Université du Québec à Montréal
Département de mathématiques

Corrigé

MAT 2080 — STATISTIQUE EN GESTION 1

EXAMEN FINAL AUTOMNE 2002

Date : Mardi 17 décembre 2002, de 18h00 à 21h00

Nom :

Prénom :

Code permanent

Groupe:

INSTRUCTIONS

1. Prendre grand soin de ne pas désassembler les feuilles du présent cahier (9 pages + formulaire + tables), qui doit être remis en entier. Seuls l'annexe, le formulaire et les tables peuvent être détachés du cahier et n'ont pas à être retournés.
2. Par mesure de précaution, inscrire lisiblement votre nom au haut de chacune des pages 2 à 9.
3. Les solutions doivent être rédigées dans les espaces prévus. Ne pas négliger d'expliquer clairement votre démarche, de donner les détails de vos calculs et d'identifier clairement les variables considérées.
4. Si l'espace est insuffisant, indiquer clairement au correcteur que la solution est continuée au verso de la page.
5. Tout texte de référence (manuel, notes de cours, notes personnelles, etc.) est interdit. ***Tout cas de plagiat ou de fraude sera sévèrement sanctionné par les hautes instances universitaires.***
6. Vous trouverez à la fin de ce cahier deux feuilles blanches, pour fins de calcul-brouillon.
7. L'usage d'une calculatrice est autorisé.
8. L'étudiant doit présenter sa carte d'étudiant (avec photo) lors de la remise de son cahier et signer la feuille de présence.

Grille à l'usage du correcteur

1-a)-e)	1-f)-g)	2	3	4	5	6
/29	/8	/10	/18	/18	/7	/10

Note finale : /100

Question 1

Un vérificateur constate que certains des 1 000 comptes à recevoir pour lesquels le paiement était dû le mois dernier sont en souffrance. Il prélève un échantillon de 20 factures afin de déterminer le montant de ces comptes. Le tableau suivant présente les montants des factures et quelques calculs.
Les données suivantes présentent les montants d'un échantillon de 20 factures tiré d'une population de 1000 factures.Les astérisques indiquent les factures impayées.

839,70*	1054,63	1576,29*	821,64	847,32	678,56*	954,58	1546,81*	1552,34*	741,43*
180,86	493,47	886,08	1280,73	791,04	773,98	533,94*	1224,76	430,9	1158,69

$\sum_{i=1}^{20} y_i$ = 18367,75 ; $\bar{y}$ = 918.3875 ; s = 383,1939; et 1orsque dans le tableau ci-dessus on remplace par 0 tous les montants *payés*, on obtient l'écart-type (corrigé) suivant : 584,8108.
Comptes impayés de l'échantillon : Nombre : 7; Somme : 7469,07; écart-type (corrigé) : 468,6592

7 pts 1-a) Déterminer un intervalle de confiance pour le nombre N_c de factures impayées dans la population.

$\hat{p} = \frac{7}{20} = 0{,}35; \quad \hat{\sigma}_{\hat{p}} = \sqrt{1-\frac{20}{1000}}\sqrt{\frac{0{,}35(0{,}65)}{19}} = 0{,}1083$

Intervalle de confiance pour p :

$0{,}35 - 2(0{,}1083) \le p \le 0{,}35 + 2(0{,}1083)$
$0{,}13335 \le p \le 0{,}56665$

Intervalle de confiance pour N_c :

$133 \le N_c \le 567$

$\boxed{133} \le N_c \le \boxed{567}$

7 pts 1-b) Déterminer un intervalle de confiance pour la valeur moyenne des comptes de la population (payés ou pas).

$\hat{\sigma}_{\bar{y}} = \sqrt{1-\frac{20}{1000}}\frac{383{,}1939}{\sqrt{20}} = 84{,}8236$

$\boxed{748{,}75} \le \mu \le \boxed{1088{,}03}$

5 pts 1-c) Estimer le montant total impayé [estimation ponctuelle seulement—pas d'intervalle de confiance]

$$N\bar{y}' = 1000\frac{7469,07}{20} = 373\ 453$$

Montant impayé : 373453

4 pts 1-d) Déterminer un intervalle de confiance pour le montant total *impayé*.

$$\hat{\sigma}_{T'} = N\sqrt{1-\frac{n}{N}}\frac{s'}{\sqrt{n}} = 1000\sqrt{1-\frac{20}{1000}}\frac{584,8108}{\sqrt{20}} = 129\ 453$$

