

Problèmes supplémentaires : Examen final

STT1700

Automne 2009

1. Dans une conserverie, un protocole de contrôle de la qualité consiste à tirer régulièrement un échantillon de 20 boîtes de conserves et de tester l'hypothèse

$$H_0 : \mu = 200 \text{ g contre } H_1 : \mu < 200 \text{ g,}$$

où μ est le poids moyen de toutes les boîtes de conserves de la population. Vous supposerez que les poids sont distribués normalement, d'écart-type connu $\sigma = 2$. On considère deux tests :

Test 1 : On rejette H_0 si $Y \geq C$, où Y est le nombre de boîtes de l'échantillon qui pèsent moins de 200 g.

Test 2 : On rejette H_0 si $\bar{X} \leq D$, où \bar{X} est le poids moyen des boîtes de l'échantillon.

- Déterminez la valeur de C dans le test 1 pour un niveau $\alpha = 0,05$. Quelle est la taille de la région critique ? Justifiez le choix de la forme $Y \geq C$ de la région critique.
 - Déterminez les valeurs de la fonction de puissance du test 1 aux points $\mu = 196 ; 197 ; 198 ; 199 ; 199,25 ; 199,50 ; 199,75 ; 199,8$.
 - Déterminez la valeur de D qui donne au test 2 la même taille que le test 1.
 - Déterminez les valeurs de la fonction de puissance aux mêmes points qu'en b) et comparez.
2. La durée des ampoules d'un certain type est distribuée normalement d'écart-type 130. Si un échantillon de $n = 50$ ampoules donne une durée moyenne de 952,2 heures, testez à 5% l'hypothèse

$$H_0 : \mu = 1000 \text{ contre } H_1 : \mu < 1000 .$$

(Trouver d'abord la région critique i.e. trouver le critère avant de l'appliquer au résultat obtenu $\bar{X} = 952,2$).

3. Pour tester l'hypothèses $H_0 : \mu \leq 180$ contre $H_1 : \mu > 180$, où μ est la moyenne d'une population normale d'écart-type $\sigma = 24$, on prélève un échantillon de taille $n = 64$ et on décide de rejeter H_0 si > 186 .
- Quelle est la probabilité de rejeter H_0 si $\mu = 175$?
 - Vous ne pouvez pas calculer la probabilité exacte d'une erreur de première espèce. Mais quelle est la probabilité *maximale* d'une erreur de première espèce ?
 - Quelle est la probabilité maximale d'une erreur de seconde espèce ?

4. Les Grecs employaient souvent dans leur architectures des "rectangles d'or". Ce sont des rectangles de dimensions a et b ($a < b$) telles que $a/b = b/(a + b)$. Dans un rectangle d'or le rapport a/b vaut environ 0,618. Ce nombre a servi d'étalon esthétique non seulement dans l'art de la Grèce antique mais dans tout l'art occidental. Un anthropologue s'est demandé si l'art d'autres peuples non touchés par la civilisation grecque était aussi conforme à cet étalon. Il prélève un échantillon de 20 rectangles trouvés dans des oeuvres artisanales des indiens Shoshoni. Il calcule le rapport de la largeur sur la longueur pour chacun de ces rectangles. Voici les résultats, multipliés par 100 :

69,3 ; 66,2 ; 69,0 ; 57,0 ; 74,9 ; 67,2 ; 62,8 ; 60,9 ; 84,4 ; 60,6
65,4 ; 61,5 ; 66,8 ; 57,6 ; 67,0 ; 60,6 ; 61,1 ; 55,3 ; 93,3 ; 60,1

Tester l'hypothèse qu'en moyenne le rapport de la largeur sur la longueur des rectangles trouvés dans l'artisanat des indiens Shoshoni est conforme à l'idéal grec.

