УНИВЕРЗИТЕТ У БЕОГРАДУ

МАТЕМАТИЧКИ ФАКУЛТЕТ

Марина Николић

ЗНАЧАЈ СТАТИСТИКЕ У МЕДИЦИНСКОМ ОБРАЗОВАЊУ. ПИТАЊЕ ИЗБОРА И ПРИМЕНЕ СТАТИСТИЧКИХ ТЕСТОВА

МАСТЕР РАД

Београд, 2012.

УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Марина Николић

ЗНАЧАЈ СТАТИСТИКЕ У МЕДИЦИНСКОМ ОБРАЗОВАЊУ. ПИТАЊЕ ИЗБОРА И ПРИМЕНЕ СТАТИСТИЧКИХ ТЕСТОВА

МАСТЕР РАД

Др. Весна Јевремовић, ментор

Др. Слободанка Јанковић, председник комисије

Мр. Марко Обрадовић, члан комисије

Београд, 2012.

Предговор:

У овом раду, пажња је посвећена значају статистике у медицинском образовању као и њеној заступљености у средњим медицинским школама. Указано је на потребе образовања из области статистике и познавања њених основних појмова, како за људе који се баве научним истраживањима у области медицине и фармације тако и за оне који се едукују за рад у здравственим установама као што је то случај са већином ученика у средњим медицинским школама.

Размотрен је наставни план и програм за математику у средњим медицинским школама у Србији чиме је сагледана заступљеност статистике и предложено је неколико основних појмова на које би требало ставити акценат када је у питању први сусрет ученика са медицинском статистиком.

Указано је и на значај статистичке писмености истраживача који се баве биомедицинским истраживањима, важност разумевања већ обрађених и представљених података као и способности за обраду података, анализу и њихово приказивање. Приказани су примери табела које се налазе у публикованим биомедицинским радовима да би се указало на то колико је довољно познавати статистику у циљу разумевања тих радова на прави начин, а колико је потребно да би се такве и сличне анализе урадиле од сопствених података.

Узевши у обзир најчешће грешке које се дешавају при обради података разматрано је и питање избора и примене статистичких тестова.

Направљен је кратак преглед најчешће примењиваних параметарских и непараметарских тестова који се користе приликом обраде података у клиничким и епидемиолошким студијама.

Садржај

1.Елементи статистике у настави за средње медицинске школе

Свакодневно се у медијима сусрећемо са великим бројем резултата разних истраживања углавном везаних за начин живота, здравље, исхрану и медицину уопште. Медицина се у зачетку сваког свог истраживања служи методама процене: посматрањем, опажањем, мерењем, и здравом логиком док статистика у помоћ прискаче својим методама које користе: статистике, моделовање, p-вредности, интервале поверења, оцењивање параметара и тестирање хипотеза. Колико је статистика значајна када су медицинска истраживања у питању говоре нам истраживачки кораци који од самог почетка истраживања, ради долажења до ваљаних резултата, консултују статистику. Већ у фази планирања истраживања неопходно је консултовати статистичара. Од циља истраживања уско зависи избор статистичких тестова, а од одабраног статистичког теста зависи обим узорка, што значи да истраживање мора бити веома прецизно испланирано још пре него што почнемо прикупљати узорке. Ваљана употреба статистичких тестова опет зависи од обима узорка као и од квалитета података. Како од резултата статистичке анализе зависи и квалитет истраживања јасно је да статистика има веома битну, ако не и суштинску улогу у фази самог планирања истраживања.

> *Консултовати статитичара, када је експеримент завршен је исто као питати га за посмртни преглед. Тада може само да утврди од чега је експеримент умро.*
>
> ***Роналд Фишер***

1.1 Значај познавања статистичких појмова

Научни радови у области биомедицине из целог света, који представљају главни ресурс литературе за научнике препуни су статистичких резултата, приказаних на разноврсне начине. Њихово тумачење је веома отежано уколико истраживач није вешт у разумевању појмова као што су најчешће помињани: средња вредност и стандардна девијација-одступање (mean ± SD), интервал поверења (CI), p вредност (p value), медијана (median), квантили (quantiles)... као и у тумачењу сложених хистограма, регресионих и корелационих графикона (plots) и сл. Стога је веома битно увести што више појмова који ће олакшавати даље усавршавање ученика медицинских школа како у даљем школовању тако и у случају рада у струци.

1.2 План и програм рада за математику у средњој медицинској школи

Као литература за четврти разред средње медицинске школе користи се само Збирка решених задатака из математике 4, Срђан Огњановић, Круг 2005 где су, пре скупа задатака из статистике, јасно уведени појмови: узорак, узорачка средина, узорачка дисперзија и узорачко стандардно одступање. За обрађивање основних појмова из вероватноће предвиђена су два школска часа на којима се уводе основне дефиниције догађаја, скупа елементарних догађаја и вероватноће, као и појам условне вероватноће и Бајесова формула. За увођење појма статистике и обраду основних појмова предвиђена су такође само два школска часа која су посвећена увођењу појма обележја, узорка, хистограма и расподеле обележја (први час), док се појам аритметичке средине и стандардне девијације уводи на другом часу, чиме се завршавају сва предавања везана за статистику. Да би се усвојило основно знање из статистике и стекао увид у њен значај јасно је да је неопходно издвојити више часова за обраду нових појмова и њихову практичну примену.

Укупан годишњи фонд часова из математике: 56

недељни фонд часова: 2, медицински техничар разред: **IV**

Тема 4: Увод у вероватноћу и статистику (12 часова)

ЦИЉ ТЕМЕ: Стицање знања о основним појмовима у вероватноћи и статистици. **ИСХОДИ ТЕМЕ:** Ученик треба да зна основне појмове вероватноће и статистике и да разуме њихов значај у анализи појава у природним и друштвеним процесима.

Литература: С.Огњановић, Збирка решених задатака из математике 4, Круг 2005 [10]

р. бр.	НАЗИВ НАСТАВНЕ ЈЕДИНИЦЕ	ТИП ЧАСА
45.	Догађај. Скуп елементарних догађаја. Дефиниције вероватноће.	**Обрада**
46.	Вероватноћа уније догађаја.	Утврђивање
47.	Вероватноћа уније догађаја.	Утврђивање
48.	Контролна вежба (вероватноћа)	Провера
49.	Условна вероватноћа. Бајесова формула.	**Обрада**
50.	Условна вероватноћа. Бајесова формула.	Утврђивање
51.	Обележје, узорак, табеле и хистограми расподеле обележја.	**Обрада**
52.	Аритметичка средина и стандардна девијација.	**Обрада**
53.	Вероватноћа и статистика.	Утврђивање
54.	Четврти писмени задатак (Вероватноћа и статистика).	Оцењивање
55.	Исправак четвртог писменог задатка.	Утврђивање
56.	Закључивање оцена	Оцењивање

1.3 Појмови из статистике који се обрађују у настави

Уводни час из статистике требало би да има тежиште на значају статистике и њеној примени у пракси како би се ученици више заинтересовали за саму материју. У те сврхе могла би да послужи шема на слици 1.3.1 која детаљно описује улогу статистике у сваком истраживању које своје резултате базира на подацима.

Слика 1.3.1

Дефиниције које се користе при увођењу статистике у медицинској школи односе се на узорак и његове нумеричке карактеристике. Узорачка средина, узорачка дисперзија уводе се на следећи начин:

Узорак обима n је n-торка $(x_1, x_2, \ldots, x_n)$ реализованих вредности обележја X.

Узорак се уводи као скуп података из којег се израчунавају нумеричке карактеристике обележја.

Узорачка средина обележја X је $\overline{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$.

Као основна нумеричка карактеристика узорка уводена је узорачка средина. Принцип израчунавања вредности узорачке средине је ученицима од раније познат као појам аритметичке средине, с том разликом што је сада формула примењена на одређени узорак и даје нам информацију о његовој равнотежној тачки.

Узорачка дисперзија (варијанса) обележја X је $\overline{S}_n^{\,2} = \frac{1}{n}\left[(x_1-\overline{x})^2+\ldots+(x_n-\overline{x})^2\right]$.

Узорачка дисперзија се уводи као величина која представља просечно квадратно одступање од просечне вредности, односно узорачке средине.

Поправљена узорачка дисперзија обележја X је $\tilde{S}_n^{\,2} = \frac{1}{n-1}\left[(x_1-\overline{x})^2+\ldots+(x_n-\overline{x})^2\right]$.

Узорачко стандардно одступање (стандардна девијација) је позитивна вредност квадратног корена узорачке дисперзије обележја X и често се означава са S.

Оно што недостаје, а што би било корисно додати при увођењу наведених појмова су неке њихове особине. Тачније, уз дефиницију узорачке средине било би корисно нагласити да се вредност $\overline{x}_n$ понаша као равнотежна тачка у скупу, док је њен недостатак то што на њену вредност утичу екстремне вредности у скупу (outliers). Требало би додати и то да се аритметичка средина изражава у истим јединицама као и основни подаци. Уз дефиницију за узорачко стандардно одступање тј. стандардну девијацију треба додати да је то величина која нам говори колико у просеку елементи скупа одступају од аритметичке средине скупа. По увођењу свих особина нумеричких карактеристика било би пожељно објаснити да нумеричке карактеристике не одређују, у потпуности, случајну величину и да реализовани узорак даје више информација о посматраном обележју него сама узорача средина као и то да више узорака истог обима из једне популације може имати и врло различите вредности узорачке средине.

1.4 Појмови које би било корисно увести у наставу

Примери који следе (**1.4.1**, **1.4.2**, **1.4.3** , **1.4.4**) представљају табеле преведене из страних научних студија и послужиће као илустрација за образложење, које једноставни статистички појмови би се могли увести у наставу средњих медицинских школа, да би се омогућило лакше разумевање овакве и сличне литературе.

Пример 1.4.1

Дијететски унос одабраних нутријената у поређењу са Нордијским дијететским препорукама за труднице [5]

Нутријент	I квартил	медијана	III квартил	препоруке за труднице
Витамин А, μg	697	979	1433	800
Витамин Д, μg	2,2	2,9	4,5	10
Витамин Е, μg	8,2	9,5	11,1	10
Рибофлавин, mg	1,7	2,1	2,5	1,6
Фолат, μg	262	305	352	400
Витамин Б12, mg	4,6	5,9	7,9	2
Витамин Ц, mg	100	143	194	70
Калцијум, mg	1101	1342	1636	900
Магнезијум, mg	296	333	394	280
Цинк, mg	10	12	14	9
Јод, μg	251	289	339	175
Селенијум, μg	47	57	66	55

У примеру 1.4.1 приказан је унос одабраних нутријената у виду првог и трећег ***квартила*** као и преко ***медијане***. Ово су појмови чија се обрада не налази у плану и програму за статистику у средњим медицинским школама а могуће их је увести на веома једноставан начин као додатне мере централне тенденције поред узорачке средине.

Медијана или медијална вредност је вредност која дели статистички скуп, у коме су подаци уређени по величини, на два једнака дела тако да се изнад и испод медијане налази једнак број података.

Ако статистички низ има непаран број података медијана је централна вредност а њен редни број је $\frac{N+1}{2}$.

Ако низ има парни број података медијана је средина између две центалне вредности у низу односно средина између вредности чији су редни бројеви $\frac{N}{2}$ и $\frac{N+2}{2}$. Главна особина медијане је да на њену вредност не утичу екстремне вредности, као што је то случај са узорачком средином, па у случају њиховог постојања не даје реалну слику о подацима. Медијана је, због наведеног својства, а за разлику од узорачке средине, такозвана робусна оцена, односно, није осетљива на измене података у низу.

Мера централне тенденције, која такође има велику примену у истраживањима, и могла би се на једноставан начин увести у наставу је мода.

Мода је вредност обележја која има највећу фреквенцију, односно, највећу заступљеност у оквиру укупне фреквенције. Када су медицинска истраживања у питању мода се користи када се на основу категоричких или нумеричких података жели представити колика је учесталост неке појаве. Неке појаве могу да имају и две модалне вредности па кажемо да су бимодалне, односно, мода не мора бити јединствена вредност скупа као што је то узорачка средина. Мода такође, као и медијана није мера осетљива на екстремне вредности.

Пример 1.4.2

Учесталост каријеса (n=100)	
Узраст	Број деце са каријесом
5	10
6	28
7	30
8	10
9	13
10	10

Табела из примера 1.4.2представља узраст у којем је каријес најзаступљенији код деце и у овом случају је то период шесте и седме године живота, односно постоје две модалне вредности: 28 и 30.

Уколико је обележје једномодално важи правило да се медијана увек налази између узорачке средине и моде.

Узорачки l-процентни квантил је број који је већи од *l* % вредности из узорка. Специјално, ако је у питању 25 % елемената из узорка, одговарајући квантил се зове ***први квартил*** и означава се са Q1, ако је у питању 50 % елемената из узорка одговарајући квантил се поклапа са медијаном, а ако је у питању 75 % елемената из узорка, одговарајући квантил се зове ***трећи квартил***, и означава се са Q3.

При оваквом приказивању података одређују се још и такозване унутрашње f_1 i f_3 и спољашње F_1 i F_3 границе узорка на следећи начин:

$$f_1 = q_1 - 1.5(q_3 - q_1), f_2 = q_1 + 1.5(q_3 - q_1),$$

$$F_1 = q_1 - 3(q_3 - q_1), F_3 = q_1 + 3(q_3 - q_1)$$

Пример 1.4.3: На датом узорку одредићемо први и трећи квартил, унутрашње и спољашње границе као и медијану тежине пацијената у килограмима.

65	72	58	69	79	70	85	67	78	83	80	67	60

Пре него што почнемо са рачунањем низ је потребно сортирати у растући поредак.

