

## 5. Testy diagnostyczne

### 5.1. Testowanie prawidłowości formy funkcyjnej modelu

- **Test Reset**

Często zdarza się, że forma funkcjonalna modelu, dla której prowadzimy estymację jest nieprawidłowa. Jak już pokazywaliśmy na rysunku 4 błąd taki może doprowadzić do pojawienia się rzekomej autokorelacji w modelu. Podobnie może on także spowodować pojawienie się heteroskedastyczności. Z tego powodu testy *DW* i *White*'a mogą być traktowane jako testy na poprawność formy funkcyjnej modelu.

Do testowania prawidłowości formy funkcyjnej możemy posłużyć się także testem *RESET* (*Regression Specification Error Test*). Szeroką klasę funkcji nieliniowych można aproksymować za pomocą wielomianów. Jeśli dodanie do zbioru zmiennych objaśniających potęg tych zmiennych istotnie poprawia dopasowanie regresji, to jest to wyraźną wskazówką, że model powinien być w rzeczywistości nieliniowy. Dodawanie potęg wszystkich zmiennych objaśniających wiązałoby się ze zbyt dużą utratą stopni swobody i z tego powodu przyjęło się używać zamiast potęg zmiennych objaśniających potęg wartości teoretycznych zmiennej  $y_i$ . Test *RESET* polega na testowaniu hipotezy o łącznej nieistotności tych potęg. Hipoteza zerową o poprawności formy funkcyjnej można więc sformułować jako  $H_0 : \alpha = \mathbf{0}$  w modelu

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + \mathbf{z}_i\boldsymbol{\alpha} + u_i$$

gdzie  $\mathbf{z}_i = [ \hat{y}_i^2 \quad \dots \quad \hat{y}_i^{p+1} ]$  a wartości teoretyczne  $\hat{y}_i$  są wyliczone na podstawie pierwotnej regresji  $y_i = \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i$ . Jako statystyki testowej można użyć statystyki  $nR^2$  z twierdzenia 3.11. Statystyka ta dla tego testu będzie miała asymptotyczny rozkład  $\chi_p^2$ . Poprawne jest także zastosowanie statystyki  $F$  dla testu na łączną nieistotność zmiennych  $\hat{y}_i^2, \dots, \hat{y}_i^{p+1}$ . Statystyka ta będzie miała asymptotycznie prawidłowy rozkład.

## 5.2. Testowanie normalności składników losowych

### • Test Jarque-Bera

Przy wyprowadzaniu rozkładów testów dla prób skończonych posługiwaliśmy się założeniem, że  $\varepsilon \sim N(0, \sigma^2 \mathbf{I})$ . Niekiedy jednak błędy losowe nie mają rozkładu normalnego. W takim przypadku rozkłady statystyk testowych w próbach skończonych będą inne od tych wyprowadzonych przy założeniu normalności. Asymptotyczne rozkłady statystyk będą w dalszym ciągu poprawne. Hipotezę zerową  $H_0 : \varepsilon \sim N(0, \sigma^2 \mathbf{I})$  można przetestować używając do tego statystyki Jarque-Bera postaci

$$LM = n \left[ \frac{\hat{\theta}_1^2}{6} + \frac{(\hat{\theta}_2 - 3)^2}{24} \right] \xrightarrow{D} \chi_2^2$$

gdzie  $\hat{\theta}_1 = \frac{\sum_{i=1}^n e_i^3/n}{\hat{\sigma}^3}$  jest wyestymowaną skośnością rozkładu reszt a  $\hat{\theta}_2 = \frac{\sum_{i=1}^n e_i^4/n}{\hat{\sigma}^4}$  a jego wyestymowaną kurtozą, przy czym  $\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}}$ . Idea testu opiera się na tym, że dla rozkładu normalnego skośność  $\theta_1 = 0$  a kurtozą  $\theta_2 = 3$ , co oznacza, że dla prawdziwej  $H_0$  wartość statystyki powinna być w przybliżeniu równa zeru. Ponieważ można pokazać, że  $\sqrt{n}\hat{\theta}_1 \xrightarrow{D} N(0, 6)$  a  $\sqrt{n}(\hat{\theta}_2 - 3) \xrightarrow{D} N(0, 24)$  i ich rozkłady są niezależne więc istotnie powyższa statystyka będzie dążyć do rozkładu  $\chi_2^2$ .

