

Biostatistiques avancées avec R

Tests de comparaison de base

Christophe Lalanne

www.aliquote.org

Synopsis

Éléments de contexte

Principes du test d'hypothèse nulle

Comparaison de deux moyennes

Analyse de la variance

Objectifs

Dans ce cours, on s'intéressera aux tests d'inférence simples visant à comparer des indices de tendance centrale calculés sur des données continues et estimés à partir de données observées à la suite d'un procédé d'échantillonnage.

On considèrera deux échantillons ou plus, indépendants ou non.

L'objectif est de démontrer le principe général de la démarche du test d'hypothèse à partir de quelques statistiques de test rencontrées fréquemment en recherche biomédicale.

Ce cours s'appuie en partie sur l'ouvrage de Everitt & Rabe-Hesketh⁽⁵⁾.

Hypothèse nulle, risque d'erreur

De manière générale, lorsque l'on s'intéresse à un effet particulier, l'idée est de postuler l'absence d'effet (**hypothèse nulle**, H_0) et de chercher à vérifier si les données observées sont compatibles ou non avec cette hypothèse. C'est le principe même de la démarche hypothético-déductive.

Pour réaliser un tel test, il est nécessaire de construire une **statistique de test**, dont la distribution d'échantillonnage est connue (ou peut être approximée) sous H_0 , qui nous permettra de répondre à la question suivante : en supposant qu'il n'existe pas d'effet dans la population, quelle est la probabilité d'observer une statistique de test au moins aussi extrême que celle estimée à partir de l'échantillon choisi aléatoirement dans cette population ?

Si cette probabilité se révèle « suffisamment petite », on conclura qu'il est vraisemblablement peu probable que le résultat observé soit dû simplement au hasard de l'échantillonnage.

Risque de première et deuxième espèce :

Supposons que l'on doive prendre une décision concernant une hypothèse nulle. Le fameux « risque alpha » (type I) est le risque de conclure à tort à l'existence d'un effet alors qu'en réalité ce dernier n'existe pas.

À ce risque est typiquement associé, de manière asymétrique, le risque (type II) de ne pas rejeter H_0 lorsque celle-ci est en réalité fautive ; le complémentaire de ce risque est appelée la puissance.

Ces risques sont effectivement asymétriques :

- dans un essai thérapeutique, si l'on doit décider si un nouveau traitement est meilleur que le traitement courant, on cherche à minimiser le risque d'une mauvaise décision ($\alpha = 0.05$);
- si à la fin de l'essai, on ne met en évidence aucune différences significatives, cela ne signifie pas que les traitements sont équivalents : il existe un risque β qu'il existe une réelle différence entre les deux.

Deux illustrations intuitives

Statisticians are applied philosophers. Philosophers argue how many angels can dance on the head of a needle ; statisticians count them. Or rather, count how many can probably dance. (...) We can predict nothing with certainty but we can predict how uncertain our predictions will be, on average that is. Statistics is the science that tells us how. — Stephen Senn

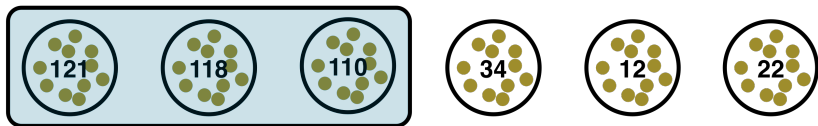
1. Expérience de biologie cellulaire
2. Jeu de pile ou face

Encore du dénombrement

On dispose de 6 lots contenant des cellules en culture (pendant 24h), dont 3 ont reçu un supplément de vitamine E (groupe expérimental). Après 10 jours, on examine les auto-radiographies pour dénombrer le nombre total de cellules dans chaque lot.

Le technicien qui apporte les résultats indique au chercheur que les étiquettes permettant d'identifier quels lots ont été traités ont été égarées⁽⁶⁾...





Si les trois premiers lots correspondent au groupe traité à la vitamine E, alors *a priori* l'expérience semble concluante : quel que soit le lot, le nombre de cellules apparaît largement supérieur à n'importe lequel des trois derniers lots.