58	60	65	67	67	69	70	72	78	79	80	83	85

Како први квартил представља вредност која одваја доњих 25 % вредности од горњих 75%, и како у овом примеру 25% чине прва три члана низа (25% од 13) то ће први квартил бити четврта вредност у низу, односно 67. Трећи квартил представља вредност која се налази изнад 75% чланова низа што је у овом случају десети члан низа односно 79. Медијану у овом случају проналазимо као средњи члан низа јер низ има непаран пркој чланова, дакле изнад и испод медијане се налази по 6 чланова низа, стога је медијана овог низа 70.

58	60	65	$\boxed{67}$	67	69	$\boxed{70}$	72	78	$\boxed{79}$	80	83	85
			Q_1			Me			Q_3			

Да бисмо илустовали израчунавање медијане за низ са парним бројем чланова посматрајмо дати сортиран низ без последњег члана, тачније, низ од 12 чланова.

58	60	65	67	67	69	70	72	78	79	80	83

Како су две централне вредности у низу шеста и седма односно 69 и 70 медијана ће бити аритметичка средина ових бројева, односно 69,5.

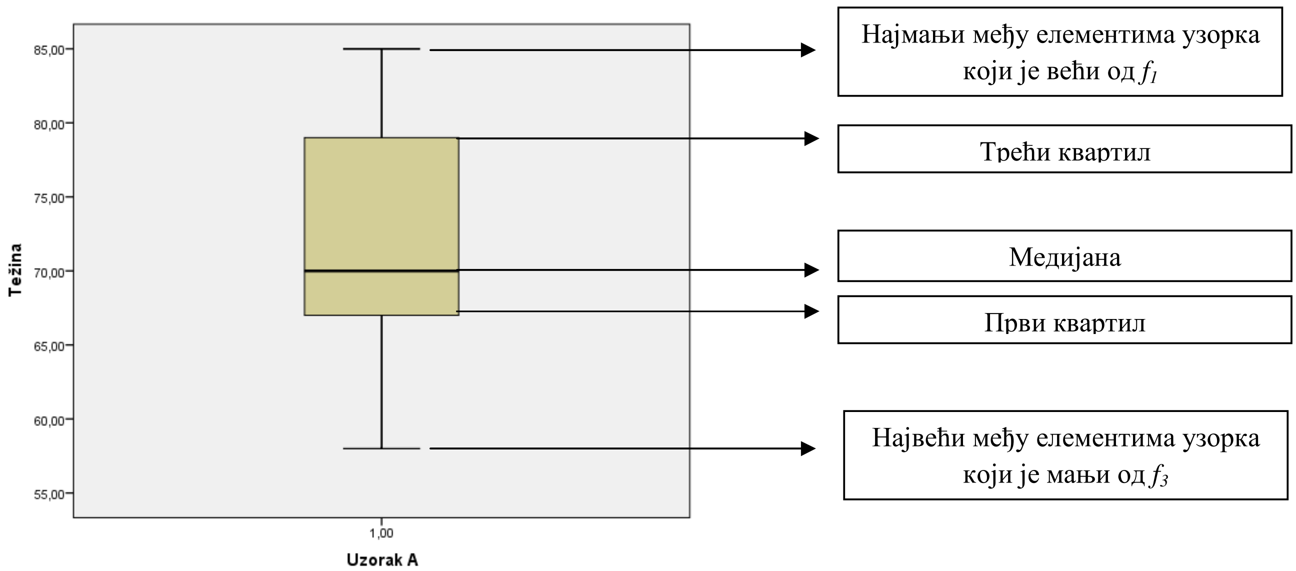
Ако заменимо прва два члана низа знатно мањим вредностима а последња два члана знатно већим вредностима добијамо следећи низ:

35	40	65	67	67	69	70	72	78	79	98	104

чија ће медијана такође бити аритметичка средина шесте и седме вредности у низу, односно 69,5 што нам говори због чега је медијана заправо робусна оцена и колико је неосетљива на измену података за разлику од узорачке средине.

Унутрашње и спољашње проналазимо на основу израчунатог првог, односно трећег квартила и добијамо вредности $f_1 = 49, f_3 = 85, F_1 = 31, F_3 = 103$.

Јаснији смисао бројева које смо израчунали даје њихов графички приказ такозваним *BOX-PLOT* дијаграмом на слици 1.4.1.



Слика 1.4.1

У примеру 1.4.4 подаци су представљени као средња вредност и стандардна девијација а њихова разлика у виду *p* вредности. У програму за средњу школу постоји јединица којом се обрађује узорачка средина и стандардна девијација док се *p* вредност као вероватноћа, да при тачној нултој хипотези, узорак припада критичној области, уопште не уводи. На основу примера 1.4.4 илустративно би се могло објаснити да са вероватноћом 1-0,024=**0,976**, на основу узорка, можемо тврдити, да деца која имају нормалан статус гвожђа имају статистички значајно мање протеина у % тоталне енергије у односу на децу са недостатком гвожђа.

Пример 1.4.4

Карактеристике које су у вези са недостатком гвожђа и нормалним статусом гвожђа код азијске деце од 12 до 24 месеца [6]			
	Недостатак гвожђа средња вредност (СД)	Нормалан статус гвожђа средња вредност (СД)	p вредност
n	12	20	
Старост (месеци)	19,5 (6,346)	20,65 (6,401)	нема значајне разлике
Тежина на рођењу	3,29 (0,314)	2,99 (0,719)	нема значајне разлике
Тренутна тежина	11,48 (2,017)	10,49 (1,829)	нема значајне разлике
Хемоглобин (g/dl)	9,78 (0,721)	11,92 (0,623)	нема значајне разлике
Феритин (μg/l)	4,42 (1,782)	22,95 (0,623)	нема значајне разлике
Енергија kJ/d	3470 (739,1)	3706 (1308,6)	нема значајне разлике
Енергија kJ/kg/d	307 (73,9)	351 (97,4)	нема значајне разлике
Протеини g/d	36 (6,4)	34 (14,4)	нема значајне разлике
Протеини g/kg/d	3,2 (0,67)	3,2 (1,10)	нема значајне разлике
Протеини као % тоталне енергије	17,6 (1,93)	15,2 (3,20)	*p = 0,024**
Масноће g/d	40 (11,2)	41 (19,9)	нема значајне разлике
Масноће као % тоталне енергије	42,7 (6,67)	40,1 (7,16)	нема значајне разлике
Енергетска густина kJ/g хране	3,4 (0,62)	4 (1,16)	нема значајне разлике
Гвожђе mg/d	2,82 (1,495)	4,14 (2,424)	нема значајне разлике
Витамин Ц mg/d	23 (19,4)	41 (39,9)	нема значајне разлике
Рибофлавин mg/d	1,76 (0,466)	1,45 (0,626)	нема значајне разлике

*Mann-Whitney U- test

Интервал поверења[1] за средњу вредност је такође веома важан и често коришћен појам у медицинској статистици. Његово увођење у план и програм за средње медицинске школе не би било сувише компликовано обзиром да се сви појмови преко којих се уводи интервал поверења већ обрађују по плану. У примеру 1.4.5 подаци су приказани као средња вредност, интервал поверења средње вредности (CI) као и p вредност за све три врсте коришћених тестова. *Пирсонов* χ^2 *тест* је непараметарски тест за проверу хипотеза о расподели обележја који је применљив на обележја и са дискретним и са непрекидним расподелама. *Ман Витни U-тест* је детаљно описан у поглављу 4.1.

[1]Појам интервал поверења је први пут увео 1934. године пољски математичар Jerzy Spława-Neyman (1894-1981)

Пример 1.4.5

Коришћење фолне киселине у репродуктивном периоду у Норвешкој према имиграционом статусу [7]			
	Норвежанке (1136) Узорачка средина (95% ИП)	Имигранти (398) Узорачка средина (95% ИП)	p вредност
Старост (у годинама)	30,6 (30,3-30,8)	28,2 (27,7-28,6)	< 0,001*
Коришћена фолна киселина током целе трудноће (%)	72,5 (69,9-75,1)	19,1 (15,2-23)	< 0,001†
Правилно коришћена фолна киселина током трудноће (%)	21,8 (20,2-25,0)	2,3 (0,8-3,8)	< 0,001†
Регион			< 0,001‡
Крајњи запад	24,5% (22,0-27,0)	10,3 % (7,3-13,3)	
Средњи запад	18,8% (16,5-21,1)	3,8% (1,9-5,7)	
Крајњи исток	29,3% (26,7-31,9)	59,3% (56,8-61,8)	
Средњи исток	22,1% (19,7-24,5)	25,1% (20,8-29,4)	
Близина Осла	5,2% (3,0-7,4)	1,5% (0,3-2,7)	

*Mann-Whitney U-test
† Fischer's exact test
‡ Pearson X^2 test
ИП-интервал поверења

Уобичајне вредности вероватноће β за коју се израчунава интервал поверења су обично вредности блиске јединици, 0.99, 0.95, или 0.90, те би се за ученике средњих школа z вредности неопходне за израчунавање граница интервала поверења, за ове вероватноће могле увести као константе:

$z_{0.90}=1.645,$

$z_{0.95}=1.96,$

$z_{0.99}=2.58$,

пошто су оне једини непознати појмови у средњој школи који су неопходни за увођење појма интервала поверења.

Ако је (a,b) интервал поверења са такозваним нивоом значајности $1-\alpha$, тада је вредност α (0.10, 0.05, или 0.01) вероватноћа да се оцењивани параметар налази ван датог интервала поверења . Тако у примеру 1.4.5 можемо тврдити са вероватноћом *p=0,95* да је између 69,9% и 75,1% жена користило фолну киселину током целе трудноће.

А једноставни изрази за израчунавање граница интревала поверења за математичко очекивање могу се увести, користећи познат појам узорачке средине и задате дисперзије обележја (σ), као границе интервала:

$$\left[\overline{x}_n - \frac{z_\beta}{\sqrt{n}}\sigma, \overline{x}_n + \frac{z_\beta}{\sqrt{n}}\sigma\right]$$

Теоријски опис интервала поверења за математичко очекивање је дат у делу 2.3.4 овог рада, након увођења нормалне расподеле.

2. Неки статистички тестови које би било корисно увести у наставу

2.1 Тестирање хипотеза:

Већина статистичких анализа у биомедицинским истраживањима укључује поређење углавном између различитих третмана и процеса или између две или више група пацијената. У пракси се издвајају две групе метода за тестирање статистичких хипотеза: параметарске и непараметарске. Обе групе су једнако заступљене у биомедицинским истраживањима у зависности од обима и расподеле обележја чији узорак посматрамо.

Основна претпоставка за параметарске тестове је познавање расподеле популације док код непараметарских тестова то није потребно, јер се сваки непараметарски тест остварује на исти начин, без обзира на расподелу обележја.

Када су медицинска истраживања у питању, често немамо довољан број података који потичу из популације са нормалном расподелом да бисмо применили неки од параметарских тестова, стога је веома битно изабрати одговарајући непараметарски тест и доћи до валидних резултата. Битно је напоменути да статистички тестови, заправо, само указују на сагласност узорака, односно, расположивих података и постављене хипотезе за дати ниво значајности.

2.2 Основни појмови:

Статистичка хипотеза је претпоставка о функцији расподеле једног или више обележја, облику, карактеристикама или вредностима непознатих параметара популације. Тестирање хипотезе подразумева постављање две статистичке хипотезе, нулте и алтернативне.

Нулта хипотеза H_0 , у случају параметарских тестова, представља хипотезу којом се тестира да се непознати параметар θ налази у скупу Θ_0 где је Θ_0 подскуп скупа допустивих вредности Θ.

У случају непараметарских тестова, H_0 представља хипотезу да обележје X има дату функцију расподеле $F_0(x)$ што се записује у облику $H_0(F(x)=F_0(x))$

Нулта хипотеза је обично она чијим погрешним одбацивањем би се направила велика грешка. На пример, ако посматрамо нулту хипотезу да је нови лек за срце на тржишту штетан, погрешно одбацивање те хипотезе би значило да лек није штетан, иако заправо јесте, па би то могло да има велике негативне последице.

Алтернативна хипотеза H_1, представља закључак који се доноси ако се одбаци нулта хипотеза, односно, у случају параметарских тестова, да се непознати параметар θ налази у комплементу скупа допустивих вредности - критичној области. У случају непараметарских тестова H_1 је хипотеза да је расподела $F(x)$ различита од $F_0(x)$, тј. $H_1(F(x) \neq F_0(x)).H_1$ је обично претпоставка коју истраживањем желимо да докажемо. Алтернативна хипотеза за претходни пример би била да нови лек за срце није штетан.

Статистички тест је правило које нам омогућава да донесемо одлуку о одбацивању или неодбацивању (прихватању) нулте хипотезе H_0 на основу добијене вредности ***тест статистике***, која представља функцију добијеног узорка.

Када у алтернативној хипотези није одређен смер ефекта, примењује се ***двострани тест***. На пример, ако није унапред познато да ли је проценат пушача код мушкараца већи или мањи у односу на жене у популацији примењује се двострани тест.

Једнострани тест примењује се када је смер ефекта одређен у алтернативној хипотези H_1. На пример, у истраживању болести од које сви нелечени пацијенти умиру па нови лек не може погоршати ситуацију, користили би смо једнострани тест.

Ако се одбацује нулта хипотеза када је она тачна, долази до грешке првое врсте, а ако се прихвата нулта хипотеза онда када је тачна алтернативна хипотеза, долази до грешке друге врсте.

