Exercices

1 Rappels

Exercice 1.1 Loi normale. L'échelle motrice de Bayley permet de mesurer le développement psychomoteur des enfants. On sait que dans la population générale des enfants normaux âgés de 6 mois, les scores de développement psychomoteur sont distribués selon une loi normale de moyenne $\mu = 100$ et d'écart-type $\sigma = 15$.

1. Définir la population et la variable étudiées. Quelle est la loi de la variable (forme, moyenne, écart-type) ?
2. Un enfant a obtenu un score de 112. Calculer son score centré réduit z.
3. Un enfant a un score qui se situe à 1,2 écart-type au-dessus de la moyenne. Quel est son score ?
4. Un enfant a un score qui se situe à 2 écart-types en-dessous de la moyenne. Quel est son score ?
5. Quelle est la proportion d'enfants normaux âgés de 6 mois ayant un score supérieur à 130 sur l'échelle motrice de Bayley ?
6. Quelle proportion de la population a un score inférieur à 82 ?
7. Quelle est la probabilité qu'un enfant, tiré au sort dans la population, obtienne un score compris entre 70 et 120 ?
8. Un enfant a obtenu un score égal à 137,5. Calculer la proportion d'enfants ayant un score plus élevé que 137,5. Peut-on dire que son score est exceptionnellement élevé ?
9. Quel est le score en dessous duquel se situent 95 % des enfants de la population ? Quel est le score au-dessus duquel se situent 1% des enfants ?
10. Calculer l'intervalle de variation au risque $\alpha = 10\%$ (resp. 5%) contenant 90% (resp. 95%) des scores de la population.

Exercice 1.2 Loi normale. Un certain test de reconnaissance d'objets donne, pour une population de personnes présentant une agnosie visuelle, une moyenne de 10 objets reconnus et une variance de 9. On sait que la distribution des scores est normale dans cette population. Un nouveau patient d'un centre de rééducation vient de passer le test et a obtenu la note de 12.
Quelle est la probabilité d'obtenir un score de 12 ou plus dans la population d'agnosiques ? Classeriez-vous cet individu comme individu « sain » ?

Exercice 1.3 Loi normale. Le score à un test d'aptitude suit une distribution normale de moyenne 150 et d'écart-type 30, dans une population d'enfants scolarisés. On dirige un enfant vers un programme de soutien scolaire si son score au test d'aptitude se situe sous le 1er décile. En dessous de quel score un enfant sera-t-il dirigé vers le programme de soutien scolaire ?

Exercice 1.4 loi normale. Un score obtenu au questionnaire d'auto-évaluation des jeunes d'Achenbach permet de mesurer le nombre total de problèmes comportementaux signalés par l'enfant, pondérés selon la gravité du problème. Un score élevé correspond à un grand nombre de problèmes comportementaux. Dans une population d'enfants donnée, la distribution de ce score est supposée normale de moyenne 50 et d'écart-type 10.

1. Quelle est la proportion d'enfants ayant un score inférieur à 74 ?
2. Quelle est la proportion d'enfants ayant un score inférieur à 34 ?
3. Quelle est la proportion d'enfants ayant un score compris entre 20 et 74 ?

4. Quel serait le score limite pour le diagnostic si l'on voulait identifier les enfants dont le score se situe dans les 2% supérieurs de la population ?

Exercice 1.5 Estimation ponctuelle. On s'intéresse aux résultats du test B101 de la batterie Standard Bonnardel chez les hommes de 20 à 30 ans. Le test consiste en une série de 16 modèles à reproduire. On mesure le temps de reproduction des modèles (en secondes).
1) Préciser la population et la variable étudiée X. Quelle est la taille n de la population ? La loi de X est-elle connue ? Connaît-on la moyenne μ et l'écart-type σ de la population ?
2) On prélève un échantillon de 18 hommes dans la population et leur fait passer le test. Les temps x_i relevés sont les suivants :

Temps x_i :	489	561	383	336	437	555	339	444	383	$\sum x_i = 7932$
	362	458	351	348	555	441	464	469	557	$\sum x_i^2 = 3\,604\,292$

a) Préciser la taille n de l'échantillon.
b) Calculer la moyenne $\bar{x}$ et l'écart-type s de l'échantillon. Donner une estimation ponctuelle du temps moyen μ de la population. Donner une estimation ponctuelle sans biais de l'écart-type des temps σ dans la population.
c) Calculer la fréquence f d'individus de l'échantillon pour lesquels le temps de reproduction des modèles est supérieur à 400 s. Quel paramètre de la population cette fréquence permet-elle d'estimer ?
3) Un second échantillon de taille 18 a été prélevé dans la population par un autre expérimentateur. Les temps sont résumés ci-dessous :
$\sum x_i = 8021 \quad \sum x_i^2 = 3\,801\,292$
Quelles sont les estimations ponctuelles de μ, σ fournies par cet échantillon ? Comparez ces estimations avec celles obtenues à partir du premier échantillon.

Exercice 1.6 Moyenne empirique. Un spécialiste en psychologie industrielle utilise auprès de la main-d'œuvre de l'entreprise Electrotek une adaptation d'un test d'aptitude pour exécuter une certaine tâche. D'après les standards établis, les résultats au test sont distribués suivant une loi normale de moyenne 150 et d'écart-type 10 et le test adapté suit les mêmes normes.
1) Définir la population et la variable étudiées. Préciser la loi de la variable, sa moyenne et son écart-type.
2) On décide de prélever un échantillon de taille $n = 25$ dans la population. Quelle est la loi de la moyenne empirique $\bar{X}_n$?
3) Calculer l'intervalle de variation contenant 95% des valeurs de $\bar{X}_n$ observées sur de tels échantillons. Donner l'intervalle de variation au même risque pour la statistique Z_n.
4) On considère qu'un échantillon est atypique de la population au seuil $\alpha = 5\%$ si sa moyenne fait partie des 5% de valeurs les plus extrêmes de $\bar{X}_n$.
Sur un échantillon de 25 individus de l'entreprise choisis au hasard on obtient un résultat moyen $\bar{x} = 154,7$. Peut-on dire que cet échantillon est atypique de la population au seuil $\alpha = 5\%$? Même question au seuil $\alpha = 1\%$.

