

федеральное государственное автономное образовательное учреждение высшего образования
**Первый Московский государственный медицинский университет им. И.М.
Сеченова Министерства здравоохранения Российской Федерации
(Сеченовский Университет)**

Институт психологии и социальной работы
Кафедра педагогики и медицинской психологии

**Методические материалы по дисциплине:
Многомерные методы анализа данных**

основная профессиональная образовательная программа
высшего образования - программа специалитета

КОД Наименование ОП: 37.05.01 Клиническая психология

**Итоговое тестирование (ЦТ) по курсу
«Многомерные методы анализа данных»**

1. С вероятностью не менее 99,7% значения нормально распределенной случайной величины находятся в диапазоне ...:

- А) $M \pm \sigma$
- Б) $M \pm 3\sigma$ +
- В) $M \pm 3\sigma$
- Г) $M \pm 3$

2. Кросстабуляция описывается в рамках...:

- А) Дисперсионного анализа
- Б) Методов сравнения
- В) Корреляционного анализа
- Г) Таблиц сопряженности +

3. К мерам рассеяния НЕ относится

- А) размах
- Б) стандартная ошибка
- В) среднее значение +
- Г) стандартная ошибка

4. К задачам факторного анализа НЕ относится

- А) изучение структуры взаимосвязи переменных
- Б) идентификация факторов как причины взаимосвязи переменных (корреляции)
- В) вычисление значений факторов, как новых переменных, с целью сокращения размерности файла
- Г) удаление переменных на основе оценок различий между объектами +

5. Что НЕ относится к способам формирования кластеров в кластерном анализе

- А) метод дробления
- Б) метод слияния
- В) метод подчинения +
- Г) межгрупповое связывание

6. Выберите НЕВЕРНОЕ утверждение относительно процедуры факторного анализа

- А) позволяет свести больше число переменных к меньшему числу независимых влияющих величин
- Б) в один фактор объединяются абсолютно разные переменные +
- В) переменные из разных факторов слабо коррелируют между собой
- Г) в один фактор объединяются переменные с высокой корреляцией

7. Статистическое наблюдение – это:

- А) научная организация регистрации информации
- Б) оценка и регистрация признаков изучаемой совокупности
- В) работа по сбору массовых первичных данных +
- Г) обширная программа статистических исследований

8. Статистический показатель – это

- А) обширная программа статистических исследований
- Б) размер изучаемого явления в натуральных единицах измерения
- В) количественная характеристика свойств в единстве с их качественной определенностью +
- Г) результат измерения свойств изучаемого объекта

9. Какие из перечисленных выборок можно охарактеризовать как зависимые?

- А) члены двух футбольных команд
- Б) ученики параллельных классов
- В) испытуемые до психотерапии и после +
- Г) страдающие алкогольной и игровой зависимостью

10. В каком из вариантов ответа указаны шкалы, относящиеся к категориальному типу:

- А) Номинальная шкала, порядковая шкала +
- Б) Шкала отношений, интервальная шкала
- В) Номинальная шкала, интервальная шкала
- Г) Порядковая шкала, шкала отношений

11. Порядковая шкала предполагает

- А) Группировку по классам (Значения неупорядоченных признаков)
- Б) Приписывание числа в зависимости от степени выраженности свойства +
- В) Обеспечение равных расстояний между градациями переменной
- Г) Включает в себя требования всех остальных видов шкал и содержит нулевую точку, в которой выраженность признака отсутствует

12. Укажите последовательность, в которой верно указано место этапа “сбора данных”:

- А) Теоретический анализ,... , выдвижение гипотезы, обработка данных, анализ данных
- Б) ...,теоретический анализ, выдвижение гипотезы, обработка данных, анализ данных
- В) Теоретический анализ, выдвижение гипотезы, ... ,обработка данных, анализ данных +
- Г) Выдвижение гипотезы, анализ данных, теоретический анализ, ... , обработка данных

13. Каких шкал не существует?

