I test di ipotesi

I test di ipotesi

- Il test delle ipotesi consente di verificare se, e in quale misura, una determinata ipotesi (di carattere sociale, biologico, medico, economico, ecc.) è supportata dall'evidenza empirica;
- L'obiettivo è decidere sulla "verità" o "falsità" di una certa ipotesi che viene formulata sulla popolazione;
- Il fenomeno studiato deve essere rappresentabile mediante una distribuzione di probabilità e l'ipotesi sulle caratteristiche del fenomeno studiato è tradotta in ipotesi su uno o più parametri della distribuzione (si parla di **test parametrici**).

- Esempi di ipotesi:

- La media calcolata su un certo campione differisce in modo significativo da un certo valore fissato?
- La differenza tra le medie di due campioni è statisticamente significativa?

I test di ipotesi

- LE IPOTESI:
- Vengono formulate due ipotesi sul valore del parametro:

- H_0 : **IPOTESI NULLA**
- È l'ipotesi di totale casualità dei risultati

- H_1 : **IPOTESI ALTERNATIVA**
- È una possibile ipotesi alternativa a quella di casualità; non sempre è possibile formularla compiutamente.

I test di ipotesi

- Le ipotesi sul valore del parametro possono essere:

- **semplici**: è specificato un solo valore (per es. $\mu = \mu_0$)
- **composte**: sono specificati uno o più intervalli di valori

- A una coda (per es. $\mu > \mu_0$)
- A due code (per es. $\mu \neq \mu_0$)

I test di ipotesi

- Un test di ipotesi inizia con la definizione del problema in termini di ipotesi sul parametro oggetto di studio (ad esempio sulla media di una data variabile nella popolazione).
- L'ipotesi da sottoporre a test è detta *ipotesi nulla* (H_0). Con questo termine si indica l'ipotesi che il risultato ottenuto sia totalmente casuale.
- Si specifica poi un'adeguata ipotesi alternativa (H_1).

Osservazione

- Quando non si rifiuta un'ipotesi nulla, non si può concludere che tale ipotesi sia "vera", ma soltanto che potrebbe essere vera (il campione non fornisce prove sufficienti a provocarne il rifiuto).
- Con l'inferenza si ha solo un'indicazione del fatto che l'ipotesi sia o meno "corroborata" dai dati disponibili.

L'ipotesi nulla e l'ipotesi alternativa - Esempio

- Vogliamo verificare se un certo farmaco (il farmaco A) migliora la capacità di concentrazione.
- Si deve partire con l'ipotesi "contraria", cioè che la differenza rispetto al gruppo dei non trattati, se rilevata, sia del tutto attribuibile al caso.
- **Questa è H_0 , l'ipotesi nulla.**
- L'ipotesi alternativa è un'ipotesi diversa rispetto all''ipotesi nulla: supponiamo che il farmaco produca un effetto sulla capacità di concentrazione, migliorandola.
- **Questa è H_1 , l'ipotesi alternativa.**

Come funziona un test di ipotesi

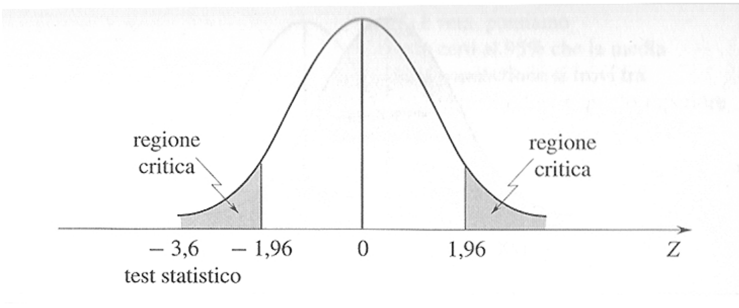
- I test di ipotesi si basano sullo studio della distribuzione campionaria di una statistica, detta statistica-test, che viene calcolata sui dati del campione.
- Prima di disporre dei dati del campione, viene definita una regola per il rifiuto o meno dell'ipotesi nulla.
- In genere, la regola consiste nel calcolare sui dati del campione la **statistica-test**, che è una statistica il cui scopo è effettuare un test di ipotesi.
- Se la statistica-test e inferiore a una certa **soglia** stabilita, non si rifiuta H_0. Se la statistica test calcolata supera la soglia, si rifiuta H_0.

