

Alina Jędrzejczak*

OCENA APROKSYMACJI ROZKŁADÓW PŁAC
WEDŁUG WYSOKOŚCI W GOSPODARCE USPOŁECZNIONEJ
ROZKŁADEM LOGARYTMICZNO-NORMALNYM

1. Wstęp

W artykule tym chcemy przedstawić aproksymację rozkładów płac ludności w gospodarce uspołecznionej w 1986 r. za pomocą rozkładu logarytmiczno-normalnego. Jako rozkład teoretyczny uwzględniono rozkład logarytmiczno-normalny, gdyż z poprzednich badań wynikało, że dobrze aproksymuje on empiryczne rozkłady płac i dochodów ludności. Wykazanie zgodności empirycznego rozkładu płac z rozkładem teoretycznym (logarytmiczno-normalnym) umożliwia przeprowadzenie dokładnej analizy struktury płac, a także ich prognozowanie.

2. Podstawowe własności rozkładu logarytmiczno-normalnego

Ogólną teorię rozkładu logarytmiczno-normalnego sformułował po raz pierwszy D. Mc Alister w XIX w. Rozkład ten został po raz pierwszy zastosowany do badania płac i dochodów przez francuskiego ekonomistę i statystyka R. Gibrata w 1931 r.

W Polsce rozkład logarytmiczno-normalny zastosował J. Wiśniewski już w 1934 r. Z jego badań wynikało, że rozkład ten może

* Mgr, asystent W Instytucie Ekonometrii i Statystyki UL.

być stosowany do badania płac pracowników biurowych. Dwaj polscy statystycy Z. Pawłowski i E. Vielrose wykazali przydatność tego rozkładu do aproksymacji płac wszystkich pracowników zatrudnionych w gospodarce uspołecznionej.

Z kolei Kordos i Stroińska badając płace w Polsce w latach 1955-1970 wykazali, że rozkład logarytmiczno-normalny dobrze opisuje rozkłady płac w różnych przekrojach (gałęziowym, działowym), a także w podziale na pracowników fizycznych i umysłowych.

Rozkład ten pozwala więc badać różne rodzaje rozkładów płac w gospodarce uspołecznionej. Może być bardziej przydatny w analizie płac i dochodów ludności niż inne znane rozkłady teoretyczne opisujące to zjawisko (K o r d o s [3]), jak rozkład Pareta, rozkład normalny, czy też krzywe Pearsona typów I, III, IV, VI. Rozkład logarytmiczno-normalny wykazuje wysoką zgodność z empirycznymi rozkładami płac (czego nie można powiedzieć o rozkładzie normalnym i rozkładzie Pareta). Wygodnie się nim posługiwać w dalszej analizie ekonomicznej, gdy większość jego charakterystyk zależy od jednego parametru (nie zaś od wielu parametrów jak w przypadku krzywych Pearsona¹). Własności rozkładu logarytmiczno-normalnego mogą być wykorzystane do budowy modeli prognoz rozkładów płac i dochodów ludności.

Podstawowe własności rozkładu logarytmiczno-normalnego zostały przedstawione w monografii A i t c h i s o n a i B r o w n a [1]. Mówimy, że zmienna losowa ma rozkład logarytmiczno-normalny, jeśli logarytmy tej zmiennej układają się według krzywej normalnej. Funkcję gęstości tego rozkładu można wyrazić za pomocą następującego wzoru:

$$(1) \quad f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} (\ln x - \mu)^2\right\}$$

dla $x > 0$,

gdzie: μ - średnia arytmetyczna logarytmów zmiennej losowej X , σ - odchylenie standardowe logarytmów zmiennej losowej X .

¹ Należy zauważyć, że rozkład Pareta jest szczególnym przypadkiem krzywej Pearsona typu III, natomiast rozkład normalny jest szczególnym przypadkiem krzywej typu VII.

Wzór ten określa rozkład logarytmiczno-normalny dwuparametrowy. Parametrami są tu: μ i σ^2 .

W badaniach rozważa się także inne warianty rozkładu logarytmiczno-normalnego: rozkład trój- i czteroparametrowy.

W rozkładzie trójparametrowym obok μ i σ^2 występuje jeszcze trzeci parametr T taki, że $x > T$. Rozkład czteroparametrowy określa cztery parametry: μ , σ^2 , T i θ ($T < x < \theta$). W naszych rozważaniach ograniczymy się do rozkładu dwuparametrowego. Cechuje go prostota, jest łatwiejszy do interpretacji i wygodniejszy do opisywania empirycznych rozkładów płac.

Średnią arytmetyczną α i wariancję β^2 rozkładu logarytmiczno-normalnego dwuparametrowego określają wzory:

$$(2) \quad \alpha = e^{\mu} + \frac{1}{2}\sigma^2, \quad \beta^2 = e^{2\mu + \sigma^2}(e\sigma^2 - 1),$$

Mediana i dominanta przyjmują postać:

$$(3) \quad Me = e^{\mu}, \quad Do = e^{\mu - \sigma^2}.$$

Kwantyle rozkładu logarytmiczno-normalnego można wyrazić za pomocą kwantyli rozkładu normalnego $N(0,1)$. Kwantyl rzędu q rozkładu logarytmiczno-normalnego wyraża wzór:

$$(4) \quad x_q = e^{\mu + U_q \sigma}$$

gdzie: U_q - kwantyl rzędu q rozkładu $N(0,1)$.