Exercice 1.7 Moyenne empirique. Le questionnaire d'auto-évaluation d'Achenbach permet de mesurer les problèmes comportementaux des jeunes (de 4 à 16 ans). On sait que dans la population générale, les scores sont distribués selon une loi normale de moyenne 50 et d'écart-type 10.
1) Définir la population et la variable étudiées. Préciser la loi de la variable, sa moyenne et son écart-type.
2) On décide de prélever un échantillon de taille $n = 24$ dans la population. Quelle est la loi de la moyenne empirique $\bar{X}_n$?
3) On considère que la moyenne d'un échantillon est, au seuil $\alpha = 5\%$, significativement plus grande que la moyenne générale si elle fait partie des 5% de valeurs les plus grandes de $\bar{X}_n$.

(a) Calculer la valeur limite au-dessus de laquelle la moyenne de l'échantillon tiré au sort sera, au seuil $\alpha = 5\%$, jugée significativement plus grande que la moyenne générale.
(b) Sur un échantillon de 24 enfants, on a observé un score moyen de 54,5. Que peut-on dire de cet échantillon ?

Exercice 1.8 Moyenne empirique. La variable « résultat à un test portant sur la richesse et la précision du vocabulaire chez les enfants de 12 ans » obéit à une loi de moyenne 60 et d'écart-type 10.
1) Définir la population et la variable étudiée.
2) On s'intéresse à la moyenne empirique $\bar{X}_n$. Indiquer l'effet de la taille de l'échantillon sur les caractéristiques de la distribution de $\bar{X}_n$:

taille de l'échantillon	moyenne empirique $\bar{X}_n$ forme de la loi	moyenne	écart-type
$n = 8$			
$n = 32$			
$n = 100$			

Evaluer la probabilité en tirant au hasard un échantillon de taille $n = 32$ d'observer un résultat moyen supérieur à 64. Même question pour $n = 100$. Que peut-on dire pour $n = 8$?

Exercice 1.9 Moyenne empirique. Dans une population de référence, les scores à un test de mémorisation (nombre d'éléments retenus) se distribuent selon une loi de moyenne 7 et d'écart-type 1,9.
On considère ici que la moyenne d'un échantillon de taille $n = 32$ est significativement plus petite que la moyenne de référence si elle fait partie des 1% de valeurs les plus petites de $\bar{X}_n$.
Sur un échantillon de 32 personnes, on a observé un résultat moyen de 6,1. Calculer la probabilité d'observer sur un échantillon de taille 32 un score moyen inférieur à celui observé. Que peut-on penser de l'échantillon ?

2 Tests sur la moyenne

Exercice 2.1 Formuler les hypothèses H_0 et H_1 pour chacune des situations suivantes :
1) On se demande si le QI moyen des enfants de 6 à 13 ans est différent de 100.
2) Le score moyen d'aptitude à l'organisation perceptive pour les enfants « normaux » de 6 à 12 ans est égal à 10. On veut tester l'hypothèse selon laquelle le score moyen des enfants hyperactifs est différent de celui des enfants normaux.
3) Des études ont montré que le niveau moyen d'anxiété pour les sujets qui ne sont pas privés de rêves est de $26,5$. On veut tester l'hypothèse selon laquelle la privation de rêves augmente le niveau d'anxiété.
4) On veut tester l'hypothèse selon laquelle le score moyen à un test de mémorisation pour les personnes très âgées est inférieur à 7 (score moyen de référence).
5) On veut tester l'efficacité d'une thérapie dans le traitement de l'anorexie chez les adolescentes. On mesure le poids des sujets au début (P_1) et à la fin (P_2) de la thérapie. On mesure l'efficacité de la thérapie par la différence de poids $P_1 - P_2$.
Dans chaque cas, on précisera la population et la variable étudiées.

Exercice 2.2 Le test du compteur de tours informe sur l'interférence d'une tâche visuelle avec une activité motrice. On admet que la variable « rendement pour la main droite » chez les enfants de 9 ans est régie par une loi normale d'écart-type $\sigma = 10$.
On applique ce test à 15 enfants de 9 ans et on obtient un rendement moyen pour la main droite de 91,3. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que le rendement moyen μ, chez les enfants de 9 ans, est différent de 95 ?

Exercice 2.3 On applique le test d'écriture en miroir à 15 sujets présentant certaines anomalies. Pour chaque sujet, on mesure le temps (en secondes) utilisé pour reproduire trois séries de chiffres et de lettres. On obtient un temps moyen $\bar{x} = 200,8$ pour un écart-type $s = 19$. On admet que la variable « temps » suit une loi normale dans la population considérée.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que le temps moyen μ est différent de 200 ?

Exercice 2.4 On soumet 15 enfants de 9 ans à un test permettant de mesurer la motricité manuelle. Les résultats (temps en minute pour enfiler 10 perles) sont résumés ci-dessous :
$\sum x_i = 47,13 \quad \sum x_i^2 = 181,95$
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que le temps moyen chez les enfants de 9 ans est inférieur à 3,5 mn ? On admettra que la variable temps est régie par une loi normale.

Exercice 2.5 Des insomniaques avaient une durée de sommeil moyenne de 2 heures par nuit avant d'être traités avec un nouveau médicament. S'il agit, ce médicament doit augmenter la durée de sommeil. Pour tester ce nouveau médicament, 11 sujets choisis au hasard ont reçu le médicament. Leurs durées de sommeil après traitement sont résumées ci-dessous :
$\sum x_i = 31,7 \quad \sum (x_i)^2 = 112,31$
On suppose que la variable « durée de sommeil » suit une loi normale. Au risque $\alpha = 5\%$, peut-on accepter l'hypothèse que le nouveau médicament est efficace ?

