TP d'introduction à R

Evan Voyles

08 Avril, 2022

Exercice 1 Loi d'une variable aléatoire discrète.
Soit X ~ $\mathcal{B}$(5, 0.6).

1. Calculer $\mathbb{P}(X = 0)$ et $\mathbb{P}(X = 1)$ avec R et vérifier par le calcul.

```
p <- 0.6
n <- 5

p_x_eq_0 <- dbinom(x = 0, size = n, prob = p)
p_x_eq_1 <- dbinom(x = 1, size = n, prob = p)
```

Nous savons que $X = 0$ est le cas si on échoue 5 fois d'affilées. Donc, $\mathbb{P}(X = 0) = (1 - 0.6)^5 = 0.01024$.

```
0.4 ^ 5
```

```
## [1] 0.01024
```

ce qui est égal à la valeur stocké dans `p_x_eq_0`:

```
p_x_eq_0
```

```
## [1] 0.01024
```

Pour $X = 1$, il s'agit d'une réussite et quatre échecs; $\mathbb{P}(X = 1) = C(5, 1) * (0.4^4)(0.6) = 0.0768$

```
choose(n = 5, k = 1) * (0.4 ^ 4) * (0.6)
```

```
## [1] 0.0768
```

ce qui est égal à la valeur stocké dans `p_x_eq_1`.

```
p_x_eq_1
```

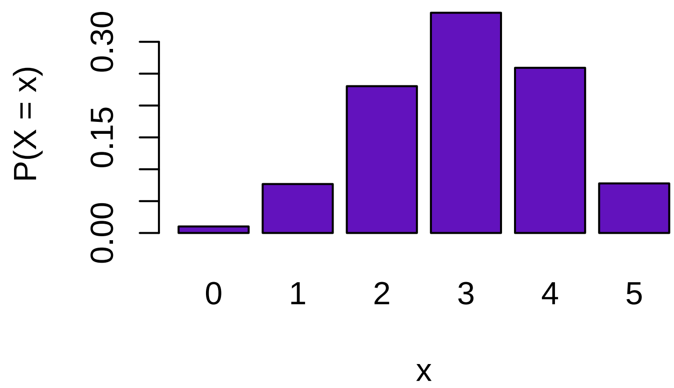
```
## [1] 0.0768
```

```
library(purrr)

X <- seq(0, 5)
P_x <- dbinom(X, 5, 0.6)

names(P_x) <- X # Associate x with P(X = x)

p <- barplot(P_x, col = "#6212bd", beside = TRUE)
title(ylab = "P(X = x)", xlab = "x")
```



3. Calculer $\mathbb{P}(X <= 1)$ et $\mathbb{P}(X <= 4)$ avec R.

```
p <- 0.6
n <- 5

pbinom(1, size = n, prob = p)
```

```
## [1] 0.08704
```

```
pbinom(4, size = n, prob = p)
```

```
## [1] 0.92224
```

```
P_x[1] + P_x[2] # P(X = 0) + P(X = 1)
```

```
## [1] 0.08704
```

```
1 - P_x[6] # 1 - P(X = 5)
```

```
## [1] 0.92224
```

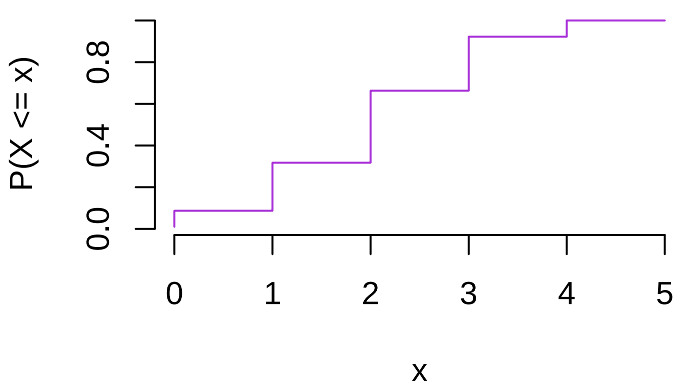
4. Représenter le graphe de la fonction de répartition de X.

```
X <- seq(0, 5)
cdf <- pbinom(X, size = 5, prob = 0.6)

cdf
```

```
## [1] 0.01024 0.08704 0.31744 0.66304 0.92224 1.00000
```

```
plot(X, cdf, bty = "n", type = "S", col = "#a830d8", ylab = "P(X <= x)", xlab = "x")
```