Est-il possible d'évaluer la **plausibilité** d'un tel résultat ?

De la nécessité de se comparer

- Il nous faut un moyen de comparer l'effet de l'adjonction de vitamine E par rapport à la situation où les lots ne sont pas traités.
- Un test statistique judicieusement choisi nous permettra de tester l'in vraisemblance d'une hypothèse appelée « hypothèse nulle ».

Comme dans un essai clinique, on cherche donc à comparer l'effet d'un traitement A à celui d'un traitement B de référence (ou placebo), qui sert de comparateur. La différence observée entre les effets de A et B peut faire l'objet d'un **test statistique**. Celui-ci permet de confronter la valeur observée à celles pouvant résulter de simples **fluctuations d'échantillonnage** (c.a.d. des différences dûes au « hasard », signifiant ici l'absence d'une réelle différence).

Si la différence observée est suffisamment grande, et on considérera que c'est le cas s'il y a moins de 5 % de chance d'observer un résultat aussi extrême, alors on conclura que celle-ci ne peut vraisemblablement pas être expliquée par de simples fluctuations d'échantillonnage et que les données observées ne sont pas compatibles avec l'hypothèse nulle d'absence d'effet, appelée H_0 .

On rejettera donc H_0 si la probabilité p d'observer, du seul fait du hasard, une différence au moins aussi grande que celle observée entre les effets de A et B est inférieure à 5 %. Le « petit p » représente cette probabilité et est appelé degré de signification.

En somme, on accepte de se tromper dans 5 % des cas en rejetant l'hypothèse d'absence de différence. D'un autre côté, il existe un risque β de ne pas être en mesure de rejeter l'hypothèse nulle lorsqu'une réelle différence existe. Le complément $1 - \beta$, appelée **puissance du test**, représente donc la probabilité de rejeter correctement l'hypothèse nulle en faveur de l'hypothèse alternative.

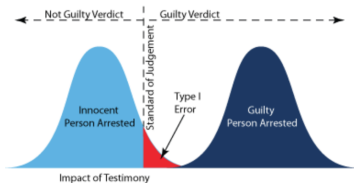


figure 3. Distribution of possible witnesses in a trial showing the probable outcomes with a single witness if the accused is innocent or obviously guilty..



figure 4. Distribution of possible witnesses in a trial showing the probable outcomes with a single witness if the accused is innocent or not clearly guilty..

Source : <http://www.intuitor.com/statistics/T1T2Errors.html>

En résumé

1. Définir une hypothèse nulle (H_0), une hypothèse alternative, et les risques associés à la prise d'une décision concernant le résultat observé à partir d'un échantillon.
2. Choisir une statistique de test, S .
3. Calculer la valeur de S .
4. Définir la distribution d'échantillonnage de S sous H_0 .
5. Conclure à partir de cette distribution.

Mise en œuvre du test d'hypothèse

Soit H_0 : « la vitamine E ne modifie pas la croissance des cultures ». Sous H_0 , les étiquettes « traité » ou « non traité » n'apportent aucune information du point de vue de la mesure considérée (tous les lots sont « échangeables »).

Il y a $\binom{6}{3} = 20$ manières de définir un groupe composé de 3 éléments pris parmi 6. Considérons la somme de l'ensemble des cellules développées dans les 3 lots définissant un même groupe. Appelons-la s . Ici, $s_{obs} = 121 + 118 + 110 = 349$.

Quelles sont les valeurs possibles de s lorsque l'on recombine les lots pour former deux groupes indépendants ?

Parmi les 20 résultats, voici les 3 premiers et les 3 derniers :

1	121	118	110	349
2	121	118	34	273
3	121	118	12	251
...
18	110	34	22	166
19	110	12	22	144
20	34	12	22	68

En fait, s prend les valeurs :

349 273 251 261 265 243 253 167 177 155 262 240 250 164
174 152 156 166 144 68

Conclusion : Parmi les 20 résultats possibles, le résultat $s_{\text{obs}} = 349$ est le plus extrême et il y a exactement $1/20 = 5\%$ de chances d'observer un résultat aussi extrême.

