SOLUTION TP n° 5

Solution 1.

1. Rappeler le contexte statistique, l'enjeu et la définition d'un T-IntConf :

   Partant d'un caractère que l'on peut modéliser par une *var* $X \sim \mathcal{N}(\mu, \sigma^2)$, le T-IntConf est un intervalle de confiance pour μ quand σ est inconnu. Sa construction repose sur la loi de Student $\mathcal{T}(\nu)$. Pour obtenir le T-IntConf, il faut calculer :

   - $\overline{x}$: la moyenne de $x_1, \ldots, x_n$,
   - s : l'écart-type corrigé de $x_1, \ldots, x_n$,
   - le réel $t_\alpha(\nu)$ vérifiant :
   $$\mathbb{P}(|T| \geq t_\alpha(\nu)) = \alpha,$$
   où $T \sim \mathcal{T}(\nu)$, $\nu = n - 1$.

   Le T-IntConf est alors donné par :
   $$i_\mu = \left[\overline{x} - t_\alpha(\nu)\frac{s}{\sqrt{n}},\ \overline{x} + t_\alpha(\nu)\frac{s}{\sqrt{n}}\right].$$

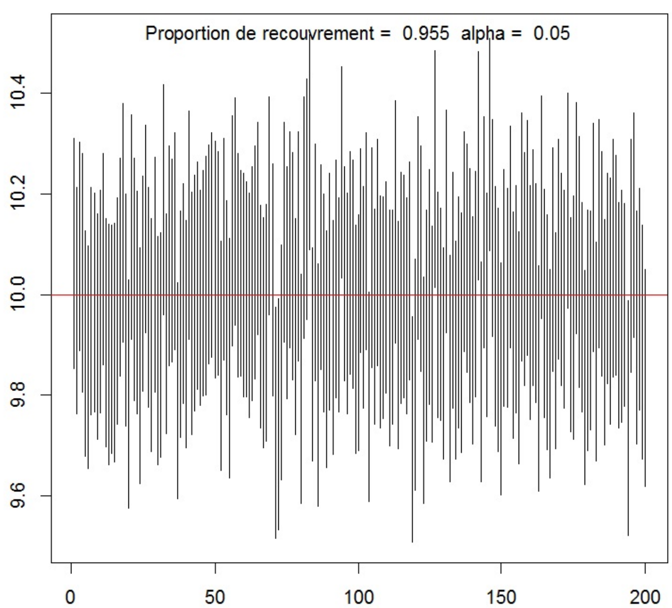
   Il y a donc $100(1-\alpha)\%$ de chances que i_μ contienne μ.

2. Construire dans R une fonction `simuTintconf` qui a 5 arguments : `mu`, `sigma` `alpha`, `n` et `k`. L'enjeu de cette fonction est de

   - représenter sur un graphique le résultat de `k` expériences : $E_1, \ldots, E_{\texttt{k}}$, avec pour tout $j \in \{1, \ldots, \texttt{k}\}$, l'expérience E_j décrite par :
     - obtenir `n` réalisations $x_1, \ldots, x_{\texttt{n}}$ d'un `n`-échantillon $(X_1, \ldots, X_{\texttt{n}})$ d'une *var* $X \sim \mathcal{N}(\texttt{mu}, \texttt{sigma}^2)$,
     - tracer un segment vertical d'ordonnées :
     $$\left(\overline{x} - t_{\texttt{alpha}}(\texttt{n}-1)\frac{s}{\sqrt{\texttt{n}}},\ \overline{x} + t_{\texttt{alpha}}(\texttt{n}-1)\frac{s}{\sqrt{\texttt{n}}}\right),$$
     (avec les notations du cours).
   - représenter une ligne horizontale rouge d'ordonnée `mu`,
   - afficher la proportion de recouvrement définie par le rapport $l/\texttt{k}$ où l désigne le nombre de fois où l'intervalle recouvre la "vraie moyenne" `mu`,
   - afficher la valeur du paramètre `alpha` considéré.

