LMAFY1101 - Solutions - Série 5

Variables aléatoires: Quelques lois usuelles
PARTI I

Variables aléatoires discrètes

Exercice 1

Soit X le nombre des réponses correctes. $X \sim Bin(20, 0.2)$

1.

$$P(X = 0) = 0.8^{20} = 0.012$$

2.

$$P(X = 20) = 0.2^{20} = 1.049 \times 10^{-14}$$

3.

$$P(X \geq 13) = \sum_{x=13}^{20} C_{20}^{x} 0.2^x 0.8^{20-x}$$

```
pbinom(12, 20, 0.2, lower.tail = FALSE)
```

```
[1] 1.52e-05
```

Exercice 2

Soit X = nbr. de contraventions à payer. $X \sim Bin(5, 0.1)$.

$$\begin{aligned} P(25X > 25) &= P(X > 1) = 1 - P(X = 0) + P(X = 1) \\ &= 1 - 0.9^5 - 5 \times 0.1 \times 0.9^4 \\ &= 0.08146. \end{aligned}$$

Excercice 3

1.

$$\begin{aligned} P(X \geq 7.5) = 1 - F(7.5) &= \exp(-\frac{1}{2}) \\ &= 0.61 \end{aligned}$$

2.

On cherche x t.q.

$$F(x) = 0.4 \Leftrightarrow 1 - \exp(-\frac{x}{15}) = 0.4 \Leftrightarrow x = 7.66$$

3.

Soit Y le nombre d'étudiants qui joue plus de $x = 7.66$ heures par semaine aux jeux vidéo. On a que $Y \sim Bin(n = 20, p = 0.6)$, $E(Y) = 20 \times 0.6 = 12$ et $Var(Y) = 20 \times 0.6 \times 0.4 = 4.8$.

Excercice 4

1.

Soit $X_H =$ Nombre d'hommes qui guérissent $\sim Bin(70, 0.6)$. On cherche $P(X_H \geq 36)$.

```
pbinom(35, 70, 0.6, lower.tail = FALSE)
```

```
[1] 0.943
```

2.

Soit $X_F =$ Nombre de femmes qui guérissent $\sim Bin(80, 0.7)$. On cherche $P(X_H > X_F)$

$$\begin{aligned} P(X_H > X_F) &= \sum_{x=0}^{80} P(X_H > x | X_F = x) P(X_F = x) \\ &= \sum_{x=0}^{80} P(X_H > x) P(X_F = x) \end{aligned}$$

```
x <- 0:80
pH <- pbinom(x, 70, 0.6, lower.tail = FALSE)
dF <- dbinom(x, 80, 0.7)
sum(pH * dF)
```

```
[1] 0.00633
```

3.

Soit $X_T = X_H + X_F =$ Nombre total d'individus qui guérissent.

Il faut calculer $P(X_T = 90)$. Cette probabilité est donnée par

$$P(X_H + X_F = 90) = \sum_{x=0}^{90} P(X_H = x) P(X_F = 90 - x)$$

```r
x <- 0:90
dH <- dbinom(x, 70, 0.6)
dF <- dbinom(90 - x, 80, 0.7)
p90 <- sum(dH * dF)
p90
```

