

Testo di riferimento:

Steven Lavenberg "Computer Performance Modeling Handbook" Academic Press,

Stime puntuali e Intervalli di confidenza

- Output della simulazione è costituito da sequenze di variabili casuali
 - o stimare media, varianza, mediana, etc.
 - o la stima consisterà in una v.c.
 - o una quantità deterministica verrà stimata mediante una v.c.

Stima di caratteristiche transienti (terminating simulation)

Eseguiamo M osservazioni (M run di simulazione) della v.c di interesse X :

Campione di dati i.i.d

$$X_1, X_2, X_3, \dots, X_M$$

$F(x) = \text{Prob}(X < x)$: distribuzione di probabilità

Media: $\mu = E[X]$;

Varianza: $\sigma^2 = E[(X - \mu)^2] = \text{Var}[X]$

Utilizziamo il campione per stimare μ .

$$\hat{\mu} = \bar{X} = \frac{1}{M} \sum_{m=1}^M X_m$$

$\hat{\mu}$ è una variabile casuale; $E[\hat{\mu}] = \mu$.

La distribuzione di $\hat{\mu}$ deve essere centrata intorno a μ

L'accuratezza della stima dipende dalla differenza

$$\hat{\mu} - \mu$$

necessità di ulteriori informazioni sulla stima.

$$\sigma^2 = \text{Var}[X]$$

$$\text{Var}[\hat{\mu}] = \sigma^2/M$$

Utilizziamo il campione per stimare anche σ^2

$$s^2 = \hat{\sigma}^2 = \frac{1}{M-1} \sum (X_m - \bar{X})^2$$

$$\hat{\mu} = \bar{X} = \frac{1}{M} \sum_{m=1}^M X_m$$

La quantità s^2/M è una stima della varianza di $\hat{\mu}$

Inoltre la variabile casuale:

$$\frac{\hat{\mu} - \mu}{s / M^{1/2}}$$

ha una distribuzione di probabilità **t-student** con $M-1$ gradi di libertà.

Fissato un livello di confidenza α e in base ai valori tabulati della **t-student** ($t_{M-1}(\alpha/2) = t_{M-1}(1-\alpha/2)$)

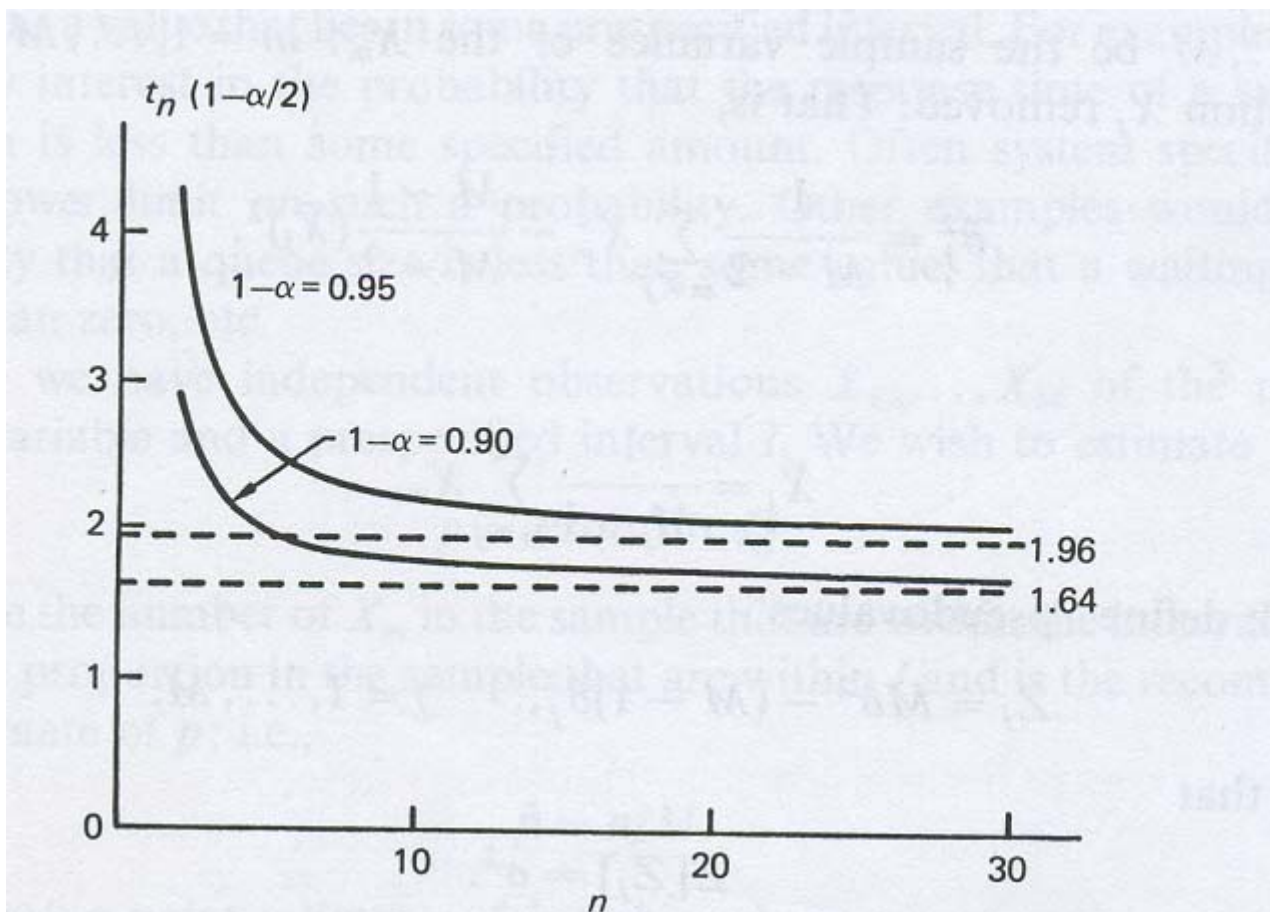
siamo in grado di calcolare intervalli di confidenza:

$$\hat{\mu} - t_{M-1}(1-\alpha/2)s / M^{1/2} \leq \mu \leq \hat{\mu} + t_{M-1}(1-\alpha/2)s / M^{1/2}$$

Con la proprietà che:

$$\text{Prob}\{\hat{\mu} - t_{M-1}(1 - \alpha/2)s/M^{1/2} \leq \mu \leq \hat{\mu} + t_{M-1}(1 - \alpha/2)s/M^{1/2}\} \approx 1 - \alpha.$$

Per valori di $(1-\alpha)$ pari a 0.95 e 0.90 i valori di $t_{n-1}(1-\alpha/2)$ sono:



Per valori di $n > 30$ i valori dei moltiplicatori diventano: 1.64 [$(1-\alpha) = 0.90$] e 1.96 [$(1-\alpha) = 0.95$]

$$\hat{\mu} - 1.64(s/M^{1/2}) \leq \mu \leq \hat{\mu} + 1.64(s/M^{1/2})$$

$$\hat{\mu} - 1.96(s/M^{1/2}) \leq \mu \leq \hat{\mu} + 1.96(s/M^{1/2})$$

Stima di caratteristiche stazionarie

- la sequenza di output converge

$$\{V_n : n = 1, 2, \dots\}$$

$$\lim_{n \rightarrow \infty} \text{Pr ob} \langle V_n \leq x \rangle = F(x)$$

Due tipi di problemi:

- a) determinare la fase transiente;
- b) eliminare la correlazione.

Fase transiente

- 1) osservare $F_n(x)$ e stabilire (mediante istogrammi) quando la fase transiente è terminata; costo elevato).

2) Esaminiamo la convergenza della media della distribuzione:

$$\mu_n = E[V_n]$$

Supponiamo di ripetere la simulazione M volte (M run di simulazione) e otteniamo N osservazioni per ogni run

- Run (1): $V_{1,1}, V_{1,2}, V_{1,3}, V_{1,4}, V_{1,5}, \dots, V_{1,N}$

- Run (2): $V_{2,1}, V_{2,2}, V_{2,3}, V_{2,4}, V_{2,5}, \dots, V_{2,N}$

- Run (3): $V_{3,1}, V_{3,2}, V_{3,3}, V_{3,4}, V_{3,5}, \dots, V_{3,N}$

-

-

-

- Run (M): $V_{M,1}, V_{M,2}, V_{M,3}, V_{M,4}, V_{M,5}, \dots, V_{M,N}$

$$\hat{\mu}_n = \bar{V}_n = \frac{1}{M} \sum_{m=1}^M V_{n,m} \quad n = 1, 2, \dots, N$$

