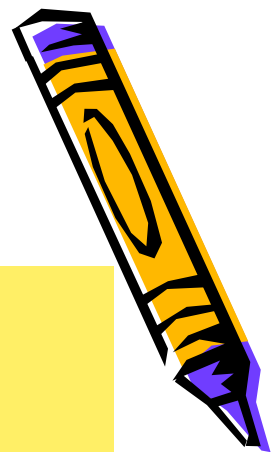


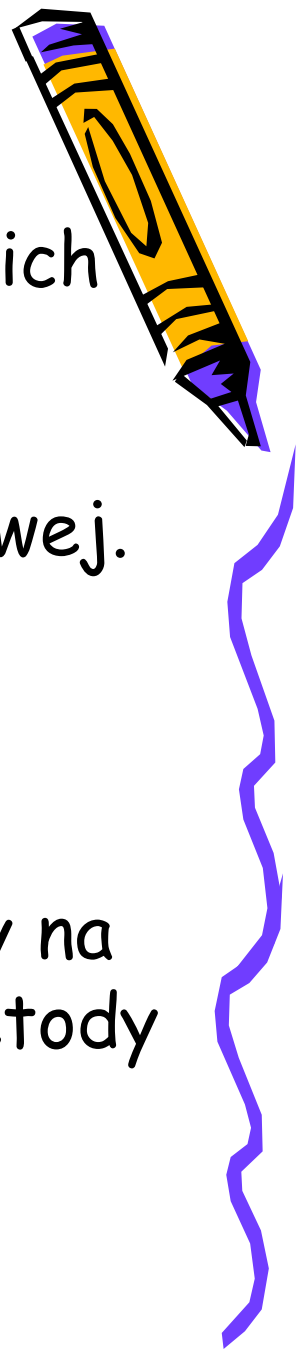
Ekonometria

Wykład 5

dr hab. Małgorzata Radziukiewicz, prof. PSW Białą Podlaska



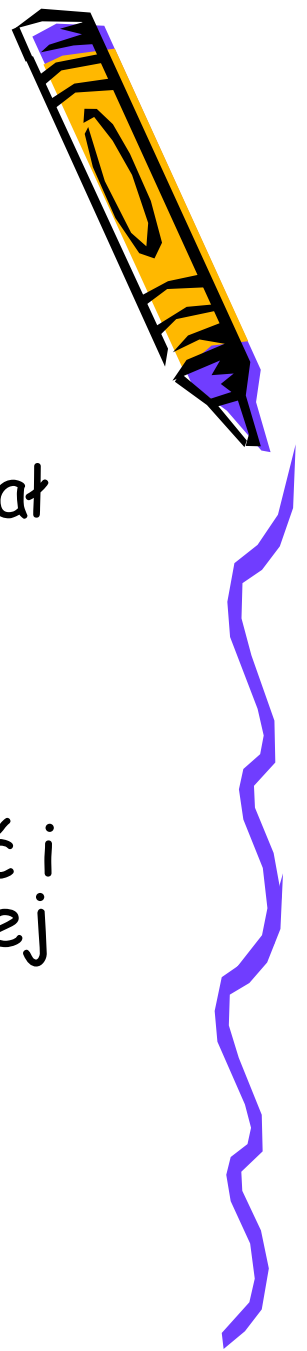
Estymacja - po co i dlaczego?



- Jeśli jesteśmy w stanie zebrać wszystkich informacji na temat interesującej nas zbiorowości wówczas do pełnego opisu wystarczą nam metody statystyki opisowej.
- W wielu jednak sytuacjach mówiąc o zbiorowości opieramy się na danych pochodzących z próby.
- Aby prawidłowo uogólniać wyniki z próby na populację generalną należy stosować metody statystyki matematycznej.



Estymacja - po co i dlaczego?

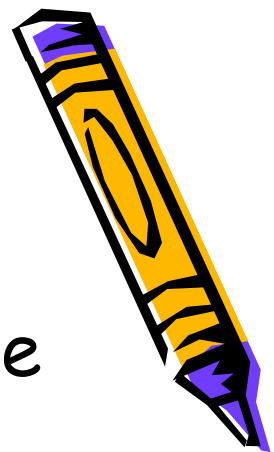


- Procedur uogólniania wyników z próby losowej na całą zbiorowość dostarcza dział wnioskowania statystycznego.
- Estymacja zatem to dział **wnioskowania statystycznego** będący zbiorem metod pozwalających na uogólnianie wyników badania **próby losowej** na **nieznaną** postać i parametry **rozkładu zmiennej losowej** całej **populacji** oraz szacowanie błędów wynikających z tego uogólnienia.

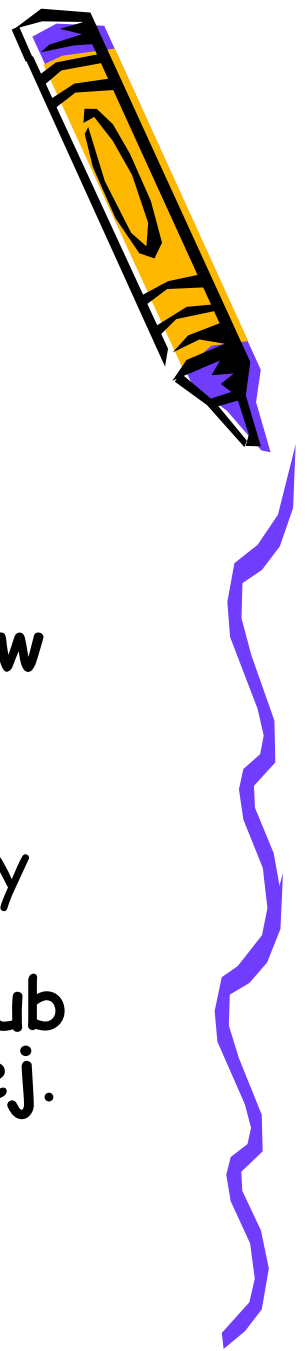


W zależności od szukanej cechy rozkładu można podzielić **metody estymacji** na dwie grupy:

- **Estymacja parametryczna** - metody znajdowania nieznanymi wartości parametrów rozkładu
- **Estymacja nieparametryczna** - metody znajdowania postaci rozkładu populacji

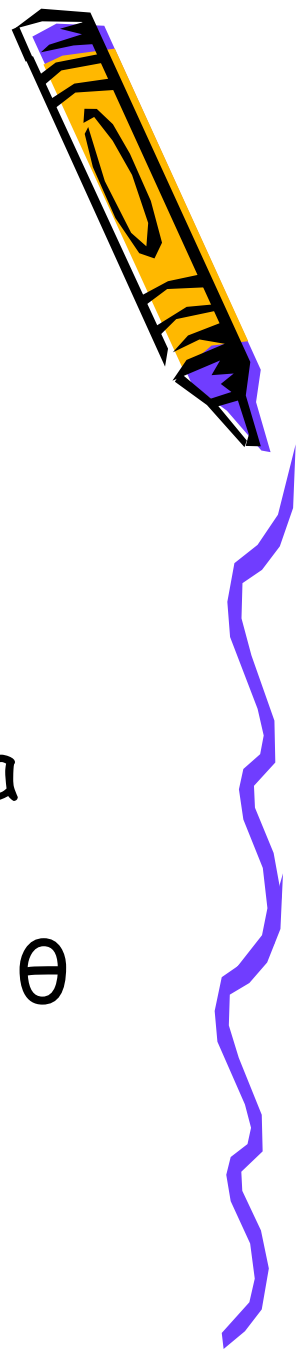


Estymacja - po co i dlaczego?



- Wnioskowanie przybiera postać:
 1. **estymacji parametrów statystycznych** czyli szacowania nieznanymi wartościami parametrów np. średniej arytmetycznej w zbiorowości generalnej, odchylenia standardowego.
 2. **testowania hipotez**, które z kolei dotyczy weryfikacji przypuszczeń odnośnie określonego poziomu zmiennej losowej lub kształtu rozkładu w populacji generalnej.





