Peatükk 2

Wilcoxoni astakmärgitest (*Wilcoxon Signed-Rank Test*)

2.1 Motivatsioon ja teststatistik

Wilcoxoni astakmärgitesti kasutatakse kahe sõltuva valimi võrdlemiseks. Oletame näiteks, et soovime võrrelda, millise väetise abil saame parema saagi. Iga katsepõld jagatakse kaheks pooleks, millest ühte poolt väetatakse ühe väetisega, teist poolt aga teise väetisega. Sügisel vaadatakse kuidas saagiga lood on. Näiteks võis vaatluse all olla kolm põldu ja olgu saadud saagid järgmised:

	I pool	II pool	erinevus
1. põld	76	78	2
2. põld	82	91	9
3. põld	80	86	6

Kahel põllul oli parem uue väetisega väetatud pool, ühel põllul andis parema tulemuse vana väetis. Kas see on ükskõik, millisel põllul oli vana väetis parem? Vaatame kahte hüpoteetilist olukorda. Ühel juhul on uus väetis paaril põllul mäekõrguselt parem vanast (6 ja 9 ühikut) ja ühel põllul oli vana väetisega väetatud pool napilt parem (2 ühikut). Teisel juhul annab uue väetisega väetatud pool kahel põllul napilt paremat saaki (2 ja 6 ühikut), üks vana väetisega väetatud põllupool on aga märkimisväärselt parem kui uue väetisega väetatud põllupool (9 ühikut). Kas mõlemal kirjeldatud juhul on andmed uue väetise kasulikkusest sama tugevad? Või arvame, et esimesena kirjeldatud juhul usume meelsamini uue väetise paremusse kui teisena kirjeldatud juhu puhul?

Sisetunnet jälgides võiks arvata, et on võimalik teha midagi paremat kui kasutada lihtsalt märgitesti. Wilcoxoni astakmärgitest võtabki arvesse lisaks erinevuse märgile ka erinevuse (absoluutväärtuse) suurust (astakut):

$$W = \sum_{i=1}^{n} sgn(X_i) rank(|X_i|)$$

Võrdluseks võib tuua märgitesti statistiku: $S(X,0) = \sum_{i=1}^{n} sgn(X_i)$.

2.2 Statistiku W jaotus nullhüpoteesi kehtides

Kui uus ja vana väetis on võrdväärselt head, siis tulenevad erinevused saakides lihtsalt põllupoolte erinevusest. Kumba põllupoolt väetatakse uue väetisega, on heas katses valitud juhuslikult (otsustatakse mündiviske abil vms). Kui vaatame saagikuste erinevusi — saak uue väetiega miinus saak vana väetisega — siis nullhüpoteesi kehtides peaksid märgid erinevuste ees olema täiesti juhuslikud. Antud näites on vaadeldud kolme põldu, seega on võimalikke märgikombinatsioone erinevuste ees $2^3 = 8$, ning järgnevad katsetulemused peaksid kõik olema võrdvõimalikud:

erinevused (x_i)			$sgn(x_i)rank(\|x_i\|)$			$W(x_i)$	$P(W = W(x_i))$
-2	-6	-9	-1	-2	-3	$W = -6$	1/8
$+2$	-6	-9	$+1$	-2	-3	$W = -4$	1/8
-2	$+6$	-9	-1	$+2$	-3	$W = -2$	1/8
$+2$	$+6$	-9	$+1$	$+2$	-3	$W = 0$	1/8
-2	-6	$+9$	-1	-2	$+3$	$W = 0$	1/8
$+2$	-6	$+9$	$+1$	-2	$+3$	$W = +2$	1/8
-2	$+6$	$+9$	-1	$+2$	$+3$	$W = +4$	1/8
$+2$	$+6$	$+9$	$+1$	$+2$	$+3$	$W = +6$	1/8

Kokkuvõttes võime välja kirjutada Wilcoxoni astakmärgitesti statistiku W jaotuse juhul, kui $n = 3$, vaata tabelit 2.1.

