MATEMATICA E STATISTICA (PARI) – CdL in BIOLOGIA

SIMULAZIONE 2

1) Studio di funzione: Ottimizzazione di un antibiotico (17 punti)

Si vuole ottimizzare un antibiotico a base di sulfametoxazolo e trimetoprim, simile al Bactrimel, usato per trattare e prevenire diverse infezioni batteriche. Sia data la seguente funzione, che modellizza la concentrazione del farmaco nel sangue in funzione del tempo:

$$C(t) = kt^2e^{-t}$$

Dove *k* è una costante positiva e non nulla interessato ed il tempo è dato in ore.

a) Ottimizzare la funzione, calcolando il valore di *k* necessario per avere una concentrazione massima C_{MAX} di antibiotico nel sangue pari a 40.6 µg/mL ed indicando dopo quanto tempo dall'assunzione, la concentrazione raggiunge tale valore. (4 punti)
b) Utilizzando il valore di *k* calcolato al punto precedente, studiare la funzione generica *C(x)*. (8 punti)
c) Tracciare il grafico dettagliato e per punti della funzione *C(t)*. (2 punti)
d) Dopo quanto tempo dall'assunzione, l'antibiotico si può dire smaltito dall'organismo? Assumere un valore soglia C_0 pari a 0.1 µg/mL. (2 punti)
e) Assumendo di prescrivere questo antibiotico ad un paziente, facendo riferimento al valore ottenuto al punto precedente, per quante volte al giorno ed ogni quante ore consigliereste di assumere il farmaco? (1 punto)

2) Proliferazione di una colonia virale (6 punti)

Viene condotto uno studio sulla proliferazione di un virus all'interno di un campione inizialmente costituito da 68 microorganismi. Si ottiene che l'andamento della velocità di proliferazione può essere descritta dalla seguente funzione:

$$v_P(t) = \frac{10t + 24}{t^2 + 5t + 6}$$

Dove il tempo t è espresso in minuti e la velocità di proliferazione v_P in numero di virus al minuto (min^{-1}).

Calcolare il numero di virus presenti nel campione dopo un'ora.

3) Statistica: Puzzle (7 punti)

Un'azienda produttrice di puzzle vuole effettuare un'indagine statistica sul legame tra il numero di vendite e la tipologia di puzzle acquistati in termini di numero di pezzi.

La statistica prende come campione di 8 000 puzzle venduti e suddivisi come segue:

Tipologia	**A**	**B**	**C**	**D**	**E**
Numero di pezzi	1000	1500	2000	5000	10000
Puzzle venduti	3020	2481	1794	514	191

a) Tenendo conto sia del peso diverso che ciascuna tipologia ha all'interno della popolazione dei puzzle venduti che della dimensione della popolazione considerata, verificare il tipo di correlazione tra il numero di vendite e la tipologia di prodotto in termini di numero di pezzi del puzzle. (3 punti)
b) Visualizzare il legame tra il numero di vendite e la relativa tipologia, sovrapponendolo ad un modello di regressione lineare. (2 punti)

Utilizzando un test adeguato, verificare l'ipotesi che il <u>numero di vendite per pezzo</u> sia mediamente pari a 1.83. (2 punti)

Valori di riferimento per i test di ipotesi

Test T				
α / v	0.10	0.05	0.01	0.001
1	6.314	12.706	63.657	636.619
2	2.920	4.303	9.925	31.599
3	2.353	3.182	5.841	12.924
4	2.132	2.776	4.604	8.610
5	2.015	2.571	4.032	6.869
6	1.943	2.447	3.707	5.959
7	1.895	2.365	3.499	5.408
8	1.860	2.306	3.355	5.041
9	1.833	2.262	3.250	4.781
10	1.812	2.228	3.169	4.587
11	1.796	2.201	3.106	4.437
12	1.782	2.179	3.055	4.318
13	1.771	2.160	3.012	4.221
14	1.761	2.145	2.977	4.140
15	1.753	2.131	2.947	4.073
16	1.746	2.120	2.921	4.015
17	1.740	2.110	2.898	3.965
18	1.734	2.101	2.878	3.922
19	1.729	2.093	2.861	3.883
20	1.725	2.086	2.845	3.850
21	1.721	2.080	2.831	3.819
22	1.717	2.074	2.819	3.792
23	1.714	2.069	2.807	3.768
24	1.711	2.064	2.797	3.745
25	1.708	2.060	2.787	3.725
26	1.706	2.056	2.779	3.707
27	1.703	2.052	2.771	3.690
28	1.701	2.048	2.763	3.674
29	1.699	2.045	2.756	3.659
30	1.697	2.042	2.750	3.646
39	1.685	2.023	2.708	3.558
49	1.677	2.010	2.680	3.500
59	1.671	2.001	2.662	3.463
69	1.667	1.995	2.649	3.437
79	1.664	1.990	2.640	3.418
89	1.662	1.987	2.632	3.403
99	1.660	1.984	2.626	3.392

