**Problem 1: Independence**

Let \((X_1, X_2)\) be multivariate normal distribution \(N(\mu, \Sigma)\) where \(\mu = (\mu_1, \mu_2)\) and \(\Sigma\) is a \(2 \times 2\) positive definite matrix. Assume that \(X_1\) and \(X_2\) are uncorrelated, i.e.,

\[
\Sigma_{12} = \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)] = 0
\]

Show that \(X_1\) and \(X_2\) are independent, i.e.,

\[
\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbb{P}(X_1 \leq x_1)\mathbb{P}(X_2 \leq x_2)
\]

Now let \(Y \sim U([0,1])\) be uniformly distributed in the interval \([0,1]\). Consider the random variables,

\[
Y_1 = \sin(2\pi Y) \\
Y_2 = \cos(2\pi Y)
\]

Show that \(Y_1\) and \(Y_2\) are uncorrelated. Are they independent?

**Solution:**

We note that \(\Sigma\) is a diagonal matrix. Then the distribution function of \((X_1, X_2)\) is

\[
f(x) = ce^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)} = ce^{-\frac{1}{2\pi\sigma^2_1} (x_1-\mu_1)^2} e^{-\frac{1}{2\pi\sigma^2_2} (x_2-\mu_2)^2}
\]

(here \(c\) is a normalizing constant). Since the distribution function factors into a product of the distribution functions for \(X_1\) and \(X_2\), they are independent.

To show that \(Y_1\) and \(Y_2\) are uncorrelated, we compute \(\mathbb{E}(Y_1) = \int_0^1 \sin(2\pi t) dt = 0\), \(\mathbb{E}(Y_2) = \int_0^1 \cos(2\pi t) dt = 0\) and \(\mathbb{E}(Y_1 Y_2) = \int_0^1 \sin(2\pi t) \cos(2\pi t) dt = 0\), and so \(Y_1\) and \(Y_2\) are uncorrelated. However, we have \(Y_1^2 + Y_2^2 = 1\) and so they cannot be independent.

**Problem 2: Linear transformation of Gaussian**

Let \(X = (X_1, \ldots, X_n)\) be a multivariate normal distribution \(N(\mu, \Sigma)\), where \(\mu = (\mu_1, \ldots, \mu_n)\) and \(\Sigma\) is an \(n \times n\) positive definite covariance matrix. For any invertible matrix \(A \in \mathbb{R}^{n \times n}\) and vector \(b \in \mathbb{R}^n\), define

\[
Y = AX + b
\]

(1) Find the distribution of \(Y\).

(2) Suppose you can sample from standard normal distribution \(N(0,1)\), how do you generate samples of random vectors from \(N(\mu, \Sigma)\)?

**Solution:**

Since \(Y\) is a linear transformation of a multivariate normal random variable, it is also multivariate normal. It has mean \(A\mu + b\) and covariance matrix \(A\Sigma A^T\).

Suppose we wish to sample from \(N(\mu, \Sigma)\). Let \(X_1\) be standard normal, then if \(X = (X_1, \ldots, X_n)\), we have \(AX + b\) is \(N(b, AA^T)\) distributed. A standard fact from linear algebra is that any symmetric positive definite matrix can be written in the form \(AA^T\) and so if we let \(b = \mu\) and pick \(A\) such that \(AA^T = \Sigma\), then \(AX + b\) has distribution \(N(\mu, \Sigma)\).

**Problem 3: Empirical variance**

Let \(X_1, X_2, \ldots, X_n\) be independent identically distributed random variables with mean \(\mu\) and variance \(\sigma^2\). Assume they also have finite fourth moment. Let \(\bar{X}_n = (X_1 + X_2 + \cdots + X_n)/n\) be the sample mean and

\[
\sigma_n^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X}_n)^2
\]
be the sample variance.

(1) Compute \( \mathbb{E}(\sigma_n^2) \).

(2) the LLN tells \( \bar{X} \rightarrow \mu \) as \( n \rightarrow \infty \), what is the limit of \( \sigma_n^2 \)? Try to write a formal proof.