Come funziona un test di ipotesi

- La regola di decisione consiste quindi nel suddividere lo spazio campionario C in **due regioni, C_0 regione di "accettazione", C_1 regione di rifiuto**, sulla base dei possibili valori della statistica-test.

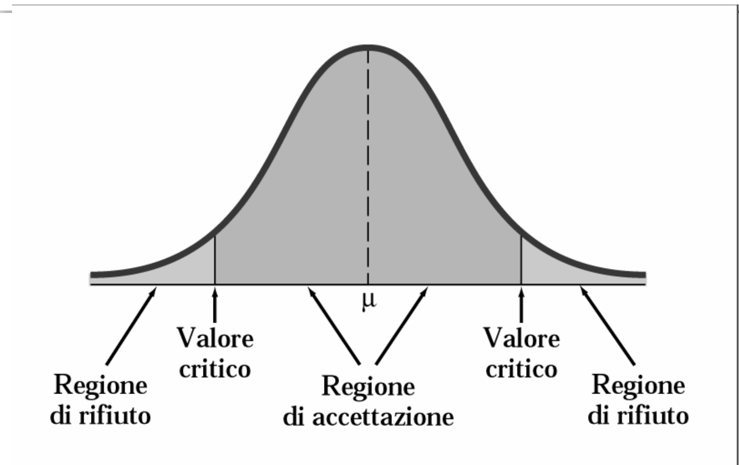
- IMPORTANTE:
- Se non si rifiuta l'ipotesi nulla, questo non significa che la teoria (espressa dall'ipotesi nulla) è provata, ma solo che ha superato una prova. In termini popperiani, possiamo dire che la teoria da cui l'ipotesi discende è stata **corroborata**.

Regione critica (di rifiuto) di un test di ipotesi (in questo caso, test a due code)



NB: il valore di Z pari a **±1,96** corrisponde al livello di significatività α=0,05 (area sottesa dalla curva)

Più in generale:



Errori di I e II tipo

In un test di ipotesi, si possono commettere due tipi di errore:

	H_0 è vera	H_0 è falsa
Non rifiuto di H_0	**DECISIONE CORRETTA**	Errore di II tipo = β
Rifiuto di H_0	Errore di I tipo = α	**DECISIONE CORRETTA** Potenza = $(1-\beta)$

Errori di I e II tipo

- **Errore di I tipo**: rifiuto un'ipotesi quando essa è vera
- **Errore di II tipo**: accetto un'ipotesi quando essa è falsa
- La **probabilità di errore di I tipo**, α, è detta **"livello di significatività"** e, come si può notare, è affine al "grado di confidenza" che abbiamo visto negli intervalli di confidenza.

Errori di I e II tipo

- α = Pr(rifiutare H_0| H_0 è vera)
- Più α sarà piccolo, minore sarà la probabilità di rifiutare erroneamente l'ipotesi nulla.
- **Probabilità di errore di II tipo**:
- β = Pr(accettare H_0| H_0 è falsa)
- **Potenza del test** = (1- β)

Errori di I e II tipo

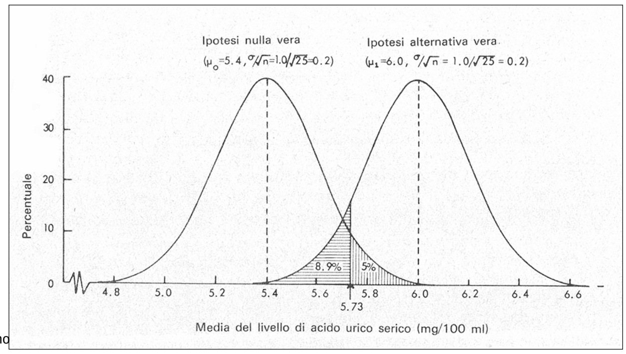
- Se rifiutiamo l'ipotesi nulla H_0 :
  - H_0 era falsa → OK!
  - **H_0 era vera → errore di 1° tipo**

- Se "accettiamo" l'ipotesi nulla H_0 :
  - **H_0 era falsa → errore di 2° tipo**
  - H_0 era vera → OK!

