**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ГРАЖДАНСКОЙ АВИАЦИИ**

Кафедра Высшей математики

Ю. И. Дементьев, И. В. Платонова

МОДЕЛИРОВАНИЕ ТРАНСПОРТНЫХ ПРОЦЕССОВ

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ
**по выполнению контрольной работы
и варианты заданий**

*для студентов IV курса
направления подготовки 23.03.01
заочной формы обучения*

Москва – 2020

Рецензент к.ф.-м.н., доцент О.Г. Илларионова.

Авторы: Ю. И. Дементьев, И. В. Платонова.

Моделирование транспортных процессов. Пособие по выполнению контрольных работ и варианты заданий для студентов четвёртого курса направления подготовки 23.03.01 заочной формы обучения – М.: МГТУ ГА, 2020.

Данное пособие издаётся в соответствии с рабочей программой дисциплины «Моделирование транспортных процессов» по учебному плану направления подготовки 23.03.01.

Рассмотрено и одобрено на заседаниях кафедры высшей математики 17.12.2019 г. и методического совета 17.03.2020 г.

Содержание

Введение

Студенты заочной формы обучения направления подготовки 23.03.01 «Технология транспортных процессов» изучают дисциплину *«Моделирование транспортных процессов»* на четвёртом курсе.

Распределение часов по видам занятий и форма контроля

Курс, семестр	Часы на дисциплину				Число конт-рольных работ	Форма контроля
	Общее число часов	Лекци-онные занятия	Практи-ческие занятия	Самостоя-тельная работа		
Курс 4 Семестр 7	180	10	4	166	1	Экзамен

Пособие содержит программу дисциплины с распределением на разделы и темы, методические указания по выполнению контрольной работы, задания контрольной работы (десять вариантов), теоретический материал по дисциплине, образец выполнения контрольной работы, приложения и список рекомендуемой литературы.

Программа дисциплины
«Моделирование транспортных процессов»

РАЗДЕЛ. 1. СТАТИСТИЧЕСКИЕ МЕТОДЫ МОДЕЛИРОВАНИЯ И АНАЛИЗА ТРАНСПОРТНЫХ ПРОЦЕССОВ

Тема 1.1. Базовые концепции методологии моделирования и анализа транспортных процессов (ТП).

Тема 1.2. Методы корреляционного анализа при исследовании взаимосвязей показателей ТП.

Тема 1.3. Однофакторное прогнозирование показателей ТП с использованием классической модели парной линейной регрессии.

Тема 1.4. Модели нелинейных регрессий для прогнозирования показателей ТП.

Тема 1.5. Моделирование и прогнозирование показателей ТП методами множественного регрессионного анализа.

Тема 1.6. Методы компонентного и факторного анализа при моделировании транспортных процессов.
Тема 1.7. Методы кластерного анализа при исследовании объектов воздушного транспорта (ВТ).

РАЗДЕЛ. 2. ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОПТИМИЗАЦИИ ТРАНСПОРТНЫХ ПРОЦЕССОВ

Тема 2.1. Моделирование и оптимизация использования дробных ресурсов на авиапредприятиях алгоритмами линейного программирования.
Тема 2.2. Экономико-математические методы оптимизации использования целочисленных ресурсов. Задача о рациональной загрузке воздушных судов (ВС).
Тема 2.3. Транспортная задача. Моделирование и оптимизация процесса стратегического использования парка ВС на сети воздушных линий (ВЛ).
Тема 2.4. Задача о назначениях. Моделирование и оптимизация процесса оперативного использования ВС, пилотов и бортпроводников.
Тема 2.5. Моделирование и оптимизация работы систем массового обслуживания (СМО).

Указания по выполнению контрольной работы

При выполнении контрольной работы необходимо строго придерживаться указанных ниже правил. Работы, выполненные без соблюдения этих правил, не зачитываются и возвращаются студенту для переработки.

1. Контрольная работа должна быть выполнена в тетради в клетку чернилами любого цвета, кроме красного. Необходимо оставлять поля для замечаний рецензента.
2. В заголовке работы на обложке тетради должны быть написаны печатными буквами фамилия, имя и отчество студента, учебный номер (шифр), название дисциплины, курс обучения и номер варианта.

3. В работу должны быть включены все задачи, указанные в задании, строго по положенному варианту. Контрольные работы, содержащие не все задачи, а также задачи не своего варианта, не зачитываются.
4. Решения задач надо располагать в порядке возрастания номеров.
5. Перед решением каждой задачи надо полностью переписать её условие. В том случае, если несколько задач, из которых студент выбирает задачи своего варианта, имеют общую формулировку, следует, переписывая условие задачи, заменить общие данные конкретными, взятыми из соответствующего номера.
6. Решения задач следует излагать подробно и аккуратно, объясняя и мотивируя все действия по ходу решения и делая необходимые чертежи и таблицы.
7. После получения прорецензированной работы, как незачтённой, так и зачтённой, студент должен исправить все отмеченные рецензентом ошибки и недочёты и выполнить все рекомендации рецензента. Если рецензент предлагает внести в решения задач те или иные исправления или дополнения, то это следует сделать в короткий срок.
8. При высылаемых исправлениях должна обязательно находиться прорецензированная работа и рецензия на неё. Поэтому рекомендуется при выполнении контрольной работы оставлять в конце тетради несколько чистых листов для всех дополнений и исправлений в соответствии с указаниями рецензента. Вносить исправления в сам текст работы после её рецензирования запрещается.
9. Прорецензированные контрольные работы вместе со всеми исправлениями и дополнениями следует сохранять. Без предъявления прорецензированной контрольной работы студент не допускается к сдаче экзамена.
10. **Номера задач, которые студент должен включить в свою контрольную работу, определяются по номеру варианта. Номер варианта совпадает с последней цифрой учебного номера (шифра) студента, при этом цифра 0 соответствует варианту 10.**

Методические указания по выполнению заданий

Методические указания по выполнению задания № 1

Парной линейной регрессией называется уравнение, описывающее корреляционную линейную связь между зависимой объясняемой переменной (результативным показателем) $\boldsymbol{y}$ и независимой объясняющей переменной (факторным признаком) $\boldsymbol{x}$: $\boldsymbol{\hat{y}=a+bx}$.

Уравнение регрессии строится на основании имеющихся исходных статистических данных конкретной выборки, содержащей $\boldsymbol{n}$ наблюдений – значений $\boldsymbol{x_i}$ и $\boldsymbol{y_i}$. В действительности в каждом отдельном случае (наблюдении) величина $\boldsymbol{y_i}$ складывается из суммы двух слагаемых: $\boldsymbol{y_i= \hat{y}_i +e_i}$, $\boldsymbol{i=1,2,...,n}$, где

$\boldsymbol{y_i}$ – фактическое значение результативного показателя;

$\boldsymbol{\hat{y}_i}$ – теоретическое значение результативного показателя, найденное по модели парной линейной регрессии (МПЛР);

$\boldsymbol{e_i}$ – случайная величина (остаток), характеризующая отклонения реального значения результативного показателя от теоретического, полученного по уравнению регрессии.

Случайная величина включает влияние неучтённых в модели случайных факторов. В связи с этим задача построения МПЛР заключается в том, чтобы по данным выборки найти неизвестные параметры $\boldsymbol{a}$ и $\boldsymbol{b}$. При этом параметры должны быть оценены так, чтобы полученное уравнение регрессии наилучшим образом описывало исходные данные. Будем находить эти оценки параметров с помощью метода наименьших квадратов (МНК), который позволяет минимизировать сумму квадратов остатков: $\boldsymbol{e_i= y_i – \hat{y}_i}$, где $\boldsymbol{\hat{y}_i= a+bx_i}$.

Таким образом, $S=\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - a - bx_i)^2 \to \min$.

Функция $\boldsymbol{S}$ достигает минимума при выполнении следующих условий:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0, \\ \dfrac{\partial S}{\partial b} = 0. \end{cases}$$

Отсюда получаем систему нормальных уравнений (СНУ):

$$\begin{cases} \mathrm{na} + \mathrm{b}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \\ \mathrm{a}\sum_{i=1}^{n} x_i + \mathrm{b}\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \end{cases}$$

Решаем систему относительно ***a*** и ***b***. Заметим, что сумма остатков $\sum_{i=1}^{n} e_i = 0$.

МНК можно применять, если остатки удовлетворяют следующим условиям:

1) остатки и объясняющая переменная не коррелируют;

2) остатки не коррелируют между собой;

3) математическое ожидание остатков равно нулю;

4) остатки должны быть гомоскедастичными, т.е. дисперсия остатков в среднем постоянна;

5) остатки должны быть распределены по нормальному закону.

В данной работе проверка указанных условий не требуется.

Замечание. *Если имеет место гетероскедастичность остатков (дисперсия остатков непостоянна), то в этом случае рекомендуется применять обобщённый метод наименьших квадратов.*

Для предварительной оценки качества полученного уравнения линейной регрессии рассчитывается коэффициент детерминации $\boldsymbol{R^2}$, который равен выборочному коэффициенту корреляции $\boldsymbol{r_{yx}^2}$. Коэффициент детерминации показывает, какая доля дисперсии результативного показателя $\boldsymbol{y}$ объясняется полученным уравнением регрессии. Соответственно, величина $\boldsymbol{1-R^2}$ характеризует долю дисперсии $\boldsymbol{y}$, вызванную влиянием случайных, неучтённых в модели факторов.

Числовые значения параметров уравнения регрессии являются выборочными оценками неизвестных параметров. Поэтому они содержат ошибку и не совпадают со значениями параметров в генеральной совокупности. Параметр уравнения линейной регрессии характеризует взаимосвязь факторного и результативного признаков (показателей). Это имеет место в том случае, если ожидаемое значение этого же параметра в генеральной совокупности с достаточно большой долей вероятности (например, 0,95) не равно нулю. В этом случае параметр считается ***значимым***. Для оценки значимости параметров ***a*** и ***b*** используется ***t-критерий Стьюдента***. Выдвигается предположение (гипотеза) о незначимости параметра. Вычисляется фактическое значение t-критерия Стьюдента как отношение значения самого параметра к его стандартной ошибке. Полученное значение надо сравнить с табличным значением (см. ***Приложение 1***) для уровня значимости ***α*** и числа степеней свободы ***df=n–m***, где ***m*** **–** число параметров в уравнении регрессии. Для парных регрессий значение ***m*** всегда равно 2. Если фактическое значение больше, чем табличное, то гипотеза отклоняется на данном уровне значимости.

Оценка значимости и адекватности уравнения в целом производится с помощью ***F-критерия Фишера***, который основан на анализе дисперсий. Общую сумму квадратов отклонений переменной ***y*** от среднего значения $\bar{y}$ можно представить в виде суммы квадратов отклонений, объяснённой регрессией и остаточной (необъяснённой) суммы квадратов отклонений.

$$\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2=\sum_{i=1}^{n}\left(\hat{y}_i-\bar{y}\right)^2+\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2.$$

Любая сумма квадратов отклонений связана с числом степеней свободы независимого варьирования признака. Так, на общую сумму квадратов отклонений приходится ***n–1*** степень свободы, на остаточную сумму – ***n–2***, а на объяснённую сумму – 1. Для определения фактического значения F-критерия Фишера рассматривают отношение объяснённой доли дисперсии в расчёте на одну степень свободы к остаточной доле в расчёте на каждую степень свободы.

$$F = \frac{\frac{R^2}{m-1}}{\frac{1-R^2}{n-m}} \quad \text{или} \quad F = \frac{R^2}{1-R^2} \times \frac{n-m}{m-1}$$, при этом число степеней свободы ***m-1*** называют числом степеней свободы числителя и обозначают df_1, а число степеней ***n–m*** **–** числом степеней свободы знаменателя и обозначают df_2. В нашем случае ***m=2***.

