भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology
Hyderabad

# Probability and Random Variables

Teaching Learning Center
IIT Hyderabad

G V V Sharma*

CONTENTS



*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in. 

***Abstract*—This book provides a simple introduction to probability and random variables. The contents are largely based on NCERT textbooks from Class 9-12.**

## 1 Sum of Independent Random Variables

### *1.1 The Uniform Distribution*

Two dice, one blue and one grey, are thrown at the same time. The event defined by the sum of the two numbers appearing on the top of the dice can have 11 possible outcomes 2, 3, 4, 5, 6, 6, 8, 9, 10, 11 and 12. A student argues that each of these outcomes has a probability $\frac{1}{11}$. Do you agree with this argument? Justify your answer.

1.1.1. *The Uniform Distribution:* Let $X_i \in \{1,2,3,4,5,6\}, i = 1,2,$ be the random variables representing the outcome for each die. Assuming the dice to be fair, the probability mass function (pmf) is expressed as

$$p_{X_i}(n) = \Pr(X_i = n) = \begin{cases} \frac{1}{6} & 1 \le n \le 6 \\ 0 & otherwise \end{cases} \tag{1.1.1.1}$$

The desired outcome is

$$X = X_1 + X_2, \tag{1.1.1.2}$$

$$\implies X \in \{1,2,\dots,12\} \tag{1.1.1.3}$$

The objective is to show that

$$p_X(n) \neq \frac{1}{11} \tag{1.1.1.4}$$

1.1.2. *Convolution:* From (1.1.1.2),

$$p_X(n) = \Pr(X_1 + X_2 = n) = \Pr(X_1 = n - X_2) \tag{1.1.2.1}$$

$$= \sum_k \Pr(X_1 = n - k | X_2 = k)\, p_{X_2}(k) \tag{1.1.2.2}$$

after unconditioning. $\because$ $X_1$ and $X_2$ are independent,

$$\Pr(X_1 = n - k | X_2 = k) = \Pr(X_1 = n - k) = p_{X_1}(n-k) \tag{1.1.2.3}$$

From (1.1.2.2) and (1.1.2.3),

$$p_X(n) = \sum_k p_{X_1}(n-k) p_{X_2}(k) = p_{X_1}(n) * p_{X_2}(n) \tag{1.1.2.4}$$

where $*$ denotes the convolution operation. Substituting from (1.1.1.1) in (1.1.2.4),

$$p_X(n) = \frac{1}{6}\sum_{k=1}^{6} p_{X_1}(n-k) = \frac{1}{6}\sum_{k=n-6}^{n-1} p_{X_1}(k) \tag{1.1.2.5}$$

$$\because p_{X_1}(k) = 0, \quad k \leq 1, k \geq 6. \tag{1.1.2.6}$$

From (1.1.2.5),

$$p_X(n) = \begin{cases} 0 & n < 1 \\ \frac{1}{6}\sum_{k=1}^{n-1} p_{X_1}(k) & 1 \leq n-1 \leq 6 \\ \frac{1}{6}\sum_{k=n-6}^{6} p_{X_1}(k) & 1 < n-6 \leq 6 \\ 0 & n > 12 \end{cases} \tag{1.1.2.7}$$

Substituting from (1.1.1.1) in (1.1.2.7),

$$p_X(n) = \begin{cases} 0 & n < 1 \\ \frac{n-1}{36} & 2 \leq n \leq 7 \\ \frac{13-n}{36} & 7 < n \leq 12 \\ 0 & n > 12 \end{cases} \tag{1.1.2.8}$$

satisfying (1.1.1.4).

1.1.3. *The Z-transform:* The Z-transform of $p_X(n)$ is defined as

$$P_X(z) = \sum_{n=-\infty}^{\infty} p_X(n) z^{-n}, \quad z \in \mathbb{C} \tag{1.1.3.1}$$

From (1.1.1.1) and (1.1.3.1),

$$P_{X_1}(z) = P_{X_2}(z) = \frac{1}{6}\sum_{n=1}^{6} z^{-n} \tag{1.1.3.2}$$

$$= \frac{z^{-1}\left(1 - z^{-6}\right)}{6\left(1 - z^{-1}\right)}, \quad |z| > 1 \tag{1.1.3.3}$$

upon summing up the geometric progression.

