

Jean-Claude Régnier

Méthodes quantitaves

Année 2002/2003-07-19 Université Lyon 2

IRSP Lausanne



Situation Problème

Étude des difficultés et des facilités d'apprentissage de la statistique au travers des méthodes quantitatives et qualitatives dans une formation à distance en vue de l'obtention d'une licence de sciences de l'éducation.

Document n°1 :***Situation d'étude à partir d'une enquête par questionnaire.***

Remarque : Ce questionnaire est exactement celui qui a été soumis aux étudiants présents à l'épreuve de méthodes quantitatives de la session 2002 —Université Lyon 2 — groupe des étudiants préparant à distance la licence de sciences de l'éducation. Nous n'avons que réduit les espaces pour recueillir les réponses textuelles pour l'intégrer à ce document. Ce questionnaire dans sa forme initiale est présenté sous la forme d'un cahier de 8 pages au format A4.

Questionnaire aux étudiants

L'objet de cette enquête par questionnaire est de construire un corpus de données à partir duquel nous étudierons la question de la formation à la statistique et nous chercherons à dégager des éléments permettant une régulation de l'approche pédagogique proposée dans le dispositif de la formation à distance dont vous êtes le sujet [apprenant] central.

Ces informations viennent compléter celles que fournissent votre copie d'examen et le dossier méthodologique. C'est pourquoi ce questionnaire ne peut être anonyme. Cependant il relève d'un travail de recherche et les réponses fournies ne peuvent en aucun cas intervenir dans l'évaluation pour l'examen final. L'anonymat sera en revanche préservé dans l'analyse et dans toute communication comme le requiert la déontologie de l'enquête.

Je vous remercie de votre coopération en tant qu'informateur, sans laquelle les informations nécessaires ne peuvent être obtenues. Le questionnaire est à me remettre personnellement.

Jean-Claude Régnier

Vous et votre parcours de formation antérieur à la licence SE

Nom patronymique :	Prénom :	[V1] Date de naissance :
Nom marital :	N° étudiant	Codage pour anonymat :

[V2] Situation professionnelle actuelle :

[V3] Comporte-t-elle des activités qui requièrent des connaissances en statistique ? :

(Entourer votre réponse)

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V4] Si des connaissances en statistique sont requises, même rarement, cela vous pose-t-il des problèmes ? : (Entourer votre réponse)

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V5] Baccalauréat : Série.....

[V6] Année d'obtention :

[V7] Vos diplômes universitaires :

	Nom	Année	Obtenu = O Non obtenu = N
Diplômes comportant de la statistique sous forme de cours, d'usage, etc.	-	-	-
	-	-	-
Diplômes ne comportant aucun cours de statistique ni aucune référence au domaine de la statistique.	-	-	-
	-	-	-
	-	-	-

[V8] Année d'entrée en licence SE (*Entourer votre réponse*) : sept. 2000 sept. 2001

[V9] Avez-vous préparé l'UE_C durant l'année 2000/2001 ? (*Entourer*): oui non

[V10] Avez-vous passé les examens de l'UE_C à la fin de l'année 2000/2001 ? :

Session 1 [mai 2001] :	oui	non	Session 2 [juin 2001] :	oui	non
note de méthodes quantitatives :	/20		note de méthodes quantitatives :	/20	
note de méthodes qualitatives :	/20		note de méthodes qualitatives :	/20	

[V11] Vos diplômes professionnels dont la préparation a requis des connaissances du domaine de la statistique :

Nom	Année	Obtenu = O Non obtenu = N
-	-	-

Vous et la statistique

[V12] A l'entrée dans cette formation en licence, qu'évoquait pour vous la statistique ?

(Donner trois mots)

--	--	--

(Compléter par une phrase)

--

[V13] A ce jour qu'évoque pour vous la statistique ?

(Donner trois mots)

--	--	--

(Compléter par une phrase)

--

[V14] A ce jour que signifient pour vous les termes suivants ?

Représentatif	
Significatif	
Estimation	
Test d'hypothèses	
Probabilité	
Sondage	

[V15] A ce jour que pensez-vous des sondages d'opinion ?

--

[V16] Dans la vie sociale actuelle, quelle utilité attribuez-vous à la statistique ?

--

Vous et l'approche pédagogique adoptée dans le dispositif de formation à la licence SE

[V17] Parmi les notions visées par le cours de méthodes quantitatives et méthodes qualitatives, quelles sont celles que vous pensez avoir (*citer au plus six*)

...assez bien comprises	... plutôt mal comprises

Dossier méthodologique

[V18] Quel niveau d'efficacité attribuez-vous à la réalisation du *dossier méthodologique*, comme médiateur d'apprentissage (*Entourer votre réponse*) ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V19] Quel niveau d'intérêt attribuez-vous à la réalisation du *dossier méthodologique* ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V20] Durée estimée (h) pour la réalisation du *dossier méthodologique* : entre ___ et ___ h

[V21] Quelles sont les trois **difficultés majeures** que vous avez dues affronter pour réaliser ce *dossier méthodologique* ?

Diff_1 :	Diff_2 :	Diff_3 :

[V22] Quels sont les trois **apports majeurs** de la réalisation du *dossier méthodologique* ?

App_1 :	App_2 :	App_3 :
---------	---------	---------

[V23] Quelles sont les trois **suggestions majeures** que vous feriez pour faciliter la réalisation du *dossier méthodologique* dans les conditions de la préparation de la licence ?

Sugg_1 :	Sugg_2 :	Sugg_3 :
----------	----------	----------

Liste de diffusion

[V24] Avez-vous eu recours aux échanges par courrier électronique sur la liste de diffusion Lyon2 pour l'étude des cours de méthodes ? (*Entourer votre réponse*)

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V25] Si oui , qu'est-ce que cela vous a apporté ?

--

Support du cours de méthodes qualitatives

[V26] Quel niveau d'efficacité attribuez-vous au **support** de cours de méthodes **qualitatives**, comme médiateur d'apprentissage ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V27] Quel niveau d'intérêt attribuez-vous au **support** de cours de méthodes **qualitatives** ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V28] Durée estimée (h) pour étudier les méthodes **qualitatives** : entre ___ et ___ h

EEE du cours de méthodes qualitatives

[V29] Avez-vous réalisé l'EEE de méthodes qualitatives ? non oui [note .../20]

[V30] Quel niveau d'efficacité attribuez-vous à l'EEE du cours de méthodes **qualitatives**, comme médiateur d'apprentissage ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V31] Quel niveau d'intérêt attribuez-vous à l'EEE de méthodes qualitatives ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V32] Durée estimée (h) pour réaliser à l'EEE de méthodes qualitatives : entre ___ et ___ h

Support du cours de méthodes quantitatives

[V33] Quel niveau d'efficacité attribuez-vous au **support** de cours de méthodes **quantitatives**, comme médiateur d'apprentissage ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V34] Quel niveau d'intérêt attribuez-vous au **support** de cours de méthodes **quantitatives** ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V35] Durée estimée (h) pour étudier les méthodes **quantitatives** : entre ___ et ___ h

EEE du cours de méthodes quantitatives

[V36] Avez-vous réalisé l'EEE de méthodes quantitatives ? non oui [note .../20]

[V37] Quel niveau d'efficacité attribuez-vous à l'EEE du cours de méthodes **quantitatives**, comme médiateur d'apprentissage ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V38] Quel niveau d'intérêt attribuez-vous à l'EEE de méthodes **quantitatives** ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V39] Durée estimée (h) pour réaliser à l'EEE de méthodes **quantitatives** :

entre ___ et ___ h

Usage des ressources diverses pour l'étude des méthodes ?

[V40] Quelles ressources avez-vous utilisées en plus du cours distribué par le CNED ?

(citer au plus 5)

--

[V41] Pensez-vous qu'il est nécessaire d'aller chercher dans d'autres ressources que le support de cours de **méthodes qualitatives** ?

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V42] Pensez-vous qu'il est nécessaire d'aller chercher dans d'autres ressources que le support de cours de **méthodes quantitatives** ?

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V43] De manière générale pour le cours de méthodes, que pensez-vous de la nécessité de compléter le support de cours et l'EEE par d'autres ressources ?

--

Formation en statistique

[V44] Que pensez-vous avoir appris durant cette année universitaire en ce qui concerne le domaine de la statistique ?

--

[V45] Les 10 objectifs énoncés sont ceux du cours de méthodes quantitatives : ranger les par ordre décroissant d'importance [1 = le plus important à 10 = le moins important]

Énoncés des objectifs	Rang attribué
1) d'expliciter les questions d'une problématique dont les réponses relèvent d'une approche statistique,	
2) de décrire, traiter, analyser des données de manière pertinente dans le cadre d'une étude en particulier dans le domaine éducatif,	
3) de faire le lien entre la réflexion analytique sur des questions relevant du champ de l'éducation, leur formalisation et leur traitement quantitatif,	
4) de lire avec un regard critique et distancié, les conclusions de diverses études statistiques apparaissant dans des rapports de recherche en Sciences de l'Education,	
5) de poursuivre de façon autonome et personnalisée un apprentissage en statistique afin d'enrichir ses acquis,	
6) de poser un regard plus positif à l'égard d'un domaine largement exploité dans les <i>media</i> , dans le sens de ne pas considérer les résultats dans l'ordre du tout ou rien mais en les replaçant judicieusement dans leur domaine de validité,	
7) d'exploiter des notions et des démarches mathématiques à des fins d'outils, et de ce fait de modifier dans un sens positif le rapport souvent négatif que nombre entretient avec cette science,	
8) de s'exercer à un raisonnement intégrant l'idée de "risque d'erreur" dans l'énoncé de ses conclusions.	
9) de s'exercer à l'interprétation de phénomènes éducatifs sur la base de données statistiques sur des "faits éducatifs" et sur des relations entre ces "faits"	
10) de s'exercer à la communication des résultats des analyses des données en distinguant clairement le modèle utilisé, de la réalité qu'il est supposé représenter, en séparant bien les traitements menés à l'intérieur du modèle, des interprétations reformulées dans le contexte du problème.	

