Statistical methods – an introduction (SS 2016)

Problem set 7

Problem 1 [4 points] *Binormal variates*

Design an algorithm (no program, just the basic algorithm) to generate binormally distributed variates ($x_1 \ldots x_N$, $y_1 \ldots y_N$) with specified $\mu_x, \mu_y, \sigma_x, \sigma_y$ and correlation coefficient ρ.

For this objective, assume that a random number generator is available delivering normally distributed variates (reduced, i.e., with expectation value 0 and unit variance).

Note: There was a typo on page 114 (left) of the lecture notes. Please look into the new (corrected) version.

Problem 2 [4 points] *Central limit theorem*

a) Design an algorithm which generates random numbers which are approximately normally distributed with mean '0' and variance σ^2, by building an *appropriate* sum of N uniformly distributed (in [0,1]) r.v.'s. Show that the algorithm becomes particularly simple if $N = 12$.

b) Write a corresponding IDL/PYTHON routine, with input parameters N and σ, which displays the approximate solution (using `my_histogram.pro` or `my_hist.py`) together with the exact Gaussian.

c) Use your routine and print the results for $\sigma = 3$ and $N = 2, 4, 8, 16, 32$. Discuss your results briefly.

Problem 3 [4 points] *Density distribution in molecular clouds*
Star formation takes place in the most extreme density enhancements of molecular clouds. These are heavily affected by the motions induced by supersonic turbulence, and the structural characteristics of molecular clouds seem to agree with theoretical predictions and numerical simulations of such turbulence (see, e.g., Sect. 2.1 in McKee & Ostriker, 2007, ARA&A 45). One particularly important statistical property is the probability distribution of densities, which is expected to take a log-normal shape in isothermal, turbulent media not significantly affected by the self-gravity of gas. From hydro-simulations, Padoan et al. (1997, MNRAS 288) showed that the standard deviation σ_ρ with respect to the pdf of the *linear* density, $f_\rho(\rho)$, grows linearly with the average Mach number M as

$$\sigma_\rho = \rho_0 \beta M$$

where β is a factor of order unity, and ρ_0 is the expectation value of ρ with respect to f_ρ, i.e.,

$$\rho_0 = E(\rho) = \int \rho f_\rho(\rho) \mathrm{d}\rho$$

and

$$\sigma_\rho^2 = Var(\rho) = \int (\rho - \rho_0)^2 f_\rho(\rho) \mathrm{d}\rho.$$

The density ρ should be log-normally distributed according to

$$f(s, \mu, \sigma_s)\mathrm{d}s = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\big(-\frac{(s-\mu)^2}{2\sigma_s^2}\big)\mathrm{d}s$$

where $s = \ln(\rho/\rho_0)$, or, alternatively w.r.t. the linear density(-ratio),

$$f(x, \mu, \sigma_s)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma_s x} \exp\big(-\frac{(\ln x - \mu)^2}{2\sigma_s^2}\big)\mathrm{d}x,$$

with $x = \rho/\rho_0$.

a) Using the relations from manuscript page 125, show that $\hat{\rho} = \rho_0$ (w.r.t. f_ρ) implies $\mu = -\sigma_s^2/2$. Hint1: Use the 'conservation of probabilities', $f_\rho(\rho)\mathrm{d}\rho = f(s)\mathrm{d}s = f(x)\mathrm{d}x$, here without the absolute value (why?). Hint2: No explict integration is required to solve this problem (though with such an integration the result can be found as well).

b) With the same strategy and using the result from a), show that $\sigma_s^2 = \ln(1 + \beta^2 M^2)$.

From a) and b) it follows that the pdf for the linear density can be described via

$$f_\rho(\rho)\mathrm{d}\rho = \frac{1}{\sqrt{2\pi}\sigma_s \rho} \exp\big(-\frac{(\ln(\rho/\rho_0) + \sigma_s^2/2)^2}{2\sigma_s^2}\big)\mathrm{d}\rho,$$

and only depends on ρ, ρ_0 and σ_s (according to the expression provided in b)).

Have fun, and much success!