WPROWADZENIE

Rozpatruje się odpowiedzi testu składającego się z m pytań jednokrotnego wyboru, udzielone przez n osób. Ocena każdej odpowiedzi x_{ij}, udzielonej przez i-tą osobę na j-te pytanie, jest dwuwartościowa:

$$x_{ij} = \begin{cases} 1 & gdy \quad odpowiedź\ prawidlowa \\ 0 & gdy \quad odpowiedź\ błędna \end{cases} \qquad i=1,..,n \quad j=1,...,m$$

Sumaryczny wynik uzyskany przez i-tą osobę

$$\mathbf{y_i} = \sum_{\mathbf{j=1}}^{\mathbf{m}} \mathbf{x_{ij}} \quad i=1,..,n$$

Liczba prawidłowych odpowiedzi dla j-tego pytania

$$\mathbf{x_j} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{x_{ij}} \quad j=1,..,m$$

ŁATWOŚĆ PYTAŃ

Łatwość j-tego pytania

$$p_j = \frac{n_j}{n} = \frac{\sum_{i=1}^{n} x_{ij}}{n} \quad j=1,...,m$$

gdzie:

n_j – liczba osób, którzy udzielili poprawnej odpowiedzi na j-te pytanie

n – liczba udzielających odpowiedzi

Ocena pytań na podstawie ich łatwości

	Dolna granica	Górna granica
Bardzo trudne	0	0,19
Trudne	0,20	0,49
Średnie / umiarkowanie trudne	0,5	0,69
Łatwe	0,7	0,89
Bardzo łatwe	0,9	1

Ze względu na wywołanie motywacji uczenia najbardziej pożądane są pytania umiarkowanie trudne i łatwe.

Na podstawie łatwości pytania określa się jego trudność

$$q_j = 1 - p_j$$

WSPÓŁCZYNNIK KORELACJI PUNKTOWO-DWUSERYJNEJ

Współczynnik wykorzystywany jest do oceny zależności korelacji pomiędzy odpowiedziami pytania testowego oraz wynikiem sumarycznym testu w wypadku, gdy ocena odpowiedzi jest dwuwartościowa.

Współczynnik dla j-tego pytania oblicza się z następującego wzoru:

$$r_j^{PD} = \frac{\bar{y}_j^p - \bar{y}_j^q}{s_y}\sqrt{p_j q_j}$$

gdzie:

$\bar{y}_j^p$ – średnia arytmetyczna sumarycznych wyników testu, uzyskanych przez osoby, które udzieliły poprawnej odpowiedzi na j-te pytanie

$\bar{y}_j^q$ – średnia arytmetyczna sumarycznych wyników testu, uzyskanych przez osoby, które udzieliły błędnej odpowiedzi na j-te pytanie

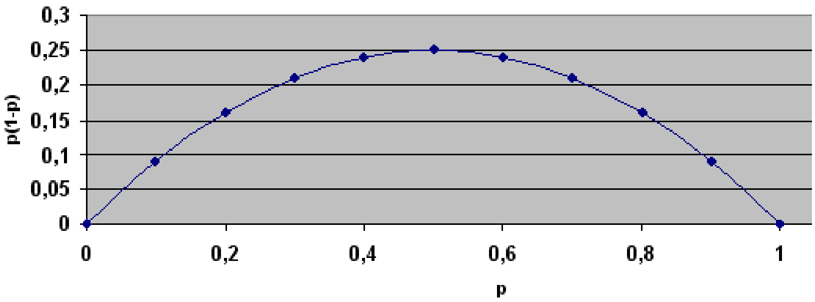
s_y – odchylenie standardowe sumarycznego wyniku testu

p_j – łatwość rozpatrywanego pytania

q_j – trudność rozpatrywanego pytania

Współczynnik przyjmuje wartości z przedziału [-1 ; 1]. Wartości współczynnika są dodatnie, jeżeli $\bar{y}_j^p > \bar{y}_j^q$, to znaczy kiedy średnia arytmetyczna sumarycznych wyników uzyskanych przez osoby, które udzieliły poprawnej odpowiedzi na j-te pytanie jest większa od średniej arytmetycznej sumarycznych wyników uzyskanych przez osoby, które udzieliły błędnej odpowiedzi na j-te pytanie.

