

Foundations of Statistical Analysis and Machine Learning - Part 2

Claire Peyran

2025-08-31

Exercise A:

Let X and Y be two independent random variables with exponential distribution of parameters λ and μ .

1. Compute the joint density $f_{X,Y}(x, y)$?

A random variable X has an exponential distribution with parameter λ when its probability density function is:

$$X \sim \text{Exponential}(\lambda)$$
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \quad \text{with } \lambda \geq 0$$

and similarly, for Y with parameter μ :

$$Y \sim \text{Exponential}(\mu)$$
$$f_Y(y) = \begin{cases} \mu e^{-\mu y}, & \text{if } y \geq 0 \\ 0, & \text{if } y < 0 \end{cases} \quad \text{with } \mu \geq 0$$

These functions are the marginal density functions (PDF) and represent the probability of each value of a random variable. Thus it is always positive and the sum of each value is 1 (i.e the integrate = 1).

And if we focus on (X, Y) :

$F_{X,Y}(x, y)$ is the joint density function, thus we have: $F_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that:

$$\forall (x, y) \in \mathbb{R}^2, \quad F_{X,Y}(x, y) \geq 0$$

and

$$\iint_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = 1$$

Then, because X and Y are **independent**, and **continuous** (exponential distribution), the joint PDF is the product of the marginal PDFs:

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

$$\Rightarrow f_{X,Y}(x, y) = \lambda e^{-\lambda x} \cdot \mu e^{-\mu y}$$

for $x \geq 0$ and $y \geq 0$

$$\Leftrightarrow f_{X,Y}(x, y) = \lambda \mu e^{-\lambda x - \mu y}, \quad x \geq 0, y \geq 0.$$

And for other values, at least one marginal PDF is zero, thus the joint PDF is zero. Thus

$$f_{X,Y}(x,y) = \begin{cases} \lambda\mu e^{-\lambda x - \mu y}, & \text{if } x \geq 0, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

We can check:

Positivity: for all x and y the function is always positive because λ and μ are ≥ 0 and exponential are always positive.

By definition, the PDF should **integrate to 1** over the whole space.

$$\int_0^\infty \int_0^\infty f_{X,Y}(x,y) dx dy = \int_0^\infty \int_0^\infty \lambda\mu e^{-\lambda x - \mu y} dx dy$$

Integrate with regards to x first (treat y as a constant):

$$\begin{aligned} \int_0^\infty \lambda e^{-\lambda x} dx &= [-e^{-\lambda x}]_0^\infty \\ &= 0 - (-1) \\ &= 1 \end{aligned}$$

Then we integrate with regards to y :

$$\begin{aligned} \int_0^\infty \mu e^{-\mu y} dy &= [-e^{-\mu y}]_0^\infty \\ &= 0 - (-1) \\ &= 1 \end{aligned}$$

Finally we have:

$$\int_0^\infty \int_0^\infty f_{X,Y}(x,y) dx dy = 1 \cdot 1 = 1$$

This confirms the joint PDF is valid.

2. Compute $P(X \leq Y)$

$P(X \leq Y)$ represents the probability that the value of X is smaller or equal to the value of Y .

So using the joint PDF from part 1,

$$f_{X,Y} = \lambda\mu e^{-\lambda x - \mu y}, \quad \text{with } x \geq 0, y \geq 0$$

We write the probability that $X \leq Y$ as a double integral over the region $\{(x,y) : 0 \leq x \leq y < \infty\}$ leading to:

$$\begin{aligned} P(X \leq Y) &= \int_{y=0}^\infty \int_{x=0}^y f_{X,Y}(x,y) dx dy \\ &= \int_{y=0}^\infty \int_{x=0}^y \lambda\mu e^{-\lambda x - \mu y} dx dy \quad (\text{from part 1.}) \end{aligned}$$

x ranges from 0 to y because of $x \leq y$.

So let's start with the **inner integrate**:

We integrate with respect to x , so y is treated as a constant.

$$\begin{aligned}
 \int_{x=0}^y \lambda \mu e^{-\lambda x - \mu y} dx &= \lambda \mu \int_0^y e^{-\lambda x - \mu y} dx \\
 &= \lambda \mu \int_0^y e^{-\lambda x} \cdot e^{-\mu y} dx \quad (\text{split the exponential}) \\
 &= \lambda \mu e^{-\mu y} \int_0^y e^{-\lambda x} dx \quad (\text{factor out } e^{-\mu y} \text{ since it does not depend on } x) \\
 &= \lambda \mu e^{-\mu y} \left[\frac{-1}{\lambda} e^{-\lambda x} \right]_0^y \\
 &= \lambda \mu e^{-\mu y} \left(\frac{-1}{\lambda} e^{-\lambda y} - \frac{-1}{\lambda} e^0 \right) \\
 &= \lambda \mu e^{-\mu y} \left(\frac{-1}{\lambda} e^{-\lambda y} + \frac{1}{\lambda} \right) \\
 &= \lambda \mu e^{-\mu y} \cdot \frac{-1}{\lambda} e^{-\lambda y} + \lambda \mu e^{-\mu y} \cdot \frac{1}{\lambda} \\
 &= \mu e^{-\mu y} - \mu e^{-\mu y - \lambda y} \\
 &= \mu e^{-\mu y} (1 - e^{-\lambda y})
 \end{aligned}$$

So now we have :

$$P(X \leq Y) = \int_0^\infty \mu e^{-\mu y} (1 - e^{-\lambda y}) dy$$

Outer integral:

To do the integration, we can separate it in two parts:

$$P(X \leq Y) = \int_0^\infty \mu e^{-\mu y} dy - \int_0^\infty \mu e^{-\mu y} \cdot e^{-\lambda y} dy \Rightarrow P(X \leq Y) = \int_0^\infty \mu e^{-\mu y} dy - \int_0^\infty \mu e^{-(\mu+\lambda)y} dy$$

Now we compute each part with limits...

