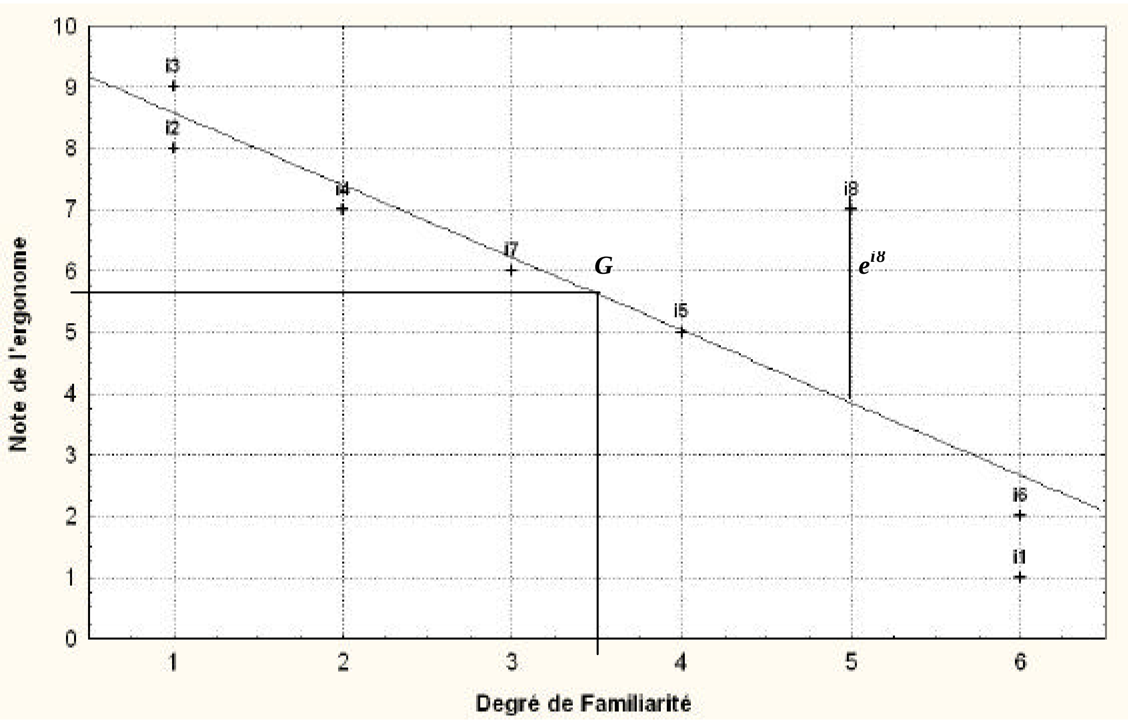
CORRIGÉ MEDI

Durée de l'épreuve : 2 heures
Aucun document n'est autorisé. Seule la calculette (sans sa documentation) est autorisée.
Attention : les exercices (encadrés) sont indépendants. Le barème donné à titre indicatif est sur 60 ; la note finale sera donnée sur 20.
La page 9 (tables et formules) peut être détachée et conservée. Indiquer les réponses exclusivement sur ce document.

DOSSIER ICÔNES (16 POINTS)

Source : données inspirées d'une étude en cours menée par Marion Wolff au Laboratoire d'Ergonomie Informatique de l'Université Paris 5.

Huit icônes informatiques (i1 à i8) ont été au préalable sélectionnées par un groupe d'informaticiens pour figurer dans un nouveau programme informatique. Ces 8 icônes ont ensuite été répertoriées suivant une échelle de familiarité allant de 1 (**symbole très familier**) à 6 (**symbole non familier**) – *Variable X* –. En dernier lieu, elles ont été évaluées par un ergonome, dont la tâche était de donner à chacune d'elles une note comprise entre 0 (icône non appropriée au contexte) et 10 (icône très appropriée au contexte) - *Variable Y* -. Les données sont représentées graphiquement ci-après. On désire analyser la relation entre le degré de familiarité de l'icône et l'appréciation finale donnée par l'ergonome.



