

Statystyka oceny parametrycznej jednostek sfery badań i wdrożeń

Czesław CEMPEL¹

1. Wstęp

Wypracowany w czasach Komitetu Badań Naukowych (KBN, od 1991 -2005r, [1]) system finansowania sfery badań i wdrożeń w kraju polegał na przedmiotowym dofinansowywaniu Jednostek Badawczo Rozwojowych, stosownie do typu jednostki i jej osiągnięć. Pod pojęciem Jednostki mieszczą się **Wydziały** Szkół Wyższych oraz niezależne organizacje i instytucje typu instytutów **PAN**, instytutów przemysłowych i Ośrodków badawczo Rozwojowych (**OBR**) różnego typu. Dla pokrycia całego spektrum aktywności badawczej tych jednostek, od teologii do badań obronnościowych, utworzono w KBN 12 zespołów merytorycznych², do zadań których należała bieżąca obsługa sfery projektów i grantów różnego typu, oraz ocena jakości badań i wdrożeń w jednostkach. Ta ocena przeprowadzana była średnio raz na cztery lata (ostatnio w 2000r) i dzieliła ogół jednostek na **kategorie jakości od I do IV**, przy czym w najlepszej pierwszej kategorii mogło być maksymalnie **20%** ogółu jednostek w danym Zespole (*prawo Pareto?*).

Krótko po tym zaczęło się w KBN myślenie o utworzeniu w miejsce KBN Ministerstwa Nauki, oczywiście z nowymi zasadami pracy i współpracy ze sferą badań i wdrożeń, w tym też z nowymi (*lepszymi !*) zasadami kategoryzacji jednostek. Stało się to faktem w roku 2005, kiedy w początku roku **KBN** przekształcono w Ministerstwo Nauki i Informatyzacji, a 4 sierpnia 2005 Minister wydał zarządzenie precyzujące nowe zasady oceny i kategoryzacji **Jednostek** na podstawie ich ankiet z lat 2001- 2004 przesyłanych corocznie do Ośrodka Przetwarzania Informacji (**OPI**). Jak się wydaje w chwili obecnej, już po zebraniu nowych Ankiet i Kart Oceny Jednostek, nie są one ulepszeniem poprzednich kryteriów ocenowych stosowanych przez **KBN**, lecz zostały napisane od nowa z silnym wypukleniem strony innowacyjno wdrożeniowej³ (*najmniejsze wdrożenie jest równoważne nowej profesurze – 30pkt*).

W nowej **Karcie Oceny** Jednostek całość ich możliwej sfery aktywności podzielono na **trzy strumienie**. Pierwszy strumień aktywności ujmuje nowo pozyskane stopnie, tytuły, laboratoria, projekty badawcze i celowe, oraz nagrody. Każdemu składnikowi opisowemu tego strumienia przypisano stosowną liczbę punktów, po to by ich sumę P_I podzielić przez liczbę zatrudnionych w jednostce N otrzymując efektywność pierwszego strumienia aktywności E_I . Podobnie został sparametryzowany drugi strumień aktywności ujmujący publikacje i monografie wszelkiego typu, z naciskiem na piśmiennictwo anglojęzyczne, co dało punkty P_{II} i efektywność E_{II} . Trzeci strumień aktywności w charakterystyce jednostki, to patenty, licencje i wdrożenia efektów badawczo rozwojowych (**B+R**) zobrażony punktowo przez P_{III} i efektywność policzoną jak poprzednio E_{III} . Te trzy strumienie aktywności mają w rozporządzeniu Ministra z sierpnia 2005 różne **wagi** w ocenie różnych jednostek. Dla jednostek powiązanych z gospodarką waga wdrożeń dochodzi nawet do **5** w skali **10** stopniowej, a z wagą **1** dla wydziałów humanistycznych szkół wyższych. Ponadto nie jest pewne, który z tak stworzonych symptomów lepiej ocenia jednostkę; punkty P czy efektywności E . A może ich kombinacja? Warto to obiektywnie zbadać.

W lutym br. członkowie Rady Nauki (**RN**) otrzymali ten całościowy materiał do oceny parametrycznej jednostek przedstawionych na nośniku CD, gdzie zawarto cząstkowe punkty P_i oraz efektywności E_i dla **832** Jednostek. Znaczy to, że weryfikację i zastosowanie wag przy ocenie łącznej liczby punktów P i łącznej efektywności E dla jednostek pozostawiono weryfikatorom z **RN**. Weryfikacja ta potrwa jeszcze trochę, bo w nadesłanym materiale wykorzystano skutecznie wiele **niejednoznaczności** pytań Ankiety, łącznie z liczbą zatrudnionych N . Weryfikacja ta zmieni położenie niektórych jednostek na **skali jakości** badań i wdrożeń (*symptomy P i E*), ale jak się

¹Komisja Nauki RN; Instytut Mechaniki Stosowanej, Politechnika Poznańska; czeslaw.cempel@put.poznan.pl

² Piszący te słowa przez lata należał do **Zespołu T07 Mechanika, Budownictwo i Architektura**.

³ Jak twierdzą naukowcy w Stanford University „**dobra nauka to taka która zmienia świat, a można to zrobić jedynie wdrażając wyniki badań**, [2].

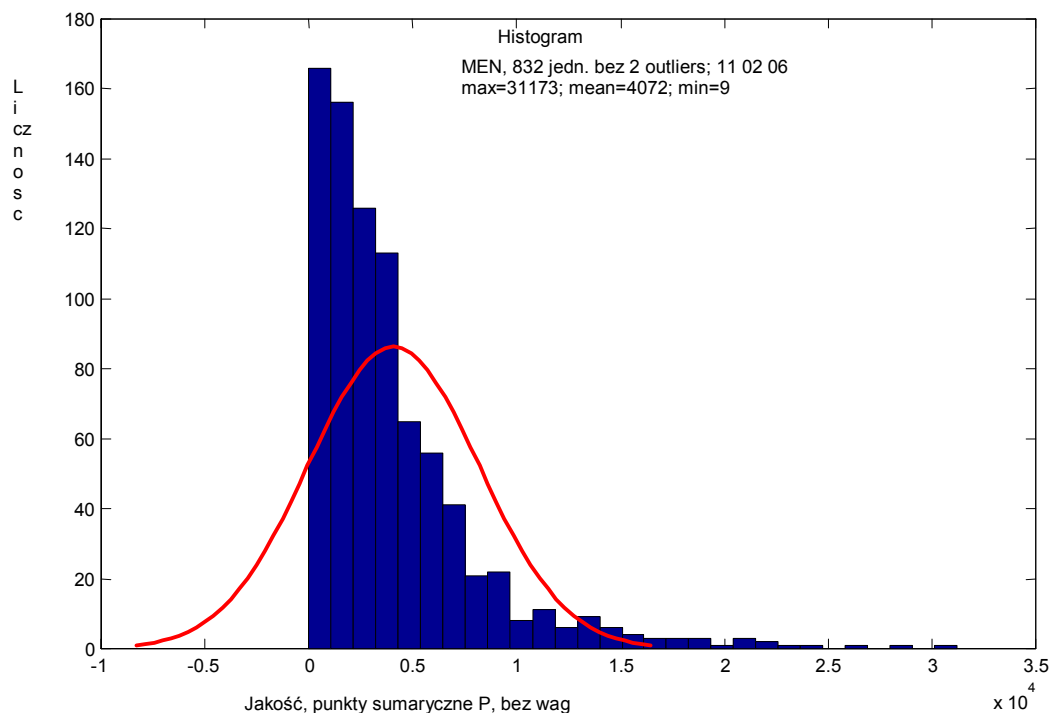
wyduje **nie zmieni** w sposób istotny własności tak licznego zbioru ocenianych jednostek. Takie założenie przyświeca poniższym rozważaniom, a ich celem jest danie przesłanek metodologicznych do zadania **kategoryzacji** jakości jednostek sfery badań i innowacji poprzez badanie własności statystycznych całego zbioru lub wyodrębnionych części.

2. Dominujący typ rozkładu ocen jakości jednostek

Jak wynika ze wstępu w bazie danych mamy generalnie dwie miary jakości jednostek (*dwa symptomy jakości*), punkty **P** oraz efektywności **E**. Ponieważ wspomniane już współczynniki wagi będą stosowane do mało licznych grup jednostek, zastosujemy w tej analizie populacji dla wszystkich jednostek wagi jednostkowe, tak jak np. wdrożenia dla wydziałów humanistycznych, czyli z **wagą 1**.

Zatem badając statystykę ocen posłużymy zwykłą się na początku sumą ocen punktowych; $P = P_I + P_{II} + P_{III}$. Z tych samych powodów podobnie postąpimy dla efektywności⁴; $E = E_I + E_{II} + E_{III}$.

