

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«Национальный государственный Университет
физической культуры, спорта и здоровья имени П.Ф.Лесгафта,
Санкт-Петербург»**

Кафедра БИОМЕХАНИКИ

Методическое обеспечение дисциплины

**СТАТИСТИЧЕСКАЯ ОБРАБОТКА, АНАЛИЗ И ПРЕДСТАВЛЕНИЕ
РЕЗУЛЬТАТОВ НАУЧНЫХ ИССЛЕДОВАНИЙ**

Основная профессиональная образовательная программа
высшего образования - программы подготовки научно-педагогических кадров в
аспирантуре по направлению подготовки кадров высшей квалификации

32.06.01 – «МЕДИКО-ПРОФИЛАКТИЧЕСКОЕ ДЕЛО»

Направленность (профиль)
Гигиена
(медицинские, биологические науки)

**квалификация ВЫПУСКНИКА - Исследователь. Преподаватель-
исследователь.**

форма обучения
очная / заочная

**Рассмотрена и утверждена на заседании
кафедры
«31» августа 2015 г., протокол № 1**

**Зав. кафедрой _____ Самсонова А.В.
Автор-разработчик:
зав. кафедрой биомеханики, д.п.н., проф.
А.В. Самсонова**

Санкт-Петербург 2015

МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРЕПОДАВАТЕЛЯ:

Преподавателю при проведении занятий по дисциплине «Статистическая обработка, анализ и представление результатов необходимо иметь следующие материалы:

1. Конспект содержания лекции.
2. Конспект содержания практических занятий.
3. Презентации содержания лекций.
4. Презентации содержания практических занятий.

КОНСПЕКТЫ ЛЕКЦИЙ

ЛЕКЦИЯ 1 СТАТИСТИЧЕСКИЕ ШКАЛЫ

Шкала – элемент счетной системы, посредством которого происходит отнесение исследуемого объекта к определенной группе.

Различают четыре типа статистических шкал:

1. Номинальная (шкала наименований)
2. Порядковая (шкала порядка, ранговая шкала)
3. Интервальная (шкала интервалов)
4. Отношений

Номинальная шкала является самой простой. В этой шкале числа служат для обнаружения и различения изучаемых объектов. Например, каждый игрок футбольной команды имеет собственный номер. Но насколько игрок под номером 1 быстрее игрока под номером 3 сказать трудно, так как этот номер дан для различения игроков.

Возможности обработки переменных, относящихся к номинальной шкале, очень ограничены. Так, например, расчет среднего арифметического не имеет смысла. В номинальной шкале можно подсчитать, насколько часто встречается то или иное число, то есть выполнить частотный анализ и определить моду.

В **порядковой шкале** каждому объекту или признаку присваиваются не только числа, но вводятся отношения между объектами по типу «больше-меньше», то есть производится операция ранжирования. Например, можно построить спортсменов по росту (ранжировать). Самому низкому спортсмену присвоить число 1 (ранг 1), рядом стоящему – ранг 2 и т.д. При этом, самый высокий спортсмен будет характеризоваться самым большим числом (высоким рангом). Однако нельзя утверждать, что *фактическое различие* между

спортсменами, имеющими ранг 1 и 2 такое же, как и между спортсменами, имеющими ранг 9 и 10. Может оказаться, если произвести измерения роста, что различие между спортсменами, имеющими ранг 1 и 2, составляет 5 см, а между спортсменами, имеющими ранг 9 и 10 – 1 см. В этом состоит недостаток шкалы порядка. Однако порядковая шкала широко используется для оценки качественных показателей спортивного мастерства, при этом результаты судейской оценки, выражаются в **баллах**.

В шкале порядка допустимы многие статистические операции: частотный анализ, вычисление моды и медианы. При проверке статистических гипотез для сравнения объектов, измеренных в порядковой шкале (например, результатов спортсменов, выраженных в баллах) можно применять ранговые критерии (Манна-Уитни, Вилкоксона), а для установления силы связи между переменными – расчет рангового коэффициента корреляции (Спирмена).

Следует отметить, что номинальная и порядковая шкалы относятся к **качественным** шкалам.

Интервальная шкала и шкала отношений относятся к количественным шкалам. Существенное их отличие от качественных шкал (номинальной и порядковой) в том, что эти шкалы обладают объективной единицей измерения (градус, метр, кг). При этом объекту присваивается число, которое эквивалентно количеству имеющегося свойства. Особенность, отличающая шкалу интервалов от шкалы отношений, состоит в том, что нулевая точка в интервальной шкале выбирается произвольно. Примером интервальной шкалы является летосчисление. Первый год в используемом во всем мире летосчислении выбран произвольно, как год рождения Иисуса Христа.

В шкале отношений нулевая точка выбирается не произвольно, а указывает на полное отсутствие измеряемого свойства. В спорте в шкале отношений измеряют массу, силу, расстояние и другие переменные. Результаты, представленные в интервальной шкале и шкале отношений можно обрабатывать любыми статистическими методами.

ЛЕКЦИЯ 2

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ. КРИТЕРИИ ЗНАЧИМОСТИ, ПРЕДСТАВЛЕНИЕ ДАННЫХ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Задачу «сжатия» информации без существенной ее потери решают **числовые характеристики выборки** — набор чисел, который характеризует выборку в разных аспектах. Следует также отметить, что в большинстве научных статей, опубликованных в журналах «Ученые записки Университета П.Ф. Лесгафта», «Адаптивная физическая культура» и «Теория и практика физической культуры» полученные экспериментальные данные представлены числовыми характеристиками выборки.

В § 5.4 были введены понятия числовых характеристик законов распределения случайной величины. Эмпирические законы распределения, ко-

торые представлены в виде выборок, имеют аналогичные числовые характеристики, называемые выборочными или статистическими. Однако, необходимо понимать, что:

- выборочные характеристики зависят от выборок; разные выборки будут иметь разные значения одной и той же числовой характеристики;
- в силу сформулированного во введении закона устойчивости частот при возрастании объёма выборки выборочные числовые характеристики приближаются к “истинным” (неизвестным) соответствующим характеристикам генеральной совокупности.

Выборочные характеристики в математической статистике принято отмечать верхней звёздочкой; например, кумулята F_x^* является выборочной функцией распределения и для другой выборки будет иметь иной вид. Определим основные **выборочные числовые характеристики**, являющиеся аналогами введённых в § 5.4 соответствующих характеристик законов распределения (попутно в равенствах приведены общепринятые обозначения этих характеристик). Эти характеристики можно разделить на три группы: **положения, рассеяния (вариативности); формы.**

положения	Числовые характеристики рассеяния	формы
1. Среднее арифметическое $M^* = \bar{x}$.	1. Размах вариации R .	1. Коэффициент асимметрии A_3^* .
2. Мода Mo^* .	2. Дисперсия $D^* = S^2$.	2. Коэффициент эксцесса E^* .
3. Медиана Me^*	3. Среднеквадратическое отклонение (стандартное отклонение) $\sigma^* = S$.	
	4. Коэффициент вариации V .	

Характеристики положения

Среднее арифметическое (M^* или \bar{x}) – одна из основных характеристик выборки. Оно представляет собой такое значение признака, сумма отклонений от которого выборочных значений равна нулю. Среднее арифметическое имеет размерность признака. Например, если признак – рост человека, выражен в метрах и среднее арифметическое выражено в метрах. Среднее арифметическое вычисляется по формуле:

$$M^* = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i ;$$

Мода (Mo^*) представляет собой значение признака, встречающееся в выборке наиболее часто. Мода, как и среднее арифметическое имеет размерность изучаемого признака.

Пример 5.4. Имеем выборку: 2, 3, 3, 4, 5, 7, 7, 7, 8. Определить моду.

Решение. В данной выборке модой является число 7, так как это число в выборке встречается три раза.

Медиана (Me^*) представляет собой такое значение признака, которое находится в середине ранжированного ряда. Медиана, как среднее арифме-

тическое и мода имеет размерность изучаемого признака.

Ранжированный ряд – ряд, в котором выборочные данные расположены в порядке возрастания или убывания.

Следует отметить, что если распределение выборочных данных близко к нормальному, значения среднего арифметического, моды и медианы достаточно близки друг к другу. Если распределение данных выборки асимметрично, численные значения характеристик положения различаются.