Tabel 2.1: Statistiku W jaotus nullhüpoteesi kehtides, $n = 3$

w	-6	-4	-2	0	2	4	6
P(W=w)	1/8	1/8	1/8	2/8	1/8	1/8	1/8

Oletame nüüd, et meie katses tulid erinevused 6 ja 9 uue väetise kasuks ning erinevus 2 vana väetise kasuks. Seega $W = -4$. Leiame olulisustõenäosuse (p-value) — milline on tõenäosus näha sedavõrd ekstremaalset või veel ekstremaalsemat statistiku väärtust? Ühepoolse hüpoteesi korral, kui uus väetis saab olla vaid kas samahea või parem kui vana väetis, oleksid sama ekstreemse ($W = -4$) või veel ekstreemsema ($W = -6$) statistiku väärtuse saamise tõenäosus nullhüpoteesi kehtides $1/8 + 1/8 = 0,25$, seega olulisustõenäosus on ühepoolse hüpoteesi korral 0,25. Kahepoolse hüpoteesi korral (kui uus väetis võib põhimõtteliselt olla ka kehvem kui vana) on sama ekstreemse või veel ekstreemsemate statistiku väärtuste hulk $\{-6, -4, 4, 6\}$ suurem ja testi olulisustõenäosuseks tuleb 0,25*2=0,5. Seega antud juhul ei suuda me alternatiivset hüpoteesi tõestatuks lugeda mistahes mõistlikku olulisuse nivood kasutades.

Võime ka tähele panna, et kolme vaatluse korral poleks ühegi katsetulemuse korral võimalik tõestada uue väetise paremust vanast olulisuse nivool 0,05!

Seega üldjuhul on võimalik Wilcoxoni astakmärgitesti olulisustõenäosust leida näiteks nii: vaatame läbi 2^n erinevat märgikombinatsiooni astakute ees, leiame iga kombinatsiooni jaoks statistiku W väärtuse ning kirjutame välja W jaotuse juhul, kui nullhüpotees kehtib. Saadud jaotust kasutades on juba lihtne leida testi olulisustõenäosust. Tasub meeles pidada, et iga valimi suuruse n korral tuleb statistiku W jaotus erinev!

Samas pole olulisustõenäosuse leidmiseks tarvis tervet jaotust välja kirjutada. Piisab, kui loeme kokku ainult nähtud W väärtuse ja temast veel ekstreemsemate väärtuste saamise tõenäosused nullhüpoteesi kehtides.

Vaatame järgmist näidet

Näide 2.1 *Ravimifirma soovis uurida, kas tema poolt välja töötatud spetsiaalset haavaplaastrit kasutades paranevad haavad kiiremini kui traditsioonilist ravimeetodit (haavade kokkuõmblemist) kasutades. Uuringusse värvati 10 katsealust (rotti) ja igale rotile tehti kaks samapikka sisselõiget. Üks haavadest õmmeldi kokku (kasutati traditsioonilist raviviisi) ja teine plaasterdati kinni. Lasti siis haavadel paar päeva paraneda. Haava paranemise headust mõõdeti paari päeva pärast järgmiselt: vaadati, kui tugevat jõudu läheb vaja haava uuesti lahti rebimiseks (seega suuremad numbrid näitavad paremini paranenud haavu). Katsetulemused on toodud tabelis 2.2.*

Leiame kogutud andmete põhjal Wilcoxoni astakmärgitesti statistiku väärtuse:

$$W = 6 + 9 - 1 - 7 + 4 + 5 + 10 + 8 + 3 + 2 (= 39).$$

Tabel 2.2: Kliiniline katse rottidel parima haavade raviviisi leidmiseks

rott	plaaster	õmblus	plaaster-õmblus	$sgn(x_i)rank(\|x_i\|)$
1	659	452	207	6
2	984	587	397	9
3	397	460	-63	-1
4	574	787	-213	-7
5	447	351	96	4
6	479	277	202	5
7	676	234	442	10
8	761	516	245	8
9	647	577	70	3
10	577	513	64	2