Rappel : 1 : symbole très familier 6 : symbole non familier

1) Commenter le nuage de points ci-dessus :
*Le nuage des 8 icônes est descendant et linéaire (**ou** allongé). Ce qui laisse envisager une corrélation forte et négative entre la variable Degré de Familiarité et la variable Note de l'ergonome : plus l'icône est jugée familière, et plus l'ergonome aurait tendance à mettre une note élevée (et vice versa). Par ailleurs, on peut constater la présence d'une valeur relativement atypique : i8 (icône jugée non familière et assez bien notée par l'ergonome).*

2) Reconstituer le tableau de données à partir du nuage de points de la page précédente (toutes les valeurs sont des entiers) :

	i1	i2	i3	i4	i5	i6	i7	i8
Degré de Familiarité *(X)*	**6**	**1**	**1**	**2**	**4**	**6**	**3**	**5**
Note de l'ergonome *(Y)*	**1**	**8**	**9**	**7**	**5**	**2**	**6**	**7**

3) Calculer le coefficient de corrélation linéaire entre Degré de Familiarité et Note de l'ergonome (indiquer seulement le résultat final, **arrondi à 4 décimales**) :

Aides à la vérification des calculs : Moy (X) = 3.50 Moy(Y) = 5.625 r_{bp} *= - 0.87*

r_{bp} = - 0.8671

Pour les exercices suivants, on donnera les résultats arrondis à 2 décimales

4) On cherche à juger, par un indice global, la qualité de la prédiction de la Note de l'ergonome par le Degré de Familiarité de l'icône.

a) Indiquer le nom usuel de cet indice :

Le coefficient de détermination : R^2

b) Donner la valeur de cet indice, en indiquant le mode de calcul :

*$R^2 = r^2 = (-0.8671)^2 = 0.75$ (ou 0.76 si calcul sur r arrondi) (**ou** $R^2 = \frac{Var\,\tilde{y}}{Var\,y}$)*

5) A partir des indices calculés en 3) et 4), que peut-on dire de la liaison entre le Degré de Familiarité de l'icône et la Note de l'ergonome pour ces 8 icônes ?

Pour ces 8 icônes, la corrélation linéaire entre Degré de Familiarité et Note de l'ergonome est négative et importante (r = - 0.87 > 0.40 valeur repère) : plus l'icône est jugée familière et plus l'ergonome aura tendance à donner une note élevée. 75% (ou 76%) de la variance est prise en compte par la régression linéaire. Ce qui peut être considéré comme important (> 16%).

6) Si y désigne la variable Note de l'ergonome :

a) Que désigne $\tilde{y}$: *la variable estimée (ou ajustée) de Note de l'ergonome*

b) Donner la valeur de la moyenne de $\tilde{y}$: $Moy\,\tilde{y} = Moy\,y = 5.63$

7) Donner l'équation de la droite de régression de "Note de l'ergonome" sur "Degré de Familiarité" (indiquer seulement le résultat final, arrondi à 2 décimales) :

$\widetilde{Note} = -1.18 Familiarité + 9.77$ *ou* $\tilde{y} = -1.18x + 9.77$

8) Représenter cette droite sur le nuage de points de la page 1.

9) Indiquer sur le graphique de la page 1, l'écart résiduel pour l'icône *i8*.

10) Donner ci-après les coordonnées du point moyen du nuage (centre de gravité G) et indiquer ce point sur le graphique de la page 1 :

$G = (\bar{x}, \bar{y}) = (3.50, 5.63)$

11) Analyse inférentielle : peut-on généraliser ces résultats à l'ensemble des icônes ayant les mêmes particularités que celles testées dans cette étude ?

a) Indiquer le nom du test que l'on pourrait mettre en œuvre ainsi que le nombre de degrés de liberté qui lui est associé :

Inférence sur une corrélation : test T de Student

ddl = n – 2 = 6

b) Est-il pertinent ici de généraliser ces résultats ? Pourquoi ? (justifier).

