

Dopasowanie funkcji liniowej i wielomianu

Dopasowanie funkcji liniowej

Zmienna Y jest liniową funkcją X

$$Y = A_0 + A_1 \cdot X$$

Wyniki pomiarów Y mają rozkład normalny o dyspersji σ_Y , która w ogólnym przypadku również zależy od X . Wartości X nie są obciążone błędami lub, w najgorszym wypadku, można je pominąć w porównaniu z błędami pomiaru wielkości Y . W wyniku wykonanych pomiarów dysponujemy zestawem wartości

$$\{x_i, y_i, \sigma_{y_i}\}$$

gdzie $i = 1, \dots, n$, a σ_{y_i} są dyspersjami wartości mierzonych y_i .

Do znalezienia estymatorów a_0 i a_1 parametrów A_0 i A_1 możemy wykorzystać Metodę Największej Wiarygodności. Odpowiednie prawdopodobieństwo będzie teraz wynosiło

$$P(a_0, a_1; \{x_i, y_i, \sigma_{y_i}\}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_{y_i}} \exp\left[-\frac{(y_i - a_0 - a_1 x_i)^2}{2\sigma_{y_i}^2}\right]$$

Maksymalizacja tego prawdopodobieństwa odpowiada minimalizacji funkcji

$$\chi^2(a_0, a_1; \{x_i, y_i, \sigma_{y_i}\}) = \sum_{i=1}^n \left(\frac{y_i - a_0 - a_1 x_i}{\sigma_{y_i}} \right)^2$$

ze względu na wartości a_0 i a_1 .

Szukamy zatem rozwiązania układu równań

$$\frac{\partial \chi^2}{\partial a_0} = -2 \sum_{i=1}^n \left[\frac{1}{\sigma_{y_i}^2} (y_i - a_0 - a_1 x_i) \right] = 0$$

$$\frac{\partial \chi^2}{\partial a_1} = -2 \sum_{i=1}^n \left[\frac{x_i}{\sigma_{y_i}^2} (y_i - a_0 - a_1 x_i) \right] = 0$$

Można je przekształcić w układ równań liniowych dla szukanych estymatorów a_0 i a_1

$$\begin{aligned} a_0 \left(\sum_{i=1}^n \frac{1}{\sigma_{y_i}^2} \right) + a_1 \left(\sum_{i=1}^n \frac{x_i}{\sigma_{y_i}^2} \right) &= \left(\sum_{i=1}^n \frac{y_i}{\sigma_{y_i}^2} \right) \\ a_0 \left(\sum_{i=1}^n \frac{x_i}{\sigma_{y_i}^2} \right) + a_1 \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_{y_i}^2} \right) &= \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_{y_i}^2} \right) \end{aligned}$$

Wprowadzamy następujące oznaczenia

$$\begin{aligned} w_i &= 1/\sigma_{y_i}^2 \\ S_w &= \sum w_i \\ S_x &= \sum w_i x_i \\ S_y &= \sum w_i y_i \\ S_{xx} &= \sum w_i x_i^2 \\ S_{xy} &= \sum w_i x_i y_i \end{aligned}$$

i przy ich pomocy zapiszemy rozwiązanie układu

$$\begin{aligned} a_0 &= \frac{S_{xx} S_y - S_x S_{xy}}{S_w S_{xx} - S_x^2} \\ a_1 &= \frac{S_w S_{xy} - S_x S_y}{S_w S_{xx} - S_x^2} \end{aligned}$$

Wariancje estymatorów $\sigma_{a_0}^2$ i $\sigma_{a_1}^2$ można ustalić podobnie jak poprzednio, korzystając z reguły propagacji wariancji

$$\begin{aligned} \sigma_{a_0}^2 &= \sum_{i=1}^n \left[\left(\frac{\partial a_0}{\partial y_i} \right)^2 \sigma_{y_i}^2 \right] \\ \sigma_{a_1}^2 &= \sum_{i=1}^n \left[\left(\frac{\partial a_1}{\partial y_i} \right)^2 \sigma_{y_i}^2 \right] \end{aligned}$$

W tym celu należy znaleźć pochodne cząstkowe a_0 i a_1 względem wszystkich y_i .

$$\frac{\partial a_0}{\partial y_i} = \frac{w_i S_{xx} - w_i x_i S_x}{S_w S_{xx} - S_x^2}$$

$$\frac{\partial a_1}{\partial y_i} = \frac{w_i x_i S_w - w_i S_x}{S_w S_{xx} - S_x^2}$$

