## Contents

### Part I Foundation in Probability and Statistics  

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Probability</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Probability Models</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>Classical Inference I: Estimators</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>Classical Inference II: Optimization</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>Classical Inference III: Confidence Intervals and Statistical Tests</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>44</td>
</tr>
<tr>
<td>7</td>
<td>Bayesian Inference</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>49</td>
</tr>
</tbody>
</table>

### Part II Measurement Techniques

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Basic Measurements</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>Event Reconstruction</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>62</td>
</tr>
<tr>
<td>10</td>
<td>Correlation Functions</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>64</td>
</tr>
<tr>
<td>11</td>
<td>The Multiple Facets of Correlation Functions</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>70</td>
</tr>
</tbody>
</table>
12 Data Correction Methods 74
Exercises 74

Part III Simulation Techniques 79

13 Monte Carlo Methods 81
Exercises 81

14 Collision and Detector Modeling 84
Exercises 84
PART I

FOUNDATION IN PROBABILITY AND STATISTICS
2.1 Demonstrate the properties listed under Eq. (2.5).

Solution: By definition, one has $A \cup \tilde{A} = S$, $A \cap \tilde{A} = \emptyset$, and $P(\emptyset) = 1$. The sets $A$ and $\tilde{A}$ being disjoint, i.e., $A \cap \tilde{A} = \emptyset$, one has $P(\tilde{A}) = 1 - P(A)$, implying that $P(A \cup \tilde{A}) = 1$. If $A = 0$, then $P(A) = 0$, and if $A = S$, then $P(A) = P(S) = 1$ and is bound by construction. Other properties can be demonstrated similarly.

2.2 Verify that the notion of conditional probability satisfies the three axioms defining the notion of probability.

Solution: By definition, $P(A|B) = P(A \cap B)/P(B)$. Provided $B \neq 0$, one has $P(B) > 0$. If $A \cap B = 0$, then $P(A \cap B) = 0$. If $A \cap B = B$, then $P(A \cap B) = P(B)$, which entails $P(A|B) = 1$. One thus concludes $0 \leq P(A|B) \leq 1$.

2.3 Verify the expression Eq. (2.18) known as law of total probability.

Solution: By definition, $S = \cup_i A_i$ and $A_i \cap A_j = \emptyset$ for $i \neq j$. Assuming all subsets $A_i$ are non-empty, one has $P(A_i) > 0$. Also, since $B \subset S$, one has $B \cap S = B$. Replace $S$ by $\cup_i A_i$ and distribute the intersection. One gets $B = B \cap \cup_i A_i = \cup_i (B \cap A_i)$. Since the subsets $A_i$ are all mutually disjoint, so will the sets $B \cap A_i$, one then concludes that $P(B) = P(\cup_i (B \cap A_i)) = \sum_i P(B \cap A_i)$ or equivalently $P(B) = \sum_i P(B|A_i)P(A_i)$, which is the law of total probability given by Eq. (2.18).

2.4 Derive the expression for the variance given by Eq. (2.80).

Solution: By definition, $\mu_2 = \text{Var}[x] = E[(x - \mu)^2]$, one can then write:

\[
\mu_2 = \int dxp(x)(x - \mu)^2 = \int dxp(x)x^2 - 2\mu \int dxp(x)x + \mu^2 \int dxp(x).
\]

Since, by definition, $\int dxp(x) = 1$, $\int dxp(x)x = \mu$, and $\int dxp(x)x^2 \equiv E[x^2]$, one gets $\mu_2 = E[x^2] - \mu^2$, which corresponds to Eq. (2.80).

2.5 Show that the skewness, $\gamma_1$, can be equivalently defined as the ratio of the third cumulant $\kappa_3$ and the third power of the square root of the second cumulant $\kappa_2^{3/2}$.

Solution: Skewness is defined as $\gamma_1 = \mu_3/\sigma^3$. The third centered moment $\mu_3$ is defined as $\mu_3 = \int p(x)(x - \mu)^3 dx$. Expanding the cube, and carrying out the distributed integrals, one gets $\mu_3 = \lambda_3 - 3\mu_2^2 \mu + 2\mu^3$. The third cumulant is defined as $\kappa_3 = \frac{d^3}{dt^3} g(t)|_{t=0}$ with $g(t) = \ln[M(t)]$, and $M(t)$ is the generating function of the moments $\mu_k$ satisfying $\mu_k = \frac{d^k}{dt^k} M(t)|_{t=0}$. Calculation of the first three cumulants
show that the excess kurtosis of a sum of independent random variables, with identical variance, is equal to the sum of the kurtosis of these variables divided by \( n^2 \) (Eq. 2.93).

Solution: Excess kurtosis is defined as \( \gamma_2 = \mu_4/\sigma^4 - 3 \) where \( \sigma^2 \) and \( \mu_4 \) and are the 2nd and 4th centered moment of the variable. Let \( z = \sum_i x_i \). We will assume the variables \( x_i \) have non-zero means, \( \langle x_i \rangle = 0 \), and equal variances, \( \sigma_i^2 = \sigma^2 \). If the variables have non-zero means, the calculation could be simply done in terms of variables \( x_i' = x_i - \langle x_i \rangle \) which by construction have null mean. Let’s proceed iteratively. First calculate \( z^2 \):

\[
    z^2 = \left( \sum_i x_i \right)^2 = \sum_i x_i \sum_j x_j = \sum_i x_i^2 + \sum_{i \neq j} x_i x_j.
\]

Since the \( x_i \) are independent and their means are null, one has \( \langle x_i x_j \rangle = \langle x_i \rangle \langle x_j \rangle \). The above becomes

\[
    \langle z^2 \rangle = \sum_i \langle x_i^2 \rangle = n\sigma^2.
\]

Next calculate the third power of \( z \). Proceeding in the same fashion

\[
    z^3 = \sum_i x_i^3 + 3 \sum_{i \neq j} x_i^2 x_j + \sum_{i \neq j \neq k} x_i x_j x_k.
\]

And the 4th power is:

\[
    z^4 = \sum_i x_i^4 + 4 \sum_{i \neq j} x_i^3 x_j + 3 \sum_{i \neq j \neq k} x_i^2 x_j x_k + 6 \sum_{i \neq j \neq k \neq m} x_i x_j x_k x_m.
\]

Accounting for the null means and the fact all the variable have equal second moments \( \sigma^2 \), the expectation value of \( z^4 \) is thus

\[
    \langle z^4 \rangle = \sum_i \langle x_i^4 \rangle + 3 \sum_{i \neq j} \sigma^4,
\]

\[
    = \sum_i \langle x_i^4 \rangle + 3n(n - 1)\sigma^4
\]
Dividing by \( \langle z^2 \rangle \) and subtracting 3, one finds
\[
\frac{\langle z^4 \rangle}{\langle z^2 \rangle^2} - 3 = \sum_i \frac{\langle x_i^4 \rangle}{n_i^2 \sigma^2} - 3n(n-1)^2 \frac{\sigma^4}{n^2 \sigma^4},
\]
\[
= \frac{1}{n^2} \sum_i \left( \frac{\langle x_i^4 \rangle}{\sigma^4} - 3 \right),
\]
which is the anticipated result.

2.7 Derive the expression Eq. 2.60 for the density \( g(q) \) obtained for a function \( q(x) \) of continuous random variable \( x \) distributed according to a PDF \( f(x) \).

Solution: The probability of an event to take place with values between \( x \) and \( x+dx \) is \( f(x)dx \) and with values between \( q \) and \( q+dq \) is \( g(q)dq \). These two probabilities are equal because one deals with the same events. Expressing \( dq \) as \( dq = q(x+dx) - q(x) = \frac{dg}{dx} \bigg|_x dx \), one can then write
\[
g(q) = \left. \frac{f(x(q))}{dq/dx} \right|_x = f(x(q)) \left| \frac{dq^{-1}}{dx} \right|_x,
\]
which is the result we sought to demonstrate.

