UNIVERSITÉ PARIS OUEST NANTERRE LA DÉFENSE
U.F.R. SEGMI
L2 Économie

Année universitaire 2016 – 2017
Cours de B. Desgraupes

Méthodes Statistiques

Exercices de TD

Séance 1

Exercice 1

Le directeur d'un magasin surveille l'évolution du nombre journalier X de clients : ce nombre fluctue aléatoirement autour de 360 avec un écart-type de 40 selon une loi normale.

1-1) Quelle est la probabilité que le nombre journalier de clients soit inférieur à 300 ?

1-2) Quelle est la probabilité que le nombre journalier de clients soit supérieur à 440 ?

1-3) Quelle est la probabilité que le nombre journalier de clients soit compris entre 300 et 440 ?

Exercice 2

Le débit journalier X en hectolitres d'une station-service est supposé suivre une loi normale de moyenne 100 et d'écart-type 20.

2-1) Quelle est la probabilité que le débit journalier soit inférieur à 130 hl ?

2-2) Quelle est la probabilité que le débit journalier soit inférieur à 70 hl ?

2-3) Quelle est la probabilité que le débit journalier soit compris entre 70 hl et 130 hl ?

2-4) Chaque jour, le gérant commande une certaine quantité d'essence qui est livrée le matin avant l'ouverture de la station-service. Quelle quantité d'essence doit commander chaque jour le gérant pour que la probabilité qu'il ne puisse pas répondre à la demande soit inférieure à $0,1$?

Exercice 3

Une usine utilise une machine automatique pour remplir des flacons contenant un certain produit en poudre. Par suite de variations aléatoires dans le mécanisme, le poids de poudre par flacons est une variable aléatoire normale de moyenne m et d'écart-type $1,1$ mg. Les flacons sont vendus comme contenant 100 mg de produit.

3-1) La machine est réglée sur $m = 101,2$ mg. Quelle est la probabilité que le poids de produit dans un flacon soit inférieur au poids annoncé de 100 mg ?

3-2) Sur quelle valeur de m faut-il régler la machine pour que la probabilité que le poids d'un flacon soit inférieur à 100 mg soit inférieure à $0,4$.

Exercice 4

On suppose que la distribution du quotient intellectuel dans une population d'enfants âgés de 7 ans suit un modèle normal de moyenne 100 et d'écart-type 15.

4-1) Quelle est la proportion d'enfants de la population qui ont un QI inférieur à 106 ?

4-2) Quelle est la proportion d'enfants de la population qui ont un QI compris entre 90 et 110 ?

4-3) Quelle est la proportion d'enfants de la population qui ont un QI compris entre 80 et 120 ?

4-4) Quel est le QI au-dessous duquel se situent 95% des enfants de 7 ans ?

4-5) Quel est le QI au-dessous duquel se situent 25% des enfants de 7 ans ?

Exercice 5

Le test de Boëhm-R (*Boëhm révisé*) est destiné à apprécier chez l'enfant la maîtrise des concepts de base, essentiels pour la compréhension des consignes verbales et fondamentaux pour la réussite scolaire. On suppose que le résultat à ce test suit une distribution normale de moyenne 60 et d'écart-type 17 dans la population des enfants de grandes sections d'une école maternelle.

5-1) Quelle est la proportion d'enfants ayant un résultat inférieur à 70 ?

5-2) Quelle est la proportion d'enfants ayant un résultat inférieur à 30 ?

5-3) Quelle est la proportion d'enfants ayant un résultat compris entre 26 et 94 ?

5-4) Si l'on décidait d'accepter au CP tous les enfants ayant un résultat supérieur à 40, quelle serait la proportion d'enfants admis en CP ?

5-5) Inversement, si l'on décidait d'accepter 99% des enfants de grande section en CP, quel serait le résultat minimum requis pour passer au CP ?

5-6) On désire regrouper les résultats en trois catégories (concepts non maîtrisés, concepts en cours d'acquisition, concepts maîtrisés) contenant chacune un tiers des enfants : calculer les deux valeurs permettant de définir les bornes de ces catégories.

Exercice 6

Dans un centre avicole, des études ont montré que la masse X d'un œuf peut être considérée comme la réalisation d'une variable aléatoire de loi normale $\mathcal{N}(\mu = 55\,;\,\sigma = 3)$.

6-1) Calculer la probabilité que la masse d'un œuf soit inférieure à $59,5$?

6-2) Calculer la probabilité que la masse d'un œuf soit supérieure à $50,5$?

6-3) En dessous de quelle masse se trouvent un tiers des œufs ?

Séance 2

Exercice 7

On considère une population constituée de 4 employés dont les salaires net en euros sont :

Individus	A	B	C	D
Salaire	1200	1000	1400	1200

Imaginons que l'on ne connaisse pas le salaire moyen de ces 4 employés, on cherche à l'estimer en appelant au hasard 2 personnes. Les appels sont effectués par des opérateurs différents, de sorte que l'on peut appeler 2 fois la même personne.

7-1) Décrire la population, l'échantillon, la variable, son type et son/ses paramètre(s).

7-2) Pour chacun des échantillons possibles, calculer la moyenne empirique du salaire.

Échantillon	(A, A)	(A, B)						
Moyenne empirique	1200	1100						

Échantillon								(D, D)
Moyenne empirique								

7-3) Combien d'échantillons donnent une estimation parfaite du salaire moyen ?

7-4) Quelle est la probabilité de tomber sur un échantillon dont la moyenne empirique est supérieure ou égale à 1300 euros ?

7-5) Représenter graphiquement la distribution de la moyenne empirique.

Exercice 8

D'après l'INSEE, l'âge moyen des Français en 2012 est de $40,4$ ans avec un écart-type de 24 ans. On fait l'approximation que la distribution de l'âge des Français suit une loi normale. On s'intéresse à l'âge moyen obtenu sur un échantillon de 55 Français tirés au sort.

8-1) Identifier la population, l'échantillon, la variable, son type et son/ses paramètre(s).

