DYNAMICZNE MODELE EKONOMETRYCZNE
VIII Ogólnopolskie Seminarium Naukowe, 9-11 września 2003 w Toruniu
Katedra Ekonometrii i Statystyki, Uniwersytet Mikołaja Kopernika w Toruniu

Jerzy Czesław Ossowski
Politechnika Gdańska

Rozkład logarytmiczno-normalny a względne i absolutne miary rozproszenia

1. Średnia arytmetyczna i geometryczna a rozkład logarytmiczno-normalny

Większość zmiennych ekonomicznych przyjmuje jedynie wartości dodatnie. Wiele z nich scharakteryzować możemy za pomocą rozkładu logarytmiczno-normalnego. Uznajmy, że zmienna losowa y należy do tej grupy zmiennych. Oznacza to, że logarytm naturalny tej zmiennej ma rozkład normalny, tym samym funkcja gęstości prawdopodobieństwa dana jest wzorem (Aitchison, Brown (1957) s.8, Pawłowski (1969) s. 170-172):

$$f(\ln y) = \left(\sigma_{\ln y}\sqrt{2\pi}\right)^{-1} e^{-(\ln y - \mu_{\ln y})^2 / 2\sigma^2_{\ln y}} \tag{1}$$

w którym parametry wartości oczekiwanej i wariancji zmiennej lny definiujemy następująco:

$$\mu_{\ln y} = E \ln y, \tag{2}$$

$$\sigma^2_{\ln y} = E(\ln y - \mu_{\ln y})^2. \tag{3}$$

Odchylenie standardowe jest równe:

$$\sigma_{\ln y} = \sqrt{E(\ln y - \mu_{\ln y})^2}. \tag{4}$$

Wykorzystując fakt, że dlny=dy/y, funkcję gęstości prawdopodobieństwa logarytmu zmiennej y przekształcić możemy w funkcję gęstości zmiennej y o następującej postaci:

$$f(\ln y) = \left(y\sigma_{\ln y}\sqrt{2\pi}\right)^{-1} e^{-(\ln y - \mu_{\ln y})^2 / 2\sigma^2_{\ln y}} \tag{5}$$

Funkcja ta wyznacza krzywą asymetryczną. Asymetria ta jest prawostronna i jej wielkość zależy od wartości oczekiwanej i wariancji logarytmu zmiennej y.

Tym samym wartość oczekiwana zmiennej y (Ey) oraz jej dominanta (Dy) i mediana (My) nie pokrywają się i wynoszą odpowiednio (patrz: Aitchison, Brown (1957), s 8-9):

$$Ey = \mu_y = e^{\mu_{\ln y} + \frac{1}{2}\sigma^2_{\ln y}}, \tag{6}$$

$$My = e^{\mu_{\ln y}}, \tag{7}$$

$$Dy = e^{\mu_{\ln y} - \sigma^2_{\ln y}}. \tag{8}$$

Z powyższego wynika, że Dy < My < Ey.
Można ponadto wykazać, że wariancja i odchylenie standardowe zmiennej y wynoszą odpowiednio:

$$\sigma_y^2 = E(y - Ey)^2 = e^{2\mu_{\ln y} + \sigma^2_{\ln y}}(e^{\sigma^2_{\ln y}} - 1), \tag{9}$$

$$\sigma_y = \sqrt{E(y - Ey)^2} = e^{\mu_{\ln y} + \frac{1}{2}\sigma^2_{\ln y}} \sqrt{e^{\sigma^2_{\ln y}} - 1}. \tag{10}$$

Przedstawione powyżej właściwości rozkładu logarytmiczno-normalnego zmiennej y wykorzystali Murti i Sastri (1957) przy formułowaniu związków pomiędzy średnią arytmetyczną i średnią geometryczną tej zmiennej losowej. Zgodnie z poczynioną przez nich umową, wartość oczekiwaną logarytmu zmiennej y oznaczymy w sposób następujący:

$$\mu_{\ln y} = \ln g. \tag{11}$$

Obecnie zgodnie z koncepcją wspomnianych autorów zdefiniujemy średnią geometryczną zmiennej y w sposób następujący:

$$g = e^{\mu_{\ln y}} = e^{E \ln y}. \tag{12}$$

Oznacza to, że średnia geometryczna zmiennej y jest zdelogarytmowaną wartością nadziei matematycznej logarytmu zmiennej y. Z drugiej strony Murti i Sastri określili średnią arytmetyczną zmiennej losowej y jako jej wartość oczekiwaną, tzn.:

$$a = \mu_y = Ey. \tag{13}$$

Wykorzystując (6) i (12) stwierdzamy, że:

$$Ey = e^{\mu_{\ln y} + \frac{1}{2}\sigma^2_{\ln y}} = e^{\mu_{\ln y}} e^{\frac{1}{2}\sigma^2_{\ln y}}.$$

W konsekwencji na podstawie (13) średnią arytmetyczną (12) zapiszemy następująco:

$$a = g\, e^{\frac{1}{2}\sigma^2_{\ln y}}. \tag{14}$$

Wykazany przez wspomnianych autorów związek funkcyjny pomiędzy średnią arytmetyczną i geometryczną zmiennej y charakteryzującej się rozkładem logarytmiczno-normalnym był swego czasu szeroko omawiany w literaturze poświęconej modelom multiplikatywnym. Funkcja ta posiada liczące się walory poznawcze i praktyczne, czemu wiele uwagi poświęcił L.R.Klein (1965 s.155-156, 221-224). Kończąc tę część rozważań zauważmy, że w przypadku zmiennej charakteryzującej się rozkładem logarytmiczno-normalnym średnia geometryczna i mediana pokrywają się ze sobą.

Przykład 1.

Niech płace miesięczne (y) pracowników pewnego sektora gospodarczego charakteryzują się rozkładem logarytmiczno normalnym. Płace wyrażone są w złotych. Załóżmy, że wartość oczekiwana i wariancja logarytmów płac równają się odpowiednio:

$$\mu_{\ln y} = 7{,}49554,$$

$$\sigma^2_{\ln y} = 0{,}16.$$

Określić: **1)** medianę (średnią geometryczną), **2)** wartość oczekiwaną (średnią arytmetyczną), **3)** dominantę płac.

Ad 1) Zgodnie z (7) i (12) otrzymujemy:

$$My = g = e^{\mu_{\ln y}} = e^{7{,}49554} = 1800 \text{ zł}$$

Zauważmy, ze logarytmy płac są wielkościami niemianowanymi. Ich delogarytmy wyrażone są w jednostkach pierwotnych. Obecnie powiemy, że średnia geometryczna płac, będąca medianą, wynosiła 1800 złotych. Oznacza to, że 50% pracowników zatrudnionych w sektorze przedsiębiorstw zarabia mniej, niż 1800 złotych. Tym samym płace 50% pracowników przekraczają 1800 złotych.

Ad 2) Zgodnie z (6) a tym samym (15) mamy:

$$Ey = a = g\, e^{\frac{1}{2}\sigma^2_{\ln y}} = 1800 e^{\frac{1}{2}0{,}16} = 1950 \text{ zł}$$

Powiemy, że średnie (arytmetyczne) miesięczne wynagrodzenie pracowników sektora przedsiębiorstw wynosiło 1950 złotych.