Каждая из характеристик положения несет определенную информацию о выборке. Какую из них выбрать для анализа зависит от статистической шкалы, в которой представлены исходные данные и задач исследования. Так, например, если данные представлены в порядковой шкале, рекомендуется для характеристики положения выборки использовать медиану. Если данные представлены в интервальной шкале и шкале отношений, рекомендуется использовать среднее арифметическое (Дж. Гласс, Дж. Стенли, 1976).

Характеристики рассеяния (вариативности)

Характеристики положения не дают полной информации о выборке, так как не учитывают ее вариации. Поэтому наряду с характеристиками положения введены характеристики рассеяния (вариативности) выборки.

Размах вариации (выборки) R вычисляется как разность между максимальной x_{\max} и минимальной x_{\min} вариантами выборки:

$$R = x_{\max} - x_{\min}.$$

Размах вариации имеет размерность изучаемого признака и вычисляется достаточно просто, однако информативность этого показателя невелика.

Дисперсия – средний квадрат отклонений значений признака от среднего арифметического. Дисперсия вычисляется по формуле:

$$D^* = S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Следует отметить, что чем больше значение дисперсии, тем больше вариативность исследуемого признака. Однако размерность дисперсии не совпадает с единицами измерения варьирующего признака. Например, у признака рост, единица измерения метр. Значение дисперсии будет представлено в квадратных метрах. Поэтому на практике используют корень квадратный из дисперсии, который называется **среднеквадратическим отклонением** (стандартным отклонением). Среднеквадратическое отклонение вычисляется по формуле:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

Коэффициент вариации (V) – безразмерная характеристика, выражается в % и вычисляется по формуле: $V = \frac{S}{\bar{x}} 100\%$

Посредством этого коэффициента сравниваются выборки величин разных размерностей. Коэффициент вариации предложен К. Пирсоном.

Примечание. Коэффициент вариации используется для оценки однородности выборки. Если $V < 10\%$, то выборка считается однородной, т.е. полученной из одной генеральной совокупности.

Характеристики формы распределения

Кривая эмпирического распределения не всегда идеально соответствует нормальной кривой. Поэтому для характеристики изменения формы кривой эмпирического распределения введены два показателя: коэффициент асимметрии и коэффициент эксцесса.

Коэффициент асимметрии (A_s) – показатель, характеризующий «скошенность» эмпирического распределения вправо или влево по отношению к нормальному распределению. Для расчета коэффициента асимметрии применяется формула:

$$A_s^* = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{nS^3}.$$

Если знак коэффициента асимметрии положительный, то асимметрия левосторонняя (рис.1), если знак коэффициента асимметрии отрицательный – асимметрия правосторонняя (рис.2). Коэффициент асимметрии является безразмерным показателем.

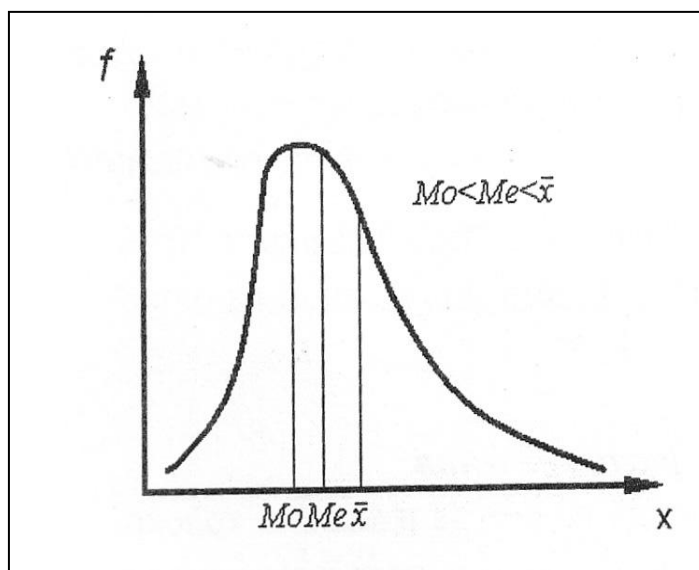


Рис.1. Левосторонняя (положительная) асимметрия (В.С. Маркарян, 2007)

Коэффициент эксцесса (E) – показатель, характеризующий «плосковершинность» или «островершинность» эмпирического распределения по отношению к нормальному распределению. Нормальное распределение характеризуется тем, что значения коэффициентов асимметрии и эксцесса равны нулю. Чем больше значения коэффициентов асимметрии и эксцесса (по

абсолютной величине), тем сильнее эмпирическое распределение отличается от нормального. Коэффициент эксцесса рассчитывается по формуле:

$$E^* = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{nS^3} - 3.$$

Как и коэффициент асимметрии, коэффициент эксцесса является безразмерным показателем. Если знак эксцесса отрицательный (-), то имеется тенденция к плосковершинности. Если знак положительный, то имеется тенденция к островершинности (рис.3).

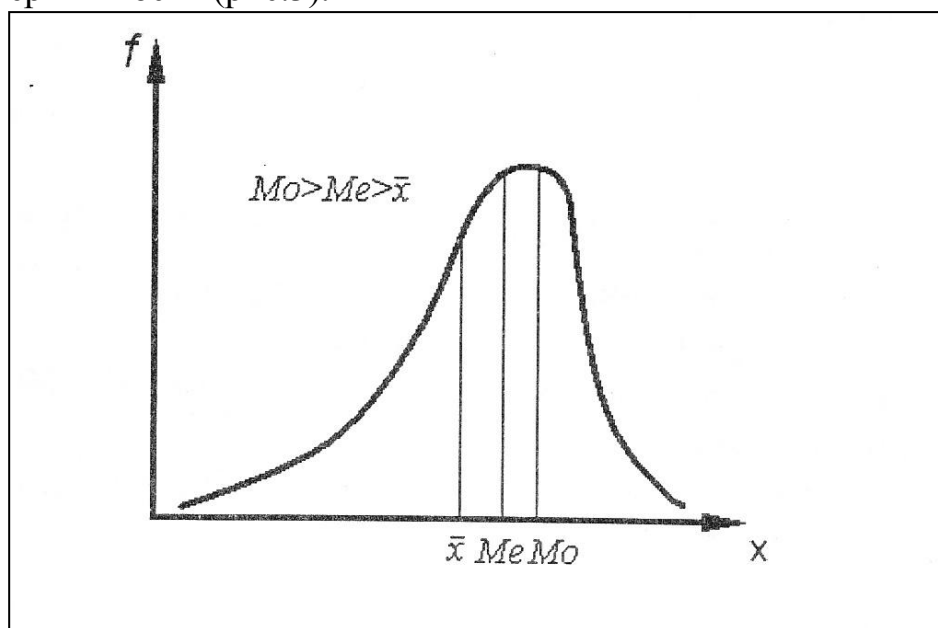


Рис.2. Правосторонняя (отрицательная) асимметрия
(В.С. Маркарян, 2007)

В математической статистике широко используются выборочные характеристики, которые не имеют аналогов в теории вероятностей. Одной из таких характеристик является **ошибка среднего арифметического** $S_{\bar{x}}$. Эта

характеристика вычисляется по формуле: $S_{\bar{x}} = \frac{S}{\sqrt{n}}$. Расчет этой характеристики связан с тем, что, как правило, среднее арифметическое и дисперсия, вычисленные на основе выборочных данных не совпадают с истинными значениями параметров μ и σ^2 генеральной совокупности. Отклонения среднего арифметического и дисперсии, вычисленных по выборочным данным от их истинных значений называются статистическими ошибками или ошибками репрезентативности. Следует отметить, что чем меньше по объему выборка, тем больше это отклонение.

Ошибка среднего арифметического $S_{\bar{x}}$ показывает, какая ошибка в среднем допускается, если использовать вместо генерального среднего μ его выборочную оценку \bar{x} . (В.С.Иванов, 1990). Поэтому вычисленное среднее

арифметическое часто указывают в виде $\bar{x} \pm S_{\bar{x}}$, чтобы указать точность оценки генерального среднего.

