

DM n°5 - Statistiques déductives

Tutorat 2020-2021 : 16 QRUS



Petit DM sur des QRUS du cours « Statistiques déductives » reprenant quelques points difficiles, ambigus ou mal compris. Certains QRUS ont été inspirés d'annales, de SDR d'années précédentes, vous les avez donc peut-être déjà croisés au détour de vos révisions. J'espère qu'avec ceci, tout sera un peu plus clair. Sinon -> FORUM. N'hésitez pas à me faire un retour sur ce DM svp (sur le forum, ou autre). Des bisous de la meilleure des matières <3

QRU 1 : On souhaite comparer l'efficacité de deux vaccins A et B contre le COVID, produits par des laboratoires pharmaceutiques différents. Pour cela, on vaccine deux groupes d'individus tirés au sort dans la population, l'un par A et l'autre par B, et on note le pourcentage de sujets qui ont contracté le virus. L'attribution des vaccins est faite par TAS. Indiquez la proposition exacte :

Vaccin A	nA = 100	% de sujets ayant contracté le COVID : 24%	IC(95%) = [19,5% ; 36,4%]
Vaccin B	nB = 200	% de sujets ayant contracté le COVID : 19%	IC(95%) = [11,2% ; 23,7%]

- A) La précision des estimations ne dépend pas de la taille des échantillons considérés
- B) On préférera utiliser un test non paramétrique pour comparer l'efficacité des deux vaccins
- C) On peut conclure que les deux vaccins ont la même efficacité au risque de 5%
- D) On peut conclure que le vaccin B est plus efficace que le vaccin A au risque de 5%
- E) Les propositions A, B, C et D sont fausses

QRU 2 : Les hémorragies sont des complications fréquentes chez les patientes ayant subi une césarienne. Un essai thérapeutique randomisé, contrôlé versus placebo et en double aveugle, a été conduit pour évaluer l'efficacité d'un traitement de Desmopressine en prévention de ces complications. Des patientes en bonne santé ont été aléatoirement réparties en deux groupes : groupe « traitement » ou groupe « placebo ». 11 patientes du groupe placebo et 2 patientes du groupe « traitement » ont développé une hémorragie post-césarienne. Après réalisation du test statistique approprié, on obtient $p < 0,001$. Indiquez la proposition exacte :

- A) Le risque de première espèce a été fixé a priori à 0,01
- B) Le nombre de sujets inclus dans l'essai était suffisant pour pouvoir conclure
- C) Si on ne rejette pas H_0 , on se trompe avec une probabilité inférieure à 0,001
- D) Si on rejette H_0 , on se trompe avec une probabilité inférieure à 0,001
- E) Les propositions A, B, C et D sont fausses

QRU 3 : A 20 ans d'intervalle (2000 et 2020), l'INSERM a réalisé une étude portant sur des enfants âgés de 3 à 11 ans. Cette étude s'intéresse à l'IMC (Indice de Masse Corporelle) qui mesure l'excès de poids de ces enfants. Lors des 2 campagnes de mesure à 20 ans d'intervalle, tous les enfants ont été tirés au sort au niveau national. Sont calculés, pour chaque année, la moyenne et l'écart-type des IMC mesurés. On obtient les résultats suivants :

Année d'étude	Effectif	Moyenne IMC (kg/m ²)	Écart-type IMC (kg/m ²)
2000	8000	20	4
2020	10000	21	5

On cherche à savoir si à 20 ans d'écart, les enfants ont un IMC comparable ? Indiquez la proposition exacte :

- A) L'hypothèse H_0 peut se noter : « Les enfants n'ont pas grossi en 20 ans »
- B) On compare deux variables quantitatives
- C) Le résultat du test statistique adapté donne 1,768 : on conclut que les enfants âgés de 3 à 11 ans n'ont pas significativement grossi en 20 ans, au risque de 5%
- D) Avec les écarts-type indiqués, on aboutit à des intervalles de confiance à 95% qui se recouperaient, donc on ne peut pas conclure
- E) Les propositions A, B, C et D sont fausses

QRU 4 : Une équipe de chercheurs niçois découvre une nouvelle molécule révolutionnaire susceptible de traiter les symptômes du COVID. Ils constituent par tirage au sort une cohorte de 24 patients qu'ils divisent en deux groupes : le groupe A avec un effectif de 11 recevra le placebo et le groupe B avec un effectif de 13 recevra la nouvelle molécule. On compte dans chaque groupe le nombre de symptômes du COVID par patients. On comparera donc le nombre de symptômes en fonction du traitement pris. Le paramètre calculé Z est de 2,7. Indiquez la proposition exacte :

