# Предельные вероятности состояний

Для эргодических однородных марковских цепей существует стационарный режим при t →∞. При стационарном режиме вероятности состояний стремятся к некоторым установившимся значениям – предельным вероятностям, которые постоянны и не зависят от начального состояния системы.

## Теорема

Если **число** состояний системы **n** конечно и из каждого состояния можно перейти в любое другое за конечное время, то существуют предельные вероятности состояний Pi = lim(Pi(t)) при t→∞. ,i=1..n;

Сумма предельных вероятностей Pi дает 1.

Предельные вероятности определяют среднее относительное время пребывания системы в данном состоянии.

Для того, чтобы вычислить предельные вероятности состояний достаточно в системе уравнений Колмогорова приравнять все левые части (производные) к нулю, поскольку в установившемся режиме все вероятности состояний постоянны. В результате получим системы однородных линейных уравнений.

-∑ (от j=1,j != n до n ) Lij \* Pi + ∑ (от j=1,j!= i до n) Lji \* Pj =0, i=1..n

Пример. Система имеет 4 состояния, известны интенсивности переходов. Нужно построить систему уравнений Колмогорова .

{

{P`1(t) = -2\*P1(t) – 3\*P1(t)+1\*P3(t)

{P`2(t) = -1\*P2(t) + 2\*P1(t) + 2\*P3(t)

{P`3(t) = -1\*P3(t) – 2\*P3(t) + 3\*P1(t) + 2\*P4(t)

{P`4(t) = -2\*P4(t) + 1\*P2(t)

{∑ (от i=1 до n=4) Pi(t)=1

{

{-5P1+P3=0

{-P2+2P1+2P3=0

{-3P3+3P1+2P4=0

{-2P4+P2=0

{P1+P2+P3+P4=1

{

Ответ: P1 =1/24, P2=1/2, P3=5/24, P4= ¼

# Процессы «гибели и размножения»

*Опр.*

Непрерывная марковская цепь называется процессом гибели и размножения, если ее граф состояний можно представить в виде одной цепочки, в которой каждое из средних состояний S1, S2, … , Sn-1 связаны прямой и обратной связью с каждым из средних состояний, а крайнее состояние S0 и Sn только с одним соседним состоянием. В общем случае граф состояния процесса гибели и размножения имеет следующий вид:

**Фото в телефоне**

Пусть для всех пар состояний известны интенсивности переходов. Составим систему равнений Колмогорова.

P`0(t) = -L01\*P0(t) +L10\*P1(t)

P`1(t) = -L12\*P1(t) + L21\*P2(t) – L10\*P1(t)+L0\*P0(t)=L01\*P0(t)-(L10+L12)\*P1(t) + L21\*P2(t)

P`2(t) = -L21\*P2(t) –L23\*P2(t) + L12\*P1(t) + L32\*P3(t) = L12\*P1(t) – (L21+L23)\*P2(t) + L32\*P3(t)

P`k(t) = -Lk,k-1\*Pk(t) – Lk,k+1\*Pk(t)+Lk-1,k\*Pk-1(t)+Lk+1,k\*Pk+1(t)

P`n(t) = -Ln,n-1\*Pn(t)+Ln-1,n\*Pn(t)

В стационарном режиме при t-> беск. имеем: Pi(t)->Pji ; P`i->0

-L01\*P0 + L10 \* P1=0

L01\*P0-(L10+L12)\*P1+L21\*P2=0

L12\*P1 – (L21+L23)\*P2+L32\*P3 = 0

…

Lk-1,k \* Pk-1 – (Lk,k-1 + Lk,k+1)\*Pk + Lk+1,k\*Pk+1=0

…

Ln-1,n\*Pn-1 – Ln,n-1\*Pn=0

{1)**L01\*P0=L10\*P1** (\*)

{2)**L01\*P0** + L21\*P2=**L10\*P1**+L12\*P1

2) L21\*P2=L12\*P1

{n)Ln,n-1\*Pn=Ln-1,n\*Pn-1

{∑(i=1.. n) Pi=1

L01\*P0=L10\*P1

P1=(L01/L10)\*P0

P2=(L12/L21)\*P1

Pk=(Lk-1,k / Lk,k-1)\*Pk-1 = (Lk-1,k / Lk,k-1) \* (Lk-2,k-1 / Lk-1,k-2)\*…\*(L01 / L10). В числителе все L12, L23 .. то есть слева на право, а в знаменателе L21, L10, то если с права на лево.

