

TD3 : Test d'hypothèses



Version web

Ex 1

Selon le choix de l'hypothèse

trhY

★ ★

Dans une grande entreprise, le salaire moyen annuel des hommes ayant entre 3 et 5 ans d'ancienneté est de 28 000 euros. Pour préparer une négociation sur la parité des salaires hommes femmes, on fait un sondage sur 10 femmes qui donne les résultats suivants (en milliers d'euros) :

24 27 31 19 26 27 22 15 33 21

A l'ouverture de la réunion, Monsieur A, délégué du personnel annonce : « Le sondage montre qu'il n'y a pas lieu de penser que le salaire des femmes est différent de celui des hommes ». Il est aussitôt interrompu par Madame B, Directrice des Ressources Humaines, qui dit « Pas du tout, le sondage prouve que le salaire des femmes est inférieur à celui des hommes ». Qu'en pensez-vous ?

Pour argumenter les réponses, on pourra se poser les questions suivantes pour Monsieur A et Madame B.

1. Que doit-on supposer sur la distribution des salaires des femmes pour pouvoir faire des tests statistiques ?
2. Si on note $\mu_0 = 28$ et $H_0: \mu = \mu_0$ l'hypothèse nulle de Monsieur A / Madame B, quelle est l'hypothèse alternative pour chacun ?
3. Déterminer la variable de décision et sa loi sous H_0 .
4. Déterminer la région critique pour un risque de première espèce $\alpha = 5\%$.
5. Calculer une réalisation de la variable de décision.
6. Peut-on dire que Monsieur A ou Madame B se trompe ?

Ex 2

Contrôle qualité

aUn2

★ ★

Désirant juger le travail d'un ouvrier ajusteur, un chef d'atelier prélève un échantillon de 50 pièces métalliques dans sa production. On note X l'épaisseur de ses pièces. L'objectif est d'avoir $\mathbb{E}(X) = 5 \text{ mm}$. Les résultats des mesures sur cet échantillon sont portés dans le tableau suivant :

n_i	5	15	20	10
x_i en mm	4.8	4.9	5.0	5.1

1. Cette vérification permet-elle de conclure que le résultat est conforme aux exigences, au seuil de confiance de 99% ?
2. Quel risque de première espèce devrait-on prendre pour que la prise de décision soit différente ?

Ex 3

Test du comportement des clients

r5Yy

Soit X le nombre d'incidents de paiements pour un crédit à la consommation observés sur la durée du prêt. On suppose que X suit une loi de Poisson de paramètre λ positif. On dispose d'un échantillon de n clients appartenant à la banque A.

1. Donner un estimateur de λ centré et convergent.
2. On désire tester l'hypothèse nulle selon laquelle les clients de la banque A sont faiblement risqués sous la forme suivante : $\begin{cases} \lambda = 1 \\ \lambda > 1 \end{cases}$.

Pour un échantillon de 100 clients de la banque A, nous observons les résultats suivants :

Incidents :	0	1	2	3	4	5	6	Total
Effectifs :	38	31	16	10	2	2	1	100

Que peut-on conclure pour un risque de première espèce de 5% ? Sur un même graphique, superposer une représentation graphique de la loi de Poisson de paramètre $\lambda = 1$ et l'histogramme des effectifs observés.

Ex 4

Contrôle qualité

3GmQ

Une entreprise fabrique des pièces en sous-traitance. Au sein d'une démarche qualité, toutes les machines ont été systématiquement révisées et on a défini une nouvelle organisation dans l'atelier : les tâches de contrôle sont réparties à chaque étape du processus de fabrication et le taux de pièces défectueuses est tombé à 1%.

Quelques mois plus tard, une opération de contrôle est effectuée pour vérifier si la norme de 1% (hypothèse H_0) de pièces défectueuses reste valable. Sur les 5 000 pièces contrôlées 100 s'avèrent défectueuses, soit 2% (hypothèse H_1).

Mme de Mainard, chef d'entreprise, décide que si l'hypothèse nulle est vérifiée, elle ne modifiera plus son processus de production (décision D_0) et au contraire, si c'est l'hypothèse alternative, elle entreprendra une action de sensibilisation des salariés de cet atelier au problème de la qualité (décision D_1).

