Tests d'hypothèses

Benchikh Tawfik

Faculté de Médecine, UDL, SBA
1ère année Médecine

06 Mars 2024

PLAN DE COURS

INTRODUCTION

- La théorie des tests consiste à formuler des hypothèses particulières sur les paramètres ou sur les lois qui interviennent dans les problèmes étudiés, puis à apporter un jugement sur ces hypothèses.
- Ce jugement est basé d'une part, sur les résultats obtenus sur un ou plusieurs échantillons extraits de la population concernée et d'autre part, sur l'acceptation d'un certain risque dans la prise de décision.

- Les tests peuvent être classés en différentes catégories:
  - tests sur une hypothèse relative à la valeur particulière d'un ou plusieurs paramètre(s) ou tests paramétriques,
  - tests de conformité de deux distributions ou tests d'ajustement entre une distribution théorique et une distribution expérimentale,
  - tests de comparaison de deux populations (comparaison des moyennes, des variances...),
  - tests d'indépendance de deux caractères quantitatifs ou qualitatifs.

EXEMPLE

- Considérons des lampes à incandescence dont la durée de vie est une variable aléatoire gaussienne de moyenne $m = 1000$ heures et d'écart-type $\sigma = 100$ heures (résultats établis expérimentalement). Un ingénieur propose un nouveau procédé de fabrication qui doit améliorer, affirme-t-il, cette durée de vie moyenne et la rendre égale à 1075 heures. Cependant, le procédé, étant plus onéreux, ne pourra être accepté que si l'amélioration est réelle.

Exemple

- Le résultat qui doit être vérifié est donc $m > 1000$ heures. Deux hypothèses sont en présence:
  - soit $m = 1000$ heures est une hypothèse encore vraie et le nouveau procédé n'a pas modifié de façon significative la durée de vie des lampes,
  - soit $m = 1075$ heures et le nouveau procédé a apporté une réelle amélioration.
- La deuxième affirmation est aussi une hypothèse car aucune expérience ne permet d'affirmer que la moyenne des durées de vie augmentera réellement ($m > 1000$ heures). En effet, il est possible que le nouveau procédé de fabrication n'améliore pas la durée de vie des ampoules de façon appréciable, il pourrait même la diminuer.
- En conclusion, rejeter l'hypothèse $m = 1000$ heures ne conduit pas nécessairement à accepter l'hypothèse $m > 1000$ heures.

EXEMPLE

- Les hypothèses ayant été formulées, il faut les accepter ou les rejeter à l'aide des données apportées par un échantillon. Une incertitude est toujours associée au jugement apporté par le statisticien, mais ce dernier doit essayer de limiter ce risque.

DÉMARCHE

1. Choisir les hypothèses à tester (H0 et H1)
2. Fixer le niveau du test α
3. Choisir une statistique de test
4. Déterminer la règle de décision (région de rejet ..)
5. Calculer la statistique (et la p-valeur)
6. Conclure

Risques d'erreurs

- Résultats possibles:

Décision Réalité	Ne pas rejeter H0 (conclure H0)	Rejeter H0 (conclure H1)
H0 vrais	ok $(1-\alpha)$	Erreur de type I (α)
H1 vrais	Erreur de type II (β)	ok $(1-\beta)$

DÉFINITIONS

- Risque de première espèce: $\alpha = \Pr$(rejeter H0 | H0 vraie) (probabilité de commettre une erreur de type I)
- Risque de seconde espece: $\beta = \Pr$(ne pas rejeter H0 | H1 vraie) (probabilité de commettre une erreur de type II)
- Puissance: $P = 1 - \beta = \Pr$(rejeter H0 | H1 vraie) (probabilité de prendre la bonne décision en rejetant H0)

Hypothèse nulle et hypothèse alternative

- **L'hypothèse nulle** (notée H0) est l'hypothèse privilégiée. C'est celle qui est supposée vraie par défaut (vérité établie) et qui sera conservée en cas de doutes (trop importants).
- L'hypothèse alternative (notée H1) contredit l'hypothèse nulle. C'est l'hypothèse que l'on cherche à montrer.

HYPOTHÈSE SIMPLE / COMPOSITE

- **Exemple d'hypothèses simples:** $\theta = \theta_0$
- Exemple d'hypothèses composites:
  - $\theta < \theta_0$
  - $\theta > \theta_0$
  - $\theta \neq \theta_0$
  - $\theta \in [a, b]$

TESTS UNILATÉRAUX / BILATÉRAUX

- **Test unilatéral (à droite):**
  - $H0$: $\theta = \theta_0$
  - $H1$: $\theta > \theta_0$
- **Test unilatéral (à gauche):**
  - $H0$: $\theta = \theta_0$
  - $H1$: $\theta < \theta_0$
- **Test Test bilatéral:**
  - $H0$: $\theta = \theta_0$
  - $H1$: $\theta \neq \theta_0$

STATISTIQUE DE TEST

- **Définition:**
  - Une **statistique de test** est une statistique (dont la loi est connue sous H0) qui permet de mesurer l'écart à l'hypothèse nulle.

Choix du niveau du test:

- Le niveau de signification du test est le risque de première espèce α consenti.
- Le niveau de signification du test est souvent fixé à 0.05 ou 0.01, mais ce seuil est arbitraire est toute autre valeur peut être choisie.

Région de rejet

- La Région de rejet est l'ensemble R des valeurs (de la statistique de test) pour lesquelles l'hypothèse nulle est rejetée.
- Démarche de Neyman Pearson: Maximiser la puissance tout en contrôlant α.

Degré de signification (p-valeur)

- Le degré de signification (ou p-valeur) est défini par:

$$p = min\{\alpha | T \in \Gamma_\alpha\},$$

  où Γ_α est la région critique.
  - Test unilatéral à droite: $p = \Pr(T > t | H0)$,
  - Test unilatéral à gauche: $p = \Pr(T < t | H0)$,
  - Test bilatéral à droite: $p = \Pr(|T| > |t| | H0)$,
- **Remarques:** La p-valeur est la probabilité d'obtenir une valeur de la statistique de test au moins aussi extrême que celle observée lorsque H0 est vraie
- En pratique, on rejette H0 lorsque $p < \alpha$.

Pourcentage:

- Variable qualitative dichotomique (Présence/Absence, Malades/Non malades, Décès/Survie, ...),
- p est le pourcentage (inconnu) d'individus présentant la caractéristique dans la population,
- p est estimé par le pourcentage p_0 observé sur un échantillon de taille n dont k individus présentent la caractéristique

$$p = \frac{k}{n}$$

Comparaison de deux pourcentages

- Comparaison d'un pourcentage observé à un pourcentage théorique (Test de conformité)
- Comparaison de deux pourcentages observés (Test d'homogénéité).
  - Échantillons indépendants,
  - Échantillons appariés.