Vos suggestions

Le questionnaire précédent donc a été soumis à 132 étudiants ayant passé les épreuves d'évaluation du cours de méthodes quantitatives lors de la session 2002 d'examen de licence de sciences de l'éducation, préparation à distance. 297 étudiants sont inscrits dans cette licence. 86 étudiants ont rendu ce questionnaire. Nous n'avons rapporté ici qu'une partie des questions. Nous n'étudierons que quelques propriétés dans le cadre actuel.

Nous rapportons ci-dessous les tableaux statistiques des variables :

(Tableau n° T1) Tableau statistique de la variable V00 = sexe

Ensemble des inscrits en licence SE préparation à distance			étudiants ayant passé au moins une épreuve de l'UE_C Méthodologie°		
Modalités de V00	homme	femme		homme	femme
effectifs	49	248		23	109

(Tableau n° T2) Tableau statistique de la variable V01 = âge à la date de l'épreuve

Valeurs de V01	[20 ; 25[[25 ; 30[[30 ; 35[[35 ; 40[[40 ; 45[[45 ; 50[[50 ; 55[
effectifs	22	19	22	8	7	5	3

(Tableau n° T4) Tableau statistique de la variable V03 : statistique dans l'activité professionnelle

Modalités de V03	1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
effectifs	48	15	18	4	0	0

(Tableau n° T5) Tableau statistique de la variable V18 = efficacité du dossier méthodologique

modalités de V18	1 le plus faible	2	3	4	5	6	7	8	9	10 le plus fort
effectifs	1	0	2	3	9	9	24	18	9	9

(Tableau n° T6) Tableau statistique de la variable V20 = durée pour l'élaboration du dossier

Valeurs de V20	[0 ; 50[[50 ; 100[[100 ; 150[[150 ; 200[[200 ; 250[
effectifs	22	28	13	3	6
Valeurs de V20	[250 ; 300[[300 ; 350[[350 ; 400[[400 ; 450[[450 ; 675[
effectifs	2	0	0	1	3

(Tableau n° T7) Tableau statistique de la variable V24 : recours à liste de diffusion

Modalités de V24	1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
effectifs	39	12	20	4	3	4

(Tableau n° T8) Tableau statistique de la variable V47B= note à l'examen

notes/20	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
effectifs	0	1	0	2	8	4	6	7	1	3	5	4	7	9	9	6	2	5	2	4	0

Document n°2 : traitement demandé

TQ1- De quoi s'agit-il ?

Q101- Après avoir pris connaissance des informations fournies dans le **document n°1**, identifiez et caractérisez "succinctement" mais clairement, les éléments du modèle statistique : **population, échantillon, individus, variables** [V1, V2, V3, V4, V5, V6, V12, V17, V18, V19, V20, V21, V22, V23, V24, V25, V44, V45]

Q102- Sur l'une des variables **qualitatives** de votre choix, réalisez une courte synthèse s'appuyant sur un graphique de votre choix afin de rendre compte des informations apportées.

Q103 Comment traiteriez-vous les réponses aux questions V12, V17, V44 ?

Q104 Quelle est votre interprétation du tableau T1?

Q105 Quelle(s) remarque(s) feriez-vous à propos du questionnaire ?

Nous vous rapportons *in extenso* l'interprétation fournie par une étudiante (F011) à laquelle cette question a été posée : « *Le tableau T1 correspond à la variable du sexe, c'est à dire la proportion d'homme et celle de femme inscrite de sciences de l'éducation. Pour ces premiers chiffres : je fais l'analyse que les femmes sont cinq fois plus nombreuses que les hommes en licence de sciences de l'éducation. De plus parmi les étudiants qui ont passé au moins une épreuve de l'UE de méthodologie, le nombre d'hommes est plus élevé proportionnellement à la population des hommes inscrits. En effet il y a environ 47% des hommes inscrits qui ont au moins une épreuve de l'UE de méthodologie. Et pour les femmes elles sont environ 44% à avoir au moins une épreuve de l'UE de méthodologie en fonction du nombre total d'inscription. Je peux en conclure que la réussite d'au moins une épreuve de l'UE de méthodologie est plus importante chez les hommes proportionnellement au nombre d'inscriptions (49).* »

TQ2- Étude de la variable V47B = note sur 20 obtenue à l'épreuve finale du cours de méthodes quantitatives (tableau T8)

Q201 Quelle est la valeur de la **note moyenne m** et celle de l'**écart-type s** échantillon ?

Q202 Quelle est la (les) **note(s) modale(s)** ?

Q203 Quelle est la **note médiane** ?

Q204 Regrouper les notes en trois classes Faible = [0 ; 8], Moyen = [9 ; 11] et Fort = [12 ; 20] de niveau de réussite et construire un diagramme de la distribution des fréquences de ces trois classes.

Q205- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

TQ3-Etude de la variable V20 = durée déclarée pour l'élaboration du dossier méthodologique

Q301 Quelle est la valeur de la note moyenne m et celle de l'écart-type $\sigma_{\text{échantillon}}$?

Q302 Construire l'histogramme de la distribution des fréquences de V20

Q303 Construire le polygone des fréquences à partir de l'histogramme

Q304 À partir du polygone des fréquences, donner une estimation de la (des) durée(s) modale(s)

Q305 À partir de l'histogramme des fréquences, donner une estimation de la durée médiane.

Q306 Donner une estimation ponctuelle s^2 de la variance de V20 sur la population à partir de $\sigma_{\text{échantillon}}^2$.

Q307 Quelle est la proportion des étudiants ayant déclaré une durée comprise l'intervalle $[m - s ; m + s]$?

Q308 Quelle serait la proportion si la variable V20 était une variable continue dont la distribution de fréquences est celle d'une variable de Laplace-Gauss de paramètres m et s ?

Q309- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

TQ4- Étude du lien entre le jugement d'efficacité du dossier méthodologique et le niveau de réussite à l'épreuve d'examen

Q401- A partir du sous-échantillon des 83 répondants, peut-on considérer qu'il y a une dépendance entre le jugement d'efficacité et le niveau de réussite à l'examen ?

Nous avons réduit le nombre de modalités de la variable V18 à trois niveaux.

		Niveau d'efficacité Dossier		
		N1 Faible 1,2,3	N2 Moyen 4,5,6,7	N3 Fort 8,9,10
Niveau de réussite à l'épreuve finale	Faible	1	15	11
	Moyen	1	5	6
	Fort	1	25	18

Testez les hypothèses correspondantes (H_0 et H_1) au seuil de 5% à l'aide du test d'indépendance du Khi-deux.

Q402- A partir de cette étude, énoncez sous une forme synthétique votre **conclusion** et votre **interprétation**.

TQ5- Étude du lien entre le niveau d'intérêt pour réaliser le dossier méthodologique et la note finale à l'épreuve de méthodes quantitatives.

Nous considérons ici V18c, la variable niveau d'intérêt pour la réalisation du dossier méthodologique, comme une variable quantitative discrète prenant ses valeurs sur $\{1; \dots; 10\}$

Le tableau ci-dessous est le tableau statistique de la distribution conjointe du couple de variables V18c et V47B.

Q501- Analyser le tableau ci-dessous.

Tableau de contingence		V47B notes à l'épreuve finale de méthodes quantitatives																				TOT.	
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		20
Niveau d'intérêt pour la réalisation du Dossier méthodologique	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	
	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	
	5	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	3
	6	0	0	0	2	1	0	2	0	0	0	0	0	0	1	1	1	0	1	0	0	0	9
	7	0	0	1	1	0	1	1	0	0	1	1	2	1	2	0	0	1	0	0	0	0	12
	8	0	0	1	1	2	2	1	0	0	3	0	2	2	3	3	1	0	1	0	0	0	22
	9	1	0	0	1	1	0	1	0	1	1	0	2	5	1	0	0	2	1	1	0	0	18
	10	0	0	0	2	0	1	1	1	1	0	3	1	1	1	2	0	1	0	1	0	0	16
TOT.	1	0	2	7	4	5	7	1	3	5	4	7	9	9	6	2	5	2	4	0	0	83	

Q502- Construire le nuage statistique.

Q503- Calculer le coefficient de corrélation linéaire de Bravais-Pearson

Q504- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

TQ6- Étude de l'intérêt personnel pour les TIC en formation

Une question concernant l'intérêt personnel pour les TICE a été posée sous le format suivant : *Quel est votre degré d'intérêt personnel relatif à l'intégration des TIC en formation ? (cocher la case correspondante)*

Très faible (f1) Plutôt faible (f2) Plutôt fort (F3) Très fort (F4) Sans avis (SA5)

(Tableau n° T9) Tableau statistique de la variable V07

Modalités de V09	f1	f2	F3	F4	SA5
Effectifs	15	13	27	15	10

Q601- A partir des informations obtenues sur l'échantillon et consignées dans le tableau T9, peut-on considérer que les points de vue sont partagés ? (il s'agit de tester l'hypothèse H_0 contre H_1 à l'aide du test du Khi-deux d'adéquation avec un niveau de risque maximum de 5%)

Q602- A partir de cette étude, énoncer sous une forme synthétique **vosre conclusion** et **vosre interprétation**.

