

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

CORSO DI LAUREA
IN STATISTICA E GESTIONE DELLE IMPRESE

TESI DI LAUREA

**INTERVALLI DI CONFIDENZA PER IL
MODELLO BINOMIALE: IL CASO DI
ZERO SUCCESSI**

RELATORE: PROF. Alessandra Salvan

LAUREANDO: TOMMASO ROSSI

ANNO ACCADEMICO 2009/2010

Indice

Introduzione	III
1 Metodi frequentisti	1
1.1 Introduzione	1
1.2 Il metodo di Clopper e Pearson	2
1.2.1 La funzione di significatività (<i>significance function</i>)	4
1.3 Il metodo <i>score</i> di Wilson	7
1.4 Il metodo di Agresti e Coull	8
1.4.1 Una prima proposta di semplificazione	8
1.4.2 L'intuizione successiva	9
1.5 La soluzione di Lancaster: il metodo mid-P	9
2 Metodi bayesiani	11
2.1 Introduzione	11
2.2 L'approccio di Jovanovich e Levy	11
2.3 Distribuzioni a priori non informative	12
2.3.1 Jeffreys	13
3 Confronto tra i metodi	15
3.1 Introduzione	15
3.2 Confronto tra i limiti superiori di confidenza	15
3.3 Livelli di copertura	18
Conclusioni	21

Riferimenti bibliografici

23

Introduzione

In questa tesi si affronta il problema dell'inferenza sul parametro ignoto di una distribuzione binomiale, nell'eventualità estrema di 0 successi in n prove. In letteratura il caso è ancora oggetto di studio e ampiamente discusso, anche a causa della sua rilevanza a livello pratico, ad esempio in campo medico. Di particolare interesse è la soluzione proposta da Clopper e Pearson, la cosiddetta regola del 3, che considera $3/n$ l'estremo superiore di un intervallo di confidenza con livello approssimato 0.95. Verrà proposto il confronto tra questo ed altri metodi, alcuni dei quali basati su tecniche bayesiane.

I risultati presentati in questa tesi sono dovuti principalmente all'articolo di Frank Tuyl, Richard Gerlach e Kerrie Mengersen *The rule of three, its variants and extensions* pubblicato su *International Statistical Review* nel 2009.

Nella trattazione verrà considerato il caso $x = 0$, ma i medesimi ragionamenti possono essere fatti anche per $x = n$.

Capitolo 1

Metodi frequentisti

1.1 Introduzione

Nei problemi di inferenza capita spesso di dover costruire intervalli di confidenza per il parametro scalare $\theta \in \Theta \subseteq \mathbb{R}$ di distribuzioni discrete. Il metodo più comunemente utilizzato, il cosiddetto intervallo alla Wald, si basa sulla normalità asintotica dello stimatore di massima verosimiglianza $\hat{\theta}$,

$$\hat{\theta} \sim N(\theta, j(\hat{\theta})^{-1}),$$

dove $j(\hat{\theta})^{-1}$ è la matrice inversa dell'informazione osservata calcolata in $\hat{\theta}$. Supponendo di condurre un esperimento in cui si ottengono x successi in n prove e assumendo x realizzazione di

$$X \sim Bi(n, \theta), \theta \in (0, 1),$$

l'intervallo con livello approssimato $1 - \alpha$ costruito a partire dall'approssimazione

$$Z = \frac{\hat{\theta} - \theta}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} \sim N(0, 1)$$

è

$$\hat{\theta} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$$

dove $z_{1-\frac{\alpha}{2}}$ indica il quantile di una normale standard di livello $1 - \alpha/2$. Si può notare che questo metodo fornisce un intervallo non degenere solo per $\hat{\theta} \in (0,1)$, mentre si riduce ad un punto nei casi estremi $\hat{\theta} \in \{0,1\}$.

1.2 Il metodo di Clopper e Pearson

Una delle prime e più importanti soluzioni al problema è dovuta a Clopper e Pearson (1934). Essi considerano

$$X \sim Bi(n, \theta)$$

con funzione di probabilità

$$p(x; \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}$$

e utilizzano il test unilaterale con $H_0 : \theta = \theta_0$ e $H_1 : \theta < \theta_0$ con livello α . Un test di questo tipo rifiuta l'ipotesi H_0 se $x^{oss} \leq c$, con c tale che

$$Pr_{\theta_0}(X \leq c) \leq \alpha,$$

ovvero se

$$p - value = Pr_{\theta_0}(X \leq x^{oss}) \leq \alpha.$$

Quindi i valori di θ che vengono accettati dal test sono tutti quelli per cui

$$Pr_{\theta}(X \leq x^{oss}) > \alpha.$$

Poiché $Pr_{\theta}(X \leq u)$ con u fissato è, nel caso in esame, una funzione continua e decrescente di θ , i valori di θ che soddisfano la disequazione sono quelli in $[0, \theta_u]$ con θ_u soluzione di

$$Pr_{\theta}(X \leq x^{oss}) = \alpha.$$

Per costruzione, l'intervallo $[0, \theta_u]$ costituisce un intervallo di confidenza esatto per θ con livello almeno $1 - \alpha$.

