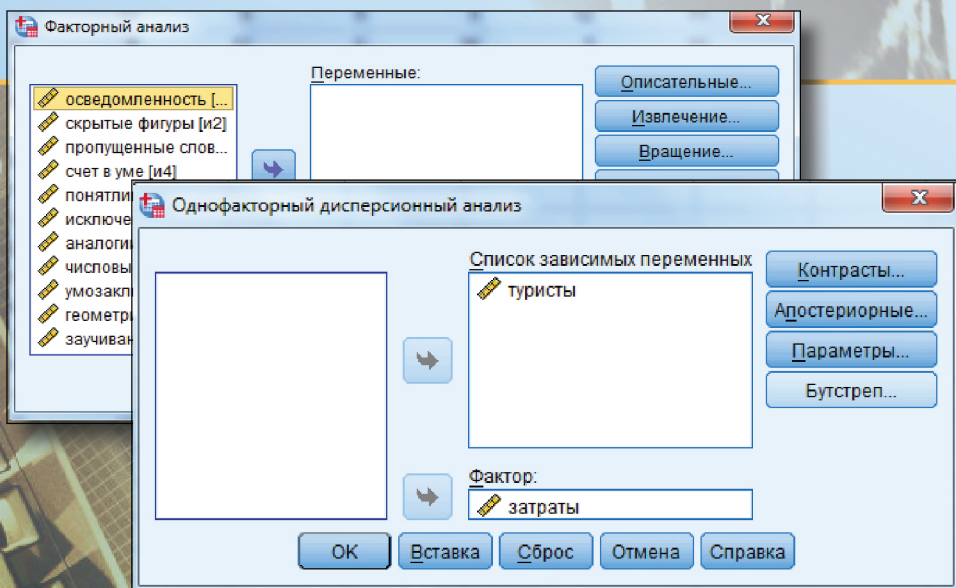


Министерство образования и науки Российской Федерации  
Тольяттинский государственный университет  
Гуманитарно-педагогический институт  
Кафедра «Теоретическая и прикладная психология»

И.В. Кулагина

# СТАТИСТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

Практикум



УДК 159.9:519.2

ББК 88в631.8

Рецензенты:

академик РАН, д-р психол. наук, профессор Тольяттинского филиала Самарской гуманитарной академии *С.В. Быков*;  
д-р психол. наук, профессор Тольяттинского государственного университета *Г.А. Виноградова*.

Кулагина, И.В. Статистические методы в психологии : практикум / И.В. Кулагина. – Тольятти : Изд-во ТГУ, 2017. – 1 оптический диск.

В практикуме раскрываются основы применения статистических методов в ходе реализации психологического исследования. Описание назначения и особенностей применения каждого метода сопровождается примерами и алгоритмами вычисления с использованием программного обеспечения. К каждой теме представлены задания и упражнения для самостоятельной работы.

Предназначен для самостоятельной и аудиторной работы студентов направления подготовки 37.04.01 «Психология», магистерские программы «Организационная психология», «Психология личности», «Психология здоровья», очной и заочной форм обучения.

Текстовое электронное издание.

Рекомендовано к изданию научно-методическим советом Тольяттинского государственного университета.

Минимальные системные требования: IBM PC-совместимый компьютер: Windows XP/Vista/7/8; PIII 500 МГц или эквивалент; 128 Мб ОЗУ; SVGA; CD-ROM; Adobe Acrobat Reader.

Редактор *Г.В. Данилова*  
Технический редактор *Н.П. Крюкова*  
Компьютерная верстка: *Л.В. Сызганцева*  
Художественное оформление,  
компьютерное проектирование: *И.И. Шишкина*

Дата подписания к использованию 10.10.2017.

Объем издания 10,3 Мб.

Комплектация издания:

компакт-диск, первичная упаковка.

Заказ № 1-76-16.

Издательство Тольяттинского государственного университета  
445020, г. Тольятти, ул. Белорусская, 14,  
тел. 8 (8482) 53-91-47, [www.tltsu.ru](http://www.tltsu.ru)

## Содержание

ВВЕДЕНИЕ .....	5
Практическое занятие 1. Параметрические и непараметрические методы сравнения выборок .....	8
Практическое занятие 2. Корреляционный анализ данных. Коэффициенты корреляции .....	23
Практическое занятие 3. Дисперсионный анализ данных .....	32
Практическое занятие 4. Регрессионный анализ данных .....	44
Практическое занятие 5. Факторный анализ данных .....	53
Практическое занятие 6. Кластерный анализ данных .....	67
Итоговый тест .....	81
БИБЛИОГРАФИЧЕСКИЙ СПИСОК .....	85
ГЛОССАРИЙ .....	88

## ВВЕДЕНИЕ

Многие психологические феномены не подлежат непосредственному измерению, что обуславливает необходимость использования математико-статистического аппарата для системного описания выделенных исследователем явлений и отношений между наблюдаемыми переменными. Современная интенсивная компьютеризация общественной жизни вносит весьма существенные коррективы в процесс и результат проведения психологических научных исследований. Востребованность компьютерной статистической обработки массива психологических эмпирических данных растет. Одним из самых распространенных статистических пакетов, используемых в исследовательской практике, является программа SPSS, имеющая на данный момент русифицированную версию и позволяющая решать основной массив психологических задач. Успешность интерпретации полученных эмпирических данных во многом зависит от уровня теоретической подготовленности исследователя, степени овладения современными статистическими пакетами и грамотности в области перевода количественных показателей на «психологический язык».

При освоении дисциплины «Статистические методы в психологии» магистрант овладевает компетенциями, знаниями, умениями и навыками, необходимыми для осуществления основных статистических процедур решения разнообразных психологических задач теоретического и прикладного характера. Профессиональное становление и психолога-исследователя, и психолога-практика предполагает не только овладение арсеналом измерительных процедур, но и способность выбирать, а в случае необходимости модифицировать стандартную измерительную процедуру согласно решаемой задаче.

Практические задания составлены с упором на приобретение умений и навыков в решении наиболее распространенных задач психологических исследований с применением статистических методов.

### *Цель и задачи изучения дисциплины*

**Цель** – формирование у магистрантов фундаментальных знаний о возможностях применения статистических методов в прикладных психологических исследованиях и профессиональной деятельности психолога.

### **Задачи:**

- формирование представления об основных статистических методах и особенностях их применения в психологических исследованиях;
- ознакомление с возможностями обработки эмпирического материала посредством функционала программы SPSS;
- ознакомление с алгоритмом осуществления статистической обработки и интерпретации данных экспериментальных исследований;
- формирование навыков правильного написания заключения на основе результатов статистического анализа.

В результате изучения дисциплины студент должен *знать* основные статистические и математические методы обработки данных, полученных при решении основных профессиональных задач; *уметь* правильно планировать исследование; обрабатывать и интерпретировать данные исследований с помощью математико-статистического аппарата; *владеть* навыками решения типовых задач профессиональной деятельности и уверенно пользоваться пакетом программ IBM SPSS.

### **Методические рекомендации по работе с практикумом:**

- ознакомиться с интерфейсом программы IBM SPSS Statistics и ее возможностями, используя приведенные литературные источники и интернет-ресурсы (в зависимости от версии IBM SPSS названия вкладок и окон могут различаться);
  - самостоятельно изучить учебные вопросы темы в научной литературе и методических рекомендациях по освоению теоретического материала к теме;
  - ознакомиться с практическими заданиями к каждой теме;
  - на основе методических рекомендаций и образцов выполнения заданий самостоятельно осуществить необходимые вычисления в программе IBM SPSS. Проанализировать полученные результаты и сделать качественный вывод на основе полученных количественных данных;
  - ответить на вопросы для самоконтроля.

### *Критерии оценки заданий и учебных мероприятий*

Практические задания оцениваются дифференцированно в зависимости от уровня сложности и количества задач по теме.

Практическое задание (занятия 1–2):

2 балла — задача решена правильно, сделан грамотный качественный вывод;

1 балл — задача решена правильно, но качественный вывод отсутствует;

0 баллов — задача не решена.

Практическое задание (занятия 3–6):

4 балла — задача решена правильно, сделан грамотный качественный вывод;

2 балла — задача решена правильно, но качественный вывод отсутствует;

0 баллов — задача не решена.

В результате освоения дисциплины «Статистические методы в психологии» студент может набрать максимум 64 балла, что эквивалентно отметке «зачтено». Чтобы добрать до 100 баллов и проверить уровень освоения теоретических основ математической статистики, магистрант может пройти тестирование (примерные тестовые вопросы приведены в заключительной части практикума).

Итоговый тест — 46 баллов (по количеству правильных ответов).

## **Практическое занятие 1**

### **Параметрические и непараметрические методы сравнения выборок**

**Цель** – ознакомление с алгоритмами применения основных параметрических и непараметрических методов сравнения эмпирических выборок.

#### **Учебные вопросы:**

- параметрические методы сравнения выборок;
- непараметрические методы сравнения выборок;
- $t$ -критерий Стьюдента для независимых выборок;
- $t$ -критерий Стьюдента для зависимых выборок;
- критерий Колмогорова – Смирнова;
- $U$ -критерий Манна – Уитни;
- $T$ -критерий Вилкоксона.

Студент должен *знать* различия в назначении и особенностях применения параметрических и непараметрических методов сравнения выборок; *уметь* выбирать метод сравнения выборок в соответствии с поставленной целью и особенностями выборки; *владеть* навыками выполнения расчетов с применением параметрических и непараметрических методов сравнения выборок.

#### **Методические рекомендации по освоению теоретического материала по теме**

К методам сравнения выборок наиболее часто относят способы проверки статистических гипотез о различии выборок по уровню выраженности признака, измеренного в количественной шкале. Сравнение двух выборок по признаку, измеренному в метрической шкале, обычно предполагает сравнение средних значений с использованием параметрического  $t$ -критерия Стьюдента. Применение критерия допустимо в отношении зависимых и независимых выборок, а также при сравнении эмпирического среднего значения с заданной величиной. При сравнении средних двух выборок проверяется ненаправленная статистическая гипотеза о равенстве средних в генеральной совокупности. При ее отклонении допустимо приня-



тие двусторонней альтернативы о конкретном направлении различий в соответствии с соотношением выборочных средних.

Непараметрические методы сравнения выборок зачастую являются аналогами параметрических методов сравнения средних значений, но значительно проще в вычислительном отношении. Параметрические методы сравнения средних значений обладают изначально большей чувствительностью к обнаружению различий, вследствие чего являются более предпочтительными, чем их непараметрические аналоги. Решение о применении непараметрических методов становится оправданным в случае, когда не выполняются исходные предположения, лежащие в основе применения параметрических методов, а именно:

- есть основания считать, что распределение значений признака в генеральной совокупности не соответствует нормальному закону;
- объем выборки слишком мал для того, чтобы судить о распределении признака в генеральной совокупности;
- не выполняется требование гомогенности дисперсии при сравнении средних значений для независимых выборок.

Структура исходных данных и интерпретация результатов применения для параметрических методов и их непараметрических аналогов являются идентичными. При сравнении выборок с использованием непараметрических критериев также проверяются ненаправленные статистические гипотезы. Нулевая гипотеза содержит утверждение об идентичности генеральных совокупностей (из которых извлечены выборки) по уровню выраженности изучаемого признака. При ее отклонении допустимо принятие двусторонней альтернативы о конкретном направлении различий в соответствии с выборочными данными [17, с. 162–169, 172–178].

## Практические задания

### Задача 1

Проведен замер тревожности у учащихся двух профильных классов: гуманитарного (выборка А) и физико-математического (выборка В). Имеются ли достоверные различия в средних уровнях тревожности учащихся этих классов?

#### *Показатели уровня тревожности*

№	А	В	№	А	В	№	А	В
1	55	35	3	60	68	5	65	45
2	40	42	4	47	40	6	30	50

### Задача 2

По методике Филлипса проведено исследование уровня школьной тревожности у школьников подросткового возраста: мальчиков (выборка А) и девочек (выборка В). Можно ли утверждать, что принадлежность к определенному полу влияет на уровень тревожности?

#### *Показатели уровня школьной тревожности*

№	А	В	№	А	В	№	А	В
1	6	4	4	5	6	7	10	9
2	3	7	5	8	7	8	3	3
3	8	5	6	9	8	9	5	10
						10	6	3

### Задача 3

Подростки двух девярых классов (выборки А и В) средней школы г. Тольятти были обследованы на склонность к асоциальному поведению (нарушению норм и правил). Можно ли утверждать, что уровни склонности к асоциальному поведению в этих классах достоверно различаются?

#### *Показатели склонности к асоциальному поведению*

№	А	В	№	А	В	№	А	В
1	48	50	3	72	57	5	50	58
2	54	56	4	49	54	6	48	53

#### Задача 4

В ходе исследования измерен уровень вербального интеллекта у респондентов возраста 16–17 лет: юношей (выборка А) и девушек (выборка В). Можно ли утверждать, что одна группа испытуемых превосходит другую по уровню вербального интеллекта?

#### *Показатели вербального интеллекта*

№	А	В	№	А	В	№	А	В
1	112	117	7	107	100	13	114	105
2	115	105	8	90	107	14	114	115
3	102	114	9	122	115	15	118	109
4	108	108	10	118	98	16	99	–
5	117	99	11	100	117	17	102	–
6	98	120	12	92	114			

#### Задача 5

У студентов вуза с помощью субшкалы методики К. Томаса выявлен уровень выраженности стиля конфликтного поведения «Соперничество», а с помощью методики Д. Роттера – тип локуса контроля. В таблице приведены данные об уровне соперничества у представителей двух типов локуса контроля: экстерналов (выборка А) и интерналов (выборка В). Можно ли утверждать, что у студентов-экстерналов уровень соперничества выше, чем у студентов с преобладанием интернального локуса контроля?

#### *Показатели уровня соперничества для испытуемых с разными типами локуса контроля*

№	А	В	№	А	В	№	А	В
1	3	7	8	9	9	15	7	4
2	7	3	9	5	9	16	2	4
3	3	10	10	6	9	17	0	3
4	10	3	11	5	8	18	1	5
5	9	6	12	6	4	19	3	4
6	7	5	13	3	10	20	7	4
7	8	3	14	3	3	21	4	2

### Задача 6

У учеников, занимающихся по системе развивающего обучения Эльконина – Давыдова, определяли уровень школьной мотивации: сначала – в первом классе (выборка А), затем – в третьем классе (выборка В). Можно ли утверждать, что система развивающего обучения способствует развитию школьной мотивации?

#### *Показатели школьной мотивации*

№	А	В	№	А	В	№	А	В
1	18	13	8	21	20	15	12	16
2	20	25	9	19	24	16	21	25
3	21	26	10	11	20	17	28	9
4	9	18	11	18	24	18	14	20
5	13	21	12	28	28	19	13	18
6	17	26	13	14	18	20	19	22
7	15	9	14	20	23	21	23	28
						22	25	28

### Задача 7

По заказу торговой фирмы проведен тренинг навыков продаж в группе продавцов-консультантов. У участников тренинговой группы измеряли уровень владения навыками продаж: до тренинга (выборка А) и после тренинга (выборка В). Наблюдается ли после тренинга достоверный сдвиг в сторону повышения уровня владения навыками продаж?

#### *Показатели уровня владения навыками продаж*

№	А	В	№	А	В	№	А	В
1	5	8	6	5	6	11	7	9
2	4	9	7	4	8	12	6	6
3	7	10	8	3	9	13	9	9
4	3	8	9	2	5	14	7	10
5	6	10	10	5	8	15	8	10

### Задача 8

У сотрудников вневедомственной охраны с помощью методики Холмса и Раге измерена степень подверженности стрессу до (выборка А) и после (выборка В) тренинга по повышению стрессоустойчивости. Снизилась ли подверженность стрессу после тренинга?

#### *Показатели подверженности стрессу*

№	А	В	№	А	В	№	А	В
1	176	161	8	154	157	15	248	165
2	230	226	9	189	163	16	216	172
3	197	200	10	207	199	17	177	172
4	279	229	11	241	220	18	318	198
5	301	302	12	266	188	19	304	215
6	320	290	13	314	210	20	226	183
7	211	184	14	210	227			

#### **Методические рекомендации и образцы выполнения заданий. Алгоритм применения параметрических и непараметрических методов сравнения выборок**

1. Ознакомьтесь с условиями задачи. Необходимо проверить, соответствуют ли эмпирические данные нормальному закону распределения, т. е. являются ли метрическими. Для этого сформулируем гипотезы:  $H_0$  – об отсутствии различий в распределении эмпирических данных и нормальном законе распределения;  $H_1$  – о наличии таких различий.

Для выполнения **проверки на нормальность** используем пакет SPSS. Применим первый и второй способы проверки на нормальность – с помощью визуальной оценки гистограммы и с помощью оценок коэффициентов асимметрии и эксцесса.

- В меню *Анализ* выберите *описательные статистики*, далее *частоты* и щелкните в нем на кнопке *Частоты: Статистические показатели*.

- Выберите переменную, выделите ее, затем щелкните на кнопке со стрелкой, чтобы перенести ее в *список тестируемых переменных*.

- Щелкните на кнопке *Статистические показатели*, чтобы открыть диалоговое окно *Частоты: Статистические показатели*.

- В группе *Показатели центральной тенденции* установите флажок *Среднее*, в группе *Распределение* – флажки *Асимметрия* и *Экссесс*, в группе *Изменчивость* – флажок *Стандартное отклонение* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *частоты*.

- В окне *частоты* выберите переменную *графики* – *гистограмма с кривой нормального распределения* – *продолжить*, в окне *частоты* нажмите *ОК*.

Распределение соответствует нормальному виду, если для соответствующей переменной абсолютные значения асимметрии и эксцесса не превышают свои стандартные ошибки. Чем ближе точки графиков к прямой линии, тем меньше отличие распределения от нормального вида.

Применим третий способ проверки на нормальность – критерий Колмогорова – Смирнова.

- В окне *редактор данных* в меню *Анализ* выберите команду *непараметрические критерии* – *одновыборочный критерий Колмогорова – Смирнова*, чтобы открыть диалоговое окно.

- Выберите переменные в список тестируемых переменных (щелкаете на переменной, выделив ее, а затем – на кнопке со стрелкой, чтобы переместить переменную в *список тестируемых переменных*).

- Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

Отклонение от нормального распределения считается существенным при значении  $p < 0,05$ ; в этом случае для соответствующих переменных следует применять непараметрические тесты. В обратном случае вероятность ошибки является незначимой, следовательно, значения переменной достаточно хорошо подчиняются нормальному распределению.

2. Определитесь с методом сравнения выборок. Сформулируйте статистические гипотезы.

3. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

В случае применения **параметрических методов сравнения выборок** с помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, таким образом, чтобы первый столбец содержал наблюдения для обеих выборок, размещенные один под другим. Второй столбец – новый, содержит 1 в клетках рядом с наблюдениями первой выборки и 2 рядом с наблюдениями второй выборки.

