Probability and Random Process

COSE382

Joint, marginal, and conditional PMF

Definition 7.1.1 (Joint CDF). The *joint CDF* of r.v.s X and Y is the function $F_{X,Y} : \mathbb{R} \times \mathbb{R} :\to \mathbb{R}$ given by

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y).$$

Definition 7.1.2 (Joint PMF). The *joint PMF* of discrete r.v.s X and Y is the function $p_{X,Y} : \mathbb{R} \times \mathbb{R} :\to \mathbb{R}$ given by

$$p_{X,Y}(x, y) = P(X = x, Y = y).$$

The joint PMF of n discrete r.v.s is defined analogously.

Definition 7.1.3 (Marginal PMF). For discrete r.v.s X and Y, the *marginal PMF* of X is

$$P(X = x) = \sum_y P(X = x, Y = y).$$

The marginal PMF of X is the PMF of X, viewing X individually rather than jointly with Y.

Definition 7.1.4 (Conditional PMF). For discrete r.v.s X and Y, the conditional PMF of Y given $X = x$ is

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)}.$$

This is viewed as a function of y for fixed x.

Example 7.1.5 (2×2 table).

Suppose we randomly sample an adult male from U.S.

- X: indicator of being a current smoker
- Y: indicator of developing lung cancer someday.

1. Joint PMF: The following table the joint PMF of X and Y.

	Y=1	Y=0	Total
X=1	$\frac{5}{100}$	$\frac{20}{100}$	$\frac{25}{100}$
X=0	$\frac{3}{100}$	$\frac{72}{100}$	$\frac{75}{100}$
Total	$\frac{8}{100}$	$\frac{92}{100}$	$\frac{100}{100}$

2. Marginal: The *marginal distribution* of X is Bern(0.25) and the *marginal distribution* of Y is Bern(0.08).

3. Conditional:

$$P(Y=1|X=1) = \frac{P(X=1, Y=1)}{P(X=1)} = \frac{5/100}{25/100} = 0.2,$$

The *conditional distribution* of Y given $X = 1$ is Bern(0.2).
The *conditional distribution* of Y given $X = 1$ is Bern(0.04).

Definition 7.1.7 (Independence of discrete r.v.s). Random variables X and Y are independent if for all x and y,

$$F_{X,Y}(x, y) = F_X(x)F_Y(y).$$

If X and Y are discrete, this is equivalent to the condition

$$p_{XY}(x, y) = p_X(x)p_Y(y)$$

for all x and y, and it is also equivalent to the condition

$$P(Y = y|X = x) = P(Y = y)$$

for all y and all x such that $P(X = x) > 0$.

Example 7.1.9 (Chicken-egg) (VD)

- N: Number of eggs a chicken lays, $N \sim \text{Pois}(\lambda)$. Each egg hatches with p and fails to hatch with $1-p$ independently.
- X: Number of eggs that hatches, $X|(N=n) \sim \text{Bin}(n,p)$
- Y: Number of eggs that does not hatch $Y|(N=n) \sim \text{Bin}(n,1-p)$
- $X+Y=N$

What is the joint PMF of X and Y?

Solution:

$$\begin{aligned}
P(X=i,Y=j) &= \sum_{n=0}^{\infty} P(X=i,Y=j|N=n)P(N=n) \\
&= P(X=i,Y=j|N=i+j)P(N=i+j) \\
&= P(X=i|N=i+j)P(N=i+j) \\
&= \binom{i+j}{i} p^i q^j \cdot \frac{e^{-\lambda}\lambda^{i+j}}{(i+j)!} \\
&= \frac{e^{-\lambda p}(\lambda p)^i}{i!} \cdot \frac{e^{-\lambda q}(\lambda q)^j}{j!} = P(X=i)P(Y=j)
\end{aligned}$$

Joint PDF

Definition 7.1.12 (Joint PDF). If X and Y are continuous with joint CDF $F_{X,Y}$, their *joint PDF* is the derivative of the joint CDF with respect to x and y

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y).$$

We require valid joint PDFs to be nonnegative and integrate to 1:

$$f_{X,Y}(x,y) \geq 0, \text{and} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1.$$