Antud juhul on erinevaid võimalusi astakute ette märke panna $2^{10} = 1024$*, kõiki neid läbi vaadata on raske. Aga vaatame, mitu märgikombinatsiooni on sellist, mille puhul tuleks teststatistiku väärtus 39 või suurem. Astakud, mille ette miinusmärgi pannes saame sama ekstreemse või veel ekstreemsema statistiku väärtuse, on ära toodud alljärgnevas tabelis:*

miinusmärkide arv	*astakud*	*juhte kokku*
0	$\emptyset$	*1*
1	*1, 2, 3, 4, 5, 6, 7, 8*	*8*
2	*(1;2), (1;3),(1;4),(1;5),(1;6),(1;7), (2;3),(2;4),(2;5),(2;6),(3;4),(3;5)*	*12*
3	*(1;2;3), (1;2;4), (1;2;5), (1;3;4)*	*4*

Kokku on seega neist 1024-st juhust selliseid, mille puhul W *väärtus tuleks nähtud väärtusest suurem* $1 + 8 + 12 + 4 = 25$*. Kui me teaksime ette (juba enne andmete kogumist/nägemist), et plaaster saab õmblusest vaid parem olla (mitte halvem), võiksime kontrollida ühepoolset hüpoteesi. Sellisel juhul oleks sedavõrd ekstreemse või veel ekstreemsema statistiku väärtuse nägemise tõenäosus nullhüpoteesi kehtides (p-value)* $25/1024 = 0.0244\ldots$*. Kui me ei saa kindlalt öelda, et "plaastriga haavad ei saa paraneda kehvemini kui õmblust kasutades", peaksime kasutama kahepoolset hüpoteesi. Sel juhul saaksime olulisustõenäosuseks* $0.0244 * 2 = 0,0488$*. Antud näite korral saame seega tõestatuks lugeda (olulisuse nivool 0,05), et plaaster parandab haavu paremini.*

Juhul kui vaatluste arv läheb suureks, tekivad siiski enamasti probleemid olulisustõenäosuse leidmisega kirjeldatud meetodi(te) abil. Lahenduseks oleks statistiku W asümptootilise jaotuse kasutamine.

2.3 Asümptootiline jaotus

Astakmärgitesti statistik on kirja pandav kui juhuslike suuruste summa, $W = \sum_{i=1}^{n} Y_i$, kus $Y_i := sgn(X_i)rank(|X_i|)$. Juhul kui igal uuritaval objektil on valitud juhuslikult millises järjekorras me töötluseid rakendame (juhuslikult valime kumba haavadest saab plaastri; valime juhuslikult põllupoole kuhu külvame sorti A jne) saame liidetavaid lugeda sõltumatuteks. Sõltumatute juhuslike suuruste summa on üldjuhul asümptootiliselt normaaljaotusega. (täpsemat tingimust vaatame kohe).

Eeldame, et nullhüpotees kehtib. Sellisel juhul tuleneb statistiku juhuslikkus puhtalt töötluste randomiseerimisest, seega suurused $rank(|X_i|)$ on käsitletavad kui fikseeritud konstandid ja juhuslikud on ainult suurused $sgn(X_i)$:

$$\begin{aligned} P(sgn(X_i) = -1) &= 0.5 \\ P(sgn(X_i) = 1) &= 0.5 \end{aligned}$$

Seega (nullhüpoteesi kehtides):

$$\begin{aligned} EW &= E\sum_{i=1}^{n} Y_i \\ &= \sum_{i=1}^{n} rank(|X_i|)E(sgn(X_i)) \\ &= \sum_{i=1}^{n} rank(|X_i|)0 \\ &= 0 \end{aligned}$$

ja

$$\begin{aligned} DW &= D\sum_{i=1}^{n} Y_i \\ &= \sum_{i=1}^{n} rank(|X_i|)^2 D(sgn(X_i)) \\ &= \sum_{i=1}^{n} rank(|X_i|)^2 \\ &= 1^2 + 2^2 + 3^2 + \ldots n^2 \\ &= n(n+1)(2n+1)/6 \end{aligned}$$