Intervalle de confiance :

$$373\ 453 - 2(129\ 453) \leq \tau_d \leq 373\ 453 + 2(129\ 453)$$

$$114\ 547 \leq \tau_d \leq 632\ 360$$

114 547 $\leq$ *Total impayé* $\leq$ 632 360

6 pts 1-e) On fait les recherches nécessaires pour déterminer précisément le nombre de factures impayées dans la population. On trouve qu'il y en a 360. Utiliser cette information pour obtenir une nouvelle estimation du montant total des factures impayées et déterminer un nouvel intervalle de confiance.

$$\hat{\tau}_d = N_d\bar{y}_d = 360\left(\frac{7469,07}{7}\right) = 384124$$

$$\hat{\sigma}_{\hat{\tau}_d} = N_d\sqrt{1-\frac{n_d}{N_d}}\frac{s_d}{\sqrt{n_d}} = 360\sqrt{1-\frac{7}{360}}\frac{468,6592}{\sqrt{7}} = 63146,13$$

Intervalle de confiance

$$384124-2(63146,13) \leq \tau_d \leq 384124+2(63146,13)$$

Estimation ponctuelle

Total impayé : 384124

Intervalle de confiance

257831 $\leq$ *Total impayé* $\leq$ 510416

$\boxed{\text{4 pts}}$ 1-f) Vous savez que le vérificateur n'a pas choisi les comptes de façon purement aléatoire et vous soupçonnez qu'il a favorisé plutôt les gros comptes. Vous faites les recherches nécessaires pour déterminer la valeur totale des comptes de la population (en souffrance ou pas) : elle est de 780 919 $. Pouvez-vous faire la preuve que l'échantillon tiré représente des montants anormalement grands?

Ce que vous savez, c'est que la vraie moyenne est $\mu = \dfrac{780919}{1000} = 780{,}919$.

Puisque l'intervalle de confiance déterminé en 1-b), $694{,}33 \le \mu \le 1\,042{,}44$, recouvre cette moyenne, notre échantillon n'a rien d'anormal. Donc non, la preuve que l'échantillon est biaisé en faveur des gros montants n'est pas faite.

$\boxed{\text{4 pts}}$ 1-g) Apprenant que la valeur totale des comptes de la population est connue (780 919 $), un analyste propose d'estimer le montant total des comptes impayés par 0,35 × 780 919 = 273 321,65 (0,35 étant la proportion de comptes impayés dans l'échantillon.) Déterminer un intervalle de confiance basé sur cet estimateur.

L'estimateur est $\hat{p}\times 780919$. On estime son écart-type par

$780919\hat{\sigma}_{\hat{p}} = 780\,919\sqrt{1-\dfrac{20}{1000}}\sqrt{\dfrac{0{,}35(0{,}65)}{19}} = 84593$. L'intervalle de confiance est donc

273 321,65-2(84593) ; 273 321,65+2(84593)
104 136 ; 442,508

Remarquez cependant que ceci est un intervalle de confiance pour le paramètre $p\tau$ et non pour τ_d. $p\tau$ et τ_d seraient égaux seulement si la moyenne des comptes impayés était égale à celle des comptes en général. C'est une supposition qu'on se voit parfois obligé de faire, faute d'information, mais qui au mieux ne donne qu'une approximation de la vraie moyenne du domaine.

$\boxed{104\ 136} \le \textit{Total impayé} \le \boxed{442\ 508}$

Problème 2

Un directeur d'école voudrait estimer la note moyenne en français des étudiants de son école. Il stratifie la population selon la filière: le général, le technique ou le professionnel. Il prélève un échantillon de chacune des strates, et fait passer à tous un test de français. Voici les données:

	Strate 1	*Strate* 2	*Strate* 3	*Total*
Taille de la strate	1300	1600	1100	4 000
Taille de l'échantillon tiré dans la strate	16	20	14	50
Note moyenne de l'échantillon	38	65	78	
Écart-type corrigé des notes de l'échantillon	14	13	11	

4 pts

2-a) Estimer la note moyenne des étudiants de la population [*Estimation ponctuelle seulement, pas d'écart-type*]

$$\bar{y}_{st} = W_1\bar{y}_1 + W_2\bar{y}_2 + W_3\bar{y}_3 = (0{,}325)(38) + (0{,}4)(65) + (0{,}275)(78) = 59{,}8$$

