MAT4681 Statistique, Été 2022

Examen Intra **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 4 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 (sauf deux), 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.

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.

1 Un vérificateur prélève un échantillon de 30 transactions dans une très grande population de transactions. Dans la population, le taux de transactions erronées est de 1 sur 200. Quelle est la probabilité d'avoir 2 transactions erronées sur les 30 tirés au hasard ?

A) environ 1 chance sur 4000

B) environ 1 chance sur 1000

C) environ 1 chance sur 200

D) environ 1 chance sur 100

E) environ 1 chance sur 20

```
> dbinom(2, 30, 1/200)
[1] 0.009450951
> dpois(2,30/200)
[1] 0.009682965
> x = rbinom(1e6, size=30, prob = 1/200)
> mean(x==2)
[1] 0.009498
```

2 X est une variable aléatoire positive, qui suit une loi exponentielle de moyenne μ.

Que vaut μ si $\mathbb{P}[X > 80] = 95\%$?

A) Moins de 1500

B) Entre 1500 et 1750, car $-\dfrac{80}{\log(.95)} \approx 1560$

C) Entre 1750 et 2000

D) Entre 2000 et 2500

E) Plus de 2500

```
> 1-pexp(80, 1/1560)
[1] 0.9500107
> x = rexp(1e6, 1/1560)
> mean(x)
[1] 1561.729
> mean(x > 80)
[1] 0.950008
```

3 Le nombre de décès par 24 heures dans la région de Saguaney suit une loi de Poisson de moyenne 4.28.

Quelle est la probabilité de n'observer aucun décès dans une journée ?

A) Moins de 0.1%

B) Entre 0.1% et 1%

C) Entre 1% et 2%, $e^{-4.28} \approx 0.0138$

D) Entre 2% et 5%

E) Plus de 5%

```
> dpois(0, 4.28)
[1] 0.01384266
> x = rpois(1e6, 4.28)
> mean(x)
[1] 4.279819
> mean(x==0)
[1] 0.013949
```

4 On trois fois une paire de dés (bien équilibrés) à six faces, chacun, et on regarde la somme des valeurs des faces apparentes.

Quelle est la probabilité d'obtenir au moins une fois 7 pour la somme des faces des deux dés, sur 3 lancers ?

A) Moins de 40%

B) Entre 40% et 45%, $1 - 125/216 = 91/216 \approx 0.4213$

C) Entre 45% et 50%

D) Entre 50% et 55%

E) Plus de 55%

```
> rdeoublede = function(n) apply(matrix(sample(1:6, 2*n, replace=TRUE),n,2),1,sum)
> rdeoublede(12)
 [1]  9 10  6  9  7  9  9  9  6 11  5  6
> x =apply(matrix(rdeoublede(3*1e6),1e6,3),1,function(x) sum(x==7))
> mean(x>0)
[1] 0.421943
```

5 On 500 fois une pièce bien équilibrée (en moyenne, on espère obtenir 250 fois "pile").

Quelle est la probabilité d'avoir entre 240 et 260 fois "pile" (sur 500 lancers) ?

A) Moins de 40%

B) Entre 40% et 50%

C) Entre 50% et 60%

D) Entre 60% et 70%, suivant que l'on inclue, ou pas, les bornes, la (vraie) probabilité est entre 60.4% et 65.2%.

Avec une approximation Gaussienne, pour la borne supérieure, on a $\Phi\left(\frac{(260-250)\sqrt{2\cdot 2}}{\sqrt{500}}\right) = \Phi\left(\frac{2}{\sqrt{5}}\right) = \Phi(0.89) \approx 81.4\%$.

Et $0.8144 - (1 - 0.8144) \approx 0.629$

E) Plus de 70%

```
> x = rbinom(1e6, 500, 1/2)
> mean((x>=240)&(x<=260))
[1] 0.652975
> mean((x>240)&(x<260))
[1] 0.605244
> pbinom(260, 500, 1/2)-pbinom(240, 500, 1/2)
[1] 0.6284125
```

6 On injecte un vaccin à 2000 personnes. La probabilité qu'une personne prise au hasard souffre d'effets secondaires est de 1 chance sur 1000.

