Devoir1\_réponses

Vincent Gagnon

September 20, 2015

Ce premier devoir a été réalisé avec la base de données Devoir1\_Cirque.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.

# Numéro 1

*Quelle est la proportion échantillonnale p des runs qui ont eu lieu aux États-Unis ?*

On peut voir dans la figure 1 qu'au niveau de l'échantillon des runs qui nous ont été fournis, 37.58% des spectacles se font aux États-Unis.  
Cela représente dans l'échantillon 102 runs aux États-Unis.

# Numéro 2

*Donnez l'intervalle de confiance de niveau 90 % associé à la proportion échantillonnale trouvée en (1), et interprétez-le.*

L'intervalle de confiance présenté dans la figure 2 est valide, car on a :

n = 323 > 30  
 np = 102 > 5   
 n(1 - p) = 221 > 5

On peut y découvrir qu'il y a 90% de chances qu'au niveau de la population, la proportion des runs effectué aux États-Unis se situe entre 28.27% et 34.89%

# Numéro 3-

*Quelle est la moyenne échantillonnale x-barre de l'appréciation par rapport aux shows ?*

La moyenne échantillonnale x est de 6,922 tel que présenté dans la figure 3.

# Numéro 4

*Donnez l'intervalle de confiance de niveau 95 % associé à la moyenne échantillonnale trouvée en (3), et interprétez-le*

L'intervalle de confiance présenté dans la figure 4 est valide. car :

n = 323 > 30

Comme l'appréciation est une variable numérique, nous pouvons nous contenter de vérifier la validité avec seulement le n > 30.  
On peut y découvrir qu'il y a 95% de chances qu'au niveau de la population, la moyenne de l'appréciation des shows se situe entre 6.81 et 7.04

# Numéro 5

*On vous dit que plus de 40 % des runs ont lieu en Amérique du Nord. Soutenez-vous cette affirmation au seuil a = 0, 05 ?*

On veut tester l'hypothèse suivante au seuil de a = 0, 05 :

H0 : p (Amérique du Nord) = 0.4  
 H1 : p (Amérique du Nord) < 0.4

Ce test est considéré comme valide, car :

n = 323 > 30  
 np = 129 > 5   
 n(1 - p) = 194 > 5

Pour rejeter l'hypothèse H0, il faudrait que p soit égale ou inférieur à 40%. Hors la proportion des runs en Amérique du Nord de notre échantillon est de 43,34%, tel que présenté dans la figure 6.

Ensuite, en effectuant un t-test sur la nouvelle colonne "AmeriqueN" créer contenant seulement des 0 sauf si la run se déroule en Amérique du Nord, on obtient la 2e partie de la figure 6.  
Étant donné que la p-value est de 2.2e-16 ce qui est nettement inférieur à 2a mais que p est plus grand que pi0, nous pouvons donc considérer H0 comme vraisemblable.

On peut donc soutenir cette affirmation.

# Numéro 6

*On vous dit qu'il y a en moyenne 59 représentations par run. Soutenez-vous cette affirmation au seuil a = 0, 10 ?*

On veut tester l'hypothèse suivante au seuil de a = 0, 10 :

H0 : mu = 59  
 H1 : mu != 59

Ce test est considéré comme valide, car :

n = 323 > 30

Cependant , étant donné que la p-value trouvé dans la figure 6 est inférieur à a = 0, 10, on rejette H0 et on considère H1 comme vraisemblable.

On ne devrait donc pas soutenir l'affirmation disant que la moyenne de représentations par run est de 59.

## Figure 1

## [1] "Nombre de runs total: 323"  
## [1] "Nombre de runs aux États-Unis: 102"

## États-Unis   
## 31.57895

## Figure 2

## p n np n1p left right   
## 0.3158000 323.0000000 102.0034000 220.9966000 0.2826539 0.3489461

## Figure 3

## [1] "Moyenne : 6.9221379389605"

## Figure 4

## moyenne n s np n1p   
## 6.922138 323.000000 1.245269 2235.850554 -1912.850554   
## left right   
## 6.807839 7.036436

## Figure 5

## [1] "Nombre de runs total: 323"  
## [1] "Nombre de runs en Amérique du Nord: 140"  
## [1] "Proportion échantillonnale en Amérique du Nord: 0.43343653250774"

##   
## One Sample t-test  
##   
## data: cirque$AmeriqueN  
## t = 13.885, df = 322, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0.05  
## 95 percent confidence interval:  
## 0.3791062 0.4877669  
## sample estimates:  
## mean of x   
## 0.4334365

## Figure 6

## [1] "La moyenne du nombre de représentations par runs est de 61.1052631578947"

##   
## One Sample t-test  
##   
## data: cirque$NbPerformances  
## t = 2.7916, df = 322, p-value = 0.005558  
## alternative hypothesis: true mean is not equal to 59  
## 90 percent confidence interval:  
## 59.86122 62.34931  
## sample estimates:  
## mean of x   
## 61.10526