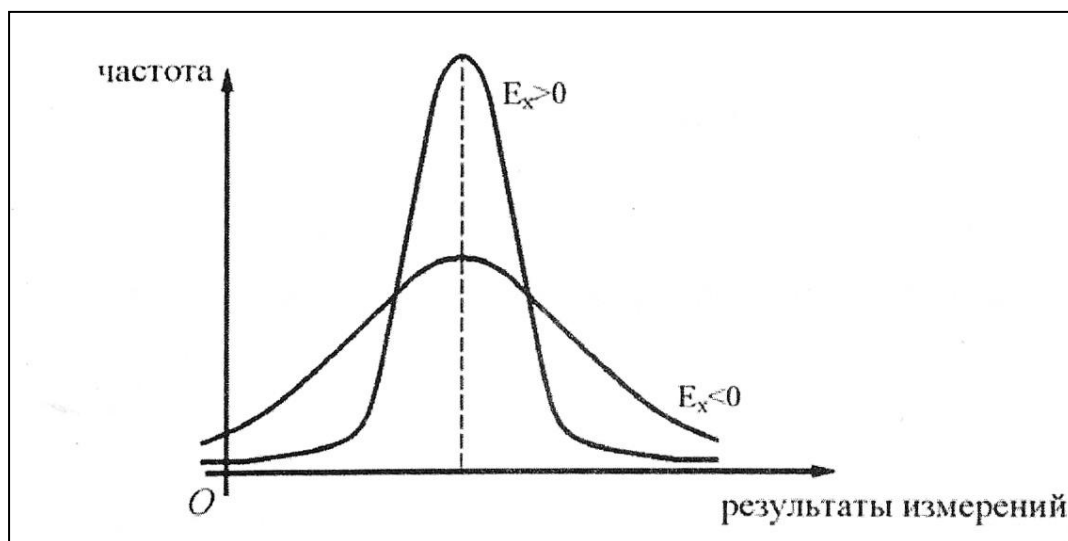


Рис.3. Островершинная и плосковершинная кривые распределения (В.С. Маркарян, 2007)

Расчет числовых характеристик выборки

Пример. 10 спортсменов, выполнили по 10 бросков баскетбольного мяча в корзину. В результате были получены следующие результаты: 2, 5, 6, 3, 4, 5, 5, 5, 2, 1. Необходимо по полученной выборке определить ее числовые характеристики или другими словами, заменить полученную выборку другими характеристиками, которые достаточно точно ее представляют.

Решение. Для расчета числовых характеристик выборки создадим расчетную таблицу (1).

Таблица 1.

Таблица для расчета числовых характеристик выборки

№	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
1	2	-1,8	3,24	- 5,832	10,4976
2	5	1,2	1,44	1,728	2,0736
3	6	2,2	4,84	10,648	23,4256
4	3	- 0,8	0,64	- 0,512	0,4096
5	4	0,2	0,04	0,008	0,0016
6	5	1,2	1,44	1,728	2,0736
7	5	1,2	1,44	1,728	2,0736
8	5	1,2	1,44	1,728	2,0736
9	2	- 1,8	3,24	- 5,832	10,4976
10	1	- 2,8	7,84	- 21,952	61,4656
	$\sum_{i=1}^n = 38$		$\sum_{i=1}^n (x_i - \bar{x})^2 = 25,6$	$\sum_{i=1}^n (x_i - \bar{x})^3 = -16,56$	$\sum_{i=1}^n (x_i - \bar{x})^4 = 114,592$

После этого переходим к вычислениям характеристик.

Среднее арифметическое $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{38}{10} = 3,8$

Медиана $Me = 4,5$

Мода $Mo = 5$

$x_{\min} = 1$

$x_{\max} = 6$

$R = x_{\max} - x_{\min} = 6 - 1 = 5$

Дисперсия $D^* = S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{10-1} 25,6 = \frac{25,6}{9} = 2,8444$

Среднеквадратическое отклонение (стандартное отклонение) –

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{2,8444} = 1,6865 ;$$

Коэффициент вариации $V = \frac{S}{\bar{x}} 100\% = \frac{1,6865}{3,8} 100 = 44,38\%$

Ошибка среднего арифметического $S_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{1,6865}{\sqrt{10}} = 0,533333$

Коэффициент асимметрии $A_s^* = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{nS^3} = \frac{-16,56}{10 \cdot 1,6865^3} = \frac{-16,56}{10 \cdot 4,797292} = -0,34519$

Коэффициент эксцесса –

$$E^* = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{nS^4} - 3 = \frac{114,592}{10 \cdot 1,6865^4} - 3 = \frac{114,595}{10 \cdot 8,090864} - 3 = -1,588369 .$$

Следует отметить, что применение персональных компьютеров позволяет проводить вычисления числовых характеристик с большим количеством знаков после запятой. Однако существуют правила записи результатов. Эти правила основываются на точности вычисления выборочных характеристик. Если объем выборки ограничен, истинное значение характеристик генеральной совокупности не может быть вычислено сколь угодно точно. Поэтому при вычислении выборочных характеристик нет смысла оставлять большое число значащих цифр.

Под **значащими цифрами** понимают все цифры, кроме нуля, а также и нуль, в том случае, если он стоит между значащими цифрами. Все нули, расположенные слева, не являются значащими. Нули, расположенные справа, могут быть как значащими, так и незначащими.

Правила записи результатов исследования:

- Результат ошибки среднего указывается двумя значащими цифрами, если первая из них 1 или 2, поскольку отбрасывание второго знака приводит к большой ошибке (до 30-50%). Если ошибка среднего начинается с цифры, равной или большей, чем 3, то в результате записи сохраняется одна значащая цифра.
- Результат измерения (среднее арифметическое) округляется до того же десятичного знака, которым оканчивается округленное значение ошибки среднего.
- Округление производится лишь в окончательном ответе.¹

Пример определить значащие цифры в следующих числах: 2,396029; 3200; 0,00267.

Решение. В числе 2,396029 все цифры значащие, в том числе и ноль, так как он стоит между значащими цифрами.

В числе 3200 все цифры значащие, в том числе и два нуля, так как они стоят справа от значащих цифр.

В числе 0,00267 — значащие только цифры 2, 6, 7. Так три нуля стоят слева от значащих цифр, они не являются значащими цифрами.

Пример. Записать результаты исследования, представленные в табл. 2 в соответствии с вышеизложенными правилами.

Решение

Таблица 2

\bar{x}	S_x	$\bar{x} \pm S_x$	Примечание
0,568321	0,014127	0,584±0,014	В записи ошибки среднего оставляем две значащие цифры,

¹ Л.В.Новицкий, И.А.Зограф Оценка погрешностей результатов измерений, Л.:Энергоатомиздат, 1985.- 248 с.

			так как первая цифра единица
48,657103	0,483127	48,7±0,5	
37,942271	0,337812	37,9±0,3	
167,37212	5,384623	167±5	

Точечное и интервальное оценивание числовых характеристик

Точечная оценка числовой характеристики представляет собой число. К точечным оценкам относятся среднее арифметическое или стандартное отклонение.

Если выборка небольшого объёма, то точечная оценка может значительно отличаться от оцениваемого параметра и её использование может привести к грубым ошибкам. Поэтому при небольшом объёме выборки в математической статистике используют другого типа оценки характеристик генеральной совокупности – интервальные.

Интервальной оценкой числовой характеристики называется интервал, который с доверительной вероятностью P (задаваемой заранее) накрывает истинное значение числовой характеристики генеральной совокупности.

Как правило, в научных исследованиях в области спорта считается достаточной доверительная вероятность $P=0,95$. В некоторых случаях, связанных с большой ответственностью при принятии решений, принимают P равной 0,99 или 0,999. Таким образом, **доверительная вероятность** – это уровень гарантии суждения о значениях генеральной характеристики на основании выборочных данных.

Вероятность $\alpha = 1 - P$ того, что построенный доверительный интервал не накроет значение генеральной характеристики, называется **уровнем значимости**; другими словами, α - вероятность ошибки.

В литературе часто обе вероятности α и P выражают в процентах, т.е. $100\alpha\%$ и $100P\%$. Приведём примеры интервальных оценок.

