Exercice 1

$$f(x) = \begin{cases} \frac{3}{4\theta}, \text{ if } x \in [0, \theta[\\ \frac{1}{4\theta}, \text{ if } x \in [0, 2\theta[\end{cases}$$

1) Is f a density function?

* For any value of x, $f(x) \geqslant 0$.
* Let's compute the integral of f over $\mathbb{R}$.

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\theta} f(x)\,dx + \int_{\theta}^{2\theta} f(x)\,dx + \int_{2\theta}^{+\infty} f(x)\,dx.$$

We know that for $x < 0$ and $x \geqslant 2\theta$, $f(x) = 0$.

We get:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{0}^{\theta} \frac{3}{4\theta}\,dx + \int_{\theta}^{2\theta} \frac{1}{4\theta}\,dx$$

$$= \frac{3}{4\theta}[x]_0^{\theta} + \frac{1}{4\theta}[x]_{\theta}^{2\theta}$$

$$= \frac{3}{4\theta}(\theta - 0) + \frac{1}{4\theta}(2\theta - \theta)$$

$$= \frac{3}{4} + \frac{1}{4}$$

$$\boxed{\int_{-\infty}^{+\infty} f(x)\,dx = 1}$$

$\forall x$, $f(x) \geqslant 0$ and $\int_{-\infty}^{+\infty} f(x)\,dx = 1$, we conclude that f is a density function.

2) <u>Determine the distribution of function F_X.</u>

We need to compute $\int_{-\infty}^{x} f(x)\,dx$.

* For $x < 0$, $f(x) = 0 \Rightarrow F_X(x) = 0$.

* For $x \in [0, \theta[$, $F_X(x) = \int_{-\infty}^{x} f(x)\,dx$

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{x} f(x)\,dx$$

$$= 0 + \int_{0}^{x} \frac{3}{4\theta}\,dx$$

$$= \frac{3}{4\theta}[x]_0^x$$

$$\forall x \in [0, \theta[, F_X(x) = \frac{3}{4\theta} x$$

* For $x \in [\theta, 2\theta[$, $F_X(x) = \int_{-\infty}^{x} f(x)\,dx$

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\theta} f(x)\,dx + \int_{\theta}^{x} f(x)\,dx$$

$$= 0 + \frac{3}{4} + \int_{\theta}^{x} \frac{1}{4\theta}\,dx$$

$$= \frac{3}{4} + \frac{1}{4\theta}[x]_{\theta}^{x}$$

$$= \frac{3}{4} + \frac{x-\theta}{4\theta}$$

$$\forall x \in [\theta, 2\theta[, \quad F_X(x) = \frac{x+2\theta}{4\theta}$$

* For $x \geq 2\theta$, $F_X(x) = \int_{-\infty}^{0} f(x)\,dx + \int_{\theta}^{2\theta} \frac{1}{4\theta}\,dx + \int_{-\infty}^{0} f(x)\,dx + \int_{2\theta}^{x} f(x)\,dx$

$$= \frac{3}{4} + \frac{1}{4\theta}[x]_{\theta}^{2\theta} + 0 + 0$$

$$= \frac{3}{4} + \frac{2\theta-\theta}{4\theta}$$

$$= \frac{3}{4} + \frac{1}{4}$$

$$\forall x \geq 2\theta, \quad F_X(x) = 1$$

the distribution function can be written:

$$F_X(x) = \begin{cases} 0 \ , \text{ if } x < 0 \\ \frac{3}{4\theta} x \ , \text{ if } x \in [0, \theta[\\ \frac{x+2\theta}{4\theta} \ , \text{ if } x \in [\theta, 2\theta[\\ 1 \ , \text{ if } x \geq 2\theta \end{cases}$$

3) Let's compute the expectation and the variance.

