

Test 08-GIU-2015

1 (L05)

In una cucciolata di 8 cuccioli quale è la probabilità di ottenere 4 maschi e 4 femmine sapendo che la probabilità dei maschi è uguale a quella delle femmine?

8 punti

2 (L15)

In un monitoraggio dei conigli selvatici in una Area Protetta tramite i segni di presenza (conteggio delle feci), sono state rinvenute 20 feci nel grano, 20 nella medica, 32 negli incolti cespugliati e 40 nella macchia mediterranea. Sapendo che gli ambienti sono tutti di 15 ettari testa se la presenza delle feci (=conigli) differisce statisticamente in relazione ai diversi ambienti e se questo avviene individua quali ambienti sono statisticamente più o meno utilizzati dell'atteso (preferiti/evitati). **12 punti**

3 (L10)

Controlla nei seguenti dati pubblicati se le differenze attestata sono effettivamente al 5%. **10 punti**

N =	10	10
Medie =	145,2 a	200,0 b
Ms-errore o Varianza-errore =	243,72	
d.s.	16,20	15,00

la probabilità dell'evento A è p la probabilità dell'evento non A è q (p+q=1) quindi se p=q; p e q =0,5. utilizzo la

$$P(s \text{ di } A \text{ e } r \text{ di non } A) = \frac{n!}{r! \cdot s!} \times p^s \times q^r$$

Combinazione		$p^s \cdot q^r$	$n!/(s! \cdot r!)$	P
maschi	femmine	p(maschi)=0,5 q(femmine)=0,5	da triangolo Tartaglia	
= s	= r			
8	0	0,00390625	1	0,0039
7	1	0,00390625	8	0,0313
6	2	0,00390625	28	0,1094
5	3	0,00390625	56	0,2188
4	4	0,00390625	70	0,2734
3	5	0,00390625	56	0,2188
2	6	0,00390625	28	0,1094
1	7	0,00390625	8	0,0313
0	8	0,00390625	1	0,0039
totale =				1,0000

0									1												
1										1		1									
2											1	2	1								
3												3	3	1							
4													6	4	1						
5														10	10	5	1				
6															15	20	15	6	1		
7																21	35	35	21	7	1
8																				8	1

1 (L05)

Risposta: la probabilità è del 27,34%

Le probabilità delle altre combinazioni sono inutili e sono state calcolate solo a fini di esercizio

2 (L15)

Calcolo il chi quadro per verificare l'ipotesi se le feci non sono distribuite in modo uniforme nei diversi ambienti

	osservato	atteso	differenze	quadrati	divisi attesi
MEDICA	20	28	-8	64	2,2857
ORZO	20	28	-8	64	2,2857
INCOLTI CESPUGLIATI	32	28	4	16	0,5714
MACCHIA MEDITERANEA	40	28	12	144	5,1429
	112	112		CHI2 CALCOLATO =	10,29
I gradi di libertà sono 3 = 4 numeri "liberi" meno 1				CHI QUADRO TABULATO 0,05 =	7,81
				CHI QUADRO TABULATO 0,01 =	11,34

15*4=60ha totali
Feci attese=
112 * 15/60 =
28

Probabilità percentuale di un valore di chi quadro maggiore del valore in tabella

gl	0,95	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01
1	0,004	0,02	0,06	0,15	0,27	0,45	0,71	1,07	1,64	2,71	3,84	5,41	6,63
2	0,10	0,21	0,45	0,71	1,02	1,39	1,83	2,41	3,22	4,61	5,99	7,82	9,21
3	0,35	0,58	1,01	1,42	1,87	2,37	2,95	3,66	4,64	6,25	7,81	9,84	11,34
4	0,71	1,06	1,65	2,19	2,75	3,36	4,04	4,88	5,99	7,78	9,49	11,67	13,28
5	1,15	1,61	2,34	3,00	3,66	4,35	5,13	6,06	7,29	9,24	11,07	13,39	15,09

10,29 è maggiore di 7,81 ma non di 11,34 quindi l'ipotesi che le feci siano distribuite in modo uniforme nei diversi ambienti è significativamente ($p < 0,05$) confutata dai risultati delle osservazioni. Procedo quindi a valutare quali sono gli ambienti più utilizzati e meno utilizzati della media

Per valutare quali ambienti sono utilizzati più o meno dell'atteso devo utilizzare la correzione di Yates avendo ognuno di loro un solo grado di libertà:

2 (L15)

	osservato	superficie	per ettaro	atteso	atteso	Yates	differenze	quadrati	divisi attesi = CHI2c
MEDICA	20	15	1,333	28	0,25	0,5	-7,5	56,25	2,0089
ORZO	20	15	1,333	28	0,25	0,5	-7,5	56,25	2,0089
INCOLTI CESPUGLIATI	32	15	2,133	28	0,25	-0,5	3,5	12,25	0,4375
MACCHIA MEDITERANEA	40	15	2,667	28	0,25	-0,5	11,5	132,25	4,7232 *
	112	60		112	1				

