

Glycémie de patients diabétiques soumis à 2 régimes alimentaires différents

2025-10-09

Partie 1 : Description des données

- Quelle est la glycémie moyenne de chaque groupe ?

```
summary(diabete)
```

```
##      groupe1      groupe2
## Min.      :0.8491  Min.      :0.986
## 1st Qu.:1.0114   1st Qu.:1.064
## Median :1.0633   Median :1.102
## Mean    :1.0525   Mean     :1.098
## 3rd Qu.:1.0977   3rd Qu.:1.126
## Max.    :1.2510   Max.      :1.230
```

La glycémie moyenne du groupe1 est égale à 1.0525 (1.05 g/L). La glycémie moyenne du groupe2 est égale à 1.098 (1.10 g/L).

- Quel est l'écart-type de la distribution de la glycémie pour le régime A et pour le régime B ?

```
# Calcul de l'ecart-type
round(sd(diabete$groupe1),2) # Régime A
```

```
## [1] 0.08
```

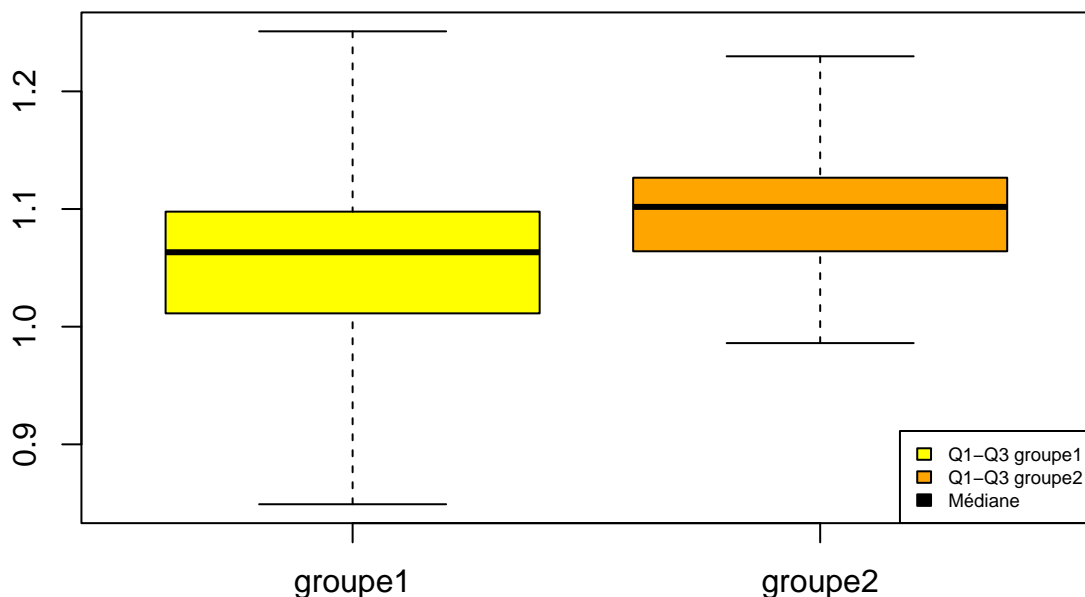
```
round(sd(diabete$groupe2),2) # Régime B
```

```
## [1] 0.05
```

L'écart-type de la distribution de la glycémie pour le groupe 1 est égale à 0.08. L'écart-type de la distribution de la glycémie pour le groupe 2 est égale à 0.05.

- Représenter la distribution de la glycémie de chaque régime sous la forme d'un graphique permettant de visualiser la médiane et les quartiles.

Distribution de la glycémie



Partie 2 : Analyses statistiques

- Quelle est la différence moyenne de glycémie entre les deux régimes ?

```
round(mean(diabete$groupe2) - mean(diabete$groupe1), 2)
```

