TESTS D'HYPOTHÈSE

Données continues – Risques et puissance

PASS Lyon Est

Dr. Nicolas ROMAIN-SCELLE

COMPARAISON DE MOYENNES

Un échantillon à référence : test de l'écart-réduit

Cadre général du problème

Une moyenne est estimée sur un échantillon lors d'une expérience

• On cherche à savoir si la moyenne estimée est significativement différente d'une moyenne de référence, à un risque α donné

Données du cas d'étude

• En 2013, la taille moyenne des hommes en Europe était estimée à 178 +/- 7 cm (référence).

 Lors d'une étude européenne, incluant 813 patients, la taille moyenne des hommes de la population dont est issue l'échantillon a été estimée à 174 +/- 5 cm.

 Existe-t-il une différence statistiquement significative entre la population de l'étude et la population générale des hommes en Europe ?

Variance de VA binaire et continue

Variable aléatoire binaire

La variance d'une VA binaire
 X dépend totalement de son espérance

•
$$Var(X) = E(X) * (1 - E(X))$$

• Pour $X \sim Bernouilli(\pi)$

Variable aléatoire continue

- La variance d'une VA continue Y n'est pas définie par son espérance
- E(Y) et Var(Y) sont deux valeurs indépendantes

Note : la variance d'une VA continue peut dépendre de son espérance, mais c'est hors du programme de PASS

Hypothèses nulle et alternative

Les hypothèses sont analogues au cas applicable au proportions

- H_0 : $\mu = \mu_0$
- $H_1: \mu \neq \mu_0$
- Notez qu'on ne pose pas d'hypothèses sur la variance de la variable étudiée.
- Dans cet exemple, elle est connue.

Modélisation du problème

- L'énoncé nous donne les valeurs suivantes :
- n = 813
- Pour la référence :
 - $\mu_0 = 178$
 - $\sigma = 7 \text{ donc } \sigma^2 = 7^2 = 49$
- Estimé dans l'échantillon :
 - m = 174
 - $s = 5 \text{ donc } s^2 = 25$

Conditions d'applications

- On souhaite utiliser le test de l'écart-réduit
- Il est nécessaire de pouvoir considérer que la variable testée suit asymptotiquement une loi normale

• Pour une variable continue, la condition est $n \ge 30$

• n = 813 > 30 donc les conditions d'application du TCL sont réunies

L'expérience sous H_0

- On considère que la VA X modélisant la taille d'un individu suit une loi : $X \sim \mathcal{L}(\mu, \sigma^2)$
- La moyenne M d'une série de n VA continues X i.i.d. suit donc asymptotiquement une loi normale (TCL) :

$$M \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Sous l'hypothèse nulle :

$$M_{H_0} \sim \mathcal{N}\left(\mu_0, \frac{\sigma^2}{n}\right)$$

La statistique de test

 Notre statistique de test est la VA Z, obtenue après centrage réduction :

$$Z = \frac{|M - \mu|}{\sqrt{\frac{\sigma^2}{n}}} \sim \mathcal{N}(0,1)$$

Sous l'hypothèse nulle :

$$Z_{H_0} = \frac{|M - \mu_0|}{\sqrt{\frac{\sigma^2}{n}}} \sim \mathcal{N}(0,1)$$

Application numérique

$$z_{H_0}^{obs} = \frac{|m - \mu_0|}{\sqrt{\frac{\sigma^2}{n}}} = \frac{|174 - 178|}{\sqrt{\frac{49}{813}}} \cong 16$$

- Pour $\alpha = 0.05$, le seuil de rejet est $z_{1-\frac{\alpha}{2}} = 1.96$
- $z_{H_0}^{obs} > z_{1-\frac{\alpha}{2}}$ donc **on rejette l'hypothèse nulle** au risque de première espèce de 5%.
- On doit s'interroger sur l'écart entre notre population étudiée et la population générale européenne

Test de l'écart-réduit pour la comparaison d'une moyenne estimée à une référence

$$Z_{H_0} = \frac{|M - \mu_0|}{\sqrt{\frac{\sigma^2}{n}}} \sim \mathcal{N}(0,1)$$

Avec

- M moyenne estimée dans l'échantillon
- μ_0 référence sous H_0
- σ² variance de la variable aléatoire testée
- n effectif de l'échantillon

• Seuil de rejet : quantile au niveau $1 - \frac{\alpha}{2}$ de la loi normale centrée réduite

• Pour $\alpha = 5\%, z_{1-\alpha/2} \cong 1,96$

Conditions :

- *n* ≥ 30
- σ^2 est connu

COMPARAISON DE MOYENNES

Deux échantillons indépendants : test de l'écart-réduit

Cadre général du problème

 On souhaite comparer deux moyennes estimées dans deux échantillons indépendants tirés dans une population

• Existe-t-il une différence statistiquement significative entre les moyennes estimées dans ces deux échantillons ?

