

**НАЦИОНАЛНА СПОРТНА АКАДЕМИЯ
“ВАСИЛ ЛЕВСКИ”**

КАТЕДРА “ТЕОРИЯ НА СПОРТА”

УЧЕБНИ ЗАПИСКИ

Статистически методи във физическото възпитание

МАГИСТЪРСКА ПРОГРАМА

Физическо възпитание

УЧЕБНИ ЗАПИСКИ

**“Статистически методи във физическото
възпитание”**

ЗА СТУДЕНТИ ОТ МАГИСТЪРСКА СТЕПЕН НА НСА

Автор : доц. В. Гигова
Сектор “Статистика”
Катедра “ТС”
2009 г.

Съдържание

Раздел I – Обща характеристика на обучението по	5
Анотация.....	5
Учебна програма	6
Указания за ползване на учебното помагало	7
Протичане на заниманията	8
Изпитни изисквания	8
Раздел II – Учебно съдържание.....	9
Тема 1 – Основни понятия.....	9
1.1. Генерална съвкупност и извадка	9
1.2. Измерителни скали.....	10
1.3. Преглед на статистическите методи, прилагани в научните изследвания в областта на училищния спорт.....	11
Тема 2 – Въвеждане на данни в SPSS	13
2.1. Обща характеристика на SPSS	13
2.2. Въвеждане на данни	16
2.3. Изчисляване на производни променливи	19
2.4. Отстраняване на грешки, възникващи при въвеждане на данни	20
Тема 3 – Вариационен анализ	23
3.1. Въведение.....	23
3.2. Показатели за средно равнище.....	23
3.3. Показатели за разсейване.....	24
3.4. Форма на разпределението	26
3.5. Вариационен анализ със SPSS	31
3.5.1. Обработка на целия масив.....	31
3.5.2. Обработка на подгрупи от базата данни.....	33
3.6. Представяне на резултати от вариационен анализ	34
Тема 4. Сигмален метод за изработване на нормативи.....	36
4.1. Основни понятия	36
4.2. Задачи за упражнение.....	40
Тема 5. Корелационен анализ	41
5.1. Въведение.....	41
5.2. Видове коефициенти на корелация	41
5.3. Статистическа значимост на коефициентите на корелация	42
5.4. Корелационен анализ със SPSS	43
5.5. Представяне на резултатите от корелационен анализ	44
Тема 6. Статистическа проверка на хипотези	46
6.1. Основни понятия	46
6.2. Алгоритъм за проверка на хипотези.....	47
6.3. Критерии за проверка на хипотези.....	49
6.4. Изчисляване на t-критерия на Стюдънт за независими извадки.....	52
6.5. Изчисляване на t-критерия на Стюдънт за зависими извадки.....	56
6.5.1. За целия масив от данни.....	56
6.5.2. За отделни подгрупи	58

Тема 7. Обработка на данни от педагогически експеримент	59
7.1. Описание на експерименталния дизайн	59
7.2. Статистически показатели, прилагани при обработка на данните	61
Литература	64
Изпитен проект	65
Задача 1.	67
Задача 2.	68
Задача 3.	69
Задача 4.	70
Задача 5.	71
Задача 6.	72
Задача 7.	73
Раздел IV - Приложения	74
Приложение 1. Данни за практическите упражнения	74
Приложение 2. Данни за упражнение 2	75
Приложение 3. Критични стойности на коефициента на асиметрия (A_s)	76
Приложение 4. Критични стойности на коефициента на ексцес (E_x)	76
Приложение 5. Нормирано нормално разпределение в %	77
Приложение 6. Критични стойности на коефициента на обикновена линейна корелация на Пирсън (r)	77
Приложение 7. Критични стойности на коефициента на рангова корелация на Спирмън	78
Приложение 8. Критични стойности на t -критерия на Стюдънт	78

Обща характеристика на обучението по модула

Анотация

Целта на учебната дисциплина е да се задълбочат познанията на студентите по отношение на приложението на статистически методи в научните изследвания във физическото възпитание. Предмет на изследването в тази област на научното познание са преди всичко физическото развитие и дееспособност на подрастващите, средствата и методите за тяхното подобряване. Признаците, които описват състоянието и развитието на двигателните способности на учащите в голямата си степен са количествени, поради което за анализ на резултатите от научните изследвания се прилагат най-вече параметричните статистически методи. Изхождайки от това съображение, в модула са застъпени вариационен, корелационен анализ и параметрични критерии за проверка на хипотези. Практическата част от обучението е насочена към овладяване на специализирания програмен продукт SPSS.

След овладяване на учебното съдържание студентите трябва да могат:

1. Да организират бази данни, да дефинират и въвеждат променливи в SPSS;
2. Да подбират и прилагат подходящ статистически анализ, съобразен с изследователските си задачи и естеството на променливите величини;

3. Да представят и интерпретират по коректен начин резултатите от статистическата обработка на данните;
4. Самостоятелно да обработят и анализират данни за изготвяне на научен продукт - статия, дипломна работа и др.

Учебна програма

№	Тема	Форма
1.	Основни понятия от общата теория на статистиката (статистически съвкупности и извадки, измерителни скали, общ преглед на статистическите методи, прилагани във физическото възпитание)	Л
2.	Обща характеристика на SPSS 15.0, дефиниране на променливи, въвеждане на данни, трансформиране на променливи.	У
3.	Характеристика на разпределението на количествени променливи (вариационен анализ).	Л, У
4.	Статистически методи за разработване на нормативи (сигмален метод, Z и Т-оценки).	Л, У
5.	Корелационен анализ.	Л, У
6.	Параметрични критерии за проверка на хипотези.	Л, У
7.	Обработка на данни от педагогически експеримент.	Л, У

Указания за ползване на учебното помагало

Настоящото учебно помагало е структурирано в съответствие с учебната програма по дисциплината „Статистически методи във физическото възпитание” за магистърска програма „Физическо възпитание”. На страниците, предвидени за всяка от темите има линкове към:

- **Теоретичния материал по темата;**
- **Задачите** за изучаване на съответния анализ със SPSS и тези, залегнали в изпитния проект;
- **Учебни филми** с решенията на задачите за изучаване. За да могат да се възпроизведат трябва да имате инсталиран Flash Player. За да се стартира даден филм е необходимо да се активира съответния линк и да се кликне на два пъти последователно Run. В началото на демонстрацията е изписана темата, за която се отнася, текста на задачата, чието решение е онагледено (това са задачите за изучаване). След това може да бъде проследен алгоритъмът за решение на задачата. В случай, че е необходимо, скоростта на появяване на отделните кадри може да се регулира с навигационните бутони в долния край на екрана. След приключване на демонстрацията, учебният филм се затваря автоматично;
- **Интерактивни тестове.** В тях е онагледено решението на задачите от изпитния проект и студентите биха могли да ги ползват във времето за самостоятелна работа. Те се активират по аналогичен на учебните филми начин, но след изписване на текста на задачата студентът трябва да реализира алгоритъма, като кликва с мишката в съответните менюта, диалогови прозорци, бутони и т.н. На местата, на които се изисква изписване на текст, същият се появява автоматично. При правилно осъществяване на стъпка от алгоритъма се появява надпис “Correct”. В случай, че стъпката е осъществена неправилно се появява съобщение “Error”. След две последователни неправилни кликания тестът продължава, но за крайния резултат се отчита грешка. Тестът се счита успешно положен, ако 75% от стъпките са осъществени правилно. Резултатът от тези тестове не участва във формирането на крайната оценка по модула.

Протичане на заниманията

Учебните занимания в **редовната форма** протичат в следната последователност – преподавателят запознава студентите с предвидения в учебната програма теоретичен материал, след което се пристъпва към практическата част на заниманието. Към всяка от темите има предвидени по 2 задачи – първата е за изучаване, а втората е част от изпитния проект. Първата се решава под ръководството на преподавателя. Студентите се запознават с начина на представяне и интерпретиране на резултатите от обработката. Втората задача се решава самостоятелно, резултатите се записват в изпитния проект (раздел III), а отговорите на поставените въпроси се попълват във времето за самостоятелна работа на студента.

При обучение в **дистанционна форма** студентите трябва да извършат самостоятелно посочените дейности, като за практическото решаване на задачите ползват алгоритмите за работа със SPSS, описани в помагалото или учебните филми и интерактивните тестове.

Изпитни изисквания

За успешно приключване на учебната дисциплина студентите трябва да владеят понятийния апарат по темите, описани в учебната програма и теоретичната част на заниманията. Уменията да обработят, представят и интерпретират научна информация магистрите демонстрират с изготвянето на изпитния проект, представен в раздел III. По избор, проектът може да се впише на ръка в настоящото помагало или да се попълни в предоставения файл с макет на изпитния проект (Ph_education_exam.doc).

Изпитът се провежда под формата на събеседване върху анализите, залегнали в изпитния проект.

Тема 1 – Основни понятия

1.1. Генерална съвкупност и извадка

Важен момент при планиране и провеждане на научно изследване е ясното дефиниране на обекта и предмета на изследване.

Обект на изследването са единиците, чрез които се проявява изучаваното явление. От статистическа гледна точка те представляват т.нар. **генерална съвкупност (ГС)**. По дефиниция генералната съвкупност е множество от обекти, които са подбрани по определени критерии и имат поне един общ признак (напр. в ГС „Ученици в 5 клас” попадат обекти, които имат два общи признака - че са ученици, че са в 5 клас). Изследването на всичките единици на съвкупността от една страна, не винаги е възможно (особено ако тя е много голяма), а от друга - създава големи организационни затруднения, изисква разходването на значителни финансови средства и труд. Поради това обикновено се изследва част от съвкупността, която се нарича **извадка**. Идеята е по данни от извадката да се направят заключения за изучаваната съвкупност. За да бъдат коректни подобни обобщения, извадката трябва достатъчно точно да възпроизвежда характеристиките на цялата генерална съвкупност. Подобна извадка се нарича **репрезентативна**.

Предмет на изследването са характеристиките на обектите, се наричат **признаци** (ръст, тегло, взривна сила на долните крайници и др.). Единиците на съвкупността си приличат или се различават по състоянието на тези си характеристики. За да могат да бъдат изучавани, същите трябва да бъдат

измерени. Измерването е процес, при който дадено състояние на признака се описва с число, в съответствие с определена измерителна скала.

1.2. Измерителни скали

В зависимост от това дали числата, получени в резултат на измерването носят информация за количеството на притежавания признак, или обозначават само принадлежност към дадена категория, измерителните скали биват количествени и качествени (таблица 1.1).

Таблица 1.1

Наименование	Определение	Пример
Качествени скали – при тях не се измерва количество на притежавания признак и няма мерни единици		
Номинална (категорийна)	Обектите се разпределят в еднородни групи между които няма различия в интензитета на признака. Принадлежността към дадена група се означава с число – код.	пол (1- мъж, 2- жена)
Ординална	Обектите се разпределят в еднородни групи, между които се наблюдават различия в интензитета на признака.	училищна степен (начален, среден, горен курс)
Рангова	Обектите се подреждат (ранжират), въз основа на определени правила, състоянието на признака се означава с число-рангов номер.	класиране в училищното първенство по баскетбол
Количествени скали– измерва се количеството на признака и се изразява в съответните мерни единици (м., сек, кг, бр.)		
Интервална	Числата носят информация за количеството на признака, но стойност 0 не означава отсъствие на признака, а е условно възприета.	гъвкавост на наклона на трупа напред
Пропорционална	Числата носят информация за количеството на признака, 0 означава отсъствие на признака	резултата в тест "Лицеви опори"

1.3. Преглед на статистическите методи, прилагани в научните изследвания в областта на физическото възпитание

Основна цел на училищното физическо възпитание е подобряване на физическото развитие, дееспособност и техническите умения на учениците със средствата на спорта. Тъй като изброените характеристики, от статистическа гледна точка са количествено измерими признаци, най-често в научни разработки се прилагат параметрични методи (вариационен, корелационен анализ, параметрични критерии за проверка на хипотези и др.).

В таблица 1.2 е направен паралел между прилагания изследователски метод, решаваната познавателна задача, статистическите методи и показатели, които се прилагат и темата в настоящата учебна дисциплина, в която те са разгледани.



.....

.....

.....

.....

Таблица 1.2

Изследователски метод	Изследователска задача	Статистически метод	Статистически показатели	Тема
Анкетно проучване	Характеристика на разпределението на променливите величини.	Честотен анализ	Брой и процент от случаите, попадащи в дадена категория	-
	Изследване на зависимости.	Корелационен анализ	Коефициент на контингенция (C);	-
	Сравняване на разпределението на променливите величини.	Статистическа проверка на хипотези	χ^2 - критерий на Пирсън	-
Спортно-педагогическо тестиране	Характеристика на разпределението на променливите величини	Вариационен анализ	$X_{\min}$, $X_{\max}$, R, $\bar{X}$, S, V, As, Ex	Тема 3
	Изработване на нормативи	Сигмален метод за изработване на нормативи	Z оценки	Тема 4
	Изследване на зависимостите между променливите величини	Корелационен анализ	Коефициент обикновена линейна корелация на Пирсън (r)	Тема 5
	Сравняване на разпределението на променливите величини по пол, възраст и др.	Проверка на хипотези	t-критерий на Стюдънт за независими и зависими извадки	Тема 6
Педагогически експеримент	Сравняване на разпределението на променливите величини	Статистически методи за описание на развитие	Абсолютен (d) и относителен (d%) прираст	Тема 7
		Проверка на хипотези	t-кр. на Стюдънт за зависими и независими извадки	

Тема 2 – Въвеждане на данни в SPSS

2.1. Обща характеристика на SPSS

Statistical Package for Social Science (SPSS – статистически пакет за социални науки) е компютърна програма, работеща в средата на операционна система Windows, която е специализирана за систематизиране, обработка и анализ на статистическа информация. Както подсказва наименованието, тя е насочена предимно към изследвания в областта на социологическите и маркетинговите проучвания, в т.ч. спортната наука. Характерна особеност програмата е, че в нея е отделено значително място на непараметричните статистически критерии. Въпреки това поради богатството на заложените методи и големите възможности за сортиране и трансформиране на променливи с успех може да се ползва във всички области на физическото възпитание и спорта.

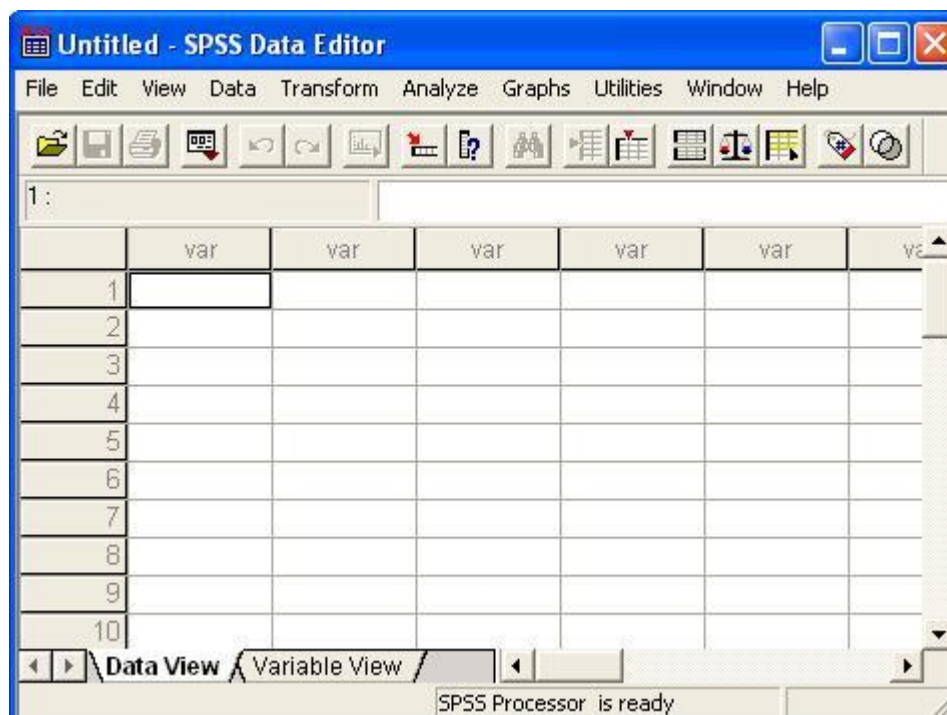
Стартирането на програмата става чрез двукратно кликуване с левия бутон на мишката върху иконата за бързо избиране или Start>All programs>SPSS for Windows.