2.8 Calculate expressions for the density \( g(a)da \) given functions \( a(x) = x \) and \( a(x) = x^4 \), assuming the continuous variable \( x \) has a PDF \( f(x) \).

Solution: According to Eq. (2.60), one can write \( g(a) = f(x(a)) \left| \frac{da}{dx} \right|^{-1} \). For \( a(x) = x, \) one has \( da/dx = 1 \), which yields \( g(a) = f(a(x)) = f(a) \). For \( a(x) = x^4, \) one has \( da/dx = 4x^3 \) and \( x = \pm a^{1/4} \). Because there are two roots, one must include a factor of two in the calculation of \( g(a) \), one thus gets \( g(a) = f(a^{1/4}/2a^{3/4}) \).

2.9 Verify that the Full Width at Half Maximum (FWHM) of a normal distribution is \( 2.35\sigma \).

Solution: The normal distribution is defined as \( f(x) = (2\pi)^{-1/2} \sigma^{-1} \exp[-(x-\mu)^2/2\sigma^2] \). At the maximum, i.e., at \( x = \mu \), one has \( f(\mu) = (2\pi)^{-1/2} \sigma^{-1} \). To get the full width at half maximum, one must calculate the value \( x_{1/2} \) such that \( f(x_{1/2}) = \frac{1}{2} (2\pi)^{-1/2} \sigma^{-1} \). That implies \( \exp \left[ -(x_{1/2} - \mu)^2/2\sigma^2 \right] = 1/2 \). Solving for \( x_{1/2} \) one finds \( x_{1/2} = \mu \pm \sqrt{2 \ln 1/2} \sigma = \mu \pm 1.177\sigma \). The full width is thus FWHM = \( 2 \times 1.177\sigma = 2.35\sigma \).

2.10 Derive a method to estimate the values \( x_{i,0} \) defined by Eq. (4.64) for PDFs \( f(x) \propto \exp(-x/\lambda) \) and \( f(x) \propto (k + x)^{-\beta} \), where \( k \) and \( \beta \) are two unknown constants. Hint: Use interpolation of the yields in bins \( i-1 \) and \( i+1 \) to estimate the constants \( \lambda \) and \( \beta \) bin by bin.

Solution: Let \( \tilde{f}_i(x_{i,0}) = h_i/\Delta x_i \). Assume the functional dependence
\[
f(x) = \lambda^{-1} \exp(-x/\lambda).
\]
We consider two methods to determine the value \( x_{i,0} \) in each bin \( i \). Both methods require one obtains a local estimate \( \lambda_i \). Once this value is known, estimate \( x_{i,0} \) as
\[
x_{i,0} = -\lambda_i \ln(\lambda_i \tilde{f}_i).
\] (2.1)
Derive an expression similar to Eq. (2.118) for Eq. (2.117):

The second technique involves performing a local fit of \( f \) in which \( h \) and \( \beta \) are jointly true. The expression \( \lambda \) is the joint probability of the variables \( x, y \) and \( m \) are marginal probabilities.

Solution: A simple way to obtain an expression for \( h_{x_1,x_2}(x_1,x_2|x_3,x_4,m) \) is to use Eq. (2.117):

\[
p(X|Y,I) = \frac{p(Y|X,I)p(X|I)}{p(Y|I)},
\]

and define statements \( X = x_1, x_2 \) and \( Y = x_3, x_4, \ldots, x_m \). \( X \) then corresponds to a statement that \( x_1 \) and \( x_2 \) are jointly true while \( Y \) represents the probability that \( x_3, x_4, \ldots, x_m \) are jointly true. The expression \( p(X|Y,I) \) becomes \( p(x_1, x_2|x_3, x_4, \ldots, x_m, I) \) and then represents the conditional probability of the conjunction \( x_1, x_2 \) given \( x_3, x_4, \ldots, x_m \) is true. One can then substitute the expressions for \( X \) and \( Y \) in Eq. (2.118) and obtain

\[
h_{x_1,x_2}(x_1,x_2|x_3,x_4,m) = \frac{h_{x_1,x_2}(x_1,x_2|x_3,x_4,m)\hat{f}_{x_1,x_2}(x_1,x_2)}{\hat{f}_{x_1,x_2}(x_1,x_2)}
\]

in which \( h_{x_1,x_2}(x_3,x_4,m|x_1,x_2) \) is the conditional probability of observing \( x_3, \ldots, x_m \), whereas \( \hat{f}_{x_1,x_2} \) and \( \hat{f}_{x_1,x_2} \) are marginal probabilities.

Show that the Pearson coefficients defined by Eq. (2.172) are bound in the range \(-1 \leq \rho_{ij} \leq 1\) by construction.

Solution: By definition,

\[
V_{ij} = \int (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) p(x_i, x_j) dx_i dx_j,
\]

\[
\sigma^2_i = \int (x_i - \langle x_i \rangle)^2 p(x_i, x_j) dx_i dx_j,
\]

where \( p(x_i, x_j) \) is the joint probability of the variables \( x_i \) and \( x_j \). Defining two variables \( u \) and \( v \) according to

\[
u = (x_j - \langle x_j \rangle) \sqrt{p(x_i, x_j)},
\]

\[
u = (x_j - \langle x_j \rangle) \sqrt{p(x_i, x_j)},
\]
one can use Schwartz inequality
\[ \int u^2 dX \int v^2 dX \geq \left( \int uv dX \right)^2 \] \hspace{1cm} (2.4)
and write
\[ \sigma_i^2 \sigma_j^2 \geq V_{ij}^2, \] \hspace{1cm} (2.5)
which implies that \( |V_{ij}/\sigma_i^2 \sigma_j^2| \leq 1 \).

2.13 Calculate the first, second, third, and fourth moments of the uniform distribution defined as follows:
\[ p_U(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases} \] \hspace{1cm} (2.6)
Solution: The calculation of the moments of \( p_U(x; \alpha, \beta) \) is accomplished by integration according to
\[ \langle x^k \rangle = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} x^k dx = \frac{1}{\beta - \alpha} \left[ \frac{x^{k+1}}{k+1} \right]_{\alpha}^{\beta}, \]
\[ = \frac{1}{(k+1)} \left( \frac{\beta^{k+1} - \alpha^{k+1}}{\beta - \alpha} \right), \]
which simplifies to \( \langle x \rangle = (\beta + \alpha) / 2, \langle x^2 \rangle = \frac{(\beta - \alpha)}{2(\beta - \alpha)} \), etc.

2.14 Calculate the first, second, third, and fourth moments of the triangular distribution defined as follows:
\[ p_T(x; \alpha, \beta) = \begin{cases} \frac{2(x - \alpha)}{(\beta - \alpha)^2} & \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases} \] \hspace{1cm} (2.7)
Solution: The calculation of the moments of \( p_T(x; \alpha, \beta) \) is accomplished by integration according to
\[ \langle x^k \rangle = \frac{2}{(\beta - \alpha)^2} \int_{\alpha}^{\beta} x^k (x - \alpha) dx = \frac{2}{(\beta - \alpha)^2} \left\{ \int_{\alpha}^{\beta} x^{k+1} dx - \alpha \int_{\alpha}^{\beta} x^k dx \right\}, \]
\[ = \frac{2}{(\beta - \alpha)^2} \left\{ \frac{1}{k+2} (\beta^{k+2} - \alpha^{k+2}) - \frac{\alpha}{k+1} (\beta^{k+1} - \alpha^{k+1}) \right\}. \]

2.15 Show that given a partition of a set \( S \) into \( n \) mutually disjoint subsets \( A_i \), with \( i = 1, \ldots, n \), the probability, \( P(B) \), of a set \( B \subset S \) may be written as follows:
\[ P(B) = \sum_i P(B|A_i)P(A_i). \] \hspace{1cm} (2.8)
Hint: Disjoints subsets have a null intersection, \( A_i \cap A_j = 0 \) for \( i \neq j \).