Wiedząc o tym popatrzymy na zbiory jednostek przez pryzmat tych dwu miar jakości, a zamierzając ustalić dobry podział na kategorie (**a na ile?**), popatrzymy na statystykę i stowarzyszone rozkłady widziane przez te miary jakości. Wpierw popatrzymy co to za rozkład, a może normalny, to by wiele uprościło. Niestety jak widać z histogramu na rys 1 jest to rozkład **prawo skośny** (np. *Pareto, Weibull*) a nie Gaussowski i nie możemy skorzystać z jego zasobów, np. kwantyli. Dla czytelności histogramu odrzucono dwa skrajnie daleko odbiegające wyniki (*jak można przypuszczać powstałe z innej interpretacji niejasnych pytań ankiety*). Jak się wydaje w ogóle można by odrzucić ich pięć, aż do liczby punktów 25 tysięcy.



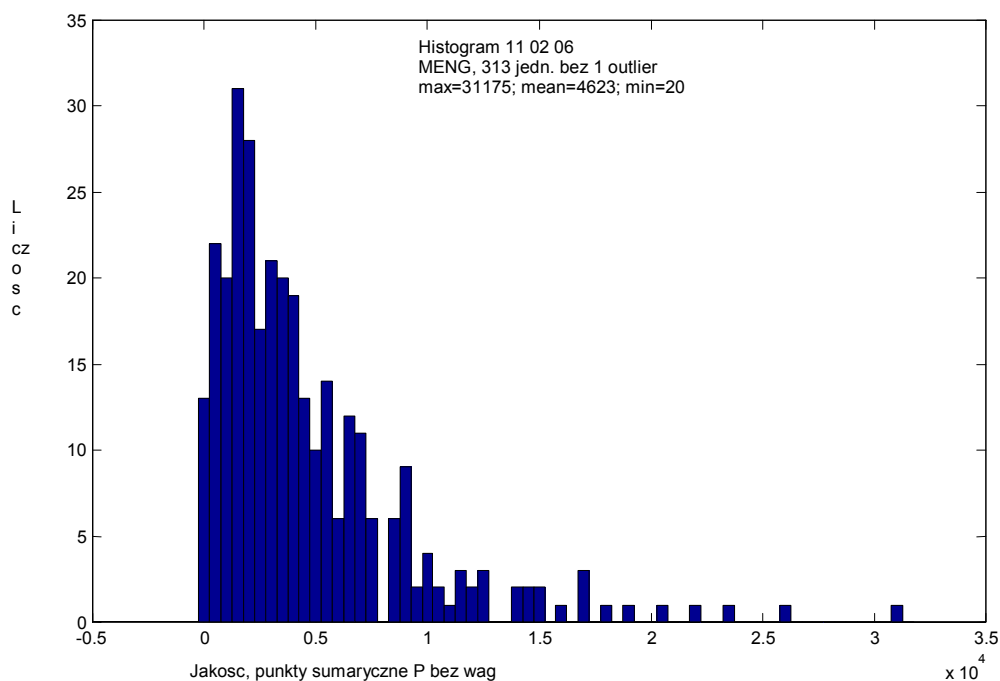
Rys. 1. Porównanie histogramu rozkładu punktów sumarycznych jednostek **P** z rozkładem Gaussowskim.

Jak widać rysunku jest to rozkład prawo skośny - Pareto podobny, z wartością minimalną liczby punktów **P** równą **9**, wartością średnią **4072** i wartością maksymalną **31173**, (przy odrzuceniu dwu skrajnych wyników z 832). Tak więc wartość maksymalna nie przekracza

⁴ Powyższe przyjęcie równoznaczne jest z założeniem równych jednostkowych wag; $w_i = 1$.

dziesięciokrotnej wartości średniej. Zatem gros ocen jakości jednostek dla symptomu **P** jest skupionych z lewej strony w zakresie niskich wartości, wrócimy do tego problemu jeszcze wielokrotnie.

W tym nurcie rozważań istotne jest pytanie czy ten sam charakter rozkładu mają podzbiory jednostek w **Komisji Nauki** (*jednostki humanistyczne i matematyczno przyrodnicze*) i w **Komisji Gospodarki** (*jednostki nauk technicznych, instytuty i ośrodki badawczo rozwojowe*). Jak się okazuje, obie podgrupy jednostek zachowują ten sam charakter statystyczny rozkładu, a ponieważ podgrupa Komisji Gospodarki jest mniejsza (*313 jednostek*) pokażemy tę własność na jej przykładzie, tak jak na rysunku 2.

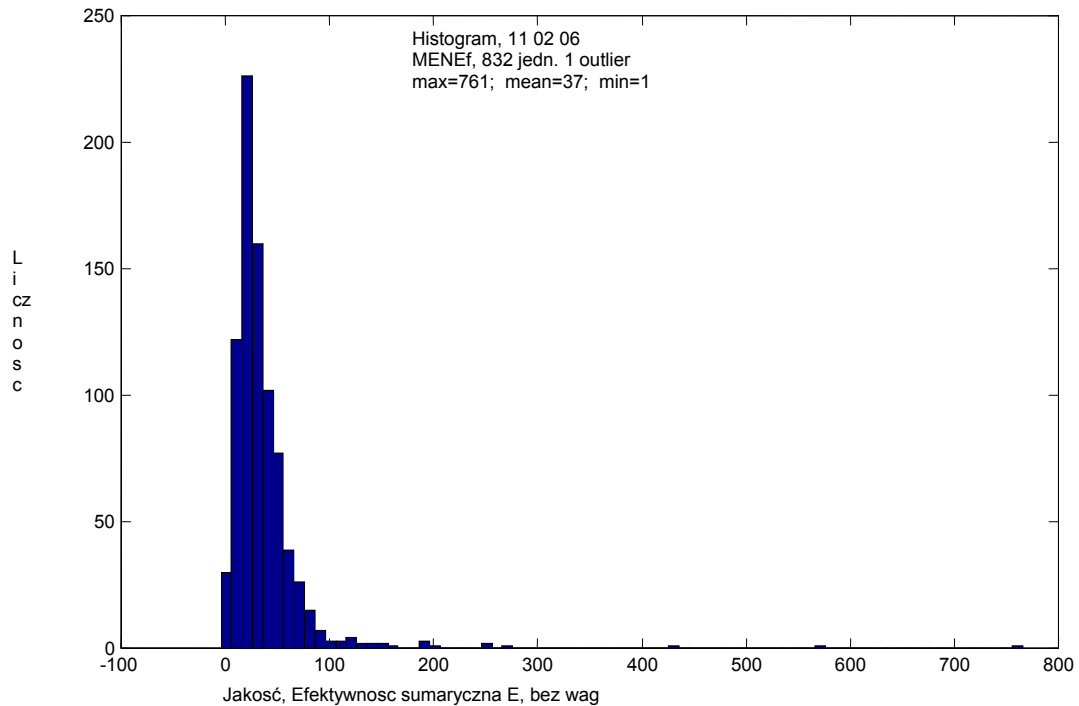


Rys. 2. Histogram punktów sumarycznych dla podzbioru jednostek naukowo innowacyjnych Komisji Gospodarki RN (313 jedn.).

Tak więc w skali punktów sumarycznych **P** mamy zawsze do czynienia z rozkładami prawo skośnymi typu **Pareto** lub Weibulla, z kilkoma jednostkami o anormalnej liczbie punktów (*ourliers*) i po ich odrzuceniu lub zweryfikowaniu, pozostała część rozkładu nadaje się dobrze do ustanowienia jednakowych przedziałów kategoryzacyjnych, co zobaczymy w następnym punkcie.

Zanim do tego dojdziemy zobaczmy czy ten sam wniosek da się wyprowadzić jeśli jakość badań w jednostkach będziemy mierzyli w skali **efektywności** sumarycznej **E**. Sytuację tę obrazuje kolejny rysunek 3, gdzie przy jednym odrzuceniu anormalnej efektywności parametry rozkładu efektywności są: minimum - **1**, średnia - **37**, maksimum - **761**. Tutaj więc wartość maksymalna skali jakości **E** przekracza **20** krotnie wartość średnią, inaczej i znacznie więcej niż w skali punktów **P**.

