University of Rhode Island

# DigitalCommons@URI

10-19-2015

# 03. Elements of Probability Theory with Applications

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

Abstract

Part three 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.
Updated with version 2 on 5/3/2016.

# Contents of this Document [ntc3]

3. Elements of Probability Theory with Applications

# Elements of Set Theory [nln4]

A **set** $S$ is a collection of **elements** $\epsilon_i$: $\quad S = \{\epsilon_1, \epsilon_2, \ldots\}$.

All elements of **subset** $A$ are also elements of $S$: $\quad A \subset S$.

The **empty set** $\emptyset$ contains no elements.

If $S$ contains $n$ elements then the number of subsets is $2^n$.

Hierarchy of subsets: $\quad \emptyset \subset A \subset A \subset S$.
    **Transitivity**: If $C \subset B$ and $B \subset A$ then $C \subset A$.
    **Equality**: $A = B$ iff $A \subset B$ and $B \subset A$.

**Union**: $\quad C = A + B$ or $C = A \cup B$.
    All elements of $S$ that are contained in $A$ or in $B$ or in both.

**Intersection**: $\quad C = AB$ or $C = A \cap B$.
    All elements of $S$ that are contained in $A$ and in $B$.

Union and intersection are **commutative**, **associative**, and **distributive**:
$$A + B = B + A, \quad AB = BA;$$
$$(A + B) + C = A + (B + C), \quad (AB)C = A(BC);$$
$$A(B + C) = AB + AC.$$

Some consequences:
$$A + A = A, \quad A + \emptyset = A, \quad A + S = S; \quad AA = A, \quad A\emptyset = \emptyset, \quad AS = A.$$

**Mutually exclusive** subsets have no common elements: $AB = \emptyset$.

**Partition** $\mathcal{P} = [A_1, A_2, \ldots]$ of $S$ into mutually exclusive subsets:
    $S = A_1 + A_2 + \ldots$ with $A_i A_j = \emptyset$ for $i \neq j$.

The **complement** $\bar{A}$ of subset $A$ has all elements of $S$ that are not in $A$.

Some consequences:
$$A + \bar{A} = S, \quad A\bar{A} = \emptyset, \quad \bar{\bar{A}} = A, \quad \bar{S} = \emptyset, \quad \bar{\emptyset} = S.$$

**DeMorgan's law**: $\quad \overline{A + B} = \bar{A}\bar{B}, \quad \overline{AB} = \bar{A} + \bar{B}$.

**Duality principle**:
    A set identity is preserved if all sets are replaced by their complements.

## [nex88] Set identities

Use De Morgan's law and the duality principle to derive the set identity on the right from the set identity on the left.

$$A(B + C) = AB + AC \ \Rightarrow \ A + BC = (A + B)(A + C).$$

**Solution:**

# Elementary Probabilities [nln42]

Probabilities can be defined (i) axiomatically, (ii) from relative frequencies of events, or (iii) by assignment to elementary events based on one of two criteria [e.g. Papoulis 1991, Sec. 1-2]:

- symmetry (somewhat circularly),
- principle of insufficient reason (somewhat dubiously).

Ambiguities may arise if the range of events is

- discrete and infinite,
- continuous (such as in Bertrand's paradox [nln41]).

In general, probability densities change under a transformation of variables. Particular assignments are often hard to justify.

In classical equilibrium statistical mechanics, uniform probability densities are assigned to canonical coordinates. This choice is most readily justifiable in the case of action-angle coordinates. Canonical transformations leave probability densities invariant. However, the existence of action-angle coordinates is limited to integrable systems. Only a tiny minority of many-body systems are integrable.

# [nex94] Probabilities of complements, unions, and intersections

Given are the probability axioms (i) $P(A) \geq 0$, (ii) $P(S) = 1$, and (iii) $P(A + B) = P(A) + P(B)$ if $AB = \emptyset$, where $S$ is the sample space and $A, B$ are events.

Derive from these axioms the following simple theorems regarding complements, unions, and intersections of events:

(a) $P(\emptyset) = 0$, (b) $P(\bar{A}) = 1 - P(A)$, (c) $P(A + B) = P(A) + P(B) - P(AB)$.

The notions $A + B$, $AB$, $\bar{A}$, $\emptyset$, mean union, intersection, complement, and empty set, respectively.

**Solution:**

# Joint and Conditional Probabilities

Joint probability: $P(AB)$   (event $A$ and event $B$)

Conditional probability: $P(A|B)$   (event $A$ if event $B$)

Relations: $P(AB) = P(A|B)P(B) = P(B|A)P(A)$

Simple consequences:

- If $A \subset B$ then $P(A|B) = P(A)/P(B)$
- If $B \subset A$ then $P(A|B) = 1$

Conditional probabilities satisfy probability axioms [nex90].