Доверительный интервал для математического ожидания μ генеральной совокупности в случае её распределения по нормальному закону и неизвестным среднеквадратическим отклонением записывается в виде:

$$\bar{x} - t_p \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + t_p \frac{S}{\sqrt{n}},$$

где t_p – так называемый квантиль распределения, который зависит от объёма выборки n и доверительной вероятности P (таблица 5.11), а $\frac{S}{\sqrt{n}}$ – ошибка среднего арифметического $S_{\bar{x}}$.

Таблица 3

Значения t_p в зависимости от объёма выборки и доверительной вероятности P .

n	P		
	0,95	0,99	0,999
10	2,265	3,250	4,781
15	2,145	2,977	4,140
20	2,093	2,861	3,883
30	2,042	2,750	3,646
40	2,021	2,704	3,551
50	2,009	2,678	3,505
60	2,000	2,660	3,505
80	1,990	2,639	3,416

В таблице 3 представлена зависимость параметра t_p от объёма выборки и доверительной вероятности.

Пример. *Определить границы 95% доверительного интервала для среднего результата в беге на 100 м ($n = 50$), если:*

$$\bar{x} = 15,38 \text{ с} \qquad S_x = 0,13 \text{ с.}$$

Решение. Из таблицы 3 для $n = 50$ и $P = 0,95$ находим значение

$t_p = 2,009$. Следовательно, доверительный интервал будет следующим:

$$15,28 - 2,009 \cdot 0,13 \leq \mu \leq 15,38 + 2,009 \cdot 0,13 \text{ или} \\ 15,12 \leq \mu \leq 15,64.$$

После округления получим итоговый результат:

$$15,1 \leq \mu \leq 15,6 \text{ с.}$$

Вывод: с доверительной вероятностью $P = 0,95$ можно утверждать, что генеральное среднее μ заключено в границах от 15,1 до 15,6 с.

Интерпретация полученных данных и их представление в научных публикациях

Очень важно, не только рассчитать числовые характеристики, но и правильно их представить в научной публикации, дипломной работе или магистерской диссертации. По мнению В.М. Зациорского (1989) минимальная статистическая обработка материала для научной публикации заключается в том, что исследователь должен указать:

- 1) числовую характеристику положения; например, среднее арифметическое;
- 2) числовую характеристику рассеивания; например, среднеквадрати-

ческое отклонение;

3) объем выборки (число наблюдений, испытуемых и т.д.).

Без указания этих числовых характеристик публикация научной ценности не имеет.

В области физической культуры и спорта в научных публикациях общепринято представление числовых характеристик выборки в виде: $\bar{x} \pm S_{\bar{x}}$ (или $M \pm m$). В качестве примера приведем фрагмент таблицы из публикации Л.И.Александровой с соавт. (2011), (табл. 4).

Таблица 4

Сравнительный анализ морфофункциональных параметров и физического развития студенток и спортсменок, $M \pm m$ (Л.И. Александрова с соавт. 2011)

Изучаемые показатели	Студентки СФУ, Красноярск n=85	Студентки КГТЭИ, Красноярск, n=65	Студентки СГУ, Сургут, n=40	Студентки циклических видов спорта, n=80
Длина тела, см	166±5	165±7	164±5	168±5
Масса тела, кг	57±8	56±6	54±7	62±8
Обхват груди, см	89±7	86±5	84±5	13±5

Если в публикациях не указывается ошибка среднего арифметического, то вместо него необходимо привести стандартное отклонение (табл.5).

Таблица 5

Возраст достижения лучшего результата и количество необходимого для этого времени от возраста начала спортивной специализации у конькобежцев, дистанция 500 м, 225 спортсменов (по Л.Н. Жданову, 1996).

Возраст начала спортивной специализации (лет)	Спортивная квалификация	Мальчики, юноши		
		Возраст лучшего результата		Количество лет с начала специализации
		\bar{x}	S	
10	МС	20,0	0,5	10,0
	КМС	17,6	0,5	7,6
	I,II	15,0	0,3	5,0

Когда необходимо сравнить вариативность признаков, измеренных в различных единицах, то кроме вышеупомянутых числовых характеристик можно привести коэффициенты вариации.

Пример Рассчитать числовые характеристики: среднее арифметическое и ошибку среднего арифметического для результатов контрольной и экспериментальной групп гимнасток в прыжке (таблица 5.14) и представить их в виде, удобном для научной публикации.

Таблица 6.

Длительность фазы полета в прыжке вверх с места у женщин-гимнасток, мс

	до эксперимента		после эксперимента	
	контрольная группа	эксперимент группа	контрольная группа	эксперимент группа
	595	585	580	630
	600	570	620	590
	605	605	610	630
	580	615	625	615
	590	605	610	650
	610	600	610	625
	560	555	550	590
	600	560	590	600
	600	595	605	620
0.	575	595	575	630
1.	570	615	575	650
2.	565	590	570	615

Решение. Расчет числовых характеристик мы предлагаем сделать самостоятельно.

Таблица 7

Предварительные результаты расчета числовых характеристик

Показатель		Контрольная группа n=12	Экспериментальная группа n=12	dk	de
Среднее арифметическое, мс	До эксперимента	587,5	590,833	5,833	29,583
	После эксперимента	593,333	620,417		
Дисперсия, мс ²	До эксперимента	284,091	399,242	16,353	12,515
	После эксперимента	542,424	397,538		
Ошибка среднего арифметического,	До эксперимента	4,86562	5,76803	4,720	3,612
	После эксперимента	6,72324	5,75571		

мс					
----	--	--	--	--	--

Примечание: dk– разности результатов, показанных контрольной группой до и после эксперимента; de- разности результатов, показанных экспериментальной группой до и после эксперимента

Представить полученные данные рекомендуем следующим образом (в данном примере незначащие цифры не округлены, а отброшены, что также допускается) (таблица 6). В заголовке таблицы указываем, какие числовые характеристики приводим ($\bar{x} \pm S_{\bar{x}}$). В скобках указываем объем выборки ($n=12$). Результаты записываем по правилам записи результатов.

Таблица 8.

Кинематические характеристики прыжка вверх с места у женщин-гимнасток до и после эксперимента, $\bar{x} \pm S_{\bar{x}}$

Показатель	Время тестирования	Контрольная группа (n=12)	Экспериментальная группа (n=12)
Длительность фазы полета в прыжке с места, мс	До эксперимента	587±5	590±6
	После эксперимента	593±7	620±6

ЛЕКЦИЯ 3 КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗ

Корреляционным анализом называется раздел математической статистики, исследующий зависимость между случайными величинами на основе различных выборочных оценок генеральных коэффициентов корреляции.

Одной из главных задач изучения случайных массовых явлений является исследование взаимозависимостей случайных событий. В математической статистике эта задача сводится к выявлению и изучению зависимостей случайных величин друг от друга. Существуют методы, позволяющие это делать. В простейшем случае рассматриваются различные статистические оценки меры указанной зависимости.

В этом параграфе будет рассмотрена зависимость случайной величины Y от случайной величины X . Две случайные величины могут быть связаны либо функциональной зависимостью (жёстко, как зависимость переменных в математическом анализе), либо так называемой статистической зависимостью, либо быть независимыми.

Статистической называют зависимость, при которой изменение одной величины влечёт изменение распределения другой. В частности, если при изменении одной из величин изменяется среднее значение другой, то говорят о корреляционной зависимости.

Итак, рассмотрим систему случайных величин $(X; Y)$. Выборочные зна-

чения этой системы (пары чисел $(x_i; y_i)$) изображаются точками на плоскости с декартовой системой координат – **корреляционном поле**.

Чем “лучше” на корреляционном поле выборочные значения укладываются в виде какой-либо кривой, тем сильнее зависимость случайных величин в системе (X;Y).

Коэффициенты корреляции и их свойства

В качестве меры зависимости случайных величин используется коэффициент корреляции. В математической статистике наиболее распространены две оценки коэффициента корреляции:

- Пирсона r и
- Спирмена r_s .

В математической статистике оценка коэффициента корреляции Пирсона вычисляется по выборочным значениям следующим образом:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) \left(\sum_{i=1}^n (y_i - \bar{y})^2\right)}}$$

Коэффициент корреляции Пирсона применяется в случае, когда предполагается, что обе переменные x и y соответствуют закону нормального распределения. Если же хотя бы одна из переменных нормальному закону не подчиняется, или же предположение о законе распределения не делается, применяется коэффициент корреляции Спирмена.