Wartości współczynnika są tym większe, im większa jest wartość iloczynu $p_j q_j$ jest większa



Największa wartość iloczynu jest dla łatwości 0,5.

Dla oceny istotności współczynnika korelacji ρ weryfikuje się hipotezy:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

W tym celu wykorzystuje się statystykę

$$t_{PD} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

gdzie:

r – współczynnik korelacji punktowo-dwuseryjnej

n – liczebność próby

Na podstawie współczynnika istotności α na podstawie tablic rozkładu Studenta wyznacza się obszar krytyczny

$$(-\infty; -t_{kr} > \cup < t_{kr}; \infty)$$

gdzie t_{kr} jest dla 1 - $\frac{\alpha}{2}$ oraz n – 2 stopni swobody.

Podejmujemy decyzję o zachowaniu lub odrzuceniu hipotezy zerowej:

- Jeżeli statystyka testowa należy do obszaru krytycznego, to odrzucamy hipotezę zerową na rzecz hipotezy alternatywnej – korelacja jest istotna.
- Jeżeli statystyka testowa nie należy do obszaru krytycznego, to nie odrzucamy hipotezy zerowej – korelacja nie jest istotna.

Współczynnik korelacji punktowo-dwuseryjnej jest szczególnym przypadkiem korelacji Pearsona wykorzystywanego do oszacowania siły związku pomiędzy wynikami pojedynczego pytania, np. j-tego, przyjmującego więcej wartości niż dwie oraz wynikiem sumarycznym y_i:

$$r_j^P = \frac{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{x_{ij}}-\bar{y}_i)(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{x_{ij}}-\bar{y}_i)^2\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}} = \frac{\sum_{i=1}^{n}(x_{ij}-\bar{y}_i)(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(y_{ik}-\bar{y}_i)^2\sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (1)$$

gdzie:

x_{ij} – ocena odpowiedzi udzielonej przez i-tą osobę na j-te pytanie

$\bar{\mathrm{y}}_\mathrm{j} = \frac{1}{n}\sum_{i=1}^{n} x_{\mathrm{ij}}$ – średnia z ocen odpowiedzi na pytanie j-te udzielonych przez n osób,

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ – średnia wyników sumarycznych uzyskanych przez n osób.

Dla wykazania tego załóżmy,

Zakładamy teraz, że ocena odpowiedzi jest dwuwartościowa oraz że n_{j1} osób udzieliło poprawnej odpowiedzi na rozpatrywane pytanie, a n_{j2} = n – n_{i1} odpowiedziało błędnie. Przy tym założeniu średni wynik pytania jest równy:

$$\bar{\mathrm{y}}_\mathrm{j} = \frac{\mathrm{n_{j1}}}{\mathrm{n}},$$

a wskaźniki zadania: łatwość zadania – p_j oraz jego trudność – q_j mają wartości:

$$\mathrm{p_j} = \frac{\sum_{i=1}^{n}\mathrm{x_{ij}}}{\mathrm{n}} = \frac{\mathrm{n_{j1}}}{\mathrm{n}} \qquad \mathrm{q_j} = 1 - \mathrm{p_j} = \frac{\mathrm{n_{j2}}}{\mathrm{n}},$$

czyli są określone przez średni wynik zadania:

$$\bar{\mathrm{y}}_\mathrm{j} = \mathrm{p_j} \qquad \mathrm{q_j} = 1 - \bar{\mathrm{y}}_\mathrm{j}.$$

Dla uproszczenia zapisu przyjmujemy, że wyniki są uporządkowane, tzn. osoby oznaczone numerami [1 ; n_{j1}] udzieliły poprawnej odpowiedzi, a osoby o numerach [n_{j1}+1 ; n] udzieliły odpowiedzi błędnej.

