Ann. Sc. Math. Québec, 1979, Vol. III, No 1, pp. 119-133

ALTERNATIVES AUX ESTIMATEURS À VRAISEMBLANCE MAXIMALE DANS UN MODÈLE DE BERNOUILLI AVEC DÉPENDANCE

Marc Moore

1. INTRODUCTION

Soient $X_1,\dots,X_n$ n variables aléatoires formant une chaîne de Markov à deux états pour laquelle

$$P(X_i = 1) = 1 - P(X_i = 0) = p \ , \quad i = 1,\dots,n \tag{1.1}$$

et la matrice de transition est

$$\begin{array}{cc} & X_i \\ & \begin{array}{cc} \quad 0 \qquad\qquad & \qquad 1 \end{array} \\ X_{i-1} \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1 - p + \pi p & (1-\pi)p \\ (1-\pi)(1-p) & (1-\pi)p + \pi \end{pmatrix} \end{array} \tag{1.2}$$

$i = 2,\dots,n$, $0 \le p \le 1$, $0 \le \pi \le 1$. Le paramètre π est un indicateur de la persistance de l'état initial, si $\pi = 0$ le modèle se réduit à n essais de Bernoulli indépendants, si $\pi = 1$ l'état initial demeure avec probabilité un. Ce modèle a été utilisé à plusieurs reprises, par exemple par Switzer [11], [12] pour définir des subdivisions aléatoires d'une partie du plan (mosaïque), et par Schachter, Rosenfeld et Davis [10] pour caractériser certaines textures. Dans le présent article nous nous intéressons à l'estimation des paramètres p et π .

* *Recherche subventionnée par le Conseil National de Recherches, subvention A-8211.*

Klotz [7] considère un modèle semblable au modèle introduit plus haut. Il arrive à la conclusion qu'il est difficile d'obtenir les estimateurs à vraisemblance maximale (e.v.m.) pour les paramètres et suggère une méthode donnant des estimateurs dont le calcul est facile, et qui sont asymptotiquement équivalents aux e.v.m. Cette méthode est critiquée par Devore [5] (cette critique contient cependant des erreurs, voir les équations 6 à 9). Price [9] a étudié le même problème d'estimation à l'aide d'une simulation.

Les difficultés mentionnées par Klotz au sujet des e.v.m. pour son modèle sont aussi présentes, mêmes amplifiées, pour le modèle donné par (1.1) et (1.2). Nous voulons ici évaluer certaines alternatives aux e.v.m. pour les paramètres p et π. Les estimateurs proposés ont l'avantage d'être de calcul facile. Leurs propriétés asymptotiques ont été étudiées en [8], elles sont identiques à celle des e.v.m. Nous nous proposons d'étudier les propriétés exactes de ces estimateurs (biais et erreur quadratique moyenne) pour le cas d'échantillons de petite dimension ($n = 10,30,50$). Cette étude est basée sur une énumération complète de l'espace échantillon. A la section 2, la loi conjointe de $X_1,\ldots,X_n$ est obtenue. Les estimateurs considérés sont présentés à la section 3. La section 4 décrit la façon suivant laquelle les comparaisons sont effectuées et finalement la section 5 contient certaines conclusions.

2. LA LOI CONJOINTE

En utilisant (1.1), (1.2) et le fait que $X_1,\ldots,X_n$ forme une chaîne de Markov, la loi conjointe de $X_1,\ldots,X_n$ peut s'écrire

$$\begin{aligned} P(X_1 = x_1,\ldots,X_n = x_n) &= p^{x_1}(1-p)^{1-x_1}\prod_{i=2}^{n}\Big\{[(1-\pi)p+\pi]^{x_{i-1}x_i}[(1-\pi)(1-p)]^{x_{i-1}(1-x_i)} \\ &\quad \times [(1-\pi)p]^{(1-x_{i-1})x_i}(1-p+\pi p)^{(1-x_{i-1})(1-x_i)}\Big\} \qquad (2.1) \\ &= p^{s-r}(1-p)^{s-r-t+1}(1-\pi)^{2s-2r-t}[(1-\pi)p+\pi]^{r}(1-p+\pi p)^{u} \end{aligned}$$

où

$$r = \sum_{i=2}^{n} x_{i-1}x_i \ , \qquad s = \sum_{i=1}^{n} x_i \ , \qquad t = x_1 + x_n \ ,$$