$$\because p_X(n) = p_{X_1}(n) * p_{X_2}(n), \tag{1.1.3.4}$$

$$P_X(z) = P_{X_1}(z) P_{X_2}(z) \tag{1.1.3.5}$$

The above property follows from Fourier analysis and is fundamental to signal processing. From (1.1.3.3) and (1.1.3.5),

$$P_X(z) = \left\{\frac{z^{-1}\left(1 - z^{-6}\right)}{6\left(1 - z^{-1}\right)}\right\}^2 \tag{1.1.3.6}$$

$$= \frac{1}{36}\frac{z^{-2}\left(1 - 2z^{-6} + z^{-12}\right)}{\left(1 - z^{-1}\right)^2} \tag{1.1.3.7}$$

Using the fact that

$$p_X(n-k) \overset{\mathcal{H}}{\longleftrightarrow} Z P_X(z) z^{-k}, \tag{1.1.3.8}$$

$$n u(n) \overset{\mathcal{H}}{\longleftrightarrow} Z \frac{z^{-1}}{\left(1 - z^{-1}\right)^2} \tag{1.1.3.9}$$

after some algebra, it can be shown that

$$\frac{1}{36}\left[(n-1)\, u(n-1) - 2\,(n-7)\, u(n-7) + (n-13)\, u(n-13)\right] \overset{\mathcal{H}}{\longleftrightarrow} Z \frac{1}{36}\frac{z^{-2}\left(1 - 2z^{-6} + z^{-12}\right)}{\left(1 - z^{-1}\right)^2} \tag{1.1.3.10}$$

where

$$u(n) = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \tag{1.1.3.11}$$

From (1.1.3.1), (1.1.3.7) and (1.1.3.10)

$$p_X(n) = \frac{1}{36}\left[(n-1)\,u(n-1) - 2\,(n-7)\,u(n-7) + (n-13)\,u(n-13)\right] \tag{1.1.3.12}$$

which is the same as (1.1.2.8). Note that (1.1.2.8) can be obtained from (1.1.3.10) using contour integration as well.

1.1.4. The experiment of rolling the dice was simulated using Python for 10000 samples. These were generated using Python libraries for uniform distribution. The frequencies for each outcome were then used to compute the resulting pmf, which is plotted in Figure 1.1.4.1. The theoretical pmf obtained in (1.1.2.8) is plotted for comparison.

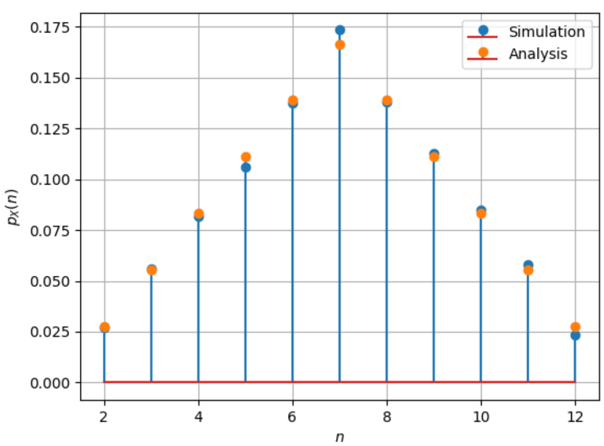


Fig. 1.1.4.1: Plot of $p_X(n)$. Simulations are close to the analysis.

1.1.5. The python code is available in

```
/codes/sum/dice.py
```

## 2 Cumulative Distribution Function

### 2.1 The Bernoulli Distribution

2.1.1. Find the probability of getting a head when a coin is tossed once. Also find the probability of getting a tail.

**Solution:** Let the random variable be $X \in \{0, 1\}$. Then

$$\Pr(X = 0) = \Pr(X = 1) = \frac{1}{2} \tag{2.1.1.1}$$

TABLE 2.1.3.1

| **Colour** | **X** | **Number** |
|---|---|---|
| Blue | 0 | $n(X = 0)$ |
| Green | 1 | $n(X = 1)$ |

The following code simulates the event for 100 coin tosses

```
codes/bernoulli/coin.py
```

2.1.2. *Bernoulli Distribution:* In general the binomial distribution is defined using the PMF

$$p_X(n) = \begin{cases} p & n = 1 \\ 1 - p & n = 0 \\ \text{otherwise} & \end{cases} \tag{2.1.2.1}$$

2.1.3. A jar contains 24 marbles, some are green and others are blue. If a marble is drawn at random from the jar, the probability that it is green is $\frac{2}{3}$. Find the number of blue balls (marbles) in the jar.