COMPARAISON D'UN POURCENTAGE OBSERVÉ À UN POURCENTAGE THÉORIQUE: TEST DE CONFORMITÉ

- **Problème:** déterminer si un pourcentage observé p_0 sur un échantillon de taille n est différent d'une valeur théorique p_{th}.
  - Comparer p à p_{th}.

Formuler une hypothèse

- Hypothèse nulle $H0$: $p = p_{th}$ où p est le pourcentage de la population dont est issu l'échantillon,
- Hypothèses alternatives $H1$:
  - Test bilatéral: $p \neq p_{th}$
  - Test unilatéral à gauche ou à droite: $p < p_{th}$ ou $p > p_{th}$.

DÉMARCHE

- Fixer le risque α,
- Choisir la statistique:
  - Test z (loi normale).
  - Test du χ^2 de conformité (loi du χ^2): Test d'indépendance.
- Conditions d'application:
  - $n * p_{th} \geq 5$ et $n * (1 - p_{th}) \geq 5$ (cas des grands échantillons).

TEST Z

- Calculer la valeur z prise par la statistique Z:
  - $z = \frac{p - p_{th}}{\sqrt{\frac{p_{th}(1-p_{th})}{n}}}$ (on a: $\chi^2_{cal} = z^2$).
  - Sous $H0$, Z suit une loi normale centrée réduite (χ^2 à 1 d.d.l est le carré d'une loi normale centrée réduite).
  - Conditions d'application: $n * p_{th} \geq 5$ et $n * (1 - p_{th}) \geq 5$.

DÉMARCHE

- Confronter z à la valeur critique z_α:
  - Test bilatéral: on rejette $H0$ si $|z| \geq z_\alpha$,
  - Test unilatéral:
    - si $H1$ s'écrit $p > p_{th}$, on rejette $H0$ si $z \geq z_{2\alpha}$;
    - si $H1$ s'écrit $p < p_{th}$, on rejette $H0$ si $z \leq -z_{2\alpha}$;

Test du χ^2 de conformité: exemple 1

- En Algérie, 7% des personnes hospitalisées contractent une infection nosocomiale dans l'établissement où elles sont soignées.
- Sur un échantillon de 250 personnes soignées à l'hôpital de Sidi-Bel-Abbès, 28 ont contracté une infection nosocomiale.
- Le pourcentage observé sur l'échantillon diffère-t-il de la référence nationale au risque $\alpha = 5\%$.

Test z : exemple 1

- Données: $n = 250$, $p_0 = \frac{28}{250} = 0.112$, $p_{th} = 0.07$.
- Hypothèse: $H0$: $p = 0.07$; $H1$: $p \neq 0.07$.
- Calcul:

$$Z = z = \frac{p - p_{th}}{\sqrt{\frac{p_{th}(1-p_{th})}{n}}} = z = \frac{0.112 - 0.07}{\sqrt{\frac{0.07(0.93)}{250}}} = 2.60\,.$$

- Conditions d'application vérifiées: $250 \times 0,07 \geq 5$ et $250 \times 0,93 \geq 5$.
- Lecture: table de l'écart-réduit (loi normale), $Z_\alpha = 1.96$.
- $z = 2,60 \geq z_{0.05} = 1,96$: rejet de $H0$ (même conclusion que test précédent).
- Degré de signification lu dans la table: $p < 0,01$

FORMULER UNE HYPOTHÈSE

- On étudie deux variables quantitatives X et Y définies sur une population P.
- On veut tester l'existence d'une liaison entre les deux variables.
  - $H0$: $p = p_{th} \Leftrightarrow$ les variables sont indépendantes
  - $H1$: $p \neq p_{th}$ (pour test bilatéral) et $p < p_{th}$ ou $p > p_{th}$ pour test unilatéral $\Leftrightarrow$ les variables sont liées (positivement, négativement):

TEST DU χ^2 DE CONFORMITÉ: TEST D'INDÉPENDANCE

- Calculer la valeur χ^2 prise par la statistique du test:
- Tableau:

Effectifs observés	$O_1 = n * p_0$	$O_2 = n * (1 - p_0)$	n
Effectifs calculés	$C_1 = n * p_{th}$	$C_2 = n * (1 - p_{th})$	n

- Conditions d'application: $C_1 \geq 5$ et $C_2 \geq 5$.
- $\chi^2_{cal} = \frac{(O_1 - C_1)^2}{C_1} + \frac{(O_2 - C_2)^2}{C_2}$.
- Sous $H0$ la statistique suit une loi du χ^2 à 1 d.d.l

Test du χ^2 de conformité

- Confronter χ^2_{cal} à la valeur seuil $\chi^2_{(1,\alpha)}$:
  - Lecture de la valeur seuil dans la table de la loi du χ^2,
  - Test bilatéral: on rejette $H0$ au risque α si $\chi^2_{cal} \geq \chi^2_{(1,\alpha)}$:
  - En pratique, si $\alpha = 5\%$, $\chi^2_{1,\alpha} = 3.84$:
    - Si $\chi^2_{cal} \geq 3.84$: rejet de $H0$.
    - si $\chi^2_{cal} < 3.84$: on accepte $H0$ (non rejet).
  - Remarque: Pour un même risque d'erreur, les valeurs seuil du χ^2_{cal} sont donc les carrés des valeurs seuil de z: $\chi^2_{cal} = (z_\alpha)^2$ $(3,84 = 1,96^2)$.

Test du χ^2 de conformité: exemple

- la variable X est la maladie (infecter ou pas), la variable Y est l'établissemant.
- Données: $n = 250$, $p_0 = \frac{28}{250} = 0.112$, $p_{th} = 0.07$.
- Hypothèse: $H0$: $p = 0.07$; $H1$: $p \neq 0.07$.
- Calcul

infection	(oui)	(non)	Total
Effectifs observés	$O_1 = 28$	$O_2 = 222$	250
Effectifs calculés	$C_1 = 17.5$	$C_2 = 232.5$	250

- Conditions d'application vérifiées: $C_1 \geq 5$ et $C_2 \geq 5$.
- $\chi^2_{cal} = \frac{(28-17.5)^2}{17.5} + \frac{(222-232.5)^2}{232.5} = 6.77$.

TEST DU χ^2 DE CONFORMITÉ: EXEMPLE

- Lecture: table de χ^2_{th}: $\chi^2_{1,\alpha} = \chi^2_{1,\,5\%} = 3.84$.
- $\chi^2_{cal} = 6.77 \geq \chi^2_{1,\alpha} = 3.84$: rejet de $H0$.
- On montre, au risque 5%, une différence significative entre le pourcentage de personnes hospitalisées contractant une infection nosocomiale à l'hôpital H et dans l'ensemble du pays $(p < 0,01)$.