P0 + P1 + P2 + … + Pn=1

P0 + (L01/L10)\*P0 + (L12 \* L01 / L21 \* L10)\*P0 + … (L01 \* ... \* Ln-1,n / L10 \* … \* Ln,n-1)\* P0 =1

P0 \* ((L01/L10) + (L12 \* L01 / L21 \* L10) + … (L01 \* ... \* Ln-1,n / L10 \* … \* Ln,n-1)) = 1

P0=1 / ((L01/L10) + (L12 \* L01 / L21 \* L10) + … (L01 \* ... \* Ln-1,n / L10 \* … \* Ln,n-1))

Пример.

Найти предельные вероятности состояний для графа следующего процесса гибели и размножения:

∑

**Фото в телефоне**

P0 = 1 + L01/L10 + L01 \* L12 / L10 \* L21 + L01 \* L12 \* L23 / L10 \* L21 \* L32 = 2/3 + 2\*1/3\*2 + 2\*1\*3 / 3\*2\*2 +1= (8+4+6+12)/12=2.5

P0 = 2/5

P1= 2/3 \* 2/5 = 4/15

P2= 2/6 \* 2/5 = 2/15

P3 = 6/12 \* 2/5= 1/5

# Простейший поток заявок

Входящий поток заявок во многом определяет характеристики производительности систем массового обслуживания. Поэтому правильное описание потока заявок, поступающих в реальную систему в случайные моменты времени, является важной задачей.

**Опр.**

Случайным потоком называется последовательность случайных моментов наступления некоторых событий. (Например, вызов на электронную станцию)

Случайный поток событий по существу является случайным процессом.

Заявки, поступающие в СМО или покидающие ее, образуют случайный поток, который можно изобразить как последовательность точек на оси времени, соответствующих случайным моментам времени получения заявок.

(23.04)

Среди свойств, которыми могут обладать случайные потоки событий выделяют свойства **стационарности**, **отсутствия последействия** и **ординарности.**

**Опр.**

**Поток событий называется стационарным**, если вероятность наступления какого-либо числа событий на интервале событий Δt зависит только от длины этого интервала и не зависит от того, где именно на оси времени взят этот интервал.

**Из этого следует:**

Условию стационарности удовлетворяет поток заявок, вероятностные характеристики которого не зависят от времени.

**Поток событий называется потоком без последействия,** если для любой пары непересекающихся интервалов времени число событий, попадающих на один из них не зависит от числа событий, попадающих на другой.

Условие отсутствия последействия означает, что заявки попадают в систему массового обслуживания в те или иные моменты времени, независимо друг от друга.

В общем случае выходной поток обслуженных заявок обычно имеет последействие, даже если входной поток его не имеет. (Пример: поток пассажиров входящих/выходящих из общественного транспорта не зависит друг от друга)

**Поток событий называется ординарным**, если пренебрежимо мала вероятность того, что на малый интервал времени Δt попадет больше одного события.

Условие ординарности означает, что заявки поступают в систему массового обслуживания по одиночке, а не группами. Однако, если в неординарном потоке заявки поступают только парами/тройками/и тд, то неординарный поток сводится к ординарному путем рассмотрения пар/троек/и тд.

**Опр.**

**Поток событий называется простейшим** или **стационарным Пуассоновским**, если он одновременно стационарен, ординарен и не имеет последействия.

