40 Alcuni Test di Student

40.1 Introduzione

Nelle Scienze Farmaceutiche, i Test *t di Student* sono uno strumento statistico fondamentale per prendere decisioni basate sui dati.

Vengono utilizzati per confrontare la media di un gruppo con un valore, oppure le medie di due gruppi, ad esempio per valutare se un nuovo farmaco ha un effetto significativamente diverso da un placebo o da un farmaco di riferimento.

Ci sono 3 tipi di Test t di Student:
– a campione singolo
– per campioni indipendenti (soggetti diversi)
– per campioni appaiati (tipicamente: prima/dopo).

40.2 Test t di Student a una coda

Nel **Test t di Student a una coda** l'obiettivo è verificare se la media di una popolazione sia *maggiore* (oppure *minore*) di un valore di riferimento prefissato μ_0. Questo tipo di test è appropriato quando, sulla base di considerazioni scientifiche o regolatorie, è rilevante una sola direzione dell'effetto.
Per esempio, nel Controllo Qualità (QC) farmaceutico, il Test t a campione singolo è usato per verificare se la concentrazione media di un principio attivo in un lotto di produzione ($\bar{x}$) sia *superiore* a un valore target minimo specificato (μ_0).
In questo contesto, le ipotesi statistiche sono formulate in modo *unilatero* come

$$H_0: \ \mu \leq \mu_0 \qquad \text{vs} \qquad H_1: \ \mu > \mu_0.$$

La *statistica di test* è

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}},$$

oppure, in forma equivalente,

$$\boxed{T = \sqrt{n}\,\frac{\bar{X} - \mu_0}{S}} \tag{159}$$

dove S è la deviazione standard campionaria (radice quadrata di S^2). Tale espressione può essere considerata la formula basale dei vari Test t di Student (ci sono variazioni nei vari tipi).
Se il valore osservato t di T supera tale quantile, si rifiuta l'ipotesi nulla; in caso contrario, non vi è evidenza statistica sufficiente per rifiutarla.
(Da un punto di vista teorico, sotto l'ipotesi nulla, la statistica di test più sopra scritta segue una distribuzione t di Student con $\nu = n - 1$ gradi di libertà).

Il criterio decisionale si basa sul **quantile critico** della distribuzione t di Student, che si determina con tavole o software, dopo aver fissato un livello di significatività $\alpha = 0.05$.

(Tecnicamente $t_{1-\alpha,\ \nu}$, che rappresenta il quantile al 95% della distribuzione t con ν gradi di libertà).

La **regione critica** del test è definita da:

$$\boxed{T > \text{quantile critico} \quad (*)} \tag{160}$$

In particolare:

- se la soprastante $(*)$ è vera si rifiuta l'ipotesi nulla H_0 e si conclude che la media è significativamente maggiore di μ_0;
- se la soprastante $(*)$ è falsa, non vi è evidenza statistica sufficiente per rifiutare H_0.

Questo approccio è largamente utilizzato in ambito farmaceutico quando si deve dimostrare che un parametro di qualità (ad esempio dissoluzione, potenza o biodisponibilità) supera una soglia minima accettabile, piuttosto che verificare una differenza bidirezionale.

40.3 Inquadramento generale

– 1: Il ruolo del Test t nelle Scienze Farmaceutiche

Nelle Scienze Farmaceutiche, il Test *t di Student* è uno strumento statistico fondamentale per prendere decisioni basate sui dati. Viene utilizzato per confrontare la media di un gruppo con un valore, oppure le medie di due gruppi, ad esempio per valutare se un nuovo farmaco ha un effetto significativamente diverso da un placebo o da un farmaco di riferimento. La sua applicazione è cruciale in tutte le fasi dello sviluppo di un farmaco, dalla ricerca pre-clinica fino agli studi clinici sull'uomo e al controllo qualità.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 2: Studi clinici e confronto Farmaco vs Placebo

Uno degli usi più comuni del Test t è negli studi clinici randomizzati e controllati (RCT), dove si confronta l'efficacia media di un nuovo trattamento con quella di un placebo. Si utilizza il Test *t per campioni indipendenti* per verificare l'ipotesi nulla H_0 che non vi sia differenza tra la media del gruppo trattato ($\bar{x}_F$) e quella del gruppo placebo ($\bar{x}_P$).