- A) On peut utiliser un test de U Mann et Whitney
- B) On peut utiliser un test t de Student pour séries appariées
- C) En utilisant le test t de Student, on conclut à H0
- D) On note : « H1 = Le nouveau traitement est plus efficace que l'ancien »
- E) Les propositions A, B, C et D sont fausses

QRU 5 : Lors de votre stage de sémiologie dans le service de gériatrie de Pasteur, vous voyez 10 patients par jour pendant 1 semaine (hors week-end). Vous avez demandé à chacun de ces patients d'attribuer une note comprise entre 0 à 10 de leur douleur. Puis vous notez leur âge. Vous vous interrogez sur le lien entre âge et la douleur. Indiquez la proposition exacte :

- A) On peut utiliser un test de comparaison de moyennes
- B) On peut extrapoler les résultats de notre étude à l'ensemble des patients du service de gériatrie de Pasteur
- C) On ne peut pas conclure au niveau de notre échantillon
- D) On utilisera un test de corrélation entre les deux données
- E) Les propositions A, B, C et D sont fausses

QRU 6 : On réalise une mesure de la pression artérielle systolique sur des patients atteint du syndrome d'Eisenmenger (*syndrome cardiaque lié à une communication anormale dans le cœur*). Une partie des patients avaient une communication pre-tricuspide et l'autre post-tricuspide. On relève la pression artérielle des patients des deux groupes. On calcule la moyenne des PA systoliques par groupe ainsi que leurs intervalles de confiance :

G1 : Pré-tricuspide $m_1=145\text{mmHg}$ [140 ; 150]

G2 : Post-tricuspide $m_2=142\text{mmHg}$ [138 ; 144]

Le but de l'étude est de déterminer si la localisation de la communication influence la pression artérielle. Indiquez la proposition exacte :

- A) On ne peut pas conclure car les deux intervalles se recoupent
- B) On conclut à H0 car les deux intervalles se recoupent, les PA systoliques ne sont donc pas significativement différentes dans les deux groupes d'Eisenmenger
- C) On conclut au risque de 5% que les deux groupes ont des PA systoliques différentes
- D) On accepte H1 car $m_2 > m_1$
- E) Les propositions A, B, C et D sont fausses

QRU 7 : Une étude visant à prouver l'efficacité d'une nouvelle crème contre l'eczéma est menée. Pour cela, on tire au sort 32 patients parmi une population de personnes atteintes d'eczéma et on répartit les sujets en deux groupes de 15 et 17 patients : le premier recevra le traitement de référence tandis que le deuxième recevra la nouvelle crème. Après analyse des résultats en double aveugle, on obtient un degré de signification $p < 0,001$. À propos de cette étude, indiquez la proposition exacte :

- A) Le degré de signification correspond au risque que l'on commet en acceptant à tort l'hypothèse nulle
- B) L'hypothèse H0 est « La nouvelle crème est plus efficace contre l'eczéma que celle de référence »
- C) Le nombre de sujets inclus dans l'étude était insuffisant pour pouvoir conclure
- D) Il y a moins de 0,1% de risque de se tromper en rejetant H0
- E) Les propositions A, B, C et D sont fausses

QRU 8 : Un groupe de 100 souris atteintes d'adénocarcinome ont été suivies afin de déterminer si la létalité était différente en fonction du stade de la tumeur. Parmi ces 100 souris, 62 avaient un cancer au stade II, tandis que les 48 souris restantes étaient au stade III. On a relevé leur durée de vie en jours en fonction du stade du cancer. Après avoir effectué le test approprié, on conclut à H1. On se demande si le stade de la tumeur influe sur la durée de survie des souris. Indiquez la proposition exacte :

Groupe de souris	Effectif	Durée de survie moyenne en jours
Stade II	62	15
Stade III	48	4

- A) Il s'agit ici de comparer deux variables qualitatives, on utilise donc un test de comparaison de pourcentages
- B) L'hypothèse H0 est : « Le stade du cancer n'influe pas sur la durée de survie des souris »
- C) L'hypothèse H1 est : « Les souris atteintes du cancer en phase III vivent moins longtemps que celles en phase II »
- D) On effectuera un test du U de Mann et Whitney
- E) Les propositions A, B, C et D sont fausses

QRU 9 : A propos du degré de signification p et du risque alpha, indiquez la proposition exacte :

