UNIVERSITÀ
DI SIENA
1240

Exponential Smoothing

Andrea Mecca[1,2]

[1]University of Florence [2]University of Siena

May 06, 2026

Contenuti

Recap sulle medie mobili

Cosa ricordate?

- Il processo generatore dei dati di una serie storica $y_1, y_2, \cdots, y_n$ per una variabile Y è descrivibile come

$$Y_t = f(t) + u_t$$

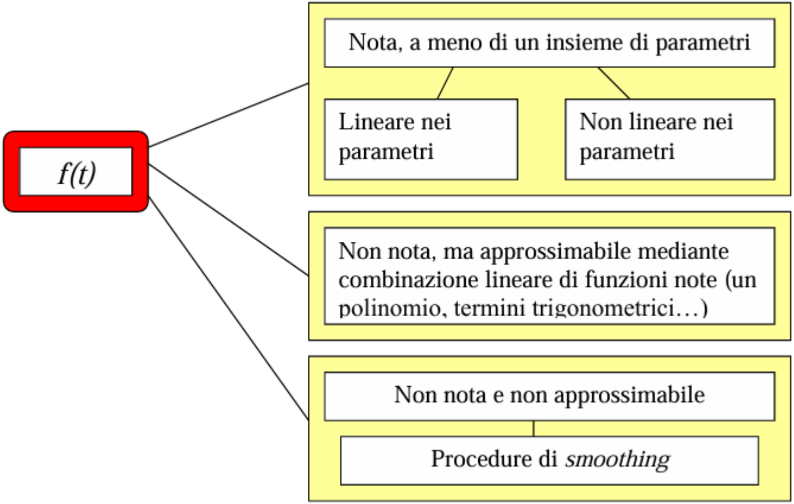
- $f(t)$ parte deterministica, u_t parte stocastica.

- Nell'**approccio classico** l'interesse è rivolto allo studio di $f(t)$.

- Si assume che la serie sia composta da:
  - **Trend:** componente di fondo del fenomeno (lungo periodo)
  - **Ciclo:** fluttuazioni dovute al succedersi nel fenomeno di fasi ascendenti e discendenti.
  - **Stagionalità:** movimenti del fenomeno che si ripetono nel medesimo periodo.
  - **Componente accidentale** $\sim$ componente erratica del modello di regressione.

- **Composizione**

In sintesi


f(t)
Nota, a meno di un insieme di parametri
Lineare nei parametri
Non lineare nei parametri
Non nota, ma approssimabile mediante combinazione lineare di funzioni note (un polinomio, termini trigonometrici…)
Non nota e non approssimabile
Procedure di smoothing

Procedure di smoothing

- Se $f(t)$ è estremamente irregolare $\to$ **Medie Mobili**

$$y_t^* = \theta_{-m_1} y_{t-m_1} + \theta_{-m_1+1} y_{t-m_1+1} + \cdots + \theta_{m_2} y_{t+m_2} = \sum_{i=-m_1}^{m_2} \theta_i y_{t+i}$$

- Permettono di individuare la componente di fondo senza conoscerne la legge sottostante.

- Caratteristihe di una media mobile:
  - **Ordine** $m_1 + m_2 + 1$
  - Una media mobile si dice **semplice** se tutti i termini hanno stesso peso.
  - Una media mobile si dice **centrata** se $m_1 = m_2 = m$. (se e solo se di **ordine dispari**)

$$y_t^* = \frac{1}{2m+1} \sum_{i=-m_1}^{m_2} y_{t+i}$$

  - **Media mobile simmetrica**

$$M = \{[5]; \left[\frac{1}{5}, \frac{1}{5}, \mathbf{\frac{1}{5}}\right]\}$$

Esercizio 1.

T	1996	1997	1998	1999	2000	2001	2002	2003	2004
Y_t	8	4	6	7	14	3	13	5	15

1. Data la seguente media mobile $y_t^* = \frac{1}{4}(y_{t-2} + y_{t-1} + y_t + y_{t+1})$, descrivere M e applicare M ai dati.
2. Applicare una media mobile centrata di ordine 3.
3. Applicare un'opportuna media mobile per non perdere i dati più recenti della serie.
4. Applicare la media mobile $M = \left\{[3]; \left[\frac{1}{4}; \mathbf{\frac{2}{4}}\right]\right\}$ e descriverla.

