

EXAMEN FINAL AUTOMNE 2005

Date : Dimanche 18 décembre 2005, de 14h00 à 17h00

INSTRUCTIONS

1. Détachez la feuille-réponses à la fin de ce cahier et inscrivez-y *immédiatement* votre nom, votre code permanent et votre numéro de groupe.
2. Seule la feuille-réponses doit être remise. Vous y inscrivez vos réponses sous la forme d'une lettre majuscule correspondant à votre choix.
3. Tout texte de référence (manuel, notes de cours, notes personnelles, etc.) est interdit. **Tout cas de plagiat ou de fraude sera soumis au Comité de discipline.**
4. L'usage d'une calculatrice est autorisé.
5. L'étudiant doit placer sa carte d'étudiant (avec photo) sur la table et signer la feuille de présence lors de la remise de sa feuille-réponses.
6. Pas de téléphone cellulaire sur la table.

Problème 1 [10 points]

On tire un échantillon aléatoire simple de n écoles parmi les N écoles (privées et publiques) d'une ville.

Dans chacun des cas énumérés dans la 2^e colonne (*Paramètre à estimer*) du tableau ci-dessous, identifier le paramètre qu'il s'agit d'estimer. Faites votre choix de réponse dans la liste de gauche.

Liste des réponses possibles	Paramètre à estimer [répondre par A, B, ... , ou H]
A: La moyenne μ_y d'une certaine variable Y	1-a) La proportion de filles parmi les élèves de la ville
B: La moyenne μ_d d'une certaine variable Y dans un domaine \mathcal{D} .	1-b) Le nombre total d'élèves dans la population
C: Le total τ_d d'une certaine variable Y dans un domaine \mathcal{D} .	1-c) Le nombre d'écoles privées
D: Le total τ_y d'une certaine variable Y	1-d) La proportion d'écoles privées
E: Le nombre N_c d'unités appartenant à une certaine classe \mathcal{C}	1-e) Le nombre moyen d'élèves par école
F: Le quotient $R = \mu_Y/\mu_X$ de deux variables Y et X	1-f) Le nombre moyen d'élèves par professeur
G: La proportion p d'unités appartenant à une certaine classe	1-g) Le nombre moyen d'élèves dans les écoles privées
H: Aucun des paramètres ci-dessus	1-h) Le nombre total d'élèves dans les écoles privées
	1-i) La note moyenne en français des élèves allophones de la population
	1-j) Le nombre total d'ordinateurs possédés par les garçons de la population

Question 2 [32 points]

On tire un échantillon de **30** comptes de dépenses d'une population de **400** comptes de dépenses soumis par des employés pour voyages d'affaires. Pour chaque compte on détermine Y le montant du compte et X la durée du voyage (en nombre de jours):

Y : Montant du compte

X : Durée du voyage

On note également si la destination est *internationale* ou *nationale*.

Les données de l'échantillon sont présentées en annexe [Veuillez utiliser les écarts-types, les moyennes et la covariance tels que présentés, sans plus de décimales]

[Choisissez vos réponses parmi celles proposées au bas de la page]

- 2-a) Estimer le nombre total de voyages internationaux dans la population
- 2-b) Estimer le montant moyen des comptes, utilisant l'estimateur par le quotient et le fait que la durée moyenne d'un voyage dans la population est de 3,9 jours.
- 2-c) Estimer le montant total des comptes pour voyages *internationaux*.
- 2-d) Estimer l'écart-type de l'estimateur en 2-a)
- 2-e) Estimer l'écart-type de l'estimateur en 2-b).
- 2-f) Estimer l'écart-type de l'estimateur en 2-c).
- 2-g) Estimer la taille de l'échantillon qu'il aurait fallu prélever pour estimer (par la moyenne) le montant moyen des comptes avec une marge d'erreur de 10 %.
- 2-h) Estimer la taille de l'échantillon qu'il aurait fallu prélever pour estimer le montant total des comptes *pour voyages internationaux* avec une marge d'erreur de 30 % [Vous supposerez qu'on ne peut pas identifier à l'avance les comptes avec voyages internationaux].

Question 3 [20 points]

Une population de commerces de détail est répartie en 3 strates. Le tableau suivant présente les tailles de ces strates ainsi qu'une estimation des écarts-types (corrigés) de la variable « nombre d'employés ».

	<i>Strate</i>		
	<i>Strate 1</i>	<i>Strate 2</i>	<i>Strate 3</i>
<i>Nombre d'unités dans la strate</i>	8000	5000	2000
<i>Écart-type corrigé de la strate</i>	25	40	120

On compte tirer un échantillon stratifié de taille 720 afin d'estimer le nombre moyen d'employés.

