

Metodi Matematici e Informatici per la Biologia

a.a. 2024-2025

(Agliari - Presilla - Simonella - Vittorio)

18 settembre 2025

Prova di Verifica III

Nome:

Cognome:

Matricola:

Esercizio 1 (2+3+2= 7 punti)

Sono state misurate le masse (mg) di 24 foglie della stessa pianta:

8.12, 10.47, 10.17, 12.26, 20.08, 12.02, 11.00, 11.53, 11.27, 12.52, 12.92, 3.55, 12.77, 12.56, 15.53, 12.62, 9.84, 13.45, 14.03, 8.39, 9.28, 11.07, 15.32, 11.75.

1.1 Calcolare minimo, massimo, range, media e deviazione standard campionarie, scrivendone le definizioni.

$x_{\min} = 3.55$, $x_{\max} = 20.08$, $\text{range} = 16.53$, $\bar{x} = 11.77$, $s = 3.07$ in mg.

1.2 Calcolare i quartili del campione, la distanza interquartile e mettere in evidenza eventuali dati outlier.

Per il campione ordinato $x_1 \leq \dots \leq x_{24}$

3.55, 8.12, 8.39, 9.28, 9.84, 10.17, 10.47, 11.00, 11.07, 11.27, 11.53, 11.75, 12.02, 12.26, 12.52, 12.56, 12.62, 12.77, 12.92, 13.45, 14.03, 15.32, 15.53, 20.08

i quartili sono $Q_1 = (x_6 + x_7)/2$, $Q_2 = (x_{12} + x_{13})/2$ e $Q_3 = (x_{18} + x_{19})/2$

$Q_1 = 10.32$ mg, $Q_2 = 11.86$ mg, $Q_3 = 12.85$ mg,

la distanza interquartile

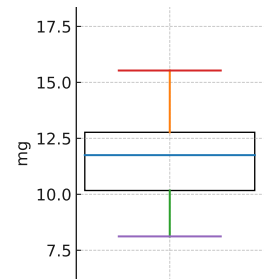
$Q_3 - Q_1 = 2.53$ mg;

l'intervallo dei non-outlier è

$[Q_1 - 1.5 * 2.53, Q_3 + 1.5 * 2.53] = [6.53, 16.63]$ mg

da cui gli outlier basso 3.55 e alto 20.08.

1.3 Rappresentare il boxplot indicando i valori di riferimento (e escludendo gli outliers).



Minimo 8.12, $Q_1 = 10.32$, mediana=11.86, $Q_3 = 12.85$, massimo= 15.53.

Esercizio 2 (2+3+2=7 punti)

La domanda mensile di un reagente utilizzato in un laboratorio di analisi è modellizzata da una variabile aleatoria normale $X \sim \mathcal{N}(\mu, \sigma^2)$ con media $\mu = 80$ e deviazione standard $\sigma = 15$ unità di reagente. Si assume che le domande in mesi diversi siano indipendenti e identicamente distribuite.

2.1 Calcolare la probabilità $\mathbb{P}(X \leq 95)$ che la domanda mensile nel mese corrente sia non superiore alle 95 unità.

$$\mathbb{P}(X \leq 95) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{95 - 80}{15}\right) = \mathbb{P}(Z \leq 1) = 0.84.$$

2.2 Calcolare la probabilità che la domanda totale nei prossimi tre mesi superi 260 unità.

Siano X_1, X_2, X_3 i.i.d. $\mathcal{N}(80, 15^2)$. Allora

$$S_3 = X_1 + X_2 + X_3 \sim \mathcal{N}(3 \cdot 80, 3 \cdot 15^2) = \mathcal{N}(240, 675)$$

dunque

$$\mathbb{P}(S_3 > 260) = \mathbb{P}\left(\frac{S_3 - 240}{\sqrt{675}} > \frac{260 - 240}{\sqrt{675}}\right) = 1 - \mathbb{P}\left(Z \leq \frac{20}{\sqrt{675}}\right) = 1 - \mathbb{P}(Z \leq 0.7698) = 1 - 0.74 = 0.26.$$

2.3 Determinare il 90-esimo percentile della domanda mensile X .

Il 90-esimo percentile è $x_{90} = \mu + \sigma z_{90}$ dove z_{90} è il 90-esimo percentile della normale standard $z_{90} = 1.28$ dunque $x_{90} = 80 + 15 \cdot 1.28 = 99.2$.