Pour choisir entre ces deux hypothèses, elle tire un échantillon de 1 500 pièces.

1. Si la chef d'entreprise se fixe un risque de 1% d'entreprendre une action de sensibilisation des salariés à tort, quel sera le taux critique de pièces défectueuses qui fera prendre une décision ?
2. Si dans l'échantillon prélevé, le nombre de pièces défectueuses est 18, quelle sera la décision de la chef d'entreprise ?
3. Calculer alors le risque de l'acheteur, c'est-à-dire ne pas modifier le processus de production alors qu'on le devrait. Comment s'appelle ce risque ?

Ex 5

Contrôle qualité : test de proportion

6YqE

★ ★ ★

Une machine de production fabrique des pièces en grande série. Lorsque la machine est bien réglée, la proportion de pièces défectueuses est $p_0 = 0.05$. Lorsqu'elle est dérégulée, cette pro-

portion passe à $p_1 = 0.10$.

Pour vérifier l'état de la machine, on prélève un échantillon aléatoire de $n = 100$ pièces dans la production, et on observe $k = 8$ pièces défectueuses.

On note $\hat{p} = \frac{k}{n}$ la fréquence observée de pièces défectueuses et on souhaite tester :

$H_0 : p = p_0 = 0.05$ (la machine est bien réglée)

$H_1 : p = p_1 = 0.10$ (la machine est dérégulée)

1. Justifier que, sous H_0 , la statistique de test

$$T_n = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

suit approximativement une loi $\mathcal{N}(0, 1)$ pour n assez grand.

2. On note π la valeur critique : on rejette H_0 si $\hat{p} > \pi$. En utilisant la définition du risque de première espèce, exprimer α en fonction de π , p_0 et n .
3. Pour un seuil $\alpha = 0.05$, déterminer la valeur critique π et en déduire la région d'acceptation de l'hypothèse H_0 en termes de fréquence observée.
4. Au vu des données ($\hat{p} = 0.08$), peut-on conclure que la machine est bien réglée au seuil $\alpha = 5\%$? Calculer également la p -valeur associée à ce test.
5. Calculer le risque de seconde espèce β , c'est-à-dire la probabilité d'accepter H_0 alors que la machine est effectivement dérégulée.
6. On souhaite désormais dimensionner le plan de contrôle. Quelle taille d'échantillon n permettrait de rejeter H_0 avec une fréquence observée de $\hat{p} = 0.08$, tout en conservant un risque de première espèce $\alpha = 5\%$?
Interpréter ce résultat : pourquoi faut-il augmenter la taille de l'échantillon?
7. **[Bonus]** Représenter graphiquement sur un même schéma :
 - la densité de probabilité de $\hat{p}$ sous H_0
 - la densité de probabilité de $\hat{p}$ sous H_1
 - la région critique
 - les risques α et β

Ce graphique permet de visualiser la *puissance* du test.

Ex 6

Contrôle qualité

aW1E

Un pharmacien de banlieue désire être approvisionné par un grossiste chaque soir avant 18h00. Le grossiste doit donc estimer le temps de trajet moyen de son coursier entre le dépôt et la pharmacie. Pour cela, il relève le temps de parcours (en minutes) pendant 13 trajets sur un itinéraire fixé :

20 30 45 39 16 25 23 19 40 21 30 19 40

On supposera que ce temps est distribué selon une loi normale.

1. Donner une estimation ponctuelle de la moyenne du temps de parcours μ et de son écart-type s en rappelant les estimateurs utilisés et leurs propriétés.
2. Le pharmacien affirme qu'il faut en moyenne 25 minutes pour faire le trajet, alors que le grossiste, quant à lui, pense qu'il faut en moyenne 30 minutes.

- (a) On cherche à tester $H_0: \mu = 25$ contre $H_1: \mu = 30$. Préciser la variable de décision et sa loi sous H_0 .
- (b) Déterminer la région critique pour un seuil $\alpha = 5\%$.
- (c) Calculer le risque de seconde espèce, c'est-à-dire la probabilité de faire une erreur sachant que H_1 est vraie.
- (d) Avec un risque de première espèce $\alpha = 5\%$, peut-on accepter H_0 à partir de ces données ?