The joint PDF of two r.v.s is the function we integrate to get the probability of a two-dimensional region. For example,

$$P(X < 3, 1 < Y < 4) = \int_{1}^{4} \int_{-\infty}^{3} f_{X,Y}(x,y) dx dy.$$

For a general set $\mathrm{A} \subseteq \mathbb{R}^2$,

$$P((X,Y) \in A) = \iint_A f_{X,Y}(x,y) dx dy.$$

Marginal and Conditional PDF

Definition 7.1.13 (Marginal PDF). For continuous r.v.s X and Y with joint PDF $f_{X,Y}$, the *marginal PDF* of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy.$$

This *is* the PDF of X, viewing X individually rather than jointly with Y.

Definition 7.1.14 (Conditional PDF). For continuous r.v.s X and Y with joint PDF $f_{X,Y}$, the *conditional PDF* of Y given $X = x$ is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

This is considered as a function of y for fixed x.

Theorem 7.1.17 (Continuous form of Bayes' rule and LOTP). For continuous r.v.s X and Y,

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$
$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dy.$$

Proof. By definition of conditional PDFs, we have

$$f_{Y|X}(y|x)f_X(x) = f_{X,Y}(x,y) = f_{X|Y}(x|y)f|Y(y).$$

The continuous version of Bayes' rule follows immediately from dividing by $f_X(x)$. The continuous version of LOTP follows immediately from integrating with respect to y:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dy.$$

Definition 7.1.18 (Independence of continuous r.v.s). Random variables X and Y are independent if for all x and y,

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

If X and Y are continuous with joint PDF $f_{X,Y}$, this is equivalent to the condition

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

for all x and y, and it is also equivalent to the condition

$$f_{Y|X}(y|x) = f_Y(y)$$

for all y and all x such that $f_X(x) > 0$.

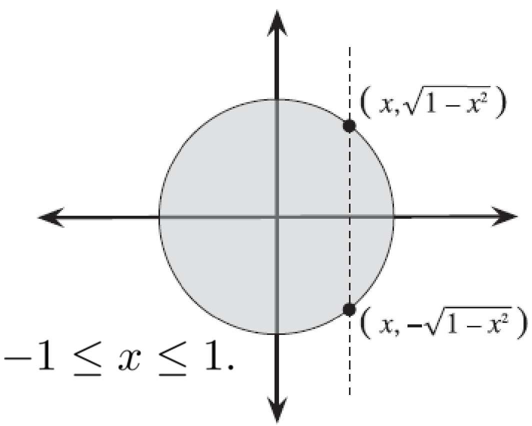
Proposition 7.1.20 Suppose that the joint PDF $f_{X,Y}$ of X and Y factors are

$$f_{X,Y}(x,y) = g(x)h(y)$$

for all x and y, where g and h are nonnegative functions. Then X and Y are independent.

Example 7.1.22 (Uniform on a region in the plane). Consider r.v.s X, Y with the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & \text{if } x^2 + y^2 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$



The marginal distributions are

$$\begin{aligned} f_X(x) &= \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1. \\ f_Y(y) &= \frac{2}{\pi}\sqrt{1-y^2}, -1 \le y \le 1. \end{aligned}$$

Suppose we observe $X = x$. The conditional distribution of Y given $X = x$ is

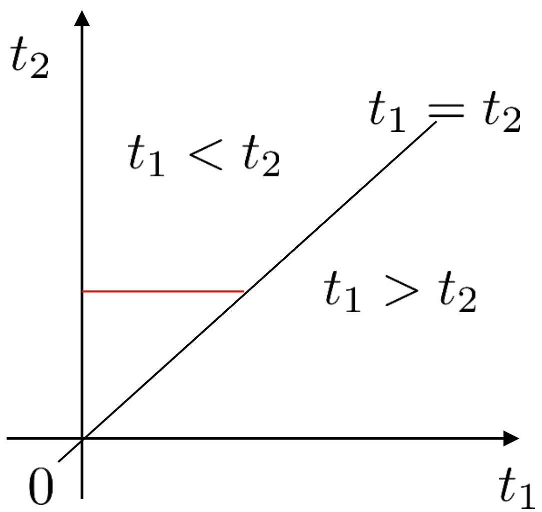
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\frac{1}{\pi}}{\frac{2}{\pi}\sqrt{1-x^2}} = \frac{1}{2\sqrt{1-x^2}}$$

for $-\sqrt{1-x^2} \le y \le \sqrt{1-x^2}$ and 0 otherwise.