- А) Дисперсионных +
- Б) Метрических
- В) Номинативных
- Г) Порядковых

14. Какого вида дисперсионного анализа не существует?

- А) Одномерный
- Б) Многомерный
- В) Многофакторный
- Г) Единичный +

15. Математическое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями называется

- А) законом распределения +
- Б) статистическим показателем
- В) дисперсией
- Г) коэффициентом вариации

16. Сила корреляции:

- А) Зависит от ее направленности
- Б) Находится в интервале от -0.5 до 0.5 включительно
- В) Равна нулю, если переменные независимы +
- Г) Все ответы неверны

17. Как называется гипотеза, предполагающая, что различий между распределениями в выборках нет?

- А) Альтернативная гипотеза
- Б) Нулевая гипотеза +
- В) Конкурирующая гипотеза
- Г) Исследовательская гипотеза

18. К какому способу проверки нормальности распределения относится критерий Колмогорова-Смирнова?

- А) К графическому
- Б) К статистическому +
- В) Ни к одному из перечисленных
- Г) К критериальному

19. Одним из методов иерархической кластеризации является

- А) Метод риска
- Б) Метод кросс-связи
- В) Метод полной связи +
- Г) Метод экстраполяции

20. Если гипотеза строится на том, что на параметр влияет несколько факторов, то речь идет (назовите вид связи)

- А) О каузальной связи
- Б) О стохастической связи +
- В) О корреляционной связи
- Г) О ковариационной связи

21. Когда возникают ложные корреляции?

- А) при совершении ошибки второго рода
- Б) когда неверно выбрал статистический критерий
- В) при совершении ошибки первого рода
- Г) когда на параметр действуют неучтенные факторы +

22. На чем основываются параметрические методы сравнения?

- А) на сравнении средних значений +
- Б) на сравнении средних рангов
- В) на сравнении статистических критериев
- Г) на сравнении дисперсий

23. Какой критерий показывает гомогенность дисперсий?

- А) U-критерий Манна-Уитни
- Б) T-критерий Уилкоксона
- В) F-критерий Ливеня +
- Г) Критерий серий

24. Какой статистический параметр будет задавать пик кривой распределения?

- А) медиана
- Б) стандартное отклонение

- В) мода
- Г) среднее значение +

25. Термин корреляция в статистике понимают как:

- А) связь, зависимость +
- Б) отношение, соотношение
- В) функцию, уравнение
- Г) тождественность

26. Термин корреляция в статистике понимают как:

- А) связь, зависимость +
- Б) отношение, соотношение
- В) функцию, уравнение

27. Для сравнения выборочных средних используется

- А) F - критерий Фишера
- Б) Критерий Манна-Уитни
- В) T - критерий Стьюдента +
- Г) критерий Уилкоксона

28. К мерам расположения случайной величины относится:

- А) Мода
- Б) Медиана
- В) Среднее
- Г) Все перечисленное верно +

29. Параметрические методы используют, как правило, если выборка испытуемых:

- А) насчитывает N меньше или равно 30
- Б) насчитывает N больше или равно 10
- В) насчитывает N больше или равно 30 +
- Г) насчитывает N больше или равно 50

30. Какая шкала позволяет дать ответ на вопрос «на сколько» одно значение превышает другое?

- А) Номинальная
- Б) Порядковая
- В) Интервальная +
- Г) Отношений

31. Что НЕ относится к мерам «центральной тенденции»?

- А) Мода

- Б) Медиана
- В) Стандартное отклонение +
- Г) Все перечисленное неверно

32. Буквой М в статистике обозначают

- А) Масштаб
- Б) Дисперсия
- В) Среднее арифметическое +
- Г) Мода

33. Для чего нужны одномерные методы статистического вывода?

