M4202C - Statistique inférentielle
Cours 2 - Tests d'hypothèses

2021/2022 - A. Ridard

Vannes
Université Bretagne Sud

A propos de ce document

- Pour naviguer dans le document, vous pouvez utiliser :
  - le menu (en haut à gauche)
  - l'icône en dessous du logo IUT
  - les différents liens
- Pour signaler une erreur, vous pouvez envoyer un message à l'adresse suivante : anthony.ridard@univ-ubs.fr

Plan du cours

1. Introduction
   - Les faiseurs de pluie
   - Principe
2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)
3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

1 Introduction

2 Tests de conformité

3 Tests de comparaison de deux échantillons indépendants (projet)

1. Introduction
   - Les faiseurs de pluie
   - Principe

2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)

3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

Des relevés effectués pendant de nombreuses années ont permis d'établir que le niveau naturel des pluies dans la Beauce en millimètres par an suit une loi normale $\mathcal{N}(600, 100)$. Des entrepreneurs, surnommés faiseurs de pluie, prétendaient pouvoir augmenter de 50 mm le niveau moyen de pluie, ceci par insémination des nuages au moyen d'iodure d'argent. Leur procédé fut mis à l'essai entre 1951 et 1959 et on releva les hauteurs de pluies suivantes :

Année	1951	1952	1953	1954	1955	1956	1957	1958	1959
mm	510	614	780	512	501	534	603	788	650

Deux hypothèses s'affrontaient :

- ou bien l'insémination était sans effet
- ou bien elle augmentait réellement le niveau moyen de pluie de 50 mm

Si m désigne la moyenne de la v.a. X égale au niveau annuel de pluie, ces **hypothèses** pouvaient se formaliser comme suit :

$$\begin{cases} H_0 : m = 600 \text{ mm} \\ H_1 : m = 650 \text{ mm} \end{cases}$$

Les agriculteurs hésitaient à opter pour le procédé onéreux des faiseurs de pluie, ils avaient besoin d'être convaincus...

? Comment pouvaient-ils décider ?

Puisqu'il s'agit de tester la valeur de m, il est naturel d'utiliser la moyenne empirique $\bar{X} \sim \mathcal{N}\left(m, \frac{\sigma}{\sqrt{n}}\right)$.

Si H_0 **est vraie**, alors $\bar{X} \sim \mathcal{N}\left(m_0, \frac{\sigma}{\sqrt{n}}\right)$ donc $W = \dfrac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$ et donc $P(W \geq 1{,}64) = 5\%$.

La règle de décision [1] est alors :

- Si $w \geq 1{,}64$, on rejette H_0 (au profit de H_1) *avec un risque de se tromper égal à* $\alpha = 5\%$
- Si $w < 1{,}64$, on conserve H_0 (faute de preuves suffisantes)

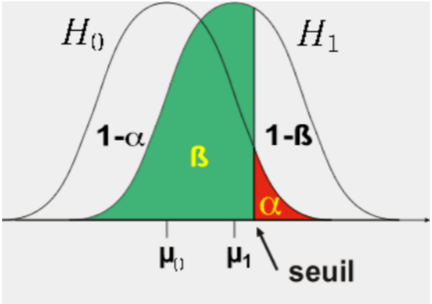
La v.a. W est appelée **variable de décision** et l'événement $\{W \geq 1{,}64\}$ la **région critique**.

Ici, les données relevées indiquent que $w = \frac{610{,}2-600}{100/3} = 0{,}306$ donc on conserve H_0 (aucune raison de rejeter H_0).

1. Ce raisonnement probabiliste est à comparer avec le raisonnement par l'absurde sauf que le résultat impossible est ici remplacé par un résultat très peu probable, et la négation de l'hypothèse de départ par l'hypthèse alternative H_1

Le risque de se tromper en conservant H_0 est :

$$\beta = P\left(U < \frac{655-650}{100/3}\right) = P(U < 0,15) = 0,56 \qquad \text{où} \quad U \sim \mathcal{N}(0,1)$$