**Solution:** Let random variable $X \in \{0, 1\}$ denote the outcomes of the experiment of drawing a marble from a jar as shown in Table 2.1.3.1 From the given information,

$$p_X(1) = \frac{2}{3} \tag{2.1.3.1}$$

$$\implies p = 1 - p_X(1) = \frac{1}{3} \tag{2.1.3.2}$$

$$n(X = 0) + n(X = 1) = 24 \tag{2.1.3.3}$$

$\because$

$$p = \frac{n(X = 0)}{n(X = 0) + n(X = 1)}, \tag{2.1.3.4}$$

from (2.1.3.4) and (2.1.3.3),

$$n(X = 0) = p\,\{n(X = 0) + n(X = 1)\} \tag{2.1.3.5}$$

$$= \frac{1}{3} \times 24 = 8. \tag{2.1.3.6}$$

The following code generates the number of blue marbles

```
codes/bernoulli/bernoulli.py
```

### 2.2 The Binomial Distribution

In a hurdle race, a player has to cross 10 hurdles. The probability that he will clear each hurdle is $\frac{5}{6}$.

What is the probability that he will knock down fewer than 2 hurdles?

2.2.1. Let $X_i \in \{0, 1\}$ represent the *ith* hurdle where 1 denotes a hurdle being knocked down. Then, $X_i$ has a bernoulli distribution with parameter

$$p = 1 - \frac{5}{6} = \frac{1}{6} \tag{2.2.1.1}$$

2.2.2. *The Binomial Distribution:* Let

$$X = \sum_{i=1}^{n} X_i \tag{2.2.2.1}$$

where $n$ is the total number of hurdles. Then $X$ has a binomial distribution. Then, for

$$p_{X_i}(n) \overset{\mathcal{Z}}{\rightleftharpoons} P_{X_i}(z), \tag{2.2.2.2}$$

yielding

$$P_{X_i}(z) = 1 - p + pz^{-1} \tag{2.2.2.3}$$

with Using the fact that $X_i$ are i.i.d.,

$$P_X(z) = \left(1 - p + pz^{-1}\right)^n \tag{2.2.2.4}$$

$$= \sum_{k=0}^{n} {}^nC_k p^k (1-p)^{n-k} z^{-k} \tag{2.2.2.5}$$

$$\implies p_X(k) = \begin{cases} {}^nC_k p^{n-k} (1-p)^k & 0 \le k \le n \\ 0 & \text{otherwise} \end{cases} \tag{2.2.2.6}$$

The cumulative distribution function of $X$ is defined as

$$F_X(r) = \Pr(X \le r) = \sum_{k=0}^{r} {}^nC_k p^k (1-p)^{n-k} \tag{2.2.2.7}$$

upon substituting from (2.2.2.6).

2.2.3. *Evaluationg the Probability:* Substituting from (2.2.1.1) in (2.2.2.7),

$$\Pr(X < 2) = F_X(1) \tag{2.2.3.1}$$

$$= \sum_{k=0}^{1} {}^nC_k \left(\frac{5}{6}\right)^{10-k} \left(\frac{1}{6}\right)^k \tag{2.2.3.2}$$

$$= 3\left(\frac{5}{6}\right)^{10} = 0.4845167486695371 \tag{2.2.3.3}$$

which is the desired probability.

2.2.4. The following code verifies the above result.

```
codes/binomial/binomial.py
```

## 3 Central Limit Theorem: Gaussian Distribution

### *3.1 Bernoulli to Gaussian*

3.1.1 *Mean :* The mean of the bernoulli distribution is

$$\mu = E(X_i) = \sum_{k=0}^{1} k p_{X_i}(k) = p = \frac{1}{6} \tag{3.1.1}$$

3.1.2 *Moment:* The moment of the distribution is defined as

$$E\left(X_i^r\right) = \sum_{k=0}^{1} k^r p_{X_i}(k) = p = \frac{1}{6} \tag{3.2.1}$$

