المجلة العراقية للعلوم الاحصائية (9) 2006
ص ص [92–101]

استخدام أنموذج سلسلة ماركوف لدراسة مرض التدرن الرئوي في محافظة صلاح الدين(1990–1999)

عبد الغفور جاسم سالم*

الملخص:

تمت في هذا البحث دراسة السلسلة الزمنية لعدد الإصابات بمرض التدرن الرئوي كمتسلسلة ماركوف ، بوضع افتراضات على عدد الإصابات لصياغة المسألة على وفق أنموذج متسلسلة ماركوف . تم إيجاد مصفوفة الانتقال من الدرجة (5×5) بالاعتماد على عدد الحالات . وبعد إيجاد صفات هذه السلسلة تبين أنها غير تخصصية (non-ergodic) ووجدنا احتمالات الاستيعاب (Absorbing Prob.) للحالات العابرة (transient states) وكذلك الزمن المتوقع للاستيعاب (The expected absorbing time) .

A Markov – Chain Model for the Tuberculosis in Salah Aldeen City (1991-1999)

ABSTRACT:

This paper studies a time series for the number of infection of Tuberculosis of the lungs as a Markov – chain. We have made hypotheses concerning the number of infections in order to set the problem according to Markov-chain. Then I'll try to find the characteristic of this chain that proved to be non-ergodic. I found the absorbing probability for the transient state and the expected absorbing time.

* أستاذ مساعد/قسم الرياضيات/ كلية التربية للبنات/ جامعة تكريت

تاريخ التسلم : 27/ 10 /2005 ــــــــــــــ تاريخ القبول : 1/ 12 /2005

المقدمة :

إن متسلسلة ماركوف(Markov-chain) هي حالة خاصة للعملية التصـــادفية (Stochastic Process) التي تشمل عــدداً مــن الحــالات (States) بمعلمــة (Parameter) زمنية . يعد العالم Markov (1906-1907) من الرواد في وضع المفاهيم الأساسية لمتسلسلة ماركوف وطور هذه المفاهيم العديد من الرياضـــيين. واستخدمت متسلسلات ماركوف في مجالات عديدة منهــا المجــال الاقتصـــادي ، المجال الزراعي ، وعلم الاجتماع.... حاول (1981) Battaglia دراسة منسوب النهر باعتباره متسلسلة ماركوف بحالتين (حالة الجفاف و حالة الفيضان) ، وهناك دراسة للسلسلة الزمنية للأمطار كمتسلسلة ماركوف بحالتين (حالة الجفاف و حالــة غزير المطر) [2] ، ودراسات عديدة أخرى في شتى المجالات .

تركز اهتمامنا في هذا البحث على صياغة عدد الإصابات بمرض التــدرن الرئوي بافتراضات بسيطة كمتسلسلة ماركوف وتم إيجاد صفات هــذه المتسلســلة حيث بينا بأنها مصفوفة غير تخصصية (Non-ergodic) ثم وجدنا الزمن المتوقع لاحتمال الاستيعاب للحالات العابرة (transient states) .

1- بعض التعاريف الأساسية :

ا- المجموعة المغلقة : لتكن Cمجموعة جزئية من فضاء الحالة S. تكــون Cمغلقة إذا كان $P_{ik} = 0$ (احتمال الانتقال من الحالــة i إلــى الحالــة k) لكــل $k \notin C, i \in C$. وإذا كانت المجموعة C المغلقة تمثل حالة واحدة (single state) فتسمى C بالحالة المستوعبة (absorbing state) [لاحظ3] .

ب- الحالة الثابتة (Persistent) : تسمى الحالــة j حالــة ثابتــة إذا كــان $P^*_{jj} = 1$. وإذا كان $P^*_{jj} < 1$ فتسمى j حالة عابرة (transient) حيث

$$P^*_{ij} = \sum_{n=1}^{\infty} P_{ii}^{(n)}$$

$$P_{ij}^{(n)} = P[X_{n+k} = j, X_{n+k-1} \neq j, ..., X_{k+1} \neq j, X_k \neq i]$$

ج-غير قابلة للتجزئة Irreducible : تكون متسلسلة ماركوف غير قابلة للتجزئة إذا وفقط إذا أمكن الانتقال من والى جميع الحالات . أي إذا كانت المجموعة المغلقة وحيدة . [4]