**Опр.**

**Интенсивностью простейшего потока** λ называют среднее число заявок, поступающих в систему в единицу времени.

Теорема 1:

Число событий, произошедших в простейшем потоке, есть случайная величина, распределенная по закону Пуассона. Вероятность того, что в простейшем потоке с интенсивностью λ за интервал времени t поступит ровно k заявок, равна:

Pk(t)=

В частности вероятность того, что за интервал времени t не поступит не одной заявки: P0(t)=e-λt

Пример: На телефонную станцию поступает простейший поток вызова с интенсивностью λ=1.2 вызовов в минуту. Найти вероятность того, что за 2 минуты:

1) не придет не одного вызова

2) придет ровно один вызов

3) придет хотя бы 1 вызов

Дано:

λ=1.2

t=2 мин

Найти:

P0(2)-?

P1(2)-?

\_\_

P0(2)-?

1) P0(2)=0.9

2)P1(2)=((1.2\*2)1/1!)\*e-1.2\*2=0.216

\_\_

P0(2) = 1-0.9 = 0.91

Теорема 2:

Интервал времени l между событиями в простейшем потоке есть случайная величина, распределенная по показательному закону.

Вероятность того, что интервал времени l между поступлениями заявок меньше некоторой величины l, равна:

P(l<t) = F(t) = 1-e-λt (t>0)

Если входящий поток заявок и все потоки обслуживания простейшие, то процесс, протекающий в системе массового обслуживания является Марковским случайным процессом.

# Время обслуживания. Время ожидания.

Одной из важнейших характеристик систем массового обслуживания, которое определяет пропускную способность всей системы является **время обслуживания,** то есть время пребывания одной заявки в канале обслуживания.

Время обслуживания tобсл является случайной величиной и зависит как от стабильности самих каналов, так и от параметров, поступающих в систему заявок. В общем случае время обслуживания может изменяться в большом диапазоне.

Случайная величина tобсл полностью характеризуется законом о распределении вероятностей, который устанавливает соответствия между возможными значениями случайной величины и вероятностями их появления.

Поскольку tобсл – непрерывная случайная величина, то вероятность того, что она примет какое-то конкретное значение в определенный момент времени равна 0. Однако можно вычислить вероятность того, что непрерывная случайная величина примет значение, принадлежащее заданному интервалу (a,b) и задать закон распределения вероятностей с помощью функции плотности распределения вероятности.

Эта функция характеризует плотность, с которой распределяются значения случайной величины в данной точке.

Время обслуживания имеет, как правило, показательный закон распределения с плотностью распределения вероятностей g(t)=μe-μt

μ=1/ tобсл - **интенсивность обслуживания.**

Показательный закон распределения времени обслуживания имеет место тогда, когда плотность распределения вероятностей резко убывает с возрастанием времени.

**Время ожидания –** время пребывания заявки в очереди, если очередь существует.

tож - непрерывная случайная величина, имеющая, как правило, показательное распределение с плотностью распределения вероятностей h(t).

h(t)=νe-ve ,где ν=1/ tож – среднее время ожидания обслуживания.

# СМО с отказами.

## Уравнения Эрланга.

В СМО с отказами заявка, поступившая в момент, когда все каналы заняты, немедленно получает отказ, покидает систему и в дальнейшем процессе обслуживания не участвует.

Рассмотрим СМО с отказами, удовлетворяющему следующим условиям:

1) Система имеет n каналов обслуживания.

2) На вход СМО поступает простейший поток заявок с интенсивностью λ.

3) Время обслуживания одной заявки – случайная величина, имеющая показательное распределение. Обслуживание заявок **каждым каналом** СМО осуществляется с интенсивностью , где tобсл – среднее время обслуживания.

4) при занятости всех n каналов вновь пришедшая заявка получает отказ и покидает СМО.

Так как поток заявок и поток освобождения заявок – простейшие потоки, то процесс, протекающий в системе будет Марковским.

Пример: **фото в телефоне**

Из рисунка видно, что процесс, протекающий в системе представляет собой частный случай процесса гибели и размножения. Пользуясь общими правилами, можно составить уравнение Колмогорова для вероятностей Pk(t).

