ივანე ჯავახიშვილის სახელობის თბილისის სახელმწიფო უნივერსიტეტი

დიმიტრი არაბიძე

ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი

მათემატიკის დეპარტამენტი

რეგრესიის ფუნქციის ინტეგრალური ფუნქციონალების შესახებ

ს ა დ ო ქ ტ ო რ ო   დ ი ს ე რ ტ ა ც ი ა

სამეცნიერო ხელმძღვანელები:

საქართველოს ეროვნული აკადემიის წევრ-კორესპონდენტი,
თბილისის სახელმწიფო უნივერსიტეტის პროფესორი,
ფიზიკა-მათემატიკურ მეცნიერებათა დოქტორი ელიზბარ ნადარაია

თბილისის სახელმწიფო უნივერსიტეტის ასოცირებული პროფესორი,
ფიზიკა-მათემატიკურ მეცნიერებათა დოქტორი გრიგოლ სოხაძე

თბილისი, 2015 წელი

შინაარსი

შესავალი

რეგრესიის ფუნქციის შეფასების აგების ამოცანა თანამედროვე მათემატიკური სტატისტიკის ერთ-ერთი უმთავრესი პრობლემაა. ვთქვათ, Y შემთხვევითი სიდიდე წარმოადგენს X შემთხვევითი სიდიდის ჩვენთვის უცნობ დეტერმინისტურ ფუნქციას. ამ სიდიდეთა (X, Y) წყვილზე (X_1, Y_1), (X_2, Y_2), . . . ,(X_n, Y_n) დამოუკიდებელ დაკვირვებათა საფუძველზე, საჭიროა აგებულ იქნეს რეგრესიის უცნობი $a(x)$ ფუნქციის ძალდებული შეფასება - $\hat{a}_n(x)$. ამასთან, რეგრესიის განტოლებას აქვს სახე

$$Y = a(X) + \varepsilon,$$

სადაც, ε შემთხვევითი შეშფოთებაა (ცდომილებაა). რეგრესიის ფუნქციის ინტენპრეტაცია შეგვიძლია პირობითი მათემატიკური ლოდინის ტერმინებში: $a(x) = E(Y \mid X = x)$.

დასმული ამოცანის ფარგლებში მრავალი, პრაქტიკულად საინტერესო ამოცანა, შეიძლება აღიწეროს. ასეთი ტიპის დამოკიდებულება გვხვდება მეცნიერების მრავალ დარგში, ფიზიკაში, ეკონომიკაში, მედიცინაში და სხვ. ცნობილია ამ პრობლემისადმი მიდგომის სხვადასხვა გზა. ამ ნაშრომში ჩვენ განვიხილავთ რეგრესიის ფუნქციის აგების სამ ყველაზე გავრცელებულ ხერხს - ნადარაია-ვატსონის, პრისტლი-ჩაოსა და გასერ-მიულერის შეფასებებს. ცხადია, ამოცანა არაპარამეტრულ სტატისტიკას მიეკუთვნება. ისტორიულად, რეგრესიის ფუნქციისა და ალბათური განაწილების სიმკვრივის შეფასებების ტექნიკა ერთმანეთის პარალელურად და მსგავსი მეთოდებით ვითარდებოდა. მეთოდი წარმოადგენს ე.წ. გულოვანი ტიპის შეფასებებს. აღნიშნული მიმართულებებით არსებობს ათასობით მონოგრაფიული თუ საჟურნალო პუბლიკაცია (იხ. მაგ. [1-7] და მათში მოცემული ლიტერატურის ჩამონათვალი).

ყველაზე გავრცელებული და პრაქტიკულად მნიშვნელოვანია ე.წ. ნადარაია-ვატსონის რეგრესიის ფუნქციის აგების მეთოდი. ეს მეთოდი შემოთავაზებულ იქნა 1964 წელს, ერთმანეთისაგან დამოუკიდებლად, ქართველი მათემატიკოსის ელიზბარ ნადარაიასა და ავსტრალიელი ჟოფრეი ვატსონის მიერ (იხ. [8-9]).

ნადარაია-ვატსონის შეფასებას აქვს სახე

$$\hat{a}_n(x) = \frac{\sum_{i=1}^{n} K\left(\frac{x - X_i}{b_n}\right) Y_i}{\sum_{i=1}^{n} K\left(\frac{x - X_i}{b_n}\right)},$$

სადაც $K(x)$ ალბათური განაწილების სიმკვრივის თვისებების მქონე ნებისმიერი ფუნქციაა, ხოლო b_n წარმოადგენს უსასრულოდ მცირე დადებით მიმდევრობას. შესაბამის პირობებში შესწავლილ იქნა ამ შეფასების ძალდებულობის, ასიმპტოტური ნორმალურობისა და სხვა თვისებები.

რეგრესიის შეფასების საინტერესო და მნიშვნელოვანი მეთოდი შემოთავაზებულ იქნა ბრიტანელი მათემატიკოსის მაურიუს პრისტლისა და ჩინელი მინ თე ჩაოს მიერ 1971 წელს (იხ. [10-11]).

პრისტლი-ჩაოს შეფასებისას $Y(t_1), Y(t_2), \dots, Y(t_n)$ დაკვირვებები დეტერმინისტულ წერტილებში გაგვაჩნია: $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq 1$. თვით რეგრესიის შეფასებას ასეთი სახე აქვს:

$$\hat{a}_n(t) = \sum_{i=1}^{n} K\left(\frac{t - t_i}{b_n}\right) \cdot \frac{t_i - t_{i-1}}{b_n} \cdot Y(t_i),$$

სადაც $K(x)$ ალბათური განაწილების სიმკვრივის თვისებების მქონე ნებისმიერი ფუნქციაა, ხოლო b_n წარმოადგენს უსასრულოდ მცირე დადებით მიმდევრობას. შესაბამის პირობებში შესწავლილ იქნა ამ შეფასების ძალდებულობის, ასიმპტოტური ნორმალურობისა და სხვა თვისებები. შეფასებები მიღებულია, აგრეთვე, რეგრესიის ფუნქციის წარმოებულებისთვისაც.

რეგრესიის ფუნქციის შეფასების საინტერესო მეთოდი შემოგვთავაზა გერმანელმა მათემატიკოსებმა თეო გასერმა და ჰანს-გეორგ მიულერმა 1979 წელს (იხ. [12-13]).

გასერ-მიულერის შეფასება იგება იმავე სქემით, როგორც პრისტლი-ჩაოს შეფასების შემთხვევაში. $Y(t_1), Y(t_2), \dots, Y(t_n)$ დაკვირვებები დეტერმინისტულ წერტილებში გაგვაჩნია: $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq 1$. ხოლო რეგრესიის შეფასებას ასეთი სახე აქვს:

$$\hat{a}_n(t) = \frac{1}{b_n} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} K\left(\frac{t - u}{b_n}\right) du \cdot Y(t_i).$$

აქაც, ისევე როგორც ზევით, $K(x)$ ალბათური განაწილების სიმკვრივის თვისებების მქონე ნებისმიერი ფუნქციაა, ხოლო b_n წარმოადგენს უსასრულოდ მცირე დადებით მიმდევრობას. რაც შეეხება ინტეგრალის საზღვრებს ისინი ასეთ პირობებს აკმაყოფილებენ:

$$0 = s_1 \leq s_2 \leq \cdots \leq s_n = 1,\ t_i \leq s_i \leq t_{i+1},\ i = 1, 2, 3, \dots, n-1 \text{ და } max_i |s_i - s_{i-1}| = O\left(\frac{1}{n}\right).$$

ამ შემთხვევაშიც სრულად იქნა გამოკვლეული შეფასების ძალდებულობისა და ასიმპტოტური ყოფაქცევის თვისებები.

როგორც ზემოთ აღვნიშნეთ, რეგრესიის გულოვანი ტიპის არაპარამეტრული მეთოდების დამუშავება მიმდინარეობდა ალბათური განაწილების სიმკვრივის გულოვანი ტიპის შეფასებების კვლევების პარალელურად. მას შემდეგ, რაც გამოკვლეულ იქნა ორივე მიმართულების შეფასებების ძირითადი თვისებები, აქტუალური გახდა ალბათური განაწილების სიმკვრივისა და

რეგრესიის ფუნქციათა ფუნქციონალების კვლევა. ეს გამოწვეულია პრაქტიკულად (და თეორიულადაც) საინტერესო ფუნქციონალების შეფასების საჭიროებით. ასეთი ფუნქციონალების რიგს მიეკუთვნებიან, მაგალითად, მომენტების ფუნქციონალები, ფიშერის ინფორმაციული ინტეგრალი, შენონის ენტროპიის ფუნქციონალი და მრავალი სხვა. საკმაო ინტენსივობით და სისრულით იქნა გამოკვლეული ალბათური განაწილების სიმკვრივის ფუნქციონალები (იხ. [14-20]). ამ მიმართულებით ყველაზე ზოგადი შედეგები მიღებულ იქნა ნაშრომებში [21-22].

რაც შეეხება რეგრესიის ფუნქციისა და მისი წარმოებულების ფუნქციონალების შეფასების პრობლემებს, აქ მიღებული შედეგები შედარებით მოკრძალებულია (იხ. [23-26]). ამიტომ აქტუალურია რეგრესიის ფუნქციისა და მისი წარმოებულების ფუნქციონალების თვისებების დადგენა. სწორედ ამ პრობლემატიკის ზოგიერთი საკითხის ანალიზს ეძღვნება წინამდებარე ნაშრომი.

ჩვენი მიდგომა ითვალისწინებს იმ გამოცდილებას, რომელიც დაგროვდა განაწილების სიმკვრივის და რეგრესიის ფუნქციისა და მათი წარმოებულების ფუნქციონალების შესწავლისას. მიღებული შედეგები მთლიანად მოიცავს ცნობილ შედეგებს ამ მიმართულებით და ანზოგადოებს მათ. ამიტომ, ჯერ ერთი, ყველგან განიხილება შეფასებები, რომლებიც დაკავშირებულია ე.წ. ”ჩასმის” (ინგლისურად მიღებულია ტერმინი “plug-in-estimator”) ოპერატორთან და მეორე მხრივ, ჩვენ ძირითად შეფასებებს განვიხილავთ სობოლევის ნორმით, რომელიც უფრო სრულფასოვან შედეგებს იძლევა და ბუნებრივადაა დაკავშირებული განხილულ პრობლემებთან.

მოვიყვანოთ ზოგიერთი შედეგი. ნაშრომში განხილულია რეგრესიის ფუნქციის ინტეგრალური სახის არაწრფივი ინტეგრალური ფუნქციონალების სტატისტიკური შეფასების პრობლემა. დასაწყისში წარმოდგენილია რეგრესიის შემდეგი ტიპის მოდელი:

$$Y(t) = a(t) + \varepsilon(t), \qquad (\mathbf{1})$$

სადაც $t \in [0,1]$, $\varepsilon(\cdot)$ არის შეშფოთება ნულოვანი საშუალოთი და სასრული და მუდმივი დისპერსიით, $E\varepsilon(t) = 0$, $E\varepsilon^2(t) = \sigma^2 < \infty$, $Y(t)$ შემთხვევითი ფუნქციაა, $a(t)$ კი - უცნობი რეგრესიის ფუნქციაა. დავუშვათ, გვაქვს n რიცხვი:

$$0 \le t_1 \le t_2 \le \cdots \le t_n \le 1,$$

სადაც თითოეული $t_k,\ k = 1,2,\dots,n$ დამოკიდებულია n-ზე, ისე, რომ $max_i|t_i - t_{i-1}| = O\left(\frac{1}{n}\right)$ და ამ წერტილებში გაგვაჩნია დაკვირვებები:

$$Y(t_1), Y(t_2), \dots, Y(t_n).$$

უცნობი რეგრესიის $a(t)$ ფუნქციის და მისი წარმოებულების შეფასებებად სხვადასხვა პარაგრაფებში წარმოდგენილია პრისტლი-ჩაოს, გასერ-მიულერის და ნადარაია-ვატსონის შეფასებები, რომლებიც განსაზღვრულია შემდეგნაირად:

ა) პრისტლი-ჩაოს შემთხვევაში:

$$\hat{a}_n(t) = \sum_{i=1}^{n} W\left(\frac{t-t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i), \tag{2}$$

$$\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} W^{(k)}\left(\frac{t-t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i) \tag{3}$$

ბ) გასერ-მიულერის შემთხვევაში:

$$\hat{a}_n(t) = \frac{1}{h_n} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i), \tag{4}$$

$$\hat{a}_n^{(k)}(t) = \frac{1}{{h_n}^{k+1}} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W^{(k)}\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i), \tag{5}$$

სადაც, $0 = s_1 \leq s_2 \leq \cdots \leq s_n = 1,\ t_i \leq s_i \leq t_{i+1},\ i = 1, 2, 3, \dots, n-1$ და $max_i|s_i - s_{i-1}| = O\left(\frac{1}{n}\right)$;

გ) ნადარაია-ვატსონის შემთხვევაში:

$$\hat{a}_n(t) = \sum_{i=1}^{n} \frac{W\left(\frac{t-t_i}{h_n}\right)}{\sum_{k=1}^{n} W\left(\frac{t-t_k}{h_n}\right)} \cdot Y(t_i), \tag{6}$$

$$\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} \alpha^{(k)}(t, t_i, h_n) \cdot Y(t_i) \tag{7}$$

სადაც

$$\frac{W\left(\frac{t-t_i}{h_n}\right)}{\sum_{k=1}^{n} W\left(\frac{t-t_k}{h_n}\right)} \equiv \alpha(t, t_i, h_n).$$

ყოველი $k = 0, 1, 2, \dots, m$-თვის. აქ და ქვემოთ ყველგან ჩავთვლით, რომ $\hat{a}_n^{(0)}(t) \doteqdot \hat{a}_n(t)$. სამივე შემთხვევაში $\{h_n,\ n = 1,2, \dots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი - ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა.

ვთქვათ, $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ უწყვეტი, შემოსაზღვრული და გლუვი ფუნქციაა. განხილულია შემდეგი სახის ინტეგრალური ფუნქციონალი:

$$I(a) = \int_{-\infty}^{\infty} \varphi\left(t, a(t), a'(t), \dots, a^{(m)}(t)\right) dt. \quad (\mathbf{8})$$

გვაქვს შერჩევა $(t_i, Y_i), i = 1, 2, \dots, n$. ეს ნიშნავს, რომ

$$Y_i = Y(t_i) = a(t_i) + \varepsilon(t_i). \quad (\mathbf{9})$$

$I(a)$-ს შეფასებისთვის ვიყენებთ ე.წ. "ჩასმის" შეფასებას. ანუ განვიხილავთ ფუნქციონალს:

$$I(\hat{a}_n) = \int_{-\infty}^{\infty} \varphi\left(t, \hat{a}_n(t), \hat{a}'_n(t), \dots, \hat{a}_n^{(m)}(t)\right) dt.$$

სამივე შემთხვევაში დამტკიცებულია ე.წ. წარმოდგენის თეორემები. ჩამოვთვალოთ ის პირობები, რომელთა შესრულებასაც მოვითხოვთ განხილული სიდიდეებისათვის.

პირობები $\boldsymbol{a}$-სათვის:

$(a1)$ $a = a(t)$ განსაზღვრულია და უწყვეტია [0, 1]-ზე, იგი მნიშვნელობებს იღებს $[-\Bbbk, \Bbbk]$ ინტერვალიდან;

$(a2)$ $a = a(t)$ ფუნქციას აქვს უწყვეტი წარმოებულები m რიგის ჩათვლით;

$(a3)$ ნებისმიერი $i = 0, 1, \dots, m$-სათვის, $a^{(i)}(\cdot) \in L_1([0, 1])$.

პირობები $\boldsymbol{\varepsilon_k = \varepsilon(t_k)}$-სათვის:

$(\varepsilon1)$ ε_k, $k = 1, 2, \dots$ დამოუკიდებელი და ერთნაირად განაწილებული შემთხვევითი სიდიდეებია;

$(\varepsilon2)$ $E\varepsilon_k = 0$, $D\varepsilon_k^2 = \sigma^2 < \infty$;

$\varphi = \varphi(x, x_0, \dots, x_m) \in C_b^2(R^{m+2})$ ფუნქციისათვის შემოვიღოთ აღნიშვნები

$$\frac{\partial\varphi}{\partial x_i} = \varphi_{(i)},\ i = 0, 1, \dots, m \quad \text{და} \quad \frac{\partial^2\varphi}{\partial x_i \partial x_j} = \varphi_{(ij)},\ i, j = 0, 1, \dots, m.$$

პირობები $\boldsymbol{\varphi}$-სათვის:

(φ) $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია უწყვეტია, შემოსაზღვრულია, ინტეგრებადია და გააჩნია მეორე რიგის ჩათვლით უწყვეტი წარმოებულები, რომელიღაც ღია, ამოზნექილ A არეში, რომელსაც შეიცავს $\mathbb{R} \times [-\Bbbk, \Bbbk]^{m+1}$ სიმრავლე; ამასთან φ-ის პირველი და მეორე წარმოებულები თანაბრად შემოსაზღვრულია A არეში, რაღაც $C_\varphi > 0$ მუდმივით.

ამ პირობების თანახმად, φ ფუნქციისათვის ნებისმიერი $i, j = 0, 1, \dots, m$-თვის გვაქვს:

$$sup\left\{\left|\varphi_{(i,j)}\right|(s, s_0, s_1, \dots, s_m): (s, s_0, s_1, \dots, s_m) \in A\right\} \le C_\varphi.$$

პირობები W-სათვის:

$(W1)$ $\int_{-\infty}^{\infty} W(t)\, dt = 1$;

$(W2)$ $W(t)$ ფუნქციას გააჩნია კომპაქტური მატარებელი $[-\tau,\ \tau]$ და ის უწყვეტად წარმოებადია $m \geq 1$ რიგის ჩათვლით. მაშინ უწყვეტი $W^{(i)}(t)$, $i = 0, 1, \ldots, m$ ფუნქციებისთვის $W^{(i)}(-\tau) = W^{(i)}(\tau) = 0$; აღვნიშნოთ $C_W > 0$ მუდმივი, რომლისთვისაც

$$sup_{t\in R}\left|W^{(i)}(t)\right| \leq C_W < \infty, \qquad i = 0, 1, \ldots, m;$$

ცხადია, ნებისმიერი $i = 0, 1, \ldots, m$-თვის $W^{(i)} \in L_1([-\tau,\ \tau])$.

პირობები h_n-სათვის:

(h_n) $\dfrac{\sqrt{\max(|logh_n|;loglogn)}}{\sqrt{n}h_n^{0,5+m}} \to 0$, როცა $n \to \infty$.

თეორემა 1: თუ შესრულებულია ჩამოთვლილი პირობები მაშინ, გასერ-მიულერის სქემისთვის სამართლიანია

$$I(\hat{a}_n) - I(a_n) = S_n(h_n) + R_n \qquad (\mathbf{10})$$

წარმოდგენა, სადაც S_n ის ტეილორის გაშლის ძირითადი შესაკრებია, რომელიც ჩვენს შემთხვევაში წარმოდგება, როგორც ერთმანეთისგან დამოუკიდებელი და ერთნაირად განაწილებული შემთხვევითი სიდიდეების ჯამი, ხოლო ნაშთით წევრს R_n-ს ალბათობით 1 აქვს რიგი

$$R_n = O\left(\frac{logn}{nh_n^{2m+2}}\right). \qquad (\mathbf{11})$$

თითოეული შემთხვევისათვის დამტკიცებულია $I(\hat{a}_n)$ შეფასების მკაცრად ძალდებულობა. მოვიყვანოთ ერთ-ერთი მათგანი.

თეორემა 2: ვთქვათ სრულდება თეორემა 1-ის პირობები. დადებითი $(h_n)_{n=1}^{\infty}$ მიმდევრობა შერჩეულია ისე, რომ როცა $n \to \infty$

$$\frac{logn}{nh_n^{2m+2}} \to 0. \qquad (\mathbf{12})$$

მაშინ ალბათობით 1

$$I(\hat{a}_n) \to I(a). \qquad (\mathbf{13})$$

შემდეგ ნაბიჯად სამივე შემთხვევისთვის დამტკიცებულია ცენტრალური ზღვარითი თეორემა:

თეორემა 3: ვთქვათ სრულდება თეორემა 1-ის პირობები. მაშინ, თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$ როცა $n \to \infty$, ადგილი აქვს კრებადობას

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2), \tag{14}$$

სადაც

$$r^2 = \sigma^2 \cdot \sum_{i=0}^{m} \left(\int_0^1 \varphi_{(i)} \left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \right)^2. \tag{15}$$

ეს თეორემა გამოყენებულია ზოგიერთი, პრაქტიკულად საინტერესო, ფუნქციონალისთვის.

ვთქვათ, $I_1(a) = \int_0^1 a^2(t)\, dt$. ამ შემთხვევაში $\varphi(t, x_0, x_1, \dots, x_m) = x_0^2$, როცა $x_0 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}]$, $b > 0$. მაშინ თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$, როცა $n \to \infty$, გვექნება შემდეგი კრებადობა:

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = 4\sigma^2 \left(\int_0^1 a(t) dt \right)^2.$$

განხილულია ფიშერის ინფორმაციული ინტეგრალი

$$I_2(a) = \int_0^1 \frac{\big(a'(t)\big)^2}{a(t)} dt.$$

ამ ფუნქციონალისთვის გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \frac{x_1^2}{x_2}$, $t \in [0,1]$, $a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. მაშინ თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$, როცა $n \to \infty$, გვექნება $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2)$, სადაც

$$r^2 = \sigma^2 \cdot \left(- \int_0^1 \left(\frac{\big(a'(t)\big)^2}{\big(a(t)\big)^2} - \frac{2a'(t)}{a(t)} \right) dt \right)^2$$
$$= \sigma^2 \left(\frac{a'(1)}{a(1)} log a(1) - \frac{a'(0)}{a(0)} log a(0) - \frac{1}{2} \cdot \frac{\big(a'(1)\big)^2}{\big(a(1)\big)^2} + \frac{1}{2} \cdot \frac{\big(a'(0)\big)^2}{\big(a(0)\big)^2} \right)^2.$$

განვიხილავთ $I_3(a) = \int_{-\infty}^{\infty} \big(a(t)\big)^s dt$ ფუნქციონალს $(s > 1)$, მაშინ $\varphi(t, x_0, x_1, \dots, x_m) = x_0^S$, $x_1 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$, როცა $n \to \infty$, მაშინ გვექნება

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = s^2\sigma^2\left(\int_0^1 a^{s-1}(t)dt\right)^2.$$

განხილულია შენონის ტიპის ენტროპიული ფუნქციონალი

$$I_4(a) = \int_{-\infty}^{\infty} a(t)loga(t)dt.$$

მაშინ გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \varphi(x_0) = x_0 log x_0,\ t \in [0,1] => a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. ამ შემთხვევაში თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$, როცა $n \to \infty$, გვექნება კრებადობა

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = \sigma^2\left(\int_0^1 a(t)\big(1 + loga(t)\big)dt\right)^2.$$

თითოეული შემთხვევისთვის დამტკიცებულია განმეორებითი ლოგარითმის თვისება. მოვიყვანოთ ერთ-ერთი მათგანი:

თეორემა 4: გასერ-მიულერის სქემაში, თუ h_n მიმდევრობა შერჩეულია ისე, რომ

$$R_n = o\left(\sqrt{\frac{loglogn}{n}}\right), \tag{16}$$

მაშინ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = r. \tag{17}$$

მომდევნო პარაგრაფში შესრულებულია რიცხვითი სიმულაციები, რომელიც გადმოცემულია ცხროლების და გრაფიკების საშუალებით. სიმულაციები ვიზუალურად ასაბუთებს თეორიულ შედეგებს.

ტექსტში ყველგან გამოყენებულია რამდენიმე ცნობილი დებულება. აქ მოგვყავს ამ დებულებების ფორმულირება მსჯელობისას ციტირების გასამარტივებლად. ეს დებულებები შეიძლება მოვიძიოთ მონოგრაფიაში [30].

მაკ-დაიარმიდის უტოლობა (McDiarmid inequality): ვთქვათ, $H(t_1, t_2, \dots, t_k)$ ისეთი ნამდვილი ფუნქციაა, რომ გარკვეული c_i-ებისთვის, ყოველი $i = 1, 2, \dots, k$-თვის სრულდება

$$|H(t_1, t_2, \dots, t_{i-1}, t_i, t_{i+1}, \dots, t_k) - H(t_1, t_2, \dots, t_{i-1}, t, t_{i+1}, \dots, t_k)| \le c_i \quad (\mathbf{18})$$

უტოლობები. ვთქვათ, $X_1, X_2, \dots, X_k$ არიან დამოუკიდებელი შემთხვევითი სიდიდეებია მნიშვნელობებით $H(t_1, t_2, \dots, t_k)$ ფუნქციის განსაზღვრის არიდან. მაშინ, ნებისმიერი დადებითი ε-თვის სრულდება უტოლობა

$$P\{|H(X_1, X_2, \dots, X_k) - EH(X_1, X_2, \dots, X_k)| > \varepsilon\} \le 2exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^{k} c_i^2}\right). \quad (\mathbf{19})$$

იენსენის უტოლობა (Jensen's inequality): ყოველი $\varphi(x)$ ამოზნექილი ფუნქციისათვის და ყოველი X შემთხვევითი სიდიდისათვის სამართლიანია შემდეგი უტოლობა:

$$\varphi(E[X]) \le E[\varphi(X)] \quad (\mathbf{20})$$

ბორელი-კანტელის ლემა (Borel-Cantelli lemma): ვთქვათ, A_n, $n = 1,2, \dots$ ხდომილობათა მიმდევრობაა რაღაც ალბათურ სივრცეზე. თუ

$$\sum_{n=1}^{\infty} P\{A_n\} < \infty,$$

მაშინ ალბათობა იმისა, რომ მოხდება უსასრულოდ ბევრი A_n, არის 0-ის ტოლი:

$$P\left\{\lim_{n\to\infty} sup\, A_n\right\} = 0.$$

აქ

$$\lim_{n\to\infty} sup\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\,.$$

ჰეფდინგის უტოლობა (Hoeffding's inequality): ვთქვათ, $X_1, X_2, \dots, X_n$ არიან ისეთი ნამდვილი დამოუკიდებელი შემთხვევითი სიდიდეები, რომ ყოველი i-სათვის X_i იღებს მნიშვნელობებს $[r_i,\ p_i]$ ინტერვალიდან. შემოვიღოთ აღნიშვნა

$$\sum_{i=1}^{n} X_i \equiv Y_n.$$

მაშინ ნებისმიერი დადებითი t-სათვის სამართლიანი იქნება შემდეგი უტოლობა:

$$P\{\left|Y - E|Y|\right| \ge t\} \le 2exp\left(-\frac{2t^2}{\sum_{k=1}^{n}(r_i - p_i)^2}\right). \quad (\mathbf{21})$$

ზემოთ, მოკლედ აღწერილი, შედეგები გამოქვეყნებულია შემდეგ სტატიებში [27-29].

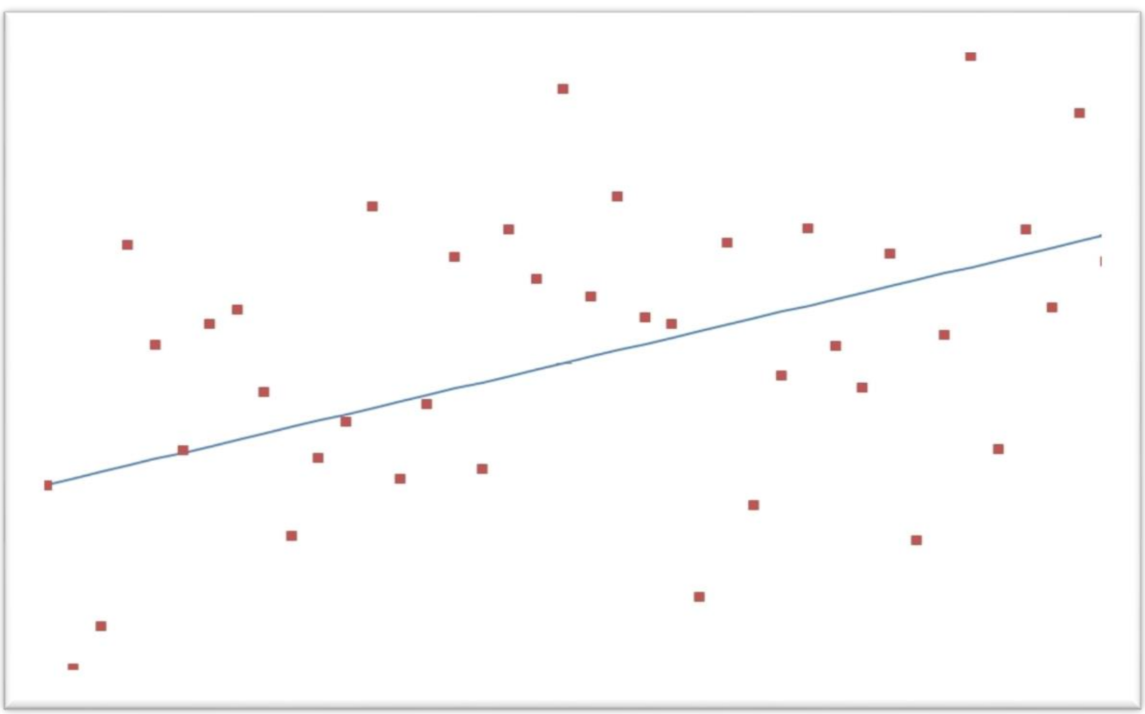
§1. რეგრესიის პრობლემის დასმა

რეგრესიის მოდელის განხილვა, გამოკვლევა და რეგრესიის ფუნქციის აგება ძალზე გავრცელებული სტატისტიკური პრობლემაა. რეგრესიის მოდელების შექმნა ძლიერი იარაღია დასაკვირვებელი ცვლადის მომავალი მნიშვნელობათა სიმრავლისა და საპროგნოზო მნიშვნელობების სანდოობის ინტერვალების შესაქმნელად, საზოგადოდ, მნიშვნელობების პროგნოზირებისათვის.

