

Miary dwuważonych ocen

Miary dwuważonych ocen – pakiet wzorów statystycznych przeznaczonych do kompleksowej analizy rezultatów badań prowadzonych za pomocą *Skal dwuważonych ocen*, dedykowanych na potrzeby *Ewaluacji splotowej*. Miary te w większości mogą być wykorzystywane również w przetwarzaniu danych z pomiarów skalą Likerta lub dyferencjałem semantycznym, pod warunkiem przypisania przeciwnych liczb do symetrycznych, przeciwstawnych wskaźników skali. Nazwę uzasadnia dwuwymiarowa analiza wskaźników z porównywaniem wartości i ważkości wypowiedzi respondentów, a także dwukierunkowa analiza zróżnicowania rozkładów, nieodzowna przy rozkładach asymetrycznych oraz dwukierunkowa, wektorowa analiza efektywności badanych procesów^[1].

Przekroje przez zmienne

Za pomocą skal Likerta wyznacza się wyniki ogólne powstałe po skumulowaniu wszystkich wskaźników (indeksów skali). W *Ewaluacji splotowej* przeprowadza się bardziej szczegółowe analizy zmiennych częściowych. W tej sytuacji zmienne losowe **X** w poniższych wzorach mogą być danymi surowymi (wagami) albo danymi wstępnie zagregowanymi (średnimi z wag). I tak rezultaty połówkowe powstają z agregacji połowy wskaźników, *czynniki* to średnie z minimum 4 wskaźników, a *składniki* to średnie z par wskaźników kontrolnych (→ indeksacja w *Ewaluacji splotowej*). Niektóre ze statystyk oblicza się tylko na ściśle wyznaczonych poziomach agregacji, np. konkluzyjność i rzetelność dla wskaźników elementarnych, a spójność dla kontrolnych par wskaźników.

Przekroje przez przypadki

W każdym badaniu istotne znaczenie ma dobór prób badawczych o odpowiedniej reprezentatywności i liczebności. W analizach porównawczych z ogółu respondentów wyodrębnią się podgrupy (→ próby badawcze w *Ewaluacji splotowej*) i dla nich oblicza rezultaty. Wskutek tego w poniższych wzorach liczebności **{n}** mogą oznaczać liczbę wszystkich respondentów albo liczbę osób tworzących podgrupę.

Relatywność wartości oczekiwanych

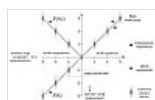
Podane tu wartości oczekiwane pełnią jedynie rolę poglądową, gdyż zależą od rodzaju wskaźników i od tego, co jest badane. Parametry te pochodzą z 7-letniej ewaluacji miękkich efektów kształcenia (→ strategia psychometryczna w *Ewaluacji splotowej*), tj. badania uczniów i studentów za pomocą *Kwestionariusza Kultury Informatycznej*^[1]. Mimo reprezentatywnych prób ~3100 respondentów, wartości oczekiwane w badaniach innymi kwestionariuszami mogą być inne.

Miary chwilowe z jednego pomiaru

Dla statystyk dwuważonych przyjęto konwencję zwartych nazw i symboli literowych uzupełnionych znakami, które charakteryzują właściwość danej miary. Znak wartości ujemnej – lub dodatniej + jest integralną składową tych parametrów, które są wyłącznie ujemne bądź wyłącznie dodatnie (**-b**, **-c**, **+c**, **-d**, **+d**). Prefiks **±** oznacza parametr, który może przyjmować wartości ujemne lub dodatnie (**±s**). Parametr porównawczy par wskaźników wyróżniono indeksem dolnym (**c_p**)^[1].