$$u = \sum_{i=2}^{n} (1-x_{i-1})(1-x_i) = n - 1 - 2s + r + t \ .$$

De (2.1), il découle que $R = \sum_{i=2}^{n} X_{i-1}X_i$, $S = \sum_{i=1}^{n} X_i$, $T = X_1 + X_n$ forment un ensemble de statistiques exhaustives pour (p,π) . Les estimateurs considérés à la section suivante sont des fonctions de R , S , T seulement. Ces estimateurs sont des alternatives aux e.v.m., $\hat{p}$, $\hat{\pi}$ qui sont donnés par les solutions simultanées des équations suivantes,

$$(s-r)/p - (s-r-t+1)/(1-p) + r(1-\pi)/[(1-\pi)p+\pi] + u(\pi-1)/(1-p+\pi p) = 0 \qquad (2.2)$$

$$(2r+t-2s)/(1-\pi) + r(1-p)/[(1-\pi)p+\pi] + up/(1-p+\pi p) = 0 \ . \qquad (2.3)$$

Ces solutions sont difficiles à obtenir; cependant il est possible de trouver, par essais successifs, une approximation au point dans $\{(p,\pi) : 0 \leq p \leq 1 \ , 0 \leq \pi \leq 1\}$ où (2.1) atteint sa valeur maximale. Il est aussi possible d'exprimer la solution pour p comme une racine d'une équation du troisième degré (l'analogue de l'équation (9) de Devore [5]).

3. LES ESTIMATEURS

Trois couples d'estimateurs sont considérés pour (p,π) .

a) Le couple d'estimateurs proposé par Devore [2] se ramène dans notre contexte aux expressions simples:

$$p_D = S/n = \overline{X}$$

$$\pi_D = \begin{cases} 1 & \text{si } S = 0 \text{ ou } n \ . \\ \max\left\{0 \ , \dfrac{(R+U)/(n-1) + 2\overline{X}(1-\overline{X}) - 1}{2\overline{X}(1-\overline{X})}\right\} & \text{si } 0 < S < n \ . \end{cases} \qquad (3.1)$$

L'estimateur pour π est suggéré par le fait que de (1.1) et (1.2) nous avons,

$$P(X_1=0\,,\,X_2=0) + P(X_1=1\,,\,X_2=1) = (1-p)^2 + p^2 + 2\pi p(1-p)$$

alors

$$\pi = [P(X_1=0\,,\,X_2=0) + P(X_1=1\,,\,X_2=1) - p^2 - (1-p)^2]/2p(1-p)$$

et en faisant l'estimation de $P(X_1=0\,,\,X_2=0) + P(X_1=1\,,\,X_2=1)$ par $(R+U)/(n-1)$ et de p par $\overline{X}$ nous obtenons (3.1). Quand $S = 0$ ou n toutes les observations sont identiques, cela suggère de donner la valeur 1 à l'estimateur de π ; l'estimateur proposé pour π peut prendre une valeur négative lorsque $0 < S < n$, dans un tel cas sa valeur devrait être posée égale à zéro.

b) En adaptant à notre modèle la procédure de Klotz [7], c'est-à-dire en faisant l'estimation de p par $\overline{X}$ et l'estimation de π par la valeur qui rend la fonction de vraisemblance maximum après y avoir substitué $\overline{X}$ à la place de p , l'estimateur obtenu pour π , dénoté $\tilde{\pi}$, prend la valeur zéro si
$m\overline{X}^2 - \overline{X}(2S-T) + R \le 0$ et

$$\frac{1}{2}\left\{\frac{U-m}{m\overline{X}} + \frac{1-2\overline{X}}{1-\overline{X}} + \frac{R}{m(1-\overline{X})} + \left[\left(\frac{m-U}{m\overline{X}} - \frac{1-2\overline{X}}{1-\overline{X}} - \frac{R}{m(1-\overline{X})}\right)^2 + 4\left(\frac{\overline{X}}{1-\overline{X}} - \frac{2S-2R-T}{m(1-\overline{X})} + \frac{R(1-2\overline{X})}{m\overline{X}(1-\overline{X})}\right)\right]^{\frac{1}{2}}\right\} \quad (3.2)$$

si $m\overline{X}^2 - \overline{X}(2S-T) + R > 0$, $m = n - 1$. Si $S = 0$ ou n la valeur 1 est donnée à l'estimateur de π .

c) Si la contribution de la première observation X_1 est négligée dans la fonction de vraisemblance, alors en appliquant la théorie générale de Billingsley ([1], p. 26) l'estimation de la probabilité de transition de l'état 1 à l'état 1 peut se faire par $R/(S-X_n)$ (en supposant $S-X_n>0$) et l'estimation de la probabilité de transition de l'état 0 à l'état 0 peut se faire par $U/(n-1-S+X_n)$ (en supposant $n-1-S+X_n>0$) . En résolvant le système

$$\begin{aligned} R/(S-X_n) &= (1-\pi)p + \pi \\ U/(n-1-S+X_n) &= 1 - p + \pi p \end{aligned} \qquad (3.3)$$

nous obtenons les solutions suivantes pour p et π respectivement.

