

# Metodi Statistici per l'Analisi Socio-Economica

## 6. Lisciamento esponenziale

# La previsione

La previsione si basa sull'idea che l'informazione sul passato ci permetta di prevedere gli eventi futuri.

Tre elementi della previsione:

- l'orizzonte temporale
  - ▶ breve
  - ▶ medio
  - ▶ lungo
- l'obiettivo
  - ▶ strumentale
  - ▶ tendenziale (neutrale)
  - ▶ condizionale
  - ▶ normativa
- il metodo
  - ▶ informale
  - ▶ serie storiche
  - ▶ regressione e modelli econometrici

# Lisciamento esponenziale

Il *lisciamento esponenziale* nasce nel 1957 come metodo pragmatico per la previsione delle serie storiche basato sulle medie mobili. In seguito esso è stato giustificato teoricamente anche nel quadro della teoria “moderna” delle serie storiche come caso particolare dei modelli ARMA/ARIMA.

- Si supponga di disporre, al tempo  $t$ , di una serie di osservazioni

$$y_{t-n}, y_{t-n+1}, \dots, y_{t-3}, y_{t-2}, y_{t-1}, y_t$$

e di voler prevedere  $y_{t+1}$

- Si potrebbe pensare di ricorrere a una media mobile “all’indietro” di alcuni termini
- L’idea alla base del lisciamento esponenziale è di modificare l’approccio delle medie mobili attribuendo più importanza alle osservazioni più recenti e in particolare all’ultima  $y_t$

# Lisciamento esponenziale semplice

Il *lisciamento esponenziale costante* o *semplice* parte dall'ipotesi che la serie sia stazionaria in media

- In prima approssimazione, dato che la serie è stazionaria in media, si potrebbe prendere come previsore in  $t + 1$  la media aritmetica delle osservazioni:

$$\hat{y}_{t+1} = \frac{\sum_{j=0}^n y_{t-j}}{n}$$

ma così si darebbe lo stesso peso a ogni osservazione.

- Il lisciamento esponenziale semplice generalizza quanto sopra assegnando a ogni osservazione un peso:

$$\hat{y}_{t+1} = \frac{\sum_{j=0}^n \omega_j y_{t-j}}{\sum_j \omega_j}$$

(nella media aritmetica è  $\omega_j = \frac{1}{n} \forall j$ )

# Determinazione dei pesi

Nel modello di lisciamento esponenziale costante si stabilisce che i pesi  $\omega_j$  decrescono esponenzialmente fino a 0 al crescere della distanza da  $t$ .

- Si impone:

$$\omega_j = (1 - \delta)\delta^j$$

con  $0 < \delta < 1$  e  $\sum_{j=0}^{\infty} \omega_j = 1$

- Sostituendo ricorsivamente ad ogni termine  $y_h$  la previsione fatta in  $h - 1$ :  $\hat{y}_h$ , si ottiene il seguente modello:

$$\hat{y}_{t+1} = (1 - \delta)y_t + \delta\hat{y}_t$$

dove  $\delta$  è chiamato *parametro di lisciamento* (smoothing).

# Il lisciamento esponenziale come correzione sequenziale degli errori di previsione

Si vede come il modello si fondi su una logica di *aggiornamento sequenziale*:

- la previsione a un passo  $\hat{y}_{t+1}$  è una media dell'ultimo termine e di tutti i precedenti, sintetizzati nella previsione precedente  $\hat{y}_t$ .
- Inoltre, riscrivendo la formula come

$$\hat{y}_{t+1} = (1 - \delta)y_t + \hat{y}_t - (1 - \delta)\hat{y}_t = \hat{y}_t + (1 - \delta)(y_t - \hat{y}_t)$$

si nota come la previsione corrente  $\hat{y}_{t+1}$  sia uguale alla precedente  $\hat{y}_t$  modificata per l'errore di previsione  $(y_t - \hat{y}_t)$  commesso al passo precedente, moltiplicato per il parametro di smussamento.

Si adotta pertanto una logica di correzione sequenziale degli errori di previsione.