Статистички тест је у потпуности одређен критичном облашћу W_n као скупом тачака $(x_1, x_2, ..., x_n)$ у n-димензионалном еуклидском простору на следећи начин:

Ако реализовани узорак $(x_1, x_2, ..., x_n)$ припада критичној области W_n тада се одбацује нулта хипотеза H_0, а ако $(x_1, x_2, ..., x_n)$ не припада W_n онда се прихвата нулта хипотеза. Према томе, вероватноћа грешке првог и другог типа су редом:

$$\alpha = P_{H0}\{(x_1, x_2, ..., x_n) \in W_n\}, \beta = P_{H1}\{(x_1, x_2, ..., x_n) \notin W_n\}.$$

Максимална дозвољена грешка првог типа се назива још и ***ниво (праг) значајности теста.***

Најмањи ниво значајности на коме би хипотеза H_0 била одбачена на основу података из датог реализованог узорка назива се ***p вредност*** или значајност вредности тест статистике. Ако је p вредност реализоване статистике теста мања од задатог нивоа значајности α, хипотеза H_0 се одбацује.

2.3 Теоријске расподеле вероватноће:

2.3.1 Нормална расподела

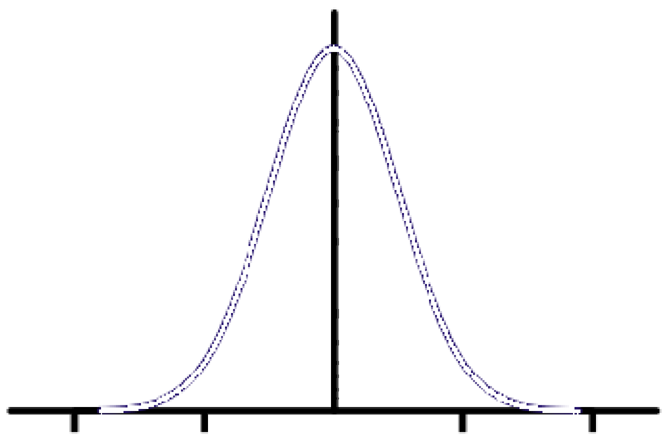
Теоријска расподела која је нашла највећу примену у медицинским истраживањима, када су у питању нумерички подаци , односно резултати добијени путем разних врста мерења, је ***нормална расподела***. Нормална расподела припада типу непрекидних теоријских расподела односно расподела вероватноће непрекидне случајне променљиве. Нормалну расподелу је први проучавао француски научник Abraham de Moivre (1667-1745), док је назив добила по немачком научнику Karlu Friedrichu Gaussu (1777-1855), који је проучавао грешке мерења.

Густина расподеле случајне величине *X* која има нормалну расподелу је :

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \text{ , x} \in \text{R, } \sigma > 0\text{, m} \in \text{R}$$

Каже се да *X* има нормалну расподелу са праметрима m и σ^2 што се означава са

X: N(m,σ^2*).*



Слика 2.3.1

График густине нормалне расподеле се назива Гаусова крива (слика 2.3.1).

На графику густине расподеле случајне величине која има нормалну расподелу *N*(m,σ^2*)* уочавамо следеће:

- Локални максимум се налази у тачки $x = m$,
- График гутине расподеле је симетричан око праве $x = m$, односно 50% вредности налази се лево (*m*) а 50% вредности налази се десно од од математичког очекивања (*m*). Око *m* се концентрише највећи број вредности, а екстремно мале и екстремно велике вредности имају најмању фреквенцију.
- Постоје две превојне тачке $x = m \pm \sigma$

- Како је крива симетрична, математичко очекивање је једнако медијани односно моди:

  $m = M_e = M_o$

- Када $x \to \pm\infty$ хоризонтална асимптота је $y = 0$.

- Максимална вредност функције густине расподеле је $\frac{1}{\sigma\sqrt{2\pi}}$

- Нормална расподела је у потпуности дефинисана математичким очекивањем (m) и дисперзијом (σ). Површина између криве густине нормалне расподеле и x-осе једнака је 1.

- Ако случајна променљива X има нормалну расподелу, тада ће и њена линеарна трансформација $Y = a+bX$ имати нормалну расподелу.

- За нормалну расподелу важе правила:

  - У интервалу $m \pm 1\sigma$ налази се 68,27% свих вредности, односно:

    $$P\ (m-1\sigma < X < m+1\sigma) = 0,6827$$

  - У интервалу $m \pm 2\sigma$ налази се 95,46% свих вредности, односно:

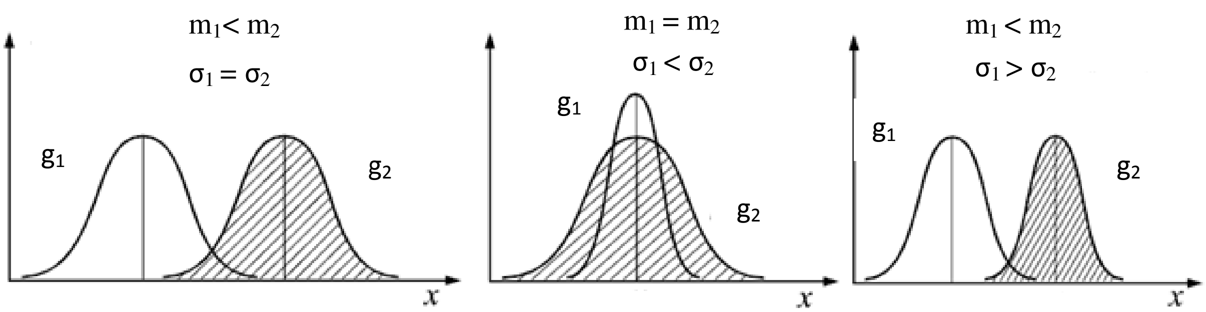
    $$P\ (m-2\sigma < X < m+2\sigma) = 0,9546$$

  - У интервалу $m \pm 3\sigma$ у односу на аритметичку средину налази се 99,73% свих вредности такозвано **3-сигма правило**, односно:

    $$P\ (m-3\sigma < X < m+3\sigma) = 0,9973$$

Највећи број вредности случајне променљиве налазе се у оквиру 6 стандардних девијација, ван овог интервала остаје занемарљив број вредности .

У зависности од тога да ли се m повећава или смањује крива се помера у десно или у лево респективно, док се при повећању стандардне девијације крива снижава и шири а при смањењу крива је све виша и ужа, што је и приказано на слици 2.3.2.



Слика 2.3.2

Специјалан случај представља случајна величина која има нормалну расподелу са параметрима $m=0$ и $\sigma=1$. За такву случајну величину се каже да има стандардизовану (нормирану) нормалну расподелу $N(0,1)$.

Ако случајна величина X има $N(m,\sigma^2)$ расподелу, тада случајна величина $Y= (X-m)/\sigma$ има $N(0,1)$ расподелу. Тачније, за рачунање вредности функције расподеле било које случајне променљиве са нормалном расподелом довољно је знати вредност функције расподеле случајне променљиве са нормалном нормираном расподелом и густином

$$g(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

2.3.2 Студентова t расподела

Познато је да је теорија нормалне расподеле базирана на великом броју података и да се не може применити у колико се располаже са мање од 30 података у узорку што је у пракси чест случај. Ако располажемо са мање од 30 података неопходно је користити нову расподелу, **која није зависна од дисперзије σ**, него само од броја података. Уочивши овај проблем, ирски хемичар Вилијам Госет[2] развио је теоријску расподелу вероватноће случајне променљиве t за нормално расподељене случајне узорке обима мањег од 30. Своју теорију је објавио 1908. године под псеудонимом "Студент" по којој је ова расподела и добила име ***Студентова t расподела***.

Када говоримо о одређеној t расподели мора се одредити и број ***степени слободе***.

---

[2] William Sealy Gosset, (1876-1937), енглески статистичар, први који се бавио проблемом малих узорака.

Нека случајна величина Z има нормалну $N(0,1)$ расподелу, и случајна величина Y има χ_n^2 расподелу. Ако су Z и Y међусобно независне, тада случајна величина:

$$\frac{Z}{\sqrt{\frac{Y}{n}}}$$

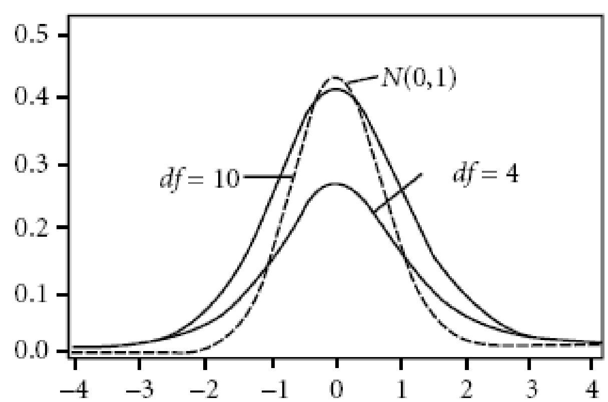
има Студентову t - расподелу са n степени слободе (Слика 2.3.3).

Густина расподеле случајне величине $X{:}t_n$ је:

$$g_n(x) = \frac{G\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,G\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{2}\right)^{-\frac{n+1}{2}}, x \in R$$

Где је $G(n)=(n-1)!,\ n \in N$ гама функција. Гама функција има и свој општији облик када n не припада скупу природних бројева али због примене у статистици најчешће посматрамо случај када n ∈ N.

- Крива t расподеле је симетрична и у облику звона као код нормалне расподеле, и дефинисана од - ∞ до +∞.
- Крива је по облику шира и спљоштенија у односу на Гаусову криву.
- Са порастом броја степени слободе крива t расподеле се ближи нормалној расподели што је приказано на слици.



Промене облика густине t расподеле у зависности од броја степени слободе у односу на Гаусову криву

Слика 2.3.3

2.3.3 Употреба статистичких таблица

У такозваним таблицама нормалне расподеле дате су вредности функције расподеле

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

за нормалну нормирану случајну промењиву.

У употреби су две врсте таблица за нормалну расподелу. Статистички софтвери и већина стране литературе користе таблицу за праву густину расподеле *F(x),* за коју важе правила:

1. *P[X< - a] = F(-a)=1-F(a)*
2. *P[a<X<b]= F(b)+ F(a)*

Међутим, у литератури се често наилази и на другу врсту таблица које дају вредности интеграла

$$\Phi(x) = \int_{0}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Како су у таквим таблицама приказане вредности функције *Φ(x)=F(x)-0,5,* правила за њихово коришћење у пракси су следећа:

3. *P[|X| < a] = 2Φ(a), a > 0*
4. *P[-b<X<a]= Φ(a)+ Φ(b), a,b>0*
5. *P[X >a]={0,5- Φ(a), a >0; 0,5+ Φ(-a),a <0.*

Таблице за Студентову расподелу дају вредности t_α за које је вероватноћа $P[|X|>t_\alpha]=\alpha$

где је α задати ниво значајности, а случајна промењива *X* има Студентову расподелу са *n* степени слободе.

2.3.4 Интервал поверења за средњу вредност:

Реализована вредност оцене параметра може доста одступати од стварне вредности параметра, стога, на основу простог случајног узорка, одређује се интервал који са унапред задатом вероватноћом садржи непознати параметар θ. Ако је (X_1, X_2,...,X_n) узорак обележја X обима n на основу добијеног узорка дефинишу се статистике f(x) и g(x) тако да важе услови :

$$P\{f \le g\}=1,$$

$$P\ \{f \le \theta \le g\} =\alpha, \qquad \alpha \in [0,1].$$

Тада се [f, g] назива ***интервал поверења*** за непознати параметар θ са нивоом поверења α.

Интервал поверења за математичко очекивање обележја X са нормалном расподелом $N(m, \sigma2)$ и **непознатом дисперзијом σ**, користи статистику $\overline{x}_n$ којом се оцењује математичко очекивање m и узорачку дисперзију $\overline{S}_n{}^2$ којим се оцењује дисперзија σ . Због симетричности густине нормалне расподеле у односу на праву $x = m$ статистике f и g бирамо симетрично у односу на $\overline{x}_n : f = \overline{x}_n - \varepsilon$, $g = \overline{x}_n + \varepsilon$. Непознато ε одредићемо, из услова

$$P\left\{\left|\overline{x}_n - m\right| \le \varepsilon\right\} = \beta$$

Како статистика $\dfrac{\overline{x}_n - m}{\tilde{S}_n}\sqrt{n}$ има студентову t_{n-1} расподелу, за дати ниво поверења могу се наћи бројеви t_1 і t_2 такви да је

$$P\left\{t_1 \le \frac{\overline{x}_n - m}{\tilde{S}_n}\sqrt{n} \le t_2\right\} = 1-\alpha,$$

одакле следи да је

$$\left(\overline{x}_n - t_{n-1;\frac{\alpha}{2}} \cdot \frac{\tilde{S}_n}{\sqrt{n}}, \overline{x}_n + t_{n-1;\frac{\alpha}{2}} \cdot \frac{\tilde{S}_n}{\sqrt{n}}\right)$$

100(1-α)% интервал поверења за непознати параметар m. Дужина оваквог интервала поверења је

$$D = 2 \cdot t_{n-1;\frac{\alpha}{2}} \cdot \frac{\tilde{S}_n}{\sqrt{n}}$$

Случајна променљива, односно мења се у зависности од реализованог узорка.

2.4 Студентов t-тест:

Када је потребно тестирати хипотезу $H_0\ (m_1=m_2)$, о једнакости средњих вредности два обележја која имају нормалну расподелу, користимо различите врсте тестова у зависности од тога да ли су обележја која поредимо независна или зависна. Студентов *t*-тест користимо онда када нам **дисперзија није позната**, обично онда када је обим узорка мањи од 30. Основни услов за коришћење *t*-теста је да узорак припада обележју са нормалном расподелом. Примена *t*-теста на узорак из популације, за коју није претходно утврђено да има нормалну расподелу доводи до грешке и до неадекватних резултата који могу да наведу истраживача на потпуно погрешне закључке.

Овде ће бити речи о две врсте t-теста, за независне и зависне узорке.