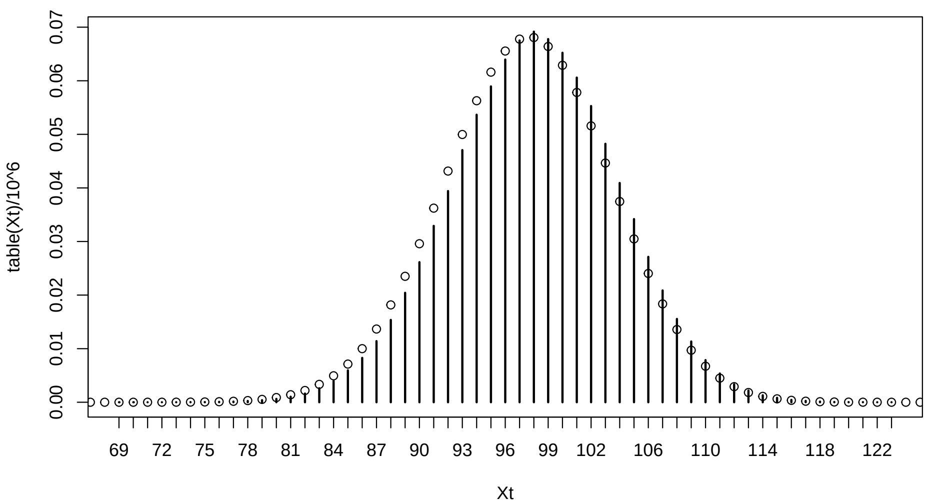
```
[1] 0.0263
```

Remarque

On'est tenté de dire que $X_T \sim Bin(150, 0.65)$ mais cela est faux, car pour que la somme de deux binomiales (ici X_H et X_F) soit une binomiale il faut que les deux variables partagent la même probabilité de succès, or cela n'est pas le cas de X_H et X_F. Rappelez-vous que, par définition d'une binomiale, la probabilité de succès ne doit pas changer d'une expérience à une autre.

Nous pouvons aussi vérifier cela via des simulations.

```r
Xh <- rbinom(10^6, 70, 0.6)
Xf <- rbinom(10^6, 80, 0.7)
Xt <- Xh + Xf
plot(table(Xt)/10^6)
points(0:150, dbinom(0:150, 150, 0.65))
```



4.

Il faut calculer $P(X_H \geq 46 | X_T = 90)$.

$$
\begin{aligned}
P(X_H \geq 46 | X_T = 90) &= \sum_{x=46}^{70} \frac{P(X_H = x, X_T = 90)}{P(X_T = 90)} \\
&= \frac{1}{P(X_T = 90)} \sum_{x=46}^{70} P(X_H = x) P(X_F = 90 - x)
\end{aligned}
$$

```
x <- 46:70
dH <- dbinom(x, 70, 0.6)
dF <- dbinom(90 - x, 80, 0.7)
sum(dH * dF)/p90
```

```
[1] 0.00616
```

Excercice 5

1.

On définit la variable aléatoire Y = nombre de clients entrant dans le magasin chaque minute.
Dans ce contexte, $Y \sim \text{Poisson}(\lambda = 1)$. On cherche $P(Y > 3)$.

```
ppois(3, 1, lower.tail = FALSE)
```

```
[1] 0.019
```

2.

On définit la variable aléatoire N = nombre de fois, parmi 10, où le seuil de 3 individus est dépassé. On a $N \sim Bin(n = 10, p)$, avec $p = P(Y > 3) = 0.019$. On cherche $P(N \geq 1)$.

```
pbinom(0, 10, 0.019, lower.tail = FALSE)
```

```
[1] 0.175
```

Excercice 6

Il suffit de remarquer que

$$E(I) = eE(N) = \lambda e$$
$$Var(I) = e^2 Var(N) = \lambda e^2$$

Excercice 7

Soit X le nombre de sinistres au cours d'une année. Puisque $X \sim Bin(5000, 1/10000) \approx Pois(0.5)$, on peut calculer les probabilités sur X en utilisant la Binomial ou la Poisson. Ici nous allons utiliser la Poisson.

$$
\begin{aligned}
P(X \geq 3) &= 1 - P(X \leq 2) \approx 1 - (1 + 0.5 + 0.5^2/2)\exp(-0.5) \\
&\approx 0.0144 \\
P(X \geq 4) &= 1 - P(X \leq 3) \approx 1 - (1 + 0.5 + 0.5^2/2 + 0.5^3/6)\exp(-0.5) \\
&\approx 0.00175
\end{aligned}
$$

Ces calculs montrent que la compagnie d'assurance sera (très) bien protégée avec (4) 3 millions d'euros de réserve.

Variables aléatoires continues

Excercice 8

On note X la v.a. désignant l'heure de visite de Caroline. $X \sim Unif(18.5, 20.75)$

$$P(19 \leq X \leq 19.5) = \frac{19.5 - 19}{20.75 - 18.5} = 2/9$$

Excercice 9

On note X la v.a. désignant la durée de la communication téléphonique entre Claire et Alice. $X \sim Unif(0, 1)$.

1.

$P(X = 0.5) = 0$

2.

$P(X \leq 0.75) = 0.75$

3.

$P(X \geq 1/6) = 1 - 1/6 = 5/6$

4.