Tester cette fonction en faisant : `simuTintconf(10, 2, 0.05, 300, 200)` :

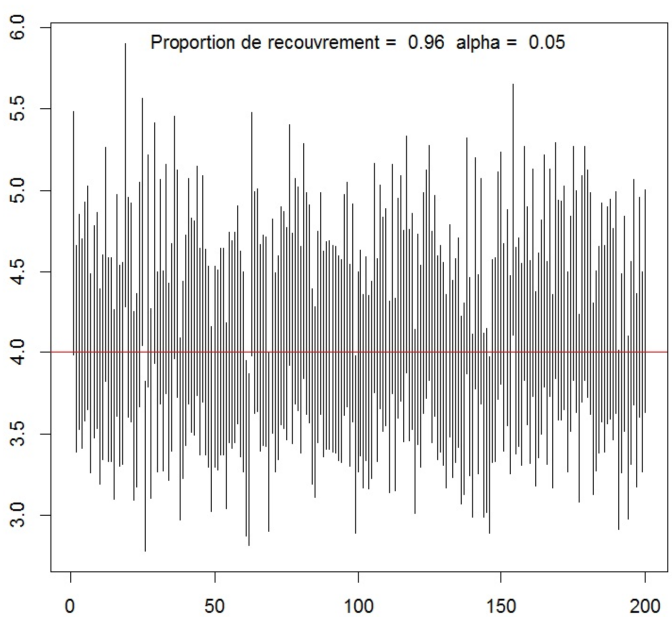
```
simuTintconf = function(mu, sigma, alpha, n, k) {
m1 = numeric(k)
m2 = numeric(k)
q = qt(1 - (alpha / 2), n - 1)
for(i in 1:k) {
tirage = rnorm(n, mu, sigma)
moy = mean(tirage)
s = sd(tirage)
e = q*s / sqrt(n)
m1[i] = moy - e
m2[i] = moy + e
}
plot(m1, type = "n", ylim = range(m1, m2), xlab = "", ylab = "")
segments(1:k, m1, 1:k, m2)
abline(h = mu, col = "red")
prop = mean((m1 <= mu) & (mu <= m2))
text(k / 2, max(m2), paste("Proportion de recouvrement = ", round(prop, 4),
"", "alpha = ", alpha))
}
simuTintconf(10, 2, 0.05, 300, 200)
```



Solution 2. Reproduire et comprendre les enjeux des commandes suivantes :

```
simuChi2intconf = function(mu, sigma, alpha, n, k) {
var1 = numeric(k)
var2 = numeric(k)
q1 = qchisq(1 - (alpha / 2), n - 1)
q2 = qchisq(alpha / 2, n - 1)
for(i in 1:k) {
tirage = rnorm(n, mu, sigma)
moy = mean(tirage)
s2 = var(tirage)
var1[i] = s2 * (n-1) / q1
var2[i] = s2 * (n-1) / q2
}
plot(var1, type="n", ylim = range(var1, var2), xlab = "", ylab = "")
segments(1:k, var1, 1:k, var2)
abline(h = sigma^2, col = "red")
prop=mean((var1 <= sigma^2) & (sigma^2 <= var2))
text(k / 2, max(var2),
paste("Proportion de recouvrement = ", round(prop, 4), "", "alpha = ", alpha))
}
simuChi2intconf(10, 2, 0.05, 300, 200)
```

Ces commandes reprennent le principe de la fonction `simuTintconf` de l'exercice 1, mais avec un Chi2IntConf ; intervalle de confiance pour la variance σ^2 d'un caractère $X \sim \mathcal{N}(\mu, \sigma^2)$.



Solution 3. On considère un échantillon de 10 bouteilles de bière blanche d'une certaine marque et on mesure le volume en litres de chacune d'entre elle. Les résultats sont :