Exercice 2.6 Une étude porte sur le temps de latence des asthmatiques au test de Rorschach. On applique le test à 65 sujets asthmatiques (choisis au hasard) et on obtient un temps de latence moyen de $15,8$ s avec un écart-type s de $8,51$ s. Le temps moyen de latence des asthmatiques est-il supérieur à 13 s, au risque $\alpha = 1\%$?

Exercice 2.7 Dans une étude sur le développement psychomoteur les enfants PRN (présentant un poids réduit à la naissance), les auteurs ont constitué un échantillon de 31 enfants PRN. Pour chaque enfant, ils ont relevé le score (IDM) de Bayley à 6 mois puis à 24 mois. Le tableau ci-dessous donne les différences d_i des scores (IDM-6 - IDM-24) :

d_i :	10	6	13	-17	12	35	25	14	25	14	14	-14	16	20	2	-4
	-9	-28	38	3	-23	-3	-2	2	9	-17	-9	-7	5	12	-9	

Des travaux antérieurs sur de tels enfants laissent penser que les scores des enfants PRN peuvent diminuer de façon sensible entre 6 et 24 mois. Réaliser un test afin de déterminer si les données observées confirment cette hypothèse (choisir un seuil de 5%).

Exercice 2.8 Le test de Janis et Field permet de mesurer le niveau d'estime de soi (un score élevé indiquant un niveau élevé). Des études ont montré que le niveau moyen des jeunes actifs est de 51,83. Dans le cadre d'une étude sur le chômage, on émet l'hypothèse que le niveau moyen des jeunes inactifs est inférieur à celui des jeunes actifs. On applique le test de Janis et Field à 37 jeunes inactifs choisis au hasard. Leurs scores (X) sont résumés ci-dessous :

$$\bar{x} = 49,38 \quad s = 7,5$$

1) Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse émise ?
2) Calculer le niveau de signification α_{obs} (p-valeur) du test. Quel risque α minimum faut-il prendre pour accepter l'hypothèse émise ?

Exercice 2.9 On s'intéresse à l'effet d'un neuroleptique sur des psychotiques. Pour cela, on considère le temps nécessaire (en heures) pour calmer une crise. On choisit au hasard 40 psychotiques en état de crise et

on leur administre le neuroleptique. On observe sur ce groupe un temps moyen $\bar{x} = 4,35$ et un écart-type $s = 1,12$.
Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que le temps moyen pour calmer une crise est supérieur à 4 heures ? Pour quel risque minimum accepterait-on cette hypothèse ? Quel serait le risque minimum à prendre dans le cas d'une hypothèse alternative bilatérale ?

Exercice 2.10 Compas et al. ont remarqué que les scores obtenus par des jeunes enfants soumis au stress à une échelle de désirabilité sociale sont étonnamment élevés. On sait que le score moyen standard de l'échelle de désirabilité sociale est égal à 3,87.
Pour un échantillon de 36 enfants soumis au stress, Compas et ses collaborateurs ont relevé une moyenne d'échantillon de 4,39 et un écart-type s de 2,61.
On veut tester l'hypothèse selon laquelle les enfants soumis au stress ont un score moyen supérieur à la norme.
Calculer le niveau de signification α_{obs} (p-valeur) du test. Que peut–on conclure à partir des données ?

Exercice 2.11 Un chercheur pense qu'il ne peut y avoir apprentissage sans renforcement (récompense). Pour prouver qu'il a raison, il réalise une expérience sur 39 rats. Il place des rats dans un labyrinthe et mesure le temps mis pour atteindre une case d'arrivée à partir d'une case de départ, ces cases étant toujours les mêmes. Une boulette de viande est placée dans la case d'arrivée.
Les temps relevés sont résumés ci-dessous :

$$\bar{x} = 23 \text{ et } s = 4,2.$$

Des études antérieures ont montré que, sans renforcement, le temps moyen mis par des rats pour parcourir le labyrinthe est de 25 s.
Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que le renforcement favorise l'apprentissage ?
Calculer le niveau de signification α_{obs} du test.

3 Test de comparaison de 2 moyennes à partir de 2 échantillons indépendants

Exercice 3.1 Le poids moyen de 50 étudiants faisant du sport est de 68,2 g pour un écart-type de 2,5. Le poids moyen de 50 étudiants ne faisant pas de sport est de 67,5 kg pour un écart-type de 2,8. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que les étudiants faisant du sport ont un poids moyen supérieur à celui des étudiants ne faisant pas de sport ? Calculer le niveau de signification du test.

Exercice 3.2 Afin de comparer la thérapie familiale et la thérapie cognitivo-comportementale dans le traitement de l'anorexie, on a constitué deux échantillons indépendants de patientes atteintes d'anorexie. Le premier groupe a suivi une thérapie cognitivo-comportementale, le second groupe a suivi une thérapie familiale. Pour chaque sujet, on a relevé le poids au début (P_1) puis à la fin d'une période de traitement (P_2) et mesuré la différence de poids (variable $X = P_2 - P_1$). Les données sont résumés ci-dessous :

groupe « thérapie cognitivo-comportementale » :	$n_1 = 13$	$\bar{x}_1 = 3,58$	$s_1 = 6,86$
groupe « thérapie familiale » :	$n_2 = 12$	$\bar{x}_2 = 7,25$	$s_2 = 7,79$

On admet que les variables différences de poids suivent des lois normales de même variance dans les deux populations de patientes.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que les différences de poids moyennes diffèrent selon les traitements ?