Solution: The set \( S \) can be decomposed into subsets \( A_i \), i.e., \( S = \bigcup_i A_i \), that are all mutually disjoint, \( A_i \cap A_j = 0 \) for \( i \neq j \). Since \( B \) is a subset of \( S \), one can write \( P(B) = P(B \cap S) = P(B|S)P(S) = P(B|S) \) since the probability of \( S \) is unity by hypothesis. Substitute \( \bigcup_i A_i \) for \( S \) to get \( P(B) = P(B| \bigcup_i A_i) = P(B \cap \bigcup_i A_i) \). The intersection of a set with union of sets is equal to the union of the intersections of the
Probability

A neural network is designed and trained to classify particles entering an electro-

Combine Bayes’ theorem with the law of total probability to obtain the following

Assuming all subsets $A_i$ are non-empty, one has $P(A_i) > 0$. Also, since $B \subset S$, one has $B \cap S = B$. Replace $S$ by $\cup_i A_i$ and distribute the intersection. One gets $B = B \cap \cup_i A_i = \cup_i (B \cap A_i)$. Since the subsets $A_i$ are all mutually disjoint, so will the sets $B \cap A_i$, one then concludes that $P(B) = P(\cup_i (B \cap A_i)) = \sum_i P(B \cap A_i)$ or equivalently $P(B) = \sum_i P(B|A_i)P(A_i)$, which is the law of total probability given by Eq. (2.18).

2.16 Combine Bayes’ theorem with the law of total probability to obtain the following expression

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}. \quad (2.9)$$

Solution: Bayes’ theorem may be written $P(A|B)P(B) = P(B|A)P(A)$ or $P(A|B) = P(B|A)P(A)/P(B)$ provided $P(B) \neq 0$. Assuming there exists a complete decomposition of $S = \cup_i A_i$, one can use the law of total probability to write $P(B) = \sum_i P(B|A_i)P(A_i)$ which thus yield the sought for result $P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}$.

2.17 A neural network is designed and trained to classify particles entering an electromagnetic calorimeter as photon (P), electron (E), or hadron (H) based on the longitudinal and transverse patterns of energy deposition they produce in the calorimeter. A Monte Carlo simulation is used to estimate the neural net performance summarized in Table 1. The notations $P_\gamma$, $E_\gamma$, and $H_\gamma$ are used to indicate that the energy deposition pattern is due to a photon, an electron, or a hadron. Data analyzed with the neural network provides for fractions 0.05, 0.15, and 0.8 of photons, electrons, and hadrons, respectively. Calculate the relative rates produced by the nuclear reaction under study.

<table>
<thead>
<tr>
<th>Table 2.1 Particle identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(P\gamma</td>
</tr>
<tr>
<td>$P(E\gamma</td>
</tr>
<tr>
<td>$P(H\gamma</td>
</tr>
</tbody>
</table>

Solution: The probability to identify a particle as a photon is the sum of the probabilities to identify a particle as photon when it is actually a photon times the probability of photons, plus the probability of identify the particle as a photon when it is actually an electron times the probability of electrons, plus the probability of identifying a hadron as a photon times the probability of hadrons. This is written $P(P\gamma) = P(P\gamma|P)P(P) + P(P\gamma|E)P(E) + P(P\gamma|H)P(H)$. Similar expressions can be written for the probabilities of accepting particles as electron and hadron. This can
then be written in a matrix form:

\[
\begin{bmatrix}
P(P_a) \\
P(E_a) \\
P(H_a)
\end{bmatrix}
= \begin{bmatrix}
P(P_a|P) & P(P_a|E) & P(P_a|H) \\
P(E_a|P) & P(E_a|E) & P(E_a|H) \\
P(H_a|P) & P(H_a|E) & P(H_a|H)
\end{bmatrix}^{-1}
\begin{bmatrix}
P(P) \\
P(E) \\
P(H)
\end{bmatrix}
\]

The production probabilities we seek can then be obtained by matrix inversion

\[
\begin{bmatrix}
P(P) \\
P(E) \\
P(H)
\end{bmatrix}
= \begin{bmatrix}
P(P_a|P) & P(P_a|E) & P(P_a|H) \\
P(E_a|P) & P(E_a|E) & P(E_a|H) \\
P(H_a|P) & P(H_a|E) & P(H_a|H)
\end{bmatrix}^{-1}
\begin{bmatrix}
P(P_a) \\
P(E_a) \\
P(H_a)
\end{bmatrix}
\]

Numerical solution with the ROOT macro shown below yields \( P(P) = P(E) = 0.0 \) and \( P(H) = 1.0 \).

// Created by Claude Pruneau on 8/11/17.
// Copyright 2017 Claude Pruneau. All rights reserved.
#include <stdio.h>

void problem217() {
    TMatrixD a(3,3);
    TVectorD b(3);
    a(0,0) = 0.98;
    a(0,1) = 0.06;
    a(0,2) = 0.05;
    a(1,0) = 0.01;
    a(1,1) = 0.90;
    a(1,2) = 0.15;
    a(2,0) = 0.01;
    a(2,1) = 0.04;
    a(2,2) = 0.80;
    b(0) = 0.05;
    b(1) = 0.15;
    b(2) = 0.80;
    TVectorD x(3);
    bool ok;
    TDecompLU lu(a);
    x = lu.Solve(b,ok);
    for (int i=0; i<3; i++) {
        for (int j=0; j<3; j++) {
            cout << i << " " << j << " " << a(i,j) << endl; }
        cout << " x =" << x(i) << endl;
    }
}
2.18 Derive the expression (2.182) giving the moments of a PDF in terms of derivatives of its characteristic function.

Solution: The characteristic function \( \phi_x(t) \) of a random variable \( x \) with PDF \( f(x) \) is defined according to

\[
\phi_x(t) = E\left[ e^{itx} \right] = \int_{-\infty}^{\infty} e^{itx} f(x) dx.
\]

Expanding the exponential in Taylor series, one gets

\[
e^{itx} = 1 + itx + \frac{(itx)^2}{2!} + \frac{(itx)^3}{3!} + \cdots + \frac{(itx)^n}{n!} + \cdots
\]

The expectation value \( E[e^{itx}] \) may then be written

\[
E[e^{itx}] = 1 + itE[x] + \frac{(it)^2 E[x^2]}{2!} + \frac{(it)^3 E[x^3]}{3!} + \cdots + \frac{(it)^n E[x^n]}{n!} + \cdots
\]

Moments of order \( n \) are then determined by taking derivatives of \( \phi_x(t) \) evaluated at \( t = 0 \), one gets

\[
\frac{d\phi_x(t)}{dt} \bigg|_{t=0} = \left( i\frac{d\phi_x(t)}{dt} \right) \bigg|_{t=0} = \mu_1',
\]

Moments of order \( n \) are obtained similarly by taking derivatives of order \( n \) w.r.t. \( it \).

2.19 Use Eq. (2.182) to calculate the moments of (a) the uniform distributions, (b) the exponential distribution, (c) the t-distributions, and (d) the \( \chi^2 \)-distribution.