В меню *Анализ* выберите команду *Сравнение средних – Т-критерий для независимых выборок* (если он соответствует сформулированным гипотезам). В этом диалоговом окне вы можете задать параметры для применения *t*-критерия.

Щелкните сначала на переменной 1, чтобы выделить ее, а затем – на верхней кнопке со стрелкой, чтобы переместить переменную в список *Проверяемые переменные*.

Щелкните на переменной 2, чтобы выделить ее, а затем – на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Группировать по:*.

Щелкните на кнопке *Задать группы*, чтобы открыть диалоговое окно. Введите число 1 в поле *Группа 1*, введите число 2 в поле *Группа 2* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Т-критерий для независимых выборок*.

Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

Рассмотрим применение ***t*-критерия для парных (зависимых) выборок**.

В меню *Анализ* выберите команду *Сравнение средних – Т-критерий для парных выборок*. На экране появится диалоговое окно *Т-критерий для парных выборок*. Процедура управляется единственным диалоговым окном и не требует указания группирующей переменной. В левой части окна находится уже знакомый список доступных переменных файла, в котором выбираются переменные для применения *t*-критерия.

Щелкните на первой переменной, чтобы выделить ее, затем щелчком по стрелке перенесите ее в правое окно. Ее имя появится в области *Парные переменные* в строке *Пара: 1* рядом с меткой *Переменная 1*.

Щелкните на второй переменной, чтобы выделить ее, затем щелчком по стрелке перенесите ее в правое окно. Ее имя появится в области *Парные переменные* в строке *Пара: 1* рядом с меткой *Переменная 2*.

Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

В случае применения **непараметрических методов сравнения выборок** с помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, таким образом, чтобы первый столбик содержал наблюдения для обеих выборок, размещенные один под другим. Второй столбец – новый, содержит 1 в клетках рядом с наблюдениями первой выборки и 2 рядом с наблюдениями второй выборки.

**Критерий Манна – Уитни**, или ***U*-критерий**, по назначению аналогичен *t*-критерию для независимых выборок. Разница заключается в том, что *t*-критерии ориентированы на нормальные и близкие к ним распределения, а критерий Манна – Уитни – на распределения, отличные от нормальных.

В меню *Анализ* выберите команду *Непараметрические критерии – Устаревшие диалоговые окна – Для двух независимых выборок*, чтобы открыть диалоговое окно.

Щелкните сначала на первой переменной, чтобы выделить ее, а затем – на верхней кнопке со стрелкой, чтобы переместить переменную в *Список проверяемых переменных*.

Щелкните на второй переменной, чтобы выделить ее, а затем – на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Группирующая переменная*.

Щелкните на кнопке *Задать группы*, чтобы открыть диалоговое окно, в поле *Группа 1* введите значение 1, в поле *Группа 2* – значение 2 и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Критерий для двух независимых выборок*.

Отметьте флажком *Тип теста – Манна – Уитни U*.

Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

***T*-критерий Уилкоксона (Вилкоксона)** для связанных пар позволяет сравнить два измерения переменной на одной выборке (например, «до» и «после») по уровню ее выраженности и аналогичен *t*-критерию Стьюдента для зависимых выборок.



В меню *Анализ* выберите команду *Непараметрические критерии – Устаревшие диалоговые окна – Для двух независимых выборок*, чтобы открыть диалоговое окно *Критерии для двух связанных выборок*. Если вы уже успели поработать с этим окном, щелкните на кнопке *Сброс*.

Щелкните на первой переменной, чтобы выделить ее, затем щелчком по стрелке перенесите ее в правое окно. Ее имя появится в области *Парные переменные* в строке *Пара: 1* рядом с меткой *Переменная 1*.

Щелкните на второй переменной, чтобы выделить ее, затем щелчком по стрелке перенесите ее в правое окно. Ее имя появится в области *Парные переменные* в строке *Пара: 1* рядом с меткой *Переменная 2*.

Отметьте флажком *Тип теста – Вилкоксон W*.

Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

### Пример

**Условие задачи.** 20 школьникам были даны тесты на наглядно-образное и вербальное мышление. Исследователями фиксировалось время решения задач в секундах. Необходимо проверить, есть ли различия во времени решения наглядно-образных и вербальных задач у учащихся разного пола и школьников, обучающихся в разных профильных классах.

### Эмпирические данные

№ п/п	Время решения наглядно-образных заданий	Время решения вербальных заданий	Пол	Профиль класса
1	19	17	М	Ф-м
2	32	7	М	Ф-м
3	33	17	М	Ф-м
4	44	28	М	Гум-й
5	28	27	М	Гум-й
6	35	31	М	Гум-й
7	39	20	М	Ф-м
8	39	17	М	Ф-м
9	44	35	М	Ф-м
10	44	43	М	Гум-й
11	24	10	Ж	Гум-й
12	37	28	Ж	Гум-й

№ п/п	Время решения наглядно-образных заданий	Время решения вербальных заданий	Пол	Профиль класса
13	29	13	Ж	Ф-м
14	40	43	Ж	Ф-м
15	42	45	Ж	Ф-м
16	32	24	Ж	Ф-м
17	48	45	Ж	Гум-й
18	42	26	Ж	Гум-й
19	33	16	Ж	Гум-й
20	47	26	Ж	Гум-й

### Алгоритм решения задачи-примера

Проверка распределений переменных на нормальность выявила их соответствие нормальному закону, следовательно, в сравнительном анализе будем использовать параметрический  $t$ -критерий Стьюдента.

А. Сравниваем время решения наглядно-образных задач и время решения вербальных задач у учащихся разного пола (будем использовать  $t$ -критерий для независимых выборок).

1. В меню *Анализ* выберите команду *Сравнение средних – T-критерий для независимых выборок*. На экране появится диалоговое окно *T-критерий для независимых выборок* (рис. 1).

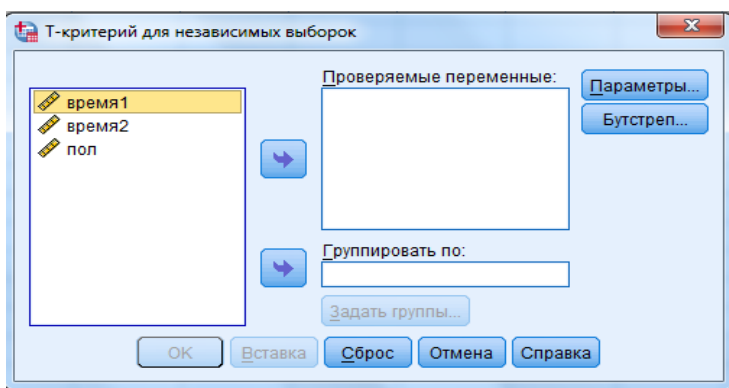


Рис. 1. Диалоговое окно  $T$ -критерий для независимых выборок

2. Выделите переменные *Время решения наглядно-образных задач* (время1) и *Время решения вербальных задач* (время2) и переместите их в список *Тестируемые переменные*.

3. Щелкните сначала на переменной *пол*, чтобы выделить ее, а затем – на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Группирующая переменная* (рис. 2).

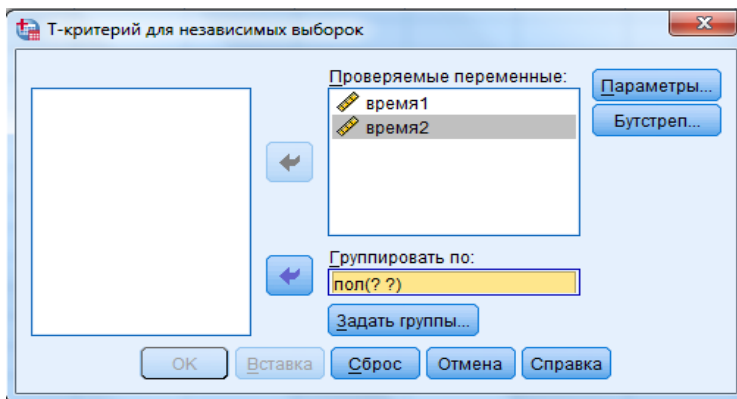


Рис. 2. Диалоговое окно *T*-критерий для независимых выборок

4. Щелкните на кнопке *Определение групп*, чтобы открыть одноименное диалоговое окно. Введите число 1 в поле *Группа 1*, введите число 2 в поле *Группа 2* (рис. 3).

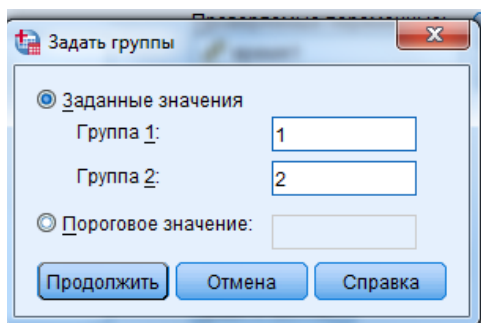


Рис. 3. Диалоговое окно *T*-критерий для независимых выборок – Задать группы

5. Щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *T*-критерий для независимых выборок. Щелкните на кнопке *OK*, чтобы открыть окно вывода.

**Результаты.** Во второй таблице *Критерий для независимых выборок* на выходном листе программы SPSS нам интересно значение критерия равенства дисперсий Левина ( $F = 0,003$  при  $p = 0,959$ , т. е.  $p > 0,05$ ) (рис. 4), оно указывает на то, что дисперсии двух распределений переменной *время решения наглядно-образных задач* (по признаку пола) статистически значимо не различаются, следовательно, можем интерпретировать результаты работы *t*-теста.

Критерий для независимых выборок										
	Критерий равенства дисперсий Левина		t-критерий равенства средних							
	F	Знач.	t	ст. св.	Значимость (2-сторонняя)	Разность средних	Стд. ошибка разности	95% доверительный интервал разности средних		
								Нижняя граница	Верхняя граница	
время1	Предполагается равенство дисперсий	,003	,959	-478	18	,639	-1,70000	3,55981		
	Равенство дисперсий не предполагается			-478	17,982	,639	-1,70000	3,55981	-9,17888	5,77888
время2	Предполагается равенство дисперсий	,273	,608	-642	18	,529	-3,4		-14,53026	7,73026
	Равенство дисперсий не предполагается			-642	17,293	,529	-3,4	Дважды щелкните, чтобы активировать	-14,56297	7,76297

Рис. 4. Таблица Критерий для независимых выборок выходного листа

По показателям *t*-критерия Стьюдента (рис. 4) для независимых выборок не обнаружено статистически значимых различий во времени решения наглядно-образных задач среди учащихся разного пола ( $t = -0,478$  при  $p = 0,639$ , т. е.  $p > 0,05$ ). Критерий равенства дисперсий Левина ( $F = 0,273$  при  $p = 0,608 > 0,05$ ) указывает на то, что дисперсии двух распределений переменной *время решения вербальных задач* статистически значимо не различаются, следовательно, можно интерпретировать результаты работы *t*-теста. По *t*-критерию Стьюдента для независимых выборок не обнаружено статистически значимых различий во времени решения вербальных задач среди учащихся разного пола ( $t = -0,642$  при  $p = 0,529$ , т. е.  $p > 0,05$ ).

**Б.** Сравниваем время решения наглядно-образных задач и время решения вербальных задач у учащихся разных профильных классов (будем использовать *t*-критерий для независимых выборок). Алгоритм решения используем тот же, с небольшими изменениями.

1. В меню *Анализ* выберите команду *Сравнение средних – T-критерий для независимых выборок*. На экране появится диалоговое окно

*T*-критерий для независимых выборок. Если окно содержит ранее введенные данные, нажмите на кнопке *Сброс* (рис. 1).

2. Выделите переменные *Время решения наглядно-образных задач* и *Время решения вербальных задач* и переместите их в список *Тестируемые переменные*.

3 Щелкните сначала на переменной *Профиль класса*, чтобы выделить ее, а затем на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Группирующая переменная* (рис. 2).

4. Щелкните на кнопке *Определение групп*, чтобы открыть одноименное диалоговое окно. Введите число 1 в поле *Группа 1*, введите число 2 в поле *Группа 2* (рис. 3) и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *T-критерий для независимых выборок*.

5. Щелкните на кнопке *OK*, чтобы открыть окно вывода.

**Результаты.** Критерий равенства дисперсий Левина (рис. 5) ( $F = 0,251$  при  $p = 0,622$ , т. е.  $p > 0,05$ ) указывает на то, что дисперсии двух распределений переменной *время решения наглядно-образных задач* (для классов с разным профилем) статистически значимо не различаются, следовательно, можно интерпретировать результаты работы *t*-теста.

**Критерий для независимых выборок**

		Критерий равенства дисперсий Левина		t-критерий равенства средних						
		F	Знач.	t	ст. св.	Значимость (2-сторонняя)	Разность средних	Стд. ошибка разности	95% доверительный интервал разности средних	
									Нижняя граница	Верхняя граница
время1	Предполагается равенство дисперсий	.251	.622	-.944	18	.358	-3,30000	3,49682	-10,64655	4,04655
	Равенство дисперсий не предполагается			-.944	17,856	.358	-3,30000	3,49682	-10,65079	4,05079
время2	Предполагается равенство дисперсий	1,090	.310	-.798	18	.436	-4,20000	5,26582	-15,26308	6,86308
	Равенство дисперсий не предполагается			-.798	17,317	.436	-4,20000	5,26582	-15,29444	6,89444

Рис. 5. Таблица Критерий для независимых выборок выходного листа

По *t*-критерию Стьюдента для независимых выборок не обнаружены статистически значимые различия во времени решения наглядно-образных задач среди учащихся разных профильных классов ( $t = -0,944$  при  $p = 0,358$ , т. е.  $p \geq 0,05$ ). Критерий равенства дисперсий Левина ( $F = 1,090$  при  $p = 0,310$ , т. е.  $p > 0,05$ ) указывает на то, что дисперсии двух распределений переменной *время решения вербальных задач* (для классов с разным профилем) статистически значимо не различаются, следовательно, можно интерпретировать

результаты работы  $t$ -теста. По  $t$ -критерию Стьюдента для независимых выборок не обнаружены статистически значимые различия во времени решения вербальных задач среди учащихся разных профильных классов ( $t = -0,798$  при  $p = 0,436$ , т. е.  $p \geq 0,05$ ).

**Вывод.** В результате проведенного статистического анализа не было выявлено статистически значимых различий во времени решения наглядно-образных и вербальных задач ни у школьников, обучающихся в разных профильных классах, ни у разнополых школьников.

### **Контрольные вопросы**

1. По каким существенным основаниям можно разграничить области применения параметрических и непараметрических методов сравнения выборок?
2. Какие конкретно параметрические критерии сравнения средних выборок вы знаете?
3. В чем состоит различие зависимых и независимых выборок?
4. Какие конкретно непараметрические критерии сравнения средних выборок вы знаете?
5. Перечислите несколько психологических задач, решение которых может потребовать применение методов сравнения выборок.

### **Рекомендуемая литература**

1. Ермолаев-Томин, О.Ю. Математические методы в психологии [Электронный ресурс] : учебник / О.Ю. Ермолаев-Томин. – 5-е изд., испр. и доп. – М. : Юрайт, 2014. – 511 с. – URL : [http://urss.ru/PDF/add\\_ru/159489-1.pdf](http://urss.ru/PDF/add_ru/159489-1.pdf) (Дата обращения 31.10.2016) (С. 91–93, 97–109, 118–133, 134–154).
2. Наследов, А. IBM SPSS Statistics 20 и AMOS [Электронный ресурс] : профессиональный статистический анализ данных / А. Наследов. – СПб. : Питер, 2013. – 416 с. – URL : <http://mexalib.com/view/29545> (Дата обращения 31.10.2016) (главы 11 и 12).
3. Некрасов, С.Д. Математические методы в психологии (MS Excel) [Электронный ресурс] : учеб. пособие / С.Д. Некрасов. – 3-е изд., испр. и доп. – Краснодар : Кубанский гос. ун-т, 2014. – 147 с. – URL : [http://docspace.kubsu.ru/docspace/bitstream/handle/1/295/Некрасов%20ММвПс\(Excel\)\\_пособие\\_2014\\_Doc2.pdf?sequence=7](http://docspace.kubsu.ru/docspace/bitstream/handle/1/295/Некрасов%20ММвПс(Excel)_пособие_2014_Doc2.pdf?sequence=7) (Дата обращения 31.10.2016) (С. 55–57, 63–68, 105–113).

## Практическое занятие 2

### Корреляционный анализ данных. Коэффициенты корреляции

**Цель** — ознакомление с алгоритмом применения основных процедур корреляционного анализа эмпирических данных.

#### Учебные вопросы:

- корреляционный анализ;
- коэффициенты корреляции;
- корреляционная матрица;
- критерий  $r$ -Пирсона;
- критерий  $r$ -Спирмена;
- критерий  $\tau$ -Кендалла.

Студент должен *знать* назначение и области применения корреляционного анализа; *уметь* выбирать коэффициент корреляции в соответствии с поставленной целью и особенностями выборки; *владеть* навыками выполнения корреляционного анализа данных.

#### Методические рекомендации по освоению теоретического материала по теме

**Корреляционный анализ** — это проверка гипотез о связях между переменными с использованием коэффициентов корреляции. **Коэффициент корреляции** — это мера прямой или обратной пропорциональности между двумя переменными, это количественная мера силы и направления вероятностной взаимосвязи двух переменных. На данный момент существует достаточно большое количество коэффициентов корреляции. Наиболее применимыми для психологических исследований являются  $r$ -Пирсона,  $r$ -Спирмена,  $\tau$ -Кендалла. Основная особенность заключается в том, что они отражают взаимосвязь двух признаков, измеренных в количественной шкале — ранговой или метрической. Корреляцию метрических переменных возможно осуществить с применением критерия  $r$ -Пирсона, гипотезу о корреляции ранговых переменных можно проверить при помощи критерия  $r$ -Спирмена или  $\tau$ -Кендалла.

Основными показателями являются сила, направление и надежность (достоверность) связи. Сила связи определяется по абсолютной величине корреляции (от 0 до 1). Направление связи определяется по знаку корреляции: положительный знак свидетельствует о прямой связи между переменными, отрицательный – о наличии обратной связи. Надежность связи определяется  $p$ -уровнем статистической значимости. Основная проверяемая статистическая гипотеза в отношении коэффициентов корреляции является ненаправленной и содержит утверждение о равенстве корреляции нулю в генеральной совокупности. При ее отклонении принимается альтернативная гипотеза о наличии положительной либо отрицательной (в зависимости от знака вычисленного коэффициента корреляции) связи.

Часто корреляционный анализ включает изучение связей не двух, а множества переменных, измеренных в количественной шкале на одной выборке. В этом случае вычисляются корреляции для каждой пары из этого множества переменных, результатом чего является корреляционная матрица.