First part:

$$\begin{aligned}
\int_0^\infty \mu e^{-\mu y} dy &= \mu \int_0^\infty e^{-\mu y} dy \\
&= \mu \left[\frac{-1}{\mu} e^{-\mu y} \right]_0^\infty \\
&= \mu \left(\lim_{y \rightarrow \infty} \left(\frac{-1}{\mu} e^{-\mu y} \right) - \frac{-1}{\mu} e^0 \right) \\
&= \mu \left(\frac{-1}{\mu} \cdot 0 - \left(-\frac{1}{\mu} e^0 \right) \right) \quad \text{since when } y \rightarrow \infty, \text{ we have } e^{-\mu y} \rightarrow 0 \\
&= \mu \cdot \left(\frac{1}{\mu} \right) \\
&= 1
\end{aligned}$$

Second part:

$$\begin{aligned}
\int_0^\infty \mu e^{-(\mu+\lambda)y} dy &= \mu \int_0^\infty e^{-(\mu+\lambda)y} dy \\
&= \mu \left[\frac{-1}{\mu+\lambda} e^{-(\mu+\lambda)y} \right]_0^\infty \\
&= \mu \left(\frac{-1}{\mu+\lambda} \cdot 0 + \frac{1}{\mu+\lambda} \cdot e^0 \right) \quad \text{since when } y \rightarrow \infty, \text{ we have } e^{-(\mu+\lambda)y} \rightarrow 0 \\
&= \frac{\mu}{\mu+\lambda}
\end{aligned}$$

... and then do the subtraction:

$$\begin{aligned}
P(X \leq Y) &= \int_0^\infty \mu e^{-\mu y} dy - \int_0^\infty \mu e^{-(\mu+\lambda)y} dy \\
&= 1 - \frac{\mu}{\mu+\lambda} \\
&= \frac{(\mu+\lambda) - \mu}{\mu+\lambda} \\
&= \frac{\lambda}{\mu+\lambda}
\end{aligned}$$

So the final result is :

$$\boxed{P(X \leq Y) = \frac{\lambda}{\mu+\lambda}}$$

3. Let $Z = \text{Max}(X, Y)$. Compute the density of Z .

We need to find the **PDF** of F . We know that $X \sim \text{Exponential}(\lambda)$ and $Y \sim \text{Exponential}(\mu)$, they are independent and $\lambda \geq 0$ and $\mu \geq 0$.

Since exponential is non negative, $X \geq 0$ and $Y \geq 0$, so $Z = \text{Max}(X, Y) \geq 0$.

For $z < 0$, $f_Z(z) = 0$.

So we only need to find values for $z \geq 0$.

Because Z is defined by a $\text{Max}()$ of the two variables X and Y is not a linear transformation. We need to first find the **distribution function** of Z , then **differentiate** to get the **density**.

- **Step 1:** find the distribution of Z : F_Z
- **Step 2:** deduce the density function f_Z by differentiation of F_Z , since it follows the rule $f_Z = F'_Z$.

Step 1: Find the distribution

So for $z \geq 0$:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P(\text{Max}(X, Y) \leq z) \\ &= P(X \leq z, Y \leq z) \\ &= P(X \leq z) \cdot P(Y \leq z) \quad \text{since } X \text{ and } Y \text{ are independent} \end{aligned}$$

The probabilities that $X \leq z$ and $Y \leq z$ are the integrates (i.e $F_X(z)$ and $F_Y(z)$), thus:

$$\begin{aligned} P(X \leq z) \cdot P(Y \leq z) &= \int_0^z f_X(z) dx \cdot \int_0^z f_Y(z) dy \\ &= \int_0^z \lambda e^{-\lambda x} dx \cdot \int_0^z \mu e^{-\mu y} dy \end{aligned}$$

Let's compute them separately:

$$\begin{aligned} \int_0^z \lambda e^{-\lambda x} dx &= \lambda \int_0^z e^{-\lambda x} dx & \int_0^z \mu e^{-\mu y} dy &= \mu \int_0^z e^{-\mu y} dy \\ &= \lambda \left[\frac{-1}{\lambda} e^{-\lambda x} \right]_0^z & &= \mu \left[\frac{-1}{\mu} e^{-\mu y} \right]_0^z \\ &= \lambda \left(\frac{-1}{\lambda} e^{-\lambda z} - \left(\frac{-1}{\lambda} e^0 \right) \right) & &= \mu \left(\frac{-1}{\mu} e^{-\mu z} - \left(\frac{-1}{\mu} e^0 \right) \right) \\ &= \lambda \left(\frac{-1}{\lambda} e^{-\lambda z} + \frac{1}{\lambda} \right) & &= \mu \left(\frac{-1}{\mu} e^{-\mu z} + \frac{1}{\mu} \right) \\ &= 1 - e^{-\lambda z} & &= 1 - e^{-\mu z} \end{aligned}$$

So if we go back to $F_Z(z)$:

$$\begin{aligned} P(X \leq z) \cdot P(Y \leq z) &= \int_0^z \lambda e^{-\lambda x} dx \cdot \int_0^z \mu e^{-\mu y} dy \\ &= (1 - e^{-\lambda z}) \cdot (1 - e^{-\mu z}) \end{aligned}$$

So, at this end of step 1 we have:

$$F_Z(z) = (1 - e^{-\lambda z}) \cdot (1 - e^{-\mu z}) \quad z \geq 0$$

Step 2:

We need to differentiate F_Z , because $f_Z(z) = F'_Z(z)$.

We use the **product rule**:

If $F_Z(z) = u(z) \cdot v(z)$, then:

$$F'_Z(z) = u'(z) \cdot v(z) + u(z) \cdot v'(z)$$

So we define:

$$\begin{aligned} u(z) &= 1 - e^{-\lambda z} & v(z) &= 1 - e^{-\mu z} \\ u'(z) &= 0 - (-\lambda) \cdot e^{-\lambda z} & v'(z) &= 0 - (-\mu) \cdot e^{-\mu z} \\ u'(z) &= \lambda e^{-\lambda z} & v'(z) &= \mu e^{-\mu z} \end{aligned}$$

So:

$$\begin{aligned} F'_Z(z) &= u'(z) \cdot v(z) + u(z) \cdot v'(z) \\ &= (\lambda e^{-\lambda z}) \cdot (1 - e^{-\mu z}) + (1 - e^{-\lambda z}) \cdot (\mu e^{-\mu z}) \\ &= \lambda e^{-\lambda z} - \lambda e^{-\lambda z} \cdot e^{-\mu z} + \mu e^{-\mu z} - \mu e^{-\mu z} \cdot e^{-\lambda z} \\ &= \lambda e^{-\lambda z} - \lambda e^{-z(\lambda+\mu)} + \mu e^{-\mu z} - \mu e^{-z(\mu+\lambda)} \\ &= \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-z(\lambda+\mu)} \end{aligned}$$

So the density function of Z is :

$$f_Z(z) = \begin{cases} \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-z(\lambda+\mu)} & \text{if } z \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Exercise B:

Let θ be a positive number.