Przekształcamy licznik wzoru (1):

$$\mathrm{L} = \sum_{i=1}^{n}(x_{ij}-\bar{y}_i)(y_i-\bar{y}) = \sum_{i=1}^{n_{j1}}(1-\frac{\mathrm{n_{j1}}}{\mathrm{n}})(y_i-\bar{y}) + \sum_{i=n_{j1}+1}^{n}(0-\frac{\mathrm{n_{j1}}}{\mathrm{n}})(y_i-\bar{y}) =$$

$$= \sum_{i=1}^{n_{j1}}(y_i-\bar{y}) - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\sum_{i=1}^{n_{j1}}(y_i-\bar{y}) - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\sum_{i=n_{j1}+1}^{n}(y_i-\bar{y}) =$$

$$= \sum_{i=1}^{n_{j1}}(y_i - \bar{y}) - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\left[\sum_{i=1}^{n_{j1}}(y_i - \bar{y}) + \sum_{i=n_{j1}+1}^{n}(y_i - \bar{y})\right] = \sum_{i=1}^{n_{j1}}(y_i - \bar{y}) - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\sum_{i=1}^{n}(y_i - \bar{y}) =$$

$$= \sum_{i=1}^{n_{j1}} y_i - n_{j1}\ \bar{y} - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\sum_{i=1}^{n} y_i + \frac{\mathrm{n_{j1}}}{\mathrm{n}}\mathrm{n}\ \bar{y} = \sum_{i=1}^{n_{j1}} y_i - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\sum_{i=1}^{n} y_i \quad (2)$$

Przekształcimy teraz podobnie pierwszy element mianownika (pod pierwiastkiem):

$$\mathrm{M_1} = \sum_{i=1}^{n}(y_{ik} - \bar{y}_i)^2 = \sum_{i=1}^{n_{j1}}\left(1 - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\right)^2 + \sum_{i=n_{j1}+1}^{n}\left(0 - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\right)^2 = \sum_{i=1}^{n_{j1}}\left(\frac{n - \mathrm{n_{j1}}}{\mathrm{n}}\right)^2 + \sum_{i=n_{j1}+1}^{n}\left(\frac{\mathrm{n_{j1}}}{\mathrm{n}}\right)^2 =$$

$$= \sum_{i=1}^{n_{j1}}\left(\frac{\mathrm{n_{j2}}}{\mathrm{n}}\right)^2 + \sum_{i=n_{j1}+1}^{n}\left(\frac{\mathrm{n_{j1}}}{\mathrm{n}}\right)^2 = \mathrm{n_{j1}}\left(\frac{\mathrm{n_{j2}}}{\mathrm{n}}\right)^2 + \mathrm{n_{j2}}\left(\frac{\mathrm{n_{j1}}}{\mathrm{n}}\right)^2 = \frac{\mathrm{n_{j1}}\mathrm{n_{j2}}}{n}\frac{\mathrm{n_{j1}} + \mathrm{n_{j2}}}{n} = \frac{\mathrm{n_{j1}}\mathrm{n_{j2}}}{n} \quad (3)$$

Wykorzystując powyższe dwa wyniki, współczynnik korelacji (1) można przedstawić w postaci:

$$r_j^P = \frac{\sum_{i=1}^{n_{j1}} y_i - \frac{\mathrm{n_{j1}}}{\mathrm{n}}\sum_{i=1}^{n} y_i}{\sqrt{\frac{\mathrm{n_{j1}}\mathrm{n_{j2}}}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Po pomnożeniu przez $1/\mathrm{n_{j1}}$ licznika i mianownika powyższego wzoru otrzymujemy:

$$r_j^P = \frac{\frac{1}{n_{j1}}\sum_{i=1}^{n_{j1}} y_i - \frac{1}{\mathrm{n}}\sum_{i=1}^{n} y_i}{\sqrt{\frac{\mathrm{n_{j2}}\,1}{n_{j1}\,n}\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{\frac{1}{n_{j1}}\sum_{i=1}^{n_{j1}} y_i - \frac{1}{\mathrm{n}}\sum_{i=1}^{n} y_i}{\sqrt{\frac{\frac{\mathrm{n_{j2}}}{n}}{\frac{n_{j1}}{n}}\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (4)$$