Nawet kiedy błędy losowe nie mają rozkładu normalnego założenia twierdzenia Gaussa-Markowa są spełnione i estymator  $\mathbf{b}_{MNK}$  jest najlepszym *liniowym i nieobciążonym* estymatorem  $\beta$  a estymator  $MNK$  macierzy wariancji-kowariancji jest dalej estymatorem nieobciążonym. W takim przypadku istnieją jednak estymatory nieliniowe, które są bardziej efektywne od estymatora  $MNK$ . Jednak dla przypadku, gdy rozkład  $\varepsilon_t$  nie jest rozkładem normalnym nie obowiązują wyprowadzone wcześniej twierdzenia na temat rozkładów estymatorów i statystyk w małych próbach.

**Literatura:** Steward (1991) str. 152-170, Green (1997) str. 549-555, Goldberger (1972) str. 312-314, Theil (1979) str. 208-212.

### 5.3. Test prognoz

W praktyce często pojawi się pytanie, czy model prawdziwy dla pewnego zbioru obserwacji jest prawdziwy dla innego zbioru obserwacji dotyczących tego samego zjawiska. Typową sytuacją jest przypadek kiedy podejrzewamy, że w momencie  $t^*$  parametry modelu uległy zmianie (np. na skutek zmiany systemu polityczno-gospodarczego). Sytuację tę ilustruje rysunek 1. Pionowymi liniami zaznaczono  $\pm 2$  teoretyczne błędy prognozy. Z rysunku widać, że prognozy odbiegają znacząco od realizacji zmiennej. Podzielmy próbę na dwie części: okres estymacji i okres prognozy:

$$\underbrace{1, \dots, n}_{\text{okres estymacji}} \quad \underbrace{n+1, \dots, n+g}_{\text{okres prognozy}}$$

Drugą możliwą zastosowanie testu prognoz dotyczy sytuacji kiedy interesuje nas, czy pewna mała podgrupa populacji (np. najwyższy decyl dochodowy) może być opisana tym samym modelem co cała populacja. Sytuację to ilustruje rysunek 2.

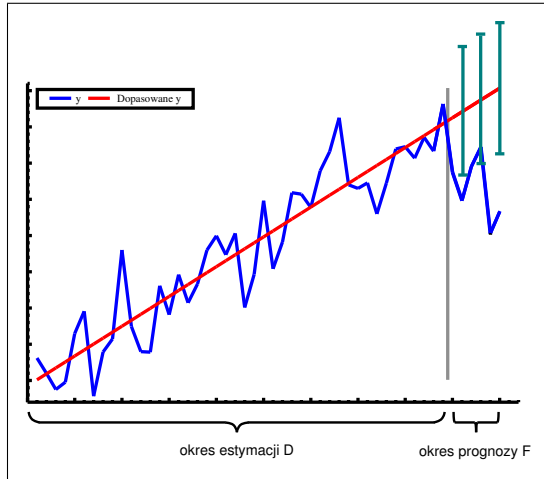
Naszą hipotezą zerową jest, że w okresie estymacji model ma te same parametry co w okresie prognozy. Formułujemy tę hipotezę w sposób następujący:

$$H_0 : \mathbf{y}_f = \mathbf{X}_f \boldsymbol{\beta} + \boldsymbol{\varepsilon}_f \quad (5.1)$$

W przeciwieństwie do poprzednio omawianych hipotez liniowych, ta hipoteza ma charakter stochastyczny, ponieważ opisuje zależność między zmiennymi losowymi. Wprowadźmy oznaczenia na prognozę

$$\hat{\mathbf{y}}_f = \mathbf{X}_f \mathbf{b}$$

i błąd prognozy



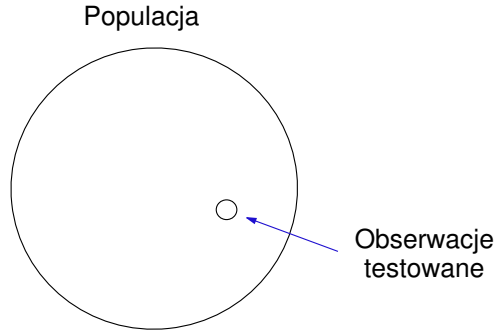
Rysunek 1: Model trendu liniowego generuje złe prognozy

$$\mathbf{f} = \mathbf{y}_f - \hat{\mathbf{y}}_f,$$

gdzie  $\mathbf{b}$  jest estymatorem *MNK* parametru  $\beta$  otrzymanym dla okresu estymacji