Po podstawieniu i wykonaniu przekształceń otrzymujemy

$$\sigma_{a_0}^2 = \sum_{i=1}^n \left[\left(\frac{\partial a_0}{\partial y_i} \right)^2 \sigma_{y_i}^2 \right] =$$

$$\sum_{i=1}^n \left[\left(\frac{w_i S_{xx} - w_i x_i S_x}{S_w S_{xx} - S_x^2} \right)^2 \frac{1}{w_i} \right] =$$

$$\frac{1}{(S_w S_{xx} - S_x^2)^2} \sum_{i=1}^n \left[\frac{w_i^2 S_{xx}^2 - 2w_i^2 x_i S_x S_{xx} + w_i^2 x_i^2 S_x^2}{w_i} \right] =$$

$$\frac{\sum_{i=1}^n (w_i S_{xx}^2 - 2w_i x_i S_x S_{xx} + w_i x_i^2 S_x^2)}{(S_w S_{xx} - S_x^2)^2} =$$

$$\frac{\sum_{i=1}^n (w_i S_{xx}^2) - 2 \sum_{i=1}^n (w_i x_i S_x S_{xx}) + \sum_{i=1}^n (w_i x_i^2 S_x^2)}{(S_w S_{xx} - S_x^2)^2} =$$

$$\frac{S_w S_{xx}^2 - 2S_x^2 S_{xx} + S_{xx} S_x^2}{(S_w S_{xx} - S_x^2)^2} =$$

$$\frac{S_{xx} (S_w S_{xx} - 2S_x^2 + S_x^2)}{(S_w S_{xx} - S_x^2)^2} = \frac{S_{xx}}{(S_w S_{xx} - S_x^2)}$$

$$\begin{aligned}
\sigma_{a_1}^2 &= \sum_{i=1}^n \left[\left(\frac{\partial a_1}{\partial y_i} \right)^2 \sigma_{y_i}^2 \right] = \\
&\sum_{i=1}^n \left[\left(\frac{w_i x_i S_w - w_i S_x}{S_w S_{xx} - S_x^2} \right)^2 \frac{1}{w_i} \right] = \\
&\frac{1}{(S_w S_{xx} - S_x^2)^2} \sum_{i=1}^n \left[\frac{w_i^2 x_i^2 S_w^2 - 2w_i^2 x_i S_w S_x + w_i^2 S_x^2}{w_i} \right] = \\
&\frac{\sum_{i=1}^n (w_i x_i^2 S_w^2 - 2w_i x_i S_w S_x + w_i S_x^2)}{(S_w S_{xx} - S_x^2)^2} = \\
&\frac{\sum_{i=1}^n (w_i x_i^2 S_w^2) - 2 \sum_{i=1}^n (w_i x_i S_w S_x) + \sum_{i=1}^n (w_i S_x^2)}{(S_w S_{xx} - S_x^2)^2} = \\
&\frac{S_w^2 S_{xx} - 2S_w S_x^2 + S_w S_x^2}{(S_w S_{xx} - S_x^2)^2} = \\
&\frac{S_w (S_w S_{xx} - 2S_x^2 + S_x^2)}{(S_w S_{xx} - S_x^2)^2} = \frac{S_w}{(S_w S_{xx} - S_x^2)}
\end{aligned}$$

Czyli wariancje estymatorów a_0 i a_1 wynoszą

$$\begin{aligned}
\sigma_{a_0}^2 &= \frac{S_{xx}}{S_w S_{xx} - S_x^2} \\
\sigma_{a_1}^2 &= \frac{S_w}{S_w S_{xx} - S_x^2}
\end{aligned}$$

Jeżeli parametry zależności Y od X są wielkościami, które mamy wyznaczyć na podstawie wyników pomiarów $\{x_i, y_i, \sigma_{y_i}\}$, to możemy zapisać:

$$a_0 = \frac{S_{xx}S_y - S_x S_{xy}}{S_w S_{xx} - S_x^2}, \quad u(a_0) = \sqrt{\frac{S_{xx}}{S_w S_{xx} - S_x^2}}$$

$$a_1 = \frac{S_w S_{xy} - S_x S_y}{S_w S_{xx} - S_x^2}, \quad u(a_1) = \sqrt{\frac{S_w}{S_w S_{xx} - S_x^2}}$$

Przykład. Rozkład Poisson'a - dopasowanie linii prostej

W pomiarach zależności intensywności promieniowania od odległości r od źródła rejestrowane szybkości zliczeń, o rozkładzie Poisson'a, zależą liniowo od odwrotności kwadratu odległości. Rejestracja przebiega w stałych odcinkach czasu i liczba zliczonych impulsów y_i ma rozkład

Poisson'a o średniej $ax_i + b$, gdzie $x_i = 1/r_i^2$. Wynika z tego, że

dyspersja i -tego pomiaru wynosi $\sigma_{y_i} = \sqrt{ax_i + b}$, a waga $w_i = \frac{1}{ax_i + b}$.