Выдвигается гипотеза H_0: $D_{\text{объясн.}} = D_{\text{небъясн}}$. Далее по таблице значений (см. ***Приложение 2***) находится критическое значение F-критерия Фишера при уровне значимости α и числе степеней свободы df_1 и df_2. Фактическое значение F-критерия Фишера сравнивается с критическим значением, найденным по таблице. Если $F > F_{\text{кр}}$, то гипотеза отклоняется и уравнение считается значимым и адекватным. В дальнейшем уравнение регрессии используется для получения точечных и интервальных прогнозных значений.

Замечание. *В рамках данной контрольной работы вычисляется только точечное прогнозное значение.*

Методические указания по выполнению задания № 2

Транспортная задача является важнейшей частной моделью класса задач линейного программирования и имеет обширные практические приложения при моделировании транспортных процессов. Особо важное значение она имеет при рассмотрении конкретных практических ситуаций, связанных с рационализацией поставок любых видов продукции и с оптимизацией планирования грузопотоков.

Постановка транспортной задачи

Найти оптимальный план перевозок некоторого однородного продукта из m исходных пунктов $A_1 \dots A_m$ к n потребителям $B_1 \dots B_n$. Известны объёмы производства продукции a_i в каждом пункте A_i $(i=1,2,\dots,m)$, спрос b_j каждого пункта B_j $(j=1,2\dots,n)$ и стоимость перевозки единицы продукции из каждого исходного пункта в каждый пункт назначения c_{ij}.

Обозначим через x_{ij} количество единиц продукции, которое необходимо доставить из пункта A_i в пункт B_j.

Математическая формулировка транспортной задачи сводится к нахождению минимального значения целевой функции $z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$ при заданных следующих ограничениях на запас и на спрос:

$$\begin{cases} \sum_{j=1}^{n} x_{ij} = a_i, \\ \sum_{i=1}^{m} x_{ij} = b_j. \end{cases}$$ При этом $\sum_{i=1}^{m} a_i = a$ и $\sum_{j=1}^{n} b_j = b$.

План перевозок с указанием запасов и потребностей удобно записывать в виде следующей таблицы, называемой таблицей перевозок.

Исходные пункты	***Пункты назначения***				***Запасы***
	$\boldsymbol{B_1}$	$\boldsymbol{B_2}$	**…**	$\boldsymbol{B_n}$	
$\boldsymbol{A_1}$	$\boldsymbol{c_{11}}$ $\boldsymbol{x_{11}}$	$\boldsymbol{c_{12}}$ $\boldsymbol{x_{12}}$	**…**	$\boldsymbol{c_{1n}}$ $\boldsymbol{x_{1n}}$	$\boldsymbol{a_1}$
$\boldsymbol{A_2}$	$\boldsymbol{c_{21}}$ $\boldsymbol{x_{21}}$	$\boldsymbol{c_{22}}$ $\boldsymbol{x_{22}}$	**…**	$\boldsymbol{c_{2n}}$ $\boldsymbol{x_{2n}}$	$\boldsymbol{a_2}$
…	**…**	**…**	**…**	**…**	**…**
$\boldsymbol{A_m}$	$\boldsymbol{c_{m1}}$ $\boldsymbol{x_{m1}}$	$\boldsymbol{c_{m2}}$ $\boldsymbol{x_{m2}}$	**…**	$\boldsymbol{c_{mn}}$ $\boldsymbol{x_{mn}}$	$\boldsymbol{a_m}$
Спрос	$\boldsymbol{b_1}$	$\boldsymbol{b_2}$	**…**	$\boldsymbol{b_n}$	$\boldsymbol{a=b}$

Если $a = b$, то модель называется сбалансированной транспортной моделью или моделью закрытого типа. Если $a \neq b$, то модель называется несбалансированной или открытого типа. В случае закрытой модели весь имеющийся в наличии продукт доставляется полностью, и спрос потребителей полностью удовлетворен. В случае открытой модели либо все потребители удовлетворены, но при этом в некоторых исходных пунктах остаются излишки продукта $(a > b)$, либо весь продукт доставлен полностью, а спрос удовлетворён только частично $(a < b)$. В данной контрольной работе предполагается решение транспортной задачи открытого типа.

Если $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$, то вводится фиктивный ***(m+1)***-й исходный пункт с запасами продукции $\boldsymbol{a_{m+1}=b-a}$ и все стоимости перевозки единицы продукции из этого пункта в пункты $\boldsymbol{B_j}$ ***(j=1,2,…,n)*** принимаются равными 0.

Если $\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$, то вводится фиктивный ***(n+1)***-й пункт назначения со спросом $\boldsymbol{b_{n+1}= a-b}$ и все стоимости перевозки единицы продукции в этот пункт назначения равны 0.

На предварительном этапе также, как и в симплекс-методе, определяется число базисных и небазисных переменных. После добавления дополнительной фиктивной строки или столбца в таблицу перевозок получаем сбалансированную модель, число уравнений в системе становится равным ***(n+m+1)***, одно из которых является зависимым. Поэтому число базисных переменных будет равно ***(n+m+1)-1=n+m***. Всего переменных $\boldsymbol{x_{ij}}$ в таблице с фиктивным исходным пунктом будет ***(m+1)n***, а с фиктивным пунктом назначения – ***m(n+1)***. В зависимости от этого определяется число небазисных переменных. Дальнейший алгоритм решения задачи состоит из следующих этапов.

1. Определить начальное допустимое базисное решение по методу минимальной стоимости (этот метод рекомендуется для задач открытого типа).
2. Из числа небазисных переменных выделить переменную, вводимую в базис, используя метод потенциалов (см. образец решения типового задания № 2). Если все небазисные переменные удовлетворяют условию оптимальности, то можно закончить вычисление; в противном случае перейти к следующему этапу 3.
3. Выбрать выводимую переменную из числа переменных текущего базиса, используя условие допустимости.
4. Найти новое базисное решение и затем вернуться к этапу 2.

Условие оптимальности. *Вводимой в базис переменной в задаче минимизации является небазисная переменная, имеющая наибольшую* *положительную оценку. В случае равенства оценок выбор делается произвольно. Если все оценки не положительны, то полученное решение является оптимальным.*

Условие допустимости. *Выводимой из базиса переменной является та переменная, которой присвоен знак «-» и которая имеет наименьшее значение.*

Методические указания по выполнению задания № 3

В классическом варианте задача о назначениях предполагает распределение ***n*** служащих на ***n*** рабочих мест. Её целью является нахождение оптимального решения (минимальной стоимости). Задача о назначениях является частным случаем транспортной задачи, в которой служащие соответствуют исходным пунктам, а рабочие места – пунктам назначения. «Стоимость» распределения служащего на рабочее место равна c_{ij}. Но особенностью этой задачи является то, что все величины спроса и запаса (предложения) равны 1 (см. Таблицу назначений). В связи с этим, задачу о назначениях можно решать совершенно другим, упрощённым алгоритмом, который называется ***венгерским методом*** (см. образец решения типового задания № 3).

Служащие	***Назначения на рабочие места***				***Предложение***
	1	***2***	**…**	***n***	
1	c_{11}	c_{12}	**…**	c_{1n}	***1***
2	c_{21}	c_{21}	**…**	c_{2n}	***1***
…	**…**	**…**	**…**	**…**	**…**
n	c_{n1}	c_{n2}	**…**	c_{nn}	***1***
Спрос	***1***	***1***	**…**	***1***	

Венгерский метод не имеет отношения к транспортной задаче, однако, стоит отметить, что он, как и метод потенциалов, основан на симплекс-методе. Задача о назначениях занимает особое место при моделировании транспортных процессов – это оптимальное распределение ВС по ВЛ, экипажей пилотов и бортпроводников по ВС.

Методические указания по выполнению задания № 4

Системы массового обслуживания (СМО) занимают особое место при моделировании транспортных процессов в гражданской авиации (ГА). Например, любой аэропорт является достаточно сложной СМО.

СМО *называется система, включающая несколько каналов (технических устройств), выполняющих одинаковые по характеру функции по обслуживанию поступающих случайным образом заявок (требований) в систему*. СМО подразделяются на СМО с отказами и СМО с ожиданием.

СМО с ожиданием *– это система, в которой в случае занятости всех каналов заявка становится в очередь и ожидает момента, когда она будет принята на обслуживание.* СМО с ожиданием делятся на системы с ограниченной и неограниченной очередью. В задании № 4 данной контрольной работы предполагается использование многоканальной модели СМО с ожиданием и неограниченной очередью. Заявки, поступающие в СМО в случайные моменты времени, образуют случайный простейший поток, который подчиняется распределению Пуассона. ***Поток называется простейшим,*** *если он обладает тремя свойствами: стационарностью, ординарностью и отсутствием последействия.*

Стационарность – это свойство заключается в том, что среднее число заявок, поступающее в систему, должно быть постоянным.

Ординарность – это свойство заключается в том, что вероятность поступления в систему двух и более заявок в течение бесконечно малого промежутка времени бесконечно мала по сравнению с поступлением одной или ни одной заявки.

Отсутствие последействия – заявки поступают в систему независимо друг от друга.

При простейшем потоке интервал времени между двумя последовательными поступлениями заявок есть случайная величина, подчинённая показательному закону распределения. Интенсивность входящего потока λ и время между последовательными поступлениями заявок t_0 связаны между собой взаимно обратной зависимостью. Время обслуживания заявки $t_{обс}$ также является случайной величиной, подчинённой показательному закону распределения. Если СМО имеет n каналов обслуживания, то интенсивность обслуживания увеличивается в n раз. В результате того, что заявки поступают в систему случайным образом и в случайное время, время обслуживания заявки является случайной величиной, то и весь процесс функционирования СМО носит случайный характер. При простейшем потоке заявок и показательном законе распределения времени обслуживания этот процесс является ***марковским***. В процессе функционирования СМО может находиться в различных дискретных состояниях. Каждое состояние системы отличается от другого состояния числом заявок, находящихся в системе. Установлено, что по истечении определённого времени, процесс функционирования СМО переходит в установившийся режим. При работе системы в установившемся режиме вероятности её состояний не зависят от времени. Входными параметрами СМО с ожиданием и неограниченной очередью являются:

n – число каналов;

λ – интенсивность входящего потока (или t_0 – среднее время между двумя последовательными поступлениями заявок);

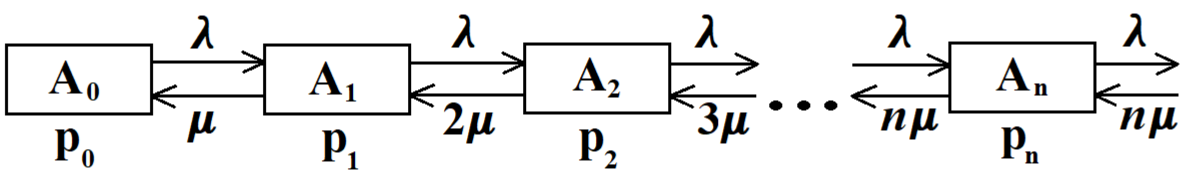
μ – интенсивность обслуживания (или $t_{обс}$ – среднее время обслуживания заявки).