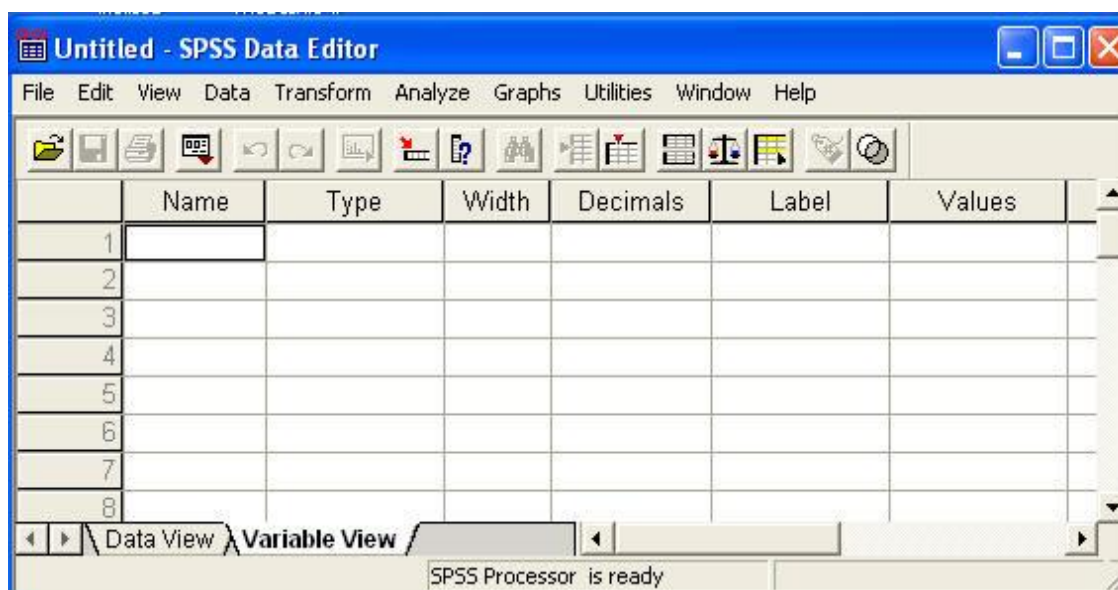
Фиг. 2. 2

SPSS разполага с два основни прозореца за управление на програмата – за въвеждане и редактиране на данни (Data Editor – фиг.2.2) и за отпечатване на резултатите от статистическата обработката - таблици и графики (Output Viewer - фиг.2.3).

В прозореца за данни (SPSS Data Editor) се дефинират, въвеждат и редактират статистическите данни. Той се отваря автоматично при стартиране на програмата. Информацията, записана в него, се съхранява като самостоятелен файл с разширение sav. Data Editor дава възможност за самостоятелен достъп до данните (в т.нар. Data View) и до наименованието и характеристиките на променливите (Variable View).

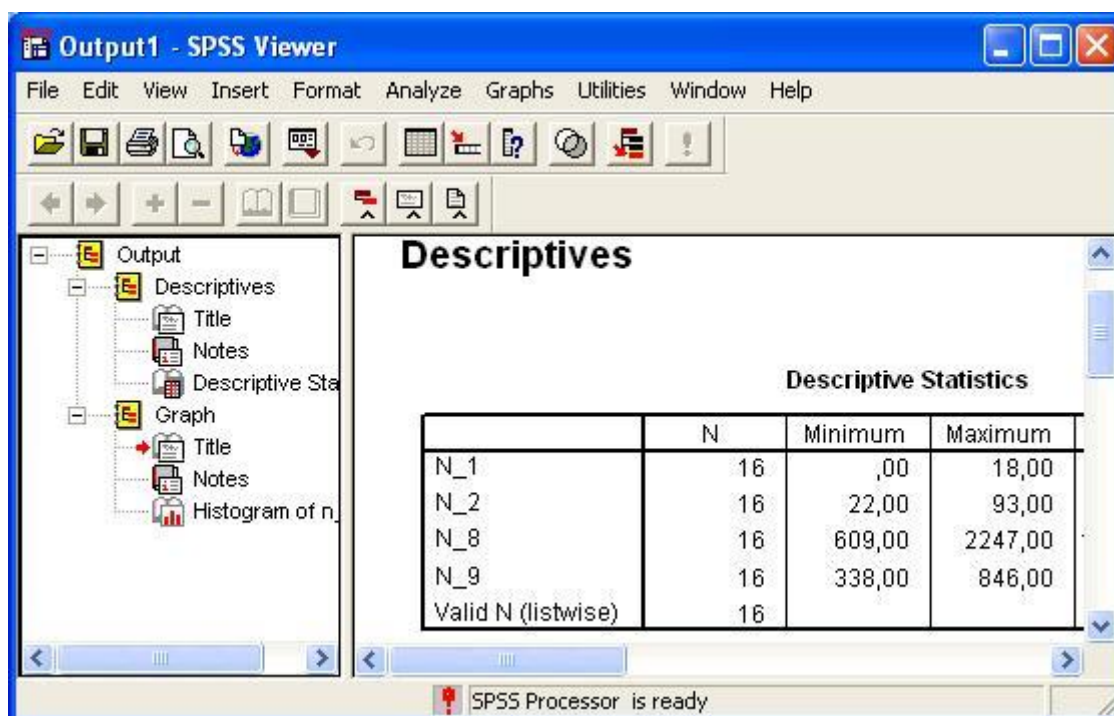


Фиг. 2.2.



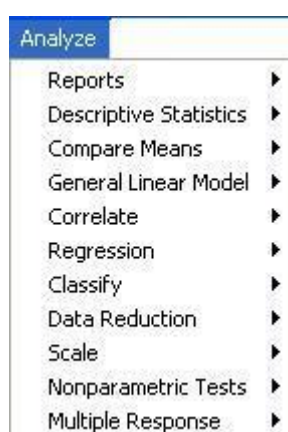
Фиг. 2.3.

Прозорецът за резултати от обработката (Output Viewer – фиг.2.4) се активира автоматично след осъществяване на статистическа обработка. В него се изписват изчислените статистически показатели или графични изображения. В случай, че е необходимо да се съхрани информацията, той трябва да бъде съхранява като самостоятелен файл с разширение .spo.



Фиг. 2. 4.

Всеки прозорец на SPSS има характерна за него лента с менюта. Активни винаги са менютата Analyze (фиг.2.5) и Graphs което дава възможност за лесен достъп до осъществяване на статистически анализи и построяване на графични изображения.



Фиг 2.5

Менюто служи за задаване на статистическите анализи. В модула, предназначен за магистърска програма "Физическо възпитание" ще се запознаем с:

- Descriptive Statistics (описателна статистика)
- Compare Means (проверка на хипотези)
- Correlate (корелационен анализ)

Част от иконите, намиращи се на лентата с менютата са еднакви с тези, характерни за всички програми, работещи под Windows, а други са специфични

за SPSS. Задържането на показалеца на мишката над всяка от тях дава кратка бърза информация за функциите, които изпълнява.

Лентата за състоянието се намира се под активните прозорци и дава информация за това активна ли е програмата, разделен ли е масива от данни на подгрупи за самостоятелна обработка и др.

[F:\Arhiv_4\myWebSites\Physical_education\index.html](F:\Arhiv_4\myWebSites\Physical_education\index.html)

2.2. Въвеждане на данни

В резултат на провеждане на научно изследване, се получава информация за състоянието на изследваните признаци (тестове, отговори на въпроси от анкета и т.н.) на обхванатите статистически единици (изследвани лица, отбори, училища и т.н.). Тези резултати се записват в таблици, известни под наименованието бази от данни

	Group	t1_1	t1_2
1	experimental	5,71	5,2
2	experimental	6,53	5,71
3	experimental	5,86	5,31
4	experimental	6,27	5,79
5	experimental	5,7	5,21
6	kontrol	6,1	6
7	kontrol	6,08	6
8	kontrol	5,89	5,9
9	kontrol	5,89	5,81
10	kontrol	6	5,95

База данни, която съдържа информация за 10 статистически единици по 3 признака.

Фиг. 2. 6

По редове се записва информация за състоянието на изследваните признаци на дадена статистическа единица. В първата колона се изписва поредният номер на наблюдението (изследваната статистическа единица), а в следващите – резултатите от измерване на отделните признаци (фиг. 2.6). Тъй като отделните единици имат различни резултати (стойностите в колоните варират), те представляват променливи величини. В по-нататъшното изложение ще ползваме наименованието променливи, като под него се разбира резултатът от измерване на даден признак, една колона от базата данни. Представената на фиг. 2.6 база данни съдържа информация за 10 наблюдения (статистически единици) по 3 променливи.

При компютърна обработка на данните е необходимо първо да се изгради **макетът на базата данни**. Това означава променливите величини се обозначават по начин, който позволява на SPSS, да ги открива и включва в различни изчислителни процедури, а резултатите от обработката да се появяват във вид, годен за разчитане. За примера, представен на фиг. 2.6 трябва да се дефинират три колони – изследователска група (експериментална и контролна) и две колони за резултати от тестове с кратко обозначение t1_1 и t1_2, в които са изписани резултатите в тест 30 м. гл. б. в началото и края на педагогически експеримент.

Задача за разучаване 2.1

Дефинирайте променливите group, t1_1 и t1_2, описани в приложение 1 и въведете стойностите от таблица 2.6. Съхранете файла с данни, като за наименование посочите факултетния си номер.

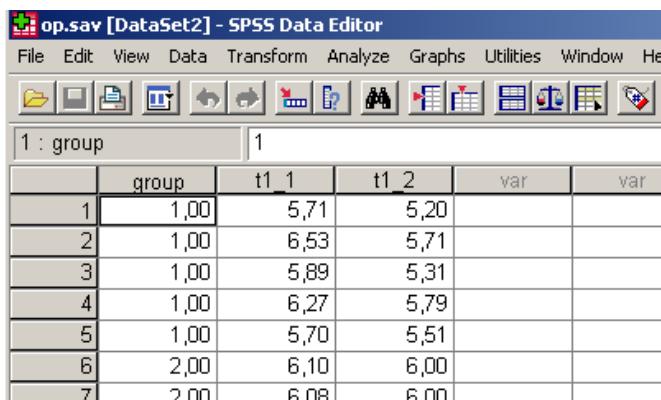
С процедурата по **дефиниране на променливите** се задават кратко и разширено наименование на променливите (в терминологията на SPSS, съответно Variable Name и Variable Label). В случай, че променливите са категорийни е необходимо да се зададат кодовете и наименованията на отделните категории (в терминологията на SPSS съответно Value и Value Label). Процедурата се осъществява в прозореца Variable View (фиг. 2.3).

- В колона **Name** се записва краткото наименование на променливата. То трябва да бъде уникално: (не може две променливи да имат едно и също име), да започва с буква, като SPSS не прави разлика между малки и главни букви, да не съдържа повече от 8 символа, като не се допускат интервали или оператори за аритметични действия (% , / , * , - , +).
- В колона **Label** се записва разширеното наименование на променливата.
- Кодовете и етикетите на категорийните променливи се задават в колона **Values**.

- В колоната Measure се задава скалата, в която е измерена променливата (Scale- количествена, Ordinal-ординална и Nominal-номинална).
- След като се дефинират променливите и се изгради т.нар. макет на таблицата с данни, се пристъпва към втория етап, а именно въвеждане на стойностите.

Въвеждането на стойностите става в прозореца **Data View**. Те се изписват с помощта на клавиатурата. Първоначално стойността се появява на реда за въвеждане над прозореца за данни. След като въвеждането се потвърди с Enter, числото се изписва на мястото, където е била активната клетка.

Стойностите на категориите променливи се въвеждат със зададените им кодове, а потребителят може да задава дали в базата данни да се появяват кодовете или етикетите (фиг. 2.7).



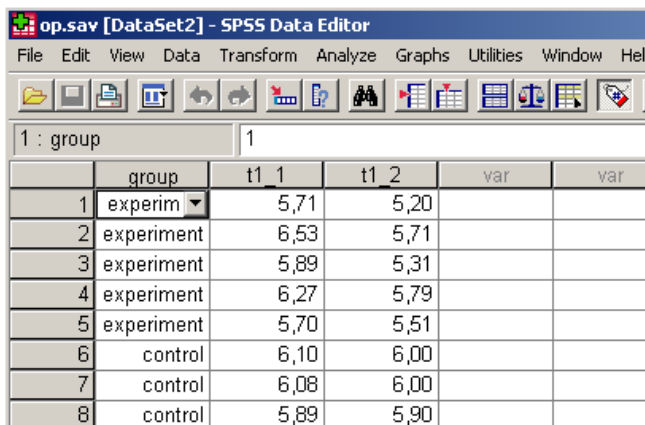
op.sav [DataSet2] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : group 1

	group	t1_1	t1_2	var	var
1	1,00	5,71	5,20		
2	1,00	6,53	5,71		
3	1,00	5,89	5,31		
4	1,00	6,27	5,79		
5	1,00	5,70	5,51		
6	2,00	6,10	6,00		
7	2,00	6,08	6,00		

Ако иконата за етикети -  не е активирана се появяват кодовете.



op.sav [DataSet2] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : group 1

	group	t1_1	t1_2	var	var
1	experim	5,71	5,20		
2	experiment	6,53	5,71		
3	experiment	5,89	5,31		
4	experiment	6,27	5,79		
5	experiment	5,70	5,51		
6	control	6,10	6,00		
7	control	6,08	6,00		
8	control	5,89	5,90		

Ако иконата е активирана се появяват етикетите на категориите.

Фиг. 2.7

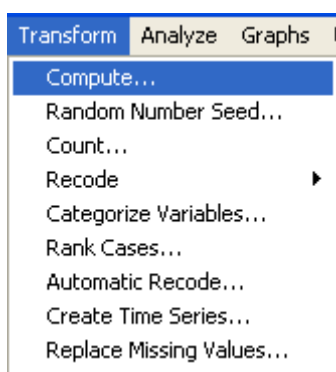
2.3. Изчисляване на производни променливи

Твърде често в научноизследователската дейност във физическото възпитание и спорта се налага преобразуване на оригиналните данни. Например, при обработка на данни от педагогически експеримент е необходимо да се изчисли прирастът на резултатите на изследваните лица. За целта е необходимо да се образува нова променлива, която да съдържа разликите между второ и първо изследване на всяко едно изследвано лице.

