Matematica e Statistica

Prova d'Esame (26/07/2010)

Università di Verona - Laurea in Biotecnologie - A.A. 2009/10

Matematica e Statistica

Prova d'Esame di $\boxed{\text{MATEMATICA}}$ (26/07/2010)

Università di Verona - Laurea in Biotecnologie - A.A. 2009/10

Cognome-Nome ______________________________ Matr. ______________

IN STAMPATELLO

(1) Dati i vettori $\vec{u} = (2, 0, -1)$ e $\vec{v} = (1, 1, 0)$ determinare un vettore $\vec{w}$ ortogonale a entrambi, e calcolare il volume del parallelepipedo compreso tra i tre vettori. Determinare poi, in forma parametrica e cartesiana, il piano Π parallelo a $\vec{w}$ e passante per i punti $P(0, 1, -1)$ e $Q(3, -1, 0)$.

(2) Studiare l'andamento di $f(x) = \log|x| + \dfrac{1}{x-2}$, e tracciarne il grafico.[1]

(3) (a) Calcolare $\displaystyle\int_0^\pi \sqrt{3 + \cos x}\, \sin 2x\, dx$.[2]

(b) Disegnare $S = \{(x, y) : y \le 2e^x\,,\ |y| \le 2 - x\,,\ x \ge -1\}$, e calcolarne l'area.

(4) (a) Data $g(x, y) = \dfrac{\sqrt{2 - xy}}{x - 1}$, determinarne dominio, zeri, segno e limiti interessanti, disegnando i risultati. Trovarne i punti stazionari.

(b) Determinare l'equazione cartesiana del piano Π tangente al grafico di g sopra il punto $(0, 1)$ del dominio. Ci sono altri punti del dominio sopra i quali il piano tangente al grafico di g è parallelo a Π?

(5) Trovare tutte le soluzioni dell'equazione differenziale $y'' + 2y' + 5y = \cos x + 2 \sin x$ e tutte quelle dell'equazione differenziale $y' - y = x$ il cui grafico passa per il punto $(0, -1)$ con pendenza 2.

---

[1] Per lo studio di zeri e segno si consiglia il metodo grafico. Non è richiesto lo studio della convessità.

[2] Si consiglia di porre $3 + \cos x = t^2$, cioè $\cos x = t^2 - 3$.

Matematica e Statistica

Prova d'Esame di $\boxed{\text{STATISTICA}}$ (26/07/2010)

Università di Verona - Laurea in Biotecnologie - A.A. 2009/10

Cognome-Nome ______________________________ Matr. ______________
IN STAMPATELLO

Per ogni calcolo effettuato scrivere anche la formula teorica da utilizzare. ▸▸▸▸ **Tabella sul retro** ▸▸▸▸

ESERCIZIO 1

Una nuova tecnica per il sequenziamento del DNA umano promette di ridurre notevolmente i tempi di elaborazione, a scapito tuttavia di una minore precisione.
In particolare, la colonna **f** della tabella seguente riporta le frequenze osservate delle 4 basi azotate con la nuova metodologia su un campione di 1000 basi, mentre la colonna **F** rappresenta quella che dovrebbe essere l'effettiva distribuzione teorica.

Base	f	F
Adenina	200	200
Citosina	315	300
Guanina	265	300
Timina	220	200

Sulla base dei risultati, si può ritenere efficace la nuova tecnica ossia considerare valida l'ipotesi di omogeneità fra le 2 distribuzioni di frequenze ad un livello di significatività del 5%?

ESERCIZIO 2

x	f
4	0
6	40
8	130
14	80

Sulla distribuzione di frequenze presentata in tabella calcolare:
(a) la media aritmetica, la media armonica e la media geometrica;
(b) la mediana e la moda;
(c) la varianza usando l'origine A=2.

ESERCIZIO 3

Una ricerca sulla relazione fra quantità assunta di un integratore a base di beta carotene e il rischio di subire un eritema solare ha dato i seguenti risultati:

Quantità beta carotene	Rischio eritema
0	50
5	35
10	15
15	0

Sui dati presentati in tabella:
(a) interpolare le due distribuzioni con una retta;
(b) calcolare il coefficiente di correlazione lineare, commentandolo brevemente;
(c) valutare la significatività del coefficiente di correlazione lineare (alpha=5%);
(d) giudicare la bontà di accostamento.

