Feuille 6 : Tests

Exercice 1 On cherche à connaître la température d'ébullition μ, en degrés Celsius, d'un certain liquide. On effectue 16 expériences indépendantes pour mesurer cette température d'ébullition. On suppose que les mesures sont distribuées selon une variable aléatoire suivant une loi Normale d'espérance μ et d'écart-type inconnu. La moyenne des mesures sur les 16 expériences est $94,32$ et l'écart-type corrigée $1,2$. Face à ces résultats on souhaite savoir si μ vaut 95 degrés Celsius ou non.

1. Décrire précisément les hypothèses à tester, proposer une règle de décision au niveau $\alpha = 5\%$ et conclure.
2. Que vaut la p-valeur de l'échantillon observé pour le test précédent ? Conclure au test en conséquences.

Exercice 2 Une usine fabrique des billes métalliques. Le diamètre en millimètres de ces billes est une variable aléatoire suivant une loi Normale. L'usine s'est engagée à fournir à un client des billes dont le diamètre moyen est de 25 millimètres. Le client réceptionne sa commande. Dans le lot reçu, il prélève un échantillon de 20 billes choisies au hasard et mesure les diamètres suivants :

$24,7$	$24,9$	$25,0$	$25,0$	$25,1$	$25,1$	$25,1$	$25,2$	$25,3$	$25,4$
$24,8$	$24,9$	$25,0$	$25,0$	$25,1$	$25,1$	$25,2$	$25,3$	$25,3$	$25,5$

Peut-on dire que l'usine a respecté ses engagements ? Faire un test d'hypothèses pour y répondre.

Exercice 3 Un fabriquant de gâteaux commercialise ses produits avec sur l'emballage la mention "Teneur moyenne en lipides inférieure ou égale à 20 grammes". On suppose que la teneur en lipides en grammes d'un gateau est une variable aléatoire réelle X suivant une loi Normale d'espérance μ. Le fabriquant veut contrôler sa production. Il prélève au hasard 7 gâteaux. Les résultats sont les suivants : 20 ; 23 ; 19 ; 23 ; 22 ; 20 ; 19.

1. Proposer une région critique au niveau $\alpha = 5\%$ pour tester $H_0 : \mu \leqslant 20$ contre $H_1 : \mu > 20$, et conclure pour les observations précédentes.
2. Proposer une région critique au niveau $\alpha = 5\%$ pour tester $H_0 : \mu \geqslant 20$ contre $H_1 : \mu < 20$, et conclure pour les observations précédentes.
3. Commenter les deux résultats précédents.
   Le fabriquant décide de prélever 5 autres gateaux pour compléter son échantillon. Il obtient les mesures : 21 ; 22 ; 23 ; 20 ; 22.
4. Calculer la p-valeur de cet échantillon de 12 valeurs pour les deux tests précédents et conclure.

Exercice 4 Un laboratoire pharmaceutique doit doser avec précision ses médicaments. Le procédé industriel utilisé garantit un dosage moyen $\mu = 1\,\text{mg}$, conforme à celui attendu, mais avec une légère variance notée σ^2. On suppose que ce dosage se distribue selon une loi normale.

1. Un surdosage du médicament peut s'avérer dangereux. Ainsi le laboratoire pharmaceutique exige qu'avec une probabilité supérieure à 0.99, le dosage ne dépasse pas 1.1 mg. A quel valeur maximal de σ cela correspond-il ? On note cette valeur σ_{max}.
2. Afin de vérifier le dosage, le laboratoire prélève 100 médicaments. Il relève une moyenne de 1.01 mg et un écart-type corrigé de 0.04. Peut-on considérer que l'écart-type est bien inférieur à σ_{max} ?

Exercice 5 Une jardinerie veut stocker ses paquets des graines végétales invendus cette année pour les vendre le printemps suivant. Toutefois, le centre s'inquiète que les graines peuvent ne pas germer au même taux après un an. Le directeur trouve un paquet de l'année dernière et plante les graines comme essai. Bien que le paquet indique un taux de germination de 92%, seulement 171 de 200 graines poussent. On souhaite savoir si ce résultat indique de façon significative que les graines ont perdu leur viabilité après une année en stockage.

1. En notant p la probabilité qu'une graine pousse, énoncer les hypothèses à tester.
2. Proposer un estimateur de p et étudier sa convergence en loi.
3. En déduire une région critique au niveau asymptotique α pour le test précédent, où $\alpha \in [0, 1]$.
4. Evaluer la p-valeur asymptotique du test pour les observations de l'énoncé et conclure.