Se il *p-value* è sufficientemente basso, si rifiuta H_0 e si conclude che il farmaco ha un effetto statisticamente significativo.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.

2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 3: Studi di bioequivalenza e disegno crossover

Gli studi di bioequivalenza dimostrano che un farmaco generico (Test, T) è equivalente al farmaco di riferimento (R). Tali studi impiegano spesso un disegno crossover, in cui ogni soggetto riceve entrambi i trattamenti in periodi diversi.

In questo contesto si utilizza il Test *t per campioni appaiati* per confrontare parametri farmacocinetici come l'AUC e la C_{max} all'interno dello stesso soggetto.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 4: Parametri farmacocinetici e intervalli di confidenza

Negli studi di bioequivalenza, l'obiettivo non è soltanto rifiutare l'ipotesi nulla, ma dimostrare che il rapporto tra le medie geometriche dei parametri farmacocinetici, ad esempio AUC_T/AUC_R, rientri in un intervallo prestabilito, tipicamente compreso tra l'80% e il 125%.

La decisione finale si basa sulla costruzione di un intervallo di confidenza al 90% per il rapporto T/R.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 5: Controllo qualità e Test t a campione singolo

Nel controllo qualità farmaceutico, il Test t a campione singolo viene utilizzato per verificare se la concentrazione media di un principio attivo in un lotto di produzione ($\bar{x}$) sia conforme a un valore target specificato (μ_0).

La statistica di test è:

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \quad \text{ovvero} \quad \sqrt{n}\,\frac{\bar{X} - \mu_0}{S}$$

dove S è la deviazione standard campionaria, radice quadrata di S^2.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 6: Confronto tra metodi analitici

Quando si sviluppa un nuovo metodo analitico per misurare la purezza o la concentrazione di un farmaco, è necessario confrontarlo con un metodo di riferimento. Analizzando lo stesso campione con entrambi i metodi, si applica il Test *t per campioni appaiati* per valutare l'eventuale presenza di un bias sistematico.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 7: Assunzione di normalità

Una delle assunzioni fondamentali del Test t è che i dati (o le differenze, nel caso appaiato) seguano una distribuzione normale. Nelle Scienze Farmaceutiche, ciò riguarda spesso variabili come livelli di metaboliti, tempi di dissoluzione o risposte biologiche.
In caso di violazione di questa assunzione, si possono applicare trasformazioni dei dati o ricorrere a test non parametrici.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 8: Omoschedasticità e correzione di Welch

Il Test t per campioni indipendenti standard assume l'uguaglianza delle varianze (omoschedasticità). Quando tale assunzione è violata, è preferibile utilizzare il Test *t di Welch*, che non richiede varianze uguali.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 9: Studi pre-clinici

Nella fase pre-clinica, il Test t è ampiamente utilizzato per confrontare l'effetto di diverse dosi o composti su modelli animali. La ridotta dimensione campionaria rende la distribuzione t di Student particolarmente appropriata.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.
2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

– 10: Quando non usare il Test t

Il Test t è limitato al confronto di due medie. In presenza di tre o più gruppi di trattamento è necessario ricorrere all'Analisi della Varianza (ANOVA). Inoltre, se le assunzioni del Test t sono fortemente violate, si utilizzano test non parametrici.

Domanda di verifica In quale fase dello sviluppo di un farmaco il Test t non è tipicamente utilizzato?

1. Fase di sintesi chimica del principio attivo.

2. È utilizzato in tutte le fasi, dalla pre-clinica al controllo qualità.
3. Fase di registrazione del brevetto.
4. Fase di marketing e promozione.

