

Università degli Studi di Padova

CORSO DI LAUREA MAGISTRALE IN  
SCIENZE STATISTICHE



Un approccio di verosimiglianza  
composita per la previsione  
del sesso dei nati

A composite likelihood approach to  
predict the babies sex

*Relatore* Prof. Bruno Scarpa

Dipartimento di Scienze Statistiche

*Co-relatore* Prof. Nicola Sartori

Dipartimento di Scienze Statistiche

*Laureando* Simone Tiberi

SESSIONE DI LAUREA 23 NOVEMBRE 2012

ANNO ACCADEMICO 2012-2013



# Indice

<b>Introduzione</b>	<b>5</b>
<b>1 Rassegna della letteratura</b>	<b>7</b>
1.1 I giorni dei rapporti sessuali . . . . .	7
1.2 Il metodo Shettles . . . . .	8
1.3 Le difficoltà di questi studi . . . . .	9
<b>2 I dati utilizzati</b>	<b>13</b>
2.1 Fertili . . . . .	13
2.1.1 Le variabili . . . . .	14
2.1.2 La stima del giorno dell'ovulazione . . . . .	14
2.2 Billings . . . . .	15
2.3 La finestra di fertilità . . . . .	17
2.4 Le diverse numerosità delle probabilità . . . . .	17
2.5 Le gerarchie . . . . .	19
2.5.1 Il fattore coppia . . . . .	20
2.5.2 I fattori centro ed istruttrice . . . . .	21
2.6 Gemelli ed aborti . . . . .	21
<b>3 La verosimiglianza composta</b>	<b>25</b>
3.1 La matrice di Godambe . . . . .	26
3.2 Proprietà asintotiche dello stimatore di massimo . . . . .	27
3.3 I test d'ipotesi . . . . .	27
3.3.1 Il test di Wald . . . . .	28
3.3.2 Il test del punteggio . . . . .	28
3.3.3 Il test del rapporto di verosimiglianza . . . . .	29
3.4 Gli intervalli di confidenza . . . . .	30
<b>4 Il modello proposto</b>	<b>31</b>
4.1 Le probabilità di concepimento . . . . .	31
4.1.1 Il modello di Barrett e Marshall . . . . .	31

4.1.2	Il modello di Schwartz, MacDonald e Heuchel . . . . .	32
4.1.3	Le due formulazioni a confronto . . . . .	32
4.2	Le probabilità dei sessi . . . . .	33
4.3	Il modello completo . . . . .	35
4.4	Il calcolo dei pesi $a$ e $b$ . . . . .	36
4.4.1	La convalida incrociata . . . . .	37
4.4.2	L'errore sui dati Billings . . . . .	39
4.5	Le stime puntuali dei parametri . . . . .	40
4.6	I Test bootstrap sulla finestra intera . . . . .	44
4.6.1	I risultati dei test . . . . .	46
4.6.2	L'analisi di sensitività . . . . .	48
4.7	I Test asintotici sulla finestra ridotta . . . . .	50
4.7.1	Il calcolo analitico delle derivate prime . . . . .	51
4.7.2	Le matrici $G(\theta)$ , $J(\theta)$ ed $H(\theta)$ . . . . .	53
4.7.3	I risultati dei test . . . . .	55
4.8	Gli intervalli di confidenza . . . . .	58
4.8.1	Le regioni trovate . . . . .	58
4.9	Un confronto tra i metodi . . . . .	60
4.9.1	Intervallo di confidenza per $\beta_{-6}$ tramite bootstrap . . .	64
4.10	Commenti ai risultati . . . . .	66
4.10.1	L'utilità della verosimiglianza composita . . . . .	66
<b>5</b>	<b>Il modello gerarchico</b>	<b>69</b>
5.1	La formulazione . . . . .	69
5.2	Le stime puntuali dei parametri . . . . .	72
5.3	Il calcolo del peso $a$ . . . . .	74
5.4	I test bootstrap . . . . .	75
5.4.1	I risultati dei test . . . . .	76
5.4.2	Un'analisi di sensitività . . . . .	78
5.5	Considerazioni sui risultati ottenuti . . . . .	81
5.5.1	Un confronto con la letteratura . . . . .	81
5.5.2	I test multipli . . . . .	82
	<b>Conclusioni</b>	<b>85</b>
	<b>A Tabelle e grafici aggiuntivi</b>	<b>87</b>
	<b>B Comandi R</b>	<b>93</b>
	<b>Bibliografia</b>	<b>113</b>

# Introduzione

Condizionare anticipatamente il sesso del nascituro è ed è sempre stato un frequente desiderio dei futuri genitori. La letteratura scientifica riporta numerosi studi atti ad indagare i fattori incidenti sulla determinazione del sesso al fine di poterlo influenzare: ciò non solo per allietare le preferenze parentali, ma per ragioni sociali e biologiche.

Se idealmente si riuscisse ad influire sul sesso del nato, si potrebbe ad esempio arginare il fenomeno dell'aborto selettivo e dell'infanticidio, conseguenza della "politica del figlio unico" che dagli anni '90 obbliga al pagamento di una tassa per ogni figlio successivo al primo e che, in un paese come la Cina, dove uomini e donne non hanno ancora raggiunto l'equiparazione sociale, ha condannato migliaia di bambine alla morte (Wikipedia, Politica del figlio unico e CIA, The World Factbook). Nel 2011 il numero di nuovi nati maschi cinesi dichiarati all'anagrafe è stato infatti di 113 ogni 100 femmine, a fronte di un dato mondiale<sup>1</sup> di 106 a 100.

Anche il sovraffollamento e lo squilibrio demografico nei sessi, in questo caso a favore delle femmine, di alcune aree del 3° mondo potrebbe giovare da questi studi: spesso molte coppie continuano a procreare fino all'ottenimento di un figlio maschio, anche in questi paesi considerato di maggior valore rispetto ad una bambina (McSweeney, 2011).

Rimarchevole sarebbe anche l'utilità medica: alcune malattie sono associate ai cromosomi sessuali X ed Y, dunque la loro trasmissione è legata al sesso. Qualora ci fosse familiarità per tali patologie, influenzare la sessualità del nascituro accrescerebbe la probabilità che questo fosse in salute. Le malattie associate al cromosoma Y non possono infatti manifestarsi su una bambina, caratterizzata da due X; quelle causate da geni presenti sull'X paterno sono invece trasmissibili solo ad una femmina, visto che un maschio riceve dal padre il cromosoma Y. In aggiunta, le patologie legate a geni recessivi posti sull'X materno sarebbero causa improbabile di malattia in una bambina, poiché ella per ammalarsi dovrebbe ricevere anche una X con la

---

<sup>1</sup>Tale valore è stato pari a 107:100 nel 2011 a causa dell'influenza della Cina stessa, ma escludendo tale paese dalla statistica si ottiene 106:100

medesima patologia dal padre, il che è infrequente. Esistono già metodi di concepimento in vitro che consentono di selezionare gli spermatozoi di uno o dell'altro sesso sino a giungere a quote del 75% dei gameti maschili desiderati (Veit e Jewelewicz, 1988). Tuttavia tali metodologie non sono applicabili su larga scala in quanto costose ed impegnative per la donna che deve sottoporsi a pesanti terapie ormonali.

Questo elaborato si pone come obiettivo quello di indagare la relazione tra il giorno in cui avviene il rapporto che determina il concepimento, considerato rispetto alla data dell'ovulazione, ed il sesso del nascituro: cercheremo di capire se si possa influenzare probabilisticamente il sesso del futuro bambino attraverso i giorni in cui avvengono gli atti sessuali.

La questione è già stata studiata più volte in precedenza (una rassegna sulla letteratura è ad esempio presente in Mauro, 2011) e si è notato che un approccio di verosimiglianza è difficilmente applicabile. Infatti, il giorno della relazione sessuale che genera il concepimento è raramente osservabile, poiché in ogni ciclo sussistono generalmente più rapporti dei quali uno soltanto è causa della fecondazione: come conseguenza è opportuno modellare congiuntamente la probabilità di concepire e quella di procreare un maschio (o una femmina) dato il concepimento. Tuttavia il sesso del nato è osservabile solo per i cicli in cui l'ovulo è fecondato: risulta quindi impossibile trattare contemporaneamente le due probabilità con una verosimiglianza congiunta. Abbiamo per tale motivo deciso di adoperare un'approssimazione a quest'ultima: la verosimiglianza composita, una pseudo verosimiglianza (Lindsay, 1998 e Besag 1974), il cui utilizzo è giustificato da buone proprietà che saranno enunciate nel seguito.

La tesi è così costruita: nel Capitolo 1 descriveremo alcuni metodi e modelli esistenti in letteratura per lo studio dei sessi dei neonati, e le difficoltà strutturali che questo tipo di analisi presenta; nel Capitolo 2 mostreremo i dati a disposizione e quali accorgimenti abbiamo adottato per trattare determinate peculiarità che questi presentano; nel Capitolo 3 introdurremo la verosimiglianza composita, spiegheremo perché essa sia utile ai fini della nostra analisi ed esporremo alcuni risultati asintotici; nel Capitolo 4 definiremo la formulazione delle probabilità trattate, illustreremo la struttura del 1° modello utilizzato, basato sull'ipotesi che le variabili di interesse tra i cicli siano i.i.d., ed esporremo i risultati ottenuti; nel Capitolo 5 correggeremo l'assunzione di indipendenza inserendo un effetto gerarchico e mostreremo come questa modifica abbia inciso sui risultati; infine trarremo le conclusioni dell'analisi, evidenziando i risultati trovati ed i relativi limiti.

# Capitolo 1

## Rassegna della letteratura

In molti si sono interrogati sull'esistenza di variabili che possano incidere sulle probabilità che il nato sia maschio o femmina. A partire da antiche credenze popolari sino a giungere a più recenti studi scientifici, svariate analisi sono state portate avanti con l'intento di provare o smentire il legame esistente tra il sesso del neonato e determinati fattori.

Alcuni tra quelli indagati sono: la durata della fase follicolare (ossia il giorno dell'ovulazione rispetto al 1° giorno del ciclo), la dieta seguita dalla donna, l'età dei genitori, la vicinanza ad eventi storici, la temperatura e l'inquinamento ambientale, il numero di mesi per cui si è tentato di concepire prima di riuscirvi ed il giorno del rapporto che genera il concepimento rispetto a quello dell'ovulazione.

In questo elaborato abbiamo deciso di studiare proprio quest'ultima variabile: la data della relazione sessuale che porta alla fecondazione rispetto all'ovulazione. Molti dei fattori sopra citati non sono infatti modificabili dalla coppia che vuole influenzare il sesso del proprio figlio, come gli eventi storici, altri semplicemente non erano studiabili poiché le informazioni a riguardo non sono presenti nei dati disponibili, come la dieta della madre.

### 1.1 I giorni dei rapporti sessuali

È stata spesso cercata l'esistenza di una relazione tra il giorno del rapporto che genera il concepimento, rispetto a quello dell'ovulazione, ed il sesso del neonato. Le ipotesi effettuate a riguardo sono notevolmente variegata e discordanti tra loro: i giorni in cui sarebbe più facile concepire un maschio (per una femmina vale l'opposto) sono stati talvolta identificati come quelli prossimi all'ovulazione, talora come i primi fertili e quindi più lontani dall'ovulazione.

Una tra le teorie più significative è costituita dal metodo Billings (Billings e Westmore, 1980); questa si focalizza sulle caratteristiche fisiche degli spermatozoi: quelli portatori del cromosoma X, che danno vita ad una femmina, sarebbero più forti, avrebbero una vita più lunga ma sarebbero anche meno veloci rispetto a quelli portatori del cromosoma Y, che generano un maschio.

In caso di rapporto vicino o subito dopo rispetto al momento dell'ovulazione, giorno stesso o quelli successivi, sarebbe dunque maggiormente probabile concepire un maschio, poiché gli spermatozoi caratterizzati dal cromosoma Y sarebbero i primi ad arrivare, essendo i più celeri. In caso di rapporto nei primi giorni di fertilità gli spermatozoi con cromosoma X, dotati di maggiore resistenza, diverrebbero invece in superiorità numerica al momento dell'ovulazione, e quindi una femmina sarebbe più probabile.

Due analisi (Dedé, 1988 e McSweeney, 1993) riportano risultati estremamente concordanti con il metodo Billings: rispettivamente l'83.6% ed il 94.9% delle coppie che hanno seguito le indicazioni sopra riportate hanno ottenuto in figlio del sesso desiderato, su un totale di 73 e 99 unità.

Ciò nonostante altri precedenti studi hanno evidenziato legami differenti del giorno del rapporto con il sesso. Ad esempio Guerrero (1970), uno studioso sudamericano che si è occupato del problema in questione negli anni '70, basandosi su 595 concepimenti, afferma che la maggiore probabilità di concepire un maschio, tra il 63% ed il 67%, si ha dai 5 ai 9 giorni prima dell'ovulazione. In un altro studio (Guerrero, 1974), dopo aver esaminato 875 fecondazioni, aggiunge che tale quantità passa dal 68% al 44% in caso di rapporto rispettivamente almeno 6 giorni prima dell'ovulazione o il giorno della stessa.

Sottolineiamo che Guerrero esamina la temperatura corporea basale per stimare la data dell'ovulazione, mentre Billings e le analisi correlate per raggiungere lo stesso obiettivo analizzano il muco cervicale (nei Paragrafi 2.1 e 2.2 i metodi verranno descritti accuratamente): tale fattore rende perciò più complesso un confronto tra questi studi.

Ulteriori analisi hanno invece smentito l'esistenza di una relazione tra le due variabili in esame (Carson, 1988 e Gray, 1991).

## 1.2 Il metodo Shettles

Nel 1960 Shettles ideò una teoria fondata in parte sulle medesime idee sulla morfologia degli spermatozoi del metodo Billings, in parte sulle caratteristiche del pH del tratto interno femminile. Secondo tale studioso un ambiente acido ucciderebbe i più deboli spermatozoi Y in gran quantità e lascerebbe una maggior quota di quelli portatori del cromosoma X, favorendo il concepimento.



mento di femmine; d'altro canto un ambiente alcalino (non acido) renderebbe più probabile la fecondazione di maschi. In base a tali considerazioni Shettles raccomanda una doccia di acqua e aceto (acido) o di acqua e bicarbonato di sodio (alcalino) subito prima del rapporto in caso si voglia concepire rispettivamente una femmina o un maschio. Shettles consiglia inoltre altre accortezze, che non illustreremo in questa sede, da adottare per influenzare il sesso relative a posizione del rapporto, penetrazione al momento dell'eiaculazione ed orgasmo femminile.

La metodica avrebbe riscontrato ottimi risultati (Rorvik e Shettles, 1970), sebbene sia doveroso sottolineare l'esigua numerosità a disposizione dello studio: 19 successi in 22 tentativi di concepire un maschio e 16 successi in 19 tentativi di procreare una femmina. Altri però (Williamson et al., 1988) hanno trovato risultati non in linea con quelli di Shettles ed hanno mostrato come il suo procedimento sia eccessivamente elaborato per essere applicato su larga scala. Segnaliamo infine come in un'analisi (Diaso e Glass, 1971) sia stato smentito il fatto che il pH dei mezzi influenzi la mobilità degli spermatozoi che li attraversano.

Nei dati in nostro possesso non vi sono le informazioni necessarie per verificare la validità del metodo descritto, ciò non rientra comunque negli obiettivi da noi prefissati. Ci concentreremo invece nell'illustrare l'influenza sul sesso del nato dell'unico aspetto di nostro interesse trattato da Shettles: i giorni dei rapporti.

Basandosi sulle stesse idee del metodo Billings, anche le indicazioni risultano analoghe: rapporti avvenuti 3 o più giorni prima dell'ovulazione aumenterebbero le possibilità di avere una femmina; l'opposto varrebbe in caso di rapporto da 2 giorni prima sino a qualche giorno dopo l'ovulazione, mentre attorno alle 48 ore precedenti l'ovulazione il tasso dei due sessi sembrerebbe essere uguale.

### 1.3 Le difficoltà di questi studi

Queste analisi sono complicate dal fatto che solitamente, tranne in rare collezioni di dati, la variabile che si vuole studiare, ossia il giorno del rapporto che genera il concepimento, è in realtà difficilmente osservabile. Soltanto quando il concepimento è il risultato di una singola giornata con rapporti nel corso dell'intero ciclo, è possibile identificare con certezza il giorno che porta alla fecondazione dell'ovulo. Nella maggior parte dei cicli tuttavia vi sono più date con rapporti, ma solo una è quella che dà origine al concepimento. In questi casi è necessario stimare la probabilità che, in un ovulo fecondato, il concepimento sia avvenuto in un determinato giorno; tale valore è pari a 0

nelle date senza rapporti, e positivo per ciascuna delle giornate in cui vi è almeno un atto sessuale.

Per semplicità ci si restringe ad una finestra di fertilità più ridotta dell'intero ciclo: questa identifica gli unici giorni ritenuti fertili e contiene quello dell'ovulazione stesso, definito come "giorno 0". Si escludono cioè le date esterne a tale intervallo poiché si considera "quasi certo" che il concepimento possa avvenire solo al suo interno.

Complesso è anche definire la finestra stessa: la vita di ovuli e spermatozoi sembrerebbe essere infatti limitata rispettivamente a circa 24 ore e 5 giornate (Wilcox et al., 1995); sarebbe quindi in linea di principio possibile considerare il solo periodo che va da 5 giorni prima dell'ovulazione ad 1 dopo. Ma naturalmente è da considerarsi la variabilità biologica presente in ciascuna donna ed il fatto che in realtà l'ovulazione avviene in un preciso istante e, dato il breve periodo in cui l'ovulo è fecondabile, tenere traccia solo del giorno in cui essa avviene è di per sé un'approssimazione: ben diverso è infatti il caso in cui l'ovulazione si verifichi al mattino o alla sera.

Inoltre il "giorno 0" è solo stimato e metodi diversi danno luogo a stime leggermente differenti quindi, in linea di principio, tale valore non dovrebbe essere trattato come una costante nota.

Un ulteriore problematica è relativa alla difficoltà di modellare la probabilità di concepire un nato di un certo sesso.

Definiamo le seguenti quantità, dove  $Z_i$  ed  $Y_i$  sono delle variabili casuali:  $Z_i = 1$  se nel ciclo  $i$ -esimo si concepisce un maschio, 0 se una femmina;  $Y_i = 1$  se il ciclo  $i$ -esimo si verifica un concepimento, 0 altrimenti;  $X_i = (X_{i1}, \dots, X_{iK})$ , dove  $K =$  numero dei giorni della finestra di fertilità, e  $X_{ik} = 1$  se nel giorno  $k$  del ciclo  $i$ -esimo si ha almeno un rapporto, 0 altrimenti.

Considerato che il sesso è osservabile solo nel caso di concepimento, l'evento  $Z_i = 1$  risulta equivalente all'evento "fecondazione dell'ovulo nel ciclo  $i$ -esimo e procreazione di un maschio nel ciclo  $i$ -esimo". Possiamo quindi scomporre la probabilità di procreare un maschio nel ciclo  $i$ -esimo come prodotto tra la probabilità di fecondare l'ovulo in tale ciclo e la probabilità di concepire un maschio condizionatamente all'avvenuta fecondazione,

$$Pr(Z_i = 1|X_i) = Pr(Y_i = 1|X_i)Pr(Z_i = 1|Y_i = 1, X_i). \quad (1.1)$$

Questa formula andrebbe moltiplicata per la probabilità che, dato un concepimento, non vi siano aborti nei primi due mesi di gravidanza sino all'osservazione del sesso. Tale quantità è tuttavia trascurabile poiché l'interesse cade sulle probabilità di procreare un maschio condizionatamente all'aver concepito in quel giorno e, di conseguenza, al fatto che il sesso sia osservabile.

I cicli che portano al concepimento in cui il sesso non è osservabile, per dato mancante o aborto precoce, non vengono considerati e si assume che questi eventi siano indipendenti dal sesso del feto.

Mentre la variabile  $Y_i$  è sempre osservabile,  $Z_i$  lo è solo per i cicli in cui si osserva un concepimento (e di cui si conosce il sesso, esclusi quindi gli aborti ed i dati mancanti). Nel seguito proporremo modelli statistici parametrici per modellare  $Pr(Z_i = 1|X_i)$  ricorrendo alla scomposizione (4).



# Capitolo 2

## I dati utilizzati

I dati a disposizione sono stati raccolti in due studi coordinati dal Professor Colombo (Dipartimento di Scienze Statistiche, Università di Padova): Fertili (Colombo e Masarotto, 2000) e Billings (Colombo et al., 2006).

Precisiamo che in questa sezione saranno solo descritte e riportate le informazioni di maggior rilevanza ai fini della nostra indagine; rimandiamo a Colombo e Masarotto (2000), Colombo et al. (2006) e Mauro (2011) per una più dettagliata analisi descrittiva dei due database.

### 2.1 Fertili

Il primo insieme di dati contiene la maggior parte delle unità a nostra disposizione ed è composto principalmente da una raccolta prospettica effettuata tra il 1992 ed il 1996 in sette istituti europei che offrivano servizi legati alla consapevolezza della fertilità e alla pianificazione familiare: Milano, Verona, Lugano, Dusseldorf, Parigi, Londra e Bruxelles.

I soggetti dello studio erano coppie sposate o in una relazione stabile. I criteri di entrata erano i seguenti: avere esperienza nella pianificazione familiare, avere tra i 18 ed i 40 anni, aver avuto almeno una mestruazione dopo la fine dell'allattamento in caso di precedente gravidanza (per la donna), non assumere farmaci o ormoni che potessero influenzare la fertilità, non essere permanentemente infertile, non essere affetti da malattie che potessero causare sub-fertilità e non alternare rapporti protetti e non.

A queste unità sono stati aggregati retrospettivamente i dati raccolti in un precedente studio prospettico condotto tra il 1979 ed il 1985 in Nuova Zelanda (Auckland) e coordinato dal Professor France. Quest'ultimo campione, sebbene nettamente più piccolo di quello europeo, è stato aggiunto per la sua appropriatezza ai fini degli studi sui sessi e sulle probabilità di

concepimento: tutte le donne dell'analisi erano obbligate ad avere un solo rapporto all'interno dei giorni fertili del ciclo. Questo vincolo consente di conoscere con esattezza il giorno del rapporto che genera il concepimento nel caso in cui si osservi una gravidanza; d'altro canto però esso ha comportato un forte abbandono delle coppie: le donne restano in media nello studio per soli 3 cicli contro un valore di quasi 10 osservato su tutti i dati a disposizione. Requisito d'ingresso era che le coppie dovessero avere una comprovata fertilità con almeno una gravidanza alle spalle.

Lo studio Fertili ha coinvolto un totale di 881 donne, delle quali 782 provenienti dai centri europei e 99 da Auckland.

### 2.1.1 Le variabili

All'entrata nello studio sono state registrate le seguenti informazioni: mese ed anno di nascita dei membri della coppia, numero di precedenti gravidanze, data dell'ultimo parto (o aborto) e fine dell'allattamento, data dell'ultimo contraccettivo assunto, data del matrimonio (per le coppie sposate); nel corso dello studio è stato anche memorizzato nella base dati il sesso dei bambini nati.

Inoltre giornalmente la coppia annotava se vi erano stati rapporti non protetti, la temperatura corporea basale della donna (BBT, basal body temperature) misurata la mattina al risveglio prima di iniziare qualsiasi attività, ed il muco cervicale, espresso su una scala quantitativa da 0 a 4 (vedi Tabella 2.1). Un concepimento è definito in presenza di una gravidanza a 60 giorni dalle ultime mestruazioni o in caso di aborto clinicamente identificato prima di questo termine.

### 2.1.2 La stima del giorno dell'ovulazione

Il giorno dell'ovulazione è stato stimato, in ogni ciclo in cui vi fosse abbastanza informazione a disposizione, tramite due metodi.

Il primo si basa sulle caratteristiche del muco cervicale e definisce il picco del muco come l'ultimo giorno in cui si identifica la migliore qualità di muco tramite la sensazione o l'apparenza, noto quindi retrospettivamente. Questo è preso come "giorno 0" e l'ovulazione è attesa entro circa due giornate.

Il secondo metodo consiste nell'esamine della temperatura corporea basale ed utilizza la regola del "tre su sei": si considera il primo periodo del ciclo in cui tre temperature consecutive superano le sei immediatamente precedenti; l'ultima data tra le sei a temperatura inferiore è identificata come rialzo termico: essa rappresenta la stima del "giorno 0".

Codice	Sensazione	Aspetto
0	Nessuna informazione	Nessuna informazione
1	Asciutto, secchezza prurito o nessuna informazione	Non si nota nulla, non è visibile muco
2	Sensazione di umidità	Non si nota nulla, non è visibile muco
3	Sensazione di umidità	Spesso, biancastro opaco, elastico, non filamentoso
4	Bagnato, scivoloso, vischioso	Trasparente, limpido, elastico, filante liquido come acqua

Tabella 2.1: Classificazione dei sintomi del muco cervicale, dati Fertili. In caso di più osservazioni del muco in un giorno, viene considerata quella dalle caratteristiche più fertili

Le numerosità dei cicli complessivi e di quelli in cui sono presenti le informazioni sulla stima del “giorno 0” sono riportate in Tabella 2.4.

## 2.2 Billings

Lo studio Billings consiste in dati provenienti da 193 donne raccolti in quattro istituti italiani tra il 1993 ed il 1997 (Milano, Parma, Saluzzo e Roma). L’obiettivo iniziale di questa indagine consisteva nell’esaminare la relazione tra le caratteristiche del muco cervicale ed i giorni fertili del ciclo.

I criteri di entrata sono i medesimi descritti nello studio Fertili, con la differenza che si fa esplicito riferimento ai metodi di ovulazione Billings per quanto riguarda la richiesta di esperienza nella pianificazione familiare.

Anche le variabili registrate sono quasi le stesse citate sopra, compresa la definizione di concepimento; le sole eccezioni sono rappresentate dalla temperatura corporea, assente, e dal muco cervicale, espresso in una scala differente in 5 categorie (vedi Tabella 2.2).

Oltre alla categorizzazione del muco anche la definizione del suo picco varia leggermente: si definisce picco del muco l’ultimo giorno del ciclo durante il quale almeno una caratteristica di elevata fertilità è osservata o percepita nel muco; considerando come sensazioni di alta fertilità una sensazione di bagnato/non viscido e/o viscido, trasparente, fluido o acquoso, a volte associato

Codice	Sensazione	Apparenza
0	Nessuna informazione	Nessuna informazione
1	Nessuna oppure asciutto	Nessun tipo di muco o di perdita
2	Sensazione di umidità	Nessun tipo di muco o di perdita
3	Sensazione di umidità	Muco denso appiccicoso, biancastro o giallastro
4	Sensazione di bagnato, scivoloso	Muco trasparente liquido come l'acqua
5	Sensazione di bagnato-lubrificato	Muco trasparente ed elastico, simile a chiara d'uovo, acquoso, talvolta tracce di sangue

Tabella 2.2: Classificazione dei sintomi del muco cervicale, dati Billings. In caso di più osservazioni del muco in un giorno, viene considerata quella dalle caratteristiche più fertili

a tracce di sangue. Questo giorno deve essere preceduto da un adeguato aumento nelle sensazioni e nell'apparenza delle caratteristiche del muco, e seguito da un cambiamento improvviso.

Come conseguenza i metodi di stima del “giorno 0” basati sul muco cervicale di Fertili e Billings risultano leggermente differenti; non è quindi opportuno pensare di aggregare i dati dei due studi (basandoci sulle informazioni relative al picco del muco) poiché le classi in cui il muco è suddiviso differiscono nei due metodi ed una classificazione non è direttamente riconducibile all'altra (vedi Tabelle 2.1 e 2.2). Sarebbe quindi un'approssimazione considerare stime differenti del “giorno 0” come provenienti dallo stesso metodo, sebbene entrambe derivate dall'analisi del muco cervicale; ciò nonostante nel seguito ricorreremo ai dati Billings come insieme di verifica del modello costruito sul database Fertili.

La Tabella 2.3 mostra il numero di cicli totali e quelli in cui sono presenti: entrambe le stime del “giorno 0”, solo il picco del muco, solo il rialzo termico e nessuno dei due. La colonna di destra, che fa riferimento al database Billings, manca dell'informazione sul rialzo termico (e quindi di entrambi gli indicatori) essendo è stato utilizzato solo il metodo del picco del muco per identificare l'ovulazione.



	Cicli ∈ Fertili (con concepimento)	Cicli ∈ Billings (con concepimento)
Totali	7594 (575)	2913 (177)
Picco identificato	6451 (523)	2282 (143)
Rialzo identificato	6098 (507)	–
Entrambi identificati	5622 (459)	–
Nessuno identificato	667 (4)	631 (34)

Tabella 2.3: Distribuzione dei cicli tra i due studi

## 2.3 La finestra di fertilità

Abbiamo già parlato della finestra di fertilità, ossia l'intervallo che contiene gli unici giorni ritenuti fertili: naturalmente tale finestra include il “giorno 0” stesso, o meglio la sua stima.

In entrambi gli studi Colombo e Masarotto (2000) e Colombo et al. (2006), si fa riferimento ad un intervallo che va da 8 date prima del “giorno 0” sino a 3 dopo, inclusi gli estremi. Questa scelta è giustificata dal fatto che essa rappresenta il minimo intervallo che consente di non avere cicli in cui l'ovulo è fecondato senza rapporti nella finestra<sup>1</sup>: per la variabilità descritta nel Paragrafo 1.3 (e talvolta forse anche per dimenticanze nel segnare i rapporti avuti), considerando intervalli più corti, si osservano cicli senza rapporti in essi ma con concepimento<sup>2</sup>. Naturalmente siffatta circostanza non è contemplata dal modello: l'obiettivo del limitare l'analisi alla finestra è proprio quello di focalizzarsi sulle sole date fertili.

Pertanto abbiamo deciso, almeno in prima istanza, di ricorrere alla medesima scelta dell'intervallo adottato da Colombo e Masarotto (2000) e Colombo et al. (2006), sebbene i giorni estremi (in particolare  $-8$ ,  $-7$  e  $3$ ) abbiano probabilità di concepimento notevolmente contenute.

## 2.4 Le diverse numerosità delle probabilità

Abbiamo visto nel Paragrafo 1.3 che la probabilità  $Pr(Z_i = 1|X_i)$  può essere composta dal prodotto di due componenti:  $Pr(Y_i = 1|X_i)$  e  $Pr(Z_i = 1|Y_i = 1, X_i)$ , dove le v.c.  $Z_i$  ed  $Y_i$  non sono sempre entrambe note; in questo elaborato le due probabilità saranno modellate tramite una forma parametrica

<sup>1</sup>In Fertili questo vale con ambedue i metodi di stima: rialzo termico e picco del muco

<sup>2</sup>Nei dati Billings è presente un concepimento con i rapporti più vicini al picco del muco nei giorni  $-10$  e  $6$ ; questo è l'unica eccezione presente nei due database

mostrata nel Capitolo 4. Affinché i cicli presenti nel database siano fruibili è necessario che si verifichino determinate condizioni.

Per stimare i parametri presenti in  $Pr(Y_i = 1|X_i)$  le uniche informazioni necessarie sono quelle relative ai rapporti nella finestra, contenute nel vettore  $X_i$ , e al concepimento, presente in  $Y_i$ .

Come mostrato nella Tabella 2.3, in alcuni cicli non è stato possibile identificare il “giorno 0”: a seconda del metodo di stima di tale data, sono stati considerati solo i cicli in cui il giorno dell’ovulazione era indicato.

Sono inoltre stati esclusi i dati in cui non si ha neppure un rapporto nel periodo di fertilità poiché non forniscono alcuna informazione né sulle probabilità di concepimento né su quelle di procreare un maschio poiché, non essendoci rapporti,  $Y_i = 0$  quasi certamente e  $Z_i$  non è osservabile.

In conclusione, per stimare i parametri di  $Pr(Y_i = 1|X_i)$  è possibile ricorrere alle numerosità espresse in Tabella 2.4, a seconda del database e del metodo di identificazione del “giorno 0” usato.

	Cicli ∈ Fertili (con concepimento)	Cicli ∈ Billings (con concepimento)
Picco identificato	3525 (523)	963 (142)
Rialzo identificato	3390 (507)	–

Tabella 2.4: Distribuzione dei soli cicli in cui si registra almeno un rapporto nella finestra  $[-8, 3]$  stimata coi metodi del Picco del muco e rialzo termico

Per quanto riguarda la stima dei parametri di  $Pr(Z_i = 1|Y_i = 1, X_i)$  sono richieste sia le stesse informazioni citate sopra sia il dato sul sesso del feto, riportato in  $Z_i$ . Quest’ultimo è disponibile solo per i concepimenti (vale a dire dove  $Y_i = 1$ ), che rappresentano una netta minoranza del totale. Vanno inoltre esclusi i casi in cui si osserva un aborto prima della rilevazione del sesso o in cui questo non è presente nel database (dati mancanti) e, come sopra, quelli in cui il “giorno 0” non è stato individuato.

Ricordiamo inoltre che, utilizzando la finestra di 12 giorni, non vi sono cicli in cui l’ovulo è fecondato senza rapporti nell’intervallo  $[-8, 3]$ ; perciò nessun dato è stato escluso dalla stima dei parametri di  $Pr(Z_i = 1|Y_i = 1, X_i)$  per tale motivazione.

Per le ragioni spiegate sopra, la numerosità effettivamente a disposizione per stimare i parametri inclusi nella probabilità di concepire un maschio è quindi limitata, utilizzando il picco del muco, a 412 in Fertili e a 112 in Billings (vedi Tabelle 2.5 e 2.6) e, ricorrendo al rialzo termico in Fertili a 409.

	M	F	Gemelli			Totale
			M, M	F, F	M, F	
Tutti i dati	238	211	1	3	1	454
Picco identificato	214	194	0	3	1	412
Rialzo identificato	216	188	1	3	1	409
Entrambi identificati	195	171	0	3	1	370

Tabella 2.5: Distribuzione dei sessi noti nei dati Fertili

	M	F	Totale
Tutti i dati	73	68	141
Picco identificato	54	59	113
Picco id. e rapporti in $[-8, 3]$	54	58	112

Tabella 2.6: Distribuzione dei sessi noti nei dati Billings, non sono presenti gemelli

Vista la maggiore numerosità campionaria del database Fertili, abbiamo deciso di utilizzare questi dati per le analisi effettuate nella tesi. Tra i due indicatori dell'ovulazione, è stato utilizzato il picco del muco per due ragioni. In primo luogo, sia i cicli totali ( $Y_i$  nota) sia quelli di cui si conosce il sesso ( $Z_i$  osservata) in cui il picco è identificato, sono leggermente più numerosi rispetto a quelli in cui il rialzo termico è individuato (vedi Tabelle 2.4 e 2.8). Inoltre ricorrere a questo metodo ci consente eventualmente di usare il data set Billings come insieme di verifica, per quanto non del tutto corretto, per ottenere valutazioni del modello indipendenti dai dati Fertili sui quali esso viene stimato; questo non sarebbe stato possibile utilizzando il rialzo termico come indicatore, poiché non presente in Billings. Abbiamo ritenuto opportuno utilizzare quest'ultimo database come insieme di verifica poiché, sebbene la classificazione del muco cervicale sia leggermente differente da quella usata in Fertili, il metodo di stima del picco è simile.

## 2.5 Le gerarchie

Nei dati esistono alcune forme di dipendenza che verranno illustrate in questo paragrafo.

### 2.5.1 Il fattore coppia

Ciascun record consiste in una serie di informazioni su un ciclo mestruale, ognuno dei quali appartiene naturalmente ad una donna.

Se a ciascuna coppia corrispondesse un solo ciclo, potremmo assumere l'indipendenza delle variabili tra i cicli, in quanto tale ipotesi risulterebbe piuttosto plausibile per le coppie. Solitamente però più cicli appartengono ad una donna e verosimilmente questi sono dipendenti tra loro. Una differenza evidente sta ad esempio nella maggiore o minore fertilità di alcune donne rispetto ad altre; lo stesso discorso vale per i relativi partner. Un'altra consiste nel modo in cui ciascuna di loro applica i criteri (comuni) di identificazione del "giorno 0": potrebbe esservi una distorsione sistematica di tale stima in alcune di esse.

La prima forma di dipendenza da considerare è pertanto identificata dal "fattore coppia".

Nel Capitolo 4 esporremo un modello che si basa sull'assunzione di indipendenza e medesima distribuzione delle variabili di interesse tra i cicli. Nel Capitolo 5 illustreremo una correzione di tale modello inserendo una gerarchia nelle probabilità di concepimento che tenga conto del "fattore coppia", mostreremo quindi se e quanto l'approssimazione effettuata nella prima analisi avesse effettivamente influenzato i risultati.