$$[1 - U/(n-1-S+X_n)]/[2 - R/(S-X_n) - U/(n-1-S+X_n)] \tag{3.4}$$

$$R/(S-X_n) + U/(n-1-S+X_n) - 1 . \tag{3.5}$$

Elles correspondent au point dans $\{(p,\pi) : 0 < p < 1 , 0 < \pi < 1\}$, quand il existe, où la fonction

$$p^{s-r-x_1}(1-p)^{s-r-x_n}(1-\pi)^{2s-2r-t}[(1-\pi)p+\pi]^r(1-p+\pi p)^u \tag{3.6}$$

atteint son maximum; cette fonction est (2.1) où la contribution de X_1 , $p^{x_1}(1-p)^{1-x_1}$, est négligée. Afin d'obtenir des estimateurs qui sont toujours bien définis et qui sont des fonctions de R , S et T seulement, à partir de (3.4) et (3.5) nous définissons les estimateurs suivants

$$\bar{p} = \begin{cases} 0 & \text{si } S = 0 \\ 1 & \text{si } S = n \\ [1 - U/(n-S)]/[2 - R/S - U/(n-S)] & \text{si } 0 < S < n \end{cases} \tag{3.7}$$

$$\bar{\pi} = \begin{cases} 1 & \text{si } S = 0 \text{ ou } n \\ \max\{0 , R/S + U/(n-S) - 1\} & \text{si } 0 < S < n . \end{cases} \tag{3.8}$$

4. COMPARAISONS

Les comparaisons entre les estimateurs considérés plus haut sont faites en considérant pour un n donné tous les éléments de l'espace échantillon munis de leur probabilité respective. De (2.1) on remarque que la loi conjointe de $X_1,\dots,X_n$ est constante pour des valeurs fixes de r , s et t . Alors il est suffisant de considérer tous les triplets (r,s,t) et d'obtenir le nombre de suites $x_1,\dots,x_n$ engendrant le même triplet. Etant donné n , r , s et t , il y a

$$\binom{2}{t}\binom{n-s-1}{s-r-t}\binom{s-1}{r} \tag{4.1}$$

suites $x_1,\ldots,x_n$ engendrant le triplet (r,s,t) (voir [6]). Alors pour p et π fixes, de (2.1) et (4.1), il découle que la probabilité portée par cette classe de suites est

$$\binom{2}{t}\binom{n-s-1}{s-r-t}\binom{s-1}{r} p^{s-r}(1-p)^{s-r-t+1}(1-\pi)^{2s-2r-t}[(1-\pi)p+\pi]^r(1-p+\pi p)^u \tag{4.2}$$

où $\binom{-1}{0} = \binom{-1}{-1} = 1$, par convention, pour les cas où $s = 0$ et $s = n$.

Afin de comparer les estimateurs pour des échantillons de petite taille, nous procédons de la façon suivante. Etant donné n , tous les triplets (r,s,t) tels que $t = 0,1,2$, $s = 1,\ldots,n-1$, $r = 0,\ldots,n-2$, $r \le s-1$, $s-r-t \ge 0$, et $n-s-1 \ge s-r-t$ sont énumérés et les triplets $(0,0,0)$, $(n-1,n,2)$ sont ajoutés. Il est possible de vérifier que le nombre de triplets ainsi obtenus est

$$\begin{aligned} &3n^2/4 - n + 2 \quad \text{si } n \text{ est pair} \\ &(3n^2+1)/4 - n + 2 \quad \text{si } n \text{ est impair} \end{aligned} \tag{4.3}$$

Pour chacun de ces triplets les estimateurs proposés pour p et π sont calculés (y compris les e.v.m.). Pour des valeurs données $\mathcal{p}$, π de p et π nous obtenons de (4.2) les probabilités correspondant aux classes de suites $x_1,\ldots,x_n$ donnant le même (r,s,t) et alors les mêmes estimateurs puisque tous les estimateurs proposés sont des fonctions de R , S et T seulement. De ces résultats nous pouvons calculer, pour chaque estimateur, des caractéristiques comme l'espérance mathématique, la variance, l'erreur quadratique moyenne, etc., ceci lorsque les vraies valeurs des paramètres sont $\mathcal{p}$ et π . Nous insistons sur le fait que puisque tous les triplets (r,s,t) sont énumérés les valeurs des caractéristiques sont exactes.

