

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro darbas

STRUKTŪRINIŲ PASIKEITIMŲ EKONOMINĖSE SISTEMOSE
TYRIMAS

THE ANALYSIS OF STRUCTURAL CHANGES IN ECONOMIC SYSTEMS

Rekutė Gražvyda

VILNIUS 2007

Ekonometrinės analizės katedra

Darbo vadovas **prof. habil. dr. A. Račkauskas** _____
(parašas)

Darbas apgintas 2007 m. birželio mėn. 04 d.

Registravimo Nr. _____
2007-05-24 _____

Struktūrinių pasikeitimų ekonominėse sistemose tyrimas

Santrauka

Ekonomiką kartais ištinka įvairios krizės, pasitaiko stichinių nelaimių ar kitokių reiškinių, kurie iš esmės paveikia ekonomikos elgesį. Šis poveikis labai gerai atsispindi ekonominių rodiklių laiko eilutėse, kurių elgesys po tam tikros krizės ar stichinės nelaimės gali pakisti kardinaliai. Toks pasikeitimas paprastai vadinamas struktūriniu pasikeitimu. Tačiau dažnai pasitaiko atveju, kai plika akimi sunku įžiūrėti esamus struktūrinius pasikeitimus, todėl šią problemą būtų galima išspręsti taikant specialius testus struktūriniams pasikeitimams aptikti. Šiame darbe nagrinėjami trys žinomi testai struktūriniams pasikeitimams nustatyti. Darbo tikslas palyginti šių testų galią ir pasiūlyti naują testą, kuris pasižymėtų patrauklesnėmis charakteristikomis.

Nagrinėdami testus, skirtus struktūriniams pasikeitimams tikrinti, matėme, kad *CUSUM* testas nėra pritaikytas epideminio struktūrinio pasikeitimo atvejui, tuo tarpu kitų trijų testų, skirtų epideminiam struktūriniams pasikeitimams tikrinti, galia mažai skiriasi, tačiau naujai išvestas testas yra pranašesnis prieš kitus tuo, kad duomenyse esančių išskirčių netraktuoja kaip galimo struktūrinio pasikeitimo.

The Analysis of Structural Changes in Economic Systems

Summary

Sometimes there happen different crisis in economics, sometimes occur natural disasters or other phenomenons which substantially affect the behaviour of the economy. This impact reflects in a time series of economic indicators whose behaviour may go into the melting-pot after some crisis or natural disaster. Usually such a change is called a structural change. Very often structural change is hardly visible then this problem can be solved by applying special tests for detecting potential structural change. In this paper three known tests for identification of structural change are examined. The aim of this work is to compare the power of these tests and propose the new one which will have more desirable features.

During the investigation of the tests that are designed for detecting structural changes it emerged that *CUSUM* test is not fitted for the case of epidemic structural change. Whereas the power of other tree tests that are designed for detecting epidemic structural changes differ narrowly, however the new test has the advantage against the other tests – it does not treat an outlier as a possible structural change.

Turinys

Įvadas	6
1. Struktūriniai pasikeitimai.....	7
1.1. Struktūrinio pasikeitimo sąvoka	7
1.2. Struktūrinių pasikeitimų tipai	8
2. Testai, skirti struktūriniam pasikeitimams tikrinti	11
2.1. Chow testas	11
2.2. Fišerio statistika	14
2.3. Bajeso statistika	15
2.4. Vienetinės šaknies tikrinimas, esant struktūriniam pasikeitimui.....	16
3. Testai, skirti epideminiam struktūriniam pasikeitimams tikrinti.....	18
3.1. CUSUM statistika	18
3.2. T_n^{\max} statistika.....	20
3.3. T_n^r statistika	30
3.4. T_n^{sum} statistika	30
4. Modeliavimo rezultatai	39
4.1. CUSUM statistika	39
4.2. T_n^{\max} statistika.....	41
4.3. T_n^r statistika	43
4.4. T_n^{sum} statistika	44
5. Testų palyginimas	46
5.1. p-reikšmių grafikai.....	46
5.2. Galios grafikai.....	49
Išvados	56
Literatūros sąrašas.....	57
PRIEDAI.....	59
1 Priedas: Vienetinės šaknies tikrinimas Lietuvos duomenims.....	59
2 Priedas: Modeliuojamų eilučių paveikslai	60
3 Priedas: Kiekvieno testo atskiri p-reikšmių ir p-reikšmių skirtumų grafikai.....	62

Įvadas

Ekonomiką kartais ištinka įvairios krizės, pasitaiko stichinių nelaimių ar kitokių reiškinių, kurie iš esmės paveikia ekonomikos elgesį. Šis poveikis labai gerai atsispindi ekonominių rodiklių laiko eilutėse, kurių elgesys po tam tikros krizės ar stichinės nelaimės gali pakisti kardinaliai. Toks pasikeitimas paprastai vadinamas struktūriniu pasikeitimu. Esant struktūriniams pasikeitimams, modeliai su pastoviais koeficientais nėra tinkami tiek prognozavimo tikslams, tiek nagrinėjant politikos pasikeitimo poveikį kitiems ekonominiams kintamiesiems, todėl sudarant modelius derėtų į juos atsižvelgti. Tačiau dažnai pasitaiko atvejų, kai plika akimi sunku įžiūrėti esamus struktūrinius pasikeitimus, todėl šią problemą būtų galima išspręsti taikant specialius testus struktūriniams pasikeitimams aptikti.

Šiame darbe nagrinėjami trys žinomi testai struktūriniams pasikeitimams nustatyti, iš kurių vienas labiau pritaikytas vieno struktūrinio pasikeitimo atvejui, o kiti du – epideminiam pasikeitimams nustatyti. Mūsų tikslas palyginti šių testų galią ir pasiūlyti naują testą, kuris pasižymėtų patrauklesnėmis charakteristikomis. Matysime, kad testo, nepritaikyto epideminiam struktūriniams pasikeitimams nustatyti, galia labai maža, tuo tarpu likusieji testai gana gerai aptinka įvairaus tipo struktūrinius pasikeitimus.

Darbo tikslas suponuoja atitinkamą darbo struktūrą. Pirmoje dalyje supažindinama su struktūrinio pasikeitimo sąvoka ir jos vaidmeniu ekonomikos kontekste. Taip pat pateikiami epideminio pasikeitimo pavyzdžiai. Antroje dalyje nagrinėjami klasikiniai testai struktūriniams pasikeitimams aptikti. Trečioje dalyje apibrėžiami darbe lyginami testai, išvedama nauja statistika, suformuluojamos reikalingos teoremos bei pateikiami jų įrodymai. Ketvirtoje dalyje pateikti modeliavimo rezultatai. Penktoje – atliekamas testų palyginimas. Darbo pabaigoje pateikiamos išvados.

Rašant darbą panaudota 18 šaltinių.

1. Struktūriniai pasikeitimai

1.1. Struktūrinio pasikeitimo sąvoka

Atliekant ekonominius tyrimus dažnai į pagalbą pasitelkiami ekonometriniai metodai, kuriuos taikant gaunamos aktualių kintamųjų prognozės, analizuojama jų raida, įgyvendinamos įvairios strategijos. Vienas iš pagrindinių ekonometrijos instrumentų – tiesinės regresijos modelis – yra plačiausiai taikomas tiriant ekonominių kintamųjų kitimą laike. Tačiau tam, kad sudarius tiesinę regresijos lygtį ir jos parametrus įvertinus mažiausių kvadratų metodu gautume pagrįstus rezultatus, turimi duomenys privalo tenkinti tam tikras prielaidas.

Viena iš pagrindinių prielaidų, kurią privalo tenkinti modeliuojamas kintamasis, yra stacionarumas¹. Tačiau daugeliui iš jų ši prielaida nėra tenkinama. Empiriniai duomenys rodo, kad daugelis ekonominių rodiklių yra integruoti dydžiai². Šią problemą galima išspręsti diferencijuojant duomenis arba modeliavimui taikant specialius metodus, kaip, pavyzdžiui, kointegravimo teorija³. Tačiau dažnai net ir diferencijuoti duomenys nėra tinkami tiesinės regresijos modeliams. Tai lemia struktūrinių pasikeitimų atsiradimas, kuriuos sąlygoja netikėti įvykiai ar reiškiniai.

Jeigu nagrinėtume pirmos eilės autoregresinį procesą:

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t, E\varepsilon_t = \sigma^2,$$

(čia α ir ρ yra nežinomi modelio parametrai, o ε – modelio paklaida), tada modelio stabilumas reikštų, kad parametrai α , ρ , ir σ^2 yra pastovūs. Jeigu bent vienas iš šių parametrų kuriuo nors laiko momentu pasikeičia, tada sakome, kad įvyko struktūrinis pasikeitimas (7, 2001).

Ekonomiką kartais ištinka įvairios krizės, pasitaiko stichinių nelaimių ar kitokių reiškinių, kurie iš esmės paveikia ekonomikos elgesį ir tokiu būdu sąlygoja struktūrinius pasikeitimus (duomenų eilutės lūžius). Jeigu atidžiai patyrinėtume, pavyzdžiui, Lietuvos ekonominių rodiklių duomenis, pamatytume,

¹ Procesas $\{X_t, t > 0\}$ yra vadinamas stacionariu siaurąja prasme, jeigu jo pasiskirstymas nepriklauso nuo postūmio laike, t.y. $F_{t_1, t_2, \dots, t_n}(x) = F_{t_1+h, t_2+h, \dots, t_n+h}(x), \forall h > 0, t_1, t_2, \dots, t_n > 0$.

Procesas $\{X_t, t > 0\}$ yra vadinamas stacionariu plačiąja prasme, jeigu $EX_t^2 < \infty$ ir jo vidurkis ir kovariacija nepriklauso nuo postūmio laike: $EX_t = const, \forall t, r(t, s) = r(t-s, 0), \forall t, s$.

² Sakoma, kad procesas $\{X_t\}$ yra integruotas d eile, žym. $X_t \sim I(d)$, jeigu jo d eilės skirtumas yra stacionarus procesas, $X_t \sim I(0)$.

³ Kointegravimo ir paklaidų korekcijos modelio teorija išsamiai išnagrinėta (9, 2002).

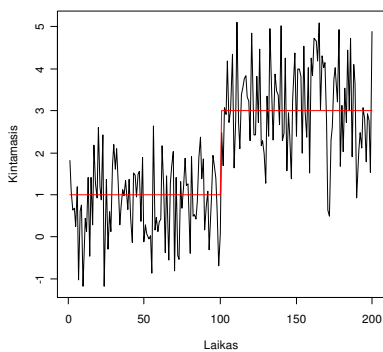
kad tam tikrų rodiklių laiko eilutės pasižymi staigiu, neįprastu pasikeitimu, t.y. kai kuriuose Lietuvos duomenyse pastebimas struktūrinis pasikeitimas Rusijos krizės metu ar įstojus į Europos Sąjungą. Todėl tam, kad modeliavimo rezultatai būtų patikimi, būtina aptikti ir atsižvelgti į galimus struktūrinius pasikeitimus.

1.2. Struktūrinių pasikeitimų tipai

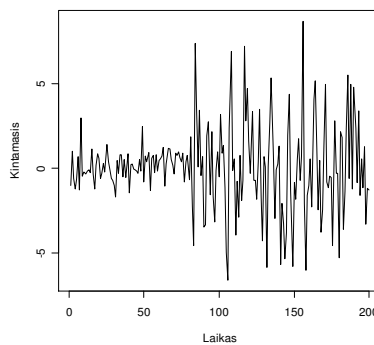
Taikydami laiko eilučių teoriją, vertindami lygtis, prognozuodami mes darome prielaidą, kad modelio parametrai yra pastovūs, t.y. mes laikome, kad tokie parametrai kaip vidurkis, dispersija ar trendas yra pastovūs, tačiau buvo pastebėta, kad, esant tokiems įvykiams kaip didžioji krizė, naftos kainų šokai, staigūs politikos pasikeitimai ir t.t., modeliai su pastoviais koeficientais nėra tinkami tiek prognozavimo tikslams, tiek nagrinėjant politikos pasikeitimo poveikį kitiems ekonominiams kintamiesiems (9, 2002; p.390).

Struktūrinis pasikeitimas gali pasireikšti įvairiai: jis gali paveikti vidurkį, dispersiją, trendą. Taip pat galimi ir mišrūs pasikeitimai, pavyzdžiui, kai pasikeičia ir vidurkis, ir dispersija. Toliau pateikti galimų struktūrinių pasikeitimų pavyzdžiai:

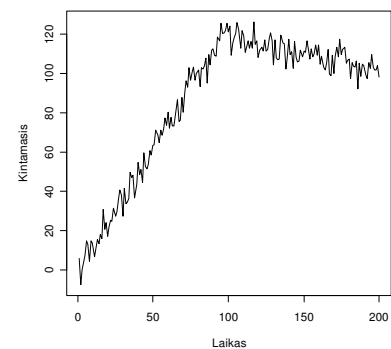
1 paveikslas: Struktūrinių pasikeitimų tipai
Pasikeičia vidurkis



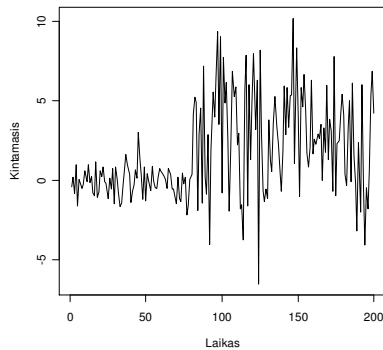
Pasikeičia dispersija



Pasikeičia trendas



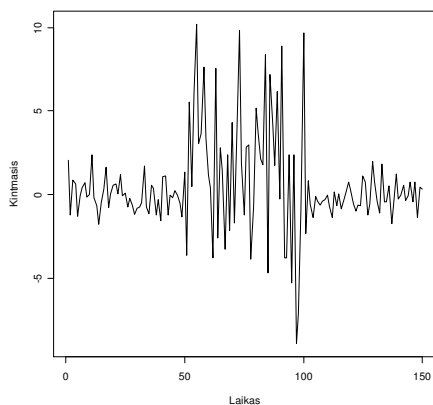
Pasikeičia vidurkis ir dispersija



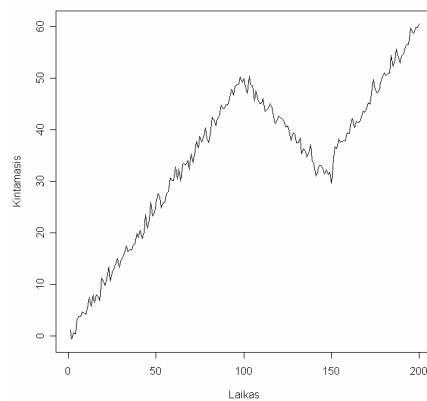
1 paveiksle pateikti pavyzdžiai, kai turimas vienas struktūrinis pasikeitimas, t.y. šiuo atveju laikoma, kad per visą nagrinėjamą periodą ekonomiką ištiko tik viena krizė ar koks kitas netikėtas įvykis, tačiau iš tikrųjų, kai nagrinėjamas ilgesnis laiko periodas, tokių krizių ar netikėtų reiškinių gali būti ne vienas ir ne du, todėl tokiu atveju jau turėtume kelis struktūrinius pasikeitimus.

Dažnai pasitaiko situacijų, kai atsitikus tam tikram įvykiui, nagrinėjamo kintamojo raida pasikeičia gana akivaizdžiai, tačiau po kurio laiko ji vėl grįžta į savo pradinę situaciją, t.y. tarkime, kad pradžioje turime tam tikrą ekonominį rodiklį, kurio elgesys yra daugiau mažiau stabilus (pavyzdžiui, imkime tam tikros šalies bendrąjį vidaus produktą, kuris auga pastoviu tempu). Tam tikru laiko momentu šalį ištinka krizė, kuri sudrumsčia įprastą rodiklio elgseną (pavyzdžiui, bendrojo vidaus produkto augimas sulėtėja), tačiau praėjus tam tikram periodui, rodiklis vėl grįžta į pradinę būseną (ilgainiui bendrojo vidaus produkto augimo tempai vėl išauga iki prieš krizę buvusio lygio). Toks pasikeitimas, kai galiausiai grįžtama į pradinę būseną, vadinamas epideminiu pasikeitimu. 2 paveiksle pateikiami epideminių pasikeitimų vidurkyje ir trende pavyzdžiai.

2 paveikslas: Epideminių pasikeitimų pavyzdžiai Dispersijos epideminis pasikeitimas



Trendo epideminis pasikeitimas



Iš pirmojo paveikslo matome, kad pradžioje duomenų išsibarstymas buvo nedidelis. Laiko momentu $t = 50$ įvyko struktūrinis pasikeitimas, t.y. duomenų išsibarstymas gerokai padidėjo, tačiau vėliau jis vėl grįžo į pradinę padėtį. Tuo tarpu antrame paveiksle matoma analogiška situacija, tik šiuo atveju pasikeitė ne duomenų išsibarstymas, bet trendo kryptis.

Taigi, sudarant lygtis bei darant išvadas svarbu atsižvelgti į duomenyse galimus struktūrinius pasikeitimus.

2. Testai, skirti struktūriniais pasikeitimams tikrinti

Šiuo metu yra daugybė statistinių testų, skirtų struktūriniais pasikeitimams tikrinti. Jie klasifikuojami į tokias kategorijas (9, 2002; p.390):

- (i) Žinomi pasikeitimo taškai – nežinomi pasikeitimo taškai;
- (ii) Vienas struktūrinis pasikeitimas – keli struktūriniai pasikeitimai;
- (iii) Vienamačiai ir daugiamačiai sąryšiai;
- (iv) Stacionarūs ir nestacionarūs kintamieji.

Vieni gali užduoti sau klausimą, ar duomenyse yra vienas ar du struktūriniai pasikeitimai, tačiau problemų iškyla ne tikrinant hipotezę, jog duomenyse yra vienas, o ne du struktūriniai pasikeitimai, bet kaip apskritai įvertinti pasikeitimų skaičių. Tai yra modelio parinkimo problema. Šį klausimą nagrinėjo daugelis autorių. Vieni jį nagrinėjo rutuliodami kelių struktūrinių pokyčių egzistavimą, kiti – vieno ir pan. (9, 2002).

Taip pat svarbu žinoti ne tik pokyčių egzistavimą, bet taip pat ir jų įvykimo momentą. Daugelis testų tepateikia informaciją apie tai, ar duomenyse yra struktūrinis pasikeitimas, tačiau nenurodo, kuriuo laiko momentu tas pokytis įvyko. Pradžioje buvo sudaromi testai laikant, kad pasikeitimo momentai yra žinomi, tačiau tokia situacija pasitaiko gana retai, t.y. paprastai mes tikslaus pasikeitimo momento nežinome. Tokiu būdu atsirado poreikis sukurti tokius testus, kurie tikrintų struktūrinio pasikeitimo egzistavimą, kai tikslus lūžio momentas nėra žinomas.

2.1. Chow testas

Vienas iš klasikinių testų – Chow testas. Tačiau atliekant šį testą būtina tiksliai žinoti pasikeitimo momentą. Chow testas veikia gana paprastai: imtis padalinama į dvi dalis ir įvertinami abiejų poabių parametrai. Jei pastarieji sutampa, tada sakoma, kad struktūrinis pasikeitimas neįvyko. Parametrų rinkiniai tikrinami naudojant klasikinę F statistiką. Pastaroji apibrėžiama tokiu būdu:

$$F = \frac{(SS_2 - SS_1)/m}{SS_1/(n-k)},$$

čia SS_1 yra liekanų kvadratų suma, gauta įvertinus lygtį su n stebėjimų, SS_2 – liekanų kvadratų suma, gauta įvertinus lygtį su $n + m$ stebėjimais. Šiuo atveju n yra pradinis stebėjimų skaičius, m – papildomi stebėjimai, o k – nepriklausomų kintamųjų skaičius. Kai teisinga nulinė hipotezė, jog abu stebėjimų rinkiniai generuojami to paties regresijos modelio su normaliai pasiskirsčiusiomis liekanomis bei tenkinamos klasikinės prielaidos, tada F statistika turės Fišerio pasiskirstymą su $(m, n-k)$ laisvės laipsniais (8, 1976).

Chow testas buvo išplėstas ir atvejams, kai imtis dalinama į daugiau nei dvi dalis, ir buvo nagrinėjami tik tie atvejai, kai režimų dydis yra mažesnis už k . Šiuo atveju tikrinama, ar yra keli struktūriniai pasikeitimai (9, 2002; p.390-391).

Tarkime, kad turime J grupių su n_j ($j=1,2,\dots,J$) stebėjimų kiekvienoje iš jų. Tegu RSS_0 žymi liekanų kvadratų sumą, kuri gauta iš regresijos su jungtiniais duomenimis, o RSS_T yra visų liekanų kvadratų suma iš rinkinių, kuriems galioja $n_j > k$. Tegu yra r tokių rinkinių ($r \leq J$). Tegu $n = \sum n_j, n_1 = \sum_r \sum n_j$ pagal r rinkinius, kuriems $n_j > k$. Tada $RSS_0 / \sigma^2 \sim \chi^2$ su $(n-k)$ l.l. ir $RSS_T / \sigma^2 \sim \chi^2$ su $(n_1 - rk)$ l.l. ir

$$\frac{(RSS_0 - RSS_T)/(n - k - n_1 + rk)}{RSS_T / (n_1 - rk)}$$

turi F pasiskirstymą su $(n - k - n_1 + rk)$, $(n_1 - rk)$ laisvės laipsniais. Remiantis šia statistika galima tikrinti hipotezę, ar režimų parametru reikšmės skiriasi.

Chow testą galima išplėsti nagrinėjant rekursines liekanas (8, 1976): nagrinėkime regresinį modelį:

$$y_j = x_j' \beta + u_j, \quad j = 1, \dots, n,$$

čia x_j yra $k \times 1$ nepriklausomų stebėjimų vektorius, β - $k \times 1$ koeficientų vektorius, y_j - priklausomas kintamasis, o u_j - triukšmas. Laikoma, kad liekanos yra normaliosios ir vienodai pasiskirsčiusios su nuliniu vidurkiu ir pastovia dispersija σ^2 .

$n - k$ rekursinių liekanų yra apibrėžiamos tokiu būdu:

$$\tilde{u}_j = \frac{y_j - x_j' b_{j-1}}{\left(1 + x_j' (X_{j-1}' X_{j-1})^{-1} x_j\right)^{1/2}}, \quad j = k + 1, \dots, n,$$

čia b_j yra koeficiento β mažiausių kvadratų įvertis, gautas naudojant pirmuosius j stebėjimų, o X_j yra $j \times k$ pilno rango matrica, sudaryta iš pirmųjų j nepriklausomų kintamųjų stebėjimų.