Jakość *Quality* Intensywność *Intensity*

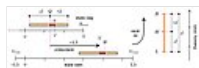
$$Q = \frac{1}{n} \sum_{i=1}^n X_i \quad I = \frac{1}{n} \sum_{i=1}^n |X_i| \quad (1)$$



Podstawowa miara jakości wypowiedzi **Q** jest średnią arytmetyczną wag **X_i** z przedziału **[-3; +3]** przypisanych opcjom wybranym przez **{n}** respondentów z 7-stopniowej *kafeterii*, ale koniecznie po odwróceniu wskaźników z tezami fałszywymi na **[+3; -3]**. Miara intensywności wypowiedzi **I** jest średnią arytmetyczną wartości bezwzględnych **|X_i|** z tych samych wag. Można te wyniki skalować procentowo: **Q%** = $\frac{1}{3}Q \cdot 100\%$, dzieląc przez najwyższą wagę skali, tu **X_{max} = 3**.

Wartość *Value* Ważkość *Weight*

$$V = Q + 2,5 \quad W = I + 2,5 \quad (2)$$



Spis treści

Miary chwilowe z jednego pomiaru

Miary dystansowe z dwóch pomiarów

Wyznaczniki norm i jakości pomiarowej

Przypisy

Bibliografia

Optymalna do publikacji porównawczych miara wartości badanych cech V , wyrażona w skali ocen akademickich ($ndst = 2$) trafnie interpretowanych przez Polaków, powstaje poprzez przeskalowanie jakości wypowiedzi Q o $+2,5$. Analogicznie miara ważkości W (pewności badanych co do swych wyborów) powstaje poprzez przeskalowanie intensywności I o $+2,5$ ($\rightarrow$ rys.). W praktyce wyniki na tej skali po uśrednieniu dla grup badawczych mieszczą się w przedziale $[2; 5]$.

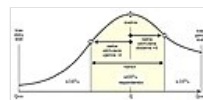
Rozziew $Bias -b$

$$-b = Q - I \quad (3)$$

Znamienna dla analizy dwuwymiarowej różnica między jakością Q a intensywnością I wypowiedzi jest miarą dysjunkcji, czyli niezgodności części respondentów w danej próbie badawczej z opinią większości. Rozziew $-b$ jest oznaką negowania wartości wyznaczonych empirycznie jako *standardy ewaluacyjne*, dlatego ma wartość ujemną. W praktyce obliczeń na poziomie elementarnych wskaźników mieści się w przedziale $[-0,2; -0,5]$.

Odchylenia realne $Real\ deviations -d +d$

$$-d = -\sqrt{\frac{1 + \frac{h}{l}}{n} \sum_{i:\{X_i < Q\}} (X_i - Q)^2} \quad +d = +\sqrt{\frac{1 + \frac{l}{h}}{n} \sum_{j:\{X_j > Q\}} (X_j - Q)^2} \quad (4)$$



Ze względu na częstą asymetrię rozkładów empirycznych, miary zróżnicowania reprezentowane są przez dwa parametry: odchylenie ujemne $-d$ oraz dodatnie $+d$ liczone osobno dla $\{l\}$ rezultatów X_i gorszych od średniej Q oraz $\{h\}$ rezultatów X_j lepszych od średniej Q przy liczebności całej próby $\{n\}$. Są to miary bazowe dla wyznaczania innych miar ($\rightarrow$ rys.), bardziej trafne i bardziej przydatne do analiz dwuważonych niż *odchylenie standardowe*.

Rozrzut $Dispersion d$

$$d = |-d| + |+d| \quad (5)$$

Rozrzut $\{d\}$ jako suma modułów obu odchylen realnych $|-d|$ oraz $|+d|$ jest miarą realnej dyspersji, tj. rozproszenia (zróżnicowania) wyników kwalifikowanej większości $\sim 68\%$ respondentów (przypadków). Miara ta ma wyrazistą interpretację geometryczną na rozkładzie rezultatów badań. Jej rozpostarcie dla wyników ogólnych w badaniach grup osiąga szerokości od $0,75$ do $1,25$ a wartość oczekiwana dla populacji dąży do 1 przy założeniu $X_{max} = 3$.