Vahemärkus: $\sum_{i=1}^n i^2 = n(n+1)(2n+1)/6$ selles veendumiseks kasuta kasvõi induktsiooni:

$$\begin{aligned}
\sum_{i=1}^n i^2 &= \sum_{i=1}^{n-1} i^2 + n^2 \\
&= (n-1)n(2n-1)/6 + n^2 \\
&= n(2n^2 - n - 2n + 1)/6 + n^2 \\
&= n(2n^2 - 3n + 1 + 6n)/6 \\
&= n(2n^2 + 3n + 1)/6 \\
&= n(n+1)(2n+1)/6
\end{aligned}$$

$$W = \sum_{i=1}^n sgn(X_i) rank(|X_i|)$$

Üritame nüüd veenduda, et suurte valimite (suur n) korral on statistiku W jaotus ligilähedaselt normaaljaotusega. Tõestusel tugineme Lindeberg-Feller'i teoreemile.

Olgu antud sõltumatute juhuslike suuruste seeriad (igas reas olevad juhuslikud suurused olgu omavahel sõltumatud, veerud võivad olla sõltuvad või osaliselt isegi kokku langeda):

$$\begin{aligned}
&X_{11} \\
&X_{21}, X_{22} \\
&X_{31}, X_{32}, X_{33} \\
&\ldots .
\end{aligned}$$

Olgu $\mathrm{E}X_{ij} = 0$, $\mathrm{D}X_{ij} = \sigma_{ij}^2$, $S_n := \sum_{j=1}^n X_{nj}$, $\mathrm{D}S_n := \sum_{j=1}^n \sigma_{nj}^2$.

Lindeberg-Felleri teoreem väidab: Kui mistahes $\varepsilon > 0$ korral

$$\frac{1}{\mathrm{D}S_n} \sum_{j=1}^n E(X_{nj}^2 I(|X_{nj}| \geq \varepsilon \mathrm{D}S_n)) \to 0, \text{ kui } n \to \infty$$

siis

$$S_n / \sqrt{\mathrm{D}S_n} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Kuna liidetavad Wilcoxoni testi puhul on (nullhüpoteesi kehtides) juhuslikud ainult märgi poolest, on $|X_{nj}|$ ja X_{nj}^2 tegelikult konstandid (astak ja astaku ruut). Kuna aga maksimaalne astak on n, siis $|X_{nj}| \leq \max_{j \leq n} |X_{nj}| = n$ ja järelikult:

$$\begin{aligned}
\frac{1}{\mathrm{D}S_n}\sum_{j=1}^{n} E(X_{nj}^2 I(|X_{nj}| \geq \varepsilon \mathrm{D}S_n)) &\leq \frac{1}{\mathrm{D}S_n}\sum_{j=1}^{n} E(X_{nj}^2 I(n \geq \varepsilon \mathrm{D}S_n)) \\
&= I(n \geq \varepsilon \mathrm{D}S_n)\frac{1}{\mathrm{D}S_n}\sum_{i=1}^{n} i^2 \\
&= I(n \geq \varepsilon \mathrm{D}S_n) \\
&= I(n/\mathrm{D}S_n \geq \varepsilon)
\end{aligned}$$

Viimane avaldis koondub aga nulliks:

$$n/\mathrm{D}S_n = \frac{n}{n(n+1)(2n+1)/6} \to 0 \text{ kui } n \to \infty.$$

Järelikult on Lindebergi tingimus rahuldatud ja nullhüpoteesi kehtides $W/\sqrt{\mathrm{D}W} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

Näide 2.2 *Proovitakse, kas päikesekreem ikka kaitseb nahka päikese eest. Katsealustel määritakse lapike nahka kreemiga kokku. Saadetakse nad siis randa. Hiljem mõõdetakse naha punetust nii kreemitatud kohast kui ka selle kõrvalt. Naha punetuse erinevused (kaitsmata koht-kreemitatud koht) on järgmised:*

-1, -3, 4, -6, 7, 8, -9, -10, 12, 14, 15, -17, 20, 25, 29, 30, -32, 40, 78, 102.

