Lecture Notes on Concepts of Statistics in Physics

Zihang Wang

September 2023

Random Variable, Probability, and Expectation Value

A random variable is an abstraction of a quantity that depends on random events, and it can be discrete or continuous. For example, the toss of a coin can result in one of two possible outcomes: "Heads" ($X = 1$) or "Tails" ($X = 0$), and $X = \{0, 1\}$ is an abstraction of this outcome. Conversely, the stopping direction of a spinning top represents a continuous random variable. In practice, we always assume the observations, therefore the associated random variables, are independent and identically distributed (i.i.d.). This assumption simplifies many statistical analyses and gives rise to important concepts and applications such as the Central Limit Theorem (CLT), hypothesis testing, Markov chains, Bayesian inference.

As an example, in quantum physics, random variables are often treated as observables, such as energy (E), position (x), momentum (k), and spin (s_z), each of which is typically associated with underlying eigenstates, denoted as $|E\rangle$, $|x\rangle$, $|k\rangle$, and $|s_z\rangle$, respectively. Loosely speaking, the density matrix, constructed from an eigenstate serves as a "random variable", e.g. $x \to |x\rangle\langle x|$.

To demonstrate this idea, let's consider a general quantum state,

$$|\Psi\rangle = \int \mathrm{dx}\psi(x)|x\rangle, \quad \int \mathrm{dx}|\psi(x)|^2 = 1, \quad \rho = |\Psi\rangle\langle\Psi|, \tag{1}$$

where $\psi(x)$ is the position wavefunction, and the second equality ensures the overall normalization. For each random variable, $|x\rangle\langle x|$, the corresponding probability density at particular position can be calculated from the trace, e.g. $\mathrm{Tr}(\rho|x\rangle\langle x|) = |\psi(x)|^2$ where $\mathrm{Tr}(\dots) = \int \mathrm{dx}\langle x|(\dots)|x\rangle$. The probability of observing a particle within in the position window $x \pm \mathrm{dx}$ is given by $P(x) = |\psi(x)|^2\mathrm{dx}$.

It is particularly useful when dealing with ensembles of quantum systems,

$$\rho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j|, \tag{2}$$

where now ρ represents a mixed state composed of multiple pure states $|\Psi_j\rangle\langle\Psi_j|$ with associated probability p_j.

An analogous concept can be found in statistical physics, which deals with the thermodynamic properties of large ensembles of microstates. In this context, macroscopic physical properties such as temperature, particle number, and energy can be viewed as random variables, depending on ensembles. Consider an ideal gas composed of N non-interacting (fermionic, bosonic, or classical) particles, and we assume there are n_0 numbers of particles that fall within a specific energy range, $E_0 \pm \epsilon$. The probability of finding a particle within the above energy window is given by the ratio,

$$\left(\frac{n_0}{N}\right)_{\text{fermi}} = \frac{1}{e^{(E_0-\mu)/kT}+1}, \quad \left(\frac{n_0}{N}\right)_{\text{bose}} = \frac{1}{e^{(E_0-\mu)/kT}-1}, \tag{3}$$

and

$$\left(\frac{n_0}{N}\right)_{\text{classical}} = \frac{e^{-(E_0-\mu)/kT}}{\sum_j e^{-(E_j-\mu)/kT}}, \tag{4}$$

where μ is the chemical potential. The expectation value of an observable is then given by,

$$\langle A\rangle = \int_{-\infty}^{\infty} A(E)\left(\frac{n_0}{N}\right)_X \mathrm{d}E, \tag{5}$$

and X labels the statistics.

Central Limit Theorem (CLT)

The CLT states that when a sufficiently large number of independent and identically distributed (i.i.d.) random variables, $X_1, X_2, \ldots X_N$ from a given (unknown) population, as the sample szie $N \to \infty$, the distribution of the sample mean (also the sum) $\bar{X} = \sum_j X_j/N$ will approximate a normal (Gaussian) distribution, regardless of the original population's distribution,

$$f(\bar{X}) \sim \frac{1}{\sqrt{2\pi\sigma^2/N}}\mathrm{Exp}\left(-\frac{(\bar{X}-\mu)^2}{2\sigma^2/N}\right), \tag{6}$$

where $f(\bar{X})$ is the probability density function (pdf) of the sample mean, μ is the mean of each X_j, and σ^2 is the variance of each X_j. We should also note that it only applies to those distributions who have well-defined mean and variance, (e.g. Cauchy distribution).

