II --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Les données sont non appariées
  – Moyenne sur 2 échan0llons indépendants

- Taille de l'échan0llon
  – Grands échan0llons ≥ 30
  – Pe0ts échan0llons <30

- La comparaison entre 2 moyennes observées mA et mB sur nA et nB individus respec0vement (promo0on PACES de 2 facultés A et B)

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Test basé sur l'écart réduit dans le cas des grands échantillons ie nA et nB ≥ 30

- $Z = \dfrac{\mu_A - \mu_B}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}} = \dfrac{m_A - m_B}{\sqrt{\dfrac{S_A^2}{n_A} + \dfrac{S_B^2}{n_B}}}$

Ecart type de la différence des moyennes

- S_A^2 et S_B^2 sont les variances estimées sur l'échantillon
- Si |Z| < 1.96 la différence n'est pas significative à 5%
- Si |Z| ≥ 1.96 la différence est significative à 5%
- Le risque Z lu dans la table de l'écart réduit fixe le degré de signification

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- La sta0s0que Z suit par approxima0on une loi normale centrée réduite

- Exemple on compare la moyenne des notes de sta0s0ques de PACES entre 2 facultés A et B
  - On 0re au sort 50 élèves (nA et nB) dans chacune des promo0ons et on compare la moyenne (note/100)
    - mA = 45/100 (d'écart type s = 3)
    - mB = 49/100 (d'écart type s = 5)

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Est---ce que la moyenne est différente entre les deux facultés (test bilatéral) ?

- Choix du test et vérifica0on des condi0ons d'u0lisa0on
  - 2 échan0llons de n ≥ 30
  - Deux variables con0nues
  - Moyennes mA et mB des notes suivent une loi normale

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Définir H0 et H1
  - H0 : μA = μB
  - H1 : μA ≠ μB

- Fixer le risque alpha et définir la règle de décision
  - Risque bilatéral 5%
  - Rejet de H0 si
    - $|Z| = \left|\frac{\mu_A - \mu_B}{\sqrt{\frac{\sigma A^2}{n_A}+\frac{\sigma B^2}{n_B}}}\right| = \left|\frac{m_A - m_B}{\sqrt{\frac{SA^2}{n_A}+\frac{SB^2}{n_B}}}\right| \geq 1.96$

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Calcul de la statistique

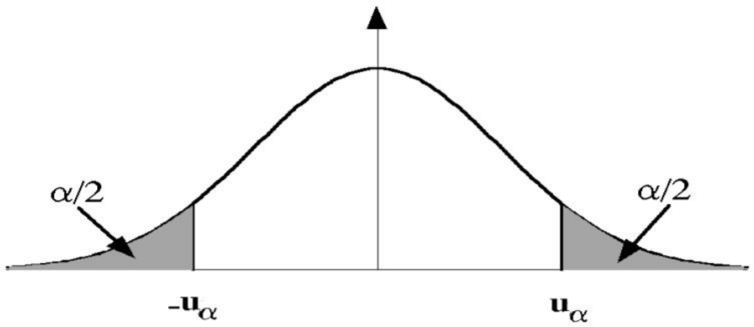
- $|Z| = \left| \frac{m_A - m_B}{\sqrt{\frac{SA^2}{n_A} + \frac{SB^2}{n_B}}} \right| = \left| \frac{45 - 49}{\sqrt{\frac{3^2}{50} + \frac{4^2}{50}}} \right| = \left| \frac{-4}{\sqrt{\frac{9}{50} + \frac{16}{50}}} \right| =$ 5.66

5.66 ≥ 1.96 on conclut que les moyennes des notes de PACES diffèrent entre la faculté A et B (ie mB significativement supérieure à mA)

- p $\leq$ 0,000 000 1

Loi normale centrée réduite

Table de l'écart réduit

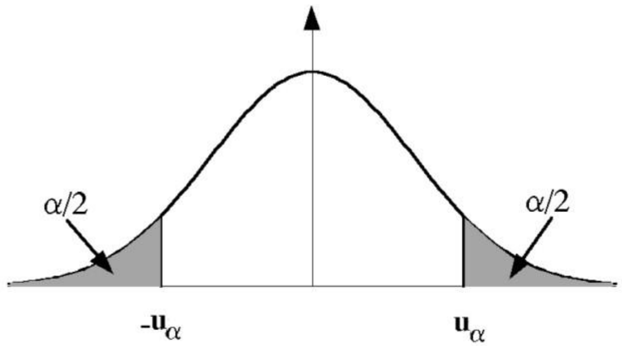


La table donne la probabilité α pour que l'écart réduit dépasse, en valeur absolue, une valeur donnée u, c'est-à-dire la probabilité de ne pas trouver z dans l'intervalle [-u ; u] centré sur 0. Chacune des 2 aires hachurées correspondent à une probabilité égales $\alpha/2$. La probabilité d'observer z dans l'intervalle]-u ; u[est évidemment 1- α.