GRILLE D'AUTO-ÉVALUATION

code	items repérés et sur lesquels le jugement est porté	validité du résultat				validité du raisonnement			
		R+	R-	E	N	R+	R-	E	N
Q101	- Analyse "succincte" de l'étude statistique	R+	R-	E	N	R+	R-	E	N
Q102	- Synthèse relative à une variable qualitative et diagramme	R+	R-	E	N	R+	R-	E	N
Q103	- Traitement V12, V17, V44	R+	R-	E	N	R+	R-	E	N
Q104	- Interprétation Tableau T1	R+	R-	E	N	R+	R-	E	N
Q105	- Remarque(s) sur le questionnaire	R+	R-	E	N	R+	R-	E	N
Q201.1	- Moyenne de V47B	R+	R-	E	N	R+	R-	E	N
Q201.2	- Variance, écart-type de V47B	R+	R-	E	N	R+	R-	E	N
Q202	- Mode(s)	R+	R-	E	N	R+	R-	E	N
Q203	- Médiane	R+	R-	E	N	R+	R-	E	N
Q204.1	- Tableau V47B regroupée	R+	R-	E	N	R+	R-	E	N
Q204.2	- Diagramme V47B regroupée	R+	R-	E	N	R+	R-	E	N
Q205	- Conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N
Q301.1	- Moyenne de V20	R+	R-	E	N	R+	R-	E	N
Q301.2	- Variance, écart-type de V20	R+	R-	E	N	R+	R-	E	N
Q302	- Histogramme	R+	R-	E	N	R+	R-	E	N
Q303	- Polygone (lissage)	R+	R-	E	N	R+	R-	E	N
Q304	- Mode(s)	R+	R-	E	N	R+	R-	E	N
Q305	- Médiane	R+	R-	E	N	R+	R-	E	N
Q306	- Estimation s^2	R+	R-	E	N	R+	R-	E	N
Q307	- Proportion dans $[m-s ; m+s]$	R+	R-	E	N	R+	R-	E	N
Q308	- Proportion dans $[m-s ; m+s]$ sous $H_0 = LG$	R+	R-	E	N	R+	R-	E	N
Q309	- Conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N
Q401.1	- Énoncé explicite de H_0 et H_1	R+	R-	E	N	R+	R-	E	N
Q401.2	- Mise en place du test (explicitation théorique, valeur critique au seuil 5%)	R+	R-	E	N	R+	R-	E	N
Q401.3	- Mise en œuvre du test (tableau des effectifs théoriques)	R+	R-	E	N	R+	R-	E	N
Q401.4	- Mise en œuvre du test (calcul du Khi-Deux)	R+	R-	E	N	R+	R-	E	N
Q402	- Conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N
Q501	- Analyse du tableau	R+	R-	E	N	R+	R-	E	N
Q502	- Nuage statistique	R+	R-	E	N	R+	R-	E	N
Q503	- Coefficient de corrélation R_{BP}	R+	R-	E	N	R+	R-	E	N
Q504	- Conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N
Q601.1	- Énoncé explicite de H_0 et H_1	R+	R-	E	N	R+	R-	E	N
Q601.2	- Mise en place du test (explicitation théorique, valeur critique au seuil 5%)	R+	R-	E	N	R+	R-	E	N
Q601.3	- Mise en œuvre du test (tableau des effectifs théoriques)	R+	R-	E	N	R+	R-	E	N
Q601.4	- Mise en œuvre du test (calcul du Khi-Deux)	R+	R-	E	N	R+	R-	E	N
Q602	- Conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N

Calcul du score

	Tous les items repérés et sur lesquels le jugement est porté	validité du résultat				validité du raisonnement			
		R+	R-	E	N	R+	R-	E	N
	modalités du jugement								
S01	nombre de...								
	résultat & raisonnement	R+	1	R-	0,5	E	0	N	0
S02	nombre de...								
S03	scores...					0		0	

$$\text{Score total obtenu } S = (0,5) \times n_{R-} + (1) \times n_{R+} =$$

$$\text{Score maximum total } S_{\max} = 36 + 36 = 72$$

$$\text{Taux de réussite } T = \frac{S}{S_{\max}} = \boxed{} \quad \text{Niveau de réussite : note} = \boxed{} / 20$$

Explicitation du codage utilisé

code des modalités du jugement	modalités du jugement	Caractérisation	pondération
R+	réussite forte	- lorsque le résultat est celui attendu par l'évaluateur - lorsque le raisonnement est logiquement vrai	+1
R-	réussite faible	Il s'agit de repérer le déjà-réussi sans laisser la place à l'ancrage de sous-compréhension. - lorsque le résultat n'est pas celui attendu par l'évaluateur sans erreur manifeste de raisonnement - lorsque le raisonnement ne comporte que des erreurs "d'esprit fatigué"	+0,5
E	échec par erreur	- lorsque le résultat est faux mathématiquement - lorsque le raisonnement est logiquement faux	0
N	échec par non-réponse	- lorsque le résultat n'est pas fourni - lorsque le raisonnement n'est pas explicité - lorsque le résultat ou le raisonnement ne sont pas interprétables par l'évaluateur	0

Évaluation des items "CONCLUSION-INTERPRETATION"

Conformément aux objectifs visés en cours, l'item "conclusion" correspond à deux niveaux :

- *niveau 1* (représenté par le point de vue "résultat") correspond à une **description** lisible et "bien" rédigée des résultats obtenus dans les traitements abordés. Ces résultats sont rapportés dans le contexte de la situation problème. (Par exemple : les nombres sont donnés avec un ordre de grandeur compatible avec le contexte, avec leurs unités, ou les modalités sont explicitées en clair...)
- *niveau 2* (représenté par le point de vue "raisonnement") correspond à l'énoncé d'une **interprétation** personnelle fondée sur les résultats de l'étude et exprimée de façon pertinente.

Document autocorrectif et auto-évaluatif

Avertissement

Ce corrigé-guide a pour fonction de vous aider à réaliser un travail d'autocorrection et d'auto-évaluation a posteriori. Il convient d'essayer de comprendre les erreurs que vous aurez pu commettre.

Le niveau de précision de ce corrigé-guide peut ne pas vous satisfaire, car votre niveau de compréhension actuel requiert un niveau de détail plus élevé. Ne vous découragez pas et utilisez le cours, les échanges par la liste de diffusion ou tout autre forme d'aide qui peut faciliter le développement votre compréhension.

Jean-Claude Régnier



Situation Problème

Étude des difficultés et des facilités d'apprentissage de la statistique au travers des méthodes quantitatives et qualitatives dans une formation à distance en vue de l'obtention d'une licence de sciences de l'éducation.

Document n°1 :**Situation d'étude à partir d'une enquête par questionnaire.**

Nous ne rapportons que l'extrait du questionnaire correspondant aux questions abordées dans le traitement.

Vous et votre parcours de formation antérieur à la licence SE

Nom patronymique :	Prénom :	[V1] Date de naissance :
Nom marital :	N° étudiant	Codage pour anonymat :

[V2] Situation professionnelle actuelle :

[V3] Comporte-t-elle des activités qui requièrent des connaissances en statistique ? :

(Entourer votre réponse)

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V4] Si des connaissances en statistique sont requises, même rarement, cela vous pose-t-il des problèmes ? : (Entourer votre réponse)

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

[V5] Baccalauréat : Série..... [V6] Année d'obtention :

Vous et la statistique

[V12] A l'entrée dans cette formation en licence, qu'évoquait pour vous la statistique ?

(Donner trois mots)

--	--	--

(Compléter par une phrase)

--

Vous et l'approche pédagogique adoptée dans le dispositif de formation à la licence SE

[V17] Parmi les notions visées par le cours de méthodes quantitatives et méthodes qualitatives, quelles sont celles que vous pensez avoir (citer au plus six)

...assez bien comprises	... plutôt mal comprises

Dossier méthodologique

[V18] Quel niveau d'efficacité attribuez-vous à la réalisation du dossier méthodologique, comme médiateur d'apprentissage (Entourer votre réponse) ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V19] Quel niveau d'intérêt attribuez-vous à la réalisation du dossier méthodologique ?

1	2	3	4	5	6	7	8	9	10
le plus faible									le plus fort

[V20] Durée estimée (h) pour la réalisation du dossier méthodologique : entre ___ et ___ h

[V21] Quelles sont les trois **difficultés majeures** que vous avez dues affronter pour réaliser ce *dossier méthodologique* ?

Diff_1 :	Diff_2 :	Diff_3 :
----------	----------	----------

[V22] Quels sont les trois **apports majeurs** de la réalisation du *dossier méthodologique* ?

App_1 :	App_2 :	App_3 :
---------	---------	---------

[V23] Quelles sont les trois **suggestions majeures** que vous feriez pour faciliter la réalisation du *dossier méthodologique* dans les conditions de la préparation de la licence ?

Sugg_1 :	Sugg_2 :	Sugg_3 :
----------	----------	----------

Liste de diffusion

[V24] Avez-vous eu recours aux échanges par courrier électronique sur la liste de diffusion Lyon2 pour l'étude des cours de méthodes ? (*Entourer votre réponse*)

1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
---------------	------------	---------------	-----------	----------------	------------

V25] Si oui , qu'est-ce que cela vous a apporté ?

--

Formation en statistique

[V44] Que pensez-vous avoir appris durant cette année universitaire en ce qui concerne le domaine de la statistique ?

--

[V45] Les 10 objectifs énoncés sont ceux du cours de méthodes quantitatives : ranger les par ordre décroissant d'importance [1 = le plus important à 10 = le moins important]

<i>Énoncés des objectifs</i>	<i>Rang attribué</i>
11) d'expliciter les questions d'une problématique dont les réponses relèvent d'une approche statistique,	
12) de décrire, traiter, analyser des données de manière pertinente dans le cadre d'une étude en particulier dans le domaine éducatif,	
13) de faire le lien entre la réflexion analytique sur des questions relevant du champ de l'éducation, leur formalisation et leur traitement quantitatif,	
14) de lire avec un regard critique et distancié, les conclusions de diverses études statistiques apparaissant dans des rapports de recherche en Sciences de l'Éducation,	
15) de poursuivre de façon autonome et personnalisée un apprentissage en statistique afin d'enrichir ses acquis,	
16) de poser un regard plus positif à l'égard d'un domaine largement exploité dans les <i>media</i> , dans le sens de ne pas considérer les résultats dans l'ordre du tout ou rien mais en les replaçant judicieusement dans leur domaine de validité,	
17) d'exploiter des notions et des démarches mathématiques à des fins d'outils, et de ce fait de modifier dans un sens positif le rapport souvent négatif que nombre entretient avec cette science,	
18) de s'exercer à un raisonnement intégrant l'idée de "risque d'erreur" dans l'énoncé de ses conclusions.	
19) de s'exercer à l'interprétation de phénomènes éducatifs sur la base de données statistiques sur des "faits éducatifs" et sur des relations entre ces "faits"	
20) de s'exercer à la communication des résultats des analyses des données en distinguant clairement le modèle utilisé, de la réalité qu'il est supposé représenter, en séparant bien les traitements menés à l'intérieur du modèle, des interprétations reformulées dans le contexte du problème.	

Le questionnaire précédent est un fragment du questionnaire complet soumis à 132 étudiants ayant passé les épreuves d'évaluation du cours de méthodes quantitatives lors de la session 2002 d'examen de licence de sciences de l'éducation, préparation à distance. 297 étudiants sont inscrits dans cette licence. 86 étudiants ont rendu ce questionnaire. Nous n'avons rapporté ici qu'une partie des questions. Nous n'étudierons que quelques propriétés dans le cadre actuel.