Données du cas d'étude

- On présente des résultats de l'étude susmentionnée, qui visait à estimer le lien entre apports nutritionnels satisfaisant dans l'enfance et taille moyenne à l'âge adulte.
- Dans le groupe A (213 sujets), apports satisfaisants, la taille moyenne des hommes adultes est estimée à 181 +/- 5 cm.
- Dans le groupe B (214 sujets), apports insuffisants, la taille moyenne des hommes adultes est estimée à 178 +/- 6 cm.

Hypothèses nulle et alternative

Les hypothèses sont analogues au cas applicable au proportions

- H_0 : $\mu_A = \mu_B = \mu_0$, $\Delta_{AB} = 0$
- $H_1: \mu_A \neq \mu_B \neq \mu_0, \ \Delta_{AB} \neq 0$
- Notez qu'on ne pose toujours pas d'hypothèses sur la variance de la variable étudiée.
- Dans cet exemple, on devra les estimer.

Modélisation du problème

Groupe A :

- $m_A = 181$
- $s_A = 5 \text{ donc } s_A^2 = 25$
- $n_A = 213$

• Groupe B:

- $m_B = 178$
- $s_B = 6 \text{ donc } s_B^2 = 36$
- $n_B = 214$

- Conditions d'applications :
 - $n_A \ge 30$ conforme
 - $n_B \ge 30$ conforme
- Rappel : S² est un estimateur non biaisé de la variance de X dans un groupe :

•
$$S^2 = \frac{1}{n_{i-1}} \sum_{i=1}^{n_i} (x_i - \overline{x})^2$$

Distribution de la différence des estimateurs

$M_A - M_B$

 Nos deux échantillons (A et B) sont indépendants, on peut donc poser :

$$M_A - M_B \sim \mathcal{N}\left(\Delta_{AB}, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right)$$

Après centrage et réduction :

$$Z = \frac{(M_A - M_B) - \Delta_{AB}}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim \mathcal{N}(0,1)$$

Distribution de la différence des estimateurs

$M_A - M_B$ sous H_0

Pour obtenir la statistique de test sous l'hypothèse nulle, on substitue

$$Z = \frac{(M_A - M_B) - \Delta_{AB}}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim \mathcal{N}(0,1)$$

$$Z_{H_0} = \frac{(M_A - M_B) - 0}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim \mathcal{N}(0,1)$$

Application numérique

$$z_{H_0}^{obs} = \frac{|m_A - m_B|}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} = \frac{|181 - 178|}{\sqrt{\frac{25}{213} + \frac{36}{214}}} \cong 5,61$$

- Pour $\alpha = 0.05$, $z_{1-\frac{\alpha}{2}} = 1.96$ donc $z_{H_0}^{obs} > z_{1-\frac{\alpha}{2}}$
- On rejette l'hypothèse nulle au risque de première espèce de 5%.

Test de l'écart-réduit pour la comparaison de deux moyennes estimées

$$Z_{H_0} = \frac{|M_A - M_B|}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} \sim \mathcal{N}(0,1)$$

- Seuil de rejet : quantile au niveau $1 \frac{\alpha}{2}$ de la loi normale centrée réduite
- Pour $\alpha = 5\%$, $z_{1-\alpha/2} = 1,96$

- Avec
 - M_A , M_B estimateurs de la moyenne
 - S_A^2 , S_B^2 estimateurs non biaisés de la variance
 - n_A , n_B effectif des échantillons

Conditions:

•
$$n_A \ge 30$$

•
$$n_B \ge 30$$

COMPARAISON DE MOYENNES

Deux échantillons indépendants : test de Student

Pourquoi un nouveau test?

 Le test de l'écart-réduit repose sur le fait que l'estimateur d'intérêt suit asymptotiquement une loi normale : si n est grand, alors la moyenne ou la somme d'une série de X suit une loi normale.

- Souvent, on peut affirmer que X suit une loi normale directement.
- Dans ce cas, l'estimateur d'intérêt suit une loi normale, indépendamment de n.

Pourquoi un nouveau test?

• Dans ce cas, on peut faire des comparaisons de moyennes sans respecter $n \ge 30$

 On doit en revanche pouvoir dire que la variable aléatoire continue étudiée est normalement distribuée (mise en pratique hors programme)

Pourquoi un nouveau test?

- La comparaison de moyennes si n < 30 pose une difficulté
- On ne peut pas simplement remplacer la variance σ^2 par son estimateur S^2 comme précédemment

- En petit effectif, $Z_{H_0}=\frac{|M_A-M_B|}{\sqrt{\frac{S_A^2+S_B^2}{n_A}+\frac{S_B^2}{n_B}}}$ n'est pas normalement distribué
- Si la variance σ² est connue => test de l'écart-réduit, mais ce n'est presque jamais le cas

Le test de Student

 Raisonnement et statistique de test très semblable au test de l'écart-réduit. Si n est grand, les deux tests sont équivalents.

• La statistique de test suit une loi de Student (notée T), et non une loi normale centrée réduite (notée Z).