{

P`0(t)=-λ P0(t)+μ P1(t)

P`1(t)=-μ P1(t)-λ P1(t)+λ P0(t)+2μ P2(t)= λP0(t)- (μ+λ)\*P1(t)+2μ P2(t)

Pk(t)=-kμ\* Pk(t)-λ Pk(t)+(k+1)μ\* Pk+1(t)=λ Pk-1(t)-(kμ+λ) Pk(t)+μ(k+1) Pk+1(t)

…

Pn(t)= -nμ Pn(t)+λ Pn-1(t)=λ Pn-1(t)-nμ Pn(t)

} (\*)

Интегрировать систему (\*) надо при начальных условиях, описывающих состояние S0

P0(0)=1

P1(0)=0

P2(0)=0

Для любого момента времени должно выполняться условие

∑ Pk(t)=1, k=0..n

Вероятности Pk(t) – **характеризуют среднюю нагрузку системы и ее изменение с течением времени**. В частности Pn(t)=Pотк

Величина q(t)=1- Pn(t) называется **относительной пропускной способностью системы**. Для данного момента времени t это есть отношение среднего числа обслуженных на единицу времени заявок к среднему числу поданных.

# Стационарный режим обслуживания СМО с отказами. Формулы Эрланга.

В начале работы в СМО, как и в любой другой динамической системе, возникает так называемый «переходный» нестационарный процесс. Однако спустя некоторое время нестационарный процесс затухнет и система перейдет на стационарный режим, вероятностные характеристики которого уже не будут зависеть от времени.

Доказано, что для любой СМО с отказами существует стационарный режим обслуживания, т.е. при t->беск. Pk(t)->Pk и P`k(t)->0, k=0..n

**Опр.**

Вероятности P0, P1, … Pn состояний СМО в стационарной режиме функционирования называют **предельными вероятностями**.

{

-λp0+μp1=0

λpk-1-( λ+kμ)pk+(k+1)μpk+1=0

λpn-1-nμpn=0

∑pi=1

i=0..n

{

Решая данную систему относительно неизвестных P0, P1, … Pn получаем выражение нахождение предельных вероятностей системы (ф. Эрланга).

P0=[∑ ]-1

pk=pk\*p0/k!

Формулы Эрланга дают предельный закон распределения числа занятых каналов в зависимости от характеристик потока заявок и производительности системы обслуживания.

Ф. Эрланга остаются справедливыми при любом законе распределения времени обслуживания, лишь бы входной поток был простейшим.

|  |  |
| --- | --- |
| Предельные характеристики | Формулы |
| Вероятность отказа, т.е. вероятность того, что поступившая заявка не будет обслужена | Pотк=pn |
| Вероятность обслуживания поступившей заявки | Pобсл=-pn |
| Абсолютная пропускная способность СМО | Q=λ\*(1-pn) |
| Относительная пропускная способность СМО | q=Q/ λ=1-Pотк=Pобсл |
| Среднее число занятых каналов | =p\*(1-pn) |

**Задачи:**

N канальная СМО представляет собой вычислительный центр с 3 взаимозаменяемыми компьютерами для решения поступающих задач. Задачи поступают на ВЦ с интенсивностью λ =1 задача в час. Средняя продолжительность обслуживания t=1.8 часа. Требуется вычислить значения:

1) вероятности числа занятых каналов ВЦ

2) вероятность отказа в обслуживании заявки

3) относительно пропускной способности ВЦ

4) абсолютной пропускной способности ВЦ

5) среднее число занятых каналов ВЦ

Дано:

n=3

λ=1 з/ч

tобсл=1.8 ч

Найти:

1) P1, P2, P3

2) Pотк

3) Q, q

4)