Skośność $Skewness \pm s$

$$\pm s = \frac{|+d| - |-d|}{|+d| + |-d|} \quad (6)$$



Stosunek różnicy do sumy wartości bezwzględnych z obu odchylen realnych $|+d|$ i $|-d|$, obliczonych wg wzorów 4, wyznacza znormalizowaną do zakresu $[-1; +1]$ miarę asymetrii rozkładu empirycznego, zwanej skośnością $\pm s$. Statystyka ta ma wartość ujemną przy asymetrii lewostronnej, a dodatnią – przy prawostronnej. Pomiary z użyciem *kafeterii Likerta* potwierdzają typową asymetrię lewostronną, gdyż skośność liczona dla wskaźników przyjmuje wartości z zakresu $[-0,1; -0,5]$ a wartość oczekiwana skośności wyników ogólnych dla populacji dąży do $-0,05$.

Poziomy $Levels\ HML$ (*high, mid, low*)

$$M = V \quad H = M + |+d| \quad L = M - |-d| \quad (7)$$

Do wizualizacji rezultatów w skali ocen akademickich wykorzystuje się trójpoziomowy wykres HML . Poziom średni M równy jest wartości V ze wzoru 2, a dodanie i odjęcie odchylen realnych ze wzorów 4 wyznacza poziomy wysoki H i niski L , jako potrzebne do analiz progi podziałów na frakcje o lepszych i gorszych wynikach. Szacunkowo w każdej z tych skrajnych frakcji znajdzie się po około 16% przypadków ($\rightarrow$ rys. do wzorów 2 i 4).

Ufność realna $Real\ confidence -c +c$

$$-c = U(\alpha) \frac{-d}{\sqrt{n}} \quad +c = U(\alpha) \frac{+d}{\sqrt{n}} \quad c = |-c| + |+c| \quad (8)$$



Dwa zakresy tolerancji z lewej $-c$ i prawej $+c$ strony średniej empirycznej wyznaczają na rozkładzie przedział ufności c , w obszarze którego z błędem nie większym niż założony poziom istotności α (alfa) mieści się średnia oczekiwana dla populacji, szacowana z próby o liczebności $\{n\}$ przy odchyleniach realnych $-d$ oraz $+d$. Współczynnik $U(\alpha)$ ma wartości $U(1\%) = 2,576$ w testach wysokiej wiarygodności, $U(5\%) = 1,96$ zwykle w badaniach społecznych, ew. $U(10\%) = 1,645$.

Spójność $Coherence c_p$

$$c_p = 1 - \frac{1}{npX_{max}} \sum_{i=1}^n \sum_{j=1}^p |X_{i;2j-1} - X_{i;2j}| \quad (9)$$

spójność par wskaźników

5	4	5	5	...
3	3	4	3	...
4	4	3	2	...
...	...	...	...	...

n przypadków

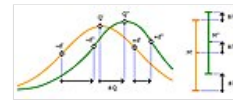
Miara wewnętrznej koherencji danych – spójność c_p – obliczana jest z porównania kontrolnych par wskaźników, które z założenia w *Skali dwuważonych ocen* są obok siebie. Przy takim założeniu zsumować należy wartości bezwzględne wszystkich $\{p\}$ różnic między nieparzystymi $\{2j-1\}$ i parzystymi $\{2j\}$ indeksami skali, dla wszystkich $\{n\}$ przypadków. Liczba par $\{p\}$ równa jest połowie pozycji w kwestionariuszu, a X_{max} to waga przypisana do górnego kresu skali, tutaj = 3. Miara ta służy do oceny spójności wypowiedzi respondentów oraz do oceny trafności treściowej bodźców-stwierdzeń. Jej wartość nie powinna być mniejsza niż 0,5.