2.5 Питање избора врсте t-теста у зависности од начина прикупљања података

Када је у питању процес прикупљања података разликујемо две методе у зависности од врсте истраживања. Прва метода односи се на две независне групе испитаника изложене различитом третману. На пример, ако две различите групе пацијената са повишеним шећером у крви третирамо различитим лековима у одређеном периоду и након тога желимо да поредимо ниво шећера у крви и испитамо који лек је делотворнији тада унапред знамо да ћемо за поређење користити t-тест за независне узорке.

Друга метода односи се на јединствену групу испитаника изложену различитим врстама третмана или једном истом третману у различитим временским периодима.

Ова метода је још позната и као метода поновљених мерења и за анализу података прикупљених на овај начин користи се *t*-тест за два зависна узорка (тест парова).

2.6 Студентов *t*-тест за независне узорке

Студентов *t*-тест који се односи на разлику између средњих вредности обележја две независне популације може се примењивати као двострани и као једнострани, уз претпоставку да оба узорка потичу из обележја са нормалном расподелом и дисперзије две популације нису познате. Израз за израчунавање тест статистике зависи од броја података у групама који ће одређивати број степени слободе.

Претпоставимо да имамо два независна узорка са задатим расподелама $X : N(m_1, \sigma^2)$ обима n_1 и $Y : N(m_2, \sigma^2)$ обима n_2. Ако су n_1 и n_2, или један од њих, бројеви мањи од 30, а дисперзије су **једнаке и непознате**, тада се користи тест заснован на статистици:

$$T = \frac{\overline{X}_{n_1} - \overline{Y}_{n_2}}{S} \text{ , где је } S = \sqrt{\frac{n_1 \overline{S}^2_{n_1} + n_2 \overline{S}^2_{n_2}}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}.$$

Расподеле случајних величина $\overline{X}_{n_1}$ и $\overline{Y}_{n_2}$ редом $N(m_1, \sigma^2/n_1)$ и $N(m_2, \sigma^2/n_2)$, а како су те случајне величине и независне , њихова разлика W= $\overline{X}_{n_1}$ - $\overline{Y}_{n_2}$ ће имати нормалну расподелу $N(m1+m2, \sigma^2/n_{1+} \sigma^2/n_2)$.

Нормирањем случајне величине W добијамо статистику:

$$W^* = \frac{\overline{X}_{n_1} - \overline{Y}_{n_1} - (m_1 - m_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} : N(0,1)$$

Како је дисперзија непозната, послужићемо се узорачким дисперзијама $\overline{S}^2_{n_1}$ и $\overline{S}^2_{n_2}$ и

чињеници да су оне независне и да $\frac{n_1 \overline{S}^2_{n_1}}{\sigma^2}$ и $\frac{n_2 \overline{S}^2_{n_2}}{\sigma^2}$ имају редом χ^2_{n1-1} и χ^2_{n2-1}

Због независности ће $U = \frac{n_1 \overline{S}_{n_1}^{\ 2}}{\sigma^2} + \frac{n_2 \overline{S}_{n_2}^{\ 2}}{\sigma^2}$ имати $\chi^2_{n1-1+n2-1}$ расподелу, па тиме долазимо до чињенице да статистика

$$\frac{W^*}{\sqrt{U/(n_1+n_2-2)}}$$

Има Студентову $t_{n1+n2-2}$ расподелу. После краће трансформације из те статистике се добија статистика T са почетка коју користимо у овом тесту.

Дакле, ако је хипотеза H_0 тачна, статистика T има Студентову расподелу са n_1+n_2-2 степени слободе.

Критичне области са нивоом значајности α за тестирање нулте хипотезе $H_0(m_1=m_2)$ против једне од алтернативних:

$$H_1(m_1 \neq m_2), \quad H_1(m_1 > m_2), \quad H_1(m_1 < m_2)$$

су редом:

$$|T| \geq t_{n_1+n_2-2,\frac{\alpha}{2}},\ T \geq t_{n_1+n_2-2,\alpha},\ T \leq -t_{n_1+n_2-2,\alpha}\ .$$

У одређеним применама у пракси не може се претпоставити да су дисперзије једнаке и тада се проблем може решити тако што ће се тест статистика израчунавати према изразу:

$$T = \frac{\overline{X}_{n_1} - \overline{Y}_{n_2}}{\sqrt{\frac{\overline{S}_{n_1}^{\ 2}}{n_1} + \frac{\overline{S}_{n_2}^{\ 2}}{n_2}}},$$

али у том случају морамо модификовати број степени слободе и израчунати га користећи израз:

$$df = \frac{\left(\frac{\overline{S}_{n_0}^{*2}}{n_1} + \frac{\overline{S}_{n_2}^{*2}}{n_2}\right)^2}{\frac{\left(\overline{S}_{n_1}^{*2} / n_1\right)^2}{n_1} + \frac{\left(\overline{S}_{n_2}^{*2} / n_2\right)^2}{n_2}}$$

Вредност добијена на овај начин се заокружује на цео број и представља модификовани број степени слободе. Из овог разлога је неопходно тестирати једнакост стандардних девијација пре употребе самог Студентовог t-теста. Статистички пакети, углавном, у својим извештајима дају информације о једнакости стандардних девјација користећи Ливенов тест који за своју нулту хипотезу има H_0 $(\sigma_1 = \sigma_2)$ и у зависности од тога да ли се она прихвата или не прихвата, резултат Студентовог t-теста се разликује , што је описано у тумачењу за пример 2.6.1).

Пример 2.6.1

Фармацеутска компанија А је избацила на тржиште нове таблете аспирина. На 30 особа које пате од главобоље испитан је узорак од 18 таблета компаније А и 12 таблета компаније Б. Мерено је време (у минутама) које је потребно да престане бол у циљу провере да ли постоји значајна разлика у деловању таблета ове две компаније. Добијени резултати у две посматране групе су:

Компанија А: $\overline{x}_1$=7,87 мин, σ=2,24 мин

Компанија Б: $\overline{x}_2$=8,07 мин, σ=1,45 мин

Нулта и алтернативна хипотеза гласе: H_0: $m_1 = \mu_2$ и H_1: $m_1 \neq \mu 2$

α = 0,05 , df = 18 + 12 – 2 = 28, таблична вредност $t_{0,05,\ 28} = 2,048$

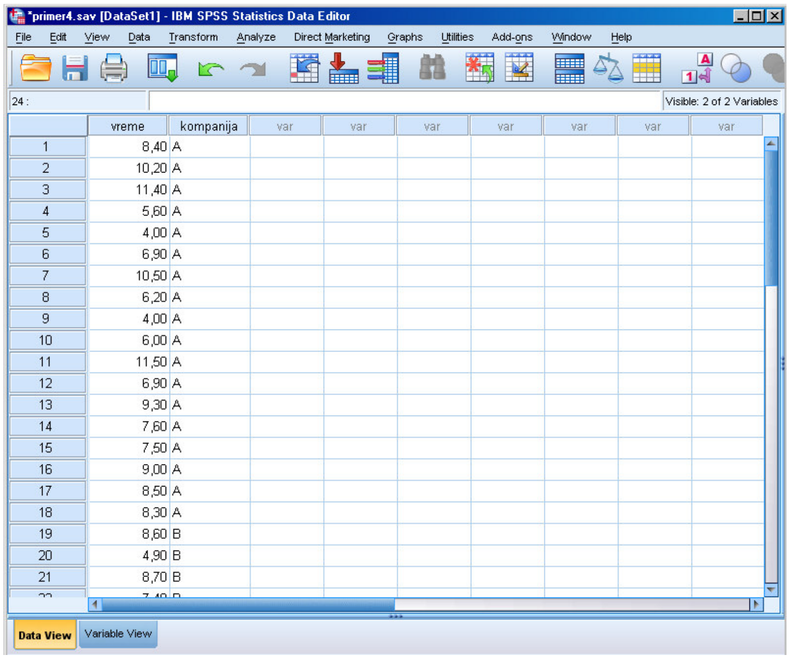
$$t = \frac{|7,87 - 8,07|}{4,12.\sqrt{\frac{18+12}{18 \cdot 12}}} = 0,26$$

Критична вредност t за $\alpha=0{,}05$ и $df=28$ је 2,048. Како је израчуната вредност t мања од ове, нулта хипотеза се прихвата и долазимо до закључка да не постоји статистички значајна разлика између средњих вредности ова два узорка.

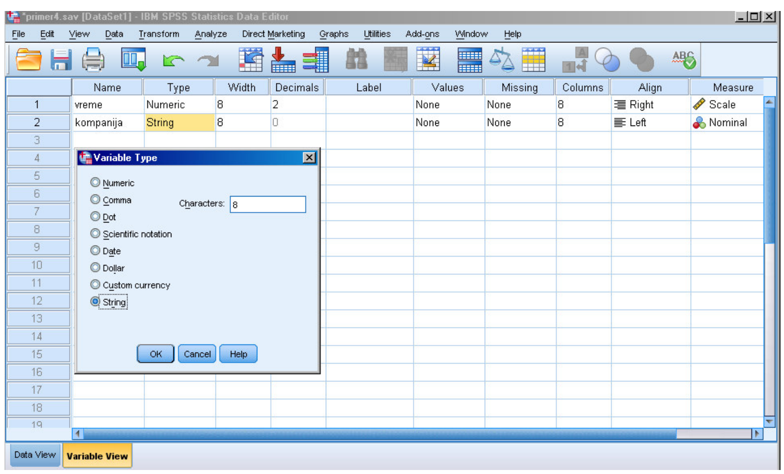
SPSS извештај за пример 2.6.1 приказан је на слици 2.6.4.

2.6.1 Употреба SPSS-а

SPSS (Statistical Package for the Social Sciences) је статистички софтвер који је веома приступачан и релативно једноставан за употребу при обради података из области биомедицине. Лак је за коришћење и интерпретацију резултата захваљујући извештају који веома детаљно приказује резултате. На слици 2.6.2 приказано је радно окружење SPSS софтвера до којег долазимо кликом на други табулатор у доњем левом углу Variable View (преглед променљивих), задужено за дефинисање променљивих тачније за номенклатуру (Name) и дефинисање типа који може бити различите нотације. Типови променљивих могу бити: нумерички (Numeric), бројеви одвојени зарезом или тачком (Comma, Dot), бројеви са научном нотацијом (Scientific notation) односно бројеви приказани у експоненцијалном облику који замењује део броја облика $E+n$ где E, које означава експонент, множи претходни број са *10* на n-ти степен, на пример број 12345678901 у научној нотацији приказан је као *1,23E+10* што је *1,23* пута *10* на *10*-ти степен. Постоје такође и подаци који приказују датум (Date) тј. редне бројеве времена и датума, променљиве изражене у одређеним валутама (Dollar, Custom currency), као и променљиве знаковног типа (String) чији се избор може видети у прозору Variable Type (врста променљиве) такође на слици 2.6.2. У простору одређеном за дефинисање и преглед променљивих може се дефинисати и број децималних места, као и ознака за различите променљиве уколико су оне категоричког типа. На слици 2.6.1 приказан је преглед самих података (Data View) који је уједно и примарно окружење које се појављује покретањем SPSS програма. Слике 2.6.1и 2.6.2 односе се на податке из примера 2.6.1. Примећује се да се подаци за обе компаније уносе у првој (једној истој) колони док се у другој колони врши њихово разврставање по узорцима. У примеру 2.6.1 за податке прве компаније ознака је А док је за податке друге компаније ознака Б. Такав принцип уношења и разврставања података захтеван је од функција за коришћење Студентовог t-теста за независне узорке. Различите статистичке анализе захтевају различит начин уноса података што ће бити објашњено уз сваки приказани тест у даљем тексту.



Слика 2.6.1

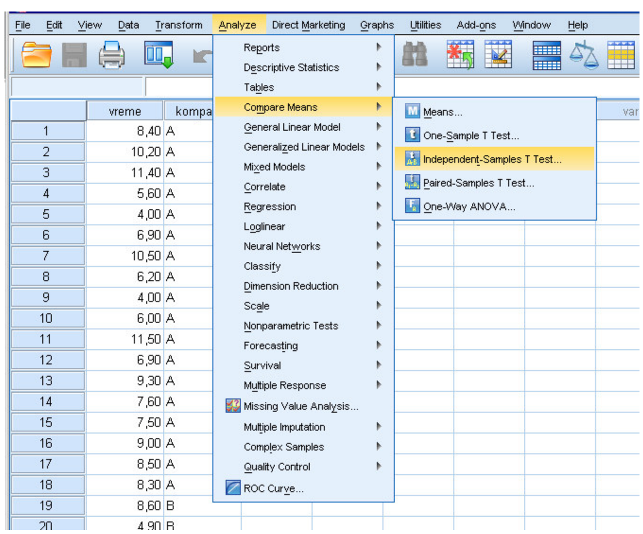


Слика 2.6

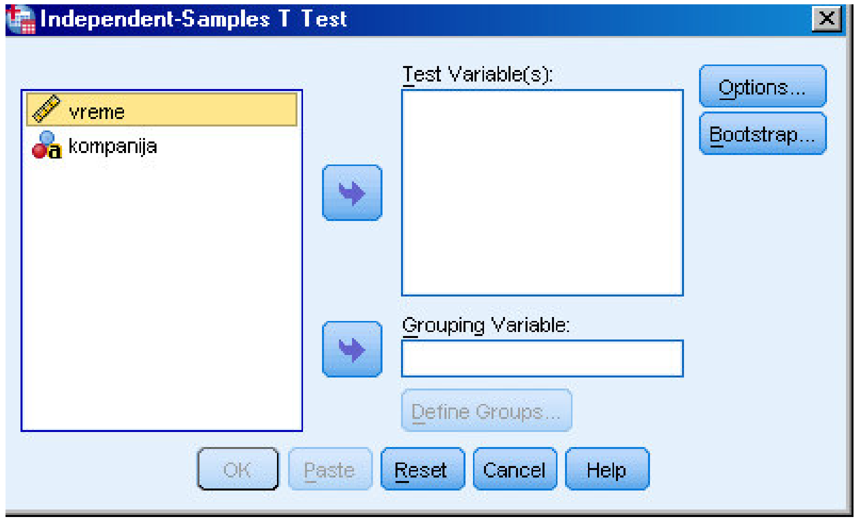
На слици 2.6.3 приказано је како у SPSS-у долазимо до t-теста за независне узорке :

Analyze > Comapare Means > Independent samples t-test. Након што одаберемо t-тест за независне узорке појављује се прозор на слици 2.6.4 где са леве стране треба да одаберемо променљиву коју желимо да тестирама и кликом на стрелицу пребацимо у прозор са десне стране (Test Variable (s)). Примећујемо да је могуће, истовремено тестирати, више променљивих које су категорисане на исти начин. Променљиву која дефинише категорије на исти начин пребацујемо у прозор (Grouping Variable) и дефинишемо груписане променљиве ознакама које смо користили и при уносу података као на слици 2.6.5.

