

Metody prognozowania: Jakość prognoz

Dr inż. Sebastian Skoczypiec
ver. 03.2012

Wprowadzenie (1)

1. Sformułowanie zadania prognostycznego:
2. Określenie przesłanek prognostycznych:
3. Zebranie danych
4. Określenie metody i reguły prognozowania
5. Wyznaczenie prognozy
- 6. Oszacowanie przypuszczalnej trafności prognozy**
4. Wykorzystanie prognoz
5. Monitorowanie prognoz

Wprowadzenie (2)

- ✓ Szczególnie istotnym etapem prognozowania jest **weryfikacja prognoz i ocena ich jakości**, która ma podstawowe znaczenie dla ewentualnego wykorzystania tej prognozy. Jeżeli spełnia określone warunki dokładności i trafności może być zastosowana.
- ✓ Podstawowe znaczenie ma zatem **analiza błędów prognozy**.

Źródła błędów prognozy (1):

- ✓ **Błąd estymacji** modelu (oszacowane wartości wektora parametrów **a** różnią się od oryginalnych wartości wektora **a**).
- ✓ **Błąd struktury** stochastycznej modelu (oszacowane parametry rozkładu składnika losowego, w szczególności wariancja resztowa Se^2 różni się od wartości rzeczywistej σ^2).
- ✓ **Błąd losowy** (wartość składnika losowego w momencie prognozy jest różna od zera).
- ✓ **Błąd specyfikacji modelu** (zastosowano nieodpowiednią postać funkcji modelu lub nieodpowiedni zestaw zmiennych objaśniających).

Źródła błędów prognozy (2):

- ✓ **Błąd warunków endogenicznych** (nastąpiło zakłócenie postaci modelu).
- ✓ **Błąd warunków egzogenicznych** (wartości – w przypadku ich szacowania – zmiennych objaśniających uwzględnione w prognozie różnią się od rzeczywistych).
- ✓ **Błąd pomiaru.**

Okres, którego dotyczy sporządzana prognoza nazywa się **okresem prognozy** (T).

Długość okresu prognozy zależy od:

- od charakteru prognozowanego zjawiska,
- od praktycznych potrzeb prognozowania.

Liczba jednostek czasu, jaka upływa od teraźniejszości do okresu prognozowania nazywa się **wyprzedzeniem prognozy** (τ).

Horyzontem prognozy jest najdłuższy okres lub moment w przyszłości, w którym prognoza jest dopuszczalna w świetle przyjętego kryterium.

Przyjmujemy następujące oznaczenia:

n – ostatni okres, dla którego dysponuje się danymi statystycznymi dotyczącymi rzeczywistych realizacji zmiennej prognozowanej,

τ – przedział czasu oddzielający okres n od najdalszego okresu w przyszłości, dla którego prognoza jest dopuszczalna,

T – okres prognozy.

Horyzontem prognozy jest zatem taki przedział czasowy $[n, n+\tau]$, w którym dla każdego okresu $t = n+1, n+2, \dots, n+\tau$ można w sposób uzasadniony sporządzać dopuszczalne prognozy badanego zjawiska.

Dopuszczalne są więc prognozy dla takich okresów T , które nie wybiegają poza okres $n + \tau$.

- ✓ Prognoza jest naukowo uzasadnionym sądem o stanie zjawiska w określonym momencie (okresie) należącym do przyszłości.
 - ✓ Słowo „sąd” sygnalizuje niepewność prognozy.
 - ✓ Prognoza jest więc sądem o nieznanym.
 - ✓ Sądy bywają fałszywe lub prawdziwe.
 - ✓ O prognozach powiemy zaś, że są:
 - ✓ **trafne** – gdy okazują się wystarczająco bliskie realizacji prognozowanej zmiennej;
 - ✓ **nietrafne** (chybione) – gdy rozbieżność prognozy i wielkości prognozowanej okazuje się zbyt wielka jak na nasze potrzeby.
- **Błąd predykcji można oszacować tylko dla prognoz ilościowych.**

dr. inż. Sebastian Skoczypiec

Metody Prognozowania: Jakość prognoz 9

Ocenę dokładności i trafności prognoz dokonujemy stosując:



dr. inż. Sebastian Skoczypiec

Metody Prognozowania: Jakość prognoz 10

Dokładność i trafność prognozy:

Błąd prognozy:

ex ante -> dokładności prognozy.

przed upływem tego czasu, na który prognoza była ustalona

ex post -> trafność prognozy

po upływie czasu, na który prognoza była ustalona (znana jest realizacja zmiennej prognozowanej)

trafność prognozy

- Trafność prognozy określa się po upływie czasu, na który prognoza była wyznaczona
- Stopień trafności prognozy ilościowej mierzy się za pomocą **błędów ex post**
- Błąd ex post to wartość odchylenia rzeczywistych realizacji zmiennej prognozowanej od obliczonych prognoz
- Błędy ex post można obliczać dla każdego momentu lub okresu należącego do przedziału czasu $[n+1, \dots, T]$

Błędy prognoz ex post

- **błąd**
- (ang. **error**)

$$E_{\tau} = y_{\tau} - y_{\tau}^P$$

- **błąd procentowy**
- (ang. **percentage error**)

$$PE_{\tau} = \frac{y_{\tau} - y_{\tau}^P}{y_{\tau}} \cdot 100$$

Różnica $Y_{\tau} - Y_{\tau}^P$ (odchylenie realizacji zmiennej prognozowanej od wartości prognozy) **jest miarą błędu prognozy dla okresu τ**

PE określa, **jaki procent rzeczywistej realizacji zmiennej prognozowanej wynosi błąd prognozy**

Błędy prognoz ex-post

Załóżmy, iż wyznaczono m prognoz wartości zmiennej Y_{τ}^P oraz znane są rzeczywiste wartości zmiennej prognozowanej Y_{τ} dla $\tau = 1, 2, \dots, m$.

średni błąd
(ang. **mean error**):

$$ME = \frac{1}{m} \sum_{\tau=1}^m (y_{\tau} - y_{\tau}^P)$$

- wartość ME powinna być równa zero lub bliska zero;
- gdy zaobserwowane odchylenie od zera jest dodatnie, wnioskujemy, że prognozy wygaste są niedoszacowane;
- gdy zaobserwowane odchylenie od zera jest ujemne, wnioskujemy, że prognozy wygaste są przeszacowane.

Błędy prognoz ex post

- **średni procentowy błąd**
- (ang. *mean percentage error*):

$$MPE = \frac{\sum_{\tau=1}^m PE_{\tau}}{m}$$

MPE informuje, jaki procent rzeczywistych realizacji zmiennej prognozowanej stanowią błędy prognozy w okresie predykcji

Błędy prognoz ex post

- **średni błąd bezwzględny**
- (ang. *mean absolute error*):

$$MAE = \frac{1}{m} \sum_{\tau=1}^m |y_{\tau} - y_{\tau}^P|$$

MAE informuje o ile średnio - w okresie predykcji - rzeczywiste realizacje zmiennej prognozowanej będą się odchyłać - co do bezwzględnej wartości - od prognoz

Błędy prognoz ex post

$$MAPE = \frac{1}{m} \sum_{\tau=1}^m \left| \frac{y_{\tau} - y_{\tau}^P}{y_{\tau}} \right| \cdot 100$$

MAPE informuje o średniej wielkości błędów prognoz dla okresu $\tau = 1, 2, \dots, m$, wyrażonych w procentach rzeczywistych wartości zmiennej prognozowanej.

$$MAPE = \frac{MAE}{|\bar{y}^P|} \cdot 100$$

Wartości MAPE pozwalają porównać dokładność prognoz otrzymywanych różnych modeli.

Błędy prognoz ex post

- **średni błąd predykcji ex post**
- pierwiastek błędu średniokwadratowego
- (ang. **root mean square error**)

$$RMSE = \sqrt{\frac{1}{m} \sum_{\tau=1}^m (y_{\tau} - y_{\tau}^P)^2}$$

RMSE mierzy, o ile średnio odchylają się realizacje zmiennej prognozowanej od obliczonych prognoz

znacząca różnica wartości między MAE i RMSE wskazuje na występowanie w okresie prognozy błędów o bardzo dużych wartościach.

