УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Кристина Вељковић

МОГУЋНОСТИ ПРИМЕНЕ НЕПАРАМЕТАРСКИХ ТЕСТОВА ЗА ТЕСТИРАЊЕ ПАРАМЕТАРСКИХ ХИПОТЕЗА

МАСТЕР РАД

др Весна Јевремовић, ментор
др Слободанка Јанковић, председник комисије
мр Милан Јовановић, члан комисије

Београд, 2010.

Предговор

У овом раду пажња је посвећена непараметарским тестовима медијане. Разматрани су тест знакова, Вилкоксонов тест означених рангова, Мен - Витнијев тест и Вилкоксонов тест суме рангова. Тестови су детаљно изложени и процедура тестирања је описана како теоријски, тако и кроз практичне примере. За обраду података који су наведени у примерима коришћен је програмски језик R. Резултати тестова добијени су помоћу уграђених функција у R - у или помоћу програма које је написао аутор и чији се кодови налазе у Прилогу рада.

У случају непрекидних симетричних расподела разматрани тестови су послужили за тестирање хипотезе о средњој вредности обележја. Тестови су упоређивани међусобно и са одговарајућим параметарским тестом - Студентовим t - тестом.

У првом поглављу изложени су основни појмови о непараметарским тестовима, тестирању статистичких хипотеза и асимптотској релативној ефикасности. Укратко су описани тест симетрије као помоћни тест и Студентов t - тест за један узорак, зависне и независне узорке.

Тестови медијане за један узорак и зависне узорке, конкретно тест знакова и Вилкоксонов тест означених рангова, описани су у другом поглављу.

У трећем поглављу изложена су два најпопуларнија теста медијане за два независна узорка - тест Мен - Витнија и Вилкоксонов тест суме рангова, као и основна својства линеарних статистика ранга.

Разматрани непараметарски тестови медијане упоређивани су са Студентовим t - тестом у четвртом поглављу. Поређења су вршена помоћу функција моћи и асимптотских релативних ефикасности за узорке из четири одабране симетричне непрекидне расподеле.

У Прилогу су наведени кодови програма коришћених у раду уз потребне коментаре. У примерима у другом поглављу употребљавани су кодови програма *sim.test* и *test.znakova* за, редом, тест симетрије и тест знакова. Приликом израчунавања функција моћи у четвртом поглављу коришћени су следећи кодови програма: $fja.moci1$ за један узорак (зависи од програма $raspodela1$, $uzorci1$, *velicina.testa*1, *znakovi.moc*, *rangovi.moc*, *t.moc*1, *moc.testa*1) и *fja.moci* за два независна узорка (зависи од програма *raspodela*, *uzorci*, *velicina.testa*, *vilkoks.moc*, *t.moc*, *moc.testa*).

Садржај

Поглавље 1

Увод

Јакоб Волфовиц[1] први је увео у статистику термин »непараметарски« 1942. године: »Ситуацију када је расподела обележја потпуно одређена знањем о коначном скупу њених параметара ћемо назвати параметарски случај, а супротан случај када је функционалан облик расподеле непознат непараметарски случај.« [5] Од тог момента непараметарска статистика дефинисана је преко онога што није: традиционална статистика заснована на познатим расподелама са непознатим параметрима. Рендлс[2], Хетманспергер[3] и Касела [4]допунили су ову дефиницију изјавом да се »непараметарска статистика може дефинисати тако да укључује сву методологију која не користи модел заснован на једнопараметарској фамилији расподела«.

Традиционалне статистичке методе засноване су на претпоставкама о параметрима расподеле тј. претпоставља се да је узорак добијен из неке познате фамилије расподела (нормална, експоненцијална, Пуасонова, итд.). Свака од ових расподела зависи од једног или више параметара, од којих је барем један непознат. Највећи нагласак је на нормалној расподели, што се пре свега оправдава централном граничном теоремом која омогућава апроксимацију расподеле узорачких средина нормалном расподелом ако су обими узорака довољно велики. Међутим, када је узорак мали и расподела обележја није нормална, потребно је ослонити се на непараметарске алтернативе познатих параматарских тестова.

У овом раду посебно ће се обрађивати непараметарски тестови који служе

[1] *Jacob Wolfowitz* **(1910-1981), амерички статистичар**
[2] *Ronald Randles* **(1942-), амерички статистичар**
[3] *Thomas Hettmansperger* **(1939-), амерички статистичар**
[4] *George Casella* **(1950-), амерички статистичар**

као алтернатива t-тесту.

ПАРАМЕТАРСКИ ТЕСТ	НЕПАРАМЕТАРСКИ ТЕСТ
t - тест за један узорак	тест знакова, Вилкоксонов тест означених рангова
t - тест за зависне узорке	тест знакова, Вилкоксонов тест означених рангова
t - тест за независне узорке	Вилкоксонов тест суме рангова, Ман-Витни тест

1.1 Основни појмови

Статистичка хипотеза је претпоставка о функцији расподеле једног или више обележја - њеном облику, карактеристикама или вредностима непознатих параметара. Конкретно, хипотезе о параметрима расподеле називају се *параметарске хипотезе.* Хипотеза се назива *проста* ако у потпуности одређује расподелу обележја, у супротном се назива *сложена хипотеза.* *Нулта хипотеза* H_0 је хипотеза која се тестира. *Алтернативна хипотеза,* у ознаци H_1 или H_A, представља закључак који се доноси ако се одбаци нулта хипотеза. *Статистички тест* је правило које нам омогућава да донесемо одлуку о одбацивању или неодбацивању H_0 на основу добијене вредности *тест статистике,* која представља функцију добијеног узорка. Расподела тест статистике када је H_0 тачна назива се *нулта расподела* тест статистике. *Критична област* или *област одбацивања* W теста је подскуп могућих вредности тест статистике који, у складу с тестом, води ка одбацивању H_0. *Критичне вредности* тест статистике су границе W. *Грешка прве врсте* се чини одбацивањем нулте хипотезе када је она тачна, а *грешка друге врсте* прихватањем нетачне нулте хипотезе. За тест статистику T и $H_0(\theta \in \Theta_0)$ против $H_1(\theta \in \Theta \setminus \Theta_0)$, где је Θ параметарски простор, вероватноће грешака прве и друге врсте су, редом, једнаке

$$\alpha(\theta) = P\{T \in W|H_0\}, \quad \beta(\theta) = P\{T \notin W|H_1\}.$$

Најмања горња граница или супремум $\alpha(\theta)$ за свако $\theta \in \Theta_0$ назива се *величина теста.* *Ниво значајности* је унапред одређена граница за $\alpha(\theta)$, која се можда неће достићи ако је расподела тест статистике дискретна. У раду ћемо користити симбол α за величину теста, ниво значајности или вероватноћу грешке прве врсте.

Вероватноћа да ће тест статистика упасти у критичну област тј. довести до одбацивања H_0 назива се *моћ теста.* Моћ теста представља вероватноћу доношења правилне одлуке и обично се рачуна при услову да је H_0 нетачна тј.

H_1 тачна, па је $P_W(\theta) = P\{T \in W|H_1\} = 1 - \beta(\theta)$. Моћ теста зависи зависи од следеће четири величине:

1. Степена нетачности H_0 тј. степена противуречности између нулте и алтернативне хипотезе.

2. Величине теста α.

3. Обима узорка.

4. Критичне области W.

Функција моћи теста добија се када се све наведене величине сем прве фиксирају. Значи унапред одређујемо величину теста, обим узорка и облик критичне области. Посматрамо моћ теста као функцију параметра θ.

Алтернативни приступ тестирању хипотеза је рачунање P - *вредности* која се дефинише (у једнодимензионом случају) као вероватноћа критичне области, при тачној H_0, чија је граница реализована вредност тест статистике, у складу с обликом алтернативне хипотезе. Ако је P - вредност мала, то значи да реализована вредност тест статистике није у складу са H_0 и према томе, треба одбацити H_0. С друге стране, ако је P - вредност велика, добијена вредност тест статистике је у складу с H_0, па се нулта хипотеза не одбацује.

Ако желимо да на основу P - вредности донесемо одлуку да ли треба одбацити H_0 или не, потребно је одабрати праг значајности α. Ако је P - вредност мања или једнака α одлучујемо да одбацимо H_0, иначе одлучујемо да не одбацимо H_0. Према томе, P - вредност је најмањи ниво значајности за који се одбацује H_0. P- вредност пружа не само начин за доношење одлуке о нултој хипотези, већ и представу о томе колико су јаки аргументи против H_0.

За статистички тест кажемо да је *постојан* за одређену алтернативу ако моћ теста, када је H_1 тачна, тежи 1, када обим узорка тежи бесконачности.

Статистичке тестове можемо поредити, тј. сматрати их упоредивим ако се користе за тестирање исте нулте и алтернативне хипотезе, са истим прагом значајности. Моћи тестова ћемо поредити помоћу њихових функција моћи.

Један од критеријума за избор између два или више упоредивих тестова је Питманова ефикасност. Код тачкастог оцењивања ефикасност две непристрасне оцене непознатог параметра се дефинише као количник њихових дисперзија. Гранична вредност овог количника, у неким случајевима, може се интерпретирати као релативни број додатних података потребних да би мање ефикасна оцена постигла исту тачност као ефикаснија оцена. Идеја ефикасности двеју тест статистика врло је блиска овоме. Моћ теста сматра се за меру тачности и тестови се морају поредити под еквивалентним условима (код тачкастог оцењивања неопходан услов је да су обе оцене непристрасне). Обично се приликом поређења два теста сви фактори од којих зависи моћ теста фиксирају, сем обима узорка.

Ефикасност моћи теста A релативно у односу на тест B, представља количник $\frac{n_B}{n_A}$, где је n_A обим узорка, потребних да би моћ теста A била једнака моћи теста B када се за тест B узима n_B података. При томе проста нулта и алтернативна хипотеза, облик критичне области и ниво значајности су исти. Ако је n_A мање од n_B, ефикасност теста A релативно у односу на тест B је већа од 1 и кажемо да је тест A ефикаснији од теста B. Преферирамо тест који захтева мањи обим узорка под истим условима, пошто мањи узорак обично значи мањи утрошак времена и новца. Пошто ефикасност моћи теста A зависи од одабраног прага значајности, хипотеза и обима узорка n_B, често су рачунски поступак и интерпретација резултата компликовани. У већини случајева ови проблеми превазилазе се рачунањем такозване граничне вредности ефикасности моћи теста.

Дефиниција 1 *Нека су A и B два постојана теста, H_0 и H_1, редом, нулта и алтернативна хипотеза и α праг значајности. Асимптотска релативна ефикасност (ARE) теста A релативно у односу на тест B је гранична вредност количника $\frac{n_B}{n_A}$, где је n_A број података потребних да би моћ теста A била једнака моћи теста B са n_B података, када $n_B \to \infty$ и $H_1 \to H_0$, односно ARE је једнака*

$$\lim_{n_B \to \infty} \frac{n_B}{n_A} = const.$$

Под записом $H_1 \to H_0$ сматрамо да се одступање између алтернативне и нулте хипотезе смањује, тј. посматрамо све мања и мања одступања H_1 од H_0.

С обзиром да гранична вредност количника обима узорака тежи константи, она неће зависити од избора нивоа значајности и моћи теста. Услов да оба теста буду постојана у односу на H_1 није никакво ограничење, пошто највећи број тестова који се разматрају за одређени тип алтернативне хипотезе је и онако постојан. Моћ постојаног теста тежи јединици када обим узорка тежи бесконачности. Према томе, пуштамо да $H_1 \to H_0$ да би се моћ сваког од тестова налазила у интервалу $(\alpha, 1)$ за коначне обиме узорака.

ARE се понекад назива локална асимптотска ефикасност, пошто се односи на моћ теста при великим узорцима у околини нулте хипотезе . У више случајева, ARE је довољно добра апроксимација за узорке средње величине и алтернативне хипотезе које се не разликују много од нулте хипотезе. Међутим, у случајевима када се разматра мали узорак, не могу се доносити смислени закључци на основу вредности ARE.

1.2 Студентов t - тест

1.2.1 t - тест за један узорак

Нека је $(X_1, X_2, \ldots, X_n)$ узорак обима n за обележје X. Претпоставља се да је расподела обележја X нормална, $X : \mathcal{N}(m, \sigma^2)$, при чему су средња вредност и дисперзија непознате. Тестирамо нулту хипотезу $H_0(m = m_0)$ са нивоом значајности α. Тест статистика је тада једнака

$$T = \frac{\overline{X}_n - m_0}{\widetilde{S}_n}\sqrt{n},$$

где је $\overline{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ узорачка средина, а $\widetilde{S}_n^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \overline{X}_n)^2$ поправљена узорачка дисперзија.

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

Ако је алтернативна хипотеза облика $H_1(m > m_0)$, критична област је $\{T \geq c\}$, где је критична вредност $c = t_{n-1;1-\alpha}$.

Ако је алтернативна хипотеза облика $H_1(m < m_0)$, критична област је $\{T \leq c\}$, где је критична вредност $c = t_{n-1;\alpha}$.

Ако је алтернативна хипотеза облика $H_1(m \neq m_0)$, критична област је $\{|T| \geq c\}$, где је критична вредност $c = t_{n-1;1-\frac{\alpha}{2}}$.

1.2.2 t - тест за зависне узорке

Нека је $((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$ узорак обима n, при чему су елементи парова X_i и Y_i, $i = 1, 2, ...$, зависни, док су парови међусобно независни. Означимо са m_X непознато математичко очекивање обележја X, а са m_Y математичко очекивање обележја Y. Претпоставља се да разлика $D = X - Y$ има нормалну расподелу, $D : \mathcal{N}(m_D, \sigma_D^2)$, где су средња вредност и дисперзија обележја D непозната. Тестирамо нулту хипотезу $H_0(m_D = m_0)$, тј. $H_0(m_X - m_Y = m_0)$ са нивоом значајности α. Тест статистика је једнака

$$T = \frac{\overline{D}_n - m_0}{\widetilde{S}_D}\sqrt{n},$$

где је $\overline{D}_n = \frac{1}{n}\sum\limits_{i=1}^{n} D_i$ узорачка средина, а $\widetilde{S}_D^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(D_i - \overline{D}_n)^2$ поправљена узорачка дисперзија, $D_i = X_i - Y_i,\ i = 1, 2, \ldots, n$.

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

За сваку од алтернативних хипотеза $H_1(m_D > m_0)$, $\quad H_1(m_D < m_0)$, $H_1(m_D \neq m_0)$, критичне области су истог облика као код тестирања хипотезе о средњој вредности код једног узорка.

1.2.3 t - тест за независне узорке

Нека су $(X_1, X_2 \ldots X_{n_1})$ и $(Y_1, Y_2 \ldots Y_{n_2})$ независни узорци, редом, обима n_1 и n_2. Претпоставља се да обележја X и Y имају нормалну расподелу, односно $X : \mathcal{N}(m_X, \sigma_X^2)$, $Y : \mathcal{N}(m_Y, \sigma_Y^2)$, при чему су њихова математичка очекивања и дисперзије непознати. Тестирамо хипотезу $H_0(m_X = m_Y)$, са прагом значајности α. Разликујемо два случаја:

1. Дисперзије σ_X^2 и σ_Y^2 су једнаке.

   Тест статистика је тада једнака

   $$T = \frac{\overline{X}_{n_1} - \overline{Y}_{n_2} - (m_X - m_Y)}{\widetilde{S}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

   где је $\widetilde{S}^2 = \frac{(n_1-1)\widetilde{S}_X^2 + (n_2-1)\widetilde{S}_Y^2}{n_1+n_2-2}$. Ако је тачна нулта хипотеза, тест статистика има Студентову расподелу са $n_1 + n_2 - 2$ степени слободе.

2. Дисперзије σ_X^2 и σ_Y^2 су различите.

   У овом случају се ради о такозваном Велчовом t - тесту. Тест статистика је једнака

   $$T = \frac{\overline{X}_{n_1} - \overline{Y}_{n_2} - (m_X - m_Y)}{\sqrt{\frac{\widetilde{S}_X^2}{n_1} + \frac{\widetilde{S}_Y^2}{n_2}}}.$$

   Ако је тачна нулта хипотеза, тест статистика има Студентову расподелу са ν степени слободе, где је

   $$\nu = \frac{(\frac{\widetilde{S}_X^2}{n_1} + \frac{\widetilde{S}_Y^2}{n_2})^2}{\frac{\widetilde{S}_X^4}{n_1^2(n_1-1)} + \frac{\widetilde{S}_Y^4}{n_2^2(n_2-1)}}$$

За сваку од алтернативних хипотеза $H_1(m_X > m_Y)$, $H_1(m_X < m_Y)$, $H_1(m_X \neq m_Y)$, критичне области су истог облика као код тестирања хипотезе о средњој вредности код једног узорка.

1.3 Тест симетрије

Нека је f густина обележја X. За расподелу обележја X кажемо да је симетрична уколико постоји тачка a за коју важи

$$f(a-x) = f(a+x),$$

за сваки реалан број x. Уколико ово не важи, расподела је асиметрична [4].

Као меру асиметрије користимо такозвани *коефицијент асиметрије,* $\gamma_1(X)$ који се дефинише као

$$\gamma_1(X) = \frac{E(X-E(X))^3}{\sigma(X)^3}.$$

За $\gamma_1(X) = 0$ расподела обележја је симетрична, ако је $\gamma_1(X) > 0$ кажемо да је расподела *позитивно асиметрична* или *асиметрична удесно,*ако је $\gamma_1(X) < 0$ кажемо да је расподела *негативно асиметрична* или *асиметрична улево.*

Код симетричних унимодалних расподела медијана, мода и средња вредност се поклапају.

Узорачки коефицијент асиметрије је једнак

$$g_1 = \frac{m_3}{\widetilde{S}_n^3},$$

где је $m_3 = \frac{n\sum\limits_{i=1}^{n}(X_i-\overline{X_n})^3}{(n-1)(n-2)}$ узорачки трећи момент, а $\widetilde{S_n^2} = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X_n})^2$ поправљена узорачка дисперзија.

Желимо да тестирамо хипотезу да је расподела обележја симетрична, односно тестирамо нулту хипотезу: $H_0(\gamma_1 = 0)$ против алтернативе $H_1(\gamma_1 \neq 0)$. Тест статистика је нешто сложенијег облика [8]

$$\begin{aligned}
\sqrt{b_1} &= \frac{(n-2)g_1}{\sqrt{n(n-1)}},\\
A &= \sqrt{b_1}\sqrt{\frac{(n+1)(n+3)}{6(n-2)}},\\
B &= \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)},\\
C &= \sqrt{2(B-1)}-1,\\
D &= \sqrt{C},\\
E &= \frac{1}{\sqrt{\ln D}},\\
F &= \frac{A}{\sqrt{\frac{2}{C-1}}}
\end{aligned}$$

Тест статистика $Z = E\ln(F+\sqrt{F^2+1})$ има приближно нормалну $\mathcal{N}(0,1)$ расподелу. Грешка прве врсте је $P_{H_0}\{|Z| \geq c\} = \alpha$, где је критична вредност $c = \Phi^{-1}(1-\frac{\alpha}{2})$, а Φ је функција расподеле нормалне $\mathcal{N}(0,1)$ расподеле.

Већ за $n \geq 9$ апроксимација расподеле тест статистике нормалном расподелом је добра (Зар[5]).

[5] *Jerrold Zar* **(1941-), амерички биостатистичар**

Поглавље 2

Тестови медијане за један узорак и за зависне узорке

2.1 Тест знакова

2.1.1 Тест знакова за један узорак

Нека је $(X_1, X_2, \ldots, X_n)$ прост случајан узорак обима n. Функција расподеле обележја X непрекидна је на целом домену и строго растућа у околини непознате медијане M, тј. обележје X има јединствену медијану. Желимо да тестирамо нулту хипотезу да узорак потиче из популације на којој је медијана посматраног обележја једнака претпостављеној вредности M_0. Нулту хипотезу можемо записати $H_0(M = M_0)$ што је еквивалентно са $H_0(\theta = P\{X > M_0\} = P\{X < M_0\} = 0.5)$ (видети [2], [3], [5], [7], [8], [10], [11]).

Ако је тачна нулта хипотеза, просечно половина вредности из узорка је мања од M_0, а половина вредности већа од M_0. Дефинишимо случајне величине - индикаторе $Z_i = I\{X_i > M_0\},\ i = 1, 2, ..., n$, тј:

$$Z_i = \begin{cases} 1, & \text{ако је } X_i > M_0 \\ 0, & \text{ако је } X_i < M_0 \end{cases}.$$

Нулте разлике $X_i = M_0$ се прескачу и смањује се n. Укупан број K елемената узорка већих од M_0, $K = \sum\limits_{i=1}^{n} Z_i$, може се користити за тестирање валидности H_0. Нулта расподела тест статистике K је биномна са параметрима n и $\theta = \frac{1}{2}$. Према томе, при тачној нултој хипотези очекује се да је приближно половина елемената узорка већа од претпостављене вредности медијане.

Ако је алтернативна хипотеза $H_1(M > M_0)$ тачна, више од половине вредности из узорка веће је од M_0. Дакле, у том случају за нулту хипотезу су критичне велике вредности тест статистике, па је критична област $\{K \geq k_\alpha\}$, где је k_α најмањи ненегативан број који задовољава

$$P_{H_0}\{K \geq k_\alpha\} = \sum_{i=k_\alpha}^{n} \binom{n}{i} 0.5^n \leq \alpha.$$

С друге стране, ако је алтернативна хипотеза $H_1(M < M_0)$ тачна, више од половине вредности из узорка мање је од M_0. Дакле, у том случају за нулту хипотезу су критичне мале вредности тест статистике, па је критична област је $\{K \leq k'_\alpha\}$, где је k'_α највећи ненегативан број који задовољава

$$P_{H_0}\{K \leq k'_\alpha\} = \sum_{i=0}^{k'_\alpha} \binom{n}{i} 0.5^n \leq \alpha.$$

Ако је алтернативна хипотеза облика $H_1(M \neq M_0)$ критична област је $\{K \geq k_{\frac{\alpha}{2}}\} \cup \{K \leq k'_{\frac{\alpha}{2}}\}$ где су $k_{\frac{\alpha}{2}}$ и $k'_{\frac{\alpha}{2}}$, редом, најмањи и највећи ненегативни бројеви који задовољавају неједнакости

$$\sum_{i=k_{\frac{\alpha}{2}}}^{n} \binom{n}{i} 0.5^n \leq \frac{\alpha}{2} \quad \text{и} \quad \sum_{i=0}^{k'_{\frac{\alpha}{2}}} n \binom{n}{i} 0.5^n \leq \frac{\alpha}{2}.$$

Важи релација: $k_{\frac{\alpha}{2}} = n - k'_{\frac{\alpha}{2}}$.

Прегледности ради представимо у следећој табели алтернативне хипотезе, одговарајуће критичне области и изразе за рачунање P-вредности за реализовану вредност тест статистике K_0.

Алтернативна хипотеза	*Критична област*	*P-вредност*
$M > M_0$	$K \geq k_\alpha$	$p_1 = \sum_{i=K_0}^{n} \binom{n}{i} 0.5^n$
$M < M_0$	$K \leq k'_\alpha$	$p_2 = \sum_{i=0}^{K_0} \binom{n}{i} 0.5^n$
$M \neq M_0$	$K \geq k_{\frac{\alpha}{2}}$ или $K \leq k'_{\frac{\alpha}{2}}$	$2 \cdot \min\{p_1, p_2\}$

У случајевима када је обим узорка $n \geq 12$ може се користити апроксимација биномне расподеле тест статистике нормалном расподелом. Пошто се дискретна расподела апроксимира непрекидном расподелом, користи се корекција непрекидности. Размотрићемо како изгледа критична област за сваку од алтернативних хипотеза.

1. Ако је $H_1(M > M_0)$, критична област је облика $\{K \geq k_\alpha\}$, где је $k_\alpha = 0.5n + 0.5 + 0.5\sqrt{n}z_\alpha$, а $z_\alpha = F^{-1}(1-\alpha)$.