  All'aumentare di alfa, beta diminuisce.

Errori di I e II tipo: li vediamo attraverso una rappresentazione grafica

- Nell'esempio riportato in figura, vengono formulate due ipotesi, H_0 e H_1, che danno luogo, nel caso si prenda α = 0,05, a una probabilità di errore di II tipo β = 0,089, e quindi a una potenza del test (1- β)= 0,911= **91,1%**



Esempio: test di ipotesi su una media (test "z") – Varianza nota

- È noto che il livello medio di acido urico serico, nei maschi sani, è pari a 5,4 mg per 100 ml (μ).
- È noto che tale variabile si distribuisce **normalmente**.
- È anche noto che, nella popolazione generale "sana", si ha (varianza nota): σ = 1,0.
- Pertanto, si ha:
- $\mu = 5{,}4$
- $\sigma = 1{,}0$

Esempio: test di ipotesi su una media (test "z") – Varianza nota

- In un campione di 25 maschi, diabetici, è stato osservato un livello medio di acido urico serico pari a 5,9.
- Formuliamo pertanto la seguente **ipotesi**: il valore medio di acido urico nel campione di maschi diabetici (m_0) non è uguale, ma è MAGGIORE della media nota e prefissata della popolazione sana (test a una coda).
- Formalmente:
- **$H_0 : m_0 = \mu$ (IPOTESI NULLA)**
- **$H_1 : m_0 > \mu$ (IPOTESI ALTERNATIVA, test a una coda)**

Esempio: test di ipotesi su una media (test “z”) – Varianza nota

- Eseguiamo pertanto il test “z” (ricordiamo che la “z” è la distribuzione normale standardizzata)
- Calcoliamo la nostra statistica-test sui dati:

$$z = \frac{m_0 - \mu}{\sigma / \sqrt{n}} = \frac{5{,}9 - 5{,}4}{1/\sqrt{25}} = \frac{0{,}5}{1/5} = \frac{0{,}5}{0{,}2} = 2{,}50$$

Fissiamo α=0,05 quale **livello di significatività**.

Osservazione

- Il livello del 5 % di α indica che vi sono 95 probabilità su 100 che il risultato ottenuto sul mio campione non sia casuale.
- Tuttavia, esistono sempre 5 probabilità su 100 che tale risultato sia del tutto casuale.
- Se p <0,05, il caso è una spiegazione improbabile dei risultati osservati. Decido pertanto di rifiuto H_0.
- Se p >0,05, il caso è invece una possibile spiegazione, e non posso rifiutare l’ipotesi nulla.

Esempio: test di ipotesi su una media (test "z") – Varianza nota

- Il valore empirico della statistica-test "z" è pari a **2,50**
- Sulla tavola della **z**, questo valore corrisponde a una probabilità (area sottesa dalla curva normale standardizzata) pari a 0,006: un valore MOLTO INFERIORE al valore-soglia (α = 0,05) che avevamo prefissato.
- Possiamo pertanto concludere con la **confutazione** dell'ipotesi nulla di uguaglianza della **media calcolata nel campione dei diabetici** rispetto a un **valore prefissato di** μ (la media nella popolazione generale sana).
- I maschi diabetici compresi nel campione hanno un valore medio **significativamente più alto** dell'acido urico (**test a una coda**) rispetto al **valore noto prefissato** della popolazione dei maschi sani.

Il "p-value"

- **NB**: il valore **esatto** della probabilità in corrispondenza della statistica-test calcolata sui dati (in questo caso, **0,006**), ottenuto da una tavola oppure attraverso un software statistico, è detto ***p*-value**.
- **NB**: da 20 anni a questa parte, grazie allo sviluppo dei software statistici, è praticamente **<u>sempre</u>** possibile ottenere il ***p*-value esatto** del test, senza dover necessariamente fare ricorso a una tavola.

Il p-value

- Il “p-value” è si basa sul principio che, data la distribuzione di probabilità di interesse, è possibile calcolare l’esatta probabilità di ottenere per solo effetto del caso il risultato osservato nel campione (partendo dall’ipotesi che H_0 sia vera).
- Il p-value dipende dall’ampiezza del campione.