Taip pat galima apibrėžti tokį rekursinį sąryšį:

$$S_j = S_{j-1} + \tilde{u}_j^2,$$

čia S_j yra liekanų kvadratų suma, gauta įvertinus regresiją, naudojant j pirmuosius stebėjimus. Iš čia seka, kad⁴

$$S_n = \sum_{j=k+1}^n \tilde{u}_j^2.$$

Tokiu būdu Chow testinę statistiką galima perrašyti tokiu būdu:

$$F = \frac{\sum_{j=n+1}^{n+m} \tilde{u}_j^2 / m}{\sum_{j=k+1}^n \tilde{u}_j^2 / (n-k)}.$$

Ši statistika prie nulinės hipotezės taip pat turės Fišerio skirstinį (8, 1976).

Ankstesniais laikais Chow testas buvo labai paplitęs, tačiau jo trūkumas yra tas, kad pasikeitimo momentai turi būti žinomi iš anksto. Šiuo atveju modeliuotojas turi du pasirinkimus: arba tiesiog pasirinkti tam tikrą lūžio tašką, arba pastarąjį nustatyti remiantis tam tikra papildoma informacija. Pirmuoju atveju Chow testas gali būti neinformatyvus, jeigu lūžio taškas buvo pasirinktas klaidingai.

⁴ Pastarieji rezultatai yra įrodyti R.L.Brown, J.Durbin, J.M.Evans straipsnyje (4, 1975).

Antruoju atveju Chow testas gali būti apgaulingas, kadangi pasirinktas lūžio taškas yra endogeninis – jis koreliuoja su duomenimis – todėl tikėtina, kad testas klaidingai įvertins lūžį, kai pastarasis galbūt išvis neegzistuoja. Be to, kadangi rezultatai labai priklauso nuo pradinio pasirinkimo, todėl skirtingi modeliuotojai gali gana nesunkiai gauti skirtingas išvadas. (7, 2001; p.118).

Matėme, kad klasikinis Chow testas nėra labai tinkamas struktūriniam pasikeitimui įvertinti, kai tiksliai nežinome pasikeitimo momento. Todėl tolesni darbai buvo plėtojami ieškant metodo struktūriniam pasikeitimams įvertinti, kai lūžio momentas nėra žinomas.

Quandt pasiūlė išplėsti Chow testą atvejams, kai pasikeitimo taško nežinome, t.y. šiuo atveju jis siūlo įvertinti daug lygčių kaskart imant vis kitą lūžio tašką ir iš visų apskaičiuotų statistikų išrinkti didžiausią (7, 2001). Tačiau šiuo atveju buvo nežinomas ribinis pasiskirstymas. Quandt parodė, kad tai tikrai nėra chi-kvadrato pasiskirstymas ir pateikė baigtinės imties kritines reikšmes konkreiems atvejams (10, 2005; p.4).

2.2. Fišerio statistika

Vienas iš dažnai naudojamų būdų struktūriniam pasikeitimams tikrinti, kai lūžio momentas yra nežinomas ir kai galimi keli struktūriniai pasikeitimai, remiasi klasikiniu tiesinės regresijos modeliu ir Fišerio statistika (17, 2003).

A. Zeileis, C. Kleiben ir kt. 2003 metų straipsnyje nagrinėjamas tiesinis regresijos modelis:

$$y_i = x_i^T \beta_i + u_i \quad (i = 1, \dots, n)$$

čia y_i yra priklausomasis kintamasis, x_i - $k \times 1$ regresorių vektorius, kurio pirmoji komponentė paprastai lygi vienetui, o β_i - $k \times 1$ regresijos koeficientų vektorius, kuris ilgainiui gali kisti.

Šiame straipsnyje tikrinama hipotezė, kad regresijos koeficientai nekinta:

$$H_0 : \beta_i = \beta_0 \quad (i = 1, \dots, n)$$

prieš alternatyvą, kad bent vienas koeficientas ilgainiui pakis. Daugelyje taikymų prasminga tarti, kad esama m struktūrinių pasikeitimų, kai koeficientai keičiasi iš vieno pastovaus regresijos sąryšio į kitą.

Taigi, yra $m+1$ segmentai, kuriuose regresijos koeficientai yra pastovūs. Tokiu būdu pradžioje apibrėžtą modelį galima perrašyti taip:

$$y_i = x_i^T \beta_j + u_i \quad (i = i_{j-1} + 1, \dots, i_j, \quad j = 1, \dots, m+1)$$

čia j yra segmento indeksas. Pagal susitarimą laikoma, kad $i_0 = 0$ ir $i_{m+1} = n$.

Praktikoje paprastai lūžio taškai nėra duoti egzogeniškai, jie yra nežinomi, todėl juos reikia įvertinti iš duomenų. Struktūriniai pasikeitimai gali būti tikrinami taikant F statistikas, kurios sudarytos specialioms alternatyvoms. Taikant F statistikas yra tikrinama alternatyva, kad yra tik vienas struktūrinis pasikeitimas, kurio pasikeitimo momentas nėra žinomas. Paprastai testai, tikrinantys tokias alternatyvas, remiasi F statistikų, apskaičiuotų kiekvienam lūžio taškui, seka: MKM liekanos $\hat{u}(i)$ iš kiekvieno segmento regresijos yra palyginamos su liekanomis \hat{u} iš visos imties modelio:

$$F_i = \frac{\hat{u}^T \hat{u} - \hat{u}(i)^T \hat{u}(i)}{\hat{u}(i)^T \hat{u}(i) / (n - 2k)}$$

Vėliau šios statistikos yra apskaičiuojamos kiekvienam $i = n_h, \dots, n - n_h$ ($n_h \geq k$) ir H_0 atmetamos, jeigu jų supremumai yra pernelyg dideli.

Struktūrinių pasikeitimų laiko momentais imami tie taškai, kurie minimizuoja įvertintos regresijos liekanų kvadratų sumą⁵.

2.3. Bajeso statistika

Kitas alternatyvus būdas struktūriniam pasikeitimams įvertinti – naudoti iš prigimties Bajeso statistiką, kuri, esant teisingai alternatyvai, priskiria svorius p_t kaip išankstines tikimybes, kad pasikeitimas įvyks laiko momentu t ($t = 1, \dots, T$). Jeigu tarsime, kad paklaidos yra normaliosios ir σ_e^2 reikšmė yra nežinoma, tada turėsime tokią statistiką:

$$Q = \hat{\sigma}_e^{-2} T^{-1} \sum_{t=1}^T p_t \left[\sum_{j=t+1}^T (y_j - \bar{y}) \right]^2,$$

⁵ Šį struktūrinių pasikeitimų radimo metodą taip pat pritaikė S. Basci, E. Basci ir A. Zaman (1).

čia $\bar{y} = T^{-1} \sum_{t=1}^T y_t$ yra imties vidurkis, o $\hat{\sigma}_e^2 = T^{-1} \sum_{t=1}^T (y_t - \bar{y})^2$ - imties dispersija (10, 2005; p.4-5).

Šios statistikos ribinis pasiskirstymas yra toks:

$$Q \Rightarrow \int_0^1 B_0(r)^2 dr,$$

čia $B_0(r) = W(r) - rW(1)$ yra Brauno tiltas.

Vėliau ši procedūra buvo išplėsta tikrinti, ar yra struktūrinis pasikeitimas trendo funkcijoje, kuri apibrėžiama tokiu būdu:

$$y_t = \sum_{i=0}^p \beta_{i,t} t^i + e_t,$$

čia $\beta_{i,t} = \beta_i + \delta_i \mathbf{I}(t > T_b)$.

Tuomet testinė statistika prie nulinės hipotezės, kad nėra struktūrinio pasikeitimo, apibrėžiama taip:

$$Q_p = \hat{\sigma}_e^{-2} T^{-2} \sum_{t=1}^T \left[\sum_{j=t+1}^T \hat{e}_j \right]^2,$$

čia $\hat{\sigma}_e^2 = T^{-1} \sum_{t=1}^T \hat{e}_t^2$. Ribinis pasiskirstymas yra:

$$Q \Rightarrow \int_0^1 B_p(r)^2 dr,$$

čia $B_p(r)$ - apibendrintas Brauno tiltas.

2.4. Vienetinės šaknies tikrinimas, esant struktūriniam pasikeitimui

Vienetinės šaknies tikrinimas yra labai paplitęs tiek ekonomikoje, tiek statistikoje. Dominuoja tvirtinimas, kad daugelis ekonominių kintamųjų yra stochastiškai, o ne deterministiškai nestacionarūs, tačiau buvo pastebėta, kad daugelis makroekonominių kintamųjų gali būti modeliuojami kaip stacionarūs kintamieji su lūžusiu (nepastoviu) trendu (11, 1989). Anot, Jiahui Wang ir Eric Zivot (15, 1999), dėl to ekonometristai ir kiti mokslininkai ėmė vis labiau tirti struktūrinio pasikeitimo egzistavimą laiko eilutės trendo funkcijoje.

Daugelis vėlesnių tyrinėjimų buvo paremti vienietinės šaknies tikrinimu, kai esamas vienas struktūrinis pasikeitimas, o pasikeitimo momentas gali būti žinomas arba nežinomas⁶. Empiriniai tyrinėjimai parodė, kad didžiajai daliai kintamųjų vienietinės šaknies hipotezė yra atmetama prieš alternatyvą, kad trendas turi vieną lūžio tašką. Tačiau apskritai galima tvirtinti, kad daugelis makroekonominių laiko eilučių gali būti deterministiškai modeliuojamos su mažiausiai vienu struktūriniu pasikeitimu. Iš tikrųjų, jeigu manoma, kad kintamasis turi lūžį, tai prielaida, kad galimas tik vienas lūžis, yra nereali ir ribota. Pavyzdžiui, trendo lūžius dažniausiai sąlygoja „dideli įvykiai“ tokie, kaip karai, naftos kainų šokai, finansinės krizės ar politinių režimų pasikeitimai, o daugelis kintamųjų apima kelis tokius įvykius (15, 1999).

Dan Ben-David ir David H. Papell straipsnyje (2, 1994) nagrinėjamas lūžusio trendo modelis ir tikrinama vienietinės šaknies hipotezė 16 valstybių. Nagrinėjami šešiolikos valstybių bendrojo vidaus produkto (BVP) duomenys ir siekiama išsiaiškinti, ar BVP eilutės turi vienietinę šaknį, ar vis dėlto tai yra stacionarūs kintamieji su lūžusiu trendu. Autoriai nagrinėja tokią lygtį:

$$\Delta y_t = \mu + \theta DU_t + \beta t + \gamma DT_t + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + \varepsilon_t,$$

$$\text{čia } DU_t = \begin{cases} 1, & t > T_B; \\ 0, & t \leq T_B. \end{cases} \quad DT_t = \begin{cases} t, & t > T_B; \\ 0, & t \leq T_B. \end{cases}$$

Modelis vertinamas imant $T_B = 2, \dots, T - 1$. Kiekvieną kartą įvertinus lygtį apskaičiuojama *ADF* t-statistika (pakoreguota Dickey-Fuller statistika). Laikoma, kad struktūrinis pasikeitimas įvyko tuo laiko momentu, kuriame statistika įgijo savo didžiausią reikšmę. Atlikus skaičiavimus autoriai gavo, kad didžiajai daliai nagrinėjamų šalių teko atmesti vienietinės šaknies hipotezę, t.y. daugumos šalių BVP eilutės buvo stacionarūs dydžiai su lūžusiu trendu⁷.

⁶ Vienietinės šaknies hipotezė, kai turimas vienas struktūrinis pasikeitimas, tikrino tokie autoriai, kaip Dan Ben-David ir David H. Papell (2, 1994), Pierre Perron (11, 1989).

⁷ Rezultatai, gauti pritaikius šį testą Lietuvos BVP duomenims, pateikti 1 priede.

3. Testai, skirti epideminiais struktūriniais pasikeitimams tikrinti

3.1. *CUSUM* statistika

Pačioje pradžioje buvo imta kurti testus, kurių paskirtis - tiesiog nustatyti, ar turimi duomenys turi struktūrinį lūžį, ar ne. Kaip buvo minėta, vienas iš pirmųjų testų buvo Chow testas, kuris skirtas nulinei hipotezei apie parametru pastovumą tikrinti prieš alternatyvą, kad yra struktūrinis pasikeitimas, kai lūžio taškas iš anksto žinomas bei daroma prielaida, kad dispersija yra pastovi. Taip pat buvo naudojama pasikeitimo regresija (angl. switching regression) tokia, kad buvo laikoma, jog stebėjimai yra generuojami dviejų skirtingų regresijų ir įvairūs kiti testai⁸. (9, 2002). Pagrindinė priežastis, kodėl pradžioje buvo tiriamas tik struktūrinio pasikeitimo egzistavimas, yra ta, kad nebuvo žinoma įvertintų lūžio taškų pasiskirstymo teorija ir pastaroji problema buvo įvertinta tik keliems atvejams (10, 2005, p.21). Daugelis testų, kurie jau buvo minėti, buvo formuojami naudojant dalines liekanų sumas, tačiau dabar apibrėšime statistiką, kuri sudaroma remiantis rekursinėmis liekanomis.

Viena iš statistikų, kuri vaidino svarbų vaidmenį struktūrinių pasikeitimų teorijoje, buvo *CUSUM* statistika⁹ (4, 1975). Šis testas paremtas rekursinių liekanų¹⁰ dalinių sumų maksimumu. Tiksliau, šią statistiką galime apibrėžti tokiu būdu: laikykime, kad turime tiesinę regresijos lygtį su k regresorių:

$$y_t = x_t' \beta + u_t$$

Tuomet *CUSUM* statistika bus apibrėžiama taip:

$$CUSUM = \max_{k+1 < r \leq T} \left| \frac{\sum_{t=k+1}^r \tilde{v}_t}{\hat{\sigma} \sqrt{T-k}} \right| / \left(1 + 2 \frac{r-k}{T-k} \right),$$

čia $\hat{\sigma}^2$ yra suderintas u_t dispersijos įvertis, o \tilde{v}_t yra rekursinės liekanos, kurių pavidalas:

⁸ Plačiau apie šį testą ir kitus, kurie skirti struktūrinių pasikeitimų egzistavimui patikrinti žr. (9, 2002).

⁹ Daugiau informacijos apie šią statistiką žr. (9, 2002), (17), (4, 1975).

¹⁰ Rekursinės liekanos buvo panaudotos ir Chow testo patobulinimui (žr. 2.1 dalis).

$$\tilde{v}_t = (y_t - x_t' \hat{\beta}_{t-1}) / f_t$$

$$f_t = (1 + x_t' (X_{t-1}' X_{t-1})^{-1} x_t)^{1/2},$$

čia X_{t-1} apima regresorių reikšmes iki momento $t-1$, o $\hat{\beta}_{t-1}$ yra parametro β mažiausių kvadratų įvertis, gautas naudojant duomenis iki momento $t-1$.

Ribinis *CUSUM* statistikos pasiskirstymas gali būti išreikštas kaip pasverto Wiener'io proceso maksimumas, t.y.

$$CUSUM \Rightarrow \sup_{0 \leq r \leq 1} \frac{|W(r)|}{1 + 2r},$$

čia $W(r)$ yra Wiener'io procesas, apibrėžtas intervale $(0,1)$. (14, 1982)

CUSUM statistika buvo taikoma nagrinėjant įvairaus tipo duomenis. R.L.Brown, J.Durbin ir J.M.Evans savo straipsnyje (4, 1975, p.157-161) pateikia tris taikymo pavyzdžius. Pirmame jų nagrinėja vietinių telefono skambučių skaičiaus augimą, kuris aprašomas tiesiniu modeliu, įtraukiant laisvąjį narį ir keturis paaiškinamuosius kintamuosius. Pritaikius *CUSUM* ir kvadratinę *CUSUM*¹¹ testą šiems duomenims buvo aiškiai pastebimas duomenų eilutės nestabilumas 1964-1965 metais¹². Antrajame pavyzdyje buvo nagrinėjamas pinigų paklausos kintamasis. Šiuo atveju atlikus testus struktūrinis pasikeitimas nebuvo aptiktas. Galiausiai, trečiame pavyzdyje tiriant organizacijos keliamus reikalavimus personalui, kurie išreikšti kaip devynių skirtingų kategorijų darbo krūvio funkcija, *CUSUM* testai rodė aiškų eilutės nestabilumą.

J.Schroeter ir A.Fenn (13, 2005) *CUSUM* testą pritaikė paneliniams duomenims. Jie nagrinėjo cigarečių paklausos stabilumą. Atlikus *CUSUM* testą buvo pastebėtas struktūrinis pasikeitimas.

Ši statistika paprastai naudojama vieno struktūrinio pasikeitimo egzistavimui patikrinti, tačiau pastarąją pritaikysime epideminiam struktūriniam pasikeitimui nustatyti ir gautus rezultatus palyginsime su tais, kuriuos gavome taikydami kitus metodus, skirtus epideminiam struktūriniam pasikeitimams aptikti.

Taikydami apibrėžtas statistikas, tikrinsime nulinę hipotezę, jog duomenyse yra epideminis struktūrinis pasikeitimas, prieš alternatyvą, kad struktūrinio pasikeitimo nėra. Naudodami ribinį statistikos pasiskirstymą, apskaičiuosime teorinius kvantilius, kuriuos naudosime tikrinant hipotezę. Jeigu apskaičiuota *CUSUM* statistika yra didesnė už kvantilio reikšmę t_α (šiuo atveju α žymi

¹¹ Apie kvadratinę *CUSUM* testą žr: (9, 2002), (17), (4, 1975).

¹² Nestabilumą parodė kvadratinis *CUSUM* testas.

reikšmingumo lygmenį), tada nulinę hipotezę atmesime ir tarsime, kad duomenyse yra struktūrinis pasikeitimas, priešingu atveju tarsime, kad struktūrinio pasikeitimo nėra.

3.2. T_n^{\max} statistika

Vieni testai buvo skirti tikrinti, ar duomenyse yra vienas ar keli struktūriniai pasikeitimai, tačiau kiti buvo labiau pritaikyti epideminių struktūrinių pasikeitimų atvejams¹³. Vienas iš jų yra A.Račkausko straipsnyje (12, 2006) apibrėžtas metodas, kuris pritaikytas epideminiams struktūriniams pasikeitimams regresiniuose modeliuose tikrinti.

Šis testas yra konstruojamas remiantis regresijos liekanų dalinėmis sumomis $\hat{S}(k), k = 0, \dots, n$. Tiksliau šis metodas apibrėžiamas taip:

Tegu $f : [0,1] \rightarrow \mathfrak{R}^d$ - duota tolydžiai diferencijuojama funkcija, $f(t) = (f_1(t), \dots, f_d(t))^{\tau}, t \in [0,1]$,

$$\mathbf{X} = \left(f\left(\frac{1}{n}\right), f\left(\frac{2}{n}\right), \dots, f\left(\frac{n}{n}\right) \right)^{\tau}.$$

Laikoma, kad galioja tokios reguliarumo sąlygos:

Matricos $A_n = n^{-1} \mathbf{X}^{\tau} \mathbf{X}$ ir $A = \int_0^1 f(t) f^{\tau}(t) dt$ yra neišsigimusios.

Nagrinėjamas toks modelis:

$$y_j = f^{\tau}(j/n) \beta + \varepsilon_j, \quad 1 \leq j \leq n, \quad \varepsilon_1, \dots, \varepsilon_n - n.v.p. (0, \sigma^2).$$

Regresijos koeficientai vertinami mažiausių kvadratų metodu:

$$\hat{\beta} = (\mathbf{X}^{\tau} \mathbf{X})^{-1} \mathbf{X}^{\tau} y, \quad y = (y_1, \dots, y_n)^{\tau}.$$

Liekanos apibrėžiamos tokiu būdu:

$$\hat{\varepsilon} = (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)^{\tau}, \quad \hat{\varepsilon}_k = y_k - \hat{y}_k = y_k - f^{\tau}(k/n) \hat{\beta}, \quad k = 1, \dots, n.$$

Tegu $\hat{S}(0) = 0$ ir $\hat{S}(k) = \sum_{t \leq k} \hat{\varepsilon}_t, \quad k = 1, \dots, n$.

Nagrinėjamas toks procesas:

$$\hat{V}_n(t) = \hat{S}(nt) + \hat{\varepsilon}_{[nt]+1}(nt - [nt]), \quad t \in [0,1].$$

¹³ Epideminių struktūrinių pasikeitimų sąvoka apibrėžiama 1.2 dalyje.

Ribinis procesas V yra išreiškiamas tokia forma:

$$V(t) = W(t) - \int_0^t \left\{ \int_0^1 f^\tau(u) A^{-1} f(v) dW(v) \right\} du,$$

kur W yra standartinis Wiener'io procesas.

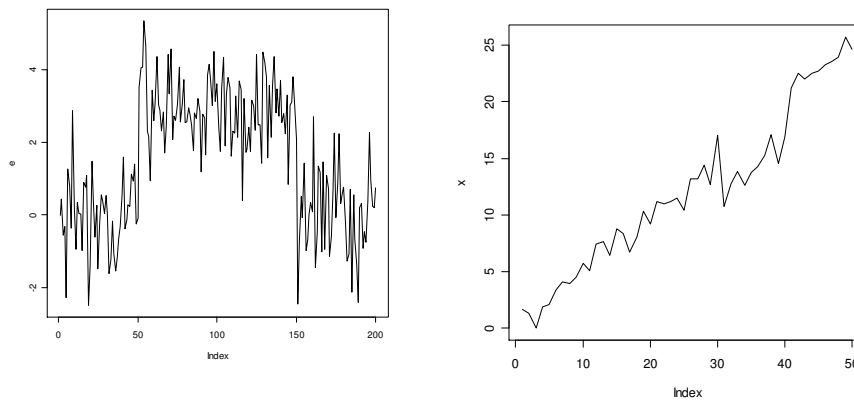
Turint \hat{V}_n ir V išraiškas, toliau galima tiksliai apibrėžti nagrinėjamą modelį, mus dominantį uždavinį ir statistiką, kurios pagalba išspręsimė suformuluotą uždavinį.