Nous rapportons ci-dessous les tableaux statistiques des variables :

(Tableau n° T1) Tableau statistique de la variable V00 = sexe

Ensemble des inscrits			étudiants (au moins une épreuve de l'UE Méthodologie°)		
Modalités de V00	homme	femme		homme	femme
effectifs	49	248		23	109

(Tableau n° T2) Tableau statistique de la variable V01 = âge à la date de l'épreuve

Valeurs de V01	[20 ; 25[[25 ; 30[[30 ; 35[[35 ; 40[[40 ; 45[[45 ; 50[[50 ; 55[
effectifs	22	19	22	8	7	5	3

(Tableau n° T4) Tableau statistique de la variable V03 : statistique dans l'activité professionnelle

Modalités de V03	1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
effectifs	48	15	18	4	0	0

(Tableau n° T5) Tableau statistique de la variable V18 = efficacité du dossier méthodologique

modalités de V18	1 le plus faible	2	3	4	5	6	7	8	9	10 le plus fort
effectifs	1	0	2	3	9	9	24	18	9	9

(Tableau n° T6) Tableau statistique de la variable V20 = durée pour l'élaboration du dossier

Valeurs de V20	[0 ; 50[[50 ; 100[[100 ; 150[[150 ; 200[[200 ; 250[
effectifs	22	28	13	3	6
Valeurs de V20	[250 ; 300[[300 ; 350[[350 ; 400[[400 ; 450[[450 ; 675[
effectifs	2	0	0	1	3

(Tableau n° T7) Tableau statistique de la variable V24 : recours à liste de diffusion

Modalités de V24	1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
effectifs	39	12	20	4	3	4

(Tableau n° T8) Tableau statistique de la variable V47B= note à l'examen

notes/20	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
effectifs	0	1	0	2	8	4	6	7	1	3	5	4	7	9	9	6	2	5	2	4	0

Document n° 3 : corrigé-guide du traitement demandé

TQ1- De quoi s'agit-il ?

Q101- Après avoir pris connaissance des informations fournies dans le **document n°1**, identifier et caractériser "succinctement" et clairement, les éléments de cette étude statistique : population, échantillon, individus, variables.

La POPULATION est constituée par les ETUDIANTS INSCRITS EN LICENCE SE A L'UNIVERSITE LYON 2. L'EFFECTIF est de N = 297. Nous nous intéressons à l'ECHANTILLON des étudiants passant l'épreuve de méthodes quantitatives à la session 2002. L'EFFECTIF est de n=132 auquel nous avons soumis un QUESTIONNAIRE. Ce sous-ensemble est un **échantillon d'observation** constitué par des individus (étudiants en situation d'évaluation) dans des circonstances sociales et institutionnelles. Ainsi cet échantillon n'est pas représentatif au sens d'une extraction aléatoire. Nous conserverons malgré tout la propriété de représentativité dans la mesure où ce sous-ensemble est un passage obligé pour les étudiants qui souhaitent obtenir leur diplôme.

Questionnaires effectifs	distribués	retournés
	132	86

Relativement à ...	Population	Échantillon «Examen»
Taux de sondage	86/297	86/132
	28,96%	65,15%

En ce qui concerne la nature des VARIABLES statistiques déterminées par les QUESTIONS :

Nature des VARIABLES				
Variables quantitatives		Variables qualitatives		
discrètes	continues	nominales	ordinales	textuelles
V45 (rang)	V1 (→ âge) V6 (→ durée depuis obtention) V20	V2	V3 V4 V6 V18 V19 V24	V2 V5 V12 V17 V21 V22 V23 V25 V44
V47B (note sur 20)		[V00] [V47B]regroupée [V18]regroupée		

Q102- Sur l'une des variables qualitatives de votre choix, réaliser une courte synthèse s'appuyant sur un graphique de votre choix afin de rendre compte des informations apportées.

Nous choisissons la variable V24 = Recours à la liste de diffusion.

Cette question obtient 82 réponses réparties comme l'indique le tableau statistique

Modalités de V24	1_Pas du tout	2_Rarement	3_Quelquefois	4_Souvent	5_Très souvent	6_Toujours
Effectifs	39	12	20	4	3	4
Fréquences	0,47561	0,14634	0,24390	0,04878	0,03659	0,04878
Fréquences (%)	47,56%	14,63%	24,39%	4,88%	3,66%	4,88%

que nous traduisons graphiquement :

Diagramme en bâtons

Recours à liste de diffusion

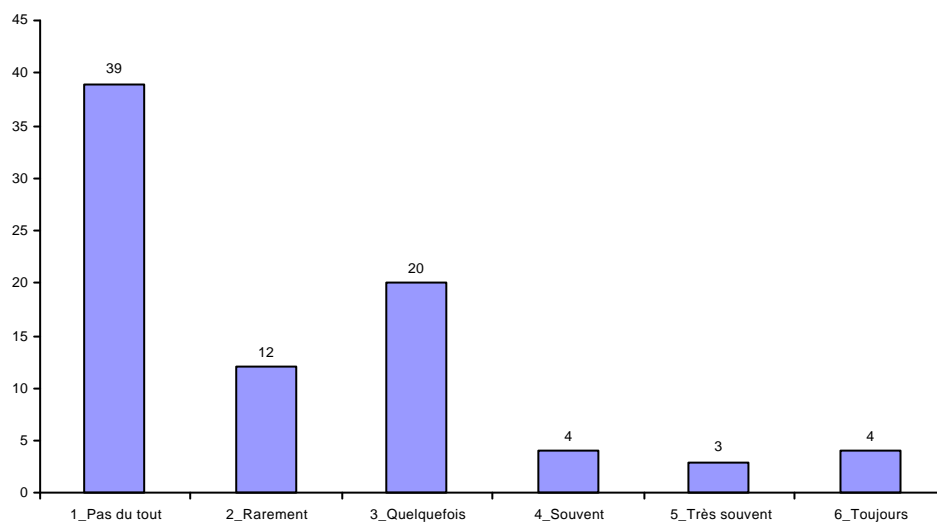
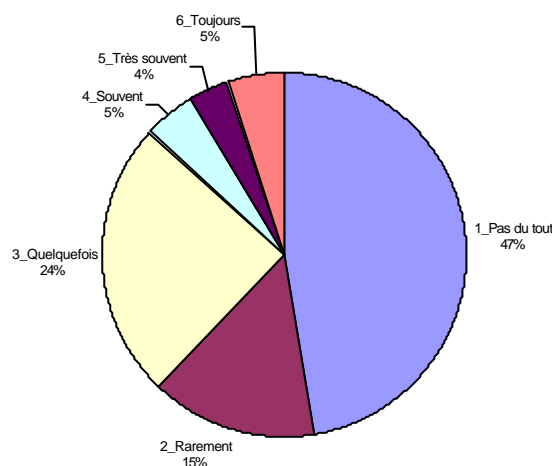


Diagramme circulaire

Recours à liste de diffusion



Il ressort que la majorité relative 47,56% (mode = pas du tout) est constituée par le groupe des étudiants n'ayant pas eu recours à la liste de diffusion pour l'étude du cours de méthodologie. La modalité médiane est "rarement".

Toutefois nous pouvons aussi dire que plus d'un étudiant sur deux (52,44%) a fait usage de cet outil de médiation qu'est la liste de diffusion.

Q103 Comment traiteriez-vous les réponses aux questions V12, V17, V44 ?

Il s'agit de variables textuelles, réponses textuelles à des questions ouvertes. Les approches développées dans le cours de méthodes qualitatives apportent des techniques tout à fait adaptées à ce type de données. Par exemple l'analyse lexicométrique (p.97-99). Elles requièrent donc un travail d'analyse de contenu.

Q104 Quelle est votre interprétation du tableau T1?

Nous pourrions nous demander si il y a une attitude à l'égard de la passation qui est déterminée par le genre. Autrement dit y a-t-il ou non homogénéité entre les hommes et les femmes face à la décision de passer ou non au moins une épreuve de l'UE_C ?

Nous pouvons alors recourir à test d'homogénéité du Khi-deux :

	homme	femme	
Passation UE_C	23	109	132
Non passation UE_C	26	139	165
	49	248	297

Sous l'hypothèse H_0 d'homogénéité, nous aurions les effectifs théoriques suivants :

	homme	femme	
Passation UE_C	22	110	132
Non passation UE_C	27	138	165
	49	248	297

Il est donc clair que cette hypothèse demeure la plus vraisemblable. La valeur empirique est de 0,147. Elle est très faible et de toute façon très inférieure à la valeur critique au niveau $\alpha = 0,05$ qui vaut 3,84. Le genre ne semble donc pas déterminé la décision de passation d'au moins une épreuve de l'UE_C à l'examen final. Nous conservons cette hypothèse avec un niveau de risque β de seconde espèce inconnu.

Q105 Quelle(s) remarque(s) feriez-vous à propos du questionnaire ?

En février 2003, nous avons soumis ce questionnaire à la critique à un échantillon d'observation de 69 étudiants (16 hommes et 53 femmes) de licence de Sciences de l'éducation. Nous vous rapportons les remarques formulées :

Individus	Réponses
F001	On peut noter que le questionnaire a plus de questions ouvertes que de questions fermées.
F004	Le questionnaire est organisé en partant de questions générales puis en progressant vers de questions plus précises. Ce qui permet que les questions soient moins orientées.
F005	Le questionnaire semble efficace puisqu'il alterne les questions fermées et les questions ouvertes. Celui-ci permet d'obtenir des réponses précises tout en laissant la possibilité aux étudiants de s'exprimer.
F009	Ce questionnaire est très précis et comporte un grand nombre de variables dont certaines sont inutilisables en tant que telle comme V12, V17 et V44.