8-2) Caractériser la distribution de l'âge moyen d'un échantillon.

8-3) Quelle est la proportion des Français dont l'âge est inférieur à 20 ans ?

8-4) Quelle est la probabilité d'observer sur un échantillon de taille 55 un âge moyen inférieur à 20 ans ?

8-5) Quelle est la probabilité d'observer sur un échantillon de taille 55 un âge moyen supérieur à 42 ans ?

8-6) Quel est le quantile à 80% pour l'âge moyen dans un échantillon aléatoire de taille 55 ?

Exercice 9

Le test de Fagerström mesure le degré de dépendance tabagique. Le score obtenu à ce test indique le degré de dépendance. Dans une population de fumeurs, le score obtenu à ce questionnaire a pour moyenne 5 et pour écart-type $4,5$. On dispose du score moyen d'un échantillon de 45 fumeurs tirés au sort dans cette population.

9-1) Identifier la population, l'échantillon, la variable, son type et son/ses paramètre(s).

9-2) Déterminer la distribution de la moyenne empirique de tels échantillons, en précisant sa forme, sa moyenne, sa variance et son écart-type.

9-3) Calculer la probabilité d'observer, sur de tels échantillons, un score moyen inférieur à 6.

9-4) Calculer la probabilité d'observer un score moyen inférieur à 3.

9-5) En considérant les scores moyens de tous les échantillons de taille 45, à quelle valeur 90% de ces scores moyens seront-ils inférieurs ?

Exercice 10

Si T est un estimateur du paramètre θ, montrer que

$$E\big((T-\theta)^2\big) = \mathrm{Var}(T) + b(\theta)^2$$

où $b(\theta)$ est le biais de l'estimateur.

Remarque : la quantité $E\big((T-\theta)^2\big)$ s'appelle le *risque quadratique*.

Exercice 11

On considère dans cet exercice le cas du tirage d'un échantillon *sans remise*.

Pour chaque individu a de la population totale, on définit la variable aléatoire de Bernoulli ε_a telle que : $\varepsilon_a = \begin{cases} 1 & \text{si } a \text{ appartient à l'échantillon;} \\ 0 & \text{sinon.} \end{cases}$

11-1) Montrer que $P(\varepsilon_a = 1) = \dfrac{n}{N}$ où n est la taille de l'échantillon et N la taille de la population.

11-2) Montrer que la moyenne empirique peut s'écrire : $\bar{X}_n = \dfrac{1}{n}\displaystyle\sum_{a=1}^{N} \varepsilon_a X_a$.

11-3) Montrer que $E\big(\bar{X}_n\big) = m$.

11-4) Montrer que $\mathrm{Var}\big(\bar{X}_n\big) = \dfrac{N-n}{N-1} \times \dfrac{\sigma^2}{n}$.

Séance 3

Exercice 12

Lors du référendum de 2005 sur la Constitution Européenne, le “non” a recueilli $54,7\%$ des suffrages exprimés. Un sondage effectué à la sortie des urnes sur 995 électeurs indique que 502 ont déclaré avoir voté “oui”.

12-1) Identifier la population, l'échantillon, la variable, son type et son/ses paramètre(s).

12-2) Lorsqu'un échantillon de taille 995 est tiré au sort dans la population, quelle est la probabilité de tomber sur une proportion au moins aussi élevée de “oui” ? Que pensez-vous alors de cet échantillon ?

Exercice 13

Une étude parue dans le *British Medical Journal* en 2005 portait sur l'écart entre les naissances de garçons et de filles en Chine entre 2000 et 2004. Sur un échantillon de 5000 enfants nés dans cette période, 2768 sont des garçons.

En biologie, le *sex-ratio* indique la proportion de mâles à la naissance dans une espèce donnée. Sachant que chez l'homme il est de 52%, quelle est la probabilité d'obtenir sur un échantillon aléatoire de 5000 naissances une proportion de garçons au moins aussi élevée que celle observée dans l'échantillon du *British Medical Journal* ?

Exercice 14

Si l'on en croit les études de marketing, 25% des consommateurs seraient influencés par la marque lors de l'achat d'un produit. On interroge 80 consommateurs choisis au hasard.

14-1) Identifier la population, l'échantillon, la variable, son type et son/ses paramètre(s).

14-2) Caractériser la distribution de la fréquence empirique de l'influence de la marque sur les échantillons de taille 80, en précisant sa forme, sa moyenne, sa variance et son écart-type.

14-3) Calculer la probabilité pour que, dans un échantillon aléatoire de 80 consommateurs, au moins 30 se déclarent influencés par la marque.

14-4) Dans quel intervalle centré autour de $0,25$ trouve-t-on 80% des fréquences empiriques des échantillons de taille 80 ?

Exercice 15

Dans une université, 45% des étudiants s'adonnent à au moins une activité physique par semaine. Sur un échantillon de 400 étudiants tirés au sort dans cette université, on dénombre 205 étudiants pratiquant au moins une activité physique par semaine.

15-1) Compléter le tableau suivant :

	Taille	*Proportion de sportifs*
Population		
Échantillon		

15-2) Quelle est la probabilité d'obtenir un échantillon comportant plus de 205 sportifs ?

Exercice 16

On considère une famille $\{X_i\}$ de variables gaussiennes indépendantes de loi $\mathcal{N}(m, \sigma)$.

16-1) Montrer que la variable $X_i - \bar{X}_n$ suit une loi normale et préciser son espérance et sa variance.

16-2) Calculer la covariance $\mathrm{Cov}(\bar{X}_n, X_i - \bar{X}_n)$.

16-3) En déduire que $\bar{X}_n$ et $X_i - \bar{X}_n$ sont indépendantes, quel que soit i.

Séance 4

Exercice 17

On suppose que la distribution X des notes obtenues en mathématiques par les étudiants en Gestion suit une loi normale d'écart-type $2,65$. Pour un échantillon d'étudiants, on a observé

$$\sum_{i=1}^{28} x_i = 291 \qquad \sum_{i=1}^{28} x_i^2 = 3091$$

17-1) Estimer ponctuellement la moyenne μ de X.