Uwzględniając w powyższej zależności wcześniejsze oznaczenia: łatwości p_i, trudności zadania q_i, średniej wyników sumarycznych $\bar{y}$, średniej wyniku sumarycznego osób, które udzieliły poprawnej odpowiedzi na i-te pytanie $\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}}$ oraz odchylenia standardowego sumarycznego wyniku dla wszystkich osób $\mathrm{s_y}$ (*pierwiastek z wariancji*) otrzymujemy pierwszą postać ***współczynnika korelacji punktowo-dwuseryjnej***:

$$r_j^P = \frac{\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - \bar{y}}{\sqrt{\frac{q_i}{p_i}s_y^2}} \quad (5)$$

W dalszych rozważaniach wykorzystamy wzór:

$$\bar{y} = p_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} + q_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{q}}$$

którego prawdziwość wynika z poniższej zależności:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n_{j1}} y_i + \frac{1}{n}\sum_{i=n_{j1}+1}^{n} y_i = \frac{n_{j1}}{n_{j1}}\frac{1}{n}\sum_{i=1}^{n_{j1}} y_i + \frac{n_{j2}}{n_{j2}}\frac{1}{n}\sum_{n_{j1}+1}^{n} y_k$$

$$= \frac{N_{i1}}{N}\frac{1}{N_{i1}}\sum_{k=1}^{N_{i1}} z_k - \frac{N_{i2}}{N}\frac{1}{N_{i2}}\sum_{k=N_{i1}+1}^{N} z_k = p_i\bar{z}_{ip} - q_i\bar{z}_{iq}$$

Tak więc

$$r_j^P = \frac{\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - p_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - q_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{q}}}{\sqrt{\frac{q_i}{p_i}s_y^2}} = \frac{\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - p_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - q_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{q}}}{\sqrt{\frac{q_i}{p_i}s_y^2}} = \frac{\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}}(1 - p_j) - q_j\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{q}}}{\sqrt{\frac{q_i}{p_i}s_y^2}} = \frac{q_j\left(\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - \bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{q}}\right)}{\sqrt{\frac{q_i}{p_i}s_y^2}} = \frac{\bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{p}} - \bar{\mathrm{y}}_{\mathrm{j}}^{\mathrm{q}}}{\mathrm{s_y}}\sqrt{\mathrm{p_j q_j}}$$

Jest to postać *współczynnika korelacji punktowo-dwuseryjnej*, podana na wstępie.

OCENA RZETELNOŚCI TESTU

Rzetelność testu jest miarą jego wewnętrznej spójności (zgodności), stopniem, oceną w jakim zakresie wszystkie pytania testu korelują ze sobą, mierzą ten sam obszar umiejętności uczniów. Na rzetelność testu wpływa jakość pytań testowych, w tym także ich wyrównanie pod względem łatwości i długości skali wyników.

Do oceny rzetelności testu można wykorzystywać współczynnik 20 Kudera Richardsona:

$$r_{K20} = \frac{\mathrm{m}}{\mathrm{m}-1}\left(1 - \frac{\sum_{j=1}^{m} p_i q_i}{s_y^2}\right)$$

gdzie:

m – liczba zadań

s_y^2 – wariancja sumarycznych wyników testu $Y = \sum_{j=1}^{m} x_i$ (x_i – ocena odpowiedzi na j-te pytanie),