Quelle est la probabilité que, sur les 2000 personnes vaccinées, aucune ne souffre d'effet secondaire ?

A) Moins de 10%

B) Entre 10% et 15%, $e^{-2000/1000} = e^{-2} \approx 0.1353$

C) Entre 15% et 20%

D) Entre 20% et 25%

E) Plus de 25%

```
> dbinom(0,2000,1/1000)
[1] 0.1351999
> dpois(0,2000/1000)
[1] 0.1353353
> x = rbinom(1e6, 2000, 1/1000)
> x
[1] 3 1 2 8 0 3 2 2 0 1 1 2 1 6 5 3 4 4 1 1 3 1 3 0 0 3 1 2 3 4
> mean(x==0)
[1] 0.135388
```

7 À un examen, on suppose que la note X (sur 100) obtenue par les élèves suit une loi normale. Les statistiques de l'an passée permettent d'affirmer que $\mathbb{E}[X] = 70$ et $\text{Var}[X] = 10^2$. Le barême des notes pour obtenir la note littéral est C+ : $[72, 75)$ et B- : $[75, 78)$ (etc).

Quelle est la probabilité d'obtenir une note supérieure à B- ?

A) Environ 25%

B) Environ 30% , $1 - \Phi\left(\frac{75-70}{\sqrt{10^2}}\right)$ où $\Phi\left(\frac{1}{2}\right) \approx 0.6914$

C) Environ 35%

D) Environ 40%

E) Environ 50%

```
> 1-pnorm(75,70,10)
[1] 0.3085375
> x = rnorm(1e6, 70, 10)
> mean(x>75)
[1] 0.307957
```

8 On lance un dé bien balancés, de manière répétée.

Quelle est la probabilité que la face "3" apparaisse pour la première fois au 5ème lancer du dé ?

A) Moins de 2%

B) Entre 2% et 5%

C) Entre 5% et 10% , $5^4/6^5 = 625/7776 \approx 0.0804$

D) Entre 10% et 15%

E) Plus de 15%

```
> x = apply(M,1, function(x) (sum(x[1:4] == 3)==0)&(x[5] == 3))
> mean(x)
[1] 0.080307
> f = function(){
+     condition = FALSE
+     n = 0
+     repeat {
+         n = n + 1
+         x = sample(1:6, size=1)
+         if (x == 3) break
+     }
+     n
+ }
> x = rep(NA, 1e6)
> for(i in 1:1e6) x[i] = f()
> mean(x==5)
[1] 0.080209
```

9 Dans la salle des profs, 60% des professeurs sont des femmes. 1 femme sur 3 et 1 homme sur 2 portent des lunettes.

Quelle est la probabilité qu'une personne portant des lunettes, prise au hasard, soit une femme ?

A) 40%

B) 45%

C) 50% , $\dfrac{0.6 \cdot 1/3}{0.6 \cdot 1/3 + 0.4 \cdot 1/2} = \dfrac{1}{2}$

D) 55%

E) 60%

```
>
```

10 On considère une loterie, avec 10,000 tickets vendus, offrant 225 prix: 200 prix de \$5, 20 prix de \$250 et 5 prix de \$1000.

Quelle est l'espérance de gain pour un ticket acheté au hasard ?