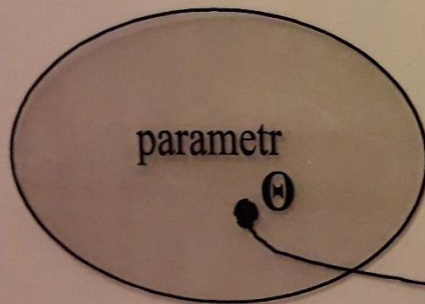
- Zatem losujemy z N -elementowej populacji generalnej n -elementową próbę losową
- Ze względu na niemożność poznania parametru θ z populacji generalnej wnioskujemy o wartości parametru θ w oparciu o zbadanie próby



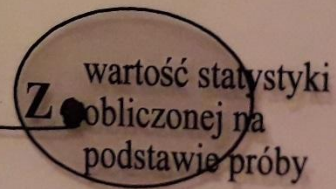
Wnioskowanie o wartości parametru Θ (lub innych parametrów populacji generalnej) w oparciu o próbę losową nosi nazwę szacowania (lub estymacji) parametrów

N
populacja generalna

n
próba losowa



Z - liczba stanowiąca
oszacowanie parametru Θ

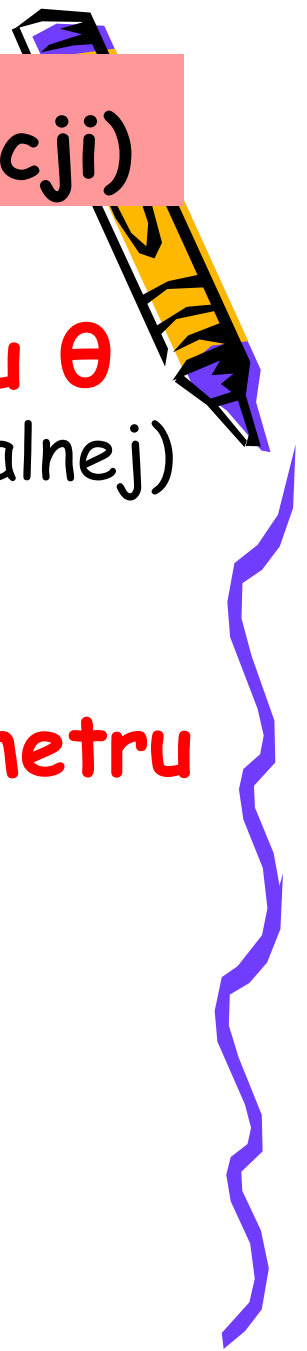


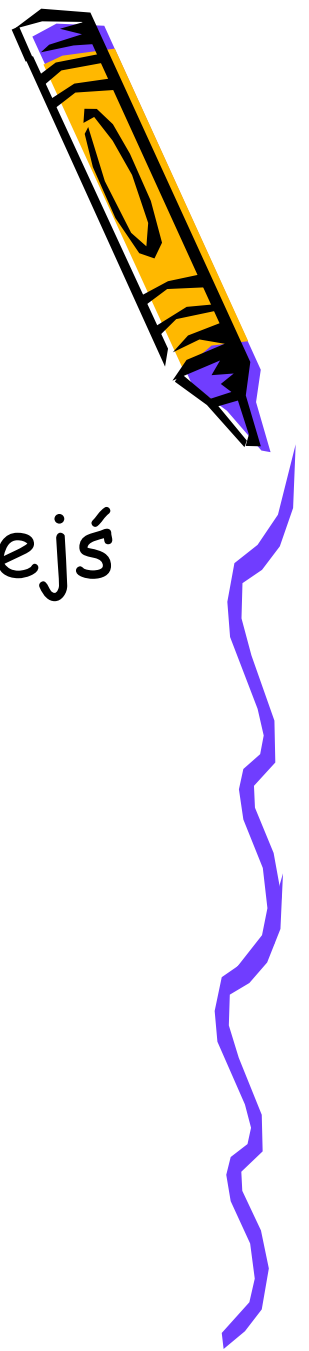
przyporządkowanie

Θ - parametr w populacji generalnej

dwa podejścia szacowania (estymacji)

- 1. **punktowe szacowanie parametru θ**
(lub innych parametrów populacji generalnej)
- podajemy jedną liczbę odpowiadającą przypuszczalnej wartości parametru
- 2. **przedziałowe szacowanie parametru**
- podajemy pewien przedział, w którym przypuszczalnie znajduje się prawdziwa wartość parametru

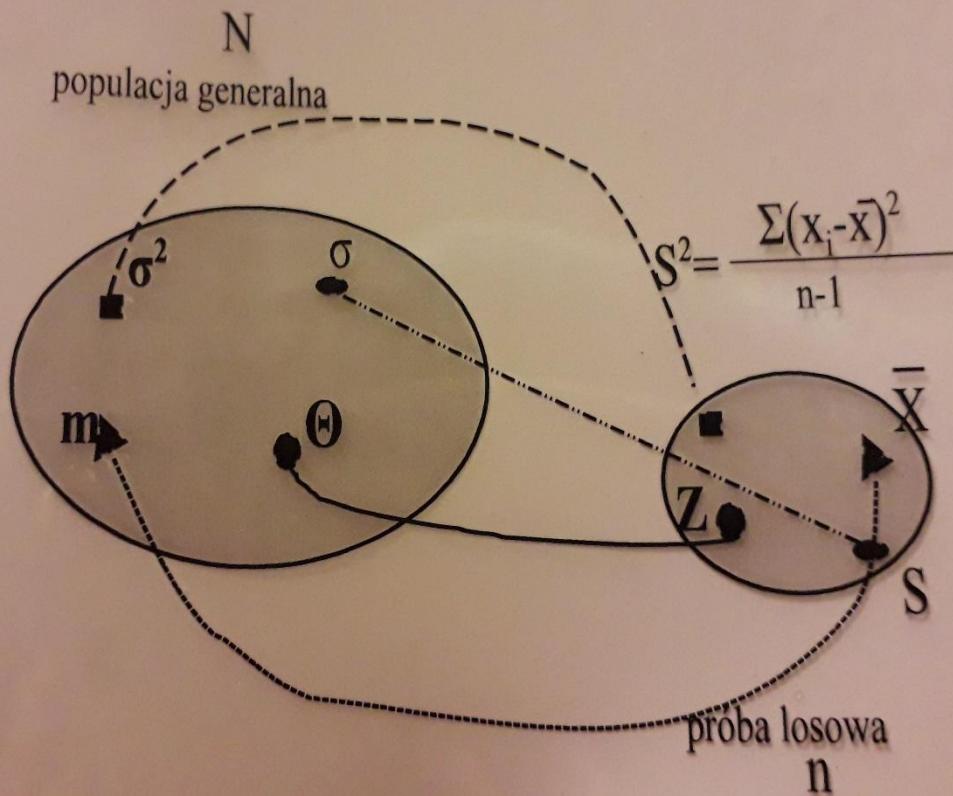




- Liczbą stanowiącą oszacowanie parametru θ musi być wartość jakiejś statystyki obliczonej na podstawie próby



Statystykę służącą do oszacowania parametru w populacji generalnej nazywamy estymatorem

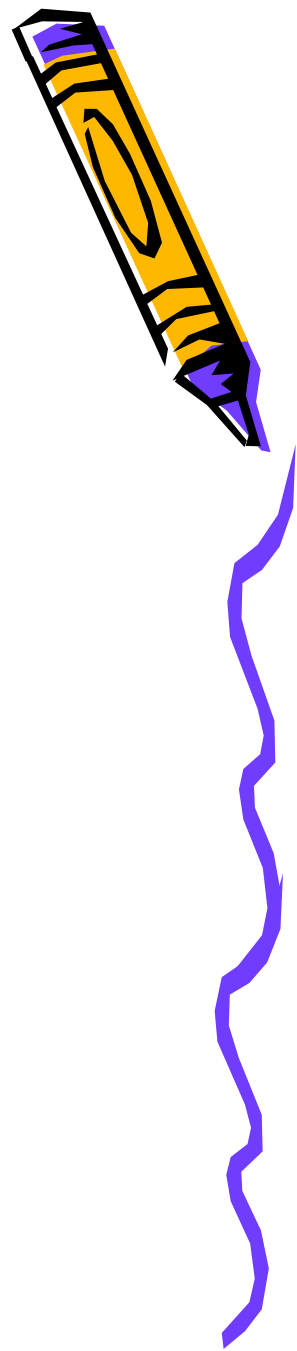


Estymator - szacowany parametr



- **Estymator** - wielkość (charakterystyka, miara), obliczona na podstawie próby, służąca do oceny wartości nieznanymi parametrów populacji generalnej.