- 3-a) Estimer l'écart-type de l'estimateur lorsqu'on utilise l'allocation proportionnelle.
- 3-b) Estimer l'écart-type de l'estimateur lorsqu'on utilise l'allocation optimale.

Choix de réponse [pour les questions 2 et 3]

A 0,07 - 0,09	B 0,19 - 0,21	C 1,5 - 1,6	D 1,9-2,1	E 26,9 - 29,0	F 41 - 42
G 74 - 76	H 79 - 81	I 91 - 93	J 128 - 130	K 189 - 192	L 646-648
M 749 - 751	N 6951 - 6954	O 16740 - 16743	P 22 955 - 22 960	Q 28 790-28 795	R 35 610 - 35 625
S 98 000 - 98 076	T 258 740 - 259 445	U 285 270 - 285 282	V 302 350-303 251	W 452 340 - 452 360	Aucune des réponses ci-dessus

Question 4 [15 points]

Une parfumerie qui produit trois fragrances d'eau de toilette — *Eaux du Nil* (20 % des ventes d'eau de toilette), *Jardins des îles* (30 % des ventes), et *Volupté* (50 % des ventes) — décide de refaire les designs des flacons. Un an plus tard, afin d'évaluer les effets, elle prélève les données suivantes sur les ventes (après le changement) de 300 flacons de cette eau de toilette.

	Flacon			Total
	Eaux du Nil	Jardins des îles	Volupté	
Effectif (Nombre de flacons vendus)	81	90	129	300

Entre autres, on voudrait déterminer si les nouveaux designs ont modifié la distribution des choix des consommateurs de ce produit.

4-a) Dites lequel ou lesquels des énoncés suivants est (sont) acceptable(s) comme hypothèse nulle: [Faites votre choix dans la liste présentée plus bas.]

H_{01} : La distribution a changé depuis l'introduction des nouveaux designs

H_{02} : La distribution n'a pas changé depuis l'introduction des nouveaux designs

H_{03} : La distribution a-t-elle changé depuis l'introduction des nouveaux designs?

H_{04} : La distribution des ventes depuis l'introduction des nouveaux designs est (en fréquences) : 81/300; 90/300 ; 129/300.

H_{05} : La distribution des ventes depuis l'introduction des nouveaux designs est (en fréquences) : 0,2; 0,3 ; 0,5.

Choix de réponse [pour la question 4-a)]

Les conclusions suivantes sont justifiées [Inscrire l'une des lettres A à J]:

A H_{01} seulement	B H_{02} seulement	C H_{01} et H_{03} seulement	D H_{01} et H_{04} seulement	E H_{02} et H_{03} seulement
F H_{02} et H_{05} seulement	G H_{03} et H_{04} seulement	H H_{03} et H_{05} seulement	I H_{01} , H_{03} et H_{04} seulement	J H_{02} , H_{03} et H_{05} seulement

4-b) Déterminer le tableau des effectifs théoriques (choisir le tableau A, B, C, D, E ou F) :

A	Eaux du Nil	Jardins des îles	Volupté
Effectif	81	90	129

B	Eaux du Nil	Jardins des îles	Volupté
Effectif	100	100	100

C	Eaux du Nil	Jardins des îles	Volupté
Effectif	60	90	150

D	Eaux du Nil	Jardins des îles	Volupté
Effectif	20	30	50

E	Eaux du Nil	Jardins des îles	Volupté
Effectif	0,2	0,3	0,5

F	Eaux du Nil	Jardins des îles	Volupté
Effectif	27	30	43

4-c) Calculer la valeur de khi-deux et tirer la conclusion qui en découle [Choisir la combinaison χ^2 /conclusion qui s'applique et inscrire A, B, ..., ou R]

Conclusion	Valeur observée de khi-deux					
	3,4-3,5	10,2-10,3	13,0-14,0	132-135	399-401	429-431
On peut conclure que la distribution a changé depuis le changement de design	A	B	C	D	E	F
On ne peut <i>pas</i> conclure que la distribution a changé depuis le changement de design	G	H	I	J	K	L
Ni l'une ni l'autre des conclusions ci-dessus n'est justifiée	M	N	O	P	Q	R

Question 5 [8 points]

Pour chacune des descriptions suivantes, dire de quel mode d'échantillonnage il s'agit. Choisir une réponse parmi les suivantes:

A: aléatoire simple

B: stratifié

C: systématique

D: par grappes avec probabilités de sélection égales

E: par grappes avec probabilités de sélection inégales.