- Ces erreurs correspondent à des risques différents :
  - le risque de première espèce α consiste à acheter un procédé d'insémination inefficace
  - le risque de deuxième espèce β consiste à perdre une occasion d'augmenter le niveau de pluie
- Dans la pratique des tests statistiques, on se fixe α ce qui fait jouer à H_0 un rôle prééminent :
  - H_0 peut être une hypothèse solidement établie n'ayant jamais été contredite par l'expérience
  - H_0 peut être une hypothèse de prudence (l'innocuité d'un vaccin, l'innocence d'une personne)
  - H_0 peut être une hypothèse à laquelle on tient pour certaines raisons
- α étant fixé, β sera alors déterminé comme résultat d'un calcul (à condition que la loi de probabilité sous H_1 soit connue). Notons cependant que β varie dans le sens contraire de α. En effet, diminuer α conduit à une règle de décision plus stricte qui aboutit à rejeter H_0 que dans des cas rarissimes et donc à conserver H_0 bien souvent à tort ce qui revient à augmenter β ou encore à diminuer la puissance du test [a] $1-\beta$.

a. La méthode de Neyman et Pearson permet de maximiser la puissance du test $1-\beta$ pour une valeur donnée de α

1 Introduction
- Les faiseurs de pluie
- Principe

2 Tests de conformité
- Moyenne
- Variance
- Proportion (hors programme)

3 Tests de comparaison de deux échantillons indépendants (projet)
- Moyennes
- Variances
- Proportions (hors programme)

Un test est un mécanisme qui permet de trancher entre deux hypothèses à l'aide d'un échantillon.

En notant H_0 et H_1 ces deux hypothèses [2], les quatre cas possibles sont représentés dans le tableau suivant :

Décision \ Vérité	H_0	H_1
H_0	$1-\alpha$	β
H_1	α	$1-\beta$

- On commence toujours par se fixer le risque de première espèce α.

2. On suppose qu'une et une seule est vraie

Voici alors la démarche d'un test :

1. Choix de H_0 et H_1
2. Détermination de la variable de décision et de sa loi sous H_0
3. Détermination de la forme de la région critique (selon H_1) et de ses bornes (selon α)
4. Calcul de la valeur expérimentale de la variable de décision
5. Décision : rejet ou non de H_0

1 Introduction

2 Tests de conformité

3 Tests de comparaison de deux échantillons indépendants (projet)

1. Introduction
   - Les faiseurs de pluie
   - Principe
2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)
3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

Supposons $X \sim \mathcal{N}(m,\sigma)$ et testons m.

Quand σ est connu

On utilise la variable de décision[3] $W = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ qui, sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$$

3. A comparer avec la fonction pivotale utilisée pour l'estimation par intervalle de confiance !

- Pour le test **unilatéral à droite** $\begin{cases} H_0 : m = m_0 \\ H_1 : m = m_1 \text{ avec } m_1 > m_0 \end{cases}$,
  la région critique est $[u_{(1-\alpha)};+\infty[$.
- Pour le test unilatéral à droite $\begin{cases} H_0 : m = m_0 \\ H_1 : m > m_0 \end{cases}$,
  la région critique est encore $[u_{(1-\alpha)};+\infty[$.
- $H_1 : m = m_1 \text{ avec } m_1 > m_0$ est une hypothèse dite **simple** alors que $H_1 : m > m_0$ est une hypothèse dite **composite**.
- Pour le test **unilatéral à gauche** $\begin{cases} H_0 : m = m_0 \\ H_1 : m < m_0 \end{cases}$,
  la région critique est $]-\infty;-u_{(1-\alpha)}]$.
- Pour le test **bilatéral** $\begin{cases} H_0 : m = m_0 \\ H_1 : m \neq m_0 \end{cases}$,
  la région critique est $]-\infty;-u_{(1-\frac{\alpha}{2})}] \cup [u_{(1-\frac{\alpha}{2})};+\infty[$.
- On fera essentiellement des tests unilatéraux avec H_1 composite.