Estimation de la note moyenne des étudiants de la population : $\boxed{59{,}8}$

6 pts

2-b) Estimer l'écart-type de l'estimateur en 2-a)

$$\hat{\sigma}^2_{\bar{y}_1} = 12{,}099231; \hat{\sigma}^2_{\bar{y}_2} = 8{,}344375; \hat{\sigma}^2_{\bar{y}_3} = 8{,}532857$$

$$\hat{\sigma}_{\bar{y}_{st}} = \sqrt{(0{,}325)^2(12{,}099231) + (0{,}4)^2(8344375) + (0{,}275)^2(8{,}532857)}$$

$$\sqrt{3{,}258379} = 1{,}805098$$

Estimation de l'écart-type de l'estimation : $\boxed{1{,}805098}$

Problème 3

Un fournisseur s'engage à remplir et expédier une commande reçue par courrier le lendemain de la réception de la commande. Afin de planifier les activités du lendemain, la compagnie souhaite prédire le nombre de commandes à partir *du poids du courrier*. Afin d'analyser le lien entre ces deux variables, on prélève pendant 30 jours des données sur le poids (en kilos) du courrier et le nombre de commandes que contient le courrier. Les données brutes sont présentées en page 8. Mais voici les calculs essentiels:

Quelques calculs : *Poids du courrier* : Somme : 322 ; $s^2 = 12{,}8402$
Nombre de commandes : Somme : 152 700; $s^2 = 967\,138$
Covariance entre *poids du courrier* et *Nombre de commandes* : 3 318

5 pts

3-a) Déterminer la droite de régression qui permettrait de prédire le nombre de commandes à partir du poids du courrier.

$$b_1 = \frac{3318}{12{,}8402} = 258{,}4072$$

$$b_o = \frac{152700}{30} - 258{,}4072\left(\frac{322}{30}\right) = 2316{,}429$$

b_o = $\boxed{2316{,}429}$; b_1 = $\boxed{258{,}4072}$

4 pts

3-b) Estimer le nombre de commandes un jour où le courrier pèse 11,2 kilos.

$$b_o + b_1(11{,}2) = 5210{,}59$$

Nombre de commandes $\boxed{5211}$

4 pts

3-c) Le courrier aujourd'hui pèse 2 kilos de plus qu'hier. Combien de commandes aura-t-on aujourd'hui de plus qu'hier?

$$2b_1 = 516{,}8124$$

Nombre de commandes $\boxed{517}$

5 pts

3-d) Le coefficient de corrélation est-il significativement différent de 0? Énoncer clairement votre conclusion

$$r = \frac{s_{xy}}{s_x s_y} = \frac{3318}{\sqrt{12{,}8402}\sqrt{967138}} = 0{,}9415562$$

$$Z = \frac{\sqrt{30-2}\,(0{,}9415562)}{\sqrt{1-(0{,}9415562)^2}} = 14{,}79044 >> 2.$$

On rejette l'hypothèse que le coefficient de corrélation est nul pour conclure que *r* est significativement différent de 0. Il y a une relation réelle entre *X* et *Y*.

Conclusion *Le coefficient de corrélation est significativement différent de 0*

Problème 4

Dans une étude sur le goût des consommateurs pour les différentes catégories de bière, on a classifié un échantillon de 150 répondants selon leur préférence et selon le sexe.

	Préférence			
	Bière légère	*Bière blonde*	*Bière rousse*	*Total*
Femmes	20	40	20	80
Hommes	40	20	10	70
Total	60	60	30	150

5 pts 4-a) On veut savoir s'il y a une différence entre hommes et femmes quant au goût pour la bière. Énoncez l'hypothèse nulle H_o.

Il n'y a pas de différence entre hommes et femmes quant au goût pour la bière.

4 pts 4-b) Déterminer les effectifs théoriques

Calculs

Effectifs théoriques

	Préférence			
	Bière légère	*Bière blonde*	*Bière rousse*	*Total*
Femmes	32	32	16	80
Hommes	28	28	14	70
Total	60	60	30	150

4 pts 4-c) Calculer la valeur de khi-deux

$$\frac{(20-32)^2}{32}+\frac{(40-32)^2}{32}+\frac{(20-16)^2}{16}+\frac{(40-28)^2}{28}+\frac{(20-28)^2}{28}+\frac{(10-14)^2}{14}$$

$$= 4,5 + 2 + 1 + 5,14 + 2,28 + 1,14 = 16,07143$$

$\chi^2 = \boxed{16,07143}$

5 pts 4-d) Compléter le test et énoncer clairement votre conclusion [*Table en page* 13]

Le point critique à 2 degrés de liberté est 5,9915. Puisque 16,07 > 5,9915, on rejette l'hypothèse nulle. On peut conclure qu'il y a une différence réelle entre hommes et femmes dans les goûts pour la bière.