Allegato: Tabella "Chi Quadrato"

Valori della variabile "Chi Quadrato" che sottendono una coda destra di ammontare alpha

G.d.l.	alpha %							
	99,5	99	97,5	95	5	2,5	1	0,5
1	0,00	0,00	0,00	0,00	3,84	5,02	6,64	7,88
2	0,01	0,02	0,05	0,10	5,99	7,38	9,21	10,60
3	0,07	0,12	0,22	0,35	7,82	9,35	11,35	12,84
4	0,21	0,30	0,48	0,71	9,49	11,14	13,28	14,86
5	0,41	0,55	0,83	1,15	11,07	12,83	15,09	16,75
6	0,68	0,87	1,24	1,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	15,51	17,54	20,09	21,96
9	1,74	2,09	2,70	3,33	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,58	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	22,36	24,74	27,69	29,82
14	4,08	4,66	5,63	6,57	23,69	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	26,30	28,85	32,00	34,27
17	5,70	6,41	7,56	8,67	27,59	30,19	33,41	35,72
18	6,27	7,02	8,23	9,39	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	30,14	32,85	36,19	38,58
20	7,43	8,26	9,59	10,85	31,41	34,17	37,57	40,00

Allegato: Tabella "*t* di Student"

Valori della "*t* di Student" che sottendono una coda destra di ammontare alpha

G.d.l.	α				
ν	*5%*	*2,5%*	*1%*	*0,5%*	*0,1%*
1	6,314	12,706	31,821	63,657	318,310
2	2,920	4,303	6,965	9,925	22,327
3	2,353	3,182	4,541	5,841	10,215
4	2,132	2,776	3,747	4,604	7,173
5	2,015	2,571	3,365	4,032	5,893
6	1,943	2,447	3,143	3,707	5,208
7	1,895	2,365	2,998	3,499	4,785
8	1,860	2,306	2,896	3,355	4,501
9	1,833	2,262	2,821	3,250	4,297
10	1,812	2,228	2,764	3,169	4,144
11	1,796	2,201	2,718	3,106	4,025
12	1,782	2,179	2,681	3,055	3,930
13	1,771	2,160	2,650	3,012	3,852
14	1,761	2,145	2,624	2,997	3,787
15	1,753	2,131	2,602	2,947	3,733
16	1,746	2,120	2,583	2,921	3,686
17	1,740	2,110	2,567	2,898	3,646
18	1,734	2,101	2,552	2,878	3,610
19	1,729	2,093	2,539	2,861	3,579
20	1,725	2,086	2,528	2,845	3,552

Soluzioni

MATEMATICA

(1) Un vettore ortogonale ai due dati vettori $\vec{u} = (2, 0, -1)$ e $\vec{v} = (1, 1, 0)$ è ad esempio il loro prodotto vettoriale $\vec{w} = \vec{u} \wedge \vec{v} = \det \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ 2 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} = (1, -1, 2)$, e il volume del parallelepipedo compreso tra i tre vettori è dato dal modulo del prodotto misto $|\vec{w} \cdot (\vec{u} \wedge \vec{v})|$, che in questo caso vale $|\vec{w} \cdot \vec{w}| = |\vec{w}|^2 = 6$. Infine, il piano Π è parallelo, oltre che a $\vec{w}$, anche al vettore $(3, -1, 0) - (0, 1, -1) = (3, -2, 1)$, dunque una forma parametrica è $\Pi = \{(0, 1, -1) + s(3, -2, 1) + t(1, -1, 2) : s, t \in \mathbb{R}\} = \{(3s + t, 1 - 2s - t, -1 + s + 2t) : s, t \in \mathbb{R}\}$; poi, eliminando i parametri si ottiene la forma cartesiana $3x + 5y + z - 4 = 0$.