Non, car ces données ne sont pas "tirées au hasard". On n'a pas de population parente proprement dite ici.

DOSSIER THÈSE (14 POINTS)

Source : données fournies par Sébastien Georges, UNIC, UPR 2191, CNRS-INAF, Gif-sur-Yvette & Équipe enseignante Metus/Medi

Dans le tableau 1 présenté ci-après figure une partie des résultats issus d'une enquête diffusée sur Internet à l'attention des étudiants actuellement en thèse, en France. Ils devaient indiquer s'ils avaient pratiqué, au cours de leurs études : une activité professionnelle à Temps-plein, à Mi-temps, des "Jobs saisonniers" ou Aucune activité professionnelle. Parmi tous les questionnaires retournés, on a pris au hasard ceux de 386 doctorants de différentes disciplines : Sciences Humaines et Sociales, Sciences "Dures", Droit/Économie, Sciences et Techniques des Activités Physiques et Sportives (STAPS), afin d'analyser les réponses concernant ces deux variables : *Activités Professionnelles* et *Disciplines*.

Tableau 1 : Effectifs observés

Disciplines / Activités Professionnelles	Sciences Humaines et Sociales	Sciences Dures	Droit / Économie	STAPS	Total
Temps plein	60	25	17	11	113
Mi-Temps	44	28	25	26	123
Jobs saisonniers	12	51	24	30	117
Aucune activité	20	4	8	1	33
Total	136	108	74	68	386

<u>Question initiale</u>: *l'exercice d'une activité professionnelle diffère-t-il selon la discipline étudiée ?*

A. Analyse descriptive

1) <u>Analyse des marges</u>. On obtient les pourcentages suivants pour la distribution marginale des Disciplines :

Tableau 2 : Distribution marginale des Disciplines

	Sciences Humaines et Sociales	Sciences Dures	Droit / Économie	STAPS
Profil Moyen	***35%***	*28%*	*19%*	***18%***

a) Compléter les 2 cases manquantes et indiquer la procédure de calcul <u>uniquement</u> pour les Sciences Humaines et Sociales ci-dessous (arrondir à l'entier le plus proche) :

$$SHS = \frac{136}{386} \times 100 = 35\%$$

b) Commenter le profil-moyen des disciplines :

<u>Parmi les 386 étudiants interrogés</u>, les <u>plus nombreux (ou plus d'1/3) sont en Sciences Humaines et Sociales (35%)</u> et en Sciences Dures (28%). Les filières où il y a le moins d'étudiants sont le Droit/économie (19%) et les STAPS (18%).

c) On nomme aussi *profil moyen* cette distribution marginale. Comment peut-on justifier l'utilisation du terme *<u>moyen</u>* ?

*Pour chaque colonne, on peut effectivement retrouver le même résultat en effectuant <u>une moyenne **pondérée**</u> par <u>le total de la ligne</u>, <u>des profils-lignes.</u>*

2) Analyse de la liaison entre les variables Activités Professionnelles et Disciplines.

Tableau 3 : Tableau des Taux de Liaison

Disciplines / Activités Professionnelles	Sciences Humaines et Sociales	Sciences Dures	Droit / Économie	STAPS
Temps plein	***+0.51***	-0.21	-0.22	-0.45
Mi-Temps	+0.02	-0.19	+0.06	*+0.20*
Jobs saisonniers	-0.71	***+0.56***	+0.07	***+0.46***
Aucune activité	***+0.72***	-0.57	+0.26	-0.83

a) Indiquer comment a été calculée la valeur 0.72 en bas à gauche du tableau 3. Rappeler les formules, et indiquer les différentes étapes de calcul. Le détail des calculs sera effectué avec une précision de 4 décimales. Indiquer le résultat final **arrondi à 4 décimales** :