Essendo

$$Pr_{\theta}(X \leq x) = \sum_{t=0}^x \binom{n}{t} \theta^t (1 - \theta)^{n-t},$$

quando $x^{oss} = 0$

$$Pr_{\theta}(X \leq 0; n, \theta) = Pr_{\theta}(X = 0; n, \theta) = (1 - \theta)^n.$$

Ponendo

$$(1 - \theta)^n > \alpha$$

e risolvendo per θ , si ha che

$$\theta < \theta_u = 1 - \alpha^{1/n}.$$

Grazie allo sviluppo di Taylor

$$\alpha^{1/n} = 1 + \frac{\ln \alpha}{n} + \frac{(\ln \alpha)^2}{2n^2} + \dots$$

si ottiene

$$1 - \alpha^{1/n} \approx -\frac{\ln \alpha}{n}.$$

Da questa approssimazione si giunge a quella che è stata definita la regola del 3. Scegliendo $\alpha = 0.05$ si ha $-\ln \alpha = 2.996 \approx 3$. L'intervallo approssimato sopra ricavato diventa quindi $[0, 3/n]$.

Nel confrontare le metodologie di seguito presentate si utilizzerà anche la definizione introdotta da Louis (1981) di

$$S = \lim_{n \rightarrow \infty} S_n = -\log(\alpha)$$

dove $S_n = n\theta_u$.

Lo stesso risultato si raggiunge considerando che, al divergere di n e per θ tendente a 0,

$$X \sim P(\lambda),$$

con $n\theta \rightarrow \lambda > 0$, di media $\lambda = n\theta$ e varianza $\lambda(1 - \theta/n) \approx \lambda$. Di nuovo

$$Pr_{\theta}(X \leq x^{oss}) = Pr_{\theta}(X \leq 0) = e^{-\lambda} > \alpha$$

quindi la disuguaglianza

$$e^{-\lambda} = e^{-n\theta} \geq \alpha,$$

da

$$-n\theta \geq \log(\alpha)$$

da cui si torna al risultato precedente

$$\theta_u = -\frac{\ln \alpha}{n}.$$

1.2.1 La funzione di significatività (*significance function*)

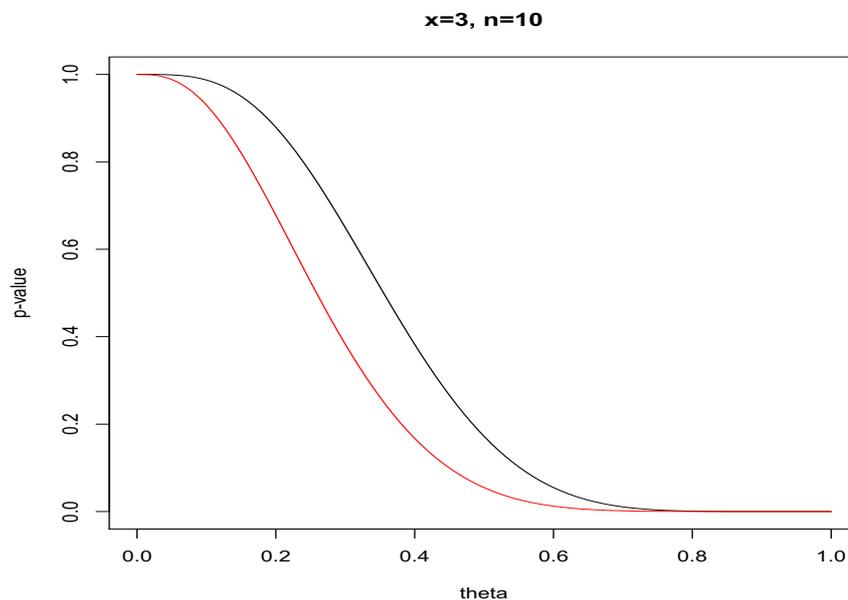


Figura 1.1: Funzione *p-value* per una distribuzione binomiale con $x=3$ successi in $n=10$ prove.

Richiamando brevemente i concetti espressi in precedenza, per definizione abbiamo che

$$p - value = Pr_{\theta_0}(X \leq x^{oss}).$$

Partendo da questo punto, Fraser, Reid e Wonk (2004) definiscono la *funzione di significatività* o *funzione p-value* come la probabilità che X sia minore o uguale a x^{oss} al variare di θ in $[0,1]$.

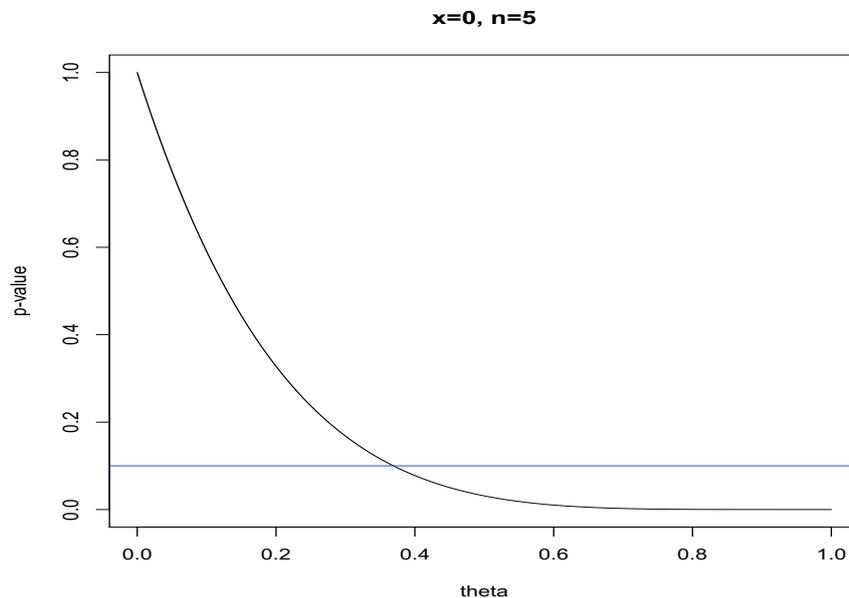


Figura 1.2: Funzione *p-value* per una distribuzione binomiale con $x=0$ successi in $n=5$ prove.

