

Statistique et aide à la décision 2023-2024

Evaluation - Session 1 (durée : 1h)

13 juin 2024

Nom Prénom :

Tous les documents sont autorisés. La calculatrice est autorisée. L'usage du téléphone portable est interdit, même pour sa fonction calculatrice.

Exercice - Conceptions de l'apprentissage

Comprendre les motivations d'un.e étudiant.e pour apprendre est un enjeu important pour aider les enseignant.e.s à concevoir des cours prenant en compte ses attentes. Cette question peut être abordée par une approche psychométrique : celle-ci s'appuie sur une *échelle*, c'est-à-dire une liste d'assertions, les *items*, pour lesquelles il est demandé aux étudiant.e.s de donner leur degré d'accord sur une échelle allant de 1 (pas d'accord) à 6 (complètement d'accord). Ces *items* sont associés à des conceptions de l'apprentissage préalablement identifiées par les chercheur.e.s en psychologie. Ainsi, par exemple, un fort degré d'accord à l'item 'Apprendre consiste à ingurgiter des faits et des informations' (item B4) associe l'étudiant.e à la conception suivante de l'apprentissage : 'Apprendre, c'est acquérir des informations'.

Les étudiant.e.s du cursus ingénieur agronome de l'Institut Agro Rennes Angers, au niveau L3 en 2017-2018, ont répondu (108 répondant.e.s) à une échelle dédiée à la mesure des conceptions de l'apprentissage, comportant 29 *items*. Elles.ils ont par ailleurs donné leur degré d'accord avec d'autres *items* caractérisant notamment leurs attentes vis-à-vis des modalités d'évaluations. Il est ainsi possible de mesurer leur préférence pour des 'examens qui me permettent de montrer que j'ai réfléchi au cours par moi-même' (correspondant, selon les psychologues, à une attente d'apprentissage **en profondeur**) ou pour des 'examens qui portent uniquement sur le contenu des cours' (correspondant, toujours selon les psychologues, à une attente d'apprentissage **superficiel**).

Par exemple, le tableau de contingence suivant donne la répartition des 108 étudiant.e.s par préférence pour des évaluations correspondant à un apprentissage en profondeur ou superficiel selon leur degré d'accord avec l'item B4 :

| | 1 | 2 | 3 | ≥ 4 |
|-------------|----|----|----|----------|
| Profondeur | 28 | 23 | 16 | 13 |
| Superficiel | 3 | 6 | 6 | 13 |

Question 1

Si le degré d'accord avec l'item B4 était indépendant de la préférence des étudiant.e.s pour un type d'évaluation, quel devrait être le nombre d'étudiant.e.s au croisement de la 2ème ligne (Superficiel) et de la 1ère colonne (accord avec item B4 = 1) du tableau ci-dessus ? (seul le mode de calcul compte, pas la valeur numérique du résultat).

Réponse

Les commandes R ci-après implémentent un test statistique sur le tableau de contingence ci-dessus :

```
chi2_test <- chisq.test(tab,correct=FALSE)
chi2_test

Pearson's Chi-squared test

data:  tab
X-squared = 12.54, df = 3, p-value = 0.00574
```

Question 2

Que déduisez-vous du test ci-dessus ? (Formulez une réponse compréhensible pour un.e non-statisticien.ne - Evitez 'On rejette/ne rejette pas l'hypothèse nulle.')

Réponse

Les tableaux ci-après aident à compléter la décision prise en réponse à la question 2 :

```
round(chi2_test$expected,digits=2)
```

| | B4 | | | |
|-------------|-------|-------|------|-------|
| Pref_Exam | 1 | 2 | 3 | >=4 |
| Profondeur | 22.96 | 21.48 | 16.3 | 19.26 |
| Superficiel | 8.04 | 7.52 | 5.7 | 6.74 |

```
round(chi2_test$residuals,digits=2)
```

| | B4 | | | |
|-------------|-------|-------|-------|-------|
| Pref_Exam | 1 | 2 | 3 | >=4 |
| Profondeur | 1.05 | 0.33 | -0.07 | -1.43 |
| Superficiel | -1.78 | -0.55 | 0.12 | 2.41 |

Question 3

D'après le tableau ci-dessus, comment décririez-vous la dépendance entre l'accord avec l'item B4 et la préférence pour des évaluations en profondeur/superficielles ? (la réponse peut tenir en une phrase)

Réponse

On cherche maintenant à savoir si le degré d'accord avec l'item B4 permet de **prédire** la préférence d'un.e étudiant.e pour un type d'évaluation. Dans la suite, on appelle x le degré d'accord avec cet item, considéré comme une variable quantitative. La préférence pour un type d'évaluation, notée Y , est une variable qualitative à deux modalités, codées '0' pour 'en profondeur' et '1' pour 'superficiel'.

Question 4

Donnez le nom usuel et l'expression mathématique du modèle permettant de répondre à la problématique ci-dessus.