COMPARAISON DE DEUX POURCENTAGES OBSERVÉ:

- **Problème:** comparer 2 proportions (p_1 et p_2) dans 2 groupes indépendants de tailles n_1 et n_2.
  - Comparer π_1 à π_2, où π_i est le pourcentage de la population dont est issu l'échantillon "i",

Comparaison de deux pourcentages observé dans le cas des grands échantillons:

- Formuler une hypothèse:
  - Hypothèse nulle $H0$: Les 2 échantillons sont issus de la même population ayant comme pourcentage π_0
    $\pi_1 = \pi_2 (= \pi_0)$ où π_1 et π_2 pourcentages de la population dont sont issus les échantillons 1 et 2.
- Hypothèses alternatives $H1$:
  - Test bilatéral: $\pi_1 \neq \pi_2$
  - Test unilatéral: $\pi_1 < \pi_2$ ou $\pi_1 > \pi_2$.

Comparaison de deux pourcentages observé:

- Fixer le risque α.
- Choisir la statistique:
  - Test du χ^2 d'homogénéité (loi du χ^2)
  - Test z (loi normale).
- Conditions d'application:
  - $n_1 * \pi_0 \geq 5$ et $n_1 * (1 - \pi_0) \geq 5$.
  - $n_2 * \pi_0 \geq 5$ et $n_2 * (1 - \pi_0) \geq 5$.

COMPARAISON DE DEUX POURCENTAGES OBSERVÉ: TEST DU χ^2

- Calculer la valeur χ^2 prise par la statistique du test:
- Tableau de contingence (tableau à 4 cases)

	Succès	Échec	Total
Groupe 1	$O_{11}(C_{11})$	$O_{12}(C_{12})$	$n_{1.}$
Groupe 2	$O_{21}(C_{21})$	$O_{22}(C_{22})$	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	N

Effectifs calculés sous $H0$: $C_{ij} = \frac{n_{i.} * n_{.j}}{N}$.

- Conditions d'application: $C_{ij} \geq 5$.
- $\chi^2_{cal} = \sum_{i,j} \frac{(O_{ij} - C_{ij})^2}{C_{ij}}$.
- Sous $H0$ la statistique suit une loi du χ^2 à 1 d.d.l.

Comparaison de deux pourcentages observé: Test du χ^2, remarque:

- Dans le cas des tableaux de contingence à 4 cases, il est possible d'utiliser la correction de continuité, surtout lorsque les valeurs attendues sont faibles (en pratique $C_{ij} < 5$)

$$\chi^2_{cal} = \sum_{i,j} \frac{(|O_{ij} - C_{ij}| - 0.5)^2}{C_{ij}}$$

- Petits échantillons: test exact de Fisher (hors programme).

COMPARAISON DE DEUX POURCENTAGES OBSERVÉ: TEST DU χ^2

- Confronter χ^2_{cal} à la valeur seuil $\chi^2_{(1,\alpha)}$:
  - Lecture de la valeur seuil dans la table de la loi du χ^2,
  - Test bilatéral: on rejette $H0$ au risque α si $\chi^2_{cal} \geq \chi^2_{1,\alpha}$:
  - En pratique, si $\alpha = 5\%$, $\chi^2_{1,\alpha} = 3.84$:
    - Si $\chi^2_{cal} \geq 3.84$: rejet de $H0$
    - si $\chi^2_{cal} < 3.84$: on accepte $H0$ (non rejet).

COMPARAISON DE DEUX POURCENTAGES OBSERVÉ: TEST DU χ^2, EXEMPLE

- On désire comparer l'efficacité de deux traitements $T1$ et $T2$ sur 100 patients atteints d'une maladie M.
- On tire au sort 2 deux groupes de 50 patients, un groupe est soumis à $T1$, le second à $T2$.
- Le pourcentage de guérison chez les patients soumis à $T1$ est de 30%, chez ceux soumis à $T2$ de 40%.
- Le taux de guérison est-il différent entre les 2 traitements ?

Comparaison de deux pourcentages observé: Test du χ^2, exemple

- Données: $n = 20$, $p_1 = 0.3$, $p_2 = 0.4$.
- Hypothèse: $H0$: $\pi_1 = \pi_2$; $H1$: $\pi_1 \neq \pi_2$.
- Calcul: Tableau de contingence (tableau à 4 cases)

	Succès	Échec	Total
Groupe T1	15 (17.5)	35 (32.5)	50
Groupe T2	20 (17.5)	30 (32.5)	50
Total	35	65	100

- Conditions d'application: $C_{ij} \geq 5$.
- $\chi^2_{cal} = \sum_{i,j} \frac{(O_{ij} - C_{ij})^2}{C_{ij}} = 1.10$.
- Lecture $\chi^2_{1,\alpha} = 3.84$: $H0$ acceptable.
- On ne met pas en évidence, au risque 5%, de différence significative entre les taux de guérison avec les 2 traitements.

COMPARAISON DE DEUX POURCENTAGES OBSERVÉ: TEST Z

- Calculer la valeur z prise par la statistique Z:
  - $z = \frac{p_1 - p_2}{\sqrt{\frac{p_0(1-p_0)}{n_1} + \frac{p_0(1-p_0)}{n_2}}}$, avec $p_0 = \frac{n_1 * p_1 + n_2 * p_2}{n_1 + n_2}$ ($\chi^2_{cal} = z^2$).
  - p_0 est l'estimation de la proportion commune π_0.
  - Z suit une loi normale centrée réduite (χ^2 à 1 d.d.l est le carré d'une loi normale centrée réduite)
  - Conditions d'application:
    - $n_1 * \pi_0 \geq 5$ et $n_1 * (1 - p_0) \geq 5$,
    - $n_2 * \pi_0 \geq 5$ et $n_2 * (1 - p_0) \geq 5$

Comparaison de deux pourcentages observé: Test Z d'homogénéité

- Confronter z à la valeur critique z_α:
  - Test bilatéral: on rejette $H0$ si $|z| \geq z_{alpha}$,
  - Test unilatéral:
    - si $H1$ s'écrit $\pi_1 > \pi_2$, on rejette $H0$ si $z \geq z_{2\alpha}$;
    - si $H1$ s'écrit $\pi_1 < \pi_2$, on rejette $H0$ si $z \leq -z_{2\alpha}$;
- Remarque: Pour un même risque d'erreur, les valeurs seuil du χ^2_{cal} sont donc les carrés des valeurs seuil de z: $\chi^2_{cal} = (z_\alpha)^2 \ (3,84 = 1,96^2)$.

