

# ПРОЕКТ

## макета учебной дисциплины (модуля)

«Многомерный статистический анализ», «Multidimensional statistical analysis»

Кафедра \_\_\_\_\_ Математических Методов Анализа Экономики\_. Ауд. 354-362\_, тел. 9392920\_\_\_\_\_

Сайт \_\_ <http://www.econ.msu.ru/departments/mmae/> \_e-mail

**Статус дисциплины:** по выбору, читается на программе бакалавров по направлению «Экономика»

### Авторы программы и преподаватели дисциплины (модуля):

Фамилия Имя Отчество

e-mail [dmitri.artamonov@mail.ru](mailto:dmitri.artamonov@mail.ru) \_\_\_\_\_

Моб. тел. \_\_ **89104452646** \_\_\_\_\_

### 1. Перечень планируемых результатов обучения

*(Соотнесение результатов обучения по дисциплине с планируемыми результатами освоения образовательной программы\*. Указываются общекультурные, общепрофессиональные и профессиональные компетенции, на достижение которых нацелена дисциплина; вид (виды) профессиональной деятельности, к которому готовит дисциплина).<sup>1</sup>*

В результате освоения курса студент должен

- Обладать способностью осуществлять **поиск**, сбор, анализ и обработку данных, необходимых для **эффективного** решения профессиональных задач (ОПК-2);
- способностью выбирать и **комбинировать** инструментальные средства для обработки экономических данных в соответствии с поставленной задачей, анализировать результаты расчетов и обосновывать полученные выводы (ОПК-3);
- способностью выполнять необходимые для составления экономических разделов планов расчеты, обосновывать их и представлять результаты работы в соответствии с принятыми в организации стандартами (ПК-3);
- способностью анализировать и интерпретировать финансовую, бухгалтерскую и иную информацию, содержащуюся в отчетности предприятий различных форм собственности, организаций, ведомств и **других источниках информации**, и использовать полученные сведения для принятия управленческих решений **на микро и макро уровне** (ПК-5);
- способностью анализировать и интерпретировать данные отечественной и зарубежной статистики о социально-экономических процессах и явлениях, выявлять тенденции изменения социально-экономических показателей (ПК-6);
- способностью, используя отечественные и зарубежные источники информации,

\* Планируемые результаты освоения образовательной программы – это компетенции обучающихся, установленные образовательным стандартом, и компетенции обучающихся, установленные организацией дополнительно к компетенциям, установленным образовательным стандартом, с учетом направленности образовательной программы (компетенции указаны в Приложениях №№ 1 и 2 письма к зав. кафедрами).

собирать необходимые данные, анализировать их и готовить информационный обзор и/или аналитический отчет (ПК-7);

- способностью использовать для решения аналитических и исследовательских задач современные технические и программные средства и информационные технологии (ПК-8);
- Способностью организовывать деятельность малой группы, созданной для реализации конкретного экономического проекта (ПК-9);
- способностью использовать для решения коммуникативных задач современные технические средства и информационные технологии (ПК-10);

### **1.1 В результате освоения дисциплины студент должен:**

*Знать\_способы основные способы анализа многомерных статистических данных, их применение к исследованию экономических процессов.* <sup>2</sup>

*Уметь\_\_исследовать данные с помощью корреляционного анализа, применять методы классификации и снижения размерности* <sup>3</sup>

*Владеть методами сбора информации, исследование ее методами перечисленными выше, получение содержательных выводов исходя из проведенного исследования.* <sup>4</sup>

### **2. Место дисциплины в структуре основной образовательной программы (ООП) подготовки бакалавра**

*Дается описание логической и содержательно-методической взаимосвязи с другими частями ООП (дисциплинами, модулями, практиками). Указываются требования к «входным» знаниям, умениям и готовностям обучающегося, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин (модулей).*

*Указываются те теоретические дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее.*

*Указывается желательный семестр ( семестры) \_\_\_\_\_*

*Для освоения данной дисциплины необходимы знания по линейной алгебра, теории вероятностей и математической статистики, эконометрике*

### **3. Желательный объем дисциплины (модуля)**

*(В зачетных единицах. 1 з.е. = 36 академических часов)*

*Общая трудоёмкость 144 часа ( 4 кредита) , в том числе лекции 16 часов, семинары 32 часа, контактные часы 18 часов, самостоятельная работа 78 часа.*

### **4. Структура и содержание дисциплины (модуля)**

*( Указываются темы по разделам, раскрывается их содержание)*<sup>5</sup>

Тема 1. Содержание и назначение многомерного статистического анализа.

