L3 Biosciences, Examen de Statistique

Julien Derr, Ghislain Durif, Laurent Modolo, Franck Picard

Janvier 2023

Note : cet examen est composé de trois exercices indépendants.

1 Réflexion sur le R2 (15 points)

Le coefficient de détermination d'un modèle est proposé par la plupart des logiciels de statistique dans le cadre des modèles linéaires (ANOVA ou régression). Malgré sa popularité, l'interprétation de cet indicateur est délicate et peut conduire à des contre-sens.

1. *(1 points)* Dans le cadre d'un modèle linéaire, donner la définition du coefficient de détermination

   Solution: R2 se définit par : la part de variance (de la réponse), expliquée par le modèle de régression (linéaire simple, ici).

2. *(2 points)* Rappeler la décomposition de la variance totale (SCT) et l'interprétation de chaque terme, et en déduire la formule du R2

   Solution: La dispersion totale des données se mesure par la somme des distances au carré, entre chaque observation et la moyenne des réponses (moyenne générale). On appelle ce paramètre la Somme des Carrés Totale (SCT). En, anglais on l'appelle Total Sum of Squares, son abréviation est SST. La variabilité totale des données est alors répartie en deux composantes : une part expliquée par le modèle de régression, une part non expliquée, qui correspond à la variance résiduelle. La formule du R2 est donc SCM / SCT.

3. *(1 points)* Quelles sont les valeurs possibles du R2 et comment s'interprète-t-il ?

   Solution: R2 est un indicateur compris entre 0 et 1. C'est un pourcentage. Il est d'autant plus élevé que le modèle explique une forte part de la variabilité observée.

4. *(2 points)* Quelle est son interprétation spécifique dans le cas de la régression simple ? Faites le lien avec l'interprétation générale

   Solution: En régression linéaire simple, le R2 est le coefficient de régression au carré, qui quantifie l'intensité d'association linéaire entre la variable réponse et la covariable. Plus cette intensité est forte mieux le modèle s'ajuste aux données.

Il semblerait que le R2 soit à interpréter avec précaution. Sa seule valeur n'est pas suffisante pour caractériser l'adéquation du modèle aux données. Francis Anscombe est un statisticien britannique (1918-2001) connu pour ses contributions à l'analyse des données. Il a notamment imaginé un jeu de données artificiel illustrant les pièges de la régression linéaire lorsque les fondements de cette méthode ne sont pas maitrisés. Ce jeu de données s'appelle le quartet d'Anscombe: 4 jeux de données très différents mais partageant des propriétés statistiques similaires.

```r
data("anscombe")
head(anscombe)
```

```
  x1 x2 x3 x4   y1   y2    y3   y4
1 10 10 10  8 8.04 9.14  7.46 6.58
2  8  8  8  8 6.95 8.14  6.77 5.76
3 13 13 13  8 7.58 8.74 12.74 7.71
4  9  9  9  8 8.81 8.77  7.11 8.84
5 11 11 11  8 8.33 9.26  7.81 8.47
6 14 14 14  8 9.96 8.10  8.84 7.04
```

On fait une régression linéaire simple pour chaque jeu de données, et voici les sorties.

5. *(3 points)* Interprétez synthétiquement les sorties des 4 modèles et concluez pour les 4 jeux de données.