Wilcoxoni astakmärgitesti statistiku väärtuseks saame:

$$\begin{aligned}
W &= -1-2+3-4+5+6-7-8+9+10+ \\
&\quad +11-12+13+14+15+16-17+18+19+20 = 108
\end{aligned}$$

Nullhüpoteesi kehtides ootaksime antud valimi suuruse juures statistiku keskväärtuseks ja dispersiooniks $EW = 0$ ja $DW = n(n+1)(2n+1)/6 = 20*21*41/6 = 2870$. Seega peaks nullhüpoteesi kehtides juhuslik suurus $(W - EW)/\sqrt{DW} = W/53,57$ olema ligikaudu standartse normaaljaotusega juhuslik suurus (nullhüpoteesi kehtides peaks vastav suurus ootuspäraselt jääma vahemikku -1,96...1,96). Antud valimi korral aga $W/53,57 = 108/53,57 = 2,016$, seega midagi tundub olevat viltu... nullhüpoteesi eeldusega. Olulisustõenäosuse saame, kui leiame, kui tõenäoliselt oleksime näinud sedavõrd ekstreemset statistiku väärtust või veel ekstreemsemat nullhüpoteesi kehtides. Tõenäosus näha veel suuremat statistiku väärtust H_0 kehtides *oleks* $1 - \Phi(2,016) = 0,0219$ *(olulisustõenäosus ühepoolse hüpoteesi korral). Kahepoolse hüpoteesi korral peaksime muidugi siia juurde liitma ka erakordselt väikese* W *saamise tõenäosuse (nullhüpoteesi kehtides). Normaaljaotuse sümmmetriat arvestades saame tulemuseks* $(1-\Phi(2,016))*2 = 0,0438$*. Seega on olulisustõenäosus väike ja me võime nullhüpoteesi (olulisuse nivool 0,05) kummutatuks lugeda. Võrdlusena: täpne olulisustõenäosus oleks 0,044054.*

2.3.1 Pidevuse parandus

Statistikas tuleb sageli ette olukordi, kus me lähendame diskreetset jaotust pideva jaotusega. Näiteks võime lähendada binoomjaotust normaaljaotusega või Wilcoxoni astakmärgitesti statistiku jaotust normaaljaotusega. Kui meie valimid on väga suured, siis võime sellist lähendamist teha üsna sirgjooneliselt ja probleemideta. Keskmise suurusega valimite korral saame aga sageli veidi täpsemaid tulemusi, kui enne arvutuste alustamist veidi mõtleme.

Vaatame näiteks juhuslikku suurust $X \sim B(10, 0.4)$. Antud binoomjaotust saab lähendada normaaljaotusega juhusliku suurusega $Y \sim N(4, \sigma^2 = 2,4)$. Kuidas leida normaaljaotuse lähendit kasutades tõenäosust, et $X = 1$, $P(X = 1) = 0,0403$? On selge, et antud tõenäosus pole ligilähedaseltki võrdne tõenäosusega $P(Y = 1) = 0$. Parem variant antud tõenäosuse leidmiseks oleks vaadata tõenäosust $P(0.5 \leq Y \leq 1.5) = 0,0414$.

Kuidas aga oleks mõistlik leida normaaljaotuse lähendit kasutades suurust $P(X \leq 1) \approx 0,04636$? Me võime mõelda ka nii: $P(X \leq 1) = P(X = 1) + P(X = 0)$. Viimast tõenäosust aga lähendab tõenäosus $P(Y \leq 1.5) = 0,0533$ palju täpsemalt kui tõenäosus $P(Y \leq 1) = 0,026$. Seega soovitatakse kasutada diskreetsete juhuslike suuruste lähendamisel pidevuse korrektsiooni (*continuity correction*): $P(X \leq c) \approx P(Y + 0,5 \leq c)$ või $P(X \geq c) \approx P(Y - 0,5 \geq c)$.