La tavola della “z” – Unilaterale, a destra (per test a una coda; per test a due code, basta raddoppiare il valore dell’area sottesa)

Tavola della Probabilita' di Z - Normale Standardizzata - Unilaterale

intero e primo decimale	secondo decimale di z									
	0,00	***0,01***	***0,02***	***0,03***	***0,04***	***0,05***	***0,06***	***0,07***	***0,08***	***0,09***
0,0	0.500	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464
0,1	0.460	0.456	0.452	0.448	0.444	0.440	0.436	0.433	0.429	0.425
0,2	0.421	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.390	0.386
0,3	0.382	0.378	0.374	0.371	0.367	0.363	0.359	0.356	0.352	0.348
0,4	0.345	0.341	0.337	0.334	0.330	0.326	0.323	0.319	0.316	0.312
0,5	0.309	0.305	0.302	0.298	0.295	0.291	0.288	0.284	0.281	0.278
0,6	0.274	0.271	0.268	0.264	0.261	0.258	0.255	0.251	0.248	0.245
0,7	0.242	0.239	0.236	0.233	0.230	0.227	0.224	0.221	0.218	0.215
0,8	0.212	0.209	0.206	0.203	0.200	0.198	0.195	0.192	0.189	0.187
0,9	0.184	0.181	0.179	0.176	0.174	0.171	0.169	0.166	0.164	0.161
1,0	0.159	0.156	0.154	0.152	0.149	0.147	0.145	0.142	0.140	0.138
1,1	0.136	0.133	0.131	0.129	0.127	0.125	0.123	0.121	0.119	0.117
1,2	0.115	0.113	0.111	0.109	0.107	0.106	0.104	0.102	0.100	0.099
1,3	0.097	0.095	0.093	0.092	0.090	0.089	0.087	0.085	0.084	0.082
1,4	0.081	0.079	0.078	0.076	0.075	0.074	0.072	0.071	0.069	0.068
1,5	0.067	0.066	0.064	0.063	0.062	0.061	0.059	0.058	0.057	0.056
1,6	0.055	0.054	0.053	0.052	0.051	0.049	0.048	0.048	0.046	0.046
1,7	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.037	0.037
1,8	0.036	0.035	0.034	0.034	0.033	0.032	0.031	0.030	0.029	0.029
1,9	0.029	0.028	0.027	0.027	0.026	0.026	0.025	0.024	0.024	0.023
2,0	0.023	0.022	0.022	0.021	0.021	0.020	0.020	0.019	0.019	0.018
2,1	0.018	0.017	0.017	0.017	0.016	0.016	0.015	0.015	0.015	0.014
2,2	0.014	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011
2,3	0.011	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.008
2,4	0.008	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006
2,5	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
2,6	0.005	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
2,7	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
2,8	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
2,9	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
3,0	0.001									

Esempio: test di ipotesi su una media (test "t") – Varianza incognita

- Spesso, nella pratica statistica, si devono fare inferenze su una media, senza però conoscere – come invece è stato nell'esempio appena visto – la varianza della popolazione. In questo caso, la varianza non nota della popolazione dovrà essere stimata, utilizzando allo scopo la varianza empirica, calcolata nel campione.

Esempio: test di ipotesi su una media (test "t") – Varianza incognita

- In questo caso, la **statistica-test** non è più la statistica "z" (normale standardizzata), ma la statistica "t", detta "t di Student" (pseudonimo di William Sealy Gosset), che segue una sua peculiare distribuzione teorica (la distribuzione omonima).

- La distribuzione "t" varia in funzione dei "gradi di libertà": gl = ν = n-1, cioè della numerosità campionaria.