Test z : exemple 2

- Données: $n = 20$, $p_1 = 0.3$, $p_2 = 0.4$.
- Hypothèse: $H0$: $\pi_1 = \pi_2$; $H1$: $\pi_1 \neq \pi_2$.
- Calcul: $p_0 = \frac{50*0.3+50*0.5}{100} = 0.35$ et

$$z = \frac{0.4 - 0.3}{\sqrt{\frac{0.35*0.65}{50} + \frac{0.35*0.65}{50}}}.$$

- Conditions d'application vérifiées: $50 * 0,35 \geq 5$ et $50 * 0,65 \geq 5$.
- Lecture $Z_\alpha = 1.96$.
- $z = 1,05 < z_{0.05} = 1,96$: $H0$ acceptable (Même conclusion que le test précédent).

Comparaison de deux proportions à l'aide d'intervalles de confiance

- La différence entre les deux fréquences observées est considérée comme significative lorsque les intervalles de confiance sont disjoints (avec le rique α). Donc avec la probabilté $1-\alpha$ les valeurs de π_1 et π_2 sont différentes.
- Comparer deux intervalles de même niveau (non disjoint) revient à faire un test bilatéral.

TESTS DE COMPARAISON DE MOYENNES: TEST DE Z OU DE L'ÉCART RÉDUIT

- Le test de Z: comparer des paramètres en testant leurs différences
- Utilisé pour comparer:
  - Une moyenne observée à une moyenne théorique,
  - Deux moyennes,
  - Deux moyennes de deux séries appariées.

Principe du test de Z

- 2 paramètres de 2 échantillons que l'on désire comparer:
  - $H0$: Les paramètres des populations d'où sont issus les 2 échantillons sont identiques.
  - $H1$ bilatérale: Les paramètres sont différents,
  - $H1$ unilatérale: un des paramètres est inférieur (ou supérieur) à l'autre.

Principe du test de Z

- On compare les 2 paramètres par leur différence Δ.
- Δ est une variable aléatoire,
- Si $H0$ est vrai alors Δ est proche de 0:
  - Si $H0$ est vrai et que l'échantillon est de taille suffisance.
  - La division de Δ par son écart type suit une loi Z normale centrée réduite de moyenne 0 et d'écart type 1.

Principe du test de Z

- Le test de Z consiste:
  - à estimer l'écart type de la différence s_d,
  - à calculer l'écart réduit $z_0 = \frac{\Delta}{s_d}$,
  - à comparer cette valeur à la distribution théorique de la loi Z.
- On utilise la table Z,
- Condition d'application: effectif de chaque échantillon ≥ 30

INTERPRÉTATION DU TEST DE Z

- α de 5%, $H1$ bilaterale:
- Si la valeur observée $z_o < 1.96$, on ne rejette pas $H0$
  - On ne peut pas affirmer que les échantillons proviennent de populations différentes.
  - la différence entre les paramètres n'est pas significative.
- Si la valeur observée $z_o > 1.96$, on rejette $H0$
  - On accepte $H1$ en affirmant que les échantillons proviennent de populations différentes.
  - On affirme que la différence entre les paramètres est significative.
  - On recherche le degré de signification.

TEST DE CONFORMITÉ: CAS σ INCONNUE

- Considérons un caractère quantitatif représenté par une variable aléatoire X d'espérance mathématique μ, d'écart-type σ, et un échantillon $\underline{X} = (X_1, \ldots, X_n)$ de taille n de cette variable.
- On compare une moyenne observée dans un échantillon à une moyenne connue dans la population de référence.
- Paramètre étudié: moyenne avec σ inconnu.
- Taille de l'échantillon ≥ 30.
- Hypothèses:
  - $H0$: $M = \mu$,
  - $H1$ bilaterale: $M \neq \mu$,
  - $H1$ unilaterale: $M > \mu$ ou $M < \mu$.

TEST DE CONFORMITÉ

- Formulation:
  - μ: moyenne théorique connue de la population de référence,
  - M: moyenne inconnue de la population d'où est issu l'échantillon,
  - $\overline{X} = m$: moyenne observée de l'échantillon,
  - s: écart type de l'échantillon,
  - n: effectif.
- Conditions d'applications: Taille ≥ 30 (sinon: test de T de student).
- On sait que $Z = \frac{\overline{X}-M}{s'/\sqrt{n}}$ suit approximativement la loi normale $\mathcal{N}(0,1)$.

PRINCIPE DU TEST:

- Si $H0$ est vrai:
  - m est l'une des valeurs possibles d'une variable normale centrée autour de M,
  - La difference Δ entre cette variable et μ suit une loi normale de moyenne 0,
  - Le rapport de Δ sur l'écart type de μ suit une loi Z normale centrée réduite,
- L'écart type de μ estimé par l'écart type de la moyenne de l'échantillon $\frac{s}{\sqrt{n-1}} = \frac{s'}{\sqrt{n}}$
- Test de Z:

$$z = \frac{|m - \mu|}{\frac{s'}{\sqrt{n}}} .$$

PRINCIPE DU TEST: REMARQUE

- Si σ est connu, alors:

$$z = \frac{|m - \mu|}{\frac{\sigma}{\sqrt{n}}}.$$

DÉCISION

- Si $H1$ bilatérale:
  - si $z < z_{\alpha/2} = z_{0.025} = 1.96$, on accepte $H0$ (interprétation: M n'est pas significativement different de μ).
  - si $z \geq 1.96$, on rejette $H0$ (interprétation: M diffère significativement de μ).
- Si $H1$ unilatérale:
  - si $z < z_{\alpha} = z_{0.05} = 1.645$, on accepte $H0$ (interprétation: M n'est pas significativement supérieur (ou inférieur) à μ).
  - si $z \geq z_{\alpha} = z_{0.05} = 1.645$, on rejette $H0$ (interprétation: M est significativement supérieur (ou inférieur) à de μ).

Test Z de conformité: exemple

- Dans une usine du secteur de l'agroalimentaire, une machine à embouteiller est alimentée par un réservoir d'eau et par une file d'approvisionnement en bouteilles vides. Pour contrôler le bon fonctionnement de la machine, on veut construire un test d'hypothèse bilatéral qui sera mis en oeuvre toutes les heures.
- Pour une production d'une heure, on suppose que la variable aléatoire X qui à toute bouteille, prise au hasard dans cette production, associe le volume d'eau (en litres) qu'elle contient, est une variable aléatoire d'espérance μ et d'écart-type σ inconnus.
- On considère que la machine est bien réglée lorsque le volume d'eau moyen dans une bouteille est $\mu = 1.5l$.
- On a prélevé un échantillon de 100 bouteilles, et on a obtenu un volume d'eau moyen de 1.495 l et un écart-type corrigé de $s' = 0,01$. Peut-on conclure, au risque 5%, que la machine est bien réglée ?