Задача за разучаване 2.2

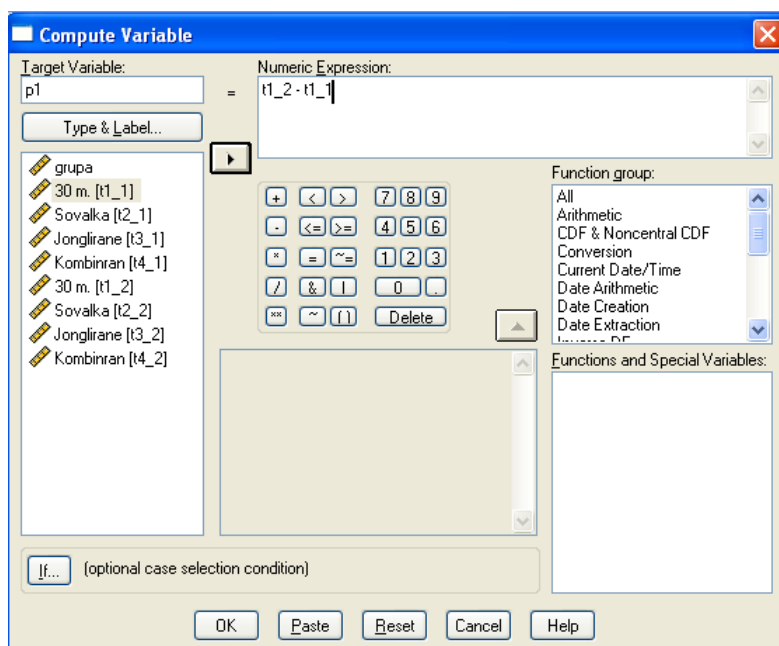
По данните от файла Ph_education.sav изчислете прираста на резултатите в тест 30 м. гл. б. За целта изчислете разликите между от резултатите от второто (t1_2) и първото изследване (t1_1). На новата променлива задайте наименование p_t1.

Изчисляването на нови променливи, въз основа на съществуващи се осъществява от менюто **Transform>Compute** (фиг. 2.8).



Фиг. 2.8

- Отваря се диалоговият прозорец Compute Variable (фиг. 2.9).



Фиг. 2.9

- От полето със списъка на променливите изберете първата променлива, която участва във формулата и с бутона  я изпратете в полето Numeric Expression (в случая резултатите от I изследване-t1_1);
- Кликнете върху оператора за аритметичното действие, което трябва да се извърши (в случая -);
- Последователно задайте следващите променливи и аритметични действия (в случая резултатите от второто изследване – t1_2);
- Когато изразът е готов, в полето Target Variable изпишете името на новата променлива;
- Потвърдете с OK и в таблицата с данни ще се появи новата променлива.

[F:\Arhiv_4\myWebSites\Physical_education\index.html](F:\Arhiv_4\myWebSites\Physical_education\index.html)

2.4. Отстраняване на грешки, възникващи при въвеждане на данни

Проблем: Изписано е погрешно число и грешката е установена преди да е натиснат клавишът Enter

Изход: Символи преди курсора се изтриват с клавиша Backspace, а тези след курсора - с клавиша Del.

Проблем Изписано е погрешно число и грешката е установена след като е натиснат клавишът Enter.

Изход:

- Активирайте клетката с погрешно въведената стойност.
- Запишете новата стойност.
- Натиснете клавиша Enter - старото число ще се подмени с новото

Проблем Погрешно е въведена стойност на ред, извън масива с данни (например изследвани са 20 лица, а е въведено число на 21 ред), SPSS автоматично създава случай за този ред. Дори да изтриете числото, редът си остава изпълнен със запетайки, т.е. програмата счита тези стойности за липсващи. Статистическите анализи, които се осъществяват впоследствие ще съдържат грешка, защото при изчисленията програмата ще работи с обем на извадката 21 души и ще отчита 1 липсваща стойност.

Изход За да изтриете погрешно въведен ред:

- Чукнете с мишката в най-лявата сива клетка на съответния ред - той се маркира.
- Натиснете клавиша Del.

Проблем: Погрешно е въведена стойност в празна колона (например изследваните показатели са 10, а е записано число в 11-а колона) SPSS автоматично създава нова променлива, означена с VAR00001. Тя може да Ви затрудни при задаване на показателите за обработка.

Изход За да изтриете погрешно въведена колона:

- Чукнете с мишката в заглавната (най-горната) сива клетка - колоната се маркира.
- Натиснете клавиша Del.

Проблем: Някои от случаите в извадката нямат резултати по част от показателите. Колебаете се, дали да ги включите в масива от данни.

Изход: Не губете информацията, която все пак имате за тези случаи. Въведете числата, като оставяте някои от клетките на съответния ред празни. Програмата счита тези стойности липсващи. По подразбиране SPSS ги замества или изключва (в зависимост от осъществявания анализ), така че липсата им не пречи да се осъществят изчислителните процедури.

[F:\Arhiv_4\myWebSites\Physical_education\index.html](F:\Arhiv_4\myWebSites\Physical_education\index.html)

Тема 3 – Вариационен анализ

3.1. Въведение

Единиците на статистическите съвкупности се различават (варират) по величината на наблюдаваните признаци. Вариацията е присъщо свойство на статистическите съвкупности. Ако тя не съществуваше, статистическите проучвания щяха да бъдат излишни. Тези различия са породени от различни по своето естество фактори, които най-общо се разделят на две категории:

- **закономерни**, които действат еднакво върху всички изучавани единици и определят типичното, закономерното състояние на признака (например типично, характерно за баскетболистите е, че имат висок ръст).
- **случайни**, които действат с различна сила върху изучаваните единици и водят до индивидуални отклонения във величината на признака (отделните баскетболисти се различават по своя ръст).

Задача на статистическото проучване е да даде обобщаващи количествени характеристики на състоянието на изследвания признак. За целта е необходимо, **на първо място**, да се установи какво е типичното за дадената съвкупност стойност на признака, за да се опише влиянието на закономерно действащите фактори. **На второ място**, от особена важност е характеризиране на разсейването, варирането на стойностите на признака, за да се отчете влиянието на случайните фактори. **Трета група** от показатели носят информация за формата на разпределението на емпиричните стойности и степента на съответствието му с някое от известните теоретични разпределения, най-често с това на нормалното разпределение.

3.2. Показатели за средно равнище

Показателите за средно равнище представляват обобщаващи количествени характеристики, с помощта на които се описва типичното, характерното

състояние на изследвания признак. Най-често употребяваните показатели са **средна аритметична величина ($\bar{X}$), мода (M_o) и медиана (Me).**

- **Средната аритметична величина ($\bar{X}$)** е най-често използваният в научно-изследователската практика показател за средно равнище. Изчислява се по формулата:

$$\bar{X} = \frac{\sum X}{n} \quad 3.1$$

Където:
 $\sum X$ - сума на стойностите
 n – брой на наблюденията

- **Модата (M_o)** е онази стойност, която има най-голяма абсолютна честота, т.е. която се повтаря най-голям брой пъти.
- **Медианата (Me)** е онази стойност на променливата величина, която заема средно положение във вариационния ред и го дели на две равни части. Броят на стойностите, които са по-малки или равни на медианата, е равен на броя на тези, които са по-големи или съвпадат с нея.

3.3. Показатели за разсейване

Показателите за разсейване дават количествена характеристика на величината на отклоняване на стойностите от типичното, средното равнище. За измерване на статистическото разсейване най-често се използват показателите **размах (R), стандартно отклонение (S) и коефициент на вариация (V).**

- **Размахът (R)** е най-елементарният измерител на статистическото разсейване. Той представлява разликата между най-голямата и най-малката стойност на променливата величина и носи информация за диапазона, в който варират значенията на признака. Изчислява се по формулата:

$$R = X_{\max} - X_{\min} \quad (3.2)$$

Където:
 $X_{\max}$ - най-голямата стойност на променливата
 $X_{\min}$ - най-малката стойност на променливата

Тъй като в изчисляването на размаха участват само двете крайни стойности на вариационния ред, той е твърде **неустойчив измерител на разсейването**.

- **Стандартното отклонение (S)** е най-прецизният и често използван показател за разсейване. Той описва големината на отклонение на стойностите на променливата от средната аритметична величина. Формулата, която го дефинира е:

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}} \quad (3.3)$$

Където:
 X_i – всяка една стойност на променливата в извадката
 $\bar{X}$ – средна аритметична величина
 n – обем на извадката

Изброените статистически показатели за разсейване изразяват степента на различията в същата мерна единица, в която е изследваният признак. Това създава затруднения при сравняване на вариацията на различни признаци.

Показателят, който дава възможност за сравняване на различията на признаци, изразени в различни мерни единици или относно различна средна стойност е **коэффициентът на вариация (V)**. Той се изчислява по формула:

$$V = \frac{S}{\bar{X}} \cdot 100 \quad (3.4)$$

Където:
 S – стандартно отклонение
 $\bar{X}$ – средна аритметична величина

Освен за сравняване на разсейването на стойностите на различни променливи, коэффициентът на вариация се ползва и за оценяване на степента на разсейването (еднородността на извадката):

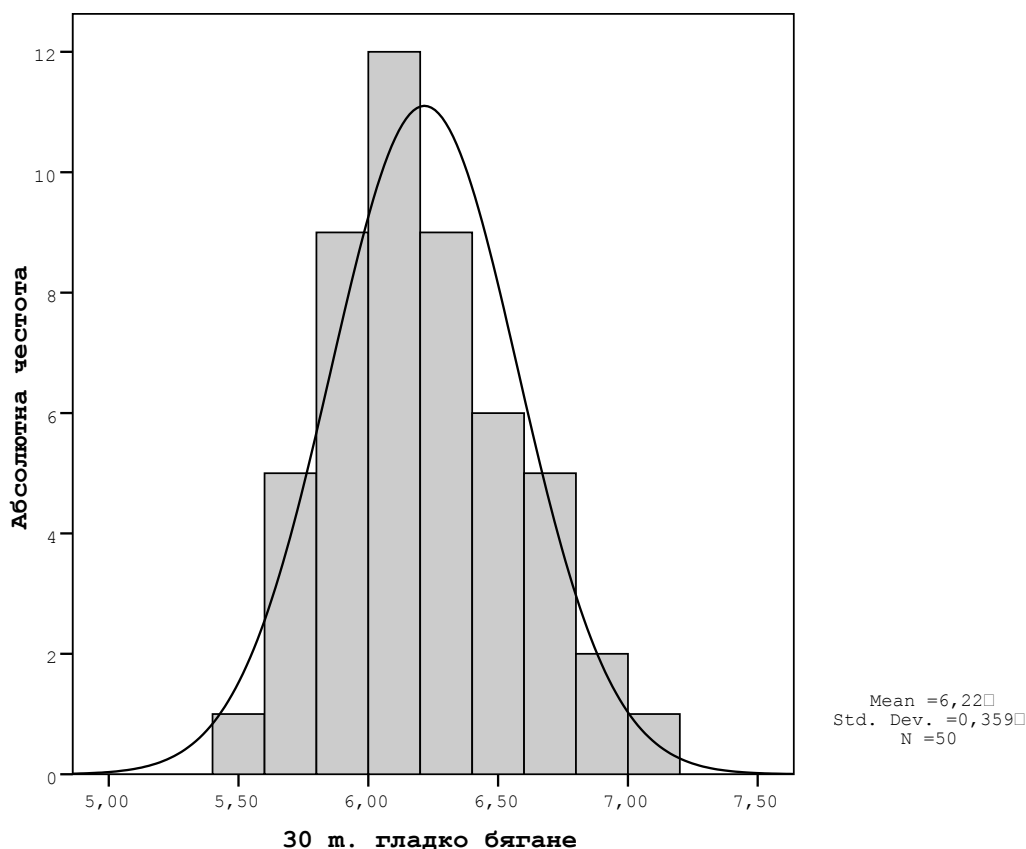
- Счита се, че разсейването на признака е малко (**извадката е еднородна**), когато стойността му е до 10-12%.
- Между 10 и 30% извадката е **приблизително еднородна**.
- Когато е над 30% разсейването на признака е голямо (**извадката е силно нееднородна**).

-

3.4. Форма на разпределението

В настоящото помагало многократно е използван терминът “разпределение на променливата величина”, с който се означава съответствието между стойностите на променливата и техните абсолютни или относителни честоти. Тъй като статистическите проучвания се основават на изследването на извадки, това разпределение се нарича емпирично. Ако то се онагледява графично, по конфигурацията на графичния образ може да се съди за закономерността, на която е подчинено изучаваното масово явление, и да се търси аналог сред т. нар. теоретични разпределения.

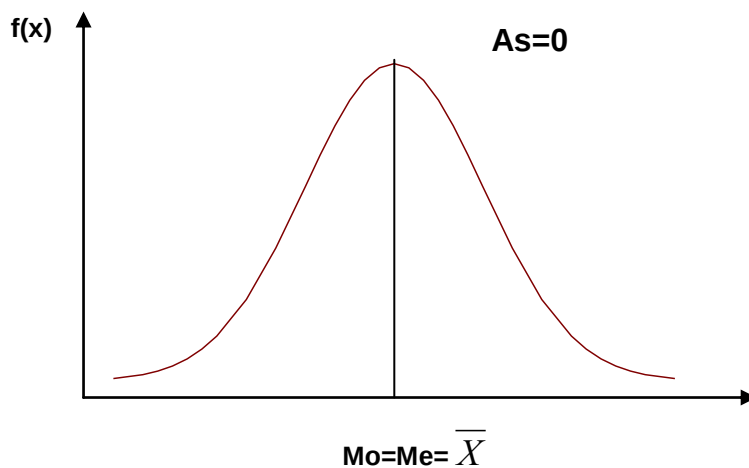
Изключително важно място в теорията на статистиката заема **нормалното разпределение**. То носи наименованието Гаус-Лапласово по имената на немския математик Карл Гаус (1777 – 1855) и френския математик Пиер Лаплас (1739 – 1827), които са го изследвали и описали теоретично. Нормалното разпределение стои в основата на изясняване на важни положения от репрезентативните статистически проучвания, интервалното оценяване, проверката на хипотези, методите за изготвяне на нормативи и др. Прилагането на параметрични методи за проверка на хипотези и за изследване на зависимости между променливите величини също изисква предварително изясняване на въпроса доколко разпределението на емпиричните данни следва закона на нормалното разпределение. Ще се спрем накратко на някои негови особености:



Фиг. 3.1

На фигура 3.1 е представена хистограма на резултатите в тест 30 м. гл. б., които се ползват за практическите упражнения в настоящото помагало. Както се вижда, разпределението на честотите следва една характерна тенденция – най-често се срещат стойностите, които се намират около средната, а онези, които са близки до минималната и максималната стойност се срещат най-рядко. На същата графика е представена и кривата на нормалното разпределение. Много ясно се вижда, че разпределението на емпиричните стойности (височината на правоъгълниците на хистограмата) в голяма степен наподобява това на нормалното разпределение. Преди да се спрем на статистическите показатели, които описват степента на приближение на емпиричното към теоретичното нормално разпределение ще разгледаме някои характерни негови особености:

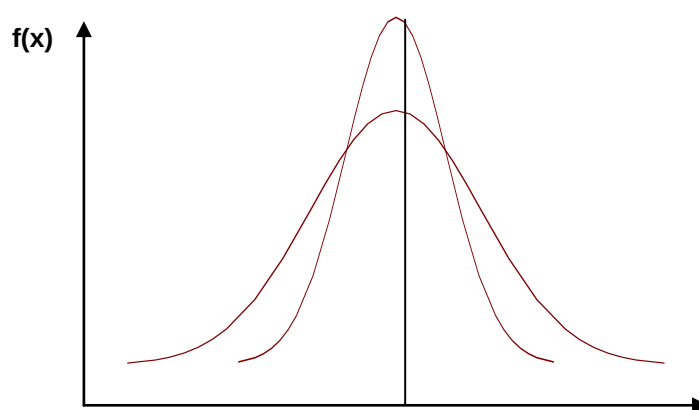
1. Кривата на разпределение на стойностите на променливата и съответните им абсолютни или относителни честоти е **камбановидна** (фигура 3.2).



