

# La distribuzione del chi quadrato $\chi^2$

Lezioni di Psicometria  
Giovanni Battista Flebus

- Se abbiamo  $k$  variabili normali, con
  - media  $=0$  e
  - varianza  $=1$

la somma dei loro quadrati si distribuisce come un chi quadrato con  $k$  gradi di libertà

- Il numero degli addendi prende il nome di **gradi di libertà**
- La distribuzione si definisce perciò un **chi quadrato con  $k$  gradi di libertà**.

La sua distribuzione di frequenza è alquanto complessa e varia al variare di  $k$ , ossia del numero di Normali standardizzate sommate

$$f(x, k) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$

$$K = N(0,1)$$

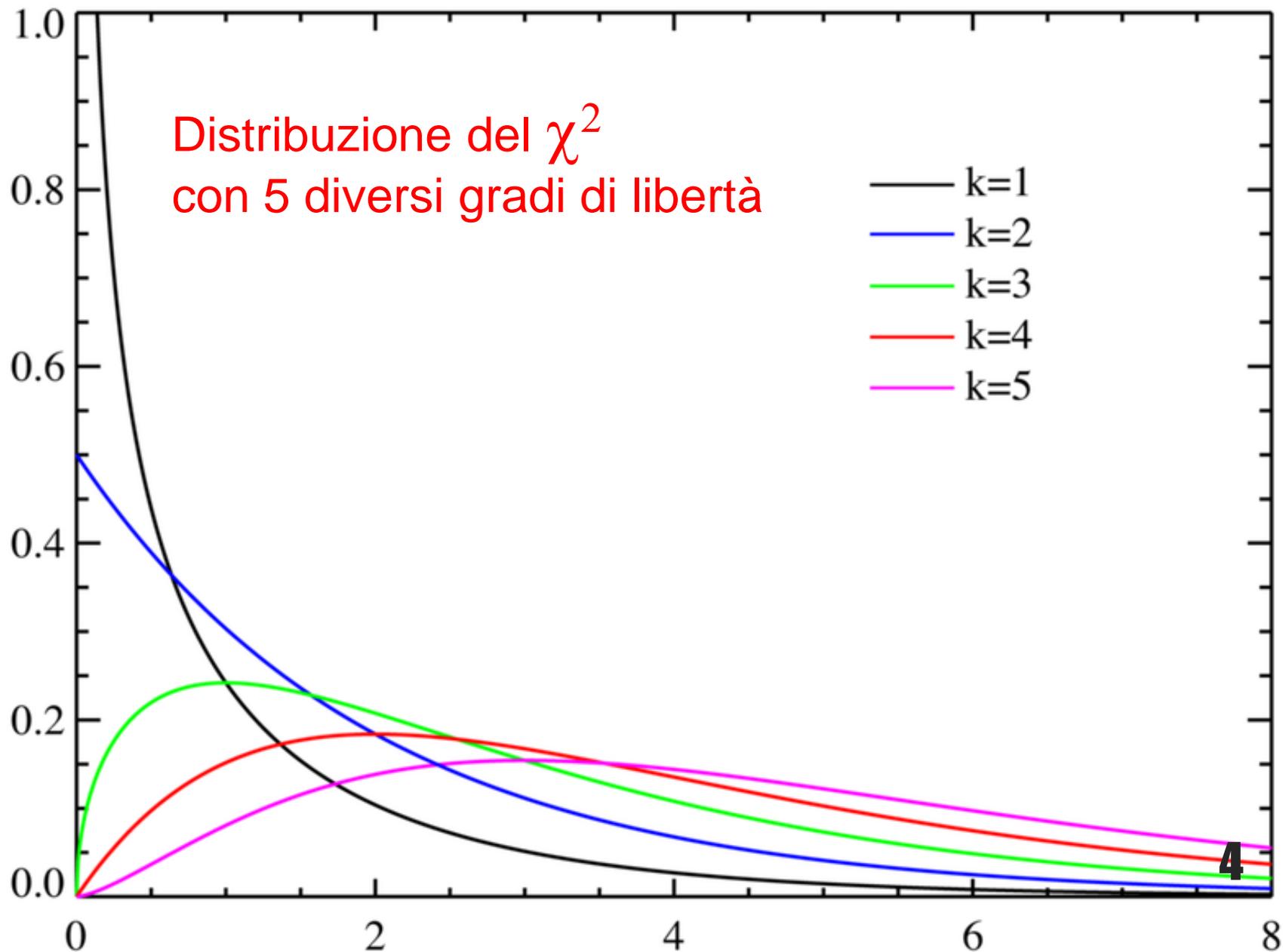
$$e = 2,7178$$

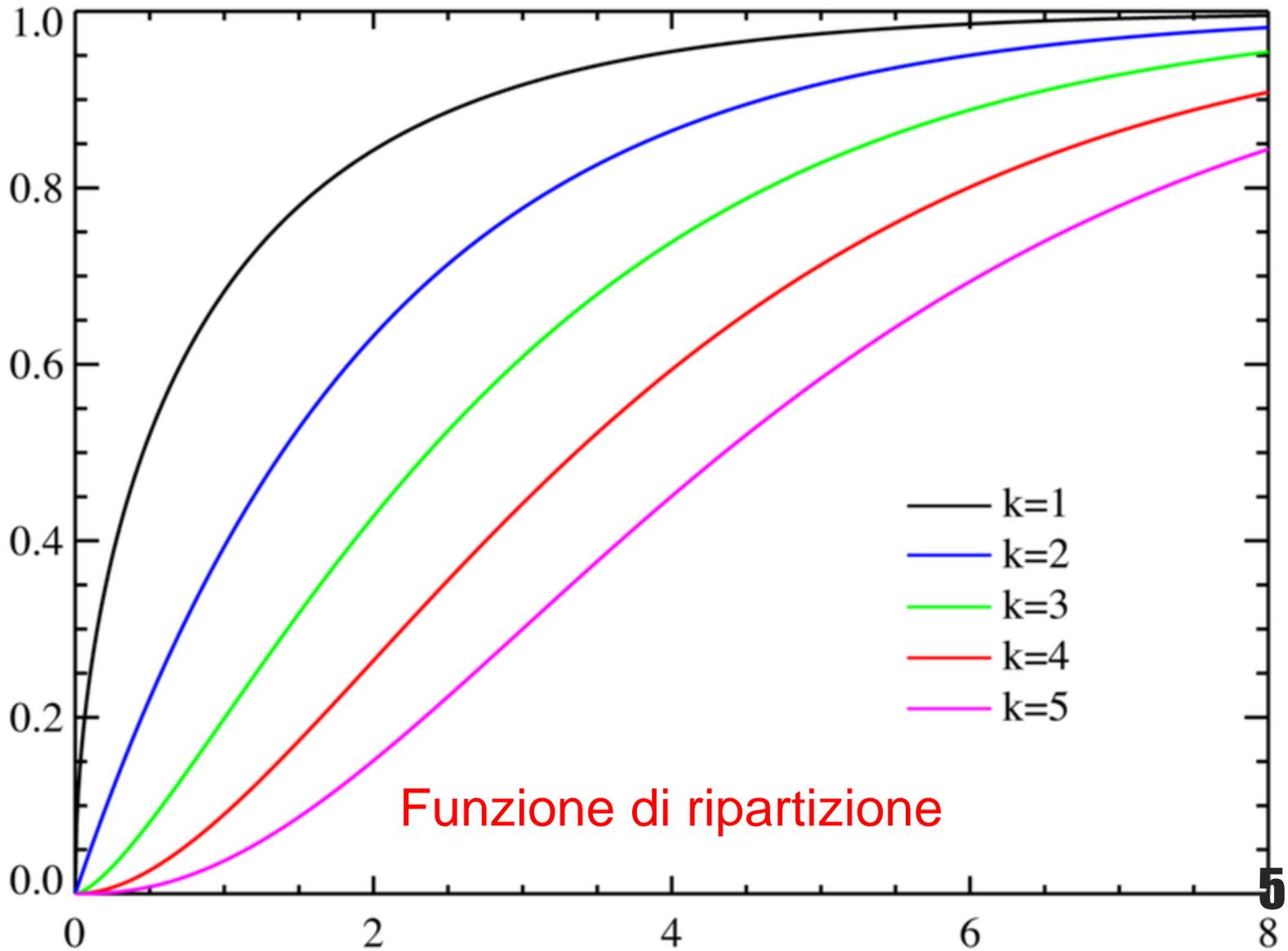
$$\text{Varianza} = 2 k$$

$$\text{Media} = k$$

$\Gamma$  = funzione gamma, fornisce il valore del fattoriale (!) per numeri non interi

Distribuzione del  $\chi^2$   
con 5 diversi gradi di libertà