Exercice 3.3 Le modèle de la mémorisation proposé par Craik et Lockhart (1972) stipule que la mémorisation d'un matériel verbal est fonction du traitement de ce matériel lors de sa présentation initiale. Un chercheur veut tester ce modèle et examiner s'il peut contribuer à expliquer certaines différences relevées entre des sujets jeunes et âgés concernant leur aptitude à se rappeler du matériel verbal.
L' étude menée par le chercheur inclut 60 sujets dont l'âge se situe entre 18 et 30 ans et 60 sujets compris dans la tranche d'âge 55-65 ans. Dans chacune des tranches d'âge, les 60 sujets sont répartis au hasard en deux groupes de même taille :
- le premier groupe a pour consigne de trouver une rime pour chaque mot d'une liste de mots (« traitement syntaxique ») ;
- le deuxième groupe doit donner un adjectif pouvant être utilisé pour modifier chaque mot de cette liste (« traitement sémantique »).
Aucun sujet ne sait qu'il lui faudra se rappeler les mots ultérieurement. A la fin de l'expérience, on relève le nombre de mots mémorisés.

1. On veut étudier chez les sujets âgés s'il existe une différence de performance entre le traitement syntaxique et le traitement sémantique, en faveur de ce dernier. Les résultats sont les suivants :

Groupe 1	$\bar{x}_1 = 6,9$	$s_1^* = 2,13$
Groupe 2	$\bar{x}_2 = 11,0$	$s_2^* = 2,49$

   Faire le test pour un risque $\alpha = 5\%$.

2. On veut savoir si pour le traitement syntaxique, il existe une différence de performance due à l'âge. Pour les deux échantillons correspondants les résultats sont les suivants :

Sujets âgés	$\bar{x}_1 = 6,9$	$s_1^* = 2,13$
Sujets jeunes	$\bar{x}_2 = 7,6$	$s_2^* = 1,95$

   Les résultats sont-ils significativement différents au niveau $\alpha = 1\%$? Pour quel risque minimum concluerait-on à une différence de performance ?

Exercice 3.4 Lors d'une expérience pédagogique, on a cherché à comparer deux pédagogies des mathématiques chez deux groupes de sujets :
Le premier groupe a suivi un enseignement selon une méthode traditionnelle (p1), le second selon une méthode nouvelle (p2). Tous les sujets ont ensuite passé une même épreuve. Les notes obtenues sont résumées ci-dessous :

p1 :	$n_1 = 12$	$\overline{x}_1 = 3,25$	$s_1 = 1,365$
p2 :	$n_2 = 10$	$\overline{x}_2 = 4,25$	$s_2 = 1,553$

Ces données expérimentales permettent-elles d'affirmer au risque $\alpha = 1\%$, que la méthode nouvelle donne de meilleurs résultats ?
On admet que les variables notes suivent des lois normales de même variance.

Exercice 3.5 Dans une étude sur le développement de l'enfant, on a émis l'hypothèse que la dépression post-natale de la mère contrarie le développement cognitif de l'enfant. Pour tester cette hypothèse, on a prélevé deux échantillons indépendants d'enfants : un échantillon de 7 enfants dont les mères ont subi une dépression post-natale et un échantillon de 7 enfants dont les mères n'ont pas souffert de cette dépression. On a mesuré le Q.I. de chaque enfant à l'âge de quatre ans :

Q.I. (mère non dépressive) :	$\sum x_i = 763$	$\sum (x_i)^2 = 84147$
Q.I. (mère dépressive) :	$\sum y_i = 699$	$\sum (y_i)^2 = 71327$

On admet que les variables Q.I. suivent des lois normales de même variance dans les deux populations d'enfants. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse émise ?

Exercice 3.6 On a appliqué à 37 travailleurs et 39 chômeurs choisis au hasard et indépendamment le uns des autres, un même test mesurant le niveau d'estime de soi. Sur l'échantillon de travailleurs, on a obtenu un niveau d'estime moyen égal à 53,83 avec un écart-type corrigé de 9,57. Sur l'échantillon de chômeurs, on a obtenu un niveau d'estime moyen égal à 43,13 avec un écart-type corrigé de 7,51.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que le niveau d'estime moyen des travailleurs est supérieur au niveau d'estime moyen des chômeurs ?

4 Tests sur une proportion

Exercice 4.1 Sur un échantillon de 300 patients traités par un certain médicament, 243 ont été guéris. Tester au risque $\alpha = 5\%$, l'hypothèse que la proportion de guérisons est supérieure à 75%.

Exercice 4.2 Dans un sondage d'opinion, 225 électeurs d'une circonscription ont été choisis au hasard et on leur a demandé s'ils pensaient que le premier ministre britannique faisait du bon travail. 75 ont répondu oui. Est-ce une présomption suffisante pour accepter l'hypothèse que moins de 50% de l'électorat pense que le premier ministre fait du bon travail ? ($\alpha = 5\%$.)

Exercice 4.3 75% des sujets normaux qui subissent le test des boulons de la Batterie Standard Bonnardel vissent au moins 36 boulons. On applique ce test à 80 ouvriers de l'automobile et parmi eux, 69 vissent au moins 36 boulons. Pour quel risque minimum, accepterait-on l'hypothèse que l'habileté manuelle des ouvriers est supérieure à la norme ?

Exercice 4.4 Lors d'une étude destinée à évaluer un programme de réinsertion professionnelle, on s'est intéressé aux chômeurs issus d'un même secteur d'activité, ayant subi un licenciement économique et ayant suivi le programme dès leur licenciement. On a prélevé un échantillon de 42 chômeurs dans la population concernée et pour chacun d'entre eux, on a relevé le nombre de mois chômés à partir de la date du licenciement (variable X). Les données x_i sont les suivantes :

1	8	3	7	5	5	4	6	2	9	6	7	7	2	8	7	4	9	5	8	4
4	7	1	2	3	7	5	9	6	5	6	7	7	8	9	9	7	7	9	2	7

Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que dans la population étudiée, la proportion de ceux qui sont restés au chômage moins de 6 mois est inférieure à 50% ? Pour quel risque minimum, accepterait-on cette hypothèse ?

