Statystyka matematyczna - wykład dziewiąty[1]
Klasyczny model regresji liniowej.
kierunek: matematyka I°
specjalność: matematyka finansowa

dr Jarosław Kotowicz

Instytut Informatyki, Uniwersytet w Białymstoku

[1] ©J.Kotowicz

Spis treści

Motywacja

Do oceny współzależności zmiennych może być użyta analiza rozkładów warunkowych zmiennych określonych w tablicy korelacyjnej.
Porównanie średnich warunkowych

1. Jeżeli $\overline{x_1} = \ldots = \overline{x_l} = \overline{x}$, to zmienna Y nie wpływa na zmienną X.
2. Jeżeli $\overline{y_1} = \ldots = \overline{y_k} = \overline{y}$, to zmienna X nie wpływa na zmienną Y.

Jeśli cechy są skorelowane, to średnie warunkowe zmiennej uznanej za zależną będą różne. Zależność jest tym silniejsza, im mocniej różne wartości przyjmowane przez cechę niezależną różnicują średni poziom wartości cechy zależnej.

Uwaga 1

Średnie warunkowe cechy zależnej możemy traktować jako funkcje wartości cechy niezależnej (funkcje regresji I rodzaju).

Warunkowe wartości oczekiwane – zmienne dyskretne. I

Definicja 1

Warunkową wartością oczekiwaną zmiennej X pod warunkiem $\{Y = y_j\}$, oznaczaną $\mathbb{E}(X|Y = y_j)$, jest liczba wyrażona wzorem

$$\sum_{i=1}^{k} x_i P(X = x_i|Y = y_j).$$

Warunkową wartością oczekiwaną zmiennej Y pod warunkiem $\{X = x_i\}$, oznaczaną $\mathbb{E}(Y|X = x_i)$, jest liczba wyrażona wzorem

$$\sum_{j=1}^{l} y_j P(Y = y_j|X = x_i).$$

Warunkowe wartości oczekiwane – zmienne dyskretne. II

Twierdzenie 1

Mamy

$$\mathbb{E}(X|Y = y_j) = \sum_{i=1}^{k} x_i p_{i|j}.$$

oraz

$$\mathbb{E}(Y|X = x_i) = \sum_{j=1}^{l} y_j p_{j|i}.$$

Warunkowe wartości oczekiwane – zmienne ciągłe. I

Definicja 2

Warunkową wartością oczekiwaną zmiennej X pod warunkiem $\{Y = y\}$, oznaczaną $\mathbb{E}(X|Y = y)$, jest liczba wyrażona wzorem

$$\int_{-\infty}^{+\infty} xf(x|y)dx.$$

Warunkową wartością oczekiwaną zmiennej Y pod warunkiem $\{X = x\}$, oznaczaną $\mathbb{E}(Y|X = x)$, jest liczba wyrażona wzorem

$$\int_{-\infty}^{+\infty} yf(y|x)dy.$$

Warunkowe wartości oczekiwane – zmienne ciągłe. II

Uwaga 2

Przypomnijmy definicję gęstości warunkowej:
Niech dwuwymiarowa zmienna losowa (X, Y) *ma rozkład ciągły z gęstością* f *tzn.* f *jest funkcją dwóch zmiennych* x *i* y*. Gęstością rozkładu warunkowego* X *pod warunkiem* $Y = y$ *nazywamy funkcję określoną dla* $x \in \mathbb{R}$ *wzorem*

$$f(x|y) = \begin{cases} \frac{f(x,y)}{f_Y(y)} & gdy\ f_Y(y) > 0 \\ 0 & w\ p.p. \end{cases}$$

gdzie $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$ *jest gęstością rozkładu brzegowego* Y*.*
Analogicznie definiujemy gęstością rozkładu warunkowego Y *pod warunkiem* $X = x$*.*

Funkcje regresji I rodzaju. I

Definicja 3

Funkcją regresji I rodzaju zmiennej X względem zmiennej Y nazywamy warunkową wartość oczekiwaną zmiennej X jako funkcję wartości zmiennej Y wyrażoną wzorem

$$m_1(y) := \mathbb{E}(X|Y = y_j) \text{ dla zmiennej } Y \text{ dyskretnej}, \quad (1)$$

$$m_1(y) := \mathbb{E}(X|Y = y) \text{ dla zmiennej } Y \text{ ciągłej}. \quad (2)$$