Ad 3) Wartość dominująca zgodnie z (8) wynosi

$$Dy = e^{\mu_{\ln y} - \sigma^2_{\ln y}} = 1800 e^{-0{,}16} = 1533{,}9 \text{ zł}$$

Oznacza to, że płace oscylujące wokół wartości 1533,9 zł były najczęściej spotykane.

Przy okazji zauważmy, że odchylenie standardowe logarytmu płac jest równe

$$\sigma_{\ln y} = \sqrt{E(\ln y - \mu_{\ln y})^2} = \sqrt{0{,}16} = 0{,}4$$

Powiemy więc, że przeciętne standardowe odchylenie logarytmu płac od wartości oczekiwanej logarytmu płac (logarytmu średniej geometrycznej płac) wynosi 0,4. Z uwagi na fakt, że zmienna losowa wyrażona jest w logarytmach jest ona tym samym niemianowana.

2. Względne i absolutne rozproszenie zmiennej losowej w relacji do średniej geometrycznej

Zastanówmy się obecnie nad możliwością określenia prawdopodobieństwa tego, że zmienna losowa y, charakteryzująca się rozkładem logarytmiczno-normalnym, przyjmie wartości z określonego przedziału. Jak pisał Z. Pawłowski, w praktyce wygodnie jest „wykorzystać fakt, że logarytm zmiennej losowej Y ma rozkład normalny N(μ,σ). Prawdopodobieństwo tego, że $a \le Y \le b$, jest równoważne prawdopodobieństwu tego, że $\ln a \le \ln Y \le \ln b$, a to ostatnie łatwo jest obliczyć korzystając z tablic dystrybuanty standaryzowanego rozkładu normalnego (Pawłowski (1969) s.172).” Z powyższego wynika, że ostatnią nierówność zapisać możemy równoważnie w sposób następujący: $e^{\ln a} \le e^{\ln Y} \le e^{\ln b}$. Obecnie zgodnie z regułą trzech sigm powiemy, że prawdopodobieństwo tego, iż logarytm zmiennej losowej y przyjmie wartość różniącą się od średniej logarytmu tej zmiennej o jedno odchylenie standardowe jest równe 0,6826, co zapiszemy następująco:

$$P(-\sigma_{\ln y} \le \ln y - \mu_{\ln y} \le \sigma_{\ln y}) = 0{,}6826\,, \tag{15}$$

lub z uwagi na (11):

$$P(-\sigma_{\ln y} \le \ln y - \ln g \le \sigma_{\ln y}) = 0{,}6826\,. \tag{16}$$

Jak wiemy odchylenie standardowe jest miarą rozproszenia zmiennej losowej wokół jej wartości przeciętnej. Jeśli zmienna losowa jest wyrażona w jednostkach rzeczywistych, tzn. nietrasformowanych, to odchylenie standardowe jest miarą zróżnicowania absolutnego. W tej sytuacji zwyczajowo interpretujemy odchylenie standardowe, jako przeciętne, standardowe odchylenie zmiennej losowej od jej wartości oczekiwanej wyrażone w jednostkach analizowanej zmiennej. Jest rzeczą oczywistą, że jeśli zmienna wyrażona jest w kilogramach to odchylenie standardowe wyrażone jest również w kilogramach, itp. Z inną nieco sytuacją mamy do czynienia w przypadku, gdy zmienna losowa i jej charakterystyki wyrażone są w logarytmach. W tej sytuacji zarówno zmienna losowa jak i odchylenie standardowe stają się jednostkami niemianowanymi. W dalszym jednak ciągu odchylenie standardowe pozostaje miarą rozproszenia. Oznacza to, że interpretując odchylenie standardowe, jako przeciętne, standardowe odchylenie logarytmu zmiennej y od wartości oczekiwanej logarytmu tej zmiennej, mamy na myśli fakt, iż dwie trzecie tych odchyleń zawierać się będzie w przedziale wyznaczonym zgodnie z (16), natomiast blisko jedna trzecia wykraczać będzie poza ten przedział. Mimo, iż takie rozumowanie odchylenia standardowego może być satysfakcjonujące w sensie statystycznym, ma jednak prawo nie zadowalać nas w sensie interpretacyjnym. Nie myślimy bowiem w kategoriach logarytmów poszczególnych zmiennych, chociaż możemy rozumieć ich istotę. Dlatego celem wzbogacenia interpretacji otrzymanych wyników przekształćmy (16) do postaci:

$$P(e^{-\sigma_{\ln y}} \le \frac{y}{g} \le e^{\sigma_{\ln y}}) = 0{,}6826 \tag{17}$$

Zauważmy, że zdefiniowana w następujący sposób zmienna:

$$v = y / g \tag{18}$$

wskazuje na stosunek zmiennej losowej y do jej średniej geometrycznej, tym samym określa udział tej zmiennej losowej w poziomie jej średniej geometrycznej. Wtedy kiedy zmienna losowa y przyjmie wartości mniejsze od średniej geometrycznej g, zmienna losowa v przyjmie wartości mniejsze od 1. W sytuacji odwrotnej, tzn. gdy zmienna losowa przyjmie wartości większe od parametru g, wówczas zmienna losowa v przyjmować będzie wartości większe od jedności. Oznacza to, że dla każdego odchylenia standardowego $\sigma_{\ln y}$, będącego dodatnim pierwiastkiem wariancji zmiennej losowej lny, spełnione są następujące nierówności:

$$0 < v_d = e^{-\sigma_{\ln y}} < 1 \tag{19}$$

$$v_g = e^{\sigma_{\ln y}} > 1 \tag{20}$$

Przy okazji zauważmy, że:

$$v_d \cdot v_g = e^{-\sigma_{\ln y}} e^{\sigma_{\ln y}} = e^0 = 1 \tag{21}$$

Powyższa właściwość jest o tyle istotna, iż średnia geometryczna zmiennej v jest równa jedności, co wynika z następującego faktu:

$$E(\ln y - \mu_{\ln y}) = E(\ln y - \ln g) = 0 \Rightarrow g_v = e^{E \ln v} = e^{E(\ln y - \ln g)} = e^0 = 1 \tag{22}$$

gdzie g_v jest średnią geometryczną zmiennej v.

Obecnie na podstawie (17) oraz po przyjęciu oznaczeń z (19) i (20) powiemy, że z prawdopodobieństwem równym 0,6826 udział zmiennej losowej y w jej średniej geometrycznej g mieścić się będzie w przedziale od v_d do v_g. Oznacza to, że dokonując interpretacji w myśl której, przeciętny udział zmiennej y w jej średniej geometrycznej waha się w granicach od v_d do v_g mamy na myśli fakt, iż jest to przeciętny udział w kategoriach standardowych, gdyż został on wyznaczony na bazie odchylenia standardowego logarytmu zmiennej y z wszelkimi wypływającymi z tego konsekwencjami stochastycznymi. Powiemy tym samym, że v_d i v_g są przeciętnymi, względnymi miarami rozproszenia zmiennej losowej y względem jej średniej geometrycznej.

