1. Let $X$ and $Y$ be two random variables whose joint density is given by

$$f(x,y) = \frac{1}{4} \left( 1 + xy(x^2 - y^2) \right), \quad |x| < 1, \quad |y| < 1.$$ 

Suppose $\Phi_Z(\cdot)$ denotes the characteristic function of random variable $Z$. Then, show that for all $\omega \in \mathbb{R}$,

$$\Phi_{X+Y}(\omega) = \Phi_X(\omega) \cdot \Phi_Y(\omega),$$

but $X$ and $Y$ are not independent. Does this contradict the sufficiency of factoring of characteristic functions for independence?

Solution:

(a) Note: In the figure shown below, the region below the line $x + y = 0.5$ from $x = 0.5$ to $x = 1$ could not be shaded due to some difficulties on tikz, but this region should be considered for computations.

It is easy to check that $f_X(x) = \frac{1}{2}$ for all $x \in (-1, 1)$. Similarly, $f_Y(y) = \frac{1}{2}$ for all $y \in (-1, 1)$. Clearly, $f(x,y) \neq f_X(x) \cdot f_Y(y)$, which implies that $X$ and $Y$ are not independent. We now compute the density of $Z = X + Y$. We have

$$P(Z \leq z) = \iint_{(x,y): x+y \leq z} f(x,y) \, dy \, dx.$$ 

The figure above depicts the region of integration for the case when $z = 0.5$. From the figure, we see that

$$P(Z \leq 0.5) = \int_{x=-1}^{-0.5} \int_{y=-1}^{1} f(x,y) \, dy \, dx + \int_{x=-0.5}^{1} \int_{y=-1}^{x-0.5} f(x,y) \, dy \, dx.$$ 

A careful evaluation of the cdf of $Z$ for all $z \in (-2, 2)$ yields the following pdf:

$$f_Z(z) = \begin{cases} \frac{2z}{4}, & -2 < z < 0 \\ \frac{2z}{4}, & 0 < z < 2 \\ 0, & \text{elsewhere}. \end{cases}$$

In other words, the pdf of $Z = X + Y$ is the result of the convolution of the pdf of $X$ and the pdf of $Y$ (although $X$ and $Y$ are not independent). This implies that the characteristic function of $Z$ is the product of the characteristic functions of $X$ and $Y$.

The sufficiency of factoring of characteristic functions for independence is not violated since this example only shows that $\Phi_{X+Y}(\omega) = \Phi_X(\omega) \cdot \Phi_Y(\omega)$ for all $\omega \in \mathbb{R}$, whereas independence requires $\Phi_U(\omega) = \Phi_X(\omega_1) \cdot \Phi_Y(\omega_2)$ for all $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$, where $U = (X,Y)$. 
2. Suppose $X$ is Gaussian $N(\mu, K)$ on $\mathbb{R}^n$, with $\det(K) > 0$. Show that there exists a matrix $B$ such that $Y = B(X - \mu)$ has $N(0, I)$ distribution, where $I$ is the $n \times n$ identity matrix.

**Solution:** We need the covariance matrix of $Y$ to be the identity matrix. Suppose such a matrix $B$ exists. Then, it should satisfy,

$$\text{Cov}(Y) = BB^T = I$$

Thus, if $K$ is diagonalizable, then we can write $K = U\Sigma U^T$ where the columns of $U$ are (orthogonal) eigenvectors of $K$ and $\Sigma$ is diagonal with corresponding eigenvalues. By appropriately normalizing $U$, we can make $\Sigma = I$ to obtain the desired matrix $B$.

3. Let $Y \sim N(0, 1)$ and $a > 0$. Define

$$Z = \begin{cases} Y, & \text{if } |Y| \leq a, \\ -Y, & \text{if } |Y| > a. \end{cases}$$

Show that $Z \sim N(0, 1)$.