7 pts **Problème 5**

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.

	Réponse (*A, B, C, D ou E*)
1) Population: l'ensemble des aspirateurs issus d'une chaîne de montage. On s'installe à l'extrémité terminale de la chaîne à un moment quelconque de la matinée, puis on tire chaque 100^{e} aspirateur	**C**
2) Population: les parents d'élèves d'une certaine école. On tire au hasard et avec remise *n* enfants dans la liste des élèves; on inclut dans l'échantillon les parents de tous les enfants sélectionnés. [La réponse est **D** si on suppose que chaque famille n'a qu'un enfant à l'école]	**E**
3) Population: l'ensemble des étudiants de l'UQAM. On dresse une liste des étudiants de chaque faculté. Puis dans chaque faculté, on tire 100 étudiants au hasard.	**B**
4) Population: l'ensemble des étudiants de l'UQAM. On tire au hasard trois programmes, et on inclut dans l'échantillon tous les étudiants des trois programmes.	**D**
5) Population: l'ensemble des étudiants de l'UQAM. On dresse une liste de tous les étudiants, puis on tire successivement et sans remise 300 étudiants.	**A**
6) Population: L'ensemble de toutes les écoles publiques du Québec. On tire au hasard 12 commissions scolaire, et on inclut dans l'échantillon toutes les écoles des commissions scolaires sélectionnées.	**D**
7) Population: les 150 succursales d'une banque. On dresse une liste des 150 succursales; on tire 10 succursales au hasard, sans remise.	**A**

Données sur la question 3

Données brutes sur un échantillon de 30 jours.

Poids du courrier	*Nombre de commandes*	*Poids du courrier*	*Nombre de commandes*	*Poids du courrier*	*Nombre de commandes*	*Poids du courrier*	*Nombre de commandes*	*Poids du courrier*	*Nombre de commandes*	*Poids du courrier*	*Nombre de commandes*
5	4000	7	4000	9	4800	10,5	5000	12,5	5700	15	6200
5,5	3200	7,5	4200	9	5300	11	5500	13	5400	15	5900
6	3500	7,5	3600	10	5400	11,5	5800	13,5	5800	16,5	6600
6,5	3800	8	4300	10	5200	12	5400	14	6100	17	6700
6,5	3700	8,5	5000	10,5	5200	12	5000	14,5	6000	17,5	6400

Problème 6 page suivante →

10 pts

Problème 6

On tire un échantillon aléatoire simple de *ménages* dans la population de tous les ménages d'un quartier afin d'estimer les quantités ci-dessous. Dans chacun des cas, identifiez formellement le paramètre qu'il s'agit d'estimer. Faites votre choix dans la liste suivante:

A. μ_y : la moyenne d'une variable quantitative Y [Identifier la variable Y]

B. τ_y : le total d'une variable quantitative Y [Identifier la variable Y]

C. μ_d : la moyenne d'une variable Y dans un domaine $\mathfrak{D}$. [Identifiez le domaine $\mathfrak{D}$ et la variable Y]

D. τ_d : le total d'une variable quantitative Y dans un domaine $\mathfrak{D}$ (sous-ensemble de la population). [Identifiez le domaine $\mathfrak{D}$ et la variable Y]

E. p : la proportion d'unités appartenant à une classe $\mathfrak{C}$. [Identifiez la classe $\mathfrak{C}$]

F. N_c : le nombre d'unités appartenant à une classe $\mathfrak{C}$. [Identifiez la classe $\mathfrak{C}$]

G. $R = \mu_Y/\mu_X$: le quotient des moyennes de deux variables quantitatives Y et X. [Identifiez Y et X]

Modèle de réponse

0. Le revenu moyen des ménages francophones

$\mathfrak{D}$: l'ensemble des ménages francophones Y : revenu du ménage	**C**

1. Le nombre moyen de téléviseurs par ménage

Y : Nombre de téléviseurs dans le ménage	**A**

2. La proportion de femmes dans le quartier

Y : Nombre de femmes dans le ménage X : Nombre de personnes dans le ménage	**G**

3. Le nombre de femmes dans le quartier

Y : Nombre de femmes dans le ménage	**B**

4. Le nombre de ménages monoparentaux

$\mathfrak{C}$: La classe des ménages monoparentaux	**F**

5. Pourcentage de femmes parmi les francophones

Y : Nombre de femmes francophones dans le ménage X : Nombre de francophones dans le ménage	**G**