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

$$= \int_0^{\theta} \frac{3}{4\theta} x \, dx + \int_{\theta}^{2\theta} \frac{1}{4\theta} x \, dx$$

$$= \frac{3}{4\theta} \left[\frac{x^2}{2} \right]_0^{\theta} + \frac{1}{4\theta} \left[\frac{x^2}{2} \right]_{\theta}^{2\theta}$$

$$= \frac{3}{4\theta} \left(\frac{\theta^2}{2} - 0 \right) + \frac{1}{4\theta} \left[\frac{4\theta^2}{2} - \frac{\theta^2}{2} \right]$$

$$E(X) = \frac{3}{8} \theta + \frac{3}{8} \theta$$

$$E(X) = \frac{6}{8} \theta$$

$$\boxed{E(X) = \frac{3}{4} \theta}$$

Let's compute the variance $V(X)$.

$$V(X) = E(X^2) - [E(X)]^2$$

Let's compute $E(X^2)$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx$$

$$= \int_0^{\theta} \frac{3}{4\theta} x^2\,dx + \int_{\theta}^{2\theta} \frac{1}{4\theta} x^2\,dx$$

$$= \frac{3}{4\theta}\left[\frac{x^3}{3}\right]_0^{\theta} + \frac{1}{4\theta}\left[\frac{x^3}{3}\right]_{\theta}^{2\theta}$$

$$= \frac{3}{4\theta}\left(\frac{\theta^3}{3} - 0\right) + \frac{1}{4\theta}\left(\frac{8\theta^3}{3} - \frac{\theta^3}{3}\right)$$

$$= \frac{1}{4}\theta^2 + \frac{7}{12}\theta^2$$

$$= \frac{3\theta^2 + 7\theta^2}{12}$$

$$= \frac{10}{12}\theta^2$$

$$E(X^2) = \frac{5}{6}\theta^2$$

We have:

$$V(X) = \frac{5}{6}\theta^2 - \left(\frac{3}{4}\theta\right)^2$$

$$= \frac{5}{6}\theta^2 - \frac{9}{16}\theta^2$$

$$\boxed{V(X) = \frac{13}{48}\theta^2}$$

4°) Determine an estimator for θ

by applying method of moments, we know that an estimator $\tilde{\theta}_n$ for θ is solution of:

$$\frac{1}{n}\sum_{i=1}^{n} X_i = h(\tilde{\theta}_n) \quad \text{with } h(t) = \frac{3}{4}t$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{3}{4}\tilde{\theta}_n \Leftrightarrow \tilde{\theta}_n = \frac{4}{3}\cdot\frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\Leftrightarrow \boxed{\tilde{\theta}_n = \frac{4}{3}\overline{X_n}}$$

5°) Is it a biased one?

$$E(\tilde{\theta}_n) = E\left(\frac{4}{3}\cdot\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \frac{4}{3}\cdot\frac{1}{n}\sum_{i=1}^{n} E(X_i) \quad \text{with } E(X_i) = \frac{3}{4}\theta$$

$$= \frac{4}{3}\cdot\frac{1}{n}\times n\times E(X_i)$$

$$= \frac{4}{3}\times\frac{3}{4}\theta$$

$$\boxed{E(\tilde{\theta}_n) = \theta}$$

$\tilde{\theta}_n$ is not biased because $E(\tilde{\theta}_n) = \theta$

6°) Provide a confidence interval for θ

We have $\hat{\theta}_n = \frac{4}{3}\bar{X}_n$ is an unbiased estimator of θ

Let's compute $V(\hat{\theta}_n)$.

$$V(\hat{\theta}_n) = V\left(\frac{4}{3}\bar{X}_n\right)$$

$$= \frac{16}{9} V(\bar{X}_n)$$

$$= \frac{16}{9} V\left(\frac{1}{n}\sum x_i\right)$$

$$= \frac{16}{9} \times \frac{1}{n^2} \sum_{i=1}^{n} V(X_i)$$

$$= \frac{16}{9} \times \frac{1}{n^2} \times n\, V(X)$$

$$= \frac{16}{9n} V(X)$$ with $V(X) = \hat{\sigma}^2_{n-1}$ estimated thanks to a dataset.