Solution: Calculation of the moments with Eq. (2.182) proceeds in the same way for all PDFs. We here demonstrate for the uniform distribution defined by Eq. (2.69). One must first evaluate the characteristic function \( \phi_x(t) \), i.e., the expectation value \( E[e^{itx}] \):

\[
\phi_x(t) \equiv E[e^{itx}] = \frac{1}{b-a} \int_{a}^{b} e^{itx} dx
\]

Moments of order \( n \) are then determined by taking derivatives of \( \phi_x(t) \) evaluated at \( t = 0 \). Given the \( 1/t \) factor in the above expression, it is clear that derivatives will contain powers of \( 1/t \). So, rather than taking derivatives of the above expression, it
is more convenient to first write the exponentials in terms of Taylor series:

\[ \phi_x(t) = (b - a)^{-1} (it)^{-1} \left[ \sum_{n=1}^{\infty} \frac{(it)^n}{n!} \right] - \left[ \sum_{n=1}^{\infty} \frac{(it)^n}{n!} \right], \]

\[ = (b - a)^{-1} \sum_{n=1}^{\infty} \frac{(it)^n}{n!} (b^n - a^n). \]

Using the above expression for \( \phi_x(t) \), the first moment is

\[ \mu_1 = i^{-1} \left. \frac{d \phi_x(t)}{dt} \right|_{t=0} = i^{-1} \left. \frac{d \phi_x(t)}{d(it)} \right|_{t=0}, \]

\[ = (b - a)^{-1} \sum_{n=1}^{\infty} \frac{(n-1)(it)^{n-2}(b^n - a^n)}{n!} \bigg|_{t=0}, \]

\[ = \frac{1}{2} \frac{b^2 - a^2}{b - a} = \frac{(a + b)}{2}. \]

Higher order moments are calculated in a similar fashion.

**2.20** Show that the difference between two independent Gaussian deviates \( \Delta x = x_1 - x_2 \) has a mean \( \mu_{\Delta x} = \mu_1 - \mu_2 \) and a variance \( \sigma_{\Delta x}^2 = \sigma_1^2 + \sigma_2^2 \).

Solution: Broadly speaking, two approaches may be used to solve this problem. The first approach involves calculating the convolution of two Gaussian PDFs in \( x_1 \) and \( x_2 \) to determine the PDF of \( z = x_1 - x_2 \). The second approach relies on the one-to-one mapping between PDFs and their characteristic function and the notion that the characteristic function of a sum of independent variables is equal to the product of the corresponding characteristic functions. The characteristic function of \( z \) is written

\[ \phi_z(t) \equiv \mathbb{E} \left[ e^{itz} \right] = \int \int \exp \left[ i(t x_1 - x_2) \right] p_G(x_1 | \mu_1, \sigma_1) p_G(x_2 | \mu_2, \sigma_2) dx_1 dx_2, \]

where \( p_G(x|\mu,\sigma) \) represent Gaussian PDFs for variable \( x \) with mean \( \mu \) and standard deviation \( \sigma \). Factorizing the exponential, one gets

\[ \phi_z(t) \equiv \mathbb{E} \left[ e^{itx} \right] = \int \exp \left[ i(t x_1) \right] p_G(x_1 | \mu_1, \sigma_1) dx_1 \int \exp \left[ i(t x_2) - i(t x_2) \right] p_G(x_2 | \mu_2, \sigma_2) dx_2, \]

\[ = \phi_{x_1}(t) \phi_{x_2}(-t), \]

where \( \phi_{x_i}(t) \) and \( \phi_{x_i}(-t) \) are, by definition, the characteristic functions of the PDFs \( p_G(x_1 | \mu_1, \sigma_1) \) and \( p_G(x_2 | \mu_2, \sigma_2) \), respectively. The characteristic function of a Gaussian variable was calculated in Eq. (2.184). One gets:

\[ \phi_z(t) = \exp \left[ i(\mu_1 t - \frac{1}{2} \sigma_1^2 t^2) \right] \exp \left[ -i(\mu_2 t - \frac{1}{2} \sigma_2^2 t^2) \right], \]

where, in the second factor of the first line, we substituted \( t \) by \(-t\). Clearly, the above expression is itself the characteristic function of a Gaussian variable (the variable \( z \) in this case). One thus concludes that the PDF of \( z \) is a Gaussian PDF with mean \( \mu_z = \mu_1 - \mu_2 \) and variance \( \sigma_z^2 = \sigma_1^2 + \sigma_2^2 \).
2.21 Extend Eq. (2.229) and find an expression for the error of a sum of several independent Gaussian deviates \( x_1, x_2, \ldots, x_n \).

Solution: Given the variables \( x_1, x_2, \ldots, x_n \) are statistically independent, one proceeds to calculate the characteristic function of \( y = x_1 + x_2 + \ldots + x_n \) as a product of the characteristic functions of the variables \( x_1, x_2, \ldots, x_n \):

\[
\phi_y(t) = \phi_{x_1}(t)\phi_{x_2}(t)\cdots\phi_{x_n}(t).
\]

The characteristic function of a Gaussian variable \( x_i \) is \( \phi_{x_i}(t) = \exp\left[i(\mu_i t - \frac{1}{2}\sigma_i^2 t^2)\right] \) in which \( \mu_i \) and \( \sigma_i \) are the mean and standard deviation of the Gaussian PDF. The characteristic function of \( y \) is thus

\[
\phi_y(t) = \exp\left[i(\mu_1 t + \mu_2 t + \cdots + \mu_n) + \frac{1}{2} \left(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2\right) t^2\right]
\]

which is itself the characteristic of a Gaussian variable. The variable \( y \) thus has a Gaussian PDF with mean \( \mu_y = \mu_1 + \mu_2 + \cdots + \mu_n \) and variance \( \sigma_y^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2 \). The error on a sum of \( n \) Gaussian variables is thus \( \sigma_y = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2} \).

2.22 Extend Eq. (2.229) and find an expression for the error of a sum of several independent Poisson deviates \( x_1, x_2, \ldots, x_n \).

Solution: Given the variables \( x_1, x_2, \ldots, x_n \) are statistically independent, one proceeds to calculate the characteristic function of \( y = x_1 + x_2 + \cdots + x_n \) as a product of the characteristic functions of the variables \( x_1, x_2, \ldots, x_n \):

\[
\phi_y(t) = \phi_{x_1}(t)\phi_{x_2}(t)\cdots\phi_{x_n}(t).
\]

The characteristic function of a Poisson variable \( x_i \), given in Table 3.3, amounts to

\[
\phi_{x_i}(t) = e^{\lambda(e^{it} - 1)}.
\]

The characteristic function of \( y \) is thus

\[
\phi_y(t) = \phi_{x_1}(t)\phi_{x_2}(t)\cdots\phi_{x_n}(t),
\]

\[
= e^{\lambda_1(e^{it} - 1)}e^{\lambda_2(e^{it} - 1)}\cdots e^{\lambda_n(e^{it} - 1)},
\]

\[
= e^{(\lambda_1 + \lambda_2 + \cdots + \lambda_n)(e^{it} - 1)},
\]

\[
= e^{\lambda y(e^{it} - 1)},
\]

where we introduce \( \lambda_y = \lambda_1 + \lambda_2 + \cdots + \lambda_n \). Since the above expression is itself the characteristic function of a Poisson variable, we conclude that \( y \) has a Poisson distribution with slope parameter \( \lambda_y \). The 2nd central moment (variance) of a Poisson distribution is equal to its slope parameter \( \lambda \). One thus find that the error on a sum of \( n \) Poisson variable is \( \lambda_{y,2} = \sqrt{\lambda_1 + \lambda_2 + \cdots + \lambda_n} \).

2.23 Show that the characteristic function of \( f(w) = \frac{1}{\sqrt{2\pi w}}e^{-w/2} \) is given by \( \phi_w(t) = (1 - 2it)^{-1/2} \).