Réponse

Les commandes R ci-dessous donnent les résultats de l'ajustement du modèle de la question précédente :

```
dta$PrefExam <- factor(Pref_Exam)
mod <- glm(PrefExam~B4,data=dta,family=binomial)
summary(mod)$coefficients
```

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|-----------|------------|----------|-------------|
| (Intercept) | -2.696234 | 0.581940 | -4.63318 | 3.60091e-06 |
| B4 | 0.609617 | 0.184582 | 3.30270 | 9.57602e-04 |

Question 5

Un.e étudiant.e a indiqué que son degré d'accord avec l'item B4 était de 1. Donnez une estimation de la probabilité $\mathbb{P}(Y = 1 \mid X = 1)$ pour que cet.te étudiant.e préfère les 'évaluations superficielles' (seul le mode de calcul compte, pas la valeur numérique du résultat).

Réponse

Lorsque l'on estime cette probabilité $\mathbb{P}(Y = 1 \mid X = x)$ pour toutes les autres valeurs x des degrés d'accord avec l'item B4, on obtient les répartitions suivantes des probabilités estimées par type d'évaluation préférée :

```
proba <- predict(mod,type="response")
proba <- round(proba,digits=2)
table(Pref_Exam,proba)
```

| | proba | | | | | |
|-------------|-------|------|-----|------|------|------|
| Pref_Exam | 0.11 | 0.19 | 0.3 | 0.44 | 0.59 | 0.72 |
| Profondeur | 28 | 23 | 16 | 9 | 4 | 0 |
| Superficiel | 3 | 6 | 6 | 9 | 3 | 1 |

Question 6

On décide de prédire qu'un.e étudiant.e préfère les évaluations superficielles si sa valeur estimée de la probabilité $\mathbb{P}(Y = 1 \mid X = x)$ est supérieure strictement à 0.3. Quelle est la valeur de l'accuracy de cette règle de classification ? (seul le mode de calcul compte, pas la valeur numérique du résultat).

Réponse

Dans la suite, les degrés d'accord avec les *items* sont des variables quantitatives notées x_1, x_2, \dots, x_{29} . Le vecteur de ces 29 degrés d'accord forme un profil noté x . On cherche désormais à prédire la préférence pour un type d'évaluation à partir de toutes ces variables, par un modèle du même type que celui proposé en réponse à la question 4. Les commandes R suivantes estiment les paramètres de ce modèle et en donnent la valeur de la déviance résiduelle :

```
mod <- glm(PrefExam~.,data=dta[,c(1:29,38)],family=binomial)
deviance(mod)

[1] 83.3531
```

Question 7

Quelle est la valeur du critère AIC pour ce modèle ? (le mode de calcul et la valeur numérique du résultat sont attendues)

Réponse

On procède maintenant à la recherche du meilleur modèle possible parmi ceux construits à partir d'un sous-sensé des 29 *items* :

```
select_f <- RcmdrMisc::stepwise(mod,direction="forward/backward",
                                criterion="AIC",trace=0)
```

```
Direction: forward/backward
Criterion: AIC
```

```
summary(select_f)
```

```
Call:
```

```
glm(formula = PrefExam ~ B4 + F2 + E3 + F1, family = binomial,
     data = dta[, c(1:29, 38)])
```

```
Coefficients:
```

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|----------|------------|---------|-----------|
| (Intercept) | 0.241 | 1.562 | 0.15 | 0.8776 |
| B4 | 0.640 | 0.196 | 3.26 | 0.0011 ** |
| F2 | -0.541 | 0.233 | -2.33 | 0.0200 * |
| E3 | -0.445 | 0.253 | -1.76 | 0.0787 . |
| F1 | 0.353 | 0.212 | 1.66 | 0.0963 . |

```
---
```

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

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 123.61  on 107  degrees of freedom
Residual deviance: 102.90  on 103  degrees of freedom
AIC: 112.9
```

Number of Fisher Scoring iterations: 4

Question 8

Pourquoi le résultat de la recherche ci-dessus n'est pas le modèle construit à partir des 29 items ?

Réponse

La table d'analyse de la déviance ci-dessous compare le sous-modèle sélectionné à celui construit à partir des 29 items :

```
anova(select_f,mod,test="Chisq")
```

Analysis of Deviance Table

Model 1: PrefExam ~ B4 + F2 + E3 + F1

Model 2: PrefExam ~ A5 + C4 + B2 + A3 + F6 + D4 + E5 + C3 + C2 + E4 +
F2 + A2 + A4 + E2 + B1 + E3 + D3 + C1 + D1 + D2 + A1 + B4 +
F4 + E1 + F5 + F3 + C5 + F1 + B3

| | Resid. Df | Resid. Dev | Df | Deviance | Pr(>Chi) |
|---|-----------|------------|----|----------|----------|
| 1 | 103 | 102.90 | | | |
| 2 | 78 | 83.35 | 25 | 19.55 | 0.77 |

Question 9

Donnez une formulation compréhensible par un.e non-statisticien.ne de l'hypothèse nulle que l'on teste ci-dessus. En quoi peut-on dire que le sous-modèle sélectionné est suffisant ?

Réponse