Se X è discreta, la funzione di significatività è data dall'intervallo

$$p(\theta) = [Pr_{\theta}(X < x^{oss}), Pr_{\theta}(X \leq x^{oss})].$$

Ad esempio, per $x^{oss} = 3$ e $n = 10$ abbiamo come risultato il grafico in Figura 1.1. Esso mostra l'intervallo $p(\theta)$ dove la curva rossa rappresenta l'estremo inferiore dell'intervallo e quella nera l'estremo superiore.

Il caso di interesse in questa tesi è quello di $x^{oss} = 0$ e viene presentato nel grafico in Figura 1.2. In questa particolare situazione l'estremo inferiore dell'intervallo è 0, quindi l'unica curva rappresentata è quella relativa all'estremo superiore $(1 - \theta)^n$. La *significance function* mostrata è quella di una distribuzione binomiale in cui $x = 0$ e $n = 5$ prove. Si osserva come al crescere del valore di θ diminuisca il *p-value*. Infatti più aumenta la probabilità θ di successo, meno è probabile ottenere $x = 0$ successi. Al contrario, per θ che si avvicina a 0, abbiamo una probabilità maggiore di osservare

$x = 0$. Ci si collega in maniera diretta al metodo di Clopper e Pearson se nel grafico si fissa un α e, dall'intersezione con la curva, si ricava il θ_u corrispondente.

Vengono inseriti anche i grafici della funzione di significatività relativi a $n = 10$ e $n = 30$. Come si può osservare, la curva si abbassa più velocemente, coerentemente con quanto già esposto.

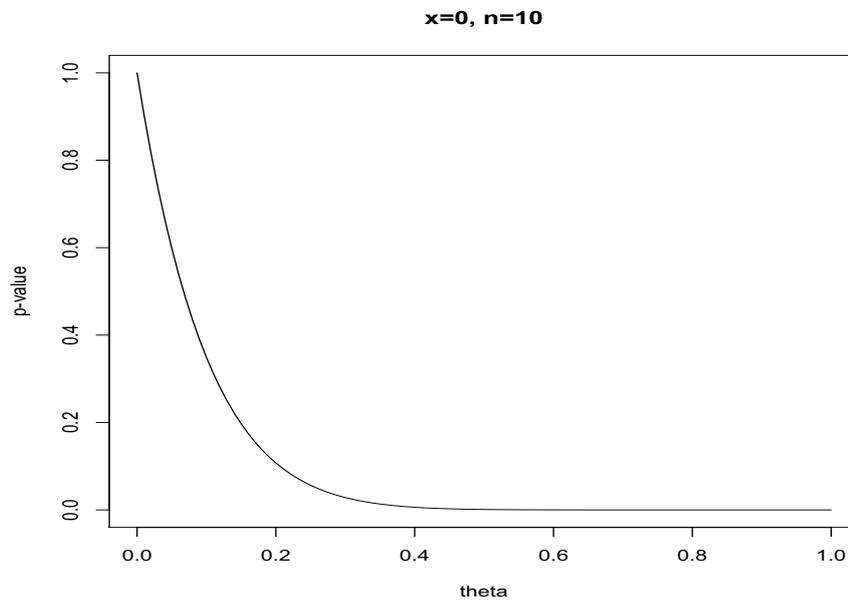


Figura 1.3: Funzione *p-value* per una distribuzione binomiale con $x=0$ successi in $n=10$ prove.

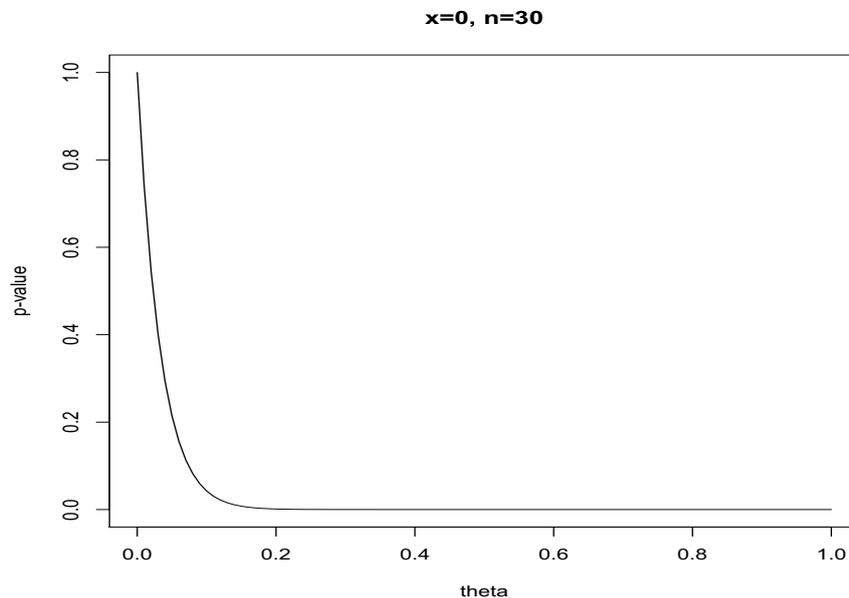


Figura 1.4: Funzione *p-value* per una distribuzione binomiale con $x=0$ successi in $n=30$ prove.

1.3 Il metodo *score* di Wilson

Un altro metodo, detto *score*, per affrontare il problema è stato proposto da Wilson (1927). Il metodo si basa sulla quantità pivotale approssimata da una normale standard

$$Z_u = \frac{l_*(\theta)}{\sqrt{i(\theta)}} = \frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}},$$

dove

$$l_*(\theta) = \frac{x}{\theta} - \frac{n-x}{1-\theta}$$

è la funzione *score* del modello e

$$i(\theta) = \frac{n}{\theta(1-\theta)}$$

rappresenta l'informazione attesa.