1ère étape : calcul des effectifs théoriques : $\hat{n}_{jk} = \frac{n_j \times n_k}{n} = \frac{136 \times 33}{386} = 11.6269$

Pour chacune des 2 étapes : 0.5 pour formule et 0.5 pour résultat (0 pour résultat TXL si pas 4 décimales)

2ème étape : calcul du Taux de liaison : $t^{jk} = \frac{n_{jk} - \hat{n}_{jk}}{\hat{n}_{jk}} = \frac{20 - 11.6269}{11.6269} = 0.7201$ ***1***

b) A partir de ce tableau 3, commenter les 4 principales attractions :

Pour les 386 doctorants, ceux qui ont tendance à n'exercer aucune activité professionnelle, sont plutôt des étudiants en SHS.Ceux qui ont tendance à favoriser les Jobs saisonniers sont étudiants en Sciences Dures et en STAPS. Ceux qui travaillent à plein temps sont plutôt issus des SHS.

3) On trouve $\phi^2 = 0.1748$. Interpréter cette valeur :

Calcul du V² de Cramer : $\frac{f^2}{f^2_{Max}} = \frac{0.1748}{3} = 0.06$ *(0.0583)*

0.04 < V² < 0.16. Liaison intermédiaire (moyenne) entre les 2 variables

4) Élaborer une conclusion descriptive :

Pour ces 386 doctorants, on peut conclure que l'exercice d'une activité professionnelle diffère selon la discipline étudiée. Le V² de Cramer indique une liaison globale moyenne (V² compris entre les valeurs repères 0.04 et 0.16).

B. Analyse inférentielle

1) On procède à un test du c^2 afin de savoir si le phénomène observé concerne l'ensemble des étudiants, actuellement en thèse, en France, dans une des disciplines concernées.

a) Indiquer la formule de calcul, développer et donner le résultat du calcul :

$c^2_{obs} = n \cdot f^2$; $c^2 = 386 \times 0.1748 = 67.47$ *(valeur exacte : 67.4825)*

b) Donner le nombre de degré de liberté (ddl) associé à ce test (rappeler la formule) :

ddl = (J – 1) (K – 1) ; ddl = (4 – 1) (4 – 1) = 9

c) Donner le résultat du test :

$c^2 = 67.47 > c^2_{[9]\ .001} = 27.88$. *Résultat significatif au seuil .001.*

2) Élaborer une conclusion inférentielle et répondre à la question initiale :

Pour l'ensemble des doctorants en France ayant répondu au questionnaire, et concernés par les disciplines SHS, Sciences Dures, Droit/économie ou STAPS (ou dans la population parente), on peut conclure que l'exercice d'une activité professionnelle diffère selon la discipline étudiée (test du Khi² significatif au seuil .001).

DOSSIER PROPRIÉTÉS (6 POINTS)

Soient une *variable X* "Note sur 15 en Psychologie Clinique" et une *variable Y* "Note sur 10 en Psychologie Différentielle". On désire connaître la relation existant entre ces 2 variables.

On trouve : *Moy (X)* = 11.00 *Moy (Y)* = 9.40 et $r_{(X,Y)}$ = + 0.88
Ety (X) = 2.00 *Ety (Y)* = 0.80

1) Si on multiplie la variable *X* par 6, puis que l'on retire 3, on obtient une nouvelle variable *X'*. Indiquer ci-après, en justifiant, les valeurs de :

a) la moyenne, *Moy (X')* : *63*

La moyenne est multipliée par 6 puis diminuée de 3 : (11 ´ 6) – 3 = 63

b) l'écart-type, *Ety (X')* : *12*

l'écart-type est multiplié par 6 : 2 ´ 6 = 12 ; le fait de retirer la constante 3 n'influence pas l'écart-type.

