Année Universitaire
2019/2020

UCAO-IEG: Licence 2

Examen: Statistique 4

Durée: 2 heures

Enseignant: Prof. YODE Armel

Exercice 1 **Questions de Cours**

(1) Donner la définition d'un échantillon de taille n.
(2) Qu'est ce qu'une statistique exhaustive pour un paramètre θ?
(3) Donner et expliquer le théorème de factorisation.
(4) Donner les deux formules de l'information de Fisher. Que représente l'information de Fisher?
(5) Qu'appelle-t-on estimateur sans biais d'un paramètre inconnu θ ?
(6) Qu'appelle-t-on borne de Cramer-Rao pour l'estimation sans biais d'un paramètre inconnu θ ?
(7) Qu'est ce qu'un estimateur efficace?
(8) Qu'appelle-t-on estimateur convergeant?
(9) Expliquer la méthode du maximum de vraisemblance.
(10) Quelles sont les différentes étapes d'un test d'hypothèses?

Exercice 2 Afin de mieux gérer les demandes de crédits de ses clients, un directeur d'agence bancaire réalise une étude relative à la durée de traitement des dossiers, supposée suivre une distribution normale. Un échantillon de 30 dossiers a donné :

Durée de taitement (en jours)	$[0, 10[$	$[10, 20[$	$[20, 30[$	$[30, 40[$	$[40, 50[$	$[50, 60[$
Effectif	3	6	10	7	3	1

(1) Déterminer les estimateurs de la moyenne m par la méthode du maximum de vraisemblance. Etudier ses propriétés.
(2) Donner une estimation de m par intervalle de confiance au seuil de risque 5%.
(3) Au seuil de 5%, tester l'hypothèse $H_0 : m = 30$ contre $H_1 : m \neq 30$. Que pouvez-vous conclure?

Exercice 3 On considère un échantillon $(X_1, \ldots, X_n)$ issu de la loi exponentielle $\mathcal{E}\left(\frac{1}{\theta}\right)$ avec $\theta > 0$ inconnu.

(1) Déterminer l'estimateur $\widehat{\theta}_n$ par la methode du maximum de vraisemblance.
(2) Montrer que $\widehat{\theta}_n$ peut être obtenu par la methode des moments.
(3) L'estimateur $\widehat{\theta}_n$ est-il efficace?

Exercice 4 Une enquête concernant l'utilisation des cartes bancaires (CB) a été effectuée en septembre 2005 auprès des personnes agées de 18 ans. Les résultats (partiels) de cette enquête sont présentés dans le tableau ci-dessous :

Description	Effectif
Personnes interrogées	501
Porteurs de CB	433
ayant effectué au moins un achat par CB	400
ayant effectué au moins un achat par CB sur Internet	144

Dans la suite, on s'intéresse à la proportion p **de personnes ayant effectué un achat par CB sur Internet parmi celles qui ont effectué au moins un achat par CB.**

(1) Donner le modèle théorique permettant l'étude de p : population, échantillon, variable aléatoire, loi.

(2) Donner un estimateur $\widehat{p}$ de p par la méthode du maximum de vraisemblance. Etudier les propriétés de l'estimateur $\widehat{p}$.

(3) Donner une estimation de p.

(4) Calculer un intervalle de confiance de niveau de confiance 95% pour p.

(5) Si on suppose constant le pourcentage de personnes interrogées ayant effectué au moins un achat par CB sur Internet, quelle devrait être la taille de l'échantillon pour connaitre p à 3% près (avec un niveau de confiance de 95%) ?

(6) En janvier 2005, une enquête similaire évaluait à 32% la part de personnes ayant effectué au moins un achat par CB sur Internet parmi celles ayant effectué au moins un achat par CB.

(a) Les données de l'enqute de septembre 2005 permettent-elles de conclure à une augmentation significative de la part de personnes utilisant leur CB sur Internet, en prenant un risque de première espèce de 1%?

(b) Quelle est la puissance du test lorsque $p = 34\%$?