Test bilatéral : lire α Test unilatéral à droite ou à gauche : diviser α par 2

α	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,00	∞	2,576	2,326	2,170	2,054	1,960	1,881	1,812	1,751	1,695
0,10	1,645	1,598	1,555	1,514	1,476	1,440	1,405	1,372	1,341	1,311
0,20	1,282	1,254	1,227	1,200	1,175	1,150	1,126	1,103	1,080	1,058
0,30	1,036	1,015	0,994	0,974	0,954	0,935	0,915	0,896	0,878	0,860
0,40	0,842	0,824	0,806	0,789	0,772	0,755	0,739	0,722	0,706	0,690
0,50	0,674	0,659	0,643	0,628	0,613	0,598	0,583	0,568	0,553	0,539
0,60	0,524	0,510	0,496	0,482	0,468	0,454	0,440	0,426	0,412	0,399
0,70	0,385	0,372	0,358	0,345	0,332	0,319	0,305	0,292	0,279	0,266
0,80	0,253	0,240	0,228	0,215	0,202	0,189	0,176	0,164	0,151	0,138
0,90	0,126	0,113	0,100	0,088	0,075	0,063	0,050	0,038	0,025	0,013

Table de l'écart réduit



α	0,001	0,000 1	0,000 01	0,000 001	0,000 000 1	0,000 000 01	0,000 000 001
u_α	3,29053	3,89059	4,41717	4,89164	5,32672	5,73073	6,10941

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Test t de Student dans le cas des petits échantillons ie nA et/ou nB < 30 (au moins l'un est petit)

  – $t = \frac{\mu_A - \mu_B}{\sqrt{\frac{\sigma^2}{n_A}+\frac{\sigma^2}{n_B}}} = \frac{m_A - m_B}{\sqrt{\frac{S^2}{n_A}+\frac{S^2}{n_B}}}$

- S^2 estimation de la variance supposée commune ie **test de l'égalité des variances au préalable**

  – $S^2 = \frac{(nA-1)\times SA^2 + (nB-1)\times SB^2}{nA+nB-2} = \frac{\sum(x-mA)^2 - \sum(x-mB)^2}{nA+nB-2}$

- Si |t| est inférieur à la valeur lue dans la table de t pour un ddl = $nA + nB - 2$

- **2 hypothèses normalité et égalité de variance** = test t robuste en première approximation

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Exemple identique au précèdent mais effectif de n_A = 9 et n_B =18

- Exemple on compare la moyenne des notes de statistiques de PACES entre 2 facultés A et B

  – On tire au sort n_A = 9 et n_B = 18 dans chacune des promotions et on compare la moyenne
  – mA = 45/100 (d'écart type $s = 3$)
  – mB = 49/100 (d'écart type $s = 5$)

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Est-ce que la moyenne est différente entre les deux facultés (test bilatéral) ?

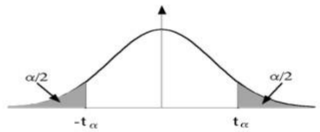
- Choix du test et vérification des conditions d'utilisation
  - 2 échantillons de n < 30
  - Deux variables continues, loi normale, on suppose l'égalité de variance
  - **On doit le faire le test d'égalité des variances test F au préalable (cf chap I)**
  - Moyennes mA et mB des notes suivent une loi normale
  - Statistique t = $\frac{m_A - m_B}{\sqrt{\frac{s^2}{n_A}+\frac{s^2}{n_B}}}$ suit une loi de Student à (9 + 18 -2) = 25 ddl

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Définir H0 et H1
  - H0 : µA = µB
  - H1 : µA ≠ µB