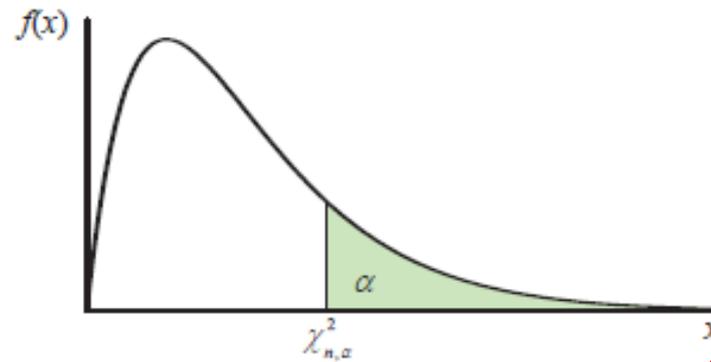




# La tavola del chi quadrato

- Poiché varia al variare del parametro  $k$ , si usa solo una piccola parte della funzione cumulativa di frequenza, detta appunto **tavola del chi quadrato**

Tavola 2 – Percentili della variabile casuale Chi-quadrato



$n$	$\alpha$									
	0.995	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.005
1	0.004393	0.005157	0.007982	0.01393	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.06	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30

- E' quel numero che divide la distribuzione di frequenza in due parti: la parte che contiene il 95% e quella che contiene il 5% dei casi.
- Serve a individuare la zona di accettazione e di rifiuto dell'ipotesi nulla
- Si considera che un valore superiore al 95 ° percentile è molto raro (una volta su 20), mentre i valori al di sotto sono considerati comuni.

# Il valore critico

---