**Корреляционная матрица** – это результат вычислений корреляций одного типа для каждой пары множества исследуемых переменных, измеренных в количественной шкале на одной выборке. Основная задача анализа корреляционной матрицы – выявление структуры взаимосвязей множества признаков. Это возможно сделать путем визуального анализа корреляционных плеяд – графического изображения структуры статистически значимых связей, если таковых не более 15. Другой способ состоит в дальнейшем применении многомерных методов, таких как множественный регрессионный, факторный или кластерный анализ [17, с. 147–160].

## Практические задания

### Задача 1

В группе менеджеров по продажам торгово-коммерческой фирмы с помощью опросника В.К. Гербачевского измерен уровень притязаний личности (выборка А) и по методике Р.С. Немова определена сила воли (выборка В). Имеется ли статистически достоверная линейная корреляционная связь между данными показателями?



### *Показатели уровня притязаний личности и силы воли*

№	Уровень притязаний	Сила воли	№	Уровень притязаний	Сила воли
1	2	21	6	5	21
2	13	27	7	12	42
3	6	33	8	10	33
4	6	9	9	3	18
5	8	33	10	10	33

#### **Задача 2**

Среди руководителей организаций определялась выраженность демократического стиля общения и демократического типа руководства. Имеется ли линейная корреляционная связь между исследованными показателями?

### *Показатели стиля общения и типа руководителя*

№	Стиль общения	Тип руководства	№	Стиль общения	Тип руководства
1	20	7,2	7	31	9,4
2	24	7,1	8	34	9,2
3	26	7,4	9	28	8,8
4	20	6,1	10	35	10,4
5	17	6,0	11	29	8,0
6	23	7,4	12	31	9,7

#### **Задача 3**

Учащиеся девятого класса общеобразовательной школы протестированы на склонность к девиантному поведению по следующим показателям: *АН* – агрессивность, насилие; *СП* – самоповреждающее поведение. Существует ли ранговая корреляционная связь между приведенными показателями?

### *Показатели склонности к девиантному поведению*

№	АН	СП	№	АН	СП
1	61	59	9	47	34
2	68	57	10	55	63
3	61	41	11	66	57
4	55	43	12	62	54

№	АН	СП	№	АН	СП
5	66	59	13	53	59
6	49	52	14	41	43
7	61	56	15	53	54
8	53	52	16	49	51

#### Задача 4

Проведено исследование показателей готовности к вузовскому обучению: умения учиться (*УУ*), навыков самоконтроля (*НС*) и успеваемости (*Усп*). Необходимо установить степень согласованности между данными показателями, используя ранговую корреляцию.

#### *Показатели психологической готовности*

№	УУ	НС	Усп	№	УУ	НС	Усп
1	50	77	5	9	80	69	5
2	70	90	5	10	90	85	5
3	60	70	5	11	75	60	5
4	90	78	4	12	50	50	3
5	50	50	5	13	59	51	4
6	60	70	5	14	60	50	4
7	30	40	4	15	50	50	3
8	20	25	3	16	80	70	4

#### Задача 5

Проведена диагностика уровня школьной тревожности (методика Филлипса) в третьих-шестых классах. Использовались следующие субшкалы: С1 – общая тревожность в школе, С2 – переживание социального стресса, С3 – фрустрация в потребности достижения успеха, С4 – страх самовыражения, С5 – страх ситуации проверки знаний, С6 – страх не соответствовать ожиданиям окружающих, С7 – низкая физиологическая сопротивляемость стрессу, С8 – проблемы и страхи в отношениях с учителями. Существует ли ранговая корреляционная связь между какими-либо показателями тревожности?

### Показатели школьной тревожности

№	C1	C2	C3	C4	C5	C6	C7	C8
1	6	4	6	3	7	4	5	5
2	5	6	4	4	6	5	6	4
3	3	5	4	6	3	8	6	5
4	6	8	4	5	3	6	3	5
5	4	4	5	3	5	9	4	6
6	5	5	4	7	3	8	7	4
7	6	6	7	4	5	7	4	4
8	5	6	5	3	5	6	4	5
9	4	6	7	4	6	7	5	4
10	4	6	5	3	4	5	6	7
11	5	6	5	7	5	7	4	3
12	4	6	5	7	3	6	7	5
13	3	5	6	4	5	8	5	6
14	5	6	7	5	3	6	5	4
15	7	5	5	4	6	7	6	4
16	6	3	6	5	7	5	4	6
17	5	4	7	6	5	9	3	7
18	6	5	3	7	6	5	6	3
19	6	5	4	7	5	7	4	5
20	4	6	4	7	6	8	4	6

#### Методические рекомендации и образцы выполнения заданий. Алгоритм осуществления корреляционного анализа

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче.

4. В меню *Анализ* выберите команду *Корреляции – Парные*. На экране появится диалоговое окно *Парные корреляции*. Чтобы выбрать переменные для вычисления корреляции, их требуется переместить в список *Переменные* при помощи кнопки со стрелкой. Если

несколько нужных переменных располагаются в исходном списке друг за другом, вы можете, наведя указатель мыши на верхнюю из них и нажав кнопку мыши, переместить указатель на нижнюю переменную и отпустить кнопку мыши, тем самым выделив сразу несколько переменных.

5. В группе *Коэффициенты корреляции* по умолчанию установлен флажок *Пирсона*. Если требуется вычислить ранговые корреляции, то следует установить флажок *r-Спирмена* и (или) *τ-Кендалла*. Можете установить все три флажка, чтобы иметь возможность сравнивать три коэффициента корреляции для различных распределений данных. В группе *Критерий значимости* по умолчанию установлен переключатель *Двухсторонний*. Если вы заранее уверены в направлении (знаке) корреляции, то можете установить переключатель *Односторонний*.

6. Флажок *Метить значимые корреляции* по умолчанию установлен. Это означает, что корреляции, вычисленные с уровнем значимости от 0,01 до 0,05, будут помечены одной звездочкой (\*), а от 0 до 0,01 — двумя звездочками (\*\*). Вне зависимости от значимости в вывод включаются коэффициенты корреляции и *p*-уровни, вычисленные с точностью до трех знаков после запятой, а также количество объектов, участвовавших в процедуре.

7. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### Пример

**Условие задачи.** 20 школьникам были даны тесты на наглядно-образное и вербальное мышление. Исследователями фиксировалось время решения задач в секундах. Выяснить, существует ли взаимосвязь между временем решения этих задач.

### Эмпирические данные

№	Время решения наглядно-образных заданий	Время решения вербальных заданий	Пол	Профиль класса
1	19	17	М	Ф-м
2	32	7	М	Ф-м
3	33	17	М	Ф-м

№	Время решения наглядно-образных заданий	Время решения вербальных заданий	Пол	Профиль класса
4	44	28	М	Гум-й
5	28	27	М	Гум-й
6	35	31	М	Гум-й
7	39	20	М	Ф-м
8	39	17	М	Ф-м
9	44	35	М	Ф-м
10	44	43	М	Гум-й
11	24	10	Ж	Гум-й
12	37	28	Ж	Гум-й
13	29	13	Ж	Ф-м
14	40	43	Ж	Ф-м
15	42	45	Ж	Ф-м
16	32	24	Ж	Ф-м
17	48	45	Ж	Гум-й
18	42	26	Ж	Гум-й
19	33	16	Ж	Гум-й
20	47	26	Ж	Гум-й

### Алгоритм решения задачи-примера

1. Выдвигаем гипотезы  $H_1: r_{xy} \neq 0$  и  $H_0: r_{xy} = 0$ .

2. Запускаем пакет SPSS, описываем переменные и заносим исходные данные в таблицу. Поскольку эмпирические данные согласованы с гипотезой о нормальности распределения (применение всех способов проверки распределения на нормальность подтвердило данный факт), то для проведения корреляционного анализа выбираем параметрический  $r$ -критерий Пирсона.

3. В меню *Анализ – Корреляции – Парные* открывается диалоговое окно *Парные корреляции* (рис. 6).

4. Выбираем переменные, выделяем их и переносим в поле *Переменные*. Флажок около коэффициента Пирсона выставлен по умолчанию.

5. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

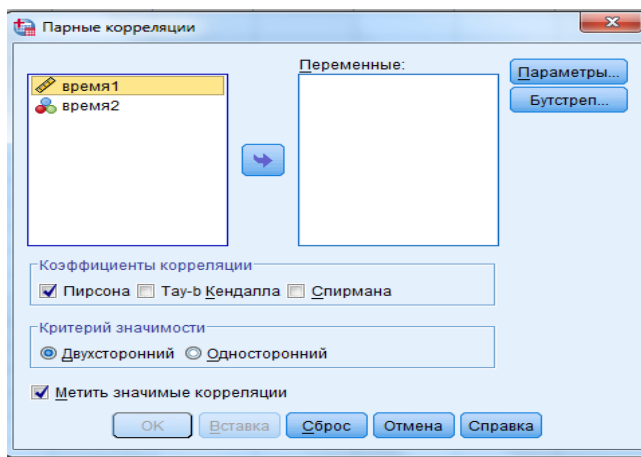


Рис. 6. Диалоговое окно Парные корреляции

**Результаты.** Выходной лист программы SPSS с вычисленными коэффициентами корреляции содержит таблицу *Корреляции* (рис. 7), в которой значками (\*\*) обозначается наличие значимой взаимосвязи.

Корреляции

		VAR00001	VAR00002
VAR00001	Корреляция Пирсона	1	,669**
	Знч. (2-сторон.)		,001
	N	20	20
VAR00002	Корреляция Пирсона	,669**	1
	Знч. (2-сторон.)	,001	
	N	20	20

\*\* Корреляция значима на уровне 0,01 (2-сторон.).

Рис. 7. Результат вычисленной корреляции

**Вывод.** В результате осуществления корреляционного анализа была обнаружена статистически значимая, прямая, умеренная взаимосвязь между временем решения наглядно-образных заданий и временем решения вербальных заданий ( $r = 0,669$ , при  $p \leq 0,01$ ), также можно отметить, что чем больше время решения наглядно-образных заданий, тем больше и время решения вербальных заданий.

### **Контрольные вопросы**

1. В чем состоит основное назначение корреляционного анализа?
2. Какие коэффициенты корреляции вам известны?
3. Какие основные показатели корреляционной связи вы знаете?
4. Перечислите несколько психологических задач, которые могут быть решены посредством применения корреляционного анализа.

### **Рекомендуемая литература**

1. Ермолаев-Томин, О.Ю. Математические методы в психологии [Электронный ресурс] : учебник / О.Ю. Ермолаев-Томин. – 5-е изд., испр. и доп. – М. : Юрайт, 2014. – 511 с. – URL : [http://urss.ru/PDF/add\\_ru/159489-1.pdf](http://urss.ru/PDF/add_ru/159489-1.pdf) (Дата обращения 31.10.2016) (С. 91–93, 97–109, 118–133, 134–154).
2. Наследов, А. IBM SPSS Statistics 20 и AMOS [Электронный ресурс] : профессиональный статистический анализ данных / А. Наследов. – СПб. : Питер, 2013. – 416 с. – URL : <http://mexalib.com/view/29545> (Дата обращения 31.10.2016) (главы 11 и 12).
3. Некрасов, С.Д. Математические методы в психологии (MS Excel) [Электронный ресурс] : учеб. пособие / С.Д. Некрасов. – 3-е изд., испр. и доп. – Краснодар : Кубанский гос. ун-т, 2014. – 147 с. – URL : [http://docspace.kubsu.ru/docspace/bitstream/handle/1/295/Некрасов%20ММвПс\(Excel\)\\_пособие\\_2014\\_Doc2.pdf?sequence=7](http://docspace.kubsu.ru/docspace/bitstream/handle/1/295/Некрасов%20ММвПс(Excel)_пособие_2014_Doc2.pdf?sequence=7) (Дата обращения 31.10.2016) (С. 55–57, 63–68, 105–113).

## Практическое занятие 3

### Дисперсионный анализ данных

**Цель** — ознакомление с алгоритмом применения основных видов дисперсионного анализа эмпирических данных.

#### Учебные вопросы:

- однофакторный дисперсионный анализ;
- многофакторный дисперсионный анализ;
- дисперсионный анализ с повторными измерениями;
- многомерный дисперсионный анализ;
- непараметрические аналоги дисперсионного анализа.

Студент должен *знать* назначение и области применения различных видов дисперсионного анализа; *уметь* выбирать вид дисперсионного анализа в соответствии с поставленной целью и особенностями выборки; *владеть* навыками выполнения дисперсионного анализа данных.

#### Методические рекомендации по освоению теоретического материала по теме

Общепринятое сокращенное обозначение дисперсионного анализа — ANOVA (от англоязычного *analysis of variance*). В соответствии с принятой классификацией, ANOVA — это метод сравнения нескольких (более двух) выборок по признаку, измеренному в метрической шкале. ANOVA решает задачу сравнения средних значений, но не двух, а нескольких выборок, также метод допускает сравнение более чем по одному основанию, в случае когда деление на выборки производится по нескольким номинативным переменным, каждая из которых имеет две и более градаций. Специфика ANOVA проявляется в двух аспектах: во-первых, этот метод использует терминологию планирования эксперимента; во-вторых, для сравнения средних значений анализируются компоненты дисперсии изучаемого признака. Типичная схема эксперимента сводится к изучению влияния независимой переменной (одной или нескольких) на зависимую. Независимая переменная представляет собой качественно определенный (номинативный) признак, имеющий две или более



градации. Каждой градации независимой переменной соответствует выборка объектов (испытуемых), для которых определены значения зависимой переменной. Независимая переменная иначе называется фактором, имеющим несколько градаций (уровней). Зависимая переменная рассматривается как изменяющаяся под влиянием независимых переменных и для реализации ANOVA должна быть представлена в метрической шкале. В зависимости от типа экспериментального плана выделяют четыре основных варианта ANOVA:

1) однофакторный ANOVA — используется при изучении влияния одного фактора на зависимую переменную;

2) многофакторный ANOVA — используется при изучении влияния двух и более независимых факторов на зависимую переменную, позволяя проверять гипотезы о влиянии каждого фактора в отдельности, а также о взаимодействии факторов;

3) ANOVA с повторными измерениями — применяется в том случае, когда различным градациям хотя бы одного фактора соответствует одна и та же выборка объектов (испытуемых);

4) многомерный ANOVA — применяется в том случае, когда зависимая переменная представляет собой множество измерений изучаемого явления (свойства).

Нулевая гипотеза в однофакторном ANOVA утверждает равенство межгрупповой и внутригрупповых составляющих дисперсии, т. е. отсутствие влияния независимой переменной на зависимую в генеральной совокупности.

Главное ограничение в использовании ANOVA связано с требованием равенства дисперсий выборок, соответствующих разным градациям фактора. Если это требование не выполняется, то можно воспользоваться непараметрическими аналогами дисперсионного анализа, такими как  $H$ -критерий Краскала – Уоллеса для независимых выборок или  $\chi^2$  Фридмана для повторных измерений [16, с. 120; 17, с. 185–190].

## Практические задания

### Задача 1

Измерялся объем восприятия геометрических фигур различного цвета: красного (К), сине-зеленого (СЗ) и черного (Ч). Предъявлялось по 10 фигур каждого цвета. Влияет ли фактор цвета фигур на объем их восприятия?

#### *Показатели объема восприятия геометрических фигур*

№	К	СЗ	Ч	№	К	СЗ	Ч
1	5	4	4	4	5	3	5
2	8	6	7	5	7	2	3
3	4	3	4	6	6	3	5

### Задача 2

В Центре планирования семьи по методике В.В. Столина исследовалось самоотношение первородящих беременных женщин и, в частности, уровень аутосимпатии. Выборки различались по фактору акушерского возраста: первая – 18–20 лет (Гр 1), вторая – 21–25 лет (Гр 2), третья – 26–30 лет (Гр 3). Влияет ли возраст женщины на уровень ее аутосимпатии?

#### *Показатели аутосимпатии*

№	Гр 1	Гр 2	Гр 3	№	Гр 1	Гр 2	Гр 3
1	11	8	5	6	10	8	7
2	10	9	5	7	12	7	4
3	13	8	7	8	13	10	9
4	13	7	8	9	11	5	8
5	12	6	7	10	13	6	7

### Задача 3

В группе испытуемых с помощью таблиц Шульце исследовали объем внимания (ОВ) по трем возрастным группам (ВГ): первая – 20–30 лет, вторая – 30–40 лет, третья – 40–50 лет. Фиксировался также пол испытуемых: 1 – мужской, 2 – женский. Определить, зависит ли объем внимания от пола и возраста.

### Показатели объема внимания, возраста и пола

№	ОВ	ВГ	Пол	№	ОВ	ВГ	Пол	№	ОВ	ВГ	Пол
1	7	1	1	11	6	2	1	21	6	3	1
2	7	1	1	12	6	2	1	22	6	3	1
3	7	1	1	13	5	2	1	23	7	3	1
4	6	1	1	14	7	2	1	24	5	3	1
5	6	1	1	15	4	2	1	25	5	3	1
6	5	1	2	16	6	2	2	26	4	3	2
7	7	1	2	17	6	2	2	27	5	3	2
8	7	1	2	18	7	2	2	28	4	3	2
9	6	1	2	19	6	2	2	29	6	3	2
10	6	1	2	20	6	2	2	30	5	3	2

#### **Методические рекомендации и образцы выполнения заданий. Алгоритм применения однофакторного дисперсионного анализа**

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, таким образом, чтобы первый столбец содержал наблюдения для всех измерений, размещенные один под другим. Второй столбец содержит единицы в клетках рядом с наблюдениями первого измерения, двойки рядом с наблюдениями второго измерения и тройки рядом с наблюдениями третьего измерения.

4. В меню *Анализ* выберите команду *Сравнение средних – Однофакторный дисперсионный анализ*. На экране появится диалоговое окно *Однофакторный дисперсионный анализ*. Структура диалогового окна *Однофакторный дисперсионный анализ* вполне типична для большинства диалоговых окон SPSS. Слева список переменных текущего файла данных. Список зависимых переменных предназначен для задания одной или нескольких зависимых переменных. Зависимые переменные должны быть метрического типа. Если

в списке указано несколько зависимых переменных, SPSS выполнит анализ для каждой из них. Под окном *Список зависимых переменных* находится поле *Фактор*, в котором нужно указать единственную независимую переменную, имеющую несколько градаций.

5. Щелкните сначала на первой переменной (содержащей значения наблюдений первого, второго и третьего измерений), чтобы выделить ее, а затем – на верхней кнопке со стрелкой, чтобы переместить переменную в *Список зависимых переменных*.

6. Щелкните на второй переменной (содержащей распределение данных по измерениям), чтобы выделить ее, а затем – на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Фактор*.

7. Нажмите кнопку *Параметры*. В открывшемся диалоговом окне отметьте флажком *Описательные статистики*, *Тест однородности дисперсии* и *График средних значений*. Нажмите *Продолжить*.

8. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

9. Нередко исследователя могут заинтересовать результаты парных сравнений уровней независимой переменной, и для этой цели в диалоговом окне *Однофакторный дисперсионный анализ* предусмотрена специальная кнопка *Апостериорные*, означающая, что эта процедура проводится после установления статистически достоверного результата однофакторного дисперсионного анализа.