17-2) Estimer cette même moyenne par un intervalle avec un seuil de confiance de $0,90$.

17-3) On suppose maintenant qu'on ne connaît pas lécart-type σ de la variable X, construire un intervalle pour la moyenne avec un seuil de confiance de $0,90$. Comparer-le à celui qui a été trouvé à la question précédente.

Exercice 18

On suppose que la distribution du prix d'un billet aller-retour à prix réduit entre New York et San Francisco suit une loi normale. On a relevé, pour un échantillon de 15 billets vendus au mois de septembre, les prix suivants :

310 260 265 255 300 310 230 250 265 280 290 240 285 250 260

18-1) Donner une estimation du prix moyen d'un billet.

18-2) Donner un intervalle de confiance d'ordre 95% pour le prix moyen d'un billet.

Exercice 19

On a mesuré le temps passé quotidiennement devant la télévision par un adolescent. Sur un échantillon de 80 personnes, on a calculé les quantités suivantes :

$$\sum_{i=1}^{80} x_i = 257,33 \qquad \sum_{i=1}^{80} x_i^2 = 874,78$$

19-1) Estimer ponctuellement la moyenne et l'écart-type du temps passé devant la télévision pour l'ensemble de la population française.

19-2) Estimer cette même moyenne par un intervalle avec un seuil de confiance de 0,95.

Exercice 20

On a mesuré, pour un échantillon de 300 enfants de 10 ans, leur poids X et obtenu :

$$\sum_{i=1}^{300} x_i = 8\ 817,6 \qquad \sum_{i=1}^{300} x_i^2 = 263\ 931,9$$

20-1) Donner une estimation ponctuelle du poids moyen μ d'un enfant de 10 ans.

20-2) Donner un intervalle de confiance à 98% sur μ.

Exercice 21

Une organisation de consommateurs souhaite estimer la distance moyenne parcourue par des pneus d'une marque X. Des tests réalisés sur un échantillon de 25 pneus ont indiqué en moyenne

45 250 kilomètres. En supposant que la distance parcourue par un pneu suit une loi normale d'écart-type 1600 kilomètres, donner un intervalle de confiance d'ordre 99% sur la distance moyenne parcourue par les pneus de la marque X. Préciser et justifier les conditions d'application du calcul.

Exercice 22

Une société souhaite mettre sur le marché un nouveau savon. Pour estimer le marché potentiel du nouveau produit, un sondage est effectué pour mesurer la consommation moyenne de savon dans la population considérée comme la cible privilégiée de ce produit (femmes actives de plus de trente cinq ans et de moins de soixante ans). La consommation mensuelle moyenne ressort à 3,73 onces (l'unité de mesure internationale utilisée par la société) sur un échantillon de quarante personnes, avec un écart-type calculé de 1,6 once.

Donner un intervalle de confiance à 95% de la consommation mensuelle moyenne de savon de la population cible.

Séance 5

Exercice 23

On a tiré un échantillon de 375 personnes parmi des hommes de 50 à 69 ans, mesuré leur taux de cholestérol X et obtenu

$$\sum_{i=1}^{375} x_i = 74\ 750 \qquad \sum_{i=1}^{375} x_i^2 = 15\ 695\ 900$$

On suppose que la distribution de X suit une loi normale.

23-1) Donner une estimation ponctuelle de la moyenne μ et de la variance σ^2 de X.

23-2) Donner un intervalle de confiance sur la moyenne de X au risque 5%.

Exercice 24

Peu de temps avant une élection, un candidat fait réaliser un sondage. Parmi 150 électeurs interrogés, 45 se disent prêts à voter pour lui.

24-1) Donner une estimation ponctuelle de la proportion p d'électeurs prêts à voter pour lui.

24-2) Donner un intervalle de confiance d'ordre 90% sur p. Préciser et justifier les conditions d'application du calcul.

Exercice 25

On considère un échantillon de 100 véhicules prélevés au hasard dans le parc des véhicules mis en service depuis six mois. Ce parc contient suffisamment de véhicules pour qu'on puisse assimiler ce tirage à un tirage avec remise. On constate que 91 véhicules de cet échantillon n'ont pas eu de sinistre.

25-1) Donner une estimation ponctuelle du pourcentage p de véhicules de ce parc qui n'ont pas eu de sinistre 6 mois après leur mise en service.

25-2) Déterminer un intervalle de confiance pour p avec un seuil de confiance de 95 %.

Exercice 26

Sur un échantillon de 1000 amateurs de café, 300 personnes disent préférer le *robusta* à l'*arabica*.

26-1) Donner un intervalle de confiance au risque 3% pour la proportion d'individus préférant le *robusta*.

26-2) Donner un intervalle de confiance au risque 3% pour la proportion d'individus préférant l'*arabica*.

Exercice 27

Pour contrôler la qualité d'une machine, on a prélevé au hasard un échantillon de 50 articles de la production récente s'élevant à 5000 articles et on compte le nombre d'articles défectueux. Sur cet échantillon, on a observé 2 articles défectueux.

27-1) Donner un intervalle de confiance au seuil 5% de la proportion d'articles défectueux produits par la machine.

27-2) On estime que la machine est défaillante si la proportion d'articles défectueux est supérieure à 10%. Les calculs de la question précédente vous permettent-ils de dire si la machine est défaillante ou non ?

27-3) Le remplacement d'une machine étant d'un coût important, on souhaiterait s'assurer de son dysfonctionnement. Avec un risque de 3%, combien d'articles doit-on prélever pour avoir une valeur approchée avec une précision de 5% de la proportion d'articles défectueux.

Exercice 28

Une enquête menée en 2003, s'intéressait à la proportion de ménages français ayant accès à l'internet (*source* : médiamétrie). Sur 2345 foyers, 657 possédaient un abonnement auprès d'un fournisseur d'accès.

