

# MAT4681 Statistique, Été 2022

## Examen Final A

L'examen dure 3 heures, toute sortie avant le fin est autorisée, et sera définitive.

Les calculatrices sont autorisées. Les documents sont en revanche interdits, sauf une copie des 14 pages d'aide mémoire envoyées avant l'examen, possiblement annotée.

Dans les feuilles qui suivent, il y a 25 questions relatives au cours de statistique. Pour chaque question, cinq réponses sont proposées. Une seule est valide, et vous ne devez en retenir qu'une,

- vous gagnez 1 point par bonne réponse
- vous ne perdez pas de points pour une mauvaise réponse
- vous ne gagnez pas de point pour plusieurs réponses

Aucune justification n'est demandée.

Il est fortement recommandé de lire attentivement la question avant de chercher la réponse.

Pour toute question donnant un ordre de grandeur ("environ...") la bonne réponse sera la plus proche.

**La page de réponses est la dernière page du lot que vous lisez présentement** : merci de décrocher ladite feuille et de ne rendre que cette dernière, après avoir inscrit votre code permanent en haut à gauche.

Merci de cocher le carré en bleu ou en noir. En cas d'erreur, vous pouvez cocher une autre case en rouge. Seule cette dernière sera alors retenue.

Le surveillant ne répondra à aucune question durant l'épreuve : en cas de soucis sur une question (interprétation possiblement fausse, typo, etc), vous pouvez mettre un court commentaire sur la feuille de réponses.

**Règlement no 18** *Tout acte de plagiat, fraude, copiage, tricherie, falsification de document ou création d'un faux document commis par une candidate, un candidat, une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen, d'un travail ou d'un stage faisant l'objet d'une évaluation ou dans toute autre circonstance, constitue une infraction au sens de ce règlement. (a. 2.1)*  
<https://r18.uqam.ca/>

- 1 Soit  $Z$  une variable aléatoire suivant une loi normale centrée réduite. Que vaut

$$\mathbb{P}[Z \in (-2.30 ; 0.14)]$$

- A) environ 0.445
- B) environ 0.545
- C) environ 0.645
- D) environ 0.745
- E) environ 0.845

- 2 Une étude est menée auprès d'un échantillon aléatoire de 2000 conducteurs québécois, pour savoir s'il y a un lien entre l'âge du conducteur et le risque d'accident. On a obtenu les statistiques suivantes

**Répartition des conducteurs de l'échantillon selon l'âge et l'implication dans un accident au cours des 12 derniers mois**

Âge du conducteur	Implication dans un accident		Total
	Oui	Non	
De 16 ans à 24 ans	14	194	208
De 25 ans à 44 ans	22	652	674
De 45 ans à 64 ans	18	782	800
65 ans et plus	6	312	318
Total	60	1 940	2 000

Source : Société de l'assurance automobile du Québec. *Bilan routier 2012*, juillet 2013.

Si les deux variables (âge du conducteur et implication dans un accident) étaient indépendantes, combien de personnes de 45 ans à 64 ans devraient être impliquées dans un accident ?

- A) environ 6
  - B) environ 12
  - C) environ 18
  - D) environ 24
  - E) environ 30
- 3 Sur 1000 billes dans un roulement à billes, on observe que le poids d'une bille est de 5.02 et l'écart-type 0.30. On prend 100 billes au hasard, quelle est la probabilité que le poids total (des 100 billes) soit plus grand que 510 ?
- A) moins de 0.5%
  - B) entre 0.5% et 1%
  - C) entre 1% et 5%
  - D) entre 5% et 20%
  - E) plus de 20%