10. Щелкните на кнопке *Апостериорные*, чтобы открыть диалоговое окно *Однофакторный дисперсионный анализ: Апостериорные множественные сравнения*.

11. Установите флажок *Шеффе* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Однофакторный дисперсионный анализ*.

12. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

13. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

## Алгоритм применения многофакторного дисперсионного анализа

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, следующим образом. Помимо измерений (первый столбец) необходимо создать две дополнительные переменные. Первая переменная (второй столбец) соответствует фактору А, вторая (третий столбец) – фактору В. В каждой из этих дополнительных переменных содержатся значения уровней факторов, при которых соответствующее измерение было сделано.

4. В меню *Анализ* выберите команду *Общая линейная модель – ОЛМ-одномерная*. На экране появится диалоговое окно *ОЛМ-одномерная*. В левой части диалогового окна *ОЛМ-одномерная* расположен список всех доступных переменных файла данных, а справа находятся поля и кнопки, с помощью которых задаются основные установки дисперсионного анализа. Поле *Зависимая переменная* предназначено для указания единственной зависимой переменной анализа, поле *Фиксированные факторы* – для имен номинальных независимых переменных или факторов.

5. Щелкните на первой переменной, чтобы выделить ее, а затем – на верхней кнопке со стрелкой, чтобы переместить переменную в поле *Зависимая переменная*.

6. Щелкните на второй переменной (обозначающей уровни фактора А), чтобы выделить ее, а затем – на второй сверху кнопке со стрелкой, чтобы переместить переменную в список *Постоянные факторы*.

7. Повторите предыдущее действие для третьей переменной (обозначающей уровни фактора В).

8. Щелкните на кнопке *Параметры*, чтобы открыть диалоговое окно *ОЛМ-одномерная: Параметры*. Установите флажки *Описатель-*

ные статистики, Оценка размера эффекта и Критерии однородности, а затем щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *ОЛМ-одномерная*.

9. Для наглядного представления результатов нажмите кнопку *Графики*. В открывшемся окне перенесите из общего поля, содержащего все факторы, название фактора, содержащего большее число уровней, в верхнее правое окно *Горизонтальная ось*.

10. Название второго фактора необходимо перенести в следующее правое окно *Отдельные линии*. Нажать кнопки *Добавить* – *Продолжить*.

11. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

12. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Алгоритм вычисления *H*-критерия Краскала – Уоллеса**

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, так, чтобы первый столбец содержал наблюдения для всех измерений, размещенные один под другим. Второй столбец содержит единицы в клетках рядом с наблюдениями первого измерения, двойки рядом с наблюдениями второго измерения и тройки рядом с наблюдениями третьего измерения.

4. В меню *Анализ* выберите команду *Непараметрические критерии – Устаревшие диалоговые окна – для *K*-независимых выборок*. В открывшемся диалоговом окне с помощью стрелок перенесите интересующие переменные (столбец 1) в верхнее поле *Список проверяемых переменных*, а данные столбца 2 в поле *Группирующая переменная*.

5. Нажмите кнопку *Задать диапазон*, в открывшемся окне укажите диапазон для группирующей переменной от минимума до

максимума (в зависимости от количества выборок, участвующих в анализе). Флажок на  $H$ -критерии Краскала – Уоллеса установлен по умолчанию.

6. Щелкните на кнопке ОК, чтобы открыть окно вывода.

7. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Алгоритм вычисления критерия $\chi^2$ Фридмана**

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, так, чтобы каждый столбец содержал количественные значения переменной, соответствующей каждому из сравниваемых условий.

4. В меню *Анализ* выберите команду *Непараметрические критерии – Устаревшие диалоговые окна – для K-зависимых выборок*. В открывшемся диалоговом окне с помощью стрелки перенести переменные, соответствующие нескольким измерениям одного и того же признака (столбец 1, 2, 3 и т. д.), в поле *Проверяемые переменные*. Флажок на критерии  $\chi^2$  Фридмана установлен по умолчанию.

5. Щелкните на кнопке ОК, чтобы открыть окно вывода.

6. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Пример**

**Условие задачи.** Имеются данные о годовых затратах туристических фирм на рекламу (млн руб.) и ежегодном количестве туристов, которых обслуживают эти фирмы. С 95 %-й доверительной вероятностью требуется выяснить, влияют ли затраты на рекламу на увеличение числа туристов, обслуживаемых фирмой.

### Эмпирические данные

Номера фирм	Затраты на рекламу, млн руб.	Количество туристов, чел	Номера фирм	Затраты на рекламу, млн руб.	Количество туристов, чел
1	8	800	11	10	920
2	8	850	12	10	1060
3	8	720	13	10	950
4	9	850	14	11	900
5	9	800	15	11	1200
6	9	880	16	11	1150
7	9	950	17	11	1000
8	9	820	18	12	1200
9	10	900	19	12	1100
10	10	1000	20	12	1000

### Алгоритм решения задачи-примера

1. Для ответа на поставленный вопрос следует использовать однофакторный дисперсионный анализ. Условием применимости дисперсионного анализа является равенство дисперсии факторной величины в различных группах. Проверить, удовлетворяют ли исходные данные этому условию, можно с помощью теста Левене. Если критическое значение статистики Левене больше эмпирического при заданном уровне значимости, то гипотеза принимается и дисперсию в группах можно считать одинаковой, а применение дисперсионного анализа оправданным.

2. Для проверки наличия взаимосвязи между вложением денег в рекламу и числом обслуженных туристов внесем данные в окно редактора SPSS таким образом, чтобы первый столбец содержал данные о затратах, а второй – о количестве туристов. Проведем дисперсионный анализ, активизировав опции *Анализ – Сравнение средних – Однофакторный дисперсионный анализ* (рис. 8).

3. В появившемся окне следует переменную *Количество туристов* перенести в окно *Список зависимых*, а переменную *Затраты на рекламу* в окно *Фактор*.

4. Для того чтобы проверить однородность дисперсии, необходимо нажать кнопку *Параметры*, в появившемся окне отметить флажком *Описательные статистики*, *Тест на однородность дисперсии*, *График средних значений* и нажать кнопку *ОК* (рис. 9).



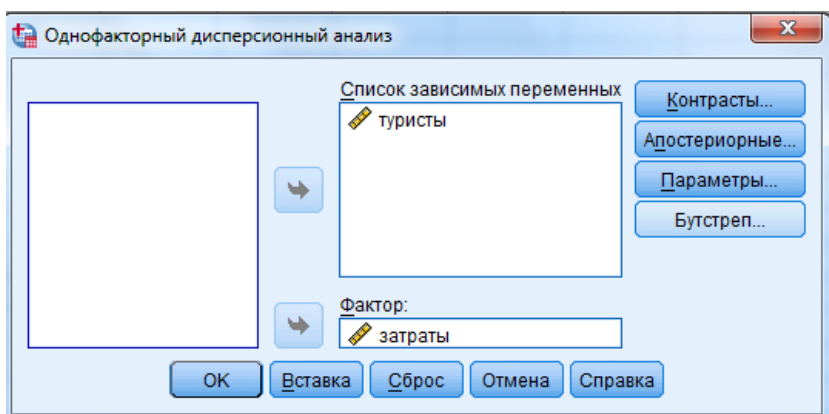


Рис. 8. Диалоговое окно Однофакторный дисперсионный анализ

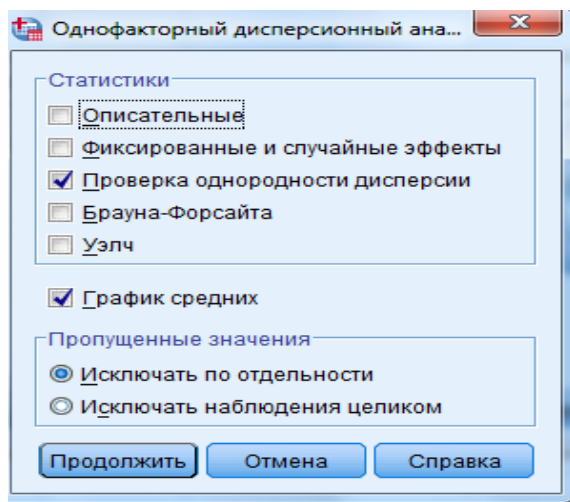


Рис. 9. Диалоговое окно Однофакторный дисперсионный анализ – Параметры

**Результат.** В окне результатов интерес представляют таблица итогов проверки входных данных на однородность дисперсии (табл. 1) и таблица итогов однофакторного дисперсионного анализа (табл. 2).

Таблица 1

## Итоги теста Левене на однородность дисперсии

Статистика Левене	Ст. св. 1	Ст. св. 2	Знач.
2,169	4	15	,122

Из приведенных данных следует, что статистически достоверных различий между дисперсиями не обнаружено, поскольку значимость статистики Левене  $p = 0,122$  превышает  $p$ -уровень  $= 0,05$ , что позволяет отклонить гипотезу о неравенстве дисперсий. Для нас представляет интерес окончательный результат: эмпирическое значение статистики Фишера в нашем примере  $F_{\text{эмп}} = 7,648$  и значимость  $p = 0,001$  (табл. 2).

Таблица 2

## Итоги однофакторного дисперсионного анализа

	Сумма квадратов отклонений	Ст. св.	Средний квадрат	$F$	Знач.
Между группами	236580,00	4	59145,00	7,648	0,001
Внутри групп	115995,00	15	7733,00		
Итого	352575,00	19			

Поскольку проверялась исходная гипотеза ( $H_0$ ) об отсутствии взаимосвязи между количеством туристов, которых обслужила фирма, и уровнем расходов на рекламу и эмпирическое значение  $F_{\text{эмп}} = 7,648$  при уровне значимости  $p = 0,05$  оказалось существенно больше критического значения  $F_{\text{кр}} = 3,06$ , то нулевую гипотезу придется отклонить и принять альтернативную ( $H_1$ ) – о наличии взаимосвязи между факторным и результативным признаками.

**Вывод.** В результате проведенного дисперсионного анализа была обнаружена статистически достоверная зависимость ( $F = 7,648$  при уровне значимости  $p \leq 0,001$ ), состоящая в увеличении количества туристов от увеличения финансовых затрат туристических фирм на рекламу.

### **Контрольные вопросы**

1. В чем состоит основное назначение дисперсионного анализа?
2. Какие виды дисперсионного анализа вам известны?
3. В чем состоит основное требование к данным, подвергающимся дисперсионному анализу?
4. Перечислите несколько психологических задач, которые могут быть решены посредством применения дисперсионного анализа.

### **Рекомендуемая литература**

1. Ермолаев-Томин, О.Ю. Математические методы в психологии [Электронный ресурс] : учебник / О.Ю. Ермолаев-Томин. – 5-е изд., испр. и доп. – М. : Юрайт, 2014. – 511 с. – URL : [http://urss.ru/PDF/add\\_ru/159489-1.pdf](http://urss.ru/PDF/add_ru/159489-1.pdf) (Дата обращения 31.10.2016) (С. 91–93, 97–109, 118–133, 134–154).
2. Наследов, А. IBM SPSS Statistics 20 и AMOS [Электронный ресурс] : профессиональный статистический анализ данных / А. Наследов. – СПб. : Питер, 2013. – 416 с. – URL : <http://mexalib.com/view/29545> (Дата обращения 31.10.2016) (главы 11 и 12).
3. Некрасов, С.Д. Математические методы в психологии (MS Excel) [Электронный ресурс] : учеб. пособие / С.Д. Некрасов. – 3-е изд., испр. и доп. – Краснодар : Кубанский гос. ун-т, 2014. – 147 с. – URL : [http://docspace.kubsu.ru/docspace/bitstream/handle/1/295/Некрасов%20ММвПс\(Excel\)\\_пособие\\_2014\\_Doc2.pdf?sequence=7](http://docspace.kubsu.ru/docspace/bitstream/handle/1/295/Некрасов%20ММвПс(Excel)_пособие_2014_Doc2.pdf?sequence=7) (Дата обращения 31.10.2016) (С. 55–57, 63–68, 105–113).

## Практическое занятие 4

### Регрессионный анализ данных

**Цель** – ознакомление с алгоритмом применения регрессионного и множественного регрессионного анализа данных.

#### **Учебные вопросы:**

- регрессионный анализ;
- множественный регрессионный анализ;
- коэффициенты регрессии;
- коэффициент детерминации.

Студент должен *знать* назначение и области применения регрессионного анализа; *уметь* выбирать вид регрессионного анализа в соответствии с поставленной целью и особенностями выборки; *владеть* навыками выполнения регрессионного анализа данных.

#### **Методические рекомендации по освоению теоретического материала по теме**

Регрессионный анализ используется для определения того, какая доля метрической переменной может быть объяснена другими переменными. Таким образом, назначение регрессионного анализа сходно с назначением дисперсионного анализа. Различие между ними состоит в возможностях использования. Дисперсионный анализ можно применять, когда независимая переменная не является метрической, осуществление же регрессионного анализа возможно только в случае, когда обе переменные, и зависимая и независимая, являются метрическими. Вид зависимости от одной переменной возможно установить посредством одномерной линейной регрессии. При изучении влияния нескольких независимых переменных на зависимую применяется множественный регрессионный анализ, распространенной целью которого является возможность предсказания некоторого результата (обучения, деятельности) по ряду предварительно измеренных характеристик [16, с. 148; 17, с. 240–245].

## Практические задания

### Задача 1

Исследователь пытается выявить взаимосвязь между количеством времени  $X$ , бесполезно потраченного студентами, и средним баллом  $Y$  их академической успеваемости, который варьируется в пределах от 2,0 до 5,0. Под потраченным без пользы временем понимается количество часов определенного соответствующего времяпровождения в неделю (например, занятого просмотром «мыльных» телесериалов). Требуется построить линейную регрессионную зависимость среднего балла успеваемости от показателя бесполезно потраченного времени, а также выполнить прогноз успеваемости для значений  $X$ , равных 20, 30 и 40 часов. В ответе к задаче приведите: а) скорректированный процент объясняемой моделью дисперсии отклика; б) оценку уровня значимости регрессии по  $F$ -критерию; в) нестандартизованные коэффициенты регрессионного уравнения и оценки их значимости по  $t$ -критерию; г) номер наблюдения, имеющего наибольший по модулю остаток; д) прогнозируемые значения отклика. Сделать содержательный вывод.

### *Средний балл успеваемости студентов и показатель бесполезно потраченного ими времени*

№	$X$	$Y$	№	$X$	$Y$	№	$X$	$Y$
1	42	2,8	4	35	3,9	7	39	3,4
2	23	4,0	5	16	4,7	8	19	4,4
3	31	3,2	6	26	4,0	9	29	3,80

### Задача 2

В консультационной фирме проведено исследование, в ходе которого у группы сотрудников информационно-аналитического отдела в шкале станайнов определялись следующие показатели: УПМ – уровень профессиональной мотивации, УКХ – суммарный уровень когнитивных характеристик, ОЭФ – объективный показатель эффективности профессиональной деятельности. Полученные данные приведены в таблице. Требуется построить линейную регрессионную зависимость эффективности изучаемой трудовой деятельности от показателей профессиональной мотивации и суммарного уровня когнитивных характеристик. Кроме этого, надо

построить графики распределения остатков, в частности — гистограмму и график с нормальной прямой. В ответе к задаче приведите: а) скорректированное значение коэффициента множественной детерминации; б) оценку уровня значимости регрессии по  $F$ -критерию; в) нестандартизованные коэффициенты регрессионного уравнения и оценки их значимости по  $t$ -критерию.

***Показатели профессиональной мотивации, когнитивных характеристик и эффективности труда***

№	УМП	УКХ	ОЭФ	№	УМП	УКХ	ОЭФ
1	4	8	5	10	3	8	3
2	5	6	5	11	5	6	5
3	6	5	6	12	6	5	6
4	7	8	8	13	7	5	7
5	5	4	4	14	8	5	7
6	8	9	8	15	5	6	6
7	5	7	6	16	6	6	6
8	4	5	4	17	8	7	7
9	8	4	6	18	7	6	6

**Задача 3**

В исследовании, посвященном методике обучения математике психологов, оценивалось время, потраченное студентами на разные формы обучения (время прослушанных лекций, время, проведенное на семинарских и практических занятиях, время самостоятельной работы, в часах), и итоговая успеваемость (по 100-балльной шкале). Одной из задач исследования было выяснение, какие виды обучения наиболее тесно связаны с итоговой успеваемостью. Были выдвинуты две основные гипотезы: 1) наиболее важными видами работы являются аудиторские занятия (лекции и семинары); 2) большое значение в общем усвоении материала имеют практические занятия, позволяющие закрепить полученные теоретические знания. Проверьте сформулированные гипотезы.

### Эмпирические данные

Количество прослушанных лекций	Количество семинаров	Количество практикумов	Время на самостоятельную работу	Успеваемость
14	16	18	20	65
16	18	14	12	50
16	20	14	17	65
16	14	14	24	75
18	14	16	15	55
18	16	14	18	60
18	16	18	21	70
20	20	16	16	55
20	20	12	14	60
20	12	14	18	60
20	18	16	16	65
20	20	16	20	80
20	16	20	18	90
22	20	18	16	65
22	18	16	20	75
22	20	22	16	85
24	18	18	16	80
16	20	16	22	70
26	20	24	20	80
26	16	18	22	90
26	20	20	22	90

### Методические рекомендации и образцы выполнения заданий. Алгоритм применения регрессионного анализа

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче.

4. В меню *Анализ* выберите команду *Регрессия – Линейная*. На экране появится диалоговое окно *Линейная регрессия*.

5. Щелкните на зависимой переменной, чтобы выделить ее, а затем на верхней кнопке со стрелкой, чтобы переместить ее в поле *Зависимая переменная*.

6. Щелкните на *Независимой переменной*, чтобы выделить ее, а затем на второй сверху кнопке со стрелкой, чтобы переместить переменную в список *Независимые переменные*. В качестве метода анализа оставьте стандартный, стоящий по умолчанию.

7. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

8. В результате программа вычислит величины  $R$  (коэффициент корреляции зависимой и независимой переменных),  $R^2$  (коэффициент, отражающий долю дисперсии зависимой переменной, обусловленную воздействием независимой переменной),  $F$  (показатель, относящийся к дисперсионному анализу), соответствующие им  $p$ -уровни,  $\beta$ -коэффициенты (коэффициенты и константы уравнения регрессии), а также стандартизированные  $\beta$ -коэффициенты, характеризующие степень зависимости между значениями исследуемых переменных.

9. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Алгоритм применения множественного регрессионного анализа**

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче так, чтобы каждый столбец содержал значения отдельной переменной (независимых и зависимой).

4. В меню *Анализ* выберите команду *Регрессия – Линейная*. На экране появится диалоговое окно *Линейная регрессия*.



5. Щелкните на зависимой переменной, чтобы выделить ее, а затем на кнопке со стрелкой, чтобы переместить ее в верхнее поле *Зависимая переменная*.