Les valeurs considérées pour n sont 10 (67 triplets), 30 (647 triplets) et 50 (1 827 triplets). Pour chaque valeur de n les vraies valeurs $\mathcal{p}$, π considérées pour p et π sont $\mathcal{p} = .1,.3,.5$ et pour chacune de ces valeurs de $\mathcal{p}$,

$\pi = .1, .3, .5, .7, .9$. Si π est fixe et p remplacé par $1 - p$, l'espérance de chaque estimateur proposé pour p (incluant l'e.v.m.) est changée en un moins l'espérance, tandis que la variance et l'erreur quadratique moyenne demeurent fixes; pour chaque estimateur de π (incluant l'e.v.m.) toutes les caractéristiques demeurent les mêmes. A cause de cette symétrie il est suffisant de considérer $p \leq .5$.

Les valeurs de deux caractéristiques, le biais relatif (B.R.) et l'erreur quadratique moyenne (E.Q.M.) sont présentées dans les tableaux I à IV . Le biais relatif est défini par

$$\text{B.R.} = \frac{100\,[\text{E(estimateur)} - \text{vraie valeur du paramètre}]}{\text{Vraie valeur du paramètre}}\ \%$$

(toutes les espérances sont considérées sous les vraies valeurs des paramètres). Il est clair que $\overline{X}$ est sans biais pour p . Pour mesurer la précision d'un estimateur, nous considérons l'erreur quadratique moyenne

$$\text{E.Q.M.} = \text{E}[(\text{estimateur} - \text{vraie valeur du paramètre})^2] .$$

p			.1					.3		
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
					n = 10					
$\overline{p}$	36.8	33.0	25.9	14.9	2.7	11.0	10.9	9.3	5.6	1.0
$\hat{p}$	.6	.8	.7	.3	.0	.4	.4	.3	.1	.0
					n = 30					
$\overline{p}$	26.0	29.9	33.4	32.9	14.7	4.7	5.9	7.7	9.6	5.5
$\hat{p}$	1.8	2.6	3.5	4.1	2.0	.0	.1	.3	.8	.7
					n = 50					
$\overline{p}$	17.1	21.1	26.6	32.6	23.9	2.9	3.7	5.0	7.5	8.3
$\hat{p}$	1.3	2.3	3.6	4.9	3.5	.0	.0	.1	.3	1.0

Quand p = .5 tous les B.R. sont zéro.

<u>Tableau I</u>

BIAIS RELATIF POUR LES ESTIMATEURS DE p

p	.1					.3					.5				
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
								n = 10							
π_D	356.0	92.0	40.6	18.3	4.9	57.3	-6.3	-11.3	-8.4	-3.1	27.4	-21.3	-22.5	-15.9	-5.7
$\tilde{\pi}$	327.0	83.4	36.1	15.7	3.8	43.7	-12.8	-16.7	-13.1	-5.7	22.5	-25.3	-27.2	-20.9	-8.8
$\bar{\pi}$	312.2	74.4	29.1	10.6	1.5	.5	-34.7	-33.1	-25.0	-10.9	-30.4	-51.6	-46.8	-35.0	-15.0
$\hat{\pi}$	328.0	83.9	36.6	16.2	4.0	45.7	-11.5	-15.5	-12.0	-5.2	24.7	-23.9	-25.7	-19.6	-8.2
								n = 30							
π_D	56.5	3.6	-.9	.7	1.9	9.7	-14.5	-13.6	-11.8	-6.3	9.5	-13.0	-11.3	-10.6	-8.1
$\tilde{\pi}$	45.9	-.1	-3.3	-1.0	1.1	9.1	-14.9	-14.2	-12.8	-7.7	9.3	-13.2	-11.6	-11.2	-9.6
$\bar{\pi}$	31.6	-7.8	-8.9	-5.2	-1.4	-14.2	-27.4	-23.6	-20.3	-13.1	-14.6	-26.3	-21.4	-19.3	-15.9
$\hat{\pi}$	80.8	13.5	5.0	3.8	2.9	9.6	-14.3	-13.1	-10.6	-4.8	9.5	-13.0	-11.2	-10.1	-6.5
								n = 50							
π_D	11.2	-12.4	-12.0	-7.4	-.5	4.2	-9.8	-8.4	-8.1	-6.4	4.0	-8.5	-6.5	-5.9	-6.6
$\tilde{\pi}$	7.1	-14.0	-13.2	-8.5	-1.2	3.9	-10.0	-8.5	-8.4	-7.2	4.0	-8.6	-6.6	-6.0	-7.3
$\bar{\pi}$	-4.5	-20.2	-17.7	-11.9	-3.3	-11.4	-18.1	-14.5	-13.3	-11.3	-11.9	-17.0	-12.5	-10.9	-11.9
$\hat{\pi}$	36.9	1.6	-.6	.5	1.8	4.0	-9.8	-8.2	-7.2	-4.4	4.0	-8.5	-6.4	-5.6	-5.1