რეგრესიის ყველაზე მარტივი მოდელია წრფივი რეგრესიის მოდელი, რომელიც მოიცემა შემდეგი სახით:

$$Y(t) = \beta_0 + \beta_1 \cdot t + \varepsilon(t), \qquad (1.1)$$

სადაც $\varepsilon(t)$ არის ცდომილება და წარმოადგენს შემთხვევით სიდიდეს, რომლის საშუალო ნულია და დისპერსია მუდმივია და სასრული: $E\varepsilon(t) = 0,\ E\varepsilon^2(t) = \sigma^2 < \infty$; $Y(t)$ შემთხვევითი ფუნქციაა, რომლისთვისაც გაგვაჩნია $Y(t_1), Y(t_2), \dots, Y(t_n)$ დაკვირვებები და ამ დაკვირვებებისათვის ასაგებია რეგრესიის ფუნქცია. თუ ეს მოდელი კარგად აღწერს რეალობას, მაშინ შესაძლებელია $\beta_i, i = 1,2$-ს შეფასების გამოთვლა, დასკვნების გაკეთება და პროგნოზი. პირველ ნახაზზე გარკვეული დაკვირვებებისათვის აგებულია წრფივი რეგრესიის ფუნქცია.



ნახაზი 1.

რა ხდება როდესაც წრფივი რეგრესია არ არის განსახილველი პროცესის ადექვატური? ამ დროს ალტერნატიული მიდგომაა არაწრფივი რეგრესიის მოდელი, რომელიც ჩაიწერება შემდეგნაირად:

$$Y(t) = a(t) + \varepsilon(t), \qquad t \in [0, 1]. \tag{1.2}$$

ცხადია, ამ ტოლობაში უცნობი რეგრესიის $a(t)$ ფუნქცია არის $Y(t)$-ს მათემატიკური ლოდინი

$$a(t) = E\big(Y(t)\big).$$

$\varepsilon(\cdot)$ არის შეშფოთება ნულოვანი საშუალოთი და სასრული, არა აუცილებლად მუდმივი, დისპერსიით, $E\varepsilon(t) = 0,\ E\varepsilon^2(t) = \sigma^2(t) < \infty$; $Y(t)$ დაკვირვებადი შემთხვევითი ფუნქციაა. დავუშვათ, გვაქვს n წერტილი:

$$0 \le t_1 \le t_2 \le \cdots \le t_n \le 1,$$

სადაც თითოეული $t_k,\ k = 1, 2, \dots, n$ დამოკიდებულია n-ზე ისე რომ $max_i|t_i - t_{i-1}| = O\left(\frac{1}{n}\right)$. ამ წერტილებში გაგვაჩნია დაკვირვებები:

$$Y(t_1), Y(t_2), \dots, Y(t_n).$$

ამ დაკვირვებების საშუალებით ხდება უცნობი რეგრესიის $a(t)$ ფუნქციის შეფასება.

(1.2) მოდელის კვლევა, პირობების მოძებნა, რომელსაც უნდა აკმაყოფილებდეს $\varepsilon(t)$ შეშფოთება და ფუნქციების იმ კლასის დადგენა, რომელშიც უნდა ვეძებოთ $a(t)$ ფუნქციის სტატისტიკური შეფასება, წარმოადგენს არაპარამეტრული სტატისტიკის ძალზე მნიშვნელოვან ამოცანას. ცნობილია ამ მოდელის შესწავლის და რეგრესიის ფუნქციის აგების რამდენიმე მიდგომა. აღვწეროთ ჩვენი კვლევის ინტერესებში შემავალი მიდგომები.

უცნობი რეგრესიის $a(t)$ ფუნქციის ერთ-ერთი შეფასება შემოღებულ იქნა პრისტლისა და ჩაოს მიერ (იხ. [10-11]) და განსაზღვრულია შემდეგნაირად:

$$\hat{a}_n(t) = \sum_{i=1}^{n} W\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i), \tag{1.3}$$

სადაც, $\{h_n,\ n = 1,2, \dots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი - ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა. მათ მიერ განსაზღვრულია რეგრესიის ფუნქციის k-ური რიგის $a^{(k)}(t)$ წარმოებულის შეფასებაც შემდეგნაირად

$$\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} W^{(k)}\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i) \tag{1.4}$$

ყოველი $k = 0, 1, 2, \dots, m$-თვის. აქ და ქვემოთ ჩავთვლით, რომ $\hat{a}_n^{(0)}(t) \doteqdot \hat{a}_n(t)$.

რეგრესიის $a(t)$ ფუნქციის კიდევ ერთი შეფასება შემოღებულ იქნა გასერისა და მიულერის მიერ (იხ. [12-13]) შემდეგნაირად:

$$\hat{a}_n(t) = \frac{1}{h_n} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i), \qquad (\mathbf{1.5})$$

სადაც, $0 = s_1 \leq s_2 \leq \cdots \leq s_n = 1,\ t_i \leq s_i \leq t_{i+1},\ i = 1, 2, 3, \ldots, n-1$ და $max_i|s_i - s_{i-1}| = O\left(\frac{1}{n}\right)$; $\{h_n,\ n = 1,2,\ldots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი - ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა. ამ შეფასებისთვისაც მათ მიერვეა განსაზღვრული უცნობი რეგრესიის ფუნქციის k-ური რიგის $a^{(k)}(t)$ წარმოებულის შეფასებაც შემდეგნაირად:

$$\hat{a}_n^{(k)}(t) = \frac{1}{{h_n}^{k+1}} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W^{(k)}\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i), \qquad (\mathbf{1.6})$$

ყოველი $k = 0, 1, 2, \ldots, m$-თვის.

ცნობილია, კიდევ ერთი შეფასება უცნობი რეგრესიის $a(t)$ ფუნქციისათვის. ეს შეფასება შემოღებულ იქნა ნადარაისა და ვატსონის მიერ (იხ. [8-9]) და განსაზღვრულია შემდეგნაირად:

$$\hat{a}_n(t) = \sum_{i=1}^{n} \frac{W\left(\frac{t-t_i}{h_n}\right)}{\sum_{k=1}^{n} W\left(\frac{t-t_k}{h_n}\right)} \cdot Y(t_i), \qquad (\mathbf{1.7})$$

სადაც, $\{h_n,\ n = 1,2,\ldots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი - ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა.

სიმარტივისთვის შემოვიღოთ აღნიშვნა:

$$\frac{W\left(\frac{t-t_i}{h_n}\right)}{\sum_{k=1}^{n} W\left(\frac{t-t_k}{h_n}\right)} \equiv \alpha(t, t_i, h_n).$$

ნადარაია-ვატსონის შეფასება გადაიწერება შემდეგნაირად:

$$\hat{a}_n(t) = \sum_{i=1}^{n} \alpha(t, t_i, h_n) \cdot Y(t_i).$$

სამართლიანია

$$\sum_{i=1}^{n} \alpha(t, t_i, h_n) = 1$$

და

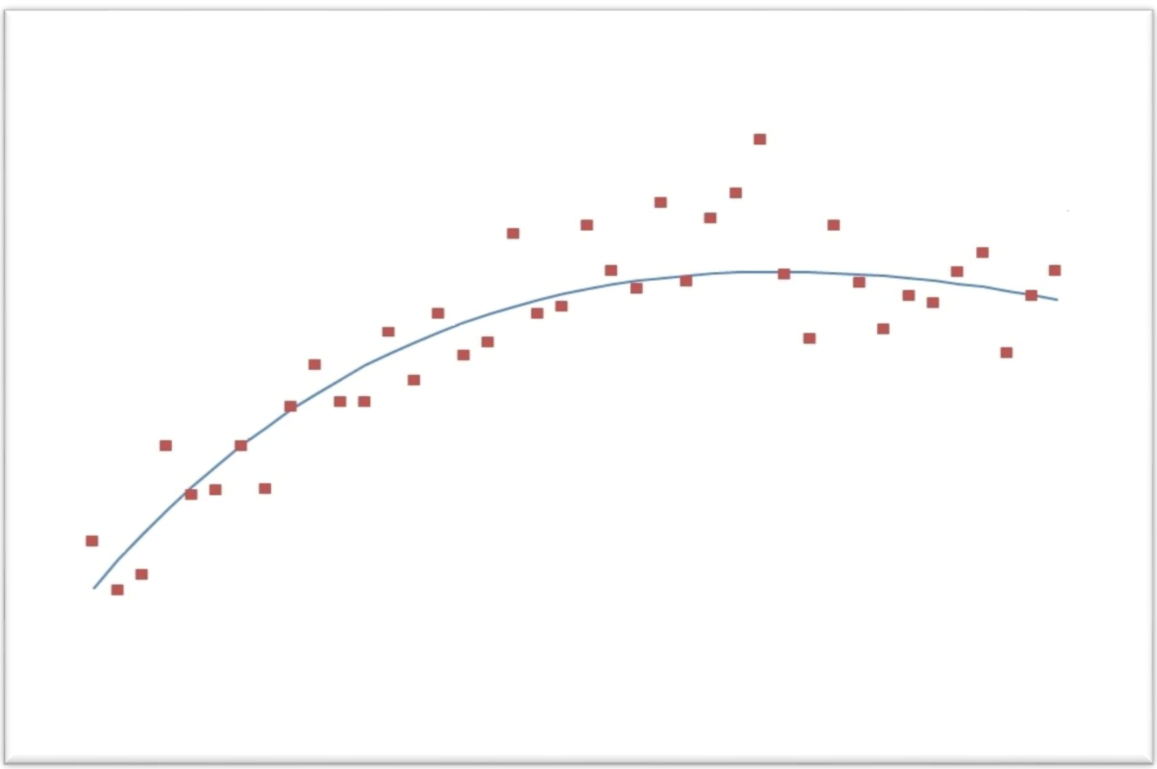
$$\sum_{i=1}^{n} \alpha^{(k)}(t, t_i, h_n) = 0$$

ტოლობები (აქ და შემდეგში ყველგან $\alpha^{(k)}(t, t_i, h_n)$-ში იგულისხმება k რიგის წარმოებული t ცვლადის მიმართ).

ასეთი აღნიშვნით ნადარაია-ვატსონის რეგრესიის ფუნქციის k-ური რიგის $a^{(k)}(t)$ წარმოებულის შეფასება შემდეგნაირად ჩაიწერება

$$\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} \alpha^{(k)}(t, t_i, h_n) \cdot Y(t_i). \tag{1.7}$$

მეორე ნახაზზე გარკვეული დაკვირვებებისათვის აგებულია არაწრფივი რეგრესიის ფუნქცია.



ნახაზი 2.

ყველა აღწერილ შემთხვევაში დადგენილია შეფასებათა ძალდებულობისა და სხვა ასიმპტოტური თვისება.

§2. პრისტლი-ჩაოს შეფასებები და ფუნქციონალების ასიმპტოტური თვისებები

ვთქვათ, $a(t)$ არის რეგრესიის ფუნქცია. კერძო შემთხვევებში შეიძლება განხილული იქნეს შემდეგი ფუნქციონალები:

$$I_1(a) = \int_{-\infty}^{\infty} a^2(t)dt, \quad I_2(a) = \int_{-\infty}^{\infty} \frac{\left(a'(t)\right)^2}{a(t)} dt,$$

$$I_3(a) = \int_{-\infty}^{\infty} \left(a(t)\right)^s dt, \quad I_4(a) = \int_{-\infty}^{\infty} a(t) log a(t) dt.$$

ეს ფუნქციონალები ცალკე შესწავლის საგანია სხვადასხვა კვლევებისათვის. ჩვენ აქ გამოვიყენებთ მათთვის ერთიან მიდგომას, რომელიც დამყარებულია ე.წ. წარმოდგენის თეორემაზე.

განვიხილოთ პრისტლი-ჩაოს რეგრესიის მოდელი:

$$Y(t) = a(t) + \varepsilon(t), \tag{2.1}$$

სადაც $t \in [0,1]$, $\varepsilon(\cdot)$ არის შეშფოთება ნულოვანი საშუალოთი და სასრული და მუდმივი დისპერსიით, $E\varepsilon(t) = 0$, $E\varepsilon^2(t) = \sigma^2 < \infty$; $Y(t)$ შემთხვევითი ფუნქციაა; $a(t)$ კი - უცნობი რეგრესიის ფუნქცია. დავუშვათ გვაქვს დაკვირვებათა n წერტილი:

$$0 \le t_1 \le t_2 \le \cdots \le t_n \le 1,$$

სადაც, თითოეული t_k, $k = 1, 2, \dots, n$ დამოკიდებულია n-ზე, ისე რომ $max_i |t_i - t_{i-1}| = O\left(\frac{1}{n}\right)$. ამ წერტილებში გაგვაჩნია დაკვირვებები: $Y(t_1), Y(t_2), \dots, Y(t_n)$.

უცნობი რეგრესიის $a(t)$ ფუნქციის ერთ-ერთი შეფასება შემოღებულ იქნა პრისტლისა და ჩაოს მიერ (იხ. [10-11]) და განსაზღვრულია შემდეგნაირად:

$$\hat{a}_n(t) = \sum_{i=1}^{n} W\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i), \tag{2.2}$$

სადაც, $\{h_n,\ n = 1,2, \dots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი - ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა. მათ მიერვე განსაზღვრულია რეგრესიის ფუნქციის k-ური რიგის $a^{(k)}(t)$ წარმოებულის შეფასებაც შემდეგნაირად

$$\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} W^{(k)}\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i) \qquad (\mathbf{2.3})$$

ყოველი $k = 0, 1, 2, \dots, m$-თვის. აქ და ქვემოთ ჩავთვლით, რომ $\hat{a}_n^{(0)}(t) \doteqdot \hat{a}_n(t)$.

ვთქვათ, $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია უწყვეტია, შემოსაზღვრულია და გლუვია. განვიხილოთ შემდეგი სახის ინტეგრალური ფუნქციონალი:

$$I(a) = \int_{-\infty}^{\infty} \varphi\left(t, a(t), a'(t), \dots, a^{(m)}(t)\right) dt. \qquad (\mathbf{2.4})$$

გვაქვს შერჩევა $(t_i, Y_i), i = 1, 2, \dots, n$. ეს ნიშნავს, რომ

$$Y_i = Y(t_i) = a(t_i) + \varepsilon(t_i). \qquad (\mathbf{2.5})$$

$I(a)$-ს შეფასებისთვის ვიყენებთ ე.წ. ”plug-in” შეფასებას. ანუ განვიხილავთ ფუნქციონალს:

$$I(\hat{a}_n) = \int_{-\infty}^{\infty} \varphi\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right) dt.$$

წარმოდგენის თეორემა

ამ ნაწილში მოვიყვანთ ”წარმოდგენის” თეორემას, საიდანაც მივიღებთ ჩვენთვის საინტერესო შედეგებს. ჩამოვაყალიბოთ პირობები განსახილველი სიდიდეებისათვის.

პირობები $\boldsymbol{a}$-სათვის:

$(a1)$ $a = a(t)$ განსაზღვრულია და უწყვეტია [0, 1]-ზე, იგი მნიშვნელობებს იღებს $[-\mathbb{k}, \mathbb{k}]$ ინტერვალიდან;

$(a2)$ $a = a(t)$ ფუნქციას აქვს უწყვეტი წარმოებულები m რიგის ჩათვლით;

$(a3)$ ნებისმიერი $i = 0, 1, \dots, m$-სათვის, $a^{(i)}(\cdot) \in L_1([0, 1])$.

პირობები $\boldsymbol{\varepsilon_k = \varepsilon(t_k)}$-სათვის:

$(\varepsilon1)$ ε_k, $k = 1, 2, \dots$ დამოუკიდებელი და ერთნაირად განაწილებული შემთხვევითი სიდიდეებია;

$(\varepsilon2)$ $E\varepsilon_k = 0$, $D\varepsilon_k^2 = \sigma^2 < \infty$.

სიმარტივისთვის შემოვიღოთ აღნიშვნები $\varphi = \varphi(x, x_0, \dots, x_m) \in C_b^2(R^{m+2})$ ფუნქციისათვის

$$\frac{\partial \varphi}{\partial x_i} = \varphi_{(i)}, \ i = 0, 1, \dots, m \text{ და } \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \varphi_{(ij)}, \ i, j = 0, 1, \dots, m.$$

პირობები $\boldsymbol{\varphi}$-სათვის:

(φ) $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია უწყვეტია, შემოსაზღვრულია, ინტეგრებადია და გააჩნია მეორე რიგის ჩათვლით უწყვეტი წარმოებულები, რომელიღაც ღია, ამოზნექილ A არეში, რომელსაც შეიცავს $\mathbb{R} \times [-\Bbbk, \Bbbk]^{m+1}$ სიმრავლე; φ-ის პირველი და მეორე წარმოებულები თანაბრად შემოსაზღვრულია A არეში $C_\varphi > 0$ მუდმივით. ამ პირობების თანახმად, φ ფუნქციისათვის ნებისმიერი $i, j = 0, 1, \dots, m$-თვის გვაქვს:

$$sup\{|\varphi_{(i,j)}|(s, s_0, s_1, \dots, s_m): (s, s_0, s_1, \dots, s_m) \in A\} \le C_\varphi. \qquad \mathbf{(2.6)}$$

პირობები $\boldsymbol{W}$-სათვის:

$(W1)$ $\int_{-\infty}^{\infty} W(t)\, dt = 1$;

$(W2)$ $W(t)$ ფუნქციას გააჩნია კომპაქტური მატარებელი $[-\tau,\ \tau]$ და ის უწყვეტად წარმოებადია $m \ge 1$ რიგის ჩათვლით. მაშინ უწყვეტი $W^{(i)}(t)$, $i = 0, 1, \dots, m$ ფუნქციებისთვის $W^{(i)}(-\tau) = W^{(i)}(\tau) = 0$; აღვნიშნოთ $C_W > 0$ მუდმივი, რომლისთვისაც

$$sup_{t \in R}|W^{(i)}(t)| \le C_W < \infty, \qquad i = 0, 1, \dots, m;$$

ცხადია, ნებისმიერი $i = 0, 1, \dots, m$-თვის $W^{(i)} \in L_1([-\tau,\ \tau])$.

პირობები $\boldsymbol{h_n}$-სათვის:

(h_n) $\quad \dfrac{\sqrt{\max(|logh_n|; loglogn)}}{\sqrt{n} h_n^{0,5+m}} \to 0,$ როცა $n \to \infty$.

$a_n(t)$-თი ავღნიშნოთ $\hat{a}_n(t)$-ს მათემატიკური ლოდინი:

$$a_n(t) = E\hat{a}_n(t) = E\left(\sum_{i=1}^{n} W\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot Y(t_i)\right) = \sum_{i=1}^{n} W\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot a(t_i).$$

ანალოგიურად,

$$a_n^{(k)}(t) = E\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} W^{(k)}\left(\frac{t - t_i}{h_n}\right) \cdot \frac{t_i - t_{i-1}}{h_n} \cdot a(t_i).$$

ვაჩვენოთ, რომ არსებობს და სასრულია გამოსახულებები: $I(a), I(a_n)$ და $I(\hat{a}_n)$.

ტეილორის ფორმულის თანახმად ნებისმიერი $(s, s_0, s_1, \dots, S_m) \in A$ წერტილისთვის და რაღაც $\tilde{s}_i \in A$-თვის შეგვიძლია დავწეროთ

$$|\varphi|(s, s_0, s_1, \dots, S_m) = \left|\sum_{i=0}^{m} \varphi_{(i)}(s, 0,0, \dots,0) s_i + \frac{1}{2}\sum_{i,j=0}^{m} \varphi_{(i,j)}(s, \tilde{s}_0, \tilde{s}_1, \dots, \tilde{s}_m) s_i s_j\right|.$$

შესაბამისად, მოიძებნება მუდმივი C ისეთი, რომ

$$|\varphi|(s, s_0, s_1, \dots, S_m) \le C\left(\sum_{i=0}^{m}|s_i| + \sum_{i=0}^{m}|s_i|^2\right).$$

აქედან გამომდინარეობს, რომ ნებისმიერი შემოსაზღვრული ზომადი $f_0(t), f_1(t), \dots, f_m(t)$ ფუნქციებისთვის $L_1(R)$-დან გვექნება

$$\int_{-\infty}^{\infty}|\varphi|\big(t, f_0(t), f_1(t), \dots, f_m(t)\big)dt < \infty. \qquad (\mathbf{2.7})$$

ასე, რომ $I(a)$ არსებობს.

პირობები, რომლებიც შემოღებულია W-თვის გვაძლევს მის შემოსაზღვრულობას და ინტეგრებადობას. $(W2)$ პირობიდან და (2.6) და (2.7) ფორმულებიდან გამომდინარეობს რომ ორივე გამოსახულება $I(a_n)$ და $I(\hat{a}_n)$, სასრულია ნებისმიერი $n \in \mathbb{N}$-თვის. ტეილორის ფორმულის თანახმად შეგვიძლია დავწეროთ

$$I(\hat{a}_n) - I(a_n) = S_n(h_n) + R_n, \qquad (\mathbf{2.8})$$

სადაც,

$$S_n(h_n) = \sum_{i=0}^{m}\int_{0}^{1}\varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)dt. \qquad (\mathbf{2.9})$$

შევნიშნოთ, რომ ნებისმიერი $h_n > 0$-თვის $s_n(h_n)$ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების ჯამს. R_n ნაშთით წევრს კი აქვს შემდეგი სახე:

$$R_n = \frac{1}{2}\cdot\sum_{i,j=1}^{m}\int_{0}^{1}\varphi_{(ij)}\left(\tilde{b}_m(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)\left(\hat{a}_n^{(j)}(t) - a_n^{(j)}(t)\right)dt. \qquad (\mathbf{2.10})$$

ამ ფორმულაში $\tilde{b}_m(t)$ არის $\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right)$ და $\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)$ წერტილების შემაერთებელი წრფე.

$(\mathbf{2.7})$ და $(\mathbf{2.10})$ ფორმულებიდან ცხადია, შეგვიძლია R_n ნაშთითი წევრის შეფასება:

$$|R_n| \le C_\varphi \int_{0}^{1}\sum_{i=0}^{m}\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)^2 dt. \qquad (\mathbf{2.11})$$

ვთქვათ W_m^2 სობოლევის სივრცეა, ფუნქციებისა რომელთაც გააჩნიათ კვადრატთან ერთად ინტეგრებადი, უწყვეტი და შემოსაზღვრული მეორე რიგის წარმოებული. ამ სივრცეში ნორმა და სკალარული ნამრავლი ასე განისაზღვრება

$$\|g\|_m = \sqrt{\sum_{i=0}^{m} \int_0^1 |g^{(i)}(t)|^2 dt}, \quad \langle g_1, g_2 \rangle_m = \sqrt{\sum_{i=0}^{m} \int_0^1 \left(g_1^{(i)}(t) \cdot g_2^{(i)}(t) \right) dt}.$$

შემოვიღოთ აღნიშვნა $r_n(m) = \|\hat{a}_n - a_n\|_m^2$. ვხედავთ, რომ

$$|R_n| \leq C_\varphi r_n(m). \qquad \mathbf{(2.12)}$$

ვთქვათ,

$$U_k = U_k(t) = W\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot [Y(t_k) - a(t_k)], \qquad k = 1, 2, \dots, n,$$

სადაც $a(t_k) = EY(t_k)$. მაშინ

$$\sum_{k=1}^{n} U_k = \sum_{k=1}^{n} W\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot [Y(t_k) - a(t_k)] =$$

$$= \sum_{k=1}^{n} W\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot Y(t_k) - \sum_{k=1}^{n} W\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot a(t_k) = \hat{a}_n(t) - a_n(t).$$

შესაბამისად,

$$r_n(m) = \left\| \sum_{k=1}^{n} U_k \right\|_m^2. \qquad \mathbf{(2.13)}$$

ნებისმიერი $k = 1, 2, \dots, n$-თვის შევაფასოთ $(\mathbf{2.13})$ ფორმულის თითოეული U_k შესაკრები $\|\cdot\|_m$ ნორმით. გვექნება

$$\|U_k\|_m = \left(\sum_{i=0}^{m} \int_0^1 \left| U_k^{(i)}(t) \right|^2 dt \right)^{\frac{1}{2}} = \left(\sum_{i=0}^{m} \int_0^1 \left| \left(W\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot [Y(t_k) - a(t_k)] \right)^{(i)} \right|^2 dt \right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=0}^{m} \int_0^1 \left| \frac{1}{h_n^i} \cdot W^{(i)}\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot [Y(t_k) - a(t_k)] \right|^2 dt \right)^{\frac{1}{2}} =$$

$$= (t_k - t_{k-1}) \cdot |Y(t_k) - a(t_k)| \left(\sum_{i=0}^{m} \int_0^1 \left| \frac{1}{h_n^{i+1}} \cdot W^{(i)}\left(\frac{t - t_k}{h_n}\right) \right|^2 dt \right)^{\frac{1}{2}} \leq \frac{|\varepsilon_k| C_W}{n} \cdot \left(\sum_{i=0}^{m} \int_0^1 \left| \frac{1}{h_n^{i+1}} \right|^2 dt \right)^{\frac{1}{2}}$$

$$= \frac{|\varepsilon_k| C_W}{n} \cdot \left(\sum_{i=0}^{m} \frac{1}{h_n^{2i+2}} \right)^{\frac{1}{2}} = \frac{|\varepsilon_k| C_W}{n} \cdot \left(\frac{1 - h_n^{2m+2}}{h_n^{2m+2}(1 - h_n^2)} \right)^{\frac{1}{2}} \leq L \cdot \frac{1}{n h_n^{m+1}} := M_m \sim O\left(\frac{1}{n h_n^{m+1}} \right) \qquad \mathbf{(2.14)}$$

საკმარისად დიდი $L > 0$-თვის.

$r_n(m)$-ის შეფასებისთვის გამოვიყენოთ მაკ-დაიარმიდის უტოლობა

$$H(U_1, U_2, \ldots, U_n) = \left\| \sum_{k=1}^{n} U_k \right\|_m$$

ფუნქციისთვის. მივიღებთ

$$\left| H(U_1, U_2, \ldots, U_{i-1}, U_x, U_{i+1}, \ldots, U_n) - H\big(U_1, U_2, \ldots, U_{i-1}, U_y, U_{i+1}, \ldots, U_n\big) \right|$$

$$= \left| \left\| \sum_{k=1, k\neq i}^{n} U_k + U_x \right\|_m - \left\| \sum_{k=1, k\neq i}^{n} U_k + U_y \right\|_m \right|$$

$$\leq \left\| \left(\sum_{k=1, k\neq i}^{n} U_k + U_x \right) - \left(\sum_{k=1, k\neq i}^{n} U_k + U_y \right) \right\|_m$$

$$\leq \left\| U_x - U_y \right\|_m \leq \|U_x\|_m + \left\|U_y\right\|_m \leq 2M_m$$

მაშასადამე (**18**) გამოსახულების მიხედვით, ყოველი $c_k \equiv 2M_m,\ k = 1, \ldots, n$ (**19**) უტოლობის და ($\mathbf{2.14}$) შეფასების თანახმად კი ნებისმიერი $\delta > 0$-თვის გვექნება

$$P\left\{ \left| \left\| \sum_{k=1}^{n} U_k \right\|_m - E \left\| \sum_{k=1}^{n} U_k \right\|_m \right| > \delta \right\} \leq 2exp\left(-\frac{2\delta^2}{\sum_{k=1}^{n} 4M_m^2} \right) = 2exp\left(-\frac{2\delta^2}{n \cdot 4M_m^2} \right)$$
$$= 2exp\left(-\frac{2\delta^2 n^2 h_n^{2m+2}}{n \cdot 4L^2} \right) = 2exp\left(-\frac{\delta^2 n h_n^{2m+2}}{2L^2} \right).$$

აქ ჩავსვათ

$$\delta = \frac{2L\sqrt{logn}}{\sqrt{n} h_n^{m+1}}$$

და მივიღებთ

$$P\left\{ \left| \left\| \sum_{k=1}^{n} U_k \right\|_m - E \left\| \sum_{k=1}^{n} U_k \right\|_m \right| > \delta \right\} \leq 2exp\left(-\frac{4L^2 logn \cdot n h_n^{2m+2}}{n h_n^{2m+2} \cdot 2L^2} \right) = 2exp(-2logn) = \frac{2}{n^2}.$$

გამოვიყენოთ ბორელი-კანტელის ლემა. მივიღებთ

$$\left\| \sum_{k=1}^{n} U_k \right\|_m = E \left\| \sum_{k=1}^{n} U_k \right\|_m + O\left(\frac{\sqrt{logn}}{\sqrt{n} h_n^{m+1}} \right). \qquad (\mathbf{2.15})$$

იენსენის უტოლობის თანახმად

$$\left(E\left\|\sum_{k=1}^{n}U_k\right\|_m\right)^2 \leq E\left\|\sum_{k=1}^{n}U_k\right\|_m^2 = E\sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1\left|\frac{1}{h_n^i}\cdot W^{(i)}\left(\frac{t-t_k}{h_n}\right)\cdot\frac{t_k-t_{k-1}}{h_n}\cdot[Y(t_k)-a(t_k)]\right|^2 dt =$$

$$=\sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 E\left|\frac{1}{h_n^i}\cdot W^{(i)}\left(\frac{t-t_k}{h_n}\right)\cdot\frac{t_k-t_{k-1}}{h_n}\cdot[Y(t_k)-a(t_k)]\right|^2 dt \leq$$

$$\leq\frac{C_W^2}{n}\cdot\sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1\frac{1}{h_n^{2i+2}}\cdot E[Y(t_k)-a(t_k)]^2 dt \leq \frac{C_W^2\sigma^2}{n}\cdot\sum_{i=0}^{m}\frac{1}{h_n^{2i+2}} \leq$$

$$\leq\frac{C_W^2\sigma^2}{n}\cdot\frac{1-h_n^{2m+2}}{h_n^{2m+2}(1-h_n^2)} \leq K\cdot\frac{1}{nh_n^{2m+2}}, \qquad (\mathbf{2.16})$$

$(\mathbf{2.12})$, $(\mathbf{2.13})$, $(\mathbf{2.15})$ და $(\mathbf{2.16})$ ფორმულებიდან შეგვიძლია დავასკვნათ, რომ $R_n = O\left(\frac{logn}{nh_n^{2m+2}}\right)$.