Exercice 4.5 Pour lutter contre la timidité, la plupart des thérapies sont inefficaces. On se demande cependant si la thérapie cognitive apporte des résultats intéressants. On teste pour cela des patients de la manière suivante.
Une première mesure des effets de la timidité (score X) est effectuée au début de la thérapie, et une seconde (score Y) est effectuée par le même questionnaire après 6 mois de thérapie. Un score est d'autant plus grand que le patient souffre de la timidité.
On note p la proportion d'améliorations que fournit la thérapie.
Sur un échantillon de 30 patients ayant suivi la thérapie, on observe 21 scores en baisse et 9 en hausse.
Faire un test sur la proportion p pour tester l'efficacité de la thérapie.

5 Test du Khi-deux d'adéquation (ajustement)

Exercice 5.1 On cherche à savoir si l'exposition répétée à un stimulus modifie l'évaluation de ce stimulus.

Pour cela, on sélectionne trois couleurs notées A, B, C. On sait que dans la population générale, si l'on demande à des sujets de choisir leur « couleur préférée » parmi les trois, la distribution des réponses est uniforme, c'est-à-dire que 1/3 des personnes choisiront chacune des couleurs.
On expose un échantillon de personnes à la couleur A sans aucune consigne pendant 10 minutes. Cela est répété 5 fois. On demande ensuite aux sujets de donner la couleur qu'ils préfèrent. On trouve :

couleur	A	B	C
effectif	46	31	37

Peut-on, au risque $\alpha = 5\,\%$, conclure à un effet de l'exposition répétée ?

Exercice 5.2 Dans une expérience célèbre, G. Mendel a trouvé qu'un échantillon de 556 pois résultant du croisement de pois jaunes et lisses et de pois verts et ridés pouvait être classé de la façon suivante :

pois	jaunes et lisses	jaunes et ridés	verts et lisses	verts et ridés
effectif n_i	315	101	108	32

Selon la théorie proposée par Mendel, les probabilités de ces classes sont respectivement : $\frac{9}{16}$ $\frac{3}{16}$ $\frac{3}{16}$ $\frac{1}{16}$.
Est-ce que cette expérience confirme cette théorie, pour un risque $\alpha = 1\%$?

Exercice 5.3 *Données provenant de l'Inserm (2004).* On souhaite évaluer l'effet éventuel de différentes psychothérapies sur la phobie sociale. On sait que dans une population particulière de patients atteints de phobie sociale, on a la distribution suivante (les patients sont classés en trois niveaux de phobie sociale, le niveau 1 étant le plus faible, le niveau 3 le plus marqué) :

niveau de phobie	niveau 1	niveau 2	niveau 3
proportion	30%	40%	30%

200 patients tirés au sort dans cette population ont suivi une thérapie cognitive-comportementale.
Les résultats post-traitement sont les suivants :

niveau de phobie	niveau 1	niveau 2	niveau 3
effectif	120	50	30

Peut on dire que la thérapie a un effet ? On prendra un risque $\alpha = 5\%$.

Exercice 5.4 En 1982, on avait appliqué un test de raisonnement aux élèves de 5ème et on avait obtenu la répartition des notes suivante (en pourcentages) :

note X	< 100	$100 \leq X < 110$	$110 \leq X < 120$	$120 \leq X < 130$	$130 \leq X$
proportion en %	17,2	19	17	17,9	28,9

En 1987, on applique le test à 50 élèves de même niveau et on observe les effectifs suivants :

note X	< 100	$100 \leq X < 110$	$110 \leq X < 120$	$120 \leq X < 130$	$130 \leq X$
effectif	14	15	4	5	12

Peut-on affirmer que la répartition des élèves en 1987 est conforme à la première ? On prendra un risque $\alpha = 5\%$.

Exercice 5.5 Lors d'une élection municipale où quatre candidats s'affrontent, on réalise un sondage d'opinion auprès de 200 habitants de la ville. Ce sondage fait apparaître la répartition suivante :

candidat	1	2	3	4
effectif	50	75	30	45

1) Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que les quatre candidats n'ont pas la même cote de popularité dans la ville ?
2) Peut-on, au risque $\alpha = 2\%$, accepter l'hypothèse que la proportion d'habitants favorables au candidat n° 2 est inférieure à 50% ?

6 Tests du Khi-deux d'indépendance et d'homogénéité

Exercice 6.1 Lors de l'embauche de pilotes, chaque candidat est soumis à un test psychologique qui le classe comme introverti ou extraverti, et à un test d'aptitude au pilotage où il peut être déclaré apte ou inapte. Sur un échantillon de 120 candidats, on a obtenu les résultats suivants :

	introverti	extraverti
apte	14	34
inapte	31	41

Au risque $\alpha = 1\%$, les résultats suggèrent-ils une association entre aptitude au pilotage et type de personnalité ?

Exercice 6.2 Le tableau suivant indique le résultat de l'examen de 528 sujets classés d'après la couleur de leurs yeux et celle de leurs cheveux :

Cheveux \ Yeux	bleus	noirs	clairs
chatains ou noirs	61	127	97
blonds	163	42	38

Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse d'une liaison entre les deux variables ?

Exercice 6.3 Dans un échantillon de 400 femmes et de 300 hommes, on observe que 25 femmes et 25 hommes développent une certaine forme de maladie mentale.
Au risque $\alpha = 1\%$, ces résultats suggèrent-ils un lien entre le sexe et le fait de développer la maladie ?

Exercice 6.4 Dans une étude sur la qualité de vie et la santé alimentaire, on a recensé les comportements alimentaires de 500 personnes ainsi que si un cancer leur était survenu ou pas. :

	Cancer	Pas de Cancer
Restauration rapide	55	70
Végétarienne	5	45
Cuisine traditionnelle	65	260

Tester l'hypothèse nulle selon laquelle le cancer est une maladie indépendante des comportements alimentaires. On prendra un risque de $\alpha = 1\%$.

