SOLUTION TP n° 7

Solution 1. On s'intéresse à la durée de vie d'un certain type de voiture. Soit X la *var* égale à la durée de vie en années d'une de ces voitures. On suppose que X suit la loi exponentielle $\mathcal{E}(\frac{1}{10})$, *i.e.* de densité

$$f(x) = \begin{cases} \frac{1}{10} e^{-\frac{x}{10}} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

1. Écrire une nouvelle fonction R équivalente à `dexp` *(on pourra utiliser une structure de contrôle* `if ... else` *pour que celle-ci soit égale à 0 quand $x < 0$)* :

```
densexp = function(x, lambda) {
if (x<0)
0
else
lambda * exp(-lambda * x)
}
```

2. Vérifier numériquement que $\int_0^{100} f(x)dx \simeq 1$:

```
integrate(function(x) dexp(x, 1 / 10), 0, 100)
```

3. Séparer l'écran graphique en 2: dans la fenêtre 1, représenter le graphe de la densité f et, dans la fenêtre 2, celui de la fonction de répartition de X :

```
par(mfrow = c(1, 2))
curve(dexp(x, 1 / 10), 0, 15, ylab = "Densité")
curve(pexp(x, 1 / 10), 0, 15, ylab = "Fonction de répartition")
```

4. Calculer la probabilité que la durée de vie d'une voiture dépasse 10.2 ans :

`1 - pexp(10.2, 1 / 10)` renvoie : `[1] 0.3605949`

Solution 2. Soit X une *var* suivant la loi gamma $\Gamma(5, 1)$, *i.e.* de densité

$$f(x) = \begin{cases} \frac{1}{24} x^4 e^{-x} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

1. Écrire une nouvelle fonction R équivalente à `dgamma` :

```
densgamma = function(x, m, lambda) {
if (x<0)
0
else
(1 / factorial(m - 1)) * lambda^m * x^(m - 1) * exp(-lambda * x)
}
```

2. Vérifier numériquement que $\int_0^{100} f(x)dx \simeq 1$:

```
integrate(function(x) dgamma(x, 5, 1), 0, 100)
```

3. Évaluer la fonction de répartition de X, *i.e.* $F(x)$, pour $x \in \{2, \ldots, 10\}$:

```
x = 2:10
pgamma(x, 5, 1)
```

renvoie :

```
[1] 0.05265302  0.18473676  0.37116306  0.55950671  0.71494350  0.82700839
[8] 0.90036760  0.94503636  0.97074731
```

qui correspondent à $F(2), \ldots, F(10)$ respectivement.

4. Déterminer le réel x vérifiant $\mathbb{P}(X \leq x) = 0.92$:

`qgamma(0.92, 5, 1)` renvoie : `[1] 8.376739`

Solution 3. Soit X une *var* suivant la loi normale $\mathcal{N}(0, 1)$, *i.e.* de densité

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}.$$

1. Écrire une nouvelle fonction R équivalente à la fonction `dnorm` :

```
densnorm = function(x) {
(1/sqrt(2 * pi)) * exp(-x^2 / 2)
}
```

2. Vérifier numériquement que $\int_{-100}^{100} f(x)dx \simeq 1$:

```
integrate(function(x) dnorm(x), -100, 100)
```

3. Séparer l'écran graphique en 2: dans la fenêtre 1, représenter le graphe de la densité f et, dans la fenêtre 2, celui de la fonction de répartition de X :

```
par(mfrow = c(1, 2))
curve(dnorm(x), -5, 5, ylab = "Densité")
curve(pnorm(x), -5, 5, ylab = "Fonction de répartition")
```

4. Calculer $\mathbb{P}(X \leq 2.2)$, $\mathbb{P}(X \geq 1.7)$, $\mathbb{P}(0.2 \leq X < 1.4)$ et $\mathbb{P}(|X| \leq 1.96)$:

   - $\mathbb{P}(X \leq 2.2)$:
     `pnorm(2.2)` renvoie : `[1] 0.9860966`
   - $\mathbb{P}(X \geq 1.7)$:
     `1 - pnorm(1.7)` renvoie : `[1] 0.04456546`
   - $\mathbb{P}(0.2 \leq X < 1.4)$:
     `pnorm(1.4) - pnorm(0.2)` renvoie : `[1] 0.3399836`