Фиг. 3.2

2. Средната аритметична величина, модата и медианата съвпадат по стойност и определят центъра разпределението.

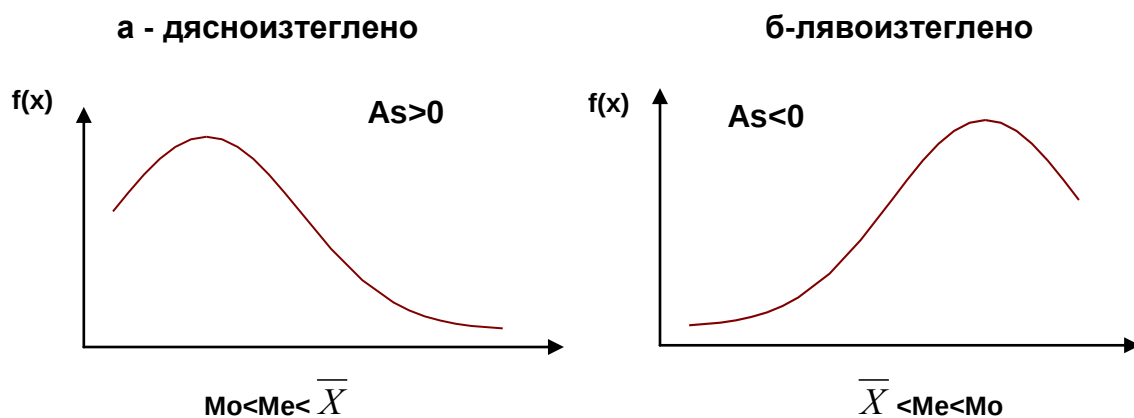
3. Стандартното отклонение (S) определя заостреността на кривата на разпределение. Колкото отклоненията от средното равнище са по-малки, толкова кривата е по-изострена и обратно. На фиг. 3.3 са представени две емпирични разпределения с еднакъв център, но с различно разсейване.



Фиг. 3.3

4. Кривата на нормалното разпределение е **симетрична** по отношение на перпендикуляра, спуснат от нейния връх към абсцисата, където се намира средната аритметична величина. На фигура 3.2 е представено симетрично

разпределение, а на фигура 3.4 две несиметрични разпределения..
Разпределение **а** е с дясно изтеглено, а **б** - с ляво изтеглено рамо.



Фиг. 3.4

Симетричността на разпределението се описва с показателя за **асиметрия (As)**, който при нормално разпределение е равен на нула (**As=0**). Когато коефициентът на асиметрия е с положителен знак, кривата на разпределение е с дясно изтеглено рамо (**а**), средната аритметична е по-голяма от модата и медианата. При разпределение с ляво изтеглено рамо (**б**) стойността на коефициента на асиметрия е с отрицателен знак, а средната аритметична е по-малка от модата и медианата.

За симетричността на емпиричното разпределение се съди по изчисления от извадката по формула 3.5 **коефициент на асиметрия (As)**.

$$As = \frac{m_3}{S^3} \quad (3.5)$$

Където:

m_3 – трети централен момент, който се изчислява по формула 3.6

S – стандартно отклонение

Третият централен момент се изчислява по формула:

$$m_3 = \frac{\sum (X_i - \bar{X})^3}{n} \quad (3.6)$$

Където:

X_i – всяка една стойност на променливата в извадката

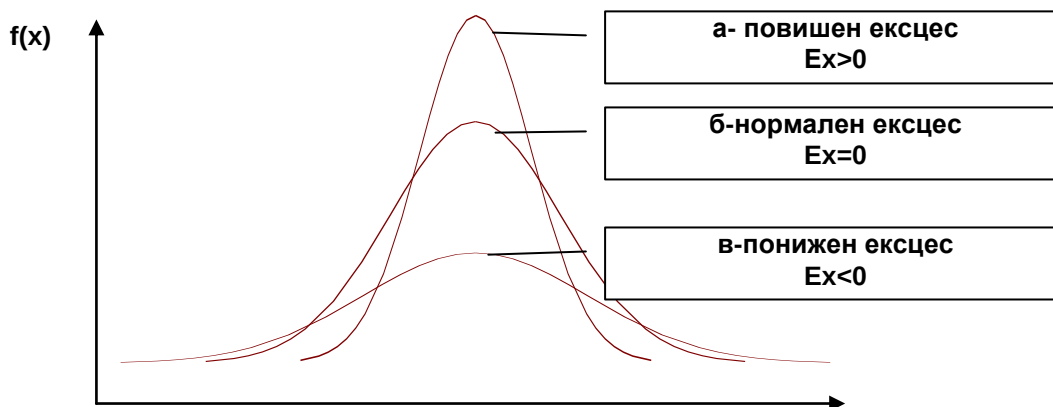
$\bar{X}$ – средна аритметична величина

n – обем на извадката

Абсолютната стойност на получения по данни от извадката коефициент се сравнява с табличната стойност на коефициента, подбрана от приложение 3 в от обема на извадката (n) и равнището на значимост (α). В случай че:

- $|As| < As_{\alpha}$ се приема, че разпределението на емпиричните данни е симетрично.
- $|As| \geq As_{\alpha}$ се счита, че разпределението е несиметрично.

5. Височината на върха на нормалното разпределение се описва с показателя ексцес (Ex), който при нормално разпределение е равен на нула. На фигура 3.5 са представени примерни разпределения с нормален (б), повишен - положителен (а) и понижен - отрицателен ексцес (в).



Фиг. 3.5

Информация за степента на съвпадение на височината на върха на емпиричното разпределение с приетия за еталон връх на нормалното разпределение ($Ex=0$) дава изчисленият по формула 3.7 **коефициент на ексцес (Ex)**.

$$Ex = \frac{m_4}{S^4} - 3 \quad (3.7)$$

Където:

m_4 – четвърти централен момент, който се изчислява по аналогичен на третия централен момент начин, като разликите между всяка една стойност на променливата и средната аритметична се повдигат на четвърта степен

S – стандартно отклонение

Абсолютната стойност на изчисления коефициент се сравнява с подбраната в зависимост от обема на извадката (n) и равнището на значимост (α) таблична стойност (приложение 4). В случай че:

- $|Ex| < Ex_{\alpha}$ се приема, че разпределението на емпиричните данни има нормален ексцес.
- $|Ex| \geq Ex_{\alpha}$ се счита, че разпределението има ненормален ексцес.

Изчислените емпирични стойности на коефициентите на асимерия и ексцес на представеното на фиг. 3.1 разпределение на резултатите в тест 30 м. гл. б. са съответно $As=0,331$ и $Ex=-0,286$. При обем на извадката $n=50$ и $\alpha=0,05$ критичните стойности са съответно $As_{0,05}=0,533$ и $Ex_{0,05}=0,848$. Тъй като емпиричните стойности са по-малки от критичните може да се приеме, че разпределението на променливата е нормално.

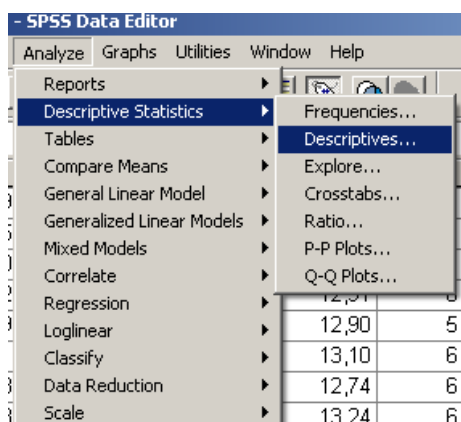
3.5. Вариационен анализ със SPSS

3.5.1. Обработка на целия масив

Задача за разучаване 3.1.

По данните от файла Ph_education.sav направете вариационен анализ на резултатите от първото тестиране.

1. Изберете менюто **Analyze> Descriptive statistics>Descriptives** (фиг. 3.6).



Фиг. 3.6

2. В диалоговия прозорец **Descriptives** (фиг.3.7) маркирайте променливите, натиснете бутона със стрелката и ги изпратете в полето Variable(s);



Фиг. 3.7

3. Натиснете бутона Options. В отворения се диалогов прозорец (фиг. 3.8) отбележете статистическите показатели, които да включи обработката.

	Средно равнище Mean - средна аритметична величина Sum - сума на стойностите
	Разсейване Std. deviation - стандартно отклонение Variance - дисперсия Range - размах Minimum - минимална стойност Maximum - максимална стойност S.E. mean - репрезентативна грешка
	Разпределение Kurtosis - ексцес Skewness - асиметрия

Фиг. 3.8

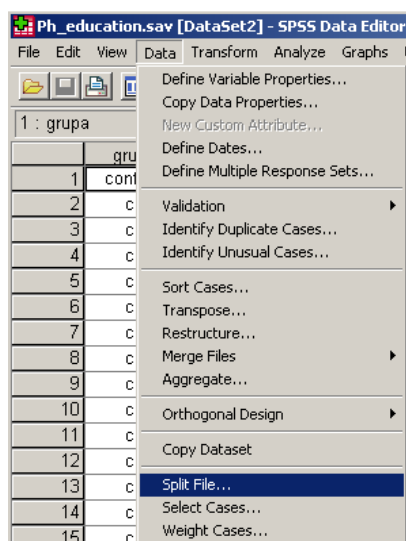
4. Натиснете бутона **Continue**, за да се върнете към основния диалогов прозорец и **OK**, за да стартирате обработката.

3.5.2. Обработка на подгрупи от базата данни

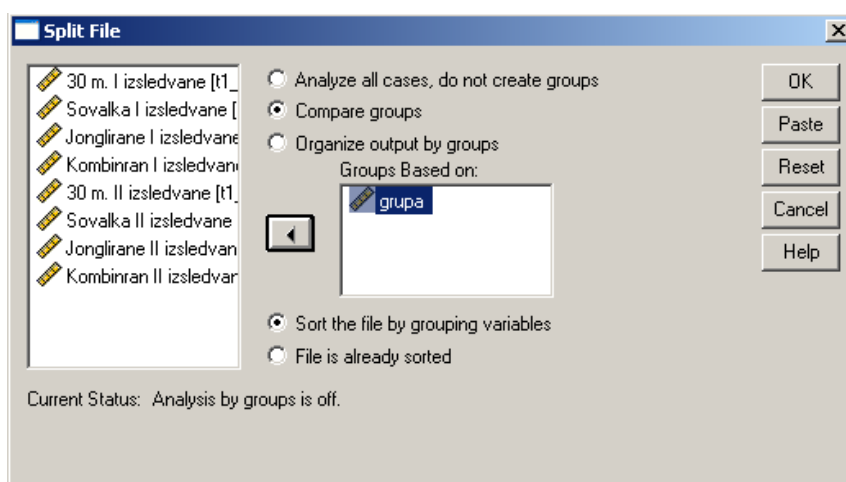
Задача за разучаване 3.2.

Разделете масива от данни на групи за самостоятелна обработка по променливата **grupa** и направете вариационен анализ за резултатите в тест «Совалка».

1. Изберете **Data>Split File** (фиг. 3.9). Появява се диалоговият прозорец, показан на фигура 3.10. От него личи, че по подразбиране програмата обработва всички случаи.



Фиг. 3.9



Фиг. 3.10

2. Задайте последователно:

- **Compare groups** (ако искате анализът за групите да се появи в една таблица) или **Organize output by groups** (ако искате анализът за всяка група да се появи в отделна таблица),
- Изпратете променливата, в която е записана принадлежността към групите, за които желаете да осъществите самостоятелна обработка в полето **Groups Based on**;
- Потвърдете с **ОК**.

3. Повторете вариационния анализ по алгоритъма, описан в 3.5.1. Като резултат обработката ще бъде осъществена за всяка от групите.

3.6. Представяне на резултати от вариационен анализ

Резултатът от обработката (задача за изучаване 3.1) се появява в разпечатката на SPSS по начина, представен на таблица 3.1:

Таблица 3.1

Descriptive Statistics										
	N	Range	Minimum	Maximum	Mean	Std.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
30 m. I izsledvane	50	1,56	5,45	7,01	6,2164	,35932	,331	,337	-,286	,662
Sovalka I izsledvane	50	2,50	13,93	16,43	14,7460	,56767	,818	,337	,066	,662
Jonglirane I izsledvane	50	2,00	2,00	4,00	2,8200	,62890	,150	,337	-,473	,662
Kombinran I izsledvane	50	4,16	10,45	14,61	13,1036	,91901	-,475	,337	,559	,662
Valid N (listwise)	50									

При публикуване в научни разработки таблицата се нуждае от редактиране, изписване на статистическите показатели с възприетите символи и допълнително изчисляване на коефициента на вариация ($V\%$). Окончателно оформената таблица изглежда по начина, представен в таблица 3.2.

Таблица 3.2.

№	Тест	Показател	n	$X_{\min}$	$X_{\max}$	R	$\bar{X}$	S	V	As	Ex
1	30 m		50	5,45	7,01	1,56	6,22	0,359	5,78	0,33	-0,29
2	Sovalka		50	13,93	16,43	2,5	14,75	0,567	3,85	0,82	0,07
3	Jonglirane		50	2	4	2	2,82	0,628	22,30	0,15	-0,47
4	Kombinran		50	10,45	14,61	4,16	13,10	0,919	7,01	-0,48	0,56

Забележка: Критичните стойности на коефициентите на асиметрия и ексцес при $n=50$ и $\alpha=0.05$ са равни на 0.673 и 1.324

Резултатите от вариационния анализ на отделните групи (задача за разучаване 3.2.) се появява по начина, представен в таблица 3.3, а редактираният и вид – в таблица 3.4.

Таблица 3.3

Descriptive Statistics											
		N	Range	Minimum	Maximum	Mean	Std.	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
experimental	Sovalka I izsledvane	25	2,50	13,93	16,43	14,7720	,63827	,990	,464	,309	,902
	Valid N (listwise)	25									
control	Sovalka I izsledvane	25	1,52	14,09	15,61	14,7200	,49912	,386	,464	-1,191	,902
	Valid N (listwise)	25									

Таблица 3.4

№	Показател	n	$X_{\min}$	$X_{\max}$	R	$\bar{X}$	S	V	As	Ex
1	Експериментална	25	13,93	16,43	2,5	14,77	0,63	4,32	0,99	0,31
3	Контролна	25	14,09	15,61	1,52	14,72	0,49	3,39	0,39	-1,19

Забележка: Критичните стойности на коефициентите на асиметрия и ексцес при $n=50$ и $\alpha=0.05$ са равни на 0.927 и 1.803



Решете задачи 1 и 2 от изпитния проект.

