Homework 1, Statistical Mechanics: Concepts and applications
2016/17 ICFP Master (first year)

Werner Krauth, Maurizio Fagotti, Olga Petrova

This exercise will not be graded, but please hand it in on Wednesday,

14 September 2016, at the start of the lecture.

NB: Sorry for the posting delay due to the current and network problems at ENS around 7 Sept 2016

(Dated: September 11, 2016)

I. CHEBYCHEV INEQUALITY: VARIATIONS ON A THEME

The Chebychev inequality:

$$P(|X - E(X)| > \epsilon) \leq \frac{\mathrm{Var}(X)}{\epsilon^2} \tag{1}$$

is one of the fundamental achievements in probability theory, and it is of great importance in statistics. In eq. (1), $\mathrm{Var}(X)$ denotes the variance of the distribution and $E(X)$ its expectation (mean value).

1. State all the conditions on the probability distribution π_ξ for the Chebychev inequality to hold. For example, does it hold for discrete distributions (a sum of δ functions), for distributions with infinite expectation yet finite variance, etc?

2. Review the proof given in Lecture 1. Can the Chebychev inequality be "sharp", that is:

   (a) Can there be a distribution π_ξ where, for some ϵ, one has

   $$P(|X - E(X)| > \epsilon) = \frac{\mathrm{Var}(X)}{\epsilon^2} \tag{2}$$

   (note the "=" sign in eq. (2) instead of the "$\leq$" in eq. (1)). If so, construct this probability distribution. Otherwise, explain why this is not possible.

   (b) Can there be a distribution π, where the inequality of eq. (1) is sharp for *all* real ϵ? If so, construct this probability distribution. Otherwise, explain why this is not possible.

II. RÉNYI'S FORMULA FOR THE SUM OF UNIFORM RANDOM NUMBERS, VARIATIONS

In tutorial 1, you derived Rényi's formula for the sum of uniform random numbers between -1 and 1:

$$\pi_n(x) = \begin{cases} \frac{1}{2^n(n-1)!}\sum_{k=0}^{\left[\frac{n+x}{2}\right]}(-1)^k\binom{n}{k}(n+x-2k)^{n-1} & \text{for } |x| < n \\ 0 & \text{else} \end{cases} \tag{3}$$

1. Compute the variance of the distribution of eq. (3) for $n = 1$, that is for uniform random numbers between -1 and 1.

2. Compute the variance of Rényi's distribution for general n (Hint: this can be computed in 1 minute, if you use a result presented in the lecture).

3. Implement eq. (3) in a computer program for general n. For your convenience, you will find such a computer program on WK's website. This program also computes $P(X > \epsilon)$. Download this program and run it (in Python2). Notice that you may change the value of n in this program.

4. Modify the program (plot) so that it compares $P_n(X > \epsilon)$ to the upper limit given by the Chebychev inequality (Attention: you may modify Chebychev's inequality to take into account that $\pi_n(x)$ is symmetric around $x = 0$). Comment.

5. Modify the program (plot) so that it compares $P_n(X > \epsilon)$ to the Cantelli inequality:

$$P(X - E(X) > \epsilon) \leq \frac{\mathrm{Var}(X)}{\mathrm{Var}(X) + \epsilon^2} \tag{4}$$

   (note that there are now no absolute values). Comment.

6. Modify the program so that it compares $P(X > \epsilon)$ to Hoeffding's inequality. Hoeffding's inequality considers a probability distribution with zero expectation and $a_i \leq X_i < b_i$ (we will later take constant bounds a and b, but in fact, they may depend on i). For every $t > 0$, it states:

$$P(\sum_{i=1}^{n} X_i \geq \epsilon) \leq \exp(-t\epsilon)\prod_{i=1}^{n}\exp\left[t^2(b_i - a_i)^2/8\right]. \tag{5}$$

   Is Hoeffding's inequality *always* sharper than the Chebychev inequality, that is, is Hoeffding with the best value of t better than Chebychev for all ϵ? What is the asymptotic behavior for $\epsilon \to \infty$ behavior of Hoeffding's inequality, and why does it satisfy such a stringent bound if the Chebychev inequality does not achieve it? Return a plot that contains, next to $\pi_n(x)$ and its integral $P_n(X > \epsilon)$, the comparison with Chebychev, Cantelli, and Hoeffding.

III. LÉVY DISTRIBUTIONS, TWO SIMPLE DEMONSTRATION

In lecture 1 and tutorial 1, we discussed and derived Lévy distributions: Universal (stable) distributions that have infinite variance. A good example for producing such random variables is from uniform random numbers between 0 and 1, $\text{ran}(0, 1)$ taken to a power $-1 < \gamma < -0.5$. Such random numbers are distributed according to a distribution

$$\pi_\xi(x) = \begin{cases} \frac{\alpha}{x^{1+\alpha}} & \text{for } 1 < x < \infty \\ 0 & \text{else} \end{cases} \tag{6}$$

where $\alpha = -1/\gamma$ (you may check this by doing a histogram, and read up on this in SMAC book).

1. Is the probability distribution of eq. (6) normalized for $\gamma = -0.8$ (that is $\alpha = 1.25$), is it normalized for $\gamma = -0.2$ (that is $\alpha = 5$)?

2. What is the expectation of the probability distribution for the above two cases, and what is the variance?

3. Write a (two-line) computer program for generating the sum of 1000 random numbers with $\gamma = -0.2$, and plot the empirical histogram of this distribution (that is, generate 1000 times the sum of 1000 such random numbers. Interpret what you observe. For your convenience, you may find a closely related program on WK's website. Modify it so that it solves the problem at hand, and adapt the range in the drawing routine. Produce output and discuss it.

4. Write a (two-line) computer program for generating the sum of 1000 random numbers with $\gamma = -0.8$, and plot the empirical histogram of this distribution. Interpret what you observe. For your convenience, please take the closely related program from WK's website. Modify it so that it solves the problem at hand, and adapt the range in the drawing routine. Produce output and discuss it.