

## Document autocorrectif et auto-évaluatif

### Q1 A propos de l'enquête par questionnaire.

Considérons une enquête par questionnaire visant la construction de données pour une étude sur la formation en statistique en licence de sciences de l'éducation.

Q101 Par une présentation sous forme d'un tableau, exposer les principales étapes de cette enquête par questionnaire et les précautions majeures à prendre.

Q102 Construire 5 questions relevant chacune d'un format différent que vous préciserez.

### Q2 A propos du traitement des données construites.

Considérons le cas suivant :

**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.**

Un questionnaire comportant 45 questions 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 que 7 questions.

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 :**

***La situation professionnelle comporte-t-elle des activités qui requièrent des connaissances en statistique ?***

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 =**

***Évaluation de l'efficacité du dossier méthodologique pour l'apprentissage***

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 =  
évaluation de la durée pour l'élaboration du dossier**

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

**(Tableau n° T7) Tableau statistique de la variable V24 :  
évaluation de la fréquence de 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 obtenue à l'examen à l'épreuve de statistique**

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

### TQ2.1- De quoi s'agit-il ?

Q201- Après avoir pris connaissance des informations fournies ci-dessus, identifiez et caractérisez "succinctement" mais clairement, les éléments du modèle statistique : **population, échantillon, individus, variables.**

**Population :**

Ensemble des étudiants (**individus**) préparant la licence de sciences de l'éducation dans un dispositif de formation à distance

**Échantillons :**

Sous-ensemble des individus ayant passé l'épreuve de méthodes quantitatives constitué de 132 individus. Partie (**sous-échantillon**) constituée des 86 étudiants ayant rendu le questionnaire de l'enquête issue des 132 étudiants de l'échantillon.

**Variables.**

Variables quantitatives		Variables qualitatives		
discrètes	continues	ordinales	nominales	Textuelles
[V47B]	[V01]→âge [V20]→durée de réalisation du dossier	[V03] [V18] [V24]	[V00] Sexe	

Q202- Sur l'une des variables **qualitatives** de votre choix, réalisez une très courte synthèse s'appuyant sur un graphique de votre choix afin de rendre compte des informations apportées.

Q203 Quelle est votre interprétation du tableau T1?

**TQ2.2-Etude de la variable V47B = note sur 20 obtenue à l'épreuve finale du cours de méthodes quantitatives**

Q301 Quelle est la valeur de la **note moyenne m** et celle de l'**écart-type s** échantillon ?

Q302 Quelle est la (les) **note(s) modale(s)** ?

Moyenne	Variance & écart-type	Modes
$m = \frac{919}{85} \approx 10,8$	$\sigma^2 \approx 21,8$ $\sigma \approx 4,7$	$m_0 = 13$ et 14

Le tableau ci-dessous fournit le détails des calculs.

valeurs (xi)	effectifs (ni)	$n_i x_i$	$n_i(x_i-m)^2$
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
14	9	126	91,4835986
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
	<i>effectif total</i>	<i>somme</i>	<i>somme</i>
	85	919	1852,98824
		<i>moyenne</i>	<i>variance</i>
		10,8117647	21,7998616
			<i>écart-type</i>
			4,66903219