Solution: First, one can verify that the normalization of the function is correct by integrating the function from 0 to \( \infty \). To carry out this integral, use \( z = w^{1/2} \),
Exercises

2.25 Derive the expression (2.223) for the covariance of functions \( y_1(\vec{x}) \) and \( y_j(\vec{x}) \).

Solution: Let us write the functions \( y_i(\vec{x}) \) as Taylor series about the mean \( \vec{\mu} = \{ \mu_i \} \).

\[
y_i(\vec{x}) = y_i(\vec{\mu}) + \sum_{k=1}^{n} \left[ \frac{\partial y_i}{\partial x_k} \right]_{\vec{x} = \vec{\mu}} (x_k - \mu_k) + O(2).
\]

The expectation value of the product \( y_i(\vec{x}) y_j(\vec{x}) \) is thus (keeping only leading terms)

\[
\mathbb{E}[y_i(\vec{x}) y_j(\vec{x})] = \mathbb{E}\left[ y_i(\vec{\mu}) + \sum_{k=1}^{n} \left[ \frac{\partial y_i}{\partial x_k} \right]_{\vec{x} = \vec{\mu}} (x_k - \mu_k) \right] \mathbb{E}\left[ y_j(\vec{\mu}) + \sum_{k'=1}^{n} \left[ \frac{\partial y_j}{\partial x_{k'}} \right]_{\vec{x} = \vec{\mu}} (x_{k'} - \mu_{k'}) \right] \\
= y_i(\vec{\mu}) y_j(\vec{\mu}) + y_i(\vec{\mu}) \sum_{k=1}^{n} \left[ \frac{\partial y_i}{\partial x_k} \right]_{\vec{x} = \vec{\mu}} \mathbb{E}[x_k - \mu_k] \\
+ y_j(\vec{\mu}) \sum_{k'=1}^{n} \left[ \frac{\partial y_j}{\partial x_{k'}} \right]_{\vec{x} = \vec{\mu}} \mathbb{E}[x_{k'} - \mu_{k'}] \\
+ \sum_{k=1}^{n} \left[ \frac{\partial y_i}{\partial x_k} \right]_{\vec{x} = \vec{\mu}} \sum_{k'=1}^{n} \left[ \frac{\partial y_j}{\partial x_{k'}} \right]_{\vec{x} = \vec{\mu}} \mathbb{E}[(x_k - \mu_k)(x_{k'} - \mu_{k'})].
\]

The expectation values \( \mathbb{E}[x_k - \mu_k] \) are null since \( \mathbb{E}[x_k] = \mu_k \) by definition. Only the first and last terms of the above expression thus remain. Calculating \( \text{Cov}[y_i(\vec{x}), y_j(\vec{x})] = \mathbb{E}[y_i(\vec{x}) y_j(\vec{x})] - \mathbb{E}[y_i(\vec{x})] \mathbb{E}[y_j(\vec{x})] \), one thus gets

\[
\text{Cov}[y_i(\vec{x}), y_j(\vec{x})] \approx \sum_{k=1}^{n} \left[ \frac{\partial y_i}{\partial x_k} \right]_{\vec{x} = \vec{\mu}} \sum_{k'=1}^{n} \left[ \frac{\partial y_j}{\partial x_{k'}} \right]_{\vec{x} = \vec{\mu}} V_{ij}, \tag{2.10}
\]

where we introduced the covariance matrix \( V_{ij} \equiv \mathbb{E}[(x_k - \mu_k)(x_{k'} - \mu_{k'})] \). The derivatives are simple coefficients, the above can be re-written

\[
\text{Cov}[y_i(\vec{x}), y_j(\vec{x})] \approx \sum_{k,k'=1}^{n} \left[ \frac{\partial y_i}{\partial x_k} \right]_{\vec{x} = \vec{\mu}} \left[ \frac{\partial y_j}{\partial x_{k'}} \right]_{\vec{x} = \vec{\mu}} V_{ij},
\]

2.25 Demonstrate the expression \( \sigma_y^2 \approx \sigma_1^2 + \sigma_2^2 + 2V_{12} \), given by Eq. (2.229), corresponding to the error on the sum of random variables \( x_1 \) and \( x_2 \).

Solution: Let \( y = x_1 + x_2 \). One can use Eq. (2.222) to estimate the error \( \sigma_y \) on \( y \).

Evidently, \( \partial y/\partial x_k = 1 \), Eq. 2.222 thus becomes

\[
\sigma_y^2 \approx \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_1} V_{11} + \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{12} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_1} V_{21} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_2} V_{22} \\
\approx V_{11} + V_{22} + 2V_{12},
\]
where we use $V_{12} = V_{21}$. Substituting $\sigma^2_{1} = V_{11}$ and $\sigma^2_{2} = V_{22}$, one finally gets

$$\sigma^2_{\bar{y}} = \sigma^2_{1} + \sigma^2_{2} + 2V_{12}.$$  

**2.26** Demonstrate the expression $\sigma^2_{\bar{y}} \approx \sigma^2_{1} + \sigma^2_{2} - 2V_{12}$ corresponding to the error on the difference of random variables $x_1$ and $x_2$.

Solution: Let $y = x_1 - x_2$. One can use Eq. (2.222) to estimate the error $\sigma_{\bar{y}}$ on $y$. Evidently, $\partial y / \partial x_1 = 1$ and $\partial y / \partial x_2 = -1$, Eq. (2.222) thus becomes

$$\sigma^2_{\bar{y}} \approx \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{11} + \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{12} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_1} V_{21} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_2} V_{22}$$

$$\approx V_{11} - V_{22} - 2V_{12},$$

where we use $V_{12} = V_{21}$. Substituting $\sigma^2_{1} = V_{11}$ and $\sigma^2_{2} = V_{22}$, one finally gets

$$\sigma^2_{\bar{y}} = \sigma^2_{1} + \sigma^2_{2} - 2V_{12}.$$  

**2.27** Demonstrate Eq. (2.230) corresponding to the error on the product of correlated random variables $x_1$ and $x_2$.

Solution: Let $y = x_1 x_2$. One can use Eq. (2.222) to estimate the error $\sigma_{\bar{y}}$ on $y$. Evidently, $\partial y / \partial x_1 = x_2$ and $\partial y / \partial x_2 = x_1$, Eq. 2.222 thus becomes

$$\sigma^2_{\bar{y}} \approx \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{11} + \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{12} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_1} V_{21} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_2} V_{22}$$

$$\approx x_1^2 V_{11} + x_2^2 V_{22} + 2y_1 y_2 V_{12},$$

where we use $V_{12} = V_{21}$. Substituting $\sigma^2_{1} = V_{11}$ and $\sigma^2_{2} = V_{22}$, and dividing by the square of $y$, one gets the sought for result:

$$\frac{\sigma^2_{\bar{y}}}{y^2} = \frac{\sigma^2_{1}}{x_1^2} + \frac{\sigma^2_{2}}{x_2^2} + 2 \frac{V_{12}}{x_1 x_2}.$$  

**2.28** Calculate the error on a ratio $y = x_1 / x_2$ obtained by dividing correlated random variables $x_1$ and $x_2$.

Solution: Let $y = x_1 / x_2$. One can use Eq. (2.222) to estimate the error $\sigma_{\bar{y}}$ on $y$. Evidently, $\partial y / \partial x_1 = 1 / x_2$ and $\partial y / \partial x_2 = -x_1 / x_2^2$, Eq. 2.222 thus becomes

$$\sigma^2_{\bar{y}} \approx \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{11} + \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} V_{12} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_1} V_{21} + \frac{\partial y}{\partial x_2} \frac{\partial y}{\partial x_2} V_{22}$$

$$\approx \frac{1}{x_2^2} V_{11} + \frac{x_1^2}{x_2^4} V_{22} - 2 \frac{1}{x_2^4} x_1 V_{12},$$

where we use $V_{12} = V_{21}$. Substituting $\sigma^2_{1} = V_{11}$ and $\sigma^2_{2} = V_{22}$, and dividing by the square of $y = x_1 / x_2$, one gets the sought for result:

$$\frac{\sigma^2_{\bar{y}}}{y^2} = \frac{\sigma^2_{1}}{x_1^2} + \frac{\sigma^2_{2}}{x_2^2} - 2 \frac{V_{12}}{x_1 x_2}.$$  

**2.29** Show that $\lambda_k = \frac{1}{2} \left[ \sigma^2_{1} + \sigma^2_{2} \pm \sqrt{\sigma^2_{1} + \sigma^2_{2} - 4 (1 - \rho^2) \sigma^2_{1} \sigma^2_{2}} \right]$ indeed satisfies Eq. (5.66), and find the eigenvectors $r_+$ and $r_-$. 