A) \$1 , $\dfrac{200 \cdot 5 + 20 \cdot 250 + 5 \times 1000}{10000} = 1.1$ (on précisait qu'il fallait retenir le plus proche)

B) \$2

C) \$3

D) \$4

E) \$5

```
>
```

11 Le tableau ci-dessous indique la fréquence du nombre X d'erreurs par pages, dans un livre de 25 pages.

Nombre d'erreurs	0	1	2	3	Total
Fréquence	0,2	0,4	0,2	0,2	1

Quelle est la “variance corrigée” s^2 de X

A) 1.0203

B) 1.0833 , $\overline{x} = 0.4 \cdot 1 + 0.2 \cdot 2 + 0.2 \cdot 3 = 1.4$, $\dfrac{25}{24}\left(0.4 \cdot (1-\overline{x})^2 + 0.2 \cdot (2-\overline{x})^2 + 0.2 \cdot (3-\overline{x})^2\right) \approx 1.083333$

C) 1.1548

D) 1.2429

E) 1.4167

```
> x= rep(0:3, 25*c(.2,.4,.2,.2))
> x
 [1] 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3
> table(x)
x
 0  1  2  3
 5 10  5  5
> var(x)
[1] 1.083333
```

12 On vous propose le jeu suivant : vous lancez trois fois un dé non truqué à 6 faces. Si “6” sort trois fois, vous gagnez \$100. Sinon vous perdez \$1. On note X le gain total après 15 participations au jeu

Que vaut l’espérance de X ?

A) une perte espérée supérieure à \$5 (ou un gain espéré inférieur à -5%) , $100 \cdot \dfrac{1}{6^3} - 1 \cdot \left(1 - \dfrac{1}{6^3}\right) = \dfrac{100-215}{216} \approx -0.5324$ (qu’il faut ensuite multiplié par 15, ≈ -7.986)

B) une perte espérée comprise entre \$1 et \$5 (ou un gain espéré compris entre -5% et -1%)

C) un gain espéré compris entre -\$1 et +\$1

D) un gain espéré compris entre \$1 et \$5

E) un gain espéré supérieur à \$5

```
> M = matrix(sample(1:6, 3*15*1e6, replace=TRUE), 15*1e6, 3)
> x = apply(M,1, function(x) (sum(x == 6)==3))
> xm = c(-1,100)[1+x]
> X = matrix(xm,1e6,15)
> s = apply(X,1, sum)
> mean(s)
[1] -8.006659
```

13 Un étudiant prend en moyenne une heure pour taper une lettre. En supposant que le temps pour taper une lettre suit une loi exponentielle, trouver la probabilité que durant sa journée de travail de 8 heures, l’étudiant réussisse à taper les 12 lettres qu’on lui a demandé de taper.

A) environ 11% , on peut utiliser une approximation sur la durée moyenne pour taper 12 lettres $\overline{X} \approx \mathcal{N}\left(1, \dfrac{1}{\sqrt{12}}\right)$, et on veut $\mathbb{P}[X \leqslant 8/12] \approx \Phi(\sqrt{12}(8/12-1)) \approx 0.124$.

B) environ 21%

C) environ 31%

D) environ 41%

E) environ 51%

```r
> M = matrix(rexp(12*1e6), 1e6, 12)
> x = apply(M,1, sum)
> mean(x<8)
[1] 0.111805
```

Les questions 14, 15 et 16 sont basées sur la même histoire

14 Deux équipes de hockey, à Québec et à Montréal s'affrontent en ronde finale des séries éliminatoires. A chaque partie, les équipes s'affrontent, et il ne peut pas y avoir de match-nul: il y a forcément un vainqueur. Selon les experts (qui ont raison), Québec a 46% de chances de remporter chacune des parties, et on peut supposer les victoires aux parties comme étant indépendantes. La première équipe à obtenir 4 victoires gagne la ronde finale.

Quelle est la probabilité que la série se termine en 4 parties ?