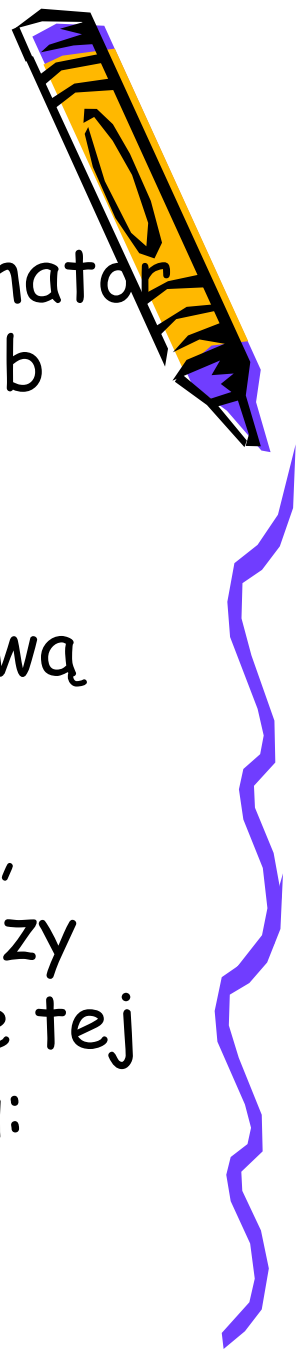




- Estymator, jak każda statystyka z próby ma pewien rozkład.
- Zadanie: - **jak dobrać estymator, aby jego rozkład gwarantował najlepsze oszacowanie?**



Własności dobrego estymatora



- Wartości, jakie może przyjmować estymator Z parametru θ są różne dla różnych prób pochodzących z tej samej populacji;
- Dlatego też nie można oczekiwać, że otrzymany estymator Z będzie prawdziwą wartością estymowanego parametru θ ;
- Powstaje więc błąd losowy parametru θ , który dla danej próby jest różnicą między oceną parametru dokonaną na podstawie tej próby a prawdziwą wartością parametru:

$$\varepsilon = Z - \theta$$



Pożądane cechy estymatora



- 1. **nieobciążoność** - aby estymator dawał gwarancję, że oszacowania nie będą w sposób systematyczny заниżane ani zawyżane;
- 2. **zgodność** - w miarę wzrostu próby (n) prawdopodobieństwo, że różnica między estymatorem a parametrem jest dowolnie mała, zbliża się do jedności;
- 3. **efektywność** - z 2-óch nieobciążonych estymatorów określonego parametru ten jest najefektywniejszy, który ma mniejszą wariancję.



Estymator nieobciążony



- Estymator nieobciążony to ten, którego przeciętna wartość jest dokładnie równa wartości szacowanego parametru tzn. zachodzi równość:

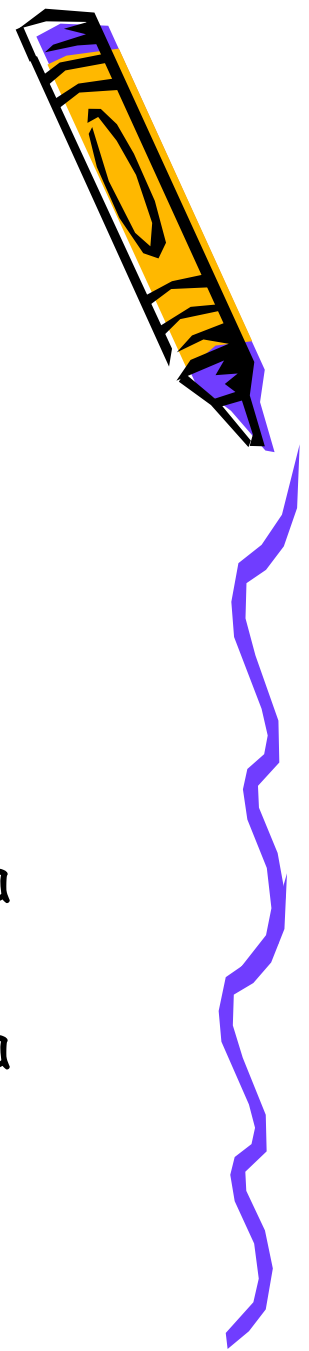
$$\gg E(Z_n) = \theta$$

- Innymi słowy, przy wielokrotnym losowaniu próby średnia z wartości przyjmowanych przez estymator nieobciążony jest równa wartości szacowanego parametru.

Obciążoność oznacza, że oszacowania dostarczone przez taki estymator są obciążone błędem systematycznym



Estymator obciążony



- **Obciążoność** oznacza, że oszacowania dostarczone przez taki estymator są obarczone błędem systematycznym.

Różnica:

$$B_n = E(Z_n) - \theta$$

nazywa się obciążeniem estymatora.

Jeżeli $B_n > 0$ to estymator Z_n daje przeciętnie za wysokie oceny parametru θ ;

Jeżeli $B_n < 0$ to estymator Z_n daje przeciętnie za niskie oceny parametru θ .





- Jeśli spełniony jest warunek:

$$\lim_{n \rightarrow \infty} B_n = 0$$

- co jest równoważne warunkowi:

$$\lim_{n \rightarrow \infty} E(Z_n) = 0$$

to estymator taki nazywa się **estymatorem asymptotycznie nieobciążonym**.

Uwaga! Postulat nieobciążoności estymatora parametru oznacza praktyczne żądanie, aby rozkład estymatora był scentrowany wokół prawdziwej wartości parametru, a więc by jego odchylenia od parametru miały charakter losowy.



Estymator - zgodność

- Estymator Z parametru θ nazywa się estymatorem zgodnym, jeśli wraz ze wzrostem liczebności próbki jest on stochastycznie zbieżny do wartości estymowanego parametru θ , tzn. jeśli jest spełniony warunek:

$$\lim_{n \rightarrow \infty} P|Z_n - \theta| < \sigma = 1$$

gdzie σ jest dowolnie małą liczbą dodatnią.

- Zgodność estymatora Z oznacza, że wraz ze wzrostem liczebności próbki n , prawdopodobieństwo dowolnie małej różnicy między wartością estymatora Z a estymowanym parametrem θ dąży do 1.
- Wynika stąd, że warto powiększyć próbkę, ponieważ przy wzroście n rośnie prawdopodobieństwo tego, że wartość estymatora parametru Z będzie się niewiele różnić od prawdziwej wartości estymowanego parametru θ , powodując tym samym mały błąd estymacji.



Estymator - efektywność



- Estymator nieobciążony, który ma najmniejszą wariancję, nazywa się estymatorem najefektywniejszym.
- Przy estymacji punktowej sytuacja jest tym korzystniejsza, im wartość Z_n oscyluje bliżej σ , a więc im wariancja jest mniejsza.
- Wyrażenie:

$$\sigma^2(Z_n) = E(Z_n - E(Z_n))^2$$

jest wariancją estymatora Z_n .

- Uwaga! Estymator jest tym efektywniejszy, im mniejsza jest jego wariancja i odchylenie standardowe.





Ze względu na formę wyniku estymacji wyróżniamy:

- **Estymacja punktowa** -gdy szacujemy liczbową wartość określonego parametru rozkładu cechy w całej populacji
- **Estymacja przedziałowa** -gdy wyznaczamy granice przedziału liczbowego, w których, z określonym prawdopodobieństwem, mieści się prawdziwa wartość szacowanego parametru.



Wprowadzenie do problematyki estymacji parametrów modeli ekonometrycznych



- Problemy estymacji należą do trudnych zagadnień;
- Nie ma jednej uniwersalnej metody estymacji;
- Strona rachunkowa metod estymacji jest zawiła, więc dla większych modeli (z wieloma zmiennymi objaśniającymi) estymacja wymaga wykorzystania komputerów;
- Estymacja jest jednym z najważniejszych działów statystyki matematycznej
- Estymacja jest o tyle ważna, że **od estymacji zależy jakość modelu ekonometrycznego i jego praktyczna użyteczność**