(2) (Figura 1) La funzione $f(x) = \log|x| + \frac{1}{x-2}$ è definita per $x \neq 0$ e $x \neq 2$, ed è derivabile infinite volte nel suo dominio. Non ha parità ne' periodicità; i limiti notevoli sono tutti determinati, e valgono $\lim_{x\to\mp\infty} f(x) = +\infty$, $\lim_{x\to 0^\mp} f(x) = -\infty$ e $\lim_{x\to 2^\mp} f(x) = \mp\infty$. Si ha poi $f(x) \geq 0$ se e solo se $\log|x| \geq -\frac{1}{x-2}$, e un confronto grafico tra le funzioni $\log|x|$ (il logaritmo simmetrizzato) e $-\frac{1}{x-2}$ (un'omografica con asintoti $y = 0$ e $x = 2$) mostra chiaramente che esiste un unico punto $\alpha \in]-2, -1[$ in cui f si annulla, e tale che $f(x) > 0$ quando $x < \alpha$ oppure $x > 2$. Le rette $x = 0$ e $x = 2$ sono asintoti verticali bilateri, ed essendo $\lim_{x\to\mp\infty} \frac{f(x)}{x} = 0$ non vi sono asintoti obliqui (in effetti, agli infiniti f tende a $+\infty$ con andamento logaritmico). La derivata $f'(x) = \frac{1}{x} - \frac{1}{(x-2)^2} = \frac{x^2-5x+4}{x(x-2)^2}$ si annulla in $x = 1$ e $x = 4$ e è positiva per $0 < x < 1$ e $x > 4$: pertanto $x = 1$ è un punto di massimo locale (con $f(1) = -1$) mentre $x = 4$ è di minimo locale (con $f(4) = 2\log 2 + \frac{1}{2} \simeq 1{,}9$). Infine, anche se non è richiesto, derivando ancora si otterrebbe $f''(x) = -\frac{x^3-5x^2+4x-8}{x^2(x-2)^3}$; la cubica $p(x) = x^3 - 5x^2 + 4x - 8$ ha massimo locale in $x = \frac{5-\sqrt{13}}{3} \sim 0{,}5$ e minimo locale in $x = \frac{5+\sqrt{13}}{3} \sim 2{,}8$ (basta guardare la derivata), entrambi con valore negativo; e poiché $p(4) = -8 < 0$ e $p(5) = 12 > 0$ essa avrà un solo zero in un certo $\beta \in]4, 5[$. Dunque vale $f''(x) > 0$ per $2 < x < \beta$, e f ha un solo flesso in $x = \beta$.

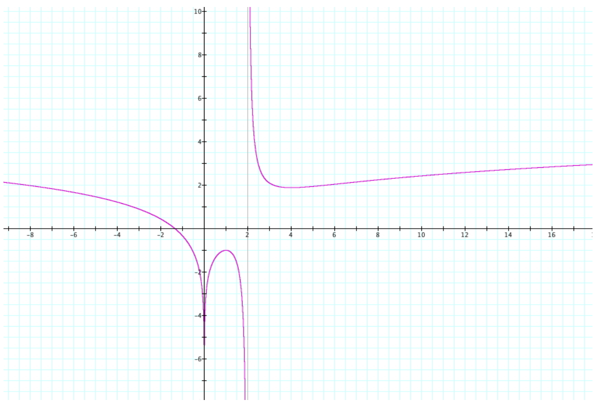
(3) (a) Posto $3 + \cos x = t^2$ si ottiene $\sin x\, dx = -2t\, dt$: perciò, ricordando che $\sin 2x = 2\sin x \cos x$, l'integrale diventa $\int_0^\pi \sqrt{3 + \cos x}\, \sin 2x\, dx = \int_2^{\sqrt{2}} \sqrt{t^2}\, (-4t(t^2 - 3))\, dt = 4\int_{\sqrt{2}}^2 t^2(t^2 - 3)\, dt = 4(\frac{1}{5}t^5 - t^3]_{\sqrt{2}}^2 = \frac{8}{5}(3\sqrt{2} - 4)$.

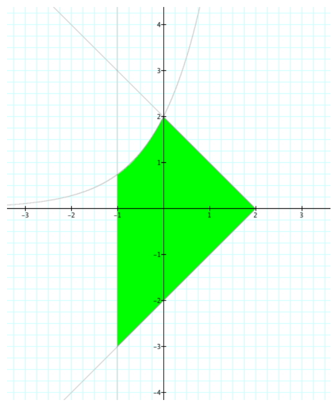
(b) (Figura 2) L'insieme $S = \{(x, y) : y \leq 2e^x,\ |y| \leq 2 - x,\ x \geq -1\}$ è rappresentato in figura; l'area risulta pertanto $\int_{-1}^0 2e^x\, dx + \int_0^2 (2 - x)\, dx + \int_2^{-1} (x - 2)\, dx = (2e^x]_{-1}^0 + (2x - \frac{1}{2}x^2]_0^2 + (\frac{1}{2}x^2 - 2x]_2^{-1} = (2) - (\frac{2}{e}) + (2) - (0) + (\frac{5}{2}) - (-2) = \frac{17}{2} - \frac{2}{e} \sim 7{,}8$.

