

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



**Analisi dei tassi di fecondità tramite la distribuzione
normale asimmetrica flessibile**

Correlatore Prof. Scarpa Bruno
Dipartimento di Scienze Statistiche

Laureando: Angelica Zenere
Matricola N 1035575

Anno Accademico 2012/2013

Indice

1	Introduzione	3
2	I dati utilizzati	5
2.1	Il tasso di fecondità	5
2.2	Reperimento dati	6
3	Modelli per i tassi di fecondità per età	11
3.1	Modelli parametrici	11
3.2	Modelli semiparametrici	12
3.3	La Normale Asimmetrica	13
3.4	La Normale Asimmetrica Flessibile	14
3.5	Modello utilizzato	15
3.6	Minimi Quadrati	17
3.7	La Massima Verosimiglianza	17
4	La fecondità nei paesi europei	19
4.1	La stima per il Regno Unito	19
4.1.1	Stima Massima Verosimiglianza	20
4.1.2	Stima ai minimi quadrati	22
4.2	Intervalli di confidenza	25
4.3	Stima parametri complessivi per paese	30
5	La fecondità in Palestina	33
5.1	Stima dei parametri	33
5.1.1	Stima parametri via massima verosimiglianza	33
5.1.2	Stima dei minimi quadrati quadrati	36
5.2	Stima complessiva dei parametri	38
6	Analisi della distribuzione Normale Asimmetrica Flessibile	45
6.1	Interpretazione della distribuzione NAF considerando un condi- zionamento	45
6.1.1	La relazione empirica tra età e mode	45
6.1.2	rappresentazione stocastica	48
6.2	Interpretazione parametri	49
6.2.1	Parametrizzazione classica	49

6.2.2	Proprietà di simmetria	56
6.2.3	Seconda parametrizzazione	58
7	Conclusioni	63
A	Grafici	65
B	Codici programma R	71

Capitolo 1

Introduzione

Il tasso di fecondità è uno degli indicatori più usati in demografia, soprattutto per confrontare l'ammontare delle nascite tra popolazioni diverse (nel tempo o nello spazio o per altre caratteristiche). Questo tasso suddiviso per fascia d'età della madre, assume un comportamento che è possibile modellare tramite semplici modifiche di funzioni di densità univariate dato che è associabile ad una forma a campana, nonostante non sia una funzione di densità. Tuttavia vi sono alcune nazioni in cui l'andamento del tasso di fecondità, specifico per ogni età, è associabile a densità bimodale.

È interessante notare come si possono associare i tassi di fecondità ad una distribuzione specifica, in modo da poter collegare il loro comportamento, suddiviso per età, sotto un unico tipo di distribuzione. In questo modo, i risultati ottenuti si potrebbero utilizzare per fornire una automatizzazione delle stime così da poterli impiegare nell'effettuare previsioni sullo sviluppo nel tempo. Per questo motivo, si cerca una struttura che possa riassumere la distribuzione dei nati dell'anno di rilevazione rispetto alle fasce d'età delle madri e, attraverso l'interpretazione dei tassi di fecondità, ottenere la probabilità che una certa donna, di una specifica età o fascia d'età, possa avere un figlio.

In questo elaborato si cercherà di considerare una struttura che possa riassumere un modo esaustivo ma semplice i dati in possesso. In questo modo si cerca di fornire delle stime in cui si vede l'evoluzione del fenomeno per un esteso arco di tempo, in modo da fornire una possibile interpretazione dell'evoluzione comportamentale femminile in merito al concepimento dei figli.

Inoltre, utilizzando il modello Normale Asimmetrico Flessibile (NAF), si cercherà di ottenere un'interpretazione di ciò che una modifica dei parametri comporta sulla struttura dei dati. Questo sarà possibile tenendo in considerazione l'andamento della distribuzione NAF al variare di specifiche componenti il cui contributo alla struttura, è dato dalla loro combinazione.

Nel capitolo 3 sarà presentato come i dati raccolti sono stati reperiti tramite le istituzioni statistiche nazionali e come sono formulati. Successivamente, nel capitolo 2, si introducono alcuni modelli che solitamente sono utilizzati in analisi demografiche, specificando poi la struttura e le caratteristiche delle distribuzioni Normale Asimmetrica e Normale Asimmetrica Flessibile che saranno usate nelle argomentazioni successive. Inoltre sono descritti i metodi di stima utilizzati.

Elaborando i tassi del Regno Unito, nel capitolo 4, si stimeranno, attraverso i Minimi Quadrati e la Massima Verosimiglianza, i parametri della Normale Asimmetrica Flessibile con i rispettivi intervalli di confidenza. In seguito si proporrà una soluzione per ridurre la numerosità dei parametri che si riscontra nella stima di diverse annualità per una medesima nazione.

Per evidenziare l'efficienza dei modelli, anche nei casi in cui i dati sono aggregati in fasce d'età più ampie, nel capitolo 5, si propongono e si utilizzano i valori Palestinesi per l'analisi eseguita nel capitolo 4.

Avendo utilizzato la Normale Asimmetrica Flessibile, nel capitolo 6.1, si sviluppano delle analisi che permettono di fornire un'interpretazione di come la caratteristica della seconda moda, per i tassi, è legata alla distribuzione della popolazione femminile. Inoltre, nel capitolo 6.2, si è cercato di dare una interpretazione di come i parametri classici, ed una loro riformulazione, influiscono sulla struttura della distribuzione.

Capitolo 2

I dati utilizzati

2.1 Il tasso di fecondità

Per il lavoro svolto in questa tesi, si considererà principalmente la numerosità della popolazione femminile, possibilmente suddivisa per età, e i relativi tassi di fecondità. In base a questa necessità, come generalmente si fa in demografia, si andranno ad utilizzare solamente i dati delle donne tra i 15 e i 49 anni a causa della quasi nulla presenza di donne che, al di fuori di questa fascia di età, hanno avuto un figlio nell'anno della rilevazione. I dati raccolti comprenderanno un arco di tempo che rispecchierà la disponibilità delle fonti. Ad esempio per i dati riguardanti la popolazione femminile e per i tassi di fecondità raccolti nell'ambito europeo, si sono considerate le annualità tra il 1960 e il 2010, mentre, per l'esempio palestinese, dove la disponibilità di rilevazioni è minore, sono stati catalogati dati per poche annualità.

I tassi di fecondità come nella definizione introdotta dall'EUROSTAT (http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Fertility_statistics), sono il numero di figli per 1000 donne. Ciò che si utilizzerà, sarà più specificatamente il tasso di fecondità per ogni età di interesse e nel caso in cui non ve sia disponibilità, sarà utilizzato il tasso di fecondità che riassumerà fasce d'età più ampie. In questo caso specifico si considereranno intervalli d'età che comprendono 5 anni.

I tassi di fecondità sono visti come il rapporto tra il numero osservato di nati, nel corso di un anno, e la dimensione della popolazione femminile, infatti si possono formulare come

$$g(x) = \frac{N(x)}{D(x)}$$

in cui $N(x)$ rappresenta il numero di nati e $D(x)$ il numero di donne di età x . La loro struttura non è riconducibile ad una densità perché, attraverso una loro integrazione, si ottiene il tasso di Fecondità totale e non una somma unitaria.

Per poter impiegare i tassi di fecondità nelle analisi c'è la necessità di ricondurli ad una struttura in distribuzione, per cui li si divide per il tasso di fecondità totale.

2.2 Reperimento dati

I dati sono stati reperiti tramite le banche dati pubblicate dagli uffici statistici di ogni nazione. In particolare per la zona europea si è utilizzato il sistema di database dell'EUROSTAT (http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database) che fornisce in modo specifico i dati di maggior interesse. Per altre nazioni, come la Palestina, per cui non è possibile reperire i valori necessari attraverso le banche dati, si è dovuto far ricorso alle pubblicazioni dei censimenti presentate durante gli anni. I dati sono stati ottenuti attraverso pubblicazioni dell'ufficio statistico palestinese (http://www.pcbs.gov.ps/PCBS_2012/Publications.aspx?CatId=8&scatId=261).

Per le annualità che non presentavano alcuna rilevazione dettagliata, si è dovuto far ricorso agli elaborati che proponevano delle proiezioni introdotte a partire dalle osservazioni negli anni in cui i valori erano noti, e all'evoluzione del comportamento della popolazione.

Nell'ambito demografico, solitamente, si utilizzano dei modelli che comportano andamenti unimodali, mentre nel problema preso in considerazione in questo caso, vi è la necessità di poter sfruttare un modello con maggior capacità di adattamento in occasioni che si differenziano per una struttura complessa, e quindi non necessariamente unimodale. Nella selezione del modello, si cercherà di considerare una flessibilità tale che permetta l'adattamento che, per ogni nazione considerata, si possa adattare a tutti gli anni rilevati, e che per nazioni come l'Irlanda e il Regno Unito possa assumere una quasi bimodalità. Nel caso specifico andremo ad utilizzare i dati Irlandesi (Figura 2.1) o del Regno Unito (Figura 2.2) per cui è visibile l'andamento desiderato necessario per sfruttare a pieno le capacità del modello.

Come già accennato, una caratteristica dei dati europei, e quindi dei dati Irlandesi e del Regno Unito è che vi sono abbastanza osservazioni da poter ottenere un'analisi che sia specifica per ogni singola fascia d'età. Questa caratteristica, spesso non è riscontrabile in alcuni paesi a causa di un mancato sviluppo ottimale del sistema. Come dimostrazione di questo, sono i paesi del medio oriente che, a causa di guerre o relativamente nuove formazioni nazionali, non hanno un sistema di rilevazione efficiente che possa fornire dati esaustivi. A questo proposito viene introdotto l'esempio della Palestina che, essendo uno stato di recente costituzione (15 novembre 1988), non ha ancora sviluppato un sistema di rilevazioni efficace.

In questo caso abbiamo le informazioni necessarie, non per ogni singola età ma a gruppi di cinque anni. Questo ci permette di mostrare che, il modello

Figura 2.1: esempio dei tassi di fecondità in Irlanda

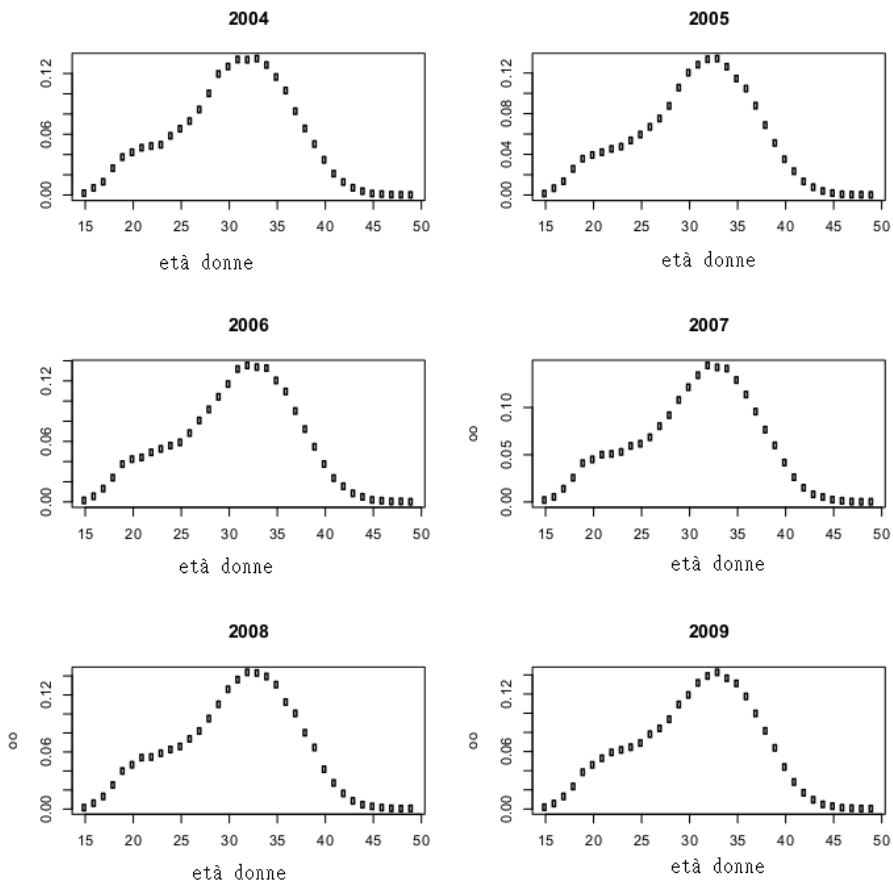
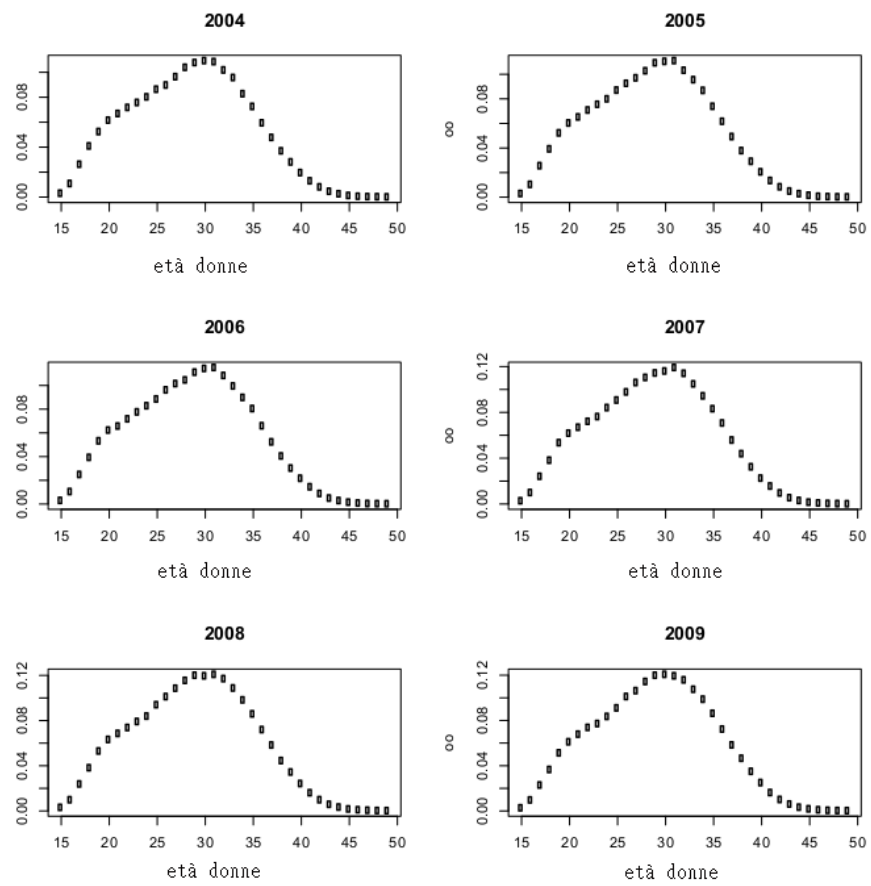


Figura 2.2: esempio dei tassi di fecondità in UK



strutturato successivamente, permette di studiare condizioni in cui i dati sono strutturati a classi di appartenenza più ampie di quelle unitarie.

Capitolo 3

Modelli per i tassi di fecondità per età

3.1 Modelli parametrici

Utilizzando come base di partenza i tassi di fecondità, la curva di fecondità può essere descritta come:

$$g(x; R, \theta_1, \dots, \theta_k) = R \cdot h(x; \theta_1, \dots, \theta_r)$$

in cui $h(\cdot; \theta_1, \dots, \theta_r)$ rappresenta una funzione di densità con k parametri, mentre R , in questo caso, rappresenta il tasso di fecondità totale (TFT).

Come suggerito da Hadwiger (1940), si possono utilizzare modelli per analizzare dati demografici. Principalmente si possono introdurre distribuzioni comuni come la Hadwiger (Gaussiana Inversa), la Gamma e la Beta.

Come prima proposta si considerano modelli come la Gamma o la Beta con cui si interpretano facilmente strutture unimodali e asimmetriche. Un esempio ne è la funzione di Hadwiger, proposta come:

$$g(x, a, b, c) = \frac{ab}{c} \left(\frac{c}{x}\right)^{\frac{3}{2}} e^{-b^2\left(\frac{x}{c} + \frac{x}{c} - 2\right)}$$

Un altro tipo di modello considerabile è quello Gaussiano in cui vi è la caratteristica principale di descrivere la simmetria delle strutture. Il modello gaussiano viene formulato come:

$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Anche se l'utilità di questi modelli, o comunque della tipologia unimodale, è presente, in alcuni casi la loro considerazione viene a mancare dato che essi non permettono la formazione degli hump o seconde mode.

Al fine di permettere la formazione di una quasi bimodalità vi è la necessità di combinare più di una distribuzione di quelle precedentemente citate. A questo scopo si introduce la mistura di modelli per ottenere una maggiore flessibilità e, considerando come base di partenza la funzione di Hadwiger, si ottiene il modello:

$$g(x, a_1, b_1, c_1, a_2, b_2, c_2) = m \frac{a_1 b_1}{c_1} \left(\frac{c_1}{x}\right)^{\frac{3}{2}} e^{-b_1^2 \left(\frac{c_1}{x} + \frac{x}{c_1} - 2\right)} + (1 - m) \frac{a_2 b_2}{c_2} \left(\frac{c_2}{x}\right)^{\frac{3}{2}} e^{-b_2^2 \left(\frac{c_2}{x} + \frac{x}{c_2} - 2\right)} \quad (3.1)$$

in cui i differenti parametri identificati con a possono essere riconosciuti come stima per R .

Come per la proposta di usare la funzione di Hadwiger in una mistura di modelli, si possono considerare le più comuni distribuzioni come la gamma, la beta e la normale da utilizzare in un contesto di mistura di modelli. Queste soluzioni hanno il vantaggio che, grazie alla presenza di molti parametri, possono fornire una diretta interpretazione per l'analisi dei dati. I differenti modelli comportano caratteristiche differenti, infatti con i modelli come quello di Hadwiger vi è la possibilità di strutturare forme asimmetriche, mentre non permettono di comporre forme simmetriche caratteristiche dei modelli Gaussiani. In controparte, la necessità di utilizzare misture di modelli per permettere una maggiore flessibilità, porta il peso di introdurre molti parametri rischiando poi di non riuscire a stimare adeguatamente il modello.

3.2 Modelli semiparametrici

Un secondo metodo per permettere un migliore adattamento ai dati, è quello di utilizzare delle basi di funzioni di tipo spline. Più precisamente si possono utilizzare basi di spline quadratiche del tipo:

$$g(x; R, \alpha, \beta, \theta_1, \theta_2, t_0, \dots, t_K) = R \cdot I(\alpha \leq x \leq \beta) \cdot \sum_{k=0}^{n_k} \theta_k (x - t_k)^2$$

in cui $I(\cdot)$ rappresenta la funzione indicatrice con limiti α e β , t_k sono i K nodi delle spline e θ_k sono i parametri. In questo caso si ha il vantaggio poter modellare le più svariate strutture rispetto alla proposta precedente, ma in controparte si ha sempre lo svantaggio di dover utilizzare molti parametri e l'interpretazione della formula non è immediata.

3.3 La Normale Asimmetrica

Considerando la necessità di un modello facilmente interpretabile, unito alla flessibilità e alla richiesta di pochi parametri, si nota che la struttura della famiglia delle distribuzioni Normali Asimmetriche fornisce una possibile alternativa. Inoltre questa distribuzione permette di esplicitare strutture sia asimmetriche che simmetriche, in quanto è possibile ricondurre alla formulazione della Normale

La Normale Asimmetrica ideata da Azzalini (1985) e formulata come

$$f(x; \mu, \omega, \alpha) = \frac{2}{\omega} \phi\left(\frac{x - \mu}{\omega}\right) \Phi\left(\alpha \left(\frac{x - \mu}{\omega}\right)\right)$$

con α parametro di forma, fornisce una struttura di base che opportunamente rielaborata permette di ottenere una forma che comprende la caratteristica di modulare l'asimmetria necessaria per l'analisi dei dati demografici.

In Azzalini (2005), vengono descritte delle proprietà della Normale Asimmetrica (SN) come:

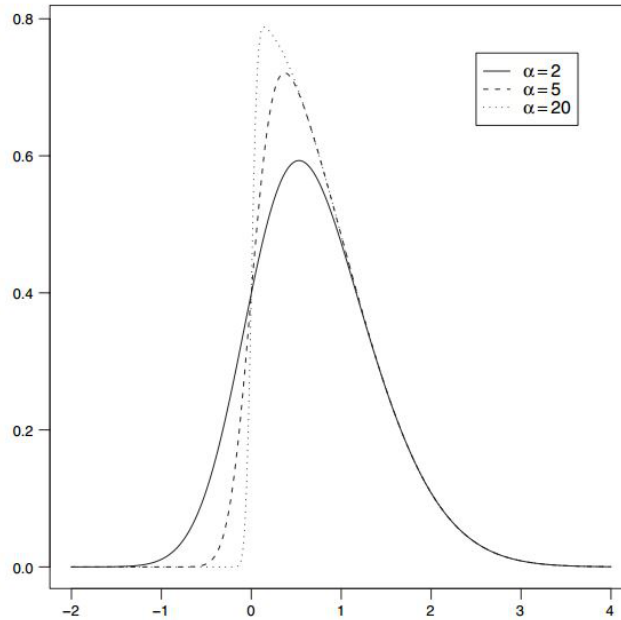
- se $\alpha = 0$ si ottiene la densità $N(0, 1)$
- se $Z \sim SN(\alpha)$, allora $-Z \sim SN(-\alpha)$
- con $\alpha \rightarrow \inf$ la distribuzione converge alla densità della metà-normale definita come $2\phi(z)$ per $z \geq 0$
- se $Z \sim SN(\alpha)$, allora $Z^2 \sim \chi_1^2$
- per un fissato α , la densità è fortemente unimodale

La distribuzione univariata SN può essere ottenuta tramite rappresentazione stocastica. Infatti la si può ritrovare considerando il modello normale bivariato

$$\begin{pmatrix} H \\ W \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \right)$$

dove le variabili H e W possono rappresentare due variabili correlate tra loro. Definendo la distribuzione di W , o una sua trasformata, condizionata ad una caratteristica d'interesse (ad esempio $U = W - E(W|H)$), si ottiene che $W|U > 0 \sim SN(\mu_2, \sigma_2^2, \delta/\sqrt{1 - \delta^2})$, dove $\delta = \sqrt{1 - \sigma_{12}^2 / (\sigma_1 \sigma_2)}$.