При этом $\lambda=1/t_0$ и $\mu=1/t_{обс}$.

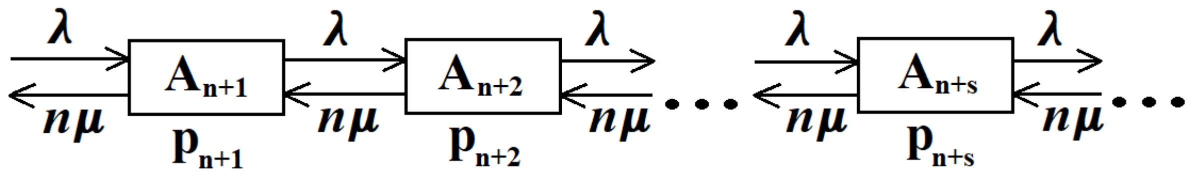
Для решения задачи целесообразно ввести безразмерную величину – приведённую интенсивность α.

Граф состояний СМО с ожиданием и неограниченной очередью

1. Состояния СМО 1 группы (когда нет очереди)



2. Состояния СМО 2 группы (когда есть очередь)



A_0 – состояние системы, в котором нет ни одной заявки и все каналы свободны, этому состоянию соответствует вероятность p_0.

A_k – состояние системы, в котором есть k заявок и k каналов заняты их обслуживанием, этому состоянию соответствует вероятность p_k, $k=1,...,n$.

A_{n+s} – состояние системы, в котором есть $n+s$ заявок, при этом n каналов заняты обслуживанием, а s заявок находятся в очереди, этому состоянию соответствует вероятность p_{n+s}, $s=1,2,3,...$

Приведённой интенсивностью входящего потока *называется среднее число заявок, поступившее в систему за среднее время обслуживания.*

$\boldsymbol{\alpha = \lambda\, t_{обс}}$ или $\boldsymbol{\alpha = \lambda / \mu}$.

Постановка задачи СМО с ожиданием и неограниченной очередью: В систему, состоящую из $\boldsymbol{n}$ каналов обслуживания, поступает простейший поток заявок с интенсивностью λ. Заявка обслуживается любым свободным каналом с интенсивностью обслуживания $\boldsymbol{\mu}$ или средним временем обслуживания $\boldsymbol{t_{обс}}$. В случае занятости всех каналов заявка становится в очередь и ожидает момента принятия на обслуживание.

Вероятностные характеристики функционирования в стационарном режиме многоканальной СМО с ожиданием и неограниченной очередью приведены в образце решения типового задания № 4.

Задания контрольной работы

Задание № 1

Постановка задачи. Построить модель парной линейной регрессии (МПЛР) $\hat{y}=a+bx$ по динамическому ряду, где зависимая переменная y – ежемесячная численность перевезенных пассажиров авиакомпанией N (млн. чел.), а x – номер периода (месяца). Провести полное исследование и всесторонний анализ качества, значимости, адекватности и точности полученной модели по следующей схеме.

1. Построить поле корреляции переменных x и y с помощью приложения Microsft Excel, провести визуальный анализ исходных данных, сделать вывод о целесообразности использования линейной регрессии.
2. Найти уравнение МПЛР. Для этого с помощью МНК решить СНУ и найти оценки параметров a и b.
3. Найти значение коэффициента корреляции r_{yx}, оценить его значимость с помощью t-критерия Стьюдента.
4. Предварительно оценить качество полученной модели с помощью коэффициента детерминации R^2.
5. Оценить значимость параметра b в уравнении регрессии при переменной x с помощью t-критерия Стьюдента.
6. Оценить значимость и адекватность уравнения в целом с помощью F-критерия Фишера, проверить тождество $|t_r|=|t_b|=\sqrt{F}$.
7. Оценить точность полученной модели. Для этого рассчитать значение относительной ошибки аппроксимации.
8. Найти прогнозное значение $\hat{y}$ для периода $x=13$, построить графики исходных, расчётных и прогнозного значений с помощью приложения Microsft Excel.
9. Сделать вывод по модели.

Все вспомогательные расчёты желательно производить с использованием приложения Microsft Excel, числа округлять не менее, чем до тысячных долей по правилам округления, принятым в математике. В противном случае возникает большая погрешность.

Исходные данные по вариантам представлены в таблице № 1.

Таблица № 1

1.1	***y***	2,05	2,28	2,22	2,23	2,20	2,31	2,48	2,35	2,42	2,39	2,45	2,55
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.2	***y***	2,15	2,38	2,32	2,39	2,4	2,42	2,48	2,46	2,51	2,49	2,53	2,56
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.3	***y***	1,35	1,48	1,52	1,53	1,51	1,6	1,71	1,65	1,72	1,69	1,75	1,82
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.4	***y***	1,56	1,69	1,73	1,74	1,72	1,81	1,92	1,86	1,93	1,9	1,96	2,03
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.5	***y***	1,26	1,49	1,43	1,5	1,51	1,53	1,59	1,57	1,62	1,6	1,64	1,67
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.6	***y***	1,66	1,79	1,73	1,8	1,81	1,83	1,89	1,87	1,92	1,9	1,94	1,97
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.7	***y***	3,03	3,26	3,2	3,21	3,18	3,29	3,46	3,33	3,4	3,37	3,43	3,53
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.8	***y***	3,14	3,36	3,31	3,38	3,39	3,41	3,47	3,45	3,5	3,48	3,52	3,55
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.9	***y***	2,66	2,89	2,83	2,9	2,91	2,93	2,98	2,95	3,02	3,00	3,04	3,07
	x	1	2	3	4	5	6	7	8	9	10	11	12
1.10	***y***	2,44	2,67	2,61	2,68	2,69	2,71	2,76	2,73	2,8	2,78	2,82	2,85
	x	1	2	3	4	5	6	7	8	9	10	11	12

Задание № 2

Постановка задачи. Авиакомпания выполняет рейсы по 4 воздушным линиям (ВЛ) на 3 типах воздушных судов (ВС). Известны:

1) c_{ij} – расходы на 1 млн. ткм на i-ом типе ВС по j-й ВЛ (ден. усл. ед./млн. ткм);

2) a_i – потенциал i-го типа ВС (млн. ткм);

3) b_j – прогноз спроса по j-й ВЛ (млн. ткм). Для ВЛ, на которых i-й тип самолета не используется, рекомендуется назначить расходы c_{ij}=100.

Исходные данные задачи c_{ij}, a_i, b_j по вариантам представлены в таблицах 2.1 – 2.10. Составить план перевозок ВС по ВЛ, дающий минимальные расходы. Начальное решение рекомендуется находить по методу минимальной стоимости, а затем решать задачу с использованием метода потенциалов.

2.1

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	6	8	4	***50***
2	2	100	6	9	***70***
3	4	7	9	100	***65***
b_j	***45***	***50***	***60***	***55***	

2.2

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	6	4	8	***60***
2	6	100	7	6	***50***
3	4	7	100	9	***50***
b_j	***45***	***55***	***40***	***60***	

2.3

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	100	4	8	***50***
2	3	6	100	6	***45***
3	4	7	6	7	***55***
b_j	***50***	***40***	***40***	***40***	

2.4

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	6	4	100	***50***
2	2	100	7	6	***60***
3	4	7	6	9	***40***
b_j	***50***	***50***	***40***	***40***	

2.5

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	100	8	4	8	***50***
2	3	5	7	100	***30***
3	5	7	6	9	***50***
b_j	***40***	***40***	***40***	***40***	

2.6

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	6	8	100	***40***
2	100	3	7	6	***60***
3	4	7	10	9	***50***
b_j	***40***	***50***	***40***	***50***	

2.7

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	7	4	8	***60***
2	100	3	7	6	***60***
3	6	7	100	9	***50***
b_j	***40***	***50***	***60***	***50***	

2.8

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	3	9	4	100	***50***
2	100	3	7	6	***60***
3	6	7	6	7	***50***
b_j	***40***	***50***	***50***	***50***	

2.9

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	5	6	4	5	***60***
2	8	3	7	100	***60***
3	4	100	9	4	***80***
b_j	***60***	***60***	***60***	***60***	

2.10

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	8	4	4	5	***50***
2	5	6	7	100	***75***
3	4	100	9	4	***75***
b_j	***60***	***60***	***60***	***60***	

Задание № 3

Постановка задачи. Авиакомпания выполняет 5 рейсов по 5 ВЛ. Задана матрица себестоимостей каждого рейса по каждой ВЛ. Найти оптимальный план назначений данных рейсов на данные ВЛ с использованием «венгерского» алгоритма. Исходные данные задачи по вариантам представлены в таблицах 3.1 – 3.10.

3.1

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	10	5	9	18	11
2	13	19	6	12	4
3	3	2	4	4	5
4	18	9	12	17	15
5	11	6	14	19	10

3.2

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	7	5	12	9	8
2	10	4	6	3	12
3	7	2	8	6	5
4	6	13	5	7	9
5	11	7	4	16	3

3.3

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	6	2	8	6	5
2	11	4	6	3	12
3	7	5	12	9	8
4	6	13	5	7	9
5	11	7	4	16	3

3.4

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	12	4	7	3	10
2	8	2	7	5	5
3	9	5	11	8	9
4	6	10	5	9	7
5	14	8	4	15	3

3.5

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	17	12	9	8	10
2	5	11	6	14	7
3	2	4	13	9	12
4	8	12	5	7	6
5	14	5	16	3	10

3.6

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	13	8	4	1	5
2	7	9	10	18	6
3	3	6	11	10	14
4	7	1	2	8	9
5	12	9	8	4	16

3.7

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	2	15	1	2	3
2	5	19	8	6	13
3	4	7	5	9	11
4	8	6	3	9	7
5	10	8	7	5	6

3.8

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	8	11	13	4	5
2	4	3	5	6	7
3	9	11	4	6	5
4	3	8	10	9	11
5	4	12	7	8	6

3.9

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	9	12	14	10	13
2	11	10	12	9	9
3	8	14	11	13	9
4	9	7	9	7	10
5	12	10	8	9	11

3.10

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	11	8	6	10	5
2	14	7	4	8	7
3	10	9	11	6	8
4	9	12	8	10	9
5	7	9	8	9	7

Задание № 4

Постановка задачи. В терминале Q аэропорта города N в течение определенного времени $\boldsymbol{T}$ открыто $\boldsymbol{n}$ стоек для регистрации пассажиров на 3 ближайших рейса. Ожидаемое число пассажиров зависит от конкретных типов ВС, готовящихся совершить вылет. Требуется определить вероятностно-статистические показатели эффективности обслуживания пассажиров в терминале Q аэропорта города N. Для решения задачи нужно использовать модель СМО с ожиданиями. Исходные данные по вариантам о периоде времени $\boldsymbol{T}$, типах ВС, коэффициенте загрузки $\boldsymbol{k}$, среднем времени регистрации пассажира $\boldsymbol{t_{обс}}$, количестве стоек регистрации $\boldsymbol{n}$ представлены в таблице № 4. Средние данные о количестве кресел ***L*** для определённых типов ВС приведены в таблице № 4А.