Quand σ est inconnu

On utilise la variable de décision $W = \dfrac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}}$ qui, sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{S^*}{\sqrt{n}}} \sim \mathcal{T}_{n-1}$$

Quand l'échantillon n'est plus gaussien mais de grande taille

Quand σ est connu, on utilise la variable de décision $W = \dfrac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ qui,

sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \to +\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Quand σ est inconnu, on utilise la variable de décision $W = \dfrac{\bar{X} - m}{\frac{S^*}{\sqrt{n}}}$ qui,

sous $H_0 : m = m_0$, vérifie :

$$W = \frac{\bar{X} - m_0}{\frac{S^*}{\sqrt{n}}} \xrightarrow[n \to +\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Lors d'une enquête sur le temps de sommeil par nuit des enfants de 2 à 3 ans dans un département français, on a trouvé une moyenne de 10,2 heures dans un groupe de 40 enfants avec un écart type de 2,1 heures. En France, la moyenne du temps de sommeil par nuit est de 11,7 heures chez les enfants de cet âge.

Est-il possible de conclure que le temps de sommeil moyen dans ce département est inférieur à celui de la France ?

1. Introduction
   - Les faiseurs de pluie
   - Principe
2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)
3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

Vannes
Université Bretagne Sud

Supposons $X \sim \mathcal{N}(m,\sigma)$ et testons σ^2.

Quand m est connue

On utilise la variable de décision $W = \dfrac{nD}{\sigma^2}$ qui, sous $H_0 : \sigma^2 = \sigma_0^2$, vérifie :

$$W = \frac{nD}{\sigma_0^2} \sim \chi_n^2$$

Quand m est inconnue

On utilise la variable de décision $W = \frac{nS^2}{\sigma^2}$ qui, sous $H_0 : \sigma^2 = \sigma_0^2$, vérifie :

$$W = \frac{nS^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

Quand l'échantillon n'est plus gaussien mais de grande taille

Si la loi de X est unimodale pas trop dissymétrique, on utilise la var. de décision $W = \dfrac{S^{*2} - \sigma^2}{\sqrt{\frac{2S^{*4}}{n-1}}}$ qui, sous $H_0 : \sigma^2 = \sigma_0^2$, vérifie :

$$W = \frac{S^{*2} - \sigma_0^2}{\sqrt{\frac{2S^{*4}}{n-1}}} \xrightarrow[n \to +\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Un agent immobilier prétend, lors d'une interview, que le prix moyen des transactions immobilières dans un quartier niçois est de 2400 euros du mètre carré avec un écart-type de 220 euros. Le journaliste chargé du dossier à paraître dans une revue spécialisée décide de vérifier ces affirmations à partir des 50 dernières transactions effectuées par quatre agences du quartier

1695	2202	2722	2534	2494	2648	2298	1997	2118	2767
2867	2391	2029	2121	2105	2565	2652	2497	2822	2713
2014	2350	2343	2398	2505	2630	2169	2661	2325	2031
2683	2328	2710	2417	2264	2299	2531	2423	2592	2577
2568	1992	2872	2603	2415	2072	2475	2089	2140	2720

On en tire : $\bar{x} = 2408.66$ et $s^* = 272.01$

En admettant que le prix de vente suive une loi normale, tester au seuil de 5% les affirmations de l'agent immobilier.

1. Introduction
   - Les faiseurs de pluie
   - Principe
2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)
3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

Supposons l'échantillon de grande taille et testons p.

On utilise la variable de décision $W = \dfrac{F-p}{\sqrt{\frac{p(1-p)}{n}}}$ qui, sous $H_0 : p = p_0$, vérifie :

$$W = \frac{F-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \xrightarrow[n\to+\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Sur un échantillon de 300 patients traités par un certain remède, 243 ont été guéris.
La proportion de guérison est-elle significativement supérieure à 75% au seuil de 5% ?

1 Introduction

2 Tests de conformité

3 Tests de comparaison de deux échantillons indépendants (projet)

1. Introduction
   - Les faiseurs de pluie
   - Principe
2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)
3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

Supposons $X_A \sim \mathcal{N}(m_A, \sigma_A)$ et $X_B \sim \mathcal{N}(m_B, \sigma_B)$ indépendantes, et comparons m_A et m_B.