- A) Le risque alpha est trouvé en utilisant les tables statistiques
- B) Le degré de signification p est défini à postériori, et vaut généralement 5%
- C) Lorsque l'on rejette l'hypothèse H₀, on détermine un degré de signification qui peut être différent d'une étude à l'autre
- D) Le risque alpha estime le risque d'accepter H₀ alors qu'elle est fausse
- E) Les propositions A, B, C et D sont fausses

QRU 10 : On cherche à savoir si le fait de porter un soutien-gorge peut accentuer le risque d'avoir un cancer au sein. Pour cela on fait un premier groupe de 100 femmes qui portent un soutien-gorge et un autre de même effectif où elles n'en portent pas. Dans le premier groupe, le pourcentage de femmes développant un cancer du sein est de 33% et dans le deuxième de 25%. Avec le test approprié on trouve un paramètre calculé de 2.68. Indiquez la proposition exacte :

- A) Le test du Khi-2 est le plus approprié pour mettre en évidence la différence entre les deux groupes
- B) A l'issue du test adapté on accepte H₀
- C) A l'issue du test adapté, on rejette H₀ avec une p-value égale de 0,05
- D) A l'issue du test adapté, on rejette H₀ avec un degré de signification inférieur à 1%
- E) Les propositions A, B, C et D sont fausses

QRU 11 : Dans un service hospitalier de 22 patients, en 1 mois on diagnostique 5 infections nosocomiales. Habituellement, le nombre de cas d'infections avec le même nombre de patients est inférieur ou égal à 3. Indiquez la proposition exacte :

- A) On étudie le lien entre des données quantitatives
- B) On ne peut pas utiliser le test de comparaison de pourcentages
- C) Si on n'accepte pas H₀, cela signifie que seul le hasard est responsable de ces chiffres
- D) Si on repousse H₀, cela signifie qu'il y a un problème d'hygiène dans ce service
- E) Les propositions A, B, C et D sont fausses

QRU 12 : Le seuil de 5% signifie que :

- A) La probabilité de se tromper en rejetant l'hypothèse nulle est inférieure à 5%
- B) La probabilité de se tromper en rejetant l'hypothèse nulle est supérieure à 5%
- C) La probabilité de se tromper en rejetant l'hypothèse nulle est inférieure à 95%
- D) La probabilité de se tromper en rejetant l'hypothèse nulle est supérieure à 95%
- E) Les propositions A, B, C et D sont fausses

QRU 13 : On cherche à savoir si le traitement permet de guérir les patients par rapport au placebo.

	Malades	Guéris	Total
Traitement	200 / 100 / 50	300 / 400 / 33,3	500
Placebo	200 / 300 / ?	? / 200 / ?	500
Total	400	600	1000

Les valeurs en gras correspondent aux effectifs théoriques et celles soulignées aux khi 2 partiels.

Indiquez la proposition exacte :

- A) Le Khi-2 partiel des malades sous le placebo est égal à 25
- B) L'effectif théorique manquant est égal à 100
- C) X² total = 166,6
- D) X² total = 142
- E) Les propositions A, B, C et D sont fausses

QRU 14 : On se demande si un nouvel hypotenseur H est plus efficace que le traitement traditionnel T. On réalise un essai thérapeutique en double aveugle où la moitié des patients (tirés au sort) reçoit H, et l'autre reçoit T. Ces deux groupes de patients sont notés H et T. Malheureusement, la pharmacie s'est trompée dans la distribution des médicaments : tous les malades (T et H) ont reçu T, ce dont ni les malades ni les médecins ne se sont aperçus car l'essai est en double aveugle. Le statisticien effectue le test de comparaison entre les groupes H et T. Lui non plus ne connaît pas l'erreur de la pharmacie. Le test est effectué, comme toujours, au risque de première espèce 5 %. La puissance de l'étude a été établie à 90 %. Indiquez la proposition exacte :

- A) Il y a 5 chances sur 100 pour que le test statistique final conduise à déclarer que T et H ont des effets différents
- B) Il y a 95 chances sur 100 pour que le test statistique final conduise à déclarer que T et H ont des effets différents
- C) Il y a 10 chances sur 100 pour que le test statistique final conduise à déclarer que T et H ont des effets différents
- D) Il y a 90 chances sur 100 pour que le test statistique final conduise à déclarer que T et H ont des effets différents
- E) Les propositions A, B, C et D sont fausses

QRU 15 : Si le degré de signification d'un test ou d'une étude est inférieur à 0,01, alors c'est :