# Come scegliere il parametro $\delta$ ?

A questo punto, rimane libero il parametro  $\delta$ :

- il criterio di *ottimalità* per la sua stima dovrà essere basato sull'impiego pratico del modello
- pertanto è naturale cercare  $\hat{\delta}$  tale da “minimizzare gli errori di previsione”, per esempio sotto forma di somma dei quadrati:

$$\min_{\hat{\delta}} SS(\hat{\delta}) = \sum_{t=1}^n (y_t - \hat{y}_t)^2$$

la stima viene ottenuta con metodi numerici (p. es. *grid search*)

- un altro problema (minore) è come *inizializzare* la serie dei valori previsti, ovvero cosa sostituire per  $\hat{y}_1$ : si può usare  $y_1$  o una media dei primi valori.

# Il metodo di Holt e Winters

Consideriamo una serie storica non stazionaria in media, che ammette

- un trend
- una componente stagionale

Se la serie storica ammette una tendenza di fondo *localmente rettilinea*, un modo di adattare il lisciamento esponenziale al caso è di scomporre il valore  $y_{t+1}$  in

- un livello medio in  $t$ ,  $L_t$ , e
- un trend  $T_t$  tra il tempo  $t$  e  $t + 1$

In generale, su un intervallo di lunghezza  $\Theta$ , il valore previsto della serie al tempo  $t + \Theta$  sarà esprimibile come

$$\hat{y}_{t+\Theta} = \hat{L}_t + \hat{T}_t\Theta$$

# Il metodo di Holt e Winters: stima - 1

Anziché stimare congiuntamente le due componenti, si scompone il procedimento utilizzando due modelli di lisciamento esponenziale:

- uno per il livello medio

$$\hat{L}_t = (1 - \alpha)y_t + \alpha(\hat{L}_{t-1} + \hat{T}_{t-1})$$

- e uno per il trend

$$\hat{T}_t = (1 - \beta)(\hat{L}_t - \hat{L}_{t-1}) + \beta\hat{T}_{t-1}$$

dove il primo modello ricostruisce, secondo il solito processo ricorsivo/di correzione dell'errore, il livello medio, il secondo il trend (il cui valore osservato  $T_t$  è definito come differenza tra i livelli medi).

## Il metodo di Holt e Winters: stima - 2

Nel caso vi fosse una componente stagionale *additiva*, ovvero

$$\hat{y}_{t+\Theta} = \hat{L}_t + \hat{T}_t\Theta + \hat{S}_{t+\Theta-s}$$

si aggiungerebbe una terza equazione:

- livello medio:

$$\hat{L}_t = (1 - \alpha)(y_t - \hat{S}_{t-s}) + \alpha(\hat{L}_{t-1} + \hat{T}_{t-1})$$

- trend:

$$\hat{T}_t = (1 - \beta)(\hat{L}_t - \hat{L}_{t-1}) + \beta\hat{T}_{t-1}$$

- stagionalità:

$$\hat{S}_t = (1 - \gamma)(y_t - \hat{L}_t) + \gamma\hat{S}_{t-s}$$

dove nell'equazione del livello medio  $(y_t - \hat{S}_{t-s})$  è il valore *destagionalizzato* di  $y_t$ , e nell'equazione della stagionalità  $(y_t - \hat{L}_t)$  è il valore *detrendizzato* di  $y_t$ .



## Il metodo di Holt e Winters: stima - 3

Nel caso vi fosse una componente stagionale *moltiplicativa*, ovvero

$$\hat{y}_{t+\Theta} = [\hat{L}_t + \hat{T}_t\Theta]\hat{S}_{t+\Theta-s}$$

- livello medio:

$$\hat{L}_t = (1 - \alpha) \frac{y_t}{\hat{S}_{t-s}} + \alpha(\hat{L}_{t-1} + \hat{T}_{t-1})$$

- trend:

$$\hat{T}_t = (1 - \beta)(\hat{L}_t - \hat{L}_{t-1}) + \beta\hat{T}_{t-1}$$

- stagionalità:

$$\hat{S}_t = (1 - \gamma) \frac{y_t}{\hat{L}_t} + \gamma\hat{S}_{t-s}$$

# Valutazione della qualità delle previsioni

Per valutare la qualità delle previsioni si fa spesso ricorso alla divisione del campione in