Коэффициент корреляции Спирмена r_s вычисляется по аналогичной формуле (вместо выборочных значений $(x_i; y_i)$ в неё подставляются их ранги $(v_i; v_i)$). Поэтому оценку r_s часто называют ранговым коэффициентом корреляции. Введённые оценки коэффициента корреляции имеют одинаковые свойства. Сформулируем основные:

- 1) обе оценки r и r_s принимают значения из сегмента $[-1;1]$;
- 2) $r = +1$ или $r = -1$ тогда и только тогда, когда случайные величины X и Y связаны жёсткой линейной зависимостью, т.е. $Y = aX + b$ (на корреляционном поле изображающие точки ложатся на какую-либо прямую);
- 3) в случае нормального распределения системы (X;Y) равенство $r = 0$ равносильно независимости случайных величин X и Y ;
- 4) если же система (X;Y) распределена не по нормальному закону, равенство $r = 0$ может выполняться и при наличии нелинейной зависимости случайных величин X и Y ;
- 5) если $r \neq 0$, то чем ближе $|r|$ к единице, тем сильнее линейная зависимость случайных величин X и Y .

На основе описанных свойств коэффициента корреляции введём следующую терминологию, связанную с величиной r :

$ r = 1$	линейная зависимость
$0,7 \leq r < 0,99$	сильная статистическая зависимость
$0,5 \leq r < 0,7$	средняя статистическая зависимость
$0,2 \leq r < 0,5$	слабая статистическая зависимость
$0,09 \leq r < 0,2$	очень слабая статистическая зависимость
$r > 0$	положительная корреляция между X и Y
$r < 0$	отрицательная корреляция между X и Y
$r = 0$	корреляции нет; случайные величины X и Y некоррелированы.

На рис. 4. (корреляционном поле) представлена графическая интерпретация коэффициентов корреляции, имеющих различные значения.

Расчет коэффициента корреляции Пирсона

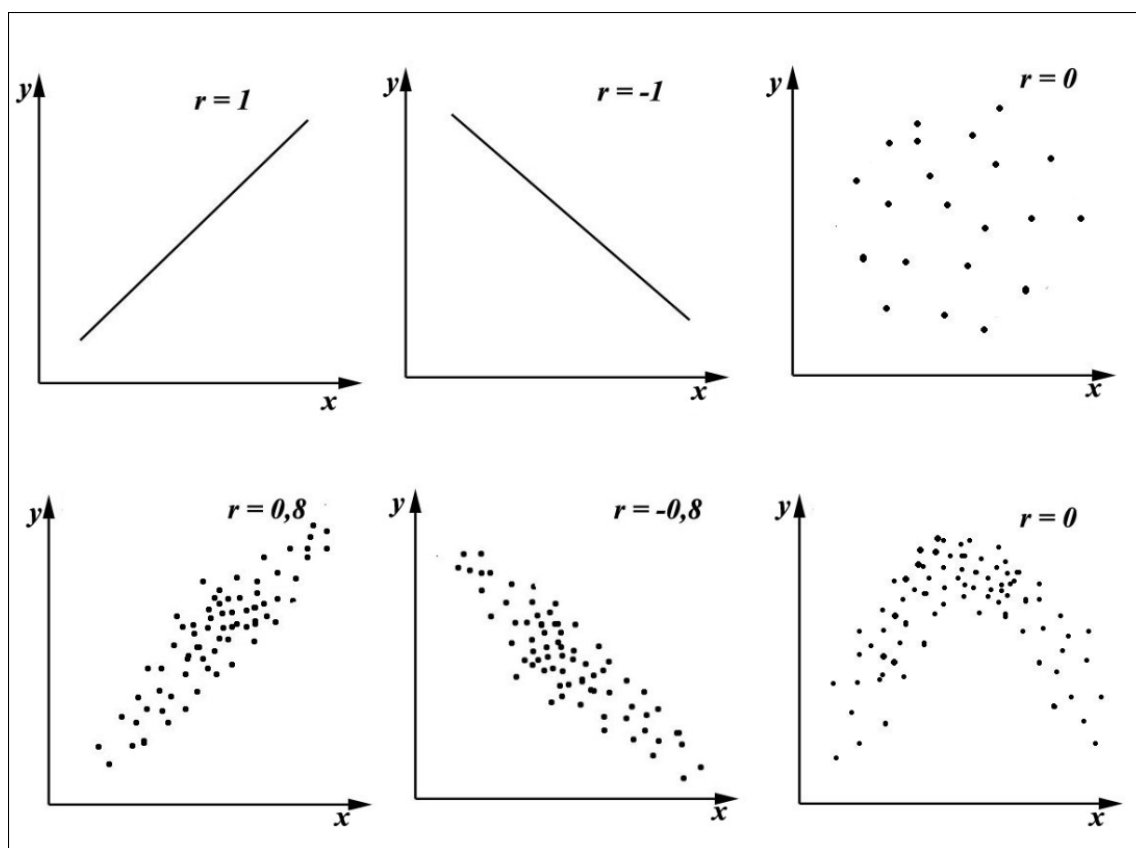


Рис.4. Виды графиков корреляционного поля

Пример Рассчитать коэффициент корреляции Пирсона для данных,

характеризующих обхват талии (X) и вес (Y) молодых женщин 20-25 лет (таблица 9) и оценить его достоверность. Сделать статистический вывод о силе зависимости между величинами X и Y .

Таблица 9

$x_i, \text{ см}$	70	74	68,5	72	76,4	83	69	75	72,5	70,5	78,5	65
$y_i, \text{ кг}$	59,1	62,3	58,6	60,2	63,4	78,6	55,4	64,9	65	63,2	72,9	56,1

Решение

1. Рассчитаем среднее арифметическое для первой и второй выборок:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{874,4}{12} = 72,9 \text{ см} \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{759,7}{12} = 63,3 \text{ кг}$$

2. Для расчета коэффициента корреляции Пирсона составляем расчетную таблицу (табл. 10).

Таблица 10

Вспомогательная таблица для расчета коэффициента корреляции Пирсона

№	x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	2	3	4	5	6	7	8
1	70	59,1	-2,9	-4,2	12,18	8,41	17,64
2	74	62,3	1,1	-1	-1,1	1,21	1
3	68,5	58,6	-4,4	-4,7	20,68	19,36	22,09
4	72	60,2	-0,9	-3,1	2,79	0,81	9,61
5	76,4	63,4	3,5	0,1	0,35	12,25	0,01
6	83	78,6	10,1	15,3	154,53	102,01	234,09
7	69	55,4	-3,9	-7,9	30,81	15,21	62,41
8	75	64,9	2,1	1,6	3,36	4,41	2,56
9	72,5	65	-0,4	1,7	-0,68	0,16	2,89
10	70,5	63,2	-2,4	-0,1	0,24	5,76	0,01
11	78,5	72,9	5,6	9,6	53,76	31,36	92,16
12	65	56,1	-7,9	-7,2	56,88	62,41	51,84
				Сумма	333,8	263,36	496,31

3. Рассчитываем коэффициент корреляции Пирсона по формуле

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) \left(\sum_{i=1}^n (y_i - \bar{y})^2\right)}} = \frac{333,8}{\sqrt{263,36 \cdot 496,31}} = 0,923$$

4. Оцениваем достоверность (значимость) коэффициента корреляции. С этой целью необходимо сравнить полученный коэффициент корреляции Пирсона равный $r = 0,923$ с критическим значением коэффициента корреляции, вели-

чина которого зависит от объёма выборки и уровня значимости. Если фактическое значение коэффициента корреляции больше, чем критическое, это означает, что коэффициент корреляции статистически достоверен. Определяем критическое значение коэффициента корреляции по таблице 5.32. В соответствии с таблицей критическое значение коэффициента корреляции при $\alpha = 0,001$ и $n=12$ составляет $r_{0,001} = 0,823$. Из этого следует, что рассчитанный коэффициент корреляции статистически достоверен.

5. Рассчитанный нами коэффициент корреляции Пирсона, равный $r = 0,923$ находится в интервале $0,7 \leq |r| < 0,99$. Следовательно, можно сделать вывод, что между обхватом талии и весом у женщин 20-25 лет существует сильная корреляционная зависимость.