Esempio: test di ipotesi su una media (test "t") – Varianza incognita

- La statistica-test è:

$$t_{n-1} = \frac{\mu - \mu_0}{s / \sqrt{n}}$$

- Dove:
- *s* non è più la radice quadrata della varianza di popolazione, ma una sua stima, cioè la d.s. calcolata empiricamente sul campione; si noti che è importante tenere conto anche del numero di gradi di libertà, che è dato da (n-1)

NB: calcolo della varianza campionaria corretta

- Ricordiamo che, per stimare il parametro σ^2, lo stimatore varianza campionaria è uno stimatore distorto.
- Tuttavia, applicando un fattore di correzione, si può calcolare una varianza campionaria corretta. La varianza campionaria corretta si ottiene con la seguente formula:

$$s^2_{corretta} = \frac{\sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2}{n-1}$$

Esempio: test di ipotesi su una media (test "t") – Varianza incognita

- In un campione di n=10 maschi è stata misurata una frequenza di pulsazione di 68,7 battiti per minuto, con **d.s.** pari a 8,67;
- Il valore "normale" è indicato in **72 bpm**.
- Si vuole saggiare l'ipotesi che il nostro campione differisca significativamente (in senso negativo) dal valore "normale" di 72 bpm (test a una coda).
- Calcoliamo la "t" empirica.

$$t_{n-1} = \frac{\mu - \mu_0}{s/\sqrt{n}} =$$

$$= \frac{72 - 68{,}7}{8{,}67/\sqrt{10}} = -\frac{3{,}3}{2{,}74} = -1{,}20$$

Esempio: test di ipotesi su una media (test "t") – Varianza incognita

- Il valore tabulato della "t" teorica, con livello di significatività α **= 0,05** e
- **n-1 = 10-1 = 9** gradi di libertà
- è pari a **-1,8331 (test a una coda)**
- **Il nostro valore empirico si trova <u>a destra</u> di quello tabulato; pertanto, l'ipotesi nulla, di media pari a 72 bpm, NON È CONFUTATA.**

La tavola della “t di Student” – Unilaterale, a destra

(per test a una coda**; per test a due code, basta considerare il valore dell’area tabulato pari a ½ del valore prescelto; ad es., se** α=0,05 **per un test a una coda, per due code si deve prendere il valore della tavola pari a** α =0,025**.)**

Gradi di libertà	Area della coda di destra					
	0.25	0.1	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.8165	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.7407	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.6800	1.3007	1.6794	2.0141	2.4121	2.6896
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6800
50	0.6794	1.2987	1.6759	2.0086	2.4033	2.6778

La relazione tra due variabili su scala nominale.

Il test “chi quadrato” (χ^2) per saggiare l’ipotesi di indipendenza distributiva in una distribuzione doppia (tabella di contingenza)

- Per tutti i test di ipotesi che abbiamo visto finora, una condizione ineludibile era che le variabili fossero QUANTITATIVE e, in particolare, la condizione per potere applicare i test era che la scala di misura dei dati da analizzare fosse – almeno – una *scala ad intervalli*.
- Nei problemi concreti di ricerca, però, spesso i dati che dobbiamo analizzare non sono né su una scala ad intervalli, né su una scala ordinale, ma sono semplicemente costituiti da “**tabelle di contingenza**”, ovvero da quelle “tabelle a doppia entrata” che si ottengono dall’**incrocio** di due variabili nominali.

Un esempio

- Nel 1969 (Dati da Jick e altri, Lancet, 1969) sono state osservate 200 donne che assumevano contraccettivi orali; di queste, 145 erano sane, mentre 55 erano affette da tromboembolia. Le 200 donne, sane e ammalate, sono state classificate secondo il gruppo sanguigno.

Tabella delle frequenze osservate

Gruppo sanguigno	Donne trombo-emboliche	Donne sane	Totale
A	32	51	**83**
B	8	19	**27**
AB	6	5	**11**
0	9	70	**79**
Totale	**55**	**145**	**200**

Per prima cosa, confrontiamo le due distribuzioni (ammalate e sane), osservando le percentuali di colonna

Gruppo sangui-gno	Donne trombo-emboliti-che	**%**	Donne sane	**%**	Totale	**%**
A	32	**58,2**	51	**35,2**	83	**41,5**
B	8	**14,5**	19	**13,1**	27	**13,5**
AB	6	**10,9**	5	**3,4**	11	**5,5**
0	9	**16,4**	70	**48,3**	79	**39,5**
Totale	**55**	**100,0**	**145**	**100,0**	**200**	**100,0**

L'ipotesi di non indipendenza

- È facile osservare che il gruppo sanguigno **A** è molto più frequente della media nel gruppo delle donne tromboembolitiche, mentre il gruppo **0** è molto meno frequente;
- L'ipotesi è che l'essere di gruppo sanguigno A sia un fattore di rischio, per donne che assumono contraccettivi orali, mentre essere del gruppo 0 sia un fattore protettivo;
- L'ipotesi, pertanto, è che le due variabili nominali (condizione di malattia e gruppo sanguigno) NON SIANO TRA LORO INDIPENDENTI.

