Université Paris II
L2 Sciences Économiques et Gestion

Statistiques 4 (5289)
Cours de M$^{\text{me}}$ Chevalier

Mai 2016

Les documents et les calculatrices ne sont pas autorisés.
Le barème est donné à titre indicatif et peut évoluer. Le sujet est noté sur 12.
Vous pouvez traiter les exercices dans l'ordre de votre choix, mais commencez si possible chaque exercice en haut d'une nouvelle page de votre copie et pensez à numéroter les pages.

Exercice 1 (3,5 points)

Soit X_1, X_2 et X_3 trois v.a. indépendantes de même loi normale standard.

1. (a) Déterminer la loi de probabilité du vecteur Y dont les composantes sont définies par
$$Y_1 = X_1 - 2X_2 + X_3 \qquad Y_2 = X_1 - X_3 \qquad Y_3 = -X_1 + 3X_2$$
   (b) Déterminer $\mathrm{Cov}(Y_1, Y_3)$.
   (c) Montrer que Y_1 suit une loi normale centrée de variance 6.
   (d) On peut montrer de même (et on l'admettra ici) que Y_2 suit une loi normale centrée $\mathcal{N}(0, \sqrt{2})$. Montrer que Y_1 et Y_2 sont indépendantes. On citera précisément le résultat du cours utilisé.
2. (a) Déduire de la question 1 la loi de probabilité de la variable $U = \frac{1}{6}(X_1 - 2X_2 + X_3)^2$.
   (b) On définit la variable aléatoire $V = Y_2/\sqrt{2U}$. Quelle est la loi de V ?

Exercice 2 (1,5 points)

Soit (X, Y) un couple de variables aléatoires dont la loi est définie par la densité
$$f(x, y) = \begin{cases} e^{-x} & \text{si } 0 \leqslant y \leqslant x \\ & 0 \text{ sinon} \end{cases}$$

1. Calculer la densité marginale de Y.
2. Pour $y \geqslant 0$, déterminer la densité conditionnelle de X sachant $Y = y$, c'est-à-dire la densité $f_X(x|Y = y)$ définie pour $x \geqslant y$.
3. Calculer la fonction de régression de X en Y, c'est-à-dire la fonction $y \mapsto E(X|Y = y)$ définie sur $\mathbb{R}_+$.
   On admettra qu'une primitive de la fonction $x \mapsto x\,e^{-x}$ est la fonction $x \mapsto -(x+1)e^{-x}$.

Exercice 3 (2,5 points)

1. Soit $(X_n)_{n\geqslant 2}$ une suite de v.a. mutuellement indépendantes, dont la loi est donnée par
$$\forall n \geqslant 2 \qquad \begin{cases} P(X_n = 0) = \dfrac{1}{4} - \dfrac{2}{n} \\ P(X_n = 1) = \dfrac{3}{4} + \dfrac{1}{n} \\ P(X_n = 2) = \dfrac{1}{n} \\ P(X_n = k) = 0 \text{ si } k \geqslant 2 \end{cases}$$
   Montrer que la suite (X_n) converge en loi vers une loi que l'on précisera.

2. Soit $(\mathrm{Y}_n)_{n\geqslant 1}$ une suite de v.a. mutuellement indépendantes, dont la loi est donnée par

$$\forall n \geqslant 1 \qquad \begin{cases} \mathrm{P}(\mathrm{Y}_n = 0) = 1 - \dfrac{1}{n} \\ \mathrm{P}(\mathrm{Y}_n = n) = \dfrac{1}{n} \\ \mathrm{P}(\mathrm{Y}_n = k) = 0 \text{ si } k \notin \{0, n\} \end{cases}$$

On admet que $\mathrm{E}(\mathrm{Y}_n) = 1$ et $\mathrm{Var}(\mathrm{Y}_n) = n - 1$ pour tout $n \in \mathbb{N}^*$.

(a) Montrer que la suite (Y_n) converge en probabilité vers 0.

(b) Montrer qu'elle ne converge pas en moyenne quadratique vers 0.

(c) Montrer que la suite $(\overline{\mathrm{Y}}_n)$ converge en probabilité vers 1.

(d) Vers quoi converge en probabilité la suite $(3\overline{\mathrm{Y}}_n + 4)$?