Lisciamento Esponenziale

Previsioni e tipologie

Previsione

Descrizione di **avvenimenti futuri** che si fonda su un **insieme coordinato di ipotesi**. È ragionevole suppore che, in alcune situazioni, le **informazioni sul passato** siano in grado di ridurre l'incertezza sul futuro.

a) Orizzonte temporale:

- breve o *congiunturale* (12-18 mesi);
- medio (5 anni);
- lungo termine (non oltre 10 anni).

b) Obiettivo:

- strumentale (avvertimento);
- tendenziale (dove finiamo con questa tendenza?);
- condizionale (se accade A allora probabilmente si verificherà B);
- normativa (come ottengo questo obiettivo).

Previsioni e tipologie

c) Metodo:

- informale o *naif* (soggettivo, saggezza accumulata);
- serie storiche (lisciamento esponenziale);
- modellistico.

- Il **lisciamento esponenziale** è un metodo semplice e flessibile da utilizzare per previsioni nel breve e brevissimo termine.
- È molto simile a quanto visto cone le medie mobili ma si tiene conto dei **valori passati osservati e previsti** con **scopo finale la previsione**.
- Consideriamo una serie $\{y_t\}_{t=1}^{n}$ e di **voler prevedere** y_{n+k} con $k \geq 1$, detto orizzonte temporale. Indichiamo la previsione fatta al tempo t con $\widehat{y}_{t+k}$.

Nozioni preliminari

- Se $y_t = a + \epsilon_t$ (trend costante con alterazioni accidentali), possiamo sfruttare l'informazione passata e ottenere la previsione come:

$$\widehat{y}_{n+1} = \frac{1}{n} \sum_{t=1}^{n} y_t = \frac{1}{n} \sum_{j=1}^{n} y_{n-j+1} \tag{1}$$

Verifica

Consideriamo per semplicità $n = 3$, avremo $\{y_t\}_{t=1}^{3} = \{y_1, y_2, y_3\}$, vogliamo prevedere y_4. Usiamo la formula in (1):

$$\widehat{y}_4 = \frac{1}{3} \sum_{t=1}^{3} y_t = \frac{y_1 + y_2 + y_3}{3} \qquad \widehat{y}_4 = \frac{1}{3} \sum_{j=1}^{3} y_{3-j+1} = \frac{y_3 + y_2 + y_1}{3}$$

- Sto utilizzando l'informazione passata **pesando ogni osservazione in egual modo** per ottenere la previsione ad orizzonte $k = 1$.

Il lisciamento esponenziale (I)

- Se ho cambiamenti più casuali, è più realistico attribuire alle osservazioni più recenti maggior peso:

$$\widehat{y}_{n+1} = \frac{\sum_{l=1}^{n} w_l y_{n-l+1}}{\sum_{l=1}^{n} w_l} \quad \text{with } w_l \geq 0; \quad w_l \leq w_{l-1} \quad l = 2, 3, \cdots, n$$

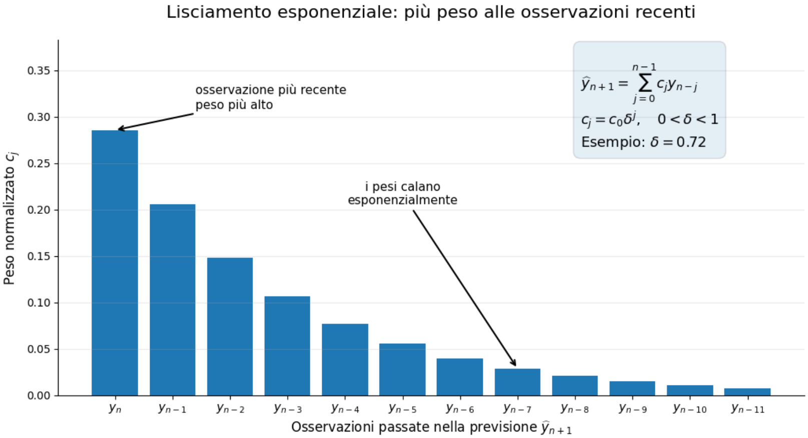
- O in maniera equivalente come

$$\widehat{y}_{n+1} = \sum_{j=0}^{n-1} c_j y_{n-j} \qquad c_j = \frac{w_{j+1}}{\sum_{l=1}^{n} w_l} \quad j = 0, 1, \cdots, n-1$$