Bayes' theorem: $P(A|B) = P(B|A)\dfrac{P(A)}{P(B)}$

Applications:

▷ Successive random picks [nex91]

▷ Heads or tails [nex93]

▷ Quantity and quality [nex76]

▷ Diagnosis of a rare disease [nex77]

▷ Event or complement [nex9]

## [nex90] Conditional probability

Given the simple probabilites $P(A), P(B)$ of two events $A, B$ and their joint probability $P(AB)$ defined as the probability of their intersection, the conditional probabilities $P(A|B), P(B|A)$ are then defined via

$$P(AB) = P(A|B)P(B) = P(B|A)P(A).$$

Show that the conditional probability $P(A|B) = P(AB)/P(B)$ is indeed a probability in the formal sense by showing that it satisfies the probability axioms: (i) $P(A|B) \geq 0$, (ii) $P(S|B) = 1$, (iii) $P(A + C|B) = P(A|B) + P(C|B)$ if $AC = \emptyset$.

**Solution:**

# Elements of Probability Theory [nln43]

Transcription from set theory:

- set $\rightarrow$ sample space
- subset $\rightarrow$ event
- element $\rightarrow$ elementary event

Sample space $S$: set of all possible outcomes in an experiment.

Event $A \subset S$: possible outcome of experiment.

Probability axioms [Kolmogorov 1933]:

- $P(A) \geq 0$ for all $A \subset S$,
- $P(S) = 1$,
- $P(A + B) = P(A) + P(B)$ if $AB = \emptyset$.

Some immediate consequences [nex94]

- $P(\emptyset) = 0$,
- $P(\bar{A}) = 1 - P(A)$,
- $P(A + B) = P(A) + P(B) - P(AB)$.

Under idealized circumstances the sample space is divisible into elementary, mutually exclusive, events to which equal probabilities can be assigned for reasons of symmetry:

$$S = A_1 + \cdots + A_N \text{ with } A_i A_j = \emptyset \text{ for } i \neq j \text{ and } P(A_i) = \frac{1}{N}.$$

# Statistical Independence

Case of two events $A, B$.

Criterion: $P(AB) = P(A)P(B)$

Simple consequences:

- $P(A|B) = P(A), \quad P(B|A) = P(B)$
- $P(A\bar{B}) + P(A)P(\bar{B}), \quad P(\bar{A}B) = P(\bar{A})P(B), \quad P(\bar{A}\bar{B}) = P(\bar{A})P(\bar{B})$

Case of three events $A, B, C$.

Criteria: pairwise statistical independence is not sufficient!

- $P(AB) = P(A)P(B)$
- $P(AC) = P(A)P(C)$
- $P(BC) = P(B)P(C)$
- $P(ABC) = P(A)P(B)P(C)$

Applications:

- ▷ Subtlety of statistical independence [nex1]
- ▷ Random train connections [nex92]
- ▷ Random inkjet printer [nex10]

# Statistical Uncertainty and Information

An experiment has $n$ possible outcomes that occur with probabilities $P_1, P_2, \ldots, P_n$.

Properties that must be satisfied by any quantitative measure of *uncertainty*:

1. The uncertainty is a function of the probabilities of all possible outcomes: $\Sigma = \Sigma(P_1, P_2, \ldots, P_n)$.

2. The uncertainty is symmetric under all permutations of the $P_i$.

3. The maximum uncertainty occurs if all $P_i$ are equal.

4. The uncertainty is zero if one of the outcomes has probability $P_i = 1$.

5. The combined outcome of two independent experiments has an uncertainty equal to the sum of the uncertainties of the outcomes of each experiment.

$$\Rightarrow \ \Sigma(P_1, P_2, \ldots, P_n) = -\sum_{i=1}^{n} P_i \ln P_i = -\langle \ln P \rangle.$$

Information comes in messages: $A_1, A_2, \ldots$. A message carries information only if it contains some news, i.e. something not completely expected.

$P(A)$: probability that message $A$ is sent.
$I(A)$: information gain if message is indeed received.

The less likely the message, the greater the information gain if the message is received:

If $P(A) < P(B)$ then $I(A) > I(B)$, if $P(A) = 1$ then $I(A) = 0$.

If two independent messages are received, then the information gain is the sum of the information gains pertaining to each individual message:

$$P(AB) = P(A)P(B) \ \Rightarrow \ I(AB) = I(A) + I(B).$$

The information content of a message is equal to the change in (statistical) uncertainty at the receiver:

$$P_1, P_2, \ldots, P_n \xrightarrow{A} \bar{P}_1, \bar{P}_2, \ldots, \bar{P}_n \ \Rightarrow \ I(A) = \Sigma(P_1, P_2, \ldots, P_n) - \Sigma(\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_n)$$

Information as used here refers only to the scarcity of events. Any aspects of usefulness and meaningfulness are disregarded.

# [tex47] Statistical concept of uncertainty

An experiment has $n$ possible outcomes that occur with probabilities $P_1, \ldots, P_n$. The uncertainty about the outcome of the experiment is defined as

$$\Sigma(P_1, \ldots, P_n) = -\sum_{i=1}^{n} P_i \ln P_i.$$

(a) Prove that the maximum uncertainty occurs if all $P_i$ are equal.