28-1) À combien estimez-vous la proportion de foyers français ayant accès à l'internet ?

28-2) Au seuil $\alpha = 5\%$, entre quelles valeurs était comprise en 2003 la proportion de ménages français ayant accès à l'internet ?

Séance 6

Tests de moyenne

Exercice 29

On a mesuré le poids de raisin produit par souche sur 10 souches prises au hasard dans un vignoble. On a obtenu les résultats suivants exprimés en kilogrammes :

2,4 3,4 3,6 4,1 4,3 4,7 5,4 5,9 6,5 6,9

On suppose que le poids de raisin produit par une souche suit une loi normale d'écart-type inconnu.

Peut-on accepter, au risque 5%, l'hypothèse que le poids moyen de raisin produit par une souche est supérieur à 4,5 kilogrammes ?

Exercice 30

On a mesuré le rendement X (en quintaux par hectare) de 60 parcelles identiques, cultivées avec la même variété de maïs et obtenu

$$\sum_{i=1}^{60} x_i = 1890,75 \quad \text{et} \quad \sum_{i=1}^{60} x_i^2 = 60\ 747,98$$

L'année dernière, le rendement moyen des parcelles était de 30 quintaux par hectare. Peut-on conclure, au seuil 4%, que le rendement moyen des parcelles est supérieur cette année ?

Exercice 31

Un chercheur émet l'hypothèse que les étudiants d'une université dépensent en moyenne 150 euros par semaine pour leurs frais de logement et de nourriture. Pour tester la validité de cette hypothèse, il sélectionne 80 étudiants et observe sur cet échantillon une dépense moyenne de $153,75$ euros avec un écart-type de $27,79$ euros. Au risque 5%, peut-on accepter l'hypothèse émise par le chercheur ?

Exercice 32

Un fabricant de tubes à essai pour laboratoires fonde sa publicité sur le fait que la durée de vie moyenne de ses tubes est de 1500 h. Un organisme de contrôle de publicité constate que, sur 100 tubes à essai, la durée de vie moyenne observée est $\bar{x} = 1485$ h avec un écart-type de 110 h. Au risque 2%, la durée de vie moyenne des tubes à essai est-elle inférieure à ce qu'annonce le fabricant ?

Exercice 33

On reprend l'exercice . On rappelle que des tests réalisés sur un échantillon de 25 pneus ont indiqué une distance moyenne parcourue de 45 250 kilomètres.

Le fabricant annonce dans une publicité que ses pneus parcourent en moyenne plus de 45 000 kilomètres. On suppose que la distance X parcourue par un pneu suit une loi normale. Par contre, on ne connaît pas l'écart-type de X. Sur l'échantillon des 25 pneus, on a observé un écart-type de 1 530 kilomètres. Peut-on, au risque 1%, conclure des observations faites sur l'échantillon que la publicité dit vrai ?

Exercice 34

Le poids moyen d'une truite d'élevage au bout d'un an est de 250 g avec un écart-type de 45 g. C'est la norme retenue pour qu'elle puisse être vendue afin de repeupler une rivière. Un échantillon de 30 truites ayant reçu, dans un bassin expérimental, une alimentation enrichie en produits végétaux indique un poids moyen de $273,6$ g. Peut-on, au seuil $\alpha = 5\%$ dire que l'alimentation a eu une influence significative sur le poids des truites.

Séance 7

Exercice 35

À la suite d'un traitement sur une variété de rongeurs, on prélève un échantillon de 15 rongeurs que l'on pèse. On observe un poids moyen $\bar{x} = 82,6$ g avec un écart-type $s = 1,86$ g. À la même époque, un grand nombre de mesures ont permis d'établir que les rongeurs non traités avaient un poids moyen de $87,6$ g.

On suppose que le poids des rongeurs suit une loi normale. Le poids moyen des rongeurs traités diffère-t-il significativement de cette norme au seuil de 3% ?

Exercice 36

Sur une portion d'autoroute où la vitesse est limitée à 110 km/h, on effectue un contrôle de vitesse sur 12 véhicules, on observe une vitesse moyenne $\bar{x} = 123,61$ km/h avec un écart-type $s = 13,47$ km/h. Au vu de ces observations, peut-on conclure, au seuil 5%, que la limitation de vitesse n'est pas respectée en moyenne ? On suppose que la vitesse d'un véhicule suit une loi normale.

Tests de variance

Exercice 37

Une méthode de mesure de référence a suggéré que le taux de magnésium d'une certaine eau de source présentait un écart-type $\sigma = 15$ (mesuré en mg/l). Lors de 10 mesures répétées du taux de magnésium, on a observé un écart-type estimé de $s_{10} = 8\ mg/l$. Ce résultat est-il significativement différent de celui de la méthode de référence ?

Exercice 38

Une machine-outil est réglée pour produire des boulons d'un poids déterminé avec un écart-type de $1,5\ g$.

On a effectué des contrôles sur 20 boulons tirés au hasard et on a obtenu un écart-type estimé de $s_{20} = 2,3\ g$. Ce résultat diffère-t-il significativement de celui fixé par la machine ?

Exercice 39

Le temps d'incubation moyen pour le développement d'un anticorps est de 13 jours. On souhaite expérimenter un traitement que l'on administre à 10 cobayes afin de diminuer ce temps d'incubation (considéré comme une variable gaussienne). Les temps d'incubation observés sont :

Cobaye	1	2	3	4	5	6	7	8	9	10
Temps d'incubation	5	11	10	13	10	8	9	9	12	12

L'action du traitement est-elle significative au seuil $\alpha = 5\%$?

Exercice 40

Il y a près de 40 ans, une enquête indiquait que les adolescents de 15 ans consacraient en moyenne $8,5$ heures par semaine à leurs devoirs à la maison (si ! si ! si !). Aujourd'hui, la même étude est réalisée sur un échantillon de 200 adolescents de 15 ans. On observe que ceux-ci consacrent en moyenne $7,1$ heures par semaine à leurs devoirs avec un écart-type de 1,1 heure.