2. Ако је $H_1(M < M_0)$, критична област је облика $\{K \leq k'_\alpha\}$, где је $k'_\alpha = 0.5n - 0.5 + 0.5\sqrt{n}z'_\alpha$, а $z'_\alpha = F^{-1}(\alpha)$.

3. Ако је $H_1(M \neq M_0)$, критична област је облика $\{K \geq k_{\frac{\alpha}{2}}\} \cup \{K \leq k'_{\frac{\alpha}{2}}\}$, где је $k_{\frac{\alpha}{2}} = 0.5n + 0.5 + 0.5\sqrt{n}z_{\frac{\alpha}{2}}$, $k'_{\frac{\alpha}{2}} = 0.5n - 0.5 - 0.5\sqrt{n}z_{\frac{\alpha}{2}}$, а $z_{\frac{\alpha}{2}} = F^{-1}(1-\frac{\alpha}{2})$.

У следећој табели представљене су алтернативне хипотезе, одговарајуће критичне области и P - вредности за реализовану вредност тест статистике K_0, за $n \geq 12$.

H_1	*Критична област*	P-*вредност*
$M > M_0$	$K \geq 0.5n + 0.5 + 0.5\sqrt{n}z_\alpha$	$p_1 = 1 - F(\frac{K_0 - 0.5 - 0.5n}{0.5\sqrt{n}})$
$M < M_0$	$K \leq 0.5n - 0.5 + 0.5\sqrt{n}z'_\alpha$	$p_2 = F(\frac{K_0 + 0.5 - 0.5n}{0.5\sqrt{n}})$
$M \neq M_0$	$K \geq 0.5n + 0.5 + 0.5\sqrt{n}z_{\frac{\alpha}{2}}$ или $K \leq 0.5n - 0.5 - 0.5\sqrt{n}z_{\frac{\alpha}{2}}$	$2 \cdot \min\{p_1, p_2\}$

Напомена 1 Поступак решавања примера наведених у овом поглављу подељен је на следеће кораке:

1. Провера нормалности расподеле разматраног обележја помоћу графика теоријских и узорачких квантила ($Q - Q$ дијаграм нормалне расподеле) и Лилифор-Колмогоров теста [12] [1]. Користиће се готове функције за цртање графика и тестирање у статистичком софтверу R. Уколико се ради о нормалној расподели обележја, сматраћемо резултате t - теста за највероводостојније, и у складу с тим изводити закључке приликом поређења са резултатима непараметарских тестова.

2. У случају да расподела обележја није нормална, тестира се симетричност расподеле обележја (користиће се код програма - фунције $sim.test$ наведен у Прилогу).

3. Код симетричних расподела биће примењени и параметарски (t - тест) и непараметарски тестови (тест знакова, Вилкоксонов тест означених рангова) за тестирање хипотеза о средњој вредности обележја.

[1] **Лилифор-Колмогоров тест је варијанта теста Колмогорова за тестирање хипотезе да посматрано обележје има нормалну расподелу, када су параметри расподеле непознати, па их је потребно оценити на основу добијеног узорка.**

4. Код асиметричних расподела биће примењен само тест знакова.

5. Извођење закључака на основу добијених резултата.

Пример 1. Током месец дана (радним данима) бележена је дужина међуградских телефонских позива (у минутима) једне мале фирме. Добијени су следећи подаци, сортирани у растућем поретку (подаци, под називом *call.time*, преузети су из базе података *Phone*, пакет *PASWR*):

0.2 0.2 0.2 0.2 0.7 0.7 0.7 0.8 1.3 1.7 2.1 2.1 2.7 4.0 4.3 5.2 5.6 6.1 6.7 7.0 9.7 9.7 12.9

Да ли можемо тврдити са прагом значајности $\alpha = 0.05$ да је медијана дужине позива већа од 2 минута ?

Обележје X представља дужину међуградских телефонских позива, вектор података назван је *duzina.poziva*. $Q-Q$ дијаграм нормалне расподеле за дате податке добија се на основу следећих наредби у R - у (помоћу функција *qqnorm* и *qqline*):

```
duzina.poziva<-Phone$call.time
qqnorm(duzina.poziva,xlab="Teorijski kvantili",ylab="Uzoracki
kvantili", main="Q-Q dijagram  normalne raspodele")
qqline(duzina.poziva)
title(sub="Duzina poziva")
```

Добијени график се налази на слици 2.1.

Код нормалне расподеле тачке $Q-Q$ дијаграма приближно припадају правој. На слици примећујемо више података који немало одступају од праве. На основу овог графика бисмо закључили да разматрано обележје нема нормалну расподелу, што ћемо проверити Лилифор-Колмогоров тестом (помоћу функције *lillie.test* из пакета *nortest*), са прагом значајности $\alpha = 0.05$.

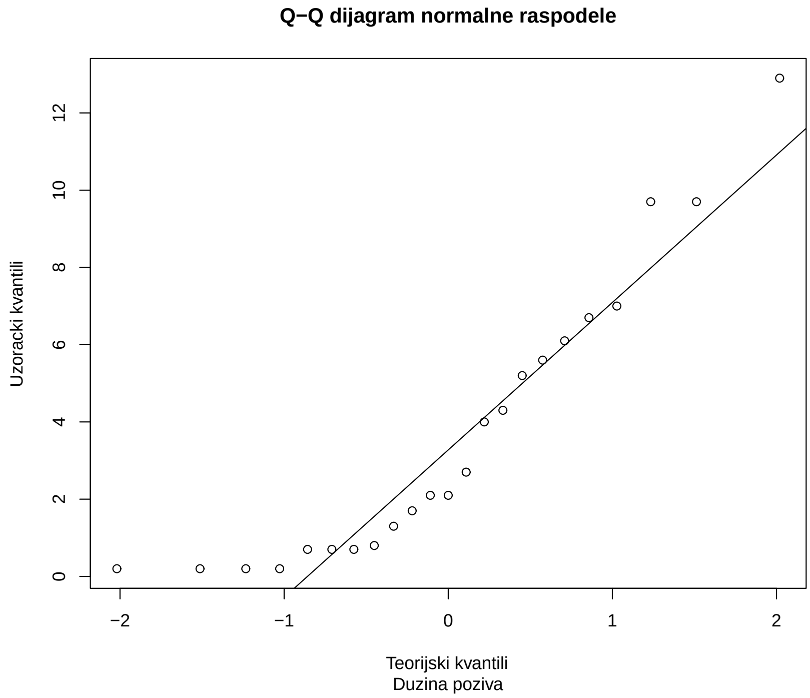
```
lillie.test(duzina.poziva)

     Lilliefors (Kolmogorov-Smirnov) normality test
data:  duzina.poziva D = 0.191, p-value = 0.0291
```

На основу P - вредности, $P = 0.0291 < 0.05$, закључујемо да обележје X нема нормалну расподелу.

Проверимо сада да ли је расподела обележја X симетрична (са прагом значајности $\alpha = 0.05$).

```
sim.test(duzina.poziva)
```



Слика 2.1: $Q-Q$ дијаграм нормалне расподеле за дужину позива

```
Uzoracki koeficijent asimetrije 1.047623
Realizovana vrednost test statistike:  2.118354
Kriticna vrednost:  1.959964
Raspodela obelezja je asimetricna
```

С обзиром да је расподела обележја асиметрична, медијана обележја се не мора поклапати са средњом вредношћу, па не одговара ни t - тест, ни Вилкоксонов тест означених рангова.

Тестираћемо нулту хипотезу $H_0(M=2)$ против алтернативне хипотезе $H_1(M>2)$ помоћу теста знакова. Укупан број елемената узорка већих од 2 је 13, па је реализована вредност тест статистике $K_0=13$. Критична област је облика $W=[k_\alpha, 23]$. Израчунајмо критичну вредност на основу неједнакости

$$\sum_{i=k_\alpha}^{23} \binom{23}{i} (0.5)^{23} \leq 0.05$$

Добија се да је $k_\alpha = 16$. Како реализована вредност тест статистике не упада у критичну област, не одбацујемо нулту хипотезу. Одредимо колико су јаки аргументи против нулте хипотезе помоћу P - вредности. За алтернативну хипотезу овог облика имамо да је

$$\sum_{i=13}^{23}\binom{23}{i}(0.5)^{23}=0.3388197,$$

што (довољно) говори у прилог неодбацивања H_0

Како је обим узорка $n=23$, можемо користити апроксимацију биномне расподеле тест статистике нормалном расподелом, уз корекцију непрекидности. P - вредност је тада једнака

$$P=1-F(\frac{13-0.5-0.5\cdot 23}{0.5\sqrt{23}})=0.3383287.$$

Дакле, можемо се уверити да се добија врло добра тачност приликом апроксимације расподеле тест статистике нормалном расподелом.

Резултате теста знакова (реализовану вредност тест статистике, P - вредност) можемо директно добити помоћу функције *test.znakova* написане у R - у.

```
test.znakova(duzina.poziva,M=2,alfa=0.05,alternativa="vece")

Vrednost test statistike je 13
P-vrednost testa:  0.3388197
Ne odbacujemo nultu hipotezu.
```

2.1.2 Тест знакова за зависне узорке

Нека је $((X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n))$ узорак обима n, при чему су елементи парова X_i и Y_i, $i=1,2,...$, зависни, док су парови међусобно независни. Формирају се разлике $D_i=X_i-Y_i$. Ако је расподела разлика непрекидна са непознатом медијаном M_D и важи $P\{D=M_D\}=0$ можемо применити на узорку разлика исту процедуру при тестирању нулте хипотезе $H_0(M_D=M_0)$ као код теста знакова за један узорак.

Пример 2. Асистент на предмету Статистика сматра да бар половина његових студената има више поена на првом него на другом колоквијуму. Да би тестирао ову тврдњу, бележи за сваког од 24 студената поене на оба колоквијума. Добијени су следећи подаци:

Први колоквијум: 0, 10, 5, 7, 1, 7, 4, 6, 3, 22, 5, 16, 10, 17, 25, 10, 6, 7, 17, 19, 16, 11, 12, 6

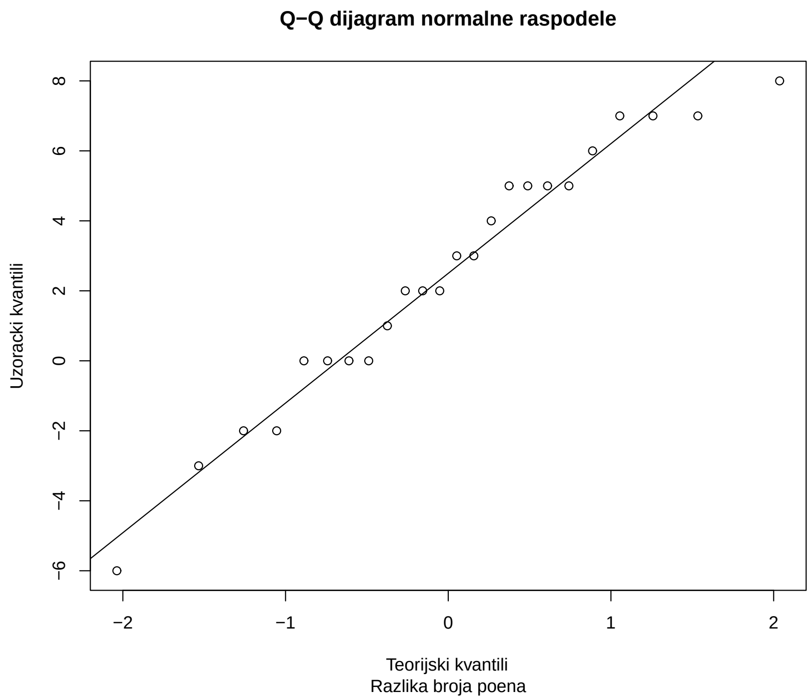
Други колоквијум: 0, 15, 13, 10, 8, 7, 10, 10, 4, 20, 12, 14, 7, 22, 25, 13, 8, 7, 24, 21, 10, 16, 17, 8

Да ли су добијени подаци у складу с хипотезом асистента, са прагом значајности $\alpha = 0.05$?

Нека је X обележје које представља број поена на другом колоквијуму, а Y број поена на првом колоквијуму. Како су вршена два мерења над истим субјектима (редом је бележен број поена са првог и другог колоквијума за сваког студента), узорци су зависни. Проверићемо да ли је расподела обележја $D = X - Y$ нормална. $Q - Q$ дијаграме нормалне расподеле добијамо помоћу следећих наредби у R - у:

```
kol.1<-c(0,10,5,7,1,7,4,6,3,22,5,16,10,17,25,10,6,7,17,19,16,
11,12,6)
kol.2<-c(0,15,13,10,8,7,10,10,4,20,12,14,7,22,25,13,8,7,24,21,10,
16,17,8)
qqnorm(kol.2-kol.1,xlab="Teorijski kvantili",ylab=
"Uzoracki kvantili",main="Q-Q dijagram normalne raspodele")
qqline(kol.2-kol.1)
title(sub="Razlika broja poena")
```

Добијени $Q - Q$ дијаграм налази се на слици 2.2.



Слика 2.2: $Q - Q$ дијаграм нормалне расподеле за разлику поена

Тачке на графику су груписане око праве, па би се могло сматрати да је расподела обележја нормална, што ћемо још проверити Лилифор - Колмогоровим тестом.

```
lillie.test(kol.2-kol.1)

    Lilliefors (Kolmogorov-Smirnov) normality test
data:  kol.2 - kol.1
D = 0.1335, p-value = 0.3271
```

Дакле, прихватамо хипотезу да је расподела разлике $D = X - Y$ нормална. Како је нормална расподела симетрична, медијана и средња вредност разлике броја поена се поклапају. Нулта хипотеза о медијани разлике $H_0(M_D = 0)$, против алтернативе $H_1(M_D > 0)$ су тада еквивалентне $H_0(m_D = 0)$ против алтернативе $H_1(m_D > 0)$, где је m_D средња вредност разлике поена. Пошто је се овом случају t - тест сматра најмоћнијим, прво ћемо њиме тестирати нулту хипотезу, па ћемо добијене резултате упоредити са резултатима теста знакова за зависне узорке.

Узорачка средина и поправљена узорачка диспезија су, редом, једнаке:

$$\overline{x}_D = 2.5, \quad \widetilde{s}^2_D = 13.13.$$

Критична област је облика $T \geq c$, где је $T = \frac{\overline{X}_D - m_0}{\widetilde{S}_D}\sqrt{n}$, а c критична вредност. Тест статистика, при тачној нултој хипотези, има Студентову расподелу са 23 степена слободе, па је критична вредност $c = t_{23;0.95} = 1.714$. Реализована вредност тест статистике је $T = 3.234$. Пошто вредност тест статистике упада у критичну област $W = [1.714, +\infty)$, одбацујемо нулту хипотезу. Прихватамо алтернативну хипотезу да бар половина студената има више поена на другом него на првом колоквијуму.

Резултате t - теста можемо директно добити помоћу функције $t.test$ у R - у.

```
t.test(kol.2,kol.1,paired=T, alternative="g")

     Paired t-test
data:  kol.2 and kol.1
t = 3.3238, df = 23, p-value=0.001478
alternative hypothesis:  true difference in means is greater than 0
95 percent confidence interval:
  1.190733   Inf
sample estimates:
mean of the differences
    2.458333
```

Дакле, на основу врло мале P - вредности $P = 0.001478 < 0.05$, не само да одбацујемо нулту хипотезу, него можемо закључити да су аргументи против ње врло јаки.

Сада ћемо тестирати нулту хипотезу помоћу теста знакова. Имамо да је низ разлика поена на другом и првом коловијуму:

0, 5, 8, 3, 7, 0, 6, 4, 1, −2, 7, −2, −3, 5, 0, 3, 2, 0, 7, 2, −6, 5, 5, 2

У низу разлика поена имамо четири нулте разлике које избацујемо и смањујемо n на 20. Укупан број позитивних разлика је једнак $K_0 = 16$. P - вредност теста је једнака $P\{K \geq 16\} = \sum\limits_{i=16}^{20} \binom{20}{i}(0.5)^{20} = 0.0059$. Дакле, на основу P - вредности одбацујемо нулту хипотезу.

Погледајмо резултате теста знакова у R - у:

```
test.znakova(kol.2-kol.1,M=0,alfa=0.05,alternativa="vece")

Vrednost test statistike je 16
P-vrednost testa:  0.005908966
Odbacujemo nultu hipotezu.
```

2.2 Вилкоксонов[2] тест означених рангова

2.2.1 Својства рангова

Нека је дат прост случајан узорак $(X_1, X_2, \ldots, X_n)$ обима n. Означимо рангове случајних величина $X_1, X_2, \ldots, X_n$, са, редом, $r(X_1),\ r(X_2), \ldots, r(X_n)$ тј. r је таква функција тако да је $r(X_i) \leq r(X_j)$, ако је $X_i \leq X_j,\ i, j = 1, 2, \ldots n$.

Многи непараметарски тестови засновани су на рангирању елемената узорка, у односу на претпостављену вредност непознатог параметра или у односу на вредности из другог узорка. Имамо да је

$$P\{r(X_i) = j\} = \frac{1}{n},\ j = 1, 2, ..., n,$$

тј. рангови имају дискретну униформну расподелу. Означимо са $R_i = r(X_i),\ i = 1, 2, ..., n$. Добијамо да су математичко очекивање и дисперзија случајних величина R_i, редом једнаки

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

$$\begin{aligned} E(R_i) &= \sum_{j=1}^{n} \frac{j}{n} = \frac{n+1}{2}, \quad E(R_i^2) = \sum_{j=1}^{n} \frac{j^2}{n} = \frac{(n+1)(2n+1)}{6}, \\ D(R_i) &= \frac{n^2-1}{12}. \end{aligned}$$

У случају да међу елементима низа који се рангирају има истих вредности најчешће се, због једноставности, примењује метода средине рангова. Метода средине рангова додељује сваком члану групе истих вредности просечну вредност рангова који би били додељени да су све вредности из групе различите. На овај начин се истим вредностима додељују исти рангови.

2.2.2 Вилкоксонов тест означених рангова за један узорак

Нека је дат прост случајан узорак $(X_1, X_2, \ldots, X_n)$ обима n и нека је расподела обележја X непрекидна и симетрична око непознате медијане M. Желимо да тестирамо нулту хипотезу да узорак потиче из популације чије обележје има медијану једнаку претпостављеној вредности M_0. Пошто претпостављамо да је расподела обележја симетрична, средња вредност и медијана се поклапају, па се нулта хипотеза о медијани $H_0(M = M_0)$ своди на хипотезу о средњој вредности m, тј. $H_0(m = m_0)$. Процедуру потребну за рачунање вредности тест статистике изложићемо у наредна четири корака (в. [2], [3], [5], [7], [8], [10], [11]).

1. Формирају се разлике $D_i = X_i - M_0,\ i = 1, 2, ..., n$.

2. Рангирају се апсолутне вредности разлика $|D_i|$. Све разлике једнаке нули прескачу се (не рангирају се и елиминишу се из даље анализе) и смањује се n на n' - број не-нултих разлика. У случају да има истих вредности у варијационом низу апсолутних разлика, примењује се метод средине рангова.

3. Знак сваке разлике ставља се испред њеног ранга.

4. Рачуна се сума свих "позитивних" рангова (рангова позитивних разлика) T^+ и сума свих "негативних" рангова (рангова негативних разлика) T^-.

Ако је тачна нулта хипотеза, просечно половина вредности из узорка већа је од M_0, а половина мања од M_0, односно очекивана вредност T^+ једнака је очекиваној вредности T^-. Како је сума свих рангова константна, тј. $T^+ + T^- = \sum\limits_{i=1}^{n} i = \frac{n(n+1)}{2}$, тест статистике које чине само T^+, T^- или $T^+ - T^-$ су линеарно зависне и према томе може се користити било која од њих.

Тест статистике T^+ и T^- можемо записати на следећи начин:

$$T^+ = \sum_{i=1}^{n} Z_i r(|D_i|), \quad T^- = \sum_{i=1}^{n} (1 - Z_i) r(|D_i|),$$

где је

$$Z_i = \begin{cases} 1, & \text{ако је } D_i > 0 \\ 0, & \text{ако је } D_i < 0 \end{cases}.$$

Добијамо да је $T^+ - T^- = \sum\limits_{i=1}^{n} 2Z_i r(|D_i|) - \frac{n(n+1)}{2}$.

Ако је тачна нулта хипотеза, случајне величине Z_i су независне и једнако расподељене, $Z_i : B(1, \frac{1}{2})$. Како је $E(Z_i) = \frac{1}{2}$ и $D(Z_i) = \frac{1}{4}$, имамо да је

$$E(T^+|H_0) = \sum_{i=1}^{n} \frac{r(|D_i|)}{2} = \frac{n(n+1)}{4}.$$

Сада ћемо показати да су Z_i и $r(|D_i|)$ независне случајне величине. Одредимо условну расподелу случајне величине $|D_i|$ $|Z_i$.

$$\begin{aligned} P\{|X_i - M_0| \leq x | Z_i = 1\} &= P\{|X_i - M_0| \leq x | X - M_0 > 0\} = \\ &= \frac{P\{0 < X_i - M_0 \leq x\}}{P\{X_i - M_0 > 0\}} = 2P\{0 < X_i - M_0 \leq x\} = \\ &= P\{|X_i - M_0| \leq x\}. \end{aligned}$$

Последња једнакост важи јер је расподела обележја симетрична. На исти начин добија се да је

$$P\{|X_i - M_0| \leq x | Z_i = 0\} = P\{|X_i - M_0| \leq x\}.$$

Дакле, случајне величине $|D_i|$ и Z_i су независне, па су $r(|D_i|)$ и Z_i такође независне случајне величине.