Tarkime, kad mūsų nagrinėjamas kintamasis yra nusakomas tokiomis lygtimis¹⁴:

$$y_i = \begin{cases} \beta \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & \text{jei } 1 \leq i \leq k^*, m^* < i \leq n; \\ \beta' \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & \text{jei } k^* < i \leq m^*. \end{cases}$$

čia $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai su nuliniu vidurkiu ir dispersija $\sigma^2 > 0$, bei baigtiniais momentais $E|\varepsilon_1|^q < \infty$, $q > 2$. β ir β' yra nežinomi modelio parametrai T.y. šiuo atveju nagrinėjame modelį su vienu epideminiu pasikeitimu, kurio ilgis lygus $l^* = m^* - k^*$. Paprastai nei pradžia k^* , nei ilgis l^* , nėra žinomi. Taigi, nagrinėjame kintamąjį, kuris periodo pradžioje buvo nusakomas pirmąja lygtimi su nežinomais parametrais β , vėliau, įvykus struktūriniam pasikeitimui, jo trajektorija pasikeitė. Nuo šio momento jis buvo nusakomas antrąja lygtimi su nežinomais parametrais β' . Tačiau galiausiai kintamasis vėl grįžo į savo pradinę trajektoriją. Žemiau pateikti tokio kintamojo pavyzdžiai¹⁵:

3 paveikslas: Kintamojo pavyzdžiai



¹⁴ Dėl paprastumo nagrinėsime vienamatį atvejį.

¹⁵ 3 paveikslo kairėje pateiktas vidurkio struktūrinis pasikeitimas, o dešinėje – trendo funkcijos pasikeitimas.

Mūsų uždavinys – patikrinti nulinę hipotezę:

$$H_0 : k^* > n$$

t.y. kad nėra struktūrinio pasikeitimo, prieš alternatyvą:

$$\bar{H} : 1 < k^* < m^* < n$$

kad yra epideminis pasikeitimas.

Kaip jau buvo minėta, regresijos koeficientai vertinami klasikiniu mažiausių kvadratų metodu. Įvertinus modelį, apskaičiuojamos regresijos liekanų dalinės sumos $\hat{S}_k, k = 0, 1, \dots, n$, kurios naudojamos testinei statistikai apibrėžti. Nagrinėjame tokią statistiką:

$$T_n^{\max} = \max_{1 < l < n} \frac{1}{l^\alpha} \max_{0 \leq k \leq n-l} \left| \hat{S}(k+l) - \hat{S}(k) - \frac{l}{n} \hat{S}(n) \right|, \quad 0 < \alpha < \frac{1}{2}$$

$$T^{\max} = \sup_{0 < h < 1} h^{-\alpha} \sup_{0 < t < 1-h} |V(t+h) - V(t) - hV(1)|, \quad 0 < \alpha < \frac{1}{2}.$$

Pirmoji yra empirinė statistika, kuri gaunama iš turimos imties. Pastaroji pagal A.Račkausko straipsnyje (12, 2006; p.202) įrodytą teoremą, jei teisinga nulinė hipotezė, pagal pasiskirstymą konverguoja į teorinę statistiką, apibrėžtą antrąja lygtimi.

Kaip ir *CUSUM* statistikos atveju, norėdami patikrinti hipotezę, pirmiausiai naudodami statistiką T^{\max} , apskaičiuojame teorinius kvantilius, kurių pagalba tikriname, ar nulinę hipotezę atmetame, ar ne: jeigu apskaičiuota statistika T_n^{\max} yra didesnė už kvantilio reikšmę t_α (šiuo atveju α žymi reikšmingumo lygmenį), tada nulinę hipotezę atmetame ir tariame, kad duomenyse yra struktūrinis pasikeitimas, priešingu atveju tarsime, kad struktūrinio pasikeitimo nėra.

Apibrėžus testinę statistiką, mus taip pat domina jos suderinamumas, kai teisinga alternatyva. Tuo tikslu, suformuluosime teoremą, kuri apibrėš sąlygas, kurioms galiojant, testas bus suderintas alternatyvos atžvilgiu. Taip pat pateiksime ir šios teoremos įrodymą.

Teorema (suderinamumas): *Jeigu tenkinamos sąlygos:*

1. $l^* \rightarrow \infty, \frac{l^*}{n} \rightarrow 0$, kai $n \rightarrow \infty$;
2. \exists tokie laiko momentai t_1 ir t_2 , kuriems teisinga: $\int_{t_1}^{t_2} f(t)dt \geq c \cdot (t_2 - t_1)$;
3. $\frac{l^{*1-\alpha}}{n^{1/2-\alpha}} \cdot |\beta' - \beta| \xrightarrow{n \rightarrow \infty} \infty$.

Tada nagrinėjiamam modeliui, kai teisinga alternatyva \bar{H} , galioja:

$$n^{-1/2+\alpha} \cdot T_n^{\max} \xrightarrow[n \rightarrow \infty]{p} \infty.$$

T.y. jeigu tenkinamos teoremos sąlygos, tada testas yra suderintas alternatyvos atžvilgiu.

Irodymas:

Statistiką T galima įvertinti tokiu būdu.

Pažymėkime $T_n^{\max} \equiv T(n; \alpha)$.

$$T(n; \alpha) = \max_{1 < l < n} \frac{1}{l^\alpha} \max_{0 \leq k \leq n-l} \left| \hat{S}(k+l) - \hat{S}(k) - \frac{l}{n} \hat{S}(n) \right| \geq \frac{1}{l^{*\alpha}} \cdot \left| \hat{S}(k^* + l^*) - \hat{S}(k^*) - \frac{l^*}{n} \hat{S}(n) \right| = |T^*(n; \alpha)|$$

Išsiveskime, kokia yra $T^*(n; \alpha)$ išraiška.

$$\begin{aligned} l^{*\alpha} \cdot T^*(n; \alpha) &= \hat{S}(k^* + l^*) - \hat{S}(k^*) - \frac{l^*}{n} \hat{S}(n) = \sum_{i \in I} \hat{\varepsilon}_i - \frac{l^*}{n} \cdot \sum_{i=1}^n \hat{\varepsilon}_i = \sum_{i \in I} \hat{\varepsilon}_i - \frac{l^*}{n} \cdot \sum_{i \in I} \hat{\varepsilon}_i - \frac{l^*}{n} \cdot \sum_{i \in I^c} \hat{\varepsilon}_i = \\ &= \left(1 - \frac{l^*}{n}\right) \cdot \sum_{i \in I} \hat{\varepsilon}_i - \frac{l^*}{n} \cdot \sum_{i \in I^c} \hat{\varepsilon}_i \end{aligned}$$

Čia $I = \{k^* + 1, \dots, k^* + l^*\}$, $I^c = \{1, \dots, k^*\} \cup \{(k^* + l^*) + 1, \dots, n\}$.

Nagrinėkime atvejį, kai yra struktūrinis pasikeitimas. Tada turėsime:

$$\hat{\varepsilon}_i = y_i - \hat{y}_i = \begin{cases} \beta' \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & i \in I \\ \beta \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & i \in I^c \end{cases} - \hat{\beta} \cdot f\left(\frac{i}{n}\right) = \begin{cases} (\beta' - \hat{\beta}) \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & i \in I \\ (\beta - \hat{\beta}) \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & i \in I^c \end{cases}$$

Iš čia gauname, kad:

$$\begin{aligned}
l^{*\alpha} \cdot T^*(n; \alpha) &= \left(1 - \frac{l^*}{n}\right) \cdot \sum_{i \in I} \left[(\beta' - \hat{\beta}) \cdot f\left(\frac{i}{n}\right) + \varepsilon_i \right] - \frac{l^*}{n} \cdot \sum_{i \in I^c} \left[(\beta - \hat{\beta}) \cdot f\left(\frac{i}{n}\right) + \varepsilon_i \right] = \\
&= \left(1 - \frac{l^*}{n}\right) \cdot \sum_{i \in I} \varepsilon_i - \frac{l^*}{n} \cdot \sum_{i \in I^c} \varepsilon_i + \left(1 - \frac{l^*}{n}\right) \cdot (\beta' - \hat{\beta}) \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot (\beta - \hat{\beta}) \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right)
\end{aligned}$$

Išveskime $\beta - \hat{\beta}$ ir $\beta' - \hat{\beta}$ išraiškas.

$$\begin{aligned}
\beta - \hat{\beta} &= \beta - \frac{\sum_{i=1}^n y_i \cdot f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} = \frac{\sum_{i=1}^n \left[\beta \cdot f^2\left(\frac{i}{n}\right) - y_i \cdot f\left(\frac{i}{n}\right) \right]}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} = \\
&= \frac{\sum_{i \in I} \left[\beta \cdot f^2\left(\frac{i}{n}\right) - \left(\beta' \cdot f\left(\frac{i}{n}\right) + \varepsilon_i \right) \cdot f\left(\frac{i}{n}\right) \right] + \sum_{i \in I^c} \left[\beta \cdot f^2\left(\frac{i}{n}\right) - \left(\beta \cdot f\left(\frac{i}{n}\right) + \varepsilon_i \right) \cdot f\left(\frac{i}{n}\right) \right]}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \\
&= \frac{\sum_{i \in I} (\beta - \beta') \cdot f^2\left(\frac{i}{n}\right) - \sum_{i=1}^n \varepsilon_i \cdot f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} = \frac{\sum_{i \in I} (\beta - \beta') \cdot f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} - V
\end{aligned}$$

$$\begin{aligned}
\beta' - \hat{\beta} &= \beta' - \frac{\sum_{i=1}^n y_i \cdot f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} = \frac{\sum_{i=1}^n \left[\beta' \cdot f^2\left(\frac{i}{n}\right) - y_i \cdot f\left(\frac{i}{n}\right) \right]}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} = \\
&= \frac{\sum_{i \in I^c} (\beta' - \beta) \cdot f^2\left(\frac{i}{n}\right) - \sum_{i=1}^n \varepsilon_i \cdot f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} = \frac{\sum_{i \in I^c} (\beta' - \beta) \cdot f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} - V
\end{aligned}$$

$$\text{Čia } V = \frac{\sum_{i=1}^n \varepsilon_i \cdot f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)}.$$

Taigi:

$$\begin{aligned}
l^{*\alpha} \cdot T^*(n; \alpha) &= \left(1 - \frac{l^*}{n}\right) \cdot \sum_{i \in I} \varepsilon_i - \frac{l^*}{n} \cdot \sum_{i \in I^c} \varepsilon_i + \left(1 - \frac{l^*}{n}\right) \cdot \left[\frac{(\beta' - \beta) \cdot \sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} - V \right] \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) - \\
&\quad - \frac{l^*}{n} \cdot \left[\frac{(\beta - \beta') \cdot \sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} - V \right] \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) = \sum_{i \in I} \varepsilon_i - \frac{l^*}{n} \cdot \sum_{i=1}^n \varepsilon_i - \left(1 - \frac{l^*}{n}\right) \cdot V \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \\
&\quad + \frac{l^*}{n} \cdot V \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) + \left(1 - \frac{l^*}{n}\right) \cdot \frac{(\beta' - \beta) \cdot \sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{(\beta - \beta') \cdot \sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) = \\
&= \sum_{i \in I} \varepsilon_i - \frac{l^*}{n} \cdot \sum_{i=1}^n \varepsilon_i - V \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot V \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right) + \\
&\quad + (\beta' - \beta) \cdot \left[\left(1 - \frac{l^*}{n}\right) \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) \right] = \\
&= \sum_{i \in I} \varepsilon_i - \frac{l^*}{n} \cdot \sum_{i=1}^n \varepsilon_i - \sum_{i=1}^n \varepsilon_i \cdot f\left(\frac{i}{n}\right) \cdot \left[\frac{\sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right] + \\
&\quad + (\beta' - \beta) \cdot \left[\left(1 - \frac{l^*}{n}\right) \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) \right]
\end{aligned}$$

Taigi, statistikos T išraišką galiausiai galima įvertinti tokiu būdu:

$$T(n; \alpha) = \max_{1 < l < n} \frac{1}{l^\alpha} \max_{0 \leq k \leq n-l} \left| \hat{S}(k+l) - \hat{S}(k) - \frac{l}{n} \hat{S}(n) \right| \geq \frac{1}{l^{*\alpha}} \cdot \left| \hat{S}(k^* + l^*) - \hat{S}(k^*) - \frac{l^*}{n} \hat{S}(n) \right| = |T^*(n; \alpha)| \geq$$

$$\begin{aligned}
&\geq \frac{1}{l^{*\alpha}} \cdot |\beta' - \beta| \cdot \left| \left(1 - \frac{l^*}{n}\right) \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) \right| + \frac{1}{l^{*\alpha}} \cdot \left| \sum_{i \in I} \varepsilon_i \right| - \\
&- \frac{1}{l^{*\alpha}} \cdot \frac{l^*}{n} \cdot \left| \sum_{i=1}^n \varepsilon_i \right| - \frac{1}{l^{*\alpha}} \cdot \left| \sum_{i=1}^n \varepsilon_i \cdot f\left(\frac{i}{n}\right) \right| \cdot \left| \frac{\sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| = \\
&= \frac{1}{l^{*\alpha}} \cdot |\beta' - \beta| \cdot |U| + \frac{1}{l^{*\alpha}} \cdot \left| \sum_{i \in I} \varepsilon_i \right| - \frac{1}{l^{*\alpha}} \cdot \frac{l^*}{n} \cdot \left| \sum_{i=1}^n \varepsilon_i \right| - \frac{1}{l^{*\alpha}} \cdot |H|
\end{aligned}$$

(*)

$$\text{Čia } U = \left(1 - \frac{l^*}{n}\right) \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right),$$

$$H = \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) \cdot \frac{\sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)}$$

Nagrinėkime pirmąjį narį. Parodysime, kad $\frac{|U|}{l^*}$ konverguoja į a , čia a yra konstanta.

$$\begin{aligned}
|U| &= \left| \left(1 - \frac{l^*}{n}\right) \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) \right| = \\
&= \left| \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) + \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) \right| \geq \\
&\geq \left| \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) \right| - \left| \frac{l^*}{n} \cdot \frac{\sum_{i \in I} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I^c} f\left(\frac{i}{n}\right) \right| - \left| \frac{l^*}{n} \cdot \frac{\sum_{i \in I^c} f^2\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) \right| \cong
\end{aligned}$$

$$\begin{aligned}
&\cong \left| n \cdot \frac{\int_0^{l^c} f^2(t) dt}{\int_0^1 f^2(t) dt} \cdot \int_I f(t) dt - \frac{l^* \cdot \int_0^1 f^2(t) dt}{\int_0^1 f^2(t) dt} \cdot \int_{I^c} f(t) dt - \frac{l^* \cdot \int_0^{l^c} f^2(t) dt}{\int_0^1 f^2(t) dt} \cdot \int_I f(t) dt \right| \geq \\
&\geq \left| n \cdot \frac{\int_0^{l^c} f^2(t) dt}{\int_0^1 f^2(t) dt} \cdot \int_I f(t) dt - \left(l^* \cdot \int_I |f^2(t)| dt \cdot \int_{I^c} |f(t)| dt \right) - \left(l^* \cdot \int_{I^c} |f^2(t)| dt \cdot \int_I |f(t)| dt \right) \right| \geq \\
&\geq \left| n \cdot c \cdot \frac{n-l^*}{n} \cdot \frac{l^*}{n} - \left(l^* \cdot M \cdot \frac{l^*}{n} \cdot M' \cdot \frac{n-l^*}{n} \right) - \left(l^* \cdot M'' \cdot \frac{n-l^*}{n} \cdot M''' \cdot \frac{l^*}{n} \right) \right|
\end{aligned}$$

(Komentaras: Antroji nelygybė išplaukia iš to, kad yra teisinga: $\left| \int_I f(x) dx \right| \leq \int_I |f(x)| dx$, trečioji nelygybė seka iš funkcijos f aprėžtumo ir iš teoremos pirmos sąlygos.

Taip pat nesunku parodyti, kad iš nelygybės $\int_{t_1}^{t_2} f(t) dt \geq c \cdot (t_2 - t_1)$ išplaukia nelygybė:

$\int_{t_1}^{t_2} f^2(t) dt \geq c' \cdot (t_2 - t_1)$. Parodykime tai:

$$\begin{aligned}
(t_2 - t_1) &\leq \int_{t_1}^{t_2} f(t) dt \leq \left(\int_{t_1}^{t_2} f^2(t) dt \right)^{1/2} \cdot \left(\int_{t_1}^{t_2} 1^2 dt \right)^{1/2} = \left(\int_{t_1}^{t_2} f^2(t) dt \right)^{1/2} \cdot (t_2 - t_1)^{1/2} \\
\Rightarrow \left(\int_{t_1}^{t_2} f^2(t) dt \right)^{1/2} &\geq (t_2 - t_1)^{1/2} \Rightarrow \int_{t_1}^{t_2} f^2(t) dt \geq (t_2 - t_1) \quad .)
\end{aligned}$$

Abi nelygybės puses padalinkime iš l^* :

$$\frac{|U|}{l^*} \geq \left| c \cdot \left(1 - \frac{l^*}{n} \right) - c' \cdot \frac{l^*}{n} \cdot \left(1 - \frac{l^*}{n} \right) - c'' \cdot \left(1 - \frac{l^*}{n} \right) \cdot \frac{l^*}{n} \right| \rightarrow a \quad \Rightarrow \quad \frac{|U|}{l^*} \xrightarrow{n \rightarrow \infty} a .$$

Taigi, turime, kad, kai $n \rightarrow \infty$, $\frac{|U|}{l^*} \rightarrow a$.

Tolesnis tikslas – parodyti, kad (*) išraiškos normuoti paskutiniai trys dėmenys konverguoja į nulį, kai $n \rightarrow \infty$.

Nagrinėkime dėmenį $\frac{1}{l^{*\alpha}} \cdot \left| \sum_{i \in I} \varepsilon_i \right|$. Normuokime jį: $\frac{n^\alpha}{\sqrt{n}} \cdot \frac{1}{l^{*\alpha}} \cdot \left| \sum_{i \in I} \varepsilon_i \right|$. Kadangi ε_i yra nepriklausomi

vienodai pasiskirstę dydžiai su vidurkiu 0 ir dispersija σ^2 , todėl pagal centrinę ribinę teoremą turėsime, kad

$$\frac{1}{\sqrt{l^*}} \cdot \sum_{i \in I} \varepsilon_i \xrightarrow[l^* \rightarrow \infty]{D} N(0, \sigma^2) \Rightarrow \frac{1}{\sqrt{l^*}} \cdot \sum_{i \in I} \varepsilon_i = O_p(1). \text{ Tokiu būdu gauname, kad}$$

$$\frac{n^\alpha}{\sqrt{n}} \cdot \frac{\sqrt{l^*}}{l^{*\alpha}} \cdot \left| \frac{1}{\sqrt{l^*}} \cdot \sum_{i \in I} \varepsilon_i \right| = \frac{l^{*1/2-\alpha}}{n^{1/2-\alpha}} \cdot O_p(1) = \left(\frac{l^*}{n} \right)^{1/2-\alpha} \cdot O_p(1) \xrightarrow[n \rightarrow \infty]{} 0.$$

Analogiškai, nagrinėdami dėmenį $\frac{1}{l^{*\alpha}} \cdot \frac{l^*}{n} \cdot \left| \sum_{i=1}^n \varepsilon_i \right|$, turėsime, kad iš centrinės ribinės teoremos

$$\frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \varepsilon_i \xrightarrow[n \rightarrow \infty]{D} N(0, \sigma^2) \Rightarrow \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \varepsilon_i = O_p(1). \text{ Todėl turėsime, kad}$$

$$\frac{n^\alpha}{\sqrt{n}} \cdot \frac{1}{l^{*\alpha}} \cdot \frac{l^*}{n} \cdot \left| \sum_{i=1}^n \varepsilon_i \right| = \frac{l^{*1-\alpha}}{n^{1-\alpha}} \cdot \left| \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \varepsilon_i \right| = \left(\frac{l^*}{n} \right)^{1-\alpha} \cdot O_p(1) \xrightarrow[n \rightarrow \infty]{} 0.$$

Galiausiai, nagrinėkime paskutinį narį: $\frac{1}{l^{*\alpha}} \cdot \left| \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) \cdot \frac{\sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right|$.

$$\left| \frac{\sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| \leq \left| \frac{\sum_{i \in I} f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| + \left| \frac{\frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| = \left| \frac{\sum_{i \in I} f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| + \left| \frac{\frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| =$$

$$= \frac{\left| \frac{1}{n} \cdot \sum_{i \in I} f\left(\frac{i}{n}\right) \right|}{\left| \frac{1}{n} \cdot \sum_{i=1}^n f^2\left(\frac{i}{n}\right) \right|} + \frac{\left| \frac{l^*}{n} \cdot \frac{1}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right) \right|}{\left| \frac{1}{n} \cdot \sum_{i=1}^n f^2\left(\frac{i}{n}\right) \right|} \cong \frac{\left| \frac{\int_0^{\frac{k^*+l^*}{n}} f(t) dt}{\frac{k^*+1}{n}} \right|}{\left| \frac{\int_0^1 f^2(t) dt}{1} \right|} + \frac{\left| \frac{l^*}{n} \cdot \int_0^1 f(t) dt \right|}{\left| \frac{\int_0^1 f^2(t) dt}{1} \right|} \leq \frac{\frac{k^*+l^*}{n}}{\frac{k^*+1}{n}} \int_0^1 |f(t)| dt + \frac{l^*}{n} \cdot \int_0^1 |f(t)| dt \leq$$

$$\leq M \int_0^{\frac{k^*+l^*}{n}} dt + M \cdot \frac{l^*}{n} \cdot \int_0^1 dt = M \cdot \left(\frac{l^*}{n} + \frac{l^*}{n} \right) = 2 \cdot M \cdot \frac{l^*}{n} \xrightarrow[n \rightarrow \infty]{} 0.$$

(Komentaras: Antroji nelygybė išplaukia iš to, kad yra teisinga: $\left| \int_I f(x) dx \right| \leq \int_I |f(x)| dx$, trečioji nelygybė – iš funkcijos f aprėžtumo.)

Beliko išsiaiškinti, kur konverguoja $\sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right)$. Reikia parodyti, kad

$\frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) = O_p(1)$. Tam pakanka įrodyti, kad $E\xi^2$ yra baigtinis.