Esercizio 3 (2+2+1+2=7 punti)

I seguenti dati rappresentano la concentrazione di azoto x in un ambiente, misurata in milligrammi per litro, e la corrispondente massa y in milligrammi di alghe misurata dopo 48 ore nello stesso ambiente :

$x \text{ (mg l}^{-1}\text{)}$	0.4	0.8	1.2	1.6	2.0	2.4	3.2	3.8
$y \text{ (mg)}$	15	31	42	55	65	71	96	109

3.1 Calcolare il coefficiente di correlazione r e dire se la correlazione è forte o debole.

Abbiamo $n = 8$ coppie di dati di medie campionarie $\bar{x} = \frac{1}{n} \sum x_i$, $\bar{y} = \frac{1}{n} \sum y_i$ e

$$S_{xx} = \sum (x_i - \bar{x})^2, \quad S_{yy} = \sum (y_i - \bar{y})^2, \quad S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

ovvero

$$\bar{x} = 1.925, \quad \bar{y} = 60.5, \quad S_{xx} = 9.595, \quad S_{yy} = 7056.0, \quad S_{xy} = 259.3.$$

Ne segue che

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = 0.997.$$

Correlazione molto forte.

3.2 Trovare a, b della retta di regressione lineare $y = ax + b$, utilizzando il metodo dei minimi quadrati.

$$a = \frac{S_{xy}}{S_{xx}} = 27.02 \text{ l} \quad b = \bar{y} - a\bar{x} = 8.48 \text{ mg}.$$

3.3 Usando la retta di regressione stimata al punto 3.2, determinare il valore di x corrispondente a una predizione di $y = 90$ mg.

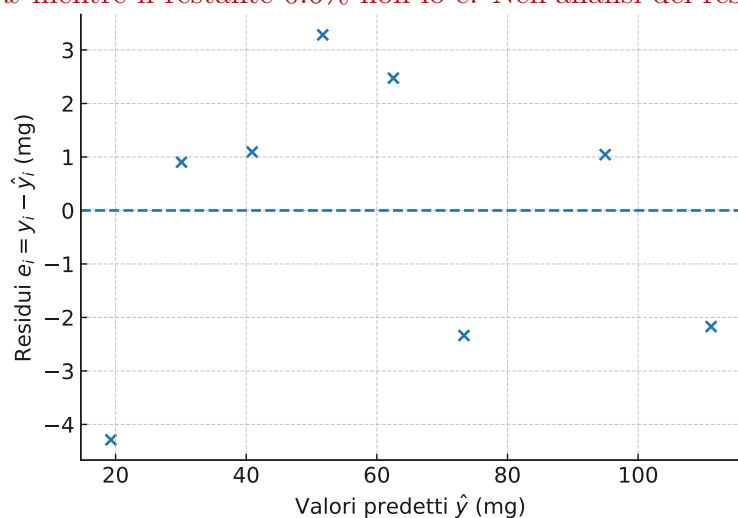
$$x = \frac{90 - b}{a} = \frac{90 - 8.48}{27.02} = 3.02 \text{ mg l}^{-1}.$$

3.4 Calcolare il coefficiente di determinazione r^2 e spiegare cosa rappresenta in termini di variazione della variabile y , complementando l'analisi con il plot dei residui.

$R^2 = r^2 = 0.994$ è dato da

$$R^2 = 1 - \frac{\sum (y_i - a - bx_i)^2}{S_{yy}}$$

e indica che circa il 99.4% della variabilità totale di y è spiegato dalla regressione lineare su x mentre il restante 0.6% non lo è. Nell'analisi dei residui non si distingue alcun pattern.