Тема 4. Сигмален метод за изработване на нормативи

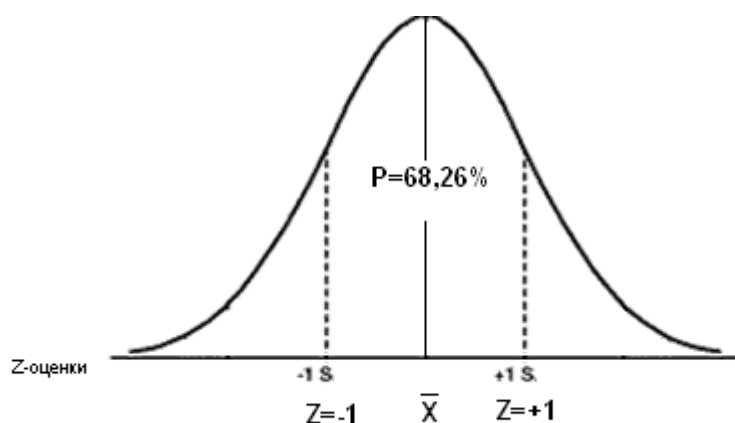
4.1. Основни понятия

Свойствата на нормалното разпределение имат голямо приложение в научно приложната дейност в спорта и физическото възпитание. Най-често на тях се основава разработването на нормативи за оценка на резултатите в спортно-педагогическите тестове. Норматив в спортната наука се нарича гранична стойност на резултата, които позволява изследваното лице да се причисли към дадена класификационна група.

Граничните стойности се изчисляват на основата на закона за нормалното разпределение, на които не се спрем накратко.

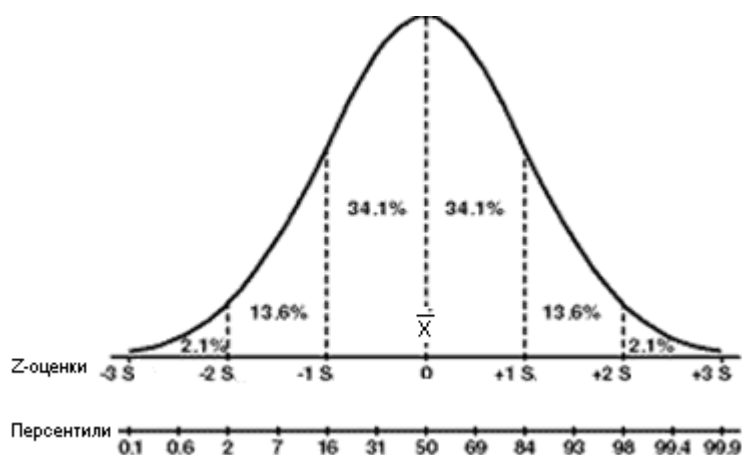
От теорията е известно, че вероятността ($P\%$), дадена стойност на променливата величина, да се намира в интервала от X_1 до X_2 може да бъде представена с площта, заключена между кривата на нормалното разпределение (фиг. 4.1), абсцисата и перпендикулярите, спуснати към нея в точки $X_1 = \bar{X} - Z.S$ и $X_2 = \bar{X} + Z.S$. Вероятността не зависи от средната аритметична и стандартното отклонение, а само от множителя Z , който се нарича стандартизирано (нормирано) отклонение и се изчислява по формула 4.1:

$$z = \frac{X_i - \bar{X}}{S} \quad (4.1)$$



Фиг. 4.1

Това важно свойство на нормалното разпределение намира израз в т.нар. “правило на трите сигми*” (фиг.4.2).



Фиг. 4.2

От фигурата се вижда, че в интервала $\bar{X} \pm 1.S$ попадат централните 68, 26% от случаите в генералната съвкупност, което означава, че вероятността за попадане на един случайно взет резултат в този интервал е 68,26%. Вероятността за попадане на даден резултат в интервала $\bar{X} \pm 2.S$ е 95,45% и т. н.

На практика това свойство на нормалното разпределение се използва за изработване на нормативи по т. нар. **сигмален метод**. В него се използват средната стойност ($\bar{X}$) и стандартното отклонение (S). Броят на степените, словесната оценка и процентът от случаите, които попадат към тях се определят от изследователя. В таблица 4.1 са дадени границите и словесните оценки на норматив по пет степенната скала на Шефко:

Таблица 4.1

Словесна оценка	Граници	Процент от случаите към всяка от оценките	Кумулативен % от случаите
Ниска	Под $-2.S$	2.27	2,27
Под средна	От $-2.S$ до $-1.S$	13.59	15,86
Средна	От $-1.S$ до $+1.S$	68.26	84,12
Над средна	От $+1.S$ до $+2.S$	13.59	97,71
Висока	Над $+2.S$	2.27	99.98

Недостатък на този оценителна скала е фактът, че твърде голям процент от случаите попадат към средната оценка, което прави норматива недостатъчно селективен. В практиката се ползва и следната седем степенна скала (таблица 4.2):

Таблица 4.2

Словесна оценка	Граници	Процент от случаите към всяка от оценките	Кумулативен % от случаите
Много ниска	Под $-2.S$	2.27	2,27
Ниска	От $-2.S$ до $-1.S$	13.59	15,86
Под средна	От $-0,5.S$ до $-1.S$	14,99	30,85
Средна	От $-0,5.S$ до $+0,5.S$	38,29	69,14
Над средна	От $+0,5.S$ до $+1.S$	14,99	84,12
Висока	От $+1.S$ до $+2.S$	13.59	97,71
Много висока	Над $+2.S$	2.27	99.98

За изчисляване на нормативи с помощта на SPSS биха могли да се ползват персентилите, защото при идеално нормално разпределение сигмалните и персентилните оценки съвпадат. За разработване на 5 степенен норматив е необходимо да се изчислят P_2 , P_{16} , P_{84} и P_{98} .¹ За разработване на норматив със 7 степенна скала е необходимо да се изчислят P_2 , P_{16} , P_{30} , P_{70} , P_{84} и P_{98} .

Законът на нормалното разпределение дава възможност да се оценява и всяка една стойност на променливата чрез т. нар. Z-оценка (виж формула 4.1).

¹ Алгоритъмът е онагледен в учебния филм.

Очевидно е, че ако оценяваната стойност е под средното равнище, нейната Z-оценка е отрицателна, ако съвпада със средната стойност $Z=0$, а ако е по-голяма от нея – Z е положително число. В програмните възможности на SPSS е заложено, успоредно с осъществяване на вариационен анализ да се изчислят и Z-оценките за всяка една стойност на X .

Оценяването на всяка една стойност със Z-оценки създава известни затруднения, защото част от оценките (тези по-ниски от средната стойност) са отрицателни. Поради това в практиката се ползват и т.нар. T-оценки. Тяхната стойност се изчислява по формула 4.2:

$$T = 50 + 10 \cdot \frac{X_i - \bar{X}}{S} = 50 + 10 \cdot Z \quad (4.2)$$

Чрез това преобразуване средната стойност се оценява с 50 точки. Оценките на резултатите варират от 0 до 100 т., но вероятността за получаване на резултат по-малък от 20 т. и по-голям от 80 т. е много малка. При нормално разпределение стойността им съвпада с персентилните оценки. Друго предимство, което имат T-оценките при изработване на нормативна база по резултати от повече тестове е, че те дават възможност за осредняване на оценките.

При разработването на нормативи трябва да се спазват следните важни условия:

- Те се съставят на базата на реално проучване на състоянието на оценяваните признаци, на основата на достатъчна по обем репрезентативна извадка.
- Съставените нормативи се отнасят за съвкупността, която е изследвана и не могат да бъдат използвани за друга съвкупност.
- Нормативите периодически трябва да се осъвременяват, защото физическото развитие и дееспособност на подрастващите от различните поколения се променят.

4.2. Задачи за упражнение

Задача за разучаване 4.1

По данните от файла Ph_education.sav изчислете необходимите стойности, за да разработите 5-степенен норматив за оценка на резултата в тест 30 м. гл. б. при I изследване. Ползвайте процентните разпределения, представени в таблица 4.1.

Изчисляването става с калкулатор, като се ползва средната стойност и стандартното отклонение на променливата. Със SPSS могат да се изчислят стандартизираните стойности (Z-оценките) на всеки резултат или да се изчислят персентилите, които съответстват на границите на норматива Analyze>Descriptive Statistics>Frequencies

Представяне на резултатите от обработката

Словесна оценка	Граници	Процент към всяка от оценките
Ниска	Под 5,50 сек.	2.27
Под средна	От 5,50 до 5,85 сек.	13.59
Средна	От 5,86 до 6,58 сек.	68.26
Над средна	От 6,59 до 6,94 сек.	13.59
Висока	Над 6,94 сек.	2.27

Забележка: Нормативът е направен на базата на средната стойност ($\bar{X} = 6,22$ сек.) и стандартното отклонение ($S = 0,36$ сек.)



Решете задача 3 от изпитния проект.

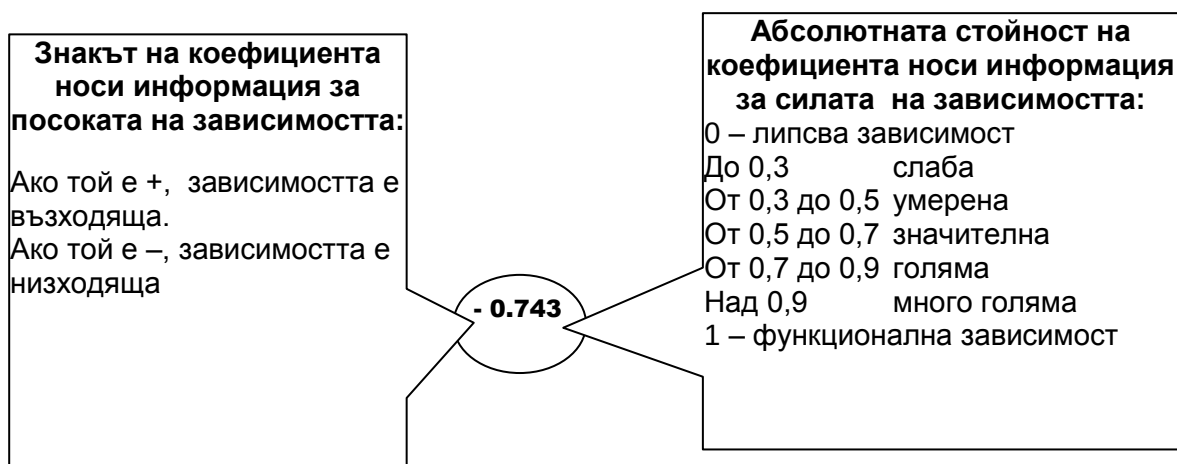
Тема 5. Корелационен анализ

5.1. Въведение

Корелационният анализ се прилага за:

- Установяване наличието на зависимост между променливите величини;
- Измерване на нейната сила (степен) и посока.

Статистическите показатели, които се ползват се наричат коефициенти на корелация. В зависимост от естеството и броя на променливите величини, те могат да бъдат различни, но стойността им винаги е в границите от -1 до +1. Интерпретират се по начина, показан на фиг. 5.1:



Фиг. 5.1

5.2. Видове коефициенти на корелация

Важен момен в прилагането на корелационния анализ е изборът на подходящ корелационен коефициент.

- В случай, че променливите величини са количествени, имат нормално разпределение и зависимостта между тях е линейна, се прилага коефициентът на **обикновена линейна корелация на Пирсън (r)**.
- В случай, че променливите са рангово скалирани или са количествени, но имат ненормално разпределение, се прилага коефициентът на **рангова корелация на Спирмън (r_s)**.

- В случай, че променливите са номинални и ординално скалирани се прилага коефициентът на **контингенция (C)**.

В практиката се прилагат и някои производни коефициенти:

- **Коефициентът на детерминация ($r^2.100$)** изразява силата на зависимостта в проценти и показва какъв процент от вариацията на едната променлива може да бъде обяснена с вариацията на другата променливата (това е т.нар. обяснена дисперсия).
- Влиянието на неизследваните фактори се описва с **коефициента на неопределеност** $k^2.100=100 - r^2.100$. Той изразява т. нар. необяснена дисперсия.

5.3. Статистическа значимост на коефициентите на корелация

Статистическата обработка се основава на данни от извадка, което предполага проверка на хипотезата доколко наблюдаваната в извадката зависимост се проявява и в изследваната генерална съвкупност. За целта се проверява статистическата значимост на коефициентите на корелация.

Абсолютната стойност на изчисления коефициент ($|r_{emp}|$) се сравнява с критичната стойност (r_{α}), която се отчита от приложение при степени на свобода $k=n-2$. Решението се взема по следния начин:

- Ако $|r_{emp}| < r_{\alpha}$ – приема се, че констатираната в извадката зависимост е случайна и не се наблюдава в генералната съвкупност;
- Ако $|r_{emp}| \geq r_{\alpha}$ – приема се, че констатирана в извадката зависимост е статистически значима и се проявява и в генералната съвкупност.

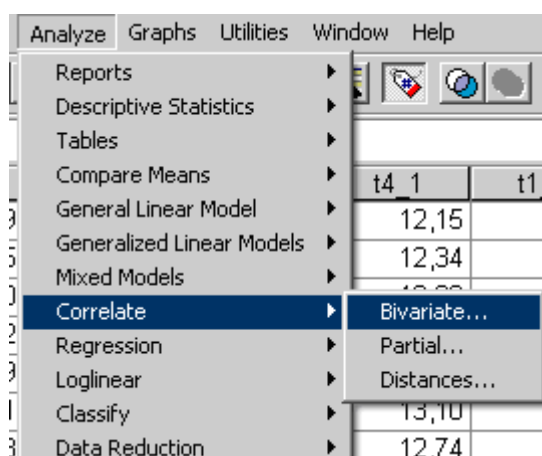
Критичните стойности на коефициента на Пирсън са представени в приложение 6, а за коефициента на Спирмън – в приложение 7.

5.4. Корелационен анализ със SPSS

Задача за разучаване 5.1.

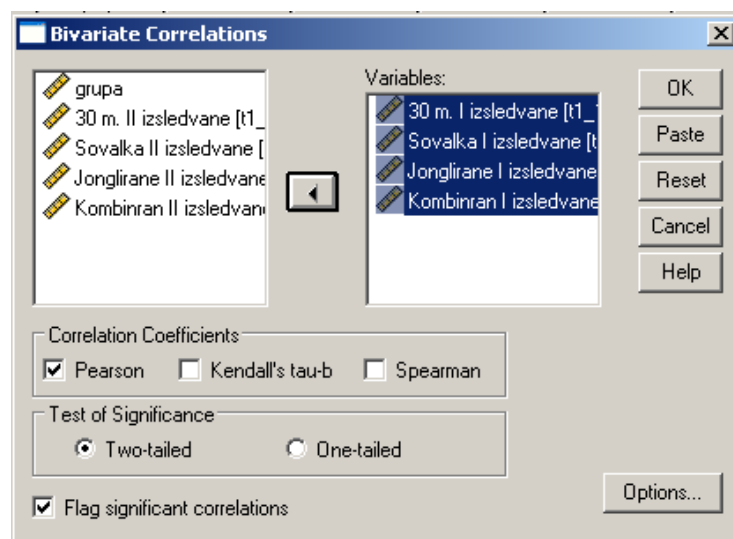
По данните от файла Ph_education.sav изчислете коефициентите на обикновена линейна корелация на Пирсън между резултатите в четирите теста при I изследване.

1. Изберете **Analyze>Correlate>Bivariate**, както е показано на фиг. 5.2.



Фиг. 5.2

2. Маркирайте променливите и ги изпратете в полето за обработка (фиг. 5.3).



Фиг. 5.3.

2. Подберете подходящия коефициент на корелация (от това меню може да се зададе изчисляване на коефициента на Пирсън и Спирман) и потвърдете с ОК.

5.5. Представяне на резултатите от корелационен анализ

Резултатите от корелационния анализ (задача за изучаване 5.1) се представят в т. нар. корелационни матрици (таблица 5.1).

Таблица 5.1.