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$$

a  $\mathbf{y}_f$ ,  $\mathbf{X}_f$ ,  $\varepsilon_f$  są to wartości zmiennych losowych w okresie prognozy a  $\mathbf{y}$ ,  $\mathbf{X}$ ,  $\varepsilon$  oznacza wartości tych samych zmiennych dla okresu estymacji. Zakładamy, że  $\varepsilon_f$  ma te same własności co  $\varepsilon$  a więc



Rysunek 2: Mała podgrupa populacji generalnej

$\varepsilon_f \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_g)$  i  $\text{Cov}(\varepsilon, \varepsilon_f) = \mathbf{0}$ , gdzie  $g$  jest ilością obserwacji w okresie prognozy. Zauważmy teraz, że

$$\mathbf{f} = \mathbf{X}_f \boldsymbol{\beta} + \varepsilon_f - \mathbf{X}_f \mathbf{b} = \mathbf{X}_f (\boldsymbol{\beta} - \mathbf{b}) + \varepsilon_f$$

Ostatni wzór w tym równaniu mówi nam, że błąd prognozy składa się z dwóch elementów: błędu estymacji i błędu losowego. Można łatwo pokazać, że te dwa źródła błędów prognozy są niezależne. Estymator *MNK* ma postać

$$\mathbf{b} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\varepsilon}$$

$$\mathbf{X}_f (\boldsymbol{\beta} - \mathbf{b}) = -\mathbf{X}_f (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \boldsymbol{\varepsilon}$$

i

$$\text{Cov} [\mathbf{X}_f (\boldsymbol{\beta} - \mathbf{b}), \boldsymbol{\varepsilon}_f] = -\mathbf{X}_f (\mathbf{X}' \mathbf{X}) \mathbf{X}' \text{Cov} (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_f) = \mathbf{0}$$

gdź

$$\text{Cov} (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_f) = \mathbf{0}$$

Policzmy teraz wartość oczekiwaną i wariancję prognozy

$$E(\mathbf{f}) = E(\mathbf{X}_f \boldsymbol{\beta} - \mathbf{X}_f \mathbf{b} - \boldsymbol{\varepsilon}_f) = \mathbf{X}_f \boldsymbol{\beta} - \mathbf{X}_f E(\mathbf{b}) - E(\boldsymbol{\varepsilon}_f) = \mathbf{0}$$

Ponieważ  $\mathbf{X}_f (\boldsymbol{\beta} - \mathbf{b})$  i  $\boldsymbol{\varepsilon}_f$  są nieskorelowane, więc

$$\text{Var}(\mathbf{f}) = [\mathbf{X}_f (\boldsymbol{\beta} - \mathbf{b})] + \text{Var}(\boldsymbol{\varepsilon}_f) = \text{Var}(\mathbf{X}_f \mathbf{b}) + \text{Var}(\boldsymbol{\varepsilon}_f)$$

a ponieważ  $\mathbf{X}_f \boldsymbol{\beta}$  jest nielosowe

$$\begin{aligned} \text{Var}(\mathbf{f}) &= \mathbf{X}_f \text{Var}(\mathbf{b}) \mathbf{X}_f' + \text{Var}(\boldsymbol{\varepsilon}_f) = \sigma^2 \mathbf{X}_f (\mathbf{X}' \mathbf{X}) \mathbf{X}_f' + \sigma^2 \mathbf{I}_g \\ &= \sigma^2 \left[ \mathbf{I}_g + \mathbf{X}_f (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_f' \right] \end{aligned}$$