- On décrit deux formes de ce test en fonction des variances estimées :
 - Variances égales ou proches entre échantillons (cas présenté)
 - Variances inégales (test t de Welch, hors programme)

Egalité/proximité des variances estimées

Critère d'égalité/proximité des variances estimées :

$$0.5 < \frac{s_A^2}{s_B^2} < 2$$

 Si cette inégalité n'est pas vraie, le test de Student n'est pas applicable

Estimateur de la variance commune

$$T_{H_0} = \frac{M_A - M_B}{\sqrt{S_c^2 \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

• S_c^2 : estimateur de la **variance commune**

$$S_c^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}$$

Statistique du test de Student

Nous avons désormais une statistique de test applicable :

$$T_{H_0} = \frac{|M_A - M_B|}{\sqrt{S_c^2 \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim T_{(n_A + n_B - 2)}$$

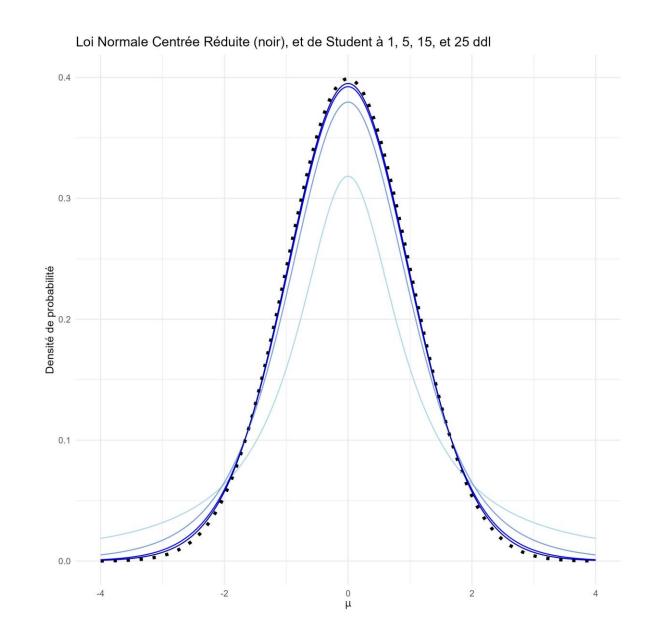
• La statistique T_{H_0} suit une loi de Student à $(n_A + n_B - 2)$ degrés de liberté

La loi de Student

Définie uniquement par son degré de liberté Symétrique autour de 0

Figure:

- Points : LNCR
- Courbes continues :
 Student à 1, 5, 15, 25
 ddl (de bas en haut)
- Plus n (donc les degrés de liberté) est élevé, plus la Student approche une LNCR

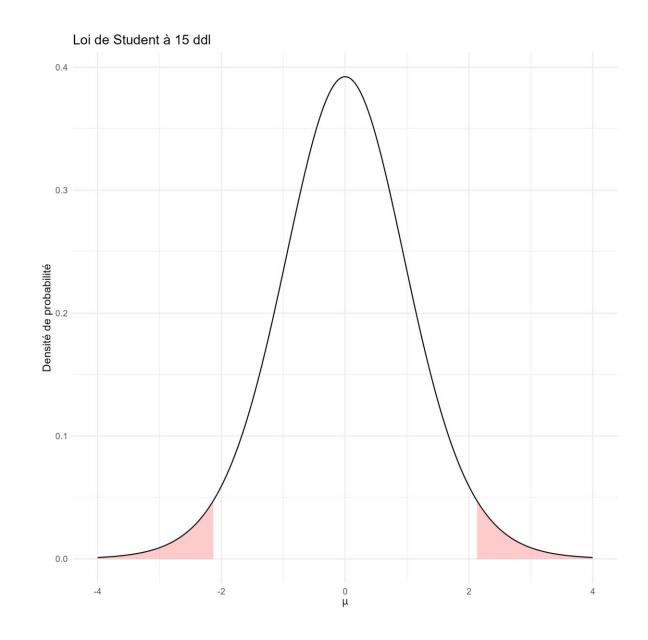


La loi de Student

Définie uniquement par son degré de liberté Symétrique autour de 0

Raisonnement du test : identique à la loi Normale

 α est distribué à gauche et à droite



Application numérique (exemple précédent)

• La statistique de test vaut $t_{H_0}^{obs} \cong 5,61$

• Le quantile $t_{1-\frac{\alpha}{2}}^{425}\cong 1{,}9656$ est le seuil de rejet

• On rejette l'hypothèse nulle au risque α de 5%

• Notez que $t_{1-\frac{\alpha}{2}}^{425}\cong z_{1-\frac{\alpha}{2}}$: le test de Student est asymptotiquement équivalent à celui de l'écart-réduit

La table de la loi de Student

- On calcule le ddl
- 2. Sur la ligne du ddl, on cherche la colonne correspondant à α
- 3. Si valeur exacte indisponible, on encadre
- 4. Si ddl > 100, dernière ligne (valeurs de la LNCR)