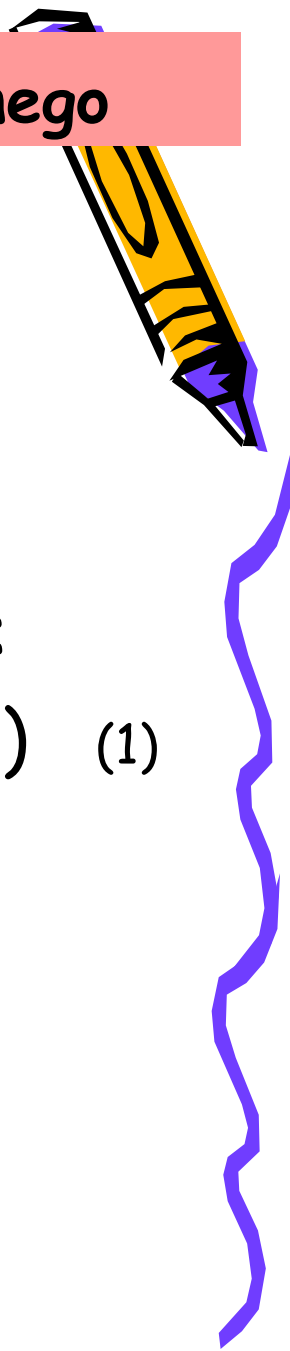
Estymacja parametrów modelu ekonometrycznego

- **Przedmiotem estymacji** w badaniu ekonometrycznym **są parametry** sformułowanych wcześniej modeli ekonometrycznych
- Ogólny zapis modelu ekonometrycznego:

$$Y = f(X_1, X_2, \dots, X_k, a_1, a_2, \dots, a_k, \xi) \quad (1)$$

gdzie:

- Y - zmienna objaśniana;
- X_1, X_2, \dots, X_k - zmienne objaśniające
- a_1, a_2, \dots, a_k - parametry strukturalne modelu
- ξ - składnik losowy



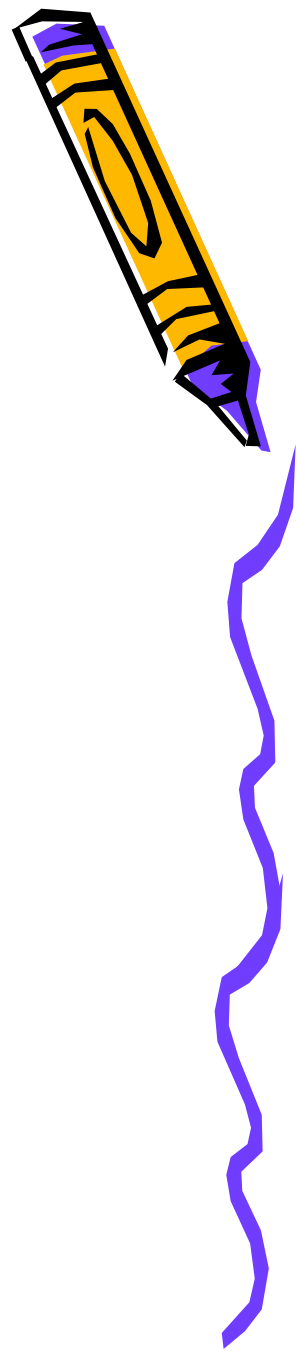
Podstawą przeprowadzenia estymacji są dane statystyczne będące zaobserwowanymi wartościami zmiennych Y, X_1, X_2, \dots, X_k . Przedstawia się je zwykle w postaci:

- $(n \times 1)$ wymiarowego wektora Y

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} \quad (2)$$

- $(n \times k)$ wymiarowej macierzy obserwacji zmiennych objaśniających

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix} \quad (3)$$



Proces estymacji polega na obliczaniu ocen parametrów strukturalnych i stochastycznych modelu.

I. Parametry strukturalne (wielkości od których zależy wartość funkcji zmiennych objaśniających) są przedstawione w postaci:

- $((k+1) \times 1)$ wymiarowego wektora parametrów strukturalnych

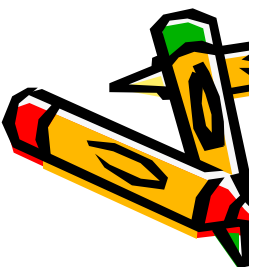
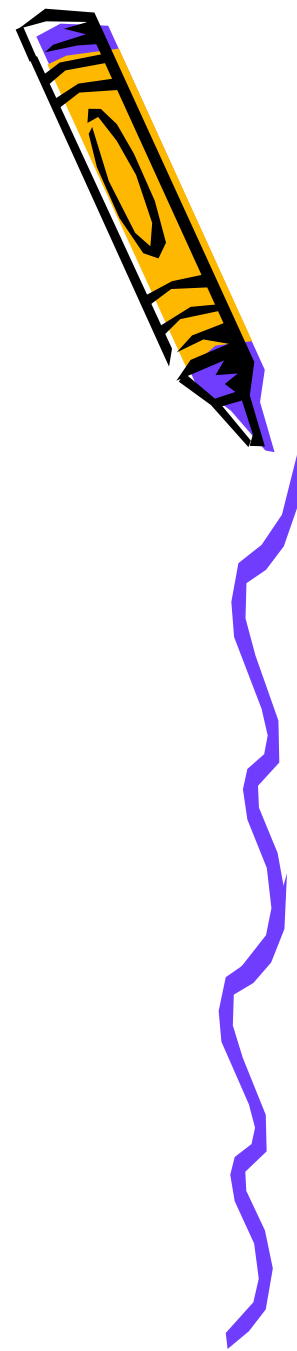
$$\alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \alpha_k \end{bmatrix} \quad (4)$$

Estymacja sprowadza się do znalezienia wektora ocen

$$a = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_k \end{bmatrix} \quad (5)$$

postać estymatora a jest różnie określana dla poszczególnych metod estymacji, ogólnie można zapisać, że wektor a jest funkcją:

$$a = f(X, y) \quad (6)$$



II. Drugą grupą parametrów modelu ekonometrycznego podlegającą estymacji są parametry struktury stochastycznej.

Informują one o rzędzie dokładności estymacji.

Można je podzielić na **2 grupy**:

- 1. parametry rozkładu składnika losowego,**
- 2. parametry estymatorów parametrów strukturalnych.**

Ad.1. **Składnik losowy modelu** jest przedstawiony w postaci:

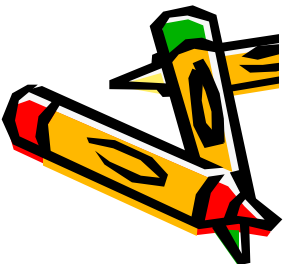
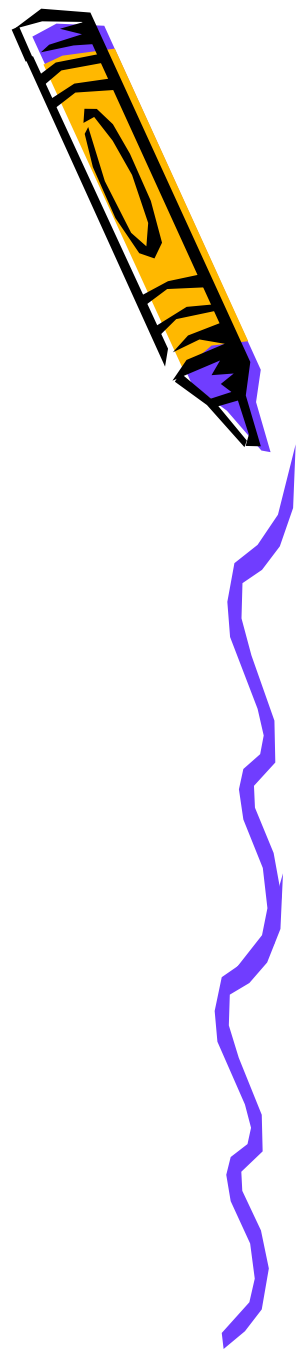
- (n x 1) wymiarowego wektora

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \cdot \\ \xi_n \end{bmatrix} \quad (7)$$

Oszacowania składników losowych noszą nazwę **reszt modelu** i przedstawiamy je w postaci:

- (n x 1) wymiarowego wektora

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_n \end{bmatrix} \quad (8)$$



Na podstawie wektora reszt przeprowadza się estymację takich parametrów rozkładu składnika losowego jak:

- wartości oczekiwanej składnika losowego $E(\xi)$, której estymatorem jest średnia reszt \bar{e}

- wariancji składnika losowego $\sigma^2(\xi)$, której estymatorem jest wariancja resztowa $S^2(e)$,

- odchylenia standardowego składnika losowego $\sigma(\epsilon)$, którego estymatorem jest odchylenie standardowe reszt $S(e)$.