Université Catholique Afrique de l'Ouest
Institut d'Economie et de Gestion (IEG)
UE: Statistique 4, L2

Année Universitaire
2019/2020

Contrôle continu: Durée: 1h30

NB: la clarté de la rédaction sera prise en compte dans la correction. Les documents sont formellement interdits

Exercice 1 (Question de cours)

(1) Soit un échantillon $(X_1, \cdots, X_n)$ issu d'un modèle statistique de paramètre $\theta \in \Theta$ Lors d'une discussion entre deux étudiants de L2 d'IEG, l'un d'eux affirme qu'il y a deux méthodes de détermination d'un estimateur de θ. Mais qu'il les a oubliés. Aider le à les retrouver et donner pour chaque méthode son principe de détermination.

(2) Dans le cas de l'estimateur par la méthode du maximum de vraisemblance. Donner les hypothèses sur le modèle statistique pour que sa détermination revienne à un simple calcul de zéro de la dérivée de la vraisemblance.

(3) On suppose que le modèle de loi $f(.,\theta)$ satisfait les hypothèses ci-dessus.
- (a) Donner la fonction de vraisemblance du modèle.
- (b) Déterminer la condition d'existence d'un estimateur du maximum de vraisemblance.
- (b) Donner la definition de la borne de Cramer-Rao.

(4) Soit un échantillon $(X_1, \cdots, X_n)$ issu d'un modèle statistique de moyenne m et de variance σ^2. Déterminer dans chaque cas l'intervalle de confiance de m de niveau β.
- (a) Le modèle est Gaussien et σ^2 est connu
- (b) Le modèle est Gaussien et σ^2 est inconnu
- (c) Le modèle est quelconque, $n > 30$ et σ^2 est inconnu.

(5) Soit un échantillon $(X_1, \cdots, X_n)$ de la loi exponentielle de paramètre θ.
- (a) Déterminer la moyenne théorique m du modèle.
- (b) Déterminer par la méthode des moment l'estimateur de m.

Donner un estimateur de la moyenne théorique

Exercice 2

On a construit un algorithme, noté $A1$, dont on souhaite déterminer les performances en terme de temps de calcul. Soit X_i le temps de réalisation de l'algorithme $A1$ sur la i-ème simulation. On suppose les X_i indépendants et de loi $\mathcal{N}(m, \sigma^2)$. On réalise $n = 41$ simulations. On obtient une moyenne empirique de 55 minutes et une variance empirique égale à 97,6 minutes^2.

(1) Donner des estimations sans biais convergentes de m et de σ.

(2) Donner l'intervalle de confiance de niveau 95% de m.

(3) On souhaite comparer avec un algorithme concurrent $A2$. Cet algorithme n'est malheureusement pas disponible et compliqué à implémenter. On décide donc de reprendre les données de l'article présentant $A2$. Il s'agit de simulations similaires aux précédentes mais la graine du générateur n'est a priori pas la même. Soient $Y_1, Y_2, \cdots, Y_n$ les temps obtenus par l'algorithme $A2$. On suppose les Y_i indépendants et de loi $\mathcal{N}(m_y, \sigma_y^2)$. On obtient une moyenne empirique $\bar{y}_n = 65$ minutes et un écart-type estimé $\sigma_y^2 = 14$ minutes.

(a) Donner l'intervalle de conance de niveau 95% de m_y.

(b) Comparer à l'intervalle donné en question 2. Est-il nécessaire d'implémenter l'algorithme $A2$?

Remarque : on ne demande pas ici de faire une procédure de test, uniquement de comparer les deux intervalles de confiance.

Fonction de répartition de la loi de normale $\mathcal{N}(0,1)$

Exemple : $P(\mathcal{N}(0,1) \leq 1{,}33) = 0{,}908\,2$.