ასე, რომ სამართლიანია შემდეგი თეორემა.

თეორემა 2.1. ვთქვათ, სრულდება $(a1)-(a3), (\varepsilon1)-(\varepsilon3), (\varphi), (W1)-(W2)$ და (h) პირობები. მაშინ, სამართლიანია $(\mathbf{2.8})$ ტოლობა, სადაც ნაშთით წევრს ალბათობით 1 აქვს შემდეგი რიგი

$$R_n = O\left(\frac{logn}{nh_n^{2m+2}}\right).$$

ძალდებულობა

ამ ნაწილში გამოვიყენებთ თეორემა 2.1-ს და დავამტკიცებთ $I(\hat{a}_n)$ შეფასების მკაცრად ძალდებულობას.

თეორემა 2.2. ვთქვათ სრულდება თეორემა 2.1-ის პირობები. დადებითი $(h_n)_{n=1}^{\infty}$ მიმდევრობა შერჩეულია ისე, რომ

$$\frac{logn}{nh_n^{2m+2}} \to 0 \text{ როცა } n\to\infty. \qquad (\mathbf{2.18})$$

მაშინ ალბათობით 1

$$I(\hat{a}_n) \to I(a) \qquad (\mathbf{2.19})$$

დამტკიცება: თეორემა 1-ის და $(\mathbf{2.8})$ ფორმულის თანახმად

$$I(\hat{a}_n) - I(a_n) = S_n(h_n) + R_n, \qquad (\mathbf{2.20})$$

სადაც $R_n = o(1)$ და

$$S_n(h_n) = \sum_{i=0}^{m} \int_0^1 \varphi_{(i)} \left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) \left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right) dt.$$

$(a1)$ პირობიდან:

$$\left\{\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) : t \in [0,1]\right\} \subset [0,1] \times [-\Bbbk, \Bbbk]^{m+1}.$$

ამიტომ, (φ) პირობის გათვალისწინებით, მოიძებნება ისეთი $C_\varphi > 0$ მუდმივი, რომ

$$sup\left\{\left|\varphi_{(i)}\right|(t, t_0, t_1, \dots, t_m) : (t, t_0, t_1, \dots, t_m) \in [0,1] \times [-\Bbbk, \Bbbk]^{m+1}\right\} \le C_\varphi.$$

შეგვიძლია დავწეროთ:

$$ES_n(h_n) = 0$$

$$DS_n(h_n) = ES_n^2(h_n) \le C_\varphi^2 \cdot \sum_{i=0}^{m} \int_0^1 E\left[\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right] dt$$

$$\le C_\varphi^2 \cdot \sum_{i=0}^{m} \int_0^1 \left|\sum_{k=0}^{n} \frac{1}{h_n^i} \cdot W^{(i)}\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot E[Y(t_k) - a(t_k)]\right|^2 dt$$

$$\le C_\varphi^2 C_W^2 \sigma^2 \cdot \sum_{k=0}^{n} (t_k - t_{k-1})^2 \cdot \sum_{u=0}^{m} \left(\frac{1}{h_n^{i+1}}\right)^2 \sim C \cdot \frac{1}{nh_n^{2m+2}} \to 0 \qquad (\mathbf{2.21})$$

რადგან $\frac{logn}{nh_n^{2m+2}} \to 0$ და ამიტომ $S_n(h_n) \to 0$ როცა $n \to \infty$.

$$Ea_n^{(k)}(t) = \int_{-\tau}^{\tau} W(u) a^{(k)}(t)\,(t - uh_n) du + O\left(\frac{1}{nh_n^k}\right).$$

აქედან შეიძლება დავასკვნათ:

i) $(\mathbf{2.17})$ ფორმულიდან გამომდინარე ნულისკენ კრებადია $\frac{1}{nh_n^k}$, ნებისმიერი $k = 0,1, \dots, m$-თვის;

ii) როცა $n \to \infty$ გვაქვს $Ea_n^{(k)}(t) \to a^{(k)}(t)$.

შევაჯამოთ ნათქვამი და მივიღებთ, რომ როცა $n \to \infty$ გვექნება

$$I(a_n) = \int_0^1 \varphi\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \to \int_0^1 \varphi\left(t, a(t), a'(t), \dots, a^{(m)}(t)\right) dt = I(a).$$

ასე, რომ $I(\hat{a}_n) - I(a_n) = o(1)$ ანუ $I(\hat{a}_n) - I(a) \to 0$. თეორემა დამტკიცებულია.

ცენტრალური ზღვარითი თეორემა

ჩვენი წარმოდგენის თეორემის გამოყენებით შეიძლება მივიღოთ ზღვარითი განაწილების თვისება ინტეგრალური ფუნქციონალისთვის

$$I(\hat{a}_n) = \int_0^1 \varphi\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right) dt.$$

განვიხილოთ სხვაობა

$$I(\hat{a}_n) - I(a_n) = s_n(h_n) + R_n \qquad (\mathbf{2.8})$$

სადაც ნებისმიერი $h_n > 0$-თვის, $s_n(h_n)$ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების ჯამს.

$$S_n(h_n) = \sum_{i=0}^{m} \int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right) dt. \qquad (\mathbf{2.9})$$

R_n ნაშთით წევრს კი აქვს შემდეგი სახე:

$$R_n = \frac{1}{2} \cdot \sum_{i,j=1}^{m} \int_0^1 \varphi_{(ij)}\left(\tilde{b}_m(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)\left(\hat{a}_n^{(j)}(t) - a_n^{(j)}(t)\right) dt. \qquad (\mathbf{2.10})$$

ცხადია, რომ

$$ES_n(h_n) = 0 \text{ და } ER_n \to 0 \text{ როცა } n \to \infty. \qquad (\mathbf{2.23})$$

გარდა ამისა

$$E\big(S_n(h_n)\big)^2 = \sigma^2 \sum_{i=0}^{m} \left(\int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \right)^2 \qquad (\mathbf{2.24})$$

და $\mathrm{Var} R_n \to 0$ როცა $n \to \infty$.

ამ პირობებში დავამტკიცოთ, რომ $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)$ სიდიდე ასიმპტოტურად ნორმალურია და გამოვთვალოთ ზღვარითი ვარიაცია. ($\mathbf{2.8}$), ($\mathbf{2.23}$) და ($\mathbf{2.24}$) ფორმულების საშუალებით უნდა ვაჩვენოთ $\sqrt{n}S_n(h_n)$ სიდიდის ასიმპტოტურად ნორმალურობა.

ამისთვის საკმარისია ეს თვისება შევისწავლოთ შემდეგი სიდიდისთვის:

$$d_k = Y(t_k) \cdot \sum_{i=0}^{m} \frac{1}{h_n^i} \cdot \int_0^1 W^{(i)}\left(\frac{t - t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \qquad (\mathbf{2.25})$$

ადვილი შესამჩნევია, რომ

$$Ed_k = a(t_k) \cdot \sum_{i=0}^{m} \frac{1}{h_n^i} \cdot \int_0^1 W^{(i)}\left(\frac{t-t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \qquad (\mathbf{2.26})$$

ანუ ვიხილავთ დამოუკიდებელი შემთხვევითი სიდიდეების მიმდევრობას:

$$f_n(k) = \alpha(n,k)\big(Y(t_k) - a(t_k)\big) = \alpha(n,k)\varepsilon_k$$

სადაც

$$\alpha(n,k) = \sum_{i=0}^{m} \frac{1}{h_n^i} \cdot \int_0^1 W^{(i)}\left(\frac{t-t_k}{h_n}\right) \cdot \frac{t_k - t_{k-1}}{h_n} \cdot \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt$$

განვიხილოთ ჯამი

$$S_n(h_n) = \sum_{k=1}^{n} \alpha(n,k)\varepsilon_k.$$

ვთქვათ, $F_{k,n}$ არის $\alpha(n,k)\varepsilon_k$ შემთხვევითი სიდიდეების ალბათური განაწილების ფუნქცია, F_ε – კი ε_k შემთხვევითი სიდიდეების ალბათური განაწილების ფუნქცია.

ლინდებერგის პირობას აქვს შემდეგი სახე $\forall\delta > 0$, $lim_{n\to\infty} L_n(\delta) = 0$ სადაც

$$L_n(\delta) = \frac{\sum_{j=1}^{n} \int x^2 J\left(|x| \geq \delta\sigma\left(\sum_{k=1}^{n} \alpha^2(n,k)\right)^{1/2}\right) dF_{k,n}(x)}{\sigma^2 \sum_{k=1}^{n} \alpha^2(n,k)}$$

სადაც $J(A)$ წარმოადგენს A სიმრავლის ინდიკატორ ფუნქციას. ადვილი დასანახია, რომ

$$L_n(\delta) \leq \frac{1}{\sigma^2} max_{1\leq j\leq n} \int x^2 J\big(|x| \geq \delta\sigma v(n,j)\big) dF_\varepsilon,$$

სადაც

$$v(n,j) = \frac{|\alpha(n,j)|}{\left(\sum_{j=1}^{n} \alpha^2(n,j)\right)^{1/2}}.$$

საჩვენებელი რჩება, რომ $max_{1\leq j\leq n} v(n,j) \to 0$ როცა $n \to \infty$. მაგრამ რადგანაც

$$max_{1\leq j\leq n} |\alpha(n,j)| = O\left(\frac{1}{n h_n^{m+1}}\right),$$

ანუ

$$max_{1\le j\le n} v(n,j) = O\left(\frac{1}{\sqrt{n}}\right).$$

ასე, რომ ლინდერბერგის პირობა სრულდება და შეგვიძლია ჩამოვაყალიბოთ

თეორემა 2.3: ვთქვათ სრულდება თეორემა 1-ის პირობები. მაშინ, თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$ როცა $n \to \infty$, სამართლიანია $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2)$. სადაც

$$r^2 = \sigma^2 \cdot \sum_{i=0}^{m}\left(\int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt\right)^2.$$

ზოგიერთი გამოყენება

განვიხილოთ ინტეგრალური ფუნქციონალი

$$I_1(a) = \int_0^1 a^2(t)\, dt$$

ამ შემთხვევაში $\varphi(t, x_0, x_1, \dots, x_m) = x_0^2$, როცა $x_0 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. მაშინ თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$ როცა $n \to \infty$, მაშინ

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = 4\sigma^2\left(\int_0^1 a(t)dt\right)^2.$$

განვიხილოთ ფიშერის ინფორმაციული ინტეგრალი.

$$I_2(a) = \int_0^1 \frac{\big(a'(t)\big)^2}{a(t)} dt.$$

ამ ფუნქციონალისთვის გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \frac{x_1^2}{x_2}$, $t \in [0,1] => a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \ge \mathbb{b}$. თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$ როცა $n \to \infty$, მაშინ

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = \sigma^2 \cdot \left(-\int_0^1 \left(\frac{\left(a'(t)\right)^2}{\left(a(t)\right)^2} - \frac{2a'(t)}{a(t)} \right) dt \right)^2$$

$$= \sigma^2 \left(\frac{a'(1)}{a(1)} loga(1) - \frac{a'(0)}{a(0)} loga(0) - \frac{1}{2} \cdot \frac{\left(a'(1)\right)^2}{\left(a(1)\right)^2} + \frac{1}{2} \cdot \frac{\left(a'(0)\right)^2}{\left(a(0)\right)^2} \right)^2.$$

ახლა განვიხილოთ

$$I_3(a) = \int_{-\infty}^{\infty} \left(a(t)\right)^s dt$$

ფუნქციონალი $(s > 1)$. მაშინ $\varphi(t, x_0, x_1, \dots, x_m) = x_0^S$, როცა $x_1 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. თუ $h_n \to 0$ და $\sqrt{n} h_n^{2m+2} \to \infty$ როცა $n \to \infty$, მაშინ

$$\sqrt{n}\left(I(\hat{a}_n) - I(a_n)\right) \to_d N(0, r^2),$$

სადაც

$$r^2 = s^2 \sigma^2 \left(\int_0^1 a^{s-1}(t) dt \right)^2.$$

ბოლოს განვიხილოთ შენონის ენტროპიის ფუნქციონალი

$$I_4(a) = \int_{-\infty}^{\infty} a(t) loga(t) dt$$

მაშინ რომელიღაც საკმარისად დიდი $b \geq \mathbb{k} > 0$-თვის თუ $0 < x_0 \leq b$ გვაქვს

$$\varphi(t, x_0, x_1, \dots, x_m) = \varphi(x_0) = x_0 log x_0.$$

განვსაზღვროთ φ ფუნქცია 0-მდე, $\varphi(x) = 0$ როცა $-b \leq x \leq 0$. მაშინ $t \in [0,1] => a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. თუ $h_n \to 0$ და $\sqrt{n} h_n^{2m+2} \to \infty$ როცა $n \to \infty$, გვექნება

$$\sqrt{n}\left(I(\hat{a}_n) - I(a_n)\right) \to_d N(0, r^2),$$

სადაც

$$r^2 = \sigma^2 \left(\int_0^1 a(t)\left(1 + loga(t)\right) dt \right)^2.$$

განმეორებითი ლოგარითმის თვისება

გამოვიყენოთ კულბსის (იხ. [31]) შედეგი განმეორებითი ლოგარითმის შესახებ და ჩამოვაყალიბოთ შემდეგი

თეორემა 2.4. თუ h_n მიმდევრობა შერჩეულია ისე, რომ

$$R_n = o\left(\sqrt{\frac{loglogn}{n}}\right),$$

მაშინ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = r.$$

დამტკიცება. მართლაც, ადვილი დასანახია, რომ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = \lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(\alpha(n,k)Y(t_k) - \alpha(n,k)a(t_k)\big)}{\sqrt{2loglogn}} = r.$$

§3. გასერ-მიულერის შეფასებები და ფუნქციონალების ასიმპტოტური თვისებები

ზემოთ მოყვანილი

$$I_1(a) = \int_{-\infty}^{\infty} a^2(t)dt, \quad I_2(a) = \int_{-\infty}^{\infty} \frac{\left(a'(t)\right)^2}{a(t)}dt,$$

$$I_3(a) = \int_{-\infty}^{\infty} \left(a(t)\right)^s dt, \quad I_4(a) = \int_{-\infty}^{\infty} a(t)loga(t)dt.$$

ფუნქციონალები საინტერესოა გასერ-მიულერის შეფასებებისთვისაც.

ვთქვათ $a(t)$ არის რეგრესიის ფუნქცია გასერ-მიულერის სქემაში. ამ შემთხვევისთვისაც გვაქვს ცნობილი რეგრესიის მოდელი:

$$Y(t) = a(t) + \varepsilon(t), \tag{3.1}$$

სადაც $t \in [0,1]$, $\varepsilon(\cdot)$ არის შეშფოთება ნულოვანი საშუალოთი და სასრული დისპერსიით, $E\varepsilon(t) = 0$, $E\varepsilon^2(t) = \sigma^2 < \infty$, $Y(t)$ შემთხვევითი ფუნქციაა, $a(t)$ კი - უცნობი რეგრესიის ფუნქცია, რომელიც არის $Y(t)$-ს მათემატიკური ლოდინი. დავუშვათ გვაქვს n წერტილი:

$$0 \le t_1 \le t_2 \le \cdots \le t_n \le 1,$$

სადაც თითოეული t_k, $k = 1,2,\dots,n$ დამოკიდებულია n-ზე. ამ წერტილებში გაგვაჩნია დაკვირვებები:

$$Y(t_1), Y(t_2), \dots, Y(t_n).$$

უცნობი რეგრესიის $a(t)$ ფუნქციის კიდევ ერთი შეფასება შემოღებულ იქნა გასერისა და მიულერს მიერ (იხ. [12-13]) და განსაზღვრულია შემდეგნაირად:

$$\hat{a}_n(t) = \frac{1}{h_n} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i), \tag{3.2}$$

სადაც, $0 = s_1 \le s_2 \le \cdots \le s_n = 1$, $t_i \le s_i \le t_{i+1}$, $i = 1,2,3,\dots,n-1$ და $max_i|s_i - s_{i-1}| = O\left(\frac{1}{n}\right)$; $\{h_n,\ n = 1,2,\dots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი -

ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა. მათ მიერ განსაზღვრულია რეგრესიის ფუნქციის k-ური რიგის $a^{(k)}(t)$ წარმოებულის შეფასებაც შემდეგნაირად

$$\hat{a}_n^{(k)}(t) = \frac{1}{{h_n}^{k+1}} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W^{(k)}\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i), \qquad (\mathbf{3.3})$$

ყოველი $k = 0, 1, 2, \ldots, m$-თვის. აქ და ქვემოთ ჩავთვლით, რომ $\hat{a}_n^{(0)}(t) \doteqdot \hat{a}_n(t)$.

ავიღოთ $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია, რომელიც იქნება უწყვეტი, შემოსაზღვრული და გლუვი და განვიხილოთ შემდეგი სახის ინტეგრალური ფუნქციონალი:

$$I(a) = \int_{-\infty}^{\infty} \varphi\left(t, a(t), a'(t), \ldots, a^{(m)}(t)\right) dt. \qquad (\mathbf{3.4})$$

გვაქვს შერჩევა $(t_i, Y_i), i = 1, 2, \ldots, n$. ეს ნიშნავს, რომ

$$Y_i = Y(t_i) = a(t_i) + \varepsilon(t_i). \qquad (\mathbf{3.5})$$

$I(a)$-ს შეფასებისთვის ვიყენებთ ე.წ. ”ჩასმის” შეფასებას, ანუ ვიხილავთ ფუნქციონალს:

$$I(\hat{a}_n) = \int_{-\infty}^{\infty} \varphi\left(t, \hat{a}_n(t), \hat{a}_n'(t), \ldots, \hat{a}_n^{(m)}(t)\right) dt.$$

აქ $\hat{a}_n^{(k)}(t)$ განსაზღვრულია $(\mathbf{3.3})$ ფორმულიდან.

წარმოდგენის თეორემა

ამ ნაწილში მოვიყვანთ ”წარმოდგენის” თეორემას, საიდანაც მივიღებთ ჩვენთვის საინტერესო შედეგებს. ჩამოვაყალიბოთ პირობები განსახილველი სიდიდეებისათვის.

პირობები $\boldsymbol{a}$-სათვის:

$(a1)$ $a = a(t)$ განსაზღვრულია და უწყვეტია [0, 1]-ზე, იგი მნიშვნელობებს იღებს $[-\mathbb{k}, \mathbb{k}]$ ინტერვალიდან;

$(a2)$ $a = a(t)$ ფუნქციას აქვს უწყვეტი წარმოებულები m რიგის ჩათვლით;

$(a3)$ ნებისმიერი $i = 0, 1, \ldots, m$-სათვის, $a^{(i)}(\cdot) \in L_1([0, 1])$.

პირობები $\boldsymbol{\varepsilon_k = \varepsilon(t_k)}$-სათვის:

$(\varepsilon 1)$ ε_k, $k = 1, 2, \ldots$ დამოუკიდებელი და ერთნაირად განაწილებული შემთხვევითი სიდიდეებია;

$(\varepsilon 2)$ $E\varepsilon_k = 0$, $D\varepsilon_k^2 = \sigma^2 < \infty$.

შემოვიღოთ აღნიშვნები $\varphi = \varphi(x, x_0, \dots, x_m) \in C_b^2(R^{m+2})$ ფუნქციისათვის

$$\frac{\partial \varphi}{\partial x_i} = \varphi_{(i)},\ i = 0, 1, \dots, m \text{ და } \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \varphi_{(ij)},\ i, j = 0, 1, \dots, m.$$

პირობები $\boldsymbol{\varphi}$-სათვის:

(φ) $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია უწყვეტია, შემოსაზღვრულია, ინტეგრებადია და მას გააჩნია მეორე რიგის ჩათვლით უწყვეტი წარმოებულები, რომელიღაც ღია, ამოზნექილ A არეში, რომელსაც შეიცავს $\mathbb{R} \times [-\Bbbk, \Bbbk]^{m+1}$ სიმრავლე; φ-ის პირველი და მეორე წარმოებულები თანაბრად შემოსაზღვრულია A არეში $C_\varphi > 0$ მუდმივით. ამ პირობების თანახმად, φ ფუნქციისათვის ნებისმიერი $i, j = 0, 1, \dots, m$-თვის გვაქვს:

$$sup\left\{\left|\varphi_{(i,j)}\right|(s, s_0, s_1, \dots, s_m): (s, s_0, s_1, \dots, s_m) \in A\right\} \le C_\varphi. \qquad (\mathbf{3.6})$$

პირობები $\boldsymbol{W}$-სათვის:

$(W1)$ $\int_{-\infty}^{\infty} W(t)\, dt = 1$;

$(W2)$ $W(t)$ ფუნქცია უწყვეტად წარმოებადია $m \ge 1$ რიგის ჩათვლით; უწყვეტ $W^{(i)}(t)$, $i = 0, 1, \dots, m$ ფუნქციებს აქვთ კომპაქტური მატარებელი $[-\tau,\ \tau]$. შესაბამისად, $W^{(i)}(-\tau) = W^{(i)}(\tau) = 0$; შესაბამისად მოიძებნება $C_W > 0$ მუდმივი, რომლისთვისაც

$$sup_{t \in R}\left|W^{(i)}(t)\right| \le C_W < \infty, \qquad i = 0, 1, \dots, m.$$

ცხადია, ნებისმიერი $i = 0, 1, \dots, m$-თვის $W^{(i)} \in L_1([-\tau,\ \tau])$.

პირობები $\boldsymbol{h_n}$-სათვის:

(h) $\quad \dfrac{\sqrt{\max(|logh_n|; loglogn)}}{\sqrt{n} h_n^{0,5+m}} \to 0$, როცა $n \to \infty$.

$a_n(t)$-თი ავღნიშნოთ $\hat{a}_n(t)$-ს მათემატიკური ლოდინი:

$$a_n(t) = E\hat{a}_n(t) = E\left(\frac{1}{h_n} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot Y(t_i)\right) = \frac{1}{h_n} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot a(t_i).$$

ანალოგიურად,

$$a_n^{(k)}(t) = E\hat{a}_n^{(k)}(t) = \frac{1}{{h_n}^{k+1}} \cdot \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} W^{(k)}\left(\frac{t-u}{h_n}\right) du \cdot a(t_i).$$

ვაჩვენოთ, რომ არსებობს და სასრულია გამოსახულებები: $I(a), I(a_n)$ და $I(\hat{a}_n)$.

ტეილორის ფორმულის თანახმად ნებისმიერი $(s, s_0, s_1, \dots, S_m) \in A$ წერტილისთვის და რაღაც $\tilde{s}_i \in A$-თვის შეგვიძლია დავწეროთ

$$|\varphi|(s, s_0, s_1, \dots, S_m) = \left| \sum_{i=0}^{m} \varphi_{(i)}(s, 0,0, \dots ,0) s_i + \frac{1}{2} \sum_{i,j=0}^{m} \varphi_{(i,j)}(s, \tilde{s}_0, \tilde{s}_1, \dots, \tilde{s}_m) s_i s_j \right|.$$

შესაბამისად, მოიძებნება მუდმივი C ისეთი, რომ

$$|\varphi|(s, s_0, s_1, \dots, S_m) \le C \left(\sum_{i=0}^{m} |s_i| + \sum_{i=0}^{m} |s_i|^2 \right).$$

აქედან გამომდინარეობს, რომ ნებისმიერი შემოსაზღვრული ზომადი $f_0(t), f_1(t), \dots, f_m(t)$ ფუნქციებისთვის $L_1(R)$-დან გვექნება

$$\int_{-\infty}^{\infty} |\varphi|\big(t, f_0(t), f_1(t), \dots, f_m(t)\big) dt < \infty. \qquad (\mathbf{3.7})$$

ასე, რომ $I(a)$ არსებობს.

პირობები, რომლებიც შემოღებულია W-თვის გვაძლევს მის შემოსაზღვრულობას და ინტეგრებადობას. $(W2)$ პირობიდან და $(\mathbf{3.6})$ და $(\mathbf{3.7})$ ფორმულებიდან გამომდინარეობს ორივე გამოსახულება $I(a_n)$ და $I(\hat{a}_n)$, სასრულია ნებისმიერი $n \in \mathbb{N}$-თვის.

ტეილორის ფორმულის თანახმად

$$I(\hat{a}_n) - I(a_n) = S_n(h_n) + R_n, \qquad (\mathbf{3.8})$$

სადაც,

$$S_n(h_n) = \sum_{i=0}^{m} \int_{0}^{1} \varphi_{(i)} \left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t) \right) \left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t) \right) dt. \qquad (\mathbf{3.9})$$

შევნიშნოთ, რომ ნებისმიერი $h_n > 0$-თვის, $s_n(h_n)$ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების ჯამს. R_n ნაშთით წევრს კი აქვს შემდეგი სახე:

$$R_n = \frac{1}{2} \cdot \sum_{i,j=1}^{m} \int_{0}^{1} \varphi_{(ij)} \left(\tilde{b}_m(t) \right) \left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t) \right) \left(\hat{a}_n^{(j)}(t) - a_n^{(j)}(t) \right) dt. \qquad (\mathbf{3.10})$$

ამ ფორმულაში $\tilde{b}_m(t)$ არის $\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t) \right)$ და $\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t) \right)$ წერტილების შემაერთებელი წრფე.

$(\mathbf{3.7})$ და $(\mathbf{3.10})$ ფორმულებიდან ცხადია, შეგვიძლია R_n ნაშთითი წევრის შეფასება:

$$|R_n| \le C_\varphi \int_0^1 \sum_{i=0}^{m} \left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)^2 dt. \qquad (\mathbf{3.11})$$

ვთქვათ, W_m^2 სობოლევის სივრცეა ფუნქციებისა რომლებიც არიან კვადრატთან ერთად ინტეგრებადი, უწყვეტი და გააჩნიათ შემოსაზღვრული მეორე რიგის წარმოებული. ამ სივრცეში ნორმა და სკალარული ნამრავლი განისაზღვრება შემდეგნაირად:

$$\|g\|_m = \sqrt{\sum_{i=0}^{m} \int_0^1 |g^{(i)}(t)|^2 dt}, \quad \langle g_1, g_2 \rangle_m = \sqrt{\sum_{i=0}^{m} \int_0^1 \left(g_1^{(i)}(t) \cdot g_2^{(i)}(t)\right) dt}.$$

შემოვიღოთ აღნიშვნა $r_n(m) = \|\hat{a}_n - a_n\|_m^2$. ვხედავთ, რომ

$$|R_n| \le C_\varphi r_n(m). \qquad (\mathbf{3.12})$$

ვთქვათ,

$$U_k = U_k(t) = \frac{1}{h_n} \cdot \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot [Y(t_k) - a(t_k)], \qquad k = 1, 2, \dots, n,$$

სადაც $a(t_k) = EY(t_k)$, მაშინ

$$\sum_{k=1}^{n} U_k = \sum_{k=1}^{n} \frac{1}{h_n} \cdot \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot [Y(t_k) - a(t_k)] =$$
$$= \sum_{k=1}^{n} \frac{1}{h_n} \cdot \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot Y(t_k) - \sum_{k=1}^{n} \frac{1}{h_n} \cdot \int_{s_{i-1}}^{s_i} W\left(\frac{t-u}{h_n}\right) du \cdot a(t_k) = \hat{a}_n(t) - a_n(t).$$

შესაბამისად,

$$r_n(m) = \left\| \sum_{k=1}^{n} U_k \right\|_m^2. \qquad (\mathbf{3.13})$$

ნებისმიერი $k = 1, 2, \dots, n$-თვის შევაფასოთ $(\mathbf{3.13})$ ფორმულის თითოეული U_k შესაკრების $\|\cdot\|_m$ ნორმა. გვექნება

$$\|U_k\|_m = \left(\sum_{i=0}^{m} \int_0^1 \left|U_k^{(i)}(t)\right|^2 dt\right)^{\frac{1}{2}} = \left(\sum_{i=0}^{m} \int_0^1 \left|\frac{1}{{h_n}^{i+1}} \int_{s_{k-1}}^{s_k} W^{(i)}\left(\frac{t-u}{h_n}\right) du \cdot [Y(t_k) - a(t_k)]\right|^2 dt\right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=0}^{m} \frac{1}{{h_n}^{2i+2}} \int_0^1 \left| h_n \cdot \int_{\frac{t-s_k}{h_n}}^{\frac{t-s_{k-1}}{h_n}} W^{(i)}\left(\frac{t-u}{h_n}\right) d\left(\frac{t-u}{h_n}\right) \right|^2 \cdot [Y(t_k) - a(t_k)]^2 dt \right)^{\frac{1}{2}} =$$

$$= |Y(t_k) - a(t_k)| \left(\sum_{i=0}^{m} \frac{1}{{h_n}^{2i}} \int_0^1 \left| \int_{\frac{t-s_k}{h_n}}^{\frac{t-s_{k-1}}{h_n}} W^{(i)}\left(\frac{t-u}{h_n}\right) d\left(\frac{t-u}{h_n}\right) \right|^2 dt \right)^{\frac{1}{2}}$$

$$\le 2|\varepsilon_k| C_W \cdot \left(\sum_{i=0}^{m} \frac{1}{{h_n}^{2i}} \int_0^1 \left| \frac{t-s_{k-1}}{h_n} - \frac{t-s_k}{h_n} \right|^2 dt \right)^{\frac{1}{2}} = 2|\varepsilon_k| C_W \cdot |s_k - s_{k-1}| \left(\sum_{i=0}^{m} \frac{1}{{h_n}^{2i}} \right)^{\frac{1}{2}} =$$

$$= \frac{|\varepsilon_k| C_W}{n} \cdot \left(\frac{1 - h_n^{2m+2}}{h_n^{2m+2}(1 - h_n^2)} \right)^{\frac{1}{2}} \le L \cdot \frac{1}{n h_n^{m+1}} := M_m \sim O\left(\frac{1}{n h_n^{m+1}}\right) \qquad (\mathbf{3.14})$$

საკმარისად დიდი $L > 0$-თვის.