- 4 On lance une pièce bien équilibrée 120 fois. Quelle est la probabilité d'avoir une proportion de 'face' qui excède  $5/8$ .
- A) moins de 0.5%
  - B) entre 0.5% et 1%
  - C) entre 1% et 5%
  - D) entre 5% et 20%
  - E) plus de 20%
- 5 On interroge 100 personnes au hasard pour savoir s'ils ont voyagé hors du Québec au cours des 5 dernières années, et 55 ont affirmé que oui. Donner un intervalle de confiance à 99% pour la probabilité qu'un Québécois ait voyagé hors du Québec au cours des 5 dernières années.
- A)  $[0.55 \pm 0.082]$
  - B)  $[0.55 \pm 0.098]$
  - C)  $[0.55 \pm 0.116]$
  - D)  $[0.55 \pm 0.128]$
  - E)  $[0.55 \pm 0.154]$
- 6 On dispose de deux échantillons d'ampoules électriques. Selon les fabricants, les ampoules de type  $A$  ont la durée de vie moyenne est 1400 heures, avec un écart-type de 120 heures; et les ampoules de type  $B$ , ont une durée de vie moyenne est de 1200 heures avec un écart-type de 80 heures. On considère deux échantillons de 125 ampoules de type  $A$ , et 125 de type  $B$ . Quelle est la probabilité que la différence entre les deux durées moyennes dépasse 240 heures.
- A) moins de 0.5%
  - B) entre 0.5% et 1%
  - C) entre 1% et 2%
  - D) entre 2% et 5%
  - E) plus de 5%
- 7 On a demandé à 400 adultes et 600 adolescents s'ils aimaient une émission de télévision diffusée hier soir : 100 adultes et 300 adolescents ont affirmé l'avoir aimé. Donner un intervalle de confiance à 95% pour la différence de pourcentage entre les deux
- A)  $25\% \pm 4\% = (21\% ; 29\%)$
  - B)  $25\% \pm 6\% = (19\% ; 31\%)$
  - C)  $25\% \pm 8\% = (17\% ; 33\%)$
  - D)  $25\% \pm 10\% = (15\% ; 35\%)$
  - E)  $25\% \pm 12\% = (13\% ; 37\%)$

- 8 On veut estimer la proportion de personnes qui utilisent internet au quotidien sous la forme d'un intervalle de confiance à 95%. Pour cela, on a prélevé un échantillon aléatoire de taille  $n = 226$  personnes, et on a observé que 196 utilisaient internet au quotidien. Déterminez la borne supérieure de l'intervalle de confiance (symétrique) à 95% de la proportion de personnes qui utilisent internet au quotidien.
- A) moins de 90%
  - B) entre 90% et 91
  - C) entre 91% et 92%
  - D) entre 92% et 93%
  - E) plus de 93%
- 9 Un pharmacien prétend qu'une molécule est efficace dans 90% des cas pour guérir une maladie dans un délai de 8 heures. Dans un échantillon de 100 personnes qui ont testé la molécule, on note  $\hat{p}$  la proportion de personnes guéries. Si on veut tester  $H_0 : p = 90\%$  quelle serait la forme de la région de rejet du test, si  $H_1 : p < 90\%$ , pour un seuil  $\alpha = 5\%$  ?
- A)  $\hat{p} < 80\%$
  - B)  $\hat{p} < 85\%$
  - C)  $\hat{p} < 90\%$
  - D)  $\hat{p} > 85\%$
  - E)  $\hat{p} > 80\%$

**Les deux questions suivantes reposent sur le même énoncé :** On cherche à estimer le moyenne du nombre d'heures de sommeil par patient dans une population de 500 patients traités par un nouveau type de somnifère. Pour cela, on a observé pour 50 patients, choisis aléatoirement dans cette population, le nombre d'heures de sommeil lors d'un enregistrement nocturne. On a obtenu, sur ce échantillon, une moyenne de 8 heures de sommeil, par patient, avec un écart-type (basé sur la variance corrigée) de 1 heure.

- 10 Quel est l'estimateur de l'écart-type de la moyenne du nombre d'heures de sommeil, par patient ?
- A) moins de 0.130 heure
  - B) entre 0.130 et 0.145 heure
  - C) entre 0.145 et 0.160 heure
  - D) entre 0.160 et 0.175 heure
  - E) plus de 0.175 heure
- 11 Déterminer la borne supérieure de l'intervalle de confiance (symétrique) à 95% pour la moyenne du nombre d'heures de sommeil, par patient, dans toute la population.
- A) moins de 8.15 heure
  - B) entre 8.15 et 8.20 heure
  - C) entre 8.20 et 8.25 heure
  - D) entre 8.25 et 8.30 heure
  - E) plus de 8.30 heure

- 12 Une étude affirme que la probabilité d'avoir un "accident majeur", pour une année donnée, sur un réacteur nucléaire est de l'ordre de 0.03%. En Europe, il y a 143 réacteurs. Quel serait un ordre de grandeur raisonnable pour qu'il ait au moins un accident majeur sur un réacteur, sur une période de 30 ans (en supposant l'indépendance entre les réacteurs, et les années).

A) environ 65%  
 B) environ 72%  
 C) environ 85%  
 D) environ 95%  
 E) plus de 99%

- 13 Un restaurateur se demande s'il existe une relation linéaire entre le temps passé dans l'établissement ( $x$ ), et le montant de la facture ( $y$ ). Sur un échantillon de  $n = 35$  clients, il a observé

$$\bar{x} = 50 \text{ minutes}, s_x = 10 \text{ minutes}, \bar{y} = 35\$, s_y = 10\$,$$

et comme les données sont appariées,  $r_{xy} = \text{cov}(x, y) = 55$ . Il se demande si la relation linéaire observée entre le temps passé dans l'établissement ( $x$ ), et le montant de la facture ( $y$ ) est significative (ou pas). On note  $t$  la statistique de test. Il fait un test de niveau de confiance  $\alpha = 5\%$ .