Ex 7

Expérience scientifique

7aqI

En 1908, en France, Binet publie une étude concernant la mesure de l'intelligence des enfants. Son échelle d'intelligence se base sur un certain nombre d'épreuves classées dans un ordre croissant de difficulté. On peut donc classer l'enfant dans une des 5 catégories suivantes : retardé de 2 ans, retardé d'1 an, régulier, avancé de 1 an et avancé de 2 ans. Voici les données, en fréquences relatives, que Binet obtient sur un échantillon de 192 enfants :

-2	-1	0	+1	+2
0,06	0,23	0,48	0,22	0,01

Des chercheurs, un américain (Goddard) et un allemand (Bobertag) font passer le test de Binet à un échantillon de 1547 enfants américains (Goddard) et 228 enfants allemands (Bobertag). Ils obtiennent les effectifs suivants :

	-2	-1	0	+1	+2
Goddard	294	309	557	325	62
Bobertag	6	40	119	57	6

Avec un niveau de confiance de 95%, peut-on dire que les données chez les enfants américains et allemands sont cohérentes avec la distribution obtenue par Binet chez les enfants français ?

Ex 8

Test d'adéquation à une loi

DkOf

★ ★

Une enquête a été réalisée auprès de 100 demandeurs d'emploi sur le nombre d'entretiens d'embauche obtenus ces deux derniers mois. Elle donne les résultats suivants :

Nombre d'entretiens :	0	1	2	3	4	5	6 et plus
Effectifs observés :	15	30	27	13	9	5	1

Peut-on dire que le nombre d'entretiens d'embauche est distribué selon une loi de Poisson ? On utilisera un seuil de significativité $\alpha = 5\%$.

Ex 9

Contrôle de qualité

XwAO

Pour simuler le tirage aléatoire uniforme avec un tableur, on peut utiliser la commande

`=ALEA.ENTRE.BORNES(1;8)`

qui génère un nombre entier aléatoire entre 1 et 8.

À l'aide d'un test du χ^2 et une simulation de 500 tirages de dé, peut-on affirmer que cette commande simule correctement un tirage uniforme? On pourra illustrer le résultat à l'aide d'un histogramme et calculer la p -valeur de ce test.

Ex 10

Test d'indépendance

5FLs

On veut étudier la liaison entre les caractères : «être fumeur» (plus de 20 cigarettes par jour, pendant 10 ans) et «avoir un cancer de la gorge», sur une population de 1000 personnes, dont 500 sont atteintes d'un cancer de la gorge. Voici les résultats observés :

<i>Observé</i>	cancer	non cancer	TOTAL
fumeur	342	258	600
non fumeur	158	242	400
TOTAL	500	500	1000

Faire un test d'indépendance pour établir la liaison entre ces caractères.

Ex 11

Comparaison d'homogénéité des évaluations

qAu0

Deux départements Mesures Physiques d'IUT utilisent le même référentiel de compétences pour évaluer leurs étudiants en statistiques. On souhaite comparer l'homogénéité de leur système de notation.

On suppose que les notes dans chaque IUT suivent une loi normale. Pour effectuer cette comparaison, on prélève :

- un échantillon de $n_1 = 10$ notes d'étudiants de l'IUT 1, qui donne un écart-type empirique $s_1 = 5$ points
- un échantillon de $n_2 = 25$ notes d'étudiants de l'IUT 2, qui donne un écart-type empirique $s_2 = 3$ points

Contexte : Une différence significative d'écart-type pourrait indiquer :

- des critères d'évaluation plus ou moins stricts
- une hétérogénéité différente dans les groupes d'étudiants
- des méthodes pédagogiques ayant des impacts différents

1. Avant de réaliser le test, calculer le coefficient de variation pour chaque IUT si on vous donne les moyennes suivantes : $\bar{x}_1 = 12.5$ et $\bar{x}_2 = 13.2$.