Let f be the following function:

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

1. Prove that f is a density function.

For f to be a density function, it needs to complete two conditions: **positivity** and it should **integrate to 1**.

Positivity:

For $x \in [0, 2\theta]$: $X \geq 0$ and $\theta \geq 0$, so $\frac{1}{4\theta^2} \geq 0$, so $\frac{1}{4\theta^2}x \geq 0$ as well.

For $x \in [2\theta, 3\theta]$, the value is $\frac{1}{2\theta}$ so it is positive, for the same reasons as above.

Everywhere else, $f(x) = 0$.

Thus, the **positivity condition is respected**.

Integrates to 1:

We integrate only between $[0, 3\theta]$, as everywhere else $f(x) = 0$.

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{2\theta} \frac{1}{4\theta^2}x dx + \int_{2\theta}^{3\theta} \frac{1}{2\theta}dx$$

We do each computation separately:

First part:

$$\begin{aligned} \int_0^{2\theta} \frac{1}{4\theta^2}x dx &= \frac{1}{4\theta^2} \int_0^{2\theta} x dx \\ &= \frac{1}{4\theta^2} \left[\frac{1}{2}x^2 \right]_0^{2\theta} \\ &= \frac{1}{4\theta^2} \left(\frac{1}{2} \cdot (2\theta)^2 - \frac{1}{2} \cdot 0^2 \right) \\ &= \frac{1}{4\theta^2} \cdot \frac{4\theta^2}{2} \\ &= \frac{1}{2} \end{aligned}$$

Second part:

$$\begin{aligned}
\int_{2\theta}^{3\theta} \frac{1}{2\theta} dx &= \frac{1}{2\theta} \int_{2\theta}^{3\theta} 1 dx \\
&= \frac{1}{2\theta} [x]_{2\theta}^{3\theta} \\
&= \frac{1}{2\theta} \cdot (3\theta - 2\theta) \\
&= \frac{1}{2\theta} \cdot \theta \\
&= \frac{1}{2}
\end{aligned}$$

We add them:

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x) dx &= \int_0^{2\theta} \frac{1}{4\theta^2} x dx + \int_{2\theta}^{3\theta} \frac{1}{2\theta} dx \\
&= \frac{1}{2} + \frac{1}{2} \\
&= 1
\end{aligned}$$

Both conditions are respected:

$f(x) \geq 0$ for all x and $\int_{-\infty}^{\infty} f(x) dx = 1$

Therefore, $f(x)$ is a **density function**.

2. Let X be a random variable with density f . Compute the expectation and variance of X .

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Since $f(x) = 0$ outside $[0, 3\theta]$, the integrale is reduce to:

$$\mathbb{E}[X] = \int_0^{2\theta} \frac{1}{4\theta^2} x^2 dx + \int_{2\theta}^{3\theta} \frac{1}{2\theta} x dx$$

Part 1: 0 to 2θ

$$\begin{aligned}
\int_0^{2\theta} \frac{1}{4\theta^2} x^2 dx &= \frac{1}{4\theta^2} \int_0^{2\theta} x^2 dx \\
&= \frac{1}{4\theta^2} \left[\frac{1}{3} x^3 \right]_0^{2\theta} \\
&= \frac{1}{4\theta^2} \left(\frac{1}{3} \cdot (2\theta)^3 - 0 \right) \\
&= \frac{1}{4\theta^2} \left(\frac{1}{3} \cdot 8\theta^3 \right) \\
&= \frac{1}{4\theta^2} \cdot \frac{8\theta^3}{3} \\
&= \frac{8\theta^3}{12\theta^2} \\
&= \frac{2}{3}\theta
\end{aligned}$$

Part 2: 2θ to 3θ :

$$\begin{aligned}
\int_{2\theta}^{3\theta} \frac{1}{2\theta} x dx &= \frac{1}{2\theta} \int_{2\theta}^{3\theta} x dx \\
&= \frac{1}{2\theta} \left[\frac{1}{2} x^2 \right]_{2\theta}^{3\theta} \\
&= \frac{1}{2\theta} \left(\frac{1}{2} \cdot (3\theta)^2 - \frac{1}{2} \cdot (2\theta)^2 \right) \\
&= \frac{1}{2\theta} \left(\frac{1}{2} \cdot 9\theta^2 - \frac{1}{2} \cdot 4\theta^2 \right) \\
&= \frac{1}{2\theta} \left(\frac{1}{2} \cdot 5\theta^2 \right) \\
&= \frac{1}{4\theta} \cdot 5\theta^2 \\
&= \frac{5}{4}\theta
\end{aligned}$$

Then, we add them:

$$\begin{aligned}
\mathbb{E}[X] &= \frac{2}{3}\theta + \frac{5}{4}\theta \\
&= \left(\frac{8}{12} + \frac{15}{12}\right)\theta \\
&= \frac{23}{12}\theta
\end{aligned}$$

So the expectation is :

$$\boxed{\mathbb{E}[X] = \frac{23}{12}\theta}$$

Let's compute the **variance**:

We will use the following formula:

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Step 1: compute $\mathbb{E}[X^2]$.

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx \\
&= \int_0^{2\theta} \frac{1}{4\theta^2} x^3 dx + \int_{2\theta}^{3\theta} \frac{1}{2\theta} x^2 dx \\
&= \frac{1}{4\theta^2} \int_0^{2\theta} x^3 dx + \frac{1}{2\theta} \int_{2\theta}^{3\theta} x^2 dx \\
&= \frac{1}{4\theta^2} \left[\frac{1}{4} x^4 \right]_0^{2\theta} + \frac{1}{2\theta} \left[\frac{1}{3} x^3 \right]_{2\theta}^{3\theta} \\
&= \frac{1}{4\theta^2} \left(\frac{1}{4} \cdot 16\theta^4 \right) + \frac{1}{2\theta} \left(\frac{1}{3} \cdot 27\theta^3 - \frac{1}{3} \cdot 8\theta^3 \right) \\
&= \frac{1}{4\theta^2} \cdot 4\theta^4 + \frac{1}{2\theta^2} \cdot \frac{19}{3} \theta^3 \\
&= \theta^2 + \frac{19\theta^3}{6\theta} \\
&= \frac{25}{6} \theta^2
\end{aligned}$$

