1. ***Основные понятия и определения ТМО***

**Теория массового обслуживания** – раздел теории вероятностей, изучающих математические модели разного рода реальных систем массового обслуживания. Эти модели представляют собой случайные процессы специального вида.

**Заявкой называется** спрос на удовлетворение какой-либо потребности.

**Обслуживание заявки** – выполнение заявки СМО называется любая система для выполнения заявок, поступающих в нее в случайные моменты времени.

**Примерами СМО могут служить**:

* Комп. сети;
* Системы связи (телефонные станции);
* Магазины, больницы и т.д.

**Каналом обслуживания называется** устройство в СМО, обслуживающее заявку.

**СМО считается заданной,** если описаны следующие параметры:

* Входной поток заявок;
* Время обслуживания заявок;
* Структура системы;
* Емкости накопителей;
* Дисциплина обслуживания.

**Входящим потоком заявок называют** последовательность событий, заключающихся в поступлении заявок в СМО.

**Классификация СМО:**

* По ограничению потока заявок (на замкнутые и открытые. В открытой СМО характеристики потока заявок не зависят от того, сколько каналов занято, а в замкнутой СМО зависят);
* По числу каналов СМО (на одноканальные и многоканальные);
* По типу обслуживаемых каналов (на однофазные и многофазные)
* По емкости накопителей (на СМО с отказами и с ожиданием. СМО с ожиданием делится на с неограниченным ожиданием и с ограниченным).

**Дисциплина обслуживания** – определяет порядок обработки заявки в системе, начиная с момента ее поступления в систему и до момента, когда она ее покидает.

1. ***Вероятностный аппарат ТМО. Случайные процессы***

**Случайный процесс X(t)** – это особого вида функция времени, характеризующаяся тем, что в любой момент времени принимаемые ею значения являются случайными величинами.

**Случайная величина X(t0)**, в которую обращается случайный процесс при фиксированном t=t0, **называется** **сечением случайного процесса**.

**Реализацией случайного процесса** называется неслучайная функция х(t), в которую превращается случайный процесс X(t) в результате испытания, т.е. конкретный вид, принимаемый случайным процессом X(t).

1. ***Вероятностные характеристики случайного процесса. Стационарные случайные процессы***

Математическим ожиданием случайного процесса называется неслучайная функция, m(t), которая при каждом значении аргумента t равна мат ожиданию соответствующего сечения случайной функции.

Геометрически мат ожидание случайного процесса можно истолковать как среднюю кривую около которой расположены другие кривые – реализации. При фиксированном значении аргумента Мат ожидание – есть среднее значение сечения.

Опр.

Дисперсией случайного процесса называют неслучайную неотрицательную функцию D(t), значение которой для каждого аргумента t равно дисперсии соответствующего сечения случайной функции.

Дисперсия характеризует степень рассеивания возможных реализаций вокруг мат ожидания случайного процесса.

**Корреляционной функцией** случайного процесса называется неслучайная функция Kx(t1\*t2), которая для каждой пары моментов времени t1, t2 равна мат ожиданию произведения соответствующих сечений X(t1) и X(t2): Kx(t1,t2)- Dx

Свойства корр функции:

1. При равных аргументах t1=t2 корреляционная функция обращается в дисперсию случайной функции. K(t1,t1)=D(t1)
2. Корреляционная функция симметрична относительно своих аргументов. K(t1,t2)=K(t2,t1)

Корреляционная функция зависит от дисперсии. Если процесс весьма мало отклоняется от своего мат ожидания, то корреляционная функция будет мала, независимо от степени связи между функциями.

**Положительная корреляция** между случайными означает, что при возрастании одной из них, другая имеет тенденции в среднем возрастать.

**Отрицательная корреляция** означает, что при возрастании одной из случайных величин, другая имеет тенденцию в среднем убывать.

Случайные процессы называют стационарными**,** если все его вероятностные характеристики не меняются при любом изменении аргумента, т.е.:

1. M(t)=m=const
2. D(t)=D=const
3. K(t,t+T)=k(T)

При Т=0 k(t,t)=D(t)=k(0)=const

Условие 3 есть единственное существенное условие, которому должная удовлетворять стационарная случайная функция.

1. ***Понятие Марковского случайного процесса. Цепь Маркова. Однородная цепь Маркова***

Случайный процесс, протекающий в некоторой системе называется Марковским или процессом без последействия, если для каждого момента времени поведение системы в будущем зависит только от состояния системы в данный момент времени и не зависит от того, каким образом система пришла в это состояние. Иными словами в марковских случайных процессах будущее состояние целиком определяется его настоящим состоянием, прошлое на него никак не влияет. Это свойство называется **СВОЙСТВОМ ОТСУТСТВИЯ ПОСЛЕДЕЙСТВИЯ** или **МАРКОВСКИМ СВОЙСТВОМ.**

Пусть имеется некоторая система S, которая может находиться в состояниях S1, S2, … , Sn. Пусть переходы системы из одного сост. в другое происходят мгновенно и только в заранее фиксированные моменты времени шаги, которые можно пронумеровать. Таким образом в системе протекает случайный процесс с дискретными состояниями и дискретным временем.