- A) Que l'on se trompe avec une probabilité inférieure à 0,01 si on ne rejette pas H_0
- B) Que même si le risque de première espèce avait été choisi à 0,01, H_0 aurait été rejetée à ce risque
- C) Que la probabilité qu'avait, sous H_0 , le paramètre du test d'indiquer un écart à H_0 plus grand que ne l'a fait la valeur observée du paramètre est supérieure à 0,01
- D) Que le risque de première espèce avait été choisi à 0,01
- E) Les propositions A, B, C et D sont fausses

QRU 16 : On étudie la relation entre le nombre d'heures hebdomadaires passées à réviser la P2 et la note aux partiels, les variables étant présentées comme suit. On note le nombre d'étudiants de la promo concernés pour chaque modalité (certains étudiants n'ayant pas souhaité participer à cette étude). Indiquez la proposition exacte :

Note aux partiels / Nb d'h	[0 ;10]]10 ;20[[20 ;30[
[11 ;20]	30	10	20
[0 ;10]	10	20	30

- A) L'effectif attendu des personnes ayant travaillé entre 0 et 10h hebdomadaires et ayant obtenu une note comprise entre 11 et 20, est égal à 15
- B) L'effectif attendu des personnes ayant travaillé entre 10 et 20h hebdomadaires et ayant obtenu une note comprise entre 0 et 10, est égal à 20
- C) Le Khi-2 partiel des personnes ayant travaillé entre 20 et 30h hebdomadaires et ayant obtenu une note comprise entre 0 et 10, est égal à 1,25
- D) Le Khi-2 partiel des personnes ayant travaillé entre 20 et 30h hebdomadaires et ayant obtenu une note comprise entre 0 et 10, est égal à 1
- E) Les propositions A, B, C et D sont fausses

Annexes :

Table t de Student

α ddl	0,90	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	1,000	1,963	3,078	6,314	12,706	31,821	63,656	636,578
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,600
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	3,768
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
26	0,127	0,684	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,684	1,057	1,314	1,703	2,052	2,473	2,771	3,689
28	0,127	0,683	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,683	1,055	1,311	1,699	2,045	2,462	2,756	3,660
30	0,127	0,683	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,681	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,678	1,043	1,292	1,664	1,990	2,374	2,639	3,416
120	0,126	0,677	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,675	1,037	1,282	1,645	1,960	2,327	2,577	3,293

CORRECTION

QRU 1 : C

- A) Faux : Plus les effectifs des échantillons sont importants, plus l'estimation sera précise
- B) Faux : On va utiliser un test paramétrique (test de comparaison de pourcentages) au vu des effectifs, ce qui sera bien plus adéquat ici qu'un test non paramétrique, s'utilisant plutôt sur des effectifs faibles
- C) Vrai : Les 2 IC se recouvrent, donc il n'existe pas de différence significative entre les 2 pourcentages, on accepte alors H_0
- D) Faux : On ne peut pas conclure ça car les 2 IC se recoupent (voir item C)
- E) Faux

QRU 2 : D

- A) Faux : à 0,05, comme d'habitude
- B) Faux : Il manque les effectifs des 2 groupes pour pouvoir répondre à cet item
- C) Faux : voir item D
- D) Vrai
- E) Faux

QRU 3 : E

- A) Faux : formulation réductrice et ambiguë, les enfants auraient pu maigrir. On préfère noter H_0 = il n'y a pas de différence d'IMC entre les enfants de 2000 et ceux de 2020
- B) Faux : une qualitative (année d'étude) et une quantitative (moyenne d'IMC)
- C) Faux : on a paramètre calculé < paramètre théorique (1,96) donc on accepte H_0
- D) Faux : on conclut à l'acceptation de H_0 cad la non-différence significative
- E) Vrai

QRU 4 : A

- A) Vrai : on peut utiliser un test de U Mann et Whitney, car il s'utilise également pour des effectifs supérieurs à 12 étant donné qu'il est robuste car non paramétrique
- B) Faux : les séries ne sont pas appariées
- C) Faux : on peut utiliser un test t de Student ici car la limite inférieure de 12 est artificielle, et ici avec ces deux effectifs, un test t de Student est tout à fait réalisable et souhaitable. $DDL = (n_1-1) + (n_2-1) = 10 + 12 = 22$. Pour $p=5\%$ on a 2,074. Donc, t calculé > t théorique, on rejette H_0 . Pour un risque d'erreur plus faible, on s'aperçoit qu'à $p = 0,01$ on a 2,819 > 2,7, donc on conclura au risque de 2%, ce qui revient à garder un degré de signification à 5% (en général, on s'intéresse uniquement à 5% ou 1%, ou 0,1% par ex).
- D) Faux : H_1 = il y a une différence significative entre les résultats de la molécule et le placebo
- E) Vrai