Solution: Au vu des résultats on conclurait que le modèle s'ajuste bien aux données R2=0.67 et que le modèle est pertinent (Fpv<0.05). Ensuite on dirait que l'intercept est significativement différent de zéro et que le coefficient de régression est significativement non nul.

```r
summary(lm(data = anscombe, y1 ~ x1))
```

```

Call:
lm(formula = y1 ~ x1, data = anscombe)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.92127 -0.45577 -0.04136  0.70941  1.83882 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)   3.0001     1.1247   2.667  0.02573 * 
x1            0.5001     0.1179   4.241  0.00217 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.237 on 9 degrees of freedom
Multiple R-squared:  0.6665,    Adjusted R-squared:  0.6295 
F-statistic: 17.99 on 1 and 9 DF,  p-value: 0.00217
```

```r
summary(lm(data = anscombe, y2 ~ x2))
```

```

Call:
lm(formula = y2 ~ x2, data = anscombe)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.9009 -0.7609  0.1291  0.9491  1.2691 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)    3.001      1.125   2.667  0.02576 * 
x2             0.500      0.118   4.239  0.00218 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.237 on 9 degrees of freedom
Multiple R-squared:  0.6662,    Adjusted R-squared:  0.6292 
```

```
F-statistic: 17.97 on 1 and 9 DF,  p-value: 0.002179
```

```r
summary(lm(data = anscombe, y3 ~ x3))
```

```

Call:
lm(formula = y3 ~ x3, data = anscombe)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.1586 -0.6146 -0.2303  0.1540  3.2411 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)   3.0025     1.1245   2.670  0.02562 * 
x3            0.4997     0.1179   4.239  0.00218 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.236 on 9 degrees of freedom
Multiple R-squared:  0.6663,	Adjusted R-squared:  0.6292 
F-statistic: 17.97 on 1 and 9 DF,  p-value: 0.002176
```

```r
summary(lm(data = anscombe, y4 ~ x4))
```

```

Call:
lm(formula = y4 ~ x4, data = anscombe)

Residuals:
   Min     1Q Median     3Q    Max 
-1.751 -0.831  0.000  0.809  1.839 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)   3.0017     1.1239   2.671  0.02559 * 
x4            0.4999     0.1178   4.243  0.00216 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.236 on 9 degrees of freedom
Multiple R-squared:  0.6667,	Adjusted R-squared:  0.6297 
F-statistic:    18 on 1 and 9 DF,  p-value: 0.002165
```

1.1 De l'importance de la représentation

6. *(3 points)* à l'aide des représentations des 4 jeux de données expliquez si selon vous les conditions d'utilisation de la régression linéaire simple sont remplies

> **Solution:** - x1,y1 : ok, relation linéaire - x2,y2 : relation quadratique, donc pas linéaire simple - x3,y3 : relation linéaire mais un point aberrant - x4,y4 : pas de variabilité sur x, donc pas vraiment adapté de faire une régression linéaire. Peut etre plus une ANOVA

7. *(3 points)* Proposez une conclusion synthétique sur l'interprétation du R2 et son lien avec la statistique de Fisher (et la p-value calculée dans les tables d'ANOVA)

> **Solution:** Il n'y a pas de lien entre le coefficient de détermination et la p-value (du test de la pente) car ces deux paramètres mesurent des choses différentes. Comme nous l'avons vu, le R2 indique la part de variabilité expliquée par le modèle. De son côté, la p-value nous permet de conclure sur la significativité de la relation linéaire entre la variable réponse et la variable prédictive. La part de variabilité expliquée par un modèle peut dépendre du contexte, elle peut être de l'ordre de 80%

dans certains domaines, et uniquement de 10% dans d'autres, comme en biologie par exemple. La part de variabilité expliquée par le modèle, ne préjuge pas de la significativité de la relation entre la réponse et la variable prédictive. Dis autrement, même si un modèle n'explique que peu e variabilité, la p-value du test de la pente demeure valide, à condition que les hypothèses du test soient validées, bien entendu.

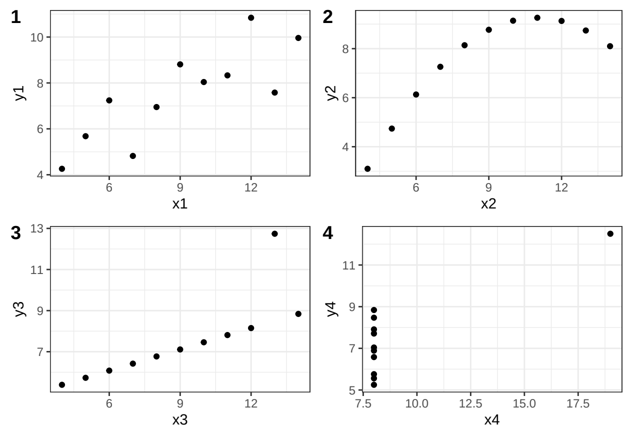

Figure 1: Representation of the Anscombe dataset

2 The Lady tasting tea (15 points)

Muriel Bristol, une scientifique de la Station de recherche de Rothamstead, affirme à Ronald Fisher qu'elle peut faire la différence entre une tasse de thé pour laquelle le lait a été versé avant le thé et une tasse de thé pour laquelle le thé a été versé avant le lait. Pour mettre à l'épreuve sa collègue Fisher réalisa une expérience qui devint célèbre, sous le nom "The lady tasty tea". Cette expérience a motivé le développement du test exact de Fisher.

