B A D A N I A   O P E R A C Y J N E   I   D E C Y Z J E

Nr 3–4 2006

Bogusław GUZIK*

ŚREDNI BŁĄD PROGNOZOWANIA DLA METODY EKSTRAPOLACJI PRZYROSTU EMPIRYCZNEGO

W artykule sformułowano standardowy układ założeń stochastycznych dla prostej metody prognozowania, polegającej na liniowej ekstrapolacji średniego przyrostu empirycznego oraz wyprowadzono wzory, określające średni błąd prognozowania (*ex ante* błąd predykcji) dla tej sytuacji. Wzory te porównano z szacunkowymi błędami średnimi dla innych metod prognozowania – metody status quo, ekstrapolacji średniej i ekstrapolacji trendu liniowego, oszacowanego klasyczną metodą najmniejszych kwadratów (mnk).

Słowa kluczowe: *średni błąd predykcji, ekstrapolacja średniego przyrostu*

1. Wstęp

Ekstrapolacja średniego przyrostu empirycznego jest jedną z prostszych metod prognozowania statystycznego na podstawie jednowymiarowego szeregu czasowego. Metoda ta, choć chętnie stosowana (np. do ustalania wstępnych wariantów prognozy), w podstawowej literaturze przedmiotu nie jest rozpatrywana z punktu widzenia szacowania *ex ante* błędu prognozy. Przypuszczalnie kryje się za tym spotykany w literaturze pogląd, iż dla tzw. „prostych” metod prognozowania (do których omawiana metoda należy) brak jest możliwości oszacowania błędu prognoz *ex ante*.

W artykule pokazano, iż taka możliwość jednak istnieje. Sformułowano standardowy układ założeń stochastycznych dla metody ekstrapolacji przyrostu empirycznego i dla (liniowej) ekstrapolacji średniego przyrostu empirycznego wyprowadzono wzory, dotyczące średniego błędu prognozowania w warunkach wyżej wspomnianych standardowych założeń stochastycznych.

* Katedra Ekonometrii, Akademia Ekonomiczna, al. Niepodległości 10, 60-967 Poznań, e-mail: b.guzik@ae.poznan.pl

Wzory te porównano ze średnim błędem prognozowania dla metody status quo, metody ekstrapolacji średniej oraz ekstrapolacji trendu szacowanego klasyczną metodą najmniejszych kwadratów.

2. Średni błąd prognozowania

Przez y_t (t = 1, ..., T) oznaczamy zaobserwowane wartości zmiennej prognozowanej Y. Prognozę na czas $\tau > T$ oznaczmy przez y_τ^*, a przez y_τ nieznaną, „prawdziwą" wartość zmiennej Y w czasie τ.

Średni błąd prognozowania[1] lub średni błąd predykcji[2] (zwany też średnim błędem prognozy[3]) wyprowadzany jest – jak wiadomo – następująco:

I. Przyjmuje się, że wartość zmiennej prognozowanej w czasie t = 1, 2, ..., T, ..., τ jest realizacją zmiennej losowej $\mathcal{Y}_t$, (którą można nazwać *możliwe wartości zmiennej prognozowanej w czasie t*), prognoza y_τ^* natomiast jest realizacją zmiennej losowej $\mathcal{Y}_\tau^*$ (predyktora losowego, który można nazwać *możliwe prognozy zmiennej Y na czas* τ).

II. W rezultacie błąd prognozy (liczba) jest realizacją zmiennej losowej:

$$\mathcal{B}_\tau = \mathcal{Y}_\tau^* - \mathcal{Y}_\tau. \tag{1}$$

(którą można nazwać *możliwe błędy prognoz na czas* τ).

III. Średnim błędem prognozowania jest

$$\mu_\tau = \sqrt{\mathrm{E}\mathcal{B}_\tau^2}\ . \tag{2}$$

3. Ekstrapolacja średniego przyrostu empirycznego

Reguła prognozowania według ekstrapolacji średniego przyrostu empirycznego ma postać

[1] Termin używany, np. w pracy B. Guzik, W. Jurek, *Ekonometria z zadaniami*, Wyd. AE Poznań, 1993, paragraf 1.11, B. Guzik, D. Appenzeller, W. Jurek, *Prognozowanie i symulacje. Wybrane zagadnienia*, Wyd. AE Poznań, 2005, paragraf 4.3..

[2] Np. Z. Pawłowski, *Teoria prognozy ekonometrycznej w gospodarce socjalistycznej*, PWN, wyd. II, 1974, paragraf 3.1; A. Zeliaś, *Teoria prognozy*, PWE, 1997, paragraf 2.5.

[3] Np. A.S. Goldberger, *Teoria ekonometrii*, PWE, 1972, paragraf 4.5; P. Dittmann, *Prognozowanie w przedsiębiorstwie*, Oficyna Ekonomiczna, Kraków, 2003, s.78.

$$y_\tau^* = y_T + a(\tau - T)\,, \tag{3}$$

gdzie a jest średnim przyrostem empirycznym z obserwacji na obserwację.

Przyjmiemy, że przyrost ten obliczany jest statystycznie[4] jako średnia z przyrostów zmiennej prognozowanej z okresu na okres w odcinku czasu $[p, k\,]$. Wówczas

$$a = \frac{y_k - y_p}{p - k} \quad (p \geq 1, \quad k \leq T). \tag{4}$$

W szczególności średni przyrost empiryczny może być obliczany na podstawie całego przedziału danych empirycznych, wtedy $p = 1$, $k = T$.