Sotto θ , $Z_u \sim N(0, 1)$ e per n sufficientemente grande, il test unilaterale con

$H_0 : \theta = \theta_0$ e $H_1 : \theta < \theta_0$ con livello α rifiuta l'ipotesi H_0 se

$$\frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \leq -z_{1-\alpha}.$$

L'estremo superiore dell'intervallo di livello $1-\alpha$ è dato dal maggiore tra le radici dell'equazione di secondo grado

$$\left(1 + \frac{z_{1-\alpha}^2}{n}\right) \theta^2 - 2 \left(\hat{\theta} + \frac{z_{1-\alpha}^2}{2n}\right) \theta + \hat{\theta}^2 = 0$$

che risultano

$$\theta_{1,2} = \frac{\hat{\theta} + \frac{z_{1-\alpha}^2}{2n} \pm z_{1-\alpha} \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n} + \frac{z_{1-\alpha}^2}{4n^2}}}{1 + \frac{z_{1-\alpha}^2}{n}}.$$

Essendo $\hat{\theta} = x/n$ e considerando il caso $x = 0$, tale estremo diventa

$$\theta_u = \frac{z_{1-\alpha}^2}{n + z_{1-\alpha}^2}.$$

1.4 Il metodo di Agresti e Coull

1.4.1 Una prima proposta di semplificazione

Prendendo spunto dall'idea di Wilson di agire sull'intervallo alla Wald per costruire intervalli quando $x = 0$, Agresti e Coull (1998) propongono di utilizzare il punto medio tra le radici sopra trovate

$$\tilde{\theta} = \frac{2x + z_{1-\alpha}^2}{2(n + z_{1-\alpha}^2)}$$

al posto del semplice $\hat{\theta}$ nell'espressione

$$\theta_u = \tilde{\theta} + z_{1-\alpha} \sqrt{\frac{\tilde{\theta}(1-\tilde{\theta})}{n}}.$$

In questo caso si ha

$$\begin{aligned} &= \frac{2x + z_{1-\alpha}^2}{2(n + z_{1-\alpha}^2)} + z_{1-\alpha} \sqrt{\frac{1}{n} \cdot \left(\frac{2x + z_{1-\alpha}^2}{2(n + z_{1-\alpha}^2)}\right) \cdot \left(1 - \frac{2x + z_{1-\alpha}^2}{2(n + z_{1-\alpha}^2)}\right)} \\ &= \frac{2x + z_{1-\alpha}^2}{2(n + z_{1-\alpha}^2)} + z_{1-\alpha} \sqrt{\frac{2x + z_{1-\alpha}^2}{4(n + z_{1-\alpha}^2)^2} \cdot \frac{2n - 2x + z_{1-\alpha}^2}{n}}, \end{aligned} \quad (1.1)$$

che per $x = 0$ e n crescente è approssimabile a

$$\theta_u = \frac{z_{1-\alpha}^2(1 + \sqrt{2})}{2(n + z_{1-\alpha}^2)}.$$

Nel seguito della trattazione questo metodo verrà indicato con $AC(z)$.

1.4.2 L'intuizione successiva

Agresti e Coull (1998) si concentrano su un'ulteriore semplificazione del loro metodo con livello di confidenza approssimato 0.95. L'intuizione è quella che 1.96 (il valore corrispondente al quantile della normale standard per un livello di confidenza 0.95 nelle ipotesi bilaterali) non è molto distante 2, il cui valore del quantile della normale standard corrispondente è 0.9544998. Aggiungendo "2 successi e 2 insuccessi" arrivano a stabilire come estremo superiore dell'intervallo

$$\frac{2 + \sqrt{2}z_{1-\alpha}}{n + 4}.$$

Questo metodo sarà chiamato $AC(2)$. Bisogna precisare che esso vale solo nel caso di intervalli riferiti ad ipotesi bilaterali. Infatti, nel caso in cui si decidesse di applicarlo ad ipotesi unilaterali, l'estremo superiore dell'intervallo si baserebbe su $\theta_{0.05} = 1.64$ e quindi cadrebbe l'ipotesi $\theta_\alpha \approx 2$. Nonostante questo, viene incluso nelle successive analisi, dal momento che i risultati ottenuti con questo metodo sembrano funzionare bene nel caso $x = 0$.

1.5 La soluzione di Lancaster: il metodo mid-P

Lancaster (1949, 1961) propone quella che è stata definita la regola del 2.3, utilizzando il metodo mid-P. Questo si basa su un procedimento analogo a quello utilizzato da Clopper e Pearson, con la sola differenza che si

considera

$$\text{mid} - p - \text{value} = \frac{1}{2}Pr_{\theta}(X = x) + \sum_{t=0}^{x-1} Pr_{\theta}(X = t)$$

e si determina θ_u imponendo

$$\text{mid} - p - \text{value} = \alpha.$$

Per $x = 0$ si ricava

$$\begin{aligned}\frac{1}{2}(1 - \theta)^n &= \alpha, \\ (1 - \theta)^n &= 2\alpha,\end{aligned}$$

da cui l'estremo superiore dell'intervallo

$$\theta_u = 1 - (2\alpha)^{1/n}.$$

Si può notare che l'estremo così costruito è più piccolo del precedente ricavato con il metodo di Clopper e Pearson. Per n grande e $\alpha = 0.05$, eseguendo le stesse approssimazioni precedenti tramite lo sviluppo di Taylor, si trova $-\ln \alpha = 2.303 \approx 2.3$.