- Fixer le risque alpha et définir la règle de décision
  - Risque bilatéral 5%
  - Rejet de H0 si
    - t = $\frac{m_A - m_B}{\sqrt{\frac{s^2}{n_A}+\frac{s^2}{n_B}}}$ t (5%, ddl = 25) = 2.06
    - Rejet si |t| > 2.06

TABLE DU t DE STUDENT



ddl \ α	0,45	0,25	0,15	0,10	0,05	0,025	0,01	0,005	0,0005	TU
ddl \ α	**0,90**	**0,50**	**0,30**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**	**0,001**	**TB**
1	0,158	1,000	1,963	3,078	6,314	12,706	31,821	63,657	636,619	
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,598	
3	0,137	0,765	1,250	1,638	2,353	3,182	4,541	5,841	12,924	
4	0,134	0,741	1,190	1,533	2,132	2,776	3,747	4,604	8,610	
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869	
6	0,131	0,718	1,134	1,440	1,943	2,447	3,143	3,707	5,959	
7	0,130	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408	
8	0,130	0,706	1,108	1,397	1,860	2,306	2,896	3,355	5,041	
9	0,129	0,703	1,100	1,383	1,833	2,262	2,821	3,250	4,781	
10	0,129	0,700	1,093	1,372	1,812	2,228	2,764	3,169	4,587	
11	0,129	0,697	1,088	1,363	1,796	2,201	2,718	3,106	4,437	
12	0,128	0,695	1,083	1,356	1,782	2,179	2,681	3,055	4,318	
13	0,128	0,694	1,079	1,350	1,771	2,160	2,650	3,012	4,221	
14	0,128	0,692	1,076	1,345	1,761	2,145	2,624	2,977	4,140	
15	0,128	0,691	1,074	1,341	1,753	2,131	2,602	2,947	4,073	
16	0,128	0,690	1,071	1,337	1,746	2,120	2,583	2,921	4,015	
17	0,128	0,689	1,069	1,333	1,740	2,110	2,567	2,898	3,965	
18	0,127	0,688	1,067	1,330	1,734	2,101	2,552	2,878	3,922	
19	0,127	0,688	1,066	1,328	1,729	2,093	2,539	2,861	3,883	
20	0,127	0,687	1,064	1,325	1,725	2,086	2,528	2,845	3,850	
21	0,127	0,686	1,063	1,323	1,721	2,080	2,518	2,831	3,819	
22	0,127	0,686	1,061	1,321	1,717	2,074	2,508	2,819	3,792	
23	0,127	0,685	1,060	1,319	1,714	2,069	2,500	2,807	3,767	
24	0,127	0,685	1,059	1,318	1,711	2,064	2,492	2,797	3,745	
25	0,127	0,684	1,058	1,316	1,708	2,060	2,485	2,787	3,725	

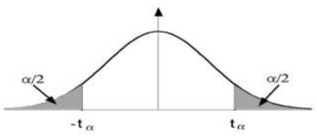
III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Calcul de la statistique

  - $S^2 = \frac{(nA-1)\times SA^2 + (nB-1)\times SB^2}{nA+nB-2} = \frac{\sum(x-mA)^2 - \sum(x-mB)^2}{nA+nB-2} = \frac{(9-1)\times 3^2+(18-1)\times 5^2}{9+18-2} = \frac{72+425}{25} = 19.88$
  - $t = \frac{m_A - m_B}{\sqrt{\frac{S^2}{n_A}+\frac{S^2}{n_B}}} = \frac{45-49}{\sqrt{\frac{19.88}{9}+\frac{19.88}{18}}} = -2.20$
  - $|t| \geq 2.06$

  - Conclusion on rejette H0 au risque alpha bilatéral de 5%
  - La moyenne des notes de faculté B est supérieure à celle de A ie elle est différente
  - p compris entre 0.05 et 0,02 (table de Student bilateral)