Vaatame pidevuse parandust eelnenud näite varal. Leidmaks tõenäosust $P(W \geq 108)$ oleks meil seega soovitav vaadata normaaljaotuse jaotusfunktsiooni kohal $(W - 1)/53,57$ (kuna W muutub meil sammuga 2), mis annaks kahepoolse hüpoteesi puhul olulisustõenäosuseks $(1 - \Phi(1,9974)) * 2 = 0,0458$. Antud näites ei andnud pidevuse paranduse kasutamine täpsemat tulemust — kui ilma pidevuse paranduseta alahindasime olulisustõenäosust, siis pidevuse parandust kasutades saime liiga suure hinnangu olulisustõenäosusele. Lähend on kõigest lähend ja vahel võib pidevuse paranduse kasutamine ka halvemaid tulemusi anda. Enamasti saadakse siiski pidevuse parandust kasutades täpsemaid tulemusi (ja ka antud näite puhul on arvatavasti konservatiivsem hinnang olulisustõenäosusele eelistatum).

2.3.2 Võrdsed astakud

Pideva tunnuse korral on kahe täpselt võrdse vaatluse saamise tõenäosus 0. Paraku ei taha aga praktika sugugi kuulata teooria sõna. Mõõtmiste ebatäpsusest, vahetulemuste ümmardamisest vms tingitult esineb ikka samasuuri väärtuseid.

Kui mõlema töötluse korral saadakse samasuur uuritava tunnuse väärtus (ei oska öelda, kumb töötlus andis parema tulemuse), ehk töötluste erinevus on 0, siis taolised vaatlused visatakse tüüpiliselt lihtsalt testimise ajaks valimist välja (vähendatakse valimi suurust).

Kui meil on kahel uuritaval indiviidil toimunud (mõõtmistäpsuse piires) samasuur muutus (muutuse absoluutväärtust arvestades) ei oska me ühele muutusele teisest kõrgemat kohta (astakut) omistada. Lahendus on muidugi üsna lihtne — anname kõigile samasuurtele muutustele ühe ja sama astaku väärtuse, milleks oleks nende astakute keskmine. Näiteks kui me ei suuda otsustada, kumb vaatlustest pretendeerib 4. ja kumb 5. kohale, siis anname mõlemale vaatlusele astakuks $(4+5)/2 = 4,5$.

Muutes taolisel viisil astakuid muutuvad muidugi ka väärtused, milliseid statistik W omandada võib. Seega muutub ka Wilcoxoni statistiu jaotus nullhüpoteesi kehtides. Kui vaatame ise paberil läbi statistiku kõik võimalikud väärtused, siis arvatavasti ei teki sellest suuremat probleemi — statistiku jaotustabel tuleb lihtsalt veidi teistsugune kui siis, kui võrdseid vahesid poleks esinenud.

Soovides kasutada statistiku W asümptootilist jaotust, peame samuti arvestama võrdsete vaatlustega — need muudavad mõnevõrra statistiku asümptootilist jaotust.

On lihtne näha, et keskväärtus ei muutu — nullhüpoteesi kehtides jätkuvalt $\mathrm{E}W = 0$. Küll aga muutub Wilcoxoni astakmärgitesti statistiku dispersioon (erinevate väärtuste arv kahaneb — järelikult kahaneb ka varieeruvus). Leiame, kuisuur on muutus.

$$\begin{aligned} DW &= D\sum_{i=1}^{n} Y_i \\ &= \sum_{i=1}^{n} \mathrm{midrank}(|X_i|)^2 D(sgn(X_i)) \\ &= \sum_{i=1}^{n} \mathrm{midrank}(|X_i|)^2 \end{aligned}$$

Oletame esmalt, et meil esineb võrdse suurusega vaatluseid vaid ühes kohas — vaatlused $(s+1)$ kuni $(s+t)$ olgu kõik samasuured. Nende vaatluste keskmiseks astakuks oleks $s + (t+1)/2$ ja taoliste astakute ruutude summa on

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

Juhul, kui võrdse suurusega vaatluseid poleks, oleks nendesamade astakute ruutude summaks olnud

