**Td Analyse de Données**

**Exo1**: On suppose que l’âge auquel apparaissent les premiers mots de vocabulaire chez l’enfant suit la loi normale de moyenne 12 mois et d’écart-type 2; 5 mois.

1. Quelle est la proportion d’enfants pour lesquels les premiers mots apparaissent avant 9 mois ?
2. Quelle est la proportion d'enfants pour lesquels les premiers mots apparaissent après 16 mois?
3. Déterminer l’âge au-dessus duquel 2% des enfants prononcent leurs premiers mots.

**Exo2**:Dans le cadre d’une étude sur la santé au travail, on a interrogé au hasard 500 salariés de différents secteurs et de différentes régions d'Afrique. 145 d’entre eux déclarent avoir déjà subi un harcèlement moral au travail.

**1.** Donner une estimation de cette proportion par un intervalle de confiance à 90%.

**2.** Si avec les mêmes données on calculait un intervalle de confiance à 95%, serait-il plus grand ou plus petit que celui trouvé à la question précédente ?

**Exo3:** En vue de réaliser un programme de rééducation, des chercheurs ont soumis un questionnaire de neuropsychologie cognitive à 150 enfants tirés au sort. Le questionnaire comporte 20 questions et les chercheurs ont recueilli pour chaque enfant le nombre xi de bonnes réponses. Les résultats ainsi obtenus sont tels que : = 1502;= 19486:

1. Donner une estimation ponctuelle du nombre moyen de bonnes réponses dans la population étudiée.
2. Donner une estimation ponctuelle de l’écart-type de la variable.
3. Estimer le nombre moyen de bonnes réponses dans la population par un intervalle de confiance au niveau 99%.
4. Quelle est la marge d’erreur dans l’estimation du nombre moyen de bonnes réponses au niveau 99%?

**Exo4**: L’inventaire de Padoue est un questionnaire portant sur les troubles obsessionnels du comportement (TOC). Chez les adultes dépressifs, le score obtenu à ce questionnaire a pour moyenne 84 avec un écart-type de 35. Des chercheurs s’intéressent alors aux scores moyens observés dans les échantillons de taille 75.

1. Caractériser la distribution de la moyenne empirique du score à l’inventaire de Padoue sur les échantillons de taille 75 (forme et valeur(s) de son/ses paramètre(s)).
2. Quelle est la probabilité d’observer sur un échantillon de taille 75 un score moyen inférieur à 90 ?
3. En dessous de quelle valeur se trouvent 95 % des scores moyens observés sur les échantillons de taille 75 ?
4. Au dessus de quelle valeur se trouvent 95 % des scores moyens observés sur les échantillons de taille75 ?
5. Pour quelle proportion d’échantillons observe-t-on un score moyen compris entre les deux valeurs déterminées aux questions 3 et 4 ?

**Ex05 :** Pour déterminer le taux d’occupation d’un métal coûteux, on emploie la méthode des « observations » : au cours de chaque mois on observe un échantillon d’instants tirés au hasard. Pour chacun de ces instants précis, un contrôleur note si le matériel est ou non utilisé. On a ainsi observé un échantillon des 500 instants au mois de Janvier et 400 instants au mois de Février. Les résultats suivants ont été obtenus : Pour un risque de 5%, existe-t-il une différence significative entre le taux d’occupation en janvier et février ?

|  |  |  |
| --- | --- | --- |
|  | Janvier | Février |
| Occupation | 400 | 300 |
| Inoccupation | 100 | 100 |

**Exo 6** : Peut-on dire que le flux l’entrée à l’université a Niamey ne dépend pas de la catégorie socioprofessionnelles des étudiants pour un risque de 4%.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **IUT** | **Sciences** | **LSH** | **Droit** | **Sc.Eco** | **Médecine** |
| Agriculteurs | 9 | 35 | 44 | 24 | 8 | 13 |
| Indépendants | 66 | 72 | 171 | 122 | 48 | 71 |
| Cadres | 77 | 139 | 380 | 195 | 69 | 233 |
| Prof.intermédiaires | 50 | 78 | 155 | 152 | 57 | 85 |
| Employés | 52 | 86 | 274 | 43 | 26 | 48 |
| Ouvriers | 55 | 103 | 191 | 40 | 25 | 46 |

**Exo7** : Une étude sur la fumée a observé un échantillon de 1000 personnes prises au hasard. Pour chaque personne on a observé les caractéristiques suivantes : F =>Fumeur ; N :=> Non fumeur ; M :=> Malade aux poumons et S :=> Sain. Voici les fréquences observées. Le fait de fumer influence-t-il l’état de santé pour un risque de 5% ?