$r_n(m)$-ის შეფასებისთვის გამოვიყენოთ მაკ-დაიარმიდის უტოლობა

$$H(U_1, U_2, \dots, U_n) = \left\| \sum_{k=1}^{n} U_k \right\|_m$$

ფუნქციისთვის.

$$\left| H(U_1, U_2, \dots, U_{i-1}, U_x, U_{i+1}, \dots, U_n) - H\big(U_1, U_2, \dots, U_{i-1}, U_y, U_{i+1}, \dots, U_n\big) \right|$$
$$= \left| \left\| \sum_{k=1, k\neq i}^{n} U_k + U_x \right\|_m - \left\| \sum_{k=1, k\neq i}^{n} U_k + U_y \right\|_m \right|$$
$$\le \left\| \left(\sum_{k=1, k\neq i}^{n} U_k + U_x \right) - \left(\sum_{k=1, k\neq i}^{n} U_k + U_y \right) \right\|_m$$
$$\le \|U_x - U_y\|_m \le \|U_x\|_m + \|U_y\|_m \le 2M_m$$

მაშასადამე (**18**) პირობებში ყოველი $c_k \equiv 2M_m,\ k = 1, \dots, n$ (**19**) უტოლობის და (**3.14**) შეფასების თანახმად, ნებისმიერი $\delta > 0$-თვის გვექნება

$$P\left\{ \left| \left\| \sum_{k=1}^{n} U_k \right\|_m - E \left\| \sum_{k=1}^{n} U_k \right\|_m \right| > \delta \right\} \le 2exp\left(-\frac{2\delta^2}{\sum_{k=1}^{n} 4M_m^2} \right) = 2exp\left(-\frac{2\delta^2}{n \cdot 4M_m^2} \right)$$

$$= 2exp\left(-\frac{2\delta^2 n^2 h_n^{2m+2}}{n \cdot 4L^2}\right) = 2exp\left(-\frac{\delta^2 n h_n^{2m+2}}{2L^2}\right).$$

აქ ჩავსვათ

$$\delta = \frac{2L\sqrt{logn}}{\sqrt{n}h_n^{m+1}}$$

და მივიღებთ

$$P\left\{\left|\left\|\sum_{k=1}^{n} U_k\right\|_m - E\left\|\sum_{k=1}^{n} U_k\right\|_m\right| > \delta\right\} \le 2exp\left(-\frac{4L^2 logn \cdot nh_n^{2m+2}}{nh_n^{2m+2} \cdot 2L^2}\right) = 2exp(-2logn) = \frac{2}{n^2}.$$

გამოვიყენოთ ბორელი-კანტელის ლემა. მივიღებთ

$$\left\|\sum_{k=1}^{n} U_k\right\|_m = E\left\|\sum_{k=1}^{n} U_k\right\|_m + O\left(\frac{\sqrt{logn}}{\sqrt{n}h_n^{m+1}}\right). \qquad (\mathbf{3.15})$$

იენსენის უტოლობის თანახმად

$$\left(E\left\|\sum_{k=1}^{n} U_k\right\|_m\right)^2 \le E\left\|\sum_{k=1}^{n} U_k\right\|_m^2 = E\sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 \left|\frac{1}{h_n{}^{i+1}} \cdot \int_{s_{k-1}}^{s_k} W^{(i)}\left(\frac{t-u}{h_n}\right) du \cdot [Y(t_k) - a(t_k)]\right|^2 dt$$

$$= \sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 E\left|\frac{1}{h_n{}^{i+1}} \cdot \int_{s_{k-1}}^{s_k} W^{(i)}\left(\frac{t-u}{h_n}\right) du \cdot [Y(t_k) - a(t_k)]\right|^2 dt$$

$$\le 2C_W^2 \cdot \sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 \frac{1}{h_n^{2i+2}} \cdot \left|\int_{s_{k-1}}^{s_k} du\right|^2 \cdot E[Y(t_k) - a(t_k)]^2 dt$$

$$= 2C_W^2\sigma^2 \cdot \frac{1 - h_n^{2m+2}}{h_n^{2m+2}(1 - h_n^2)} \cdot \sum_{k=1}^{n}(s_k - s_{k-1})^2 \le K \cdot \frac{1}{nh_n^{2m+2}}, \qquad (\mathbf{3.16})$$

$(\mathbf{3.12})$, $(\mathbf{3.13})$, $(\mathbf{3.15})$ და $(\mathbf{3.16})$ ფორმულებიდან შეგვიძლია დავასკვნათ, რომ

$$R_n = O\left(\frac{logn}{nh_n^{2m+2}}\right).$$

ასე, რომ სამართლიანია შემდეგი თეორემა.

თეორემა 3.1. ვთქვათ სრულდება $(a1)-(a3),(\varepsilon 1)-(\varepsilon 3),(\varphi),(W1)-(W2)$ და (h) პირობები. მაშინ, სამართლიანია $(\mathbf{3.8})$ ტოლობა, სადაც ნაშთით წევრს ალბათობით 1 აქვს შემდეგი რიგი

$$R_n = O\left(\frac{logn}{nh_n^{2m+2}}\right).$$

ძალდებულობა

ამ ნაწილში გამოვიყენებთ თეორემა 3.1-ს და დავამტკიცებთ $I(\hat{a}_n)$ შეფასების მკაცრად ძალდებულობას.

თეორემა 3.2. ვთქვათ, სრულდება თეორემა 3.1-ის პირობები. დადებითი $(h_n)_{n=1}^{\infty}$ მიმდევრობა შერჩეულია ისე, რომ

$$\frac{logn}{nh_n^{2m+2}} \to 0, \text{ როცა } n \to \infty. \qquad (\mathbf{3.18})$$

მაშინ ალბათობით 1

$$I(\hat{a}_n) \to I(a). \qquad (\mathbf{3.19})$$

დამტკიცება: თეორემა 1-ის და $(\mathbf{2.8})$ ფორმულის თანახმად

$$I(\hat{a}_n) - I(a_n) = S_n(h_n) + R_n, \qquad (\mathbf{3.20})$$

სადაც $R_n = o(1)$ და

$$S_n(h_n) = \sum_{i=0}^{m} \int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)dt.$$

$(a1)$ პირობიდან:

$$\left\{\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right): t \in [0,1]\right\} \subset [0,1] \times [-\Bbbk, \Bbbk]^{m+1}.$$

ამიტომ, $(\varphi 2)$ პირობის გათვალისწინებით, მოიძებნება ისეთი $C_\varphi > 0$ მუდმივი, რომ

$$sup\{\left|\varphi_{(i)}\right|(t, t_0, t_1, \dots, t_m): (t, t_0, t_1, \dots, t_m) \in [0,1] \times [-\Bbbk, \Bbbk]^{m+1}\} \le C_\varphi.$$

შეგვიძლია დავწეროთ:

$$ES_n(h_n) = 0$$

$$DS_n(h_n) = ES_n^2(h_n) \leq C_\varphi^2 \cdot \sum_{i=0}^{m} \int_0^1 E\left[\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right] dt$$

$$\leq C_\varphi^2 \cdot \sum_{i=0}^{m} \int_0^1 \left| \sum_{k=0}^{n} \frac{1}{h_n^{i+1}} \cdot \int_{s_{k-1}}^{s_k} W^{(i)}\left(\frac{t-u}{h_n}\right) du \cdot E[Y(t_k) - a(t_k)] \right|^2 dt$$

$$\leq C_\varphi^2 C_W^2 \sigma^2 \cdot \sum_{k=0}^{n} (s_k - s_{k-1})^2 \cdot \sum_{i=1}^{m} \left(\frac{1}{h_n^{i+1}}\right)^2 \sim C \cdot \frac{1}{nh_n^{2m+2}} \to 0 \qquad \mathbf{(3.21)}$$

რადგან $\frac{logn}{nh_n^{2m+2}} \to 0$ და ამიტომ $S_n(h_n) \to 0$ როცა $n \to \infty$.

$$Ea_n^{(k)}(t) = \int_{-\tau}^{\tau} W(u) a^{(k)}(t)(t - uh_n) du + O\left(\frac{1}{nh_n^k}\right).$$

აქედან შეიძლება დავასკვნათ:

i) $(\mathbf{3.17})$ ფორმულიდან გამომდინარე ნულისკენ კრებადია $\frac{1}{nh_n^k}$, ნებისმიერი $k = 0,1, \dots, m$-თვის;

ii) როცა $n \to \infty$ გვაქვს $Ea_n^{(k)}(t) \to a^{(k)}(t)$.

შევაჯამოთ ნათქვამი და მივიღებთ, რომ $n \to \infty$-თვის

$$I(a_n) = \int_0^1 \varphi\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \to \int_0^1 \varphi\left(t, a(t), a'(t), \dots, a^{(m)}(t)\right) dt = I(a).$$

ასე, რომ $I(\hat{a}_n) - I(a_n) = o(1)$ ანუ $I(\hat{a}_n) - I(a) \to 0$. თეორემა დამტკიცებულია.

ცენტრალური ზღვარითი თეორემა

ჩვენი წარმოდგენის თეორემის გამოყენებით შეიძლება მივიღოთ ზღვარითი განაწილების თვისება ინტეგრალური ფუნქციონალისთვის

$$I(\hat{a}_n) = \int_0^1 \varphi\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right) dt.$$

განვიხილოთ სხვაობა

$$I(\hat{a}_n) - I(a_n) = s_n(h_n) + R_n \qquad \mathbf{(3.8)}$$

სადაც ნებისმიერი $h_n > 0$-თვის, $s_n(h_n)$ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების ჯამს.

$$S_n(h_n) = \sum_{i=0}^{m} \int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right) dt. \qquad (\mathbf{3.9})$$

R_n ნაშთით წევრს კი აქვს შემდეგი სახე:

$$R_n = \frac{1}{2} \cdot \sum_{i,j=1}^{m} \int_0^1 \varphi_{(ij)}\left(\tilde{b}_m(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)\left(\hat{a}_n^{(j)}(t) - a_n^{(j)}(t)\right) dt. \qquad (\mathbf{3.10})$$

ცხადია, რომ

$$ES_n(h_n) = 0 \text{ და } ER_n \to 0 \text{ როცა } n \to \infty. \qquad (\mathbf{3.23})$$

გარდა ამისა

$$E\big(S_n(h_n)\big)^2 = \sigma^2 \sum_{i=0}^{m} \left(\int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \right)^2 \qquad (\mathbf{3.24})$$

და $\mathrm{Var}R_n \to 0$ როცა $n \to \infty$.

ამ პირობებისთვის უნდა დავამტკიცოთ $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)$ სიდიდის ასიმპტოტურად ნორმალურობა და უნდა გამოვთვალოთ ზღვარითი ვარიაცია. აგრეთვე, ($\mathbf{3.8}$), ($\mathbf{3.23}$) და ($\mathbf{3.24}$) ფორმულების საშუალებით უნდა ვაჩვენოთ $\sqrt{n}S_n(h_n)$ სიდიდის ასიმპტოტურად ნორმალურობა. ამისთვის საკმარისია ეს თვისება შევისწავლოთ შემდეგი სიდიდისთვის:

$$d_k = Y(t_k) \cdot \sum_{i=0}^{m} \frac{1}{h_n^{i+1}} \cdot \int_0^1 \int_{s_{k-1}}^{s_k} W^{(i)}\left(\frac{t-u}{h_n}\right) \cdot \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt du \qquad (\mathbf{3.25})$$

ადვილი შესამჩნევია, რომ

$$Ed_k = a(t_k) \cdot \sum_{i=0}^{m} \frac{1}{h_n^{i+1}} \cdot \int_0^1 \int_{s_{k-1}}^{s_k} W^{(i)}\left(\frac{t-u}{h_n}\right) \cdot \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt du \qquad (\mathbf{3.26})$$

ანუ ვიხილავთ დამოუკიდებელი შემთხვევითი სიდიდეების მიმდევრობას:

$$f_n(k) = \alpha(n,k)\big(Y(t_k) - a(t_k)\big) = \alpha(n,k)\varepsilon_k$$

სადაც

$$\alpha(n,k)=\sum_{i=0}^{m}\frac{1}{h_n^{i+1}}\cdot\int_0^1\int_{s_{k-1}}^{s_k}W^{(i)}\left(\frac{t-u}{h_n}\right)\cdot\varphi_{(i)}\left(t,a_n(t),a_n'(t),\dots,a_n^{(m)}(t)\right)dtdu$$

განვიხილოთ ჯამი $S_n(h_n)=\sum_{k=1}^{n}\alpha(n,k)\varepsilon_k$.

ვთქვათ, $F_{k,n}$ არის $\alpha(n,k)\varepsilon_k$ შემთხვევითი სიდიდეების ალბათური განაწილების ფუნქცია, F_ε – კი ε_k შემთხვევითი სიდიდეების ალბათური განაწილების ფუნქცია. ლინდებერგის პირობას აქვს შემდეგი სახე $\forall\delta>0$, $lim_{n\to\infty}L_n(\delta)=0$ სადაც

$$L_n(\delta)=\frac{\sum_{j=1}^{n}\int x^2J\left(|x|\geq\delta\sigma(\sum_{k=1}^{n}\alpha^2(n,k))^{1/2}\right)dF_{k,n}(x)}{\sigma^2\sum_{k=1}^{n}\alpha^2(n,k)}.$$

აქ $J(A)$ წარმოადგენს A სიმრავლის ინდიკატორ ფუნქციას. ადვილი დასანახია, რომ

$$L_n(\delta)\leq\frac{1}{\sigma^2}max_{1\leq j\leq n}\int x^2J\big(|x|\geq\delta\sigma v(n,j)\big)dF_\varepsilon$$

სადაც

$$v(n,j)=\frac{|\alpha(n,j)|}{\left(\sum_{j=1}^{n}\alpha^2(n,j)\right)^{1/2}}$$

საჩვენებელი რჩება, რომ $max_{1\leq j\leq n}v(n,j)\to 0$ როცა $n\to\infty$. მაგრამ რადგანაც

$$max_{1\leq j\leq n}|\alpha(n,j)|=O\left(\frac{1}{nh_n^{m+1}}\right),$$

ანუ

$$max_{1\leq j\leq n}v(n,j)=O\left(\frac{1}{\sqrt{n}}\right)$$

ასე, რომ ლინდერბერგის პირობა სრულდება და შეგვიძლია ჩამოვაყალიბოთ

თეორემა 3.3: ვთქვათ, სრულდება თეორემა 3.1-ის პირობები. მაშინ, თუ $h_n\to 0$ და $\sqrt{n}h_n^{2m+2}\to\infty$ როცა $n\to\infty$, სამართლიანია $\sqrt{n}\big(I(\hat{a}_n)-I(a_n)\big)\to_d N(0,r^2)$. სადაც

$$r^2=\sigma^2\cdot\sum_{i=0}^{m}\left(\int_0^1\varphi_{(i)}\left(t,a_n(t),a_n'(t),\dots,a_n^{(m)}(t)\right)dt\right)^2.$$

ზოგიერთი გამოყენება

განვიხილოთ ინტეგრალური ფუნქციონალი

$$I_1(a) = \int_0^1 a^2(t)\, dt.$$

ამ შემთხვევაში $\varphi(t, x_0, x_1, \dots, x_m) = x_0^2$, როცა $x_0 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. მაშინ თუ $h_n \to 0$ და $\sqrt{n} h_n^{2m+2} \to \infty$ როცა $n \to \infty$, გვექნება კრებადობა

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2).$$

სადაც

$$r^2 = 4\sigma^2 \left(\int_0^1 a(t) dt \right)^2.$$

ახლა განვიხილოთ ფიშერის ინტეგრალური ფუნქციონალი

$$I_2(a) = \int_0^1 \frac{\big(a'(t)\big)^2}{a(t)} dt.$$

ფუნქციონალისთვის გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \frac{x_1^2}{x_2}$. თუ $t \in [0,1]$მაშინ $a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. თუ $h_n \to 0$ და $\sqrt{n} h_n^{2m+2} \to \infty$ როცა $n \to \infty$, მაშინ გვექნება $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2)$ კრებადობა, სადაც

$$r^2 = \sigma^2 \cdot \left(-\int_0^1 \left(\frac{\big(a'(t)\big)^2}{\big(a(t)\big)^2} - \frac{2a'(t)}{a(t)} \right) dt \right)^2$$
$$= \sigma^2 \left(\frac{a'(1)}{a(1)} log a(1) - \frac{a'(0)}{a(0)} log a(0) - \frac{1}{2} \cdot \frac{\big(a'(1)\big)^2}{\big(a(1)\big)^2} + \frac{1}{2} \cdot \frac{\big(a'(0)\big)^2}{\big(a(0)\big)^2} \right)^2.$$

ახლა განვიხილოთ

$$I_3(a) = \int_{-\infty}^{\infty} \big(a(t)\big)^s dt$$

ფუნქციონალი $(s > 1)$. მაშინ, $\varphi(t, x_0, x_1, \dots, x_m) = x_0^S$, როცა $x_1 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. თუ $h_n \to 0$ და $\sqrt{n} h_n^{2m+2} \to \infty$ როცა $n \to \infty$ მაშინ გვექნება $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2)$ კრებადობა, სადაც

$$r^2 = s^2\sigma^2\left(\int_0^1 a^{s-1}(t)dt\right)^2.$$

ბოლოს განვიხილოთ შენონის ინტეგრალური ენტროპია

$$I_4(a) = \int_{-\infty}^{\infty} a(t)loga(t)dt.$$

მაშინ, რომელიღაც საკმარისად დიდი $b \geq \mathbb{k} > 0$-თვის, თუ $0 < x_0 \leq b$ გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \varphi(x_0) = x_0 logx_0$. დამატებით განვსაზღვროთ φ ფუნქცია 0-მდე, $\varphi(x) = 0$ როცა $-b \leq x \leq 0$. მაშინ $t \in [0,1] => a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m+2} \to \infty$ როცა $n \to \infty$, ამ შემთხვევაში

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = \sigma^2\left(\int_0^1 a(t)\big(1 + loga(t)\big)dt\right)^2.$$

განმეორებითი ლოგარითმის თვისება

გამოვიყენოთ კულბსის ცნობილი განმეორებითი ლოგარითმის შედეგი (იხ. [30]) და ჩამოვაყალიბოთ შემდეგი

თეორემა 3.4. თუ h_n მიმდევრობა შერჩეულია ისე, რომ

$$R_n = o\left(\sqrt{\frac{loglogn}{n}}\right),$$

მაშინ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = r.$$

დამტკიცება.მართლაც, ადვილი დასანახია, რომ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = \lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(\alpha(n,k)Y(t_k) - \alpha(n,k)a(t_k)\big)}{\sqrt{2loglogn}} = r.$$

§4. ნადარაია-ვატსონის შეფასებები და ფუნქციონალების ასიმპტოტური თვისებები

ვთქვათ, $a(t)$ არის რეგრესიის ფუნქცია. აქ გამოვიყენებთ ნადარაია-ვატსონის შეფასებას იმ კერძო ვარიანტში, როცა რეგრესორად დეტერმინისტული წერტილებია აღებული. კერძო შემთხვევებში განვიხილავთ შემდეგ ფუნქციონალებს:

$$I_1(a) = \int_{-\infty}^{\infty} a^2(t)dt, \quad I_2(a) = \int_{-\infty}^{\infty} \frac{\left(a'(t)\right)^2}{a(t)}dt,$$

$$I_3(a) = \int_{-\infty}^{\infty} \left(a(t)\right)^s dt, \quad I_4(a) = \int_{-\infty}^{\infty} a(t)loga(t)dt.$$

ეს ფუნქციონალები ცალკე შესწავლის საგანია სხვადასხვა კვლევებისათვის. ჩვენ აქ გამოვიყენებთ მათთვის ერთიან მიდგომას, რომელიც დამყარებულია ე.წ. წარმოდგენის თეორემაზე.

განვიხილოთ შემდეგი ტიპის რეგრესიის მოდელი:

$$Y(t) = a(t) + \varepsilon(t), \tag{4.1}$$

სადაც $t \in [0,1]$, $\varepsilon(\cdot)$ არის შეშფოთება ნულოვანი საშუალოთი და სასრული დისპერსიით, $E\varepsilon(t) = 0,\ E\varepsilon^2(t) = \sigma^2 < \infty$, $Y(t)$ შემთხვევითი ფუნქციაა, $a(t)$ კი - უცნობი რეგრესიის ფუნქცია.

დავუშვათ გვაქვს n წერტილი:

$0 \le t_1 \le t_2 \le \cdots \le t_n \le 1$, სადაც თითოეული $t_k,\ k = 1,2,\dots,n$ დამოკიდებულია n-ზე, $max_i|t_i - t_{i-1}| = O\left(\frac{1}{n}\right)$. ამ წერტილებში გაგვაჩნია დაკვირვებები $Y(t_1), Y(t_2), \dots, Y(t_n)$.

უცნობი რეგრესიის $a(t)$ ფუნქციის ერთ-ერთი შეფასება შემოღებულ იქნა ნადარაიასა და ვატსონის მიერ (იხ. [9-10]) და განსაზღვრულია შემდეგნაირად:

$$\hat{a}_n(t) = \sum_{i=1}^{n} \frac{W\left(\frac{t - t_i}{h_n}\right)}{\sum_{k=1}^{n} W\left(\frac{t - t_k}{h_n}\right)} \cdot Y(t_i), \tag{4.2}$$

სადაც, $\{h_n,\ n = 1,2,\dots\}$ მონოტონურად, 0-სკენ კრებადი, დადებითი რიცხვების მიმდევრობაა, $W(t)$ კი - ალბათური სიმკვრივის თვისებების მქონე ფუნქციაა.

სიმარტივისთვის შემოვიღოთ აღნიშვნა:

$$\frac{W\left(\frac{t-t_i}{h_n}\right)}{\sum_{k=1}^{n} W\left(\frac{t-t_k}{h_n}\right)} \equiv \alpha(t, t_i, h_n).$$

ნადარაია-ვატსონის შეფასება გადაიწერება შემდეგნაირად:

$$\hat{a}_n(t) = \sum_{i=1}^{n} \alpha(t, t_i, h_n) \cdot Y(t_i).$$

სამართლიანია $\sum_{i=1}^{n} \alpha(t, t_i, h_n) = 1$ და $\sum_{i=1}^{n} \alpha^{(k)}(t, t_i, h_n) = 0$ ტოლობები (აქ და შემდეგში ყველგან $\alpha^{(k)}(t, t_i, h_n)$-ში იგულისხმება k რიგის წარმოებული t ცვლადის მიმართ).

ასეთი აღნიშვნით რეგრესიის ფუნქციის k-ური რიგის $a^{(k)}(t)$ წარმოებულის შეფასება შემდეგნაირად ჩაიწერება

$$\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} \alpha^{(k)}(t, t_i, h_n) \cdot Y(t_i) \qquad (\mathbf{4.3})$$

ყოველი $k = 0, 1, 2, \dots, m$-თვის.

ვთქვათ, $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია უწყვეტია, შემოსაზღვრულია და გლუვია. განვიხილოთ შემდეგი სახის ინტეგრალური ფუნქციონალი:

$$I(a) = \int_{-\infty}^{\infty} \varphi\left(t, a(t), a'(t), \dots, a^{(m)}(t)\right) dt. \qquad (\mathbf{4.4})$$

გვაქვს შერჩევა $(t_i, Y_i), i = 1, 2, \dots, n$. ეს ნიშნავს, რომ

$$Y_i = Y(t_i) = a(t_i) + \varepsilon(t_i). \qquad (\mathbf{4.5})$$

$I(a)$-ს შეფასებისთვის ვიყენებთ ე.წ. ”plug-in” შეფასებას. ანუ განვიხილავთ ფუნქციონალს:

$$I(\hat{a}_n) = \int_{-\infty}^{\infty} \varphi\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right) dt.$$

წარმოდგენის თეორემა

ამ ნაწილში მოვიყვანთ ”წარმოდგენის” თეორემას, საიდანაც მივიღებთ ჩვენთვის საინტერესო შედეგებს. ჩამოვაყალიბოთ პირობები განსახილველი სიდიდეებისათვის.

პირობები $\boldsymbol{a}$-სათვის:

$(a1)$ $a = a(t)$ განსაზღვრულია და უწყვეტია [0, 1]-ზე, იგი მნიშვნელობებს იღებს $[-\Bbbk, \Bbbk]$ ინტერვალიდან;

$(a2)$ $a = a(t)$ ფუნქციას აქვს უწყვეტი წარმოებულები m რიგის ჩათვლით;

$(a3)$ ნებისმიერი $i = 0, 1, \dots, m$-სათვის, $a^{(i)}(\cdot) \in L_1([0,1])$.

პირობები $\boldsymbol{\varepsilon_k = \varepsilon(t_k)}$-სათვის:

$(\varepsilon1)$ ε_k, $k = 1, 2, \dots$ დამოუკიდებელი და ერთნაირად განაწილებული შემთხვევითი სიდიდეებია;

$(\varepsilon2)$ $E\varepsilon_k = 0$, $D\varepsilon_k^2 = \sigma^2 < \infty$.

სიმარტივისთვის შემოვიღოთ აღნიშვნები $\varphi = \varphi(x, x_0, \dots, x_m) \in C_b^2(R^{m+2})$ ფუნქციისათვის

$$\frac{\partial\varphi}{\partial x_i} = \varphi_{(i)}, \ i = 0, 1, \dots, m \text{ და}$$

$$\frac{\partial^2\varphi}{\partial x_i \partial x_j} = \varphi_{(ij)}, \ i, j = 0, 1, \dots, m.$$

პირობები $\boldsymbol{\varphi}$-სათვის:

(φ) $\varphi: \mathbb{R}^{m+2} \to \mathbb{R}$ ფუნქცია უწყვეტია, შემოსაზღვრულია, ინტეგრებადია და მას გააჩნია მეორე რიგის ჩათვლით უწყვეტი წარმოებულები, რომელიღაც ღია, ამოზნექილ A არეში, რომელსაც შეიცავს $\mathbb{R} \times [-\Bbbk, \Bbbk]^{m+1}$ სიმრავლე; φ-ის პირველი და მეორე წარმოებულები თანაბრად შემოსაზღვრულია A არეში $C_\varphi > 0$ მუდმივით.

ამ პირობების თანახმად, φ ფუნქციისათვის ნებისმიერი $i, j = 0, 1, \dots, m$-თვის გვაქვს:

$$sup\left\{\left|\varphi_{(i,j)}\right|(s, s_0, s_1, \dots, s_m): (s, s_0, s_1, \dots, s_m) \in A\right\} \le C_\varphi. \qquad (\mathbf{4.6})$$

პირობები $\boldsymbol{W}$-სათვის:

$(W1)$ $\int_{-\infty}^{\infty} W(t)\, dt = 1$;

$(W2)$ $W(t)$ ფუნქცია უწყვეტად წარმოებადია $m \ge 1$ რიგის ჩათვლით. უწყვეტ $W^{(i)}(t)$, $i = 0, 1, \dots, m$ ფუნქციებს აქვთ კომპაქტური მატარებელი $[-\tau, \ \tau]$ ინტერვალში და ცხადია $W^{(i)}(-\tau) = W^{(i)}(\tau) = 0$;

$(W3)$ მოიძებნება $0 < C_{1W}, C_{2W} < \infty$ მუდმივები, რომლებისთვისაც

$$inf_{t\in R}\left(W^{(i)}(t)\right) \geq C_{1,W},$$

$$sup_{t\in R}\left(W^{(i)}(t)\right) \leq C_{2,W}$$

ყოველი $i = 0, 1, \dots, m$ −თვის.

$(W4)$ ნებისმიერი $i = 0, 1, \dots, m$-თვის $W^{(i)} \in L_1([-\tau,\ \tau])$.

პირობები $\boldsymbol{h_n}$-სათვის:

(h) $\quad \frac{\sqrt{\max(|logh_n|; loglogn)}}{\sqrt{n}h_n^{0,5+m}} \to 0,$ როცა $n \to \infty$.

შემდეგში დაგვჭირდება $\left|\alpha^{(k)}(t, t_i, h_n)\right|$-ის შეფასება და ვაჩვენოთ, რომ ყოველი $k = 0, 1, \dots, m$ −თვის სამართლიანია

$$\left|\alpha^{(k)}(t, t_i, h_n)\right| \leq \frac{1}{n \cdot h_n^k} \cdot C_k$$

უტოლობა, სადაც $0 < C_k < \infty$ გარკვეული მუდმივებია.