40.4 Note euristiche sui Test di Student

Consideriamo la formula (159)-(160).
– Supponiamo dapprima fissato il livello α.
Con n grande, T tende a superare il quantile anche con piccola superiorità di $\bar{X}_n$ su μ_0.
Con S^2 grande, T stenta a superare il quantile anche con grande superiorità di $\bar{X}_n$ su μ_0. (Dati inconsistenti, si contraddicono).
– Ma anche α partecipa, e infatti è nel quantile.
Più α è piccolo, ovvero più è grande il livello di confidenza $1-\alpha$ ricercato, più cresce il quantile ovvero la soglia da superare (col triplice contributo combinato di grande n, grande sopravanzare di $\bar{X}_n$ su μ_0, e/o piccolo S^2).

Nota. Di questa Lezione ci si dovrebbe aspettare che lo studente conosca l'inquadramento teorico della questione trattata e conosca a memoria le formule contemporaneamente numerate *et* riquadrate.

Il Corso continua e finisce alla pagina seguente.

40.5 Come misurare l'esposizione?

Inseriamo in questa Lezione una questione di portata molto generale in Medicina e Farmacia.
Supponiamo di voler testare, magari proprio col gold standard, un trial clinico randomizzato in doppio cieco, un farmaco per grandi fumatori, magari per ridurre gli episodi di bronchite, o quant'altro. Ma chi sono i "grandi fumatori", o anche semplicemente i fumatori? Chi fuma più di 10 sigarette al giorno, oppure 20, oppure... quante? Qua si apre un immenso e catastrofico capitolo della Medicina e della Farmacia: come misurare l'esposizione, per esempio all'amianto, oppure al fumo? Considereremo fumatore chi fuma più di 0 o di 10 sigarette al giorno? È chiaro che modulando questi parametri si può largamente pilotare il risultato, eventualmente a favore del finanziatore delle ricerche. E infatti negli articoli scientifici spesso vi è una dichiarazione, eventualmente fasulla, di mancanza di conflitto di interessi. Fumatore, consumatore di carne o pesce o alcol o tranquillanti o qualsiasi cosa... in tutto si potranno scegliere arbitrariamente soglie. Per esempio, qualche anno fa è diventato famosissimo un articolo scientifico che trovava che mangiare molta carne rossa aumenterebbe la mortalità. Ma chi deve essere classificato grande mangiatore di carne rossa, e chi scarso mangiatore di carne rossa, con cui fare il confronto?
E così pure – con meno arbitrarietà perché su queste cose vi sono alcuni standard di enti autorevoli – per la carenza di vitamina D o B12 o quant'altro.

40.6 Bottom line – Legge pseudo-magica della Statistica

Se c'è un interesse che un dato si formi quel dato tenderà a formarsi

FINE DEL CORSO

40.7 Complementi – Esempio di test a 2 code

Leggiamo (10 gennaio 2023) in
https://www.jmp.com/it_it/statistics-knowledge-portal/t-test.html

> La frequenza cardiaca media di un gruppo di persone è uguale a 65

Il test è a 2 code perchè rifiuteremo l'ipotesi nulla dell'uguaglianza sia se troveremo una media troppo minore di 65 sia troppo maggiore.

40.8 Complementi – Esempio estremo per fissare le idee

Con $n = 2$, due soli dati, è $n - 1 = 1$ e ci serve $t_\alpha(1)$ da trovarsi con le tavole o i software o addirittura coi calcoli perché $f(t;1)$ è la densità di Cauchy:

$$f(t;1) = \frac{1}{\pi}\frac{1}{1+t^2}$$

$$F(x) = \int_{-\infty}^{x} f(t;1) = \int_{-\infty}^{x} \frac{1}{\pi}\frac{1}{1+t^2}\,dt =$$

$$= \left[\frac{1}{\pi}\arctan t\right]_{-\infty}^{x} = \frac{1}{\pi}\arctan x - \frac{1}{\pi}\arctan(-\infty) =$$

$$= \frac{1}{2} + \frac{1}{\pi}\arctan x$$

e la sua inversa dà i quantili di student $t_\alpha(1)$ risolvendo

$$F(x) \overset{EQ}{=} \alpha \qquad > 0.5 \text{ (“grande")}$$

$$\frac{1}{2} + \frac{1}{\pi}\arctan x = \alpha \quad \Big/ + \left(-\frac{1}{2}\right)$$