A)  $t \geq 1.5$  et on ne peut pas conclure qu'il y a une relation linéaire entre le temps passé et le montant dépensé  
 B)  $t \leq 1.5$  et on ne peut pas conclure qu'il y a une relation linéaire entre le temps passé et le montant dépensé  
 C)  $t \leq 2.5$  et on peut conclure qu'il y a une relation linéaire entre le temps passé et le montant dépensé  
 D)  $t \geq 2.5$  et on ne peut pas conclure qu'il y a une relation linéaire entre le temps passé et le montant dépensé  
 E)  $t \geq 3.5$  et on peut conclure qu'il y a une relation linéaire entre le temps passé et le montant dépensé

- 14 Afin de tester l'efficacité d'un sérum, on constitue un échantillon de 200 personnes : 100 se voient prescrire le sérum (groupe A) et 100 se voient prescrire un placebo (groupe B)

Frequencies Observed			
	Recover	Do Not Recover	TOTAL
Group A (using serum)	75	25	100
Group B (not using serum)	65	35	100
TOTAL	140	60	200

On cherche à tester  $H_0$ , le sérum n'a pas d'effet (pas davantage qu'un placebo) par un test du chi-deux. Que vaut la statistique de test ?

A) moins de 2.5  
 B) entre 2.5 et 3.5  
 C) entre 3.5 et 4.5  
 D) entre 4.5 et 5.5  
 E) plus de 5.5

- 15 Afin de tester l'efficacité d'un sérum, on constitue un échantillon de 200 personnes : 100 se voient prescrire le sérum (groupe A) et 100 se voient prescrire un placebo (groupe B). On obtient le tableau de contingence suivant, avec  $x \in \{0, 1, \dots, 30\}$

Frequencies Observed			
	Recover	Do Not Recover	TOTAL
Group A (using serum)	$70 + x$		100
Group B (not using serum)			100
TOTAL	140	60	200

On cherchait à tester  $H_0$  correspondant à l'hypothèse que le sérum n'a pas d'effet (pas davantage qu'un placebo) par un test du chi-deux. On sait que  $H_0$  est rejeté au seuil  $\alpha = 5\%$ . Que vaut  $x$  ?

- A) 3 ou moins
- B) 7 ou moins
- C) 10 ou moins
- D) 3 ou plus
- E) 7 ou plus

- 16 On dispose des observations suivantes, supposées suivre une distribution  $\mathcal{U}([-\theta, \theta])$ , avec  $\theta > 0$ ,

$$\mathbf{x} = \{-0.94, -0.51, 0.29, 1.63, -1.19, 1.59\}$$

Que vaut l'estimateur  $\hat{\theta}$  obtenu par la méthode des moments ?

- A) moins de 1.7
- B) entre 1.7 et 1.95
- C) entre 1.95 et 2.25
- D) entre 2.25 et 2.4
- E) plus de 2.4

- 17 En faisant un calcul rapide, un étudiant propose un intervalle de confiance pour la moyenne de la forme  $[-2.053749; 2.053749]$ , en supposant la variance connue. En se souvenant qu'il n'avait que 30 observations, il envisage une correction, pour tenir compte du fait que la variance est inconnue. Que deviendra alors l'intervalle de confiance (pour le même niveau de confiance) ?

- A)  $[-2.04 ; 2.04]$
- B)  $[-2.15 ; 2.15]$
- C)  $[-2.31 ; 2.31]$
- D)  $[-2.46 ; 2.46]$
- E)  $[-2.62 ; 2.62]$

- 18 Une étude menée auprès d'un échantillon de 450 hommes et 500 femmes indique que  $\hat{p}_H = 17\%$  des hommes et  $\hat{p}_F = 13\%$  des femmes dorment moins de 6.5 heures par nuit. Au seuil de signification de 5%, quelle différence minimale entre les deux proportions doit-on avoir pour conclure que le pourcentage de personnes qui dorment moins de 6.5 heures par nuit est plus élevé chez les hommes que chez les femmes ?