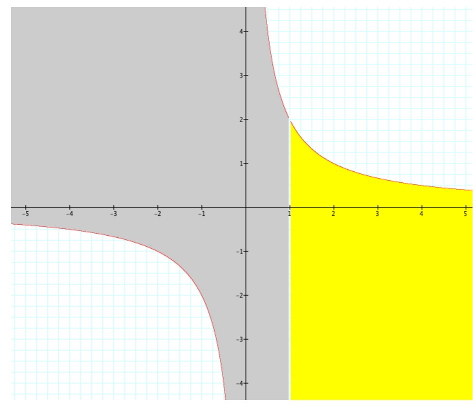
(4) (a) (Figura 3) Il dominio di $g(x, y) = \frac{\sqrt{2-xy}}{x-1}$ è dato da $xy \leq 2$ (si tratta della parte di piano compresa tra i due rami, inclusi, dell'iperbole equilatera $xy = 2$) meno i punti della retta $x = 1$, parallela all'asse y. La funzione si annulla sui punti dei due rami dell'iperbole (tranne ovviamente il punto $A(1, 2)$, che è fuori dal dominio), è positiva per $x > 1$, ed è continua in tutti i punti del suo dominio. Gli unici limiti interessanti sono in A e in ∞_2, e nessuno dei due limiti esiste: in entrambi i casi basta notare che sui rami dell'iperbole la funzione è nulla, mentre avvicinandosi alla retta $x = 1$ la funzione tende a ∞. Le derivate parziali sono $\frac{\partial g}{\partial x} = -\frac{\frac{-y}{2\sqrt{2-xy}}(x-1) - \sqrt{2-xy}}{(x-1)^2} = -\frac{xy+y-4}{2(x-1)^2\sqrt{2-xy}}$ e $\frac{\partial g}{\partial y} = \frac{1}{x-1}\frac{-x}{2\sqrt{2-xy}} = -\frac{x}{2(x-1)\sqrt{2-xy}}$, e il sistema $\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y} = 0$ dà il solo punto stazionario $P(0, 2\sqrt{2})$.

(b) Come noto, l'equazione cartesiana del piano tangente al grafico di g sopra il punto $(0, 1)$ del dominio è data da $z = g(0, 1) + \frac{\partial g}{\partial x}(0, 1) \cdot (x - 0) + \frac{\partial g}{\partial y}(0, 1) \cdot (y - 1)$: essendo $g(0, 1) = -\sqrt{2}$, $\frac{\partial g}{\partial x}(0, 1) = -\frac{3\sqrt{2}}{4}$ e $\frac{\partial g}{\partial y}(0, 1) = 0$ si ottiene $z = -\sqrt{2} - \frac{3\sqrt{2}}{4}x$, ovvero $3\sqrt{2}\,x + 4z + 4\sqrt{2} = 0$. Eventuali altri punti (x_0, y_0) del dominio sopra i quali il piano tangente al grafico di g è parallelo a quello trovato dovranno soddisfare il sistema $(\frac{\partial g}{\partial x}(x_0, y_0), \frac{\partial g}{\partial y}(x_0, y_0)) = (-\frac{3\sqrt{2}}{4}, 0)$, ma è facile vedere che l'unica soluzione di tale sistema è per l'appunto $(0, 1)$. Dunque non vi sono altri punti con questa proprietà.

