

Statistica Applicata
Corso di Laurea in Scienze Naturali
a. a. 2016/2017

prof. Federico Plazzi

4 Settembre 2017

Nome: _____

Cognome: _____

Matricola: _____

Alcune indicazioni:

- La prova è costituita da cinque esercizi; dopo ogni esercizio c'è lo spazio in cui scrivere la risposta o le risposte. In caso questo spazio non sia sufficiente, si può continuare a rispondere sul retro del foglio, avendo cura di indicare il numero dell'esercizio a fianco della continuazione della risposta.
- Alcuni esercizi richiedono semplici calcoli, per i quali è consentito l'uso di una calcolatrice ed eventualmente la consultazione di una o più delle tabelle allegate.
- Altri esercizi richiedono invece la lettura dei dati: verrà valutata in questo caso l'argomentazione che giustifica l'interpretazione fornita.
- La durata massima della prova è di 60 minuti.
- Si prega di non scrivere nulla sulle tabelle allegate.

1 Dati

Alcuni malacologi (ossia zoologi specializzati nel phylum dei molluschi) decidono di studiare il dimorfismo sessuale della specie *Ruditapes philippinarum* (Adams & Reeve, 1850). A questo scopo, effettuano quattro campionamenti in laguna durante la stagione riproduttiva, misurano la lunghezza della conchiglia (la distanza estremità posteriore - estremità anteriore) di tutti gli esemplari raccolti e poi determinano il sesso osservando al microscopio il contenuto della gonade. I risultati sono mostrati nella tabella 1.

Tabella 1: Sesso e lunghezza (mm) degli esemplari catturati nei quattro campionamenti in altrettanti siti lagunari.

Campionamento 1			Campionamento 2			Campionamento 3			Campionamento 4		
Sesso	Lunghezza										
1	F	47,16	1	F	61,17	1	F	63,49	1	F	53,73
2	M	58,85	2	M	57,78	2	F	58,42	2	F	52,86
3	M	47,26	3	F	57,90	3	M	75,16	3	M	52,99
4	M	49,64	4	M	60,14	4	F	55,99	4	M	55,96
5	F	72,32	5	M	66,12	5	F	66,76	5	M	65,36
6	M	68,25	6	F	62,09	6	M	52,52	6	F	59,64
7	M	52,15	7	M	55,18	7	M	53,67	7	F	58,21
8	F	52,06	8	M	60,97	8	F	62,82	8	F	55,12
9	F	59,47	9	M	63,83	9	F	65,46	9	F	56,43
10	M	49,18	10	F	61,32	10	F	55,98	10	F	57,55
11	F	54,22	11	M	61,55	11	F	53,98	11	F	61,60
12	F	69,36	12	F	56,90	12	F	62,41	12	M	61,22
13	F	61,41	13	F	60,89	13	M	59,65	13	F	53,93
14	M	62,05	14	F	59,91	14	F	48,83	14	F	47,11
15	M	59,75	15	F	49,54	15	M	55,94	15	F	65,28
16	M	57,43	16	F	58,84	16	F	54,83	16	M	66,75
17	M	60,76	17	F	60,15	17	F	46,48	17	M	60,14
18	F	58,97	18	F	59,48	18	M	61,76	18	F	56,89
19	M	51,13	19	F	58,99	19	M	57,77	19	M	58,31
20	F	59,86	20	F	54,37	20	M	55,74	20	F	59,63
21	F	61,11	21	F	57,35	21	F	57,84	21	M	56,84
			22	M	61,52				22	F	52,90
			23	M	66,79				23	F	52,61
			24	F	51,14				24	F	52,86
									25	F	58,41

2 Esercizi

2.1 Statistiche di base

Calcolare il numero medio di femmine raccolte per campione. Calcolare anche la varianza e la deviazione standard di questo valor medio.

<i>Media</i>	<i>Varianza</i>	<i>Deviazione standard</i>
13,75	6,6875	2,58602

2.2 Rapporto sessi

La prima domanda riguarda il rapporto sessi: maschi e femmine sono sostanzialmente in ugual numero o ci sono deviazioni significative da una sostanziale parità? Qual è il test statistico più adatto e quale sarebbe il risultato?