Rappel : le coefficient de variation est $CV = \frac{s}{\bar{x}} \times 100\%$. Que peut-on en conclure intuitivement ?

2. On souhaite tester si les deux IUT ont la même dispersion de notes.
 - Formuler les hypothèses H_0 et H_1 de ce test
 - Quelle statistique de test doit-on utiliser ? Préciser sa loi sous H_0
 - Ce test nécessite-t-il des conditions particulières ?

3. Au seuil de signification $\alpha = 5\%$, peut-on considérer que les enseignants des deux IUT ont la même manière de noter les étudiants ?
4. Construire un intervalle de confiance à 95% pour le rapport $\frac{\sigma_1^2}{\sigma_2^2}$. Que peut-on en déduire concernant la comparaison des variances ?
5. Pour aller plus loin :
 - Que se passerait-il si on augmentait la taille des échantillons (par exemple $n_1 = 30$ et $n_2 = 50$) tout en conservant les mêmes écarts-types ?
 - Quelle serait la taille d'échantillon minimale nécessaire pour l'IUT 1 (en gardant $n_2 = 25$ et les mêmes écarts-types) pour pouvoir rejeter H_0 au seuil de 5% ?
 - Dans le contexte pédagogique, quelles pourraient être les causes d'une différence d'écart-type si elle était avérée ?

Ex 12

Comparaison de consommations énergétiques

rhUy

Un fournisseur d'énergie souhaite comparer la consommation de chauffage durant les mois d'hiver (décembre à février) entre deux régions géographiques A et B afin d'optimiser ses approvisionnements et son réseau de distribution.

On suppose que la consommation mensuelle de chauffage (en kWh) dans chaque région suit une loi normale. Pour effectuer cette comparaison, on prélève un échantillon aléatoire de ménages dans chaque région et on mesure leur consommation moyenne sur la période.

Résultats observés :

Région A : $n_1 = 200$ ménages, moyenne $\bar{x}_1 = 600$ kWh, variance empirique $s_1^2 = 20\,000$ kWh² (donc $s_1 \approx 141.4$ kWh)

Région B : $n_2 = 100$ ménages, moyenne $\bar{x}_2 = 500$ kWh, variance empirique $s_2^2 = 160\,000$ kWh² (donc $s_2 = 400$ kWh)

Question générale : L'écart de consommation entre les deux régions est-il statistiquement significatif au seuil de 5% ?

1. Avant tout test, analyser les données de manière descriptive :
 - Calculer la différence observée entre les moyennes
 - Calculer les coefficients de variation pour chaque région
 - Que peut-on dire intuitivement sur la dispersion dans chaque région ?
2. Avant de comparer les moyennes, doit-on vérifier l'égalité des variances ? Réaliser un test de Fisher au seuil de 5% pour comparer σ_1^2 et σ_2^2 .
3. Formuler les hypothèses pour tester si la consommation moyenne est différente entre les deux régions. S'agit-il d'un test bilatéral ou unilatéral ? Justifier.
4. Les variances étant inégales, on utilise le test de Welch. La statistique de test est :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

qui suit approximativement une loi de Student à ν degrés de liberté, où :

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

Calculer la statistique de test observée et le nombre de degrés de liberté.

5. Au seuil de signification $\alpha = 5\%$, l'écart de consommation entre les deux régions est-il significatif? Calculer également la p -valeur et interpréter le résultat dans le contexte.
6. Construire un intervalle de confiance à 95% pour la différence des moyennes $\mu_1 - \mu_2$. Que peut-on en déduire?
7. Synthèse :
 une analyste propose d'utiliser le test de Student classique (avec hypothèse de variances égales) au lieu du test de Welch.
 - Expliquer pourquoi ce serait inapproprié dans ce cas
 - Calculer néanmoins la statistique de test qu'on obtiendrait avec le test de Student classique (variance poolée)
 - Comparer les deux approches : quelle différence observe-t-on dans la conclusion?
 - Quel enseignement peut-on tirer sur l'importance de vérifier les hypothèses d'un test?