Symbolem Z będziemy oznaczać zbiór obserwacji, na podstawie których dokonywane są oszacowania. Zbiór ten zawiera kolejne obserwacje, przynajmniej $t = p$, $p + 1, ..., k$, z tym że nie zawiera on obserwacji $t = T$. Tak więc $Z \subset \{1, 2, ..., T - 1\}$[5].

Założenia. Proponujemy następujący *standardowy* układ założeń dla prognozowania przez ekstrapolację średniego przyrostu empirycznego:

I. **Predyktor** ma postać

$$\mathcal{Y}_\tau^* = \mathcal{Y}_T + \mathcal{A}\,(\tau - T), \tag{5}$$

gdzie:

$\mathcal{Y}_T$ – zmienna losowa generująca wartości prognozowanej zmiennej w ostatniej obserwacji empirycznej,

$\mathcal{A}$ – zmienna losowa (estymator), której realizacją jest średni przyrost empiryczny a.

Jest to predyktor liniowy o (losowym) współczynniku kierunkowym $\mathcal{A}$ oraz losowym wyrazie wolnym równym $\mathcal{Y}_T$[6].

II. Zmienna losowa $\mathcal{Y}_\tau$ generująca wartość zmiennej prognozowanej w czasie τ jest określona jako

$$\mathcal{Y}_\tau = \mathcal{Y}_T + \alpha(\tau - T) + \mathcal{E}_\tau\,, \tag{6}$$

gdzie α – nielosowy, „prawdziwy” współczynnik kierunkowy[7].

---

[4] Obliczenia statystyczne to jedna z możliwych dróg. Przyrost ten może być liczony inaczej, a nawet może być ustalony *a priori*.

[5] Dlaczego do zbioru Z nie zaliczmy $t = T$, wyjaśni się za chwilę.

[6] Predyktor ma taką właśnie postać w ślad za regułą prognozy (3), a zwłaszcza w ślad za tym, że y_T oraz średni przyrost empiryczny a to wielkości statystyczne, które należy traktować jako realizacje odpowiednich zmiennych losowych.

[7] Którego oceną jest a, zaś estymatorem jest $\mathcal{A}$. Równanie (6) to funkcja liniowa z losowym wyrazem wolnym $\mathcal{Y}_T$ oraz nielosowym współczynnikiem kierunkowym α, na które addytywnie nałożono składnik losowy.

III. Wartość zmiennej prognozowanej w obserwacji T jest realizacją zmiennej losowej $\mathcal{Y}_T$.

IV. Wartość zmiennej prognozowanej w obserwacji $t \in Z$ jest realizacją zmiennej losowej

$$\mathcal{Y}_t = \mathcal{Y}_T + \alpha(t - T) + \mathcal{E}_t \quad (t \in Z). \tag{7}$$

V. O składnikach losowych na zbiorze Z przyjmujemy założenia klasyczne; mianowicie, że mają one zerowe wartości oczekiwane, że ich wariancje są identyczne i równe σ^2 oraz że są one wzajemnie nieskorelowane:

$$E\mathcal{E}_\tau,\ \mathrm{E}\mathcal{E}_t = 0; \tag{8}$$

$$\mathrm{Var}\mathcal{E}_\tau,\ \mathrm{Var}\mathcal{E}_t = \sigma^2 \text{ (z uwagi na (8) jest } E(\mathcal{E}_\tau)^2,\ E(\mathcal{E}_t)^2 = \sigma^2); \tag{9}$$

$$\mathrm{Cov}\mathcal{E}_\tau\,\mathcal{E}_t,\ \mathrm{Var}\mathcal{E}_t\,\mathcal{E}_l = 0 \text{ (z uwagi na (8) jest } E(\mathcal{E}_\tau\,\mathcal{E}_t),\ E(\mathcal{E}_t\,\mathcal{E}_l)^2 = 0); \tag{10}$$

$$\text{dla } t, l \in Z;\ t \neq l.$$

Wartości oczekiwane oraz wariancje wszystkich zmiennych losowych są skończone.

Komentarze do równań określających zmienne losowe $\mathcal{Y}_t$ oraz $\mathcal{Y}_\tau$

1. W ślad za szczególną postacią predyktora, w którym wyrazem wolnym jest zmienna losowa $\mathcal{Y}_T$ generująca wartości zmiennej prognozowanej w ostatniej obserwacji empirycznej, zmienna ta występuje też jako wyraz wolny równań, określających zmienne losowe $\mathcal{Y}_t$ $(t \in Z)$ oraz zmienną losową $\mathcal{Y}_\tau$.

2. Zmienne te zależą od *dwóch*_zmiennych losowych: $\mathcal{Y}_T$ oraz – „standardowo" – od składnika losowego. Zależność od składnika losowego jest niezbędna, gdyż jego pominięcie prowadzi do niedorzeczności, np. różnica $\mathcal{Y}_k - \mathcal{Y}_p$, o której powiemy za chwilę, stawałaby się jakoby nielosowa.

3. Zmienna losowa $\mathcal{Y}_T$ ma szczególny charakter: generuje inne zmienne losowe, a ponadto jest niejako „egzogeniczna" – nie zależy od innych, ani od składnika losowego itp. Gdybyśmy napisali dla tej zmiennej równanie analogiczne do (7), tj.

$$\mathcal{Y}_t = \mathcal{Y}_T + \alpha(t - T) + \mathcal{E}_t \quad \text{dla} \quad t = T,$$

wówczas otrzymalibyśmy

$$\mathcal{Y}_T = \mathcal{Y}_T + \mathcal{E}_T,$$

co jest niedorzeczne.