$$\boxed{V(\hat{\theta}_n) = \frac{16}{9n}\hat{\sigma}^2_{n-1}}$$

by applying the central limit theorem (CLT), we have

$$\sqrt{n}\,\frac{\hat{\theta}_n - \theta}{\sigma_\theta} \rightsquigarrow \mathcal{N}(0,1)$$, with σ_θ the true standard deviation of θ.

$$\Leftrightarrow \quad \hat{\theta}_n \rightsquigarrow \mathcal{N}\left(\theta, \frac{\sigma_\theta}{n}\right)$$

$$\Leftrightarrow \quad \frac{4}{3}\bar{X}_n \rightsquigarrow \mathcal{N}\left(\theta, \frac{\sigma_\theta}{n}\right)$$

we have to find r_n and m_n such that:

$$P\left(-r_n \leq \frac{4}{3}\bar{X}_n - \theta \leq m_n\right) = 1-\alpha$$

$$P\left(-\frac{\sqrt{n}\,.\,r_n}{\sigma_\theta} \leq \sqrt{n}\,.\,\frac{\frac{4}{3}\bar{X}_n - \theta}{\sigma_\theta} \leq \frac{\sqrt{n}\,.\,m_n}{\sigma_\theta}\right) = 1-\alpha.$$

σ is unknown, we are going to replace σ_θ by an estimator equal to: $\sqrt{V(\hat{\theta}_n)} = \sqrt{\frac{16}{9n}\hat{\sigma}^2_{n-1}} = \frac{4}{3\sqrt{n}}\hat{\sigma}_{n-1}$.

so $\frac{\sigma_\theta}{\sqrt{n}}$ is replaced by $\frac{4}{3\sqrt{n}}\hat{\sigma}_{n-1}$

we have:

$$P\left(-\frac{\sqrt{n}\,.\,r_n}{\frac{4}{3}\hat{\sigma}_{n-1}} \leq \sqrt{n}\,.\,\frac{\frac{4}{3}\bar{X}_n - \theta}{\frac{4}{3}\hat{\sigma}_{n-1}} \leq \frac{\sqrt{n}\;m_n}{\frac{4}{3}\hat{\sigma}_{n-1}}\right) = 1-\alpha$$

we can rewrite:

$$\sqrt{n}\,.\,\frac{\frac{4}{3}\bar{X}_n - \theta}{\frac{4}{3}\hat{\sigma}_{n-1}} = \sqrt{n}\,.\,\frac{\frac{4}{3}\bar{X}_n - \theta}{\sigma_\theta} \Big/ \sqrt{\frac{(n-1)\frac{16}{9}\hat{\sigma}^2_{n-1}}{(n-1)\,\sigma_\theta^2}}$$

the CLT tells us that $\sqrt{n}\,.\,\frac{\frac{4}{3}\bar{X}_n - \theta}{\sigma_\theta} \sim \mathcal{N}(0,1)$.

by applying slutsky's lemma we have:

$$\frac{16}{9}\hat{\sigma}_{n-1} \xrightarrow[P]{} \sigma_\theta^2$$

so to end the construction of the confidence interval, we say:

$$\sqrt{n}\,.\,\frac{\frac{4}{3}\bar{X}_n - \theta}{\frac{4}{3}\hat{\sigma}_{n-1}} \sim \mathcal{N}(0,1).$$

We make an approximation to determine η_n and m_n !