(5) L'equazione differenziale $y'' + 2y' + 5y = \cos x + 2\sin x$ è lineare del secondo ordine a coefficienti costanti. L'equazione caratteristica $t^2 + 2t + 5 = 0$ ha soluzioni complesse coniugate $-1 \mp 2i$, dunque le soluzioni dell'equazione omogenea associata sono del tipo $Ae^{-x}\cos 2x + Be^{-x}\sin 2x$ con $A, B \in \mathbb{R}$. Una soluzione particolare per $b(x) = \cos x + 2\sin x$ sarà del tipo $\tilde{y}(x) = a\cos x + b\sin x$, e imponendo che $\tilde{y}'' + 2\tilde{y}' + 5\tilde{y} = \cos x + 2\sin x$ si ottiene $(a, b) = (0, \frac{1}{2})$. Pertanto tutte le soluzioni della prima equazione differenziale sono $y(x) = Ae^{-x}\cos 2x + Be^{-x}\sin 2x + \frac{1}{2}\sin x$, con $A, B \in \mathbb{R}$. • L'equazione differenziale $y' - y = x$ è lineare del primo ordine, e può essere vista nella forma $y' + p(x)\,y = q(x)$ con $p(x) = -1$ e $q(x) = x$. Poiché $P(x) = \int p(x)\,dx = -x$ e $\int e^{P(x)} q(x)\,dx = \int x e^{-x}\,dx = -(x+1)e^{-x}$, la soluzione generale si può scrivere come $y(x) = e^{x}(-(x+1)e^{-x} + k) = ke^{x} - x - 1$ al variare di $k \in \mathbb{R}$. • Chiedere che il grafico di $y(x)$ passi per il punto $(0, -1)$ con pendenza 2 equivale a chiedere che $y(0) = -1$ e $y'(0) = 2$. Per la prima equazione differenziale ciò è niente altro che un dato di Cauchy, che darà luogo a una e una sola soluzione: in effetti, imponendo tali condizioni si ottiene $(A, B) = (-1, \frac{1}{4})$, ovvero $y(x) = e^{-x}(\frac{1}{4}\sin 2x - \cos 2x) + \frac{1}{2}\sin x$. Quanto alla seconda, la condizione $y(0) = -1$ dà già l'unica possibilità $k = 0$, ovvero $y(x) = -x - 1$: ma questa soluzione soddisfa $y'(0) = -1 \neq 2$, dunque non vi sono soluzioni con le caratteristiche cercate.







1. Il grafico della funzione dell'ex. 2. **2.** L'insieme dell'ex. (3.b). **3.** Ex. 4: zeri (rosso), segno positivo (giallo) e negativo (grigio) della funzione g.

ESERCIZIO 1

Una nuova tecnica per il sequenziamento del DNA umano promette di ridurre notevolmente i tempi di elaborazione, a scapito tuttavia di una minore precisione.

In particolare, la colonna *f* della tabella seguente riporta le frequenze osservate delle 4 basi azotate con la nuova metodologia su un campione di 1000 basi, mentre la colonna *F* rappresenta quella che dovrebbe essere l'effettiva distribuzione teorica.

Sulla base dei risultati, si può ritenere efficace la nuova tecnica ossia considerare valida l'ipotesi di omogeneità fra le 2 distribuzioni di frequenze ad un livello di significatività del 5%?

base	**f**	**F**	**f(-F)^2/F**
Adenina	200	200	0,0000
Citosina	315	300	0,7500
Guanina	265	300	4,0833
Timina	220	200	2,0000
	1000	**1000**	**6,8333**

alpha=5%

Calcolo del Chi Quadrato:

ChiQc = **6,8333**

Si individua sulle tavole del Chi Quadrato il valore teorico da confrontare:

ni=n-1=4-1= **3** gdl alpha = 5%

ChiQt = **7,82**

Poiché ChiQc<=ChiQt si accetta ipotesi di omogeneità fra le due distribuzioni

ESERCIZIO 2

Sulla distribuzione di frequenze presentata in tabella, calcolare:
a) la media aritmetica, la media armonica e la media geometrica;
b) la mediana e la moda;
c) la varianza usando l'origine A=2.