მართლაც $k = 0$ −თვის

$$\left|\alpha^{(0)}(t, t_i, h_n)\right| = \left|\frac{W\left(\frac{t-t_i}{h_n}\right)}{\sum_{k=1}^n W\left(\frac{t-t_k}{h_n}\right)}\right| \leq \frac{C_{2,W}}{n \cdot C_{1,W}} \equiv \frac{1}{n \cdot h_n^0} \cdot C_0.$$

$k = 1 -$ სთვის

$$\left|\alpha^{(1)}(t, t_i, h_n)\right| = \left|\frac{\frac{1}{h_n} \cdot W'\left(\frac{t-t_i}{h_n}\right) \cdot \sum_{k=1}^n W\left(\frac{t-t_k}{h_n}\right) - W\left(\frac{t-t_i}{h_n}\right) \cdot \frac{1}{h_n} \cdot \sum_{k=1}^n W'\left(\frac{t-t_k}{h_n}\right)}{\left(\sum_{k=1}^n W\left(\frac{t-t_k}{h_n}\right)\right)^2}\right|$$

$$\leq \frac{1}{h_n} \cdot \frac{\left|W'\left(\frac{t-t_i}{h_n}\right) \cdot \sum_{k=1}^n W\left(\frac{t-t_k}{h_n}\right)\right| + \left|W\left(\frac{t-t_i}{h_n}\right) \cdot \sum_{k=1}^n W'\left(\frac{t-t_k}{h_n}\right)\right|}{\left(\sum_{k=1}^n W\left(\frac{t-t_k}{h_n}\right)\right)^2}$$

$$\leq \frac{n \cdot C_{2,W}^2 + n \cdot C_{2,W}^2}{h_n \cdot \left(n \cdot C_{1,W}\right)^2} = \frac{1}{n \cdot h_n^1} \cdot \frac{2C_{2,W}^2}{C_{1,W}^2} \equiv \frac{1}{n \cdot h_n^1} \cdot C_1.$$

ინდუქციით ადვილად მივიღებთ რომ

$$\left|\alpha^{(k)}(t, t_i, h_n)\right| \leq \frac{1}{n \cdot h_n^k} \cdot C_k.$$

C_k-ებს შორის უდიდესს თუ ავღნიშნავთ C_α-თი, მივიღებთ, რომ ყოველი $k = 0, 1, \dots, m$ −თვის სამართლიანია უტოლობა

$$\left|\alpha^{(k)}(t, t_i, h_n)\right| \leq \frac{1}{n \cdot h_n^k} \cdot C_\alpha.$$

$a_n(t)$-თი ავღნიშნოთ $\hat{a}_n(t)$-ს მათემატიკური ლოდინი:

$$a_n(t) = E\hat{a}_n(t) = E\left(\sum_{i=1}^{n} \alpha(t, t_i, h_n) \cdot Y(t_i)\right) = \sum_{i=1}^{n} \alpha(t, t_i, h_n) \cdot a(t_i).$$

ანალოგიურად,

$$a_n^{(k)}(t) = E\hat{a}_n^{(k)}(t) = \frac{1}{h_n^k} \cdot \sum_{i=1}^{n} \alpha^{(k)}(t, t_i, h_n) \cdot a(t_i).$$

ვაჩვენოთ, რომ არსებობს და სასრულია გამოსახულებები: $I(a), I(a_n)$ და $I(\hat{a}_n)$.

ტეილორის ფორმულის თანახმად ნებისმიერი $(s, s_0, s_1, \dots, S_m) \in A$ წერტილისათვის და რაღაც $\tilde{s}_i \in A$-თვის შეგვიძლია დავწეროთ

$$|\varphi|(s, s_0, s_1, \dots, S_m) = \left|\sum_{i=0}^{m} \varphi_{(i)}(s, 0,0, \dots, 0)s_i + \frac{1}{2}\sum_{i,j=0}^{m} \varphi_{(i,j)}(s, \tilde{s}_0, \tilde{s}_1, \dots, \tilde{s}_m)s_i s_j\right|.$$

შესაბამისად, მოიძებნება მუდმივი C ისეთი, რომ

$$|\varphi|(s, s_0, s_1, \dots, S_m) \leq C\left(\sum_{i=0}^{m} |s_i| + \sum_{i=0}^{m} |s_i|^2\right).$$

აქედან გამომდინარეობს, რომ ნებისმიერი შემოსაზღვრული ზომადი $f_0(t), f_1(t), \dots, f_m(t)$ ფუნქციებისთვის $L_1(R)$-დან გვექნება

$$\int_{-\infty}^{\infty} |\varphi|\big(t, f_0(t), f_1(t), \dots, f_m(t)\big)dt < \infty. \qquad (\mathbf{4.7})$$

ასე, რომ $I(a)$ არსებობს.

პირობები, რომლებიც შემოღებულია W-თვის გვაძლევს მის შემოსაზღვრულობას და ინტეგრებადობას. $(\mathbf{4.6})$ და $(\mathbf{4.7})$ ფორმულებიდან გამომდინარეობს ორივე გამოსახულება $I(a_n)$ და $I(\hat{a}_n)$, სასრულია ნებისმიერი $n \in \mathbb{N}$-თვის.

ტეილორის ფორმულის გამოყენებით შეგვიძლია დავწეროთ

$$I(\hat{a}_n) - I(a_n) = S_n(h_n) + R_n, \qquad (\mathbf{4.8})$$

სადაც,

$$S_n(h_n) = \sum_{i=0}^{m} \int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right) dt. \qquad (\mathbf{4.9})$$

შევნიშნოთ, რომ ნებისმიერი $h_n > 0$-თვის $s_n(h_n)$ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების ჯამს. R_n ნაშთით წევრს კი აქვს შემდეგი სახე:

$$R_n = \frac{1}{2} \cdot \sum_{i,j=1}^{m} \int_0^1 \varphi_{(ij)}\left(\tilde{b}_m(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)\left(\hat{a}_n^{(j)}(t) - a_n^{(j)}(t)\right) dt. \qquad (\mathbf{4.10})$$

ამ ფორმულაში $\tilde{b}_m(t)$ არის $\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right)$ და $\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)$ წერტილების შემაერთებელი წრფე.

გამოვიყენოთ სტანდარტული პროცედურა, $(\mathbf{4.7})$ და $(\mathbf{4.10})$ ფორმულებიდან შევძლებთ R_n ნაშთითი წევრის შეფასებას:

$$|R_n| \le C_\varphi \int_0^1 \sum_{i=0}^{m} \left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)^2 dt. \qquad (\mathbf{4.11})$$

ვთქვათ W_m^2 სობოლევის სივრცეა. ამ სივრცეში ნორმა და სკალარული ნამრავლი ასე განისაზღვრება

$$\|g\|_m = \sqrt{\sum_{i=0}^{m} \int_0^1 |g^{(i)}(t)|^2 dt}, \quad \langle g_1, g_2 \rangle_m = \sqrt{\sum_{i=0}^{m} \int_0^1 \left(g_1^{(i)}(t) \cdot g_2^{(i)}(t)\right) dt}.$$

შემოვიღოთ აღნიშვნა $r_n(m) = \|\hat{a}_n - a_n\|_m^2$. ვხედავთ, რომ

$$|R_n| \le C_\varphi r_n(m). \qquad (\mathbf{4.12})$$

ვთქვათ, $U_k = U_k(t) = \alpha(t, t_i, h_n) \cdot [Y(t_k) - a(t_k)],\ k = 1, 2, \dots, n,$

სადაც $a(t_k) = EY(t_k)$, მაშინ

$$\sum_{k=1}^{n} U_k = \sum_{k=1}^{n} \alpha(t, t_i, h_n) \cdot [Y(t_k) - a(t_k)] =$$

$$= \sum_{k=1}^{n} \alpha(t, t_i, h_n) \cdot Y(t_k) - \sum_{k=1}^{n} \alpha(t, t_i, h_n) \cdot a(t_k) = \hat{a}_n(t) - a_n(t).$$

შესაბამისად,

$$r_n(m) = \left\| \sum_{k=1}^{n} U_k \right\|_m^2. \qquad (\mathbf{4.13})$$

$(\mathbf{4.13})$ ფორმულაში ნებისმიერი $k = 1, 2, \dots, n$-თვის შევაფასოთ თითოეული U_k შესაკრები $\|\cdot\|_m$ ნორმით. გვექნება

$$\|U_k\|_m = \left(\sum_{i=0}^{m} \int_0^1 \left| U_k^{(i)}(t) \right|^2 dt \right)^{\frac{1}{2}} = \left(\sum_{i=0}^{m} \int_0^1 \left| (\alpha(t, t_i, h_n) \cdot [Y(t_k) - a(t_k)])^{(i)} \right|^2 dt \right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=0}^{m} \int_0^1 \left| \alpha^{(i)}(t, t_i, h_n) \cdot [Y(t_k) - a(t_k)] \right|^2 dt \right)^{\frac{1}{2}} =$$

$$= |Y(t_k) - a(t_k)| \cdot \left(\sum_{i=0}^{m} \int_0^1 \left| \alpha^{(i)}(t, t_i, h_n) \right|^2 dt \right)^{\frac{1}{2}} \le \frac{|\varepsilon_k| C_\alpha}{n} \cdot \left(\sum_{i=0}^{m} \int_0^1 \left| \frac{1}{h_n^i} \right|^2 dt \right)^{\frac{1}{2}} = \frac{|\varepsilon_k| C_\alpha}{n} \cdot \left(\sum_{i=0}^{m} \frac{1}{h_n^{2i}} \right)^{\frac{1}{2}}$$

$$= \frac{|\varepsilon_k| C_\alpha}{n} \cdot \left(\frac{1 - h_n^{2m+2}}{h_n^{2m}(1 - h_n^2)} \right)^{\frac{1}{2}} \le L \cdot \frac{1}{n h_n^m} = M_m \sim O\left(\frac{1}{n h_n^m} \right) \qquad (\mathbf{4.14})$$

საკმარისად დიდი $L > 0$-თვის.

$r_n(m)$-ის შეფასებისთვის გამოვიყენოთ მაკ-დაიარმიდის უტოლობა და იგი გამოვიყენოთ

$$H(U_1, U_2, \dots, U_n) = \left\| \sum_{k=1}^{n} U_k \right\|_m$$

ფუნქციისთვის.

$$\left| H(U_1, U_2, \dots, U_{i-1}, U_x, U_{i+1}, \dots, U_n) - H\big(U_1, U_2, \dots, U_{i-1}, U_y, U_{i+1}, \dots, U_n\big) \right|$$
$$= \left| \left\| \sum_{k=1, k \ne i}^{n} U_k + U_x \right\|_m - \left\| \sum_{k=1, k \ne i}^{n} U_k + U_y \right\|_m \right|$$
$$\le \left\| \left(\sum_{k=1, k \ne i}^{n} U_k + U_x \right) - \left(\sum_{k=1, k \ne i}^{n} U_k + U_y \right) \right\|_m$$
$$\le \left\| U_x - U_y \right\|_m \le \|U_x\|_m + \left\| U_y \right\|_m \le 2M_m$$

მაშასადამე ($\mathbf{18}$) პირობებში ყოველი $c_k \equiv 2M_m,\ k = 1, \dots, n$. ($\mathbf{19}$) უტოლობის და ($\mathbf{4.14}$) შეფასების თანახმად კი ნებისმიერი $\delta > 0$-თვის გვექნება

$$P\left\{\left|\left\|\sum_{k=1}^{n} U_k\right\|_m - E\left\|\sum_{k=1}^{n} U_k\right\|_m\right| > \delta\right\} \le 2exp\left(-\frac{2\delta^2}{\sum_{k=1}^{n} 4M_m^2}\right) = 2exp\left(-\frac{2\delta^2}{n\cdot 4M_m^2}\right) = 2exp\left(-\frac{2\delta^2 n^2 h_n^{2m}}{n\cdot 4L^2}\right)$$

$$= 2exp\left(-\frac{\delta^2 n h_n^{2m}}{2L^2}\right).$$

აქ ჩავსვათ

$$\delta = \frac{2L\sqrt{logn}}{\sqrt{n}h_n^m}$$

მივიღებთ

$$P\left\{\left|\left\|\sum_{k=1}^{n} U_k\right\|_m - E\left\|\sum_{k=1}^{n} U_k\right\|_m\right| > \delta\right\} \le 2exp\left(-\frac{4L^2 logn\cdot nh_n^{2m}}{nh_n^{2m}\cdot 2L^2}\right) = 2exp(-2logn) = \frac{2}{n^2}.$$

გამოვიყენოთ ბორელი-კანტელის ლემა. მივიღებთ

$$\left\|\sum_{k=1}^{n} U_k\right\|_m = E\left\|\sum_{k=1}^{n} U_k\right\|_m + O\left(\frac{\sqrt{logn}}{\sqrt{n}h_n^m}\right). \qquad (\mathbf{4.15})$$

იენსენის უტოლობის თანახმად

$$\left(E\left\|\sum_{k=1}^{n} U_k\right\|_m\right)^2 \le E\left\|\sum_{k=1}^{n} U_k\right\|_m^2 = E\sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 \left|\alpha^{(i)}(t,t_i,h_n)\cdot[Y(t_k)-a(t_k)]\right|^2 dt$$

$$= \sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 E\left|\alpha^{(i)}(t,t_i,h_n)\cdot[Y(t_k)-a(t_k)]\right|^2 dt \le C_\alpha^2\cdot\sum_{k=1}^{n}\sum_{i=0}^{m}\int_0^1 \frac{1}{n^2\cdot h_n^{2i}}\cdot E[Y(t_k)-a(t_k)]^2 dt$$

$$\le \frac{C_W^2\sigma^2}{n}\cdot\sum_{i=0}^{m}\frac{1}{h_n^{2i}} = \frac{C_W^2\sigma^2}{n}\cdot\frac{1-h_n^{2m+2}}{h_n^{2m}(1-h_n^2)} \le K\cdot\frac{1}{nh_n^{2m}}, \qquad (\mathbf{4.16})$$

($\mathbf{4.12}$), ($\mathbf{4.13}$), ($\mathbf{4.15}$) და ($\mathbf{4.16}$) ფორმულებიდან შეგვიძლია დავასკვნათ, რომ

$$R_n = O\left(\frac{logn}{nh_n^{2m}}\right).$$

ასე, რომ სამართლიანია შემდეგი დებულება.

თეორემა 4.1. ვთქვათ სრულდება $(a1)-(a3),(\varepsilon1)-(\varepsilon2),(\varphi),(W1)-(W4)$ და (h) პირობები. მაშინ, სამართლიანია

$$I(\hat{a}_n)-I(a_n)=s_n(h_n)+R_n$$

ტოლობა, სადაც ნაშთით წევრს ალბათობით 1 აქვს შემდეგი რიგი

$$R_n=O\left(\frac{logn}{nh_n^{2m}}\right). \qquad \mathbf{(4.17)}$$

ძალდებულობა

ამ ნაწილში გამოვიყენებთ თეორემა 4.1-ს და დავამტკიცებთ $I(\hat{a}_n)$ შეფასების მკაცრად ძალდებულობას.

თეორემა 4.2. ვთქვათ, სრულდება თეორემა 4.1-ის პირობები. დადებითი $(h_n)_{n=1}^{\infty}$ მიმდევრობა შერჩეულია ისე, რომ როცა $n\to\infty$

$$\frac{logn}{nh_n^{2m}}\to 0. \qquad \mathbf{(4.18)}$$

მაშინ ალბათობით 1

$$I(\hat{a}_n)\to I(a) \qquad \mathbf{(4.19)}$$

დამტკიცება: თეორემა 4.1-ის და $(\mathbf{4.8})$ ფორმულის თანახმად

$$I(\hat{a}_n)-I(a_n)=S_n(h_n)+R_n \qquad \mathbf{(4.20)}$$

სადაც $R_n=o(1)$ და

$$S_n(h_n)=\sum_{i=0}^{m}\int_0^1\varphi_{(i)}\left(t,a_n(t),a_n'(t),\dots,a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t)-a_n^{(i)}(t)\right)dt.$$

$(a1)$ პირობიდან:

$$\left\{\left(t,a_n(t),a_n'(t),\dots,a_n^{(m)}(t)\right):t\in[0,1]\right\}\subset[0,1]\times[-\Bbbk,\Bbbk]^{m+1}.$$

ამიტომ, $(\varphi2)$ პირობის გათვალისწინებით, მოიძებნება ისეთი $C_\varphi>0$ მუდმივი, რომ

$$sup\left\{\left|\varphi_{(i)}\right|(t,t_0,t_1,\dots,t_m):(t,t_0,t_1,\dots,t_m)\in[0,1]\times[-\Bbbk,\Bbbk]^{m+1}\right\}\le C_\varphi.$$

შეგვიძლია დავწეროთ:

$$ES_n(h_n) = 0$$

$$DS_n(h_n) = ES_n^2(h_n) \le C_\varphi^2 \cdot \sum_{i=0}^{m} \int_0^1 E\left[\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right]^2 dt$$

$$\le C_\varphi^2 \cdot \sum_{i=0}^{m} \int_0^1 \left|\sum_{k=0}^{n} \alpha^{(i)}(t, t_i, h_n) \cdot E[Y(t_k) - a(t_k)]\right|^2 dt$$

$$\le C_\varphi^2 C_\alpha^2 \sigma^2 \cdot \sum_{i=0}^{m} \left(\frac{1}{nh_n^i}\right)^2 \sim C \cdot \frac{1}{nh_n^{2m}} \to 0. \qquad (\mathbf{4.21})$$

რადგან

$$\frac{logn}{nh_n^{2m}} \to 0$$

ამიტომ $S_n(h_n) \to 0$ როცა $n \to \infty$.

$$Ea_n^{(k)}(t) = \int_{-\tau}^{\tau} W(u) a^{(k)}(t)\,(t - uh_n) du + O\left(\frac{1}{nh_n^k}\right) \qquad (\mathbf{4.22})$$

აქედან შეიძლება დავასკვნათ:

i) $(\mathbf{4.17})$ ტოლობიდან გამომდინარეობს, რომ ნულისკენ კრებადია $\frac{1}{nh_n^k}$, ნებისმიერი $k = 0,1, \dots, m$-თვის;

ii) როცა $n \to \infty$ გვაქვს $Ea_n^{(k)}(t) \to a^{(k)}(t)$.

შევაჯამოთ ყველა ნათქვამი და მივიღებთ, რომ $n \to \infty$-თვის

$$I(a_n) = \int_0^1 \varphi\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \to \int_0^1 \varphi\left(t, a(t), a'(t), \dots, a^{(m)}(t)\right) dt = I(a).$$

ასე, რომ $I(\hat{a}_n) - I(a_n) = o(1)$ ანუ $I(\hat{a}_n) - I(a) \to 0$.

თეორემა დამტკიცებულია.

ცენტრალური ზღვარითი თეორემა

ჩვენი წარმოდგენის თეორემის გამოყენებით შეიძლება მივიღოთ ზღვარითი განაწილების თვისება ინტეგრალური ფუნქციონალისთვის

$$I(\hat{a}_n) = \int_0^1 \varphi\left(t, \hat{a}_n(t), \hat{a}_n'(t), \dots, \hat{a}_n^{(m)}(t)\right) dt.$$

განვიხილოთ სხვაობა

$$I(\hat{a}_n) - I(a_n) = s_n(h_n) + R_n, \tag{4.8}$$

სადაც ნებისმიერი $h_n > 0$-თვის, $s_n(h_n)$ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების ჯამს

$$S_n(h_n) = \sum_{i=0}^{m} \int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right) dt. \tag{4.9}$$

R_n ნაშთით წევრს კი აქვს შემდეგი სახე:

$$R_n = \frac{1}{2} \cdot \sum_{i,j=1}^{m} \int_0^1 \varphi_{(ij)}\left(\tilde{b}_m(t)\right)\left(\hat{a}_n^{(i)}(t) - a_n^{(i)}(t)\right)\left(\hat{a}_n^{(j)}(t) - a_n^{(j)}(t)\right) dt. \tag{4.10}$$

ცხადია, რომ

$$ES_n(h_n) = 0 \text{ და } ER_n \to 0 \text{ როცა } n \to \infty. \tag{4.23}$$

გარდა ამისა

$$E\big(S_n(h_n)\big)^2 = \sigma^2 \sum_{i=0}^{m} \left(\int_0^1 \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt\right)^2 \tag{4.24}$$

და $\mathrm{Var}R_n \to 0$ როცა $n \to \infty$.

ამ პირობებისთვის უნდა დავამტკიცოთ, რომ $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)$ სიდიდე ასიმპტოტურად ნორმალურია და უნდა გამოვთვალოთ ზღვარითი ვარიაცია. აგრეთვე, თეორემიდან და (**4.8**), (**4.23**) და (**4.24**) ფორმულებიდან უნდა ვაჩვენოთ $\sqrt{n}S_n(h_n)$ სიდიდის ასიმპტოტურად ნორმალურობა. (**4.10**) ფორმულიდან გამომდინარე საკმარისია ეს თვისება შევისწავლოთ შემდეგი სიდიდისთვის:

$$d_k = Y(t_k) \cdot \sum_{i=0}^{m} \int_0^1 \alpha^{(i)}(t, t_i, h_n) \cdot \varphi_{(i)}\left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \tag{4.25}$$

ადვილი შესამჩნევია, რომ

$$Ed_k = a(t_k) \cdot \sum_{i=0}^{m} \int_0^1 \alpha^{(i)}(t, t_i, h_n) \cdot \varphi_{(i)} \left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \qquad (\mathbf{4.26})$$

ანუ ვიხილავთ დამოუკიდებელი შემთხვევითი სიდიდეების მიმდევრობას:

$$f_n(k) = \alpha(n,k)\big(Y(t_k) - a(t_k)\big) = \alpha(n,k)\varepsilon_k,$$

სადაც

$$\alpha(n,k) = \sum_{i=0}^{m} \int_0^1 \alpha^{(i)}(t, t_i, h_n) \cdot \varphi_{(i)} \left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt.$$

განვიხილოთ ჯამი $S_n(h_n) = \sum_{k=1}^{n} \alpha(n,k)\varepsilon_k$.

ვთქვათ, $F_{k,n}$ არის $\alpha(n,k)\varepsilon_k$ შემთხვევითი სიდიდეების ალბათური განაწილების ფუნქცია, F_ε – კი ε_k შემთხვევითი სიდიდეების ალბათური განაწილების ფუნქცია. ლინდებერგის პირობას აქვს შემდეგი სახე $\forall\delta > 0$, $lim_{n\to\infty} L_n(\delta) = 0$, სადაც

$$L_n(\delta) = \frac{\sum_{j=1}^{n} \int x^2 J\left(|x| \geq \delta\sigma(\sum_{k=1}^{n} \alpha^2(n,k))^{1/2}\right) dF_{k,n}(x)}{\sigma^2 \sum_{k=1}^{n} \alpha^2(n,k)}$$

აქ $J(A)$ წარმოადგენს A სიმრავლის ინდიკატორ ფუნქციას. ადვილი დასანახია, რომ

$$L_n(\delta) \leq \frac{1}{\sigma^2} max_{1\leq j\leq n} \int x^2 J\big(|x| \geq \delta\sigma v(n,j)\big) dF_\varepsilon$$

სადაც

$$v(n,j) = \frac{|\alpha(n,j)|}{\left(\sum_{j=1}^{n} \alpha^2(n,j)\right)^{1/2}}.$$

საჩვენებელი რჩება, რომ $max_{1\leq j\leq n} v(n,j) \to 0$ როცა $n \to \infty$. მაგრამ რადგანაც

$$max_{1\leq j\leq n} |\alpha(n,j)| = O\left(\frac{1}{nh_n^m}\right),$$

ამიტომ

$$max_{1\leq j\leq n} v(n,j) = O\left(\frac{1}{\sqrt{n}}\right).$$

ასე, რომ ლინდერბერგის პირობა სრულდება და შეგვიძლია ჩამოვაყალიბოთ

თეორემა 4.3: ვთქვათ სრულდება თეორემა 4.1-ის პირობები. მაშინ, თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m} \to \infty$ როცა $n \to \infty$, სამართლიანია $\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2)$. სადაც

$$r^2 = \sigma^2 \cdot \sum_{i=0}^{m} \left(\int_0^1 \varphi_{(i)} \left(t, a_n(t), a_n'(t), \dots, a_n^{(m)}(t)\right) dt \right)^2 .$$

ზოგიერთი გამოყენება

განვიხილოთ ინტეგრალური ფუნქციონალი

$$I_1(a) = \int_0^1 a^2(t)\, dt$$

ამ შემთხვევაში $\varphi(t, x_0, x_1, \dots, x_m) = x_0^2$ როცა $x_0 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. მაშინ, თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m} \to \infty$ როცა $n \to \infty$, გვექნება კრებადობა

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = 4\sigma^2 \left(\int_0^1 a(t) dt \right)^2 .$$

განვიხილოთ ფიშერის ინფორმაციული ინტეგრალი

$$I_2(a) = \int_0^1 \frac{\big(a'(t)\big)^2}{a(t)} dt.$$

გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \frac{x_1^2}{x_2}$. მაშინ $t \in [0,1] => a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. თუ $h_n \to 0$ და $\sqrt{n}h_n^{2m} \to \infty$ როცა $n \to \infty$, მაშინ

$$\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big) \to_d N(0, r^2),$$

სადაც

$$r^2 = \sigma^2 \cdot \left(-\int_0^1 \left(\frac{\left(a'(t)\right)^2}{\left(a(t)\right)^2} - \frac{2a'(t)}{a(t)} \right) dt \right)^2$$

$$= \sigma^2 \left(\frac{a'(1)}{a(1)} loga(1) - \frac{a'(0)}{a(0)} loga(0) - \frac{1}{2} \cdot \frac{\left(a'(1)\right)^2}{\left(a(1)\right)^2} + \frac{1}{2} \cdot \frac{\left(a'(0)\right)^2}{\left(a(0)\right)^2} \right)^2 .$$

ახლა განვიხილოთ

$$I_3(a) = \int_{-\infty}^{\infty} \left(a(t)\right)^s dt$$

ფუნქციონალი ($s > 1$). აქ $\varphi(t, x_0, x_1, \dots, x_m) = x_0^S$, როცა $x_1 \in [-b, b] \supset [-\mathbb{k}, \mathbb{k}], b > 0$. თუ$h_n \to 0$ და $\sqrt{n} h_n^{2m} \to \infty$ როცა $n \to \infty$, მაშინ

$$\sqrt{n}\left(I(\hat{a}_n) - I(a_n)\right) \to_d N(0, r^2),$$

სადაც

$$r^2 = s^2 \sigma^2 \left(\int_0^1 a^{s-1}(t) dt \right)^2 .$$

ბოლოს განვიხილოთ შენონის ინტეგრალური ენტროპია

$$I_4(a) = \int_{-\infty}^{\infty} a(t) loga(t) dt.$$

საკმარისად დიდი $b \geq \mathbb{k} > 0$-თვის თუ $0 < x_0 \leq b$ გვაქვს $\varphi(t, x_0, x_1, \dots, x_m) = \varphi(x_0) = x_0 log x_0$. დამატებით განვსაზღვროთ φ ფუნქცია 0-მდე, $\varphi(x) = 0$ როცა $-b \leq x \leq 0$. მაშინ $t \in [0,1] \Rightarrow a(t) \in [\mathbb{a}, \mathbb{b}], \mathbb{b} > a > 0$ და $b \geq \mathbb{b}$. თუ $h_n \to 0$ და $\sqrt{n} h_n^{2m} \to \infty$ როცა $n \to \infty$, მაშინ

$$\sqrt{n}\left(I(\hat{a}_n) - I(a_n)\right) \to_d N(0, r^2),$$

სადაც

$$r^2 = \sigma^2 \left(\int_0^1 a(t)\left(1 + loga(t)\right) dt \right)^2 .$$

განმეორებითი ლოგარითმის თვისება

გამოვიყენოთ კიულბსის ცნობილი განმეორებითი ლოგარითმის თვისება და ჩამოვაყალიბოთ შემდეგი

თეორემა 4.4. თუ h_n მიმდევრობა შერჩეულია ისე, რომ

$$R_n = o\left(\sqrt{\frac{loglogn}{n}}\right),$$

მაშინ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = r.$$

დამტკიცება. მართლაც, ადვილი დასანახია, რომ

$$\lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(I(\hat{a}_n) - I(a_n)\big)}{\sqrt{2loglogn}} = \lim_{n\to\infty} sup \pm \frac{\sqrt{n}\big(\alpha(n,k)Y(t_k) - \alpha(n,k)a(t_k)\big)}{\sqrt{2loglogn}} = r.$$

§5. სიმულაციები

ამ პარაგრაფში წარმოდგენილია რიცხვითი სიმულაციები ცხრილების და გრაფიკების სახით. პირველ და მეორე ცხრილებში მოცემულია რიცხვითი სიმულაციის შედეგები იმ შეთხვევებისთვის, როცა $n = 10000$, $h_n = \frac{1}{\sqrt{n}}$, $W(t)$ ალბათური სიმკვრივის თვისებების მქონე ფუნქცია $[0, 1]$ შუალედში იღებს $W(t) = -6t(t-1)$ მნიშვნელობებს, სხვაგან კი უდრის 0-ს. შედეგზე დაკვირვებისთვის აღებულია წინასწარ ცნობილი $a(t) = 36t\left(t-\frac{1}{3}\right)\left(t-\frac{2}{3}\right)$ ფუნქცია.

ცხრილი 1.

t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
ზუსტი	0.000	0.476	0.448	0.132	-0.25	-0.50	-0.38	0.308	1.792	4.284	8.000	
P-Ch რეგრესია	0.000	0.685	0.340	0.213	-0.21	-0.52	-0.42	0.192	1.586	4.072	7.776	
Error P-Ch	0.000	0.209	0.108	0.081	0.038	0.028	0.045	0.116	0.206	0.212	0.224	1.26653
N-W რეგრესია	0.000	0.685	0.340	0.213	-0.21	-0.52	-0.42	0.192	1.586	4.072	7.777	
Error N-W	0.000	0.209	0.108	0.081	0.038	0.028	0.045	0.116	0.206	0.212	0.223	1.26530

ცხრილი 2.

t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
ზუსტი	0.000	0.476	0.448	0.132	-0.25	-0.50	-0.38	0.308	1.792	4.284	8.000	
P-Ch რეგრესია	0.000	0.491	0.389	-0.03	-0.28	-0.50	-0.55	0.181	1.716	4.107	7.820	
Error P-Ch	0.000	0.015	0.059	0.169	0.026	0.004	0.166	0.127	0.076	0.177	0.180	0.999138
N-W რეგრესია	0.000	0.492	0.389	-0.03	-0.28	-0.50	-0.55	0.18	1.716	4.107	7.821	
Error N-W	0.000	0.016	0.059	0.169	0.026	0.004	0.166	0.127	0.076	0.177	0.179	0.997903

მესამე და მეოთხე ცხრილებში მოცემული რიცხვითი სიმულაციის შედეგების მისაღებად აღებულია შეთხვევები, როცა $n = 10000,\ h_n = \frac{1}{\sqrt{n}}$, $W(t)$ ალბათური სიმკვრივის თვისებების მქონე ფუნქცია $[0, 1]$ შუალედში იღებს $W(t) = -6t(t-1)$ მნიშვნელობებს, სხვაგან კი უდრის 0-ს. შედეგის ანალიზისათვის აღებულია წინასწარ ცნობილი $a(t) = e^t$ ფუნქცია.

