Problems on *statistics and data analysis (MVComp2)*

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

Problem sheet 10

To be handed in during the exercise group on 11.July.2016

1. *extreme value statistics* *(10 points)*

We've already encountered the Gram-Charlier distribution,

$$p(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\times\left[1+\frac{\kappa_3}{3!\sigma^3}H_3\left(\frac{x}{\sigma}\right)+\frac{\kappa_4}{4!\sigma^4}H_4\left(\frac{x}{\sigma}\right)\right]\mathrm{d}x \tag{I}$$

together with the expression for the cumulative distribution,

$$P(x) = \int_{-\infty}^{x}\mathrm{d}x'\;p(x') = \Phi\left(\frac{x}{\sigma}\right) - \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\left[\frac{\kappa_3}{3!\sigma^2}H_2\left(\frac{x}{\sigma}\right)+\frac{\kappa_4}{4!\sigma^3}H_3\left(\frac{x}{\sigma}\right)\right] \tag{II}$$

which are both written in terms of Hermite-polynomials,

$$H_1(x) = x,\quad H_2(x) = x^2-1,\quad H_3(x) = x^3-3x,\quad H_4(x) = x^4-6x^2+3. \tag{III}$$

The Gram-Charlier distribution is parameterised by the cumulants κ_3 and κ_4 and is able to reflect weak non-Gaussianity. Please derive the maximum distribution $p_+(x)\mathrm{d}x$ and the minimum distribution $p_-(x)\mathrm{d}x$ in n draws. Then, provide plots for the cases

(a) $\kappa_3 = 10^{-2}$, $\kappa_4 = 0$, and

(b) $\kappa_3 = 0$, $\kappa_4 = 10^{-3}$,

for $n = 10^3$ and show that the extreme value distributions are sensitive to small deviations from Gaussianity. You may set $\sigma = 1$. The cumulative distribution for a Gaussian is given by the Φ-function,

$$\Phi\left(\frac{x}{\sigma}\right) = \frac{1}{2}\left(1+\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right)\right), \tag{IV}$$

written in terms of the standard error function,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x\mathrm{d}t\;\exp(-t^2). \tag{V}$$

2. *python exercise* *(10 points)*

Please generate samples with n Gaussian distributed random numbers with zero mean and unit variance, of which you keep the largest and the smallest value. Confirm that the distribution of the maxima and minima corresponds to what you expect from the extreme value distribution for $n = 10$ and $n = 100$.

3. *extreme values from the uniform distribution* *(10 points)*

A random variable x follows the uniform distribution $p(x) = \mathrm{d}x/2$ in the range $x = -1 \ldots +1$.

(a) Please derive the extreme value statistics $p_+(x)\mathrm{d}x$ and $p_-(x)\mathrm{d}x$ for n samples.

(b) Show the symmetry $p_+(x) = p_-(-x)$.

(c) Please derive the scaling of the maximum position of $p_\pm(x)\mathrm{d}x$ with n.

(d) Would a similar scaling hold for the expectation values $\int \mathrm{d}x\, x p_\pm(x)$?

4. *Gumbel-distribution* *(10 points)*

The standard Gumbel-distribution $p(x)\mathrm{d}x$ has the form

$$p(x) = \exp\left(-x - \exp\left(-x\right)\right). \tag{VI}$$

Please show that the moments $\langle x^n \rangle$ can be written as

$$\langle x^n \rangle = \int \mathrm{d}x\, x^n p(x) = \int_0^\infty \mathrm{d}z\ \ln^n(z) \exp(-z) \tag{VII}$$

after suitable substitution of a variable z for x, and that this expression corresponds to

$$\langle x^n \rangle = (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}z^n} \Gamma(z)\bigg|_{z=1}, \tag{VIII}$$

with Euler's Γ-function, $\Gamma(x) = \int_0^\infty \mathrm{d}t\, t^{x-1} \exp(-t)$.