Attention : la table donne la probabilité d'être audessus du quantile combinée à gauche et à droite, on doit donc chercher $p=\alpha$ et non $p=1-\alpha/2$

Loi de Student

Soit T une variable aléatoire suivant une loi de Student à n degrés de liberté. Pour une probabilité p donnée, la table donne la valeur de t telle que P(|T|>t)=p

ddl	р	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,005	0,001
1		0,1584	0,3249	0,5095	0,7265	1,0000	1,3764	1,9626	3,0777	6,3138	12,7062	31,8205	63,6567	127,3213	636,6192
2		0,1421	0,2887	0,4447	0,6172	0,8165	1,0607	1,3862	1,8856	2,9200	4,3027	6,9646	9,9248	14,0890	31,5991
3		0,1366	0,2767	0,4242	0,5844	0,7649	0,9785	1,2498	1,6377	2,3534	3,1824	4,5407	5,8409	7,4533	12,9240
4		0,1338	0,2707	0,4142	0,5686	0,7407	0,9410	1,1896	1,5332	2,1318	2,7764	3,7469	4,6041	5,5976	8,6103
5		0,1322	0,2672	0,4082	0,5594	0,7267	0,9195	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	4,7733	6,8688
6		0,1311	0,2648	0,4043	0,5534	0,7176	0,9057	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	4,3168	5,9588
7		0,1303	0,2632	0,4015	0,5491	0,7111	0,8960	1,1192	1,4149	1,8946	2,3646	2,9980	3,4995	4,0293	5,4079
8		0,1297	0,2619	0,3995	0,5459	0,7064	0,8889	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	3,8325	5,0413
9		0,1293	0,2610	0,3979	0,5435	0,7027	0,8834	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,7809
10		0,1289	0,2602	0,3966	0,5415	0,6998	0,8791	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	3,5814	4,5869
11		0,1286	0,2596	0,3956	0,5399	0,6974	0,8755	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	3,4966	4,4370
12		0,1283	0,2590	0,3947	0,5386	0,6955	0,8726	1,0832	1,3562	1,7823	2,1788	2,6810	3,0545	3,4284	4,3178
13		0,1281	0,2586	0,3940	0,5375	0,6938	0,8702	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	3,3725	4,2208
					·	·	·	·	·	·					·

80	0,1261	0,2542	0,3867	0,5265	0,6776	0,8461	1,0432	1,2922	1,6641	1,9901	2,3739	2,6387	2,8870	3,4163
90	0,1260	0,2541	0,3866	0,5263	0,6772	0,8456	1,0424	1,2910	1,6620	1,9867	2,3685	2,6316	2,8779	3,4019
100	0,1260	0,2540	0,3864	0,5261	0,6770	0,8452	1,0418	1,2901	1,6602	1,9840	2,3642	2,6259	2,8707	3,3905
Inf	0,1257	0,2533	0,3853	0,5244	0,6745	0,8416	1,0364	1,2816	1,6449	1,9600	2,3263	2,5758	2,8070	3,2905
									•			•	•	•

Test de Student pour la comparaison de deux moyennes estimées

$$T_{H_0} = \frac{|M_A - M_B|}{\sqrt{S_c^2 \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim T_{(n_A + n_B - 2)}$$

$$S_c^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}$$

- Avec
 - M_A , M_B estimateurs de la moyenne
 - S_A^2 , S_B^2 estimateurs non biaisés de la variance
 - n_A , n_B effectif des échantillons

- Seuil de rejet : quantile au niveau $1-\frac{\alpha}{2}$ de la loi de Student (chercher $p=\alpha$ et non $p=1-\frac{\alpha}{2}$ dans la table)
- Conditions :
 - Variables normalement distribuées
 - Variances égales ou proches
- Ce test est dans votre formulaire

RISQUES ASSOCIÉS AU TEST

Les risques de faire erreur

 Lors de la réalisation d'un test d'hypothèse, deux risques existent de faire erreur dans la conclusion :

- Rejeter H₀ alors que celle-ci est vraie : le risque de première espèce (α)
- Ne pas rejeter H_0 alors que celle-ci est fausse : le risque de seconde espèce (β)

Les risques en une slide

Cf. Wikipedia EN, « Type I and Type II errors »

		L'hypothèse nulle (H_0) est	
		Vraie	Fausse
Décision sur l'hypothèse nulle (<i>H</i> ₀)	Ne pas rejeter	Inférence correcte (vrai négatif) (probabilité = 1-α)	Erreur de type II (faux négatif) (probabilité = β)
	Rejeter	Erreur de type I (faux positif) (probabilité = α)	Inférence correcte (vrai positif) (probabilité = 1-β)

Risque de première espèce, α

 Probabilité fixée par l'investigateur, en amont de la réalisation du test

Objectif : contrôler le risque de dire à tort qu'une différence existe

Sert donc de règle de décision pour le test

Risque de seconde espèce, β

- Probabilité non connue, dépendante de la vraie valeur du paramètre étudié
- Le risque de seconde espèce correspond au risque, sachant que l'hypothèse nulle est fausse, de ne pas la rejeter
- Par essence, on ne peut pas le calculer : il serait nécessaire de savoir quelle valeur du paramètre, parmi toutes celles possibles sous H₁, est vraie.
- Le risque β est d'intérêt dans le calcul de la puissance d'un test, en amont de celui-ci.