5. Déterminer le réel

$$q = \inf\{k, \mathbb{P}(X \leq k) \geq 0.25\}$$

```r
pred <- function(k) { pbinom(k, 5, 0.6) >= 0.25 }
q <- min(keep(X, pred))
```

On vient de calculer la 0.25 quantile de X pour une loi **discrète**.

```r
qbinom(0.25, size = 5, prob = 0.6)
```

```
## [1] 2
```

```r
q
```

```
## [1] 2
```

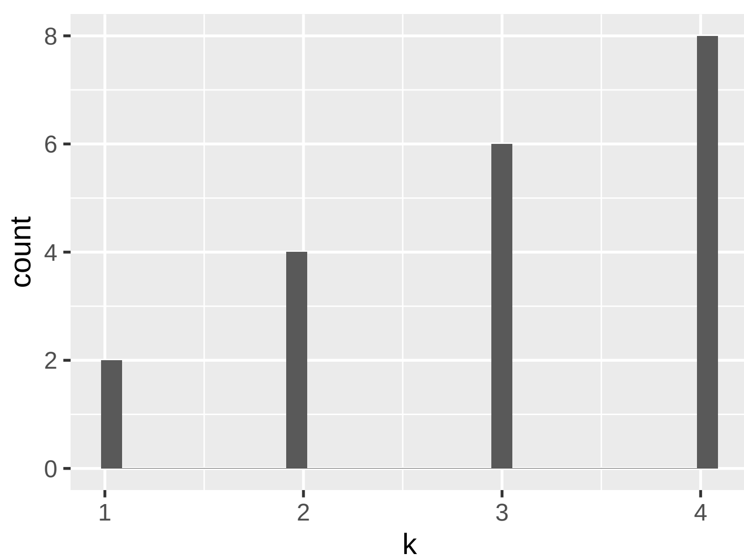
6. Simuler 20 réalisations de X.

```r
sim <- rbinom(20, size = 5, prob = 0.6)

df <- tibble(sim)

df |> ggplot(aes(sim)) + geom_histogram() + labs(x = "k")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

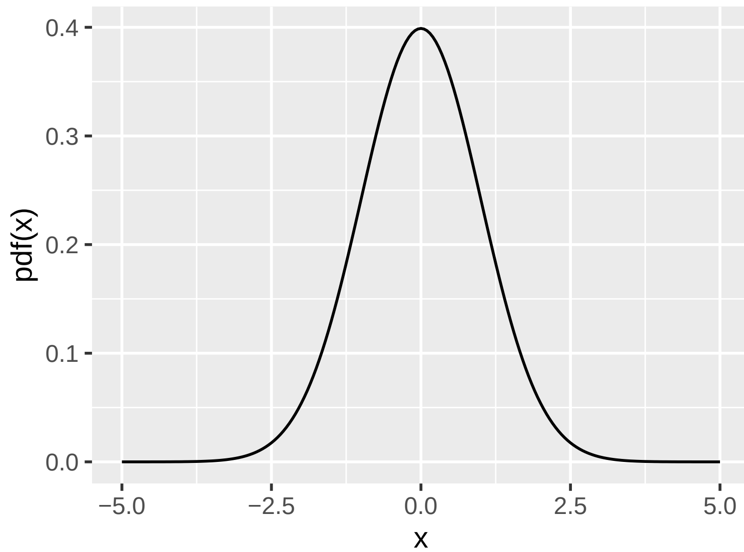


Exercice 2 Loi d'une variable aléatoire continue.

1. Soit $U \sim \mathcal{N}(0, 1)$
   (a) Représenter le graphe de la densité de U entre -5 et 5.

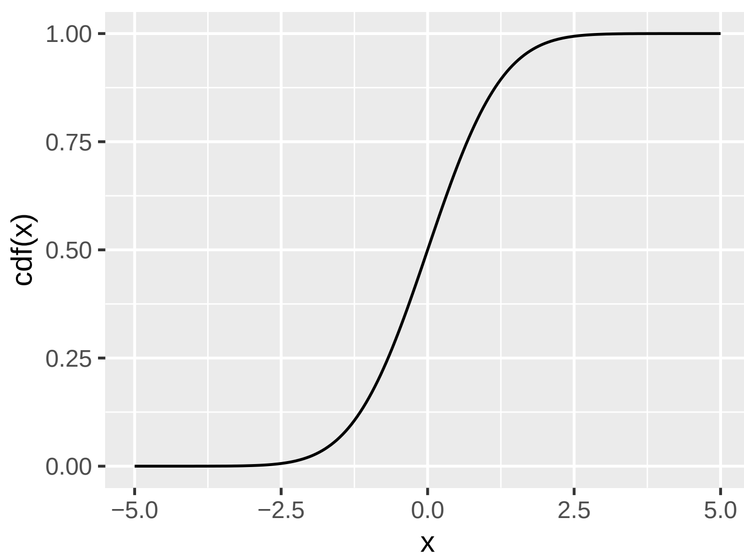
```r
x <- seq(-5, 5, 0.05)
pdf <- dnorm(x)
cdf <- pnorm(x)