Celem dalszego wzbogacenia interpretacji omawianej przez nas względnej miary rozproszenia, dokonajmy przekształcenia nierówności równoczesnej ujętej w (17) poprzez odjęcie stronami wartości 1, wykorzystując jednocześnie oznaczenia przyjęte w (19) i (20). W rezultacie tego działania otrzymujemy:

$$P(v_d - 1 \le \frac{y}{g} - 1 \le v_g - 1) = 0{,}6826 \tag{23}$$

Przemnażając powyższą nierówność stronami przez 100, otrzymany wynik wyrażamy w procentach, co zapiszemy następująco:

$$P[(v_d-1)100 \le (\frac{y-g}{g})100 \le (v_g-1)100] = 0{,}6826 \qquad (24)$$

W sensie standardowym przeciętnie, zmienna losowa y odchyla się od jej średniej geometrycznej w przedziale od (v_d-1)100% do (v_g-1)100%. W analizowanym przypadku odchylenia te będą zawierać się w wyznaczonych granicach dla 2/3 wszystkich przypadków. Można zadać pytanie, dlaczego wyznaczamy dolne i górne przedziały przeciętnych odchyleń, zamiast powiedzieć wprost, o ile procent przeciętnie zmienna y odchyla się od jej średniej geometrycznej? Odpowiedź jest prosta i wynika z asymetrii rozkładu logarytmiczno-normalnego. Można bowiem udowodnić, rozpisując w szereg Maclaurina wyrażenia (19) i (20), iż spełniona jest następująca nierówność:

$$(v_d-1)+(v_g-1) > 0 \qquad (25)$$

Zauważmy ponadto, że z reguły trzech sigm wynika, iż

$$P(e^{-2\sigma_{\ln y}} \le \frac{y}{g} \le e^{2\sigma_{\ln y}}) = 0{,}9545 \qquad (26)$$

$$P(e^{-3\sigma_{\ln y}} \le \frac{y}{g} \le e^{3\sigma_{\ln y}}) = 0{,}9973 \qquad (27)$$

W celu utrzymania się w przyjętej konwencji oznaczeń umówmy się, że:

$$v_{2d} = e^{-2\sigma_{\ln y}} \ \text{ i } \ v_{2g} = e^{2\sigma_{\ln y}} \qquad (28)$$

$$v_{3d} = e^{-3\sigma_{\ln y}} \ \text{ i } \ v_{3g} = e^{3\sigma_{\ln y}} \qquad (29)$$

Przykład 2

Na podstawie danych z przykładu 1 określić przedziały udziałów płac w ich średniej geometrycznej (medianie) realizowane z prawdopodobieństwem: **1)** P_1=0,6826, **2)** P_2=0,9545, **3)** P_3=0,9973.

Ad 1) Na podstawie (19) i (20) otrzymujemy:

$$v_d = e^{-\sigma_{\ln y}} = e^{-0,4} = 0{,}67$$

$$v_g = e^{\sigma_{\ln y}} = e^{0,4} = 1{,}49$$

Z uwagi na (17) powiemy, że w analizowanym przypadku prawdopodobieństwo tego, że udział płac w średniej geometrycznej (medianie) mieścić się będzie w przedziale od 0,67 do 1,49 wynosi 0,6826. Tym samym, zgodnie z (24) powiemy, że płace odchylają się od średniej geometrycznej (mediany) w przedziale od –33% do 49% w 2/3 przypadków.

Ad 2) Na podstawie (28) otrzymujemy:

$$v_{2d} = e^{-2\sigma_{\ln y}} = e^{-0,8} = 0{,}449$$

$$v_{2g} = e^{2\sigma_{\ln y}} = e^{0,8} = 2{,}225$$

Z uwagi na (26) powiemy, że w analizowanym przypadku prawdopodobieństwo tego, że udział płac w średniej geometrycznej (medianie) mieścić się będzie w przedziale od 0,449 do 2,225 wynosi 0,9545. Tym samym płace odchylają się od średniej geometrycznej (mediany) w przedziale –55,1% do 122,5% w około 95 przypadków na sto.

Ad 3) Obecnie na podstawie (29) stwierdzamy, że

$$v_{3d} = e^{-3\sigma_{\ln y}} = e^{-1,2} = 0,301$$

$$v_{3g} = e^{3\sigma_{\ln y}} = e^{1,2} = 3,32$$

Z uwagi na (27) powiemy, że w analizowanym przypadku z prawdopodobieństwem 0,9973 udział płac w średniej geometrycznej (medianie) mieścić się będzie w przedziale od 0,301 do 3,32. Oznacza to, że prawdopodobieństwo tego, iż płace odchylają się od średniej geometrycznej (mediany) w przedziale –69,88% do 232,01% wynosi 0,9973.

Należy sądzić, iż wcześniej przeprowadzone rozważania oraz otrzymane powyżej wyniki upoważniają do stwierdzenia, że w sensie standardowym przeciętny udział płac w ich średniej geometrycznej (medianie) waha się w granicach od 0,67 do 1,49, czyli płace przeciętnie odchylają się od ich średniej geometryczne (mediany) w przedziale od –33% do 49%.

Wnioski wypływające z przeprowadzonej powyżej analizy dotyczącej względnego rozproszenia zmiennej losowej y wokół jej średniej geometrycznej są szczególnie przydatne przy określeniu właściwości stochastycznych modeli multiplikatywnych. W praktyce statystycznej, w przypadku rozpatrywania zmiennych losowych o rozkładzie logarytmiczno-normalnym, szczególnie interesujące mogą być miary rozproszenia wyrażone w jednostkach absolutnych. Celem ich wyznaczenia przekształćmy (15) do następującej postaci:

$$P(\mu_{\ln y} - \sigma_{\ln y} \le \ln y \le \mu_{\ln y} + \sigma_{\ln y}) = 0,6826\,. \tag{30}$$

Po zdelogarytmowaniu stronami wyrażenia zapisanego w nawiasie otrzymujemy:

$$P(e^{\mu_{\ln y}} e^{-\sigma_{\ln y}} \le y \le e^{\mu_{\ln y}} e^{\sigma_{\ln y}}) = 0,6826\,, \tag{31}$$

co po uwzględnieniu przyjętych wcześniej oznaczeń zapiszemy w następujący sposób:

$$P(g \cdot v_d \le y \le g \cdot v_g) = 0,6826 \tag{32}$$

Obecnie powiemy, że prawdopodobieństwo tego, iż zmienna losowa y przyjmuje wartości w granicach od $g \cdot v_d$ do $g \cdot v_g$, jest równe 0,6826. Zauważmy, że zarówno zmienna losowa y jak i wyznaczone granice przedziałów wyrażone są w jednostkach rzeczywistych analizowanej zmiennej. Aby wyjaśnić istotę asymetrii wyznaczonej tutaj absolutnej miary rozproszenia zauważmy, że ponieważ logarytm zmiennej y ma rozkład normalny, więc spełniona musi być następująca równość:

$$P(\mu_{\ln y} - \sigma_{\ln y} \le \ln y \le \mu_{\ln y}) = P(\mu_{\ln y} \le \ln y \le \mu_{\ln y} + \sigma_{\ln y}) = 0,341,$$

co po zdelogarytmowaniu wyrażeń ograniczonych nawiasami i przyjęciu wcześniej przyjętych oznaczeń zapiszemy następująco:

$$P(g\cdot v_d \leq y \leq g) = P(g \leq y \leq g\cdot v) = 0{,}341. \qquad (33)$$

Z uwagi na (25) stwierdzamy, że:

$$|g\cdot v_d - g| \langle |g - g\cdot v_g| \qquad (34)$$

Oznacza to, że analizowane absolutne rozproszenie zmiennej y odnosi się do średniej geometrycznej (mediany) zmiennej y. Rozproszenie to charakteryzuje się tym, iż jednakowemu prawdopodobieństwu realizacji zdarzeń odpowiada, co do wartości bezwzględnej, mniejszy przedział dolny i większy przedział górny odchyleń zmiennej y od jej średniej geometrycznej (mediany).