TABLE DU t DE STUDENT



ddl \ α	0,45	0,25	0,15	0,10	0,05	0,025	0,01	0,005	0,0005	TU
ddl \ α	0,90	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001	TB
1	0,158	1,000	1,963	3,078	6,314	12,706	31,821	63,657	636,619	
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,598	
3	0,137	0,765	1,250	1,638	2,353	3,182	4,541	5,841	12,924	
4	0,134	0,741	1,190	1,533	2,132	2,776	3,747	4,604	8,610	
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869	
6	0,131	0,718	1,134	1,440	1,943	2,447	3,143	3,707	5,959	
7	0,130	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408	
8	0,130	0,706	1,108	1,397	1,860	2,306	2,896	3,355	5,041	
9	0,129	0,703	1,100	1,383	1,833	2,262	2,821	3,250	4,781	
10	0,129	0,700	1,093	1,372	1,812	2,228	2,764	3,169	4,587	
11	0,129	0,697	1,088	1,363	1,796	2,201	2,718	3,106	4,437	
12	0,128	0,695	1,083	1,356	1,782	2,179	2,681	3,055	4,318	
13	0,128	0,694	1,079	1,350	1,771	2,160	2,650	3,012	4,221	
14	0,128	0,692	1,076	1,345	1,761	2,145	2,624	2,977	4,140	
15	0,128	0,691	1,074	1,341	1,753	2,131	2,602	2,947	4,073	
16	0,128	0,690	1,071	1,337	1,746	2,120	2,583	2,921	4,015	
17	0,128	0,689	1,069	1,333	1,740	2,110	2,567	2,898	3,965	

p compris entre 0.05 et 0.02 (table du t de Student bilateral avec 2.20 et 25 ddl

22	0,127	0,686	1,061	1,321	1,717	2,074	2,508	2,819	3,792	
23	0,127	0,685	1,060	1,319	1,714	2,069	2,500	2,807	3,767	
24	0,127	0,685	1,059	1,318	1,711	2,064	2,492	2,797	3,745	
25	0,127	0,684	1,058	1,316	1,708	2,060	2,485	2,787	3,725	

III --- Comparaison de 2 moyennes observées sur 2 échanRllons indépendants

- Remarque
  - **Non avons supposé ici que les variances étaient égales**
  - **En realité, il faut tester ce\e hypothèse d'égalité des variances**
  - **Si les variances ne sont pas égales et ou si elles les variables ne suivent pas une loi normale**

↓

- Tests non paramétriques Wilcoxon / Mann et whitney

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Les échan0llons ne sont pas indépendants

  – 2 examinateurs corrigent les copies de 100 étudiants (ie même échan0llon de copie)

  – Les tests précédents ne sont plus valables car ils présupposent l'indépendance des échan0llons = ils corrigent les copies des mêmes étudiants

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Les échantillons appariés avec des effectifs ≥ 30

  - Pour comparer les moyennes de deux échantillons appariés, on forme pour chaque paire la différence des deux mesures ***di*** et on compare la moyenne des n différences di à 0 par l'écart réduit

  - $Z = \frac{\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{m}{\frac{s}{\sqrt{n}}}$

  - m et S désignent la moyenne et l'écart type estimés sur l'échantillon des n différences

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Si |Z| < 1.96 les moyennes ne diffèrent pas significa0vement au seuil de 5 %

- Si |Z| ≥ 1.96 les moyennes différent significa0vement et le risque correspondant à Z lu dans la table de l'écart réduit fixe le degré de significa0on

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Les échantillons appariés avec des effectifs < 30

  – Pour comparer les moyennes de deux séries appariées de faible effectif, on forme pour chaque paire la différence des deux mesures et on compare la moyenne des différences à 0 par le rapport

  – $t = \frac{\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{m}{\frac{s}{\sqrt{n}}} = \frac{m-0}{\frac{s}{\sqrt{n}}}$

  – m et S désignent la moyenne et l'écart type estimés sur l'échantillon des n différences di

  – $S^2 = \frac{\sum_{i=1}^{n} d_i^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n}}{n-1}$

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Si |t| est inférieur à la valeur lue dans la table de t pour le nombre de degrés de liberté (n---1) et le risque 5 % les moyennes ne diffèrent pas significa0vement au seuil de 5 %

- Si |t| est supérieur les moyennes différent significa0vement et le risque indiqué par la table pour la valeur |t| trouvée fixe le degré de significa0on

- Formule applicable si la différence est distribuée selon une loi normale

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- On mesure l’effet du stress lié au examen PACES sur la glycémie de 9 étudiants
  - La glycémie suit une loi normale dans la popula0on dont sont issues les étudiants
  - Pour chaque étudiant 2 mesures sont faites avant et après les examens
- Résultats

étudiants	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
avant	5.5	4.3	6.5	4.5	5.2	4.3	5.0	5.4	5.2
après	5.4	6.7	6.5	6.0	5.2	5.0	4.8	4.7	4.5

- Le stress du à l’examen modifie t il la glycémie des étudiants ?