Tableau II

BIAIS RELATIF POUR LES ESTIMATEURS DE π

p	.1					.3					.5				
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
								n = 10							
$\bar{X}$	1078	1561	2340	3740	6549	2515	3643	5461	8726	15280	2994	4337	6501	10388	18191
$\bar{p}$	1823	2483	3300	4439	6704	2435	3623	5400	8782	12478	2496	3706	5811	9857	18068
$\hat{p}$	1113	1598	2318	3600	6384	2548	3646	5334	8364	14892	2999	4298	6314	9942	17728
								n = 30							
$\bar{X}$	364	550	860	1544	3976	850	1271	2007	3604	9278	1012	1514	2389	4290	11045
$\bar{p}$	477	786	1344	2411	4693	821	1240	1989	3669	9398	946	1408	2195	3893	10609
$\hat{p}$	353	542	874	1545	3640	851	1267	1976	3436	8327	1012	1509	2356	4082	9858
								n = 50							
$\bar{X}$	219	330	526	964	2775	511	770	1226	2249	6476	609	916	1460	2678	7709
$\bar{p}$	252	410	730	1499	3745	500	757	1212	2257	6638	585	877	1385	2491	7205
$\hat{p}$	204	307	501	953	2569	511	768	1214	2177	5804	609	915	1446	2602	6872

Tableau III

ERREUR QUADRATIQUE MOYENNE $\times 10^5$ POUR LES ESTIMATEURS DE p

p			.1					.3					.5		
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
								$n = 10$							
π_D	33350	26328	18661	10334	2773	6729	9805	12192	11167	4999	3819	6944	10466	11212	5706
$\tilde{\pi}$	32840	26910	20027	12117	4175	6495	10225	13783	14340	8179	3730	7181	11817	14616	9456
$\bar{\pi}$	32120	26407	20357	13415	5463	5009	9733	15159	17640	11205	2151	6868	13683	18649	13066
$\hat{\pi}$	32943	27050	20128	12099	4076	6669	10438	13926	14256	7945	3888	7384	11951	14500	9177
								$n = 30$							
π_D	7093	10316	12104	10109	3159	1781	3432	4516	5439	4037	1656	2976	3299	3557	3886
$\tilde{\pi}$	6960	10573	12810	11195	4108	1767	3446	4661	6016	5607	1652	2981	3342	3828	5518
$\bar{\pi}$	6522	10238	12797	11588	4746	1400	3507	5168	6939	7756	1307	3102	3911	4835	7137
$\hat{\pi}$	7259	9441	10226	8121	2457	1783	3442	4437	4912	3151	1660	2993	3335	3475	3088
								$n = 50$							
π_D	2537	5462	8253	9024	3703	1234	2207	2449	2822	3238	1159	1893	1782	1571	2428
$\tilde{\pi}$	2512	5586	8612	9704	4445	1231	2211	2475	2957	4021	1158	1895	1789	1600	3061
$\bar{\pi}$	2269	5448	8664	9988	4878	1065	2282	2716	3382	4773	1006	1981	2007	1970	3816
$\hat{\pi}$	3157	5036	5769	4806	1765	1234	2215	2439	2425	1725	1160	1898	1790	1550	1458

Tableau IV

ERREUR QUADRATIQUE MOYENNE $\times 10^5$ POUR LES ESTIMATEURS DE π

5. CONCLUSIONS

Rappelons que le but de la présente étude est d'analyser le comportement de certains estimateurs simple à calculer en tant qu'alternatives aux e.v.m., dont le calcul est beaucoup plus complexe. En gardant ce but à l'esprit, nous pouvons faire les observations suivantes à partir des tableaux I à IV.

Estimation de p : Au niveau du biais, $\bar{X}$ l'emporte évidemment sur $\bar{p}$ comme alternative à $\hat{p}$. On remarque que $\hat{p}$ n'est que légèrement biaisé tandis que $\bar{p}$ l'est passablement. Pour ce qui est de l'E.Q.M., afin de faciliter l'examen du tableau III, on peut par exemple convenir de dire qu'un estimateur A est meilleur qu'un estimateur B (ou B est pire que A) si
(E.Q.M. de B - E.Q.M. de A) > (0,05) E.Q.M. de B (le 5% est arbitraire). Le tableau V présente pour chaque couple (p,π) l'estimateur qui est meilleur que les deux autres lorsqu'il y en a un et l'estimateur qui est pire que les deux autres lorsqu'il y en a un. On remarque que $\bar{X}$ n'est jamais meilleur que les deux autres, ni jamais pire que les deux autres. De plus, le tableau III illustre que

dans plusieurs cas il existe un des estimateurs $\overline{X}$ ou $\overline{p}$ qui a une plus petite E.Q.M. que celle de l'e.v.m. (c'est $\overline{X}$ quand p est petit ou grand). Le nombre de ces cas semble diminuer lorsque n augmente mais alors les différences diminuent aussi. Pour un p fixe, dans chaque cas, les E.Q.M. augmentent quand π augmente, i.e. quand on s'éloigne de l'indépendance.