- In generale, la funzione del test "chi quadrato" consiste nel valutare se vi sia (o meno) una **differenza statisticamente significativa** tra due distribuzioni, una empirica e una "teorica"
- Ricordiamo che una **tabella di contingenza** esprime **una distribuzione**, seppure operata tramite la combinazione di due caratteri.
- Di queste due distribuzioni, una è formata dai valori "empirici", ovvero dai valori **reali** del campione, mentre l'altra "obbedisce" ad un qualche "modello" teorico di riferimento, oppure a una certa condizione teorica.
- Quest'ultima distribuzione **non è, pertanto, una distribuzione reale, ma "obbedisce" ad una qualche ipotesi che viene formulata sui dati**.
- Uno dei "modelli" teorici più utilizzati esprime semplicemente l'idea che i due fenomeni siano tra loro ***indipendenti***.

- Nel nostro esempio, il problema è quello di valutare la **relazione** tra i due fenomeni (**diagnosi di tromboembolia** delle 200 donne e **gruppo sanguigno di appartenenza**), ovvero di saggiare l'ipotesi di ***non indipendenza*** tra le due variabili su scala nominale.
- A tale scopo, si rende necessario stabilire come *sarebbe* conformata la distribuzione nel caso, appunto teorico, di ***indipendenza*** tra i due caratteri.

Calcolo delle frequenze attese

- Applicando una nota regola del calcolo delle probabilità (la regola del prodotto), le **frequenze attese** (teoriche) nel caso di *indipendenza* tra i due caratteri si possono calcolare con la seguente formula:

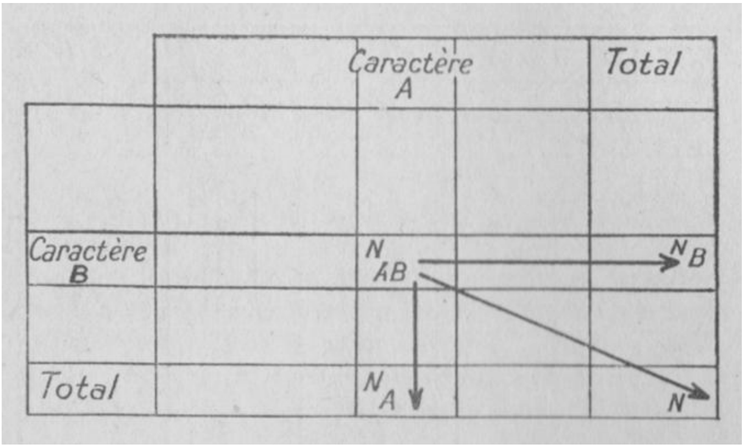
$$FA_{ij} = \frac{FO_{i\cdot} \times FO_{\cdot j}}{n}$$

Calcolo delle frequenze attese

- Dove $FO_{i.}$ e $FO_{.j}$ sono, rispettivamente, i totali **di riga** e **di colonna** (i cosiddetti **totali marginali**) della tabella di contingenza, mentre ***n*** è la numerosità totale del campione.

- Il calcolo delle frequenze "attese" secondo l'ipotesi di indipendenza stocastica avviene moltiplicando tra loro i due totali marginali, dividendo il tutto per *n*; può essere utile lo schema rappresentato nella figura che segue:

Calcolo delle frequenze attese



Calcolo delle frequenze attese

- Utilizzando la formula appena vista, si può “riscrivere” la tabella di contingenza, come se fosse la tabella “attesa” che si otterrebbe nel caso di TOTALE INDIPENDENZA (indipendenza stocastica) tra le due variabili nominali.
- Quella che segue è la tabella teorica, calcolata nel modo appena visto.