Miary dystansowe z dwóch pomiarów

Do wyznaczania statystyk dystansowych wykorzystuje się powyższe parametry chwilowe, oznaczone dodatkowo znakami *prim* ' i *bis* ", pochodzące odpowiednio z pierwszego i drugiego pomiaru. Dla wyróżnienia symboli miar porównawczych dodano nad literami znaki diakrytyczne: wektorów $(\vec{A}, \vec{F})$ na oznaczenie zbiorów wieloelementowych bądź „daszków” dla parametrów będących pojedynczą liczbą $(\pm\hat{b}, \pm\hat{c}_p, \pm\hat{d}, -\hat{e}, +\hat{e}, \hat{e}, \pm\hat{e}, \pm\hat{i}, \pm\hat{r}, \pm\hat{t})^{[1]}$.

Zmiany *Alterations* $\vec{A}$

$$\vec{A} = \{(\pm H_1; \pm M_1; \pm L_1); (\pm H_2; \pm M_2; \pm L_2); \dots; (\pm H_k; \pm M_k; \pm L_k)\} \quad (10)$$

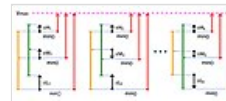


$$\text{gdzie } \pm H = H'' - H' \quad \pm M = M'' - M' \quad \pm L = L'' - L'$$

Zmiany $\vec{A}$ to zbiór przesunięć w czasie trójelementowych parametrów HML (wzór 7), uzyskanych jako wyniki zbiorcze $\{k\}$ prób badawczych (grup). Różnice między poziomami na końcu i na początku badanego procesu służą do różnicowej analizy zmian jakie zaszły w porównywanych grupach. Analiza samych różnic nie wystarcza, trzeba je oszacowywać w odniesieniu do poziomów średnich M' na wejściu ($\rightarrow$ rys.), dlatego łatwiejsze do eksploracji są znormalizowane miary trendu i zawłości.

Trend *Trend* $\pm\hat{t}$

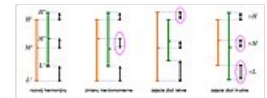
$$\pm\hat{t} = \frac{(\pm H) + (\pm M) + (\pm L)}{3V_{max} - \min(H'; H'') - \min(M'; M'') - \min(L'; L'')} \quad (11)$$



Znormalizowana miara trendu $\pm\hat{t}$ wyznaczana jest z sumy trzech przesunięć na poziomach HML odniesionych do potencjalnie możliwych zmian względem kresu skali ocen $V_{max} = 5,5$. Dostępną przestrzeń do zmian wyznaczają funkcje $\min()$, wybierające niższe wartości z par poziomów progowych osobno H, M i L z obu pomiarów ($\rightarrow$ rys.). Wzorem tym wyznacza się parametry trendu oddzielnie dla każdej z $\{k\}$ prób badawczych. W ewaluacji procesów edukacyjnych (bez wskaźników przyrostu wiedzy) wartością oczekiwaną tendencji wyników ogólnych jest +0,06 (zmiana o 6%). Trend dodatni < 3% to rezultaty niskie, a > 10% to wysokie.

Zawłóść *Intricacy* $\pm\hat{i}$

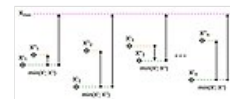
$$\pm\hat{i} = \frac{(\pm H) + (\pm M)}{2V_{max} - \min(H'; H'') - \min(M'; M'')} - \frac{(\pm L) + (\pm M)}{2V_{max} - \min(L'; L'') - \min(M'; M'')} \quad (12)$$



Zawłóść $\pm\hat{i}$ to miara nierównomierności przemian zachodzących w górnych H i dolnych L frakcjach rezultatów lepszych i gorszych od średnich M ($\rightarrow$ rys.), normalizowana względem zmian potencjalnie możliwych na skali ocen (por. trend). Liczona jest osobno dla każdej grupy i służy do oceny trafności oddziaływań. W procesach edukacyjnych wartością oczekiwaną jest 0,02 (oddziaływanie dopasowane do możliwości), > 0,08 to program zbyt ambitny, a < 0 zbyt łatwy.