c) le coefficient de corrélation entre *X'* et *Y*, $r_{(X',Y)}$:

le coefficient de corrélation reste inchangé (au signe près) pour toute transformation affine de l'une ou des 2 variables : $r_{(X',Y)} =$ *+ 0.88*

2) Si on transforme les variables *X* et *Y* en variables centrées réduites, on obtient 2 nouvelles variables *X"* et *Y"*. Indiquer ci-après les valeurs de :

a) la moyenne, *Moy (Y") :* 0

b) l'écart-type, *Ety (Y")* : *1*

c) le coefficient de corrélation entre *X"* et *Y"*, $r_{(X'',Y'')}$ *: + 0.88*

DOSSIER CONSOMOTO1 (12 POINTS)

Source : données fournies par Sébastien Georges, UNIC, UPR 2191, CNRS-INAF, Gif-sur-Yvette & Équipe enseignante Metus/Medi

Lors d'une étude, réalisée par un magazine "deux roues", visant à comparer la consommation de carburant, 62 motocyclettes à moteur 4 temps, de puissances identiques et de marques diverses, ont été choisies au hasard parmi les modèles les plus vendus. Deux groupes de motos de cylindrées différentes ont été constitués : 28 sont à quatre cylindres (groupe 1 : *g1*) et les 34 autres sont des bi-cylindres (groupe 2 : *g2*). On a relevé pour chacune d'entre elles la consommation moyenne pour un même parcours de 1000 kilomètres constitué d'un tracé autoroutier, d'une partie urbaine et d'une partie sur route montagneuse. Les consommations moyennes sont exprimées en litres pour une distance de 100 kilomètres.

Hypothèse: *on suppose que le type de cylindrée a un effet sur la consommation de carburant : les modèles de type 4 cylindres consommeraient moins que les modèles bi-cylindres.*

Critère sémantique : *on jugera un effet important s'il est supérieur ou égal à 2 litres de carburant consommés pour 100 km parcourus, et un effet faible s'il est inférieur ou égal à 1 litre de carburant consommé pour 100 km parcourus.*

Voici les données pour les deux types de cylindrées : groupe des motocyclettes 4 cylindres (*g1*) et groupe des motos bi-cylindres (*g2*) :

n_{g1} = 28	n_{g2} = 34	
m^{g1} = 8.40	m^{g2} = 9.75	*Moyenne Générale* : 9.140
Var^{g1} = 0.6236	Var^{g2} = 4.6801	*Variance Totale* : 3.300

-0.5 pour l'ensemble de l'exercice si pas respect du nb de décimales demandé

A. Analyse descriptive

1) Calculer l'effet moyen du facteur "Type de cylindrée" :

dobs = 9.75 – 8.40 = 1.35

2) Élaborer une conclusion descriptive :

Pour ces 62 motos, on peut constater que la consommation d'essence est moins élevée pour les motos 4 cylindres que pour les motos 2 cylindres. La différence observée (dobs) est de 1.35 l pour 100 km parcourus, et cette différence peut être considérée comme intermédiaire (moyenne) selon le critère sémantique donné (1 < 1.35 < 2).

Tous les calculs et résultats suivants seront indiqués avec une précision de 4 décimales

3) Calculer la variance inter (*Vinter*) de ce protocole et indiquer la procédure de calcul :

$$Vinter = \frac{28 \times (8.40 - 9.140)^2 + 34 \times (9.75 - 9.140)^2}{(28 + 34)} = \frac{27.9842}{62} = 0.4514$$

4) En déduire la valeur de la variance intra (*Vintra*), et justifier :

Vintra = Vtot – Vinter = 3.300 – 0.4514 = 2.8486

5) On a trouvé η² = 0.14.

a) Indiquer la formule qui a permis de calculer cet indice, et donner la valeur arrondie à 4 décimales.

$$h^2 = \frac{V\,int\,er}{Vtot} = \frac{0.4514}{3.300} = 0.1368$$

b) Interpréter η² :

Le facteur "type de cylindrée" prend en compte (explique) 14% de la dispersion (variance) des consommations, ce qui peut être considéré comme moyen (intermédiaire) selon les valeurs repères données (0.04 < 0.14 < 0.16).