- Il lisciamento esponenziale sfrutta questa idea e **calcola la previsione come media ponderata** di tutte le osservazioni passate con **pesi che seguono la successione esponenziale**, i.e.

$$c_j = c_0 \delta^j, \quad j = 0, 1, \cdots, n-1, \quad 0 < \delta < 1$$

Un esempio



Lisciamento esponenziale: più peso alle osservazioni recenti
osservazione più recente
peso più alto
$\widehat{y}_{n+1} = \sum_{j=0}^{n-1} c_j y_{n-j}$
$c_j = c_0\delta^j, \quad 0 < \delta < 1$
Esempio: $\delta = 0.72$
i pesi calano
esponenzialmente
Peso normalizzato c_j
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
y_n y_{n-1} y_{n-2} y_{n-3} y_{n-4} y_{n-5} y_{n-6} y_{n-7} y_{n-8} y_{n-9} y_{n-10} y_{n-11}
Osservazioni passate nella previsione $\widehat{y}_{n+1}$

Il lisciamento esponenziale (II)

- La previsione con costante di lisciamento δ è quindi ottenuta come

$$\widehat{y}_{n+1} = (1-\delta)\sum_{j=0}^{n-1} \delta^j y_{n-j} = \delta \widehat{y}_n + (1-\delta) y_n \tag{2}$$

- **È una media ponderata di tutte le osservazioni con pesi che diminuiscono all'aumentare della distanza dal tempo** n.
- Si tratta di una **procedura ricorsiva**.
- La costante δ è detta *vischiosità* del sistema (importanza della vecchia previsione):
  - per $\delta \to 0$, più peso ai nuovi dati y_n, meno alla previsione precedente previsione flessibile;
  - per $\delta \to 1$, peso quasi nullo ai nuovi dati, previsione simile alla precedente previsione rigida.

Le scelte

- Per ottenere il risultato dell'equazione (2) occore:

  1. Conoscere δ
  2. Avere un valore per la previsione iniziale $\widehat{y}_1$

- Per il problema 1. si pone generalmente $0.05 \leq \delta \leq 0.3$, oppure con i minimi quadrati si cerca:

$$\widehat{\delta} = \arg\min_{\delta} \sum_{t=m}^{n-1} (y_{t+1} - \widehat{y}_{t+1})^2$$

- Per il problema 2., se n è elevato e/o δ è piccolo la scelta è ininfluente, altrimenti:
  - $\widehat{y}_1 = y_1$ se $\delta \to 0$.
  - $\widehat{y}_1 = (y_1 + y_2)/2$
  - $\widehat{y}_1 = \frac{1}{n} \sum_{t=1}^{n} y_t$ se $\delta \to 1$

I metodi di Holt-Winters

In breve

- Il lisciamento esponenziale presenta delle limitazioni. La previsione ≈ **costante che meglio si adatta in prossimità di** n → trend regolare.
- Metodi di Holt-Winters sono una generalizzazione, appliccano il l.e. a trend e stagionalità dopo averle stimate.
- Idea è che in prossimità di n, una retta $L_n + T_n(t-n)$ sia preferibile ad una costante. Le equazioni di aggiornamento sono:

$$\text{peso attuale vs prevista} \quad \widehat{L}_n = \alpha(\widehat{L}_{n-1} + \widehat{T}_{n-1}) + (1-\alpha)y_n \quad 0 < \alpha < 1 \tag{3}$$

$$\text{peso al cambio nel livello} \quad \widehat{T}_n = \beta\widehat{T}_{n-1} + (1-\beta)(\widehat{L}_n - \widehat{L}_{n-1}) \quad 0 < \beta < 1 \tag{4}$$

- Considerando la stagionalità, la serie storica in prossimità di n è espressa come: $L_n + T_n(t-n) + S_n$ (metodo additivo) o $L_n + T_n(t-n) \times S_n$ (metodo moltiplicativo)

Valutare una previsione

Nozioni preliminari

1. *Date due o più serie di previsioni, quale ha fornito risultati migliori?*

2. *In che misura una serie di previsioni può considerarsi soddisfacente?*

- In genere si considera la distanza tra previsione e la sua realizzazione, si pone molta attenzione sui **punti di svolta**

  $\rightarrow$ punti in cui la serie cambia direzione, salto o rottura del trend.