Le plan d'expérience est le suivant :

- on prépara 8 tasses de thé :
  - 4 tasses avec le lait versé en premier
  - 4 tasses avec le thé versé en premier.
- Fisher demanda à Madame Bristol de gouter chaque tasse et d'assigner 4 tasses à chaque ordre de remplissage.

Pour formaliser cette expérience, on note $X_i = 1$ si le thé a été versé avant le lait pour l'expérience $i = 1, ..., 8$, $X_i = 0$ sinon, et $Y_i = 1$ si l'expérimentatrice (M. Bristol) estime que le thé a été versé avant le lait pour l'expérience i, $Y_i = 0$ sinon. On notera $p_X = \mathbb{P}(X_i = 1)$, $p_Y = \mathbb{P}(Y_i = 1)$, et $p_{XY} = \mathbb{P}(X_i = 1, Y_i = 1)$.

1. *(1 point)* Formulez l'hypothèse nulle d'indépendance, en termes mathématiques et en la formulant à l'aide d'une phrase appliquée au contexte de l'expérience de la dégustation de thé.

> **Solution:** Le test d'indépendance consiste à tester l'hypothèse nulle $\mathcal{H}_0 : \{p_{XY} = p_X p_Y\}$, ce qui signifie qu'on fait l'hypothèse que la décision de Muriel Bristol est indépendante de la manière dont les tasses de thé ont été préparées.

Ensuite, on compte le nombre d'expériences telles que:

- Le thé a été versé avant le lait et Muriel Bristol le détecte: $N_{1,1} = \sum_{i=1}^{8} X_i Y_i$
- Le thé a été versé après le lait et Muriel Bristol le détecte: $N_{0,0} = \sum_{i=1}^{8}(1 - X_i)(1 - Y_i)$
- Le thé a été versé avant le lait et Muriel Bristol ne le détecte pas: $N_{1,0} = \sum_{i=1}^{8} X_i(1 - Y_i)$
- Le thé a été versé après le lait et Muriel Bristol ne le détecte pas: $N_{0,1} = \sum_{i=1}^{8}(1 - X_i)(1 - Y_i)$

On appellera les marges, le nombre d'expériences telles que:

- le thé a été versé avant le lait : $N_{1,+} = \sum_{i=1}^{8} X_i$
- le thé a été versé après le lait : $N_{0,+} = \sum_{i=1}^{8}(1 - X_i)$
- Muriel Bristol dit que le thé a été versé avant le lait : $N_{+,1} = \sum_{i=1}^{8} Y_i$
- Muriel Bristol dit que le thé a été versé après le lait : $N_{+,0} = \sum_{i=1}^{8}(1 - Y_i)$

On note enfin $N_{+,+}$ le nombre total d'expériences.

2. *(1 points)* A l'aide des notations proposées, construire la table de contingence de l'expérience.

> **Solution:** La table de contingence présente les effectifs croisés par modalité:

	$Y = 1$	$Y = 0$	$Marge$
$X = 1$	$N_{1,1}$	$N_{1,0}$	$N_{1,+}$
$X = 0$	$N_{0,1}$	$N_{0,0}$	$N_{0,+}$
$Marge$	$N_{+,1}$	$N_{+,0}$	$N_{+,+}$

3. *(1 points)* A l'aide des notations précédentes, proposez un estimateur des paramètres $p_X = \mathbb{P}(X_i = 1)$, $p_Y = \mathbb{P}(Y_i = 1)$, $p_{XY} = \mathbb{P}(X_i = 1, Y_i = 1)$, notés $\widehat{p}_X$, $\widehat{p}_Y$, $\widehat{p}_{XY}$.

> **Solution:** Les estimateurs sont les fréquences empiriques:
>
> $$\widehat{p}_X = \frac{N_{1,+}}{N_{+,+}}, \quad \widehat{p}_Y = \frac{N_{+,1}}{N_{+,+}}, \quad \widehat{p}_{XY} = \frac{N_{1,1}}{N_{+,+}},$$

Le tableau suivant donne les résultats de cette expérience:

	$Y = 1$	$Y = 0$	$Marge$
$X = 1$	3	1	4
$X = 0$	1	3	4
$Marge$	4	4	8

4. *(1 points)* Donnez les estimations des paramètres p_X, p_Y, p_{XY}.