Solution:
We have \( \sigma_n^2 = \frac{1}{n-1} \sum X_j^2 - 2X_j \bar{X} + \overline{X_n^2} = \frac{1}{n-1} (\sum X_j^2) - \frac{n}{n-1} \overline{X_n^2} \). Taking expectations, we have \( E(\sum X_j^2) = n(\sigma^2 + \mu^2) \) and \( E(\overline{X_n^2}) = Var(\overline{X_n}) + \mu^2 = \sigma^2/n + \mu^2 \). Combining these calculations gives \( E(\sigma_n^2) = \frac{n}{n-1} (\sigma^2 + \mu^2 - \sigma^2/n - \mu^2) = \sigma^2 \).

For convergence in probability, simply note that \( \frac{1}{n} \sum X_j^2 \rightarrow \sigma^2 + \mu^2 \) and \( \overline{X_n} \rightarrow \mu \) in probability, which implies that \( \sigma_n^2 \rightarrow \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \) in probability.

Problem 4: Poisson approximation

Let \( X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)} \) be independent Bernoulli random variables with \( P(X_j^{(n)} = 1) = p_n \) and \( P(X_j^{(n)} = 0) = 1 - p_n \). That is, for each \( n \) they are independent and identically distributed but the probability of “success” \( p_n \) depends on \( n \). Let \( Z_n = X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)} \). Suppose that \( p_n = \lambda/n \) with \( \lambda \) a positive number. Show that \( Z_n \) converges in distribution to a Poisson random variable \( Z \), where

\[
\mathbb{P}(Z = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots
\]

(Hint: compute the characteristic function of \( Z_n \) using the fact that \( E[e^{itX + Y}] = E[e^{itX}]E[e^{itY}] \) if \( X, Y \) are independent, and show that the characteristic function of \( Z_n \) converges to that of \( Z \).)

Solution:
We have by independence

\[
E(e^{itZ_n}) = \prod E(e^{itX_i^{(n)}}) = (1 + \frac{(e^{it} - 1)\lambda}{n})^n \rightarrow e^{\lambda(e^{it} - 1)}
\]

On the other hand, \( E(e^{itZ}) = \sum_{k=0}^{\infty} e^{itk} \frac{e^{-\lambda} \lambda^k}{k!} = e^{\lambda(e^{it} - 1)} \) and so as the characteristic functions converge for all \( t \), \( Z_n \) converges in distribution to \( Z \).

Problem 5: Estimating volatility

The Brownian motion process \( B_t \), \( 0 \leq t \leq T \), is a family of random variables with the following property:

Let \( \{t_i\}_{i=0}^N \) be any partition of the interval \([0, T]\) with \( t_0 = 0, t_N = T \) and \( t_i - t_{i-1} = T/N = \Delta N \). Then \( \{B_{t_i} - B_{t_{i-1}}\}_{i=1}^N \) are i.i.d. normal random variables with mean 0 and variance \( \Delta N \).

Consider the process \( X_t = \sigma B_t \) with \( \sigma > 0 \) representing the volatility. We observe \( \{X_t\}_{i=0}^N \) and want to estimate the volatility \( \sigma \). Define \( \Delta_i X = X_{t_i} - X_{t_{i-1}} \). Show that

\[
\frac{1}{\sqrt{\Delta N}} \left( \sum_{i=1}^N |\Delta_i X|^2 - T\sigma^2 \right) \text{ weakly } \Rightarrow \text{ Normal}(0, 2T\sigma^4)
\]

as \( N \rightarrow \infty, \Delta N \rightarrow 0, N\Delta N = T \). This also means \( \frac{1}{T} \sum_{i=1}^N |\Delta_i X|^2 \) is a consistent estimate of \( \sigma^2 \). (Hint: you may want to directly apply CLT and also compute the value of \( \mathbb{E}[Y^4] \) when \( Y \sim \text{Normal}(0, 1) \))