- А) Для проверки исследовательских гипотез +
- Б) Для представления эмпирических данных в доступном для интерпретации виде
- В) Для автоматизации основных мыслительных операций
- Г) Для многократного преобразования переменных

34. В условиях нормального распределения какой процент наблюдений позволит описать среднее значение в сумме с одним стандартным отклонением?

- А) 50%
- Б) 95,2%
- В) 99,7%
- Г) 68,26% +

35. Что обозначается буквой "а" в регрессионном анализе?

- А) Асимметрия
- Б) Коэффициент регрессии
- В) Константа +
- Г) Зависимая переменная

36. Какого способа проверки гипотез не существует?

- А) Графического
- Б) Экспозиционного +
- В) Критериального
- Г) Статистического

37. Регрессионный анализ позволяет...

- А) выявить шкалы как критерии поляризации сравниваемых объектов
- Б) предсказать значения метрической зависимой переменной по известным значениям независимых переменных +
- В) выявить структуры переменных по совокупности факторов

Г) разбить множество на группы схожих объектов по измеренным характеристикам

38. Определить вклад каждой из независимых переменных в вариацию зависимой переменной позволяет...

- А) регрессионный анализ +
- Б) дисперсионный анализ
- В) факторный анализ
- Г) кластерный анализ

39. Для определения линии регрессии программным обеспечением используется...

- А) метод обратных квадратов
- Б) метод расширенных квадратов
- В) метод наибольших квадратов
- Г) метод наименьших квадратов +

40. Что представляет собой диаграмма рассеяния

- А) График для анализа связи между двумя переменными, на котором каждое наблюдение представляет собой точку +
- Б) График для анализа одной переменной
- В) Диаграмма разброса, включающая сдвиги точек от линии регрессии по вертикальной оси.
- Г) Диаграмма, позволяющая выбрать число факторов в факторном анализе на основе критерия каменистой осыпи Р. Кеттелла.

41. Что описывает Лямбда Уилкса

- А) Отношение внутригрупповой суммы квадратов к общей сумме квадратов, характеризующее дисперсию оценок дискриминантной функции, не обусловленную различиями между двумя группами +
- Б) Критерий равенства дисперсионно-ковариационных матриц, основанный на близости значений определителей матриц ковариаций двух или более групп.
- В) Критерий многомерной нормальности для распределения переменных.
- Г) Все ответы верные

42. Завершающим этапом кластерного анализа является:

- А) Интерпретация +
- Б) Оценка достоверности кластеризации
- В) Формулировка проблемы
- Г) Выбор меры расстояния

43. Дать оценку достоверности сдвига в значениях исследуемого признака можно с помощью критерия

- А) Т – критерия Вилкоксона +
- Б) χ^2 критерия Пирсона
- В) F – критерий Фишера
- Г) Коэффициент конкордации

44. Основным критерием, который используется в дисперсионном анализе выступает:

- А) коэффициент Фишера +
- Б) критерий Манна-Уитни
- В) критерий Стьюдента
- Г) критерия Пирсона

45. Статистический метод, применяемый к данным, для которых закон распределения данных не известен, можно охарактеризовать как:

- А) Непараметрический +
- Б) Параметрический
- В) Экспоненциальный
- Г) Нет правильного ответа

46. Коэффициент корреляции $r = -0,6$ свидетельствует о наличии между переменными:

- А) Средней обратной связи +
- Б) Высокой обратной связи
- В) Высокой прямой связи
- Г) Низкой обратной связи

47. Расставьте правильную последовательность факторного анализа (ФА)

- А) Вычисление корреляционной матрицы-вращение факторов-извлечение факторов-интерпретация факторов
- Б) Извлечение факторов-вычисление корреляционной значимости-вращение факторов-интерпретация факторов
- В) Вычисление корреляционной значимости-извлечение факторов-вращение факторов-интерпретация факторов +
- Г) Интерпретация факторов-вычисление корреляционной значимости-вращение факторов-извлечение факторов