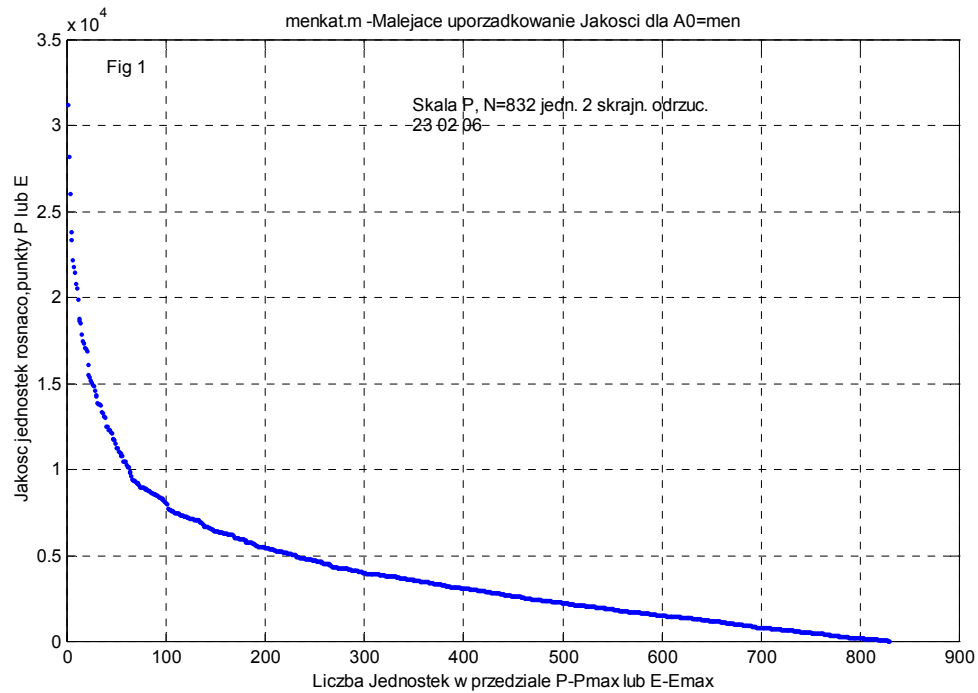
Podobnie też dzieje się dla podgrup nauki (**n**) czy też gospodarki (**g**). Zatem efektywnościowa miara jakości ma większe tendencje grupowania większości jednostek dla małych wartości **E**, a nie narusza anormalności jednostek skrajnych. Wiąże się to prawdopodobnie normalizacją do liczby zatrudnionych **N**, inną dla każdej jednostki. Wyobraźmy sobie zatem jak będzie wyglądał sytuacja w rozkładzie **E** unormowanym do wartości maksymalnej w zbiorze (*tak jak to nakazuje wspomniane zarządzenie Ministerstwa*)! W skali liniowej będzie to jedna duża plama w okolicy zera i kilka punktów na drodze do jedynki. Jedyna wtedy rada na rozróżnienie jednostek to używanie **logarytmicznej** skali jakości **E**.



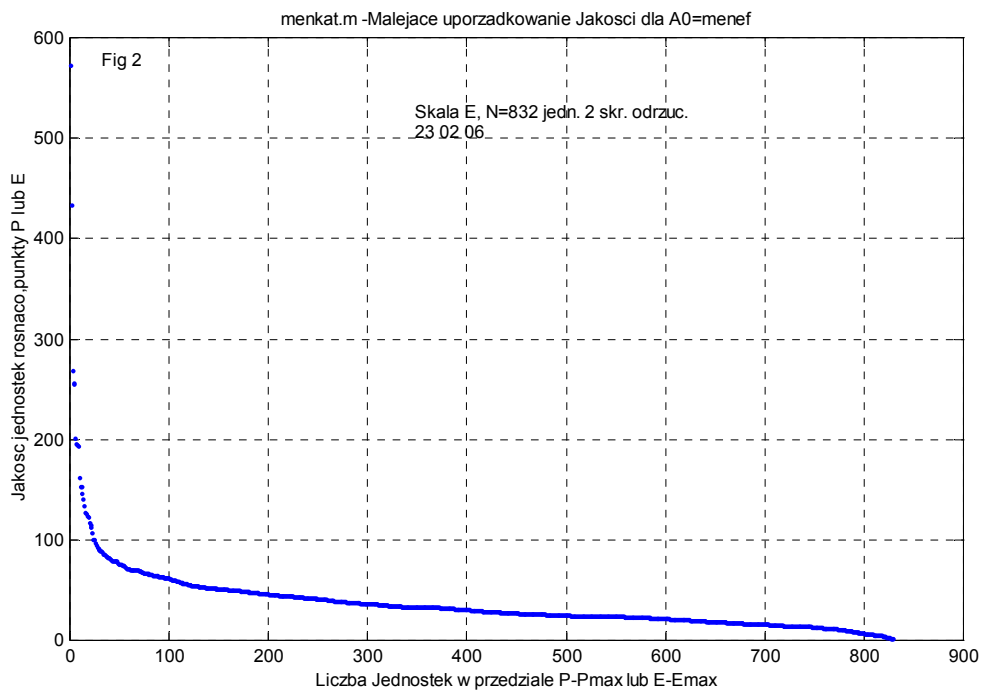
Rys. 3. Histogram łącznej efektywności badań i innowacji E dla całego zbioru jednostek z odrzuceniem jednego skrajnego wyniku.

3. Rozkłady licznosci jednostek w przedziałach jakości i percentyle

Zmierzając do uchwycenia istoty podziału jednostek na kategorie jakości uporządkujemy jednostki malejąco wg posiadanych punktów sumarycznych P a także efektywności E . W ten sposób uzyskamy łatwy do uchwycenia rozkład licznosci jednostek w zadanym przedziale $P-P_{max}$ lub $E-E_{max}$, tak jak na rysunkach 4 i 5. Jak widać z rysunku na ogólną liczbę jednostek **832** ujętych w wykazie **100** z nich ma więcej niż około 7 tysięcy punktów P . Podobnie dla efektywności E jako miary jakości, **100** jednostek ma efektywność sumaryczną większą niż **55**. Porównując rysunki 5 i 6 widać również, iż rozkład E jest bardziej stromy i podział na kategorie jednostek może być trudniejszy niż dla punktów P .

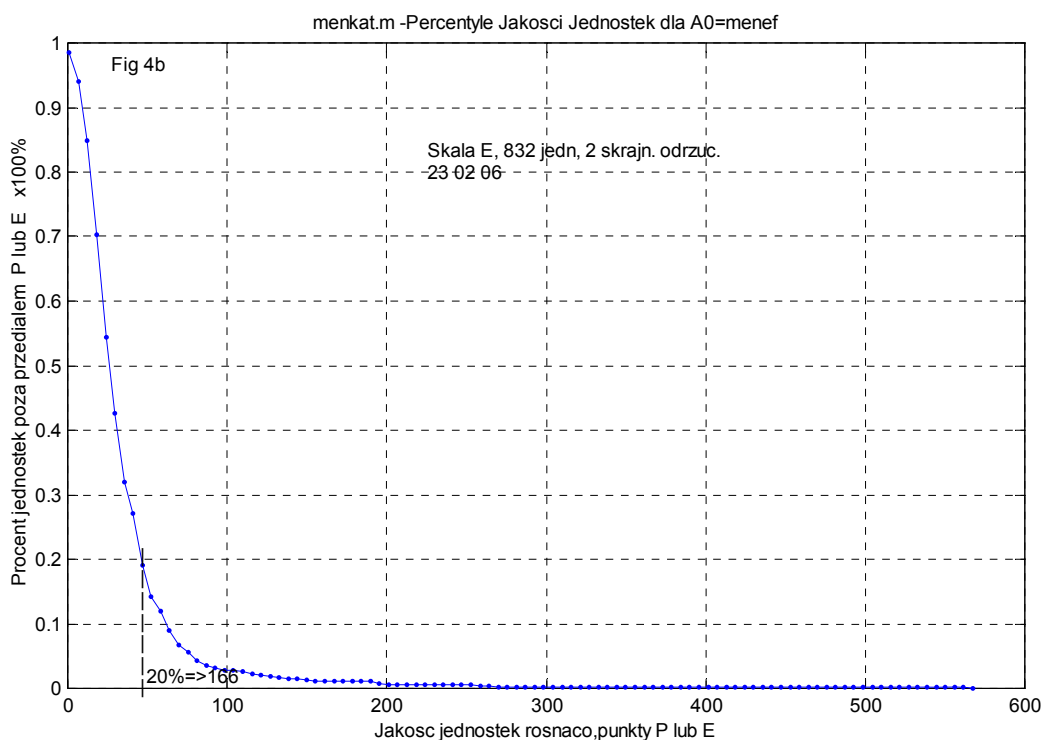
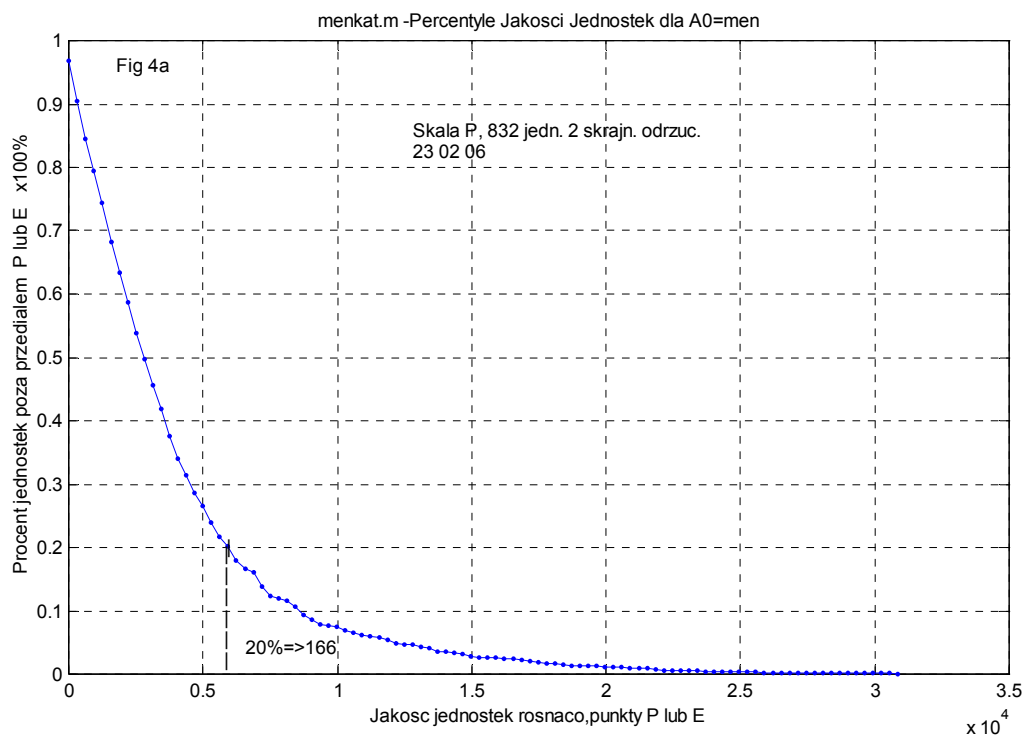