Funkcje regresji I rodzaju. II

Definicja 4

Funkcją regresji I rodzaju zmiennej Y względem zmiennej X nazywamy warunkową wartość oczekiwaną zmiennej Y jako funkcję wartości zmiennej X wyrażoną wzorem

$$m_2(x) := \mathbb{E}(Y|X = x_i) \text{ dla zmiennej } X \text{ dyskretnej}, \tag{3}$$

$$m_2(x) := \mathbb{E}(Y|X = x) \text{ dla zmiennej } X \text{ ciągłej}. \tag{4}$$

Uwaga 3

1. *Empiryczna krzywa regresji cechy X względem cechy Y jest to łamana łącząca punkty $(\overline{x_j}, y_j)$ dla $j = 1, \ldots, l$.*
2. *Empiryczna krzywa regresji cechy Y względem cechy X jest to łamana łącząca punkty $(x_i, \overline{y_i})$ dla $i = 1, \ldots, k$.*

Stosunki korelacyjne zmiennych

Jeżeli regresja zmiennych jest nieliniowa, to do pomiaru siły zależności cech wykorzystuje się tzw. stosunki (wskaźniki) korelacyjne wykorzystujące funkcję regresji.
Wykorzystując w tym celu zależności

$$\begin{aligned}\mathbb{E}((Y-\mathbb{E}(Y))^2) &= \mathbb{E}((m_2(x)-\mathbb{E}(Y))^2)+\mathbb{E}((Y-m_2(x))^2),\\ \mathbb{E}((X-\mathbb{E}(X))^2) &= \mathbb{E}((m_1(y)-\mathbb{E}(X))^2)+\mathbb{E}((X-m_1(y))^2).\end{aligned}$$

mamy

$$\begin{aligned}\eta_{xy} &:= \sqrt{\frac{\mathbb{E}(m_1(y)-\mathbb{E}(X))^2}{\mathbb{D}^2(X)}}=\sqrt{1-\frac{\mathbb{E}(X-m_1(y))^2}{\mathbb{D}^2(X)}},\\ \eta_{yx} &:= \sqrt{\frac{\mathbb{E}(m_2(x)-\mathbb{E}(Y))^2}{\mathbb{D}^2(Y)}}=\sqrt{1-\frac{\mathbb{E}(Y-m_2(x))^2}{\mathbb{D}^2(Y)}}.\end{aligned}$$

Motywacja

W celu stwierdzenia, czy istnieje zależność między cechami i określenia jej siły dla zmiennej uznanej za zależną można przeprowadzić analizę wariancji.

Dla zmiennej X.

Dokonujemy podziału całkowitej sumy kwadratów odchyleń od średniej na sumę kwadratów międzygrupową i wewnątrz grupową tj.

$$\sum_{i=1}^{k}(x_i - \overline{x})^2 \cdot n_{i\cdot} = \sum_{j=1}^{l}(\overline{x_j} - \overline{x})^2 \cdot n_{\cdot j} + \sum_{j=1}^{l}\sum_{i=1}^{k}(x_i - \overline{x_j})^2 \cdot n_{ij}, \qquad (5)$$

gdzie $\overline{x}$ średnia w brzegowym rozkładzie cechy X, $\overline{x_j}$ średnie w warunkowych rozkładach cechy X.
Pierwszy składnik sumy po prawej stronie jest to zróżnicowanie wyjaśniające regresje, a drugi to zróżnicowanie nie wyjaśniane regresją.

Dla zmiennej Y.

$$\sum_{j=1}^{l}(y_j - \overline{y})^2 \cdot n_{\cdot j} = \sum_{i=1}^{k}(\overline{y_i} - \overline{y})^2 \cdot n_{i\cdot} + \sum_{i=1}^{k}\sum_{j=1}^{l}(y_j - \overline{y_i})^2 \cdot n_{ij}. \qquad (6)$$

Uwaga 4

Zróżnicowanie nie wyjaśniane regresją reprezentuje rozrzut indywidualnych wartości zmiennej uznawanej za zależną wokół empirycznej linii regresji.