Correlations					
		30 m. I izsledvane	Sovalka I izsledvane	Jonglirane I izsledvane	Kombinran I izsledvane
30 m. I izsledvane	Pearson Correlation	1	,066	,075	,139
	Sig. (2-tailed)		,648	,606	,337
	N	50	50	50	50
Sovalka I izsledvane	Pearson Correlation	,066	1	-,242	,303*
	Sig. (2-tailed)	,648		,090	,032
	N	50	50	50	50
Jonglirane I izsledvane	Pearson Correlation	,075	-,242	1	-,442**
	Sig. (2-tailed)	,606	,090		,001
	N	50	50	50	50
Kombinran I izsledvane	Pearson Correlation	,139	,303*	-,442**	1
	Sig. (2-tailed)	,337	,032	,001	
	N	50	50	50	50

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

На първия ред и в първата колона на таблицата са представени зададените за обработка променливи. В клетките на таблицата са изписани последователно:

- Коефициентът на корелация (Pearson Correlation);
- Съответстващото му равнище на значимост (Sig 2-tailed). Статистически значими са корелационните коефициенти, чието равнище на значимост е по-малко от 0,05. Същите корелационни коефициенти са означени със звездичка;
- Брой на наблюденията (N).

В клетките, които са по диагонала на таблицата се намират коефициентите на корелация между променливата сама със себе си, поради което стойността е равна на единица. Стойностите в клетките, които са под и над този диагонал са еднакви, поради което за окончателното оформяне на таблицата се преписва само половината от разпечатката на SPSS и корелационната матрица придобива следния вид:

Таблица 5.2

	30 m. I izsledvane	Sovalka I izsledvane	Jonglirane I izsledvane	Kombinran I izsledvane
30 m. I izsledvane	1,000			
Sovalka I izsledvane	0,066	1,000		
Jonglirane I izsledvane	0,075	-0,242	1,000	
Kombinran I izsledvane	0,139	0,303	-0,442	1,000

Забележка: Критичната стойност на коефициента на Пирсън при $n=50$ и $\alpha=0,05$ е равна на 0,27.

За открояване на статистически значимите корелационни коефициенти е уместно под корелационната матрица да се изпише критичната стойност на коефициента. Статистически значими в примера (таблица 5.2), са само коефициентите на корелация на „Комбиниран тест” със „Совалка” ($r=0,303$) и „Жонглиране с топка” ($r=-0,442$). Същите са обозначени със звездичка в разпечатката на SPSS (таблица 5.1).



Решете задача 4 от изпитния проект.

Тема 6. Статистическа проверка на хипотези

6.1. Основни понятия

Голяма част от научните изследвания се отнасят до сравняване на разпределението на две или повече променливи величини. Така например типични за изследванията в областта на физическото възпитание са проблемите за сравняване на нивото на двигателните качества на различни категории подрастващи, за съпоставяне на ефективността на различни методики на подготовка и др.

Особеното на тези сравнения е, че изводите, които се правят, трябва да се отнасят до целите съвкупности, а данните, с които разполага изследователят, обхващат само извадка от нея. Поради това първоначално се формулират предположения – хипотези, като впоследствие се прави проверка дали данните от извадката ги потвърждават или отхвърлят.

- **Нулевата или работна (H_0)** - тя твърди, че няма статистически достоверна разлика в сравняваните статистически показатели. Въпреки, че в извадките може да се наблюдава известна разлика, тя е случайна и не може да бъде обобщена за генералните съвкупности.

- **Алтернативна хипотеза (H_1)**, която твърди, че констатираната разлика в сравняваните статистически показатели в извадките е статистически достоверна и може да бъде обобщена за генералните съвкупности.

Решенията, които се вземат при проверката на статистически хипотези, имат вероятностен характер. Това се дължи на факта, че изследванията са репрезентативни, т.е. изводите за генералната съвкупност се правят въз основа на изследване на относително малка, но **репрезентативна** извадка. Приемането или отхвърлянето на нулевата хипотеза се прави с определено степен на сигурност, като в същото време се допуска възможността за грешка на статистическите изводи.

Степента на сигурност, с която се приема за вярна алтернативната хипотеза, се нарича **гаранционна вероятност (P)**.

Рискът да се допусне грешка, като се приеме за вярна алтернативната хипотеза се нарича **равнище на значимост (α)**.

В практиката се използват следните стойности за гаранционната вероятност (P) и равнище на значимост (α):

- P= 95%, на която съответства $\alpha = 0,05$ (5% възможност за грешка).
- P= 99%, на която съответства $\alpha = 0,01$ (1% възможност за грешка).
- P=99,9%, на която съответства $\alpha=0,001$ (0,1% възможност за грешка).

6.2. Алгоритъм за проверка на хипотези

Общата процедура за проверка на статистически хипотези предполага преминаването през няколко поредни стъпки:

I стъпка - Формулиране на нулевата и алтернативната хипотеза.

Както вече бе споменато, нулевата (H_0) хипотеза е предположение за нулева разлика, нулев ефект.. Ако емпиричните данни сочат известни различия - те се дължат на случайни фактори. Противоположно е твърдението на алтернативната хипотеза (H_1), според която наблюдаваните в емпиричните данни различия са резултат на закономерно действащи фактори.

II стъпка - Избор на подходящ критерий за проверка на хипотезата.

Хипотезите се проверяват със специфични критерии. За целта е необходимо да се познават условията, на които трябва да отговарят променливите величини, естеството на извадките и някои други особености, които са разгледани в следващия раздел.

III стъпка - Изчисляване на емпиричната стойност на критерия (Критерий_{emp}).

За всеки един критерий има разработена процедура за изчисляването му по данни от извадката.

IV стъпка – Определяне на табличната стойност на критерия (Критерий α).

Тя се взема от приложение, в зависимост от степените на свобода (k) и равнището на значимост (α). Към статистическата литература има приложени таблици с критични стойности на критериите, към които е описан и начинът на изчисляване на степените на свобода. В настоящото ръководство тези таблици се намират, приложени към съответната тема.

V стъпка – Вземане на решение.

За целта табличната (теоретичната) стойност на критерия се сравнява с емпиричната (изчислената по данни от извадката). Именно това сравнение дава основание за формулиране на извод коя от хипотезите (нулевата или алтернативната) е вярна. Възможностите са две:

- **Емпиричният критерий да е по-малък от теоретичния.** В такива случаи за правдоподобна се приема нулевата хипотеза (H_0).
- **Емпиричният критерий да е по-голям или равен на теоретичния.** В такива случаи нулевата хипотеза (H_0) се отхвърля и за вярна се приема алтернативната хипотеза (H_1).

Вземането на решение може да стане и въз основа на равнището на значимост (α), което съответства на емпиричната стойност на критерия, чиято стойност се изписва в разпечатките на SPSS:

- **Равнището на значимост (α) да е по-голямо от възприетото ($\alpha > 0,05$ (0,01 или 0,001)).** В такъв случай няма основание за отхвърляне на нулевата хипотеза (H_0).
- **Равнището на значимост (α) да е по-малко или равно на възприетото ($\alpha \leq 0,05$ (0,01 или 0,001)).** Това означава, че нулевата хипотеза (H_0) се отхвърля в полза на алтернативната.

Както вече бе споменато, аналогична информация носи гаранционната вероятност $P\%$, която при известно равнище на значимост се изчислява по формулата:

$$P\% = (1 - \alpha) * 100$$

Възможностите придобиват следния вид:

- **Гаранционната вероятност да е по-малка от възприетата (95% или 99%).** В такъв случай за вярна се приема нулевата хипотеза (H_0).
- **Гаранционната вероятност да е равна или по-голяма от 95% (99%).** В такъв случай се отхвърля нулевата и за вярна се приема алтернативната хипотеза.

Обобщение на начините за вземане на решение е представено в таблица 6.1.

Таблица 6.1

Хипотеза	Решение, основано на стойността на критериите за проверка на хипотези	Решение, основано на достигнатото равнище на значимост	Решение, основано на гаранционната вероятност
Приемане на H_0	Критерий _{emp} < Критерий α	$\alpha > 0,05$ (0,001; 0,001)	$P\% < 95\%$ (99%; 99,9%)
Приемане на H_1	Критерий _{emp} $\geq$ Критерий α	$\alpha \leq 0,05$ (0,001; 0,001)	$P\% \geq 95\%$ (99%; 99,9%)

Въпреки голямото разнообразие от статистически критерии за проверка на хипотези, тази обща принципна схема се спазва винаги. Особено важен момент в процедурата е изборът на критерия. Непознаването на вида на променливите, вида на извадките, както и случаите, в които се прилага съответният критерий, може да доведе до некоректно ползване на статистическите методи, а оттам до утвърждаване на неверни научни изводи.

6.3. Критерии за проверка на хипотези

Изборът на критерий за проверка на дадена хипотезата зависи от:

1. Използвания статистически метод и показатели, които се сравняват (за сравняване на средни аритметични, дисперсии, относителни дялове и др. се ползват различни критерии).

2. Естеството на променливите величини, които подлежат на обработка – на първо място е важно дали променливите са количествени или категоријни

(качествени). На второ място изборът на критерий трябва да се съобрази и с вида на разпределение на променливите (нормално или не). От тази гледна точка са разработени две големи групи от статистически критерии:

- **Параметрични** – прилагат се при количествени (интервално и пропорционално скалирани) признаци, когато предварително е установено, че разпределението на променливите е нормално.

- **Непараметрични** – прилагат се при качествени (номинално и рангово) скалирани признаци, а така също, когато разпределението на количествен признак не е известно или не е нормално.

3. Вида на извадките, които подлежат на сравнение. От гледна точка на начина, по който са съставени, те биват:

- **Независими** - когато изборът на единиците в едната извадка не е предопределен от избора на единиците в другата извадка (с други думи изследваните лица от двете извадки са различни и няма никаква връзка между тях). Експерименти за сравняване на експериментална и контролна група са типичен пример за независими извадки.

- **Зависими** са извадките, когато единиците от едната извадка предопределят тези във втората. Типичен случай са изследванията за сравняване на променливи преди и след някакво въздействие. Очевидно е, че експерименти от такова естество са коректни само в случай, че изследваните лица (единици) от първото и второто изследване са едни и същи.

4. Броя на извадките, които се сравняват.

В таблица 6.2 е направена систематизация на основите критерии, които се ползват при научни изследвания във физическото възпитание.

Таблица 6.2

Променливи величини	Брой на извадките	Вид на извадките	Критерии за проверка на хипотези
Количествени, които имат нормално разпределение	Две	Независими	t-критерий на Стюдънт за независими извадки
		Зависими	t-критерий на Стюдънт за зависими извадки
	Три и повече	Независими	F-критерий на Фишер (дисперсионен анализ)
		Зависими	Дисперсионен анализ за повтарящи се наблюдения
Алтернативни	Две	Независими	U-критерий на Стюдънт
		Зависими	χ^2 -критерий на Мак Немар
Номинални	Две и повече	-	χ^2 -критерий на Пирсън
Рангови и количествени, които имат ненормално разпределение	Две	Независими	U-критерий на Ман Уитни
		Зависими	T-критерий на Уилкоксън
	Три и повече	Независими	Критерий на Кръскал Уолис
		Зависими	χ^2 -критерий на Фридман

t-критерият на Стюдънт за независими извадки се прилага за проверка на нулевата хипотеза за сравняване на средни аритметични величини, когато:

- Променливите са количествени и имат нормално разпределение;
- Извадките са две и са независими.

Емпиричната стойност на критерия се изчислява (t_{emp}) по формула 6.1, а табличната му стойност се отчита от приложение 8:

$$t_{emp} = \frac{|\bar{X}_2 - \bar{X}_1|}{\sqrt{\frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 n_2}}} \quad (6.1)$$

където:

$\bar{X}_1$ и $\bar{X}_2$ - ср.

аритметични в

извадките

S_1^2 и S_2^2 - дисперсии в

извадките

n_1 и n_2 - обем на

извадките

t-критерият на Стюдънт за зависими извадки се използва за сравняване на средни аритметични величини, когато:

- Признаците са метрично скалирани и имат нормално разпределение.
- Извадките са две и са зависими.

Емпиричната му стойност е изчислява по формула 6.2, а табличната се отчита от приложение 8.

$$t_{emp} = \frac{|\bar{d}|}{\sqrt{\frac{\sum d^2 - n(\bar{d})^2}{n^2 - n}}}$$

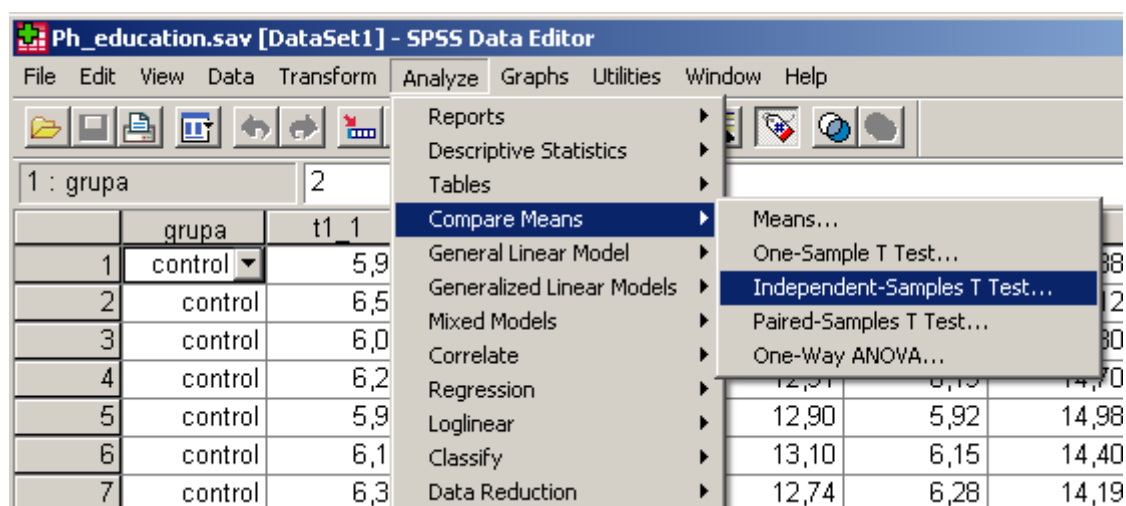
(6.2) където:
d – разлики между I и II
изследване за всяко изсл. лице
 $\bar{d}$ - средна аритметична на тези
разлики
n – обем на извадката

6.4. Изчисляване на t-критерия на Стюдънт за независими извадки

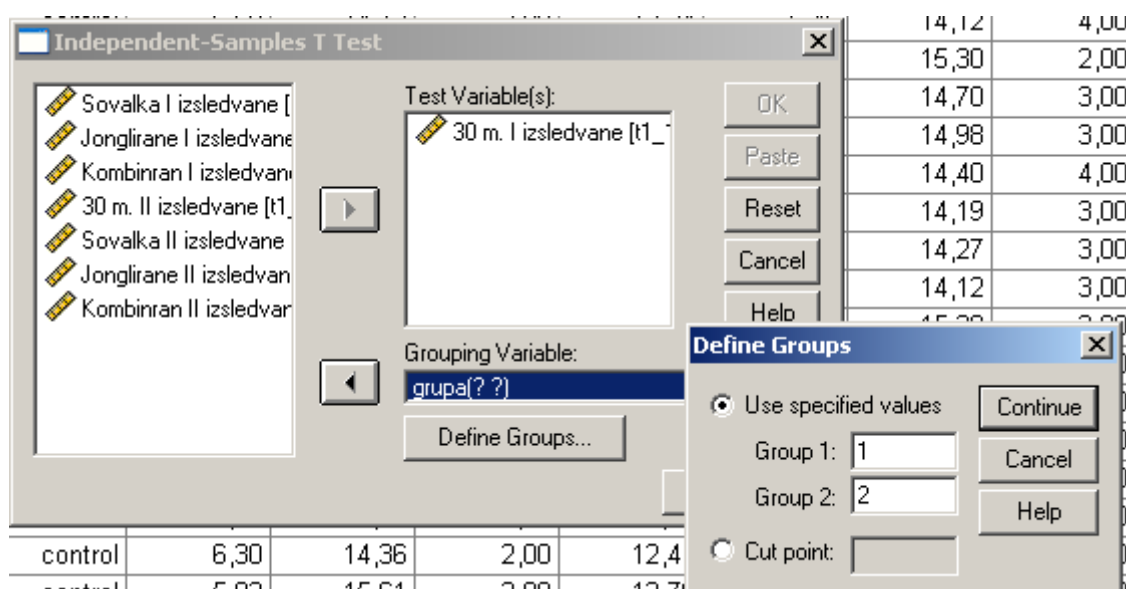
Задача за изучаване 6.1.