$$\frac{1}{\pi}\arctan x = \alpha - \frac{1}{2} \quad \Big/ \cdot \pi$$

$$\arctan x = \pi\left(\alpha - \frac{1}{2}\right)$$

un'attenta analisi mostrerebbe che possiamo applicare la tangente

$$\Big/ \tan \qquad \text{(che “mangia" l'arcotangente)}$$

$$x = \tan\left(\pi\left(\alpha - \frac{1}{2}\right)\right)$$

che è il quantile di Student cercato,

$$t_\alpha(1) = \tan\left(\pi\left(\alpha - 0.5\right)\right) \qquad \alpha > 0.5$$

e alcuni valori li avremo in forma esatta:
($t_{0.5}(1) = 0$, basta che $\bar{X}_n = \bar{X}_2 > \mu_0$, ma l'affermazione è irrilevante per troppo basso livello di confidenza, 50%)
$t_{\frac{2}{3}}(1) = \frac{\sqrt{3}}{3} \approx 0.577$
$t_{\frac{5}{6}}(1) = \sqrt{3} \approx 1.73$
($t_1(1) = +\infty$, cioè mai: livello di confidenza voluto troppo alto)
ottendendosi questa tavola dei quantili di Student

α	0.5	0.667	0.833	1
$t_\alpha(1)$	0	0.577	1.73	$+\infty$

che fanno capire chiaramente la situazione. Ma nessuno di quei 4 quantili è usato nella pratica. Più significativamente, con una calcolatrice che ci dia la tangente di

$\pi\left(0.95 - 0.5\right)$
$\pi\left(0.99 - 0.5\right)$
$\pi\left(0.995 - 0.5\right)$

otteniamo questa tavola dei quantili di Student

α	0.95	0.99	0.995
$t_\alpha(1)$	6.31375	31.8205	63.6567

Li troviamo confermati nella tavola della Lezione (**??**). Indicati con “one tail", cioè test “a una coda", perché indaghiamo sul $\leq$ e non $\neq$, che sarebbe a 2 code.

Nel seguente paragrafo diamo una seguente più completa tavola dei quantili di Student.

40.9 Complementi – Tavola dei quantili di Student

Dal sito del NIST in https://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm traiamo la seguente tavola dei quantili di Student.

ν	0.90	0.95	0.975	0.99	0.995	0.999
1.	3.078	6.314	12.706	31.821	63.657	318.313
2.	1.886	2.920	4.303	6.965	9.925	22.327
3.	1.638	2.353	3.182	4.541	5.841	10.215
4.	1.533	2.132	2.776	3.747	4.604	7.173
5.	1.476	2.015	2.571	3.365	4.032	5.893
6.	1.440	1.943	2.447	3.143	3.707	5.208
7.	1.415	1.895	2.365	2.998	3.499	4.782
8.	1.397	1.860	2.306	2.896	3.355	4.499
9.	1.383	1.833	2.262	2.821	3.250	4.296
10.	1.372	1.812	2.228	2.764	3.169	4.143
11.	1.363	1.796	2.201	2.718	3.106	4.024
12.	1.356	1.782	2.179	2.681	3.055	3.929
13.	1.350	1.771	2.160	2.650	3.012	3.852
14.	1.345	1.761	2.145	2.624	2.977	3.787
15.	1.341	1.753	2.131	2.602	2.947	3.733
16.	1.337	1.746	2.120	2.583	2.921	3.686
17.	1.333	1.740	2.110	2.567	2.898	3.646
18.	1.330	1.734	2.101	2.552	2.878	3.610
19.	1.328	1.729	2.093	2.539	2.861	3.579
20.	1.325	1.725	2.086	2.528	2.845	3.552
21.	1.323	1.721	2.080	2.518	2.831	3.527
22.	1.321	1.717	2.074	2.508	2.819	3.505
23.	1.319	1.714	2.069	2.500	2.807	3.485
24.	1.318	1.711	2.064	2.492	2.797	3.467
25.	1.316	1.708	2.060	2.485	2.787	3.450
26.	1.315	1.706	2.056	2.479	2.779	3.435
27.	1.314	1.703	2.052	2.473	2.771	3.421
28.	1.313	1.701	2.048	2.467	2.763	3.408
29.	1.311	1.699	2.045	2.462	2.756	3.396
30.	1.310	1.697	2.042	2.457	2.750	3.385
31.	1.309	1.696	2.040	2.453	2.744	3.375
32.	1.309	1.694	2.037	2.449	2.738	3.365
33.	1.308	1.692	2.035	2.445	2.733	3.356
34.	1.307	1.691	2.032	2.441	2.728	3.348
35.	1.306	1.690	2.030	2.438	2.724	3.340
36.	1.306	1.688	2.028	2.434	2.719	3.333
37.	1.305	1.687	2.026	2.431	2.715	3.326
38.	1.304	1.686	2.024	2.429	2.712	3.319
39.	1.304	1.685	2.023	2.426	2.708	3.313
40.	1.303	1.684	2.021	2.423	2.704	3.307
41.	1.303	1.683	2.020	2.421	2.701	3.301
42.	1.302	1.682	2.018	2.418	2.698	3.296
43.	1.302	1.681	2.017	2.416	2.695	3.291