Solution: See the solution of problem 2.33.
2.30 Show that if the intersection of two subsets $A$ and $B$ is null (i.e., for $A \cap B = 0$), the subsets cannot be independent, and determine the value of $P(A \cap B)$.

Solution: The condition $A \cap B = 0$ implies that if an event is within $A$, it cannot be within $B$, and conversely, if an event is in $B$, it cannot be in $A$. One can then write that the probability of having an event in $A$ given that it is known to be in $B$ is null. This is written $P(A|B) = 0$. This, in turn, implies that $P(A \cap B) = 0$ which obviously does not satisfy the condition $P(A \cap B) = P(A)P(B)$ unless one of the two sets is empty: the two sets cannot be statistically independent if knowledge of an event being in one set implies the event cannot be in the other set.

2.31 Given PDFs $g(x)$ and $h(y)$ for random variables $x$ and $y$ respectively, calculate the PDF of variable $z$ defined as

- $z^2 = x^2 + y^2$.
- $\tan^{-1}(x/y)$.

Solution: Let $z^2 = x^2 + y^2$. The PDF of $z$, which we denote $f_z(z)$, is calculated by summing over all combinations of $x$ and $y$ that satisfy $z = \sqrt{x^2 + y^2}$:

$$f_z(z)dz = \int_{-\infty}^{\infty} f_x(x)dx \int_{-\infty}^{\infty} f_y(y)\delta\left(z - \sqrt{x^2 + y^2}\right)dy.$$ 

Consider the delta function as a function $\delta(g(y))$, one can then write

$$\delta(g(y)) = \sum_i \frac{\delta(y - y_i)}{|dg(y)/dy|_{y_i}},$$

where the sum proceeds on the roots $y_i$ of $g(y)$. If the variable $x$, $y$, and $z$ are all three defined over the real axis (i.e., $]-\infty, \infty[$), then there are two roots $y_i = \pm \sqrt{z^2 - x^2}$. The derivative of $g(y)$ w.r.t. $y$ evaluated at these roots is

$$\left.\frac{dg(y)}{dy}\right|_{y_i} = \frac{y}{\sqrt{x^2 + y^2}} = \pm \frac{\sqrt{z^2 - x^2}}{z}.$$ 

Substituting in the integral for $f_z(z)dz$, one finds

$$f_z(z)dz = z \int_{-\infty}^{\infty} f_x(x)f_y\left(\frac{\sqrt{z^2 - x^2}}{\sqrt{z^2 - x^2}}\right)dx + z \int_{-\infty}^{\infty} f_x(x)f_y\left(-\frac{\sqrt{z^2 - x^2}}{\sqrt{z^2 - x^2}}\right)dx.$$ 

(2.11)

Evaluation of the integral should be carried at fixed $z$ and evidently requires knowledge of the functions $f_x$ and $f_y$. A similar reasoning is required for the PDF of $\tan^{-1}(x/y)$.

2.32 Verify by direct substitution of the eigenvectors $\vec{r}$ into $\mathbf{R}$ defined by Eq. (2.231) that $\mathbf{R}$ satisfies $\mathbf{R}^{-1} = \mathbf{R}^T$ (i.e., its inverse is equal to its transpose), and that it is consequently an orthogonal transformation of the vector $\vec{x}$ which leaves its norm invariant.

Solution: The eigenvectors $\vec{r}^{(i)}$ are defined according to

$$\mathbf{V}\vec{r}^{(i)} = \alpha_i\vec{r}^{(i)}.$$
The matrix $\mathbf{R}$ then consists of the vectors $\vec{r}^{(i)}$ which substituted in $\mathbf{R}\mathbf{R}^T$ yields

$$
\begin{bmatrix}
\vec{r}^{(1)T} \\
\vec{r}^{(2)T} \\
\vdots \\
\vec{r}^{(n)T}
\end{bmatrix}
= \begin{bmatrix}
\vec{r}^{(1)} & \vec{r}^{(2)} & \ldots & \vec{r}^{(n)} \\
\vec{r}^{(2)T} & \vec{r}^{(1)T} & \vec{r}^{(1)T} & \ldots & \vec{r}^{(1)T} & \vec{r}^{(n)} \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
\vec{r}^{(n)T} & \vec{r}^{(1)T} & \vec{r}^{(2)T} & \ldots & \vec{r}^{(2)T} & \vec{r}^{(n)T} \\
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix} = \delta_{ij}
$$

One thus indeed verify that the inverse of the matrix $\mathbf{R}$ is equal to its transpose.

**2.33** Consider a two-dimensional covariance matrix $\mathbf{V}$ defined as follows:

$$
\mathbf{V} = \begin{pmatrix}
\sigma_1^2 & \rho\sigma_1\sigma_2 \\
\rho\sigma_1\sigma_2 & \sigma_2^2
\end{pmatrix}. \quad (2.10)
$$

Solve the eigenvalue equation $(\mathbf{V} - \lambda \mathbf{I})\vec{r} = 0$ and show the eigenvalues may be written

$$
\lambda_{\pm} = \frac{1}{2} \left[ \sigma_1^2 + \sigma_2^2 \pm \sqrt{(\sigma_1^2 + \sigma_2^2 - 4(1-\rho^2)\sigma_1^2\sigma_2^2)} \right] \quad (2.11)
$$

with eigenvectors given by

$$
\begin{pmatrix}
\cos \theta \\
\sin \theta
\end{pmatrix} = \begin{pmatrix}
\cos \theta \\
\sin \theta
\end{pmatrix}
$$

such that

$$
\theta = \frac{1}{2} \tan^{-1} \left( \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \right)
$$

and

$$
\mathbf{A} = \begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}
$$

**Solution:** The eigenvalues are obtained by solving $(\mathbf{V} - \lambda \mathbf{I})\vec{r} = 0$. This requires

$$
\begin{vmatrix}
\sigma_1^2 - \lambda & \rho\sigma_1\sigma_2 \\
\rho\sigma_1\sigma_2 & \sigma_2^2 - \lambda
\end{vmatrix} = 0
$$

which can be written

$$
0 = (\sigma_1^2 - \lambda)(\sigma_2^2 - \lambda) - \rho^2\sigma_1^2\sigma_2^2
$$

$$
= \sigma_1^2\sigma_2^2 - \lambda (\sigma_1^2 + \sigma_2^2) + \lambda^2 - \rho^2\sigma_1^2\sigma_2^2
$$

$$
= \lambda^2 - \lambda (\sigma_1^2 + \sigma_2^2) + \sigma_1^2\sigma_2^2(1 - \rho^2)
$$

Consider a system with two random (or fluctuating) observables. Next calculate the second moment of independent, i.e., $E[\langle x \rangle^2]$ yielding the variance of $x$ as $\sigma_x^2$. The variance of $x$ and $y$ might yield (a) independent variables, (b) correlated variables with a positive covariance, (c) correlated variables with a negative covariance, and (d) correlated variables with a Pearson coefficient equal to unity. Assume the random variables $r_1$ and $r_2$ have null expectation values, $E[r_1] = E[r_2] = 0$, and nonvanishing variances $\sigma_1^2$ and $\sigma_2^2$, respectively.