Wyprowadzenie wzorów

Estymator współczynnika α

Na podstawie definicji (4) estymatorem współczynnika α jest

$$A = \frac{Y_k - Y_p}{k - p}. \tag{11}$$

Przypadek 1. Wyznaczając współczynnik a, nie wykorzystano informacji z ostatniej obserwacji, czyli jest $k < T$.

Wtedy estymator parametru α wyraża się jako

$$A = \frac{Y_k - Y_p}{k - p} = \frac{[\alpha(k-T) + Y_T + E_k] - [\alpha(p-T) + Y_T + E_p]}{k - p},$$

czyli

$$A = \alpha - \frac{E_k - E_p}{k - p}. \tag{12}$$

Przypadek 2. Wyznaczając średni przyrost a, uwzględniono ostatnią obserwację empiryczną, czyli $k = T$. Wówczas

$$A = \frac{Y_T - Y_p}{T - p} = \frac{Y_T - [\alpha(p-T) + Y_T + E_p]}{T - p},$$

tak więc

$$A = \alpha - \frac{E_p}{T - p}. \tag{13}$$

Dalej, o ile nie powiemy inaczej, zakładamy, że $k < T$ (przypadek 1). Do przypadku drugiego wrócimy pod koniec tej części artykułu.

Zmienna losowa generująca błędy prognoz określona jest wzorem:

$$B_\tau = Y_\tau^* - Y_\tau = [Y_T + A(\tau - T)] - [Y_T + \alpha(\tau - T) + E_\tau] = (A - \alpha)(\tau - T) - E_\tau,$$

$$B_\tau = w(E_k - E_p) - E_\tau,$$

gdzie

$$w = \frac{\tau - T}{k - p}.$$

Kwadrat średniego błędu prognozowania:

$$E(B_\tau)^2 = E[w(E_k - E_p)]^2 - 2w\, E(E_k - E_p)E_\tau + E(E_\tau)^2.$$

Środkowy składnik jest zerowy, bowiem składniki losowe są wzajemnie nieskorelowane. Ostatecznie otrzymujemy

$$E(\mathcal{B}_\tau)^2 = \sigma^2\,(2w^2+1) \quad (\text{gdy } k < T)\,. \tag{14}$$

Można sprawdzić, że gdy $k = T$, kwadrat średniego błędu prognozowania jest mniejszy i wynosi:

$$E(\mathcal{B}_\tau)^2 = w^2\,E(\mathcal{E}_p)^2 - 2w\,E\mathcal{E}_p\,\mathcal{E}_\tau + E(\mathcal{E}_\tau)^2,$$

$$E(\mathcal{B}_\tau)^2 = \sigma^2(w^2+1) \quad (\text{gdy } k = T). \tag{15}$$

Średni błąd prognozowania dla procedury ekstrapolacji średniego przyrostu empirycznego (3), (4) w warunkach założeń (5)–(11) wynosi:

$$\mu_\tau = \begin{cases} \sigma\sqrt{2w^2+1}, & \text{jeżeli} \quad k < T, \\ \sigma\sqrt{w^2+1}, & \text{jeżeli} \quad k = T, \end{cases} \tag{16}$$

gdzie

$$w = \frac{\tau - T}{k - p}\,. \tag{17}$$

Uwagi

1. Średni błąd prognozowania jest większy od σ, gdyż wyrażenie podpierwiastkowe jest większe od 1.

2. Średni błąd prognozowania jest mniejszy w przypadku drugim ($k = T$), gdyż wtedy zmienna losowa generująca błędy prognoz zależy od mniejszej liczby zmiennych losowych (nie zależy od $\mathcal{E}_k$). Ta własność sugeruje, aby licząc średni przyrost empiryczny, uwzględniać ostatnią obserwację empiryczną, czyli brać dane z przedziału czasowego $[p, T]$.

Dalej zakładamy ten przypadek i analizujemy średni błąd prognozowania:

$$\mu_\tau = \sigma\sqrt{1+w^2}\,, \quad \text{gdzie } w = \frac{\tau - T}{k - p}\,. \tag{18}$$

3. Jeśli dokonuje się ekstrapolacji tylko jednego (np. ostatniego) przyrostu, czyli gdy $a = y_k - y_{k-1}$, (np. $k = T$), to średni błąd prognozowania:

$$\mu_\tau = \sigma\sqrt{1+(\tau - T)^2}\,.$$

Oznacza to, że każdy przypadek przyjęcia za współczynnik a tylko przyrostu między *sąsiednimi* obserwacjami ($k - p = 1$) jest najgorszy z punktu widzenia wartości średniego błędu prognozowania (błąd ten jest największy).

4. Najlepiej, gdy różnica $k - p$ jest możliwie największa, a więc gdy $k = T$, $p = 1$, tzn. gdy współczynnik a obliczany jest przez przyrównanie wartości w ostatniej i w pierwszej obserwacji.

5. Średni błąd prognozowania, co oczywiste, rośnie względem odchylenia standardowego składników losowych oraz względem odległości ekstrapolacji $\tau - T$.