Rys. 4. Rozkład liczości wszystkich jednostek badawczych (832) w zadanym przedziale $P-P_{max}$



Rys.5. Rozkład liczości wszystkich jednostek badawczych (832) w zadanym przedziale $E-E_{max}$

Do celów klasyfikacji jednostek bardzo przydatna byłaby prezentacja miary jakości **P** lub **E** w skali procentu ogólnej liczby jednostek w zadanym przedziale jakości. Można to uzyskać jeśli z histogramów (rys.2 i 3) policzymy sumy kumulacyjne w przedziale i zaprezentujemy je w znormalizowanej skali procentowej liczości jednostek, tak jak na rysunkach 6a i 6b.



Rys.6. Procent całej liczby jednostek (830) jako funkcja przedziału punktów $P-P_{max}$ (a), oraz przedziały efektywności $E-E_{max}$. (b).

Jak widać oba rysunki 6 przedstawiają tę samą możliwość podziału jednostek na równo liczne grupy, np. 166 jednostek co stanowiłoby około **20 procentowe** przedziały w liczności jednostek, co na rysunkach zaznaczyliśmy jako **20% => 166**. Pierwsze **20%** najlepszych jednostek (**ca 166**) leży w przedziale punktowym powyżej ca **6 tys.** punktów **P** (Rys.6a), natomiast w skali efektywności

zakres ten rozciąga się powyżej $E = 45$ (Rys.6b). Dodać trzeba, że znowu wykres percentyli efektywności E jest bardzo stromy co może utrudniać klasyfikacje (*mała rozróżnialność jednostek*), odwrotnie niż w skali punktów P .

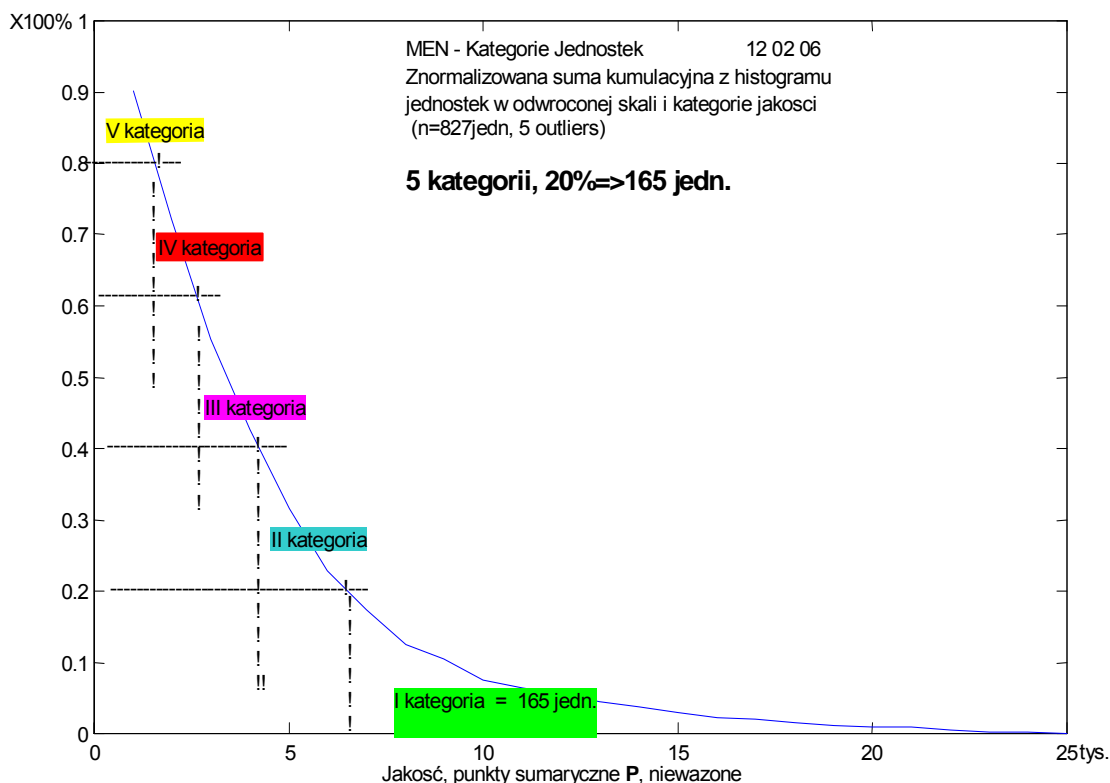
Interesująca będzie odpowiedź na pytanie, jak te same charakterystyki percentyli przedstawiają się w podzbiorach Komisji Nauki (**n**) i Komisji Gospodarki (**g**). Jak wynika z podobnych obliczeń 104 najlepsze jednostki Komisji Nauki (**20%** z ogólnej liczby **518**) posiada powyżej 5 tysięcy punktów P , natomiast w skali efektywności E pierwsze **20%** ma efektywność E większą od około 40.

To samo pytanie zadane danym Komisji Gospodarki daje podobną odpowiedź. Tutaj liczba jednostek jest mniejsza (**318**) a przedziały zmienności miar są inne niż poprzednio. Jednostek. I tak **20%** najlepszych jednostek (**63 jedn.**) posiada więcej niż **7 tysięcy** punktów P . Prawdopodobnie liczącym się powodem tego wzrostu granicy punktów są wdrożenia w trzecim strumieniu aktywności. Podobnie pierwsze 20% (**63 jedn.**) najlepszych jednostek uszeregowane w skali efektywności E zaczynają się od liczby $E=60$.

4. Koncepcja kategoryzacji jednostek

Jak widać z powyższego miary P i E zachowują te same prawidłowości rozkładów kumulacyjnych, a także podobieństwo cech obu komisji zależnie od przyjętej symptomów jakości P lub E jest zachowane. Tak więc dla punktów sumarycznych P krzywe rozkładów percentyli są bardziej prostoliniowe dla małej liczby punktów, co umożliwi lepszy podział na kategorie jakości. Percentyle łącznej efektywności w obu podgrupach nauki i gospodarki są podobne, niestety bardziej strome dla małych E , tam gdzie jest duże zagęszczenie jednostek. Utrudni to rozróżnienie jednostek, a także zmniejszy siłę racjonalnych argumentów na rzecz takiego rozróżnienia.

Dla pokazania samej istoty proponowanej kategoryzacji jakości jednostek weźmy zbiór wszystkich jednostek w skali jakości punktów P , ułatwiających rozróżnienie, jak to już pokazaliśmy niejednokrotnie. Podzielmy cały zakres jakości $0 - P_{\max}$ na pięć **20%** przedziałów, co nam da kategorie po około **165** jednostek, tak jak na rysunku 7.