Q303- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

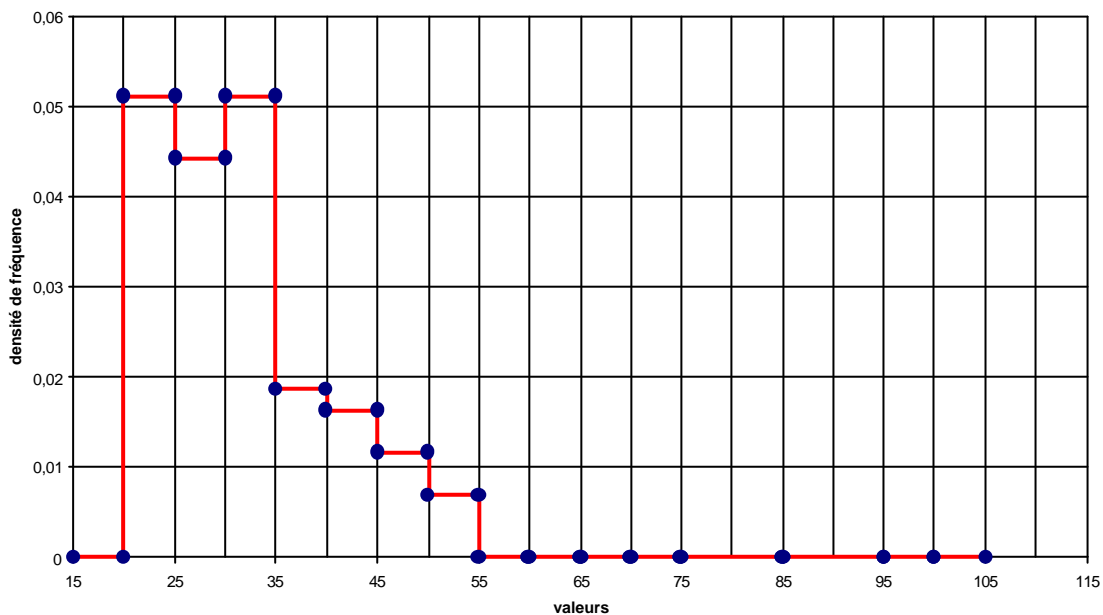
**TQ2A.3-Etude de la variable V01 = âge à la date de l'épreuve.**

Q401 Quelle est la valeur de l'âge moyen  $m$  et celle de l'écart-type  $s_{\text{échantillon}}$  ?

valeurs centrales (xi)	effectifs (ni)	$n_i x_i$	$n_i(x_i-m)^2$
22,5	22	495	1856,435911
27,5	19	522,5	332,9367226
32,5	22	715	14,57544619
37,5	8	300	270,4164413
42,5	7	297,5	818,5911303
47,5	5	237,5	1250,405625
52,5	3	157,5	1299,661979
<b>somme</b>	<b>86</b>	<b>2725</b>	<b>5843,023256</b>
		<b>moyenne</b>	<b>variance</b>
		<b>31,686</b>	<b>67,94213088</b>
			<b>écart-type</b>
			<b>8,242701674</b>

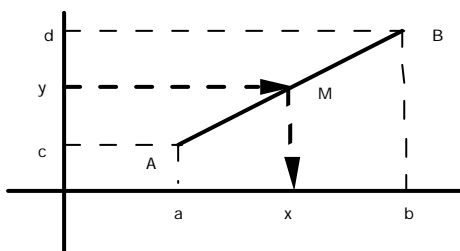
Q402 Construire l'histogramme de la distribution des fréquences.

Histogramme de la variable V02



Q403 Quelle la valeur de la **âge médian** ?

La valeur médiane se situe dans l'intervalle [30 ; 35]. L'estimation de l'âge médian par interpolation linéaire donne approximativement **30,45** soit environ **30 ans 6 mois** . Cette démarche est fondée sur le raisonnement suivant



$$\frac{d-c}{b-a} = \frac{y-c}{x-a} = \frac{d-y}{b-x}$$

ainsi ici  $\frac{43 - 41}{x - 30} = \frac{22}{5}$

donne  $x = 30 + \frac{10}{22} \approx 30,45$

Q404 Quelle est la proportion estimée des individus dans l'intervalle  $[m - s_{\text{échantillon}} ; m + s_{\text{échantillon}}]$  ?

Il s'agit de réaliser une estimation de la proportion des individus dont l'âge est dans l'intervalle  $[23,444 ; 39,928]$  à partir de l'histogramme. Nous procédons par interpolation linéaire.

Pour ce faire, il faut estimer la proportion des individus dont l'âge est inférieur à 39,928 et lui retrancher la proportion des individus dont l'âge est inférieur à 23,444.