44.	1.301	1.680	2.015	2.414	2.692	3.286
45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265
50.	1.299	1.676	2.009	2.403	2.678	3.261
51.	1.298	1.675	2.008	2.402	2.676	3.258
52.	1.298	1.675	2.007	2.400	2.674	3.255
53.	1.298	1.674	2.006	2.399	2.672	3.251
54.	1.297	1.674	2.005	2.397	2.670	3.248
55.	1.297	1.673	2.004	2.396	2.668	3.245
56.	1.297	1.673	2.003	2.395	2.667	3.242
57.	1.297	1.672	2.002	2.394	2.665	3.239
58.	1.296	1.672	2.002	2.392	2.663	3.237
59.	1.296	1.671	2.001	2.391	2.662	3.234
60.	1.296	1.671	2.000	2.390	2.660	3.232
61.	1.296	1.670	2.000	2.389	2.659	3.229
62.	1.295	1.670	1.999	2.388	2.657	3.227
63.	1.295	1.669	1.998	2.387	2.656	3.225
64.	1.295	1.669	1.998	2.386	2.655	3.223
65.	1.295	1.669	1.997	2.385	2.654	3.220
66.	1.295	1.668	1.997	2.384	2.652	3.218
67.	1.294	1.668	1.996	2.383	2.651	3.216
68.	1.294	1.668	1.995	2.382	2.650	3.214
69.	1.294	1.667	1.995	2.382	2.649	3.213
70.	1.294	1.667	1.994	2.381	2.648	3.211
71.	1.294	1.667	1.994	2.380	2.647	3.209
72.	1.293	1.666	1.993	2.379	2.646	3.207
73.	1.293	1.666	1.993	2.379	2.645	3.206
74.	1.293	1.666	1.993	2.378	2.644	3.204
75.	1.293	1.665	1.992	2.377	2.643	3.202
76.	1.293	1.665	1.992	2.376	2.642	3.201
77.	1.293	1.665	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195
81.	1.292	1.664	1.990	2.373	2.638	3.194
82.	1.292	1.664	1.989	2.373	2.637	3.193
83.	1.292	1.663	1.989	2.372	2.636	3.191
84.	1.292	1.663	1.989	2.372	2.636	3.190
85.	1.292	1.663	1.988	2.371	2.635	3.189
86.	1.291	1.663	1.988	2.370	2.634	3.188
87.	1.291	1.663	1.988	2.370	2.634	3.187
88.	1.291	1.662	1.987	2.369	2.633	3.185
89.	1.291	1.662	1.987	2.369	2.632	3.184
90.	1.291	1.662	1.987	2.368	2.632	3.183
91.	1.291	1.662	1.986	2.368	2.631	3.182
92.	1.291	1.662	1.986	2.368	2.630	3.181
93.	1.291	1.661	1.986	2.367	2.630	3.180
94.	1.291	1.661	1.986	2.367	2.629	3.179