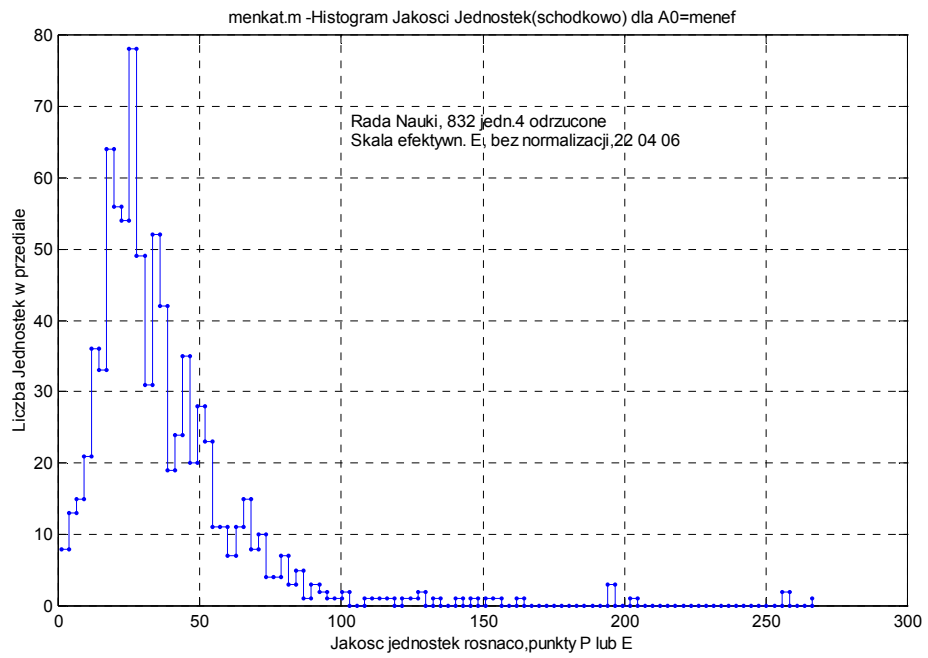
Rys. 7. Propozycja podziału jakości jednostek na pięć równych przedziałów 20% po około 165 jednostek

Jak widać z rysunku 7, prezentującego przedziały jakości w skali sumarycznych punktów **P**, koncepcja ta jest do przeprowadzenia, czyli jakość jednostek jest do rozróżnienia. Oczywiście jest to bardziej wykonalne jeśli liczba kategorii zmniejszy się np. do czterech, czyli do **25%** przedziałów po około **208** jednostek. Wymaga to jednak zastanowienia i symulacji w mniejszych grupach jednorodnych Komisji Gospodarki oraz Komisji Nauki.

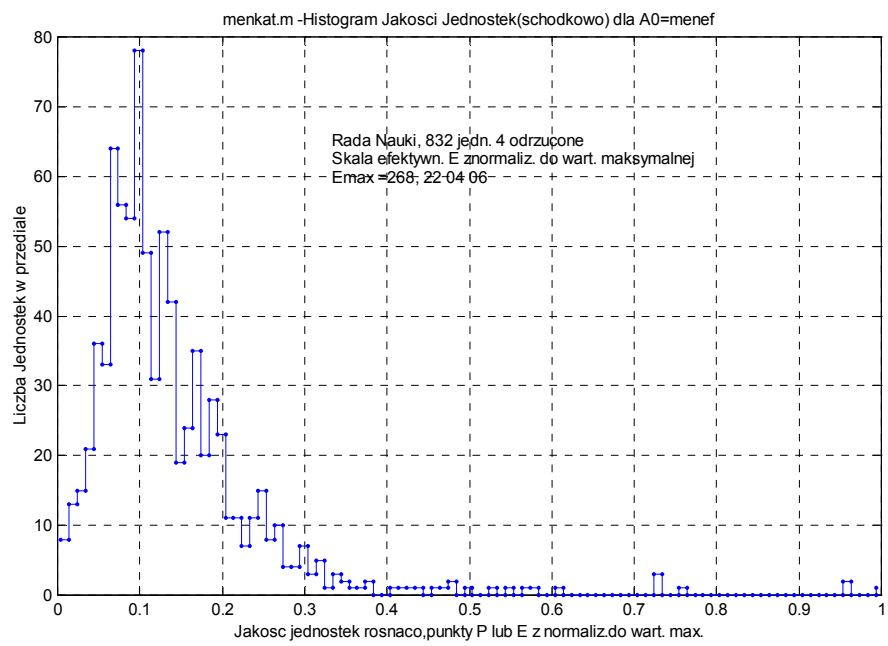
5. Skalowanie rozkładów

Wspomnieliśmy już kilkakrotnie, że zarządzenie Ministra z dnia 05 08 2005 wymaga normalizacji efektywności **E** do E_{max} w danej grupie jednorodnej, oczywiście po stosownym obliczeniu efektywności **ważonej**. Popatrzmy zatem jak się może zmienić nasza rozróżnialność jednostek zależnie od zastosowanego sposobu normalizacji w danej grupie zakładając, że wykonujemy to dla efektywności **E** i jak zwykle z wagą $w=1$. Kolejny rysunek 8 pokazuje trzy histogramy tego samego zbioru danych (**menef.txt**) z czterema odrzuconymi skrajnymi jednostkami, przedstawione w trzech różnych skalach; oryginalna skala efektywności **E** (obrazek a), z normalizacją do wartości maksymalnej w zbiorze $E_{max}=268$ (obrazek b), oraz z normalizacją zgodną z wymogami statystyki, czyli do wartości średniej $E_m=35,75$ (obrazek c). Przy sporządzaniu rysunku cały zakres $0-E_{max}$ podzieliliśmy na **100** przedziałów.