Pažymėkime: $\xi = \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right)$.

$$\begin{aligned} E\xi^2 &= E \left[\frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) \right]^2 = \frac{1}{n} \cdot \sum_{i=1}^n f^2\left(\frac{i}{n}\right) \cdot E\varepsilon_i^2 = \frac{\sigma^2}{n} \cdot \sum_{i=1}^n f^2\left(\frac{i}{n}\right) \leq \\ &\leq \frac{\sigma^2}{n} \cdot n \cdot \sup_{1 \leq i \leq n} \left| f^2\left(\frac{i}{n}\right) \right| = \sigma^2 \cdot \sup_{1 \leq i \leq n} \left| f^2\left(\frac{i}{n}\right) \right| = \sigma^2 \cdot M^2 < \infty \end{aligned}$$

(Komentaras: Nelygybė išplaukia iš to, kad f – aprėžta; $E\varepsilon_i^2 = \sigma^2, \forall i = \overline{1, n}$.)

Kadangi $E\xi^2$ yra baigtinis, todėl turime, kad

$$P \left(\left| \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) \right| > A \right) \rightarrow 0, \text{ kai } A \rightarrow \infty \Rightarrow \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) = O_p(1).$$

Tokiu būdu gavome, kad:

$$\frac{n^\alpha}{\sqrt{n}} \cdot \frac{1}{l^{*\alpha}} \cdot \left| \sum_{i=1}^n \left(\varepsilon_i \cdot f\left(\frac{i}{n}\right) \right) \right| \left| \frac{\sum_{i \in I} f\left(\frac{i}{n}\right) - \frac{l^*}{n} \cdot \sum_{i=1}^n f\left(\frac{i}{n}\right)}{\sum_{i=1}^n f^2\left(\frac{i}{n}\right)} \right| \xrightarrow{n \rightarrow \infty} 0.$$

Taigi, parodėme, kad (*) išraiškos paskutiniai trys nariai konverguoja į 0, kai $n \rightarrow \infty$. Todėl iš čia išplaukia, kad $n^{-\frac{1}{2}+\alpha} \cdot T(n; \alpha) \xrightarrow[n \rightarrow \infty]{p} \infty$, t.y. statistika T yra suderinta alternatyvos atžvilgiu, jeigu

$$\frac{n^\alpha}{\sqrt{n}} \cdot \frac{1}{l^{*\alpha}} \cdot a \cdot l^* \cdot |\beta' - \beta| = a \cdot \frac{l^{*1-\alpha}}{n^{\frac{1}{2}-\alpha}} \cdot |\beta' - \beta| \xrightarrow{n \rightarrow \infty} \infty.$$

□

Iš teoremos matome, kad kuo statistika greičiau artėja į begalybę, tuo didesnė tikimybė, kad bus aptiktas struktūrinis pasikeitimas.

Vėliau bus pateikti modeliavimo rezultatai, kai nagrinėjami keli atvejai: kai kinta epideminio struktūrinio pasikeitimo ilgis, ir kai parametro α reikšmė yra arti nulio ir arti $\frac{1}{2}$. Iš rezultatų matysime, kaip pasikeitimo ilgis ir parametras α paveikia tikimybę aptikti struktūrinį pasikeitimą.

3.3. T_n^r statistika

Dar vienas testas, kurio pagalba bandysime aptikti epideminius struktūrinius pasikeitimus yra analogiškas tik ką apibrėžtam, tik šiuo atveju vietoj paprastų modelio liekanų $\varepsilon_i, i = \overline{1, n}$, imamos rekursinės liekanos, kurios, kaip buvo minėta anksčiau, apibrėžiamos tokiu būdu:

Jeigu nagrinėsime tokį regresinį modelį:

$$y_t = x_t' \beta + u_t, \quad t = 1, \dots, n,$$

Tada $n - k$ rekursinių liekanų bus apibrėžiamos tokiu būdu¹⁶:

$$\tilde{u}_t = \frac{y_t - x_t' b_{t-1}}{\left(1 + x_t' (X_{t-1}' X_{t-1})^{-1} x_t\right)^{1/2}}, \quad t = k + 1, \dots, n,$$

Gautas liekanas naudojame dalinėms sumoms apskaičiuoti bei testinei statistikai sudaryti. Hipotezės tikrinimas atliekamas analogiškai kaip ir pirmų dviejų testų atveju.

3.4. T_n^{sum} statistika

Aukščiau pateikėme tris statistikas struktūriniams pasikeitimams aptikti, tačiau pagrindinis šio darbo tikslas – išvesti naują statistiką, kuri pasižymėtų tam tikromis patrauklesnėmis savybėmis.

Naujas testas, kaip ir T_n^{\max} statistikos atveju, konstruojamas remiantis regresijos liekanų dalinėmis sumomis $\hat{S}(k), k = 0, \dots, n$. Visi žymėjimai ir sąlygos sutampa su anksčiau apibrėžtomis (žr. 3.2. dalis).

Vėl gi tariame, kad mūsų nagrinėjamas kintamasis yra nusakomas tokiomis lygtimis:

¹⁶ Pažymėjimai paaiškinami 3.2. dalyje.

$$y_i = \begin{cases} \beta \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & \text{jei } 1 \leq i \leq k^*, m^* < i \leq n; \\ \beta' \cdot f\left(\frac{i}{n}\right) + \varepsilon_i, & \text{jei } k^* < i \leq m^*. \end{cases}$$

čia $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai su nuliniu vidurkiu ir dispersija $\sigma^2 > 0$, bei baigtiniais momentais $E|\varepsilon_1|^q < \infty$, $q > 2$. β ir β' yra nežinomi modelio parametrai.

Mūsų uždavinys – patikrinti nulinę hipotezę:

$$H_0 : k^* > n$$

t.y. kad nėra struktūrinio pasikeitimo, prieš alternatyvą:

$$\bar{H} : 1 < k^* < m^* < n$$

kad yra epideminis pasikeitimas.

Kaip jau buvo minėta, regresijos koeficientai vertinami klasikiniu mažiausių kvadratų metodu. Įvertinus modelį, apskaičiuojamos regresijos liekanų dalinės sumos $\hat{S}_k, k = 0, 1, \dots, n$, kurios naudojamos testinei statistikai apibrėžti. Šiuo atveju naujoji statistika atrodo taip:

$$T_n^{sum} = \frac{1}{n^2} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \left| \hat{S}(k+l) - \hat{S}(k) - \frac{l}{n} \cdot \hat{S}(n) \right| \cdot \left(\frac{n}{l} \right)^\alpha, \quad 0 < \alpha < \frac{1}{2}$$

$$T^{sum} = \int_0^{1-h} \int_0^h \frac{|V(s+h) - V(s) - h \cdot V(1)|}{h^\alpha} ds dh, \quad 0 < \alpha < \frac{1}{2}$$

Pirmoji yra empirinė statistika, kuri pagal žemiau apibrėžtą išvadą pagal pasiskirstymą konverguoja į teorinę statistiką, apibrėžtą antrąja lygtimi.

Hipotezės tikrinimas atliekamas analogiškai kaip ir abiejų testų atveju.

Toliau formuluojame teoremą apie statistikos konvergavimą bei pateikiame jos įrodymą.

Žymėjimai:

$$S(0) = 0, S(k) = \sum_{i \leq k} \varepsilon_i$$

$$\tilde{T}_n^{sum} = \frac{1}{n^2} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l) - S(k)| \cdot \left(\frac{n}{l}\right)^\alpha$$

$$T(x) = \int_0^1 \int_0^{1-h} \frac{|x(s+h) - x(s)|}{h^\alpha} ds dh, \quad W_n(t) = \sum_{k \leq nt} \varepsilon_k + \varepsilon_{[nt]+1} \cdot (n \cdot t - [n \cdot t]) = S(n \cdot t) + \varepsilon_{[nt]+1} \cdot (n \cdot t - [n \cdot t])$$

$W = \{W(t), t \in [0,1]\}$ – Wiener'io procesas.

2 Teorema: Jeigu $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai, $E\varepsilon_1 = 0$, ir

tenkina sąlygą: $\lim_{t \rightarrow \infty} t \cdot P(|\varepsilon_1| \geq t^{1/2-\alpha}) = 0$, tada:

$$P\left(n^{-1/2} \cdot \sigma^{-1} \cdot \tilde{T}_n^{sum} \leq a\right) \xrightarrow{n \rightarrow \infty} P(T(W) \leq a)$$

kiekviename ribinės funkcijos tolydumo taške a .

Irodymas:

Išsiveskime $T(W_n)$ išraišką.

$$\begin{aligned} T(W_n) &= \int_0^1 \int_0^{1-h} \frac{|W_n(s+h) - W_n(s)|}{h^\alpha} ds dh = \\ &= \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{|S(k+l-2) + \varepsilon_{k+l-1} \cdot (n \cdot (s+h) - (k+l-2)) - S(k-1) - \varepsilon_k \cdot (n \cdot s - (k-1))|}{h^\alpha} ds dh = \\ &= \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{\left| S(k+l-2) - S(k-1) + \frac{1}{n} \cdot \varepsilon_{k+l-1} \cdot \left((s+h) - \frac{k+l-2}{n} \right) - \frac{1}{n} \cdot \varepsilon_k \cdot \left(s - \frac{k-1}{n} \right) \right|}{h^\alpha} ds dh \end{aligned}$$

Apibrėžkime T_n išraišką tokiu būdu:

$$T_n = \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l) - S(k)| \cdot a_{k,l}(n), \quad \text{čia } a_{k,l}(n) = \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{1}{h^\alpha} ds dh.$$

Įrodykime, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |T(W_n) - T_n|$ konverguoja į nulį pagal tikimybę, kai $n \rightarrow \infty$.

Pažymėkime:

$$\tilde{T}_n = \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l-2) - S(k-1)| \cdot a_{k,l}(n)$$

$$B = \frac{1}{n} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left| \frac{\varepsilon_{k+l-1} \cdot \left((s+h) - \frac{k+l-2}{n} \right) - \varepsilon_k \cdot \left(s - \frac{k-1}{n} \right)}{h^\alpha} \right| dsdh$$

Tuomet:

$$|T(W_n) - T_n| = |T(W_n) - \tilde{T}_n + \tilde{T}_n - T_n| \leq |T(W_n) - \tilde{T}_n| + |\tilde{T}_n - T_n|.$$

Taip pat akivaizdu, kad: $|T(W_n) - \tilde{T}_n| \leq B$.

Parodysime, kad $\frac{1}{\sqrt{n} \cdot \sigma} B$ konverguoja į nulį pagal tikimybę, kai $n \rightarrow \infty$.

$$\begin{aligned} \frac{1}{\sqrt{n} \cdot \sigma} \cdot B &= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \frac{1}{n} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left| \frac{\varepsilon_{k+l-1} \cdot \left((s+h) - \frac{k+l-2}{n} \right) - \varepsilon_k \cdot \left(s - \frac{k-1}{n} \right)}{h^\alpha} \right| dsdh \leq \\ &\leq \frac{1}{\sigma} \cdot \frac{1}{n^{3/2}} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{|\varepsilon_{k+l-1}| \cdot \left| (s+h) - \frac{k+l-2}{n} \right| + |\varepsilon_k| \cdot \left| s - \frac{k-1}{n} \right|}{h^\alpha} dsdh \leq \\ &\leq \frac{1}{\sigma} \cdot \frac{1}{n^{3/2}} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{|\varepsilon_{k+l-1}| \cdot \frac{1}{n} + |\varepsilon_k| \cdot \frac{1}{n}}{h^\alpha} dsdh = \\ &= \frac{1}{\sigma} \cdot \frac{1}{n^{5/2}} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{|\varepsilon_{k+l-1}| + |\varepsilon_k|}{h^\alpha} dsdh = \frac{1}{\sigma} \cdot \frac{1}{n^{7/2}} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \int_{\frac{l-1}{n}}^{\frac{l}{n}} \frac{|\varepsilon_{k+l-1}| + |\varepsilon_k|}{h^\alpha} dh \leq \\ &\leq \frac{1}{\sigma} \cdot \frac{1}{n^{7/2}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \left(\frac{n}{l-1} \right)^\alpha \cdot \frac{1}{n} \cdot (|\varepsilon_{k+l-1}| + |\varepsilon_k|) + \frac{1}{\sigma} \cdot \frac{1}{n^{7/2}} \cdot \sum_{k=1}^{n-1} \int_0^{\frac{1}{n}} \frac{|\varepsilon_k| + |\varepsilon_k|}{h^\alpha} dh = \\ &= \frac{1}{\sigma} \cdot \frac{1}{n^{9/2-\alpha}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \frac{1}{(l-1)^\alpha} \cdot (|\varepsilon_{k+l-1}| + |\varepsilon_k|) + \frac{1}{\sigma} \cdot \frac{2}{n^{9/2-\alpha}} \cdot \sum_{k=1}^{n-1} \frac{1}{1-\alpha} \cdot |\varepsilon_k| = Z_1 + Z_2 = Z \end{aligned}$$

Toliau naudosimės Čebyšovo nelygybe: $P(Z > e) \leq \frac{1}{e} \cdot EZ$. Jeigu parodysime, kad $EZ \rightarrow 0$, kai $n \rightarrow \infty$,

tada įrodysime, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot B$ konverguoja į nulį pagal tikimybę, kai $n \rightarrow \infty$. Pažymėkime: $E|\varepsilon_1| = \mu$.

Tuomet:

$$\begin{aligned} EZ &= E(Z_1 + Z_2) = \frac{1}{\sigma} \cdot \frac{1}{n^{9/2-\alpha}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \frac{1}{(l-1)^\alpha} \cdot (E|\varepsilon_{k+l-1}| + E|\varepsilon_k|) + \frac{1}{\sigma} \cdot \frac{2}{n^{9/2-\alpha}} \cdot \sum_{k=1}^{n-1} \frac{1}{1-\alpha} \cdot E|\varepsilon_k| = \\ &= \frac{1}{\sigma} \cdot \frac{2 \cdot \mu}{n^{9/2-\alpha}} \cdot \sum_{l=2}^n \frac{n-l}{(l-1)^\alpha} + \frac{1}{\sigma} \cdot \frac{2\mu}{n^{9/2-\alpha} \cdot (1-\alpha)} \cdot (n-1) \leq \frac{1}{\sigma} \cdot \frac{2 \cdot \mu}{n^{9/2-\alpha}} \cdot n^2 + \frac{1}{\sigma} \cdot \frac{2\mu}{n^{7/2-\alpha} \cdot (1-\alpha)} \rightarrow 0, n \rightarrow \infty \\ &\Rightarrow \boxed{\frac{1}{\sqrt{n} \cdot \sigma} \cdot B \xrightarrow[n \rightarrow \infty]{p} 0} \end{aligned}$$

Nagrinėkime skirtumą $|\tilde{T}_n - T_n|$. Parodysime, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n - T_n|$ konverguoja į nulį pagal tikimybę,

kai $n \rightarrow \infty$.

$$\begin{aligned} \frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n - T_n| &= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \left| \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l-2) - S(k-1)| \cdot a_{k,l}(n) - \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l) - S(k)| \cdot a_{k,l}(n) \right| = \\ &= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \left| \sum_{l=1}^n \sum_{k=1}^{n-l} a_{k,l}(n) \cdot (|S(k+l-2) - S(k-1)| - |S(k+l) - S(k)|) \right| \leq \\ &\leq \frac{1}{\sqrt{n} \cdot \sigma} \cdot \left| \sum_{l=1}^n \sum_{k=1}^{n-l} a_{k,l}(n) \cdot |S(k+l-2) - S(k-1) - S(k+l) + S(k)| \right| = \\ &= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \left| \sum_{l=1}^n \sum_{k=1}^{n-l} a_{k,l}(n) (\varepsilon_k - \varepsilon_{k+l-1} - \varepsilon_{k+l}) \right| \leq \frac{1}{\sqrt{n} \cdot \sigma} \cdot \left| \sum_{l=1}^n \sum_{k=1}^{n-l} a_{k,l}(n) \cdot (|\varepsilon_k| + |\varepsilon_{k+l-1}| + |\varepsilon_{k+l}|) \right| \end{aligned}$$

[vertinkime integralą: $a_{k,l}(n) = \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{1}{h^\alpha} ds dh = \frac{1}{n} \cdot \int_{\frac{l-1}{n}}^{\frac{l}{n}} \frac{1}{h^\alpha} dh \leq \frac{1}{n^2} \cdot \frac{n^\alpha}{(l-1)^\alpha}$, kai $l = 2, 3, \dots, n$.

$$\begin{aligned} \frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n - T_n| &\leq \frac{1}{\sigma} \cdot \frac{1}{n^{5/2-\alpha}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \frac{1}{(l-1)^\alpha} \cdot (|\varepsilon_k| + |\varepsilon_{k+l-1}| + |\varepsilon_{k+l}|) + \frac{1}{\sqrt{n} \cdot \sigma} \cdot \sum_{k=1}^{n-1} a_{k,1}(n) \cdot (|\varepsilon_k| + |\varepsilon_k| + |\varepsilon_{k+1}|) = \\ &= \frac{1}{\sigma} \cdot \frac{1}{n^{5/2-\alpha}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \frac{1}{(l-1)^\alpha} \cdot (|\varepsilon_k| + |\varepsilon_{k+l-1}| + |\varepsilon_{k+l}|) + \frac{1}{\sqrt{n} \cdot \sigma} \cdot \sum_{k=1}^{n-1} \left(\frac{1}{n}\right)^{1-\alpha} \cdot \frac{1}{1-\alpha} \cdot (2 \cdot |\varepsilon_k| + |\varepsilon_{k+1}|) = \end{aligned}$$

$$= \frac{1}{\sigma} \cdot \frac{1}{n^{5/2-\alpha}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \frac{1}{(l-1)^\alpha} \cdot (|\varepsilon_k| + |\varepsilon_{k+l-1}| + |\varepsilon_{k+l}|) + \frac{1}{\sigma} \cdot \frac{1}{n^{3/2-\alpha}} \cdot \frac{1}{(1-\alpha)} \sum_{k=1}^{n-1} (2 \cdot |\varepsilon_k| + |\varepsilon_{k+1}|) = V_1 + V_2 = V$$

Vėl naudosimės Čebyšovo nelygybe: $P(V > e) \leq \frac{1}{e} \cdot EV$. Jeigu parodysime, kad $EV \rightarrow 0, n \rightarrow \infty$, tada

gausime, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n - T_n|$ konverguoja į nulį pagal tikimybę, kai $n \rightarrow \infty$. Taip pat vėl pažymėkime:

$E|\varepsilon_1| = \mu$. Tuomet:

$$\begin{aligned} EV = E(V_1 + V_2) &= \frac{1}{\sigma} \cdot \frac{1}{n^{5/2-\alpha}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \frac{1}{(l-1)^\alpha} \cdot (E|\varepsilon_k| + E|\varepsilon_{k+l-1}| + E|\varepsilon_{k+l}|) + \\ &+ \frac{1}{\sigma} \cdot \frac{1}{n^{3/2-\alpha}} \cdot \frac{1}{(1-\alpha)} \sum_{k=1}^{n-1} (2 \cdot E|\varepsilon_k| + E|\varepsilon_{k+1}|) = \\ &= \frac{1}{\sigma} \cdot \frac{3 \cdot \mu}{n^{5/2-\alpha}} \cdot \sum_{l=2}^n \frac{n-l}{(l-1)^\alpha} + \frac{1}{\sigma} \cdot \frac{3\mu}{n^{3/2-\alpha}} \cdot (n-1) \leq \frac{1}{\sigma} \cdot \frac{3 \cdot \mu}{n^{5/2-\alpha}} \cdot n^2 + \frac{1}{\sigma} \cdot \frac{3\mu}{n^{1/2-\alpha}} \cdot (1-\alpha) = \\ &= \frac{1}{\sigma} \cdot \frac{3 \cdot \mu}{n^{1/2-\alpha}} + \frac{1}{\sigma} \cdot \frac{3\mu}{n^{1/2-\alpha}} \cdot (1-\alpha) \rightarrow 0, n \rightarrow \infty, \quad \text{kai } \alpha < 1. \end{aligned}$$

$$\Rightarrow \boxed{\frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n - T_n| \xrightarrow[n \rightarrow \infty]{p} 0.}$$

$$\Rightarrow \boxed{\frac{1}{\sqrt{n} \cdot \sigma} \cdot |T(W_n) - T_n| \xrightarrow[n \rightarrow \infty]{p} 0.}$$

Norime gauti statistiką: $\tilde{T}_n^{sum} = \frac{1}{n^2} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l) - S(k)| \cdot \left(\frac{n}{l}\right)^\alpha$.

Turime tokią statistiką: $T_n = \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l) - S(k)| \cdot a_{k,l}(n)$.

