33-765 — Statistical Physics

Department of Physics, Carnegie Mellon University, Spring Term 2019, Deserno

Problem sheet #3

8. An application of the transformation law of probability densities *(4 points)*

Let X and Y be two independent continuous random variables on $\mathbb{R}$ which are both distributed according to a Gaussian with mean zero and variance one. Define the new random variable $Z = X/Y$. What is the probability density of Z?

9. Simplifying ratios *(4 points)*

Let's say we have two *independent* continuous random variables X and Y. Consider the following claim:

$$\left\langle \frac{X}{Y} \right\rangle \stackrel{?}{=} \frac{\langle X \rangle}{\langle Y \rangle} . \tag{1}$$

Is this true? If you think yes, then prove it. If you think it's not true, then prove that it's not true.

10. Transformation theorem for probability densities—prescription strength *(6 points)*

Let $X > 0$ be a continuous random variable with probability density $p_X(x) = \mathrm{e}^{-x}$. What is the p-density $p_Y(y)$ of $Y = \sin X$? *Hints: You will find it useful to distinguish the cases $y > 0$ and $y < 0$. Plot the argument of the delta-function in the transformation law and figure out, how to systematically enumerate its zeros. Try to write the answer as nicely as you can!*

11. Characteristic functions and the amazing Central Limit Theorem *(6 points)*

The Fourier transform $\tilde{p}(k)$ of a probability density (henceforth: "p-density") $p(x)$ is also called the "*characteristic function*":

$$\tilde{p}(k) \;=\; \left\langle \mathrm{e}^{\mathrm{i}kx} \right\rangle \;=\; \int \mathrm{d}x \; p(x) \, \mathrm{e}^{\mathrm{i}kx} \qquad \left[\text{and hence:} \quad p(x) = \frac{1}{2\pi} \int \mathrm{d}k \; \tilde{p}(k) \, \mathrm{e}^{-\mathrm{i}kx} \right] . \tag{2}$$

1. Let X be a random variable whose p-density $p_X(x)$ has moments $\mu_n = \langle X^n \rangle$. If these moments μ_n exists, prove that

$$\mu_n \;=\; \mathrm{i}^{-n} \left[\frac{\partial^n}{\partial k^n} \tilde{p}_X(k) \right]_{k=0} . \tag{3}$$

2. If $\tilde{p}_{aX}(k)$ is the characteristic function of the random variable aX (with some $a \in \mathbb{R}$), show that $\tilde{p}_{aX}(k) = \tilde{p}_X(ak)$.

3. Let X and Y be two *independent* random variables with p-densities $p_X(x)$ and $p_Y(y)$. Let $p_{X+Y}(z)$ be the p-density of $Z = X + Y$. Prove that $p_{X+Y}(z) = \int \mathrm{d}x \; p_X(x) \, p_Y(z - x)$ and that $\tilde{p}_{X+Y}(k) = \tilde{p}_X(k) \, \tilde{p}_Y(k)$.

4. Let $X_1, \ldots, X_n$ be n *independent* random variables with *identical distribution* $p_X(x)$, which has mean μ_1 and *finite variance* $\sigma^2 = \mu_2 - \mu_1^2$. Consider the centered and normalized random variables $Y_i = (X_i - \mu_1)/\sigma$ (which obviously have zero mean and unit variance) and the new (and seemingly curiously normalized) sum random variable

$$Z_n \;=\; \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i \;=\; \frac{X_1 + X_2 + \cdots + X_n \;-\; n\,\mu_1}{\sigma\,\sqrt{n}} . \tag{4}$$

If $\tilde{p}_{Z_n}(k)$ is the characteristic function of (the p-density of) Z_n, show that in the limit of large n you get

$$\lim_{n\to\infty} \tilde{p}_{Z_n}(k) \;=\; \mathrm{e}^{-\frac{1}{2}k^2} \qquad \text{and hence} \qquad p_{Z_n}(x) \;\longrightarrow\; \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} \;\equiv\; \mathcal{G}_{(0,1)}(x) . \tag{5}$$

Hint: The proof follows swiftly from what you've worked out so far; you will also need a cute representation for the exponential function: $\lim_{n\to\infty}[1 + x/n + o(x/n)]^n = \mathrm{e}^x$, where $o(z)$ is any term that satisfies $\lim_{z\to 0} o(z)/z = 0$.

This is (a version of) the amazing **Central Limit Theorem**: The distribution of the $\sqrt{n}$-normalized sum of the centered X_i becomes a Gaussian with zero mean and unit variance, *independent of the actual distribution of the* X_i (as long as their variance is finite). It also implies that for increasing n the p-density of the arithmetic mean, $\overline{X_i} = \frac{1}{n}(X_1 + \cdots + X_n) = \mu + \frac{\sigma}{\sqrt{n}} Z_n$ converges against, $\mathcal{G}_{(\mu,\sigma/\sqrt{n})}(x)$, a Gaussian centered around μ with variance $\sigma/\sqrt{n}$. Hence, the *error of the mean* also becomes Gaussian and decreases like $1/\sqrt{n}$. The Central Limit Theorem explains why the Gaussian distribution is "normal": It naturally emerges once you do averaging. This is also why it appears all over the place.