu_X)^\top\}.$$

Prove that the covariance matrix of a vector valued variable $Y = AX + b$ with constant $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ is given by

$$\text{cov}(Y) = A \cdot \text{cov}(X) \cdot A^\top.$$

(2 points)

4. Show that if two scalar random variables are statistically independent, then they are uncorrelated. Conversely, can you find an example of two variables that are statistically dependent but uncorrelated? (2 points)
5. Computer exercise with MATLAB. We consider the following experimental setup:



The resistor R has a current i_0 flowing through it which is determined by the current source. The goal of the experiment is to estimate the value of R using Ohm's law. Since the value of i_0 is not known exactly, we use an ammeter to measure the current $i(k)$ and a voltmeter to measure $u(k)$. Since several measurements are taken, k represents the measurement number. We assume that the measurements are noisy:

$$i(k) = i_0 + n_i(k)$$

and

$$u(k) = u_0 + n_u(k)$$

where u_0 is the true value of the voltage across the resistor and $n_i(k)$ and $n_u(k)$ are the values of the noise.

Let us now investigate the behaviour of the following three different estimators:

- $\hat{R}_{SA}(N) = \frac{1}{N} \sum_{k=1}^N \frac{u(k)}{i(k)}$
- $\hat{R}_{LS}(N) = \frac{\frac{1}{N} \sum_{k=1}^N u(k)i(k)}{\frac{1}{N} \sum_{k=1}^N i(k)^2}$
- $\hat{R}_{EV}(N) = \frac{\frac{1}{N} \sum_{k=1}^N u(k)}{\frac{1}{N} \sum_{k=1}^N i(k)}$

We will use MATLAB to simulate the behavior of these estimators for $i_0 = 5$ and $u_0 = 10$ and the following measurement noise: $n_i(k)$ and $n_u(k)$ are i.i.d. with normal distribution, zero mean, and $\sigma = 2$.

For each of the three estimators, carry out the following tasks (useful MATLAB commands are `help`, `plot`, `for`, `randn`, `mean`, `std`)

- (a) Simulate the outcome of $M=200$ experiments with $N_{\max}=1000$ measurements each, i.e. $k = 1, \dots, N_{\max}$;
- (b) For all 200 experiments plot the function $\hat{R}_*(N)$, $N = 1, \dots, N_{\max}$ (* can be either SA, LS or EV); for each estimator; to see the stochastic variations, plot all these functions in one graph using `hold on`.
- (c) Compute the mean of $\hat{R}_*(N)$ over all experiments and plot it for N from 1 to N_{\max} ;
- (d) Compute the standard deviation of $\hat{R}_*(N)$ over all experiments and plot it for N from 1 to N_{\max} ;
- (e) Plot a histogram containing all values of $\hat{R}_*(N_{\max})$;
- (f) Set $\epsilon = 0.01$ and plot the experimental probability $P(|\hat{R}_*(N) - R| \geq \epsilon)$ w.r.t. N (use $R = u_0/i_0 = 2$).

(6 points)

This sheet gives in total 12 points (plus 4 bonus points from the previous sheet)