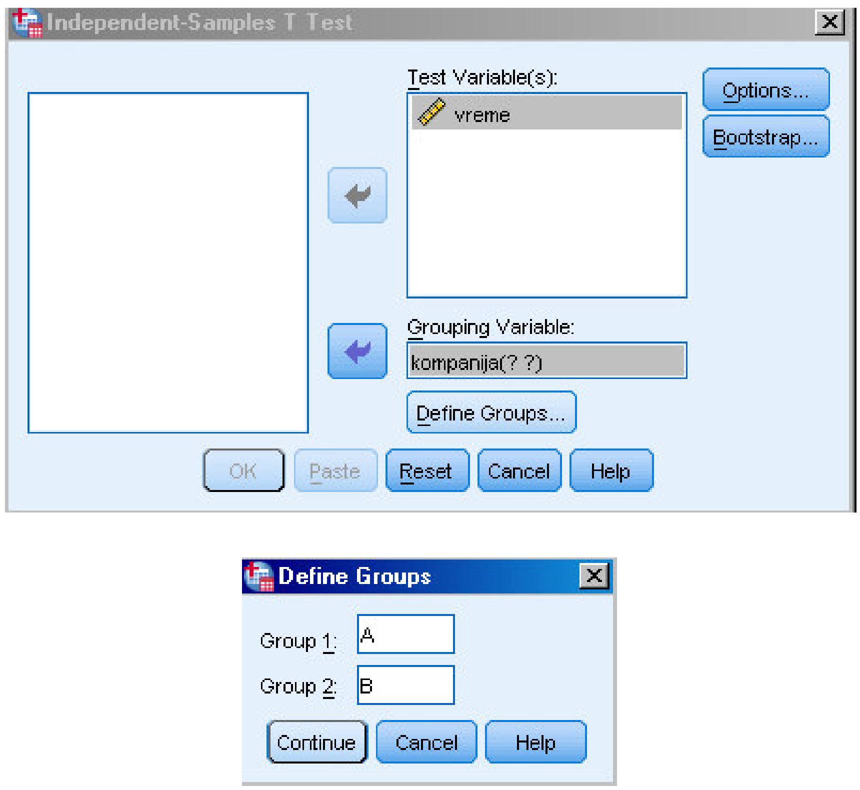
Резултат теста из примера 2.6.1 у SPSS извештају изгледа као на слици 2.6.6, одакле истраживач јасно може да прочита све параметре теста, неопходно је само знати правилно протумачити извештај



Слика 2.6.3



Слика 2.6.4



Слика 2.6.5

Group Statistics

	kompanija	N	Mean	Std. Deviation	Std. Error Mean
vreme	A	18	7,8778	2,24462	,52906
	B	12	8,0750	1,45735	,42070

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
vreme	Equal variances assumed	2,731	,110	-,268	28	,791	-,19722	,73535	-1,70352	1,30908
	Equal variances not assumed			-,292	27,996	,773	-,19722	,67594	-1,58183	1,18738

Слика 2.6.6

Тумачење:

У првој табели на слици 2.6.6 приказани су основни статистички параметри за оба узорка из примера 2.6.1 као што су (N – обим узорка, Mean- узорачка средина, Std. Deviation- стандардна девијација и Std. Error Mean- стандардна грешка- Std. Deviation/√n) за оба узорка. У доњој табели представљени су параметри два теста. Први је Ливенов тест о једнакости дисперзије за који је дата F статистика као и значајност теста која је у овом примеру 0.110 што значи да нема статистички значајне разлике између дисперзија у узорцима и у овом случају параметре t теста гледамо у првом реду (Equal variances assumed). Други тест је t тест за независне узорке (t – test for Equality of Means). У другом реду доње таблице приказани су параметри t теста који нам служе за тумачење: тест статистика (t), број степени слободе (Degrees of Freedom – у извештају df), статистичка значајност-двострана (Statitical Significance – у извештају Sig. (2-tailed)), разлика између средњих вредности узорака (Mean Difference), разлика између стандардних грешака узорака (Standard Error Difference -$[(S_{n1}^2/n_1) + (S_{n2}^2/n_2)]$*) као и доња и горња граница 95% интервала поверења за разлику између средњих вредности узорака (95% Confidence Interval of the Difference – Lower and Upper) респективно. У примеру 2.6.1, статистичка значајност је 0,791- Sig. (2-tailed), на основу које закључујемо да прихватамо нулту хипотезу тј. да нема статистички значајне разлике у деловању таблета две различите компаније.*

2.7 Студентов *t*-тест за два зависна узорка (тест парова)

Студентов *t*-тест за два зависна узорка се у биомедицини примењује када процењујемо неки фактор утицаја (време, лекови, разне врсте терапија и сл.). У тим случајевима се различити узорци обрађују пре и после деловања испитиваних фактора, а значајност утицаја тог фактора се оцењује израчунавањем вредности тест статистике.

Нека је $((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$ узорак обима n, и нека су елементи парова X_i и Y_i, i=1,2,..., зависни, док су парови међусобно независни.

Нека X и Y имају редом нормалне расподеле: $N(m_1, \sigma_1^2)$ и $N(m_2, \sigma_2^2)$. Тада и случајна величина $D = X-Y$ такође има нормалну расподелу $D\ (m_D, \sigma_D^2)$ где су математичко очекивање и дисперзија обележја D непознате. Ако желимо да тестирамо нулту хипотезу $H_0\ (m_D = m_0)$ са нивоом значајности α, тест статистика је једнака: $T = \dfrac{\bar{D}n - m_0}{\tilde{S}_D}\sqrt{n}$, где је $\bar{D}_n = \dfrac{1}{n}\sum_{i=1}^{n} D_i$ узорачка средина, и

$$\tilde{S}_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D}_n)^2$$

је узорачка дисперзија за $D_i = X_i - Y_i,\ i=1,2,\ldots n.$

Како важи:

$$\bar{D}_n \sim N\left(m_{D,}\ \frac{\sigma_D^2}{n}\right),\ X^* = \frac{\bar{D}_n - m_D}{\frac{\sigma_D}{\sqrt{n}}} \sim N(0,1)$$

и како $W = \dfrac{n-1}{\sigma_D^2}\tilde{S}_D^2 : \chi_{n-1}^2$

после краћег рачуна долазимо до једнакости

$$\frac{X^*}{\sqrt{\frac{W}{n-1}}} = \frac{\overline{D}_n - m_D}{\tilde{S}_D}\sqrt{n} \sim t_{n-1}.$$

Ако је тачна нулта хипотеза, тест статистика Т има Студентову расподелу са *n-1* степени слободе.

Критичне области са нивоом значајности α за тестирање нулте хипотезе $H_0(m_1=m_2)$ против једне од алтернативних:

$$H_1(m_1 \neq m_2), \quad H_1(m_1 > m_2), \quad H_1(m_1 < m_2)$$

су редом:

$$|T| \geq t_{n-1,1-\frac{\alpha}{2}},\ T \geq t_{n-1,1-\alpha},\ T \leq t_{n-1,\alpha}.$$

Пример 2.7.1

Групи од 13 пацијената дат је један стимуланс. Испитаћемо да ли постоји статистички значајна разлика у вредности пулса пре и после давања стимуланса са нивоом значајности $\alpha=0,01$.

Пре: 65 72 58 69 79 70 85 67 78 83 80 67 60

После: 71 80 65 67 77 71 82 65 70 90 78 70 67

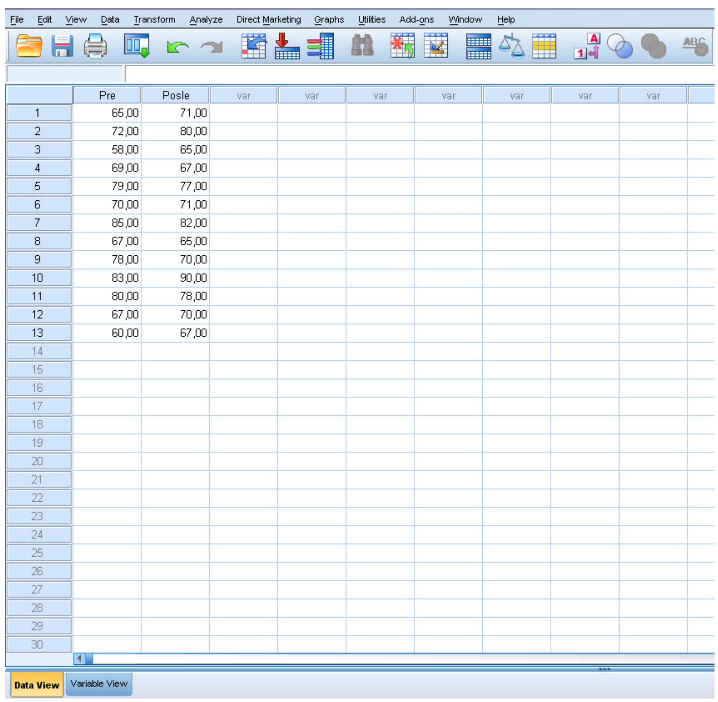
$H_0 : \mu_{\text{пре стимулације}} = \mu_{\text{после стимулације}}$ и $H_1 : \mu_{\text{пре стимулације}} \neq \mu_{\text{после стимулације}}$

df = 13-1=112, α=0,01

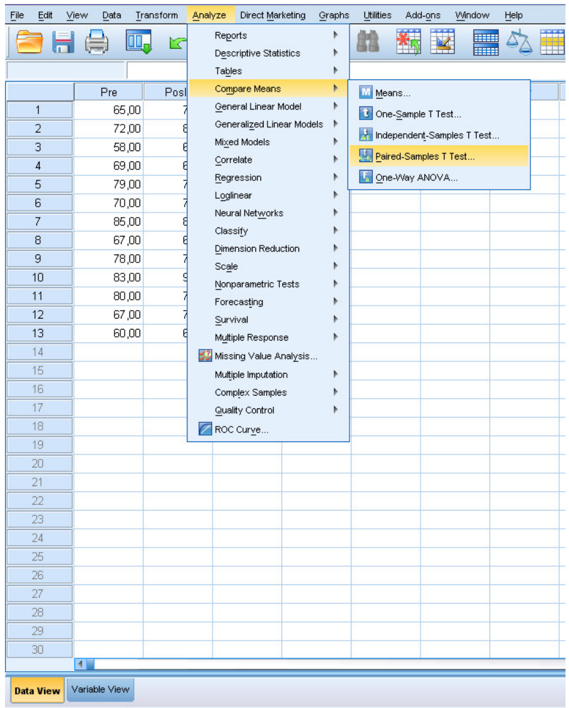
$$\overline{D}_n = \frac{20}{13} = 1.53,\ \tilde{S}_D = \sqrt{\frac{1}{12}\cdot 315,77} = 5,12,\ t = \frac{1.53}{\frac{5.12}{\sqrt{13}}} = 1.082$$

према таблици $t_{0,01,\ 12} = 2.681$. Како је израчуната вредност t мања од табличне за број степени слободе 13 и ниво значајности $0,01$. Закључак је да се нулта хипотеза прихвата тј. да не постоји значајна статистичка разлика вредности пулса пре и после давања стимуланса за дати узорак и дати ниво значајности.

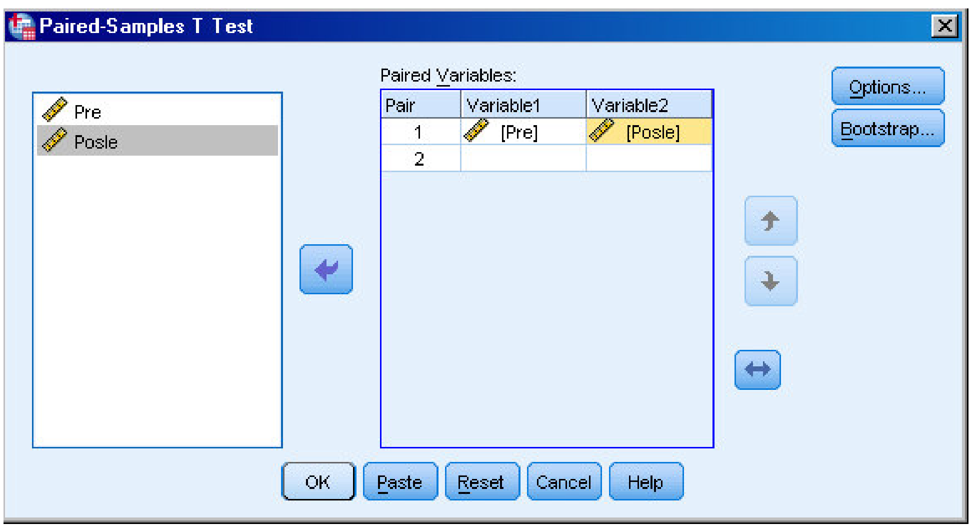
За t-тест парова, захтева се нешто другачији унос података од оног у претходном примеру. Подаци се за свако мерење уносе у различитим колонама, као што је то за пример 2.7. 1 приказано на слици 2.7.1. До t-теста за зависне узорке долазимо истим путем као у примеру 2.6.1 само што у последњем падајућем подпрозору бирамо опцију Paired-Samples T test (слика 2.7.2). Прозор за одабир променљивих које тестирамо захтева уношење једног или више парова променљивих које тестирамо које пребацујемо са леве на десну стану уз помоћ стрелице као у претходном примеру (слика 2.7.3). Кликом на ОК добијамо извештај као на слици 2.7.4.