Aby wyprowadzić rozkład testu prognoz zauważmy, że

$$\mathbf{f} = \mathbf{X}_f (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}_f$$

a więc błąd prognozy jest liniową funkcją błędów w okresie estymacji i prognozy. Jeśli założymy, że założenia *KMRL* są spełnione zarówno dla okresu estymacji jak i prognozy, to błędy losowe są nieskorelowane i mają łączny rozkład normalny, i błąd prognozy ma rozkład normalny o rozkładzie

$$\mathbf{f} \sim N\left(\mathbf{0}, \sigma^2 \left[ \mathbf{I}_g + \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_f \right]\right)$$

Oznacza to, że statystyka

$$\frac{\mathbf{f}' \left[ \mathbf{I}_g + \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_f \right]^{-1} \mathbf{f}}{\sigma^2} \sim \chi_g^2$$

Podana powyżej statystyka jest mało użyteczna, ponieważ zależy od nieznanego parametru  $\sigma^2$ . Zauważmy jednak, że  $\frac{\mathbf{e}'\mathbf{e}}{(T-K-1)} \sim \chi_{T-K-1}$  i dystrybuanty  $\mathbf{f}$  i  $\mathbf{e}$  są niezależne, ponieważ

$$\begin{aligned} \text{Cov}(\mathbf{f}, \mathbf{e}) &= \text{Cov} \left\{ \left[ \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}_f \right], \mathbf{M}\boldsymbol{\varepsilon} \right\} \\ &= \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \text{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \mathbf{M} + \text{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}_f') \mathbf{M} \\ &= \sigma^2 \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{M} = \mathbf{0}, \end{aligned}$$

gdzie dwie ostatnie równości wynikają z tego, że  $\boldsymbol{\varepsilon}$  i  $\boldsymbol{\varepsilon}_f$  są niezależne i  $\mathbf{M}\mathbf{X} = \mathbf{0}$ . Konkludując

$$\frac{\mathbf{f}' \left[ \mathbf{I}_g + \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_f \right]^{-1} \mathbf{f}}{\frac{g}{\frac{\mathbf{e}'\mathbf{e}}{(n-K-1)}}} \sim F(g, n - K - 1)$$

Forma kwadratowa  $\mathbf{f}' \left[ \mathbf{I}_g + \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_f \right]^{-1} \mathbf{f}$  jest potencjalnie trudna do policzenia. Pokazano jednak, że

$$S_{D+F} - S_D = \mathbf{f}' \left[ \mathbf{I}_g + \mathbf{X}_f (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_f \right]^{-1} \mathbf{f}$$

gdzie  $S_{D+F}$  jest sumą kwadratów reszt gdy model został wyestymowany na pełnej próbie, a  $S_D$

jest sumą kwadratów reszt, gdy model został wyestymowany na danych jedynie z okresu estymacji. Możemy teraz sformułować twierdzenie dotyczące rozkładu statystyki testującej hipotezę (5.1).

**Twierdzenie 5.1** *Hipotezę (5.1) można testować za pomocą statystyki*

$$\frac{(S_{D+F} - S_D)}{\frac{g}{S_D}} \sim F(g, n - K - 1)$$
$$\frac{S_D}{(n - K - 1)}$$

gdzie  $S_{D+F}$  jest sumą kwadratów reszt z regresji na całej próbie a  $S_D$  sumą kwadratów reszt z regresji na danych z okresu estymacji.

**Literatura:** Steward (1991) str. 82-88.

## 5.4. Test Chowa

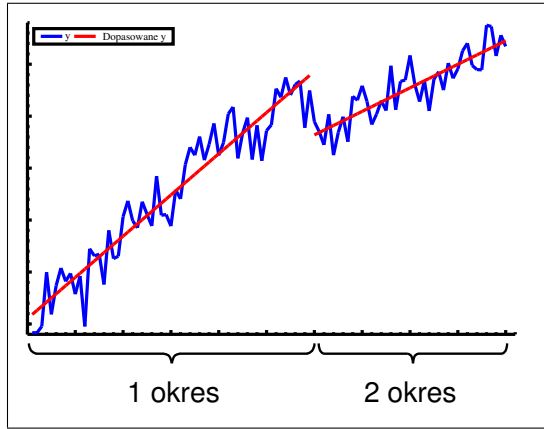
Test Chowa jest testem na stabilność parametrów. Testujemy za jego pomocą, czy w dwóch próbkach parametry modelu są takie same. Najczęstsze przykłady zastosowania testu Chowa to badanie stabilności parametrów w czasie i badanie stabilności modelu w podpróbkach (kategoriach) próby przekrojowej. Prosty przypadek, kiedy załamanie następuje w parametrach trendu liniowego, wraz z dopasowanymi dla każdego z okresów regresjami, zilustrowany jest na rysunku 3.