Soit, il y a 22 individus dans l'intervalle  $[20 ; 25[$  ( nous pouvons estimer qu'il y en a un peu plus de la moitié de 22 dans l'intervalle  $[20 ; 23,444]$  ) Par interpolation, nous obtenons la

proportion dans l'intervalle  $[20 ; 23,444]$  : 
$$\frac{y - 0}{23,444 - 20} = \frac{22 - 0}{25 - 20}$$

ainsi  $y \approx 3,444 \times 4,4$  et donc une proportion de  $\frac{15,1536}{86} \approx 0,17962$

En d'autres termes : nous estimons cette proportion à **17,96%**

Soit, il y a 63 individus dans l'intervalle  $[20 ; 35[$ . Par interpolation, nous obtenons la proportion dans l'intervalle  $[20 ; 39,928[$  :

$$\frac{y - 63}{39,928 - 35} = \frac{71 - 63}{40 - 35}$$

ainsi  $y \approx 63 + 4,928 \times 1,6$  et donc une proportion de  $\frac{70,8848}{86} \approx 0,8242$

En d'autres termes : nous estimons cette proportion à **82,42%**

**Finalement la proportion recherchée est de 64,44%**

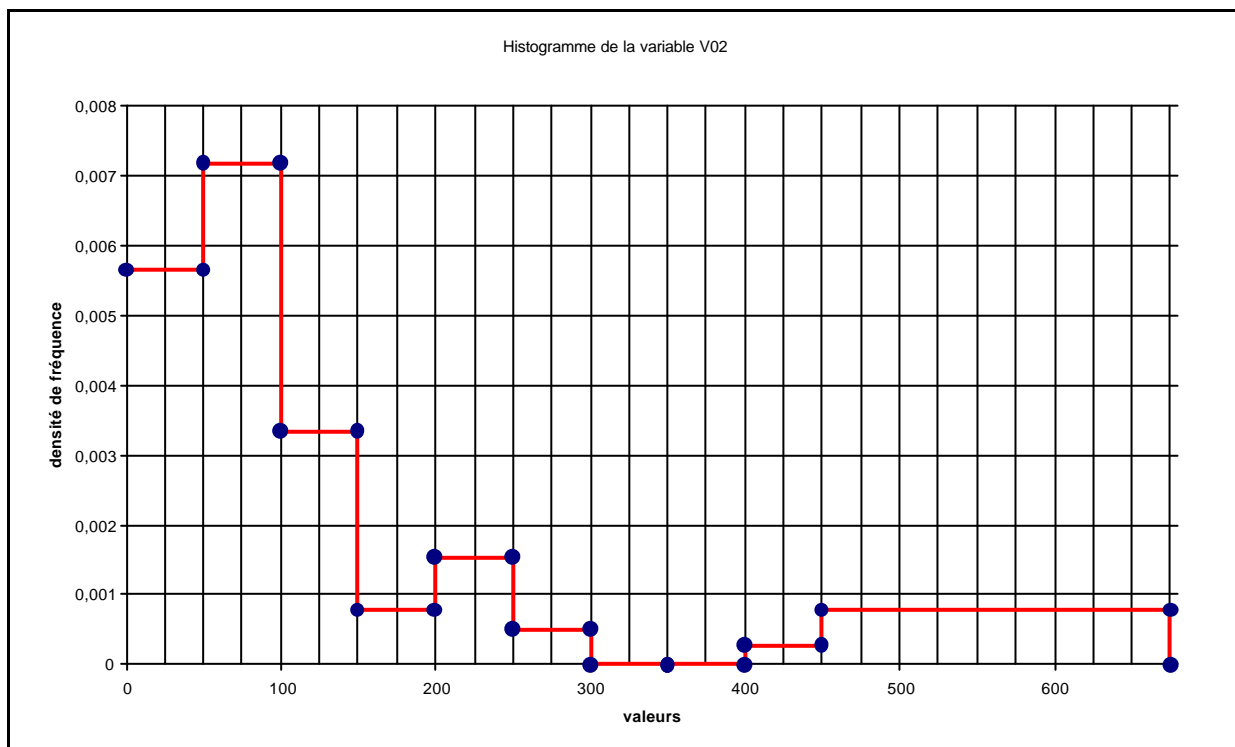
Q405- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