Fluktuacja *Fluctuation* $\vec{F}$

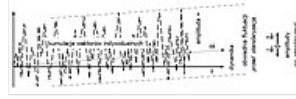
$$\vec{F} = \left\{ \frac{X''_1 - X'_1}{X_{max} - \min(X'_1; X''_1)}; \frac{X''_2 - X'_2}{X_{max} - \min(X'_2; X''_2)}; \dots; \frac{X''_k - X'_k}{X_{max} - \min(X'_k; X''_k)} \right\} \quad (13)$$



Fluktuacja $\vec{F}$ to podzbiór $\{k\}$ elementarnych wahań wypowiedzi respondentów, zawierający wyłącznie niezerowe, obliczalne ilorazy, tj. bez przypadków $X' = X''$. Wahania to różnice między wskaźnikami z pomiaru drugiego X'' i pierwszego X' odniesione do potencjalnie możliwych przesunięć względem kresu skali X_{max} . Przestrzeń do przesunięć wyznacza funkcja minimum wybierająca niższy wskaźnik, lecz tutaj tylko z obliczalnych przypadków $\min(X'; X'') \neq X_{max}$. Elementy tego podzbioru potrzebne są do obliczenia dynamiki przemian metodą wektorowej analizy fluktuacji ($\rightarrow$ rys.).

Dynamika *Dynamics* $\pm \hat{d}$

$$\pm \hat{d} = \frac{1}{k} \sum_{i: \{X'_i \neq X''_i\}}^k \frac{X''_i - X'_i}{X_{max} - \min(X'_i; X''_i)} \quad (14)$$



Dynamika $\pm \hat{d}$ to zbiorcza miara rezultatów badań, liczona jako średnia arytmetyczna z $\{k\}$ obliczalnych, niezerowych elementów fluktuacji wypowiedzi (wzór 13). Jej wartość empiryczna jest zależna od jakości oddziaływań, od rodzaju i czasu trwania procesu oraz od właściwości skali pomiarowej i nade wszystko od właściwości tego, co jest badane (np. od względnej stałości postaw i cech wolicjonalnych). W ogólności – oczekuje się dynamiki przemian rzędu kilku do kilkunastu procent.

Spadek *Decrease* $-\hat{e}$ Przyrost *Increase* $+\hat{e}$

$$-\hat{e} = \frac{\sum_{i: \{X''_i < X'_i\}} (X''_i - X'_i)}{\sum_{i: \{X''_i \leq X'_i\}} (X_{max} - X''_i)} \quad +\hat{e} = \frac{\sum_{j: \{X''_j > X'_j\}} (X''_j - X'_j)}{\sum_{j: \{X''_j \geq X'_j\}} (X_{max} - X'_j)} \quad (15)$$

W analizie dwuważonej oblicza się osobno amplitudy przemian negatywnych $-\hat{e}$ i pozytywnych $+\hat{e}$. Miary te są precyzyjniejsze, gdyż ich wartości są redukowane przez przypadki braku zmian mimo istnienia przestrzeni do zmian ($X' = X''$) $< X_{max}$. W licznikach wzorów sumuje się odpowiednio przypadki $\{l\}$ spadków i $\{h\}$ przyrostów, a w mianownikach sumuje wektory potencjalnych zmian dla tych samych przypadków oraz dodatkowo dla $\{m\}$ przypadków braku zmian.

Efekt *Effect* $\pm \hat{e}$

$$\pm \hat{e} = |+\hat{e}| - |-\hat{e}| \quad (16)$$

Znormalizowana miara efektywności przemian $\pm \hat{e}$ zachodzących w badanym procesie obliczana jest z różnicy modułów przyrostu $|+\hat{e}|$ i spadku $|-\hat{e}|$ ze wzorów 15. W *Ewaluacji splotowej* miękkich efektów procesów edukacyjnych (tj. bez wskaźników przyrostu wiedzy) wartością oczekiwaną dla wyników ogólnych jest **+0,06**. W badaniach społeczno-ekonomicznych jest to miara skuteczności interwencji.