B. Analyse inférentielle : on aimerait généraliser les résultats obtenus sur cet échantillon à la population parente.

1) Indiquer quelle est la population parente :

Toutes les motos de type 4 cylindres et 2 cylindres

2) On a utilisé un test pour cette analyse inférentielle. La valeur trouvée pour ce test est 3.08.

a) Indiquer le nom du test que l'on a utilisé, la formule, et le détails des calculs qui a permis de trouver ce résultat :

Test T de Student ; $t = \sqrt{n-2} \times \sqrt{\frac{V_{inter}}{V_{intra}}} = \sqrt{62-2} \times \sqrt{\frac{0.4514}{2.8486}} = 3.08$

b) Quel est le nombre de degrés de liberté (ddl) associé à ce test ?

n – 2 = 62 – 2 = 60

c) Indiquer le résultat du test :

tobs = 3.08 > $t_{[60].01}$ *= 2.66 ; test significatif au seuil .005* ***unilatéral***

d) Élaborer une conclusion inférentielle :

D'une manière générale, les motos 4 cylindres consomment moins d'essence que les motos bicylindres (test T de Student significatif au seuil .005 unilatéral).

DOSSIER CONSOMOTO2 (6 POINTS)

Source : données fournies par Sébastien Georges, UNIC, UPR 2191, CNRS-INAF, Gif-sur-Yvette & Équipe enseignante Metus/Medi

Quelques mois plus tard, le même magazine cité page 5 (*voir Dossier Consomoto1*) apporte un complément d'informations concernant la consommation de carburant des motocyclettes. Le journaliste prend en compte cette fois-ci, en plus des quatre cylindres (groupe 1 : *g1*) et des bi-cylindres (groupe 2 : *g2*), les motos équipées d'un moteur trois cylindres (groupe 3 : *g3*). Vingt et une motocyclettes de ce nouveau groupe (*g3*) ont été choisies au hasard. Ces dernières répondent aux mêmes critères que celles des deux autres catégories concernant les caractéristiques moteur, et sont également de marques diverses. On a relevé pour chaque type de cylindrée la consommation moyenne pour un même parcours de 1000 kilomètres constitué d'un tracé autoroutier, d'une partie urbaine et d'une partie sur route montagneuse. Les consommations moyennes sont exprimées en litres pour une distance de 100 kilomètres.

Les statistiques élémentaires de ces 3 groupes sont indiquées ci-après :

$n_{g1} = 28$	$n_{g2} = 34$	$n_{g3} = 21$
$m^{g1} = 8.400$	$m^{g2} = 9.750$	$m^{g3} = 9.000$
$Var^{g1} = 0.6236$	$Var^{g2} = 4.6801$	$Var^{g3} = 4.000$

$Vintra = 3.1396$ $\quad$ $Vinter = 0.3409$

Hypothèse *: on suppose que la consommation d'essence diffère selon le type de cylindrée utilisé (4 cylindres, 2 cylindres ou 3 cylindres).*

Les résultats ci-après seront présentés arrondis à 4 décimales.

-0.5 pour l'ensemble si pas présentation à 4 décimales

A. Analyse descriptive

1) Calculer la variance totale de ces 3 groupes (donner le détail des calculs) :

Vintra + Vinter = Vtot ***ou*** *3.1396 + 0.3409 = 3.4805*

2) On cherche à savoir si le facteur "Type de cylindrée" pourrait expliquer la dispersion des consommations de carburant.

a) Indiquer ci-après l'indice pertinent à calculer, donner la procédure de calcul et donner le résultat :

$$\text{Eta}^2 = \eta^2 = \frac{Vinter}{Vtot} = 0.0979 \quad \textbf{ou} \quad \eta^2 = \frac{0.3409}{3.4805} = 0.0979$$

b) Élaborer une conclusion descriptive :

Pour les 83 motos testées, on peut conclure que 10% de la dispersion des consommations est expliqué par le facteur "Type de cylindrée" (certaines motos consomment plus que d'autres, du fait qu'elles n'ont pas la même cylindrée). Ce pourcentage de dispersion pris en compte par ce facteur peut être considéré comme moyen en regard des valeurs-repères données (0.04 < 0.10 < 0.16).

