

SPÓJNOŚĆ I PORÓWNYWALNOŚĆ DANYCH JAKO KRYTERIA OCENY JAKOŚCI STATYSTYKI DZIAŁALNOŚCI GOSPODARCZEJ

Andrzej Młodak



Urząd Statystyczny w Poznaniu

Ośrodek Statystyki Małych Obszarów

Europejska Praktyka Statystyczna

– Zasada 14

European Statistics Code of Practice – Principle 14

- **Statystyka europejska winna być spójna wewnątrznie oraz względem czasu a także porównywalna pomiędzy regionami i krajami**
- **Wskaźniki statystyczne**
 - są wewnątrznie spójne i porównywalne
 - są spójne lub zbieżne w rozsądnym okresie czasu
 - są opracowywane w oparciu o jednolite standardy
 - jeśli pochodzą z różnych źródeł to są porównywalne i zbieżne
 - okresowa wymiana doświadczeń w ramach ESS i poza nim zapewnia porównywalność międzynarodową.

Układ prezentacji

■ Spójność

- definicja spójności; spójność a precyzja
- typy spójności
- źródła niespójności
- pomiar spójności

■ Porównywalność

- porównywalność a spójność
- źródła braku porównywalności
- typy i ocena porównywalności

■ Kompleksowa ocena obu aspektów.

Spójność

■ Definicja spójności

- odnosi się do odpowiedniości pomiędzy zbiorem zmiennych statystycznych opisujących parametry dla skończonej populacji w kontekście wzajemnie ze sobą powiązanych zjawisk społeczno – gospodarczych
- poziom spójności informuje nas czy i w jakim stopniu odnośne dane statystyczne mogą być analizowane łącznie i jak wskazać ich „optymalne” wielkości, tzn. jeśli dwie zmienne są powiązane metodologicznie i zmiana jednej zmiennej wpływa na drugą, to obie mogą być rozpatrywane łącznie.

Spójność

■ Definicja spójności (dok.)

- według Eurostatu spójność dwu lub więcej wyników statystycznych odnosi się do stopnia w jakim procesy statystyczne, w których te informacje wytworzono oparte zostały na tych samych koncepcjach (klasyfikacjach, definicjach i populacjach docelowych) oraz zharmonizowanych metodach
- pojęcie spójności stosuje się zazwyczaj do oceny zakresu, w jakim rezultaty różnych procesów statystycznych mają potencjał do tego, aby być rzetelnie stosowane w powiązaniu ze sobą.

Spójność

■ Spójność a precyzja

- precyzja jest jednym z aspektów spójności i odnosi się do oceny jakości poszczególnych procesów bez uwzględniania analizy możliwości ich łączenia
- precyzja jest mierzona w oparciu o metadane operacyjne (np. wskaźniki błędów), podczas gdy spójność odnosi się do metadanych opisowych (np. koncepcje i metody)
- relacja pomiędzy spójnością a precyzją zależy od
 - *spójności logicznej*
 - *błędów estymacji zachodzących w procesach statystycznych.*

Spójność

■ Typy spójności

- spójność wewnętrzna – spójność w ramach jednolitego zbioru statystyk gospodarczych lub pomiędzy danymi pochodzącymi z różnych źródeł (m.in. definicyjna)
- spójność zewnętrzna – spójność pomiędzy danymi dotyczącymi działalności gospodarczej a informacjami makroekonomicznymi
- inne typy spójności według Eurostatu
 - *spójność pomiędzy danymi rocznymi a krótkookresowymi*
 - *spójność z rachunkami narodowymi*
 - *spójność z innymi danymi statystycznymi.*

Spójność

- **Źródła niespójności (Körner i Puch (2011)):**
 - różnice w zakresie koncepcji (populacja docelowa, okres odniesienia, definicja pozycji analitycznych, itp.)
 - różnice w zakresie stosowanych metod (np. metody i procedury gromadzenia danych, podejście w zakresie przetwarzania informacji, schemat losowania, itp.)
 - różnice łączone (np. błąd próbkowania stanowi kombinację efektów schematu losowania i odchyień w ich praktycznym wdrażaniu).

Spójność

■ Pomiar spójności

- najprostszy sposób: wyznaczenie różnic pomiędzy wynikami analizowanych badań
- w statystyce gospodarczej spójność zależy od integracji informacji uzyskanych z różnych źródeł
- nie wszystkie aspekty spójności są mierzalne
- porównanie szacunków obrotów uzyskanych na podstawie badania z odpowiednim oszacowaniem dokonany z połączenia jego wyników z danymi administracyjnymi (Casciano *i in.* (2012))
- miara braku precyzji winna obejmować wszystkie źródła danych, nie tylko próby.

