hw3

May 14, 2020

1 Problem 3

1.1 Question a)

Let us write the likelihood function in this setting :

$$L(\alpha, \lambda) = f_{\alpha,\lambda}(T_1) \times f_{\alpha,\lambda}(T_2) \times \mathbb{P}(T > 100K) \times \mathbb{P}(T > T_4) \times \mathbb{P}(T > T_5)$$

After calculations, the negative log-likelihood is equal to :

$$-l(\alpha, \lambda) = -[2\log(\alpha\lambda) + (\alpha - 1)\log(\lambda^2 T_1 T_2) - (\lambda)^\alpha (T_1^\alpha + T_2^\alpha)] + (\lambda)^\alpha (T_3^\alpha + T_4^\alpha + T_5^\alpha)$$

where $T_1 = 44K, T_2 = 26K, T_3 = 100K, T_4 = 19K, T_5 = 45K$.

We can solve the MLE numerically with `scipy` module `optimize`.

[23]:
```
import numpy as np
from scipy import optimize
from scipy.stats import weibull_min
from math import gamma

# Problem 3
T1 = 44 * 10 ** 3
T2 = 26 * 10 ** 3
T3 = 100 * 10 ** 3
T4 = 19 * 10 ** 3
T5 = 45 * 10 ** 3


def neg_log_L(gamma, alpha) -> float:
    A = 2 * np.log(alpha / gamma) + (alpha - 1) * np.log(T1 * T2 / (gamma **␣
↪2)) - \
        ((1 / gamma) ** alpha) * ((T1 ** alpha) + (T2 ** alpha))
    B = -((1 / gamma) ** alpha) * (T3 ** alpha + T4 ** alpha + T5 ** alpha)
    return -A - B


## Numerically solve MLE equations.
```

```
epsilon = 1 * 10 ** -12
x0 = [10, 10]
bnds = ((epsilon, np.inf), (epsilon, np.inf))
fun = lambda x: neg_log_L(x[0], x[1])
solver = optimize.minimize(fun, x0=x0, bounds=bnds)
print(" {} iterations \n".format(solver.nit),
      "lambda = {} and alpha = {}".format(1 / solver.x[0], solver.x[1]))
```

```
171 iterations
lambda = 1.0823481787828909e-05 and alpha = 1.535248714292919
```

1.2 Question b)

We want to compute a $(1-\delta)\%$ confidence interval for $\mathbb{E}(T)$. Theoretically :

$$\mathbb{E}(T) = \frac{1}{\lambda}\Gamma(1 + \frac{1}{\alpha})$$

So we have the plug-in estimate of $\mathbb{E}(T)$ by using $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$.

Let us compute a parametric bootstrap confidence interval.

- 1) Generate $(\mu_n^{(1)}, \ldots, \mu_n^{(m)})$ where $\mu_n^{(i)} = \frac{1}{n}\sum_{k=1}^{n} T_k^{(i)}$
- 2) Find x and y such that $\mathbb{P}(\mu_n - \mu_* < x) = 1 - \delta/2$ and $\mathbb{P}(\mu_n - \mu_* < y) = \delta/2$ where μ_* is the plug-in estimate obtained with the MLE.

[24]:
```
## Parametric Boostrap (1-delta) confidence interval
lmda = 1 / solver.x[0]
alpha = solver.x[1]
# reference parameter
mu_star = (1 / lmda) * gamma(1 + 1 / alpha)
print(" In average, {} kms before a reliability problem occurs (plug-in␣
 ↪estimation with MLE)".format(mu_star))

m = 100  # number of iterations to aggregate
bootstrap_estimates = []
for i in range(m):
    T_bootstrap = weibull_min.rvs(c=alpha, scale=1 / lmda, size=100)
    bootstrap_estimates.append(np.mean(T_bootstrap))

delta = 0.1
print(" delta = {}".format(delta), "\n", "m = {} ; ".format(m), "n = 100")
# Upper bound : we find x s.t prob(estimate - mu_star < x) = 1-delta/2
x = 2000
count = 0
while (count / m != 1 - delta / 2):
    count = 0
    for i in range(m):
```

```
        if bootstrap_estimates[i] - mu_star < x:
            count += 1
    # print(count / m)
    x += 10
print(" Prob(mu_n - mu_estimate < {}) = {}".format(x, count / m))

# Lower bound : we find y s.t prob(estimate - mu_star < y ) = delta/2
y = -10000
count = 0
while (count / m != delta / 2):
    count = 0
    for i in range(m):
        if bootstrap_estimates[i] - mu_star < y:
            count += 1
    # print(count / m)
    y += 10
print(" Prob(mu_n - mu_estimate < {}) = {}".format(y, count / m))

print(" Parametric Boostrap {}% confidence interval : [{},{}]".format((1 -␣
↪delta)*100, -x + np.mean(bootstrap_estimates),
                                                                     - y + np.
↪mean(bootstrap_estimates)))
```

```
 In average, 83182.41571001986 kms before a reliability problem occurs (plug-in
estimation with MLE)
 delta = 0.1
 m = 100 ;  n = 100
 Prob(mu_n - mu_estimate < 8170) = 0.95
 Prob(mu_n - mu_estimate < -9780) = 0.05
 Parametric Boostrap 90.0% confidence interval :
[74749.82839853111,92699.82839853111]
```