48. Выборка – это

- А) выбранная с целью исследования часть генеральной совокупности +
- Б) те, кто готов принять участие в исследовании

В) часть генеральной совокупности (30 и > человек), не участвующих в исследовании

Г) те, кто не готов принять участие в исследовании

49. Если доверительный интервал равен 68%, то ошибка равна...:

А) 21%

Б) 32% +

В) 38%

Г) 0%

50. Дисперсия — это

А) разброс значений относительно медианы

Б) разброс значений относительно среднего +

В) разница между максимальным и минимальным значениями ряда

Г) отклонение от нормального распределения

51. Нулевая гипотеза с направленной альтернативной

А) $H_0: X_1 = X_2$, $H_1: X_2 > X_1$ +

Б) $H_0: X_1 = X_2$, $H_1: X_2 \geq X_1$

В) $H_0: X_1 \neq X_2$, $H_1: X_2 \geq X_1$

Г) $H_0: X_1 \neq X_2$, $H_1: X_2 > X_1$

52. F-критерий Фишера используется

А) для проверки нормальности распределения

Б) проверки значимости выборочного коэффициента корреляции

В) сравнения двух зависимых выборок по уровню выраженности признака

Г) сравнения дисперсий двух выборок +

53. Положительное значение квадратного корня из дисперсии — это

А) разброс

Б) мода

В) стандартное отклонение +

Г) эксцесс

54. Упорядоченное множество значений выборочной характеристики

$X = \{1, 2, 3, 4, 4, 10\}$. Выборочное среднее равно

А) 3,5

Б) 4 +

В) 4,5

Г) 5,5

55. Упорядоченное множество значений выборочной характеристики

$X = \{1, 2, 3, 4, 4, 10\}$. Мода равна

А) 3,5

Б) 4 +

В) 4,5

Г) 6

56. Упорядоченное множество значений выборочной характеристики

$X = \{1, 2, 3, 4, 4, 10\}$. Медиана равна

А) 3,5 +

Б) 4

В) 4,5

Г) 6

57. Верным утверждением относительно выборочного среднего является

А) Чем больше отклонение от симметричности, тем меньше расхождение между значениями медианы, моды и среднего арифметического

Б) Среднее весьма чувствительно к выбросам - экстремально малым или большим значениям переменной +

В) Сумма всех отклонений от среднего никогда не равна нулю

Г) Для метрических переменных, распределение которых унимодальное и симметричное, мода, медиана и среднее не совпадают

58. Для наглядного представления характера корреляционной связи двух А) переменных используется

А) круговая диаграмма

Б) столбиковая диаграмма

В) пузырьковая диаграмма

Г) диаграмма рассеивания +

59. Основное тождество дисперсионного анализа формулируется

А) общая дисперсия равна разности внутригрупповой и межгрупповой дисперсий

Б) внутригрупповая дисперсия равна сумме общей и межгрупповой дисперсий

В) общая дисперсия равна сумме внутригрупповой и межгрупповой дисперсий +

Г) внутригрупповая дисперсия равна общей и равна межгрупповой дисперсиям

60. Такой признак как пол описывается шкалой

А) отношений

- Б) интервалов
- В) порядка
- Г) наименований +

61. Если имеется два ряда, упорядоченных по убыванию значений, то коэффициент корреляции между ними близок к

- А) 0
- Б) 1 +
- В) -1
- Г) 0,5

62. Коэффициент корреляции Спирмена неприменим, когда признаки А) принадлежат шкале

- А) отношений
- Б) интервалов
- В) порядка
- Г) наименований +

63. Мера изменчивости пропорциональная сумме квадратов разностей значений случайной величины и среднего называется

- А) размах
- Б) дисперсия +
- В) разброс
- Г) вариативность

64. Кластерный анализ представляет собой процедуру упорядочивания объектов в на основании их попарного сравнения по предварительно измеренным признакам

- А) ранжированные континуумы
- Б) однородные классы +
- В) упорядоченные группы
- Г) формальные ранги