Capitolo 2

Metodi bayesiani

2.1 Introduzione

Nel presente capitolo si presentano alcune soluzioni bayesiane al problema oggetto di questa tesi. È doveroso precisare che tutt'ora non si è giunti ad un parere unanime sull'argomento, nonostante i metodi proposti siano molti e con caratteristiche differenti. Il punto d'attrito principale sembra essere quello riguardante l'informazione a priori da utilizzare e quanto essa debba essere informativa, anche se, nei casi in cui n è molto grande, si è visto come quest'ultima problematica non sussista, poiché i dati attutiscono l'influenza dell'informazione a priori.

2.2 L'approccio di Jovanovich e Levy

Ottenere $x=0$ successi ci fa intuire che θ sarà piccolo e quest'informazione viene usata da Jovanovic e Levy (1997) per costruire una versione della regola del 3 bayesiana. Si considera una distribuzione a priori

$$\theta \sim \text{beta}(a, b)$$

con densità

$$\pi(\theta; a, b) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a, b)}, \theta \in (0, 1), a, b > 0,$$

dove $B(a, b)$ rappresenta la funzione *beta* di parametri $a, b > 0$, e l'aggiornamento a posteriori

$$\theta|x \sim \text{beta}(x + a, n - x + b).$$

Ora, osservando che è ragionevole pensare che θ sia prossimo allo 0, essi pongono $a=1$ e $b > 1$, valori per i quali la densità a priori ha moda in 0. Inoltre, valori per a diversi da 1 porterebbero la densità ad avere un massimo locale positivo, che non può essere giustificato senza ulteriori informazioni.

L'estremo superiore dell'intervallo $[0, \theta_u]$ si ricava da

$$F(\theta_u|x = 0, a = 1, b) = 1 - \alpha,$$

$$1 - (1 - \theta_u)^{n+b} = 1 - \alpha,$$

$$(1 - \theta_u)^{n+b} = \alpha,$$

per cui

$$\begin{aligned} \theta_u &= 1 - \alpha^{\frac{1}{n+b}} \\ &\approx 1 - \left(1 + \frac{\ln \alpha}{n+b} + \frac{(\ln \alpha)^2}{2(n+b)^2} + \dots\right) \\ &\approx -\frac{\ln \alpha}{n+b} \\ &\approx \frac{3}{n+b}. \end{aligned} \tag{2.1}$$

Una delle critiche principali mosse agli Autori da Tuyl, Gerlach e Mengersen (2008) è quella di non fornire sufficienti spiegazioni sul perché a non possa essere minore di 1.

2.3 Distribuzioni a priori non informative

Da una diversa prospettiva si muovono quanti ritengono più adatto che l'approccio bayesiano si fondi su distribuzioni a priori non informative. La

loro convinzione è che una distribuzione non informativa dovrebbe poggiare su una distribuzione iniziale basata non sulle informazioni a disposizione di chi conduce l'esperimento, ma tenendo conto solamente della struttura statistica dell'esperimento. La simmetria è quindi una caratteristica imprescindibile delle distribuzioni a priori non informative.

Secondo Tuyl, Gerlach e Mengersen (2009) l'unica distribuzione $beta(a, b)$ in accordo con la regola del 3 è la $beta(1,1)$. L'ipotesi di porre $a = b = 1$ fu di Bayes e successivamente confermata da Laplace. Il motivo di questa scelta è che la distribuzione a priori deve essere, secondo loro, uniforme, poiché a priori ogni evento deve essere considerato equiprobabile ed infatti $beta(1, 1) \sim U(0, 1)$. Essi definiscono, pertanto, il limite superiore dell'intervallo $3/(n + 1)$.

2.3.1 Jeffreys

L'ultimo metodo che viene confrontato con i precedenti è quello di Jeffreys, il quale ponendo $a=1/2$ e $b=1/2$ giunge a definire la regola del 1.9.

A rigore, questo metodo va catalogato tra le distribuzioni a priori non informative, ma diversi autori criticano questa posizione. Ad esempio Tuyl, Gerlach e Mergensen (2008) sostengono che la scelta di $a < 1$ sia eccessivamente informativa perché sposta l'estremo superiore dell'intervallo di credibilità vicino allo 0, riducendo l'importanza dei dati osservati.

Capitolo 3

Confronto tra i metodi

3.1 Introduzione

In questo capitolo verranno confrontati i metodi precedentemente descritti per tentare di stabilire se uno di essi sia il più conveniente.

Nella tabella sottostante, proposta anche da Tuyl, Gerlach e Mergensen (2009), vengono presentati i risultati ottenuti con i diversi metodi fin qui presentati e si fornisce l'espressione di S per ognuno di essi e i limiti superiori corrispondenti a diversi livelli di confidenza, prendendo come riferimento il metodo di Clopper e Pearson.