95.	1.291	1.661	1.985	2.366	2.629	3.178
96.	1.290	1.661	1.985	2.366	2.628	3.177
97.	1.290	1.661	1.985	2.365	2.627	3.176
98.	1.290	1.661	1.984	2.365	2.627	3.175
99.	1.290	1.660	1.984	2.365	2.626	3.175
100.	1.290	1.660	1.984	2.364	2.626	3.174
$+\infty$	1.282	1.645	1.960	2.326	2.576	3.090

Si noti nell'ultima riga il limite all'infinito e in particolare il valore

$$1.96 \approx \phi_{0.975} = t_{0.975}(+\infty)$$

Esempio. Dal classico libro *Matematica e Statistica. Le basi per le Scienze della vita*, di Marco Abate.

> *Il proprietario dell'azienda vinicola da te preferita teme che il tasso alcolico del suo vino quest'anno possa non essere più pari al 12.5% indicato in etichetta e ti chiede d'investigare. Misurando il tasso alcolico di 6 bottiglie, ottiene i seguenti valori:*
>
> $$11.5,\ 11,\ 12.5,\ 13.1,\ 12.7,\ 12.4\ .$$
>
> *Supponendo che il tasso alcolico nel vino segua una distribuzione normale, verifica o smentisci il timore del proprietario.* L'ipotesi nulla è che il tasso alcolico medio del vino sia 12.5%. Siccome abbiamo supposto una distribuzione normale, possiamo applicare il test T con $\nu = 6 - 1 = 5$ gradi di libertà. La media dei campioni è $M_6 = 12.2$ mentre la deviazione standard campionaria è $s_6 \simeq 0.79$. Quindi il valore del test è:
>
> $$|T_5| = \frac{|12.2 - 12.5|}{0.79}\sqrt{6} \simeq 0.93.$$
>
> (...) il valore di soglia al livello di affidabilità 0.1 (QUINDI PER NOI $\alpha = 0.9$) (...) è 2.015, ben più alto del valore che abbiamo ottenuto. Quindi i dati statistici da te raccolti non smentiscono l'ipotesi nulla e non confermano il timore del proprietario.

Si noti che noi abbiamo agito come un'autorità di controllo: vogliamo dimostrare che non è 12.5 e allora mettiamo come ipotesi alternativa proprio quella. Solo al 90% è stato fatto, poco.
Il valore 2.015 è il quantile $t_{\frac{1+\alpha}{2}}(n-1)$ cioè $t_{\frac{1+0.9}{2}}(6-1)$ cioè $t_{0.95}(5)$.

Esempio. Del libro di Paolo Baldi già citato:

> L'altezza media delle reclute alla visita di leva del 1970 era di 169 cm; 121 reclute vengono scelte a caso nel 1980. I valori di media varianza del campione sono:
>
> $$\bar{X} = 171 \quad \leftarrow \textit{intende } \bar{X}_n \textit{ con } n{:=}121$$
>
> $$S^2 = 85 \quad \leftarrow \textit{intende } S_n^2 \textit{ con } n{:=}121$$
>
> Si può affermare che l'altezza media delle reclute è aumentata?

(Siccome si vuole dimostrare che l'altezza è aumentata, si metta come ipotesi che non è aumentata, e siccome nulla è detto sul livello del test si usi la classica soglia del 5%).

Applichiamo la (**??**):

$$\begin{cases} H : \mu \leq 169 \\ A : \mu > 169 \end{cases} \qquad rifiuta\ H\ se\ \sqrt{121}\,\frac{171-169}{\sqrt{85}} > t_{0.95}(120)$$

$$11\,\frac{2}{9.22} \approx 2.386 > 1.645 = t_{0.95}(+\infty) \approx t_{0.95}(120)$$

Sulla tavola data prima, abbiamo

$t_{0.95}(100) \approx 1.660$

$t_{0.95}(+\infty) \approx 1.645$

ed entrambi possono essere presi come approssimazione di $t_{0.95}(120)$, concludendo analogamente.
Si respinge al livello 0.05 l'ipotesi che l'altezza non sia aumentata.