Porównywalność

■ Porównywalność a spójność

- o porównywalności mówimy, gdy oceniamy w jakim stopniu wytwory (nominalne) tego samego procesu statystycznego ale dla różnych okresów czasu lub regionów mogą być efektywnie użyte łącznie - spójność dotyczy możliwości łączenia różnych zmiennych dla tego samego regionu lub okresu czasu
- porównywalność jest pojęciem słabszym od spójności
- porównywalność odnosi się do możliwości dokonywania efektywnych porównań pomiędzy jednostkami i zmiennymi w zakresie różnych aspektów i kryteriów.

Porównywalność

■ Źródła braku porównywalności (Eurostat)

- różnice dotyczące koncepcji
 - *populacja docelowa (jednostki i pokrycie)*
 - populacje dla różnych edycji badania mogą być odmienne (miesięczne badanie przedsiębiorstw ma inną populację niż roczne, różne zakresy definicyjne pracujących w Europie, itp.)
 - *pokrycie geograficzne*
 - pewne typy obszarów mogą być ujmowane w jednych a pomijane w innych krajach
 - *okres odniesienia*
 - w różnych krajach te same informacje od przedsiębiorstw pozyskiwane są według stanu na trzeci poniedziałek lub na pierwszy dzień roboczy w miesiącu
 - *definicje i klasyfikacje dotyczące danych.*

Porównywalność

■ Źródła braku porównywalności (Eurostat)

- różnice dotyczące metod gromadzenia, przetwarzania i analizy danych
 - *operat losowania*
 - faktyczna populacja objęta badaniem może się różnić od populacji źródłowej, np. małe firmy są badane reprezentatywnie raz w roku, duże – w sposób pełny – miesięcznie
 - *źródła danych i schemat losowania*
 - źródła danych lub schematy losowania mogą się różnić
 - *gromadzenie danych, pokrycie informacyjne i edycja*
 - odmienne występowanie problemów dotyczących braku odpowiedzi i ich redukcji w tym samym badaniu
 - *imputacja i estymacja*
 - metody zależą od miejsca i czasu przeprowadzania badania.

Porównywalność

■ Źródła braku porównywalności – przykład

- badanie działalności gospodarczej (DG-1) – miesięczny obowiązek sprawozdawczy dla wszystkich przedsiębiorstw zatrudniających ponad 49 osób i wylosowanych o mniejszej liczbie pracujących
- uogólnienia podstawowych danych ilościowych dokonywane przy pomocy wskaźnika

$$u_d \equiv \frac{\sum_{j \in N_{dr}} \vartheta_j + \sum_{j \in N_d \setminus N_{dr}} \theta_j}{\sum_{j \in N_{dr}} \vartheta_j}$$

- N_d - zbiór jednostek na poziomie agregacji d , N_{dr} – zbiór jednostek na poziomie d , które złożyły sprawozdanie, ϑ_j – zmienna liczba pracujących dla j -tej jednostki (z meldunku), θ_j – stała liczba pracujących dla j -tej jednostki (z rejestru).

Porównywalność

■ Źródła braku porównywalności – przykład

- model uogólnienia

$$\hat{\varphi}_{(d)} \equiv u_d \sum_{i \in N_d} \varphi_i$$

- φ_i – uogólniana wielkość dla i -tej jednostki (np. przychody ze sprzedaży wyrobów i usług, obrót, itp.)
- problemy z brakiem spójności i porównywalności:
 - liczba pracujących w rejestrze może się istotnie różnić od faktycznej
 - uogólnienie zależy od poziomu agregacji, zatem np. dla województwa możemy otrzymać dwie różne liczby, jeśli estymujemy odrębnie na poziomie grupy/działu PKD i sektorów własności – a takie są założenia badania.