2 Problem 6

Written part is in the written section of the report.

[25]:
```
# Problem 4
## Posterior distribution
X = np.array([6.00, 4.82, 3.35, 2.38, 3.59, 4.12, 4.98, 2.69, 6.24, 6.77,
              6.22, 5.42, 5.42, 3.10, 4.65, 4.24, 4.53, 4.62, 5.36, 2.57])

Sigma = (1 + (0.5) * np.sum(X[0:19]**2))**-1
X_i_1 = X[0:19]
X_i = X[1:20]
Mu = Sigma * ( (0.5) * np.sum(X_i_1*(X_i-2)) + 0.5)
print("rho*|X follows a normal distribution with mean = {} and variance = {}".
↪format(Mu, Sigma))
```

```python
## Prob(X_23 > 4)
n = 100000
X23_estimates = []
count = 0
for i in range(n) :
    rho = np.random.normal(Mu, Sigma, 1)
    epsilon = np.random.normal(0,2,3)
    X23 = (rho**3) * X[19] + 2 * (1 + rho + rho**2) + epsilon[0] + epsilon[1] *␣
 ↪rho + epsilon[2] * rho**2
    X23_estimates.append(X23)
    if X23 >= 4 :
        count += 1

print("X23 estimate = {}".format(np.mean(X23_estimates)))
print("Prob(X_23 > 4) = {}".format(count/n))
```

```
rho*|X follows a normal distribution with mean = 0.5160781386089336 and variance
= 0.0045245286006764165
X23 estimate = 3.9179892166389694
Prob(X_23 > 4) = 0.4861
```

Homework 3 – T. Biuyelle

Problem 1

a) First, let us note that X is non-random and observable.

We have :

$$\hat{\beta} = \left[(X^TX)^{-1}X^T + D\right](X\beta + \varepsilon)$$

$$= \beta + DX\beta + C\varepsilon .$$

Then $\mathbb{E}[\hat{\beta}] = (1 + DX)\beta + C\underbrace{\mathbb{E}[\varepsilon]}_{=0}$

In order to have $\hat{\beta}$ unbiased, we must have :

$$\boxed{DX = 0}$$

b) $\mathbb{E}(\hat\beta \hat\beta^T) = C\, \mathbb{E}(YY^T) C^T$

$= C \left[(X\beta)(X\beta)^T + 1 \right] C^T$

since $YY^T = (X\beta)(X\beta)^T + X\beta\varepsilon^T + \varepsilon\beta^T X^T + \varepsilon\varepsilon^T$

<u>So</u> :

$$\mathbb{E}(\hat\beta\hat\beta^T) - \overbrace{C(X\beta)(X\beta)^T C^T}^{= \mathbb{E}[\hat\beta] \times \mathbb{E}[\hat\beta^T]}$$

$$= \left((X^TX)^{-1}X^T + D \right)\left((X^TX)^{-1}X^T + D \right)^T$$

$$= (X^TX)^{-1} X^T X (X^TX)^{-1} +$$

$$(X^TX)^{-1} X^T \underbrace{D^T}_{=0} + \underbrace{DX}_{=0} (X^TX)^{-1} + DD^T$$

$$= \underbrace{(X^TX)^{-1}} + DD^T$$

$$= \mathbb{E}(\hat\beta_* \hat\beta_*^T) - \underbrace{\mathbb{E}(\hat\beta_*) \cdot \mathbb{E}[\hat\beta_*^T]}_{= \beta\beta^T}$$