x	f	x*f	f/x	ln(x)	ln(x)*f	x-2	$(x-2)^2$	$(x-2)^2*f$
4	50	200	12,50	1,3863	69,3147	2	4	200
6	40	240	6,67	1,7918	71,6704	4	16	640
8	130	1040	16,2500	2,0794	270,327	6	36	4680
14	80	1120	5,7143	2,6391	211,1246	12	144	11520
	300	**2600**	**41,1310**	**7,8966**	**622,4371**			**17040**

a) Calcolo della media aritmetica, armonica e geometrica:

$$\mathbf{M(X)} = \frac{\Sigma\, x * f}{\Sigma\, f} = \frac{2600}{300} = \mathbf{8,66667}$$

$$\mathbf{Ma(X)} = \frac{\Sigma\, f}{\Sigma\, f/x} = \frac{300}{41,1310} = \mathbf{7,294}$$

$$\ln(Mg(X)) = \frac{\Sigma\, \ln(x) * f}{\Sigma\, f} = \frac{622,4371}{300} = \mathbf{2,0748} \qquad \mathbf{Mg(X)} = e^{2,0748} = \mathbf{7,963}$$

b) Calcolo della mediana e della moda:

x150 =< mediana =< x151 : **me = 8**

moda = 8

c) Calcolo della varianza usando l'origine A=2:

$$V_A(X) = \frac{17040}{300} = \mathbf{56,8}$$

$$V(X) = V_A(X) - (M(X)-2)^2 = \mathbf{12,3556}$$

ESERCIZIO 3

Una ricerca sulla relazione fra quantità assunta di un integratore a base di beta carotene e il rischio di subire un eritema solare ha dato i seguenti risultati:

X	Y	X * Y	X^2	Y^2
0	50	0	0	2500
5	35	175	25	1225
10	15	150	100	225
15	0	0	225	0
30	**100**	**325**	**350**	**3950**

Sui dati presentati in tabella:
a) interpolare le due distribuzioni con una retta;
b) calcolare il coefficiente di correlazione lineare, commentandolo brevemente;
c) valutare la significatività del coefficiente di correlazione lineare (alpha=5%);
d) giudicare la bontà di accostamento.

a) Calcolo dei parametri della retta interpolante Y'=a+bX :

Calcolo attraverso le formule dirette (ma si poteva anche sviluppare il sistema):

$$b = \frac{Cov(X;Y)}{V(X)} \qquad a = M(Y) - bM(X)$$

$$M(X) = \frac{30}{4} = \mathbf{7,5}$$

$$M(Y) = \frac{100}{4} = \mathbf{25}$$

$$Cov(X;Y) = M(X*Y) - M(X)*M(Y)= \frac{325}{4} - 7,5 * 25 = \mathbf{-106,25}$$

$$V(X) = M(X^2) - M(X)^2 = \frac{350}{4} - 7,5\text{^}2 = \mathbf{31,25}$$

$$\mathbf{b =} \frac{Cov(X;Y)}{V(X)} = \frac{-106,25}{31,25} = \mathbf{-3,4}$$

$$\mathbf{a =} \quad M(Y) - bM(X) = \quad 25 - 3,1036*7,5= \quad \mathbf{50,5}$$

b) Calcolo del coefficiente di correlazione lineare e suo breve commento:

$$r = \frac{Cov(X;Y)}{\sigma(X)\ \sigma(Y)}$$

$$V(Y) = \frac{3950}{4} - 25\text{^}2 = \mathbf{362{,}5}$$

$\sigma(Y) =$ RADQ(362,5) = **19,0394**

$\sigma(X) =$ RADQ(31,25) = **5,5902**

$$\mathbf{r} = \frac{-106{,}25}{3{,}9607 * 12{,}4775} = \mathbf{-0{,}9983}$$

Si registra una forte relazione lineare indiretta

c) Giudicare la significatività del coefficiente di correlazione lineare:

H_0: $\rho = 0$ H_1: $\rho \neq 0$

Calcolo del *t*:

$$t_c = r\sqrt{\frac{1}{1-r^2}(n-2)} = \mathbf{-24{,}042}$$

Individuo dalle tavole della *t* di Student i 2 estremi dell'intervallo del *t teorico*:

ni = n - 2 = 2 gdl alpha = 5%

$$t_t = \pm t_{ni;\frac{\alpha}{2}} = \mathbf{\pm 4{,}303}$$

Poiché il *t* calcolato è esterno all'intervallo compreso fra i due valori del *t* teorico, si rifiuta l'ipotesi H_0 e si considera **significativo** il coefficiente di correlazione calcolato.

d) Giudicare la bontà di accostamento del modello teorico:

Calcolo il coefficiente di determinazione r^2

$r^2 = 0{,}9852^2 =$ **0,9966** **Il modello teorico riesce a spiegare quasi completamente le variazioni di Y**