Solution:
We write
\[
\frac{1}{\sqrt{\Delta N}} \left( \sum_{i=1}^{N} |\Delta_i X|^2 - T\sigma^2 \right) = \sqrt{N} \left( \sum_{i=1}^{N} \frac{N}{\sqrt{T N}} |\Delta_i X|^2 - \sqrt{T}\sigma^2 \right)
\]

and note that \( \frac{N}{\sqrt{T N}} |\Delta_i X|^2 \) has distribution independent of \( N \) (it’s distributed as the square of a \( N(0, \sqrt{T\sigma^2}) \) random variable). Also, we compute \( E(\sum_{i=1}^{N} \frac{N}{\sqrt{T N}} |\Delta_i X|^2) = \sqrt{T}\sigma^2 \) and \( \text{Var}(\sqrt{N} \sum_{i=1}^{N} \frac{N}{\sqrt{T N}} |\Delta_i X|^2) = \frac{N^2}{T} (E(|\Delta_i X|^4)) - (E(|\Delta_i X|^2))^2 = 2T\sigma^4 \) where we use the fact that for \( Z \sim N(0, 1) \) we have \( E(Z^4) = 3 \).

Then we can apply the central limit theorem, which tells us that the expression converges weakly to a \( N(0, 2T\sigma^4) \) distribution.

**Problem 6: Monte Carlo integration**

Let \( f(x) = e^{-x^3} \).

(1) Compute \( \int_{0}^{1} f(x) dx \) using Monte Carlo. Compare with the true value (you may need to use some numerical integration package or just simply ask WolframAlpha).

(2) Do the simulation many times, and plot the histogram of the errors. Rescale your error in a way so that your new histogram looks like a bell-shape curve \( \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \) (you may need WolframAlpha again to compute some variance).

(3) Construct \( \alpha \)-confidence intervals for your Monte Carlo result for \( \alpha = 50\%, 75\%, 95\% \). How do you check (again by simulation) that the interval you constructed is indeed the \( \alpha \)-confidence interval? Is your way of checking becoming harder as \( \alpha \) increases?

**Solution:**

For part 1, you generate uniform random variables \( X_i \sim Unif(0, 1) \) and compute \( \sum_{i=1}^{n} \exp(-X_i^3) \) and compare with \( \int_{0}^{1} \exp(-x^3) dx = 0.807511 \).

For part 2, we repeat part 1 multiple times, generating estimates \( Y_i \), and plot a histogram of \( (Y_i - \bar{Y})/s \) where \( s \) is the sample variance of the \( Y_i \).

For part 3, we use the formula for the confidence interval \( (\overline{X} - \frac{c_{\alpha} \sigma}{\sqrt{n}}, \overline{X} + \frac{c_{\alpha} \sigma}{\sqrt{n}}) \). Here, the exact value of the variance should be calculated (although using the sample variance is okay for a large sample size), which can be done by computing \( \text{Var}(\exp(-X_i^3)) = \int_{0}^{1} e^{-2x^3} dx - (\int_{0}^{1} e^{-x^3} dx)^2 = 0.0386574 \). The values of \( c_\alpha \) for \( \alpha = 0.5, 0.75, 0.95 \) are 0.67, 1.15, 1.96.

We can test whether these confidence intervals are accurate by repeating this process many times and checking whether the true value lies in the confidence interval. This should occur with probability \( \alpha \). This method of checking gets harder as \( \alpha \) gets closer to 1, for example if \( \alpha \) were 0.999999 we would need to repeat the process many times to observe one instance of the true value lying outside the interval.

Attached below is some sample code and output for part 2.
Figure 1: Histogram of normalized errors

Sample code for part 2 (in R)

```r
error<-c()

for (i in 1:1000){
  u<-exp(-runif(1000)^3)
  est<-mean(u)
  error<-c(error, est+0.807511)
}

std.error<-(error-mean(error))/sd(error)
hist(std.error, freq=FALSE)

curve(dnorm(x,0,1),add=TRUE, col="red")
```