Peut-on conclure, au risque 1%, que les adolescents de 15 ans d'aujourd'hui consacrent en moyenne moins de temps à leurs devoirs qu'il y a 40 ans ?

Séance 8

Tests de proportion

Exercice 41

Une enquête réalisée sur un échantillon de 500 ménages a montré que 419 d'entre eux possédaient une connexion internet. Peut-on dès lors rejeter, à l'aide d'un test bilatéral au seuil $\alpha = 5\%$, l'hypothèse que 80 % des ménages ont une connexion internet ?

Exercice 42

Pour contrôler un important stock commercial comprenant des dizaines de milliers d'articles, on sélectionne au hasard un échantillon de 800 articles. Parmi eux, 36 sont défectueux. Le directeur des ventes est tenu de respecter une proportion d'articles défectueux strictement inférieure à 3 %. Acceptera-t-il le stock (au seuil 1%) ?

Exercice 43

Une société est spécialisée dans le crédit à la consommation. Ses dirigeants craignent que plus de 10 % des créances soient douteuses (c'est-à-dire ne soient pas honorées). Un sondage est organisé. On tire au hasard 50 créances et on observe 8 créances douteuses dans l'échantillon. Au niveau 5%, les craintes des dirigeants sont-elles confirmées par cette observation ?

Exercice 44

Sur 4000 naissances dans une maternité de Chine, on a relevé 2075 garçons.

44-1) Cette observation est-elle conforme avec l'hypothèse que la probabilité théorique qu'il naisse un garçon est de 0,5 ? On prendra successivement $\alpha = 5\%$ et $\alpha = 1\%$. Commenter la différence de conclusion entre les deux cas.

44-2) L'exercice indiquait une proportion de 52% (et non pas 50%). Les données observées s'accordent-elles avec cette valeur au seuil $\alpha = 5\%$?

Exercice 45

Sur un échantillon de 200 individus d'une commune, 30% sont favorables à l'implantation d'un centre commercial. Au niveau 1%, ceci contredit-il l'hypothèse selon laquelle un habitant sur trois y est favorable ?

Exercice 46

La moitié des bacheliers soumis à un test de connaissances en géographie obtiennent un score supérieur ou égal à 13. Sur un échantillon de 70 bacheliers scientifiques, on observe que 48 obtiennent un score supérieur ou égal à 13. Au risque 2%, peut-on en conclure que les connaissances en géographie des bacheliers scientifiques sont au-dessus du standard ?

Séance 9

Exercice 47

Une entreprise souhaite tester l'efficacité d'un programme de formation. Pour cela, elle sélectionne au hasard un échantillon de 10 employés n'ayant pas suivi le programme de formation et un autre échantillon de 12 employés ayant suivi le programme de formation. Chacune des personnes passe un test dont les questions portent sur les thèmes présentés dans le programme de formation. Pour l'échantillon des 10 employés n'ayant pas suivi la formation, on observe une note moyenne de $13,4$ avec un écart-type de $2,98$. Pour l'échantillon des 12 employés ayant suivi la formation, on observe une note moyenne de $14,6$ avec un écart-type de $2,64$. Les notes des personnes ayant suivi la formation et des personnes n'ayant pas suivi la formation sont supposées être distribuées selon une loi normale de même variance.

Au risque 5%, peut-on conclure que le programme est efficace ?

Exercice 48

Une société de location de voitures met en place une expérience afin de décider si les consommations d'essence sont différentes pour deux types de pneus. 7 voitures choisies au hasard sont conduites sur un parcours précis avec des pneus de type A. Les pneus sont alors remplacés par ceux de type B et ces voitures sont de nouveau conduites sur le même parcours. Les consommations en km/litre des voitures en question sont supposées être distribuées selon une loi normale de même variance. On observe les résultats suivants :

Voiture	1	2	3	4	5	6	7
Pneus A	4,2	4,7	6,6	7	6,7	4,5	5,7
Pneus B	4,1	4,9	6,2	6,9	6,8	4,4	5,7

Au risque 2%, peut-on conclure que les consommations d'essence sont différentes pour les deux types de pneus ?

Exercice 49

Dans une coopérative agricole, on désire tester l'effet d'un engrais sur la production de blé. Pour cela, on choisit au hasard 20 parcelles de terrain de même superficie. La moitié de ces parcelles est traitée avec l'engrais, et l'autre ne l'est pas. Pour les 10 parcelles non traitées, on observe $\sum x_i = 6,16$ et $\sum x_i^2 = 29,22$ où x_i représente la production en tonnes de la i-ème parcelle. Pour les 10 parcelles traitées, on observe $\sum y_j = 6,68$ et $\sum y_j^2 = 34,35$ où y_j représente la production en tonnes de la j-ème parcelle. Les productions en tonnes des parcelles traitées et non traitées sont supposées être distribuées selon une loi normale de même variance.

Au risque 10%, peut-on conclure que l'engrais a un effet sur la production d'une parcelle ?

Exercice 50

Neuf malades présentant des symptômes d'anxiété sont choisis au hasard et reçoivent un traitement. On évalue l'état du malade avant et après traitement par un indice spécifique que le médecin calcule. Si le traitement est efficace, l'indice doit diminuer. Les valeurs observées de cet indice sont les suivantes :

Avant	1,83	0,5	1,62	2,48	1,68	1,88	1,55	3,06	1,30
Après	0,88	0,65	0,60	2,05	1,06	1,29	1,06	3,14	1,29

Au seuil 1%, peut-on conclure que le traitement est efficace ?

Séance 10

Exercice 51

Dans un échantillon de 300 personnes choisies au hasard dans la population d'une ville A, on observe que 36 d'entre elles fument au moins deux paquets de cigarettes par jour. Dans une autre ville B, pour un échantillon de 100 personnes choisies au hasard dans la population de la ville, on observe que 8% fument au moins deux paquets par jour. Au risque 1%, peut-on conclure que les proportions de personnes qui fument au moins deux paquets par jour sont différentes dans les deux villes ?