**Solution:** Consider the characteristic function of $Z$,

$$\phi_Z(\omega) = \mathbb{E}[e^{i\omega Z}] = \mathbb{E}[e^{i\omega Y}1_{|Y|\leq a} + e^{-i\omega Y}1_{|Y|>a}]$$

Now, since the distribution of $Y$ is the same as that of $-Y$, we have,

$$\phi_Z(\omega) = \mathbb{E}[e^{i\omega Y}1_{|Y|\leq a} + e^{-i\omega Y}1_{|Y|>a}]$$

thereby allowing us to conclude that $\phi_Z(\omega) = \phi_Y(\omega)$. By uniqueness of characteristic functions, $Z \sim N(0, 1)$.

4. Let $X_1, \ldots, X_n$ be jointly Gaussian. Show that

$$X_j \text{ and } X_k \text{ are independent} \iff X_j \text{ and } X_k \text{ are uncorrelated}$$

for all $1 \leq j < k \leq n$.

**Solution:** The “only if” part is clear. Assume that $X_j$ and $X_k$ are uncorrelated with variances $\sigma_j$ and $\sigma_k$ respectively. WLOG, assume that $X_j$ and $X_k$ have zero mean. Define a Gaussian random vector

$$Y = [X_j \ X_k]$$

By hypothesis,

$$\text{Cov}(Y) = \begin{bmatrix} \sigma_j^2 & 0 \\ 0 & \sigma_k^2 \end{bmatrix}.$$ 

Now, the joint distribution of $Y$ is given by

$$\mathbb{P}_Y(y) = \frac{1}{2\pi\sigma_j\sigma_k} \exp\left\{ -\frac{1}{2} \left( \frac{x_j^2}{\sigma_j^2} + \frac{x_k^2}{\sigma_k^2} \right) \right\}$$

where $y = (x_j, x_k)$ which factors as

$$\mathbb{P}_Y(y) = P_{X_j}(x_j)P_{X_k}(x_k)$$

thereby allowing us to conclude that $X_j$ and $X_k$ are independent.
5. Let $(X, Y)$ be bivariate Normal $N(0, K)$ where

\[ K = \begin{bmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}. \]

Suppose that $|\rho| < 1$ (and hence $K$ is invertible). Then $(X, Y)$ has a density $f$. Show that the conditional density $f_{Y|X}(y|x)$ is the density of a univariate normal with mean $x \frac{\sigma_Y}{\sigma_X} \rho$ and variance $\sigma_Y^2 (1 - \rho^2)$.

**Solution:** The joint distribution of $(X, Y)$ is given by

\[ f_{XY}(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp\left\{ -\frac{1}{2} \left[ \left( y - x \frac{\sigma_Y}{\sigma_X} \rho \right)^2 + x^2 \frac{\sigma_Y^2}{\sigma_X^2} (1 - \rho^2) \right] \right\} \frac{1}{\sigma_Y^2 (1 - \rho^2)} \]

Now, by marginalizing over $y$, we get

\[ f_X(x) = \frac{1}{\sqrt{2\pi \sigma_X}} \exp\left\{ -\frac{1}{2} \cdot x^2 \frac{\sigma_Y^2}{\sigma_X^2} (1 - \rho^2) \right\} \]

Now, since $f_{Y|X}(y|x) = f_{XY}(x, y)/f_X(x)$, we have

\[ f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi (1 - \rho^2) \sigma_Y}} \exp\left\{ -\frac{1}{2} \cdot \left( y - x \frac{\sigma_Y}{\sigma_X} \rho \right)^2 \right\} \frac{1}{\sigma_Y^2 (1 - \rho^2)} \]

Thus, the conditional density $f_{Y|X}(y|x)$ is the density of a univariate normal with mean $x \frac{\sigma_Y}{\sigma_X} \rho$ and variance $\sigma_Y^2 (1 - \rho^2)$.

6. Let $X_n$ be a sequence of independent Bernoulli random variables defined by

\[ P(X_n = 1) = p_n, \quad P(X_n = 0) = 1 - p_n. \]

For the following cases, determine if the sequence converges in the mean squared sense, in probability and a.s.