A) moins de 10%

B) entre 10% et 20%, $0.46^4 + 0.54^4 \approx 0.1298051$

C) entre 20% et 30%

D) entre 30% et 40%

E) plus de 40%

```r
> M = matrix(sample(c("Q","M"),
                   size = 7*1e6,
                   replace = TRUE, prob = c(.46,.54)), 1e6, 7)
> vainqueurM = matrix(apply(M,1,function(x) cumsum(x == "M")), 7, 1e6)
> vainqueurQ = matrix(apply(M,1,function(x) cumsum(x == "Q")), 7, 1e6)
> vM = apply(vainqueurM, 2, function(x) min(which(x == 4)))
> vQ = apply(vainqueurQ, 2, function(x) min(which(x == 4)))
> (sum(vM==4) + sum(vQ==4))/1e6
[1] 0.130103
```

15 On sait que la ronde finale s'est terminée en 4 parties.

Quelle est la probabilité que Québec ait été le vainqueur de la ronde finale ?

A) moins de 10%

B) entre 10% et 20%

C) entre 20% et 30%

D) entre 30% et 40%, $\dfrac{0.46^4}{0.46^4 + 0.54^4} \approx 0.34493$

E) plus de 40%

```r
> (sum(vQ==4))/(sum(vM==4)+sum(vQ==4))
[1] 0.3445193
```

16 Quelle est la probabilité que Québec gagne la série en 7 parties ?

A) moins de 10%

B) entre 10% et 20% , $\dfrac{6!}{3!3!}0.46^3 \cdot 0.54^3 \cdot 0.46 \approx 0.1410076$

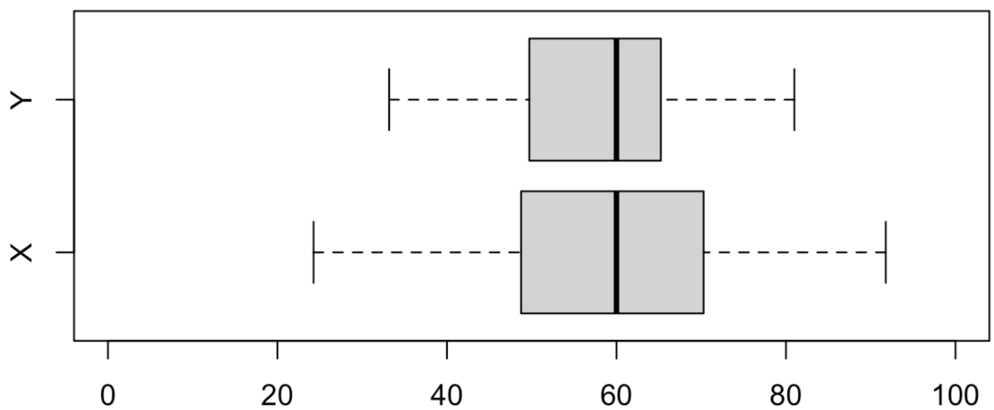
C) entre 20% et 30%

D) entre 30% et 40%

E) plus de 40%

```
> mean((vQ==7)&(vM>vQ))
[1] 0.141177
```

17 Un cours est donné à deux groupes, X et Y, et on obtient la distribution des notes finales, par groupe, sous la forme de boîtes à moustaches (ou *box plot*)



On a les quatre propositions suivantes (en sachant que la note de passage est de 60)

(1) le pourcentage d'échecs est (strictement) supérieur dans le groupe X

(2) la médiane est la même dans les deux groupes

(3) il y a moins d'individus dans le groupe Y

(4) le groupe X est plus homogène

Parmi les propositions, lesquelles sont vraies

A) aucune proposition n'est correcte

B) seulement (2)

C) (2) et (1)

D) (2) et (3)

E) toutes les propositions sont correctes

18 On considère une variable aléatoire continue X, de densité $f(x) = \dfrac{1}{2x}$ sur l'intervalle $\left[\dfrac{1}{e}, e\right]$ (où avec les notations usuelles $e = \exp(1)$). On dispose de trois observations $\{x_1, x_2, x_3\}$, indépendantes, de densité f. Quelle est la probabilité qu'une observation soit plus petite que 1 et que les deux autres soient plus grande que 1 ?

