Freie Universität Berlin
Fachbereich Physik
April 24th 2018

Prof. Dr. Roland Netz
Douwe Bonthuis
Jan Daldrop
Philip Loche

Statistical Physics and Thermodynamics (SS 2018)

Problem sheet 2

Hand in: Friday, May 11 during the lecture (note: one week later than usual)

http://www.physik.fu-berlin.de/en/einrichtungen/ag/ag-netz/lehre

1 Characteristic Functions (8 points)

Consider the normalized probability distributions

$$P_1(x) = A_1\delta(x - x_0), \tag{1}$$
$$P_2(x) = A_2 \begin{cases} 1 & \text{if } 1 < |x| < 2 \\ 0 & \text{else,} \end{cases} \tag{2}$$
$$P_3(x) = A_3\, \mathrm{e}^{-a|x|}, \tag{3}$$
$$P_4(x) = A_4\, \frac{1}{1+x^2}, \tag{4}$$

where $\delta(x)$ is the Dirac-delta distribution, $x, x_0, a \in \mathbb{R}$ and $a > 0$.

a) Determine the normalization constants $A_1, \dots A_4$. **(1 point)**

b) Calculate the characteristic function (moment generating function) $G(k) = \langle e^{-ikx}\rangle = \int_{-\infty}^{\infty} P(x)\mathrm{e}^{-ikx}\mathrm{d}x$ and the cumulant generating function $\ln G(k)$ for all distributions. **(4 points)**

Hint: For calculating $G_4(k)$ you can either use the "Residue theorem" or the relation betwee $P_3(x)$ and $G_3(k)$.

c) Calculate the moments $\langle x^n\rangle$ and the cumulants $\langle x^n\rangle_c$ for $n = 1, 2$ from the generating functions for all the distributions from part a). What are the relations among moments, cumulants, mean and variance? **(3 points)**

2 Central limit theorem for binomial distributions (8 points)

In this exercise, we will derive that the probability density function describing the outcome of a binomial experiment approaches the normal distribution for a large number of experiments. Suppose we do an experiment with two possible outcomes: A, occurring with probability p, and B, occurring with probability q. When we do N experiments, we are interested in the total number of times X_N that the outcome of the experiment is A.

a) Write down the probability $P(X_N = k)$ in terms of p, q, N and k and calculate the average $\langle X_N\rangle$ and the variance $\langle X_N^2\rangle - \langle X_N\rangle^2$. **(1 point)**

Hint: To calculate $\langle X_N\rangle = \sum_{k=0}^{N} kP(X_N = k)$, define $l = k-1$, $M = N-1$ and rewrite the sum. A similar procedure can be used for the variance.

b) Consider the ratio of the probabilities of having $X_N = k+1$ and $X_N = k$, and show that it can be written as

$$\frac{P(X_N = k+1)}{P(X_N = k)} = \frac{(N-x)p}{(k+1)q}. \quad \textbf{(2 points)} \tag{5}$$

c) For convenience, we shift and normalize X_N by the following linear transformation $Z_N = (X_N - Np)/\sqrt{Npq}$. Show that the ratio of part b) can be written as

$$\frac{P(Z_N = z + 1/\sqrt{Npq})}{P(Z_N = z)} = \frac{1 - zp/\sqrt{Npq}}{1 + zq/\sqrt{Npq} + q/(Npq)}, \tag{6}$$

with z being the continuous variable $z = (k - Np)/\sqrt{Npq}$. **(1 point)**

d) Assume there exists a smooth probability density function $f(z)$ such that for large N the probability $P(Z_N = z)$ can be approximated with the differential $P(Z_N = z) \approx f(z)dz$. Defining $\Delta = 1/\sqrt{Npq}$ and taking the logarithm of the expression of part c), show that

$$\ln f(z + \Delta) - \ln f(z) = \ln(1 - zp\Delta) - \ln(1 + zq\Delta + q\Delta^2). \quad \textbf{(1 point)}$$

e) Divide both sides of part d) by Δ and take the limit of $N \to \infty$ to show that

$$\frac{d \ln f(z)}{dz} = -z. \tag{7}$$

Integrate this expression and normalize to find an expression for $f(z)$. **(2 points)**

f) Use your result for $f(z)$ to approximate the probability of exactly (i) 3 heads in 5 coin tosses and (ii) 60 heads in 100 coin tosses ($p = q = 1/2$ in both cases). How accurate are your results? **(1 point)**

3 Estimating probabilities (4 points)

a) Suppose we draw, with equal probabilities, n integer numbers k_i from the list $\{-1, 0, 1\}$, and calculate $X_n = \sum_{i=1}^{n} k_i$. Sketch the probability distributions of X_n for $n = 1$, 2 and 3. **(2 points)**

b) b) Given a large number of experiments of which the continuous outcomes x follow an unknown probability distribution with finite mean μ and variance σ^2. Estimate the probability of (i) finding a value within 1 standard deviation from the mean and (ii) finding a value more than 2 standard deviations from the mean. **(2 points)**

Hint: Use a table of the integrated normal distribution based on the value $z = (x - \mu)/\sigma$, which can be found anywhere in statistics books or on the internet.