Найти значения следующих показателей.

1. Приведённую интенсивность (округлить до десятых по правилам округления, принятым в математике).
2. Вероятность, что все стойки регистрации свободны.
3. Вероятности состояний системы 1 и 2 групп.
4. Вероятности ожидания и немедленного обслуживания пассажиров.
5. Среднее число пассажиров в очереди.
6. Среднее время ожидания регистрации пассажиром в очереди.
7. Среднее время ожидания регистрации и собственно процедуры регистрации пассажира у стойки.
8. Среднее число занятых стоек регистрации.
9. Написать вывод.

Таблица № 4

Исходные данные

№ варианта	***T (час.)***	***Тип ВС***	***k***	$t_{обс}$ ***(сек.)***	***n***
4.1	00-01	А-319; А-320; В-767	0,75	50	6
4.2	01-02	А-320; А-321; В-787	0,8	48	6
4.3	02-03	В-757; В-767; В-787	0,75	44	7
4.4	03-04	А-330; А-320; А-321	0,8	45	7
4.5	04-05	А-320; А-321; В-757	0,77	52	6
4.6	05-06	А-319; А-320; В-787	0,85	46	6
4.7	06-07	А-330; А-320; В-757	0,76	45	7
4.8	07-08	А-319; А-321; В-757	0,75	55	6
4.9	08-09	А-320; В-757; В-767	0,81	48	7
4.10	09-10	А-320; А-321; В-757	0,82	52	7

Таблица № 4А

Тип ВС	**Количество кресел *L***
А-319	156
А-320	140
А-321	160
А-330	360
В-757	200
В-767	260
В-787	240

Образец выполнения контрольной работы

Типовое задание № 1

Постановка задачи. Построить (МПЛР) $\hat{y}=a+bx$ по динамическому ряду, где зависимая переменная y – ежемесячная численность перевезенных пассажиров авиакомпанией N (млн. чел.), x – номер периода (месяца).

y	4,05	4,18	4,16	4,24	4,22	4,35	4,32	4,46	4,52	4,6	4,58	4,66
x	1	2	3	4	5	6	7	8	9	10	11	12

Провести полное исследование и всесторонний анализ качества, значимости, адекватности и точности полученной модели по схеме, изложенной в задании № 1.

Решение задачи

1. Проведем визуализацию исходных данных с помощью приложения Micrositt Excel. Для этого построим поле корреляции между переменными x и y.

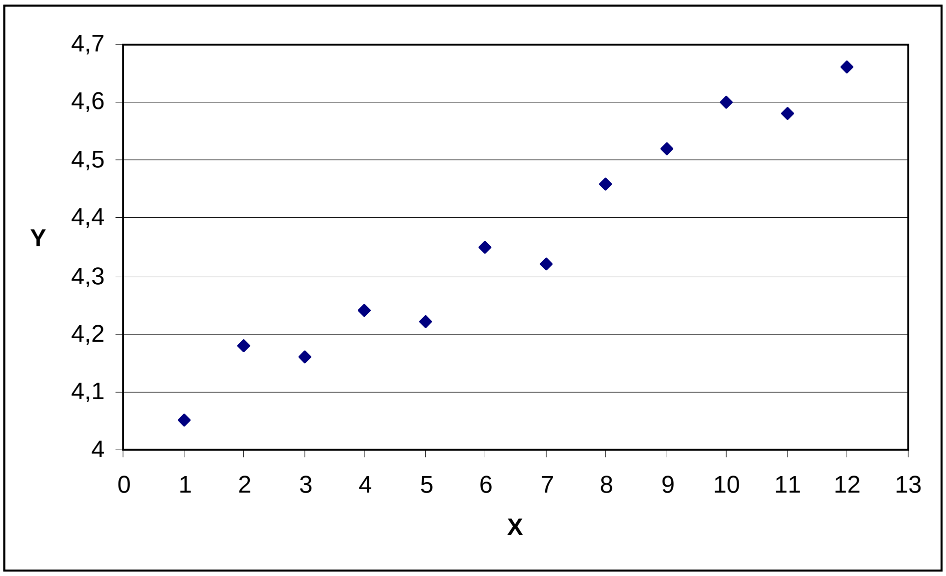


Рис. 1. Динамика численности перевезенных пассажиров

Точки на графической иллюстрации как бы выстраиваются вдоль некоторой прямой. Таким образом, визуальный анализ подтверждает целесообразность использования МПЛР.

2. Найдём уравнение МПЛР $\hat{y}=a+bx$. Для этого с помощью МНК решим СНУ:

$$\begin{cases} na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \end{cases}$$

По условию задания число наблюдений ***n=12***. Для определения параметров ***a*** и ***b*** найдём значения сумм:

$$\sum_{i=1}^{12} x_i = 1+2+3+4+5+6+7+8+9+10+11+12 = 78;$$

$$\sum_{i=1}^{12} y_i = 4{,}05+4{,}18+4{,}16+4{,}24+4{,}22+4{,}35+4{,}32+4{,}46+4{,}52+4{,}6+4{,}58+4{,}66 = 52{,}34;$$

$$\sum_{i=1}^{12} x_i^2 = 1^2+2^2+3^2+4^2+5^2+6^2+7^2+8^2+9^2+10^2+11^2+12^2 = 650;$$

$$\sum_{i=1}^{12} x_i y_i = 1\cdot 4{,}05+2\cdot 4{,}18+3\cdot 4{,}16+4\cdot 4{,}24+5\cdot 4{,}22+6\cdot 4{,}35+7\cdot 4{,}32+$$
$$+8\cdot 4{,}46+9\cdot 4{,}52+10\cdot 4{,}6+11\cdot 4{,}58+12\cdot 4{,}66 = 347{,}95.$$

Подставим полученные значения в СНУ:

$$\begin{cases} 12a + 78b = 52{,}34, \\ 78a + 650b = 347{,}95. \end{cases}$$

Разделим обе части первого уравнения на 12, а второго – на 78, получим следующую систему:

$$\begin{cases} a + 6{,}5b = 4{,}362, \\ a + 8{,}333b = 4{,}461. \end{cases}$$

В вычислениях производим округления до 3 – 4 значащих цифр. Для того, чтобы найти значение оценки параметра ***b***, вычтем из второго уравнения системы первое уравнение:

$a + 8{,}333b – a – 6{,}5b=4{,}461– 4{,}362;$

$1{,}833b = 0{,}099;$

$b = 0{,}054.$

Найдем значение оценки параметра ***a:*** $a = 4{,}362 - 6{,}5 \cdot 0{,}054 = 4{,}011.$

Таким образом, получим решение системы:

$$\begin{cases} a = 4{,}011, \\ b = 0{,}054. \end{cases}$$

Тогда искомое уравнение МПЛР имеет вид $\hat{y} = 4{,}011 + 0{,}054x$.

Оценку параметра $\boldsymbol{b}$ называют коэффициентом регрессии. Его значение, равное 0,054, показывает, что в течение каждого месяца численность перевезённых пассажиров увеличивается, в среднем, на $0{,}054 \cdot 1000000 = 54000$ человек.

3. Найдём значение выборочного коэффициента корреляции $\boldsymbol{r_{yx}}$ и оценим его значимость с помощью t-критерия Стьюдента. В дальнейшем, будем называть выборочный коэффициент корреляции $\boldsymbol{r_{yx}}$ коэффициентом корреляции. Воспользуемся следующей формулой для расчёта коэффициента корреляции:

$$r_{yx} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}.$$

Найдём значение суммы $\sum_{i=1}^{n} y_i^2$.

$$\sum_{i=1}^{n} y_i^2 = 4{,}05^2 + 4{,}18^2 + 4{,}16^2 + 4{,}24^2 + 4{,}22^2 + 4{,}35^2 + 4{,}32^2 + 4{,}46^2 + 4{,}52^2 + 4{,}6^2 + 4{,}58^2 + 4{,}66^2 = 228{,}7254.$$

Найдём значение коэффициента корреляции, для этого подставим все значения сумм в формулу.

$$r_{yx} = \frac{12 \cdot 347{,}95 - 78 \cdot 52{,}34}{\sqrt{12 \cdot 650 - 78^2} \cdot \sqrt{12 \cdot 228{,}7254 - 52{,}34^2}} = 0{,}9805.$$

Если значение r_{yx} по модулю больше 0,7, то, говорят, что между переменными $\boldsymbol{x}$ и $\boldsymbol{y}$ наблюдается тесная взаимосвязь. Если $r_{yx} > 0,$ то взаимосвязь называют прямой, если $r_{yx} < 0,$ то взаимосвязь называют обратной. Если $r_{yx} \approx 0,$ то между переменными $\boldsymbol{x}$ и $\boldsymbol{y}$ отсутствует линейная взаимосвязь.

В нашем случае полученное значение $r_{yx} = 0,9805$ указывает на то, что между переменными ***x*** и ***y*** наблюдается тесная прямая взаимосвязь.

Оценим значимость коэффициента корреляции. Для этого выдвинем гипотезу H_0 об отсутствии линейной взаимосвязи между переменными и альтернативную гипотезу H_1.

H_0: $r_{yx} = 0$, H_1: $r_{yx} \neq 0$.

Проверим нулевую гипотезу H_0 при конкурирующей гипотезе H_1 на уровне значимости $\boldsymbol{\alpha=0,05}$. Для этого найдём фактическое значение t-критерия, которое обозначим t_r.

$t_r = \dfrac{r_{yx}}{\sigma_r}$, где σ_r – ошибка коэффициента корреляции.

$$\sigma_r = \frac{\sqrt{1-r_{yx}^2}}{\sqrt{n-2}}, \quad \sigma_r = \frac{\sqrt{1-0,9805^2}}{\sqrt{12-2}} = 0,06214.$$

$$t_r = \frac{0,9805}{0,06214} = 15,779.$$

Найдём критическое значение $\boldsymbol{t_{кр}}$ двусторонней критической области по таблице значений t-критерия Стьюдента на уровне значимости $\boldsymbol{\alpha=0,05}$ и числе степеней свободы ***df=n-2***, где ***n=12*** – число наблюдений (см. Приложение 1).

$t_{кр}(\alpha=0,05;\ df=10)=2,23$.

Сравним фактическое значение $\boldsymbol{t_r}$, взятое по модулю, с критическим $\boldsymbol{t_{кр}}$. Если $|t_r| < t_{кр}$, то H_0 не отклоняется на данном уровне значимости, а если $|t_r| \geq t_{кр}$, то отклоняется.

Величина $\boldsymbol{t_r=15,779}$ значительно превышает табличное значение 2,23, следовательно, полученное значение коэффициента существенно отличается от нуля и зависимость является значимой.

4. Оценим качество полученной модели с помощью коэффициента детерминации $\boldsymbol{R^2}$, который определяет, какая доля общей вариации анализируемой результативной переменной ***y*** обусловлена изменением

объясняющей переменной. Для МПЛР коэффициент детерминации можно вычислить с помощью возведения в квадрат коэффициента корреляции.

$$R^2 = r_{yx}^2; \quad R^2 = (0,9805)^2 = 0,961.$$

Полученное значение коэффициента детерминации показывает, что 96% доли общей дисперсии объясняется полученным уравнением линейной регрессии, а оставшиеся 4% приходятся за счет воздействия случайных и неучтенных факторов.