Definicja stosunków korelacyjnych. I

Definicja 5

Wskaźnikiem (stosunkiem) korelacyjnym zmiennej zależnej X względem zmienne Y nazywamy liczbę e_{xy} taką, że

$$e_{xy}^2 := \frac{\sum_{j=1}^{l} (\overline{x_j} - \overline{x})^2 \cdot n_{.j}}{\sum_{i=1}^{k} (x_i - \overline{x})^2 \cdot n_{i.}} = 1 - \frac{\sum_{j=1}^{l} \sum_{i=1}^{k} (x_i - \overline{x}_j)^2 n_{ij}}{\sum_{i=1}^{k} (x_i - \overline{x})^2 \cdot n_{i.}}. \tag{7}$$

Definicja stosunków korelacyjnych. II

Definicja 6

Wskaźnikiem (stosunkiem) korelacyjnym zmiennej zależnej Y względem zmienne X nazywamy liczbę e_{yx} taką, że

$$e_{yx}^2 := \frac{\sum_{i=1}^{k} (\overline{y_i} - \overline{y})^2 \cdot n_{i\cdot}}{\sum_{j=1}^{l} (y_j - \overline{y})^2 \cdot n_{\cdot j}} = 1 - \frac{\sum_{i=1}^{k} \sum_{j=1}^{l} (y_j - \overline{y}_i)^2 n_{ij}}{\sum_{j=1}^{l} (y_j - \overline{y})^2 \cdot n_{\cdot j}}. \tag{8}$$

Uwaga 5

1. *Ponieważ $e_{xy}^2 \in [0,1]$ i $e_{yx}^2 \in [0,1]$, więc $e_{xy}, e_{yx} \in [0,1]$.*
2. *Stosunki korelacyjne e_{xy} i e_{yx} nie muszą być równe.*

Testowanie hipotez o niezależności jednej zmiennej od drugiej. I

Będziemy weryfikować hipotezę, że zmienna losowa X w populacji generalnej jest niezależna od zmiennej Y.
Hipotezę traktuje się, jako przypuszczenie, że stosunek korelacyjny zmiennej X względem zmiennej Y w populacji generalnej równe są zero.
Mamy więc

$$H_0: \quad \eta_{xy} = 0$$
$$H_1: \quad \eta_{xy} > 0.$$

W celu weryfikacji tej hipotezy stosowana jest statystyka

$$F = \frac{\frac{e_{xy}^2}{l-1}}{\frac{1-e_{xy}^2}{n-l}},$$

Testowanie hipotez o niezależności jednej zmiennej od drugiej. II

gdzie n jest liczebnością próby, l liczba warunkowych rozkładów zmiennej X, e_{xy} stosunek korelacyjny z próby.
Jeżeli hipoteza zerowa jest prawdziwa, to statystyka F ma rozkład Fishera-Snedecora o $l-1$ stopniach swobody licznika i $n-l$ stopniach swobody mianownika.
Dla poziomu istotności α obszar krytyczny określony jest przez zależność

$$P(\{F \geqslant F_\alpha\}) = \alpha.$$

Testowanie hipotez o niezależności jednej zmiennej od drugiej. III

Uwaga 6

W przypadku weryfikowani hipotezy, że zmienna losowa Y w populacji generalnej jest niezależna od zmiennej X mamy następujące hipotezy (zerowa i alternatywna):

$$H_0: \quad \eta_{yx} = 0,$$
$$H_1: \quad \eta_{yx} > 0.$$

A w celu weryfikacji tej hipotezy stosowana jest statystyka

$$F = \frac{\frac{e_{yx}^2}{k-1}}{\frac{1-e_{yx}^2}{n-k}},$$

Testowanie hipotez o niezależności jednej zmiennej od drugiej. IV

gdzie n jest liczebnością próby, k liczba warunkowych rozkładów zmiennej Y, e_{yx} stosunek korelacyjny z próby.

Miara siły korelacji nieliniowej

W przypadku nieliniowej zależności wartość współczynnika korelacji jest niższa niż powinno to wynikać z siły związku między cechami.
Miarą siły korelacji nieliniowej są stosunki korelacyjne:

$$\widehat{m}_{xy} = e_{xy}^2 - r^2 \tag{9}$$

$$\widehat{m}_{yx} = e_{yx}^2 - r^2, \tag{10}$$

gdzie $\widehat{m}_{xy}, \widehat{m}_{yx} \in [0, 1]$.
Im bliższe jedności są wartości tych wskaźników, tym bardziej związki między cechami odchylają się od zależności liniowej.

