04. Random Variables: Concepts

Gerhard Müller
University of Rhode Island, gmuller@uri.edu

Follow this and additional works at: http://digitalcommons.uri.edu/nonequilibrium_statistical_physics

Part of the Physics Commons

Abstract
Part four of course materials for Nonequilibrium Statistical Physics (Physics 626), taught by Gerhard Müller at the University of Rhode Island. Entries listed in the table of contents, but not shown in the document, exist only in handwritten form. Documents will be updated periodically as more entries become presentable.

Recommended Citation
http://digitalcommons.uri.edu/nonequilibrium_statistical_physics/4

This Course Material is brought to you for free and open access by the Physics Course Materials at DigitalCommons@URI. It has been accepted for inclusion in Nonequilibrium Statistical Physics by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.
4. Random Variables: Concepts

- Probability distributions
- Moments, variance, standard deviation
- Moment expansion and characteristic function
- Cumulant expansion
- Factorial moments and cumulants, generating function
- Multivariate distributions [nln7]
- Transformation of random variables
- Propagation of statistical uncertainty [nex24]
- Chebyshev’s inequality [nex6]
- Law of large numbers [nex7]
- Binomial, Poisson, and Gaussian distribution [nln8]
- Binomial to Poisson distribution [nex15]
- De Moivre - Laplace limit theorem [nex21]
- Central limit theorem [nln9]
- Multivariate Gaussian distribution
- Robust probability distributions [nex19]
- Stable probability distributions [nex81]
- Exponential distribution [nln10]
- Waiting time problem [nln11]
- Pascal distribution [nex22]
Multivariate Distributions

Let $X = (X_1, \ldots, X_n)$ be a random vector variable with $n$ components.

Joint probability distribution: $P(x_1, \ldots, x_n)$.

Marginal probability distribution:

$$P(x_1, \ldots, x_m) = \int dx_{m+1} \cdots dx_n P(x_1, \ldots, x_n).$$

Conditional probability distribution:

$$P(x_1, \ldots, x_n) = P(x_1, \ldots, x_m | x_{m+1}, \ldots, x_n) P(x_{m+1}, \ldots, x_n).$$

Moments: $\langle X_{m_1} \cdots X_{m_n} \rangle = \int dx_1 \cdots dx_n x_1^{m_1} \cdots x_n^{m_n} P(x_1, \ldots, x_n)$.

Characteristic function: $\Phi(k) = \langle e^{ikX} \rangle$.

Moment expansion: $\Phi(k) = \sum_{m_1, \ldots, m_n} \frac{(ik_1)^{m_1} \cdots (ik_n)^{m_n}}{m_1! \cdots m_n!} \langle X_{m_1}^{m_1} \cdots X_{m_n}^{m_n} \rangle$.

Cumulant expansion: $\ln \Phi(k) = \sum_{m_1, \ldots, m_n} \frac{t}{m_1! \cdots m_n!} \langle \langle X_{m_1}^{m_1} \cdots X_{m_n}^{m_n} \rangle \rangle$.

Covariance matrix: $\langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle$.

(i = j: variances,  j ≠ j: covariances).

Correlations: $C(X_i, X_j) = \frac{\langle X_i X_j \rangle}{\sqrt{\langle X_i^2 \rangle \langle X_j^2 \rangle}}$.

Statistical independence of $X_1, X_2$: $P(x_1, x_2) = P_1(x_1)P_2(x_2)$.

Equivalent criteria for statistical independence:

- all moments factorize: $\langle X_1^{m_1} X_2^{m_2} \rangle = \langle X_1^{m_1} \rangle \langle X_2^{m_2} \rangle$;
- characteristic function factorizes: $\Phi(k_1, k_2) = \Phi_1(k_1)\Phi_2(k_2)$;
- all cumulants $\langle \langle X_1^{m_1} X_2^{m_2} \rangle \rangle$ with $m_1m_2 \neq 0$ vanish.

If $\langle X_1 X_2 \rangle = 0$ then $X_1, X_2$ are called uncorrelated.

This property does not imply statistical independence.
Transformation of statistical uncertainty.