Informację o zależności błędu (wagi) pomiaru od zmiennej niezależnej można wykorzystać na dwa sposoby. Można te wagi wstawić do poprzednio ustalonego wyrażenia na χ^2 , otrzymanego dla rozkładu normalnego (zakładając, że liczby zliczeń są na tyle duże, że rozkład Poisson'a jest bliski normalnemu),

$$\chi^2(a, b) = \sum_{i=1}^n \frac{(y_i - ax_i - b)^2}{ax_i + b}$$

dostając po zróżniczkowaniu układ dwóch równań

$$\frac{\partial \chi^2}{\partial a} = \sum_{i=1}^n x_i - \sum_{i=1}^n \frac{x_i y_i^2}{(ax_i + b)^2} = 0$$

$$\frac{\partial \chi^2}{\partial b} = n - \sum_{i=1}^n \frac{y_i^2}{(ax_i + b)^2} = 0$$

które są nieliniowe ze względu na parametry a i b , i rozwiązać je można tylko metodami numerycznymi (kolejnych przybliżeń).

Drugi sposób polega na zastąpieniu rozkładu normalnego rozkładem Poisson'a o wartości średniej $\mu_i = y(x_i) = a \cdot x_i + b$ i ponownym skonstruowaniu funkcji rozkładu gęstości prawdopodobieństwa $p(a, b; \{x_i, y_i\})$ otrzymania zestawu wyników $\{x_i, y_i\}$ pod warunkiem, że parametry określające wartość średnią μ mają wartości a i b

$$p(a, b) = \prod_{i=1}^n \left(\frac{[y(x_i)]^{y_i}}{y_i!} e^{-y(x_i)} \right)$$

i bezpośrednim zastosowaniu MNW. Zamiast maksymalizować prawdopodobieństwo można równie dobrze maksymalizować jego logarytm, co znacznie upraszcza wzory

$$\begin{aligned} \ln p(a, b) &= \sum_{i=1}^n [y_i \ln y(x_i)] - \sum_{i=1}^n y(x_i) - \sum_{i=1}^n y_i! = \\ &= \sum_{i=1}^n [y_i \ln y(x_i)] - \sum_{i=1}^n y(x_i) + \text{const.} \end{aligned}$$

Po zróżniczkowaniu względem a otrzymujemy

$$\frac{1}{p} \frac{\partial p}{\partial a} = \sum_{i=1}^n \left[y_i \frac{1}{y(x_i)} \frac{\partial y(x_i)}{\partial a} \right] - \sum_{i=1}^n \left[\frac{\partial y(x_i)}{\partial a} \right].$$

Warunkiem zerowania się pochodnej $\frac{\partial p}{\partial a}$ jest zerowanie się prawej

strony. Biorąc pod uwagę, że $\frac{\partial y(x_i)}{\partial a} = x_i$ otrzymujemy pierwsze z układu równań

$$\sum_{i=1}^n x_i - \sum_{i=1}^n \frac{x_i y_i}{ax_i + b} = 0$$

Różniczkowanie logarytmu $\ln p(a, b)$ względem b daje:

$$\frac{1}{p} \frac{\partial p}{\partial b} = \sum_{i=1}^n \left[y_i \frac{1}{y(x_i)} \frac{\partial y(x_i)}{\partial b} \right] - \sum_{i=1}^n \left[\frac{\partial y(x_i)}{\partial b} \right]$$

a uwzględniając, że $\frac{\partial y(x_i)}{\partial b} = 1$ otrzymujemy drugie równanie

$$n - \sum_{i=1}^n \frac{y_i}{ax_i + b} = 0$$

Ostatecznie układ równań przyjmuje w tym przypadku postać

$$\sum_{i=1}^n x_i - \sum_{i=1}^n \frac{x_i y_i}{ax_i + b} = 0$$

$$n - \sum_{i=1}^n \frac{y_i}{ax_i + b} = 0$$

bardzo podobną do poprzedniego układu.

