

Examen de Probabilités et Statistique du 15 Mai 2014

Durée : 4h

Les calculatrices sont autorisées

1 Statistique

1.1 Régression polynômiale

Soient p et n deux entiers tels que $0 \leq p < n$. Pour tout $1 \leq j \leq n$ on pose $x_j = j - \frac{n+1}{2}$. On considère $\psi_0, \psi_1, \dots, \psi_p$ polynômes fixés tels que :

- degré(ψ_i) = i , $\forall 0 \leq i \leq p$
- $\sum_{j=1}^n \psi_i(x_j)\psi_k(x_j) = 0$, $\forall 0 \leq i \neq k \leq p$

On a en vue d'étudier le modèle de régression

$$Y_j = \sum_{i=0}^p \lambda_i \psi_i(x_j) + \varepsilon_j, \quad 1 \leq j \leq n$$

où $\varepsilon_1, \dots, \varepsilon_n$ sont indépendantes et de même loi $\mathcal{N}(0, \sigma^2)$ et $\lambda_0, \dots, \lambda_p, \sigma^2$ sont des paramètres inconnus.

- a) Estimer les paramètres de ce modèle.
- b) Ecrire un test de niveau α de " $\lambda_p = 0$ " contre " $\lambda_p \neq 0$ ".

Le revenu net par action de la compagnie Gillette pour les années 57 à 64 est le suivant :

Année (z_j)	:	1957	1958	1959	1960	1961	1962	1963	1964
Revenu en \$ ($Y_j(\omega)$)	:	0,93	0,99	1,11	1,33	1,52	1,60	1,47	1,33

On pose $x_j = z_j - 1960,5$ puis $\psi_0(x) = 1$, $\psi_1(x) = 2x$ et $\psi_2(x) = x^2 - 21/4$.

c) Vérifier brièvement que le choix ci-dessus de $x_j, \psi_0, \psi_1, \psi_2$ rentre dans le cadre décrit plus haut.

d) En supposant que $Y_j = \sum_{i=0}^2 \lambda_i \psi_i(x_j) + \varepsilon_j$, $1 \leq j \leq 8$ avec $\varepsilon_1, \dots, \varepsilon_8$ indépendantes et de même loi $\mathcal{N}(0, \sigma^2)$, estimer $\lambda_0, \lambda_1, \lambda_2$ et σ^2 . Tester si $\lambda_2 = 0$.

e) Faire une prévision pour le revenu net par action en 1965.

N.B. Pour faciliter les calculs, on indique les valeurs numériques suivantes :

$$\sum_j \psi_1^2(x_j) = \sum_j \psi_2^2(x_j) = 168 \quad \sum_j Y_j^2(\omega) = 13,65 \quad \sum_j \psi_0(x_j)Y_j(\omega) = 10,28$$

$$\sum_j \psi_1(x_j)Y_j(\omega) = 6,86 \quad \sum_j \psi_2(x_j)Y_j(\omega) = -4,1.$$

1.2 Loi multinomiale

Soit n et l deux entiers naturels non nuls, on considère $N := (N_i)_{i=1, \dots, l}$ un vecteur aléatoire de loi multinomiale de paramètre n et p^* où $p^* := (p_i^*)_{i=1, \dots, l}$ avec $p_i^* > 0$, $i = 1, \dots, l$ et $\sum_{i=1}^l p_i^* = 1$.

1. Ecrire la log-vraisemblance $l_n(p)$ du modèle. Montrer que l'estimateur du maximum de vraisemblance de p^* est $\hat{p} = \frac{N}{n}$.

2. A partir de maintenant, on suppose que $l = 2r$ ($r \in \mathbb{N}_*$). On suppose que $p_{2i}^* = 2p_i^*$, $i = 1 \dots, r$. Déterminer l'estimateur du maximum de vraisemblance \widehat{p} de p^* .
3. Toujours sous l'hypothèse $p_{2i}^* = 2p_i^*$, $i = 1 \dots, r$, déterminer la distribution asymptotique de

$$T := \sum_{i=1}^{2r} \frac{(N_i - n\widehat{p}_i)^2}{n\widehat{p}_i}.$$

4. Expliquer comment tester l'hypothèse H_0 $p_{2i}^* = 2p_i^*$, $i = 1 \dots, r$. On suppose $r = 2$ et $n = 100$. On a observé $N_1^{\text{obs}} = 17$, $N_2^{\text{obs}} = 20$, $N_3^{\text{obs}} = 30$, $N_4^{\text{obs}} = 33$. Que décide-t-on ?