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

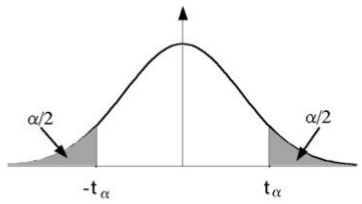
- Choix du test statistique et vérification des conditions d'utilisation
  - Données appariées, un seul échantillon, loi normale
  - Test de la différence à zéro
  - $t = \frac{\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{m}{\frac{s}{\sqrt{n}}} = \frac{m-0}{\frac{s}{\sqrt{n}}}$ suit une loi de Student à 9-1 = 8 ddl

- Définir les hypothèses
  - H0 : $\mu = 0$
  - H1 : $\mu \neq 0$

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Fixer le risque alpha et définir la règle de décision
  - Alpha 5 % bilatéral
  - Zone de rejet de H0 : 2.306 sur la table de Student à 8 ddl

TABLE DU t DE STUDENT



ddl \ α	0,45	0,25	0,15	0,10	0,05	0,025	0,01	0,005	0,0005	TU
ddl \ α	**0,90**	**0,50**	**0,30**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**	**0,001**	**TB**
1	0,158	1,000	1,963	3,078	6,314	12,706	31,821	63,657	636,619	
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,598	
3	0,137	0,765	1,250	1,638	2,353	3,182	4,541	5,841	12,924	
4	0,134	0,741	1,190	1,533	2,132	2,776	3,747	4,604	8,610	
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869	
6	0,131	0,718	1,134	1,440	1,943	2,447	3,143	3,707	5,959	
7	0,130	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408	
8	0,130	0,706	1,108	1,397	1,860	2,306	2,896	3,355	5,041	

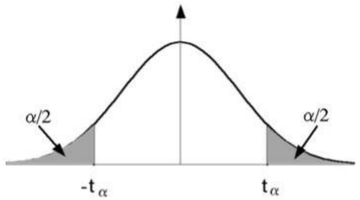
IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Calcul de la statistique

étudiants	1	2	3	4	5	6	7	8	9	
avant	5.5	4.3	6.5	4.5	5.2	4.3	5.0	5.4	5.2	
après	5.4	6.7	6.5	6.0	5.2	5.0	4.8	4.7	4.5	
di	---0.1	2.4	0	1.5	0	0.7	---0.2	---0.7	---0.7	2.9
di^2	0.01	5.76	0	2.25	0	0.49	0.04	0.49	0.49	9.53

- $S^2 = \frac{\sum_{i=1}^{n} di^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n}}{n-1} = \frac{9.53 - \frac{8.41}{9}}{8} = 1.0744$
- m = 2.9 / 9 = 0.32
- t = $\frac{0.32}{\frac{1.036}{\sqrt{9}}}$ =0.93
- p compris entre 0.50 et 0.30

TABLE DU t DE STUDENT



ddl \ α	0,45	0,25	0,15	0,10	0,05	0,025	0,01	0,005	0,0005	TU
ddl \ α	**0,90**	**0,50**	**0,30**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**	**0,001**	**TB**
1	0,158	1,000	1,963	3,078	6,314	12,706	31,821	63,657	636,619	
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,598	
3	0,137	0,765	1,250	1,638	2,353	3,182	4,541	5,841	12,924	
4	0,134	0,741	1,190	1,533	2,132	2,776	3,747	4,604	8,610	
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869	
6	0,131	0,718	1,134	1,440	1,943	2,447	3,143	3,707	5,959	
7	0,130	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408	
8	0,130	0,706	1,108	1,397	1,860	2,306	2,896	3,355	5,041	

- $t = \frac{0.32}{\frac{1.036}{\sqrt{9}}} = 0.93$

IV ---Comparaison de 2 moyennes observées sur 2 échanRllons appariés

- Appliquer les règles de décisions
  - t < 2.306 , on ne rejese pas H0 : la différence de glycémie n'est pas différent de zéro

- Conclusion le stress lié aux examens ne semble pas agir sur la glycémie

Conclusion

- Vérifier les condi0ons d’applica0on des tests:
  - Distribu0on de la variable
  - Effec0fs
  - Indépendance/appariement
- Ques0on à formulée sous la forme d’hypothèses
  - Unilatérale/ bilatérale
- Choisir les bonnes tables sta0s0ques