د- المتسلسلة التخصصية Ergodic : تكون متسلسلة ماركوف تخصصية (Ergodic) إذا كانت بدائية (Primitive) وغير قابلة للتجزئة . [2]
أي هي العملية التي يكون فيها الانتقال ممكنا من حالة واحدة الى بقية الحالات ولا يكون ضروريا على هذا الانتقال بخطوة واحدة . [3]

2- التحليل الإحصائي :

من خلال ملاحظة طبيعة البيانات والمؤشرات الأساسية المراد الوقوف عليها ، نجد أن الأسلوب المناسب للدراسة هو باستخدام متسلسلة ماركوف (Markov – chain) .
الخطوة الأولى للمسألة هي تعريف الخطوة (step) والحالة (state) لعدد المصابين لتستخدم في متسلسلة ماركوف . ومن ثم وضع عدة افتراضات لتشمل الحالات (States) وكيفية حركة هذه الحالات التي تمثل أعداد المرضى لتكوين مصفوفة الانتقال (transition matrix) .
في هذه المسألة عرفنا الحالة بعدد المرضى لفترة عشرين سنة ولفترة أكثر من مائة عام التي تمثل ذروة عمر الإنسان أما الخطوة فهي زيادة عدد المرضى من فترة زمنية (عشرون سنة) إلى أخرى .
الجدول الآتي يمثل عدد المرضى (ذكور وإناث) المصابين بالتدرن الرئوي والذي تم الحصول عليها ميدانيا من مستشفى صلاح الدين 1991-1999 بعد تبويبها .

	S_1	S_2	S_3	S_4	S_5
1991	38	62	51	51	7
1992	41	89	56	36	6
1993	67	91	76	48	5
1994	47	107	96	56	4
1995	401	137	88	47	4
1996	67	120	85	52	3
1997	49	160	89	50	3
1998	71	136	88	62	2
1999	91	161	91	58	2

حيث S_i $i = 1,2,3,4,5$ تمثل الحالات المختلفة لمتسلسلة ماركوف (فئات العمر).

إن حركة المرضى خلال الفئات التي تمثل كل منها عشرين سنة باعتبارها عملية عشوائية فقد وضع عمر المريض في واحدة من هذه الفئات الاتية :

الجدول 1: البيانات المبوبة

State	Class
S_1	0-19
S_2	20-39
S_3	40-59
S_4	60-79
S_5	80 and more

لتكوين مصفوفة الاحتمالات P (Transition probability Matrix) التي هي مصفوفة الانتقال من أية حالة إلى أخرى في وحدة زمنية واحدة ، لذا نحتاج إلى بيانات تصف حركة المرضى بصورة منفردة عبر الزمن وهذه البيانات غير متوفراة لعدة أسباب منها :

1- عدم مراجعة المريض في الأدوار الأولى للمرض. ومراجعة المريض لأكثر من طبيب .

2- تخوف المريض وإحجامه عن المراجعة .

3- شحة الأدوية والمستلزمات الطبية .

أما المتوافر من البيانات فيعطي معلومات عن العدد الكلي للمرضى بفئات العمر المختلفة واتخذت الفترة الزمنية (سنة واحدة) كفترة أساسية ملائمة لمصفوفة التحول التي يمكن تكوينها بوضع افتراضات ملائمة لوصف حركة المرضى بين فئات العمر المختلفة وهي :

1- أي مريض يصل إلى الحالة S_5 يبقى ضمنها .

2- أي مريض يزداد عمره سينتقل إلى المستوى الذي يكون أعلى منه ، فالزيادة في عدد المرضى في أية حالة S_i تأتي من الحالة السابقة لها مباشرة S_{i-1} .

3- التناقص في عدد المرضى ناتج عن انتقالهم إلى الحالة S_0 المتمثلة بحالة الشفاء .

وباستخدام هذه الافتراضات والبيانات المتوافرة لفترة سنة واحدة (فترة الأساسية) لذا يمكننا تبيان الحركة التقديرية للمرضى من حالة إلى أخرى خلال(1991-1992) كما موضحة في الجدول (2) .

الجدول (2): (1991-1992)

1991/1992	S_0	S_1	S_2	S_3	S_4	S_5	مجموع الصفوف
S_0							
S_1	3	38					41
S_2	16		62	11			89
S_3				40	16		56
S_4					35	1	36
S_5						6	6
مجموع الأعمدة	19	38	62	51	51	7	

States	S_1	S_2	S_3	S_4	S_5
1991	38	62	51	51	7
1992	41	89	56	36	6

وبنفس الطريقة يتم تكوين الجداول الأخرى الخاصة بحركة المرضى من حالة إلى أخرى خلال (93-94).....(98-99). وعندما نجمع هذه الجداول وفقاً لقاعدة جمع المصفوفات نحصل على الجدول التالي (الجدول 3) الذي يوضح التحـــولات (الحركات) التقديرية للمرضى من عام (1991) لغاية (1999) .