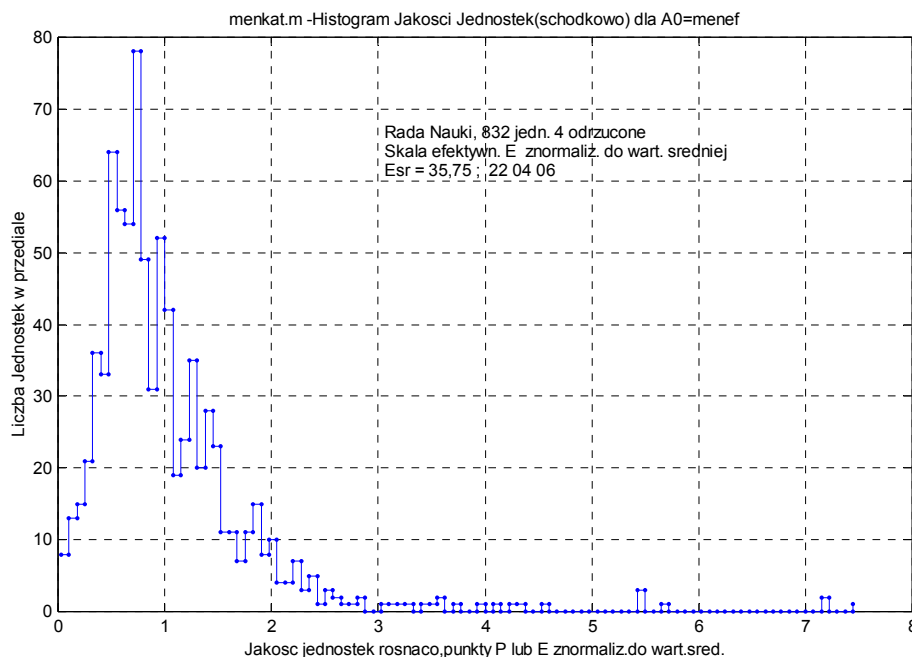
A



B



C



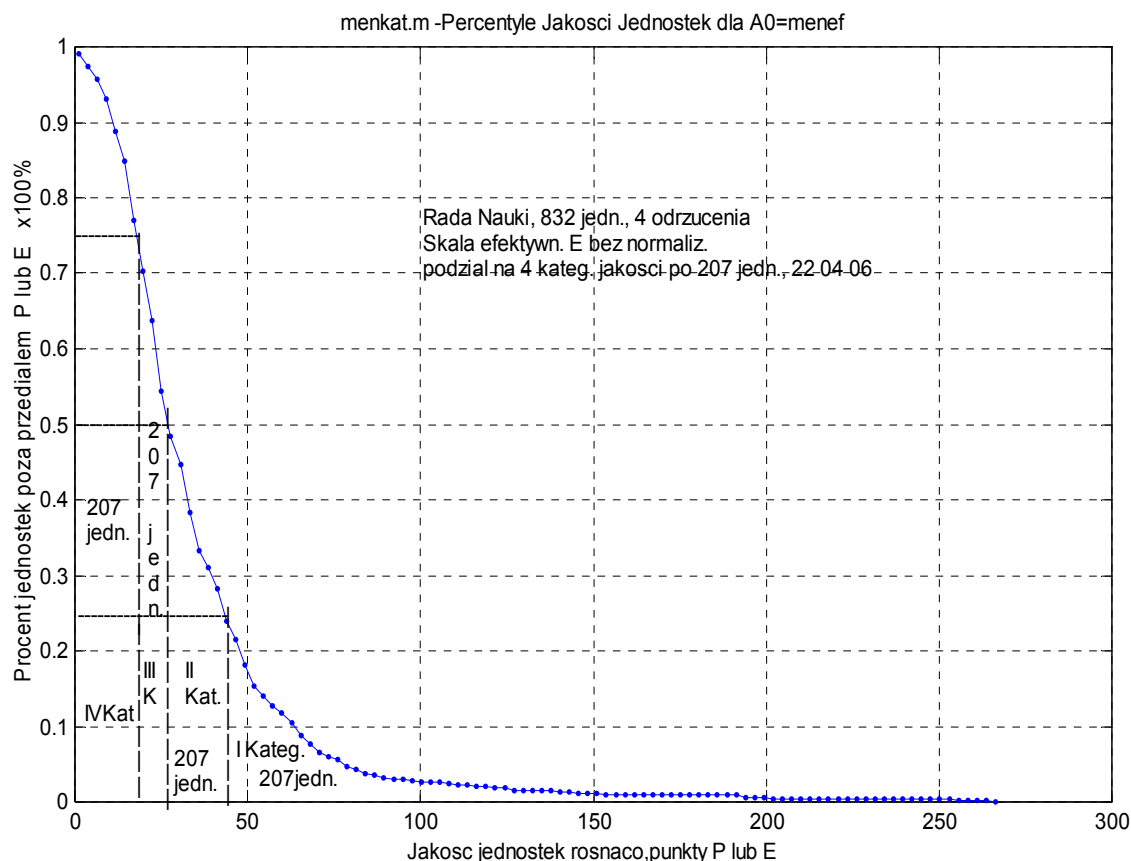
Rys. 8. Trzy różne skale prezentacji histogramu efektywności tego samego zbioru jednostek całej Rady Nauki

Komentując rysunek 8 z punktu widzenia ostatecznego celu, czyli kategoryzacji, zwróćmy uwagę na **szczyt histogramu** z lewej zawierający prawie **80** jednostek w przedziale. Te **80** jednostek na rysunku 8a (*bez normalizacji*), rozmieszczone jest w przedziale $E_{max}/100 = 2,68$ punktu skali efektywności. Natomiast na rys. 8b, przy normalizacji do wartości maksymalnej $E_{max} = 268$, ta sama liczba jednostek rozmieszczona jest w przedziale **0,01**. Z kolei zaś dla normalizacji zgodnej z wymogami statystyki, czyli do wartości średniej zbioru, te **80** jednostek mieści się w przedziale **0,08** unormowanej skali efektywności.

Zatem jaką metodę normalizacji wybrać by najlepiej rozróżnić jednostki ?

Odpowiedź jest jedna, **nie normalizować** wyniku, a wtedy rozróżnialność jednostek będzie najlepsza i tym samym przynależność do danej kategorii jakości najlepiej i **obiektywnie umotywowana**.

A jak by wyglądała tak kategoryzacja przy założonym podziale na 4 kategorie jakości wg efektywności jako symptomu jakości jednostki i bez żadnej normalizacji ? Rysunek 9 pokazuje taką hipotetyczną możliwość dla całej Rady Nauki (832 jednostki).



Rys. 9. Czysto algorytmiczny podział jednostek całej Rady Nauki na 4 kategorie jakości po 25%, czyli po 207 jednostek.

Jak widać z rysunku 9 zastosowane równe przedziały 25 procentowe generują nierówne przedziały w skali efektywności E, lecz zdrowo rozsądkowa ingerencja przywróci tu z pewnością zamierzony ład.

6. Optymalizacja opisu zbioru jednostek

Dążąc do prostoty opisu zagadnienia wzięliśmy wyżej pod uwagę najprostszy wektor opisujący stan jednostek, mianowicie wektor o dwu symptomach (*miarach*) jakości $\{E, P\}$ a nawet traktowaliśmy te symptomy alternatywnie; rozważając oddzielnie E lub P . Zastanówmy się obecnie nad lepszym opisem tego zbioru jednostek, a więc na wektorze cech o większej liczbie symptomów. A może wśród tej zwiększonej wymiarowości opisu znajdziemy nowy zbiór **cech niezależnych** (*ortogonalnych*) charakteryzujących każdą jednostkę?

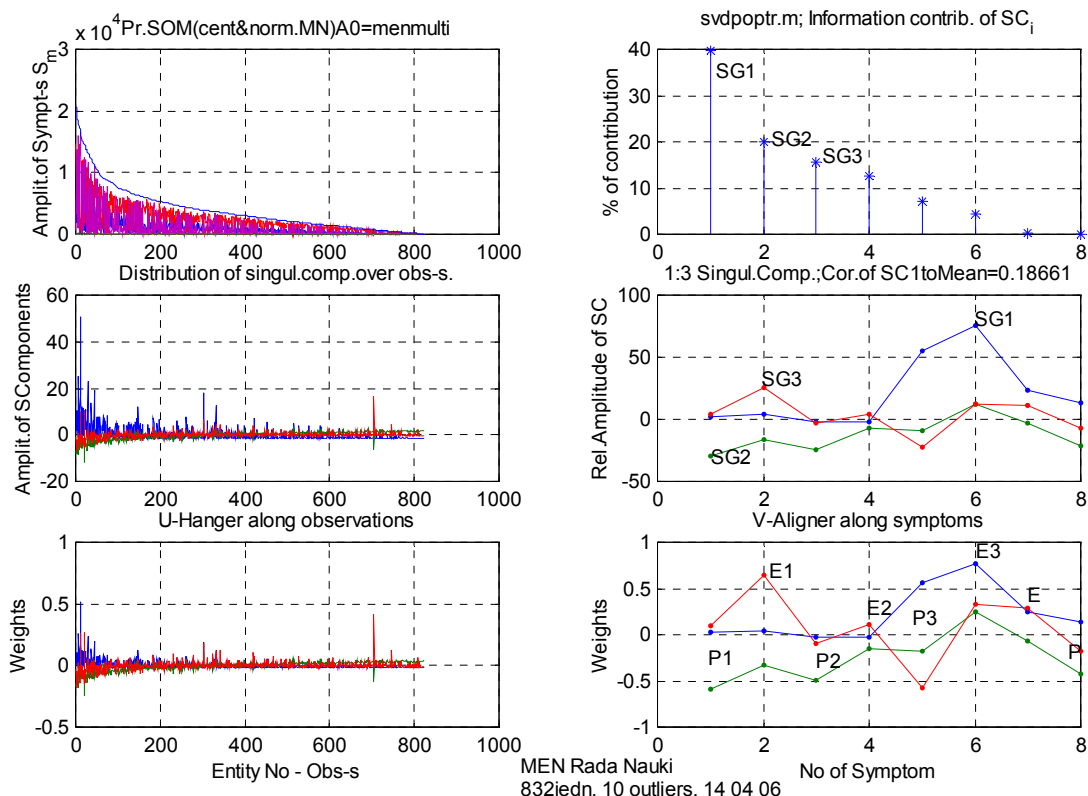
Zarówno P jak i E są sumami odnośnych punktów P_i i efektywności E_i ($i=1,2,3$). Weźmy zatem jako składowe wektora opisującego jednostkę jej punkty i efektywności składowe i sumaryczne, co maksymalnie da nam wektor symptomów S opisujący jednostkę o ośmiu składowych

$$S_0 = \{P_1, E_1, P_2, E_2, P_3, E_3, P, E\} \quad (1)$$

Dla zbadania istotności tych symptomów jakości i możliwości prostszego optymalnego opisu jednostki zastosujemy koncepcje składowych głównych (*principal component*) znaną w analizie wielowymiarowych populacji (*PCA – Principal Component Analysis*) [5]. Operacja ta staje się prostsza i lepiej interpretowana jeśli dodatkowo zastosujemy rozkład naszej macierzy obserwacji (tworzą ją kolejne 832 realizacje wektora cech S_0 dla każdej jednostki) wg składowych

szczególnych, tzw. w nomenklaturze angielskiej **SVD** (*Singular Value Decomposition*). Program taki wykorzystałem już do badania populacji ocen ekspertów dotyczących wniosków o wsparcie badań przez **Fundację Nauki Polskiej** [4], stąd też z drobnymi zmianami posłużę się nim tutaj.