Sequenza random che converge:

$$E[\hat{\mu}_n] = \mu_n$$

Esempio:

μ_2 : Tempo medio di permanenza nella CODA 2

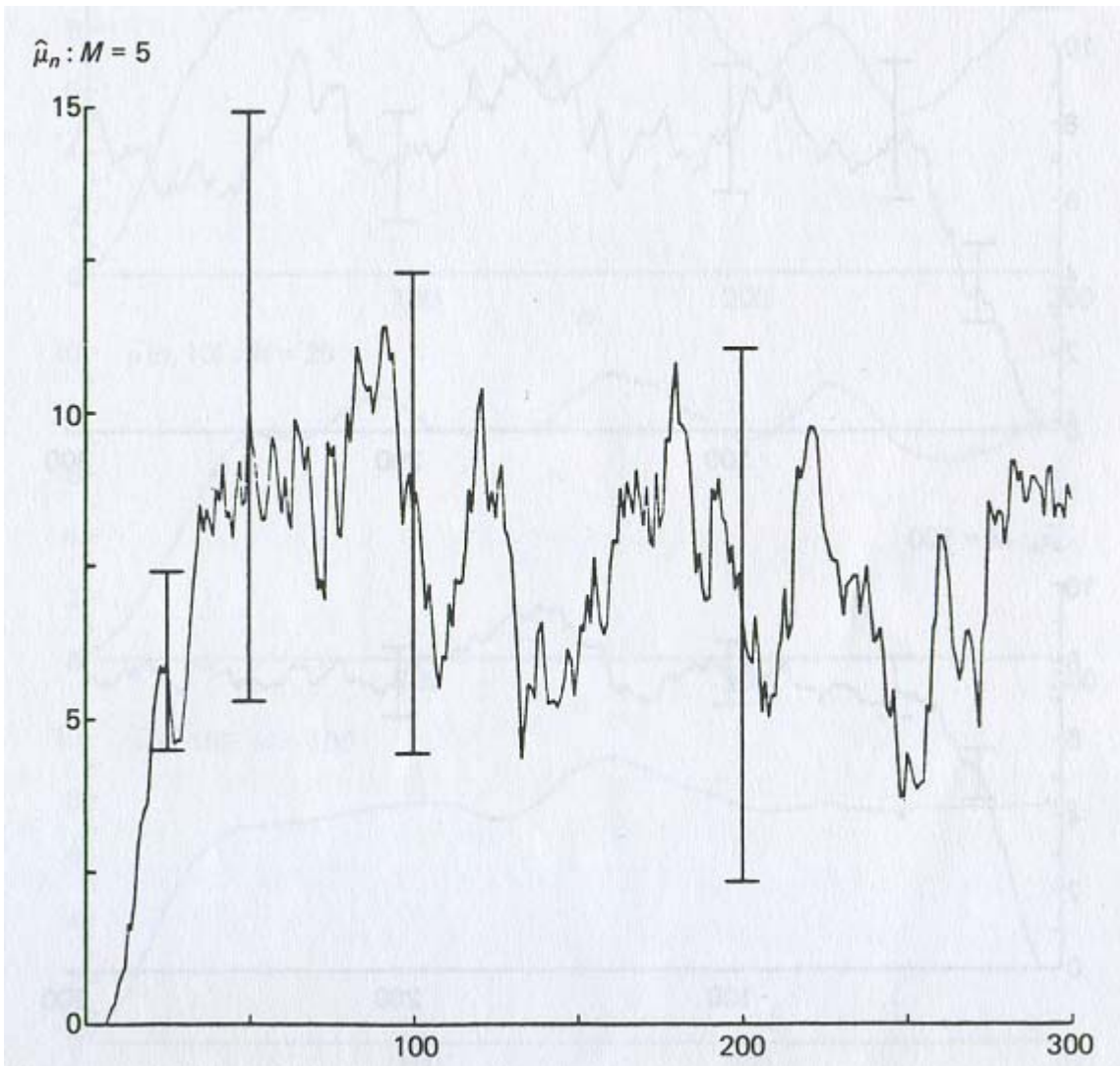
$W_{2,n}$: Tempo di attesa dell'n-ennesimo utente
nella CODA 2.