Step 2: compute $(\mathbb{E}[X])^2$

$$\begin{aligned}
(\mathbb{E}[X])^2 &= \left(\frac{23}{12}\theta\right)^2 \\
&= \frac{529}{144}\theta^2
\end{aligned}$$

And then, we just need to subtract them to compute the variance:

$$\begin{aligned}
Var(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
&= \frac{25}{6}\theta^2 - \frac{529}{144}\theta^2 \\
&= \left(\frac{25}{6} - \frac{529}{144}\right)\theta^2 \\
&= \left(\frac{600 - 529}{144}\right)\theta^2 \\
&= \frac{71}{144}\theta^2
\end{aligned}$$

Finally we have:

$$\boxed{\mathbb{E}[X] = \frac{23}{12}\theta} \quad \text{and} \quad \boxed{Var(X) = \frac{71}{144}\theta^2}$$

3. Find an estimator $\hat{\theta}$ of θ using the method of moments.

We consider the random variable X with density f depending on the parameter $\theta > 0$. We observed n independent realizations X_1, X_2, \dots, X_n with all the same distribution as X .

An estimator of θ is a function of the observed data X_1, \dots, X_n for which we can compute a value thanks to the observations.

Moments are quantitative measures that describe the shape and characteristics of a distribution (first moment = mean, second moment = variance, etc). The **method of moments** matches sample moments (that we can measure from observations) to theoretical moments (with infinite population, we can't really compute).

Then, for the first moment we have:

- the theoretical first moment is the **expectation** of X :

$$\mathbb{E}[X] = \frac{23}{12}\theta$$

- the first sample moment is the **sample mean**:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

An estimator $\hat{\theta}$ for θ would be a solution of the moment equation:

$$\mathbb{E}[X] = \bar{X}$$

$$\Leftrightarrow \frac{23}{12}\theta = \bar{X}$$

Thus, an estimator would be :

$$\begin{aligned} \hat{\theta}_1 &= \frac{12}{23}\bar{X} \\ \Leftrightarrow \hat{\theta}_1 &= \frac{12}{23} \cdot \frac{1}{n} \sum_{i=1}^n X_i \end{aligned}$$

We could also have another estimator for the second moment :

- the theoretical second moment (the variance) of X :

$$Var(X) = \frac{71}{144}\theta^2$$

- the sample variance S^2 :

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

The estimator would be obtained thanks to the moment equation:

$$Var(X) = S^2$$

$$\Leftrightarrow \frac{71}{144}\theta^2 = S^2$$

Thus, the estimator would be :

$$\hat{\theta}_2^2 = \frac{144}{71}S^2$$

$$\Leftrightarrow \hat{\theta}_2 = \sqrt{\frac{144}{71} \cdot \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

4. Define a confidence interval for θ .

We search a confidence interval for θ , which is a range of the possible values for the unknown parameter θ , based on our sample X_1, X_2, \dots, X_n . Thus, we need to have an idea of the variability of the estimator.

$$\text{Estimator} \pm \text{Margin of error}$$

$\hat{\theta}$ is the estimator of θ , obtained using the method of moments based on the sample mean.

$$\hat{\theta} = \frac{12}{23}\bar{X} \quad \text{and} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

The margin of error depends on the standard error : $SE(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$.

The theoretical expectation and variance of X are:

$$\mathbb{E}[X] = \frac{23}{12}\theta \quad \text{and} \quad Var(X) = \frac{71}{144}\theta^2$$

Here we see that the variance depends on θ , which is an unknown parameter that we are trying to estimate and for which we are searching the confidence interval. As a consequence, the variance remains unknown and we can just estimate it. We are thus in a context of **not Gaussian settings and unknown variance**.

The confidence interval is such that :

$$1 - \alpha = P(\theta \in [A, B]) \quad \text{with } 1 - \alpha \text{ the confidence level}$$

We can write :

$$A = \hat{\theta} - \text{margin} \quad \text{and} \quad B = \hat{\theta} + \text{margin}$$

So :

$$1 - \alpha = P(-\text{margin} \leq \hat{\theta} \leq \text{margin})$$

Even if X is not Gaussian, thanks to the **Central Limit Theorem**, we assume that the sample mean \bar{X} is approximately normal for large n . Because $\hat{\theta}$ is a linear function of the sample mean: $\hat{\theta} = \frac{12}{23}\bar{X}$, $\hat{\theta}$ is approximately normal.

So for large n :

$$\hat{\theta} \approx \mathcal{N}\left(\theta, Var(\hat{\theta})\right)$$

So we have:

$$\begin{aligned} Var(\hat{\theta}) &= Var\left(\frac{12}{23}\bar{X}\right) \\ &= \left(\frac{12}{23}\right)^2 Var(\bar{X}) \quad \text{because if } Y = aX, \text{ with } a \text{ constant, then } Var(Y) = a^2 Var(X) \end{aligned}$$

Then the variance of the mean:

$$Var(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n Var(X_i)$$

$$= \frac{nVar(X)}{n^2} \quad \text{because the variance of the sum of independent variables is the sum of}$$

and there are identically distributed

$$= \frac{Var(X)}{n}$$

Thus, all put together we have :

$$\begin{aligned} Var(\hat{\theta}) &= \left(\frac{12}{23}\right)^2 Var(\bar{X}) \\ &= \left(\frac{12}{23}\right)^2 \cdot \frac{1}{n} \cdot Var(X) \\ &= \left(\frac{12}{23}\right)^2 \cdot \frac{1}{n} \cdot \frac{71}{144} \theta^2 \\ &= \frac{144}{529} \cdot \frac{71}{144} \cdot \frac{1}{n} \theta^2 \\ &= \frac{71}{529n} \theta^2 \end{aligned}$$

The variance of θ is still unknown because it depends on θ which is unknown. But now we can deduce the standard error of $\hat{\theta}$

$$\begin{aligned} SE(\hat{\theta}) &= \sqrt{Var(\hat{\theta})} \\ &= \sqrt{\frac{71}{529n} \theta^2} \\ &= \frac{\theta \sqrt{71}}{23\sqrt{n}} \end{aligned}$$

It still depend on θ which is unknown but the Slutsky's lemma allows us to replace θ by its estimator $\hat{\theta}$ for large n , since $\hat{\theta} \xrightarrow{p} \theta$ (converge in probability).