Finally we have :

$$\mathbb{E}(\hat{\beta}\hat{\beta}^T) = \mathbb{E}(\hat{\beta}_*\hat{\beta}_*^T) + \underbrace{C(X\beta)(X\beta)^T C^T + \beta\beta^T + DD^T}_{\text{non-negative definite}},$$

Hence we have

$$\boxed{\mathbb{E}(\hat{\beta}\hat{\beta}^T) \succcurlyeq \mathbb{E}(\hat{\beta}_*\hat{\beta}_*^T)}$$

Problem 4

a) If $\bar{X}_m \leq \bar{Y}_n$ then the MLE for (μ_1, μ_2) under the constraints $\hat{\mu}_1 \leq \hat{\mu}_2$ are given by:

$$\hat{\mu}_1 = \frac{1}{m}\sum_{i=1}^{m} X_i$$

$$\boxed{\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n} Y_i}$$

Else if $\bar{X}_m > \bar{Y}_n$, then we must look for $\hat{\mu}_1 = \hat{\mu}_2$.
Consequently we solve:

$$\frac{\partial \ell(X, Y, \theta)}{\partial \theta} = 0 \text{ where } \theta = \mu$$

So:

$$\boxed{\begin{gathered}\hat{\mu} = \frac{1}{m+n}\left(\sum_{i=1}^{m} X_i + \sum_{j=1}^{n} Y_j\right) \\ \text{and } \hat{\mu}_1 = \hat{\mu}_2 = \hat{\mu}.\end{gathered}}$$

b) Let us assume that $\mu_1 < \mu_2$.

Then $\mathbb{P}(X_n > Y_m) \longrightarrow 0$
as $n \to +\infty$ and $m \to +\infty$.

So $\mathbb{1}_{\{\bar{X}_m \leq \bar{Y}_n\}} \xrightarrow[m+n \to +\infty]{\mathbb{P}} 1$.

Besides the CLT ensures that:

$$\begin{cases} \sqrt{m}\left(\bar{X}_m - \mu_1\right) \xrightarrow[m \to +\infty]{\mathcal{D}} \mathcal{N}(0,1) \\ \sqrt{n}\left(\bar{Y}_n - \mu_2\right) \xrightarrow[n \to +\infty]{\mathcal{D}} \mathcal{N}(0,1) \end{cases}$$

By Slutsky's theorem:

$$\left[\left(\mathbb{1}_{\{\bar{X}_m \leq \bar{Y}_n\}}, \mathbb{1}_{\{\bar{X}_m > \bar{Y}_n\}}\right), \left(\sqrt{n}(\bar{X}_n - \mu_1), \sqrt{m}(\bar{Y}_m - \mu_2)\right)\right]$$

$$\xrightarrow[n+m \to +\infty]{\mathcal{D}} \left[(1,0), (Z, Z)\right]$$

where $Z \sim \mathcal{N}(0,1)$.

So finally by using the appropriate mapping function,

$$\left[\sqrt{n}(\hat{\mu}_1-\mu_1), \sqrt{n}(\hat{\mu}_2-\mu_2)\right] \xrightarrow{\mathcal{D}} \left[N(0,1), N(0,1)\right]$$

c) Let us assume that $\mu_2 = \mu_1$

Hence : $\mathbb{P}(\bar{X}_m = \bar{Y}_n) \underset{n+m \to +\infty}{\longrightarrow} 1$

We also have : $\mathbb{P}(\bar{X}_m < \bar{Y}_n) \underset{n+m \to +\infty}{\longrightarrow} 0$

And then for n or m large enough :

$$\hat{\mu}_1 = \hat{\mu}_2 \approx \frac{1}{n+m}\left[\sum_{i=1}^{m} X_i + \sum_{i=1}^{n} Y_i\right]$$

So we also have (CLT) :

$$\sqrt{m+n}\,(\hat{\mu}_1 - \mu_1) \xrightarrow{\mathcal{D}} N(0,1)$$

$$\begin{cases} \sqrt{m+n}\,(\hat{\mu}_1 - \mu_1) \xrightarrow[m+n \to +\infty]{\mathcal{D}} \mathcal{N}(0, \ldots) \\ \sqrt{m+n}\,(\hat{\mu}_2 - \mu_1) \xrightarrow[m+n \to +\infty]{\mathcal{D}} \mathcal{N}(0, M) . \end{cases}$$

Problem 5

a) Let us consider:

$$\alpha_n^{(1)} = \frac{1}{n} \sum_{i=1}^{n} H_i W_i .$$

$(H_1, W_1), \ldots, (H_n, W_n)$ are iid therefore $H_1W_1, \ldots, H_nW_n$ are iid observations.

