TD N°2 : Tests d'hypothèses

TESTS SUR UNE CARACTERISTIQUE

Exercice 1

Une machine produit des billes de roulement de diamètre fixe. Si elle fonctionne normalement, il y a une proportion de 5% de billes défectueuses. Si elle est déréglée, la proportion de billes défectueuses est de 10%.
Avant d'envoyer une commande à son client, il teste un lot de 500 billes. Soit X le variable aléatoire représentant le nombre de billes défectueuses sur ce lot. On sait que X suit une loi binomiale b(500,p) où p=0.05 si la machine fonctionne correctement et p=0.1 si la machine est déréglée.

1. Il décide de ne pas livrer son client si la machine est déréglée, c'est-à-dire si le nombre de billes défectueuses est supérieur ou égal à 50 (=500×10%).
   (a) Exprimer les hypothèses H_0 et H_1 ainsi que les risques de $1^{ère}$ et $2^{ème}$ espèces.
   (b) Calculer le risque de $1^{ère}$ espèce.
   (c) Calculer la puissance du test et en déduire si un client doit acheter ou non un lot ayant subi un tel test.
2. Le fabricant re-définit son test en imposant un risque de $1^{ère}$ espèce de 1%.
   (a) Déterminer la région critique du test.
   (b) Calculer sa puissance.
   (c) Enoncer les règles de décision avec les erreurs associées.
3. Le client décide lui aussi d'effectuer son test avec un risque de $1^{ère}$ espèce de 1%.
   (a) Exprimer les hypothèses H_0 et H_1 ainsi que les risques de $1^{ère}$ et $2^{ème}$ espèces.
   (b) Déterminer la région critique du test.
   (c) Calculer sa puissance.
   (d) Enoncer les règles de décision avec les erreurs associées.
   (e) Sur un même lot, on trouve 35 billes défectueuses. Que se passe-t'il ?
4. Le fabricant et le client se mette d'accord pour effectuer chacun leur test de risque de $1^{ère}$ espèce de 1% sur un échantillon de taille *n*.
   (a) Calculer la taille de l'échantillon afin que les deux tests donnent le même seuil critique.
   (b) Calculer les risques de $2^{ème}$ espèce.

Exercice 2

Les habitants d'une région aéroportuaire se plaignent que le bruit des avions dépasse la limite autorisée de 80 décibels en moyenne imposée par la législation. On admet que l'intensité du bruit causé par les avions est une variable aléatoire X de loi gaussienne d'espérance μ et de variance 64.

On mesure un échantillon journalier de n=16 variables aléatoires indépendantes $X_1, \dots, X_n$ de l'intensité du bruit, et on effectue le test statistique suivant.

$$\begin{cases} H_0 : \mu=\mu_0=80 \text{ décibels} \\ H_1 : \mu=\mu_1=85 \text{ décibels} \end{cases}$$

1) Expliciter les risques de première et deuxième espèces. De quel point de vue est fait ce test ? Celui des habitant ou celui des responsables de l'aéroport ?
2) Quelle variable de décision faut-il choisir et quelle est sa loi ?
3) Déterminer graphiquement l'allure de la région critique et représenter sur le graphique les erreurs de première et deuxième espèces.
4) Calculer le seuil de la région critique pour un risque α=5%.
5) Calculer la puissance du test.
6) Enoncer les règles de décision avec les probabilités d'erreur.
7) La moyenne calculée sur l'échantillon est $\bar{x} = 83$ décibels. Les habitants ont-ils raison de se plaindre ? Le test d'hypothèses ainsi établi leur est-il favorable ou défavorable ?
8) Combien faudrait-il faire de relevés journaliers, pour que le risque de deuxième espèce soit de 5% ?
9) Quelle serait alors le seuil de décision ?

Exercice 3

Un fabricant de conserves de petits pois produit des boîtes où l'étiquette annonce un poids net égoutté de 560gr. Il souhaite construire un test pour s'assurer, d'une part qu'il n'aura pas d'ennui à l'issu d'un contrôle éventuel, et d'autre part, que le poids moyen des boîtes n'est pas excédentaire. Pour ce faire, il compte prélever un lot de 25 boites et relever le poids moyen ainsi que l'écart-type.