Parodysime, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n^{sum} - T_n| \xrightarrow[n \rightarrow \infty]{p} 0$.

$$\frac{1}{\sqrt{n} \cdot \sigma} \cdot E|\tilde{T}_n^{sum} - T_n| = \frac{1}{\sqrt{n} \cdot \sigma} \cdot E \left| \sum_{l=1}^n \sum_{k=1}^{n-l} |S(k+l) - S(k)| \cdot \left(\frac{1}{n^2} \cdot \left(\frac{n}{l}\right)^\alpha - a_{k,l}(n) \right) \right| =$$

$$\begin{aligned}
&= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \left| \sum_{l=1}^n \sum_{k=1}^{n-l} E|S(k+l) - S(k)| \cdot \left(\frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - a_{k,l}(n) \right) \right| \leq \\
&\leq \frac{1}{\sqrt{n} \cdot \sigma} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} [E(S(k+l) - S(k))^2]^{1/2} \cdot \left| \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - a_{k,l}(n) \right| = \\
&= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \left[E \left(\sum_{i=k+1}^{k+l} \varepsilon_i \right)^2 \right]^{1/2} \cdot \left| \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - a_{k,l}(n) \right| = \\
&= \frac{1}{\sqrt{n} \cdot \sigma} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \sigma \cdot \sqrt{l} \cdot \left| \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - a_{k,l}(n) \right|
\end{aligned}$$

Įvertinkime integralą $a_{k,l}(n)$:

$$\begin{aligned}
\int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{1}{h^\alpha} ds dh &= \frac{1}{n} \cdot \int_{\frac{l-1}{n}}^{\frac{l}{n}} \frac{1}{h^\alpha} dh \geq \frac{1}{n} \cdot \left(\frac{n}{l} \right)^\alpha \cdot \frac{1}{n} = \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \\
\frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - a_{k,l}(n) &\leq \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha = 0 \Rightarrow \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \leq a_{k,l}(n) \\
\Rightarrow \left| \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha - a_{k,l}(n) \right| &= a_{k,l}(n) - \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha = \int_{\frac{l-1}{n}}^{\frac{l}{n}} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \frac{1}{h^\alpha} ds dh - \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \leq \\
&\leq \frac{1}{n^2} \cdot \left(\frac{n}{l-1} \right)^\alpha - \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha, \text{ kai } l = 2, 3, \dots, n
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\sqrt{n} \cdot \sigma} \cdot E|\tilde{T}_n^{sum} - T_n| &\leq \frac{1}{\sqrt{n}} \cdot \sum_{l=2}^n \sum_{k=1}^{n-l} \sqrt{l} \cdot \left(\frac{1}{n^2} \cdot \left(\frac{n}{l-1} \right)^\alpha - \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \right) + \frac{1}{\sqrt{n}} \cdot \sum_{k=1}^{n-1} \left[\frac{1}{n \cdot (1-\alpha)} \cdot \left(\frac{1}{n} \right)^{1-\alpha} - \frac{1}{n^2} \cdot n^\alpha \right] \leq \\
&\leq \frac{1}{\sqrt{n}} \cdot \sum_{l=2}^n (n-l) \cdot \sqrt{l} \cdot \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \cdot \left(\left(\frac{l}{l-1} \right)^\alpha - 1 \right) + \frac{n}{\sqrt{n}} \cdot \left[\frac{1}{(1-\alpha) \cdot n \cdot n^{1-\alpha}} - \frac{1}{n^2} \cdot n^\alpha \right] \leq \\
&\leq \frac{1}{\sqrt{n}} \cdot \sum_{l=2}^n (n-l) \cdot \sqrt{l} \cdot \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \cdot \left(\frac{l}{l-1} - 1 \right)^\alpha + \frac{1}{n^{-1/2}} \cdot \frac{1}{n^{2-\alpha}} \cdot \left[\frac{1}{1-\alpha} - 1 \right] \leq * \\
&\leq \frac{1}{\sqrt{n}} \cdot \sum_{l=2}^n n \cdot \sqrt{l} \cdot \frac{1}{n^2} \cdot \left(\frac{n}{l} \right)^\alpha \cdot \left(\frac{1}{l-1} \right)^\alpha + \frac{1}{n^{3/2-\alpha}} \cdot \frac{\alpha}{1-\alpha} = \frac{1}{n^{3/2-\alpha}} \cdot \sum_{l=2}^n l^{1/2-2\alpha} + \frac{1}{n^{3/2-\alpha}} \cdot \frac{\alpha}{1-\alpha} \leq
\end{aligned}$$

$$\leq \frac{n^{1/2-2\alpha}}{n^{3/2-\alpha}} + \frac{1}{n^{3/2-\alpha}} \cdot \frac{\alpha}{1-\alpha} = \frac{1}{n^{\alpha+1}} + \frac{1}{n^{3/2-\alpha}} \cdot \frac{\alpha}{1-\alpha} \xrightarrow{n \rightarrow \infty} 0$$

(Komentaras: žvaigždute pažymėta nelygybė išplaukia iš nelygybės:

$$a^\alpha - b^\alpha \leq |a-b|^\alpha, \quad 0 < \alpha \leq 1, \quad a, b > 0.)$$

Taigi parodėme, kad $\boxed{\frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n^{sum} - T_n| \xrightarrow[n \rightarrow \infty]{p} 0}$.

Taigi turime, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |T(W_n) - T_n| \xrightarrow[n \rightarrow \infty]{p} 0$. Kadangi įrodėme, jog $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |\tilde{T}_n^{sum} - T_n| \xrightarrow[n \rightarrow \infty]{p} 0$, todėl

teisinga: $\frac{1}{\sqrt{n} \cdot \sigma} \cdot |T(W_n) - \tilde{T}_n^{sum}| \xrightarrow[n \rightarrow \infty]{p} 0$. Žinome, kad $\frac{1}{\sqrt{n} \cdot \sigma} \cdot W_n$ konverguoja pagal pasiskirstymą į

Wiener'io procesą W Hölder'io erdvėje $H_\alpha^0[0,1]$ (12, 2006, p.195).

Taip pat nesunku patikrinti, kad funkcija $T : H_\alpha^0[0,1] \rightarrow \mathfrak{R}$ yra tolydi:

- pirmiausiai, nusakykime Hölder'io erdvę¹⁷. Tegu $\alpha \in (0,1)$. Tada Hölder'io erdvė su eksponente α

$$\text{apibrėžiama tokiu būdu: } H_\alpha^0[0,1] = \left\{ x : [0,1] \rightarrow \mathfrak{R} : \sup_{\substack{t,s \in [0,1] \\ 0 < t-s < \delta}} \frac{|x(t) - x(s)|}{|t-s|^\alpha} < \infty \right\} \text{ ir}$$

$$\|x\|_\alpha = |x(0)| + \sup_{\substack{t,s \in [0,1] \\ 0 < t-s < \delta}} \frac{|x(t) - x(s)|}{|t-s|^\alpha}.$$

- Parodykime, kad funkcija T yra tolydi. Tegu $x, y \in H_\alpha^0[0,1]$, tada

$$\begin{aligned} |T(x) - T(y)| &= \left| \int_0^{1-h} \int_0^{1-h} \frac{|x(s+h) - x(s)|}{h^\alpha} ds dh - \int_0^{1-h} \int_0^{1-h} \frac{|y(s+h) - y(s)|}{h^\alpha} ds dh \right| \leq \\ &\leq \int_0^{1-h} \int_0^{1-h} \frac{|(x-y)(s+h) - (x-y)(s)|}{h^\alpha} ds dh \leq \int_0^{1-h} \int_0^{1-h} \sup_{0 < h \leq 1} \frac{|(x-y)(s+h) - (x-y)(s)|}{h^\alpha} ds dh = \|x-y\|_\alpha \cdot \int_0^{1-h} \int_0^{1-h} ds dh = \end{aligned}$$

$$= \|x-h\|_\alpha \cdot \int_0^1 (1-h) dh = \frac{1}{2} \cdot \|x-h\|_\alpha \quad \Rightarrow |T(x) - T(y)| \leq c \cdot \|x-y\|_\alpha \quad \Rightarrow T \text{ tolydi funkcija.}$$

¹⁷ Išsamiau apie Hölder'io erdvę žr. (6).

Kadangi funkcija $T : H_\alpha^0[0,1] \rightarrow \mathfrak{R}$ yra tolydi, tai, remiantis tolydžiojo atvaizdžio principu (1 išvada, žr.

3, 1977; p.49), teisinga, kad $T\left(\frac{1}{\sqrt{n} \cdot \sigma} \cdot W_n\right) \xrightarrow[n \rightarrow \infty]{D} T(W)$, bet $T\left(\frac{1}{\sqrt{n} \cdot \sigma} \cdot W_n\right) = \frac{1}{\sqrt{n} \cdot \sigma} \cdot T(W_n)$.

(Arba kitaip konvergavimas pagal pasiskirstymą reiškia, kad: $P\left(\frac{1}{\sqrt{n} \cdot \sigma} \cdot T(W_n) \leq a\right) \rightarrow P(T(W) \leq a)$

kiekviename ribinės funkcijos tolydumo taške a .)

Be to, kadangi teisinga

$$\frac{1}{\sqrt{n} \cdot \sigma} \cdot |T(W_n) - \tilde{T}_n^{sum}| \xrightarrow[n \rightarrow \infty]{p} 0 \Rightarrow \boxed{P\left(\frac{1}{\sqrt{n} \cdot \sigma} \cdot \tilde{T}_n^{sum} \leq a\right) \rightarrow P(T(W) \leq a)}$$

tolydumo taške a .

□

Iš šios teoremos išplaukia tokia išvada:

Žymėjimai:

$$T_n^{sum} = \frac{1}{n^2} \cdot \sum_{l=1}^n \sum_{k=1}^{n-l} \left| \hat{S}(k+l) - \hat{S}(k) - \frac{l}{n} \cdot \hat{S}(n) \right| \cdot \left(\frac{n}{l}\right)^\alpha, \quad 0 < \alpha < \frac{1}{2}$$

$$T^{sum} = \int_0^1 \int_0^{1-h} \frac{|V(s+h) - V(s) - h \cdot V(1)|}{h^\alpha} ds dh, \quad 0 < \alpha < \frac{1}{2}$$

$$V(t) = W(t) - \int_0^t \left\{ \int_0^1 f^\tau(u) \cdot A^{-1} \cdot f(v) dW(v) \right\} du \quad (\text{žymėjimai yra tokie pat kaip apibrėžta 3.2. dalyje});$$

Išvada: Jeigu $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai ir tenkina sąlygą:

$$\lim_{t \rightarrow \infty} t \cdot P(|\varepsilon_1| \geq t^{1/2-\alpha}) = 0, \text{ tada:}$$

$$P(n^{-1/2} \cdot \sigma^{-1} \cdot T_n^{sum} \leq a) \xrightarrow[n \rightarrow \infty]{} P(T^{sum} \leq a)$$

kiekviename ribinės funkcijos tolydumo taške a .

Įrodymas analogiškas 2 teoremos įrodymui.

4. Modeliavimo rezultatai

Tam, kad patikrintume, kaip gerai aukščiau apibrėžti testai aptinka epideminį struktūrinį pasikeitimą, naudosisimės modeliavimo principu. Naudodami statistinį paketą R, modeliuosime n skirtingų kintamųjų su įvairaus tipo epideminiais pasikeitimais. Kadangi mus domina vidurkio arba trendo pasikeitimas, todėl modeliuosime tik kintamuosius, kurių tam tikru laikotarpiu pakinta vidurkis arba trendo funkcija¹⁸.

4.1. CUSUM statistika

Pirmiausiai nagrinėkime CUSUM testą, kuris, kaip buvo minėta, labiau pritaikytas vieno struktūrinio pasikeitimo atvejui. Šiuo atveju mūsų tikslas – parodyti, kad jis netinka epideminiam struktūriniam pasikeitimui aptikti. Tam tikslui, modeliuojame 300 imčių po 300 stebėjimų su vidurkio ir trendo pasikeitimu. Nagrinėjame tokius modelius:

<p>Vidurkio pasikeitimas</p> $y_t = \begin{cases} \alpha + \varepsilon_t, & 1 \leq t \leq k^* \text{ ir } m^* < t \leq n, \\ \alpha' + \varepsilon_t, & k^* < t \leq m^*. \end{cases}$	<p>Trendo funkcijos pasikeitimas</p> $y_t = \begin{cases} \beta \cdot t + \varepsilon_t, & 1 \leq t \leq k^* \text{ ir } m^* < t \leq n, \\ \beta' \cdot t + \varepsilon_t, & k^* < t \leq m^*. \end{cases}$
---	---

T.y. nagrinėjant vidurkio pasikeitimą nuo laiko momento k^* iki m^* modelio laisvojo nario reikšmė lygi α' , o likusiame periode α . Tuo tarpu trendo pasikeitimas gaunamas keičiant parametro β reikšmę. Nuo laiko momento k^* iki m^* ji lygi β' , o likusiame periode – β .

Naudodami CUSUM statistiką patikrinsime hipotezę, jog yra epideminis struktūrinis pasikeitimas, prieš alternatyvą, kad struktūrinio pasikeitimo nėra. Kiekvienai imčiai apskaičiuojame statistiką ir gautą reikšmę lyginame su kritine reikšme. Jeigu statistika T yra didesnė už kritinę reikšmę, tada nulinę hipotezę atmetame ir sakome, kad duomenyse yra struktūrinis pasikeitimas. Toliau lentelėje pateikiami gauti rezultatai, t.y. lentelės skiltyse pateikiama, kiek kartų iš 300 galimų buvo atmesta nulinė hipotezė:

¹⁸ Kaip atrodo sumodeliuoti kintamieji, pateikta 2 priede.

1 lentelė

	Vidurkio pasikeitimas (didelis)	Vidurkio pasikeitimas (mažas)	Trendo pasikeitimas (didelis)	Trendo pasikeitimas (mažas)
Psk.periodas [100]	300	207	300	300
Psk.periodas [50]	300	152	300	300
Psk.periodas [30]	300	77	300	300
Psk.periodas [20]	290	46	300	300
Psk.periodas [10]	112	28	300	274
Psk.periodas [5]	38	15	292	79
Psk.periodas [3]	20	12	110	30

* Kai trende nėra struktūrinio pasikeitimo, tada nulinė hipotezė atmetama 13 kartų iš 300.

Vidurkio ir trendo pasikeitimai nagrinėjami dvejopu didumo. Vidurkio atveju pradžioje nagrinėjamas didesnis pasikeitimas, kai vidurkis pasikeičia iš 0 į 3, o vėliau imamas mažesnis pasikeitimas: iš 0 į 1. Analogiškai nagrinėjama ir trendo funkcija. Tik šiuo atveju keičiasi trendo funkcijos nuolydžio koeficientas. Pradžioje pastarasis keičiasi iš 0,5 į 0,7, o vėliau nagrinėjamas atvejis, kai nuolydis keičiasi iš 0,5 į 0,55. Be to, imami ir skirtingi epideminio pasikeitimo ilgiai. Nagrinėjami atvejai, kai epideminis pasikeitimas lygus 100, 50, 30, 20, 10, 5 ir 3 stebėjimams. Iš viso nagrinėjami 28 atvejai.

Iš 1 lentelės matome, kad esant dideliame vidurkio ir trendo pasikeitimui, kai pasikeitimo periodas pakankamai didelis, struktūrinis pasikeitimas yra aptinkamas 100 procentu. Pasikeitimo periodo ilgiui mažėjant aptinkamumo procentas taip pat mažėja. Tuo tarpu, kai pasikeitimo mastas nėra didelis, tada vidurkio atveju, net ir esant dideliame pasikeitimo periodui (100 stebėjimų), epideminis struktūrinis pasikeitimas aptinkamas tik 69 procentų tikslumu. Periodui mažėjant aptinkamumo procentas sumažėja net iki 4 procentų (3 stebėjimai). Tuo tarpu trendo atveju net ir esant nedideliame pasikeitimui, kai pasikeitimo periodas pakankamai didelis, struktūrinis pasikeitimas aptinkamas 100 procentų tikslumu. Analogiškai, periodo ilgiui mažėjant, aptinkamumo tikslumas mažėja. Taip pat svarbu pastebėti, kad šis testas geriau atpažįsta trendo funkcijos pasikeitimą nei vidurkio. Vidutiniškai vidurkio pasikeitimas atpažįstamas 45 procentų tikslumu, tuo tarpu trendo funkcijos pasikeitimas atpažįstamas vidutiniškai net 83 procentų tikslumu.

Žemiau esančioje lentelėje pateikti rezultatai, gauti tikrinant hipotezę, kai vidurkio pasikeitimas yra ganėtinai mažas. Nagrinėjami trys atvejai: kai vidurkis pasikeičia iš 0 į 0,5; iš 0 į 0,3 ir iš 0 į 0,1. Iš

lentelės matome, kad esant labai mažam vidurkio pasikeitimui testas labai prastai aptinka epideminį struktūrinį pasikeitimą. Vidutinis atpažinimo procentas lygus 9.

2 lentelė

	Vidurkio pasikeitimas (0⇒0.5)	Vidurkio pasikeitimas (0⇒0.3)	Vidurkio pasikeitimas (0⇒0,1)
Psk.periodas [100]	53	31	17
Psk.periodas [50]	43	27	15
Psk.periodas [30]	30	24	13

4.2. T_n^{\max} statistika

Toliau nagrinėkime kitą metodą, kuris pritaikytas epideminiam struktūriniam pasikeitimui aptikti. Turėdami 3.2. dalyje apibrėžtas statistikas galime tikrinti hipotezę, ar duomenyse yra epideminis struktūrinis pasikeitimas. Kaip ir anksčiau mus domina patikrinti, kaip gerai šis testas aptinka struktūrinį pasikeitimą, t.y. taikydami modeliavimo principą iširsime, kiek kartų yra atmetama nulinė hipotezė, kai teisinga alternatyva. Tuo tikslu, kaip ir anksčiau, modeliuojame 300 imčių po 300 stebėjimų su vidurkio ir trendo pasikeitimu¹⁹ ir naudodami 3.2. dalyje apibrėžtą statistiką tikriname hipotezę, jog yra epideminis struktūrinis pasikeitimas. Toliau 3 lentelėje pateikiami gauti rezultatai.

3 lentelė

	Vidurkio pasikeitimas (didelis)			Vidurkio pasikeitimas (mažas)			Trendo pasikeitimas (didelis)			Trendo pasikeitimas (mažas)		
	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$
Psk.periodas [100]	300	300	300	300	300	300	300	300	300	300	300	300
Psk.periodas [50]	300	300	300	300	300	300	300	300	300	300	300	300
Psk.periodas [30]	300	300	300	260	285	290	300	300	300	300	300	300
Psk.periodas [20]	300	300	300	135	224	246	300	300	300	300	300	300
Psk.periodas [10]	287	300	300	35	51	78	300	300	300	300	300	300
Psk.periodas [5]	60	225	294	14	15	24	300	300	300	195	300	300
Psk.periodas [3]	21	41	194	13	12	13	300	300	300	40	283	300

* Kai nėra struktūrinio pasikeitimo, tada: (12, 11, 7).

¹⁹ Kaip ir *CUSUM* testo atveju, kintamieji modeliuojami naudojant tas pačias lygtis.

Kaip ir *CUSUM* statistikos atveju vidurkio ir trendo pasikeitimai nagrinėjami dvejetainio didumo, t.y. vidurkio pasikeitimas nagrinėjamas iš 0 į 3 (didesnis pasikeitimas) ir iš 0 į 1 (mažesnis pasikeitimas). Analogiškai nagrinėjama ir trendo funkcija: pradžioje funkcijos nuolydis keičiasi iš 0,5 į 0,7 (didesnis pasikeitimas), o vėliau nagrinėjamas atvejis, kai nuolydis keičiasi iš 0,5 į 0,55 (mažesnis pasikeitimas). Be to, imami ir skirtingi epideminio pasikeitimo ilgiai. Nagrinėjami atvejai, kai epideminis pasikeitimas lygus 100, 50, 30, 20, 10, 5 ir 3 stebėjimams. Parametro α reikšmė imama: 0,0625, 0,25, 0,375. Iš viso nagrinėjami 84 atvejai.

Iš 3 lentelės matome, kad nagrinėjant šią statistiką rezultatai yra geresni nei *CUSUM* statistikos atveju. Trendo pasikeitimo atveju beveik visais atvejais struktūrinis pasikeitimas yra aptinkamas su šimto procentų tikslumu (t.y. nulinė hipotezė, jog nėra struktūrinio pasikeitimo, yra atmetama 300 atveju iš 300 nagrinėtų, kai teisinga alternatyva). Tik esant nedideliam trendo funkcijos nuolydžio pasikeitimui, kai pasikeitimo periodas yra 5 arba 3 stebėjimai, o parametras α yra artimas nuliui, testas prasčiau aptinka epideminį pasikeitimą²⁰. Tuo tarpu vidurkio pasikeitimas aptinkamas prasčiau nei trendo atveju. Iš 3 lentelės matome, kad pasikeitimas aptinkamas šimto procentų tikslumu tuo atveju, kai pasikeitimo periodas yra pakankamai didelis (šiuo atveju, kai jis lygus 100, 50, 30 stebėjimų). Pasikeitimo periodui mažėjant aptinkamumo tikslumas mažėja. Taip pat vertėtų paminėti tai, kad kuo parametras α yra arčiau $\frac{1}{2}$, tuo geriau atpažįstamas epideminis struktūrinis pasikeitimas. Vidutiniškai vidurkio pasikeitimas atpažįstamas apytiksliai 70 procentų tikslumu, tuo tarpu trendo funkcijos pasikeitimas atpažįstamas vidutiniškai net 97 procentų tikslumu.

Žemiau esančioje lentelėje pateikti rezultatai, gauti tikrinant hipotezę, kai vidurkio pasikeitimas yra ganėtinai mažas. Nagrinėjami trys atvejai: kai vidurkis pasikeičia iš 0 į 0,5; iš 0 į 0,3 ir iš 0 į 0,1. Iš lentelės matome, kad kaip ir *CUSUM* statistikos atveju, kai vidurkio pasikeitimas mažėja, aptinkamumo procentas taip pat mažėja. Tačiau šiuo atveju, kai vidurkio pasikeitimas yra iš 0 į 0,5 ir pasikeitimo periodas didelis (100 ir 50 stebėjimų), tada tikimybė atpažinti struktūrinį pasikeitimą yra didesnė nei 50 procentų, tuo tarpu pasikeitimo periodui ir vidurkio pasikeitimui dar labiau sumažėjus, tikimybė aptikti pasikeitimą gerokai sumažėja. Vidutiniškai vidurkio pasikeitimas yra atpažįstamas 28,7 procentų tikslumu.

²⁰ Tuo atveju, kai trendo nuolydis pasikeičia iš 0,5 į 0,55, pasikeitimo periodas lygus trim stebėjimams, o parametras α reikšmė lygi 0,0625, tada pasikeitimas aptinkamas tik su apytiksliai 13 procentų tikslumu.

4 lentelė

	Vidurkio pasikeitimas (0⇒0.5)			Vidurkio pasikeitimas (0⇒0.3)			Vidurkio pasikeitimas (0⇒0,1)		
	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$
Psk.periodas [100]	280	275	257	136	130	91	23	19	14
Psk.periodas [50]	172	193	181	55	66	51	13	10	11
Psk.periodas [30]	68	93	91	27	26	21	12	10	8

4.3. T_n^r statistika

2.1. dalyje buvo pateiktas pakoreguotas Chow testas, kai pastarasis konstruojamas imant rekursines liekanas. Pastarąjį žingsnį pritaikysime aukščiau apibrėžtam metodui ir patikrinsime, ar įvedus rekursines liekanas, testas geriau aptinka epideminius struktūrinius pasikeitimus.

Vėlgi, modeliuojame 300 imčių po 300 stebėjimų su vidurkio ir trendo pasikeitimu bei tikriname nulinę hipotezę, kad nėra struktūrinio pasikeitimo, prieš alternatyvą, kad vis dėlto struktūrinis pasikeitimas egzistuoja. 5 lentelėje pateikti gauti rezultatai:

5 lentelė

	Vidurkio pasikeitimas (didelis)			Vidurkio pasikeitimas (mažas)			Trendo pasikeitimas (didelis)			Trendo pasikeitimas (mažas)		
	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$
Psk.periodas [100]	300	300	300	300	300	300	300	300	300	300	300	300
Psk.periodas [50]	300	300	300	300	300	300	300	300	300	300	300	300
Psk.periodas [30]	300	300	300	262	273	280	300	300	300	300	300	300
Psk.periodas [20]	300	300	300	167	199	215	300	300	300	300	300	300
Psk.periodas [10]	286	300	300	52	67	86	300	300	300	300	300	300
Psk.periodas [5]	105	211	294	23	23	22	300	300	300	244	300	300
Psk.periodas [3]	36	52	184	16	14	18	300	300	300	77	273	300

* Kai nėra struktūrinio pasikeitimo, tada: (9, 15, 19).