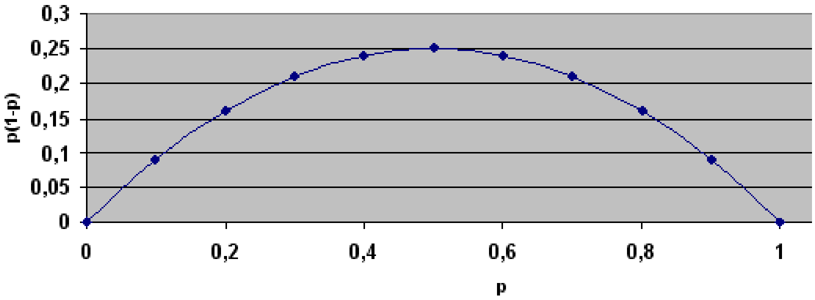
p_j – łatwość j-tego pytania

q_j – trudność rozpatrywanego pytania, równa 1 - p_j

Współczynnik ten wykorzystywany jest w przypadku, gdy ocena odpowiedzi jest dwuwartościowa (tak jak w przypadku współczynnika korelacji punktowo-dwuseryjnej).

Współczynnik 20 Kudera-Richardsona przyjmuje wartości z przedziału [0,+1], przy czym:

- Duża wartość współczynnika (bliska 1) występuje wtedy, gdy pytania są albo bardzo łatwe, albo bardzo trudne – bo wtedy $\sum_{i=1}^{M} p_i \cdot q_i \approx 0$.
- Mała wartość współczynnika (bliska zeru) występuje wtedy, gdy wielkość $\sum_{i=1}^{M} p_i \cdot q_i$ jest duża, tzn. jeżeli łatwość pytań jest na poziomie 50% .



Wykres wartości funkcji p(1-p)

Przyjmuje się, że test posiada wystarczającą rzetelność, jeżeli współczynnik 20 Kudera-Richardsona nie jest mniejszy niż 0,70.

Jeżeli wartość współczynnika alfa Cronbacha lub współczynnika 20 Kudera-Richardsona jest mała, to należy przeprowadzić analizę czynnikową.

Ocena testu na podstawie współczynnika KR_{20}

	Dolna granica	Górna granica	Wniosek
Test nierzetelny	0	0,4	Test nie nadaje się do analiz
Test rzetelny	0,41	0,8	Test pozwala analizować wyniki, ale nie można wystawiać ocen
Test bardzo rzetelny	0,81	0,9	Test daje rzetelny obraz wiedzy i umiejętności

Współczynnik 20 Kudera-Richardsona jest szczególnym przypadkiem ***współczynnika α alfa Cronbacha*** (współczynnika rzetelności alfa Cronbacha):

$$\alpha = \frac{m}{m-1}\left(1 - \frac{\sum_{j=1}^{m} s_j^2}{s_y^2}\right)$$

gdzie: m – liczba pytań w teście,

s_j^2 – wariancja wyników j-tego pytania,

s_y^2 – wariancja wyniku sumarycznego

Wynika to z zależności określającej wariancję:

$$s_j^2 = (1-p_j)^2 \cdot p_j + (0-p_j)^2 \cdot (1-p_j) = (1-p_j)^2 \cdot p_j + p_j^2 \cdot (1-p_j) = 1-p_j)[(1-p_j) \cdot p_j + p_j^2] =$$
$$= (1-p_j)\,(p_j - p_j^2 + p_j^2) = (1-p_j) \cdot p_j = p_j \cdot q_j$$

w której uwzględniono, że średnia wyników j-pytania

$$\bar{x}_{j=} = 1 \cdot pj + 0 \cdot (1\text{-}p_j) = p_j$$

Postać współczynnika alfa Cronbacha można skomentować uwzględniając, że wyrażenie na wariancję wyniku sumarycznego testu $Y = \sum_{j=1}^{m} X_j$ ma postać:

$$\sigma_y^2 = D^2Y = D^2\left(\sum_{j=1}^{m} X_i\right) = E\left(\sum_{j=1}^{m} X_i - E\sum_{j=1}^{m} X_i\right)^2 = \sum_{j=1}^{m} \sigma_j^2 + \sum_{j=1}^{m} \sum_{\substack{j=1 \\ k \neq j}}^{m} \rho_{jk} \sigma_j \sigma_k$$