Puissance, $1 - \beta$

La puissance, complément du risque de seconde espèce,
 « mesure » la capacité d'un test à détecter une différence significative spécifique (pour une « vraie valeur » donnée).

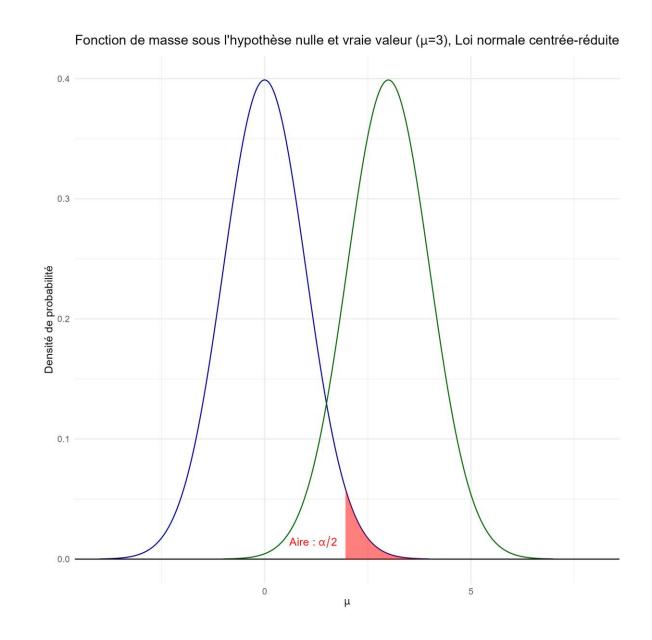
 Elle est utilisée en calcul du nombre d'effectifs (hors cadre de ce cours) : « De combien de sujets ai-je besoin dans mon expérience pour prouver une différence avec une puissance de 80% ? »

Rappel : α est une probabilité définie sous H_0 , matérialisée par l'aire en rouge sur cette figure

Le quantile $z_{1-\frac{\alpha}{2}}$ définit à gauche cette aire sous la courbe.

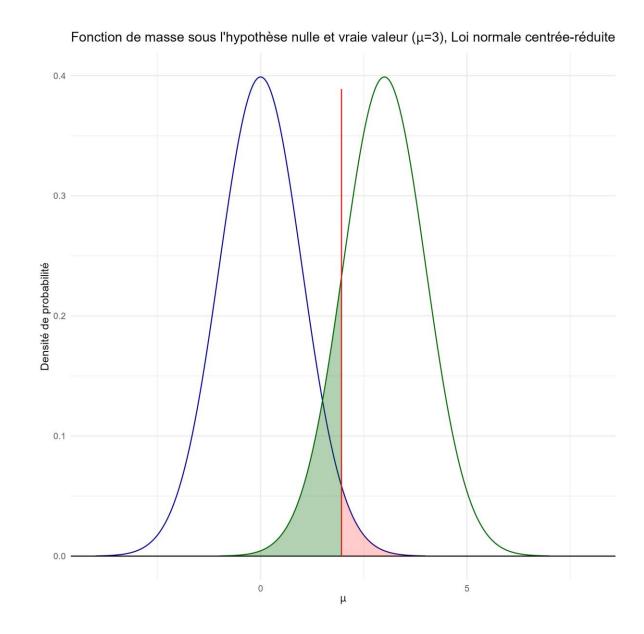
Courbes:

- Bleue : distribution du résultat expérimental sous H₀
- Verte : distribution du résultat expérimental réel (vraie valeur)



Sous la vraie distribution, on peut définir une aire **par la droite** avec le seuil de rejet.

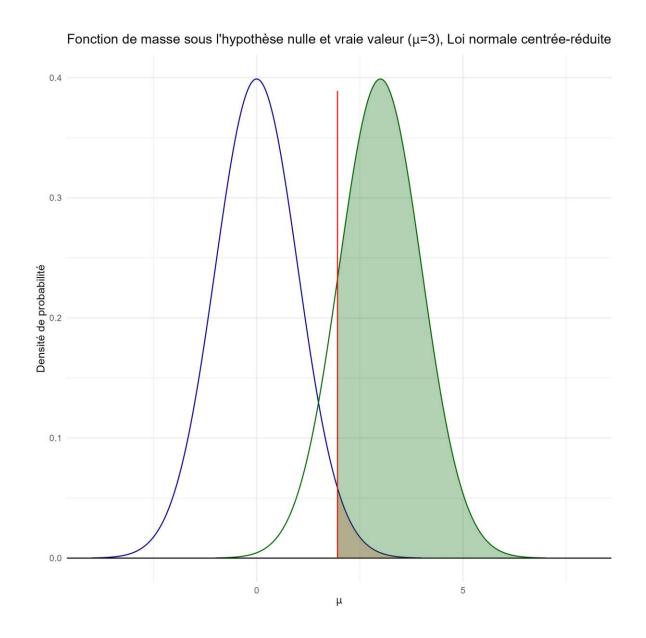
Cette aire correspond au risque de seconde espèce, β



La puissance est le complément de β , donc l'aire sous la vraie distribution définie par la gauche par le seuil de rejet.