Esercizio 4 (3+3=6 punti)

Sono stati misurati i tempi di fermentazione (in ore) di $n = 48$ colture:

12.38, 7.08, 9.68, 14.59, 9.45, 6.18, 17.91, 14.25, 14.57, 10.93,
 13.54, 10.48, 13.47, 11.79, 8.36, 16.23, 12.09, 13.14, 7.63, 17.73,
 19.22, 12.86, 12.45, 12.47, 8.94, 10.16, 9.01, 12.73, 11.43, 12.88,
 11.66, 18.01, 8.25, 13.22, 14.41, 15.03, 14.60, 11.56, 11.93, 12.33,
 13.69, 9.35, 6.32, 11.96, 15.37, 7.93, 10.92, 13.32.

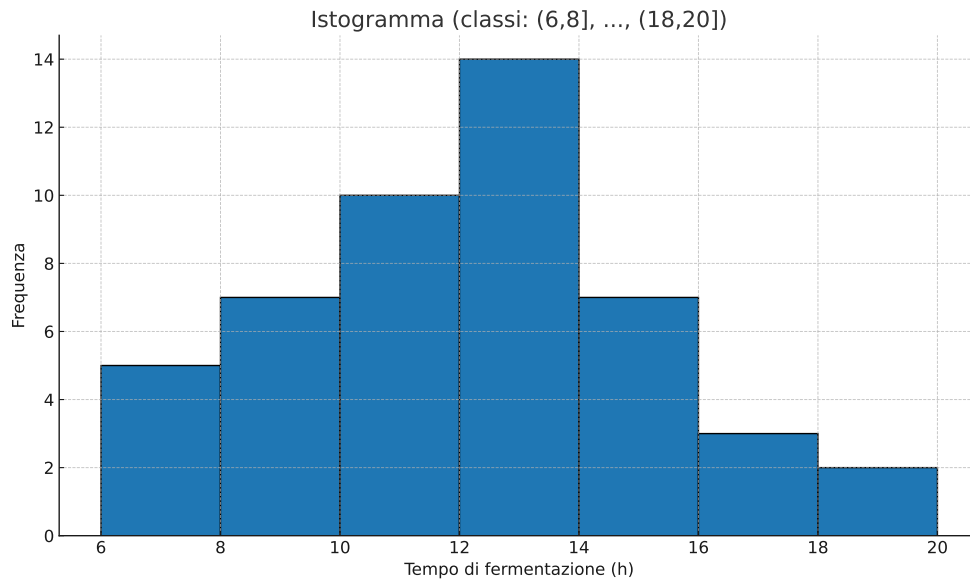
4.1 Calcolare le frequenze dei dati del campione usando classi di ampiezza 2, e quindi produrre il relativo istogramma.

Dato che il minimo è 6.18 e il massimo 19.22, usando ad esempio le classi

$(6, 8]$, $(8, 10]$, $(10, 12]$, $(12, 14]$, $(14, 16]$, $(16, 18]$, $(18, 20]$

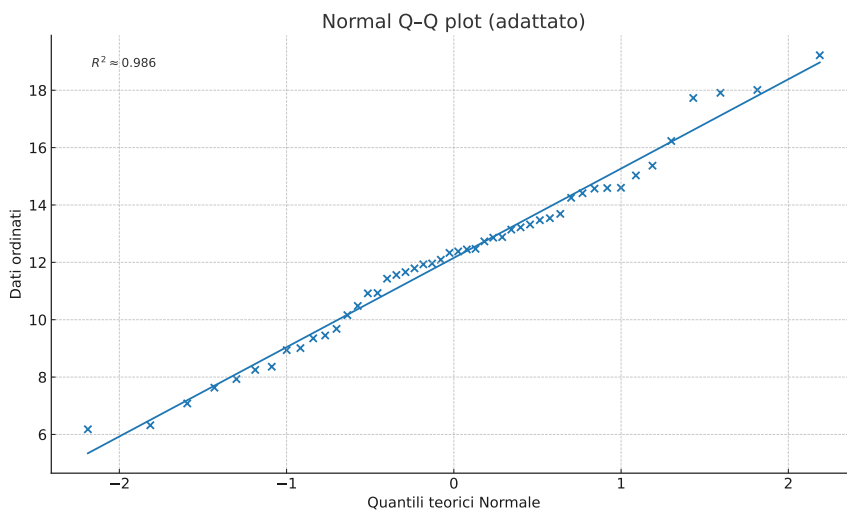
le frequenze osservate sono

Classe	(6, 8]	(8, 10]	(10, 12]	(12, 14]	(14, 16]	(16, 18]	(18, 20]
Frequenza	5	7	10	14	7	3	2