6. Oszacowaniem wartości σ^2 może być dotycząca zbioru Z suma kwadratów reszt oszacowanego trendu $\hat{Y} = y_T + (t - T)a$, dzielona przez liczebność zbioru.

4. Metoda ekstrapolacji średniego przyrostu a metoda status quo

A. Założenia porównań z innymi metodami prognozowania

W tym i w dwóch następnych paragrafach porównamy średni błąd prognozowania dla metody ekstrapolacji średniego błędu prognozowania ze średnim błędem prognozowania dla metod: status quo, ekstrapolacji średniej arytmetycznej oraz ekstrapolacji trendu liniowego oszacowanego klasyczną metodą najmniejszych kwadratów.

Spośród dwóch podanych wyżej wzorów, dotyczących prognozowania według średniego przyrostu empirycznego, będziemy rozpatrywać postać bardziej efektywną, którą określa druga część wzoru (16), czyli

$$\mu_\tau^{EP} = \sqrt{\sigma^2 \left[\left(\frac{\tau - T}{T - 1} \right)^2 + 1 \right]}. \tag{19}$$

Jak już powiedziano, metoda ekstrapolacji średniego przyrostu bezwzględnego to metoda prognozowania zjawiska charakteryzującego się liniowym trendem wartości oczekiwanych, por. założenia (6), (7). Trend ten ma postać

$$\Psi = \beta + (t - T)\alpha, \quad t = 1, 2, ..., T, \tau,$$

gdzie β jest wartością oczekiwaną zmiennej losowej Y_T.

To samo, a więc to, iż wartości oczekiwane zmiennych losowych Y_t wykazują trend liniowy, trzeba zakładać w odniesieniu do innych porównywanych metod – w tym wypadku: metody status quo, ekstrapolacji średniej arytmetycznej i ekstrapolacji trendu liniowego oszacowanego według klasycznej mnk, gdyż inaczej porównywano by metody przy innych założeniach modelowych.

Założenie o liniowym trendzie wartości oczekiwanych dla metody ekstrapolacji oszacowanego trendu liniowego jest oczywiste. W dwóch pierwszych – nie, gdyż prognozowanie status quo oraz na poziomie średniej, „modelowo” związane jest z brakiem trendu. Niemniej jednak obliczenie prognozy status quo lub prognozy na poziomie średniej w wypadku trendu jest formalnie możliwe. Oznacza jedynie niezgodność interpretacyjną reguły prognozowania z założeniami modelowymi, a więc oznacza popełnienie błędu w specyfikacji procedury prognozowania – zamiast procedury uwzględniającej trend, stosuje się procedurę pomijającą go[8].

B. Średni błąd prognozowania dla metody status quo

W metodzie status quo:

$$y_\tau^* = y_T . \tag{20}$$

Rozpatrujemy wariant metody opisany następującymi założeniami:

I. Zmienne losowe generujące wartości zmiennej prognozowanej są określone wzorem:

$$Y_T = \alpha_T + E_T \quad \text{oraz} \quad Y_\tau = \alpha_\tau + E_\tau . \tag{21}$$

Wielkości α_T oraz α_τ są nielosowe.

II. Współczynniki α_T oraz α_τ wykazują trend liniowy:

$$\alpha_T = \alpha T + \beta \quad \text{oraz} \quad \alpha_\tau = \alpha\,\tau + \beta . \tag{22}$$

III. Predyktorem jest zmienna losowa:

$$Y_\tau^* = Y_T = \alpha_T + E_T. \tag{23}$$

IV. Składniki losowe spełniają klasyczne założenia:

mają one zerowe wartości oczekiwane,
ich wariancje są jednakowe, a kowariancje wynoszą zero. (24)

W tej sytuacji średnim błędem prognozowania dla metody *status quo* (20), przy założeniach (21)–(24) jest:

$$\mu_\tau^{SQ} = \sqrt{2\sigma^2 + \alpha^2 (T-\tau)^2}\ ^{9}. \tag{25}$$

C. Porównania średniego błędu prognozowania

Porównując wyrażenia podpierwiastkowe wzorów (25) oraz (19) widać, że $\mu_\tau^{EP} < \mu_\tau^{SQ}$, gdy

$$\sigma^2 \left[\left(\frac{\tau - T}{T-1} \right)^2 + 1 \right] < 2\sigma^2 + \alpha^2 (\tau - T)^2 .$$

[8] Jest to zupełnie analogiczne na przykład do ekstrapolacji trendu liniowego w sytuacji, gdy dane i teoria mówią, że przebieg jest wykładniczy.

[9] Dodajmy, że jeśli przyjmiemy założenia metody status quo, tzn: I, III, IV, i jednocześnie przyjmiemy „standardowe" założenie II: współczynniki α_T, α_τ są równe (co oznacza brak trendu), to dla metody status quo przy standardowych założeniach stochastycznych $\mu_\tau = \sigma\sqrt{2}$.

1. Średni błąd prognozowania dla metody ekstrapolacji średniego przyrostu empirycznego będzie mniejszy od średniego błędu prognozowania metody status quo, gdy odległość prognozowania $\tau - T$ nie przekracza liczby obserwacji $T - 1$[10].

Dla „rozsądnej" odległości ekstrapolacji (co najwyżej na $T - 1$ jednostek czasu) metoda średniego przyrostu empirycznego jest więc – w sensie średniego błędu prognozowania – „na pewno" lepsza od metody *status quo*.