Thus:

$$SE(\hat{\theta}) = \frac{\theta \sqrt{71}}{23\sqrt{n}} \approx \frac{\hat{\theta} \sqrt{71}}{23\sqrt{n}}$$

For now, we know that $\hat{\theta}$ is approximately normal when n large, but the variance and standard error still depends on θ , and we don't know the shape of its distribution. Thus, we standardize $\hat{\theta}$ to have something that follow approximately a standard normal distribution:

The standardized estimator is:

$$Z_n = \frac{\hat{\theta} - \theta}{SE(\hat{\theta})}$$

Now :

$$Z_n \approx \mathcal{N}(0, 1)$$

And we can write:

$$P\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} \leq z_{1-\frac{\alpha}{2}}\right) \approx 1 - \alpha$$

where $z_{1-\frac{\alpha}{2}}$ is the **standard normal critical value** which can be find in specific tables (that is why we also need to know the distribution).

For example, for a risk of 5% ($\alpha = 0.05$) we have a 95% confidenc interval ($1 - \alpha = 0.95$) and $z_{1-\frac{\alpha}{2}} = z_{0.975} \approx 1.96$.

We multiply everything my $SE(\hat{\theta})$ and rearranging :

$$P\left(\hat{\theta} - z_{1-\frac{\alpha}{2}} \cdot SE(\hat{\theta}) \leq \theta \leq \hat{\theta} + z_{1-\frac{\alpha}{2}} \cdot SE(\hat{\theta})\right) \approx 1 - \alpha$$

And finally we can write the confidence interval of θ :

$$\theta \in \left[\hat{\theta} - z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\theta}\sqrt{71}}{23\sqrt{n}}, \quad \hat{\theta} + z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\theta}\sqrt{71}}{23\sqrt{n}} \right]$$

5. Simulate 1000 observations of X .

We want to simulate 1000 random values that follow the distribution of X . Since X does not follow a standard distribution, we cannot easily simulate it. We are going to use the **inverse transform sampling method**. The idea is that any **continuous random variable** can be generated from a **uniform random variable** on $[0, 1]$, by applying the **inverse of its cumulative distribution function (CDF)**.

Basically, we are going to:

- **step 1**: compute the CDF of $f(x)$: $F(x)$,
- **step 2**: compute the inverse of $F(x)$: $F^{-1}(x)$,
- **step 3**: generate a uniform random variable $U \sim Uniform(0, 1)$,
- **step 4**: transform them using $X = F^{-1}(U)$.

With this method we should obtain the 1000 observations following the distribution of X .

X has for density:

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

Step 1: Compute the CDF of $f(x)$: $F(x)$

By definition:

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_{-\infty}^x f(x) dx \end{aligned}$$

But because the distribution of $f(x)$ is piecewise, so we integrate it in piece:

When $x < 0$: $f(x) = 0$, thus $F(x) = 0$

When $0 \leq x \leq 2\theta$:

$$\begin{aligned}
 F(x) &= \int_0^x f(x) dx \\
 &= \int_0^x \frac{1}{4\theta^2} x dx \\
 &= \frac{1}{4\theta^2} \left[\frac{1}{2} x^2 \right]_0^x \quad \text{see computation details above} \\
 &= \frac{1}{4\theta^2} \left(\frac{1}{2} x^2 - 0 \right) \\
 &= \frac{1}{4\theta^2} \cdot \frac{1}{2} x^2
 \end{aligned}$$

$$\text{So} \quad F(x) = \frac{x^2}{8\theta^2} \quad \text{when } 0 \leq x \leq 2\theta$$

When $2\theta \leq x \leq 3\theta$:

$$\begin{aligned}
 F(x) &= \int_0^{2\theta} f(x) dx + \int_{2\theta}^x f(x) dx \quad \text{Because we are in CDF, we accumulate the probabilities} \\
 &= \int_0^{2\theta} \frac{1}{4\theta^2} x dx + \int_{2\theta}^x \frac{1}{2\theta} dx \\
 &= \frac{1}{4\theta^2} \left[\frac{1}{2} x^2 \right]_0^{2\theta} + \frac{1}{2\theta} [x]_{2\theta}^x \\
 &= \frac{1}{4\theta^2} \left(\frac{1}{2} (2\theta^2) \right) + \frac{1}{2\theta} (x - 2\theta) \\
 &= \frac{1}{4\theta^2} \cdot \frac{1}{2} \cdot 4\theta^2 + \frac{x - 2\theta}{2\theta} \\
 &= \frac{1}{2} + \frac{x - 2\theta}{2\theta} \\
 &= \frac{\theta + x - 2\theta}{2\theta} \\
 &= \frac{x - \theta}{2\theta}
 \end{aligned}$$

When $x \geq 3\theta$: $F(x) = 1$ because we are above the possible values of x and we thus already reached 1.

Thus the final CDF is :

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{x^2}{8\theta^2} & \text{if } 0 \leq x \leq 2\theta \\ \frac{x-\theta}{2\theta} & \text{if } 2\theta \leq x \leq 3\theta \\ 1 & \text{if } x \geq 3\theta \end{cases}$$

Step 2: find the invert of $F(x)$: $F^{-1}(x)$.

The inverse function is such that $F^{-1}(F(x)) = x$.

In fact, we want the function that gives x for a given cumulative probability $t \in [0, 1]$.

When $0 \leq x \leq 2\theta$:

$$F(x) = \frac{x^2}{8\theta^2} = u$$

$$\Rightarrow x^2 = u8\theta^2$$

$$\Rightarrow x = \sqrt{u8\theta^2}$$

$$= \sqrt{u8\theta^2}$$

$$= \sqrt{u \cdot 2 \cdot 4\theta^2}$$

$$= 2\theta\sqrt{2u}$$

When $2\theta \leq x \leq 3\theta$:

$$F(x) = \frac{x-\theta}{2\theta} = u$$

$$\Rightarrow x - \theta = u2\theta$$

$$\Rightarrow x = 2\theta u + \theta$$

$$= (2u + 1)\theta$$

We also change the limits of u for $F^{-1}(x)$ because it is now defined on $[0, 1]$.