Consideriamo $V_{m,n}$ valore assunto da $W_{2,n}$ durante
l'm-esimo run di simulazione;

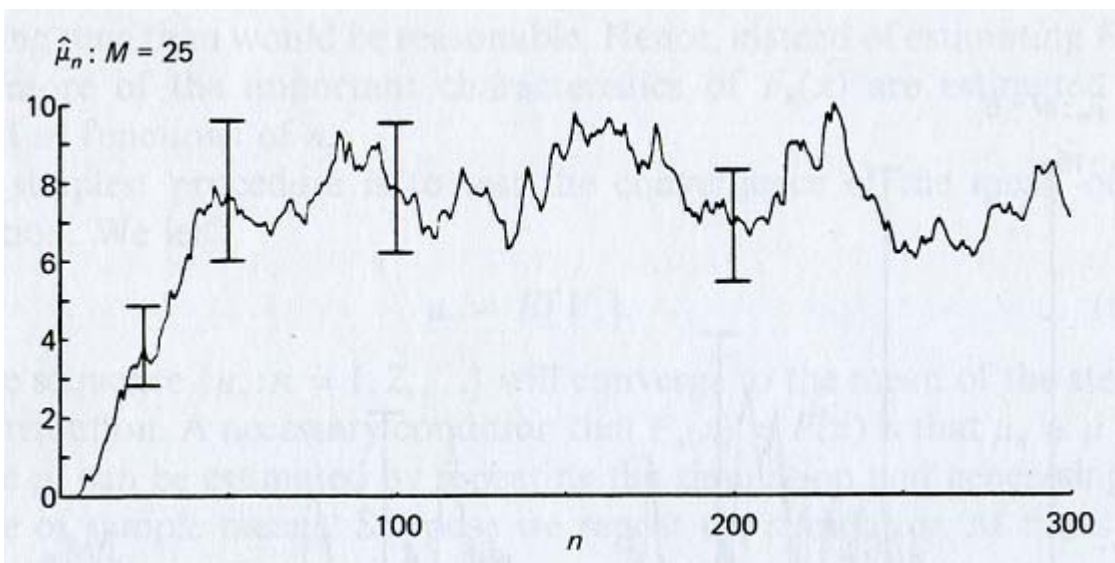
$$m = 1, 2, \dots, M;$$

$$n = 1, 2, \dots, N.$$

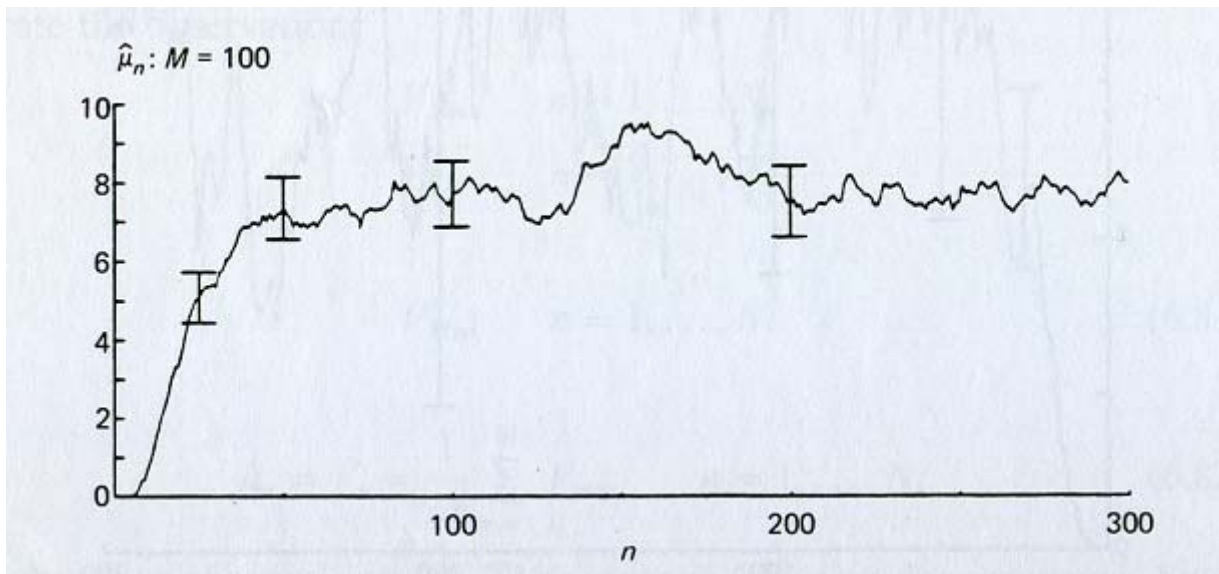