Inny przypadek, kiedy możemy zastosować test Chowa dotyczy sytuacji kiedy podejrzewamy, że populacja dzieli się na kilka różnych grup (np. emeryci i pozostali), których zachowanie powinno być opisane zupełnie różnymi modelami. Sytuację taką ilustruje rysunek 4.

Załóżmy, że mamy następujące modele dla podpróbek:

$$\mathbf{y}_s = \mathbf{X}_s \boldsymbol{\beta}_s + \boldsymbol{\varepsilon}_s, \quad \boldsymbol{\varepsilon}_s \sim N(\mathbf{0}, \sigma_s^2 \mathbf{I}_{n_s}) \quad (5.2)$$

gdzie  $s = 1, \dots, m$  oznacza numer próbki a  $n_s$  liczebność danej próbki. W każdej z podpróbek obserwacje dotyczą tych samych zmiennych egzogenicznych i endogenicznych przy czym podejrzewamy, że parametry modelu są różniące się dla poszczególnych podpróbek. Jeśli istotnie tak jest,



Rysunek 3: Załamanie w parametrach trendu liniowego

to powinniśmy wyestymować  $m$  modeli, osobny dla każdej z podpróbek. Dla szczególnego przypadku, kiedy  $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2 = \sigma^2$  estymację tych  $m$  modeli można sprowadzić do estymacji jednego modelu. Zdefiniujmy indeks  $i$  przebiegający po wszystkich obserwacjach,  $i = 1, \dots, n$  i  $n = \sum_{s=1}^m n_s$  oraz następującą zmienną zero-jedynkową:

$$Q_{s,i} = \begin{cases} 1 & \text{jeśli obserwacja } i \text{ należy do próbki } s \\ 0 & \text{jeśli obserwacja } i \text{ nie należy do próbki } s \end{cases}$$

i sprowadźmy model (5.2) do następującej postaci

$$y_i = \sum_{s=1}^m Q_{s,i} \mathbf{x}_i \boldsymbol{\beta}_s + \varepsilon_i, \quad \varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

Po zdefiniowaniu zmiennej  $\mathbf{x}_i^* = (Q_{1,i}\mathbf{x}_i, \dots, Q_{m,i}\mathbf{x}_i)$  i  $\boldsymbol{\beta}^* = (\beta'_1, \dots, \beta'_m)'$  model ten sprowadzi się do typowego modelu liniowego

$$y_i = \mathbf{x}_i^* \boldsymbol{\beta}^* + \varepsilon_i$$

lub

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\varepsilon}$$

Testowanie hipotezy zerowej o tym, że parametry modeli szacowanych w różnych próbkach są sobie równe sprowadza się do testowania hipotezy zerowej

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_m \quad (5.3)$$

Test taki można łatwo przeprowadzić w pakietach ekonometrycznych, które umożliwiają testowanie hipotez złożonych (np. PCGIVE). Druga możliwa strategia oparta jest na obserwacji, że suma kwadratów reszt w modelu bez ograniczeń równa jest sumie sum kwadratów reszt z modeli wyestymowanych na podpróbkach  $S = \sum_{i=1}^m S_i$ . Suma kwadratów reszt w modelu z ograniczeniami  $S_R$  równa jest sumie kwadratów reszt w modelu wyestymowanym na całej próbie

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Ilość ograniczeń jest równa  $(K - 1)(m - 1)$  i możemy sformułować następujące twierdzenie:

**Twierdzenie 5.2** *Hipotezę (5.3) dla modelu (5.2) można przetestować za pomocą stystyki testu Chowa*

$$F = \frac{\frac{(S_R - S)}{(K+1)(m-1)}}{\frac{S}{n-m(K+1)}} = \frac{\frac{(S_R - \sum_i S_i)}{(K+1)(m-1)}}{\frac{\sum_i S_i}{n-m(K+1)}} \sim F_{(K+1)(m-1), n-m(K+1)}$$

gdzie  $S_R$  jest sumą kwadratów reszt z regresji na całej badanej próbie a  $S_i$  jest sumą kwadratów reszt z regresji na  $i$ -tej podpróbcie.