Relativamente alle probabilità di concepire un maschio non è invece stata considerata alcuna gerarchia poiché a quasi ogni coppia corrisponde un solo ciclo con sesso noto (vedi Tabella 2.7): nei dati disponiamo quindi di troppe poche ripetizioni per ciascuna di esse; questo è invece possibile nelle probabilità di concepimento poiché ad ogni donna appartengono mediamente 4.5 cicli con picco del muco identificato nei dati Fertili e almeno un rapporto nella finestra  $[-8, 3]$ .

Studio	Picco identificato	Cicli (con concepimento)	Donne	Cicli medi per Donna
Fertili	Tutti	3525 (523)	777	4.54
Fertili	Sesso noto	413	400	1.03
Billings	Tutti	963 (142)	178	5.41
Billings	Sesso noto	112	98	1.14

Tabella 2.7: Cicli totali e con sesso noto e numero di donne da cui provengono, cicli con picco del muco identificato e rapporti in  $[-8, 3]$ . Sono riportati 413 sessi e non 412 come in Tabella 2.5 perché abbiamo considerato le numerosità utilizzate dopo aver trattato i gemelli come esposto nel Paragrafo 2.6

### 2.5.2 I fattori centro ed istruttrice

I record di entrambi gli studi provengono da diversi istituti: in Fertili da 7 centri europei più Auckland (Nuova Zelanda), in Billings da 4 italiani. La differenza a questo livello però non è tanto nelle caratteristiche morfologiche delle coppie; possiamo infatti immaginare che non vi siano marcate influenze territoriali a riguardo. La peculiarità degli istituti risiede invece nel metodo con cui si raccolgono le informazioni sul muco cervicale e sulla temperatura basale e di conseguenza nella maniera di identificare il “giorno 0”. Naturalmente le linee guida dello studio per tale operazione sono le medesime, eppure è plausibile che in centri diversi siano presenti leggere divergenze relative all’applicazione di tali istruzioni.

Inoltre, anche in uno stesso centro possono esistere peculiarità attribuibili alle istruttrici che hanno insegnato alle donne questi metodi: spesso vi erano infatti più insegnanti per istituto.

Malgrado ciò i livelli di gerarchia descritti in questa sezione non sono stati considerati, essendo verosimilmente meno marcati del “fattore coppia”.

## 2.6 Gemelli ed aborti

Due casi che richiedono una differente trattazione sono rappresentati dai gemelli e dagli aborti che non permettono il conoscimento del sesso. Con questi ultimi ci si riferisce agli aborti avvenuti tra la rilevazione della gravidanza e l’osservazione del sesso del feto: quelli che avvengono ancor prima che il concepimento sia riscontrato, e dei quali la donna è ignara, sono trattati come dei non concepimenti.

Nel campione Fertili sono presenti 5 gemelli e 49 aborti (su un totale di 575 gravidanze); questi sono pari a 0 e 14 in Billings (su 177 gravidanze in totale).

Per quanto concerne i gemelli è necessario anzitutto fare una precisazione: nel nostro campione erano inclusi sia gemelli omo-zigote sia di-zigote; in natura i secondi sono circa 2 volte più frequenti dei primi<sup>3</sup>. I gemelli omo-zigote nascono dallo stesso ovulo e spermatozoo, presentano quindi il medesimo sesso (condividono infatti il 100% del patrimonio genetico); quelli di-zigote sono invece formati da due ovuli e due spermatozoi distinti e quindi i loro sessi sono totalmente indipendenti (come nel caso di due fratelli non gemelli hanno il 50% del DNA in comune). Questa distinzione è fondamentale ai fini del nostro studio.

---

<sup>3</sup>Wikipedia, Gemelli (biologia)

Nello stimare i parametri di  $Pr(Y_i = 1|X_i)$ , sia i gemelli sia gli aborti sono stati considerati esattamente come gli altri cicli che danno luogo ad un concepimento: l'interesse è solo nella fecondazione, non nella prosecuzione della gravidanza né nel sesso del nato.

Al contrario, nell'ottimizzazione dei parametri di  $Pr(Z_i = 1|Y_i = 1, X_i)$ , gli aborti sono stati esclusi poiché l'informazione sul sesso non è osservabile.

Relativamente ai gemelli, essendo il nostro obiettivo la previsione del sesso dei singoli bambini, in linea di principio quelli di-zigote andrebbero considerati entrambi poiché indipendenti e provenienti da due spermatozoi differenti, mentre di quelli omo-zigote andrebbe incluso un solo record, in quanto lo spermatozoo che ha dato origine ai due bambini è lo stesso ed i loro sessi sono dipendenti al 100%.

Nei dati a disposizione non possediamo l'informazione relativa alla tipologia di gemelli. Abbiamo considerato, nel calcolo dei parametri di  $Pr(Z_i = 1|Y_i = 1, X_i)$ , entrambi i record appartenenti ai gemelli con sessi differenti (1 in Fertili), poiché (per quanto detto sopra) necessariamente di-zigote. Tuttavia i gemelli di-zigote possono derivare da relazioni sessuali in giorni differenti ma probabilmente i rapporti che hanno portato alla fecondazione dei due ovuli coincidono o sono molto prossimi tra loro. Considerandoli entrambi dovremmo teoricamente includere una penalità nella funzione obiettivo per ogni giorno di distanza tra i due rapporti che hanno generato i concepimenti. Tuttavia, considerando che nei dati a disposizione vi è una sola coppia di gemelli con sessi diversi e quindi necessariamente di-zigote, abbiamo deciso di considerare entrambi i record senza imporre alcun vincolo o penalità.

Per ciò che concerne invece i gemelli con lo stesso sesso è stata calcolata (tramite il teorema di Bayes) la probabilità che ciascuno di essi fosse di-zigote<sup>4</sup>

$$Pr(\text{diz.} \mid \text{stesso sesso}) = \frac{Pr(\text{stesso sesso} \mid \text{diz.})Pr(\text{diz.})}{Pr(\text{stesso sesso} \mid \text{diz.})Pr(\text{diz.}) + Pr(\text{stesso sesso} \mid \text{omoz.})Pr(\text{omoz.})}. \quad (2.1)$$

Partendo dalle quote di omo-zigote rispetto ai di-zigote in natura, circa 1:2, abbiamo che  $Pr(\text{di-zigote})$  è  $2/3$ ; mentre  $Pr(\text{omo-zigote})$  è  $1/3$ . Inoltre per quanto chiarito sopra:

$$Pr(\text{stesso sesso} \mid \text{di-zigote}) = 1/2 \text{ e } Pr(\text{stesso sesso} \mid \text{omo-zigote}) = 1.$$

Sostituendo tali valori nella (2.1) otteniamo un risultato pari a  $1/2$ .

In virtù di quanto detto, due modi di trattare i gemelli consisterebbero nell'includere entrambi i record dei gemelli dello stesso sesso (inserendo mediamente 2 unità in più del dovuto, rispetto alle 4 dei dati Fertili) oppure nel

---

<sup>4</sup>Nella formula diz. indica di-zigote

selezionare solo uno dei due (escludendo in media 2 sessi in Fertili): abbiamo deciso di optare per la seconda soluzione.

A seguito di queste considerazioni la distribuzione dei sessi in Fertili esposta in Tabella 2.5, diviene (ai fini della trattazione del modello) quella in Tabella 2.8; quelle dei dati Billings rimangono invece invariate in quanto non includono gemelli.

	M	F	Totale
Tutti i dati	240	215	455
Picco identificato	215	198	413
Rialzo identificato	218	192	410
Entrambi identificati	196	175	371

Tabella 2.8: Distribuzione dei sessi noti nei dati Fertili, considerando i gemelli come descritto nel Paragrafo 2.6





## Capitolo 3

# La verosimiglianza composita

Il calcolo della verosimiglianza può talora essere computazionalmente proibitivo, a causa della complessa struttura del modello; altre volte, come nel nostro caso, la distribuzione congiunta dei dati non è nota e quindi la verosimiglianza congiunta non è definita (Pace, Salvan e Sartori, 2011; Molenberghs e Verbeke, 2005, 189-202).

In tutti questi casi l'utilizzo della verosimiglianza composita può rappresentare una valida soluzione: essa è una pseudo verosimiglianza che consiste combinare delle verosimiglianze basate su sottoinsiemi dei dati, opportunamente pesate.

$$L_C(\theta; y) = \prod_{k=1}^K L_k(\theta; y)^{w_k} = \prod_{k=1}^K \prod_{i=1}^n L_k(\theta; y_i)^{w_k}, w_k \geq 0, \quad (3.1)$$

dove  $L_k(\theta; y)$  rappresenta la componente  $k$ -esima della verosimiglianza composita e  $w_k$  il relativo peso. È necessario che ogni componente abbia almeno un parametro in comune con almeno un'altra componente: ciò è essenziale affinché  $L_C$  sia una verosimiglianza composita. Se non ci fossero quantità ignote in comune tra le parti, ottimizzare  $L_C$  equivarrebbe a massimizzare i singoli  $L_k$  separatamente, poiché i  $K$  insiemi di parametri sarebbero indipendenti tra loro.

Analogamente a quanto accade per la verosimiglianza, sotto le condizioni di regolarità (Molenberghs e Verbeke, 2005, 189-202), la massimizzazione di  $L_C$  avviene eguagliando a zero la derivata prima di  $l_C(\theta; y) = \log L_C(\theta; y)$ , ossia:

$$U_C(\theta) = \frac{\partial l_C(\theta; y)}{\partial \theta}. \quad (3.2)$$

L'equazione di stima risulta quindi essere

$$U_C(\theta) = \sum_{k=1}^K w_k \frac{\partial l_k(\theta; y)}{\partial \theta} = \sum_{k=1}^K \sum_{i=1}^n w_k \frac{\partial l_k(\theta; y_i)}{\partial \theta} = 0, \quad (3.3)$$

dove  $l_k(\theta; y) = \log L_k(\theta; y)$ . La soluzione della (3.3) è detta stima di massima verosimiglianza composita e viene indicata col simbolo  $\hat{\theta}_c$ .

La verosimiglianza composita non è una verosimiglianza nonostante sia la composizione di verosimiglianze proprie. Ciò nonostante, essendo ciascuna componente della (3.1) una densità condizionata o marginale ben definita, l'equazione di stima (3.3) risulta non distorta. Tuttavia le verosimiglianze  $L_k(\theta; y)$  sono moltiplicate tra loro: come conseguenza, che esse siano indipendenti o meno, l'inferenza che ne deriva ha le medesime proprietà di quella ottenuta da un modello non correttamente specificato. Come conseguenza la seconda identità di Bartlett non vale: è perciò necessario distinguere tra le matrici di sensibilità  $H(\theta)$  e di variabilità  $J(\theta)$ ,

$$H(\theta) = E_\theta \left( -\frac{\partial U_C(\theta)}{\partial \theta^T} \right) \text{ e } J(\theta) = E_\theta(U_C(\theta)U_C(\theta)^T) = \text{Var}_\theta(U_C(\theta)).$$

### 3.1 La matrice di Godambe

Attraverso la teoria delle equazioni di stima non distorte, si dimostra che la distribuzione asintotica di  $\hat{\theta}_c$  è normale di media  $\theta$  e varianza  $G(\theta)^{-1}$ , dove  $G(\theta)$  è l'informazione di Godambe

$$G(\theta) = H(\theta)J(\theta)^{-1}H(\theta).$$

Se  $L_C(\theta; y)$  fosse una vera verosimiglianza,  $H$  e  $J$  coinciderebbero e sarebbero quindi uguali alla matrice di informazione di Fisher. Essendo lo stimatore  $\hat{\theta}_c$  non distorto, la disuguaglianza di Cramér-Rao implica che  $G(\theta)^{-1} - I(\theta)^{-1}$  è semi-definita positiva; questa è invece strettamente definita positiva se  $\hat{\theta}_c$  non è funzione di una statistica sufficiente minimale; nel qual caso tale stimatore è sempre meno efficiente del corrispettivo di massima verosimiglianza ( $\hat{\theta}_{MV}$ ). Da un punto di vista pratico  $G(\theta)$  è solitamente calcolata in  $\hat{\theta}_c$ ; inoltre se il valore atteso rispetto a  $\theta$  è complesso da calcolare, si possono usare delle approssimazioni empiriche.

## 3.2 Proprietà asintotiche dello stimatore di massimo

In caso di campione indipendente ed identicamente distribuito da una v.c. in  $\mathbb{R}^m$ , e sotto le usuali condizioni di regolarità delle componenti della verosimiglianza, i seguenti risultati asintotici sono validi (Molenberghs e Verbeke, 2005, 189-202):

1. lo stimatore  $\hat{\theta}_c$  converge in probabilità al vero ed ignoto valore del parametro,

$$\hat{\theta}_c \xrightarrow{p} \theta;$$

2. vale inoltre la teoria asintotica sulle equazioni di stima non distorte, portando al risultato asintotico di normalità (convergenza in distribuzione) dello stimatore di massima verosimiglianza composita,

$$\sqrt{n}(\hat{\theta}_c - \theta) \xrightarrow{d} N_p \{0, G(\theta)^{-1}\}. \quad (3.4)$$

Tale risultato rimane valido se si calcola  $G(\theta)$  in  $\hat{\theta}_c$  o in un altro stimatore consistente di  $\theta$ .

## 3.3 I test d'ipotesi

In questo paragrafo saranno illustrate delle modifiche che consentono l'applicazione al campo della verosimiglianza composita dei classici test d'ipotesi per definire ipotesi nulle puntiformi contro ipotesi alternative bidirezionali del tipo

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases} \quad (3.5)$$

Per quanto riguarda il test di Wald si può usare la 3.4 mentre per la statistica punteggio basta considerare che  $Var_{\theta}(U_C(\theta)) = J(\theta)$ ; si arriva quindi alle usuali distribuzioni asintotiche  $\chi^2$ . Diversamente, il test del rapporto di verosimiglianza si presenta nella solita forma, ma con una differente distribuzione asintotica, vale a dire una somma pesata di  $\chi_1^2$  indipendenti (si veda ad esempio Varin et al., 2011). Esistono però numerose correzioni del test che consentono di fare riferimento, come per gli altri due, al  $\chi_q^2$  (Pace, Salvan e Sartori, 2011); nel seguito del capitolo sarà illustrata in dettaglio quella adottata.

Nel caso in cui si sia interessati solo ad un sottovettore  $\psi$  di  $\theta = (\psi, \tau)$ , di dimensione  $d$ , abbiamo a che fare con un test profilo, dove il sistema d'ipotesi

diventa

$$\begin{cases} H_0 : \psi = \psi_0 \\ H_1 : \psi \neq \psi_0 \end{cases} \quad (3.6)$$

Definiamo quindi le seguenti quantità che saranno utili nel prosieguo del paragrafo:

- $\hat{\theta}_{c\psi_0} = (\psi_0, \hat{\tau}_{c\psi_0})$ , la stima vincolata sotto  $H_0$ ;
- $H(\theta) = \begin{bmatrix} H_{\psi\psi} & H_{\psi\tau} \\ H_{\tau\psi} & H_{\tau\tau} \end{bmatrix}$  e  $H(\theta)^{-1} = \begin{bmatrix} H^{\psi\psi} & H^{\psi\tau} \\ H^{\tau\psi} & H^{\tau\tau} \end{bmatrix}$ ,

dove, ad esempio,  $H_{\psi\tau} = E_{\theta}(-\frac{\partial U_{C\psi}(\theta)}{\partial \tau^T})$ . Una notazione analoga vale per le matrici  $J(\theta)$  e  $G(\theta)$ .

### 3.3.1 Il test di Wald

Grazie alla normalità asintotica di  $\hat{\theta}_c$ , sotto  $H_0$  vale

$$cw^e(\theta) = (\hat{\theta}_c - \theta_0)^T G(\theta) (\hat{\theta}_c - \theta_0) \sim \chi_q^2.$$

La formulazione corrispondente al test d'ipotesi profilo (3.6) è invece

$$cw_P^e(\psi) = (\hat{\psi}_c - \psi_0)^T (G^{\psi\psi})^{-1} (\hat{\psi}_c - \psi_0) \sim \chi_d^2.$$

Come nel caso della classica verosimiglianza il test di Wald non è invariante rispetto a riparametrizzazioni; inoltre le regioni di confidenza che ne derivano presentano forma ellittica, vincolo stringente e non sempre adeguato rispetto alla forma della verosimiglianza composita.

### 3.3.2 Il test del punteggio

La sua applicazione si basa sulla normalità asintotica della funzione  $U_C(\theta)$  (Rotnitzky e Jewell, 1990). Il test ha la seguente forma

$$cw^u(\theta) = U_C(\theta)^T J(\theta)^{-1} U_C(\theta),$$

con distribuzione asintotica  $\chi_q^2$  sotto  $H_0$  (Geys et al., 1999).

Il medesimo test nel caso in cui il sistema d'ipotesi coinvolga solo una parte dei parametri, diviene

$$cw_P^u(\psi) = U_{C\psi}(\hat{\theta}_{c\psi_0})^T H^{\psi\psi}(\hat{\theta}_{c\psi_0}) \left\{ G^{\psi\psi}(\hat{\theta}_{c\psi_0}) \right\}^{-1} H^{\psi\psi}(\hat{\theta}_{c\psi_0}) U_{C\psi}(\hat{\theta}_{c\psi_0}) \sim \chi_d^2.$$

Tale statistica ha i vantaggi di poter essere ottenuta lavorando solo con il modello nullo e di essere invariante rispetto a riparametrizzazioni monotone; ciò nonostante potrebbe essere instabile numericamente nell'ambito della verosimiglianza composta e più in generale in quello delle equazioni di stima generalizzate (Rotnitzky e Jewell, 1990).

### 3.3.3 Il test del rapporto di verosimiglianza

Entrambi i test sopra descritti presentano evidenti limiti; risulta quindi più attraente il test del rapporto di verosimiglianza. La formulazione non varia rispetto alla sua applicazione nell'ambito classico della verosimiglianza. Senza e con parametri di disturbo, facenti rispettivamente riferimento ai sistemi d'ipotesi (3.5) e (3.6), il test è uguale a

$$cw(\theta) = -2\{l_C(\theta_0) - l_C(\hat{\theta}_c)\} \quad \text{e} \quad cw_P(\psi) = -2\{l_C(\hat{\theta}_{c\psi_0}) - l_C(\hat{\theta}_c)\}.$$

La differenza rispetto alla classica verosimiglianza risiede nella sua distribuzione asintotica nulla; per quanto riguarda  $cw(\theta)$ , essa è eguale a:  $\sum_{\alpha=1}^q \mu_\alpha(\theta) X_\alpha$ , dove  $X_1, \dots, X_q$  sono v.c. i.i.d.  $\chi_1^2$ , con  $\mu_1(\theta), \dots, \mu_q(\theta)$  autovalori di:  $J(\theta)H(\theta)^{-1} = H(\theta)G(\theta)^{-1}$  (Geys et al., 1999).

Per  $cw_P(\psi)$ ,  $q$  (la dimensione di  $\theta$ ) è sostituito da  $d$  (la lunghezza di  $\psi$ ) e  $\mu_1(\theta), \dots, \mu_q(\theta)$  da  $\nu_1(\theta), \dots, \nu_d(\theta)$ , autovalori di  $(H^{\psi\psi})^{-1}G^{\psi\psi}$ .

È possibile calcolare gli autovalori, nel primo caso in  $\theta_0$ , e nel secondo in  $\hat{\theta}_{c\psi_0}$ ; tuttavia, sebbene il massimo vincolato sarebbe più appropriato, i risultati asintotici sono ugualmente validi se viene utilizzato in entrambe  $\theta = \hat{\theta}_c$ , ossia la stima del modello non ristretto, in quanto stimatore consistente di  $\theta$ .

Per applicare il test è quindi necessario stimare, attraverso una simulazione, la sua distribuzione asintotica nulla; alternativamente è possibile applicare una correzione affinché questa sia ricondotta al classico  $\chi^2$ , con  $q$  (o  $d$ ) gradi di libertà.

#### La correzione del test

Sono state proposte una serie di correzioni. In particolare esponiamo quella presentata in Pace, Salvan e Sartori (2011) e che adotteremo nel seguito

$$cw(\theta)_{INV} = \frac{U_C(\theta)^T J(\theta)^{-1} U_C(\theta)}{U_C(\theta)^T H(\theta)^{-1} U_C(\theta)} cw(\theta) = \frac{cw^u(\theta)}{U_C(\theta)^T H(\theta)^{-1} U_C(\theta)} cw(\theta).$$

Essa risulta invariante e riporta il test  $cw(\theta)_{INV}$  ad avere distribuzione asintotica nulla  $\chi_q^2$ ; inoltre gli IC relativi presentano buona copertura. L'analogo in caso di parametri di disturbo  $\tau$ , con solo interesse nel sottoinsieme  $\psi$  di  $\theta$  diviene

$$cw_P(\psi)_{INV} = \frac{U_{C\psi}(\hat{\theta}_{c\psi_0})^T H^{\psi\psi}(\hat{\theta}_{c\psi_0}) \left\{ G^{\psi\psi}(\hat{\theta}_{c\psi_0}) \right\}^{-1} H^{\psi\psi}(\hat{\theta}_{c\psi_0}) U_{C\psi}(\hat{\theta}_{c\psi_0})}{U_{C\psi}(\hat{\theta}_{c\psi_0})^T H^{\psi\psi}(\hat{\theta}_{c\psi_0}) U_{C\psi}(\hat{\theta}_{c\psi_0})} cw_P(\psi) = \frac{cw_P^u(\psi)}{U_{C\psi}(\hat{\theta}_{c\psi_0})^T H^{\psi\psi}(\hat{\theta}_{c\psi_0}) U_{C\psi}(\hat{\theta}_{c\psi_0})} cw_P(\psi),$$

che ha distribuzione approssimata  $\chi_d^2$  sotto  $H_0$ .

Nel caso in cui il parametro di interesse  $\psi$  sia unidimensionale i test globale e profilo si semplificano rispettivamente nelle forme

$$cw(\theta)_{INV} = \frac{H(\theta)}{J(\theta)} cw(\theta) \quad \text{e} \quad cw_P(\psi)_{INV} = \frac{H^{\psi\psi}(\hat{\theta}_{c\psi_0})}{G^{\psi\psi}(\hat{\theta}_{c\psi_0})} cw_P(\psi).$$

### 3.4 Gli intervalli di confidenza

Illustriamo la struttura delle regioni di confidenza derivate dai test appena descritti; concentriamoci sui soli IC profilo unidimensionali poiché sono stati gli unici utilizzati nell'analisi corrente: definiamo  $\psi$  come una costante ignota.

L'IC basato sul test di Wald si fonda sul risultato (3.4); dal quale segue

$$IC_{1-\alpha}^{Wald}(\psi) = \left[ \hat{\psi} \pm z_{1-\alpha/2} \sqrt{G(\theta)_{\psi\psi}} \right],$$

dove  $z_{1-\alpha/2}$  indica il quantile di livello  $1 - \alpha/2$  della v.c.  $N(0, 1)$ .

L'IC derivato dal test del punteggio è invece ottenuto numericamente, cercando i valori soglia che definiscono il minor intervallo

$$IC_{1-\alpha}^{Punteggio}(\psi) = [\psi^{LB}, \psi^{UB}] \quad \text{t.c.} \quad cw_P^u(\psi) < \chi_{1,1-\alpha}^2 \forall \psi \in [\psi^{LB}, \psi^{UB}].$$

Quello relativo al rapporto di verosimiglianza sostituisce  $cw_P(\psi)$  a  $cw_P^u(\psi)$ .

# Capitolo 4

## Il modello proposto

Abbiamo già illustrato in (1.1) come la probabilità di procreare un maschio nell' $i$ -esimo ciclo,  $Pr(Z_i = 1|X_i)$ , possa essere scissa nel prodotto di due componenti:  $Pr(Y_i = 1|X_i)$  e  $Pr(Z_i = 1|Y_i = 1, X_i)$ ; nei prossimi paragrafi illustreremo le forme parametriche che abbiamo adottato per stimarle.

### 4.1 Le probabilità di concepimento

#### 4.1.1 Il modello di Barrett e Marshall

Esiste una letteratura molto vasta che tratta le probabilità giornaliere di concepire; uno dei primissimi modelli proposto da Barrett e Marshall nel 1969, contempla come unici parametri le probabilità giornaliere di fecondare in caso di rapporto nel solo giorno  $k$ , definite come  $\lambda_k$ .

Possiamo osservare solo l'*outcome* finale del ciclo  $Y_i$ : concepito o non concepito; nel primo caso non è possibile indicare con certezza in quale giorno sia il rapporto che ha fecondato l'ovulo. La probabilità modellata è quindi quella di concepire nel ciclo  $i$ -esimo dati i rapporti in esso presenti, cioè  $Pr(Y_i = 1|X_i)$ .

Consideriamo ora la probabilità dell'evento opposto, ossia di non concepire nel ciclo  $i$ -esimo:  $Pr(Y_i = 0|X_i) = 1 - Pr(Y_i = 1|X_i)$ . Un ovulo non è fecondato se, ogni giorno  $k$  in cui vi è almeno un rapporto ( $X_{ik} = 1$ ) non porta al concepimento, e la probabilità di questo evento è

$$(1 - \lambda_k)^{X_{ik}} = \begin{cases} 1 & \text{se } X_{ik} = 0, \text{ ossia se non vi sono rapporti} \\ 1 - \lambda_k & \text{se } X_{ik} = 1, \text{ ossia se vi sono rapporti} \end{cases}$$

Il modello assume che queste quantità siano indipendenti nei vari giorni del ciclo: conseguentemente la probabilità che non si osservi un concepimento

nel ciclo risulta pari al loro prodotto sui giorni  $k$  della finestra di fertilità.

Pertanto  $Pr(Y_i = 0|X_i) = \prod_{k=1}^K (1 - \lambda_k)^{X_{ik}}$ , da cui otteniamo che la probabilità di concepire nel ciclo  $i$ -esimo è pari a

$$Pr(Y_i = 1|X_i) = 1 - \prod_{k=1}^K (1 - \lambda_k)^{X_{ik}}. \quad (4.1)$$

Dove  $K$  rappresenta il numero di giorni della finestra, nella nostra analisi inizialmente fissato uguale a 12.

### 4.1.2 Il modello di Schwartz, MacDonald e Heuchel

Schwartz, MacDonald e Heuchel, propongono nel 1980 una modifica del modello di Barrett e Marshall che inserisce un parametro in più in  $Pr(Y_i = 1|X_i)$ : essi considerano il fatto che non tutti i cicli sono “vitali” (ossia fertili), talvolta infatti l’ovulo femminile non è fecondabile, e che in caso di concepimento seguito da un aborto precoce non sempre la gravidanza è rilevata. Tali fenomeni riguardano tutti i giorni dei suddetti cicli; Schwartz et al. li separano in una quantità comune ad ogni giorno:  $\omega$ . Questo parametro rappresenta quindi il prodotto tra la probabilità che un ciclo sia fertile e quella che, in caso di concepimento, la gravidanza prosegua senza aborto fino al momento della sua rilevazione. La formulazione da loro proposta, che nel seguito chiameremo SMH, è la seguente

$$Pr(Y_i = 1|X_i) = \omega \left\{ 1 - \prod_{k=1}^K (1 - \lambda_k)^{X_{ik}} \right\}. \quad (4.2)$$

In questo modello i parametri  $\lambda_k$  non hanno più lo stesso significato di quelli della (4.1); essi rappresentano ora la probabilità di concepire in un ciclo fertile, in caso di rapporto nel giorno  $k$ . Per poter confrontare i risultati ottenuti con quelli del modello di Barrett e Marshall è necessario moltiplicare le quantità giornaliere per  $\omega$ , pari a  $\omega\lambda_k$ , rifacendosi così alle probabilità di concepimento dato un rapporto nel giorno  $k$  senza condizionarsi alla vitalità del ciclo né al fatto che si osservi una gravidanza in atto.

### 4.1.3 Le due formulazioni a confronto

Entrambi i modelli si basano su due forti assunzioni: la probabilità di concepimento in un giorno è costante tra i cicli e i gruppi di spermatozoi provenienti



da diversi rapporti della finestra procedono indipendentemente nel processo di fertilizzazione dell'ovulo. I modelli (4.1) e (4.2) sono molto simili tra loro: la differenza risiede solo nel considerare come un parametro separato o meno, la probabilità che un ciclo sia fertile e che in caso di concepimento non vi sia un aborto prima del suo riscontro. È stato notato che spesso il modello di SMH non è in grado di discernere tra la quota comune ( $\omega$ ) e quella quotidiana ( $\lambda$ ); di conseguenza esso risulta debolmente identificabile ed instabile (Dunson e Stanford, 2005).

### I problemi riscontrati

Ottimizzando attraverso metodi numerici  $L_C$  ricorrendo in  $L_1$  alla (4.1) ed alla (4.2) con intervalli di ammissibilità dei parametri  $[0.0001, 0.9999]$ , abbiamo ottenuto rispettivamente tre e quattro stime sulla frontiera del supporto: in questo modo cadono le assunzioni alla base dei risultati asintotici mostrati nel Paragrafo 3.3.

Mentre i parametri stimati sulla frontiera utilizzando la (4.1) sono solo relativi ai giorni più esterni della finestra ( $-7$  e  $3$ ), con la (4.2) osserviamo stime pari ai limiti di ammissibilità anche nel giorno  $-2$ . Se nel primo caso è possibile pensare di restringere la finestra a  $[-6, 2]$  eliminando solo parametri corrispondenti a giorni a bassissima probabilità di concepire, col secondo modello non potremmo agire analogamente escludendo  $-2$  dall'analisi poiché un giorno centrale dell'intervallo di fertilità.

Con la formulazione di SMH abbiamo allora percorso una seconda via ottimizzando il modello rispetto a due trasformazioni: quella logit e l'inversa della funzione probit; ciò nonostante, in entrambe,  $\lambda_{-2}$  è risultato pressoché pari ad 1.

Abbiamo deciso di adottare la formulazione di Barrett e Marshall per il prosieguo delle analisi, considerato che è l'unica delle due a consentirci di ricorrere a test ed IC asintotici, sebbene solo sulla finestra ridotta rispetto a quella che ci eravamo prefissati di utilizzare.

## 4.2 Le probabilità dei sessi

Consideriamo  $Pr(Z_{ik} = 1 | Y_i = 1, X_i)$ , la probabilità di concepire un maschio nel giorno  $k$  del ciclo  $i$ -esimo dato che il suo ovulo è stato fecondato. Questo accade se e solo se, dato un concepimento nel ciclo  $i$ , questo avviene nel suo giorno  $k$  ed inoltre si procrea un maschio in tale data, ossia  $(Z_{ik} = 1 \cap Y_{ik} = 1) | (Y_i = 1, X_i)$ . La probabilità di tale evento può essere espressa

come prodotto di  $Pr(Z_{ik} = 1|Y_{ik} = 1)$ , ossia la probabilità di procreare un maschio nel giorno  $k$  dato che in esso si è concepito, e  $Pr(Y_{ik} = 1|Y_i = 1, X_i)$ , cioè la probabilità che, tra tutti i giorni della finestra, sia stato quello  $k$ -esimo a fecondare l'ovulo; quest'ultima è naturalmente solo positiva per i giorni  $k : X_{ik} = 1$ , e 0 altrimenti.

La probabilità  $Pr(Z_{ik} = 1|Y_{ik} = 1)$  rappresenta proprio la quantità di interesse nello studio corrente, che chiamiamo  $\beta_k$ .

Invece  $Pr(Y_{ik} = 1|Y_i = 1, X_i)$  è pari al rapporto tra la probabilità di concepire nel giorno  $k$ ,  $\lambda_k X_{ik}$ , e la somma di queste stesse quantità su tutti i giorni della finestra:  $\frac{\lambda_k X_{ik}}{\sum_{h=1}^K \lambda_h X_{ih}}$ , che definiamo  $\lambda_{ik}^*$ . A contribuire alla sommatoria sono solo le giornate in cui si registrano rapporti.

Abbiamo quindi che la  $Pr(Z_{ik} = 1|Y_i = 1, X_i) = \beta_k \frac{\lambda_k X_{ik}}{\sum_{h=1}^K \lambda_h X_{ih}} = \beta_k \lambda_{ik}^*$ . Questa rappresenta la probabilità di concepire un maschio nel giorno  $k$  del ciclo  $i$ -esimo. Quella invece relativa all'intero ciclo risulta pari alla probabilità dell'unione di questi eventi:  $Pr(Z_i = 1|Y_i = 1, X_i) = Pr(\bigcup_{k=1}^K Z_{ik} = 1|Y_i = 1, X_i)$ . Dato che la fecondazione avviene solo in un giorno, non è possibile concepire un maschio in più giornate e quindi gli eventi  $Z_{ik} = 1$ ,  $k = 1, \dots, K$ , sono incompatibili; quindi la probabilità dell'unione risulta eguale alla somma delle probabilità:

$$Pr(Z_i = 1|Y_i = 1, X_i) = \quad (4.3)$$

$$= Pr\left(\bigcup_{k=1}^K Z_{ik} = 1|Y_i = 1, X_i\right) \quad (4.4)$$

$$= \sum_{k=1}^K Pr(Z_{ik} = 1|Y_i = 1, X_i) \quad (4.5)$$

$$= \sum_{k=1}^K \beta_k \frac{\lambda_k X_{ik}}{\sum_{h=1}^K \lambda_h X_{ih}} \quad (4.6)$$

$$= \sum_{k=1}^K \beta_k \lambda_{ik}^* \quad (4.7)$$

Facciamo notare che nel modello di SMH i  $\lambda_k$  rappresentano quantità differenti da quelle indicate qui; per ricondurci ai parametri espressi nella (4.3), dovremmo moltiplicare i  $\lambda_k$  della formulazione (4.2) per  $\omega$ . Tutta-

via ciò non risulta necessario poiché  $\omega$ , essendo uguale per ciascun giorno, sarebbe presente sia al numeratore sia al denominatore di  $\lambda_{ik}^*$ , quindi si semplificherebbe.

### 4.3 Il modello completo

Dopo aver esplicitato  $Pr(Y_i = 1|X_i)$  e  $Pr(Z_i = 1|Y_i = 1, X_i)$  possiamo scrivere l'intera formulazione del modello adottato nelle analisi di questo capitolo.

Assumiamo una distribuzione di Bernoulli per la v.c.  $Z_i$  con probabilità di successo (ossia di concepire un maschio) pari a

$$p_i(\omega, \lambda, \beta) = p_{1_i}(\omega, \lambda)p_{2_i}(\lambda, \beta),$$

dove  $\lambda = (\lambda_1, \dots, \lambda_K)$ ,  $\beta = (\beta_1, \dots, \beta_K)$  e  $p_{1_i}(\omega, \lambda) = Pr(Y_i = 1|X_i)$ , espressa dalla (4.1) o (4.2), e  $p_{2_i}(\lambda, \beta) = Pr(Z_i = 1|Y_i = 1, X_i)$ , modellata nella (4.3).

Sotto l'ipotesi di cicli i.i.d. la verosimiglianza sarebbe pari a

$$L(\theta; z) = \prod_{i=1}^n p_i(\theta)^{Z_i} (1 - p_i(\theta))^{1-Z_i},$$

dove  $\theta = (\omega, \lambda, \beta)$ ; per ottenere la formulazione di Barrett e Marshall è sufficiente imporre  $\omega \equiv 1$ .  $L(\theta; z)$  risulta tuttavia impossibile da calcolare e perciò massimizzare: come già detto, il problema risiede nel fatto che mentre la v.c.  $Y_i$  è sempre nota,  $Z_i$  è osservabile solo per il sottoinsieme di cicli con concepimento.

Per risolvere tale inconveniente abbiamo deciso di ricorrere al metodo della verosimiglianza composita. Scindiamo la verosimiglianza globale in due componenti, ossia in due verosimiglianze, che presentano una parte dei parametri in comune.

Le due verosimiglianze in questione sono relative alle variabili  $Y_i$  e  $Z_i|Y_i = 1$  trattate ciascuna come una v.c. Bernoulli di parametri rispettivamente  $p_{1_i}(\omega, \lambda)$  e  $p_{2_i}(\lambda, \beta)$ ,

$$L_1(\omega, \lambda; y) = \prod_i^{n_1} p_{1_i}(\omega, \lambda)^{Y_i} (1 - p_{1_i}(\omega, \lambda))^{1-Y_i},$$

$$L_2(\lambda, \beta; z) = \prod_i^{n_2} p_{2_i}(\lambda, \beta)^{Z_i} (1 - p_{2_i}(\lambda, \beta))^{1-Z_i}.$$

La verosimiglianza composita è definita come il prodotto pesato (tramite due esponenti:  $a$  e  $b$ ) di tali verosimiglianze,

$$L_C(\theta; z, y) = L_1(\omega, \lambda; y)^a L_2(\lambda, \beta; z)^b. \quad (4.8)$$

Notiamo come  $\lambda$  compaia in entrambe le componenti: ciò è essenziale affinché  $L_C$  sia una verosimiglianza composita. Se non ci fossero quantità ignote in comune tra  $L_1$  ed  $L_2$ , massimizzare  $L_C$  equivarrebbe a massimizzare le due parti separatamente, poiché i parametri dell'una sarebbero indipendenti da quelli dell'altra.