**TQ2B.3-Etude de la variable V20 = durée estimée pour élaborer le dossier méthodologique.**

Q401 Quelle est la valeur de la **durée moyenne m** et celle de l'**écart-type s<sub>échantillon</sub>** ?

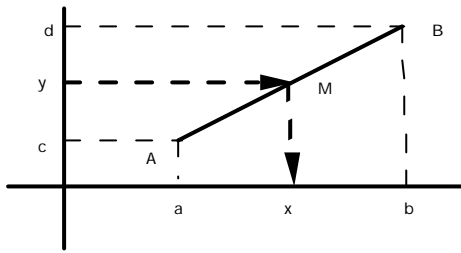
valeurs centrales (xi)	effectifs (ni)	$n_i x_i$	$n_i(x_i-m)^2$
25	22	550	170293,5466
75	28	2100	40391,08728
125	13	1625	1878,004808
175	3	525	11539,15496
225	6	1350	75289,84837
275	2	550	52500,46228
325	0	0	0
375	0	0	0
425	1	425	97356,00037
562,5	3	1687,5	606202,6165
<b>somme</b>	<b>78</b>	<b>8812,5</b>	<b>1055450,721</b>
		<b>moyenne</b>	<b>variance</b>
		<b>112,981</b>	<b>13531,4195</b>
			<b>écart-type</b>
			<b>116,3246298</b>

Q402 Construire l'histogramme de la distribution des fréquences.



Q403 Quelle la valeur de la **durée médiane** ?

La valeur médiane se situe dans l'intervalle [50 ; 100]. L'estimation de la durée médiane par interpolation linéaire donne approximativement **80,35 minutes** soit environ **1h 21 mn** . Cette démarche est fondée sur le raisonnement suivant



$$\frac{d-c}{b-a} = \frac{y-c}{x-a} = \frac{d-y}{b-x}$$

ainsi ici  $\frac{39-22}{x-50} = \frac{28}{100-50}$

donne  $x = 50 + \frac{850}{28} \approx 80,35$

Q404 Quelle est la proportion estimée des individus dans l'intervalle  $[m-S_{\text{échantillon}} ; m+S_{\text{échantillon}}]$  ?

Il s'agit de réaliser une estimation de la proportion des individus dont la durée est dans l'intervalle  $[-3,343 ; 229,305]$  à partir de l'histogramme. Nous procédons par interpolation linéaire. Cela revient à la proportion des individus dont la durée est inférieure à 229,305.

Soit, il y a 66 individus dans la durée est inférieure à 200, c'est à dire  $\frac{66}{78}$ . Par interpolation, nous obtenons la proportion dans l'intervalle  $[0 ; 229,305]$ , en raisonnant sur

l'intervalle  $[200 ; 229,305]$  :

$$\frac{y - 66}{229,301 - 200} = \frac{72 - 66}{250 - 200}$$

ainsi  $y \approx 66 + 29,301 \times 0,12$  et donc une proportion de  $\frac{69,516}{78} \approx 0,8912$

En d'autres termes : nous estimons cette proportion à **89,12%**

Q405- A partir de cette étude, énoncer sous une forme synthétique votre **conclusion** et votre **interprétation**.