(a) $p_n = \frac{1}{n}$
(b) $p_n = \frac{1}{n^2}$

**Solution:** Since the probability that $X_n$ assumes the value 0 increases with $n$, it is intuitively appealing to consider the identically 0 random variable as the limit. Assuming this, for any $\varepsilon > 0$, we have

\[ P(|X_n - 0| > \varepsilon) = \begin{cases} 0, & \varepsilon \geq 1 \\ p_n, & 0 < \varepsilon < 1. \end{cases} \]

Thus, it is clear that $P(|X_n - 0| > \varepsilon)$ goes to 0 as $n \to \infty$ for every $\varepsilon > 0$. Hence, $X_n \overset{L^2}{\to} 0$. Using this and the fact that $P(|X_n| \leq 1) = 1$, we can conclude that $X_n \overset{p}{\to} 0$.

In part (a), we observe that \( \sum_{n=1}^{\infty} P(X_n = 1) = +\infty \), and since $X_n$'s are independent, by Borel-Cantelli Lemma, we can conclude that $P(X_n = 1 \text{ i.o.}) = 1$, which implies that $X_n$ does not converge to 0 almost surely.

In part (b), we observe that \( \sum_{n=1}^{\infty} P(X_n = 1) < \infty \), and by using the Borel-Cantelli Lemma, we can conclude that $P(X_n = 1 \text{ i.o.}) = 0$, which implies that $X_n$ converges to 0 almost surely.
7. Let \((X_n, n \geq 1)\) be a sequence of iid Bernoulli random variables with \(P(X_n = 0) = P(X_n = 1)\) for all \(n \geq 1\).

(a) Determine if the sequence converges in distribution.

(b) Determine if the sequence converges in probability.

Solution:

(a) Since \(X_n\)'s are identically distributed, the sequence converges in distribution to that of a \(\text{Ber}(0.5)\) distributed random variable.

(b) Assume that \(X\) is a \(\text{Ber}(0.5)\) random variable, independent of \(X_n\) for all \(n \geq 1\). Then, \(X_n \overset{d}{\rightarrow} X\).

However,

\[
P(|X_n - X| > 0.5) = P(X_n = 1, X = 0) + P(X_n = 0, X = 1) = 0.5,
\]

which shows that the sequence \(X_n\) does not converge in probability.

8. Let \((X_n, n \geq 1)\) be a sequence of random variables with zero mean and finite variance. Define \(S_n = \sum_{k=1}^{n} X_k\).

(a) If \(X_i\)'s are iid, then show that \(S_n/n\) converges to 0 in \(L^2\) and in probability.

(b) Suppose that \(X_i\)'s are such that \(\text{Var}(S_n) = n^\alpha\), where \(0 < \alpha < 1\) is a fixed constant. Then, show that \(S_n/n\) converges to 0 almost surely.

Solution:

(a) Assume that the variance of \(X_i\)'s is \(\sigma^2\). Then, for any \(\varepsilon > 0\), by Chebyshev’s inequality, we have

\[
P \left( \left| \frac{S_n}{n} - 0 \right| > \varepsilon \right) \leq \frac{E[S_n^2]}{n^2 \varepsilon^2}.
\]

Since the mean of \(S_n\) is zero, \(E[S_n^2] = \text{Var}(S_n) = n \sigma^2\) by using the iid property. Plugging this into the RHS of (1) and observing that the RHS decreases to 0 as \(n \rightarrow \infty\) leads us to conclude that \(S_n/n\) converges to 0 in probability.

Also, we have

\[
E \left( \left( \frac{S_n}{n} - 0 \right)^2 \right) = \frac{\sigma^2}{n}
\]

which converges to 0 as \(n \rightarrow \infty\). Hence, \(S_n/n\) converges to 0 in \(L^2\).

(b) Since \(\text{Var}(S_n) = n^\alpha\), where \(0 < \alpha < 1\), for any \(\varepsilon > 0\), we have

\[
P \left( \left| \frac{S_n}{n} - 0 \right| > \varepsilon \right) \leq \frac{1}{n^{2 - \alpha} \varepsilon^2},
\]

which implies that \(\sum_{n=1}^{\infty} P \left( \left| \frac{S_n}{n} - 0 \right| > \varepsilon \right) < \infty\) for all \(\varepsilon > 0\). Thus, by using Borel-Cantelli Lemma, we conclude that \(S_n/n\) converges to 0 almost surely.