- $\mathbb{P}(|X| \leq 1.96) = 2\mathbb{P}(X \leq 1.96) - 1$:
  `2 * pnorm(1.96) - 1` renvoie : `[1] 0.9500042`

5. Déterminer le réel x vérifiant $\mathbb{P}(X \leq x) = 0.98$:
   `qnorm(0.98)` renvoie : `[1] 2.053749`

Solution 4. On suppose que le poids d'un foie gras est une *var* X suivant la loi normale $\mathcal{N}(550, 100^2)$, l'unité étant le gramme. Quelle est la probabilité qu'un foie gras pèse

- moins de 650 grammes :
  `pnorm(650, 550, 100)` renvoie : `[1] 0.8413447`
- plus de 746 grammes ?
  `1 - pnorm(746, 550, 100)` renvoie : `[1] 0.0249979`
  (ou `pnorm(746, 550, 100, lower.tail=F)`).
- entre 550 grammes et 600 grammes :
  `pnorm(600, 550, 100) - pnorm(550, 550, 100)` renvoie : `[1] 0.1914625`

Solution 5. Soit X une *var* suivant la loi normale $\mathcal{N}(0, 1)$. La fonction de répartition de X, *i.e.* $F(x) = \mathbb{P}(X \leq x)$, avec $x \geq 0$, peut s'écrire comme

$$F(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \sum_{n=0}^{\infty} \frac{x^{2n+1}}{\prod_{m=0}^{n}(2m+1)}.$$

Utiliser cette expression pour écrire une nouvelle fonction R équivalente à la fonction `pnorm` *(on tronquera la somme infinie à la valeur* 50*)* :

```
foncrep = function(x) {
n = 1 + 2 * 0:50
0.5 + (1/sqrt(2 * pi)) * exp(-x^2 / 2) * sum(x^n / cumprod(n))
}
```

Par exemple, `foncrep(1.2)` renvoie : `[1] 0.8849303`, ce qui correspond exactement au résultat de `pnorm(1.2)`.

Solution 6. Commenter le résultat des commandes suivantes :

```
curve(dweibull(x, 2, 2), -5, 5, col = "red")
```

renvoie un graphique où est tracé en rouge la densité associée à la loi de Weibull $\mathcal{W}(2, 2)$, *i.e.*

$$f(x) = \begin{cases} \frac{x}{2} e^{-(\frac{x}{2})^2} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

```
curve(dt(x, 3), -5, 5, col = "blue", add = T)
```

renvoie le tracé en bleu de la densité associée à la loi de Student $\mathcal{T}(3)$, *i.e.*

$$g(x) = \frac{2}{\pi\sqrt{3}} \frac{1}{\left(1 + \frac{x^2}{3}\right)^2}, \qquad x \in \mathbb{R}.$$

```
curve(df(x, 2, 4), -5, 5, col = "orange", add = T)
```

renvoie le tracé en orange de la densité associée à la loi de Fisher $\mathcal{F}(2,4)$, *i.e.*

$$h(x) = \begin{cases} \dfrac{1}{\left(1+\frac{1}{2}x\right)^3} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

```
curve(dcauchy(x, 2, 3), -5, 5, col = "magenta", add = T)
```

renvoie le tracé en magenta de la densité associée à la loi de Cauchy $\mathcal{C}(2,3)$, *i.e.*

$$i(x) = \frac{1}{3\pi\left(1+\left(\frac{x-2}{3}\right)^2\right)}, \qquad x \in \mathbb{R}.$$

Solution 7. La loi binomiale $\mathcal{B}(n,p)$ peut être approchée par la loi normale $\mathcal{N}(np, np(1-p))$ si $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$, *i.e.*

$$\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p)), \qquad n \geq 30,\ np \geq 5,\ n(1-p) \geq 5.$$

L'objectif de cet exercice est d'étudier cette approximation avec le logiciel R.