B. Analyse inférentielle : on cherche à généraliser les résultats obtenus sur les échantillons à l'ensemble des motos ayant les mêmes caractéristiques que celles testées.

1) Indiquer précisément le nom du test à mettre en œuvre pour répondre à cette question :

Test F de Fisher-Snedecor

a) Indiquer les degrés de liberté associé à ce test (donner le détail des calculs) :

ddl1 = K – 1 = 3 – 1 = 2 ; ddl2 = n – K = 83 – 3 = 80

b) La valeur trouvée pour ce test est : *4.34*. Indiquer la formule qui a permis de trouver ce résultat :

$$F = \frac{(n-K)}{(K-1)} \times \frac{V_{inter}}{V_{intra}} \quad \textbf{ou} \quad F = \frac{80}{2} \times \frac{0.3409}{3.1396} = 4.34$$

c) Donner le résultat du test :

$F_{obs} = 4.34 > F_{[2,\, 80].05} = 3.11$; test significatif au seuil .05

2) Élaborer une conclusion inférentielle :

Pour l'ensemble des motos bicylindres, à 3 cylindres ou à 4 cylindres, on peut conclure que la consommation de carburant diffère selon la cylindrée utilisée (test F significatif au seuil .05).

DOSSIER MINCEUR (6 POINTS)

On demande à 150 femmes, échantillonnées au hasard en région parisienne parmi un ensemble de volontaires, de tester pendant *une semaine* une nouvelle crème amincissante spéciale "zones difficiles", selon différents critères : odeur et texture, hydratation, raffermissement des tissus, début d'amincissement … Au terme de cette semaine, on relève l'appréciation globale des utilisatrices : plutôt positive (+) ou plutôt négative (-). Un mois après cette première utilisation, on leur demande de tester à nouveau cette même crème, mais cette fois pendant *3 semaines*. Au terme de ce deuxième test, on relève leur nouvelle appréciation (plutôt positive ou négative). Ci-après, figure le tableau de résultats :

		2nd Test +	**2nd Test** -	
1er Test	+	65	38	*103*
	-	32	15	*47*
		97	*53*	150

Question *: la durée de la période d'utilisation de la crème amincissante a-t-elle un effet sur l'appréciation du produit ?*

Critère sémantique *: on jugera un effet important s'il est supérieur ou égal à 10 points de pourcentage, faible s'il est inférieur ou égal à 5 points de pourcentage.*

1) Analyse descriptive

a) Pour ces 150 femmes, évaluer l'effet "durée de la période d'utilisation de la crème amincissante " :

$f_{t1+} = \frac{65 + 38}{150} = \frac{103}{150} = 0.687\ (0.69)$; $f_{t2+} = \frac{65 + 32}{150} = \frac{97}{150} = 0.647\ \ (0.65)$

$d_{obs} = 0.687 - .0.647 = 0.04$ » .4%

[ou $d_{obs} = \frac{38 - 32}{150} = 0.04$ » 4%]

b) Rédiger une conclusion descriptive :

Pour ces 150 femmes volontaires, on peut constater que la fréquence des personnes satisfaites a légèrement diminué entre le 1er test (» 69%) et le 2nd test (» 65%). La différence est de 4 points de pourcentage et peut être considérée comme faible en regard du critère sémantique donné (4% < 5%). ***(OK si conclusion avec les insatisfaites)***

2) Analyse inférentielle : on désire généraliser ces résultats à la population parente.

a) Quelle est la population parente ?