5. La teneur en nicotine de deux marques de cigarettes a un écart-type de $\sigma_1 = 1,2 \text{ mg}$ pour la première marque et $\sigma_2 = 1,4 \text{ mg}$ pour la deuxième. Pour tester l'hypothèse que $\mu_1 - \mu_2 = 2$, on a prélevé un échantillon de chacune des deux marques. Le premier échantillon, de taille 50, avait en moyenne $26,6 \text{ mg}$ (\bar{X}) ; le deuxième, de taille 40, avait en moyenne $23,8 \text{ mg}$ (\bar{Y}).
- Tester à 5% l'hypothèse $H_0 : \mu_1 - \mu_2 = 2$ contre $H_1 : \mu_1 - \mu_2 \neq 2$
 - Tester à 5% l'hypothèse $H_0 : \mu_1 - \mu_2 = 2$ contre $H_1 : \mu_1 - \mu_2 > 2$
 - Calculer la puissance au point $\mu_1 - \mu_2 = 2,5$ pour le test en a) et le test en b).
6. Une enquête menée en 1950 et à nouveau en 1975 a montré qu'en 1950 la taille moyenne de 400 garçons de 10 ans était $135,13 \text{ cm}$ avec un écart-type de $6,1 \text{ cm}$ (c.-à.-d. $\bar{X} = 135,13$ et $S_1 = 6,1$), alors qu'en 1975 la taille moyenne de 500 garçons de 10 ans était de $136,1 \text{ cm}$ avec un écart-type de $6,4 \text{ cm}$ (c.-à.-d. $\bar{Y} = 136,1$ et $S_2 = 6,4$). Tester à 5% l'hypothèse nulle que la population visée a connu un accroissement de taille de 1 cm contre l'alternative que l'accroissement a été inférieur à 1 cm .
7. Pour comparer deux marques de peinture, on emploie 4 bidons de chaque marque et on note la superficie qu'ils peuvent couvrir. Les 4 bidons de la première marque ont couvert $47,6 \text{ m}^2$ en moyenne avec un écart-type de $2,9 \text{ m}^2$; les 4 bidons de la deuxième marque ont couvert en moyenne $45,7 \text{ m}^2$ avec un écart-type de $2,41 \text{ m}^2$. Tester à 5% l'hypothèse nulle $\mu_1 = \mu_2$ contre l'alternative $\mu_1 - \mu_2 > 0$.
8. Pour déterminer si les habitants de deux îles du Pacifique sont de même race, un anthropologue détermine les indices céphaliques de 6 adultes de sexe masculin de chacune des îles. Il obtient les résultats suivants :

$$\bar{X}_1 = 77,4, \bar{X}_2 = 72,2$$

$$S_1 = 3,3, S_2 = 2,1$$

Tester à 1% l'hypothèse qu'il n'y a pas de différence réelle entre les deux séries. (Supposez que les populations sont normales et qu'elles ont la même variance).

9. Les planchers des hôpitaux devraient-ils être tapissés ? Un tapis propre est-il aussi propre qu'un plancher lavé ? Si l'on pouvait être sûr que non, l'administrateur de l'hôpital disposerait d'un élément d'information important. Dans un hôpital du Montana, une expérience a été faite pour répondre à cette question. Seize chambres de malades ont été testées, dont huit avaient des tapis et huit autres n'en avaient pas. Les chambres étaient comparables autrement. En forçant de l'air à un taux constant à travers un vase de pétri, et en comptant ensuite le nombre de colonies de bactéries qui s'y forment, on a pu mesurer la quantité de bactéries en suspension dans les chambres. Voici les résultats :

<i>No de la chambre avec tapis</i>	#212	#216	#220	#223	#225	#226	#227	#228
<i>Nombre de colonies par m³</i>	1,10	0,76	0,66	1,21	1,00	0,94	1,36	1,30
<i>No de la chambre sans tapis</i>	#210	#214	#215	#217	#221	#222	#224	#229
<i>Nombre de colonies par m³</i>	1,12	0,77	0,35	0,67	1,11	1,03	0,94	1,27

Quelle est votre conclusion ? Faites un test unilatéral à 5%.

10. Le rôle qu'à joué le célèbre humoriste Mark Twain dans la guerre civile a longtemps été un sujet de controverse. Certains savants ont présenté comme évidence des exploits militaires de Mark Twain une série de 10 lettres publiées dans le New Orleans Daily Crescent durant l'année 1861. Les lettres constituaient une chronique des aventures de l'auteur, un dénommé "Quintus Curtis Snodgrass". Les historiens admettent que les événements décrits dans ces lettres sont authentiques, mais on n'a trouvé aucune trace du dénommé Quintus Curtis Snodgrass. Selon certains critiques, cependant, le style et

l'esprit de ces lettres étaient typiques de ceux de Mark Twain. La question est donc : Mark Twain était-il l'auteur des lettres de Quintus Curtis Snodgrass ? Pour répondre à cette question, on a choisi 8 lettres écrites par Mark Twain et on se proposait de les comparer de façon quantitative aux lettres de Quintus Curtis Snodgrass. Un style est caractérisé par plusieurs variables. Pour les fins de cette étude, on en a choisi une seule : la fréquence des mots de 3 lettres. Pour les 8 lettres de Mark Twain, les fréquences sont les suivantes :