5-a) Population: l'ensemble des formulaires soumis aux autorités douanières par des passagers arrivant à un aéroport. À partir d'un moment choisi au hasard, on tire chaque 100 ^e formulaire.
5-b) Population: l'ensemble des PME d'un pays. On divise le pays en 35 régions; puis on tire 5 régions au hasard. L'échantillon comprend toutes les PME des 5 régions choisies.
5-c) Population: l'ensemble des employés d'une compagnie. On dresse une liste de tous les employés, puis on tire successivement et sans remise 300 employés.
5-d) Population: l'ensemble des comptes mensuels de dépenses soumis par les employés d'une compagnie au courant d'une année. On tire trois comptes pour chacun des 12 mois (donc 36 comptes au total).
5-e) Population : l'ensemble des PME d'un pays. On divise le pays en 35 régions; puis on tire 2 PME dans chaque région.
5-f) Population: l'ensemble des étudiants de l'UQÀM. On tire au hasard trois programmes, et on inclut dans l'échantillon tous les étudiants inscrits dans ces trois programmes
5-g) Population: l'ensemble des logements d'une petite ville. On divise la ville en 5 secteurs, puis on tire au hasard 15 logements dans chacun des 5 secteurs.
5-h) Population: l'ensemble des logements d'une ville. On divise la ville en quartiers, puis, sur un plan de la ville, on choisit au hasard n points. L'échantillon est l'ensemble des logements situés dans les quartiers touchés par les points.

Question 6 [5 points]

Lesquelles des affirmations suivantes sont vraies?

P_1 : L'estimation de la moyenne dans un échantillon systématique n'a pas de variance

P_2 : Un estimateur sans biais est un estimateur qui est égal au paramètre

P_3 : Un échantillon stratifié est toujours plus précis qu'un échantillon aléatoire simple, quelle que soit l'allocation

Choix de réponses [pour la question 6]

A : Aucune n'est vraie	B : Seule P_1 est vraie	C : Seule P_2 est vraie	D : Seule P_3 est vraie
E : Seules P_1 et P_2 sont vraies	F : Seules P_1 et P_3 sont vraies	G : Seules P_2 et P_3 sont vraies	H : Toutes sont vraies

Question 7 [5 points]

Lesquelles des affirmations suivantes sont vraies?

P_1 : L'allocation optimale dans un échantillon stratifié est parfois moins précise que l'allocation proportionnelle

P_2 : L'estimateur par le quotient n'est pas sans biais

P_3 : L'estimateur par la différence n'est pas sans biais

Choix de réponses [pour la question 7]

A : Aucune n'est vraie	B : Seule P_1 est vraie	C : Seule P_2 est vraie	D : Seule P_3 est vraie
E : Seules P_1 et P_2 sont vraies	F : Seules P_1 et P_3 sont vraies	G : Seules P_2 et P_3 sont vraies	H : Toutes sont vraies

Question 8 [5 points]

Choisir dans la colonne de gauche l'expression mathématique qui complète correctement la phrase dans la colonne centrale. Inscire l'une des lettres A à O sur la feuille-réponses.

A : $\mu_{\hat{\sigma}} = \sigma$	8-a) Dire que l'estimateur de μ par la différence $\hat{\mu}_{y_d}$ est meilleur que la moyenne arithmétique, c'est dire que...
B : $\mu_{\hat{\sigma}_1} < \mu_{\hat{\sigma}_2}$	
C : $\hat{\mu}_{y_d} < \bar{y}$	8-b) On a deux estimateurs sans biais, $\hat{\sigma}_1$ et $\hat{\sigma}_2$, de l'écart-type σ de la population. Dire que $\hat{\sigma}_1$ est le meilleur des deux c'est dire que...
D : $\mu_{\bar{y}} = \mu$	
E : $\mu_{\hat{\mu}_{y_d}} = \mu$	8-c) Dire qu'un estimateur $\hat{\sigma}$ de σ (l'écart-type de la population) est sans biais, c'est dire que...
F : $\mu_{\hat{\sigma}_1} = \mu_{\hat{\sigma}_2}$	
G : $\hat{\mu}_{y_d} < \mu$	8-d) Dire que la moyenne \bar{y} de l'échantillon est sans biais pour la moyenne μ , c'est dire que...
H : $\sigma_{\hat{\sigma}_1} < \mu_{\hat{\sigma}_2}$	
I : $\sigma_{\hat{\sigma}_1} < \sigma_{\hat{\sigma}_2}$	8-e) Dire que $\hat{\sigma}_{\bar{y}}^2$ est un estimateur sans biais de la variance de \bar{y} c'est dire que...
J : $\mu_{\hat{\sigma}} = \mu$	
K : $\mu_{\hat{\sigma}_{\bar{y}}}^2 = \sigma_{\bar{y}}^2$	
L : $\sigma_{\hat{\mu}_{y_d}} < \sigma_{\bar{y}}$	
M : $\sigma_{\hat{\sigma}_{\bar{y}}} = \sigma_{\bar{y}}$	
N : $\hat{\sigma}_{\bar{y}} = \sigma_{\bar{y}}$	
O : Aucun de ces énoncés	

Annexe

Données sur un échantillon de taille **30** tiré d'une population de **400** comptes de dépenses.