4.2 Giudicare il carattere normale dei dati del campione realizzando un normal plot.

I punti aderiscono bene alla linea retta e l'approssimazione normale è buona.



Esercizio 5 (4+3=7 punti)

I seguenti dati riportano il numero di incidenti stradali con feriti registrati, per giorno della settimana, in un capoluogo di regione nell'arco di diversi anni:

Giorno	Domenica	Lunedì	Martedì	Mercoledì	Giovedì	Venerdì	Sabato
Numero casi	92	78	101	88	85	112	104

5.1 Al livello di significatività del 5%, testare l'ipotesi che un incidente sia equamente

probabile in ciascuno dei 7 giorni della settimana.

Effettuiamo un test di χ^2 . Secondo l'ipotesi nulla H_0 che la probabilità sia la stessa per ciascuno dei 7 giorni, le frequenze attese sono tutte pari a

$$E_i = \frac{N}{7} = \frac{660}{7} = 94.286$$

$i = 1, \dots, 7$ essendo $N = 92+78+101+88+85+112+104 = 660$. La statistica di test

$$ST = \sum_{i=1}^7 \frac{(O_i - E_i)^2}{E_i} = 0.055 + 2.813 + 0.478 + 0.419 + 0.915 + 3.328 + 1.001 = 9.009 .$$

Essendo i gradi di libertà $7 - 1 = 6$, il valore critico al 5% è $\chi^2 = 12.592 > ST$ e la conclusione è che non possiamo rifiutare H_0 .

5.2 Spiegare che cosa cambia se due soli dati sono modificati come segue: Lunedì 70 incidenti (anziché 78) e Venerdì 120 incidenti (anziché 112). Determinare in particolare il nuovo livello di fiducia al quale l'ipotesi è non rifiutabile.

Il calcolo della statistica di test si modifica con i due nuovi contributi

$$ST = 0.055 + 6.26 + 0.478 + 0.419 + 0.915 + 7.01 + 1.001 = 16.14.$$

Confrontando con la tabella di valori di χ^2 a 6 gradi di libertà troviamo i valori $14.449 < 16.14 < 16.812$ che corrispondono a un valore $p \in (0.01, 0.025)$. Rifiutiamo dunque H_0 al livello 5% ma non rifiutiamo al livello 1%.

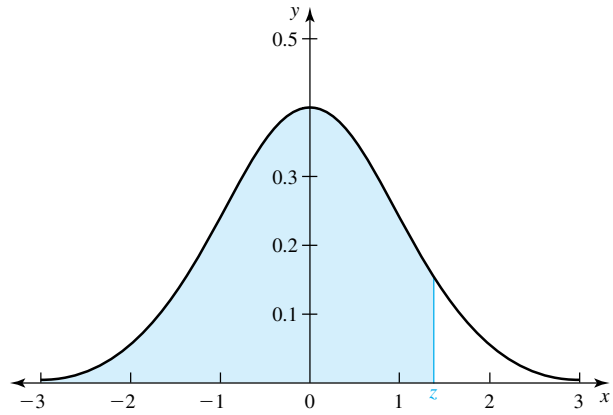
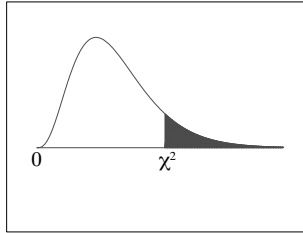


Figure B.1 Areas under the standard normal curve from $-\infty$ to z .

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5754
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7258	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7518	.7549
0.7	.7580	.7612	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7996	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_{\alpha}$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169