Probabilità percentuale di un valore di chi quadro maggiore del valore in tabella

gl	0,95	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01
1	0,004	0,02	0,06	0,15	0,27	0,45	0,71	1,07	1,64	2,71	3,84	5,41	6,63
2	0,10	0,21	0,45	0,71	1,02	1,39	1,83	2,41	3,22	4,61	5,99	7,82	9,21
3	0,35	0,58	1,01	1,42	1,87	2,37	2,95	3,66	4,64	6,25	7,81	9,84	11,34
4	0,71	1,06	1,65	2,19	2,75	3,36	4,04	4,88	5,99	7,78	9,49	11,67	13,28
5	1,15	1,61	2,34	3,00	3,66	4,35	5,13	6,06	7,29	9,24	11,07	13,39	15,09

RISPOSTA: Solo 4,7232 è maggiore di 3,84 ma non di 6,63 quindi l'unico ambiente significativamente più utilizzato rispetto all'atteso è la macchia mediterranea.

Impiego l'analisi della varianza:

3 (L10)

n	10	10	20
media	145,2	200,0	
d.s.	16,20	15,00	
Mserrore	243,72		

Somma totale dati = 3452 = 10*145,2+10*200 somma totale2 = 11916304

Somma dei quadrati degli scarti fra gruppi $n_a * x_a^2 + n_b * x_b^2 - \frac{-(somma\ totale)^2}{n_a + n_b}$

$10*145,2^2 + 10*200,0^2 - 3452^2 / 20$

$210.830,40 + 400.000,00 - 595.815,20 = 15.015,20$

Somma dei quadrati degli scarti entro gruppi = MS errore*gradi di libert  = 243,72 * 18 = 4.386,96

Somma dei quadrati degli scarti entro gruppi $ds_a^2*(n_a-1) + ds_b^2*(n_b-1)$

$16,20^2*9 + 15,00^2*9$

$2.361,96 + 2.025,00 = 4.386,96$

Sorgenti di variazione	Somme quadrati Scarti SS	gradi di libert� gl o df	Varianze MS	Rappo rti F
Fra gruppi TRATTAME NTI	15.015,20	1	15.015,20	41,552
entro gruppi ERRORE	4.386,96	18	243,72	
TOTALE	19.402,16	19		

F da tabelle	p
4,414	0,05
8,285	0,01

il valore di f calcolato   altamente significativo

Risposta: la differenza fra le medie non   solo significativa ma altamente significativa

3 (L10)

probabilità =0,05 di un valore più elevato di F.

g.l.	1	2	3	4
1	161,448	199,5	215,707	224,583
2	18,513	19	19,164	19,247
3	10,128	9,552	9,277	9,117
4	7,709	6,944	6,591	6,388
5	6,608	5,786	5,409	5,192
6	5,987	5,143	4,757	4,534
7	5,591	4,737	4,347	4,12
8	5,318	4,459	4,066	3,838
9	5,117	4,256	3,863	3,633
10	4,965	4,103	3,708	3,478
12	4,747	3,885	3,49	3,259
14	4,6	3,739	3,344	3,112
16	4,494	3,634	3,239	3,007
18	4,414	3,555	3,16	2,928
20	4,351	3,493	3,098	2,866
25	4,242	3,385	2,991	2,759
30	4,171	3,316	2,922	2,69
40	4,085	3,232	2,839	2,606

probabilità =0,01 di un valore più elevato di F.

g.l.	1	2	3	4
1	4052,181	4999,5	5403,352	5624,583
2	98,503	99	99,166	99,249
3	34,116	30,817	29,457	28,71
4	21,198	18	16,694	15,977
5	16,258	13,274	12,06	11,392
6	13,745	10,925	9,78	9,148
7	12,246	9,547	8,451	7,847
8	11,259	8,649	7,591	7,006
9	10,561	8,022	6,992	6,422
10	10,044	7,559	6,552	5,994
12	9,33	6,927	5,953	5,412
14	8,862	6,515	5,564	5,035
16	8,531	6,226	5,292	4,773
18	8,285	6,013	5,092	4,579
20	8,096	5,849	4,938	4,431
25	7,77	5,568	4,675	4,177
30	7,562	5,39	4,51	4,018
40	7,314	5,179	4,313	3,828

3 (L10)

Impiego il test t di Student:

$$d = (\bar{X}_A - \bar{X}_B)$$

= t

$$\sqrt{\frac{\sum(x - \bar{x})_A^2 + \sum(x - \bar{x})_B^2}{n_A + n_B - 2} * \frac{n_A + n_B}{n_A * n_B}}$$

Risposta: la differenza fra le medie non è solo significativa ma altamente significativa (t calcolato maggiore del t tabulato sia a 0,05 che 0,01)

	A	B
n =	10	10
media =	145,2	200,0
SS=	2.361,96	2.025,00
d.s. =	16,2	15,0
e.s. =	5,1228898	4,7434165
d (A-B)=	54,8	
$n_A + n_B =$	20	
$n_A * n_B =$	100	
$gl=n_A + n_B -2 =$	18	
$ds_d^2 =$	48,744	
$ds_d =$	6,98169	
t calcolato =	7,85	
$T_{tabulato}_{0,05} =$	2,10	
$T_{tabulato}_{0,01} =$	2,88	