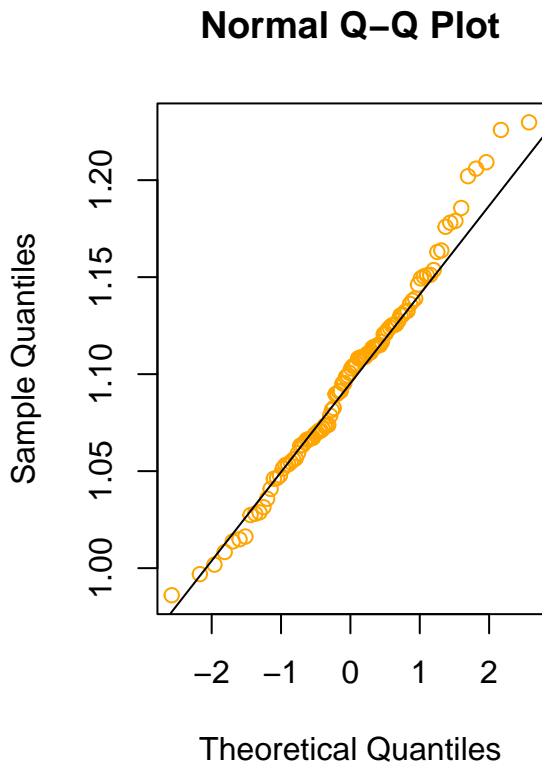
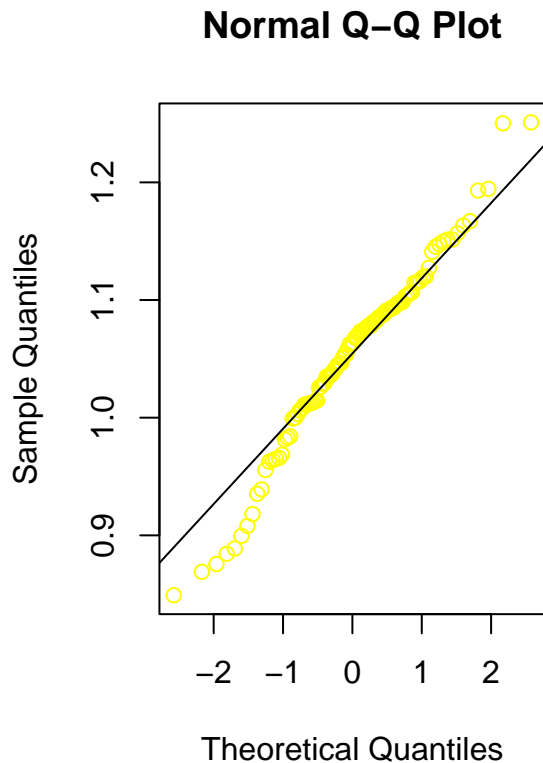
```
## [1] 0.05
```

La différence moyenne de glycémie entre les deux régimes est égale à 0.05 g/L.

- Cette différence est-elle significative ?

Pour savoir si cette différence est statistiquement significative, on va effectuer un test statistique. Cependant, plusieurs critères sont à valider avant de pouvoir réaliser le test. Tout d'abord, nos deux groupes doivent contenir un assez grand nombre de sujets à comparer (>30) et leurs distributions doivent suivre une loi normale. On sait que les 2 groupes contiennent chacun 100 patients diabétiques. Un des critères est donc déjà valide. Pour vérifier si nos deux groupes suivent une loi normale, on peut le vérifier avec un diagramme de normalité.

```
# Affichage de mes graphiques côte à côte
par(mfrow= c(1,2))
# Diagramme groupe1
qqnorm(diabete$groupe1, col = "yellow");qqline(diabete$groupe1)
# Diagramme groupe2
qqnorm(diabete$groupe2, col = "orange");qqline(diabete$groupe2)
```



On constate sur les diagrammes des 2 groupes que les points sont majoritairement proches de la ligne droite. Cela signifie que les données suivent une distribution normale. Le second critère est également validé. Il ne reste qu'un dernier critère à valider pour effectuer notre test. Il faut que les variances soient identiques dans les deux groupes. Pour cela, on va faire un test d'égalité des variances.

```
var.test(diabete$groupe1, diabete$groupe2)
```

```
##
##  F test to compare two variances
##
## data:  diabete$groupe1 and diabete$groupe2
## F = 2.3643, num df = 99, denom df = 99, p-value = 2.588e-05
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.590824 3.513956
## sample estimates:
## ratio of variances
##           2.364336
```

Interprétation du test :

H0 : les variances sont égales (hypothèse nulle)

H1 : les variances ne sont pas égales (hypothèse alternative)

On rejette H0 ($p\text{-value} < 0.05$) et H1 est validé. Donc, les variances des deux distributions ne sont pas égales.

Malgré que les variances des deux distributions ne soient pas égales, il est possible d'effectuer notre test statistique. Dans cette situation, on utilise un test de Welch pour comparer les moyennes.

```
t.test(diabete$groupe1, diabete$groupe2, var.equal=FALSE)

##
##  Welch Two Sample t-test
##
## data:  diabete$groupe1 and diabete$groupe2
## t = -4.8508, df = 170.04, p-value = 2.764e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.06413844 -0.02703532
## sample estimates:
## mean of x mean of y
##  1.052534  1.098121
```

Interprétation du test :

H0 : les moyennes sont égales (hypothèse nulle)

H1 : les moyennes ne sont pas égales (hypothèse alternative)

On rejette H0 ($p\text{-value} < 0.05$) et H1 est validé. Donc, les moyennes des deux distributions ne sont pas égales.

A partir du résultat du test statistique, on peut en conclure que la différence moyenne de glycémie entre les deux régimes est significative.