Zasięg *Extent* $\hat{e}$

$$\hat{e} = |+\hat{e}| + |-\hat{e}| \quad (17)$$

Zasięg $\hat{e}$ to suma modułów przyrostu $|+\hat{e}|$ i spadku $|-\hat{e}|$ ze wzorów 15. Ta obwiednia fluktuacji wypowiedzi respondentów interpretowana jest jako miara siły oddziaływań na badane grupy w procesach edukacyjnych lub miara wyrazistości wpływu interwencji w badaniach społeczno-ekonomicznych.

Bilans *Balance* $\pm \hat{b}$

$$\pm \hat{b} = \frac{|-\hat{b}'|}{d'} - \frac{|-\hat{b}''|}{d''} \quad (18)$$

Bilans $\pm \hat{b}$ to miara równoważenia opinii respondentów, będąca różnicą modułów rozrzewów $|-\hat{b}'|$ i $|-\hat{b}''|$ w obu pomiarach (wzór 3), normalizowanych względem rozrzutów d' i d'' (wzór 5). Dla wyników ogólnych statystyka ta jest bliska zeru. Ewentualna wartość ujemna jest zjawiskiem negatywnym, wymagającym analiz.

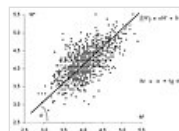
Zbieżność *Convergence* $\pm \hat{c}_p$

$$\pm \hat{c}_p = c''_p - c'_p \quad (19)$$

Zbieżność $\pm \hat{c}_p$ to miara redukcji niespójności wypowiedzi respondentów, obliczana z różnicy między spójnością w drugim $\hat{c}''_p$ i pierwszym $\hat{c}'_p$ pomiarze (wzór 9). Zazwyczaj w pomiarze końcowym spójność jest tylko nieco większa od początkowej, dlatego w praktyce zbieżność przyjmuje niewielkie wartości z przedziału **[0; 0,1]**. Ewentualna wartość ujemna świadczy o niespójności, co wymaga rozpoznania przyczyn.

Korelat *Correlate* $\pm \hat{r}$

$$\pm \hat{r} = \frac{\sum_{i=1}^n \{(X'_i - Q')(X''_i - Q'')\}}{\sqrt{\sum_{i=1}^n (X'_i - Q')^2 \cdot \sum_{i=1}^n (X''_i - Q'')^2}} \quad (20)$$



Standaryzowaną miarę zależności pomiędzy rezultatami pierwszego i drugiego pomiaru nazwano tu *korelatem* $\pm\hat{r}$. Wykorzystano współczynnik korelacji Pearsona, przy czym ma on tutaj wyrazisty sens autokorelacji między przesuniętymi w czasie, dwukrotnymi wypowiedziami tych samych respondentów. Jest to iloraz kowariancji i średniej geometrycznej z wariancji wskaźników X'_i i X''_i odniesionych do średnich Q' i Q'' odpowiednio przed ' i po " oddziaływaniu. Dla zmiennych ogólnych M' i M'' korelat wykazuje zależność monotoniczną ($\rightarrow$ rys.). Wartość przeciętna dla populacji to **0,6**, wartości **<0,5** to słaba, a **>0,75** to silna autokorelacja.

Wyznaczniki norm i jakości pomiarowej

Wyznaczniki pełnią szczególne role. Baza i Norma służą jako poziomy odniesień dla wyników lokalnych. Selektowność umożliwia testowanie istotności różnic między poziomami średnimi. Konkluzyjność służy do weryfikacji wskaźników i doskonalenia bodźców-stwierdzeń. Trafność określa poziom wiarygodności wyników. Rzetelność oznacza jakość danych empirycznych^[1]. Dla odróżnienia od innych miar, wyznaczniki uzupełnia diakrytyczny haczek $\checkmark$.