F012	On peut remarquer que ce questionnaire a plus de questions ouvertes que semi-ouvertes.
F021	Ce questionnaire est convenablement construit car il intègre à la fois des questions ouvertes et fermées. Le discours est adapté, des mots simples et clairs sont utilisés. L'ordre des questions est bien établi, c'est à dire les questions fermées au départ ainsi que celles qui impliquent peu l'enquêté puis les questions plus délicates. La longueur du questionnaire est favorable. De plus il y a des questions qualitatives et quantitatives.
F022	Globalement ce questionnaire semble être plutôt complet. Il est en effet très intéressant de voir l'évolution du début à la fin de la licence.
F026	Le questionnaire met en évidence toutes les sortes de variables et de non-variables possibles. On remarque, en outre, qu'il ne permet pas de réponse moyenne ou, plutôt, de non-réponse. En outre, les effectifs varient à chaque variable.
F029	Évaluation d'une pédagogie avec l'objectif éventuel d'une adaptation.
F030	Ce qui me semble aberrant, c'est que les questionnaires ne sont pas anonymes. Ce qui n'est pas le signe d'une grande éthique. Où se trouve le secret de l'anonymat ?
F031	Les questions ouvertes permettent plus difficilement de traiter les réponses (notamment V12,V17,V44)
F033	Ce questionnaire progresse des questions les plus générales de l'utilisation de la statistique aux questions plus précises comme les savoirs acquis, par exemple. Il me semble que cette progression est pertinente dans le sens où l'enquêté n'est pas conduit à orienter ses réponses. Dans le sens inverse, l'enquêté saurait déjà exactement comment définir la statistique en trois mots, par exemple.
F038	Il s'agit d'un questionnaire où les questions fermées ou ouvertes sont mélangées. Il apparaît difficile à traiter.
F039	Ce questionnaire comporte de nombreuses questions ouvertes. Il n'est pas représentatifs de l'ensemble des étudiants qui ont passé la licence, en effet sur les 297 étudiants seulement 86 ont rendu le questionnaire. Une autre démarche de recueil serait nécessaire afin de collecter des données réelles sur la formation en apprentissage en statistique.
F041	Ce questionnaire mêle des questions ouvertes, semi-ouvertes et quelques questions fermées. Ce choix méthodologique me paraît rendre très long et délicat le traitement des données apportées par ce questionnaire. Il donne lieu à plus de variables qualitatives que quantitatives.
F043	Le questionnaire composé au total de 45 questions peut s'avérer long à compléter. Les questions extraites font apparaître des éléments relatifs à l'état civil, à la situation professionnelle, à l'intérêt pratique que peut avoir la personne quant aux statistiques dans son travail. Il permet d'obtenir, outre des éléments quantitatifs, des éléments qualitatifs quant à l'intérêt pour les statistiques et les cours de méthodes quantitatives. Il permet par les questions semi-ouvertes et ouvertes, aux enquêtés de donner leur opinion personnelle sans aspect restrictif.
F045	Je trouve ce questionnaire difficile d'approche de par sa forme. En effet, les titres des parties sont écrits en minuscule (exemple : vous et votre parcours de formation antérieur à la licence SE). De ce fait, il y a une confusion entre les questions, et cela ne donne pas envie de remplir ce questionnaire. Par contre, sur le fond, ce questionnaire suit une certaine logique (passé, présent, avenir). D'autre part, les questions faciles qui ne demandent pas une réflexion sont alternées avec des questions plus complexes.
F046	Ce questionnaire me semble avoir été fait pour que le professeur de statistique puisse évaluer la pédagogie et la méthode de travail afin qu'il puisse davantage les adapter aux étudiants de la session suivante

F047	<p>Le questionnaire procède par différentes phases. Il permet dans un premier temps à l'aide des premières questions de cibler la population par le sexe, l'âge et la situation professionnelle.</p> <p>Il permet aussi de visualiser le niveau des étudiants ainsi que leur motivation à utiliser ou non l'outil statistique. Ce questionnaire permet aussi à l'enquêteur de pouvoir faire un parallèle entre les représentations de l'étudiant face à l'outil statistique et à la réalité. Cependant je crois qu'il sera difficile ou plutôt délicat de traiter des questions comme V12, V17, V44 car elles me semblent trop ouvertes et peuvent ainsi amener l'enquêteur à devoir traiter trop de données différentes.</p> <p>De plus ce questionnaire basé sur les représentations fait appel à la subjectivité des étudiants, ce qui comme chacun sait est très personnel et risque donc d'induire chez l'enquêteur des résultats pouvant être contestés.</p>
F048	C'est un questionnaire qui propose beaucoup de questions ouvertes. Les réponses peuvent être difficiles à regrouper et donc à coder.
F051	Le questionnaire proposé part de la spécificité de chaque personne interrogée (Nom, sexe, date de naissance) pour aller vers des questions aussi bien d'ordre général avec les questions quantitatives mais en laissant un mode d'expression grâce aux variables qualitatives. Il permet ainsi de voir d'ensemble de ce que pensent les enquêtés sur la formation et peut amener une amélioration de l'enseignement.
F052	Nous pouvons constater que pour chaque variable, l'effectif total de l'une à l'autre variable est différent : V01 = 86 ; V03 = 85 ; V18 = 84, etc. D'autre part, il comprend beaucoup de variables qualitatives et de questions ouvertes, ce qui complique l'analyse.
F053	Ce questionnaire permet de prendre en compte plusieurs types de variables et donc d'être très précis. Mais certaines variables sont inutilisables d'un point de vue statistique.
F054	Nous pourrions dire que le questionnaire, partant des questions les plus générales vers les plus précises n'oriente pas la personne interrogée dans sa réponse, étant donné le fait que le chercheur part de questions générales.
F055	Évaluer la pédagogie en vue de l'adapter
F056	Ce questionnaire, ou tout du moins cet extrait, risque d'être long à traiter en raison du nombre de questions faisant intervenir, de la part de l'individu, le qualitatif ou des phrases qu'il doit proposer de lui-même. Un certain nombre de ces données risque d'autre part de ne pas être utilisables, car, dans ce cas, des réponses souvent peuvent s'avérer inadaptées. D'autre part, la proportion des individus qui s'abstiennent à ces questions risque d'être plus élevée que pour les autres questions, car cela demande un effort de réflexion et un temps plus important. Malgré tout, de telles questions ouvertes permettent une vue plus globale car elles n'évincent pas des réponses qui l'auraient été si les catégories avaient été pré-fixées. Peut-être, pour une enquête de grande envergure, aurait-il été judicieux de procéder à une pré-enquête permettant de créer des catégories qui soient les plus représentatives possible.
F059	Je peux dire qu'il y a homogénéité entre l'échantillon total et l'échantillon interrogé. De plus, ce questionnaire met en jeu toute sorte de variables et des questions d'ordre personnel.
F060	86 étudiants ont rendu ce questionnaire, cependant tous n'ont pas répondu à toutes les questions, si bien que n=86 mais que les effectifs de V03, V18, V20, V24 et V47B sont respectivement de 85, 84, 78, 82 et 85.
F062	Le questionnaire me paraît pertinent car il met en avant toute sorte de variables et des questions d'ordre personnel.
F063	Le questionnaire a plus de questions de type ouvert que semi-ouvert.
F065	Le questionnaire me semble pertinent car il met en jeu toute sorte de variables et des questions d'ordre personnel.
F068	Questions se rapportant à un échantillon peut-être un peu trop réduit de par le nombre de réponses aux questionnaires.
H014	Le questionnaire représente toutes les formes de modalités statistiques : qualitatif et quantitatif. Il est à la fois ordonné et ouvert à des réponses libres.

H017	Ce questionnaire fait apparaître différents types de variables et semble intéressant quant à la somme des données qu'il permet de recueillir. Mais sa longueur et son « accessibilité » pas nécessairement aisée peuvent expliquer un nombre conséquent de non-réponses
H018	Ce questionnaire semble assez complet. Ce qui est intéressant, c'est la vision du contenu de la licence au début puis à la fin.
H020	Je crois que ce questionnaire est très intéressant pour le formateur qui le fait passer car il pourra s'en servir pour évaluer son intervention. Néanmoins, je pense qu'il sera difficilement interprétable car les variables qualitatives et les questions ouvertes utilisées sont difficiles à regrouper dans ce genre d'étude. À la lecture de cette enquête, le formateur retiendra certaines suggestions ou certaines difficultés (V21, V23) qui peuvent éventuellement revenir plusieurs fois mais qui peuvent aussi être complètement différente d'un questionnaire à l'autre.
H027	C'est un questionnaire qui vise à rendre compte de l'intérêt, des difficultés et des observations des étudiants sur l'enseignement reçu dans le cadre de leur formation en statistique. On peut noter le souci du commanditaire du questionnaire de s'assurer de la cohérence et du bien-fondé de la statistique dans le cadre de cette formation. Cohérence observée ici à partir de l'avis des principaux intéressés : les étudiants. On peut observer que seulement 86 étudiants sur 132 ont rendu le questionnaire. On pourrait donc se poser la question des non-rendus, puisque les réponses à ce questionnaire ne relevait apparemment pas d'un caractère obligatoire.
H028	Il y a homogénéité entre l'échantillon total et l'échantillon interrogé. Je pense que, par contre, le questionnaire pour être plus précis aurait dû prendre en compte les 86 personnes ayant rendu ce questionnaire. Ce qui représente 29% des étudiants en licence de sciences de l'éducation. D'autre part, l'échantillon testé sur l'échantillon global ne représente que 44,5% donc pas majoritaire.
H032	Il y a beaucoup de questions ouvertes et qualitatives, ce qui rend difficile le traitement des données et donc l'interprétation de manière quantitative et statistique. Aussi l'effectif total varie selon les variables V01>86 ; V03>85 et V18>84
H042	Ce questionnaire comporte beaucoup de questions ouvertes. Ceci ne facilite pas le traitement.
H061	Il apparaît que les individus de l'échantillon se répartissent de manière bimodale sur l'échelle de réponses. Les réponses extrêmes sont plus nombreuses que les réponses moyennes. L'intérêt suscité par cette session de formation est néanmoins homogène pour la majorité des individus de l'échantillon. Les techniques d'apprentissage et l'efficacité semblent aussi homogènes. Des questions sur l'intérêt de cette formation au-delà des objectifs auraient pu apporter des réponses et aboutir à des analyses complémentaires.
H067	Le questionnaire comporte un nombre important de questions ouvertes ce qui peut être plus compliqué, il me semble, au niveau du traitement. Mais c'est également ce point qui doit en faire sa richesse.
H069	Dans sa globalité, ce questionnaire semble être assez complet. Il est surprenant et donc intéressant de voir le contenu de la licence au début puis à la fin.