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

Leiame, kui palju muutus astakute ruutude summa:

$$\begin{aligned}
\text{muutus} &= ts^2 + 2st(t+1)/2 + t(t+1)(2t+1)/6 - t\left(s^2 + s(t+1) + \frac{(t+1)^2}{4}\right) \\
&= t(t+1)(2t+1)/6 - t(t+1)^2/4 \\
&= t(t+1)[(2t+1)*2 - (t+1)*3]/12 \\
&= t(t+1)(t-1)/12 \\
&= t(t^2-1)/12.
\end{aligned}$$

Olgu meil valimis N' erinevat väärtust, kusjuures j. unikaalset väärtust esinegu t_j korda. Sellisel juhul saame Wilcoxoni astakmärgistatistiku dispersiooni (nullhüpoteesi kehtides) leida järgmise valemi abil:

$$DW = n(n+1)(2n+1)/6 - \sum_{j=1}^{N'} t_j(t_j^2-1)/12.$$

Kui võrdseid vaatluseid on palju, võib dispersioon märkimisväärselt väheneda. NB! Mitte kõik programmid ei pruugi võrdse suurusega vaatluseid kohates kõige mõistlikumalt käituda. Kui võrdseid astakuid on palju, siis võib Wilcoxoni astakmärgitesti statistiku jaotus koonduda normaaljaotuseks aeglaselt (vajame tõesti suurt valimit, enne kui võime kasutada lähendit!)

2.4 Erinevuse kirjeldamine

Mis saab siis, kui me nullhüpoteesi ümber lükkame? On ju tore, kui saame öelda: üus ravi on parem vanast"või üus väetis on parem kui vana", aga alati tekib küsimus: Kui palju parem? Kui tänu uuele väetisele saame hektarilt täiendavalt 1 grammi jagu rohkem saaki, siis ei tasu uue väetise eest kallimat hinda maksta. Aga kui tänu uuele ja kallimale väetisele saaksime tonni jagu rohkem saaki, tasuks muutus meile ära. Mõlema kirjeldatud juhu puhul võib Wilcoxoni astaksummatest võtta vastu alternatiivse hüpoteesi (piisavalt suure valimi korral).

Üks võimalus töötlustevahelise erinevuse suuruse kirjeldamiseks oleks järgmine. Me võime arutleda, kui palju me peaksime ühe töötlusega saadud tulemustest maha lahutama, et saada teise töötlusega võrdväärseid (sarnaseid) tulemusi.

Täpsemalt, otsitakse sellist Δ väärtust, et Wilcoxoni teststatistik kinnitaks peale Δ juurdeliitmist uuele töötlusele võimalikult hästi nullhüpoteesi ehk

$$W(Y - X + \Delta) = 0.$$

Antud võrrandil on sageli rohkem kui üks lahend, millisel juhul on sobilik raporteerida lahenditeks sobivate väärtuste lõigu keskpunkti.

Näide 2.3 *Vaatlustulemused (töötluste erinevused) on järgmised: -2,3,6,10,20. Wilcoxoni teststatistiku väärtuseks on* $W = -1 + 2 + 3 + 4 + 5 = 13$. *Selleks, et saaksime tulemuseks* $W = 0$, *peaksime kõigist erinevustest lahutama 6.5. Sellisel juhul oleksid meie vaatlusteks -8.5, -3.5, -0.5, 3.5, 13.5 ja statistiku väärtus* $W = -4 - 2.5 - 1 + 2.5 + 5 = 0$. *Seega* $\Delta = 6.5$.

Märkus: Otsitav väärtus, Δ, ei ole vahede (töötluste erinevuste) mediaan!

Tehniline märkus: Kuidas leida sobivat Δ väärtust? Üks võimalus oleks järgmine: leia kõikvõimalikud vahede keskmised:

$$[(-2)+(-2)]/2, [(-2)+3]/2, \ldots, [(-2)+20]/2, [3+3]/2, \ldots, [3+20]/2, \ldots [20+20]/2$$

ning leia siis nende $5 * (5 + 1)/2$ keskmise mediaan. Viimane ongi otsitavaks Δ väärtuseks! Kas oskad põhjendada, miks?