df <- tibble(x, pdf, cdf)
df |> ggplot(aes(x, pdf)) + geom_line() + labs(y = "pdf(x)")
```



(b) Représenter le graphe de la fonction de répartition de U.

```
df |> ggplot(aes(x, cdf)) + geom_line() + labs(y = "cdf(x)")
```



(c) Retrouver avec R, les valeurs lues en TD dans la table de la loi normale centrée réduite.

- $\mathbb{P}(U \leq 0.11)$

```
pnorm(0.11)
```

```
## [1] 0.5437953
```

- $\mathbb{P}(U \leq -0.51)$

```
pnorm(-0.51)
```

```
## [1] 0.3050257
```

- $\mathbb{P}(0.5 \leq U \leq 1.5)$

```
pnorm(1.5) - pnorm(0.5)
```

```
## [1] 0.2417303
```

et retrouver les quantiles q_1, q_2, q_3 tels que

- $\mathbb{P}(U \leq q_1) = 0.975$

```
qnorm(0.975)
```

```
## [1] 1.959964
```

- $\mathbb{P}(U \leq q_2) = 0.2358$

```
qnorm(0.2358)
```

```
## [1] -0.7198782
```

- $\mathbb{P}(|U| \leq q_3) = 0.33$

On cherche donc un intervalle centrée autour de 0 qui contient 33% de la densité de la loi normale. $p_{gauche} = \frac{1-0.33}{2}$