5. Оценим значимость параметра ***b*** в уравнении регрессии при переменной ***x*** с помощью t-критерия Стьюдента. Рассмотрим гипотезу H_0 о незначимом отличии от нуля выборочного значения параметра ***b*** и альтернативную гипотезу H_1 о неравенстве нулю анализируемого параметра.

H_0: $b = 0$, H_1: $b \neq 0$.

Проверим нулевую гипотезу H_0 при конкурирующей гипотезе H_1 на уровне значимости ***α=0,05***. Для этого найдем фактическое значение t-критерия, которое обозначим t_b.

$t_b = \dfrac{b}{\sigma_b}$, где σ_b – оценка стандартной ошибки параметра ***b***.

$\sigma_b = \dfrac{\sigma_y \cdot \sigma_r}{\sigma_x}$, где σ_y–оценка среднеквадратического отклонения показателя y, σ_x – оценка среднеквадратического отклонения признака x.

$\sigma_y = \sqrt{\dfrac{1}{n} \cdot \sum_{i=1}^{n} y_i^2 - \bar{y}^2}$, $\quad \sigma_x = \sqrt{\dfrac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 - \bar{x}^2}$, где $\bar{x}$, $\bar{y}$ – выборочные средние значения признаков x и y.

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i.$$

Используемые при подсчётах значения были найдены в пунктах 2 и 3.

$$\bar{x} = \frac{78}{12} = 6,5, \quad \bar{y} = \frac{52,34}{12} = 4,3617, \quad \sigma_y = \sqrt{\frac{1}{12} \cdot 228,7254 - (4,3617)^2} = 0,1898,$$

$$\sigma_x = \sqrt{\frac{1}{12} \cdot 650 - (6,5)^2} = 3,452, \quad \sigma_b = \frac{0,1898 \cdot 0,06214}{3,452} = 0,00342.$$

$$t_b = \frac{0,054}{0,00342} = 15,789.$$

Найдём критическое значение $t_{кр}$ двусторонней критической области по таблице значений t-критерия Стьюдента на уровне значимости $\alpha=0,05$ и числе степеней свободы $df=n-2$, где $n=12$ – число наблюдений (см. Приложение 1).

$t_{кр}(\alpha=0,05;\ df=10)=2,23$.

Сравним фактическое значение t_b, взятое по модулю, с критическим $t_{кр}$. Если $|t_b| < t_{кр}$, то H_0 не отклоняется на данном уровне значимости, а если $|t_r| \geq t_{кр}$, то отклоняется.

Величина $t_b=15,789$ значительно превышает табличное значение 2,23, следовательно, полученное значение параметра b существенно отличается от нуля и зависимость является значимой.

6. Оценим качество, значимость и адекватность уравнения в целом с помощью F-критерия Фишера.

Для этого рассмотрим гипотезу H_0 о равенстве «объяснённой» и «необъяснённой» дисперсии: H_0: $D_{объясн.} = D_{небъясн.}$

Найдём значение F-критерия, которое определяется как отношение «объяснённой» дисперсии к «необъяснённой» в расчёте на одну степень свободы по формуле:

$$F = \frac{R^2}{1} : \frac{1-R^2}{n-2} \quad \text{или} \quad F = \frac{R^2}{1-R^2} \times \frac{n-2}{1}.$$

$$F = \frac{0,9805^2}{1-0,9805^2} \times \frac{12-2}{1} = \frac{0,9805^2}{1-0,9805^2} \times 10 = 248,935.$$

Найдём критическое значение $F_{кр}$ по таблице значений F-критерия Фишера (см. Приложение 2) на уровне значимости $\alpha=0,05$, числе степеней свободы числителя $df_1=m-1$, числе степеней свободы знаменателя $df_2=n-m$, где $n=12$ – число наблюдений, а $m=2$ – число параметров в уравнении регрессии. Получаем, $F_{кр}(\alpha=0,05;\ df_1=1;\ df_2=10)=4,96$. Сравним расчетное значение F – с критическим $F_{кр}$. Если $F < F_{кр}$, то гипотеза H_0 не отклоняется на данном уровне значимости, а если $F > F_{кр}$, то отклоняется. Величина $F=248,935$ значительно превышает табличное значение 4,96, следовательно, нулевую гипотезу следует отклонить, а полученное уравнение линейной регрессии считать значимым.

Проверим правильность выполненных расчётов. Если все расчёты выполнены верно, то должны выполняться тождества $|t_r|=|t_b|=\sqrt{F}$.

$15{,}779 \approx 15{,}789 \approx \sqrt{248{,}935}$; $15{,}779 \approx 15{,}789 \approx 15{,}778.$

Отличия в расчётных значениях критериев из-за ошибок округления оказались несущественными, это подтверждает правильность вычислений.

Примечание. «Объяснённую» дисперсию также называют факторной, а «необъяснённую» – остаточной.

7. Этап верификации очень важен для процесса моделирования. После построения модели нужно оценить её точность. Это можно сделать с помощью вычисления значения относительной ошибки аппроксимации, которая показывает, насколько отличаются полученные расчетные значения результативного показателя $\hat{y}$ от исходных (реальных) значений y.

$$A\% = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \cdot 100\%.$$

Найдём все расчётные значения $\hat{y}_i$, для этого воспользуемся полученным в пункте 2 уравнением линейной регрессии:

$\hat{y}_1 = 4{,}011 + 0{,}054 \cdot 1 = 4{,}065;$ $\hat{y}_2 = 4{,}011 + 0{,}054 \cdot 2 = 4{,}119;$

$\hat{y}_3 = 4{,}011 + 0{,}054 \cdot 3 = 4{,}173;$ $\hat{y}_4 = 4{,}011 + 0{,}054 \cdot 4 = 4{,}227;$

$\hat{y}_5 = 4{,}011 + 0{,}054 \cdot 5 = 4{,}281;$ $\hat{y}_6 = 4{,}011 + 0{,}054 \cdot 6 = 4{,}335;$

$\hat{y}_7 = 4{,}011 + 0{,}054 \cdot 7 = 4{,}389;$ $\hat{y}_8 = 4{,}011 + 0{,}054 \cdot 8 = 4{,}443;$

$\hat{y}_9 = 4{,}011 + 0{,}054 \cdot 9 = 4{,}497;$ $\hat{y}_{10} = 4{,}011 + 0{,}054 \cdot 10 = 4{,}551;$

$\hat{y}_{11} = 4{,}011 + 0{,}054 \cdot 11 = 4{,}605;$ $\hat{y}_{12} = 4{,}011 + 0{,}054 \cdot 12 = 4{,}659.$

Найдём значение $\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}$.

$$\sum_{i=1}^{12}\frac{|y_i-\hat{y}_i|}{y_i}=\frac{|4{,}05-4{,}065|}{4{,}05}+\frac{|4{,}18-4{,}119|}{4{,}18}+\frac{|4{,}16-4{,}173|}{4{,}16}+\frac{|4{,}24-4{,}227|}{4{,}24}+\frac{|4{,}22-4{,}281|}{4{,}22}+$$
$$\frac{|4{,}35-4{,}335|}{4{,}35}+\frac{|4{,}32-4{,}389|}{4{,}32}+\frac{|4{,}46-4{,}443|}{4{,}46}+\frac{|4{,}52-4{,}497|}{4{,}52}+\frac{|4{,}6-4{,}551|}{4{,}6}+\frac{|4{,}58-4{,}605|}{4{,}58}+$$
$$+\frac{|4{,}66-4{,}659|}{4{,}66}\approx 0{,}0836.$$

$$A\%=\frac{1}{12}\cdot 0{,}0836\cdot 100\%\approx 0{,}7\%.$$

Значения относительной ошибки аппроксимации, не превосходящие 7%, свидетельствуют о хорошем соответствии линии регрессии исходным данным. В нашем случае относительная ошибка составила менее 1%.

8. Найдём значение точечного прогноза $\hat{y}$ для периода $x=13$:

$$\hat{y}_{13}=4{,}011+0{,}054\cdot 13=4{,}713.$$

Таким образом, можно сделать вывод, что ожидаемое число авиапассажиров в 13 периоде (месяце) составит 4713000 человек.

Построим графики исходных, расчётных и прогнозного значений с помощью приложения Microsft Excel (Рис.2). По принятым в мировой практике правилам, отрезок, соединяющий крайнее расчётное значение с прогнозным, изображается пунктирной линией.

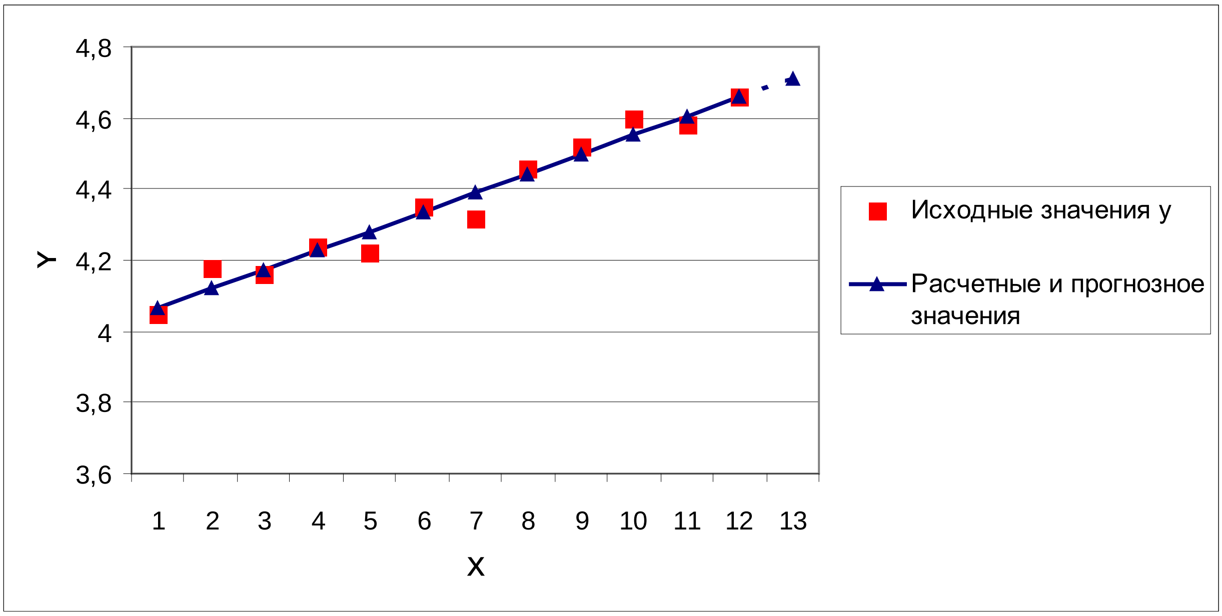


Рис. 2. Графики исходных, расчётных и прогнозного значений

9. Был проведён всесторонний корреляционно-регрессионный анализ полученной модели линейной регрессии. С использованием МНК было

выведено уравнение тренда, отражающее положительную динамику численности перевезённых авиапассажиров компанией N. В среднем ежемесячный прирост пассажиров составил 54 000 человек. Оценка статистической значимости коэффициента корреляции и параметра $\boldsymbol{b}$ при переменной $\boldsymbol{x}$ в уравнении регрессии проводилась с помощью t-критерия Стьюдента на уровне значимости $\boldsymbol{\alpha=0,05}$. Качество, значимость и адекватность модели в целом оценивались с использованием F-критерия Фишера. Коэффициент корреляции, параметр $\boldsymbol{b}$ и уравнение регрессии были признаны значимыми на уровне $\boldsymbol{\alpha=0,05}$. Для оценки точности полученного уравнения регрессии проводилось сопоставление реальных и модельных данных с помощью расчета относительной ошибки аппроксимации, которая не превысила 1%. Всё это свидетельствует о том, что данная модель может быть использована для прогнозирования численности авиапассажиров на определённом этапе транспортного процесса. Стоит отметить о необходимости осуществления регулярного мониторинга ситуации на рынке авиаперевозок с целью выявления происходящих изменений.