Obecnie możemy powiedzieć, że przeciętne, w sensie standardowym, odchylenie zmiennej losowej y od jej średniej geometrycznej (mediany) waha się w granicach od $g\cdot v_d$ do $g\cdot v_g$. Jest to, jak się wydaje, w miarę poprawny sposób określenia przeciętnej, absolutnej miary rozproszenia zmiennej losowej y w stosunku do jej wartości średniej w sytuacji, gdy zmienna ta charakteryzuje się asymetrycznym rozkładem. Uzupełniając rozważania dotyczące absolutnego rozproszenia zmiennej losowej y względem jej średniej geometrycznej (mediany) zauważmy, że zgodnie z regułą trzech sigm otrzymujemy:

$$P(g\cdot v_{2d} \leq y \leq g\cdot v_{2g}) = 0{,}9545 \qquad (35)$$

$$P(g\cdot v_{3d} \leq y \leq g\cdot v_{3g}) = 0{,}9945 \qquad (36)$$

Przykład 3

Na podstawie danych z przykładu 1 i 2 wyznaczyć absolutne przedziały rozproszenia płac w stosunku średniej geometrycznej (mediany) realizowane z prawdopodobieństwem: **1)** P_1=0,6826, **2)** P_2=0,9545, **3)** P_3=0,9973. Wyniki zinterpretować charakteryzując równocześnie asymetrię badanych miar rozproszenia.

Na wstępie zauważmy, że średnia geometryczna płac (mediana płac) wynosi: g =1800 zł

Ad 1) Elementy zawarte w nawiasie w wyrażeniu (32) wynoszą odpowiednio:

$g\cdot v_d$ = 1800·0,67 = 1206 zł,

$g\cdot v_g$ = 1800·1,49 = 2682 zł.

Powiemy, że prawdopodobieństwo tego, iż płace pracowników odchylają się od ich średniej geometrycznej (mediany) w przedziale od 1206 zł do 2682 zł wynosi 0,6826. Z drugiej strony prawdopodobieństwo tego, iż płace będą zawarte w przedziale od 1206 zł do 1800 zł jest równe prawdopodobieństwu tego, iż płace będą zawarte w przedziale od 1800 zł do 2682 zł. W jednostkach absolutnych odchylenie dolne płac od ich średniej geometrycznej wynosi –594 zł, natomiast odchylenie górne jest równe 882 zł.

Ad 2) Elementy zawarte w nawiasie w wyrażeniu (35) wynoszą odpowiednio:

$g\cdot v_{2d}$ = 1800·0,449 = 808 zł,

$g\cdot v_{2g}$ = 1800·2,225 = 4005 zł.

Tak więc prawdopodobieństwo tego, iż płace pracowników odchylają się od ich średniej geometrycznej (mediany) w przedziale od 808 zł do 4005 zł wynosi

0,9545. Z drugiej strony prawdopodobieństwo tego, iż płace będą zawarte w przedziale od 808 zł do 1800 zł jest równe prawdopodobieństwu tego, iż płace będą zawarte w przedziale od 1800 zł do 4005 zł. W jednostkach absolutnych odchylenie dolne płac od ich średniej geometrycznej wynosi –992 zł, natomiast odchylenie górne jest równe 2205 zł.

Ad 3) Elementy zawarte w nawiasie w wyrażeniu (36) wynoszą odpowiednio:

$g \cdot v_{3d} = 1800 \cdot 0{,}301 = 542$ zł,

$g \cdot v_{3g} = 1800 \cdot 3{,}320 = 5976$ zł.

Tym samym prawdopodobieństwo tego, iż płace pracowników odchylają się od ich średniej geometrycznej (mediany) w przedziale od 542 zł do 5976 zł wynosi 0,9973. Z drugiej strony prawdopodobieństwo tego, iż płace będą zawarte w przedziale od 542 zł do 1800 zł jest równe prawdopodobieństwu tego, iż płace będą zawarte w przedziale od 1800 zł do 5976 zł. W jednostkach absolutnych odchylenie dolne płac od ich średniej geometrycznej wynosi –1258 zł, natomiast odchylenie górne jest równe 4176 zł.

3. Absolutne i względne rozproszenie zmiennej losowej w relacji do średniej arytmetycznej

Omawiane dotychczas miary rozproszenia odnosiły się do średniej geometrycznej. W prowadzonych rozważaniach pomijaliśmy miarę rozproszenia odnoszącą się do wartości oczekiwanej y, czyli średniej arytmetycznej tej zmiennej. Skoncentrowanie się na średniej geometrycznej uznać jednak należy za zasadne, z uwagi na fakt, że logarytm średniej geometrycznej jest jednocześnie wartością oczekiwaną w rozkładzie normalnym logarytmu zmiennej y. Powstaje jednak pytanie, jakimi właściwościami charakteryzuje się rozproszenia zmiennej y wokół jej średniej arytmetycznej? Zauważmy, że rozproszenie to mierzone odchyleniem standardowym (10) w relacji do wartości oczekiwanej zmiennej y (6) w sensie odległości jest symetryczne. Natomiast, z uwagi na charakter rozkładu logarytmiczno-normalnego, jest ono asymetryczne w sensie prawdopodobieństwa realizacji zdarzeń zachodzących w odchyleniu dolnym i górnym od średniej arytmetycznej. Z uwagi na fakt, że średnia arytmetyczna jest położona bardziej na prawo od średniej geometrycznej, będącej medianą w rozpatrywanym rozkładzie, musi zajść następująca nierówność:

$$P_d(\mu_y - \sigma_y < y < \mu_y) > P_g(\mu_y < y < \mu_y + \sigma_y) \tag{37}$$

gdzie występujące w nawiasach parametry rozkładu zmiennej y zdefiniowano w (6) i w (10), natomiast P_d i P_g są odpowiednio prawdopodobieństwami zajścia zdarzeń w przedziale dolnym i górnym odchyleń zmiennej losowej y od jej średniej arytmetycznej o wielkość odchylenia standardowego. Obecnie utrzymując oznaczenia w myśl których:

- $a = \mu_y$, jest średnią arytmetyczną zmiennej y,

- $\ln g = \ln \mu_{\ln y}$, jest logarytmem średniej geometrycznej zmiennej y,

przekształcimy (37) do pożądanej standaryzowanej postaci prawdopodobieństw w następujący sposób:

$$P_d(a-\sigma_y < y < a) > P_g(a < y < a+\sigma_y), \tag{38}$$

$$P_d[\ln(a-\sigma_y) < \ln y < \ln a) > P_g(\ln a < \ln y < \ln(a+\sigma_y)], \tag{39}$$

$$P_d(z_1 < z < z_0) > P_g(z_0 < z < z_2), \tag{40}$$

gdzie:

$$z = (\ln y - \ln g)/\sigma \ln y = (\ln(y/g))/\sigma_{\ln y} \tag{41}$$

$$z_0 = (na - \ln g)/\sigma_{\ln y} = (\ln a/g)/\sigma_{\ln y} \tag{42}$$

$$z_1 = (\ln(a-\sigma_y) - \ln g)/\sigma_{\ln y} = (\ln[(a-\sigma_y)/g])/\sigma_{\ln y} \tag{43}$$

$$z_1 = (\ln(a+\sigma_y) - \ln g)/\sigma_{\ln y} = (\ln[(a+\sigma_y)/g])/\sigma_{\ln y} \tag{44}$$

Na podstawie (38), (39), (40), (41) i (42) potrafimy określić prawdopodobieństwo zajścia zdarzeń opisanych przez (37).

Przykład 4

Na podstawie danych z przykładu 1 i 2 obliczyć i zinterpretować:

1. odchylenie standardowe płac (y) względem ich średniej arytmetycznej a = μ_y ,
2. prawdopodobieństwo tego, że płace nie będą mniejsze o wielkość jednego odchylenia standardowego od płacy wyznaczonej na poziomie średniej arytmetycznej,
3. prawdopodobieństwo tego, że płace nie będą większe o wielkość jednego odchylenia standardowego od płacy wyznaczonej na poziomie średniej arytmetycznej,
4. prawdopodobieństwo tego, że płace będą mniejsze (większe) od płacy wyznaczonej na poziomie średniej arytmetycznej,

Ad 1) Wiedząc, że średnia arytmetyczna płac a=1950 zł oraz wariancja logarytmu płac wynosi 0,16 odchylenie standardowe zmiennej y od jej średniej arytmetycznej wyznaczymy na podstawie (10) i (14). Wynosi ono:

$$\sigma_y = e^{\mu_{\ln y} + \frac{1}{2}\sigma^2_{\ln y}} \sqrt{e^{\sigma^2_{\ln y}} - 1} = 1950\sqrt{e^{0,16} - 1} = 812,27 \text{ zł.}$$

Wstępnie powiemy, że płace przeciętnie odchylają się od płacy średniej, wynoszącej 1950 zł, o około 812,27 zł.

Aby w pełni poprawnie zinterpretować odchylenie standardowe, w przypadku rozkładu asymetrycznego powinniśmy określić prawdopodobieństwo odchyleń dolnych i górnych

Przed określeniem prawdopodobieństw z punktów 2, 3 i 4 obliczamy zgodnie z (42), (43) i (44) wartości z_0, z_1, i z_2. Wartości te równają się odpowiednio:

$z_0 = [\ln(a/g)]/\sigma_{\ln y} = [\ln(1950/1800)]/0,4 = 0,20$

$z_1 = \{\ln[(a-\sigma_y)/g)]\}/\sigma_{\ln y} = \{\ln[(1950-812,27)/1800]\}/0,4 = -1,15$

$z_2 = \{\ln[(a+\sigma_y)/g)]\}/\sigma_{\ln y} = \{\ln[(1950+812,27)/1800]\}/0,4 = 1,07$

Ad 2) Postępując zgodnie z odpowiednią procedurą (patrz: [10] s. 91-93) wyznaczamy prawdopodobieństwo tego, że płace nie będą mniejsze o jedno odchylenie standardowe od płacy wyznaczonej na poziomie średniej arytmetycznej. Prawdopodobieństwo to wynosi:

P_d (a-σ_y < y < a) = P_d (z_1 <z < z_0) = P_d(-1,15<z<0,20) = 0,3749 + 0,0793 = 0,454

Powiemy, że prawdopodobieństwo tego, że płace nie będą mniejsze o 812,27 zł (odchylenie standardowe) od 1950 zł (średnia arytmetyczna) jest równe 0,454.

Ad 3) Postępując zgodnie z odpowiednią procedurą (patrz: Kmenta (1990) s. 91-93) wyznaczamy prawdopodobieństwo tego, że płace nie będą większe o jedno odchylenie standardowe od płacy wyznaczonej na poziomie średniej arytmetycznej. Prawdopodobieństwo to wynosi:

P_g (a < y < a+σ_y) = P_g (z_0 <z < z_2) = P_g(0,2<z<1,07) = -0,0793 + 0,3577= 0,2784

Oznacza to, że prawdopodobieństwo tego, że płace nie będą większe o 812,27 zł (odchylenie standardowe) od 1950 zł (średnia arytmetyczna) jest równe 0,2784.

Obecnie precyzując interpretację odchylenia standardowego z Ad 1) powiemy, że płace przeciętnie odchylają się od płacy średniej wynoszącej 1950 zł o około 812,7 zł, z prawdopodobieństwem P_d+P_g = 0,7324.

Ad 4) Przechodząc do wyznaczenia prawdopodobieństwa tego, że płace będą mniejsze od średniej arytmetycznej wynoszącej a=1950 zł zauważmy, że:

P(0<y<a) = P_1(0<y<g) + P_2(g<y<a) = 0,5 + P_2(g<y<a).

Obecnie logarytmując i standaryzując zmienne występujące w nawiasie przy P_2 i postępując zgodnie z procedurą wyznaczania prawdopodobieństwa dla standaryzowanych zmiennych losowych o rozkładzie normalnym otrzymujemy:

P_2(lng<lny<lna) = P_2[(lng-lng)/σ_{lny})<(lny-lng)/σ_{lny})<(lna-lng)/σ_{lny})] = P_2(0<z<z_0) = 0 +0,0793 = 0,0793

Ostatecznie stwierdzamy, że:

P(0<y<a) = P_1 + P_2 = 0,5 + 0,0793 = 0,5793

Powiemy więc, że w 58 przypadkach na 100 płace pracowników analizowanego sektora będą niższe od ich średniej arytmetycznej, tym samym w w 42 przypadkach na 100 płace pracowników analizowanego sektora będą wyższe od ich średniej arytmetycznej

Aby wyznaczyć i poprawnie zinterpretować względną miarę rozproszenia zmiennej losowej y w relacji do średniej arytmetycznej wykorzystajmy nierówność (38). Na jej podstawie powiemy, że:

$$P(a-\sigma_y < y < a+\sigma_y) = P_d(a-\sigma_y < y < a) + P_g(a < y < a+\sigma_y)\,. \quad (45)$$

Po podzieleniu stronami elementów zawartych w nawiasach przez wielkość średniej arytmetycznej (a) otrzymujemy:

$$P(1-\frac{\sigma_y}{a} < \frac{y}{a} < 1+\frac{\sigma_y}{a}) = P_d(1-\frac{\sigma_y}{a} < \frac{y}{a} < 1) + P_g(1 < \frac{y}{a} < 1+\frac{\sigma_y}{a})\,, \quad (46)$$

co równoważnie możemy zapisać następująco:

$$P(-\frac{\sigma_y}{a} < \frac{y}{a} - 1 < \frac{\sigma_y}{a}) = P_d(-\frac{\sigma_y}{a} < \frac{y}{a} - 1 < 0) + P_g(0 < \frac{y}{a} - 1 < \frac{\sigma_y}{a}). \quad (47)$$

gdzie E(y/a) =1, tym samym E[(y-a)/a] = 0.