Test Z				
α	0.10	0.05	0.01	0.001
	1.645	1.960	2.576	3.291

Test χ^2				
α / v	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.828
2	4.605	5.991	9.210	13.816
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.467
5	9.236	11.070	15.086	20.515
6	10.645	12.592	16.812	22.458
7	12.017	14.067	18.475	24.322
8	13.362	15.507	20.090	26.124
9	14.684	16.919	21.666	27.877
10	15.987	18.307	23.209	29.588
11	17.275	19.675	24.725	31.264
12	18.549	21.026	26.217	32.909
13	19.812	22.362	27.688	34.528
14	21.064	23.685	29.141	36.123
15	22.307	24.996	30.578	37.697
16	23.542	26.296	32.000	39.252
17	24.769	27.587	33.409	40.790
18	25.989	28.869	34.805	42.312
19	27.204	30.144	36.191	43.820
20	28.412	31.410	37.566	45.315
21	29.615	32.671	38.932	46.797
22	30.813	33.924	40.289	48.268
23	32.007	35.172	41.638	49.728
24	33.196	36.415	42.980	51.179
25	34.382	37.652	44.314	52.620
26	35.563	38.885	45.642	54.052
27	36.741	40.113	46.963	55.476
28	37.916	41.337	48.278	56.892
29	39.087	42.557	49.588	58.301
30	40.256	43.773	50.892	59.703

$$Z^* = \frac{|x - \mu|}{\sigma}\sqrt{n} \qquad T^*_{n-1} = \frac{|x - \mu|}{s}\sqrt{n}$$

$$\chi^2_{(n-1)(m-1)} = \sum_{i=1}^{k} \frac{\left(f_i^{expected} - f_i^{observed}\right)^2}{f_i^{expected}}$$

Retta di regressione lineare (generica): $m = \frac{\sigma_{XY}}{{\sigma_X}^2}$ $\quad q = \bar{y} - m\bar{x}$

SOLUZIONE ESERCIZIO 1

a) Per trovare il punto di concentrazione massima, bisogna calcolare la derivata prima ed individuare il punto in cui essa si annulla:

$$C'(t) = k\,(2t\,e^{-t} - t^2\,e^{-t}) = k\,(2-t)\,t\,e^{-t}$$

$$C'(t) = 0 \rightarrow \begin{cases} t = 0 \\ \boldsymbol{t = 2} \end{cases}$$

Escludendo ovviamente la soluzione *t = 0* (che comunque rappresenta indica comunque la presenza di un estremo relativo nell'origine degli assi e del quale si dovrà tenere conto in sede di studio di funzione), la concentrazione massima viene raggiunta dopo 2 ore dall'assunzione del farmaco.

Imponendo le condizioni $C_{MAX} = 40.6\ \mu g/mL$ e $t_{CMAX} = 2$ *ore*, si può scrivere:

$$C(t_{CMAX}) = k\,{t_{CMAX}}^2 e^{-t_{CMAX}} \;\rightarrow\; k = \frac{C(t_{CMAX})}{{t_{CMAX}}^2 e^{-t_{CMAX}}} \;\rightarrow\; k = \frac{40.6}{4e^{-2}} = 75$$

b) La funzione ammette un minimo per x = 0 ed un massimo per x = 2.

La derivata seconda risulta:

$$C''(x) = k(x^2 - 4x + 2)e^{-x} = 75(x^2 - 4x + 2)e^{-x}$$

Che si annulla in corrispondenza delle radici del polinomio $x^2 - 4x + 2$:

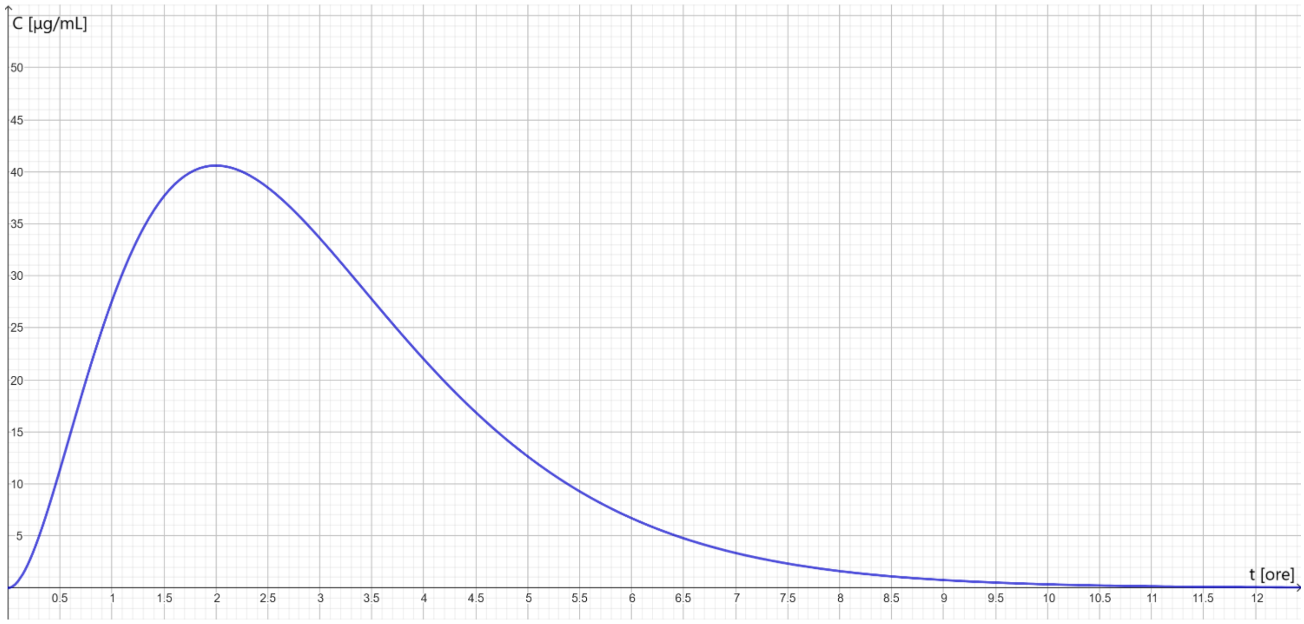
$$x^2 - 4x + 2 = 0 \;\rightarrow x = 2 \pm \sqrt{2}$$

Vi sono quindi due punti di flesso: $F_1\ (2 - \sqrt{2};\ 14.33)$ e $F_2\ (2 + \sqrt{2};\ 28.77)$

Il grafico di *C(x)* è il seguente:

c)

t [ore]	C [µg/mL]
0	0
0.25	3.65
0.5	11.37
0.75	19.93
1	27.59
1.5	37.65
2	**40.60**
2.5	38.48
3	33.61
3.5	27.74
4	21.98
4.5	16.87
5	12.63
5.5	9.27
6	6.69
6.5	4.76
7	3.35
7.5	2.33
8	1.61
8.5	1.10
9	0.75
9.5	0.51
10	0.34
10.5	0.23
11	0.15
11.5	**0.10**
12	0.07



d) Dalla tabella di cui al punto precedente, si evince come il valore di soglia venga raggiunto dopo 11h 30min dall’assunzione.

e) L’antibiotico può essere assunto per due volte al giorno, cioè ogni 12 ore.

SOLUZIONE ESERCIZIO 2

N_0 = 32 batteri; t_x = 1 hr = 60 min

La funzione che dà il numero di batteri in funzione del tempo è data dall'integrale indefinito di $v_P(t)$:

$$N(t) = \int v_P(t)\, dt$$

Per calcolare l'integrale indefinito della funzione, si deve scomporre la funzione fratta mediante il metodo dei fratti semplici:

$$\int v_P(t)\, dt = \int \frac{10t + 24}{t^2 + 5t + 6} dt = 2 \int \frac{5t + 12}{(t+2)(t+3)} dt$$

$$\frac{5t + 12}{(t+2)(t+3)} = \frac{A}{t+2} + \frac{B}{t+3} = \frac{(A+B)t + (3A + 2B)}{(t+2)(t+3)}$$

$$\begin{cases} A + B = 5 \\ 3A + 2B = 12 \end{cases} \rightarrow \begin{cases} A = 2 \\ B = 3 \end{cases}$$

Percio:

$$\frac{5t + 12}{(t+2)(t+3)} = \frac{2}{t+2} + \frac{3}{t+3}$$

Possiamo così svolgere l'integrale:

$$\int v_P(t)\, dt = 2 \int \left(\frac{2}{t+2} + \frac{3}{t+3} \right) dt = 2[2\ln(t+2) + 3\ln(t+3)] + c =$$

$$= 4\ln(t+2) + 6\ln(t+3) + c$$

Per calcolare il numero di virus dopo un'ora (60 minuti), si deve svolgere l'integrale definito tenendo conto della popolazione iniziale:

$$N_x = N_0 + \int_0^{t_x} v_P(t)\, dt$$

$$N_x = N_0 + 4\ln(t_x + 2) + 6\ln(t_x + 3) - 4\ln 2 - 6\ln 3 = 100\ virus$$

SOLUZIONE ESERCIZIO 3

Tipologia	A	B	C	D	E
Numero di pezzi	1000	1500	2000	5000	10000
Puzzle venduti	3020	2481	1794	514	191

a)

Avendo un numero di unità (puzzle venduti) pari a 8 000, NON si deve far riferimento agli indicatori di tipo campionario, bensì a quelli classici. Inoltre, come suggerito nel testo, ad ogni tipologia corrisponde un diverso numero di puzzle, perciò la media per la tipologia di puzzle dovrà essere calcolata come media ponderata piuttosto che aritmetica.