$P(1/3 \leq X \leq 2/3) = 2/3 - 1/3 = 1/3$

Excercice 10

1.

$$
\begin{aligned}
E(4\sqrt{1-X^2}) &= \int_{-\infty}^{\infty} 4\sqrt{1-x^2} f(x)dx \\
&= 4\int_0^1 \sqrt{1-x^2}dx = \text{ l'aire d'un disque de rayon 1} \\
&= 1^2 \times \pi = \pi
\end{aligned}
$$

2.

On peut approximer $E(4\sqrt{1-X^2})$ par la moyenne empirique: $\frac{1}{n}\sum_{i=1}^{n} 4\sqrt{1-x_i^2}$, où les x_i sont des observations qui proviennent d'une variable de distribution $Unif[0,1]$. Pour calculer une telle moyenne en R, il suffit de taper

```
x <- runif(10^6)
mean(4 * sqrt(1 - x^2))
```

```
[1] 3.14
```

Excercice 11

$X \sim N(90, 12^2)$

1.

On cherche $P[80 \leq X \leq 100] = P(X \leq 100) - P(X < 80)$.

```
pnorm(100, 90, 12) - pnorm(80, 90, 12)
```

```
[1] 0.595
```

2.

On cherche $P[X \geq 120]$

```
pnorm(120, 90, 12, lower.tail = FALSE)
```

```
[1] 0.00621
```

4.

On cherche x tel que $P[X \geq x] = 0.2$

```r
qnorm(0.8, 90, 12)
```

```
[1] 100
```

```r
# ou
qnorm(0.2, 90, 12, lower.tail = FALSE)
```

```
[1] 100
```

Excercice 12

Soit X la durée de vie. $X \sim N(10, 12.5)$.
On cherche le plus grand x tel que $P(X \leq x) \leq p$, ce qui revient à chercher x tel que $P(X \leq x) = p$ (pourquoi ?).

- Pour $p = 0.04$

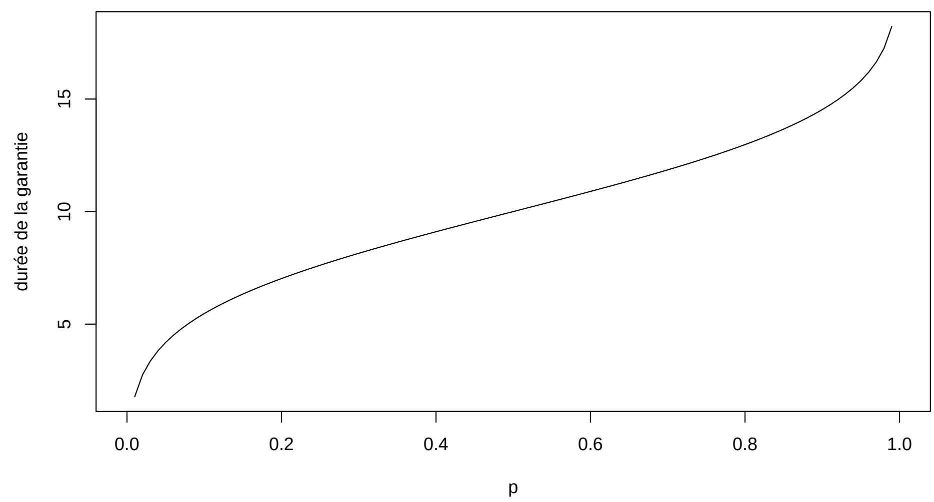
```r
qnorm(0.04, mean = 10, sd = sqrt(12.5))
```

```
[1] 3.81
```

$\rightarrow$ 4 ans de garanties.

- Pour p quelconque

```r
curve(qnorm(x, mean = 10, sd = sqrt(12.5)), from = 0, to = 1,
  xlab = "p", ylab = "durée de la garantie")
```



Excercice 13

Soit $Y = X_2 - X_1 \sim N(1.5 - 1.48, 0.0009 + 0.0016) = N(0.02, 0.0025)$. On cherche $P(Y < 0)$.

```
pnorm(0, 0.02, sqrt(0.0025))
```

```
[1] 0.345
```