$$M = 5; N = 300.$$



$M = 25; N = 300.$

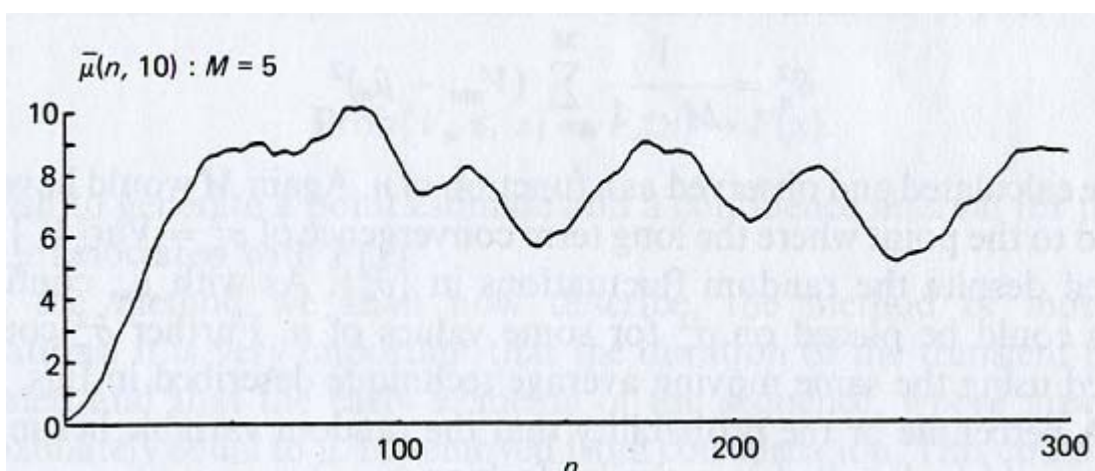


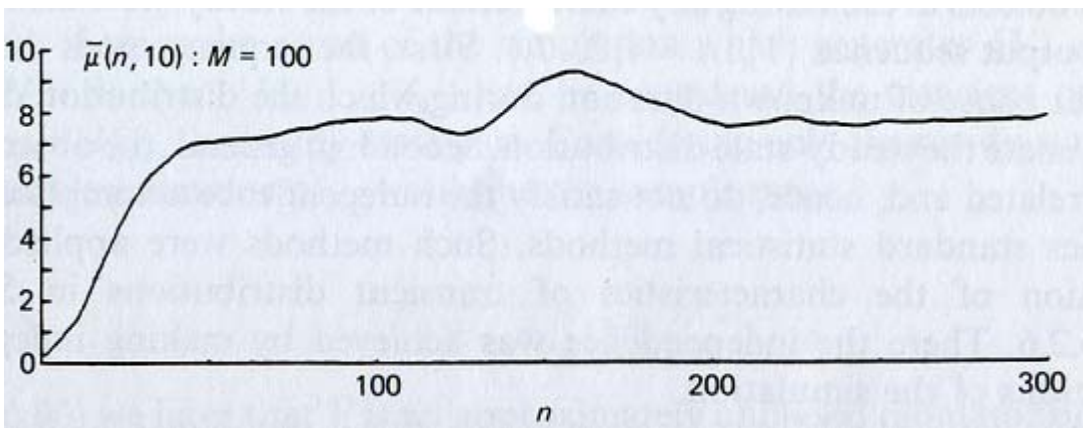
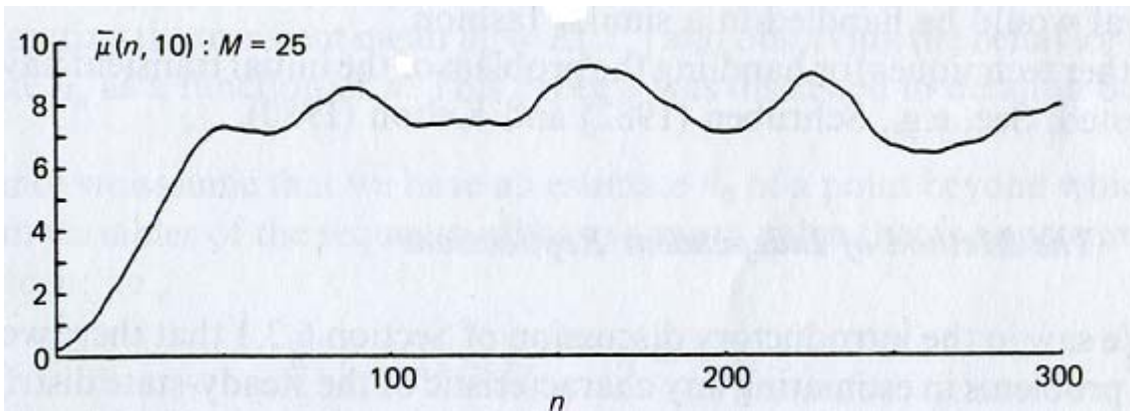
$M = 100; N = 300.$