|  |  |  |
| --- | --- | --- |
|  | F | N |
| M | 200 | 200 |
| S | 135 | 465 |

**Exo8** : On désire étudier l’effet de quatre systèmes de présentation (A1, A2, A3 et A4). Nous voulons tester si les quatre systèmes ont le même effet sur la vente. Aider nous à décider pour un risque de 1%.

|  |  |  |  |
| --- | --- | --- | --- |
| **A1** | **A2** | **A3** | **A4** |
| 120 | 122 | 116 | 112 |
| 118 | 120 | 108 | 114 |
| 122 | 132 | 116 | 122 |
| 110 | 124 | 116 | 122 |
| 130 | 112 | 124 | 130 |
| 120 | 122 | 116 | 122 |

**Exo 9 :**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | L1 | L2 | L3 | L4 |
| P1 | 15 | 15 | 15 | 15 |
| P2 | 25 | 12 | 12 | 16 |
| P3 | 16 | 15 | 24 | 20 |
| P4 | 9 | 15 | 10 | 24 |

Deux amis sont confrontés à un problème d’usage de la langue et du nombre de pays visités. L’un dit que, plus tu visites des pays, plus tu parles de langue. L’autre Dit : Cher ami je peux rester chez moi et apprendre par Internet beaucoup des langues : donc le nombre des langues parlées n’a rien à voir avec le nombre de pays visités !

1. Départagez-les à l’aide d’une méthode d’analyse de données pour un risque de 5%
2. pour quel(s) risque(s) peut-on prendre une décision contraire à celle prise en 1. ?

**Exo 10 :** La note obtenue parles étudiants dépend de leur préparation au devoir, ainsi X désigne les notes obtenues sans préparation et Y les notes obtenues après une laborieuse préparation.

1. Peut-on considérer que les notes obtenues après préparation sont meilleures pour un risque de 5%
2. Pour le même risque peut-on considérer qu’en moyenne les notes sont différentes ?

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 12 | 5 | 15 | 19 | 8 | 7 | 18 | 16 |
| Y | 12 | 15 | 19 | 18 | 8 | 12 | 19 | 15 |

**Exo 11 :** Comparez les moyennes des quatre échantillons E1 ; E2 : E3 et E4

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | E1 | E2 | E3 | E4 |
|  | 15 | 20 | 105 | 39 |
|  | 18 | 90 | 30 | 16 |
|  | 19 | 45 | 26 | 120 |
|  | 9 | 15 | 10 | 24 |

**Exo12:**

1. T suit une loi normale centrée réduite N(0,1). Calculera) P(-1.96 < T < 1.96) ; b) P(T> 1.96) ;P(T< 1.96)
2. X une variable aléatoire normale N(3,2). Calculer P(X <4 ) ; P(-3 < X< 11) et P(X >8).
3. On trouve que la note X obtenue par les étudiants, lors d’un examen suit une loi Normale N(m,σ) telle que P(X >11,9)=0,0869 et P(X < 6,8)=0,2483
4. Déterminer les paramètres m et σ de cette Variable aléatoire
5. Calculer le pourcentage des étudiants ayant plus de 10 de moyenne.
6. Soit Y la loi de Poisson *P*(15), calculer P(Y< 10σ) .

**Exo13:**

1. Déterminer les droites d’ajustement de Y=aX+b et X=a’Y+b’, ainsi que leur coefficient de corrélation r.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 12 | 5 | 15 | 19 | 8 | 7 | 18 | 16 |
| Y | 12 | 15 | 19 | 18 | 8 | 12 | 19 | 15 |

**Exo:14**

Soient 10 individus représentés par le tableau suivant

|  |  |  |
| --- | --- | --- |
| Individus | X | Y |
| I1 | 16,00 | 17,50 |
| I2 | 19,00 | 18,00 |
| I3 | 14,00 | 18,25 |
| I4 | 17,00 | 13,00 |
| I5 | 14,00 | 15,50 |
| I6 | 13,00 | 13,00 |
| I7 | 13,00 | 11,00 |
| I8 | 14,00 | 17,00 |
| I9 | 14,00 | 17,00 |
| I10 | 12,00 | 17,00 |
| I11 | 13,00 | 15,00 |
| I12 | 14,50 | 17,00 |
| I13 | 18,25 | 18,75 |
| I14 | 18,25 | 13,00 |
| I15 | 14,00 | 18,00 |
| I16 | 18,00 | 19,50 |
| I17 | 18,75 | 19,00 |
| I18 | 12,00 | 16,50 |
| I19 | 16,50 | 16,00 |
| I20 | 16,50 | 16,00 |

1. Effectuez l’ajustement de Y par X
2. Comparez les deux moyennes pour un risque α=0.03