Il est donc peu probable que les résultats observés (« les trois premiers lots sont ceux qui ont été traités ») puissent s'expliquer simplement par les fluctuations d'échantillonnage.

Pile ou face ?

On lance une pièce 10 fois et on observe la séquence de résultats suivants :

P P P P F F F P F P

Question générale : la pièce est-elle truquée ?



Si l'on suppose une pièce bien équilibrée et des lancers indépendants, le nombre attendu de « Face » est $10 \times 0,5 = 5$. La fréquence observée de « Face » dans l'expérience est de $4/10 = 0.4$.

Nous pouvons formuler une hypothèse nulle selon laquelle $p = 0,5$, et l'hypothèse alternative est $p \neq 0,5$. En utilisant un **test binomial**, il est possible de vérifier si la proportion observée diffère de celle attendue théoriquement, en considérant un risque de 5 % de prendre une mauvaise décision en rejetant l'hypothèse nulle.

```
> binom.test(4, 10)
```

```
Exact binomial test
```

```
data: 4 and 10
```

```
number of successes = 4, number of trials = 10, p-value = 0.8
```

```
alternative hypothesis: true probability of success is not equal to 0.5
```

```
95 percent confidence interval:
```

```
0.122 0.738
```

```
sample estimates:
```

```
probability of success
```

```
0.4
```

Le résultat suggère que cette séquence de Pile/Face n'est pas incompatible avec l'hypothèse d'équi-distribution des deux côtés de la pièce.

Autres cadres de raisonnement

In sum, the idea is to confront a single hypothesis with the data, through a designed experiment, with falsification as the only « truth. » This approach follows from Popper's philosophical development and was implemented by Fisher, and Neyman & Pearson's **NHST framework**⁽⁷⁾

We would rather like to know $P(H_0 \mid \text{data})$ than $P(|S| > |s|)$ under the null—even if **the earth is round** ($p < .05$).

- **Likelihood approach** : Use the data to arbitrate between two models. Given the data and a mathematical formulation of two competing models, we can ask, « How likely are the data, given the model ? »
- **Bayesian approach** : Use external information that allows to judge *a priori* which model is more likely to be true, i.e. use a prior probability that can be « updated » to yield a posterior probability, given the data.

References : ⁽²⁾ for a review in biomedical research, and [A Good P-value is Hard to Find: Why I'm a Bayesian When Time Allows](#) (FE Harrell Jr, 2013).

Comparer deux moyennes

Le test de Student est un test paramétrique permettant de tester si deux moyennes de groupe (indépendants ou appariés) peuvent être considérées comme significativement différentes en considérant un seuil d'erreur α .

Exemples d'application : comparer un dosage biologique entre deux groupes de patients, comparer des mesures physiologiques avant et après traitement chez les mêmes patients.

Conditions d'application : normalité des distributions parentes, homogénéité des variances, indépendance (dans le cas du test t pour échantillons indépendants).

Formalisation

Le test t de Student

La statistique de test est définie comme suit :

$$t_{\text{obs}} = \frac{\bar{x}_1 - \bar{x}_2}{s_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

où les \bar{x}_i et n_i sont les moyennes et effectifs des deux échantillons, et $s_c = [((n_1 - 1)s_{x_1}^2 + (n_2 - 1)s_{x_2}^2) / (n_1 + n_2 - 2)]^{1/2}$ est la *variance commune*.

Sous H_0 , cette statistique de test suit une loi de Student à $n_1 + n_2 - 2$ degrés de liberté. Un intervalle de confiance à $100(1 - \alpha)\%$ pour la différence $\bar{x}_1 - \bar{x}_2$ peut être construit comme suit :

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha, n_1 + n_2 - 2} s_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

avec $P(t < t_{\alpha, n_1 + n_2 - 2}) = 1 - \alpha/2$.