6. Le nombre moyen de téléviseurs par personne

Y : Nombre de téléviseurs dans le ménage X : Nombre de personnes dans le ménage	**G**

7. Le revenu moyen par personne

Y : Revenu total du ménage X :Nombre de personnes dans le ménage	**G**

8. Le revenu total des ménages monoparentaux

Y : Revenu total du ménage $\mathfrak{D}$: L'ensemble des ménages monoparentaux	**D**

9. Le revenu moyen des femmes dans le quartier

Y : Revenu total des femmes dans le ménage X : Nombre de femmes dans le ménage	**G**

10. Le revenu total des femmes du quartier

Y : Revenu total des femmes dans le ménage	**B**

Table de la loi normale

Surfaces à gauche du point z

z	0,09	0,08	0,07	0,06	0,05	0,04	0,03	0,02	0,01	0,00
-4,00	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
-3,90	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
-3,80	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,70	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,60	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002
-3,50	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3,40	0,0002	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003
-3,30	0,0003	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0005	0,0005	0,0005
-3,20	0,0005	0,0005	0,0005	0,0006	0,0006	0,0006	0,0006	0,0006	0,0007	0,0007
-3,10	0,0007	0,0007	0,0008	0,0008	0,0008	0,0008	0,0009	0,0009	0,0009	0,0010
-3,00	0,0010	0,0010	0,0011	0,0011	0,0011	0,0012	0,0012	0,0013	0,0013	0,0013
-2,90	0,0014	0,0014	0,0015	0,0015	0,0016	0,0016	0,0017	0,0018	0,0018	0,0019
-2,80	0,0019	0,0020	0,0021	0,0021	0,0022	0,0023	0,0023	0,0024	0,0025	0,0026
-2,70	0,0026	0,0027	0,0028	0,0029	0,0030	0,0031	0,0032	0,0033	0,0034	0,0035
-2,60	0,0036	0,0037	0,0038	0,0039	0,0040	0,0041	0,0043	0,0044	0,0045	0,0047
-2,50	0,0048	0,0049	0,0051	0,0052	0,0054	0,0055	0,0057	0,0059	0,0060	0,0062
-2,40	0,0064	0,0066	0,0068	0,0069	0,0071	0,0073	0,0075	0,0078	0,0080	0,0082
-2,30	0,0084	0,0087	0,0089	0,0091	0,0094	0,0096	0,0099	0,0102	0,0104	0,0107
-2,20	0,0110	0,0113	0,0116	0,0119	0,0122	0,0125	0,0129	0,0132	0,0136	0,0139
-2,10	0,0143	0,0146	0,0150	0,0154	0,0158	0,0162	0,0166	0,0170	0,0174	0,0179
-2,00	0,0183	0,0188	0,0192	0,0197	0,0202	0,0207	0,0212	0,0217	0,0222	0,0228
-1,90	0,0233	0,0239	0,0244	0,0250	0,0256	0,0262	0,0268	0,0274	0,0281	0,0287
-1,80	0,0294	0,0301	0,0307	0,0314	0,0322	0,0329	0,0336	0,0344	0,0351	0,0359
-1,70	0,0367	0,0375	0,0384	0,0392	0,0401	0,0409	0,0418	0,0427	0,0436	0,0446
-1,60	0,0455	0,0465	0,0475	0,0485	0,0495	0,0505	0,0516	0,0526	0,0537	0,0548
-1,50	0,0559	0,0571	0,0582	0,0594	0,0606	0,0618	0,0630	0,0643	0,0655	0,0668
-1,40	0,0681	0,0694	0,0708	0,0721	0,0735	0,0749	0,0764	0,0778	0,0793	0,0808
-1,30	0,0823	0,0838	0,0853	0,0869	0,0885	0,0901	0,0918	0,0934	0,0951	0,0968
-1,20	0,0985	0,1003	0,1020	0,1038	0,1056	0,1075	0,1093	0,1112	0,1131	0,1151
-1,10	0,1170	0,1190	0,1210	0,1230	0,1251	0,1271	0,1292	0,1314	0,1335	0,1357
-1,00	0,1379	0,1401	0,1423	0,1446	0,1469	0,1492	0,1515	0,1539	0,1562	0,1587
-0,90	0,1611	0,1635	0,1660	0,1685	0,1711	0,1736	0,1762	0,1788	0,1814	0,1841
-0,80	0,1867	0,1894	0,1922	0,1949	0,1977	0,2005	0,2033	0,2061	0,2090	0,2119
-0,70	0,2148	0,2177	0,2206	0,2236	0,2266	0,2296	0,2327	0,2358	0,2389	0,2420
-0,60	0,2451	0,2483	0,2514	0,2546	0,2578	0,2611	0,2643	0,2676	0,2709	0,2743
-0,50	0,2776	0,2810	0,2843	0,2877	0,2912	0,2946	0,2981	0,3015	0,3050	0,3085
-0,40	0,3121	0,3156	0,3192	0,3228	0,3264	0,3300	0,3336	0,3372	0,3409	0,3446
-0,30	0,3483	0,3520	0,3557	0,3594	0,3632	0,3669	0,3707	0,3745	0,3783	0,3821
-0,20	0,3859	0,3897	0,3936	0,3974	0,4013	0,4052	0,4090	0,4129	0,4168	0,4207
-0,10	0,4247	0,4286	0,4325	0,4364	0,4404	0,4443	0,4483	0,4522	0,4562	0,4602
0,00	0,4641	0,4681	0,4721	0,4761	0,4801	0,4840	0,4880	0,4920	0,4960	0,5000