We choose $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$, we then have $\eta_n = m_n$ such that

$$P\left(Y \leq -\frac{\eta_n \sqrt{n}}{\frac{4}{3}\hat{\sigma}_{n-1}}\right) = \frac{\alpha}{2}, \text{ with } Y = \sqrt{n}.\frac{\frac{4}{3}\bar{X}_n - \theta}{\frac{4}{3}\hat{\sigma}_{n-1}}$$

$$\Leftrightarrow -\frac{\sqrt{n}.\eta_n}{\frac{4}{3}\hat{\sigma}_{n-1}} = -z_{1-\frac{\alpha}{2}}$$

We get:

$$P\left(-z_{1-\frac{\alpha}{2}} \leq \sqrt{n}.\frac{\frac{4}{3}\bar{X}_n - \theta}{\frac{4}{3}\hat{\sigma}_{n-1}} \leq z_{1-\frac{\alpha}{2}}\right) \xrightarrow[n\to+\infty]{} 1-\alpha$$

$$I = P\left(-\frac{4\hat{\sigma}_{n-1}}{3\sqrt{n}} z_{1-\frac{\alpha}{2}} \leq \frac{4}{3}\bar{X}_n - \theta \leq \frac{4\hat{\sigma}_{n-1}}{3\sqrt{n}}.z_{1-\frac{\alpha}{2}}\right)$$

$$= P\left(-\frac{4}{3}\bar{X}_n - \frac{4}{3}.\frac{\hat{\sigma}_{n-1}}{\sqrt{n}}.z_{1-\frac{\alpha}{2}} \leq -\theta \leq -\frac{4}{3}\bar{X}_n + \frac{4}{3}\frac{\hat{\sigma}_{n-1}}{\sqrt{n}}.z_{1-\frac{\alpha}{2}}\right)$$

$$= P\left(\frac{4}{3}\bar{X}_n - \frac{4}{3}.\frac{\hat{\sigma}_{n-1}}{\sqrt{n}}.z_{1-\frac{\alpha}{2}} \leq \theta \leq \frac{4}{3}\bar{X}_n + \frac{4}{3}.\frac{\hat{\sigma}_{n-1}}{\sqrt{n}}.z_{1-\frac{\alpha}{2}}\right)$$

An asymptotic confidence interval for θ with level $100(1-\alpha)\%$ is

$$\left[\frac{4}{3}\bar{X}_n - \frac{4}{3}\frac{\hat{\sigma}_{n-1}}{\sqrt{n}}.z_{1-\frac{\alpha}{2}} ; \frac{4}{3}\bar{X}_n + \frac{4}{3}.\frac{\hat{\sigma}_{n-1}}{\sqrt{n}}.z_{1-\frac{\alpha}{2}}\right]$$

7. Let's determine F inverse

for x $\in [0, \theta[, y = \frac{3 \cdot \theta \cdot x}{4} \Leftrightarrow x = \frac{4 \cdot \theta \cdot y}{3}$,
with y $\in [0, \frac{3}{4}[$
for x $\in [\theta, 2 \cdot \theta[, y = \frac{x + 2 \cdot \theta}{4 \cdot \theta} \Leftrightarrow x = 2 \cdot \theta \cdot (2 \cdot y - 1)$
with y $\in [\frac{3}{4}, 1[$