Application

Données sur les poids à la naissance⁽⁸⁾.

```
> data(birthwt, package="MASS")
> birthwt$smoke <- factor(birthwt$smoke,
+                          labels=c("No", "Yes"))
> t.test(bwt ~ smoke, data=birthwt, var.equal=TRUE)
```

Two Sample t-test

data: bwt by smoke

t = 3, df = 200, p-value = 0.009

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

72.8 494.8

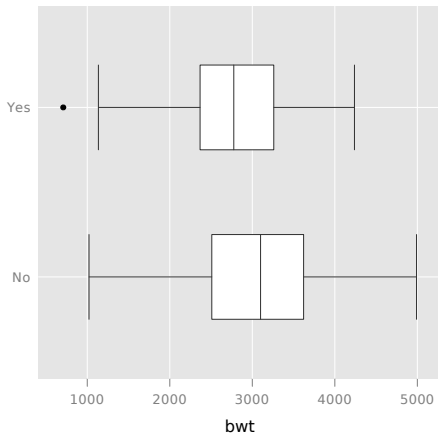
sample estimates:

mean in group No mean in group Yes

3056

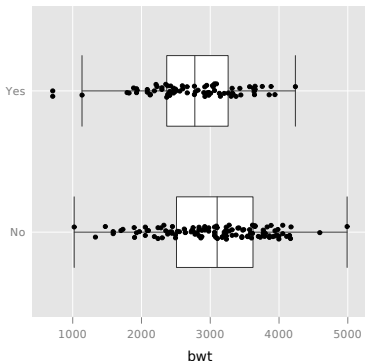
2772


```
> bwplot(smoke ~ bwt, data=birthwt)
```



```
> ## qqmath(~ bwt, data=birthwt, group=smoke)
```

```
> bwplot(smoke ~ bwt, data=birthwt,  
+        panel = function(x, y, ...) {  
+          panel.bwplot(x, y, ...)  
+          panel.points(x, jitter(as.numeric(y),  
+                             amount = .05), ...) })
```



Cas des données non indépendantes

Dans le cas où les échantillons ne sont pas indépendants, p. ex. sujets mesurés à deux reprises ou données appariées, le même type de test peut être utilisé, et la différence de moyennes est comparée à 0. On se retrouve dans la même situation que le **test de Student pour un échantillon** (comparaison à 0).

Le plus souvent ce sont les mêmes unités statistiques qui servent de support à la comparaison, mais n'importe quelle forme d'appariement, pourvu qu'elle fasse sens, peut justifier le recours à un test apparié.

Formalisation

Le test t de Student pour données appariées

Si l'on note les valeurs observées pour la i ème paire x_{1i} et x_{2i} , la différence $d_i = x_{1i} - x_{2i}$ est supposée suivre une loi normale.

L'hypothèse nulle est $\mu_d = 0$. La statistique de test est donnée par :

$$t = \frac{\bar{d}_i}{s_d/\sqrt{n}}$$

où \bar{d}_i la différence moyenne, et s_d est la déviation standard de d_i .

Ici, on peut exploiter la corrélation intra-unité car si X_1 et X_2 ne sont pas indépendants, alors on sait que

$\mathbb{V}(X_1 - X_2) = \mathbb{V}(X_1) + \mathbb{V}(X_2) - 2\text{Cov}(X_1, X_2)$. Cette statistique de test suit alors une loi de Student à $n - 1$ degrés de liberté. Un intervalle de confiance à $100(1 - \alpha)\%$ pour la différence d_i peut être construit comme suit :

$$\bar{d}_i \pm t_{\alpha, n-1} s_d/n,$$

avec $P(t < t_{\alpha, n-1}) = 1 - \alpha/2$.

Application

Effet de somnifères sur le temps de sommeil⁽¹⁰⁾.

```
> t.test(extra ~ group, data=sleep, paired=TRUE)
```

Paired t-test

```
data: extra by group
```

```
t = -4, df = 9, p-value = 0.003
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-2.46 -0.70
```

```
sample estimates:
```

```
mean of the differences
```

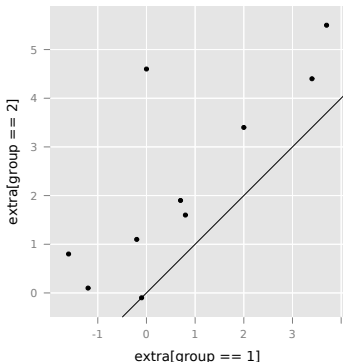
```
-1.58
```

On constate une diminution de la durée de sommeil lorsque le médicament 1 est administré ($\Delta = -1.6$ h $[-2.5; -0.7]$, $p < 0.01$).