<i>No de la lettre</i>	1	2	3	4	5	6	7	8
<i>Fréquence</i>	0,225	0,262	0,217	0,240	0,230	0,229	0,235	0,217

Pour les 10 lettres de Quintus Curtis Snodgrass, les fréquences sont :

<i>No de la pièce</i>	1	2	3	4	5	6	7	8	9	10
<i>Fréquence</i>	0,229	0,205	0,196	0,210	0,202	0,207	0,224	0,223	0,220	0,201

Testez l'hypothèse que les lettres signées « Quintus Curtis Snodgrass » ont été écrites par Mark Twain.

- Un étudiant de Pennsylvania State University s'est donné pour projet de maîtrise l'étude de l'effet des boissons au Cola sur ceux qui en consomment en quantités abusives. Il s'est trouvé 57 étudiants qui consommaient quotidiennement de 48 à 111 onces de Cola. Comme groupe témoin, il a pris 112 étudiants qui ne consommaient pas plus de 8 onces par jour. Dans le premier groupe, 58% ont déclaré qu'ils se sentent tendus et nerveux quelques fois. Dans le groupe témoin, le pourcentage n'était que de 40%. Cette différence est-elle significative?
- On estime à environ 46 % le pourcentage des amputés qui succombaient aux infections post chirurgicales lorsqu'on n'employait pas de désinfectant dans les salles d'opérations. Joseph Lister, un chirurgien anglais, s'inspirant de l'œuvre de Louis Pasteur, a effectué 40 amputations en se servant du phénol comme désinfectant. De ces 40, seulement 6 sont morts d'une infection postopératoire. Peut-on conclure que le phénol est bénéfique ?
- Une étude qui portait sur le taux de mortalité de 92 patients atteints de troubles cardiaques sérieux a rapporté les observations suivantes : 53 des patients possédaient un animal familier ; 3 de ces patients moururent au courant de la première année d'observation. Parmi les 39 patients qui ne possédaient pas d'animal familier, 11 sont décédés durant la même période de temps. L'écart entre les deux taux de décès est-il significatif ?
- Une étude des décisions rendues par des jurys dans des cas de vols par effraction où l'accusé était de race noire a révélé les faits suivants : parmi les 28 cas où les victimes étaient de race noire, l'accusé a été trouvé coupable dans 12 cas ; parmi les 36 cas où la victime était de race blanche, l'accusé a été trouvé coupable dans 23 cas. Peut-on conclure que les jurys ont une plus forte tendance à déclarer coupables ceux qui sont accusés d'avoir commis des vols contre des Blancs ?
- Les enfants qui ont de graves problèmes d'apprentissage et de comportement ont souvent le même type d'électroencéphalogramme que celui qu'on retrouve chez les enfants souffrant du petit mal. Ceci a mené à la conjecture que l'éthosuximide, un anticonvulsif qui sert à traiter le petit mal épileptique, pourrait faciliter l'apprentissage. Une expérience a été faite récemment pour tester cette théorie. Dix enfants, âgés de 8 à 14 ans ont servi de sujets. L'expérience a duré 6 semaines. Pendant 3 semaines, l'enfant recevait un placebo, et pendant 3 autres semaines il était à l'éthosuximide. A la fin de chacune des deux périodes de 3 semaines, l'enfant composait plusieurs parties du test d'intelligence de Wechsler. De façon purement aléatoire on choisissait l'ordre dans lequel les traitements était donnés : certains enfants recevaient l'éthosuximide durant les 3 premières semaines et le placebo ensuite ; et d'autres les recevaient dans l'ordre inverse. Voici les scores pour chaque sujet. L'ordre dans lequel chaque traitement a été donné n'est pas indiqué.

<i>Patient</i>	1	2	3	4	5	6	7	8	9	10
<i>Sous placebo</i>	97	106	106	95	102	111	115	104	90	96
<i>Sous médication</i>	113	113	101	119	111	122	121	106	110	126

Que concluez-vous ?