Jak wynika z tych badań najlepsze rezultaty w ekstrakcji optymalnego opisu populacji otrzymuje się kiedy symptomową macierz obserwacji (**SOM**) przekształcimy przez centrowanie i normalizację jej kolumn (*symptomów*) do wartości średniej. Podobna transformacja SOM^5 sprawdza się i w naszym przypadku co widać z rysunku 10, gdzie kolejne obrazki przedstawiają ważne etapy i rezultaty obliczeń.



Rys. 10. Składowe główne opisu całej populacji jednostek przy 8miu symptomach opisujących jakość jednostki przedstawione wzdłuż osi zbioru jednostek (z lewej) i zbioru symptomów (z prawej).

I tak, obrazek lewy górny przedstawia w zarysie całą macierz obserwacji z 8mioma symptomami w skali jednostek w liczbie **832**. Jak widać, poza sumaryczną wartością punktów **P** (największym z symptomów), niewiele tu można zauważyć, co najwyżej wyciągnąć wniosek o dużym skomplikowaniu zagadnienia. Obrazek prawy górny pokazuje liczbę niezależnych składowych informacyjnych zawartych w naszej macierzy obserwacji i jak widać mimo 8miu składowych mamy tylko sześć niezależnych źródeł informacji (dwa słupki zerowe a dwa dalsze mniejsze niż 10%). Domyśleć się można z własności ortogonalności rozkładów **SVD** oraz sposobów tworzenia naszych symptomów, że w naszym wektorze (1) sumy punktów **P** i efektywności **E** można odrzucić jako symptomy zależne, już raz uwzględnione w poprzez **P_i** oraz **E_i**.

Obrazek środkowy lewy przedstawia trzy pierwsze składowe główne **SG_i**, (lub szczególne), przedstawione w skali wszystkich **832** jednostek. Jak widać da się tu z trudem odróżnić trzy kolory, ale liczba jednostek nie ułatwia rozróżnienia. Natomiast te same składowe główne przedstawione są w skali pierwotnych symptomów wektora **S₀** na obrazku prawy środkowym, który jest dobrze rozróżnialny i niezwykle istotny dla naszej analizy. Widać na nim trzy niezależne charakterystyki naszej populacji i sposób ich pozyskania tak jak na obrazku prawym dolnym pokazującym wagi

⁵ W naszym przypadku każda jednostka to wiersz ośmiokolumnowej SOM ze swoimi wartościami wektora cech. Zatem nasza SOM zawiera 832 wiersze i 8 kolumn opisujących jakość jednostki.

pierwotnych symptomów P_i , E_i . Charakterystykom tym (*składowym głównym SG_i*) odpowiadają oczywiście proporcjonalne wysokości słupków na rysunku prawym górnym. Tak więc pierwszą niezależną charakterystykę naszej populacji jednostek **$SG1$** (*40% informacji*) uzyskamy z prostej ważonej sumy P_3 i E_3 , jeśli pominiemy E i P z uwagi na ich zależność od składników cząstkowych, o czym już mówiliśmy. Kolejna **20%** charakterystyka **$SG2$** , jak widać z obrazka prawego dolnego, to liniowa kombinacja stosownych wag i wszystkich symptomów P_i i E_i . Wreszcie trzecia składowa główna **$SG3$** , najmniejsza, tworzona jest przede wszystkim jako liniowa kombinacja symptomów E_1 , P_3 i E_3 , bo udziały innych symptomów (*obrazek dolny prawy*) są prawie zerowe. Co jest ważne, że te same wnioski o składowych głównych można wyciągnąć analizując podzbiory jednostek Komisji Nauki (*518 jednostek*) oraz Komisji Gospodarki (*313 jednostek*), czego już nie pokazujemy.

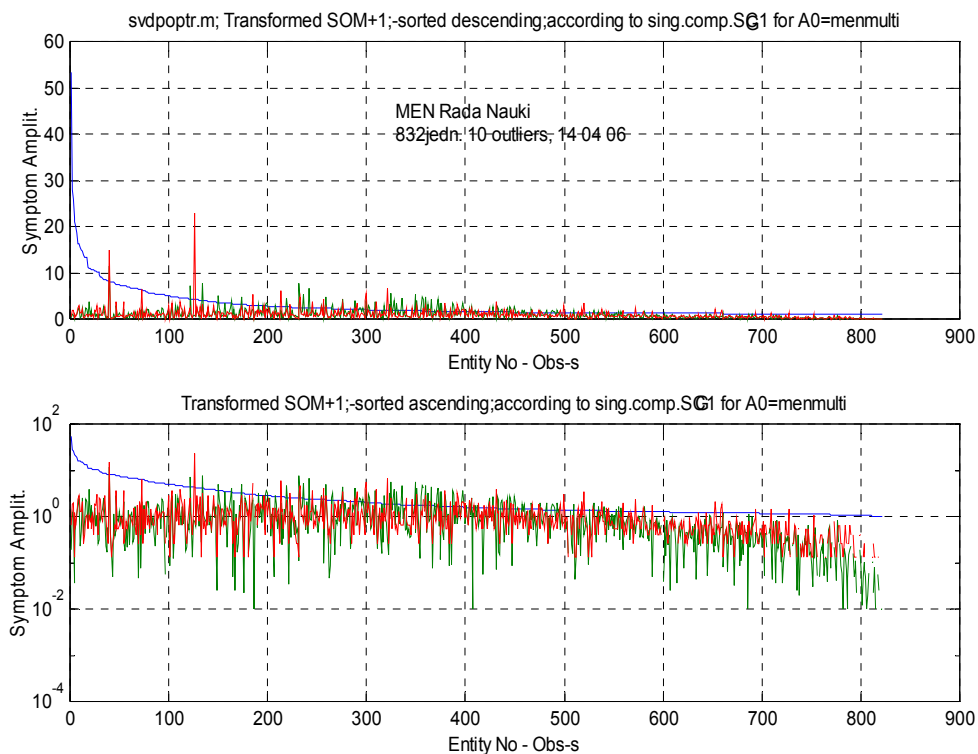
Zatem analiza nasza pokazała, że potrafimy wyróżnić trzy maksymalnie informacyjne cechy (*symptomy*) naszej populacji jednostek;

$SG1$ jako opis aktywności wdrożeniowej (*głównie $P3$, $E3$*),

$SG2$ jako ważona sumę wszystkich strumieni aktywności jednostek (*$P1$, $E1$, $P2$, $E2$, $E3$*),

$SG3$ jako osiągnięcia naukowo wdrożeniowe (*$E1$, $P3$, $E3$*).

Zachodzi jeszcze pytanie czy potrafimy dokonać kategoryzacji podług tak wyróżnionych składowych optymalnego opisu. To znaczy czy potrafimy pogrupować jednostki zgodnie z wybraną składową główną. Okazuje się że tak, jak to wynika z rysunku 11 gdzie przykładowo pogrupowano wszystkie jednostki zgodnie z malejącą wartością **$SG1$** .