الجدول (3): (1991–1999)

S_i / S_{i-1}	S_0	S_1	S_2	S_3	S_4	S_5	مجمـــــوع الصفوف
S_0	0	0	0	0	0	0	0
S_1	134	309	39	0	0	0	489
S_2	26	120	846	30	0	0	1022
S_3	40	0	0	518	84	1	643
S_4	18	0	0	14	368	4	404
S_5	0	0	0	0	0	29	29
مجموع الأعمدة	218	429	885	562	452	34	

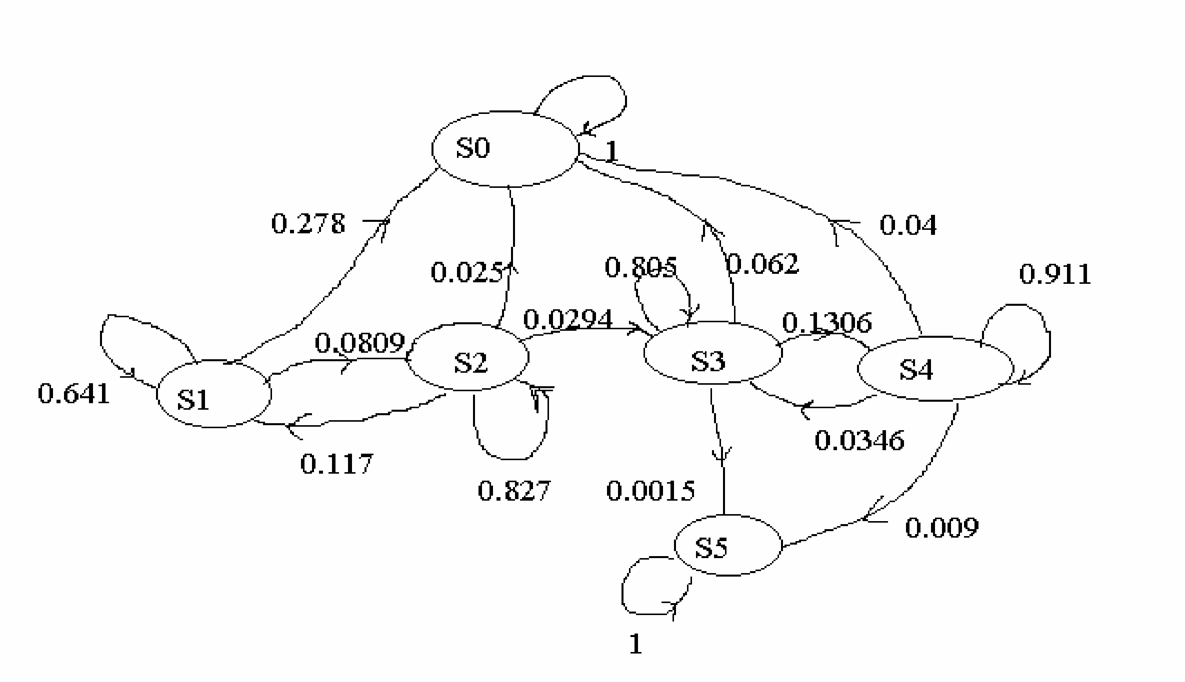
من الجدول الذي في أعلاه نقسم عناصر كل صف على المجمـــوع الكلـــي للصفوف الذي تقع فيه، فينتج عن ذلك مصـــفوفة عشـــوائية تســـتخدم كمصـــفوفة التحولات الاحتمالية للفترات (1991–1999) . وهي تعكس لنا الافتراضات الأولية عن حركة المرضى حيث S_0 و S_6 حالتان منتهيتان وحصلنا على الجدول الآتي (الجدول 4) .

الجدول (4): مصفوفة الانتقال

P=

	S_0	S_1	S_2	S_3	S_4	S_5
S_0	1	0	0	0	0	0
S_1	0.278	0.641	0.081	0	0	0
S_2	0.025	0.117	0.828	0.030	0	0
S_3	0.062	0	0	0.806	0.130	0.002
S_4	0.045	0	0	0.034	0.911	0.010
S_5	0	0	0	0	0	1

ويمكن تمثيل مصفوفة الانتقال P بمخطط شجيري (Digraph) الـــذي يوضـــح صفات متسلسلة ماركوف الممثلة بالمصفوفة P



المخطط الشجيري لمصفوفة الانتقال

نلاحظ من المخطط الشجيري ومن تعريف الحالة المغلقة أن الحالتان S_0 و S_5 مغلقتان وبما أن كل منهما يمثل حالة واحدة (single – state) فتكون كل من S_0 و S_5 حالة مستوعبة (absorbing – state) .