Udowadnia się to dla trzech pytań:

$$\sigma_y^2 = D^2Y = D^2\left(\sum_{j=1}^{3} X_j\right) = E\left[\sum_{j=1}^{3} X_j - E\sum_{j=1}^{3} X_j\right)]^2 = E(X_1 + X_2 + X_3 - EX_1 - EX_2 - EX_3)^2 =$$
$$= E[(X_1 - EX_1) + (X_2 - EX_2) + (X_3 - EX_3)]^2$$

Po wprowadzeniu oznaczenia $Q_j = X_j - EX_j$ otrzymuje się:

$$D^2Y = E(Q_1 + Q_2 + Q_3)^2 = E(Q_1^2 + Q_2^2 + Q_3^2 + 2Q_1Q_2 + 2Q_1Q_3 + 2Q_2Q_3) =$$
$$\sum_{j=1}^{3} EQ_j^2 + \sum_{j=1}^{3} \sum_{\substack{k=1 \\ k \neq j}}^{3} E(Q_jQ_k).$$

Otrzymany wynik, po uwzględnieniu oznaczeń, ma postać:

$$D^2Y = \sum_{j=1}^{3} D^2 X_j + \sum_{j=1}^{3} \sum_{\substack{k=1 \\ k \neq j}}^{3} \left[E(X_j - EX_j) \cdot (X_k - EX_k)\right],$$

Drugi składnik w powyższej równości jest kowariancją pary pytań testu (X_j, X_k)

$$\sum_{j=1}^{3}\sum_{\substack{k=1\\k\neq j}}^{3} E\left(X_j - EX_j\right)\cdot(X_k - EX_k) = \mu_{jk},$$

która określa współczynnik korelacji tych pytań $\rho_{jk} = \frac{\mu_{jk}}{\sigma j_j \sigma_k}$.

Stąd otrzymuje się ostatecznie $D^2Y = \sum_{j=1}^{3}\sigma_j^2 + \sum_{j=1}^{3}\sum_{\substack{k=1\\k\neq j}}^{3}\rho_{jk}\cdot\sigma_j\cdot\sigma_k$

cbdw

Tak więc wariancja wyniku sumarycznego testu składa się z sumy wariancji odpowiedzi na poszczególne pytania $\sum_{i=1}^{M}\sigma_i^2$ oraz wielkości zależnej od współczynników korelacji pytań ρ_{jk} i ich wariancji σ_j^2.

Tak więc współczynnik alfa Cronbacha można przedstawić w postaci (element m/(m-1) jest unormowaniem):

$$\alpha = \frac{m}{m-1}\left(1 - \frac{\sum_{j=1}^{m} s_j^2}{s_y^2}\right) = \frac{m}{m-1}\left(1 - \frac{\sum_{j=1}^{m}\sigma_j^2}{\sum_{j=1}^{m}\sigma_j^2 + \sum_{j=1}^{m}\sum_{\substack{j=1\\k\neq j}}^{m}\rho_{jk}\sigma_j\sigma_k}\right).$$

Z postaci tej wynika, że:

- Jeżeli współczynniki korelacji odpowiedzi pytań są równe 0, to $\sum_{j=1}^{m}\sum_{\substack{j=1\\k\neq j}}^{m}\rho_{jk}\sigma_j\sigma_k$ drugi element w mianowniku jest równy zeru i w konsekwencji

$$\alpha = \frac{m}{m-1}\left(1 - \frac{\sum_{j=1}^{m}\sigma_j^2}{\sum_{j=1}^{m}\sigma_j^2}\right) = \frac{m}{m-1}(1-1) = 0$$

- Jeżeli współczynniki korelacji wyników pytań są bliskie 1 oraz odchylenia standardowe wyników pytań są duże, to wyraz w mianowniku $\sum_{j=1}^{m}\sum_{\substack{j=1\\k\neq j}}^{m}\rho_{jk}\sigma_j\sigma_k$ jest duży i w konsekwencji współczynnik α Cronbacha przyjmuje wartość bliską 1.