TQ2- Étude de la variable V47B = note sur 20 obtenue à l'épreuve finale du cours de méthodes quantitatives (tableau T8)

Q201 Quelle est la valeur de la note moyenne m et celle de l'écart-type $s_{\text{échantillon}}$?

Q202 Quelle est la (les) note(s) modale(s) ?

Q203 Quels sont les quartiles Q_1 , Q_2 (note médiane) et Q_3 ?

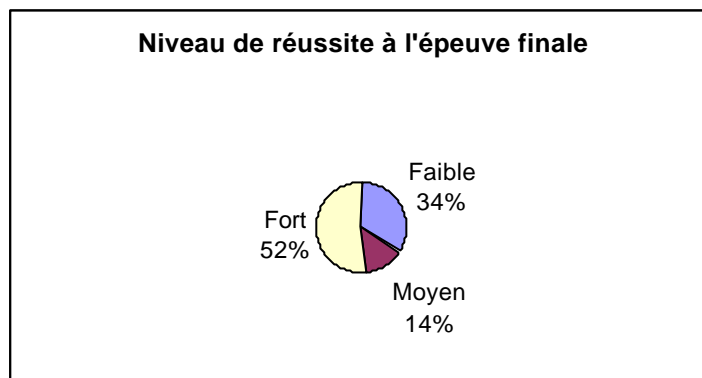
valeurs (x_i)	effectifs (n_i)	$n_i x_i$	$n_i (x_i - m)^2$	
0	0	0	0	
1	1	1	96,2707266	
2	0	0	0	
3	2	6	122,047336	
4	8	32	371,201107	
5	4	20	135,106436	
6	6	36	138,918478	
7	7	49	101,706851	
8	1	8	7,90602076	
9	3	27	9,84747405	
10	5	50	3,29480969	
11	4	44	0,1417301	
12	7	84	9,8833218	
13	9	117	43,0953633	mode
14	9	126	91,4835986	mode
15	6	90	105,247889	
16	2	32	53,8355709	
17	5	85	191,47128	
18	2	36	103,341453	
19	4	76	268,188789	
20	0	0	0	
Somme	85	919	1852,98824	
		moyenne 10,8117647	variance 21,7998616	
			écart-type 4,66903219	

Q1= quartile1= 6,5	Q2= médiane= 12	Q3= quartile3= 14	
Valeur centrale de l'intervalle défini par la valeur de rang 21 et celle de rang 22	Valeur de rang 43	Valeur centrale de l'intervalle défini par la valeur de rang 64 et celle de rang 65	<i>Remarque :</i> 85 = 4 x 21 + 1

Q204 Regrouper les notes en trois classes Faible = [0, 8], Moyen = [9 ; 11] et Fort = [12 ; 20] de niveau de réussite et construire un diagramme de la distribution des fréquences de ces trois classes.

Tableau statistique de la variable V47B= note à l'examen

Niveau de réussite finale notes/20	Faible								Moyen			Fort									
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
effectifs	0	1	0	2	8	4	6	7	1	3	5	4	7	9	9	6	2	5	2	4	0
effectifs	29								12			44									



Q205- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

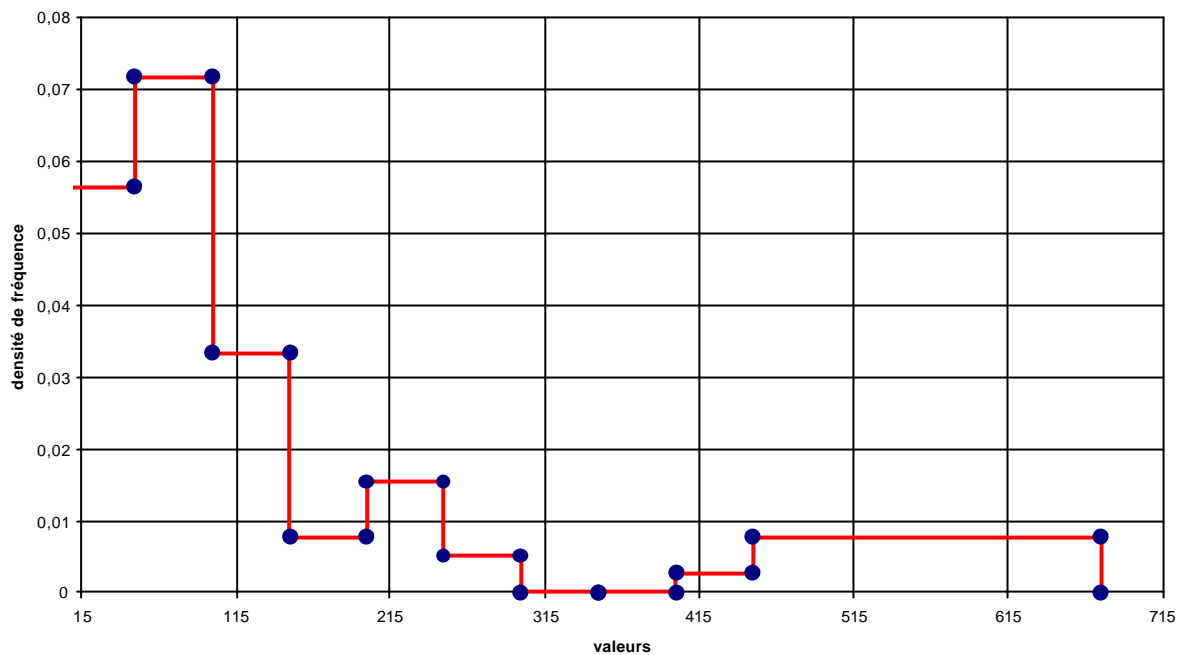
TQ3-Etude de la variable V20 = durée déclarée pour l'élaboration du dossier méthodologique

Q301 Quelle est la valeur de la note moyenne m et celle de l'écart-type $s_{\text{échantillon}}$?

	valeurs centrales (x_i)	effectifs (n_i)	$n_i x_i$	$n_i (x_i - m)^2$
[0 ; 50[25	22	550	170293,5466
[50 ; 100[75	28	2100	40391,08728
[100 ; 150[125	13	1625	1878,004808
[150 ; 200[175	3	525	11539,15496
[200 ; 250[225	6	1350	75289,84837
[250 ; 300[275	2	550	52500,46228
[300 ; 350[325	0	0	0
[350 ; 400[375	0	0	0
[400 ; 450[425	1	425	97356,00037
[450 ; 675[562,5	3	1687,5	606202,6165
	Somme	78	8812,5	1055450,721
			moyenne 112,981	variance 13531,4195
				écart-type 116,3246298

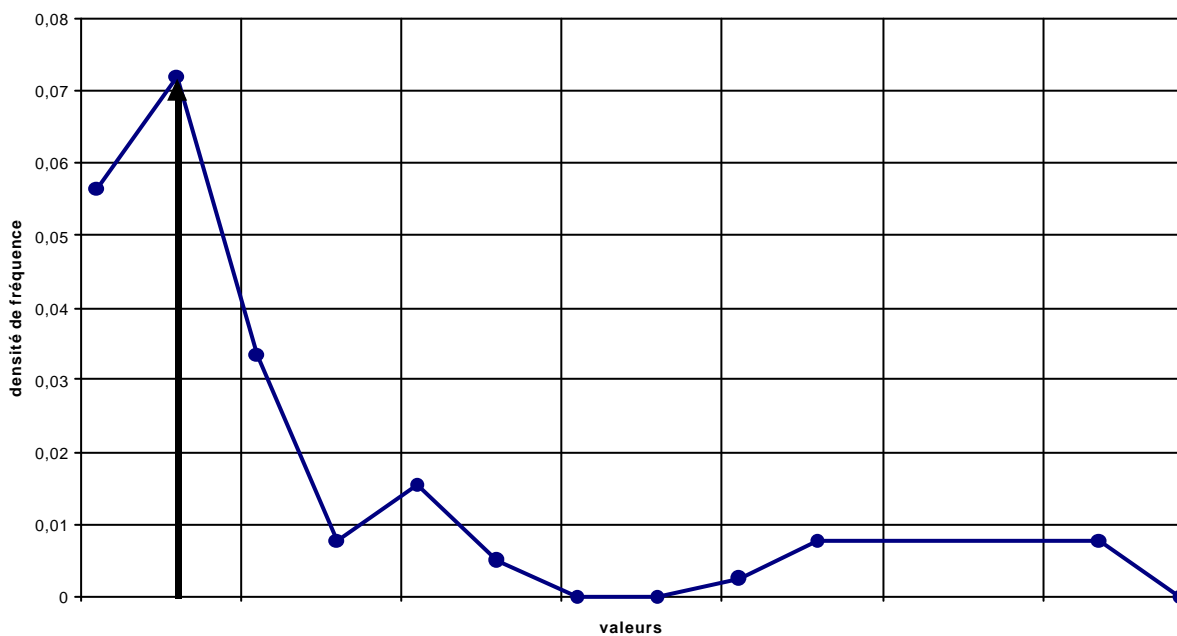
Q302 Construire l'histogramme de la distribution des fréquences de V20

Histogramme de la variable Durée déclarée de réalisation du dossier



Q303 Construire le polygone des fréquences à partir de l'histogramme

Histogramme lissé



Q304 À partir du polygone des fréquences, donner une estimation de la (des) durée(s) modale(s)

La valeur modale peut être estimée par la valeur de densité de fréquence maximum repérable sur l'histogramme : sommet de cette courbe. Nous estimons cette valeur à 75

Q305 À partir de l'histogramme des fréquences, donner une estimation de la durée médiane.

Il s'agit de trouver une valeur qui partage l'histogramme en deux parties dont les surfaces ont même aire. Cette valeur est située entre 50 et 100. Nous pouvons l'estimer en procédant par interpolation linéaire.

Q306 Donner une estimation ponctuelle s^2 de la variance de V20 sur la population à partir de $s^2_{\text{échantillon}}$.

$$S^2 = \frac{n}{n-1} \sigma_{\text{échantillon}}^2 = \frac{78}{77} 13531,4195 = 13707,15222$$

$$\text{d'où } s = 117,0775479$$

Q307 Quelle est la proportion des étudiants ayant déclaré une durée comprise l'intervalle $[m - s ; m + s]$?