3. Metody szacowania parametrów rozkładu logarytmiczno-normalnego

Do najczęściej stosowanych metod estymacji parametrów rozkładu logarytmiczno-normalnego należą:

- metoda największej wiarygodności,
- metoda momentów,
- metoda najmniejszych kwadratów,
- metoda kwantyli.

W metodzie największej wiarygodności estymatory m i s^2 (parametrów μ i σ^2) przyjmują postać:

$$(5a) \quad m = \frac{1}{n} \sum_{i=1}^n \ln x_i,$$

$$(5b) \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (\ln x_i - m)^2.$$

Oba estymatory są nieobciążone i najbardziej efektywne. W przypadku danych pogrupowanych konieczne są jednak modyfikacje wymienionych wzorów (K o r d o s [3], s. 134). Poza tym, w przypadku gdy skrajne grupy są nieograniczone i liczne, możemy popełnić duży błąd przy obliczaniu średnich grupowych.

Metoda momentów opiera się na wykorzystaniu momentów zwykłych pierwszego i drugiego rzędu. Estymatory m i s^2 otrzymujemy porównując średnią i moment zwykły drugiego rzędu z próby do odpowiednich momentów rozkładu logarytmiczno-normalnego (za μ i σ^2 przyjmujemy ich estymatory):

$$(6a) \quad \frac{1}{n} \sum_{i=1}^n x_i = e^{m + \frac{1}{2}s^2},$$

$$(6b) \quad \frac{1}{n} \sum_{i=1}^n x_i^2 = e^{2m + 2s^2},$$

stąd

$$(7a) \quad m = 2 \ln\left(\frac{1}{n} \sum_{i=1}^n x_i\right) - \frac{1}{2} \ln\left(\frac{1}{n} \sum_{i=1}^n x_i^2\right),$$

$$(7b) \quad s^2 = \ln\left(\frac{1}{n} \sum_{i=1}^n x_i^2\right) - 2 \ln\left(\frac{1}{n} \sum_{i=1}^n x_i\right).$$

Oba estymatory są zgodne, jednakże efektywność ich gwałtownie spada wraz ze wzrostem σ^2 . W przypadku danych pogrupowanych napotykamy tu liczne problemy. Konieczne jest stosowanie tzw. poprawek Shepparda. Mogą także wystąpić trudności związane z obliczaniem średniej arytmetycznej dla grup skrajnych (podobnie jak w metodzie największej wiarygodności).

Jeśli chcemy korzystać z danych pogrupowanych to wygodniej stosować metodę najmniejszych kwadratów. W metodzie tej nie występują trudności związane ze zagregatowaną formą przedstawienia

danych. Aby otrzymać oceny parametrów μ i σ^2 metodą najmniejszych kwadratów musi być spełniony warunek:

$$(8) \quad \sum_{i=1}^{n-1} (\ln x_i - \sigma u_i - \mu)^2 = \min,$$

gdzie: u_i - argument znormalizowanej dystrybuanty rozkładu normalnego, odpowiadającej prawdopodobieństwu równemu i/n ($i = 1, \dots, n-1$).

Estymatory m i s przyjmują postać:

$$(9a) \quad m = \frac{\sum_{i=1}^{n-1} u_i^2 \sum_{i=1}^{n-1} \ln x_i - \sum_{i=1}^{n-1} u_i^2 \sum_{i=1}^{n-1} u_i \ln x_i}{(n-1) \sum_{i=1}^{n-1} u_i^2 - \left(\sum_{i=1}^{n-1} u_i \right)^2},$$

$$(9b) \quad s = \frac{(n-1) \sum_{i=1}^{n-1} u_i \ln x_i - \sum_{i=1}^{n-1} u_i \sum_{i=1}^{n-1} \ln x_i}{(n-1) \sum_{i=1}^{n-1} u_i^2 - \left(\sum_{i=1}^{n-1} u_i \right)^2}.$$

W przypadku danych pogrupowanych n oznacza liczbę grup, a x_i - górne granice przedziałów.

Metodą stosowaną najczęściej do oceny parametrów μ i σ^2 rozkładu logarytmiczno-normalnego jest metoda kwantyli. Pozwala ona na oszacowanie tych parametrów z danych indywidualnych i pogrupowanych (podobnie jak metoda najmniejszych kwadratów). Poza tym cechuje ją stosunkowo prosta procedura obliczeniowa. Dlatego też tę właśnie metodę wykorzystaliśmy w naszych obliczeniach (dysponujemy danymi pogrupowanymi, skrajne przedziały klasowe są nieograniczone).

Aby oszacować parametry μ i σ^2 metodą kwantyli. Wybieramy dwa kwantyle rzędów q_1 i q_2 (za μ i σ^2 przyjmujemy ich estymatory):

$$(10a) \quad x_{q_1} = e^{m + u_{q_1} s},$$

$$(10b) \quad (q_1 < q_2),$$

$$x_{q_2} = e^{m + u_{q_2} s}$$

stąd otrzymujemy estymatory:

$$(11a) \quad m = \frac{u_{q_2} \ln x_{q_1} - u_{q_1} \ln x_{q_2}}{u_{q_2} - u_{q_1}},$$

$$(11b) \quad s = \frac{\ln x_{q_2} - \ln x_{q_1}}{u_{q_2} - u_{q_1}}.$$

Może istnieć wiele wariantów tej metody zależnych od doboru pary kwantyli. Jednak badając efektywności obu estymatorów wykazano, że największą efektywność m (81%) uzyskujemy stosując kwantyle rzędów: 0,27 i 0,73. Natomiast największą efektywność s^2 (65%) można otrzymać obliczając kwantyle: 0,07 i 0,97.