По данните от файла **Ph_education.sav.** изчислете t-критерия на Стюдънт за независими извадки и установете съществуват ли статистически значими разлики между резултатите на експериментална и контролна група в тест 30 м. гладко бягане (t1_1) при I-во изследване.

1. Изберете **Analyze>Compare Means>Independent Samples T Test** (фиг. 6.1).



Фиг. 6.1



Фиг. 6.2

2 Изпратете променливата (фиг. 6.2), в която са записани резултатите на двете групи в полето Test Variable(s).

3. Изпратете в полето Grouping Variable променливата, в която е записана принадлежността към сравняваните категории (групи). Натиснете бутона Define Groups и в отворилия се прозорец задайте кодовете на сравняваните групи.

4. Потвърдете с Continue>OK .

Резултатът от обработката е в 2 части:

В първата се изписва информация за разпределението на стойностите в двете извадки (таблица 6.4). Във втората част е изписана информация за

статистическата значимост на разликите между средните стойности в двете извадки (таблица 6.5). Използвани са следните означения:

- N- брой на наблюденията във всяка от извадките
- Mean - средна аритметична
- Std. Deviation – стандартно отклонение
- t – емпирична стойност на t-критерия на Стюдънт за независими извадки
- df – степени на свобода $k=n_1+n_2-2$
- Sig. (2-tailed) – равнище на значимост
- Mean differences – разликата между емпиричните средни

Таблица 6.4.

Group Statistics

grupa	N	Mean	Std. Deviation	Std. Error Mean
30 m. I izsledvane experimental	25	6,1704	,37225	,07445
control	25	6,2624	,34730	,06946

Таблица 6.5

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
30 m. I izsledvane	Equal variances assumed	,288	,594	-,904	48	,371	-,09200	,10182	-,29673	,11273
	Equal variances not assumed			-,904	47,771	,371	-,09200	,10182	-,29675	,11275

Резултатите от проверката на хипотези за сравняване на средни стойности на две групи се представят по начина, представен на таблица 6.6. За удобство под таблицата може да се изпише и критичната стойност на критерия.

Таблица 6.6

Показател	Експериментална			Контролна			d	t	P(t)
	n_1	$\bar{X}_1$	S_1	n_2	$\bar{X}_2$	S_2			
30 м.	25	6,17	0,37	25	6,26	0,35	0,09	0,90	62,93

Забележка: Критичната стойност на t-критерия на Стюдънт при $k=n_1+n_2-2=48$. и $\alpha=0.05$ е равна на 2,01 (виж приложение 8).

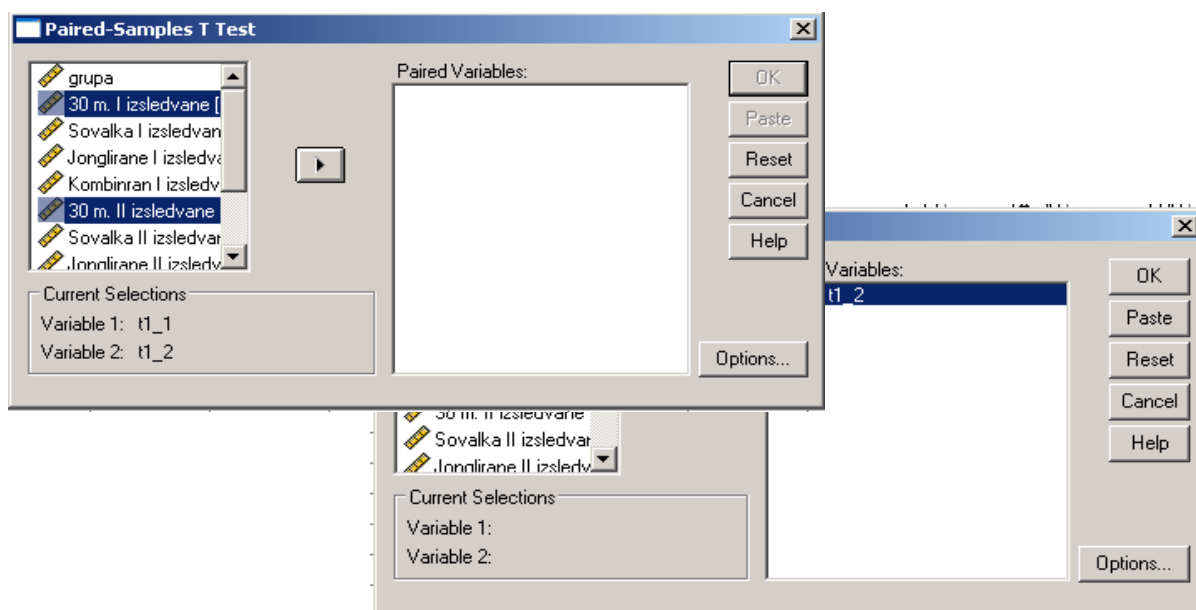
6.5. Изчисляване на t-критерия на Стюдънт за зависими извадки

6.5.1. За целия масив от данни

Задача за разучаване 6.2.

По данните от файла **Ph_education.sav**, изчислете t-критерия на Стюдънт за зависими извадки и установете статистическата достоверност на прираста на резултатите в тест 30 м. гл. б. (**t1_1** и **t1_2**).

1. Изберете Analyze>Compare Means>Paired Samples T Test (фиг. 6.1).



Фиг. 6.3

2. В отворения се диалогов прозорец (фиг. 6.3) изберете променливата, в която се намират резултатите от първото изследване. Името и се изписва на реда Variable 1 в полето Current Selections. Изберете и променливата с резултатите от второто изследване. Изпратете двойката променливи в прозореца Paired Variables и потвърдете с ОК.

Разпечатката съдържа 3 части:

В първата е представена информация за броя на наблюденията, средните стойности и вариативността на изследваните показатели в двете извадки;

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	30 m. I izsledvane	6,2164	50	,35932	,05082
	30 m. II izsledvane	5,8974	50	,45688	,06461

Във втората са представени коефициентите на корелация между I и II изследване, с оглед да се провери дали наистина извадките са зависими (корелирани).

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	30 m. I izsledvane & 30 m. II izsledvane	50	,764	,000

В третата е характеризирана големината (Mean) и статистическата значимост (t и α) на разликите между средните стойности на двете извадки.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	30 m. I izsledvane - 30 m. II izsledvane	,31900	,29504	,04172	,23515	,40285	7,645	49	,000

Резултатите от сравнението се представят в научни разработки по начина, представен в таблица 6.7.

Таблица 6.7

Показател		n	I изследване		II изследване		d	t	P(t)
			$\bar{X}_1$	S_1	$\bar{X}_2$	S_2			
1	30 m. I izsledvane 30 m. II izsledvane	50	6,22	0,36	5,90	0,46	-0,32	7,65	100,00

Забележка: Критичната стойност на *t*-критерия на Стюдънт при $k=n-1=49$ и $\alpha=0.05$ е равна на 2,01 (виж приложение 8).



Решете задача 5 от изпитния проект.

6.5.2. За отделни подгрупи

Задача за изучаване 6.3

По данните от файла **Ph_education.sav**, изчислете *t*-критерия на Стюдънт за зависими извадки и установете статистическата достоверност на прираста на резултатите на експериментална и контролна група в тест „Совалка”. (**t2_1** и **t2_2**).

Когато се обработват данни от педагогически експеримент с две групи (опитна и контролна) е необходимо да се установи статистическата значимост на прираста на резултатите на двете групи отделно. По тази причина файла се разделя на групи за самостоятелна обработка по начина, описан в 3.5.2. След това се прилага алгоритъма, представен в настоящия раздел.

Краткото описание на процедурите е, както следва:

1. Data>Split File и задайте разделяне на масива за самостоятелна обработка по променливата group.
2. Analyze>Compare Means>Pared Samples T Test

Резултатът от обработката след редактиране на разпечатката на SPSS е представен в таблица 6.8.

Таблица 6.8

Тест “Совалка”	I изследване		II изследване		d	t	P(t)
	$\bar{X}_1$	S_1	$\bar{X}_2$	S_2			
Експериментална	14,77	0,64	13,36	0,63	-1,41	14,51	100,00
Контролна	14,72	0,50	14,61	0,42	-0,11	3,52	99,80



Решете задача 6 от изпитния проект.

Тема 7. Обработка на данни от педагогически експеримент

7.1. Описание на експерименталния дизайн

Голяма част от научните изследвания в областта на физическото възпитание са свързани с експериментирание на нови средства и методи за въздействие, които да оказват благоприятен ефект върху физическото развитие и дееспособност на подрастващите. Най-често използваният експериментален дизайн включва формирането на две групи – експериментална (с която се прилага новото въздействие) и контролна (която прилага традиционни методи на въздействие). Състоянието на признаците, върху които се въздейства се измерва най-малко два пъти – в началото на прилагане на методиките (I изследване) и след определен период от време, който изследователят счита че е необходим, за да се доловят настъпилите промени (II изследване).

Обработката на данните би следвало да реши следните познавателни задачи:

1. **Да се установи ефектът от прилаганите въздействия във всяка от групите (схема 7.1).** За ефекта на новото въздействие се съди от прираста на резултатите (в абсолютни или относителни единици) в експерименталната група (**сравнение I**) и неговата статистическа значимост. По аналогичен начин може да се проследят промените, настъпили в контролната група (**сравнение II**).

Показател Група	I изследване		II изследване		$d = \overline{X}_2 - \overline{X}_1$	$d\% = \frac{\overline{X}_2 - \overline{X}_1}{\overline{X}_1} \cdot 100$	Критерий за проверка	P(%)
	$\overline{X}_1$	S ₁	$\overline{X}_2$	S ₂				
Експериментална	Сравнение I							
Контролна	Сравнение II							
$d = \overline{X}_e - \overline{X}_k$								
Критерий за проверка								
P(%)								

Схема 7.1

2. Да се сравни ефективността на въздействията, прилагани от експерименталната и контролната група. За целта се осъществяват следните сравнения (схема 7.2):

- **Сравнение III** - Установяване на статистическата значимост на разликите между ЕГ и КГ при I изследване. Това сравнение се прави с оглед да се установи доколко е съпоставимо по-нататъшното развитие на резултатите. Известно че, че не може да се сравнява ефективността на методики за развитие на дадено качество при качествено различно състояние на началните резултати (при висококвалифицирани и подрастващи спортисти, например). Ако групите са на едно и също ниво,, може да продължи по-нататъшното сравнение на ефективността на двете методики.
- **Сравнение IV** - Установяване на статистическата значимост на разликите между двете групи при II изследване и на тази база формиране на изводи за различия в ефекта на двете въздействия;

Показател Група	I изследване		II изследване		$d = \bar{X}_2 - \bar{X}_1$	$d\% = \frac{\bar{X}_2 - \bar{X}_1}{\bar{X}_1} \cdot 100$	Критерий за проверка	P(%)
	$\bar{X}_1$	S1	$\bar{X}_2$	S2				
Експериментална		Сравнение III		Сравнение IV	Сравнение V			
Контролна								
$d = \bar{X}_e - \bar{X}_k$								
Критерий за проверка								
P(%)								

Схема 7.2

- **Сравнение V** - Сравняване на прираста (d_e и d_k) на резултатите на двете групи и установяване на статистическата значимост на разликите между тях. За целта в SPSS първо се изчисляват нови променливи за прираста на резултатите във всеки от тестовете (виж задача за разучаване 2.2). След това с подходящ статистически критерий се проверява достоверността на разликите между тях.

7.2. Статистически показатели, прилагани при обработка на данните

Развитието на показателите се описва с показателите абсолютен (d) и относителен прираст (d%):

- **Абсолютен прираст** - $d = \bar{X}_2 - \bar{X}_1$.

Той се изразява в мерните единици на променливата величина и показва с колко и в каква посока се е променила средната стойност на съответната група. Не е подходящ за сравнения на резултати в тестове, изразени в различни мерни единици или които имат различна средна стойност.

$$d\% = \frac{\bar{X}_2 - \bar{X}_1}{\bar{X}_1} \cdot 100$$

- **Относителен прираст** –

Той се изразява в проценти, което дава възможност да се прави сравнение на развитието на показатели, изразени в различни мерни единици.

Статистическата значимост на прираста (сравнение I и II) се проверява с подходящ критерий за сравняване на средни стойности за **зависими извадки**. В случай, че признаците са количествени и имат нормално разпределение се ползва t-критерият на Стюдънт за зависими извадки, в противен случай – с T-критерия на Уилкоксън. И в двата случая, по равнището на достигнатата гаранционна вероятност ($P\%$) се взема решение за достоверността на прираста. Ако $P(\%) \geq 95\%$ се приема, че той е статистически значим и констатираните в извадките промени се дължат на приложените въздействия.

Статистическата значимост на разликите между експериментална и контролна група (сравнения III и IV) се проверява с подходящ критерий за сравняване на **две независими извадки**. В случай, че признаците са количествени и имат нормално разпределение се ползва t-критерият на Стюдънт за независими извадки, в противен случай – с U-критерия на Ман Уитни. Вземането на решение става въз основа на достигнатата гаранционна вероятност ($P\%$).

Най-съществен критерий за различията в ефекта на експерименталното и контролното въздействие е сравняването на прирастите на двете групи (сравнение V), което също се осъществява с подходящ критерий за независими извадки.

В таблица 7.1 са представени резултати от обработката на данните за тест 30 м. гл.б. от файла Ph_education.sav. Средната стойност на експериментална група при I изследване е 6,17 сек. В хода на експеримента е реализиран прираст от -0,59 сек. (9,54%). Той е подкрепен с близка до 100% гаранционна вероятност, което недвусмислено показва положителния ефект, който има новата методика върху скоростните възможности на учениците.

От статистическа гледна точка реализираният от контролната група прираст от -0,05 сек. (0,79%) също е достоверен ($P/t \approx 100\%$), но от практическа гледна точка той е твърде малък, за да считаме, че се наблюдава осезаемо развитие на скоростните възможности.

Хипотезата за по-голяма ефективност на тренировъчните въздействия, прилагани от ЕГ се потвърждава от сравнението на прирастите на двете групи ($d_e = -0,59$ сек. и $d_k = -0,05$ сек.). Разликата от 0,54 сек. е статистически значима, защото е подкрепена с гаранционна вероятност близка до 100%

Резултати от педагогическия експеримент в тест 30 м. гл. б.