Quand les variances sont connues

On utilise la variable de décision $W = \dfrac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}} \sim \mathcal{N}(0,1)$$

Quand les variances sont inconnues et supposées égales

On utilise la variable de décision $W = \dfrac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\dfrac{S_p^2}{n_A} + \dfrac{S_p^2}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\dfrac{S_p^2}{n_A} + \dfrac{S_p^2}{n_B}}} \sim \mathscr{T}_{n_A+n_B-2}$$

où $S_p^2 = \dfrac{(n_A - 1)S_A^{*2} + (n_B - 1)S_B^{*2}}{n_A + n_B - 2}$ est la variance de pool [4].

4. S_p^2, appelée aussi variance combinée, n'est rien d'autre que la moyenne des variances corrigées des échantillons, pondérées par les tailles des échantillons diminuées de 1

Quand les variances sont inconnues et supposées différentes

On utilise la variable de décision $W = \dfrac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\dfrac{S_A^{*2}}{n_A} + \dfrac{S_B^{*2}}{n_B}}}$ qui, sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\dfrac{S_A^{*2}}{n_A} + \dfrac{S_B^{*2}}{n_B}}} \sim \mathcal{T}_m$$

où $m = \dfrac{1}{\dfrac{c^2}{n_A-1} + \dfrac{(1-c)^2}{n_B-1}}$ avec $c = \dfrac{\dfrac{S_A^2}{n_A-1}}{\dfrac{S_A^2}{n_A-1} + \dfrac{S_B^2}{n_B-1}}$.

Quand les échant. ne sont plus gaussiens mais de grandes tailles

Quand les variances sont connues

On utilise la variable de décision $W = \dfrac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}}$ qui,

sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}} \xrightarrow[n \to +\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Quand les échant. ne sont plus gaussiens mais de grandes tailles

Quand les variances sont inconnues et supposées égales

On utilise la variable de décision $W = \dfrac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\dfrac{S_p^2}{n_A} + \dfrac{S_p^2}{n_B}}}$ qui,

sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\dfrac{S_p^2}{n_A} + \dfrac{S_p^2}{n_B}}} \xrightarrow[n\to+\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Quand les échant. ne sont plus gaussiens mais de grandes tailles

Quand les variances sont inconnues et supposées différentes

On utilise la variable de décision $W = \dfrac{(\bar{X}_A - \bar{X}_B) - (m_A - m_B)}{\sqrt{\frac{S_A^{*2}}{n_A} + \frac{S_B^{*2}}{n_B}}}$ qui,

sous $H_0 : m_A = m_B$, vérifie :

$$W = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_A^{*2}}{n_A} + \frac{S_B^{*2}}{n_B}}} \xrightarrow[n\to+\infty]{\mathscr{L}} \mathscr{N}(0,1)$$

Le test de comparaison des moyennes doit être précédé par celui des variances dès que les variances sont inconnues

1. Introduction
    - Les faiseurs de pluie
    - Principe

2. Tests de conformité
    - Moyenne
    - Variance
    - Proportion (hors programme)

3. Tests de comparaison de deux échantillons indépendants (projet)
    - Moyennes
    - Variances
    - Proportions (hors programme)

Supposons $X_A \sim \mathcal{N}(m_A, \sigma_A)$ et $X_B \sim \mathcal{N}(m_B, \sigma_B)$ indépendantes avec $S_A^* > S_B^*$, et comparons σ_A^2 et σ_B^2.

Quand les moyennes sont connues

On utilise la variable de décision $W = \dfrac{\frac{D_A}{\sigma_A^2}}{\frac{D_B}{\sigma_B^2}}$ qui, sous $H_0 : \sigma_A^2/\sigma_B^2 = 1$, vérifie :

$$W = \frac{D_A}{D_B} \sim \mathscr{F}(n_A, n_B)$$ [5]

5. La Loi de Fisher de degrés de liberté n, p est définie par :

$$\mathscr{F}(n,p) = \frac{\chi_n^2/n}{\chi_p^2/p}$$

Quand les moyennes sont inconnues

On utilise la variable de décision $W = \dfrac{\frac{S_A^{*2}}{\sigma_A^2}}{\frac{S_B^{*2}}{\sigma_B^2}}$ qui, sous $H_0 : \sigma_A^2/\sigma_B^2 = 1$, vérifie :

$$W = \frac{S_A^{*2}}{S_B^{*2}} \sim \mathscr{F}(n_A - 1, n_B - 1)$$

Le test étant ici $\begin{cases} H_0 : \sigma_A^2/\sigma_B^2 = 1 \\ H_1 : \sigma_A^2/\sigma_B^2 > 1 \end{cases}$, la région critique est toujours $[f_{(1-\alpha)}; +\infty[$.