3.2 Confronto tra i limiti superiori di confidenza

Con riferimento alla Tabella 3.1 si possono fare alcune osservazioni riguardo ai differenti metodi. Ad esempio si vede come entrambi gli approcci di Agresti e Coull portino ad ottenere, a parità di livello di confidenza scelto, un limite superiore più elevato rispetto a quello di riferimento di Clopper e Pearson (eccezion fatta per $1 - \alpha = 0.90$, dove il secondo metodo di Agresti e Coull fornisce un limite pari a 1.98 rispetto al 2.30 di Clopper e Pearson e per $1 - \alpha = 0.999$ dove il primo metodo fornisce un limite di 6.37, leggermente inferiore al 6.91 di Clopper e Pearson). La differenza è

Metodo	S	0.90	0.95	0.975	0.99	0.999
Clopper e Pearson, beta(1,1)	$-\log(\alpha)$	2.30	3.00 (0.95)	3.69	4.61	6.91
Score	z_α^2	1.64	2.71 (0.93)	3.84	5.41	9.55
Agresti e Coull (2)	$2 + \sqrt{2}z_\alpha$	3.81	4.33 (0.98)	4.77	5.29	6.37
Agresti e Coull (z)	$z_\alpha^2(1 + \sqrt{2})/2$	1.98	3.27 (0.96)	4.64	6.53	11.53
Mid-P	$-\log(2\alpha)$	1.61	2.30 (0.90)	3.00	3.91	6.21
Jeffreys	$z_{\alpha/2}^2/2$	1.35	1.92 (0.85)	2.51	3.32	5.41

Tabella 3.1: Limiti superiori di confidenza relativi ai metodi descritti. $S = \lim_{n \rightarrow \infty} S_n$. In evidenza il metodo di Clopper e Pearson. Tra parentesi sono indicati i livelli di copertura effettivi per n divergente.

particolarmente evidente per $\alpha = 0.001$ dove il metodo Agresti e Coull(z) genera un limite maggiore sia a quello di Clopper e Pearson, sia a quello di ogni altro metodo presentato.

Per quanto riguarda il metodo *score* di Wilson si nota che fornisce limiti superiori maggiori solo per $1 - \alpha > 0.9694$.

L'unico metodo frequentista che fornisce degli estremi inferiori a quelli trovati con il metodo Clopper e Pearson per ogni livello di confidenza $1 - \alpha$ è quello basato sul *mid-P*.

Infine, si vede come il metodo di Jeffreys si comporti in maniera forse troppo radicale. I limiti superiori forniti sono al di sotto di quelli presi come riferimento, ma la differenza tra i due metodi è forse eccessiva.

Secondo Tuyl, Gerlach e Mengersen (2009) ciò che è ragionevole chiedere ad un metodo approssimato è che, per ogni x , il limite superiore dell'intervallo sia minore di quello fornito dal metodo di Clopper e Pearson e per questo le tecniche adottate da Agresti e Coull e da Wilson, che si basano su approssimazioni alla normale, possono essere considerate inadeguate quando $x = 0$.

Alle stesse conclusioni si giunge osservando il grafico in figura 3.1, dove viene mostrato S come una funzione continua del livello di confidenza

$1 - \alpha$. Si vede chiaramente come la curva corrispondente al primo metodo di Agresti e Coull sia costantemente sopra la curva di riferimento di Clopper e Pearson, intersecandola solo nella parte destra, in corrispondenza di $1 - \alpha \approx 0.9974$. È facile notare anche la rapida crescita della curva relativa al metodo Agresti e Coull(z) che interseca prima la curva di Clopper e Pearson vicino a $1 - \alpha = 0.9322$ e poi quella di Agresti e Coull(2) a destra rispetto a $1 - \alpha = 0.977$. Allo stesso modo è evidente la distanza tra la curva di Jeffreys e quella di riferimento.

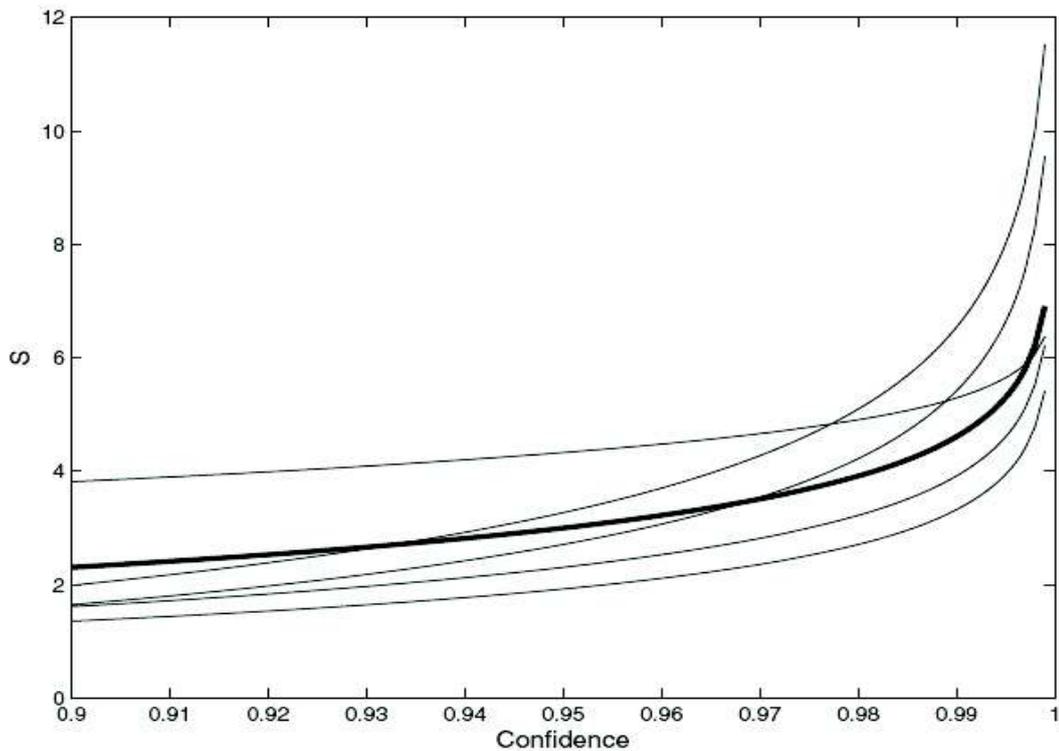


Figura 3.1: S come funzione del livello di confidenza ($1 - \alpha$). Dall'alto verso il basso ($1 - \alpha = 0.95$): AC(2), AC(z), Clopper e Pearson/Bayes e Laplace (in evidenza), score, mid-P, Jeffreys. Il grafico è tratto da Tuyl, Gerlach e Mengersen (2009)

3.3 Livelli di copertura

L'analisi condotta da Tuyl, Gerlach e Mengersen e fin qui presentata si concentra sul caso specifico di $x = 0$. Tuttavia, può risultare interessante valutare il livello di copertura di alcuni metodi dei metodi discussi.