Różnicy między tymi dwoma sposobami można się spodziewać dla mniejszych liczb zliczonych impulsów, kiedy różnice między y_i a $a \cdot x_i + b$ stają się relatywnie większe. Wykorzystanie funkcji χ^2 , uwarunkowane normalnym rozkładem gęstości prawdopodobieństwa, wymaga większych liczb impulsów lub grupowania pomiarów. Uniknięcie problemów związanych z grupowaniem wyników i tym samym zmniejszeniem liczby punktów pomiarowych oraz ustalaniem właściwych wag jest możliwe przez bezpośrednie wykorzystanie Metody Największej Wiarygodności, jak w powyższym przykładzie.

Dopasowanie wielomianu

Jeżeli linia prosta nie daje się dobrze dopasować do danych pomiarowych to można spróbować dopasować bardziej złożoną funkcję, to jest wielomian stopnia m

$$y(x) = \sum_{k=0}^m a_k x^k$$

lub w bardziej ogólnej formie

$$y(x) = \sum_{k=0}^m a_k f_k(x)$$

gdzie $f_k(x)$ są dowolnymi funkcjami nie zawierającymi parametrów a_k .

Jeżeli wyniki pomiarów y_i mają rozkład normalny o dyspersji σ_{y_i} , to maksymalizacja funkcji gęstości prawdopodobieństwa

$$P(a_0, \dots, a_m; \{x_i, y_i, \sigma_{y_i}\}) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi}\sigma_{y_i}} \right) \cdot \exp \left[-\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_{y_i}^2} \left(y_i - \sum_{k=0}^m a_k f_k(x_i) \right)^2 \right]$$

sprowadza się do minimalizacji funkcji

$$\chi^2(a_0, \dots, a_m; \{x_i, y_i, \sigma_{y_i}\}) = \sum_{i=1}^n \left[\frac{1}{\sigma_{y_i}^2} \left(y_i - \sum_{k=0}^m a_k f_k(x_i) \right)^2 \right]$$

względem parametrów a_k .

Różniczkowanie χ^2 względem parametru a_l daje

$$\begin{aligned} \frac{\partial \chi^2}{\partial a_l} &= -2 \sum_{i=1}^n \left[\frac{1}{\sigma_{y_i}^2} \left(y_i - \sum_{k=0}^m a_k f_k(x_i) \right) f_l(x_i) \right] = \\ &= -2 \sum_{i=1}^n \left[y_i \frac{f_l(x_i)}{\sigma_{y_i}^2} \right] + 2 \sum_{i=1}^n \left[\frac{1}{\sigma_{y_i}^2} \left(\sum_{k=0}^m [a_k f_l(x_i) f_k(x_i)] \right) \right] = \\ &= -2 \sum_{i=1}^n \left[y_i \frac{f_l(x_i)}{\sigma_{y_i}^2} \right] + 2 \sum_{k=0}^m \left(a_k \sum_{i=1}^n \left[\frac{f_l(x_i) f_k(x_i)}{\sigma_{y_i}^2} \right] \right) \end{aligned}$$

W rezultacie otrzymujemy układ $m + 1$ równań liniowych (dla wskaźnika l przebiegającego wartości od 0 do m).

$$\sum_{i=1}^n \left[y_i \frac{f_l(x_i)}{\sigma_{y_i}^2} \right] = \sum_{k=0}^m \left(a_k \sum_{i=1}^n \left[\frac{f_l(x_i) f_k(x_i)}{\sigma_{y_i}^2} \right] \right)$$

Układ równań można zastąpić równoważnym mu równaniem macierzowym

$$\boldsymbol{\beta} = \mathbf{a} \boldsymbol{\alpha}$$

gdzie \mathbf{a} i $\boldsymbol{\beta}$ są macierzami jednowymiarowymi o $m + 1$ elementach (wektorami), a $\boldsymbol{\alpha}$ jest kwadratową i symetryczną macierzą $(m + 1) \times (m + 1)$. Elementami macierzy \mathbf{a} są szukane parametry dopasowania a_k , a elementy macierzy $\boldsymbol{\beta}$ i $\boldsymbol{\alpha}$ wynoszą odpowiednio

$$\beta_k = \sum_{i=1}^n \left[\frac{y_i f_k(x_i)}{\sigma_{y_i}^2} \right]$$

$$\alpha_{lk} = \alpha_{kl} = \sum_{i=1}^n \left[\frac{f_l(x_i) f_k(x_i)}{\sigma_{y_i}^2} \right]$$