CLT has important applications in physics. For example, let's say we want to track the how far and how fast on average a particle move over time T. We record the random variable X_{t_i} (relative position with respect to $X_{t_{i-1}}$) at each time interval t_i,

$$\bar{X}_T = \sum_j X_{t_i}, \quad \bar{V}_T = \frac{\bar{X}_T}{T}. \tag{7}$$

The CLT states that as $T \to \infty$, the distribution of both $\bar{X}_T$ and $\bar{V}_T$ (average position and velocity of the particle) will approach a normal distribution. This phenomenon is known as Brownian motion, a random process governed by the diffusion equation.

It is worth noting the distinction between the solution of the diffusion equation and CLT. The diffusion equation is derived from the principle of conservation of particle number and provides information about the local density of particles at a specific time. In contrast, the random variable used in the CLT characterizes the (average) positions of a single particle. These two concepts, while related, serve distinct purposes in modeling random processes. Although we will not discuss in details, Ergodicity is a concept that relates the behavior of a single-particle time average to many-particle ensemble average. In an ergodic system, the behavior of a single trajectory (e.g., the motion of a single particle) over time is representative of the behavior of the entire ensemble of particles.

Lorentzian Distribution in Physics

Resonance behavior is critical for understanding the energy dissipation and absorption of a system. As an example, let's consider a general dissipative, driven harmonic oscillator in both time and frequency domain,

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \gamma \frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = F e^{i\omega_f t}, \quad -\omega^2 \tilde{x} + i\gamma\omega\tilde{x} + \omega_0^2 \tilde{x} = F\delta(\omega - \omega_f), \tag{8}$$

where $\tilde{x}$ and x are related via Fourier transform. The natural and driven frequency are labelled by ω_0 and ω_f. γ is the damping factor. Then, we have the solution in frequency and time domain,

$$\tilde{x} = \frac{F\delta(\omega - \omega_f)}{\omega_0^2 - \omega^2 + i\gamma\omega}, \quad x = \mathcal{F}^{-1}(\tilde{x}) = \frac{e^{i\omega_f t}}{\omega_0^2 - \omega_f^2 + i\gamma\omega_f} = A e^{i\omega_f t + i\Sigma}, \tag{9}$$

where,

$$A = \frac{1}{\sqrt{(\omega_0^2 - \omega_f^2)^2 + \gamma^2\omega_f^2}}, \quad \tan(\Sigma) = \frac{-\gamma\omega_f}{\omega_0^2 - \omega_f^2}. \tag{10}$$

We finally have the energy absorption per cycle as a Lorentzian,

$$\begin{aligned} E &= AF\mathrm{Re}\left\{ \int e^{-i\omega_f t} \frac{\mathrm{d}}{\mathrm{d}t} e^{i\omega_f t + i\Sigma} \mathrm{d}t \right\} = i\omega_f AF\mathrm{Re}\left\{ \int_0^{\pi/\omega_f} e^{i\Sigma} \mathrm{d}t \right\} \\ &= -\frac{\pi}{\omega_f} AF\omega_f \sin(\Sigma) = \pi A^2 F\gamma\omega_f = F\pi \frac{\gamma\omega_f}{(\omega_0^2 - \omega_f^2)^2 + \gamma^2\omega_f^2}, \end{aligned} \tag{11}$$

and we use the fact $\sin(\Sigma) = -A\gamma\omega_f$. The universal absorption profile of the driven harmonic oscillator, characterized by the Lorentzian distribution, holds profound implications across various fields of physics. Although we will not discuss it in detail, it's worth mentioning that this represents a specific case of the spectral function, a fundamental idea in many-body physics.