РОЗДІЛ 3
АНАЛІЗ ТЕОРЕТИЧНОЇ ЕФЕКТИВНОСТІ ПОЛІНОМІАЛЬНИХ ОЦІНОК ПАРАМЕТРІВ РЕГРЕСІЇ

3.1 Асимптотичні дисперсії ММПл-оцінок векторного параметра при неоднаково розподілених даних

Відомо, що основним критерієм ефективності оцінок інформативних параметрів регресії виступають їх дисперсії. Як вже зазначалося раніше, оцінки векторних параметрів, які знаходяться із застосуванням методу максимізації поліномів, в загальному випадку володіють властивістю слушності і є асимптотично незміщеними. Для отримання аналітичних виразів, що описують дисперсії ММПл-оцінок векторного параметру $\boldsymbol{\theta}$, використовується матриця кількості добутої інформації про компоненти параметру при застосуванні стохастичних поліномів загального виду (2.6) порядку S [58]. При моментному описі така матриця $\mathbf{J}_S(\boldsymbol{\theta})$ складається із елементів

$$J_{SN}^{(p,q)} = \sum_{v=1}^{N}\sum_{i=1}^{S}\sum_{j=1}^{S} k_{iv}^{(p)} k_{jv}^{(q)} F_{(i,j)v} = \sum_{v=1}^{N}\sum_{i=1}^{S} k_{i,v}^{(p)} \frac{\partial}{\partial a_q}\alpha_{iv}, \;\; p,q = \overline{0,Q-1}. \quad (3.1)$$

У статистичному сенсі кількість добутої інформації концептуально є поняттям близьким до кількості інформації по Фішеру. Показано, що дисперсії $\sigma^2_{(a_p)S}$ ММПл-оцінок складових векторного параметру $\boldsymbol{\theta}$ в асимптотичному випадку (при $N \to \infty$) можуть бути отримані як елементи головної діагоналі варіаційної матриці $\mathbf{V}_S(\boldsymbol{\theta})$, яка є оберненою до матриці, складеної із елементів матриці (3.1)

$$\mathbf{V}_{PMMS}(\boldsymbol{\theta}) = \begin{bmatrix} J_{SN}^{(1,q)} & J_{SN}^{(2,q)} & \ldots & J_{SN}^{(1,q)} \\ J_{SN}^{(2,1)} & J_{SN}^{(2,2)} & \ldots & J_{SN}^{(2,q)} \\ \ldots & \ldots & \ldots & \ldots \\ J_{SN}^{(p,1)} & J_{SN}^{(p,2)} & \ldots & J_{SN}^{(p,q)} \end{bmatrix}^{-1}. \qquad (3.2)$$

Ще одна важлива властивість ММПл полягає у тому, що при збільшенні числа членів S стохастичного поліному (1) дисперсія оцінок зменшується, оскільки кількість добутої інформації асимптотично (при $S \to \infty$) прямує до інформації по Фішеру [58].

3.2 Порівняльний аналіз теоретичної ефективності ММПл-оцінок із МНК-оцінками

3.2.1. Точність лінійних ММПл-оцінок параметрів регресії при моментному описі

Синтезовані у розділі 2 обчислювальні алгоритми оцінювання параметрів регресії базуються на застосуванні як базисних функції МММл степеневих перетворень (2.11). Це визначає необхідність використання для ймовірнісного опису випадкових складових регресійних моделей сукупність відповідних початкових моментів випадкових послідовностей (2.12).

Використовуючи результати підрозділів 2.3.2, 2.4.2, 2.5.1 щодо застосування лінійного варіанта метода максимізації поліномів можемо записати вираз для кількості добутої інформації (при степені S=1) про значення оцінок параметрів для лінійної багатофакторної регресії

$$J_{1N}^{(p,q)} = \frac{1}{\mu_2} \sum_{v=1}^{N} \left[x_{v,p} x_{v,q} \right], \; p,q = \overline{0,Q-1}, \qquad (3.3, a)$$

для поліноміальної регресії

$$J_{1N}^{(p,q)} = \frac{1}{\mu_2}\sum_{v=1}^{N}\left[x_v^{p+q}\right],\ p,q=\overrightarrow{0,Q-1} \qquad (3.3,\ б)$$

та нелінійної регресії

$$J_{1N}^{(p,q)} = \frac{1}{\mu_2}\sum_{v=1}^{N}\left[\frac{\partial}{\partial a_p}R_v\left(\boldsymbol{\theta},\mathbf{X}\right)\frac{\partial}{\partial a_q}R_v\left(\boldsymbol{\theta},\mathbf{X}\right)\right],\ p,q=\overrightarrow{0,Q-1}, \qquad (3.3,\ в)$$

де μ_2 - центральний момент 2-го порядку (дисперсія) випадкових помилок ξ_v відповідної регресійної моделі.