Zauważmy, że iloraz odchylenia standardowego i średniej arytmetycznej wyrażony w postaci ułamkowej lub w procentach nazywany jest współczynnikiem zmienności losowej. Jak piszą Kendall i Buckland (1975, s. 270) został on po raz pierwszy zaproponowany przez K. Pearsona w 1895 roku w celu porównania dyspersji rozkładów częstości. Oznaczając współczynnik zmienności losowej dużą literą V mamy:

$$V = \frac{\sigma_y}{a}. \quad (48)$$

Wykorzystując powyższą definicję wyrażenie (47) zapiszemy następująco:

$$P(-V < \frac{y-a}{a} < V) = P_d(-V < \frac{y-1}{a} < 0) + P_g(0 < \frac{y-a}{a} < V) \quad (49)$$

gdzie, zgodnie z (38) $P_d > P_g$.

Na podstawie powyższego stwierdzamy, że w przypadku rozkładu logarytmiczno-normalnego zmiennej losowej y:

- przeciętnie zmienna losowa y odchyla się od średniej arytmetycznej o (V·100)%,
- prawdopodobieństwo tego, iż zmienna losowa przeciętnie odchyli się od średniej arytmetycznej o -(V·100)% jest większe od prawdopodobieństwo tego, że zmienna ta przeciętnie odchyli się od średniej arytmetycznej o (V·100)%, co wynika z asymetrii rozkładu zmiennej losowej y.

Przykład 5

Na podstawie danych z przykładu 1 i 2 oraz informacji z przykładu 4 obliczyć i zinterpretować współczynnik zmienności losowej płac względem ich średniej arytmetycznej.

Z uwagi na fakt, że:

a = 1950,00 zł,

σ_y= 812,27 zł,

otrzymujemy:

V=812,27/1950 =0,4165.

Powiemy więc, ze przeciętnie płace pracowników analizowanego sektora odchylają się od ich średniej arytmetycznej wynoszącej 1950 zł o około 41,65%.

Celem doprecyzowania powyższej interpretacji, wykorzystując zapis (48) oraz informacje z przykładu 4 stwierdzamy, że:

P_d\{-0,4165<[(y-1950)/1950]<0\} = 0,454,

P_g\{0<[(y-1950)/1950]<0,4165\} = 0,2785,

P\{-0,4165<[(y-1950)/1950]<0,4165\} = 0,7324.

Oznacza to, że w 73,24 przypadkach na 100 płace nie będą niższe lub wyższe o więcej niż 41,65% od ich średniej arytmetycznej. Uściślając powiemy, że w 45,4 przypadkach na 100 płace nie będą niższe od ich średniej arytmetycznej o więcej niż 41,65% i jednocześnie o tę wielkość nie będą wyższe od ich średniej arytmetycznej w 27,85 przypadkach na 100.

Podsumowując tę część rozważań powiemy, że jednakowemu rozproszeniu absolutnemu i względnemu zmiennej losowej w relacji do średniej arytmetycznej odpowiada większe prawdopodobieństwo odchyleń ujemnych oraz mniejsze prawdopodobieństwo odchyleń dodatnich.

4. Dwa równoważne twierdzenia dotyczące rozkładu logarytmiczno-normalnego

W dotychczas prowadzonych rozważaniach wartość oczekiwaną (6) i wariancję zmiennej y (9) wyrażaliśmy w kategoriach wartości oczekiwanej i wariancji logarytmu zmiennej y. W sensie poznawczym za pożyteczne należy uznać odwrócenie sytuacji poprzez wyrażenie wartości oczekiwanej oraz wariancji logarytmu zmiennej y w kategoriach wartości oczekiwanej zmiennej y. W tym celu, zgodnie z propozycją Teekens'a i Koerts'a (1972), wyraźmy (9) w następującej postaci:

$$\sigma_y^2 = e^{(\mu_{\ln y} + \sigma_{\ln y}^2)2}(e^{\sigma_{\ln y}^2} - 1) \tag{50}$$

Wykorzystując (6) powyższe wyrażenie zapiszemy następująco:

$$\sigma_y^2 = \mu_y^2(e^{\sigma_{\ln y}^2} - 1) = \mu_y^2 e^{\sigma_{\ln y}^2} - \mu_y^2 \tag{51}$$

Przekształcając (51) otrzymujemy ostatecznie:

$$\sigma_{\ln y}^2 = E(\ln y - E\ln y)^2 = \ln\left(\frac{\sigma_y^2}{\mu_y^2} + 1\right), \tag{52}$$

co należało wykazać.

Z drugiej strony na podstawie (6) mamy:

$$\mu_y = e^{\mu_{\ln y}} e^{\frac{1}{2}\sigma_{\ln y}^2}. \tag{53}$$

Wprowadzając w powyższym wyrażeniu w miejsce wariancji logarytmu zmiennej y wariancję zdefiniowaną w (52) w wyniku prostych przekształceń dochodzimy do następującej postaci:

$$\mu_{\ln y} = \ln\mu_y - \tfrac{1}{2}\ln\left(\frac{\sigma_y^2}{\mu_y^2} + 1\right). \tag{54}$$

Obecnie sformułować możemy dwa równoważne względem siebie twierdzenia dotyczące rozkładu logarytmiczno-normalnego zmiennej losowej y. Przy pierwszym z nich wykorzystamy zdefiniowania ujęte w (2), (3), (6) oraz (9) i powiemy:

TWIERDZENIE 1. Jeżeli logarytm zmiennej losowej y ma rozkład normalny $N(\mu_{\ln y}, \sigma^2_{\ln y})$ to zmienna y ma rozkład logarytmiczno-normalny $\Lambda[e^{\mu_{\ln y} + \frac{1}{2}\sigma^2_{\ln y}}, \quad e^{2\mu_{\ln y} + \sigma^2_{\ln y}}(e^{\sigma^2_{\ln y}} - 1)]$

Z kolei wykorzystując (52) i (54) oraz ogólne zdefiniowanie wartości oczekiwanej i wariancji zmiennej losowej y powiemy:

TWIERDZENIE 2. Jeżeli zmienna losowa y ma rozkład logarytmiczno-normalny $\Lambda(\mu_y, \sigma^2_y)$ to logarytm zmiennej losowej y ma rozkład normalny

$$N[\ln \mu_y - \tfrac{1}{2}\ln\left(\frac{\sigma^2_y}{\mu^2_y} + 1\right), \quad \ln\left(\frac{\sigma^2_y}{\mu^2_y} + 1\right)]$$