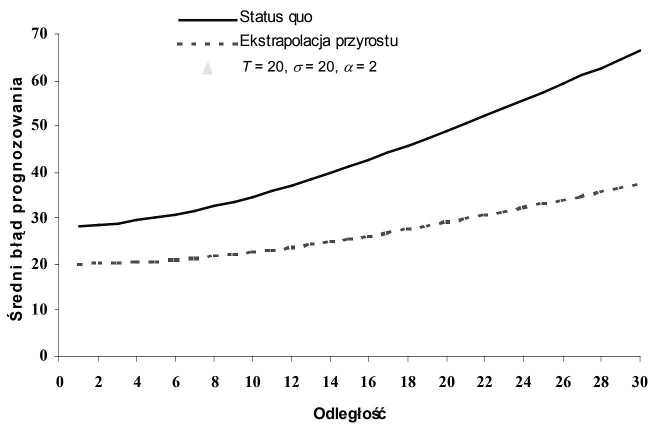
2. Przy większych odległościach prognozowania może być różnie. Przy danej liczbie obserwacji T zależy to od relacji między odchyleniem standardowym składników losowych σ a modułem współczynnika kierunkowego trendu α.

Po podzieleniu obu stron powyższego warunku przez σ^2 i wprowadzeniu oznaczenia $r^2 = \sigma^2/\alpha^2$ otrzymamy następujący warunek na to, by średni błąd prognozowania był mniejszy dla metody ekstrapolacji przyrostu empirycznego niż dla metody status quo:

$$\left(\frac{1}{T-1}\right)^2 - \left(\frac{1}{\Delta}\right)^2 < r^2, \quad \Delta = \tau - T, \tag{26}$$

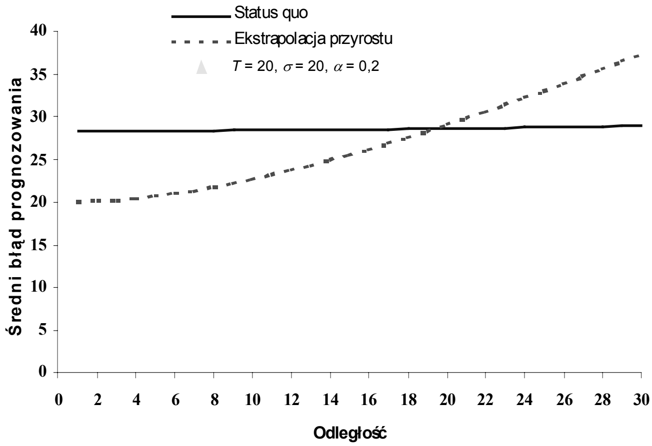
gdzie r^2 określa tę część wariancji σ^2, którą stanowi α^2.

Jeśli trend jest odpowiednio stromy względem dyspersji składników losowych (a więc, gdy r jest odpowiednio duże), metoda przyrostu empirycznego daje mniejszy średni błąd prognozowania (zob. rys. 1, na tym rysunku $r = 0{,}1$). Jeśli trend jest prawie płaski w tym sensie, że współczynnik kierunkowy stanowi niewielką część dyspersji (r jest odpowiednio małe), to mogą się pojawiać sytuacje, gdy metoda przyrostu empirycznego jest gorsza (zob. rysunku 2, na którym $r = 0{,}001$).



Rys. 1. Średni błąd prognozowania a odległość

[10] Wyrażenie po lewej stronie nie przekracza wtedy $2\sigma^2$, a wyrażenie po prawej przekracza.



Rys. 2. Średni błąd prognozowania a odległość $\alpha = 0{,}2$

5. Metoda ekstrapolacji średniego przyrostu a ekstrapolacja średniej arytmetycznej

A. Średni błąd prognozowania dla ekstrapolacji zwykłej średniej arytmetycznej

Prognozę oblicza się jako

$$y^*_\tau = \overline{y}\,, \tag{27}$$

gdzie $\overline{y}$ – średnia arytmetyczna obliczona na podstawie zbioru S danych z przedziału $[1, ..., T\,]$, czyli $S \subset \{1, 2, ..., T\}$. Liczebność zbioru S oznaczymy przez M[11].

Rozpatrujemy następujący wariant założeń stochastycznych:

I. Zmienne losowe generujące wartości zmiennej prognozowanej są określone wzorem:

$$Y_t = \alpha_t + E_t \quad (t \in S) \quad \text{oraz} \quad Y_\tau = \alpha_\tau + E_\tau\,. \tag{28}$$

Wielkości α_T oraz α_τ są nielosowe.

II. Współczynniki α_t oraz α_τ charakteryzują się trendem liniowym:

$$\alpha_t = \alpha\, t + \beta \quad (t \in S) \quad \text{oraz} \quad \alpha_\tau = \alpha\, \tau + \beta\,. \tag{29}$$

[11] Niekoniecznie muszą to być dane z kolejnych momentów czasu.

III. Predyktorem jest zmienna losowa:

$$Y_\tau^* = \bar{Y} = \bar{\alpha} + \bar{E}\ , \tag{30}$$

gdzie:

$$\bar{Y} = \frac{1}{M}\sum_{t\in S} Y_t = \frac{1}{M}\sum_{t\in S}(\alpha_t + E_t)\,, \tag{31}$$

$$\bar{\alpha} = \frac{1}{M}\sum_{t\in S}\alpha_t \tag{32}$$

(średnia z wartości oczekiwanych zmiennych Y_t na zbiorze S);

$$\bar{E} = \frac{1}{M}\sum_{t\in S} E_t \tag{33}$$

(średnia składników losowych E_t na zbiorze S).