In Figura 3.1 si vede come si presenta la distribuzione della SN al variare di α . In questo caso, con asimmetria positiva (per valori $\alpha > 0$), si vede come da una distribuzione quasi simmetrica rappresentabile da $\alpha = 2$, la curtosi aumenta fino a portare la densità ad una metà-normale $\alpha = 20$.

Figura 3.1: densità della funzione $SN(\alpha)$ per diversi valori di α 

3.4 La Normale Asimmetrica Flessibile

Data l'incapacità della Normale Asimmetrica di formare un hump, vi è la necessità di generalizzare la formulazione in modo da poter configurare un'opportuna forma che permetta la doppia modalità. Questa necessità è accolta dalla formulazione della densità Normale Asimmetrica Flessibile (NAF) introdotta da Ma e Genton (2004).

Il modello NAF è definito come :

$$NAF(\mu, \omega, \alpha, \beta) = \frac{2}{\omega} \phi\left(\frac{x-\mu}{\omega}\right) \Phi\left(\alpha \frac{x-\mu}{\omega} + \beta \left(\frac{x-\mu}{\omega}\right)^3\right)$$

dove i valori di μ e ω rappresentano rispettivamente i parametri di posizione e di scala, mentre i valori di α e β sono riconducibili ai parametri di forma. Queste ultime due componenti hanno la caratteristica di poter strutturare la funzione in base alle differenti combinazioni dei valori assunti. Infatti avendo $\alpha = 0$ e $\beta = 0$ si ottiene una funzione di densità pari alla Normale (μ, ω^2) , mentre con $\alpha \neq 0$ e $\beta = 0$ si riduce alla formulazione della distribuzione Normale Asimmetrica. Anche in questo caso, come per la Normale Asimmetrica, il valore del parametro α definisce principalmente l'asimmetria della funzione.

Ma e Genton (2004) hanno dimostrato che la distribuzione Normale Asimmetrica Flessibile, comporta al più la presenza di una seconda moda o di un "hump". Questa caratteristica è discussa nel capitolo 6.

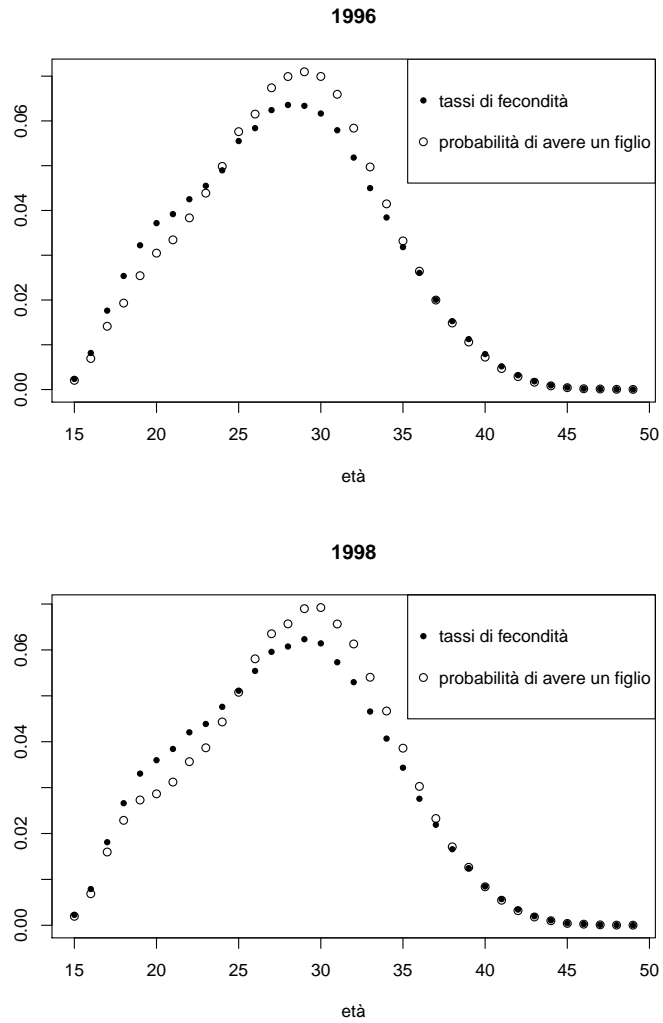
La proprietà di poter assumere una doppia modalità del modello Normale Asimmetrico Flessibile, è essenziale per alcune situazioni che altrimenti necessiterebbero di una modellazione alternativa con modelli di misture di distribuzioni. Queste strutture vengono scartate a causa della presenza di numerosi parametri. Grazie a questa capacità si può, infatti, riassumere e cogliere in pieno la struttura che ritroviamo dai tassi di fecondità e quindi la probabilità che una donna possa concepire un figlio data la sua età.

3.5 Modello utilizzato

L'analisi che si vuole sviluppare è strutturata sul concetto che si può determinare la probabilità di avere un figlio per una donna data un'età dal tasso di fecondità corrispettivo. In questo modo la struttura che ne risulterà sarà un modello che comprende tutta l'informazione sulle donne in un'unica volta, al posto di formulare un modello per ogni singola età.

Come già anticipato, si conosce che per l'Irlanda e per Regno Unito vi è la necessità di un modello che possa assumere una seconda moda per la struttura dei valori di fecondità. A questo proposito Mazzuco e Scarpa (2011) introducono il modello Normale Asimmetrico Flessibile che si adatta bene ai tassi di fecondità. Questo modello inoltre permette di ottenere una certa flessibilità in modo da renderlo adattabile a situazioni differenti così da ottenere una distribuzione unica che possa adeguarsi alle differenti situazioni.

Successivamente si andranno ad osservare due modalità di stima per i parametri dei modelli, per cui è doveroso specificare che le due proposte non riassumono esattamente lo stesso concetto. Infatti confrontando i tassi e la distribuzione delle probabilità di nascita appaiono differenti. Una dimostrazione di ciò è visibile in Figura 3.2 in cui si vede la differenza per due anni nei dati derivanti dalla Gran Bretagna. Questa differenza può essere attribuita al fatto che i tassi di fecondità sono espressi universalmente per un numero fissato di donne (1000) mentre, nella realtà, la popolazione femminile non è costante per tutte le età. Questo comporta che se vi è una notevole discrepanza di numerosità tra differenti fasce d'età, la similarità tra tassi di fecondità e probabilità di nascita, viene a mancare. Inoltre nell'ambito della distribuzione per le probabilità di nascita, queste rispecchiano il concetto di ricavare la probabilità di avere un figlio nella fascia d'età considerata.

Figura 3.2: differenza tra i tassi e le probabilità in UK

3.6 Minimi Quadrati

La metodologia dei minimi quadrati è una tecnica di ottimizzazione volta a determinare una funzione analitica che approssimi un insieme di dati senza necessariamente passare per i dati stessi (interpolazione), o meglio che si avvicini più possibile ad un'interpolazione di un insieme di dati (tipicamente punti del piano).

Sia $y = f(x; a, b, c, \dots, k)$ una funzione dall'andamento regolare che approssimi un numero finito di coppie di valori (x, y) appartenenti a due variabili X e Y . Così facendo, però, si commettono degli errori, per difetto o per eccesso a seconda che il punto si trovi sopra o sotto la curva. Il fine è quello di rendere questi errori quanto più piccoli possibile per ottenere una funzione che meglio si adatti ai dati originali. Per fare ciò, bisogna minimizzare gli scostamenti tra il valore osservato e quello teorico ma, per evitare compensazioni tra valori positivi e negativi, occorre minimizzare i quadrati di queste differenze. Ciò si attua determinando dei parametri reali a, b, c, \dots, k in modo tale che sia minima la funzione:

$$\phi(a, b, c, \dots, k) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

In particolare la funzione trovata deve essere quella che minimizza la somma dei quadrati delle distanze dai punti.

Nella stima dei Minimi Quadrati la funzione obiettivo è costituita dai tassi di fecondità, per cui, la formulazione del problema, si riconduce a dover stimare i parametri di una distribuzione rimodellata tramite il tasso di fecondità totale attraverso il confronto con i tassi. Questo è formulabile come:

$$g(x; R, \mu, \omega, \alpha, \beta) = R \cdot \text{NAF}(x; \mu, \omega, \alpha, \beta)$$

in cui $g(x; \dots)$ sono i tassi di fecondità e la Normale Asimmetrica Flessibile ($\text{NAF}(x; \dots)$), come distribuzione da modellare, viene moltiplicata per il tasso di fecondità totale (R).

In questo modo si stimano i parametri $\mu, \omega, \alpha, \beta$ e R che permettono il miglior adattamento della distribuzione sui dati.

3.7 La Massima Verosimiglianza

Dato che la presenza dei dati è caratterizzata da una struttura a raggruppamento per fasce di età, si potrà considerare la distribuzione Multinomiale per analizzare il numero di nascite per donna.

Il metodo di stima, in questo caso, è tramite la Massima Verosimiglianza. Per cui l'associazione tra il numero di nuovi nati rispetto l'età della madre è strutturata come :

$$\text{Multinom}(n.\text{don.tot}, \psi_{15} \dots \psi_{49}) = \frac{n.\text{don.tot}!}{\text{nati.et15}! \dots \text{nati.et49}!} \psi_{15}^{\text{don.et15}} \dots \psi_{49}^{\text{don.et49}}$$

Considerato questo modello, si vuole cercare una struttura per definire le probabilità per le singole età. Questa conformazione può essere facilmente associabile ad una funzione di distribuzione, dato che assumono un comportamento uguale: necessita di valori strettamente tra 0 e 1 e che abbiano somma unitaria.

Differentemente dal caso precedente, ora si formula il problema portando come funzione obiettivo il numero dei nati che saranno stimati grazie ai valori dei tassi di fecondità moltiplicati per il numero di donne in una fascia d'età.

Si pone quindi $N(x) \sim \text{Multinomiale}(\psi_1, \dots, \psi_j)$ dove j indica il numero di classi di età (annuali o quinquennali) considerate per le donne.

Ciascuna $\psi_j, j = 1, \dots, J$, misura la probabilità che una donna di età i abbia un figlio e, considerando il modello NAF per la curva di fecondità viene ottenuto come

$$\psi_j = \int_k^{k+h} \frac{2}{\omega} \phi\left(\frac{x-\mu}{\omega}\right) \Phi\left(\alpha \frac{x-\mu}{\omega} + \beta \left(\frac{x-\mu}{\omega}\right)^3\right) dx$$

dove abbiamo indicato con j la classe di età tra k e $k+h$ anni dove $h = 1$ per il Regno Unito e $h = 5$ per la Palestina.

indicando queste ψ_j nella funzione di probabilità Multinomiale otteniamo la funzione di verosimiglianza per μ, ω, α e β che una volta massimizzata forniscono la stima di Massima Verosimiglianza per i quattro parametri.

Capitolo 4

La fecondità nei paesi europei

In questo capitolo applichiamo il modello NAF ai dati del Regno Unito, come esempio delle fecondità dei paesi europei. Stimeremo i parametri del modello usando le due procedure discusse nei paragrafi 3.6 e 3.7

Volendo quantificare l'adattamento del modello ai dati, si andranno a confrontare i valori derivati dalla distribuzione Normale Asimmetrica Flessibile con le osservazioni. Nel verificare la bontà del modello e delle stime, come già accennato, si deve tener conto del fatto che, attraverso la stima di massima verosimiglianza e dei minimi quadrati, si formulano due confronti differenti. Infatti nel primo caso, la distribuzione si adatta direttamente sulla probabilità che una donna di una data età abbia un figlio, mentre nel secondo caso si adatta alla struttura dei tassi di fecondità. L'errore di stima, è quindi considerato, come la differenza tra i valori osservati e i valori ottenuti con il modello avente i parametri ottenuti con le due distribuzioni differentemente parametrizzate. Ovvero si può dire che il confronto avviene tra le probabilità di concepimento e i tassi di fecondità corretti con il tasso di fecondità totale in modo da poter comparare le due differenti quantità.

4.1 La stima per il Regno Unito

Un primo approccio ai dati è quello di introdurre una stima anno per anno. In questo modo, i valori utilizzati per formare il modello, si concentrano sui 35 dati ottenuti per età.

Inizialmente le stime riguarderanno i singoli anni, in modo da fornire una stima per ogni anno di rilevazione, così da permettere una visione dettagliata a specifica.

4.1.1 Stima Massima Verosimiglianza

La stima parametrica attraverso la Massima Verosimiglianza può sembrare la via più diretta per stimare una distribuzione. In questo caso, dato che le probabilità utilizzate nella Multinomiale sono definite da una distribuzione, ciò che realmente si andrà a stimare saranno i parametri che meglio adatteranno la distribuzione della Normale Asimmetrica Flessibile alla struttura dei figli corrispettivi per ogni fascia d'età femminile.

In Figura 4.1 si vede come la Normale Asimmetrica Flessibile si adatta alle probabilità di concepimento nelle varie fasce d'età.

Nella figura Figura 4.2 sono riportati i valori dei parametri per tutti gli anni di rilevazione. Si può notare come il parametro di posizione tenda a crescere; solo nella prima metà degli anni '80 c'è una lieve e temporanea tendenza a diminuire. Nei riquadri riguardanti la stima dei parametri di posizione μ e di forma α si notano che ci sono dei salti nelle stime dei due parametri. Questo comportamento può essere giustificato dal cambiamento di asimmetria, infatti nello stesso anno i valori del parametro α passano da valori positivi a valori negativi.

Nei riquadri relativi alla stima dei parametri di forma (α e β), si può notare che le loro stime sono spesso allo stesso livello, infatti concordano in segno, solo nei primi anni delle rilevazioni, fino al 1980, si può notare che la caratteristica della formazione del hump è accennata dalla differenza in grandezza del parametro β .

L'errore di adattamento nella scala delle probabilità del numero di nati risulta essere pari a 0.00777.

Figura 4.1: esempio di adattamento delle curve con parametri stimati con Massima Verosimiglianza

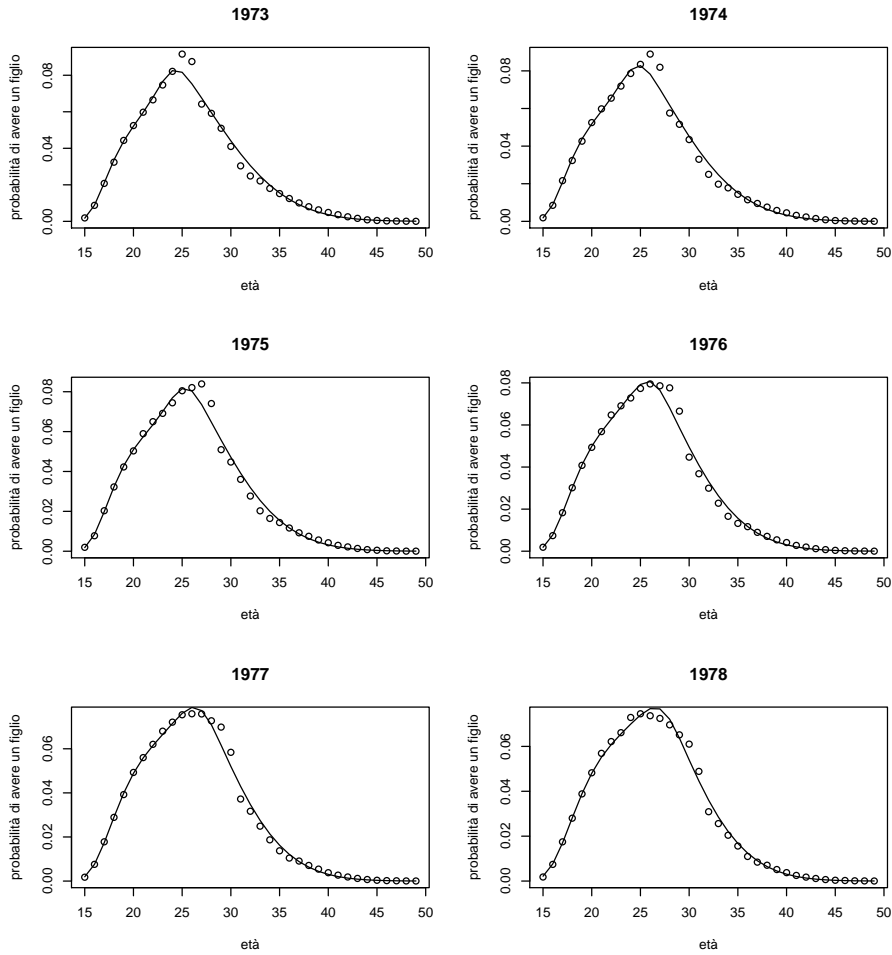
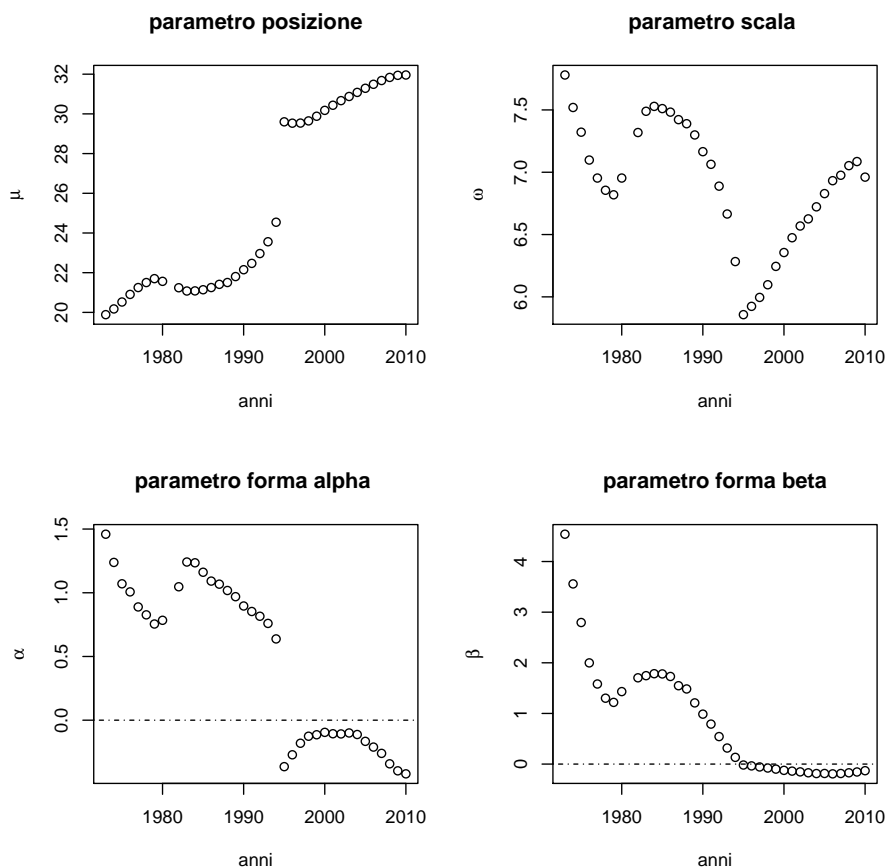


Figura 4.2: stima parametri via Massima Verosimiglianza



4.1.2 Stima ai minimi quadrati

Attraverso la stima ai minimi quadrati, si cerca di associare la distribuzione Normale Asimmetrica Flessibile ai tassi di fecondità. In questo caso si ha la necessità di stimare il valore del tasso totale di fecondità a posteriori, dato che si deve ricondurre la struttura dei tassi (non necessariamente in un intervallo 0 e 1) ad una distribuzione, e quindi ad una struttura di probabilità.