Test Z de conformité: exemple

- Population: bouteilles produites.
- Variable X: volume d'eau, variable aléatoire de moyenne μ et d'écart-type σ inconnu.
- Échantillon $\underline{X} = (X_1, \ldots, X_n)$ de taille $n = 100$ de X.
- Observation de l'échantillon: $e = (x_1, x_2, \ldots, x_n)$
- Estimations ponctuelles: $\overline{X} = m = 1,495$ et $s' = 0.01$.
- On a $n = 100 > 30$ donc un grand échantillon.
- On effectue un test (bilatéral) de $H0$: $M = \mu$ contre $H1$: $M \neq \mu$.
- On sait que $Z = \frac{\overline{X}-\mu}{s'/\sqrt{n}}$ suit approximativement la loi normale $\mathcal{N}(0, 1)$.

Test Z de conformité: exemple

- On calcule $z = \frac{|\bar{x}-\mu_0|}{\frac{0.01}{\sqrt{100}}} = 5$.
- On détermine $z_{\alpha/2}$ tel que $\Pr(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$: pour $\alpha = 0.05$, on trouve $z_\alpha = 1.96$.
- Comme $z \notin]-z_{\alpha/2}, z_{\alpha/2}[$ ($z = 5 > z_{\alpha/2} = 1.96$), on rejette $H0$ avec une probabilité α de se tromper: la machine n'est pas bien réglée.

TEST Z D'HOMOGÉNIÉTÉ

- On veut comparer les moyennes observées dans deux échantillons.
- Paramètre étudié: moyennes.
- Taille de l'échantillon ≥ 30 par échantillon.
- Hypothèses:
  - $H0$: $\mu_1 = \mu_2$,
  - $H1$ bilaterale: $\mu_1 \neq \mu_2$,
  - $H1$ unilaterale: $\mu_1 > \mu_2$ ou $\mu_1 < \mu_2$.

TEST Z D'HOMOGÉNÉITÉ

- Formulation:
  - μ_1 et μ_2: moyennes inconnues des deux populations d'où sont tirés les échantillons,
  - m_1 et m_2: moyennes observées des 2 échantillons,
  - s_1^2 et s_2^2: variances des 2 échantillons,
  - n_1 et n_2: effectifs des 2 échantillons.
- Conditions d'applications: Taille ≥ 30 (sinon: test de T de student).
- Sous l'hypothèse $H0$, $Z = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ suit approximativement la loi normale $\mathcal{N}(0, 1)$. Le résultat reste valable si on remplace σ_1^2 et σ_2^2 par leurs estimations s_1^2 et s_2^2.

PRINCIPE DU TEST Z D'HOMOGÉNIÉTÉ:

- Si $H0$ est vrai:
  - m_1 et m_2 sont deux valeurs possible d'une variable normale centrée autour de la moyenne commune au deux populations,
  - La difference $\Delta = m_1 - m_2$ suit une loi normale de moyenne 0,
  - Le rapport de cette différence sur son écart type suit une loi de Z,
- Calcul:

$$z = \frac{|m_1 - m_2|}{\sqrt{\frac{s_1'^2}{n_1} + \frac{s_2'^2}{n_2}}}.$$

PRINCIPE DU TEST Z D'HOMOGÉNIÉTÉ: DÉCISION

- Si $H1$ bilatérale ($\alpha = 0.05$):
  - si $z < z_{\alpha/2} = z_{0.025} = 1.96$, on accepte $H0$ (interprétation: μ_1 n'est pas significativement different de μ_2).
  - si $z \geq 1.96$, on rejette $H0$ (interprétation: μ_1 diffère significativement de μ_2).
- Si $H1$ unilatérale:
  - si $z < z_{\alpha} = z_{0.05} = 1.645$, on accepte $H0$ (interprétation: μ_1 n'est pas significativement supérieur (ou inférieur) à μ_2).
  - si $z \geq z_{\alpha} = z_{0.05} = 1.645$, on rejette $H0$ (interprétation: μ_1 est significativement supérieur (ou inférieur) à de μ_2).

PRINCIPE DU TEST Z D'HOMOGÉNIÉTÉ: EXEMPLE

- On désire comparer la pression artérielle diastolique d'un groupe de sujets sains et d'un groupe de sujets atteints de drépanocytose. Une étude donne les résultats suivants:
  - Formulez les hypothèses $H0$ et $H1$
  - Que concluez vous ?
-

	Effectif (n)	Pression artérielle diastolique	Variance (s'^2)
Sujets sains	88	70.1	10.8
Sujets drépanocytaires	85	61.8	6.9

Principe du test Z d'homogénéité: exemple

- $H0$: les pressions artérielles sont identiques $\mu_1 = \mu_2$.
- $H1$: la pression artérielle est différente chez les sujets drépanocytaires $\mu_1 \neq \mu_2$.
- $s_d = \sqrt{\frac{s_1'^2}{n_1} + \frac{s_2'^2}{n_2}} = \sqrt{\frac{10.8}{88} + \frac{6.9}{85}} = 0.45$
- $z = \frac{|70.1 - 61.8|}{0.45} = 18.4$.
- $18,4 > 1,96$: on rejette $H0$.
- La pression artérielle des sujets drépanocytaires est significativement différente de celle des sujets sains.
- $p < 10^{-5}$.

OBJECTIF

- Lorsque la taille des échantillons est faible (n<30) le rapport entre les différences de leurs moyennes et l'écart type ne suit pas une loi normale centrée réduite Z
- On utilise alors le test T de Student.
- Le test de Student sert à comparer:
  - Une moyenne observée à une moyenne théorique (test de conformité).
  - Les moyennes de 2 petits échantillons (test d'homogénéité).
  - Les moyennes de 2 petites séries appariées (non traité).

PRINCIPE DU TEST T

- Principe : idem à Z
- On calcule la difference Δ entre les moyennes.
- On estime l'écart type s_d de la difference Δ.
- On calcule $t_o = |\Delta|/s_d$.
- On compare cette valeur à la distribution théorique de la loi T de Student.
- On utilise la table de la loi T.

Interprétation du test T

- $H1$ bilatérale:
  - $t_o < T_{5\%}$ la différence entre les paramètres n'est pas significative.
  - $t_o > T_{5\%}$ la différence entre les paramètres est significative.
- $H1$ unilatérale:
  - La valeur seuil pour un risque de 5% est donné par la valeur de $T_{10\%}$.
  - $t_0 < T_{10\%}$ la différence entre les paramètres n'est pas significative.
  - $t_0 > T_{10\%}$ la différence entre les paramètres est significative. L'un des paramètres est inférieur (ou supérieur) à l'autre.

CONDITIONS D'APPLICATION DE TEST T:

- Utilisable si petits effectifs.
- Mais la distribution de la variable dans les populations doit être normale.
- Et les populations doivent avoir des variances identiques:
  - Soit on le sait.
  - Soit on le teste (test de F de comparaison des variances).

