

## **Worksheet n°2**

**Exercise 1.** To target customers for a new consumer product, a company surveys 321 people. Interest in the product is rated as "no interest", "minor interest" or "major interest". Family situation (at least one dependent child: yes or no) is also noted. The aim is to check whether interest in the product depends on family situation. The results are as follows:

*Pour cibler la clientèle d'un nouveau produit de consommation, une entreprise fait un sondage auprès de 321 personnes. L'intérêt dans le produit est noté par "aucun intérêt", "un intérêt mineur" ou un "intérêt important". La situation familiale (au moins un enfant à charge : oui ou non) est notée également. On cherche à vérifier si l'intérêt dans le produit dépend de la situation familiale. Les résultats sont les suivants :*

| Enfant | No interest | minor | major |
|--------|-------------|-------|-------|
| Yes    | 10          | 12    | 3     |
| No     | 7           | 38    | 9     |

So we have 79 respondents. We want to check whether there is a link between the two measures at the 5% level.

*On a donc 79 personnes qui répondent. On veut vérifier s'il y a un lien entre les deux mesures au niveau 5%.*

**Exercise 2.** A fleet of buses is equipped with 4 types of tire (A, B, C, D). The mileage covered before tire wear is measured. We construct 3 mileage classes (in thousands)  $<20$ ,  $[20,30]$ ,  $>30$ . The following results are obtained:

*Une flotte d'autobus est équipée de 4 types de pneus (A, B, C, D). On mesure le kilométrage parcouru avant usure du pneu. On construit 3 classes de kilométrage (en milliers)  $<20$ ,  $[20,30]$ ,  $>30$ . On a obtenu les résultats suivants :*

| Observations | A          | B          | C          | D          | Total      |
|--------------|------------|------------|------------|------------|------------|
| $<20$        | 26         | 23         | 15         | 32         | <b>96</b>  |
| $[20,30]$    | 118        | 93         | 116        | 121        | <b>448</b> |
| $>30$        | 56         | 84         | 69         | 47         | <b>256</b> |
| <b>Total</b> | <b>200</b> | <b>200</b> | <b>200</b> | <b>200</b> | <b>800</b> |

Are the two variables independent?

*Les deux variables sont-elles indépendantes?*

**Exercise 3.** A researcher wants to check whether two universities have the same rating scale. To do this, he selects a sample of 21,000 students from the two universities and looks at the ratings assigned to students in 2001:

*Un chercheur veut vérifier si deux universités ont un même barème pour l'attribution des côtes. Pour ce faire il choisit un échantillon de 21000 étudiants provenant des deux universités et il regarde les cotes attribuées aux étudiants de 2001 :*

| Rating scale  | A    | B    | C    | D    | E   |
|---------------|------|------|------|------|-----|
| Université I  | 605  | 1400 | 1789 | 300  | 70  |
| Université II | 2014 | 4178 | 8032 | 2005 | 607 |

In fact, we want to check whether the distribution of ratings is dependent on the universities, i.e. whether the variables "university" and "rating" are independent r.v. at the 5% level.

*En fait, on cherche à vérifier si la répartition des cotes est dépendante des universités c'est-à-dire si les variables "universités" et "cote" sont des v.a. indépendantes au niveau 5%.*

**Exercise 4.** IQ is measured for each member of a couple consisting of a father and son. We considered a sample of 12 fathers and their sons.

*On mesure le QI pour chacun des membres d'un couple formé du père et de son fils. Nous avons considéré un échantillon de 12 pères et de leur fils.*

|                           |     |     |     |     |    |     |     |     |     |     |     |     |
|---------------------------|-----|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|
| Couple (father, son) no i | 1   | 2   | 3   | 4   | 5  | 6   | 7   | 8   | 9   | 10  | 11  | 12  |
| Father's IQ $x_i$         | 123 | 144 | 105 | 110 | 98 | 138 | 131 | 90  | 119 | 109 | 125 | 100 |
| son's IQ $y_i$            | 102 | 138 | 126 | 133 | 95 | 146 | 115 | 100 | 142 | 105 | 130 | 120 |

Is there a link between the father's IQ (X) and the son's IQ (Y)?

- If the variables are assumed to be normal,
- If we have no information on the distribution of the two variables.

*Existe-t-il une liaison entre le QI du père (X) et le QI du fils (Y) ?*

- Si on suppose que les variables sont normales,*
- Si on n'a aucune information sur la loi des deux variables.*

**Exercise 5.** The distribution of flight times for the 670 Algiers-Paris flights is shown in the table below:  
*La répartition des durées de vols de 670 vols Alger-Paris est donnée dans le tableau suivant :*

|                 |            |            |            |            |            |            |              |
|-----------------|------------|------------|------------|------------|------------|------------|--------------|
| <b>Duration</b> | [1.9,1.95[ | [1.95,2[   | [2,2.05[   | [2.05,2.1[ | [2.1,2.15[ | [2.15,2.2[ | [2.2,2.25[   |
| <b>Numbers</b>  | 19         | 19         | 39         | 48         | 87         | 94         | 104          |
| <b>Duration</b> | [2.25,2.3[ | [2.3,2.35[ | [2.35,2.4[ | [2.4,2.45[ | [2.45,2.5[ | [2.5,2.55[ | <b>Total</b> |
| <b>Numbers</b>  | 92         | 57         | 44         | 28         | 26         | 13         | <b>670</b>   |

Test the normality of this distribution using:

- The Henry line
- A parametric test
- A non-parametric test

*Tester la normalité de cette distribution en utilisant :*

- La droite de Henry*
- Un test paramétrique*
- Test non paramétrique.*

**Exercise 6.** The number of customers visiting a store was counted for 100 10-minute intervals. The results obtained are summarized in the following table:

*On a compté, pendant 100 intervalles de 10 minutes, le nombre de clients se présentant à un magasin. Les résultats obtenus ont été résumés dans le tableau suivant :*

|                           |    |    |    |    |    |    |    |              |
|---------------------------|----|----|----|----|----|----|----|--------------|
| <b>Number of customer</b> | 5  | 6  | 7  | 8  | 9  | 10 | 11 |              |
| <b>Intervals</b>          | 1  | 0  | 1  | 2  | 1  | 3  | 5  |              |
| <b>Number of customer</b> | 12 | 13 | 14 | 15 | 16 | 17 | 18 |              |
| <b>Intervals</b>          | 6  | 9  | 10 | 11 | 12 | 8  | 9  |              |
| <b>Number of customer</b> | 19 | 20 | 21 | 22 | 23 | 24 | 25 | <b>Total</b> |
| <b>Intervals</b>          | 7  | 5  | 4  | 3  | 1  | 1  | 1  | <b>100</b>   |

Check whether the character follows a Poisson distribution using an appropriate test.

*Vérifier si le caractère suit une loi de Poisson en utilisant un test approprié.*