3.1.3 *Variance :* The variance of the bernoulli distribution is defined as

$$\sigma^2 = E(X - E(X))^2 = E\left(X^2\right) - E^2(X) \tag{3.3.1}$$

$$= p - p^2 = p(1-p) = \frac{5}{36} \tag{3.3.2}$$

The standard deviation

$$\sigma = \sqrt{p(1-p)} \tag{3.3.3}$$

3.1.4 *The Gaussian Distribution:* Define

$$G = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_i - \mu}{\sigma} \tag{3.4.1}$$

3.1.5 *Approximating Binomial Using Gaussian:* From (3.4.1) and (2.2.2.1),

$$X \approx \sigma\sqrt{n}G + n\mu \tag{3.5.1}$$

$$\implies F_X(k) = \Pr\left(\sigma\sqrt{n}G + n\mu \le k\right) \tag{3.5.2}$$

$$= F_G\left(\frac{k - n\mu}{\sigma\sqrt{n}}\right) \approx \phi\left(\frac{k - n\mu}{\sigma\sqrt{n}}\right) \tag{3.5.3}$$

where

$$\phi_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < \infty \tag{3.5.4}$$

3.1.6 The probability density function (PDF) of $G$ is

$$p_G(x) = \frac{d}{dx}F_X(x) \tag{3.6.1}$$

$$= \frac{1}{\sigma\sqrt{n}}\phi'\left(\frac{k-n\mu}{\sigma\sqrt{n}}\right) \tag{3.6.2}$$

For large $n$, $G$ is a continuous distribution with probability density function (PDF)

$$p_G(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), -\infty < x < \infty, \tag{3.6.3}$$

3.1.7 *Evaluationg the Probability:* From 3.5.3 and 3.6.2,

$$\Pr(X \le 1) = F_G(1) = p_G(0) + p_G(1) \tag{3.7.1}$$

$$\approx 0.41299463887797094 \tag{3.7.2}$$

which is close to (2.2.3.3).

### *3.2 Uniform to Gaussian*

3.2.1 Generate $10^6$ samples of the random variable

$$X = \sum_{i=1}^{12} U_i - 6 \tag{3.1.1}$$

using a C program, where $U_i, i = 1, 2, \ldots, 12$ are a set of independent uniform random variables between 0 and 1 and save in a file called gau.dat
**Solution:** Download the following files and execute the C program.

```
codes/cdf/exrand.c
codes/cdf/coeffs.h
```

3.2.2 Load gau.dat in python and plot the empirical CDF of $X$ using the samples in gau.dat. What properties does a CDF have?
**Solution:** The CDF of $X$ is plotted in Fig. 3.2

3.2.3 Load gau.dat in python and plot the empirical PDF of $X$ using the samples in gau.dat. The PDF of $X$ is defined as

$$p_X(x) = \frac{d}{dx}F_X(x) \tag{3.3.1}$$

What properties does the PDF have?
**Solution:** The PDF of $X$ is plotted in Fig. 3.3 using the code below

```
codes/clt/pdf_plot.py
```

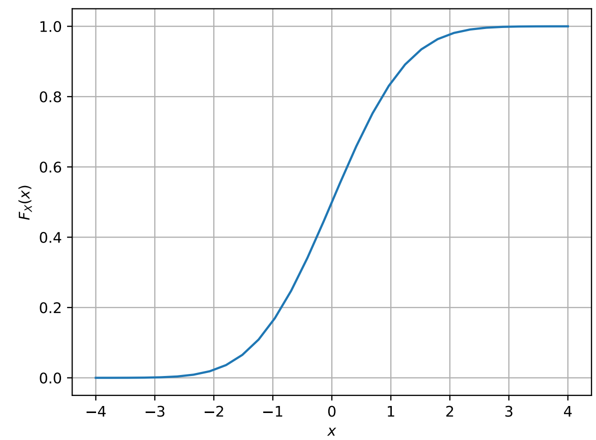


Fig. 3.2: The CDF of $X$

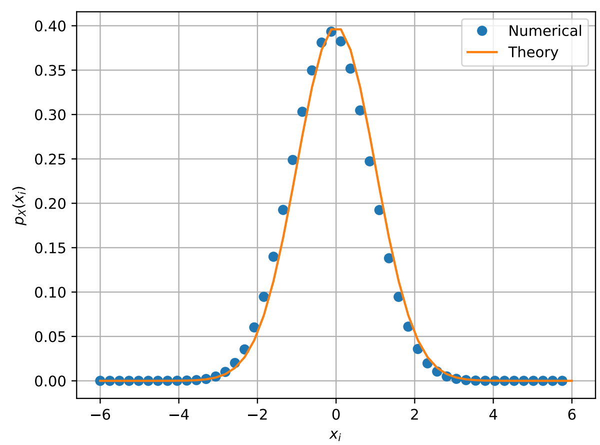


Fig. 3.3: The PDF of $X$

3.2.4 Find the mean and variance of $X$ by writing a C program.

3.2.5 Given that

$$p_X(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), -\infty < x < \infty, \tag{3.5.1}$$

repeat the above exercise theoretically. Let $U$ be a uniform random variable between 0 and 1.