QRU 5 : D

- A) Vrai : On ne peut pas utiliser la comparaison de moyennes car on est en présence d'une variable pseudo-quantitative, à n classes, et on n'a pas le droit d'effectuer le calcul de moyennes
- B) Faux : On ne peut pas extrapoler les résultats car il n'y a pas eu de TAS des patients dans le service ++ on reste au niveau de l'échantillon uniquement
- C) Faux : Si (étant donné qu'il n'y a pas eu de gros biais de sélection, une conclusion au niveau de l'échantillon sera possible), mais au niveau de l'échantillon seulement ++
- D) Vrai : En considérant le score de douleur comme une variable pseudo-quantitative, il est tout à fait possible de réaliser un test de corrélation (avec l'âge comme autre variable quantitative)
- E) Faux

QRU 6 : B

- A) Faux : On peut tout à fait conclure
- B) Vrai : On accepte H_0 lorsque les 2 intervalles se recoupent
- C) Faux
- D) Faux : Il faut regarder les intervalles de confiance, et faire le bon test
- E) Faux

QRU 7 : D

- A) Faux : En rejetant à tort H_0
- B) Faux : H_0 = il n'y a pas de différence significative entre les 2 groupes
- C) Faux : il était suffisant
- D) Vrai ++
- E) Faux

QRU 8 : B

- A) Faux : On compare une variable qualitative (stade du cancer) et une variable quantitative (durée de survie en jours)
B) Vrai
C) Faux : Lors de la formulation des hypothèses, on ne peut pas prédire quel groupe sera « meilleur » que l'autre
D) Faux : Non, car les effectifs sont trop élevés. On pourrait l'utiliser, mais il vaut mieux privilégier le test de comparaison de moyennes
E) Faux

QRU 9 : C

- A) Faux : Non, il est défini à priori, c'est p qui est trouvé à postériori dans les tables
B) Faux : c'est alpha qui vaut généralement 5%
C) Vrai
D) Faux : C'est le risque de se tromper en rejetant à tort H0
E) Faux

QRU 10 : D

- A) Faux : Le test de comparaison de pourcentages est le plus approprié même si on PEUT faire un Khi-2
B) Faux : $Z_c > Z_t = 1,96$, donc on rejette H0
C) Faux : on rejette H0 mais $2,68 > 2,6$ donc $p < 1\%$
D) Vrai
E) Faux

QRU 11 : D

- A) Faux : qualitatives (mois à 5 infections = mois actuel / ou les autres mois)
B) Faux : on fait un test de comparaison de pourcentages (comparaison de % de patients infectés)
C) Faux : si on accepte H0
D) Vrai
E) Faux

QRU 12 : A

- A) Vrai : Si on veut valider une hypothèse, fixer un seuil de signification de 5% signifie que la probabilité de se tromper en rejetant cette hypothèse est inférieure à 5%. Choisir un seuil de signification inférieur à la p value conduit à accepter l'hypothèse, alors que le choisir supérieur à la valeur p conduit à rejeter l'hypothèse. La valeur p est donc la valeur limite du seuil de signification à partir duquel l'observation conduit à rejeter l'hypothèse.
B) Faux
C) Faux
D) Faux
E) Faux

QRU 13 : C

- A) Faux : $(300 - 200)^2 / 200 = 50$
B) Faux : $ET = (600 \times 500) / 1000 = 300$
C) Vrai : $\text{Khi 2 manquant} = (200 - 300)^2 / 300 = 33,3$
et $\text{Khi-2 total} = 50 + 50 + 33,3 + 33,3 = 166,6$
D) Faux
E) Faux

QRU 14 : A

- A) Vrai
B) Faux
C) Faux
D) Faux
E) Faux

QRU 15 : B

- A) Faux : si on rejette H0
B) Vrai
C) Faux : inférieure à 0,01
D) Faux : pas du tout, et généralement, alpha est fixé à 5%
E) Faux

QRU 16 : D

A) Faux = $30 \times 60 / 120 = 20$

B) Faux = $30 \times 60 / 120 = 15$

C) Faux ET = $50 \times 60 / 120 = 25$

Khi2 partiel = $(25-20)^2 / ET = 25/25 = 1$

D) Vrai (voir item C)

E) Faux