16. Soit S_1^2 et S_2^2 les variances de deux échantillons indépendants issus de populations normales de même variance σ^2 .

- a) Montrez que $a S_1^2 + b S_2^2$ est un estimateur sans biais de σ^2 si et seulement si $a + b = 1$.
- b) Montrez que parmi tous les estimateurs sans biais de la forme $a S_1^2 + b S_2^2$, celui qui minimise la variance est $a S_1^2 + (1 - a) S_2^2$ où $a = (n_1 - 1) / (n_1 + n_2 - 2)$.

17. On a souvent affirmé que la lune a un effet sur la santé mentale de certaines personnes (ce qui explique le terme "lunatique"). Des études récentes ont tenté d'établir un lien entre les phases de la lune et le taux de suicides, la pyromanie, et même l'épilepsie. Une étude plus modeste a été faite dans une clinique psychiatrique en Virginie. On a noté le nombre d'admissions par jour durant la pleine lune pendant 12 mois consécutifs — du mois d'août 1971 au mois de juillet 1972. Voici les nombres de patients admis :

Mois	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'admissions	5	13	14	12	6	9	13	16	25	13	14	20

On sait que durant le reste de l'année, le nombre d'admissions moyen est de 11,3 par jour. Testez l'hypothèse que la lune n'a pas d'effet sur la santé mentale contre l'alternative selon laquelle la pleine lune provoque la maladie mentale.

18. Voici la répartition, selon le mois, de 1 855 décès choisis au hasard parmi tous les décès qui ont eu lieu en Amérique du Nord en 1976.

Mois	Jan	Fév.	Mar s	Avril	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.	Tota l
Effectif	157	142	158	152	156	153	158	159	153	157	152	158	1855

- a) Peut-on attribuer les différences uniquement au hasard?
- b) Supposons que ces données vous aient été fournies par un assistant en qui vous n'avez pas trop confiance. Les données ont-elles l'air d'avoir été manipulées?

19. Le tableau ci-dessous donne, pour une population de 87 627 mariages qui ont eu lieu au Canada en 1974, la distribution de la religion de l'époux et de la religion de l'épouse.

		Religion de l'épouse				
		Anglicane	Baptiste	Catholique romaine	Église Unie	
Religion de L'époux	Anglicane	7504	701	4731	4 680	17 616
	Baptiste	710	2222	992	1 050	4 974
	Catholique romaine	4589	914	26919	6 044	38 466
	Église Unie	4247	932	6143	15 249	26 571
	Total	17 050	4 769	38 785	27 023	87 627

- a) Dressez un tableau qui expose clairement la forte dépendance entre ces deux variables.
- b) Testez l'hypothèse que 50% des femmes Baptistes épousent des coreligionnaires.
- c) Testez l'hypothèse que 50% des mariages entre catholiques et baptistes sont des mariages entre un homme catholique et une femme baptiste.
- d) Testez l'hypothèse que la probabilité qu'un homme épouse une coreligionnaire est la même pour les 4 religions représentées.

20. Une usine fabrique des toiles métalliques pour des usines de pâtes et papier. Afin de mieux répartir son personnel, le gérant aimerait prévoir le temps, T , requis pour la finition des toiles. Ce temps pourrait être lié, entre autres variables, à la surface de la toile, S . On a obtenu les données du tableau 11.3:
- Quelle variable doit-on utiliser comme variable dépendante? (Justifier ce choix).
 - Déterminer l'équation de régression correspondante.
 - Quel est le temps moyen de finition pour une toile de 20 m^2 ?
 - La régression est-elle utilisable pour des fins de prévision? (Faire un test à 5%).
 - Quel est le pourcentage de variation expliquée par la droite de régression?
 - Faire un graphique des données. Tracer la droite de régression. Le modèle est-il raisonnable?

Temps de finition d'une toile (T) et surface de la toile (S)

<i>i</i>	<i>T</i>	<i>S</i>
1	5,50	9,30
2	5,90	13,50
3	5,80	11,10
4	6,30	14,90
5	7,00	16,70
6	7,50	23,20
7	5,50	11,10
8	7,20	20,40
9	6,50	15,80
10	6,50	14,90
11	7,10	18,60
12	7,00	15,80
13	6,90	16,70
14	6,80	15,80
15	6,60	16,70
Totaux	98,10	234,50