Exercice 4 (4,5 points)

Les dirigeants d'une entreprise souhaitent évaluer le nombre d'incidents par jour et par employé à une chaîne de montage. On suppose que ce nombre d'incidents peut être modélisé par une variable aléatoire X qui suit une loi de Poisson de paramètre λ, où λ est un réel positif inconnu qu'on cherche à estimer. Rappelons qu'une telle loi est définie par

$$\forall x \in \mathbb{N} \qquad \mathrm{P}(\mathrm{X} = x) = \mathrm{e}^{-\lambda}\frac{\lambda^x}{x!}$$

Le nombre d'incidents par jour et par employé observés pendant n journées choisies au hasard fournit ainsi un échantillon $(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ de la variable aléatoire X.

1. Construction d'un estimateur du paramètre λ

   (a) Montrer qu'un estimateur T_n de λ construit par la méthode des moments est $\mathrm{T}_n = \overline{\mathrm{X}}_n$.

   (b) Montrer que l'estimateur de λ obtenu par la méthode du maximum de vraisemblance est également T_n.

   (c) Montrer que l'information de Fisher est égale à $\mathrm{I}_n(\lambda) = n/\lambda$.

   (d) Montrer que T_n est sans biais et convergent.

   (e) Est-il efficace ?

   (f) Sur une période d'observation de $n = 400$ jours, les observations $(x_1, \ldots, x_n)$ sont telles que $\sum_{i=1}^{n} x_i = 1600$. Donner l'estimation de λ obtenue avec ces valeurs.

2. Construction d'un intervalle de confiance pour le paramètre λ

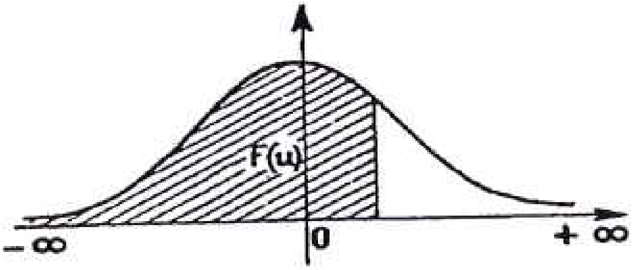
   (a) Quelle approximation peut-on utiliser pour la loi de T_n si n est suffisamment grand ?

   (b) En déduire un intervalle de confiance bilatéral symétrique au niveau de confiance $1 - \alpha$ pour le paramètre λ.

   On remplacera pour cela l'écart-type de la variable X *dans les bornes de l'intervalle par un estimateur.*

   (c) On choisit un seuil de confiance $1 - \alpha = 90\%$. On rappelle que sur une période de $n = 400$ jours, les observations $(x_1, \ldots, x_n)$ sont telles que $\sum_{i=1}^{n} x_i = 1600$. Donner la réalisation de l'intervalle de confiance obtenue avec ces valeurs (en arrondissant à deux chiffres après la virgule).

FONCTION DE REPARTITION DE LA LOI NORMALE REDUITE
(Probabilité de trouver une valeur inférieure à u)



u	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Table pour les grandes valeurs de u

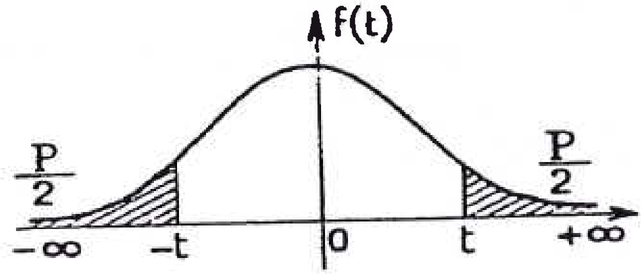
u	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9
F(u)	0,99865	0,99904	0,99931	0,99952	0,99966	0,99976	0,999841	0,999928	0,999968	0,999997

Nota – La table donne les valeurs de F(u) pour u positif. Lorsque u est négatif il faut prendre le complément à l'unité de la valeur lue dans la table.

Exemple — pour u = 1,37 — F(u) = 0,9147
pour u = -1,37 — F(u) = 0,0853

TABLE DE DISTRIBUTION DE t
(Loi de Student)

Valeurs de t ayant la probabilité P d'être dépassées en valeur absolue



ν \ P	0,900	0,800	0,500	0,200	0,100	0,050	0,020	0,010	0,001
1	0,158	0,325	1,000	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,816	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,765	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,741	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,727	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,711	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,706	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,695	1,358	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,694	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,692	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,691	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,690	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,681	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,679	1,295	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,677	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,674	1,282	1,645	1,960	2,326	2,576	3,291