Esempio. I test di Student sono usati amplissimamente in Medicina e Farmacia, ma non è facile trovare un test che si possa fare in

aula; non possiamo chiedere glicemie o sideremie, sconosciute agli studenti, nè il peso o l'altezza, che potrebbero creare imbarazzo. Tuttavia qualcosa si può tentare. È noto che in Italia il calcio è più seguito dai maschi che dalle femmine, e potremmo cercare di verificarlo anche a livello degli studenti dell'aula.
Per il seguire il calcio, facciamo ragionevolmente così: misureremo l'interesse per il calcio dal numero di anni passati di cui la persona ritiene di conoscere la squadra vincitrice del Campionato di serie A. Per esempio per lo scrivente l'interesse per il calcio, così misurato, è 0: non sa chi ha vinto nemmeno l'ultimo campionato. Ma qualcuno potrebbe sapere anche i risultati degli ultimi 10 o perfino 100 campionati, se è più interessato.
Un'altro possibile indicatore potrebbe essere – e si otterrebbero risultati diversi – il numero di calciatori noti per nome e cognome nella squadra di calcio nazionale italiana. Ma si noti che una ragazza magari è appassionatissima di calcio e conosce tutti i nomi della squadra femminile, ma non di quella maschile. E insomma non se ne viene fuori: ci sono diverse possibili scelte per misurare l'interesse per il calcio, come per l'esposizione al fumo di sigaretta o all'inquinamento o a qualsiasi cosa.
La questione in oggetto potrebbe essere banalmente risolta al livello di Statistica Descrittiva: chiediamo il numero detto a tutti i maschi e facciano la media, e poi similmente con tutte le femmine. Ma invece vogliamo fare Statistica Inferenziale: chiederemo il numero a tutti i maschi presenti (in generale molto pochi), faremo la media e quello sarà il parametro μ_0 di riferimento. Invece per le femmine (in generale molto più numerose) prenderemo un campione, per esempio di $n = 10$ soggetti, trovando lo stimatore $\bar{X}_{10}$ dell'interesse femminile per il calcio nell'aula. Qua c'è un problema perchè il test ha come alternativa $\mu > \mu_0$ e non $<$. Per superare questo problema, applicheremo il test a 1 coda (**??**) applicato ai valori *opposti*, cioè l'interesse, poniamo, 3.2, equivale a un disinteresse -3.2 per il calcio:

— rifiuta l'ipotesi nulla che le femmine si interessino al calcio di più dei maschi (ovvero in qualche modo quasi “accetta" che le femmine si interessino di meno dei maschi al calcio, frase più comprensibile anche se meno esatta)

$$se \ \sqrt{10}\,\frac{-\bar{X}_{10}+\mu_0}{S_{10}} > t_{0.95}(10-1) \approx 1.833$$

(Naturalmente possiamo usare WolframAlpha). Con il passaggio ai valori cambiati di segno, se respingeremo l'ipotesi nulla, avremo verificato statisticamente il *minor* interesse delle femmine per il calcio rispetto ai maschi; quello che supponevamo e volevamo dimostrare statisticamente con un campionamento. Il risultato così ottenuto si riferisce solo all'aula, ovviamente, non a tutta Italia, e già là viene solo da un campione e non da tutta la popolazione dell'aula. E poi potrebbe essere verificato con la Statistica Descrittiva, chiedendo a *tutte* le femmine.

40.10 Complementi – Test di Student per confronto di medie

Un problema classico della Statistica è determinare se 2 popolazioni hanno la stessa media, a partire da campioni di esse.

Per fissare le idee consideriamo il problema:
il tempo di attesa in 2 farmacie è uguale?
Definiamo, in via semplificata 2 variabili aleatorie:
X = tempo di attesa nella farmacia x
Y = tempo di attesa nella farmacia y.
Naturalmente sono variabili aleatorie: non c'è in nessuna delle 2 farmacie un tempo di attesa fisso e costante, con ogni cliente si produce un valore che è una determinazione della variabile aleatoria X o Y rispettivamente.