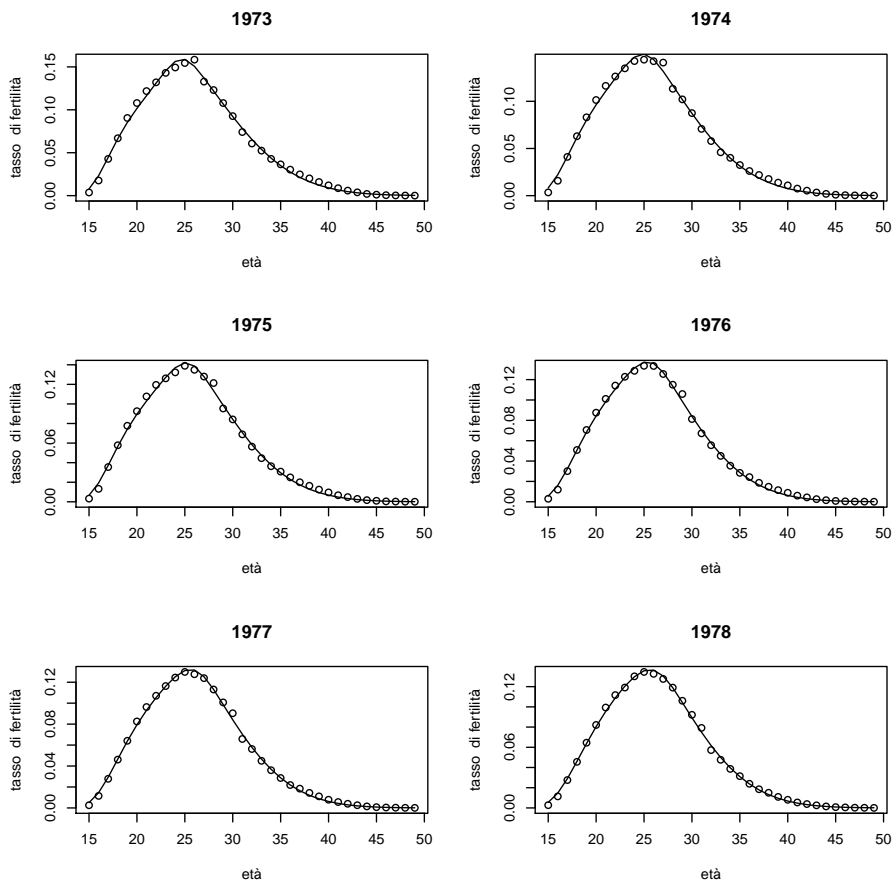
La formulazione usata è quindi

$$\text{NAF}(\mu, \omega, \alpha, \beta, R) = \frac{2}{\omega} \phi\left(\frac{x - \mu}{\omega}\right) \Phi\left(\alpha \frac{x - \mu}{\omega} + \beta \left(\frac{x - \mu}{\omega}\right)^3\right) R$$

Già precedentemente si era parlato della differenza dei dati su cui la distribuzione Normale Asimmetrica Flessibile viene imposta. Questa discrepanza viene

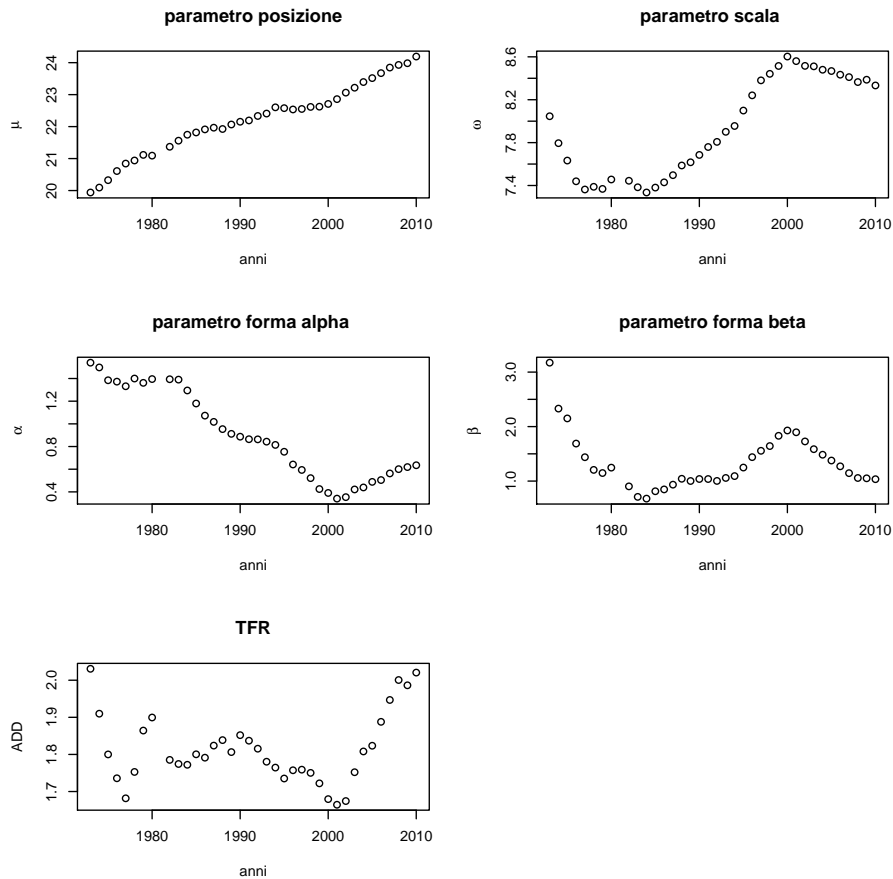
confermata dalla figura Figura 4.3, in cui si nota la differenza dei tre riquadri a destra rispetto a quelli mostrati precedentemente. In questo esempio Figura 4.3 si vede come la Normale Asimmetrica Flessibile si adatta ai tassi di fecondità alle varie età.

Figura 4.3: esempio di adattamento con parametri di Stima di Massima Verosimiglianza



Nei riquadri in Figura 4.4 sono riportate le stime ai Minimi Quadrati dei parametri per diversi anni, e si nota la differenza delle stime dei parametri attraverso i Minimi Quadrati rispetto alle stime derivate tramite Massima Verosimiglianza e come, la discrepanza tra tassi di fecondità e probabilità di avere un figlio, abbia una influenza sia nella rappresentazione della distribuzione (vedi Figura 3.2) che nella stima dei parametri. Differentemente dalle stime

Figura 4.4: stima parametri via Minimi Quadrati



precedenti, ora non si notano salti per i valori dei parametri, infatti il parametro di forma α non definisce alcun cambio di asimmetria.

Nonostante questo, l'interpretazione fornita per i parametri stimati con la Massima Verosimiglianza rimane valida. L'età delle madri è sempre maggiore e, fino ai primi anni '80, la differenza tra il parametro β e il parametro α è marcata. L'errore di adattamento in questo caso è di 0.004495.

Dal confronto di queste due metodologie, sembra apparire come migliore la stima attraverso i Minimi Quadrati, infatti vi è un errore di adattamento minore rispetto alla stima data dalla Massima Verosimiglianza. Il confronto tra queste due valutazioni può essere eseguito in quanto i valori confrontati con le distribuzioni stimate, sono riportati sulla stessa scala di osservazione.

4.2 Intervalli di confidenza

Una volta ottenuta la stima dei parametri è interessante ottenere una stima dell'intervallo di confidenza, ad esempio, per verificare se la necessità di avere parametri aggiuntivi, come α e β , sia confermata.

Come riportato ad esempio in Azzalini (2004), per ottenere le stime sono state utilizzate le verosimiglianze profilo per ogni parametro d'interesse, per cui si è ottenuto un intervallo di confidenza basato sul TRV (test del rapporto di verosimiglianza) a livello 0.95.

Una via semplice per reperire la verosimiglianza profilo per un parametro di interesse γ consiste nel sostituire, nella verosimiglianza complessiva $L(\gamma, \delta)$, il parametro di disturbo $\hat{\delta}_\gamma$ con la sua stima di massima verosimiglianza per γ fissato.

In un modello statistico parametrico con funzione di verosimiglianza $L(\Theta)$, in cui $\Theta = (\gamma, \delta)$ e γ è il parametro d'interesse, si definisce la verosimiglianza profilo per γ la funzione

$$LP(\gamma) = L(\gamma, \hat{\delta}_\gamma)$$

dove $\hat{\delta}_\gamma$ è la Stima di Massima Verosimiglianza di δ nel sotto-modello con γ fissato e si indica con $lP(\gamma) = \log LP(\gamma)$ la funzione di log-verosimiglianza profilo.

L'intervallo di confidenza, basato sul Test rapporto di verosimiglianza, viene creato considerando i valori di $\hat{\gamma}$ che soddisfano la disequazione

$$2(lP(\hat{\gamma}) - lP(\gamma)) < \chi_{0.95,1}^2$$

La stima di un intervallo di confidenza viene eseguita solo per le stime basate sulla verosimiglianza, in quanto, l'intervallo è estratto tramite verosimiglianza profilo e, anche se differenti, i parametri ottenuti con i Minimi Quadrati, mantengono il medesimo significato interpretativo.

Nei grafici Figura 4.5 e Figura 4.6, che propongono le stime degli intervalli di confidenza per i parametri di posizione e di scala, non vi è una visibile distinzione tra la stima del parametro e la stima dell'intervallo di confidenza perché la verosimiglianza profilo dei parametri è molto ristretta. Gli intervalli di confidenza dimostrano che c'è poca variabilità nelle stime dei parametri dovuta all'altissima numerosità di osservazioni considerate

Nei grafici di Figura 4.7 e Figura 4.8 vi è la dimostrazione che il modello sfrutta a pieno la necessità di poter formare un hump, come in questo caso. Infatti gli intervalli di confidenza stimati scartano pienamente l'ipotesi della nullità dei due coefficienti.

Figura 4.5: stima verosimiglianza profilo per μ con intervallo confidenza basati su test rapporto verosimiglianza

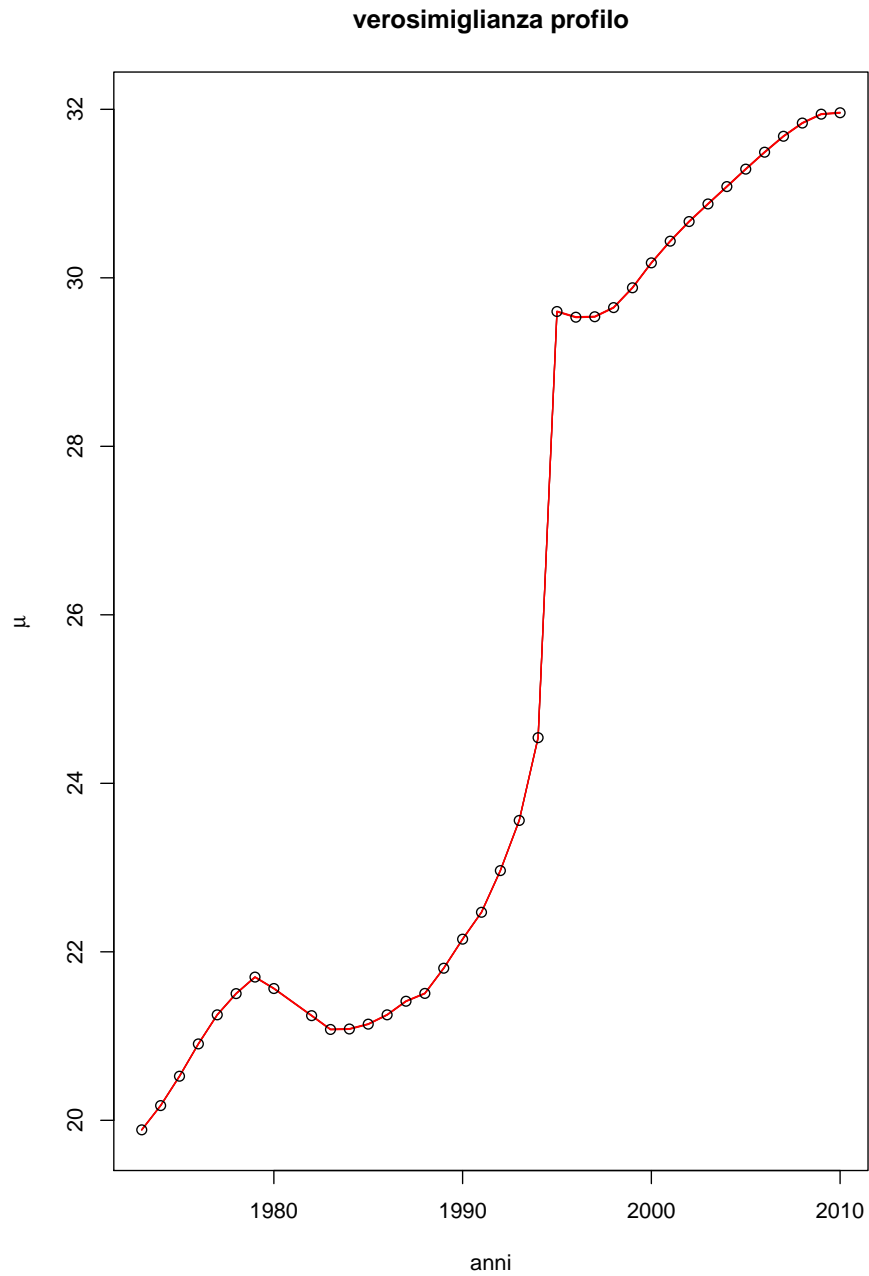


Figura 4.6: stima verosimiglianza profilo per ω con intervallo confidenza basati su test rapporto verosimiglianza

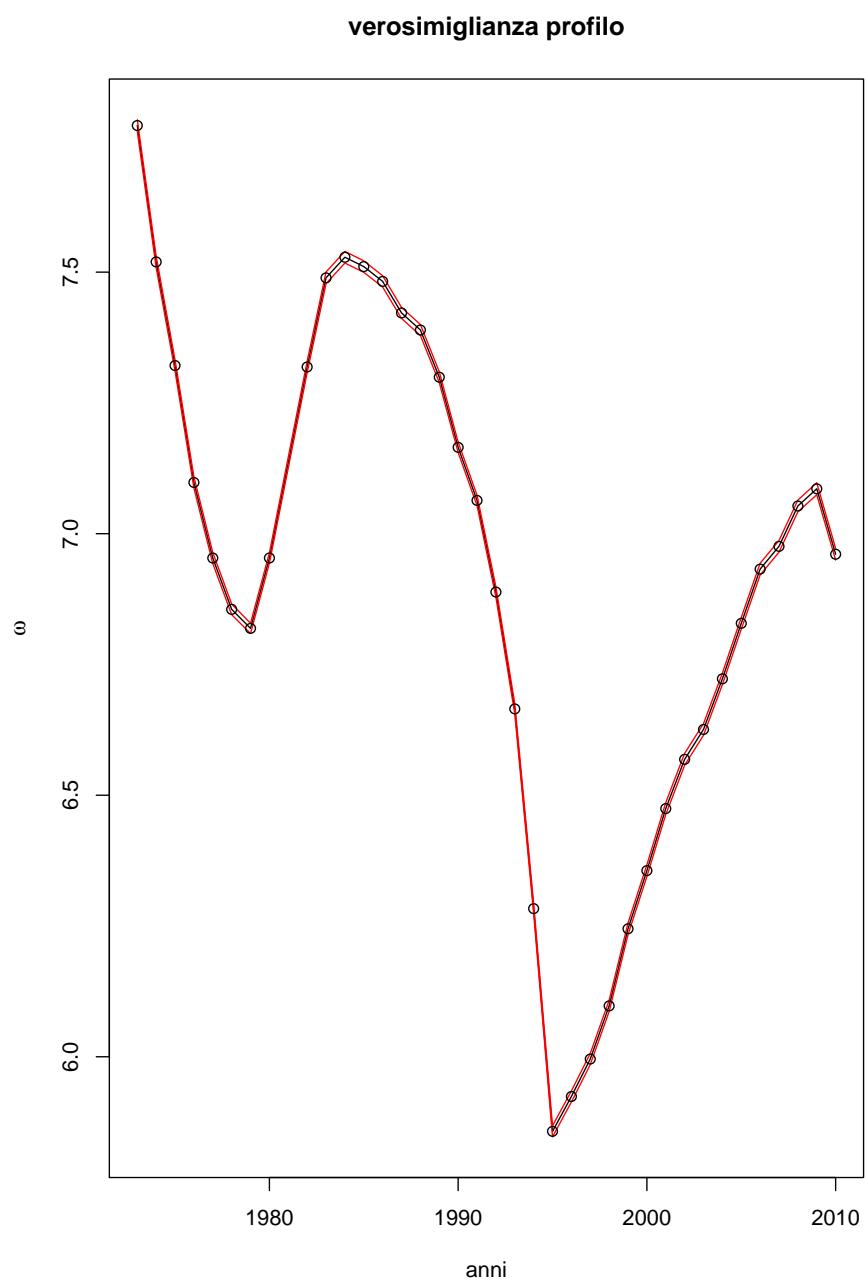


Figura 4.7: stima verosimiglianza profilo per α con intervallo confidenza basati su test rapporto verosimiglianza

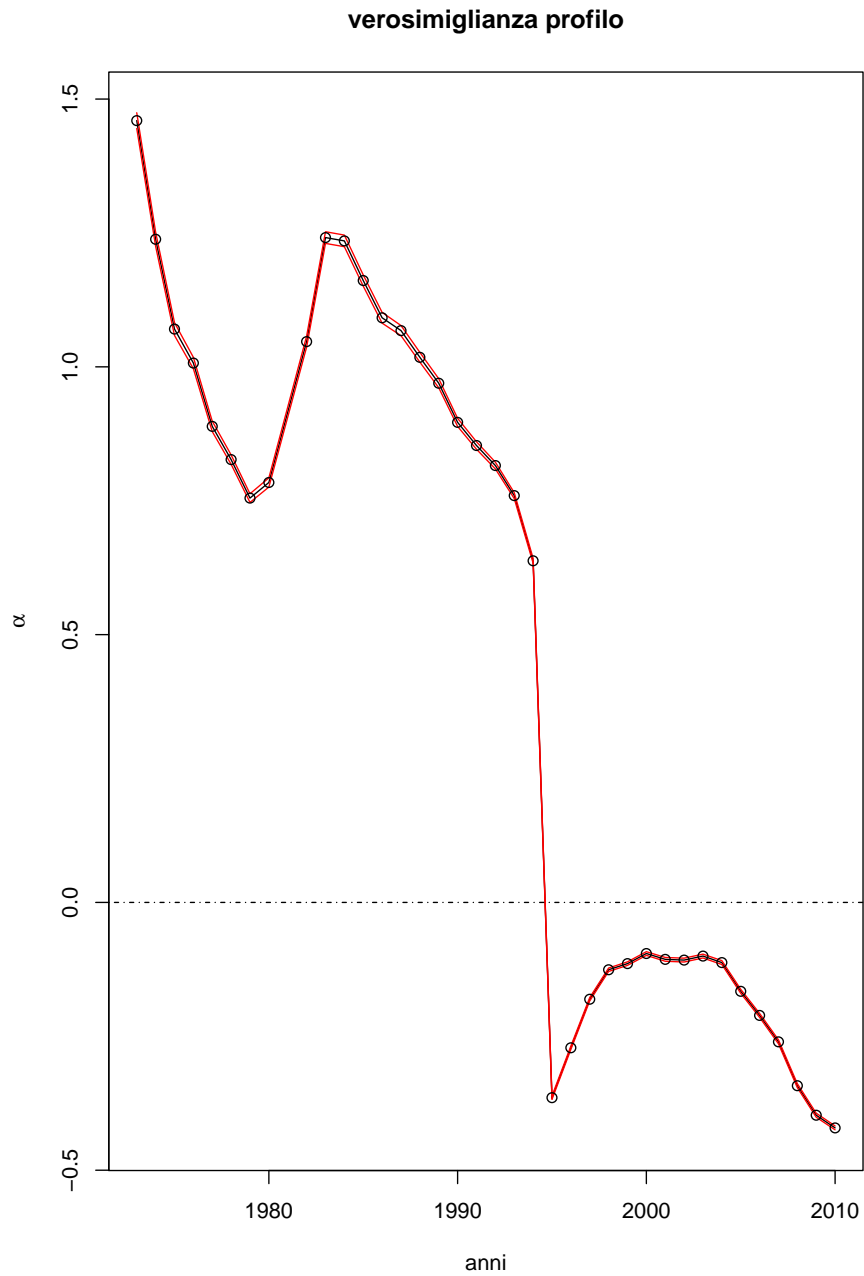
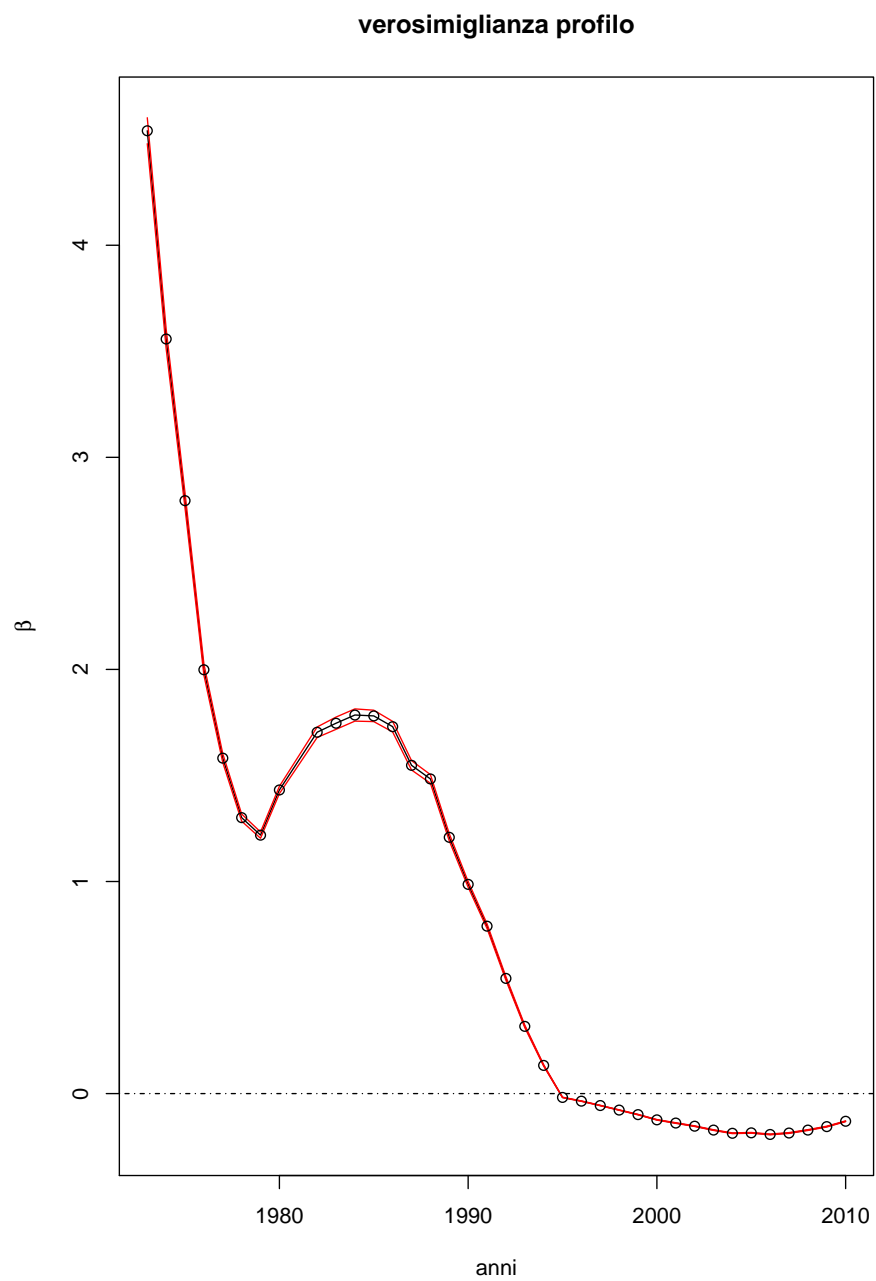


Figura 4.8: stima verosimiglianza profilo per β con intervallo confidenza basati su test rapporto verosimiglianza



4.3 Stima parametri complessivi per paese

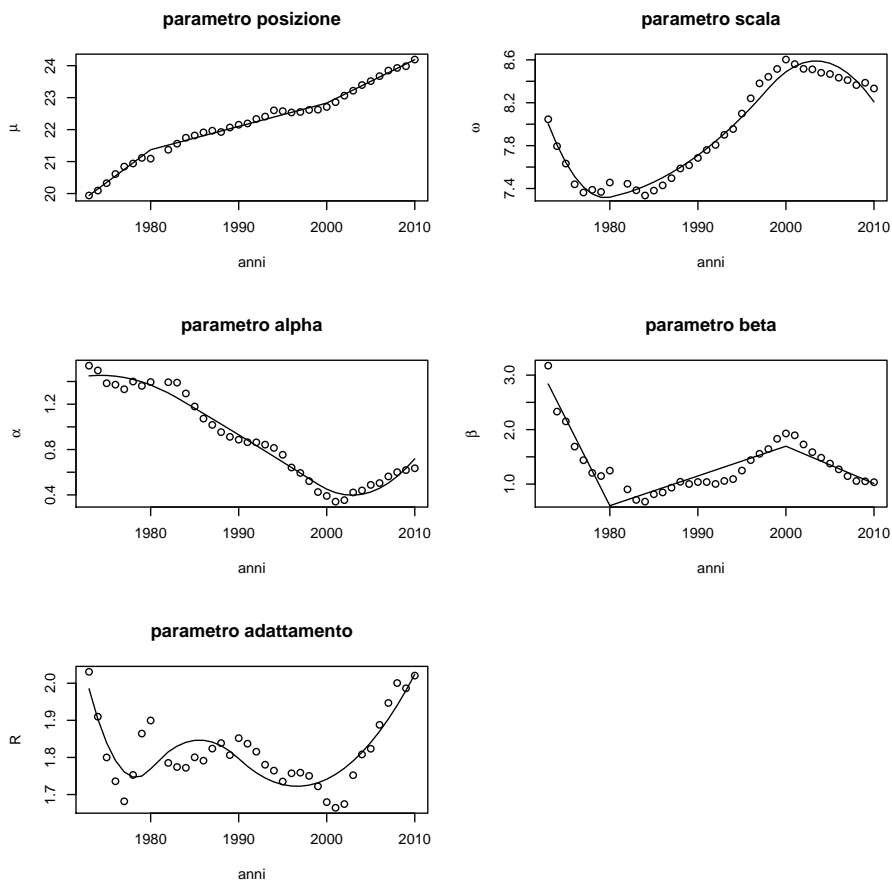
La stima annuale presenta degli svantaggi, infatti per ogni anno si devono stimare 4 o 5 parametri a seconda della stima considerata. Tenendo conto che si lavora con 37 anni (1973-2010), alla fine si ottengono fino a 185 parametri da stimare.