So now we have :

$$F^{-1}(u) = \begin{cases} 2\theta\sqrt{2u} & \text{if } 0 \leq u \leq 0.5 \\ (2u + 1)\theta & \text{if } 0.5 < u \leq 1 \end{cases} \quad \text{because } F(2\theta) = \frac{2\theta^2}{8\theta^2} = 0.5$$

Step 3: generate a uniform random variable $U \sim Uniform(0, 1)$

```
set.seed(42) # for reproducibility
n <- 1000
U <- runif(n, min = 0, max = 1)
```

Step 4: transform them using $X = F^{-1}(U)$.

```
# Create the inverse function:
invF <- function(u, theta) {
  ifelse( u <= 0.5,
    2 * theta * sqrt(2 * u),
    (2 * u + 1) * theta )
}
```

Now we can transform our 1000 observations generated with the uniform distribution:

```
# define theta:
theta <- 1
X <- invF(U, theta)
```

6. Show with a graph that your answer is correct.

Let's plot an histogram to visualize the data we simulated:

We can plot the curve for the theoretical PDF to see if it matches the simulated distribution :

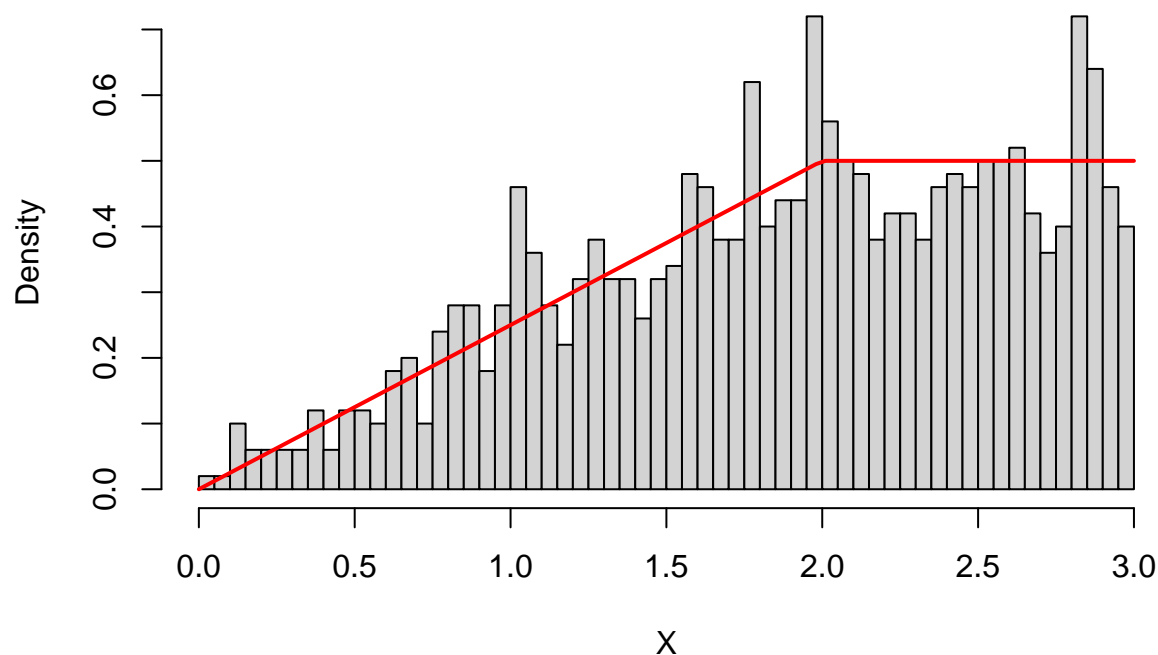
$$f(x) = \begin{cases} \frac{1}{4\theta^2}x & \text{if } x \in [0, 2\theta] \\ \frac{1}{2\theta} & \text{if } x \in [2\theta, 3\theta] \\ 0 & \text{otherwise} \end{cases}$$

```
# Theoretical density function
f <- function(x, theta) {
  ifelse(x >= 0 & x <= 2*theta,
    (1/(4*(theta^2))) * x,
    ifelse(x >= 2*theta & x <= 3*theta,
      1/(2*theta),
      0))
}
```

Now let's plot the histogram and the PDF curve:

```
hist(X, probability = TRUE, breaks = 50,
  main = paste0("Simulated X (n = ", n, " with theta = ", theta, ")") )
curve(f(x, theta), add = TRUE, col = "red", lwd = 2)
```

Simulated X (n = 1000 with theta = 1)



We can see that the PDF curve matches quite well the data we simulated.

Also, now that we have defined θ we can also compare the sample mean and the expectation and the variances:

```
sample_mean <- round(mean(X), 2)
expectation <- round((23/12) * theta ,2)

sample_var <- round(var(X), 2)
Var_X <- round((71/144) * theta^2, 2)

cat("Sample mean = ", sample_mean, "\n")
```

```
## Sample mean = 1.89
```

```
cat("Expectation = ", expectation, "\n")
```

```
## Expectation = 1.92
```

```
cat("Sample variance = ", sample_var, "\n")
```

```
## Sample variance = 0.51
```

```
cat("Var(X) = ", Var_X, "\n")
```

```
## Var(X) = 0.49
```

Not so bad !

Exercise C:

Let X and Y be two independent standard Gaussian random variables.
 α a positive real number.

Let ϵ be a random variable, independent from X and Y taking values $-\alpha$ and α with same probability $1/2$.

Let $Z = \epsilon X$.

1. Prove that Z is Gaussian random variable.

Let $t \in \mathbb{R}$.

With the law of total probability:

$$\begin{aligned} F_Z(t) &= P(Z \leq t) \\ &= P(\epsilon X \leq t) \\ &= P(\epsilon X \leq t, \epsilon = \alpha) + P(\epsilon X \leq t, \epsilon = -\alpha) \quad \text{because } \epsilon \text{ can take only } \alpha \text{ or } -\alpha \\ &= P(\alpha X \leq t, \epsilon = \alpha) + P(-\alpha X \leq t, \epsilon = -\alpha) \\ &= P(\alpha X \leq t) \cdot P(\epsilon = \alpha) + P(-\alpha X \leq t) \cdot P(\epsilon = -\alpha) \quad \text{because } \epsilon \text{ and } X \text{ are independent.} \\ &= \frac{1}{2}P(\alpha X \leq t) + \frac{1}{2}P(-\alpha X \leq t) \quad \text{because } P(\epsilon = \alpha) = P(\epsilon = -\alpha) = \frac{1}{2} \\ &= \frac{1}{2}P(X \leq \frac{t}{\alpha}) + \frac{1}{2}P(X \geq -\frac{t}{\alpha}) \quad \text{divide by } \alpha \text{ because } \alpha > 0 \\ &= \frac{1}{2} \left(P(X \leq \frac{t}{\alpha}) + P(X \geq -\frac{t}{\alpha}) \right) \end{aligned}$$