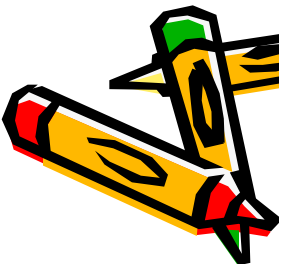
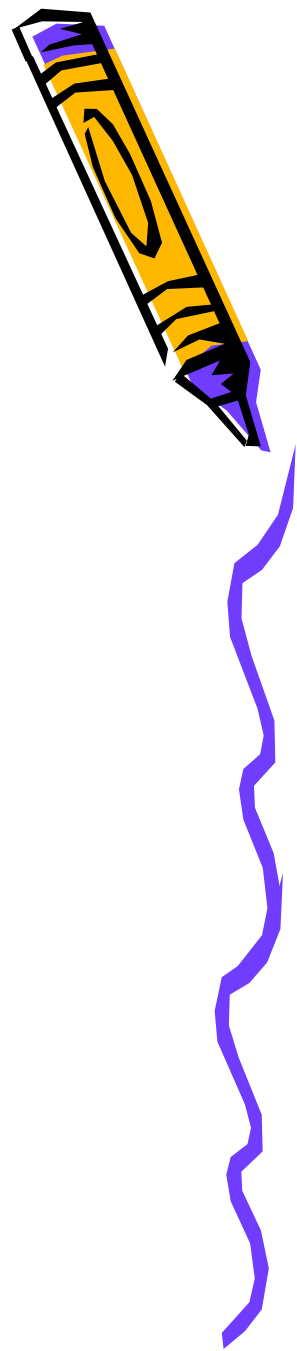
Innymi parametrami rozkładu składnika losowego są kowariancje składników losowych $\sigma_{ij}^r(\xi) = E(\xi_i, \xi_j)$ t.l = 1,2,...,n przedstawione w postaci:

- (n x n) wymiarowej macierzy

$$D^2(\xi) = \begin{bmatrix} \sigma_{11}(\xi) & \sigma_{12}(\xi) & \dots & \sigma_{1n}(\xi) \\ \sigma_{21}(\xi) & \sigma_{22}(\xi) & \dots & \sigma_{2n}(\xi) \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{n1}(\xi) & \sigma_{n2}(\xi) & \dots & \sigma_{nn}(\xi) \end{bmatrix} \quad (9)$$

których estymatorami są elementy macierzy:

$$S^2(e) = \begin{bmatrix} s_{11}(e) & s_{12}(e) & \dots & s_{1n}(e) \\ s_{21}(e) & s_{22}(e) & \dots & s_{2n}(e) \\ \cdot & \cdot & \cdot & \cdot \\ s_{n1}(e) & s_{n2}(e) & \dots & s_{nn}(e) \end{bmatrix} \quad (10)$$



Ad.2. Parametry estymatorów parametrów strukturalnych są przedstawione

w postaci:

- $((k+1) \times (k+1))$ wymiarowej macierzy wariancji i kowariancji estymatorów parametrów strukturalnych

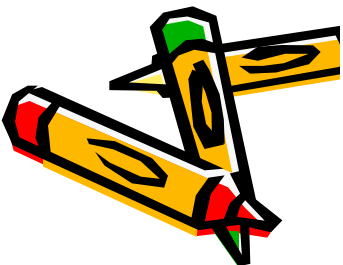
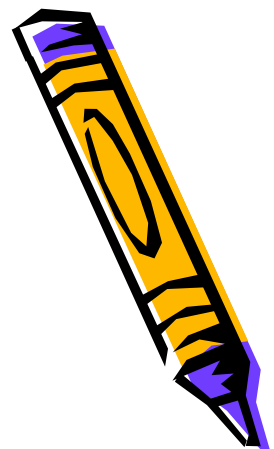
$$D^2(\mathbf{a}) = \begin{bmatrix} \sigma_{00}(\mathbf{a}) & \sigma_{01}(\mathbf{a}) & \dots & \sigma_{0k}(\mathbf{a}) \\ \sigma_{10}(\mathbf{a}) & \sigma_{11}(\mathbf{a}) & \dots & \sigma_{1k}(\mathbf{a}) \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{k0}(\mathbf{a}) & \sigma_{k1}(\mathbf{a}) & \dots & \sigma_{kk}(\mathbf{a}) \end{bmatrix} \quad (11)$$

w której element na głównej przekątnej $\sigma_{ii}(\mathbf{a}) = \sigma_i^2(\mathbf{a})$ oznacza wariancję estymatora a_i parametru strukturalnego α_i , a element poza główną przekątną $\sigma_{ij}(\mathbf{a})$ gdzie $i \neq j$ oznacza kowariancję estymatorów a_i oraz a_j parametrów strukturalnych α_i i α_j .

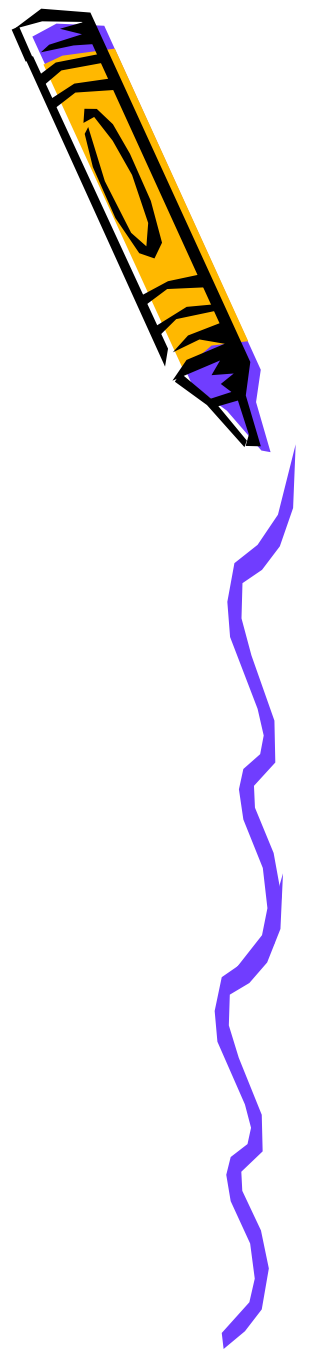
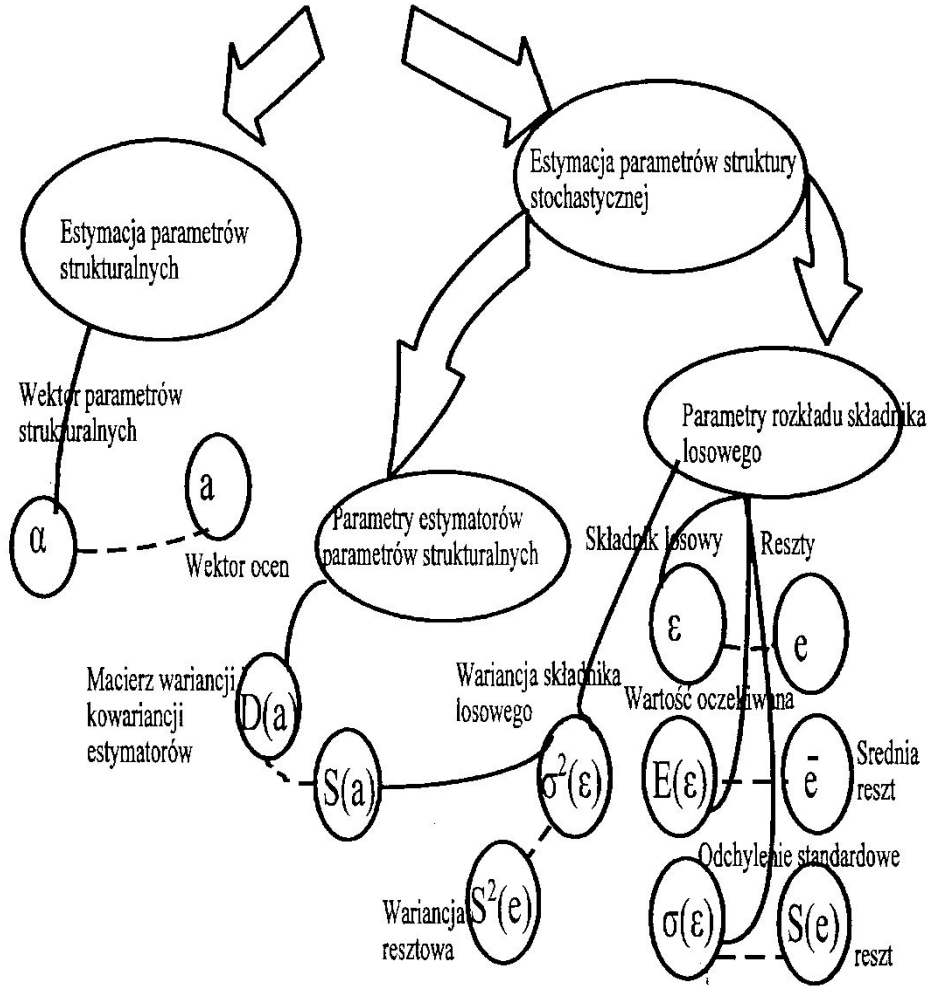
Estymator macierzy $D^2(\mathbf{a})$ będziemy oznaczać jako:

$$\hat{S}(\mathbf{a}) = \begin{bmatrix} s_{00}(\mathbf{a}) & s_{01}(\mathbf{a}) & \dots & s_{0k}(\mathbf{a}) \\ s_{10}(\mathbf{a}) & s_{11}(\mathbf{a}) & \dots & s_{1k}(\mathbf{a}) \\ \cdot & \cdot & \cdot & \cdot \\ s_{k0}(\mathbf{a}) & s_{k1}(\mathbf{a}) & \dots & s_{kk}(\mathbf{a}) \end{bmatrix} \quad (12)$$