Tabella delle frequenze attese: i valori in rosso sono i valori "attesi"

Gruppo sanguigno	Donne trombo-emboliche	Donne sane	Totale
A	**22,8**	**60,2**	83
B	**7,4**	**19,6**	27
AB	**3,0**	**8,0**	11
0	**21,7**	**57,3**	79
Totale	**55**	**145**	**200**

Il test "chi-quadrato"

- Il valore della statistica-test "chi quadrato", se si dispone di una tabella di contingenza formata da *p* righe e *k* colonne, si ottiene applicando la formula che segue :

$$\chi^2 = \sum_{i=1}^{p \times k} \frac{\left(FO_i - FA_i\right)^2}{FA_i}$$

- Dove :
- FO_i è la frequenza osservata (empirica) che si trova nella i-esima casella della tabella ;
- FA_i è la frequenza attesa (teorica) che si trova nella i-esima casella.

Il test “chi-quadrato”

- Nel nostro esempio, il valore empirico della statistica-test “chi quadrato” è pari a **19,46**.
- Come possiamo, ora, stabilire se il test si può considerare statisticamente significativo oppure no ?
- Per prima cosa, dovremo stabilire quanti **gradi di libertà** vanno considerati.
- Nel caso del test “chi quadrato”, la formula per il calcolo dei gradi di libertà è la seguente:
- $Gl = (p - 1) \times (k - 1)$
- Dove *p* e *k* sono, rispettivamente, il numero di righe e il numero di colonne della tabella di contingenza.

Il test “chi-quadrato”

- Nel caso del nostro esempio, visto che la tabella di contingenza ha **4** righe e **2** colonne, avremo *tre* gradi di libertà.
- Una volta stabilito il numero di gradi di libertà, come abbiamo fatto per il test “t” di Student, dovremo scegliere il **livello di significatività** che vogliamo attribuire al nostro test.
- Anche per il test “chi quadrato” vale la stessa logica: più piccolo sarà il valore che sceglieremo, più “garantita” sarà la significatività del test; più piccolo sarà il valore del livello di significatività, α, meno sarà probabile la **confutazione erronea** dell’ipotesi nulla.

Il test "chi-quadrato"

- Giunti a questo punto, andiamo a consultare la tavola dei **valori teorici** della distribuzione "chi quadrato".
- Per meglio valutare la "forza" della nostra ipotesi (di non indipendenza), scegliamo un valore del livello di significatività molto "sicuro", ovvero molto basso: α **= 0,001**.
-
- Come possiamo vedere nella tavola, riprodotta di seguito, il valore di "chi quadrato" corrispondente a tale livello (e a **3** gradi di libertà) è **16,27**.

Nella prima colonna sono indicati i gradi di libertà (*gl*), nella prima riga i possibili livelli di significatività (*p*).
All'incrocio tra il numero di gradi di libertà e il livello di significatività si trova il valore critico del χ^2.

	Probabilità *p* di essere nella zona di rifiuto individuata dal valore critico tabulato							
gl	0.500	0.250	0.100	0.050	0.025	0.010	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	4.35	6.63	9.24	11.07	12.83	15.09	16.75	20.51
6	5.35	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	6.35	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	7.34	10.22	13.36	15.51	17.53	20.09	21.95	26.12
9	8.34	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	9.34	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	10.34	13.70	17.28	19.68	21.92	24.73	26.76	31.26
12	11.34	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	12.34	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	13.34	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	14.34	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	15.34	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	16.34	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	17.34	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	18.34	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	19.34	23.83	28.41	31.41	34.17	37.57	40.00	45.31
21	20.34	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	21.34	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	22.34	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	23.34	28.24	33.20	36.42	39.36	42.98	45.56	51.18
25	24.34	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	25.34	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	26.34	31.53	36.74	40.11	43.19	46.96	49.65	55.48
28	27.34	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	28.34	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	29.34	34.80	40.26	43.77	46.98	50.89	53.67	59.70
31	30.34	35.89	41.42	44.99	48.23	52.19	55.00	61.10
32	31.34	36.97	42.58	46.19	49.48	53.49	56.33	62.49
33	32.34	38.06	43.75	47.40	50.73	54.78	57.65	63.87
34	33.34	39.14	44.90	48.60	51.97	56.06	58.96	65.25
35	34.34	40.22	46.06	49.80	53.20	57.34	60.27	66.62
36	35.34	41.30	47.21	51.00	54.44	58.62	61.58	67.98
37	36.34	42.38	48.36	52.19	55.67	59.89	62.88	69.35
38	37.34	43.46	49.51	53.38	56.90	61.16	64.18	70.70
39	38.34	44.54	50.66	54.57	58.12	62.43	65.48	72.06
40	39.34	45.62	51.81	55.76	59.34	63.69	66.77	73.40
50	49.33	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	59.33	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	69.33	77.58	85.53	90.53	95.02	100.43	104.21	112.32
80	79.33	88.13	96.58	101.88	106.63	112.33	116.32	124.84
90	89.33	98.65	107.57	113.15	118.14	124.12	128.30	137.21
100	99.33	109.14	118.50	124.34	129.56	135.81	140.17	149.45