Example 7.1.23 (Comparing Exponentials of different rates) (VD).

Let T_1 ~Expo(λ_1) and T_2 ~Expo(λ_2) be independent. Find $P(T_1 < T_2)$.

Solution: We just need to integrate the joint PDF of T_1 and T_2 over the appropriate region, which is all (t_1, t_2) with $t_1 > 0, t_2 > 0$, and $t_1 < t_2$. This yields



$$\begin{aligned}
P(T_1 < T_2) &= \int_{\{t_1<t_2\}\subset\mathbb{R}^2} f_{T_1T_2}(t_1,t_2)dt_1dt_2 \\
&= \int_{t_2=0}^{t_2=\infty}\int_{t_1=0}^{t_1=t_2} \lambda_1 e^{-\lambda_1 t_1}\lambda_2 e^{-\lambda_2 t_2}dt_1dt_2 \\
&= \int_{t_2=0}^{t_2=\infty}\left(\int_{t_1=0}^{t_1=t_2} \lambda_1 e^{-\lambda_1 t_1}dt_1\right)\lambda_2 e^{-\lambda_2 t_2}dt_2 \\
&= \int_{t_2=0}^{t_2=\infty}(1-e^{-\lambda_1 t_2})\lambda_2 e^{-\lambda_2 t_2}dt_2 \\
&= 1-\int_{t_2=0}^{t_2=\infty}\lambda_2 e^{-(\lambda_1+\lambda_2)t_2}dt_2 \\
&= 1-\frac{\lambda_2}{\lambda_1+\lambda_2} = \frac{\lambda_1}{\lambda_1+\lambda_2}
\end{aligned}$$

2D LOTUS

Theorem 7.2.1 (2D LOTUS). Let $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. If X and Y are discrete, then

$$E(g(X,Y)) = \sum_x \sum_y g(x,y) P(X=x, Y=y).$$

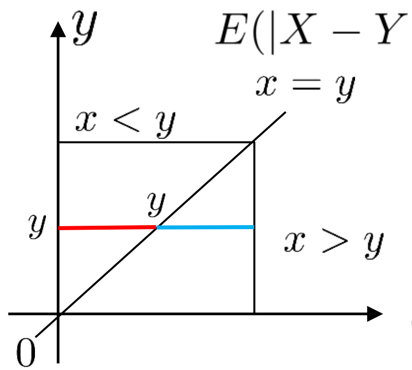
If X and Y are continuous with joint PDF $f_{X,Y}$, then

$$E(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy.$$

Example 7.2.2 (Expected distance between two Uniforms). (VD)

For $X, Y \overset{\text{i.i.d}}{\sim} \text{Unif}(0,1)$, find $E(|X-Y|)$.

Solution: Since $f_{X,Y}(x,y) = 1$ on the unit square $(x,y) : x, y \in [0,1]$, 2D LOTUS gives



$$\begin{aligned}
E(|X-Y|) &= \int_{y=0}^{y=1}\int_{x=0}^{x=1} |x-y| dxdy \\
&= \int_{y=0}^{y=1}\int_{x=y}^{x=1} (x-y)dxdy + \int_{y=0}^{y=1}\int_{x=0}^{x=y} (y-x)dxdy \\
&= 2\int_{y=0}^{y=1}\int_{x=y}^{x=1} (x-y)dxdy = 1/3.
\end{aligned}$$

Let $M = \max(X,Y)$ and $L = \min(X,Y)$. Note that $M+L = X+Y$ and $M-L = |X-Y|$,

$$\begin{aligned}
E(M+L) &= E(X+Y) = 1, \\
E(M-L) &= E(|X-Y|) = 1/3.
\end{aligned}$$

We have $E(M) = 2/3$ and $E(L) = 1/3$.