Stosując kwantyle najbardziej efektywne otrzymujemy następujące estymatory:

$$(12a) \quad m = \frac{1}{2} (\ln x_{0,73} - \ln x_{0,27}),$$

$$(12b) \quad s = \frac{1}{2 u_{0,93}} (\ln x_{0,93} - \ln x_{0,07}).$$

Gdy dane są tak pogrupowane, że nie można obliczyć kwantyli najbardziej efektywnych, to wskazana jest ich interpolacja, bądź też zastosowanie krańców przedziałów najbliższych kwantylom najbardziej efektywnym do wzorów (11).

Tablica 1 przedstawia wyniki, które otrzymano szacując parametry rozkładów logarytmiczno-normalnych dwuparametrowych metodą kwantyli, na podstawie empirycznych rozkładów płac opublikowanych przez WUS w Łodzi (1987). Rozkłady płac były badane według poszczególnych działów gospodarki narodowej z uwzględnieniem podziału na pracowników fizycznych i umysłowych.

T a b l i c a 1

Oceny parametrów rozkładu logarytmiczno-normalnego
w poszczególnych działach gospodarki narodowej dla pracowników
fizycznych i umysłowych

Działy gospodarki narodowej	m			s		
	ogółem	pracownicy		ogółem	pracownicy	
		fizyczni	umysłowi		fizyczni	umysłowi
1	2	3	4	5	6	7
Razem	10,0163	10,0217	10,0057	0,3362	0,3463	0,3230
Sfera produkcji material- nej	10,0664	10,0686	10,0674	0,2995	0,3014	0,3071
przemysł	10,0990	10,0891	10,1351	0,2799	0,2799	0,2719
budownictwo	10,1295	10,1071	10,1809	0,3191	0,3088	0,3347
rolnictwo	10,1344	10,0937	10,1686	0,2922	0,3046	0,2585
leśnictwo	9,9412	9,9012	9,9967	0,2077	0,2117	0,1285
transport	10,0141	10,0293	9,9718	0,2606	0,2718	0,2320
łączyność	9,8930	9,8421	9,5491	0,2199	0,2144	0,1893
handel	9,8946	9,8476	9,9151	0,2710	0,3012	0,2748
Pozostałe branże produkcji materialnej	9,9111	9,8012	10,0065	0,2649	0,2964	0,2549
Sfera poza produkcją mate- rialną	9,8336	9,6465	9,9199	0,3612	0,3433	0,3334

Tablica 1 (cd.)

1	2	3	4	5	6	7
gospodarka mieszkaniowa	9,9674	9,9065	10,0728	0,2821	0,3063	0,2654
nauka i technika	10,1212	10,1235	10,1182	0,2736	0,2966	0,2513
oświata i wychowanie	9,7296	9,4957	9,8172	0,3194	0,2284	0,3206
kultura i sztuka	9,7489	9,6416	9,7911	0,2948	0,3053	0,3034
ochrona zdrowia i opieka społeczna	9,7857	9,5437	9,9060	0,3474	0,2717	0,3391
kultura fizyczna, turystyka i wypoczynek	10,0019	9,9065	10,0995	0,2435	0,2399	0,2015
Pozostałe branże usług niematerialnych	9,4872	9,4284	10,1155	0,2734	0,2125	0,2149
finanse i ubezpieczenia	10,0606	-	10,0226	0,3161	-	0,2610
administracja i wymiar sprawiedliwości	10,0230	-	10,0416	0,3265	-	0,3240

Źródło: Obliczenia własne.

Aby oszacowania te mogły być przydatne w dalszej analizie badanych rozkładów płac, należy ocenić zgodność rozkładów empirycznych z teoretycznymi.

4. Miara zgodności rozkładów

Można wyróżnić wiele miar badających stopień zgodności rozkładów empirycznych z teoretycznymi. Najprostszą metodą jest badanie różnic między charakterystykami obliczonymi dla rozkładu empirycznego a odpowiednimi oszacowaniami tych charakterystyk. Najczęściej porównuje się tak: średnią, medianę, dominantę, skrajne kwantyle i decyle.

Przyjmuje się, że jeśli obliczone w ten sposób błędy nie przekraczają 5% wartości oszacowanych parametrów (lub innych charakterystyk), to rozkład teoretyczny dostatecznie dokładnie opisuje rozkład empiryczny.

Wygodniejsze w zastosowaniu są syntetyczne mierniki zgodności rozkładów. Można tu wymienić często stosowane odchylenia standardowe różnic częstości względnych:

$$(13) \quad S = \sqrt{\frac{1}{k} \sum_{i=1}^k (f_i - \hat{f}_i)^2},$$

gdzie: f_i - częstość względna w grupie "i" rozkładu empirycznego, $\hat{f}_i$ - częstość względna w grupie "i" rozkładu teoretycznego, k - liczba grup.

Wskaźnik ten informuje, ile średnio różnią się od siebie częstości względne badanych rozkładów. Przyjmuje się, że zgodność rozkładów jest wysoka, gdy $0 < S < 0,005$, natomiast gdy $S > 0,01$, to odchylenia między nimi można uznać za znaczne.