Jeigu palygintume gautus rezultatus su tais, kurie buvo gauti prieš tai, kai į modelį įtraukiamos ne rekursinės liekanos, matytume, kad rezultatai labai nežymiai skiriasi. Vienu atveju pirmas testas geriau aptinka struktūrinį pasikeitimą, kitu atveju – antras, tačiau bendrai paėmus, aptinkamumo procentas yra panašus. Vidutiniškai vidurkio pasikeitimas atpažįstamas apytiksliai 70,5 procentų tikslumu, tuo tarpu trendo funkcijos pasikeitimas atpažįstamas vidutiniškai net 97,6 procentų tikslumu.

Iš žemiau esančios lentelės, kurioje pateikti rezultatai, gauti tikrinant hipotezę, kai vidurkio pasikeitimas yra ganėtinai mažas, matome, kad rezultatai yra artimi tiems, kuriuos gavome, kad taikėme metodą su paprastomis liekanomis. Šiuo atveju vidutiniškai vidurkio pasikeitimas yra atpažįstamas 28,8 procentų tikslumu.

6 lentelė

	Vidurkio pasikeitimas (0⇒0.5)			Vidurkio pasikeitimas (0⇒0.3)			Vidurkio pasikeitimas (0⇒0,1)		
	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$
Psk.periodas [100]	262	252	225	126	115	92	26	27	23
Psk.periodas [50]	164	164	140	71	74	58	26	26	18
Psk.periodas [30]	82	94	86	41	47	33	19	18	16

4.4. T_n^{sum} statistika

3.4. dalyje išvedėme naują testinę statistiką epideminiams struktūriniais pasikeitimams aptikti. Vėlgi, mus domina patikrinti, kaip gerai šis testas aptinka epideminį struktūrinį pasikeitimą, t.y. mus domina iširti šio testo galią. Taikydami modeliavimo principą tikriname, kiek kartų iš 300 buvo atmesta nulinė hipotezė alternatyvos atžvilgiu. Kintamuosius modeliuojame taikydami tokį pat modelį, kaip buvo apibrėžta 4.1. dalyje. 7 lentelėje pateikti gauti rezultatai:

7 lentelė

	Vidurkio pasikeitimas (didelis)			Vidurkio pasikeitimas (mažas)			Trendo pasikeitimas (didelis)			Trendo pasikeitimas (mažas)		
	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$
Psk.periodas [100]	300	300	300	300	300	300	300	300	300	300	300	300
Psk.periodas [50]	300	300	300	299	299	299	300	300	300	300	300	300
Psk.periodas [30]	300	300	300	233	236	239	300	300	300	300	300	300
Psk.periodas [20]	300	300	300	126	130	130	300	300	300	300	300	300
Psk.periodas [10]	229	236	237	39	43	43	300	300	300	299	299	300
Psk.periodas [5]	57	59	60	21	24	24	300	300	300	160	166	168
Psk.periodas [3]	30	31	33	16	17	18	280	282	280	46	49	50

* Kai nėra struktūrinio pasikeitimo, tada: (12, 12, 12).

Iš 7 lentelės matome, kad esant dideliame vidurkio ir trendo pasikeitimui, kai pasikeitimo periodas pakankamai didelis, struktūrinis pasikeitimas yra aptinkamas 100 procentu. Pasikeitimo periodo ilgiui mažėjant aptinkamumo procentas taip pat mažėja. Tuo tarpu, kai pasikeitimo mastas nėra

didelis, tada vidurkio atveju, net ir esant pakankamai dideliam pasikeitimo periodui (20 stebėjimų), epideminis struktūrinis pasikeitimas aptinkamas tik apytiksliai 43 procentų tikslumu. Periodui mažėjant aptinkamumo procentas sumažėja net iki 8 procentų (3 stebėjimai). Tuo tarpu trendo atveju, kai trendo funkcijos pasikeitimas yra ganėtinai didelis, net ir esant trumpam pasikeitimo periodui struktūrinis pasikeitimas aptinkamas 94 procentų tikslumu. Tačiau, kai trendo funkcijos koeficientai pasikeičia mažesniu dydžiu, tada esant trumpam pasikeitimo periodui (5, 3 stebėjimai) struktūrinis pasikeitimas yra aptinkamas atitinkamai 55 ir 16 procentų tikslumu. Taip pat svarbu pastebėti, kad kaip ir kitų testų atveju, šis testas taip pat geriau atpažįsta trendo funkcijos pasikeitimą nei vidurkio.

Žemiau esančioje lentelėje pateikti rezultatai, kurie gauti tikrinant, kiek kartų iš 300 buvo atmesta nulinė hipotezė alternatyvos atžvilgiu, kai vidurkio pasikeitimas yra ganėtinai mažas. Iš lentelės matome, kad rezultatai artimi tiems, kuriuos gavome, kai taikėme metodą epideminiam struktūriniam pasikeitimui aptikti su paprastomis liekanomis, t.y. vidurkio pasikeitimui mažėjant, modelio galia mažėja.

8 lentelė

	Vidurkio pasikeitimas (0⇒0.5)			Vidurkio pasikeitimas (0⇒0.3)			Vidurkio pasikeitimas (0⇒0,1)		
	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$	$\alpha=0.0625$	$\alpha=0.25$	$\alpha=0.375$
Psk.periodas [100]	276	276	277	138	136	136	24	23	23
Psk.periodas [50]	162	164	164	59	61	61	21	20	20
Psk.periodas [30]	67	68	70	32	35	36	17	18	18

5. Testų palyginimas

5.1. p-reikšmių grafikai

Siekdami ištirti testų, skirtų hipotezėms tikrinti savybes, naudojame modeliavimo metodus. Lentelėse pateikėme gautus rezultatus bei glaustą jų analizę, tačiau norėtuši šiuos rezultatus pateikti tokia forma, kuri būtų kompaktiška ir lengvai įsisąmoninama. R.Davidson ir J.G.MacKinnon (5, 1997) straipsnyje aprašomi grafiniai metodai, simuliacinio rezultatus apibendrinti, kuriuos pritaikysime mūsų nagrinėjamiems testams palyginti.

Vienas iš įprastinių būdų tam tikro eksperimento rezultatams pateikti – sudaryti lenteles, kuriose pateikiama, kiek kartų statistikos T reikšmė viršija kritinę reikšmę²¹. Tačiau šis būdas turi mažiausiai du rimtus trūkumus. Pirmiausiai, lentelėse pateikiama informacija tik apie kelis T pasiskirstymo taškus. Antra, lenteles paprastai sunku interpretuoti ir iš jų sunku pamatyti, kaip imties dydis, laisvės laipsnių skaičius ir kiti veiksniai įtakoja testo dydį. (5, 1997; p.1)

Norėdami patikrinti, kaip dažnai mūsų nagrinėjami testai priima ar atmets nulinę hipotezę, brėšime p-reikšmių grafikus ir p-reikšmių skirtumų grafikus. p-reikšmių grafikas leidžia atskirti, kurios testinės statistikos pernelyg dažnai atmets nulinę hipotezę, kurios priešingai – pernelyg dažnai priima ir kurios atmets tiek, kiek turėtų atmesti. Tačiau, jeigu p-reikšmių grafikas rodo, kad visos testinės statistikos elgiasi teisingai, tada šie grafikai tampa nelabai reikšmingi lyginant testines statistikas. Tuomet prasminga brėžti p-reikšmių skirtumų grafiką.

Be to, p-reikšmių grafikas ir p-reikšmių skirtumų grafikas yra dažnai pranašesnis už dažnių lenteles tuo, kad čia neapsiribojama vienu įprastiniu reikšmingumo lygmeniu, kaip pavyzdžiui 0,05. Apskritai paėmus, 5 procentų reikšmingumo lygmuo niekuo ne ypatingas, tiesiog jis dažnai naudojamas. Tyrėjai gali panorėti nagrinėti 0,01 ar net 0,001 reikšmingumo lygmenį.

Tarkime, turime statistiką T , kuri yra pasiskirsčiusi pagal tam tikrą pasiskirstymo funkciją F , t.y. $T \sim F$, $F(t) = P(T \leq t)$. Kadangi mūsų atveju tikro pasiskirstymo nežinome, todėl jį aproksimuojame

empirine pasiskirstymo funkcija $\hat{F}(t) = \frac{1}{N} \cdot \sum_{k=1}^N I[L_k \leq t]$, kur L_1, L_2, \dots, L_N yra N statistikos T analogų.

Naudodami modeliavimo metodą generuojame n testinės statistikos T realizacijų $T_j, j=1, \dots, n$. Šiuo atveju

²¹ Būtent tai pateikėme aukščiau esančiose lentelėse.

statistikos T_j p-reikšmė bus tikimybė, kad T reikšmė bus kiek galima labiau nutolusi nuo T_j , atsižvelgiant į tam tikrą pasiskirstymą F , t.y.

$$p_j = 1 - P(T \leq T_j) = P(T > T_j) = 1 - F(T_j) \approx 1 - \hat{F}(T_j) = 1 - \frac{1}{N} \cdot \sum_{k=1}^N I[L_k \leq T_j] = \frac{1}{N} \cdot \sum_{k=1}^N I[L_k > T_j]$$

p-reikšmių grafikas ir p-reikšmių skirtumų grafikas yra paremtas statistikos T_j p-reikšmių empirine funkcija, kuri apibrėžiama tokiu būdu:

$$\hat{F}_0(\alpha_i) \equiv \frac{1}{n} \cdot \sum_{j=1}^n I[p_j \leq \alpha_i]$$

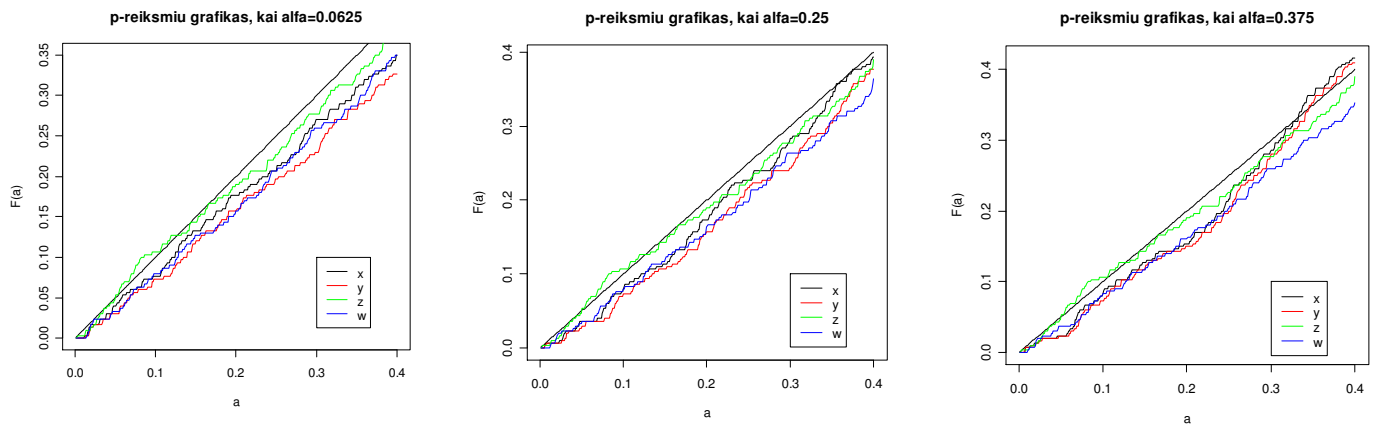
Šiuo atveju $\alpha_i, i = 1, \dots, m$, yra pasirinkti reikšmingumo lygmenys α .

Šiuo konkrečiu atveju reikšmingumo lygmens α vektorių pasirinksime tokį:

$$\alpha = (0,001; 0,002; \dots; 0,400), \text{ kur } m = 400.$$

p-reikšmių grafikas yra brėžiamas x -ų ašyje atidedant pasirinktą α vektorių, o y -ų ašyje empirinės pasiskirstymo funkcijos $\hat{F}_0(\alpha_i)$ reikšmes. Toliau pateikti visų nagrinėtų testų p-reikšmių grafikai.

4 paveikslas: p-reikšmių grafikai



Pastaba: Paveiksle naudojami tokie pažymėjimai:

$x - T_n^{\max}$ statistika;

$y - T_n^r$ statistika;

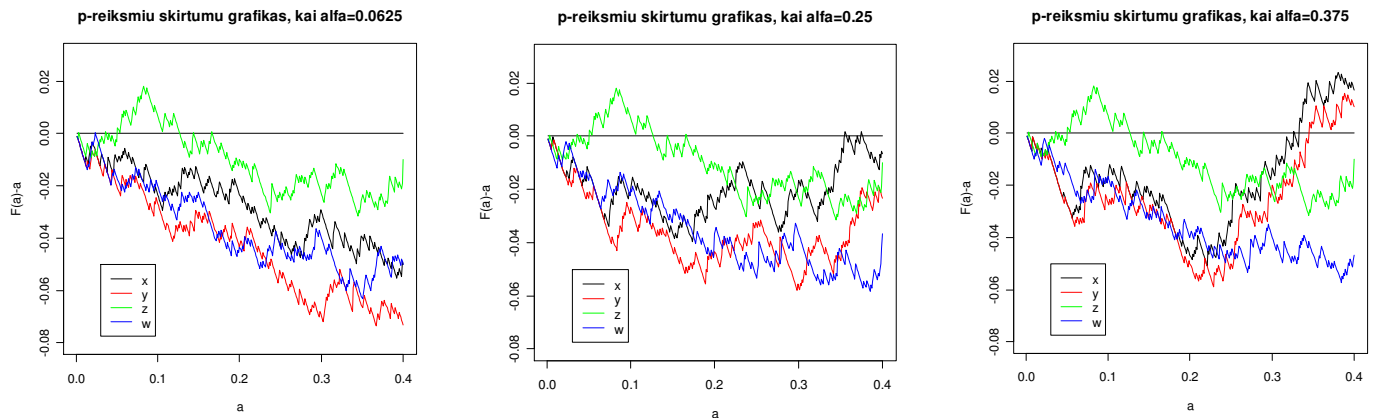
$z - CUSUM$ statistika;

$w - T_n^{sum}$ statistika.

Iš paveikslo matome, kad visi testai prie nulinės hipotezės elgiasi pakankamai gerai, t.y. visu testų p-reikšmių grafikai yra netoli nutolę nuo 45° tiesės. Atidžiau pažvelgus į grafikus matyti, kad geriausiai prie nulinės hipotezės elgiasi *CUSUM* testas, t.y. žalia spalva pažymėta kreivė svyruoja apie 45° tiesę. Tai rodo, kad *CUSUM* testo atveju nulinė hipotezė yra atmetama tinkamą kiekį kartų. Tuo tarpu kiti testai beveik visiems reikšmingumo lygmenims per retai atmeta nulinę hipotezę, t.y. kitų testų kreivės išsidėsčiusios žemiau 45° tiesės. Tik, kai parametras α^{22} lygus 0,375, tada T_n^{\max} ir T_n^r statistikos per dažnai atmeta nulinę hipotezę, kai reikšmingumo lygmuo yra didesnis nei 0,3.

p-reikšmių skirtumų grafikai brėžiami x-ų ašyje atidedant pasirinktas vektoriiaus α reikšmes, o y-kų ašyje – p-reikšmių empirinės pasiskirstymo funkcijos ir α skirtumus, t.y. $\hat{F}_0(\alpha_i) - \alpha_i$. Gavome tokius rezultatus:

5 paveikslas: p-reikšmių skirtumų grafikai



Šiuo atveju aiškiau matosi, kurie testai geriau, o kurie prasčiau elgiasi prie nulinės hipotezės. Iš p-reikšmių skirtumų grafikų vėlgi matyti, kad prie nulinės hipotezės geriausiai elgiasi *CUSUM* testas. Kai reikšmingumo lygmuo α mažas, tada *CUSUM* testas šiek tiek per dažnai atmeta nulinę hipotezę, tuo tarpu reikšmingumo lygmeniui didėjant, testas nulinę hipotezę atmeta rečiau nei derėtų. Kaip jau buvo minėta, kitų testų atveju nulinė hipotezė yra per dažnai priimama. Pirmame paveiksle visą laiką kreivės yra žemiau nulio. Tik antrojo ir trečiojo paveikslo dešinėje pusėje, kai reikšmingumo lygmuo yra didesnis už 0,3, kreivės kerta x-sų ašį, t.y. testai šiek tiek per dažnai atmeta nulinę hipotezę.

²² Šiuo atveju parametras α yra ne reikšmingumo lygmuo, bet testinės statistikos parametras.

5.2. Galios grafikai

Šiame darbe nagrinėjome testus, skirtus epideminiam struktūriniam pasikeitimams aptikti. Pateikėme keturis skirtingus testus, iš kurių trys specialiai pritaikyti tokio tipo struktūriniam pasikeitimams nustatyti, o kitas, t.y. *CUSUM* testas, labiau skirtas vieno struktūrinio pasikeitimo atvejams. Parodysime, kad pastarasis iš tikrųjų nėra tinkamas epideminiam struktūriniam pasikeitimams aptikti bei palyginsime kitus testus.

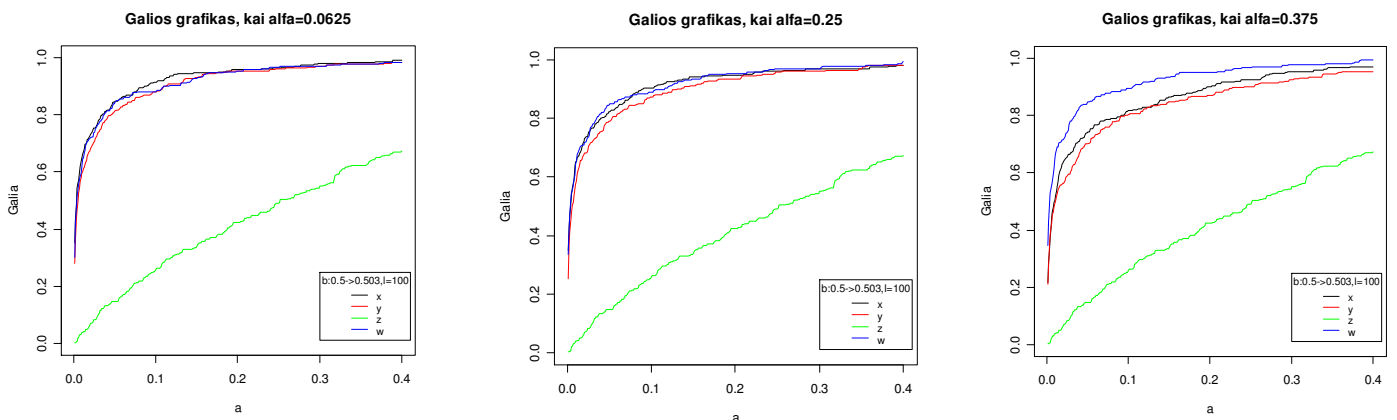
Vienas iš būdų palyginti statistinių testų efektyvumą, apskaičiuoti kiekvieno iš jų galią. Statistinio testo galia yra tikimybė, kad testas atmes nulinę hipotezę, kai ji yra klaidinga. Šiuo atveju statistinė galia yra testo gebėjimas aptikti pasikeitimą, kai jis iš tikrųjų egzistuoja. Kuo testo galia didesnė, tuo geriau jis aptiks epideminį struktūrinį pasikeitimą.

Pirmiausiai nagrinėkime trendo funkcijos pasikeitimą. 4 dalyje pateiktose lentelėse matėme, kad visi testai gana gerai aptinka trendo funkcijos pasikeitimus, t.y. parametro β pasikeitimus. Tačiau mus labiau domina, nuo kada testas pradeda blogai atskirti struktūrinį pasikeitimą, t.y. nagrinėsime, kokio dydžio turi būti parametro pasikeitimas ir epidemijos ilgis, kad testo galia imtų mažėti. Taigi, pateiksime grafikus, kuriuose matysime, kaip kinta testo galia, kai keičiasi parametro β pasikeitimo dydis ir kai kinta epideminio pasikeitimo ilgis. Pateiksime kelis variantus.

Gavome tokius rezultatus:

Pirmiausiai nagrinėjame atvejį, kai epidemijos ilgis yra didelis (100 stebėjimų iš 300), o parametro β pasikeitimas yra labai mažas: $0,5 \rightarrow 0,503$.

6 paveikslas: Galios grafikai, kai β keičiasi $0,5 \rightarrow 0,503$, o $l = 100$

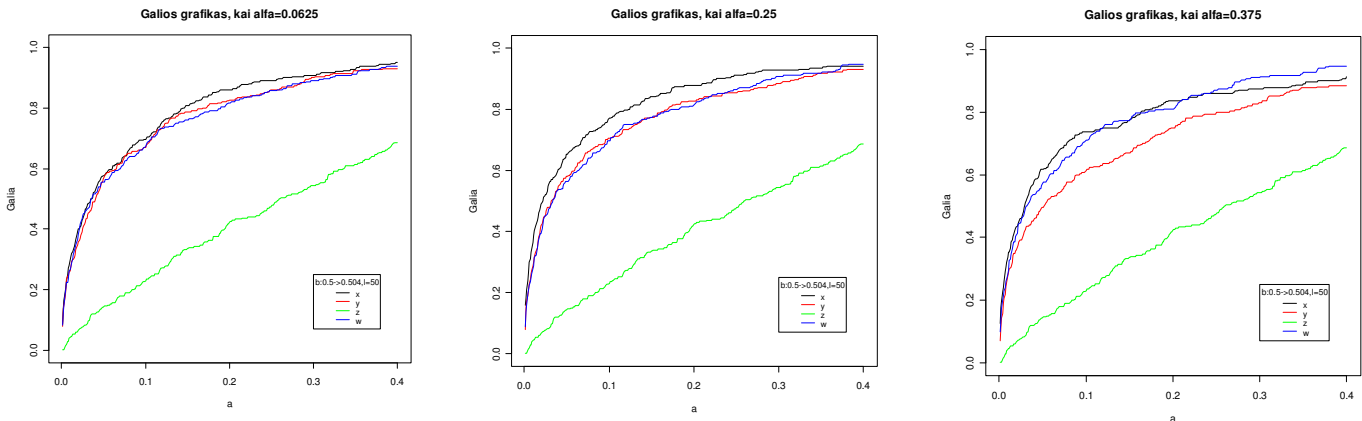


Iš pateiktų grafikų galima pastebėti tokius dalykus:

- ✓ *CUSUM* statistikos galia labai maža, t.y. šis testas labai prastai aptinka epideminį pasikeitimą;
- ✓ Kitų nagrinėjamų testų galia beveik sutampa, t.y. šie testai gana gerai aptinka net ir tokį nežymų parametro β pasikeitimą;
- ✓ Parametro α reikšmei didėjant, T_n^{sum} statistikos galia išlieka beveik nepakitusi, tuo tarpu T_n^{max} ir T_n^r statistikų galia sumažėja. Tai rodo trečiasis grafikas, kuriame aiškiai matyti, kad T_n^{sum} statistikos galios kreivė (mėlyna linija) išlieka panaši į pirmame ir antrame grafikuose pateiktas kreives, tuo tarpu T_n^{max} ir T_n^r statistikų galios kreivės pasislenka žemyn.