Nom de l'étudiant:

n° :

code	items repérés et sur lesquels le jugement est porté	validité précision				validité pertinence			
Q101_1	- Étapes d'une enquête par questionnaire : construction	R+	R-	E	N	R+	R-	E	N
Q101_2	- Étapes d'une enquête par questionnaire : passation	R+	R-	E	N	R+	R-	E	N
Q101_3	- Étapes d'une enquête par questionnaire : traitement	R+	R-	E	N	R+	R-	E	N
Q102_1	- Question 1 (format1 = fermée nominale ; pertinence)	R+	R-	E	N	R+	R-	E	N
Q102_2	- Question 2 (format2 = fermée ordinale; pertinence)	R+	R-	E	N	R+	R-	E	N
Q102_3	- Question 3 (format3 =; pertinence)	R+	R-	E	N	R+	R-	E	N
Q102_4	- Question 4 (format1 ; pertinence)	R+	R-	E	N	R+	R-	E	N
Q102_5	- Question 5 (format1 ; pertinence)	R+	R-	E	N	R+	R-	E	N
code	items repérés et sur lesquels le jugement est porté	validité du résultat				validité du raisonnement			
Q201_1	- Identification : population, individus	R+	R-	E	N	R+	R-	E	N
Q201_2	- Identification : échantillon, individus	R+	R-	E	N	R+	R-	E	N
Q201_3	- Identification : variable 1 = sexe = qualitative nominale (T1)	R+	R-	E	N	R+	R-	E	N
Q201_4	- Identification : variable 2 = âge = quantitative continue (T2)	R+	R-	E	N	R+	R-	E	N
Q201_5	- Identification : variable 3 = V03 = qualitative ordinale (T4)	R+	R-	E	N	R+	R-	E	N
Q201_6	- Identification : variable 4 = V18 = qualitative ordinale (T5) ou quantitative discrète	R+	R-	E	N	R+	R-	E	N
Q201_7	- Identification : variable 5 = V20 = quantitative continue (T6)	R+	R-	E	N	R+	R-	E	N
Q201_8	- Identification : variable 6 = V24 = qualitative ordinale (T7)	R+	R-	E	N	R+	R-	E	N
Q201_9	- Identification : variable 7 = V47B =note = quantitative discrète (T8)	R+	R-	E	N	R+	R-	E	N
Q202_1	- choix d'une variable qualitative (1 ; 3 ; 4; 6)	R+	R-	E	N	R+	R-	E	N
Q202_2	- choix d'un graphique et réalisation	R+	R-	E	N	R+	R-	E	N
Q202_3	- synthèse relative à une variable qualitative	R+	R-	E	N	R+	R-	E	N
Q203	- Interprétation du Tableau T1	R+	R-	E	N	R+	R-	E	N
Q301.1	- moyenne de V47B	R+	R-	E	N	R+	R-	E	N
Q301.2	- variance de V47B	R+	R-	E	N	R+	R-	E	N
Q301.3	- écart-type de V47B	R+	R-	E	N	R+	R-	E	N
Q302	- mode(s)	R+	R-	E	N	R+	R-	E	N
Q303	- conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N
Q401.1	- moyenne de V20 ou V01	R+	R-	E	N	R+	R-	E	N
Q401.2	- variance de V20 ou V01	R+	R-	E	N	R+	R-	E	N
Q401.3	- écart-type de V20 ou V01	R+	R-	E	N	R+	R-	E	N
Q402	- Histogramme	R+	R-	E	N	R+	R-	E	N
Q403	- durée médiane ou âge médian	R+	R-	E	N	R+	R-	E	N
Q404	- proportion estimée dans $[m \cdot S_{\text{échantillon}} ; m \cdot S_{\text{échantillon}}]$	R+	R-	E	N	R+	R-	E	N
Q405	- conclusion et interprétation personnelle	R+	R-	E	N	R+	R-	E	N

Tous les items repérés et sur lesquels les jugements sont portés		validité du résultat				validité du raisonnement					
	modalités du jugement	R+	R-	E	N	R+	R-	E	N		
S01	nombre de...										
	résultat & raisonnement	R+	1	R-	0,5	E	0	N	0		
S02	nombre de...									Score total /66	Note/20
S03	scores...										