Questa significativa presenza di parametri potrebbe essere ridotta andando a riassumere la serie storica di valori ad esempio attraverso delle funzioni spline al più di second'ordine. Attraverso un'approssimazione dei valori con spline di regressione, con due nodi, si ottiene una significativa diminuzione dei valori da stimare.

Volendo focalizzare l'analisi sui tassi di fecondità piuttosto che sulla probabilità di avere un figlio, per le prossime considerazioni, si utilizzeranno le stime delle annualità dei parametri ottenute attraverso i Minimi Quadrati. Nella rappresentazione delle spline è necessario fissare alcuni nodi. Per ogni parametro vengono identificati, tramite ispezione grafica, i due anni che saranno utilizzati come punti di variazione.

Si andranno a formare delle basi di funzione di tipo spline di secondo grado per i parametri ω , α , R , con due nodi rispettivamente negli anni 1980-1998, 1983-1999 e 1980-1990. Mentre per i parametri μ e β è stata costruita una spline di primo grado sempre con due nodi negli anni 1980-2000 in entrambi i casi. I parametri, quindi, si sono ridotti alla numerosità di 23 elementi.

Figura 4.9: stima dei parametri complessiva con spline



Capitolo 5

La fecondità in Palestina

5.1 Stima dei parametri

L'esempio dello stato Palestinese fornisce statistiche sulla popolazione riunite in fasce d'età di cinque anni.

I dati sono strutturati come in precedenza, ovvero vengono considerati i tassi di fecondità per fascia d'età e il numero di donne corrispondente per l'età e anno. Dato che non è stato possibile ottenere valori completi per più anni di fila, si è tenuto conto solamente delle annualità 1999, 2003, 2006 e 2010 in quanto vi è un'equa distanza temporale tra di loro e i dati forniti sono completi sia per la fecondità che per la popolazione femminile. In questo particolare caso, i valori di interesse sono stati raccolti tramite pubblicazioni di rilevazioni effettuate in anni differenti e non consecutivi.

Come precedentemente sono stati implementati entrambi i metodi di stima. Sono state effettuate le stime per le classi d'età 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49.

5.1.1 Stima parametri via massima verosimiglianza

Anche in questo caso, tramite la stima di massima verosimiglianza, si ottengono le stime dei quattro parametri della Normale Asimmetrica Flessibile. Come precedentemente si confronta la stima del modello parametrico con le effettive probabilità del modello Multinomiale (Figura 5.1).

Considerando le quattro annualità, le stime dei parametri sono riportate in Figura 5.2.

Da queste stime si nota come, con il passare degli anni l'età media delle donne che hanno un figlio, si alza. I parametri di forma α e β indicano una asimmetria positiva con la tendenza ad avere una seconda moda con l'avanzare dell'età. Questa caratteristica non è accentuata dai grafici di Figura 5.1 a causa della poca quantità di dati rispetto alla numerosità dei parametri (28 rilevazioni

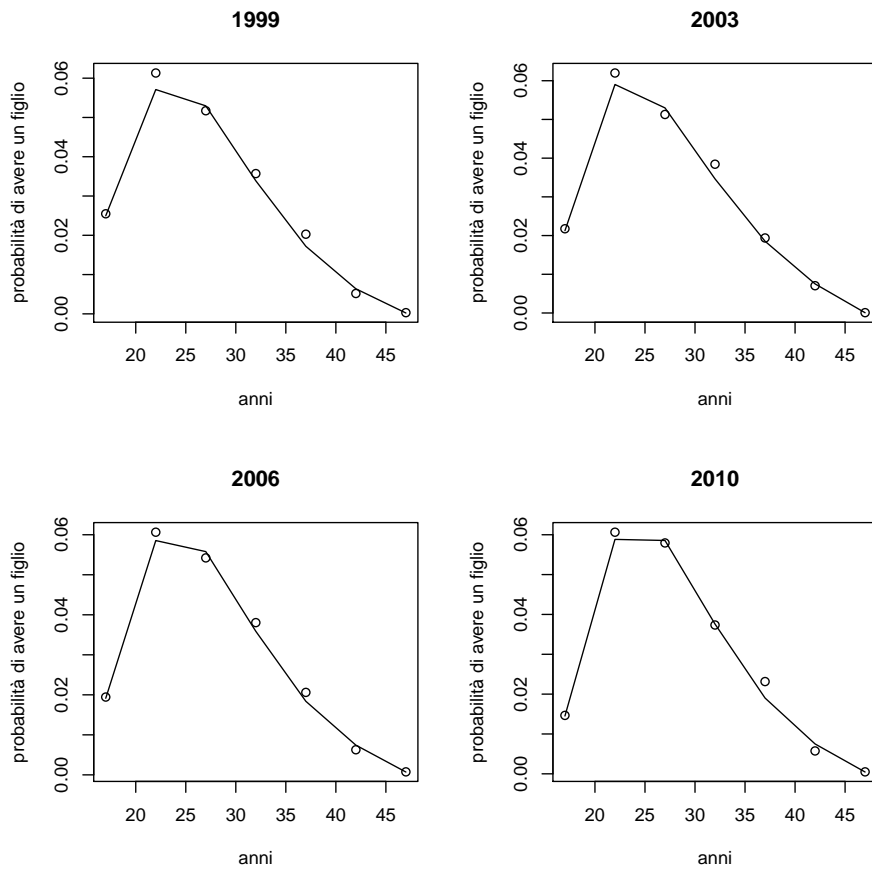
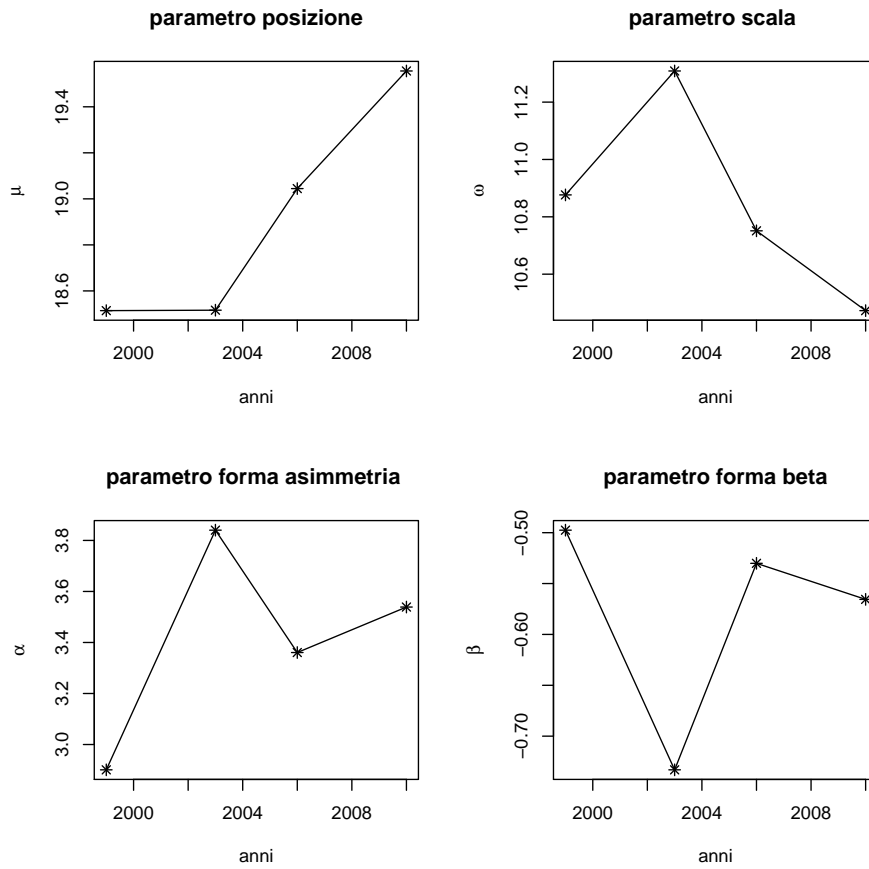
Figura 5.1: esempio di adattamento con parametri di Stima Massima Verosimiglianza

Figura 5.2: stima tramite la massima verosimiglianza

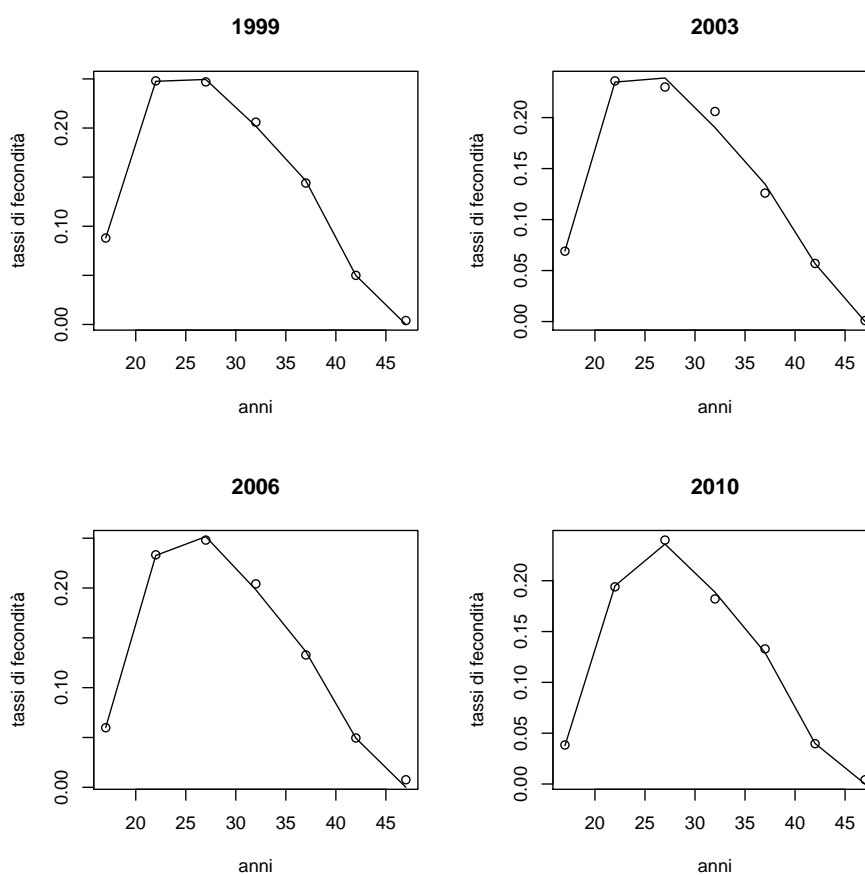
rispetto a 20 parametri).

L'errore di adattamento riscontrato è: 0.000103

5.1.2 Stima dei minimi quadrati quadrati

Anche in questo caso si nota la diversità dei dati su cui andiamo ad adattare la Normale Asimmetrica Flessibile, infatti la differenza si nota in quanto i valori delle età 25 e 30 della Figura 5.3 hanno i valori per le osservazioni indicate visibilmente differenti rispetto quelle della Figura 5.1.

Figura 5.3: esempio di adattamento con parametri di Stima con i Minimi Quadrati



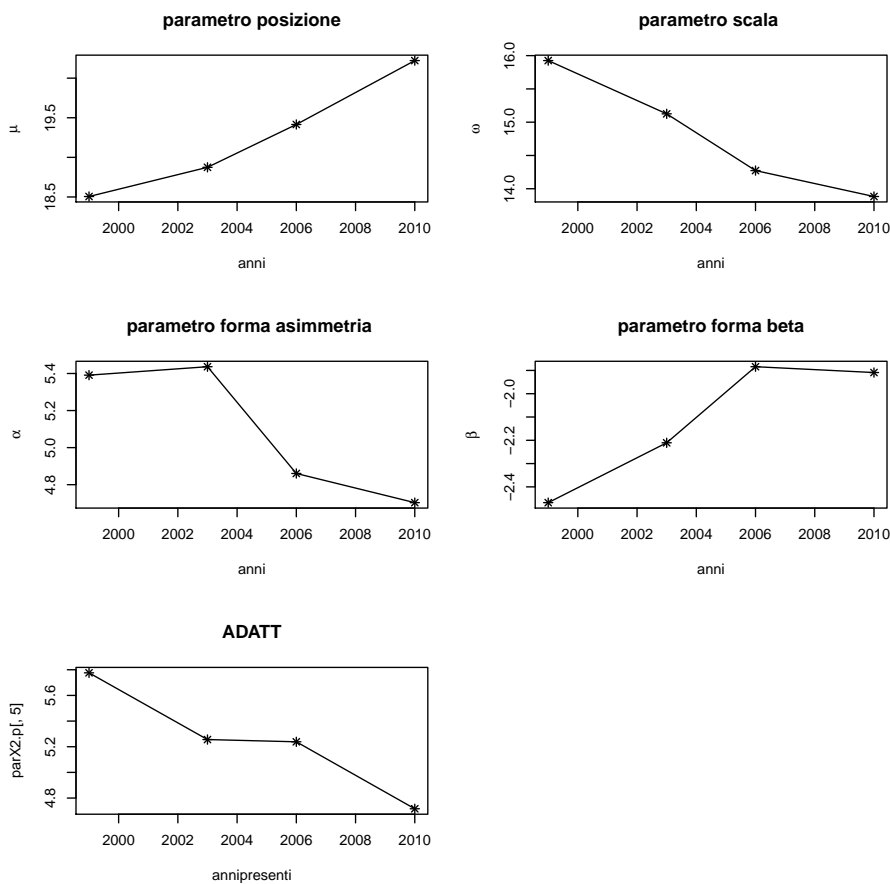
La stima dei parametri (Figura 5.4), anche se con valori differenti dai precedenti, mantiene tutte le caratteristiche che sono presentate per la Massima Verosimiglianza. Oltre a contrassegnare l'incremento dell'età media nell'aver figli, c'è la conferma che vi è la propensione per la scelta di un modello che permette la presenza di una doppia moda o di un hump. La dimostrazione di ciò è fornita dalle stime dei parametri α e β che per ogni annualità i segni dei

due parametri sono contrapposti.

Non vi è l'evidenza di hump a causa della scarsità delle rilevazioni, ma la rappresentazione della sua presenza si può notare particolarmente per gli anni 1999 e 2010 in cui la presenza di una gobba è accennato dall'attenuarsi della pendenza del lato destro della distribuzione.

L'errore di adattamento è : 0.000024

Figura 5.4: stima tramite minimi quadrati



Anche in questo caso l'adattamento migliore è fornito dalla stima di dei Minimi Quadrati.

5.2 Stima complessiva dei parametri

Dai precedenti modelli si devono stimare 20 parametri avendo non più di 28 osservazioni. In questo caso la necessità di un modello che possa ridurre la numerosità dei valori stimati è ancora più evidente dato che, la differenza di numerosità delle osservazioni e la numerosità dei parametri, è ancor più esigua rispetto all'esempio del Regno Unito.

In questo caso, vista l'esiguità delle osservazioni, scegliamo di adattare delle spline di primo grado con un solo nodo, per permettere i cambiamenti di pendenza. I risultati di adattamento mostrati nei grafici in Figura 5.5, mentre nella tabella sottostante sono riportate le stime per i parametri relativi alle basi delle splines ottenute tramite i Minimi Quadrati.

Stimando i valori delle spline attraverso la minimizzazione dei quadrati risulta che lo scarto tra i valori osservati e quelli stimati è di 0.025.

Andando a creare degli intervalli di confidenza basati sul TRV a livello 0.95 si ottiene : Figura 5.6, Figura 5.7, Figura 5.8 e Figura 5.9. Con questi grafici si nota come la variabilità dei valori è nulla. Soprattutto α e β sono parametri necessari per modellare le curve osservate, infatti è scartata l'ipotesi che vi sia lo zero tra i veri valori dei due parametri.

Figura 5.5: stima dei parametri con spline

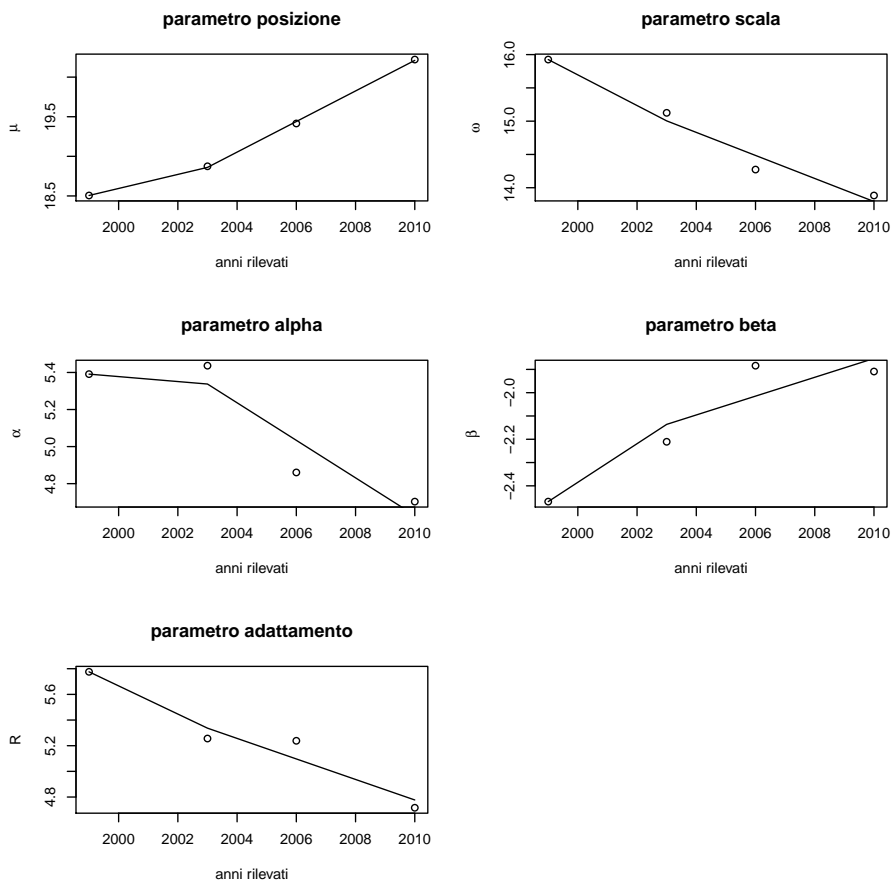


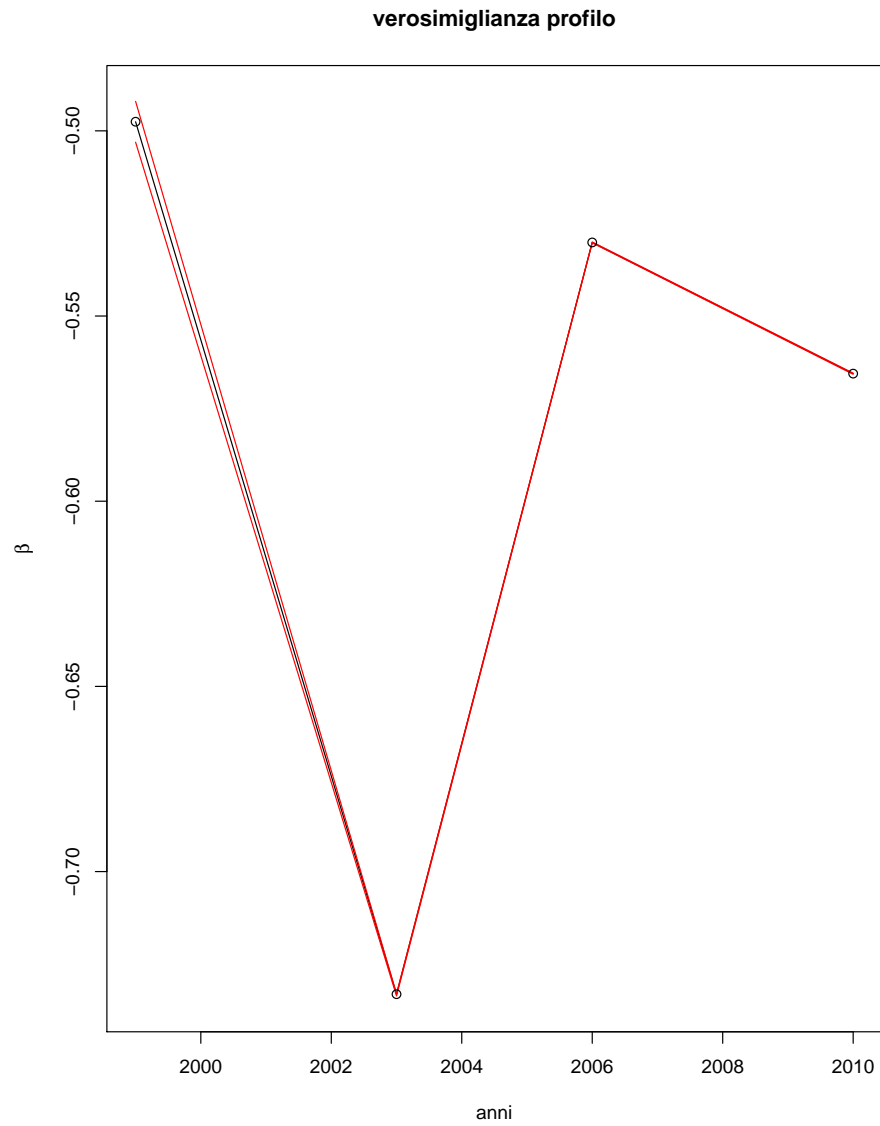
Figura 5.9: stima parametro β con intervalli di confidenza livello 0.95