Слика 2.7.1



Слика 2.7.2



Слика 2.7.3

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	VAR00001	71,7692	13	8,61350	2,38895
	VAR00002	73,3077	13	7,55408	2,09512

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	VAR00001 & VAR00002	13	,807	,001

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	VAR00001 - VAR00002	-1,5385	5,12535	1,42152	-4,6357	1,5588	-1,082	12	,300

Слика 2.7.4

Тумачење:

На слици 2.7.1 приказан је SPSS извештај који се односи на пример 2.7.1. У првој табели приказана је дескриптивна статистика података за оба узорка, као што су средње вредности узорака, број елемената у узорцима, стандардне девијације и стандардне грешке узорака. Друга табела представља корелацију (ниво дирктне или индиректне, у овом случају директне, повезаности) два зависна узорка, односно коефицијент корелације од 0,807 (Correlation) који указује на значајну корелацију између ова два узорка што потврђује и статистичка значајност од 0,001 (Sig.).Трећа табела односи се на сам t тест парова као и на вредности које описују разлике пара зависних променљивих (Paired Differences). Поред средње вредности, стандардне девијације и стандардне грешке разлике између парова приказан је интервал поверења разлике парова. Последње три колоне у табели представљају t вредност, број степени слободе као и двострану статистичку значајност од 0,300 (Sig.(2-tailed)) на основу које и доносимо закључак о прихватању нулте хипотезе тј. закључујемо да не постоји статистички значајна разлика (са нивоом значајности 0,01) у вредности пулса пре и после давања стимуланса.

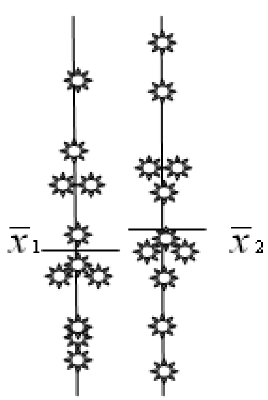
3. Утицаји различитих фактора на тестирање хипотезе Студентовим t-тестом

На крајњи закључак код тестирања хипотезе t тестом утичу:

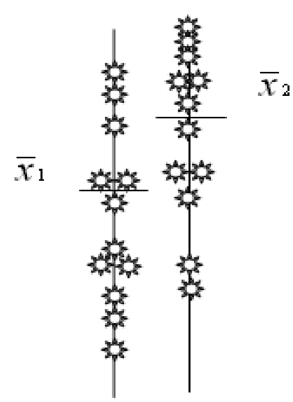
- величина разлике између средње вредности узорка и средње вредности популације, односно две средње вредности
- величина узорка
- варијација у узорку (величина стандарде девијације)

3.1 Разлика између средњих вредности

Велика разлика између средњих вредности даје велику вредност t (већу од критичне вредности за одабрани ниво значајности), што значи да се са великом разликом у мањем броју случајева долази до закључка о прихватању нулте хипотезе . Ако се вратимо на изразе за тест статистике Студентовог t-теста за зависне и независне узорке, приметићемо да се са смањењем разлике између средњих вредности смањује и t вредност што директно повећава шансе за прихватањем нулте хипотезе за одабрани ниво значајности.



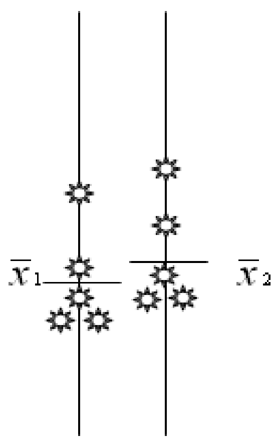
Мала разлика између група – веће шансе за прихватањем нулте хипотезе



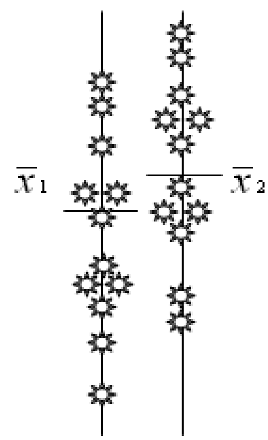
Велика разлика између група – мање шансе за прихватањем нулте хипотезе

3.2 Величина узорка

Величина узорка такође утиче на крајњи закључак код тестирања хипотезе. При тестирању већих узорка, у мањем броју случајева се долази до закључка о прихватању нулте хипотезе за одабрани ниво значајности, јер велики узорак даје велику вредност t (већу од критичне вредности за изабрани ниво значајности). За тест статистике Студентовог t-теста за зависне и независне узорке такође примећујемо да је t вредност директно пропорционална величини узорка/ака.



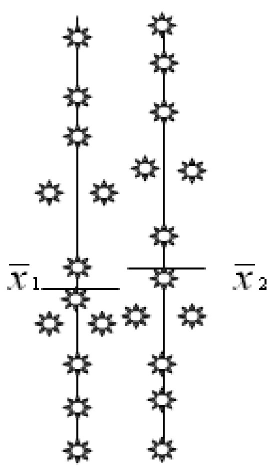
Мали број података у групи – веће шансе за прихватањем нулте



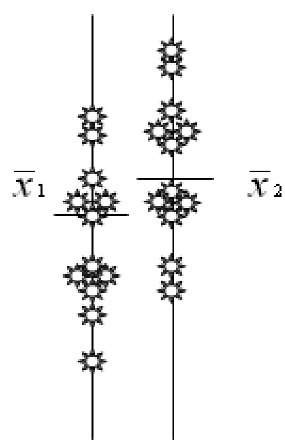
Велики број података у групи – мање шансе за прихватањем нулте хипотезе

3.3 Величина стандардне девијације у узорку

Велика стандардна девијација (велика варијација) у узорку даје малу вредност t (мању од критичне вредности за изабрани ниво значајности), што значи да се са великом стандардном девијацијом, у већем броју случајева, долази до закључка о прихватању нулте хипотезе за одабрани ниво значајности. Обрнута пропорционалност између стандардне девијације и t вредности приказана је такође у формулама за тест статистике Студентовог t-теста за зависне и независне узорке.



Велика варијација унутар група – веће шансе за прихватањем нулте хипотезе



Мала варијација унутар група – мање шансе за прихватањем нулте хипотезе

4. Непараметарски тестови

4.1 Ман - Витни U тест (непараметарски тест за независне узорке)

До сада приказане методе за упоређивање две групе података заснивају се на претпоставци да су подаци нормално расподељени, и увек се у њима појављује потреба за оцењивањем појединих параметара (средње вредности, стандардне девијације). Међутим, када не можемо са сигурношћу тврдити да је расподела једне групе података нормална, израчунавање појединих параметара и примена параметарских метода дају врло непоуздане закључке. У тим случајевима се примењују непараметарске методе. Ман - Витни[3] U тест је на неки начин еквивалентан параметарском t тесту јер се њиме такође пореде две групе података, али преко медијана. Нулта и алтернативна хипотеза гласе:

$$H_0: Me_1 = Me_2 \qquad H_1: Me_1 \neq Me_2 \; (Me_1 > Me_2)$$

Принцип поређења се састоји у томе да се вредности група уреде по величини, а затим обележе редним бројевима (ранговима) посматрајући обе групе истовремено. Ако се вредности у групама разликују, онда ће се у једној групи налазити већи а у другој мањи рангови (јер се у том случају у једној групи налазе претежно веће а у другој претежно мање вредности), па се самим тим и збирови рангова разликују. Насупрот овом случају, ако се вредности у две групе не разликују, онда су рангови равномерно распоређени, тачније у обе групе се налазе и мањи и већи рангови, односно суме рангова су приближно једнаке. Дакле, код непараметарских тестова, сами подаци из група не служе за поређење, већ нам за то служе њихови рангови или суме њихових рангова у појединим групама. Међутим, главни недостатак непараметарских тестова је да често долази до губитка информација из узорка, па они могу бити мање ефикасни него параметарске методе.

[3] Henry Berthold Mann (1905-2000), амерички статистичар, Donald Ransom Whitney (1915-2001), амерички статистичар

Изрази којима се упоређују суме рангова зависе од броја података у групама. Овај тест може да се користи и када групе садрже врло мали број података (мање од 9), али су закључци који се том приликом добијају непоуздани, јер су изведени из малог броја података.

Ман Витни U тест за $n_1 \geq 9$ и $n_2 \leq 20$

Израз за израчунавање статистике U за Ман Витни тест је:

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1$$

Односно,

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2$$

где је :

$\sum R_1$ - сума рангова у првом узорку обележја X обима n_1

$\sum R_2$ - сума рангова у узорку обележја Y обима n_2.

Однос између ове две вредности је исказан формулом:

$$U_1 = n_1 n_2 - U_2$$

Овај израз служи и за једноставније израчунавање, тако да уместо израчунавања обе вредности U_1 і U_2 према горе наведеним изразима довољно је израчунати само једну, а друга се израчунава из наведеног односа.

Мања од добијених вредности U_1 и U_2 упоређује са вредношћу u_α из табеле за критичне вредности U Ман Витни теста (у прилогу) за жељени ниво вероватноће и одговарајући број података и ако је мања од критичне вредности можемо тврдити да је разлика између две групе статистички значајна. Статистички заначајна разлика показује да посматране групе не припадају истој популацији.

Дакле, ако је критична вредност са нивоом значајности α за тестирање нулте хипотезе $H_0(Me_1 = Me_2)$ против једне од алтернативних $H_1(Me_1 < Me_2)$, $H_1(Me_1 > Me_2)$, $H_1(Me_1 \neq Me_2)$ мања од израчунате вредности мин (U_1, U_2) прихвата се нулта хипотеза, док се у супротном прихвата алтернативна хипотеза.

Пример 4.1.1:

Код 16 здравих особа и код 15 особа које су биле у акутном нападу инфаркта миокарда одређена је концентрација арахидонске киселине методом течне хроматографије. Показати да ли између добијених вредности постоји значајна разлика, узимајући у обзир ниво вероватноће 0,05 (резултати су изражени у μmol/L)

Рангирање, односно обележавање вредности редним бројевима врши се посматрањем обе групе истовремено. Најнижа вредност у обе групе је 1, и она добија редни број 1, следећа је 1,4 и она добија редни број 2 итд. Вредност 1,8 се појављује два пута и она ће у оба случаја имати редни број 5,5 која представља средњу вредност између бројева 5 и 6.

У случају да се неки број појављује више пута примењује се такозвано дељење рангова тј његов редни број се израчунава као средња вредност редних бројева које оне заузимају. У нашем примеру вредност 1,8 се појавила два пута и заузела редне бројеве 5 и 6 тако да следећа вредност по величини 2, заузима редни број 7. Нумерисање се тако изводи до последње највеће вредности, и ако није направљена грешка, последња највећа вредност имаће редни број који је једнак збиру у обе групе. У нашем примеру је то 31.

Табела 4.1.1

Група 1	Р.бр. (R1)	Група 2	Р.бр. (R2)
12,4	30	1,8	5,5
2,1	8,5	1,7	4
8,7	29	3,7	20
7,4	27	2	7
5,2	23,5	1,6	3
4,9	21,5	2,9	14,5
2,5	11	1,8	5,5
2,9	14,5	1	1
2,8	12,5	4,9	21,5
3,4	18,5	5,2	23,5
8,1	28	2,8	12,5
6	25	13,7	31
3	16	3,2	17
1,4	2	2,1	8,5
2,2	10	3,4	18,5
		6,8	26
Укупно	$\sum$ 277		$\sum$ 219

$$U_1 = 15 \cdot 16 + \frac{15 \cdot (15+1)}{2} - 277 = 83$$

$$U_2 = 15 \cdot 16 - 83 = 157$$

Критична вредност *U* за вероватноћу $p=0.05$ и број података $n_1= 15$ *и* $n_2=16$ износи 70, а како је израчуната вредност $U_1 = 83$, односно већа од критичне, значи да разлика између ове две групе пацијената није статистички значајна.

У SPSS-у није није предвиђена могућност тестирања Ман-Витнијевим тестом за узорке малог обима па тако за Ман Витни тест посматраног обима у нашем примеру SPSS даје вредности за Z, уместо за U, што у овом примеру не прави разлику, али у неком другом би могло доћи до грешке. На слици 4.1.1 приказан је SPSS извештај из примера 4.1.1.