Таблица 11

Критические значения выборочного коэффициента корреляции Пирсона
(В.С. Иванов, 1990).

Уровни значимости,							
n	0,05	0,01	0,001	n	0,05	0,01	0,001
3	0,9969	0,999877	0,99999877	26	0,388	0,496	0,607
4	0,950	0,9900	0,9990	27	0,381	0,487	0,597
5	0,878	0,9597	0,99114	28	0,374	0,479	0,588
6	0,811	0,9172	0,9741	29	0,367	0,470	0,579
7	0,754	0,875	0,9509	30	0,361	0,463	0,570
8	0,707	0,834	0,9244	32	0,349	0,449	0,554
9	0,666	0,798	0,898	35	0,332	0,435	0,539
10	0,632	0,765	0,872	37	0,325	0,418	0,519
11	0,602	0,735	0,847	40	0,312	0,402	0,501
12	0,576	0,708	0,823	42	0,304	0,393	0,490
13	0,553	0,684	0,801	45	0,292	0,384	0,416
14	0,532	0,661	0,780	47	0,288	0,372	0,465
15	0,544	0,641	0,760	50	0,279	0,361	0,451
16	0,497	0,623	0,742	52	0,273	0,354	0,443
17	0,482	0,606	0,725	60	0,254	0,330	0,414
18	0,468	0,590	0,708	80	0,220	0,286	0,380
19	0,456	0,575	0,693	100	0,196	0,258	0,324
20	0,444	0,561	0,679	125	0,175	0,230	0,286
21	0,433	0,549	0,665	150	0,160	0,210	0,249
22	0,423	0,537	0,652	250	0,124	0,163	0,207
23	0,413	0,526	0,641	500	0,088	0,115	0,147
24	0,404	0,515	0,629	1000	0,062	0,081	0,104
25	0,396	0,505	0,618				

Расчет коэффициента корреляции Спирмена

Пример. Рассчитать коэффициент Спирмена для данных, представленных в табл.9 и оценить его достоверность. Сделать статистический вывод о силе зависимости между величинами X и Y.

Решение. Для расчета коэффициента корреляции Спирмена (r_s) используем следующую формулу:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

где d_i – разность рангов. С этой целью составляем таблицу 5.33.

Таблица 12

Вспомогательная таблица расчета коэффициента корреляции Спирмена (r_s)

	X	Y	Ранги, X	Ранги, Y	d	d ²
1	2	3	4	5	6	7
1	65	56,1	1	2	-1	1
2	68,5	58,6	2	3	-1	1
3	69	55,4	3	1	2	4
4	70	59,1	4	4	0	0
5	70,5	63,2	5	7	-2	4
6	72	60,2	6	5	1	1
7	72,5	65	7	10	-3	9
8	74	62,3	8	6	2	4
9	75	64,9	9	9	0	0
10	76,4	63,4	10	8	2	4
11	78,5	72,9	11	11	0	0
12	83	78,6	12	12	0	0
				Сумма	0	28
	вес среднее	63,3				
	талия, среднее	72,87				

1. Находим ранги значений (столбцы 4 и 5, таблица 12)
2. Вычисляем разности рангов d (столбец 6)
3. Возводим разности рангов в квадрат и вычисляем сумму (столбец 7).
4. Затем по формуле определяем коэффициент корреляции Спирмена

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot 28}{12 \cdot (12^2 - 1)} = 1 - \frac{168}{12 \cdot 143} = 1 - \frac{168}{1716} = 1 - 0,0979 = 0,9021$$

5. Оцениваем достоверность коэффициента корреляции. Для этого определяем критическое значение коэффициента корреляции по таблице 5.34. Для объема выборки равной 12 критическое значение коэффициента корреляции при α

= 0,001 составляет: $r_{0,001} = 0,8182$. Фактическое значение коэффициента корреляции больше критического. Из этого следует, что рассчитанный коэффициент корреляции статистически достоверен.

6. Рассчитанный нами коэффициент корреляции Спирмена, равный $r = 0,902$ находится в интервале $0,7 \leq |r| < 0,99$, следовательно, можно сделать вывод, что между обхватом талии и весом у женщин 20-25 лет существует сильная корреляционная зависимость.

Таблица 14

Критические значения коэффициента ранговой корреляции Спирмена
(n- объем выборки) (В.С.Иванов, 1990)

n	Уровни значимости			n	Уровни значимости		
	0,05	0,01	0,001		0,05	0,01	0,001
5	0,8000	0,9000	-	18	0,3994	0,5480	0,6904
6	0,7714	0,8857	-	19	0,3895	0,5333	0,6737
7	0,6786	0,8571	0,9643	20	0,3789	0,5203	0,6586
8	0,5952	0,8095	0,9286	21	0,3688	0,5078	0,6455
9	0,5833	0,7667	0,9000	22	0,3597	0,4963	0,6318
10	0,5515	0,7333	0,8667	23	0,3578	0,4852	0,6186
11	0,5273	0,7000	0,8455	24	0,3435	0,4748	0,6070
12	0,4965	0,6713	0,8182	25	0,3362	0,4654	0,5962
13	0,4780	0,6429	0,7912	26	0,3299	0,4564	0,5856
14	0,4593	0,6220	0,7670	27	0,3236	0,4481	0,5757
15	0,4429	0,6000	0,7464	28	0,3175	0,4401	0,5660
16	0,4265	0,5824	0,7265	29	0,3113	0,4320	0,5567
17	0,4118	0,5637	0,7083	30	0,3059	0,4251	0,5479

Интерпретация и представление результатов корреляционного анализа в научных публикациях

В научной публикации очень важно не только описать результаты корреляционного анализа, но и правильно представить их в наглядной форме, для того, чтобы читатель мог получить максимум информации.

Часто корреляционная зависимость между результатами с указанием достоверности коэффициентов корреляции приводится в самом тексте, что не всегда удобно для восприятия.

Результаты корреляционного анализа нагляднее всего представлять в виде корреляционной матрицы, в которой представлены коэффициенты корреляции и указана их достоверность. Чаще всего значения коэффициентов корреляции приводятся с точностью до трех знаков после запятой.

Анализ публикаций свидетельствует о том, что форма представления результатов корреляционного анализа разная. Иногда в корреляционной матрице приводятся только значимые (достоверные) коэффициенты корреляции, а вместо недостоверных ставятся прочерки.

В другой форме представления результатов (табл.15) в заголовке таблицы указывается объем выборки и критические значения коэффициента корреляции при разных уровнях значимости. По нашему мнению это самая удобная форма представления результатов.

В зарубежных научных публикациях результаты часто представлены в следующей форме (табл.16). В скобках в названии таблицы указан объем выборки, $n = 13$. В таблице достоверность коэффициента корреляции обозначается различным количеством звездочек, которые помещаются в правом верхнем углу ячейки, в которой расположен коэффициент. Внизу таблицы всегда помещается примечание. Достоверность представленных в табл. 16 данных оценивалась по таблице 14. Следует отметить, что в современных статистических пакетах, таких как STATGRAPHICS и SPSS помимо расчета значений коэффициентов корреляции приводится их значимость. В связи с этим отпадает необходимость применения дополнительных источников информации в виде таблиц критических значений.

Таблица 15.