6. Щелкните на каждой из независимых переменных, чтобы выделить их, а затем на второй сверху кнопке со стрелкой, чтобы переместить их в список поля *Независимые переменные*. В качестве метода анализа можно выбрать любой из раскрывающегося списка пункта *Метод*, оставим метод, установленный по умолчанию — *Принудительное включение*.

7. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

8. В окне вывода результатов наиболее информативными являются следующие.

В таблице *Сводка для модели*:

- коэффициент множественной корреляции  $R$  — отражает связь зависимой переменной с совокупностью независимых переменных;
- коэффициент множественной детерминации  $R^2$  — отражает процент дисперсии зависимой переменной, который обусловлен влиянием совокупности независимых переменных.

В таблице *Коэффициенты*:

- стандартные коэффициенты регрессии  $\beta$  отражают относительную степень влияния каждого из предикторов (независимых переменных) при их статистической значимости ( $p < 0,005$ ). Также  $\beta$ -коэффициенты используются для предсказания значений зависимой переменной.

9. Проанализировав результат анализа, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Пример**

**Условие задачи.** При низкой возбудимости результаты экзамена должны быть низкими, поскольку излишнее спокойствие снижает потребность студента в подготовке. С увеличением возбудимости результат до определенного момента должен улучшаться, однако слишком возбудимые студенты вряд ли способны сконцентрироваться и показать хороший результат. Поэтому наилучших показателей должны добиваться те студенты, чей уровень возбудимости

является промежуточным. В таблице приведены показатели тревожности и результаты теста, необходимо проверить регрессионную зависимость результатов теста от уровня тревожности.

### *Эмпирические данные*

№	Трев.	Тест	№	Трев.	Тест	№	Трев.	Тест
1	8	13	13	2	10	25	10	12
2	3	12	14	8	14	26	7	18
3	2	10	15	8	12	27	3	12
4	7	14	16	2	17	28	4	12
5	6	15	17	6	16	29	6	15
6	5	12	18	3	9	30	7	17
7	9	13	19	9	12	31	6	16
8	9	12	20	6	16	32	7	17
9	4	10	21	6	14	33	5	14
10	3	8	22	4	10	34	8	15
11	6	16	23	8	14	35	5	14
12	7	15	24	7	16	36	7	17

### **Алгоритм решения задачи-примера**

1. Внесите данные задачи в окно редактора SPSS. Обозначьте имена переменных. Сформулируйте гипотезы.

2. В меню *Анализ* выберите команду *Регрессия – Линейная*. На экране появится диалоговое окно *Линейная регрессия*. В списке, находящемся в левой части диалогового окна, содержатся имена двух переменных – *трев* и *тест*. Щелкните сначала на переменной *тест*, чтобы выделить ее, а затем на верхней кнопке со стрелкой, чтобы переместить ее в поле *Зависимая переменная*.

3. Щелкните на переменной *трев*, чтобы выделить ее, а затем – на второй сверху кнопке со стрелкой, чтобы переместить ее в список *Независимые переменные* (рис. 10).

4. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

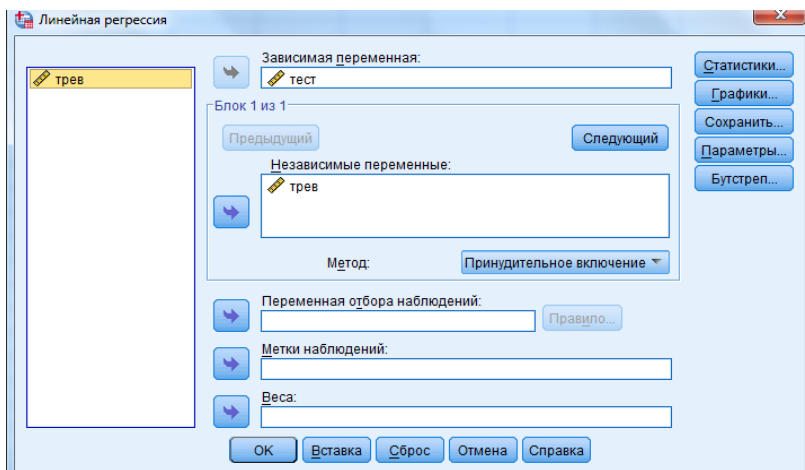


Рис. 10. Диалоговое окно Линейная регрессия

**Результат.** На выходном листе программы SPSS представляют интерес таблицы *Сводка для модели* и *Коэффициенты* (рис. 11).

#### Сводка для модели

Модель	R	R-квадрат	Скорректированный R-квадрат	Стд. ошибка оценки
1	,546 <sup>а</sup>	,298	,278	2,306

а. Предикторы: (конст) трев

#### Коэффициенты<sup>а</sup>

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Знач.
	B	Стд. Ошибка	Бета		
1 (Константа)	9,311	1,118		8,327	,000
	,675	,177	,546	3,803	,001

а. Зависимая переменная: тест

Рис. 11. Результаты регрессионных исчислений

Необходимо обратить внимание на значительный коэффициент корреляции зависимой переменной *тест* с независимой переменной *трев*. Величина  $R^2 = 0,298$  означает, что 29,8 % дисперсии переменной *тест* обусловлено влиянием со стороны переменной *трев*.  $\beta$ -коэффициент, равный 0,546, характеризует вклад переменной *трев* в регрессионную модель.

**Вывод.** При осуществлении регрессионного анализа была выявлена некоторая связь результатов теста с уровнем тревожности студентов, но полученные результаты не позволяют утверждать, что с увеличением уровня тревожности будет повышаться и успешность прохождения тестирования.

### Контрольные вопросы

1. В чем состоит основное назначение регрессионного анализа?
2. Какие виды регрессионного анализа вам известны?
3. Каково основное требование к данным, подвергающимся регрессионному анализу?
4. Перечислите несколько психологических задач, которые могут быть решены посредством применения регрессионного анализа.

### Рекомендуемая литература

1. Ермолаев-Томин, О.Ю. Математические методы в психологии [Электронный ресурс] : учебник / О.Ю. Ермолаев-Томин. – 5-е изд., испр. и доп. – М. : Юрайт, 2014. – 511 с. – URL : [http://urss.ru/PDF/add\\_ru/159489-1.pdf](http://urss.ru/PDF/add_ru/159489-1.pdf) (Дата обращения 31.10.2016) (С. 91–93, 97–109, 118–133, 134–154).
2. Наследов, А. IBM SPSS Statistics 20 и AMOS [Электронный ресурс] : профессиональный статистический анализ данных / А. Наследов. – СПб. : Питер, 2013. – 416 с. – URL : <http://mexalib.com/view/29545> (Дата обращения 31.10.2016) (главы 11 и 12).

## Практическое занятие 5

### Факторный анализ данных

**Цель** — ознакомление с алгоритмом осуществления основных процедур факторного анализа эмпирических данных.

#### Учебные вопросы:

- анализ главных компонент и факторный анализ;
- проблема числа факторов и проблема общности;
- методы факторного анализа;
- проблема вращения и интерпретации;
- проблема оценки значений факторов.

Студент должен *знать* назначение и особенности применения факторного анализа; *уметь* выбирать метод факторного анализа в соответствии с поставленной целью и особенностями выборки; *владеть* навыками выполнения факторного анализа данных.

#### Методические рекомендации по освоению теоретического материала по теме

**Факторный анализ** представляет собой совокупность методов, призванных определить, насколько связанные (коррелирующие) переменные могут быть сгруппированы так, чтобы каждую группу можно было рассматривать как одну составную переменную или фактор, а не как ряд отдельных переменных. **Фактор** интерпретируется как причина совместной изменчивости нескольких исходных переменных. С помощью факторного анализа исследуется взаимосвязь измеряемых индикаторов (наблюдаемых зависимых переменных) и латентных факторов (независимых переменных) между собой. Задача состоит в группировке этого большого числа наблюдаемых переменных в существенно меньшее число групп. Каждая группа трактуется как латентная переменная, которая предполагается как независимая и детерминирующая измеряемые переменные. Сила детерминации выражается в факторных нагрузках. Чем больше абсолютное значение факторной нагрузки, тем в большей степени измеряемая переменная характеризует соответствующую латентную. Задача интерпретации факторного решения заключается в содержательном объяснении статистически полученной группировки, что предполагает сопостав-

ление отдельных смыслов каждой вошедшей в фактор переменной и выделение общего инварианта.

Особенность факторного анализа заключается в неопределенности решения его основных проблем, таких как проблема числа факторов; общности; вращения и интерпретации; оценки значений факторов. Отсутствие четких критериев эффективности решения этих проблем трансформирует процесс факторного анализа в пошаговую процедуру, предполагающую последовательное принятие решений о дальнейших действиях на каждом шаге. Совокупность действий при реализации факторного анализа можно представить следующим образом.

1. Выбор исходных данных.
2. Предварительное решение проблемы числа факторов.
3. Факторизация матрицы интеркорреляций.
4. Вращение факторов и их предварительная интерпретация.
5. Принятие решения о качестве факторной структуры.
6. Вычисление факторных коэффициентов и оценок.

В зависимости от поставленных целей исследователь сам решает вопросы о том, какие из этапов могут быть пропущены, какова должна быть степень проработки каждого этапа, а также вопрос количества повторений последовательности этапов факторного анализа данных [16, с. 168–176; 17, с. 251–277].

## Практические задания

### Задача 1

На коммерческом предприятии с помощью опросника Р. Кеттелла 16-PF измерены (в стэхах) личностные свойства менеджеров по персоналу. Выявите общие факторы данных свойств, применив метод главных компонент и упростив факторную структуру с помощью вращения типа *Varimax Normalized*. В качестве критерия количества факторов следует использовать количество собственных значений, превышающих единицу. Дополнительно в ходе решения надо построить график «каменистая осыпь».

В ответе к задаче укажите: а) количество выделенных факторов и процент объясняемой дисперсии; б) собственные значения, соответствующие факторам; в) для каждого фактора – личностные

свойства, по которым факторные нагрузки превышают по модулю 0,7; г) интерпретацию выделенных факторов, основанную на матрице факторных нагрузок.

### *Личностные характеристики менеджеров*

													1	2	3	4
1	4	10	9	8	5	3	5	6	8	6	8	5	5	6	9	6
2	6	6	4	6	6	3	6	7	7	4	9	6	6	6	8	5
3	9	6	7	5	4	3	6	4	7	2	10	9	3	7	10	8
4	6	5	7	3	9	2	4	5	4	3	8	6	7	4	8	5
5	7	4	8	5	6	5	5	6	5	9	7	5	9	6	7	6
6	8	6	7	6	5	3	5	8	6	5	7	9	5	3	5	7
7	7	6	6	8	6	6	10	5	7	3	10	7	6	9	7	6
8	7	7	6	5	4	1	9	4	7	7	6	6	4	7	9	7
9	5	6	8	6	4	3	7	5	8	6	9	5	6	7	8	5
10	7	6	7	5	5	3	8	4	9	5	8	6	5	5	6	5
11	7	6	7	5	5	3	6	5	8	4	8	3	5	8	7	5
12	7	7	5	5	6	4	4	6	7	3	4	5	7	7	8	9
13	8	9	6	7	7	3	5	7	6	5	7	4	5	8	6	6
14	7	6	5	6	6	2	5	6	5	4	3	5	6	7	3	6
15	8	6	7	5	6	3	5	7	8	3	5	6	5	7	4	5
16	7	6	7	6	5	4	8	4	9	5	7	3	6	7	8	5
17	7	5	7	9	5	1	6	5	8	4	7	6	5	4	7	5
18	7	3	9	6	5	2	8	4	9	5	8	4	6	7	8	5
19	7	6	7	5	5	2	6	5	8	4	7	6	5	6	9	5
20	7	5	8	6	3	3	8	4	9	5	6	7	5	5	7	5
21	7	8	7	4	5	2	6	8	8	4	8	6	4	8	7	5

### **Задача 2**

Привлеченные к исследованию эксперты (15 человек) выделили 14 наиболее важных, на их взгляд, профессиональных качеств консультанта кризисного центра по оказанию психологической помощи женщинам, оказавшимся в ситуации насилия. В состав данных качеств вошли: К1 – эмпатия, К2 – коммуникабельность, К3 – артистизм, К4 – суггестивность, К5 – толерантность, К6 – наблюдательность, К7 – конформизм, К8 – эмоциональная уравновешенность, К9 – интуитивность, К10 – инструментальный профессионализм в области психологического консультирования, К11 – экспрессив-

ность, К12 – концентрация внимания, К13 – способность к психической саморегуляции, К14 – профессиональная креативность.

Затем эксперты оценили (в баллах по шкале от 1 до 7) степень важности каждого из выделенных качеств для профессиональной деятельности консультанта. Требуется выявить в выборке измеренных качеств общие факторы, используя метод главных факторов. В качестве критерия количества факторов следует использовать количество собственных значений, превышающих единицу. Кроме того, необходимо упростить факторную структуру, применив вращение факторов типа *Varimax Normalized*, после чего попытаться дать интерпретацию выделенных факторов, исходя из полученной матрицы факторных нагрузок.

**В ответе к задаче укажите:** а) количество выделенных факторов и процент объясняемой с их помощью дисперсии; б) собственные значения, соответствующие данным факторам; в) для каждого фактора – параметры, по которым факторные нагрузки по данному фактору превышают по модулю 0,7.

***Экспертные оценки профессионально важных качеств  
консультанта кризисного центра***

№	К1	К2	К3	К4	К5	К6	К7	К8	К9	К10	К11	К12	К13	К14
1	4	4	2	6	5	7	2	7	6	4	4	4	7	3
2	5	5	2	5	5	6	3	6	6	5	3	3	6	4
3	4	4	3	5	6	5	2	7	5	6	1	2	7	6
4	5	5	4	6	5	7	2	7	6	7	2	4	7	7
5	6	5	2	6	4	7	2	6	6	6	2	4	5	5
6	4	6	2	5	5	6	3	7	5	7	2	3	7	7
7	5	5	2	6	5	7	2	7	6	6	2	4	7	6
8	5	4	4	5	6	6	1	6	5	6	3	3	7	6
9	6	5	2	6	5	7	2	7	6	7	2	4	7	7
10	5	5	3	6	5	7	2	7	6	7	4	5	7	6
11	5	5	2	6	6	5	2	7	5	7	2	4	7	7
12	5	5	2	7	5	7	2	7	6	7	2	4	7	6
13	5	4	3	6	5	7	2	5	6	7	3	4	6	7
14	5	5	2	5	4	6	2	7	6	5	2	4	7	7
15	6	6	2	6	5	7	3	7	7	7	2	3	6	6



## **Методические рекомендации и образцы выполнения заданий. Алгоритм осуществления факторного анализа**

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче, так, чтобы каждая строка содержала наблюдение (например, испытуемый, респондент), а столбцы – показатели (например, ответы на вопросы) – это переменные, которые факторизируются в ходе анализа. Таким образом, получается матрица первичных данных.

4. В меню *Анализ* выберите команду *Снижение размерности – Факторный анализ*. На экране появится диалоговое окно *Факторный анализ*. Оно почти ничем не отличается от большинства диалоговых окон команд статистических операций SPSS: список доступных переменных находится в левой части, список *Переменные*, предназначенный для указания имен переменных, участвующих в анализе, – в центре, а 5 кнопок, с помощью которых задаются параметры анализа, – в правой.

5. Щелкните на первой переменной, нажмите клавишу Shift и, не отпуская ее, щелкните на последней переменной. В результате окажутся выделенными все промежуточные переменные.

6. Щелкните на верхней кнопке со стрелкой, чтобы переместить выделенные переменные в список *Переменные*.

7. Щелкните на кнопке *Описательные*, чтобы открыть диалоговое окно *Факторный анализ: Описательные*.

8. В группе *Статистики* установите флажок *Одномерные описательные*, в группе *Корреляционная матрица* – флажки *Коэффициенты* и *Уровень значимости* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ*.

9. Щелкните на кнопке *Извлечение*, чтобы открыть диалоговое окно *Факторный анализ: Выделение факторов*.

10. Установите флажок *График собственных значений* в группе *Отображать* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ*.

11. Щелкните на кнопке *Вращение*, чтобы открыть диалоговое окно *Факторный анализ: Вращение*.

12. В группе *Метод* установите переключатель *Варимакс*, в группе *Отображать* — флажок *Диаграммы нагрузок* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ*.

13. Щелкните на кнопке *Параметры*, чтобы открыть диалоговое окно *Факторный анализ: Параметры*.

14. Установите флажки *Отсортировать по величине* и *Не выводить коэффициенты с низкими значениями*; в окне *Абсолютное значение меньше* напротив установите вместо 0,10 величину 0,7. Щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ*.

15. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

16. Проанализировав результат, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Пример**

**Условие задачи.** У исследователя есть данные об 11 переменных интеллекта, полученных на выборке из 46 школьников. Тест включал 11 субтестов (переменные 1, 2, ..., 11). Предполагалось, что эти 11 субтестов позволят измерить 3 и более обобщенные интеллектуальные характеристики: математические, вербальные и невербальные (образные). Необходимо объединить переменные в факторы с включением в вывод одномерных описательных статистик всех переменных и коэффициентов корреляции. Для извлечения факторов необходимо использовать метод главных компонентов, а для отображения — график собственных значений. Вращение факторов должно производиться методом *Varimax*, с отображением факторной структуры после вращения.

*Эмпирические данные*

Пер. 1	Пер. 2	Пер. 3	Пер. 4	Пер. 5	Пер. 6	Пер. 7	Пер. 8	Пер. 9	Пер. 10	Пер. 11
12	9	11	8	8	11	13	8	12	10	11
10	12	12	11	10	12	9	12	8	11	11
11	8	9	11	11	12	9	11	9	11	8
14	12	14	13	13	9	9	9	12	11	12
12	12	9	10	9	12	5	10	3	11	9
10	12	12	8	13	12	9	9	11	8	11
9	2	6	10	7	4	8	10	7	5	9
14	5	13	11	13	13	13	9	14	9	14
14	11	11	10	8	12	13	8	12	9	9
15	14	11	11	12	12	14	11	12	11	16
13	7	3	9	8	7	7	6	7	3	9
9	8	7	12	14	12	7	6	8	13	11
16	14	15	11	11	11	10	10	12	12	12
14	12	11	10	10	12	10	11	13	14	12
11	6	9	9	9	7	14	8	11	6	11
7	7	15	11	9	9	10	8	12	8	10
13	12	15	9	11	7	8	12	15	11	11
8	9	9	9	12	12	9	9	15	13	9
12	11	10	13	11	8	10	10	10	14	11
14	9	11	13	12	13	13	14	13	11	11
11	12	12	12	11	12	14	9	12	10	11
11	9	10	7	7	8	10	5	12	9	9
11	11	11	12	8	11	10	10	14	11	10
10	12	10	12	12	9	13	12	12	12	10
9	12	8	13	12	9	14	12	18	12	6
10	10	9	11	10	12	12	10	12	13	7
9	6	10	13	8	7	10	12	11	9	8
10	7	8	12	10	11	10	10	14	12	9
11	16	12	13	10	9	16	11	12	16	9
9	12	7	10	12	8	10	10	14	12	6
10	12	9	14	10	7	14	11	13	12	11
9	9	11	12	8	11	9	10	16	13	6
11	10	13	13	13	9	17	11	14	12	9
9	8	11	13	8	8	13	12	13	10	7
10	9	7	12	7	11	13	15	12	10	8
9	10	13	13	10	11	13	11	14	12	10

Пер. 1	Пер. 2	Пер. 3	Пер. 4	Пер. 5	Пер. 6	Пер. 7	Пер. 8	Пер. 9	Пер. 10	Пер. 11
13	10	8	15	12	7	16	9	15	18	5
7	11	7	9	12	11	9	8	13	11	11
7	11	8	11	11	9	9	6	12	11	6
13	9	13	14	8	11	13	15	15	13	12
7	4	6	8	8	2	4	6	6	12	9
11	9	8	6	7	8	8	5	7	10	6
11	7	12	11	9	9	14	11	14	12	11
11	11	10	12	11	13	13	12	14	13	11
11	11	12	11	10	11	9	9	15	12	9
8	12	9	11	8	9	9	12	14	13	12

### Алгоритм решения задачи-примера

1. Внесите данные задачи в окно редактора SPSS. Сформулируйте гипотезы.