Because $X \sim \mathcal{N}(0, 1)$ it is symmetric and we can say:

$$P(X \leq \frac{t}{\alpha}) = P(X \geq -\frac{t}{\alpha})$$

So:

$$\begin{aligned}
F_Z(t) &= \frac{1}{2} \left(P(X \leq \frac{t}{\alpha}) + P(X \geq -\frac{t}{\alpha}) \right) \\
&= \frac{1}{2} \left(2 \cdot P(X \leq \frac{t}{\alpha}) \right) \\
&= P(X \leq \frac{t}{\alpha})
\end{aligned}$$

The distribution of $F_Z(t)$ is the same as the one of X but scaled by α .
Thus Z is a **Gaussian random variable**:

$$Z \sim \mathcal{N}(0, \alpha^2)$$

2. Determine the distribution of the random vector $U = (Y, Y - Z, Y + Z)$.

$$Y \sim \mathcal{N}(0, 1) \quad \text{and} \quad Z \sim \mathcal{N}(0, \alpha^2)$$

$Z = \epsilon X$ and X and Y are independent. So because Z is a function of X , we can deduce that Z and Y are independent.

We have :

$$U = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} = \begin{pmatrix} Y \\ Y - Z \\ Y + Z \end{pmatrix}$$

Because U is a vector made with **Gaussian random variables** and **linear transformation of independent Gaussian random variables**, we can say that U is a **multivariate random vector**.

To find its distribution, we just need to find $\mathbb{E}[U]$ and $Var(U)$.

Expectation:

If the expectation of Y and Z exist, thus the expectation of U exists.

Also, $\mathbb{E}[Y] = 0$ and $\mathbb{E}[Z] = 0$, so we have :

$$\mathbb{E}[U] = \begin{pmatrix} \mathbb{E}[U_1] \\ \mathbb{E}[U_2] \\ \mathbb{E}[U_3] \end{pmatrix} = \begin{pmatrix} \mathbb{E}[Y] \\ \mathbb{E}[Y - Z] \\ \mathbb{E}[Y + Z] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Variance:

$$Var(U) = \begin{pmatrix} cov(U_1, U_1) & cov(U_1, U_2) & cov(U_1, U_3) \\ cov(U_2, U_1) & cov(U_2, U_2) & cov(U_2, U_3) \\ cov(U_3, U_1) & cov(U_3, U_2) & cov(U_3, U_3) \end{pmatrix}$$

As it is a lot of computations let's do first :

$$cov(Y, Z) = cov(Z, Y) = \mathbb{E}[YZ] - \mathbb{E}[Y] \cdot \mathbb{E}[Z] = 0$$

$$cov(Y, Y) = Var(Y) = 1$$

$$cov(Z, Z) = Var(Z) = \alpha^2$$

Let's do it step by step:

$$\text{cov}(U_1, U_1) = \text{Var}(U_1) = \text{Var}(Y) = 1$$

$$\text{cov}(U_1, U_2) = \text{cov}(Y, Y - Z) = \text{cov}(Y, Y) - \text{cov}(Y, Z) = \text{Var}(Y) - \text{cov}(Y, Z) = 1 - 0 = 1$$

$$\text{cov}(U_1, U_3) = \text{cov}(Y, Y + Z) = \text{cov}(Y, Y) + \text{cov}(Y, Z) = \text{Var}(Y) + \text{cov}(Y, Z) = 1 + 0 = 1$$

$$\text{cov}(U_2, U_1) = \text{cov}(U_1, U_2) = 1$$

$$\text{cov}(U_2, U_2) = \text{Var}(U_2) = \text{Var}(Y - Z) = \text{Var}(Y) + \text{Var}(Z) - 2\text{cov}(Y, Z) = 1 + \alpha^2 - 0 = 1 + \alpha^2$$

$$\text{cov}(U_2, U_3) = \text{cov}(Y - Z, Y + Z) = \text{cov}(Y, Y) + \text{cov}(Y, Z) - \text{cov}(Z, Y) - \text{cov}(Z, Z) = 1 + 0 - 0 - \alpha^2 = 1 - \alpha^2$$

$$\text{cov}(U_3, U_1) = \text{cov}(U_1, U_3) = 1$$

$$\text{cov}(U_3, U_2) = \text{cov}(U_2, U_3) = 1 - \alpha^2$$

$$\text{cov}(U_3, U_3) = \text{Var}(U_3) = \text{Var}(Y + Z) = \text{Var}(Y) + \text{Var}(Z) + 2\text{cov}(Y, Z) = 1 + \alpha^2 + 0 = 1 + \alpha^2$$

So we have:

$$\text{Var}(U) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 + \alpha^2 & 1 - \alpha^2 \\ 1 & 1 - \alpha^2 & 1 + \alpha^2 \end{pmatrix}$$

So the distribution of the vector U is:

$$U \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 + \alpha^2 & 1 - \alpha^2 \\ 1 & 1 - \alpha^2 & 1 + \alpha^2 \end{pmatrix} \right)$$

3. Given a sample of Z of size n , determine an unbiased estimator of α^2 and derive a confidence interval for α^2 .

Estimator of α^2 :

Let $\hat{\alpha}^2$ the estimator of α^2 .

By definition, if $\hat{\alpha}^2$ is unbiased, we need:

$$\mathbb{E}[\hat{\alpha}^2] = \alpha^2$$

We know that $\text{Var}(X) = \alpha^2$. Thus, we search the estimator for the second moment. The estimator is obtained thanks to the moment equation that matches the theoretical variance $\text{Var}(X)$ to the sample variance S^2 : The theoretical second moment (the variance) of X :

$$Var(X) = S^2$$

$$\Leftrightarrow \alpha^2 = \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{Z})^2$$

With the equation moment for the estimator for the first moment we know that:

$$\mathbb{E}[Z] = \bar{Z} = 0$$

So we have:

$$\alpha^2 = \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{Z})^2$$

$$= \frac{1}{n} \sum_{i=1}^n (Z_i - 0)^2$$

$$= \frac{1}{n} \sum_{i=1}^n Z_i^2$$

So the estimator for α^2 is:

$$\hat{\alpha}^2 = \frac{1}{n} \sum_{i=1}^n Z_i^2$$

Now we need to check if it is unbiased:

$$\mathbb{E}[\hat{\alpha}^2] = \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n Z_i^2 \right]$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Z_i^2]$$

By definition we have :

$$Var(Z) = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2$$

$$\Leftrightarrow \mathbb{E}[Z^2] = Var(Z) + (\mathbb{E}[Z])^2$$

$$= Var(Z) + 0$$

$$= \alpha^2$$

So:

$$\begin{aligned}
\mathbb{E}[\hat{\alpha}^2] &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[Z_i^2] \\
&= \frac{1}{n} \sum_{i=1}^n \alpha^2 \\
&= \frac{1}{n} \cdot n \cdot \alpha^2 \\
&= \alpha^2
\end{aligned}$$

So the estimator $\hat{\alpha}^2$ is **unbiased**.