Toliau mažiname epidemijos ilgį iki $l = 50$. Šiuo atveju parametras β pasikeičia didesniu dydžiu: 0,5->0,504.

7 paveikslas: Galios grafikai, kai β keičiasi 0,5->0,504, o $l = 50$

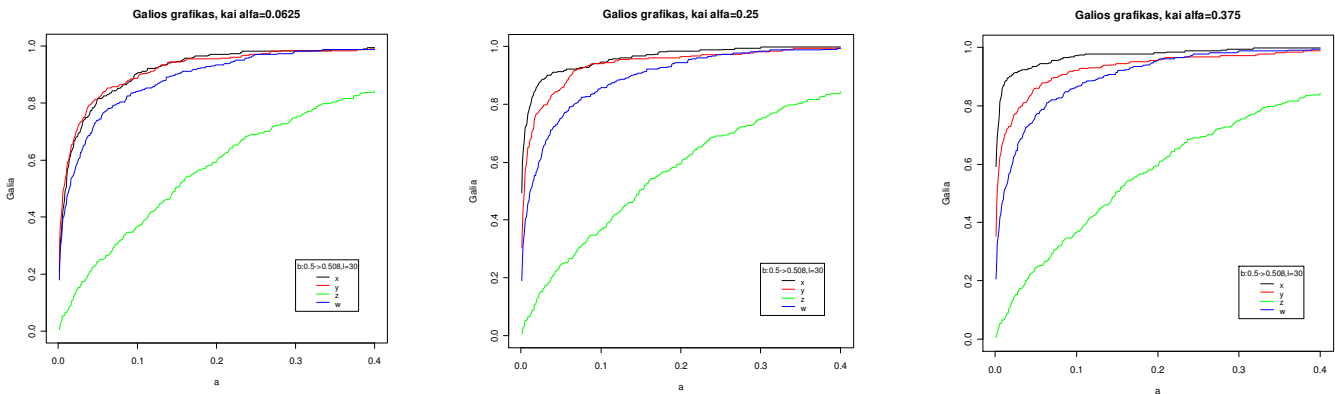


Gavome tokius rezultatus:

- ✓ *CUSUM* statistikos galia taip pat labai maža, t.y. šis testas labai prastai aptinka epideminį pasikeitimą;
- ✓ Kitų nagrinėjamų testų, kurie pritaikyti epideminiam struktūriniam pasikeitimams tikrinti, galia beveik sutampa; net ir perpus sumažėjus epidemijos ilgiui testai gerai aptinka labai nežymų parametro β pasikeitimą;
- ✓ Parametrui α didėjant, T_n^r statistikos galia mažėja. T_n^{max} ir T_n^{sum} statistikų galia, galima sakyti, išlieka beveik nepakitusi.

Imkime dar mažesnę epidemijos ilgį: $l = 30$ ir didesnę parametro β pasikeitimą: $0,5 \rightarrow 0,508$.

8 paveikslas: Galios grafikai, kai β keičiasi $0,5 \rightarrow 0,508$, o $l = 30$

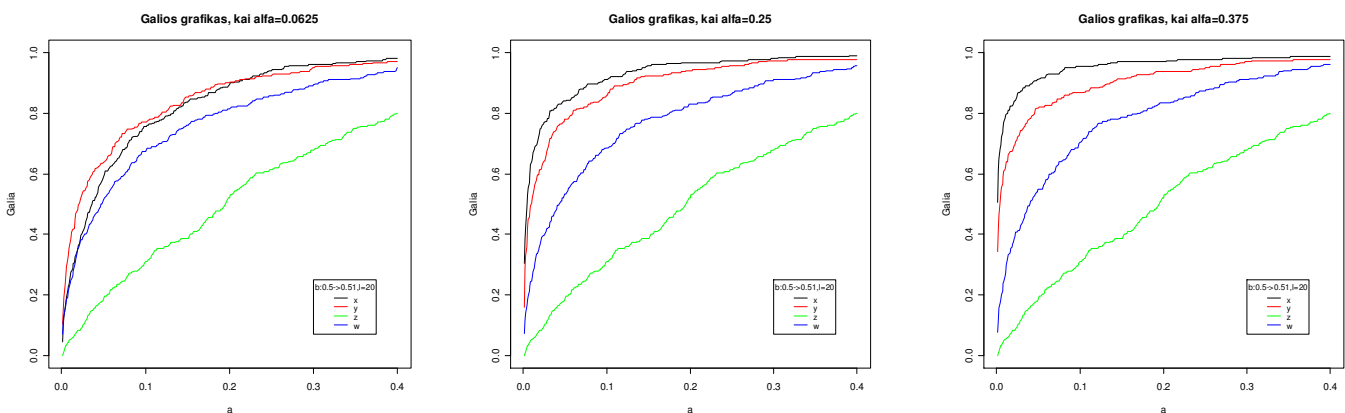


Iš 8 paveikslo galime pastebėti tokias tendencijas:

- ✓ Kaip ir ankstesniais atvejais, *CUSUM* statistikos galia išlieka labai maža. Tai dar kartą įrodo, kad šis testas nėra tinkamas epideminiams struktūriniais pasikeitimams tikrinti;
- ✓ Kai parametro α reikšmė artima nuliui, kitų testų galios artimos viena kitai; parametro α reikšmei augant, T_n^{\max} statistikos galia didėja, o T_n^r ir T_n^{sum} statistikų galia išlieka beveik nepakitusios.

Nagrinėkime atvejį, kai epideminio pasikeitimo ilgis lygus 20, o parametras β pasikeičia nuo 0,5 iki 0,51.

9 paveikslas: Galios grafikai, kai β keičiasi $0,5 \rightarrow 0,51$, o $l = 20$



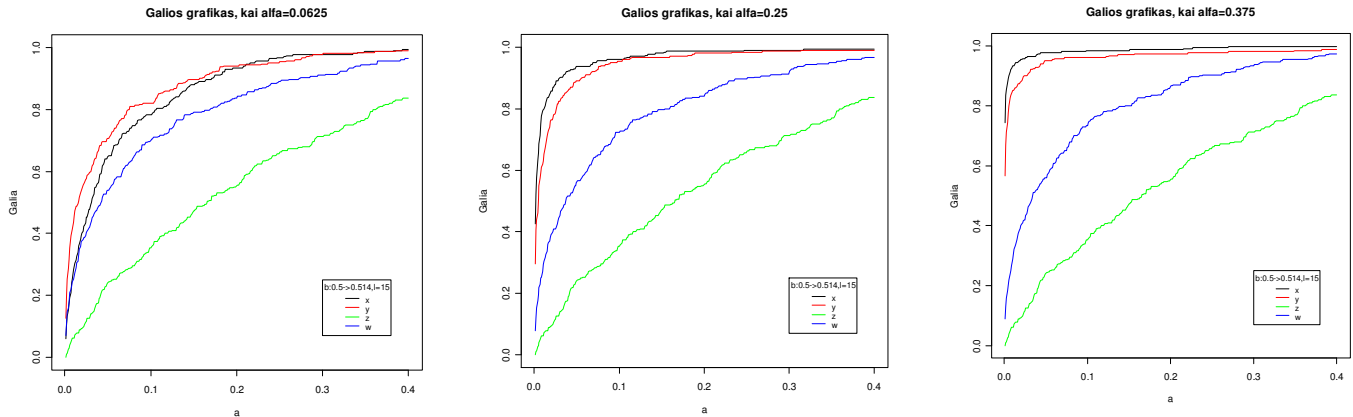
Iš 9 paveikslo galime padaryti tokias išvadas:

- ✓ *CUSUM* testas prastai aptinka struktūrinį pasikeitimą;
- ✓ T_n^{sum} statistikos galia, kintant parametru α , išlieka beveik nepakitusi, t.y. bet kokiai α reikšmei šis testas vienodai aptinka struktūrinį pasikeitimą;

- ✓ T_n^{\max} ir T_n^r statistikų galia yra didesnė už T_n^{sum} statistikos galią, tačiau kintant parametro α reikšmei T_n^r statistikos galia mažai kinta, o T_n^{\max} statistikos galia didėja didėjant α reikšmei.

Galiausiai, nagrinėkime atvejį, kai pasikeitimo ilgis lygus tik 15 stebėjimų. Šiuo atveju parametras β keičiasi nuo 0,5 iki 0,514.

10 paveikslas: Galios grafikai, kai β keičiasi 0,5->0,514, o $l = 15$



Gavome tokius rezultatus:

- ✓ *CUSUM* testo galia ir vėl labai prasta;
- ✓ T_n^{sum} statistikos galia, keičiantis parametru α , beveik nekinta, t.y. bet kokiai parametro α reikšmei testas vienodai atpažįsta struktūrinį pasikeitimą;
- ✓ T_n^{\max} ir T_n^r statistikų atveju, parametru α didėjant, testų galia taip pat didėja.

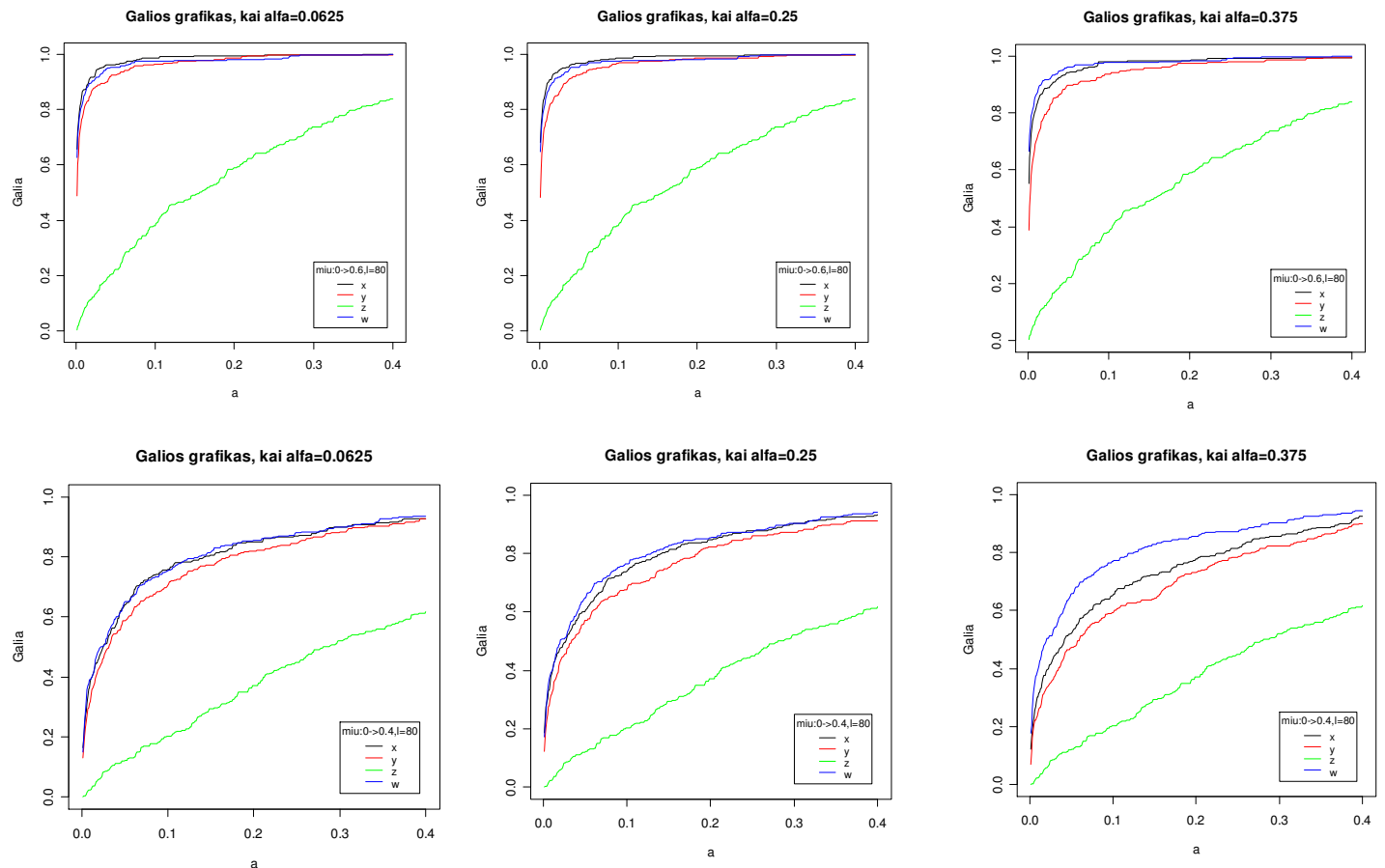
Taigi, ištyrę įvairius variantus, t.y. įvairius epideminio pasikeitimo ilgius bei skirtingo dydžio β pasikeitimus, galime padaryti tokias apibendrintas išvadas:

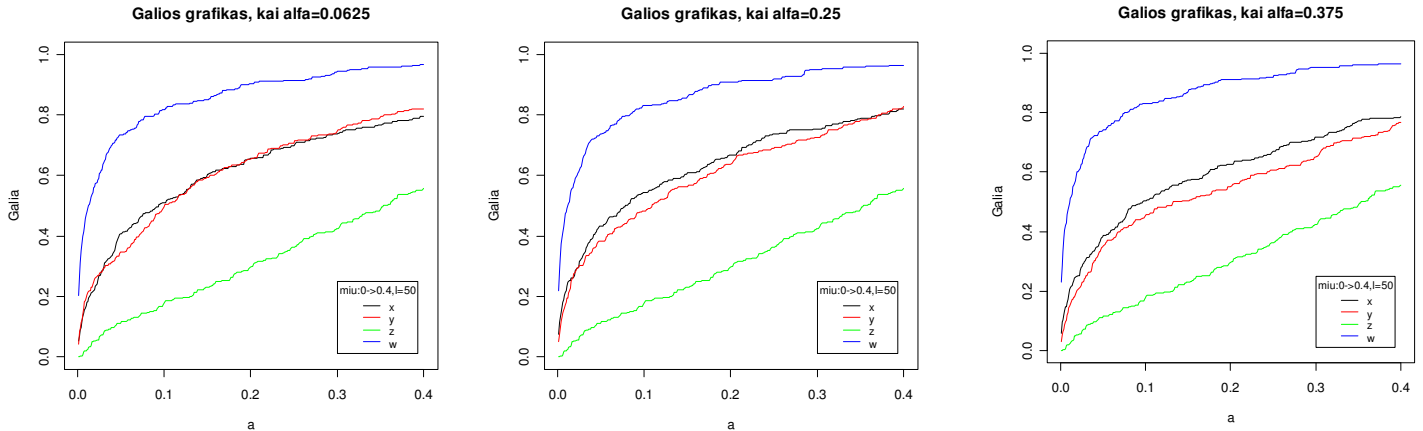
1. *CUSUM* testas akivaizdžiai nėra pritaikytas epideminiam struktūriniam pasikeitimui aptikti.
2. Kitų testų galia tam tikrais atvejais beveik sutampa, tačiau bendrai paėmus T_n^{\max} statistika geriausiai aptinka struktūrinį pasikeitimą, t.y. šio testo galia yra didžiausia. Tuo tarpu, antra pagal prastumą yra T_n^{sum} statistika, kuri rečiau aptinka epideminį struktūrinį pasikeitimą.
3. Iš pateiktų paveikslų matyti, kad kuo epideminio pasikeitimo ilgis didesnis, tuo geriau aptinkamas itin mažas parametro β pasikeitimas. Tuo tarpu epidemijos ilgiui mažėjant, modelių galia mažėja,

t.y. kai pasikeitimo ilgis nedidelis, tada modeliai geriau aptinka didesnį parametro β pasikeitimą, o itin mažo pasikeitimo testai neužfiksuoja.

Toliau nagrinėjame vidurkio pasikeitimą. Šiuo atveju mus domina patikrinti, kokia yra testų galia tikrinant hipotezę, jog duomenyse yra vidurkio pasikeitimas. Tam, kad iširtume, kaip kinta modelių galia priklausomai nuo esamo epideminio pasikeitimo ilgio bei vidurkio pasikeitimo dydžio, nagrinėsime tris skirtingus atvejus: pirmuoju atveju nagrinėjame epideminį pasikeitimą, kurio ilgis lygus 80 stebėjimų, o vidurkis pakinta nuo 0 iki 0,6. Antruoju atveju, epidemijos ilgis išlieka toks pat, tuo tarpu vidurkio pasikeitimas sumažėja, t.y. vidurkis pakinta nuo 0 iki 0,4. Ir paskutiniu atveju kinta ne vidurkio pasikeitimas, bet epidemijos ilgis, t.y. šiuo atveju vidurkis vėl kinta nuo 0 iki 0,4, o epideminio pasikeitimo ilgis sumažėja iki 50. Taigi, gavome tokius rezultatus:

11 paveikslas: Galios grafikai, kai nagrinėjamas vidurkio pasikeitimas





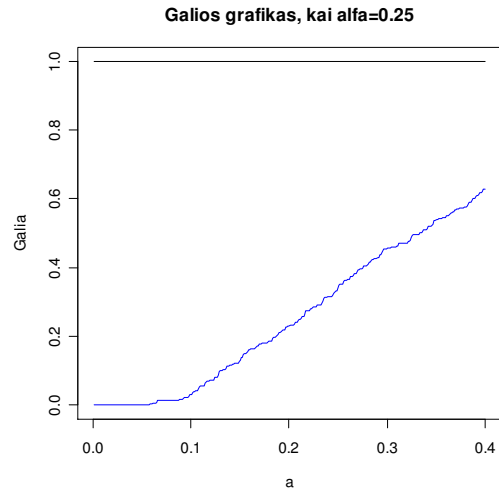
Iš 11 paveikslo galime padaryti tokias išvadas:

1. *CUSUM* testo galia yra maža ir vidurkio pasikeitimo atveju, t.y. šis testas nėra tinkamas ir epideminiam vidurkio pasikeitimui tikrinti;
2. Vidurkio pasikeitimui didėjant, testų galia taip pat didėja;
3. Epideminio pasikeitimo ilgiui mažėjant, T_n^{\max} ir T_n^r statistikų galia mažėja, tuo tarpu T_n^{sum} statistikos galia išlieka beveik nepakitusi;
4. Testų galia, keičiant parametro α reikšmę, taip pat beveik nepakinta;
5. Tikrinant vidurkio pasikeitimą, T_n^{sum} statistikos galia yra didžiausia.

Iš galios grafikų matome, kad T_n^{\max} statistikos galia trendo funkcijos pasikeitimo atveju yra didesnė nei naujai išvestos statistikos T_n^{sum} , tačiau pastebėsime vieną faktą, kuris parodys, kad pastaroji statistika turi pranašumą prieš T_n^{\max} statistiką.

Žemiau esančiame paveiksle pateikti šių dviejų statistikų galios grafikai. Juoda kreivė žymi T_n^{\max} statistikos galią, o mėlyna - T_n^{sum} statistikos galią. Šiuo atveju nagrinėjama situacija, kai epideminio pasikeitimo ilgis lygus tik vienam stebėjimui, o pats pasikeitimas didelis. Nagrinėjame trendo nuolydžio pasikeitimą, t.y. parametras β pasikeičia nuo 0,5 iki 0,9. Tačiau tokio tipo pasikeitimo mes negalime traktuoti kaip struktūrinio pasikeitimo, nes tai yra ne kas kita kaip tiesiog išskirtis. Todėl mes tikimės, kad mūsų testas priims nulinę hipotezę. Tačiau iš X paveikslo matome, kad T_n^{\max} statistika atmeta nulinę hipotezę 300 kartų iš 300 galimų. Todėl tai rodo, kad pastarasis testas turi tam tikrų trūkumų. Tuo tarpu T_n^{sum} statistikos galios grafikas rodo, kad šis testas priima nulinę hipotezę, t.y. šis

testas paprastos išskirties netraktuoja kaip galimo struktūrinio pasikeitimo. Tuo šis testas yra pranašesnis prieš T_n^{\max} statistiką.



Išvados

Nagrinėdami testus, skirtus struktūriniais pasikeitimams tikrinti, matėme, kad vieni testai geriau, o kiti prasčiau aptinka duomenyse esančius struktūrinius pasikeitimus. Kadangi *CUSUM* testas nėra pritaikytas epideminio struktūrinio pasikeitimo atvejui, todėl pastarasis rodė labai prastus rezultatus. Iš galios grafikų matėme, kad šio testo galia tiek vidurkio, tiek trendo funkcijos pasikeitimo atveju, buvo itin maža.

Tuo tarpu kiti trys testai gana gerai aptiko epideminius pasikeitimus:

T_n^{\max} statistikos, skirtos epideminiams struktūriniais pasikeitimams aptikti, galia tiek vidurkio, tiek trendo pasikeitimo atveju buvo gana didelė, tačiau iš paveikslų galime matyti, kad didžiausia galia šis testas pasižymi trendo funkcijos pasikeitimo atveju. Beveik visais nagrinėjamais atvejais T_n^{\max} statistikos galios kreivė buvo labiausiai nutolusi į viršų.

T_n^r statistikos galia artima T_n^{\max} statistikos galiai. Abiejų testų galios kreivės mažai viena nuo kitos nutolusios, tačiau tiek trendo, tiek vidurkio pasikeitimo atveju pirmosios statistikos galios kreivė beveik visais atvejais buvo žemiau T_n^{\max} statistikos galios kreivės. Tai rodo, kad vis dėl to paprastų liekanų pakeitimas rekursinėmis liekanomis nepagerino testo savybių.