L'approssimazione effettuata rende possibile il calcolo della pseudo verosimiglianza  $L_C(\theta; z, y)$  poiché ora sia  $L_1(\omega, \lambda; y)$  sia  $L_2(\lambda, \beta; z)$  sono quantificabili in quanto ricorrono a due insiemi di dati differenti. La prima utilizza i cicli di cui si conosce la  $Y_i$  ( $n_1$ ), ossia la fecondabilità; la seconda usa un sottinsieme di questi ultimi, rappresentato da quelli con concepimento (ossia con  $Y_i = 1$ ) e per i quali è noto il sesso del nato ( $n_2$ ).

Ricorrendo alla finestra di fertilità  $[-8, 3]$ , le numerosità a disposizione per l'ottimizzazione della verosimiglianza composita sono quindi  $n_1 = 3525$  in  $L_1$  ed  $n_2 = 413$  in  $L_2$  (vedi Tabelle 2.4 e 2.8).

Aniché ricorrere alla verosimiglianza composita avremmo potuto stimare i due parametri  $\lambda$  e  $\beta$  separatamente: il primo su  $L_1$ , ed il secondo su  $L_2$  con  $\lambda$  costante fissato pari al valore di massimo trovato su  $L_1$ . Tuttavia stimando congiuntamente i due insiemi di parametri in  $L_C$ , includiamo nel modello anche l'incertezza relativa alle probabilità di concepimento giornaliere: di conseguenza questa variabilità si ripercuote su  $\beta$ .

## 4.4 Il calcolo dei pesi $a$ e $b$

Dopo aver specificato interamente la verosimiglianza composita è stato necessario, prima di procedere alla stima dei parametri, scegliere i pesi  $a$  e  $b$  relativi alle componenti  $L_1$  ed  $L_2$ .

In base agli obiettivi prefissati, l'interesse della nostra analisi cade principalmente sui parametri  $\beta_k$ : abbiamo quindi ritenuto opportuno, almeno in prima istanza, pesare maggiormente la seconda parte della verosimiglianza composita che include tali quantità, poiché ci aspettavamo che così facendo le prestazioni sarebbero state migliori.

Tuttavia abbiamo deciso di non stabilire dei valori a priori, bensì di stimarli dai dati sia attraverso il metodo della convalida incrociata *leave-one-out* sia ricorrendo al database Billings come insieme di verifica: in entrambi i casi abbiamo fissato  $b = 1$  e misurato l'errore del modello su una griglia di valori

di  $a$  tra 0 (escluso) ed 1. Porre  $a = 1$  equivale ad attribuire la medesima importanza alle due parti, mentre  $a = 0$  corrisponderebbe all'esclusione dal modello di  $L_1$  e quindi al solo studio della verosimiglianza  $L_2$ , tale modello non sarebbe più una verosimiglianza composita né sarebbe ottimizzabile rispetto a  $\theta = (\lambda, \beta)$ .

#### 4.4.1 La convalida incrociata

Date le diverse numerosità delle due verosimiglianze abbiamo dovuto stabilire se, per effettuare la convalida incrociata, era opportuno levare uno alla volta dall'insieme di stima i cicli totali ( $n_1$ ) oppure solo quelli di cui è noto il sesso ( $n_2$ ). Considerato che per ogni iterazione del metodo si necessita la massimizzazione della (4.8), la quale coinvolge  $2K$  parametri (cioè 24, con una finestra di 12 giorni), e che l'interesse in  $L_1$  è solo indiretto, abbiamo limitato l'operazione agli  $n_2$  cicli che compaiono in  $L_2$ .

In particolare, fissata una coppia di valori  $(a, b) \equiv (a_0, 1)$ , ad ogni iterazione, il record selezionato (che indichiamo con  $x$ ), viene rimosso dall'insieme di stima di  $L_C(\theta; z_{-x}, y_{-x})$ ; le numerosità di  $L_1$  ed  $L_2$  divengono quindi rispettivamente  $n_1 - 1$  ed  $n_2 - 1$ .

Una volta ottimizzati i parametri del modello su questi dati, si calcola

$$\hat{p}_{2i} = \hat{Pr}(Z_i = 1 | Y_i = 1, X_i) = \sum_{k=1}^K \hat{\beta}_k \hat{\lambda}_{ik}^*.$$

A partire da questa quantità è ottenuto un indicatore dell'errore della previsione dei sessi sull' $i$ -esima osservazione, non usata per la stima

$$\hat{\epsilon}_i = |Z_i - \hat{z}_i|, \text{ con } \hat{z}_i = \begin{cases} 1 & \text{se } \hat{p}_{2i} \geq 0.5 \\ 0 & \text{se } \hat{p}_{2i} < 0.5 \end{cases}$$

Tale operazione viene ripetuta per ciascuna unità tra le  $n_2$  di cui è noto il sesso, infine si perviene ad una media dell'errore:  $\hat{\epsilon} = \sum_{i=1}^{n_2} \hat{\epsilon}_i / n_2$ .

Questo è stato calcolato solamente su  $\hat{Pr}(Z_i = 1 | Y_i = 1, X_i)$ , senza tenere in considerazione  $\hat{Pr}(Y_i = 1 | X_i)$ : l'interesse è infatti nell'ottenere la massima precisione per ciò che concerne la stima delle probabilità dei sessi; quelle di concepimento sono funzionali solo indirettamente, poiché rientrano nella formulazione di  $L_2$ .

Così facendo abbiamo ottenuto, per ciascuna coppia di pesi  $(a, b) \equiv (a_0, 1)$ , un indicatore dell'errore di stima del modello; è quindi stato possibile scegliere quello che ha presentato le migliori prestazioni.

Contrariamente a quanto ci saremmo attesi, abbiamo notato che  $\hat{\epsilon}$  cala al crescere di  $a$  fino a raggiungere il suo minimo in 1, il limite superiore della griglia utilizzata. Questo fenomeno pare sottolineare l'importanza della stima congiunta dei parametri  $\lambda$  e  $\beta$ , per la previsione del sesso del feto. Abbiamo perciò allargato tale intervallo sino a giungere a 2; i risultati sono visibili in Figura 4.1.

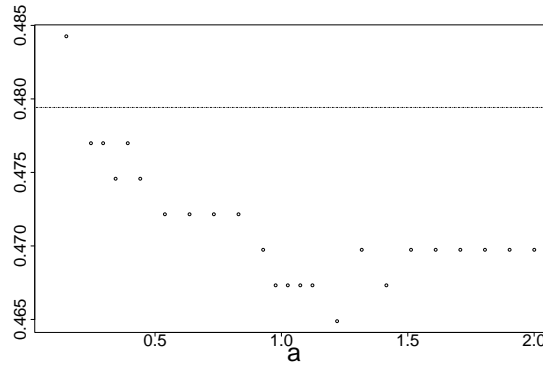


Figura 4.1: Errore di classificazione dei sessi al variare di  $a$ , stimato col metodo della convalida incrociata

Dall'analisi del grafico emerge come l'errore di classificazione cali al crescere di  $a$  tra 0 ed 1, fino a stabilizzarsi attorno a questa quantità, per poi oscillare superato tale valore e aumentare leggermente mentre  $a$  si avvicina a 2.

Facciamo notare come l'errore di previsione sia notevolmente elevato (in 1 esso è pari al 46.73%) se consideriamo che il “modello nullo”, il quale classifica tutti i sessi come maschi, essendo la moda del campione, presenta un tasso d'errata classificazione del 47.94%, evidenziato nel grafico da una linea tratteggiata orizzontale.

Calcolando inoltre, con  $a = 1$ , la  $\hat{Pr}(Z_i = 1|Y_i = 1, X_i)$  media per i cicli in cui  $Z_i = 1$  e per quelli in cui  $Z_i = 0$ , ossia la probabilità media di procreare un maschio separatamente per i cicli in cui è stato concepito un bambino e per quelli in cui si è concepita una bambina, otteniamo rispettivamente 51.94% contro 52.51%: il modello attribuisce quindi in media una probabilità leggermente maggiore di concepire un maschio ai cicli che hanno portato ad una femmina.

### 4.4.2 L'errore sui dati Billings

Abbiamo inoltre stimato la coppia che minimizzasse l'errore di stima dei sessi sull'insieme di dati Billings. Per ciascun valore di  $a$  tra 0 (escluso) e 2, abbiamo ottimizzato il modello, stimando i parametri sul database Fertili per poi misurare l'errore di previsione del sesso sui dati Billings. Questo si è reso possibile sebbene le classificazioni delle caratteristiche del muco cervicale siano leggermente differenti (vedi Paragrafo 2.1 e 2.2), in quanto entrambi i metodi di stima del "giorno 0" sono basati sul picco di tale muco, e per questo risultano simili.

L'errore misurato sulle 112 unità a sesso noto in Billings è stato costante sull'intera griglia di  $a$ . Già sui dati Fertili, con il metodo della convalida incrociata, avevamo notato come esso fosse piuttosto stabile; tuttavia si evolveva leggermente, soprattutto per valori del peso inferiori all'unità. Visto che nel precedente esempio, con diversi  $a$ , osservavamo pochi errori di stima in più o in meno su 413 record (il numero di sessi noti usufruibili in Fertili), era prevedibile una minore variabilità su soli 112 casi (i sessi noti utilizzabili in Billings), ma la totale stabilità su tutta la griglia ci ha sorpresi. Abbiamo quindi ampliato tale intervallo sino a 6 e siamo giunti ancora al medesimo numero di errori di classificazione in ogni punto. Sembra quindi che l'effetto di  $a$  sia estremamente contenuto, verosimilmente i  $\beta$  rimangono stabili a prescindere da quale sia il peso di  $L_2$  rispetto ad  $L_1$  (sempre nell'ambito di valori ragionevoli), mentre i  $\lambda$  sono comunque determinati principalmente da  $L_1$ , che ha quasi 9 volte la numerosità di  $L_2$ : le stime finali non paiono quindi risentire dell'importanza attribuita all'una o all'altra componente.

Un altro dato notevole è stato il valore dell'errore di classificazione, 49.11%, prossimo a 0.5 e persino maggiore di quello che avremmo ottenuto classificando tutti i nati come maschi, 48.21%.


Se consideriamo, con  $a = 1$ , la media delle  $\hat{Pr}(Z_i = 1|Y_i = 1, X_i)$ , calcolata separatamente per i cicli nei quali sono stati concepiti maschi e per quelli che hanno portato a femmine sui dati Billings, otteniamo un valore leggermente maggiore per i secondi, 0.5068 contro 0.4844, al contrario di quanto sarebbe logico; queste indicazioni confermano quanto già visto con la convalida incrociata sul campione Fertili. Tale fenomeno sembrerebbe indicare l'incapacità del modello nel prevedere il sesso del nato basandosi sui soli giorni dei rapporti.

Data la stabilità al variare di  $a$  dell'errore di classificazione misurato sui dati Billings, ci siamo basati sui soli risultati ottenuti dalla convalida incrociata sul database Fertili per stabilire il valore di tale peso. Nonostante nella Figura 4.1 il minimo sia raggiunto in 1.2, in 1.3 l'errore è nuovamente salito ed è maggiore di quanto fosse in 1. Considerando come questo indice sembri

variare attorno ad 1.2, mentre pare meno instabile in 1 (l'errore assume lo stesso valore anche in 0.95, 1.05 ed 1.1), nonché il fatto che sarebbe in linea di principio più logico attribuire maggiore importanza a  $L_2$  piuttosto che a  $L_1$ , abbiamo deciso di porre  $a = 1$ , pesando ugualmente le due componenti della verosimiglianza composita, sebbene esso non sia esattamente il punto di minimo del grafico. Inoltre tra  $a = 1$  ed  $a = 1.2$  esiste una differenza minima, pari ad una sola errata classificazione su 413.

## 4.5 Le stime puntuali dei parametri

Ottimizzando la formulazione (4.8), ricorrendo al modello di Barrett e Marshall in  $Pr(Y_i = 1|X_i)$ , abbiamo ottenuto i risultati in Tabella 4.1. In questa fase abbiamo deciso di riportare anche le stime ottenute utilizzando la formulazione di SMH in  $L_1$ , osservabili anch'esse in Tabella 4.1. Sono inoltre riportate le derivate prime di  $l_C$  rispetto a ciascun parametro calcolate in  $\hat{\theta}_c$ . La Figura 4.2 mostra graficamente le probabilità di concepire e quelle di procreare un maschio in caso di concepimento nel giorno  $k$ -esimo.

L'algoritmo utilizzato è *nminb* del pacchetto *numDeriv* (Gilbert P, 2012) del software statistico  (R Core Team, 2012); entrambe le ottimizzazioni hanno raggiunto la convergenza.

Dall'analisi della Tabella 4.1 e della Figura 4.2, notiamo come le stime delle probabilità giornaliere di concepimento utilizzando l'una o l'altra formulazione differiscano per ordine di grandezza. Le due quasi coincidono per quanto concerne gli estremi dell'intervallo, con probabilità inferiori all'1% (indici  $-8$ ,  $-7$  e  $3$ ); invece al centro della finestra la differenza si fa più marcata: le stime derivate ricorrendo alla formula di SMH sono maggiori di quelle ottenute col modello di Barrett e Marshall, sebbene entrambe siano caratterizzate dallo stesso andamento. Probabilmente è anche questa similitudine a far sì che tale differenza non si ripercuota sulle stime di  $\beta$ : esse risultano quasi sovrapposte in ogni giorno dell'intervallo. In  $L_2$  compaiono infatti le probabilità quotidiane di concepire standardizzate, ossia come rapporto tra le probabilità giornaliere di fecondare l'ovulo e la somma delle stesse probabilità su tutti i giorni in cui vi sono rapporti:  $\lambda^*$  della (4.3); quindi idealmente se i due modelli fornissero vettori delle probabilità di concepire giornaliere identici a meno di una costante, il loro contributo in  $L_2$  sarebbe analogo.

Rileviamo come vi siano dei parametri relativi alle probabilità giornaliere di procreare un maschio che spiccano rispetto agli altri: nei giorni  $-7$  e  $3$ ,  $\beta$  assume valore 0.0001, pari al limite inferiore dell'intervallo di ammissibilità fissato nell'ottimizzazione; ciò indica che in quei giorni il modello attribuisce probabilità quasi nulla di concepire un maschio. Tali valori sono però da



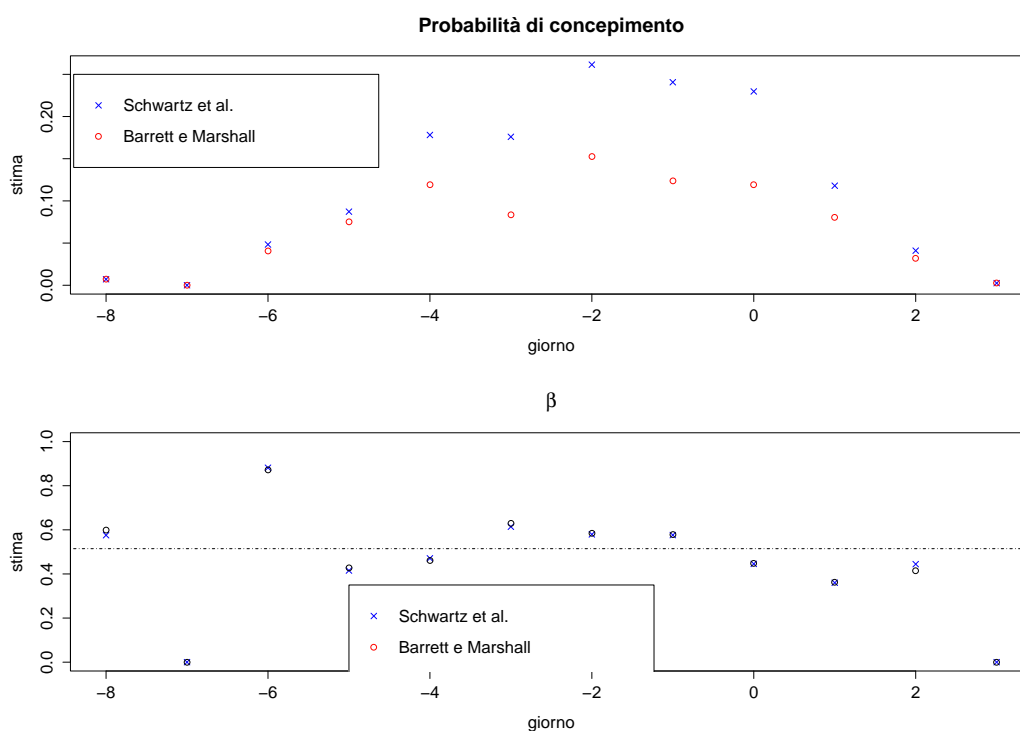


Figura 4.2: Stime delle probabilità giornaliere di concepimento (alto),  $\omega\lambda$  per SMH e  $\lambda$  per Barrett e Marshall, e  $\beta$  (basso)

considerarsi congiuntamente alle probabilità di concepimento stimate nelle stesse giornate: 0.0001 e 0.0026 con la formulazione (4.1) e 0 e 0.0025 con la (4.2). Considerato che i dati su cui si basano  $\hat{\beta}_{-7}$  e  $\hat{\beta}_3$  sono pressoché nulli, visto che il modello considera quasi impossibile concepire in quelle date, ci aspettiamo che nel seguito la variabilità di tali stime risulterà essere estremamente elevata.

Due risultati che paiono essere maggiormente interessanti sono relativi ai giorni  $-6$  e  $1$ , in cui troviamo i due punti (esclusi i precedenti) più distanti dalla linea tratteggiata indicante il rapporto atteso  $106/206$ . Questi valori sembrerebbero più rilevanti dei precedenti per il fatto che le relative probabilità di concepire stimate sono pari al  $4\%$  e  $8\%$  ricorrendo al modello di Barrett e Marshall e all' $8\%$  e  $12\%$  con quello di SMH. Naturalmente saranno i test e gli intervalli di confidenza applicati nel seguito a stabilire se tali diversità sono significative o attribuibili al caso; per ora ci limitiamo solamente ad indicare queste due giornate come quelle che sembrano distinguersi maggiormente dal tasso nullo di mascolinità alla nascita.

	SMH		Barrett e Marshall	
	Stima	Derivata	Stima	Derivata
$\lambda_{-8}$	0.0272	-0.0003	0.0073	0.0038
$\lambda_{-7}$	0.0001	-90.4891	0.0001	-416.1636
$\lambda_{-6}$	0.1847	0.0002	0.0407	0.0039
$\lambda_{-5}$	0.3337	-0.0000	0.0752	0.0011
$\lambda_{-4}$	0.6812	0.0000	0.1193	0.0013
$\lambda_{-3}$	0.6728	-0.0000	0.0834	-0.0028
$\lambda_{-2}$	0.9999	4.0184	0.1525	-0.0004
$\lambda_{-1}$	0.9203	0.0002	0.1239	-0.0008
$\lambda_0$	0.8785	-0.0001	0.1194	0.0014
$\lambda_1$	0.4512	-0.0002	0.0804	0.0007
$\lambda_2$	0.1569	-0.0003	0.0320	0.0029
$\lambda_3$	0.0097	-0.0011	0.0026	-0.0061
$\beta_{-8}$	0.5755	0.0000	0.5988	-0.0001
$\beta_{-7}$	0.0001	-0.0007	0.0001	-0.0020
$\beta_{-6}$	0.8815	0.0002	0.8725	0.0000
$\beta_{-5}$	0.4153	0.0000	0.4278	0.0002
$\beta_{-4}$	0.4709	-0.0000	0.4618	-0.0001
$\beta_{-3}$	0.6140	0.0002	0.6289	-0.0000
$\beta_{-2}$	0.5792	-0.0003	0.5842	-0.0001
$\beta_{-1}$	0.5762	0.0002	0.5781	0.0006
$\beta_0$	0.4456	0.0005	0.4477	-0.0001
$\beta_1$	0.3601	-0.0000	0.3621	0.0001
$\beta_2$	0.4445	-0.0000	0.4146	0.0000
$\beta_3$	0.0001	-0.2314	0.0001	-0.2366
$\omega$	0.2615	0.0006	—	—
$\omega\lambda_{-8}$	0.0071	—	—	—
$\omega\lambda_{-7}$	0.0000	—	—	—
$\omega\lambda_{-6}$	0.0483	—	—	—
$\omega\lambda_{-5}$	0.0872	—	—	—
$\omega\lambda_{-4}$	0.1781	—	—	—
$\omega\lambda_{-3}$	0.1759	—	—	—
$\omega\lambda_{-2}$	0.2615	—	—	—
$\omega\lambda_{-1}$	0.2406	—	—	—
$\omega\lambda_0$	0.2297	—	—	—
$\omega\lambda_1$	0.1180	—	—	—
$\omega\lambda_2$	0.0410	—	—	—
$\omega\lambda_3$	0.0025	—	—	—

Tabella 4.1:  $\hat{\theta}_c$  e derivate di  $l_C$  rispetto ai singoli parametri calcolate in  $\hat{\theta}_c$

In Figura 4.3 sono visibili le log-verosimiglianze profilo dei  $\beta_k$ , ottenute adottando la formulazione di Barrett e Marshall in  $L_1$ . Nei parametri relativi ai giorni tra  $-6$  e  $2$  è anche riportata la linea tratteggiata corrispondente agli  $IC_{0.95}$  profilo unidimensionali, che saranno ottenuti nel Paragrafo 4.7 tramite il metodo del RV. In appendice, in Figura A.1, sono riportati i grafici relativi alle log-verosimiglianze profilo dei parametri  $\lambda_k$ , anch'essi ricorrendo alla formulazione di Barrett e Marshall in  $L_1$ .

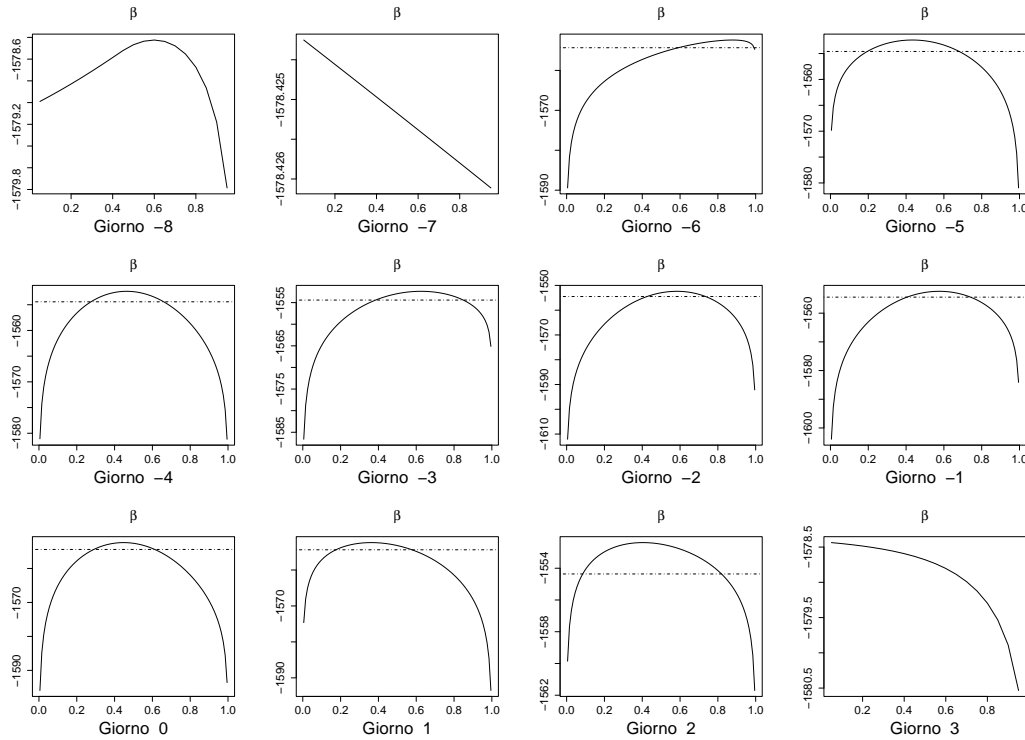


Figura 4.3: Log-verosimiglianze profilo di  $\beta_k$  con la formulazione di Barrett e Marshall in  $L_1$ , quelle relative ai giorni  $-8$ ,  $-7$  e  $3$  calcolate sulle finestra  $[-8, 3]$ , le altre sulla finestra  $[-6, 2]$ . Ove presenti, le linee tratteggiate indicano l' $IC_{0.95}$  profilo unidimensionale ottenuto col TRV

Le ipotesi alla base dei risultati asintotici dei test esposti nel Capitolo 3 includono il fatto che la stima del parametro non si trovi sulla frontiera del supporto (Molenberghs e Verbeke, 2005, 189-202): nel nostro caso questa situazione non è rispettata poiché  $\hat{\lambda}_{-7}$ ,  $\hat{\beta}_{-7}$  e  $\hat{\beta}_3$  eguagliano il limite inferiore dell'intervallo di ammissibilità fissato (0.0001) e, ricorrendo alla formulazione di SMH,  $\hat{\lambda}_{-2}$  è pari a quello superiore (0.9999).

Conseguentemente i test così come sono stati descritti nel Paragrafo 3.3 non sono applicabili, per cui abbiamo percorso due strade parallele.

Da un lato abbiamo adottato un approccio di simulazione tramite bootstrap parametrico per stimare la distribuzione nulla del test del rapporto di verosimiglianza (TRV). Abbiamo scelto tale test poiché, il fatto che la stima di alcuni parametri sia sulla frontiera fa sì che la matrice  $H$  (vedi Capitolo 3) calcolata in  $\hat{\theta}_c$  non sia definita positiva: il TRV è l'unico dei tre illustrati che non contempla  $H$  nella sua formulazione.

Dall'altro lato, abbiamo ristretto la finestra di fertilità ai giorni  $[-6, 2]$ , escludendo le tre date (con la formulazione di Barrett e Marshall in  $L_1$ ) con parametri stimati sulla frontiera del supporto, che peraltro presentavano probabilità stimate di concepire minori (inferiori all'1%), ed applicato su quest'intervallo ridotto i test illustrati nel Paragrafo 3.3.

D'ora in avanti per riferirsi all'ottimo vincolato sotto  $H_0$  utilizzeremo la notazione  $\hat{\theta}_{H_0}$ , e non  $\hat{\theta}_0$ , essendo quest'ultima già utilizzata per indicare la stima dei parametri relativi al giorno 0.

In entrambi i contesti saranno prima illustrati i test profilo globali atti a saggiare il seguente sistema d'ipotesi

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \dots = \beta_K \\ H_1 : \text{almeno un'uguaglianza non verificata} \end{cases} \quad (4.9)$$

Questi saranno seguiti dai test profilo individuali sui singoli  $\beta_k$ , per ogni giorno  $k$  della finestra, verificando se questi differiscono significativamente dal tasso nullo di mascolinità alla nascita 106/206 (CIA, The World Factbook) ricorrendo al sistema d'ipotesi

$$\begin{cases} H_0 : \beta_k = 106/206 \\ H_1 : \beta_k \neq 106/206 \end{cases} \quad (4.10)$$

Per i motivi spiegati nel Paragrafo 4.1.3, per le analisi che seguiranno faremo solo riferimento alla verosimiglianza composita che ricorre alla formulazione di Barrett e Marshall in  $L_1$ .

## 4.6 I Test bootstrap sulla finestra intera

Per ottenere la distribuzione nulla del TRV sotto  $H_0$  e calcolare il livello di significatività del valore osservato sui dati originali sulla finestra  $[-8, 3]$ , abbiamo ricorso alla tecnica del bootstrap parametrico.

Una prima fonte di errore del metodo riguarda il fatto che la stima della distribuzione bootstrap si basa su un numero limitato di campioni bootstrap

$B$ ; tuttavia quest'errore può essere valutato attraverso degli IC dei valori  $p$  stimati nonché limitato ricorrendo, all'occorrenza, ad un elevato numero di campioni simulati. Nella nostra analisi non è stato possibile innalzare eccessivamente  $B$  a causa dell'elevato peso computazionale della procedura, la quale prevede due ottimizzazioni numeriche ad ogni iterazione, e del fatto che è stata ripetuta 13 volte (una per il test globale e 12 per quelli individuali). Le numerosità scelte sono quindi state pari a 1000 per il test globale e 700 per quelli individuali; successivamente abbiamo innalzato tale valore sino a 4000 per il calcolo della distribuzione nulla del test su  $\beta_{-6}$ , unico parametro ad essersi avvicinato alla significatività con la precedente numerosità.

Una seconda fonte di errore è relativa ai parametri della distribuzione nulla da cui si simulano i campioni bootstrap: sia la (4.9) sia la (4.10) sono sistemi d'ipotesi composti; non si fissano quindi tutte le quantità ignote, come conseguenza i campioni bootstrap non sono simulati dal vero  $\theta_{H_0}$ . L'insieme di parametri sotto  $H_0$  da cui si simula è quindi in parte fissato, i  $\beta$  vincolati sotto  $H_0$ , ed in parte stimato, i  $\lambda$  ed i  $\beta$  non vincolati sotto  $H_0$  (nel caso di test su un singolo parametro): è perciò necessario massimizzare  $L_C$  rispettando le ipotesi nulle dei sistemi (4.9) e (4.10), ottenendo il massimo vincolato  $\hat{\theta}_{H_0}$ ; da questo vengono generati i campioni bootstrap.

Il processo di simulazione si articola nelle seguenti fasi: all'inizio si stima una sola volta e sotto  $H_0$ ,

$$Pr(Y_i = 1|X_i, \hat{\lambda}_{H_0}) = \hat{p}_{1i} = 1 - \prod_{k=1}^K (1 - \hat{\lambda}_{H_0k})^{X_{ik}}.$$

Ad ogni iterazione  $b = 1, \dots, B$  si ricrea l'insieme di dati, lasciando invariate le  $X_i$  e simulando, prima le  $Y_i$ , e poi le  $Z_i$  solo dove è stato generato un concepimento. Si estrae ciascun record  $Y_i^b$ ,  $i = 1, \dots, n_1$ , da una Bernoulli di parametro  $\hat{p}_{1i}$ ; solo per i cicli  $i : Y_i^b = 1$  si simulano i sessi  $Z_i^b$ , anch'essi estratti da v.c. Bernoulli, stavolta di parametri

$$Pr(Z_i^b = 1|X_i, Y_i^b = 1, \hat{\lambda}_{H_0}, \hat{\beta}_{H_0}) = \hat{p}_{2i} = \sum_{k=1}^K \hat{\beta}_{H_0k} \frac{\hat{\lambda}_{H_0k} X_{ik}}{\sum_{h=1}^K \hat{\lambda}_{H_0h} X_{ih}}.$$

Il processo di generazione del campione bootstrap non termina però qui: è infatti necessaria un'ulteriore accortezza. Nei dati a disposizione si osserva il sesso solo di 413 concepimenti su un totale di 523. Se includessimo in  $Z^b$  tutti i record in cui si simula un concepimento, i campioni bootstrap dei sessi noti sarebbero mediamente più numerosi di quello osservato: si modificherebbe quindi la struttura del database originale.

Per emulare il processo generatore dei dati è stato quindi necessario includere in  $Z^b$  ciascun sesso  $Z_i^b$  con probabilità pari a  $413/523 = 79\%$ .

A questo punto, sul campione  $b$ -esimo appena creato si ottimizza  $L_C$  sotto le ipotesi nulla ed alternativa e, a partire dai valori di massimo trovati nelle due ipotesi, viene calcolato il TRV, che chiamiamo  $test_b$ . Dopo  $B$  iterazioni otteniamo una quantità di valori del test che ci consente di approssimare la sua distribuzione nulla ed ottenere una stima della significatività osservata:

$$\hat{\alpha}^{oss} = \hat{Pr}(TRV_{H_0} \geq t^{oss}) = \frac{1}{B} \sum_{b=1}^B I(test_b \geq t^{oss}).$$

Per fornire una misura dell'errore di simulazione dovuto al limitato numero di campioni  $B$ , abbiamo associato ad ogni livello di significatività stimato un relativo  $IC_{0.95}$ , calcolato in modo esatto basandoci sulla distribuzione binomiale.

### 4.6.1 I risultati dei test

Dall'analisi delle quantità in Tabella 4.2 e delle Figure 4.4 e 4.5 viene confermata la nostra intuizione relativa ai parametri che avrebbero potuto risultare significativamente differenti dagli altri.

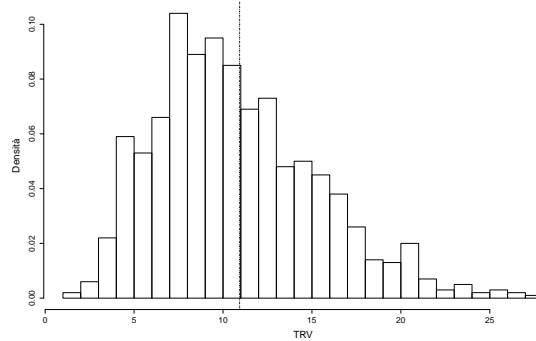


Figura 4.4: Distribuzione bootstrap sotto  $H_0$  del TRV profilo globale; la linea verticale tratteggiata indica il valore osservato del test

I valori osservati dei test profilo individuali sono quasi tutti compatibili con la rispettiva distribuzione nulla bootstrap compresi, come avevamo già intuito, quelli delle date  $-7$  e  $3$ , sebbene presentino una stima di  $\beta$  pari a  $0.0001$ .

L'unico test a spiccare rispetto alla sua distribuzione nulla è quello relativo al giorno  $-6$  Il suo valore  $p$  stimato è significativo al livello del  $5\%$  e pari a

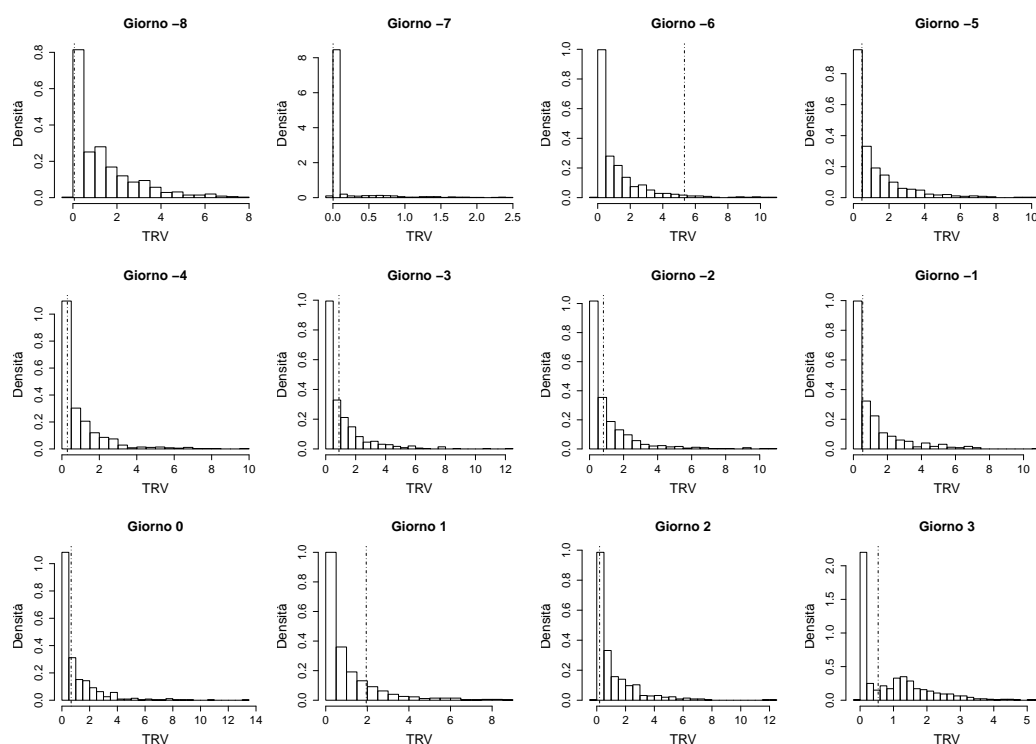


Figura 4.5: Distribuzioni bootstrap sotto  $H_0$  dei TRV profilo individuali; le linee verticali tratteggiate indicano i valori osservati dei test

0.03. Il secondo valore  $p$  (in ordine crescente) è relativo al giorno 1 (24 ore dopo il picco del muco), nonostante sia ben lontano dall'essere significativo essendo pari a 0.16.

Per quanto riguarda il test profilo globale su tutti i  $\beta$ , anch'esso non è significativo: il valore osservato, pari al percentile di livello 0.42, è solo di poco maggiore della mediana della distribuzione nulla stimata. Quest'informazione rispecchia le indicazioni ricevute dai test individuali.