Excercice 14

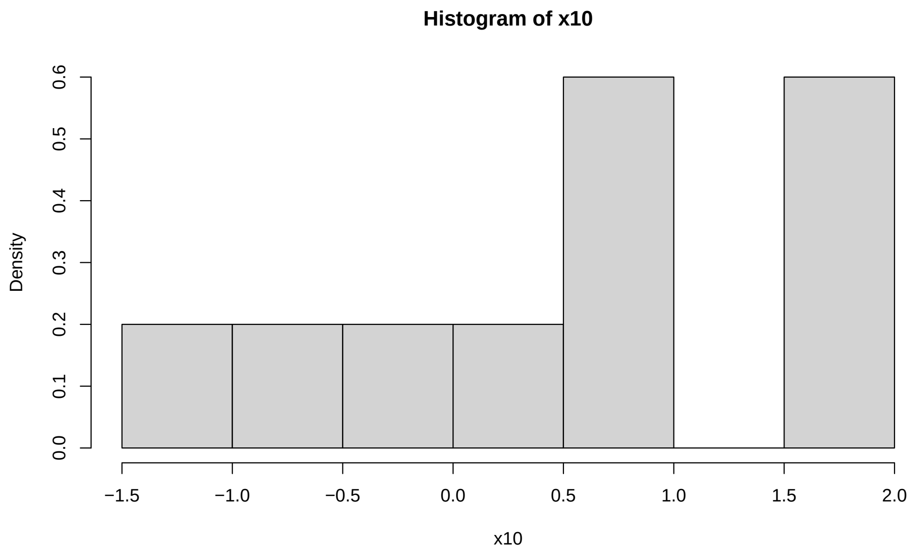
1.

```
x10 <- rnorm(10)
x10
```

```
 [1]  0.217 -0.542  0.891  0.596  1.636  0.689 -1.281 -0.213  1.897  1.777
```

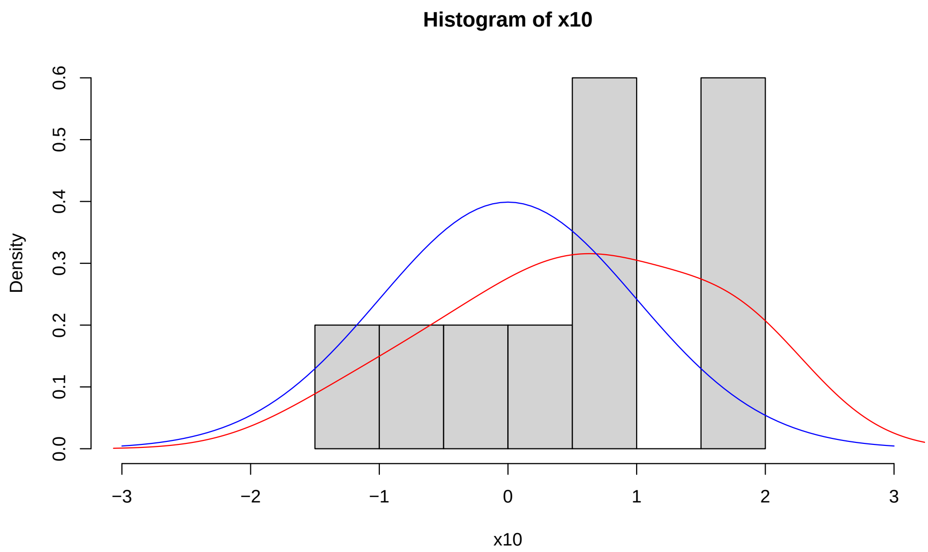
2.

```
hist(x10, freq = FALSE)
```



3.

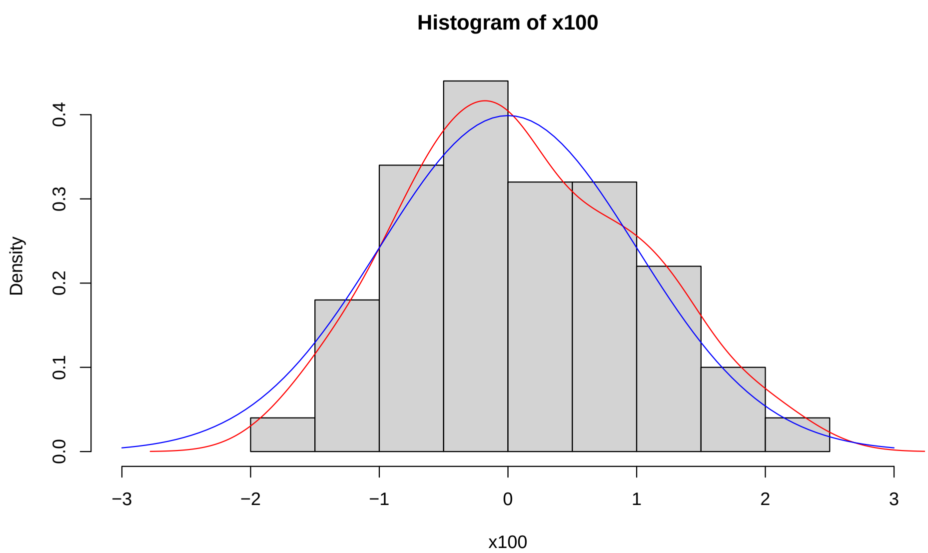
```
hist(x10, freq = FALSE, xlim = c(-3, 3))
lines(density(x10), col = "red")
curve(dnorm, from = -3, to = 3, add = TRUE, col = "blue")
```