Exercice 6 On modélise la durée de vie X d'un composant électronique par une loi exponentielle de paramètre $\lambda > 0$. On appelle demi-vie la valeur T telle que $P(X < T) = P(X > T) = 1/2$. Dans un précédent exercice, on a établi que $T = \ln 2/\lambda$ et qu'un intervalle de confiance au niveau $1 - \alpha$ pour T, basé sur un n-échantillon $(X_1, \dots, X_n)$ est

$$IC = \left[\ln 2 \frac{\sum_{i=1}^n X_i}{\gamma_{(n,1)}(1 - \frac{\alpha}{2})} ; \ln 2 \frac{\sum_{i=1}^n X_i}{\gamma_{(n,1)}(\frac{\alpha}{2})} \right]$$

où pour tout $p \in [0, 1]$, $\gamma_{(n,1)}(p)$ désigne le quantile d'orde p d'une loi $\Gamma(n, 1)$.

1. Soit $T_0 > 0$ fixé. On désire tester $H_0 : T = T_0$ contre $H_1 : T \neq T_0$. Pour cela, on propose la règle de décision suivante : si $T_0 \notin IC$, alors on rejette H_0 et si $T_0 \in IC$, alors on ne rejette pas H_0.
   Quel est le niveau de ce test ?
2. On désire à présent tester $H_0 : T \leqslant T_0$ contre $H_1 : T > T_0$. Pour cela on propose la règle de décision suivante : si $T_0 < \ln 2 \frac{\sum_{i=1}^n X_i}{\gamma_{(n,1)}(1-\frac{\alpha}{2})}$, alors on rejette H_0, sinon on ne rejette pas H_0.
   Quel est le niveau de ce test ?
3. En déduire un test de niveau 5% pour tester $H_0 : T \leqslant T_0$ contre $H_1 : T > T_0$.

Exercice 7 Deux candidats, A et B, se confrontent lors d'une élection.

1. Un institut de sondage effectue une étude auprès de 1000 personnes tirées aléatoirement avec remise. Il observe que 48% des personnes interrogées voteront pour le candidat A, tandis que 52% voteront pour le candidat B. Proposer un intervalle de confiance au niveau asymptotique 95% pour estimer la proportion d'individus dans la population totale qui voteront pour A.
2. On désire tester si le candidat A obtiendra moins de 50% des votes ou non. Déduire du résultat précédent une décision. Quel est le niveau du test utilisé ?
3. Un autre institut de sondage effectue la même étude auprès de 1000 personnes et obtient une proportion de 47.5% de votes pour A. Effectuer le même test que ci-dessus à partir de cet échantillon.
4. En regroupant les deux échantillons ci-dessus, quelle est la proportion de votes estimée pour A ? Que devient le résultat du test précédent ?

Exercice 8 Le revenu annuel en euro d'un individu d'une certaine population est une variable aléatoire réelle X de densité

$$f(x) = \begin{cases} \theta x^{-(\theta+1)} & \text{si } x \geqslant 1, \\ 0 & \text{sinon.} \end{cases}$$

Le paramètre $\theta > 1$ est un réel inconnu. Soient $n \in \mathbb{N}^*$, et $(X_1, \ldots, X_n)$ un n-échantillon de X. On pose $T_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \ln X_i$.

1. Montrer que T_n est un estimateur efficace de $\dfrac{1}{\theta}$.
2. Étudier la convergence en loi de $\left(\sqrt{n}\,(\theta T_n - 1)\right)_{n \in \mathbb{N}^*}$.
3. En déduire une région critique au niveau asymptotique $\alpha \in [0, 1]$ pour tester
$$H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta \neq \theta_0,$$
basée sur l'échantillon $(X_1, \ldots, X_n)$, où θ_0 est un réel fixé.

Exercice 9 Le nombre d'absences par semaine dans une entreprise est une variable aléatoire réelle X suivant la loi de Poisson de paramètre λ. Donc

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k \in \mathbb{N}.$$

Ici, $\lambda > 0$ est un réel inconnu. Soient $n \in \mathbb{N}^*$ et $(X_1, \ldots, X_n)$ un n-échantillon de X. On considère le test d'hypothèses :

$$H_0 : \lambda = 1 \text{ contre } H_1 : \lambda \neq 1.$$

1. Déterminer l'estimateur du maximum de vraisemblance $\hat{\lambda}$ de λ.
2. Quel est la loi de $n\hat{\lambda}$ lorsque $\lambda = 1$?

3. En déduire une région critique au niveau $\alpha \in [0, 1]$ pour le test précédent, en supposant que le niveau α est admissible.
4. Que vaut la p-valeur du test précédent pour un échantillon $(X_1, \ldots, X_n)$?
5. On dispose des données suivantes, où n_i est le nombre de semaines pour lesquelles on a relevé i absences :

i	0	1	2	3
n_i	5	6	2	3

Calculer la p-valeur du test avec ces données et conclure.

Exercice 10 On désire tester si un médicament a une influence sur le comportement psychomoteur. On choisit au hasard 20 sujets qu'on répartit au hasard en deux groupes : le groupe témoin et le groupe expérimental. On leur fait subir la même expérience psychomotrice. On a administré auparavant le médicament aux sujets du groupe expérimental et un placebo au groupe témoin. Les résultats sont les suivants :

Groupe témoin	166	167	169	170	174	173	172	170	166	173
Groupe expérimental	167	162	165	168	162	160	164	158	165	169

On suppose que dans chaque groupe les résultats sont distribués selon une loi gaussienne, que la variance est la même pour les deux groupes et que les performances des sujets sont indépendantes. Tester au niveau 0.05 l'hypothèse selon laquelle le médicament n'a aucun effet sur le comportement psychomoteur.