Application

On peut utiliser un diagramme de dispersion, de Tukey (`tmd()`) ou de Bland-Altman⁽¹⁾.

```
> xyplot(extra[group==2] ~ extra[group==1], data = sleep,  
+        type = c("p", "g"), aspect = "iso",  
+        abline = list(a = 0, b = 1))
```



Alternative non paramétrique

Dans le test t, on fait une hypothèse sur la nature de la distribution de la réponse mesurée dans la population parente. Si l'on relaxe cette hypothèse et que l'on exige simplement que les échantillons proviennent de populations ayant des distributions à peu près comparables en terme de forme, alors on peut utiliser des tests dits non-paramétriques. Dans le cadre de la comparaison de la tendance centrale de deux échantillons, l'alternative au test t est le **test de Wilcoxon** (`wilcox.test()`).

Ce type de test est souvent utilisé dans le cas des petits effectifs lorsque les données disponibles ne permettent pas réellement de vérifier la normalité des distributions parentes, comme dans le test t. On retiendra toutefois que le test de Wilcoxon a une puissance relative de 80 % par rapport au test de Student. Ce type de test présente en outre l'avantage d'être moins sensible à la présence de valeurs extrêmes.

Formalisation

Le test des rangs de (Mann-Whitney-)Wilcoxon

L'hypothèse nulle est que les deux échantillons comparés proviennent de deux distributions ayant le même paramètre de position. La statistique de test est construite comme la plus petite somme des rangs d'un des deux échantillons ; quand $n_1, n_2 > 15$ et qu'il n'y a pas d'ex-æquo, on a l'approximation suivante pour la statistique de test :

$$Z = \frac{S - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \sim \mathcal{N}(0; 1),$$

où S est la statistique de test pour l'échantillon de taille n_1 .

Pour deux échantillons appariés, le test des rangs signés est utilisé : on calcule la somme T^+ des rangs des différences $z_i = x_{1i} - x_{2i}$ en valeurs absolues positifs ; dans le cadre asymptotique, la statistique de test correspondante est : $Z = \frac{T^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \sim \mathcal{N}(0; 1)$.

Exercices

1. Simuler répétitivement une différence de moyennes empiriques pour deux échantillons tirés aléatoirement parmi des populations de paramètres (moyenne et variance) fixes, et représenter graphiquement la distribution de ces différences de moyennes (`rnorm()`, `replicate()`).
2. À partir de l'étude « Framingham »^(9,4) (`Framingham.csv`), réaliser un test de Student comparant la pression systolique moyenne entre les personnes ayant un âge inférieur ou égal à l'âge médian *versus* celles ayant un âge supérieur à l'âge médian.
3. Faire une représentation graphique des données. Comparer avec un test de Wilcoxon.

Comparer plus de deux moyennes

L'ANOVA constitue une extension naturelle au cas où plus de deux moyennes de groupe sont à comparer. Attention, avec k échantillons, l'hypothèse nulle se lit :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k,$$

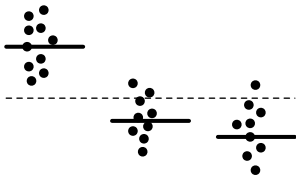
alors que l'alternative est l'existence d'au moins une paire de moyennes qui diffèrent (négation logique de H_0). Si l'on exprime chaque observation comme une déviation par rapport à sa propre moyenne de groupe, $y_{ij} = \bar{y}_i + \varepsilon_{ij}$, on voit que la variabilité totale peut se décomposer comme suit :

$$\underbrace{(y_{ij} - \bar{y})}_{\text{totale}} = \underbrace{(\bar{y}_i - \bar{y})}_{\text{groupe}} + \underbrace{(y_{ij} - \bar{y}_i)}_{\text{résiduelle}}.$$

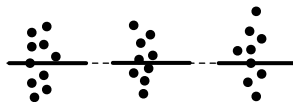
Conditions d'application : normalité des distributions parentes pour chaque groupe, homogénéité des variances, indépendance des observations.