EXAM FSML PART 2

Jean-Luc BOA THIEMELE

2024-01-14

7. Produce simulated observation of X

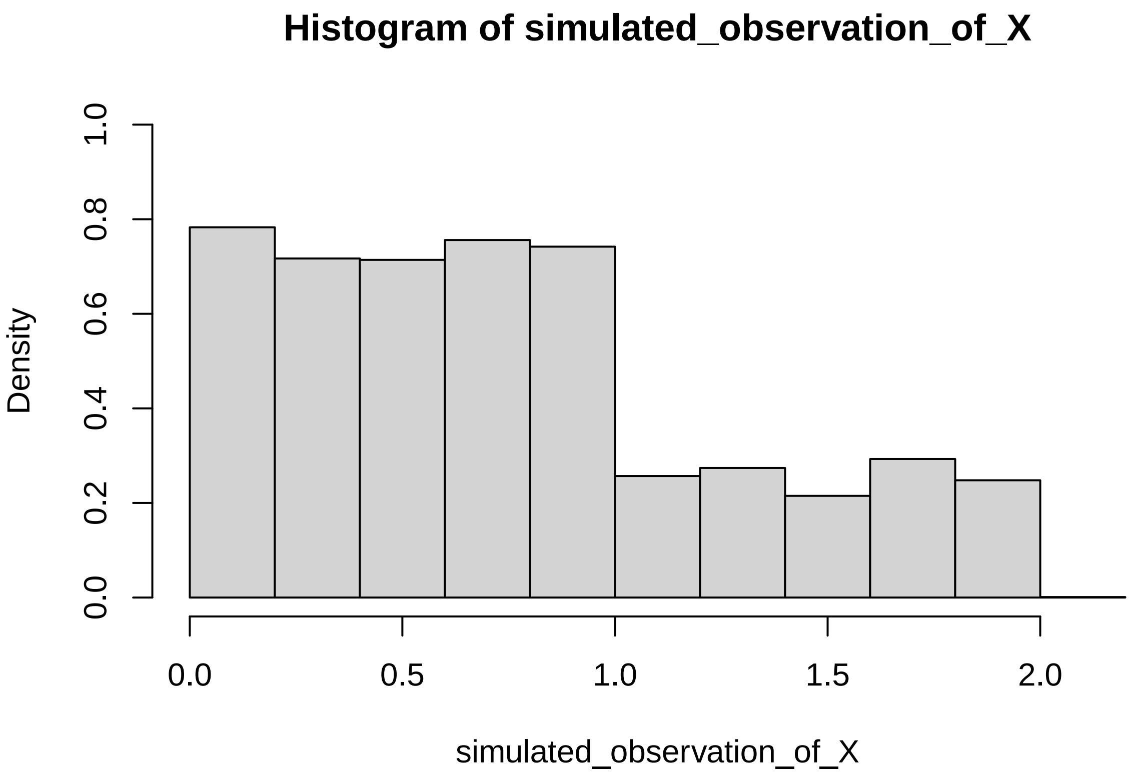
```r
# Step 1 : Let's define F_inverse
F_inverse <- function(x, theta) {
  if (x<0){
    result <- 0
  }
  else if (0 <= x && x < (3/4)) {
    result <- (4 * theta)*x / 3
  } else if (x >= 3/4 && x<1){
    result <- 2 * theta * (2 * x - 1)
  }
  else{
    result <- 1
  }
  return(result)
}
```

```r
#Step 2 : Generate 5000 uniform random number and compute the estimate of theta

# Set the seed for reproducibility
set.seed(42)

# Number of simulations
n <- 5000

#Generate n uniform random numbers
uniform_numbers <- runif(n,min = 0, max = 1)

#Setting working directory
setwd("/Users/jlbt/Downloads")

#Load data in order to compute theta_hat and adding a column.
X = read.table("data.txt", header=FALSE)
col_names <- c("Observation_of_X")  # Replace with your column names
colnames(X) <- col_names

#Computing the estimate of theta

theta_hat =  (4/3)*mean(X[,1])

cat('The estimate of theta (theta_hat) is equal to:', theta_hat, '\n')
```

```
## The estimate of theta (theta_hat) is equal to: 1.000419
```

```r
# Step 3 : Computing the inverse of every uniform_numbers thanks to F_inverse
simulated_observation_of_X <- sapply(uniform_numbers, F_inverse, theta = theta_hat)
```

```r
#printng the result
hist(simulated_observation_of_X, freq = FALSE, xlim=c(0,2.2), ylim = c(0,1))
```



8. Computing the probability for X to be between 0.3 and 0.6 thanks to the simulations

```r
# Count the number of element between 0.3 and 0.6 in the inverse_resuls
interval_count <- sum(simulated_observation_of_X >= 0.3 & simulated_observation_of_X <= 0.6)

# We compute the probability
probability_estimate <- interval_count / n
cat('The probability that X takes values between 0.3 and 0.6 is :', probability_estimate, '\n')
```

```
## The probability that X takes values between 0.3 and 0.6 is : 0.2144
```