#	Nombre de jours	Montant dépensé	Destination
1	11	1599,11	Internationale
2	9	1208,56	Internationale
3	8	1031,36	Internationale
4	7	1350,02	Internationale
5	7	1007,72	Internationale
6	7	1158,85	Internationale
7	6	993,47	Nationale
8	6	451,57	Nationale
9	5	790,13	Nationale
10	5	712,35	Nationale
11	5	493,33	Nationale
12	4	920,09	Nationale
13	4	957,7	Nationale
14	4	651,86	Nationale
15	4	363,88	Nationale
16	3	313,53	Nationale
17	3	217,47	Nationale
18	3	821,42	Nationale
19	3	593,07	Nationale
20	3	596,24	Nationale
21	3	567,72	Nationale
22	3	390,57	Nationale
23	3	913,16	Nationale
24	2	677,14	Nationale
25	2	259,46	Nationale
26	2	572,98	Nationale
27	2	548,42	Nationale
28	2	438,82	Nationale
29	2	408,44	Nationale
30	1	387,56	Nationale
Moyennes	4,3	713,2	
Écart-types (corrigés)	2,39	341,94	
Covariance	654,91		

Formulaire MAT2080 Examen final

Résumé des paramètres, leur estimateur, l'écart-type de l'estimateur, et l'estimateur de l'écart-type de l'estimateur.

Paramètre	Estimateur	Écart-type de l'estimateur	Estimateur de l'écart-type de l'estimateur
Moyenne μ	\bar{y}	$\sigma_{\bar{y}} = \sqrt{1-f} \frac{S}{\sqrt{n}}$	$\hat{\sigma}_{\bar{y}} = \sqrt{1-f} \frac{s}{\sqrt{n}}$
Proportion p	$\hat{p} = \frac{X}{n}$	$\sigma_{\hat{p}} = \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{p(1-p)}{n}}$	$\hat{\sigma}_{\hat{p}} = \sqrt{1-f} \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}$
Un quotient $R = \frac{\mu_y}{\mu_x}$	$\hat{R} = \frac{\bar{y}}{\bar{x}}$	$\sigma_{\hat{R}} \approx \frac{\sqrt{1-f}}{\mu_x} \frac{\sqrt{S_y^2 + R^2 S_x^2 - 2RS_{xy}}}{\sqrt{n}}$	$\hat{\sigma}_{\hat{R}} = \frac{\sqrt{1-f}}{\bar{x}} \frac{\sqrt{s_y^2 + \hat{R}^2 s_x^2 - 2\hat{R}s_{xy}}}{\sqrt{n}}$
Moyenne μ Estimation par la différence	$\hat{\mu}_{yd} = \mu_x + (\bar{y} - \bar{x})$	$\sigma_{\hat{\mu}_{yd}} = \sqrt{1-f} \frac{\sqrt{S_y^2 + S_x^2 - 2S_{xy}}}{\sqrt{n}}$	$\hat{\sigma}_{\hat{\mu}_{yd}} = \sqrt{1-f} \frac{\sqrt{s_y^2 + s_x^2 - 2s_{xy}}}{\sqrt{n}}$
Moyenne μ Estimation par le quotient	$\hat{\mu}_{yq} = \mu_x \hat{R}$	$\sqrt{1-f} \frac{\sqrt{S_y^2 + R^2 S_x^2 - 2RS_{xy}}}{\sqrt{n}}$	$\sqrt{1-f} \frac{\sqrt{s_y^2 + \hat{R}^2 s_x^2 - 2\hat{R}s_{xy}}}{\sqrt{n}}$
Moyenne μ_d d'un domaine \mathfrak{D}	\bar{y}_d : Moyenne du domaine dans l'échantillon		$\sqrt{1-\frac{n_d}{N_d}} \frac{s_d}{\sqrt{n_d}}$ ou $\sqrt{1-\frac{n}{N}} \frac{s_d}{\sqrt{n_d}}$ selon que N_d est connu ou pas
Total $\tau_d = N_d \mu_d$ d'un domaine (N_d connu)	$T_d = N_d \bar{y}_d$		$N_d \sqrt{1-\frac{n_d}{N_d}} \frac{s_d}{\sqrt{n_d}}$
Total $\tau_d = N_d \mu_d$ d'un domaine (N_d inconnu)	$\hat{T}_d = \hat{N}_d \bar{y}_d = N \bar{y}'$ où $\hat{N}_d = \frac{n_d}{n} N$		$N \sqrt{1-f} \frac{s'}{\sqrt{n}}$