Inną miarą zgodności jest współczynnik podobieństwa (Vielrose [5]). Liczy się go jako sumę mniejszych liczebności przy zestawieniu liczebności empirycznych i teoretycznych dla tych samych grup dochodowych. Wartość współczynnika podobieństwa waha się od 0 do 1, przy czym wartość 1 przyjmuje on, gdy rozkłady są całkowicie zgodne. Współczynnik podobieństwa jest bardzo wy-

Wybrane parametry charakteryzujące
empirycznych z teoretycznymi dla poszczególnych

Działy gospodarki narodowej	Ogółem		
	odchylenie standardo- we różnic	współczynnik	
		deter- minacji	korelacji
Razem	0,0048	0,9832	0,9924*
Sfera produkcji materialnej	0,034	0,9931	0,9968*
przemysł	0,0030	0,9920	0,9962*
budownictwo	0,0083	0,9567	0,9802*
rolnictwo	0,0223	0,5748	0,7967
leśnictwo	0,0475	0,4566	0,6996
transport	0,0047	0,9901	0,9953*
łączność	0,0126	0,9523	0,9783*
handel	0,0086	0,9543	0,9818*
Pozostałe branże produkcji materialnej	0,0239	0,7948	0,8925*
Sfera poza produkcją materialną	0,0023	0,9942	0,9972*
gospodarka mieszkaniowa	0,0094	0,9494	0,9747*
nauka i technika	0,0130	0,9125	0,9572*
oświata i wychowanie	0,0100	0,8762	0,9446*
kultura i sztuka	0,0216	0,4104	0,7068
ochrona zdrowia i opieka społeczna	0,0063	0,9658	0,9839*
kultura fizyczna, turystyka i wypoczynek	0,0332	0,6369	0,7991
Pozostałe branże usług niematerialnych	0,0309	0,6724	0,8232
finanse i ubezpieczenia	0,0465	-0,0306	0,4313
administracja i wymiar sprawiedliwości	0,0124	0,8600	0,9289*

* Zgodność rozkładów jest zadowalająca.

Źródło: Jak w tabl. 1.

stopień zgodności rozkładów
działów gospodarki narodowej

Pracownicy fizyczni			Pracownicy umysłowi		
odchylenie standardo- we różnic	współczynnik		odchylenie standardo- we różnic	współczynnik	
	deter- minacji	korelacji		deter- minacji	korelacji
0,0081	0,9510	0,9767*	0,0035	0,9958	0,9911*
0,0062	0,9774	0,9894*	0,0062	0,9879	0,9756*
0,0059	0,9816	0,9911*	0,0127	0,9579	0,9174*
0,0092	0,9513	0,9774*	0,0103	0,9623	0,9188*
0,0298	0,0212	0,5560	0,0457	0,5568	0,3063
0,0901	-0,4758	-0,0201	0,0696	0,3678	0,0050
0,0045	0,9898	0,9853*	0,0130	0,9716	0,9411*
0,0217	0,8726	0,9426*	0,0291	0,7723	0,5917
0,0090	0,9306	0,9697*	0,0090	0,9782	0,9563*
0,0728	-0,2092	0,1747	0,0262	0,9089	0,8085*
0,0091	0,9134	0,9571*	0,0033	0,9960	0,9914*
0,0134	0,8729	0,9356*	0,0215	0,9039	0,8152*
0,0230	0,5155	0,7711	0,0142	0,9623	0,9177*
0,0130	0,9789	0,9672*	0,0110	0,9400	0,8808*
0,0317	-1,0119	-0,1042	0,0205	0,7384	0,4913
0,0149	0,8199	0,9072*	0,0114	0,9583	0,9116*
0,0304	0,6929	0,8355	0,0537	-0,0903	-0,5372
0,0298	0,8011	0,9022*	0,1475	-0,2867	-0,1404
-	-	-	0,0340	0,7432	0,5499
-	-	-	0,0121	0,9394	0,8824*

godny w zastosowaniu ze względu na swą prostotę, a także łatwy do interpretacji (określa część wspólną badanych rozkładów). Jest on jednak uzależniony od szczegółowości klasyfikacji i nie powinien być stosowany, gdy porównujemy szeregi o różnej liczbie przedziałów. Nie mogliśmy więc zastosować go w tym opracowaniu, gdyż różnice między empirycznymi szeregami płac są tu znaczne.

Wiele trudności przysparza także zastosowanie klasycznych testów zgodności (χ^2 i λ -kołmogorowa) do badania zgodności rozkładów płac. Testy te są przystosowane do badania zgodności rozkładów uzyskanych z próbek losowych, podczas gdy badania płac są na ogół badaniami pełnymi. Zbyt duża liczba jednostek wzięta do próby wpływa na to, że nawet minimalne odchylenia rozkładów empirycznych i teoretycznych uznawane są przez test jako istotne. Inne testy, pozwalające zweryfikować zgodność empirycznych rozkładów płac z rozkładem logarytmiczno-normalnym dla bardzo dużych prób, zaproponowane przez É l t é t ö [2] i S i l b e r m a n a [4] są bardziej złożone i na tym etapie badania nie zostały uwzględnione.

W tym opracowaniu zastosowaliśmy trzy wielkości charakteryzujące stopień zgodności rozkładów empirycznych z teoretycznymi: odchylenie standardowe różnic częstości względnych, a także ogólnie znane - współczynnik korelacji r -Pearsona (pomiędzy f_i a $\hat{f}_i$) oraz współczynnik determinacji R^2 . Wielkości te zestawione są w tabl. 2.