Porównywalność

■ Źródła braku porównywalności – przykład

■ metody rozwiązywania problemów

- *oszacowanie wpływu zastąpienia nieznannej zmiennej liczby pracujących przez dane rejestrowe (na podstawie dostępnych danych)*

$$v_d \stackrel{\text{def}}{=} \left| \frac{1}{\sum_{j \in N_{dr}} \theta_j} - \frac{1}{\sum_{j \in N_d \setminus N_{dr}} \theta_j} \right| \frac{\sum_{j \in N_d \setminus N_{dr}} \theta_j}{u_d}$$

- dla uzyskania oceny kompleksowej wielkości te można zsumować lub uśredniać po poziomach agregacji w ramach danego obszaru
- *dla usunięcia różnic w uogólnieniach na różnych poziomach agregacji konieczna jest imputacja. Powyższy wskaźnik może pomóc zidentyfikować, gdzie jest ona najbardziej potrzebna. Prace w tym zakresie właśnie się toczą.*

Porównywalność

■ Typy porównywalności

- porównywalność w czasie
- porównywalność pomiędzy regionami
- porównywalność między innymi domenami (np. płeć)
- wewnętrzna porównywalność (dla segmentów procesu przy pojedynczym okresie czasu i regionie)
- porównywalność statystyk rocznych i dla okresów śródrocznych
- porównywalność z rachunkami narodowymi
- porównywalność z innymi statystykami (pochodzącymi z innych badań).

Porównywalność

■ Ocena porównywalności

■ Podejście Ogólne

- *ma na celu wyjaśnienie wszystkich czynników mających wpływ na brak porównywalności (zmian legislacyjnych, standardów międzynarodowych, definicji, itp.)*
- *kroki*
 - I. Systematyczna ocena możliwych przyczyn braku porównywalności
 - II. Prognoza wpływu tych różnic na kształt danych wynikowych
 - III. Agregacja i podsumowanie możliwych efektów w ujęciu ogólnym.

Porównywalność

■ Ocena porównywalności (cd.)

- jakościowa kontrola zmiennych
 - *kontrola typu wielowymiarowego*
 - czy suma wartości różnych zmiennych równa się odpowiedniej wartości ogółem
 - *kontrola typu hierarchicznego*
 - czy suma wartości dla niższych poziomów agregacji jest równa odpowiedniej wartości dla poziomu wyższego
 - *detekcja obserwacji odstających*
 - klasyczny przedział zmienności $\bar{x} \pm z \cdot \sigma_x$,
 - medianowy przedział zmienności $med(X) \pm z \cdot mad(X)$
 - *analiza wskaźników wzrostu (tolerancja, wypełnianie luk).*

Porównywalność

■ Ocena porównywalności (cd.)

- porównywalność w czasie
 - *dokładne wskazanie czy są zmiany a jeśli tak, to czy powodują załamanie szeregu czasowego czy też nie*
 - *możliwe działania:*
 - wyczerpujące: konwersja danych sprzed zmiany przy zastosowaniu warunków po zmianie i podanie obu wariantów
 - mniej kosztowne: dostarczenie użytkownikom danych wskaźników przejścia umożliwiających im analizy
 - najmniej kosztowne: prosty opis zmian, które miały miejsce i dostarczenie jakościowej oceny ich prawdopodobnego wpływu na wielkość oszacowań.

Porównywalność

■ Ocena porównywalności (cd.)

- porównywalność między regionami
 - *może być oceniana poprzez porównywanie metadanych parami według regionów lub w ramach danego regionu z przyjętymi standardami (np. normami ESS)*
 - *można tu rozpatrywać sytuację, gdy użyto tych samych bądź różnych rodzajów procesów*
 - *miara kompleksowa*
 - *suma różnic cząstkowych wyrażonych w wartościach bezwzględnych lub na skali porządkowej (od 5 do 0, w malejącym porządku istotności) według opinii ekspertów*
 - *ważenie odzwierciedlające wpływ poszczególnych elementów metadanych na porównywalność.*

Porównywalność

■ Ocena porównywalności (cd.)

- porównywalność dla innych domen
 - *dotyczy ona np. grup przedsiębiorstw według wielkości, rodzaju działalności gospodarczej i może być oceniana analogicznie jak w przypadku regionów*
- spójność i porównywalność wewnętrzna
 - *ocena różnic arytmetycznych oraz logicznych*
- porównywalność danych rocznych i krótkookresowych
 - *metody oceny:*
 - jeśli dane roczne i śródroczne mierzą to samo zjawisko w postaci innej niż indeksy dynamiki to agregaty roczne skonstruowane w oparciu o dane krótkookresowe mogą być porównywane z rocznymi wartościami ogółem
 - wskaźniki wzrostu mogą być porównywane w ujęciu rok do roku.

Porównywalność

■ Ocena porównywalności (dok.)