На основу претходно добијених резултата, дисперзија тест статистике T^+ је једнака:

$$D(T^+|H_0) = \sum_{i=1}^{n} \frac{r(|D_i|)^2}{4} = \frac{n(n+1)(2n+1)}{24}.$$

Тест статистика Вилкоксоновог теста означених рангова може се представити и на следећи начин:

$$T^+ = \sum_{1 \le i \ne j \le n}\sum T_{ij},$$

где је

$$T_{ij} = \begin{cases} 1, & \text{ако је } D_i + D_j > 0 \\ 0, & \text{иначе} \end{cases}.$$

При тачној нултој хипотези, случајне величине D_i имају исту расподелу. Дефинишимо за различите i, j, k вероватноће

$$\begin{aligned} p_1 &= P\{D_i > 0\}, & p_2 = P\{D_i + D_j > 0\}, \\ p_3 &= P\{D_i > 0, D_i + D_j > 0\}, & p_4 = P\{D_i + D_j > 0, D_i + D_k > 0\}. \end{aligned}$$

Математичко очекивање, дисперзија и коваријација случајних величина - индикатора T_{ij} за различите i, j, k, m тада су једнаки

$$\begin{aligned} E(T_{ii}) &= p_1, & E(T_{ij}) = p_2, \\ D(T_{ii} &= p_1 - p_1^2, & D(T_{ij}) = p_2 - p_2^2, \\ cov(T_{ii}, T_{ik}) &= p_3 - p_1 p_2, & cov(T_{ij}, T_{ik}) = p_4 - p_2^2, \quad cov(T_{ij}, T_{mk}) = 0. \end{aligned}$$

Према томе, математичко очекивање и дисперзија тест статистике T^+ једнаки су

$$\begin{aligned} E(T^+) &= n \cdot E(T_{ii}) + \frac{n \cdot (n-1)}{2} E(T_{ij}) = np_1 + \frac{n \cdot (n-1)}{2} \cdot p_2, \\ D(T^+) &= n \cdot D(T_{ii}) + \binom{n}{2} D(T_{ij}) + 2n(n-1)\, cov(T_{ii}, T_{ik}) + \\ &+ 2n\binom{n-1}{2} cov(T_{ij}, T_{ik}) + \binom{n}{4}\, cov(T_{ij}, T_{mk}) = \\ &= np_1(1-p_1) + \frac{n(n-1)}{2} p_2(1-p_2) + 2n(n-1)(p_3 - p_1 p_2) + \\ &+ n(n-1)(n-2)(p_4 - p_2^2). \end{aligned}$$

Под претпоставком да је расподела обележја симетрична и да је нулта хипотеза тачна, добијамо

$$
\begin{aligned}
p_1 &= P\{D_i > 0\} = \frac{1}{2}, \\
p_2 &= P\{D_i + D_j > 0\} = \int_{-\infty}^{+\infty} \int_{-u}^{+\infty} f_D(u) f_D(t) du dt = \\
&= \int_{-\infty}^{+\infty} (1 - F_D(-u)) f_D(u) du = \int_{-\infty}^{+\infty} F_D(u) f_D(u) du = \frac{1}{2}, \\
p_3 &= P\{D_i > 0, D_i + D_j > 0\} = \int_{0}^{+\infty} \int_{-u}^{+\infty} f_D(u) f_D(t) du dt = \\
&= \int_{0}^{+\infty} (1 - F_D(-u)) f_D(u) du = \int_{0}^{+\infty} F_D(u) f_D(u) du = \frac{3}{8}, \\
p_4 &= P\{D_i + D_j > 0, D_i + D_k > 0\} = P\{0 < D_i + D_j < D_i + D_k\} + \\
&+ P\{0 < D_i + D_k < D_i + D_j\} = 2P\{-D_i < D_j < D_k\} = \\
&= 2 \int_{-\infty}^{+\infty} \int_{-u}^{+\infty} \int_{v}^{+\infty} f_D(u) f_D(v) f_D(t) du dv dt = \frac{1}{3}.
\end{aligned}
$$

Када заменимо израчунате вероватноће у изразе за математичко очекивање и дисперзију статистике T^+, добијамо

$$
E(T^+|H_0) = \frac{n(n+1)}{4}, \ D(T^+|H_0) = \frac{n(n+1)(2n+1)}{24}.
$$

Екстремне вредности тест статистике T^+ које се јављају када све разлике D_i имају исти знак, негативан или позитиван, редом су 0 и $\frac{n(n+1)}{2}$. Како је тест статистика потпуно одређена индикаторима Z_i, за простор свих исхода може се сматрати скуп свих могућих n-торки $(Z_1, Z_2, ..., Z_n)$ којих је укупно 2^n. Ако је тачна нулта хипотеза свака од тих n-торки нула и јединица има исту вероватноћу појављивања. Према томе, при тачној нултој хипотези расподела вероватноћа T^+ је

$$
P\{T^+ = k\} = \frac{t(k)}{2^n},
$$

где је $t(k)$ број свих начина да природним бројевима од 1 до n додели позитиван или негативан знак тако да је сума бројева са позитивним знаком једнака k.

Како се свака од тих додела јавља с истом вероватноћом, расподела T^+ при тачној нултој хипотези симетрична је око своје средње вредности $\frac{n(n+1)}{4}$.

Пример 3. Нађимо закон расподеле тест статистике T^+ за $n = 4$. Због својства симетричности расподеле тест статистике, довољно је одредити само половину вредности T^+.

Вредности T^+	Рангови позитивних разлика	Број додела
10	$1, 2, 3, 4$	1
9	$2, 3, 4$	1
8	$1, 3, 4$	1
7	$1, 2, 4$ или $3, 4$	2
6	$1, 2, 3$ или $2, 4$	2
5	$1, 4$ или $2, 3$	2

Добијамо да је $P\{T^+ = k\} = \begin{cases} \frac{1}{16}, & k = 0, 1, 2, 8, 9, 10 \\ \frac{2}{16}, & k = 3, 4, 5, 6, 7 \end{cases}$

Велике вредности T^+ одговарају малим вредностима T^-. Статистике T^+ и T^- једнако су расподељене, пошто је

$$\begin{aligned} P\{T^+ \geq c\} &= P\{T^+ - \frac{n(n+1)}{4} \geq c - \frac{n(n+1)}{4}\} \overset{3}{=} \\ &= P\{\frac{n(n+1)}{4} - T^+ \geq c - \frac{n(n+1)}{4}\} = P\{\frac{n(n+1)}{2} - T^+ \geq c\} = \\ &= P\{T^- \geq c\} \end{aligned}$$

Ако је t_α такав број да је $P\{T \leq t_\alpha\} = \alpha$, одговарајуће критичне области величине α при тестирању $H_0(M = M_0)$ су:

1. Алтернативна хипотеза је облика $H_1(M > M_0)$. Ако се прихвати H_1 значи да је узорак добијен из популације на којој медијана обележја X има већу вредност од претпостављене, биће $T^+ > T^-$, па је критична област облика $\{T^+ \geq c\}$ (или $\{T^- \leq c\}$).

2. Алтернативна хипотеза је облика $H_1(M < M_0)$. Ако се прихвати H_1 значи да је узорак добијен из популације на којој медијана обележја X има мању вредност од претпостављене, биће $T^+ < T^-$, па је критична област облика $\{T^+ \leq c\}$.

[3] Једнакост важи због симетричности расподеле T^+ око $\frac{n(n+1)}{4}$.

3. Алтернативна хипотеза је облика $H_1(M \neq M_0)$. За нулту хипотезу су тада критичне мале или велике вредности тест статистике, па је критична област облика $\{T^+ \leq c_1\} \cup \{T^+ \geq c_2\}$.

Представимо у следећој табели алтернативне хипотезе, одговарајуће критичне области и P - вредности на основу реализоване вредности t^+ тест статистике.

Алтернативна хипотеза	*Критична област*	*P-вредност*
$M > M_0$	$T^+ \geq t_\alpha$	$p_1 = P\{T^+ \geq t^+\}$
$M < M_0$	$T^+ \leq t'_\alpha$	$p_2 = P\{T^+ \leq t^+\}$
$M \neq M_0$	$T^+ \geq t_{\frac{\alpha}{2}}$ или $T^+ \leq t'_{\frac{\alpha}{2}}$	$2 \cdot \min\{p_1, p_2\}$

Пример 4. Професор Икс тврди да просечно половина његових студената има барем 90 поена на његовом испиту. Сумњичави студент жели да испита професорову тврдњу и бележи поене 10 студената који су положили испит код професора Икс. Добијени су следећи подаци:

$$75,\ 100,\ 92,\ 94,\ 86,\ 77,\ 99,\ 66,\ 64,\ 71$$

Са прагом значајности $\alpha = 0.05$ тестирати да ли је студент у праву.

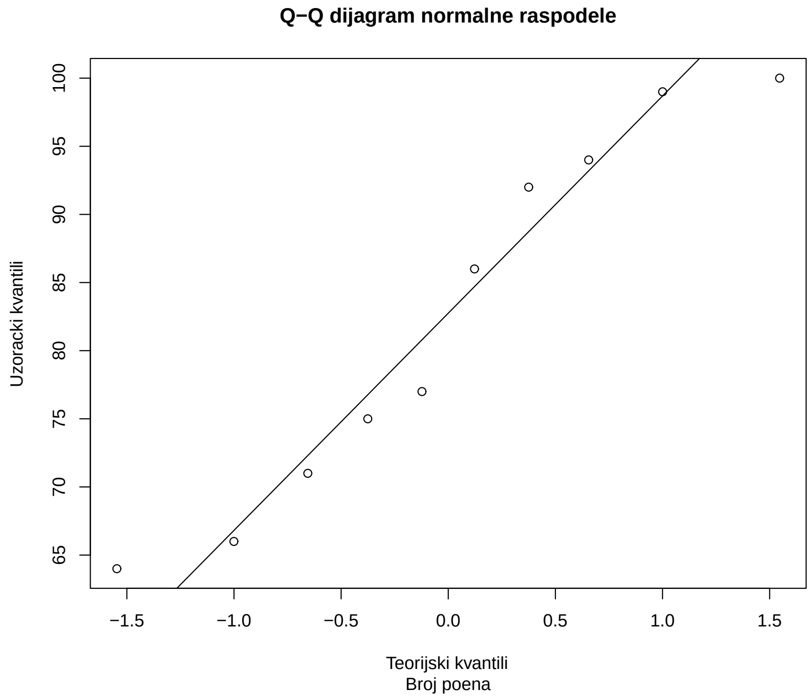
Испитаћемо да ли обележје X - број поена на испиту професора Икс има нормалну расподелу. $Q-Q$ дијаграм нормалне расподеле добијамо помоћу истог низа наредби као у претходним примерима.

```
poeni<-c(75,100,92,94,86,77,99,66,64,71)
qqnorm(poeni, xlab="Teorijski kvantili", ylab="Uzoracki
kvantili", main="Q-Q dijagram normalne raspodele")
qqline(poeni)
title(sub="Broj poena")
```

Добијени график се налази на слици 2.3.

Примећујемо на графику да су тачке концентрисане око праве, односно нема значајних одступања тачака од праве. Дакле, на основу графика бисмо донели закључак да обележје X има нормалну расподелу што ћемо проверити и Лилифор-Колмогоровим тестом.

```
lillie.test(poeni)
    Lilliefors (Kolmogorov-Smirnov) normality test
data:  poeni
D = 0.1611, p-value = 0.653
```



Слика 2.3: $Q - Q$ *дијаграм нормалне расподеле за број поена*

На основу P - вредности, можемо прихватити хипотезу да обележје X има нормалну расподелу. У овом случају t - тест се сматра најмоћнијим. Тестираћемо нулту хипотезу $H_0(m = 90)$ против алтернативе $H_0(m < 90)$.

Узорачка средина и поправљена узорачка дисперзија броја поена су, редом, једнаке:

$$\overline{x}_n = 82.4, \;\; \widetilde{s}_n^2 = 182.93.$$

Критична област је облика $\{T \leq c\}$, где је $T = \frac{\overline{X}_n - 90}{\widetilde{S}_n}\sqrt{n}$ тест статистика, а c критична вредност. Ако је тачна нулта хипотеза, тест статистика има приближно Студентову расподелу са 9 степени слободе, па је критична вредност једнака $c = t_{9;0.05} = -1.833$. Реализована вредност тест статистике је $T = -1.777$. Како вредност тест статистике не упада у критичну област, не одбацујемо нулту хипотезу.

Погледајмо резултате t - теста у R - у.

```
t.test(poeni,mu=90,alternative="l")
  One Sample t-test
data:  poeni t = -1.7769, df = 9, p-value = 0.05465
alternative hypothesis:  true mean is less than 90
```

```
95 percent confidence interval:
  -Inf 90.24035
sample estimates:
mean of x
  82.4
```

Дакле, P-вредност је једнака $P = 0.055 > 0.05$, на основу чега не можемо одбацити нулту хипотезу, али сматрамо да би било потребно спровести ново истраживање.

Упоредимо резултате t - теста са резултатима теста знакова и Вилкоксоновог теста означених рангова. Користићемо опет функцију *test.znakova*. Добија се:

```
   test.znakova(poeni,90,0.05,"manje")
Vrednost test statistike je 4
P-vrednost testa:  0.3769531
Ne odbacujemo nultu hipotezu.
```

Представимо потребне кораке у израчунавању реализоване вредности тест статистике следећом табелом:

x_i	75	100	92	94	86	77	99	66	64	71
$d_i = x_i - 90$	-15	10	2	4	-4	-13	9	-24	-26	-19
$\lvert d_i \rvert$	15	10	2	4	4	13	9	24	26	19
ранг$\lvert d_i \rvert$	7	5	1	2.5	2.5	6	4	9	10	8
рангови са знаком	-7	5	1	2.5	-2.5	-6	4	-9	-10	-8

Добијамо да је сума рангова са позитивним знаком једнака $t^+ = 12.5$, а сума рангова са негативним знаком $t^- = 42.5$. Пошто је лакше радити са мањим сумама, да би илустровали рачунање критичне вредности узећемо T^+ за тест статистику. Критична област је тада облика $\{T^+ \leq c\}$

Вредности T^+	Рангови позитивних разлика	Број додела
0		1
1	1	1
2	2	1
3	$3;\ 1,2$	2
4	$4;\ 1,3$	2
5	$5;\ 2,3;\ 1,4$	3
6	$6;\ 2,4;\ 1,2,3;\ 1,5$	4
7	$7;\ 1,6;\ 2,5;\ 3,4;\ 1,2,4$	5
8	$8;\ 1,7;\ 2,6;\ 3,5;\ 1,3,4;\ 1,2,5$	6
9	$9;\ 1,8;\ 2,7;\ 3,6;\ 4,5;\ 2,3,4;\ 1,3,5$	7
10	$10;\ 1,9;\ 2,8;\ 3,7;\ 4,6;\ 1,4,5;\ 1,2,7;$	10
	$1,2,3,4;\ 2,3,5;\ 1,3,6$	
11	$11;\ 1,10;\ 2,9;\ 3,8;\ 4,7;\ 5,6;\ 1,2,8;$	12
	$2,3,6;\ 2,4,5;\ 1,3,7;\ 1,4,6;\ 1,2,3,5$	
12	$12;\ 1,11;\ 2,10;\ 3,9;\ 4,8;\ 1,2,9;\ 1,3,8$	13
	$1,4,7;\ 1,5,6;\ 1,2,3,6;\ 1,2,4,5;\ 2,3,7\ 2,4,6$	

Како је $P\{T^+ \leq 10\} = \frac{42}{1024} = 0.041 < 0.05$, а $P\{T^+ \leq 11\} > 0.05$, критична вредност $c = 10$. С друге стране, P -вредност теста је једнака $P\{T^+ \leq 12.5\} = P\{T^+ \leq 12\} = 0.0654$.

Резултате Вилкоксоновог теста означених рангова можемо директно добити помоћу функције *wilcox.exact* (пакет *exactRankTests*).

```
wilcox.exact(poeni,mu=90,alternative="l")

Exact Wilcoxon signed rank test
data:  poeni
V = 12.5, p-value = 0.06934
alternative hypothesis:  true mu is less than 90
```

Према томе, ако сматрамо t - тест за најверодостојнији у овом случају, видимо да је P - вредност Вилкоксоновог теста означених рангова блиска P - вредности t - теста, док је P - вредност теста знакова доста већа у односу на њих и више говори у прилог неодбацивања нулте хипотезе.Дакле, до врло сличних закључака бисмо дошли употребом t - теста и Вилкоксоновог теста означених рангова на овом примеру, док би нас резултати теста знакова довољно убедили у одлуку у неодбацивању нулте хипотезе.

Када је потребно одредити расподелу тест статистике може се користити једноставна рекурзивна релација за генерисање вероватноћа. Означимо са T_n^+ суму рангова придружених позитивним разликама D_i, $i = 1, 2, ..., n$, за узорак обима n. Размотримо низ $n - 1$ апсолутних разлика $|D_i|$ којима су додељени рангови $1, 2, ..., n-1$, за које нулта расподела T_{n-1}^+ позната. Да бисмо одредили расподелу T_n^+, додајемо разлику D_n за n-ти елемент узорка и, без губитка општости, можемо претпоставити да је $|D_n| > |D_i|$, за свако $i \leq n - 1$. Ранг разлике $|D_n|$ је тада једнак n. Ако је $D_n > 0$, вредност T_n^+ ће бити већа од T_{n-1}^+ за n, а ако је $D_n < 0$, T_n^+ ће бити једнака T_{n-1}^+. Дакле, добијамо

$$\begin{aligned} P\{T_n^+ = k\} &= \frac{u_n(k)}{2^n} = \frac{u_{n-1}(k-n)P\{D_n > 0\} + u_{n-1}(k)P\{D_n < 0\}}{2^{n-1}} = \\ &= \frac{u_{n-1}(k-n) + u_{n-1}(k)}{2^n}. \end{aligned}$$

Пример 5. Да бисмо илустровали рачунање вероватноћа на основу рекурзивне релације, узмимо обим узорка $n = 8$ и израчунајмо вероватноћу $P\{T^+ = 12\}$.

Имамо да је

$$\begin{aligned} P\{T^+ = 12\} &= \frac{u_8(12)}{2^8} = \frac{u_7(4) + u_7(12)}{2^8} = \frac{2 + u_6(5) + u_6(12)}{2^8} = \\ &= \frac{5 + u_5(6) + u_5(12)}{2^8} = \frac{10}{2^8} = 0.039. \end{aligned}$$

Приликом израчунавања коришћено је да се, за $n = 7$ збир 4 може добити од бројева 1 и 3 или 4, па је $u_7(4) = 2$; за $n = 6$ збир 5 од бројева $1, 4$ или $2, 3$ или 5, па је $u_6(5) = 3$; за $n = 5$ збир 6 од бројева $1, 5$ или $1, 2, 3$ или $2, 4$, а збир 12 од бројева $3, 4, 5$ или $1, 2, 4, 5$, па је $u_5(6) = 3$ и $u_5(12) = 2$.

Ако је обим узорка $n \geq 15$, Вилкоксонова T статистика ће имати приближно нормалну расподелу. У следећој табели ћемо приказати алтернативне хипотезе, критичне области и P - вредности када се користи апроксимација расподеле тест статистике нормалном расподелом уз корекцију непрекидности.

H_1	*Критична област*	*P-вредност*
$M > M_0$	$T^+ \geq \frac{n(n+1)}{4} + 0.5 + \sqrt{\frac{n(n+1)(2n+1)}{24}}t_\alpha$	$p_1 = 1 - F\left(\frac{t^+ - 0.5 - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}\right)$
$M < M_0$	$T^+ \leq \frac{n(n+1)}{4} - 0.5 + \sqrt{\frac{n(n+1)(2n+1)}{24}}t'_\alpha$	$p_2 = F\left(\frac{t^+ + 0.5 - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}\right)$
$M \neq M_0$	$T^+ \geq \frac{n(n+1)}{4} + 0.5 + \sqrt{\frac{n(n+1)(2n+1)}{24}}t_{\frac{\alpha}{2}}$ или $T^+ \leq \frac{n(n+1)}{4} - 0.5 - \sqrt{\frac{n(n+1)(2n+1)}{24}}t_{\frac{\alpha}{2}}$	$2 \cdot \min\{p_1, p_2\}$

2.2.3 Проблем истих вредности у низу апсолутних разлика

Приликом примене тестова рангова обично се претпоставља да је расподела обележја непрекидна. Када се начини ова претпоставка, вероватноћа да било које две случајне величине из узорка узму исту вредност је нула. Међутим, у пракси се исте вредности јављају, било зато што је расподела обележја заправо дискретна, било због грешака начињених приликом заокруживања. Расподела тест статистике није иста када има истих рангова, али је њихов ефекат мали и корекција је потребна само ако је присутан велики број истих вредности. Код узорака обима $n \geq 15$ где се користи стандардизована тест статистика врши се корекција дисперзије. Претпоставимо да има t истих вредност у узорку. Да су све ове вредности различите, били би им додељени рангови $s+1, s+2, ..., s+t$. Њихов просечан ранг је тада $s + \frac{t+1}{2}$ и сума квадрата ових рангова је тада једнака:

$$t(s + \frac{(t+1)^2}{2}) = t(s^2 + s(t+1) + \frac{(t+1)^2}{4}).$$

У случају различитих рангова, њихова сума квадрата би била једнака

$$\sum_{i=1}^{t}(s+i)^2 = ts^2 + st(t+1) + \frac{t(t+1)(2t+1)}{6}.$$

Присуство ових t истих вредности смањује суму квадрата рангова за

$$\frac{t(t+1)(2t+1)}{6} - \frac{t(t+1)^2}{6} = \frac{t(t+1)(t-1)}{12}.$$

Према томе, редукована дисперзија је једнака:

$$D(T^+|H_0) = \frac{n(n+1)(2n+1)}{24} - \frac{\sum\limits_{t} t(t^2-1)}{48}.$$

2.2.4 Вилкоксонов тест означених рангова за зависне узорке

Добијен је узорак $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ обима n. Елементи сваког пара X_i и Y_i су зависни, док су парови међусобно независни. Желимо да тестирамо хипотезу да је медијана разлике M_D једнака M_0, тј. $H_0(M_D = M_0)$. Претпостављамо да је расподела обележја $X - Y$ симетрична. Ако тестирамо хипотезу да је $M_D = 0$ тада је $m_D = 0$, тј. $H_0(m_X = m_Y)$, где су m_X, m_Y, редом, средње вредности обележја X и Y. Формирамо разлике $D_i = X_i - Y_i - M_0$ и рангирамо апсолутне разлике $|D_i|$, од најмање до највеће вредности, додељујући им бројеве од $1, 2, ...n$. Затим, као у случају Вилкоксоновог теста рангова за један узорак, рачунамо суму рангова придружених негативним разликама и суму рангова придружених негативним разликама. Критична област се одређује као у случају једног узорка.

Пример 6. У примеру 2 смо тестирали хипотезу $H_0(M_D = 0)$, где је $D = X - Y$ разлика поена студената на другом и првом колоквијуму. Алтернативна хипотеза је била облика $H_1(M_D > 0)$ тј. да бар половина студената има више поена на другом него на првом колоквијуму. Који резултати се добијају тестирањем нулте хипотезе помоћу Вилкоксоновог теста означених рангова за зависне узорке?

Желимо да тестирамо хипотезу $H_0(m_D = 0)$ (пошто је расподела обележја $X - Y$ симетрична) против алтернативне хипотезе $H_1(m_D > 0)$. Имамо низ разлика:

$0,\ 5,\ 8,\ 3,\ 7,\ 0,\ 6,\ 4,\ 1,\ -2,\ 7,\ -2,\ -3,\ 5,\ 0,\ 3,\ 2,\ 0,\ 7,\ 2,\ -6,\ 5,\ 5,\ 2$

Пошто су се јавиле четири нулте разлике, редукујемо $n = 20$. Дакле, имамо нови низ апсолутних разлика:

$5,\ 8,\ 3,\ 7,\ 6,\ 4,\ 1,\ 2,\ 7,\ 2,\ 3,\ 5,\ 3,\ 2,\ 7,\ 2,\ 6,\ 5,\ 5,\ 2$

Формирамо варијациони низ апсолутних разлика:

$1,\ 2,\ 2,\ 2,\ ,2,\ 2,\ 3,\ 3,\ 3,\ 4,\ 5,\ 5,\ 5,\ 5,\ 6,\ 6,\ 7,\ 7,\ 7,\ 8$

Рангирамо вредности варијационог низа:

$1,\ 4,\ 4,\ 4,\ 4,\ 4,\ 8,\ 8,\ 8,\ 10,\ 12.5,\ 12.5,\ 12.5,\ 12.5\ 15.5,\ 15.5,\ 18,\ 18,\ 18,\ 20$

Добијамо низ означених рангова:

$$12.5,\ 20,\ 8,\ 18,\ 15.5,\ 10,\ 1, -4,\ 18, -4, -8,\ 12.5,\ 8,\ 4,\ 18,\ 4, -15.5,\ 12.5,\ 12.5,\ 4$$

Добијамо да је сума рангова са позитивним знаком једнака $T^+ = 178.5$, а сума рангова са негативним знаком $T^- = 31.5$. Пошто је једноставније радити са мањим сумама, узмимо T^- за тест статистику. Критична област је облика $\{T^- \leq c\}$, P - вредност можемо добити у R - у помоћу функције *psignrank* која рачуна вредност функције расподеле тест статистике.

```
psignrank(31.5,20)
[1] 0.002110481
```

На основу добијене P - вредности, одбацујемо нулту хипотезу. Како је обим узорка већи од 15 а присутан је велики број истих вредности у низу апсолутних разлика, користићемо апроксимацију расподеле тест статистике нормалном расподелом уз редукцију дисперзије услед истих вредности и корекцију непрекидности. Математичко очекивање и дисперзија тест статистике су редом, једнаки

$$E(T^-) = 105, \ \ D(T^-) = 717.5.$$

Означимо са t_i број вредности у варијационом низу апсолутних разлика једнаких i. Добијамо $t_2 = 5, \ t_3 = 3, \ t_5 = 4, \ t_6 = 2, \ t_7 = 3$, па је сума за коју редукујемо дисперзију једнака $\sum\limits_t (t^3 - t) = 234$. Редукована дисперзија је тада једнака $D_{red}(T^-) = 483.5$, па је вредност стандардизоване тест статистике:

$$Z = \frac{T^- - E(T^-) + 0.5}{\sqrt{D_{red}(T^-)}} = -3.32,$$

P - вредност је тада једнака $P = F(-3.32) = 0.00045,$. Добија се знатно нижа P вредност него пре апроксимације.

