33-341 — Thermal Physics I

Department of Physics, Carnegie Mellon University, Fall Term 2018, Deserno

Problem sheet #5

15. Averaging helps. Until it doesn't *(6 points, due on Monday)*

While measuring some observable in an experiment, you collect a sequence of numbers $\{X_1, X_2, X_3, \ldots\}$, which you may think of as independent random variables drawn from some p-density $p_X(x)$. You are interested in the mean of that distribution. To reduce the noise in each individual measurement, you decide to calculate the arithmetic average of N measurements:

$$\overline{X}_N := \frac{1}{N}\sum_{i=1}^{N} X_i \ . \tag{1}$$

Being the sum of N random variables, *this is still a random variable*! It is obvious that $\langle \overline{X}_N \rangle_{p_X} = \langle X_i \rangle_{p_X}$, but you hope that $\overline{X}_N$ scatters less around that mean than each individual X_i. Let us see whether your hope is justified.

1. Assume the $\{X_i\}$ come from a normal distribution with zero mean and unit variance, $p(x) = \mathrm{e}^{-x^2/2}/\sqrt{2\pi}$. Extend the program from problem (12) so that it creates a histogram of $\overline{X}_N$ values, for various values of N. That is, draw N random numbers X_i, calculate $\overline{X}_N$, repeat this process `trials` number of times, and make a histogram of the `trials` $\overline{X}_N$ values thus calculated. Plot this histogram together with the p-density of a single X_i. Is the histogram more narrow?

2. Same as (1), but assume that the $\{X_i\}$ instead come from a Cauchy-Lorentz distribution $p(x) = \frac{1}{\pi}\frac{1}{1+x^2}$.

16. Error propagation *(7 points, due on Wednesday)*

Consider a collection of random variables $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, from which we calculate a function of interest, $F(\boldsymbol{X})$. Assume we know all expectation values $\langle X_i \rangle$ and all covariances $C_{ij} := \mathrm{Cov}(X_i, X_j) = \big\langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \big\rangle$.

1. Taylor-expand F to *second order* around $\langle \boldsymbol{X} \rangle$. Now take the average and show how $\langle F(\boldsymbol{X}) \rangle$ differs from $F(\langle \boldsymbol{X} \rangle)$.

2. For the special case of $n = 1$ and a convex F, show that your result is consistent with Jensen's inequality!

3. The variance of F is given by $\sigma_F^2 = \big\langle [F(\boldsymbol{X}) - \langle F(\boldsymbol{X}) \rangle]^2 \big\rangle$. Simplify this by replacing F with its *first order* Taylor expansion. Show further that if all X_i are uncorrelated, you end up with the "standard" formula for error propagation!

17. Sum of random variables and the characteristic function *(7 points, due on Friday)*

Assume you have two continuous independent random variable X and Y with p-densities $p_X(x)$ and $p_Y(y)$, respectively. Define the random variable $Z = X + Y$ and call its p-density $p_Z(z)$.

1. Show that the p-density $p_Z(z)$ is the *convolution* of the p-densities $p_X(x)$ and $p_Y(y)$:

$$p_Z(z) = \int_{-\infty}^{\infty} \mathrm{d}x \; p_X(x)\, p_Y(z-x) \ . \tag{2}$$

2. The *characteristic function* $\tilde{p}_X(k)$ of a random variable X is defined to be the Fourier transform of its p-density $p_X(x)$:

$$\tilde{p}_X(k) := \left\langle \mathrm{e}^{\mathrm{i}kX} \right\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \; p_X(x)\, \mathrm{e}^{\mathrm{i}kx} \ . \tag{3}$$

   a) If the moments of X are written as $\mu_n := \langle X^n \rangle$, show that the Taylor expansion of $\tilde{p}_X(k)$ is given by

$$\tilde{p}_X(k) = 1 + \mathrm{i}\mu_1 k - \frac{1}{2}\mu_2 k^2 - \frac{1}{6}\mathrm{i}\mu_3 k^3 + \frac{1}{24}\mu_4 k^4 \cdots \ . \tag{4}$$

   b) Show that the characteristic function of $Z = X + Y$ is the product of the characteristic functions of X and of Y:

$$\tilde{p}_Z(k) = \tilde{p}_X(k)\, \tilde{p}_Y(k) \ . \tag{5}$$