> **Solution:** Au vu des résultats de l'expérience, on a
>
> $$\widehat{p}_X = \widehat{p}_Y = 4/8 = 1/2, \quad \widehat{p}_{XY} = 3/8,$$

5. *(2 points)* Au vu des valeurs des marges, quel serait les effectif attendus $(N_{1,1}^{attendu}, N_{1,0}^{attendu}, N_{0,1}^{attendu}, N_{0,0}^{attendu})$ pour la table de contingence sous l'hypothèse nulle d'indépendance ?

> **Solution:** Sous l'hypothèse d'indépendance, $p_{XY} = p_X p_Y$, donc on aurait un effectif attendu de
>
> $$N_{1,1}^{attendu} = N_{++}\widehat{p}_X\widehat{p}_Y = 8 \times 1/2 \times 1/2 = 2$$
>
> On a aussi que $N_{1,1}^{attendu} = N_{1,0}^{attendu} = N_{0,1}^{attendu} = N_{0,0}^{attendu} = 2$.

6. *(1 point)* Donnez la formule de la statistique du χ^2 permettant de tester l'hypothèse nulle. Quel est sont degré de liberté ? Donnez sa valeur sur cet exemple.

> **Solution:** La statistique du χ^2 est construite telle que:
>
> $$T = \sum_{i=0}^{1}\sum_{j=0}^{1} \frac{\left(N_{i,j} - N_{i,j}^{attendu}\right)^2}{N_{i,j}^{attendu}}$$
>
> C'est une statistique du χ^2 à $(2-1) \times (2-1) = 1$ degré de liberté. La statistique est donc : $(3-2)^2/2 + (1-2)^2/2 + (1-2)^2/2 + (3-2)^2/2 = 2$

7. *(1 point)* La p-value donnée par le test du χ^2 est de 0.1573. Qu'en concluez vous ?

> **Solution:** Au vu du résultat de l'expérience, on ne peut pas rejeter l'hypothèse nulle (même si on prenait $\alpha = 10\%$). On peut donc en conclure que Muriel Bristol n'arrive pas à distinguer le mode de préparation de son thé (au lait) !

8. *(1 point)* La loi de la statistique du χ^2 est-elle une loi exacte? La statistique du χ^2 suit-elle une loi continue ou discrète ?

Solution: La statistique du Chi2 suit une loi continue, c'est une distance, et sa distribution provient d'une approximation asymptotique. Il faut donc qu'on ait suffisamment d'observations pour l'utiliser.

9. *(2 points)* Si on on considère le même nombre de tasses de chaque préparation est fixée à 4 et que Mme Bristol ne goute que 4 tasses de chaque mode de remplissage (marges fixées), quel est le nombre de réalisations possibles pour cette expérience ? Donnez les tables de contingence possibles.

Solution: Il y a 5 tables de contingence possibles:

	Y=1	Y=0
X=1	4	0
X=0	0	4

	Y=1	Y=0
X=1	3	1
X=0	1	3

	Y=1	Y=0
X=1	2	2
X=0	2	2

	Y=1	Y=0
X=1	1	3
X=0	3	1

	Y=1	Y=0
X=1	0	4
X=0	4	0

10. *(1 point)* . Que pouvez vous en conclure quant à l'utilisation de la loi du χ^2 sur cette expérience ?

Solution: Etant donné le faible nombre d'échantillons et le faible nombre de valeurs possibles pour la statistique, l'approximation asymptotique de la loi de la statistique apparait comme peu appropriée à ce contexte.

Au lieu d'utiliser la statistique du χ^2 pour ces données, Fisher utilisa la distribution hypergéométrique. Ce test calcule la probabilité d'avoir de 0 à 4 tasses dans chacune des cases du tableau quand toutes les marginales sont 4. C'est une probabilité exacte pour des marginales données fixées. Fisher proposa ce test comme une solution optimale au problème.

11. *(1 point)* Pourquoi Fisher qualifie ce test d'optimal par rapport à l'utilisation de la statistique du χ^2 ?

> **Solution:** Fisher développe un test exact qui ne repose pas sur une approximation asymptotique de la distribution de la statistique de test.

12. *(1 point)* La p-value du test exact de Fisher est de 0.4857. Qu'en concluez-vous ?

> **Solution:** Le test exact de Fisher corrobore les conclusions du test du χ^2 comme quoi Muriel Bristol n'a pas pu détecter la méthode de préparation du thé.