le meilleur / le pire	p = .1					.3					.5				
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
n = 10 (le meilleur)										$\overline{p}$	$\overline{p}$	$\overline{p}$	$\overline{p}$		
n = 10 (le pire)	$\overline{p}$	$\overline{p}$	$\overline{p}$	$\overline{p}$											
n = 30 (le meilleur)					$\hat{p}$					$\hat{p}$	$\overline{p}$	$\overline{p}$	$\overline{p}$		$\hat{p}$
n = 30 (le pire)	$\overline{p}$	$\overline{p}$	$\overline{p}$												
n = 50 (le meilleur)	$\hat{p}$	$\hat{p}$			$\hat{p}$					$\hat{p}$					
n = 50 (le pire)			$\overline{p}$	$\overline{p}$											

<u>Tableau V</u>

CHOIX D'UN ESTIMATEUR POUR p

Estimation de π : Tous les estimateurs considérés pour π sont biaisés et souvent l'ordre de grandeur des B.R. pour π_D , $\tilde{\pi}$ et $\hat{\pi}$ sont les mêmes. En ce qui concerne l'E.Q.M., le tableau VI présente pour π l'analogue du tableau V pour p . On remarque que pour n = 10 , sauf dans le cas où π est petit, π_D est le meilleur des estimateurs. Lorsque n augmente, il arrive souvent que $\hat{\pi}$ soit le meilleur des estimateurs mais alors, parmi les alternatives π_D , $\tilde{\pi}$, $\overline{\pi}$, il est fréquent que π_D soit la meilleure des trois. De plus, si $.2 < p < .8$ et $\pi < .8$ les différences entre les E.Q.M. de π_D et de $\hat{\pi}$ ne sont pas très grandes.

le meilleur	p	.1					.3					.5				
le pire	π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
n = 10				π_D	π_D	π_D	$\bar{\pi}$		π_D	π_D	π_D	$\bar{\pi}$		π_D	π_D	π_D
					$\bar{\pi}$	$\bar{\pi}$			$\bar{\pi}$	$\bar{\pi}$	$\bar{\pi}$			$\bar{\pi}$	$\bar{\pi}$	$\bar{\pi}$
n = 30		$\bar{\pi}$	$\hat{\pi}$	$\hat{\pi}$	$\hat{\pi}$	$\hat{\pi}$	$\bar{\pi}$			$\hat{\pi}$	$\hat{\pi}$	$\bar{\pi}$				$\hat{\pi}$
									$\bar{\pi}$	$\bar{\pi}$	$\bar{\pi}$			$\bar{\pi}$	$\bar{\pi}$	$\bar{\pi}$
n = 50		$\bar{\pi}$	$\hat{\pi}$	$\hat{\pi}$	$\hat{\pi}$	$\hat{\pi}$	$\bar{\pi}$			$\hat{\pi}$	$\hat{\pi}$	$\bar{\pi}$				$\hat{\pi}$
		$\hat{\pi}$				$\bar{\pi}$			$\bar{\pi}$	$\bar{\pi}$	$\bar{\pi}$			$\bar{\pi}$	$\bar{\pi}$	$\bar{\pi}$

<u>Tableau VI</u>

CHOIX D'UN ESTIMATEUR POUR π

Tel que mentionné à la section 1, les propriétés asymptotiques des estimateurs considérés ici sont les mêmes que celles des e.v.m. Plus précisément, en utilisant le théorème 3.5.11 de Devore [2], en notant que le modèle (2.1) est un cas particulier des modèles markoviens généraux considérés par Billingsley [1], que sa condition 5.1 est satisfaite ici, qu'ainsi ses théorèmes 2.1 et 2.2 sont applicables et en utilisant des arguments parallèles à ceux utilisés par Klotz [7] nous pouvons démontrer que pour chaque $(p,\pi) \in \{(p,\pi) : 0 < p < 1\ ,\ 0 < \pi < 1\}$ tous les estimateurs considérés plus haut sont convergents et que les vecteurs aléatoires $\sqrt{n}(\hat{p}-p\ ,\ \hat{\pi}-\pi)$, $\sqrt{n}(\bar{X}-p\ ,\ \pi_D-\pi)$, $\sqrt{n}(\bar{X}-p\ ,\ \tilde{\pi}-\pi)$, $\sqrt{n}(\bar{p}-p\ ,\ \bar{\pi}-\pi)$ convergent tous en loi vers une loi binormale avec vecteur moyenne (0,0) et matrice de covariance

$$\Sigma = \begin{pmatrix} p(1-p)(1+\pi)/(1-\pi)\ , & (1-2p)\pi \\ & \\ (1-2p)\pi & ,\ (1-\pi)[p(1-p)(1-3\pi)+\pi]/p(1-p) \end{pmatrix} . \qquad (5.1)$$

