

Devoir 7: Analyse de la variance

Visseho Adjiwanou, PhD.

25 November 2020

EXERCICE

QUESTION 1

Supposons que vous souhaitiez en savoir plus sur la manière dont la vie sur le campus en première année influe sur la satisfaction générale à l'égard de l'expérience universitaire. Vous avez un échantillon de 30 personnes: 10 qui vivaient sur le campus, 10 qui vivaient à proximité du campus et 10 qui se rendaient sur le campus. Vous demandez ensuite aux répondants d'indiquer leur satisfaction globale à l'égard de leur expérience universitaire sur une échelle de 1 à 10. Vos données donnent les informations suivantes:

Sommes des carrés intergroupe = 93.3 Somme des carrés intra-groupe = 29,7 Somme totale des carrés = 123,0

1. Quelle est la valeur de la variance intergroupe?
2. Quelle est la valeur de la variance intra-groupe?
3. Quelle est la valeur de F?
4. Quelle est la valeur du ratio de corrélation ?
5. Les résultats sont-ils significatifs au seuil de signification de 0.05? de 0.01?

Remarque: Essayer de répondre à ces questions en créant des objets dans R et en utilisant R comme une calculatrice

QUESTION 2

Utiliser les données de l'enquête canadienne sur le revenu de 2015 pour répondre aux questions suivantes. On se limitera à une base de données réduites de cette base. Les variables retenues dans cette base sont:

- PERSONID: Person identifier (identifiant)
- AGEGRP: Person's age group as of December 31 of reference year (groupe d'âge)
- SEX: Sex (Sexe du répondant)
- MARST: Marital status (Statut matrimonial)
- PROV: Province
- CFATINC: CF - After-tax income (Revenu familial après taxe)

Cette étape est déjà faite pour vous ici.

```
library(tidyverse)
library(summarytools)

cis_2015 <- read_csv("CIS-72M0003-E-2015_F1.csv")

cis_short <-
  cis_2015 %>%
```

```

select(PERSONID, AGEGP, SEX, MARST, PROV, CFATINC) %>%
mutate(Province = case_when(
  PROV == 10 ~ "Terre-Neuve-et-Labrador",
  PROV == 11 ~ "Ile du prince Edouard",
  PROV == 12 ~ "Nouvelle-Écosse",
  PROV == 13 ~ "Nouveau-Brunswick",
  PROV == 24 ~ "Québec",
  PROV == 35 ~ "Ontario",
  PROV == 46 ~ "Manitoba",
  PROV == 47 ~ "Saskatchewan",
  PROV == 48 ~ "Alberta",
  PROV == 59 ~ "Colombie Britannique"
),
region = case_when(
  PROV <= 13 ~ "Atlantique",
  PROV == 24 ~ "Québec",
  PROV == 35 ~ "Ontario",
  PROV >= 46 & PROV <= 48 ~ "Ouest",
  PROV == 59 ~ "Colombie Britannique"
),
sexe = if_else(SEX == 1, "Male", "Female"))

#sort(cis_short$CFATINC)

freq(cis_short$MARST)

```

1. Calculer le revenu moyen dans chaque province, et globalement
2. Quelles sont les tailles des échantillons dans chaque province et globalement?
3. Présenter un diagramme à moustache (boxplot) du revenu selon la province. Commentez.
4. Calculer les statistiques suivantes (vous devez utiliser au besoin la fonction summarise de manière ingénieuse:
 - 4a. la somme des carrés intragroupe (à l'intérieur de chaque province)
 - 4b. la somme des carrés intergroupe (entre les provinces)
5. Déterminer les degrés de liberté intragroupe et intergroupe
6. Calculer la statistique F.
7. Est-ce qu'il existe une association significative entre le revenu et la province de résidence au seuil de significativité de 5% (Donner la valeur du F lu)?
8. Calculer le ratio et corrélation et commenter.
9. Utiliser maintenant la fonction aov de R pour tester la différence de revenu entre les provinces.
10. Finalement, on veut savoir s'il y a inégalité de revenu entre les hommes et les femmes au Canada. Cependant, on ne peut pas utiliser le revenu familial pour répondre à cette question (pourquoi?). On va alors s'intéresser à un groupe particulier, ceux et celles qui sont encore célibataires.
 - 10a. Sélectionner la nouvelle base de données que vous appelez **cis_single** à partir de la base de données cis_short en sélectionnant uniquement les célibataires (voyez le codebook de l'enquête pour vous référer).
 - 10b. calculer le revenu moyen des femmes célibataires et des hommes célibataires.
 - 10c. Utiliser la fonction **t.test** pour tester le fait que le revenu moyen des femmes célibataires est différent du revenu moyen des hommes célibataires

Réponse 3