Il s'agit alors de donner une estimation de la proportion des étudiants ayant déclaré une durée dans l'intervalle $[-4,096778671 ; 230,0583171]$ soit donc entre les valeurs -4 et 230 .

Entre 0 et 200, nous observons une proportion de $66/78$, donc entre 0 et 230, nous en déduisons par interpolation linéaire, une proportion de $69,6/78$ c'est à dire environ $89,2\%$ des étudiants.

Q308 Quelle serait la proportion si la variable V20 était une variable continue dont la distribution de fréquences est celle d'une variable de Laplace-Gauss de paramètres m et s ?

Si la variable était une variable $LG(112,98 ; 117,07)$, alors nous pourrions espérer une proportion d'environ $68,28\%$ de réponses situées dans l'intervalle $[m - s ; m + s]$

Q309- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

TQ4- Étude du lien entre le jugement d'efficacité du dossier méthodologique et le niveau de réussite à l'épreuve d'examen

Q401- A partir du sous-échantillon des 83 répondants, peut-on considérer qu'il y a une dépendance entre le jugement d'efficacité et le niveau de réussite à l'examen ?

Nous avons réduit le nombre de modalités de la variable V18 à trois niveaux.

Tableau de contingence		Niveau d'efficacité Dossier			
		N1 Faible 1,2,3	N2 Moyen 4,5,6,7	N3 Fort 8,9,10	
Niveau de réussite à l'épreuve finale	Faible	1	15	11	27
	Moyen	1	5	6	12
	Fort	1	25	18	44
		3	45	35	83

L'hypothèse H_0 s'énonce ainsi « le niveau de performance à l'épreuve finale de méthodes quantitatives est indépendant du niveau d'efficacité attribué à la réalisation du dossier méthodologique »

L'hypothèse alternative H_1 s'énonce par négation de H_0 « le niveau de performance à l'épreuve finale de méthodes quantitatives est dépendant du niveau d'efficacité attribué à la réalisation du dossier méthodologique »

Sous l'hypothèse H_0 , au seuil de risque de première espèce $\alpha = 0,05$ la valeur critique déterminée à partir de la distribution du Khi-deux à ddl = 4 vaut 9,49. La valeur empirique est de 1,576. Elle est donc très inférieure à la valeur critique. Nous conservons donc l'hypothèse d'indépendance avec un risque de seconde espèce β de niveau inconnu.

L'ensemble de la procédure détaillée est présentée ci-dessous :

Tableau de contingence (effectifs observés)

	N1	N2	N3	
Faible	1	15	11	27
Moyen	1	5	6	12
Fort	1	25	18	44
	3	45	35	83

Tableau des fréquences

	N1	N2	N3	
Faible	0,01204819	0,18072289	0,13253012	0,3253012
Moyen	0,01204819	0,06024096	0,07228916	0,14457831
Fort	0,01204819	0,30120482	0,21686747	0,53012048
	0,03614458	0,54216867	0,42168675	1

Tableau des fréquences conditionnelles

Profil ligne

	N1	N2	N3	
Faible	0,03703704	0,55555556	0,40740741	1
Moyen	0,08333333	0,41666667	0,5	1
Fort	0,02272727	0,56818182	0,40909091	1
	0,03614458	0,54216867	0,42168675	1

Tableau des **fréquences conditionnelles***Profil colonne*

	N1	N2	N3	
Faible	0,33333333	0,33333333	0,31428571	0,3253012
Moyen	0,33333333	0,11111111	0,17142857	0,14457831
Fort	0,33333333	0,55555556	0,51428571	0,53012048
	1	1	1	1

Tableau des **effectifs théoriques** sous Ho

	N1	N2	N3	
Faible	0,97590361	14,6385542	11,3855422	27
Moyen	0,43373494	6,5060241	5,06024096	12
Fort	1,59036145	23,8554217	18,5542169	44
	3	45	35	83

Tableau des **contributions absolues**

	N1	N2	N3	
Faible	0,00059497	0,00892459	0,0130554	0,02257496
Moyen	0,7392905	0,34861669	0,17452668	1,26243386
Fort	0,21914932	0,05491664	0,01655453	0,29062049
	0,95903479	0,41245791	0,2041366	1,57562931

Tableau des **contributions relatives**

	N1	N2	N3	
Faible	0,0378%	0,5664%	0,8286%	1,4328%
Moyen	46,9203%	22,1256%	11,0766%	80,1225%
Fort	13,9087%	3,4854%	1,0507%	18,4447%
	60,8668%	26,1773%	12,9559%	1

Valeurs lues dans la table du Khi-deux (ddl=4) niveau α du risque de 1 ^{ère} espèce	Valeurs critiques k	La valeur empirique $d^2 = 1,576$ très inférieure aux valeurs critiques
0,01	13,28	Par ailleurs si nous cherchons à l'aide de la fonction LOI.KHIDEUX(1,576 ;4) sous Excel, nous trouvons que la probabilité d'obtenir une valeur supérieure à 1,576 avec la variable du Khi-deux (ddl=4) vaut :
0,05	9,49	Prob($D^2 > 1,576$) = 0,81316548
0,10	7,78	

Q402- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

La valeur empirique $d^2 = 1,576$ est très inférieure à la valeur critique du Khi-deux au niveau de risque 5 % : 9.49 . On conserve l'hypothèse Ho d'indépendance de ces deux variables en prenant un risque de seconde espèce de niveau β inconnu.

TQ5- Étude du lien entre le niveau d'intérêt pour réaliser le dossier méthodologique et la note finale à l'épreuve de méthodes quantitatives.

Nous considérons ici V18c, la variable niveau d'intérêt pour la réalisation du dossier méthodologique, comme une variable quantitative discrète prenant ses valeurs sur $\{1; \dots; 10\}$

Le tableau ci-dessous est le tableau statistique de la distribution conjointe du couple de variables V18c et V47B.

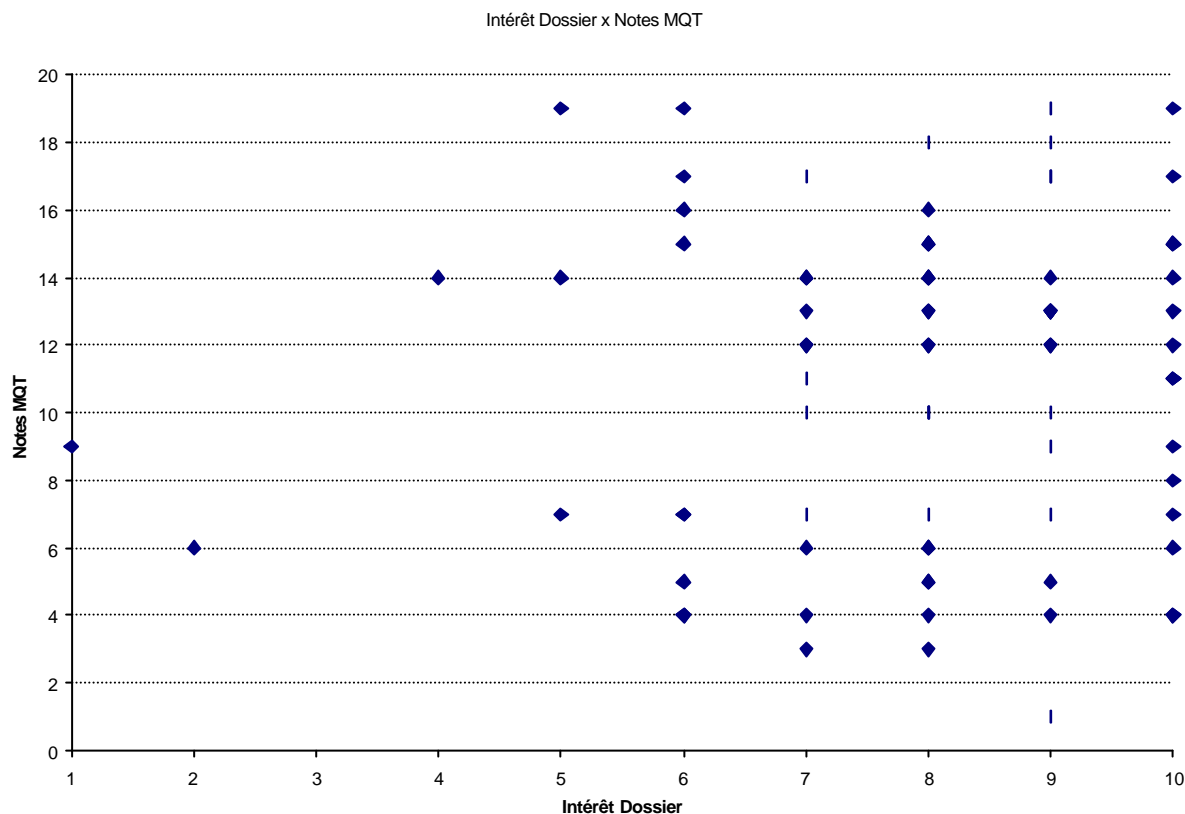
Q501- Analyser le tableau ci-dessous.

Tableau de contingence		V47B notes à l'épreuve finale de méthodes quantitatives																				TOT.
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Niveau d'intérêt pour la réalisation du Dossier méthodologique	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
	5	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	3
	6	0	0	0	2	1	0	2	0	0	0	0	0	0	0	1	1	1	0	1	0	9
	7	0	0	1	1	0	1	1	0	0	1	1	2	1	2	0	0	1	0	0	0	12
	8	0	0	1	1	2	2	1	0	0	3	0	2	2	3	3	1	0	1	0	0	22
	9	1	0	0	1	1	0	1	0	1	1	0	2	5	1	0	0	2	1	1	0	18
	10	0	0	0	2	0	1	1	1	1	0	3	1	1	1	2	0	1	0	1	0	16
TOT.	1	0	2	7	4	5	7	1	3	5	4	7	9	9	6	2	5	2	4	0	83	

Un premier niveau d'exploration consiste à griser les cases dont l'effectif est non nul. L'ensemble ainsi obtenu est une représentation du nuage statistique. La forme peut suggérer une tendance allant dans le sens d'une corrélation positive traduisant le fait que « plus l'intérêt est fort, plus la performance est forte ». Toutefois une difficulté provient du fait les cases n'ont pas toutes la même importance puisque les effectifs varient de 1 à 5.