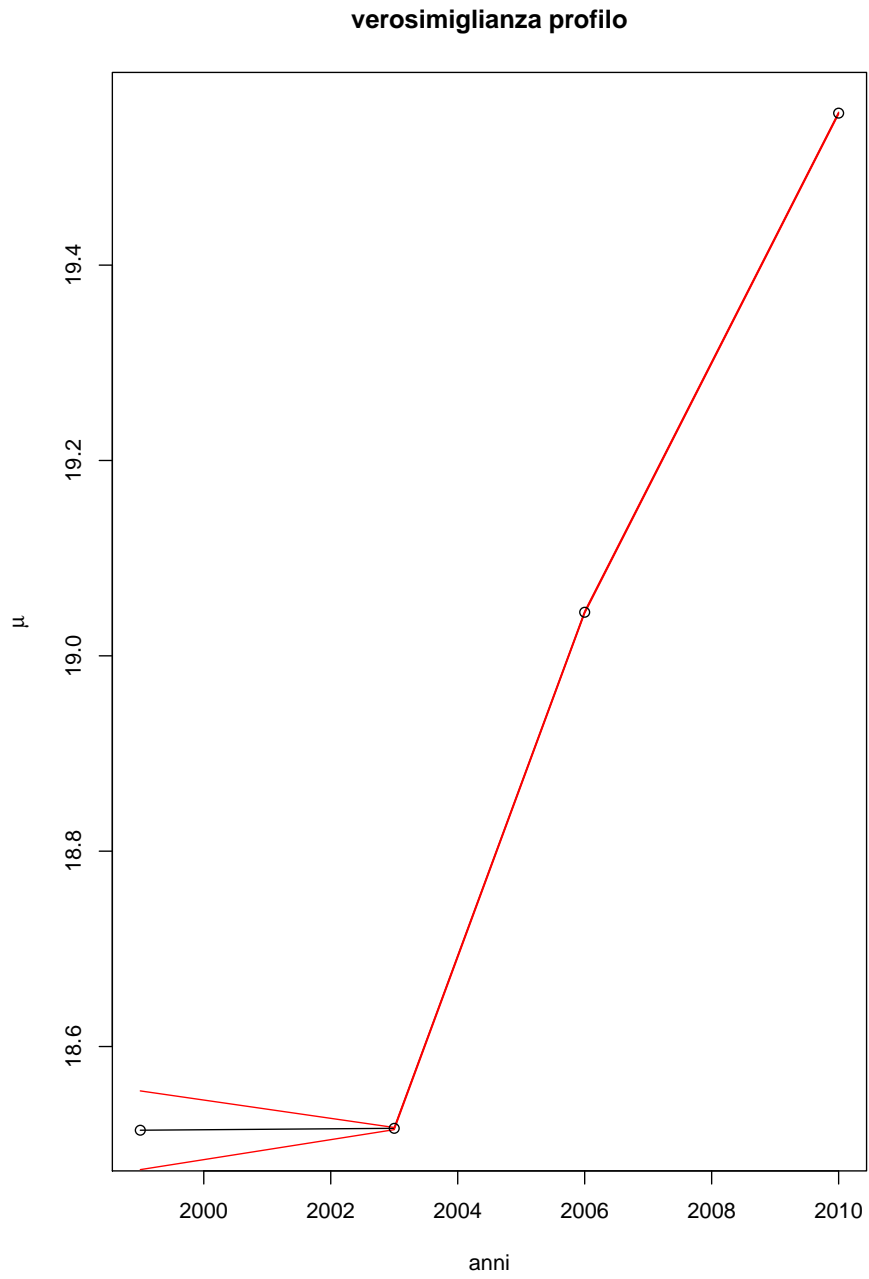
Figura 5.6: stima parametro μ con intervalli di confidenza livello 0.95

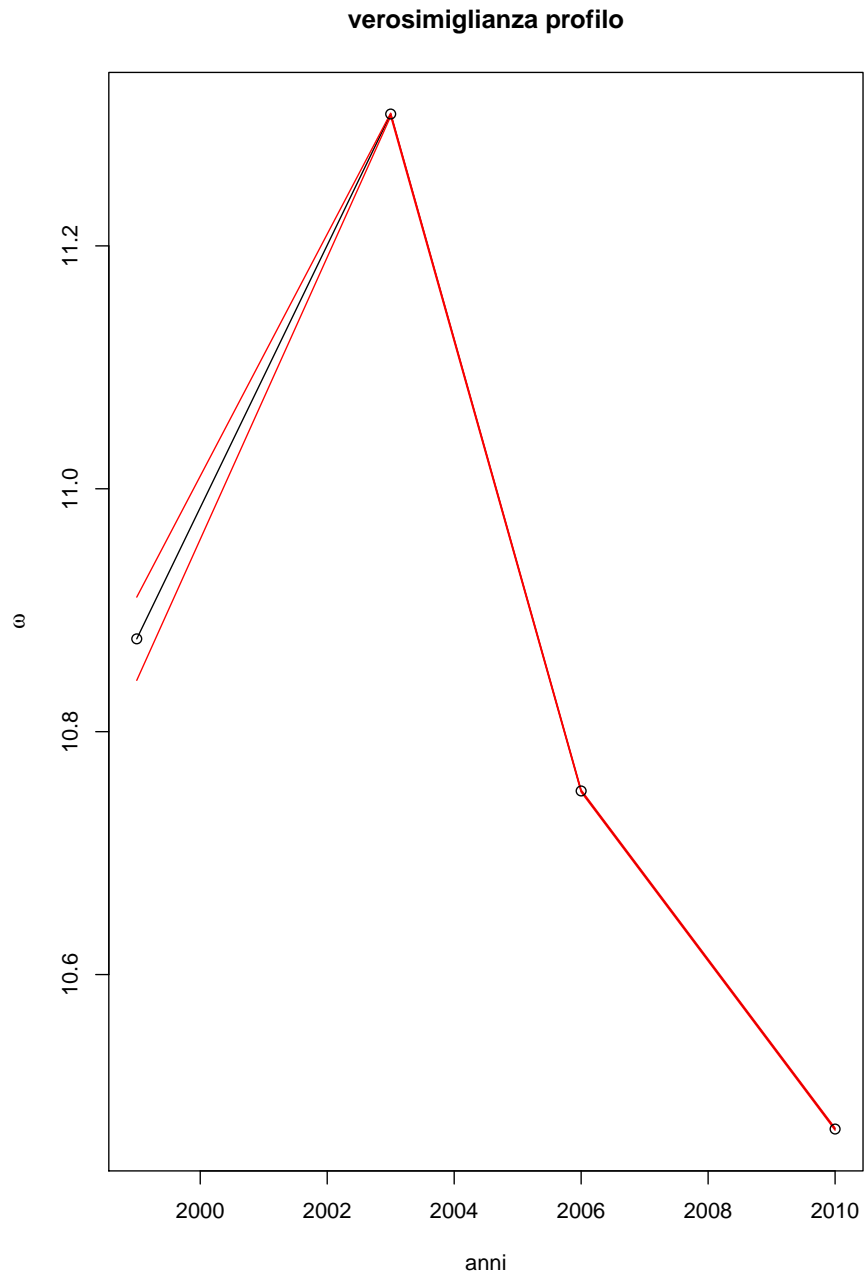
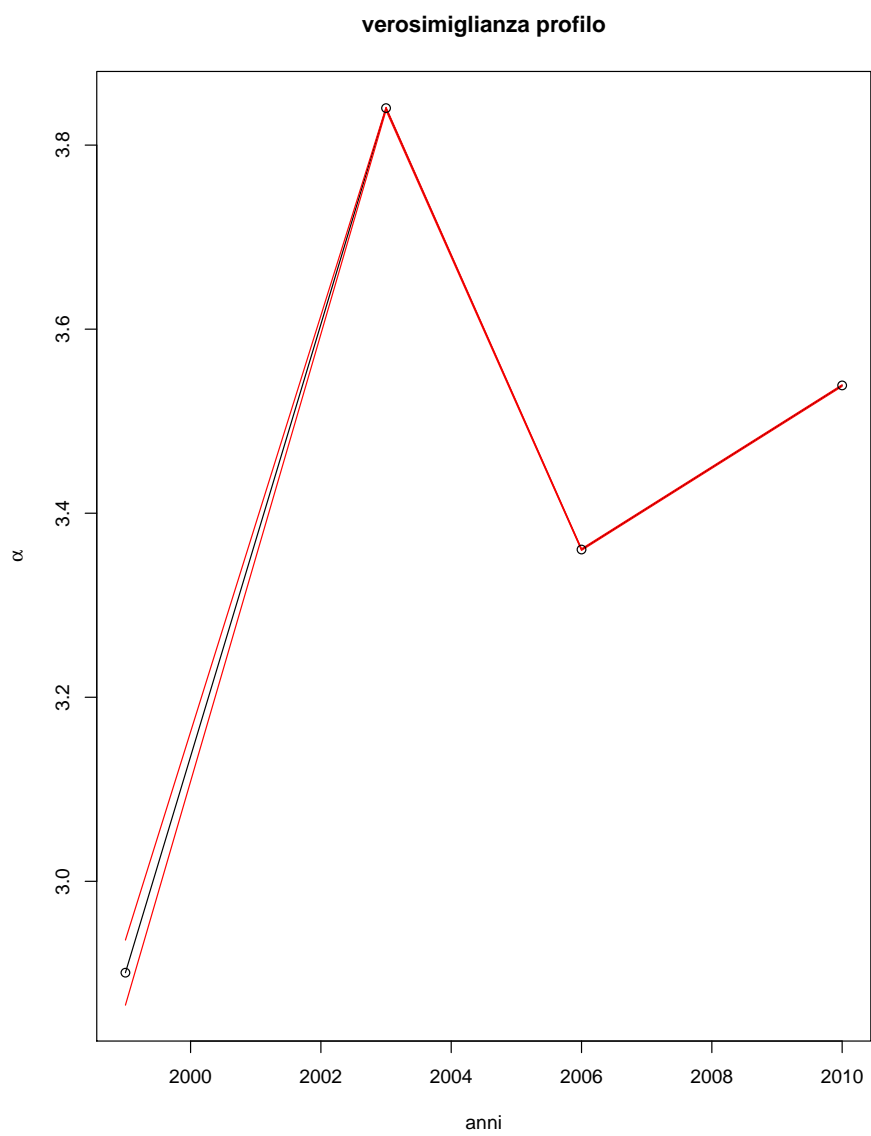
Figura 5.7: stima parametro ω con intervalli di confidenza livello 0.95

Figura 5.8: stima parametro α con intervalli di confidenza livello 0.95

Capitolo 6

Analisi della distribuzione Normale Asimmetrica Flessibile

6.1 Interpretazione della distribuzione NAF considerando un condizionamento

Guardando le distribuzioni dei tassi di fecondità è facile chiedersi da dove derivi una tale struttura. Si può quindi cercare di dare una giustificazione teorica della presenza di una seconda moda.

Come in Azzalini (2005) vi è la giustificazione per la struttura della Normale Asimmetrica con il condizionamento di una normale bivariata, si può ipotizzare che anche in questo caso vi sia una rappresentazione stocastica che aiuti ad interpretare il modello.

6.1.1 La relazione empirica tra età e mode

Nell'accostare la densità dell'età della popolazione femminile e i tassi di fecondità, si può percepire un collegamento tra la presenza di una moda per l'età e la posizione della seconda moda nella distribuzione dei tassi. Per verificare la relazione, si confrontano le medie delle distribuzioni delle età delle donne e la seconda moda identificata per le distribuzioni dei tassi.

Nel seguito sono riportati gli esempi di Irlanda, Spagna e Stati Uniti. Sono state scelte queste tre nazioni perché offrono un rilevante numero di confronti per analizzare un'eventuale relazione tra gli elementi.

Nei grafici Figura 6.1, Figura 6.2 e Figura 6.3 viene rappresentata la relazione che intercorre tra la moda delle età per le donne e la moda secondaria (con la

densità minore). In tutti e tre i grafici si vede che è possibile un'interpolazione dei punti attraverso una retta.

Figura 6.1: relazione mode per Irlanda

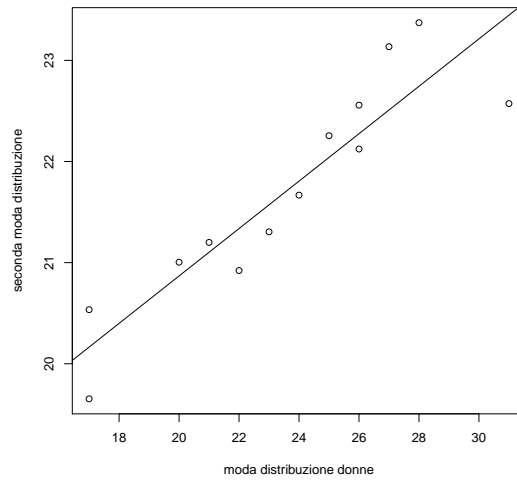
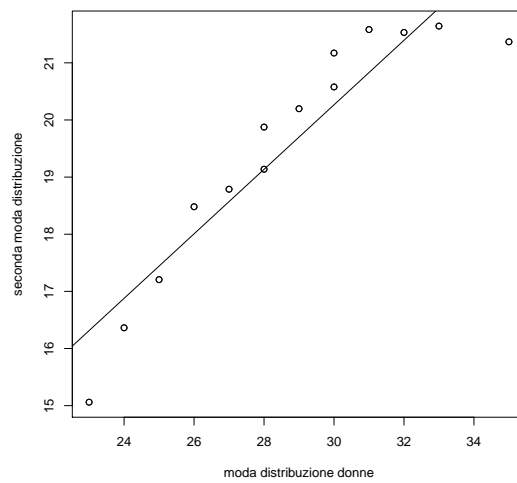
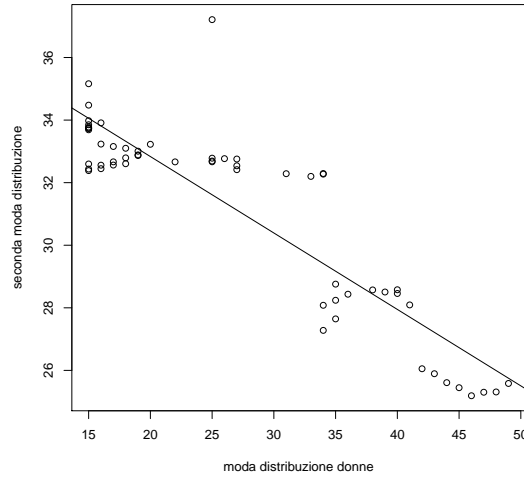


Figura 6.2: relazione mode per Spagna



6.1. INTERPRETAZIONE DELLA DISTRIBUZIONE NAF CONSIDERANDO UN CONDIZIONAMENTO47

Figura 6.3: relazione mode per Stati Uniti



Le interpolazioni lineari, ricavate tramite i minimi quadrati, dimostrano che la relazione supposta è presente e significativa.

Irlanda

Coefficienti

	Stima	Errore Standard	Valore t	Pr(> t)	
Intercetta	16.7783	0.9026	18.589	1.73e-08	***
Mediana	0.2107	0.0370	5.694	0.000297	***

—
Signif.: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

Spagna

Coefficienti

	Stima	Errore Standard	Valore t	Pr(> t)	
Intercetta	3.32901	1.82064	1.828	0.0924	.
Mediana	0.56453	0.06313	8.942	1.18e-06	***

—
Signif.: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

Stati Uniti					
Coefficienti					
	Stima	Errore Standard	Valore t	Pr(> t)	
Intercetta	37.72108	0.47852	78.83	< 2e - 16	***
Mediana	-0.24431	0.01647	-14.84	< 2e - 16	***
—					
Signif.: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Dai parametri stimati si vede come per ogni nazione si differenzia il comportamento della relazione. Ampliando i casi che dimostrano questa relazione vi è la possibilità di generalizzare il comportamento, fornendo un'indicazione sulla struttura della combinazione dei due fattori presi in considerazione.

6.1.2 rappresentazione stocastica

Rifacendosi a Arellano Valle, Branco e Genton (2006) e potendo rappresentare i tassi di fecondità (X) come $X =^d (V|U \in C)$ ci si riconduce alla struttura

$$f_x(x) = f_v(x) \frac{P(U \in C|V = x)}{P(w(u) \in C)}$$

con cui si può spiegare l'utilizzo di una Normale Asimmetrica Flessibile tramite un condizionamento attraverso la trasformazione di una distribuzione correlata con i dati. La trasformazione ricercata è identificabile nella funzione dispari di $\omega(x) = \alpha x + \beta x^3$.

In questo caso però non si comprende cosa rappresentano le quantità di U, V in quanto per avere una $x \sim \text{NAF}(x; \mu, \omega, \alpha, \beta)$, devono assumere distribuzioni normali come introdotto in Arellano Valle, Branco e Genton (2006).

Nonostante ciò è ipotizzabile che la doppia modalità della distribuzione dei tassi rispecchi la moda principale della distribuzione delle donne.

6.2 Interpretazione parametri

In questa sezione si studia una distribuzione Normale Asimmetrica Flessibile con valori, del parametro di posizione e del parametro di scala, fissati per semplicità a:

$$\mu = 0 \quad \omega = 1$$

6.2.1 Parametrizzazione classica

Precedentemente si sono utilizzati modelli che coinvolgono la Normale Asimmetrica Flessibile, la cui caratteristica principale è la possibile presenza di una seconda moda identificata da un cambiamento di concavità o una "gobba" più o meno evidente.

In questo paragrafo, si cerca di dare una valutazione della forma della densità al variare dei parametri in modo da avere un'interpretazione diretta e semplificata dell'effetto che le modifiche di questi comportano sulla forma della distribuzione.

La densità della variabile casuale Normale Asimmetrica Flessibile (NAF) è formulata come

$$2\phi(z)\Phi(\alpha z + \beta z^3)$$

dove

- ϕ identifica la funzione di distribuzione di una normale standard,
- Φ identifica la sua funzione di ripartizione
- z identifica la una variabile standardizzata ovvero: $z = \frac{x-\mu}{\omega}$

In questa formulazione, oltre ai parametri noti di posizione (μ) e di scala (ω), troviamo due parametri di forma α e β che coinvolgono rispettivamente z e z^3 .

Nella formulazione base della distribuzione Normale Asimmetrica proposta da Azzalini (1985) si ha la presenza solamente del parametro α che, in questo caso, è facilmente identificabile come il parametro di forma che regola sia l'asimmetria che la curtosi della distribuzione. Nella distribuzione introdotta da Ma e Genton non si ha una diretta interpretazione dei due parametri dato che, la distribuzione è determinata dalle diverse combinazioni dei due valori.

Per cogliere una possibile influenza effettiva dei parametri sono state rilevate quattro caratteristiche :

- momento primo
- asimmetria ottenuta tramite momento terzo della distribuzione
- distanza euclidea dei due punti identificati come moda Figura 6.4
- presenza del tipo di mode che si distingue con le strutture di

- singola moda, come in Figura 6.5
- due mode ben marcate (che per semplicità chiameremo tipo 1), come in Figura 6.6
- seconda moda presente come un cambiamento di concavità (che per semplicità chiameremo tipo 2), come in Figura 6.7

Nell'identificazione per la quarta caratteristica rilevabile vengono considerati dei pesi differenti per ogni struttura identificata. Dopo un'analisi e, attraverso i valori che annullano la derivata prima, di questi valori vengono identificati, come massimi relativi presenti nella funzione, se restituiscono un valore positivo quando sono inseriti nella funzione della derivata seconda della distribuzione. Ai parametri in cui viene rilevato il massimo, viene assegnato un punteggio base (es. 3). Così, per ogni funzione determinata da una combinazione di α e β , si avrà il peso 3 se vi è soltanto un massimo, mentre si otterrà 6 se ve ne sono due e sarà identificato come seconda moda di tipo 1.

Ma e Genton (2004) dimostrano che, per la NAF, vi sono al più due massimi, per cui non vi è la necessità di analizzare ulteriormente questo tipo di distribuzione.

Successivamente si considereranno le strutture in cui è stato rilevato un solo massimo. Per cercare la presenza di hump si isolano i punti in cui vi è un cambiamento di concavità. Attraverso la derivata seconda, escludendo i valori esterni dovuti al naturale cambio di concavità della distribuzione, si ottengono i valori che la annullano. Se viene rilevato un hump viene assegnato e sommato un peso differente dal precedente alla combinazione di α e β . Identificando come 1.5 il peso minore si otterrà quindi che, una funzione che presenta un hump, avrà come peso 4.5 e sarà identificata come seconda moda di tipo 2.

Alle funzioni che risulteranno aventi una moda soltanto sarà assegnato il singolo peso di 3 e quindi sarà identificata come una funzione a moda singola.

Con questo algoritmo, tenendo in considerazione anche i punti rilevati per l'identificazione delle mode e degli hump, si è potuto esaminare anche la distanza euclidea citata nelle caratteristiche rilevate. Questo è stato possibile grazie alla differenza del valore della distribuzione per i due punti rilevati e alla loro distanza.