9. Give an estimation of the biais and of the mean quadratic error.

```r
#An estimation of the biais when fixing the tetha = tetha_hat is:
biais =  mean((4/3)*mean(simulated_observation_of_X)) - theta_hat

cat('An estimation of the biais is :', biais, '\n')
```

```
## An estimation of the biais is : 0.009110019
```

```r
# The estimate of the MSE is given by :
MSE = mean(((4/3)*simulated_observation_of_X-theta_hat)^2)

cat('An estimation of the mean quadratic error is :', MSE, '\n')
```

```
## An estimation of the mean quadratic error is : 0.4913655
```

Exercice 2:

Let $X \sim N(1,4)$

1. $P(X \in [0;1,5])$:

To compute this probability we need to read the table of standard gaussian random variable.

We say that X is a standard gaussian if $X \sim N(0,1)$

This is not the case, so we are going to transform it:

We have $X \sim N(1,4)$

We consider $Y = \frac{X-1}{\sqrt{4}}$

$= \frac{X-1}{2}$

So we have:

$$P(0 \leq X \leq 1,5) = P(0-1 \leq X-1 \leq 1,5-1)$$

$$= P\left(-0,5 \leq \frac{X-1}{2} \leq 0,25\right)$$

$$= P(-0,5 \leq Y \leq 0,25)$$

$$= F_Y(0,25) - F_Y(-0,5)$$

$$= P(Y \leq 0,25) - P(Y \geq 0,5)$$

$$= P(Y \leq 0,25) + P(Y \leq 0,5) - 1$$

$$= 0,5987 + 0,6915 - 1$$

$$\boxed{P(0 \leq X \leq 1,5) = 0,2902}$$

2°) $P(X \geq 0{,}7)$

$$P(X \geq 0{,}7) = 1 - P(X \leq 0{,}7)$$

$$= 1 - P\left(\frac{X-1}{2} \leq \frac{0{,}7-1}{2}\right)$$

$$= 1 - P(Y_2 \leq -0{,}15) \quad \text{with } Y_2 = \frac{X-1}{2}$$

$$= 1 - P(Y_2 \geq 0{,}15) \quad \text{as } P(X \leq -t) = P(X \geq t)$$

$$= 1 - 1 + P(Y_2 \leq 0{,}15)$$

$$= P(Y_2 \leq 0{,}15)$$

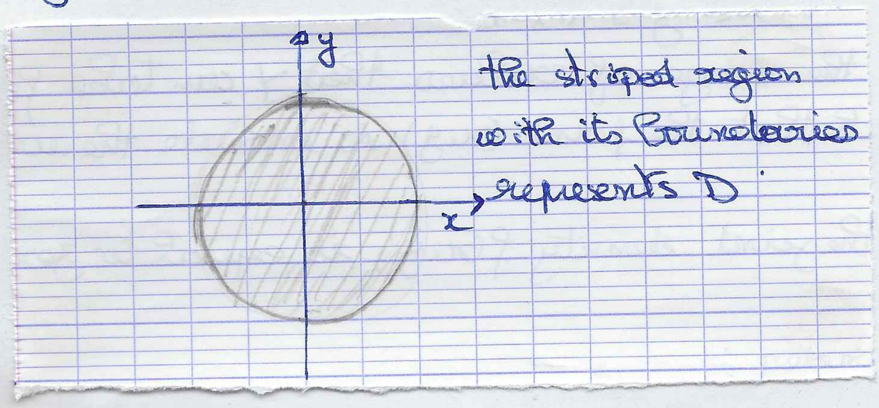
$$\boxed{P(X \geq 0{,}7) = 0{,}5596}$$

Exercice 3

1. give the density function

$D = \{(x,y) \in \mathbb{R}^2,\ x^2+y^2 \leq 1\}$

D defines the set of points (x,y) in $\mathbb{R}^2$ such that $x^2+y^2 \leq 1$.
In other words, D represents the unit disk centered at the origin with a radius of 1



The area of D is given by πr^2 with $r=1$, we have
$A = \pi \times 1^2 = \pi$.

for any point (x,y) within D, the probability is given by $\frac{1}{\pi}$
so we can deduce that:

$$f(x,y) = \begin{cases} \frac{1}{\pi}, & \text{for } x^2+y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

2°) Determine the density of the margin.