Saadud Δ väärtuse interpreteerimine võib ja ei pruugi raskuseid valmistada. Kui töötlused erineksid selle poolest, et üks töötlus lisaks uuritava tunnuse väärtusele alati Δ ühikut juurde (nihutaks uuritava tunnuse väärtuseid), ehk $Y_i = X_i + \Delta$ siis on meie saadud Δ väärtus hinnanguks sellele nihkele. Enamasti pole muidugi töötluse mõju kõigile ühesugune (väetamine aitab mõnda põldu enam kui teist; tänu ravile paraneb mõni inimene kibekiiresti ja teine (haigusest rohkem kurnatud) jääb veel mitmeks päevaks voodisse jne. Seega on igal inimesel/põllul oma isiklik töötluse mõju Δ_i. Juhul, kui juhuslike suuruste Δ_i jaotus oleks sümmeetriline, võiksime enda poolt leitud Δ väärtust käsitleda kui hinnangut juhuslike suuruste Δ_i'de mediaanile. Aga kas töötluste erinevuste jaotus on ikka sümmeetriline? Märkusena olgu mainitud: kui töötlused on eristamatud, siis on erinevuste jaotus $X - Y$ sümmeetriline.

Heal lapsel on ikka mitu nime. Meie poolt vaadeldud Δ't võib kirjanduses leida nii pseudomediaani (pseudomedian) kui ka Hodges-Lehmann'i hinnangu nime all.

Kui märgitesti statistikut kasutades saime leida usaldusintervalli mediaanile, nii saab Wilcoxoni astakmärgitesti kasutades leida usaldusintervalli pseudomediaanile. Tuleb lihtsalt leida nihete võimalike väärtuste hulk, mille puhul Wilcoxoni astakmärgitest jääb (antud olulisuse nivool) veel nullhüpoteesi juurde.

2.5 Ülesanded

1. Vahel tuntakse Wilcoxoni astakmärgitesti teststatistikuna ka statistikut
$$W_+ = \sum_{i=1}^{n} I(X_i > 0)\mathrm{rank}(|X_i|).$$
Näita, et Wilcoxoni astakmärgitesti olulisustõenäosus (ja testi otsus) milleni jõuame W_+ kasutamisel langeb kokku teststatistiku W kasutamisel saadud tulemustega. Eeldame, et uuritav tunnus on pidev, st $P(X_i = 0) = 0$.

2. Peale ainurakse jagunemist kaheks võetakse üks jagunemisel saadud ainurakne ja antakse talle kerge kuumashokk (kuumutatakse teda veidi). Teine jagunemisel saadud rakk elab senikaua normaaltingimustes. Seejärel pannakse mõlemad ainuraksed külma keskkonda. Jälgitakse, kuna algab eriliste "külmageenide"ekspressioon (kuna neid geene hakatakse kasutama). Soovitakse teada, kas eelnevalt kuumashoki saanud ainuraksetes lülituvad külmageenid sisse samakiiresti kui normaaltingimustes elanud rakul või on hilisemas reaktsioonikiiruses toimunud muutus. Eksperimendi käigus kogutud andmed on järgmised:

ainuraksete paari nr	kuumashoki saanu reaktsioonikiirus (ms)	reaktsioonikiirus normaaltingimustes (ms)
1	123	125
2	240	275
3	156	160
4	108	112
5	132	140
6	138	137
7	142	151
8	186	200

Kas suudad Wilcoxoni astakmärgitesti abil tõestada erinevust reaktsioonikiirustes? Milline on (kahepoolse) testi olulisustõenäosus (näita arvutusi) ja milline oleks sinu otsus. Seega võime nullhüpoteesi kummutada — ainuraksete reaktsioonikiirus tõuseb, kui neid eelnevalt kuumashokiga ehmatada (kui teisiti pole mainitud, kasutame tavapärast olulisuse nivood 0,05).