MEK4350, fall 2016
Exercises VII

Problem 1 – normal or Gaussian distribution $X \sim N(\mu, \sigma)$

The probability density of a normal or Gaussian stochastic variable X is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

1. Show that $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$.

   Hint: Let $I = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$. Evaluate $I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(y)\,\mathrm{d}x\,\mathrm{d}y$ by employing polar coordinates.

2. Show that $E[X] = \mu$.

3. Compute the first four central moments $\mu_n = E[(X-\mu)^n]$ for $n = 1, 2, 3, 4$.

   Hint: For the odd moments, use the symmetry properties of the integrand. For the even moments, first derive the identity $\int_{-\infty}^{\infty} \mathrm{e}^{-\alpha x^2}\,\mathrm{d}x = \sqrt{\pi/\alpha}$, then differentiate this identity a few times with respect to α.

4. Find the variance σ^2, standard deviation $\sigma = \sqrt{\sigma^2}$, skewness $\gamma_1 = \gamma$, kurtosis κ and excess of kurtosis $\gamma_2 = \kappa - 3$ of the Gaussian distribution.

5. Show that the characteristic function is

   $$\phi(k) = E[\mathrm{e}^{\mathrm{i}kX}] = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}+\mathrm{i}kx}\,\mathrm{d}x = \mathrm{e}^{\mathrm{i}\mu k - \frac{1}{2}\sigma^2 k^2}$$

   Hint 1: Complete the square of the exponent in the integral such that it becomes $-a(x-b)^2 + c$.

   Hint 2 (requires some knowledge of complex analysis): Construct a rectangular contour in the complex $z = x + \mathrm{i}y$ plane going along the real x-axis from $-L$ to L, then parallel to the imaginary y-axis to somewhere related to b, and close the contour. The integrand has no singularities within the contour, thus the integral around the closed contour is zero (Cauchy's theorem). Show that the contributions from the "vertical" parts at $x = \pm L$ vanish in the limit $L \to \infty$. Show that the integral along the straight line parallell to the real x-axis is known in accordance with the previous problems.

6. The error function is defined as

   $$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2}\,\mathrm{d}t$$

   Express the cumulative Gaussian distribution function by means of the error function.

7. Compute the probability that a Gaussian stochastic variable is within n standard deviations from the mean, $P\{|X - \mu| \leq n\sigma\}$ for $n = 1, 2, 3, \ldots, 8$.

   Do this by showing how to express the result in terms of the error function, and then find the numerical value by invoking `erf` on the computer.

Problem 2 – empirical distribution function

1. Consider 100 numbers $\{x_j\}$ for $j = 1, 2, \ldots, 100$ drawn from a Gaussian distribution. This can be achieved on the computer with the command `randn`. To construct the empirical distribution function, sort the numbers in increasing order $x_{min} \leq \ldots \leq x_{max}$. This can be done on the computer with the command `sort`. Then construct a staircase with steps of size 1/100 with the steps located at the sorted values of x_j.

   Overplot in the same coordinate system the theoretical cumulative distribution function. Do this on the computer by calling `erf`.

2. Repeat the above exercise, but this time for the uniform distribution. Uniformly distributed numbers can be drawn on the computer with the command `rand`.

Problem 3 — central limit theorem

We consider N uniformly distributed stochastic variables X_n for $n = 1, 2, \ldots, N$. Draw J numbers for each of these N variables $\{x_{n,j}\}$ where n is the index of the stochastic variable, and $j = 1, 2, \ldots, J$ is the index of the number drawn. Let $X = \sum_n X_n$. Do an experiment on the computer to decide how big N should be in order that the cumulative distribution function of X be "sufficiently" Gaussian.

Hint: You may let $J = 100$ as in problem 2, and most likely N does not have to be very large?

Problem 4

Let $\phi(k)$ be the characterictic function of a stochastic variable (of any distribution). Show that the characteristic function is maximum at the origin

$$|\phi(k)| \leq \phi(0)$$

and find the value at the origin $\phi(0)$.