Facciamo notare che nei test profilo individuali alcuni valori della distribuzione simulata sotto  $H_0$  sono negativi, sebbene la verosimiglianza (composita) del modello non vincolato (sotto  $H_1$ ) è necessariamente maggiore o uguale di quella del modello vincolato (sotto  $H_0$ ), trattandosi di modelli annidati. Il fenomeno osservato è dovuto al fatto che, essendo l'ottimizzazione numerica e non analitica, quando il valore dell'ottimo del  $\beta_k$  vincolato è particolarmente vicino a 106/206 (il che è plausibile visto che i campioni bootstrap sono simulati imponendo tale vincolo) il calcolatore potrebbe individuare un valore della log-verosimiglianza (composita) leggermente maggiore sotto  $H_0$ .

Test	Significatività osservata	Campioni bootstrap	$IC_{0.95}$	
			LB	UB
Globale	0.424	1000	0.393	0.455
$\beta_{-8}$	0.790	700	0.758	0.820
$\beta_{-7}$	0.900	700	0.875	0.921
$\beta_{-6}$	0.030	4000	0.025	0.036
$\beta_{-5}$	0.531	700	0.494	0.569
$\beta_{-4}$	0.570	700	0.532	0.607
$\beta_{-3}$	0.356	700	0.320	0.392
$\beta_{-2}$	0.377	700	0.341	0.414
$\beta_{-1}$	0.474	700	0.437	0.512
$\beta_0$	0.393	700	0.356	0.430
$\beta_1$	0.163	700	0.136	0.192
$\beta_2$	0.677	700	0.641	0.712
$\beta_3$	0.491	700	0.454	0.529

Tabella 4.2: Significatività del TRV stimate tramite bootstrap e rispettivi  $IC_{0.95}$ , sulla finestra di 12 giorni con  $a = 1$

Tali anomalie dovute alla massimizzazione numerica sono tuttavia rare e di minima entità: solo 11 valori del TRV su 8400 simulati sono negativi e la distanza massima dallo 0 è di 1 millesimo.

### 4.6.2 L'analisi di sensitività

Come illustrato nel Paragrafo 4.4, il peso  $a$  della prima componente della verosimiglianza composita è stato scelto attraverso il metodo della convalida incrociata come quel valore che ottimizza la previsione dei sessi calcolata senza restrizioni sui parametri (cioè sotto  $H_1$ ).

In questo modo i pesi di  $L_1$  ed  $L_2$  sono tali da preferire il modello non vincolato, per la sola ragione di essere stati scelti per migliorarne le prestazioni.

L'influenza di  $a$  risiede nel fatto che i campioni bootstrap sono simulati da  $\hat{\theta}_{H_0}$  e non da  $\theta_{H_0}$ , dove  $\hat{\theta}_{H_0}$  è ottenuto massimizzando  $L_C$ , sotto  $H_0$ , con  $a$  fissato pari alla quantità che minimizza<sup>1</sup> l'errore di previsione del sesso sotto  $H_1$  (ossia  $a = 1$ ).

In linea di principio dovremmo perciò inserire la scelta del peso tramite convalida incrociata nel metodo bootstrap, fissando  $a$  in ogni campione si-

<sup>1</sup>In realtà l'errore è minimo in 1.2, ma  $a = 1$  rappresenta quasi il minimo dell'errore



mulato come quel valore che migliora la stima non vincolata (ossia sotto  $H_1$ ) di  $Z_i^b$ . Tuttavia, data l'elevata mole computazionale di ciascuna iterazione del metodo bootstrap, ognuna delle quali consta di due ottimizzazioni di  $L_C$  (una vincolata, sotto  $H_0$ , e una non, sotto  $H_1$ ), non sarebbe stato fattibile inserire anche la scelta del peso di  $L_1$  con convalida incrociata, essendo tale procedimento estremamente dispendioso.

Per verificare se questa imprecisione possa aver influenzato i risultati abbiamo applicato una analisi di sensitività, calcolando il valore osservato del test e stimando la sua distribuzione nulla tramite il metodo bootstrap con  $(a, b)$  fissati indipendentemente dai dati utilizzati per la stima: abbiamo scelto a priori  $(a, b) \equiv (1/2, 1)$ , in modo da pesare maggiormente, ma non troppo,  $L_2$  rispetto a  $L_1$ .

Vista l'elevata complessità computazionale del metodo bootstrap, abbiamo deciso di non applicarlo su tutti i sistemi d'ipotesi, ma di riservare tale operazione al solo test globale e a quello individuale su  $\beta_{-6}$ , con un numero di campioni rispettivamente pari a 1000 e 2000.

Osservando i risultati in Tabella 4.3 possiamo notare come questi siano particolarmente vicini a quelli ottenuti fissando  $a = 1$ : il test globale presenta significatività analoga; mentre quello su  $\beta_{-6}$  è anch'esso prossimo al caso precedente, sebbene l' $IC_{0.95}$  del valore  $p$  sia leggermente più ampio a causa del ridotto numero di campioni bootstrap (vedi Tabella 4.2).

Test	Significatività osservata	Campioni bootstrap	$IC_{0.95}$	
			LB	UB
Globale	0.443	1000	0.412	0.474
$\beta_{-6}$	0.031	2000	0.023	0.039

Tabella 4.3: Significatività del TRV stimate tramite bootstrap e rispettivi  $IC_{0.95}$ , sulla finestra di 12 giorni con  $a = 0.5$

L'aver stimato il peso di  $L_1$  sui dati stessi sotto il modello non vincolato non sembra quindi aver modificato i risultati ottenuti applicando il metodo bootstrap. Inoltre, da quanto osservato misurando l'errore di previsione al variare di  $a$  sia sui dati Fertili con la convalida incrociata sia sul database Billings come insieme di verifica, sembra che il valore assunto da tale quantità non influenzi particolarmente le stime del modello.

## 4.7 I Test asintotici sulla finestra ridotta

In questa sezione saranno mostrati i risultati dei test profilo descritti nel Paragrafo 3.3, applicati ai sistemi d'ipotesi (4.9) e (4.10), con la finestra di 9 giorni:  $[-6, 2]$ .

A seguito della restrizione dell'intervallo di fertilità sarebbe, in linea di principio, più preciso ricalcolare la coppia di pesi  $(a, b)$  che minimizza l'errore di previsione del sesso. Ciò nonostante abbiamo deciso di adottare la medesima stima della finestra intera, tenendo in considerazione la notevole complessità computazione della procedura, il fatto che l'unica modifica rispetto alla precedente formulazione risiede nei giorni considerati, mentre il modello rimane invariato, e i minimi cambiamenti dell'errore di classificazione dei sessi in funzione di  $a$  osservati in precedenza.

Analogamente a quanto accaduto col precedente intervallo di fertilità, sono stati eliminati tutti i record senza relazioni sessuali in  $[-6, 2]$  poiché non contribuiscono a  $L_C$ ; così facendo le numerosità espresse in Tabella 2.4 e 2.8 si sono ridotte considerevolmente, soprattutto nella prima componente della verosimiglianza composita  $L_1$ . I cicli complessivi su cui si basa questa componente (ossia con picco identificato ed almeno un rapporto in finestra) sono passati da  $n_1 = 3525$  (con 523 concepimenti) a  $n_1 = 2588$  (con 519 concepimenti).

Tra questi vi sono 4 concepimenti (di cui 3 con sesso noto): le distribuzioni dei sessi variano quindi minimamente: i maschi passano da 215 a 214 e le femmine da 198 a 196;  $n_2$  cala perciò di sole tre unità, da 413 a 410.

In Tabella 4.4 notiamo le numerosità relative alla nuova finestra dei cicli totali, dei sessi noti e delle donne da cui tali record provengono.

Picco identificato	Cicli (Concepiti)	Donne	Cicli medi per Donna
Tutti	2588 (519)	707	3.66
Sesso noto	410	397	1.03

Tabella 4.4: Numerosità delle donne rispetto ai cicli totali e a quelli con sesso noto, dati Fertili con picco identificato e rapporti in  $[-6, 2]$

Come possiamo osservare dalla Tabella 4.5, le differenze delle stime di massimo dei parametri rispetto alla finestra di 12 giorni sono minime.

	Stima vincolata	Derivata
$\lambda_{-6}$	0.0422	-0.0032
$\lambda_{-5}$	0.0773	0.0038
$\lambda_{-4}$	0.1199	-0.0032
$\lambda_{-3}$	0.0838	-0.0002
$\lambda_{-2}$	0.1531	0.0032
$\lambda_{-1}$	0.1247	-0.0027
$\lambda_0$	0.1192	0.0026
$\lambda_1$	0.0813	-0.0009
$\lambda_2$	0.0329	0.0060
$\beta_{-6}$	0.8768	-0.0004
$\beta_{-5}$	0.4331	-0.0009
$\beta_{-4}$	0.4637	-0.0001
$\beta_{-3}$	0.6254	0.0002
$\beta_{-2}$	0.5832	0.0002
$\beta_{-1}$	0.5756	-0.0001
$\beta_0$	0.4477	-0.0013
$\beta_1$	0.3622	-0.0003
$\beta_2$	0.4032	-0.0001

Tabella 4.5: Finestra di 9 giorni: stime di massimo della verosimiglianza composta e derivata della  $l_C$  rispetto ai singoli parametri calcolata in  $\hat{\theta}_c$

#### 4.7.1 Il calcolo analitico delle derivate prime

Per applicare i test descritti nel Paragrafo 3.3 è necessario disporre del vettore delle derivate prime della log-verosimiglianza composta  $U_C(\theta)$  e delle matrici  $J(\theta) = E_\theta(U_C(\theta)U_C(\theta)^T)$  e  $H(\theta) = E_\theta(-\frac{\partial U_C(\theta)}{\partial \theta^T})$ , utilizzate per ottenere  $G(\theta)$ .

In primo luogo abbiamo calcolato analiticamente

$$U_C(\theta)^T = \left( \frac{\partial l_C(\theta; z, y)}{\partial \lambda_1}, \dots, \frac{\partial l_C(\theta; z, y)}{\partial \lambda_K}, \frac{\partial l_C(\theta; z, y)}{\partial \beta_1}, \dots, \frac{\partial l_C(\theta; z, y)}{\partial \beta_K} \right)$$

Nel seguito chiameremo per semplicità

$$Pr(Y_i = 1|X_i) = p_{1_i}(\lambda), \text{ vedi (4.1), e } Pr(Z_i = 1|Y_i = 1, X_i) = p_{2_i}(\lambda, \beta), \text{ vedi (4.3) .}$$

Riportiamo i risultati del calcolo della derivata di  $l_C(\theta; z, y)$  rispetto al gene-

rico  $\lambda_k$ .

$$U_{C\lambda_k}(\theta) = \frac{\partial l_C(\theta; z, y)}{\partial \lambda_k} = a \sum_{i=1}^{n_1} \frac{\partial l_1(\lambda; y_i)}{\partial \lambda_k} + b \sum_{i=1}^{n_2} \frac{\partial l_2(\lambda, \beta; z_i)}{\partial \lambda_k}$$

dove

$$\begin{aligned} \frac{\partial l_1(\lambda; y_i)}{\partial \lambda_k} &= y_i \frac{\partial \ln(p_{1_i}(\lambda))}{\partial \lambda_k} + (1 - y_i) \frac{\partial \ln(1 - p_{1_i}(\lambda))}{\partial \lambda_k} \\ &= y_i \frac{(\partial/\partial \lambda_k)p_{1_i}(\lambda)}{p_{1_i}(\lambda)} - (1 - y_i) \frac{(\partial/\partial \lambda_k)p_{1_i}(\lambda)}{1 - p_{1_i}(\lambda)}, \end{aligned}$$

$$\begin{aligned} \frac{\partial l_2(\lambda, \beta; z_i)}{\partial \lambda_k} &= z_i \frac{\partial \ln(p_{2_i}(\lambda, \beta))}{\partial \lambda_k} + (1 - z_i) \frac{\partial \ln(1 - p_{2_i}(\lambda, \beta))}{\partial \lambda_k} \\ &= z_i \frac{(\partial/\partial \lambda_k)p_{2_i}(\lambda, \beta)}{p_{2_i}(\lambda, \beta)} - (1 - z_i) \frac{(\partial/\partial \lambda_k)p_{2_i}(\lambda, \beta)}{1 - p_{2_i}(\lambda, \beta)}; \end{aligned}$$

con

$$(\partial/\partial \lambda_k)p_{1_i}(\lambda) = X_{ik} \prod_{j \neq k} (1 - \lambda_j)^{X_{ij}},$$

$$(\partial/\partial \lambda_k)p_{2_i}(\lambda, \beta) = -X_{ik} \left[ \frac{\beta_k \sum_k X_{ik} \lambda_k - \sum_k \beta_k X_{ik} \lambda_k}{\sum_k X_{ik} \lambda_k \{ \sum_k X_{ik} \lambda_k - X_{ik} \lambda_k \}} \right].$$

Illustriamo ora lo svolgimento della derivata di  $l_C(\theta; z, y)$  rispetto a  $\beta_k$ .

$$U_{C\beta_k}(\theta) = \frac{\partial l_C(\theta; z, y)}{\partial \beta_k} = a \sum_{i=1}^{n_1} \frac{l_1(\lambda; y_i)}{\partial \beta_k} + b \sum_{i=1}^{n_2} \frac{\partial l_2(\lambda, \beta; z_i)}{\partial \beta_k}$$

dove

$$\frac{l_1(\lambda; y_i)}{\partial \beta_k} = 0, \text{ essendo } l_1 \text{ dipendente dal solo parametro } \lambda$$

$$\begin{aligned} \frac{\partial l_2(\lambda, \beta; z_i)}{\partial \beta_k} &= z_i \frac{\partial \ln(p_{2_i}(\lambda, \beta))}{\partial \beta_k} + (1 - z_i) \frac{\partial \ln(1 - p_{2_i}(\lambda, \beta))}{\partial \beta_k} \\ &= z_i \frac{(\partial/\partial \beta_k)p_{2_i}(\lambda, \beta)}{p_{2_i}(\lambda, \beta)} - (1 - z_i) \frac{(\partial/\partial \beta_k)p_{2_i}(\lambda, \beta)}{1 - p_{2_i}(\lambda, \beta)}, \end{aligned}$$

con

$$(\partial/\partial \beta_k)p_{2_i}(\lambda, \beta) = \lambda_{ik}^*.$$

### 4.7.2 Le matrici $G(\theta)$ , $J(\theta)$ ed $H(\theta)$

Lo svolgimento analitico delle derivate prime, sebbene ci consenta di pervenire a risultati esatti e quindi più precisi di quelli numerici, non era in realtà strettamente necessario: esistono infatti algoritmi, estremamente precisi ed affidabili, che consentono di approssimare le derivate prime in un punto, ai quali avremmo potuto ricorrere per stimare le matrici nei valori di interesse (massimo globale,  $\hat{\theta}_c$ , e vincolato,  $\hat{\theta}_{H_0}$ ).

La formulazione analitica risulta però essenziale per il calcolo della matrice  $J(\theta) = E_\theta(U_C(\theta)U_C(\theta)^T)$ : è infatti necessario possedere le componenti di ciascuna unità delle derivate prime della  $l_C$ ; in questo modo l'operatore valore atteso  $E_\theta(\cdot)$  viene sostituito dalla media empirica sul campione a disposizione. La quantità così ottenuta rappresenta la stima di  $J(\theta)$  per un singolo record (essendo la media sui dati), cioè di ordine  $O(1)$ : per ottenere un'ordine  $O(n)$  è necessario ri-moltiplicare la matrice per  $n$ ; essendo interessati alla matrice relativa all'intero campione, evitiamo di dividere per  $n$ .

Risulta perciò che  $\hat{J}(\theta) = \sum_{i=1}^n U_C(\theta; z_i, y_i)U_C(\theta; z_i, y_i)^T$ .

Abbiamo parlato nel Paragrafo 2.5 di una gerarchia nei dati relativa alla coppia. Sebbene la pseudo verosimiglianza sia costruita sotto l'assunzione che i cicli siano i.i.d. (il che rappresenta un'approssimazione), sia  $J$  sia  $H$  sono delle medie su  $\theta$ , ossia rispetto al vero modello: è quindi appropriato modificare la precedente espressione di  $J$  in modo da tenere conto della dipendenza esistente tra i cicli. È infatti stato osservato in letteratura che, anche in modelli che si basano su errate specificazioni, calcolare le matrici di sensibilità e di variabilità sotto le corrette assunzioni consente di ottenere stime consistenti di tali indicatori (Varin et al., 2011).

Possiamo assumere che le coppie siano effettivamente i.i.d.: per stimare la media su cui si basa  $J$  sommiamo rispetto a queste componenti indipendenti; sostituiamo perciò gli indici della sommatoria da  $i = 1, \dots, n$  (relativo al ciclo) a  $j = 1, \dots, N$  (riferito alla coppia), dove  $N$  è eguale al numero di donne (e quindi di coppie) da cui provengono i cicli. Ne consegue che

$\hat{J}(\theta) = \sum_{j=1}^N U_C(\theta; z_j, y_j)U_C(\theta; z_j, y_j)^T$ , dove le singole componenti sono ottenute come somma delle derivate su tutti i cicli  $C(j)$  appartenenti alla coppia

$$j: U_C(\theta; z_j, y_j) = \sum_{i:i \in C(j)} U_C(\theta; z_i, y_i).$$

Data la particolare struttura del modello adottato e delle diverse numerosità delle sue componenti,  $N$  varia a seconda dei parametri in base ai quali si deriva: mentre infatti  $U_{C\lambda}(\theta; z_i, y_i)$  è calcolabile per i cicli  $i = 1, \dots, n_1$ ;

$U_{C\beta}(\theta; z_i, y_i)$  è disponibile solo per il sottoinsieme di questi con concepimento e per i quali è noto il sesso del nato, ossia  $n_2 \subset n_1$ . Analogamente, anche nei blocchi della matrice relativi alle derivate incrociate rispetto ai due parametri sono disponibili solo le  $n_2$  unità comuni ad entrambe le verosimiglianze  $L_1$  ed  $L_2$ .

Se scomponiamo  $J(\hat{\theta})$  in una matrice a quattro blocchi otteniamo

$$J(\hat{\theta}) = \begin{bmatrix} \hat{J}_{\lambda\lambda} & \hat{J}_{\lambda\beta} \\ \hat{J}_{\beta\lambda} & \hat{J}_{\beta\beta} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{N_1} U_{C\lambda}(\theta; z_j, y_j) U_{C\lambda}(\theta; z_j, y_j)^T & \sum_{j=1}^{N_2} U_{C\lambda}(\theta; z_j, y_j) U_{C\beta}(\theta; z_j, y_j)^T \\ \sum_{j=1}^{N_2} U_{C\beta}(\theta; z_j, y_j) U_{C\lambda}(\theta; z_j, y_j)^T & \sum_{j=1}^{N_2} U_{C\beta}(\theta; z_j, y_j) U_{C\beta}(\theta; z_j, y_j)^T \end{bmatrix}$$

Dove  $N_1$  sono il numero di donne per le quali si dispone di almeno un ciclo ed  $N_2$  quelle con almeno un concepimento con sesso noto; abbiamo quindi  $\sum_{j=1}^{N_1} \sum_{i:i \in C_1(j)} 1 = n_1$  e  $\sum_{j=1}^{N_2} \sum_{i:i \in C_2(j)} 1 = n_2$ , con  $C_1(j)$  pari al numero totale di cicli disponibili per la donna  $j$ -esima, mentre  $C_2(j)$  rappresenta il sottoinsieme di cicli di  $C_1(j)$  con concepimento e per i quali è noto il sesso del nato.

Per quanto riguarda il calcolo di  $H(\theta)$ , essendo tale matrice pari alla media, rispetto a  $\theta$ , delle derivate seconde di  $l_C(\theta, z, y)$ , essa andrebbe calcolata come la somma<sup>2</sup> di queste quantità misurate sulle donne del campione (come fatto per  $J$ ):  $\hat{H}(\theta) = \sum_{j=1}^N -\frac{\partial U_C(\theta; z_j, y_j)}{\partial \theta^T}$ . Questa è invece stata stimata

numericamente<sup>3</sup>: tuttavia, sebbene il calcolatore non tenga conto della gerarchia dei dati e tratti i cicli come se fossero i.i.d., otteniamo ugualmente la formulazione corretta di  $H$ . Essendo infatti  $U_C(\theta; z_j, y_j) = \sum_{i:i \in C(j)} U_C(\theta; z_i, y_i)$ ,

deriva che le due formulazioni coincidono

$$\hat{H}(\theta) = \sum_{j=1}^N \sum_{i:i \in C(j)} -\frac{\partial U_C(\theta; z_i, y_i)}{\partial \theta^T} = \sum_{i=1}^n -\frac{\partial U_C(\theta; z_i, y_i)}{\partial \theta^T}.$$

Per ciò che concerne le diverse numerosità che compongono i blocchi della

<sup>2</sup>Somma, e non media, per le medesime ragioni spiegate sopra per  $J$

<sup>3</sup>Facendo ricorso al metodo *hessian* del pacchetto *numDeriv* del software 

matrice a seconda dei parametri rispetto cui si deriva, rimane valido quanto detto sopra per  $J$

L'unica approssimazione al calcolo di  $H$  risulta quindi, non di costruzione, ma numerica (che assumiamo trascurabile) e, come per  $J$ , di calcolo empirico sui dati del valore atteso rispetto a  $\theta$ .

### 4.7.3 I risultati dei test

Una volta definito il metodo di calcolo di  $H(\theta)$  e  $J(\theta)$ , abbiamo proceduto all'applicazione dei test illustrati nel Paragrafo 3.3.3: i risultati sono visibili nelle Tabelle 4.6, 4.7 e 4.8. Ciascuno dei tre test è stato calcolato sia sostituendo  $\hat{\theta}_c$  (stima non vincolata) a  $\theta$ , sia ricorrendo a  $\hat{\theta}_{H_0}$  (stima vincolata sotto  $H_0$ ): le due varianti hanno portato a risultati molto simili, anche se delle differenze notevoli si sono evidenziate ricorrendo al metodo di Wald e alla statistica punteggio nel test profilo globale ed in quello individuale su  $\beta_{-6}$ . Tale fenomeno non è accaduto col TRV, che è sembrato essere il più stabile dei tre; tale test è stato applicato nel sistema d'ipotesi globale, sia con la correzione descritta nel Paragrafo 3.3.3, sia stimando la distribuzione asintotica esatta (simulando 1 milione di dati): possiamo osservare l'istogramma di quest'ultima nel grafico in Figura 4.6. I due calcoli (TRV corretto e con distribuzione asintotica esatta) coincidono nei test individuali, poiché la distribuzione asintotica esatta (che è una somma pesata di  $p$   $\chi_1^2$  indipendenti<sup>4</sup>) si riduce ad un singolo  $\chi_1^2$ , moltiplicato da una costante.

I test individuali sono stati confrontati con un  $\chi_1^2$ , mentre quelli globali (escluso il TRV con distribuzione asintotica esatta) con un  $\chi_8^2$ : gli 8 gradi di libertà sono dovuti al fatto che abbiamo vincolato i 9 parametri  $\beta_k$  ad assumere lo stesso valore, senza però fissarlo: abbiamo quindi 8 vincoli.

Test	Matrici calcolate in $\hat{\theta}_{H_0}$		Matrici calcolate in $\hat{\theta}_c$	
	valore	valore $p$	valore	valore $p$
Wald	9.58	0.295	15.62	0.048
Punteggio	10.10	0.258	7.22	0.513
Rapporto Corretto	10.43	0.236	10.60	0.227
Rapporto Distr. Asintotica	10.82	0.228	10.82	0.232

Tabella 4.6: Valori dei test profilo globali osservati e relative significatività, sulla finestra di 9 giorni

<sup>4</sup>Dove  $p$  è pari a  $K - 1 = 8$  nella finestra di 9 giorni, vedi Paragrafo 3.3.3

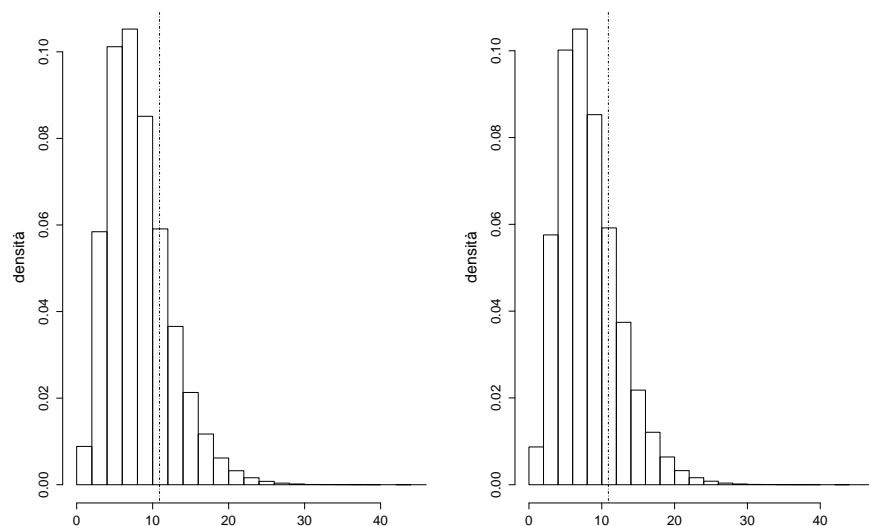


Figura 4.6: Distribuzione sotto  $H_0$  asintotica esatta del TRV profilo globale, stimata con 1 milione di simulazioni, sulla finestra di 9 giorni: a sinistra matrici calcolate in  $\hat{\theta}_{H_0}$ , a destra in  $\hat{\theta}_c$

Tutti i test globali, sia con  $J$  ed  $H$  calcolate in  $\hat{\theta}_c$  sia in  $\hat{\theta}_{H_0}$ , sono risultati ben lontani dalla significatività, ad eccezione del test Wald con matrici calcolate in  $\hat{\theta}_c$  che ha presentato un valore  $p$  di poco inferiore a 0.05. Nonostante questa indicazione più estrema rispetto alle altre riportate in Tabella 4.6, l'unica significativa, i risultati globalmente ottenuti sembrano essere perfettamente compatibili con l'ipotesi nulla.

Vi sono tuttavia forti diversità a seconda del test utilizzato e, per uno stesso test, se  $J$  ed  $H$  (e quindi  $G$ ) sono calcolate in  $\hat{\theta}_c$  o in  $\hat{\theta}_{H_0}$ . Quest'ultimo fenomeno non si osserva col TRV: i valori derivati dall'utilizzo di tali matrici calcolate nel massimo globale ( $H_1$ ) o vincolato ( $H_0$ ) sono molto prossimi tra loro; non vi sono inoltre delle differenze sostanziali ricorrendo al metodo della correzione del test o calcolando la relativa distribuzione nulla asintotica esatta. I livelli di significatività sono compresi tra il 22% ed il 24%; comunque ben lontani da 0.05, sebbene inferiori rispetto a quanto riportato nel precedente paragrafo ricorrendo al metodo bootstrap (0.42 e 0.44 rispettivamente con  $a$  pari ad 1 e 0.5).

Esaminando invece le Tabelle 4.7 e 4.8 relative ai test individuali notiamo come i risultati siano molto simili, sia al variare del test, sia passando da  $\hat{\theta}_c$  a  $\hat{\theta}_{H_0}$  nelle matrici  $J$  ed  $H$ . L'unica eccezione è rappresentata da  $\beta_{-6}$ : non solo



	Wald		Punteggio		Rapporto	
	valore	valore $p$	valore	valore $p$	valore	valore $p$
$\beta_{-6}$	11.01	0.001	5.40	0.020	5.74	0.017
$\beta_{-5}$	0.42	0.517	0.42	0.517	0.42	0.515
$\beta_{-4}$	0.27	0.604	0.27	0.605	0.27	0.606
$\beta_{-3}$	0.83	0.362	0.81	0.368	0.82	0.365
$\beta_{-2}$	0.77	0.379	0.76	0.382	0.76	0.383
$\beta_{-1}$	0.50	0.479	0.50	0.479	0.50	0.479
$\beta_0$	0.67	0.413	0.66	0.418	0.65	0.419
$\beta_1$	2.09	0.148	1.94	0.164	1.95	0.162
$\beta_2$	0.27	0.600	0.28	0.597	0.27	0.601

Tabella 4.7: Test profilo individuali e relative significatività, ricorrendo alle matrici calcolate in  $\hat{\theta}_{H_0}$ , sulla finestra di 9 giorni

vi sono evidenti disparità tra i test, ma anche al variare del valore in base al quale si calcolano  $H$  e  $J$ . Questo indica che le matrici ottenute fissando un solo parametro  $\beta_k$  sono molto simili alle medesime nel caso non vincolato, tranne quando ad essere bloccato è la quantità più diversa da 106/206, ossia  $\beta_{-6}$ . Infatti esso è l'unica probabilità a risultare significativamente differente dal tasso nullo di mascolinità alla nascita: tutti i valori  $p$  sono attorno all'1-3%. Diversamente il test del punteggio presenta un comportamento anomalo: cambiando le matrici come descritto sopra, esso varia la sua significatività da 0.02 a 0.11.

Per quanto riguarda gli altri parametri, nessun valore  $p$  si avvicina a 0.05 o 0.10: solo quelli relativi a  $\beta_1$  sono inferiori al 20% (oscillanti tra 0.15 e 0.18).

Concludiamo questa sezione indicando come i risultati discussi in questo capitolo, sulla finestra di 9 giorni utilizzando risultati asintotici, siano particolarmente prossimi a quelli mostrati nel precedente, sull'intervallo di 12 date ricorrendo al metodo bootstrap: entrambe le analisi mostrano che solamente nel giorno  $-6$  sembra essere significativamente più probabile concepire un maschio, mentre in tutti gli altri i  $\beta_k$  sono perfettamente compatibili con l'ipotesi nulla.

Completiamo queste indicazioni illustrando gli intervalli di confidenza profilo unidimensionali per i parametri trattati.

	Wald		Punteggio		Rapporto	
	valore	valore $p$	valore	valore $p$	valore	valore $p$
$\beta_{-6}$	4.96	0.026	2.62	0.106	5.96	0.015
$\beta_{-5}$	0.43	0.512	0.42	0.517	0.42	0.518
$\beta_{-4}$	0.26	0.608	0.26	0.608	0.27	0.605
$\beta_{-3}$	0.83	0.362	0.81	0.367	0.82	0.365
$\beta_{-2}$	0.75	0.386	0.75	0.387	0.76	0.382
$\beta_{-1}$	0.50	0.480	0.49	0.483	0.50	0.481
$\beta_0$	0.64	0.423	0.65	0.421	0.66	0.416
$\beta_1$	1.92	0.166	1.81	0.178	1.97	0.161
$\beta_2$	0.26	0.610	0.24	0.626	0.26	0.610

Tabella 4.8: Test profilo individuali, ottenuti ricorrendo alle matrici calcolate in  $\hat{\theta}_c$ , sulla finestra di 9 giorni

## 4.8 Gli intervalli di confidenza

Sempre lavorando sulla finestra di 9 giorni, in base ai valori osservati dei test statistici e alle distribuzioni asintotiche degli stessi, abbiamo stimato delle regioni di confidenza profilo per i parametri giornalieri  $\beta_k$ .

Per ottenere numericamente gli intervalli di confidenza relativi ai test del punteggio e del RV abbiamo utilizzato una griglia di 199 valori: uno ogni 5 millesimi da 0.005 a 0.995. Di conseguenza l'errore dovuto all'approssimazione numerica è sempre inferiore a mezzo centesimo; esso è inoltre "conservativo" poiché che i veri limiti dell'IC sono leggermente più stretti di quelli stimati.

Sono state utilizzate le matrici  $J$  ed  $H$  calcolate in  $\hat{\theta}_c$ , poiché valutarle in  $\hat{\theta}_{H_0}$  avrebbe richiesto il calcolo numerico di  $H$  in ognuna delle 199 iterazioni, per ciascun parametro. Inoltre abbiamo notato che, quando si calcolano queste matrici per valori dei parametri diversi dal massimo globale  $\hat{\theta}_c$ , i risultati dei test non sono sempre ottenibili se  $H$  non è definita positiva. Abbiamo inoltre appurato quanto i due metodi diano risultati analoghi quando applicati sui singoli parametri, soprattutto col rapporto di verosimiglianza: per tale ragione l'intervallo derivato dal TRV è stato quello che abbiamo tenuto maggiormente in considerazione.

### 4.8.1 Le regioni trovate

Gli IC, essendo basati sui valori dei test ottenuti nel precedente paragrafo, hanno naturalmente fornito informazioni analoghe. Dall'analisi dei Grafici

4.7 e 4.8 e delle Tabelle 4.9 e 4.10 emerge che tutti i parametri contengono anche nell' $IC_{0.90}$ , con ciascuno dei tre metodi, il tasso nullo di mascolinità alla nascita (linea tratteggiata orizzontale), ad eccezione di  $\beta_{-6}$ .

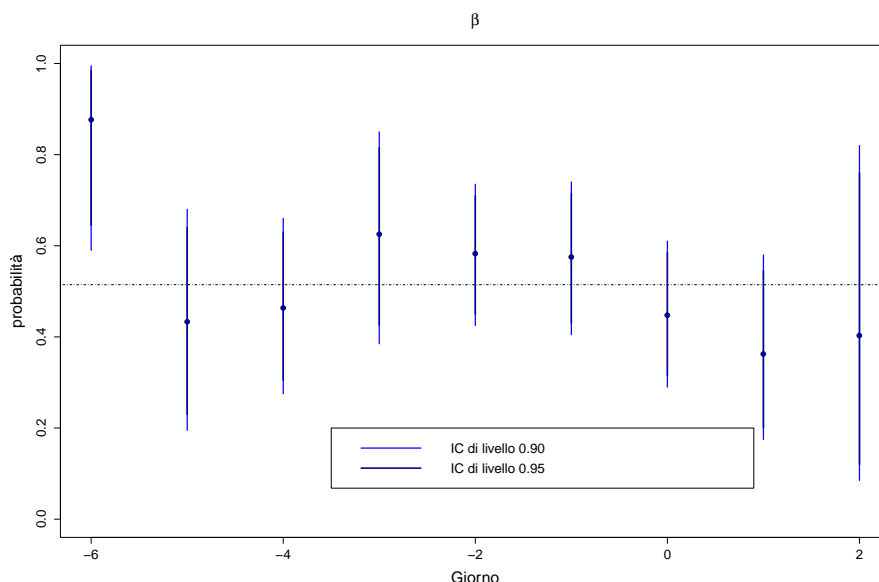


Figura 4.7: Stime dei  $\beta$  con relativi IC trovati tramite il metodo del RV

Come nel paragrafo precedente, i risultati del test del punteggio discorrono da quelli degli altri due; l' $IC_{1-\alpha/2}$  include  $106/206 = 51.46\%$ , anche ad un livello di confidenza del 90% (infatti la significatività in Tabella 4.8 è pari a 0.11).

Al contrario, l'IC di Wald e quello del rapporto di verosimiglianza escludono abbondantemente questo valore, anche ad un livello del 95%. L'IC di Wald è il meno adeguato dei tre a causa del supporto limitato delle probabilità e, per il giorno in esame, il suo limite superiore è maggiore di 1. Di conseguenza, il limite inferiore è eccessivamente elevato, poiché dovrebbe essere spostato verso lo 0 per tener conto del fatto che l'area oltre 1 dell'IC è in realtà a probabilità nulla.

L' $IC_{0.95}$  di  $\beta_{-6}$  basato sul TRV è lontano da  $106/206$ .

Nelle Figure A.2, A.3 e A.4 possiamo osservare le log-verosimiglianze profilo composite dei parametri sulla finestra di 9 giorni, nonché i livelli di significatività dei test del punteggio e del rapporto di verosimiglianza al variare dei valori di  $\beta_{H_0}$ . Notiamo come i grafici relativi al giorno  $-6$  siano i più decentrati rispetto al valore nullo, rispecchiando le indicazioni di test ed IC; notiamo come il valore  $p$  del TRV cali molto rapidamente scendendo

sotto le linee del 10% e 5% prima di 106/206, al contrario di quello punteggiato, che infatti esclude tali cut-off da entrambi gli IC.