Exercice 52

Soit p_1 la probabilité de guérison d'une maladie grâce au traitement T_1. Un groupe de 50 malades choisis au hasard est soumis à ce traitement et on observe que 28 guérissent. On dispose d'un autre traitement T_2 permettant de soigner cette maladie. Sur 60 malades choisis au hasard soumis à ce nouveau traitement, 24 guérissent. Au risque 10%, peut-on conclure qu'il y a une différence entre T_1 et T_2 quant à leur efficacité ?

Exercice 53

On reprend les données de l'exercice .

On note X_1 la consommation en km/litre des voitures équipées de pneus A et X_2 la consommation en km/litre des voitures équipées de pneus B. On suppose que les variables X_1 et X_2 sont distribuées selon une loi normale. Au risque 10%, peut-on conclure que les variances de X_1 et X_2 sont égales ?

Exercice 54

On reprend les données de l'exercice .

On note X_1 la production en tonnes d'une parcelle traitée avec l'engrais et X_2 la production en tonnes d'une parcelle qui n'est pas traitée avec l'engrais. On suppose que les deux variables X_1 et X_2 sont distribuées selon une loi normale. Au risque 10%, peut-on conclure que les variances de X_1 et X_2 sont égales ?

Séance 11

Exercice 55

Voici les 200 premières décimales du nombre π :

1415926535 8979323846 2643383279 5028841971 6939937510 5820974944 5923078164 0628620899 8628034825 3421170679 8214808651 3282306647 0938446095 5058223172 5359408128 4811174502 8410270193 8521105559 6446229489 5493038196

En dénombrant combien de fois apparaît chaque chiffre, on arrive à la table suivante :

Chiffres	0	1	2	3	4	5	6	7	8	9
Effectifs	19	20	24	19	22	20	16	12	25	23

Peut-on considérer, au vu de cet échantillon, que les décimales sont équiréparties ?

Exercice 56

On a noté à 200 reprises le nombre N de particules β émises par un échantillon de potassium radioactif pendant une durée d'une minute. On a obtenu les résultats suivants :

Nombre de particules	0	1	2	3	4	5	6
Nombre d'observations	8	36	45	50	30	25	6

On fait l'hypothèse que N suit une loi de Poisson de paramètre inconnu λ. On rappelle que pour une loi de Poisson, on a

$$P(N = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

56-1) Donner une estimation de λ.

56-2) Au moyen d'un test du χ^2, vérifier si l'hypothèse d'ajustement à une loi de Poisson est acceptable au seuil $\alpha = 5\%$?

Exercice 57

On a relevé le nombre de filles dans 500 familles de quatre enfants :

Nombre de filles	0	1	2	3	4
Nombre de familles	28	142	169	136	25

57-1) Quelle est la loi de probabilité suivie par la variable "*nombre de filles*" ?
57-2) Quelle est la moyenne empirique du nombre de filles ?
57-3) Peut-on rejeter, au seuil $\alpha = 5\%$, l'hypothèse d'équiprobabilité des sexes ?

Exercice 58

Une enquête sur la répartition des groupes sanguins en fonction du sexe a donné les résultats suivants :

	A	B	AB	O
H	51	12	6	50
F	38	6	5	32

58-1) Calculer les distributions marginales.
58-2) Faire un test d'indépendance de Fisher au seuil 5% pour déterminer si la répartition des groupes sanguins est indépendante du sexe.

Exercice 59

Les 35 étudiants d'un Master1 ont choisi un enseignement optionnel parmi les trois possibilités O_1, O_2, O_3 qui leur sont offertes. L'option O_1 a été choisie par 10 étudiants, l'option O_2 par 12 étudiants et l'option O_3 par 13 étudiants. On veut tester si les résultats obtenus sont indépendants de l'option choisie. Les notes observées dans ces trois options sont les suivantes :

	Notes sur 20												
O_1	5,5	7,5	11,5	18	4	18	19	13	12,5	1			
O_2	4	3,5	13,5	7,5	15,5	10	14,5	20	7,5	15,5	18,5	4	
O_3	13	2,5	5,5	7,5	0,5	7,5	17,5	7	9,5	12	10	3,5	16,5

59-1) Formuler l'hypothèse nulle H_0 du test.
59-2) Calculer le nombre de réussites et d'échecs pour chacun des trois groupes dans un tableau ayant la forme suivante *(on considère qu'il y a réussite si la note est supérieure ou égale à 10)* :

Observations	O_1	O_2	O_3
Réussite			
Échec			

59-3) Compléter le tableau précédent en indiquant les sommes marginales.
59-4) Calculer, dans un tableau similaire, les valeurs théoriques attendues si H_0 est vraie.
59-5) Effectuer un test d'indépendance de Fisher pour comparer les valeurs observées et les valeurs théoriques au seuil $\alpha = 5\%$. Préciser bien le nombre de degrés de liberté de ce test.

Exercice 60

Dans le cadre du développement d'un nouveau jeu video, on teste un générateur de nombres au hasard produisant des nombres entiers de 0 à 9. Un échantillon de 1000 observations est réparti selon la table suivante :

Chiffres	0	1	2	3	4	5	6	7	8	9
Effectifs	115	93	119	93	107	103	92	105	94	79

Peut-on, au seuil $\alpha = 5\%$, rejeter l'hypothèse que ces chiffres sont équirépartis ?

Exercice 61

Un test de quotient intellectuel réalisé auprès de 200 enfants a donné la répartition suivante :

Q.I.	$[75, 85[$	$[85, 95[$	$[95, 105[$	$[105, 115[$	$[115, 125[$
Effectifs	4	51	88	47	10

Chaque intervalle d'amplitude 10 sera représenté par son milieu (par exemple, l'intervalle $[75, 85[$ correspond à un *Q.I.* de 80).

61-1) Calculer la moyenne et l'écart-type empiriques de l'échantillon.