Po przeanalizowaniu wszystkich miar zgodności (zob. tabl. 2) zauważamy, że dopasowanie rozkładów płac ogółem jest podobne do dopasowania rozkładów płac w podziale na pracowników fizycznych i umysłowych. Należy jednak dodać, że nieco gorsze wyniki uzyskano dla pracowników fizycznych. W większości przypadków (ok. 2/3) dopasowanie zaobserwowanych i oszacowanych rozkładów płac jest zadowalające. Słabą lub bardzo słabą zgodność zauważamy w następujących działach gospodarki narodowej: rolnictwo, leśnictwo, kultura i sztuka, kultura fizyczna, turystyka i wypoczynek oraz finanse i ubezpieczenia.

Przyczyną takiego stanu wydaje się być małe zróżnicowanie kwalifikacji i stanowisk w tych dziedzinach. Wśród pracowników fizycznych niska zgodność występuje jeszcze w nauce i technice oraz

w pozostałych branżach produkcji materialnej, natomiast wśród pracowników umysłowych w pozostałych branżach usług niematerialnych.

Jakkolwiek porównanie kilku mierników zgodności rozkładów pozwala ocenić przydatność danego rozkładu teoretycznego do opisywania rozkładów płac, to jednak dla jednoznacznej oceny konieczne jest stosowanie odpowiednich testów zgodności. Przy pomocy tych testów można wykazać, czy zaobserwowane różnice między rozkładem empirycznym i teoretycznym są istotne (przy danym poziomie istotności α), czy też są losowe. Dlatego też w dalszych badaniach dotyczących zgodności rozkładów płac należy wziąć pod uwagę także wyniki tych testów.

5. Analiza wyników

Wykazując zgodność empirycznych rozkładów płac z rozkładem logarytmiczno-normalnym uzyskujemy znaczne korzyści przy analizie wyników, tym bardziej, że niektóre charakterystyki rozkładów teoretycznych posiadają interpretacje ekonomiczne.

Dużym ułatwieniem jest także fakt, że większość charakterystyk rozkładu logarytmiczno-normalnego zależy od jednego parametru σ . Stwarza to możliwość ich stabilizowania (K o r d o s [3], s. 260-261) i znacznie przyspiesza analizę.

Do charakterystyk rozkładu logarytmiczno-normalnego zależnych jedynie od σ należą:

- współczynnik zmienności

$$(14) \quad v = \sqrt{e^2 - 1},$$

- współczynnik skośności

$$(15) \quad q_1 = v^3 + 3v,$$

- współczynnik spłaszczenia

$$(16) \quad g_2 = v^8 + 6v^6 + 15v^4 + 16v^2,$$

- stosunek średniej do mediany

$$(17) \quad \frac{\alpha}{M_e} = e^{\frac{1}{2} \sigma^2},$$

-- stosunek średniej do dominanty

$$(18) \quad \frac{\alpha}{D_o} = e^{\frac{3}{2} \sigma^2},$$

- frakcja obserwacji o wartościach nie większych od średniej

$$(19) \quad N\left(\frac{\sigma}{2}; 0,1\right),$$

- współczynnik koncentracji Lorenza

$$(20) \quad L = 2N\left(\frac{\sigma}{2}; 0,1\right) - 1,$$

gdzie: $N(u; 0,1)$ - dystrybuanta rozkładu normalnego o parametrach $(0,1)$.

Podstawowe mierniki charakteryzujące rozkłady płac w gospodarce społecznej we wrześniu 1986 r. zostały zamieszczone w tabl. 3, 4, 5.

Analizując te tablice zauważamy, że wśród pracowników fizycznych, umysłowych oraz wziętych razem średnia arytmetyczna i mediana były najwyższe w budownictwie (najwyższe wartości występowały wśród pracowników umysłowych - średnia arytmetyczna wynosiła 27 915,72 tys. zł, natomiast mediana 26 394,44 tys. zł). Dominanty dla pracowników fizycznych i umysłowych były także najwyższe w budownictwie, jednakże dla pracowników fizycznych i umysłowych wziętych razem najwyższa wartość wystąpiła w "nauce i technice". Najniższe wartości średnie dla pracowników umysłowych i wziętych razem zauważamy w "oświacie i wychowaniu". Dla pracowników fizycznych najniższe płace obserwujemy w "pozostałych branżach usług niematerialnych" a zaraz potem w "oświacie i wychowaniu". W gałęziach tych od dawna występują bardzo niskie płace, co może budzić wiele zastrzeżeń. Najwyższy udział płac poniżej średniej

występował w "ochronie zdrowia i opiece społecznej" i wynosił 57%. Świadczy to o złej sytuacji płacowej w tej gałęzi, tym bardziej, że miary średnie należały tu do najniższych (średnia arytmetyczna wynosiła 18 982,88 tys. zł).

Przechodząc do przeciętnego zróżnicowania płac w stosunku do średniej arytmetycznej łatwo zauważyć, że płace pracowników fizycznych były bardziej zróżnicowane niż płace pracowników umysłowych. Charakterystyczny jest także fakt, że zróżnicowanie płac było znacznie wyższe w działach poza produkcją materialną.