Example 7.2.3 (Expected distance between two Normals) (VD).

For $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ find $E(|X-Y|)$.

Solution:

We could again use 2D LOTUS, giving

$$E(|X-Y|) = \int_{y=-\infty}^{y=\infty} \int_{x=-\infty}^{x=\infty} |x-y| \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dx dy,$$

but an easier solution uses the fact that the sum or difference of independent Normals is Normal, as we proved using MGFs. Then $X - Y \sim \mathcal{N}(0, 2)$. For $Z \sim \mathcal{N}(0, 1)$

$$E(|Z|) = \int_{z=-\infty}^{z=\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 2 \int_{z=0}^{z=\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \sqrt{\frac{2}{\pi}},$$

Hence, $E(|X-Y|) = E(\sqrt{2}\,|Z|) = \dfrac{2}{\sqrt{\pi}}$.

Example 7.2.4 (Linearity via 2D LOTUS).

Let X and Y be continuous r.v.s

$$\begin{aligned}
E(X+Y) &= \int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty}(x+y)f_{X,Y}(x,y)dxdy \\
&= \int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty}xf_{X,Y}(x,y)dxdy + \int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty}yf_{X,Y}(x,y)dxdy. \\
&= E(X)+E(Y).
\end{aligned}$$

where we have used

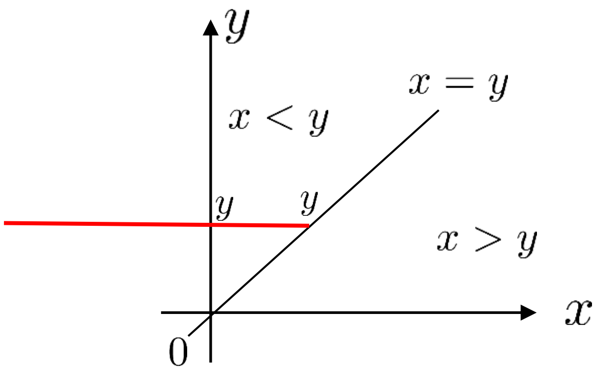
$$\begin{aligned}
\int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty}yf_{X,Y}(x,y)dxdy &= \int_{y=-\infty}^{y=\infty}y\int_{x=-\infty}^{x=\infty}f_{X,Y}(x,y)dxdy \\
&= \int_{y=-\infty}^{y=\infty}yf_Y(y)dy \\
&= E(Y),
\end{aligned}$$

Example (Probability of tie for continuous r.v.) (VD).

Let X and Y be i.i.d. t continuous r.v.s, to prove $P(X=Y)=0$ it is sufficient to show $P(X<Y)=1/2$, since $P(Y<X)=1/2$ by symmetry.

$$1 = P(X<Y) + P(Y<X) + P(X=Y)$$

$$\begin{aligned} P(X<Y) &= \int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=y} f_{X,Y}(x,y)dxdy = \int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=y} f_X(x)f_X(y)dxdy \\ &= \int_{y=-\infty}^{y=\infty} f_X(y)\int_{x=-\infty}^{x=y} f_X(x)dxdy = \int_{y=-\infty}^{y=\infty} f_X(y)F_X(y)dy \\ &= E(F_X(y)) = E(F_Y(y)) = \frac{1}{2} \end{aligned}$$



We have used the fact $F_Y(Y) \sim \text{Unif}(0,1)$
(Theorem 5.3.1, Universality of Unifrom)

Covariance and correlation

Definition 7.3.1 (Covariance). The *covariance* between r.v.s X and Y is

$$\mathrm{Cov}(X,Y) := E((X-EX)(Y-EY)) = \mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y).$$

- X and Y are *uncorrelated* if $\mathrm{Cov}(X,Y) = 0$

Theorem 7.3.2. If X and Y are independent, then they are uncorrelated.

