Problems on *statistics and data analysis (MVComp2)*

lecturer: Björn Malte Schäfer
head tutor: Alessio Spurio Mancini
summer term 2016

Problem sheet 2

To be handed in during the exercise group on 09.May.2016

1. *Bernoulli distribution* *(10 points)*

Show that the Bernoulli distribution approaches a Gaussian distribution for large n, even if p is not equal to 1/2. Do this by computing the centralised third moment $\langle(x - x_{mean})^3\rangle$ and normalise it to the variance $\langle(x - x_{mean})^2\rangle$, taken to the power of $3/2$.

2. *Moments* *(10 points)*

Under which conditions does a series of moments converge towards the characteristic function? Say if you have convergence for

(a) constant moments $\langle x^n\rangle = c$,

(b) moments $\langle x^n\rangle = n!$,

(c) moments $\langle x^n\rangle = c^n n!$.

If there is convergence only for a finite interval in t, estimate how large it would be.

3. *Moments of the exponential distribution* *(10 points)*

Show that the moments $\langle x^n\rangle$ of the exponential distribution $p(x) = \exp(-x)$, $x = 0 \ldots \infty$, are equal to $n!$, both by direct integration (use induction to show the general case) and by differentiation of the moment generating function.

4. *Cumulants and moments* *(10 points)*

Show the relation between cumulants and moments: differentiate the relation $\phi(t) = \exp(K(t))$, where $\phi(t) = \sum_n \langle x^n\rangle \frac{(it)^n}{n!}$ and $K(t) = \sum_n \kappa_n \frac{(it)^n}{n!}$.

5. Python exercise *(10 points)*

- Write a Python script to simulate a 1D random walk, in which a person starts off at a point 0 and at each step randomly picks a direction (left or right) and moves 1 step in that direction. Take a positive integer n and terminate the simulation when the walk reaches n or $-n$.

  Decide the number of runs that you want to consider and, by averaging over this number, report an estimated average number of steps needed for the walk to terminate.

  Repeat this for different values of n, and plot the results to show qualitatively how rapidly the walk terminates, as a function of n.

- Write a Python script to draw a vector $\vec{X}$ of n random numbers from a multivariate Gaussian distribution, whose covariance matrix is

$$C = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

  Execute the task by first generating a vector $\vec{Y}$ of n independent normal random variables with 0 mean and unit variance, then using the Cholesky decomposition of C to derive $\vec{X}$.