

La variabile casuale F di Snedecor (o variabile F di Fisher)

Giovanni Battista Flebus
Lezioni di Psicometria

La v.c. F di Snedecor

- Ha come dominio solo valori positivi, a cui corrispondono delle aree, considerate misure di probabilità
- Si usa per confrontare due varianze, per stabilire se sono o no uguali.
- Simile al valore t di Student o al chi quadrato, l'F di Fisher ha **due** parametri (gl_1 e gl_2) che servono a identificare la curva di probabilità;
- I **due parametri** sono le numerosità dei campioni da cui sono ricavate le varianze



George W. Snedecor (1882 -1974)

*Autore del libro
Statistical Methods
(1937). (più di
125,000 copie
vendute)*



•R. A. Fisher (1890-1962)

Il nome F

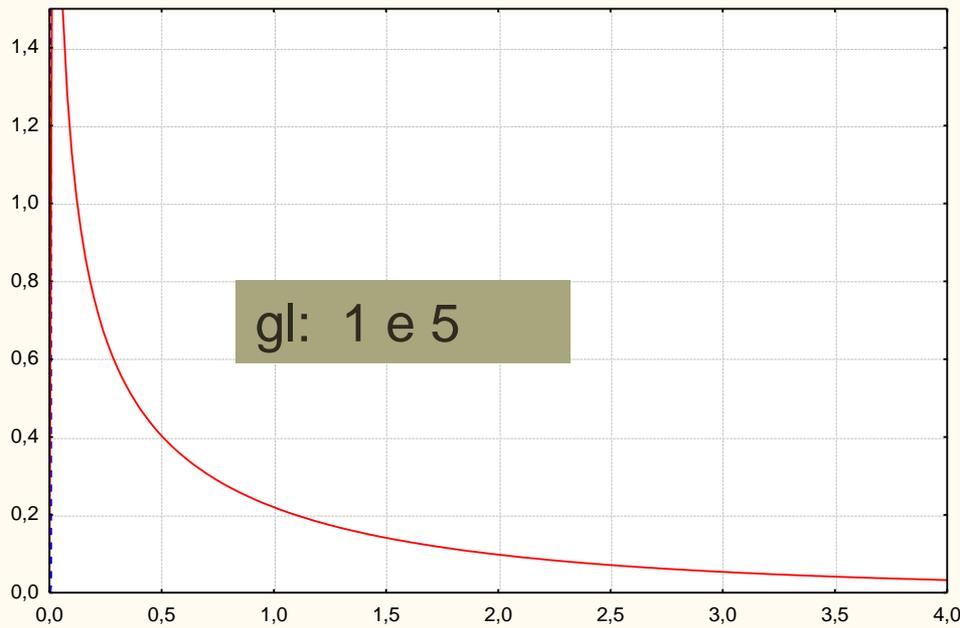
- Lo statistico Snedecor ha nominato questa importantissima variabile casuale in onore del suo maestro Ronald Fisher