A) 1/3

B) 3/8 , $F(x)=\int_{1/e}^{x}\frac{dt}{2t}=\left[\frac{\log(t)}{2}\right]_{1/e}^{x}$ et $F(x)=\frac{\log(x)+1}{2}$, donc $F(1)=\frac{1}{2}$. $\mathbb{P}[X_1<1,X_2>,,X_3>1]=\frac{1}{2^3}$ donc $\frac{3}{2^3}=\frac{3}{8}$.

C) 1/4

D) 1/2

E) 3/4

```
> integrate(function(x) .5/x, exp(-1),1)
0.5 with absolute error < 5.6e-15
> Q = function(u) exp(2*u-1)
> M = matrix(Q(runif(3e6)), 1e6, 3)
> X = (M>1)*1
> x = apply(X,1,function(x) sum(x)==1)
> mean(x)
[1] 0.375149
> 3/8
[1] 0.375
```

19 Dans une salle, il y a 30 places numérotées (de 1 à 30). Il y a 18 élèves à lunettes, et 12 élèves sans lunettes. En supposant une répartition des élèves dans la salle indépendantes du fait de porter des lunettes ou pas, quelle est la probabilité que le pupitre 18 soit occupé par une personne à lunettes ?

A) 1/30

B) 2/5

C) 1/2

D) 3/5, $\frac{18}{30}=\frac{3}{5}$

E) 2/3

```
> sample(rep(c("avec L","sans L"),c(18,12)))[18]=="avec L"
[1] TRUE
> f = function() sample(rep(c("avec L","sans L"),c(18,12)))[18]=="avec L"
> x = rep(NA, 1e6)
> for(i in 1:1e6) x[i]=f()
> mean(x)
[1] 0.600495
> 3/5
[1] 0.6
```

Les questions 20 et 21 sont basées sur la même histoire

20 Une laiterie emplit des bouteilles de lait avec X millilitres de lait, par bouteille. On suppose que X suit une loi normale de moyenne 1000 et d'écart-type 40.

Parmi les milliers de bouteilles qui sortent de l'usine tous les jours, quelle proportion des bouteilles contiennent entre 950 et 1050 millilitres ?

A) moins de 50%

B) entre 50% et 60%

C) entre 60% et 75%

D) entre 75% et 85% , $1-2\Phi\left(\frac{1000-950}{40}\right)=1-2\Phi\left(-1.25\right)\approx 0.7887$

E) plus de 85%

```
> integrate(function(x) dnorm(x,1000,40),950,1050)
0.7887005 with absolute error < 8.8e-15
> pnorm(1050,1000,40)-pnorm(950,1000,40)
[1] 0.7887005
> x = rnorm(1e6,1000,40)
> mean((x>950)&(x<1050))
[1] 0.788141
```

21 Au moment de renouveler le stock de bouteilles vides, on veut des bouteilles ayant une contenance de telle sorte que 99% des bouteilles auront une quantité de lait inférieure à k.

Quelle est la valeur de k ?

A) 1051 millilitres

B) 1067 millilitres

C) 1071 millilitres

D) 1083 millilitres

E) 1093 millilitres , $40 \cdot \Phi^{-1}(.99) = 40 \cdot 2.326348 =\approx 93.0539$

```
> qnorm(.99,1000,40)
[1] 1093.054
> x = rnorm(1e6,1000,40)
> quantile(x,.99)
    99%
1093.34
```

22 Une boulangerie vend des petits pains artisanaux. On suppose que le nombre de petits pains vendus chaque heure suit une loi de Poisson de paramètre $\lambda = 14.2$. On suppose que les ventes chaque heure sont indépendantes.

Soit X le nombre de petits pains vendus entre 8 heures du matin et midi. Quel est l'écart-type de X ?