13. *(1 point)* Si la pvalue du test exact de Fisher avait été $< \alpha$ qu'en auriez vous déduit par rapport au résultat du test du χ^2 ?

> **Solution:** Si la décision prise à l'issue du test de Fisher était différente de celle prise à l'issue du test du chi2, on aurait privilégié le test de Fisher car il est exact, et adapté aux faibles effectifs, alors que le test du chi2 est un test asymptotique, peu adapté à la situation présente.

3 Lutte contre les insectes par traitement thermique (15 points)

Le traitement thermique est une alternative populaire aux pesticides synthétiques pour désinfecter les installations de stockage alimentaires et les entrepôts à grains vides. *Sitophilus zeamais* Mostchulsky est l'un des insectes les plus destructeurs trouvés dans ce type d'installations alimentaires. Dans l'article "*The Effect of Acclimation to Sublethal Temperature on Subsequent Susceptibility of Sitophilus zeamais Mostchulsky (Coleoptera: Curculionidae) to High Temperatures*", Lü et Zhang, PLOS One (2016), on réalise l'expérience suivante en deux phases:

- Les insectes sont chauffés à une température non léthale de (36°C) pendant une phase plus ou moins longue d'acclimatation (*acclimation time*)
- Les insectes sont ensuite re-chauffés à des températures léthales (43°C ou 55°C) pendant une phase plus ou moins longue d'exposition (*exposition time*)

Un temps d'acclimatation de 0 correspond au contrôle, les insectes ne subissent pas la première phase d'acclimatation à 36°C.

L'article étudie l'effet de ces deux phases sur la mortalité des populations de *Sitophilus zeamais* Mostchulsky. La mortalité des adultes *S. zeamais* exposés à différentes durée d'exposition à une température létale est affichée dans les Tables~6 et 7, ainsi que les tables d'ANOVA respectives correspondant aux analyses.

1. *(1 point)* Le protocole est-il équilibré ? Pourquoi ? Quelle est l'implication de cette propriété du plan d'expérience sur les analyses ?

> **Solution:** Le plan est équilibré car chaque croisement des modalités exposure/acclimatation est mesurée sur le même nombre de répétitions (3). Dans ce cadre, la contribution de chacun de facteurs peut être distinguée (la décomposition des sommes de carré est unique).

2. *(2 points)* En considérant les facteurs temps d'exposition et temps d'acclimatation comme des facteurs qualitatifs, écrivez le modèle d'analyse de la variance à deux facteurs avec interaction, ainsi que ses hypothèses. Interprétez les facteurs du modèle proposé.

> **Solution:** On note Y_{ijk} la mortalité des individus ayant reçu le temps d'acclimatation i ($i = 1, \cdots, 8$) et le temps d'exposition j ($j = 1, \cdots, 4$) à la répétition k ($k = 1, \cdots, 3$). Le modèle d'ANOVA est tel que:
> $$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2).$$
> On considère que les erreurs sont indépendantes et identiquement distribuées, et que σ^2 est constante. Les paramètres du modèles sont μ, la mortalité moyenne, α_i l'effet du temps d'acclimatation i sur la mortalité, β_j l'effet du temps d'exposition j sur la mortalité, et γ_{ij} l'interaction (ou effet conjoint) des temps d'acclimatation et d'exposition.

3. *(2 points)* Expliquez les degrés de liberté dans les tables 6 et 7.

> **Solution:** Le premier facteur (temps d'acclimatation) présente 4 modalités (0h, 1h, 3h, 5h). On a donc $I = 4$ et le degré de liberté pour tester l'effet acclimiation est donc $I - 1 = 3$. On a 8 modalités pour le temps d'exposition (t1,...t8), donc $J = 8$ et le degré de liberté pour tester l'effet est donc $J - 1 = 7$. L'interaction est mesurée sur toutes les combinaisons de niveaux des deux facteurs, donc on a $IJ = 4 \times 8$ combinaisons, et le degré de liberté pour tester l'interaction est $(I - 1) \times (J - 1) = 21$. Pour le degré de liberté de la somme des carrés résiduelle, étant donné qu'on a un dispositif équilibré, c'est $I \times J \times (r - 1)$ avec $r = 3$ le nombre de réplicats (ce qui fait bien 64). Au total, on a bien $I \times J \times r$ données enregistrées ($n = 96$).

