Classical Statistical Models - Spring 2016

Bielefeld University

Lecture/Tutorial: Wolfgang Unger Office: E6-118 wunger@physik.uni-bielefeld.de

Exercise Nr. 1

Discussion on April 25th, 15:00-16:00, Room C01-148

1) Probability Distributions *(2+4 points)*

a) Compute the moments of the Poisson distribution μ_i for all $1 \leq i \leq 3$. Does your result generalize to all $i > 0$?

b*) Compute the hypergeometric distribution via a computer programm based on random numbers, and use the corresponding urn model: draw without replacement n balls from an urn with a total of N balls, with K white and $N - K$ black balls. The hypergeometric probability distribution for drawing k white balls is $P(\underline{X} = k) = f(k; N, K; n) \equiv \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$. Crosscheck that your statistical computer experiment reproduces this probability distribution, for $N = 100$ fixed and various K, n.

c) Prove that in the limit $N, K \to \infty$, $\frac{K}{N} = p$ fixed, the hypergeometric distribution approximates the Binomial distribution. Why is that expected?

2) Absorbing Markov Chains *(2+4 points)*

a) Prove that for an absorbing Markov chain, the probability that the process will be absorbed in some absorbing state, starting from any transient state, is 1.

b) Consider the example of a drunkard's walk for n blocks. The absorbing states are s_0 and s_n, the transient states are s_i, $i = 1, \ldots n-1$, and the non-zero transition probabilities are $p_{i,i+1} = p_{i,i-1} = \frac{1}{2}$, and $p_{00} = p_{nn} = 1$. Prove that the expected time to absorption, starting at state s_i, is equal to $i(n-i)$

3) Ehrenfest Urn Model *(2+2+4 points)*

The Ehrenfest urn model is a simple model for the diffusion of gases. Assume two urns and N balls. The Markov states are given by the number of balls in the first urn, i.e. we have $N+1$ states s_i, $i = 0 \ldots N$. (s_0: all balls in second urn, s_N: all balls in first urn). At each step, one of the N balls is chosen by random and moved to the other urn. For $N = 4$, the transition matrix is

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

a) Generalize the tranisition matrix for arbitrary N.

b) Prove that the transition matrix is ergodic, but not regular.

c*) Implement a Markov Chain Monte Carlo simulation for this model, and determine the equilibrium distribution for $N = 4, 8, 16$.

Andrey Andreyevich Markov
(June 14, 1856 - July 20, 1922) was a Russsian mathematician.

Russian mathematician who helped to develop the theory of stochastic processes, especially those called Markov chains. Based on the study of the probability of mutually dependent events, his work has been developed and widely applied in the [physical,] biological and social sciences.

As a child Markov had health problems and used crutches until he was 10 years old. In 1874 he enrolled at the University of St. Petersburg (now St. Petersburg State University), where he earned a bachelor's degree (1878), a master's degree (1880), and a doctorate (1884).

In 1883, as his station in life improved, he married his childhood sweetheart, the daughter of the owner of the estate that his father managed. Markov became a professor at St. Petersburg in 1886 and a member of the Russian Academy of Sciences in 1896. Although he officially retired in 1905, he continued to teach probability courses at the university almost to his deathbed. While his early work was devoted to number theory and analysis, after 1900 he was chiefly occupied with probability theory. As early as 1812 the French mathematician Pierre-Simon Laplace had formulated the first central limit theorem, which states, roughly speaking, that probabilities for almost all independent and identically distributed random variables converge rapidly (with sample size) to the area under an exponential function. [...] In 1887 Markov's teacher Pafnuty Chebyshev outlined a proof of a generalized central limit theorem. Using a different approach, Chebyshev's student Aleksandr Lyapunov proved the theorem under weakened hypotheses in 1901. Eight years later Markov succeeded in proving the general result rigorously using Chebyshev's method. While working on this problem, he extended both the law of large numbers (which states that the observed distribution approaches the expected distribution with increasing sample size) and the central limit theorem to certain sequences of dependent random variables forming special classes of what are now known as Markov chains. These chains of random variables have found numerous applications in modern physics. One of the earliest applications was to describe Brownian motion, the small, random fluctuations or jiggling of small particles in suspension. Another frequent application is to the study of fluctuations in stock prices, generally referred to as random walks.

[from Encyclopaedia Britannica Online Academic Edition]