Per un'analisi approfondita riguardo i livelli di copertura si rimanda ai lavori di Agresti e Coull (1998) e Brown, Cai e DasGupta (2001).

Quando si parla di probabilità di copertura di un intervallo per il parametro θ si intende la probabilità che il parametro cada nell'intervallo con livello di confidenza nominale $1 - \alpha$.

Costruendo un grafico dei livelli di copertura relativi al metodo di Wald, al secondo metodo di Agresti e Coull e al metodo esatto, si rendono abbastanza evidenti le differenze tra essi.

I grafici relativi al metodo di Agresti e Coull indicano che aggiungere 2 successi e 2 insuccessi migliora il metodo suggerito da Wald in maniera sensibile; infatti, mentre prima il metodo di Wald si posizionava costantemente al di sotto del livello nominale 0.95 e falliva per livelli di θ estremi, con l'accorgimento proposto da Agresti e Coull il grafico di copertura oscilla attorno al valore nominale senza mai scendere sotto 0.92 per $n > 10$ (Agresti e Coull, 1998). La copertura per θ prossimo ai valori 0 e 1 è migliorata, anche se restano i valori verso i quali la copertura oscilla maggiormente e raggiunge i livelli più bassi. Per valori di θ molto prossimi a 0, sembra che anche il metodo di Agresti e Coull sia conservativo in accordo con quanto discusso nel paragrafo 3.2. Per quanto riguarda il metodo esatto, il grafico si trova interamente al di sopra del livello nominale 0.95. Ciò significa che sarebbe possibile alzare il livello nominale oltre 0.95 e continuare ad avere un livello di copertura sufficiente. Al fine di effettuare un confronto, vengono riportati anche i grafici relativi a $n = 10$ e $n = 50$ esclusivamente per il metodo esatto e per quello di Agresti e Coull.

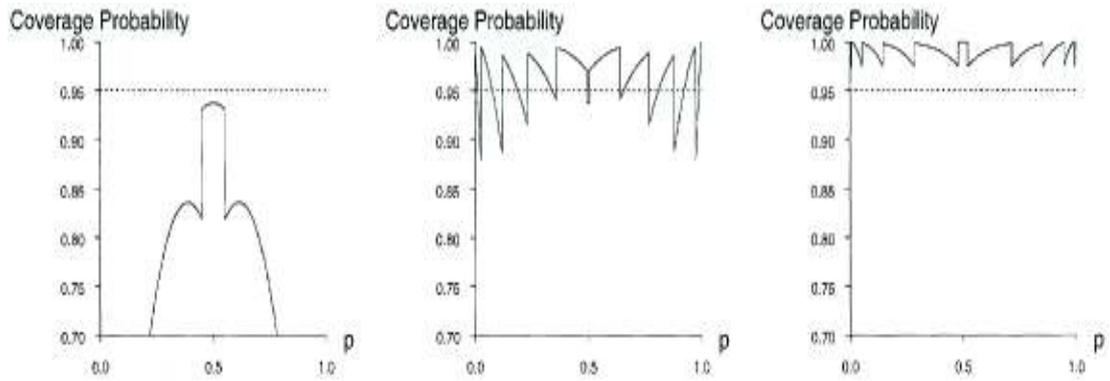


Figura 3.2: Grafici di copertura relativi a $n=5$. Da Sinistra verso destra: Wald, $AC(2)$, esatto.

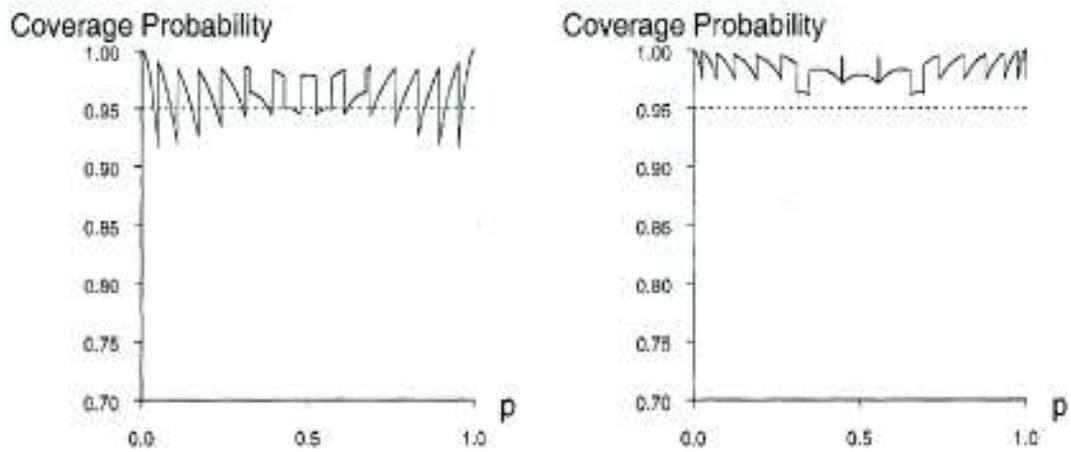


Figura 3.3: Grafici di copertura relativi a $n=10$. Da Sinistra verso destra: $AC(2)$, esatto.