Besides, $\mathbb{E}(\alpha_n^{(1)}) = \mathbb{E}(H_1 W_1)$
$= \mathbb{E}(H_1)\,\mathbb{E}(W_1)$
$= h \cdot w$

(again because H_1 and W_1 are independently observed).

$$\mathbb{V}(\alpha_n^{(1)}) = \frac{1}{n^2} \times n \times \mathbb{V}(H_1 W_1)$$
$$= \frac{1}{n} \cdot \left[\mathbb{E}(H_1^2 W_1^2) - h^2 w^2 \right]$$
$$= \frac{1}{n} \left[\mathbb{E}(H_1^2) \cdot \mathbb{E}(W_1^2) - h^2 w^2 \right].$$

Then by using the CLT:

$$\boxed{n^{1/2}(\alpha_n^{(1)} - hw) \xrightarrow[n \to +\infty]{\mathcal{D}} \mathcal{N}(0, \operatorname{var}(H_1 W_1))}$$

b) Let us consider g such that

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R} \qquad . \quad g \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R}).$$
$$(x, y) \longmapsto xy$$

$$\forall (x,y) \in \mathbb{R}^2, \quad \nabla g(x,y) = \begin{bmatrix} y \\ x \end{bmatrix}.$$

We notice that:

$$\alpha_n^{(2)} = g(\overline{H}_n, \overline{W}_n).$$

Given the fact that $H_1, \dots, H_m$ and $W_1, \dots, W_m$ are all iid and independently observed,

- $\mathrm{var}(\bar{H}_m) = \frac{1}{m} \mathrm{var}(H_1)$
- $\mathrm{var}(\bar{W}_m) = \frac{1}{m} \mathrm{var}(W_1)$
- $\mathrm{cov}(\bar{H}_m, \bar{W}_m) = \frac{1}{m^2} \mathrm{cov}\left(\sum_i H_i, \sum_k W_k\right)$

$$= \frac{1}{m^2} m \,\mathrm{cov}(H_1, W_1)$$

$$= \frac{1}{m} \mathrm{cov}(H_1, W_1).$$

$$= 0.$$

Thus by the CLT:

$$\sqrt{m}\left((\bar{H}_m, \bar{W}_m) - (h, w) \right) \longrightarrow \mathcal{N}(0, \Sigma')$$

where

$$\Sigma' = \begin{bmatrix} \mathrm{var} H_1 & 0 \\ 0 & \mathrm{var} W_1 \end{bmatrix}$$

Then the Delta Method ensures that:

Then the delta method ensures that:

$$\sqrt{n}\Big(\alpha_n^{(2)} - \overbrace{hw}^{=g(h,w)}\Big) \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}\big(0, \nabla g^T(\theta)\Sigma_1' \nabla g(\theta)\big)$$

$$\Sigma_1' \nabla g(\theta) = \begin{bmatrix} \operatorname{var} H_1 & 0 \\ 0 & \operatorname{var} W_1 \end{bmatrix} \begin{bmatrix} w \\ h \end{bmatrix}$$

$$= \begin{bmatrix} \operatorname{var} H_1 \cdot w \\ \operatorname{var} W_1 \cdot h \end{bmatrix}$$

Then $\nabla^T g(\theta) \Sigma_1' \nabla g(\theta) = \begin{bmatrix} w & h \end{bmatrix} \begin{bmatrix} \operatorname{var} H_1 \cdot w \\ \operatorname{var} W_1 \cdot h \end{bmatrix}$

$$= \operatorname{var} H_1\, w^2 + \operatorname{var} W_1\, h^2$$

Finally:

$$\boxed{n^{1/2}\big(\alpha_n^{(2)} - hw\big) \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}\big(0, \operatorname{var} H_1 w^2 + \operatorname{var} W_1 h^2\big)}$$

c) We have:

$$\operatorname{var}(H_1 W_1) = h^2 \operatorname{var}(W_1) + w^2 \operatorname{var}(H_1)$$

$$+ \operatorname{var}(H_1) \cdot \operatorname{var}(W_1) .$$

So :

$$\operatorname{var}(\alpha_m^{(1)}) = \operatorname{var}(\alpha_m^{(2)}) + \operatorname{var} H_1 \cdot \operatorname{var} W_1$$

$$> \operatorname{var}(\alpha_m^{(2)}) .$$

$\hookrightarrow$ more accurate !