Utilisation de la table T

- La table de T est plus difficile à utiliser que la table de Z
- Il y a autant de table de T que de degré de liberté.
- "d.d.l" c'est l'effectif d'un échantillon-1.
  - Pour 1 échantillon : ddl = n-1
  - Pour 2 échantillons : ddl = $(n_1 - 1) + (n_2 - 1)$.
  - En ligne les valeurs possible de d.d.l.
  - En colonne les valeurs de α.

- Considérons un caractère quantitatif représenté par une variable aléatoire X d'espérance mathématique μ, d'écart-type σ, et un échantillon $\underline{X} = (X_1, \ldots, X_n)$ de taille n de cette variable.
- On compare une moyenne observée dans un échantillon de petite taille à une moyenne connue dans la population de référence.
- Paramètre étudié: moyenne avec σ inconnu.
- Taille de l'échantillon < 30.
- Hypothèses:
  - $H0$: $M = \mu$,
  - $H1$ bilaterale: $M \neq \mu$,
  - $H1$ unilaterale: $M > \mu$ ou $M < \mu$.

TEST T DE CONFORMITÉ

- Formulation:
  - μ: moyenne théorique connue de la population de référence,
  - M: moyenne inconnue de la population d'où est issu l'échantillon,
  - $\overline{X} = m$: moyenne observée de l'échantillon,
  - s: écart type de l'échantillon,
  - n: effectif.
  - ddl : degré de liberté.
- Conditions d'applications: la distribution de la variable doit être supposée normale dans la population d'où est issu l'échantillon.

Principe du test T:

- Si $H0$ est vrai, le rapport de la difference $\Delta = |m - \mu|$ sur l'écart type de μ suit une loi T de student de n-1 d.d.l.
- Test T de student:

$$t = \frac{|m - \mu|}{\frac{s'}{\sqrt{n}}} \quad \text{avec } d.d.l = n - 1 \, .$$

DÉCISION

- Si $H1$ bilatérale:
  - si $t < T_\alpha = T_{5\%}$, on accepte $H0$ (interprétation: M n'est pas significativement different de μ).
  - si $t \geq T_\alpha = T_{5\%}$, on rejette $H0$ (interprétation: M diffère significativement de μ).
- Si $H1$ unilatérale:
  - si $t < T_{10\%}$, on accepte $H0$ (interprétation: M n'est pas significativement supérieur (ou inférieur) à μ).
  - si $t \geq T_{10\%}$, on rejette $H0$ (interprétation: M est significativement supérieur (ou inférieur) à de μ).

Test T de conformité: exemple

- Dans un échantillon de 18 sujets suspects d'être atteints de trypanosomiase, on mesure la quantité de protéines dans le liquide céphalorachidien. On trouve dans ce groupe une protéinorachie moyenne de 460 mg/l avec un écart type de 280 mg/l.
- Dans la population générale, la protéinorachie est en moyenne de 300 mg/l.
- On se demande si ce groupe de sujet présente une protéinorachie différente de normale ?
  - Formulez les hypothèses $H0$ et $H1$.
  - Quel test utilisez-vous ? Justifiez la réponse
  - Que concluez vous ?

Test T de conformité: exemple

- $H0$: la proteinorachie des sujets atteints de trypanosomiase ne diffère pas de celle de la population générale.
- $H1$: la proteinorachie des sujets atteints de trypanosomiase est différente de celle de la population
- $n < 30$: Test de T.
- Condition d'application: on suppose que la proteinorachie est distribuée normalement chez les sujets atteints de trypanosomiase
- $t = (460 - 300)/(280/\sqrt{18}) = 2.4$.
- d.d.l =17.
- $T_{5\%}$ pour 17 d.d.l = 2.11.
- $2.4 > 2.11$: on rejette $H0$.
- la proteinorachie des sujets atteints de trypanosomiase est significativement différente de celle de la population $p < 0.03$.

TEST T D'HOMOGÉNIÉTÉ (1)

- On veut comparer les moyennes dans 2 échantillons de petite taille (extraits de populations gaussiennes).
- On suppose que $n_1 \leq 30$ ou $n_2 \leq 30$, et que les échantillons E_1 et E_2 sont indépendants.
- On suppose de plus que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1, \sigma_1)$ et $\mathcal{N}(\mu_2, \sigma_2)$, et que $\sigma_1 = \sigma_2 = \sigma$ inconnus.
- Paramètre étudié: moyennes.
- Taille de l'échantillon: au moins un inférieur à 30.
- Hypothèses:
  - $H0$: $\mu_1 = \mu_2$,
  - $H1$ bilaterale: $\mu_1 \neq \mu_2$,
  - $H1$ unilaterale: $\mu_1 > \mu_2$ ou $\mu_1 < \mu_2$.

Test T d'homogénéité (2)

- Formulation:
  - μ_1 et μ_2: moyennes inconnues des deux populations d'où sont tirés les échantillons,
  - m_1 et m_2: moyennes observées des 2 échantillons,
  - s_1^2 et s_2^2: variances des 2 échantillons,
  - n_1 et n_2: effectifs des 2 échantillons.
  - ddl: degré de liberté

Test T d'homogénéité: Conditions d'applications (3)

- Les distributions de la variable dans les populations d'où sont tirés les échantillons doivent être normales.
- Les variances des deux populations d'où sont tirés les échantillons doivent être égales.
- Sous l'hypothèse $H0$, $T = \frac{\overline{X_1} - \overline{X_2}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ suit approximativement la loi de Student à $n_1 + n_2 - 2$ degrés de liberté.
- Comme on ne connait pas $\sigma_1 = \sigma_2$, on doit d'abord tester l'égalité des variances $\sigma_1^2 = \sigma_2^2$.

- Si cette hypothèse est retenue, alors cette valeur commune σ (estimation de la variance commune aux 2 échantillons) peut être estimer par:

$$s'^2 = \frac{(n_1 - 1)s_1'^2 + (n_2 - 1)s_2'^2}{n_1 + n_2 - 2}.$$

Principe du test T d'homogénéité:(4)

- Si $H0$ est vraie, le rapport de la difference $\mu_1 - \mu_2$ sur son écart type suit une loi de T de Student lorsque les effectifs sont faibles.
- Écart type de la différence $\Delta = \mu_1 - \mu_2$ est donné par:

$$s_d = \sqrt{\frac{s'^2}{n_1} + \frac{s'^2}{n_1}}\,.$$

- Test T de Student:

$$t = \frac{|m_1 - m_2|}{s_d} \quad \text{avec } d.d.l = n_1 + n_2 - 2\,.$$

INTERPRÉTATION DU TEST T (5)

- $H1$ bilatérale:
  - $t_o < T_{5\%}$ μ_1 n'est pas significativement différent de μ_2.
  - $t_o > T_{5\%}$ μ_1 diffère significativement de de μ_2.
- $H1$ unilatérale:
  - La valeur seuil pour un risque de 5% est donné par la valeur de $T_{10\%}$.
  - $t_0 < T_{10\%}$ μ_1 n'est pas significativement superieur (ou inferieur) à μ_2.
  - $t_0 > T_{10\%}$ μ_1 est significativement supérieur (ou inférieur) à μ_2.