De là, il découle que tous les estimateurs proposés sont asymptotiquement sans biais. Cette propriété doit être retenue quand en examinant les tableaux I et II on constate que parfois pour un couple (p,π) fixe la variation du B.R. n'est pas monotone en fonction de n ce qui est dû au fait que les estimateurs considérés sont tronqués.

En comparant la variance d'un estimateur avec sa variance asymptotique, nous pouvons avoir une idée de sa vitesse de convergence. A cette fin, nous avons calculé, tableaux VII et VIII, la déviation relative de chaque estimateur, celle-ci étant définie par

$$\text{D.R.} = \frac{100\,[\text{variance limite} - n \times \text{variance de l'estimateur}]}{n \times \text{variance de l'estimateur}}\ \%$$

les variances limites étant celles apparaissant dans la matrice (5.1). Les valeurs obtenues indiquent que $\overline{X}$ semble converger plus rapidement que $\hat{p}$ mais que $\hat{\pi}$ semble converger plus vite que π_D. A $n = 100$, $\overline{X}$ est très près de son comportement asymptotique tandis que π_D ne l'est pas.

p	.1					.3					.5				
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
								n = 10							
$\overline{X}$	2.0	7.0	15.4	36.4	161	2.0	7.0	15.4	36.4	161	2.0	7.0	15.4	36.4	161
$\overline{p}$	-34.8	-29.6	-16.5	21.5	155	10.5	14.2	18.4	35.9	219	22.4	25.3	29.1	43.7	163
$\hat{p}$	-1.1	4.6	16.5	41.7	168	.8	7.0	18.1	42.3	168	1.9	8.0	18.8	42.5	168
								n = 30							
$\overline{X}$	.7	2.2	4.6	10.1	43.3	.7	2.2	4.6	10.1	43.3	.7	2.2	4.6	10.1	43.3
$\overline{p}$	-10.4	-20.0	-27.0	-26.2	22.0	6.8	7.6	8.5	10.6	41.9	7.7	9.9	13.9	21.3	49.2
$\hat{p}$	4.0	2.9	3.1	10.2	56.6	.6	2.6	6.3	15.5	59.7	.6	2.5	6.1	15.7	60.6
								n = 50							
$\overline{X}$	.4	1.3	2.7	5.8	23.2	.4	1.3	2.7	5.8	23.2	.4	1.3	2.7	5.8	23.2
$\overline{p}$	-1.2	-8.5	-18.1	-26.8	-7.3	4.1	4.7	5.9	7.9	21.4	4.5	5.9	8.3	13.7	31.9
$\hat{p}$	8.0	8.9	8.0	7.3	33.2	.4	1.5	3.8	9.3	37.5	.3	1.5	3.7	8.9	38.2

<u>Tableau VII</u>

DÉVIATION RELATIVE POUR LES ESTIMATEURS DE p

p			.1					.3					.5		
π	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9	.1	.3	.5	.7	.9
								n = 10							
π_D	-21.1	28.5	73.8	130	222	65.4	9.5	-20.8	-38.0	-39.7	164	39.3	-18.5	-48.9	-65.1
$\tilde{\pi}$	-26.4	16.4	50.7	83.6	104	68.0	6.2	-28.1	-50.4	-67.3	169	37.8	-24.8	-59.1	-78.5
$\bar{\pi}$	-27.1	12.2	38.6	56.3	125	120	23.7	-24.3	-54.0	-74.8	381	103	-8.6	-59.7	-83.1
$\hat{\pi}$	-26.6	16.0	50.7	85.1	110	63.9	3.7	-29.4	-50.6	-66.5	159	32.4	-27.1	-59.5	-78.0
								n = 30							
π_D	-19.8	-22.3	-30.4	-33.9	-11.6	56.1	9.9	-22.6	-53.3	-76.8	100	7.4	-16.0	-43.4	-81.1
$\tilde{\pi}$	-19.5	-24.3	-34.1	-40.3	-32.5	101	9.9	-24.6	-57.2	-83.2	101	7.4	-16.8	-47.2	-86.7
$\bar{\pi}$	-15.4	-21.3	-33.1	-41.7	-41.5	156	26.0	-17.0	-54.6	-86.5	157	22.4	-9.6	-43.5	-87.6
$\hat{\pi}$	-17.8	-13.6	-17.1	-17.0	15.7	98.9	9.4	-21.8	-48.8	-70.9	99.8	6.7	-17.3	-42.9	-76.9
								n = 50							
π_D	29.1	-9.7	-35.9	-54.2	-55.0	71.7	.9	-17.2	-46.4	-82.2	71.1	-.4	-10.5	-27.2	-81.6
$\tilde{\pi}$	30.0	-11.2	-38.2	-57.2	-62.5	72.2	.8	-18.0	-48.7	-85.6	71.2	-.4	-10.8	-28.4	-85.6
$\bar{\pi}$	43.8	-5.4	-35.8	-56.9	-65.3	101	7.7	-14.1	-46.7	-86.2	99.6	5.8	-7.2	-26.5	-85.8
$\hat{\pi}$	7.9	-4.5	-12.4	-16.6	-4.6	71.8	.5	-17.1	-38.3	-67.1	70.9	-.7	-11.0	-26.8	-69.6