Galiausiai, išvesta nauja statistika T_n^{sum} pasižymi didesne galia vidurkio pasikeitimo atveju. Iš visų nagrinėtų testų ji geriausiai aptinka vidurkio pasikeitimą. Nors T_n^{\max} statistikos galia trendo atveju yra didesnė už T_n^{sum} statistikos galią, tačiau pastaroji yra pranašesnė tuo, kad nereaguoja į duomenyse esančias išskirtis.

Be galios grafikų mus taip pat domino, kaip šie testai elgiasi prie nulinės hipotezės. Tam tikslui brėžėme p-reikšmių grafikus, kai teisinga nulinė hipotezė. Iš paveikslų matėme, kad visų testų p-reikšmių kreivės buvo artimos 45° tiesei, o tai rodo, kad visi testai gerai elgiasi prie nulinės hipotezės, t.y. jie atmeta nulinę hipotezę tiek kartų, kiek reikia.

Apibendrinant, būtų galima teigti, kad visų trijų testų, skirtų epideminiams struktūriniais pasikeitimams tikrinti, galia mažai skiriasi, tačiau naujai išvestas testas yra pranašesnis prieš kitus tuo, kad duomenyse esančių išskirčių netraktuoja kaip galimo struktūrinio pasikeitimo.

Literatūros sarakšas

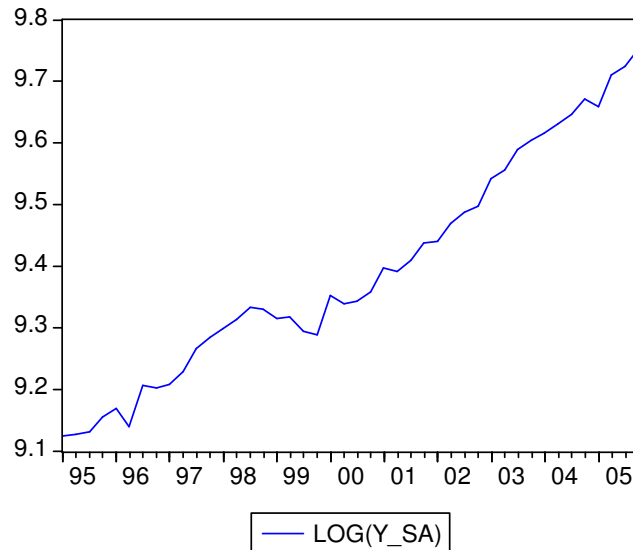
- 1 1 Basci S., Basci E. ir Zaman A., *A Method for Detecting Structural Breaks and an Application to the Turkish Stock Market*, <http://www.econturk.org/ye2.pdf> .
- 2 Ben-David D., Papell D.H., 1994, *The Great Wars, The Great Crash, and the Unit Root Hypothesis: Some New Evidence about an Old Stylized Fact*, National Bureau of Economic Research, Working Paper No.4752, <http://www.nber.org/papers/w4752.pdf> .
- 3 Биллингсли П., 1977, *Сходимость вероятностных мер*, „Наука“, Москва.
- 4 Brown R.L., Durbin J., Evans J.M., 1975, *Techniques for Testing the Constancy of Regression Relationships over Time*, Journal of the Royal Statistical Society. Series B (Methodological), Vol.37, No.2, p.149-192, <http://www.jstor.org/view/00359246/di993185/99p0205i/0> .
- 5 Davidson R., MacKinnon J.G., 1997, *Graphical Methods for Investigating the Size and Power of Hypothesis Tests*.
- 6 Hamadouche D., *Convergence de Processus Stochastiques à Trajectoires Hölderiennes*, Université des Sciences et Technologies de Lille, U.F.R. de Mathématiques Pures et Appliquées.
- 7 Hansen B.E., 2001, *The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity*, Journal of Economic Perspectives, Vol.15, No.4, p.117-128, <http://www.jstor.org/view/08953309/di021427/02p0046e/0> .
- 8 Harvey A.C., 1976, *An Alternative Proof and Generalization of a Test for Structural Change*, The American Statistician, Vol.30, No.3, p.122-123, <http://www.jstor.org/view/00031305/di020547/02p0089c/0> .
- 9 Maddala G. S., Kim In-Moo, 2002, *Unit Roots, Cointegration, and Structural Change*, ketvirtas leidimas, United Kingdom: Cambridge university press.

- 10 Perron P., 2005, *Dealing with Structural Breaks*, Palgrave Handbook of Econometrics, Boston University, Vol.1, <http://people.bu.edu/perron/papers/dealing.pdf> .
- 11 Perron P., 1989, *The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis*, *Econometrica*, Vol.57, No.6, <http://www.jstor.org/view/00129682/di952704/95p00332/0> .
- 12 Račkauskas A., 2006, *Hölderian properties of partial sums of regression residuals*, *Metrika* 63, p. 191-205, <http://www.springerlink.com/content/f77330xk3773w283/fulltext.pdf> .
- 13 Schroeter J., Fenn A., 2005, *Structural change in cigarette demand: cusum tests using panel data*, *Economics Bulletin*, Vol.9, No.8, p.1-11, <http://economicsbulletin.vanderbilt.edu/2005/volume9/EB-05I10004A.pdf> .
- 14 Sen P.K., 1982, *Invariance Principles for Recursive Residuals*, *The Annals of Statistics*, Vol.10, No.1., <http://www.jstor.org/view/00905364/di983913/98p0289k/0> .
- 15 Wang J. ir Zivot E., 1999, *A Time Series Model of Multiple Structural Changes in Level, Trend and Variance*, <http://129.3.20.41/eps/em/papers/9903/9903002.pdf> .
- 16 Zeileis A., Kleiber Ch., Krämer W. ir Hornik K., 2003, *Testing and dating of Structural Changes in Practice*, „Computational Statistics and data Analysis“, 44, 109-123 psl. <http://statmath.wu-wien.ac.at/~zeileis/papers/Zeileis+Kleiber+Kraemer-2003.pdf> .
- 17 Zeileis A., Leisch F., Hornik K. ir Kleiber Ch., *strucchange: An R package for Testing for Structural Change in Linear Regression Models*, <http://www.jstatsoft.org/v07/i02/strucchange.pdf> .
- 18 Zivot E. ir Andrews D.W.K., 1992, *Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis*, *Journal of Business & Economic Statistics*, July, Vol.10, No.3, <http://cowles.econ.yale.edu/P/cp/p08a/p0811.pdf> .

PRIEDAI

1 Priedas: Vienetinės šaknies tikrinimas Lietuvos duomenims

Lietuvos bendrojo vidaus produkto grafikas:



Atlikus Lietuvos BVP vienetinės šaknies testą, nulinė hipotezė, jog kintamasis turi vienetinę šaknį, buvo priimta, tuo tarpu įvedus vieną struktūrinį pasikeitimą ir dar kartą atlikus vienetinės šaknies testą, buvo gauti tokie rezultatai:

Dependent Variable: DLOG(Y_SA)
 Method: Least Squares
 Date: 10/29/06 Time: 20:45
 Sample(adjusted): 1995:3 2005:4
 Included observations: 42 after adjusting endpoints

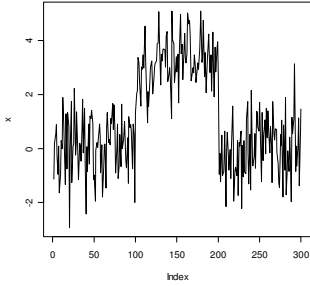
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.424699	1.128519	-3.920803	0.0004
DU28	0.129043	0.047920	2.692870	0.0107
T	-0.005281	0.001390	-3.799947	0.0005
T28	-0.004310	0.001533	-2.811580	0.0079
LOG(Y_SA)	0.486920	0.123722	3.935586	0.0004
DLOG(Y_SA(-1))	-0.419296	0.131742	-3.182707	0.0030
R-squared	0.411602	Mean dependent var		0.014916
Adjusted R-squared	0.329880	S.D. dependent var		0.020663
S.E. of regression	0.016915	Akaike info criterion		-5.189687
Sum squared resid	0.010300	Schwarz criterion		-4.941449
Log likelihood	114.9834	F-statistic		5.036606
Durbin-Watson stat	1.348578	Prob(F-statistic)		0.001339

Iš lentelės matome, kad šiuo atveju koeficientas prie $\log(y_{sa})$ yra reikšmingas, todėl nulinę hipotezę, jog yra vienetinė šaknis, atmetame. Taigi, gavome, kad Lietuvos bendrasis vidaus produktas yra stacionarus dydis su lūžusiu trendu.

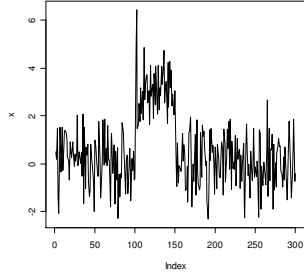
2 Priedas: Modeliuojamų eilučių paveikslai

Didelis vidurkio pasikeitimas:

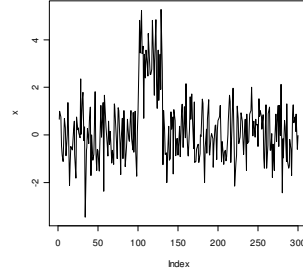
Vidurkis_d: [100]



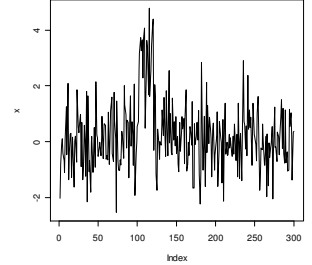
Vidurkis_d: [50]



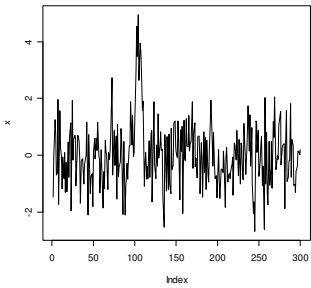
Vidurkis_d: [30]



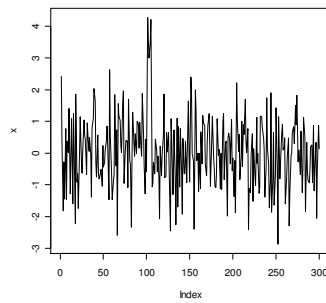
Vidurkis_d: [20]



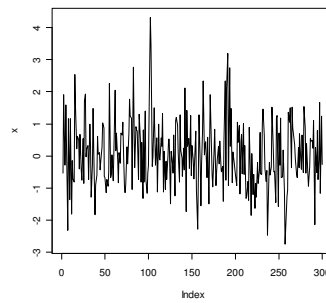
Vidurkis_d: [10]



Vidurkis_d: [5]

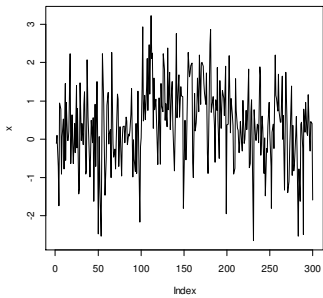


Vidurkis_d: [3]

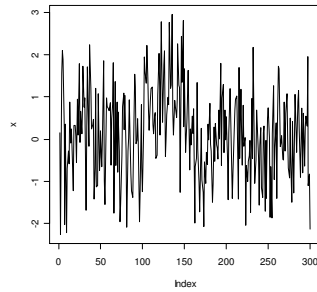


Mažas vidurkio pasikeitimas:

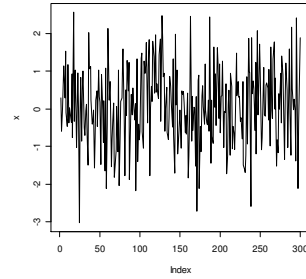
Vidurkis_m: [100]



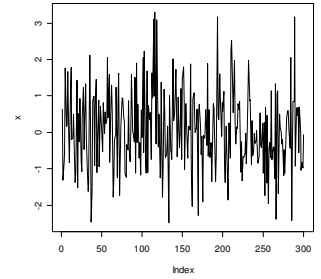
Vidurkis_m: [50]



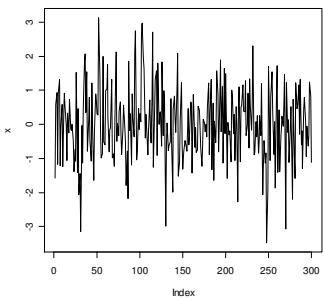
Vidurkis_m: [30]



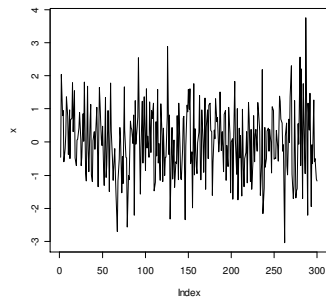
Vidurkis_m: [20]



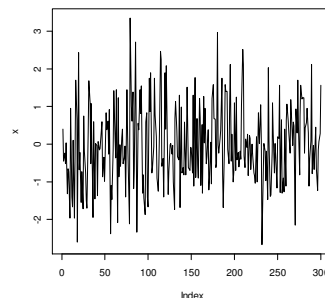
Vidurkis_m: [10]



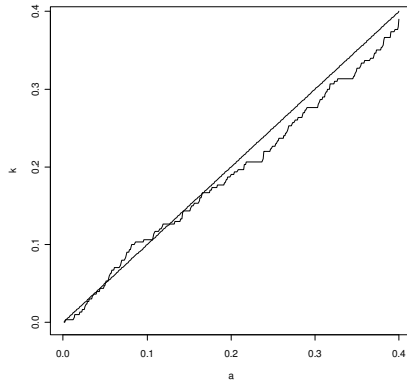
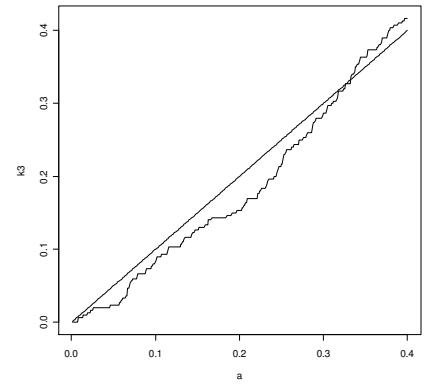
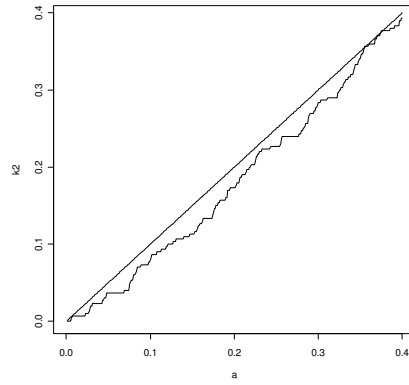
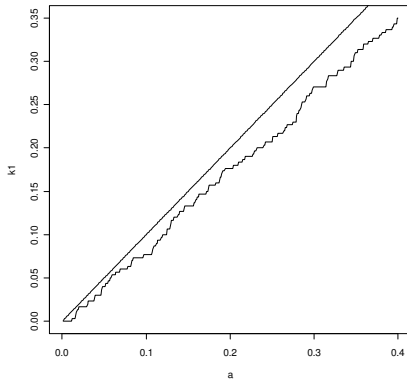
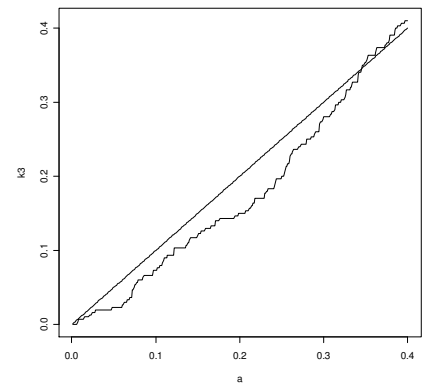
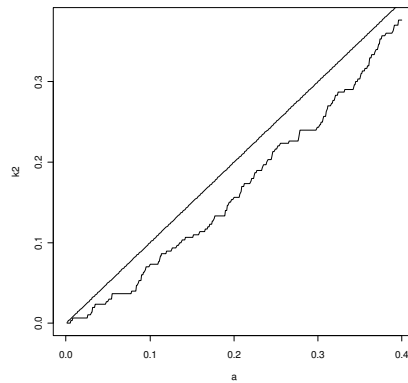
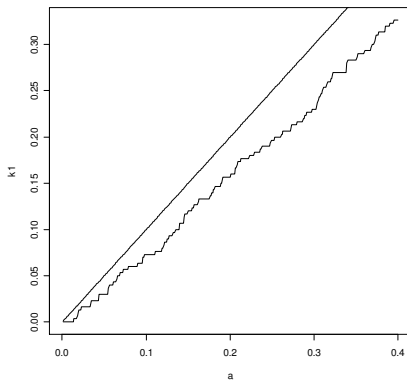
Vidurkis_m: [5]



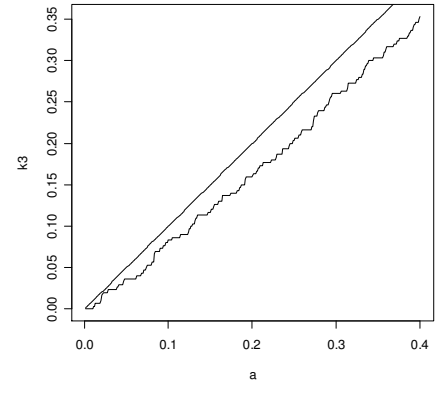
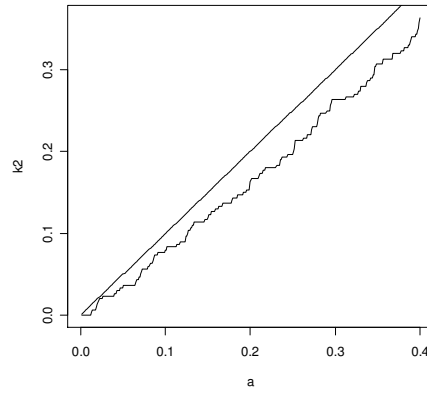
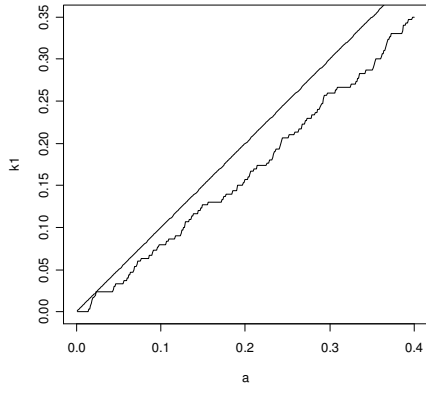
Vidurkis_m: [3]



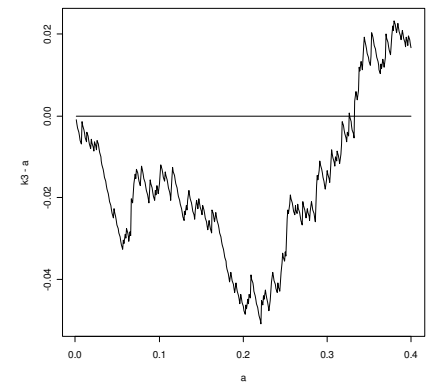
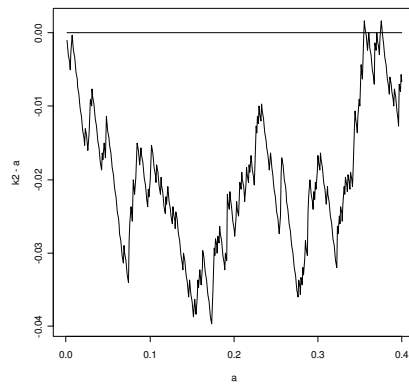
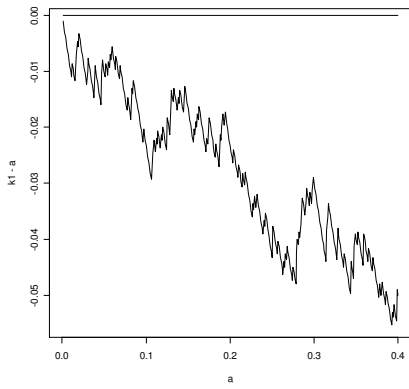
3 Priedas: Kiekvieno testo atskiri p-reikšmių ir p-reikšmių skirtumų grafikai

CUSUM statistika **T_n^{\max} statistika** **T_n^r statistika**

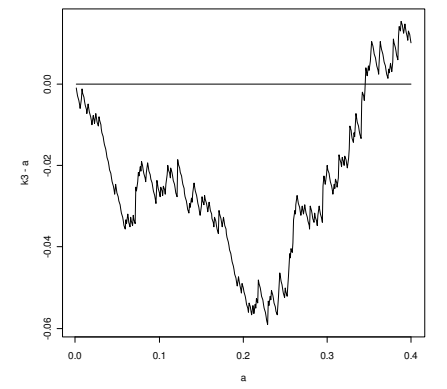
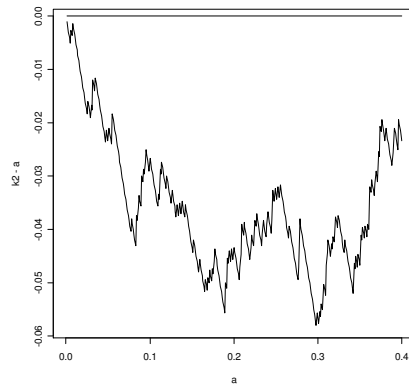
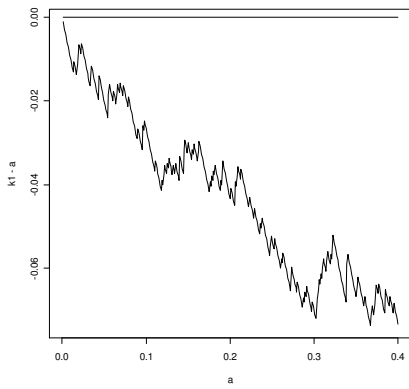
T_n^{sum} statistika



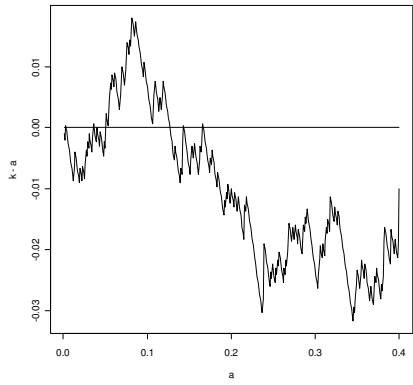
T_n^{max} statistika



T_n^r statistika



CUSUM statistika



T_n^{sum} statistika