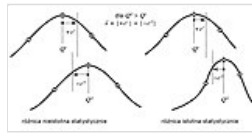
Baza *Base* B' **Norma** *Norm* N''

$$B' = \frac{1}{k} \sum_{i=1}^k M'_i \quad N'' = \frac{1}{k} \sum_{i=1}^k M''_i \quad (21)$$

Miary służące do kalibracji danej skali i wytyczające standardy ewaluatywne to Baza B' (stan oczekiwany na wejściu) oraz Norma N'' (stan oczekiwany na wyjściu). Obliczane są jako średnie ze średnich poziomów M' i M'' (ze wzoru 7) osiąganych przez każdą z $\{k\}$ elementarnych grup badawczych, tj. wielu klas z licznej próby minimum kilkuset przypadków, odpowiednio w dwukrotnych pomiarach tych samych respondentów przed ' i po ". W praktyce ewaluacji procesów edukacyjnych wartość Normy N'' dla wyniku ogólnego powinna mieścić się w przedziale **[3,7; 4,0]**.

Selektowność *Selectivity* $\check{s}$

$$\check{s} = \begin{cases} U(\alpha) \cdot \left(\frac{|+d'| + |-d''|}{\sqrt{n}} \right) & \text{dla } Q' < Q'' \\ U(\alpha) \cdot \left(\frac{|-d'| + |+d''|}{\sqrt{n}} \right) & \text{dla } Q' > Q'' \end{cases} \quad (22)$$



Miara służąca do testowania istotności różnic między średnimi rezultatami z dwóch pomiarów tych samych $\{n\}$ respondentów, oparta jest na rozstępie przedziałów ufności realnej (wzory 8). Zależnie od tego, która ze średnich Q' i Q'' jest większa, sumuje się z odpowiednich stron zakresy tolerancji ($\rightarrow$ rys.). Różnica między średnimi większa od selektowności $\check{s}$ jest istotna statystycznie na poziomie istotności α (alfa). Wartości współczynników $U(\alpha)$ podano w opisie wzorów 8.

$$\check{s}_{ab} = \begin{cases} U(\alpha) \cdot \left(\frac{|+d_a|}{\sqrt{n_a}} + \frac{|-d_b|}{\sqrt{n_b}} \right) & \text{dla } Q_a < Q_b \\ U(\alpha) \cdot \left(\frac{|-d_a|}{\sqrt{n_a}} + \frac{|+d_b|}{\sqrt{n_b}} \right) & \text{dla } Q_a > Q_b \end{cases} \quad (23)$$

Zmodyfikowany wzór na selektowność $\check{s}_{ab}$ stosuje się do testowania różnic między rezultatami dwóch różnych grup $\{a\}$ oraz $\{b\}$ o liczebnościach $\{n_a\}$ i $\{n_b\}$. Interpretacja wzorów 22 i 23 jest identyczna, przy czym miary te można liczyć z ufności realnej (wzory 8) albo wprost z odchyżeń realnych (wzory 4).

Konkluzyjność *Conclusiveness* $\check{c}$

$$\check{c} = \frac{\frac{1}{2}(d' + d'') - |-b'| - |-b''| - |\pm s'| \cdot |\pm s''|}{X_{max}} \quad (24)$$

Konkluzyjność $\check{c}$ to wyznacznik jakości i mocy wnioskotwórczej elementarnych wskaźników. Jest kluczowym selektorem doboru i optymalizacji bodźców-stwierdzeń (mikrotez) w tworzeniu i pilotażowej weryfikacji kwestionariusza. *Skalę dwuważonych ocen* tworzą mikrotezy, wobec których wypowiedzi są znacznie zróżnicowane (duże rozrzuty d' i d'' ze wzoru 5), ale przy możliwie niskiej niezgodności co do standardów ewaluatywnych (małe rozziwy $-b'$ i $-b''$ ze wzoru 3). Parametr ten jest redukowany przy dużych skośnościach rozkładów $\pm s'$ i $\pm s''$ (wzór 6) oraz normalizowany względem kresu skali X_{max} . Dobre jakościowo elementarne wskaźniki powinny mieć konkluzyjność $\check{c} \geq 0,7$.