Spis treści

Motywacja

Ostateczny cel analizy regresji to narzędzie predykcji, czyli przewidywanie jakie wartości przyjmie zmienna zależna przy ustalonych wartościach zmiennej (zmiennych) uznanych za niezależną (niezależne).
Stosowana jest konstrukcja tzw. **modeli regresji**, które wyjaśnią w sposób analityczny kształtowanie się wartości jednej zmiennej losowej pod wpływem innej lub innych zmiennych losowych.
Spośród wielu możliwych postaci modelu regresji podstawowe znacznie ma tzw. **model klasycznej regresji liniowej**.

Klasyczny model regresji liniowej – przypadek dwuwymiarowy

Model

Dla każdej ustalonej wartości jednej zmiennej losowej (np. X - zmienna niezależna) druga zmienna losowa (Y - zmienna zależna) ma warunkowy rozkład z wartością oczekiwaną

$$\mathbb{E}(Y|X = x) = \beta_1 x + \beta_0, \tag{11}$$

gdzie funkcja regresji I-go rodzaju zmiennej Y względem zmiennej X jest liniowa (β_1 – współczynnik regresji liniowej)
oraz stałą wariancję

$$\mathbb{D}^2(Y|X = x) = \sigma^2 \tag{12}$$

niezależna od x.

Uwaga

1. Zmienna Y traktujemy jako zmienną zależną, a zmienną X jako niezależną.
2. Współczynnik regresji liniowej β_1 jest wielkością o jaką zmienia się warunkowa wartość oczekiwana zmiennej zależnej Y, gdy x wzrasta o jednostkę.
3. Istotą klasycznego podejścia do zagadnienia regresji jest traktowanie wartości zmiennej niezależnej, jako wartości z góry ustalonych, czyli nielosowych.

Klasyczny model normalnej regresji liniowej

Oprócz klasycznego modelu regresji liniowej będziemy rozważać jeszcze jeden model.

Definicja 7

Jeżeli oprócz warunków (11) *i* (12) *będziemy zakładać, że rozkłady warunkowe zmiennej* Y *są normalne, tzn.* Y *dla* $X = x$ *ma rozkład* $\mathcal{N}(\beta_1 x + \beta_0, \sigma)$*, to będziemy mówili wtedy o klasycznym modelu normalnej regresji liniowej.*

Alternatywne sformułowanie klasycznego modelu regresji liniowej. I

Niech ciąg par $(x_1, Y_1), (x_2, Y_2), \ldots, (x_n, Y_n)$ będzie n - elementową próbą losową z populacji dwuwymiarowej, stanowiącą podstawę estymacji parametrów badanej zależności (wartości zmiennej X są w próbie ustalone).
Kształtowanie się wartości Y_i w próbie można wyjaśnić następująco

$$Y_i = \mathbb{E}(Y|X = x_i) + \varepsilon_i = \beta_1 x_i + \beta_0 + \varepsilon_i, \tag{13}$$

gdzie $i \in \overline{1, n}$ i ε_i są zmiennymi losowymi takimi, że

$$\forall_{i \in \overline{1,n}} \quad \mathbb{E}(\varepsilon_i) = 0 \tag{14}$$

$$\forall_{i \in \overline{1,n}} \quad \mathbb{D}^2(\varepsilon_i) = \sigma^2, \tag{15}$$

$$\operatorname{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ dla dowolnych } i \neq j. \tag{16}$$

Alternatywne sformułowanie klasycznego modelu regresji liniowej. II

Jest to alternatywne sformułowanie klasycznego modelu regresji liniowej Y względem X.

Alternatywne sformułowanie klasycznego modelu normalnej regresji liniowej

Jeżeli warunki (13), (14), (15), (16) uzupełnimy o założenie, że ε_i, dla $i \in \overline{1, n}$, maja rozkład $\mathcal{N}(0, \sigma)$, to otrzymujemy klasyczny model normalnej regresji liniowej zmiennej Y względem zmiennej X.

