La logistica: una curva semplice con molte applicazioni

Francesco Galvagno
Relatore: Franco Pastrone

Università degli studi di Torino
Scuola di Studi Superiori di Torino

Torino, 27 giugno 2017

SSST

Introduzione: modelli matematici

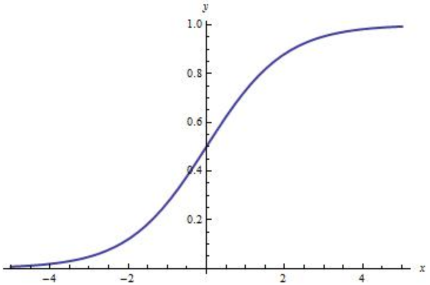
- Un **modello matematico** è una costruzione artificiale per riprodurre alcune proprietà di situazioni reali mediante gli strumenti della matematica.
- Grande applicabilità a diversi ambiti scientifici e non solo.

Nel lavoro di tesi e in questa presentazione:

- Si analizzano alcuni aspetti dell'**equazione differenziale logistica**.
- Si delineano alcuni semplici esempi di applicazione in alcuni settori specifici, ovvero biologia, medicina e statistica.

Equazione logistica

- Si tratta di un'equazione differenziale ordinaria del primo ordine:
$$y'(x) = y(x)\big(1 - y(x)\big)\,.$$
- Soluzione con condizione al contorno $y(0) = 1/2$:
$$y(x) = \frac{1}{1 + e^{-x}}$$
- Crescita esponenziale, poi rallentamento, cambio di concavità e saturazione asintotica a $y = 1$.



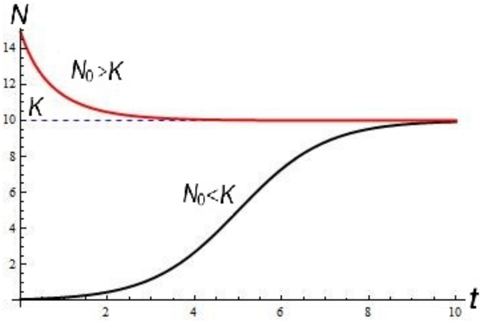
Dinamiche di popolazione (Verhulst,1838)

La curva sigmoide è nata come strumento nei modelli demografici in un ambiente con **risorse limitate**:

$$N'(t) = \alpha N(1 - N/K)$$

$$N(t) = \frac{KN_0}{N_0 + (K - N_0)e^{-\alpha t}}$$

dove N_0 è il numero iniziale di individui, K è la capacità dell'ambiente, α è il tasso di crescita.

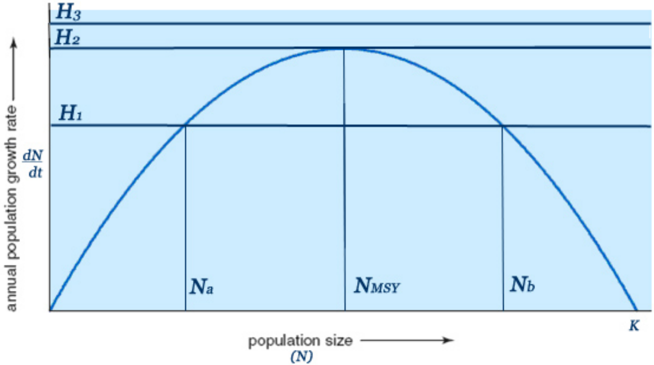


Utilizzo sostenibile di risorse biologiche

Tramite un modello logistico si studia la dinamica del livello di abbondanza di una risorsa $N(t)$ sottoposta ad un **tasso di raccolta** H:

$$N'(t) = F(N) - H$$

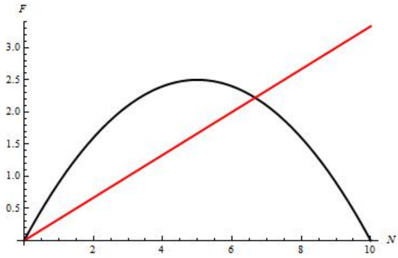
dove $F(N) = \alpha N(1 - N/K)$



Esempio: Risorse ittiche (Schaefer,1954)