Supposons trois groupes indépendants. Voici quelques scénarios imaginaires concernant la distribution des scores individuels :

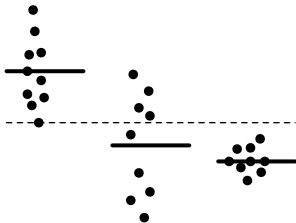
Variances homogènes



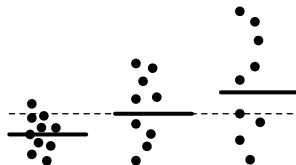
Moyennes constantes



Variances non-homogènes



Variance dépendant de la moyenne



Décomposition de la variance

On peut résumer les sources de variations observées au niveau de la variable réponse comme suit :

Source	SC	dl	CM	F
Inter (B)	$\sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$	$k - 1$	$SC_b / dl (1)$	$(1)/(2)$
Intra (W)	$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$	$N - k$	$SC_w / dl (2)$	
Total	$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$	$N - 1$		

(SC = somme de carrés, CM = carré moyen)

On rappelle qu'une variance se calcule comme la somme des écarts quadratiques à la moyenne, i.e. $\sum_j (y_j - \bar{y})^2 / (n - 1)$.

La SC_b représente la variance des moyennes de groupe (fluctuation autour de la moyenne générale), et la SC_w (« résiduelle ») représente la moyenne des variances des mesures individuelles autour de la moyenne de groupe (dans les deux cas, les moyennes sont pondérées).

Application

Polymorphisme et gène du récepteur estrogène⁽⁴⁾.

```
> library(foreign)
> d <- read.dta("../data/polymorphism.dta")
> head(d)
```

```
  id age genotype
1  1  43  1.6/1.6
2  2  47  1.6/1.6
3  3  55  1.6/1.6
4  4  57  1.6/1.6
5  5  61  1.6/1.6
6  6  63  1.6/1.6
```

```
> xtabs(~ genotype, data = d)
```

```
genotype
1.6/1.6 1.6/0.7 0.7/0.7
      14      29      16
```

```
> summary(d)
```

id	age	genotype
Min. : 1.0	Min. :33.0	1.6/1.6:14
1st Qu.:15.5	1st Qu.:50.0	1.6/0.7:29
Median :30.0	Median :62.0	0.7/0.7:16
Mean :30.0	Mean :60.6	
3rd Qu.:44.5	3rd Qu.:72.0	
Max. :59.0	Max. :87.0	

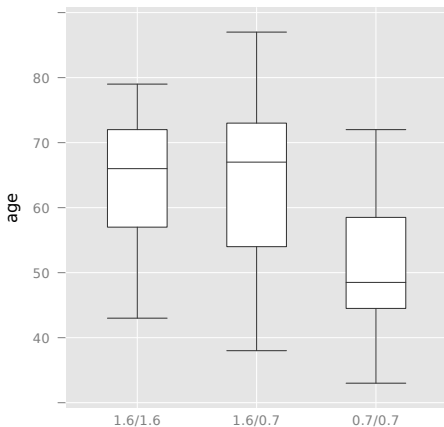
```
> fm <- age ~ genotype
```

```
> aggregate(fm, data = d, mean)
```

	genotype	age
1	1.6/1.6	64.6
2	1.6/0.7	64.4
3	0.7/0.7	50.4

```
> ## aggregate(fm, data = d, Hmisc::smean.sd)
```

```
> bwplot(fm, data = d)
```



```
> m <- aov(fm, data = d)
```

```
> summary(m)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
genotype	2	2316	1158	7.86	0.00098 ***
Residuals	56	8246	147		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Le test F indique que l'on peut rejeter l'hypothèse nulle d'absence de différence d'âge entre les trois génotypes. Il est possible d'utiliser le rapport $2316 / (2316 + 8246) = 0,219$ (η^2) pour quantifier la **part de variance expliquée**. On pourrait poser la question de savoir quelles paires de moyennes diffèrent significativement (comparaisons post-hoc).