Błędy prognoz ex ante

- Ponieważ w chwili wyznaczania prognozy nie jest znana wartość rzeczywista zmiennej prognozowanej błąd prognozy ex ante może być tylko oszacowany.
- Wartość błędu ex ante przynosi informacje o oczekiwanych przeciętnych odchyleniach realizacji zmiennej prognozowanej od prognoz w czasie $t > n$.
- Błąd ex ante służy określeniu dokładności prognozy.

Błędy „ex ante” (1)

Błąd prognozy ex ante jest funkcją różnicy wartości prognozowanej i oczekiwanej rzeczywistej tj.

$$D_t = f(y_t - y_t^*)$$

Postać funkcji f zależy od metody prognozowania.

Wariancja to w [statystyce](#) klasyczna [miara zmienności](#). Intuicyjnie utożsamiana ze zróżnicowaniem zbiorowości jest [średnią arytmetyczną kwadratów](#) odchyłeń (różnic) poszczególnych wartości [cechy](#) od [wartości oczekiwanej](#).

Błędy „ex ante” (2)

1. Bezwzględny błąd *ex ante*:
Jest pierwiastkiem z wariancji prognozy

$$D_t = \sqrt{D_t^2}, \quad t > n$$

Informuje, jakich przeciętnych wahań zmiennej prognozowanej wokół jej wartości oczekiwanej można spodziewać

Tak otrzymany błąd prognozy jest wyrażony w tych jednostkach miary co zmienna prognozowana. Jest on wystarczający do wyboru spośród kilku modeli tej zmiennej takiego modelu, który daje najlepszą prognozę, czyli ma najwyższą wartość prognostyczną.

Błędy „ex ante” (2)

1. Względny błąd *ex ante*:
Jest pierwiastkiem z wariancji prognozy

$$V = \frac{D_t}{y_t^*} * 100\%, \quad t > n$$

Błąd ten stosuje się do porównywania dokładności prognoz różnych zmiennych.

dopuszczalność prognozy

- Prognoza jest dopuszczalna, gdy jest obdarzona przez jej odbiorcę stopniem zaufania wystarczającym do tego, by mogła być wykorzystana do celu, dla którego została ustalona.
- Dopuszczalność prognozy jest określona w tym samym czasie, w którym wyznacza się prognozę.

Błędy prognoz ex ante

- **średni błąd predykcji ex ante**

$$V_{\tau} = \sqrt{S^2(e) + x_{\tau} S^2(a_j) x_{\tau}^T}$$

gdzie:

- $S^2(e)$ – wariancja resztowa
- $D^2(a_j)$ – ocena wariancji estymatorów a_j (ocen parametrów strukturalnych)

- **względny błąd predykcji ex ante**

$$\eta_{\tau} = \frac{V_{\tau}}{y_{\tau}^P} \cdot 100$$

- wartość V_{τ} przynosi informację o **oczekiwanych przeciętnych odchyleniach realizacji zmiennej prognozowanej od prognoz w czasie $t > n$**
- wartość η_{τ} informuje jak wielki będzie w chwili $t > n$ oczekiwany błąd V_{τ} (odchylenie liczone w procentach wartości prognoz)

Kryteria dopuszczalności prognoz

- **subiektywne** kryteria dopuszczalności formułowane przez odbiorców prognozy;
- prognoza jest dopuszczalna, gdy spełniona jest jedna z poniższych relacji:

$$V_{\tau} \leq V_{\tau}^*, \quad t > n$$

lub

$$\eta_{\tau} \leq \eta_{\tau}^*, \quad t > n$$

gdzie: V_{τ}^* i η_{τ}^* to progowe wartości błędów zadane np. przez odbiorcę prognozy

Kryteria dopuszczalności prognoz

- **obiektywne** – przyjmuje się, że jeżeli względny miernik dokładności predykcji ex ante (lub ex post) spełnia nierówność:
 - $V_{\tau} \leq 3\%$, to prognozy są **bardzo dokładne**;
 - $3\% < V_{\tau} \leq 5\%$, to prognozy uznajemy za **dokładne**;
 - $5\% < V_{\tau} \leq 10\%$, to prognozy mogą być **dopuszczalne**;
 - $V_{\tau} > 10\%$, to prognozy są **niedopuszczalne**.