Problem 2

a) The likelihood can be written as :

$$L(X_{1:n}, \mu) = \left[\prod_{i=1}^{m} f_{\mathcal{N}(\mu,1)}(x_i) \right] \times p(m+1, n)$$

to be precised

let $i \in [\![m+1, n]\!]$:

$$\mathbb{P}(X_i \leq 0) = \mathbb{P}(x_i - \mu \leq -\mu)$$

$$= \Phi_{\mathcal{N}(0,1)}(-\mu) := \alpha_\mu .$$

And $\mathbb{P}(X_i \geqslant 0) = 1 - \Phi_{N(0,1)}(-\mu)$.

Then: $\qquad n-(m+1)+1 = n-m$

$$p(m+1,n) = \binom{n-m}{k} \alpha_\mu^{k} (1-\alpha_\mu)^{n-m-k}$$

number of negative signs among $X_{m+1},\dots,X_n$.

To keep it simple, let us write:

$$p(m+1,n) = \prod_{\substack{i=m+1\\ X_i \leqslant 0}}^{n} \alpha_\mu \times \prod_{\substack{i=m+1\\ X_i \geqslant 0}}^{n} (1-\alpha_\mu)$$

So μ^* maximizes:

$$\log\left[\prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(X_i-\mu)^2} \times \prod_{\substack{i=m+1\\ X_i \leqslant 0}}^{n} \Phi(-\mu) \right.$$

$$\left. \times \prod^{n} \left[1-\Phi(-\mu)\right] \right]$$

$$\sum_{\substack{i=m+1 \\ x_i \geq 0}}$$

$$= \sum_{i=1}^{m} \log\left((2\pi)^{-\frac{1}{2}}\right) - \frac{1}{2}(X_i - \mu)^2 +$$

$$\sum_{\substack{i=m+1 \\ x_i \leq 0}}^{n} \log(\Phi(-\mu)) + \sum_{\substack{i=m+1 \\ X_i \geq 0}}^{n} \log(1 - \Phi(-\mu)) \Big)$$

So μ^* satisfies:

$$-\sum_{i=1}^{m}(X_i - \mu^*) - \sum_{\substack{i=m+1 \\ x_i < 0}}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mu^{*2}} \times \frac{1}{\Phi(-\mu^*)}$$

$$+ \sum_{\substack{i=m+1 \\ X_i \geq 0}}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^{*2}}{2}} \times \frac{1}{1 - \Phi(-\mu^*)} = 0.$$

$$\Rightarrow \sum^{m}(X_i - \mu^*) + \sum^{n} \ldots (X_i) \frac{e^{-\frac{\mu^{*2}}{2}}}{} \ldots \frac{1}{} = 0$$

$$\Rightarrow \sum_{i=1} (X_i - \mu) + \sum_{i=m+1} \operatorname{sgn}(X_i) \frac{}{\sqrt{2\pi}} \frac{}{\Phi(\operatorname{sign} X_i \mu^*)} = 0$$

b) let us assume that:

$$\frac{m}{n} \longrightarrow \alpha \in [0,1] \text{ as}$$

$m \to +\infty$ and $n \to +\infty$.

- $$\frac{1}{m} \sum_{i=1}^{m} (X_i - \mu^*) \xrightarrow[\substack{n \to +\infty \\ m \to +\infty}]{\mathcal{D}} \mathcal{N}(0,1).$$

- $$\frac{1}{n} \sum_{i=m+1}^{n} \operatorname{sgn}(X_i) \frac{e^{-\frac{\mu^*}{2}}}{\sqrt{2\pi}} \cdot \frac{1}{\Phi(\operatorname{sign} X_i \mu^*)}$$

$$= \frac{1}{n-m} \times (1-c) \sum_{i=m+1}^{n} \operatorname{sgn}(X_i) \frac{e^{-\frac{1}{2}\mu^{*2}}}{\sqrt{2\pi}} \cdot \frac{1}{\Phi(\operatorname{sgn} X_i \mu^*)}$$