2. В меню *Анализ* выберите команду *Снижение размерности – Факторный анализ*. На экране появится диалоговое окно *Факторный анализ* (рис. 12).

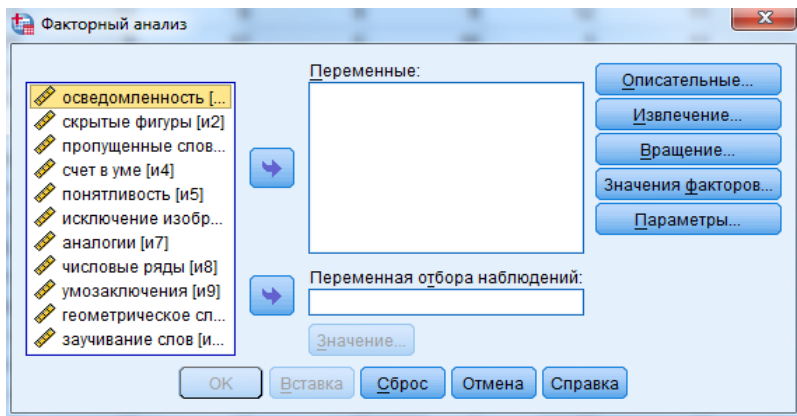


Рис. 12. Диалоговое окно Факторный анализ

3. Щелкните на переменной 1, нажмите клавишу Shift и, не отпуская ее, щелкните на переменной 11. В результате окажутся выделенными все промежуточные переменные, начиная от переменной 1 и заканчивая переменной 11.

4. Щелкните на верхней кнопке со стрелкой, чтобы переместить выделенные переменные в список *Переменные*.

5. Щелкните на кнопке *Описательные статистики*, чтобы открыть диалоговое окно *Факторный анализ: Описательные статистики*.

6. В группе *Статистики* установите флажок *Одномерные описательные статистики*, в группе *Корреляционная матрица* – флажки *Коэффициенты* и *Критерии КМО и сферичности Бартлетта* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ* (рис. 13).

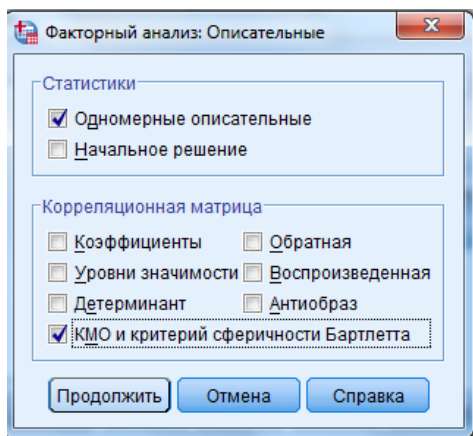


Рис. 13. Диалоговое окно Факторный анализ – Описательные

7. Щелкните на кнопке *Извлечение*, чтобы открыть диалоговое окно *Факторный анализ: Извлечение*.

8. Установите флажок *График собственных значений* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ* (рис. 14).

9. Щелкните на кнопке *Вращение*, чтобы открыть диалоговое окно *Факторный анализ: Вращение*.

10. В группе *Метод* установите переключатель *Varimax*, в группе *Отображать* – флажок *Графики нагрузок* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ* (рис. 15).

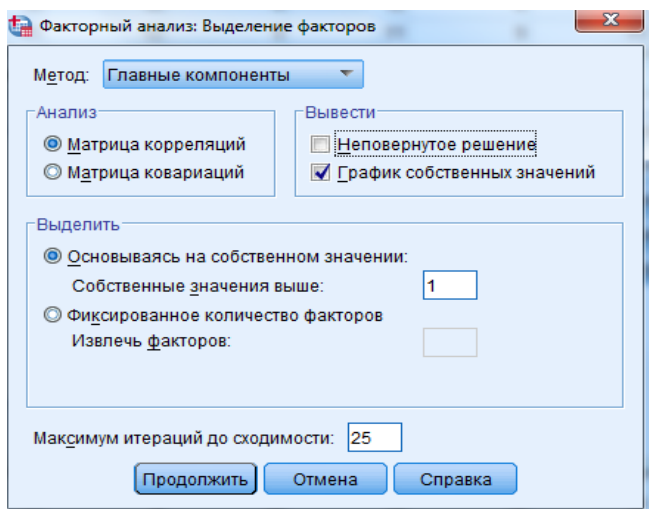


Рис. 14. Диалоговое окно Факторный анализ – Извлечение

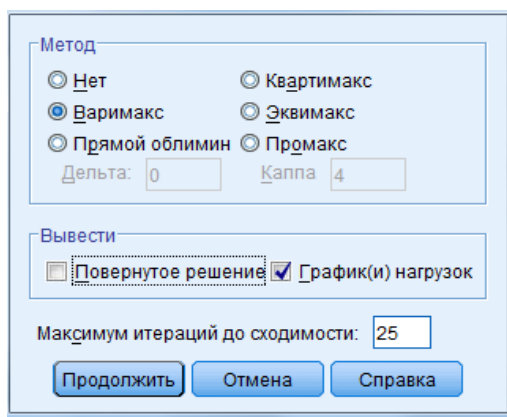


Рис. 15. Диалоговое окно Факторный анализ – Вращение

11. Щелкните на кнопке *Параметры*, чтобы открыть диалоговое окно *Факторный анализ: Параметры*.

12. Установите флажок *Отсортировать по величине* и *Не выводить коэффициенты с низкими значениями*. В поле *Абсолютное значение меньше* вместо 0,10 задайте 0,4. Щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Факторный анализ* (рис. 16).

13. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

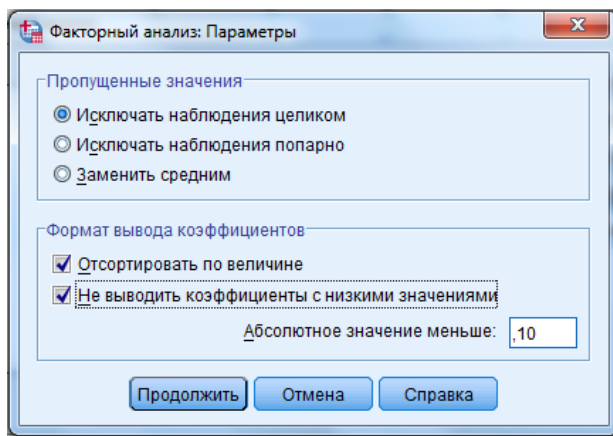


Рис. 16. Диалоговое окно Факторный анализ – Параметры

**Результат.** На выходном листе с результатами для нас представляет интерес таблица *Мера адекватности и критерий Бартлетта* (рис. 17), в которой отражены величина *КМО*, демонстрирующая приемлемую адекватность выборки для факторного анализа, а также критерий сферичности Бартлетта, который показывает статистически достоверный результат ( $p < 0,05$ ): корреляции между переменными значимо отличаются от 0.

#### Мера адекватности и критерий Бартлетта

Мера выборочной адекватности Кайзера-Мейера-Олкина.		,704
Критерий сферичности Бартлетта	Прибл. хи-квадрат	146,069
	ст.св.	55
	Знч.	,000

Рис. 17. Таблица выходного листа SPSS – Мера адекватности и критерий Бартлетта

В таблице *Полная объясненная дисперсия* (рис. 18) содержатся характеристики факторов: суммы квадратов нагрузок, процент общей дисперсии, обусловленной фактором, и соответствующий кумулятивный (накопленный) процент.

Чем больше процент дисперсии, обусловленной фактором, тем больший вес имеет данный фактор. А чем больше кумулятивный

процент, накопленный к последнему фактору, тем более состоятельным является факторное решение. Если этот накопленный процент менее 50 %, следует либо сократить количество переменных, либо увеличить количество факторов. В данном случае накопленный процент дисперсии вполне приемлем.

Полная объясненная дисперсия

Компонента	Суммы квадратов нагрузок вращения		
	Итого	% Дисперсии	Кумулятивный %
1	2,649	24,084	24,084
2	2,089	18,989	43,073
3	2,000	18,180	61,253

Метод выделения: Анализ главных компонент.

Рис. 18. Результаты факторного анализа

Приведенная диаграмма называется графиком собственных значений или диаграммой «каменистой осыпи» (рис. 19). Она представляет точки, соответствующие собственным значениям, в пространстве двух координат. Этот тип диаграммы обычно используется при определении достаточного числа факторов перед вращением. При этом руководствуются следующим правилом: оставлять нужно лишь те факторы, которым соответствуют первые точки на графике до того, как кривая станет более пологой.

По умолчанию SPSS вращает все факторы, чьи собственные значения превышают 1; в нашем случае число таких факторов равно 3, а достаточно было бы не трех, а четырех факторов. Преобразованная матрица факторных нагрузок после вращения (рис. 20) является главным итогом факторного анализа и подлежит содержательной интерпретации. Благодаря установке флажка *Отсортировать по величине* в окне *Факторный анализ: Параметры* факторные нагрузки отсортированы следующим образом:

- наибольшие значения нагрузок для каждого фактора сортируются в отдельных блоках;
- внутри каждого блока нагрузки факторов упорядочены по убыванию.

Кроме того, установка флажка *Не выводить коэффициенты с низкими значениями* и задание в поле *Абсолютное значение меньше значения 0,4* позволили вывести только существенные для интерпретации величины факторных нагрузок.



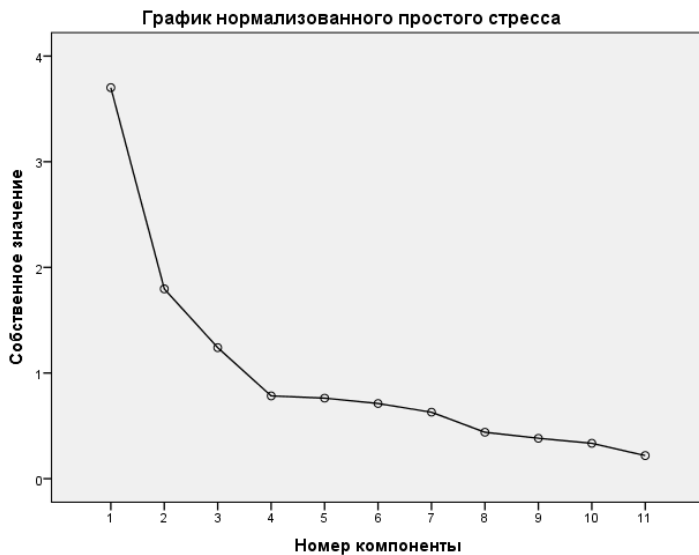


Рис. 19. Диаграмма «каменистой осыпи»

**Матрица повернутых компонент<sup>а</sup>**

	Компонента		
	1	2	3
счет в уме	,804		
анalogии	,769		
числовые ряды	,762		
умозаключения	,696		
заучивание слов		,829	
осведомленность		,746	
пропущенные слова		,630	
понятливость			,739
скрытые фигуры			,737
геометрическое сложение	,421		,692
исключение изображений		,487	,506

Метод выделения: Анализ методом главных компонент  
 Метод вращения: Варимакс с нормализацией Кайзера  
 а. Вращение сошло за 5 итераций

Рис. 20. Матрица факторных нагрузок после вращения

**Вывод.** В результате применения факторного анализа исходные 11 субтестов были сгруппированы в три фактора. Первый из факторов соответствует предполагаемым математическим способностям, так как объединяет субтесты «счет в уме», «анalogии», «числовые ряды» и «умозаключения». Во второй фактор попали три субтеста, относящиеся к вербальным способностям: «заучивание слов», «осведомленность», «пропущенные слова», в третий фактор – три субтеста, относящиеся к невербальным способностям: «скрытые фигуры», «геометрическое сложение», «исключение изображений».

? К «странностям» результатов можно отнести разве что распределение переменной «исключение изображений» между вторым и третьим факторами и попадание переменной «понятливость» в третий фактор. Подобные отклонения обычно требуют отдельного изучения. В частности, можно увеличить число факторов или исключить «неопределенные» переменные и повторить анализ.

### **Контрольные вопросы**

1. В чем состоит основное назначение факторного анализа?
2. Какие методы факторного анализа вам известны? В чем их принципиальное различие?
3. С какими основными проблемами сталкивается исследователь в ходе осуществления факторного анализа? Каким образом их можно решать?
4. Перечислите несколько психологических задач, которые могут быть решены посредством применения факторного анализа.

### **Рекомендуемая литература**

1. Ермолаев-Томин, О.Ю. Математические методы в психологии [Электронный ресурс] : учебник / О.Ю. Ермолаев-Томин. – 5-е изд., испр. и доп. – М. : Юрайт, 2014. – 511 с. – URL : [http://urss.ru/PDF/add\\_ru/159489-1.pdf](http://urss.ru/PDF/add_ru/159489-1.pdf) (Дата обращения 31.10.2016) (С. 91–93, 97–109, 118–133, 134–154).
2. Наследов, А. IBM SPSS Statistics 20 и AMOS [Электронный ресурс] : профессиональный статистический анализ данных / А. Наследов. – СПб. : Питер, 2013. – 416 с. – URL : <http://mexalib.com/view/29545> (Дата обращения 31.10.2016) (главы 11 и 12).

## Практическое занятие 6

### Кластерный анализ данных

**Цель** — ознакомление с алгоритмом применения кластерного анализа эмпирических данных.

#### **Учебные вопросы:**

- кластерный анализ;
- методы кластерного анализа;
- графическое представление результатов кластерного анализа.

Студент должен *знать* назначение и области применения кластерного анализа; *уметь* выбирать методы кластерного анализа в соответствии с поставленной целью и особенностями выборки; *владеть* навыками выполнения кластерного анализа данных.

#### **Методические рекомендации по освоению теоретического материала по теме**

Кластерный анализ решает задачу построения классификации, т. е. разделения исходного множества объектов на группы (классы, кластеры). Предполагается, что у исследователя нет исходных допущений ни о составе классов, ни об их отличии друг от друга. Приступая к кластерному анализу, исследователь располагает лишь информацией о характеристиках (признаках) объектов, позволяющей судить о сходстве / различии этих объектов, или только данными об их попарном сходстве / различии.

**Кластерный анализ** — это процедура упорядочивания объектов в сравнительно однородные классы на основе попарного сравнения этих объектов по предварительно определенным и измеренным критериям. Существует большое количество вариантов кластерного анализа, но наиболее распространенными в психологических исследованиях являются методы, объединенные общим названием **иерархический кластерный анализ**. Применение кластерного анализа наиболее адекватно при решении следующих исследовательских задач:

- разбиение совокупности испытуемых на группы по измеренным признакам с целью дальнейшей проверки причин межгрупповых различий по внешним критериям;
- группировка признаков на основе их корреляции;

– классификация объектов на основе непосредственных оценок различий между ними.

Осуществление кластерного анализа предполагает последовательность из следующих относительно самостоятельных шагов:

- отбор объектов для кластеризации. В качестве таковых могут выступать испытуемые; объекты, оцененные испытуемыми; признаки испытуемых;
- определение множества переменных, по которым будут различаться объекты кластеризации;
- определение меры различия между объектами кластеризации;
- выбор и применение метода классификации для создания групп сходных объектов;
- проверка достоверности разбиения на классы.

Графически результаты кластеризации можно представить в виде дендрограммы или диаграммы накопления [16, с. 194–197; 17, с. 329–336].

## **Практические задания**

### **Задача 1**

На выборке специалистов экономического профиля для трудового поста, занимаемого испытуемым на предприятии, по специальной методике определены девять информационно-психологических параметров. Данные параметры, измеряемые в интервальной шкале (в станайнах), характеризуют трудовой пост в плане информационных требований, предъявляемых им к специалисту, и имеют следующую интерпретацию: X1 и X2 – соответственно содержательно-управленческий и информационно-технологический потенциал информационного взаимодействия; X3, X4 и X5 – соответственно учетно-аналитическая, социономическая и знаково-техническая направленность труда; X6 – коллективность субъекта труда, X7 – неординарность деятельности в классе трудовых постов экономиста, X8 – интенсивность труда, X9 – социальная ценность труда для работника. Требуется: 1) построить дендрограммы иерархической кластеризации обследованной выборки, используя евклидово межкластерное расстояние и метод Уорда; 2) определить, чему равно расстояние между случаями с номерами 1 и 8.