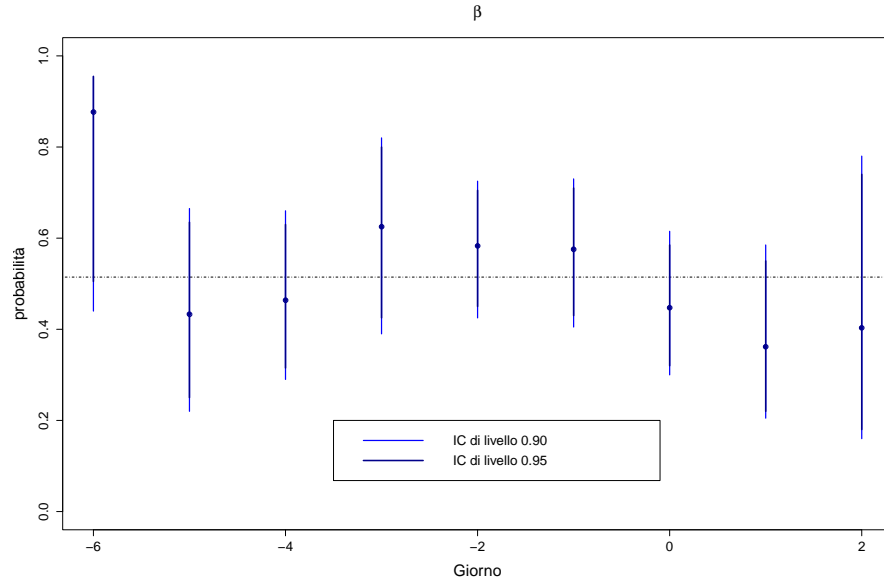


Figura 4.8: Stime dei  $\beta$  con relativi IC trovati tramite il test del punteggio

## 4.9 Un confronto tra i metodi

Le conclusioni cui siamo giunti ricorrendo al metodo bootstrap e ai risultati asintotici per ottenere la distribuzione nulla del TRV, sono estremamente simili; tuttavia notiamo delle differenze in ciò che concerne i test profilo globale e relativo al parametro  $\beta_{-6}$ : quelli basati sul bootstrap forniscono valori  $p$  maggiori.

Illustriamo ora i limiti dei due metodi nell'applicazione corrente.

I test bootstrap commettono l'imprecisione, che è un'approssimazione asintotica, di simulare i campioni da  $\hat{\theta}_{H_0}$  e non dal vero  $\theta_{H_0}$ : vincolando solo alcuni parametri, gli altri vengono ottimizzati. Sussiste inoltre l'approssimazione dovuta al fatto che la distribuzione bootstrap è solo stimata attraverso un numero limitato di campioni  $B$ ; quest'ultima è però valutabile (attraverso degli IC delle significatività stimate) e limitabile incrementando  $B$ , come accaduto per il test globale e per quello su  $\beta_{-6}$ .

Nonostante ciò i test che si basano sul bootstrap dovrebbero in linea di principio essere più accurati di quelli che ricorrono a distribuzioni asintotiche:

	Punteggio		Rapporto		Wald	
	Limite Inferiore	Limite Superiore	Limite Inferiore	Limite Superiore	Limite Inferiore	Limite Superiore
$\beta_{-6}$	0.505	0.955	0.645	0.985	0.697	1.056
$\beta_{-5}$	0.250	0.635	0.230	0.640	0.226	0.640
$\beta_{-4}$	0.315	0.630	0.305	0.630	0.302	0.625
$\beta_{-3}$	0.425	0.800	0.425	0.815	0.426	0.825
$\beta_{-2}$	0.450	0.705	0.450	0.710	0.455	0.712
$\beta_{-1}$	0.430	0.710	0.430	0.715	0.434	0.717
$\beta_0$	0.320	0.585	0.315	0.585	0.314	0.582
$\beta_1$	0.220	0.550	0.200	0.545	0.189	0.536
$\beta_2$	0.180	0.740	0.120	0.760	0.053	0.753

Tabella 4.9: Regioni di confidenza di livello 0.9, con  $J$  ed  $H$  calcolate in  $\hat{\theta}_c$

in Aerts e Claeskens (1999) è stato mostrato da un punto di vista teorico che, nell'ambito della verosimiglianza composita, il bootstrap parametrico porta ad uno stimatore consistente della distribuzione nulla del TRV; questa procedura corregge inoltre per l'incompleta specificazione della distribuzione congiunta della verosimiglianza. Nello stesso studio si mostra attraverso simulazioni che i test asintotici basati sulla distribuzione  $\chi^2$  soffrono spesso di un aumentato errore del 1° tipo; questo fenomeno non è invece presente ricorrendo al bootstrap. Tuttavia il bootstrap parametrico potrebbe fallire nel caso in cui la verosimiglianza fosse altamente mal specificata (Aerts et al., 2002); nel nostro caso però le errate specificazioni non ci sembrano particolarmente marcate, considerando che il commento di Aerts et al. fa riferimento al contesto di pseudo verosimiglianza in cui ogni modello si basa su errate specificazioni.

Nella nostra analisi i metodi asintotici, oltre a ricorrere a distribuzioni asintotiche, si basano anche su altre approssimazioni: le matrici  $H$  e  $J$  sono delle medie su  $\theta$ ; queste sono state invece stimate sui dati (tre blocchi su quattro sono una media su 410 unità). Inoltre negli IC, per i problemi notati quando  $H$  non era calcolata in un massimo, il vero valore del parametro è stato sostituito non da  $\hat{\theta}_{H_0}$  bensì da  $\hat{\theta}_c$ , sebbene il primo sarebbe più appropriato.

Col fine di comprendere se le diversità nei risultati forniti dalle due procedure fossero attribuibili al cambiamento di finestra<sup>5</sup> o ai metodi in sé, abbia-

<sup>5</sup>Il metodo bootstrap è stato applicato sull'intervallo di fertilità  $[-8, 3]$ , mentre quelli asintotici su  $[-6, 2]$

	Punteggio		Rapporto		Wald	
	Limite Inferiore	Limite Superiore	Limite Inferiore	Limite Superiore	Limite Inferiore	Limite Superiore
$\beta_{-6}$	0.440	0.955	0.590	0.995	0.663	1.091
$\beta_{-5}$	0.220	0.665	0.195	0.680	0.187	0.680
$\beta_{-4}$	0.290	0.660	0.275	0.660	0.271	0.656
$\beta_{-3}$	0.390	0.820	0.385	0.850	0.387	0.863
$\beta_{-2}$	0.425	0.725	0.425	0.735	0.430	0.736
$\beta_{-1}$	0.405	0.730	0.405	0.740	0.407	0.745
$\beta_0$	0.300	0.615	0.290	0.610	0.288	0.608
$\beta_1$	0.205	0.585	0.175	0.580	0.156	0.569
$\beta_2$	0.160	0.780	0.085	0.820	-0.013	0.820

Tabella 4.10: Regioni di confidenza di livello 0.95, con  $J$  ed  $H$  calcolate in  $\hat{\theta}_c$

mo simulato tramite bootstrap<sup>6</sup> sull'intervallo di 9 giorni, dove entrambi sono applicabili, le distribuzioni sotto  $H_0$  del TRV globale ed unidimensionale per  $\beta_{-6}$ , confrontandole con quelle derivate dalle teorie asintotiche.

La Figura 4.9 mostra, attraverso l'istogramma, quelle calcolate tramite bootstrap e, con le linee rosse e blu, le densità di quelle ottenute con le correzioni della distribuzione asintotica  $\chi_8^2$  basate sulle matrici  $H$  e  $J$  (calcolate sia in  $\hat{\theta}_{H_0}$  sia in  $\hat{\theta}_c$ ); per il test globale è riportata anche la distribuzione asintotica esatta<sup>7</sup>, anch'essa calcolata in base alle precedenti matrici. Quest'ultima manca nel TRV individuale poiché nel caso unidimensionale coincide con il metodo della correzione.

In Tabella 4.11 sono riportati i valori del test e i rispettivi livelli i significatività relativi alle metodologie appena descritte.

Notiamo come, sia per la correzione del test sia per il calcolo della distribuzione asintotica esatta, calcolare le matrici in  $\hat{\theta}_c$  o in  $\hat{\theta}_{H_0}$  non comporti differenze evidenti: le due si sovrappongono in entrambe le figure. La distribuzione corretta è inoltre estremamente prossima a quella bootstrap sebbene, in entrambe le immagini, l'istogramma superi le linee blu e rossa nella parte destra del grafico: il bootstrap sembra quindi stimare delle code leggermente più pesanti; i valori  $p$  corrispondenti ai test sono infatti appena maggiori (vedi Tabella 4.11). Anche la distribuzione asintotica esatta è in linea con le

<sup>6</sup>Ripetendo sulla finestra ridotta la procedure già eseguita nel Paragrafo 4.6

<sup>7</sup>Esatta poiché, al contrario delle precedenti, non è ottenuta correggendo il test riportandolo ad un  $\chi_8^2$  ma stimando la vera distribuzione nulla asintotica, pari ad una somma pesata di 8  $\chi_1^2$ , vedi Paragrafo 3.3.3

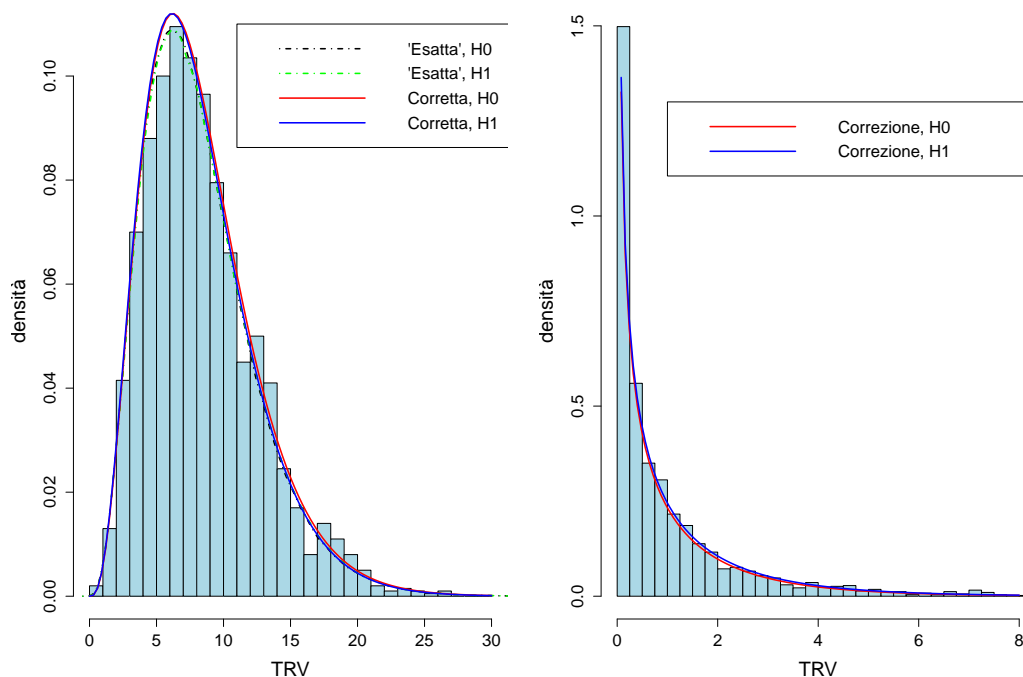


Figura 4.9: Distribuzioni sotto  $H_0$  dei TRV profilo globale (sinistra) ed individuale su  $\beta_{-6}$  (destra) stimate tramite bootstrap e relative distribuzioni asintotiche con correzioni basate su  $H$  e  $J$ . In legenda  $H_0$  ed  $H_1$  indicano che tali matrici sono state calcolate rispettivamente in  $\hat{\theta}_{H_0}$  e  $\hat{\theta}_c$

precedenti.

Ancora più rimarchevole è il fatto che i valori  $p$  di entrambi i test bootstrap risultino notevolmente inferiori rispetto a quelli osservati sulla finestra di 12 giorni.

La distanza tra i risultati mostrati nei paragrafi 4.6 e 4.7 sembra quindi attribuibile in minima parte alle differenze tra i due metodi e in maggiore alla diversità nella definizione della finestra.

Abbiamo notato che nei dati vi sono 8 cicli con rapporti solo nel giorno  $-6$  nella finestra di 9 giorni; in 7 di questi di è concepito un maschio: all'interno di queste 8 unità abbiamo quindi un rapporto bambini a bambine di 7:1. Il ciclo che ha portato alla nata femmina non presenta atti sessuali al di fuori di  $-6$  neppure nella finestra di 12 date; al contrario, dei 7 cicli relativi ai nati maschi, 4 hanno almeno un rapporto nelle tre date escluse ( $-8$ ,  $-7$  e  $3$ ). Nonostante la probabilità di concepire in tali giornate sia estremamente

Test	Metodo	TRV		valore $p$	
		$H_0$	$H_1$	$H_0$	$H_1$
Globale	Bootstrap, 2000 campioni	10.82		0.238 (0.220, 0.257)	
Globale	Correzione	10.43	10.60	0.236	0.227
Globale	Distr. Asintotica Esatta	10.82		0.228	0.232
Su $\beta_{-6}$	Bootstrap, 2000 campioni	5.86		0.023 (0.017, 0.031)	
Su $\beta_{-6}$	Correzione	5.74	5.96	0.017	0.015

Tabella 4.11: TRV profilo globale ed individuale su  $\beta_{-6}$  e relative significatività, sulla finestra di 9 giorni. Per la stima bootstrap tra parentesi è riportato l' $IC_{0.95}$  del valore  $p$  stimato

contenuta, l'informazione sul sesso di questi bambini va condivisa tra i giorni in cui si registrano atti sessuali: il  $\lambda_{-6}^*$  della (4.3) medio su questi 7 cicli non è più 1 come nella finestra di 9 giornate, ma 0.93. La quantità appena citata rappresenta la probabilità di avere fecondato l'ovulo nel giorno  $-6$  della finestra in caso di concepimento: esso indica quanto il sesso osservato nel ciclo viene "attribuito" a  $-6$ , ossia è la probabilità che il bambino nato sia stato concepito in tale giornata.  $\lambda_{-6}^*$  è pari ad 1 solo se  $-6$  è l'unico giorno della finestra con rapporti; il calo medio di tale parametro sulle 7 unità considerate si traduce quindi in una perdita di informazione: dei 7 maschi poco meno di 6.5 sono ora attribuiti a  $-6$ . Sebbene in modo minimo, lo sbilanciamento a sfavore delle femmine tra gli 8 cicli con soli rapporti in  $-6$  nella finestra di 9 date cala leggermente e diviene 6.5:1.

Essendo tali unità le più rilevanti nella stima di  $\beta_{-6}$  poiché l'informazione sui sessi dei nati è solo (nella finestra di 9 giorni) o quasi solo (in quella di 12) attribuita al giorno  $-6$ , è possibile che il test risenta del calo di informazione descritto e la significatività aumenti leggermente sulla finestra maggiore: per il test bootstrap passa infatti da 0.02 a 0.03.

Considerando che nel campione si osservano dei cicli che portano al concepimento con rapporti nei soli  $-8$ ,  $-7$  e  $3$ , si ritiene possibile fecondare in tali date: sarebbe preferibile l'uso di questo intervallo poiché così facendo non si attribuiscono ad altri giorni le probabilità di aver concepito in  $-8$ ,  $-7$  e  $3$ .

#### 4.9.1 Intervallo di confidenza per $\beta_{-6}$ tramite bootstrap

Viste le differenze notate in termini di significatività relativi ai test sul parametro  $\beta_{-6}$ , nonché la maggiore affidabilità del bootstrap spiegata sopra, abbiamo deciso di ricorrere ancora una volta a questo metodo per costrui-



re una regione di confidenza alternativa per  $\beta_{-6}$ . Considerato quanto detto sopra sui risultati derivati dai due intervalli di fertilità, abbiamo deciso di usare la finestra originale. La procedura è stata eseguita solo su  $-6$  poiché sarebbe stato notevolmente dispendioso, da un punto di vista computazionale, applicarla su 12 giorni e per il fatto che abbiamo già ottenuto degli IC per le probabilità di concepire un maschio e tutti (ad eccezione appunto di  $\beta_{-6}$ ) includono abbondantemente 106/206 anche ad un livello del 90%.

Abbiamo costruito una griglia per  $\beta_{H_0}$ , valore nullo del parametro relativo al giorno  $-6$ , e ripetendo su ogni suo punto la procedura già vista nel Paragrafo 4.6, abbiamo simulato la distribuzione nulla del TRV per ciascun  $\beta_{H_0}$ ; in base a questa abbiamo stimato il livello di significatività del test. Oltre alla distribuzione sotto  $H_0$ , anche il valore osservato del TRV varia naturalmente al modificarsi della quantità cui viene vincolato il parametro in studio.

Abbiamo quindi cercato i due  $\beta_{H_0}$  che portassero ad una significatività del 5%; definendo così i limiti superiore ed inferiore dell' $IC_{0.95}$ . Considerando che in precedenza, col metodo bootstrap, avevamo ottenuto un valore  $p$  pari al 3% saggiando il valore nullo  $106/206=51.46\%$ , sapevamo già che l'intervallo di livello 0.95 avrebbe escluso questo punto.

Siamo ricorsi ad un numero di campioni pari a 2000 per ciascun punto della griglia, aumentando successivamente tale quantità, in corrispondenza di significatività del test prossime a 0.05, sino a 8000 per il limite inferiore e 6000 per quello superiore. Le simulazioni nei diversi valori di  $\beta_{H_0}$  sono state effettuate fissando il seme del generatore casuale del software  $\mathbb{R}$  in modo tale da avere uniformità su tutta la griglia rispetto all'errore di simulazione.

Non abbiamo applicato alcuna analisi di sensitività, facendo variare il peso di  $L_1$ , sia per il dispendio computazionale della procedura, sia per il fatto che abbiamo notato come le significatività dei test bootstrap rimanessero stabili passando da  $a = 1$  ad  $a = 0.5$  (vedi Paragrafo 4.6.2); inoltre anche le precedenti figure dell'errore di previsione mostrano la poca incidenza di tale quantità sul modello.

L' $IC_{0.95}$  trovato è stato pari a  $[0.550, 0.982]$ : il limite inferiore si è posizionato tra il 51.46%, tasso nullo di mascolinità alla nascita, e 0.59, limite inferiore dell' $IC_{0.95}$  ottenuto col TRV ricorrendo alla distribuzione asintotica; il limite superiore invece è stato solo di poco minore di quello osservato nel precedente paragrafo. Sebbene questo intervallo sia più vicino del precedente a 106/206, lo esclude; anche i risultati ottenuti col bootstrap sembrerebbero quindi confermare la significativa differenza di  $\beta_{-6}$  dal valore nullo, nonostante l'implausibilità di tale quantità sia meno marcata rispetto a quanto evidenziato dai metodi asintotici.

## 4.10 Commenti ai risultati

Sebbene vi sia un solo parametro significativamente diverso da 106/206, notiamo come sia presente un certo andamento nelle probabilità di procreare un maschio al variare dei giorni, prima crescente poi calante; inoltre ogni  $\hat{\beta}_k$  è prossimo a quelli dei giorni a lui vicino, almeno nella finestra di 9 date. Tuttavia questo non accade per  $\hat{\beta}_{-6}$ : il suo valore è totalmente discordante dai due successivi. Ci aspetteremmo invece che anche nel giorno o nei due giorni dopo  $-6$ ,  $\hat{\beta}_k$  fosse maggiore di 106/206 (anche se non significativamente): sarebbe logico che i relativi IC includessero più valori superiori al 51.46% che minori.

Abbiamo quindi cercato di comprendere le ragioni di tali risultati.

Come spiegato nel precedente paragrafo, dei 410 cicli con sesso noto (nella finestra di 9 giorni) ce ne sono 8 con rapporti solo in  $-6$ ; questi sono caratterizzati da una netta preponderanza di maschi: l'87.5%. Definendo la v.c.  $X \sim Bin(n = 8, p)$ , con  $x^{oss} = 7$ , otteniamo degli IC di livello 0.9 e 0.95 per  $p$  pari a:  $[0.53, 0.99]$  e  $[0.47, 1]$ . Nonostante essi siano più ampi di quelli esaminati in questo capitolo relativi a  $\beta_{-6}$  (ad eccezione di quello derivato dal metodo del punteggio), già l' $IC_{0.9}$  basato solo su queste 8 unità esclude il valore nullo; bisogna inoltre considerare che le regioni di confidenza per  $\beta_{-6}$  osservate nei precedenti paragrafi si basano su una maggiore evidenza empirica, che include le 8 unità in esame.

Abbiamo inoltre costruito il grafico con le curve di livello (tramite il TRV col metodo della correzione) di  $\beta_{-6}$  congiuntamente col relativo  $\lambda_{-6}$  sulla finestra di 9 giorni: osservando la Figura 4.10 notiamo come, in corrispondenza di  $\hat{\lambda}_{-6}$ , linea verticale tratteggiata, la curva di livello 0.95 sia circa sovrapposta a 106/206, linea orizzontale tratteggiata, mentre questo valore è ampiamente escluso dall' $IC_{0.95}$  ottenuto col TRV (calcolando le matrici in  $\hat{\theta}_c$ ) che tiene conto solo della log-verosimiglianza profilo di  $\beta_{-6}$ . Sembrerebbe quindi che, considerare congiuntamente i due parametri  $\lambda_{-6}$  e  $\beta_{-6}$ , aumenti l'ampiezza della regione di confidenza più di quanto tenga già conto il passaggio da 1 a 2 gradi di libertà della distribuzione nulla asintotica  $\chi^2$ .

### 4.10.1 L'utilità della verosimiglianza composita

Col fine di confrontare i risultati ottenuti ricorrendo alla verosimiglianza composita con quelli ottenibili con un approccio *plug-in*, abbiamo stimato i parametri  $\lambda$  e  $\beta$  separatamente: i primi tramite  $L_1$ , i secondi tramite  $L_2$  considerando  $\lambda$  come una costante pari alla stima ottenuta su  $L_1$ .

La Tabella 4.12 riporta sia i valori di massimo dei parametri sia le significatività osservate del TRV applicato ai sistemi d'ipotesi (4.9) e (4.10):

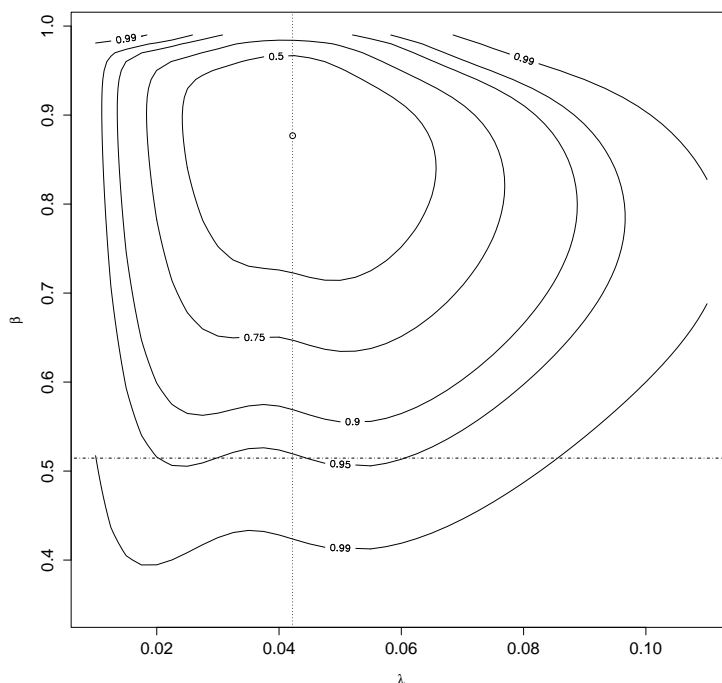


Figura 4.10: Curve di livello di  $(\lambda_{-6}, \beta_{-6})$  calcolate tramite la correzione del TRV (matrici  $J$  ed  $H$  calcolate in  $\hat{\theta}_c$ )

sebbene i risultati siano molto vicini tra loro, possiamo notare come le significatività osservate restituite dal modello composito risultino maggiori, soprattutto per il test globale e quello su  $\beta_{-6}$ . Sarebbe perciò che la verosimiglianza composita porti a stime puntuali pressoché eguali a quelle del modello non composito ma aumenti la variabilità relativa ai  $\beta$  portando a valori  $p$  leggermente maggiori.

Concludiamo il capitolo ricordando che il presente modello si basa sull'assunzione di indipendenza delle variabili tra i cicli; quanto questa abbia effettivamente influenzato le conclusioni cui siamo giunti sarà mostrato nel prosieguo della trattazione. Nel prossimo capitolo illustreremo infatti un miglioramento della formulazione utilizzata inserendo un'effetto gerarchico relativo al "fattore coppia"; descriveremo quindi come i risultati si siano modificati a seguito di questo cambiamento.

	Stima composita		Stima non composita	
	Massimo	valore $p$ , matrici in $\hat{\theta}_{H_0}$ (bootstrap, se presente)	Massimo	valore $p$
$\lambda_{-6}$	0.0422	–	0.0422	–
$\lambda_{-5}$	0.0773	–	0.0793	–
$\lambda_{-4}$	0.1199	–	0.1191	–
$\lambda_{-3}$	0.0838	–	0.0850	–
$\lambda_{-2}$	0.1531	–	0.1536	–
$\lambda_{-1}$	0.1247	–	0.1233	–
$\lambda_0$	0.1192	–	0.1196	–
$\lambda_1$	0.0813	–	0.0801	–
$\lambda_2$	0.0329	–	0.0320	–
$\beta_{-6}$	0.8768	0.017 (0.023)	0.8762	0.016
$\beta_{-5}$	0.4331	0.515	0.4339	0.497
$\beta_{-4}$	0.4637	0.606	0.4641	0.603
$\beta_{-3}$	0.6254	0.365	0.6236	0.361
$\beta_{-2}$	0.5832	0.383	0.5829	0.373
$\beta_{-1}$	0.5756	0.480	0.5756	0.473
$\beta_0$	0.4477	0.419	0.4476	0.411
$\beta_1$	0.3622	0.162	0.3619	0.156
$\beta_2$	0.4032	0.601	0.4050	0.613
Globale	–	0.235 - 0.307 (0.238)	–	0.215

Tabella 4.12: Finestra  $[-6, 2]$ : stime della verosimiglianza composita e non e significatività del TRV. Per il test globale su tutti i  $\beta$ : a sinistra il metodo della correzione, a destra quello del calcolo esatto della distribuzione nulla

# Capitolo 5

## Il modello gerarchico

In questa sezione descriveremo una modifica del modello presentato nel precedente capitolo che includa un effetto gerarchico relativo alla coppia sulle probabilità di concepimento giornaliero. L'esiguo numero di sessi dei nati disponibile per ciascuna coppia, in media 1.03, non ci ha consentito di indagare una possibile gerarchia relativa ai  $\beta_k$ ; al contrario questo si è reso possibile per i  $\lambda_k$  poiché in  $L_1$  ad ogni donna appartengono mediamente 3.7 cicli con la finestra di 9 giorni e 4.5 con quella di 12. Così facendo viene corretta l'assunzione di indipendenza delle variabili tra i cicli su cui si basava il precedente modello.

### 5.1 La formulazione

Abbiamo inserito un effetto casuale  $\xi_j$  nei  $\lambda_k$ , che chiamiamo ora  $\lambda_{jk}$ , poiché questi dipendono non solo dal giorno del ciclo ma anche dalla coppia; analogamente anche le v.c.  $Y_i$  e  $Z_i$  nonché il vettore dei rapporti  $X_i$  divengono  $Y_{ij}$ ,  $Z_{ij}$  ed  $X_{ij}$ . Modifichiamo la notazione del precedente capitolo indicando con  $j = 1, \dots, N$  la donna e con  $i = 1, \dots, n_j$  il ciclo della donna *j-esima*<sup>1</sup>.

Seguendo Dunson e Stanford (2005) riscriviamo la probabilità di concepire nel ciclo  $i$ -esimo della donna *j-esima* come

$$Pr(Y_{ij} = 1 | X_{ij}, \xi_j) = 1 - \prod_{k=1}^K (1 - \lambda_{jk})^{X_{ijk}},$$

---

<sup>1</sup> $n_j$  sarà sostituito da  $n_{1j}$  ed  $n_{2j}$  riferendosi rispettivamente ai cicli totali, in  $L_1$ , e a quelli con sesso noto, in  $L_2$ , della donna *j-esima*

dove questa volta,

$$\lambda_{jk} = 1 - \exp\{-\xi_j \exp(\alpha_k)\}, \text{ con } \xi_j \sim G(\phi, \phi),$$

dove  $G(a, b)$  indica la v.c. Gamma di parametri  $a$  e  $b$ , con media  $a/b$  e varianza  $a/b^2$ .

L'effetto  $\xi_j$  rappresenta un moltiplicatore di fecondabilità per la coppia  $j$ -esima, di media 1 e varianza  $1/\phi$ . La media unitaria evita la non identificabilità dei parametri tra la media di  $\xi_j$  e le componenti giornaliere  $\alpha_k$ ;  $\sqrt{1/\phi} = \sigma_{\xi_j}$  indica invece la variabilità tra coppie.

Una proprietà notevole di questa formulazione è che la probabilità marginale di concepire, ottenuta integrando rispetto a  $f_{G(\phi, \phi)}(\xi_j)$  (la densità di  $\xi_j$ ), ha forma chiusa,

$$\begin{aligned} Pr(Y_{ij} = 1 | X_{ij}) = p_{1_{ij}}(\phi, \lambda) &= \int_0^\infty \left\{ 1 - \prod_{k=1}^K (1 - \lambda_{jk})^{X_{ijk}} \right\} f_{G(\phi, \phi)}(\xi_j) d\xi_j = \\ &= 1 - \left( \frac{\phi}{\phi + \sum_{k=1}^K X_{ijk} \exp(\alpha_k)} \right)^\phi. \end{aligned} \quad (5.1)$$

Da questa è possibile calcolare la probabilità di fecondare l'ovulo in caso di rapporto nel solo giorno  $k$ , imponendo  $X_{ij} \equiv e_k$  (il vettore di tutti 0, ed eccezione di un 1 in posizione  $k$ -esima),

$$Pr(Y_{ij} = 1 | X_{ij} = e_k) = 1 - \left( \frac{\phi}{\phi + \exp(\alpha_k)} \right)^\phi. \quad (5.2)$$

La prima componente della verosimiglianza composta assume quindi la seguente forma

$$L_1(\phi, \alpha; y) = \prod_{j=1}^{N_1} \prod_{i=1}^{n_{1j}} p_{1_{ij}}(\phi, \lambda)^{Y_{ij}} (1 - p_{1_{ij}}(\phi, \lambda))^{1-Y_{ij}}.$$

Contrariamente al modello precedente, ipotizzare l'indipendenza tra i cicli non è più un'errata assunzione: attraverso l'effetto gerarchico modelliamo già la dipendenza tra record appartenenti alla stessa coppia.

La nuova formulazione dei  $\lambda_{jk}$  entra anche nella probabilità di procreare

un maschio dato un concepimento,

$$Pr(Z_{ij} = 1|Y_{ij} = 1, X_{ij}, \xi_j) = \sum_{k=1}^K \beta_k \frac{\lambda_{jk} X_{ijk}}{\sum_{h=1}^K \lambda_{jh} X_{ijh}} = \sum_{k=1}^K \beta_k \lambda_{ijk}^* \quad (5.3)$$

La probabilità marginale rispetto all'effetto casuale è quindi

$$Pr(Z_{ij} = 1|Y_{ij} = 1, X_{ij}) = p_{2_{ij}}(\phi, \lambda, \beta) = \int_0^\infty \sum_{k=1}^K \beta_k \lambda_{ijk}^* f_{G(\phi, \phi)}(\xi_j) d\xi_j. \quad (5.4)$$

Non è però possibile risolvere tale integrale ottenendo un risultato in forma chiusa come fatto per la (5.1).

Una soluzione a questo inconveniente consiste nel risolvere numericamente l'integrale della (5.4); tale metodo va però applicato su ciascuno degli  $n_2$  cicli presenti in  $L_2$ . Essendo il massimo della verosimiglianza composta individuato numericamente, il processo descritto è estremamente dispendioso poiché la stima degli  $n_2$  integrali va effettuata in ciascuna iterazione dell'ottimizzazione.

Questa scelta ci porta ad una serie di problematiche: i tempi di ottimizzazione estremamente elevati non ci consentono di applicare procedure iterative come il bootstrap; inoltre le derivate seconde della  $l_C$  in  $\hat{\theta}_c$  in cui si deriva prima rispetto a  $\phi$ , ossia  $\frac{\partial^2 l_C(\theta)}{\partial \phi \partial \theta^T} = \frac{\partial U_{C\phi}(\theta)}{\partial \theta^T}$ , non sono ottenibili numericamente e il calcolo analitico non conduce al risultato cercato perché prevede un integrale rispetto a  $\xi_j$  non in forma chiusa che non converge numericamente. L'integrale rispetto a  $\xi_j$  presente nella  $\frac{\partial p_{2_{ij}}(\phi, \lambda, \beta)}{\partial \phi}$  non è infatti calcolabile con limite inferiore di integrazione eguale a 0 e, partendo da punti ad esso vicino come  $10^{-8}$  o  $10^{-10}$ , restituisce risultati differenti poiché in  $0^+$  questa funzione sembra tendere a infinito; tale inconveniente si è ripetuto sia utilizzando la finestra intera sia quella ridotta. In questo modo non è possibile saggiare la significatività dei parametri del modello tramite i test asintotici illustrati nel Paragrafo 3.3 né ricorrere al bootstrap per stimare la distribuzione nulla del test, essendo eccessivamente dispendioso sul modello costruito. Un'alternativa più veloce computazionalmente e che consente di aggirare le difficoltà appena citate consiste nel sostituire, solo in  $L_2$ ,  $\lambda_{jk}$  con una quantità che non contiene l'effetto gerarchico e che ci permette di evitare l'integrazione numerica della (5.4) e delle sue derivate. Sebbene questa rappresenti un'imprecisione poiché esprime la stessa quantità  $\lambda_{jk}$  in modi differenti nelle due componenti, l'approssimazione non è inverosimile poiché

in  $L_2$  quasi ogni donna possiede un solo record: 413 cicli appartengono a 400 donne<sup>2</sup>; queste unità contribuiscono quindi solo marginalmente alla stima del parametro  $\phi$ .

Un possibile candidato alla sostituzione è  $\lambda_{jk}$  marginalizzato rispetto all'effetto casuale, ossia la probabilità  $Pr(Y_{ij} = 1 | X_{ij} = e_k)$  espressa in formula (5.2). Un'altra possibilità è rappresentata da  $\lambda_{jk}$  con il parametro gerarchico fissato eguale alla sua media: ponendo  $\xi_j \equiv 1$  otteniamo la quantità  $1 - \exp\{-\exp(\alpha_k)\}$ . Un'ulteriore alternativa consiste nel considerare la media rispetto a  $\xi_j$  di  $\lambda_{jk}$ :  $E_{\xi_j}(\lambda_{jk}) = \int_0^\infty \{1 - \exp(-\xi_j \exp(\alpha_k))\} f_{G(\phi, \phi)}(\xi_j) d\xi_j$ ; tuttavia quest'integrale non è in forma chiusa: dovremmo quindi ricorrere nuovamente a metodi di integrazione numerica per il calcolo della (5.4) e delle sue derivate. Le prime due approssimazioni presentano perciò un notevole vantaggio computazionale consentendo di effettuare operazioni dispendiose come il bootstrap, non applicabili (se non in tempi enormi) nell'ultima citata.

Abbiamo scelto di ottimizzare  $L_C$  sia ricorrendo alle due formulazioni approssimate scelte in  $L_2$  sia rifacendosi a quella originale di  $\lambda_{jk}$ , col fine di confrontare i risultati e verificare quanto le modifiche apportate incidessero sulle stime. Successivamente abbiamo applicato i test statistici su una delle due formule approssimate, vista l'impossibilità di questa procedura su quella iniziale.

A prescindere dall'utilizzo di  $\lambda_{jk}$  con l'effetto gerarchico o di una sua approssimazione in  $p_{2_{ij}}(\phi, \lambda, \beta)$ , la seconda componente della verosimiglianza composita è espressa come:

$$L_2(\phi, \alpha, \beta; z) = \prod_{j=1}^{N_2} \prod_{i=1}^{n_{2j}} p_{2_{ij}}(\phi, \lambda, \beta)^{Z_{ij}} (1 - p_{2_{ij}}(\phi, \lambda, \beta))^{1-Z_{ij}},$$

dove  $\phi$  è assente nel caso in cui si utilizza l'approssimazione che fissa  $\xi_j \equiv 1$ .

La verosimiglianza composita è perciò uguale a:

$$L_C(\theta; z, y) = L_1(\phi, \alpha; y)^a L_2(\phi, \alpha, \beta; z)^b,$$

dove  $\theta = (\phi, \alpha, \beta)$ .

## 5.2 Le stime puntuali dei parametri

Fissando il peso della prima componente di  $L_C$ ,  $a = 1$ , abbiamo massimizzato, sull'intervallo di fertilità di 12 date, la log-verosimiglianza composita con

<sup>2</sup>Nella finestra di 12 date; in quella di 9 abbiamo 410 cicli provenienti da 397 donne



le tre formulazioni di  $\lambda_{jk}$  in  $L_2$ . La Figura 5.1 mostra le probabilità giornaliere di fecondare l'ovulo dato un solo giorno con rapporti nella finestra e quelle di procreare un maschio in caso di concepimento in quella data; la Tabella A.1 in appendice riporta le stime di tutti i parametri nonché le derivate prime della  $l_C$  rispetto ad essi calcolate in  $\hat{\theta}_C$ .