2 Chaîne de Markov

2.1 Chaîne de Markov à k états

Soit (X_n) une chaîne de Markov d'ensemble d'états E fini stationnaire de transition P . On pose :

$$Q = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n P^j$$

1. On suppose qu'il existe $\omega \geq 1$ tel que $P^{n+\omega} = P^n$ pour $n \geq 1$; montrer que

$$Q = \frac{1}{\omega} \sum_{j=1}^{\omega} P^j.$$

2. On suppose P primitive, c'est-à-dire qu'il existe $r \geq 1$ tel que $p_{ij}^{(r)} > 0$, $\forall i, j$. Montrer alors les propriétés suivantes (faire d'abord $r = 1$) :
 - (a) Si $Px = x$, $x \in \mathbb{R}^k$, $x \neq 0$ alors il existe x_1 tel que $x = x_1 e$ avec $e = (1, \dots, 1)^T$.
 - (b) Si $(P - I)^2 x = 0$ alors $(P - I)x = 0$ (on écrira $P = I + R$ et $P^n x = x + nRx$). En déduire que l'espace caractéristique associé à la valeur propre 1 est la droite engendrée par e .
 - (c) Si λ est une valeur propre de P telle que $\lambda \neq 1$, alors $|\lambda| < 1$.
 - (d) En déduire que P^n converge, à vitesse exponentielle vers Q , et que Q a toutes ses lignes égales, à coefficients strictement positifs.

2.2 File d'attente $G/G/1$

Des clients arrivent à une caisse de magasin. L'intervalle de temps entre l'arrivée du n -ème client et du $(n+1)$ -ème client est une v.a. A_n . Le temps de service du n -ème client est une v.a. B_n . Les v.a. (A_n) (resp. (B_n)) sont supposées de même loi, et toutes ces v.a. sont supposées indépendantes entre elles. Soit W_n le temps d'attente du n -ème client.

- 1) Trouver une relation entre W_{n+1} , W_n , A_n et B_n . Montrer que (W_n) est une chaîne de Markov. Donner la transition de la chaîne.
- 2) On pose $S_n = \sum_{k=1}^n (B_k - A_k)$ et $I = \inf S_n$. Montrer que $P(I = -\infty)$ vaut 0 ou 1 suivant que 0 est récurrent ou transitoire pour la chaîne (W_n) , (Etablir en premier lieu que $W_{n+1} = S_n - \inf_{0 \leq p \leq n} S_p$ $S_0 = 0$).

Table de la loi de Fisher 0,95

La table donne les valeurs des fractiles d'une loi de Fisher à (n_1, n_2) degrés de liberté pour une probabilité 0,95 i.e. les valeurs u telles que $P(X \leq u) = \int_{-\infty}^u f_X(x)dx = 0,95$ où X suit une loi de Fisher à (n_1, n_2) degrés de liberté, dont on rappelle ici l'expression de la densité :

$$f_X(x) = \frac{\Gamma[(n_1 + n_2)/2]}{\Gamma(n_1/2) \Gamma(n_2/2)} \cdot \left(\frac{n_1}{n_2}\right)^{n_1/2} \cdot \frac{x^{(n_1-2)/2}}{(1 + n_1 x/n_2)^{(n_1+n_2)/2}}.$$

La table donne donc les valeurs $F_{n_1, n_2}^{-1}(0,95)$ où F_{n_1, n_2}^{-1} est la fonction réciproque de la fonction de répartition d'une Fisher à (n_1, n_2) degrés de liberté. Par exemple, $F_{10, 20}^{-1}(0,95) = 2,35$. Remarque : $F_{n_1, n_2}^{-1}(\nu) = 1/F_{n_2, n_1}^{-1}(1 - \nu)$.

Avec Excel, on utilise la fonction =INVERSE.LOIF(1- ν ; n_1 ; n_2) pour obtenir $F_{n_1, n_2}^{-1}(\nu)$.

Avec R, la fonction : `qf`(ν, n_1, n_2) donne la valeur cherchée.

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	12	14
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.90	245.36
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.42
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.71
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.87
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.64
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.96
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.53
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.24
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.03
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.86
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.74
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.64
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.55
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.48
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.42
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.37
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.33
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.29
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.26
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.22
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.20
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.17
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.15
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.13
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.11
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.09
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.08
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.06
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.05
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.04
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.07	2.01
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.05	1.99
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.03	1.98
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.02	1.96
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.95
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.89
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.86
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.89	1.84
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.88	1.82
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.86	1.80
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.85	1.79
∞	3.84	3.00	2.61	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.69