4. *(3 points)* Les Figures 2 et 3 représentent les tendances moyennes des mortalités en fonction du temps d'exposition et du temps d'acclimatation. Commentez l'impact du temps d'exposition sur la mortalité (à 43 puis 55C). Commentez également l'effet du temps d'acclimatation. Au vu des graphiques, de manière préliminaire, que pensez vous de l'existence d'une interaction entre les facteurs ? A 43C ? A 55C ? Pourquoi ?

   Solution: En Figure 2 l'augmentation de la mortalité dépend du temps d'acclimatation : la pente est plus forte pour un temps d'acclimatation nul, et est la plus faible pour un temps d'acclimatation élevée (5h). Ce qui signifie qu'à cette température d'exposition, lorsque les animaux se sont acclimatés, l'exposition à une température léthale semblerait avoir moins d'impact sur la mortalité. En revanche la Figure 3 montre que globalement les pentes sont comparables (sauf pour le temps d'acclimatation 5h qui présente une pente légèrement différente). D'après les informations graphiques, il semblerait que l'interaction soit moint forte quand la température d'exposition est de 55C.

5. *(2 points)* Grâce à l'étude des tables d'ANOVA (6 et 7), expliquez quel facteur a le plus d'impact sur la mortalité des insectes, après une exposition à 43C et à 55C. Justifiez votre choix en expliquant l'indicateur que vous utilisez.

   Solution: Si on considère la statistique de Fisher qui compare la somme des carrés du facteur à la somme des carrés résiduelle (en normalisant par les degrés de libertés), alors on s'aperçoit qu'à 43C c'est le temps d'acclimatation qui a le plus d'impact (F-value = 81.841), alors que c'est le temps d'exposition à 55C (F-value= 57.839). Les deux statistiques de Fisher sont significatives à au moins $\alpha = 0.001$. La conclusion c'est que lorsqu'on chauffe fort (55C), le temps d'acclimatation a moins d'impact (mais il en a un petit quand même).

6. *(3 points)* Grâce à l'étude des tables d'ANOVA (6 et 7), interprétez l'impact de l'interaction entre les deux temps d'exposition sur la mortalité.

   Solution: Les tables d'ANOVA montrent qu'à 43C l'interaction est significative (F-value=6.147, p-value<0.001) , alors qu'elle ne l'est pas à 55C (F-value=0.611, p-value=0.896). Ce résultat signifie que lorsque le silot est chauffé à 55C c'est le temps d'exposition qui explique majoritairement la variabilité de la mortalité. Au vu du tableau de données, on pourrait suggérer que la meilleure des stratégies consiste à chauffer directement le silot à une température léthale sans période d'acclimatation. Le fait que l'interaction soit significative à 43C et pas à 55C suggère qu'il ne faut pas laisser de période d'acclimatation, sinon les insectes s'y adaptent et sont moins sensibles à l'exposition à une température léthale.

7. *(2 points)* Au vu des Figures 2 et 3, proposez une modélisation alternative à l'ANOVA à deux facteurs pour analyser ces données.

   Solution: Etant donné la nature continue des covariables (temps d'acclimatation et exposition) on pourrait proposer un modèle de régression linéaire. N'ayant que peu de niveaux pour le temps d'acclimatation on pourrait proposer un modèle de régression linéaire simple avec le temps d'exposition comme variable explicative et le temps d'acclimatation comme facteur qualitatif. Un tel modèle s'appelle une analyse de la covariance, qui mélange facteurs quantitatifs et qualitatifs.

Exposure time at 43°C	Acclimation time at 36°C (hour)			
	0	1	3	5
t1	3.5 ± 1.67	5.0 ± 2.89	0.0 ± 0.00	1.7 ± 1.67
t2	5.0 ± 2.89	3.33 ± 3.33	3.3 ± 1.67	1.7 ± 1.67
t3	18.3 ± 3.30	10.0 ± 5.77	6.7 ± 3.33	1.7 ± 1.67
t4	41.6 ± 11.67	15.0 ± 7.64	13.3 ± 3.33	8.3 ± 4.41
t5	45.0 ± 15.28	25.0 ± 5.00	21.7 ± 0.00	10.0 ± 6.67
t6	65.0 ± 2.89	38.3 ± 8.82	30.0 ± 8.66	20.0 ± 2.89
t7	86.7 ± 7.26	40.0 ± 8.66	26.7 ± 4.41	21.7 ± 4.41
t8	95.0 ± 5.00	53.3 ± 9.28	35.0 ± 1.67	33.0 ± 5.00