Exercice 6.5 Howell et Huessy (1981) ont utilisé une échelle d'évaluation pour classer un échantillon d'enfants de deuxième année en deux catégories selon qu'ils manifestent ou non le comportement communément associé au trouble de déficit de l'attention. Lorsque les enfants ont atteint la fin de la dernière année, les chercheurs ont examiné les bulletins scolaires et ont relevé quels enfants suivaient des cours de rattrapage en anglais.
Les données suivantes rassemblent en un groupe (appelé Déficit de l'attention) tous les enfants qui ont été classés, à un moment ou à un autre, dans la catégorie des sujets manifestant un comportement associé au trouble de déficit de l'attention :

	Cours de rattrapage en Anglais	Pas de cours de rattrapage
Normaux	22	187
Déficit de l'attention	19	74

Le comportement manifesté à l'école primaire influence-t-il la répartition en classe à l'école secondaire ? On prendra $\alpha = 5\,\%$.

Exercice 6.6 Le tableau qui suit (Rosenthal 1990) donne le nombre de crises cardiaques observées chez des personnes qui ont consommé pendant longtemps et quotidiennement soit de l'aspirine, soit un placebo :

	crise cardiaque	pas de crise cardiaque
aspirine	104	10933
placebo	189	10845

Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse d'une liaison entre les deux variables ?

Exercice 6.7 On s'intéresse à la participation à un club sportif dans deux collèges A et B. On a prélevé au hasard 50 élèves dans le collège A et 60 dans le collège B. On observe que 12 des 50 élèves du collège A et 26 des 60 élèves du collège B sont membres d'un club sportif.

Au risque $\alpha = 1\%$, peut-on dire que les résultats sont significativement différents dans les deux collèges ?

Exercice 6.8 Une enquête sur les joueurs adeptes du jeu de rôle en ligne Everquest a été réalisée auprès de joueurs de différents pays ayant répondu à un questionnaire. Un échantillon de 82 adolescents et un échantillon de 436 adultes ont été constitués pour cette étude.
Pour chaque joueur on a relevé la « durée hebdomadaire de jeu » , c'est-à-dire le nombre moyen d'heures de jeu par semaine, regroupée en trois classes : « moins de 15 heures » , « entre 15 et 30 heures » et plus de « 30 heures ».
Les effectifs observés sont donnés dans le tableau suivant :

Durée \Age	adolescent	adulte
« moins de 15 heures »	23	119
« entre 15 et 30 heures »	38	216
« plus de 30 heures »	21	101

Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que la durée hebdomadaire de jeu diffère selon l'âge des joueurs ?
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que plus de 25 % des adolescents jouent (en moyenne) plus de 30 heures par semaine ?
Pour quel risque minimum peut-on accepter cette hypothèse ?

Exercice 6.9 La deuxième épreuve du test Droite-Gauche de Piaget a pour but de reconnaître sur autrui, placé face à face, la main gauche de la main droite. On veut tester l'hypothèse que le taux de réussite à l'épreuve n'est pas le même chez les droitiers et chez les gauchers. Pour cela, on fait passer l'épreuve à 50 gauchers et 50 droitiers de 8 ans choisis au hasard.
Parmi les gauchers, 45 enfants ont réussi et chez les droitiers, 33 enfants ont réussi. Au risque $\alpha = 1\%$, peut-on accepter l'hypothèse émise ?

7 Exercices de révision

Exercice 7.1 L'épreuve des signes arithmétiques comporte 26 égalités dans lesquelles manquent les signes arithmétiques comme par exemple « 6...2 = 8 ». On applique ce test à 24 filles de 10 ans. Chacune doit compléter les égalités et on compte le nombre de bonnes réponses. On observe sur cet échantillon, un nombre moyen de bonnes réponses $\bar{x} = 10$ avec un écart-type $s^* = 2,1$.
On suppose que la variable « nombre de bonnes réponses » suit une loi normale dans la population des filles de 10 ans. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que le nombre moyen de bonnes réponses est différent de 11 ?

Exercice 7.2 Une épreuve de dextérité manuelle vise à déterminer le degré de facilité à manipuler des objets très petits, à l'aide d'un instrument. L'épreuve consiste à loger dans un tube de verre des rondelles à l'aide d'une pince fine. On relève le temps nécessaire (en mn) pour loger 50 rondelles.
On fait passer l'épreuve à 37 garçons de 13 à 15 ans choisis au hasard. Les résultats sont les suivants :

$$\bar{x} = 1,42 \qquad s = 0,24.$$

On émet l'hypothèse que le temps moyen pour les garçons de 13 à 15 ans est inférieur à 1,5 mn. Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse émise ? Quel risque minimum doit-on prendre pour accepter cette hypothèse ?

Exercice 7.3 On lance un dé 450 fois et on enregistre le nombre de fois où chacune des faces apparaît :

face	1	2	3	4	5	6
effectif n_i	62	50	76	68	111	83

Peut-on, au risque $\alpha = 5\%$, rejeter l'hypothèse que le dé est bien équilibré ?

Exercice 7.4 On souhaite étudier l'effet de la privation de rêves sur le niveau d'anxiété d'un sujet. Des études antérieures ont montré que le niveau moyen d'anxiété pour les sujets qui ne sont pas privés de rêves est de $26,5$.
Lors d'une expérience, 40 sujets choisis au hasard ont été privés de rêves pendant trois nuits consécutives. On leur a administré ensuite un test mesurant leur niveau d'anxiété. Les résultats sont les suivants :
$\bar{x} = 28,25 \quad s = 8,81$.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que la privation de rêves augmente le niveau d'anxiété ? Quel risque α minimum doit-on prendre pour accepter cette hypothèse ?