3.2.6 Load the uni.dat file into python and plot the empirical CDF of $U$ using the samples in uni.dat. The CDF is defined as

$$F_U(x) = \Pr(U \le x) \tag{3.6.1}$$

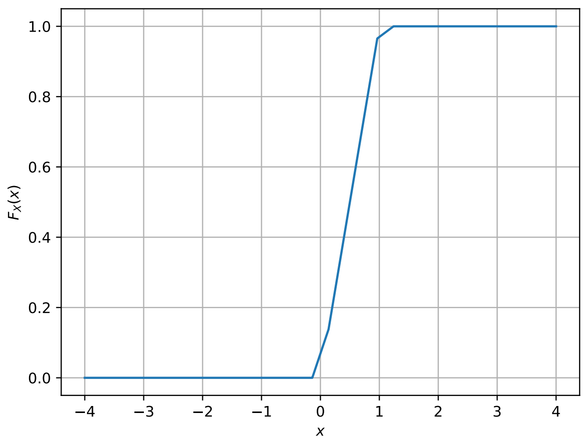


Fig. 3.6: The CDF of $U$

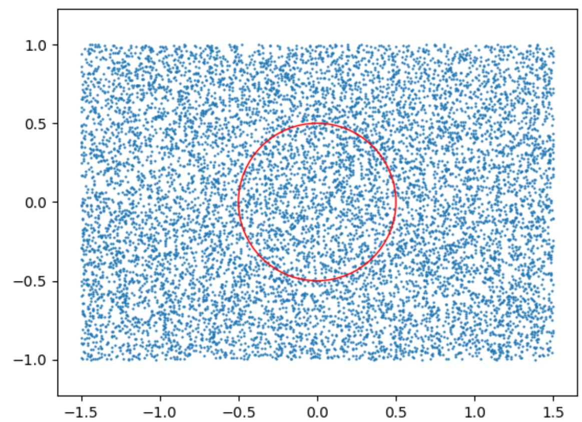

Fig. 4.1.1

**Solution:** The following code plots Fig. 3.6

```
codes/cdf/cdf_plot.py
```

3.2.7 Find a theoretical expression for $F_U(x)$.

3.2.8 The mean of $U$ is defined as

$$E[U] = \frac{1}{N}\sum_{i=1}^{N} U_i \tag{3.8.1}$$

and its variance as

$$\text{var}[U] = E[U - E[U]]^2 \tag{3.8.2}$$

Write a C program to find the mean and variance of $U$.

3.2.9 Verify your result theoretically given that

$$E\left[U^k\right] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{3.9.1}$$

## 4 Stochastic Geometry

Suppose you drop a die at random on the rectangular region shown in Fig. 4.1.1. What is the probability that it will land inside the circle with diameter 1m?

4.1. In Fig. 4.1.1, the sample size $S$ is the area of the rectangle given by

$$S = 3 \times 2 = 6m^2 \tag{4.1.1}$$

The event size is the area of the circle given by

$$E = \pi\left(\frac{1}{2}\right)^2 = \frac{\pi}{4}m^2 \tag{4.1.2}$$

The probabilty of the dice landing in the circle is

$$\Pr(E) = \frac{E}{S} = \frac{\pi}{24} \tag{4.1.3}$$

4.2. The python code is available in

```
/codes/stochastic/rect.py
```

The python code generates 10,000 points uniformly within the rectangle of dimensions 3×2 and checks for the number of points within the circle of radius 0.5. The ratio of these is close to $\frac{\pi}{24}$. Note that each time the code is run, the ratio will change, but will still be close to $\frac{\pi}{24}$.