<u>Tableau VIII</u>

DÉVIATION RELATIVE POUR LES ESTIMATEURS DE π

De ce qui précède, nous conclusons qu'il n'est pas nécessaire de calculer les e.v.m. Dans le cas de p, $\bar{X}$ pourrait être utilisé parce qu'il est sans biais et que son E.Q.M. est plus petite quand p est petit (ou grand) et que les différences sont relativement importantes; pour les autres valeurs de p souvent l'E.Q.M. de $\bar{p}$ est plus petite mais les différences sont relativement moins importantes. Pour le paramètre π, π_D semble une bonne alternative à l'e.v.m. Si une information à priori au sujet de l'ordre de grandeur des paramètres est disponible, alors l'information résumée dans les tableaux V et VI peut être utilisée afin de choisir le meilleur estimateur lorsque celui-ci existe et éviter le pire.

L'auteur désire remercier messieurs N. Beaumier et H.Q. Lam pour leur aide dans l'écriture des programmes, le professeur Jérôme Klotz pour une conversation stimulante et les arbitres dont les commentaires ont permis la clarification de certaines parties du texte. Les formules (4.3) sont dues à un des arbitres. Ce travail fut revisé durant un séjour de l'auteur au Centre de recherches mathématiques de l'Université de Montréal.

BIBLIOGRAPHIE

[1] BILLINGSLEY, P., Statistical Inference for Markov Processes, The University of Chicago Press, Chicago (1961).

[2] DEVORE, J.L., Noisy Markov Chains, Technical Report No. 26, Department of Statistics, Stanford University (1971).

[3] DEVORE, J.L., Reconstructing a Noisy Markov Chain, Journal of the American Statistical Association, 68, 394-398 (1973).

[4] DEVORE, J.L., Reconstructing a Noisy Markov Chain Using Near-Neighbor Rules, Journal of the American Statistical Association, 68, 599-601 (1973).

[5] DEVORE, J.L., A note on the Estimation of Parameters in a Bernoulli Model with Dependence, The Annals of Statistics, 4, 990-992 (1976).

[6] KLOTZ, J., Markov Chain Clustering of Births by Sex, Proceeding of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume IV: Biology and Health, University of California Press, Berkeley (1972).

[7] KLOTZ, J., Statistical Inference in Bernoulli Trials with Dependence, The Annals of Statistics, 1, 373-379 (1973).

[8] MOORE, M., Comparison of Estimators for the Parameters in Bernoulli Trials with a Persistence Indicator, Rapport technique EP76-R-8, Ecole polytechnique, Montréal (1976).

[9] PRICE, B., A Note on Estimation in Bernoulli Trials with Dependence, Communication in Statistics, A5(7), 661-671 (1976).

[10] SCHACHTER, J., ROSENFELD, A. and DAVIS, L.S., Random Mosaic Models for Textures, Computer Science Technical Report Series No. 462, University of Maryland (1976).

[11] SWITZER, P., Reconstructing Patterns from Sample Data, The Annals of Mathematical Statistics, 38, 138-154 (1967).

[12] SWITZER, P., Mapping a Geographically Correlated Environment, Statistical Ecology, Proceedings of the 1969 International Symposium on Statistical Ecology, Yale University, 2, 235-270 (1971).

Département de mathématiques appliquées
Ecole polytechnique
C.P. 6079, Succ. A
Montréal, Québec
H3C 3A7

Manuscrit reçu le 15 septembre 1977.
Revisé le 26 mai 1978.