From a given stochastic variable $X$ with probability distribution $P_X(x)$ we can calculate the probability distribution of the stochastic variable $Y = f(X)$ via the relation

$$P_Y(y) = \int dx \, P_X(x) \delta (y - f(x)).$$

Show by systematic expansion that if $P_X(x)$ is sufficiently narrow and $f(x)$ sufficiently smooth, then the mean values and the standard deviations of the two stochastic variables are related to each other as follows:

$$\langle Y \rangle = f(\langle X \rangle), \quad \sigma_Y = |f'(\langle X \rangle)| \sigma_X.$$

Solution:
Chebyshev’s inequality

Chebyshev’s inequality is a rigorous relation between the standard deviation $\sigma_X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}$ of the random variable $X$ and the probability of deviations from the mean value $\langle X \rangle$ greater than a given magnitude $a$.

$$P[|x - \langle X \rangle|^2 > a^2] \leq \left( \frac{\sigma_X}{a} \right)^2$$

Prove Chebyshev’s inequality starting from the following relation, commonly used for the transformation of stochastic variables (as, for example, in [nex24]):

$$P_Y(y) = \int dx \, \delta(y - f(x)) \, P_X(x) \text{ with } f(x) = (x - \langle X \rangle)^2.$$ 

Solution:
Law of large numbers

Let \( X_1, \ldots, X_N \) be \( N \) statistically independent random variables described by the same probability distribution \( P_X(x) \) with mean value \( \langle X \rangle \) and standard deviation \( \sigma_X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} \). These random variables might represent, for example, a series of measurements under the same (controllable) conditions. The law of large numbers states that the uncertainty (as measured by the standard deviation) of the stochastic variable \( Y = (X_1 + \cdots + X_N)/N \) is

\[
\sigma_Y = \frac{\sigma_X}{\sqrt{N}}.
\]

Prove this result.

Solution:
Consider a set of $N$ independent experiments, each having two possible outcomes occurring with given probabilities.

<table>
<thead>
<tr>
<th>events</th>
<th>$A + B = S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>probabilities</td>
<td>$p + q = 1$</td>
</tr>
<tr>
<td>random variables</td>
<td>$n + m = N$</td>
</tr>
</tbody>
</table>

**Binomial distribution:**

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}.$$  
Mean value: $\langle n \rangle = Np$.

Variance: $\langle \langle n^2 \rangle \rangle = Npq$. [nex15]

In the following we consider two different asymptotic distributions in the limit $N \to \infty$.

**Poisson distribution:**

Limit #1: $N \to \infty$, $p \to 0$ such that $Np = \langle n \rangle = a$ stays finite [nex15].

$$P(n) = \frac{a^n}{n!} e^{-a}.$$  
Cumulants: $\langle \langle n^m \rangle \rangle = a$.

Factorial cumulants: $\langle \langle n^m \rangle \rangle_f = a \delta_{m,1}$. [nex16]

Single parameter: $\langle n \rangle = \langle \langle n^2 \rangle \rangle = a$.

**Gaussian distribution:**

Limit #2: $N \gg 1$, $p > 0$ with $Np \gg \sqrt{Npq}$.

$$P_N(n) = \frac{1}{\sqrt{2\pi \langle \langle n^2 \rangle \rangle}} \exp \left( -\frac{(n - \langle n \rangle)^2}{2\langle \langle n^2 \rangle \rangle} \right).$$  
Derivation: DeMoivre-Laplace limit theorem [nex21].

Two parameters: $\langle n \rangle = Np$, $\langle \langle n^2 \rangle \rangle = Npq$.

Special case of central limit theorem [nlh9].
Binomial to Poisson distribution

Consider the binomial distribution for two events $A, B$ that occur with probabilities $P(A) \equiv p$, $P(B) = 1 - p \equiv q$, respectively:

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n},$$

where $N$ is the number of (independent) experiments performed, and $n$ is the stochastic variable that counts the number of realizations of event $A$.

(a) Find the mean value $\langle n \rangle$ and the variance $\langle n^2 \rangle$ of the stochastic variable $n$.

(b) Show that for $N \to \infty$, $p \to 0$ with $Np \to a > 0$, the binomial distribution turns into the Poisson distribution

$$P_\infty(n) = \frac{a^n}{n!} e^{-a}.$$

Solution:
De Moivre–Laplace limit theorem.

Show that for large $Np$ and large $Npq$ the binomial distribution turns into the Gaussian distribution with the same mean value $\langle n \rangle = Np$ and variance $\langle \langle n^2 \rangle \rangle = Npq$:

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad \rightarrow \quad P_N(n) \simeq \frac{1}{\sqrt{2\pi\langle \langle n^2 \rangle \rangle}} \exp \left( -\frac{(n - \langle n \rangle)^2}{2\langle \langle n^2 \rangle \rangle} \right).$$

Solution:
Central Limit Theorem

The central limit theorem is a major extension of the law of large numbers. It explains the unique role of the Gaussian distribution in statistical physics.