Марковский случайный процесс с дискр. сост-ми и дискр. Временем называют **ДИСКРЕТНОЙ МАРКОВСКОЙ ЦЕПЬЮ** или **ЦЕПЬЮ МАРКОВА.** Обозначим через Ski событие, состоящее в том, что на k-том шаге система будет находиться в состоянии Si. (K=1,2.., I=1..n)

Вероятность события Ski, состоящего в том, что система на k-ом шаге будет находиться в сост. Si называют **вероятностью i-го состояния на k-ом шаге и обозначают Pi(k).** Вероятности состояния системы в совокупности образуют вектор P(k)=(P1(k),P2(k),..,Pn(k))

Вектор P(0)=P1(0),P2(0),..,Pn(0), Где Pi(0) –вероятность появления сост. Si в начальном испытании, называют **вектором начальных вероятностей**.

Очевидно, что на любом шаге процесс может находиться в одном и только одном из n возможных состояний S1,S2, ... ,Sn, следовательно при любом k событие Sk1,Sk2,..Skn будут единственно возможны и несовместны, то есть образуют полную группу. Поэтому для каждого шага k имеет место равенство **P1(k)+P2(k)+...+Pn(k)=1**.

Переходными вероятностями называют вероятности перехода системы из любого состояния в любое состояние за один шаг. Некоторые из переходных вероятностей могут быть равны нулю, если непосредственный переход за один шаг невозможен.

Марковскую цепь называют однородной, если переходные вероятности не зависят от номера шага. В противном случае марковскую цепь называют **неоднородной**. Будем считать марковскую цепь однородной. Обозначим через Pij вероятность перехода системы **за один шаг** из состояния Si в сост Sj.

Матрицей перехода за один шаг называют матрицу, которая содержит все переходные вероятности этой системы за один шаг.

1. ***Равенство Маркова***

Пусть событие А может произойти только при условии появления одного из событий H1, H2, …, Hn, образующих полную группу т.к. заранее неизвестно, какое именно из этих событий наступит, их называют гипотезами. Тогда вероятности наступления события А равно P(A)=P(H1) \* P(A) + P(H2) \* P(A) + ... + P(Hn) \* P(A).

Введем обозначения: А – за k шагов система перейдет из сост Si в сост Sj. Следовательно вероятность события P(A)=Pij(k). Hr – за m шагов система перейдет из сост Si в сост Sr. Следовательно P(Hr)=Pir(m).

1. ***Классификация состояний Марковской цепи. Достаточное условие эргодичности***

## Достаточное условие эргодичности.

При рассмотрении цепей Маркова важно изучить поведение системы на большом интервале времени, когда число переходов стремится к бесконечности.

Марковские цепи классифицируются в зависимости от возможности переходов из одних состояний в другие.

Состояние Sj называют достижимым из состояния Si, если вероятность перехода системы из i в j за конечное число шагов положительна.

**Состояния** Si и Sj, достижимые друг из друга, называют **сообщающимися**.

Цепь Маркова называют неприводимой, если все ее состояния сообщаются друг с другом.

Состояние Si называют возвратным, если вероятность того, что система, выйдя из этого состояния, вернется в него за конечное число шагов, равна 1. Если же эта вероятность меньше 1, то состояние называют **невозвратным**.

Возвратное состояние, для которого возвращение возможно лишь через число шагов, кратное d (d!=1), называют **периодическим**.

Эргодическим состоянием называют возвратное непериодическое состояние, для которого среднее время возвращение конечно.

Эргодическая Марковская цепь – Марковская цепь, все состояния которой являются эргодическими.

Процесс, порождаемый эргодической цепью, начавшись в некотором состоянии никогда не завершается, а последовательно переходит из одного состояния в другое, попадая в различные состояния с разной частотой, зависящей от переходных вероятностей.

Состояние Si называют поглощающим, если вероятность перехода из этого состояния в любое другое равна 0, т.е. Pij=1.

Эргодическая Марковская цепь обладает важными свойствами:

При неограниченным увеличением числа шагов, переходные вероятности стремятся к некоторым постоянным числам.

## Теорема:

Если существует такое число m>0, при котором все элементы матрицы Пm переходов за m шагов положительны, то существуют такие постоянные числа Pj (j=1,2..n),что

Lim pj(k)=pj

k→∞

1. ***Непрерывные цепи Маркова***

**Процесс**, протекающий в системе, **называется процессом с непрерывным временем**, если переход системы из одного состояния в другое возможен в любые, заранее неизвестные, случайные моменты времени.

Марковский процесс с дискретными состояниями и непрерывным временем называется **непрерывной марковской цепью.**

1. ***Уравнения Колмогорова***

{

{

1. ***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.

1. ***Процессы «гибели и размножения»***

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:

1. ***Простейший поток заявок***

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.

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

**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.

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

**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)=

**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.

1. ***Время обслуживания. Время ожидания***

**Время обслуживания -** 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.

1. ***Система массового обслуживания с отказами. Уравнения Эрланга. Стационарный режим обслуживания СМО с отказами. Формулы Эрланга***

## 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.

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.

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.

1. ***Система массового обслуживания с ожиданием. Система массового обслуживания с ограниченным временем ожидания***

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

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.

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 →∞. 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..∞. 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