Institut für Biologische Physik
J. Berg
U. Michel, S. Kleinbölting

Many phenomena in nature are stochastic, i.e. not deterministic but governed by randomness. For example, it is impossible to predict the trajectory of a micron-sized colloidal particle in an aqueous solution. We may repeat an experiment many times over though and make statements about relative frequencies of events. The limit of these frequencies when the number of repetitions goes to infinity is called probability. Stochastics is the mathematical discipline to describe random systems. This sheet is a refresher in basic stochastics.

1 Combinatorics

a) How many permutations are there for M numbered (i.e. distinguishable) particles?

b) In how many ways can one distribute M distinguishable particles on $N \geq M$ sites?

c) In how many ways can one distribute M *in*distinguishable particles on $N \geq M$ sites without double occupancy?

d) How many ways are there to place M indistinguishable particles into $N \leq M$ boxes? Double-occupancy and empty boxes are allowed.

2 Probability distributions

A *random variable* X is a quantity that takes values in a state space Ω randomly. The measurements of a colloidal particles's postition are an example. One discriminates two cases for the state space.

i.) $\Omega = \{x_1, x_2, \ldots\}$ is countable (finitely or infinitely)

Let p_i or $p(x_i)$ denote the probility that the system is in state i:

$$p_i = p(x_i) = \mathrm{Prob}(X = x_i).$$

From the definition based on relative frequencies it follows immediately: $p_i \geq 0$ and $\sum_i p_i = 1$. We call $\{p_i\}$ a *discrete probability distribution.*

ii.) $\Omega = \mathbb{R}^d$. Now the probability that X is contained in a volume V is given by:

$$\mathrm{Prob}(X \in V) = \int_V p(\mathbf{x}) \mathrm{d}^d x.$$

The function $p(\mathbf{x})$ is called *probability density* and we say X is distributed *absolutely continuous* with density $p(\mathbf{x})$. For example, in $d = 1$ the following notation is common: $p(x)\mathrm{d}x = \mathrm{Prob}(X \in [x, x + \mathrm{d}x))$. Again, from the definition of probability: $p(\mathbf{x}) \geq 0$ (up to isolated points) and $\int_{\mathbb{R}^d} p(\mathbf{x}) \mathrm{d}^d x = 1$.

The *expectation value* of a random variable is defined as

$$\langle X \rangle = \sum_i p_i x_i, \quad \text{or} \quad \langle X \rangle = \int_{\mathbb{R}^d} \mathbf{x} p(\mathbf{x}) \mathrm{d}^d x$$

and its *variance*, which is a measure for the spread of a distribution, as

$$\mathrm{Var}(X) = \langle (X - \langle X \rangle)^2 \rangle$$

a) Calculate expectation and variance of the binomial distribution where

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \;\text{ für }\; k = 0, 1, \dots, n; n \in \mathbb{N}, p \in (0,1)$$

b) Calculate expectation and variance of the Poisson distribution where

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda} \;\text{ für }\; k \in \mathbb{N}_0, \lambda > 0$$

c) Calculate expectation and variance of the normal distribution with density

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \;\text{ für }\; x, \mu \in \mathbb{R}, \sigma > 0$$

d) Calculate expectation and variance of the exponential distribution with density

$$p(x) = \lambda e^{-\lambda x} \;\text{ für }\; x > 0, \lambda > 0$$

3 The weak law of large numbers Consider a collection of random variables $X_1, X_2, \dots$ with distributions $p_1(x_i), p_2(x_i), \dots$ or densities $p_1(x), p_2(x), \dots$ respectively. Consider the vector of the joined random variable $\mathbf{X} = (X_1, X_2, \dots, X_N)$ where $N \in \mathbb{N}$. The random variables are called *independent*, if the distribution of $\mathbf{X}$ factorises, i.e. if

$$\text{Prob}(\mathbf{X} = (x_1, x_2, \dots, x_N)) = p_1(x_1)p_2(x_2)\dots p_N(x_N),$$

or

$$p((x_1, x_2, \dots, x_N)) = p_1(x_1)p_2(x_2)\dots p_N(x_N)$$

respectively.

a) Show that for $a, b \in \mathbb{R}$: $\text{Var}(a(X+b)) = a^2 \text{Var}(X)$.

b) Show that for independent random variables with finite variance:

$$\left\langle \sum_{i=1}^N X_i \right\rangle = \sum_{i=1}^N \langle X_i \rangle \;\text{ and }\; \text{Var}\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^N \text{Var}(X_i)$$

(the first equality is valid even for dependent variables).

c) Assume now that all random variables have the same expectation μ and variance σ^2. Use parts a) and b) to show that the arithmetic mean converges to the expectation:

$$\left\langle \left(\frac{1}{N} \sum_{i=1}^N X_i - \mu \right) \right\rangle = 0 \;\text{ and }\; \lim_{N\to\infty} \text{Var}\left(\frac{1}{N} \sum_{i=1}^N X_i - \mu \right) = 0$$