A) moins de 10 , $\sqrt{14.2 \cdot 4} \approx 7.536577$

B) entre 10 et 20

C) entre 20 et 30

D) entre 30 et 40

E) plus de 40

```
> x = rpois(1e6, 14.2)+rpois(1e6, 14.2)+rpois(1e6, 14.2)+rpois(1e6, 14.2)
> sd(x)
[1] 7.532612
> x = rpois(1e6, 4*14.2)
> sd(x)
[1] 7.535747
```

Les questions 23 et 24 sont basées sur la même histoire

23 Un grand nombre de candidats postulent à un emploi, et on fait passer une rapide entrevue pour déterminer si la personne a les exigences minimales pour le poste. On sait que 25% qui postulent ont les exigences minimales.

Quelle ait la probabilité qu'au bout de 4 entrevues, on ait au moins un candidat acceptable ?

A) moins de 50%

B) entre 50% et 65%

C) entre 65% et 70% , $1 - 0.75^4 \approx 0.6835938$

D) entre 70% et 85%

E) plus de 85% e

```
> M = matrix(rbinom(4e6,1, 1/4), 1e6, 4)
> x = apply(M, 1, sum)
> mean(x>0)
[1] 0.683241
```

24 Quelle ait la probabilité qu'au bout de 12 entrevues, on ait exactement 3 candidats acceptables ?

A) moins de 30%, $\dfrac{12!}{(12-3)!3!}0.25^3 \cdot 0.75^{12-3} \approx 0.2581036$

B) entre 30% et 40%

C) entre 40% et 50%

D) entre 50% et 60

E) plus de 60%

```
> M=matrix(rbinom(12*1e6,1, 1/4),1e6, 12)
> x = apply(M,1,sum)
> mean(x==3)
[1] 0.258301
```

25 Dans un examen à choix multiples, comprenant 25 questions, avec 5 choix de réponse pour chaque question (1 bonne réponse et 4 mauvaises), on gagne 4 points par bonne réponse, et toute réponse incorrecte rapporte 0. Un étudiant passe beaucoup de temps sur 15 questions, où il estime avoir, pour chaque question 2 chances sur 3 d'avoir la bonne réponse (et c'est effectivement le cas). Et pris par le temps, il décide de répondre aux 10 questions restantes au hasard. Soit X la note obtenue par l'élève a final.

Que vaut l'espérance de X ?

A) moins de 45

B) entre 45 et 50, $4\left(15\dfrac{2}{3} + 10\dfrac{1}{5}\right) = 48$

C) entre 50 et 55

D) entre 55 et 60

E) plus de 60

Code permanent : ABCD12345678 Sujet :

question 1	□ A	□ B	□ C	■ D	□ E
question 2	□ A	■ B	□ C	□ D	□ E
question 3	□ A	□ B	■ C	□ D	□ E
question 4	□ A	■ B	□ C	□ D	□ E
question 5	□ A	□ B	□ C	■ D	□ E
question 6	□ A	■ B	□ C	□ D	□ E
question 7	□ A	■ B	□ C	□ D	□ E
question 8	□ A	□ B	■ C	□ D	□ E
question 9	□ A	□ B	■ C	□ D	□ E
question 10	■ A	□ B	□ C	□ D	□ E
question 11	□ A	■ B	□ C	□ D	□ E
question 12	■ A	□ B	□ C	□ D	□ E
question 13	■ A	□ B	□ C	□ D	□ E
question 14	□ A	■ B	□ C	□ D	□ E
question 15	□ A	□ B	□ C	■ D	□ E
question 16	□ A	■ B	□ C	□ D	□ E
question 17	□ A	■ B	□ C	□ D	□ E
question 18	□ A	■ B	□ C	□ D	□ E
question 19	□ A	□ B	□ C	■ D	□ E
question 20	□ A	□ B	□ C	■ D	□ E
question 21	□ A	□ B	□ C	□ D	■ E
question 22	■ A	□ B	□ C	□ D	□ E
question 23	□ A	□ B	■ C	□ D	□ E
question 24	■ A	□ B	□ C	□ D	□ E
question 25	□ A	■ B	□ C	□ D	□ E