IV. Składniki losowe spełniają klasyczne założenia:

ich wartości oczekiwane są zerowe,
wariancje są jednakowe, a kowariancje wynoszą zero. (34)

Średnim błędem prognozowania dla metody ekstrapolacji zwykłej średniej arytmetycznej w warunkach założeń (28)–(34) jest

$$\mu_\tau^{ŚA} = \sqrt{\sigma^2(1+\frac{1}{M}) + \alpha^2(\tau - \bar{t}\ )^2}\ ^{12}. \tag{35}$$

gdzie $\bar{t}$ – średnia wartości t na zbiorze S.

Trudno przesądzić, czy lepiej, gdy liczebność M zbioru, na podstawie którego liczono średnią jest większa, czy też gdy jest mniejsza, bowiem przy wzroście M zmniejsza się wprawdzie pierwszy składnik, ale jednocześnie może się zwiększać drugi składnik, bo wzrastać może średnia $\bar{t}$ [13].

B. Porównanie średniego błędu prognozowania

Dzieląc wyrażenia podpierwiastkowe wzorów (35) oraz (19) przez σ^2 oraz wprowadzając oznaczenie $\alpha^2 = r^2\sigma^2$ (moduł współczynnika kierunkowego trendu stanowi r-tą część odchylenia standardowego składników losowych), otrzymujemy następujące wnioski:

[12] Jeśli przyjąć „standardowe” założenie, że wartości oczekiwane zmiennych Y_t nie odznaczają się trendem, co znaczy, że zamiast (29) przyjmuje się, że α_t, $\alpha_\tau = \beta$ ($t\in Z$), to średni błąd prognozowania na podstawie średniej arytmetycznej M okresowej wynosi $\mu_\tau^{ŚA} = \sigma\sqrt{\left(1+\frac{1}{M}\right)}$.

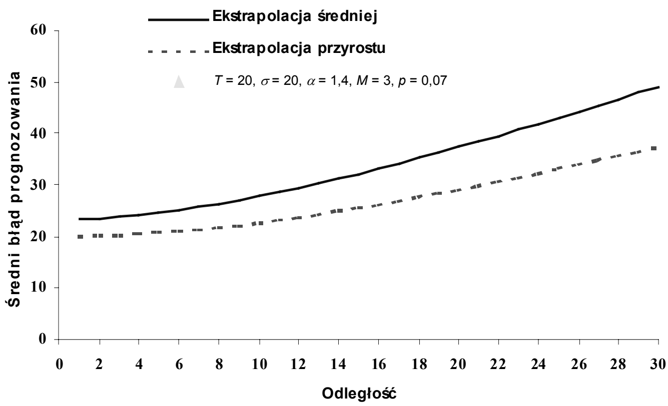
[13] Np. jeśli zbiór S zawiera M ostatnich obserwacji.

1. Średni błąd prognozowania jest mniejszy dla metody ekstrapolacji przyrostu empirycznego niż dla metody ekstrapolacji średniej, gdy

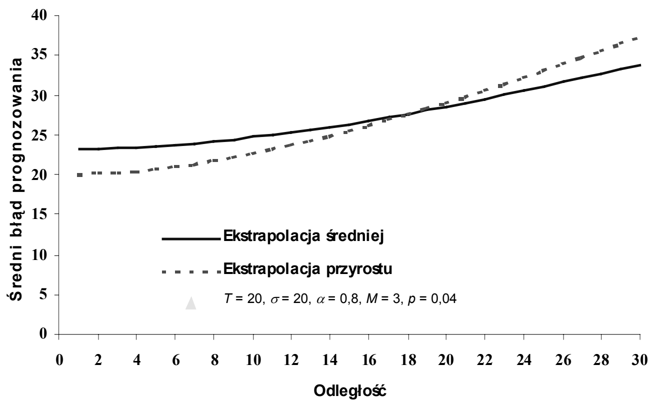
$$\frac{(\tau-T)^2}{(T-1)^2} - r^2(\tau-\bar{t})^2 < \frac{1}{M}. \tag{36}$$

2. Nierówność (36) na pewno jest spełniona, gdy lewa strona jest ujemna. Ponieważ $(\tau-\bar{t})^2 > (\tau-T)^2$ (bo $\bar{t} < T < \tau$, lewa strona nierówności (36) jest więc ujemna, gdy

$$\frac{1}{T-1} < p. \tag{37}$$



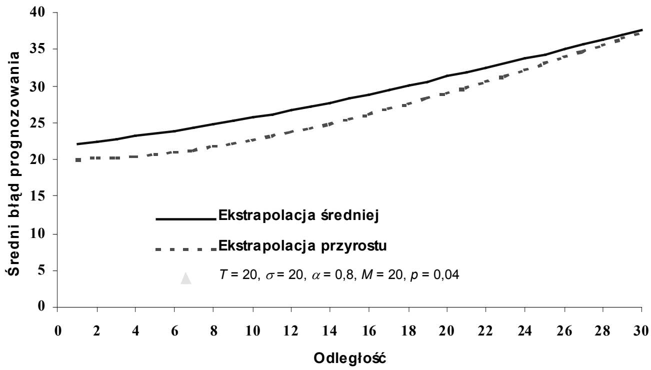
Rys. 3. Średni błąd prognozowania, p = 0,07