61-2) La distribution peut-elle être ajustée par une loi normale (au seuil $\alpha = 0,05$) ?

Exercice 62

On lance une paire de dés et on note la somme S des scores. En répétant cette opération 180 fois, on obtient la table suivante qui indique combien de fois chaque somme possible a été réalisée.

Somme	2	3	4	5	6	7	8	9	10	11	12
Effectifs	3	11	12	23	28	32	31	15	12	8	5

62-1) Écrire la loi de probabilité exacte de la variable S. On calculera explicitement les probabilités $P(S = k)$ pour chaque valeur de k.

62-2) Ces résultats sont-ils en adéquation avec ce qu'on attend si les deux dés sont parfaitement équilibrés ?

Tables

Fonction de répartition de la loi normale $\mathcal{N}(0,1)$

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500	0,504	0,508	0,512	0,516	0,520	0,524	0,528	0,532	0,536
0,1	0,540	0,544	0,548	0,552	0,556	0,560	0,564	0,567	0,571	0,575
0,2	0,579	0,583	0,587	0,591	0,595	0,599	0,603	0,606	0,610	0,614
0,3	0,618	0,622	0,626	0,629	0,633	0,637	0,641	0,644	0,648	0,652
0,4	0,655	0,659	0,663	0,666	0,670	0,674	0,677	0,681	0,684	0,688
0,5	0,691	0,695	0,698	0,702	0,705	0,709	0,712	0,716	0,719	0,722
0,6	0,726	0,729	0,732	0,736	0,739	0,742	0,745	0,749	0,752	0,755
0,7	0,758	0,761	0,764	0,767	0,770	0,773	0,776	0,779	0,782	0,785
0,8	0,788	0,791	0,794	0,797	0,800	0,802	0,805	0,808	0,811	0,813
0,9	0,816	0,819	0,821	0,824	0,826	0,829	0,831	0,834	0,836	0,839
1,0	0,841	0,844	0,846	0,848	0,851	0,853	0,855	0,858	0,860	0,862
1,1	0,864	0,867	0,869	0,871	0,873	0,875	0,877	0,879	0,881	0,883
1,2	0,885	0,887	0,889	0,891	0,893	0,894	0,896	0,898	0,900	0,901
1,3	0,903	0,905	0,907	0,908	0,910	0,911	0,913	0,915	0,916	0,918
1,4	0,919	0,921	0,922	0,924	0,925	0,926	0,928	0,929	0,931	0,932
1,5	0,933	0,934	0,936	0,937	0,938	0,939	0,941	0,942	0,943	0,944
1,6	0,945	0,946	0,947	0,948	0,949	0,951	0,952	0,953	0,954	0,954
1,7	0,955	0,956	0,957	0,958	0,959	0,960	0,961	0,962	0,962	0,963
1,8	0,964	0,965	0,966	0,966	0,967	0,968	0,969	0,969	0,970	0,971
1,9	0,971	0,972	0,973	0,973	0,974	0,974	0,975	0,976	0,976	0,977
2,0	0,977	0,978	0,978	0,979	0,979	0,980	0,980	0,981	0,981	0,982
2,1	0,982	0,983	0,983	0,983	0,984	0,984	0,985	0,985	0,985	0,986
2,2	0,986	0,986	0,987	0,987	0,987	0,988	0,988	0,988	0,989	0,989
2,3	0,989	0,990	0,990	0,990	0,990	0,991	0,991	0,991	0,991	0,992
2,4	0,992	0,992	0,992	0,992	0,993	0,993	0,993	0,993	0,993	0,994
2,5	0,994	0,994	0,994	0,994	0,994	0,995	0,995	0,995	0,995	0,995
2,6	0,995	0,995	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996
2,7	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997
2,8	0,997	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998
2,9	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,999	0,999	0,999

Quantiles de la loi de Student

$\nu \backslash P$	0,6	0,7	0,8	0,9	0,95	0,975	0,99	0,995
1	0,3249	0,7265	1,3764	3,0777	6,3138	12,7062	31,8205	63,6567
2	0,2887	0,6172	1,0607	1,8856	2,9200	4,3027	6,9646	9,9248
3	0,2767	0,5844	0,9785	1,6377	2,3534	3,1824	4,5407	5,8409
4	0,2707	0,5686	0,9410	1,5332	2,1318	2,7764	3,7469	4,6041
5	0,2672	0,5594	0,9195	1,4759	2,0150	2,5706	3,3649	4,0321
6	0,2648	0,5534	0,9057	1,4398	1,9432	2,4469	3,1427	3,7074
7	0,2632	0,5491	0,8960	1,4149	1,8946	2,3646	2,9980	3,4995
8	0,2619	0,5459	0,8889	1,3968	1,8595	2,3060	2,8965	3,3554
9	0,2610	0,5435	0,8834	1,3830	1,8331	2,2622	2,8214	3,2498
10	0,2602	0,5415	0,8791	1,3722	1,8125	2,2281	2,7638	3,1693
11	0,2596	0,5399	0,8755	1,3634	1,7959	2,2010	2,7181	3,1058
12	0,2590	0,5386	0,8726	1,3562	1,7823	2,1788	2,6810	3,0545
13	0,2586	0,5375	0,8702	1,3502	1,7709	2,1604	2,6503	3,0123
14	0,2582	0,5366	0,8681	1,3450	1,7613	2,1448	2,6245	2,9768
15	0,2579	0,5357	0,8662	1,3406	1,7531	2,1314	2,6025	2,9467
16	0,2576	0,5350	0,8647	1,3368	1,7459	2,1199	2,5835	2,9208
17	0,2573	0,5344	0,8633	1,3334	1,7396	2,1098	2,5669	2,8982
18	0,2571	0,5338	0,8620	1,3304	1,7341	2,1009	2,5524	2,8784
19	0,2569	0,5333	0,8610	1,3277	1,7291	2,0930	2,5395	2,8609
20	0,2567	0,5329	0,8600	1,3253	1,7247	2,0860	2,5280	2,8453
21	0,2566	0,5325	0,8591	1,3232	1,7207	2,0796	2,5176	2,8314
22	0,2564	0,5321	0,8583	1,3212	1,7171	2,0739	2,5083	2,8188
23	0,2563	0,5317	0,8575	1,3195	1,7139	2,0687	2,4999	2,8073
24	0,2562	0,5314	0,8569	1,3178	1,7109	2,0639	2,4922	2,7969
25	0,2561	0,5312	0,8562	1,3163	1,7081	2,0595	2,4851	2,7874
26	0,2560	0,5309	0,8557	1,3150	1,7056	2,0555	2,4786	2,7787
27	0,2559	0,5306	0,8551	1,3137	1,7033	2,0518	2,4727	2,7707
28	0,2558	0,5304	0,8546	1,3125	1,7011	2,0484	2,4671	2,7633
29	0,2557	0,5302	0,8542	1,3114	1,6991	2,0452	2,4620	2,7564
30	0,2556	0,5300	0,8538	1,3104	1,6973	2,0423	2,4573	2,7500