Równoważność warunków modeli

$$\begin{aligned}\mathbb{E}(Y_i) &= \mathbb{E}(\beta_1 x_i + \beta_0 + \varepsilon_i) = \beta_1 x_i + \beta_0 + \mathbb{E}(\varepsilon_i) \\ &= \beta_1 x_i + \beta_0,\end{aligned}$$

gdzie w pierwszej równości skorzystaliśmy z (13), w drugiej z liniowości wartości oczekiwanej i faktu, że wartości zmiennej niezależnej są nielosowe (deterministyczne), a w trzecie z (14).
Podobnie wykorzystując (15) otrzymujemy

$$\mathbb{D}^2(Y_i) = \mathbb{E}\left[Y_i - \mathbb{E}(Y_i)\right]^2 = \mathbb{E}\left[Y_i - \beta_1 x_i + \beta_0\right]^2 = \mathbb{E}(\varepsilon_i^2) = \sigma^2.$$

Estymacja parametrów β_1 i β_0. I

Załóżmy, że w populacji dwuwymiarowej (X, Y) pobieramy n - elementową próbę $(x_1, Y_1), (x_2, Y_2), \ldots, (x_n, Y_n)$. Wyniki konkretnej próby $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ można przedstawić w układzie współrzędnych otrzymując w ten sposób wykres rozrzutu punktów empirycznych. Szukamy wykresy prostej *"najlepiej dopasowanej"* do otrzymanych punktów, stosując metodę najmniejszych kwadratów (**MNK**). Będziemy minimalizować funkcję

$$SSE \equiv SSE(\beta_1, \beta_0) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [Y_i - (\beta_1 x_i + \beta_0)]^2 . \qquad (17)$$

Uwaga 7

Nazwa funkcji SSE jest skrótem angielskiej od „Sum of Squares of Errors".

Estymacja parametrów β_1 i β_0. II

Licząc pochodne cząstkowe i przyrównując je do zera (warunek konieczny istnienia ekstremum) otrzymujemy

$$\begin{cases} \frac{\partial}{\partial \beta_1} SSE = -2 \sum_{i=1}^{n} x_i (Y_i - \beta_1 x_i - \beta_0) = 0 \\ \frac{\partial}{\partial \beta_0} SSE = -2 \sum_{i=1}^{n} (Y_i - \beta_1 x_i - \beta_0) = 0 \end{cases}.$$

Zastępujemy parametry β_1 i β_0 ich estymatorami $\widehat{\beta_1}$ i $\widehat{\beta_0}$ otrzymując

$$\begin{cases} \sum_{i=1}^{n} x_i Y_i = \widehat{\beta_1} \sum_{i=1}^{n} x_i^2 + \widehat{\beta_0} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} Y_i = \widehat{\beta_1} \sum_{i=1}^{n} x_i + n\widehat{\beta_0} \end{cases}.$$

Estymacja parametrów β_1 i β_0. III

Pierwsze równanie przekształcamy, a z drugiego równania wyznaczamy parametr $\widehat{\beta_0}$ mamy

$$\begin{cases} \sum_{i=1}^{n} x_i Y_i = \widehat{\beta_1} \sum_{i=1}^{n} x_i^2 + \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} Y_i\right) - \widehat{\beta_1}\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 \\ \widehat{\beta_0} = \overline{Y} - \widehat{\beta_1}\overline{x} \end{cases}.$$

Z pierwszego równania wyznaczamy parametr $\widehat{\beta_1}$ otrzymując

$$\begin{cases} \widehat{\beta_1} = \dfrac{\sum_{i=1}^{n} x_i Y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} \\ \widehat{\beta_0} = \overline{Y} - \widehat{\beta_1}\overline{x} \end{cases}.$$

Estymacja parametrów β_1 i β_0. IV

Ostatecznie otrzymujemy

$$\begin{cases} \widehat{\beta_1} = \dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} \\ \widehat{\beta_0} = \overline{Y} - \widehat{\beta_1}\overline{x} \end{cases}.$$

Dzieląc przez n licznik i mianownik w pierwszym z równań możemy zapisać inaczej rozwiązania

$$\begin{cases} \widehat{\beta_1} = \dfrac{cov(x,Y)}{s_X^2} \\ \widehat{\beta_0} = \overline{Y} - \widehat{\beta_1}\overline{x} \end{cases}.$$

Otrzymane wzory przedstawiają estymatory parametrów β_1 i β_0 metodą **MNK**.

Twierdzenie Gaussa-Markowa

Własności estymatorów parametrów β_1 i β_0 przedstawia twierdzenie Gaussa-Markowa.