- Formalmente, sia p_t il valore previsto al tempo t e r_t il valore realizzato. Si considerano i vettori dei valori p_t e r_t $\forall t = 1, \cdots, n$ e si costuisce la tabella a doppia entrata:

Valori previsti p_t	Valori realizzati r_t		
	p.s.	No p.s.	Tot
p.s.	n_{11}	n_{12}	$n_{1.}$
No p.s.	n_{21}	n_{22}	$n_{2.}$
Tot	$n_{.1}$	$n_{.2}$	n

Indici (I)

- Sulla diagonale principale della Tabella 19 si hanno le **previsioni rivelate esatte** $n_{11} + n_{22}$, n_{12} è il numero di **errori di prima specie** ps previsto e non realizzato mentre n_{21} è il numero di **errori di seconda specie** ps non previsto e realizzatosi.

- Indice degli errori di prima specie: $E_1 = \frac{n_{12}}{n_{1.}}$

- Indice degli errori di seconda specie: $E_2 = \frac{n_{21}}{n_{.1}}$

- Una prima valutazione/stima della bontà del modello previsivo utilizzato.

- Ma possiamo fare di meglio, consideriamo $e_t = p_t - r_t$ con $t = 1, \cdots, n$, un indice generale è l'**errore medio di previsione**, una media potenziata degli errori di previsione:

$$I_s = \left(\frac{1}{n} \sum_{t=1}^{n} |e_t|^s \right)^{1/s}$$

Indici (II)

- per $s = 1$, Errore Assoluto Medio di previsione:

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

- Spesso affiancato da Errore Medio di previsione:

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

- **Se i due sono molto vicini, c'è sistematicità degli errori** $\rightarrow$ MALE!!
- Media Quadratica degli Errori di previsione, per $s = 2$:

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

Indici (III)

- L' MQE può essere scomposto in tre componenti

$$ES = \frac{(\bar{p} - \bar{r})^2}{MQE^2} \quad \text{errore sistematico parto da livelli diversi} \tag{5}$$

$$EV = \frac{(\sigma_p - \sigma_r)^2}{MQE^2} \quad \text{errore nella variabilità livelli simili ma diversa variabilità} \tag{6}$$

$$EC = \frac{2\sigma_p \sigma_r (1 - \rho_{pr})}{MQE^2} \quad \text{errore nella covarianza dovrebbero coincidere, vorrei } \rho = 1 \tag{7}$$

  tali che $ES + EV + EC = 1$, vediamo da cosa e quanto dipende l'errore di previsione commesso.

- Per confrontare più previsioni relative a più serie storiche, **Indice di Theil**

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

Indici (IV)

- Infine, l'**Indice di Giano** viene utilizzato per valutare le previsioni in **due archi di tempo successivi**.

- Consideriamo che il periodo di n tempi sia suddiviso in due sottoperiodi di lunghezza n_1 e n_2 tali che $n_1 + n_2 = n$, l'indice è calcolato come:

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

- $J = 0$ quando a previsioni non tutte esatte nel primo sottoperiodo corrispondono previsioni perfette nel secondo, viceversa $J \to \infty$.

Esercizio 2.

Considerata la seguente serie storica per la variabile Y

T	**1996**	**1997**	**1998**	**1999**	**2000**
Y_n	8	4	6	7	14

Vogliamo **prevedere il valore di Y per l'anno 2001**, posti $\delta = 0.05$ e $\widehat{Y}_{1996} = 7.8$.

Grazie dell'attenzione!

Domande?

andrea.mecca2@unisi.it