# Usi del chi quadrato

- Il chi quadrato è molto importante perché molte statistiche seguono questa distribuzione. Lo vedremo presto in uso nel test di associazione di una tavola di contingenza.
- Una distribuzione derivata dal chi quadrato è l'F di Snedecor che si studia nell'analisi della varianza

## Simulazione con SPSS con 1000 casi fittizi

```
set seed = 234567.
```

```
COMPUTE pasquale1=RV.NORMAL(0,1).
```

```
COMPUTE pasquale2=RV.NORMAL(0,1).
```

```
COMPUTE pasquale3=RV.NORMAL(0,1).
```

```
compute chiqua1= pasquale1*pasquale1.
```

```
compute chiqua2= pasquale2*pasquale2.
```

```
compute chiqua3= pasquale3*pasquale3.
```

```
compute somchi2= chiqua1 +chiqua2.
```

```
compute somchi3= chiqua1+chiqua2+chiqua3.
```

```
FREQUENCIES
```

```
    VARIABLES=pasquale1 pasquale2 pasquale3  
    chiqua1 chiqua2 chiqua3 somchi2 somchi3  
/FORMAT=NOTABLE  /PERCENTILES=90 95 99  
/STATISTICS=MEAN MEDIAN  
/ORDER=ANALYSIS.
```