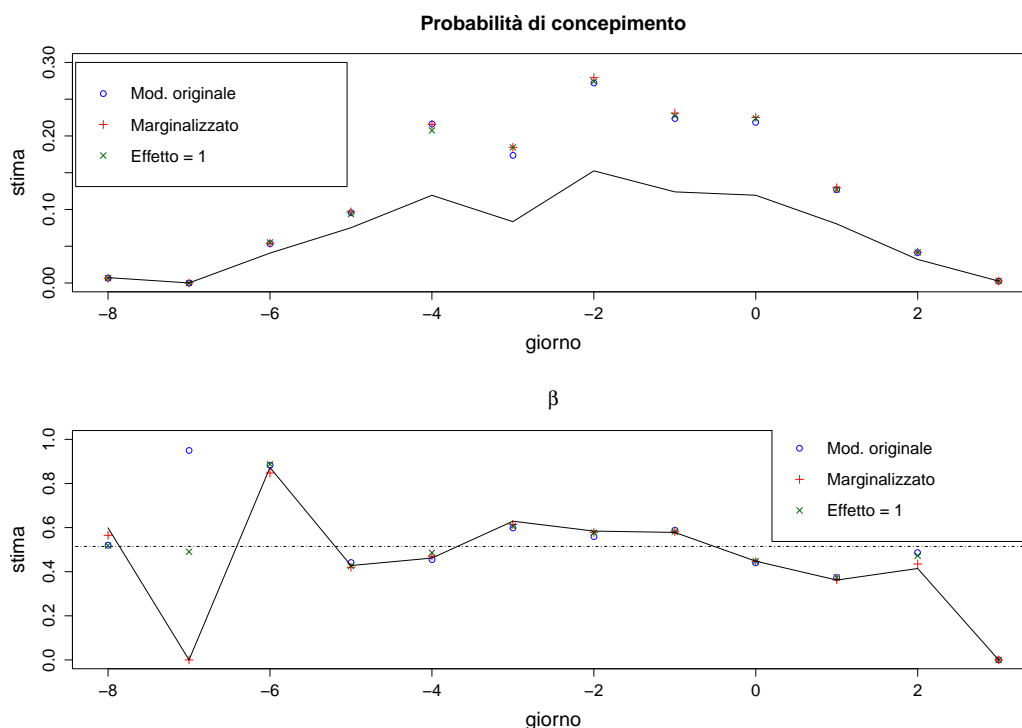


Figura 5.1: Stime delle probabilità giornaliere di concepimento (alto) e di  $\beta$  (basso) con le tre formulazioni di  $\lambda_{jk}$  in  $L_2$ . La linea nera continua rappresenta le stime del modello non gerarchico, quella tratteggiata (basso) 106/206

È evidente come i tre metodi forniscano stime analoghe, talvolta persino sovrapposte le une alle altre nella figura; l'unica eccezione è rappresentata da  $\beta_{-7}$  per il fatto che la probabilità stimata di concepire in tale giorno è pressoché nulla: la sua influenza nella verosimiglianza composta è perciò minima. Le stime dei parametri  $\beta$  sono inoltre prossime a quelle ottenute con la formulazione utilizzata nel precedente capitolo; quelle delle probabilità giornaliere di concepimento invece sono piuttosto differenti: sebbene presentino lo stesso andamento, nel modello corrente sono nettamente maggiori. Questo dato rispecchia quanto già visto ricorrendo alla formulazione di SMH in  $L_1$  (vedi Figura 4.2), risultato inoltre maggiormente concorde con quelli

solitamente riportati in letteratura. Analogamente a come accaduto in quel caso, questa differenza non si ripercuote nei  $\hat{\beta}$ , verosimilmente per il fatto che essi dipendono dal  $\lambda_{ijk}^*$  della (5.3) in cui ogni diversità di ordine di grandezza è ininfluente.

Per il prosieguo delle analisi tra le due approssimazioni abbiamo scelto di utilizzare quella che sostituisce, in  $L_2$ ,  $\lambda_{ij}$  con la formulazione (5.2) poiché, per evitare l'integrale che coinvolge l'effetto casuale, sembra logico ricorrere alla stessa probabilità, ossia alla probabilità di concepire con rapporti solo in un giorno della finestra, però in  $L_1$  condizionata a  $\xi_j$  ed in  $L_2$  marginalizzata rispetto ad esso.

Abbiamo inoltre utilizzato la finestra  $[-6, 2]$  per problemi algoritmici notati su quella originale che invece non si presentano nell'intervallo ridotto.

### 5.3 Il calcolo del peso $a$

Analogamente a quanto effettuato nel precedente capitolo, abbiamo valutato l'errore di previsione del sesso su una griglia di valori (uno ogni decimo) di  $a$  tra 0 (escluso) e 2, per stabilire quale peso della prima componente adottare.

Abbiamo in primo luogo stimato il modello sui dati Fertili e previsto i sessi dei nati sul database Billings; quest'operazione è risultata poco informativa: l'errore è stato costante tra 0.2 e 2 (e pari al 50.93%) e leggermente inferiore in 0.1 (49.07%).

Siamo quindi ricorsi al metodo della convalida incrociata sul data set Fertili sperando in una maggiore variabilità: in Figura 5.2 sono riportati i risultati trovati.

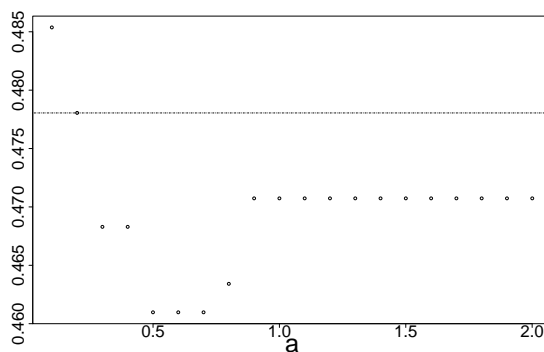


Figura 5.2: Errore di classificazione al variare di  $a$ , stimato col metodo della convalida incrociata

Abbiamo ottenuto un andamento prima calante poi crescente tra 0 ed 1, ed una totale stabilità oltre l'unità: il minimo è raggiunto in 0.5, 0.6 e 0.7; abbiamo deciso di porre  $a = 0.6$  essendo il valore centrale tra i tre. Rispetto al modello precedente facciamo notare come l'errore sia tendenzialmente minore ed in particolare, nei valori scelti dei pesi<sup>3</sup>, passa dal 46.73% a 46.10%; inoltre, contrariamente a quanto accadeva nel precedente capitolo, la probabilità media di concepire un maschio stimata con la convalida incrociata è di poco maggiore nei cicli che portano a bambini rispetto ai cicli in cui si concepiscono bambine.

Queste indicazioni sembrano suggerire che il cambio di formulazione abbia aumentato leggermente le prestazioni del modello; sembra inoltre più sensato che l'errore raggiunga il suo minimo pesando maggiormente la seconda componente della verosimiglianza composita ove si modella la probabilità di concepire un maschio.

Adottando la finestra ridotta ed il nuovo valore del peso di  $L_1$ , le stime di massimo dei parametri rimangono sostanzialmente stabili (vedi Tabella A.2 in appendice).

Rispetto invece alle quantità ottenute sull'intervallo di 12 giorni fissando  $a = 1$  nel modello originale (ossia con la corretta formulazione di  $\lambda_{jk}$  in  $L_2$ ) le probabilità di concepimento differiscono tutte per valori inferiori a 0.01, mentre per i  $\beta_k$  i cambiamenti sono minori di 0.02 ad eccezione di  $\hat{\beta}_2$  che varia di 6 centesimi. Sembra quindi non vi siano sostanziali differenze a seguito dell'approssimazione effettuata e delle modifiche di  $a$  nonché delle finestre; in particolare, relativamente al parametro di maggior interesse dati i risultati del precedente capitolo, le due stime di  $\beta_{-6}$  risultano notevolmente vicine: la loro distanza è inferiore a 2 millesimi. Il confronto è stato effettuato naturalmente sui soli 9 parametri reattivi ai giorni comuni ad entrambe le formulazioni.

## 5.4 I test bootstrap

In virtù di quanto discusso nel Paragrafo 4.9, abbiamo deciso di adottare un approccio di tipo bootstrap per stimare la distribuzione nulla del TRV dei sistemi d'ipotesi saggiati, preferendo tale metodo a quello asintotico. Abbiamo valutato i sistemi d'ipotesi (4.9) e (4.10), già affrontati nel precedente capitolo, ed il seguente, atto a saggiare la significatività dell'effetto casuale inserito,

$$\begin{cases} H_0 : 1/\phi = 0 \\ H_1 : 1/\phi > 0 \end{cases} \quad (5.5)$$

---

<sup>3</sup>Ossia con  $a = 1$  nel precedente modello e  $a = 0.6$  nella notazione corrente

Dove  $1/\phi = \text{Var}(\xi_j) = 0$  corrisponde ad assumere che l'effetto casuale sia una v.c. degenera in 1.

Per ciò che concerne i primi due test, abbiamo simulato la distribuzione nulla in modo simile a quanto fatto nel Paragrafo 4.6, con delle modifiche per adattarsi al modello corrente. La probabilità  $Pr(Y_{ij} = 1|X_{ij}, \hat{\lambda}_{H_0})$  è stata stimata attraverso la formula (5.1), mentre  $\hat{\lambda}_{H_0k}$  è stato sostituito dalla (5.2) in  $Pr(Z_{ij}^b = 1|X_{ij}, Y_{ij}^b = 1, \hat{\lambda}_{H_0}, \hat{\beta}_{H_0})$ ; in entrambe la (5.1) e la (5.2)  $(\lambda, \phi) = (\hat{\lambda}_{H_0}, \hat{\phi}_{H_0})$ .

Per quanto riguarda il test su  $\xi_j$ , trattandosi non solo di un modello differente ma anche di un parametro diverso, i cambiamenti sono maggiori. Essendo  $\xi_j$  fissato, i  $\lambda_{jk}$  sono nuovamente delle costanti ignote che dipendono solo dall'indice del giorno  $k$ ; ne segue che

$$Pr(Y_{ij} = 1|X_{ij}, \hat{\lambda}_{H_0}) = 1 - \prod_{k=1}^K (1 - \hat{\lambda}_{H_0k})^{X_{ijk}},$$

con  $\hat{\lambda}_{H_0k} = 1 - \exp(-\exp(\hat{\alpha}_{H_0k}))$ . Analogamente  $\hat{\lambda}_{H_0}$  assume tale forma anche in  $Pr(Z_{ij}^b = 1|X_{ij}, Y_{ij}^b = 1, \hat{\lambda}_{H_0}, \hat{\beta}_{H_0})$ .

Chiariamo che, nonostante l'approssimazione effettuata in  $L_2$ , i modelli sotto l'ipotesi nulla ed alternativa sono comunque annidati ed inoltre, sotto  $H_0$ , la formulazione approssimata di  $\lambda_{jk}$  equivale a quella esatta; infatti

$$\lim_{\phi \rightarrow +\infty} 1 - \left( \frac{\phi}{\phi + \exp(\alpha_k)} \right)^\phi = 1 - \exp\{-\exp(\alpha_k)\}.$$

### 5.4.1 I risultati dei test

La Tabella 5.1 riporta le significatività osservate dei test ed gli  $IC_{0.95}$  dei valori  $p$ ; le Figure 5.3 e 5.4 mostrano le distribuzioni nulle bootstrap stimate dei TRV applicati.

Il test su  $\phi$  è risultato altamente significativo: dei 1000 campioni bootstrap simulati nessuno ha raggiunto il valore osservato del TRV, pari a 42.93; il massimo del test su questi campioni è stato infatti 7.75.

Dall'istogramma notiamo che molti valori, circa la metà, sono negativi; tuttavia la grande maggioranza di questi ultimi (il 93%) sono compresi tra  $-0.001$  e  $0$  e solo 4 sono inferiori a  $-0.1$  ma comunque maggiori di  $-0.4$ . Sebbene il TRV sia per costruzione positivo, il fenomeno osservato è dovuto ad approssimazioni numeriche di  $l_C$ : se ottenessimo il massimo della verosimiglianza composita in maniera analitica tutti i valori sarebbero positivi; tuttavia ottimizzando numericamente la funzione può accadere che il massi-

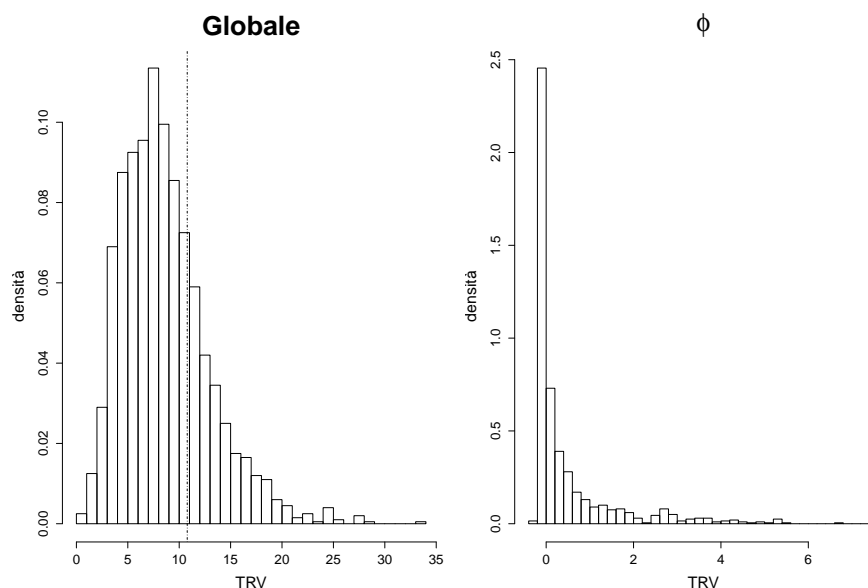


Figura 5.3: Distribuzione bootstrap sotto  $H_0$  del TRV profilo globale (sinistra) e su  $\phi$  (destra), finestra di 9 giorni con  $a = 0.6$ . La linea tratteggiata indica il valore osservato del test, per  $\phi$  non compare poichè pari a 42.93

mo di  $l_C$  vincolato (sotto  $H_0$ ) risulti leggermente maggiore (in modo infimo) dello stesso senza vincoli (sotto  $H_1$ ), soprattutto quando il campione simulato rende  $H_0$  particolarmente “plausibile”. Abbiamo infatti notato che in molti campioni bootstrap la stima di  $\phi$  del modello non vincolato è pari a qualche decina o centinaio, il che indica che la varianza dell’effetto gerarchico è estremamente contenuta; in questi casi spesso l’ottimizzazione del modello non vincolato ha maggiore difficoltà ad individuare l’ottimo con un parametro in più che sostanzialmente non aggiunge informazione.

I test individuali sui  $\beta_k$  hanno fornito indicazioni analoghe a quelle riportate nel Capitolo 4: le significatività osservate sono tutte prossime a quelle viste in precedenza ed in particolare l’unico parametro significativamente diverso da  $106/206$  è  $\beta_{-6}$  con un valore  $p$  pari al 2.1%.

Il test globale su  $\beta$  è non significativo, inoltre il valore  $p$  del test è estremamente prossimo a quello osservato nel precedente capitolo quando abbiamo applicato il metodo bootstrap per calcolare la distribuzione nulla del TRV globale nella finestra di 9 giorni (vedi Tabella 4.11).

Per quanto l’effetto gerarchico risulti altamente significativo, l’aver modellato la dipendenza delle probabilità giornaliere di concepimento tra i cicli non ha quindi modificato i risultati ottenuti.

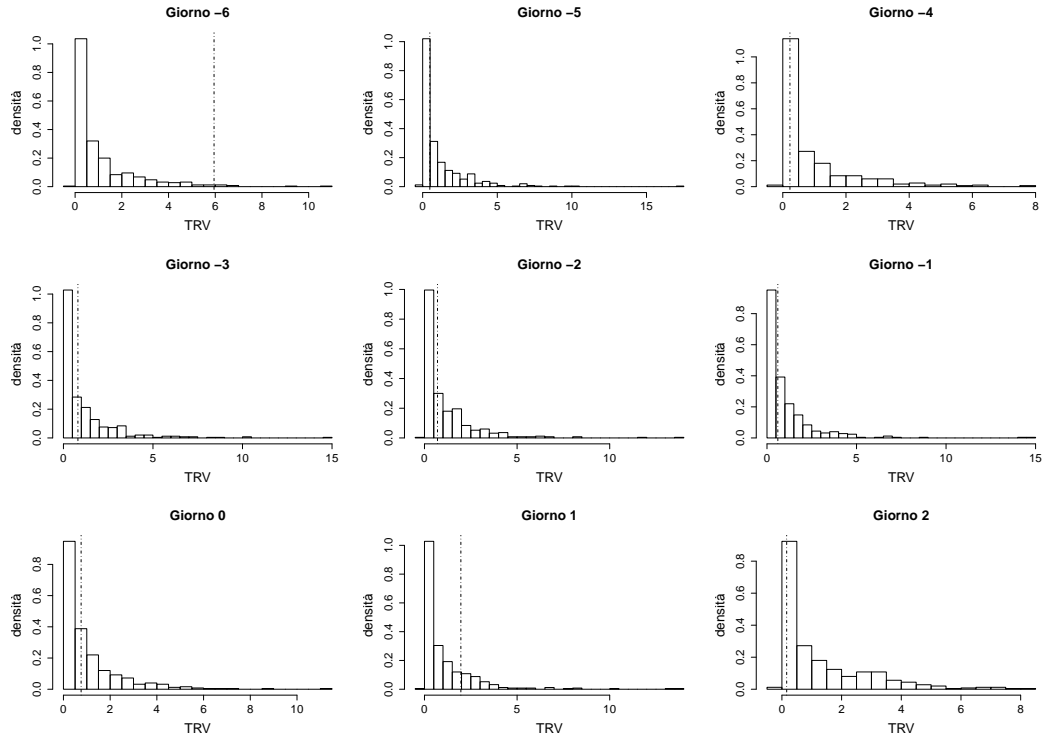


Figura 5.4: Distribuzione bootstrap sotto  $H_0$  del TRV profilo individuali; la linea verticale tratteggiata indica il valore osservato del test,  $a = 0.6$

## 5.4.2 Un'analisi di sensitività

Analogamente a quanto effettuato nel precedente capitolo, abbiamo applicato un'analisi di sensitività col fine di valutare se i risultati dei test bootstrap fossero stati influenzati dall'aver scelto  $a = 0.6$  come quel valore che minimizza l'errore di previsione del sesso sotto l'ipotesi non vincolata ( $H_1$ ).

Abbiamo quindi nuovamente saggiato alcuni tra i test profilo applicati sopra, scegliendo il peso di  $L_1$  indipendentemente dai dati: in questo caso abbiamo posto  $a = 1$ .

Abbiamo inoltre applicato una seconda analisi di sensitività per comprendere se l'approssimazione adottata nella formulazione di  $\lambda_{jk}$  in  $L_2$  avesse condizionato i risultati ottenuti, con particolare riferimento al test su  $\phi$ .

### Il peso di $L_1$ $a = 1$

Abbiamo ricalcolato il test globale sui  $\beta$  e quelli individuali su  $\phi$  e sul solo  $\beta_{-6}$ , essendo l'unico  $\beta_k$  significativo della finestra; in Figura 5.5 possiamo

Test	Significatività osservata	Campioni bootstrap	$IC_{0.95}$	
			LB	UB
Globale	0.257	2000	0.238	0.276
$\phi$	0.000	1000	0.000	0.004
$\beta_{-6}$	0.024	2500	0.019	0.031
$\beta_{-5}$	0.496	500	0.451	0.541
$\beta_{-4}$	0.592	500	0.548	0.635
$\beta_{-3}$	0.386	500	0.343	0.430
$\beta_{-2}$	0.440	500	0.396	0.485
$\beta_{-1}$	0.468	500	0.424	0.513
$\beta_0$	0.402	500	0.359	0.447
$\beta_1$	0.180	500	0.147	0.217
$\beta_2$	0.722	500	0.681	0.761

Tabella 5.1: Significatività del TRV stimate tramite bootstrap e rispettivi  $IC_{0.95}$ , sulla finestra di 9 giorni con  $a = 0.6$

osservare le distribuzioni nulle bootstrap dei TRV applicati ed in Tabella 5.2 le relative significatività osservate.

Test	Significatività osservata	Campioni bootstrap	$IC_{0.95}$	
			LB	UB
Globale	0.256	2000	0.237	0.275
$\phi$	0.000	1000	0.000	0.004
$\beta_{-6}$	0.022	2000	0.016	0.029

Tabella 5.2: Significatività del TRV stimate tramite bootstrap e rispettivi  $IC_{0.95}$ , sulla finestra di 9 giorni con  $a = 1$

Il test globale su  $\beta$  ha fornito pressoché la medesima significatività del caso precedente.

Quello su  $\phi$ , analogamente a prima, è caratterizzato dal fatto che il TRV osservato è nettamente maggiore dei valori simulati sotto  $H_0$ , e che quasi la metà di questi ultimi sono compresi tra -0.001 e 0.

Anche il test su  $\beta_{-6}$  ha differito in modo infimo rispetto a quando  $a = 0.6$ : la significatività è passata dal 2.1% all'1.8%; questa minima differenza è verosimilmente attribuibile alla variabilità della stima dei due valori  $p$ , dovuta al non eccessivo numero di campioni bootstrap.

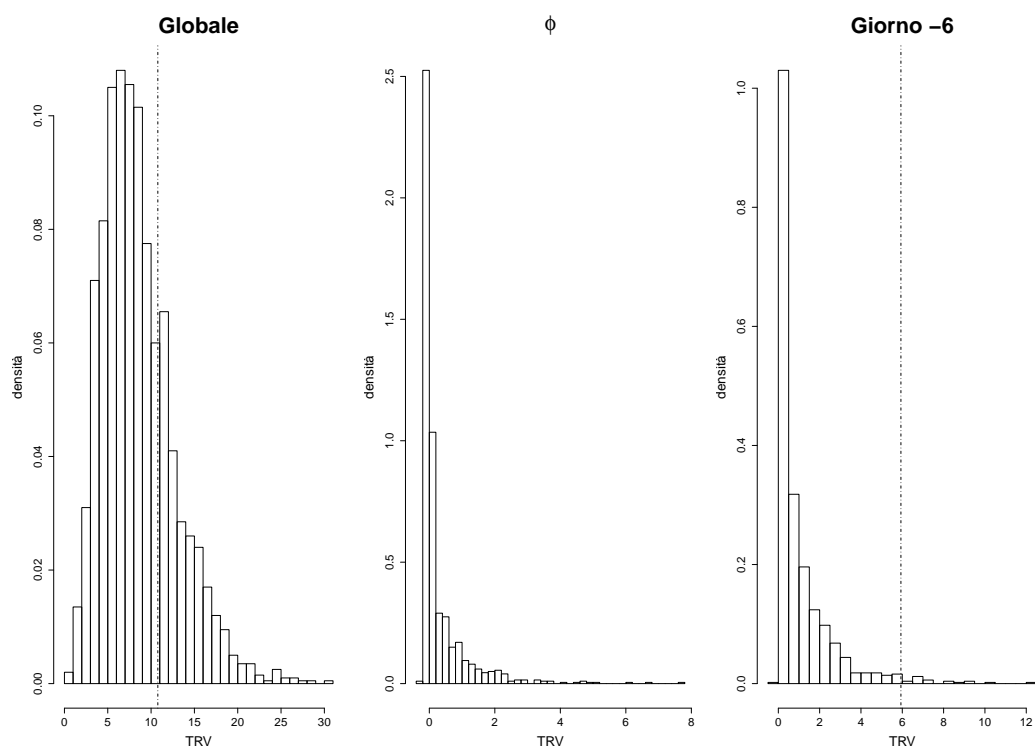


Figura 5.5: Distribuzione bootstrap sotto  $H_0$  del TRV profilo globale (sinistra) e su  $\phi$  (destra), finestra di 9 giorni con  $a = 1$ . La linea tratteggiata indica il valore osservato del test, per  $\phi$  non compare poichè pari a 71.59

### L'approssimazione $\xi_j = 1$ in $L_2$

Abbiamo ripetuto i soli test profilo su  $\phi$  e  $\beta_{-6}$ , gli unici parametri significativi nella precedente analisi; in Figura 5.6 sono raffigurate le distribuzioni nulle bootstrap dei TRV ed in Tabella 5.3 sono riportati i rispettivi valori  $p$  stimati.

Il test su  $\phi$  presenta le medesime caratteristiche dei due osservati con l'altra approssimazione di  $\lambda_{jk}$  in  $L_2$ .

Quello su  $\beta_{-6}$  ha significatività stimata leggermente maggiore, ma pur sempre sotto il 3%.

Non sembra quindi che né il valore del peso  $a$  né l'approssimazione di  $\lambda_{jk}$  in  $L_2$  abbiano influenzato i risultati. Inoltre, sebbene il parametro  $\phi$ , e di conseguenza l'effetto gerarchico, sia altamente significativo, non riscontriamo modifiche sostanziali rispetto al modello esaminato nel precedente capitolo né nelle stime puntuali dei  $\beta_k$  né nelle significatività di tali parametri nei test considerati. L'effetto casuale inserito non pare quindi aver



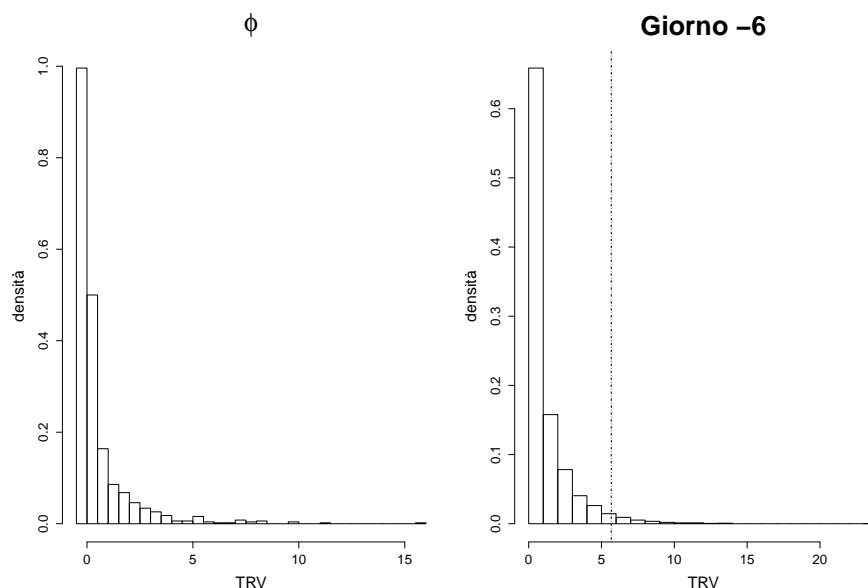


Figura 5.6: Distribuzione bootstrap sotto  $H_0$  del TRV profilo su  $\phi$  (sinistra) e su  $\beta_{-6}$  (destra), finestra di 9 giorni con  $\xi_j \equiv 1$  in  $L_2$  ed  $a = 1$ . La linea tratteggiata indica il valore osservato del test, per  $\phi$  non compare poichè pari a 71.09

alterato particolarmente né il valore di massimo né la variabilità dei  $\beta_k$ .

## 5.5 Considerazioni sui risultati ottenuti

### 5.5.1 Un confronto con la letteratura

Sebbene nella nostra analisi globalmente non siano stati rilevati effetti significativi, nel giorno  $-6$  sembra essere più probabile concepire un maschio: abbiamo quindi voluto effettuare un breve confronto tra questa indicazione ed i risultati della letteratura.

Guerrero nel 1974 stima, sulla base di 875 concepimenti, che la probabilità di procreare un maschio concependo almeno 6 dì prima del “giorno 0” sia pari al 68%, mentre diminuisca sino a giungere al 44% in caso di rapporto fecondo nel “giorno 0” stesso (vedi Paragrafo 1.1). Ciò nonostante, le stime giornaliere dei  $\beta_k$  calano gradualmente nelle analisi di Guerrero: non si osserva quindi uno stacco netto tra una data e quelle vicine, come accade nella nostra analisi con  $\beta_{-6}$ . Bisogna inoltre adottare cautela nel confrontare i nostri risultati

Test	Significatività osservata	Campioni bootstrap	$IC_{0.95}$	
			LB	UB
$\phi$	0.000	1000	0.000	0.004
$\beta_{-6}$	0.029	4000	0.024	0.035

Tabella 5.3: Significatività del TRV stimate tramite bootstrap e rispettivi  $IC_{0.95}$ , sulla finestra di 9 giorni con  $\xi_j \equiv 1$  in  $L_2$  ed  $a = 1$

con quelli ottenuti da tale studioso per il fatto che egli ricorre all'analisi della temperatura corporea basale per stimare il "giorno 0".

Il metodo Billings si basa invece sul picco del muco per determinare il "giorno 0"; le sue teorie sono tuttavia opposte rispetto agli esiti del nostro studio: l'elevata probabilità di concepire un maschio 6 giorni prima dell'ovulazione è infatti totalmente contraria a quanto sostenuto da Billings (Billings e Westmore, 1980), il quale ritiene più probabile concepire un maschio il giorno del picco del muco e quelli successivi e più probabile concepire una femmina prima del "giorno 0" (vedi Paragrafo 1.1). Anche il metodo Shettles è affine a quello Billings in ciò che concerne l'influenza dei giorni dei rapporti sul sesso del nato.

### 5.5.2 I test multipli

Bisogna tenere in considerazione che sinora abbiamo applicato numerosi test e questo aumenta la probabilità che almeno uno di loro risulti significativo, anche quando i dati provengono effettivamente da  $H_0$ .

Escludendo il test globale che dipende dagli altri e rappresenta quasi il riassunto di quelli individuali, abbiamo applicato  $K$  test "quasi" indipendenti: possiamo pensare ai test sui singoli parametri come quasi indipendenti poiché esaminano ogni giorno separatamente dagli altri, sebbene ci sia un legame tra date vicine. Infatti un valore elevato della stima in una data è probabile sia dovuto ad un fenomeno biologico che riguarda quel periodo del ciclo; inoltre, anche se esso fosse limitato a quella giornata, il fatto che il "giorno 0" sia stimato fa sì che in realtà alcuni sessi concepiti in quello che è indicato come giorno  $k$ , siano in realtà stati procreati in quelli limitrofi. Tuttavia ragioniamo per semplicità in termini di test indipendenti, ipotesi comunque parzialmente plausibile. Essendo le indicazioni dei vari test (Wald, punteggio e rapporto di verosimiglianza) quasi le medesime, possiamo focalizzarci solo su uno di essi, cioè il TRV che è parso essere il più stabile calcolando le matrici  $H$  e  $J$  in  $\hat{\theta}_c$  o  $\hat{\theta}_{H_0}$ .

Consideriamo ad esempio il risultato ottenuto nel precedente capitolo sulla finestra intera dove la significatività di  $\beta_{-6}$  era 0.03: la probabilità che sotto  $H_0$  almeno uno su 12 test indipendenti dia valore  $p$  uguale o inferiore a quello osservato è  $1 - (1 - 0.03)^{12} = 0.31$ .

Un'analogia indicazione è ottenibile ragionando in termini di soglia di significatività: la probabilità di osservare almeno un valore  $p$  su 12 sotto il 5% se vale  $H_0$  è  $1 - 0.95^{12} = 0.46$ .

Applicando numerosi test indipendenti non è quindi così improbabile osservare significatività come quelle ottenute, anche sotto  $H_0$ .



# Conclusioni

Attraverso quest'elaborato abbiamo cercato di fornire un contributo alla letteratura, focalizzando la nostra attenzione sui giorni dei rapporti sessuali; nonostante vi siano svariate proposte per ciò che concerne il tipo di relazione esistente tra questo fattore ed il sesso del nascituro, non vi sono altrettanti lavori di verifica basati su dati esterni ed indipendenti da quelli degli studi che le hanno elaborate. Questo genere di analisi è complicato da problematiche intrinseche degli studi sui sessi.

In prima istanza, il giorno dell'ovulazione è normalmente ignoto, ed è pertanto stimato: metodologie di stima diverse determinano risultati con differenze non trascurabili (Mauro, 2011). A ciò si somma il fatto che l'ovulazione si verifica in un preciso istante e che l'ovulo è fecondabile per un breve lasso di tempo: considerare dunque il giorno in cui essa avviene e non il momento esatto, è di per sé un'approssimazione.

Inoltre, la definizione della finestra di fertilità non è univoca: nella nostra analisi abbiamo notato come lievi modifiche possano variare i risultati (vedi Paragrafo 4.9).

Infine, nella maggior parte delle raccolte di dati, il giorno del rapporto che genera il concepimento è raramente osservabile, poiché in ogni ciclo sussistono generalmente più relazioni sessuali di cui una soltanto è causa della fecondazione: come conseguenza è opportuno modellare congiuntamente le probabilità giornaliere di concepire e quelle di procreare un maschio (o una femmina) dato il concepimento. Tuttavia, un approccio di verosimiglianza che coinvolga congiuntamente entrambi i fattori è difficilmente applicabile: per superare questa difficoltà siamo ricorsi alla verosimiglianza composita.

Oltre alle difficoltà riguardanti gli studi sui sessi, sono da considerarsi le imprecisioni proprie della nostra analisi. In primo luogo, i test e le regioni di confidenza derivati dai metodi asintotici utilizzati nel modello del Capitolo 4 sulla finestra di 9 giorni si basano su alcune approssimazioni, come descritto nel Paragrafo 4.9, sebbene anche con il metodo bootstrap sulla finestra intera siamo pervenuti alle medesime conclusioni. Inoltre il primo modello costruito assume l'indipendenza delle variabili tra i cicli, che abbiamo verifi-

cato essere implausibile, anche se i risultati del modello gerarchico relativi ai test sui  $\beta$  hanno fornito indicazioni analoghe. Infine nonostante nel modello gerarchico abbiamo corretto tale approssimazione, non abbiamo introdotto variabili, come l'età o il numero di precedenti gravidanze, per modellare la diversa distribuzione delle probabilità di concepire tra donne: questa potrebbe eventualmente essere una prosecuzione del nostro studio.

Per quanto concerne i risultati della nostra analisi, non abbiamo evidenziato una chiara influenza sul sesso del nascituro da parte dei giorni dei rapporti, tenendo in considerazione sia la non significatività dei test globali ed individuali sui  $\beta_k$  (ad eccezione di  $\beta_{-6}$ ) sia l'elevato tasso di errore di previsione del sesso di entrambi i modelli.

Nel valutare la significatività di  $\beta_{-6}$  si deve considerare il fatto che, sotto  $H_0$ , un test ogni 20 risulta significativo al livello del 5%, e che inoltre sono stati applicati numerosi test: di conseguenza la probabilità di ottenere significatività come quelle osservate aumenta considerevolmente. Pur volendo usufruire di questa indicazione per influenzare il sesso del nato nel caso in cui si volesse procreare un maschio, bisognerebbe tenere in considerazione il fatto che il giorno del picco del muco è identificato retrospettivamente: l'individuazione del giorno  $-6$  in anticipo sarebbe pertanto soggetta ad un errore di previsione ancora maggiore di quello già presente nella stima del "giorno 0". Dato che i giorni attorno a  $-6$  sono caratterizzati da stime non significative a favore dei maschi e anzi inferiori a 106/206 (sebbene non significativamente), anche qualora nel giorno  $-6$  fosse più probabile concepire un maschio sarebbe complesso riuscire a trarre beneficio da questa informazione. Inoltre la discontinuità tra le stime in tale giorno ed in quelli attorno è implausibile biologicamente: anche se ci fosse un fenomeno biologico per il quale nel giorno  $-6$  fosse più probabile concepire un maschio, questo riguarderebbe verosimilmente un periodo più duraturo.

Infine, da un punto di vista medico, non vi sono ragioni per supporre che 6 giorni prima del picco del muco sia più probabile concepire un bambino; le motivazioni biologiche sulle quali si basano le teorie di Billings e Shettles contraddicono anzi tale fenomeno.

In virtù di quanto detto sinora concludiamo la nostra analisi affermando che non sembra possibile influenzare neanche probabilisticamente il sesso del nato tramite i giorni dei rapporti.