65. Корреляционный анализ используется для

- А) сравнения двух выборок по уровню выраженности признака
- Б) оценки связи двух количественных переменных +
- В) изучения источников вариативности признака
- Г) оценки влияния латентных факторов на вариативность признака

66. К метрическим шкалам относятся шкалы

- А) отношений и порядка
- Б) интервалов и порядка

- В) порядка, интервалов и отношений
- Г) интервалов и отношений +

67. Значение признака, которое делит ранжированное множество значений пополам, называется

- А) среднее
- Б) центральный момент
- В) медиана +
- Г) мода

68. Для оценки влияния некоторого фактора(ов) на среднее значение изучаемого признака используется

- А) регрессионный анализ
- Б) факторный анализ
- В) дисперсионный анализ +
- Г) кластерный анализ

69. Для обнаружения латентной структуры некоторого явления и описания ее при помощи небольшого (по сравнению с исходным) множества измеряемых переменных используется

- А) регрессионный анализ
- Б) факторный анализ +
- В) дисперсионный анализ
- Г) кластерный анализ

70. Ошибкой первого рода называется

- А) отвержение нулевой гипотезы, когда она верна +
- Б) не отвержение нулевой гипотезы, когда она не верна
- В) неправомерный выбор критерия проверки нулевой гипотезы
- Г) ошибка, обусловленная неправильным выбором уровня значимости

71. Ошибкой второго рода называется

- А) отвержение нулевой гипотезы, когда она верна
- Б) не отвержение нулевой гипотезы, когда она не верна +
- В) неправомерный выбор критерия проверки нулевой гипотезы
- Г) ошибка, обусловленная неправильным выбором уровня значимости

72. График плотности нормального распределения – это кривая

- А) симметричная относительно нуля
- Б) симметричная относительно единицы
- В) симметричная относительно произвольного значения случайной величины +

Г) несимметричная относительно любого значения случайной величины

73. Нулевая гипотеза отвергается, если

А) значение статистического критерия не попадает в критическую область на заданном уровне значимости

Б) значение статистического критерия попадает в критическую область на заданном уровне значимости +

В) значение статистического критерия равно заданному уровню значимости

Г) значение статистического критерия не равно заданному уровню значимости

74. Для наглядного представления кластерного анализа используется

А) гистограмма

Б) полигон

В) кумулята

Г) дендрограмма +

75. Одним из условий, определяющих возможность проведения дисперсионного анализа, является знание или доказательство того, что выборка

А) нормально распределена +

Б) бимодальна

В) репрезентативна

Г) однородна

76. Вероятность совершения ошибки второго рода

А) не связана с уровнем значимости

Б) возрастает при уменьшении уровня значимости +

В) уменьшается при уменьшении уровня значимости

Г) всегда одинакова

77. Интервал числовой оси, в котором в пределах которого с некоторой вероятностью находится параметр генеральной совокупности, называется

А) предельным

Б) вероятностным

В) доверительным +

Г) достоверным

78. Для сравнения средних значений двух нормально распределенных выборок используется

А) F-критерий Фишера

Б) U-критерий Манна-Уитни

- В) t-критерий Стьюдента +
- Г) T-критерий Вилкоксона

79. Множество значений выборочной статистики, определенное на основании выбранного уровня значимости, позволяющих принять решение об ошибочности H_0 называется

- А) областью оценивания
- Б) критической областью +
- В) областью вывода
- Г) областью альтернатив

80. Число повторений одинаковых результатов в составе вариационного ряда – это

- А) статистическая значимость
- Б) статистический вес
- В) частота значения переменной +
- Г) выброс

81. Коэффициент корреляции отражает ... связь между переменными

- А) геометрическую
- Б) причинно-следственную
- В) линейную +
- Г) гипотетическую

82. Математическая модель n независимых испытаний Бернулли – это

- А) биномиальное распределение +
- Б) непрерывное распределение
- В) геометрическое распределение
- Г) гипергеометрическое распределение