Quantiles de la loi du χ^2

Probabilités sur la première ligne

Degré de libertés sur la première colonne

$\nu \backslash P$	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99
1	0,000	0,001	0,004	0,016	2,706	3,841	5,024	6,635
2	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210
3	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345
4	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277
5	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086
6	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812
7	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475
8	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090
9	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666
10	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209
11	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725
12	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217
13	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688
14	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141
15	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578
16	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000
17	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409
18	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805
19	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191
20	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566

Quantiles de la loi de Fisher à $0,95$

Degré de libertés du numérateur sur la première ligne

Degré de libertés du dénominateur sur la première colonne

	1	2	3	4	5	6	7	8	9	10
1	161,45	18,51	10,13	7,71	6,61	5,99	5,59	5,32	5,12	4,96
2	199,50	19,00	9,55	6,94	5,79	5,14	4,74	4,46	4,26	4,10
3	215,71	19,16	9,28	6,59	5,41	4,76	4,35	4,07	3,86	3,71
4	224,58	19,25	9,12	6,39	5,19	4,53	4,12	3,84	3,63	3,48
5	230,16	19,30	9,01	6,26	5,05	4,39	3,97	3,69	3,48	3,33
6	233,99	19,33	8,94	6,16	4,95	4,28	3,87	3,58	3,37	3,22
7	236,77	19,35	8,89	6,09	4,88	4,21	3,79	3,50	3,29	3,14
8	238,88	19,37	8,85	6,04	4,82	4,15	3,73	3,44	3,23	3,07
9	240,54	19,38	8,81	6,00	4,77	4,10	3,68	3,39	3,18	3,02
10	241,88	19,40	8,79	5,96	4,74	4,06	3,64	3,35	3,14	2,98
11	242,98	19,40	8,76	5,94	4,70	4,03	3,60	3,31	3,10	2,94
12	243,91	19,41	8,74	5,91	4,68	4,00	3,57	3,28	3,07	2,91
13	244,69	19,42	8,73	5,89	4,66	3,98	3,55	3,26	3,05	2,89
14	245,36	19,42	8,71	5,87	4,64	3,96	3,53	3,24	3,03	2,86
15	245,95	19,43	8,70	5,86	4,62	3,94	3,51	3,22	3,01	2,85
16	246,46	19,43	8,69	5,84	4,60	3,92	3,49	3,20	2,99	2,83
17	246,92	19,44	8,68	5,83	4,59	3,91	3,48	3,19	2,97	2,81
18	247,32	19,44	8,67	5,82	4,58	3,90	3,47	3,17	2,96	2,80
19	247,69	19,44	8,67	5,81	4,57	3,88	3,46	3,16	2,95	2,79
20	248,01	19,45	8,66	5,80	4,56	3,87	3,44	3,15	2,94	2,77

	11	12	13	14	15	16	17	18	19	20
1	4,84	4,75	4,67	4,60	4,54	4,49	4,45	4,41	4,38	4,35
2	3,98	3,89	3,81	3,74	3,68	3,63	3,59	3,55	3,52	3,49
3	3,59	3,49	3,41	3,34	3,29	3,24	3,20	3,16	3,13	3,10
4	3,36	3,26	3,18	3,11	3,06	3,01	2,96	2,93	2,90	2,87
5	3,20	3,11	3,03	2,96	2,90	2,85	2,81	2,77	2,74	2,71
6	3,09	3,00	2,92	2,85	2,79	2,74	2,70	2,66	2,63	2,60
7	3,01	2,91	2,83	2,76	2,71	2,66	2,61	2,58	2,54	2,51
8	2,95	2,85	2,77	2,70	2,64	2,59	2,55	2,51	2,48	2,45
9	2,90	2,80	2,71	2,65	2,59	2,54	2,49	2,46	2,42	2,39
10	2,85	2,75	2,67	2,60	2,54	2,49	2,45	2,41	2,38	2,35
11	2,82	2,72	2,63	2,57	2,51	2,46	2,41	2,37	2,34	2,31
12	2,79	2,69	2,60	2,53	2,48	2,42	2,38	2,34	2,31	2,28
13	2,76	2,66	2,58	2,51	2,45	2,40	2,35	2,31	2,28	2,25
14	2,74	2,64	2,55	2,48	2,42	2,37	2,33	2,29	2,26	2,22
15	2,72	2,62	2,53	2,46	2,40	2,35	2,31	2,27	2,23	2,20
16	2,70	2,60	2,51	2,44	2,38	2,33	2,29	2,25	2,21	2,18
17	2,69	2,58	2,50	2,43	2,37	2,32	2,27	2,23	2,20	2,17
18	2,67	2,57	2,48	2,41	2,35	2,30	2,26	2,22	2,18	2,15
19	2,66	2,56	2,47	2,40	2,34	2,29	2,24	2,20	2,17	2,14
20	2,65	2,54	2,46	2,39	2,33	2,28	2,23	2,19	2,16	2,12