ცხრილი 3.

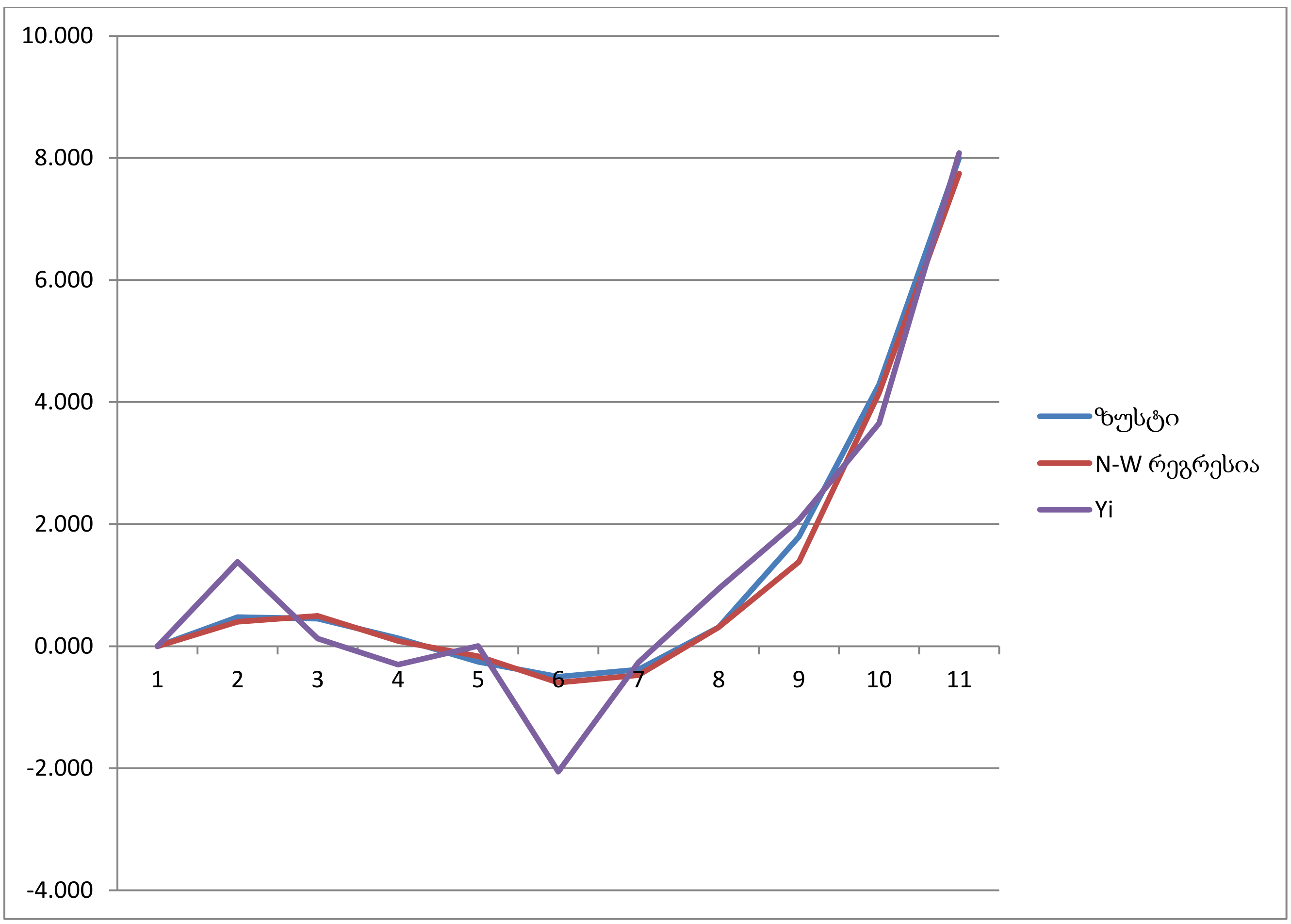
t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
ზუსტი	1.000	1.105	1.221	1.350	1.492	1.649	1.822	2.014	2.226	2.460	2.718	
P-Ch რეგრესია	1.000	1.203	1.292	1.366	1.542	1.666	1.765	2.076	2.217	2.478	2.714	
Error P-Ch	0.000	0.098	0.071	0.016	0.050	0.017	0.057	0.062	0.008	0.018	0.004	0.40161
N-W რეგრესია	1.000	1.203	1.292	1.366	1.542	1.666	1.765	2.076	2.218	2.478	2.715	
Error N-W	0.000	0.098	0.071	0.016	0.050	0.017	0.057	0.062	0.008	0.018	0.004	0.40200

ცხრილი 4.

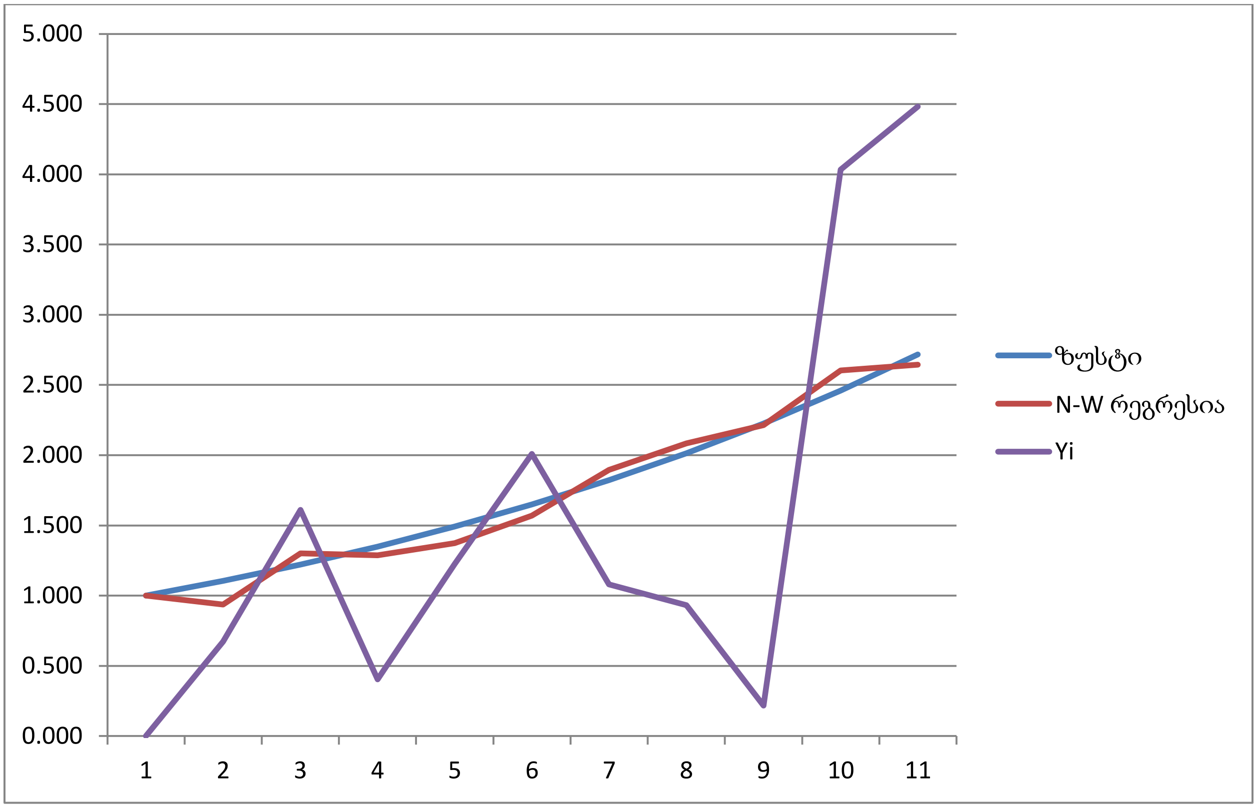
t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
ზუსტი	1.000	1.105	1.221	1.350	1.492	1.649	1.822	2.014	2.226	2.460	2.718	
P-Ch რეგრესია	1.000	1.155	1.162	1.326	1.299	1.476	1.824	2.038	2.129	2.414	2.698	
Error P-Ch	0.000	0.050	0.059	0.024	0.193	0.172	0.002	0.024	0.096	0.046	0.021	0.68672
N-W რეგრესია	1.000	1.156	1.162	1.326	1.299	1.477	1.824	2.038	2.129	2.414	2.698	
Error N-W	0.000	0.050	0.059	0.024	0.193	0.172	0.002	0.024	0.096	0.045	0.020	0.68587

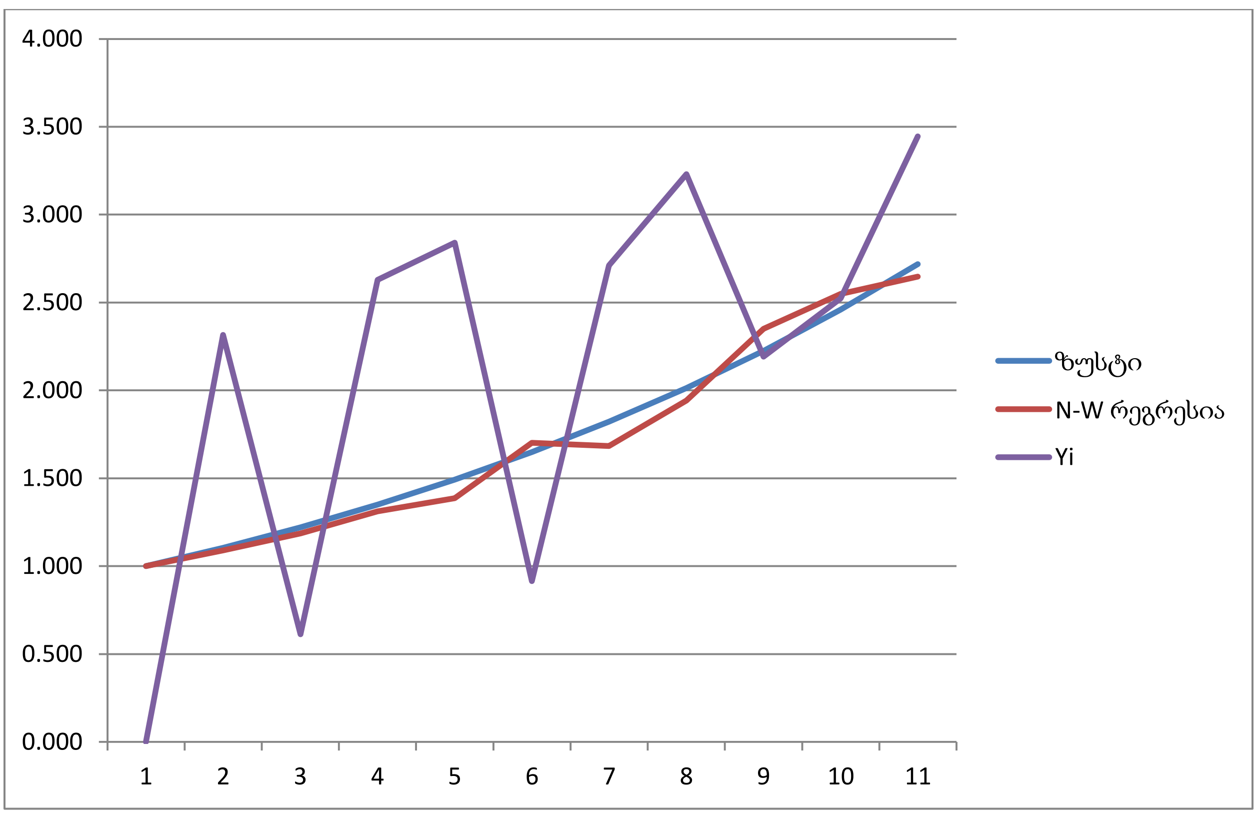
პირველ გრაფიკის შედეგების მისაღებად მოცემული რიცხვითი სიმულაციისათვის აღებულია $n = 10000$, $h_n = \frac{1}{\sqrt{n}}$, $W(t)$ ალბათური სიმკვრივის თვისებების მქონე ფუნქცია $[0, 1]$ შუალედში იღებს $W(t) = -6t(t-1)$ მნიშვნელობებს, სხვაგან კი უდრის 0-ს. შედეგზე დაკვირვებისთვის აღებულია წინასწარ ცნობილი $a(t) = 36t\left(t - \frac{1}{3}\right)\left(t - \frac{2}{3}\right)$ ფუნქცია, რომელიც გამოსახულია ლურჯი ფერით. y_i დაკვირვებები აღებულია იასამნისფრად, ნადარაია-ვოტსონის შედეგი კი - წითლად. მეორე და მესამე გრაფიკების შედეგების მისაღებად კი აღებულია წინასწარ ცნობილი $a(t) = e^t$ ფუნქცია.

გრაფიკი 1.



გრაფიკი 2-3.


5.000
4.500
4.000
3.500
3.000
2.500
2.000
1.500
1.000
0.500
0.000
1 2 3 4 5 6 7 8 9 10 11
ზუსტი
N-W რეგრესია
Yi

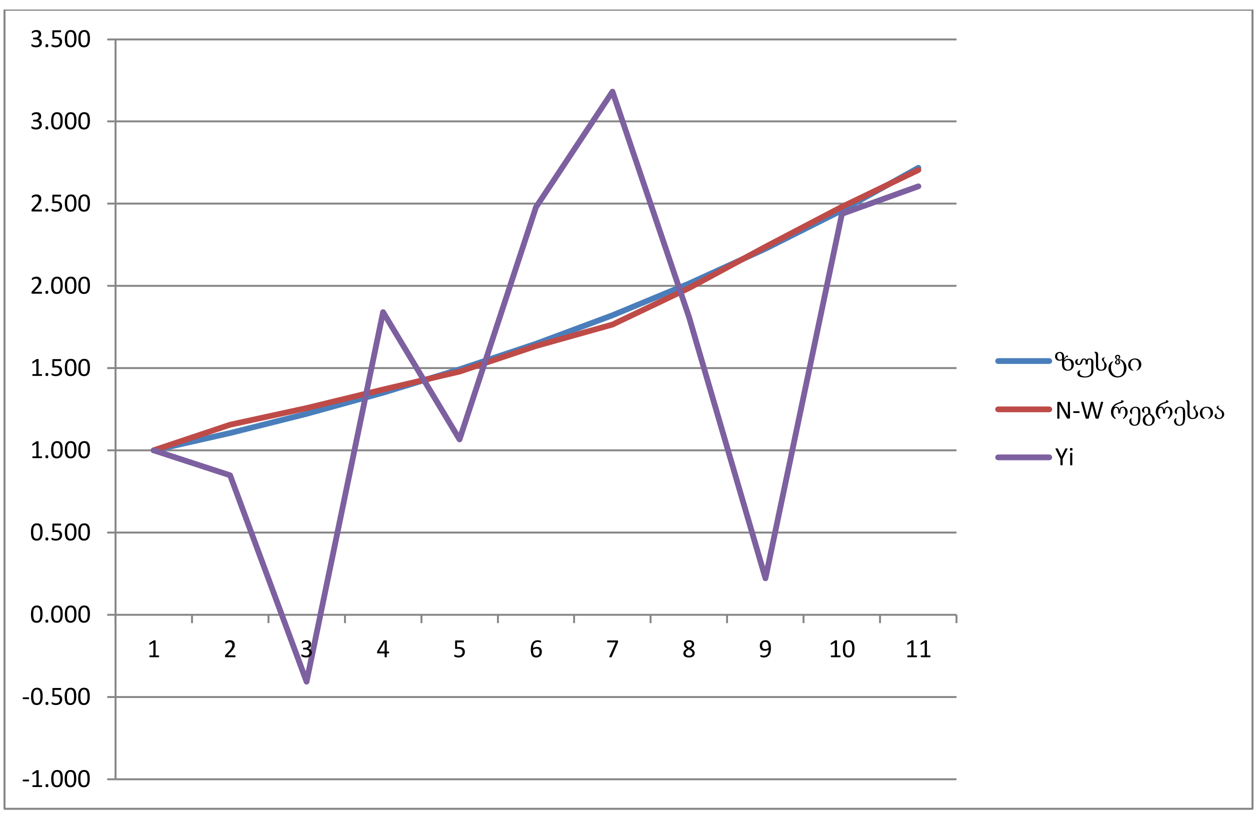

4.000
3.500
3.000
2.500
2.000
1.500
1.000
0.500
0.000
1 2 3 4 5 6 7 8 9 10 11
ზუსტი
N-W რეგრესია
Yi

მეხუთე ცხრილის და მეოთხე გრაფიკის შედეგების მისაღებად მოცემული რიცხვითი სიმულაციისათვის ისევ აღებულია $n = 10000$, $h_n = \frac{1}{\sqrt{n}}$, $W(t) = e^{-\frac{t^2}{2}}$. შედეგზე დაკვირვებისთვის აღებულია წინასწარ ცნობილი $a(t) = e^t$ ფუნქცია, რომელიც გამოსახულია ლურჯი ფერით. y_i დაკვირვებები აღებულია იასამნისფრად, ნადარაია-ვოტსონის შედეგი კი - წითლად.

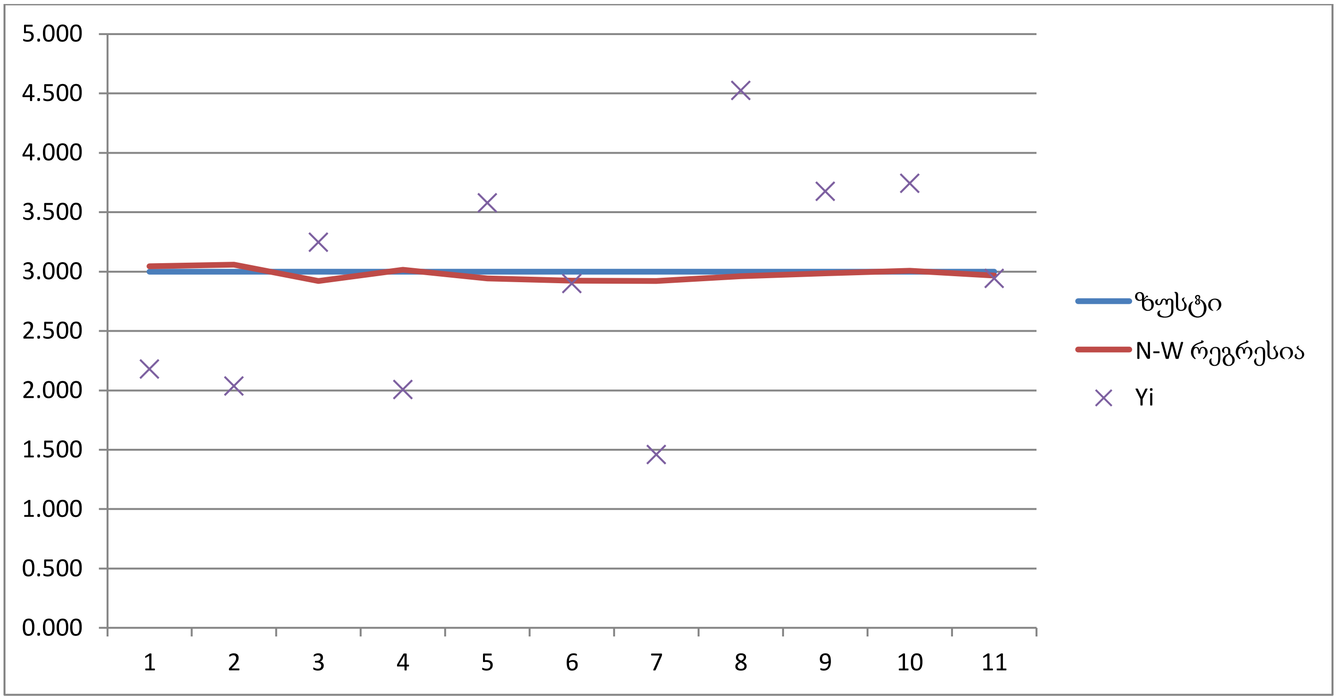
ცხრილი 5.

t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
ზუსტი	1.000	1.105	1.221	1.350	1.492	1.649	1.822	2.014	2.226	2.460	2.718	
N-W რეგრესია	1.000	1.156	1.257	1.369	1.478	1.634	1.764	1.988	2.237	2.482	2.706	
Error N-W	0.000	0.051	0.035	0.019	0.014	0.015	0.058	0.026	0.012	0.022	0.013	0.265342

გრაფიკი 4.



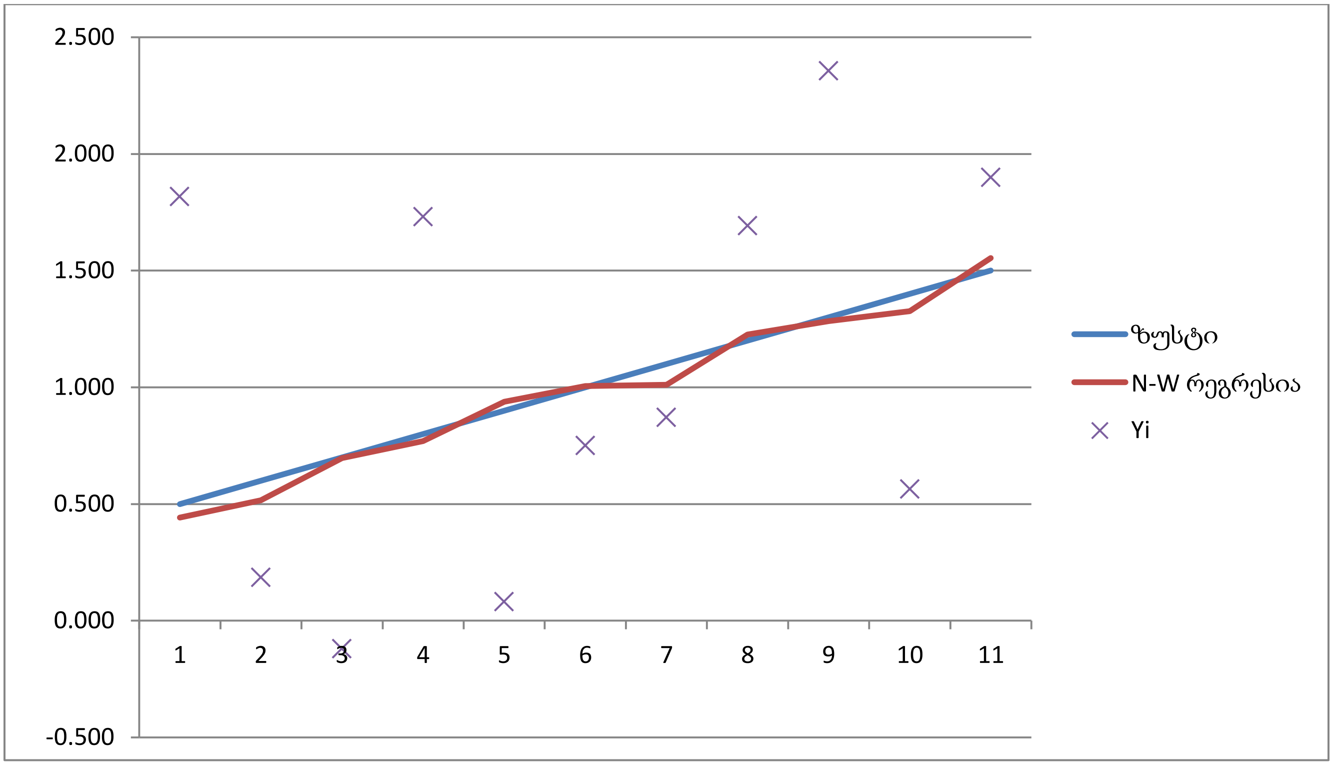
შემდეგში ყველგან აღებულია $n = 10000$, $h_n = \frac{1}{\sqrt{n}}$, $W(t) = e^{-\frac{t^2}{2}}$. მომდევნო სიმულაციებში შეიცვლება მხოლოდ $a(t)$ ფუნქცია. აქ $a(t) = 3$. დათვლილია ინტეგრალები.



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000	3.000
N-W რეგრესია	2.994	2.977	2.945	2.973	3.013	2.944	3.030	2.953	2.968	3.016	3.064
Error N-W	0.006	0.023	0.055	0.027	0.013	0.056	0.030	0.047	0.032	0.016	0.064
Yi	5.9807	2.9326	4.5246	1.4562	3.513	4.1436	3.6009	2.6189	3.3913	2.0717	1.7357

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	3.000	3.000	3.000
ინტეგრალის მიახლოებითი მნიშვნელობა	2.988	2.985	2.981
Error	0.012	0.015	0.019

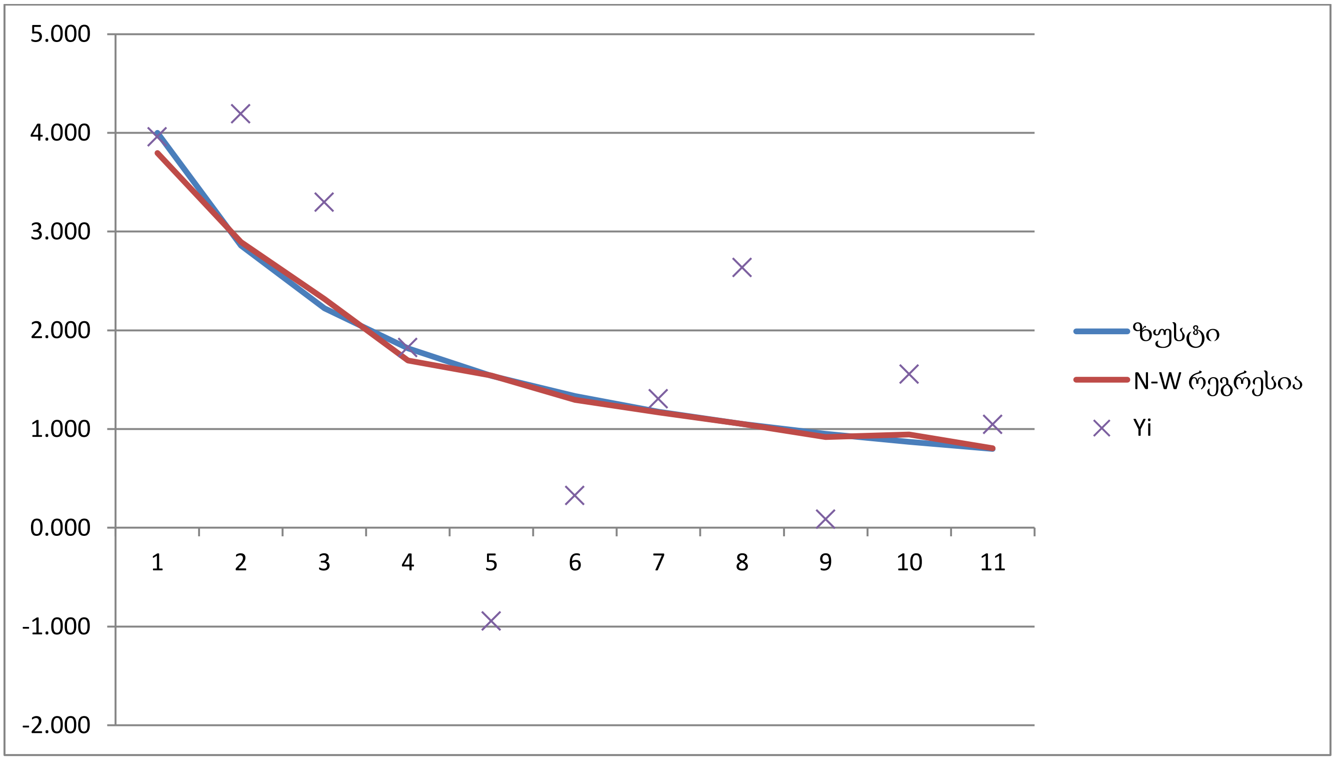
$$a(t) = t + \frac{1}{2}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.500	0.600	0.700	0.800	0.900	1.000	1.100	1.200	1.300	1.400	1.500
N-W რეგრესია	0.512	0.579	0.668	0.904	0.871	1.044	1.025	1.207	1.305	1.286	1.401
Error N-W	0.012	0.021	0.032	0.104	0.029	0.044	0.075	0.007	0.005	0.114	0.099
Yi	0.6519	1.2085	2.041	3.0673	0.323	1.0871	1.1709	1.8588	1.6032	5.0602	2.3988

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	1.000	1.000	1.000
ინტეგრალის მიახლოებითი მნიშვნელობა	1.029	0.985	0.991
Error	0.029	0.015	0.009