{

-λp0+μp1=0

λp0-μp1- λp1+ 2μp2=0

λp1- 2μp2- λp2+ 3μp3=0

λp2-3μp3=0

{

∑pi=1

i=0..n

q=λ/μ=1/1/tобсл=tобсл=1.8

p0=[]-1=(1+1.8+1.02+0.972)-1=0.185

p1=1.81\*0.185=0.334

p2=1.82\*0.185/6=0.3

p3=1.83\*0.185/6=0.18

pn=pотк=p3=0.18

Q=λ\*Pобсл=1\*0.82=0.82

Pобсл=1-p3=1-0.18=0.82

q=Q/ λ=0.82

=p(1-p3)=0.82\*0.82=1.476

**Пример:**

Телефонная станция имеет n=4 линии связи. На станцию поступает простейший поток заявок с интенсивностью λ=3 вызова в минуту. Вызов, поступивший в момент, когда все линии заняты, получает отказ. Средняя длительность разговора tобсл=2 мин.

Найти: **1)** среднюю долю времени, в течение которой телефонная станция вообще не загружена. **2)** вероятность отказа

Дано:

n=4

λ=3 в/м

tобсл=2 мин

Найти:

p0, pотк

p= λ/μ=3\*2=6

p0=[]-1=(1+6+13+216/6 +1296/24)-1=0,009

pотк=p4=p4/4! \* p0=1296/24 \* 0.009=0.467

# СМО с ожиданием

**Опр.**

**СМО** называется **системой с ожиданием**, если заявка, заставшая все каналы занятыми, становится в очередь и ждет, пока не освободится какой-нибудь канал.

**Опр.**

Если ожидание заявки в очереди ничем не ограничено, то система называется **чистой системой с ожиданием**. Если она ограничена какими-то условиями, то система называется системой **смешенного типа**.

Ограничения, наложенные на ожидание, могут быть различного типа. **Часто ограничение накладывается на время ожидания заявки в очереди**, при этом ограничивается только срок ожидания в очереди, а начатое обслуживание доводится до конца не зависимо от того, сколько времени продолжалось ожидание. В других задачах **естественнее наложить ограничение на общее время пребывания заявки в системе**.

Ограничение может заключаться в том, что **заявка становится в очередь только в том случае, если длина очереди не слишком велика**. Здесь ограничение накладывается на длину очереди.

В системе с ожиданием существенную роль играет дисциплина очереди. Ожидающие заявки могут вызываться на обслуживание как в порядке очереди, так и в случайном порядке.

Существуют системы с «преимуществами», в которых некоторые заявки обслуживаются предпочтительнее других.

Рассмотри СМО смешенного типа с ограниченным временем ожидания при следующих допущениях:

1) Система имеет n канало обслуживания.

2) На вход СМО поступает простейший поток заявок с интенсивностью λ.

3) Время обслуживания одной заявки представляет собой случайную величину, имеющую показательное распределение. Обслуживание заявок **каждым каналом** СМО осуществляется с интенсивностью μ=1/tобсл, где tобсл –среднее время обслуживания

4) Заявка, заставшая все каналы занятыми, становится в очередь и ожидает обслуживания. Число мест в очереди не ограничено, но время ожидания ограничено некоторым сроком. Если до истечения этого срока заявка не будет принята к обслуживанию, то она покидает очередь и становится необслуженной.

5) **Срок ожидания** заявки в очереди представляет собой случайную величину, имеющую показательное распределение. На каждую заявку, состоящую в очереди, действует «поток уходов из очереди» с интенсивностью v=1/tож, где tож – среднее время ожидания.

6) Все потоки событий, приводящие к изменениям состояний СМО, носят пуассоновский характер. Тогда процесс, протекающий в системе, будет Марковским.

Будем нумеровать состояния системы не по числу занятых каналов, а **по числу связанных с системой заявок**.

**Заявку** будем называть **связанной с системой**, если она либо находится в состоянии обслуживания, либо ожидает очереди.

Состояния:

S0 – все каналы свободны, очереди нет

S1 – 1 канал занят, очереди нет

S2 – 2 канала заняты, очереди нет

…

Sn – все n каналов заняты, очереди нет

Sn+1 - n каналов занято, 1 заявка в очереди

…

Sn+L - n каналов занято, L заявок в очереди

…

и т.д.