La puissance varie en fonction de :

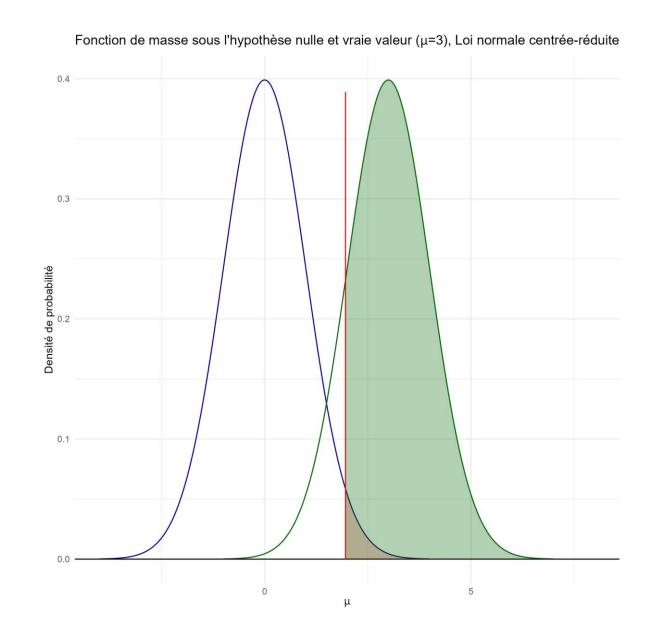
- α
- $\mu_0 \mu$
- (
- n via l'erreur-type



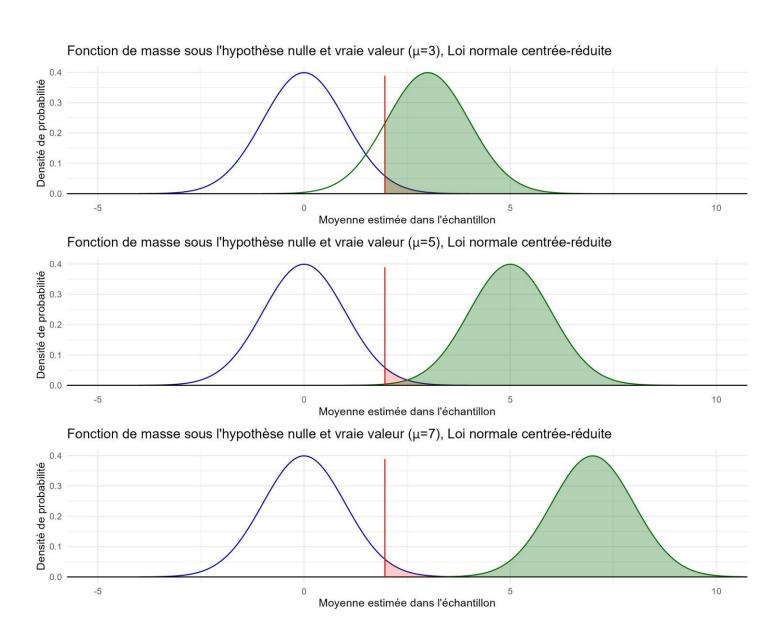
Connaître la puissance requiert de connaître la vraie valeur du paramètre.

Donc on ne connaît jamais la puissance effective d'un test.

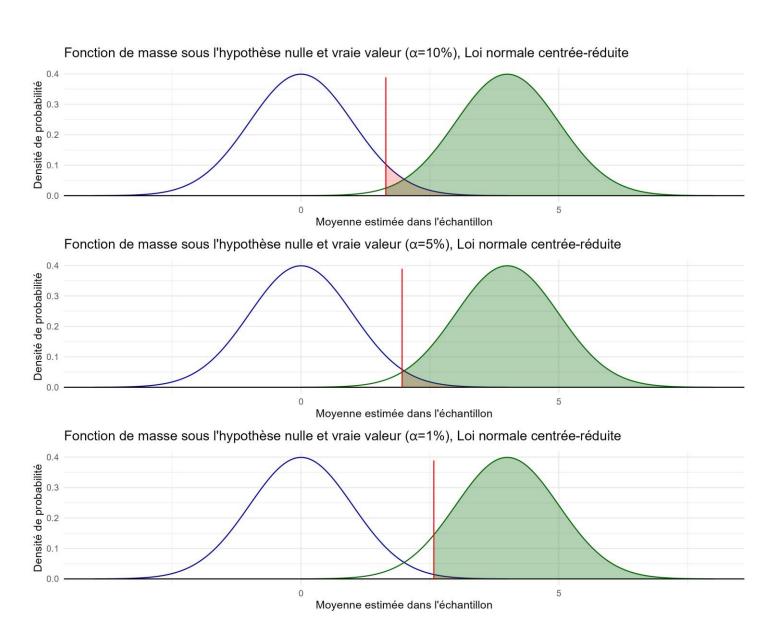
Le seul moment où la puissance est connue avec certitude, c'est lors du calcul du nombre de sujets nécessaire, où l'on présuppose une « vraie valeur ».