Типовое задание № 2

Постановка задачи. Авиакомпания Q выполняет рейсы по n=4 воздушным линиям (ВЛ) на m=3 типах воздушных судов (ВС). Известны:

1) $\boldsymbol{c_{ij}}$ – транспортные расходы на 1 млн. ткм на i-ом типе ВС по j-ой ВЛ (ден. усл. ед./млн. ткм);

2) $\boldsymbol{a_i}$ – потенциал i-ого типа ВС (млн. ткм);

3) $\boldsymbol{b_j}$ – спрос по грузообороту по j-ой ВЛ (млн. ткм).

На ВЛ1 не используется 1 тип ВС, а на ВЛ2 – 2 тип ВС. В этом случае расходы $\boldsymbol{c_{11}}$=100 и $\boldsymbol{c_{22}}$ =100. Исходные данные задачи $\boldsymbol{c_{ij}}$, $\boldsymbol{a_i}$, $\boldsymbol{b_j}$ представлены в таблице. Требуется составить оптимальный план перевозок ВС по ВЛ, (т.е. найти все значения $\boldsymbol{x_{ij}}$ (млн. ткм)), дающий минимальные транспортные расходы. Начальное решение рекомендуется находить по методу минимальной стоимости, а затем решать задачу с использованием метода потенциалов.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	a_i
1	100	10	6	9	***60***
2	7	100	6	11	***45***
3	8	6	10	12	***55***
b_j	***40***	***50***	***40***	***50***	

Математическая постановка задачи:

Найти минимальное значение целевой функции $z = \sum_{i=1}^{3}\sum_{j=1}^{4} c_{ij}x_{ij}$ и все значения грузооборота x_{ij} (млн. ткм), если заданы ограничения:

$$\begin{cases} \sum_{j=1}^{4} x_{ij} = a_i, \\ \sum_{i=1}^{3} x_{ij} \le b_j, \end{cases} \quad x_{ij} \ge 0, \; i = 1,2,3, \quad j = 1,2,3,4.$$

Решение задачи

Данная задача является несбалансированной (задачей открытого типа), т.к. $\sum_{i=1}^{3} a_i = 60 + 45 + 55 = 160,$ $\sum_{j=1}^{4} b_i = 40 + 50 + 40 + 50 = 180$ и $160 \ne 180$.

Если $\sum_{i=1}^{3} a_i < \sum_{j=1}^{4} b_j,$ то вводится фиктивное ВС, а если $\sum_{i=1}^{3} a_i > \sum_{j=1}^{4} b_j,$, то вводится фиктивная ВЛ. В нашем случае $\sum_{i=1}^{3} a_i < \sum_{j=1}^{4} b_j,$ поэтому введём фиктивное ***ВС 4***. Для этого добавим в таблицу одну строку. Тогда a_4 =180-160=20, а расходы c_{41}, c_{42}, c_{43}, c_{44} по всем ВЛ равны 0. Следует обратить внимание, что в этом случае число ВС равно ***m+1=4*** (***это важно***).

Тип ВС	*ВЛ1*	*ВЛ2*	*ВЛ3*	*ВЛ4*	a_i
1	100	10	6	9	***60***
2	7	100	6	11	***45***
3	8	6	10	12	***55***
4	0	0	0	0	***20***
b_j	***40***	***50***	***40***	***50***	

Найдём начальное решение (опорный план) методом минимальной стоимости, для этого сначала определим число базисных (Чбп) и небазисных переменных (Чнп).

Чбп = Число ВС + Число ВЛ – 1 = (m+1)+n-1= m+n;   Чбп =(3+1)+4-1=7;
Чнп = Число ВС х Число ВЛ – Чбп = (m+1)n-(m+n);   Чнп =16-7=9.

В процессе решения таблицу транспортных расходов и таблицу перевозок будем объединять в одну. Метод заключается в том, что сначала в таблице выбирается ячейка с наименьшими расходами c_{ij}. Далее переменной x_{ij} присваивается наибольшее значение, допускаемое ограничениями на прогноз спроса и потенциал ВС. При получении начального решения нужно следить за балансом базисных и небазисных переменных. Если количество базисных переменных оказалось меньше, чем небазисных, то такой опорный план называется вырожденным. В этом случае одной из переменных $\mathbf{x_{ij}}$ нужно присвоить значение, равное 0.

Также при составлении начального опорного плана для задачи с фиктивным ВС (или ВЛ) и запретными для посадки аэропортами нужно учитывать следующие отличительные особенности, которые не возникают при использовании обычного метода минимальной стоимости.

1. Необходимо выбирать ячейку с наименьшей стоимостью только среди реальных ВС (с ячейками, неравными нулю), а потенциал фиктивного ВС распределять в последнюю очередь.

2. Сначала заполняются строки с «запретными» ячейками (там, где значения 100).

Тип ВС	***ВЛ1***	***ВЛ2***	***ВЛ3***	***ВЛ4***	a_i
1	100	10	6 **40**	9 **20**	***60***
2	7 **40**	100	6	11 **5**	***45***
3	8	6 **50**	10	12 **5**	***55***
4	0	0	0	0 **20**	***20***
b_j	***40***	***50***	***40***	***50***	

В первой строке переменной x_{13} присвоим значение, равное 40, а переменной x_{14} =20. Во второй строке переменная x_{21} =40, а x_{24} =5. Далее заполняем ячейки x_{32} =50, x_{34} =5 и x_{44} =20.

Базисные переменные: x_{13}, x_{14}, x_{21}, x_{24}, x_{32}, x_{34}, x_{44}.

Небазисные переменные: x_{11}, x_{12}, x_{22}, x_{23}, x_{31}, x_{33}, x_{41}, x_{42}, x_{43}.

Соотношение базисных (7) и небазисных (9) переменных выполняется. При этом значения всех небазисных переменных равны 0.

Найдём начальное решение целевой функции z_0:

$$z_0 = 40 \cdot 6 + 20 \cdot 9 + 40 \cdot 7 + 5 \cdot 11 + 50 \cdot 6 + 5 \cdot 12 + 20 \cdot 0 = 1115 \ .$$

Проверим оптимальность полученного плана с помощью метода потенциалов. Всем ВС поставим в соответствие потенциалы $\boldsymbol{u_i}$, а ВЛ – потенциалы $\boldsymbol{v_j}$ (не следует путать метод потенциалов с потенциалом ВС $\boldsymbol{a_i}$).

Потенциалы находим из условий для базисных переменных $\boldsymbol{u_i + v_j = c_{ij}}$. Одному из потенциалов (как правило, это $\boldsymbol{u_1}$) присваиваем произвольное значение, например, 0. Пусть $\boldsymbol{u_1=0}$. Последовательно находим потенциалы по значениям стоимостей ячеек x_{13}, x_{14}, x_{24}, x_{21}, x_{34}, x_{32}, x_{44}:

$u_1 + v_3 = 6$; $v_3 = 6 - 0$; $\boldsymbol{v_3=6}$;

$u_1 + v_4 = 9$; $v_4 = 9 - 0$; $\boldsymbol{v_4=9}$;

$u_2 + v_4 = 11$; $u_2 = 11 - 9$; $\boldsymbol{u_2=2}$;

$u_2 + v_1 = 7$; $v_1 = 7 - 2$; $\boldsymbol{v_1=5}$;

$u_3 + v_4 = 12$; $u_3 = 12 - 9$; $\boldsymbol{u_3=3}$;

$u_3 + v_2 = 6$; $v_2 = 6 - 3$; $\boldsymbol{v_2=3}$;

$u_4 + v_4 = 0$; $u_4 = 0 - 9$; $\boldsymbol{u_4=-9}$.

Для всех небазисных переменных найдём оценки из условия $\boldsymbol{c_{lk}^* = u_l + v_k - c_{lk}}$.

x_{11}: $c_{11}^* = u_1 + v_1 - c_{11} = 0 + 5 - 100 = -95$;

x_{12}: $c_{12}^* = u_1 + v_2 - c_{12} = 0 + 3 - 10 = -7$;

x_{22}: $c_{22}^* = u_2 + v_2 - c_{22} = 2 + 3 - 100 = -95$;

x_{23}: $c_{23}^* = u_2 + v_3 - c_{23} = 2 + 6 - 6 = 2$;

x_{31}: $c_{31}^* = u_3 + v_1 - c_{31} = 3 + 5 - 8 = 0$;

x_{33}: $c_{33}^* = u_3 + v_3 - c_{33} = 3 + 6 - 10 = -1$;

x_{41}: $c_{41}^* = u_4 + v_1 - c_{41} = -9 + 5 - 0 = -4$;

x_{42}: $c_{42}^* = u_4 + v_2 - c_{42} = -9 + 3 - 0 = -6$;

x_{43}: $c_{43}^* = u_4 + v_3 - c_{43} = -9 + 6 - 0 = -3$.

Условие оптимальности выполняется, если все оценки небазисных переменных неположительны. В данной ситуации условие оптимальности не выполнено. В этом случае в базис будет вводиться та небазисная переменная, которая имеет наибольшую положительную оценку. В рассматриваемой задаче такая переменная только одна – переменная x_{23} с оценкой, равной 2, поэтому именно она и будет введена в базис.

На следующем этапе необходимо определить переменную, которая будет выводиться из базиса, иначе соотношение базисных и небазисных переменных будет нарушено. Для этого построим замкнутый цикл, который начинается и заканчивается в ячейке с вводимой в базис переменной x_{23}. Контур состоит из последовательности горизонтальных и вертикальных стрелок любой длины, соединяющих ячейки, соответствующие текущим базисным переменным, и ячейку с вводимой в базис переменной. Вновь вводимой переменной присвоим знак «+», знаки остальных переменных будут чередоваться. В контур включаем максимально возможное число базисных переменных. Получаем контур: $x_{23} \rightarrow x_{13} \rightarrow x_{14} \rightarrow x_{24} \rightarrow x_{23}$.

Тип ВС	*ВЛ1*	*ВЛ2*	*ВЛ3*	*ВЛ4*	a_i
1	100	10	6 **40**	+ 9 **20**	***60***
2	7 **40**	100	+ 6	- 11 **5**	***45***
3	8	6 **50**	10	12 **5**	***55***
4	0	0	0	0 **20**	***20***
b_j	***40***	***50***	***40***	***50***	

По условию допустимости выводимой из базиса переменной является та переменная, которой присвоен знак «-» и которая имеет наименьшее значение. Это переменная x_{24} (значение 5 меньше значения 40 в ячейке x_{13}). Построим новый улучшенный план перевозок. К переменным, которым присвоен знак «+», прибавим число 5, а от переменных, которым присвоен знак «-», отнимем 5.