Il test "chi-quadrato"

- Si tratta di un valore, seppure non di molto, inferiore a quello empirico (**19,46**). Ricordiamo, poi, che abbiamo scelto un valore molto basso di α (α=0,001, cioè uno per mille), che praticamente ci mette in una "botte di ferro".
- Possiamo pertanto considerare **confutata** l'ipotesi nulla.
- Possiamo, quindi, concludere che, tra le donne che fanno uso di contraccettivi orali, *vi è una certa dipendenza*, espressa in modo **statisticamente significativo**, tra diagnosi di tromboembolia e gruppo sanguigno di appartenenza.

Intervalli di confidenza e test di ipotesi per il coefficiente di correlazione

- Anche sul coefficiente di correlazione, *r*, (come del resto anche sul coefficiente di regressione, *b*) è possibile fare inferenze.
- Per il coefficiente di correlazione, esiste la possibilità di ricavare i relativi **intervalli di confidenza** e di effettuare i relativi **test di ipotesi**;
- Di solito, l'**ipotesi nulla** che viene saggiata è quella di uguaglianza a zero del coefficiente di correlazione nella popolazione di riferimento, cioè, in simboli:
- $\mathbf{H_0: \rho = 0}$

Un esempio

- L'articolo, del 1997, dal titolo:
- **Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group**
- è tratto dall'**American Journal of Epidemiology**

- Nella tabella sono riportate, secondo alcune caratteristiche socio-demografiche, le correlazioni – in 754 bambini affetti da leucemia – tra **età gestazionale riferita telefonicamente dalle madri** ed **età gestazionale riportata nelle corrispondenti cartelle cliniche**.

- Nella tabella sono riportati gli intervalli di confidenza, al **98%**, per *r*.

TABLE 3. Validity and reliability of gestational age within specific demographic subgroupings among participating members from a United States and Canadian cooperative clinical trials group and matched controls, 1983–1988

	Correlation of gestational age	98% CI*	Kappa statistic†
All gestational ages	0.839	0.817–0.859	0.62
Case/control status			
Cases	0.849	0.813–0.878	0.63
Controls	0.835	0.805–0.861	0.61
Education			
<High school	0.694	0.553–0.797	0.51
High school	0.833	0.790–0.868	0.63
>High school	0.835	0.804–0.861	0.62
Household income			
<$22,000	0.791	0.734–0.837	0.59
$22,000–$34,999	0.882	0.849–0.908	0.62
≥$35,000	0.843	0.800–0.877	0.65
Unknown	0.745	0.641–0.823	0.60
Time (years) from delivery to interview			
<2	0.896	0.862–0.921	0.64
2–3.9	0.821	0.784–0.852	0.63
4–5.9	0.828	0.775–0.869	0.61
6–8	0.852	0.734–0.920	0.42
Maternal age (years)			
<25	0.822	0.773–0.861	0.64
25–29	0.889	0.862–0.912	0.63
30–34	0.760	0.694–0.813	0.57
≥35	0.888	0.824–0.930	0.64
Birth order			
First born	0.880	0.853–0.903	0.67
Second born	0.815	0.778–0.846	0.57
≥Third born	0.632	0.416–0.781	0.52
Maternal race			
White	0.846	0.822–0.866	0.64
Other	0.782	0.680–0.855	0.42

* CI, confidence interval.