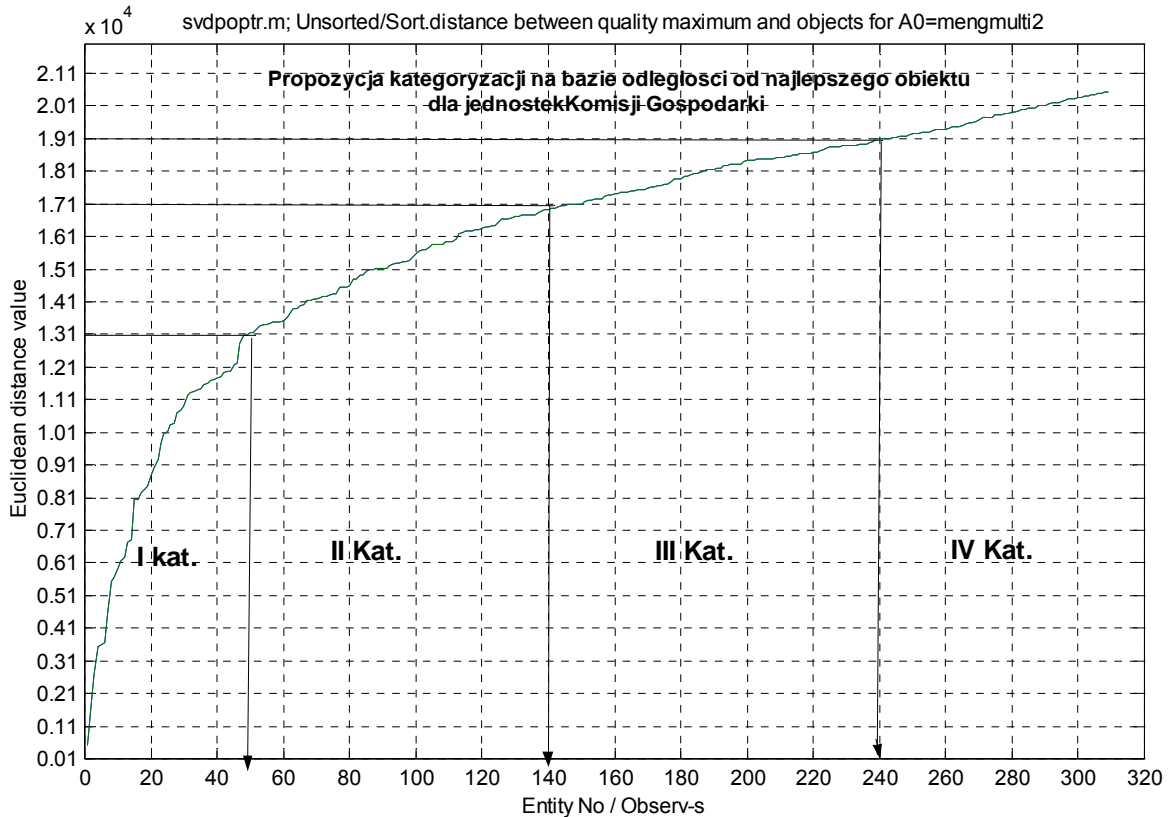


Rys.11 Przykład uszeregowania pierwszych trzech kolumn macierzy obserwacji wszystkich 832 jednostek zgodnie z malejącą wartością pierwszej składowej głównej $SG1$, w skali liniowej (górze) i logarytmicznej (dół).

Widać z powyższego rysunku, iż różnica w wartościach symptomu uogólnionego **$SG1$** , klasyfikującego jednostki rozciąga się od **0** do **50**, a skala logarytmiczna ułatwiła by nam tu zadanie wielce (*obrazek dolny*).

Jednak w obecnej kategoryzacji jednostek możemy dysponować tym, co mamy i to najlepiej w minimalnym rozmiarze wektora symptomów. Analizując obrazek prawy dolny rysunku 13 zauważmy, że najbardziej istotne są tam składowe wektora symptomów $S=\{E_1, E_2, E_3, P_c\}$. Wybierzmy zatem taki minimalny opis do dalszej analizy i zamiast rozważać kwestie jego

optymalizacji popatrzymy na ułatwienie możliwości klasyfikacji (*kategoryzacji jednostek*). Zatem w naszym programie obliczeniowym stworzymy sztucznie **najlepszą jednostkę**, taką co ma maksymalne wszystkie cechy, tzn. nasz wzorzec jakości będzie: $S_{\max} = \{E_{1\max}, E_{2\max}, E_{3\max}, P_{c\max}\}$ i obliczymy odległości istniejących jednostek od tego idealnego wzorca. Oczywiście zasada jest, że im bliżej wzorca jakości znajduje się dana jednostka tym lepsza jej kategoria. Taką propozycję kategoryzacji dla jednostek Komisji Gospodarki przedstawia niżej kolejny rysunek 12.



Rys. 12. Propozycja kategoryzacji jednostek na podstawie odległości od obiektu najlepszego w czterowymiarowej przestrzeni opisu efektywności cząstkowych i punktów całkowitych $\{E_1, E_2, E_3, P_c\}$

Konieczne dane do takich obliczeń znajdują się wszystkie w arkuszu kalkulacyjnym **Excel**, można by zatem w jednej dodatkowej kolumnie obliczyć odległość D_i dla każdej jednostki, wg prostego wzoru na odległość Euklidesową (*lub inną*);

$$D_i = \{ (E_{1\max} - E_{1i})^2 + (E_{2\max} - E_{2i})^2 + (E_{3\max} - E_{3i})^2 + (P_{c\max} - P_{ci})^2 \}^{1/2}. \quad (2)$$

I to już wszystko, bo **Excel** pogrupuje następnie rosnąco jednostki i pozostaje tylko zastosować zdrowy rozsądek do ustanowienia granic takich kategorii jakości (*patrz rys. 12*).

7. Podsumowanie

Metodologia kategoryzacji jednostek podjęta wyżej jest frapującym problemem aplikacji statystycznej teorii decyzji. Analizując całość danych, początkowo w kategoriach dwu zmiennych opisujących (*symptomów jakości jednostek*) pokazaliśmy, że cała populacja jednostek i podgrupy jednostek Komisji Nauki i Komisji Gospodarki mają ten sam charakter **Pareto** podobny rozkładów prawo skośnych. Analizując zaś rozkłady kumulacyjne jednostek pokazaliśmy, że można pogrupować jednostki na pięć kategorii po 20% każda (*rys. 10*). Poruszyliśmy również problem normalizacji wyniku efektywności każdej jednostki do wartości reprezentatywnej dla danej podgrupy. Okazuje się z tego, że najlepsze rozróżnienie jednostek daje brak normalizacji. Następnie

podjęliśmy temat optymalnego opisu populacji jednostek formułując wektor obserwacji o większym wymiarze, ośmiu symptomów składowych: \mathbf{P} , \mathbf{E} , \mathbf{P}_i , \mathbf{E}_i , $i=1,2,3$. Za pomocą analizy składowych głównych pokazaliśmy, że do optymalnej charakterystyki jednostek wystarczy sześć składowych wektora $\{\mathbf{P}_i, \mathbf{E}_i\}$, $i=1,2,3$. Pokazano także wstępnie, że da się klasyfikować jednostki za pomocą wyłonionych przez PCA składowych głównych \mathbf{SG}_i .

Nie mając jednak zamiaru robić rewolucji, np. przez propozycję nowych pytań ankiety⁶, zaproponowano prostą metodę klasyfikacji jako **odległość jednostki** od obiektu idealnego dla danego zbioru jednostek, co automatycznie na podstawie obecnych danych (*zweryfikowanych*) mógłby wyliczać program Excel.

Jakie zatem praktyczne wnioski wynikają z powyższego dla aktualnie przeprowadzonego procesu kategoryzacji? Jak się wydaje można je ująć w kilku punktach jak niżej.

1. Paretoowski charakter populacji implikuje, że pierwsza kategoria nie powinna być liczniejsza niż **20%** danej grupy, bądź też co najwyżej **25%**.
2. Najlepszą rozróżnialność jakości jednostek dają nieznormalizowane punkty \mathbf{P} lub \mathbf{E} .
3. W większości przypadków analizy rozkładów populacji jak i w PCA istotną rolę odgrywały efektywności cząstkowe \mathbf{E}_i jak i punkty cząstkowe \mathbf{P}_i . **Nie da się** więc powiedzieć jednoznacznie, że jakość jednostek opisują tylko efektywności \mathbf{E}_i , bądź ich ważone sumy, trzeba zatem w kategoryzacji używać punktów \mathbf{P} jak i efektywności \mathbf{E} .
4. Dążąc do ulepszenia efektów, kategoryzacji można by ostrożnie próbować kategoryzacji według pokazanej koncepcji składowych głównych \mathbf{SG}_i , lecz wymagałoby to dalszych obliczeń i symulacji, oczywiście dla zweryfikowanych już danych wejściowych.
5. Nie mając obecnie czasu i możliwości zaprojektowania i przeprowadzenia ponownie akcji oceny jednostek (*nowa ankieta*), można by przyjąć klasyfikację jakości jako odległość od zdefiniowanej jednostki najlepszej, posiadającej najlepsze wartości swych cech opisujących. Jak wynika z obliczeń mogą to być cech opisane wektorem $\mathbf{S}=\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{P}_c\}$, a odległość Euklidesowska jednostek w każdej zdefiniowanej subpopulacji jest do policzenia w ramach programu **Excel**.

8. Literatura

1. Strona internetowa Komitetu Badań Naukowych, **Internet**, <http://www.kbn.gov.pl>, kwiecień 2006.
2. Bendyk E., Potęga Google, **Polityka**, No 13, 1 kwiecień 2006., s78-80.
3. Dane zawarte w CD, **System oceny parametrycznej**, 2001-2004, **OPI**, Warszawa, luty 2006.
4. Cempel C., Rozkład symptomowej macierzy obserwacji populacji jako pomoc w ocenie jakości wniosków i obiektów, **Diagnostyka**, 35, 2005, s7-12.
5. Principal Component Analysis, **Internet**, <http://en.wikipedia.org/Wiki/>, kwiecień 2006
6. Singular Value Decomposition, **Internet**, <http://en.wikipedia.org/Wiki/>, kwiecień 2006.
7. Wachowiak P., (red), **Pomiar Kapitału intelektualnego przedsiębiorstwa**, Wyd. SGH, 2005, s 167.

Poznań kwiecień/maj 06

⁶ Ustanowionych np. wg metod pomiaru kapitału intelektualnego [7,rozd.7].