Таблица 7.1

Група \ Показател	I изследване		II изследване		d	d %	t	P(t)
	$\bar{X}_I$	S_I	$\bar{X}_2$	S_2				
Експериментална	6,17	0,37	5,58	0,33	-0,59	-9,54	19,46	100,00
Контролна	6,26	0,35	6,21	0,33	-0,05	-0,79	4,10	100,00
Разлика	-0,092		-0,631		-0,54			
t	0,90		6.74		16,54			
P/t/	62.90		100		100,00			



Решете задача 7 от изпитния проект.

Литература

1. Аркадиев, Д. (2003) Статистика – Учебник за студенти от икономически специалности, Стара Загора
2. Брогли, Я., Л. Петкова (1983) “Статистически методи в спорта”, МиФ, С
3. Гигова, В. (1999) “Статистическа обработка и анализ на данни”, учебно помагало за студентите от магистърска степен на НСА, С.
4. Дамянова, Р., В. Гигова (2000) “Статистически методи в спорта”, ръководство за студентите от бакалавърска степен на НСА, С.
5. Зациорски, В. М. (1982) Основи на спортната метрология, МиФ, София
6. Калоянов, Т. (2004) Статистика, „Тракия”, С.
7. <http://1.nsa-virtualeducation.com/index.htm>
8. <http://bg.wikipedia.org/wiki/Статистика>
9. <http://www.fmi.uni-sofia.bg/fmi/statist/lectures/stat1/stat.htm>
10. <http://www.psychstat.smsu.edu/>
11. <http://www.statsoft.ru/home/>
12. <http://www.sportsci.org/>
13. <http://cmamucmuka.hit.bg/>
14. <http://wiki.stat.ucla.edu/socr/index.php/EBook>
15. Vincent, W. (1994) Statistics in Kinesiology, Human Kinetics, Northridge

Изпитен проект

“Статистически методи във физическото възпитание”

Магистърска програма

Физическо възпитание

Изготвил проекта:

Фак. №

Форма на обучение:

Проверил:

Дата:

Задача 1.



Направете вариационен анализ на резултатите от второто тестиране

	Показател	n	Xmin	Xmax	R	$\bar{X}$	S	V	As	Ex
1										
2										
3										
4										

Забележка: Критичните стойности на коефициентите на асиметрия и ексцес при $n=.....$ и $\alpha=....$ са равни на $As=....$ и $Ex=.....$



Отговорете на въпросите:

1. На колко е равно най-доброто постижение в тест 30 м. гл. б.?

.....
2. На колко е равна е средната стойност на постиженията в тест "Совалка"

.....
3. В какъв диапазон варира резултатът в жонглирането с топка?

.....
4. В какви мерни единици е изразено стандартното отклонение на "Комбиниран тест"?

.....
5. Може ли да се счита, че резултатите в изследваните 4 теста са еднородни и защо?

.....
6. Може ли да се счита, че разпределението на тези показатели е нормално и защо?

.....

Въпросите за магистрите могат да бъдат и по-трудни!

Задача 2.



Направете вариационен анализ на резултатите на експериментална и контролна група в тест „Жонглиране с топка” при I изследване.

	Жонглиране с топка	n	Xmin	Xmax	R	$\bar{X}$	S	V	As	Ex
1	Експериментална									
3	Контролна									



Отговорете на въпросите:

1. В какъв диапазон варират постиженията на експерименталната група?

.....

2. На колко е равна е средната стойност на постиженията в контролната група?

.....

3. Голяма ли е вариацията на стойностите?

.....

4. Нормално ли е разпределението на постиженията и защо?

.....

5. В рамките на 5-6 изречения анализирайте състоянието на резултатите в тест „Жонглиране с топка” на една от групите (по избор).

.....

.....

.....

Задача 3.



Изработете норматив за оценка на резултата в „Совалка” при I изследване, като използвате сигмалния метод и по-специално 5 степенната скала на Шефко.

$\bar{X}$ = ..., S=

Словесна оценка	Граници	P%
Ниска		
Под средна		
Средна		
Над средна		
Висока		



Отговорете на въпросите:

1. Какви са условията за прилагане на сигмалния метод за изработване на нормативи?
.....
2. Колко процента от изследваните лица биха получили средна оценка?
.....
3. Изберете едно от изследваните лица и оценете резултата му по разработения от Вас норматив.
.....
4. Изчислете и анализирайте Z и T-оценката на резултата на същото изследвано лице.
.....

Задача 4



Изследвайте зависимостите между тестовете при II изследване

	30 м. гл.б.	Совалка	Жонглиране	Комбиниран
30 м. гл.б.	1			
Совалка		1		
Жонглиране			1	
Комбиниран				1

Забележка: Критичната стойност на коефициента на обикновена линейна корелация при $n=...$ и $\alpha=0.05$ е равна на



Отговорете на въпросите:

- В каква скала са измерени променливите величини и кой коефициент на корелация е коректно да се приложи?
.....
- Колко силна е зависимостта между резултатите в тестове 30 м. гл. б. и „Совалка”?.....
- Анализирайте корелационните зависимости на резултатите от комбинирания тест с останалите спортно педагогически тестове. Кое от изследваните качества корелира в най-голяма степен с комбинирания тест?
.....

Задача 5.



Сравнете резултатите в тест “Совалка” на експериментална и контролна група при I-во и II-ро изследване.

Показател		I изследване			II изследване		
		n_1	$\bar{X}_I$	S_1	n_2	$\bar{X}_{II}$	S_2
1	Експериментална група						
2	Контролна група						
	Разлика						
	t						
	P(t)						

Забележка: Критичната стойност на t-критерия на Стюдънт при $k=.....$ и $\alpha=0.05$ е равна на



Отговорете на въпросите:

- Кой критерий за проверка на хипотези е коректно да се приложи за решаване на поставената изследователска задача и защо?
.....
- На колко са равни средните стойности на двете групи при I-во изследване? Достоверна ли е разликата между тях и защо?
.....
.....
- Може ли да се счита, че средните стойности на експерименталната и контролната група при II изследване се различават и защо?
.....
.....
.....

Задача 6.



Осъществете необходимите изчислителни процедури, за да установите настъпили ли са статистически значими промени в резултатите в тестове „Жонглиране с топка” и „Комбинирана тест”.

№	Показател	n	I изследване		II изследване		d	t	P(t)
			$\bar{X}_1$	S_1	$\bar{X}_2$	S_2			
1	Жонглиране с топка								
2	Комбиниран тест								

Забележка: Критичната стойност на t-критерия на Стюдънт за извадки при $k=.....$ и $\alpha=0.05$ е равна на



Отговорете на въпросите:

- Опишете промените, настъпили в резултата в тест “Жонглиране с топка?”
.....
- С кой критерий е коректно да се провери хипотезата за достоверност на разликите между I-во и II-ро изследване и защо?
.....
- На колко е равна емпиричната стойност на t-критерия за тест “Жонглиране”?
.....
- Достоверен ли е прирастът на резултатите в тест “Жонглиране и защо?”
.....
- Интерпретирайте резултатите от комбинирания тест:
.....
.....
.....

Задача 7.



Направете необходимите изчисления, за да попълните таблицата с резултатите от педагогическия експеримент в тест „Совалка” и отговорете на поставените въпроси.

Тест “Совалка”	I изследване		II изследване		d	d %	t	P(t)
	$\bar{X}_1$	S_1	$\bar{X}_2$	S_2				
Експериментална								
Контролна								
Разлика								
t								
P(t)								



Отговорете на въпросите:

1. Направете анализ на резултатите от педагогическия експеримент в тест „Совалка”.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Приложения

Приложение 1. Данни за практическите упражнения

Проведен е педагогически експеримент за установяване ефективността на нова методика за физическо възпитание, включваща завишен обем на упражнения от футбола. Формирани са две групи – експериментална (с която се прилага новата методика) и контролна, която работи по традиционен начин. В началото и края на експеримента са регистрирани резултатите в 4 теста. Данните от изследването са записани във файла Ph_education.sav), като са използвани следните имена и етикети на променливите:

Име на променливата Name	Етикет на променливата Label	Етикети на категориите Values	Измерителна скала Measure	Мерна единица	Посока
grupa		Value - Labe. 1 – enperimental 2 - kontrol	Номинална		
t1_1	30 m. I izsledvane		Метрично скалирана	Сек.	-
t2_1	Sovalka I izsledvane		Метрично скалирана	Сек.	-
t3_1	Jonglirane I izsledvane		Метрично скалирана	Брой	+
t4_1	Kombinran I izsledvane		Метрично скалирана	Сек.	-
t1_2	30 m. II izsledvane		Метрично скалирана	Сек.	-
t2_2	Sovalka II izsledvane		Метрично скалирана	Сек.	-
t3_2	Jonglirane II izsledvane		Метрично скалирана	Брой	+
t4_2	Kombinran II izsledvane		Метрично скалирана	Сек.	-

Приложение 2. Данни за упражнение 2

За упражнението и тестовия контрол към тема 2 са ползвани следните данни:

Данни за задачата за разучаване 2.1.

Дефинирането на променливите group, t1_1 и t1_2 и въвеждането на данните е онагледено в демонстрация 2.1, а изчисляването на прираста на резултатите в теста (p_t1) – в демонстрация 2.2.

	Group	t1_1	t1_2	p_t1
1	experimental	5,71	5,2	
2	experimental	6,53	5,71	
3	experimental	5,86	5,31	
4	experimental	6,27	5,79	
5	experimental	5,7	5,21	
6	kontrol	6,1	6	
7	kontrol	6,08	6	
8	kontrol	5,89	5,9	
9	kontrol	5,89	5,81	
10	kontrol	6	5,95	

Данни за тестовете към тема 2.

Тест 2.1. – Дефинирайте променливите group, t2_1 и t2_2.

Тест 2.2. Изчислете прираста на резултатите в тест Совалка (p_t2), като разлика между t2_2 и t2_1.

	Group	t2_1	t2_2	P_t2
1	experimental	15,11	14,8	
2	experimental	14,3	14,03	
3	experimental	14,69	14,96	
4	experimental	14,69	14,63	
5	experimental	14,3	14,21	
6	kontrol	14,13	12,07	
7	kontrol	14,11	12,23	
8	kontrol	15,66	13,81	
9	kontrol	14,84	14,01	
10	kontrol	14,13	13,05	

Приложение 3. Критични стойности на коефициента на асиметрия (A_s) при $\alpha=0,05$

Обем на извадката (n)	$A_{s,0,05}$	Обем на извадката (n)	$A_{s,0,05}$
10	1,374	75	0,555
15	1,160	80	0,538
20	1,024	85	0,522
25	0,927	90	0,508
30	0,854	95	0,495
35	0,795	100	0,483
40	0,748	150	0,396
45	0,707	200	0,344
50	0,673	250	0,308
55	0,643	300	0,281
60	0,617	350	0,261
65	0,594	400	0,244
70	0,574		

**Приложение 4. Критични стойности на коефициента на ексцес (E_x)
При равнище на значимост $\alpha=0,05$**

Обем на извадката (n)	$E_{x,0,05}$	Обем на извадката (n)	$E_{x,0,05}$
10	2,668	75	1,096
15	2,242	80	1,064
20	1,985	85	1,034
25	1,803	90	1,006
30	1,665	95	0,980
35	1,556	100	0,957
40	1,465	150	0,787
45	1,389	200	0,684
50	1,324	250	0,614
55	1,267	300	0,561
60	1,217	350	0,520
65	1,172	400	0,487
70	1,133		

Приложение 5. Нормирано нормално разпределение в %

Z	P%	Z	P%	Z	P%
0.1	7.97	1.1	72.87	2.1	96.43
0.2	15.85	1.2	76.99	2.2	97.22
0.3	23.58	1.3	80.64	2.3	97.86
0.4	31.08	1.4	83.85	2.4	98.36
0.5	38.29	1.5	86.64	2.5	98.76
0.6	45.15	1.6	89.04	2.6	99.07
0.7	51.61	1.7	91.09	2.7	99.31
0.8	57.63	1.8	92.81	2.8	99.49
0.9	63.19	1.9	94.26	2.9	99.63
1.0	68.27	2.0	95.45	3.0	99.73

**Приложение 6. Критични стойности на коефициента на обикновена
линейна корелация на Пирсън (r)**

Степени на свобода ($k=n-2$)	Равнище на значимост (α)		Степени на свобода (k)	Равнище на значимост (α)	
	0,05	0,01		0,05	0,01
5	0,75	0,87	25	0,38	0,49
6	0,71	0,83	26	0,37	0,48
7	0,67	0,80	27	0,37	0,47
8	0,63	0,77	28	0,36	0,46
9	0,60	0,74	29	0,36	0,46
10	0,58	0,71	30	0,35	0,45
11	0,55	0,68	35	0,33	0,42
12	0,53	0,66	40	0,30	0,39
13	0,51	0,64	45	0,29	0,37
14	0,50	0,62	50	0,27	0,35
15	0,48	0,61	60	0,25	0,33
16	0,47	0,59	70	0,23	0,30
17	0,46	0,58	80	0,22	0,28
18	0,44	0,56	90	0,21	0,27
19	0,43	0,55	100	0,20	0,25
20	0,42	0,54	200	0,11	0,18
21	0,41	0,53			
22	0,40	0,52			
23	0,40	0,51			
24	0,39	0,50			

Приложение 7. Критични стойности на коефициента на рангова корелация на Спирмън

Степени на свобода ($k=p-2$)	Равнище на значимост (α)		Степени на свобода (k)	Равнище на значимост (α)	
	0,05	0,01		0,05	0,01
6	0,89	-	21	0,44	0,56
7	0,79	0,93	22	0,43	0,55
8	0,74	0,88	23	0,42	0,54
9	0,68	0,83	24	0,41	0,53
10	0,65	0,79	25	0,40	0,52
11	0,61	0,77	26	0,39	0,51
12	0,59	0,75	27	0,38	0,50
13	0,56	0,71	28	0,38	0,49
14	0,54	0,69	29	0,37	0,48
15	0,52	0,66	30	0,36	0,47
16	0,51	0,64	31	0,36	0,46
17	0,49	0,62	32	0,36	0,45
18	0,48	0,61	33	0,34	0,45
19	0,46	0,60	35	0,33	0,43
20	0,45	0,58	40	0,31	0,40

Приложение 8. Критични стойности на t-критерия на Стюдънт

Степени на свобода (k)	Равнище на значимост (α)		Степени на свобода (k)	Равнище на значимост (α)	
	0,05	0,01		0,05	0,01
1	12,71	63,60	21	2,08	3,82
2	4,30	9,93	22	2,07	3,79
3	3,18	5,84	23	2,07	3,77
4	2,78	4,60	24	2,06	3,75
5	2,57	4,03	25	2,06	3,73
6	2,45	3,71	26	2,06	3,71
7	2,37	3,50	27	2,05	3,69
8	2,31	3,36	28	2,05	3,67
9	2,26	3,25	29	2,04	3,66
10	2,23	3,17	30	2,04	3,65
11	2,20	3,11	40	2,02	3,55
12	2,18	3,06	50	2,01	3,50
13	2,16	3,01	60	2,00	3,46
14	2,15	2,98	80	1,98	3,42
15	2,13	2,95	100	1,98	3,39
16	2,12	2,92	120	1,97	3,37
17	2,11	2,90	200	1,96	3,34
18	2,10	2,88	500	1,96	3,31
19	2,09	2,86			
20	2,09	2,85			

Забележка: При зависими извадки степените на свобода са равни на $k=n-1$

При независими извадки степените на свобода са равни на $k=n_1+n_2-2$