Taille de l'échantillon :

- Échantillonner aléatoirement 10 patients de chaque groupe (groupe1b et groupe2b)

```
groupe1b <- sample(diabete$groupe1, 10)
groupe1b
```

```
## [1] 1.0742182 0.9181026 1.1065905 1.0905786 1.0982516 1.0145580 1.1409902
## [8] 1.0539530 1.1930430 0.9657683
```

```
groupe2b <- sample(diabete$groupe2, 10)
groupe2b
```

```
## [1] 1.027377 1.069795 1.081488 1.103012 1.225885 1.073339 1.125125 1.111069
## [9] 1.069019 1.105261
```

- Répéter les questions des parties 1 et 2 pour ces sous-échantillons.

Glycémie moyenne de chaque groupe :

```
summary(groupe1b)
summary(groupe2b)
```

La glycémie moyenne du groupe1b est égale à 1.064 (1.06 g/L).

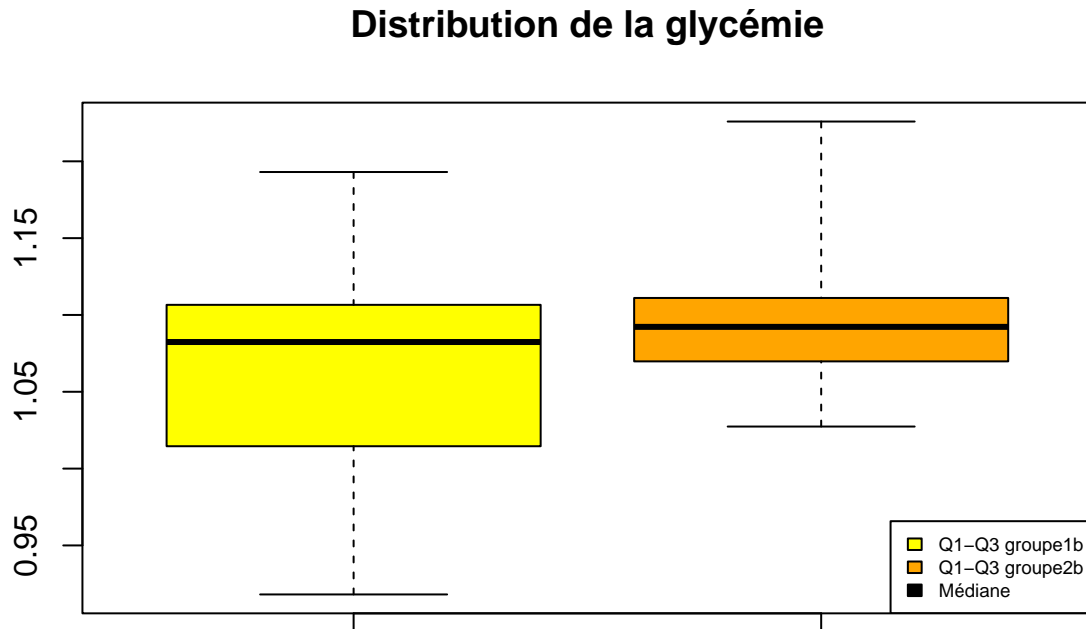
La glycémie moyenne du groupe2b est égale à 1.128 (1.13 g/L).

Ecart-type de la distribution de la glycémie pour le régime A et pour le régime B :

```
# Calcul de l'ecart-type
round(sd(groupe1b),2) # Régime A
round(sd(groupe2b),2) # Régime B
```

L'écart-type de la distribution de la glycémie pour le groupe 1b est égale à 0.06 g/L. L'écart-type de la distribution de la glycémie pour le groupe 2b est égale à 0.05 g/L.

Représentation graphique :



Différence moyenne de glycémie entre les deux régimes pour les sous-groupes :

```
round(mean(groupe2b) - mean(groupe1b),2)
```

La différence moyenne de glycémie entre les deux régimes est égale à 0.06 g/L.

Cette différence est-elle significative ?

Pour savoir si cette différence est statistiquement significative, il faut de nouveau effectuer un test statistique. Cependant, l'échantillon est trop petit pour utiliser un test paramétrique (test t de Student, test de Welch). Toutefois, il existe une alternative. On peut effectuer un test de Mann-Whitney (ou test de Wilcoxon).

```
wilcox.test(groupe1b, groupe2b, var.equal=FALSE)
```

Interprétation du test :

H0 : les moyennes sont égales (hypothèse nulle)

H1 : les moyennes ne sont pas égales (hypothèse alternative)

H_0 est rejeté et H_1 est validé ($p\text{-value} < 0.05$). Donc, les moyennes des deux sous-groupes ne sont pas égales.

A partir du résultat du test statistique, on peut en conclure que la différence moyenne de glycémie entre les deux régimes des deux sous-groupes est significative.

- Que remarquez-vous en comparant les résultats obtenus pour des groupes de 100 ou 10 patients ?

En comparant les résultats obtenus pour des groupes de 100 patients et de 10 patients, on constate que leurs résultats sont différents. Cependant, il faut prendre en compte que les données récoltés à partir de petits échantillons sont moins fiables que celles de grands échantillons car l'influence du hasard est accrue. Ainsi, les données sont susceptibles d'être moins précises.