L'andamento logistico di risorse biologiche è stato applicato nella gestione di **attività di pesca.**
Il tasso di raccolta H si modellizza come $H = H(N) = qEN$ dove:

- q (*catchability coefficient*) misura la facilità con cui il pesce viene pescato;
- E (*Effort*) è una misura delle risorse (navi,reti...) utilizzate nella pesca.



L'equazione diventa:

$$N'(t) = \alpha N(1 - N/K) - qEN$$

Dal punto di intersezione tra le due curve leggiamo la stima del livello di popolazione N all'equilibrio.

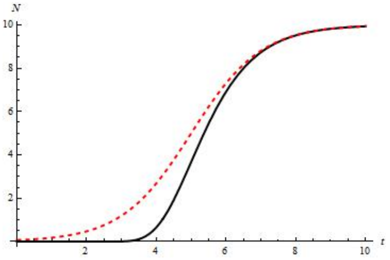
Funzione di Gompertz e crescita dei tumori (Laird, 1964)

Si parte da due presupposti:

- Tumore = popolazione cellulare che si sviluppa in uno spazio finito.
- Tumori non crescono con tasso costante, ma la crescita rallenta all'aumentare della dimensione della massa tumorale.

Quindi si utilizza una funzione logistica generalizzata (**curva di Gompertz**) che raggiunge la saturazione più lentamente.

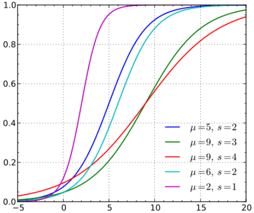
$$N(t) = \exp\left[\left(\log(N_0/K)\right)e^{-\alpha t}\right]$$



Funzione logistica come distribuzione statistica

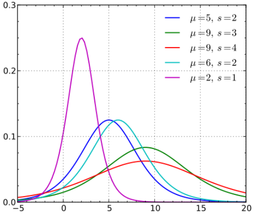
La funzione logistica rappresenta un ottimo esempio di **funzione cumulativa** per una variabile casuale x:

$$F(x) = \frac{1}{1 + e^{-\frac{x-\mu}{s}}}$$



La densità di probabilità assume la forma:

$$f(x) = F'(x) = \frac{e^{-\frac{x-\mu}{s}}}{s\left(1 + e^{-\frac{x-\mu}{s}}\right)^2}$$



Distribuzione simile a quella Gaussiana, con maggiore spessore nelle code $\Rightarrow$ eventi distanti dal valor medio μ più probabili.

Applicazione: regressione logistica

La curva logistica interpola bene set di dati dove la variabile dipendente è di tipo **categorico** (set discreto di possibili valori).

Esempio: gruppo di studenti che impiegano da 0 a 6 ore di studio per superare un esame.

- Regressione lineare in questo caso priva di significato.
- Punti in blu: set di dati plausibili (probabilità alta per > 4 ore, bassa per < 1.5 ore di studio)
- Responso binario: $0 =$ bocciatura, $1 =$ superamento.

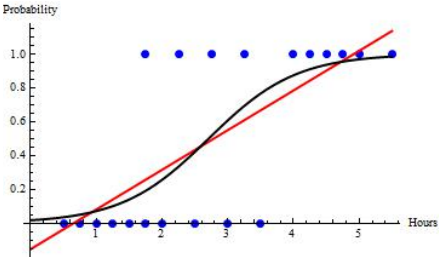


Figura: Probabilità di passare un esame rispetto alle ore di studio: confronto tra regressione logistica (curva in nero) e regressione lineare (curva in rosso).

Grazie per l'attenzione