Exercice 7.5 On veut étudier les changements structuraux du cerveau et en particulier l'atrophie de l'hippocampe chez les personnes atteintes de schizophrénie. Pour cela, on a mesuré par IRM le volume (en cm^3) de l'hippocampe gauche de 15 sujets atteints de schizophrénie. Les volumes (x_i) sont résumés ci-dessous :
$\sum x_i = 23,4 \quad \sum x_i^2 = 37,78$.
On admet que le volume suit une loi normale dans la population étudiée.

Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que le volume moyen de l'hippocampe gauche chez les sujets atteints de schizophrénie est inférieur à 1,7 cm^3 ?

Exercice 7.6 Dans un hôpital psychiatrique, on se demande s'il y a un lien entre la saison et la présence ou l'absence de réaction à un certain médicament. Sur un échantillon de 490 patients, les observations sont les suivantes :

Saison\ Réaction	Absence	Présence
P	55	64
E	59	60
A	52	63
H	60	77

Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse d'un lien ?

Exercice 7.7 On émet l'hypothèse que chez les ouvriers travaillant en usine, le nombre d'accidents subis dans l'année suit la loi théorique suivante :

Nombre d'accidents subis	0	1	2	3 ou plus
proportion	0,57	0,32	0,08	0,03

Sur un échantillon de 220 ouvriers travaillant en usine, on a relevé les nombres de personnes ayant subi 0, 1, 2, 3... accidents dans l'année :

Nombre d'accidents subis	0	1	2	3 ou plus
effectif observé	154	50	15	1

Tester l'hypothèse émise pour un risque $\alpha = 5\%$.

Exercice 7.8 Dans quelle mesure l'alcool affecte-t-il le développement cérébral prénatal ? Pour étudier cette question, on s'est intéressé aux mères d'un enfant, présentant les mêmes caractéristiques (éducation, statut conjugal, etc.). On a retenu 6 femmes qui ont été des alcooliques chroniques durant leur grossesse et, indépendamment, on a constitué un groupe témoin de 12 femmes n'ayant eu aucune tendance à l'alcoolisme durant leur grossesse. Dans les deux groupes, chaque enfant a passé le même test de Q.I. à l'âge de 7 ans. Les scores relevés sur ces 2 échantillons sont les suivants :

groupe « alcoolique » : $n_1 = 6$ $\quad \overline{x}_1 = 78$ $\quad s_1 = 8,18$
groupe témoin **:** $n_2 = 12$ $\quad \overline{x}_2 = 99$ $\quad s_2 = 10,13$

On suppose que dans les deux populations étudiées, les scores suivent une loi normale, de même variance. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que l'alcool contrarie le développement cérébral prénatal ?

Exercice 7.9 La distribution d'effectifs suivante a été établie par Haberman (1978) à partir de données fournies par le National Opinion Research Center de l'Université de Chicago, et recueillies auprès d'un échantillon de 776 individus. Pour chaque individu, on a relevé le nombre d'années de scolarité (variable X) et l'attitude face à l'avortement (variable Y) :

$X\backslash Y$	Pour	Indifférent	Contre
Inférieur à 8 ans	31	23	56
Entre 8 et 12 ans	171	89	177
Supérieur à 12 ans	116	39	74

1) Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse d'une liaison entre la durée de la scolarité et l'opinion face à l'avortement ?
2) On s'intéresse ici aux individus dont la durée de scolarité est inférieure à 8 ans. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que le rejet de l'avortement est majoritaire chez ces individus ? Quel risque minimum faudrait-il prendre pour accepter cette hypothèse ?

Exercice 7.10 Le directeur d'une caisse d'assurance maladie souhaite étudier l'influence de la liberté de choix du médecin traitant sur le degré de satisfaction des patients. Pour réaliser cette étude, 30 patients choisis au hasard sont répartis au hasard en deux groupes A et B de même taille. Les 15 patients du groupe A peuvent choisir leur médecin sur une liste qui leur est proposée. Les 15 patients du groupe B se voient imposer leur médecin traitant. (L'ensemble des médecins est le même pour les deux groupes).
Au bout d'un an, on fait passer aux 30 patients un même test mesurant leur niveau de satisfaction (un score élevé correspondant à un haut niveau de satisfaction). Les scores relevés sont résumés ci-dessous :

groupe A (médecin choisi) : $n_1 = 15$ $\overline{x}_1 = 65,07$ $s_1 = 15,22$
groupe B (médecin imposé) : $n_2 = 15$ $\overline{x}_2 = 52,47$ $s_2 = 14,58$

Dans les deux cas, les variables « scores » suivent une loi normale de même variance. Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que la possibilité de choisir son médecin traitant a une influence bénéfique sur le niveau de satisfaction ?

Exercice 7.11 Lors d'une enquête, 200 habitants d'une ville, qui ont déclaré vouloir prendre des vacances en France cet été, ont indiqué la région de leur choix selon un découpage en cinq modalités : Ouest, Sud, Est, Nord et Centre. La répartition observée est la suivante :

Régions R_i	O	S	E	N	C
effectifs n_i	50	91	22	15	22

1) Au risque $\alpha = 5\%$, tester l'hypothèse selon laquelle, pour les habitants de la ville prenant leurs vacances en France cet été, les proportions d'entre-eux ayant choisi ces diverses régions valent respectivement 22%, 33%, 17%, 10% et 18%.
2) Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que les habitants de la ville prenant leurs vacances en France cet été ont minoritairement choisi le Sud ? Pour quel risque minimum, accepterait-on cette hypothèse ?