Solution: Let $\langle r_1 \rangle \equiv E[r_1] = 0$ and $\langle r_2 \rangle \equiv E[r_2] = 0$. Then $\langle (r_1 - \langle r_1 \rangle)^2 \rangle \equiv E[(r_1 - \langle r_1 \rangle)^2] = E[r_1^2] = \sigma_1^2$ given the problem statement. Similarly for $r_2$, one has $\langle r_2^2 \rangle = \sigma_2^2$. Additionally, consider that the variables $r_1$ and $r_2$ are statistically independent, i.e., $E[r_1 r_2] = E[r_1] E[r_2] = 0$. First calculate the means of $x$ and $y$:

$$\langle x \rangle = a + b_1 r_1 + b_2 r_2$$

$$\langle y \rangle = c + d_1 r_1 - d_2 r_2$$

where $a$, $b_1$, $b_2$, $c$, $d_1$, and $d_2$ are arbitrary constants. Calculate the mean and variance of the observable $x$ and $y$ as well as the covariance of $x$ and $y$. Discuss conditions under which $x$ and $y$ might yield (a) independent variables, (b) correlated variables with a positive covariance, (c) correlated variables with a negative covariance, and (d) correlated variables with a Pearson coefficient equal to unity. Assume the random variables $r_1$ and $r_2$ have null expectation values, $E[r_1] = E[r_2] = 0$, and nonvanishing variances $\sigma_1^2$ and $\sigma_2^2$, respectively.

Next calculate the second moment of $x$:

$$\langle x^2 \rangle = E[(a + b_1 r_1 + b_2 r_2)^2]$$

$$= E[a^2 + b_1^2 r_1^2 + b_2^2 r_2^2 + 2ab_1 r_1 + 2ab_2 r_2 + 2b_1 b_2 r_1 r_2]$$

$$= a^2 + b_1^2 E[r_1^2] + b_2^2 E[r_2^2] + 2ab_1 E[r_1] + 2ab_2 E[r_2] + 2b_1 b_2 E[r_1 r_2]$$

$$= a^2 + b_1^2 \sigma_1^2 + b_2^2 \sigma_2^2.$$

The variance of $x$ is thus $\text{Var}[x] = b_1^2 \sigma_1^2 + b_2^2 \sigma_2^2$. Calculation of $\langle y^2 \rangle$ proceeds similarly and yields

$$\langle y^2 \rangle = c^2 + d_1^2 \sigma_1^2 + d_2^2 \sigma_2^2.$$
and the variance of $y$ is $\text{Var}[y] = d_1^2\sigma_1^2 + d_2^2\sigma_2^2$. Next, calculate the covariance of $x$ and $y$:

$$\text{Cov}[x, y] = E[xy] - \langle x \rangle \langle y \rangle.$$ 

Consider the first term

$$E[xy] = E[(a + b_1 r_1 + b_2 r_2)(c + d_1 r_1 - d_2 r_2)].$$

$$= E[ac + ad_1 r_1 - ad_2 r_2 + b_1 cr_1 + b_1 d_1 r_1^2 - b_1 d_2 r_1 r_2 + b_2 cr_2 + b_2 d_1 r_1 r_2 - b_2 d_2 r_2^2],$$

$$= ac + b_1 d_1 E[r_1^2] - b_2 d_2 E[r_2^2],$$

$$= ac + b_1 d_1 \sigma_1^2 - b_2 d_2 \sigma_2^2.$$ 

The covariance is thus

$$\text{Cov}[x, y] = b_1 d_1 \sigma_1^2 - b_2 d_2 \sigma_2^2.$$ 

We can now answer the questions. (a) A necessary condition for $x$ and $y$ to be considered independent variables is for the covariance to be null. This shall be the case if $b_1 d_1 \sigma_1^2 = b_2 d_2 \sigma_2^2$. (b) the two variables shall have a positive correlation when $b_1 d_1 \sigma_1^2 > b_2 d_2 \sigma_2^2$, and (c) a negative covariance when $b_1 d_1 \sigma_1^2 < b_2 d_2 \sigma_2^2$. The Pearson coefficient $\rho$ is

$$\rho = \frac{\text{Cov}[x, y]}{\sqrt{\text{Var}[x] \text{Var}[y]}},$$

$$= \frac{b_1 d_1 \sigma_1^2 - b_2 d_2 \sigma_2^2}{\sqrt{b_1^2 \sigma_1^4 + b_2^2 \sigma_2^4} \sqrt{d_1^2 \sigma_1^4 + d_2^2 \sigma_2^4}},$$

$$= \frac{b_1 d_1 \sigma_1^2 - b_2 d_2 \sigma_2^2}{\sqrt{b_1^2 d_1^2 \sigma_1^4 + (b_2^2 d_2^2 + b_2 d_1^2) \sigma_1^2 \sigma_2^2 + b_2^2 d_2^2 \sigma_2^4}}.$$ 

It is then easy to verify that $\rho = 1$ (perfect correlation) if $b_1 d_2 + b_2 d_1 = 0$.

Verify that the first order and second order moments associated with a two-dimensional random-walk are given by Eqs. (2.249) and (2.251).

Solution: First, let’s consider a non directed random walk. Let $x_i = \cos \phi_i$ and $y_i = \sin \phi_i$ be the projections of the unit step size in the the $x-y$ plane along the $x$ and $y$ axes, respectively. All directions $\phi_i$ being equally probable, we write $P(\phi_i) = (2\pi)^{-1}$ in the range $[0, 2\pi]$. Calculate basic moments:

$$\langle x_i \rangle = \langle \cos \phi_i \rangle = (2\pi)^{-1} \int_0^{2\pi} \cos \phi_i d\phi_i = 0,$$

$$\langle y_i \rangle = \langle \sin \phi_i \rangle = (2\pi)^{-1} \int_0^{2\pi} \sin \phi_i d\phi_i = 0,$$

$$\langle x_i^2 \rangle = \langle \cos^2 \phi_i \rangle = (2\pi)^{-1} \int_0^{2\pi} \cos^2 \phi_i d\phi_i = 1/2.$$ 

Define $S_{x,n} = \sum_{j=1}^{n} x_j$ and $S_{y,n} = \sum_{j=1}^{n} y_j$ and a displacement vector $\vec{S}_n = (S_{x,n}, S_{y,n})$. 
The moments of $S_{x,n}$ are

$$\langle S_{x,n} \rangle = \sum_{j=1}^{n} (x_j) = 0,$$

$$\langle S_{x,n}^2 \rangle = \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{j=1}^{n} x_j \right),$$

$$= \sum_{i=1}^{n} (x_i^2) + \sum_{i \neq j=1}^{n} (x_i x_j),$$

$$= \sum_{i=1}^{n} 1/2 = n/2,$$

where in the last line, we use the fact that each new step is independent of the previous steps and consequently $\langle x_i x_j \rangle = 0$ for $i \neq j$. Calculation of the moments of $S_{y,n}$ proceeds in the same way. The variances are then

$$\text{Var}[S_{x,n}] = \langle S_{x,n}^2 \rangle - \langle S_{x,n} \rangle^2 = n/2,$$

$$\text{Var}[S_{y,n}] = \langle S_{y,n}^2 \rangle - \langle S_{y,n} \rangle^2 = n/2.$$