Q502- Construire le nuage statistique.

Nous pouvons aussi réaliser une représentation géométrique plus précise en recourant à un repère orthonormé. L'axe des abscisses représente le niveau d'intérêt pour le dossier et l'axe des ordonnées, le niveau de performance à l'épreuve finale. Toutefois les points géométriques demeurent des points pondérés par les effectifs. La difficulté d'interprétation n'est pas complètement levée.



Q503- Calculer le coefficient de corrélation linéaire de Bravais-Pearson

Pour calculer le coefficient de corrélation linéaire de Bravais-Pearson, il y a nécessité de calculer préalablement :

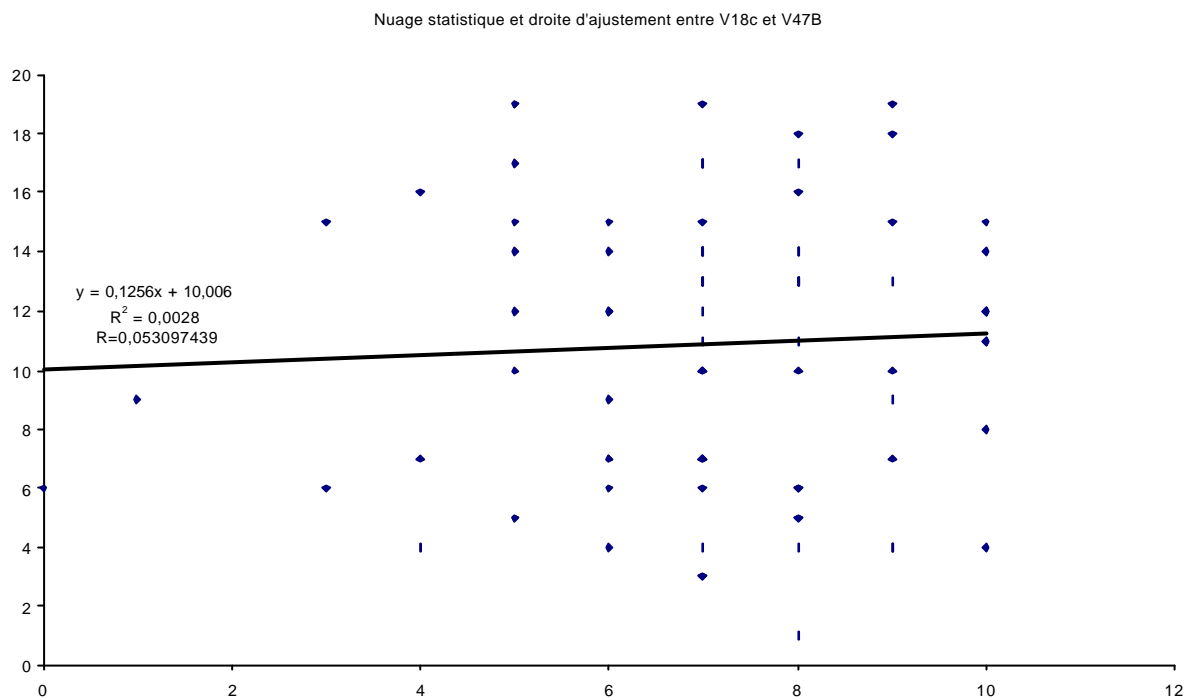
Le niveau moyen d'intérêt pour le dossier, la performance moyenne à l'épreuve finale, puis les indicateurs de fluctuation : variance et écart-type de V18c et de V47B, enfin la covariance de V18c et V47B.

Nous obtenons les résultats suivants :

Moyenne de V18c	7,070588235
Variance de V18c	3,806782007
Écart type de V18c	1,95109764
Moyenne de V47B	10,81176471
Variance de V47B	21,79986159
Écart type de V47B	4,66903219
Covariance (V18c ; V47B)	0,482568027

De là nous pouvons calculer la valeur du coefficient de Bravais Pearson empirique :

$$R_{BP}(V18c; V47B) = \frac{Cov(V18c; V47B)}{\sigma(V18c)\sigma(V47B)} = \frac{0,482568027}{(1,95109764)(4,66903219)} = 0,053097439$$



Nous pouvons chercher à déterminer une droite d'ajustement. En utilisant les fonctions du logiciel Excel, nous obtenons :

Q504- *A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.*

Si le signe du coefficient confirme l'hypothèse d'une corrélation positive, sa faible valeur ne paraît cependant pas s'éloigner significativement de 0. Il semble difficile de tenir pour vraisemblable l'existence d'une corrélation linéaire entre les deux variables. Nous serions tenté de conserver l'hypothèse d'absence de lien.

TQ6- Étude de l'intérêt personnel pour les TIC en formation

Une question concernant l'intérêt personnel pour les TICE a été posée sous le format suivant : *Quel est votre degré d'intérêt personnel relatif à l'intégration des TIC en formation ? (cocher la case correspondante)*

Très faible (f1) Plutôt faible (f2) Plutôt fort (F3) Très fort (F4) Sans avis (SA5)

(Tableau n° T9) Tableau statistique de la variable V07

Modalités de V09	f1	f2	F3	F4	SA5
Effectifs	15	13	27	15	10

Q601- *A partir des informations obtenues sur l'échantillon et consignées dans le tableau T9, peut-on considérer que les points de vue sont partagés ? (il s'agit de tester l'hypothèse H_0 contre H_1 à l'aide du test du Khi-deux d'adéquation avec un niveau de risque maximum de 5%)*

L'hypothèse H_0 peut être ainsi formulée "la distribution des fréquences des diverses opinions f_1, f_2, f_3, f_4 et f_5 est uniforme" contre l'hypothèse alternative H_1 "cette distribution n'est pas uniforme et comporte une opinion dominante f_3 ".

Pour résoudre ce problème, nous recourons à la procédure suivante :

Un test d'adéquation consiste à expliciter une règle de décision relative à une hypothèse portant sur la **comparaison d'une distribution empirique à une distribution théorique précisée a priori ou ajustée à partir des observations recueillies.**

Conditions d'utilisation:

- Prélever aléatoirement dans la population P inconnue un échantillon de n observations indépendantes réparties en s classes,
- Choisir une loi de distribution théorique dont les r paramètres sont soit fixés a priori, soit estimés à partir de l'échantillon.
- Aucun effectif théorique n'est inférieur à 5.

Critère, variable de décision:

On choisit la variable $D^2 = \sum_{i=1}^{i=s} \frac{(O_i - T_i)^2}{T_i}$ où O_i = effectif observé et T_i = effectif théorique sous la

contrainte de la loi de référence choisie pour chacune des s classes.

Si aucun des effectifs théoriques n'est inférieur à 5, la variable D^2 suit approximativement une loi du Khi-Deux à $s-1-e$ degrés de liberté (où e désigne le nombre de paramètres estimés parmi les r paramètres).

Si un des effectifs est inférieur à 5, on procède à des regroupements de classes pour se mettre dans les conditions d'application.

Conclusion:

Choisir la valeur α du risque de première espèce. Rechercher dans une table du Khi-Deux à $s-1-e$ degrés de liberté, la valeur critique k telle que $\text{Prob}\{\chi^2_{(s-1-e)} > k\} = \alpha$. Calculer d^2 la valeur empirique de D^2 avec les s valeurs O_i observées sur l'échantillon extrait.

Si $d^2 < k$, c'est à dire si $d^2 \in A = [0, k]$, région d'acceptation, alors on ne rejette pas l'hypothèse nulle H_0 , on accepte alors H_0 en prenant un risque β d'erreur de seconde espèce

Si $d^2 > k$, c'est à dire si $d^2 \in K = [k, +\infty[$, région critique, alors on rejette l'hypothèse nulle H_0 et on accepte H_1 .

Pour ce qui concerne notre cas , le dépouillement fait apparaître les résultats suivants que nous consignons dans le tableau statistique :

(Tableau n° T9) Tableau statistique de la variable V07

Modalités de V09 = opinions	f1	f2	F3	F4	SA5	
Effectifs observés	15	13	27	15	10	80
Effectifs théoriques sous Ho	16	16	16	16	16	80

Calcul de la valeur d^2 prise par la variable $D^2 = \sum_{i=1}^{i=s} \frac{(O_i - T_i)^2}{T_i}$ où ici $s = 5$.

$$d^2 = \frac{(15 - 16)^2}{16} + \frac{(13 - 16)^2}{16} + \frac{(27 - 16)^2}{16} + \frac{(15 - 16)^2}{16} + \frac{(10 - 16)^2}{16} = \frac{168}{16} = 10,5$$

Comme nous avons choisi la valeur $\alpha = 0,05$ du risque de première espèce, nous recherchons alors dans une table du Khi-Deux à $s-1-e = 5-1-0 = 4$ degrés de liberté, la valeur critique k telle que $\text{Prob}\{\chi^2_{(4)} > k\} = 0,05$, nous trouvons $k = 9,48$.

Nous constatons que d^2 est dans la région de rejet de H_0 , $d^2 = 10,5 > 9,48$, alors on rejette l'hypothèse nulle H_0 , avec un risque de première de $\alpha = 0,05$

Si nous choisissons un niveau de risque $\alpha = 0,01$ alors la valeur critique k telle que $\text{Prob}\{\chi^2_{(4)} > k\} = 0,01$, nous trouvons $k = 13,27$. Comme nous constatons alors que d^2 est dans la région $A = [0, k]$, région d'acceptation, $d^2 = 10,5 < 13,27$, alors on ne rejette pas l'hypothèse nulle H_0 , et l'on accepte en prenant un risque β d'erreur de seconde espèce dont la valeur nous est inconnue.

Q602- A partir de cette étude, énoncer sous une forme synthétique **vosre conclusion** et **vosre interprétation**.

Au niveau de risque $\alpha = 0,05$ on peut rejeter l'affirmation selon laquelle les opinions sont "partagées" à l'égard de l'intérêt pour les TIC en formation et admettre comme vraisemblable l'existence d'un point de vue dominant autour d'un niveau d'intérêt plutôt fort..