- dataset di *training*
- dataset di *test*

In generale, processi formali di valutazione della bontà delle previsioni necessitano di:

- un vettore di valori *previsti*  $p = [p_1, \dots, p_n]'$
- un vettore di valori *realizzati*  $r = [r_1, \dots, r_n]'$

# Previsione dei punti di svolta

In prima approssimazione, risulta interessante valutare la frequenza di corretta previsione dei *punti di svolta*. Definiamo *punto di svolta superiore* per una serie  $x_t$  il periodo  $t$ :

$$x_{t-1} < x_t, x_t > x_{t+1}$$

e *punto di svolta inferiore* per una serie  $x_t$  il periodo  $t$ :

$$x_{t-1} > x_t, x_t < x_{t+1}$$

Preso il totale, consideriamo la frequenza delle previsioni corrette e quelle degli errori

- di “prima specie” (svolta prevista ma non realizzata)  $n_{12}$
- di “seconda specie” (viceversa)  $n_{21}$

e costruiamo una tabella di contingenza.

## Previsione dei punti di svolta - 2

$p_t$ :	$r_t$ : p.s.	no p.s.	Totale
p.s.	$n_{11}$	$n_{12}$	$n_1$
no p.s.	$n_{21}$	$n_{22}$	$n_2$
Totale	$n_1$	$n_2$	$n$

Definiamo *indice relativo degli errori di I specie* il rapporto tra questi e il totale dei punti di svolta previsti:

$$E_1 = \frac{n_{12}}{n_1}$$

e *indice relativo degli errori di II specie* il rapporto tra questi e il totale dei punti di svolta realizzati:

$$E_1 = \frac{n_{21}}{n_2}$$

# Indici dell'errore medio di previsione

Per *errore di previsione* intendiamo  $e_t = p_t - r_t$ ; per sintetizzare gli  $e_t$  usiamo una funzione monotona non decrescente degli stessi che assuma valore 0 solo se  $e_t = 0 \forall t$ .

Per esempio, l'*errore medio assoluto* di previsione

$$EAM = \frac{1}{n} \sum_{t=1}^n |e_t|$$

che rispetto al semplice *errore medio*  $EM = \frac{1}{n} \sum_{t=1}^n e_t$  evita la compensazione di positivi e negativi: solo se vi è sistematicità negli errori di previsione sarà  $EM \sim EAM$ .

# Mean square error (MQE)

La *media quadratica degli errori di previsione* (internazionalmente, Mean Square Error o MSE)

$$\sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}$$

è un caso particolare della media potenziata  $I_s = \left(\frac{1}{n} \sum_{t=1}^n e_t^2\right)^{1/s}$   
Il quadrato del MQE/MSE si lascia scomporre in

$$MQE^2 = (\bar{p} - \bar{r})^2 + (\sigma_p - \sigma_r)^2 + 2\sigma_p\sigma_r(1 - \rho_{pr})$$

dove si vede che l'errore di previsione è dovuto a:

- diverse medie di  $p$  e  $r$  (*errore sistematico*)
- diverse variabilità di  $p$  e  $r$  (*errore nelle variabilità*)
- imperfetta correlazione lineare tra  $p$  e  $r$  (*errore nelle covarianze*)

# Confronti tra indici di bontà previsiva

Per confrontare tra loro indici di bontà previsiva bisogna standardizzarli, p. es. usando l'*indice di Theil*:

$$U = \frac{MQE}{\sqrt{\frac{1}{n} \sum_{t=1}^n r_t^2}} = \sqrt{\frac{\sum_{i=1}^n (p_t - r_t)^2}{\sum_{i=1}^n r_t^2}}$$

Per confrontare invece le performance di un previsore su due sotto-intervalli si usa l'*indice di Giano*:

$$J = \frac{\frac{1}{n_2} \sum_{t=n_1+1}^{n_2} e_t^2}{\frac{1}{n_1} \sum_{t=1}^{n_1} e_t^2}$$

che assume valori maggiori di 1 se la capacità previsiva peggiora.

