Statistical methods – an introduction (SS 2016)

Problem set 3

Problem 1 [3 points] *Expectation value by Taylor expansion*

To very low accuracy, the expectation value of a function (with respect to a given distribution) can be approximated by

$$E(f(x)) \approx f(\hat{x}) \tag{1}$$

where $\hat{x}$ is the expectation value of x. This approximation can be improved considerably, by using higher moments.

a) Perform a Taylor expansion of $f(x)$ around $\hat{x}$, until fourth order.

b) Calculate the corresponding expectation value, and express the powers of $(x - \hat{x})$ in terms of σ, γ_1 and γ_2 values of the distribution. Show that Eq. (1) corresponds to *first order* accuracy.

c) Let
$$f(x) = 1 + \sin x$$
Approximate $E(f(x))$ according to b) with respect to the exponential distribution, $g(x) = \exp(-x)$ with $x \geq 0$. Compare your approximation with the exact value,
$$E_g(f(x)) = \int_0^\infty f(x)g(x)\mathrm{d}x = 3/2.$$
What would be the result if you use only the first order approximation, Eq. (1)?

Problem 2 [1.5 points] *Expectation value und variance of a sample mean*

Assume a random variable x to be distributed according to a certain pdf, with well-defined expectation value $E(\mathrm{x})$ and standard deviation $\sigma(\mathrm{x})$. A sample of N independent variates x_i is drawn from the distribution, and the sample (= arithmetic) mean $\bar{\mathrm{x}}$ is calculated.

Show from a *simple* calculation (roughly two times one line; more complicate derivations will be discarded) that $E(\bar{\mathrm{x}}) = E(\mathrm{x})$, and that $\sigma(\bar{\mathrm{x}}) = \frac{1}{\sqrt{N}}\sigma(\mathrm{x})$.

Problem 3 [3 points] *Characteristic functions*

a) Let $\mathrm{y} = a\mathrm{x} + b$, with $a, b \in \mathbb{R}$ and $a \neq 0$.
Express the characteristic function $\Phi_y(t)$ in terms of $\Phi_x(t')$, where t' depends on t. Here, Φ_y and Φ_x shall denote the characteristic functions with respect to the *distribution* of y and x, respectively.
(Hint: What is the relation between $g(y(x))\mathrm{d}y$ and $f(x)\mathrm{d}x$, if y is distributed according to g, and x according to f?)

b) Use the characteristic function of the Poisson distribution (see script p. 101),

$$\Phi_P(t) = \exp\bigl[\lambda(e^{it} - 1)\bigr],$$

with parameter λ, to calculate the expectation value of the distribution.

c) Repeat problem 3b), but now by using the corresponding cumulant(s). Calculate also the variance of the Poisson distribution from the corresponding cumulant(s), and check whether your results are correct.

Problem 4 [4.5 points] *Cumulants and the Central Limit Theorem*

a) Prove that the cumulants κ_n are homogeneous of degree n, i.e., that for a random variable x and an arbitrary constant $c \in \mathbb{R}$,

$$\kappa_n(c\mathrm{x}) = c^n \kappa_n(\mathrm{x})$$

Hint: Use one of the results from problem 3a), namely that if $\mathrm{y} = a\mathrm{x}$, then $\Phi_y(t) = \Phi_x(at)$

b) The pdf of a normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\bigl(-\frac{(x-\mu)^2}{2\sigma^2}\bigr),$$

with expectation value μ and standard deviation σ. By using the basic definition, it is straightforward to show that the corresponding characteristic function is given by

$$\Phi_f(t) = \exp\bigl(i\mu t - t^2\sigma^2/2\bigr).$$

Show that only the first two cumulants of the normal distribution have non-vanishing values. How do they read?

c) Assume, in close analogy to problem 2, a set of N independent, identically distributed (i.i.d.) random variables x_i, with expectation value $E(\mathrm{x_i}) = \mu$ and standard deviation $\sigma(\mathrm{x_i}) = \sigma$. Let

$$\mathrm{y} = \sum_{i=1}^{N} \mathrm{x}_i$$

Normalize the random variable y in such a way that the new variable y' has

$$Var(\mathrm{y}') = \sigma^2.$$

By using the general properties of cumulants, show that for $N \to \infty$ only the first two cumulants of the y'-distribution, $\kappa_1(\mathrm{y}')$ and $\kappa_2(\mathrm{y}')$, do *not* vanish. What are their values? What do you conclude from this problem?

Have fun, and much success!