Given are a large number of statistically independent random variables \(X_i, i = 1, \ldots, N\) with equal probability distributions \(P_X(x_i)\). The only restriction on the shape of \(P_X(x_i)\) is that the moments \(\langle X_i^n \rangle = \langle X^n \rangle\) are finite for all \(n\).

Goal: Find the probability distribution \(P_Y(y)\) for the random variable \(Y = (X_1 - \langle X \rangle + \cdots + X_N - \langle X \rangle)/N\).

\[
P_Y(y) = \int dx_1 P_X(x_1) \cdots \int dx_N P_X(x_N) \delta \left( y - \frac{1}{N} \sum_{i=1}^N [x_i - \langle X \rangle] \right).
\]

Characteristic function:

\[
\Phi_Y(k) \equiv \int dy e^{iky} P_Y(y), \quad P_Y(y) = \frac{1}{2\pi} \int dk e^{-iky} \Phi_Y(k).
\]

\[
\Rightarrow \Phi_Y(k) = \int dx_1 P_X(x_1) \cdots \int dx_N P_X(x_N) \exp \left( i \frac{k}{N} \sum_{i=1}^N [x_i - \langle X \rangle] \right)
\]

\[
= \left[ \Phi \left( \frac{k}{N} \right) \right]^N,
\]

\[
\Phi \left( \frac{k}{N} \right) \equiv \int dx e^{i(k/N)(x-\langle X \rangle)} P_X(x) = \exp \left( -\frac{1}{2} \left( \frac{k}{N} \right)^2 \langle \langle X^2 \rangle \rangle + \cdots \right)
\]

\[
= 1 - \frac{1}{2} \left( \frac{k}{N} \right)^2 \langle \langle X^2 \rangle \rangle + O \left( \frac{k^3}{N^3} \right),
\]

where we have performed a cumulant expansion to leading order.

\[
\Rightarrow \Phi_Y(y) = \left[ 1 - \frac{k^2 \langle \langle X^2 \rangle \rangle}{2N^2} + O \left( \frac{k^3}{N^3} \right) \right]^N \xrightarrow{N \to \infty} \exp \left( -\frac{k^2 \langle \langle X^2 \rangle \rangle}{2N} \right).
\]

where we have used \(\lim_{N \to \infty} (1 + z/N)^N = e^z\).

\[
\Rightarrow P_Y(y) = \sqrt{\frac{N}{2\pi \langle \langle X^2 \rangle \rangle}} \exp \left( -\frac{Ny^2}{2\langle \langle X^2 \rangle \rangle} \right) = \frac{1}{\sqrt{2\pi \langle \langle Y^2 \rangle \rangle}} e^{-y^2/2\langle \langle Y^2 \rangle \rangle}
\]

with variance \(\langle \langle Y^2 \rangle \rangle = \langle \langle X^2 \rangle \rangle/N\)

Note that regardless of the form of \(P_X(x)\), the average of a large number of (independent) measurements of \(X\) will be a Gaussian with standard deviation \(\sigma_Y = \sigma_X/\sqrt{N}\).
Robust probability distributions

Consider two independent stochastic variables $X_1$ and $X_2$, each specified by the same probability distribution $P_X(x)$. Show that if $P_X(x)$ is either a Gaussian, a Lorentzian, or a Poisson distribution,

(i) $P_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$,  
(ii) $P_X(x) = \frac{1}{\pi} \frac{a}{x^2 + a^2}$,  
(iii) $P_X(n) = \frac{a^n}{n!} e^{-a}$.

then the probability distribution $P_Y(y)$ of the stochastic variable $Y = X_1 + X_2$ is also a Gaussian, a Lorentzian, or a Poisson distribution, respectively. What property of the characteristic function $\Phi_X(k)$ guarantees the robustness of $P_X(x)$?

Solution:
Stable probability distributions

Consider $N$ independent random variables $X_1, \ldots, X_N$, each having the same probability distribution $P_X(x)$. If the probability distribution of the random variable $Y_N = X_1 + \cdots + X_N$ can be written in the form $P_Y(y) = P_X(y/c_N + \gamma_N)/c_N$, then $P_X(x)$ is stable. The multiplicative constant must be of the form $c_N = N^{1/\alpha}$, where $\alpha$ is the index of the stable distribution. $P_X(x)$ is strictly stable if $\gamma_N = 0$.

Use the results of [nex19] to determine the indices $\alpha$ of the Gaussian and Lorentzian distributions, both of which are both strictly stable. Show that the Poisson distribution is not stable in the technical sense used here.

Solution:
Exponential distribution

Busses arrive randomly at a bus station. The average interval between successive bus arrivals is $\tau$.