Погледајмо још резултате тестирања у R - у. Функција *wilcox.exact* узима у обзир присуство нултих разлика и истих вредности у низу апсолутних разлика. Добијамо:

```
wilcox.exact(poeni2,poeni1,mu=0,alternative= "g",paired=T)

    Exact Wilcoxon signed rank test
data:  poeni2 and poeni1
V = 178.5, p-value = 0.002086
alternative hypothesis:  true mu is greater than 0
```

Дакле, у примеру 2 смо добили да су P- вредности t - теста и теста знакова редом једнаке 0.001478 и 0.0059, док код Вилкоксоновог теста означених рангова се добија $P = 0.002$ (без апроксимације нормалном расподелом). Према томе, можемо закључити, као у случају тестирања код једног узорка, да је P - вредност Вилкоксоновог теста означених рангова ближа P - вредности t - теста (када је t - тест најмоћнији), него што је P - вредност теста знакова.

Поглавље 3

Тестови медијане за два независна узорка

Нека су $(X_1, X_2, \ldots, X_m)$ и $(Y_1, Y_2, \ldots, Y_n)$ независни узорци. Желимо да тестирамо хипотезу да су медијане обележја X и Y једнаке тј. $H_0(M_X = M_Y)$. Формира се заједнички варијациони низ обима $N = m + n$.

3.1 Тест Ман-Витнија

Ман[1] - Витнијев[2]тест првенствено је конструисан за тестирање хипотезе о идентичним расподелама обележја X и Y тј. $H_0(F_X(x) = F_Y(x))$, али се може применити и за тестирање једнакости њихових медијана. Мен-Витнијев тест заснован је на идеји да одређена правилност која се уочава када се вредности обележја X и Y поређају у заједнички варијациони низ пружа информације о односу њихових расподела. Тако, у случајевима када је већина вредности обележја X у узорку већа од вредности обележја Y, или обратно, имамо очигледан показатељ против њиховог случајног поретка у варијационом низу и у прилог одбацивању нулте хипотезе о идентичним расподелама (в. [2], [3], [5], [7], [8], [10], [11]).

Мен-Витнијева статистика дефинише се као број пута да вредности обележја Y претходе вредностима обележја X у заједничком варијационом низу. Претпостављамо да су расподеле оба обележја непрекидне, па није потребно разматрати могућност $X_i = Y_j,\ i = 1, 2, ..., m,\ j = 1, 2, ..., n$. Дефинишимо случајне величине - индикаторе $D_{ij},\ i = 1, 2, ..., m,\ j = 1, 2, ..., n$ на следећи начин:

$$D_{ij} = \begin{cases} 1, & \text{ако је } Y_j < X_i \\ 0, & \text{ако је } Y_j > X_i \end{cases} . \qquad (*)$$

[1] *Henry Berthold Mann* **(1905-2000), амерички статистичар**

[2] *Donald Ransom Whitney* **(1915-2001), амерички статистичар**

Ман - Витнијева тест статистика је

$$U = \sum_{i=1}^{m}\sum_{j=1}^{n} D_{ij}$$

.

Када је

$$F_Y(x) \leq F_X(x), \text{ за свако } x,$$

$$F_Y(x) < F_X(x), \text{ за неко } x,$$

кажемо да је обележје Y стохастички веће од обележја X, што се означава са $Y \overset{ST}{>} X$. Према томе, за алтернативну хипотезу да је обележје Y стохастички веће од обележја X нулта хипотеза се одбацује при врло малим вредностима тест статистике U. Дефинишимо вероватноћу p на следећи начин:

$$p = P\{Y < X\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{x} f_X(x)f_Y(y)dxdy = \int_{-\infty}^{+\infty} F_Y(x)dF_X(x),$$

где су f_X и f_Y, редом, густине расподела обележја X и Y.

Сада можемо дефинисати алтернативне хипотезе у односу на вредност параметра p. Ако је $H_0(F_X(x) = F_Y(x))$ тачна за свако x, тада имамо да је

$$p = \int_{-\infty}^{+\infty} F_X(x)dF_X(x) = 0.5.$$

Ако је нпр. алтернативна хипотеза облика $H_1(F_Y(x) \leq F_X(x))$ што се може означити са $Y \overset{ST}{>} X$, тада је $H_1(p < 0.5)$. Према томе, параметарска варијанта нулте хипотезе о идентичним расподелама је $H_0(p = 0.5)$ против $H_1(p < 0.5)$.

Случајне величине D_{ij} дефинисане у (*) су Бернулијеве случајне величине са математичким очекивањем $E(D_j) = p$ и дисперзијом $D(D_{ij}) = p(1-p)$. Случајне величине D_{ij} су зависне, кад год су X_i и Y_j заједнички, па имамо

$cov(D_{ij}, D_{lk}) = 0,\ i \neq l,\ j \neq k,$

$cov(D_{ij}, D_{ik}) = E(D_{ij}D_{ik}) - E(D_{ij})E(D_{ik}) = p_1 - p^2,$

где је p_1

$$\begin{aligned} p_1 &= E(D_{ij}D_{ik}) = P\{Y_j < X_i, Y_k < X_i\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{x}\int_{-\infty}^{x} f_X(x)f_Y(y)f_Y(z)dxdydz = \\ &= \int_{-\infty}^{+\infty} (F_Y(x))^2 dF_X(x) \end{aligned}$$

$cov(D_{ij}, D_{lj}) = p_2 - p^2,$

где је p_2

$$\begin{aligned} p_2 &= P\{X_i > Y_j, X_i > Y_k\} = \int\limits_{-\infty}^{+\infty}\int\limits_{y}^{+\infty}\int\limits_{y}^{+\infty} f_X(x) f_Y(y) f_Y(z) dx dy dz = \\ &= \int\limits_{-\infty}^{+\infty} (1 - F_X(y))^2 dF_Y(y) \end{aligned}$$

Како смо дефинисали тест статистику U као линеарну комбинацију mn случајаних величина D_{ij}, њено математичко очекивање и дисперзија су једнаки

$$\begin{aligned} E(U) &= \sum_{i=1}^{m}\sum_{j=1}^{n} E(D_{ij}) = mn \cdot p \\ D(U) &= \sum_{i=1}^{m}\sum_{j=1}^{n} D(D_{ij}) + \sum_{i=1}^{m}\sum_{1\le j\neq k\le n}\sum cov(D_{ij}, D_{ik}) \sum_{j=1}^{n}\sum_{1\le i\neq l\le m}\sum cov(D_{ij}, D_{lj}) + \\ &+ \sum_{1\le i\neq l\le m}\sum\sum_{1\le j\neq k\le n}\sum cov(D_{ij}, D_{lk}) = \\ &= mn\cdot (1-p) + mn(n-1)(p_1 - p^2) + mn(m-1)(p_2 - p^2) = \\ &= mn(p - p^2(N-1) + (n-1)p_1 + (m-1)p_2) \end{aligned}$$

Непристрасна и постојана оцена параметра p је $\frac{U}{mn}$. Представимо у следећој табели алтернативне хипотезе и одговарајуће критичне области.

Алтернативна хипотеза		*Критична област*
$p < 0.5$	$F_Y(x) \le F_X(x)$	$U - \frac{mn}{2} < k_1$
$p > 0.5$	$F_Y(x) \ge F_X(x)$	$U - \frac{mn}{2} > k_2$
$p \neq 0.5$	$F_Y(x) \neq F_X(x)$	$U - \frac{mn}{2} > k_3$

Да бисмо одредили величину теста, потребно је наћи нулту расподелу тест статистике U. Ако је тачна H_0, сваки од $\binom{m+n}{n}$ распореда елемената узорака у заједничком варијационом низу јавља се са истом вероватноћом, па је

$$P\{U = u\} = \frac{r_{m,n}(u)}{\binom{m+n}{m}},$$

где је $r_{m,n}(u)$ укупан број могућих распореда m вредности обележја X (m X-ова) и n вредности обележја Y (n Y-а), тако да у заједничком варијационом низу Y претходи X u пута.

Крајње вредности тест статистике U су 0 и mn које се добијају када свако X претходи Y и обратно. Нулта расподела тест статистике U симетрична је око средње вредности $\frac{mn}{2}$ што се може закључити на следећи начин. За сваки одређени распоред z који се састоји од m x-ова и n y-а, дефинишимо конјуговани распоред z' као низ z написан уназад. Свако y које претходи x у низу z је испред x у низу z', тако да ако је вредност Мен-Витнијеве статистике једнака u за распоред z, $mn-u$ је вредност за z'. Према томе, ако је тачна нулта хипотеза, имамо да је $r_{m,n}(u) = r_{m,n}(mn-u)$, или еквивалентно

$$P\{U-\frac{mn}{2}=u\} = P\{U=\frac{mn}{2}+u\} = P\{U=mn-(\frac{mn}{2}+u)\} = P\{U-\frac{mn}{2}=-u\}.$$

Због овог својства симетричности расподеле, потребно је наћи само критичне вредности левог репа, за било за коју алтернативну хипотезу. Дефинишимо статистику U' као укупан број пута да X претходи Y у заједничком варијационом низу или

$$U' = \sum_{i=1}^{m}\sum_{j=1}^{n}(1-D_{ij})$$

. Представимо сада критичне области у следећој табели.

Алтернативна хипотеза		*Критична област*
$p < 0.5$	$F_Y(x) \le F_X(x)$	$U \le c_\alpha$
$p > 0.5$	$F_Y(x) \ge F_X(x)$	$U' \le c'_\alpha$
$p \ne 0.5$	$F_Y(x) \ne F_X(x)$	$U \le c_{\frac{\alpha}{2}}$ или $U' \le c'_{\frac{\alpha}{2}}$

Да бисмо одредили критичну вредност c_α за свако m и n, можемо набројати све случајеве почевши од $u=0$ све док не набројимо барем $\alpha \cdot \binom{m+n}{m}$ случајева.

Пример 7. Прикажимо у следећој табели распореде X-ова и Y-а којим се добијају најмање вредности u, где су обими узорака, редом, $m=4$ и $n=5$.

Распоред	*Вредности*	*Бр.распореда*
$XXXXYYYYY$	0	1
$XXXYXYYYY$	1	1
$XXYXXYYYY$, $XXXYYXYYY$	2	2
$XYXXXYYYY$, $XXYXYXYYY$, $XXXYYYXYY$	3	3

Према томе, имамо да је $P\{U \le 0\} = \frac{1}{126} = 0.008$, $P\{U \le 1\} = \frac{2}{126} = 0.016$, $P\{U \le 2\} = \frac{4}{126} = 0.032$, $P\{U \le 3\} = \frac{7}{126} = 0.056$. Дакле, критичне области за алтернативну хипотезу $H_1(F_Y(x) \leq F_X(x))$ са праговима значајности 0.01 и 0.05 би биле, редом, $\{U \leq 0\}$ и $\{U \leq 2\}$.

Како се повећавају обими узорака m и n, укупан број свих распореда расте, што може довести да се неки од распореда пропусти. Према томе, за конструкцију табела критичних вредности потребан је систематичнији метод од горе наведеног метода набрајања могућих распореда X - ова и Y - а. У ту сврху извешћемо једну врло једноставну и корисну рекурентну релацију за закон расподеле Ман - Витнијеве тест статистике. Размотримо низ $m+n$ X - ова и Y - а који се добио додавањем једног X - а или Y - а на крај низа од $m+n-1$ X - ова и Y - а. Ако низ дужине $m+n-1$ садржи m X - ова и $n-1$ Y - а, овај додатни члан низа мора бити Y. У том случају, укупан број пута да Y претходи X остаје непромењен, тј. једнак u. С друге стране, ако у низу дужине $m+n-1$ има $m-1$ X - ова и n Y - а, додатни члан низа ће бити X што значи да ће сви остали Y-и претходити овом новом X и вредност тест статистике ће бити $u+n$. Ове две могућности међусобно су искључиве. Према томе, рекурентну релацију можемо изразити на следећи начин.

$$r_{m,n}(u) = r_{m,n-1}(u) + r_{m-1,n}(u-n)$$

и

$$\begin{aligned} p_{m,n}(u) &= P\{U=u\} = \frac{r_{m,n-1}(u) + r_{m-1,n}(u-n)}{\binom{m+n}{m}} = \\ &= \frac{n}{m+n} \frac{r_{m,n-1}(u)}{\binom{m+n-1}{n-1}} + \frac{m}{m+n} \frac{r_{m-1,n}(u-n)}{\binom{m+n-1}{m-1}} \end{aligned}$$

или

$$(m+n) \cdot p_{m,n}(u) = n \cdot p_{m,n-1}(u) + m \cdot p_{m-1,n}(u-n).$$

Ова рекурзивна релација важи за све $u = 0, 1, ..., mn$ и све $m, n \in N$, при чему су почетне вредности, за све $i = 1, 2, ..., m,\ j = 1, 2, ..., n$ дефинисани са

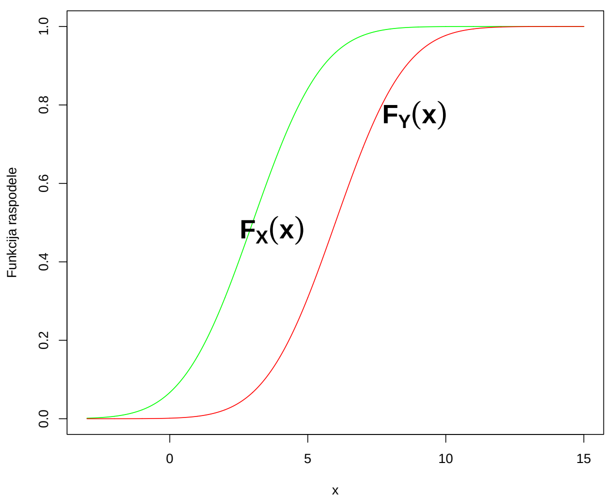
$$\begin{aligned} r_{i,j}(u) &= 0 \text{ за свако } u < 0, \\ r_{i,0}(0) &= 1,\ r_{0,j}(0) = 1, \\ r_{i,0}(u) &= 0 \text{ за свако } u \neq 0, \\ r_{0,j}(u) &= 0 \text{ за свако } u \neq 0. \end{aligned}$$

Даље, једноставности ради, узорак мањег обима може се сматрати да представља вредности обележја X, па су таблице потребне само за $m \leq n$ и критичне вредности левог репа. Када су обими узорака m и n велики, може се користити апроксимација расподеле тест статистике U нормалном расподелом, на основу централне граничне теореме за зависне случајне величине. При тачној нултој хипотези стандардизована статистика U ће имати приближно нормирану нормалну расподелу када је $\frac{m}{n}$ константно, $m, n \to \infty$.

При тачној нултој хипотези $H_0(F_Y(x) = F_X(x))$ тј. $H_0(p = 0.5)$ имамо да су $p_1 = p_2 = \frac{1}{3}$, па добијамо $E(U) = \frac{mn}{2}$, $D(U) = \frac{mn(N+1)}{12}$. Према томе, већ за узорке обима 6 (једнаки обими узорка) може се користити апроксимација расподеле тест статистике нормалном расподелом (Ман - Витни), уз корекцију непрекидности.

Нултој хипотези $H_0(F_X(x) = F_Y(x))$ је еквивалентна хипотеза да су медијане два узорка једнаке, тј. $M_X = M_Y$. Такође, исто важи и за алтернативне хипотезе, односно у случају алтернативне хипотезе $H_1(F_X(x) \leq F_Y(x))$ имамо да је $H_1(M_X > M_Y)$, односно $H_1(M_X < M_Y)$ у случају $H_1(F_X(x) \geq F_Y(x))$. Према томе, тест Ман -Витнија може се користити и за тестирање једнакости медијана код два независна узорка.

Пример 8 Представимо графички функцију расподеле случајне величине X са нормалном расподелом $\mathcal{N}(3, 2)$ и функцију расподеле случајне величине Y са нормалном расподелом $\mathcal{N}(6, 2)$. Можемо уочити да је $F_X(x) \geq F_Y(x)$.



Слика 3.1: X је стохастички мање од Y

Пример 9. Мерена је дневна енергетска потрошња (у MJ) 13 случајно одабраних мршавих и 9 дебелих жена. Добијени су следећи резултати:

Мршаве жене: 7.53, 7.48, 8.08, 8.09, 10.15, 8.40, 10.88, 6.13, 7.90, 7.05, 7.48, 7.58, 8.11

Дебеле жене: 9.21, 11.51, 12.79, 11.85, 9.97, 8.79, 9.69, 9.68, 9.19

Са прагом значајности $\alpha = 0.05$ тестирати хипотезу да просечно половина дебелих жена троши више енергије.

Нека обележје X представља енергетску потрошњу мршавих жена, а обележје Y енергетску потрошњу дебелих жена. Проверимо да ли су расподеле обележја нормалне. Помоћу наредби у статистичком софтверу R - у, као у претходним примерима добијамо $Q-Q$ дијаграм нормалне расподеле за енергетске потрошње мршавих и дебелих жена.

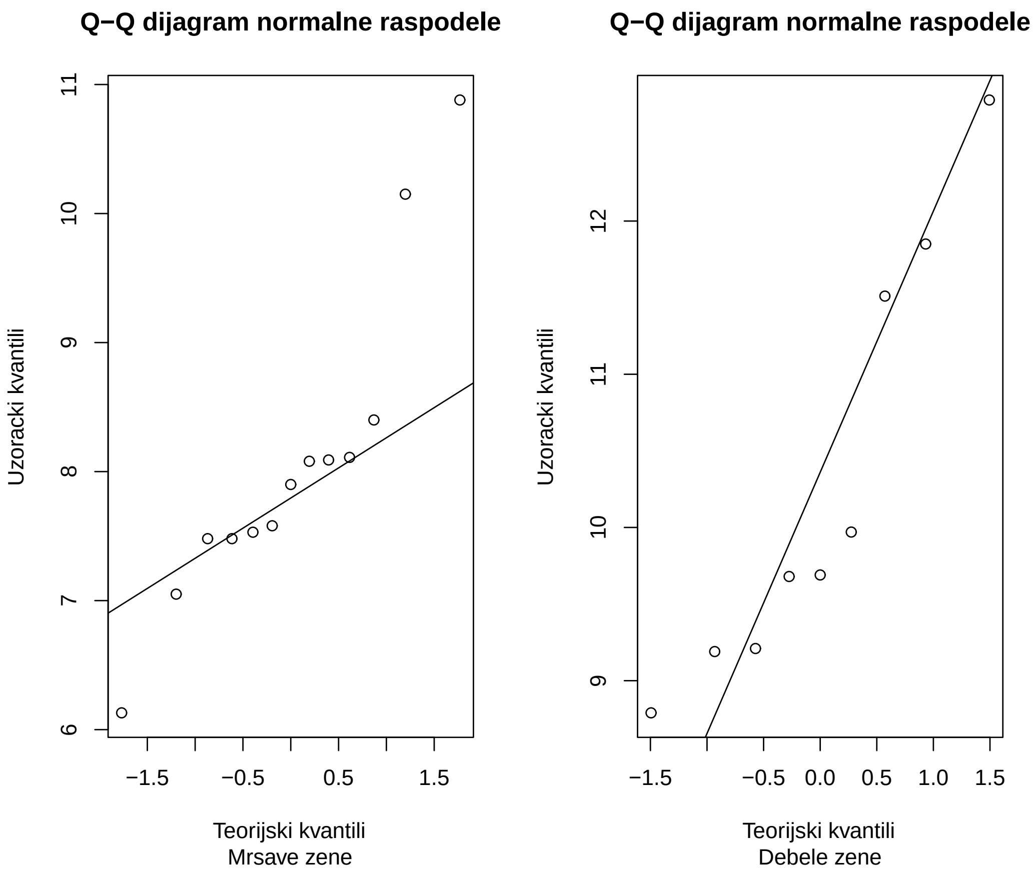
```
attach(energy)
mrsave<-expend[stature=="lean"]
debele<-expend[stature=="obese"]
par(mfrow=c(1,2))
qqnorm(mrsave,xlab="Teorijski kvantili",ylab=
"Uzoracki kvantili",main="Q-Q dijagram normalne raspodele")
qqline(mrsave) title(sub="Mrsave zene")
qqnorm(debele,xlab="Teorijski kvantili",ylab=
"Uzoracki kvantili",main="Q-Q dijagram normalne raspodele")
qqline(debele)
title(sub="Debele zene")
```

Добијени графици налазе се на слици 3.2.

Примећујемо на графику енергетске потрошње мршавих жена да је већина тачака је сконцентрисана око праве, али, с друге стране, неколико тачака значајно одступа од праве. Што се тиче графика енергетске потрошње дебелих жена,тачке нису тако близу праве него код мршавих жена, али нема ни толиких одступања. Значи, не очекује се добро слагање података са нормалном расподелом, али с обзиром да су узорци мали, поставља се питање о валидности закључака. Ослонићемо се на тест Лилифор - Колмогорова .

```
lillie.test(mrsave)

    Lilliefors (Kolmogorov-Smirnov) normality test
data:  mrsave
D = 0.2551, p-value = 0.02048
```



Слика 3.2: $Q - Q$ дијаграм нормалне расподеле за енергетску потрошњу

```
lillie.test(debele)

    Lilliefors (Kolmogorov-Smirnov) normality test

data:  debele
D = 0.2594, p-value = 0.08203
```

Дакле, на основу добијених вредности, у случају мршавих жена одбацујемо нулту хипотезу да је расподела обележја X нормална, док у случају дебелих жена прихватамо да расподела Y има нормалну расподелу, али са резервом.

Да бисмо закључили да ли има смисла применити t - тест, проверићемо, помоћу функције *sim.test* да ли су расподеле обележја симетричне.

```
sim.test(mrsave)

Uzoracki koeficijent asimetrije 1.161499
Realizovana vrednost test statistike:  1.871536
Kriticna vrednost:  1.959964
```

```
Raspodela obelezja je simetricna

sim.test(debele)

Uzoracki koeficijent asimetrije 0.8490838
Realizovana vrednost test statistike:  1.201384
Kriticna vrednost:  1.959964
Raspodela obelezja je simetricna
```

На основу добијених резултата прихватамо хипотезу да су расподеле обележја симетричне. Према томе, тестирамо хипотезу да су средње вредности ова два обележја једнаке тј. $H_0(m_X = m_Y)$ против алтернативне хипотезе $H_1(m_X < m_Y)$. Прво ћемо применити тест Ман - Витнија, а затим t - тест за тестирање нулте хипотезе.

Формирамо заједнички варијациони низ и уписујемо X ако је елемент узорка за обележје X, а Y ако је елемент узорка за обележје Y. Добијамо низ X - ова и Y - а

$X, X, X, X, X, X, X, X, X, X, X, Y, Y, Y, Y, Y, Y, X, X, Y, Y, Y$

Укупан број пута да Y-и претходе X - овима је $U = 6 + 6 = 12$. Критична област је облика $\{U \leq c\}$, где је c критична вредност коју можемо добити помоћу функције *qwilcox* у *R*-у.

```
qwilcox(0.05,13,9)
[1] 34
```

Дакле, пошто вредност тест статистике упада у критичну област $W = (0, 34]$, одбацујемо нулту хипотезу о једнакости средњих вредности обележја X и Y.

Резултате можемо директно добити помоћу функције *wilcox.exact* у R - у:

```
wilcox.exact(mrsave,debele,alternative="l")

    Exact Wilcoxon rank sum test
data:  mrsave and debele
W = 12, p-value = 0.0005287
alternative hypothesis:  true mu is less than 0
```

На основу P - вредности, не само да одбацујемо нулту хипотезу већ и закључујемо да у аргументи против ње врло јаки.