## Statistiche

	Statistiche				
	Media	Mediana	Percentili		
			90	95	99
pasquale1	,03	,01	1,41	1,78	2,49
pasquale2	-,01	,00	1,27	1,60	2,26
pasquale3	-,02	,00	1,18	1,63	2,37
chiqua1	1,10	,54	3,05	4,30	7,54
chiqua2	,97	,43	2,71	3,59	6,90
chiqua3	,97	,46	2,65	4,00	6,09
somchi2	2,08	1,47	4,77	6,18	9,33
somchi3	3,04	2,38	6,27	7,80	12,24

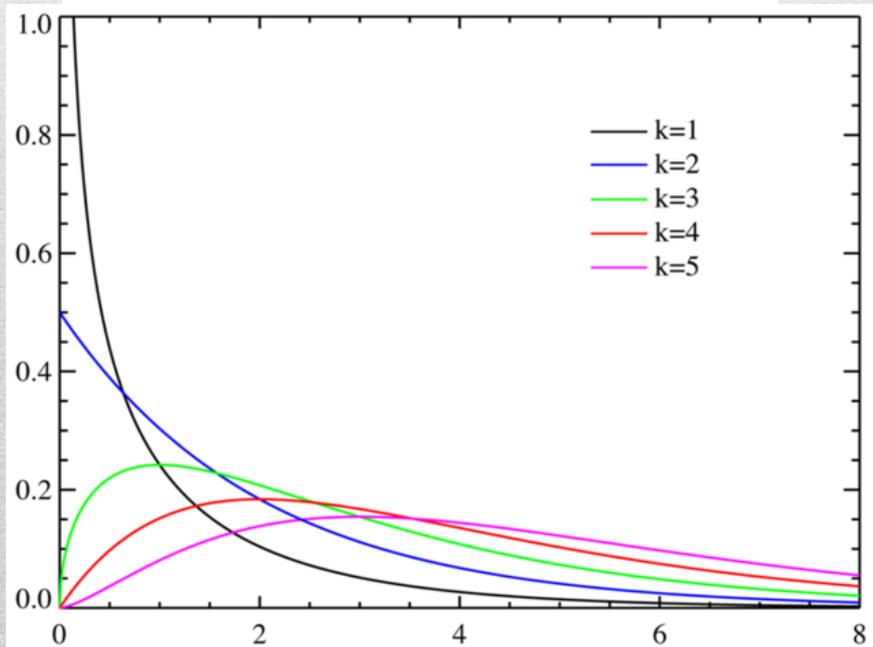
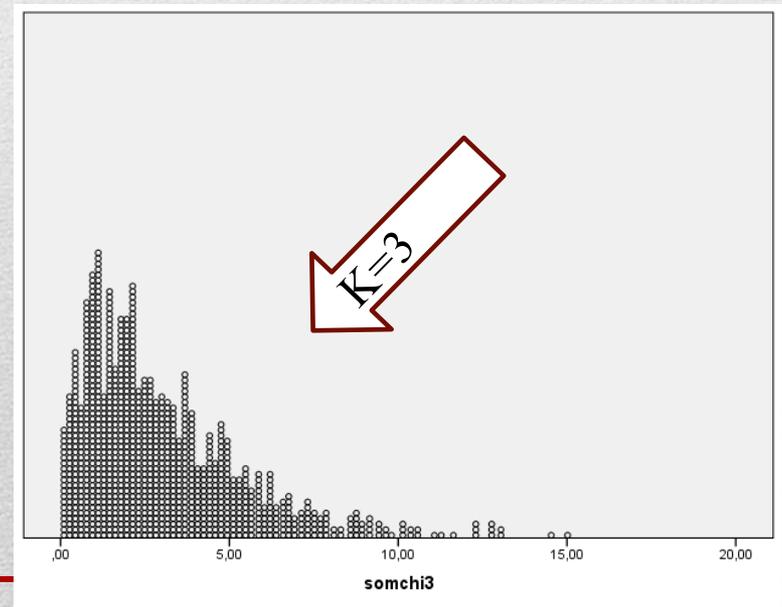
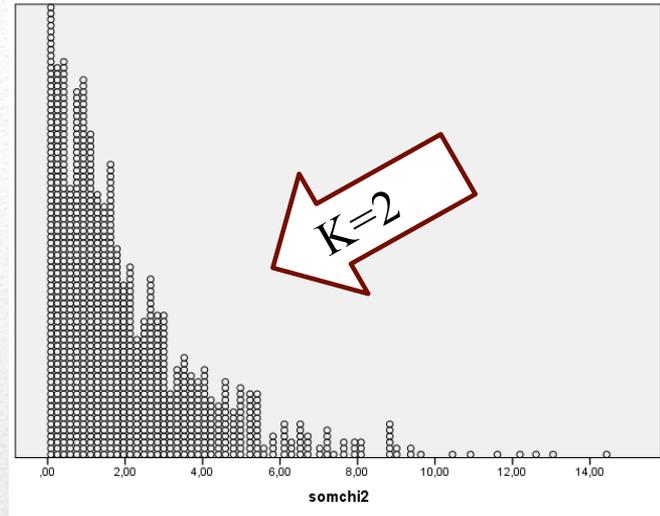
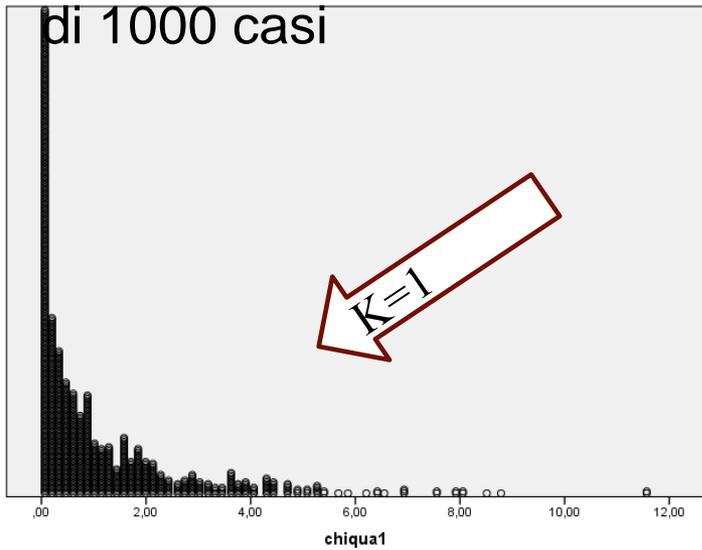
## Statistiche

	Statistiche				
	Media	Mediana	Percentili		
			90	95	99
pasquale1	,03	,01	1,41	1,78	2,49
pasquale2	-,01	,00	1,27	1,60	2,26
pasquale3	-,02	,00	1,18	1,63	2,37
chiqua1	1,10	,54	3,05	4,30	7,54
chiqua2	,97	,43	2,71	3,59	6,90
chiqua3	,97	,46	2,65	4,00	6,09
somchi2	2,08	1,47	4,77	6,18	9,33
somchi3	3,04	2,38	6,27	7,80	12,24

0.050
3.84
5.99
7.81

# Esempio di distribuzione empirica

di 1000 casi



- Anche il foglio di Excel offre la funzione del chi quadrato con questo comando
- =DISTRIB.CHI.QUAD(valore;gl;vero)
- Dove Valore si riferisce al valore che vogliamo valutare, gl è il numero di gradi di libertà, e VERO è un valore logico (opposto a FALSO) per affermare che vogliamo la funzione cumulativa di frequenza.
- Questa funzione sostituisce completamente le tavole del chi quadrato.

# La funzione Chi quadrato di Excel