	Media	Varianza	Dev Std
Tipologia (pezzi)	1 851	15 437 633	3 929
Puzzle venduti	1 600	1 198 975	1 095

Si possono ora ottenere covarianza e coefficiente di correlazione (con N = 5):

$$\sigma_{xy} = \frac{(x_i - \bar{x})(y_i - \bar{y})}{N} = -3\,378\,100$$

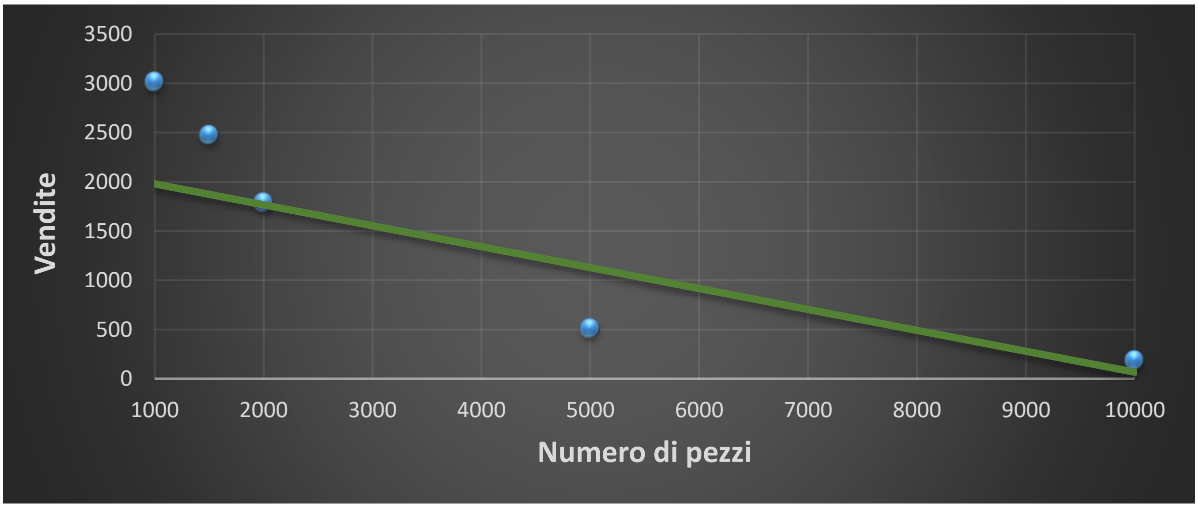
$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = -0.762$$

La correlazione è quindi di natura forte e le grandezze sono inversamente correlate.

b)

La retta di regressione risulta essere:

$$\boldsymbol{y = -0.21\,x + 2\,191}$$



c)

Il test fa riferimento al numero di vendite per pezzo di puzzle, quindi si dovrà ottenere la relativa serie di dati mediante il rapporto tra il numero di puzzle venduti ed il numero di pezzi del puzzle:

Tipologia	A	B	C	D	E
Numero di pezzi	1000	1500	2000	5000	10000
Puzzle venduti	3020	2481	1794	514	191
Vendite per pezzo	3.02	1.654	0.897	0.1028	0.0191

Il test da effettuare dipende sostanzialmente dal numero di campioni considerati che, in questo caso è pari a $n = 8000$. Per questo motivo, dovrà essere effettuato un test di tipo *Z* e dovrà essere calcolata la seguente quantità pivotale:

$$\boxed{Z^* = \frac{X - \mu}{\sigma}\sqrt{n}}$$

E' necessario quindi calcolare media e deviazione standard della serie di dati ottenuta:

Media (ponderata)	**1.86**
Varianza	1.76
Dev Std	1.327
Pivot T*	**2.104**

Test Z				
α	0.10	0.05	0.01	0.001
	1.645	1.960	2.576	3.291

Confrontando il valore ottenuto con quelli tabulati, si ottiene che l'ipotesi nulla H_0 secondo la quale il numero di vendite per pezzo di puzzle sia pari a **1.83** è <u>negabile al 95% ma non al 99%</u>.