Сада ћемо тестирати нулту хипотезу помоћу t - теста за независне узорке.

Добијамо да су узорачке средине и дисперзије, редом, једнаке

$$\overline{x}_m = 8.067, \quad \widetilde{s}^2_m = 1.533, \quad \overline{y}_n = 10.298, \quad \widetilde{s}^2_n = 1.954.$$

Тестирајмо прво једнакост дисперзија помоћу Фишеровог теста (у R - у се користи функција *var.test*).

```
var.test(mrsave, debele)

    F test to compare two variances
data:  mrsave and debele F = 0.7844, num df = 12, denom df = 8,
p-value = 0.6797
alternative hypothesis:  true ratio of variances is not
equal to 1 95 percent confidence interval:
 0.1867876 2.7547991
sample estimates:
ratio of variances
    0.784446
```

Дакле, прихватамао хипотезу о једнакости дисперзија, па ћемо применити t - тест за независне узорке са једнаким дисперзијама.

```
t.test(mrsave, debele, var.equal=T, alternative="l")

   Two Sample t-test
data:  mrsave and debele t = -3.9456, df = 20, p-value = 0.0003995
alternative hypothesis:  true difference in means is
less than 0 95 percent confidence interval:
   -Inf -1.256118
sample estimates:
mean of x mean of y
  8.066154 10.297778
```

Дакле, P - вредности оба теста су врло блиске, па бисмо донели исту одлуку о одбацивању нулте хипотезе на основу оба теста.

3.2 Линеарне статистике ранга

Многи тестови рангова за два узорка представљају линеарне комбинације одређених индикатора за заједнички уређени узорак. Такве функције називају се *линеарне статистике ранга* [2] .

Нека су $(X_1, X_2, \ldots, X_m)$ и $(Y_1, Y_2, \ldots, Y_n)$ два независна узорка и нека су функције расподела обележја, редом, F_X и F_Y. Претпоставља се да су расподеле обележја непрекидне. Желимо да тестирамо хипотезу да оба узорка потичу из исте популације, тј. $H_0(F_X = F_Y)$. Формира се варијациони низ од $m + n = N$ елемената оба узорка. Нека је $Z = (Z_1, Z_2, \ldots, Z_N)$, где је $Z_i = 1$ ако је i-та случајна величина у варијационом низу X, а $Z_i = 0$ ако је Y.

Пример 10. Имамо два узорка $(x_1, x_2, x_3, x_4) = (2, 9, 3, 4)$ и $(y_1, y_2, y_3) = (1, 6, 10)$. Заједнички варијациони низ је $(1, 2, 3, 4, 6, 9, 10)$ или $(y_1, x_1, x_3, x_4, y_2, x_2, y_3)$ и одговарајући вектор Z узима вредности $(0, 1, 1, 1, 0, 1, 0)$.

Означићемо са $T_N(Z) = \sum\limits_{i=1}^{N} a_i Z_i$ линеарну статистику ранга. Навешћемо сада три теореме о својствима расподеле линеарне статистике ранга, које ће нам бити потребне у даљем раду.

Теорема 1 *Ако је тачна нулта хипотеза* $H_0(F_X(x) = F_Y(x) = F(x))$ *за свако* x, *тада за свако* $i = 1, 2, ..., N$ *важи:*

$$E(Z_i) = \frac{m}{n}, \quad D(Z_i) = \frac{mn}{N^2}, \quad cov(Z_i, Z_j) = -\frac{mn}{N^2(N-1)}.$$

Доказ: Како $Z_i : \left(\begin{array}{cc} 0 & 1 \\ \frac{n}{N} & \frac{m}{N} \end{array} \right)$, тј. $Z_i : B(1, \frac{m}{N})$, имамо да је

$$E(Z_i) = \frac{m}{n}, \quad D(Z_i) = \frac{mn}{N^2}.$$

$$E(Z_i Z_j) = P\{Z_i = 1, Z_j = 1\} = \frac{\binom{m}{2}}{\binom{N}{2}} = \frac{m(m-1)}{N(N-1)},$$

па је

$$\begin{aligned} cov(Z_i, Z_j) &= E(Z_i Z_j) - E(Z_i)E(Z_j) = \\ &= \frac{m(m-1)}{N(N-1)} - (\frac{m}{N})^2 = -\frac{mn}{N^2(N-1)}. \end{aligned}$$

Теорема 2 *Ако је тачна нулта хипотеза* $H_0(F_X(x) = F_Y(x) = F(x))$ *за свако* x, *тада важи:*

$$E(T_N) = m\sum_{i=1}^{N}\frac{a_i}{N}, \quad D(T_N) = \frac{mn}{N^2(N-1)}(N\sum_{i=1}^{N}a_i^2 - (\sum_{i=1}^{N}a_i)^2).$$

Доказ: На основу теореме 1 имамо да је $E(T_N) = \sum\limits_{i=1}^{N} a_i E(Z_i) = m\sum\limits_{i=1}^{N}\frac{a_i}{N}$.

$$\begin{aligned} D(T_N) &= \sum_{i=1}^{N}a_i^2 D(Z_i) + \sum_{i,j}a_i a_j cov(Z_i, Z_j) = \\ &= \frac{mn}{N^2}\sum_{i=1}^{N}a_i^2 - \frac{mn}{N^2(N-1)}\sum_{i,j}a_i a_j = \\ &= \frac{mn}{N^2(N-1)}(N\sum_{i=1}^{N}a_i^2 - \sum_{i=1}^{N}a_i^2 - \sum_{i,j}a_i a_j) = \\ &= \frac{mn}{N^2(N-1)}(N\sum_{i=1}^{N}a_i^2 - (\sum_{i=1}^{N}a_i)^2). \end{aligned}$$

Како је $\binom{m+n}{m} = \binom{N}{m}$ укупан број свих вектора Z, тј. различитих распореда m јединица и n нула, укупно има $\binom{N}{m}$ свих могућих исхода, при чему се сваки исход јавља са истом вероватноћом $\frac{1}{\binom{N}{m}}$. Имамо да је

$$P\{T_N(Z) = k\} = \frac{r_{m,n}(k)}{\binom{N}{m}},$$

где је $r_{m,n}(k)$ број могућих распореда m вредности обележја X и n вредности обележја Y тако да је $T_N(Z) = k$.

Расподела статистике $T_N(Z)$ симетрична је око своје средње вредности μ, ако је, за свако $k \neq 0 : P\{T_N(Z) - \mu = k\} = P\{T_N(Z) - \mu = -k\}$.

Ако за сваки вектор Z који се састоји од m јединица и n нула постоји коњуговани вектор Z' m јединица и n нула такав да кад год је $T_N(Z) = \mu + k$ имамо $T_N(Z') = \mu - k$, расподела статистике $T_N(Z)$ је симетрична. Другим речима, услов за симетричност расподеле линеарне статистике ранга је $T_N(Z) + T_N(Z') = 2\mu$.

Теорема 3 *Ако је тачна нулта хипотеза, расподела статистике* $T_N(Z)$ *је симетрична око своје средње вредности* $\mu = \frac{m}{N}\sum\limits_{i=1}^{N} a_i$ *ако тежински коефицијенти* a_i *задовољавају услов* $a_i + a_{N-i+1} = c, c = const.$ *за* $i = 1, 2, ..., N$.

Доказ: За сваки вектор $Z = (Z_1, Z_2, \ldots, Z_N)$ m јединица и n нула дефинишимо конјуговани вектор $Z' = (Z'_1, Z'_2, \ldots Z'_N)$, где је $Z'_i = Z_{N-i+1}$. Тада је

$$\begin{aligned} T_N(Z) + T_N(Z') &= \sum_{i=1}^{N} a_i Z_i + \sum_{i=1}^{N} a_i Z_{N-i+1} = \\ &= \sum_{i=1}^{N} a_i Z_i + \sum_{j=1}^{N} a_{N-j+1} Z_j = \\ &= \sum_{i=1}^{N} (a_i + a_{N-i+1}) Z_i = c \sum_{i=1}^{N} Z_i = cm. \end{aligned}$$

Како је $ET_N(Z) = ET_N(Z')$, имамо да је $cm = 2\mu$ или $c = \frac{2\mu}{m} = 2\sum_{i=1}^{N} \frac{a_i}{N}$.

3.3 Вилкоксонов тест суме рангова

Рангови вредности обележја X у заједничком варијационом низу углавном ће бити већи од рангова вредности обележја Y ако је медијана обележја X већа од медијане обележја Y. Према томе, Вилкоксон је 1945. године је предложио тест на основу кога се прихвата алтернативна хипотеза $M_X > M_Y$ ако је сума рангова вредности обележја X сувише велика, односно алтернативна хипотеза $M_X < M_Y$ ако је сума рангова вредности обележја X сувише мала (в. [2], [3], [5], [7], [8], [10], [11]).

Вилкоксонова статистика је линеарна статистика ранга, где је $a_i = i$, тј.

$$W_N = \sum_{i=1}^{N} i \cdot Z_i.$$

На основу теореме 2, ако нема истих вредности у заједничком варијационом низу, имамо да је

$$E(W_N) = \frac{m(N+1)}{2}, \quad D(W_N) = \frac{mn(N+1)}{12}.$$

Ако је $m \leq n$, статистика W_N има минималну вредност $\sum_{i=1}^{m} i = \frac{m(m+1)}{2}$ и максималну вредност $\sum_{i=N-m+1}^{N} i = \frac{m(2N-m+1)}{2}$.

На основу теореме 3, како је $a_i + a_{N-i+1} = N + 1$, за $i = 1, 2, ..., N$, тест статистика је симетрична око своје средње вредности.

На основу претходно реченог о расподели линеарних статистика ранга, имамо да је нулта расподела тест статистике W_N

$$p_{m,n}(k) = P\{W_N = k\} = \frac{r_{m,n}(k)}{\binom{N}{m}}.$$

У следећој табели су, за дати праг значајности α, представљене алтернативне хипотезе, одговарајуће критичне области и P - вредности за реализовану вредност тест статистике W_0.

Алтернативна хипотеза	*Критична област*	P *- вредност*
$M_X > M_Y$	$W_N \geq w_\alpha$	$p_1 = P\{W_N \geq W_0\}$
$M_X < M_Y$	$W_N \leq w'_\alpha$	$p_2 = P\{W_N \leq W_0\}$
$M_X \neq M_Y$	$W_N \geq w_{\frac{\alpha}{2}}$ или $W_N \leq w'_{\frac{\alpha}{2}}$	$2\min\{p_1, p_2\}$

Пример 11. Дат је узорак обима $m = 3$ за обележје X и узорак обима $n = 4$ за обележје Y. Одредимо расподелу тест статистике. Имамо $\binom{7}{3}$ различитих распореда 3 јединице и 4 нуле у вектору Z. Тест статистика ће узимати вредности између 6 и 18, симетрично око 12. Добијамо

Вредности W_7	Рангови X-ова	Број додела $r_{3,4}$
18	$5, 6, 7$	1
17	$4, 6, 7$	1
16	$3, 6, 7;\ 4, 5, 7$	2
15	$2, 6, 7;\ 3, 5, 7;\ 4, 5, 6$	3
14	$1, 6, 7;\ 2, 5, 7;\ 3, 4, 7\ 3, 5, 6$	4
13	$1, 5, 7;\ 2, 4, 7;\ 2, 5, 6;\ 3, 4, 6$	4
12	$1, 4, 7;\ 2, 3, 7;\ 1, 5, 6;\ 2, 4, 6;\ 3, 4, 5$	5

.

Дакле, закон расподеле тест статистике је

$$P\{W_7 = k\} = \begin{cases} \frac{1}{35}, & k = 6,7,17,18 \\ \frac{2}{35}, & k = 8,16 \\ \frac{3}{35}, & k = 9,15 \\ \frac{4}{35}, & k = 10,11,13,14 \\ \frac{5}{35}, & k = 12 \end{cases}.$$

Тако, нпр. $P\{W_7 \geq 16\} = \frac{4}{35} = 0.114$.

За генерисање расподеле тест статистике можемо користити рекурзивну релацију изведену код Вилкоксоновог теста означених рангова и Ман - Витни теста. Имамо да важи

$$r_{m,n}(k) = r_{m-1,n}(k-N) + r_{m,n-1}(k)$$

и

$$p_{m,n}(k) = \frac{r_{m-1,n}(k-N) + r_{m,n-1}(k)}{\binom{N}{m}}$$

или

$$(m+n)p_{m,n}(k) = mp_{m-1,n}(k-N) + np_{m,n-1}(k).$$

За веће обиме узорка, $m \leq n, N \geq 12$, може се користити апроксимација расподеле тест статистике нормалном расподелом уз корекцију непрекидности. Уколико је у заједничком варијационом низу присутан већи број истих вредности, врши се корекција дисперзије тест статистике. Код Вилкоксоновог теста означених рангова видели смо да исте вредности доводе до смањења суме рангова за $\sum\limits_t \frac{t(t^2-1)}{12}$, где је t број истих вредности у заједничком варијационом низу. Заменом овог резултата у дисперзију тест статистике добијамо редуковану дисперзију

$$\begin{aligned} D(W_N) &= \frac{mn}{N^2(N-1)}[N(\frac{N(N+1)(2N+1)}{6} - \frac{\sum\limits_t t(t^2-1)}{12}) - (\frac{N(N+1)}{2})^2] = \\ &= \frac{mn(N+1)}{12} - \frac{mn\sum\limits_t t(t^2-1)}{12N(N-1)}. \end{aligned}$$

Вилкоксонова тест статистика суме рангова еквивалентна је Ман - Витнијевој тест статистици, пошто постоји линеарна веза између ове две тест статистике.

Ман - Витнијева тест статистика U дефинисана је као укупан број пута да Y претходи X у заједничком варијационом низу тј.

$$U=\sum_{i=1}^{m}\sum_{j=1}^{n}D_{ij}=\sum_{i=1}^{m}(D_{i1}+D_{i2}+...+D_{in}),$$

где је

$$D_{ij}=\begin{cases}1, & \text{ако је } Y_j<X_i\\ 0, & \text{ако је } Y_j>X_i\end{cases}.$$

Дакле, $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}D_{ij}$ једнака је или суми вредности j за које је $Y_j<X_i$ или рангу X_i редукованом за n_i, број X- ова који су мањи или једнаки од X_i. Према томе, имамо да је

$$\begin{aligned}U &= \sum_{i=1}^{m}(r(X_i)-n_i)=\sum_{i=1}^{m}r(X_i)-(n_1+n_2+...+n_m)=\\ &= \sum_{i=1}^{N}i\cdot Z_i-(1+2+...+m)=\sum_{i=1}^{N}i\cdot Z_i-\frac{m(m+1)}{2}=W_N-\frac{m(m+1)}{2}.\end{aligned}$$

Пример 12. У примеру 9 смо тестирали нулту хипотезу да је $H_0(m_X\neq m_Y)$, где обележја X и Y представљају енергетску потрошњу мршавих и дебелих жена. Алтернативна хипотеза је била облика $(m_X<m_Y)$, односно да дебеле жене просечно троше више енергије. Који резултати се добијају тестирањем нулте хипотезе помоћу Вилкоксоновог теста суме рангова?

Заједнички варијациони низ обједињеног узорка је

$$6.13,\ 7.05,\ 7.48,\ 7.48,\ 7.53\ 7.58,\ 7.90,\ 8.08,\ 8.09,\ 8.11,\ 8.40,\ 8.79,$$

$$9.19,\ 9.21,\ 9.68,\ 9.69,\ 9.97,\ 10.15,\ 10.88,\ 11.51,\ 11.85,\ 12.79$$

Према томе, елементима варијационог низа придружујемо рангове и изнад сваког ранга пишемо X ако се односи на елемент из узорка за обележје X, а са Y ако се односи на елемент из узорка за обележје Y.

$$\overset{X}{1},\ \overset{X}{2},\ \overset{X}{3.5},\ \overset{X}{3.5},\ \overset{X}{5},\ \overset{X}{6},\ \overset{X}{7},\ \overset{X}{8},\ \overset{X}{9},\ \overset{X}{10},\ \overset{X}{11},\ \overset{Y}{12},\ \overset{Y}{13},\ \overset{Y}{14},\ \overset{Y}{15},\ \overset{Y}{16},\ \overset{Y}{17},\ \overset{X}{18},\ \overset{X}{19},\ \overset{Y}{20},\ \overset{Y}{21},\ \overset{Y}{22}$$

Када саберемо рангове елемената из узорка за обележје X, што је реализована вредност тест статистике, добијамо $W_{22}=103$. Критична област је облика $\{W_N\leq c\}$, где се критична вредност c може добити на основу везе са статистиком Ман - Витнија. У примеру 9 добили смо да је критична област за тест Ман - Витнија једнака 34. Према томе, $c=34+\frac{13\cdot 14}{2}=125$. Како вредност тест статистике упада у критичну област, одбацујемо нулту хипотезу.

Поглавље 4

Поређење тестова

Један од циљева овог рада је приказати одређене непараметарске тестове као одговарајуће алтернативе t - тесту за тестирање хипотезе о средњој вредности код непрекидних симетричних расподела. Да бисмо то постигли, поредићемо функције моћи и ефикасност ових тестова за узорке из различитих непрекидних симетричних расподела [1].

4.1 Рачунање функције моћи теста

Узорци су генерисани помоћу Монте Карло метода коришћењем готових функција у програмском језику R. За добијене узорке рачунате су реализоване вредности тест статистика разматраних тестова, критичне вредности и бележена је одлука о одбацивању или неодбацивању нулте хипотезе. Пошто је средња вредност расподеле одабрана тако да нулта хипотеза није тачна, релативни удео узорака који су довели до одбацивања нулте хипотезе представљао је моћ теста. Вредности функције моћи тестова рачунате су помоћу програма чији код се налази у Прилогу. Затим су функције моћи тестова графички представљене за различите расподеле и обиме узорака.

4.2 Рачунање асимптотске релативне ефикасности теста

У уводу било је речи о асимптотској релативној ефикасности тестова у основним цртама. Сада ћемо се позабавити детаљима око израчунавања ARE. У наставку наводимо дефиницију и теореме потребне за даљи рад.

Нека су дате две тест статистике T_n и T_n^* за узорак обима n и нека су обе статистике постојане за тестирање нулте хипотезе $H_0(\theta=\theta_0)$ против алтернативне хипотезе $H_1(\theta\in\Theta\backslash\{\theta_0\})$ тј.

$$\lim_{n\to\infty} P_W[T_n(\theta)]=1,\quad \lim_{n\to\infty} P_W[T_n^*(\theta)]=1$$

Претпостављамо даље да важи $\{\theta_1,\theta_2,\ldots,\theta_n\ldots\}\subset\Theta$ и $\lim\limits_{n\to\infty}\theta_n=\theta_0$ (алтернативне хипотезе су облика $\theta=\theta_i,\ i=1,2,\ldots)$. Тако у случају алтернативне хипотезе $(\theta>\theta_0)$, узимамо опадајући низ бројева $(\theta_1,\theta_2,\ldots)$ који конвергира ка θ_0. Ако свако θ_i одређује расподелу тест статистике, можемо рећи да се расподела при тачној алтернативној хипотези све више приближава нултој расподели када $n\to\infty$. При овим условима наводимо формалну дефиницију ARE теста T^* у односу на T.

Дефиниција 2 . *Нека су $P_{W_n}(\theta)$ и $P^*_{W_n}(\theta)$ функције моћи двају тестова T и T^*, против фамилије алтернативних хипотеза облика $(\theta=\theta_i)$ и нека је θ_0 вредност параметра θ на основу нулте хипотезе. Нека T и T^* имају исти праг значајности α. Размотримо низ алтернатива (θ_n) и низ $(n^*)=(h(n))$ природних бројева, тако да је*

$$\lim_{n\to\infty} P_{W_n}(\theta_n)=\lim_{n\to\infty} P^*_{W_{n^*}}(\theta_n),$$

где се претпоставља да обе граничне вредности постоје и нису једнаке нули или јединици. Тада је асимптотска релативна ефикасност ARE теста T^* у односу на тест T једнака

$$ARE(T^*,T)=\lim_{n\to\infty}\frac{n}{n^*},$$

под условом да гранична вредност постоји и да не зависи од низова $\{\theta_n\},\{n\},\{n^*\}$.

Другим речима, ARE је реципрочна вредност количника обима узорака потребних за добијање моћи γ тестова T^* и T,када обими узорака теже бесконачности и низ алтернатива тежи θ_0, при чему оба теста имају исти ниво значајности. Према томе, ARE је мера асимптотске и локалне ефикасности моћи теста.

Претпоставимо да су T_n и T_n^* постојани тестови величине α за тестирање хипотезе

$$H_0(\theta=\theta_0)\quad \text{против}\quad H_1(\theta>\theta_0),$$

и имамо, редом, критичне области облика

$$T_n \in W \quad \text{за} \quad T_n \geq t_{n,\alpha} \quad \text{и} \quad T_n^* \in W \quad \text{за} \quad T_n^* \geq t_{n,\alpha}^*,$$

где се $t_{n,\alpha}$ и $t_{n,\alpha}^*$ налазе из

$$P\{T_n \geq t_{n,\alpha} | \theta = \theta_0\} = \alpha \quad \text{и} \quad P\{T_n^* \geq t_{n,\alpha}^* | \theta = \theta_0\} = \alpha$$

Следећи услови регуларности за тест статистику T_n (и аналогно за T_n^*) морају да буду испуњени [2], [6]:

(У1) Први извод $\frac{dE(T_n)}{d\theta}$ постоји и непрекидан је у θ_0. Сви изводи вишег реда $\frac{d^r E(T_n)}{d\theta^r}$, $r = 2, 3, \ldots$ једнаки су 0 у θ_0.

(У2) Постоји позитивна константа c тако да важи:

$$\lim_{n \to \infty} \frac{dE(T_n)/d\theta|_{\theta=\theta_0}}{\sqrt{n}\sigma(T_n)|_{\theta=\theta_0}} = c$$

(У3) Постоји низ алтернатива $\{\theta_n\}$ тако да за неку коннстанту $d > 0$ важи

$$\theta_n = \theta_0 + \frac{d}{\sqrt{n}}$$

$$\lim_{n \to \infty} \frac{dE(T_n)/d\theta|_{\theta=\theta_n}}{dE(T_n)/d\theta|_{\theta=\theta_0}} = 1$$

$$\lim_{n \to \infty} \frac{\sigma(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_0}} = 1$$

(У4) $\lim\limits_{n \to \infty} P\{\frac{T_n - E(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_n}} \leq z | \theta = \theta_n\} = F(z)$

(У5) $\lim\limits_{n \to \infty} P\{T_n \geq t_{n,\alpha} | \theta = \theta_0\} = \alpha, \ \ 0 < \alpha < 1$

Теорема 4 *Ако важе услови регуларности (У1)- (У5), гранична вредност моћи теста T_n је једнака*

$$\lim_{n \to \infty} P_W\{T_n | \theta = \theta_n\} = 1 - F(z_\alpha - dc),$$

где је z_α број за који важи $1 - F(z_\alpha) = \alpha$.

Доказ: Гранична вредност моћи теста T_n је

$$\begin{aligned}\lim_{n\to\infty} P\{T_n \geq t_{n,\alpha}|\theta=\theta_n\} &= \lim_{n\to\infty} P\{\frac{T_n - E(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_n}} \geq \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_n}}\} = \\ &= 1 - F(z).\end{aligned}$$

на основу (У4), где је

$$\begin{aligned} z &= \lim_{n\to\infty} \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_n}} = \lim_{n\to\infty} \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_0}} \cdot \frac{\sigma(T_n)|_{\theta=\theta_0}}{\sigma(T_n)|_{\theta=\theta_n}} = \\ &= \lim_{n\to\infty} \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_n}}{\sigma(T_n)|_{\theta=\theta_0}}.\end{aligned}$$

на основу (У3).

