

**Contrôle Terminal – 2<sup>ème</sup> session – 28/06/2019 – 2h00**  
**Statistiques – L1 de psychologie**

Veuillez rendre ce sujet et votre copie.

Numéro d'anonymat :

*Le formulaire et la calculatrice sont autorisés. Merci d'indiquer dans la case Numéro d'anonymat, ci-dessus, un numéro que vous reporterez aussi sur votre copie. Vous rendrez l'énoncé et votre copie, et pouvez soit répondre sur l'énoncé, soit détailler certaines questions sur la copie si vous avez besoin de plus de place. Le soin de la rédaction entrera en compte dans la notation mais dans les questions où des détails ne sont pas explicitement demandés, un résultat correct, donné sans détails de calcul sera accepté.*

**Exercice 1 : Chaleur estivale et millésimes viticoles**

On cherche dans cet exercice à déterminer si la qualité d'un millésime est liée aux températures estivales mesurées cette année là. On relève donc les notes attribuées aux vins rouges bourguignons des années 2005 à 2014 sur le site d'une enseigne de vente de vin. De plus, on considère la température quotidienne (mesurée à Dijon au moment le plus chaud de la journée) dont on fait la moyenne sur les mois de juillet et août de ces mêmes années. On obtient les données suivantes :

Année	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Température	25,8 °C	26,8 °C	23,8 °C	24,8 °C	26,5 °C	25,1 °C	24,1 °C	25,6 °C	26,8 °C	23,9 °C
Note attribuée à ce millésime	9	6	5	7	9	7	7	7	6	8

1. Quel est le coefficient de corrélation linéaire entre cette note et la température ?

2. Une fois que l'on connaîtra les températures moyennes de l'été 2019, serait-il pertinent d'utiliser une droite de régression pour en déduire une prédiction de la note qu'obtiendrait le millésime 2019 ? Si oui, quelle droite considérer ?



2. (a) Estimer le temps de lecture moyen du texte modifié pour l'ensemble des enfants de 15 ans.

*Vous calculerez un intervalle de confiance, associé au risque d'erreur  $\alpha = 0,01$ .*

- (b) Conclure : peut-on affirmer, avec la confiance 99%, que les enfants de 15 ans mettent en moyenne plus longtemps à lire le texte modifié que l'original ?

**Exercice 3 : Anxiété en milieu rural ou urbain.**

En France, 20% de la population vit en milieu rural. On souhaite savoir si ces habitants de milieux ruraux sont plus sujets à l'anxiété que le reste de la population française (sachant que chaque année en France, 15% de population environ a un trouble anxieux).

On considère pour cela un échantillon aléatoire de personnes ayant eu un trouble anxieux au cours de l'année 2019. On note par  $X$  le nombre d'individus de l'échantillon qui habitent en milieu rural.

1. On considère un échantillon aléatoire de 62 français-e-s ayant eu un trouble anxieux en 2019. Montrer que si l'on suppose que les habitants de milieux ruraux et urbain ont la même probabilité de développer un trouble anxieux, on peut décrire la loi de  $X$  par une loi binomiale (vous préciserez laquelle).

2. Justifiez que cette loi peut s'approcher par une loi normale, dont vous préciserez la moyenne et l'écart type.

3. Quelle serait alors la probabilité  $\mathbb{P}[1 \leq X \leq 23]$  ?

*On utilisera la loi normale de la question précédente et une correction de continuité*

4. Sur un échantillon aléatoire de 62 français·e·s ayant eu un trouble anxieux en 2019, si l'on constatait qu'il y ait 29 habitants de milieux ruraux, cela serait-il compatible avec l'hypothèse selon laquelle les habitants de milieux ruraux et urbain ont la même probabilité de développer un trouble anxieux ?

#### **Exercice 4 : Genre et goût pour les films d'action**

Pendant une durée d'une semaine, la gérante d'un cinéma décide de collecter des statistiques sur le genre (Femme / Homme) de ses client·e·s. Elle constate alors que parmi les 500 billets vendus au cours de cette semaine, 51% ont été vendus à des femmes et 49% à des hommes.

De plus, elle constate que parmi ces 500 billets vendus, il y en a 160 qui ont été vendus pour des films d'action. Parmi ces billets pour des films d'actions, 45% ont été vendus à des femmes, et 55% à des hommes.

Pour simplifier les questions posées, on suppose que les clients étaient 500 personnes distinctes (c'est à dire qu'aucune personne n'a acheté plusieurs billets cette semaine là).

1. (a) Parmi les femmes qui (au cours de cette semaine) ont acheté un billet dans ce cinéma, combien ont opté pour un film d'action ?

(b) Parmi les hommes (qui ont acheté un billet dans ce cinéma au cours de cette semaine), quelle proportion a opté pour un film d'action ?

2. On souhaite désormais déduire, à partir des observations sur cet échantillon de clients, des proportions au sein de l'ensemble des clients de cinéma en France.

(a) Estimer, parmi l'ensemble des billets de cinéma vendus en France à des hommes, la proportion qui correspondent à des films d'action.

*Vous calculerez un intervalle de confiance, associé à la confiance  $c = 90\%$ .*

(b) Estimer, parmi l'ensemble des billets de cinéma vendus en France à des femmes, la proportion qui correspondent à des films d'action.

*Vous calculerez un intervalle de confiance, associé à la confiance  $c = 90\%$ .*

(c) À partir de ces estimations, peut-on conclure (avec la confiance 90%) que lorsqu'elles vont au cinéma, les femmes ont moins tendance que les hommes à aller voir un film d'action ?