Do pełnej analizy zróżnicowania płac należałoby oczywiście zastosować także inne miary, jak: współczynnik zróżnicowania decylowego, który informuje o skali rozpiętości relatywnie wysokich i niskich wynagrodzeń, wskaźnik wahanía kwantylowego, który bada zróżnicowanie 50% środkowych wartości płac lub też wskaźnik wahanía decylowego informujący o zróżnicowaniu w grupach 80% środkowych wartości płac. Należy podkreślić, że badanie zróżnicowania płac ma bardzo duże znaczenie z punktu widzenia ich funkcji ekonomicznych. Płace mało zróżnicowane tracą funkcje motywacyjne, co prowadzi do wytworzenia się mechanizmu samoograniczającego wzrost wydajności pracy. Z miarami dyspersji silnie związane są miary koncentracji płac.

Na podstawie tabl. 3, 4, 5 możemy ocenić kształtowanie się współczynników koncentracji Lorenza oraz współczynników spłaszczenia. Współczynnik koncentracji Lorenza charakteryzuje stopień nierównomierności podziału ogólnego funduszu płac pomiędzy poszczególnych pracowników, natomiast współczynnik spłaszczenia mierzy koncentrację płac wokół średniej arytmetycznej. Obie miary są zgodne co do tego, że najbardziej skoncentrowane są płace w "budownictwie" a zaraz potem w "gospodarce mieszkaniowej". Płace pracowników fizycznych są bardziej skoncentrowane niż płace pracowników umysłowych (miara Lorenza dla pracowników fizycznych wynosi 0,195 natomiast dla umysłowych 0,179).

Bardziej szczegółowa analiza badanych rozkładów płac nie mieści się w ramach tego opracowania. Chcielibyśmy tylko zasygnalizować szerokie możliwości wykorzystania charakterystyk rozkładu logarytmiczno-normalnego do badania struktury płac.

Charakterystyki rozkładu logarytmiczno-normalnego

Gałęzie gospodarki narodowej	Średnia arytmetyczna	Odchylenie standardowe	Mediana	Dominanta
Razem	23 689,07	8 194,11	22 387,58	19 995,18
Sfera produkcji materialnej				
przemysł	24 619,75	7 542,80	23 539,76	21,519,83
budownictwo	25 288,97	7 218,70	24 317,66	22 485,52
rolnictwo*	26 380,23	8 635,90	25 071,02	22 644,32
leśnictwo*	26 293,95	7 850,85	25 194,42	23 132,13
transport	21 220,67	4 455,31	20 767,89	19 891,10
łączność	23 110,18	6 126,65	22 338,52	20 871,65
handel	20 276,36	4 512,56	19 792,13	18 858,09
Pozostałe branże produkcji materialnej	20 564,99	5 676,65	19 823,62	18 420,10
Sfera poza produkcją materialną	21 509,41	5 800,39	20 767,55	19 359,70
gospodarka mieszkaniowa	19 907,41	7 430,59	18 650,55	16 369,88
nauka i technika	22 184,47	6 385,11	21 319,01	19 688,05
oświata i wychowanie	25 813,68	7 195,85	24 865,57	23 072,64
kultura i sztuka*	17 687,42	5 795,93	16 808,02	15 178,20
ochrona zdrowia i opieka społeczna	17 895,53	5 391,56	17 134,76	15 708,88
kultura fizyczna, turystyka i wypoczynek*	18 882,88	6 762,08	17 777,37	15 756,72
Pozostałe branże usług niematerialnych*	22 732,84	5 618,68	22 068,79	20 798,26
finanse i ubezpieczenia*	13 691,79	3 813,85	13 189,66	12 239,96
administracja i wymiar sprawiedliwości	24 601,61	7 975,64	23 402,52	21 176,87
	23 773,16	7 972,29	22 539,54	20,261,02

* Wiarygodność oszacowania niewystarczająca do interpretacji pirycznych z teoretycznymi).

^a Zastosowano współczynnik skośności $g_1 = \frac{e_3}{S(x)^3}$, gdzie $e^3 =$

Ź r ó d ł o: Jak w tabl. 1.

w poszczególnych działach gospodarki narodowej

Współczynnik			Stosunek średniej do		Miara Lorenza	Frakcja poniżej średniej
zmienności	skośności ^a	spłaszczenia	mediany	dominanty		
0,3459	1,0791	2,1396	1,0581	1,1847	0,190	0,568
0,3064	0,9479	1,6390	1,0459	1,1440	0,168	0,560
0,2854	0,8796	1,4066	1,0399	1,1247	0,157	0,556
0,3274	1,0172	1,8944	1,0522	1,1650	0,179	0,564
0,2986	0,9224	1,5500	1,0436	1,1367	0,163	0,560
0,2100	0,6391	0,7349	1,0218	1,0668	0,118	0,542
0,2651	0,8140	1,2007	1,0345	1,1073	0,146	0,552
0,2226	0,6787	0,8300	1,0245	1,0752	0,124	0,544
0,2760	0,8491	1,3089	1,0374	1,1164	0,151	0,554
0,2697	0,8286	1,2452	1,0357	1,1110	0,146	0,552
0,3733	1,1718	2,5369	1,0674	1,2161	0,201	0,571
0,2878	0,8873	1,4318	1,0406	1,1268	0,157	0,556
0,2788	0,8579	1,3368	1,0381	1,1188	0,151	0,554
0,3277	1,0182	1,8986	1,0523	1,1653	0,179	0,564
0,3013	0,9312	1,5805	1,0444	1,1392	0,163	0,558
0,3581	1,1202	2,3114	1,0622	1,1984	0,195	0,570
0,2472	0,7566	1,0348	1,0301	1,0930	0,135	0,540
0,2785	0,8573	1,3346	1,0381	1,1186	0,151	0,554
0,3242	1,0066	1,8544	1,0512	1,1617	0,179	0,564
0,3359	1,0438	1,9977	1,0547	1,1733	0,185	0,566

merytorycznej (wynika to ze zbyt niskiej zgodności rozkładów em-
trzeci moment centralny, $S(x)$ - odchylenie standardowe.