Користећи Тејлоров развој $E(T_n)|_{\theta=\theta_n}$ у θ_0 и на основу (У1) добијамо

$$E(T_n)|_{\theta=\theta_n} = E(T_n)|_{\theta=\theta_0} + (\theta_n - \theta_0)\frac{dE(T_n)}{d\theta}|_{\theta=\theta_0^*},\ \ \theta_0 < \theta_0^* < \theta_n.$$

Када заменимо ово у горњи израз за z добијамо, на основу (У1) - (У3)

$$\begin{aligned} z &= \lim_{n\to\infty} \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_0}}{\sigma(T_n)|_{\theta=\theta_0}} - \frac{(\theta_n - \theta_0)dE(T_n)/d\theta|_{\theta=\theta_0^*}}{\sigma(T_n)|_{\theta=\theta_0}} = \\ &= \lim_{n\to\infty} \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_0}}{\sigma(T_n)|_{\theta=\theta_0}} - dc = z_\alpha - dc.\end{aligned}$$

На основу (У4) и (У5) имамо да је

$$\begin{aligned} \alpha &= \lim_{n\to\infty} P\{T_n \geq t_{n,\alpha}|\theta=\theta_0\} = \\ &= \lim_{n\to\infty} P\{\frac{T_n - E(T_n)|_{\theta=\theta_0}}{\sigma(T_n)|_{\theta=\theta_0}} \geq \frac{t_{n,\alpha} - E(T_n)|_{\theta=\theta_0}}{\sigma(T_n)|_{\theta=\theta_0}}\} = \\ &= 1 - F(z_\alpha).\end{aligned}$$

Теорема 5 . *Ако тестови T и T^* задовољавају услове регуларности (У1) - (У5), ARE теста T^* релативно у односу на T је*

$$ARE(T^*, T) = \lim_{n\to\infty} (\frac{dE(T_n^*)/d\theta|_{\theta=\theta_0}}{dE(T_n)/d\theta|_{\theta=\theta_0}})^2 \cdot \frac{\sigma^2(T_n)|_{\theta=\theta_0}}{\sigma^2(T_n^*)|_{\theta=\theta_0}}.$$

Доказ: Из теореме 4, граничне вредности моћи тестова T и T^* су, редом,

$$1 - F(z_\alpha - dc) \quad \text{и} \quad 1 - F(z_\alpha - d^*c^*)$$

Ове две вредности једнаке су ако је

$$z_\alpha - dc = z_\alpha - d^*c^* \quad \Leftrightarrow \quad \frac{d}{d^*} = \frac{c^*}{c}.$$

Из (У3), низови алтернатива су исти ако је

$$\theta_n = \theta_0 + \frac{d}{\sqrt{n}} = \theta_n^* = \theta_0 + \frac{d^*}{\sqrt{n^*}} \quad \Leftrightarrow \quad \frac{d}{\sqrt{n}} = \frac{d^*}{\sqrt{n^*}} \text{ или } \frac{d}{d^*} = (\frac{n}{n^*})^{\frac{1}{2}},$$

односно

$$\frac{n}{n^*} = (\frac{c^*}{c})^2$$

Пошто је ARE гранична вредност количника обима узорака када су граничне вредности моћи и низови алтерантива за оба теста једнаки, када искористимо једнакост из (У2) за константе c и c^* добијамо

$$\begin{aligned} ARE(T^*, T) &= \lim_{n\to\infty} (\frac{dE(T_n^*/d\theta|_{\theta=\theta_0}}{\sqrt{n}\sigma(T_n^*)|_{\theta=\theta_0}} \cdot \frac{\sqrt{n}\sigma(T_n)|_{\theta=\theta_0}}{dE(T_n)/d\theta|_{\theta=\theta_0}})^2 = \\ &= \lim_{n\to\infty} \frac{\frac{(dE(T_n^*)/d\theta)^2}{\sigma^2(T_n^*)}|_{\theta=\theta_0}}{\frac{(dE(T_n)/d\theta)^2}{\sigma^2(T_n)}|_{\theta=\theta_0}} \end{aligned} \tag{4.1}$$

Из израза 4.1 видимо да, када су услови регуларности испуњени, ARE се може интерпретирати као гранична вредност количника две величине

$$ARE(T^*, T) = \lim_{n\to\infty} \frac{e(T_n^*)}{e(T_n)} \tag{4.2}$$

где се $e(T_n)$ назива ефикасност статистике T_n када се тестира хипотеза ($\theta = \theta_0$) која је једнака

$$e(T_n) = \frac{(dE(T_n)/d\theta)^2|_{\theta=\theta_0}}{\sigma^2(T_n)|_{\theta=\theta_0}} \tag{4.3}$$

Приметимо да је резултат за ARE у Теореми 5 независан и од прага значајности α и од моћи теста γ. Међутим, ARE је само апроксимација релативне ефикасности за било који коначан узорак и/или алтернативну хипотезу која није у околини нулте хипотезе.

У случају тестирања нулте хипотезе о једнакости средњих вредности код два независна узорка обима m и n ,исте теореме се могу користити, при чему се претпоставља да када m и n теже бесконачности, однос m/n је константан.

4.3 Проблем једног узорка

Тестираћемо нулту хипотезу $H_0(\theta = 0)$, где је θ средња вредност расподеле обележја X, против алтернативне хипотезе $H_1(\theta > 0)$ (тј. $\theta = c > 0$). Можемо говорити о проблему параметра локације код једног узорка са моделом

$$F_X(x) = F(x - \theta),$$

за неку непрекидну симетричну расподелу F са средњом вредношћу једнаком нули, а F_X је расподела обележја X симетрична око θ.

Дефиниција 3 *Коефицијент спљоштености $\gamma_2(X)$ је дефинисан са*

$$\gamma_2(X) = \frac{E(X - E(X))^4}{D(X)^2} - 3,$$

За $\gamma_2(X) > 0$ кажемо да је расподела са дебелим реповима, за $\gamma_2(X) < 0$ кажемо да је расподела са лаким реповима, док за $\gamma_2(X) = 0$ кажемо да расподела има репове исте дебљине као нормална расподела.

Изабраћемо четири непрекидне симетричне расподеле различите дебљине репа.

1. Нормална расподела $\mathcal{N}(\theta, \sigma^2)$ са густином расподеле

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{(x-\theta)^2}{2}},\ x \in \mathbb{R}.$$

   Студентов t - тест сматра се најмоћнијим за узорке из нормалне расподеле, па ће бити интересантно упоредити његов учинак са учинком Вилкоксоновог теста означених рангова и теста знакова.

2. Униформна расподела $\mathcal{U}(-\frac{1}{2} + \theta, \frac{1}{2} + \theta)$ са густином расподеле

$$f_X(x) = 1,\ -\frac{1}{2} + \theta < x < \frac{1}{2} + \theta$$

   Ова расподела одабрана је као пример симетричне расподеле лаког репа.

3. Логистичка расподела $\mathcal{LGS}(a, b)$ са густином расподеле

$$f_X(x) = \frac{e^{-\frac{x-a}{b}}}{b \cdot (1 + e^{-\frac{x-a}{b}})^2},\ x \in \mathbb{R},$$

   где је $a \in \mathbb{R}$ параметар локације а $b > 0$ параметар скалирања. Логистичка расподела узета је као пример симетричне расподеле средње тешког репа.

4. Лапласова расподела $\mathcal{L}_2(\lambda, a)$, са густином расподеле

$$f_X(x) = \frac{\lambda}{2} e^{-\lambda|x-a|},\ x \in \mathbb{R},$$

где је параметар локације, а $\lambda > 0$ параметар скалирања. Ова расподела представља добар пример симетричне расподеле тешког репа.

4.3.1 Функција моћи t - теста, Вилкоксоновог теста означених рангова и теста знакова

Разматрани су узорци малог, средњег и великог обима. Узети су узорци обима $n = 9, 18, 36$ и 108. За сваки обим узорка рачуната је функција моћи разматраних тестова. За различите вредности константе $c = \theta$ генерисано је по $N = 1000$ узорака обима n из сваке од поменутих расподела и израчуната је вредност функције моћи теста.

Ради поређења моћи t - теста са непараметарским тестовима, узето је да вероватноће које одговарају критичним вредностима (односно вероватноће добијања вредности веће или једнаке критичној вредности) ових тестова буду једнаке, тј. да су тестови исте величине. Тако, ако означимо са W_n, K_n, T_n тест статистике, редом, Вилкоксоновог теста означених рангова, теста знакова и t - теста из услова

$$P\{W_n \geq w_c\} \leq \alpha \ \text{ и } \ P\{K_n \geq k_c\} \leq \alpha$$

рачунамо за изабрани праг значајности α критичне вредности w_c и k_c и затим величине тестова

$$p_1 = P\{W_n \geq w_c\} \ \text{ и } \ p_2 = P\{K_n \geq k_c\}.$$

Приликом рачунања критичне вредности t - теста узимамо

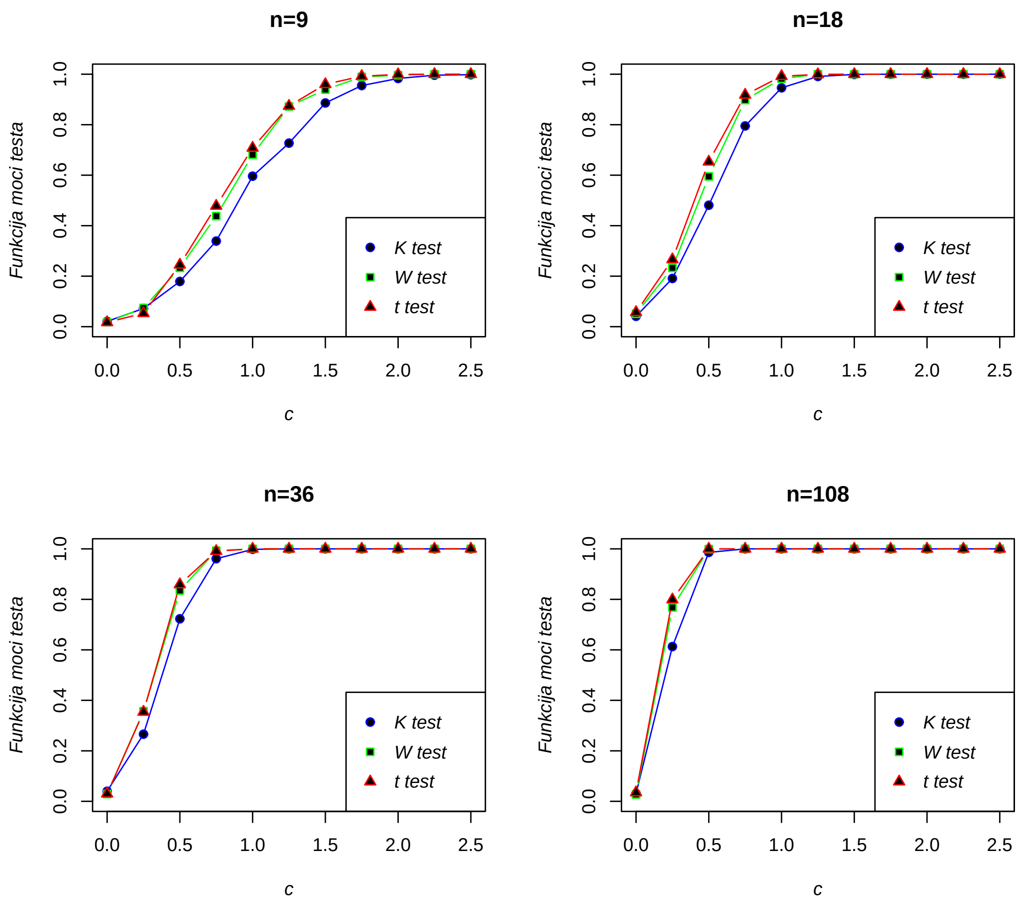
$$\alpha^{'} = \min(p_1, p_2) \ \text{ и } \ P\{T_n \geq t_c\} = \alpha^{'},$$

па рачунамо нове критичне вредности $w_c^{'}$ и $k_c^{'}$ тако да је

$$P\{W_n \geq w_c^{'}\} \leq \alpha^{'} \ \text{ и } \ P\{K_n \geq k_c^{'}\} \leq \alpha^{'}.$$

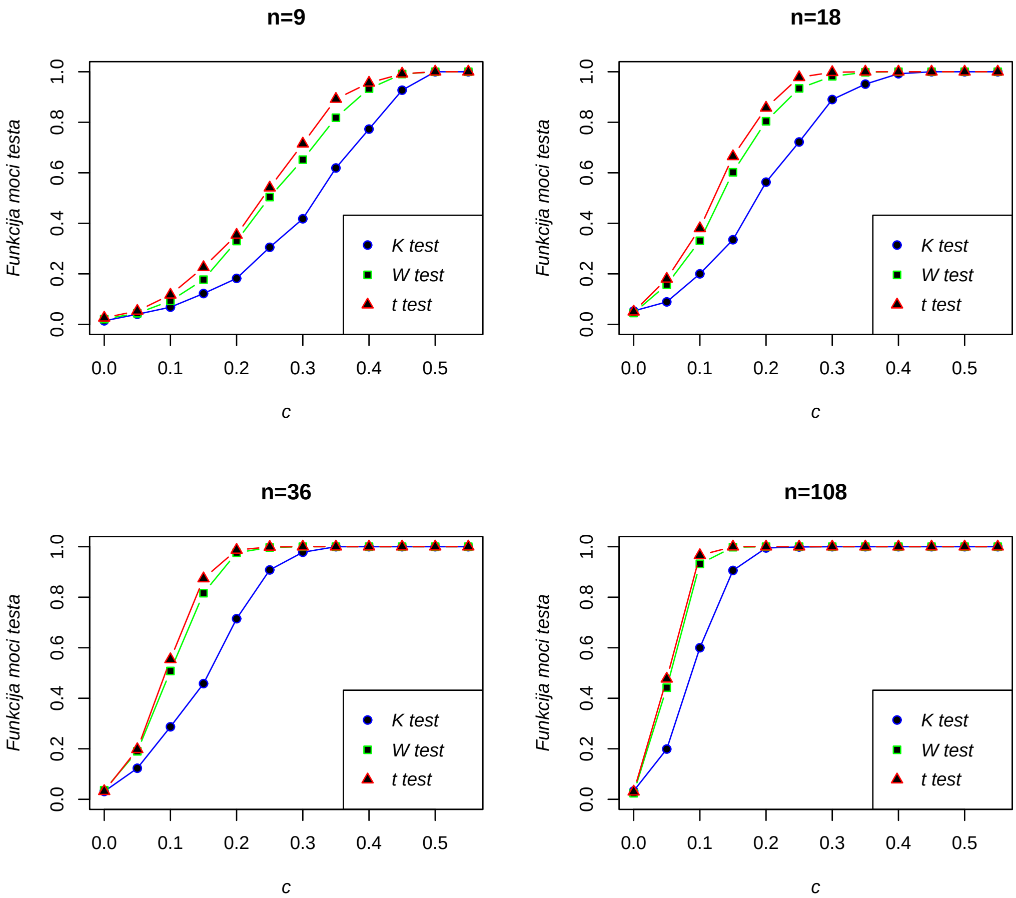
Ова два теста ће бити величине $\alpha^{'}$ или приближно $\alpha^{'}$.

На следеће четири слике представљене су функције моћи теста знакова (у легенди графика означен као K тест), Вилкоксоновог теста означених рангова (у легенди означен као W тест) и t - теста за узорке различитих обима из нормалне, униформне, логистичке и Лапласове расподеле.



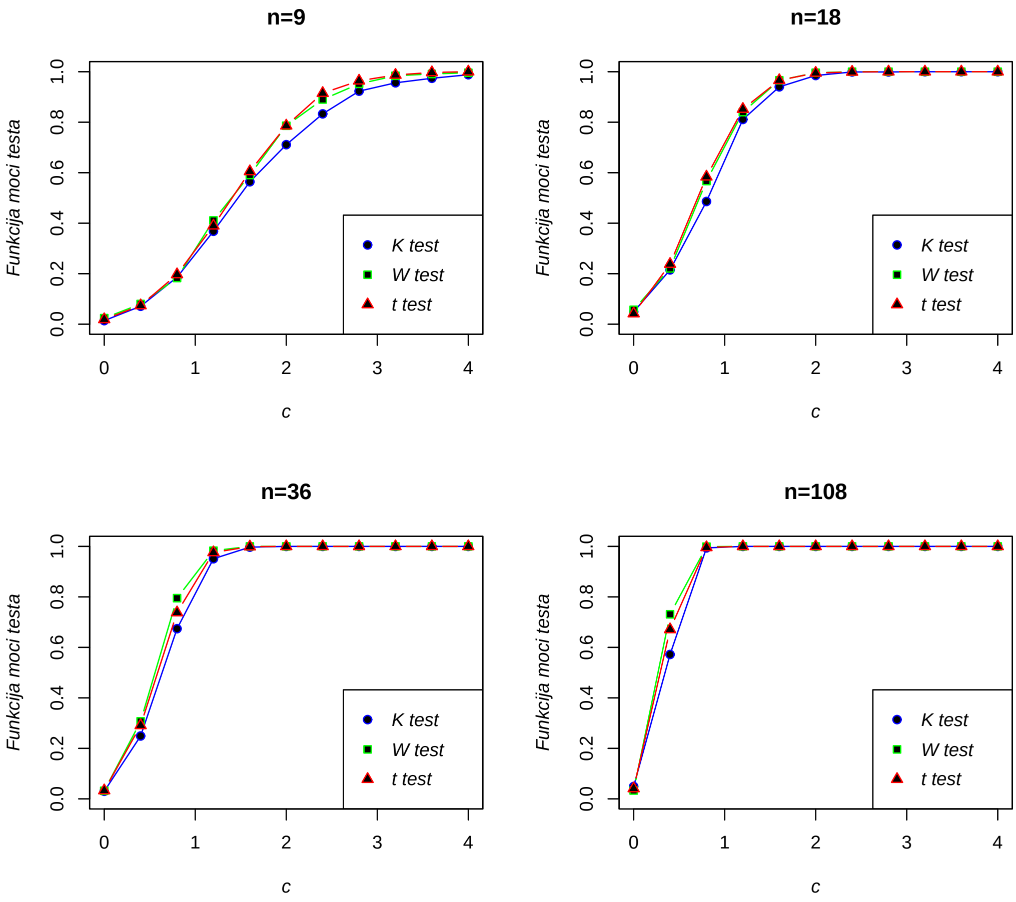
Слика 4.1: Функције моћи Вилкоксоновог теста означених рангова, теста знакова и t - теста за узорке из нормалне расподеле

Посматрајући графике на слици 4.1 , примећујемо да, за узорке из нормалне расподеле, Вилкоксонов тест означених рангова нешто слабије пролази од t - теста, разлика између њихових функција моћи је врло мала, за веће обиме узорака ($n = 36$ и $n = 108$) се практично поклапају. Тест знакова слабије пролази, поклапа се са ова два теста тек при већим одступањима од нулте хипотезе.



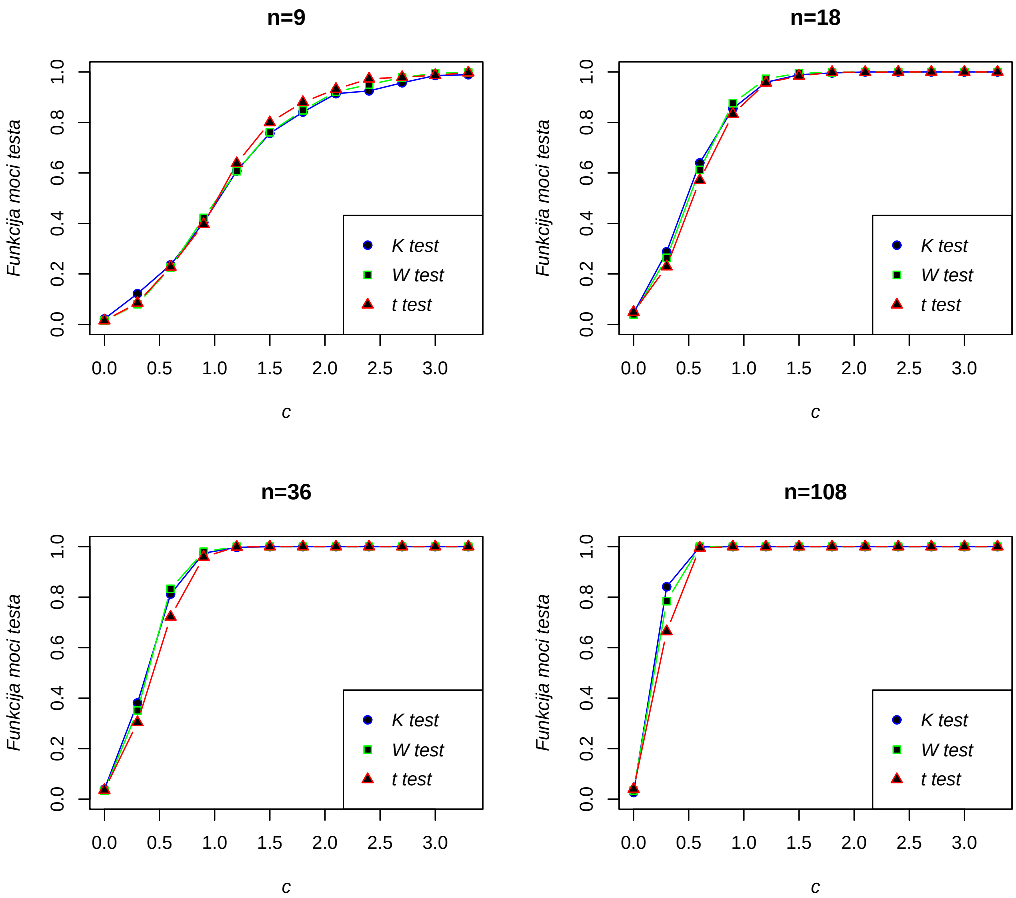
Слика 4.2: Функције моћи Вилкоксоновог теста означених рангова, теста знакова и t - теста за узорке из униформне расподеле

На слици 4.2, за узорке из униформне расподеле, уочавамо да је тест знакова знатно слабији од t - теста и Вилкоксоновог теста означених рангова. Студентов t - тест је најмоћнији, са порастом обима узорка Вилкоксонов тест означених рангова му се све више приближава и достиже га за $n = 108$.



Слика 4.3: Функције моћи Вилкоксоновог теста означених рангова, теста знакова и t - теста за узорке из логистичке расподеле

На основу графика на слици 4.3, за узорке из логистичке расподеле, примећујемо да је тест знакова најслабији, док за мањи обим узорка ($n = 9, 18$) Вилкоксонов тест означених рангова је нешто слабији од t - теста. Како се обим узорка повећава (од $n = 36$), Вилкоксонов тест означених рангова благо престиже t - тест.



Слика 4.4: Функције моћи Вилкоксоновог теста означених рангова, теста знакова и t - теста за узорке из Лапласове расподеле

За функције моћи теста на слици 4.4 , за узорке из Лапласове расподеле, можемо рећи да се преплићу за мали обим узорка ($n = 9$),тј. за неке вредности константе c сваки од тестова нешто бољи од осталих. Са повећањем обима узорка, за $n = 18$ и $n = 36$, уочавамо да се функције моћи теста знакова и Вилкоксоновог теста поклапају (тест знакова је за нијансу бољи), а да је t - тест слабији у односу на њих. За $n = 108$ најбоље пролази тест знакова, затим Вилкоксонов тест означених рангова, а најслабији је t - тест.

4.3.2 Асимптотска релативна ефикасност

За прост случајан узорак $(X_1, X_2, \ldots, X_n)$ обима n из било које непрекидне симетричне расподеле F_X са средњом вредношћу θ и дисперзијом σ^2, израчунаћемо прво ефикасност статистике t - теста, па затим ефикасност статистика теста знакова и Вилкоксоновог теста означених рангова. Као што већ раније речено, тестирамо нулту хипотезу $H_0(\theta = 0)$ против алтернативе $H_1(\theta > 0)$. Услови регуларности су испуњени за сваку разматрану тест статистику, што се једноставно проверава.