Variation en fonction de l'écart entre l'hypothèse nulle et la vraie valeur.

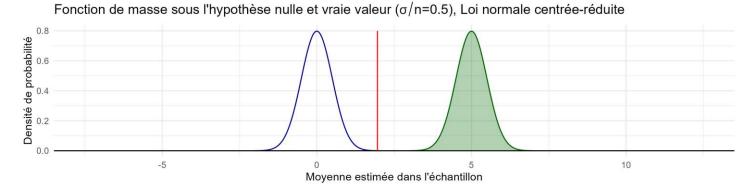


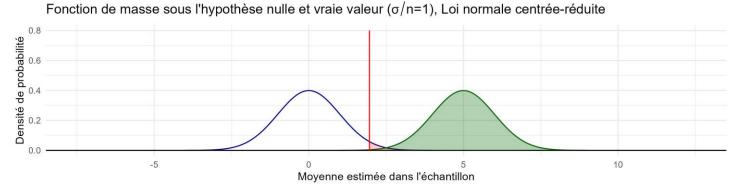
Variation en fonction de α .

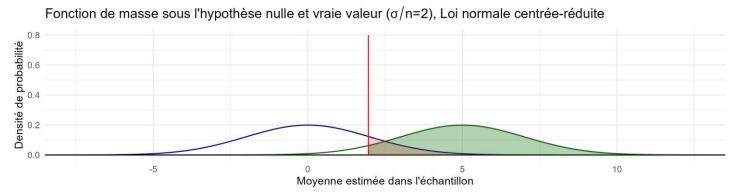


Variation en fonction de l'erreur-type (σ/\sqrt{n}) .

- Si l'écart-type augmente ou n diminue, l'erreur type augmente
- Si l'écart-type estimé diminue ou n augmente, l'erreur type diminue







LATÉRALITÉ DU TEST

Bilatéral ou unilatéral?

- La latéralité d'un test est intégralement définie par les hypothèses du test
- Un test bilatéral pose H_0 : $\mu_A = \mu_B$ et H_1 : $\mu_A \neq \mu_B$
- Un test unilatéral peut poser, par exemple H_0 : $\mu_A = \mu_B$ et H_1 : $\mu_A > \mu_B$
- Ce test unilatéral ne s'intéresse qu'au cas où la moyenne A est supérieure à la moyenne B => il est unilatéral

Bilatéral ou unilatéral?

La répartition du risque α sur la distribution de la statistique de test se fait ainsi :

- Dans le cas d'un test bilatéral :
 - A gauche **et** à droite de l'espérance pour le Z-test et le t-test
- Dans le cas d'un test unilatéral :
 - A gauche ou à droite de l'espérance pour le Z-test et le t-test
- Distinction sans objet pour le χ^2

Bilatéral ou unilatéral?

Les tests unilatéraux ne sont pas au programme

- En conséquence, dans les tests de l'écart-réduit et du Student, le risque α est distribué à gauche et à droite
- Dans le test du Chi-deux, le risque α est intégralement à droite

RÉCAPITULATIF DES TESTS

Variables catégorielles

Test du Chi-deux

- Deux proportions estimées entre elles (échantillons indépendants) (équivalent de l'écart-réduit)
- Deux variables catégorielles à K modalités entre deux échantillons indépendants (échantillons indépendants)