Exercice 11 Une usine fabrique une fibre synthétique dans deux ateliers. Un échantillon de 25 spécimens de fibres provenant de l'atelier 1 et de 30 provenant de l'atelier 2 donnent une résistance moyenne à la rupture égale à $64 kg/cm^2$ pour le premier et à $61 kg/cm^2$ pour le second. Par ailleurs, la variance empirique corrigée pour le premier échantillon vaut 7, tandis qu'elle vaut 8 pour le second. On suppose que la résistance à la rupture est une variable aléatoire gaussienne et que les résistances des fibres sont indépendantes entre elles. Peut-on conclure au niveau 0.05 qu'il n'y a pas de différence concernant la résistance à la rupture des fibres produites par les deux ateliers ?

Exercice 12 On désire tester l'efficacité d'un médicament contre l'hypertension. On a mesuré la tension de 20 patients avant et après la prise du médicament. Tester l'efficacité du médicament.

Avant	17.8	16.1	17.4	15.3	14.8	16.3	17.6	16.6	18.1	15.4	17.7	16.9
Après	16.2	15.0	15.4	14.0	14.5	17.3	15.6	16.7	18.1	14.9	17.4	16.5

Exercice 13 Deux cent jets d'un même dé ont donné les résultats suivants :

j : numéro obtenu	1	2	3	4	5	6
n_j : nombre d'apparitions de j	30	26	29	30	40	45

On se demande si le dé est pipé.

1. Quelle statistique de test utiliseriez vous pour répondre à cette question ? (Contentez vous d'indiquer les calculs numériques que vous auriez à faire, mais ne les faites pas).
2. Tous calculs faits on trouve que la valeur prise dans cette experience par la statistique de test est 8.26. Dans quelle table lisez vous la conclusion de votre test ? Pourquoi ? Quelle est cette conclusion ?

Exercice 14 La marque Smorties produit des bonbons au chocolat de six couleurs. Le responsable de la communication affirme que la proportion de chaque couleur est très précisemment de 30% pour le brun (B), 20% pour le jaune (J), 20% pour le rouge (R), 10% pour l'orange (O), 10% pour le vert (V) et 10% pour le doré (D) dans tout échantillon de grande taille. Une expérience réalisée sur un échantillon de 370 bonbons donne les comptages suivants :

couleur	B	J	R	O	V	D
nombre de bonbons	84	79	75	49	36	47

Peut-on conclure, au niveau 5%, que le responsable de la communication a tort ?

Exercice 15 Un cabinet de communication est chargé de mettre en place un dispositif d'information juridique portant sur le harcèlement sexuel des femmes au travail. Avant de lancer une campagne nationale, ce cabinet décide de tester différents formats de communication et souhaite évaluer les éléments retenus par les femmes à qui ils sont présentés. Deux cent dix-sept femmes acceptent d'y prendre part. Comme il est important de bien contrôler le niveau des participantes à ce pré-test, le cabinet souhaite vérifier si l'échantillon possède les mêmes caractéristiques de formation scolaire que celles que l'on trouve dans la population féminine générale. À l'aide d'une classification en six catégories de diplômes, on observe les répartitions suivantes :

	pourcentage national*	nombre de femmes
aucun diplôme ou certificat d'étude primaire	$20,56\%$	51
BEPC seul	$7,65\%$	8
CAP, BEP, ou autre diplôme de ce niveau	$24,05\%$	60
Bac, BP, ou autre diplôme de ce niveau	$17,61\%$	35
Bac + 2	$15,57\%$	38
diplôme supérieur à Bac +2	$14,56\%$	25

*Source : Insee, enquête Emploi de 2003. France métropolitaine, individus de 25 à 54 ans.

Peut-on affirmer que la répartition des femmes de l'échantillon respecte la répartition nationale ?

Exercice 16 Le tableau ci-dessous donne le nombre d'étudiants qui ont été brillants et médiocre devant trois examinateurs : examinateur A, examinateur B et examinateur C.

résultat \ examinateur	A	B	C
brillants	50	47	56
médiocres	5	14	8

Peut-on affirmer, au niveau 5%, que le résultat d'un individu dépend de l'examinateur ?

Exercice 17 Le tableau suivant indique le classement de 124 personnes selon la couleur de leurs yeux et la couleur de leurs cheveux. Peut on conclure à l'indépendance de ces deux caractères ?

yeux \ cheveux	Blonds	Bruns	Noirs	Roux	Total
Bleus	25	9	3	7	44
Gris ou verts	13	17	10	7	47
Marrons ou noirs	7	13	8	5	33
Total	45	39	21	19	124