TEST D'HOMOGÉNIÉTÉ: EXEMPLE (6)

- On a mesuré un marqueur biologique chez 2 séries de sujets, l'une composée de sujets sains, l'autre de sujets atteints d'hépatite alcoolique. L'étude a trouvé les résultats suivants:

	Effectif (n)	Moyenne du marqueur (g/l)	Écart type s'
Sujets sains	15	1,6	0,19
Sujets alcooliques	12	1.4	0.21

  - Formuler les hypothèses.
  - Quel test choisissez vous ?
  - Quelles en sont les conditions d'application
  - Que concluez vous ?

Test T d'homogénéité: exemple (7)

- $H0$: la valeur moyenne du marqueur est identique dans les 2 populations
- $H1$: la valeur moyenne du marqueur est différente chez les sujets atteints d'hépatite alcoolique.
- $n < 30$: test de T
- Condition d'application : on suppose que:
  - le marqueur se distribue normalement dans les 2 populations
  - Les variances des 2 populations sont égales (sinon on effectue un test de cochran (hors programme)).
- $s'^2 = \frac{(15-1)(0.19)^2+(12-1)(0.21)^2}{15+12-2} = 0.04$
- $s_d = \sqrt{\frac{0.04}{15} + \frac{0.04}{12}} = 0.077, \quad t = \frac{1.6-1.4}{0.077} = 2.6$.

Test T d'homogénéité: exemple (8)

- ddl = 15+12-2 = 25.
- $T_{5\%}$ pour 25 ddl = 2.06.
- 2,6 > 2,06 : on rejette $H0$.
- Les malades atteints d'hépatite alcoolique présentent une valeur du marqueur significativement différente de celle des sujets sains $p < 0,02$.

TEST DE D'HOMOGÉNIÉTÉ: (1)

- Les échantillons E_1 et E_2 sont supposés indépendants.
- On suppose de plus que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1, \sigma_1)$ et $\mathcal{N}(\mu_2, \sigma_2)$.
- Les variances σ_1 et σ_2 des deux populations étudiées sont inconnues.
- On les estime à partir des échantillons en calculant $s_1'^2$ et $s_1'^2$.
- Test de $H0$: $\sigma_1 = \sigma_2$ contre $H1$: $\sigma_1 \neq \sigma_2$.
- On les compare avec un test de F de snédécor.

TEST DE D'HOMOGÉNIÉTÉ: (2)

- Sous l'hypothèse $H0$, $F = \frac{S_1'^2}{S_2'^2}$ suit la loi de Snedecor à $(n_1 - 1, n_2 - 1)$ degrés de liberté qui est tabulée en fonction de α, ν_1 et ν_2:
  - ν_1 degrés de liberté de la variance du numérateur = taille de l'échantillon le plus grand -1,
  - ν_2 degrés de liberté de la variance du numérateur = taille de l'échantillon le plus petit -1.
- L'indicateur calculé est

$$f = \frac{s_1'^2}{s_2'^2}$$

Si nécesaire, on permute les échantillons de sorte que $f \geq 1$.

Test de d'homogénéité: (3)

- On détermine $f_\alpha(n_1 - 1, n_2 - 1)$ tel que

$$\Pr(F > f_\alpha(n_1 - 1, n_2 - 1)) = \frac{\alpha}{2}$$

(table de Fisher-snédécor), et on décide que:

  - si $f < f_\alpha(n_1 - 1, n_2 - 1)$, alors on ne peut rejeter $H0$;
  - si $f \geq f_\alpha(n_1 - 1, n_2 - 1)$, alors on rejette $H0$ et on accepte $H1$ avec une probabilité α de se tromper.

TEST CONFORMIÉ: (1)

- Considérons un caractère quantitatif représenté par une variable aléatoire X d'espérance mathématique μ, d'écart-type σ, et un échantillon $\underline{X} = (X_1, \ldots, X_n)$ de taille n de cette variable.
- Alors $Y = \frac{n-1}{\sigma} S'^2$ suit la loi de khi-deux à $n-1$ degrés de liberté.

TEST CONFORMIÉ: (2)

- **Test (bilatéral):** $H0$: $\sigma^2 = \sigma_0^2$ contre $H1$: $\sigma^2 = \sigma_0^2$.
- On calcule $y = \frac{n-1}{\sigma} s'^2$.
- On détermine a_α et b_α tels que $\Pr(Y^2 \geq a_\alpha = 1 - \frac{\alpha}{2}$ et $\Pr(Y^2 \geq b_\alpha = \frac{\alpha}{2}$ (table de χ^2).
- On a donc $\Pr(Y^2 \in]a_\alpha, b_\alpha[) = 1 - \alpha$ et et $\Pr(Y^2 \notin]a_\alpha, b_\alpha[) = \alpha$.
- On décide que:
  - si $y^2 \in]a_\alpha, b_\alpha[$, alors on ne peut rejeter $H0$;
  - si $y^2 \notin]a_\alpha, b_\alpha[$, alors on rejette $H0$ avec une probabilité α de se tromper.

TEST CONFORMIÉ: (3)

- Test (unilatéral) à droite: $H0$: $\sigma^2 = \sigma_0^2$ contre $H1$: $\sigma^2 > \sigma_0^2$. On détermine b'_α tel que $\Pr(Y^2 \geq b'_\alpha) = \alpha$, $b'_\alpha = b_{2\alpha}$ et on décide que:
  - si $y^2 < b_\alpha$, alors on ne peut rejeter $H0$;
  - si $y^2 \geq b_\alpha$, alors on rejette $H0$ avec une probabilité α de se tromper.
- Test (unilatéral) à gauche: $H0$: $\sigma^2 = \sigma_0^2$ contre $H1$: $\sigma^2 < \sigma_0^2$. On détermine a''_α tel que $\Pr(Y^2 \geq a''_\alpha) = 1 - \alpha$, $a''_\alpha = a_{2\alpha}$ et on décide que:
  - si $y^2 > a''_\alpha$, alors on ne peut rejeter $H0$;
  - si $y^2 \leq a''_\alpha$, alors on rejette $H0$ avec une probabilité α de se tromper.

Test d'homogénéité: exemple

- On a mesuré les dimensions d'une tumeur chez des souris traitées ou non avec une substance antitumorale. On a obtenu les résultats suivants:
- Souris témoins: $n_1 = 20$, $\overline{X}_1 = 7.075cm^2$ et $s'_1 = 0.576cm^2$.
- Souris traitées: $n_2 = 18$, $\overline{X}_2 = 5.850cm^2$ et $s'_2 = 0.614cm^2$.
- La différence observée est-elle significative ?