Table de la loi normale

Surfaces à gauche du point z

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,10	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,20	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,30	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,40	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,50	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,60	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,70	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,80	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,90	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,00	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,10	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,20	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,30	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,40	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,50	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,60	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,70	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,80	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,90	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,00	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,10	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,20	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,30	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,40	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,50	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,60	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,70	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,80	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,90	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,00	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,10	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,20	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,30	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,40	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,50	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,60	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,70	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,80	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,90	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
4,00	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Formulaire MAT2080 Examen final

<u>Résumé des paramètres, leur estimateur, l'écart-type, et l'estimateur de l'écart-type.</u>

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 $= \frac{\mu_y}{\mu_x}$	$\hat{R} = \frac{\bar{y}}{\bar{x}}$	$\sigma_{\hat{R}} \approx \frac{\sqrt{1-f}}{\mu_x}\frac{\sqrt{S_y^2+R^2S_x^2-2RS_{xy}}}{\sqrt{n}}$	$\hat{\sigma}_{\hat{R}} = \frac{\sqrt{1-f}}{\bar{x}}\frac{\sqrt{s_y^2+\hat{R}^2s_x^2-2\hat{R}s_{xy}}}{\sqrt{n}}$
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	$\hat{\mu}_{yq} = \mu_x \hat{R}$	$\sqrt{1-f}\ \frac{\sqrt{S_y^2+R^2S_x^2-2RS_{xy}}}{\sqrt{n}}$	$\sqrt{1-f}\ \frac{\sqrt{s_y^2+\hat{R}^2s_x^2-2\hat{R}s_{xy}}}{\sqrt{n}}$
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*)	$T_d = N_d\,\bar{y}_d$		$Nd\sqrt{1-\frac{n_d}{N_d}}\ \frac{s_d}{\sqrt{n_d}}$
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* $\hat{N}_d = \frac{n_d}{n}N$		$N\sqrt{1-f}\ \frac{s'}{\sqrt{n}}$

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

<u>Taille d'échantillon</u>

Estimation de la moyenne La taille d'échantillon nécessaire pour que la marge d'erreur absolue soit égale à E est $n = \frac{n_o}{1+\frac{n_o}{N}}$ 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 $n = \frac{n_o}{1+\frac{n_o}{N}}$

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 $n = \dfrac{n_o}{1+\dfrac{n_o}{N}}$ où $n_o = \dfrac{4p(1-p)}{E^2}$.

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 $\sigma_{\bar{y}_h}^2 = (1\text{-}f_h)\dfrac{S_h^2}{n_h}$ et $f_h = n_h/N_h$.

L'allocation optimale pour l'estimation d'une moyenne dans u échantillon stratifié est donnée par n_h proportionnels aux $W_h S_h$.

Test du khi-deux

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

Régression simple

Les estimateurs de β_1 et β_0 sont $b_1 = \dfrac{s_{xy}}{s_x^2}$ et $b_o = \bar{y} - b_1 \bar{x}$, $r = \dfrac{s_{xy}}{s_x s_y}$; $Z = \dfrac{\sqrt{n-2}\ r}{\sqrt{1-r^2}}$

Points crtiques (α = 5 %) d'une loi khi-deux

ν	χ^2_ν	ν	χ^2_ν	ν	χ^2_ν	ν	χ^2_ν
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

Brouillon

Brouillon