utilizziamo un metodo per ottenere grafici che attenuino il contributo delle fluttuazioni a breve termine:

$$\bar{\mu}(n; K) = \begin{cases} (2K + 1)^{-1} \sum_{k=-K}^K \hat{\mu}_{n+k} & \text{if } n \geq K + 1, \\ (2n - 1)^{-1} \sum_{k=-(n-1)}^{n-1} \hat{\mu}_{n+k} & \text{if } n < K + 1. \end{cases}$$





- 3) stima delle caratteristiche steady state mediante il metodo delle prove ripetute.
- suddividere i dati raccolti in ogni run in fase transiente e fase stazionaria;
 - per ogni singolo run si ottiene, utilizzando i dati relativi alla fase stazionaria, una stima puntuale della caratteristica di interesse;
 - la stima puntuale finale si ottiene come media delle stime puntuali ottenute;

d) si calcolano gli intervalli di confidenza.

Stima della media di una v.c. stazionaria

$$\{V_n : n = 1, 2, \dots\}$$

$$\text{Prob}\{V_n \leq x\} = F_n(x) \rightarrow F(x).$$

Si vuole stimare la media μ della associata ad $F(x)$.

a) eliminazione del transiente:

è necessario calcolare il valore \hat{n}_0 tale che:

$$E[V_{\hat{n}_0+n}] \approx \mu, \quad n = 1, 2, \dots$$

I dati raccolti:

$$\{V_1, \dots, V_N\}$$

vengono suddivisi in due insiemi:

$$\text{Fase transiente } \{V_1, \dots, V_{\hat{n}_0}\}$$

$$\text{Fase stazionaria } \{V_{\hat{n}_0+1}, \dots, V_N\}$$

b) stima puntuale della media (singolo run)

indichiamo con $V_{m,n}$ l'n-esimo elemento del run m-esimo; $m = 1, 2, \dots, M$; $n = 1, 2, \dots, N$

Stimiamo per ogni run la media:

$$\hat{\mu}_m = \frac{1}{N - \hat{n}_0} \sum_{n=\hat{n}_0+1}^N V_{mn}.$$

$$\hat{\mu}_m, m = 1, \dots, M,$$

utilizzando questi ultimi valori siamo in grado di ottenere una stima puntuale:

$$\hat{\mu} = \frac{1}{M} \sum_{m=1}^M \hat{\mu}_m.$$

Stimiamo la varianza campionaria:

$$s^2(\hat{\mu}_m) = \frac{1}{M-1} \sum_{m=1}^M (\hat{\mu}_m - \hat{\mu})^2 = \frac{1}{M-1} \sum_{m=1}^M \hat{\mu}_m^2 - \frac{M}{M-1} \hat{\mu}^2$$

Sappiamo che la v.c.

$$\frac{\hat{\mu} - \mu}{s(\hat{\mu}_m)/M^{1/2}}$$

ha una distribuzione t-student ed è quindi possibile calcolare, fissato un livello di confidenza α , gli intervalli di confidenza:

$$\text{Prob}\{\hat{\mu} - t_{M-1}(1 - \alpha/2)s(\hat{\mu}_m)/M^{1/2} \leq \mu \leq \hat{\mu} + t_{M-1}(1 - \alpha/2)s(\hat{\mu}_m)/M^{1/2}\} \\ \approx 1 - \alpha$$