Ex 13**Test d'indépendance**

DzLc

Dans le cadre d'une enquête sur le SIDA réalisée en Allemagne durant l'été 1990 (A. Hahn, W.H. Eirmbter et R. Jacob), on a interrogé 2089 personnes. Le questionnaire comportait notamment l'item suivant : Le sida est un péril omniprésent. Indiquez si vous êtes : d'accord, indécis, pas d'accord. Le croisement de la réponse du sujet avec son âge donne le tableau de contingence suivant :

Classe d'âge	d'accord	indécis	pas d'accord	Total
18 à < 30	43	116	365	524
30 à < 40	36	116	273	425
40 à < 50	32	95	217	344
50 à < 60	38	114	167	319
60 et plus	67	160	250	477
Total	216	601	1272	2089

1. Réaliser un test du χ^2 permettant de répondre à la question suivante : « Les réponses des sujets dépendent-elles de leur âge ? »
2. Comparer de même à l'aide d'un test du χ^2 les réponses des deux dernières classes d'âge, puis les réponses des moins de 30 ans à celles des 60 ans et plus.

Ex 14**Analyse de la précision d'un système d'arme**

QeZr

★★★★

Dans le cadre de la qualification d'une section d'infanterie, on évalue la précision du tir au FAMAS sur cible à 200 mètres. Le référentiel de l'armée de terre fixe un écart-type maximal de $\sigma_0 = 8$ cm pour les impacts par rapport au centre de la cible (tous tireurs confondus) pour qu'une section soit considérée comme opérationnelle.

Un chef de section souhaite vérifier si sa section respecte ce standard. Lors d'une séance de tir, 25 soldats effectuent chacun un tir, et on mesure la distance de l'impact au centre de la cible (en cm).

Distances observées (en cm) :

5.2	12.3	8.7	6.5	15.1	9.2	7.4	11.8	4.9	13.6	9.8
4.1	11.2	7.8	13.5	6.9	10.5	8.3	14.2	5.7	7.2	10.1
8.9	12.5	6.3								

Statistiques calculées (pour vérification) : $\bar{x} = 9.27$ cm, $s^2 = 9.83$ cm², $s = 3.13$ cm

On suppose que la distance des impacts suit une loi normale.

1. Avant de réaliser les tests statistiques, visualiser les données :
 - Représenter les distances par un histogramme ou un diagramme en boîte
 - Cette représentation est-elle cohérente avec l'hypothèse de normalité ?
 - Y a-t-il des valeurs aberrantes qui pourraient indiquer un problème (tireur mal formé, incident de tir, etc.) ?
2. Le chef de section souhaite tester si sa section respecte le standard de l'armée.
 - Formuler les hypothèses H_0 et H_1 de ce test
 - Quel type de test doit-on utiliser ? (test bilatéral ou unilatéral ?)
 - Justifier le choix de la statistique de test et donner sa loi sous H_0
3. Au seuil de signification $\alpha = 5\%$, peut-on conclure que la section respecte le standard de précision de l'armée de terre ?
4. Calculer un intervalle de confiance à 95% pour l'écart-type σ de la population. Interpréter ce résultat dans le contexte militaire.
5. Le commandant de compagnie examine les résultats et fait la remarque suivante :

« Votre section montre d'excellents résultats avec un écart-type de 3.1 cm, bien meilleur que le standard de 8 cm. Cependant, je constate que la moyenne des impacts est à 9.3 cm du centre. Ne faudrait-il pas aussi vérifier que le tir n'est pas systématiquement décentré ? »

 - Le commandant a-t-il raison de s'interroger sur ce point ?
 - Proposer et réaliser un test statistique pour vérifier si le tir de la section présente un biais systématique (hypothèse : le tir devrait être centré, donc $\mu = 0$)
 - Que recommanderiez-vous au chef de section suite à cette analyse complète ?
6. Un officier du bureau emploi propose de comparer cette section avec une autre section de la même compagnie qui a obtenu les résultats suivants (20 tireurs) : $\bar{y} = 8.1$ cm, $s_y^2 = 15.3$ cm².
 Les deux sections ont-elles le même niveau de précision ? Réaliser un test de comparaison de variances au seuil de 5%.