Зазначимо, отриманий у розділі 2 факт, що результуючі вирази для знаходженні ММПл-оцінки параметрів регресійних моделей при степені S=1 є еквівалентними МНК-оцінкам. А отже у цьому випадку співпадає і точність оцінок обох методів. Тому отримані вирази (3.3) можуть бути використані у подальшому для порівняльного аналізу ефективності ММПл-оцінок, які отримуються із застосуванням поліномів біль високого порядку.

3.2.2. Відносна ефективність квадратичних ММПл-оцінок при асиметрично-розподілених моделях регресійних помилок

Наведені результати підрозділів 2.3.3, 2.4.3, 2.5.2 свідчать, що застосування квадратичного варіанта метода максимізації поліномів для знаходження оцінок параметрів регресії є доцільним лише у випадку асиметрії розподілу випадкової складової регресійних помилок. Одним із математичнх критеріїв цієї властивості може виступати відмінність величини центрального момента 3-го порядку від нуля $\mu_3 = 0$.

Використовуючи отримані вирази для оптимальних коефіцієнтів $k_{i,v}^{(p)}$, $i=\overrightarrow{1,2}$, що забезпечують мінімізацію дисперсії оцінок компонентів шуканого параметра при степені S=2, можемо записати відповідні вирази для кількості добутої інформації про значення оцінок параметрів для лінійної багатофакторної регресії

$$J_{2N}^{(p,q)} = \frac{\mu_2^3 - \mu_2\mu_4}{\mu_2\left(\mu_2^3 - \mu_2\mu_4 + \mu_3^2\right)} \sum_{v=1}^{N} \left[x_{v,p} x_{v,q} \right], \; p,q = \overline{0,Q-1}, \qquad (3.4, a)$$

для поліноміальної регресії

$$J_{2N}^{(p,q)} = \frac{\mu_2^3 - \mu_2\mu_4}{\mu_2\left(\mu_2^3 - \mu_2\mu_4 + \mu_3^2\right)} \sum_{v=1}^{N} \left[x_v^{p+q} \right], \; p,q = \overline{0,Q-1} \qquad (3.4, б)$$

та нелінійної регресії

$$J_{2N}^{(p,q)} = \frac{\mu_2^3 - \mu_2\mu_4}{\mu_2\left(\mu_2^3 - \mu_2\mu_4 + \mu_3^2\right)} \sum_{v=1}^{N} \left[\frac{\partial}{\partial a_p} R_v\left(\boldsymbol{\theta}, \mathbf{X}\right) \frac{\partial}{\partial a_q} R_v\left(\boldsymbol{\theta}, \mathbf{X}\right) \right], \; p,q = \overline{0,Q-1}, \qquad (3.4, в)$$

де μ_2, μ_3, μ_4 - центральні моменти випадкових помилок ξ_v відповідної регресійної моделі.

Зіставлення відповідних виразів (3.3) і (3.4) дозволяє зробити два основні висновки:

1) вирази, які описують кількість добутої інформації при степені S=2 для різних типів регресійних моделей відрізняться між собою однаковим ваговим коефіцієнтом, який залежить від центральних моментів випадкової складової регресійної моделі;

2) ефективність квадратичних ММПл-оцінок відносно лінійних ММПл-оцінок (а відповідно і МНК-оцінок) параметрів регресії залежить від значень статистик вищих порядків (до 4-го порядку).

Таким чином, можна записати співвідношення для відповідних варіаційних матриць загального виду (3.2), що містять асимптотичні дисперсії оцінок параметрів регресії, у вигляді

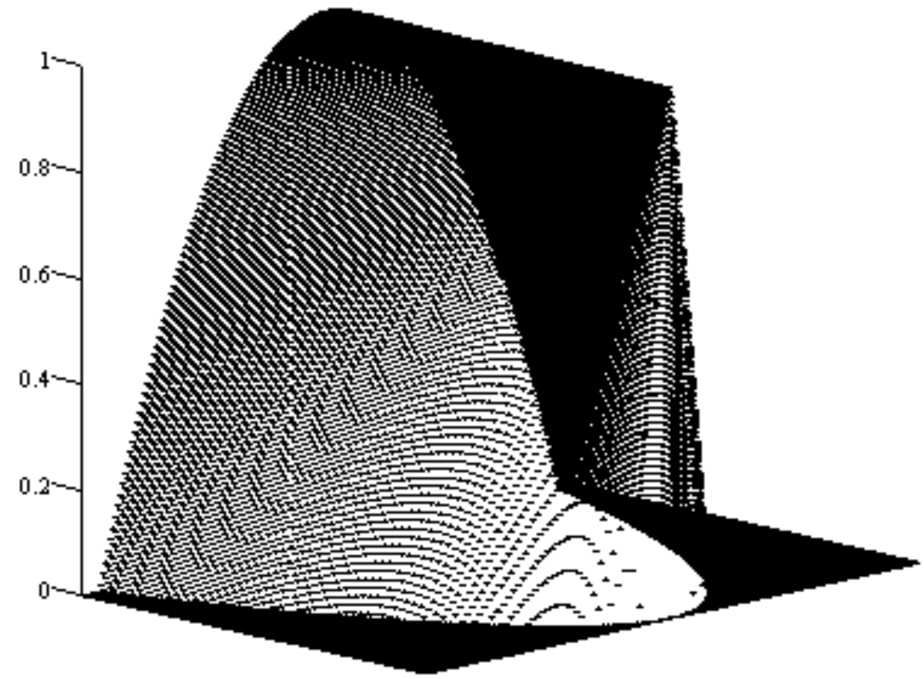
$$\mathbf{V}_{(PMM2)} = g_{PMM2/PMM1}\mathbf{V}_{(PMM1)} = g_{PMM2/OLS}\mathbf{V}_{(PMM1)}. \tag{3.5}$$