- spójność i porównywalność z rachunkami narodowymi
 - *efektywne narzędzie wychwytywania niespójności pomiędzy danymi gospodarczymi z różnych źródeł*
 - *ocena własności danych i dopasowań dla uzyskania bilansu*
- statystyka zwierciadlana
 - *dane dla jednego regionu (kraju) mają swój odpowiednik w innym (np. przepływy ludności czy kapitału)*
 - *niezgodność może wskazywać na niedo- lub przeszacowania*
- porównywalność z innymi statystykami i benchmarking
 - *ocena stopnia możliwego łączenia danych z różnych źródeł*
 - *analiza i dokumentowanie i dobrych praktyk dla tworzenia nowych, lepiej zharmonizowanych rozwiązań oraz technik.*

Kompleksowa ocena obu aspektów

■ Wskaźnik spójności/porównywalności

- wskaźnik różnic metodologicznych

$$I_{M(r)} \stackrel{\text{def}}{=} \frac{1}{pq_0(t-1)} \sum_{i=1}^p \sum_{j=1}^q w_{ij} \sum_{k=2}^t d_{ijk_r}$$

- $d_{ijk_r} \in [0,1]$ - poziom spójności definicji r -tej zmiennej w k -tej edycji badania w stosunku do edycji $k-1$ w kontekście j -tego aspektu w opinii i -tego eksperta (0 – zupełny brak, 1 – pełna)
- $w_{ij} \in [0,1]$ - waga przypisana przez i -tego eksperta efektowi j -tego aspektu, przy czym $\sum_{j=1}^q w_{ij} = 1$ dla każdego $i=1,2,\dots,p$.
- wskaźnik przyjmuje wartości z przedziału $[0,1]$. Zmienne wystarczająco spójne: $S_M \stackrel{\text{def}}{=} \{X_r : r \in \{1,2,\dots,m\} \wedge I_{M(r)} \geq \theta\}$ (θ – próg spójności, np. 0,6).

Kompleksowa ocena obu aspektów

■ Wskaźnik spójności/porównywalności (cd.)

- wskaźnik spójności porównywalności w czasie i według innych dziedzin

$$I_{D(r)} \stackrel{\text{def}}{=} \frac{1}{ap(t-1)} \sum_{i=1}^p \left(w_{ir(T)} \sum_{k=2}^t d_{ikr(T)} + w_{ir(U)} \sum_{k=2}^t d_{ikr(U)} \right)$$

- $w_{ir(T)}$ i $w_{ir(U)}$ - wagi (z przedziału $[0,1]$) przyporządkowane przez i -tego eksperta znaczeniu porównywalności w czasie (T) i według innych domen (D) dla j -tej zmiennej należącej do zbioru S_M .
- $d_{ikr(T)}, d_{ikr(U)} \in [0,1]$ - odpowiednie oceny stopnia porównywalności w czasie (T) i według domen (D)
- wskaźnik przyjmuje wartości z przedziału $[0,1]$.

Kompleksowa ocena obu aspektów

■ Wskaźnik spójności/porównywalności (dok.)

- wskaźnik kompleksowy

$$I_C \stackrel{\text{def}}{=} \frac{\text{med}_{r=1,2,\dots,m} I_{M(r)} + b \min_{r \in \{1,2,\dots,m\}: X_r \in S_M} I_{D(r)}}{2} \in [0,1]$$

- $b=1$ gdy $|S_M| \geq [n/2] + 1$ oraz $b=0$ w przeciwnym razie

■ Ujednolicanie miar i zakresu wartości

- normalizacja
- benchmarking:

$$z_{ij} = \frac{|x_{ij} - \varphi_j|}{d_i}$$

- x_{ij} – wartość cechy j dla obiektu i , φ_j – wartość j -tej cechy dla wzorca, d_i – odległość obiektu i od wzorca.

Kompleksowa ocena obu aspektów

■ Pozostałe narzędzia

- mierniki kompleksowe (wg Z. Hellwiga lub pozycyjne)
- charakterystyki precyzji
 - *błędy pomiaru dla jednostek wybieranych z prawdopodobieństwem 1*
 - *różnice pomiędzy podstawowymi charakterystykami dla dwóch różnych badań*
 - *poziom braku odpowiedzi jako ocena jakości wskaźników wtórnych (redukcja pewnych efektów przez ilorazy)*
 - *precyzja estymacji, przedziały niepewności*
- metoda koordynowanych statystyk wynikowych
 - *jeden zbiór szacunków jest wzorcem dla oceny innego, np. statystyki roczne dla statystyk krótkookresowych.*

Dziękuję za uwagę