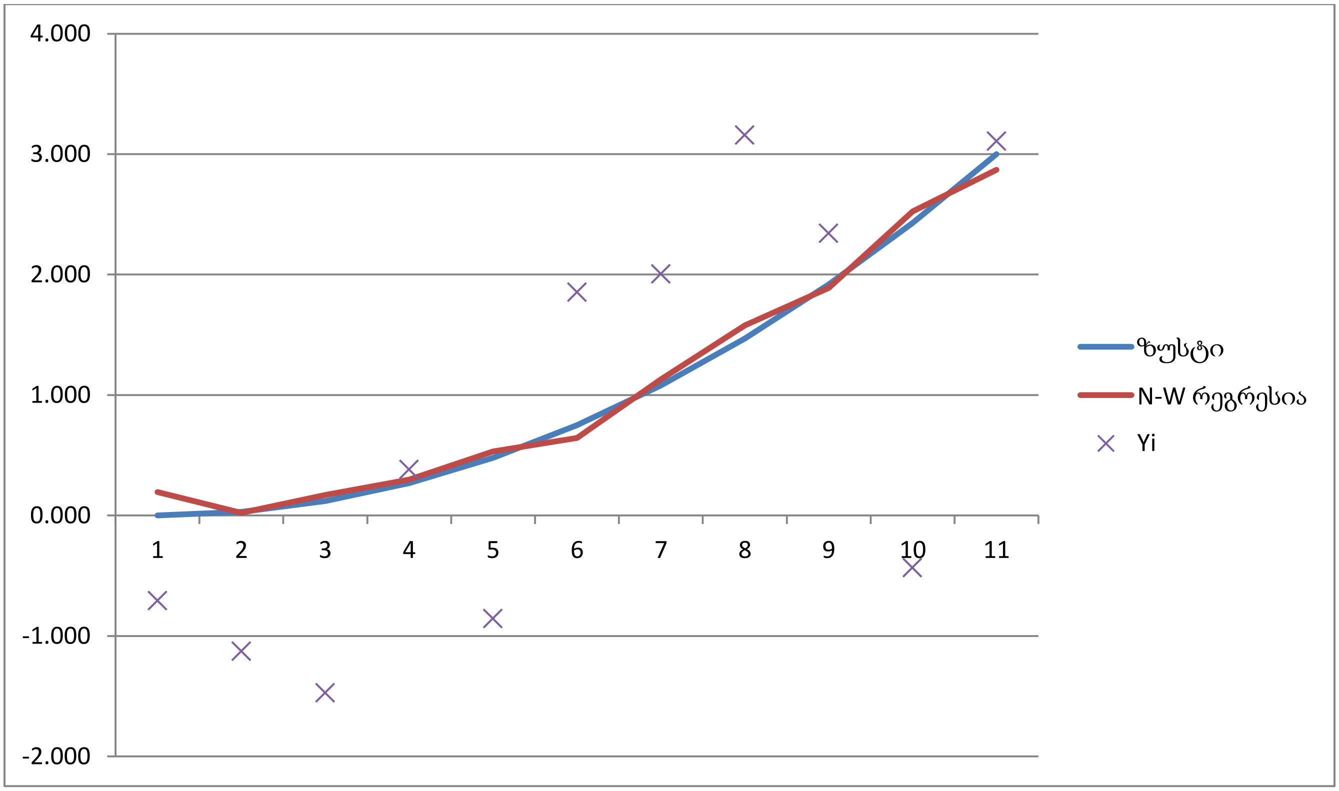
$$a(t) = \frac{1}{t + 0.25}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	3.998	2.857	2.222	1.818	1.538	1.333	1.176	1.053	0.952	0.870	0.800
N-W რეგრესია	3.777	2.827	2.231	1.794	1.510	1.384	1.268	1.023	1.009	0.838	0.837
Error N-W	0.221	0.030	0.009	0.024	0.028	0.051	0.091	0.030	0.057	0.032	0.037
Yi	3.7327	3.1541	2.4448	1.4732	1.4324	2.8709	2.085	0.1121	1.9688	1.0938	-0.32

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	1.609	1.609	1.609
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.472	1.619	1.604
Error	0.137	0.010	0.006

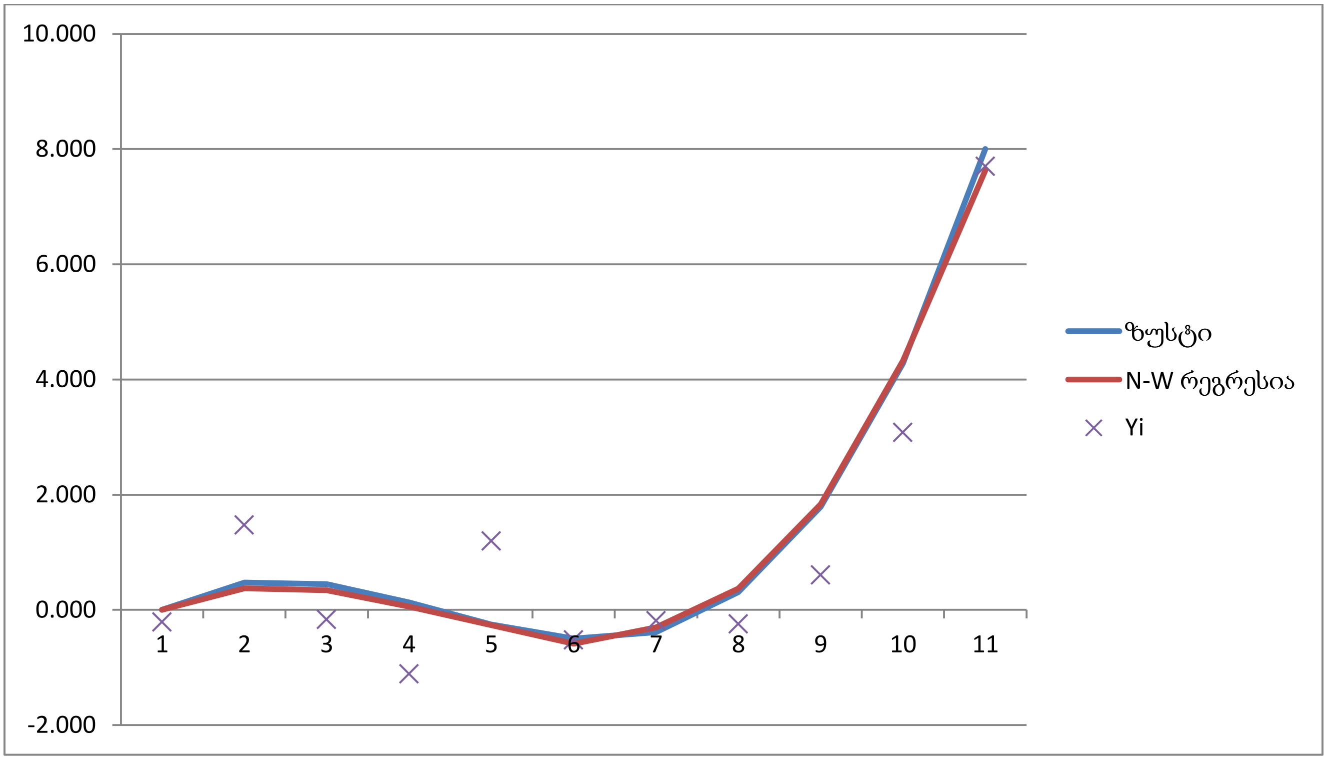
$$a(t) = 3t^2$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.000	0.030	0.120	0.270	0.480	0.750	1.080	1.470	1.920	2.430	3.000
N-W რეგრესია	-0.085	-0.078	0.162	0.360	0.537	0.773	1.072	1.406	1.946	2.427	3.033
Error N-W	0.085	0.108	0.042	0.090	0.057	0.023	0.008	0.064	0.026	0.003	0.033
Yi	0.6367	0.7697	0.1491	0.5213	0.2029	0.2029	0.5358	2.5384	2.078	4.0377	1.4507

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	1.000	1.000	1.000
ინტეგრალის მიახლოებითი მნიშვნელობა	1.110	0.976	0.970
Error	0.110	0.024	0.030

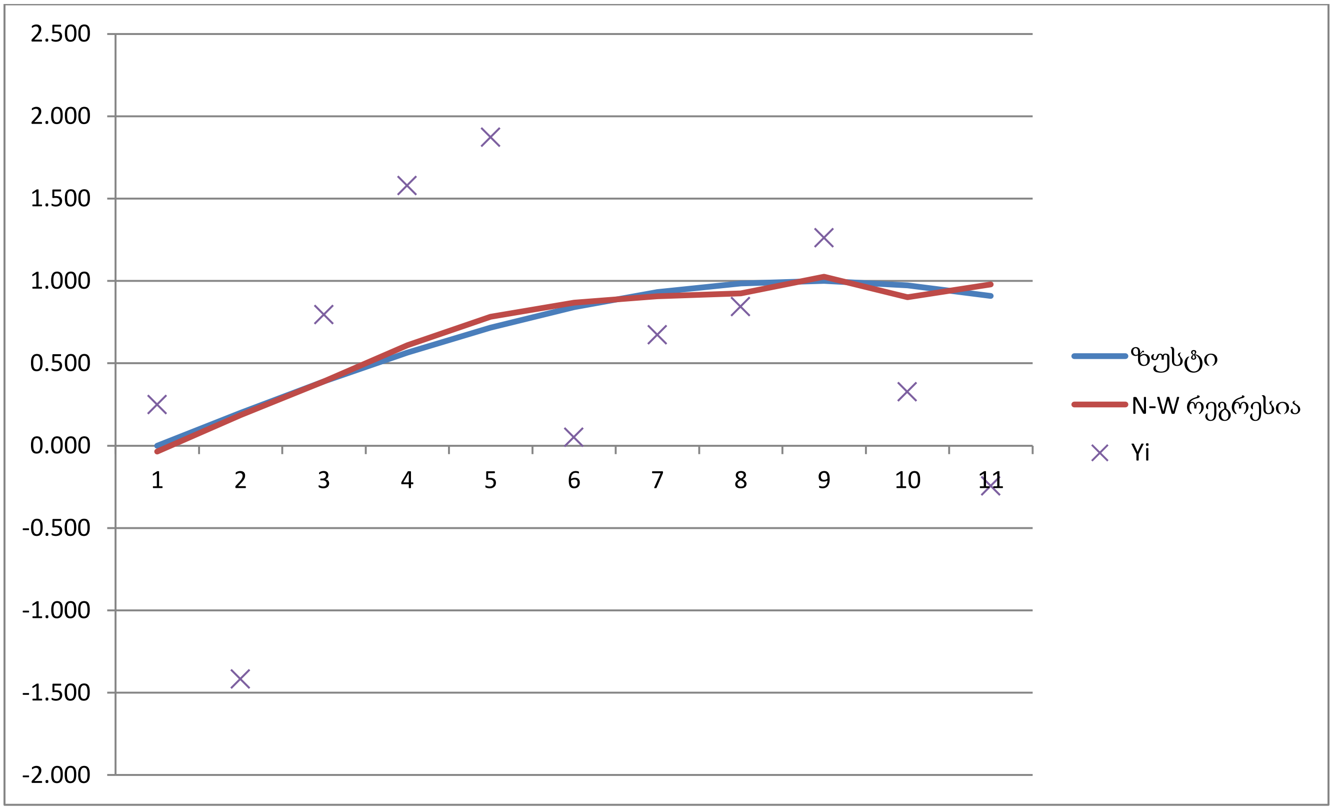
$$a(t) = 35t\left(t-\frac{1}{3}\right)\left(t-\frac{2}{3}\right)$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.001	0.476	0.448	0.132	-0.25	-0.50	-0.384	0.308	1.792	4.284	8.000
N-W რეგრესია	0.155	0.469	0.441	0.160	-0.22	-0.42	-0.406	0.248	1.849	4.317	7.577
Error N-W	0.154	0.007	0.007	0.028	0.036	0.075	0.022	0.060	0.057	0.033	0.423
Yi	0.7145	0.2767	0.9731	-0.83	-0.91	-0.23	1.0173	1.7485	0.8874	3.6088	7.9592

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	1.000	1.000	1.000
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.401	1.030	1.005
Error	0.401	0.030	0.005

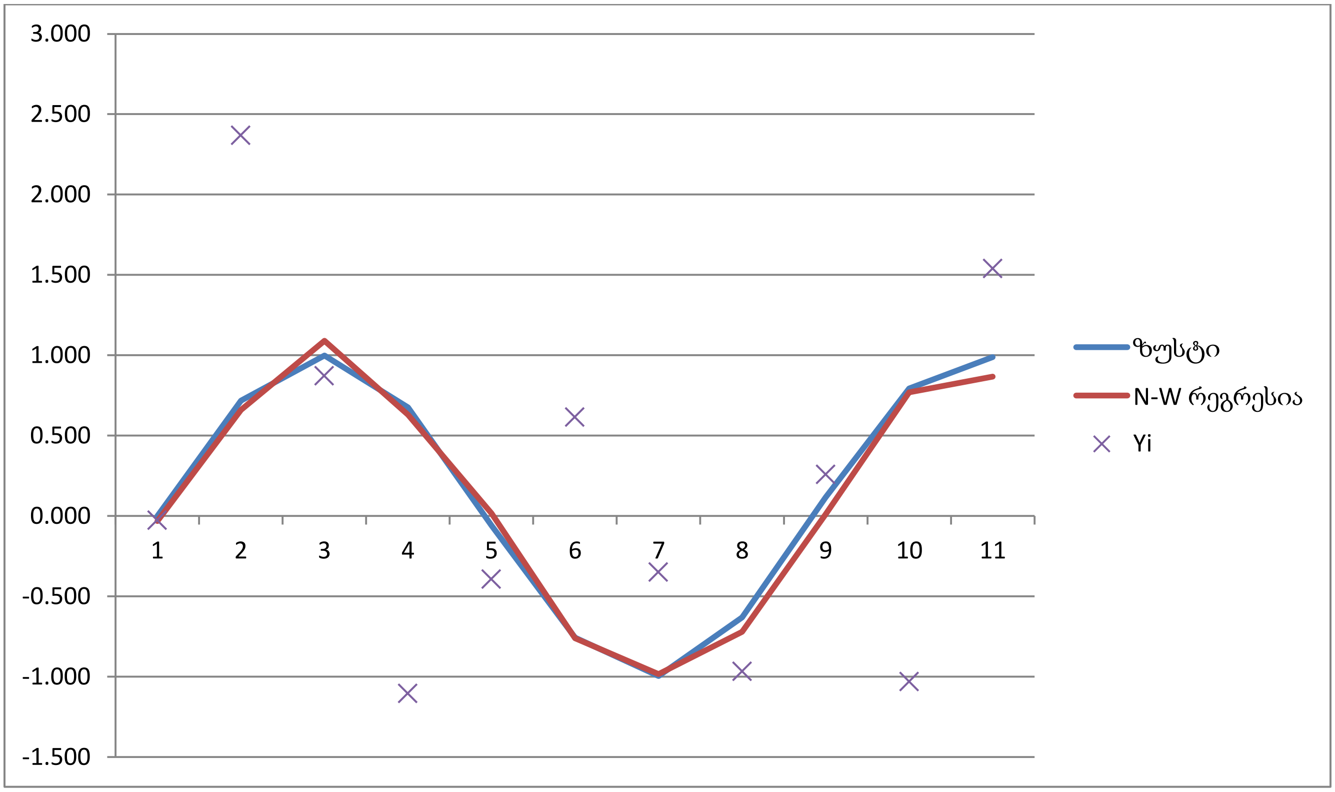
$$a(t) = sin2t$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.00 0	0.199	0.389	0.565	0.717	0.841	0.932	0.985	1.000	0.974	0.909
N-W რეგრესია	-0.00	0.203	0.487	0.493	0.611	0.920	1.051	0.902	1.007	0.981	0.999
Error N-W	0.00 4	0.004	0.097	0.072	0.106	0.078	0.119	0.083	0.007	0.008	0.090
Yi	-0.63	-0.03	-0.30	1.238	0.752	1.376	2.132	1.840	1.091	1.115	-0.44

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	0.708	0.708	0.708
ინტეგრალის მიახლო-ებითი მნიშვნელობა	0.765	0.715	0.710
Error	0.057	0.007	0.002

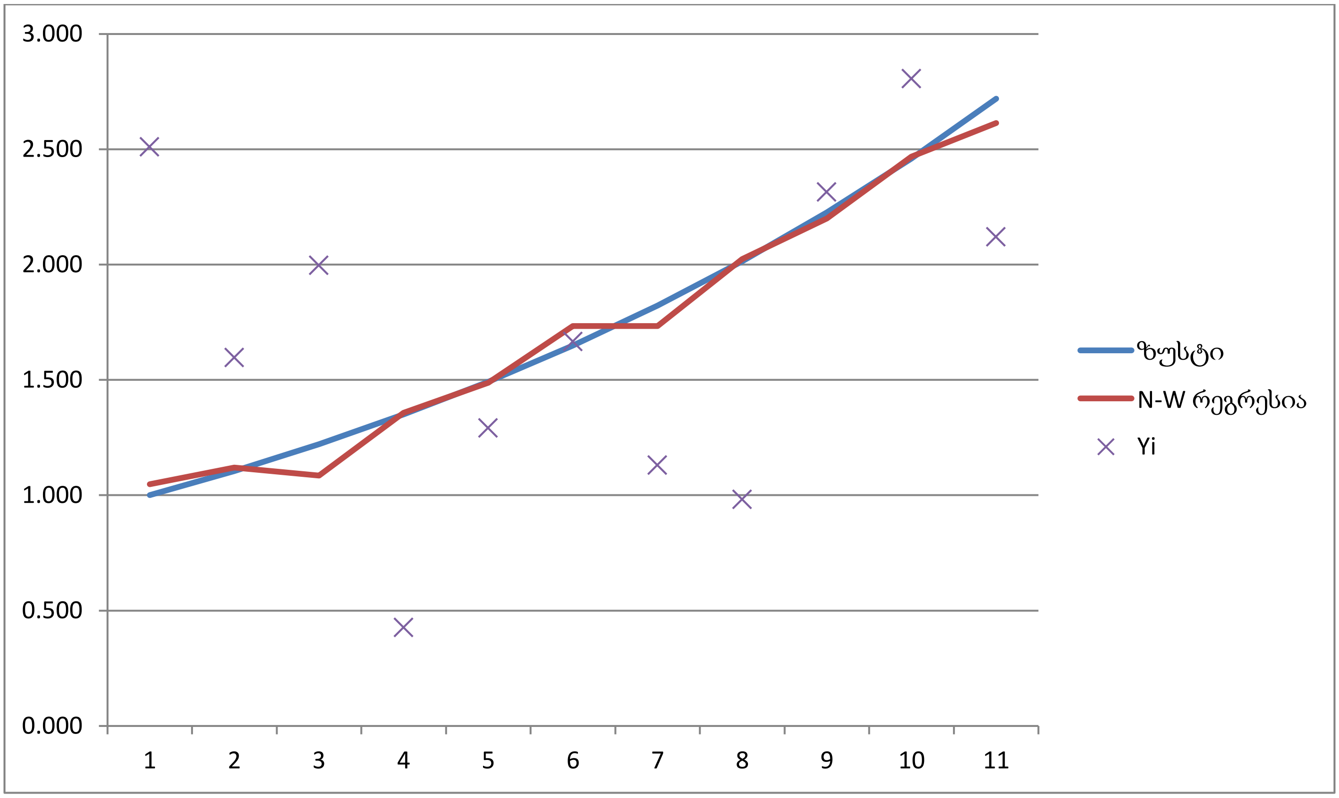
$$a(t) = sin8t$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.001	0.717	1.000	0.675	-0.058	-0.757	-0.99	-0.631	0.117	0.794	0.989
N-W რეგრესია	0.078	0.810	0.968	0.672	-0.024	-0.817	-0.99	-0.613	0.230	0.729	1.041
Error N-W	0.077	0.093	0.031	0.003	0.035	0.060	0.001	0.018	0.113	0.065	0.051
Yi	-2.34	2.3544	0.2278	-0.46	0.8752	0.6589	-1.23	1.7239	-0.36	2.1472	2.0379

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	0.143	0.143	0.143
ინტეგრალის მიახლო-ებითი მნიშვნელობა	0.177	0.126	0.132
Error	0.034	0.017	0.011

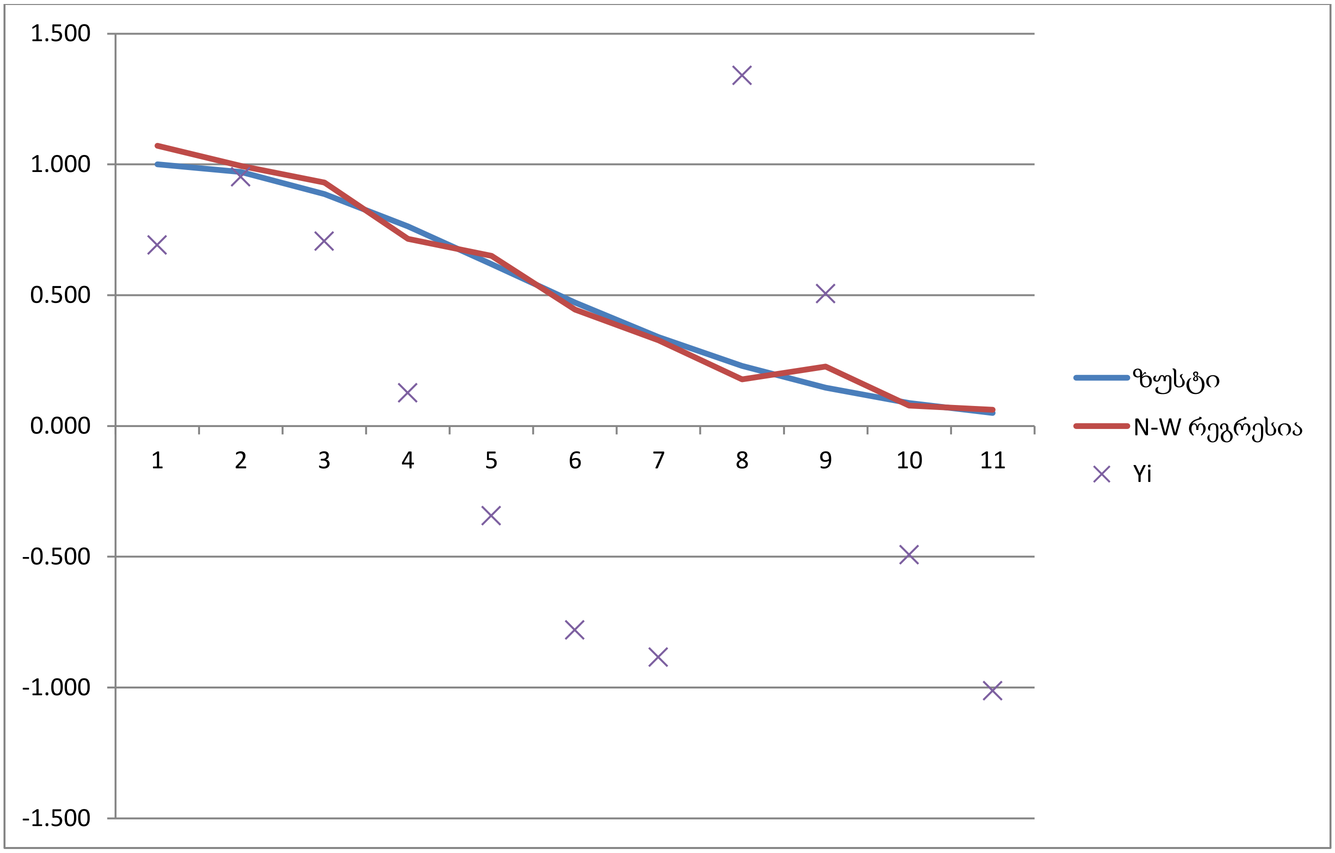
$$a(t) = e^t$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	1.000	1.105	1.221	1.350	1.492	1.649	1.822	2.014	2.226	2.460	2.718
N-W რეგრესია	0.984	1.168	1.252	1.289	1.402	1.710	1.788	2.015	2.252	2.479	2.833
Error N-W	0.016	0.063	0.031	0.061	0.090	0.061	0.034	0.001	0.027	0.019	0.115
Yi	1.6574	0.2806	0.0911	0.8953	3.2928	1.8217	2.9214	0.6838	3.1579	0.9861	2.8968

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ზუსტი ინტეგრალი	1.718	1.718	1.718
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.806	1.726	1.718
Error	0.088	0.007	0.001

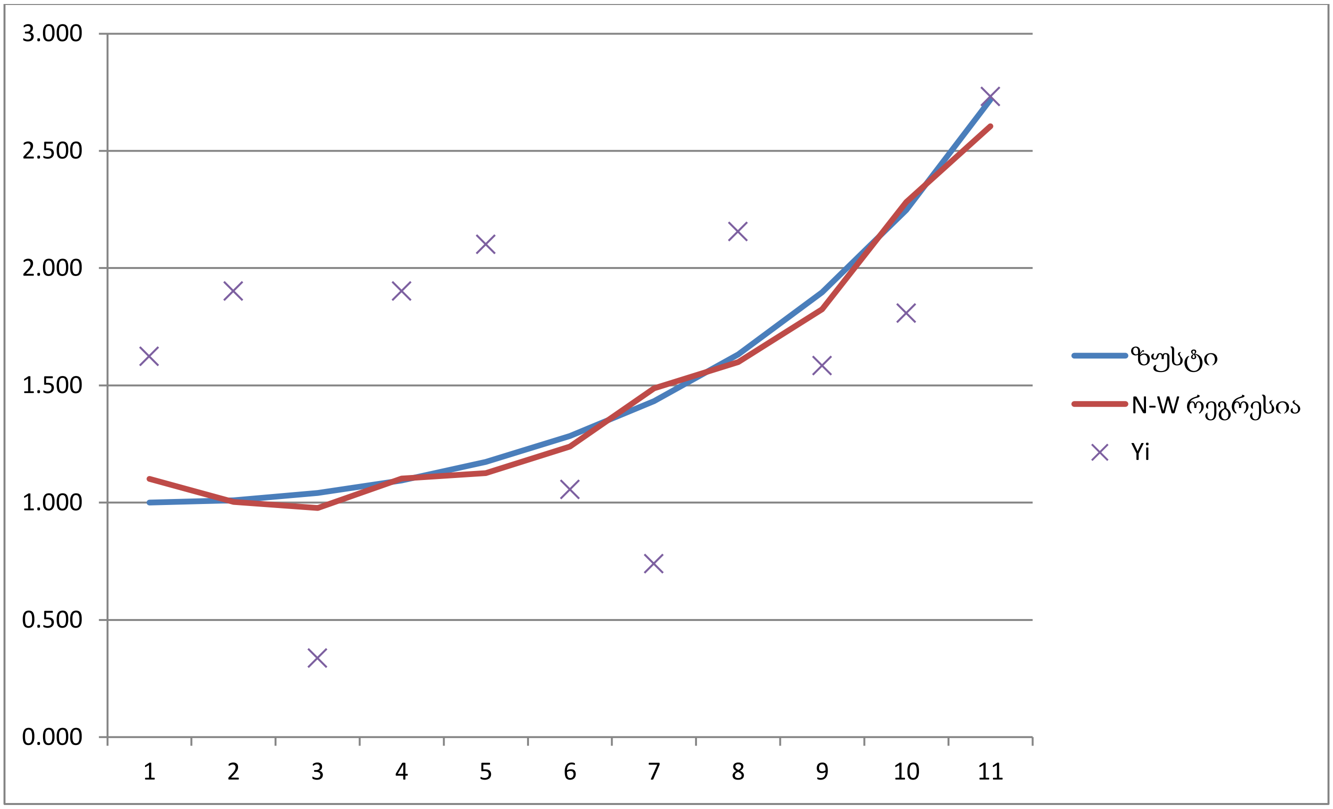
$$a(t) = e^{-3t^2}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	1.000	0.970	0.887	0.763	0.619	0.472	0.340	0.230	0.147	0.088	0.050
N-W რეგრესია	0.927	0.832	0.807	0.704	0.571	0.503	0.313	0.209	0.205	0.076	-0.02
Error N-W	0.073	0.139	0.080	0.059	0.047	0.031	0.027	0.021	0.058	0.012	0.073
Yi	0.8762	0.2675	0.3225	1.8612	1.1181	1.9325	1.5344	0.2879	-0.46	-1.36	-0.45

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	0.450	0.506	0.509

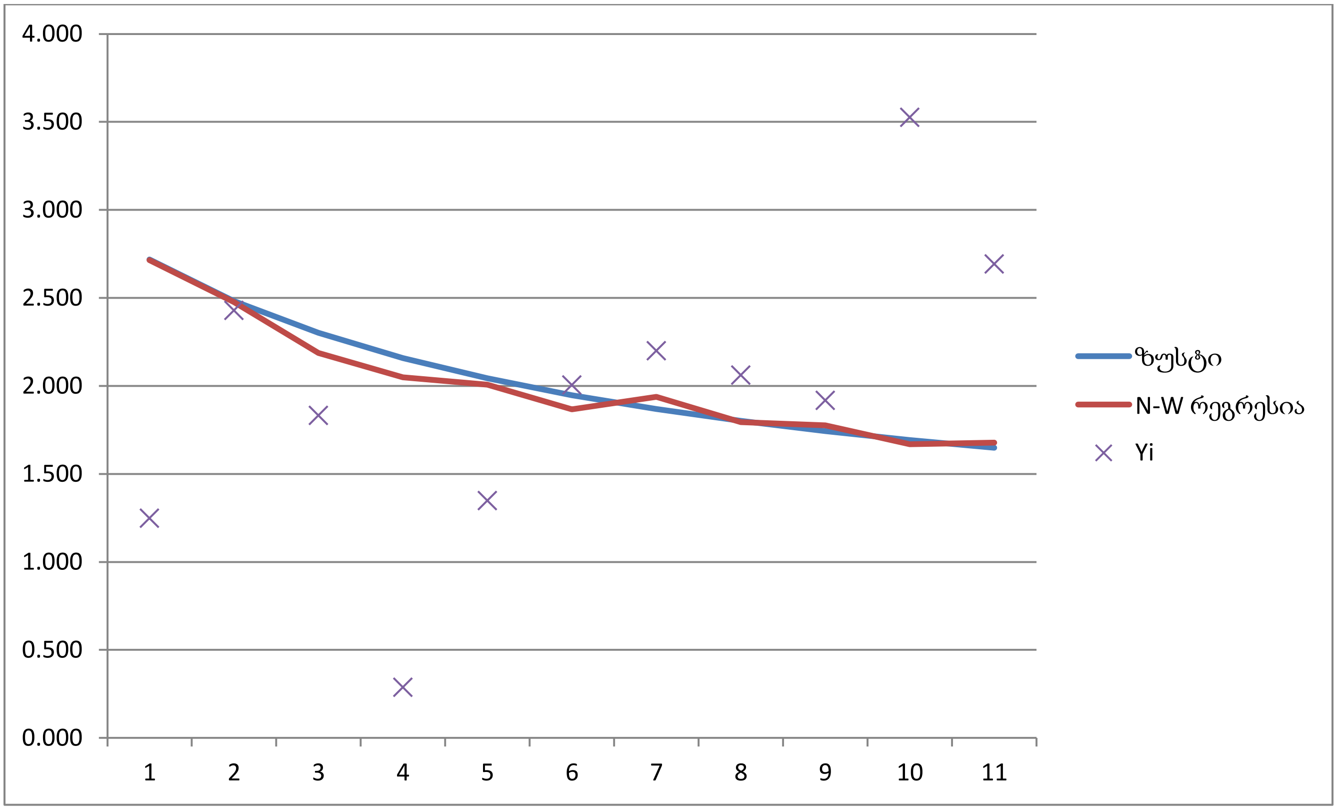
$$a(t) = e^{t^2}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	1.000	1.010	1.041	1.094	1.174	1.284	1.433	1.632	1.896	2.248	2.718
N-W რეგრესია	0.906	1.015	1.139	1.138	1.089	1.233	1.437	1.653	1.894	2.357	2.570
Error N-W	0.094	0.004	0.099	0.044	0.085	0.051	0.004	0.021	0.002	0.109	0.149
Yi	1.5987	0.8072	-0.27	0.539	1.3379	1.0771	1.3451	0.6849	0.2615	2.3447	2.8123