(b) The $n^2$ combined outcomes of two independent experiments have probabilities $P_{ij} = P_i^I P_j^{II}$. Show that the uncertainty about the combined outcome of the two independent experiments is equal to the sum of the uncertainties of the outcomes of each experiment: $\Sigma(\{P_{ij}\}) = \Sigma(\{P_i^I\}) + \Sigma(\{P_i^{II}\})$.

**Solution:**

## [nex9] Event or complement? That is the question.

Ten people throw two dice each, once per minute, at ten tables. When any person hits a 12, he/she leaves the room. Find the probability that not everybody has left the room after one hour.

**Solution:**

## [nex91] Successive random picks

A bin contains three white balls and two red balls. We remove at random two balls in succession.
What is the probability that the first ball is white and the second is red?

**Solution:**

## [nex93] Heads or tails

We have two coins, one of which is fair and the other two-headed. Now we pick one coin at random and toss it twice. Heads shows both times. Find the probability that we had picked the fair coin.

**Solution:**

## [nex76] Quantity and Quality

A company has two plants producing identical cars. In a particular year, plant 1 produced 60%
of the cars, 35% of which were defective. Of the remaining 40% of cars produced in plant 2, only
25% were defective.
(a) What percentage of cars in that year were defective?
(b) What is the probability that a defective car produced in that year came from plant 1?

**Solution:**

## [nex77] Diagnosis of a rare deadly disease

A fatal disease is known to afflict one person in a million during his or her life time. There exists a test which detects the presence of the deadly virus in a blood sample with 99% reliability. With what probability is a person who tests positive indeed afflicted by the disease?

**Solution:**

## [nex1] Subtlety of statistical independence

(i) Given that events $A, B$ are statistically independent, $P(AB) = P(A)P(B)$, show that the event pairs $\bar{A}, B$ and $\bar{A}, \bar{B}$ are also statistically independent: $P(\bar{A}B) = P(\bar{A})P(B)$, $P(\bar{A}\bar{B}) = P(\bar{A})P(\bar{B})$.

(ii) Consider three events $A, B, C$ occurring with probabilities $P(A), P(B), P(C)$, respectively, and satisfying the relations $P(AB) = P(A)P(B)$, $P(BC) = P(B)P(C)$, $P(CA) = P(C)P(A)$. Show that these relations are compatible with the relation $P(ABC) \neq P(A)P(B)P(C)$, in which case the three events are not statistically independent.

**Solution:**

## [nex92] Random train connections

Trains $X$ and $Y$ reach a station at random and independently of each other between 8:00am and 8:20am. Train $X$ stops for 4 minutes and train $Y$ for 5 minutes.
(a) Find the probability that train $X$ arrives before train $Y$.
(b) Find the probability that the two trains meet at the station.
(c) Find the probability that train $Y$ pulls into the station while train $X$ is already there.


**Solution:**

## [nex10] Random inkjet printer

A number $N$ of black inkdots, all with area $A_D$, are printed randomly on a sheet of paper with area $A_S$. Assume that $N$ and $A_S$ are sufficiently large, so that edge effects can be neglected.

(a) Calculate the average blackened area $\langle A_B \rangle$ on the sheet as a function of $N$, $A_D$, and $A_S$.

(b) Show that for very large $N$ and finite $N A_D / A_S$ the asymptotic expression,

$$\langle A_B \rangle \rightsquigarrow A_S \left[ 1 - e^{-N A_D / A_S} \right],$$

ensues, telling us that some white area is likely to persist in the face of a multitudes of random inkdots.

**Solution:**

## [tex48] Statistical uncertainty and information

The number of bird species living on some continent is known to be 100. An ornithologist visits a small island off the coast of that continent to find out how many of the 100 bird species have migrated to the island.

One month after her arrival on the island she sends a first message to the Ornithological Society, stating that there exist only five of the 100 bird species on the island.

A month later she sends a second message stating that the relative abundance of the five bird populations identified previously is 80%, 10%, 5%, 3%, 2%.

Determine the numerical value of the information contained in each message.

**Solution:**

## [tex61] Information of sequenced messages

In which months of the year do Ellen, Nancy, and Susan have their birthdays? Three messages $X, Y, Z$ about their birthdays are received by persons $a, b, c$:

Person $a$ receives the messages in the sequence $X, Y, Z$.

Person $b$ receives the messages in the sequence $Z, X, Y$.

Person $c$ receives the messages in the sequence $Y, Z, X$.

The three messages are the following:

$X$: Nancy's birthday is in April.

$Y$: Ellen's birthday is in a later month than Nancy's birthday.

$Z$: Susan's birthday is in the same month as Ellen's birthday.

Find the numerical value of the information contained in the messages $X, Y, Z$ as received by each person $a, b, c$.

**Solution:**