Figura 6.4: distanza punti di moda

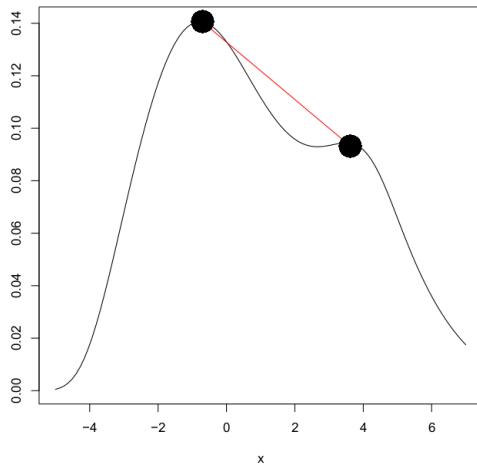


Figura 6.5: densità con un punto di moda

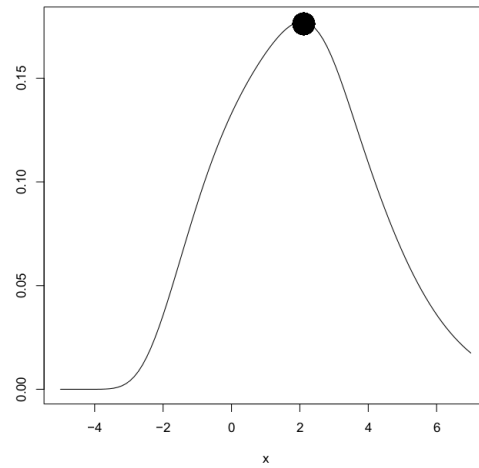


Figura 6.6: densità con due punti di moda tipo 1

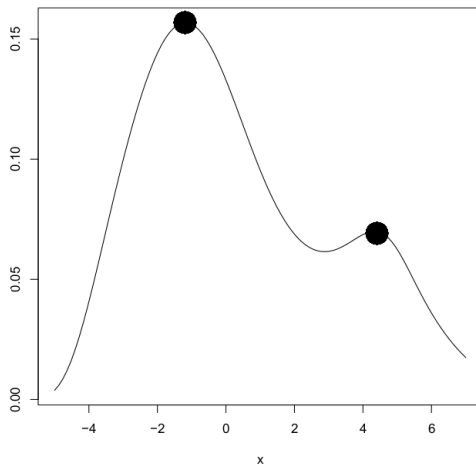
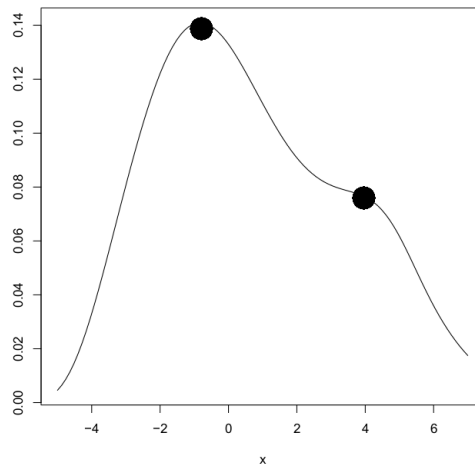


Figura 6.7: densità con due punti di moda tipo 2



La formulazione classica per la Normale Asimmetrica Flessibile, con i parametri originali, assume la struttura

$$\text{NAF}(\mu, \omega, \alpha, \beta) = \frac{2}{\omega} \phi\left(\frac{x-\mu}{\omega}\right) \Phi\left(\alpha\left(\frac{x-\mu}{\omega}\right) + \beta\left(\frac{x-\mu}{\omega}\right)^3\right)$$

abbiamo che le misure precedentemente elencate sono visibili nei grafici seguenti.

Figura 6.8: distanza tra i punti di moda

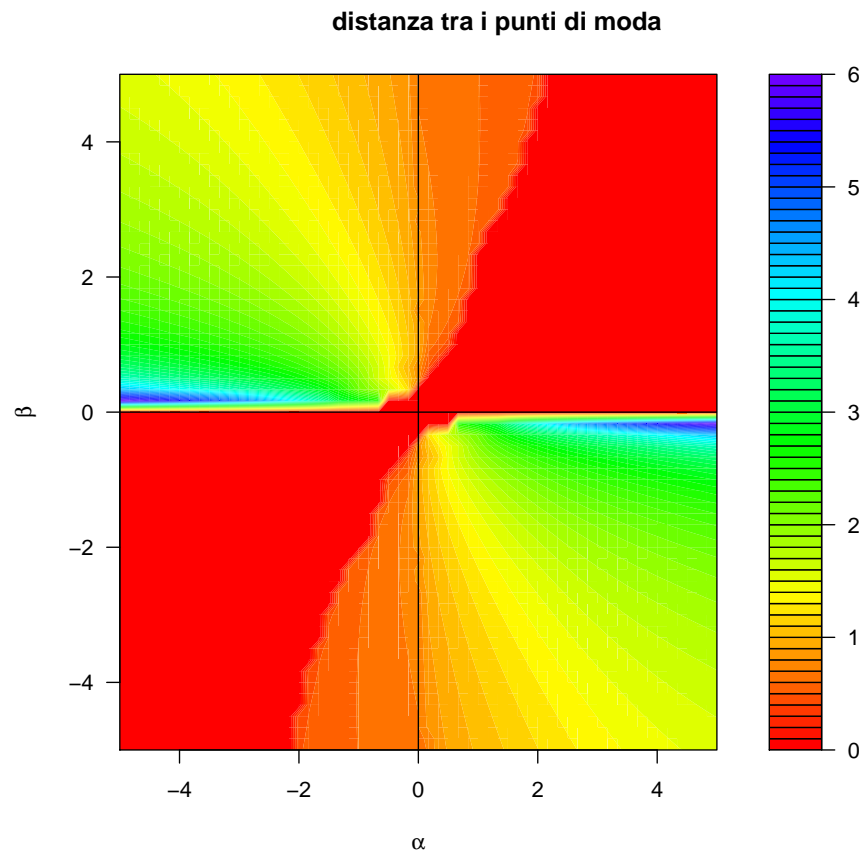
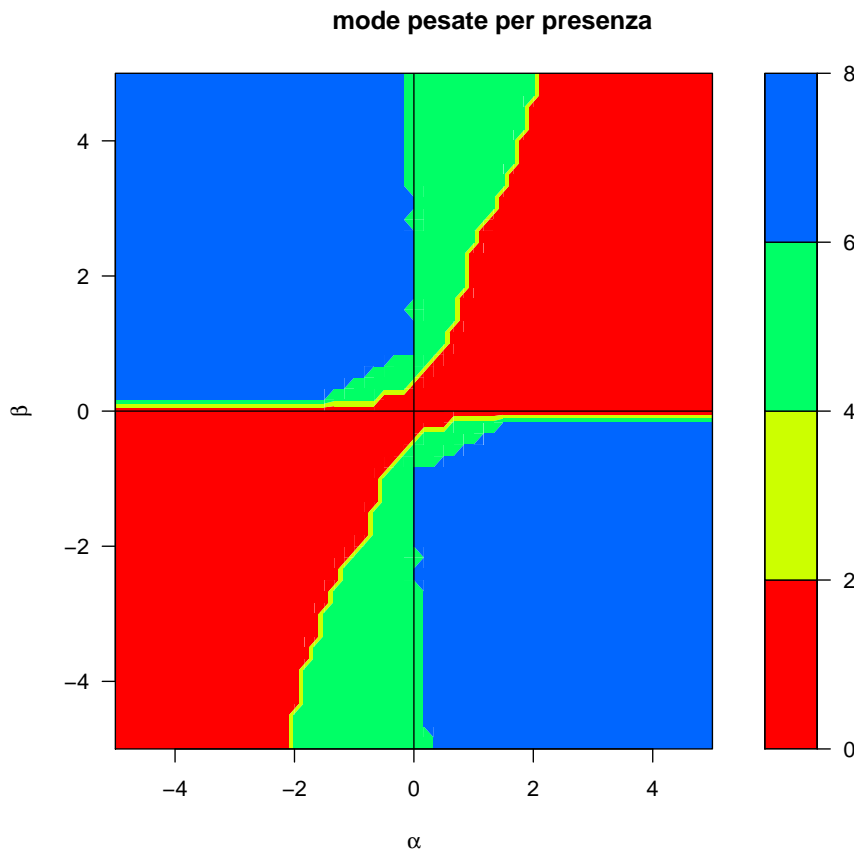


Figura 6.9: densità rilevazione mode attraverso un sistema di pesi



Dai grafici Figura 6.8 e Figura 6.9, si inizia a notare l'influenza dei parametri sulla distribuzione, infatti si vede che le mode sono fortemente caratterizzate dalla contrapposizione dei segni dei due parametri di forma, inoltre si nota che il secondo tipo di distribuzione bimodale (Figura 6.7) risulta dal rapporto elevato tra β e α anche se dello stesso segno. L'area contrassegnata di rosso definisce le combinazioni di α e β che riconducono la distribuzione alla struttura unimodale che caratterizza la Normale Asimmetrica

Dai grafici in Figura 6.10 e Figura 6.10 si vede come non vi sono determinate combinazioni che denotino lo scostamento della media e l'asimmetria della distribuzione. Nonostante ciò, la concordanza dei segni dei due parametri interessati, porta ad un effettivo scostamento dal valore che si otterrebbe con una distribuzione Normale, questo si denota nel grafico in Figura 6.10 dove i valori sono differenti da zero. Questo porta a dedurre che la concordanza dei

segni positivi di α e β , tende ad alzare l'età media delle madri, e conseguentemente, la concordanza dei segni negativi ne porta un abbassamento. Nell'area contraddistinta con il colore verde (identificabile tramite la bisettrice 2° e 4° quadrante) vi sono combinazioni di valori che permettono di strutturare le caratteristiche di centralità in zero e asimmetria nulla. Queste due peculiarità derivano dalla compensazione delle due mode quando i valori di α e β non sono nulli, e dalla riconducibilità della distribuzione alla Normale quando sono pari a zero.

Figura 6.10: lo scostamento della media della distribuzione dal parametro di posizione

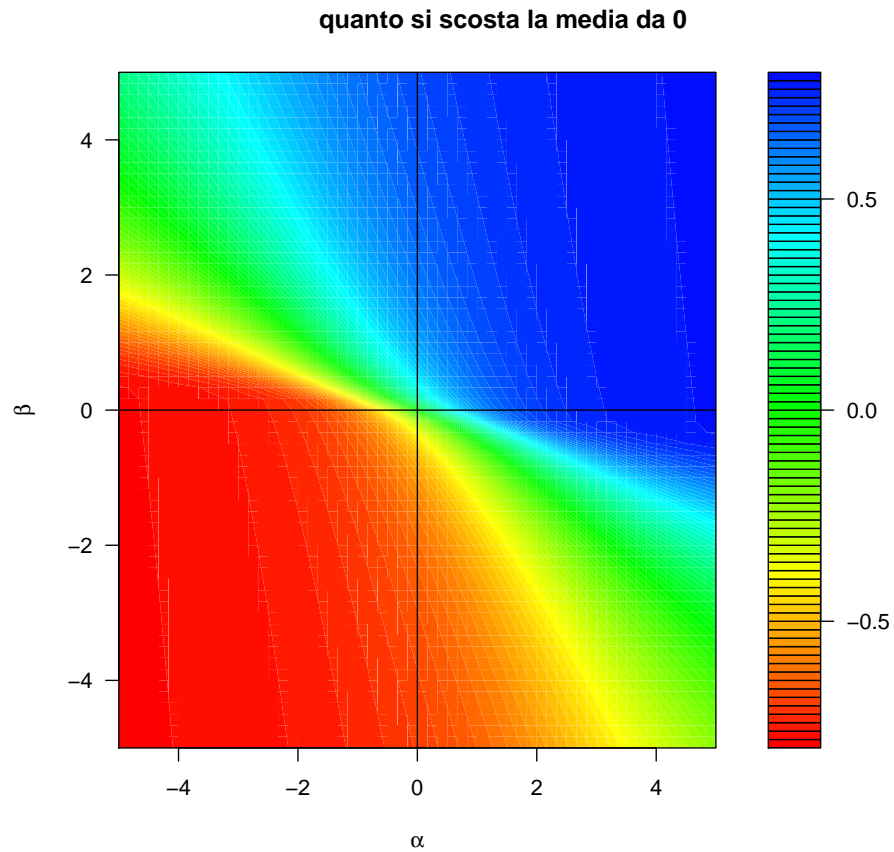
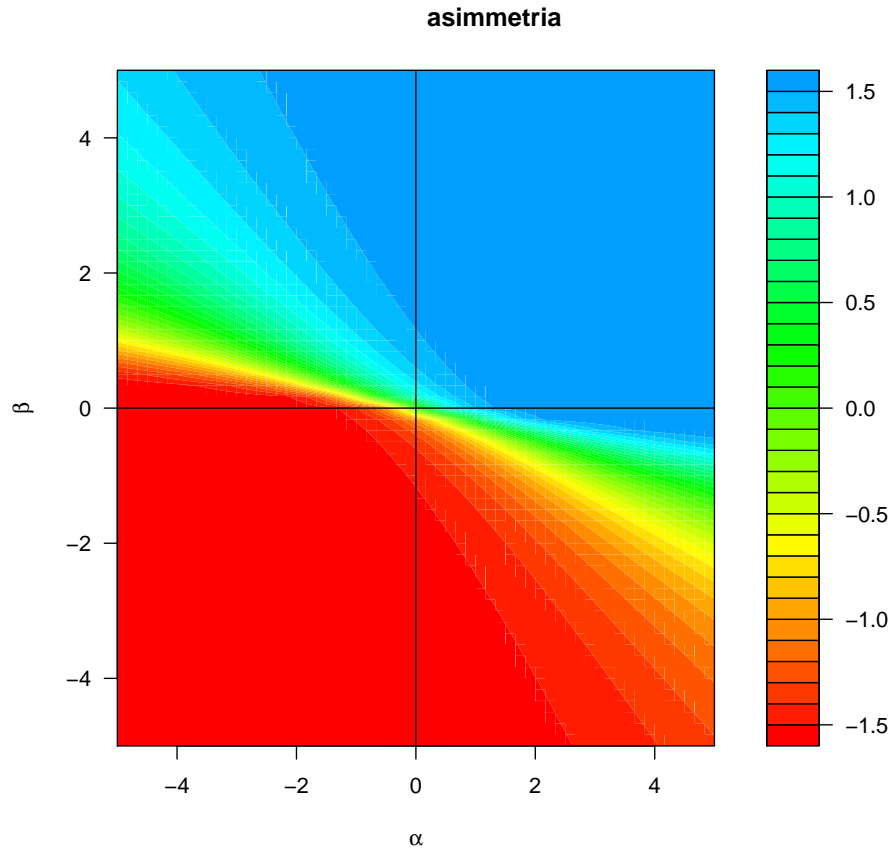


Figura 6.11: l'asimmetria della distribuzione



6.2.2 Proprietà di simmetria

Dai primi due grafici, Figura 6.8 e Figura 6.9, si osserva una simmetria della distribuzione tra le due coppie di parametri di forma, α e β , che comportano segni opposti. Nella ricerca della motivazione si isolano i due valori che vanno ad influire sulla densità all'interno della funzione di ripartizione. Si nota che, al variare di segno di questi due elementi (ad esempio da $\alpha = 3$ e $\beta = -4$ ad $\alpha = -3$ e $\beta = 4$) si ottiene un mutamento della funzione di ripartizione del tipo $\Phi(-k)$ in $\Phi(k)$ in cui k è $(\alpha z + \beta z^3)$, α e β sono di segno opposto e Φ è la funzione di ripartizione di una normale.

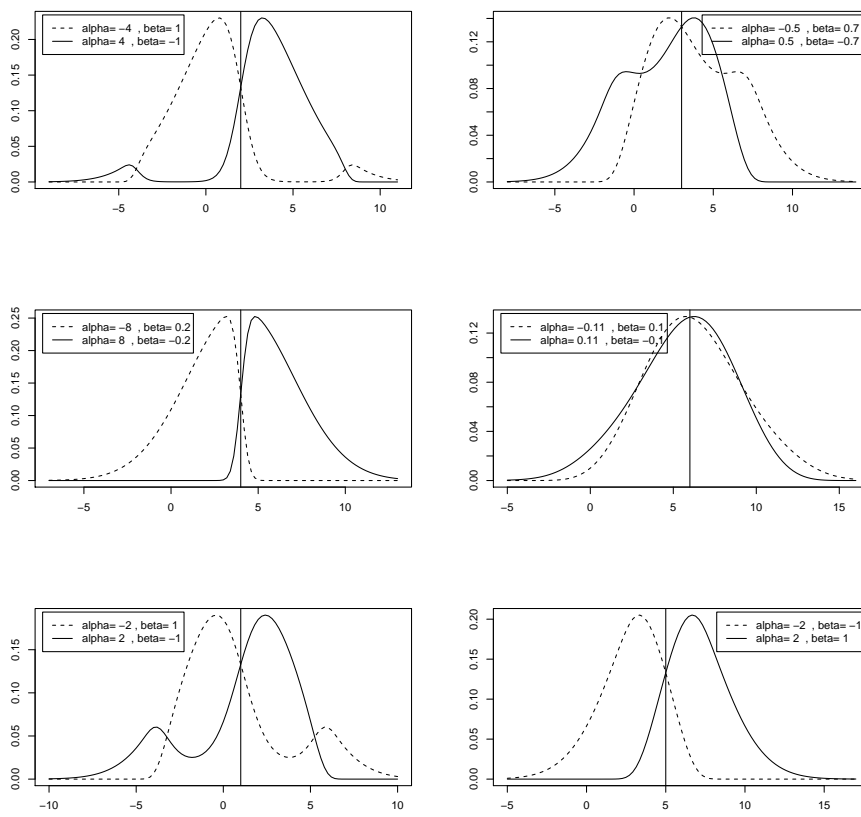
Considerando ciò, ci si riconduce alla proprietà di simmetria della funzione di ripartizione e di distribuzione di una normale:

$$\Phi(d) = 1 - \Phi(-d)$$

per cui avvalendosi del fatto che si ottiene la stessa percentuale a destra e a sinistra dato un quantile q e $-q$, si può dire che per valori di segno opposto per α e β si ottengono due distribuzioni simmetriche, e quindi che hanno lo stesso comportamento anche se contrapposto.

Questa caratteristica fornisce quindi alla distribuzione la proprietà che se $Z \sim \text{NAF}(\alpha, \beta)$, allora $-Z \sim \text{NAF}(-\alpha, -\beta)$ (Figura 6.12). La caratteristica di simmetria riscontrata non vale per le variazioni di segno che coinvolgono solamente uno dei due parametri, infatti, da come si nota in Figura 6.9 nei riquadri inferiori del grafico ($\text{NAF}(\alpha, \beta)$ a sinistra e $\text{NAF}(-\alpha, \beta)$ a destra), vi è un differente comportamento della distribuzione.

Figura 6.12: proprietà simmetria distribuzioni



6.2.3 Seconda parametrizzazione

Volendo dare una formulazione alternativa alla densità per avere una più diretta interpretazione dei parametri, si riscrive la densità come

$$\text{NAF}(\mu, \omega, \alpha, b) = \frac{2}{\omega} \phi\left(\frac{x-\mu}{\omega}\right) \Phi\left[\alpha\left(\frac{x-\mu}{\omega} + b\left(\frac{x-\mu}{\omega}\right)^3\right)\right]$$

dove $a = \alpha$ e $b = \frac{\beta}{\alpha}$.