La distribuzione di F

- Varia molto quando i gradi di libertà sono piccoli

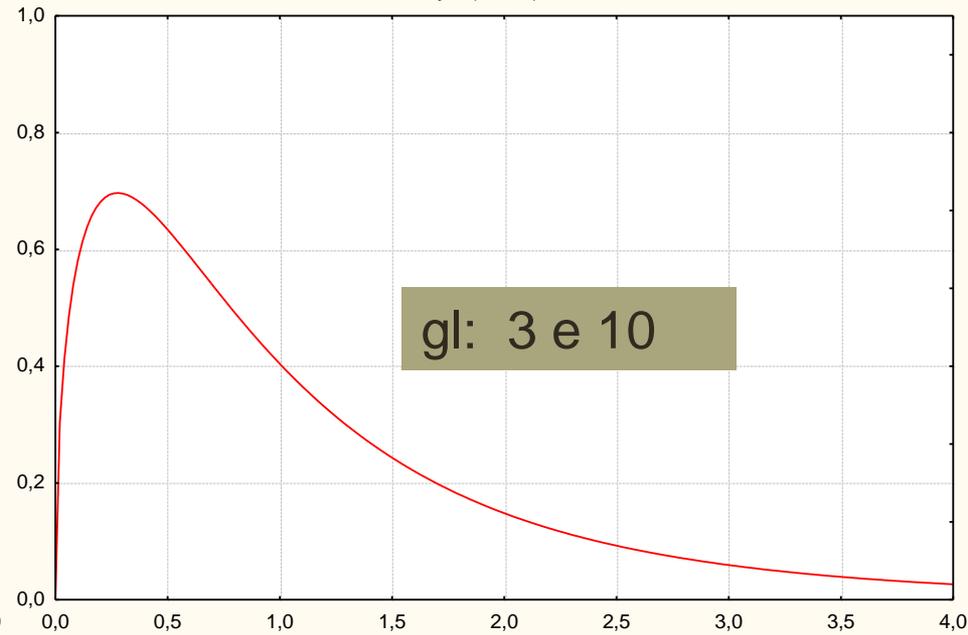
Funzione di Densità di Probabilità

$$y=F(x;1;5)$$



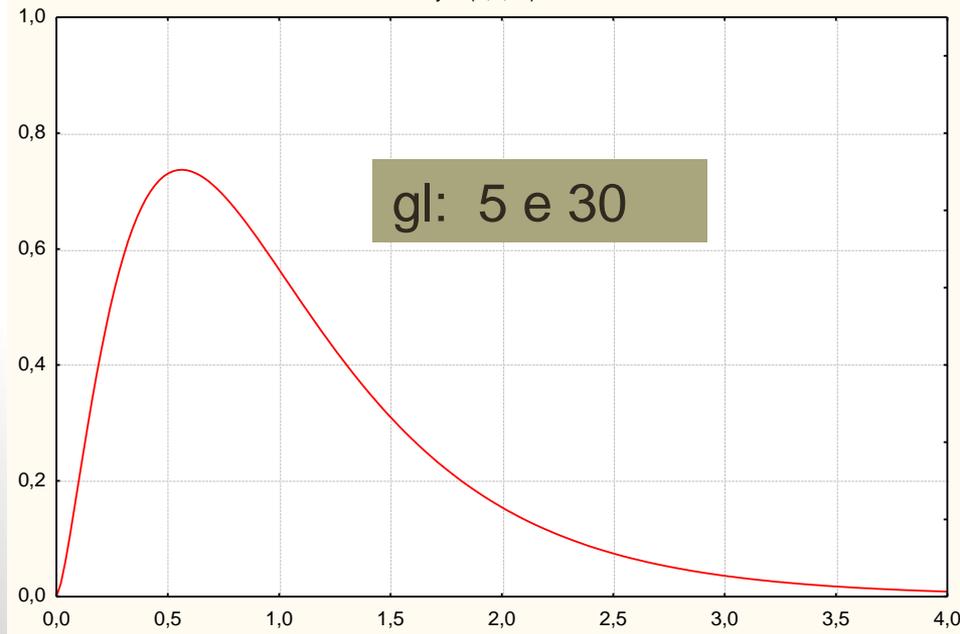
Funzione di Densità di Probabilità

$$y=F(x;3;10)$$



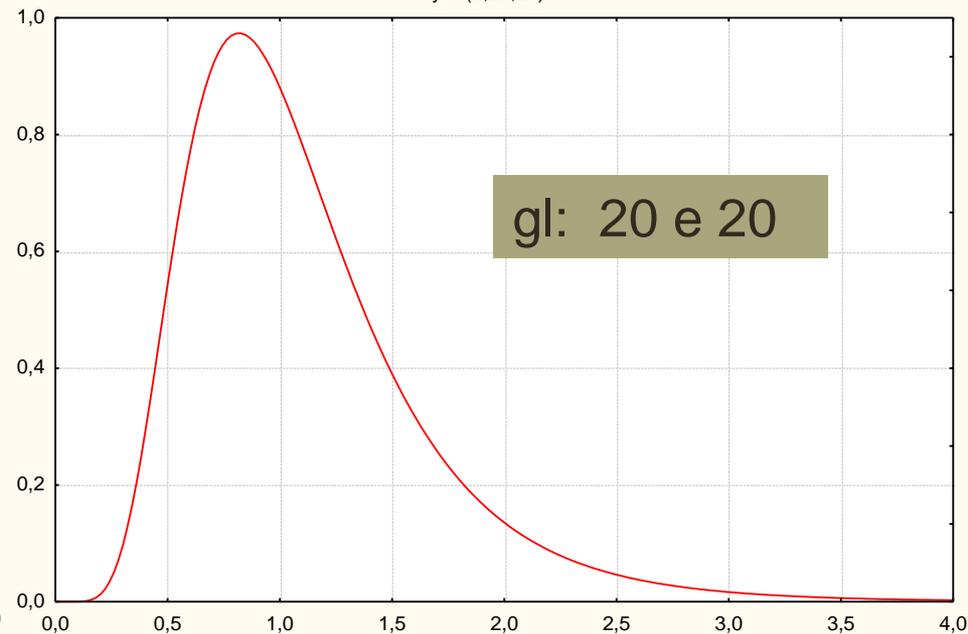
Funzione di Densità di Probabilità

$$y=F(x;5;30)$$



Funzione di Densità di Probabilità

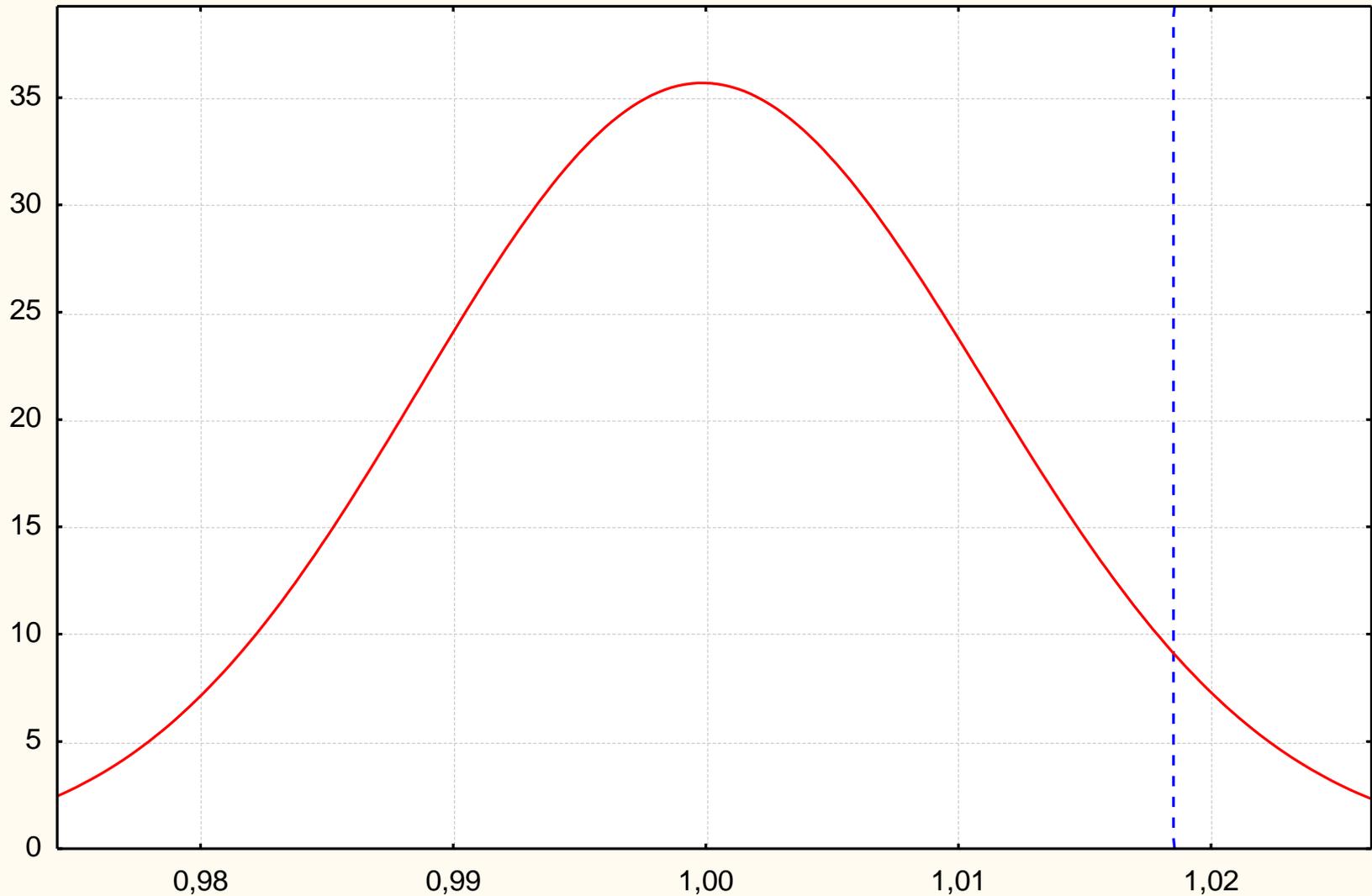
$$y=F(x;20;20)$$



Tende alla gaussiana se entrambi i due valori di gl tendono all'infinito