Exercice 7.12 Dans une étude sur les aspects cliniques de la dépression chez les adolescents, on a constitué deux échantillons indépendants d'adolescents âgés de 15 ans : un échantillon de 358 adolescents de mères dépressives et un échantillon de 458 adolescents de mères non dépressives.
Dans chaque échantillon, on a relevé le nombre d'adolescents dépressifs (épisode dépressif majeur ou dysthymie). Les résultats sont les suivants :

	Mère dépressive	Mère non dépressive
Adolescent dépressif	65	45
Adolescent non dépressif	293	413

1) On s'intéresse ici aux adolescents de mères dépressives.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que dans cette population, la proportion de dépressifs est supérieure à 17 % ? Pour quel risque minimum accepterait-on cette hypothèse ?
2) On s'intéresse ici aux adolescents dépressifs.
On émet l'hypothèse que pour les adolescents dépressifs, les premiers troubles apparaissent chez les enfants de mères dépressives plus tôt que chez les enfants de mères non dépressives. Pour les deux échantillons d'adolescents dépressifs concernés, on a relevé l'âge auquel les premiers troubles se sont déclarés. Les âges (en années) sont résumés dans le tableau suivant :

Adolescents dépressifs de mères dépressives : $n_1 = 65$ $\bar{x}_1 = 12,22$ $s_1 = 2,65$
Adolescents dépressifs de mères non dépressives : $n_2 = 45$ $\bar{x}_2 = 12,6$ $s_2 = 2,38$

Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse émise ?
3) On veut comparer la proportion d'adolescents dépressifs dans les deux populations d'adolescents.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que la proportion d'adolescents dépressifs diffère selon l'état de santé (dépressif, non dépressif) de la mère ?
(Source : Constance Hammen, Journal of Consulting and Clinical Psychology, 2001)

Exercice 7.13 Une entreprise souhaite modifier l'organisation du travail en créant des équipes de travail autonomes.
Un sondage d'opinion est réalisé auprès de 3 échantillons indépendants de salariés choisis au hasard dans l'entreprise : un groupe de 285 ouvriers, un groupe de 75 cadres moyens et un groupe de 40 cadres supérieurs. Chacun des 400 salariés ainsi interrogés doit dire s'il est favorable ou bien opposé à la réforme. Les résultats sont les suivants :

	ouvriers	cadres moyens	cadres supérieurs
favorable	184	49	31
opposé	101	26	9

Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que dans l'entreprise, l'opinion vis à vis de la réforme diffère suivant le type d'emploi ?

Examen de Statistique, janvier 2006

Exercice 1 Un psychologue du travail souhaite étudier les effets des contrats de travail à durée déterminée (CDD) sur le moral. Pour son étude, il a constitué deux échantillons de travailleurs :
– un échantillon de 13 travailleurs en CDI (contrat de travail à durée indéterminée) ;
– un échantillon de 10 travailleurs en CDD.
Il a évalué ensuite leur niveau d'anxiété par un score, un score élevé traduisant un niveau d'anxiété élevé. On admet que les scores obéissent à des lois normales de même variance dans les deux populations étudiées. Les scores relevés sur les deux échantillons sont résumés ci-dessous :

Scores du groupe CDI :	$n_1 = 13$	$\bar{x}_1 = 27,3$	$s_1 = 7,21$
Scores du groupe CDD :	$n_2 = 10$	$\bar{x}_2 = 34,2$	$s_2 = 7,09$

Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que les contrats de travail à durée déterminée ont un effet négatif sur le moral ?

Exercice 2 Des chercheurs désiraient savoir si l'obésité était perçue négativement chez les jeunes enfants. Pour cela, ils ont réalisé une expérience afin de voir si les enfants préfèrent majoritairement choisir comme compagnon de jeu, un enfant de corpulence normale.
Ils ont montré à un groupe d'enfants entre 5 et 7 ans deux images : l'une représentant un enfant de cette classe d'âge de corpulence normale et l'autre représentant un enfant de la même classe d'âge et obèse.
Les chercheurs ont demandé aux enfants de choisir parmi ces deux images celle qui représentait l'enfant avec lequel ils voudraient jouer.
Sur les 60 enfants participant à l'étude, 42 ont préféré l'enfant de corpulence normale.
Peut-on, au risque $\alpha = 5\%$, accepter l'hypothèse que l'obésité est perçue négativement chez les enfants de cette classe d'âge ?
Calculer le niveau de signification α_{obs} du test.

Exercice 3 Dans une étude destinée à évaluer l'efficacité d'une thérapie cognitive de groupe s'adressant à des hommes ayant subi des agressions sexuelles durant l'enfance (ASE), les chercheurs ont constitué un échantillon de 34 hommes ayant vécu des ASE.
L'échelle TSC-40 a été utilisée pour mesurer le niveau de symptômes traumatiques de chaque individu.

Plus le score sur l'échelle est élevé et plus les symptômes traumatiques sont sévères.
Pour chaque individu, on a relevéle score au début de la thérapie (score T_1), le score à la fin de la thérapie (score T_2), puis on a calculé la différence entre les deux scores (variable $X = T_1 - T_2$).
Les résultats sont les suivants :

$$\text{Différences des scores } x_i : \quad n = 34 \quad \bar{x} = 4,257 \quad s = 11,9$$

Peut-on, au risque $\alpha = 1\%$, accepter l'hypothèse que la différence moyenne des scores est positive dans la population étudiée ? Calculer le risque minimum à prendre pour accepter cette hypothèse. Que peut-on en conclure sur l'efficacité de la thérapie ?

Exercice 4 Dans une étude, des chercheurs ont émis l'hypothèse que chez les adolescents, les mauvais traitements ont une influence sur l'apparition des comportements à risque tels que le fait de fuguer.
Ils ont constitué un échantillon de 142 adolescents. La répartition des adolescents selon les antécédents de mauvais traitements et le fait de fuguer (oui, non) est donnée dans le tableau suivant :

Fugue	Adolescents maltraités	Adolescents non maltraités
oui	29	5
non	58	50

La p-valeur obtenue est $\alpha_{\text{obs}} = 0,097\%$. Que peut-on en conclure pour un risque $\alpha = 5\%$? Détaillez la procédure du test. On pourra prendre la décision à l'aide de la p-valeur en énonçant la règle de décision adéquate.