Charakterystyki rozkładu logarytmiczno-normalnego w poszczególnych

Dział gospodarki narodowej	Średnia arytmetyczna	Odchylenie standardowe	Mediana	Domnanta
Razem	23 318,15	7 739,08	22 151,97	19 957,40
Sfera produkcji materialnej	24 701,67	7 768,54	23 563,15	21 442,25
przemysł	26 161,30	7 247,68	25 211,69	23 414,61
budownictwo	27 915,42	9 612,41	16 394,44	23 596,59
rolnictwo*	26 957,91	7 087,02	26 072,02	24 386,61
leśnictwo*	22 134,79	2 855,99	21 952,81	21 593,32
transport	21 998,67	5 173,74	21 414,40	20 292,02
łącznie*	14 285,58	2 729,11	14 031,82	13 537,75
handel	21 012,46	5 886,05	20 233,61	18 761,43
Pozostałe branże produkcji materialnej	22 902,07	5 932,81	22 170,25	20 776,81
Sfera poza produkcją materialną	21 493,98	7 370,54	20 331,80	18 192,56
gospodarka mieszkaniowa	24 539,70	6 629,84	23 690,34	22 078,79
nauka i technika	25 585,20	6 533,33	24 789,74	23 272,24
oświata i wychowanie	19 313,73	6 353,64	18 346,49	16 554,90
kultura i sztuka*	18 715,18	5 811,63	17 873,26	16 301,34
ochrona zdrowia i opieka społeczna	21 237,37	7 412,94	20 050,99	17 873,36
kultura fizyczna, turystyka i wypoczynek*	24 830,57	5 055,13	24 331,46	23 363,13
Pozostałe branże usług niematerialnych*	25 300,35	5 501,49	24 722,61	23 606,43
finanse i ubezpieczenia*	23 310,23	6 189,95	22 529,43	21 045,41
administracja i wymiar sprawiedliwości	24 198,67	8 051,74	22 960,99	20 672,31

* Wiarygodność oszacowania niewystarczająca do interpretacji rzeczywistych z teoretycznymi).

^a Zastosowano współczynnik skośności $g_1 = \frac{e_3}{S(x)^3}$, gdzie e_3 -

Ź r ó d ł o: Jak w tabl. 1.

działach gospodarki narodowej dla pracowników umysłowych

Współczynnik			Stosunek średniej do		Miara Lorenza	Fracja poniżej średniej
zmienności	skośności ^a	spłaszczenia	mediany	dominanty		
0,3316	1,0313	1,9485	1,0535	1,1694	0,179	0,564
0,3145	0,9746	1,7352	1,0483	1,1520	0,174	0,562
0,2770	0,8524	1,3191	1,0377	1,1173	0,151	0,554
0,3443	1,0739	2,1182	1,0576	1,1830	0,185	0,566
0,2629	0,8068	1,1794	1,0340	1,1054	0,146	0,552
0,1290	0,3892	0,2706	1,0083	1,0251	0,073	0,526
0,2352	0,7186	0,9319	1,0273	1,0841	0,129	0,546
0,1910	0,5801	0,6042	1,0181	1,0552	0,107	0,538
0,2801	0,8623	1,3508	1,0385	1,1200	0,151	0,554
0,2591	0,7945	1,1431	1,0330	1,1023	0,140	0,550
0,3429	1,0691	2,0988	1,0572	1,1815	0,185	0,566
0,2702	0,8302	1,2501	1,0350	1,1115	0,151	0,554
0,2554	0,7827	1,1088	1,0321	1,0994	0,140	0,550
0,3290	1,0225	1,9150	1,0527	1,1666	0,179	0,564
0,3105	0,9615	1,6878	1,0471	1,1481	0,168	0,560
0,3491	1,0897	2,1831	1,0592	1,1882	0,190	0,568
0,2036	0,6192	0,6893	1,0205	1,0628	0,113	0,540
0,2174	0,6626	0,7907	1,0234	1,0718	0,118	0,542
0,2655	0,8154	1,2050	1,0347	1,1076	0,146	0,552
0,3327	1,0350	1,0635	1,0539	1,1706	0,179	0,564

merytorycznej (wynika to ze zbyt niskiej zgodności rozkładów empirycznych z teoretycznymi, $S(x)$ - odchylenie standardowe).