Число L заявок, стоящих в очереди может быть сколь угодно большим. Таким образом система имеет бесконечное, хотя и счетное множество состояний.

21.05

Фото в телефоне

Фото уравнений (\*) в телефоне

Уравнения (\*) являются естественным обобщением уравнений Эрланга на случай системы смешанного типа с ограниченным временем ожидания. Параметры μ, λ, ν могут быть как постоянными, так и переменными. При интегрировании системы (\*) нужно учитывать, что хотя теоретически число возможных состояний бесконечно, но на практике вероятности Pn+r ­при возрастании r становятся пренебрежимо малыми и соответствующие уравнения могут быть отброшены.

При t -> бескон. Pk(t)->Pk, а Pk`(t)->0 (условие стационарности)

{

μP1-λP0=0

λP0-(λ+μ)P1+2μP2=0

…

λPk-1-(λ+kμ)Pk+(k+1)μPk+1=0

…

λPn-1-(λ+nμ)Pn+nμPn+1=0

…

λPn+r-1-(λ+(n+r)μ)Pn+r+(n+r+1)μPn+r+1=0

…

∑i=0..беск. Pi=0

{

Решая полученную систему уравнений найдем значение предельных вероятностей.

P0=

где ,

Параметры p и B выражают соответственно среднее число заявок и среднее число ухода заявок, стоявших в очереди, приходящееся на среднее время обслуживания одной заявки. Очевидно, что при одинаковых значениях λ и μ пропускная способность СМО с ожиданием будет всегда выше, чем пропускная способность системы с отказами, так как в случае наличия ожидания необслуженными уходят не все заявки, заставшие каналы занятыми, а только некоторые. Пропускная способность увеличивается при увеличении среднего времени ожидания.

1) Вероятность отказа вычисляется по следующей формуле:

Pотк=

2) Вероятность того, что любая заявка будет обслужена

Pобсл=1-Pотк

3) Среднее число заявок, находящихся в очереди

4) Абсолютная пропускная способность СМО

5) Относительная пропускная способность СМО

6) Среднее число занятых каналов

Пусть B 🡪. Тогда система с ожиданием превратится в **систему с отказами** (заявки мгновенно уходят из очереди).

Пусть B 🡪0. Тогда получим **чистую систему с ожиданием.** В такой системе заявки вообще не уходят из очереди, и каждая заявка рано или поздно дождется обслуживания. Однако в чистой системе с ожиданием не всегда имеется стационарный режим при t🡪. Доказано, что такой режим существует при p<n.Т.е. когда среднее число заявок, приходящееся на время обслуживания одной заявки, не выходит за пределы возможностей n-канальной системы. Если же p>=n, то число заявок, стоящих в очереди будет с течением времени неограниченно возрастать.

1) Пусть p<n. Все заявки будут обслужены и очередь не будет возрастать до inf. Можно найти значения предельных вероятностей.

(1<=k<=n)

(l>=1)

Вероятностные характеристики:

1) вероятность отказа Pотк=0

2) вероятность того, что любая заявка будет обслужена Pобсл=1

3) абсолютная пропускная способность Q=λ

4) относительная пропускная способность q=1

5) среднее число заявок в очереди

6) среднее число занятых каналов Nкан =Q/μ

7) среднее число заявок в СМО Nсист=Nоч+p

8) среднее время пребывания заявки в СМО tсист=Nсист/λ

9) среднее время пребывания заявки в очереди tоч= Nоч/λ

n=3. Интенсивность λ=4 з/ч. tср=30 мин. Определить, существует ли стационарный режим обслуживания, если да, то найти вероятности P0, P1, P2, P3, вер-ть наличия очереди, длину очереди.

Найти: p, P0, P1, P2, P3, Pоч, Nоч

p=λ/μ

μ=1/t

p=2

2<3 . Стационарный режим сущ-т.

P0=(1+2+2+8/6 + 16/6)-1=1/9

P1=2/9

P2=2/9

P3=8/6 \* 1/9 = 4/27

Pоч=1-(P0+P1+P2+P3)=8/27

Nоч=8/9