Il sesso è una variabile qualitativa, per cui il test più indicato in questo caso è il χ^2 . In tutti e quattro i casi il risultato del test non è significativo: non si può quindi rifiutare l'ipotesi nulla e bisogna assumere un rapporto sessi bilanciato, senza particolari deviazioni. Il risultato del test, peraltro, non risulta significativo neanche unendo tutti i dati in un'unica tabella di contingenza 4×2 , per cui non si sarebbero neanche potute ravvisare particolari differenze tra campioni in merito al rapporto sessi.

2.3 Distribuzione della variabile

Per prima cosa, è necessario capire come si distribuisca la variabile “lunghezza”. A questo scopo, si esegue il test di Shapiro e Wilk su tutti gli individui insieme; i risultati sono mostrati di seguito. Cosa si può concludere?

Shapiro-Wilk test

```
data: c(Length1, Length2, Length3, Length4)
W = 0.98463, p-value = 0.3623
```

Il test di Shapiro e Wilk è un test di normalità e l'ipotesi nulla è che la variabile in questione segua una distribuzione normale. Dato che il test risulta non significativo (p -value = 0,3623), questa ipotesi nulla non si può rifiutare e quindi possiamo considerare i dati a distribuzione normale.

2.4 Differenze tra i campioni

Gli studiosi sono interessati ora a capire se ci siano differenze nelle dimensioni dei quattro campioni; di seguito sono presentati i risultati dell'ANOVA. Si tratta dell'approccio corretto?

Perché? Cosa possiamo concludere?

Tabella 2: ANOVA sulle lunghezze ripartite per campionamento.

	g.l.	D	σ^2	F	p-value
<i>tra</i>	3	56,26	18,753	0,6019	0,6155
<i>entro</i>	87	2710,34	31,153		

L'ANOVA è l'approccio corretto, visto che i dati sono a distribuzione normale e i campioni più o meno della stessa dimensione. Non ha dato risultati significativi (p-value = 0,6155), per cui non si evidenziano differenze significative di dimensioni tra i quattro campioni.

2.5 Dimorfismo sessuale

A questo punto, è possibile confrontare le dimensioni di tutte le femmine campionate con tutti i maschi. A questo scopo, sono mostrati di seguito quattro test possibili. Qual è quello più indicato e perché? Cosa se ne conclude a proposito del dimorfismo sessuale di questa specie?

- Test t a campioni indipendenti

Two Sample t-test

```
data: Lengths[Sexes == "F"] and Lengths[Sexes == "M"]
t = -0.97249, df = 69.233, p-value = 0.3342
alternative hypothesis: true difference in means is not equal
to 0
```

- Test t a campioni appaiati

Paired t-test

```
data: Lengths[Sexes == "F"] and Lengths[Sexes == "M"]
t = -0.98038, df = 27, p-value = 0.3356
alternative hypothesis: true difference in means is not equal
to 0
```

- Test di Mann e Whitney

Wilcoxon rank sum test

```
data: Lengths[Sexes == "F"] and Lengths[Sexes == "M"]
W = 877, p-value = 0.3612
alternative hypothesis: true location shift is not equal to 0
```

- Test di Wilcoxon

Wilcoxon signed rank test

```
data: Lengths[Sexes == "F"] and Lengths[Sexes == "M"]
V = 183, p-value = 0.2094
alternative hypothesis: true location shift is not equal to 0
```

Il test corretto è il test t a campioni indipendenti, visto che la variabile è a distribuzione normale e che non c'è nessuna ragione per confrontare una femmina con uno specifico maschio; oltretutto, femmine e maschi non sono nello stesso numero. Dato il valore del p-value (0,3342), il test non risulta significativo: non si può rigettare l'ipotesi nulla. Di conseguenza, non si hanno prove di dimorfismo sessuale in questa specie, almeno in termini di lunghezza della conchiglia.