1. Reproduire et analyser les commandes suivantes :

   ```
   plot(0:100, dbinom(0:100, 100, 0.5), type = "h", xlim = c(20, 80))
   ```

   renvoie l'histogramme associé à la loi binomiale $\mathcal{B}(100, 0.5)$,

   ```
   curve(dnorm(x, 100 * 0.5, sqrt(100 * 0.5 * (1 - 0.5))), col = "red", add = T)
   ```

   renvoie la courbe de la loi normale $\mathcal{N}(50, 25)$. On constate que cette courbe ajuste bien l'histogramme, cela illustre le fait que $\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p))$ avec $n = 100$ et $p = 0.5$. Notons que les conditions théoriques sont vérifiées :

   $$n = 100 \geq 30, \qquad np = 50 \geq 5, \qquad n(1-p) = 50 \geq 5.$$

2. Pour chaque $n \in \{5, 10, 20, 50, 100\}$, évaluer le maximum de l'erreur commise lorsque l'on approche la loi binomiale par la loi de Poisson (avec correction de continuité), *i.e.* calculer

   $$E_n = \max_{k \in \{0,\ldots,n\}} |\mathbb{P}(X = k) - \mathbb{P}(k - 1/2 \leq Y \leq k + 1/2)|,$$

   où X est une *var* suivant la loi binomiale $\mathcal{B}(n, 5/n)$ et Y est une *var* suivant la loi normale $\mathcal{N}(5, 5(1-5/n))$:

   ```
   for (n in c(5, 10, 20, 50, 100)) {
   En = max(abs(dbinom(0:100, n, 5 / n) - (pnorm(0:100 + 0.5, 5,
   sqrt(5 * (1 - 5 / n))) - pnorm(0:100 - 0.5, 5, sqrt(5 * (1 - 5 / n))))))
   print(En) }
   ```

   renvoie : `[1] 0.002719793 0.0129594 0.0181115 0.01960167`

   Commenter les résultats obtenus :

   On remarque que cette erreur s'amenuise à mesure que n croît, illustrant encore l'approximation $\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p))$ avec $n = 100$ et $p = 0.5$.

3. *Application.* Un disquaire a observé qu'un client sur 20 achète un cd de Jazz. En un mois, 800 clients se présentent au guichet du magasin de disques. Soient X la *var* égale au nombre de clients qui achètent un cd de Jazz durant ce mois.

   (a) Déterminer la loi de X :
   On a $X(\Omega) = \{0, \ldots, 800\}$. Par définition, X est égale au nombre de réalisations de $A=$ "le client achète un cd de Jazz" en 800 expériences indépendantes. On en déduit que X suit la loi binomiale $\mathcal{B}(n, p)$ avec $n = 800$ et $p = \mathbb{P}(A) = \frac{1}{20}$. Donc

   $$\mathbb{P}(X = k) = \binom{800}{k} \left(\frac{1}{20}\right)^k \left(1 - \frac{1}{20}\right)^{800-k}, \qquad k \in \{0, \ldots, 800\}.$$

   (b) Par quelle loi qui peut-on approcher la loi de X ?
   Comme

   $$n = 800 \geq 30, \qquad np = 40 \geq 5, \qquad n(1-p) = 760 \geq 5,$$

   on a $\mathcal{B}(n, p) \approx \mathcal{N}(np, np(1-p))$ avec $np = 40$ et $np(1-p) = 38$.
   Donner une approximation de $\mathbb{P}(35 \leq X \leq 45)$ (avec correction de continuité), puis comparer avec la vraie valeur :
   Soit Y une *var* suivant la loi $\mathcal{N}(40, 38)$. Avec la correction de continuité, on a

   $$\mathbb{P}(35 \leq X \leq 45) \simeq \mathbb{P}\left(35 - \frac{1}{2} \leq Y \leq 45 + \frac{1}{2}\right)$$

   `pnorm(45.5, 40, sqrt(38)) - pnorm(34.5, 40, sqrt(38))` renvoie :
   `[1] 0.6277238`
   La vraie valeur est donnée par la commande
   `sum(dbinom(35:45, 800, 1 / 20))` renvoie : `[1] 0.6279832`
   On constate que les 2 valeurs obtenues sont très proches.