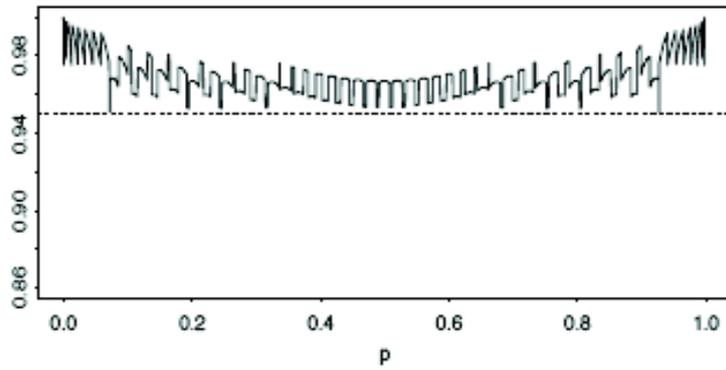


Figura 3.4: *Grafico di copertura per il metodo esatto relativo a $n=50$.*

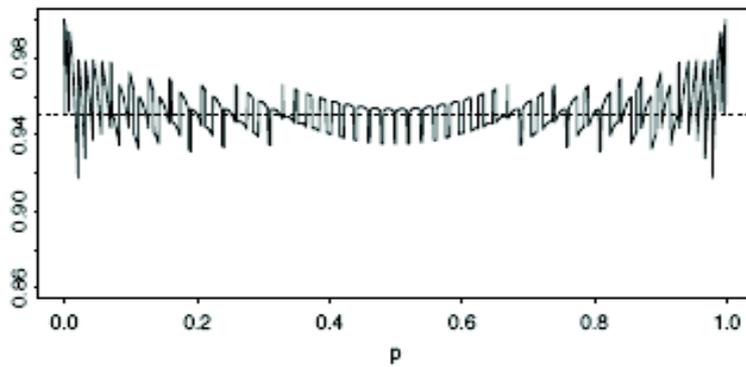


Figura 3.5: *Grafico di copertura per il metodo AC2 relativo a $n=50$.*

Conclusioni

Dall'analisi svolta si può concludere che, per il caso $x^{oss} = 0$, il metodo più ampiamente adottato è il metodo esatto, come era possibile aspettarsi. Tuttavia si è visto come applicando tecniche bayesiane e scegliendo una distribuzione a priori con $a = b = 1$, si ottengano risultati analoghi. Inoltre, se si lavora con n molto grande, risultano validi anche metodi che utilizzano $a = 1, b > 1$ perché, come detto, le informazioni a priori vengono attenuate dai dati.

Per quanto riguarda la probabilità di copertura è emerso che, soprattutto per numerosità campionarie piccole, gli intervalli alla Wald risultano insoddisfacenti. Significativamente migliore appare il metodo proposto da Agresti e Coull, che però presenta ancora dei punti in cui la probabilità di copertura risulta inferiore al livello nominale, cosa che non accade, invece, per il metodo esatto, che presenta un grafico di copertura costantemente al di sopra dello 0.95. Va detto che, anche in questo caso, i risultati migliorano all'aumentare di n . Agresti e Coull (1998) si interrogano, senza poter giungere ad una conclusione precisa, su quale dei risultati sia da preferire, cioè se utilizzare un approccio che garantisca un livello di *almeno* 0.95 oppure, come essi propongono, un metodo che sia abbastanza vicino al nominale.

Riferimenti bibliografici

- Agresti, A. e Coull, B.A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, **52**, 119-126.
- Brown, L.D., Cai, T.T. e DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, **16**, 101-133.
- Fraser, D.A.S, Reid, N. e Wong, A.C.M. (2004). Inference for bounded parameters. *Physical Review D*, **69**, 033002.
- Jovanovic, B.D. e Levy, P.S. (1997). A look at the rule of three. *The American Statistician*, **51**, 137-139.
- Lancaster, H.O. (1961). Significance tests in discrete distributions. *Journal of the American Statistical Association*, **56**, 223-234.
- Lancaster, H.O. (1949). The combination of probabilities arising from data in discrete distributions. *Biometrika*, **36**, 370-382.
- Louis, T.A. (1981). Confidence intervals for a binomial parameter after observing no successes. *The American Statistician*, **35**.
- Olivier, J. e May, W.L. (2007). A discussion of binomial interval estimators on the boundary. *Journal of the Mississippi Academy of Sciences*, **52**, 154.
- Tuyl, F., Gerlach, R. e Mengersen, K. (2009). The rule of three, its variants and extensions. *International Statistical Review*, **77**, 266-275.

Tuyl, F., Gerlach, R. e Mengersen, K. (2008). A comparison of Bayes-Laplace, Jeffreys, and other priors: the case of zero events. *The American Statistician*, **62**, 40-44.

Winkler, R.L., Smith, J.E. e Fryback, D.G. (2002). The role of informative priors in zero-numerator problems: being conservative versus being candid. *The American Statistician*, **56**, 1-4.

Per alcune definizioni e per lo svolgimento di alcune dimostrazioni si sono utilizzati

Azzalini, A e Vedaldi, R. (1985). *Introduzione all'Inferenza Statistica Parametrica*. Cleup editore, Padova.

Pace, L e Salvan, A. (2001). *Introduzione alla Statistica. II Inferenza, Verosimiglianza, Modelli*. Cedam, Padova.