$f(t)dt$: probability that the interval is between $t$ and $t + dt$.

$P_0(t) = \int_t^\infty dt' f(t')$: probability that the interval is larger than $t$.

Relation: $f(t) = -\frac{dP_0}{dt}$.

Normalizations: $P_0(0) = 1$, $\int_0^\infty dt f(t) = 1$.

Mean value: $\langle t \rangle \equiv \int_0^\infty dt t f(t) = \tau$.

Start the clock when a bus has arrived and consider the events $A$ and $B$.

Event $A$: the next bus has not arrived by time $t$.

Event $B$: a bus arrives between times $t$ and $t + dt$.

Assumptions:

1. $P(AB) = P(A)P(B)$ (statistical independence).

2. $P(B) = cdt$ with $c$ to be determined.

Consequence: $P_0(t + dt) = P(\overline{A}B) = P(A)P(\overline{B}) = P_0(t)[1 - cdt]$.

$\Rightarrow \frac{d}{dt}P_0(t) = -cP_0(t) \Rightarrow P_0(t) = e^{-ct} \Rightarrow f(t) = ce^{-ct}$.

Adjust mean value: $\langle t \rangle = \tau \Rightarrow c = \frac{1}{\tau}$.

**Exponential distribution:** $P_0(t) = e^{-t/\tau}$, $f(t) = \frac{1}{\tau} e^{-t/\tau}$.

Find the probability $P_n(t)$ that $n$ busses arrive before time $t$.

First consider the probabilities $f(t')dt'$ and $P_0(t - t')$ of the two statistically independent events that the first bus arrives between $t'$ and $t' + dt'$ and that no further bus arrives until time $t$.

Probability that exactly one bus arrives until time $t$:

$P_1(t) = \int_0^t dt' f(t')P_0(t - t') = \frac{t}{\tau} e^{-t/\tau}$.

Then calculate $P_n(t)$ by induction.

**Poisson distribution:** $P_n(t) = \int_0^t dt' f(t')P_{n-1}(t - t') = \frac{(t/\tau)^n}{n!} e^{-t/\tau}$. 
Waiting Time Problem

Busses arrive more or less randomly at a bus station. Given is the probability distribution $f(t)$ for intervals between bus arrivals.

Normalization: $\int_{0}^{\infty} dt f(t) = 1$.

Probability that the interval is larger than $t$: $P_0(t) = \int_{t}^{\infty} dt' f(t')$.

Mean time interval between arrivals: $\tau_B = \int_{0}^{\infty} dt t f(t) = \int_{0}^{\infty} dt P_0(t)$.

Find the probability $Q_0(t)$ that no arrivals occur in a randomly chosen time interval of length $t$.

First consider the probability $P_0(t'+t)$ for this to be the case if the interval starts at time $t'$ after the last bus arrival. Then average $P_0(t'+t)$ over the range of elapsed time $t'$.

$\Rightarrow Q_0(t) = c \int_{0}^{\infty} dt' P_0(t'+t)$ with normalization $Q_0(0) = 1$.

$\Rightarrow Q_0(t) = \frac{1}{\tau_B} \int_{t}^{\infty} dt' P_0(t')$.

Passengers come to the station at random times. The probability that a passenger has to wait at least a time $t$ before the next bus is then $Q_0(t)$:

Probabilty distribution of passenger waiting times:

$g(t) = -\frac{d}{dt}Q_0(t) = \frac{1}{\tau_B} P_0(t)$.

Mean passenger waiting time: $\tau_P = \int_{0}^{\infty} dt t g(t) = \int_{0}^{\infty} dt Q_0(t)$.

The relationship between $\tau_B$ and $\tau_P$ depends on the distribution $f(t)$. In general, we have $\tau_P \leq \tau_B$. The equality sign holds for the exponential distribution.
Pascal distribution.

Consider the quantum harmonic oscillator in thermal equilibrium at temperature $T$. The energy levels (relative to the ground state) are $E_n = n\hbar\omega$, $n = 0, 1, 2, \ldots$

(a) Show that the system is in level $n$ with probability

$$P(n) = (1 - \gamma)^n \gamma^n, \quad \gamma = \exp(-h\omega/k_B T).$$

$P(n)$ is called Pascal distribution or geometric distribution.

(b) Calculate the factorial moments $\langle n^m \rangle_f$ and the factorial cumulants $\langle(n^m)\rangle_f$ of this distribution.

(c) Show that the Pascal distribution has a larger variance $\langle n^2 \rangle$ than the Poisson distribution with the same mean value $\langle n \rangle$.

Solution: