

Werk

Label: Article

Jahr: 1959

PURL: https://resolver.sub.uni-goettingen.de/purl?31311157X_0084|log75

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

✉ info@digizeitschriften.de

ОБ АСИМПТОТИЧЕСКОМ ПОВЕДЕНИИ ВЕРОЯТНОСТЕЙ В ГРУППАХ СОСТОЯНИЙ ОДНОРОДНОЙ ЦЕПИ МАРКОВА

Петр МАНДЛ (Petr Mandl), Прага

(Поступило в редакцию 20/III 1958 г.)

DT: 519.217

В статье исследуется предельное поведение распределения вероятностей в классе состояний однородной цепи Маркова при условии, что пребывание системы в данном классе не нарушалось.

Пусть R — однородная цепь Маркова с конечным числом состояний, определенная матрицей вероятностей перехода и начальным распределением вероятностей. Под регулярностью вероятностей $p_j^{(n)}$, т. е. вероятностей того, что система будет во времени n находиться в состоянии j , мы будем понимать существование предела $\lim_{n \rightarrow \infty} p_j^{(n)}$ и независимость этих вероятностей от начального распределения. Вопрос регулярности вероятностей может быть дополнен вопросами об асимптотическом поведении величин $P(x_n = \omega/T^n)$, определенных следующим образом:

Если T — какая-то группа состояний цепи R и ω — состояние группы T , то выражение $P(x_n = \omega/T^n)$ представляет вероятность того, что система будет во времени n находиться в состоянии ω при условии, что она не выйдет за пределы класса T . Значит, T^n обозначает явление, что система осталась пребывать в классе T по крайней мере до времени n . Эта проблема является частным случаем вопроса, который поставил Я. Гаек (Математический институт ЧСАН), которому я обязан за ценные указания.

Можем, например, предполагать, что система может пребывать в состояниях двух видов: в таких, за которыми можно наблюдать, и в таких, которые не поддаются наблюдениям. Будем, следовательно, изучать распределение вероятностей при условии, что в течение продолжительного времени наблюдения за системой не велись.

Определение. Группу S состояний однородной цепи Маркова назовем *регулярной*, если матрица P вероятностей перехода в группе S неразложима и непериодична.

Теорема 1. В случае, когда T — регулярный класс, для каждого ω существует $\lim_{n \rightarrow \infty} P(x_n = \omega | T^n) = P_T(\omega) > 0$ и не зависит от начального распределения вероятностей.

Доказательство. Умножим строки матрицы P на положительные числа так, чтобы получилась стохастическая матрица. Используя известные теоремы о стохастических матрицах (см., напр. [3]), легко обнаружим, что вследствие предположения о матрице P существует n_0 такое, что все элементы P^{n_0} положительны. По теореме Перрона (см., напр., [1]) существует характеристический корень Θ матрицы P^{n_0} , который имеет следующие свойства:

1. Он является положительным и простым, 2. превосходит по абсолютной величине остальные корни, 3. матрица $\text{Adj}(\Theta E - P^{n_0})$ имеет только положительные элементы, равно как и 4. в качестве элементов соответствующего характеристического вектора можно взять сплошь положительные числа.

Итак, матрица P имеет корень $\varrho = \Theta^{\frac{1}{n_0}}$, который, очевидно, обладает свойствами 1, 2, 4.

Займемся представлением матрицы P^n при помощи формулы Перрона ([3]):

$$P^n = \varrho^n B_{11} + \sum_{\mu=2}^s \sum_{v=0}^{s_\mu-1} \binom{n}{v} \lambda_\mu^{n-v} B_{v,\mu} = \varrho^n B_{11} + o(\varrho^n),$$

где λ_μ — характеристические корни матрицы P , s_μ — их кратности, $B_{v,\mu}$ — матрицы и $o(\varrho^n)$ — матрица, элементы которой суть бесконечно малые величины порядка выше ϱ^n . В частности, матрица B_{11} имеет вид

$$B_{11} = \lim_{\lambda \rightarrow \varrho} (\lambda - \varrho)(\lambda E - P)^{-1} = \text{Adj}(\varrho E - P) \psi_P(\varrho)^{-1},$$

где $\psi_P(x) = (x - \varrho)^{-1} |xE - P|$. Очевидно, что должно быть $\psi_P(\varrho) > 0$, ибо в противном случае было бы ϱ или многократным корнем многочлена $|xE - P|$, или этот многочлен имел бы еще один действительный корень, который был бы больше ϱ .

Сравним выражение для матрицы P^n с аналогичным выражением для степеней матрицы P^{n_0} . Имеем

$$\varrho^{n \cdot n_0} B_{11} + o(\varrho^{n \cdot n_0}) = (P^{n_0})^n = \Theta^n \text{Adj}(\Theta E - P^{n_0}) \psi_{P^{n_0}}(\Theta)^{-1} + o(\Theta^n).$$

Итак, мы видим, что характеристический корень ϱ имеет свойство 3, т. е. матрица $B_{11} \psi_P(\varrho) = \text{Adj}(\varrho E - P)$ имеет только положительные элементы.

Пусть, далее, p — вектор начального распределения вероятностей (во времени $n = 0$), e — вектор-столбец, в котором все элементы равны единице,

e_ω — вектор-столбец, в котором единица стоит на том месте, где в матрице P состояние ω , на остальных местах нули. Имеем

$$\begin{aligned}\lim_{n \rightarrow \infty} P(x_n = \omega | T^n) &= \lim_{n \rightarrow \infty} \frac{P(\{x_n = \omega\} \cap T^n)}{P(T^n)} = \lim_{n \rightarrow \infty} \frac{pP^n e_\omega}{pP^n e} = \\ &= \lim_{n \rightarrow \infty} \frac{\varrho^n pB_{11}e_\omega + o(\varrho^n)}{\varrho^n pB_{11}e + o(\varrho^n)} = \frac{pB_{11}e_\omega}{pB_{11}e} = P_T(\omega),\end{aligned}$$

так как $pB_{11}e > 0$ ввиду того, что выполняется свойство 3.

Теперь можем писать

$$[P_T(\omega_1), \dots, P_T(\omega_s)] P = \lim_{n \rightarrow \infty} \frac{pP^{n+1}}{pP^n e} = \varrho \frac{pB_{11}}{pB_{11}e} = \varrho [P_T(\omega_1), \dots, P_T(\omega_s)]. \quad (1)$$

Значит, $[P_T(\omega_1), \dots, P_T(\omega_s)]$ есть собственный вектор матрицы P с суммой, равной единице. Следовательно, он обладает свойством 4 и ввиду того, что ϱ — простой корень, дефект матрицы $\varrho E - P$ равен единице; следовательно, $P_T(\omega)$ определено однозначно.

Замечание. Если $\{P_T(\omega)\}$ представляет первоначальное распределение вероятностей, то распределение вероятностей в группе T при условии, что система не выйдет за пределы группы T , не зависит от n . Время N выхода из группы T в таком случае подчинено, если считать $\varrho < 1$, распределению Паскаля, т. е. $P(N = n) = \varrho^{n-1}(1 - \varrho)$. Это распределение является предельным распределением в том смысле, что

$$\begin{aligned}\lim_{n \rightarrow \infty} P(N > n + m | N > n) &= \lim_{n \rightarrow \infty} P(T^{n+m} | T^n) = \\ &= \lim_{n \rightarrow \infty} \frac{pP^{n+m}e}{pP^n e} = [P_T(\omega_1), \dots, P_T(\omega_s)] P^m e = \varrho^m \sum_{\omega \in T} P_T(\omega) = \varrho^m.\end{aligned}$$

Доказательство теоремы 1 оправдывает введение следующего определения:

Определение. *Характеристическим числом* регулярной группы S назовем максимальное по абсолютной величине характеристическое число матрицы, образованной из вероятностей перехода внутри класса S .

Теперь предположим, что группа состояний T разбита на подклассы $T_1, \dots, T_l$ такие, что система, не выходя за пределы T , может с положительной вероятностью перейти из класса T_j в класс T_{j+1} и только в этот класс. Пусть P_{jj} означает матрицу вероятностей перехода внутри класса T_j , а P_{jj+1} — матрицу вероятностей перехода из состояний класса T_j в состояния класса T_{j+1} .

Теорема 2. В случае, когда классы T_j регулярны и $P(x_0 \in T_1) = 1$, для каждого $\omega \in T_1$ будет

$$\lim_{n \rightarrow \infty} \frac{P(\{x_n = \omega\} \cap T^n)}{\varrho^n n^{k-1}} = \frac{\varrho^{-(k-1)}}{(k-1)!} p A_{11} P_{12} A_{22} \dots P_{l-1, l} A_{ll} e_\omega,$$

где $\varrho = \max_{j=1, \dots, l} \varrho_j$, ϱ_j равно характеристическому числу класса T_j и число k указывает, сколько раз встретится характеристическое число ϱ среди чисел ϱ_j . Далее,

$$A_{jj} = (\varrho E - P_{jj})^{-1}, \quad \text{если } \varrho_j < \varrho,$$

$$A_{jj} = \lim_{\lambda \rightarrow \varrho} (\lambda - \varrho)(\lambda E - P_{jj})^{-1}, \quad \text{если } \varrho_j = \varrho$$

(A_{jj} есть коэффициент у ϱ^n в формуле Перрона).

Доказательство. Имеем $P(\{x_n = \omega\} \cap T^n) = \sum_{\Sigma n_i = n+1} p P_{11}^{n_1-1} P_{12} \dots P_{ll}^{n_l-1} e_\omega$. Обозначим $\sum_{\varrho_i < \varrho} n_i = \Sigma_1$, $\sum_{\varrho_i = \varrho} n_i = \Sigma_2$; ввиду того, что $(\varrho E - P_{jj})^{-1} = \frac{1}{\varrho} \sum_{n=0}^{\infty} \left(\frac{P_{jj}}{\varrho} \right)^n$ можем писать, положив сначала $\frac{P_{jj}}{\varrho} = Q_{jj}$, $\frac{\varrho^{-(k-1)}}{(k-1)!} p A_{11} P_{12} \dots A_{22} \dots P_{l-1, l} A_{ll} = \frac{\varrho^{-(l-1)}}{(k-1)!} \sum_{\Sigma_1 = l-k}^{\infty} p \dots Q_{jj}^{n_j-1} P_{jj+1} \dots B_{mm} P_{mm+1} \dots$, запись

слагаемого в этом соотношении означает, что при $\varrho_j < \varrho$ умножаем на матрицу $Q_{jj}^{n_j-1}$; при $\varrho_m = \varrho$ умножаем на коэффициент в соответствующей формуле Перрона B_{mm} . Итак, нам надо доказать, что при достаточно больших n выражение

$$\left| \frac{\varrho^{-(l-1)}}{(k-1)!} \sum_{\Sigma_1 = l-k}^{\infty} p \dots Q_{jj}^{n_j-1} P_{jj+1} \dots B_{mm} P_{mm+1} \dots e_\omega - \frac{\varrho^{-(l-1)}}{n^{k-1}} \sum_{\Sigma_1 = l-k}^{n+1-k} \sum_{\Sigma_2 = n+1-\Sigma_1} p Q_{11}^{n_1-1} P_{12} Q_{22}^{n_2-1} \dots Q_{ll}^{n_l-1} e_\omega \right|,$$

как угодно мало. Первая сумма во втором члене предписывает суммирование по показателям n_i классов с характеристическими числами, меньшими ϱ , а вторая сумма — с характеристическими числами, равными ϱ .

Возьмем натуральные числа N, M . Символ Σ^M будет обозначать суммирование только таких слагаемых, для которых $n_m > M$, а символ $\Sigma^{\bar{M}}$ — суммирование тех слагаемых, для которых по крайней мере одно $n_m \leq M$.

Записывая все еще короче, имеем

$$\left| \frac{1}{(k-1)!} \sum_{\Sigma_1 = l-k}^{\infty} \dots Q_{jj}^{n_j-1} \dots B_{mm} \dots - \frac{1}{n^{k-1}} \sum_{\Sigma_1 = l-k}^{n+1-k} \sum_{\Sigma_2 = n+1-\Sigma_1} p Q_{11}^{n_1-1} \dots \right| \leq$$

$$\begin{aligned} &\leq \left| \frac{1}{(k-1)!} \sum_{\Sigma_1=l-k}^N \dots Q_{jj}^{n_1-1} \dots B_{mm} \dots - \frac{1}{n^{k-1}} \sum_{\Sigma_1=l-k}^N \sum_{\Sigma_2=n+1-\Sigma_1}^M p Q_{11}^{n_1-1} \right| + \\ &+ \left| \frac{1}{(k-1)!} \sum_{\Sigma_1=N+1}^{\infty} \dots Q_{jj}^{n_1-1} \dots B_{mm} \right| + \left| \frac{1}{n^{k-1}} \sum_{\Sigma_1=N+1}^{n+1-k} \sum_{\Sigma_2=n+1-\Sigma_1} p Q_{11}^{n_1-1} \right| + \\ &+ \left| \frac{1}{n^{k-1}} \sum_{\Sigma_1=l-k}^N \sum_{\Sigma_2=n+1-\Sigma_1}^M p Q_{11}^{n_1-1} \right| = I_1 + I_2 + I_3 + I_4. \end{aligned}$$

Прежде всего покажем, что выбрав надлежащим образом N , можем добиться того, чтобы выражение I_3 было меньше $\varepsilon > 0$ для всех n . Обозначим $\Theta = \max_{\varrho_k < \varrho} \varrho_k$. Ввиду того, что $Q_{jj}^n = \left(\frac{\varrho_j}{\varrho}\right)^n B_{jj} + o\left[\left(\frac{\varrho_j}{\varrho}\right)^n\right]$, легко обнаружим, что существует постоянная C_1 такая, что $p Q_{11}^{n_1-1} \dots Q_{ii}^{n_i-1} e_{\omega} \leq C_1 \left(\frac{\Theta}{\varrho}\right)^{\Sigma_1}$. Теперь установим число слагаемых при суммировании $\sum_{\Sigma_1=\text{konst.}} \sum_{\Sigma_2=n+1-\Sigma_1}$. Легко видно, что оно равно $\binom{\Sigma_1-1}{l-k-1} \binom{n-\Sigma_1}{k-1} \leq C_2 (\Sigma_1)^{l-k-1} n^{k-1}$. Следовательно, $I_3 \leq C_1 C_2 \sum_{\Sigma_1=N+1}^{\infty} (\Sigma_1)^{l-k-1} \left(\frac{\Theta}{\varrho}\right)^{\Sigma_1}$ и выражение в правой части можно, выбрав надлежащим образом N , сделать как угодно малым (меньше ε). То же можно сказать и о выражении I_2 , потому что рассматриваемый ряд сходится.

Теперь аналогичным приемом оценим выражение I_4 . Имеем $p Q_{11}^{n_1-1} \dots Q_{ii}^{n_i-1} e_{\omega} \leq C_1$ и число слагаемых в сумме $\sum_{\Sigma_2=n+1-\Sigma_1}^M$ меньше $k \sum_{i=1}^M \binom{n-\Sigma_1-i}{k-2} \leq C_3 n^{k-2}$. Итак, мы видим, что для достаточно больших n также $I_4 < \varepsilon$.

О выражении I_1 будем рассуждать следующим образом: если T_j имеет характеристическое число ϱ , то $\lim_{n \rightarrow \infty} Q_{jj}^n = B_{jj}$. Значит, выбрав подходящим способом число M , можно добиться того, чтобы при фиксированных показателях, соответствующих классам с немаксимальными характеристическими числами, каждое из слагаемых суммы $\sum_{\Sigma_2=n+1-\Sigma_1}^M$ отличалось от соответствующего ему $\dots Q_{jj}^{n_j-1} \dots B_{mm} \dots e_{\omega}$ на как угодно малое число. Следовательно, достаточно, обозначив через $S_n(M, \Sigma_1)$ число слагаемых суммы $\sum_{\Sigma_2=n+1-\Sigma_1}^M$, убедиться в том, что $\lim_{n \rightarrow \infty} \frac{S_n(M, \Sigma_1)}{n^{k-1}} = \frac{1}{(k-1)!}$. Однако $S_n(M, \Sigma_1) = \binom{n-\Sigma_1-kM}{k-1}$. Этим теорема полностью доказана.

Перейдем теперь к произвольному классу T состояний цепи. Разобьем T на подклассы T_j состояний, сообщающихся внутри T ; это значит следующее: Если $\omega_1 \in T_j$, то система может перейти с положительной вероятностью из состояния ω_1 в состояние ω_2 и обратно, не выходя при этом за пределы класса T , в том и только в том случае, когда также $\omega_2 \in T_j$. Эти подклассы можно как известно, частично упорядочить следующим образом: $T_k < T_j$ означает, что внутри класса T возможен переход из класса T_k в класс T_j . Разбиение на подклассы T_j не должно обязательно исчерпать весь класс T . Ради простоты обозначений будем однако предполагать, что каждое состояние является сообщаемым с самим собой. Из хода доказательств станет ясным, что следовало бы изменить в противном случае. Также будем предполагать, что положительна вероятность того, что система будет во времени 0 пребывать в таких классах T_k , для которых не существует $T_j < T_k$, $T_j \neq T_k$. (Условие А.)

Теорема 3. Пусть классы T_j регулярны. Пусть $\rho = \max_{T_j} \rho_j$, пусть k — наибольшее число, обладающее тем свойством, что существуют классы $T_{\alpha_1} < T_{\alpha_2} < \dots < T_{\alpha_k}$, для которых $\rho_{\alpha_1} = \rho_{\alpha_2} = \dots = \rho_{\alpha_k} = \rho$. Тогда, если еще выполняется условие А, существует

$$\lim_{n \rightarrow \infty} \frac{P(T^n)}{n^{k-1} \rho^n} > 0.$$

Доказательство. Пусть P_{jj} — матрицы, образованные из вероятностей перехода внутри T_j , пусть P_{jk} — матрицы, образованные из вероятностей перехода из состояний класса T_j в состояния класса T_k . Через p_j обозначим ту часть вектора начального распределения, которая соответствует классу T_j . Если случайные явления $\{x_n = \omega\} \cap T^n$ разложить на непересекающиеся явления, руководствуясь тем, через какие классы прошла система, пока она не находилась в состоянии ω , получим

$$P(T^n) = \sum_{T_{\alpha} < T_{\beta} < \dots < T_{\omega}} \left(\sum_{\Sigma n_v = n+1} p_{\alpha} P_{\alpha\alpha}^{n_{\alpha}-1} P_{\alpha\beta} P_{\beta\beta}^{n_{\beta}-1} \dots P_{\omega\omega}^{n_{\omega}-1} e \right).$$

Применяя результат теоремы 2, мы видим, что порядок выражений в скобках не меньше $n^{k-1} \rho^n$ и что там существует по крайней мере один член такой, что частное от деления его на $n^{k-1} \rho^n$ имеет положительный предел.

В следующей теореме символ A_{jj} означает те же матрицы, как в теореме 2.

Теорема 4. При условиях теоремы 3 справедливо утверждение: Для каждого $\omega \in T$ существует предел $\lim_{n \rightarrow \infty} P(x_n = \omega | T^n) = P_T(\omega)$; он будет положительным тогда и только тогда, когда ω будет принадлежать такому классу T_j , для которого существуют $T_{s_1} < T_{s_2} < \dots < T_{s_k} \leq T_j$, характеристические числа которых равны ρ . В таком случае

$$P_T(\omega) = C \sum^* p_{r_1} A_{r_1 r_1} P_{r_1 r_2} A_{r_2 r_2} \dots A_{j j} e_{\omega},$$

где C — постоянная, не зависящая от j , и $\sum^*$ означает суммирование по всем последовательностям классов $T_{\tau_1} < T_{\tau_2} < \dots < T_j$, среди характеристических чисел которых число ϱ встретится k раз.

Доказательство. Имеем

$$P(x_n = \omega | T^n) = \frac{P(\{x_n = \omega\} \cap T^n)}{P(T^n)} = \\ = P(T^n)^{-1} \sum_{T_\alpha < T_\beta \dots < T_j, \sum n_\nu = n+1} \left(\sum p_\alpha P_{\alpha\alpha}^{n_\alpha-1} P_{\alpha\beta} \dots P_{jj}^{n_j-1} e_\omega \right).$$

Разделим числитель и знаменатель на $n^{k-1}\varrho^n$. Из теоремы 2 вытекает, что в случае, когда T_j не обладает перечисленными в теореме 4 свойствами, все выражения в скобке являются бесконечно малыми величинами порядка, большего $n^{k-1}\varrho^n$. Аналогично, если класс T_j удовлетворяет условиям теоремы, то по теореме 2 видно, что порядок $n^{k-1}\varrho^n$ имеют именно те выражения, которые суммируются по символу $\sum^*$. В этой же теореме приведено и предельное выражение. Если теперь воспользоваться теоремой 3, то получим, что

$$C = \frac{\varrho^{-(k-1)}}{(k-1)!} \left(\lim_{n \rightarrow \infty} \frac{P(T^n)}{n^{k-1}\varrho^n} \right)^{-1}.$$

Иначе эта постоянная определена условием $\sum_{\omega \in T} P_T(\omega) = 1$.

Предел вероятности можем привести еще в несколько ином виде. Пусть T_{n_k} — класс с характеристическим числом ϱ , для которого существуют $T_{n_1} < T_{n_2} < \dots < T_{n_{k-1}} < T_{n_k}$, характеристические числа которых также равны ϱ . По предположению класс T_{n_k} регулярен. Из доказательства теоремы 1 вытекает, что для матрицы $A_{n_k n_k}$, которая является коэффициентом при ϱ^n в формуле Перрона, справедливо соотношение (1), т. е.: $p A_{n_k n_k} = C_p \pi_{n_k}$, где π_{n_k} — характеристический вектор, соответствующий корню ϱ матрицы $P_{n_k n_k}$. Если в качестве вектора p брать последовательно векторы, одна компонента которых равна единице а остальные нулю, то окажется, что строки матрицы $A_{n_k n_k}$ суть кратные π_{n_k} , т. е. $A_{n_k n_k} = u_{n_k} \pi_{n_k}$, где u_{n_k} — вектор-столбец с одними положительными компонентами. Итак, каждому классу можно поставить в соответствие число

$$C_{n_k} = C \sum p_{s_1} A_{s_1 s_1} P_{s_1 s_1} A_{s_2 s_2} \dots A_{s_j s_j} P_{s_j n_k} u_{n_k},$$

причем на этот раз требуется, чтобы слагаемые содержали матрицы $k-1$ классов с характеристическим числом ϱ .

Предельные вероятности в классе T_j можно тогда писать в виде

$$P_T(\omega) = \sum C_{s_1 s_1} P_{s_1 s_1} A_{s_2 s_2} \dots A_{j j} e_\omega, \quad (2)$$

где суммируются произведения для всех классов T_{s_1} , удовлетворяющих только что высказанным условиям, и для классов следующих за ними.

О регулярности условных вероятностей справедлива при условиях теоремы 3.

Теорема 5. Если существует одна единственная k -членная последовательность классов $T_{s_1} < T_{s_2} < \dots < T_{s_k}$ с характеристическими числами ϱ , то $P_T(\omega)$ одна и та же для всех начальных распределений, удовлетворяющих условию А.

Доказательство легко вытекает из соотношения (2). Постоянная опять-таки определена условием $\sum_{\omega \in T} P_T(\omega) = 1$.

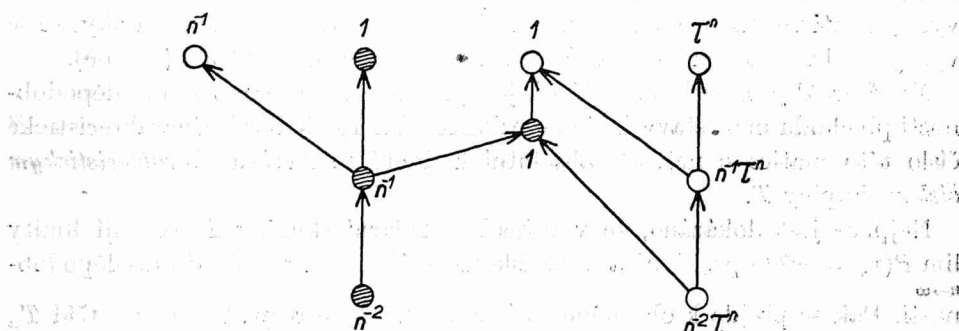


Рис. 1.

Замечание. Если $\varrho = 1$, то и $k = 1$, ибо, как легко обнаружить, в регулярному классу T_j , характеристическое число которого равно единице, не существует T_k такое, чтобы было $T_j < T_k$. Предельное распределение в таких классах T_j можно представить в виде $C_j \pi_j$. В остальных классах предел равен 0.

Добавление. Предположим, что характеристическое число подклассов равно ϱ или Θ , $\varrho > \Theta$, и что эти подклассы частично упорядочены таким образом, как показано на рисунке 1; классы с характеристическим числом ϱ обозначены полными кружками. Если положить $\frac{\Theta}{\varrho} = \tau$ то из теорем 2 и 3 вытекает, что величины $P(x_n = \omega | T^n)$ имеют в отдельных классах такой порядок, какой написан у каждого кружка на рисунке.

ЛИТЕРАТУРА

- [1] Ф. Г. Гантмахер, М. Г. Крейн: Осцилляционные матрицы и ядра и малые колебания механических систем, Москва-Ленинград 1950.
- [2] Т. А. Сарымсаков: Основы теории процессов Маркова. Москва 1954.
- [3] L. Truksa: Statistická dynamika, Praha 1956.