де коефіцієнт зменшення дисперсій оцінок, який описує залежність відносної точності, може бути представлений у вигляді

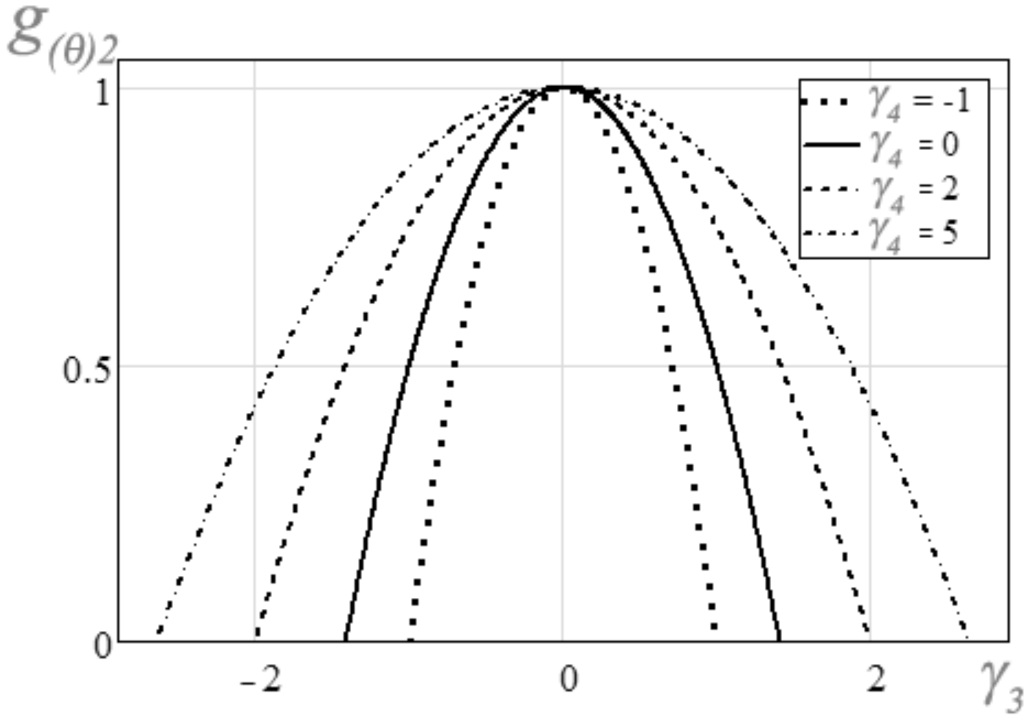
$$g_{PMM2/OLS} = 1 - \frac{\gamma_3^2}{2+\gamma_4}. \tag{3.6}$$

Перехід у виразі (3.6) від моментного опису до кумулянтного обумовлений тим фактом, що відхилення значень кумулянтних коефіцієнтів вищих порядків γ_r від нуля показує міру відмінності від гауссового розподілу. Крім того, кумулянтні коефіцієнти вищих порядків не можуть приймати довільних значень. Зокрема, для кумулянтних коефіцієнтів асиметрії та ексцесу відомою є нерівністю $\gamma_4 + 2 \geq \gamma_3^2$ [46]. З урахуванням цієї нерівності, на основі аналізу (3.6) можна зробити висновок, що коефіцієнт зменшення дисперсії $g_{PMM2/OLS}$ є безрозмірною величиною, що належить діапазону $(0;1]$.

На рис.3.1 зображена залежність коефіцієнту зменшення дисперсії $g_{PMM2/OLS}$ від кумулянтних коефіцієнтів асиметрії γ_3 та ексцесу γ_4. Оскільки коефіцієнт зменшення дисперсії є функціями двох змінних, то множина їх значень представляє собою поверхню (рис.3.1, *а*). Для більшої наочності крім 3D-графіка на рис.3.1,*б* також представлені розрізи, побудовані як залежності від коефіцієнта асиметрії γ_3 при фіксованих значеннях γ_4.