- **Test *CUSUM* i *CUSUMQ***

W omówionych dotąd testach zakładaliśmy, że moment w którym ulegają zmianie wielkości parametrów jest z góry znany. Pojawia się jednak pytanie, czy można przetestować istnienie załamania w wielkościach parametrów nawet wtedy, kiedy nieznan jest moment tego załamania. Na podstawie tego co wcześniej powiedzieliśmy na temat wariancji błędu prognozy, wiemy, że dla

$$e_t = y_t - \mathbf{x}_t \mathbf{b}_{t-1}$$

gdzie  $\mathbf{b}_{t-1}$  jest estymatorem *MNK* policzonym na podstawie  $t - 1$  pierwszych obserwacji, wariancja błędu wyniesie

$$\sigma_{ft}^2 = \sigma^2 [1 + \mathbf{x}_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1}) \mathbf{x}'_t]$$

Zdefiniujmy

$$w_t = \frac{e_t}{\sqrt{1 + \mathbf{x}_t (\mathbf{X}'_{t-1} \mathbf{X}_{t-1}) \mathbf{x}'_t}} \sim N(0, \sigma^2)$$

W przypadku, kiedy tak znormalizowane reszty będą znacznie przekraczać  $\pm 2\sigma^2$  dla któregoś z  $t$ , będziemy mogli podejrzewać, że dla tego  $t$  nastąpiło zmiana wielkości parametrów (załamanie strukturalne). Problem z taką procedurą polega na tym, że postępując tak wielokrotnie testujemy hipotezy proste, co jak wcześniej powiedziano utrudnia formalne testowanie ze względu na trudności z uzyskaniem prawdziwego poziomu istotności (obciążenia Lavella).

Inną procedurą testowania jest test *CUSUM*. Zdefiniujmy następującą statystykę

$$W_t = \sum_{i=K+1}^{r=t} \frac{w_i}{\hat{\sigma}}$$

gdzie

$$\hat{\sigma}^2 = \frac{\sum_{i=K+1}^{r=t} (w_i - \bar{w})^2}{T - K - 1}$$

$$\bar{w} = \frac{\sum_{i=K+1}^{r=t} w_i}{T - K}$$

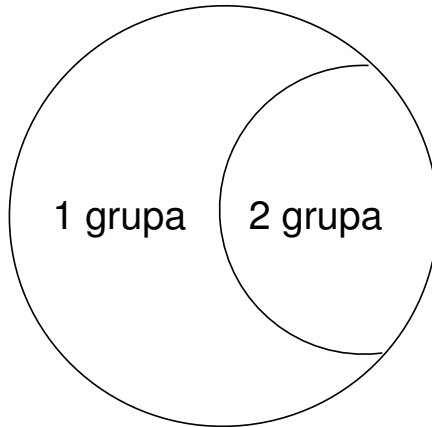
Test przeprowadzamy porównując wartości statystyki  $W_t$  z wartościami krytycznymi. Zazwyczaj można to zrobić za pomocą wykresu generowanego automatycznie przez pakiet ekonometryczny. Wartości  $W_t$  przekraczające wartość krytyczną będą nas skłaniać do odrzucenia hipotezy zerowej o stabilności parametrów. W przypadku testu *CUSUMQ* badamy odpowiednio znormalizowaną sumę kwadratów błędów prognoz:

$$S_r = \frac{\sum_{i=K+1}^{r=t} w_i^2}{\sum_{i=K+1}^{r=T} w_i^2}$$

Podobnie jak w przypadku testu *CUSUM* wysokie wartości tej statystyki będą nas skłaniać do odrzucenia hipotezy zerowej o stabilności parametrów.

**Literatura:** Steward (1991) str. 101-102, Green (1997) str. 349-353

Populacja



Rysunek 4: Podgrupy o podobnej wielkości