Un altro esempio ci viene da Wikipedia:

> suppose we are evaluating the effect of a medical treatment, and we enroll 100 subjects into our study, then randomly assign 50 subjects to the treatment group and 50 subjects to the control group. In this case, we have two independent samples and would use the unpaired form of the t-test. https://en.wikipedia.org/wiki/Student%27s_t-test, letto il 10 gennaio 2023.

Un altro esempio in
https://www.jmp.com/it_it/statistics-knowledge-portal/t-test.html, letto il 10 gennaio 2023:

> Le frequenze cardiache medie di due gruppi di persone sono uguali

(Fra maschi/femmine, sportivi/non sportivi, fumatori/non fumatori...) Ciascuna delle 2 variabili aleatorie ha una sua distribuzione, che non conosceremo mai esattamente, con una sua speranza matematica, che ugualmente non conosceremo mai esattamente, ma è proprio il confronto statistico di quelle 2 speranze matematiche che ci interessa. Le chiameremo μ_X e μ_Y.

La Statistica fra poco ci insegnerà come rifiutare o non rifiutare l'ipotesi nulla dell'uguaglianza con ragionevole certezza statistica, purchè disponiamo di un buon numero n di determinazioni di X e un buon numero m di determinazioni di Y. La questione si dirime con uno stimatore da confrontare con un quantile di Student, e ha una formula molto complicata, ma

<u>noi considereremo solo il caso semplice in cui $n = m$</u>

cioè – nell'esempio di prima – misuriamo un ugual numero di tempi di attesa nelle 2 farmacie.
Prima di avere le misurazioni ovvero i valori numerici, abbiamo 2 campioni aleatori

$X_1,\ldots,X_n$
$Y_1,\ldots,Y_n$.

Dopo la sperimentazione, avremo 2 dataset numerici

$x_1,...,x_n$
$y_1,...,y_n$.

Teorema. Siano $X_1,...,X_n$ un campione aleatorio (gaussiano oppure con n sufficientemente grande) di media μ_X e varianza σ_X^2 sconosciute e $Y_1,..,Y_n$ un campione aleatorio (gaussiano oppure con n sufficientemente grande) di media μ_Y e varianza $\sigma_Y^2 = \sigma_X^2 =: \sigma^2$ (uguale a quella del primo campione e sconosciuta), con le X_i e le Y_i indipendenti fra loro.

Si ha questo test (bilatero) al livello $1-\alpha$ (e altri dicono al livello α; in ogni caso α è "grande", tipicamente 0.95):

$$\begin{cases} H : \mu_X = \mu_Y \\ A : \mu_X \neq \mu_Y \end{cases} \qquad rifiuta\ H\ se\ \left| \frac{\bar{X}_n - \bar{Y}_n}{\sqrt{S_X^2 + S_Y^2}} \sqrt{n} \right| > t_{\frac{1+\alpha}{2}}(2\,n-2) \tag{161}$$

(La quantità col valore assoluto è la statistica del test, e il suo argomento si potrà trovare indicato con T).

Nota. Purtroppo come si vede c'è la condizione molto limitativa che X e Y abbiano uguali varianze.

Esempio. Non è facile trovare un esempio che si possa fare in un'aula universitaria. Comunque, supponiamo che qualcuno abbia la strana idea che alla nascita le mamme degli studenti maschi e femmine avessero un'età media diversa (cosa relativamente improbabile in Italia) pur rimanendo simili le varianze.

X età mamma dei maschi, determinazioni $x_1,...,x_n$

Y età mamma delle femmine, determinazioni $y_1,...,y_n$

Con WolframAlpha si calcolerà T, con `Mean`, `Variance`, eccetera. Quello che c'è da aspettarsi, ragionevolmente, è che le medie campionarie verranno diverse ma non così tanto da respingere l'ipotesi nulla che le "vere" medie siano uguali.
Se per esempio $n = 5$, cioè si considerano 5 maschi e 5 femmine, il

quantile da usare al solito livello del 95% è

$$t_{\frac{1+\alpha}{2}}(2 \cdot 5 - 2) = t_{\frac{1+0.95}{2}}(10 - 2) = t_{0.975}(8) \approx 2.306$$