а)



б)

Рисунок 3.1. Залежність коефіцієнту зменшення дисперсії $g_{PMM2/OLS}$ від кумулянтних коефіцієнтів асиметрії γ_3 та ексцесу γ_4

Таким чином, відносне зменшення асимптотичної дисперсії PMM-оцінок є однаковим для всіх компонентів векторного параметра. Загалом відносна ефективність залежить від ступені відмінності розподілу регресійних помилок від гаусового закону, що чисельно виражається перш за все відмінністю величини абсолютного значення (модуля) коефіцієнта асиметрії γ_3. При її збільшенні величина відносного зменшення дисперсії може бути достатньо суттєвою і навіть асимптотично прагнути до нуля при наближенні абсолютного значення коефіцієнту асиметрії до межі області допустимих значень, тобто $|\gamma_3| \to \sqrt{2+\gamma_4}$.

3.2.3. Відносна ефективність кубічних ММПл-оцінок при симетрично-розподілених моделях регресійних помилок

Використовуючи результати підрозділів 2.3.4, 2.4.4, 2.5.3 щодо застосування кубічного варіанта метода максимізації поліномів для випадку симетрично розподіленої випадкових регресійних помилок, можемо записати відповідні вирази для кількості добутої інформації при S=3 про значення оцінок параметрів для лінійної багатофакторної регресії у вигляді

$$J_{2N}^{(p,q)} = \frac{9\mu_2^3 - 6\mu_2\mu_4 + \mu_6}{\mu_2\mu_6 - \mu_4^2} \sum_{v=1}^{N} \left[x_{v,p} x_{v,q} \right], \; p,q = \overline{0, Q-1}, \tag{3.7, a}$$

для поліноміальної регресії

$$J_{2N}^{(p,q)} = \frac{9\mu_2^3 - 6\mu_2\mu_4 + \mu_6}{\mu_2\mu_6 - \mu_4^2} \sum_{v=1}^{N} \left[x_v^{p+q} \right], \; p,q = \overline{0, Q-1} \tag{3.7, $б$}$$

та нелінійної регресії

$$J_{2N}^{(p,q)} = \frac{9\mu_2^3 - 6\mu_2\mu_4 + \mu_6}{\mu_2\mu_6 - \mu_4^2} \sum_{v=1}^{N} \left[\frac{\partial}{\partial a_p} R_v(\boldsymbol{\theta}, \mathbf{X}) \frac{\partial}{\partial a_q} R_v(\boldsymbol{\theta}, \mathbf{X}) \right], \; p,q = \overline{0, Q-1},$$

(3.7, *в*)

де μ_2, μ_4, μ_6 - парні центральні моменти випадкових помилок ξ_v відповідної регресійної моделі.

Таким чином, як і в попередньому випадку ефективність кубічних ММПл-оцінок відносно лінійних ММПл-оцінок (а відповідно і МНК-оцінок) не відрізняться для різних типів регресійних моделей і визначається однаковим ваговим коефіцієнтом

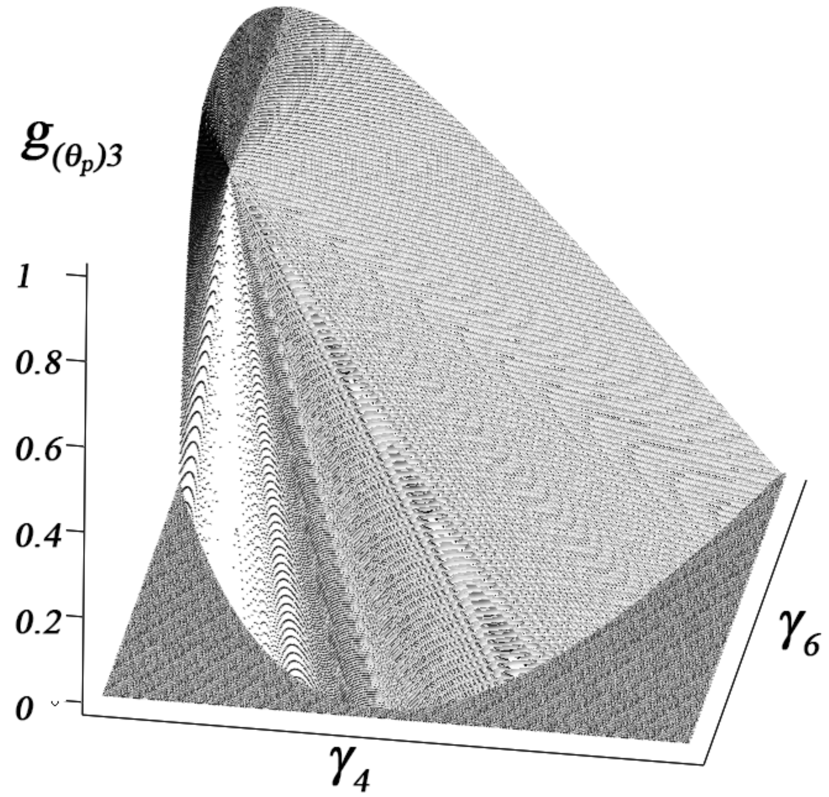
$$\mathbf{V}_{(PMM3)} = g_{PMM3/PMM1} \mathbf{V}_{(PMM1)} = g_{PMM3/OLS} \mathbf{V}_{(PMM1)}. \tag{3.8}$$