- A)  $\hat{p}_H - \hat{p}_F > 1.2\%$
- B)  $\hat{p}_H - \hat{p}_F > 2.6\%$
- C)  $\hat{p}_H - \hat{p}_F > 3.2\%$
- D)  $\hat{p}_H - \hat{p}_F > 3.8\%$
- E)  $\hat{p}_H - \hat{p}_F > 4.6\%$

- 19 Lors d'un test qualité, on a interrogé 200 personnes pour savoir s'ils préféraient les croustilles A ou les croustilles B. On a eu un peu plus de personne qui ont préféré B, de telle sorte qu'on a le comptage suivant, avec  $x \in \{0, 1, 2, \dots, 100\}$ ,

A	B
100-x	100+x

Un test statistique (avec un seuil de 5%) n'a pas permis de les départager. Quelle est la valeur maximale de  $x$  ?

- A)  $x \leq 8$
- B)  $x \leq 14$
- C)  $x \leq 18$
- D)  $x \leq 22$
- E)  $x \leq 26$

- 20 Après avoir lancé 600 fois un dé, on a obtenu le comptage suivant. Les deux premiers comptages manquent (pour 1 et 2), mais on sait qu'il y a eu plus de 2 que de 1. Et comme la somme vaut 200, on sait qu'on peut écrire  $100 - x$  et  $100 + x$  respectivement avec  $x \in \{0, 1, 2, \dots, 100\}$ .

1	2	3	4	5	6
100-x	100+x	95	101	98	106

Un test du chi-deux, avec un seuil de 5% conclut que le dé est équilibré. Quelle est la valeur maximale de  $x$  ?

- A)  $x \leq 8$
- B)  $x \leq 14$
- C)  $x \leq 18$
- D)  $x \leq 22$
- E)  $x \leq 26$

- 21 On prélève un échantillon de 40 enfants afin d'estimer le nombre de moyen  $\mu$  de caries dentaires parmi les 1300 élèves d'une école. Les résultats de l'examen dentaire sont présentés dans le tableau suivant

Nombre de caries	0	1	2	3	4	Total
Effectifs (nombre d'enfants)	28	7	2	2	1	40

Déterminez un intervalle de confiance à 95% pour  $\mu$ .

- A) [0.17 ; 0.87]
- B) [0.21 ; 0.83]
- C) [0.33 ; 0.72]
- D) [0.47 ; 0.57]
- E) [0.51 ; 0.54]

**Les deux questions suivantes portent sur le même énoncé:** On cherche à déterminer le poids d'athlètes à partir de leur taille. Sur un échantillon de  $n = 38$  sportifs, on obtient les statistiques suivantes, pour la taille ( $x$ , en centimètres) et pour le poids ( $y$ , en kilogrammes),

$$\bar{x} = 183 \text{ cm}, s_x = 8 \text{ cm}, \bar{y} = 70 \text{ kg}, s_y = 3 \text{ kg},$$

et comme les données sont appariées,  $r_{xy} = \text{cov}(x, y) = 22.8$ . On note  $y = a + bx$  l'équation de la droite de régression obtenue par moindres carrés.

- 22 Que vaut  $b$  ?

- A) moins de 0.35
- B) entre 0.35 et 0.40
- C) entre 0.40 et 0.45
- D) entre 0.45 et 0.50
- E) plus de 0.50

- 23 Que vaut l'écart-type de  $b$  ?

- A) moins de 0.018
- B) entre 0.018 et 0.020
- C) entre 0.020 et 0.022
- D) entre 0.022 et 0.024
- E) plus de 0.024



**Les deux questions suivantes portent sur le même énoncé:** Un sondage a été mené auprès de 1000 travailleurs québécois. On leur a demandé leur opinion quant à une diminution de dépenses publiques dans les programmes d'aide sociale. On leur a aussi demandé s'ils étaient syndiqués, ou pas, et les données ont été résumées dans le tableau de comptage suivant

	<i>Pour une diminution des dépenses</i>			Total
	En accord	Indifférent	En désaccord	
Syndiqué	200	160	150	510
Non syndiqué	170	140	180	490
Total	370	300	330	1000

- 24 Que vaut la statistique du chi-deux du test d'indépendance entre l'opinion quant à une baisse des dépenses publiques et le fait d'être syndiqué
- A) moins de 3.5
  - B) entre 3.5 et 4.5
  - C) entre 4.5 et 5.5
  - D) entre 5.5 et 6.5
  - E) plus de 6.5
- 25 Que vaut la  $p$ -value du test
- A) environ 0.5%
  - B) environ 1%
  - C) environ 2%
  - D) environ 5%
  - E) environ 8%



Code permanent :

Sujet : A

question 1	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 2	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 3	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 4	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 5	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 6	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 7	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 8	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 9	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 10	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 11	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 12	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 13	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 14	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 15	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 16	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 17	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 18	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 19	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 20	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 21	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 22	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 23	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 24	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E
question 25	<input type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D	<input type="checkbox"/> E