- S’agissant de recherche si la substance antitumorale a, ou non, un effet de diminution sur la tumeur, nous utiliserons un test unilatéral. Les hypothèses à tester l’une contre l’autre sont:
  - $H0$: $\mu_1 = \mu_2$,
  - $H1$: $\mu_1 > \mu_2$, en appelant μ_1 la moyenne de la population dont est tiré l’échantillon témoin, μ_2 la moyenne de la population dont est tiré l’échantillon 2.
- Les tailles sont trop petites (inférieures à 30).
- Comme les variances des populations sont inconnues, pour comparer les moyennes des deux populations, on peut utiliser un test de Student: il faut faire l’hyptohèse que les populations sont normales.
- De plus, il faut teste l’homogénéité des variances. S’il s’avère possible de faire une estimation commune:

$$s'^2 = \frac{(n_1 - 1)s_1'^2 + (n_2 - 1)s_2'^2}{n_1 + n_2 - 2}.$$

de la variance des deux populations,

Exemple: Test d'homogénéité des variances.

- Le test T:

$$t = \frac{|m_1 - m_2|}{s_d} \quad \text{avec } d.d.l = n_1 + n_2 - 2\,.$$

- Dans l'hypothèse (H0), T suit une loi de Student à $(n_1 + n_2 - 2)$ degrés de liberté.
- Pour tester l'homogénéité des variances, on calcul $f = \frac{s_1'^2}{s_2'^2} = 1.136$
- La table de la loi F de Fischer-Snedecor à $(17; 19)$ degrés de liberté donne une valeur seuil comprise entre 2.62 et 2.51.
- Donc, de toutes façons, la valeur observée appartient à l'intervalle d'acceptation (au seuil de 5%): on accepte l'hypothèse d'égalité des variances.
- On peut alors avoir une estimation ponctuelle de la variance commune:

$$s'^2 = \frac{(n_1 - 1)s_1'^2 + (n_2 - 1)s_2'^2}{n_1 + n_2 - 2} = 0.3531 \qquad s' = 0.5942\,.$$

Exemple: Test d'égalité des moyennes.

- $t_0 = \frac{7.075-5.850}{0.5942\sqrt{\frac{1}{20}+\frac{1}{18}}} = 6.345$
- La table de Student à 36 degrés de liberté donne une idée de la probabilité critique correspondante:

$$p = 2(1 - F_{36}(6.345)) < 2 * 10^{-3}$$

Cette probabilité critique est inférieure à 5%: la différence est significative.

EXERCICE-1

- Pour mesurer de pH d'une solution, on utilise un pH-mètre qui affiche un résultat dont la loi est $\mathcal{N}(m; (0.05)^2)$, où m est la vraie valeur du pH de la solution. On a mesuré le pH d'une solution A par 12 mesures indépendantes et trouvé une moyenne de 7.04 .

  1. L'intervalle de confiance de niveau 90% de m contient la valeur: $(A)\,6.5,\ (B)\,6.8,\ (C)\,7,\ (D)\,7.05, (E)\,7.08$.
  2. L'intervalle de confiance de niveau 95% de m est: $(A)\,[6.15718 - 7.12282,],\ (B)\,[7.02153 - 7.05847],\ (C)\,[7.00824 - 7.07176],\ (D)\,[7.051265 - 8.089976], (E)$ la taille de l'échantillon ne permet pas d'estimer.

EXERCICE-1: LA SUITE

- On a mesuré le pH d'une solution B par 10 mesures indépendantes et trouvé une moyenne de 7.05.

    3. Peut-on considérer que les deux solutions ont même pH ? Pour cela on effectue:
    (A.) une comparaison de deux pourcentages observés.
    (B.) une comparaison de deux moyennes d'échantillons indépendants.
    (C.) une comparaison entre un pourcentage observé et un pourcentage théorique.
    (D.) une comparaison entre une moyenne observée et une moyenne théorique
    (E.) une comparaison d'effectifs.
    4. Le résultat du test (variable de décision) est:
    $(A)\ 0.1342,\ (B)\ Z = 0.29461,\ (C)\ 0.46707\ (D)\ t = 2.38409$,
    $(E) 4.65781$.

EXERCICE-1: LA SUITE

5. Parmi conditions de validité du test, il y a:
   $(A.)$ Les variances des 2 populations sont égales
   $(B.)$ Les distributions de la variable dans les populations d'où sont tirés les échantillons doivent suivent une loi de Poisson.
   $(C.)$ $n * p_{th} \geq 5$ et $n * (1 - p_{th}) \geq 5$
   $(D.)$ $n_1 * \pi_1 \geq 5$ et $n_1 * (1 - p_1) \geq 5$, $n_2 * \pi_2 \geq 5$ et $n_2 * (1 - p_2) \geq 5$
   $(E.)$ Utilisable si petits effectifs.

6. Que peut-on conclure (niveau de confiance $= 95\%$):
   (A) les deux échantillons sont homogènes
   (B) les deux échantillons sont hétérogènes
   (C) les résulats de l'échantillon ne sont conforme aux résultats théoriques
   (D) les résulats de l'échantillon sont conforme aux résultats théoriques
   (E) les conditions de validité du test ne sont pas remplies.

EXERCICE-1: LA SUITE

- Pour mesurer de pH d'une solution, on utilise un nouveau pH-mètre qui affiche un résultat dont la loi est $\mathcal{N}(m; s^2)$, où m est la vraie valeur du pH de la solution et où s n'a pas été déterminé. On a mesuré le pH d'une solution A par 12 mesures indépendantes et trouvé une moyenne de 7,04 et un écart-type empirique de 7.04 , et le pH d'une solution B par 10 mesures indépendantes et trouvé une moyenne de 7.05 et un écart-type empirique de 0.08. Peut-on considérer que les deux solutions ont même pH ? Peut-on considérer que les deux solutions ont même pH ?

  7. Pour cela on effectue:
     (A.) une comparaison de deux pourcentages observés.
     (B.) une comparaison de deux moyennes d'échantillons indépendants.
     (C.) une comparaison entre un pourcentage observé et un pourcentage théorique.
     (D.) une comparaison entre une moyenne observée et une moyenne théorique
     (E.) une comparaison d'effectifs.

EXERCICE-1: LA SUITE

8. Quelle est la valeurs de la statistique théorique:(niveau de confiance $= 95\%$)
   (A) 2.71 (B) 3.84 (C) 4.61 (D) 5.99 (E) 12.59

9. Que peut-on conclure (niveau de confiance $= 95\%$):
   $(A.)$ Il y a une relation entre les deux traitements
   $(B.)$ Les résultats de l'échantillon ne sont pas conforme aux résultats de la population
   $(C.)$ Les résultats donnent une différence significative
   $(D.)$ Les résultats donnent une différence non significative
   $(E.)$ Les conditions de validité du test ne sont pas remplies.