* Marginal density of X.

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy.$$

We have :

$$x^2 + y^2 \leq 1$$

$$y^2 \leq 1 - x^2$$

$$-\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}$$

This represent the range of values that y can take for a given x such that the point (x,y) is inside the unit disk D.

We then consider:

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy.$$

$$= \frac{1}{\pi}\left[y\right]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}$$

$$= \frac{1}{\pi}\left(\sqrt{1-x^2} + \sqrt{1-x^2}\right).$$

$$f_X(x) = \frac{2}{\pi}\sqrt{1-x^2}.$$

$$f_X(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2}, \text{ if } x \in [-1,1] \\ 0, \text{ otherwise.} \end{cases}$$

* Marginal density of y.

by following the same steps as for x, we find that:

$$f_Y(y) = \begin{cases} \frac{2}{\pi}\sqrt{1-y^2}, & \text{if } y \in [-1,1] \\ 0, & \text{otherwise.} \end{cases}$$

3. Are X and Y independant?

X and Y are independant only if $f(x,y) = f_X(x) \times f_Y(y)$

• We have $f(x,y) = \frac{1}{\pi}$, if $x^2+y^2 \leq 1$.

• $f_X(x) \times f_Y(y) = \frac{2}{\pi}\sqrt{1-x^2} \times \frac{2}{\pi}\sqrt{1-y^2}$

$$= \frac{4}{\pi^2}\sqrt{1-x^2} \times \sqrt{1-y^2}$$

So $f_{X,Y}(x,y) \neq f_X(x) . f_Y(y)$.

We conclude that X and Y are not independant

4.) Compute the covariance.

$cov(X,Y) = E(XY) - E(X)E(Y)$

X and Y have the same distribution as they have the same density function. So $E(X) = E(Y)$

Let's compute $E(X)$

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

f_X is only define if $x \in [-1, 1]$; we have

$$E(X) = \int_{-1}^{1} x \times \frac{2}{\pi} \sqrt{1-x^2}\, dx$$

$$= \frac{1}{\pi} \int_{-1}^{1} 2x \sqrt{1-x^2}\, dx$$

$$= -\frac{1}{\pi} \int_{-1}^{1} -2x \,(1-x^2)^{1/2}\, dx$$

$$= -\frac{1}{\pi} \left[\frac{2}{3} (1-x^2)^{3/2} \right]_{-1}^{1}$$

$$= -\frac{1}{\pi} (0 - 0)$$

$\underline{E(X) = 0}$

So we have $E(X) = E(Y) = 0$.

Let's compute $E(XY)$.

$$E(XY) = \int_{-1}^{1} \int_{-1}^{1} xy \times \frac{1}{\pi}\, dx . dy$$

$$= \frac{1}{\pi} \int_{-1}^{1} x\, dx \times \int_{-1}^{1} y\, dy$$

$$= \frac{1}{\pi} \left(\left[\frac{x^2}{2} \right]_{-1}^{1} \right) \times \left(\left[\frac{y^2}{2} \right]_{-1}^{1} \right)$$

$$= \frac{1}{\pi} \left(\frac{1}{2} - \frac{1}{2} \right) \times \left(\frac{1}{2} - \frac{1}{2} \right)$$

$E(XY) = 0$

So, $\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y)$

$$= 0 - 0 \times 0$$

$$\boxed{\mathrm{cov}(X,Y) = 0}$$