Proof.

$$\begin{aligned}
E(XY) &= \int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty} xy f_X(x) f_Y(y)\,dx\,dy = \int_{y=-\infty}^{y=\infty} y f_Y(y)\left(\int_{x=-\infty}^{x=\infty} x f_X(x)\,dx\right)dy \\
&= \int_{x=-\infty}^{x=\infty} x f_X(x)\,dx \int_{y=-\infty}^{y=\infty} y f_Y(y)\,dy \\
&= E(X)E(Y).
\end{aligned}$$

Properties of Covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$
2. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
3. $\text{Cov}(X, c) = 0$ for any constant c
4. $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$ for any constant a
5. $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$
6. $\text{Cov}(X + Y, Z + W) = \text{Cov}(X, Z) + \text{Cov}(X, W) + \text{Cov}(Y, Z) + \text{Cov}(Y, W)$
7. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$

$$\text{Var}(X_1 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) + 2\sum_{i<j} \text{Cov}(X_i, X_j)$$

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)$$

Definition 7.3.4 (Correlation). The *correlation* between r.v.s X and Y is

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

(This is undefined in the degenerate cases $\mathrm{Var}(X) = 0$ or $\mathrm{Var}(Y) = 0$.)

Theorem 7.3.5 (Correlation bounds). For any r.v.s X and Y,

$$-1 \le \mathrm{Corr}(X,Y) \le 1.$$

Proof. Without loss of generality we can assume X and Y have variance 1, since scaling does not change the correlation. Let $\rho = \mathrm{Corr}(X,Y) = \mathrm{Cov}(X,Y)$, then

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) = 2 + 2\rho \ge 0,\\ \mathrm{Var}(X-Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X,Y) = 2 - 2\rho \ge 0.\end{aligned}$$

Thus, $-1 \le \rho \le 1$.

Example 7.3.6 (Exponential max and min) (VD).

Let X and Y be i.i.d. Expo(1) r.v.s. Find the correlation between $\max(X,Y)$ and $\min(X,Y)$.

Solution: Let $M = \max(X,Y)$ and $L = \min(X,Y)$. By the memoryless property and previous results, we know that $L \sim \text{Expo}(2)$, $M - L \sim$Expo(1), and $M - L$ is independent of L. We now have

$$\begin{aligned} \text{Cov}(M,L) &= \text{Cov}(M-L+L,L) = \text{Cov}(M-L,L) + \text{Cov}(L,L) = 0 + \text{Var}(L) = \frac{1}{4}, \\ \text{Var}(M) &= \text{Var}(M-L+L) = \text{Var}(M-L) + \text{Var}(L) = 1 + \frac{1}{4} = \frac{5}{4}, \end{aligned}$$

and

$$\text{Corr}(M,L) = \frac{\text{Cov}(M,L)}{\sqrt{\text{Var}(M)\text{Var}(L)}} = \frac{\frac{1}{4}}{\sqrt{\frac{5}{4}\cdot\frac{1}{4}}} = \frac{1}{\sqrt{5}}.$$

It makes sense that the correlation is positive because M is constrained to be at least as large as L.

Joint MGF: MGF for joint distribution

Definition 7.5.6 (Joint MGF). The *joint MGF* of a random vector $\mathbf{X} = (X_1, \ldots, X_k)$ is the function which takes a vector of real values $\mathbf{t} = (t_1, \ldots, t_k)$ and returns

$$M_{\mathbf{X}}(\mathbf{t}) = E(e^{\mathbf{t}\mathbf{X}^T}) = E(e^{t_1 X_1 + \cdots + t_k X_k})$$

Especially, for two random variables X and Y, the joint MGF is given as

$$M_{X,Y}(t_1, t_2) = E(e^{t_1 X + t_2 Y}) = \int\int e^{t_1 x + t_2 y} f_{XY}(x, y) dx dy$$

Property When $\mathbf{X}$ is a random vector of independent random variables, then the joint MGF is given as a product of MGF's.