83. Уровнем значимости принято называть

- А) вероятность совершения ошибки первого рода +
- Б) вероятность совершения ошибки второго рода
- В) мощность критерия
- Г) надежность критерия

84. Коэффициент эксцесса стандартного нормального распределения равен

- А) 0 +
- Б) 1
- В) -1
- Г) π

85. Математическое ожидание стандартного нормального распределения равно

- А) 0 +
- Б) 1
- В) -1
- Г) π

86. Дисперсия стандартного нормального распределения равна

- А) 0
- Б) 1 +
- В) -1
- Г) π

87. Разность между фактическим и предсказываемым значением параметра называется

- А) стандартным отклонением
- Б) случайным отклонением
- В) ошибкой оценки +
- Г) ошибкой параметра

88. Форма записи выборки, где все значения переменной расположены в один ряд в порядке возрастания или убывания называется

- А) вариативная последовательность
- Б) монотонный ряд
- В) распределение
- Г) вариационный ряд +

89. Экспериментальные данные, представленные по градациям фактора, называются комплексом

- А) дисперсионным +
- Б) интервальным
- В) факторным
- Г) вариационным

90. Какой коэффициент корреляции применяется, если x и y измерены в шкалах порядка?

- А) Пирсона
- Б) Кендалла +
- В) рангово-бисериальный
- Г) точечно-бисериальный

91. Какой из критериев нет необходимости учитывать при выборе метода статистической обработки?

- А) размер(ы) выборки(ок)
- Б) типы шкал, в которых измерены интересующие исследователя признаки
- В) отношения между выборками
- Г) надежность процедур сбора данных +

92. В шкале ... учитывается лишь символическое значение цифр, а все арифметические операции (сложение, вычитание, умножение, деление) не имеют смысла

- А) отношений
- Б) наименований
- В) порядка +
- Г) интервалов

93. Непараметрическим критерием сравнения двух зависимых выборок является

- А) F-Фишера
- Б) T-Вилкоксона +
- В) t-Стьюдента
- Г) r-Пирсона

94. Правило трех сигм гласит, что

- А) с вероятностью не менее 99,7% значения нормально распределенной случайной величины находятся в диапазоне $M \pm 3\sigma$ +
- Б) с вероятностью не менее 99,7% значения произвольно распределенной случайной величины находятся в диапазоне $M \pm 3\sigma$
- В) с вероятностью не менее 99,7% значения нормально распределенной случайной величины находятся в диапазоне $M \pm 1,5\sigma$
- Г) с вероятностью не менее 99,7% значения произвольно распределенной случайной величины находятся в диапазоне $M \pm 1,5\sigma$

95. Использование непараметрического критерия сравнения предполагает

- А) сравнение двух или более выборок по значениям их средних
- Б) сравнение двух или более выборок по значениям их дисперсий
- В) сравнение двух или более выборок по всему ряду значений характеристик +
- Г) сравнение двух или более выборок по максимальным и минимальным значениям характеристик

96. Т-критерий Вилкоксона используется

- А) для проверки нормальности распределения

- Б) проверки значимости выборочного коэффициента корреляции
- В) сравнения двух зависимых выборок по уровню выраженности признака +
- Г) сравнения дисперсий двух выборок

97. Какую гипотезу не проверяет дисперсионный анализ:

- А) о совокупном влиянии факторов
- Б) о влиянии фактора
- В) о совокупном влиянии переменных
- Г) о совокупном влиянии наблюдений +

98. Уровень значимости в математической статике — это?

- А) Вероятность ошибочного отклонения верной нуль-гипотезы +
- Б) Вероятность ошибочного принятия неверной альтернативной гипотезы
- В) Ошибка в выборе генеральной совокупности
- Г) Вероятность возникновения ошибки при проверке соответствия теоретической модели эмпирическим результатам