Корреляционная матрица, характеризующая зависимость между технической и функциональной подготовленностью мальчиков-пловцов 13-14 лет
($n = 13$, $r_{0,05} = 0,553$; $r_{0,01} = 0,684$; $r_{0,001} = 0,801$)

	Темп (гребк/мин)	Время прохождения дистанции кролем на груди, с			
		25 м	50 м	100 м	200 м
1	1	0,135	0,167	0,371	0,105
2		1	0,911	0,679	0,739
3			1	0,861	0,969
4				1	0,865
5					1

Таблица 16

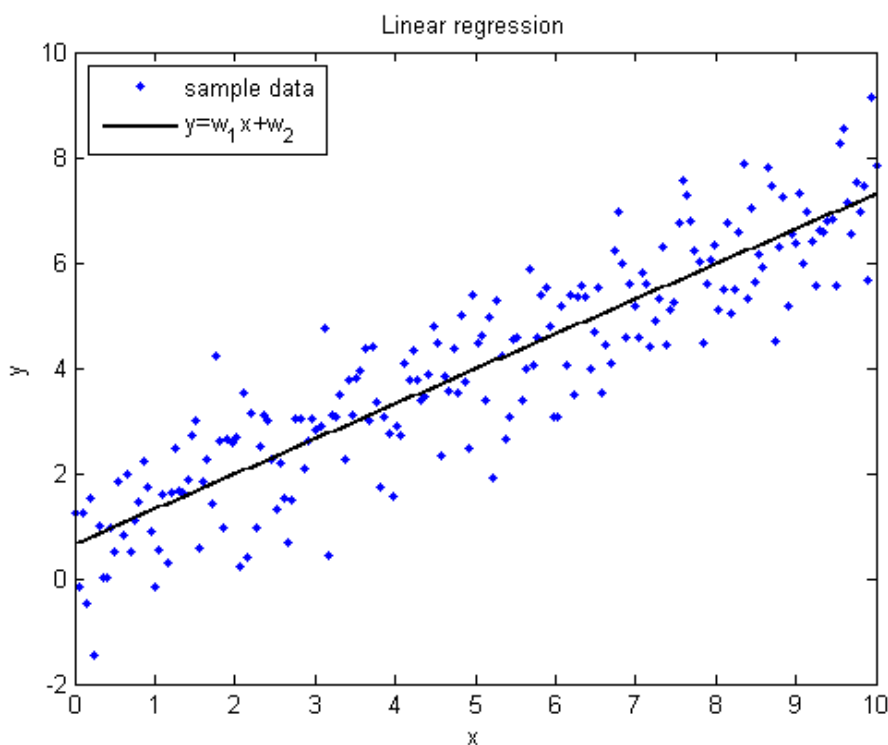
Корреляционная матрица, характеризующая зависимость между технической и функциональной подготовленностью мальчиков-пловцов 13-14 лет
($n = 13$)

	Темп (гребк/мин)	Время прохождения дистанции кролем на груди, с			
		25 м	50 м	100 м	200 м
1	1	0,135	0,167	0,371	0,105
2		1	0,911***	0,679*	0,739**
3			1	0,861***	0,969***
4				1	0,865***
5					1

Примечание: * – коэффициент корреляции достоверен при $p \leq ,05$;
 ** – коэффициент корреляции достоверен при $p \leq 0,01$;
 *** – коэффициент корреляции достоверен при $p \leq 0,001$.

Регрессионный анализ – статистический метод исследования влияния одной или нескольких независимых переменных X_1, X_2, \dots, X_p на зависимую переменную Y . Независимые переменные иначе называют предикторами, а зависимые переменные — критериальными.

Терминология *зависимых* и *независимых* переменных отражает лишь математическую зависимость переменных, а не причинно-следственные отношения.



Регрессионный анализ устанавливает формы зависимости между случайной величиной Y и значениями одной или нескольких переменных величин, причем значения последних считаются точно заданными. Такая зависимость обычно определяется некоторой математической моделью (уравнением регрессии, содержащей несколько неизвестных параметров) (рис. 1).

В спорте регрессионный анализ используется для прогноза рекордов, прогноза результатов отдельных спортсменов, в том случае, когда полное тестирование затруднено (например, в марафонском беге) и для оценки значений признака, который напрямую измерить достаточно трудно.

ЛЕКЦИЯ 4 МНОГОМЕРНЫЙ АНАЛИЗ

Многомерный анализ данных – раздел математической статистики, в котором исследуются эксперименты с множественными наблюдениями.

При проведении педагогических исследований в области физической

культуры и спорта довольно часто измеряется большое количество переменных (признаков). Для обработки таких данных, применяются многомерные методы: факторный, кластерный, дискриминантный, а также канонический корреляционный анализ. Эти методы позволяют классифицировать, визуализировать или определять взаимосвязи и закономерности в большом объеме полученной информации.

Факторный анализ (ФА) – это процедура, с помощью которой большое число переменных, сводится к меньшему количеству независимых влияющих величин, называемых факторами. При этом в один фактор объединяются переменные (признаки), сильно коррелирующие (связанные) между собой. Переменные (признаки), относящиеся к разным факторам слабо коррелируют между собой.

Факторный анализ используется при исследовании сложных объектов и систем (например, в психологии, биологии, социологии) в том случае, когда напрямую невозможно измерять величины, определяющие свойства этих объектов (так называемые факторы). Однако для измерения доступны другие величины (переменные, признаки), которые зависят от этих факторов. Иными словами, фактор, который мы не можем измерить напрямую, проявляется в изменении нескольких переменных. В области физической культуры и спорта в роли переменных могут выступать результаты тестирования уровня технической, физической, тактической подготовленности спортсменов, а также результаты медико-биологических и педагогических исследований.

Кластерный анализ предназначен для разбиения множества объектов на заданное или неизвестное число классов на основании некоторого критерия качества классификации.

Критерий качества классификации отражает следующие неформальные требования:

- Внутри групп объекты должны быть тесно связаны между собой;
- Объекты разных групп должны быть далеки друг от друга.
- При прочих равных условиях распределение объектов по группам должно быть равномерным.

Дискриминантным анализом называют раздел многомерного статистического анализа, содержащий статистические методы классификации многомерных наблюдений по одной или нескольким категориям или совокупностям.

Дискриминантный анализ является статистическим методом, который позволяет изучать различия между двумя или более группами объектов по нескольким переменным одновременно. Этот раздел статистики обширен и разнообразен. Этот метод часто бывает полезен в социальных науках.

Дискриминантный анализ – общий термин, относящийся к нескольким тесно связанным статистическим процедурам. Имеются определенные вариации применения классического дискриминантного анализа в различных пакетах данных. В пакете STATGRAPHICS реализован классический вариант

дискриминантного анализа, основанный на определении канонических направлений в исходном пространстве признаков.

Канонический корреляционный анализ — метод нахождения канонической корреляции, основанный на построении таких линейных комбинаций признаков (в двух заданных группах признаков), что обычный коэффициент парной корреляции между этими комбинациями достигает наибольшего значения. Каноническая корреляция - обобщение парной корреляции, которое используется для определения взаимосвязи между двумя группами признаков.

В отличие от обычного корреляционного анализа, в котором количественно выражается взаимосвязь между двумя различными случайными величинами, канонический корреляционный анализ позволяет находить максимальные связи между двумя различными группами величин с совместным распределением. В каждой группе отыскиваются линейные комбинации исходных признаков, имеющие максимальные корреляции. Сначала ищется первая пара таких линейных комбинаций с самой сильной зависимостью, затем вторая пара, некоррелированная с первой и т.д.

КОНСПЕКТ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

Практическое занятие №1

Расчет числовых характеристики выборки. Представление экспериментальных данных

1. Сообщить занимающимся сведения о числовых характеристиках выборки.
2. Показать с использованием статистического пакета STATGRAPHICS CENTURION возможности выполнения расчетов с одной выборкой, несколькими выборками и с выборками имеющими различный объем.
3. Для расчетов использовать сборник примеров.
4. Обратить внимание на представление статистических данных по правилам ГОСТ 7.05-2008 (оформление диссертации и автореферата).

Практическое занятие №2

Эмпирические распределения. Построение гистограммы, полигона и кумуляты.

1. Сообщить занимающимся сведения об эмпирических распределениях; правилах построения вариационного ряда.
2. Показать с использованием статистического пакета STATGRAPHICS CENTURION возможности построения гистограммы, полигона и кумуляты.
3. Для расчетов использовать сборник примеров.

4. Обратить внимание на представление графиков по правилам ГОСТ 7.05-2008 (оформление диссертации и автореферата).

Практическое занятие №3.

Проверка нормальности распределения экспериментальных данных с использованием критериев согласия. Представление экспериментальных данных

1. Сообщить занимающимся сведения критериях проверки нормальности распределения экспериментальных данных.
2. Показать с использованием статистического пакета STATGRAPHICS CENTURION возможности проверки гипотезы о распределении экспериментальных данных с использованием критериев: хи-квадрат, лямбда, Шапиро-Уилки.
3. Для расчетов использовать сборник примеров.
4. Обратить внимание на представление статистических данных по правилам ГОСТ 7.05-2008 (оформление диссертации и автореферата).