Twierdzenie 2

W klasycznym modelu regresji liniowej najefektywniejszym nieobciążonym estymatorami współczynników regresji są estymatory uzyskane metodą najmniejszych kwadratów.

Odchylenia standardowe estymatorów $\widehat{\beta_1}$ i $\widehat{\beta_0}$

Miarą wielkości błędu losowego przy estymacji parametru przy pomocy estymatora jest odchylenie standardowe estymatora nazywane również **standardowym błędem oceny** parametru.
W naszym przypadku mamy

$$\mathbb{D}(\widehat{\beta_1}) = \sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}, \tag{18}$$

$$\mathbb{D}(\widehat{\beta_0}) = \sqrt{\frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{n \sum_{i=1}^{n}(x_i - \overline{x})^2}}. \tag{19}$$

Uwagi

1. Obie wielkości zależą od σ^2.
2. Obie wielkości zależą od liczebności próby n.
3. Obie wielkości zależą od rozproszenia w próbie obserwowanych wartości zmiennej niezależnej $\sum_{i=1}^{n}(x_i - \overline{x})^2$
4. Mogą być oszacowane dopiero po oszacowaniu σ^2.

Teoretyczne wartości zmiennej Y i reszty modelu. I

Liniowa funkcja regresji po oszacowaniu parametrów na podstawie próby wyraża się wzorem

$$\widehat{Y}_i = \widehat{\beta_1} x_i + \widehat{\beta_0}. \tag{20}$$

Definicja 8

*Wartości $\widehat{Y}_i$ nazywamy **teoretycznymi wartościami zmiennej** Y.*

Definicja 9

Zmienne losowe e_i, dla $i \in \overline{1, n}$, zadane warunkiem

$$e_i := Y_i - \widehat{Y}_i$$

*nazywamy **resztami modelu** lub **residuami**.*

Teoretyczne wartości zmiennej Y i reszty modelu. II

Biorąc równanie

$$\sum_{i=1}^{n} Y_i = \widehat{\beta_1} \sum_{i=1}^{n} x_i + n\widehat{\beta_0},$$

widzimy, że

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \widehat{Y_i}. \tag{21}$$

Stąd suma reszt model spełnia równanie

$$\sum_{i=1}^{n} e_i = 0. \tag{22}$$

Uwaga

Rozważając równanie określające estymator $\widehat{\beta_0}$ otrzymujemy następujący fakt

Fakt 1

Wykres funkcji regresji z próby przechodzi przez punkty $(\overline{x}, \overline{Y})$.

Estymacja σ^2

Podstawą estymacji wariancji składników losowych σ^2 są reszty $e_i = Y_i - \widehat{Y}_i$.
Obliczając

$$\begin{aligned} \mathbb{E}(\sum_{i=1}^{n} e_i^2) &= \mathbb{E}(\sum_{i=1}^{n} Y_i^2 - 2\sum_{i=1}^{n} Y_i \widehat{Y}_i + \sum_{i=1}^{n} \widehat{Y}_i^2) = \dots \\ &= \sigma^2(n-2). \end{aligned}$$

Tak więc nieobciążonym estymatorem parametru σ^2 jest wariancja reszt

$$S_e^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-2}. \tag{23}$$

Estymacja $\mathbb{D}(\widehat{\beta_1}), \mathbb{D}(\widehat{\beta_0})$

Natomiast odchylenie standardowe reszt

$$S_e = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-2}}$$

można wykorzystać do estymacji standardowych błędów ocen parametrów β_1 i β_0, czyli $\mathbb{D}(\widehat{\beta_1})$ i $\mathbb{D}(\widehat{\beta_0})$.
Otrzymujemy wtedy

$$S_{\widehat{\beta_1}} = \sqrt{\frac{S_e^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}} \tag{24}$$

$$S_{\widehat{\beta_0}} = \sqrt{\frac{S_e^2 \sum_{i=1}^{n} x_i^2}{n \sum_{i=1}^{n}(x_i - \overline{x})^2}}. \tag{25}$$

Uwagi

1. Dokładność estymacji parametrów β_1 i β_0 jest tym większa,
   - im mniejsza jest wariancja reszt,
   - im większa jest próba,
   - im większy zakres zmienności zmiennej niezależnej X.