Rys. 4. Średni błąd prognozowania, p = 0,04

Średni błąd prognozowania dla metody ekstrapolacji średniego przyrostu empirycznego „na pewno” jest więc mniejszy niż dla metody ekstrapolacji średniej, gdy moduł współczynnika kierunkowego trendu w stosunku do odchylenia standardowego składników losowych jest relatywnie wysoki, nie mniejszy niż $1/(T-1)$ część odchylenia σ. Tego przypadku dotyczy rysunek 3. Jeśli natomiast trend jest prawie płaski w tym sensie, że współczynnik kierunkowy α jest mały i stanowi mniej niż $1/(T-1)$ część wartości σ, może być różnie – dla pewnych odległości prognozowania lepsza w sensie wielkości średniego błędu prognozowania, będzie metoda ekstrapolacji przyrostu empirycznego, a dla innych – metoda ekstrapolacji średniej (por. rys. 4).

3. Przyjmijmy, że zbiór S, na podstawie którego liczona jest średnia, stanowi M ostatnich obserwacji. Wtedy wzrostowi liczebności zbioru S odpowiada wzrost wartości $(\tau-\bar{t})^2$ i łatwiej spełnić warunek ujemności lewej strony (gwarantujący bardziej korzystny średni błąd prognozowania metody ekstrapolacji średniego przyrostu). Ilustruje to rysunek 5, nawiązujący do rysunku 4 (na którym $M = 3$).



Rys. 5. Średni błąd prognozowania, $M = 20$

6. Metoda ekstrapolacji średniego przyrostu a ekstrapolacja trendu liniowego oszacowanego klasyczną metodą najmniejszych kwadratów

Ekstrapolacja średniego przyrostu empirycznego jest pewną „prostą” alternatywą ekstrapolacji trendu liniowego:

$$Y^{\mathrm{Kmnk}} = b^K + ta^K, \tag{38}$$

którego parametry wyznaczono klasyczną metodą najmniejszych kwadratów. Na ogół jednak parametry oszacowanego trendu z klasycznej mnk nie są identyczne z parametrami trendu, wyznaczonego za pomocą przyrostów empirycznych:

$$Y^{EP} = (y_T - Ta) + ta\ .$$

Na ogół bowiem wyraz wolny trendu według klasycznej mnk nie jest równy (y_T – Ta) i na ogół wyznaczony według klasycznej mnk współczynnik kierunkowy różni się od współczynnika wyznaczonego (metodą momentów) według wzoru (4).

A. Średni błąd prognozowania dla ekstrapolacji oszacowanego trendu liniowego

Przyjmujemy założenia klasycznej regresji liniowej odnośnie do mechanizmu generującego przeszłe, obecne i przyszłe wartości zmiennej prognozowanej. Jeśli w warunkach tych założeń parametry oszacowanego trendu liniowego wyznaczono klasyczną mnk, to – jak wiadomo – średni błąd prognozowania wyraża się wzorem

$$\mu_\tau^{\mathrm{Kmnk}} = \sigma\sqrt{1 + \mathbf{x}_\tau'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_\tau}\ . \tag{39}$$

Z uwagi na możliwość uproszczeń przyjmijmy, że zmienną czasową t tak przekształcono, iż jej suma = 0 dla wszystkich obserwacji empirycznych[14]. Dodatkowo zakładamy, że przyrosty tak określonej zmiennej czasowej z obserwacji na obserwację są jednostkowe.

W takim wypadku:

$$\mathbf{x}_\tau = \begin{bmatrix} 1 \\ \tau \end{bmatrix},\quad (\mathbf{X'X})^{-1} = \frac{1}{T\sum t^2}\begin{bmatrix} \sum t^2 & 0 \\ 0 & T \end{bmatrix}. \tag{40}$$

Wtedy

$$\mu_\tau^{\mathrm{Kmnk}} = \sigma\sqrt{1 + \frac{1}{T} + \frac{\tau^2}{\sum t^2}}\ . \tag{41}$$

We wzorach tych T jest liczbą obserwacji empirycznych, τ jest wartością „przekształconej" zmiennej czasowej dla prognozowanego momentu czasu:

$$\tau = t_T + \Delta\ ,$$

gdzie:

t_T – wartość zmiennej czasowej w ostatniej obserwacji empirycznej,

Δ – odległość między momentem prognozowanym a momentem, którego dotyczy ostatnia obserwacja empiryczna (różnica odpowiednich wartości zmiennej czasowej).

Zatem

$$\mu_\tau^{\mathrm{Kmnk}} = \sigma\sqrt{1 + \frac{1}{T} + \frac{(t_T + \Delta)^2}{\sum t^2}}\ . \tag{42}$$

---

[14] Np. przy nieparzystej liczbie obserwacji t = , ..., –3, –2, –1, 0, 1, 2, 3, ...,

Przypomnijmy, że dla ekstrapolacji średniego przyrostu empirycznego

$$\mu_{\tau}^{EP} = \sigma \sqrt{\left(\frac{\Delta}{T-1}\right)^2 + 1]} . \tag{43}$$

B. Porównanie średniego błędu prognozowania

1. Na podstawie wzorów (42) oraz (43) stwierdzamy, że średni błąd prognozowania jest mniejszy dla metody ekstrapolacji średniego przyrostu empirycznego niż dla metody ekstrapolacji trendu liniowego oszacowanego klasyczną mnk, gdy

$$\left(\frac{\Delta}{T-1}\right)^2 < \frac{1}{T} + \frac{(t_T + \Delta)^2}{\sum t^2} . \tag{44}$$