який залежить від парних центральних моментів випадкової складової регресійної моделі (до 6-го порядку).

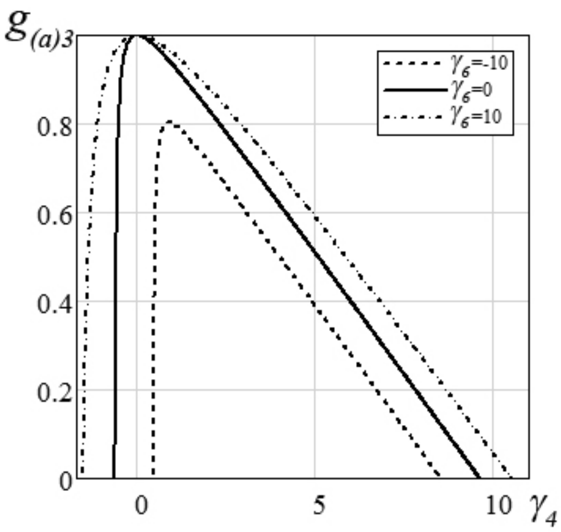
Для компактності запису та більш зручного трактування результатів аналогічно до (3.6) здійснимо перехід від моментного до кумулянтного опису і представимо коефіцієнт зменшення дисперсії МММл оцінок при S=3 у вигляді

$$g_{PMM3/OLS} = 1 - \frac{\gamma_4^2}{6 + 9\gamma_4 + \gamma_6}. \tag{3.9}$$

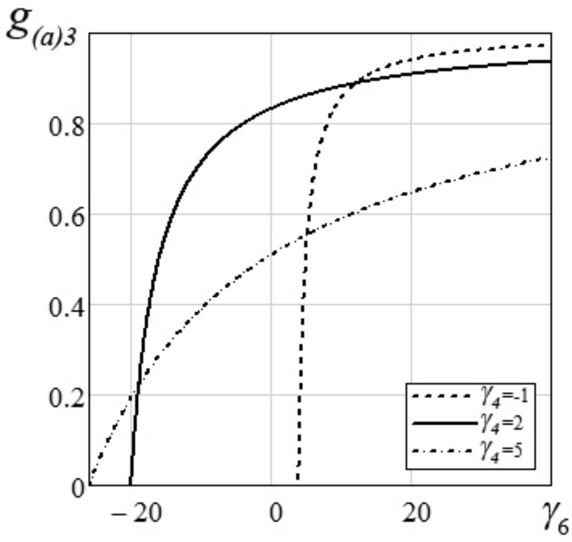
Як вже було зазначено вище, для значеннь кумулянтних коефіцієнтів вищого порядку існують певні області допустимих значень. Зокрема, для симетрично розподілених випадкових величин обмеження для кумулянтних коефіцієнтів 4-го і 6-го порядку визначається нерівностями: $\gamma_4 > -2$ та $\gamma_6 + 9\gamma_4 + 6 > \gamma_4^2$ [46]. Враховуючи ці нерівності, можна зробити висновок, що коефіцієнт зменшення дисперсії $g_{PMM3/OLS}$ також є безрозмірною величиною, що належить діапазону $(0;1]$.



а)



a)



b)

б) *в)*

Рисунок 3.2. Залежність коефіцієнту зменшення дисперсії $g_{PMM3/OLS}$ від кумулянтних коефіцієнтів γ_4 та γ_6

На рис.3.2,*б* зображена залежність (у вигляді 3D-графіка) коефіцієнту зменшення дисперсії $g_{PMM3/OLS}$ від кумулянтних коефіцієнтів 4-го γ_4 та 6-го γ_6 порядків. Для більшої наочності на рис.3.2,*б* представлені розрізи, побудовані як залежності від коефіцієнта γ_4 при фіксованих значеннях γ_6, а на рис.3.2,*в* залежності від коефіцієнта γ_6 при фіксованих значеннях γ_4.