W praktyce korzysta się jedynie z elementów głównej przekątnej macierzy (12) tzn. z elementów $s_{ii}(\mathbf{a}) = s_i^2(\mathbf{a})$, będących estymatorami wariancji estymatorów parametrów strukturalnych α_i ($i=0,1,2,..k$). Pierwiastki tych elementów $s_i(\mathbf{a})$ noszą nazwę **średnich błędów szacunku (estymacji) parametrów strukturalnych**.



Estymacja parametrów modelu ekonometrycznego

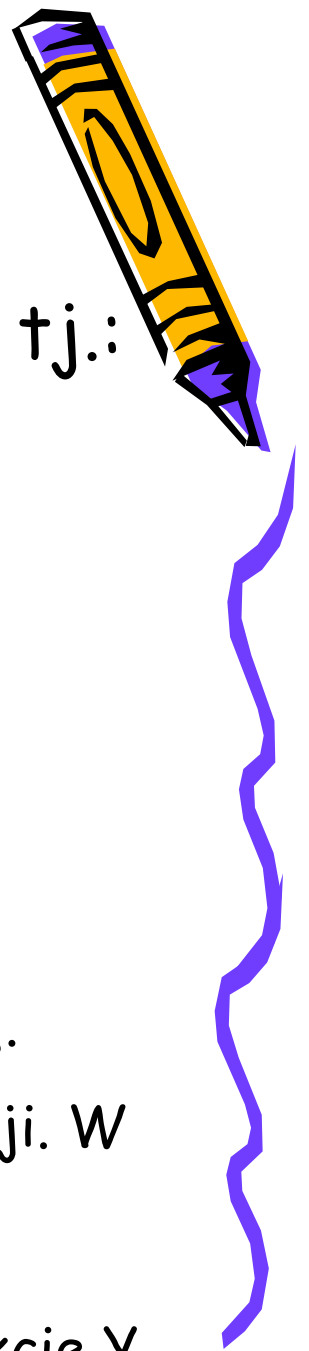


Estymacja parametrów modelu ekonometrycznego

- Z reguły estymatory uzyskuje się w wyniku zastosowania procedury numerycznej zwanej **metodą najmniejszych kwadratów**.
- Estymatory mają wówczas pożądane własności, o ile spełnione są pewne istotne założenia.
- Założenia te dotyczą głównie:
 - - specyfikacji modelu i
 - - własności składnika losowego.



Założenia: model i dane



- Założenie 1

Model jest liniowy względem parametrów tj.:

$$Y_t = a_0 + a_1 X_{1t} + a_2 X_{2t} + \dots + a_k X_{kt} + \xi_t$$

gdzie $t = 1, 2, \dots, n$

- Założenie 2

- Zmienne objaśniające są nielosowe

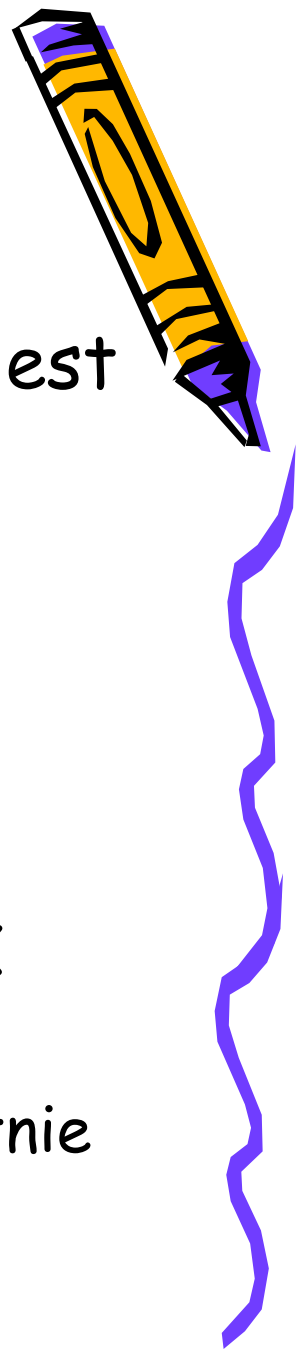
Zmienna Y jest losowa, bowiem jest funkcją losowego ξ .

Przyjmijmy Y - koszt produkcji, X - wartość produkcji. W modelu mogą zmieniać się rolami.



Uwaga! Niekonsekwencja klasycznej ekonometrii - w efekcie Y

Założenia: model i dane



- Założenie 3
- Liczba obserwacji n (wielkość próby n) jest większa od liczby parametrów do oszacowania:

$$n > k+1$$

- Parametrów jest $k+1$:
wyraz wolny + k parametrów przy zmiennych X
- W praktyce żądamy aby n była liczbą kilkakrotnie większą od $k+1$ (np. dwukrotnie)



Założenia: model i dane



- Założenie 4
- **Żadna ze zmiennych nie jest kombinacją liniową innych zmiennych objaśniających** (włączając w ten zbiór także „sztuczną” zmienną $X_0 = 1$, która „stoi” przy wyrazie wolnym modelu)
- Jest to założenie o braku współliniowości.
- Nie istnieje zależność liniowa między wartościami z próby dla jakichkolwiek 2-óch, lub większej ilości zmiennych objaśniających.
- Chodzi to, aby żadna ze zmiennych nie wносиła do modelu tych informacji które już są wniesione przez inne zmienne.



Założenia: składnik losowy modelu



- Założenie 5
- Składnik losowy ξ jest zmienną losową
- **Składnik losowy ma wartość oczekiwaną równa zero** dla wszystkich $i=1,2,\dots,n$:

$$E(\xi_i) = 0$$

- Oznacza to, że czynniki nie uwzględnione w modelu nie oddziałują w systematyczny sposób na średnią wartość zmiennej Y :
 - wpływy dodatnie (+) i wpływy ujemne(-) „znoszą się” i w sumie efekt jest zerowy.



Założenia: składnik losowy modelu



- Założenie 6
- Składnik losowy ξ jest zmienną losową
- **Wariancja zmiennej losowej ξ_i jest taka sama dla wszystkich obserwacji**

$$D^2(\xi_i) = \sigma^2$$

dla $i=1,2,\dots,n$:

- Przyjmujemy, że zmienne losowe mają jednakową dyspersję. Oznacza to, że wpływy na Y czynników nie ujętych w modelu mają takie same rozproszenie (niezależnie od numeru obserwacji)
- Założenie o jednakowych wariancjach nosi nazwę założenia o **homoscedastyczności**.