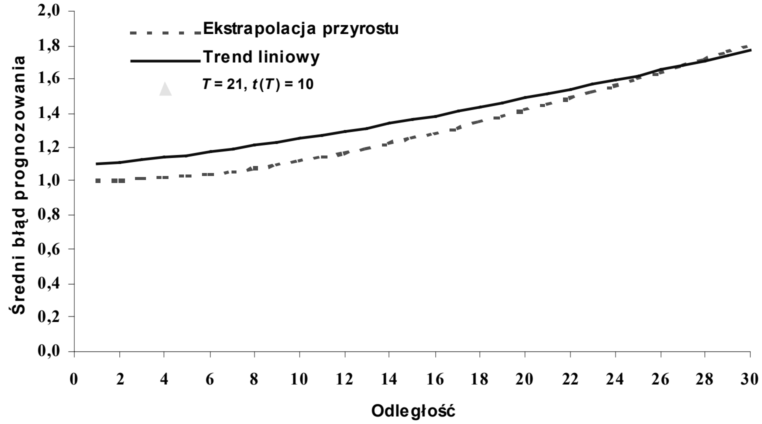
2. Średni błąd prognozowania dla ekstrapolacji średniego przyrostu empirycznego „na pewno" jest mniejszy od średniego błędu prognozowania dla ekstrapolacji trendu liniowego oszacowanego klasyczną mnk, gdy

$$\left(\frac{\Delta}{T-1}\right)^2 < \frac{1}{T} ,$$

co prowadzi do warunku

$$\Delta / (T-1) < \sqrt{1/T} , \tag{45}$$

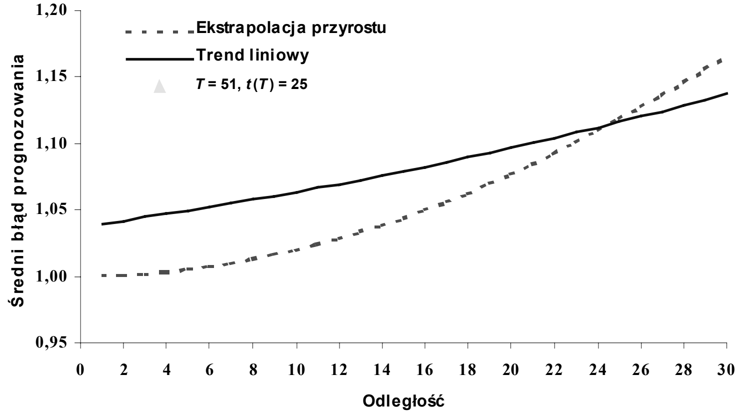
żeby udział odległości prognozowania Δ w wartości $T-1$ był relatywnie mały.



Rys. 6. Średni błąd prognozowania, $T = 21$

3. Przy odpowiednio niewielkich odległościach prognozowania, spełniających warunek (45), „na pewno" ekstrapolacja średniego przyrostu empirycznego daje mniejszy średni błąd prognozowania (por. rys. 6)[15].

4. Przy większych odległościach może być różnie (rys. 7). Dla pewnych – relatywnie mniejszych odległości – lepsza jest metoda ekstrapolacji średniego przyrostu empirycznego.



Rys. 7. Średni błąd prognozowania, $T = 51$

Dla relatywnie większych odległości prognozowania lepsza jest metoda ekstrapolacji trendu liniowego, oszacowanego klasyczną metodą najmniejszych kwadratów.

Bibliografia

[1] DITTMANN P., *Prognozowanie w przedsiębiorstwie*, Oficyna Ekonomiczna, Kraków 2003.
[2] GOLDBERGER A.S., *Teoria ekonometrii*, PWE, 1972.
[3] GUZIK B., JUREK W., *Ekonometria z zadaniami*, Wyd. AE, Poznań 1993.
[4] GUZIK B., APPENZELLER D., JUREK W., *Prognozowanie i symulacje. Wybrane zagadnienia*, Wyd. AE, Poznań 2005.
[5] PAWŁOWSKI Z., *Teoria prognozy ekonometrycznej w gospodarce socjalistycznej,* PWN, wyd. II, 1974.
[6] ZELIAŚ A., *Teoria prognozy*, PWE, 1997, paragraf 2.5.

[15] Jest to jednak warunek słaby. Nawiązując do sytuacji przedstawionej na rysunku 6, metoda ekstrapolacji średniego przyrostu byłaby lepsza dla $\Delta < 5$, bowiem $T = 21$. Na rysunku 6 widzimy, że byłoby tak dla $\Delta < 27$.

Mean error of prediction for a method of empirical growth extrapolation

The objective of this paper is to formulate a standard set of stochastic assumptions for a prediction method which consists in a linear extrapolation of the mean empirical growth.

The author shows how to derive formulas for the mean error of prediction (the ex ante prediction error). These formulas are then compared to the prediction errors of the following methods: the status quo method, the mean extrapolation method and the extrapolation of the linear trend function estimated by the least-squares method. This paper shows that the extrapolation of the mean empirical growth is more efficient than the status quo method and under some assumptions (that are defined in this article) is more efficient than the mean extrapolation method or the extrapolation of the linear trend function estimated by the least-squares method.

Keywords: *mean error of prediction, extrapolation of empirical growth*