Charakterystyki rozkładu logarytmiczno-normalnego w poszczególnych

Działy gospodarki narodowej	Średnia arytmetyczna	Odchylenie standardowe	Mediana	Dominanta
Razem	23 901,05	8 532,24	22 509,77	19 965,45
Sfera produkcji materialnej	24 685,78	7 613,13	23 589,45	21 540,68
przemysł	25 040,37	7 149,91	24 078,30	22 263,65
budownictwo	25 712,48	8 132,95	24 515,35	22 285,71
rolnictwo*	25 338,82	7 900,22	24 190,32	22 047,14
leśnictwo*	20 406,95	4 368,61	19 954,83	19 080,42
transport	23 533,91	6 517,00	22 680,35	21 065,00
łączność	19 245,62	4 173,50	18 808,46	17 963,70
handel	19 789,72	6 098,13	18 912,19	17 272,13
Pozostałe branże produkcji materialnej*	18 865,39	5 716,14	18 054,81	16 536,63
Sfera poza produkcją materialną	16 406,23	5 801,83	15 467,55	13 748,22
gospodarka mieszkaniowa	21 023,01	6 593,13	20 059,67	18 263,39
nauka i technika*	26 041,68	7 896,10	24 921,28	22 823,01
oświata i wychowanie	13 654,06	3 159,24	13 302,63	12 626,65
kultura i sztuka*	16 126,91	5 040,66	15 392,46	14 022,51
ochrona zdrowia i opieka społeczna	14 481,85	4 008,59	13 957,04	12 963,77
kultura fizyczna, turystyka i wypoczynek*	20 646,26	5 025,83	20 060,46	18 938,25
Pozostałe branże usług niematerialnych	12 720,68	2 733,86	12 436,71	11 887,64

* Wiarygodność oszacowania niewystarczająca do interpretacji ryczynnych z teoretycznymi).

^a Zastosowano współczynnik skośności $g_1 = \frac{e_3}{S(x)^3}$, gdzie $e_3 =$

Ź r ó d ł o: Jak w tabl. 1.

T a b l i c a 5

działach gospodarki narodowej dla pracowników fizycznych

Współczynnik			Stosunek średniej do		Miara Lorenza	Fracja poniżej średniej
zmienności	skośności ^a	spłaszczenia	mediany	dominanty		
0,3570	1,1164	2,2953	1,0618	1,1971	0,195	0,570
0,3084	0,9545	1,6627	1,0465	1,1460	0,168	0,560
0,2855	0,8798	1,4071	1,0400	1,1247	0,157	0,556
0,3163	0,9806	1,7570	1,0488	1,1538	0,174	0,562
0,3118	0,9657	1,7027	1,0475	1,1493	0,168	0,560
0,2141	0,6520	0,7653	1,0227	1,0695	0,118	0,542
0,2769	0,8520	1,3179	1,0376	1,1172	0,151	0,544
0,2169	0,6608	0,7862	1,0232	1,0714	0,118	0,542
0,3081	0,9537	1,6597	1,0464	1,1458	0,168	0,560
0,3030	0,9368	1,6000	1,0449	1,1408	0,168	0,560
0,3536	1,1051	2,2475	1,0607	1,1933	0,190	0,568
0,3136	0,9717	1,7246	1,0480	1,1511	0,174	0,562
0,3032	0,9375	1,6025	1,0450	1,1410	0,168	0,560
0,2414	0,7065	0,9005	1,0264	1,0814	0,129	0,546
0,3126	0,9682	1,7120	1,0477	1,1501	0,174	0,562
0,2768	0,8516	1,3167	1,0376	1,1172	0,151	0,554
0,2434	0,7447	1,0020	1,0292	1,0902	0,135	0,548
0,2149	0,6547	0,7716	1,0228	1,0701	0,118	0,542

merytorycznej (wynika to ze zbyt niskiej zgodności rozkładów empi-

- trzeci moment centralny, $S(x)$ - odchylenie standardowe.

6. Uwagi końcowe

Wyniki przedstawione w niniejszym artykule potwierdzają dotychczasowe wnioski o celowości kontynuowania badań nad aproksymacją płac przy pomocy rozkładu logarytmiczno-normalnego. Większość badanych rozkładów płac wykazało wysoką zgodność z rozkładem teoretycznym. W przyszłych badaniach należałoby zwrócić uwagę na zastosowanie innych metod aproksymacji, co mogłoby zwiększyć stopień zgodności rozkładów empirycznych z teoretycznymi - dotyczy to szczególnie działów, w których zgodność ta była słaba. Celowym wydaje się także szukanie odpowiedzi na pytanie, jaki rozkład teoretyczny mógłby opisać te rozkłady płac, które charakteryzowały się całkowitym brakiem zgodności z rozkładem logarytmiczno-normalnym.

Literatura

- [1] A i t c h i s o n J., B r o w n J. A. C. (1957): The Log-normal Distribution, Cambridge.
- [2] E l t é t ö O. (1965): Large-Sample Lognormality Tests Based on New Inequality Measures, 35 Sses. Inter. Stat. Inst., Belgrad.
- [3] K o r d o s J. (1973): Metody analizy i prognozowania rozkładów płac i dochodów ludności, Warszawa.
- [4] S i l b e r m a n I. H. (1963): Application of the Log-normal Distribution to Industrial Concentration, Econometrica, 31, 745-746.
- [5] V i e l r o s e E. (1960): Rozkład dochodów według wielkości, Warszawa.
- [6] W i ś n i e w s k i J. (1934): Rozkład dochodów według wysokości, Warszawa.
- [7] Zatrudnienie w gospodarce współczesnej według wysokości wynagrodzenia miesięcznego we wrześniu 1986 roku, WOS (1987), Łódź.

Alina Jędrzejczak

ESTIMATION OF APPROXIMATION OF WAGES DISTRIBUTION
IN THE SOCIALIZED SECTOR BY THEIR AMOUNT USING
LOGARITHMIC NORMAL DISTRIBUTION

The article shows approximation of wages distribution in the socialized sector in 1986 by means of logarithmic normal distribution. Most examined wages distributions proved high consistency with the logarithmic normal distributions.