Funzione di Densità di Probabilità

$$y=F(x;31999;31999)$$



Verifica dell'ipotesi

- H0: $\sigma_1^2 = \sigma_2^2$
- H1: $\sigma_1^2 \neq \sigma_2^2$

- con $gl = n_1 - 1$ e $n_2 - 1$

Ipotesi da verificare

- Se le due varianze sono **simili o uguali**, il loro rapporto è vicino all'unità
- Se le due varianze sono **diverse**, il loro rapporto è molto superiore all'unità

Nota: al **numeratore** si usa la varianza con il numero di gradi di libertà più **piccoli**

- La statistica F stabilisce la probabilità di occorrenza del rapporto calcolato
- Il confronto con l'F critico (che rileva i valori con meno del 5% di probabilità di accadere), permette di trarre la conclusione

Un rapporto e non una differenza fra due varianze

- Come avrete intuito, per confrontare due varianze si usa un rapporto (varianza A diviso varianza B) e non una differenza, come si fa con le medie (Media A meno Media B).
- Se le due varianze sono simili, il rapporto è vicino a 1.
- Se le due varianze sono diverse, il rapporto è molto superiore a 1.
- (potrebbe anche essere inferiore a 1, ma per questioni teoriche questa evenienza non è per niente interessante ai fini della verifica di ipotesi).

Esempio di calcolo

Da due campioni di 8 e 10 casi si ottiene:

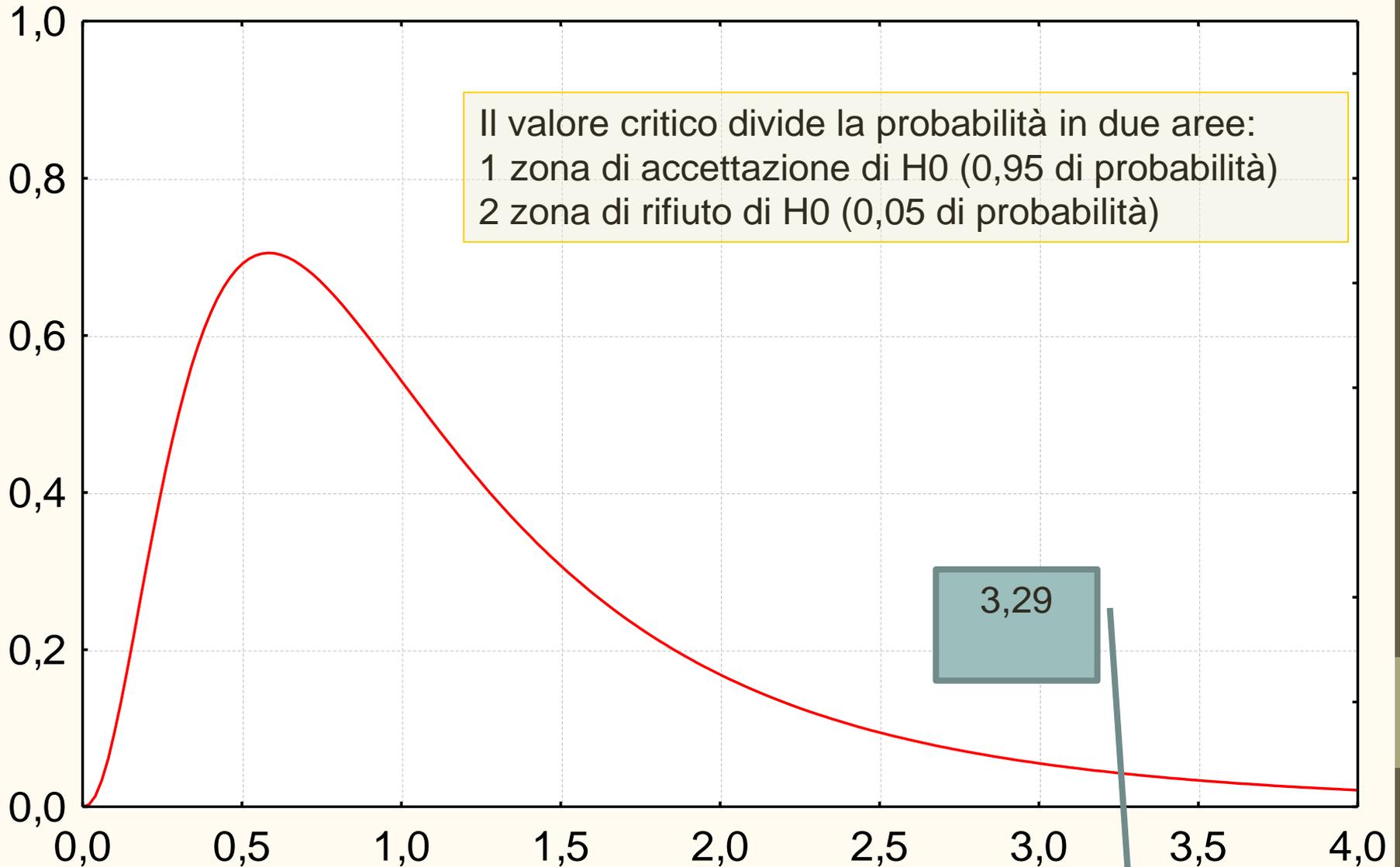
- campione 1 $s_1^2 = 56$ $gl = n-1 = 7$
campione 2 $s_2^2 = 24$ $gl = n-1 = 9$
- L'ipotesi è la seguente: le due varianze sono uguali (tratte della stessa popolazione) oppure sono diverse?