Означимо t - тест са T^*. Његова тест статистика је једнака

$$T_n^* = \frac{(\overline{X}_n - \theta) \cdot \sqrt{n}}{\widetilde{S_n}} = \frac{(\overline{X}_n - \theta) \cdot \sqrt{N}}{\sigma} \frac{\sigma}{\widetilde{S_N}},$$

где је $\overline{X}_n = \frac{1}{n} \sum\limits_{i=1}^{n} X_i$ узорачка средина, а $\widetilde{S_n^2} = \frac{1}{n-1} \sum\limits_{i=1}^{n} (X_i - \overline{X_n})^2$ поправљена узорачка дисперзија.

Пошто је $\lim\limits_{N \to \infty} \frac{\widetilde{S_N}}{\sigma} = 1$, тест T^* је еквивалентан Z тесту за познато σ.

Математичко очекивање и дисперзија T_n^*, за велико n тада су једнаки

$$E(T_n^*) = \frac{\sqrt{n}\theta}{\sigma} \quad \text{и} \quad D(T_n^*) = \frac{n \cdot D(\overline{X}_n)}{\sigma^2} = 1$$

па се добија

$$\frac{dE(T_n^*)}{d\theta}|_{\theta=0} = \frac{\sqrt{n}}{\sigma}$$

Користећи једнакост 4.3 добијамо да је ефикасност статистике Студентовог t - теста једнака

$$e(T_n^*) = \frac{n}{\sigma^2}.$$

Означимо са K_n статистику теста знакова. Пошто K_n има биномну расподелу, $K_n : \mathcal{B}(n, p)$, њено математичко очекивање и дисперзија једнаки су

$$E(K_n) = np, \quad D(K_n) = np(1-p),$$

где је $p = P\{X > 0\} = 1 - F_X(0)$.

Даље, имамо да је

$$\begin{aligned}
\frac{E(K_n)}{d\theta}|_{\theta=0} &= n \cdot \frac{dp}{d\theta}|_{\theta=0} = n \cdot \frac{d}{d\theta}(1 - F_X(0))|_{\theta=0} = n \cdot \frac{d}{d\theta}(1 - F(-\theta))|_{\theta=0} = \\
&= n \cdot f(\theta)|_{\theta=0} = n \cdot f(0) = n \cdot f_X(\theta).
\end{aligned}$$

Према томе, добијамо да је ефикасност статистике теста знакова једнака:

$$e(K_n) = \frac{n^2 \cdot f_X(\theta)}{n/4} = 4n \cdot f_X^2(\theta) = 4n \cdot f^2(0).$$

Израчунаћемо сада ефикасност Вилкоксоновог теста означених рангова. Означимо са T_n^+ тест статистику. Ради лакшег рачунања, погодније је радити са статистиком $V_n^+ = \frac{T_n^+}{\binom{n}{2}}$. Пошто је тест заснован на статистици T_n^+ еквивалентан тесту заснованом на статистици V_n^+, ове две статистике ће имати исту ефикасност. Имамо да је математичко очекивање тест статистике T_n^+ једнако

$$E(T_n^+) = n \cdot p_1 + \frac{n \cdot (n-1)}{2} \cdot p_2,$$

где су вероватноће p_1 и p_2 једнаке

$$p_1 = P\{D_i > 0\}, \quad p_2 = P\{D_i + D_j > 0\},$$

а $D_i = X_i - \theta \;\; i = 1, 2, \dots, n$. Дакле, математичко очекивање статистике V_n^+ је

једнако

$$\begin{aligned}
E(V_n^+) &= \frac{2}{n-1} P\{D_i > 0\} + P\{D_i + D_j > 0\} = \\
&= \frac{2}{n-1}(1 - F(-\theta)) + \int_{-\infty}^{+\infty} \int_{-x}^{+\infty} f_D(x) f_D(t) dx dt = \\
&= \frac{2}{n-1}(1 - F(-\theta)) + \int_{-\infty}^{+\infty} (1 - F(-x - \theta)) dF(x - \theta) = \\
&= \frac{2}{n-1}(1 - F(-\theta)) + \int_{-\infty}^{+\infty} (1 - F(-y - 2\theta)) dF(y).
\end{aligned}$$

Према томе, имамо да је

$$\frac{dE(V_n^+)}{d\theta} = \frac{2}{n-1} f(-\theta) + 2 \int_{-\infty}^{+\infty} f(-y - 2\theta) dF(y).$$

Пошто је густина расподеле f симетрична око нуле, тј. $f(y) = f(-y)$, добијамо

$$\frac{dE(V_n^+)}{d\theta}\Big|_{\theta=0} = \frac{2}{n-1} f(0) + 2 \int_{-\infty}^{+\infty} f^2(y) dy = 2\left(\frac{f(0)}{n-1} + I\right),$$

где је $I = \int\limits_{-\infty}^{+\infty} f^2(y)dy$.

Дисперзија тест статистике T_n^+ ако је тачна нулта хипотеза је једнака

$$D(T_n^+) = \frac{n(n+1)(2n+1)}{24},$$

па је дисперзија статистике V_n+

$$D(V_n^+) = \frac{(n+1)(2n+1)}{6n(n-1)^2}.$$

Према томе, ефикасност тест статистике Вилкоксоновог теста означених рангова је на основу једнакости 4.3 једнака

$$e(T_n^+) = \frac{24(f(0)/(n-1)+I)^2 n(n-1)^2}{(n+1)(2n+1)}.$$

На основу израчунатих ефикасности тест статистика теста знакова, t-теста и Вилкоксоновог теста означених рангова можемо на основу 4.2 израчунати асимптотске релативне ефикасности за било која два од ових тестова.

ARE теста знакова у односу на t тест је

$$ARE(K, T^*) = 4f^2(0)\sigma^2.$$

ARE теста знакова у односу на Вилкоксонов тест означених рангова је

$$ARE(K, T^+) = \frac{f^2(0)}{3I^2}.$$

ARE Вилкоксоновог теста означених рангова у односу на t-тест је

$$ARE(T^+, T^*) = 12\sigma^2 I^2.$$

Израчунаћемо сада асимптотске релативне ефикасности ових тестова за нормалну, униформну, логистичку и Лапласову расподелу.

1. Нормална расподела $\mathcal{N}(0, \sigma^2)$
   $f(0) = \frac{1}{\sqrt{2\pi}\sigma}$, $I = \frac{1}{2\sigma\sqrt{\pi}}$
   $ARE(K, T^*) = \frac{2}{\pi}$,
   $ARE(K, T^+) = \frac{2}{3}$
   $ARE(T^+, T^*) = \frac{3}{\pi}$

2. Униформна расподела $\mathcal{U}(-1/2, 1/2)$
   $f(0) = 1,\ \ I = 1$
   $ARE(K, T^*) = \frac{1}{3}$,
   $ARE(K, T^+) = \frac{1}{3}$
   $ARE(T^+, T^*) = 1$

3. Логистичка расподела $\mathcal{LGS}(0, b)$
   $f(0) = \frac{1}{4b},\ \ I = \frac{1}{6b}$
   $ARE(K, T^*) = \frac{\pi^2}{12}$,
   $ARE(K, T^+) = \frac{3}{4}$
   $ARE(T^+, T^*) = \frac{\pi^2}{9}$

4. Лапласова расподела $\mathcal{L}_2(0, \lambda)$
   $f(0) = \frac{\lambda}{2},\ \ I = \frac{\lambda}{4}$
   $ARE(K, T^*) = 2$,
   $ARE(K, T^+) = \frac{4}{3}$
   $ARE(T^+, T^*) = \frac{3}{2}$

Представимо добијене резултате за све четири расподеле у следећој табели.

Расподела	$ARE(T^+, T^*)$	$ARE(K, T^*)$	$ARE(K, T^+)$
Нормална	$3/\pi = 0.955$	$2/\pi = 0.637$	$2/3$
Униформна	1	$1/3$	$1/3$
Логистичка	$\pi^2/9 = 1.097$	$\pi^2/12 = 0.822$	$3/4$
Лапласова	$3/2$	2	$4/3$

Табела 1. Вредности $ARE(T^+, T^*), ARE(K, T^*)$ и $ARE(K, T^+)$ за неке симетричне расподеле

На основу израчунатих асимптотских релативних ефикасности можемо извести следеће закључке. За нормалну расподелу, за коју је t-тест оптималан тест, врло мало се губи, у смислу ефикасности теста, ако се уместо њега користи Вилкоксонов тест означених рангова. За симетричне расподеле лаког или тешког репа, Вилкоксонов тест означених рангова супериорнији је у смислу да је ARE барем 1. Хоџиз и Леман [9] су доказали да је асимптотска релативна ефикасност Вилкоксоновог теста означених рангова у односу на t тест најмање 0.864 за било коју непрекидну симетричну расподелу.

Видимо да је тест знакова мање ефикасан од t-теста за расподеле лаког до средње тешког репа. Код нормалне расподеле, тест знакова је 64% ефикасан у односу на t-тест, што не изненађује, с обзиром да тест знакова користи врло мало информација из узорка. Интересантно је да је тест знакова супериорнији (и у односу на t - тест и Вилкоксонов тест означених рангова) у случају расподела са тешким репом (Лапласова расподела). Хоџиз и Леман [9] су доказали да је асимптотска релативна ефикасност теста знакова у односу на Студентов t - тест најмање $1/3$ за било коју непрекидну симетричну расподелу.

Када упоредимо добијене резултате са функцијама моћи ових тестова за узорке различитог обима, уочавамо да се вредности асимптотске релативне ефикасности не могу применити на узорке малог обима. Тако, код униформне расподеле, иако је асимптотска релативна ефикасност Вилкоксоновог теста означених рангова у односу на t - тест једнака 1, за узорке обима $n = 9$ и $n = 18$, t - тест је моћнији. Код Лапласове расподеле, за узорке обима $n = 18$ и $n = 36$, функције моћи Вилкоксоновог теста означених рангова и тест знакова се поклапају, док је асимптотска релативна ефикасност теста знакова у односу на Вилкоксонов тест означених рангова у овом случају једнака $4/3$.

Дакле, на основу добијених резултата можемо закључити да је Вилкоксонов тест означених рангова врло добра алтернатива Студентовом t - тесту. Тест знакова је обично мање ефикасан, можемо рећи да је једноставност његове употребе превагнула над његовим учинком.

4.4 Проблем два независна узорка

Тестираћемо нулту хипотезу $H_0(\theta = 0)$, где је $\theta = m_x - m_Y$ разлика средњих вредности обележја X и Y, против алтернативне хипотезе $H_1(\theta > 0)$. Можемо разматрати модел локације који претпоставља да се расподеле обележја X и Y разликују само за померај θ, односно

$$F_X(x) = F_Y(x - \theta), \quad \text{за свако } x.$$

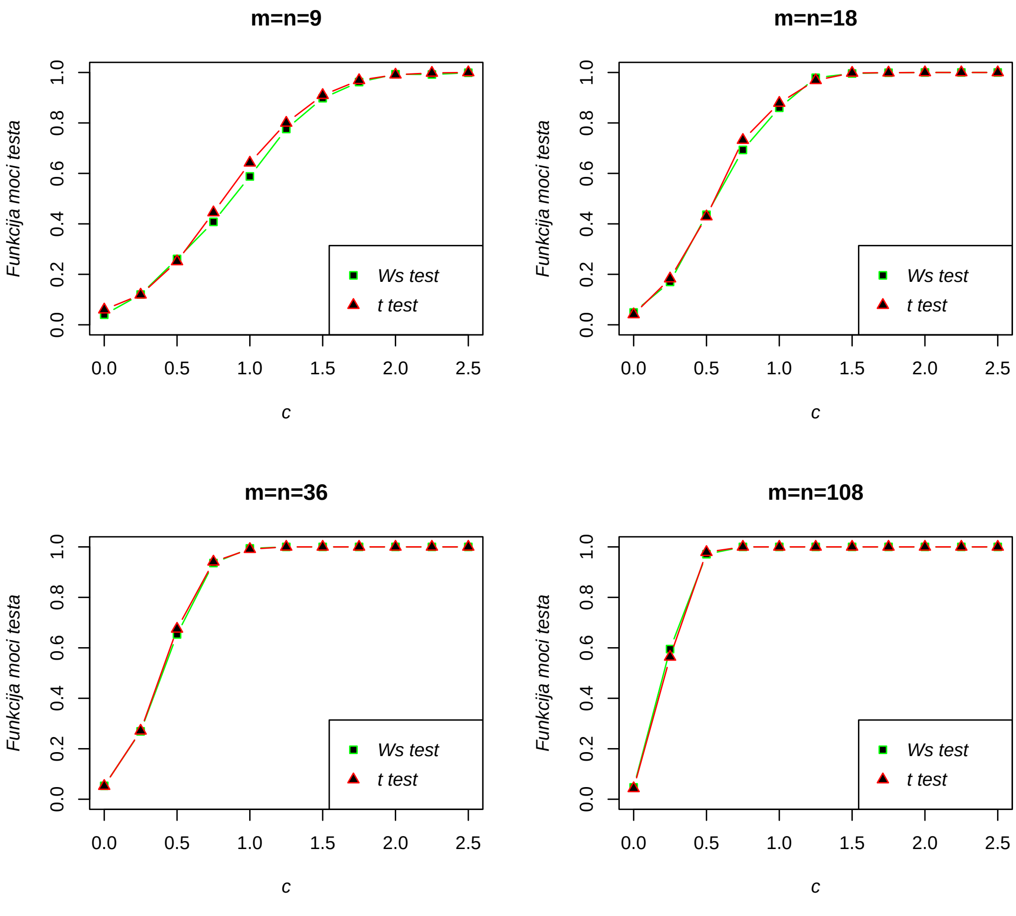
У том случају X и $Y + \theta$ имају исту расподелу. Расподела обележја X иста је као расподела обележја Y ако је $\theta = 0$, померена је удесно ако је $\theta > 0$ и померена је улево ако је $\theta < 0$. Под овим претпоставкама расподеле обележја X и Y имају исти облик и исту дисперзију. Као у случају једног узорка разматраћемо независне узорке из нормалне, униформне, логистичке и Лапласове расподеле.

4.4.1 Функција моћи t - теста и Вилкоксоновог теста суме рангова

Разматрани су независни узорци обима 9, 18, 36 и 108. У програмском језику R генерисано је 1000 независних узорака обима m и n из поменутих непрекидних симетричних расподела (са средњом вредношћу једнаком 0). Узорцима обима m додата је константа c у складу са алтернативном хипотезом. За изабране обиме узорака и вредност константе c рачуната је вредност функција моћи Студентовог t - теста и Вилкоксоновог теста суме рангова.

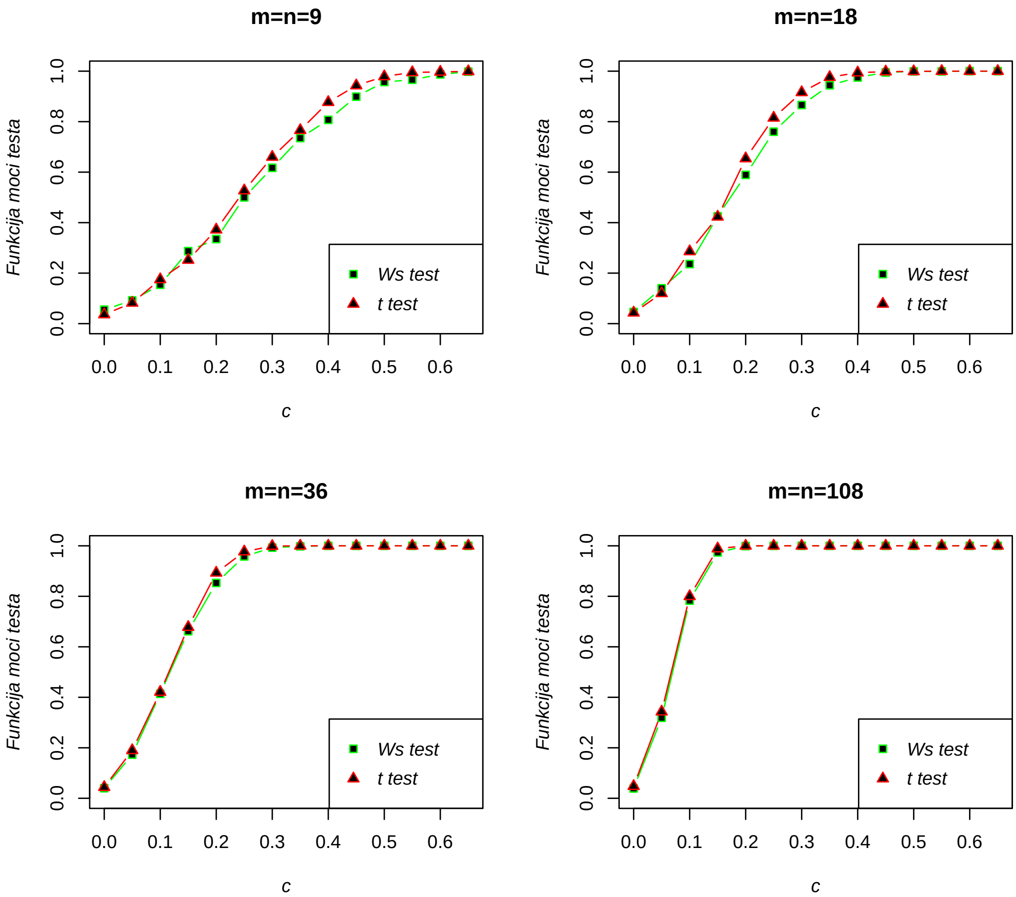
Ради поређења функција моћи ова два теста, узете су иста нулта и алтернативна хипотеза, обими узорака, праг значајности, као и иста величина тестова.

На следеће четири слике графички су представљене функције моћи Вилкоксоновог теста суме рангова (у легенди графика означен као Ws тест) и t - теста за узорке различитих обима из разматраних расподела.



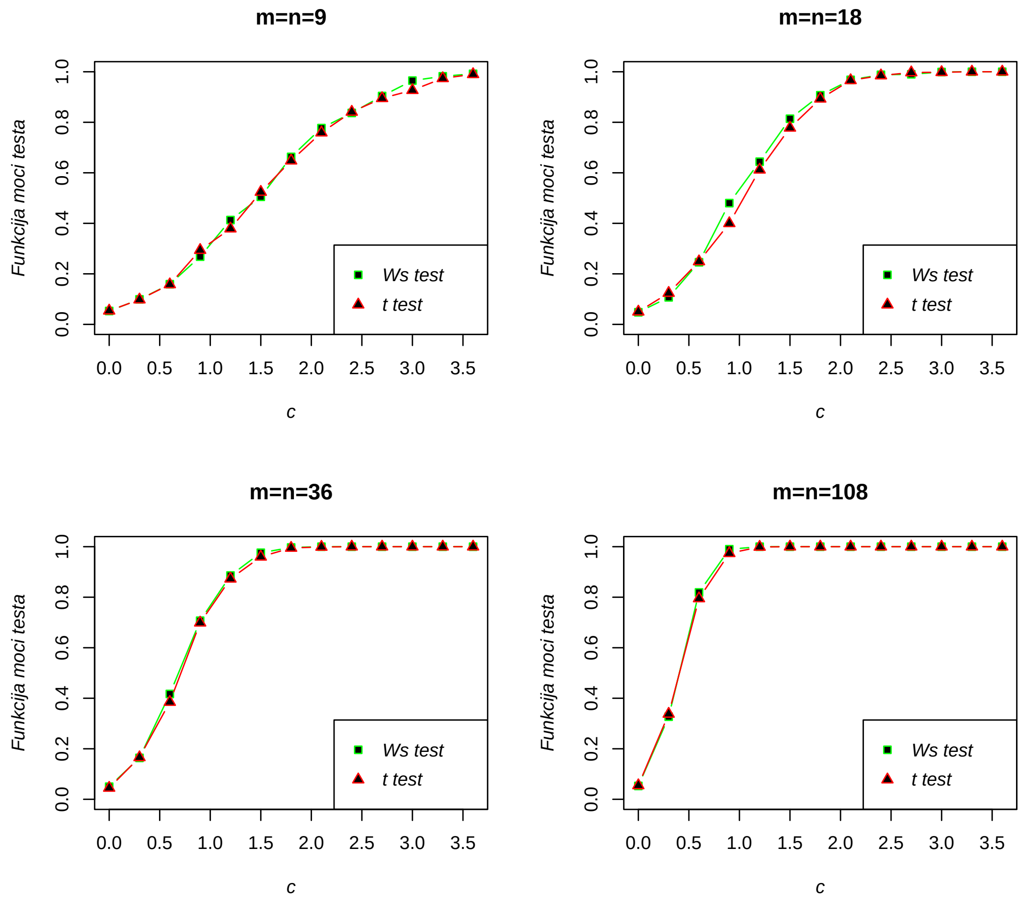
Слика 4.5: Функције моћи Вилкоксоновог теста суме рангова, и t - теста за узорке из нормалне расподеле

На слици 4.5, за узорке из нормалне расподеле, уочавамо да је t - тест нешто бољи од Вилкоксоновог теста суме рангова за $m = n = 9$, док за веће обиме узорка се њихове функције моћи практично поклапају.



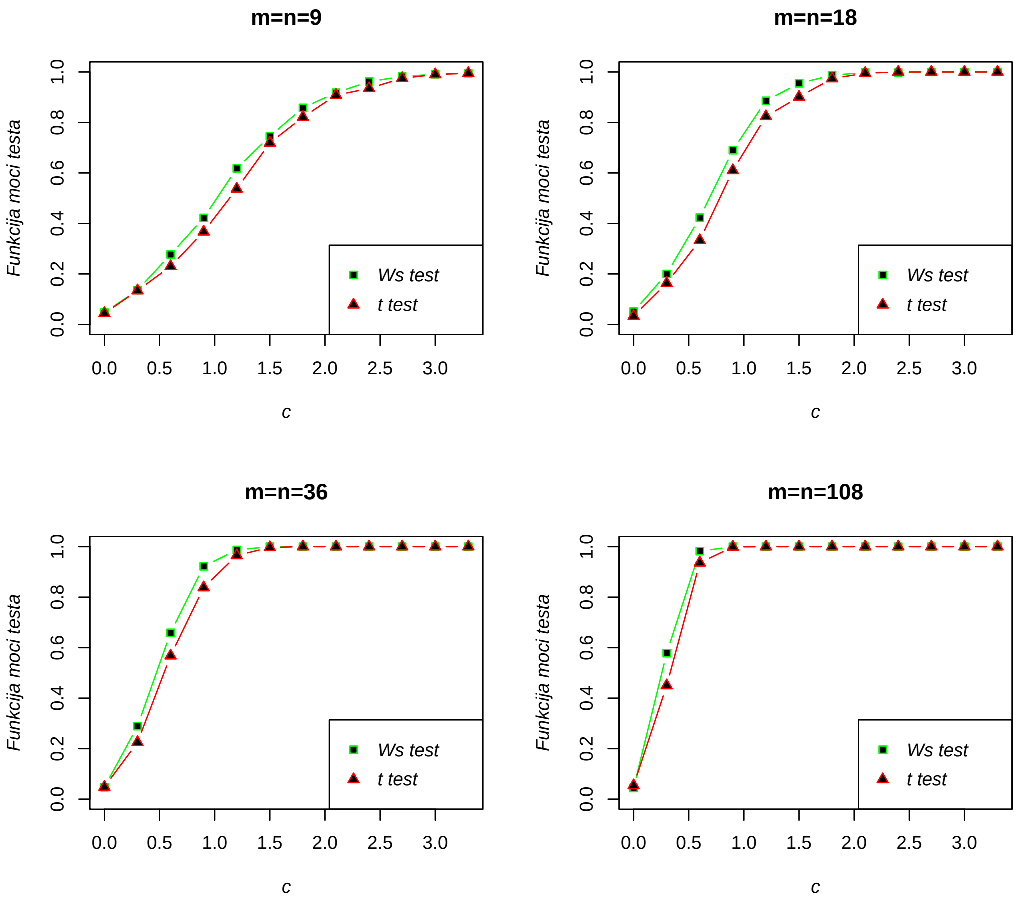
Слика 4.6: Функције моћи Вилкоксоновог теста суме рангова, и t - теста за узорке из униформне расподеле

На основу графика на слици 4.6, примећујемо да је t - тест моћнији за узорке обима $m = n = 9$, и $m = n = 18$. Са повећањем обима узорка смањује се разлика између функција моћи t - теста и Вилкоксоновог теста суме рангова и за $m = n = 108$ ови тестови се поклапају.