Les QI de 9 enfants d'un quartier d'une grande ville ont une moyenne de 107 avec un écart-type de 10. Les QI de 12 enfants d'un autre quartier ont une moyenne de 112 avec un écart-type de 9. On suppose que la variable aléatoire associée au QI suit une loi Normale.

Y a-t-il une différence significative au seuil de 5% entre les QI moyens des 2 quartiers ?

Quand les échant. ne sont plus gaussiens mais de grandes tailles

Si les lois sont unimodales pas trop dissymétriques,

on utilise $W = \dfrac{(S_A^{*2} - S_B^{*2}) - (\sigma_A^2 - \sigma_B^2)}{\sqrt{\frac{2S_A^{*4}}{n_A-1} + \frac{2S_B^{*4}}{n_B-1}}}$ qui, sous $H_0 : \sigma_A^2 = \sigma_B^2$, vérifie :

$$W = \frac{S_A^{*2} - S_B^{*2}}{\sqrt{\frac{2S_A^{*4}}{n_A-1} + \frac{2S_B^{*4}}{n_B-1}}} \xrightarrow[n\to+\infty]{\mathscr{L}} \mathcal{N}(0,1)$$

Des essais cliniques sont menés auprès de 137 patients atteints d'une maladie pulmonaire sans gravité afin de tester l'efficacité d'un traitement à la pulmotrycine.
Le protocole est le suivant :

- Des exercices respiratoires sont prescrits à 67 patients choisis au hasard ainsi qu'un placebo (groupe témoin).
- Les mêmes exercices respiratoires sont precrits aux 70 autres patients ainsi que de la pulmotricyne (groupe traité).
- Au bout de trois mois, l'amélioration de la capacité pulmonaire de chaque patient est mesurée sur une échelle de 0 (pas d'amélioration) à 10 (récupération totale).

Voici les résultats obtenus :

Amélioration	Groupe témoin (A)	Groupe traité (B)
0	2	0
1	8	0
2	4	3
3	7	0
4	14	10
5	9	14
6	5	13
7	4	17
8	7	10
9	2	3
10	5	0

On en tire : $\bar{x}_A = 4.78$, $s_A^2 = 7.37$, $\bar{x}_B = 6$ et $s_B^2 = 2.66$

L'amélioration moyenne du groupe traité est supérieure à celle du groupe témoin (de combien ?) mais cette différence doit-elle être attribuée aux bienfaits de la pulmotrycine ou aux fluctuations d'échantillonage (le même protocole sur d'autres individus n'aurait sans doute pas donné les mêmes résultats). Autrement dit, la supériorité observée entre les deux moyennes empiriques est-elle significative au seuil de 2% ?

1. Introduction
   - Les faiseurs de pluie
   - Principe
2. Tests de conformité
   - Moyenne
   - Variance
   - Proportion (hors programme)
3. Tests de comparaison de deux échantillons indépendants (projet)
   - Moyennes
   - Variances
   - Proportions (hors programme)

Supposons les deux échantillons de grandes tailles et indépendants, et comparons p_A et p_B.

On utilise la variable de décision $W = \dfrac{(F_A - F_B) - (p_A - p_B)}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_A} + \frac{1}{n_B})}}$ qui,

sous $H_0 : p_A = p_B$, vérifie :

$$W = \frac{F_A - F_B}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_A} + \frac{1}{n_B})}} \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

où $\hat{p} = \dfrac{n_A f_A + n_B f_B}{n_A + n_B}$.

Lors des primaires d'une campagne présidentielle, des sympathisants d'un parti sont interrogés sur leur opinion à propos d'un candidat avant et après un débat télévisé. Avant le débat, 64% des 980 personnes interrogées déclarent avoir une opinion positive sur le candidat. Après le débat, cette proportion n'est plus que de 61% chez 1001 autres personnes interrogées.

Cette baisse est-elle significative au seuil de 5% ?

Il existe des tests pour comparer des échantillons dépendants...