Devoir2\_réponses

Vincent Gagnon

30 septembre 2015

Ce deuxième devoir a été réalisé avec la base de données ImpactPromo\_900.sav et l'analyse a été faite avec Rstudio. Ce présent document .docx a été produit à l'aide de RStudio et le R markdown.

Dans ce devoir, on suppose que la population est la base de données contenant les 64 000 clients et que la base de données contenant 900 clients est l'échantillon. Toutes les questions sont considéré avec un seuil de 0.05.

# Numéro 1

*Est-ce que l'échantillon est représentatif de la population pour ce qui est de la répartition des clients dans les 5 régions ?*  
*Vous devez utiliser la base de données contenant les 64 000 observations pour obtenir l'information au niveau de la population.*

On pose les hypothèse suivantes :

H0 : Les proportions de l'échantillon représentent bien les proportions de la population.  
H1 : Les proportions de l'échantillon ne représentent pas les proportions de la population.

En utilisant le test de Khi-Deux, tel qu'illustrer dans la Figure 1, on constate que : x2 =4.7258 < 9.4877$  
On ne rejette donc pas H0. On considère le test du Khi-deux fiable étant donné que toutes les fréquences sont supérieur à 5.

# Numéro 2

*Est-ce que la variable Collections\_speciales a un impact sur la visite du site dans la semaine suivant l'envoi des offres ?*  
*Étudiez cette relation selon les étapes suivantes, en justifiant chacune de vos réponses :*

On pose les hypothèse suivantes :

H0 : Dans la population, les collections spéciales n'ont pas d'impact sur les visites du site web.  
H1 : Dans la population, les collections spéciales ont un impact sur les visites du site web.

## Numéro 2 a)

*Testez s'il y a un lien, et dites si ce test est valide.*

Dans la figure 2, on peut voir la table de contingence dont nous nous servirons pour ce numéro.

Étant donnée que la p-value est de 6.596928e-15 , ce qui est nettement inférieur au seuil de 0.05, on rejette H0

Donc, au risque de se tromper 1 fois sur 20, on peut affirmer que les collections spéciales on un impact sur les visites du site web.

## Numéro 2 b)

*Quelle est la force du lien ?*

Dans la Figure 3, on peut voir que le V de Cramer est inférieur à 0.30 donc on peut considérer la relation comme faible.

## Numéro 2 c)

*Faites l'interprétation du tableau, mais seulement pour la ligne du oui (donc la ligne qui correspond aux clients ayant visité le site)*

*Analyse descriptive selon la Figure 2*  
71,5% des gens ont visité le site web alors qu'ils n'ont pas acheté de la marchandise des trois collections spéciales. Ensuite, 13.9% des gens ont visité le site et ont profité une fois des collections spéciales. Finalement, autant de gens on visité le site web, qu'ils aient profité 2fois ou 3 fois des collections spéciales, soit 7,3%.

*Analyse de mise en évidence selon la Figure 2*  
Les visiteurs n'ayant pas acheté de la marchandise des trois collections spéciales sont une absences **marquée** sur le site internet. Cependant, ceux qui on profité une ou 2 fois des collections spécials ont une présence **significative** sur le site web. Les clients qui ont acheté de la marchandise dans les trois collections spéciales ont une présence **marquée** sur les visites du site web.

## Figure 1

## re-encoding from CP1252  
## re-encoding from CP1252

## [1] "Frequence projetté de l'échantillon"  
## Région 1 Région 2 Région 3 Région 4 Région 5   
## 120 214 200 192 174   
## [1] "Total: 900"  
## [1] "Frequence de la population"  
## Région 1 Région 2 Région 3 Région 4 Région 5   
## 134.4797 205.2984 179.8594 199.3641 180.9984   
## [1] "Total: 900"  
## [1] "Loi de Khi-deux"  
## [1] "Pour chaque région on obtient un écart de :"  
## Région 1 Région 2 Région 3 Région 4 Région 5   
## 1.5590559 0.3688152 2.2553441 0.2720120 0.2705997   
## [1] "Total: 4.72582692046385"  
## [1] " X^2 -> p=0,05 et k=4"  
## [1] 9.487729

## Figure 2

## re-encoding from CP1252

##   
## Cell Contents  
## |-------------------------|  
## | Count |  
## | Expected Values |  
## | Column Percent |  
## | Std Residual |  
## |-------------------------|  
##   
## Total Observations in Table: 900   
##   
## | impactEchant$Visite\_site   
## impactEchant$Collections\_speciales | non | oui | Row Total |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Non | 662 | 118 | 780 |   
## | 637.000 | 143.000 | |   
## | 90.068% | 71.515% | |   
## | 0.991 | -2.091 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Une fois sur les trois | 54 | 23 | 77 |   
## | 62.883 | 14.117 | |   
## | 7.347% | 13.939% | |   
## | -1.120 | 2.364 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Deux fois sur les trois | 18 | 12 | 30 |   
## | 24.500 | 5.500 | |   
## | 2.449% | 7.273% | |   
## | -1.313 | 2.772 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## A acheté de la marchandise dans les trois collections spéciales | 1 | 12 | 13 |   
## | 10.617 | 2.383 | |   
## | 0.136% | 7.273% | |   
## | -2.951 | 6.229 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Column Total | 735 | 165 | 900 |   
## | 81.667% | 18.333% | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
##   
##   
## Statistics for All Table Factors  
##   
##   
## Pearson's Chi-squared test   
## ------------------------------------------------------------  
## Chi^2 = 69.11689 d.f. = 3 p = 6.596928e-15   
##   
##   
##   
## Minimum expected frequency: 2.383333   
## Cells with Expected Frequency < 5: 1 of 8 (12.5%)

## Figure 3

## re-encoding from CP1252

##   
##   
## Cell Contents  
## |-------------------------|  
## | N |  
## | Expected N |  
## | N / Col Total |  
## |-------------------------|  
##   
##   
## Total Observations in Table: 900   
##   
##   
## | impactEchant$Visite\_site   
## impactEchant$Collections\_speciales | non | oui | Row Total |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Non | 662 | 118 | 780 |   
## | 637.000 | 143.000 | |   
## | 0.901 | 0.715 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Une fois sur les trois | 54 | 23 | 77 |   
## | 62.883 | 14.117 | |   
## | 0.073 | 0.139 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Deux fois sur les trois | 18 | 12 | 30 |   
## | 24.500 | 5.500 | |   
## | 0.024 | 0.073 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## A acheté de la marchandise dans les trois collections spéciales | 1 | 12 | 13 |   
## | 10.617 | 2.383 | |   
## | 0.001 | 0.073 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
## Column Total | 735 | 165 | 900 |   
## | 0.817 | 0.183 | |   
## ----------------------------------------------------------------|-----------|-----------|-----------|  
##   
##   
## Statistics for All Table Factors  
##   
##   
## Pearson's Chi-squared test   
## ------------------------------------------------------------  
## Chi^2 = 69.11689 d.f. = 3 p = 6.596928e-15   
##   
##   
##   
## [1] "Le v de cramer= 0.277121903054891"