Alternatywny sposób zdefiniowania rozkładu logarytmiczno-normalnego sugeruje możliwość alternatywnego względem (14) zdefiniowania związku pomiędzy średnią arytmetyczną i geometryczną. W tym celu zdelogarytmujmy stronami wyrażenie (54), w wyniku czego otrzymujemy:

$$e^{\mu_{\ln y}} = e^{\ln \mu_y} e^{-\frac{1}{2}\left(\frac{\sigma^2_y}{\mu^2_y} + 1\right)} \tag{55}$$

Zauważmy, ze wyrażenie z lewej strony równania jest średnia geometryczna zmiennej y, natomiast $\mu_y = \exp(\ln \mu_y)$ jest średnią arytmetyczną (a). W konsekwencji tego otrzymujemy:

$$g = a / \sqrt{(\sigma^2_y / a^2) - 1} \tag{56}$$

Po prostym przekształceniu (56), alternatywny względem (14) związek funkcyjny pomiędzy średnią geometryczną i arytmetyczną, przedstawia się następująco:

$$g = a^2 / \sqrt{\sigma^2_y + a^2} \tag{57}$$

Przykład 6

Niech płace miesięczne (y) pracowników pewnego sektora gospodarczego charakteryzują się rozkładem logarytmiczno-normalnym. Płace wyrażone są w złotych. Załóżmy, że wartość oczekiwana (średnia arytmetyczna) i odchylenie standardowe płac równają się odpowiednio:

$a = \mu_y = 1950$ zł,

$\sigma_y = 812{,}27$ zł.

Określić: **1)** medianę (średnią geometryczną) płac **2)** względne rozproszenie płac wokół średniej geometrycznej.

Na wstępie zauważmy, że w omawianym przypadku mamy:

$a^2 = 1950^2 = 3\,802\,500$

$\sigma^2_y = 812{,}27^2 = 659\,782{,}55$

Ad 1) Na podstawie (57) otrzymujemy:

$$g = 3\,802\,500 / \sqrt{659\,722{,}52 + 3802\,500} = 1800 \text{ zł},$$

co jest zgodne z założeniami do przykładu 1.
Ad 2) Wariancję logarytmu zmiennej y względem logarytmu zmiennej geometrycznej obliczymy w warunkach alternatywnych na podstawie (52). W rezultacie otrzymujemy:

$$\sigma^2_{\ln y} = \ln[(659782{,}55/3802500)+1] = 0{,}16\,,$$

co znajduje potwierdzenie w danych do przykładu 1.
Oczywiście odchylenie standardowej logarytmu zmiennej y jest równe: $\sigma_{\ln y}$=0,4. Na jego podstawie obliczymy przeciętne w sensie standardowym, względne rozproszenie zmiennej y w stosunku do średniej geometrycznej (mediany) płac. Zgodnie z (19) i (20) mamy:

v_d = exp(-0,4) = 0,67
v_d = exp(0,4) = 1,4918

Obecnie tak jak w przykładzie 2 powiemy, że w sensie standardowym przeciętny udział płac w ich średniej geometrycznej (medianie) waha się w granicach od 0,67 do 1,49, czyli płace przeciętnie odchylają się od ich średniej geometryczne (mediany) w przedziale od –33% do 49%. Rozpatrywany przykład unaocznia równoważność Twierdzenia 1 i 2.

Kończąc tę część rozważań zauważmy, że mimo równoważności Twierdzenia 1 i 2 uznać należy Twierdzenie 2 za pochodne względem Twierdzenia 1. Świadczy o tym sposób dochodzenia do Twierdzenia 2 wynikający z przekształceń parametrów rozkładu normalnego logarytmu zmiennej losowej y.

5. Krótkie wprowadzenie do modeli multiplikatywnych

Rozpatrywany rozkład logarytmiczno-normalny w naturalny sposób związany jest z modelami multiplikatywnymi. Aby to wyjaśnić załóżmy, że zmienna v wyraża stosunek pewnej, przyjmującej jedynie wartości dodatnie, zmiennej losowej y do pewnego dodatniego, nielosowego parametru μ:

$$v = (y/\mu) > 0\,. \tag{58}$$

Oznacza to, że

$$y = \mu \cdot v\,. \tag{59}$$

Jeśli obecnie założymy, iż wartość oczekiwana zmiennej losowej v jest równa jeden, tzn.:

$$\mu_v = Ev = 1, \tag{60}$$

wówczas stwierdzamy, że

$$Ey = \mu \cdot Ev = \mu = a\,. \tag{61}$$

Oznacza to, że wtedy gdy spełniony jest warunek w myśl którego Ev=1, to parametr μ w równaniu (59) jest średnią arytmetyczną zmiennej losowej y, co zasygnalizowano za pomocą symbolu a, zgodnie z wcześniej przyjętą umową.

Jednocześnie zmienna losowa v wyraża sobą stosunek zmiennej losowej y do jej średniej arytmetycznej zgodnie z (58).

Po zlogarytmowaniu stronami równania (59) otrzymujemy:

$$\ln y = \ln \mu + u \tag{62}$$

gdzie:

$$u = \ln v \,. \tag{63}$$

Zauważmy, że jeśli zmienna u ma rozkład normalny to zmienna v ma rozkład logarytmiczno-normalny. Jeśli obecnie założymy, że wartość oczekiwana u jest równa zero, tzn.:

$$\mu_u = Eu = 0 \,, \tag{64}$$

wówczas stwierdzamy, że

$$E \ln y = E \ln \mu + Eu = \ln \mu \,. \tag{65}$$

Ponieważ wartość oczekiwana logarytmu zmiennej y jest równa logarytmowi parametru μ, więc parametr μ jest średnią geometryczną zmiennej y, jako że

$$e^{\ln y} = e^{\ln \mu} = \mu = g \tag{66}$$

Oznacza to, że wtedy gdy spełniony jest warunek w myśl którego Eu=lnv=0, to parametr μ w równaniu (59) jest średnią geometryczną zmiennej losowej y, co zasygnalizowano za pomocą symbolu g, zgodnie z wcześniej przyjętą umową. Jednocześnie zmienna losowa v wyraża sobą stosunek zmiennej losowej y do jej średniej geometrycznej zgodnie z (58)[1].

Obecnie na bazie dwu wcześniej sformułowanych alternatywnych twierdzeń, sformułować możemy dwa następne dotyczące związków pomiędzy parametrami rozkładu zmiennych losowych zdefiniowanych jako u i v.

TWIERDZENIE 3. Jeżeli w warunkach (63) zmienna losowa u ma rozkład normalny $\mathrm{N}(0, \sigma_u^2)$, to v jest zmienną losową o rozkładzie logarytmiczno-normalnym $\Lambda[e^{\frac{1}{2}\sigma_u^2}, \quad e^{\sigma_u^2}(e^{\sigma_u^2} - 1)]$

TWIERDZENIE 4[2]. Jeżeli w warunkach (63) zmienna v jest zmienna losową o rozkładzie logarytmiczno-normalnym $\Lambda(1, \sigma_v^2)$, to u jest zmienną losową o rozkładzie normalnym $\mathrm{N}[-\frac{1}{2}\ln(\sigma_v^2 + 1), \quad \ln(\sigma_v^2 + 1)]$

Podsumowując zauważmy, że parametr μ może być uznany za warunkową średnią arytmetyczną lub geometryczną zmiennej losowej y w sytuacji, gdy średnia arytmetyczna a tym samym średnia geometryczna tej zmiennej ulegać będzie zmianie pod wpływem czynników nielosowych. Zmiana bowiem

[1] Przedstawione tutaj założenia i wypływające z nich wnioski stanowiły podstawę rozważań dotyczących modeli multiplikatywnych w artykułach Bołt, Ossowski (1992), Ossowski (1988) i monografii Ossowski (1989).