# Appendice A

## Tabelle e grafici aggiuntivi

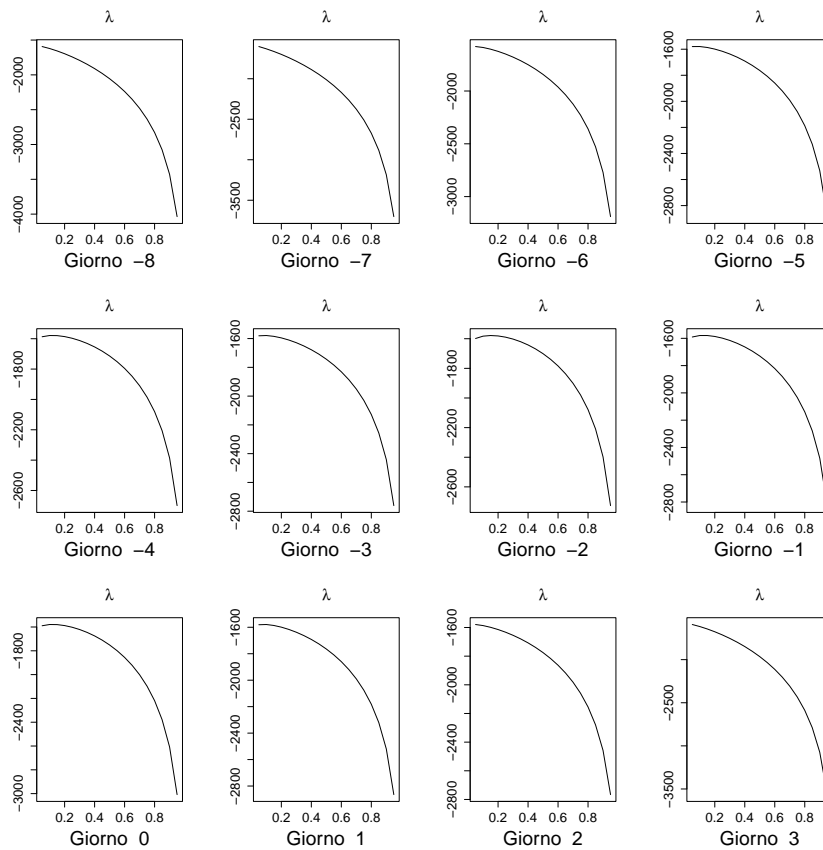


Figura A.1: Log-verosimiglianze profilo di  $\lambda_k$  del modello non gerarchico, con la formulazione di Barrett e Marshall in  $L_1$  calcolate sulle finestra  $[-8, 3]$

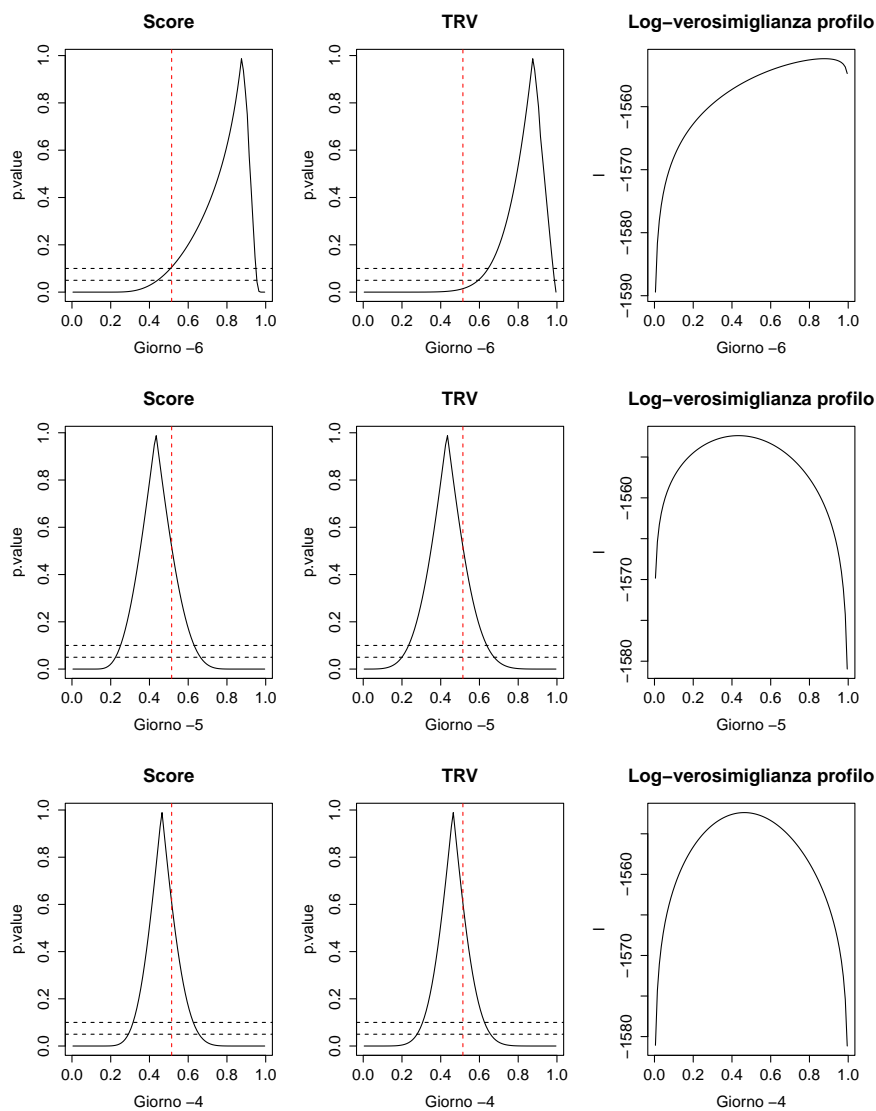


Figura A.2: Significatività del test del punteggio e del rapporto di verosimiglianza al variare del valore dei singoli parametri giornalieri  $\beta_k$  e relative log-verosimiglianze profilo, per  $\beta_{-6}$ ,  $\beta_{-5}$  e  $\beta_{-4}$ . Le linee tratteggiate orizzontali marcano i livelli di significatività del 5% e 10%, quella verticale indica il tasso atteso 106/206



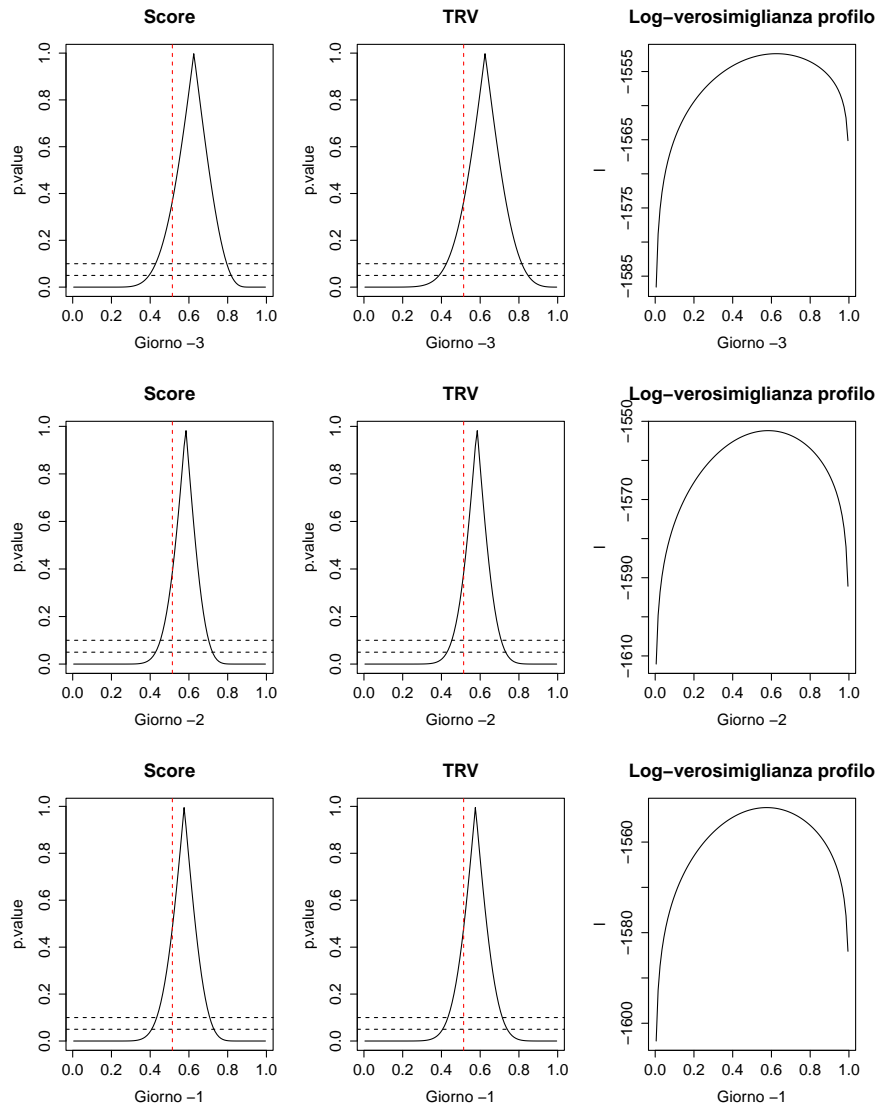


Figura A.3: Significatività del test del punteggio e del rapporto di verosimiglianza al variare del valore dei singoli parametri giornalieri  $\beta_k$  e relative log-verosimiglianze profilo per  $\beta_{-3}$ ,  $\beta_{-2}$  e  $\beta_{-1}$ . Le linee tratteggiate orizzontali marcano i livelli di significatività del 5% e 10%, quella verticale indica il tasso atteso 106/206

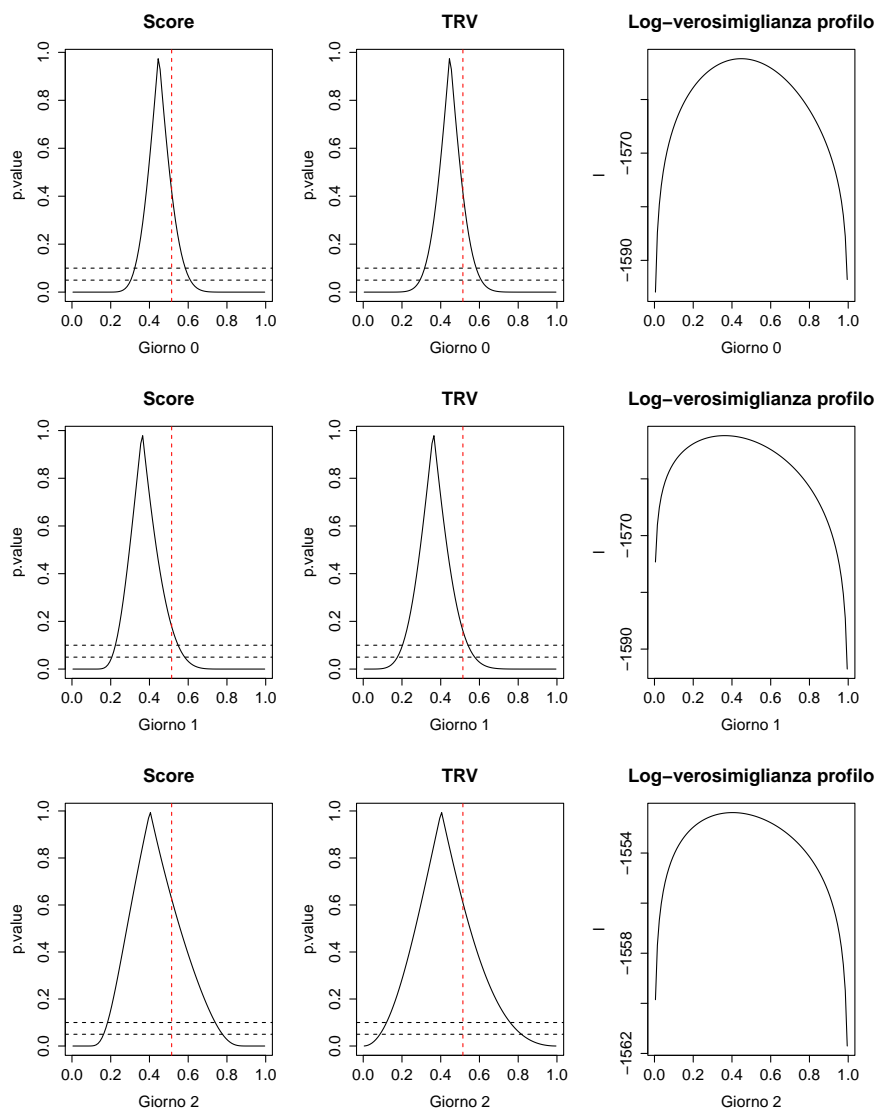


Figura A.4: Significatività del test del punteggio e del rapporto di verosimiglianza al variare del valore dei singoli parametri giornalieri  $\beta_k$  e relative log-verosimiglianze profilo per  $\beta_0$ ,  $\beta_1$  e  $\beta_2$ . Le linee tratteggiate orizzontali marciano i livelli di significatività del 5% e 10%, quella verticale indica il tasso atteso  $106/206$

	Modello originale		$\lambda_{jk}$ marginale in $l_2$		$\xi_j \equiv 1$ in $l_2$	
	Stima	Derivata	Stima	Derivata	Stima	Derivata
$\alpha_{-8}$	-4.9374	-0.0001	-4.9527	-0.0000	-4.9298	0.0000
$\alpha_{-7}$	-21.3157	0.0000	-21.4642	-0.0000	-19.4074	-0.0000
$\alpha_{-6}$	-2.2816	0.0001	-2.2971	0.0000	-2.1176	0.0000
$\alpha_{-5}$	-1.0893	-0.0000	-1.0844	0.0000	-0.9259	0.0000
$\alpha_{-4}$	1.9947	-0.0000	1.8251	0.0000	2.3764	0.0000
$\alpha_{-3}$	0.8842	-0.0000	1.0867	0.0000	1.6876	-0.0000
$\alpha_{-2}$	3.5170	-0.0000	3.5670	0.0000	4.4710	0.0000
$\alpha_{-1}$	2.1849	-0.0000	2.3214	-0.0000	2.9849	-0.0000
$\alpha_0$	2.0503	-0.0000	2.1974	-0.0000	2.8769	0.0000
$\alpha_1$	-0.2911	-0.0001	-0.2807	0.0000	0.0381	0.0000
$\alpha_2$	-2.7028	0.0000	-2.6588	0.0000	-2.6040	0.0000
$\alpha_3$	-5.8913	0.0006	-5.9250	-0.0000	-5.8786	-0.0000
$\phi$	0.0485	-0.0094	0.0479	0.0078	0.0419	0.0076
$\beta_{-8}$	0.5195	-0.0000	0.5640	0.0001	0.5186	0.0001
$\beta_{-7}$	0.9497	-0.0000	0.0001	-0.0000	0.4904	0.0000
$\beta_{-6}$	0.8829	-0.0002	0.8813	-0.0001	0.8868	-0.0000
$\beta_{-5}$	0.4412	0.0000	0.4204	-0.0003	0.4257	-0.0000
$\beta_{-4}$	0.4556	0.0000	0.4683	-0.0001	0.4859	-0.0001
$\beta_{-3}$	0.5979	-0.0000	0.6150	-0.0005	0.6064	0.0000
$\beta_{-2}$	0.5594	0.0003	0.5785	-0.0001	0.5769	-0.0000
$\beta_{-1}$	0.5884	-0.0002	0.5801	-0.0002	0.5815	0.0000
$\beta_0$	0.4399	-0.0000	0.4459	-0.0002	0.4471	-0.0000
$\beta_1$	0.3742	-0.0001	0.3623	-0.0001	0.3738	0.0000
$\beta_2$	0.4865	-0.0000	0.4333	-0.0000	0.4727	-0.0001
$\beta_3$	0.0001	-0.2123	0.0001	-0.1628	0.0001	-0.1979
Pr(conc.)						
-8	0.0067	-	0.0066	-	0.0066	-
-7	0.0000	-	0.0000	-	0.0000	-
-6	0.0535	-	0.0527	-	0.0551	-
-5	0.0956	-	0.0951	-	0.0936	-
-4	0.2164	-	0.2080	-	0.2075	-
-3	0.1736	-	0.1798	-	0.1844	-
-2	0.2720	-	0.2711	-	0.2739	-
-1	0.2236	-	0.2264	-	0.2274	-
0	0.2185	-	0.2219	-	0.2239	-
1	0.1269	-	0.1262	-	0.1273	-
2	0.0412	-	0.0422	-	0.0417	-
3	0.0027	-	0.0026	-	0.0027	-


Tabella A.1: Modello gerarchico: stime di massimo dei parametri della verosimiglianza composta con le tre formulazioni di  $\lambda_{jk}$  utilizzate in  $l_2$ , seguite dalla derivata della  $l_C$  rispetto ai singoli parametri calcolata nel massimo. Nella tabella Pr(conc.) fa riferimento alla probabilità di concepire dato un solo rapporto nel giorno  $k$

	$\lambda_{jk}$ marginale in $l_2$	
	Stima	Derivata
$\alpha_{-6}$	-2.2214	0.0499
$\alpha_{-5}$	-1.0270	-0.0732
$\alpha_{-4}$	2.1326	0.0128
$\alpha_{-3}$	1.2842	-0.0303
$\alpha_{-2}$	3.9559	-0.0081
$\alpha_{-1}$	2.6425	0.0206
$\alpha_0$	2.4741	-0.0143
$\alpha_1$	-0.1060	0.0462
$\alpha_2$	-2.5842	0.0281
$\phi$	0.0450	-0.6027
$\beta_{-6}$	0.8846	-0.0001
$\beta_{-5}$	0.4236	0.0001
$\beta_{-4}$	0.4688	-0.0001
$\beta_{-3}$	0.6139	0.0002
$\beta_{-2}$	0.5778	0.0001
$\beta_{-1}$	0.5787	-0.0001
$\beta_0$	0.4457	0.0006
$\beta_1$	0.3623	0.0001
$\beta_2$	0.4234	0.0000
Pr(conc.)		
-6	0.0537	-
-5	0.0939	-
-4	0.2099	-
-3	0.1795	-
-2	0.2720	-
-1	0.2278	-
0	0.2219	-
1	0.1280	-
2	0.0433	-

Tabella A.2: Modello gerarchico con finestra di 9 giorni e peso  $a = 0.6$ : stime di massimo dei parametri della verosimiglianza composta con  $\lambda_{jk}$  sostituito  $Pr(Y_{ij} = 1|X_{ij} = e_k)$  in  $l_2$ , seguite dalla derivata della  $l_C$  rispetto ai singoli parametri calcolata nel massimo. Nella tabella Pr(conc.) fa riferimento alla probabilità di concepire dato un solo rapporto nel giorno  $k$

# Appendice B

## Comandi R

Tutte le analisi sono state effettuate ricorrendo al software , versione 2.15.1. Riportiamo nel seguito i principali comandi utilizzati.

### L'estrazione dei database

Riportiamo le sole operazioni di estrazione dei dati Fertili, essendo quelle relative a Billing analoghe.

```
#leggo le informazioni sui cicli in Fertili:
ciclo.old <- read.fwf("fertili//ciclo.txt", widths= c(9, 1, 3, 1, 8, 2, 2, 2, 2, 1,
rep(c(4, 1, 1, 1), 99)), col.names=c("codice", "ingr", "n.cart", "p.spezz", "da-
ta", "lun.tot", "lun.pre", "picco", "period", "qualifi", paste(c("T","M","D","R"),
rep(1:99,each=4),sep="")))
#levo i cicli con picco non identificato e definisco Y:
ciclo <- ciclo.old[ciclo.old$picco != 0,]
Y <- ifelse(ciclo$lun.tot ==99, 1, 0)

#funzione per calcolare X (la matrice contenente i rapporti) su una fine-
stra variabile:
finestra <- function(x, inizio = -8, fine = 3){
  picco <- as.integer(x["picco"]) #picco del muco
  ris <- as.integer(x[paste("R",((picco + inizio ): (picco + fine)),sep="")])
  ris
}
#ora applico la funzione finestra con la finestra di 12 giorni e salvo sia il
codice della donna sia quello del ciclo:
X.conc_12 <- apply(ciclo, 1, finestra, inizio = -8, fine = 3)
X.conc_12 <- t(X.conc_12)
codice.ciclo <- ciclo$codice
```

```

ingresso.ciclo <- ciclo$ingr

#leggo le informazioni sulla donna in Fertili:
donna.old <-read.fwf("fertili//donna.txt", widths=c(9, 1, 1, 5, 6, 2, 8,
1, 1, 8, 8, 1, 1, 8, 1, 8), col.names=c("codice", "ingr", "tot_spez", "na-
scita", "nas_par", "grav_pre", "d_ul_ev", "tipo_u_e", "contr", "d_contr",
"d_us_pr", "mot_usc", "grav_con", "d_test, sesso, d_matr))
#escludo gli NA nei sessi e gli aborti che non mi consentono di conoscere
il sesso
donna <- donna.old[ (donna.old$sesso != 0 & donna.old$sesso != 3),]
#salvo le informazioni su X e Z per i soli cicli concepiti, ossia in  $L_2$ :
ciclo.incinte <- ciclo[ (ciclo$lun.tot == 99), ]
inizio = -8; fine = 3; ciclo = ciclo.incinte; n = 412;
X <- matrix(NA, nrow = n, ncol = fine - inizio + 1); Z <- rep(NA, n);
picco <- rep(NA, n); codice.sesso <- rep(NA,n); ingresso.sesso <- rep(NA,n);
k <- 1
for(i in 1:nrow(donna)){
  y <- donna[i,]
  riga <- NA
#cerco a quali cicli in  $L_1$  corrisponde il codice della donna presente in  $L_2$  (1
o più)
  indice <- which( ciclo$codice == donna$codice[i])
  if(length(indice) >0 ) #SE LA DONNA DI SESSO NON e IN CICLO
  for(j in 1:length(indice))
  if( (ciclo$lun.tot[indice[j]] ==99) & (ciclo$ingr[indice[j]] == donna$ingr[i])
)
  #dei cicli della stessa donna, prendo quello che ha portato al concepimento
(Y==1) e con stessa entrata: ciò mi identifica univocamente il ciclo
corrispondente al sesso in  $L_2$ .
  # riga rappresenta la posizione del ciclo  $i$ -esimo del database donna (con
i cicli di  $L_2$ ) nel database cicli (con i cicli di  $L_1$ )
  riga <- indice[j]
}
}
x <- ciclo[riga,] #salvo i rapporti del ciclo corrispondente
picco[k] <- as.integer(x["picco"]) # salvo il picco del muco
Z[k] <- as.integer(y["sesso"])
codice.sesso[k] <- as.integer(y["codice"])
ingresso.sesso[k] <- as.integer(y["ingr"])
for(colonne in 1:(fine - inizio + 1)){
if(picco[k] + inizio + colonne - 1 > 0){

```

```

    X[k,colonne] <- as.integer(x[paste(R,(picco[k] + inizio + colonne - 1),
sep="")])
  }
  else{
    X[k,colonne] <- 0
    print(picco nel giorno:); print(picco[i])
  }
}
k <- k+1
}
#i gemelli di 2 sessi (necessariamente indipendenti) vengono trsformati
in 2 record: 1 M ed 1 F.
X.sesso.na <- X; Z <- c(Z,0,1)
X.sesso.na <- rbind(X.sesso.na, X.sesso.na[Z.un == 6,], X.sesso.na[Z.un ==
6,])
codice.sesso <- c(codice.sesso, codice.sesso[Z.un == 6], codice.sesso[Z.un ==
6])
ingresso.sesso <- c(ingresso.sesso, ingresso.sesso[Z.un == 6], ingresso.sesso[
Z.un == 6])
X.sesso.na <- X.sesso.na[(Z!=6),]
codice.sesso <- codice.sesso[(Z!=6)]
ingresso.sesso <- ingresso.sesso[(Z!=6)]
Z <- Z[(Z!=6)]
#delle 3 coppie di gemelle dello stesso sesso (F) prendo solo 1 record:
trasformo in Z il valore 5 (2 gemelle) in 0, come se avessi una sola F.
Z <- ifelse(Z == 5, 0, Z)
#della coppia di gemelli dello stesso sesso (M) prendo solo 1 record: tra-
sformo in Z il valore 4 (2 gemelli) in 1, come se avessi un solo M.
Z <- ifelse(Z == 4, 1, Z)
#levo i dati senza rapporti nella finestra (zero) tra i cicli totali di L1:
X.conc.na <- X.conc_12; Y.na <- Y
zero <- apply(X.conc.na, 1, sum)
X.conc.na.sel <- X.conc.na[zero != 0,]
Y.na.sel <- Y.na[zero != 0]
codice.ciclo <- codice.ciclo[zero != 0]
ingresso.ciclo <- ingresso.ciclo[zero != 0]
#levo i dati senza rapporti nella finestra (zero) tra i cicli con sesso noto
di L2:
Z.na <- Z
zero.sesso <- apply(X.sesso.na, 1, sum)
X.sesso.na.sel <- X.sesso.na[zero.sesso != 0,]

```

```
Z.na.sel <- Z.na[zero.sesso != 0]
codice.sesso <- codice.sesso[zero.sesso != 0]
ingresso.sesso <- ingresso.sesso[zero.sesso != 0]
```

## La definizione della verosimiglianza composta

Riportiamo il solo codice relativo alla finestra di 12 giorni, essendo per altre è analogo:

```
# l1: ω non fornito: di default calcola il modello di SMH; se ω = 1,
calcola il modello di Barrett e Marshall.
vero_conc <- function(omega, l1, l2, l3, l4, l5, l6, l7, l8, l9, l10, l11, l12, X
= X.conc.na.sel, Y = Y.na.sel){
  p <- omega * (1 - ((1-l1)^X[,1]) * ((1-l2)^X[,2]) * ((1-l3)^X[,3]) * ((1-
l4)^X[,4]) * ((1-l5)^X[,5]) * ((1-l6)^X[,6]) * ((1-l7)^X[,7]) * ((1-l8)^X[,8]) *
((1-l9)^X[,9]) * ((1-l10)^X[,10]) * ((1-l11)^X[,11]) * ((1-l12)^X[,12]) )
  l <- sum(Y * log(p) + (1-Y) * log(1 - p) )
  l
}

# l2:
vero_sesso <- function(l1, l2, l3, l4, l5, l6, l7, l8, l9, l10, l11, l12, b1, b2, b3,
b4, b5, b6, b7, b8, b9, b10, b11, b12, X.sesso = X.sesso.na.sel, Z = Z.na.sel){
#w sono i λ* standardizzati che vanno a moltiplicare i β
DEN <- (X.sesso[,1] * l1 + X.sesso[,2] * l2 + X.sesso[,3] * l3 + X.sesso[,4]
* l4 + X.sesso[,5] * l5 + X.sesso[,6] * l6 + X.sesso[,7] * l7 + X.sesso[,8] * l8 +
X.sesso[,9] * l9 + X.sesso[,10] * l10 + X.sesso[,11] * l11 + X.sesso[,12] * l12)
w1 <- (X.sesso[,1] * l1) / DEN
w2 <- (X.sesso[,2] * l2) / DEN
w3 <- (X.sesso[,3] * l3) / DEN
w4 <- (X.sesso[,4] * l4) / DEN
w5 <- (X.sesso[,5] * l5) / DEN
w6 <- (X.sesso[,6] * l6) / DEN
w7 <- (X.sesso[,7] * l7) / DEN
w8 <- (X.sesso[,8] * l8) / DEN
w9 <- (X.sesso[,9] * l9) / DEN
w10 <- (X.sesso[,10] * l10) / DEN
w11 <- (X.sesso[,11] * l11) / DEN
w12 <- (X.sesso[,12] * l12) / DEN
theta <- b1 * w1 + b2 * w2 + b3 * w3 + b4 * w4 + b5 * w5 + b6 * w6
+ b7 * w7 + b8 * w8 + b9 * w9 + b10 * w10 + b11 * w11 + b12 * w12
l <- sum(Z * log(theta) + (1-Z) * log(1-theta))
```



```

    l
  }
  # l_C:
  vero_composita <- function(omega, l1, l2, l3, l4, l5, l6, l7, l8, l9, l10, l11,
  l12, X = X.conc.na.sel, Y = Y.na.sel, b1, b2, b3, b4, b5, b6, b7, b8, b9, b10,
  b11, b12, X.sesso = X.sesso.na.sel, Z = Z.na.sel, a = 1, b = 1, vero_conc =
  vero_conc, vero_sesso = vero_sesso){
    ris <- (a * vero_conc(omega, l1, l2, l3, l4, l5, l6, l7, l8, l9, l10, l11, l12,
  X, Y) + b* vero_sesso(l1, l2, l3, l4, l5, l6, l7,l8, l9, l10, l11, l12, b1, b2, b3,
  b4, b5, b6, b7, b8, b9, b10, b11, b12, X.sesso, Z))
    ris
  }
  # funzione da ottimizzare che richiama l_C; sotto la relativa ottimizzazio-
  ne:
  vero_BM <- function(par, X = X.conc.na.sel, Y = Y.na.sel, X.sesso =
  X.sesso.na.sel, Z = Z.na.sel, a = 1, b = 1, vero_composita = vero_composita,
  vero_conc = vero_conc, vero_sesso = vero_sesso){
    l1 <- par[1]; l2 <- par[2]; l3 <- par[3]; l4<- par[4]; l5 <- par[5]; l6 <-
  par[6]; l7 <- par[7]; l8 = par[8]; l9 = par[9]; l10 = par[10]; l11 = par[11]; l12
  = par[12];
    b1 <- par[13]; b2 <- par[14]; b3 <- par[15]; b4<- par[16]; b5 <- par[17];
  b6 <- par[18]; b7 <- par[19]; b8 = par[20]; b9 = par[21]; b10 = par[22]; b11
  = par[23]; b12 = par[24];
    omega <- 1;
    - vero_composita(omega, l1, l2, l3, l4, l5, l6, l7, l8, l9, l10, l11, l12, X, Y,
  b1, b2, b3, b4, b5, b6, b7, b8, b9, b10, b11, b12, X.sesso, Z, a, b, vero_conc
  = vero_conc, vero_sesso = vero_sesso)
  }

  k <- ncol(X.sesso.na.sel)
  ottimizza <- nlminb(start = rep(0.5,2*k), objective = vero_BM, a = 1, b =
  1, lower= rep(10^-4, 2*k), upper = rep(1 - 10^-4, 2*k),
  control = list(iter.max = 10000, abs.tol = 1e-20, eval.max = 10000))

```

## La stima di $a$ con convalida incrociata

Presentiamo solo il codice relativo alla convalida incrociata applicata sul primo modello, non gerarchico, in quanto quella adottata nel secondo, gerarchico, presenta solo minime modifiche.

#la funzione errore prende gli insiemi di dati con l'indicazione del codice del ciclo sia in  $L_1$  sia in  $L_2$  da escludere dall'insieme di stima; l'identificazione di tali record è effettuata dalla funzione leave, descritta sotto. In un altro ciclo, riportato in basso sotto la funzione leave, creo una griglia di  $a$  e per ogni valore di  $a$ , applico la convalida incrociata solo sugli  $n_2$  cicli in  $L_2$  e levo anche il corrispondente ciclo in  $L_1$ .

```
errore <- function(X.conc, Y, X.sesso, Z, A, riga, i){
  #definisco insieme di stima (.cv) e quello di verifica (.test)
  X.sesso.test <- X.sesso[i,]; Z.test <- Z[i]
  X.cv <- X.conc[-riga,]; Y.cv <- Y[-riga]; X.sesso.cv <- X.sesso[-i,]; Z.cv
  <- Z[-i]; k <- ncol(X.sesso.cv) #numero di giorni della finestra
  #ottimizzo il modello sul solo insieme di stima: con  $n_1 - 1$  ed  $n_2 - 1$  unità
  in  $L_1$  ed  $L_2$ 
  ottimizza <- nlminb(start = rep(0.5,2*k), objective = vero_BM, a = A,
  b = 1, X = X.cv, Y = Y.cv, X.sesso = X.sesso.cv, Z = Z.cv,
  lower= rep(10^-4, 13), upper = rep(1 - 10^-4, 13),
  control = list(iter.max = 10000, eval.max = 10000, abs.tol = 1e-20) )
  lambda <- ottimizza$par[1:k] # la stima di  $\lambda$  ottenuta dal modello meno
  l' $i$ -esima unita
  beta <- ottimizza$par[(k+1):(2*k)] # la stima di  $\beta$  ottenuta dal modello
  meno l' $i$ -esima unita
  somma_beta_X_lambda <- t(X.sesso.test) %*% (lambda * beta)
  somma_X_lambda <- t(X.sesso.test) %*% lambda
  thetai <- somma_beta_X_lambda / somma_X_lambda # la probabili-
  tà di concepire un M
  err <- abs(Z.test - thetai) # l'errore assoluto di stima, ancora non cate-
  gorizzato
  return(err)
}
```

#Funzione che seleziona il ciclo in  $L_1$  da escludere dalla convalida incrociata corrispondente a quello in  $L_2$  (ossia l'indice X o i nella funzione), quindi richiama la funzione errore indicando il codice, sia in  $L_1$  sia in  $L_2$  dell'unità componente l'insieme di verifica e non facente parte di quello di stima

```
leave <- function(X, A, codice.ciclo = codice.ciclo, codice.sesso = codice.sesso,
  ingresso.ciclo = ingresso.ciclo, ingresso.sesso = ingresso.sesso){
```

```
  i <- X
  riga <- NA #cerco a quali cicli in  $L_1$  corrisponde il codice della donna
  presente in  $L_2$  (1 o più)
  indice <- which(codice.ciclo == codice.sesso[i])
  for(j in 1:length(indice)){
```

```

    if( (Y[indice[j]] == 1) & (ingresso.ciclo[indice[j]] == ingresso.sesso[i]) ) {
      #dei cicli della stessa donna, prendo quello che ha portato al concepimento (Y==1) e con stessa entrata: ciò mi identifica univocamente il ciclo corrispondente al sesso in  $L_2$ .
      riga <- indice[j]
    }
  }
  #ora che ho identificato il ciclo da escludere sia da  $L_2$  (i) sia da  $L_1$  (riga);
  li levo ed applico la funzione errore
  if(is.na(riga)){
    print(riga e un NA, posizione i nel ciclo:)
    print(i)
    return(NA)# è un controllo per verificare che la funzione abbia individuato il ciclo.
  }
  else{
    return( errore(X.conc, Y, X.sesso, Z, A, riga, i) )
  }
}
#Applico sulle  $n_2$  (chiamato n nel ciclo) unità di  $L_2$  la funziona leave che calcola l'errore di stima con convalida incrociata sull' $i$ -esima unità esclusa dall'insieme di stima; quindi ciclo sui valori di  $a$  (chiamato A nel ciclo) tra 0.1 e 2 (1 ogni decimo).
errore <- matrix(NA, nrow= n, ncol=20)
for(A in seq(0.1, 2, length = 20) )
  errore[, A*20 ] <-sapply(1:n , leave, A = A)

```

## Il calcolo delle derivate prime analitiche e di J

Riportiamo le funzioni per il calcolo analitico delle derivate prime della log-verosimiglianza composta in  $\theta$  e di  $J(\theta)$ .

```

#Calcolo la derivata prima di  $l_C$  rispetto a  $\beta$ :
l_prima_beta <- function(par, X.sesso = X.sesso.na.sel, Z = Z.na.sel, a=1,
b=1, fine = 3, inizio = -8){
  n <- fine - inizio + 1
  lambda <- par[1:n]
  beta <- par[ (n+1): (2*n)]
  lambda_star <- matrix(NA, nrow = nrow(X.sesso), ncol= n)
  for(j in 1:n){
    lambda_star[,j] <- X.sesso[,j] * lambda[j]
  }
}

```

```

}
lambda_star <- lambda_star / apply(lambda_star, 1, sum) # calcolo
lambda_star
beta_lambda <- matrix(beta %*% t(lambda_star), nrow = length(Z),
ncol= n )# replicò la quantità calcolata e la inserisco in una matrice, per
facilità di calcolo nelle successive operazioni
contributi <- b * Z * lambda_star / beta_lambda - (1-Z) * lambda_star / (1 - beta_lambda)
# b è il peso di L2, nelle analisi sempre fissato ad 1
ris <- apply(contributi, 2, sum)
ris
}
# Calcolo la derivata prima di l_C rispetto a lambda:
l_prima_lambda <- function(par, omega = 1, X = X.conc.na.sel, Y =
Y.na.sel, X.sesso = X.sesso.na.sel, Z = Z.na.sel, a=1, b=1, fine = 3, inizio =
-8){
n <- fine - inizio + 1
lambda <- par[1:n]
beta <- par[(n+1):(2*n)]
l_X <- t((1-lambda)^t(X))
prod_2 <- apply(l_X, 1, prod)
prod_2 <- matrix(prod_2, nrow=length(Y), ncol=n, byrow=F)
prod_1 <- prod_2 / l_X
Y_mat <- matrix(Y, nrow= length(Y), ncol= n, byrow=F)
#il prodotto su j, ripetuto uguale in ogni riga, viene moltiplicato per
l'inverso del prodotto su k, ottenendo il prodotto su j ≠ k
# componenti individuali della derivata di l1:
pi <- omega * (1 - prod_2)
p_i <- X * omega * prod_1 #la derivata di pi
prima <- Y_mat * p_i/pi - (1-Y) * p_i/(1-pi)
# componenti individuali della derivata di l2:
beta_ <- matrix(beta, nrow= length(Z), ncol = n, byrow = T)
somma_beta_X_lambda <- matrix(X.sesso %*% (lambda * beta), nrow=
nrow(X.sesso), ncol= ncol(X.sesso), byrow=F)
somma_X_lambda <- matrix(X.sesso %*% lambda, nrow= nrow(X.sesso),
ncol= ncol(X.sesso), byrow=F)
thetai <- somma_beta_X_lambda / somma_X_lambda
theta_i <- X.sesso * ( (beta_ * somma_X_lambda -
somma_beta_X_lambda) / (somma_X_lambda^2) ) #la derivata di thetai
seconda <- Z * theta_i/thetai - (1-Z) * theta_i/(1-thetai)
#sommo i contributi individuali da l1 ed l2:

```

```

contributi <- a * apply(prima,2,sum) + b * apply(seconda,2,sum)
ris <- contributi
ris
}
# Calcolo  $J(\theta)$ :
J <- function(par, omega = 1, X = X.conc.na.sel, Y = Y.na.sel, X.sesso =
X.sesso.na.sel, Z = Z.na.sel, a=1, b=1, fine = 3, inizio = -8, ciclo_codice
= codice.ciclo, ciclo_ingresso = ingresso.ciclo, donna_codice = codice.sesso,
donna_ingresso = ingresso.sesso){ n<- fine - inizio + 1; lambda <- par[1:n];
beta <- par[(n+1):(2*n)]
# calcolo la derivata prima di  $l_C$  rispetto a  $\beta$ , calcoli analoghi a quelli
riportati sopra nella funzione l_prima_beta
lambda_star <- matrix(NA, nrow = nrow(X.sesso), ncol= n) for(j in
1:n){ lambda_star[,j] <- X.sesso[,j] * lambda[j] } lambda_star <- lambda_star / apply(lambda_star, 1, sum)
beta_lambda <- matrix(beta %*% t(lambda_star), nrow = length(Z),
ncol= n )
contributi <- b * Z * lambda_star / beta_lambda - (1-Z) * lambda_star/(1 - beta_lambda)
# fino a qui i calcoli sono identici a quelli in l_prima_beta
#ora calcolo la somma dei contributi appartenenti alle stesse donne (identificati dalla variabile codice)
nd <- length(table(factor(donna_codice))) # il numero di donne distinte
deri_beta <- matrix(NA, ncol = n, nrow = nd)
for(j in 1:n){
deri_beta[,j] <- tapply(contributi[,j], donna_codice, sum) #somma sugli
 $n_j$  cicli della donna  $j$ -esima
}
#adesso calcolo il prodotto matriciale delle componenti (vettori di lunghezza  $k$ ) su ciascuna delle  $n_d$  righe, quindi sommo le matrici ottenute:
beta_per_beta <- 0
for(i in 1:nrow(deri_beta)){
beta_per_beta <- beta_per_beta + deri_beta[i,] %*% t(deri_beta[i,])
#questa è una matrice, di dimensioni  $k * k$ 
} beta_per_beta <- beta_per_beta #questa è la parte della matrice  $J$ 
relativa alle componenti  $\beta\beta$ :  $\hat{J}_{\beta\beta}$ 
# ora calcolo invece la derivata prima di  $l_C$  rispetto a  $\lambda$ , calcoli analoghi
a quelli riportati sopra nella funzione l_prima_lambda
l_X <- t((1-lambda)^t(X))
prod_2 <- apply(l_X, 1, prod)
prod_2 <- matrix(prod_2, nrow=length(Y), ncol=n, byrow=F)

```

```

prod_1 <- prod_2 / l_X
Y_mat <- matrix(Y, nrow= length(Y), ncol= n, byrow=F)
#il prodotto su j, ripetuto uguale in ogni riga, viene moltiplicato per
l'inverso del prodotto su k, ottenendo il prodotto su j ≠ k
# componenti individuali della derivata di l1:
pi <- omega* (1 - prod_2)
p_i <- X * omega * prod_1 #la derivata di pi
prima <- Y_mat * p_i/pi - (1-Y) * p_i/(1-pi)
# componenti individuali della derivata di l2:
beta_ <- matrix(beta, nrow= length(Z), ncol = n, byrow = T)
somma_beta_X_lambda <- matrix(X.sesso %*% (lambda * beta), nrow=
nrow(X.sesso), ncol= ncol(X.sesso), byrow=F)
somma_X_lambda <- matrix(X.sesso %*% lambda, nrow= nrow(X.sesso),
ncol= ncol(X.sesso), byrow=F)
thetai <- somma_beta_X_lambda / somma_X_lambda
theta_i <- X.sesso * ( (beta_ * somma_X_lambda -
somma_beta_X_lambda) / (somma_X_lambda^2) ) #la derivata di thetai
seconda <- Z * theta_i/thetai - (1-Z) * theta_i/(1-thetai)
# fino qui ho calcolato le derivate di l1 ed l2 rispetto a λ
# ora devo sommare rispetto alle componenti appartenenti alla stessa
donna in ambedue le derivate
#infine sommare, solo per donne presenti sia in l1 sia in l2, le rispettive
componenti
#per l1:
nc <- length(table(factor(ciclo_codice))) # numero di donne distinte in
l1
deri_lambda <- matrix(NA, ncol = n, nrow = nc)
for(j in 1:n){
deri_lambda[,j] <- a * tapply(prima[,j], ciclo_codice, sum) #somma sugli
nj cicli in l1 della donna j-esima
}
#per l2:
nd <- length(table(factor(donna_codice))) # il numero di donne distinte
in l2
deri_lambda2 <- matrix(NA, ncol = n, nrow = nd)
for(j in 1:n){
deri_lambda2[,j] <- b * tapply(seconda[,j], donna_codice, sum) #somma
sui nj cicli in l2 della donna j-esima
}
#scrivo il codice della donna:
ciclo_codice_lambda <- names(tapply(prima[,1], ciclo_codice, sum))

```

```

donna_codice_lambda <- names(tapply(seconda[,1], donna_codice, sum))
#Ora devo sommare le due quantità appena calcolate per le donne in
comune tra  $l_1$  ed  $l_2$ : il ciclo è solo rispetto alle  $n_d$  donne in  $l_2$ ; se una donna
non ha cicli in  $l_2$  ma solo in  $l_1$  il suo contributo è solo nella prima componente
e non vi è nulla da aggiungere dalla seconda.
for(i in 1:nd){
  indice <- which(ciclo_codice_lambda == donna_codice_lambda[i])
  deri_lambda[indice,] <- deri_lambda[indice,] + deri_lambda2[i,]
}
#adesso calcolo il prodotto matriciale delle componenti (vettori di lun-
ghezza  $k$ ) su ciascuna delle  $n_d$  righe, quindi sommo le matrici ottenute:
lambda_per_lambda <- 0
for(i in 1:nrow(deri_lambda)){
  lambda_per_lambda <- lambda_per_lambda + deri_lambda[i,] %*%
t(deri_lambda[i,])#questa è una matrice, di dimensioni  $k * k$ 
}
lambda_per_lambda <- lambda_per_lambda #questa è la parte della
matrice  $J$  relativa alle componenti  $\lambda\lambda$ :  $\hat{J}_{\lambda\lambda}$ 
#Ora calcolo la parte di  $J$  che coinvolge  $\lambda$  e  $\beta$  congiuntamente, sommando
rispetto alle sole donne in comune tra  $l_1$  ed  $l_2$ 
ciclo_codice_lambda <- names(tapply(prima[,1], ciclo_codice, sum))
donna_codice_lambda <- names(tapply(seconda[,1], donna_codice, sum))
beta_per_lambda <- 0
for(i in 1:length(donna_codice_lambda)){#sommo solo sulle  $n_2$  donne
con cicli in  $l_2$ 
  indice <- which(ciclo_codice_lambda == donna_codice_lambda[i])
  for(j in 1:length(indice)){
    if( (Y[indice[j]] == 1) & (ciclo_ingresso[indice[j]] == donna_ingresso[i])){
      beta_per_lambda <- beta_per_lambda + deri_lambda[indice[j],] %*%
t(deri_beta[i,])
    }
  }
}
beta_per_lambda <- beta_per_lambda # questo è il blocco  $\hat{J}_{\lambda\beta}$ 
#ora pongo i blocchi calcolati in  $J$ :
matrice <- matrix(NA, nrow = 2*n, ncol = 2*n)
matrice[1:n, 1:n] <- lambda_per_lambda
matrice[(n+1):(2*n) , (n+1):(2*n)] <- beta_per_beta
matrice[1:n, (n+1):(2*n)] <- beta_per_lambda
matrice[(n+1):(2*n), 1:n] <- t(beta_per_lambda)

```

```

    matrice
  }

```

## I test asintotici nel modello non gerarchico

Riportiamo i codici relativi all'applicazione dei test asintotici sul sistema d'ipotesi globale (4.9); quelli relativi ai sistemi d'ipotesi individuali (4.10) sono analoghi con le dovute modifiche.

### Il test di Wald

```

beta0 <- rep(ottimizza_H0$par[k+1],k) #  $\hat{\beta}_{H_0}$ 
beta <- ottimizza$par[(k+1):(2*k)] #  $\hat{\beta}_c$ 
par <- ottimizza$par #  $\hat{\theta}_c$ 
G1 <- G(par, fine = 2, inizio = 6) #Calcolo la matrice  $\hat{G}(\theta)$  in  $\hat{\theta}_c$  (la
funzione G non è mostrata) sulla finestra [-6, 2]
par0 <- c(ottimizza_H0$par[1:k], beta0) #  $\hat{\theta}_{H_0}$ 
G0 <- G(par0, fine = 2, inizio = -6) #Calcolo la matrice  $\hat{G}(\theta)$  in  $\hat{\theta}_{H_0}$ 
sulla finestra [-6, 2]
wald1 <- t(beta - beta0) %*% solve(solve(G1)[(k+1):(2*k),(k+1):(2*k)])
%*% (beta - beta0) # test di Wald con G calcolata in  $\hat{\theta}_c$ 
wald0 <- t(beta - beta0) %*% solve(solve(G0)[(k+1):(2*k),(k+1):(2*k)])
%*% (beta - beta0) # test di Wald con G calcolata in  $\hat{\theta}_{H_0}$ 
#Valori p osservati:
1-pchisq(wald1, k-1)
1-pchisq(wald0, k-1)

```

### Il test del punteggio

```

deri_beta_H0 <- l_prima_beta( par0 , fine = 2, inizio = -6) # derivata di
 $l_C$  rispetto a  $\beta$  calcolata in  $\hat{\theta}_{H_0}$ 
H_theta0 <- hessian(vero_BM, par0) #  $\hat{H}(\theta)$  calcolata in  $\hat{\theta}_{H_0}$ 
H_theta <- hessian(vero_BM, par) #  $\hat{H}(\theta)$  calcolata in  $\hat{\theta}_c$ 
score_tot <- t(deri_beta_H0)
%*% solve(H_theta0)[(k+1):(2*k),(k+1):(2*k)]
%*% solve(solve(G0)[(k+1):(2*k),(k+1):(2*k)])
%*% solve(H_theta)[(k+1):(2*k),(k+1):(2*k)] %*% deri_beta_H0 #Test
del puteggio con matrici calcolate in  $\hat{\theta}_{H_0}$ 
score_tot2 <- t(deri_beta_H0)
%*% solve(H_theta)[(k+1):(2*k),(k+1):(2*k)]
%*% solve(solve(G1)[(k+1):(2*k),(k+1):(2*k)])

```



```

%% solve(H_theta)[(k+1):(2*k),(k+1):(2*k)]
%% deri_beta_H0 #Test del puteggio con matrici calcolate in  $\hat{\theta}_c$ 
#Valori  $p$  osservati:
1 - pchisq(score_tot, k-1)
1 - pchisq(score_tot2, k-1)

```

## Il test del rapporto di verosimiglianza

```

#TRV calcolato col metodo della correzione con matrici calcolate in  $\hat{\theta}_{H_0}$ :
  correzione_tot <- score_tot / (t(deri_beta_H0) %%
solve(H_theta0)[(k+1):(2*k),(k+1):(2*k)] %% deri_beta_H0)
  rapporto_tot <- -2 * ( ottimizza$objective - ottimizza_H0$objective )
* correzione_tot # il - all'inizio è per il fatto che i valori delle ottimizzazioni
sono riferiti a  $-l_C$ 
#TRV calcolato col metodo della correzione con matrici calcolate in  $\hat{\theta}_c$ :
  correzione_tot2 <- score_tot2 /
(t(deri_beta_H0) %% solve(H_theta)[(k+1):(2*k), (k+1):(2*k)] %% de
ri_beta_H0)
  rapporto_tot2 <- -2 * ( ottimizza$objective -
ottimizza_H0$objective ) * correzione_tot2 # il - all'inizio è per il fatto
che i valori delle ottimizzazioni sono riferiti a  $-l_C$ 
#relativi valori  $p$  osservati:
1- pchisq(rapporto_tot, k-1)
1- pchisq(rapporto_tot2, k-1)
#Alternativamente al metodo della correzione, calcolo la distribuzione
nulla asintotica esatta; anzitutto calcolo gli autovalori sui quali si basa tale
distribuzione:
  mat <- solve( solve(H_theta0)[(k+1):(2*k),(k+1):(2*k)] ) %%
(solve(G0)[(k+1):(2*k),(k+1):(2*k)])
  auto <- eigen(mat)$values
  mat2 <- solve( solve(H_theta)[(k+1):(2*k),(k+1):(2*k)] ) %%
(solve(G1)[(k+1):(2*k),(k+1):(2*k)])
  auto2 <- eigen(mat2)$values
#stimo la distribuzione nulla asintotica esatta con un campione di nume
rosità  $10^6$ , ossia 1 milione, ricorrendo agli autovalori derivati sia dalle matrici
calcolate in  $\hat{\theta}_{H_0}$  sia in  $\hat{\theta}_c$ 
  distr.nulla <- distr.nulla.rapporto(auto, N = 10^6)
  distr.nulla2 <- distr.nulla.rapporto(auto2, N = 10^6)
#dove la funzione distr.nulla.rapporto è uguale a:
distr.nulla.rapporto <- function(autovalori, N = 10^6){
  n <- length(autovalori)

```

```

ris <- rep(NA, N)
ris <- sapply(ris, function(x) { t(autovalori) %*% rchisq(n, df=1) })
ris
}
rapporto_tot <- -2 * ( ottimizza$objective - ottimizza_H0$objective )
#quindi confronto il TRV non corretto con tale distribuzione, ottenendo
le significatività osservate:
mean( c(distr.nulla, rapporto_tot) >= rapporto_tot)
mean( c(distr.nulla2, rapporto_tot) >= rapporto_tot)
#Comandi relativi al grafico in Figura 4.6
par(mfrow= c(1,2))
hist(distr.nulla, main = distribuzione nulla del test del rapporto di vero-
simiglianze matrici calcolate sotto H0)
abline(v = rapporto_tot, lty =4, pch=3)
hist(distr.nulla2, main = distribuzione nulla del test del rapporto di ve-
rosimiglianze matrici calcolate sotto H1)
abline(v = rapporto_tot, lty =4, pch=3)

```

## Il TRV bootstrap nel modello non gerarchico

Riportiamo i codici relativi all'applicazione del solo test bootstrap globale, essendo quelli individuali similari.

```

k <- ncol(X.sesso.na.sel); valore.nullo <- mean(Z.na.sel) # 0.5205811
#ottimizzo il modello nullo (sotto  $H_0$ ); la formulazione di vero_BM_H0
non è riportata poiché analoga a quella di vero_BM, dove però si impone
l'uguaglianza di tutti i  $\beta_k$ 
ottimizza_H0 <- nlminb(start = rep(0.5,k + 1), objective = vero_BM_H0,
a = 1, b = 1, lower= rep(10^-4, 13), upper = rep(1 - 10^-4, 13), control =
list(iter.max = 10000, abs.tol = 1e-20, eval.max = 10000))
#ottimizzo il modello non vincolato (sotto  $H_1$ ):
ottimizza <- nlminb(start = rep(0.5,2*k+1), objective = vero_BM, a =
1, b = 1, lower= rep(10^-4, 2*k), upper = rep(1 - 10^-4, 2*k), control =
list(iter.max = 10000, abs.tol = 1e-20, eval.max = 10000) )
#Salvo i parametri di  $\hat{\lambda}_{H_0}$ :
lambda_0 <- ottimizza_H0$par[1:k]; n1 <- nrow(X.conc.na.sel); k <-
ncol(X.conc.na.sel);
lambda_ <- matrix(lambda_0, nrow= n1, ncol =k, byrow = T) # replico
il valore del parametro in una matrice per semplicità di cacoli:
p <- 1 - ( apply( (1-lambda_)^X.conc.na.sel, 1, prod ) )
freq <- length(Z.na.sel) / sum(Y.na.sel)

```

```

#Definisco la funzione che sarà richiamata sotto e che, a partire dalle
probabilità stimate sotto  $H_0$  di concepire  $\hat{p}_{H_0}$  simula il campione dei concepimenti
Y.boot, quindi solo per i cicli nei quali si è simulato un concepimento
(Y.boot == 1) simula i sessi Z.boot con probabilità di procreare un maschio
fissata in ogni giorno pari alla frequenza di maschi osservata nel camione
(0.52); infine calcola il TRV su tale campione bootstrap.
rapporto.boot <- function(X, valore.nullo, p, X.conc = X.conc.na.sel){
  n1 <- nrow(X.conc)
  Y.boot <- rbinom(n = n1, size = 1, prob = p)
  cond <- (Y.boot == 1) & (rbinom(n = n1, size = 1, prob = freq) == 1)
  X.sesso <- X.conc[ cond ,]
  n2 <- sum(cond)
  Z.boot <- rbinom(n = n2, size = 1, prob = valore.nullo)
  # Ottimizzazione del modello non vincolato (sotto  $H_1$ ) sul campione
bootstrap appena costruito:
  ottimizza.boot <- nlm(bf = function(x) {
    start = rep(0.5, 2*k), objective = vero_BM, a
= 1, b = 1, Y=Y.boot, Z = Z.boot, X.sesso = X.sesso, lower= rep(10^-4,
2*k), upper = rep(1 - 10^-4, 2*k), control = list(iter.max = 10000, abs.tol
= 1e-20, eval.max = 10000) )
  # Ottimizzazione del modello vincolato (sotto  $H_0$ ) sul campione boo
tstrap appena costruito:
  ottimizza_H0.boot <- nlm(bf = function(x) {
    start = rep(0.5, k + 1), objective = ve
ro_BM_H0, a = 1, b = 1, Y=Y.boot, Z = Z.boot, X.sesso = X.sesso, lo
wer= rep(10^-4, k+1), upper = rep(1 - 10^-4, k+1), control = list(iter.max
= 10000, abs.tol = 1e-20, eval.max = 10000) )
  -2 * ( ottimizza.boot$objective - ottimizza_H0.boot$objective ) # il -
all'inizio è per il fatto che i valori restituiti dall'ottimizzazione sono riferiti a
 $-l_C$ 
}
  R <- ... #il numero di campioni bootstrap
  dati.boot.sel <- sapply( 1:R, FUN = rapporto.boot, valore.nullo = valo
re.nullo, p = p) # richiamo la funzione appena creata R volte
  rapp.tot <- -2 * ( ottimizza$objective - ottimizza_H0$objective ) #
TRV osservato
  p.value <- mean( dati.boot.sel >= rapp.tot) # la significatività osservata
  binom.test(p.value * length(dati.boot.sel), length(dati.boot.sel)) # l' $IC_{0.95}$ 
per il valore  $p$  calcolato con il metodo esatto basato sulla distribuzione
binomiale.
#Codice del grafico in Figura 4.9:
lwd <- 2

```

```

hist(dati.boot.sel,freq=F, breaks = 2*(0:30/2),col = light blue, main =
“, ylab = densità, xlab = TRV, cex.lab=1.2)
lines(density(distr.nulla), lwd=lwd, col =black)
lines(density(distr.nulla2), lwd=lwd, col =green)
curve( dchisq(x * correzione_tot, k-1) , xlim=c(1,30), lwd=lwd, add=T,
col = red)
curve( dchisq(x * correzione_tot2, k-1) , xlim=c(1,30), lwd=lwd, add=T,
col = blue)
legend(x=15, y = .1, lwd=lwd, col = c(black,green,red, blue ),legend
=c(Distribuzione 'esatta', H0,Distribuzione 'esatta', H1,Correzione, H0,Cor-
rezione, H1) )

```

## La formulazione del modello gerarchico

Illustriamo ora la formulazione del modello gerarchico e delle sue approssimazioni.

```

#l1:
vero_conc_ger <- function(par, X = X.conc.na.sel, Y = Y.na.sel){
  n <- nrow(X); k <- ncol(X); alpha <- par[1:k]; phi <- par[k+1]
  somma <- X %*% exp(alpha)
  p <- 1 - ( phi/(phi + somma) )^phi #la probabilità che ciclo i-esimo
della donna donna j-esima sia concepito, calcolata con la formula (5.1)
  l <- sum(Y * log(p) + (1-Y) * log(1 - p) )
  l
}
#funzione che calcola le probabilità di concepire giornaliera, in caso di
rapporto nel solo giorno k:
trasforma.lambda <- function(par, k = 12){
  phi <- par[k+1]; alpha <- par[1:k]
  l <- ( phi / (phi + exp(alpha)) )^phi
}
#funzione che calcola le probabilità di concepire individuali (in base ai
loro rapporti) ricorrendo alla (5.1):
p_1 <- function(par, X = X.conc.na.sel, k=k){
  phi <- par[k+1]; alpha <- par[1:k]
  somma <- X %*% exp(alpha)
  p <- 1 - ( phi/(phi + somma) )^phi
  p
}

```

```

#l2:
vero_sesso_ger <- function(par, X = X.conc.na.sel, Y = Y.na.sel, X.sesso
= X.sesso.na.sel, Z = Z.na.sel){
  n <- nrow(X.sesso); k <- ncol(X.sesso); alpha <- par[1:k]; phi <- par[k+1];
beta <- par[(k+2):(2*k+1)]
  #scrivo la funzione che contiene l'integrale della (5.4), da integrare nu-
mericamente:
  f <- function(x, X){
    xi <- x
    cond <- X # l'unità rispetto alla quale calcolo l'integrale
    lambda <- 1 - exp(-xi * exp(alpha) )
    X.sesso[cond, ] %*% (beta * lambda) / ( X.sesso[cond, ] %*% lambda )
* xi^(phi-1) * exp(-phi * xi)
  }
  integrale <- sapply(X = 1:n, FUN = function(X){
    int <- integrate(f = Vectorize(f), lower = 0, upper = Inf, X=X, stop.on.error
= F)
    if(int$message != OK){ print(c(X, int$message, int$value)) } # restitui-
sce l'integrale anche in caso di non convergenza, nel qual caso stampa un
messaggio d'errore
    return(int$value) })
  # ho appena calcolato e salvato in integrale l'integrale della (5.4).
  theta <- integrale * phi^phi / gamma(phi) # calcolo la probabilità di
procreare un maschio per ciascuna unità
  l <- sum(Z * log(theta) + (1-Z) * log(1-theta))
  l
}
# -lC:
vero_ger <- function(par, X = X.conc.na.sel, Y = Y.na.sel, X.sesso = X.sesso.na.sel,
Z = Z.na.sel, a = 1, b = 1){
  k <- ncol(X.sesso)
  ris <- (a * vero_conc_ger(par, X, Y) + b* vero_sesso_ger(par, X.sesso,
Z))
  - ris
}
# l2 con λjk sostituito dalla (5.2):
vero_sesso_marginale <- function(par, X = X.conc.na.sel, Y = Y.na.sel,
X.sesso = X.sesso.na.sel, Z = Z.na.sel){
  n <- nrow(X.sesso); k <- ncol(X.sesso); alpha <- par[1:k]; phi <- par[k+1];
beta <- par[(k+2):(2*k+1)]

```

lambda <- 1 - (phi / (phi + exp(alpha))) ^ phi #qui  $\lambda_{jk}$  viene sostituito dalla (5.2)

theta <- ifelse( apply(X.sesso,1,sum)==1, X.sesso %\*% beta, ((X.sesso) %\*% (lambda \* beta)) / ((X.sesso) %\*% lambda) ) #Utilizzo questa formulazione per arginare i problemi di underflow osservati nel corso dell'ottimizzazione

```
l <- sum(Z * log(theta) + (1-Z) * log(1-theta))
l
```

```
}
```

#Questa funzione si sostituisce poi a vero\_sesso\_ger in vero\_ger.

#  $l_2$  con  $\xi_j$  fissato pari ad 1:

```
vero_sesso_xi1 <- function(par, X = X.conc.na.sel, Y = Y.na.sel, X.sesso
= X.sesso.na.sel, Z = Z.na.sel){
```

```
  n <- nrow(X.sesso); k <- ncol(X.sesso); alpha <- par[1:k]; phi <- par[k+1];
beta <- par[(k+2):(2*k+1)]
```

```
  lambda <- 1 - exp(- exp(alpha) ) #fisso  $\xi_j = 1$ 
```

```
  somma_beta_X_lambda <- (X.sesso) %*% (lambda * beta)
```

```
  somma_X_lambda <- (X.sesso) %*% lambda
```

```
  theta <- somma_beta_X_lambda / somma_X_lambda
```

```
  l <- sum(Z * log(theta) + (1-Z) * log(1-theta))
```

```
  l
```

```
}
```

#Questa funzione si sostituisce poi a vero\_sesso\_ger in vero\_ger.

## Il TRV bootstrap nel modello gerarchico

Riportiamo ora il test bootstrap individuale applicato al sistema d'ipotesi (4.10) con  $a = 0.6$ , utilizzando la prima formulazione approssimata cui si è ricorso in  $l_2$ , ossia con  $\lambda_{jk}$  sostituito dalla (5.2).

#Definiamo  $l_C$  sotto  $H_0$ :

```
vero_marginale_H0_ind <- function(par, X = X.conc.na.sel, Y = Y.na.sel,
X.sesso = X.sesso.na.sel, Z = Z.na.sel, a = 1, b = 1, K){
```

```
  k <- ncol(X.sesso);
```

```
  if(K < k){
```

```
    par0 <- c(par[1:(k+K)], 106/206, par[(k+K+1):(2*k)])
```

```
  }else{
```

```
    par0 <- c(par[1:(k+K)], 106/206)
```

```
  } #fisso il parametro vincolato al valore 106/206, gli altri vengono ottimizati
```

```
  ris <- (a * vero_conc_ger(par0, X, Y) + b * vero_sesso_marginale(par0,
X, Y, X.sesso, Z))
```

```

- ris
}
k <- ncol(X.sesso.na.sel)
trv_marginale_ind <- rep(NA, k)
#ottimizzo il modello non vincolato (sotto  $H_1$ ):
ottimizza_marginale <- nlminb(start = rep(0.5, 2*k +1), objective =
vero_marginale, a = 0.6, b = 1, lower= c(rep(-Inf, k), 10^-4, rep(10^-4,
k)), upper= c(rep(Inf, k), Inf ,rep( 1 - 10^-4, k)), control = list(iter.max =
1000000, eval.max = 1000000))
#Ottimizzo i  $K = 9$  modelli vincolati (sotto  $H_0$ ) e salvo i relativi TRV
osservati:
ottimizza_marginale_ind_H0 <- vector(list, k)
for(i in 1:k){
ottimizza_marginale_ind_H0[[i]] <- nlminb(start = rep(0.5, 2*k ), ob-
jective = vero_marginale_H0_ind, a = 0.6, b = 1, K=i, lower= c(rep(-Inf,
k), 10^-4, rep(10^-4, k-1)), upper= c(rep(Inf, k), Inf ,rep( 1 - 10^-4, k-1)),
control = list(iter.max = 1000000, eval.max = 1000000))
trv_marginale_ind[i] <- -2* (ottimizza_marginale$objective - ottimizza_
marginale_ind_H0[[i]]$objective)
}
#Definisco la funzione che sarà richiamata sotto e che, a partire dalle
probabilità stimate sotto  $H_0$  di concepire  $\hat{p}_{H_0}$  simula il campione dei concepimenti
Y.boot, quindi solo per i cicli nei quali si è simulato un concepimento
(Y.boot == 1) simula i sessi Z.boot con probabilità di procreare un maschio
calcolata in base ai rapporti nella finestra seguendo la (5.3) con  $\lambda_{jk}$  sostituito
dalla (5.2) e con  $(\phi, \alpha, \beta)$  stime del modello vincolato sotto  $H_0$ ; infine calcola
il TRV.
TRV.uni <- function(X, X.conc, k, i, lambda, beta, p){
n1 <- nrow(X.conc)
Y.boot <- rbinom(n = n1, size = 1, prob = p)
cond <- (Y.boot == 1) & (rbinom(n = n1, size = 1, prob = freq) == 1)
X.sesso <- X.conc[ cond ,]
n2 <- sum(cond)
somma_beta_X_lambda <- X.sesso
somma_X_lambda <- X.sesso
prob <- somma_beta_X_lambda / somma_X_lambda
Z.boot <- rbinom(n = n2, size = 1, prob = prob)
# Ottimizzazione del modello non vincolato (sotto  $H_1$ ) sul campione
bootstrap appena costruito:
ottimizza.boot <- nlminb(start = rep(0.5, 2*k +1), objective =
vero_marginale, a = 0.6, b = 1, Y=Y.boot, Z = Z.boot, X.sesso = X.sesso,

```

```

lower= c(rep(-Inf, k), 10^-4, rep(10^-4, k)), upper= c(rep(Inf, k), Inf ,rep(
1 - 10^-4, k)), control = list(iter.max = 1000, eval.max = 1000))
  # Ottimizzazione del modello vincolato (sotto  $H_0$ ) sul campione boo-
strap appena costruito:
  ottimizza_H0.boot <- nlmnb(start = rep(0.5, 2*k ), objective =
vero_marginale_H0_ind, a = 0.6, b = 1, K=i,Y=Y.boot, Z = Z.boot,
X.sesso = X.sesso, lower= c(rep(-Inf, k), 10^-4, sarep(10^-4, k-1)), up-
per= c(rep(Inf, k), Inf,rep( 1 - 10^-4, k-1)), control = list(iter.max = 1000,
eval.max = 1000))
  trv <- -2 * ( ottimizza.boot$objective - ottimizza_H0.boot$objective )
  trv
}
  freq <- length(Z.na.sel) / sum(Y.na.sel) #frequenza di sessi osservati
rispetto al totale di concepimenti: sulla finestra di 12 giorni pari a 413/523
  R <- 500 #il numero di campioni bootstrap
  boot.ger.uni <- matrix(NA, nrow=R, ncol =k)
for(i in 1:k){
  par <- ottimizza_marginale_ind_H0[[i]]$par
  if(i < k){
    par0 <- c(par[1:(k+i)], 106/206, par[(k+i+1):(2*k)])
  }else{
    par0 <- c(par[1:(k+i)], 106/206)
  }
  alpha0 <- par0[1:k]; phi0 <- par0[k+1]; beta0 <- par0[(k+2):(2*k + 1)];
  somma0 <- X.conc.na.sel
  p <- 1 - (phi0/(phi0 + somma0 ))^phi0 #per beta fissati
  lambda0 <- 1 - (phi0/(phi0 + exp(alpha0) ))^phi0
  # richiamo la funzione appena creata R volte per ciascun parametro  $\beta_k$ 
vincolato sotto  $H_0$  boot.ger.uni[1:R,i] <- sapply(X = 1:R, FUN = TRV.uni,
X.conc =
X.conc.na.sel, k=k, i=i, lambda = lambda0, beta = beta0, p = p)
}
  print( sapply(1:k, function(x) mean( boot.ger.uni[,x] >= trv_marginale_ind[x]
) ) # calcolo i valori  $p$  osservati
#codice per il grafico in Figura 5.4:
  par(mfrow = c(3,3))
for(i in 1:k){
  hist(boot.ger.uni[,i], freq=F, breaks = 25,xlab = TRV, ylab = densità,
main = paste(Giorno ,i-7,sep="") ,cex.lab = 1.2)
  abline(v=trv_marginale_ind[i],lty =4, pch=3)
}
}

```



# Bibliografia

- Aerts M, Claeskens G (1999). Bootstrapping pseudolikelihood models for clustering binary data. *Annals of the Institute of Statistical Mathematics* **51**, 515-530.
- Aerts M, Claeskens G, Hens N, Molenberghs G (2002). Local multiple imputation. *Biometrika* **89**, 375-388.
- Barrett JC, Marshall J (1969). The risk of conception on different days of the menstrual cycle. *Population Studies* **23**, 455-461.
- Besag J (1974). Spatial interaction and the statistical analysis of lattice systems (with discussions). *Journal of the Royal Statistical Society, Series B* **34**, 192-236.
- Billings E, Westmore A (1980). The Billings method: controlling fertility without drugs or devices. *O'Donovan*.
- Carson SA (1988). Sex selectio: the ultimate in family planning. *Fertility and Sterility* **50**, 16-19.
- CIA, The World Factbook. *Central Intelligence Agency*. URL <https://www.cia.gov/library/publications/the-world-factbook/index.html> (01/11/2012)
- Colombo B, Masarotto G (2000). Daily fecundability: first results from a new data base. *Demographic Research* **3**, 5.
- Colombo B, Milon A, Passarin K, Scarpa B (2006). Cervical mucum symptom and daily fecundability: first results from a new database. *Statistical Methods in Medical Research* **15**, 161:180.
- Dedé A (1988). Il metodo Billings: marcatore della fertilità per la predeterminazione del sesso del concepito. *La regolazione naturale della fertilità oggi, Certezze e dubbi*, Congresso Internazionale Milano, 485-488
- Diasio RB, Glass RH (1971). Effects of pH on the migration of X and Y sperm. *Fertility and Sterility* **22**, 303-305.
- Dunson DB, Stanford JB (2005). Bayesian inferences on predictors of conception probabilities. *Biometrics* **61**, 126-133.
- Geys H, Molenberghs G, Ryan L (1999). Pseudolikelihood modeling of multivariate outcomes in developmental toxicology. *Journal of the American*

- Statistical Association* **94**, 734-745.
- Gilbert P (2012). numDeriv: Accurate Numerical Derivatives. *R package version 2012.3-1*. URL <http://CRAN.R-project.org/package=numDeriv>
- Gray RH (1991). Natural family planning and sex selection: fact or fiction?. *American Journal of Obstetrics and Gynecology* **165-6**, 2, 1982-1984.
- Guerrero R (1970). Sex ratio: a statistical association with the type and time of insemination in the menstrual cycle. *International Journal of Fertility* **15**, 221-225.
- Guerrero R (1974). Association of the type and time of insemination within the menstrual cycle with the human sex ratio at birth. *The New England Journal of Medicine* **291**, 1056-1059.
- Lindsay BG (1988). Composite likelihood methods. *Contemporary mathematics* **80**, 221-239.
- Mauro T (2011). Modelli statistici per la previsione del sesso di un neonato. URL <http://tesi.cab.unipd.it/35370/>
- McSweeney L (1993). A prospective study of sex preselection in Ondo, Nigeria, using the Billings ovulation method of natural family planning. *Bulletin of the Natural Family Planning Council of Victoria* **20**, 9-16.
- McSweeney L (2011). Successful sex pre-selection using natural family planning. *African Journal of Reproductive Health* **15**, 79-84.
- Molenberghs G, Verbeke G (2005). Models for discrete longitudinal data. *Springer Series in Statistics*, 189-202.
- Pace L, Salvan A, Sartori N (2011). Adjusting composite likelihood ratio statistics. *Statistica Sinica* **21**, 129-148.
- R Core Team (2012). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Rorvik DM, Shettles LB (1970). Your baby's sex: now you can choose. *Dodd, Mead (Ney York)*.
- Rotnitzky A, Jewell NP (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for clustering correlated data. *Biometrika* **77**, 485-497.
- Shettles LB (1960), Ovum Humanum.
- Schwartz D, MacDonald PMD, Heuchel V (1980). Fecundability, coital frequency and the viability of ova. *Population Studies* **34**, 397-400.
- Varin C, Reid N, Firth D (2011). An overview of composite likelihood methods. *Statistica Sinica* **21**, 5-42.
- Veit CR, Jewelewicz R (1988). Gender preselection: facts and myths. *Fertility and Sterility* **49**, 937-940.
- Wilcox AJ, Weinberg CR, Baird DD (1995). Timing of sexual intercourse in relation to ovulation. Effects on the probability of conception, survival of

the pregnancy and sex of the baby. *The New England Journal of Medicine* **333**, 1517-1521.

Williamson NE, Lean TH, Vengadasalam D (1978). Evaluation of an unsuccessful sex preselection clinic in Singapore. *Journal of Biosocial Science* **10**, 375-388.

Wikipedia, Gemelli (biologia). *Wikipedia, l'enciclopedia libera*.

URL [http://it.wikipedia.org/wiki/Gemelli\\_\(biologia\)](http://it.wikipedia.org/wiki/Gemelli_(biologia)) (01/11/2012)

Wikipedia, Politica del figlio unico. *Wikipedia, l'enciclopedia libera*.

URL [http://it.wikipedia.org/wiki/Politica\\_del\\_figlio\\_unico](http://it.wikipedia.org/wiki/Politica_del_figlio_unico) (01/11/2012)