**Информационно-психологические параметры трудовых постов  
специалистов экономического профиля**

№	X1	X2	X3	X4	X5	X6	X7	X8	X9
1	3	2	5	7	3	4	3	6	3
2	4	2	6	7	4	3	4	4	4
3	3	1	5	8	4	4	3	5	4
4	4	2	7	7	4	3	2	4	5
5	4	2	6	9	3	2	4	5	4
6	4	3	5	7	4	3	4	6	3
7	5	2	5	8	2	4	3	5	4
8	5	6	5	4	4	7	6	5	7
9	5	6	6	4	5	5	5	3	4
10	5	6	5	4	5	5	5	5	7
11	5	5	5	4	6	5	5	4	4
12	5	6	6	4	6	6	6	5	5
13	5	6	6	5	6	7	7	5	7
14	4	5	5	5	6	7	7	4	4
15	6	7	6	5	5	5	5	5	6
16	6	7	5	5	6	7	7	5	5
17	4	7	4	6	6	5	5	4	5
18	6	7	6	4	6	6	5	5	5
19	4	5	4	6	6	5	5	4	5
20	6	7	4	5	6	7	5	5	7
21	5	6	4	5	5	6	7	5	6
22	5	7	6	5	4	5	5	5	7
23	5	6	6	4	5	5	5	4	5
24	4	5	5	5	5	6	5	3	7
25	4	5	5	4	6	5	6	5	5
26	6	7	5	4	4	7	6	5	7
27	5	6	6	4	5	5	5	3	4
28	6	7	5	5	5	5	6	4	7
29	5	5	5	4	6	5	5	4	4
30	5	6	6	4	6	6	6	4	5
31	6	5	5	4	6	7	6	4	7
32	7	5	5	6	4	7	6	7	7
33	7	5	5	6	2	5	6	6	6
34	6	5	7	7	4	5	6	6	5
35	8	5	6	5	2	5	5	6	7
36	8	4	5	7	4	4	7	8	6

№	X1	X2	X3	X4	X5	X6	X7	X8	X9
37	7	7	8	6	2	5	5	7	7
38	6	7	5	5	3	7	8	8	6
39	7	5	8	7	2	5	5	6	7
40	8	6	6	7	4	6	8	8	5
41	8	4	5	7	3	6	6	7	6
42	6	7	5	5	2	5	8	8	5
43	8	6	5	6	2	7	8	6	5
44	6	6	7	6	3	4	8	6	7
45	7	7	5	5	2	4	7	8	6
46	8	4	5	8	4	4	8	8	5
47	6	6	6	5	5	6	7	8	6
48	7	5	5	6	3	7	7	8	7
49	6	5	7	7	4	5	6	6	6
50	8	5	6	8	3	5	5	6	7
51	8	4	5	7	4	4	7	8	7
52	7	5	8	6	5	5	5	7	7

### Задача 2

По данным таблицы и алгоритму кластерного анализа провести классификацию объектов иерархическим методом (древовидная кластеризация). В качестве расстояния между объектами принять «обычное евклидово расстояние», а расстояния между кластерами измерять по принципу «ближайшего соседа». В ответе указать: а) количество кластеров; б) страны, вошедшие в каждый кластер.

### Эмпирические данные

№ п/п	Страны	Число специалистов психологического профиля на 10000 населения	Психопатологическая симптоматика на 100000 населения	ВВП по паритету покупательной способности, в % к США	Расходы на здравоохранение, в % к США
		X1	X2	X3	X4
1	Россия	44,5	84,98	20,4	3,2
2	Австралия	32,5	30,58	71,4	8,5
3	Австрия	33,9	38,42	78,7	9,2
4	Азербайджан	38,8	60,34	12,1	3,3
5	Армения	34,4	60,22	10,9	3,2
6	Беларусь	43,6	60,79	20,4	5,4

№ п/п	Страны	Число специалистов психологического профиля на 10000 населения	Психопатологическая симптоматика на 100000 населения	ВВП по паритету покупательной способности, в % к США	Расходы на здравоохранение, в % к США
		X1	X2	X3	X4
7	Бельгия	41	29,82	79,7	8,3
8	Болгария	36,4	70,57	17,3	5,4
9	Великобритания	17,9	34,51	69,7	7,1
10	Венгрия	32,1	64,73	24,5	6
11	Германия	38,1	36,63	76,2	8,6
12	Греция	41,5	32,84	44,4	5,7
13	Грузия	55	62,64	11,3	3,5
14	Дания	36,7	34,07	79,2	6,7
15	Ирландия	15,8	39,27	57	6,7
16	Испания	40,9	28,46	54,8	7,3
17	Италия	49,4	30,27	72,1	8,5
18	Казахстан	38,1	69,04	13,4	3,3
19	Канада	27,6	25,42	79,9	10,2
20	Киргизия	33,2	53,13	11,2	3,4

### **Методические рекомендации и образцы выполнения заданий. Алгоритм осуществления кластерного анализа**

1. Ознакомьтесь с условиями задачи. Сформулируйте статистические гипотезы.

2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.

3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче.

4. В меню *Анализ* выберите команду *Классификация – Иерархическая кластеризация*. Откроется диалоговое окно *Иерархический кластерный анализ*.

5. Щелкните сначала на первой по очереди переменной, а затем – на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Метить значениями*.

6. Щелкните на второй по очереди переменной и, нажав на клавиатуре кнопку Shift, щелкните на последней переменной. В результате окажутся выделенными все оставшиеся в списке переменные. Щелкните на верхней кнопке со стрелкой, чтобы переместить выделенные переменные в список *Переменные*.

7. Щелкните на кнопке *Графики*, чтобы открыть диалоговое окно *Иерархический кластерный анализ: Графики*.

8. Установите флажок *Дендрограмма*. Выберите предпочтительную ориентацию диаграммы: вертикальную или горизонтальную. Щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Иерархический кластерный анализ*.

9. Щелкните на кнопке *Метод*, чтобы открыть диалоговое окно *Иерархический кластерный анализ: Метод*.

10. В списке *Метод* выберите тот, что требуется условиями задачи, в списке *Стандартизация* выберите пункт *z-значения* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Иерархический кластерный анализ*.

11. Щелкните на кнопке *Сохранить*, чтобы открыть диалоговое окно *Иерархический кластерный анализ: Сохранить*.

12. Установите переключатель *Одно решение*, введите в расположенное рядом поле значение количества предполагаемых вами кластеров и щелкните на кнопке *Продолжить*, чтобы вернуться в *Иерархический кластерный анализ*.

13. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

14. Проанализировав результат, сделайте статистический вывод о том, какая из гипотез подтвердилась. После этого сформулируйте качественный вывод и оформите ответ на задание.

### **Пример**

**Условие задачи.** Имеются данные о результатах тестирования 15 представителей автомобильных концернов (закодированы маркой автомобиля) на предмет профессиональной пригодности ( $X_1$ ,  $X_2$ ) и удовлетворенности трудом ( $X_3$ ,  $X_4$ ). Проведите кластерный анализ с несколькими дополнительными параметрами. В качестве идентификатора используйте переменную *марка*, которая указывает на принадлежность респондента к определенному автомобилю-



ному концерну. Все остальные переменные файла задействуйте для вычисления расстояния между объектами. Включите в выводимые результаты последовательность слияния и дендрограмму. В качестве расстояния между объектами задайте квадрат евклидова расстояния, а в качестве метода кластеризации – межгрупповое связывание.

### *Эмпирические данные*

Марка	X1	X2	X3	X4
Audi	16350	7	3	66000
BMW	14500	8	4	92500
Buick	8950	9	8	92500
Chrysler	8950	6	6	92500
Dodge	8450	5	6	92500
Honda	9850	7	4	118500
Mazda	12650	7	8	58000
Mercedes	17250	8	5	92500
Mitsubishi	8950	3	5	136000
Nissan	9850	6	4	150000
Pontiac	8950	7	6	110000
Saab	14950	7	4	101000
Toyota	11700	8	7	57500
VW	8450	5	6	110000
Volvo	13100	8	4	75000

### **Алгоритм решения задачи-примера**

1. Ознакомьтесь с условиями задачи. Сформулируйте гипотезы.
2. Запустите программу SPSS при помощи значка на рабочем столе или выбрав в главном меню Windows команду *Пуск – Программы – IBM SPSS Statistics*. После выполнения этого шага на экране появится окно редактора данных SPSS.
3. С помощью копирования или ручным набором внесите в окно редактора первичные данные, представленные в задаче.
4. В меню Анализ выберите команду *Классификация – Иерархическая кластеризация*. Откроется диалоговое окно Иерархический кластерный анализ (рис. 21).

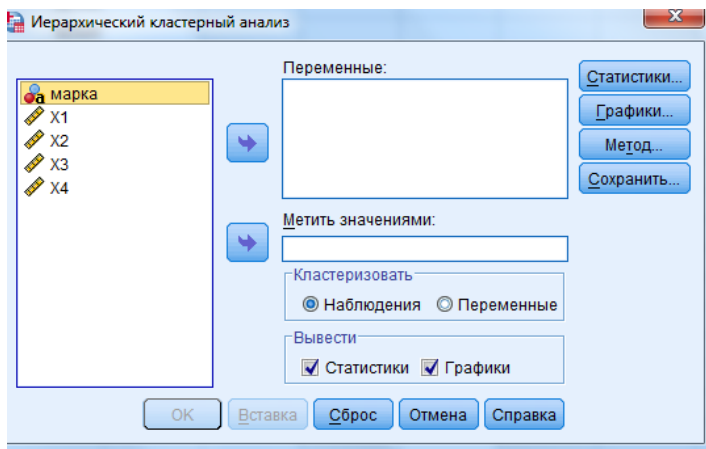


Рис. 21. Диалоговое окно Иерархический кластерный анализ

5. Щелкните сначала на переменной *марка*, чтобы выделить ее, а затем – на нижней кнопке со стрелкой, чтобы переместить переменную в поле *Метить значениями*.

6. Нажмите кнопку мыши на переменной X1 и, не отпуская кнопки, перетащите указатель на переменную X4, затем кнопку мыши отпустите. В результате окажутся выделенными все оставшиеся в списке переменные. Щелкните на верхней кнопке со стрелкой, чтобы переместить выделенные переменные в список *Переменные* (рис. 22).

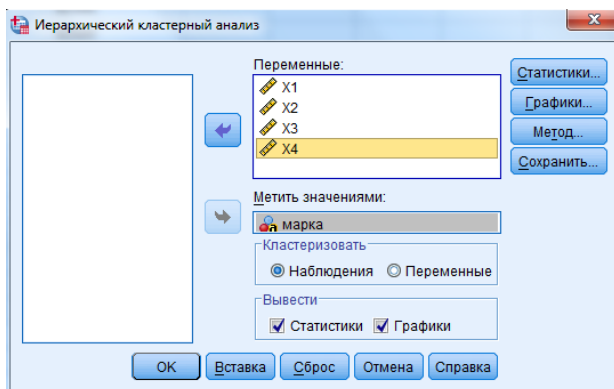


Рис. 22. Диалоговое окно Иерархический кластерный анализ

7. Щелкните на кнопке *Графики*, чтобы открыть диалоговое окно *Иерархический кластерный анализ: Графики*.

8. Установите флажок *Дендрограмма* и переключатель *Нет* в группе *Сосульчатая диаграмма*. Щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Иерархический кластерный анализ* (рис. 23).

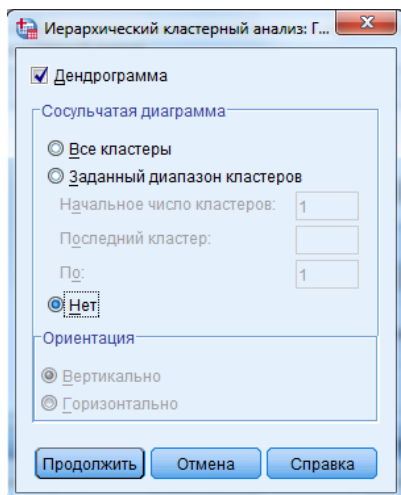


Рис. 23. Диалоговое окно *Иерархический кластерный анализ – Графики*

9. Щелкните на кнопке *Метод*, чтобы открыть диалоговое окно *Иерархический кластерный анализ: Метод*.

10. В списке *Метод кластеризации* оставьте выбранным пункт *Межгрупповые связи*, в списке *Стандартизация* выберите пункт *z-значения* и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Иерархический кластерный анализ* (рис. 24).

11. Щелкните на кнопке *Сохранить*, чтобы открыть диалоговое окно *Иерархический кластерный анализ: Сохранение новых переменных*.

12. Установите переключатель в поле *Одно решение* и введите в расположенное рядом поле значение 3 и щелкните на кнопке *Продолжить*, чтобы вернуться в диалоговое окно *Иерархический кластерный анализ* (рис. 25).

13. Щелкните на кнопке *ОК*, чтобы открыть окно вывода.

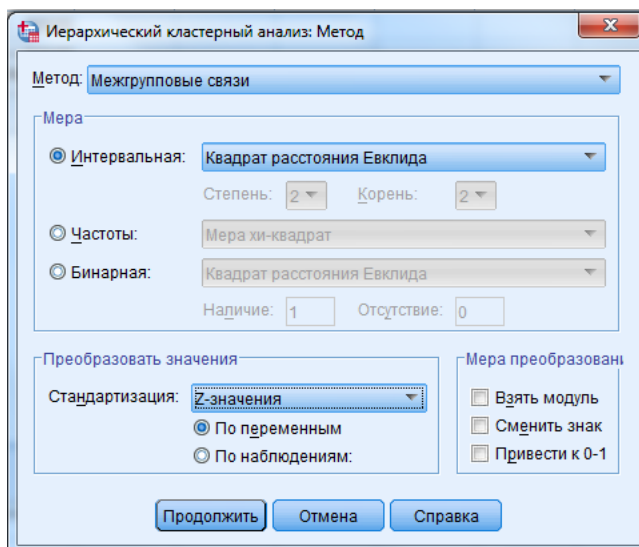


Рис. 24. Диалоговое окно Иерархический кластерный анализ – Метод

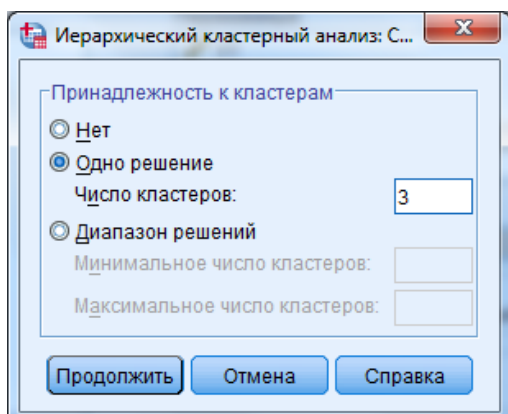


Рис. 25. Диалоговое окно Иерархический кластерный анализ – Сохранить

**Результат.** Файл с выходными данными примера является небольшим, поэтому удобно включить все выходные данные в вывод и использовать их при интерпретации результатов кластерного анализа. Частью выводимых результатов является таблица последовательности слияния – *Шаги агломерации* (рис. 26). В этой таблице вторая колонка *Кластер объединен с* содержит первый и второй столбцы *Кластер 1* и *Кластер 2*

2, которые соответствуют номерам кластеров, объединяемых на данном шаге. После объединения кластеру присваивается номер, соответствующий номеру в колонке Кластер 1. Так, на первом шаге происходит объединение объектов 5 и 14 и кластеру присваивается номер 5, далее на шаге 3 этот кластер объединяется с элементом 4 и новому кластеру присваивается номер 4 и т. д. Следующая колонка *Коэффициенты* содержит значение расстояния между кластерами, которые объединяются на данном шаге. Так, расстояние между объектами 5 и 14 является наименьшим (0,439). На втором этапе SPSS снова подсчитывает расстояния между объектами и объединяет в кластер пару наиболее близких объектов, в результате может получиться либо один кластер из трех объектов, либо два кластера из двух объектов. Процесс слияния продолжается до тех пор, пока все объекты не попадут в один кластер.

Шаги агломерации

Этап	Кластер объединен с		Коэффициенты	Этап первого появления кластера		Следующий этап
	Кластер 1	Кластер 2		Кластер 1	Кластер 2	
1	5	14	,439	0	0	3
2	2	12	,550	0	0	5
3	4	5	,671	0	1	6
4	7	13	,942	0	0	10
5	2	15	1,203	2	0	7
6	4	11	1,586	3	0	11
7	2	8	1,810	5	0	9
8	6	10	1,847	0	0	11
9	1	2	2,471	0	7	13
10	3	7	4,136	0	4	13
11	4	6	4,492	6	8	12
12	4	9	5,914	11	0	14
13	1	3	9,656	9	10	14
14	1	4	10,498	13	12	0

Рис. 26. Таблица Шаги агломерации

Колонка *Этап первого появления кластера* показывает, на каком шаге до этого появлялись первый и второй объединяемые кластеры.

В данном случае ни один из двух объектов не принадлежит какому-либо кластеру, о чем свидетельствуют нули в столбцах Кластер 1 и Кластер 2 колонки *Этап первого появления кластера*.

Последняя колонка *Следующий этап* показывает, на каком шаге снова появится кластер, образованный на этом шаге. Следующим этапом для данного кластера, судя по столбцу *Следующий этап*, является этап 3, на котором к кластеру присоединяется объект 4. На этапе 13 происходит объединение кластеров, содержащих объекты 1 и 3. Объект 1 был объединен с кластером, содержащим объект 2 на этапе 9, а объект 3 с объектами 7 и 13 на этапе 10. Расстояние между объединяемыми на этом этапе кластерами равно 9,656. Образованный на этом этапе кластер появляется далее на следующем шаге.

По таблице *Шаги агломерации* (рис. 26) можно предварительно оценить число кластеров. Для этого необходимо проследить динамику увеличения расстояний по шагам кластеризации и определить шаг, на котором отмечается резкое возрастание расстояний. Оптимальному числу классов соответствует разность между числом объектов и порядковым номером шага, на котором было обнаружено резкое возрастание расстояний. В данном случае резкое возрастание различий обнаруживается при переходе от шага 12 к шагу 13. Следовательно, наиболее оптимальное количество кластеров должно быть получено на шаге 12 или 13. Это количество равно численности объектов минус номер шага, то есть  $15 - 12 = 3$  или  $15 - 13 = 2$ , то есть 3 или 2 кластера. Выбор того или иного решения зависит от конкретной ситуации.

На дендрограмме (рис. 27) любое решение характеризуется вертикальной линией, число точек пересечения которой с деревом соответствует количеству кластеров на текущем этапе.

Для нашего примера эту линию следует расположить на уровне 15–20: между шагами 12 и 13 кластеризации. В этом случае получается три кластера. Для того чтобы установить состав каждого кластера, необходимо вернуться к корням дерева и выяснить соответствующие номера объектов. На листе с входными данными в SPSS появилась новая переменная, содержащая сведения о принадлежности каждого объекта к одному из трех кластеров.

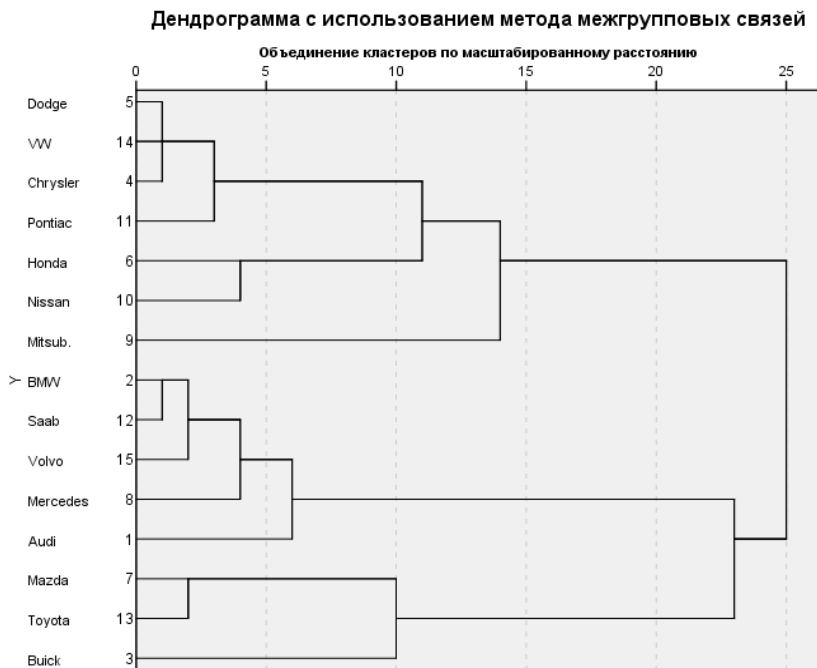


Рис. 27. Дендрограмма

**Вывод.** Представители автомобильных концернов после проведения тестирования на профессиональную пригодность и удовлетворенность трудом могут быть сгруппированы в три кластера.