(for n and m large enough)

$$\xrightarrow[n \to +\infty]{\mathcal{D}} \mathcal{N}\left(0, (1-c)^2 \cdot \frac{e^{-\mu^{*2}}}{2\pi} \left[\frac{1}{\Phi(\mu^*)} + \frac{1}{\Phi(-\mu^*)}\right]\right)$$

$\mu \to +\infty$

$\mu \to +\infty$

Problem 6

a) $f(\rho^* | X) \propto f(X|\rho^*) \, f(\rho^*)$

posterior distribution — likelihood function — prior distribution

$$f(X|\rho^*) = f(X_1|\rho^*) \times \prod_{i=2}^{n} f(X_i | X_{i-1}, \rho^*)$$

$$= f(X_1|\rho^*) \cdot \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{\frac{n-1}{2}} \cdot e^{-\frac{1}{2\sigma_\varepsilon^2}\sum_{i=2}^{n}(X_i - \rho^* X_{i-1} - 2)^2}$$

So finally :

$$f(\rho^* | X) \propto e^{-\frac{1}{2\sigma_\varepsilon^2}\sum_{i=2}^{n}(X_i - \rho^* X_{i-1} - 2)^2 - \frac{1}{2}(\rho^* - \mu)^2}$$

$f(\varphi^* \mid X) \propto e$

Then:

$$-\frac{1}{2\sigma_\varepsilon^2}\sum_{i}^{n}{}'\left(X_i-\varphi^* X_{i-1}-2\right)^2-\frac{1}{2}\left(\varphi^*-\mu\right)^2$$

$$=-\frac{1}{2}\left[\frac{1}{\sigma_\varepsilon^2}\sum_{i}{}'\left(X_i-2\right)^2-2X_{i-1}\left(X_i-2\right)\varphi^* + X_{i-1}^2\varphi^{*2}\right]-\frac{1}{2}\left[\varphi^{*2}-2\mu\varphi^*+\mu^2\right]$$

$$=-\frac{1}{2}\Bigg[\left(1+\frac{1}{\sigma_\varepsilon^2}\sum_{i=2}^{n}{}'X_{i-1}^2\right)\varphi^*$$

$$-2\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=2}^{n}{}'X_{i-1}\left(X_i-2\right)+\mu\right)\varphi^*$$

$$+\frac{1}{\sigma_\varepsilon^2}\sum_{i}{}'\left(X_i-2\right)^2+\mu^2\Bigg]$$

Let us note :

$$\boxed{\Sigma := \left[1 + \frac{1}{\sigma_\varepsilon^2}\sum_{i=2}^{m} X_{i-1}^2\right]^{-1}}$$

$$\boxed{M := \frac{\frac{1}{\sigma_\varepsilon^2}\sum_{i=2}^{m} X_{i-1}(X_i - \alpha) + \mu}{1 + \frac{1}{\sigma_\varepsilon^2}\sum_{i=2}^{m} X_{i-1}^2}}$$

Then we have

$$f(\varphi^+ | X) \propto e^{-\frac{1}{2\Sigma}(\varphi^+ - M)^2}$$

So we guess that :

$$\varphi^+ | X \sim \mathcal{N}(M, \Sigma)$$

b) We have :

$$X_{23} = \rho^*(\overbrace{\rho^* X_{21} + 2 + \varepsilon_{22}}^{= X_{22}}) + 2 + \varepsilon_{23}$$

$$= \rho^{*2}\left[\rho^* X_{20} + 2 + \varepsilon_{21}\right]$$

$$+ \ 2(1+\rho^*) + \varepsilon_{23} + \rho^*\varepsilon_{22}$$

$$\boxed{= \rho^{*3} X_{20} + 2(1+\rho+\rho^{*2}) + \varepsilon_{23} + \rho^*\varepsilon_{22} + \rho^{*2}\varepsilon_{21}}$$

- In order to compute $\mathbb{P}(X_{23} > 4)$, I chose to generate a sequence $X_{23}^{(1)}, \dots, X_{23}^{(n)}$ with $n \in \mathbb{N}^*$ and then use a Monte-Carlo estimator :

$$\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{X_{23}^{(i)} \geq 4\}$$

$$\longrightarrow \mathbb{P}(X_{23} > 4)$$

$$\xrightarrow[n \to +\infty]{} \mathbb{P}(X_{23} = 1) \quad a.s.$$

- Similarly, the best estimator of $\hat{X}_{23}$ is

$$\boxed{\frac{1}{n} \sum_{i=1}^{n} X_{23}^{(i)}}$$