NPar Tests

Mann-Whitney Test

Ranks

	GRUPE	N	Mean Rank	Sum of Ranks
KISELINA	1,00	15	18,47	277,00
	2,00	16	13,69	219,00
	Total	31		

Test Statistics[b]

	KISELINA
Mann-Whitney U	83,000
Wilcoxon W	219,000
Z	-1,464
Asymp. Sig. (2-tailed)	,143
Exact Sig. [2*(1-tailed Sig.)]	,151[a]

a. Not corrected for ties.

b. Grouping Variable: GRUPE

Слика 4.1.1

Тумачење:

На слици 4.1.1 приказан је SPSS извештај који се односи на пример 4.1.1 и непараметарски Ман Витни U test. У првој табели (Ranks / рангови) приказани су дескриптивни подаци о групама, величини узорака (n), средњој вредности редних бројева односно рангова (Mean Rank) као и суме рангова за оба узорка (Sum of Ranks). У другој табели (Test statistics /Тест статистике) приказана је вредност U за Mann-Whitney, као и вредност W за Вилкоксон тест, z вредност као и асимтотска двострана статистичка значајност (Asimp. Sig.(2-tailed)) на основу које и доносимо закључак прихватању нулте хипотезе јер је израчуната статистичка значајност већа од 0,05. Такође је приказана и одређена статистичка значајност непоправљена за вредности који се поклапају - (Exact Sig.(2(1-tailed Sig.)), not corrected for ties)*

Ман Витни U тест за $n_1 \geq 9$ и $n_2 \geq 20$

Када се број података у групама повећава, расподела вредности U се врло брзо приближава нормалној расподели $N(n_1n_2/2,\ n_1n_2(n_1+n_2+1))$, тако да се значајност израчунате вредности U може оценити помоћу израза:

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Вредност z има нормалну раподелу са средњом вредношћу нула и стандардном девијацијом 1, а значајност израчунате вредности z се упоређује са вредностима из табеле нормалне расподеле за жељени ниво значајности.

Вредност U се израчунава према раније наведеним изразима, а у израз за z може да се уврсти U_1 и U_2 . Од тога да ли ће се уврстити U_1 или U_2 у израз за z зависи само знак добијеног резултата, али не и његова апсолутна вредност.

Ако је критична вредност са нивоом значајности α за тестирање нулте хипотезе $H_0(Me_1 = Me_2)$ против једне од алтернативних $H_1(Me_1 < Me_2)$, $H_1(Me_1 > Me_2)$, $H_1(Me_1 \neq Me_2)$ већа од израчунате вредности z_α прихватамо нулту хипотезу, у супротном прихватамо алтернативну хипотезу.

Пример 4.1.2

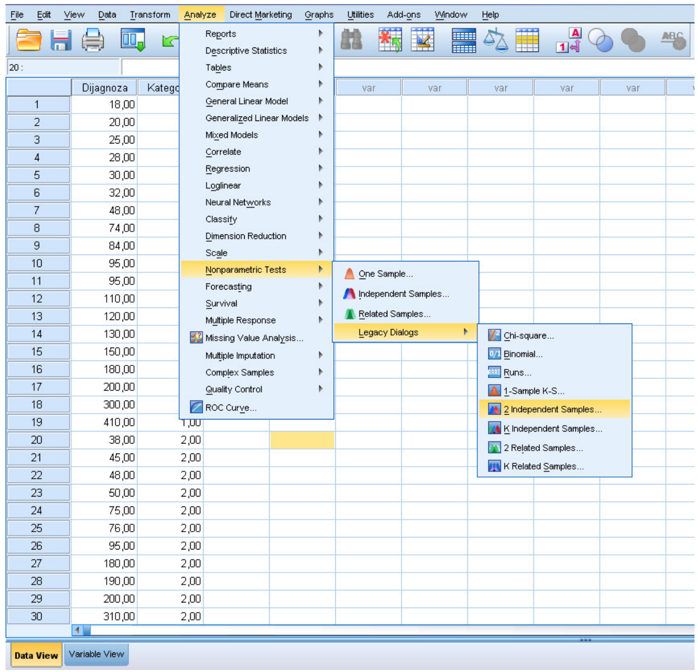
Показати да ли се значајно разликују вредности АЛТ код пацијената са цирозом јетре (N_1=19) и код пацијената са хроничним хепатитисом (N_2=21).

Група 1 (цироза)	Р.бр. (R1)	Група 2 (хепатитис)	Р.бр. (R2)
18	1	38	7
20	2	45	8
25	3	48	9,50
28	4	50	11
30	5	75	13
32	6	76	14
48	9,50	95	17
74	12	180	23,5
84	15	190	25
95	17	200	26,5
95	17	310	29
110	19	340	30
120	20	380	31
130	21	390	32
150	22	420	34
180	23	435	35
200	26	450	36
300	28	470	37
410	33	520	38
		720	39
		790	40
	$\sum R_1$=284,5		$\sum R_2$=535,5

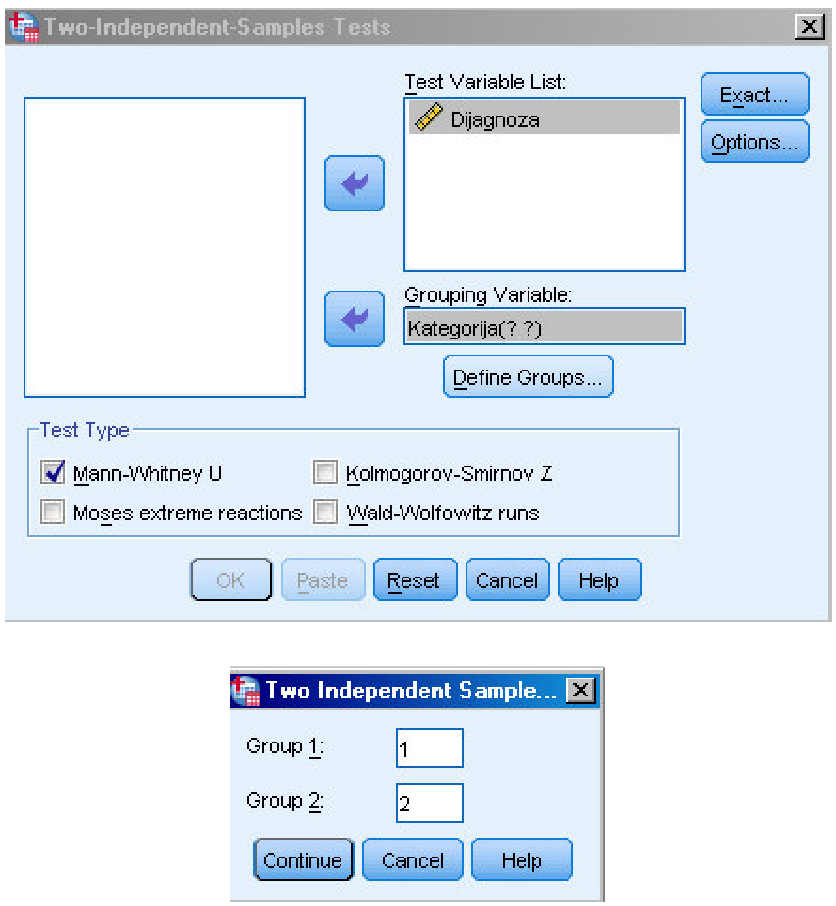
Најнижа вредност у скупу је 18 и она носи редни број 1, следећа је 20 са редним бројем 2 и тако све до вредности 48 која се појављује два пута. До вредности 48 заузето је 8 места, односно осам редних бројева, тако да се вредности 48 распоређују на места 9 и 10 и носе редни број 9,5. Следећа вредност је 50 са редним бројем 11 (9 и 10 су већ искоришћени) и скуп се даље рангира до вредности 95 која се појављује три пута (два пута у групи један и једанпут у групи два). До вредности 95 рангирано је 15 вредности, тако да три вредности 95 треба да заузму редна места 15, 17 и 18. Њих ћемо обележити средњом вредношћу ова три редна броја, а то је 17. Следећа вредност у скупу је 110 и она носи редни број 19 (јер су 16, 17 и 18 искоришћени). Рангирање вредности у скупу настављамо до последње, највеће вредности 790 која има редни број 40, што одговара збиру података у обе групе.

Користећи формуле за U_1 и U_2 добијамо да је U = 304,5 а користећи формулу за z добијамо да је z = 2,844. Ако се у израз за z уврсти U_2 добиће се резултат са истом апсолутном вредношћу, али супротног знака, што не утиче на тумачење значајности разлике. Израчуната вредност z је већа од критичне 1,96 која се добија из табеле нормалне расподеле за вероватноћу p = 0,05, што говори да је разлика између ове две групе статистички значајна, што се може уочити и из SPSS извештаја са слике 4.1.3.

Предност Ман Витни U теста у односу на Студентов t-тест је у томе што на крајњи резултат не утиче расипање резултата унутар групе, што што се види у примеру 4.1.2. Обе посматране групе имају велику стандардну девијацију јер је разлика између најниже и највише вредности велика, што би се одразило на вредност t, која би због тога била мања од критичне вредности за одговарајући број степени слободе. Како је Ман Витни тест непараметарски пандан Студентовом t-тесту за независне узорке то је и унос података исти као што је описано у примеру 2.6.1. На слици 4.1.2 приказано је како стижемо до Ман Витни теста. Уместо Compare means, сада бирамо Nonparametric tests > Legacy Dialogs > 2 Independent Samples. Задавање променљивих је такође слично као код Студентовог t-теста за независне узорке само је још неопходно означити поље именом непараметарког теста који желимо да урадимо (Слика 4.1.3). На слици 4.1.4 приказан је СПСС извештај за пример 4.1.2.



Слика 4.1.2



Слика 4.1.3

Mann-Whitney Test

Ranks

	Kategorija	N	Mean Rank	Sum of Ranks
Dijagnoza	1,00	19	14,97	284,50
	2,00	21	25,50	535,50
	Total	40		

Test Statistics[b]

	Dijagnoza
Mann-Whitney U	94,500
Wilcoxon W	284,500
Z	-2,845
Asymp. Sig. (2-tailed)	,004
Exact Sig. [2*(1-tailed Sig.)]	,004[a]

a. Not corrected for ties.
b. Grouping Variable: Kategorija

Слика 4.1.4

Тумачење: *На слици 4.1.4 је приказан SPSS извештај који се односи на пример 4.1.2 односно непараметарски Ман Витни U тест са бројем узорака* $n_1 \geq 9$ *и* $n_2 \geq 20$. *У првој табели (Ranks / рангови) приказани су дескриптивни подаци о групама, величини узорака (n), средњој вредности редних бројева односно рангова (Mean Rank) као и суме рангова за оба узорка (Sum of Ranks). У другој табели (Test statistics /Тест статистике) приказана је вредност U за Mann-Whitney, као и вредност W за Вилкоксонов тест, z вредност као и асимтотска двострана статистичка значајност (Asimp. Sig.(2-tailed)) на основу које и доносимо закључак о одбацивању нулте хипотезе, у овом случају јер је израчуната статистичка значајност 0,004. Такође је приказана и одређена статистичка значајност непоправљена за вредности које се поклапају - (Exact Sig.(2*(1-tailed Sig.)), not corrected for ties)*

4.2 Вилкоксонов[4] тест означених рангова (непараметарски тест за зависне узорке)

Вилкоксонов тест означених рангова дизајниран је и намењен за анализу података добијених из поновљених мерења, тачније када се једна величина мери под различитим условима, или у два различита временска интервала (мери се на пример унос витамина Ц у летњим па поново у зимским месецима код истих пацијената). Овај тест заправо представља непараметарску алтернативу за параметарски t-тест парова. За разлику од Ман Витни U теста не приступамо рангирању података одмах, већ рачунамо разлику између прве и друге групе података које желимо да поредимо и разлици додељујемо редне бројеве. Том приликом узимамо у обзир знак разлике који ће касније приликом рангирања понети и вредности редног броја. Ако је разлика једнака нули ту вредност искључујемо из тестирања. Сходно томе, суме рангова одвајамо на позитивне и негативне те тако добијамо две суме рангова T_+ (сума позитивних рангова) и T_- (сума негативних рангова). Мања од ове две вредности се користи за поређење. Да бисмо израчунали статистичку значајност вредности T користићемо вредности $\bar{T}$ и SE_T које су у зависности само од броја субјеката и израчунавају се помоћу израза:

$$\bar{T} = \frac{n(n+1)}{4}$$

$$SE_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Као и код Ман Витни U теста, сада лако можемо израчунати вредност z помоћу

израза: $$Z = \frac{\min(T_{+,} T_-) - \bar{T}}{SE_T}$$

[4] Frank Wilcoxon (1892-1965), амерички статистичар и хемичар

Ако је критична вредност са нивоом значајности α за тестирање нулте хипотезе $H_0(Me_1 = Me_2)$ против једне од алтернативних $H_1(Me_1 < Me_2)$, $H_1(Me_1 > Me_2)$, $H_1(Me_1 \neq Me_2)$ већа од израчунате вредности z_α прихвата се нулту хипотезу, у супратном прихвата се алтернативна хипотеза.

Пример 4.2.1

Мерен је степен депресије код зависника од алкохола и зависника од пушења прво у недељу па затим поново у среду. Испитујемо да ли постоји статистички значајна разлика у нивоу депресије између недеље и среде.

Табела 4.2.1

Недеља	Среда	Разлика	Знак разлике	Р. Бр	Позитиван Р. бр.	Негативан Р. бр.
Цигарете						
15	28	13	+	2,5	2,5	
35	35	0	Одбацује се			
16	35	19	+	6	6	
18	24	6	+	1	1	
19	39	20	+	7	7	
17	32	15	+	4,5	4,5	
27	27	0	Одбацује се			
16	29	13	+	2,5	2,5	
13	36	23	+	8	8	
20	35	15	+	4,5	4,5	
				Укупно =	36	0
Алкохол						
16	5	-11	-	9		9
15	6	-9	-	7		7
20	30	10	+	8	8	
15	8	-7	-	3,5		3,5
16	9	-7	-	3,5		3,5
13	7	-6	-	2		2
14	6	-8	-	5,5		5,5
19	17	-2	-	1		1
18	3	-15	-	10		10
18	10	-8	-	5,5		5,5
				Укупно =	8	47

Из табеле 4.2.1 видимо да је сума позитивних редних бројева за цигарете T_{+}=36 док је сума негативних редних бројева T_{-}= 0, тачније нама негативних редних бројева јер су и све вредности разлике позитивне. За алкохол T_{+}= 8 док је T_{-}= 47. Како за вредност T бирамо мању од две добијене, користићемо T_{-}= 0 за цигарете и T_{+}= 8 за алкохол.