نستنتج من ذلك بأن المصفوفة P قابلة للتجزئة (reducible) ، وعليه تكون متسلسلة ماركوف الممثلة بالمصفوفة P غير تخصصية (non-ergodic)

احتمال الاستيعاب والقيمة المتوقعة لزمن الاستيعاب :

Absorbing probability and expected absorbing time

تعد الحركة (الانتقال) من الحالات العابرة (transient-states) إلى الحالات المغلقة الجزئية من الحالات الثابتة (persistent-states) من المسائل المهمة لمتسلسلات ماركوف غير التخصصية (non-ergodic) .[3].

وفي هذا البحث تم استخدام طريقة المصفوفة الأساسية (لاحظ 3) التي تتضمن أيجاد مصفوفة Q تتوافق مع الحالات العابرة .

ولحساب احتمالات الاستيعاب (absorbing probability) ، نجد المصفوفة N حيث $N=(I-Q)^{-1}$ عناصر المصفوفة N تمثل احتمالات الاستيعاب .

أما القيمة المتوقعة لزمن الاستيعاب (Expected absorbing time) فيتم إيجاده من جمع عناصر كل صف في المصفوفة N ليمثل القيمة المتوقعة لزمن الاستيعاب من الحالة المقابلة لذلك الصف .

بما أن الحالات العابرة (transient-states) هي S_1,S_2,S_3,S_4 وعليه تكون المصفوفة Q التي تمثل تلك الحالات

$$Q=\begin{bmatrix} 0.641 & 0.081 & 0 & 0 \\ 0.117 & 0.828 & 0.03 & 0 \\ 0 & 0 & 0.806 & 0.130 \\ 0 & 0 & 0.034 & 0.911 \end{bmatrix}$$

وبما أن $N=(I-Q)^{-1}$ فان

$$N=\begin{bmatrix} 3.291 & 1.550 & 0.322 & 0.470 \\ 2.238 & 0.868 & 1.428 & 2.085 \\ 0 & 0 & 6.928 & 10.120 \\ 0 & 0 & 2.647 & 15.102 \end{bmatrix}$$

وعليه يكون الزمن المتوقع للاستيعاب من تلك الحالات

الزمن المتوقع للاستيعاب	الحالات
5.633	S_1
12.619	S_2
17.048	S_3
17.749	S_4

أي أن الفترة الزمنية للذهاب إلى حالة الشفاء S_0 من S_1 هي مـــا يقـــارب الســـتة أشهر. وأن S_2 تبقى ما يقارب السنة وشهر لكي تتحول إلى الحالة S_0 المستوعبة. و S_3 تبقى ما يقارب السبعة عشر شهراً لكـــي تتحـــول إلـــى إحـــدى الحـــالات المستوعبة S_0 أو S_5. و S_4تبقى ما يقارب الثمانية عشر شهراً لكي تتحول إلى إحدى الحالات المستوعبة S_0 أو S_5.

الاستنتاجات :

من خلال وضع السلسلة الزمنية للإصابة بمرض التدرن الرئوي كمتسلســلة ماركوف ، تم إيجاد صفاتها وتبين أنها متسلسلة غير تخصصية (non-ergodic) لاحتوائها على حالتين مستوعبتين (absorbing-states) . وعند إيجــاد الــزمن المتوقع للاستيعاب من الحالات العابرة (transient-states) لاحظنـــا أن أســـرع الحالات انتقالاً للحالة المستوعبة S_1 التي تمثل العمر (0-19) سنة وأكثر الحالات بقاءً هي S_4 التي تمثل العمر (79-100) سنة لكي تنتقل إلى حالة مستوعبة .

المصادر :

1- Battaglia, F. (1981) "Up crossing of discrete hydrologic processes and Markov chains", J. of Hydrology, 53, 1-16.
2- Cox, D. R. and Miller, H. D. (1965) " The Theory of stochastic processes", Methuen, London.
3- Dean, L. Isaacson (1976) "Markov chains Theory and application", John Wiley and Sons, Inc.
4- Iosifescu, M. (1980) " Finite Markov processes and their applications" John Wiley and Sons, Inc.