$$f = \frac{n}{N}$$

Taille d'échantillon

Estimation de la moyenne

La taille d'échantillon nécessaire pour que la marge d'erreur absolue soit égale à E est

$$n = \frac{n_o}{1 + \frac{n_o}{N}} \quad \text{où } n_o = \left(\frac{2S}{E} \right)^2.$$

La taille d'échantillon nécessaire pour que la marge d'erreur relative soit égale à R est

$$n = \frac{n_o}{1 + \frac{n_o}{N}} \quad \text{où } n_o = \left(\frac{2S}{R\mu} \right)^2.$$

Estimation d'une proportion p

Pour estimer une proportion p de telle sorte que la marge d'erreur absolue soit égale à E , la taille approximative de l'échantillon qu'il faut tirer est donnée par $n = \frac{n_o}{1 + \frac{n_o}{N}}$ où $n_o = \frac{4p(1-p)}{E^2}$.

Pour estimer une proportion p de telle sorte que la marge d'erreur *relative* soit égale à R , la taille approximative de l'échantillon qu'il faut tirer est donnée par $n = \frac{n_o}{1 + \frac{n_o}{N}}$ où $n_o = \frac{4(1-p)}{R^2 p}$.

Échantillonnage par strates

L'estimateur de la moyenne dans un échantillon stratifié est $\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$.

Son écart type est $\sigma_{\bar{y}_{st}} = \sqrt{\sum_{h=1}^L W_h^2 \sigma_{\bar{y}_h}^2}$ où $\sigma_{\bar{y}_h}^2 = (1-f_h) \frac{S_h^2}{n_h}$ et $f_h = n_h/N_h$.

L'estimateur d'une proportion dans un échantillon stratifié est $\hat{p}_{st} = \sum_{h=1}^L W_h \hat{p}_h$.

Son écart-type est estimé par $\hat{\sigma}_{\hat{p}_{st}} = \sqrt{1-f_h} \sqrt{\frac{\hat{p}_h(1-\hat{p}_h)}{n_h-1}}$.

L'allocation optimale pour l'estimation d'une moyenne dans un échantillon stratifié est donnée par

$$n_h \text{ proportionnels aux } W_h S_h$$

Test du khi-deux

$$\chi^2 = \sum \frac{(O_i - T_i)^2}{T_i}$$

Points critiques ($\alpha = 5\%$) d'une loi khi-deux

v	χ_v^2	v	χ_v^2	v	χ_v^2	v	χ_v^2
1	3,8415	6	12,5916	11	19,6751	16	26,2962
2	5,9915	7	14,0671	12	21,026	17	27,5871
3	7,8147	8	15,5073	13	22,362	18	28,8693
4	9,4877	9	16,919	14	23,6848	19	30,1435
5	11,0705	10	18,307	15	24,9958	20	31,4104

Feuille-réponses

/100

Ne rien écrire ici

MAT 2080 **Examen final automne 2005** **Méthodes statistiques**

Dimanche 18 décembre 2005, de 14h00 à 17h00

Nom :

Prénom :

Code permanent :

Groupe:

<i>Question</i>	<i>Réponse</i>	
1-a)	F	1
1-b)	D	1
1-c)	E	1
1-d)	G	1
1-e)	A	1
1-f)	F	1
1-g)	B	1
1-h)	C	1
1-i)	F	1
1-j)	D	1
2-a)	H	4
2-b)	L	4
2-c)	S	4
2-d)	E	4
2-e)	F	4
2-f)	R	4
2-g)	G	4
2-h)	J	4
3-a)	D	10
3-b)	C	10

<i>Question</i>	<i>Réponse</i>	
4-a)	F	5
4-b)	C	5
4-c)	B	5
5-a)	C	1
5-b)	D	1
5-c)	A	1
5-d)	B	1
5-e)	B	1
5-f)	D	1
5-g)	B	1
5-h)	E	1
6	A	5
7	C	5
8-a)	L	1
8-b)	I	1
8-c)	A	1
8-d)	D	1
8-e)	K	1