Jeżeli macierz $\boldsymbol{\alpha}$ jest nieosobliwa (ma wyznacznik różny od zera), to możemy obie strony równania macierzowego pomnożyć prawostronnie przez macierz $\boldsymbol{\varepsilon}$ odwrotną do macierzy $\boldsymbol{\alpha}$ ($\boldsymbol{\varepsilon} = \boldsymbol{\alpha}^{-1}$, $\boldsymbol{\alpha} \boldsymbol{\alpha}^{-1} = \mathbf{1}$). Otrzymujemy wtedy

$$\boldsymbol{\beta} \boldsymbol{\varepsilon} = \mathbf{a} \boldsymbol{\alpha} \boldsymbol{\varepsilon} = \mathbf{a} \boldsymbol{\alpha} \boldsymbol{\alpha}^{-1} = \mathbf{a}$$

co daje

$$\mathbf{a} = \boldsymbol{\beta} \boldsymbol{\varepsilon} = \boldsymbol{\beta} \boldsymbol{\alpha}^{-1}$$

Rozwiązanie w formie macierzowej można przedstawić inaczej w formie skalarnej

$$a_l = \sum_{k=0}^m \beta_k \varepsilon_{kl}$$

Macierz $\boldsymbol{\alpha}$ nazywana jest macierzą krzywizny funkcji χ^2 w przestrzeni parametrów a_k . Powodem jest związek elementów α_{lk} z drugimi

pochođnymi funkcji χ^2 .

$$\alpha_{lk} = \sum_{i=1}^n \left[\frac{f_l(x_i) f_k(x_i)}{\sigma_{y_i}^2} \right] = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_l \partial a_k}$$

Macierz ε jest nazywana macierzą (wariancji-)kowariancji. Macierz ε jest również symetryczna. Jej elementy diagonalne są równe wariacjom estymatorów $a_0 \dots a_m$, a pozostałe kowariancjom między odpowiednimi parami estymatorów.

$$\sigma_{a_j}^2 = \varepsilon_{jj}$$

$$\sigma_{a_j a_l} = \varepsilon_{jl}$$

Jeżeli w szczególności dopasowywane funkcje $f_k(x)$ tworzą układ ortogonalny, to w macierzach α i ε znikają elementy poza główną przekątną. Oznacza to brak korelacji między parametrami i ułatwia rozszerzanie modelu opisanego równaniem $y(x) = \sum a_k f_k(x)$ o kolejne składniki.

Uwaga o linearyzacji modelu zależności między wielkościami mierzonymi

Przedstawione powyżej techniki dopasowania funkcji do wyników pomiarów były ograniczone tylko do funkcji, które są liniowe względem szukanych parametrów. W przypadku kiedy funkcja nie jest liniowa względem swoich parametrów stosuje się często przekształcenie zmiennych w taki sposób, żeby funkcja opisująca zależność przekształconych zmiennych była już liniowa względem parametrów. Jeżeli jednak przekształcana jest zmienna zależna (przekształceniem nieliniowym), to zniekształceniu ulegają różnice między pomiarami i modelem i rozwiązanie problemu dla modelu zlinearyzowanego nie odpowiada minimum χ^2 dla modelu oryginalnego.