Користећи формуле за $\bar{T}$ и SE израчунавамо:

$\bar{T}_{цигарете}=18$, $SE_{T цигарете}=7.14$;

$\bar{T}_{алкохол}=27.50$, $SE_{T алкохол}=9.81$,

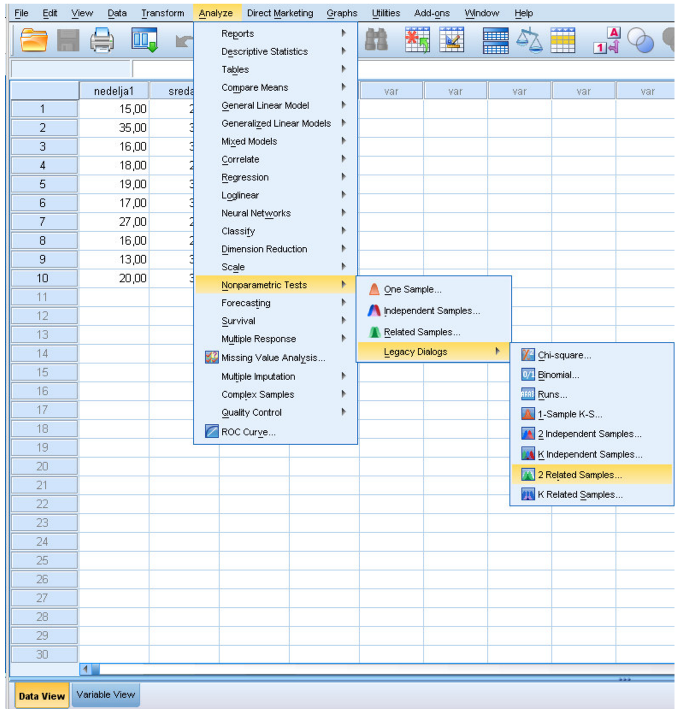
и на крају, користећи формулу за вредности z која нам омогућава да оценимо статистичку значајност разлике добијамо:

$z_{цигарете} = -2.52$;

$z_{алкохол} = -1.99$.

Како су обе апсолутне вредности z веће од 1.96 долазимо до закључка да постоји статистички значајна разлика у нивоу депресије између недеље и среде и код једне и код друге врсте зависника.

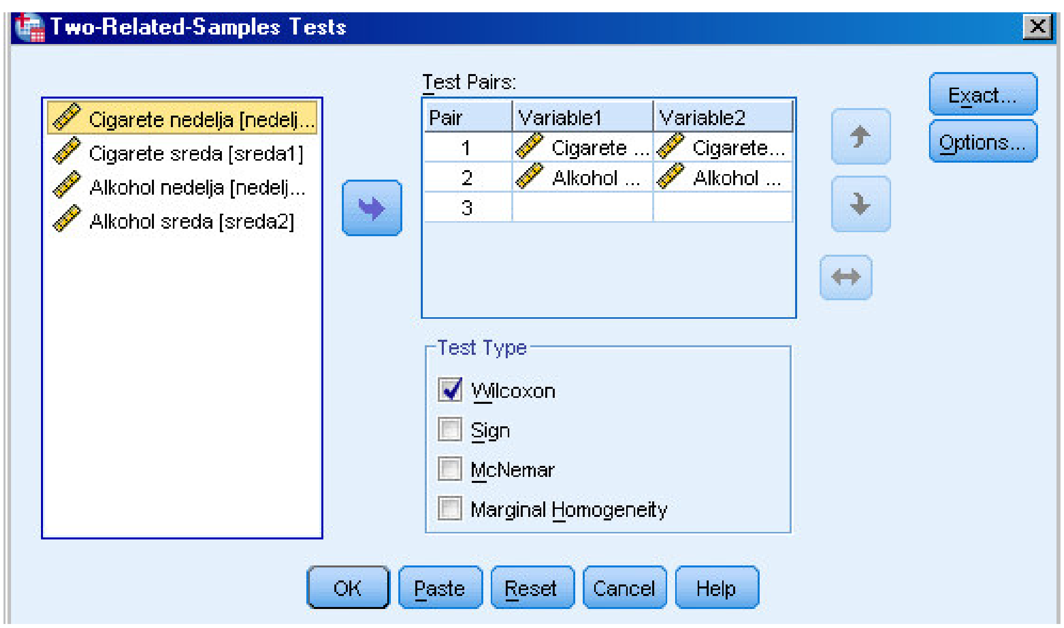
Унос података за Вилкоксонов тест је идентичан као за Студентов t-тест парова и описан је у примеру 2.7.1 док се до Вилкоксоновог теста долази кликом на Analize > Nonparametric Tests > Legacy Dialogs > 2 Realted Samples (Слика 4.2.1).



Слика 4.2.1

На слици 4.2.2 приказан је прозор који такође подсећа на слику 2.7.2 с том разликом што и овде морамо назначити који непараметарски тест желимо да користимо.

SPSS извештај за пример 4.2.1 приказан је на слици 4.2.3.



Слика 4.2.2

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Cigarete sreda - Cigarete nedelja	Negative Ranks	0[a]	,00	,00
	Positive Ranks	8[b]	4,50	36,00
	Ties	2[c]		
	Total	10		
Alkohol sreda - Alkohol nedelja	Negative Ranks	9[d]	5,22	47,00
	Positive Ranks	1[e]	8,00	8,00
	Ties	0[f]		
	Total	10		

a. Cigarete sreda < Cigarete nedelja
b. Cigarete sreda > Cigarete nedelja
c. Cigarete sreda = Cigarete nedelja
d. Alkohol sreda < Alkohol nedelja
e. Alkohol sreda > Alkohol nedelja
f. Alkohol sreda = Alkohol nedelja

Test Statistics[c]

	Cigarete sreda - Cigarete nedelja	Alkohol sreda - Alkohol nedelja
Z	-2,527[a]	-1,990[b]
Asymp. Sig. (2-tailed)	,012	,047

a. Based on negative ranks.
b. Based on positive ranks.
c. Wilcoxon Signed Ranks Test

Слика 4.2.3

Тумачење: *На слици 4.2.3 приказан је SPSS извештај за непараметарски Вилкоксонов тест означених рангова (Wicoxon Signed Ranks Test) из примера 4.2.1. У првој табели (Ranks/Рангови) представљени су основни подаци о ранговима, тачније број позитивних и негативних рангова, средња вредност рангова (Mean Rank), као и сума рангова (Sum of Ranks). У другој табели (Test Statistics/ Тест Статистика) за Вилкоксонов тест означених рангова базиран на негативним ранговима за цигарете и позитивним ранговима за алкохол приказане су вредности за z као и статистичке значајности на основу којих и доносимо закључке о прихватању нулте хипотезе, односно да постоји разлика у нивоу депресије мерене у среду у односу на недељу.*

Закључак:

Разматрано увођење нових статистичких појмова у план и програм за математику средњих медицинских школа веома би олакшало разумевање статистике и њене примене у каснијем раду и едукацији младих здравствених и научних радника у Србији. Подизање свести младих људи и будућих научника о важности правилног коришћења и тумачења података у циљу унапређења јавног здравља је изузетно значајно и требало би да буде актуелно питање у будућим реформама плана и програма. Приближавајући младим људима једноставне појмове као што су мере централне тенденције и интервал поверења, преко сликовитих примера, открива им се значај и важност информација које се добијају интерпретацијом тих величина и буди интересовање за статистиком уопште.

Када су биомедицинска истраживања и одабир тестова у питању обрађени су најчешће коришћени, и у биомедицини најпознатији, параметарски и непараметарски тестови примењиви на мале узорке (< 30) јер је обично јако тешко и скупо прикупити више узорака људског порекла. Показано је да је кључна ствар за одабир теста и његову правилну примену, сам начин прикупљања података и постављање хипотезе. У колико се провером у SPSS-у утврди нормална расподела података, избор је у већини случајева један од параметарских тестова. Од начина прикупљања података зависи да ли ће се применити t-тест за независне узорке или t-тест парова. Уколико нисмо сигурни у нормалност расподеле избор ће се свести на неки од непараметарских тестова за које се одлучујемо такође у зависности од начина прикупљања података и величине узорка. SPSS је веома лак за коришћење и корисници га прихватају лакше од осталих софтвера на тржишту зато што не захтева додтну обуку за коришћење, бар не већу од оне за Excel. Остали популарни статистички пакети као што су SAS, STATA, S PLUS, R и други захтевају додатно програмерско знање, мада не превише компликовано. Њихова предност је у томе што истраживач, много детаљније, може да задаје команде при обради својих података и да притом дизајнира анализу на много различитије начине него што SPSS то дозвољава. Ови софтвери се масовно користе агенцијама за истраживање тржишта, банкама, осигурањима, великим научно истраживачким центрима итд.

Литература

[1] Andy Field, Discovering statistics using SPSS, SAGE Publications London

[2] Julie Pallant, SPSS Survival Manual. A Step by Step Guide to Data Analysis using SPSS for Windows. Mc Graw Hill Open University Press.

[3] Весна Јевремовић, Вероватноћа и Статистика, Универзитет у Београду, Математички Факултет, Београд, 2009.

[4] Kristina Veljković, Mogućnost Primene Neparametarskih Testova na Testiranje Parametarskih Hipoteza. Univerzitet u Beogradu, Matematički Fakultet,Beograd 2010, Master rad.

[5] M. Erkkola, M. Kavppinen, A. Jarvinen , Folate, Vitamin D and iron intakes are low among pregnant Finnish women. European Journal of Clinical Nutrition (1998) 52, 742-748

[6] Lynn Harbottle and Maureen B. Duggan, Comparative study of the dietary characteristics of Asian toddlers with iron deficiency in Sheffild. Journal of Human Nutrition and Dietetics, The University of Sheffield (1992), 5, 351-361

[7] K. Braekke, A. Cathrine Staf, Periconceptional use of folic acid supplements in Oslo, Acta Obstet Gyneal Scand (2003) 82, 620-627

[8] Douglas G. Altman Practical Statistics for medical Research, Chapman & Hall New York 1999

[9] http://www.medicinskazemun.edu.rs/Udzbenici.htm

[10] Srdjan Ognjanović, Zbirka zadataka za četvrti razred gimnazija i tehničkih škola, Krug 2005

[11] Ljiljana Petrović, Teorijska statistika. Teorija statističkog zaključivanja, Centar za izdavačku delatnost ekonomskog fakulteta, Beograd 2010.

[12] Павле Младеновић, Вероватноћа и статистика, Математички факултет, Београд 2008.

[13] Весна Јевремовић, Јован Малишић, Статистичке методе у метеорологији и инжењерству, Савезни хидрометеоролошки завод, Београд 2002

Прилог

Вредности U за Ман-Витни U тест

Табела 1. Двострано тестирање

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	--	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	--	--	--	--	--	0	0	0	1	1	1	2	2	2	2	3	3
4	.05		0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01		--	--	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05			2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01			0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05				5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01				2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05					8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01					4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05						13	15	17	19	22	24	26	29	31	34	36	38	41
	.01						7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05							17	20	23	26	28	31	34	37	39	42	45	48
	.01							11	13	16	18	20	22	24	27	29	31	33	36
10	.05								23	26	29	33	36	39	42	45	48	52	55
	.01								16	18	21	24	26	29	31	34	37	39	42
11	.05									30	33	37	40	44	47	51	55	58	62
	.01									21	24	27	30	33	36	39	42	45	48
12	.05										37	41	45	49	53	57	61	65	69
	.01										27	31	34	37	41	44	47	51	54
13	.05											45	50	54	59	63	67	72	76
	.01											34	38	42	45	49	53	57	60
14	.05												55	59	64	69	74	78	83
	.01												42	46	50	54	58	63	67
15	.05													64	70	75	80	85	90
	.01													51	55	60	64	69	73
16	.05														75	81	86	92	98
	.01														60	65	70	74	79
17	.05															87	93	99	105
	.01															70	75	81	86
18	.05																99	106	112
	.01																81	87	92
19	.05																	113	119
	.01																	93	99
20	.05																		127
	.01																		105

Табела 2. Једнострано тестирање

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
	.01	--	--	--	--	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	.05		1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	.01		--	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	.05			4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	.01			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	.05				7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	.01				3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	.05					11	13	15	17	19	21	24	26	28	30	33	35	37	39
	.01					6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	.05						15	18	20	23	26	28	31	33	36	39	41	44	47
	.01						9	11	13	15	17	20	22	24	26	28	30	32	34
9	.05							21	24	27	30	33	36	39	42	45	48	51	54
	.01							14	16	18	21	23	26	28	31	33	36	38	40
10	.05								27	31	34	37	41	44	48	51	55	58	62
	.01								19	22	24	27	30	33	36	38	41	44	47
11	.05									34	38	42	46	50	54	57	61	65	69
	.01									25	28	31	34	37	41	44	47	50	53
12	.05										42	47	51	55	60	64	68	72	77
	.01										31	35	38	42	46	49	53	56	60
13	.05											51	56	61	65	70	75	80	84
	.01											39	43	47	51	55	59	63	67
14	.05												61	66	71	77	82	87	92
	.01												47	51	56	60	65	69	73
15	.05													72	77	83	88	94	100
	.01													56	61	66	70	75	80
16	.05														83	89	95	101	107
	.01														66	71	76	82	87
17	.05															96	102	109	115
	.01															77	82	88	93
18	.05																109	116	123
	.01																88	94	100
19	.05																	123	130
	.01																	101	107
20	.05																		138
	.01																		114