Аналіз представлених залежностей показує, що точність PMM-оцінок при $S = 3$ може може бути суттєво більшою за точність МНК-оцінок. Зокрема, при наближенні величини кумулянтних коефіцієнтів до кордону області допустимих значень $|\gamma_4| \to \sqrt{\gamma_6 + 9\gamma_4 + 6}$ дисперсія PMM-оцінок асимптотично прагне до нуля. Відносне зменшення дисперсій оцінок не спостерігається лише у тому випадку, коли коефіцієнт ексцесу тотожний нулю $\gamma_4 = 0$, що зокрема відповідає гаусовому розподілу регресійних помилок.

3.3 Порівняльний аналіз теоретичної ефективності ММПл-оцінок із ММП-оцінками

Як вже зазначалося раніше, однією із основних альтернатив методу найменших квадратів в регресійному аналізі є підхід, що базується на методі максимальної правдоподібності. Його застосування потенційно є більш точним, оскільки дозволяє враховувати ймовірнісні властивості випадкової складової регресійної моделі. Але платою за це є необхідність вирішення додаткових задач ідентифікації щільності розподілу регресійних помилок та оцінювання її параметрів.

3.3.1. Застосування експоненційного степеневого розподілу для опису регресійних помилок

Здійснимо порівняльний аналіз ефективності цього підходу використавши як ймовірнісну модель регресійних помилок експоненційний степеневий розподіл (ЕСР). Вибір цієї моделі обумовлений тим фактором, що вона адекватно підходить для опису великого числа процесів та явищ, які спостерігаються в реальності. Про її універсальність говорить той факт, що розподіл Лапласа, гаусовий та рівномірний розподіли є окремими випадками ЕСР при певних значеннях його параметру форми. Тому в різних літературних джерелах можна зустріти й альтернативні назви: «узагальнений нормальний розподіл», «узагальнений гаусовий розподіл», «узагальнений розподіл помилок», «розподіл Суботіна» та ін. [95–97].

ЕСР застосовується як ймовірнісна модель при вирішення великої кількості різноманітних практичних задач: для опису похибок вимірювальних приладів і результатів вимірювання [19, 98], випадкових шумів та завад при опрацюванні радіо [99], відео [100] та аудіо [101] сигналів, статистичних даних медичного [102], біологічного [103], економічного [104] походження, в управлінні ризиками [105] тощо. Подібне широке поширення ЕСР спонукало до розробки ряду спеціалізованих програмних продуктів, орієнтованих на комп'ютерне статистичне моделювання. Серед них виділимо бібліотеку «normalp» [31] для мови програмування R, що використовується у даній роботі.

З урахуванням центрованості (нульового математичного сподівання) регресійних помилок модель ЕСР описується функцією виду

$$w(\xi)=\frac{1}{2\sigma_\beta p^{1/\beta}\Gamma(1+1/\beta)}\exp\left(-\frac{|\xi|^\beta}{\beta\sigma_\beta^\beta}\right), \tag{3.10}$$

де значення параметрів масштабу σ_β і форми β.

На рис.3.3 наведено графіки щільності ЕСР, побудовані при різних значеннях параметру форми β. Очевидно, що при $\beta=2$ розподіл (3.10) відповідає гаусовому закону. Для значень $\beta<2$ функція (3.10) має лептокуртичний (гостровершинний) характер з пологими спадами, при $\beta=1$

відповідає розподілу Лапласа. Із зростанням величини $\beta > 2$ розподіл стає платокуртичним, тобто більш плосковершинним, а в граничному випадку (при $\beta \to \infty$) трансформується в рівномірний закон розподілу.

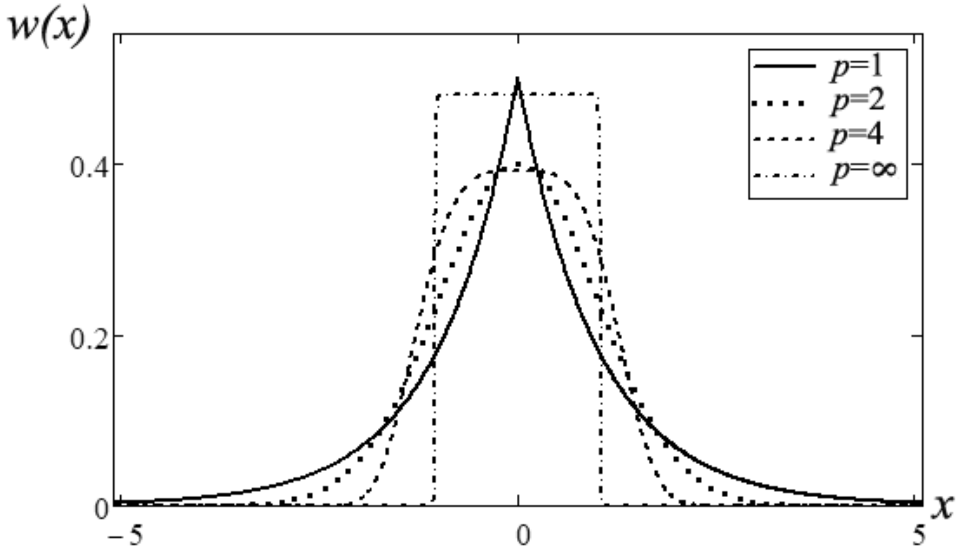


Рисунок 3.3. Експоненційний степеневий розподіл при різних значеннях параметра форми β.