In questo modo si può avere una interpretazione dei parametri, ovvero:

- se $a = 0$ si ha una distribuzione normale
- se $b = 0$ si ha una distribuzione Skew Normal
- se $b < 0$ si ha una parametrizzazione con segni opposti
- se $b > 0$ si ha una parametrizzazione con segni concordanti

per cui, teoricamente, dove prima avevamo una rappresentazione grafica dei parametri del tipo Figura 6.13, ora, con la nuova parametrizzazione, si ottiene che i quadranti per $\alpha < 0$ si invertono. Figura 6.14

Figura 6.13: parametri standard

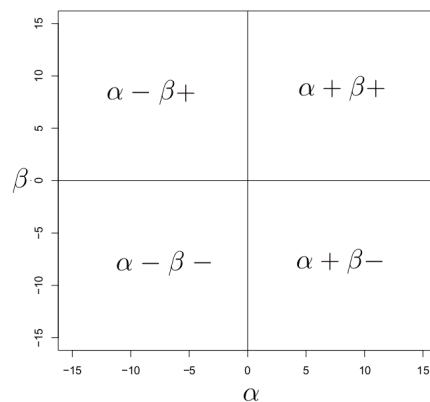
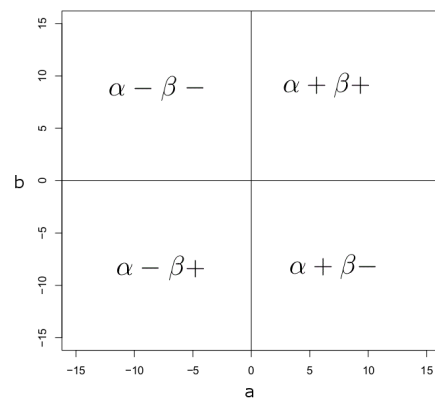


Figura 6.14: parametri combinati



Da questo risultato si può presupporre che l'effetto di $b < 0$ è identificativo di una doppia moda. Infatti quello che risulta dalle rilevazioni grafiche è riportato nei grafici di Figura 6.15 e Figura 6.16

Figura 6.15: distanza tra i punti di moda nella riparametrazione

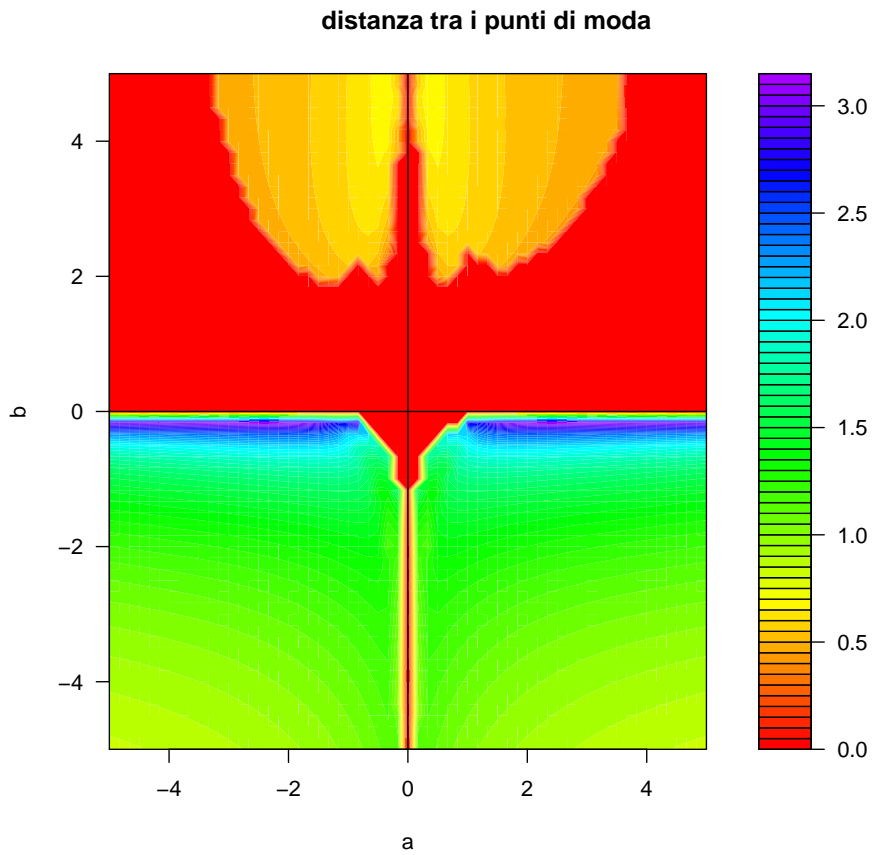
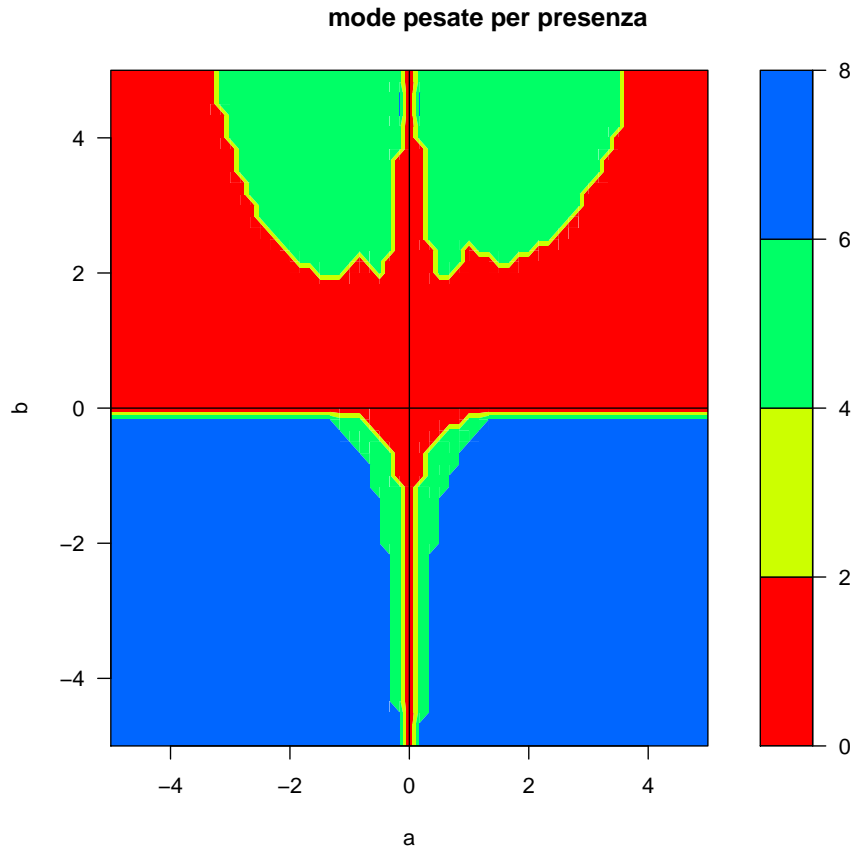


Figura 6.16: densità rilevazione mode attraverso un sistema di pesi nella riparametrazione

Nei due grafici di Figura 6.15 e Figura 6.16 si vede come la teoria, precedentemente proposta, viene rispettata, infatti si nota che ad un valore di $b < 0$, la distribuzione si caratterizza per avere una doppia modalità ben definita (area in cui vi sono i valori massimi in Figura 6.16). Mentre per un suo elevato valore si ottiene la moda caratterizzata solamente dal cambio di concavità e senza inversione della decrescenza. Anche questa peculiarità si può riscontrare con ciò che è stato detto nel capitolo precedente, infatti una il rapporto di β su α , ovvero b , molto grande genera la presenza di una seconda moda.

Questo comportamento può essere riassunto dal verificarsi della condizione che $b < f(a)$, dove $f(a)$ è una funzione del parametro a rappresentata in Figura 6.9 (area evidenziata quando $b > 0$) per $a \neq 0$. Ovvero l'hump si verifica quando il valore del rapporto è all'interno di una parabola.

Figura 6.17: lo scostamento della media della distribuzione dal parametro di posizione nella riparametrazione

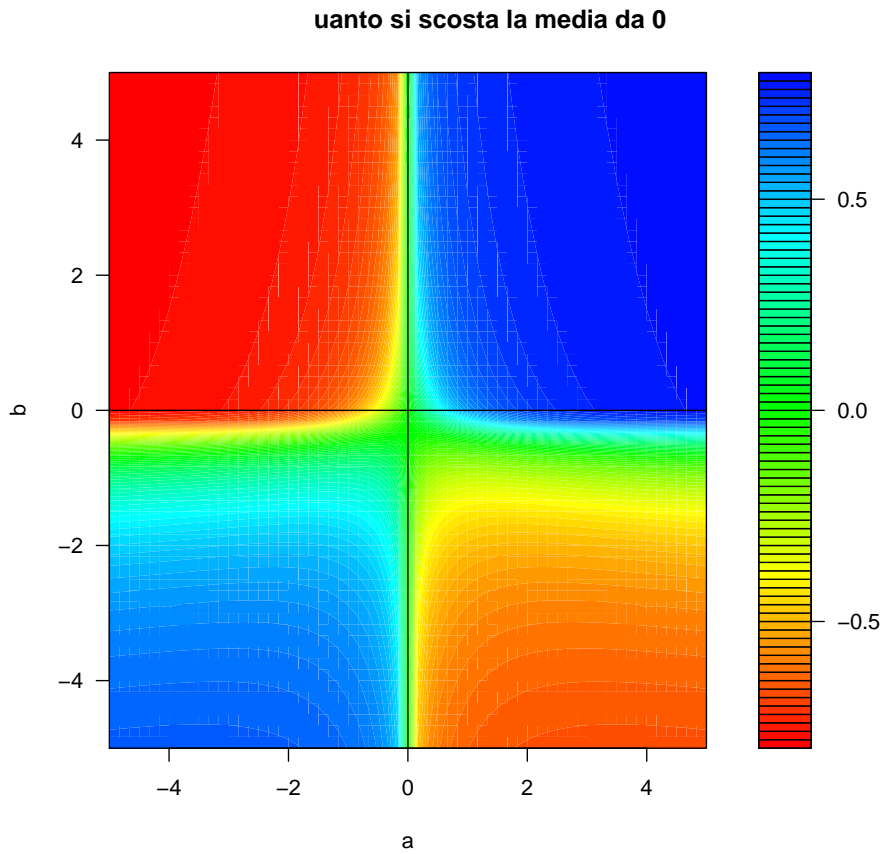
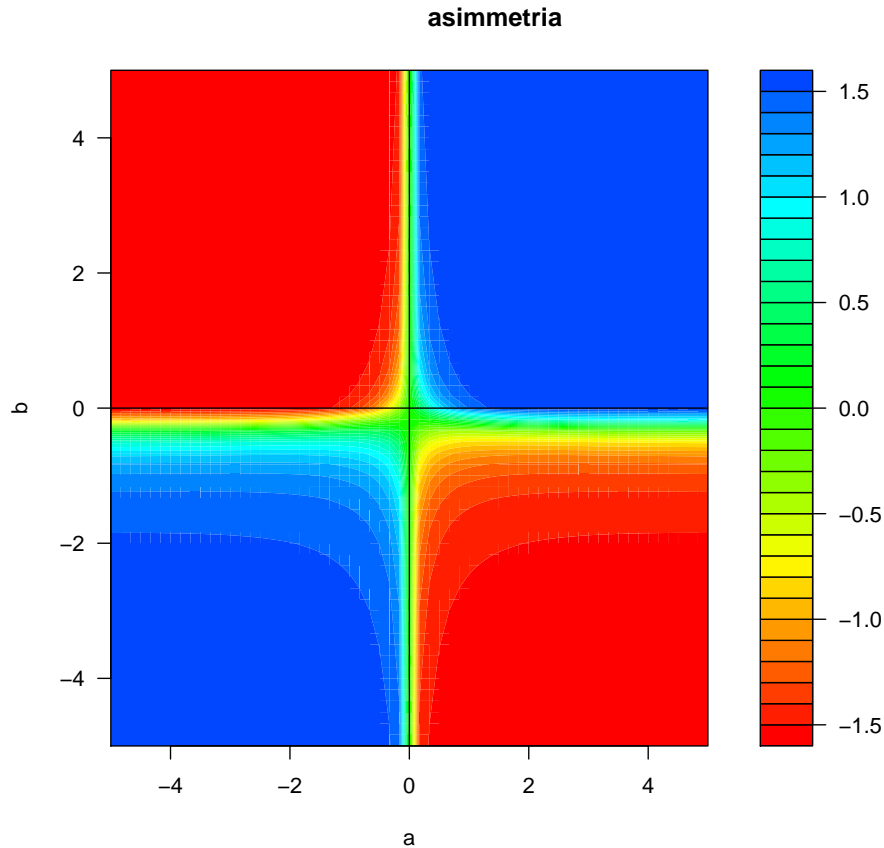


Figura 6.18: l'asimmetria della distribuzione nella riparametrizzazione



In Figura 6.17 e Figura 6.18 si nota che, il tipo di asimmetria e, lo scostamento della moda, è guidato dalla concordanza o discordanza dei segni. Per questa parametrizzazione troviamo che, a segni contrastanti di a e b , si genera un'asimmetria negativa che concorda con uno spostamento su valori negativi della moda, mentre con a e b positivi si ha sia un'asimmetria positiva che una moda positiva.

Come preannunciato si ha che, per $a = 0$, si ottiene il comportamento di una normale con parametri $\mu = 0$ e $\omega = 1$, infatti si nota che si ha un accentramento della distribuzione attorno al parametro di posizione e l'asimmetria è nulla. Inoltre si osserva che il comportamento assunto quando si ha $a = 0$, si ottiene con b quando questo è leggermente inferiore a 0 ovvero $b = \left(\frac{x-\mu}{\omega}\right)^{-2}$

Capitolo 7

Conclusioni

Sfruttando la possibilità di creare una seconda moda, si possono ampliare i casi in cui è possibile riassumere il comportamento dei dati tramite una distribuzione più flessibile senza avere la necessità di introdurre misture di modelli che appesantiscono la stima con troppi parametri. Per questo si può evitare l'utilizzo di combinazioni di due o più funzioni che aumenterebbero esponenzialmente il numero dei parametri da analizzare. Grazie a questo, nell'elaborato si è visto come la Normale Asimmetrica Flessibile si possa adattare bene ai tassi di fecondità sia nei casi in cui vi è la disponibilità di dati esaustivi che nei casi in cui non vi siano.

Si è potuto constatare, inoltre, come la Normale Asimmetrica Flessibile sia influenzata dalle diverse combinazioni dei parametri di forma introdotti e come, attraverso una loro trasformazione, vi è una migliore interpretazione di ciò che essi rappresentano, giustificando così come una variazione dei tassi di fecondità possa risultare più facilmente interpretabile da una modifica introdotta ai parametri del modello. Oltre a ciò è stato introdotto come, grazie ad una variabilità congiunta dei dati e attraverso un condizionamento, ci si può ricondurre dalla formulazione di base della normale standardizzata ad una condizione di distribuzione bimodale definita dalla Normale Asimmetrica Flessibile.

Appendice A

Grafici

Figura A.1: distanza in altezza per i due punti di moda nella parametrizzazione classica

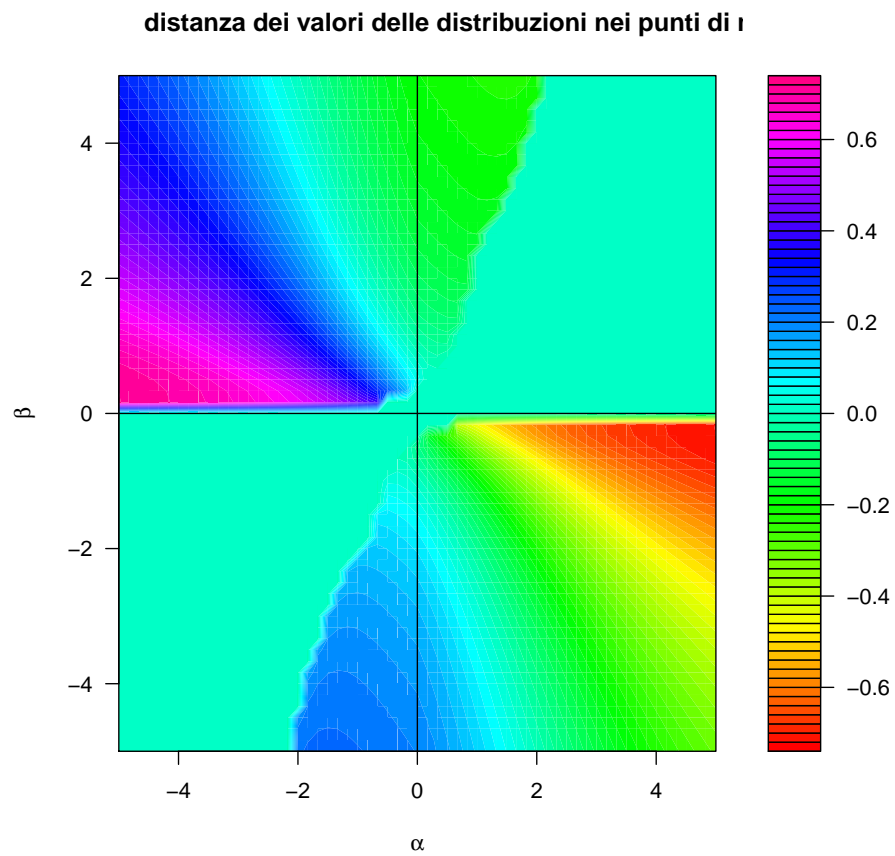


Figura A.2: distanza in larghezza per i due punti di moda nella parametrizzazione classica

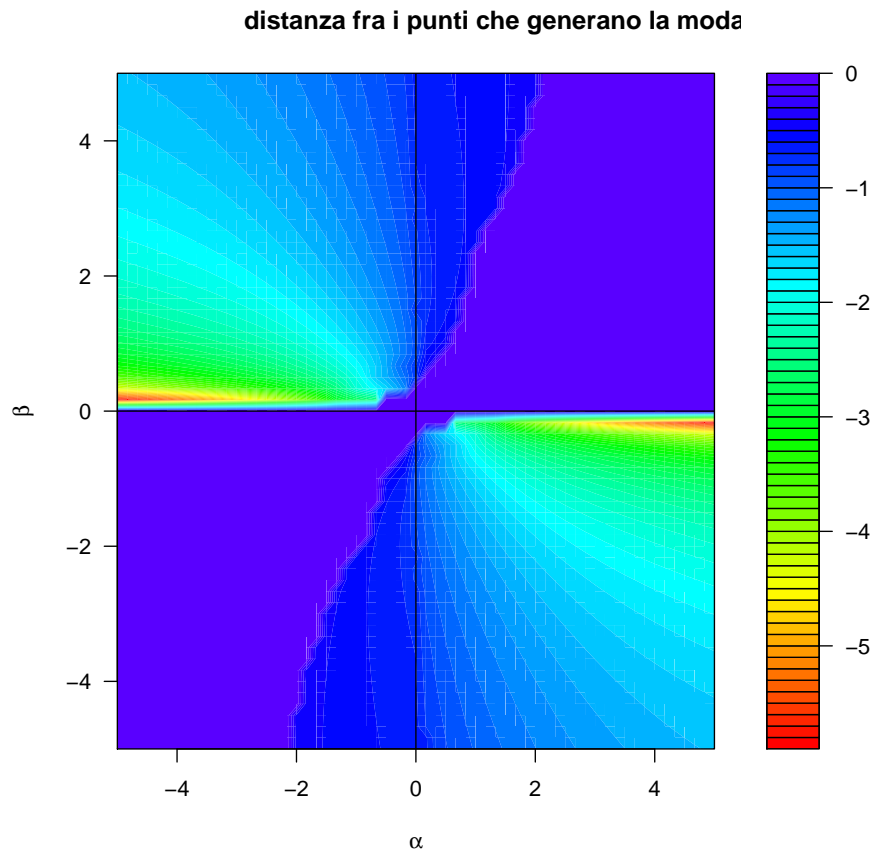


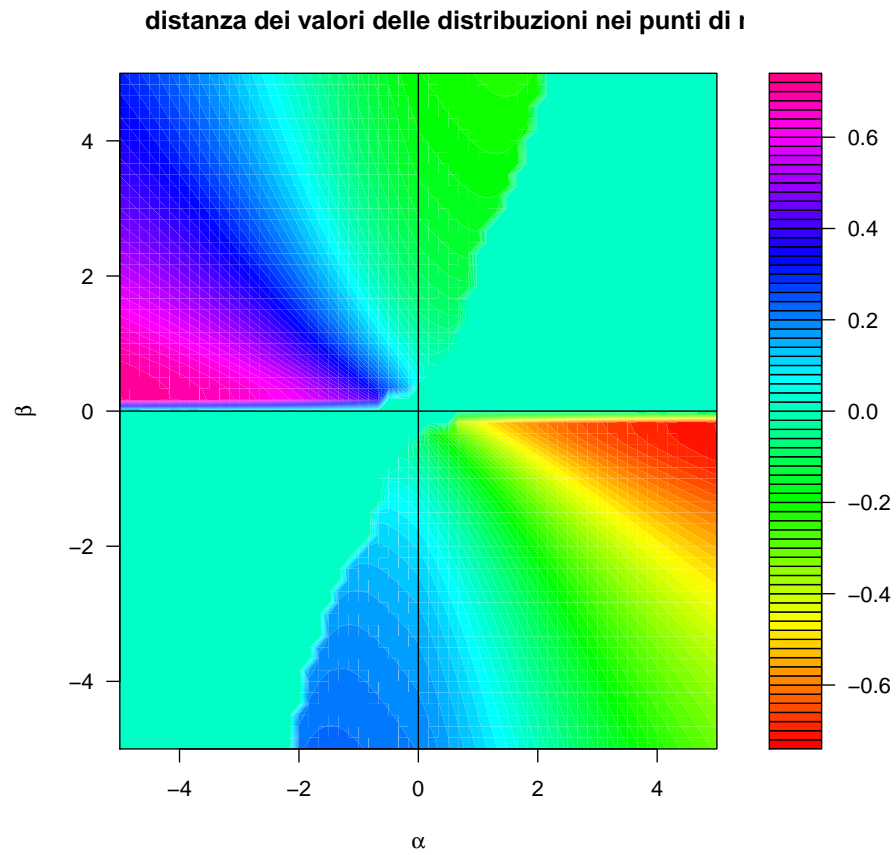
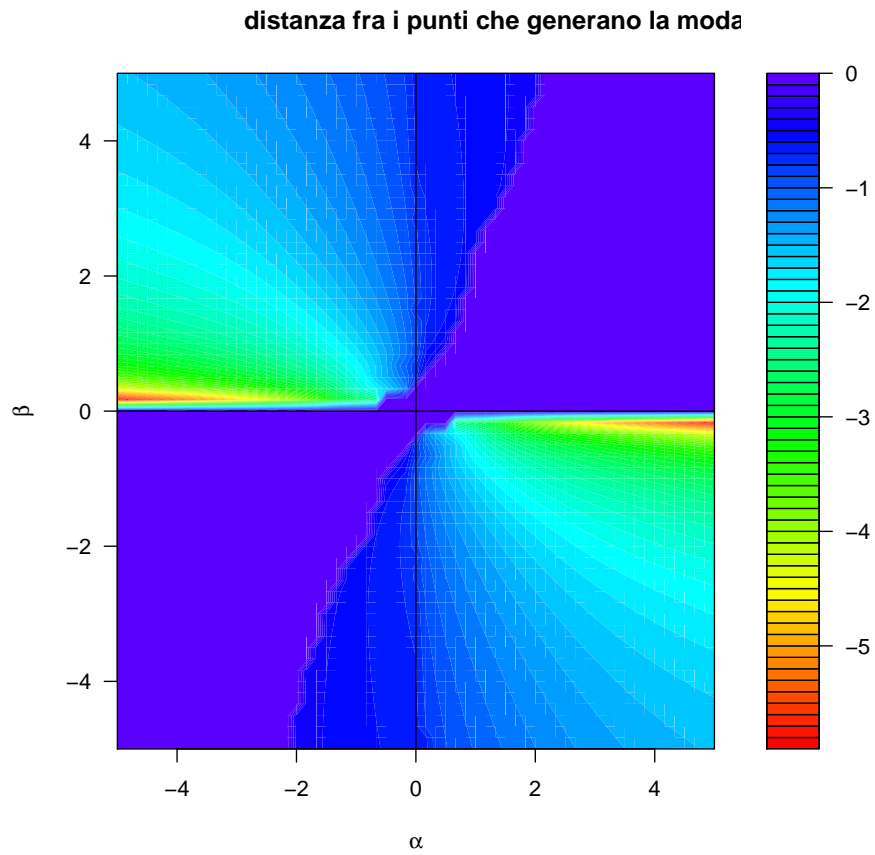
Figura A.3: distanza in larghezza per i due punti di moda nella parametrizzazione modificata