Les femmes de la région parisienne (volontaires)

b) Donner précisément le nom du test que l'on va utiliser, ainsi que le nombre de degrés de liberté associé à ce test :

Test du Khi² corrigé de Mac Nemar à 1 ddl

c) Ce test donne *un résultat non significatif*. Que peut-on en conclure ?

Bien que l'on ait constaté pour l'échantillon une légère diminution de personnes satisfaites entre le 1er test et le 2nd test, quant à l'appréciation de la nouvelle crème amincissante, on ne peut pas conclure qu'il en ira de même (ou constat d'ignorance) pour la population parente (l'ensemble des femmes de région parisienne) (test du Khi² corrigé de McNemar non significatif au seuil .05).

- Extrait de la table des valeurs critiques de la variable c^2 :

a / **ddl**	**.05**	**.01**	**.001**
1	3.84	6.63	10.83
2	5.99	9.21	13.82
3	7.81	11.34	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.12
9	16.92	21.67	27.88
10	18.31	23.21	29.59

- Extrait de la table des valeurs critiques de la variable *T* de Student :

a **/2** a **ddl**	**.025** **.05**	**.005** **.01**	**.0005** **.001**
25	2.060	2.787	3.725
26	2.056	2.779	3.707
27	2.052	2.771	3.690
28	2.048	2.763	3.674
29	2.045	2.756	3.659
30	2.042	2.750	3.646
40	2.021	2.704	3.551
60	2.000	2.660	3.460
120	1.980	2.617	3.373
30000	1.960	2.576	3.291

- Extrait de la table des valeurs critiques de la variable *F* de Snedecor :

Ddl1 / ***Ddl2***	*a*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
50	.05	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.92	1.89	1.87
	.01	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.63	2.56	2.51	2.46	2.42
	.001	12.2	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82	3.67	3.55	3.44	3.35	3.27	3.20
60	.05	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.89	1.86	1.84
	.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.44	2.39	2.35
	.001	12.0	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69	3.54	3.42	3.32	3.23	3.15	3.08
80	.05	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.91	1.88	1.84	1.82	1.79
	.01	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.64	2.55	2.48	2.42	2.36	2.31	2.27
	.001	11.7	7.54	5.97	5.12	4.58	4.20	3.92	3.70	3.53	3.39	3.27	3.16	3.07	3.00	2.93
100	.05	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.82	1.79	1.77
	.01	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.43	2.37	2.31	2.27	2.22
	.001	11.5	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.44	3.30	3.18	3.07	2.99	2.91	2.84

- Quelques formules... $t_{obs} = \sqrt{n-2} \times \sqrt{\dfrac{V_{inter}}{V_{intra}}}$ $\qquad$ $\boldsymbol{t_{obs}} = \sqrt{n-2} \times \dfrac{r}{\sqrt{1-r^2}}$

$$F = \frac{(n-K)}{(K-1)} \times \frac{V_{inter}}{V_{intra}} \qquad dd1 = K - 1 \qquad ddl2 = n - K$$

$$r = \frac{Cov(x\ ,y)}{Ety\ x\ \ Ety\ y} \qquad Cov(x,y\ ) = \frac{\sum_i (x^i\ y^i)}{n} - \bar{x}\ \bar{y} \qquad a = \frac{Cov(x\ ,y\ )}{Var\ x} = r \times \frac{Ety\ y}{Ety\ x} \qquad b = \bar{y} - a\ \bar{x}$$

$$t^{jk} = \frac{n_{jk} - \hat{n}_{jk}}{\hat{n}_{jk}} \qquad F^2 = \sum_{J,K} Cta_{jk} \qquad c^2_{obs} = \frac{(|n_{+-} - n_{-+}| - 1)^2}{n_{+-} + n_{-+}}$$

Les différents calculs ont été réalisés à l'aide des logiciels Statistica et DS3Win