Тип ВС	*ВЛ1*	*ВЛ2*	*ВЛ3*	*ВЛ4*	a_i
1	100	10	6 **35**	9 **25**	***60***
2	7 **40**	100	6 **5**	11	***45***
3	8	6 **50**	10	12 **5**	***55***
4	0	0	0	0 **20**	***20***
b_j	***40***	***50***	***40***	***50***	

Получили, что новой базисной переменной стала переменная x_{23}, а небазисной – переменная x_{24}, остальные базисные и небазисные переменные не изменились.

Найдём новое базисное решение z_1:

$$z_1 = 35 \cdot 6 + 25 \cdot 9 + 40 \cdot 7 + 5 \cdot 6 + 50 \cdot 6 + 5 \cdot 12 + 20 \cdot 0 = 1105.$$

Решение улучшилось на 10 ден. усл. ед. Проверим его оптимальность.

Найдём значения потенциалов для нового опорного плана и оценки небазисных переменных. Пусть $\boldsymbol{u_1=0}$, тогда

$u_1 + v_3 = 6$; $v_3 = 6 - 0$; $\boldsymbol{v_3=6}$;

$u_1 + v_4 = 9$; $v_4 = 9 - 0$; $\boldsymbol{v_4=9}$;

$u_2 + v_3 = 6$; $u_2 = 6 - 6$; $\boldsymbol{u_2=0}$;

$u_2 + v_1 = 7$; $v_1 = 7 - 0$; $\boldsymbol{v_1=7}$;

$u_3 + v_4 = 12$; $u_3 = 12 - 9$; $\boldsymbol{u_3=3}$;

$u_3 + v_2 = 6$; $v_2 = 6 - 3$; $\boldsymbol{v_2=3}$;

$u_4 + v_4 = 0$; $u_4 = 0 - 9$; $\boldsymbol{u_4=-9}$.

Для всех небазисных переменных найдем оценки из условия $\boldsymbol{c_{lk}^* = u_l + v_k - c_{lk}}$.

x_{11}: $c_{11}^* = u_1 + v_1 - c_{11} = 0 + 7 - 100 = -93$;

x_{12}: $c_{12}^* = u_1 + v_2 - c_{12} = 0 + 3 - 10 = -7$;

x_{22}: $c_{22}^* = u_2 + v_2 - c_{22} = 0 + 3 - 100 = -97$;

x_{24}: $c_{24}^* = u_2 + v_4 - c_{24} = 0 + 9 - 11 = -2$;

x_{31}: $c_{31}^* = u_3 + v_1 - c_{31} = 3 + 7 - 8 = 2$;

x_{33}: $c_{33}^* = u_3 + v_3 - c_{33} = 3 + 6 - 10 = -1$;

x_{41}: $c_{41}^* = u_4 + v_1 - c_{41} = -9 + 7 - 0 = -2$;

x_{42}: $c_{42}^* = u_4 + v_2 - c_{42} = -9 + 3 - 0 = -6$;

x_{43}: $c_{43}^* = u_4 + v_3 - c_{43} = -9 + 6 - 0 = -3$.

Тип ВС	*ВЛ1*	*ВЛ2*	*ВЛ3*	*ВЛ4*	a_i
1	100	10	- 6 **35**	+ 9 **25**	***60***
2	- 7 **40**	100	+ 6 **5**	11	***45***
3	+ 8	6 **50**	10	- 12 **5**	***55***
4	0	0	0	0 **20**	***20***
b_j	***40***	***50***	***40***	***50***	

Условие оптимальности не выполняется. В базис будем вводить переменную x_{31}, которая является единственной, имеющей положительную оценку. Берём контур: $x_{31} \to x_{21} \to x_{23} \to x_{13} \to x_{14} \to x_{34} \to x_{31}$.

Выводимой из базиса переменной является переменная, которой присвоен знак «-» и которая имеет наименьшее значение. Это переменная x_{34} (в ячейке x_{34} стоит знак «-» и значение 5 меньше значений 40 и 35 в ячейках x_{21} и x_{13} соответственно). К переменным, которым присвоен знак «+», прибавим число 5, а от переменных, которым присвоен знак «-», отнимем 5. Получим новый улучшенный план перевозок.

Тип ВС	***ВЛ1***	***ВЛ2***	***ВЛ3***	***ВЛ4***	a_i
1	100	10	6 **30**	9 **30**	***60***
2	7 **35**	100	6 **10**	11	***45***
3	8 **5**	6 **50**	10	12	***55***
4	0	0	0	0 **20**	***20***
b_j	***40***	***50***	***40***	***50***	

Найдём новое базисное решение z_2:

$$z_2 = 30\cdot 6 + 30\cdot 9 + 35\cdot 7 + 10\cdot 6 + 5\cdot 8 + 50\cdot 6 + 20\cdot 0 = 1095.$$

Решение улучшилось на 10 ден. усл. ед. Проверим его оптимальность.

Найдём значения потенциалов u_i и v_j. Пусть $u_1=0$, тогда

$u_1 + v_3 = 6$; $v_3 = 6 - 0$; $v_3=6$;

$u_1 + v_4 = 9$; $v_4 = 9 - 0$; $v_4=9$;

$u_2 + v_3 = 6$; $u_2 = 6 - 6$; $u_2=0$;

$u_2 + v_1 = 7$; $v_1 = 7 - 0$; $v_1=7$;

$u_3 + v_1 = 8$; $u_3 = 8 - 7$; $u_3=1$;

$u_3 + v_2 = 6$; $v_2 = 6 - 1$; $v_2=5$;

$u_4 + v_4 = 0$; $u_4 = 0 - 9$; $u_4=-9$.

Для всех небазисных переменных найдём оценки c_{lk}^*.

x_{11}: $c_{11}^* = u_1 + v_1 - c_{11} = 0 + 7 - 100 = -93$;

x_{12}: $c_{12}^* = u_1 + v_2 - c_{12} = 0 + 5 - 10 = -5$;

x_{22}: $c_{22}^* = u_2 + v_2 - c_{22} = 0 + 5 - 100 = -95$;

x_{24}: $c_{24}^* = u_2 + v_4 - c_{24} = 0 + 9 - 11 = -2$;

x_{33}: $c_{33}^* = u_3 + v_3 - c_{33} = 1 + 6 - 10 = -3$;

x_{34}: $c_{34}^* = u_3 + v_4 - c_{34} = 1 + 9 - 12 = -2$;

x_{41}: $c_{41}^* = u_4 + v_1 - c_{41} = -9 + 7 - 0 = -2$;

x_{42}: $c_{42}^* = u_4 + v_2 - c_{42} = -9 + 5 - 0 = -4$;

x_{43}: $c_{43}^* = u_4 + v_3 - c_{43} = -9 + 6 - 0 = -3$.

Положительных оценок нет, поэтому условие оптимальности выполняется. Значит, найденное выше решение z_2, является оптимальным. Минимальные транспортные расходы составят 1095 ден. усл. ед.

Из последней таблицы делаем следующие выводы. Потенциал ВС 1 предполагается направить по ВЛ 3 и ВЛ 4 в равных количествах по 30 млн. ткм. Потенциал ВС 2 – по ВЛ 1 в количестве 35 млн. ткм и ВЛ 3 в количестве 10 млн. ткм. Потенциал ВС 3 – по ВЛ 1 и ВЛ 2 в количествах, соответственно, 5 млн. ткм и 50 млн. ткм. Отметим, что спрос по грузообороту для ВЛ 4 останется частично неудовлетворённым, будет недопоставлено 20 млн. ткм.

Типовое задание № 3

Постановка задачи. Авиакомпания S выполняет 5 рейсов ВС по 5 ВЛ. Задана матрица себестоимостей каждого рейса по каждой ВЛ. Найти оптимальный план назначений данных рейсов на данные ВЛ с использованием «венгерского» алгоритма. Исходные данные задания представлены в таблице.

Тип ВС	***ВЛ 1***	***ВЛ 2***	***ВЛ 3***	***ВЛ 4***	***ВЛ 5***
1	5	9	14	12	7
2	7	8	13	10	6
3	6	10	14	14	8
4	9	11	12	9	7
5	8	10	15	11	6

Решение задачи

В данной матрице себестоимостей определим в каждой строке минимальный элемент и вычтем его из других элементов строки.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*	*Минимальный элемент строки*
1	5	9	14	12	7	***5***
2	7	8	13	10	6	***6***
3	6	10	14	14	8	***6***
4	9	11	12	9	7	***7***
5	8	10	15	11	6	***6***

Получим следующую матрицу.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	0	4	9	7	2
2	1	2	7	4	0
3	0	4	8	8	2
4	2	4	5	2	0
5	2	4	9	5	0

В полученной матрице в каждом столбце найдём минимальный элемент.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	0	4	9	7	2
2	1	2	7	4	0
3	0	4	8	8	2
4	2	4	5	2	0
5	2	4	9	5	0
Минимальный элемент столбца	***0***	***2***	***5***	***2***	***0***

Вычтем найденные минимальные элементы из других элементов соответствующего столбца. Получим матрицу, у которой в каждой строке и в каждом столбце есть нулевые элементы.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	0	2	4	5	2
2	1	0	2	2	0
3	0	2	3	6	2
4	2	2	0	0	0
5	2	2	4	3	0

Расположение нулевых элементов в последней матрице не позволяет назначить каждому ВС одну ВЛ, так как, для ВС 1 и ВС 3 можно определить только ВЛ 1, что не удовлетворяет условию задачи. Это указывает на то, что оптимальное решение не получено. В этом случае в последней матрице проведём ***минимальное число*** (это важно) горизонтальных и вертикальных прямых по строкам и столбцам таким образом, чтобы вычеркнуть в матрице все нулевые элементы.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	0	2	4	5	2
2	<u>1</u>	0	2	2	<u>0</u>
3	0	2	3	6	2
4	<u>2</u>	2	0	0	<u>0</u>
5	2	2	4	3	0

Найдём ***наименьший*** невычеркнутый (невыделенный) элемент, вычтем его из оставшихся невычеркнутых элементов и прибавим к элементам, находящимся на пересечении проведенных на предыдущем шаге прямых. Наименьший невычеркнутый элемент равен 2. Вычтем его из чисел, расположенных в невыделенных ячейках. К подчеркнутым числам в выделенных ячейках прибавим 2. Остальные выделенные числа оставим без изменений.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	0	0	2	3	2
2	3	0	2	2	2
3	0	0	1	4	2
4	4	2	0	0	2
5	2	0	2	1	0

Полученное распределение не является оптимальным, т.к., например, для ВС 1 и ВС 3 получаются совпадающие назначения на ВЛ 1. Продолжим решение. Снова проведем в последней матрице минимальное число горизонтальных и вертикальных линий по строкам и столбцам, чтобы в матрице все нулевые элементы оказались вычеркнутыми.

Тип ВС	*ВЛ 1*	*ВЛ 2*	*ВЛ 3*	*ВЛ 4*	*ВЛ 5*
1	0	0	2	3	2
2	3	0	2	2	2
3	0	0	1	4	2
4	<u>4</u>	<u>2</u>	0	0	<u>2</u>
5	2	0	2	1	0

Наименьший невыделенный элемент равен 1. Вычтем его из элементов, стоящих в невыделенных ячейках. К подчеркнутым элементам, находящимся на пересечении вычеркнутых (выделенных) строк и столбцов, прибавим число 1. Получим новое решение.