Понятие многомерного наблюдения и многомерного признаку, сущность МСА. Основные задачи МСА (исследования зависимостей, классификация объектов, снижение размерности). Основные типы задач, решаемые с помощью МСА (разведочный анализ, визуализация, сжатие больших данных, диагностика финансово-экономического состояния, анализ экспертных оценок.

Тема 2. Способы описание поведения многомерного признака. Многомерные, частные и условные законы распределения, Основные числовые характеристики и их

статистические оценки. Измерение степени тесноты связи между компонентами многомерного статистического распределения (корреляционный анализ многомерной выборки)

Тема 3. Классификация многомерных наблюдений. Общая постановка задачи классификации. Две основные формы задания исходной информации. Основные типы расстояний между объектами и классами объектов, иерархическая классификация. Классификация методом к-средних. Параметрические и непараметрические методы дискриминантного анализа.

Тема 4. Снижение размерности. Общая постановка задачи снижения размерности, основные методы снижения размерности (метод главных компонент, метод экстремальной группировки признаков). Использование метода главных компонент при построении моделей регрессии, при двухшаговом методе наименьших квадратов. Методы факторного анализа. Методы снижения размерности, ориентированные на внешние критерии информативности (экспертно-статистический метод построения сводного латентного показателя эффективности функционирования)

### 5. Перечень учебно-методического обеспечения для самостоятельной работы

*Литература:*<sup>6</sup>

1. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. – М.: Финансы и статистика, 2011
2. Айвазян С.А., Мхитарян В.С. Прикладная статистика в задачах и упражнениях. – М.: ЮНИТИ, 2001
3. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: исследование зависимостей. – М.: Финансы и статистика, 1985.
4. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Классификация и снижение размерности. - М: Финансы и статистика, 1989, 607 с.

### 6. Фонд оценочных средств для проведения промежуточной аттестации по дисциплине

*(Указываются: а) перечень компетенций с указанием этапов их формирования в процессе обучения; б) типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций; в) методические материалы, определяющие процедуры оценивания)*

*а) перечень компетенций с указанием этапов их формирования в процессе обучения*

	ОПК-2	ОПК-3	ПК-3	ПК-5	ПК-6	ПК-7	ПК-8	ПК-9	ПК-10
Исследовательский проект	X	X	X	X	X	X	X	X	X
Задачи на семинарах		X	X				X		X
Контрольная		X							
Домашнее задание		X	X			X	X		X
Экзамен		X							

*б) типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций*

### Образец домашнего задания 1

В файле - различные характеристики банков.

Требуется сделать классификацию

- с помощью иерархической процедуры
- с помощью метода k-средних

В первом случае в отчете поместить дерево, выбранное вами разумное разбиение на кластеры, указать выбор расстояния

Во втором случае указать полученное разбиение на группы, результаты дисперсионного анализа (и прокомментировать их)

### Образец домашнего задания 2.

В файле - различные характеристики авиакомпаний

Требуется сделать классификацию

- с помощью иерархической процедуры
- с помощью метода k-средних

По сделанной классификации осуществить дискриминантный анализ. Привести критерии качества разбиения. Классифицировать 1-ую авиакомпанию.

### Образец контрольной работы.

**Вопросы:** 1. Дискриминантный анализ, функции потерь и вероятности неправильной классификации. Построение оптимальных процедур.

**Задачи:** 1. По эффективности работы выделены две группы, состоящие из  $n_1 = 4$  и  $n_2 = 5$  хозяйств. Для этих групп по двум показателям: производительность труда (тыс. руб./чел.) объем реализованной продукции растениеводства с одного гектара (тыс. руб./га), получены оценки векторов средних

$$\bar{x} = \begin{pmatrix} 5,5 \\ 1,2 \end{pmatrix} \quad \bar{y} = \begin{pmatrix} 3,8 \\ 2,9 \end{pmatrix}$$

и ковариационных матриц

$$S_1 = \begin{pmatrix} 1,1 & 0,2 \\ 0,2 & 0,1 \end{pmatrix} \quad S_2 = \begin{pmatrix} 1,3 & 0,3 \\ 0,3 & ,04 \end{pmatrix}$$

Требуется вычислить значение дискриминантной функции для хозяйства с показателями,  $z_v = (5,9 \quad 0,8)$ . Определить к какому классу относится это наблюдение.