Quale è la forma di F?

- La curva della variabile casuale F con 7 e 9 gradi di libertà ha questa forma, e il suo valore critico è...

Funzione di Densità di Probabilità

$$y=F(x;7;9)$$

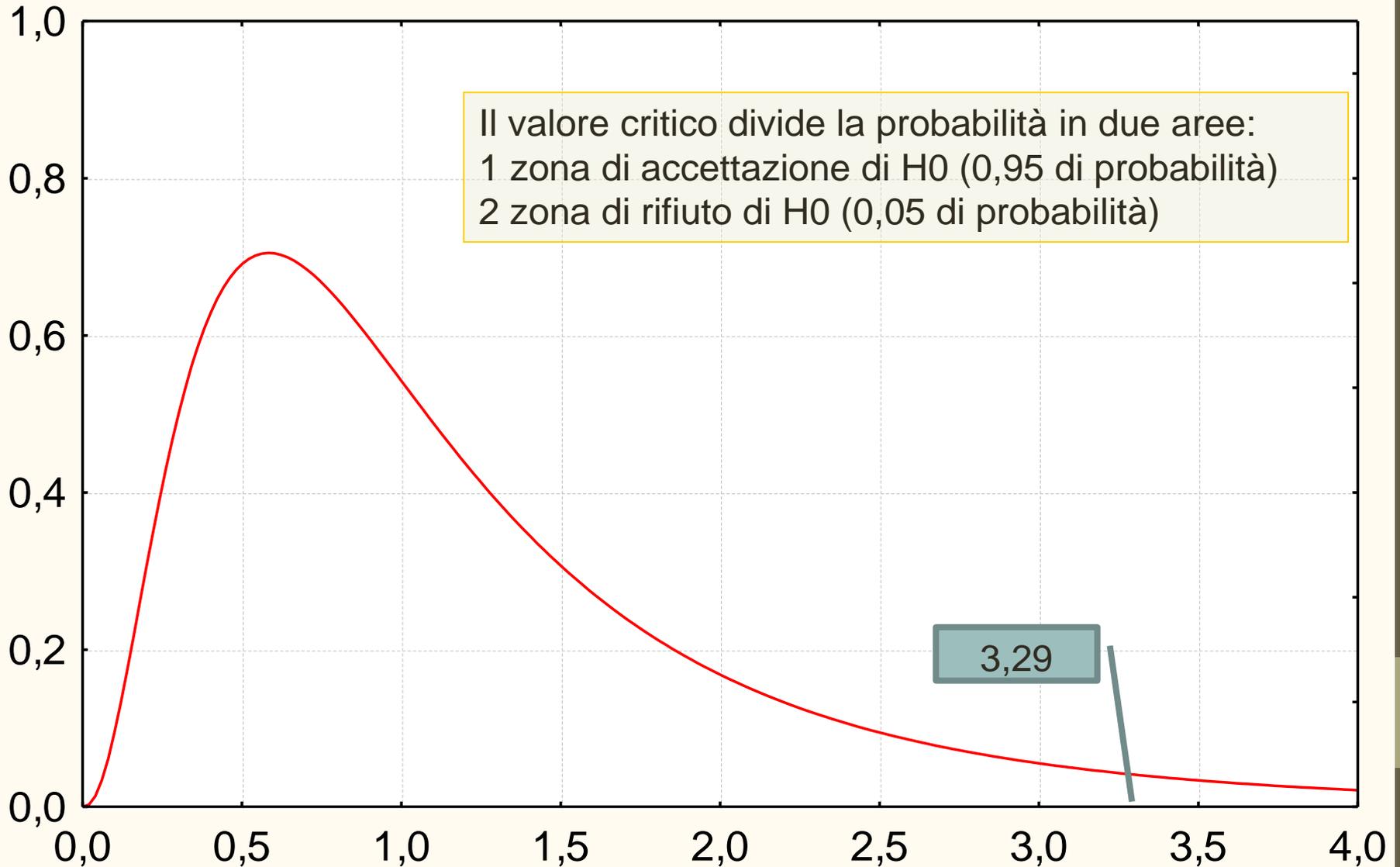


Il calcolo della statistica F dà...