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= E(e^{\mathbf{t}\mathbf{X}^T}) = E(e^{t_1 X_1 + \cdots + t_k X_k}) = E(\prod_{i=1}^{k} e^{t_k X_k}) \\ &= \prod_{i=1}^{k} E(e^{t_k X_k}) = \prod_{i=1}^{k} M_{X_k}(t_k) \end{aligned}$$

Furthermore, if a joint MGF is given as a product of marginal MGFs, the marginal random variables are indpendent.

Multivariate Normal (Jointly Gaussian)

Definition 7.5.1 (Jointly Gaussian distribution).

A random vector $\mathbf{X} = (X_1, \ldots, X_k)$ is said to have a *jointly Gaussian or Multivariate Normal* (MVN) distribution if every linear combination of the X_j has a Gaussian distribution. That is,

$$t_1 X_1 + \cdots + t_k X_K$$

is a Gaussian distribution for any choice of constants $t_1, \ldots, t_k$.

- A Special case is $k = 2$; this distribution is called the *Bivariate Normal* (BVN).
- If $(X_1, \ldots, X_k)$ is jointly Gaussian, then each X_i is Gaussian.
- $(X_1, \ldots, X_k)$ may not jointly Normal, even if $X_1, \ldots, X_k$ are Gaussian.

Example 7.5.2 (Non-example of jointly Gaussian). Example of two Gaussian r.v.s X and Y whose joint distribution is not jointly Gaussian. Let $X \sim \mathcal{N}(0,1)$, and let

$$S = \begin{cases} 1 & \text{with probabilty } 1/2 \\ -1 & \text{with probability} 1/2 \end{cases}$$

be a *random sign* independent of X. Then $Y = SX$ is a standard Normal r.v.;

$$\begin{aligned} P(Y < t) &= P(SX < t) = P(X < t|S = 1)\frac{1}{2} + P(-X < t|S = -1)\frac{1}{2} \\ &= \Phi(t)\frac{1}{2} + P(X > -t)\frac{1}{2} = \Phi(t) \end{aligned}$$

However, (X, Y) is not jointly Gaussian because $P(X + Y = 0) = P(S = -1) = 1/2$, which implies that $X + Y$ can't be Gaussian (or, any continuous distribution).

Theorem 7.5.7. Within an MVN random vector, uncorrelated implies independent. If (X, Y) is Bivariate Normal and $\text{Corr}(X, Y) = 0$, then X and Y are independent.

Proof. Let (X, Y) be Bivariate Normal with $E(X) = \mu_1, E(Y) = \mu_2$, $\text{Var}(X) = \sigma_1^2$, $\text{Var}(Y) = \sigma_2^2$, and $\text{Corr}(X, Y) = \rho$. The joint MGF is

$$\begin{aligned} M_{X,Y}(s,t) = E(e^{sX+tY}) &= \exp\left(s\mu_1 + t\mu_2 + \frac{1}{2}\text{Var}(sX+tY)\right) \\ &= \exp\left(s\mu_1 + t\mu_2 + \frac{1}{2}(s^2\sigma_1^2 + t^2\sigma_2^2 + 2st\sigma_1\sigma_2\rho)\right). \end{aligned}$$

If $\rho = 0$, the joint MGF reduces to

$$M_{X,Y}(s,t) = \exp\left(s\mu_1 + t\mu_2 + \frac{1}{2}(s^2\sigma_1^2 + t^2\sigma_2^2)\right) = M_X(s)M_Y(t)$$

which implies $E(e^{sX}e^{tY}) = E(e^{sX})E(e^{tY})$. Therefore X and Y are independent.

Example 7.5.8 (Independence of sum and difference). Let $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ Find the joint distribution of $(X + Y, X - Y)$.

Solution: Since $(X + Y, X - Y)$ is Bivariate Normal and

$$\mathrm{Cov}(X + Y, X - Y) = \mathrm{Var}(X) - \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, X) - \mathrm{Var}(Y) = 0,$$

$X + Y$ is independent of $X - Y$. Furthermore, they are i.i.d. $\mathcal{N}(0,\ 2)$.

Similarly, we have that if $X \sim \mathcal{N}(\mu_1, \sigma^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma^2)$ are independent (with the same variance), then $X + Y$ is independent of $X - Y$.