[2] Twierdzenie to stanowiło podstawę do poszukiwań nieobciążonego estymatora dla składnika systematycznego w modelu multiplikatywnym w przypadku gdy składnik ten wyznacza warunkowe średnie arytmetyczne zmiennej objaśnianej (por. Bradu, Mundlak (1970), Golberger (1968), Heien (1968), i Teekens, Koerts (1972).

warunkowej średniej geometrycznej zmiennej y, charakteryzującej się rozkładem logarytmiczno-normalnym, gwarantuje ściśle określone zmiany warunkowej średniej arytmetycznej i odwrotnie, zmiana warunkowej średniej arytmetycznej zmiennej y gwarantuje ściśle określone zmiany warunkowej średniej geometrycznej. O tym, czy parametr μ ma być warunkową średnią arytmetyczną lub geometryczną decydujemy przyjmując określone założenia dotyczące parametrów rozkładu zmiennej losowej v a tym samym zmiennej u. Należy podkreślić z całą mocą, że zmienna losowa y, jeśli charakteryzuje się rozkładem logarytmiczno-normalnym, ma ściśle określone parametry rozkładu, takie jak średnia arytmetyczna (wartość oczekiwana), średnia geometryczna (mediana), dominanta, wariancja zmiennej y lub wariancja logarytmu zmiennej y. Tym samym korzystanie z twierdzenia 3 lub 4 w niczym nie narusza charakterystyki rozkładu zmiennej losowej y. Decyduje jedynie o wyborze parametru odniesienia w stosunku do innych parametrów rozkładu zmiennej y, a tym samym o sposobie zapisu wszystkich pozostałych parametrów rozkładu zmiennej y. Parametry te bowiem możemy zapisywać w kategoriach zmiennej y lub jej logarytmu a tym samym w kategoriach zmiennej u lub v.

6. Wnioski końcowe

W artykule wykazano, że w przypadku, gdy zmienna losowa y ma rozkład logarytmiczno-normalny $N(\mu_{\ln y}, \sigma^2_{\ln y})$ to:

- parametry $v_d=\exp(-\sigma_{\ln y})$ i $v_g=\exp(\sigma_{\ln y})$ wyznaczają przedział dla przeciętnego w sensie standardowym, względnego rozproszenia zmiennej y w stosunku do jej średniej geometrycznej $g=\exp(\mu_{\ln y})$,
- parametry $g\cdot v_d$ do $g\cdot v_g$ wyznaczają przedział dla przeciętnego w sensie standardowym, absolutnego rozproszenia zmiennej y względem jej średniej geometrycznej.

Stwierdzono:

- dla przypadku średniej geometrycznej zmiennej losowej y, iż jednakowemu prawdopodobieństwu realizacji zdarzeń odpowiada, co do wartości bezwzględnej, mniejszy przedział dolny i większy przedział górny odchyleń zmiennej y od jej średniej geometrycznej (mediany).
- dla przypadku średniej arytmetycznej zmiennej losowej y, iż jednakowemu rozproszeniu absolutnemu i względnemu zmiennej losowej y w relacji do średniej arytmetycznej odpowiada większe prawdopodobieństwo odchyleń ujemnych oraz mniejsze prawdopodobieństwo odchyleń dodatnich.

Na podstawie przeprowadzonej analizy dotyczącej rodzajów miar dyspersji wnioskujemy, że:

- miarę względnego rozproszenia zmiennej losowej uznać można za naturalną miarę rozproszenia zmiennej losowej w relacji do średniej geometrycznej; natomiast miarę absolutnego rozproszenia zmiennej losowej względem średniej geometrycznej uznać należy za miarę pochodną względem miary rozproszenia względnego.

- miarę absolutnego rozproszenia zmiennej losowej uznać można za naturalną miarę rozproszenia zmiennej losowej w relacji do średniej arytmetycznej; natomiast miarę względnego rozproszenia zmiennej losowej względem średniej arytmetycznej uznać należy za miarę pochodną względem miary rozproszenia absolutnego.

Wykazano ponadto, że w zakresie związków pomiędzy średnią arytmetyczną (a) i geometryczną (g) zmiennej losowej y o rozkładzie logarytmiczno-normalnym, funkcja: $g=a^2(\sigma^2_y+a^2)^{-(1/2)}$ jest równoważna funkcji: $g=a\cdot\exp[(1/2)\ \sigma^2_{lny}]$, gdzie parametr σ^2_y jest wariancją zmiennej losowej y, natomiast parametr σ^2_{lny} jest wariancją logarytmu zmiennej losowej y.

Literatura

Aitchison, J., Brown, A. (1957), *The Lognormal Distribution*, Cambridge University Press, Cambidge.

Bołt, T.W., Ossowski, J. (1992), Prognozowanie na podstawie modeli logarytmiczno-liniowych, *Przegląd Statystyczny*, z. 3-4 s.327-340.

Bradu, D., Mundlak, Y. (1970), Estimation in Lognormal Linear Models, *Journal of the American Staistical Association*, nr 65, s.198-211.

Bronsztejn, J.N., Siemiendiajew, K.A. (1976), *Matematyka. Poradnik encyklopedyczny*, PWN, Warszawa

Goldberger, A.S. (1972), *Teoria ekonometrii*, PWE, Warszawa.

Golberger, A.S. (1968), The Interpretation and Estimation of Cobb-Douglas Functions, *Econometrica*, nr 35, s. 464-472.

Heien, D.M. (1968), Not on Log-linear Regression, *Journal on the American Statistical Associacion*, nr 63, s.1034-1038

Kendall, M. Bucland, W.R. (1975), *Słownik terminów statystycznych,* PWE, Warszawa.

Klein, L.R. (1965), *Wstęp do ekonometrii,* PWE, Warszawa.

Kmenta, J.(1990), *Elements of Econometrics, Second Edition*, Macmillan Publishing Company, New York.

Murti, V.N., Sastry, V.K. (1957), Production Functions for Indian Industry, *Econometrica*, nr 25, s. 205-221.

Ossowski, J. (1988), Własności interpretacyjne składnika losowego w modelu multiplikatywnym, *Przegląd Statystyczny*, z.2, s.131-142.

Ossowski, J. (1989), *Modele klasy logarytmiczno-liniowej w analizie efektywności procesu produkcji,* Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk, Zeszyty Naukowe, Rozprawy i Monografie 130.

Pawłowski, Z. (1969), *Wstęp do statystyki matematycznej,* PWN, Warszawa.

Teekens, R., Koerts, J. (1972), Some Statistical Implications of the Log Transformations of Multiplicative Models, *Econometrica*, nr 5 , s. 793-819.

Theil, H. (1979), *Zasady ekonometrii*, PWN, Warszawa.