2. По иерархическому агломеративному алгоритму провести классификацию  $n=4$  предприятий по двум показателям:

Номер предприятия	1	2	3	4
$x_{i1}$	3	10	9	2
$x_{i2}$	2	6	5	1

В качестве расстояния между объектами принято обычное Евклидово расстояние, а расстояние между кластерами определить по принципу дальнего соседа.

### Образец экзамена

**Вопросы:** 1. Параметрический дискриминантный анализ в случае нормального распределения внутри классов. Число классов  $p=2$ .

2. Расщепление смесей вероятностных распределений и их идентифицируемость

- Даны векторы  $\bar{X}$  и  $\bar{Y}$ , матрицы  $S_x$  и  $S_y$ , число объектов в выборках  $n_x$  и  $n_y$ , одно классифицируемое наблюдение  $Z$ . Требуется определить, к какому из множеств  $(X, Y)$  следует отнести объект  $Z$ .
- Дана матрица коэффициентов парной корреляции  $R$  размерности  $3 \times 3$ . Вычислить один из частных коэффициентов корреляции, проверить его значимость и построить интервальную оценку

### 7. Форма проведения самостоятельной работы и текущего контроля успеваемости

*Указываются обязательные формы (например, изучение материалов учебника, лекции) и по выбору студента (например, написание реферата, подготовка доклада, презентации, коллективного или индивидуального проекта, участие в конференции, подготовка кейса и пр.) формы самостоятельной работы и способы проверки, трудоемкость выполнения самостоятельной работы.*

Изучение материалов учебника (5 часов), лекции (15 часов), выполнение домашних заданий (15 часов), написание группового проекта (20 часов). Контрольная работа (5 часов), Экзамен (25 часов)

### 8. Образовательные технологии

---

*(Указываются образовательные технологии, используемые при реализации различных видов учебной работы: компьютерные симуляции, деловые и ролевые игры, разбор конкретных ситуаций, психологические и иные тренинги в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся. В рамках учебных курсов должны быть предусмотрены встречи с представителями российских и зарубежных компаний, государственных и общественных организаций, мастер-классы экспертов и специалистов.*

Лекции, семинарские занятия в компьютерных классах, анализ данных с использованием статистических пакетов SPSS, STATISTICA, R

9. **Балльная система оценки знаний** (Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкалы оценивания с учетом, принятой на факультете балльно-рейтинговой системой)

- *Домашнее задание 1.*
- *Домашнее задание 2.*
- *Проект*
- *Контрольная работа*
- *Экзамен*

Критерии оценок: Отлично — не менее 80%, хорошо — не менее 60%, удовлетворительно — не менее 40%.

**10. Материально-техническое обеспечение дисциплины (модуля) \_\_ программы \_\_  
SPSS, STATISTICA,  
R**

*Указывается материально-техническое обеспечение данной дисциплины (модуля), например, программное обеспечение, специализированные аудитории и пр.*

<sup>1</sup> *Планируемые результаты обучения по дисциплине* – это знания, умения, навыки и (или) опыт деятельности, характеризующие этапы формирования компетенций и обеспечивающие достижение планируемых результатов освоения образовательной программы. (Виды профессиональной деятельности и компетенции указаны в Приложениях №№1 и 2 письма к зав.кафедрами)

<sup>2</sup> *Знать* – воспроизводить и объяснять учебный материал с требуемой степенью научной точности и полноты.

<sup>3</sup> *Уметь* - решать типичные задачи на основе воспроизведения стандартных алгоритмов решения;

<sup>4</sup> *Владеть* - решать усложненные задачи на основе приобретенных знаний, умений и навыков, с их применением в нетипичных ситуациях. «Иметь навык» – многократно применять «умение», довести «умение» до автоматизма

<sup>5</sup> Содержание должно вытекать из параметров, указанных в п. 1, не более 3000 знаков с пробелами

<sup>6</sup> Не более 5 источников