Jego przeciwieństwem jest założenie o **heteroscedastyczności** (nierówna dyspersja)



Założenia: składnik losowy modelu



- Założenie 7
- Składnik losowy ξ jest zmienną losową
- **Zmienne losowej ξ_i są nieskorelowane, czyli nie występuje autokorelacja składników losowych):**

$$\text{cov}(\xi_i, \xi_j) = \sigma_{i,j}(\xi) = 0 \quad \text{dla } i \neq j$$

$i=1,2,\dots,n; j=1,2,\dots,n:$

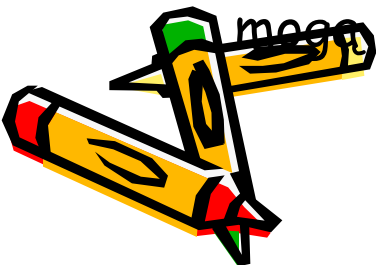
- Oznacza to, że wpływy na Y czynników nie ujętych w modelu są nieskorelowane pomiędzy różnymi obserwacjami
- Jest to założenie często niespełnione w modelach trendu



Założenia: składnik losowy modelu



- Założenie 8
- **Każdy ze składników losowych ξ_i ma rozkład normalny.**
- Biorąc pod uwagę założenia 4 i 5 oznacza to, że ξ_i **ma rozkład $N(0, \sigma^2)$ dla $i = 1, 2, \dots, n$**
- Niekiedy założenia 1-7 uzupełnia się o założenie 8 a model określa się wówczas mianem **klasycznego modelu normalnej regresji liniowej**
- Założenie 8 ułatwia konstruowanie hipotez statystycznych służących weryfikacji modelu
- Założenia dotyczące składnika losowego są nieznane, sprawdzone mogą być dopiero po oszacowaniu parametrów modelu





Model jest liniowy względem parametrów tj.:

$$Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \dots + \alpha_k X_{kt} + \xi_t$$

gdzie $t = 1, 2, \dots, n$

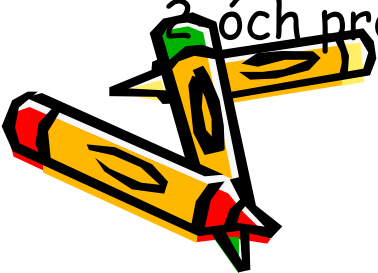
Wielkości parametrów α_j ($j = 0, 1, 2, \dots, k$) w modelu liniowym są niewiadomymi'

Po to by uzyskać wiedzę na temat wielkości parametrów modelu musimy posłużyć się danymi empirycznymi Y i X_k ($k = 1, 2, \dots, n$).

Na podstawie danych szacujemy nieznane parametry α_j na podstawie reakcji zmiennej zależnej na zmiany wielkości zmiennych niezależnych zaobserwowanych w próbie.

To co uzyskujemy na podstawie danych jest jedynie szacunkiem i będzie mniej lub bardziej dokładnym przybliżeniem prawdziwych wielkości parametrów α_j .

W rezultacie oszacowania parametrów uzyskane na podstawie dwóch prób z reguły będą różne.



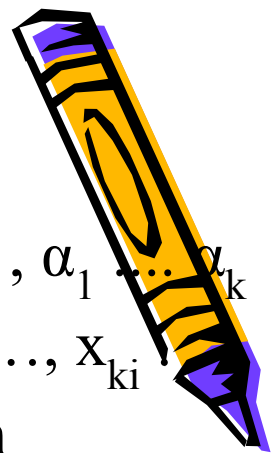


- **Wniosek:**
- Oszacowania nielosowych parametrów są losowe.
- Będąc jedynie niedokładnym przybliżeniem prawdziwych wielkości parametrów mogą różnić się w zależności od wylosowanej próby.
- Niedokładności w oszacowaniach wielkości parametrów wynikają z zaburzeń losowych (ξ), które uniemożliwiają dokładne zmierzenie parametrów modelu.



Wartości dopasowane i reszty

- Znajdowanie estymatorów (oszacowań) parametrów $\alpha_0, \alpha_1, \dots, \alpha_k$ ($j=0,1,2,\dots,k$) określamy mianem regresji liniowej y_i na x_{1i}, \dots, x_{ki}
- Zgodnie z przyjętą konwencją *oszacowania* nieznanymi parametrów $\alpha_0, \alpha_1, \dots, \alpha_k$ uzyskanych za pomocą MNK oznaczamy zwykle $\alpha_0, \alpha_1, \dots, \alpha_k$.
- Przewidywane na podstawie oszacowanego modelu wartości zmiennej zależnej Y **nazywamy wartością teoretyczną** (dopasowaną):
 - $$\hat{y} = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k$$
- Wartości dopasowane różnią się od rzeczywistych wartości Y , ponieważ w modelu oszacowanym zamiast prawdziwych (nieznanych) wartości parametrów $\alpha_0, \alpha_1, \dots, \alpha_k$ używamy ich oszacowań $\alpha_0, \alpha_1, \dots, \alpha_k$ i pomijamy błąd losowy



Wartości dopasowane i reszty

- Reszty definiujemy jako różnicę między wartością zaobserwowaną zmiennej zależnej (objaśnianej) Y , a wartością dopasowaną tej zmiennej:

$$e = Y - (a_0 + a_1 X_1 + a_2 X_2 + \dots + a_k X_k)$$

$$e = Y - a_0 - a_1 X_1 - a_2 X_2 - \dots - a_k X_k$$

$$e = Y - \hat{Y}$$

Relację między resztami, obserwacjami i oszacowaniami parametrów można zapisać w sposób następujący:

$$Y = \hat{Y} + e \\ = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_k X_k + e$$

Taki zapis pokazuje „pokrewieństwo” między $\alpha_0, \alpha_1 \dots \alpha_k$ i $a_0, a_1 \dots a_k$ oraz między ξ i e .

Tak jak $a_0, a_1 \dots a_k$ są oszacowaniami $\alpha_0, \alpha_1 \dots \alpha_k$ tak reszty e stanowią oszacowania składnika losowego ξ .

Uwaga! Reszty e nie są równe ξ



Wartości dopasowane i reszty

- Model jest tym lepiej dopasowany, im mniejsza jest odległość wartości teoretycznych od wartości obserwowanych

$$e = Y - \hat{Y}$$

- Najlepiej dopasowanym jest ten model, w którym reszty są - co do wartości bezwzględnych – najmniejsze.
- Estymator MNK znajdujemy, szukając takich $a_0, a_1 \dots a_k$ dla których łączna odległość $e = Y - \hat{Y}$ jest najmniejsza



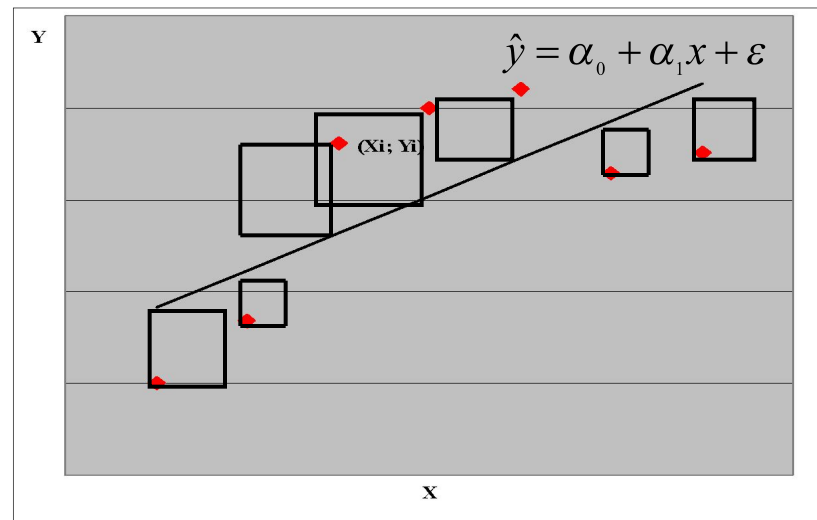
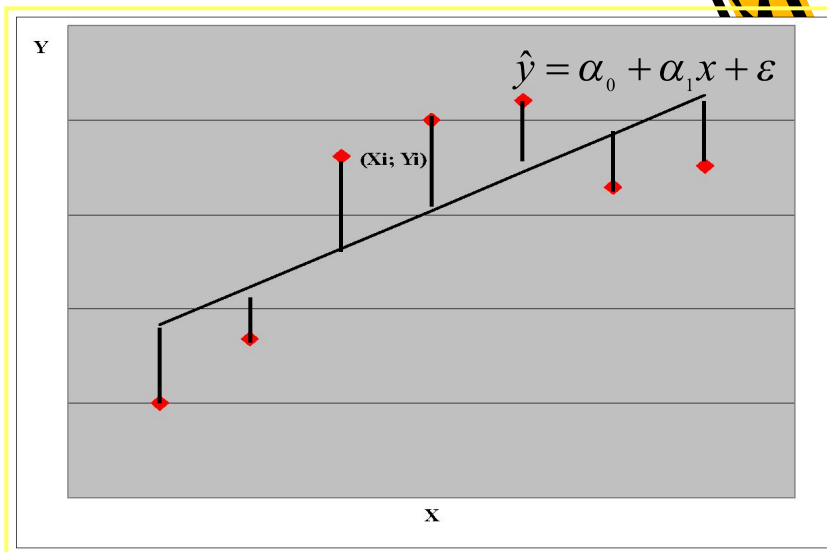
Reasumując:

- Do poszukiwania najlepiej dopasowanej prostej stosuje się kryterium minimalizacji sumy kwadratów odchyłek.
- Metoda wyznaczania parametrów prostej oparta na tym kryterium nosi nazwę **metody najmniejszych kwadratów (MNK)**.
- Stosując MNK wyznacza się na podstawie danych (x_i, y_i) , $i=1,2,\dots, n$, parametry α_0 i α_1 prostej tak, by suma kwadratów odchyłek y_i od $\alpha_0 + \alpha_1 x_i$ była najmniejsza:

$$S = \psi = \sum_{i=1}^n (y_i - \hat{y}_i)^2 =$$

$$\sum_{i=1}^n (y_i - a_0 - a_1 x_i)^2 \rightarrow \min$$

Rysunek 1 i 2. Ilustracja metody najmniejszych kwadratów



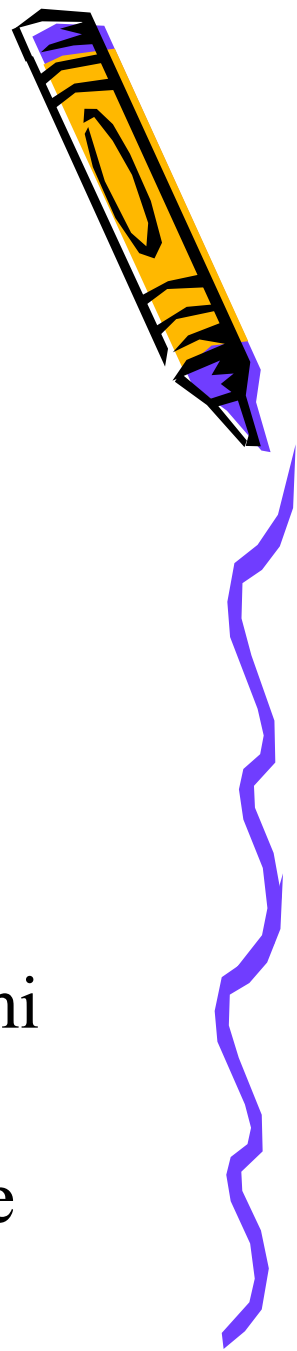
Mamy model liniowy z jedną zmienną objaśniającą

$$Y = \alpha_0 + \alpha_1 X_1 + \xi$$

Wielkości parametrów α_i ($i=0,1$) w modelu liniowym są niewiadomymi.

Po to, by uzyskać wiedzę na temat wielkości parametrów modelu musimy posłużyć się danymi empirycznymi.

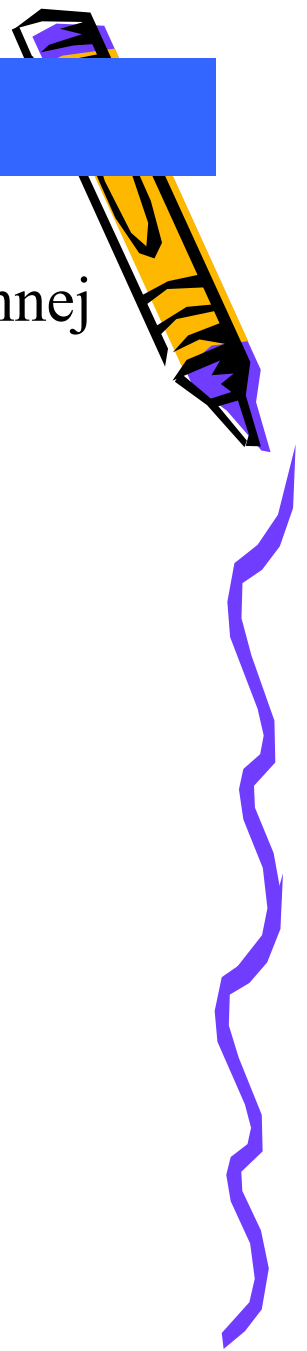
Parametry α_i ($i=0,1$) szacujemy na podstawie danych:



Estymacja

- Y jest wektorem zaobserwowanych wartości zmiennej objaśnianej:

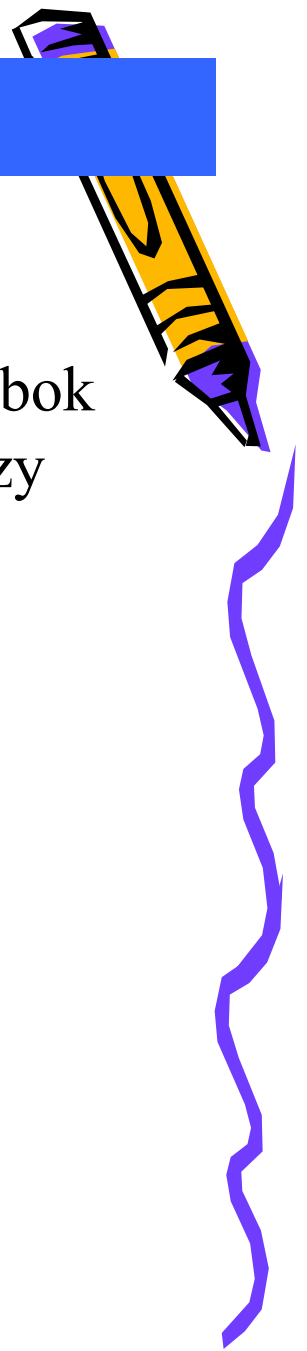
$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{bmatrix}$$

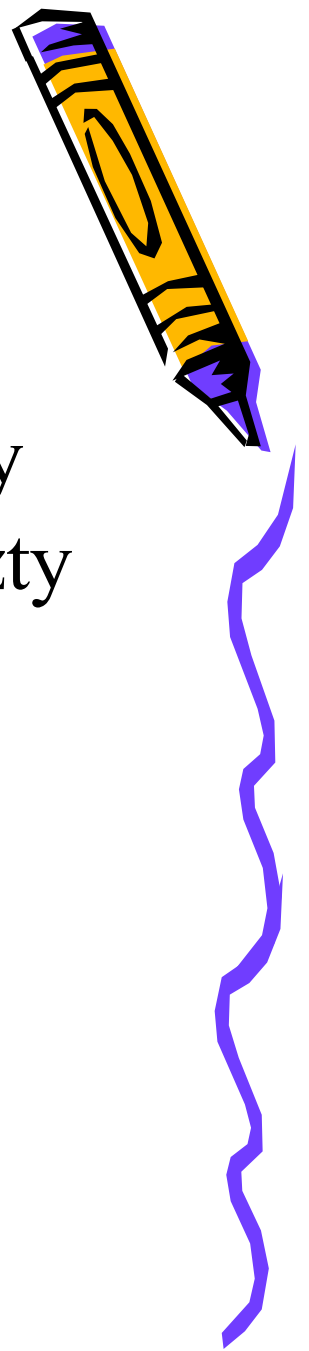


Estymacja

- X jest macierzą zaobserwowanych wartości zmiennych objaśniających, przy czym przyjmuje się, że w modelu obok wymienionych zmiennych występuje zmienna $x_{01}=1$ (przy parametrze α_0), a więc:

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & x_{1n} \end{bmatrix}$$





- **Funkcja kryterium** (minimalizujemy sumę kwadratów reszt e , przy czym reszty to odchylenia wartości teoretycznych od wartości empirycznych y) w zapisie skalarnym ma postać:

$$\psi = \sum_{t=1}^{\hat{n}} (y - \hat{y})^2 = \sum_{t=1}^n (y - \hat{y})^2 = \min$$



Estymacja

- Wektor ocen a parametrów strukturalnych α otrzymujemy obliczając pochodną funkcji ψ względem wektora a i przyrównując ją do zera.
- Wzór na wektor ocen parametrów strukturalnych przybiera ostatecznie postać:

$$a = (X^T X)^{-1} X^T y$$

- Podstawiając do wzoru:

$$X^T X = \begin{bmatrix} n & \sum x_{1t} \\ \sum x_{1t} & \sum x_{1t}^2 \end{bmatrix} \quad \text{oraz} \quad X^T y = \begin{bmatrix} \sum y_t \\ \sum x_{1t} y_t \end{bmatrix}$$



Estymacja

- otrzymamy wektor ocen parametrów strukturalnych funkcji liniowej:

$$a = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