## 5 Transformation of Variables

### 5.1 Using Definition

5.1.1. Let $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim \mathcal{N}(0,1)$. Plot the CDF and PDF of

$$V = X_1^2 + X_2^2 \tag{5.1.1.1}$$

5.1.2. If

$$F_V(x) = \begin{cases} 1 - e^{-\alpha x} & x \geq 0 \\ 0 & x < 0, \end{cases} \tag{5.1.2.1}$$

find $\alpha$.

5.1.3. Plot the CDF and PDf of

$$A = \sqrt{V} \tag{5.1.3.1}$$

5.1.4. Find an expression for $F_A(x)$ using the definition. Plot this expression and compare with the result of problem 5.1.3.

5.1.5. Find an expression for $p_A(x)$.

### 5.2 Using Jacobian

5.2.1. Evaluate the joint PDF of $X_1, X_2$, given by

$$p_{X_1,X_2}(x_1, x_2) = p_{X_1}(x_1)\, p_{X_2}(x_2) \tag{5.2.1.1}$$

5.2.2. Let

$$X_1 = \sqrt{V}\cos\theta \tag{5.2.2.1}$$

$$X_2 = \sqrt{V}\sin\theta. \tag{5.2.2.2}$$

Evaluate the Jacobian

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial v} & \frac{\partial x_2}{\partial v} \\ \frac{\partial x_1}{\partial \theta} & \frac{\partial x_2}{\partial \theta} \end{vmatrix} \tag{5.2.2.3}$$

5.2.3. Find

$$p_{V,\Theta}(v, \theta) = |J|\, p_{X_1,X_2}(x_1, x_2) \tag{5.2.3.1}$$

5.2.4. Find $p_V(v)$.
5.2.5. Find $p_\Theta(\theta)$.
5.2.6. Are $Y$ and $\Theta$ independent?
5.2.7. Find $p_A(x)$ using the Jacobian.

## 6 Conditional Probability

6.1. Plot

$$P_e = \Pr\left(\hat{X} = -1|X = 1\right) \tag{6.1.1}$$

for

$$Y = AX + N, \tag{6.1.2}$$

where $A$ is Raleigh with $E\left[A^2\right] = \gamma, N \sim \mathcal{N}(0, 1), X \in (-1, 1)$ for $0 \le \gamma \le 10$ dB.

6.2. Assuming that $N$ is a constant, find an expression for $P_e$. Call this $P_e(N)$

6.3. For a function $g$,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) p_X(x)\, dx \tag{6.3.1}$$

Find $P_e = E[P_e(N)]$.

6.4. Plot $P_e$ in problems 6.1 and 6.3 on the same graph w.r.t $\gamma$. Comment.

## 7 Two Dimensions

7.1. Let

$$\mathbf{y} = A\mathbf{x} + \mathbf{n}, \tag{7.1.1}$$

where

$$x \in (\mathbf{s}_0, \mathbf{s}_1), \mathbf{s}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{s}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.1.2}$$

$$\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}, n_1, n_2 \sim \mathcal{N}(0, 1). \tag{7.1.3}$$

7.2. Plot

$$\mathbf{y}|\mathbf{s}_0 \text{ and } \mathbf{y}|\mathbf{s}_1 \tag{7.2.1}$$

on the same graph using a scatter plot.

7.3. For the above problem, find a decision rule for detecting the symbols $\mathbf{s}_0$ and $\mathbf{s}_1$.

7.4. Plot

$$P_e = \Pr(\hat{\mathbf{x}} = \mathbf{s}_1|\mathbf{x} = \mathbf{s}_0) \tag{7.4.1}$$

with respect to the SNR from 0 to 10 dB.

7.5. Obtain an expression for $P_e$. Verify this by comparing the theory and simulation plots on the same graph.

## 8 Transform Domain

Let $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$.

8.1. Find $M_X(s) = E\left[e^{-sX}\right]$.

8.2. Let

$$N = n_1 - n_2, \quad n_1, n_2 \sim \mathcal{N}(0, 1). \tag{8.2.1}$$

Find $M_N(s)$, assuming that $n_1$ and $n_2$ are independent.

8.3. Show that $N$ is Gaussian. Find its mean and variance. Comment.

## 9 Uniform to Other

9.1. Generate samples of

$$V = -2\ln(1 - U) \tag{9.1.1}$$

and plot its CDF. Comment.

9.2. Generate the Rayleigh distribution from Uniform. Verify your result through graphical plots.