### Контрольные вопросы

1. В чем состоит основное назначение кластерного анализа?
2. Какие виды иерархического кластерного анализа вам известны?
3. Какие проблемы сопровождают процедуру проведения кластерного анализа?
4. Перечислите несколько психологических задач, которые могут быть решены посредством применения кластерного анализа.

### Рекомендованная литература

1. Ермолаев-Томин, О.Ю. Математические методы в психологии [Электронный ресурс] : учебник / О.Ю. Ермолаев-Томин. – 5-е изд., испр. и доп. – М. : Юрайт, 2014. – 511 с. – URL : [http://urss.ru/PDF/add\\_ru/159489-1.pdf](http://urss.ru/PDF/add_ru/159489-1.pdf) (Дата обращения 31.10.2016) (С. 91–93, 97–109, 118–133, 134–154).
2. Наследов, А. IBM SPSS Statistics 20 и AMOS [Электронный ресурс] : профессиональный статистический анализ данных / А. Наследов. – СПб. : Питер, 2013. – 416 с. – URL : <http://mexalib.com/view/29545> (Дата обращения 31.10.2016) (главы 11 и 12).



## Итоговый тест

1. Дисперсия измеряет

- а) разброс значений относительно медианы
- б) разброс значений относительно среднего
- в) разницу между максимальным и минимальным значениями ряда

2. Уровень значимости — это

- а) ошибка первого рода
- б) ошибка второго рода
- в) мощность критерия
- г) критическое значение

3. Коэффициент корреляции измеряет

- а) зависимость  $x$  от  $y$
- б) зависимость  $y$  от  $x$
- в) одновременно зависимость  $x$  от  $y$  и  $y$  от  $x$

4. В каком диапазоне может изменяться коэффициент корреляции?

- а) от  $-1$  до  $1$
- б) от  $0$  до  $1$
- в) от  $0$  до  $100$

5. Проверить статистическую гипотезу означает

- а) проверить, равно ли значение параметра генеральной совокупности определенному числу
- б) проверить, равно ли значение параметра выборки определенному числу
- в) сделать выводы о распределении параметра генеральной совокупности
- г) сделать выводы о распределении параметра выборки

6.  $t$ -критерий Стьюдента является

- а) параметрическим
- б) непараметрическим

7. Дисперсионный анализ — это

- а) проверка гипотезы о корреляции между градациями фактора
- б) проверка гипотезы о различиях математических ожиданий градаций

- в) проверка гипотезы о различиях в уровнях градаций фактора
- г) проверка гипотезы о различиях в сдвиге значений для разных градаций

**8. Регрессионный анализ предназначен**

- а) для моделирования стохастических процессов
- б) построения функциональной зависимости детерминированных процессов
- в) выявления степени взаимосвязи между переменными

**9. Факторная нагрузка — это**

- а) субъективная трудность выполнения теста или тестового задания для испытуемого
- б) степень влияния определенного свойства (фактора) на наблюдаемую переменную
- в) коэффициент статистической связи между двумя латентными переменными
- г) вклад отдельного вопроса в общую дисперсию свойства

**10. Факторный анализ предназначен**

- а) для прогноза интегрального показателя по его составным частям
- б) объединения исходных признаков в группы на основании их близости
- в) объединения объектов выборки в группы по степени близости
- г) прогноза принадлежности объекта к заранее выделенному классу объектов

**11. Результатом метода иерархической классификации является**

- а) дерево классификации
- б) цепочка объектов
- в) пространство объектов

**12. Кластерный анализ используется**

- а) для объединения исходных признаков в группы на основании их близости
- б) объединения объектов выборки в группы по степени близости
- в) прогноза интегрального показателя по его составным частям
- г) прогноза принадлежности объекта к заранее выделенному классу объектов

**13.** К параметрическим критериям оценки значимости различий на уровне исследуемого признака относится

- а)  $Q$ -критерий Розенбаума
- б)  $t$ -критерий Стьюдента
- в)  $U$ -критерий Манна – Уитни

**14.** Если необходимо выявить взаимосвязь между несколькими независимыми переменными, измеренными в ранговой шкале, то можно использовать

- а) коэффициент корреляции Пирсона
- б) биссерийальный коэффициент корреляции
- в) коэффициент корреляции Спирмена

**15.** Если эмпирическое значение  $t$ -критерия Стьюдента меньше значения табличного при  $p = 0,05$ , то различия между выборками статистически

- а) достоверны
- б) недостоверны
- в) «зона неопределенности»

**16.** Если необходимо выявить причинно-следственные связи между двумя и более признаками, то можно использовать

- а) дисперсионный анализ
- б) корреляционный анализ
- в) угловое преобразование Фишера

**17.** Выявить, соответствует ли эмпирическое распределение результатов нормальному распределению, возможно с помощью

- а) критерия Колмогорова – Смирнова для парных выборок
- б) критерия Мозеса
- в) критерия Колмогорова – Смирнова для одной выборки

**18.** Статистически определить причинно-следственные связи позволяет

- а) коэффициент корреляции
- б) определение значимости различий на уровне исследуемого признака
- в) определение значимости сдвига на уровне исследуемого признака
- г) факторный анализ

**19.** Если необходимо выявить достоверность различий между результатами измерений, проведенных с помощью одной и той же методики на одной и той же группе испытуемых, при условии, что распределение признака не выходит за рамки нормального, то можно использовать

- а)  $t$ -критерий Стьюдента для парных выборок
- б)  $t$ -критерий Стьюдента для самостоятельных выборок
- в)  $t$ -критерий Стьюдента для одной выборки

**20.** Древовидная фигура, где более близкие ветви соответствуют более близким объектам или группам объектов, называется

- а) дендрограмма
- б) гистограмма
- в) кластер

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

### *Основная литература*

1. Дьячук, А.А. Математические методы в психологических и педагогических исследованиях : учебное пособие / А.А. Дьячук. — Красноярск : Красноярский гос. пед. ун-т им. В.П. Астафьева, 2013. — 347 с.
2. Ермолаев-Томин, О.Ю. Математические методы в психологии : учебник / О.Ю. Ермолаев-Томин. — 5-е изд., испр. и доп. — М. : Юрайт, 2014. — 511 с.
3. Кричевец, А.Н. Математическая статистика для психологов : учебник для студ. учреждений высш. проф. образования / А.Н. Кричевец, А.А. Корнеев, Е.И. Рассказова. — М. : Академия, 2012. — 400 с.
4. Лупандин, В.И. Математические методы в психодиагностике / В.И. Лупандин. — Екатеринбург : Изд-во Уральского ун-та, 2012. — 86 с.
5. Наследов, А.Д. IBM SPSS Statistics 20 и AMOS : профессиональный статистический анализ данных / А.Д. Наследов. — СПб. : Питер, 2013. — 416 с.
6. Остапенко, Р.И. Многомерный анализ данных для психологов : учеб.-метод. пособие / Р.И. Остапенко. — Воронеж : ВГПУ, 2012. — 72 с.
7. Романов, В.П. Вероятностно-статистический метод психолого-педагогических исследований / В.П. Романов, Н.А. Соколова. — М. : Ладомир, 2012. — 144 с.
8. Слепко, Ю.Н. Анализ данных и интерпретация результатов психологического исследования : учебное пособие / Ю.Н. Слепко, Т.В. Ледовская. — Ярославль : Изд-во ЯГПУ, 2013. — 136 с.

### *Дополнительная литература*

9. Гусев, А.Н. Психологические измерения. Теория. Методы : учеб. пособие / А.Н. Гусев, М.С. Уточкин. — М. : Аспект-Пресс, 2011. — 320 с.
10. Двоерядкина, Н.Н. Использование статистических методов в психологии и социологии : учеб. пособие / Н.Н. Двоерядкина, А.Н. Киселева, Т.А. Юрьева. — Благовещенск : Изд-во Амур. гос. ун-та, 2010. — 116 с.

11. Дорогонько, Е.В. Обработка и анализ социологических данных с помощью пакета SPSS : учеб.-метод. пособие / Е.В. Дорогонько. – Сургут : СурГУ, 2010. – 60 с.
12. Ермолаев, О.Ю. Математическая статистика для психологов : учебник / О.Ю. Ермолаев. – 5-е изд. – М. : МПСИ: Флинта, 2011. – 336 с.
13. Карманов, Ф.И. Статистические методы обработки экспериментальных данных : учеб. пособие / Ф.И. Карманов, В.А. Острейковский. – М. : Абрис, 2011. – 208 с.
14. Кутейников, А.Н. Математические методы в психологии : учеб.-метод. комплекс / А.Н. Кутейников. – СПб. : Речь, 2008. – 172 с.
15. Леонов, Н.И. Многомерные статистические методы анализа данных в психологических исследованиях / Н.И. Леонов, М.М. Главатских. – М. : МПСИ, 2011. – 128 с.
16. Митина, О.В. Математические методы в психологии : практикум / О.В. Митина. – М. : Аспект Пресс, 2008. – 238 с.
17. Наследов, А.Д. Математические методы психологического исследования : анализ и интерпретация данных / А.Д. Наследов. – СПб. : Речь, 2004 – 1-е изд., 2011 – 4-е изд. – 390 с.
18. Наследов, А.Д. SPSS 19 : профессиональный статистический анализ данных / А.Д. Наследов. – СПб. : Питер, 2011. – 400 с.
19. Остапенко, Р.И. Математические основы психологии : учеб.-метод. пособие / Р.И. Остапенко. – Воронеж : ВГПУ, 2010. – 76 с.
20. Погребницкая, М.В. Математические методы в психологии : учеб.-метод. пособие / М.В. Погребницкая. – Петропавловск : СКГУ, 2004. – 228 с.
21. Сидоренко, Е.В. Методы математической обработки в психологии / Е.В. Сидоренко. – СПб. : Речь, 2003. – 350 с.
22. Титкова, Л.С. Математические методы в психологии : учеб. пособие / Л.С. Титкова. – Владивосток : ТИДОТ ДВГУ, 2002. – 140 с.
23. Чашкин, Ю.Р. Математическая статистика. Анализ и обработка данных : учеб. пособие / Ю.Р. Чашкин ; ред. С.Н. Смоленский. – Ростов н/Д : Феникс, 2010. – 236 с.

### *Интернет-ресурсы*

1. <http://www.twirpx.com/file/1236468/> – файлообменник с возможностью бесплатного скачивания учебных изданий.
2. <http://www.klex.ru/8bz> – книжный архив.
3. <http://wiki.uspi.ru/index.php/> – возможности программы SPSS.
4. <http://padabum.com/d.php?id=59256> – с помощью этой ссылки можно скачать и установить программу SPSS 19.
5. <http://www.psychology-online.net/314/> – книги по математическим методам в психологии on-line.
6. [http://www.hse.ru/data/2011/01/16/1208066856/method\\_MODP\\_SPSS\\_2011-01-16.pdf](http://www.hse.ru/data/2011/01/16/1208066856/method_MODP_SPSS_2011-01-16.pdf) – электронное методическое пособие по работе в программе SPSS.
7. <http://www.learnspss.ru/aufgabe.htm> – типичные задачи и примеры их решения в программе SPSS.
8. <http://any-book.org/download/80404.html> – электронное обучающее пособие по работе в программе SPSS.
9. <http://rutracker.org/forum/viewtopic.php?t=3767645> – ссылка на скачивание IBM SPSS Statistics 20.
10. <http://www.pdfs.me/lib> – учебные материалы по статистическим методам в психологии.
11. <http://www.spsstools.ru/spss.htm> – он-лайн практикумы по SPSS.

## ГЛОССАРИЙ

***F*-критерий** — в дисперсионном анализе отношение межгруппового среднего квадрата к внутригрупповому среднему квадрату. Данная величина позволяет сравнить межгрупповую дисперсию с внутригрупповой дисперсией. Если первая окажется значительно выше второй, это будет означать наличие значимого различия между группами. В множественном регрессионном анализе *F*-критерий позволяет определить значимость множественной корреляции.

***R*** — множественный коэффициент корреляции между зависимой переменной и двумя или более независимыми переменными. Значение *R* лежит в пределах от 0 до 1 и интерпретируется по аналогии с обычным (двухмерным) коэффициентом корреляции. ***R*<sup>2</sup>** — квадрат коэффициента множественной корреляции (коэффициент детерминации), доля дисперсии зависимой переменной, обусловленная воздействием двух или более независимых переменных.

***t*-критерий** — критерий для определения статистической значимости различия двух средних. ***t*-критерий в регрессионном анализе** — критерий, определяющий статистическую значимость корреляций, равен отношению коэффициента *B* к своей стандартной ошибке. ***t*-критерий для зависимых выборок** — критерий, сравнивающий средние значения двух распределений для одной и той же выборки. ***t*-критерий для независимых выборок** — критерий, сравнивающий средние значения одной и той же переменной для двух независимых выборок.

**Альтернативная гипотеза** — одна из статистических гипотез. Формулируется как противоположная нулевой гипотезе. Принимается в случае отклонения основной гипотезы. Может быть направленной, когда важно установить, в каком направлении происходит отклонение от равенства сопоставляемых параметров, и ненаправленной, когда специально не проверяется предположение о направлении.

**Выборка** — подмножество объектов из некоторой генеральной совокупности, выбранное для статистических выводов относительно свойств всей совокупности.

**Гистограмма** — столбиковая диаграмма для отображения распределения частот по категориям (диапазонам значений) переменной. Горизонтальная ось графика соответствует значениям переменной, а вертикальная — частотам.

**Дендрограмма** — диаграмма древовидной структуры, иллюстрирующая процесс кластеризации в кластерном анализе.



**Дисперсионный анализ (ANOVA)** – статистический анализ, устанавливающий статистическую значимость различий между средними значениями для трех или более выборок. **Многомерный дисперсионный анализ (MANOVA)**. Отличие многомерного дисперсионного анализа от одномерного заключается в том, что число зависимых переменных в нем может быть теоретически любым.

**Зависимые выборки** – выборки, в которых каждому объекту одной выборки поставлен в соответствие по определенному критерию объект другой выборки. Имеют одинаковый объем.

**Значимость ( $p$ -уровень)** – мера случайности полученного результата, равная вероятности того, что в генеральной совокупности этот результат (различия, связь) отсутствует. Чем меньше эта вероятность (значение  $p$ -уровня), тем выше статистическая значимость результата. Результат считается статистически достоверным (значимым), если  $p$ -уровень не превышает 0,05.

**Измерение** – процедура установления взаимно-однозначного соответствия между множеством объектов и множеством символов. В классическом понимании измерение – это процедура сравнения измеряемого объекта с эталоном, при котором он получает численное выражение в определенном масштабе или шкале.

**Кластер** – группа схожих, однородных объектов, которая может рассматриваться как самостоятельная единица, обладающая определенными свойствами. Результат кластерного анализа.

**Кластерный анализ** – процедура, на основе заданного правила объединяющая объекты или переменные в группы, называемые кластерами.

**Корреляция** – мера степени и направления связи между значениями двух переменных. **Коэффициент корреляции** – мера связи двух переменных, обозначаемая символом  $r$  и принимающая значения от  $-1$  до  $+1$ .

**Корреляционная матрица** – это результат вычисления корреляций одного типа для каждой пары из множества  $P$  переменных, измеренных на одной выборке. Применяется в корреляционных исследованиях и факторном анализе.

**Манна – Уитни и Вилкоксона критерий ранговых сумм** – непараметрический аналог  $t$ -критерия, определяющий различие между двумя выборками на основе рангов.

**Метод главных компонент** — метод, применяемый SPSS по умолчанию в факторном анализе для извлечения факторов.

**Метод иерархического слияния** — метод, используемый в кластерном анализе, при выполнении которого объекты объединяются в кластеры по одному на каждом шаге до тех пор, пока не будет образован единственный кластер, охватывающий все объекты.

**Мощность критерия** — величина, характеризующая способность критерия отклонять нулевую гипотезу, когда она неверна.

**Независимые выборки** — выборки, объекты которых не поставлены ни в какое соответствие друг с другом. Вероятность отбора объектов в одну выборку не зависит от вероятности отбора объекта в другую выборку. Могут иметь разные объемы.

**Непараметрические критерии** — серия критериев, каждый из которых применяется без предварительных допущений относительно нормальности распределения. Непараметрические критерии основаны на ранжировании, попарных сравнениях и других средствах, не требующих нормальности распределения переменных.

**Нормальное распределение** — распределение частот (вероятностей), графически представляемое в виде симметричной кривой, имеющей пик в центре, и асимптотически приближающееся к горизонтальной оси по краям. Идеальное нормальное распределение характеризуется нулевыми значениями асимметрии и эксцесса.

**Нулевая гипотеза** — основная гипотеза, проверяемая на выборочной совокупности о неизвестных параметрах случайной величины. Рассматривается как наиболее вероятное решение о случайной величине на основе априорной модели, свидетельствует о сходстве параметров выборочной совокупности и модели случайной величины.

**Параметры распределения** — числовые характеристики распределения, которые описывают данный вид распределения. Выделяют следующие параметры: математическое ожидание или среднее арифметическое, дисперсию, асимметрию и эксцесс. Нормальное распределение является двухпараметрическим, так как его можно описать через два параметра: среднее значение и среднее квадратичное отклонение.

**Параметрические критерии** — критерии, применяемые в предположении о нормальном распределении переменных в генеральной совокупности.

**Ранг** — место значения в упорядоченном ряду значений, величина значения относительно других значений в совокупности. **Ранжирование** — определение порядка согласно рангу.

**Регрессионный анализ** — инструмент статистики, позволяющий прогнозировать значения зависимой переменной с помощью известных значений независимых переменных. **Множественный регрессионный анализ** — метод, позволяющий спрогнозировать значения зависимой переменной на основе известных значений независимых переменных.

**Фактор** — в факторном анализе объединение нескольких переменных, чья взаимная корреляция исчерпывает определенную долю общей дисперсии. После процедуры вращения каждый фактор интерпретируется как некоторая общая причина взаимосвязи группы переменных.

**Факторный анализ** — метод, позволяющий свести большое количество исходных переменных к значительно меньшему числу факторов, каждый из которых объединяет исходные переменные, имеющие сходный смысл.

**Факторная матрица** — таблица, описывающая факторные нагрузки для анализируемых переменных относительно выделенных факторов.