3.3.2. Відносна точність оцінок параметрів регресії при експоненційному степеневому розподілі помилок

Використовуючи (3.10) логарифмічну функцію правдоподібності, яка приймає максимум в околі істинних значень інформативних параметрів $\boldsymbol{\theta}$ детермінованої складової $R_v\left(\boldsymbol{\theta}, \mathbf{X}\right)$ регресійної моделі, можна записати у вигляді

$$L\left(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}\right) = -N\log\left(2\right) + N\log\left(\sigma\right) - \sum_{v=1}^{N}\left[\frac{\Gamma\left(1+1/\beta\right)\left|y_v - R_v\left(\boldsymbol{\theta}, \mathbf{X}\right)\right|}{\sigma}\right]^{\beta}, \quad (3.11)$$

де σ нормалізоване значення середньоквдратичного відхилення

$$\sigma = \sqrt{\mu_2} = \beta^{1/\beta}\sigma_\beta \Gamma\left(1 + 1/\beta\right). \tag{3.12}$$

Для аналітичного розрахунку величини дисперсії ММП-оцінок, використаємо величину кількості інформації по Фішеру, яка розрахована для регресійних моделей з ЕСР в роботі [106]

$$J_N^{(p,q)} = \frac{\Gamma\left(1/\beta\right)\Gamma\left(2-1/\beta\right)}{\sigma^2}\sum_{v=1}^{N}\left[\frac{\partial}{\partial a_p}R_v\left(\boldsymbol{\theta},\mathbf{X}\right)\frac{\partial}{\partial a_q}R_v\left(\boldsymbol{\theta},\mathbf{X}\right)\right],\ p,q=\overline{0,Q-1}. \tag{3.11}$$

Із урахуванням (3.3) відповідна варіаційна матриця може бути представлена у вигляді $\mathbf{V}_{(MLE)} = g_{MLE/OLS}\mathbf{V}_{(OLS)}$, де

$$g_{MLE/OLS} = \frac{\Gamma\left[1+\beta^{-1}\right]^2}{\Gamma\left[2-\beta^{-1}\right]\Gamma\left[3\beta^{-1}\right]}. \tag{3.12}$$

Оскільки ЕСР має симетричний характер, то для отримання оцінок регресійних параметрів необхідно застосовувати кубічну модифікацію методу максимізації поліномів. Для отримання аналітичних виразів, що описують дисперсії ММПл-оцінок, використаємо взаємозвязок між центральними моментами μ_i і параметрами ЕСР [31]:

$$\mu_i = \Gamma\left[\left(i+1\right)\beta^{-1}\right]\Gamma\left[\beta^{-1}\right]^{\frac{i-2}{2}}\Gamma\left[3\beta^{-1}\right]^{-\frac{i}{2}}\sigma_\beta^2. \tag{3.13}$$

Використовуючи вирази (3.13) та (3.7) для кількості добутої інформації можна показати, що елементи варіаційної матриці ММПл-оцінок при $S=3$ відрізняються від елементів матриці лінійних МНК-оцінок (3.3) на коефіцієнт, який може бути представлено як функція від параметру форми ЕСР

$$g_{PMM3/OLS} = \frac{\Gamma\left[p^{-1}\right]^2\left(\Gamma\left[5\beta^{-1}\right]^2 - \Gamma\left[3\beta^{-1}\right]\Gamma\left[7\beta^{-1}\right]\right)}{\Gamma\left[3\beta^{-1}\right]\left(9\Gamma\left[3\beta^{-1}\right]^3 - 6\Gamma\left[\beta^{-1}\right]\Gamma\left[3\beta^{-1}\right]\Gamma\left[5\beta^{-1}\right] + \Gamma\left[\beta^{-1}\right]^2\Gamma\left[7\beta^{-1}\right]\right)} . \quad (3.14)$$

Аналіз залежностей, представлених на рис.3.4, свідчить, що для лептокуртичних (гостровершинних) розподілів при $\beta < 2$ відносна ефективність ММП може бути суттєво вищою за інші методи.