Confidence interval for α^2 :

We are in a context of **Gaussian settings but unknown variance** (since α^2 is the parameter we want to estimate).

We want to find the confident interval for α^2 such that :

$$1 - \gamma = P(\alpha^2 \in [A, B]) \quad \text{with } 1 - \gamma \text{ the level of confidence of the interval}$$

$$1 - \gamma = P(A \leq \alpha^2 \leq B)$$

where A and B are the bounds of the confidence interval. The bounds need to be function of the data, but we need to know the distribution.

We know that :

$$Z \sim \mathcal{N}(0, \alpha^2)$$

So if we scale it by dividing by α we obtain a standard Gaussian distribution :

$$\frac{Z}{\alpha} \sim \mathcal{N}(0, 1)$$

If we use only $\frac{Z}{\alpha}$, we know the distribution but the problem is that it does not directly depend on α^2 so we cannot have a probabilistic statement about α^2 .

So, we are going to use the squares: $\left(\frac{Z}{\alpha}\right)^2$ because we know that this follows a χ^2 distribution. Indeed, by definition, the square of a standard normal random variable follows a χ^2 distribution, with 1 degree of freedom.

Thus:

$$\left(\frac{Z}{\alpha}\right)^2 \sim \chi^2$$

Let Z_1, \dots, Z_n be observations of Z , now we have:

$$\left(\frac{Z_i}{\alpha}\right)^2 \sim \chi_1^2 \Leftrightarrow \sum_{i=1}^n \left(\frac{Z_i}{\alpha}\right)^2 \sim \chi_n^2 \Leftrightarrow \frac{1}{\alpha^2} \cdot \sum_{i=1}^n (Z_i)^2 \sim \chi_n^2$$

So we have something where we know the distribution, it depends on the data and we can isolate α^2 .

We now have :

$$P\left(\chi_{n,\gamma_1}^2 \leq \frac{1}{\alpha^2} \cdot \sum_{i=1}^n (Z_i)^2 \leq \chi_{n,1-\gamma_2}^2\right) \text{ with } \gamma_1, \gamma_2 > 0 \text{ and } \gamma_1 + \gamma_2 = \gamma$$

So we can finally isolate α^2 !

$$P\left(\chi_{n,\gamma_1}^2 \leq \frac{1}{\alpha^2} \cdot \sum_{i=1}^n (Z_i)^2 \leq \chi_{n,1-\gamma_2}^2\right) \Leftrightarrow P\left(\chi_{n,\gamma_1}^2 \cdot \alpha^2 \leq \sum_{i=1}^n (Z_i)^2 \leq \chi_{n,1-\gamma_2}^2 \cdot \alpha^2\right)$$

$$\chi_{n,\gamma_1}^2 \cdot \alpha^2 \leq \sum_{i=1}^n (Z_i)^2 \quad \sum_{i=1}^n (Z_i)^2 \leq \chi_{n,1-\gamma_2}^2 \cdot \alpha^2$$

$$\alpha^2 \leq \frac{\sum_{i=1}^n (Z_i)^2}{\chi_{n,\gamma_1}^2} \quad \frac{\sum_{i=1}^n (Z_i)^2}{\chi_{n,1-\gamma_2}^2} \leq \alpha^2$$

So at the end we have:

$$P\left(\frac{\sum_{i=1}^n (Z_i)^2}{\chi_{n,1-\gamma_2}^2} \leq \alpha^2 \leq \frac{\sum_{i=1}^n (Z_i)^2}{\chi_{n,\gamma_1}^2}\right)$$

so the confidence interval for α^2 is :

$$\left[\frac{\sum_{i=1}^n (Z_i)^2}{\chi_{n,1-\gamma_2}^2} , \frac{\sum_{i=1}^n (Z_i)^2}{\chi_{n,\gamma_1}^2} \right]$$

4. Compute the interval numerically with the attached data for a confidence interval of level 95%.

First, we need to load the data.

```
load("D:/_DSTI/A24_Foundations_Statistical_Analysis_ML_Part2/Exam/dataExC.RDATA")
hist(ZZ,)
```



We want a 95% confidence interval for α^2 .

This means:

$$1 - \gamma = 0.95 \quad \Rightarrow \quad \gamma = 0.05, \quad \gamma_1 = \gamma_2 = \frac{\gamma}{2} = 0.025$$

Then, we need the critical values of the χ^2 distribution. These values can be obtained through specific tables and depend on level of confidence we want and the degree of freedom, which depend on the sample size. However, here the degree of freedom is high ($n = 1000$) and most of tables stop at 100. We can thus use the R function `qchisq()`.

```
n <- length(ZZ) # sample size
gamma <- 0.05 # confidence level

lower_chi2 <- qchisq(1 - gamma/2, df = n)
upper_chi2 <- qchisq(gamma/2, df = n)
```

Now we can compute the confidence interval of level 95%.

```
# Compute the sum of square:
sum_sq <- sum(ZZ^2)

# Compute the bounds of the confidence interval:
lower_bound <- (sum_sq / lower_chi2)
upper_bound <- (sum_sq / upper_chi2)
```

Then, we can check if the variance α^2 is in the confidence interval we computed:

```
alpha_sq <- var(ZZ)
```

```
cat("Sample variance:", round(alpha_sq, 4), "\n")
```

```
## Sample variance: 4.094
```

```
cat("95% Confidence interval for variance = [",  
    round(lower_bound, 4), ", ",  
    round(upper_bound, 4), "]\n")
```

```
## 95% Confidence interval for variance = [ 3.7613 ,  4.4823 ]
```

We see that the variance is indeed in the confidence interval.