```
p_gauche <- (1 - 0.33) / 2
q3 <- -qnorm(p_gauche)
q3
```

```
## [1] 0.426148
```

```
# Ici, on verifie le calcul
pnorm(q3) - pnorm(-q3)
```

```
## [1] 0.33
```

2. Soit $X \sim \mathcal{N}(3, (0.2)^2)$

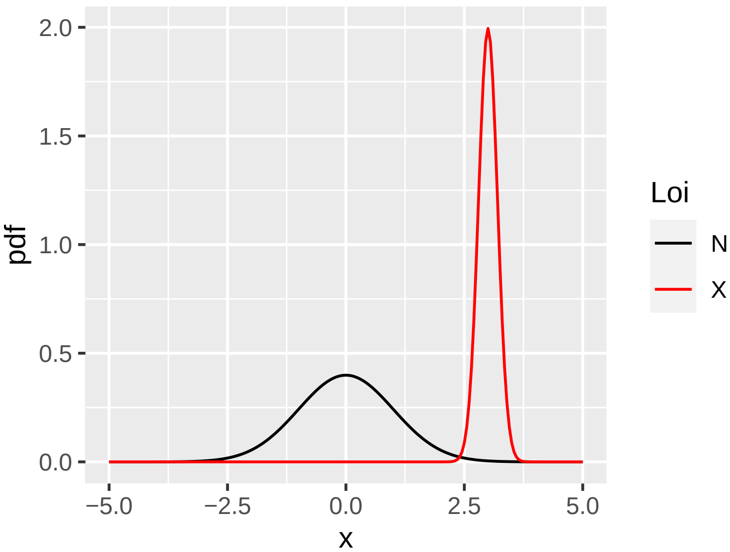
    (a) Superposer (en rouge) le graphe de la densité de X sur celui de la densité de U.

```
library(tidyverse)

x <- seq(-5, 5, 0.05)
pdf_n01 <- dnorm(x)
pdf_n304 <- dnorm(x, 3, 0.2)

df <- tibble(x, pdf_n01, pdf_n304)
df <- df |> pivot_longer(c(pdf_n01, pdf_n304), names_to = "curve", values_to = "pdf")

df |> ggplot(aes(x, pdf)) + geom_line(aes(col = curve)) +
    scale_color_manual(values = c("black", "red"), labels = c("N", "X")) +
    labs(col = "Loi")
```



(b) Retrouver avec R, $\mathbb{P}(X \leq 3.29)$ et a tel que $\mathbb{P}(X \leq a) = 0.975$

```
pnorm(3.29,  mean = 3, sd = 0.2) # P(X <= 3.9)
```

```
## [1] 0.9264707
```

```
qnorm(0.975, mean = 3, sd = 0.2) # a, tq P(X <= a) = 0.975
```

```
## [1] 3.391993
```

Exercice 3 Approximation de la loi Binomiale par ... La loi Binomial $\mathcal{B}(n, p)$ peut-être approchée lorsque $n \to \infty$ par la loi de Poisson ou par la noi normale en fonction des valeurs de p.

1. Approximation par la loi normale lorsque p est fixé et $n \to \infty$.

   (a) Ecrire une fonction `approx_bin_normale` qui prend pour arguments les paramètres n et p et qui renvoie sur la même figure, le diagramme en bâton de la loi $\mathcal{B}(n, p)$ et la densité de la loi $\mathcal{N}(np, np(1-p))$.

```
library(pracma)
```

```
# Approximation of the binomial distribution with size n and probability of success p
# by a normal distribution
approx_bin_normale <- function(n, p) {

    mu  <- n * p
    sig <- sqrt(n * p * (1 - p))

    # We want the values from 0 to n
    # x <- linspace(0, n) #
    k <- seq(0, n)
    # x <- linspace(0, n, 1000)
    x <- k

    pdf.norm <- dnorm(x, mean = mu,  sd = sig)
    pdf.bin  <- dbinom(x = k, size = n, prob = p)

    # names(pdf.bin) <- k

    # df.bar <- barplot(pdf.bin)
```

```
    # lines(x = df.bar, pdf.norm)
    # points(pdf.norm)

    # df.bin <- tibble(k, pdf = pdf.bin)
    # df.norm <- tibble(x, pdf = pdf.norm)

    # plt <- df.bin |> ggplot(aes(k, pdf.bin)) + geom_col()
    # print(length(df.norm$x))
    # plt + geom_line(aes(data = df.norm, x, pdf.norm))
    # p <- df.norm |> ggplot(aes(x, pdf)) + geom_line()
    # p + geom_col(aes(x = df.bin$k, y = df.bin$pdf))
}
```