Variables continues

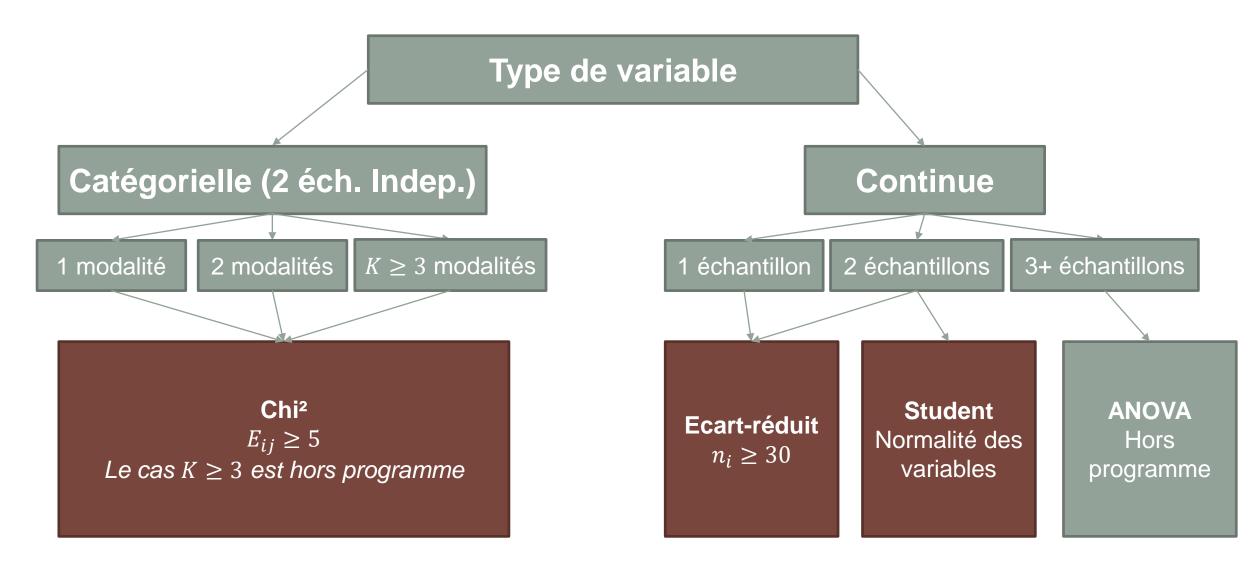
Test de l'écart-réduit

- Une moyenne estimée à référence
- Deux moyennes estimées entre elles (échantillons indépendants, variance égale ou proche)

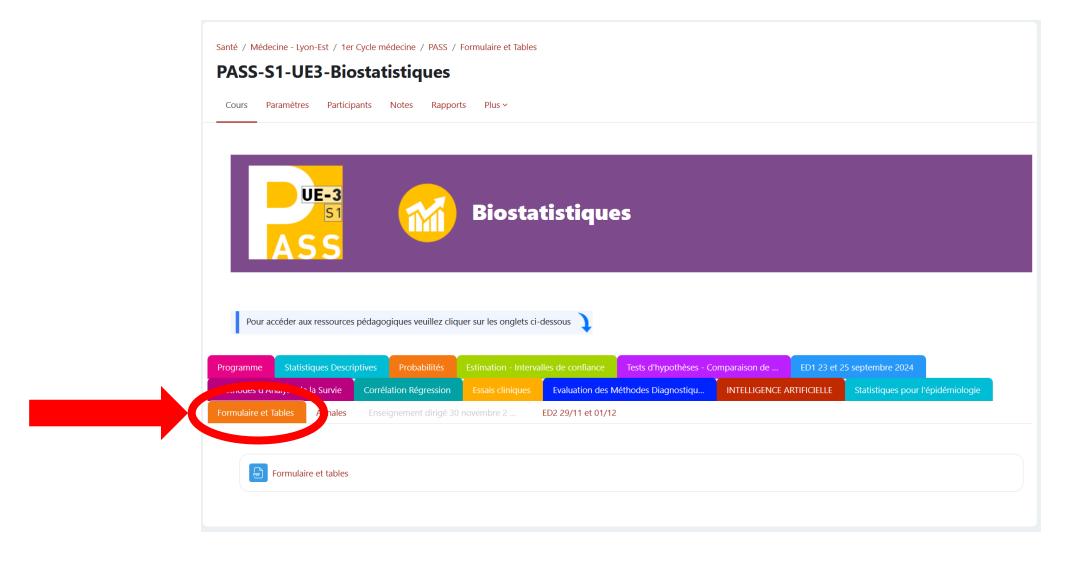
Test de Student

 Deux moyennes estimées entre elles (échantillons indépendants, variance égale ou proche)(asymptotiquement équivalent à l'écart-réduit)

Arbre décisionnel



Formulaire



Formulaire

Tests statistiques

$$\chi^{2} = \frac{(O_{2} - E_{2})^{2}}{\sum_{i=1}^{k} v_{i}}$$

$$\chi^{2}_{a} = \frac{(O_{2} - E_{2})^{2}}{E_{2}} + \frac{(O_{1} - E_{1})^{2}}{E_{1}}$$

$$T = \frac{(M_{1} - M_{2}) - 0}{\sqrt{\frac{(n_{1} - 1)s_{1}^{2} + (n_{2} - 1)s_{2}^{2}}{n_{1} + n_{2} - 2}} \sqrt{\frac{1}{n_{1}} + \frac{1}{n_{2}}}$$

CONCLUSIONS

Notions clés

- Quels tests sont applicables face aux variables catégorielles ?
 Continues ?
- Quelles sont les conditions d'application du test de l'écart-réduit
 ? De Student ?
- Comment retrouver les formules de calcul des statistiques de test ?
- Comment j'identifie le seuil de décision ?
- Quelle information je tire de la statistique calculée et le seuil de décision ?

MERCIPOUR VOTRE ATTENTION