Тип ВС	ВЛ 1	ВЛ 2	ВЛ 3	ВЛ 4	ВЛ 5
1	***0***	0	1	2	2
2	3	***0***	1	1	2
3	0	0	***0***	3	2
4	5	3	0	***0***	3
5	2	0	1	0	***0***

Оптимальное решение найдено. Оно является единственным. В таблице назначения оптимальное решение отмечено выделенными жирным шрифтом нулями. Таким образом, ВС 1 надо назначить на ВЛ 1, ВС 2 – на ВЛ 2, ВС 3 – на ВЛ 3, ВС 4 – на ВЛ 4 и ВС 5 – на ВЛ 5. Суммарная минимальная себестоимость данных назначений равна 5+8+14+9+6=42. Это отражено в последней таблице.

Тип ВС	ВЛ 1	ВЛ 2	ВЛ 3	ВЛ 4	ВЛ 5
1	***5***	9	14	12	7
2	7	***8***	13	10	6
3	6	10	***14***	14	8
4	9	11	12	***9***	7
5	8	10	15	11	***6***

Типовое задание № 4

Постановка задачи. В терминале В аэропорта города N в период $\boldsymbol{T}$ с 12.00 до 13.00 открыто $\boldsymbol{n=6}$ стоек для регистрации пассажиров на 3 ближайших рейса. Пассажиры готовятся совершить вылет на ВС А-321, В-757, В-787. Коэффициент загрузки $\boldsymbol{k}$=0,75. Среднее время регистрации пассажира $\boldsymbol{t_{обс}}$=45 сек. Средние данные о количестве кресел $\boldsymbol{L}$ для определённых типов ВС приведены в таблице № 4А. Определить вероятностно-статистические показатели эффективности обслуживания пассажиров в терминале В аэропорта города N. Для решения задачи использовать модель СМО с ожиданием.

Решение задачи

Будем решать задачу по схеме, изложенной в постановке задания № 4. Сначала определим входные параметры СМО: $\boldsymbol{\lambda, t_{обс}, n}$.

Интенсивность входящего потока $\lambda = \dfrac{L \cdot k}{T}$. T=1 час.;

$\lambda = ((160+200+240)\cdot 0,75):1 = 450$ пасс./час.; $t_{обс} = \frac{45}{3600} = \frac{1}{80}$ час./пасс.; n=6.

1. Вычислим значение приведенной интенсивности ***α***, которая является безразмерной величиной: $\alpha = \lambda \cdot t_{обс}$; $\alpha = 450 \cdot \frac{1}{80} = 5{,}625 \approx 5{,}6.$

При решении задачи с использованием модели СМО с ожиданием необходимо следить за выполнением условия ***n > α*** (это важно). В данной практической ситуации это условие выполнено. Если условие ***n > α*** не выполнено, то очередь нарастает лавинообразно и не может быть решена.

2. Найдём вероятность того, что все стойки регистрации будут свободны:

$$p_0 = \left(\sum_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \cdot \frac{\alpha}{n-\alpha} \right)^{-1};$$

$$p_0 = \left(\frac{5{,}6^0}{0!} + \frac{5{,}6^1}{1!} + \frac{5{,}6^2}{2!} + \frac{5{,}6^3}{3!} + \frac{5{,}6^4}{4!} + \frac{5{,}6^5}{5!} + \frac{5{,}6^6}{6!} + \frac{5{,}6^6}{6!} \cdot \frac{5{,}6}{6-5{,}6} \right)^{-1} \approx 0{,}0013.$$

3. Найдём вероятности состояний данной СМО 1 и 2 групп.

Сначала определим вероятности состояний системы 1 группы (когда нет очереди).

Вероятность того, что в системе k пассажиров и k стоек регистрации заняты обслуживанием: $p_k = \frac{\alpha^k}{k!} \cdot p_0$, где $k = 1,\ldots,n.$

$p_1 = \frac{5{,}6^1}{1!} \cdot 0{,}0013 \approx 0{,}0073;$ $p_2 = \frac{5{,}6^2}{2!} \cdot 0{,}0013 \approx 0{,}0204;$

$p_3 = \frac{5{,}6^3}{3!} \cdot 0{,}0013 \approx 0{,}0381;$ $p_4 = \frac{5{,}6^4}{4!} \cdot 0{,}0013 \approx 0{,}0533;$

$p_5 = \frac{5{,}6^5}{5!} \cdot 0{,}0013 \approx 0{,}0597;$ $p_6 = \frac{5{,}6^6}{6!} \cdot 0{,}0013 \approx 0{,}0557.$

Состояниям СМО 2 группы (когда есть очередь), соответствуют вероятности p_{n+s}, где s – число авиапассажиров в очереди.

$$p_{n+s} = \frac{\alpha^n}{n!} \cdot \left(\frac{\alpha}{n} \right)^s \cdot p_0, \text{ где } s \in N.$$

Найдем значения некоторых вероятностей для состояний системы 2 группы.

Вероятность того, что 6 пассажиров обслуживаются и 1 находится в очереди:

$$p_{6+1} = \frac{5{,}6^6}{6!} \cdot \left(\frac{5{,}6}{6}\right)^1 \cdot 0{,}0013 \approx 0{,}052;$$

Вероятность того, что 6 пассажиров обслуживаются и 2 находятся в очереди:

$$p_{6+2} = \frac{5{,}6^6}{6!} \cdot \left(\frac{5{,}6}{6}\right)^2 \cdot 0{,}0013 \approx 0{,}0485 \text{ и т.д.}$$

Вероятность того, что 6 пассажиров обслуживаются и 6 находятся в очереди:

$$p_{6+6} = \frac{5{,}6^6}{6!} \cdot \left(\frac{5{,}6}{6}\right)^6 \cdot 0{,}0013 \approx 0{,}0368.$$

4. Определим вероятности ожидания и немедленного обслуживания авиапассажиров.

Вероятность ожидания $\boldsymbol{P_{ож}}$ найдём по формуле: $P_{ож} = \frac{\alpha^n}{n!} \cdot \frac{n}{n-\alpha} \cdot p_0$.

$$P_{ож} = \frac{5{,}6^6}{6!} \cdot \frac{6}{6-5{,}6} \cdot 0{,}0013 \approx 0{,}8353.$$

Авиапассажир в любом случае будет обслужен у стойки регистрации, поэтому практический интерес представляет вероятность немедленного обслуживания $\boldsymbol{P_{н.обс.}}$

$$P_{н.обс} = 1 - \frac{\alpha^n}{n!} \cdot \frac{n}{n-\alpha} \cdot p_0 \text{ или } P_{н.обс} = 1 - P_{ож}.$$

Найдем значение этой вероятности: $P_{н.обс} = 1 - 0{,}8353 = 0{,}1647.$

5. Среднее число пассажиров в очереди можно определить по формулам:

$$m_s = P_{ож} \cdot \frac{\alpha}{n-\alpha} \quad \text{или} \quad m_s = \frac{\alpha^n}{n!} \cdot \frac{\alpha n}{(n-\alpha)^2} \cdot p_0.$$

$$m_s = 0{,}8385 \cdot \frac{5{,}6}{6-5{,}6} = 11{,}694 \approx 12 \text{ чел.}$$

6. Вычислим среднее время ожидания регистрации пассажиром в очереди:

$t_s = \frac{m_s \cdot t_{обс}}{n}$; $t_s = \frac{12 \cdot 45}{6} = 90$ сек. или t_s=1,5 мин.

7. Среднее время ожидания регистрации и процедуры регистрации пассажира у стойки $t_{рег} = t_{обс} + t_s = 0,75 + 1,5 = 2,25$ мин = 2 мин 15 сек.
8. Среднее число занятых стоек регистрации $\xi_k = \alpha$. По смыслу задачи оно может быть только целым, поэтому $\xi_k = 5,6 \approx 6$.

Вывод

Вероятность того, что все стойки регистрации окажутся свободными, достаточно мала и составляет 0,0013. Все остальные вероятности состояний системы (состояний $p_k, k = 1, 2, 3, \ldots$) не превосходят 0,06. Вероятности немедленного обслуживания и ожидания равны, соответственно, 0,1647 и 0,8353. Это означает, что, в среднем, из 100 человек 16 пассажирам не придётся стоять в очереди на регистрацию. Длина очереди небольшая и составляет в среднем 12 человек. Время ожидания освобождения стойки регистрации 1,5 минуты, что свидетельствует о достаточно быстрой разгрузке системы. Среднее время, которое придётся затратить пассажиру на всю процедуру регистрации, составит 2 мин 15 сек. Все стойки регистрации функционируют в полном объёме.

Рекомендуемая литература

1. Гармаш А. Н., Орлова И. В., Федосеев В. В. Экономико-математические методы и прикладные модели. Учебник для бакалавриата и магистратуры – М.: Юрайт, 2019.
2. Ниворожкина Л. И., Арженовский С. В. Многомерные статистические методы в экономике. Учебик для вузов – М.: РИОР: ИНФРА-М, 2018.
3. Елисеева И. И., Курышева С. В., Гордеенко Н. М. и др. Практикум по эконометрике. – М.: Финансы и статистика, 2008.
4. Вентцель Е.С., Овчаров Л.А. Исследование операций: задачи, примеры, методология. Учебное пособие. – М.: Дрофа, 2006.
5. Письменный Д.Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам. – М.: Айрис-пресс, 2008.

Приложения

Приложение 1

Таблица значений t-критерия Стьюдента $t_\alpha = t(\alpha, df)$ ***на уровне значимости α и числе степеней свободы df***

α / ***df***	***0,1***	***0,05***	***0,01***	*α* / ***df***	***0,1***	***0,05***	***0,01***
1	6,31	12,71	63,66	***18***	1,734	2,101	2,878
2	2,92	4,30	9,92	***19***	1,729	2,093	2,861
3	2,35	3,18	5,84	***20***	1,725	2,086	2,845
4	2,13	2,78	4,60	***21***	1,721	2,080	2,831
5	2,01	2,57	4,03	***22***	1,717	2,074	2,819
6	1,94	2,45	3,71	***23***	1,714	2,069	2,807
7	1,89	2,37	3,50	***24***	1,711	2,064	2,797
8	1,86	2,31	3,36	***25***	1,708	2,060	2,787
9	1,83	2,26	3,25	***26***	1,706	2,056	2,779
10	1,81	2,23	3,17	***27***	1,703	2,048	2,771
11	1,80	2,20	3,11	***28***	1,701	2,045	2,763
12	1,78	2,18	3,06	***29***	1,699	2,045	2,756
13	1,77	2,16	3,01	***30***	1,697	2,042	2,750
14	1,76	2,15	2,98	***40***	1,684	2,021	2,705
15	1,75	2,13	2,95	***60***	1,671	2,000	2,660
16	1,75	2,12	2,92	***120***	1,658	1,980	2,617
17	1,74	2,11	2,90	∞	1,645	1,960	2,576

Приложение 2

Таблица значений F-критерия Фишера на уровне значимости α= 0,05 и числе степеней свободы df_1 и df_2

Число степеней свободы df_2	***Число степеней свободы df_1***									
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
1	161	200	216	225	230	234	237	239	241	242
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,94	4,88	4,82	4,78	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76
13	4,67	3,81	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28

24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26
25	4,24	3,88	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83