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.553	1.469	1.473

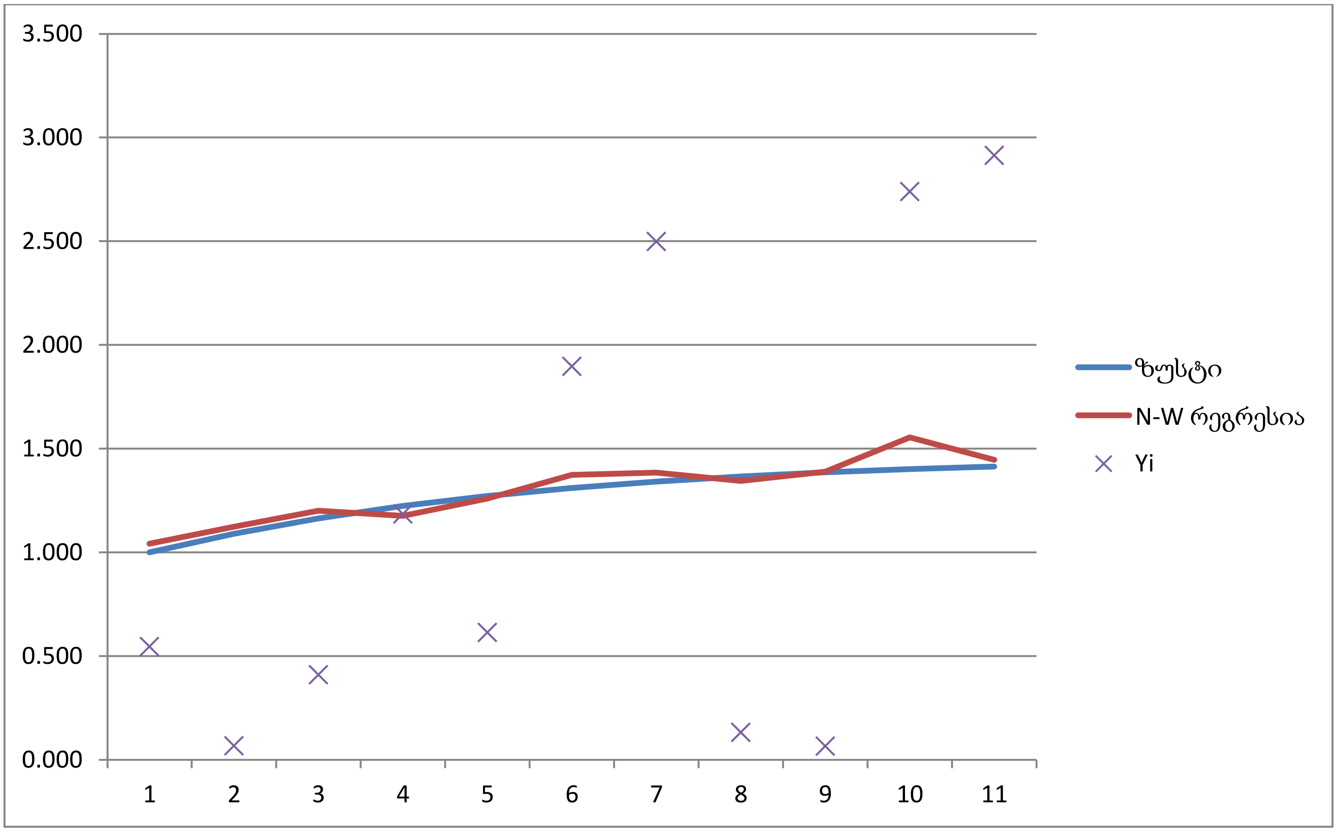
$$a(t) = e^{\frac{1}{t+1}}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	2.718	2.482	2.301	2.158	2.043	1.948	1.868	1.801	1.743	1.693	1.649
N-W რეგრესია	2.801	2.414	2.277	2.147	2.150	1.896	1.937	1.686	1.758	1.689	1.542
Error N-W	0.083	0.069	0.024	0.011	0.108	0.052	0.069	0.115	0.015	0.004	0.107
Yi	1.6644	2.6423	1.945	2.7868	1.6106	0.634	1.2314	2.9151	1.4594	1.4861	0.1838

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.950	2.013	1.997

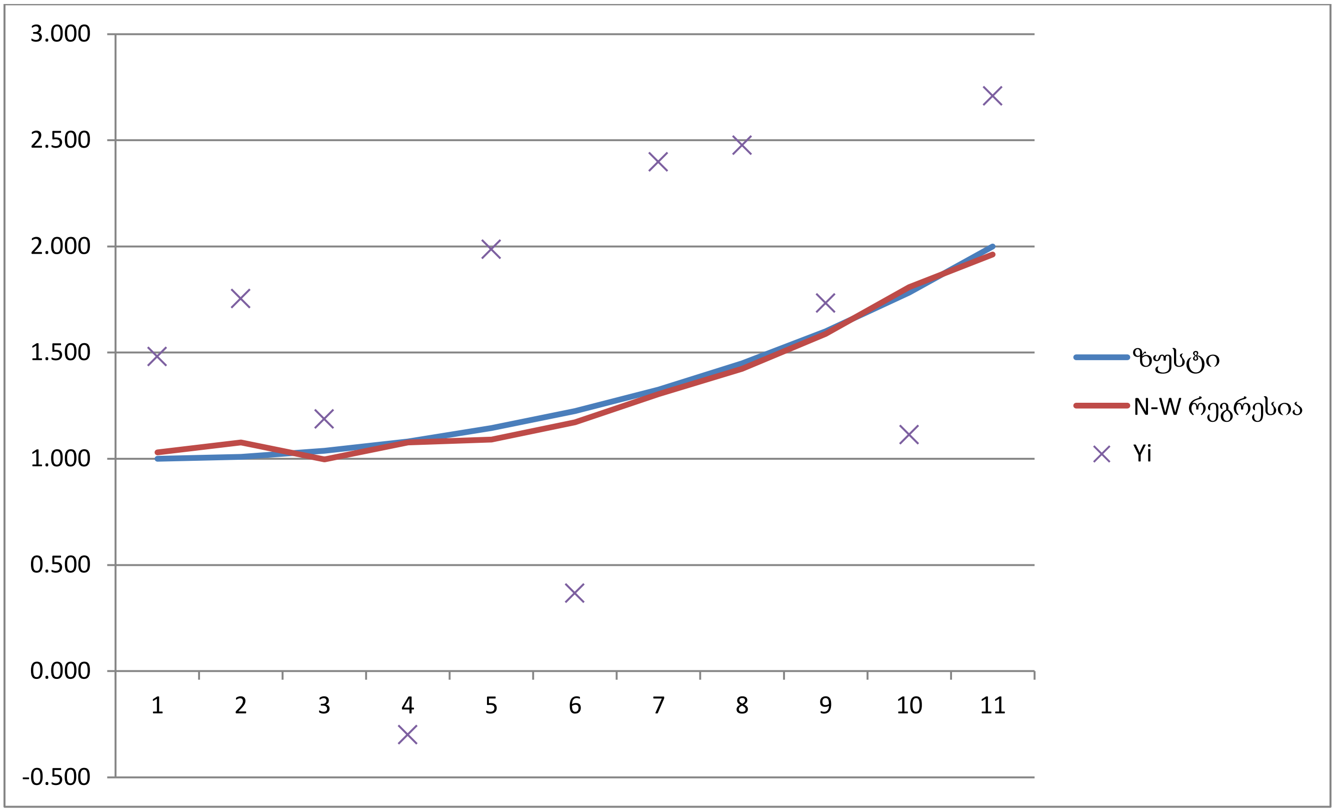
$$a(t) = (t+1)^{\frac{1}{t+1}}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	1.000	1.091	1.164	1.224	1.272	1.310	1.341	1.366	1.386	1.402	1.414
N-W რეგრესია	0.962	1.080	1.195	1.134	1.314	1.331	1.357	1.409	1.472	1.347	1.482
Error N-W	0.038	0.011	0.031	0.090	0.043	0.021	0.015	0.043	0.085	0.055	0.068
Yi	-0.01	1.0649	1.4317	1.042	0.2106	-0.16	1.8008	1.2367	2.3937	1.6349	2.7734

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.291	1.272	1.271

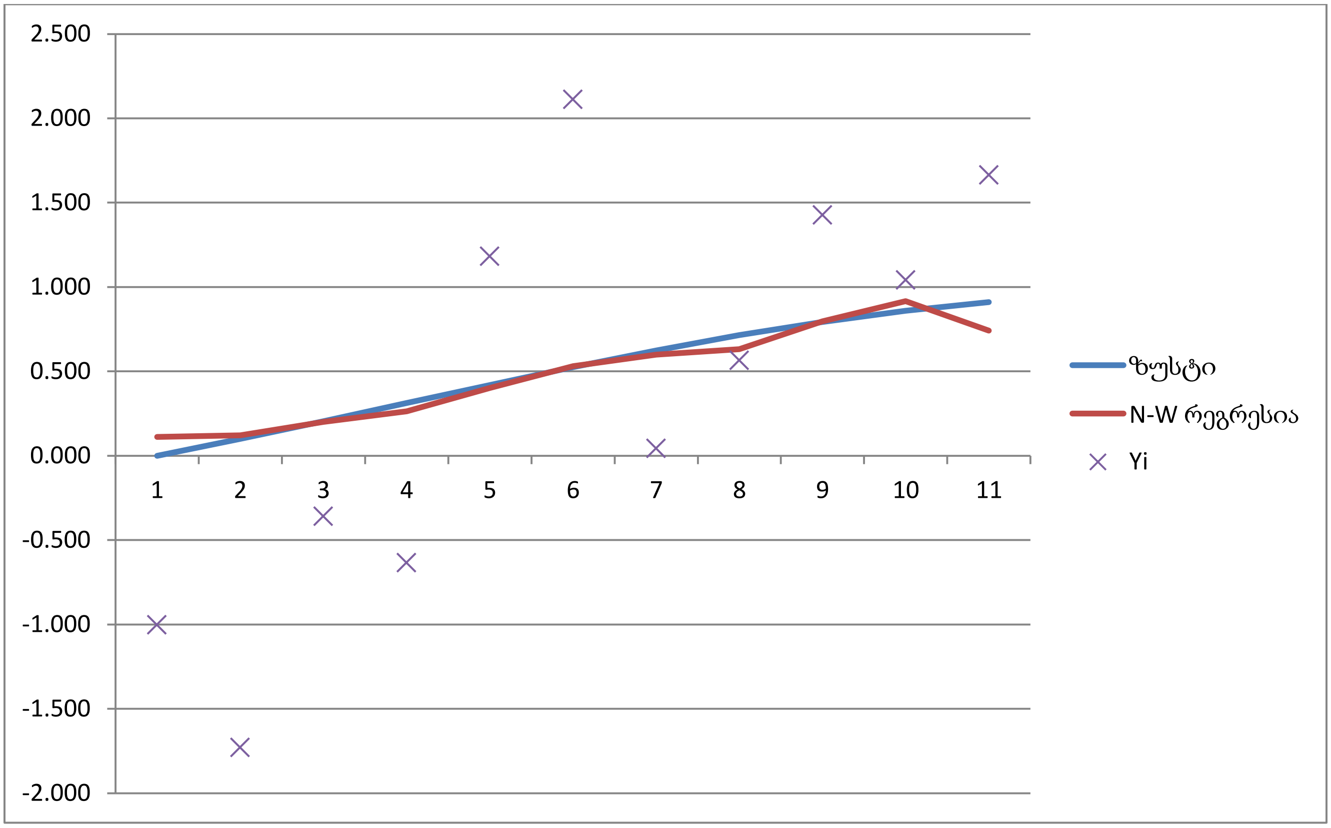
$$a(t) = (t+1)^t$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	1.000	1.010	1.037	1.082	1.144	1.225	1.326	1.450	1.600	1.782	2.000
N-W რეგრესია	0.978	1.060	1.032	1.089	1.128	1.252	1.360	1.450	1.622	1.862	2.004
Error N-W	0.022	0.050	0.005	0.007	0.016	0.028	0.035	0.000	0.022	0.080	0.004
Yi	2.0773	1.3621	1.2373	0.3312	1.2836	-0.14	3.4734	0.7422	-0.56	0.8036	2.4705

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	1.386	1.335	1.337

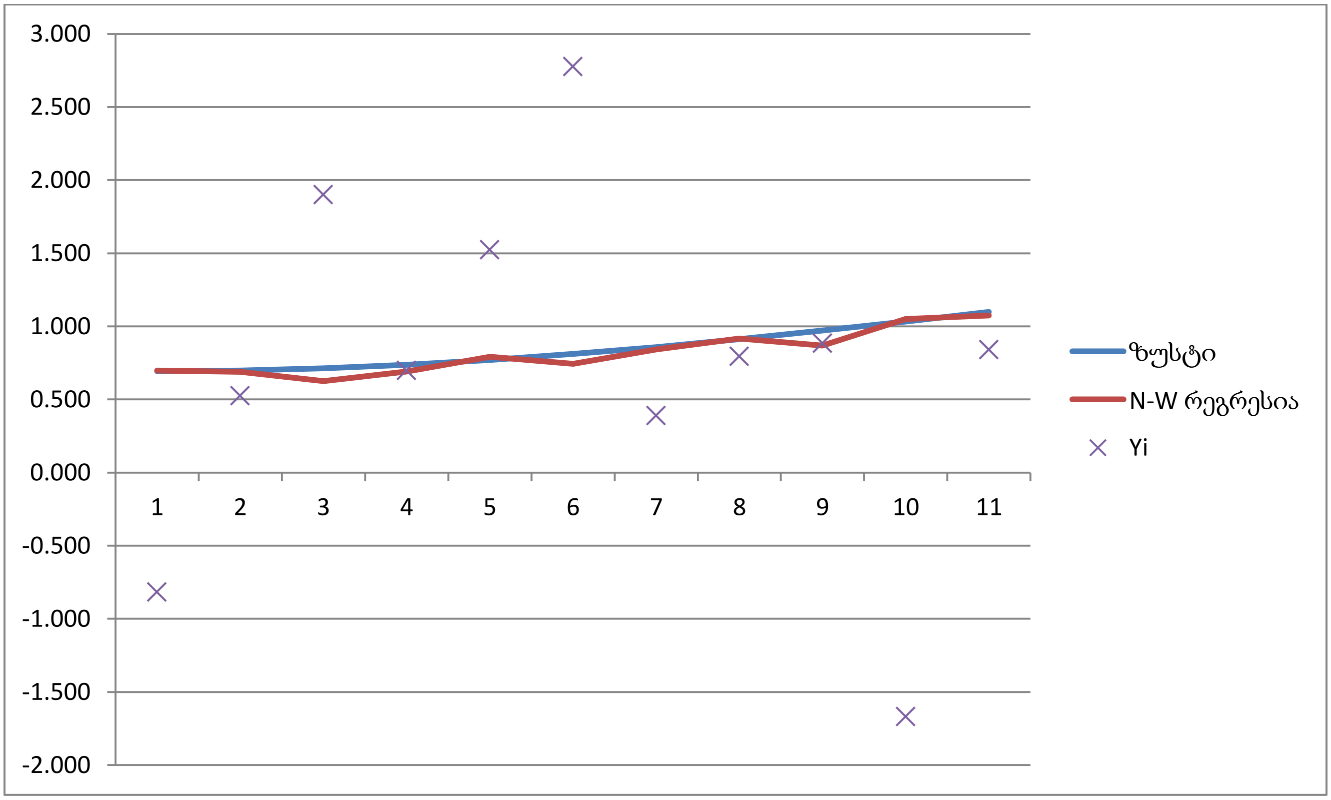
$$a(t) = (sint)^{cost}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.000	0.101	0.205	0.312	0.420	0.525	0.624	0.714	0.793	0.859	0.911
N-W რეგრესია	0.069	0.150	0.207	0.302	0.427	0.514	0.651	0.735	0.787	0.857	0.901
Error N-W	0.069	0.049	0.002	0.010	0.008	0.011	0.028	0.021	0.006	0.002	0.009
Yi	-0.48	-0.35	1.652	-0.01	-0.72	-0.72	1.337	1.8265	2.4849	0.6231	3.6094

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	0.553	0.512	0.512

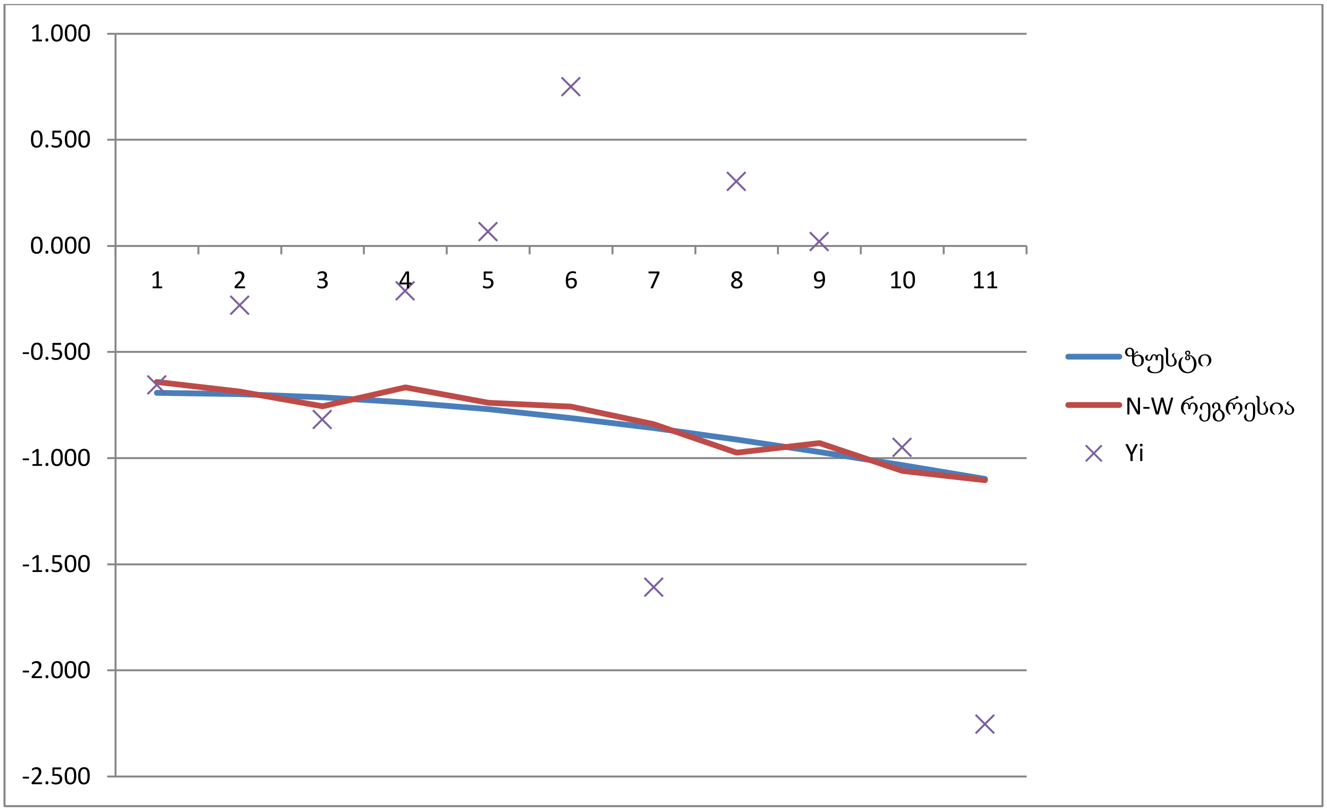
$$a(t) = ln(t^2 + 2)$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	0.693	0.698	0.713	0.737	0.770	0.811	0.859	0.912	0.971	1.033	1.099
N-W რეგრესია	0.683	0.682	0.778	0.868	0.757	0.835	0.849	0.931	0.879	1.052	1.137
Error N-W	0.010	0.016	0.065	0.131	0.013	0.024	0.010	0.019	0.092	0.019	0.038
Yi	-1.79	0.628	0.5732	-0.82	2.6965	2.4319	1.1004	3.2283	0.8776	1.6891	0.8981

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	0.877	0.854	0.861

$$a(t) = ln\frac{1}{t^2+2}$$



t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ზუსტი	-0.69	-0.698	-0.713	-0.73	-0.77	-0.81	-0.85	-0.91	-0.97	-1.03	-1.09
N-W რეგრესია	-0.63	-0.751	-0.691	-0.81	-0.87	-0.81	-0.81	-0.86	-1.00	-1.03	-1.05
Error N-W	0.062	0.053	0.022	0.074	0.102	0.003	0.041	0.050	0.031	0.002	0.041
Yi	-1.06	0.0008	0.8782	-0.97	-0.34	-0.97	-0.25	-2.79	-1.23	-0.56	0.971

	მართკუთხედის ფორმულა	ტრაპეციის ფორმულა	სიმპსონის ფორმულა
ინტეგრალის მიახლო-ებითი მნიშვნელობა	-0.872	-0.850	-0.852

§6. დასკვნა

1. ნაშრომში განხილულია ყველაზე გავრცელებული და აქტუალური რეგრესიის ფუნქციების შეფასების პრისტლი-ჩაოს, გასერ-მიულერისა და ნადარაია-ვატსონის მიდგომები;

2. ყველა განხილულ შემთხვევაში მიღებულია ე.წ. წარმოდგენის თორემები, რომელთა არსი იმაში მდგომარეობს, რომ რეგრესიის ფუნქციისა და მისი წარმოებულების ინტეგრალური ფუნქციონალი წარმოდგება ორი შესაკრების სახით. მათგან პირველი არის დამოუკიდებელ შემთხვევით სიდიდეთა ჯამი, ხოლო მეორე ნაშთითი წევრი;

3. ნაშთითი წევრისათვის, სამივე წევრისათვის, მიღებულია ნულისკენ კრებადობის რიგი;

4. საკმაოდ ზოგად პირობებში ინტეგრალური ფუნქციონალისათვის დამტკიცებულია ძალდებულობის თეორემები;

5. ყველა გახილულ შემთხვევაში მიღებულია პირობები, რომელთა შესრულებისას ადგილი აქვს ასიმპტოტურად ნორმალურობის თვისებას, დადგენილია მანორმალებელი წევრები;

6. სამივე შემთხვევაში დადგენილია ე.წ. განმეორებითი ლოგარითმის თვისება;

7. მოყვანილია რამდენიმე კონკრეტული გამოყენებითი მაგალითი ყველა განხილული ინტეგრალური ფუნქციონალისათვის, მათ შორის ფიშერის ინფორმაციული ინტეგრალისათვის, შენონის ინტეგრალური ენტროპიისა და რეგრესიის ფუნქციის სხვა ფუნციონალებისათვის;

8. შესრულებულია მიღებული შედეგების სიმულაციური კვლევის შედეგები. მოყვანილი ცხრილები და გრაფიკები გვიჩვენებენ თეორიული შედეგების პრაქტიკულ მნიშვნელობას.

გამოყენებული ლიტერატურა

1. Efromovich, S. (1999), Nonparametric Curve Estimation: Methods, Theory and Applications. Springer-Verlag. New York, NY.
2. Hardle, W. (1990), Applied Nonparametric Regression. Cambridge University Press. Cambridge.
3. Loader, C. (1999), Local Regression and Likelihood. Springer-Verlag. New York, NY.
4. Nadaraya, E. (1988), Nonparametric Curve Estimation: Methods, Theory and Applications. Kluwer, the Lenguag of Science.
5. Simonoff, J. S. (1996), Smoothing Methods in Statistics. Springer-Verlag. New York, NY.
6. Venables, W. N. and Ripley, B. D. (2002), Modern Applied Statistics with S. Springer-Verlag. New York, NY.
7. Wasserman, L. (2005), All of Nonparametric Statistics, Springer-Verlag. New York, NY.
8. Nadaraya, E.A. (1964), On estimating regression. Theory of Probability and its Applications, 9 pp. 141-142.
9. Watson, G.S. (1964), Smooth regresseion analysis. Sankhya, Ser. A, 26, pp. 359-372.
10. Priestley, M. B. and Chao, M. T. (1971), Non-parametric function fitting. Internal Bell Telephone Laboratories memo. 15 p.
11. Priestley, M. B. and Chao, M. T. (1972), Non-Parametric Function Fitting. Journal of the Royal Statistical Society. Series B (Methodological), 34, 3. pp. 385-392.
12. Gasser, T. and Muller, H.G. (1979), Kernel Estimation of Regression Functions, In: Th. Muller and M. Rosenblatt, eds. Lecture Notes in Mathematics 757, Springer-Verlag, Berlin. pp. 23-68.
13. Gasser, T. and Muller, H.G. (1984), Estimating Regression Functions and Their Derivatives by the Kernel Method, Scandinavian Journal of Statistics Vol. 11, No. 3. pp. 171-185.
14. Gine, E. and Mason, D. M. (2008), Uniform in bandwidth estimation of integral functionals of the density function. Scand. J. Statist. 35. pp. 739-761.
15. Mason, D. M. and Nadaraya, E. and Sokhadze, G. (2010), Integral functionals of the density. Nonparametrics and robustness in modern statistical inference and time series analysis: a Festschrift in honor of Professor Jana Jureckova, Inst. Math. Stat. Collect., 7, Inst. Math. Statist., Beachwood, OH, pp. 153-168.
16. Babilua, P. K. and Nadaraya E. A. and Patsatsia, M. B. and Sokhadze G. A. (2008), On the integral funciotnals of a kernel estimation of a distribution density. Proc. I. Vekua Institute of Applied Mathematics. 58, 110. pp. 6-14.
17. Birge, L. and Massart, P. (1995), Estimation of integral functionals of a density. Ann. Statist. 23. pp. 11-29.
18. Laurent, B. (1996), Efficient estimation of integral functionals of a density. Annals of Statistics.24, No. 2. pp. 659-681.

19. Mason, D. M. (2003), Representations for integral functionals of kernel density estimators. Austrian J. Statistics. 32. pp. 131-142.
20. Dmitriev, Yu. G. and Tarasenko, F. P. (1974), On estimation of functionals of the probability density function and its derivatives. Theory of Probability and its Applications. 18, Issue 3. pp. 628–633.
21. Nadaraya, E. and Sokhadze, G. (2015), Integral Functionals of a Density. Communications in Statistics - Theory and Methods. (in press).
22. Nadaraya, E. and Sokhadze, G. (2015), On Functionals of a Density. Proceedings of Razmadze Mathematical Institute. (in press).
23. Goldstein, L. and Messer, K. (1992), Optimal Plug-in estimator for nonparametric functional estimation. Ann. Statist, **20**, 3. pp. 1306-1328.
24. Bissantz, N. and Holzmann, H. (2007), Estimation of a quadratic regression functional using the sinc kernel. J. Statist. Plann. Inference, **137**, 3, pp. 712-719.
25. Levit, B. Ya. (1978), Asymptotically efficient estimation of nonlinear functionals. (Russia) Problems Inform. Transmission **14**, 3, pp. 65-72.
26. Hlavka, Z. (2011), On Nonparametric estimators of location of maximum. Acta Univ. Carolina – Math. Phys. **52**, 1, pp. 5-13.
27. Arabidze, D. and Babilua, P. and Nadaraya, E. and Sokhadze, G. and Tkeshelashvili, A. (2015), Integral Functionals of the Gasser-Muller Regression Function. Ukrainian Mathematical Journal. Vol. 67, Issue 4, pp. 435-446.
28. Arabidze, D. (2014) Integral Functionals of the Nadaraya-Watson Regression Function. International Journal of Engineering and Innovative Technology (IJEIT), Vol 4. Issuue 1, pp 26-32
29. Arabidze, D. (2014) Integral Functionals of the Priestley-Chao Regression Function. Bulletin of the Georgian National Academy of Scienses, Vol 8, no 2, pp 29-40.
30. Devroye, L. (1991), Exponential Inequalities in Nonparametric Estimation. Nonparametric Functional Estimation and Related Topics. (Spetses, 1990), NATO ASI Series 335, 31-44. Kluwer Acad. Publ., Dortrecht.
31. Kuelbs, J. (1976), The law of the iterated logarithm and related strong convergence theorems for Banach space valued random variables. Ecoled'Eté de Probabilités de Saint-Flour V-1975. Lecture Notes in Mathematics, 539, 224-314. Springer, Berlin.