0.499	0.509	0.501	0.494	0.498	0.497	0.504	0.506	0.502	0.496

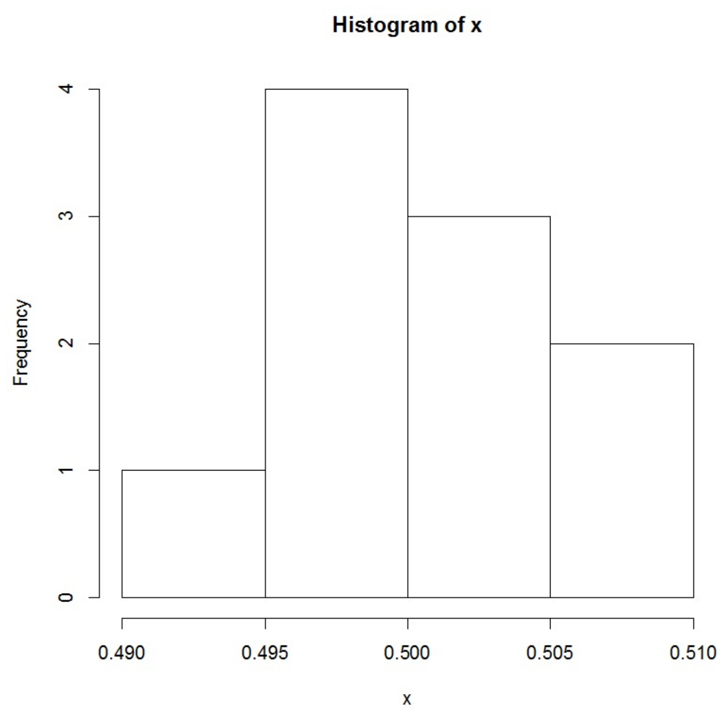
On suppose que le volume en litres d'une telle bouteille peut être modélisée par une *var* $X \sim \mathcal{N}(\mu, \sigma^2)$ avec μ et σ inconnus.

1. Calculer les paramètres statistiques de base :

```
x = c(0.499, 0.509, 0.501, 0.494, 0.498, 0.497, 0.504, 0.506, 0.502, 0.496)
summary(x)
```

2. Tracer un histogramme des fréquences :

```
hist(x)
```



3. Déterminer un intervalle de confiance pour μ au niveau 95% avec la commande `t.test` :

```
t.test(x, conf.level = 0.95)$conf.int
```

4. Déterminer un intervalle de confiance pour μ au niveau 99% :

```
t.test(x, conf.level = 0.99)$conf.int
```

Solution 4. Une ferme de Bay of Plenty en Nouvelle-Zélande produit des kiwis. On a mesuré les masses de 16 kiwis provenant de cette ferme. Les résultats, en grammes, sont :

65.06	71.44	67.93	69.02	67.28	62.34	66.23	64.16
68.56	70.45	64.91	69.90	65.52	66.75	68.54	67.90

On suppose que la masse en grammes d'un kiwi de cette ferme peut être modélisée par une *var* $X \sim \mathcal{N}(\mu, \sigma^2)$, avec μ et σ inconnus.

1. Déterminer un intervalle de confiance pour μ au niveau 95% :

```
x = c(65.06, 71.44, 67.93, 69.02, 67.28, 62.34,  66.23, 64.16, 68.56,
70.45, 64.91, 69.90, 65.52,  66.75,  68.54, 67.90)
t.test(x, conf.level = 0.95)$conf.int
```

2. Déterminer un intervalle de confiance pour σ^2 au niveau 98% *(on pourra utiliser la fonction* `interval_var1` *de la librairie* `OneTwoSamples`*)* :

```
library(OneTwoSamples)
interval_var1(x, alpha = 0.02)
```

Solution 5. Afin d'étudier le pourcentage de consommateurs satisfaits par la boisson énergisante Funny Tiger, on en a interrogé 112 choisis au hasard. Parmi eux, 54 se disent satisfaits.

Déterminer un intervalle de confiance pour la proportion inconnue de consommateurs satisfaits au niveau 99% avec :

- la commande `prop.test` sans la correction de Yates :

  `prop.test(54, 112, conf.level = 0.99, corr = FALSE)$conf.int`

- la commande `prop.test` avec la correction de Yates :

  `prop.test(54, 112, conf.level = 0.99)$conf.int`

- la commande `binom.test` :

  `binom.test(54, 112, conf.level = 0.99)$conf.int`

Solution 6. Un radar situé sur une route départementale enregistre chaque jour la vitesse des véhicules qui passent entre deux sorties. On a regroupé les résultats en classes de 5 kilomètres heure. Le tableau suivant indique les effectifs observés le 25 janvier 2011. Tous les véhicules dont la vitesse dépasse strictement 90 kilomètres heure sont en infraction.

Classe	$]0,75]$	$]75,80]$	$]80,85]$	$]85,90]$	$]90,95]$	$]95,100]$	$]100,105]$	$]105,\infty[$
Effectif	150	450	1000	2000	50	20	10	3

Déterminer un intervalle de confiance pour la proportion inconnue de véhicules en infraction un jour donné au niveau 95% *(on utilisera la commande* `binom.test`*)* :

`binom.test(83, 3683, conf.level = 0.95)$conf.int`