Практическое занятие №4.

Критерии проверки статистических гипотез (независимые выборки). Табличное и графическое представление данных проверки статистических гипотез.

1. Сообщить занимающимся сведения критериях статистических гипотез (Стьюдента и Манна-Уитни)
2. Дать таблицу обоснования выбора статистического критерия.
3. Показать с использованием статистического пакета STATGRAPHICS CENTURION возможности проверки статистических гипотез
4. Для расчетов использовать сборник примеров.
5. Обратить внимание на представление статистических данных по правилам ГОСТ 7.05-2008 (оформление диссертации и автореферата).

Практическое занятие №5.

Критерии проверки статистических гипотез (связанные выборки). Табличное и графическое представление данных проверки статистических гипотез.

1. Сообщить занимающимся сведения критериях статистических гипотез (Стьюдента и Вилкоксона)
2. Дать таблицу обоснования выбора статистического критерия.
3. Показать с использованием статистического пакета STATGRAPHICS CENTURION возможности проверки статистических гипотез
4. Для расчетов использовать сборник примеров.
5. Обратить внимание на представление статистических данных по правилам ГОСТ 7.05-2008 (оформление диссертации и автореферата).

Практическое занятие №6

Корреляционный анализ. Представление данных корреляционного анализа.

1. Использование корреляционного анализа для решения научных задач в области ФКиС
2. Расчет коэффициентов Пирсона и Спирмена в пакете STATGRAPHICS CENTURION
3. Представление данных корреляционного анализа в диссертационной работе

Практическое занятие №7

Регрессионный анализ. Представление данных регрессионного анализа.

1. Использование регрессионного анализа для решения научных задач в области ФКиС
2. Простая регрессия в пакете STATGRAPHICS CENTURION
3. Полиномиальная регрессия в пакете STATGRAPHICS CENTURION
4. Множественная регрессия в пакете STATGRAPHICS CENTURION
5. Представление данных регрессионного анализа в диссертационной работе

Практическое занятие №8

Факторный анализ. Представление данных факторного анализа.

1. Использование факторного анализа для решения научных задач в области ФКиС
2. Выполнение факторного анализа в пакете STATGRAPHICS CENTURION
3. Интерпретация данных факторного анализа.
4. Представление данных факторного анализа в диссертационной работе

Презентации содержания лекций и практических занятий имеются на кафедре биомеханики.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ АСПИРАНТА: ОЧНАЯ ФОРМА ОБУЧЕНИЯ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ АНАЛИТИЧЕСКОЙ РАБОТЫ

Требования к выполнению работы 1.

Необходимо представить ксерокопии двух публикаций одной – на русском языке, другой – на английском языке по тематике своей научной работы (сроки публикации – не более 5 лет), в которых представлены результаты

первичной обработки экспериментальных данных (в числовой и графической форме), а также результаты проверки статистических гипотез. Помимо публикаций необходимо представить анализ правильности представления результатов.

Анализ представления результатов должен быть основан на:

- наличия (или отсутствии) представления объема выборки в таблицах;
- правильности представления записи среднего и ошибки среднего;
- правильности использования критериев проверки статистических гипотез;
- наличия (или отсутствии) границ доверительных интервалов на графиках.

Анализ публикаций представляется в печатной форме на листах формата А-4. Титульный лист должен содержать название работы, год обучения, группу, фамилию, инициалы автора.

Аналитическая работа должна быть представлена в сроки, согласованные с преподавателем.

По результатам выполнения аналитической работы аспирант может получить оценку от 3 до 5.

Требования к выполнению работы 2.

Необходимо представить ксерокопии двух публикаций одной – на русском языке, другой – на английском языке по тематике своей научной работы (сроки публикации – не более 5 лет) в которых представлены результаты корреляционного анализа. Помимо публикаций необходимо представить анализ правильности представления результатов.

Анализ представления результатов должен быть основан на:

- наличия (или отсутствии) представления объема выборки в таблицах;
- правильности представления коэффициентов корреляции;
- форме записи значимости коэффициентов корреляции;
- наличие (или отсутствию) графического представления информации (наличие корреляционного поля).

Анализ публикаций представляется в печатной форме на листах формата А-4. Титульный лист должен содержать название работы, год обучения, группу, фамилию, инициалы автора.

Аналитическая работа должна быть представлена в сроки, согласованные с преподавателем.

По результатам выполнения аналитической работы аспирант может получить оценку от 3 до 5.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ РАСЧЕТНО-ГРАФИЧЕСКОЙ РАБОТЫ

Необходимо представить:

1. Исходные данные, полученные при проведении собственных исследований (не менее двух тестов для двух групп (контрольной и экспериментальной)).
2. Результаты первичной обработки экспериментальных данных и проверки статистических гипотез (в виде таблиц)
3. Обоснование выбора критерия для проверки данных;
4. Графики, на которых представлены результаты собственных исследований с указанием границ доверительных интервалов.

Расчетно-графическая работа представляется в печатной форме на листах формата А-4. Титульный лист должен содержать название работы, год обучения, группу, фамилию, инициалы автора.

По результатам выполнения аналитической работы аспирант может получить оценку от 3 до 5.

Расчетно-графическая работа должна быть представлена в сроки, согласованные с преподавателем.

ЗАОЧНАЯ ФОРМА ОБУЧЕНИЯ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ПО ВЫПОЛНЕНИЮ ПЕРВОЙ АНАЛИТИЧЕСКОЙ РАБОТЫ.

Требования к выполнению работы.

Необходимо представить ксерокопии двух публикаций одной – на русском языке, другой – на английском языке по тематике своей научной работы (сроки публикации – не более 5 лет), в которых представлены результаты первичной обработки экспериментальных данных (в числовой и графической форме), а также результаты проверки статистических гипотез. Помимо публикаций необходимо представить анализ правильности представления результатов.

Анализ представления результатов должен быть основан на:

- наличия (или отсутствию) представления объема выборки в таблицах;
- правильности представления записи среднего и ошибки среднего;
- правильности использования критериев проверки статистических гипотез;
- наличия (или отсутствию) границ доверительных интервалов на графиках.

Анализ публикаций представляется в печатной форме на листах формата А-4. Титульный лист должен содержать название работы, год обучения, группу, фамилию, инициалы автора.

По результатам выполнения аналитической работы аспирант может получить оценку от 3 до 5.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ПО ВЫПОЛНЕНИЮ ВТОРОЙ АНАЛИТИЧЕСКОЙ РАБОТЫ

Требования к выполнению работы

Необходимо представить ксерокопии двух публикаций одной – на русском языке, другой – на английском языке по тематике своей научной работы (сроки публикации – не более 5 лет) в которых представлены результаты корреляционного анализа. Помимо публикаций необходимо представить анализ правильности представления результатов.

Анализ представления результатов должен быть основан на:

- наличию (или отсутствию) представления объема выборки в таблицах;
- правильности представления коэффициентов корреляции;
- форме записи значимости коэффициентов корреляции;
- наличию (или отсутствию) графического представления информации (наличия корреляционного поля).

Анализ публикаций представляется в печатной форме на листах формата А-4. Титульный лист должен содержать название работы, год обучения, группу, фамилию, инициалы автора.

По результатам выполнения аналитической работы аспирант может получить оценку от 3 до 5.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ПО ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ

Требования к выполнению контрольной работы

Необходимо представить:

1. Исходные данные, полученные при проведении собственных исследований (не менее двух тестов для двух групп (контрольной и экспериментальной).
2. Результаты первичной обработки экспериментальных данных и проверки статистических гипотез (в виде таблиц)
3. Обоснование выбора критерия для проверки данных;
4. Графики, на которых представлены результаты собственных исследований с указанием границ доверительных интервалов.

Расчетно-графическая работа представляется в печатной форме на листах формата А-4. Титульный лист должен содержать название работы, год обучения, группу, фамилию, инициалы автора.

По результатам выполнения аналитической работы аспирант может получить оценку от 3 до 5.

Результаты анализа публикаций и контрольная работа могут быть также представлены аспирантом заочной формы обучения в межсессионный период, в часы консультаций преподавателя.