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500 0	0,504 0	0,508 0	0,512 0	0,516 0	0,519 9	0,523 9	0,527 9	0,531 9	0,535 9
0,1	0,539 8	0,543 8	0,547 8	0,551 7	0,555 7	0,559 6	0,563 6	0,567 5	0,571 4	0,575 3
0,2	0,579 3	0,583 2	0,587 1	0,591 0	0,594 8	0,598 7	0,602 6	0,606 4	0,610 3	0,614 1
0,3	0,617 9	0,621 7	0,625 5	0,629 3	0,633 1	0,636 8	0,640 6	0,644 3	0,648 0	0,651 7
0,4	0,655 4	0,659 1	0,662 8	0,666 4	0,670 0	0,673 6	0,677 2	0,680 8	0,684 4	0,687 9
0,5	0,691 5	0,695 0	0,698 5	0,701 9	0,705 4	0,708 8	0,712 3	0,715 7	0,719 0	0,722 4
0,6	0,725 7	0,729 1	0,732 4	0,735 7	0,738 9	0,742 2	0,745 4	0,748 6	0,751 7	0,754 9
0,7	0,758 0	0,761 1	0,764 2	0,767 3	0,770 4	0,773 4	0,776 4	0,779 4	0,782 3	0,785 2
0,8	0,788 1	0,791 0	0,793 9	0,796 7	0,799 5	0,802 3	0,805 1	0,807 8	0,810 6	0,813 3
0,9	0,815 9	0,818 6	0,821 2	0,823 8	0,826 4	0,828 9	0,831 5	0,834 0	0,836 5	0,838 9
1,0	0,841 3	0,843 8	0,846 1	0,848 5	0,850 8	0,853 1	0,855 4	0,857 7	0,859 9	0,862 1
1,1	0,864 3	0,866 5	0,868 6	0,870 8	0,872 9	0,874 9	0,877 0	0,879 0	0,881 0	0,883 0
1,2	0,884 9	0,886 9	0,888 8	0,890 7	0,892 5	0,894 4	0,896 2	0,898 0	0,899 7	0,901 5
1,3	0,903 2	0,904 9	0,906 6	0,908 2	0,909 9	0,911 5	0,913 1	0,914 7	0,916 2	0,917 7
1,4	0,919 2	0,920 7	0,922 2	0,923 6	0,925 1	0,926 5	0,927 9	0,929 2	0,930 6	0,931 9
1,5	0,933 2	0,934 5	0,935 7	0,937 0	0,938 2	0,939 4	0,940 6	0,941 8	0,942 9	0,944 1
1,6	0,945 2	0,946 3	0,947 4	0,948 4	0,949 5	0,950 5	0,951 5	0,952 5	0,953 5	0,954 5
1,7	0,955 4	0,956 4	0,957 3	0,958 2	0,959 1	0,959 9	0,960 8	0,961 6	0,962 5	0,963 3
1,8	0,964 1	0,964 9	0,965 6	0,966 4	0,967 1	0,967 8	0,968 6	0,969 3	0,969 9	0,970 6
1,9	0,971 3	0,971 9	0,972 6	0,973 2	0,973 8	0,974 4	0,975 0.	0,975 6	0,976 1	0,976 7
2,0	0,977 2	0,977 8	0,978 3	0,978 8	0,979 3	0,979 8	0,980 3	0,980 8	0,981 2	0,981 7
2,1	0,982 1	0,982 6	0,983 0	0,983 4	0,983 8	0,984 2	0,984 6	0,985 0	0,985 4	0,985 7
2,2	0,986 1	0,986 4	0,986 8	0,987 1	0,987 5	0,987 8	0,988 1	0,988 4	0,988 7	0,989 0
2,3	0,989 3	0,989 6	0,989 8	0,990 1	0,990 4	0,990 6	0,990 9	0,991 1	0,991 3	0,991 6
2,4	0,991 8	0,992 0	0,992 2	0,992 5	0,992 7	0,992 9	0,993 1	0,993 2	0,993 4	0,993 6
2,5	0,993 8	0,994 0	0,994 1	0,994 3	0,994 5	0,994 6	0,994 8	0,994 9	0,995 1	0,995 2
2,6	0,995 3	0,995 5	0,995 6	0,995 7	0,995 9	0,996 0	0,996 1	0,996 2	0,996 3	0,996 4
2,7	0,996 5	0,996 6	0,996 7	0,996 8	0,996 9	0,997 0	0,997 1	0,997 2	0,997 3	0,997 4
2,8	0,997 4	0,997 5	0,997 6	0,997 7	0,997 7	0,997 8	0,997 9	0,997 9	0,998 0	0,998 1
2,9	0,998 1	0,998 2	0,998 2	0,998 3	0,998 4	0,998 4	0,998 5	0,998 5	0,998 6	0,998 6

t	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,8	4,0	4,5
P	0,998 65	0,999 04	0,999 31	0,999 52	0,999 66	0,999 76	0,999 841	0,999 928	0,999 968	0,999 997

FIGURE 1.