Source	df	Type III SS	mean Square	F-value	p-value
Acclimation time	3	15409.115	5136.372	81.841	0.000
Exposure time	7	31635.156	4519.308	72.009	0.000
Acclimation time x exposure time	21	8101.302	385.776	6.147	0.000
Error	64	40136.667	62.760		
Total	96	117425.000			

Table 6: The effect of acclimation to sublethal high temperature on mortality (%) of S. zeamais exposed to 43°C. Top : Data are average mortality rates ± SE of three replicates, and exposure times are in increasing order ($t1 < t2 \cdots < t8$). Bottom: ANOVA table

Exposure time at 55°C	Acclimation time at 36°C (hour)			
	0	1	3	5
t1	1.7 ± 2.89	0.0 ± 0.00	0.0 ± 0.00	1.7 ± 1.67
t2	5.0 ± 0.00	0.0 ± 0.00	0.0 ± 0.00	1.7 ± 1.67
t3	21.7 ± 29.30	20.0 ± 17.56	3.3 ± 3.33	1.7 ± 1.67
t4	31.7 ± 24.66	35.0 ± 12.58	25.0 ± 22.55	3.3 ± 3.33
t5	56.7 ± 20.21	33.3 ± 10.92	28.0 ± 9.23	8.3 ± 4.41
t6	78.3 ± 33.29	70.0 ± 10.00	65.0 ± 7.64	55.0 ± 13.23
t7	86.7 ± 23.09	83.3 ± 16.67	76.7 ± 16.67	75.0 ± 2.89
t8	96.7 ± 5.77	96.3 ± 1.67	81.7 ± 1.67	81.7 ± 3.33

Source	df	Type III SS	mean Square	F-value	p-value
Acclimation time	3	4869.531	1623.177	6.153	0.001
Exposure time	7	106805.990	15256.999	57.839	0.000
Acclimation time x exposure time	21	3382.552	161.074	0.611	0.896
Error	64	16883.333	263.802		
Total	96	272625.00			

Table 7: The effect of acclimation to sublethal high temperature on mortality (%) of S. zeamais exposed to 55°C. Top : Data are average mortality rates ± SE of three replicates, and exposure times are in increasing order ($t1 < t2 \cdots < t8$). Bottom: ANOVA table

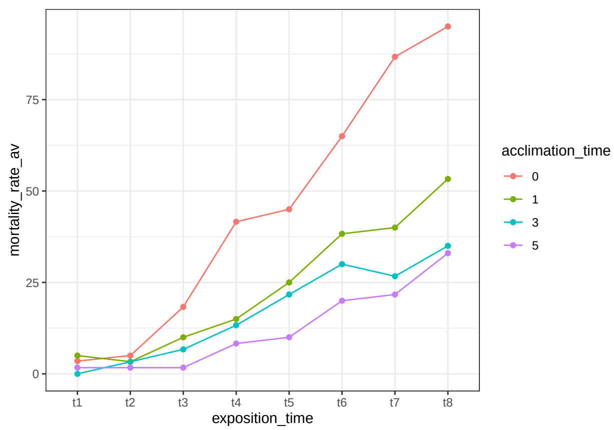

Figure 2: Plot at temp 43

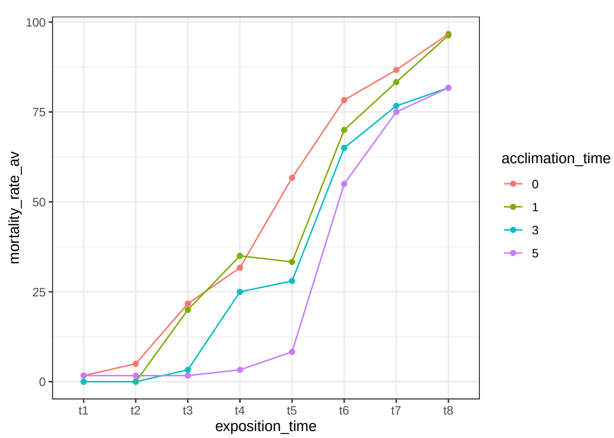

Figure 3: Plot at temp 55