4.

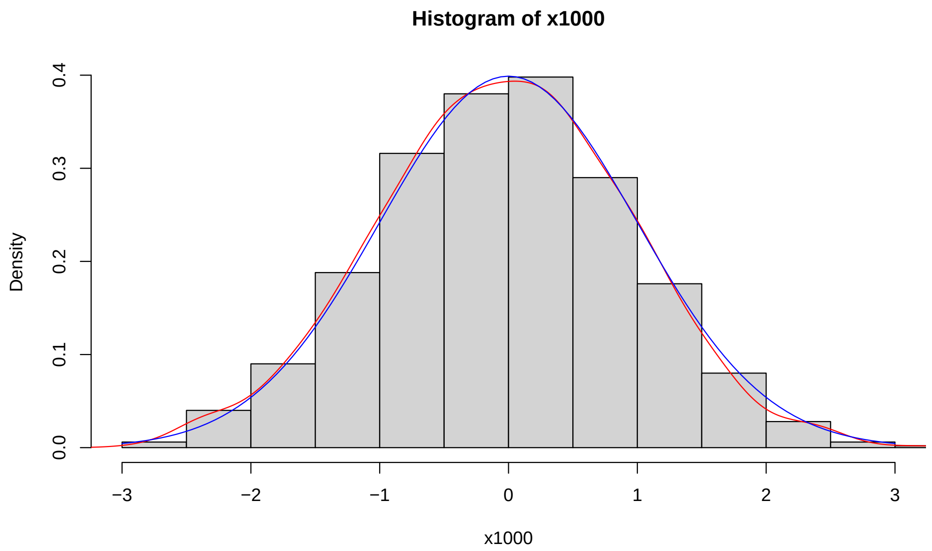
- $n = 100$

```
x100 <- rnorm(100)
hist(x100, freq = FALSE, xlim = c(-3, 3))
lines(density(x100), col = "red")
curve(dnorm, from = -3, to = 3, add = TRUE, col = "blue")
```



- $n = 1000$

```
x1000 <- rnorm(1000)
hist(x1000, freq = FALSE, xlim = c(-3, 3))
density(x1000) |> lines(col = "red")
curve(dnorm, from = -3, to = 3, add = TRUE, col = "blue")
```



L'estimation de la densité qu'on obtient avec la fonction `density` deviens de plus en plus précise lorsque n augmente.

Excercice 15

1.

Soit $W = X - Y \sim N(0, 10)$.

$$\begin{aligned} P(W^2 > 5) &= P(|W| > \sqrt{5}) \\ &= P(W > \sqrt{5})) + P(W < -\sqrt{5}) \\ &= 2P(W < -\sqrt{5}) \end{aligned}$$

```
2 * pnorm(-sqrt(5), 0, sqrt(10))
```

```
[1] 0.48
```

2.

```
x <- rnorm(10^6, 0, 1)
y <- rnorm(10^6, 0, sqrt(9))
mean((x - y)^2 > 5)
```

```
[1] 0.48
```

Excercice 16

1.

```r
X <- rnorm(10^6, 0, 6)
Y <- rnorm(10^6, 0, 6)
D <- sqrt(X^2 + Y^2)
mean(D <= 10)
```

```
[1] 0.75
```

2.

```r
quantile(D, c(0.25, 0.5, 0.75))
```

```
  25%   50%   75%
 4.55  7.07 10.00
```