Trafność *Aptness* $\check{a}$

$$\check{a} = 1 - |(\pm\hat{t}) - (\pm\hat{e})| \quad (25)$$

Wyznacznik trafności pomiarowej powstaje z porównania rezultatów dwóch metod: różnicowej analizy zmian $\pm \hat{t}$ (wzór 11) z wektorową analizą fluktuacji $\pm \hat{e}$ (wzór 16). Trafność $\hat{a}$ jest kwantyfikatorem kumulacyjnym, toteż jego obliczanie i interpretowanie zasadne jest tylko na wyższych poziomach agregacji indeksów w skali, a więc co najmniej dla czynników, a głównie dla wyników połówkowych i ogólnych. Wyznacznik ten dąży do jedności i w praktyce powinien osiągać wartości $\geq 0,95$.

Rzetelność *Reliability* $\hat{r}$

$$\hat{r} = \frac{p}{p-1} \left(1 - \frac{\sum_{k=1}^p \sum_{j=1}^n \left(X_{jk} - \frac{1}{n} \sum_{j=1}^n X_{jk} \right)^2}{\sum_{j=1}^n \left(\sum_{k=1}^p X_{jk} - \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^p X_{jk} \right)^2} \right) \quad (26)$$

Kohärenz – spójność p zmiennych

	4	5	5	...	4
3	3	4	3	...	3
1	2	2	1	...	1
4	4	3	2	...	3
...	...	...	...	...	...
2	2	1	1	...	2

Rzetelność $\hat{r}$ – jako wyznacznik jakości macierzy danych empirycznych – określa dla liczby $\{p\}$ pozycji w kwestionariuszu spójność wszystkich wskaźników X_{jk} indeksujących zmienną ogólną, przy jednoczesnym zróżnicowaniu przypadków, wynikającym z osobniczych cech $\{n\}$ respondentów ($\rightarrow$ rys.). Jest to alfa Cronbacha, przy czym wskaźniki do *Skali dwuważonych ocen* dobierane są ze względu na ich wysoką konkluzyjność i wieloaspektowość, dlatego spójność może tu być niższa. W praktyce wystarczy gdy osiąga wartości $\geq 0,75$.

Założenia metodologiczne i metrologiczne

PATRZ *Ewaluacja splotowa* oraz *Skala dwuważonych ocen*.

Przypisy

1. Stanisław Ubermanowicz: *Ewaluacja splotowa InfoKultury. Skala dwuważonych ocen* (<https://repozytorium.amu.edu.pl/bitstream/10593/946/3/EwaluacjaSplotowa.pdf>), Poznań: Wydawnictwo Naukowe UAM, 2005.

Bibliografia

- Frankfort-Nachmias Chava, Nachmias David: *Metody badawcze w naukach społecznych*. Poznań: Wydawnictwo Zysk i S-ka, 2001. ISBN 83-7150-702-X.
- Józwiak Janina, Podgórski Jarosław: *Statystyka od podstaw*. Warszawa: Państwowe Wydawnictwo Ekonomiczne, 1994. ISBN 83-208-0971-1.
- Ubermanowicz Stanisław: *Ewaluacja splotowa InfoKultury. Skala dwuważonych ocen* (<https://repozytorium.amu.edu.pl/bitstream/10593/946/3/EwaluacjaSplotowa.pdf>), Poznań: Wydawnictwo Naukowe UAM, 2005. ISBN 83-232-1608-8.

Źródło: „[https://pl.wikipedia.org/w/index.php?title=Miały dwuważonych ocen&oldid=64759123](https://pl.wikipedia.org/w/index.php?title=Mia%C5%82y_dwuw%C4%85żonych_ocen&oldid=64759123)”

Tę stronę ostatnio edytowano 10 paź 2021, 10:21. Tekst udostępniany na licencji Creative Commons: uznanie autorstwa, na tych samych warunkach (<http://creativecommons.org/licenses/by-sa/3.0/deed.pl>), z możliwością obowiązywania dodatkowych ograniczeń. Zobacz szczegółowe informacje o warunkach korzystania (http://foundation.wikimedia.org/wiki/Warunki_korzystania).