Слика 4.7: Функције моћи Вилкоксоновог теста суме рангова, и t - теста за узорке из логистичке расподеле

За функције моћи теста на слици 4.7, за узорке из логистичке расподеле, уочавамо да за $m = n = 9$ слично пролазе оба теста. Како се обими узорака повећавају, Вилкоксонов тест суме рангова боље пролази од t - теста, за $m = n = 36$ и $m = n = 108$ скоро се поклапају.



Слика 4.8: Функције моћи Вилкоксоновог теста суме рангова, и t - теста за узорке из Лапласове расподеле

На основу графика на слици 4.8, за узорке из Лапласове расподеле, примећујемо да је Вилкоксонов тест суме рангова моћнији од t - теста за све изабране обиме узорака.

4.4.2 Асимптотска релативна ефикасност Вилкоксоновог теста суме рангова у односу на t - тест

Нека су $(X_1, X_2, \ldots, X_m)$ и $(Y_1, Y_2, \ldots, Y_n)$ два независна узорка обима m и n из било које непрекидне симетричне расподеле са непознатим средњим вредностима, редом, m_X и m_Y и непознатом дисперзијом σ^2.

Тест статистика Студентовог t - теста за два независна узорка са непознатим али једнаким дисперзијама је

$$T^*_{m,n} = \sqrt{\frac{mn}{m+n}} \cdot \frac{\overline{X}_m - \overline{Y}_n - (m_X - m_Y)}{\widetilde{S}_{m,n}} = \sqrt{\frac{mn}{m+n}} \cdot \frac{\overline{X}_m - \overline{Y}_m - \theta}{\sigma} \frac{\sigma}{\widetilde{S}_{m,n}},$$

где је $\widetilde{S}^2_{m,n} = \frac{\sum_{i=1}^{m}(X_i - \overline{X}_m)^2 + \sum_{j=1}^{n}(Y_j - \overline{Y}_n)^2}{m+n-2}$ оцена дисперзије σ^2.

Како $\frac{\widetilde{S}_{m,n}}{\sigma} \to 1$, када $m, n \to \infty$, $\frac{m}{n} \to \lambda$, математичко очекивање и дисперзија тест статистике $T^*_{m,n}$ за велико n су једнаки

$$E(T^*_{m,n}) = \frac{\theta}{\sigma}\sqrt{\frac{mn}{m+n}}, \quad D(T^*_{m,n}) = \frac{mn}{m+n}\frac{\sigma^2/m + \sigma^2/n}{\sigma^2} = 1.$$

Према томе,

$$\frac{dE(T_{m,n})}{d\theta} = \sqrt{\frac{mn}{m+n}} \cdot \frac{1}{\sigma},$$

па је ефикасност статистике Студентовог t - теста једнака

$$e(T^*_{m,n}) = \frac{mn}{\sigma^2(m+n)}.$$

Сада ћемо израчунати ефикасност статистике $T_{m,n}$ Вилкоксоновог теста суме рангова. Ради лакшег извођења, радићемо са статистиком теста Ман - Витнија $U_{m,n} = T_{m,n} - \frac{m(m+1)}{2}$. Математичко очекивање тест статистике Ман - Витнија је

$$E(U_{m,n}) = mnp,$$

где је

$$p = P\{Y < X\} = \int_{-\infty}^{+\infty} F_Y(x) dF_X(x) = \int_{-\infty}^{+\infty} F_X(x-\theta) f_X(x) dx.$$

Дакле, имамо да је

$$\frac{dE(U_{m,n})}{d\theta}|_{\theta=0} = mn\frac{dp}{d\theta}|_{\theta=0} = -mn\int_{-\infty}^{+\infty} f^2_X(x)dx.$$

Када је тачна нулта хипотеза, $p = 0.5$ и дисперзија статистике $U_{m,n}$ је

$$D(U_{m,n}) = \frac{mn(m+n+1)}{12},$$

па је ефикасност статистике Ман-Витнијевог теста (самим тим и Вилкоксоновог теста суме рангова)

$$e(U_{m,n}) = e(T_{m,n}) = \frac{12mnI}{m+n+1},$$

где је $I = \int\limits_{-\infty}^{+\infty} f_X^2(x)dx$.

Према томе, асимптотска релативна ефикасност Вилкоксоновог теста суме рангова у односу на t - тест је

$$ARE(T, T^*) = \lim_{m,n\to\infty} \frac{e(T_{m,n})}{e(T^*_{m,n})} = 12I^2\sigma^2,$$

што се поклапа са асимптотском релативном ефикасносношћу Вилкоксоновог теста означених рангова у односу на t - тест за један узорак.

Дакле, асимптотска релативна ефикасност Вилкоксоновог теста суме рангова у односу на t - тест за сваку непрекидну симетричну расподелу једнака је асимптотској релативној ефикасности Вилкоксоновог теста означених рангова у односу на t - тест.

Представимо добијене резултате за све четири расподеле у следећој табели.

Расподела	$ARE(T, T^*)$
Нормална	$3/\pi = 0.955$
Униформна	1
Логистичка	$\pi^2/9 = 1.097$
Лапласова	$3/2$

Табела 2. Вредности $ARE(T, T^*)$ за неке симетричне расподеле

Дакле, на основу израчунатих асимптотских релативних ефикасности можемо извести исте закључке као у случају Вилкоксоновог теста означних рангова. Већ за нормалну расподелу Вилкоксонов тест суме рангова је врло добра замена за t - тест. Такође важи да, за симетричне расподеле лаког или тешког репа, Вилкоксонов тест суме рангова је боља алтернатива у смислу да је ARE барем 1. Асимптотска релативна ефикасност Вилкоксоновог теста суме рангова у односу на t тест најмање је 0.864 за било коју непрекидну симетричну расподелу [9].

Као у случају једног узорка, када упоредимо добијене резултате ARE са функцијама моћи ових тестова за узорке различитог обима, уочавамо да се вредности асимптотске релативне ефикасности не могу применити на узорке малог обима. Код нормалне, логистичке и Лапласове расподеле се резултати поклапају за различите обиме узорака, али код униформне расподеле, за узорке обима $n = 9$ и $n = 18$, t - тест је моћнији. Дакле, вредности асимптотске релативне ефикасности Вилкоксоновог теста суме рангова у односу на t - тест нам не могу априори пружити критеријум за избор једног од ова два теста за узорке малог (мањег) обима из било које непрекидне симетричне расподеле.

Закључак

Разматрани непараметарски тестови представљају најпопуларније тестове медијане у смислу да се најчешће срећу у литератури о непараметарској статистици, теоријској и примењеној. За изабране симетричне непрекидне расподеле Вилкоксонов тест означених рангова се показао као сјајна алтернатива Студентовом t - тесту за тестирање хипотезе о средњој вредности обележја. Када су обими узорака мали, асимптотске релативне ефикасности Вилкоксоновог теста означених рангова у односу на t - тест не поклапају се увек са закључцима добијеним на основу функција моћи ова два теста. Међутим, из наведених теоријских резултата следи да, ако има губитка, врло је мали ако се увек определимо за Вилкоксонов тест означених рангова. Као што смо видели, исти закључци се односе и на Вилкоксонов тест суме рангова за тестирање хипотезе о једнакости средњих вредности код два обележја за независне узорке из непрекидних симетричних расподела. У случају тестова за један узорак, тест знакова се показао као слабија алтернатива t - тесту, тако да се може препоручити његова употреба само за тестирање хипотезе о медијани када расподела обележја није симетрична.

Када је расподела обележја позната (или се може утврдити) непрекидна симетрична расподела а обим узорка је мали, препорука која се може начинити при избору између Студентовог t - теста и Вилкоксоновог теста означених рангова (за независне узорке: Вилкоксоновог теста суме рангова) је следећа: генерисати узорке тог обима из уочене расподеле и посматрати добијене функције моћи ова два теста, понављати поступак израчунавања функција моћи тестова и цртања графика велики број пута и на основу запажања о њиховом односу начинити избор између ова два теста за конкретне податке.

С друге стране, у литератури о непараметарској статистици се могу наћи и други непараметарски тестови медијане који нису постали толико познати. Питање које се јавља и на које ће се тражити одговор у даљем раду је који је разлог популарности овде разматраних тестова. Да ли је у питању једноставност њихове примене, њихова дугогодишња употреба или пак, заиста представљају најмоћније тестове у класи тестова о медијани обележја?

Прилог

Тест симетрије

```
sim.test<-function(niz,alfa=0.05) #niz predstavlja dobijeni uzorak
{n<-length(niz) # obim uzorka
xn<-mean(niz)  #uzoracka sredina
m3<-n*sum((niz-xn)^3)/((n-1)*(n-2))  #treci uzoracki moment
s3<-(sd(niz))^3
g1<-m3/s3   #uzoracki koeficijent asimetrije

sqb1<-(n-2)*g1/sqrt(n*(n-1))
A<-sqb1*sqrt((n+1)*(n+3)/(6*(n-2)))
B<-3*(n^2+27*n-70)*(n+1)*(n+3)/((n-2)*(n+5)*(n+7)*(n+9))
C<-sqrt(2*(B-1))-1
#if(C==1)
# C<-C+0.00001
D<-sqrt(C)
E<-1/sqrt(log(D))
F<-A/sqrt(2/(C-1))
Z<-E*log(F+sqrt(F^2+1))  #test statistika

c<-qnorm(1-alfa/2)  #kriticna vrednost

cat("Uzoracki koeficijent asimetrije",g1,"\n")
cat("Realizovana vrednost test statistike:",Z,"\n")
cat("Kriticna vrednost:",c,"\n")
if(abs(Z)>=c)
cat("Raspodela obelezja je asimetricna \n")
else
cat("Raspodela obelezja je simetricna \n") }
```

Тест знакова

```
test.znakova<-function(
niz,m0,alfa, alternativa=c("vece","manje","razlicito")){
#parametri f-je su:  dobijeni uzorak, pretpostavljena vrednost
# medijane, prag znacajnosti, alternativna hipoteza
m<-length(niz)  #obim uzorka
k<-c()
j<-1
for(i in 1:m)
{if(niz[i]>m0)
{k[j]<-1
#ako je element niza veci od medijane upisujemo u vektor k
#vrednost 1, a ako je manji upisujemo nulu
j<-j+1}
else if(niz[i]<m0)
{k[j]<-0
j<-j+1}
# ako ima clanova niza jednakih m0, u vektor k se nista ne upisuje
}
test.stat<-sum(k) # vrednost test statistike
p.vred<-0
cat("Vrednost test statistike je",test.stat,"\n")
n<-length(k) #broj clanova niza k

# racunanje P-vrednosti u zavisnosti od oblika alt.hipoteze
if(alternativa=="vece")
{p.vred<-1-pbinom(test.stat-1,size=n,prob=0.5)}
else if (alternativa=="manje")
{p.vred<-pbinom(test.stat,size=n,prob=0.5)}
else
{p.vred<-2*min(1-pbinom(test.stat-1,size=n,prob=0.5),
                  pbinom(test.stat,size=n,prob=0.5))}
cat("P-vrednost testa:",p.vred,"\ n")

if(p.vred<alfa) {cat("Odbacujemo nultu hipotezu.\n")}
else {cat("Ne odbacujemo nultu hipotezu.\n")}
}
```

Рачунање функција моћи теста знакова, Вилкоксоновог теста означених рангова и Студентовог t - теста за један узорак

```
#generisanje uzorka obima n iz izabrane raspodele,
#uzorku se dodaje konstanta c, u skladu sa alternativnom hipotezom
#za generisanje vrednosti iz Laplasove raspodele koriscena je
#f-ja rlaplace koja se nalazi u paketu VGAM
raspodela1<-function(n,c,izbor=c("normalna","laplasova","logisticka",
         "uniformna")) {
if(izbor=="normalna")
{x<-rnorm(n,mean=c,sd=1)}
else if(izbor=="laplasova")
{x<-rlaplace(n,location=c,scale=1)}
else if(izbor=="logisticka")
{x<-rlogis(n,location=c,scale=1)}
else {x<-runif(n,min=-1/2+c,max=1/2+c)}
return(x)
}


#matrica N uzoraka iz odabrane raspodele
#uzorci se smestaju u vrste:  N vrsta,n kolona
uzorci1<-function(n,c,N,izbor) {
z<-matrix(,nrow=N,ncol=n,byrow=T)
for(i in 1:N){z[i,]<-raspodela1(n,c,izbor)}
return(z)
}


#Racunanje minimalne velicine W i testa znakova koja ce se
#koristiti kod sva tri testa prilikom racunanja kriticne vrednosti
velicina.testa1<-function(n,alfa) {
zkrit.vrednost<-qbinom(1-alfa,size=n,prob=0.5)+1
p1<-(1-pbinom(zkrit.vrednost-1,size=n,prob=0.5))
wkrit.vrednost<-qsignrank(1-alfa,n)+1
p2<-(1-psignrank(wkrit.vrednost-1,n))
```

```
if(abs(p1-p2)>0.001)
{return(min(p1,p2))}
else
{if(runif(1)<=0.5)
return(p1)
else
return(p2)}
}


#Racunanje moci, redom, testa znakova, Vilkoksonovog testa
#oznacenih rangova (W testa),i t-testa.
#Moc oba testa se racuna kao
#broj odbacivanja nulte hipoteze/ukupan broj uzoraka N
znakovi.moc<-function(n,niz,N,alfa=0.05) {
ind<-rep(0,N)
z<-rep(0,N)
zkrit.vrednost<-qbinom(1-velicina.testa1(n,alfa),size=n,prob=0.5)+1
for(k in 1:N){
x<-niz[k,]
for(i in 1:n){
if(x[i]>0) {z[k]<-z[k]+1} }

if(z[k]>=zkrit.vrednost) {ind[k]<-1}
}
moc<-sum(ind)/N
return(moc)
}

rangovi.moc<-function(n,niz,N,alfa=0.05) {
ind<-rep(0,N)
T.plus<-rep(0,N)
rkrit.vrednost<-qsignrank(1-velicina.testa1(n,alfa),n)+1
for(k in 1:N){
x<-niz[k,]
d<-x
aps.razlike<-abs(d)
rangovi<-rank(aps.razlike)
```

```
for(i in 1:n){
if(d[i]>0) {T.plus[k]<-T.plus[k]+rangovi[i]} }

if(T.plus[k]>=rkrit.vrednost) {ind[k]<-1}
}
moc<-sum(ind)/N
return(moc)
}

t.moc1<-function(n,z,N,alfa=0.05) {
ind<-rep(0,N)
T<-rep(0,N)
Tkrit.vrednost<-qt(1-velicina.testa1(n,alfa),n-1)
for(k in 1:N){
x<-z[k,]
sigma<-sd(x)
T[k]<-mean(x)*sqrt(n)/sigma
if(T[k]>=Tkrit.vrednost) {ind[k]<-1}
}
moc<-sum(ind)/N
return(moc)
}


#Objedinjavanje prethodne tri funkcije za racunanje moci W,
#testa znakova i t-testa u jednu funkciju
moc.testa1<-function(n,c,N,izbor,alfa,test=c("znakovi","rangovi","t1"))
{
sim.niz<-uzorci1(n,c,N,izbor)
moc<-0
if(test=="znakovi")
{moc<-znakovi.moc(n,sim.niz,N,alfa)}
else if(test=="rangovi") {moc<-rangovi.moc(n,sim.niz,N,alfa)}
else {moc<-t.moc1(n,sim.niz,N,alfa)}
return(moc)
}
```

```
#Graficko predstavljanje moci sva tri testa.
#Korisnik bira broj tacaka na grafiku,ukupan broj uzoraka N
# za koje se vrsi simulacija iz odabrane raspodele,
#razliku izmedju nulte i alternativne hipoteze, gresku prve vrste
fja.moci1<-function(br.tacaka,n,c=rep(0,br.tacaka),N,izbor,alfa){
znakovi.moci<-rep(0,br.tacaka)
rangovi.moci<-rep(0,br.tacaka)
t.moci<-rep(0,br.tacaka)
for(i in 1:br.tacaka)
{znakovi.moci[i]<-moc.testa1(n,c[i],N,izbor,alfa,"znakovi")
rangovi.moci[i]<-moc.testa1(n,c[i],N,izbor,alfa,"rangovi")
t.moci[i]<-moc.testa1(n,c[i],N,izbor,alfa,"t1")
}
plot(c,znakovi.moci[1:br.tacaka],pch=21,type="o",xlim=c(0,c[br.tacaka]),
      ylim=c(0,1),col="blue",bg="black",xlab="c",
      ylab="Funkcija moci testa",font.lab=3)
lines(c,rangovi.moci[1:br.tacaka],pch=22,type="b",col="green",
      bg="black")
lines(c,t.moci[1:br.tacaka],pch=24,type="b",col="red",bg="black")
legend("bottomright",legend<-c(expression(italic("K test")),
      expression(italic("W test")),expression(italic("t test"))),
      col=c("blue","green","red"),pch=c(21,22,24),pt.bg="black")
}
```

Рачунање моћи Вилкоксоновог теста суме рангова и Студентовог t - теста за два независна узорка

```
#generisanje dva uzorka iz izabrane raspodele, obima m i n,
#drugom uzorku se dodaje konstanta c, u skladu sa alternativnom hipotezom
raspodela<-function(m,n,c,izbor=c("normalna","laplasova",
         "logisticka","uniformna")){
if(izbor=="normalna"){
y<-rnorm(n,mean=0,sd=1)
x<-rnorm(m,mean=0,sd=1)+c}
else if(izbor=="laplasova"){
y<-rlaplace(n,location=0,scale=1)
x<-rlaplace(m,location=0,scale=1)+c}
else if(izbor=="logisticka"){
y<-rlogis(n,location=0)
x<-rlogis(m,location=0)+c}
else{
y<-runif(n,min=-1/2,max=1/2)
x<-runif(m,min=-1/2,max=1/2)+c}
return(c(x,y))
}


#matrica N uzoraka iz odabrane raspodele
#uzorci se smestaju u vrste:  N vrsta,m+n kolona
uzorci<-function(m,n,c,N,izbor) {
z<-matrix(,nrow=N,ncol=m+n,byrow=T)
for(i in 1:N){z[i,]<-raspodela(m,n,c,izbor)}
return(z)
}


#Racunanje moci Vilkoksonovog testa sume rangova i
#t-testa za nezavisne uzorke
#t-test koristi istu velicinu testa dobijenu kod W-testa
#Moc oba testa se racuna kao
```

```
# broj odbacivanja nulte hipoteze/ukupan broj uzoraka N
vilkoks.moc<-function(m,n,z,N,alfa=0.05) {
ind<-rep(0,N)
w<-rep(0,N)
Wkrit.vrednost<-qwilcox(1-alfa,m,n)+1+m*(m+1)/2

for(k in 1:N){
x<-z[k,1:m]
y<-z[k,(m+1):(m+n)]
sort.niz<-sort(c(x,y))
xs<-sort(x)
for(i in 1:(m+n))
{for(j in 1:m)
{if(sort.niz[i]==xs[j]) {w[k]<-w[k]+i}
}}

if(w[k]>=Wkrit.vrednost) {ind[k]<-1}
}
moc<-sum(ind)/N
return(moc)
}

velicina.testa<-function(m,n,alfa){
krit.vrednost<-qwilcox(1-alfa,m,n)+1
return(1-pwilcox(krit.vrednost-1,m,n))
}

t.moc<-function(m,n,z,N,alfa=0.05) {
ind<-rep(0,N)
T<-rep(0,N)
Tkrit.vrednost<-qt(1-velicina.testa(m,n,alfa),m+n-2)
for(k in 1:N){
x<-z[k,1:m]
y<-z[k,(m+1):(m+n)]
sigma<-sqrt(((m-1)*var(x)+(n-1)*var(y))/(m+n-2))
T[k]<-(mean(x)-mean(y))/(sigma*sqrt(1/m+1/n))
if(T[k]>=Tkrit.vrednost) {ind[k]<-1}
}
```

```
moc<-sum(ind)/N
return(moc)
}



#Objedinjavanje funkcija za racunanje moci W i t-testa
#u jednu funkciju
moc.testa<-function(m,n,c,N,izbor, alfa,test=c("t","w")){
sim.niz<-uzorci(m,n,c,N,izbor)
moc<-0
if(test=="w")
{moc<-vilkoks.moc(m,n,sim.niz,N,alfa)}
else {moc<-t.moc(m,n,sim.niz,N,alfa)}

return(moc)
}



#Graficko predstavljanje moci oba testa.
#Korisnik bira broj tacaka na grafiku, broj razlicitih obima uzoraka,
#ukupan broj uzoraka N za koje se vrsi simulacija iz odabrane
raspodele,
# razliku izmedju nulte i alternativne hipoteze, gresku prve vrste
fja.moci<-function(br.tacaka,m,n,c=rep(0,br.tacaka),N,izbor,alfa){
w.moci<-rep(0,br.tacaka)
t.moci<-rep(0,br.tacaka)
for(i in 1:br.tacaka){
w.moci[i]<-moc.testa(m,n,c[i],N,izbor,alfa,"w")
t.moci[i]<-moc.testa(m,n,c[i],N,izbor,alfa,"t")
}
plot(c,w.moci[1:br.tacaka],pch=22,type="b",xlim=c(0,c[br.tacaka]),
      ylim=c(0,1),col="green",xlab="c",
      ylab="Funkcija moci testa",bg="black",font.lab=3)
lines(c,t.moci[1:br.tacaka],pch=24,type="b",col="red",bg="black")
legend("bottomright",legend<-c(expression(italic("Ws test")),
      expression(italic("t test"))),col=c("green","red"),pch=c(22,24),
      pt.bg="black")
}
```

Литература

[1] Blair, R.Clifford and Higgins, James J.(1980). A comparison of the power of Wilcoxon's rank - sum statistics to that of Student's t statistic under various non - normal distributions, *Journal of Educational Statistics,* **5,** 309-335.

[2] Gibbons, Jean Dickinson and Chakraborti, Subhabrata (2003). Nonparametric Statistical Inference, Mercel Dekker, Inc., New York.

[3] Daniel, Wayne W. (1978). Applied Nonparametric Statistics, Houghton Mifflin Company, Boston.

[4] Ђорић, Драган, Јевремовић, Весна, Малишић, Јован, Николић-Ђорић, Емилија (2007). Атлас расподела, Грађевински факултет, Београд.

[5] Kvam, Paul H. and Vidakovic, Brani (2007). Nonparametric Statistics with Applications to Science and Engineering, John Wiley & Sons, New York.

[6] Noether, Gottfried E. (1955). On a Theorem of Pitman, *The Annals of Mathematical Statistics,* **26,** 64-68.

[7] Sprent, Peter and Smeeton, Nigel C. (2001). Applied Nonparametric Statistical Methods, Chapman & Hall/CRC, Boston.

[8] Sheskin, David J. (2004). Handbook of Parametric and Nonparametric Statistical Procedures, Chapman & Hall/CRC, Boca Raton.

[9] Hodges, J.L. and Lehmann, E.L. (1956). The efficiency of some nonparametric competitors of the t - test, *Annals of Mathematical Statistics,* **27,** 324-335.

[10] Hollander, Myles and Wolfe, Douglas A. (1973). Nonparametric Statistical Methods, John Wiley & Sons, New York.

[11] Conover, W.J. (1980). Practical Nonparametric Statistics, John Wiley&Sons, New York.

[12] Cohen, Yosef and Cohen, Jeremiah Y. (2008). Statistics and Data with R, John Wiley & Sons, New York.