Figura A.4: distanza in larghezza per i due punti di moda nella parametrizzazione modificata



Appendice B

Codici programma R

```
myf_dens<-function(param,dat=15:49) {
z<-(dat-param[1])/(param[2])
2/(param[2])*(dnorm(z))*
  pnorm(param[3]*z+param[4]*z^3)#*param[5]
}

myf_dens.2<-function(param,dat=15:49) {
z<-(dat-param[1])/(param[2])
2/(param[2])*(dnorm(z))*
  pnorm(param[3]*z+param[4]*z^3)*param[5]
}

#####distribuzione
log_v<-function(param,dat_don,dat_tass,dat=15:49){
  phi<-myf_dens(param,dat)
-dmultinom(round(dat_don*dat_tass,0),sum(round(dat_don*dat_tass,0)),phi,log=T)
}

err<-err2<-rep(NA,length(annipresenti))
par(mfrow=c(3,2))
for( i in 1:length(annipresenti)){
oo<- x_tass[,nazione,paste(annipresenti)[i]]
parr<-c(20,5,-0.5,1)
parr2<-c(20,5,-0.5,1,1)
  if(i!=1){parr<-mle3$par[1:4];parr2<-mle4$par}
messag="false convergence (8)"
messag2="false convergence (8)"

while (sum(messag==mex)!=0){
messag="false convergence (8)"
```



```

mle3=nlminb(parr,function(x)(log_v(x,x_donne[,nazione,paste(annipresenti)
[i]],x_tass[,nazione,paste(annipresenti)[i]])), lower=c(15,1e-10,-10,-20),
upper=c(59,60,60,60),control = list(step.max=0.5,iter.max=1500,eval.max=600))
messag<-mle3$message
parr<-mle3$par+runif(4,-3,3)
if (parr[2]<0)parr[2]=0
err[i]<-sum(( hist(rep(15:49,round(x_donne[,nazione,paste(annipresenti)[i]]
*x_tass[,nazione,paste(annipresenti)[i]],0)),breaks=14:49,plot=F)$density-
myf_dens(mle3$par))^2)
}
plot(hist(rep(15:49,round(x_donne[,nazione,paste(annipresenti)[i]]*x_tass
[,nazione,paste(annipresenti)[i]],0)),breaks=14:49,plot=F)$density~c(15:49)
,main=(annipresenti[i]),xlab="età",ylab="probabilità di avere un figlio")
lines(myf_dens(mle3$par)~c(15:49))#mle3$par

while ( sum(messag2==mex)!=0){
messag2="false convergence (8)"

mle4=nlminb(parr2,function(x)sum((oo-myf_dens.2(x))^2), lower=c(15,1e-10,-10,-20
,1e-10),upper=c(59,60,60,60,60))

messag2<-mle4$message
parr2<-mle4$par+runif(5,-3,3)
if (parr2[2]<0)parr2[2]=1
if (parr2[5]<0)parr2[5]=1
err2[i]<-sum((oo*x_donne[,nazione,paste(annipresenti)[i]]-myf_dens(mle4$par)
*x_donne[,nazione,paste(annipresenti)[i]])^2)
}
hist(rep(15:49,round(x_donne[,nazione,paste(annipresenti)[i]]*x_tass[,nazione
,paste(annipresenti)[i]],0)),breaks=14:49,freq=T)

plot(oo~c(15:49),main=(annipresenti[i]),xlab="età",ylab="tasso di fertilità")
lines(myf_dens.2(mle4$par)~c(15:49))#mle4$par

readline(prompt = "Pause. Press <Enter> to continue...")

if (i==1)parX<-mle3$par
if (i!=1)parX<-rbind(parX,mle3$par)
if (i==1)parX2<-mle4$par
if (i!=1)parX2<-rbind(parX2,mle4$par)

}

par(mfrow=c(3,2))

```

```

{
##med
ere<-c(1973, 1974 ,1975 ,1976 ,1977 ,1978 ,1979 ,1980 ,1982, 1983 ,1984 ,1985 ,
1986 ,1987 ,1988, 1989 ,1990 ,1991, 1992, 1993, 1994 ,1995 ,1996, 1997 ,1998 ,1999
,2000, 2001, 2002 ,2003,2004 ,2005 ,2006 ,2007 ,2008 ,2009 ,2010)
prov<- lm(parX2[,1]~bs(ere, knots = c(1980,2000), degree = 1, intercept = T)-1)
summary(prov)
parr[1:4]<-prov$coeff
summary(prov)
prov1<-predict(prov)
length(ere)
length(prov1)
plot((prov1)~ere,xlab="anni",ylab=expression(mu),type="l",ylim=c(min(parX2[,1]),
,max(parX2[,1])))
points(parX2[,1]~ere);title("parametro posizione")

##var

prov<- lm(parX2[,2]~bs(ere, knots = c(1980,1998), degree = 2, intercept = T)-1)
parr[5:9]<-prov$coeff
summary(prov)
prov1<-predict(prov)
length(ere)
length(prov1)
plot((prov1)~ere,xlab="anni",ylab=expression(omega),type="l",ylim=c(min(parX2[,2])
,max(parX2[,2])))
points(parX2[,2]~ere);title("parametro scala")

##alpha
prov<- lm(parX2[,3]~bs(ere, knots = c(1983,1999), degree = 2, intercept = T)-1)
parr[10:14]<-prov$coeff
summary(prov)
prov1<-predict(prov)
length(ere)
length(prov1)
plot((prov1)~ere,xlab="anni",ylab=expression(alpha),type="l",ylim=c(min(parX2[,3])
,max(parX2[,3])))
points(parX2[,3]~ere);title("parametro alpha")

##beta
prov<- lm(parX2[,4]~bs(ere, knots = c(1980,2000), degree = 1, intercept = T)-1)
parr[15:18]<-prov$coeff
summary(prov)
prov1<-predict(prov)
length(ere)
length(prov1)
plot((prov1)~ere,xlab="anni",ylab=expression(beta),type="l",ylim=c(min(parX2[,4])

```

```

,max(parX2[,4]))
points(parX2[,4]~ere);title("parametro beta")

##R

prov<- lm(parX2[,5]~bs(ere, knots = c(1980,1990), degree = 2, intercept = T)-1)
parr[19:23]<-prov$coeff
summary(prov)
prov1<-predict(prov)
length(ere)
length(prov1)
plot((prov1)~ere,type="l",xlab="anni",ylab=expression(R),ylim=c(min(parX2[,5])
,max(parX2[,5])))
points(parX2[,5]~ere);title("parametro adattamento")
}

myf_dens.t<-function(param,ere=1960:2010,dat=15:49) {
##dovo estrarre una matrice

paramet<-array(NA,5)
aa<-array(NA,dim=c(35,length(ere)),dimnames=list(as.character(dat),as.character(ere)))

paramet<-array(NA,dim=c(5,length(ere)),dimnames=list(c("mu","sig","alp","bet"
,"r"),as.character(ere)))
paramet[1,]<-apply((bs(ere, knots = c(1980,2000), degree = 1, intercept = T))
,1,function(x)param[1:4]%%x)#5
paramet[2,]<-apply((bs(ere, knots = c(1980,1998), degree = 2, intercept = T))
,1,function(x)param[5:9]%%x)#5
paramet[3,]<-apply((bs(ere, knots = c(1983,1999), degree = 2, intercept = T))
,1,function(x)param[10:14]%%x)#5
paramet[4,]<-apply((bs(ere, knots = c(1980,2000), degree = 1, intercept = T))
,1,function(x)param[15:18]%%x)#4
paramet[5,]<-apply((bs(ere, knots = c(1980,1990), degree = 2, intercept = T))
,1,function(x)param[19:23]%%x)#5

for (i in 1:length(ere)){
z<-(dat-paramet[1,i])/(paramet[2,i])
aa[,i]<-2/(paramet[2,i])*(dnorm(z))*
pnorm(paramet[3,i]*z+paramet[4,i]*z^3)*paramet[5,i]
}
aa}

```

```

#pseudo verosimiglianza
{

log_v_prof_mu<- function(param, dat_don,dat_tass,parr) {##mu
mle=nlminb(parr,function(x)(log_v(c(param,x),x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],dat=annidon)), lower=c(1e-10,-10,-20)
,upper=c(60,60,60),control = list(step.max=0.5,iter.max=1500,eval.max=600))
ll1<-log_v(c(param,mle$par[1:3]),dat_don,dat_tass,dat=annidon)
list(ver=ll1,x=mle$par,mes=mle$message,r2=mle$convergence)}

log_v_prof_se<- function(param, dat_don,dat_tass,parr) {##mu
mle=nlminb(parr,function(x)(log_v(c(x[1],param,x[2:3]),x_donne[ ,nazione,
annipresenti[i]],x_tass[,nazione,annipresenti[i]],dat=annidon)), lower=c(14,-10,-20)
,upper=c(49,60,60))
ll1<-log_v(c(mle$par[1],param,mle$par[2:3]),dat_don,dat_tass,dat=annidon)
list(ver=ll1,x=mle$par,mes=mle$message,r2=mle$convergence)}

log_v_prof_alp<- function(param, dat_don,dat_tass,parr) {##mu
oo<-dat_tass
mle=nlminb(parr,function(x)(log_v(c(x[1:2],param,x[3]),x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],dat=annidon)), lower=c(14,1e-10,-20),upper=c(49,60,60)
,control = list(step.max=0.5,iter.max=1500,eval.max=600))
ll1<-log_v(c(mle$par[1:2],param,mle$par[3]),dat_don,dat_tass,dat=annidon)
list(ver=ll1,x=mle$par,mes=mle$message,r2=mle$convergence)}

log_v_prof_beta <- function(param, dat_don,dat_tass,parr) {##beta
oo<-dat_tass
mle=nlminb(parr,function(x)(log_v(c(x,param),x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],dat=annidon)), lower=c(14,1e-10,-20),upper=c(49,60,60)
,control = list(step.max=0.5,iter.max=1500,eval.max=600))
ll1<-log_v(c(mle$par[1:3],param),dat_don,dat_tass,dat=annidon)
list(ver=ll1,x=mle$par,mes=mle$message,r2=mle$convergence)}

}
par(mfrow=c(4,1))
#stima
mle_mu<-array(NA,length(annipresenti))
se_mu1<-array(NA,length(annipresenti))
se_mu2<-array(NA,length(annipresenti))
for( i in 1:length(annipresenti)){
rr=mes2="false convergence (8)"
r=r2=1
param_iniz4<-parX[i,-1]
parr<-mle$par<-parX[i,1]

```

```

while(r!=0 |r2==1| sum(rr==mex)!=0 | sum(mes2==mex)!=0){
mle=nlminb(parr,function(x)sum(log_v_prof_mu(x,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$ver),lower=14
,upper=49)##come data devo dare gli y
rr<-mle$message;r<-mle$convergence
param_iniz4<- log_v_prof_mu(mle$par,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$x
parr=parr+runic(1,-1,1)
mes2<- log_v_prof_mu(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass[,nazione
,annipresenti[i]],param_iniz4)$mes
r2<-log_v_prof_mu(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass[,nazione
,annipresenti[i]],param_iniz4)$r2

}

lrt.ci1<-uniroot(function(x) -log_v(c(x,param_iniz4),x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],annidon)+ log_v(c(mle$par,param_iniz4),x_donne[ ,nazione
,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(16,mle$par))$roo

lrt.ci2<-uniroot(function(x) -log_v(c(x,param_iniz4),x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],annidon)+ log_v(c(mle$par,param_iniz4),x_donne[ ,nazione
,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(mle$par,39))$roo

mle_mu[i]<-mle$par
se_mu1[i]<-lrt.ci1
se_mu2[i]<-lrt.ci2
}

mle_sig<-array(NA,length(annipresenti))
se_sig1<-array(NA,length(annipresenti))
se_sig2<-array(NA,length(annipresenti))
for( i in 1:length(annipresenti)){
parr<-mle$par<-parX[i,2]
param_iniz4<-parX[i,-2]
r=r2=1
rr=mes2="false convergence (8)"

while(r!=0 |r2==1| sum(rr==mex)!=0 | sum(mes2==mex)!=0){
mle=nlminb(parr,function(x)log_v_prof_se(x,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$ver,lower=1e-20,upper=20)
r<-mle$convergence ;rr<-mle$message
parr=parr+runic(1,-1,1)
param_iniz4<- log_v_prof_se(mle$par,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$x
mes2<- log_v_prof_se(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass[
,nazione,annipresenti[i]],param_iniz4)$mes

```

```

r2<-log_v_prof_se(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass[
,nazione,annipresenti[i]],param_iniz4)$r2

}

lrt.ci1<-uniroot(function(x) -log_v(c(param_iniz4[1],x,param_iniz4[-1]),x_donne[
,nazione,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+ log_v(
c(param_iniz4[1],mle$par,param_iniz4[-1]),x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(3,mle$par))$root

lrt.ci2<-uniroot(function(x) -log_v(c(param_iniz4[1],x,param_iniz4[-1]),x_donne[
,nazione,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+ log_v(
c(param_iniz4[1],mle$par,param_iniz4[-1]),x_donne[ ,nazione,annipresenti[i]],x_tass
[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(mle$par,15))$root

mle_sig[i]<-mle$par
se_sig1[i]<-lrt.ci1
se_sig2[i]<-lrt.ci2
}

mle_alp<-array(NA,length(annipresenti))
se_alp1<-array(NA,length(annipresenti))
se_alp2<-array(NA,length(annipresenti))
for( i in 1:length(annipresenti)){
parr<-mle$par<-parX[i,3]
param_iniz4<-parX[i,-3]
r=r2=1
rr=mes2="false convergence (8) "
while(r==1 | r2==1 | sum(rr==c(mex,"X-convergence (3) " ))!=0 | sum(mes2==c
(mex,"X-convergence (3) " ))!=0){
mle=nlminb(parr,function(x)(log_v_prof_alp(x,x_donne[ ,nazione,annipresenti
[i]],x_tass[,nazione,annipresenti[i]],param_iniz4)$ver) ,lower=-1,upper=2)
r<-mle$convergence ;rr<-mle$message
parr=parr+runif(1,-0.1,0.1)

param_iniz4<- log_v_prof_alp(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass
[,nazione,annipresenti[i]],param_iniz4)$x
mes2<- log_v_prof_alp(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass
[,nazione,annipresenti[i]],param_iniz4)$mes
r2<- log_v_prof_alp(mle$par,x_donne[ ,nazione,annipresenti[i]],x_tass
[,nazione,annipresenti[i]],param_iniz4)$r2
}
lrt.ci1<-uniroot(function(x)-log_v(c(param_iniz4[1:2],x,param_iniz4[3]),x_donne
[ ,nazione,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+ log_v(
c(param_iniz4[1:2],mle$par,param_iniz4[3]),x_donne[ ,nazione,annipresenti[i]],
x_tass[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(-2,mle$par))$root

```

```
lrt.ci2<-uniroot(function(x) -log_v(c(param_iniz4[1:2],x,param_iniz4[3]),x_donne[
,nazione,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+ log_v(c(param_iniz4[1:2]
,mle$par,param_iniz4[3]),x_donne[ ,nazione,annipresenti[i]],x_tass[,nazione,annipresenti[i]]
,annidon)+qchisq(0.95,1)/2,c(mle$par,5))$root
```

```
mle_alp[i]<-mle$par
se_alp1[i]<-lrt.ci1
se_alp2[i]<-lrt.ci2
}
```

```
mle_bet<-array(NA,length(annipresenti))
se_bet1<-array(NA,length(annipresenti))
se_bet2<-array(NA,length(annipresenti))
for( i in 1:length(annipresenti)){
parr<-mle$par<-parX[i,4]
param_iniz4<-parX[i,-4]
r=r2=1
rr=mes2="false convergence (8)"

while(r!=0|r2==1 | sum(rr==mex)!=0 | sum(mes2==mex)!=0){
mle=nlminb(parr,function(x)(log_v_prof_beta(x,x_donne[ ,nazione,annipresenti[i]
],x_tass[,nazione,annipresenti[i]],param_iniz4)$ver),lower=-10,upper=49)
r<-mle$convergence ;rr<-mle$message
parr=parr+runif(1,-1,1)
param_iniz4<- log_v_prof_beta(mle$par,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$x
mes2<- log_v_prof_beta(mle$par,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$mes
r2<- log_v_prof_beta(mle$par,x_donne[ ,nazione,annipresenti[i]]
,x_tass[,nazione,annipresenti[i]],param_iniz4)$r2
}
}
```

```
lrt.ci1<-uniroot(function(x) -log_v(c(param_iniz4,x),x_donne[ ,nazione,
annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+ log_v(
c(param_iniz4,mle$par),x_donne[ ,nazione,annipresenti[i]],x_tass
[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(-1,parX[i,4]))$root
```

```
lrt.ci2<-uniroot(function(x)-log_v(c(param_iniz4,x),x_donne[ ,nazione
,annipresenti[i]],x_tass[,nazione,annipresenti[i]],annidon)+ log_v
(c(param_iniz4,mle$par),x_donne[ ,nazione,annipresenti[i]],x_tass
[,nazione,annipresenti[i]],annidon)+qchisq(0.95,1)/2,c(parX[i,4],20))$root
```

```

mle_bet[i]<-mle$par
  se_bet1[i]<-lrt.ci1
  se_bet2[i]<-lrt.ci2
}

####identificazione mode

prova<-function(media=5,varianza=3,al.val,bet.val,from=-8,to=20){
  punti<-rep(NA,2)
  mio<-rep(NA,2)
  dens<-rep(NA,2)
  punti_di_maxmin<-array(NA,50)
  maxi<-array(NA,length(punti_di_maxmin))

  ff<-function(x)grad(function(x)densita(x,c(media,varianza,al.val,bet.val)),x)
  ff<-Vectorize(ff, vectorize.args = "x")
  gg<-function(x)grad(function(x)ff(x),x)
  gg<-Vectorize(gg,"x")
  # curve(ff(x),from=-4,to=4)
  punti_di_maxmin<-uniroot.all(ff, interval = c(from, to))
    curve( densita(x,c(media,varianza,al.val,bet.val)),from=from,to=to,main=str_c
      (round(al.val,1),round(bet.val,1),sep = "  "))#c()
  # curve(ff(x),from=-4,to=4)
  for(k in 1:length(punti_di_maxmin) ){
    if (!is.na(punti_di_maxmin[k])){
      maxi[k]<-gg(punti_di_maxmin[k])<c(-1e-10)
      # print(hessian(function(x)densita(x,c(0,1,al.val[i],bet.val[j]))
      ,punti_di_maxmin[k,]))
    }
  }

  punti<-punti_di_maxmin[maxi][which(!is.na(punti_di_maxmin[maxi]))]
  if (length(punti)==1)punti[2]<-punti[1]
  for(d in 1:2){
    if (!is.na(punti[d])){
      abline(v=punti[d],col=d+1)
    }
  }
  dens<-densita(punti[which(!is.na(punti))],c(media,varianza,al.val,bet.val))
  mio[1]<-3
  mio[2]<-3
  if( length(unique(punti))==1){

    ###troppo tempo
    # curve(gg,-int,int)

    # curve(gg,-7, 7); abline(h=0)
    ptt<-uniroot.all(gg, c(from, to))
  }
}

```



```

ptt<-ptt[which(densita(ptt,c(media,varianza,al.val,bet.val))>0.000003)]

if(length(ptt)==3){
if (ptt[2]<(punti[1])){
pto=(ptt[1]);punti[1]=pto
                                abline(v=(pto),col=4)
mio[1]<-1.5
}
if (ptt[2]>=(punti[2])){
pto=(ptt[3]);punti[2]=pto
                                abline(v=(pto),col=5)
mio[2]<-1.5
}
}
  if(length(ptt)==4){
if (ptt[2]<(punti[1])){
pto=(ptt[2]);punti[1]=pto
mio[1]<-1.5
                                abline(v=(pto),col=6)
}
if (ptt[2]>=(punti[2])){
pto=(ptt[3]);punti[2]=pto
                                abline(v=(pto),col=7)
mio[2]<-1.5
} } }
dens<-densita(punti,c(media,varianza,al.val,bet.val))
med<-integrate(function(x) x*densita(x,c(media,varianza,al.val,bet.val))
,lower=-Inf,upper=Inf)$value
m3<-integrate(function(x) (((x-media)/varianza)^3)*densita(x,c(media,varianza
,al.val,bet.val)),lower=-Inf,upper=Inf)$value

c(dens=dens,punt=punti,pes=mio,m1=med,m3=m3)
}

densitaclassica<-function(x,param) {
z<-(x-param[1])/(param[2])
(2/(param[2]))*(dnorm(z))*
pnorm(param[3]*z+(z^3)*param[4])
}

densitariparametrizzata<-function(x,param) {
z<-(x-param[1])/(param[2])
(2/(param[2]))*(dnorm(z))*
pnorm(param[3]*(z+(z^3)*param[4]))
}

```

Bibliografia

Azzalini (1985), A class of distributions which includes the normal ones, *Scandinavian journal of statistics*, pag. 171–178

Azzalini (2004), Inferenza statistica: Una presentazione basata sul concetto di verosimiglianza, *Springer*

Azzalini (1985), The Skew-normal Distribution and Related Multivariate Families, *Scandinavian journal of statistics*, vol. 32, n. 2 pag.159–188

Hadwiger, Hugo (1940), Eine analytische Reproduktionsfunktion für biologische Gesamtheiten, *Scandinavian Actuarial Journal*, n. 3-4, pag. 101–113

S. Mazzucco, B. Scarpa (2011) , Fitting age-specific fertility rates by a skew-symmetric probability density function, *Working Paper Series*, N. 10

Ma and Genton(2004), Flexible Class of Skew-Symmetric Distributions, *Scandinavian Journal of Statistics*, vol. 31, N. 3, pag. 459–468

Eurostat (2013), Fertility statistics - Statistics Explained, http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Fertility_statistics

Palestinian Central Bureau of Statistics (2013), Search Database , http://www.pcbs.gov.ps/PCBS_2012/Publications.aspx?CatId=8&scatId=261

Human Fertility Database (2013), Search Database , <http://www.humanfertility.org/cgi-bin/country.php?country=USA&tab=asfr&t1=3&t2=4>

Eurostat (2013), Search Database ,http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database