$$F = \frac{56,0}{24,0} = 2,33$$

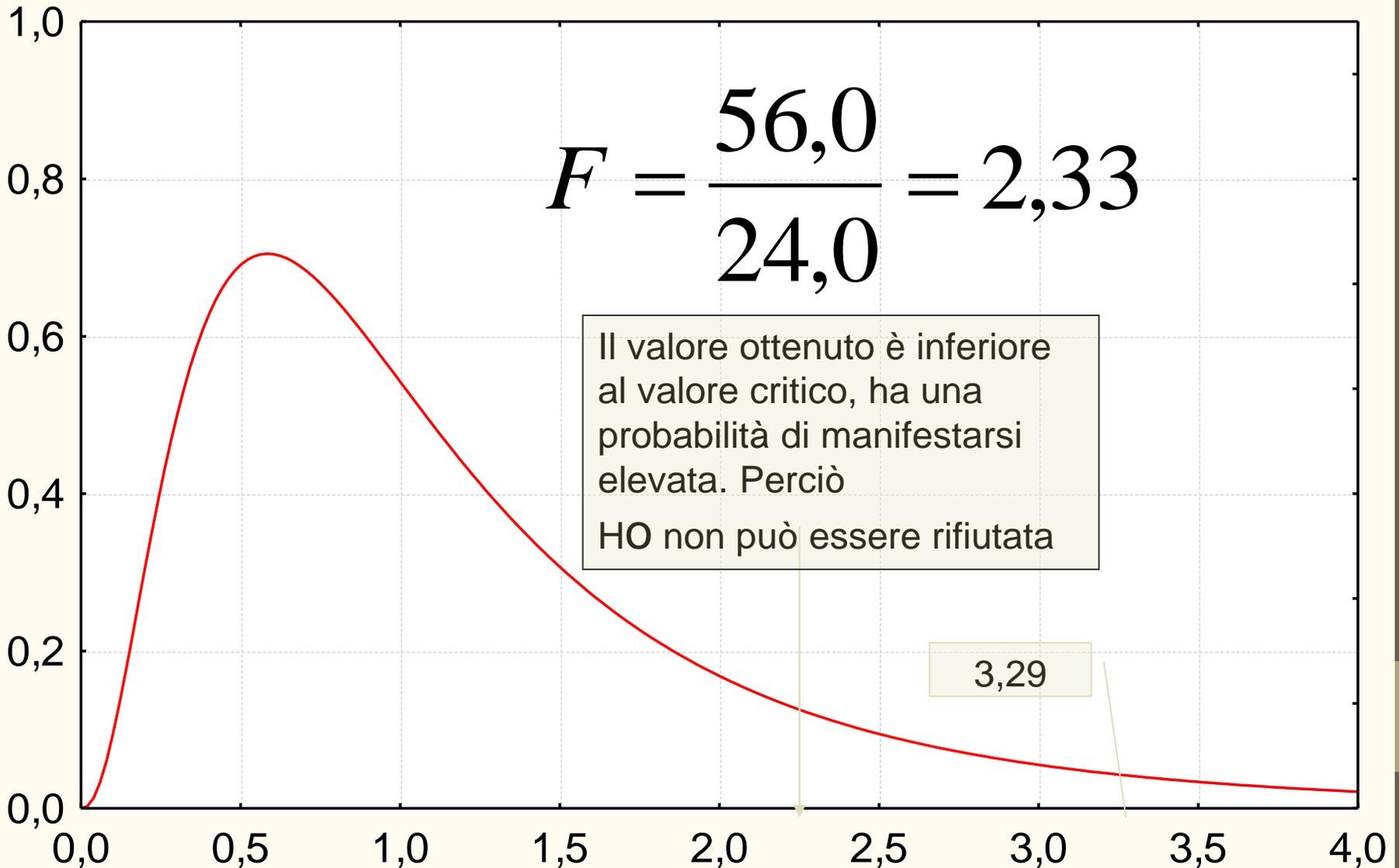
Funzione di Densità di Probabilità

$$y=F(x;7;9)$$



Funzione di Densità di Probabilità

$$y = F(x; 7; 9)$$



$$F = \frac{56,0}{24,0} = 2,33$$

Il valore ottenuto è inferiore al valore critico, ha una probabilità di manifestarsi elevata. Perciò HO non può essere rifiutata

3,29

Requisiti per l'uso di F

- Le due varianze devono provenire da popolazioni distribuite **normalmente**
- I campioni devono essere estratti in modo **indipendente** (non ci devono essere legami fra le osservazioni e l'attribuzione ad un gruppo)

Il ricorso alle tavole

- Come per la curva gaussiana e le tavole del t di Student, esistono delle tavole per indicare i percentili per $p = 0,05$ o $0,01$ di probabilità, tenendo conto dei g.l. delle due forme.

Senza Tavole ?

- Si può fare a meno delle tavole, se abbiamo Excel a disposizione (o un altro foglio elettronico a disposizione)...

Excel

- Il foglio elettronico ha molte funzioni disponibili, che possono essere inserite in una cella, digitando prima il segno di uguale e scrivendo il nome della funzione.

-

La distribuzione F con Excel

- **=DISTRIB.F.DS(X;G_libertà1;G_libertà2)**
- Primo: il valore F calcolato
- Secondo, i gradi di libertà al numeratore (sopra)
- Terzo, i gradi di libertà denominatore (sotto)
- Il risultato è l'area **a destra** del valore inserito

Esempio

- =DISTRIB.f.ds(3,5 ; 1 ; 5)

0,120

Il valore F pari a 3,5 con 1 e 5 gradi di libertà è divide l'area in due parti, una a sinistra pari a 0,880 e l'altra pari a 0,120

La variabile F con SPSS

SPSS fornisce direttamente la probabilità cumulativa **a destra** della curva di F assieme ai gradi di libertà delle due stime al numeratore e al denominatore

Uso di F

- Nella ricerca psicologica si usa raramente questa statistica per verificare se due campioni differiscono fra di loro per la varianza.
- Invece, l'uso di F per confrontare **due varianze** calcolate in **modo diverso su sub-campioni che si suppongono provenire dalla stessa popolazione** è un'operazione comunissima nella ricerca psicologica.

Interpretazione di F

L'interpretazione della probabilità è generalmente la seguente:

E' la probabilità di ottenere una stima della varianza della popolazione con due stime indipendenti che secondo l'ipotesi nulla provengono dalla stessa popolazione.

Perciò...

Un valore **molto basso** di probabilità porta a escludere **che le due stime siano tratte dalla stessa popolazione (si rifiuta H_0)**

Un **valore alto (superiore a 0,05)**...

che i due metodi di calcolare la varianza sono equivalenti, perché sono tratti dalla stessa popolazione, o anche, che i gruppi non sono distinguibili fra di loro (si accetta H_0)