Przykłady rachunkowe

Pomiary spadku napięcia wzdłuż drutu oporowego z prądem

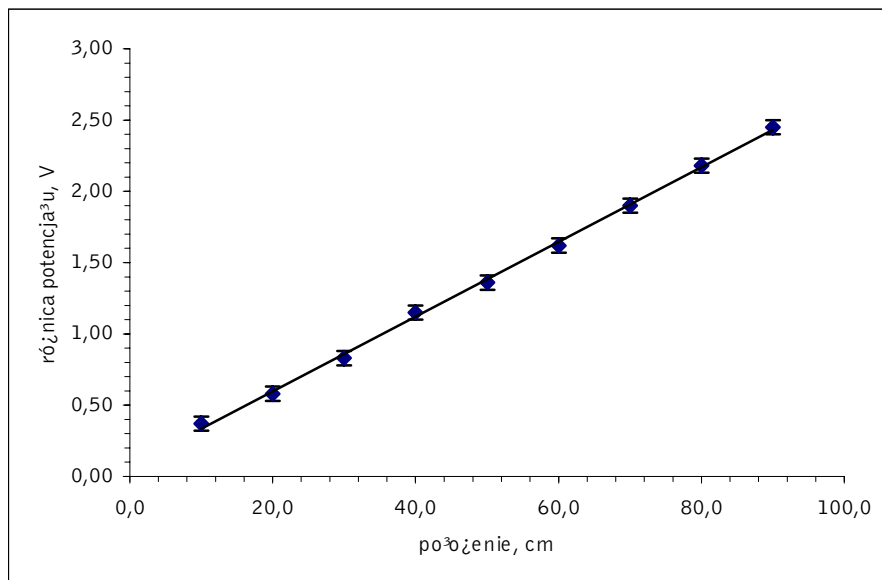
$$U = ax + b$$

Nr pomiaru	położenie cm	różnica potencjału V	niepewność pomiaru napięcia V	x_i^2	$x_i U_i$	dopasowane napięcie $ax_i + b$	kwadraty różnic
1	10,0	0,37	0,05	100	3,70	0,33	0,001328
2	20,0	0,58	0,05	400	11,60	0,60	0,000247
3	30,0	0,83	0,05	900	24,90	0,86	0,000778
4	40,0	1,15	0,05	1600	46,00	1,12	0,000897
5	50,0	1,36	0,05	2500	68,00	1,38	0,000494
6	60,0	1,62	0,05	3600	97,20	1,64	0,000595
7	70,0	1,90	0,05	4900	133,00	1,91	0,000043
8	80,0	2,18	0,05	6400	174,40	2,17	0,000127
9	90,0	2,45	0,05	8100	220,50	2,43	0,000365
$N = 9$	$S_x = 450,0$	$S_U = 12,44$		$S_{xx} = 28500$	$S_{xU} = 779,30$		$\Sigma = 0,004874$

$$\Delta = N \sum x_i^2 - (\sum x_i)^2 = 54000$$

$$b = (\sum x_i^2 \sum U_i - \sum x_i \sum x_i U_i) / \Delta = 7,14 \cdot 10^{-2} \quad \sigma_U^2 \sum x_i^2 / \Delta = 1,319 \cdot 10^{-3} \quad u(b) = 3,63 \cdot 10^{-2} \text{ V}$$

$$a = (N \sum x_i U_i - \sum x_i \sum U_i) / \Delta = 262 \cdot 10^{-4} \quad N \sigma_U^2 / \Delta = 4,167 \cdot 10^{-7} \quad u(a) = 6,45 \cdot 10^{-4} \text{ V/cm}$$



Liczby zliczeń impulsów licznika GM w jednakowych odcinkach czasu w funkcji odległości źródła cząstek od licznika

$$N = a \frac{1}{d^2} + b$$

Nr pomiaru i	odległość d (cm)	$x = 1/d^2$	zliczenia N	$u(N)$	w_i	$w_i x_i$	$w_i x_i^2$	$w_i N_i$	$w_i x_i N_i$	dopasowana liczba zliczeń $ax_i + b$	kwadrat różnicy
1	20	25,00	901	30,0	0,00111	0,0277	0,6937	1,0	25,00	886,94	0,22
2	25	16,00	652	25,5	0,00153	0,0245	0,3926	1,0	16,00	610,66	2,62
3	30	11,11	443	21,0	0,00226	0,0251	0,2787	1,0	11,11	460,58	0,70
4	35	8,16	339	18,4	0,00295	0,0241	0,1966	1,0	8,16	370,09	2,85
5	40	6,25	283	16,8	0,00353	0,0221	0,1380	1,0	6,25	311,36	2,84
6	45	4,94	281	16,8	0,00356	0,0176	0,0868	1,0	4,94	271,09	0,35
7	50	4,00	240	15,5	0,00417	0,0167	0,0667	1,0	4,00	242,29	0,02
8	60	2,78	220	14,8	0,00455	0,0126	0,0351	1,0	2,78	204,77	1,05
9	75	1,78	180	13,4	0,00556	0,0099	0,0176	1,0	1,78	174,07	0,20
10	100	1,00	154	12,4	0,00649	0,0065	0,0065	1,0	1,00	150,19	0,09
$n=10$			Sumy		$S_w=0,03570$	$S_x=0,1868$	$S_{xx}=1,9122$	$S_y=10,0$	$S_{xy}=81,02$		$S_s=10,95$

$$\Delta = S_w S_{xx} - (S_x)^2 = 0,03339$$

$$b = S_y S_{xx} - S_x S_{xy} = 119,50 \quad u(b) = \sqrt{S_{xx} / \Delta} = 7,57$$

$$a = S_w S_{xy} - S_x S_y = 30,70 \quad u(a) = \sqrt{S_w / \Delta} = 1,03$$