Next, let’s repeat the above treatment for a directed random walk. Let $x_i = r \cos \phi_i$ and $y_i = r \sin \phi_i$ be the projections of the fix step size $r$ in the the $x - y$ plane along the $x$ and $y$ axes, respectively. The motion is now directed with probability $P(\phi_i) = (2\pi)^{-1} \{ 1 + 2v_m \cos[m(\phi_i - \psi)] \}$ in the range $[0, 2\pi]$. Calculate basic moments:

$$\langle x_i \rangle = \langle \cos \phi_i \rangle = (2\pi)^{-1} \int_0^{2\pi} \cos \phi_i \{ 1 + 2v_m \cos[m(\phi_i - \psi)] \} \, d\phi_i = v_m r \cos(m\psi),$$

$$\langle y_i \rangle = \langle \sin \phi_i \rangle = (2\pi)^{-1} \int_0^{2\pi} \sin \phi_i \{ 1 + 2v_m \cos[m(\phi_i - \psi)] \} \, d\phi_i = v_m r \sin(m\psi),$$

$$\langle x_i^2 \rangle = \langle \cos^2 \phi_i \rangle = \frac{r^2}{4\pi},$$

$$\langle y_i^2 \rangle = \langle \sin^2 \phi_i \rangle = \frac{r^2}{4\pi}.$$

The variances are

$$\text{Var}[x_i] = \langle x_i^2 \rangle - \langle x_i \rangle^2 = \frac{r^2}{4\pi} - r^2 v_m^2 \cos^2(m\psi) \approx \frac{r^2}{4\pi},$$

$$\text{Var}[y_i] = \langle y_i^2 \rangle - \langle y_i \rangle^2 = \frac{r^2}{4\pi} - r^2 v_m^2 \sin^2(m\psi) \approx \frac{r^2}{4\pi},$$

where the approximations hold for $v_m^2 \ll \frac{1}{4\pi}$. Define once again $S_{x,n} = \sum_{j=1}^{n} x_j$ and $S_{y,n} = \sum_{j=1}^{n} y_j$ and a displacement vector $\vec{S}_n = (S_{x,n}, S_{y,n})$. The first moments of $S_{x,n}$
and \( S_{yn} \) are

\[
\langle S_{xn} \rangle = \sum_{j=1}^{n} \langle x_j \rangle = nrv_m \cos(m\psi),
\]

\[
\langle S_{yn} \rangle = \sum_{j=1}^{n} \langle y_j \rangle = nrv_m \sin(m\psi).
\]

The second moments are

\[
\langle S_{2xn} \rangle = \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{j=1}^{n} x_j \right),
\]

\[
= \sum_{i=1}^{n} \langle x_i^2 \rangle + \sum_{i \neq j=1}^{n} \langle x_i x_j \rangle,
\]

\[
= \sum_{i=1}^{n} 1/2 = n^2 \frac{r^2}{4\pi},
\]

\[
\langle S_{2yn} \rangle = n^2 \frac{r^2}{4\pi},
\]

where we use the fact that each new step is independent of the previous steps and consequently \( \langle x_i x_j \rangle = 0 \) for \( i \neq j \).

2.36 Derive Eq. (2.260) from Eq. (2.258) for the expression of the probability density \( \frac{dN}{dS_{xn} dS_{yn}} \) of the sum vector \( \vec{S}_n \).

Solution: One defines the cumulative displacement vector \( \vec{S}_n = (S_{xn}, S_{yn}) \). The mean of this vector is \( \langle \vec{S}_n \rangle = (\langle S_{xn} \rangle, \langle S_{yn} \rangle) \) and the square of \( \vec{S}_n - \langle \vec{S}_n \rangle \) is

\[
(S_n - \langle S_n \rangle)^2 = (S_{xn} - \langle S_{xn} \rangle)^2 + (S_{yn} - \langle S_{yn} \rangle)^2
\]

In the large \( n \) limit, the components along \( x \) and \( y \) are independent and tend to Gaussian deviates. One can then write

\[
\frac{1}{N} \frac{dN}{dS_{xn} dS_{yn}} = \frac{1}{2\pi} \frac{1}{\sigma_{xn} \sigma_{yn}} \exp \left[ -\frac{(S_{xn} - \langle S_{xn} \rangle)^2}{2\sigma_{xn}^2} \right] \exp \left[ -\frac{(S_{yn} - \langle S_{yn} \rangle)^2}{2\sigma_{yn}^2} \right]
\]

In the large \( n \) limit, one also has \( \sigma_{xn} \approx \sigma_{yn} \equiv \sigma_n \). The above becomes

\[
\frac{1}{N} \frac{dN}{dS_{xn} dS_{yn}} = \frac{1}{2\pi} \frac{1}{\sigma_n^2} \exp \left[ -\frac{(S_n - \langle S_n \rangle)^2}{2\sigma_n^2} \right].
\]

Finally, rather than measure \( S_n^2 \) in terms of cartesian coordinates, one can switch to polar coordinates

\[
\frac{1}{N} \frac{d^2N}{dS_{n} d\theta} = \frac{1}{N} \frac{dN}{dS_{xn} dS_{yn}} \left| \frac{\partial(S_{xn}, S_{yn})}{\partial(S_n, \theta)} \right|.
\]
The Jacobian is equal to $S_n$. We thus get the desired result:

\[
\frac{1}{N} \frac{d^2N}{S_n dS_n d\theta} = \frac{1}{2\pi \sigma_n^2} \exp \left[ -\frac{(\vec{S}_n - \langle \vec{S}_n \rangle)^2}{2\sigma_n^2} \right].
\]

2.37 Verify the expressions (2.262) and (2.263) by explicitly integrating Eq. (2.260).

Hints:

\[
I_o(z) = \frac{1}{\pi} \int_0^\pi e^{az \cos(\theta)} d\theta
\]

\[
I_n(z) = \frac{1}{\pi} \int_0^\pi e^{az \cos(n\theta)} d\theta
\]

Solution: Let $\chi = \langle S_n \rangle / \sigma_n$, $\bar{z} = \chi \cos \theta$, $y = S_n / \sigma_n$, and $dy = dS_n / \sigma_n$. Eq. (2.260) can then be written

\[
\frac{1}{N} \frac{d^2N}{S_n dS_n d\theta} = \frac{1}{2\pi \sigma_n^2} \exp \left[ -\frac{(\vec{S}_n - \langle \vec{S}_n \rangle)^2}{2\sigma_n^2} \right],
\]

\[
= \frac{1}{2\pi \sigma_n^2} \exp \left[ -\frac{(S_n^2 - 2S_n \langle S_n \rangle \cos \theta + \langle S_n \rangle^2)}{2\sigma_n^2} \right],
\]

\[
= \frac{1}{2\pi \sigma_n^2} \exp \left[ -\frac{1}{2} \left( y^2 - 2y\chi \cos \theta + \chi^2 \right) \right],
\]

\[
= \frac{1}{2\pi \sigma_n^2} \exp \left[ -\frac{1}{2} \left( y^2 + \chi^2 \right) \right] \exp \left[ y\chi \cos \theta \right].
\]

Integration over the angle $\theta$ is readily achieved with the expression

\[
I_o(z) = \frac{1}{\pi} \int_0^\pi e^{az \cos(\theta)} d\theta,
\]

and one gets Eq. (2.263).

2.38 Show that the expression $U = \mathbf{A} \mathbf{V} \mathbf{A}^T$ is equivalent to Eq. (2.223).

Solution: The matrix equation stands for

\[
U_{ij} = \sum_{k=1}^n \sum_{k'=1}^n A_{ik} V_{kk'} A_{jk'}.
\]

The matrix $\mathbf{A}$ is defined as

\[
A_{ik} = \frac{\partial y_i}{\partial x_k} \bigg|_{\vec{x}=\vec{\mu}}.
\]

Substituting in the equation for $U_{ij}$ yields

\[
U_{ij} = \sum_{\alpha \beta = 1}^n \frac{\partial y_i}{\partial x_{k}} \frac{\partial y_j}{\partial x_{\alpha}} \bigg|_{\vec{x}=\vec{\mu}} V_{k\beta},
\]

which is equivalent to Eq. (2.223).