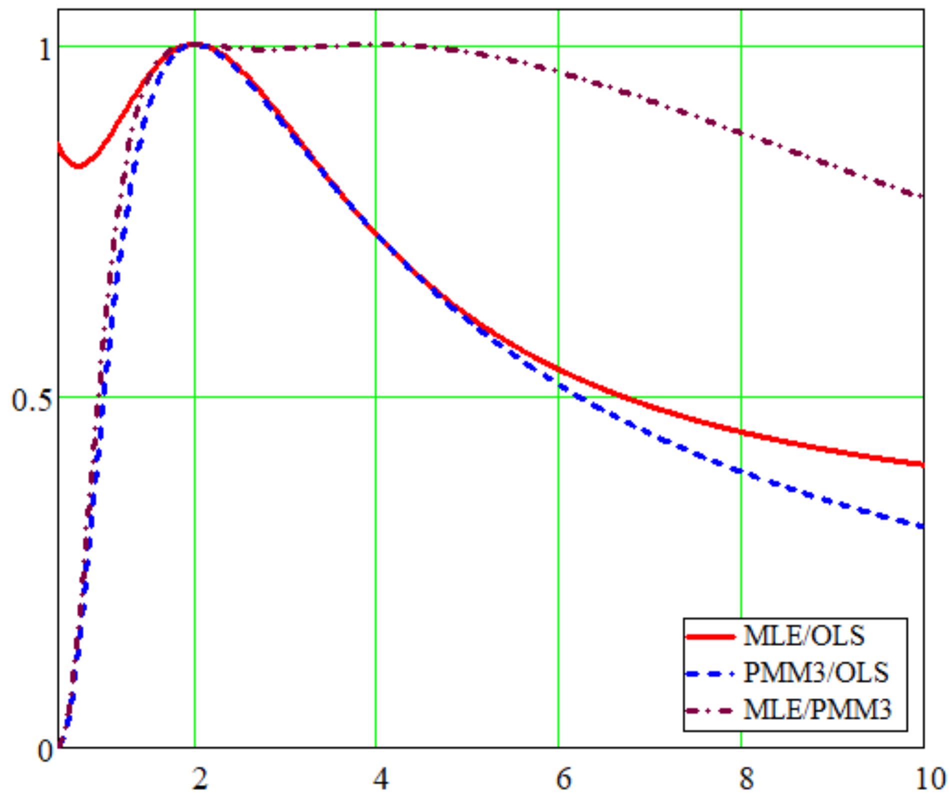


Рисунок 3.4. Залежність коефіцієнту зміни дисперсії $g_{MLE/OLS}$, $g_{PMM3/OLS}$ і $g_{MLE/PMM3}$ від параметра форми β ЕСР.

У випадку гаусового розподілу при $\beta = 2$ точність усіх трьох методів (ММП, ММПл і МНК) є однаковою. При зростанні величини параметру форми $\beta > 2$ для достатньо широкого діапазону значень теоретична ефективність

ММП і ММПл між собою фактично не відрізняється, але може суттєво перевищувати МНК. Для суттєво платокуртичних (плосковершинних) розподілів при $\beta > 5$ відносна ефективність ММП щодо ММПл поступово зростає.

3.4. Висновки

Проведено теоретичний аналіз ефективності поліноміальних оцінок параметрів регресії порівняно із оцінками методу найменших квадратів та максимальної правдоподібності. Отримані наступні основні результати:

1. Дисперсії оцінок параметрів регресійних залежностей, отриманих із застосуванням методу максимізації поліномів, в загальному випадку не перевищують, а в цілому можуть бути значно менші за дисперсії оцінок, знайдених методом найменших квадратів.

2. Ступінь ефективності ММП-оцінок (відносно МНК) не залежить від типу регресійних моделей та є однаковою для всіх складових векторного параметру, що оцінюється.

3. Теоретична величина коефіцієнтів зменшення дисперсії ММПл-оцінок залежить від ступеня негаусовості випадкової складової регресійної моделі.

4. При використанні стохастичних поліномів степені S=2 і асиметричному характері регресійних помилок коефіцієнт зменшення дисперсії (відносно МНК) залежить лише від величини їх коефіцієнтів асиметрії та ексцесу.

5. При симетричному характері регресійних помилок і використанні стохастичних поліномів степені S=3 коефіцієнт зменшення дисперсії (відносно МНК) залежить від величини їх коефіцієнтів 4-го та 6-го порядків.

6. Зростання відносної точності ММПл-оцінок може бути достатньо суттєвим при наближенні величини кумулянтних коефіцієнтів до меж своїх областей допустимих значень.

7. При експоненційному степеневому розподілі помилок теоретична точність ММПл-оцінок параметрів регресії у цілому є меншою за точність ММП-оцінок. Проте при платокуртичному (плосковершинному) характері розподілів для достатньо широкого діапазону зміни величини параметру форми ЕСР теоретична точність ММПл-оцінокі ММП-оцінок фактично не відрізняється.