```
> model.tables(m)
```

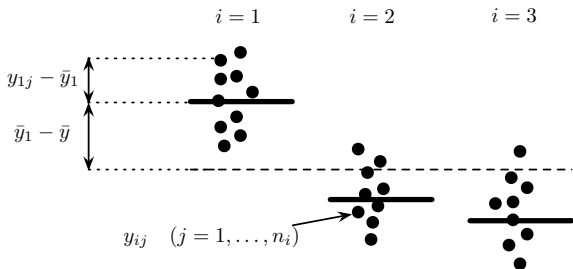
Tables of effects

genotype

```
1.6/1.6 1.6/0.7 0.7/0.7
```

```
3.999 3.735 -10.27
```

```
rep 14.000 29.000 16.00
```

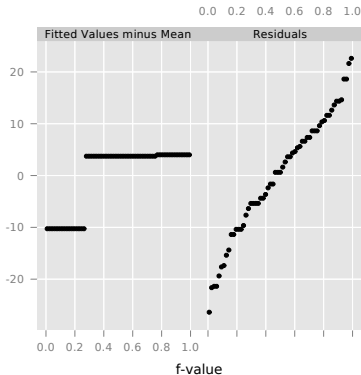


Diagnostic du modèle

Essentiellement, il faut vérifier la distribution des résidus

- à l'aide d'un QQ plot (**normalité des résidus**);
- selon les valeurs prédites (**homoscédasticité**).

> rfs(m)



Représentation graphique

Les diagrammes en barres ne sont pas très économes en termes de « data-ink ratio ». On peut leur préférer des diagrammes en points⁽³⁾, générés à partir de `dotplot()` ou `segplot()` (`latticeExtra`).

Idée : représenter l'âge moyen par génotype et un indicateur de la variabilité autour de cet âge moyen.

Quel indicateur choisir ?

Exercices

1. Calculer des intervalles de confiance à 95 % pour les 3 moyennes de groupe en utilisant la variance commune de l'ANOVA ou d'un test de Student. Comparer avec les IC à 95 % reposant sur une loi normale (`qnorm(0.975)`).
2. Réaliser un test de Student pour comparer les deux groupes homozygotes et comparer le résultat à une ANOVA incluant le facteur genotype mais restreint à ces deux mêmes groupes (cf. `subset=`).
3. Réaliser des comparaisons de l'ensemble des paires de moyenne à l'aide de tests de Student. Arrive-t-on à la même conclusion concernant les différences observées. Comparer avec les résultats produits par `pairwise.t.test()`.

Références I

1. DG Altman and JM Bland. Measurement in medicine : the analysis of method comparison studies. *Statistician*, 32 :307–317, 1983.
2. D Ashby. Bayesian statistics in medicine : a 25 year review. *Statistics in Medicine*, 25(21) :3589–3631, 2006.
3. WS Cleveland. *Visualizing Data*. Hobart Press, 1993.
4. WD Dupont. *Statistical Modeling for Biomedical Researchers*. Cambridge University Press, 2ème édition, 2009.
5. B Everitt and S Rabe-Hesketh. *Analyzing Medical Data Using S-PLUS*. Springer, 2001.
6. P Good. *Permutation, Parametric, and Bootstrap Tests of Hypotheses*. New York, Springer, 2005.
7. R Hilborn and M Mangel. *The ecological detective. Confronting models with data*. Princeton University Press, 1997.
8. D Hosmer and S Lemeshow. *Applied Logistic Regression*. New York : Wiley, 1989.
9. D Levy. 50 years of discovery : Medical milestones from the national heart, lung, and blood institute's Framingham Heart Study. Hackensack, N.J. : Center for Bio-Medical Communication Inc., 1999.
10. Student. The probable error of a mean. *Biometrika*, 6(1) :1–25, 1908.

Index des commandes

aggregate, 38
aov, 40
bwplot, 25, 26, 39
dotplot, 43
factor, 24
head, 37

library, 37
model.tables, 41
pairwise.t.test,
44
qnorm, 44
qqmath, 25

read.dta, 37
replicate, 33
rfs, 42
rnorm, 33
segplot, 43
subset, 44

summary, 38, 40
t.test, 24, 29
tmd, 30
xyplot, 30