1) Déterminer les hypothèses et expliciter les risques de 1ère et 2ème espèces.
2) Quelle est la variable de décision ? Préciser sa loi.
3) Déterminer graphiquement l'allure de la région critique.
4) Calculer les seuils de la région critique sachant que le risque de 1ère espèce est 10%.
5) Peut-on calculer la puissance du test ?
6) Il prélève un lot de 25 boites et il pèse un poids moyen de $\bar{x} = 556$gr avec un écart-type empirique $s^* = 10$gr.
   Quelle décision doit-il prendre ?

Exercice 4

Sur un échantillon de 900 naissances, on constate qu'il y a 470 garçons. Un généticien décide d'utiliser ces données pour effectuer le test suivant relatif aux proportions p et 1-p de naissances respectivement masculines et féminines :

$$H_0 : p=0.5$$
$$H_1 : p=0.55$$

1) Construire un test pour ces hypothèses avec un risque α=5%. Peut-on être satisfait du test ? Si non comment peut-on l'améliorer ?

2) Ce généticien effectue une nouvelle étude sur un échantillon de même taille. Il souhaite cette fois tester les hypothèses :

$$H_0 : p=0.5$$
$$H_1 : p\neq 0.5$$

<u>Exercice supplémentaire</u>

Un fabricant produit des piles dont la durée de vie suit une loi normale d'espérance annoncée à 80h et d'écart-type 6.44h. Suite à des réclamations, il veut vérifier si la qualité a baissé ou non. Dans le cas où la moyenne de la durée de vie est inférieure à 75h, il sera obligé de baisser le prix de vente.

1) Déterminer les hypothèses et expliciter les risques de $1^{\text{ère}}$ et $2^{\text{ème}}$ espèces.
2) Quelle est la variable de décision ? Préciser sa loi.
3) Déterminer graphiquement l'allure de la région critique.
4) Enoncer la règle de décision sachant que le fabricant souhaite maîtriser les deux erreurs avec $\alpha=0.025$ et $\beta=0.05$.

TESTS DE COMPARAISON D'ECHANTILLONS

<u>Exercice 5</u>

Une entreprise fabrique un médicament sur deux chaînes de production. On s'intéresse aux variations de la quantité d'une certaine substance A contenue dans chaque médicament. On a contrôlé le dosage de la substance A avec un échantillon de 100 médicaments à la sortie de chacune des deux chaînes de fabrication. On a trouvé un dosage moyen de 10.75mg pour la première chaîne et 10.70mg pour la deuxième. On sait par ailleurs que l'écart-type des chaînes de production est le même et est égal à 0.2mg.

Construire un test à 1% permettant de savoir si la différence des moyennes observées est due à des fluctuations de l'échantillonnage ou bien si la chaîne de fabrication n°1 produit des médicaments contenant davantage de substance A que la chaîne n°2.

<u>Exercice 6</u>

Un sondage effectué auprès de 2000 personnes indique que 19% d'entre elles connaissent la marque de lessive Omopaic. Après une campagne publicitaire, un sondage analogue auprès de 1000 personnes montre que 230 d'entre elles connaissent cette marque. Peut-on considérer que la campagne a été efficace ?

TESTS D'ADEQUATION ET D'INDEPENDANCE

Exercice 7

Pendant 200 minutes, on note toutes les minutes le nombre de voitures arrivant au poste de péage sur l'autoroute.

Nombre de voitures	0	1	2	3	4	5	6	7	8	9	10	11
Effectif observé	1	15	30	46	38	30	16	13	5	3	2	1

Déterminer la loi de probabilité de cet échantillon.

Exercice 8

On considère les données d'un essai visant à déterminer la solidité d'une corde d'escalade. Un morceau de 1 m corde est mis sous tension jusqu'à cassure. On se demande si la corde peut casser à n'importe endroit. On obtient les résultats suivants (en mètre):

0,1	0,4	0,4	0,6	0,7	0,7	0,8	0,9	0,9	0,9

1) Tracer la fonction de répartition empirique.
2) Ajouter sur le graphique la fonction de répartition de la loi U([0,1]).
3) La corde peut-elle casser à n'importe quel endroit ?
4) Déterminer la p-valeur du test.

Exercice 9

Un traitement est administré à trois doses différentes D_1, D_2, D_3, à un groupe de sujets atteints d'une même maladie. L'expérimentation est faite en double aveugle. On compte le nombre de guérisons pour chaque dose. Les résultats sont les suivants :

	Sujets guéris	Sujets non guéris	Total
D_1	30	30	60
D_2	42	35	77
D_3	58	31	89
Total	130	96	226

L'efficacité du traitement est-elle liée à la dose utilisée ?