

ESERCIZIO 1

La seguente tabella riporta il prezzo, il contenuto di grassi, di calorie e di sodio per diciotto panini serviti al fast-food MangiaVelocemente. I panini sono distinti in due tipologie, come indicato sulla prima colonna. Il nuovo gestore del fast-food vuole analizzare la relazione esistente tra il contenuto di grassi dei panini e il loro prezzo.

Type	Price	Total Fat	Calories	Sodium
Burger	2.09	19	410	920
Burger	2.79	31	580	1500
Burger	2.24	34	590	1310
Burger	1.49	35	570	860
Burger	2.59	39	640	1180
Burger	2.34	39	680	940
Burger	2.29	43	660	1260
Chicken	2.99	7	300	740
Chicken	2.39	9	290	870
Chicken	2.99	15	350	880
Chicken	2.99	16	420	930
Chicken	2.89	16	430	750
Chicken	2.94	18	450	970
Chicken	2.99	22	450	940
Chicken	2.89	25	550	1110
Chicken	2.79	27	550	1180
Chicken	4.49	33	750	1860
Chicken	2.94	39	660	1330

A partire dalle seguenti informazioni:

PANINI DI TIPO BURGER	$\bar{x} = 34.29$ $\bar{y} = 2.26$ $\sum (x_i - \bar{x})^2 = 365.42$ $\sqrt{\sum (y_i - \hat{y}_i)^2 / (n-2)} = 0.4477$	$\sum (x_i - \bar{x})(y_i - \bar{y}) = 2.50$ $\sum (y_i - \bar{y})^2 = 1.01$
PANINI DI TIPO CHICKEN	$\bar{x} = 20.64$ $\bar{y} = 3.03$ $\sum x_i^2 = 5639$ $\sqrt{\sum (y_i - \hat{y}_i)^2 / (n-2)} = 0.4891$	$\sum x_i y_i = 710.48$ $\sum y_i^2 = 103.40$
TUTTI I PANINI	NOTA: sfruttare le proprietà della media e la decomposizione della varianza per ricavare le informazioni su tutti i panini a partire da quelle disponibili per i gruppi $\sum (x_i - \bar{x})(y_i - \bar{y}) = -18.66$ $\sqrt{\sum (y_i - \hat{y}_i)^2 / (n-2)} = 0.6133$	

a) Costruire il diagramma di dispersione della variabile Fat (variabile esplicativa) contro la variabile Price (variabile di risposta) per tutti i panini e separatamente per i due tipi:

TUTTI I PANINI	
-----------------------	--

PANINI DI TIPO BURGER	
PANINI DI TIPO CHICKEN	

b) Che tipo di informazioni può desumere il gestore del fast-food dall'analisi dei tre grafici?

--

c) Stimare i coefficienti della retta di regressione della variabile Fat (variabile esplicativa) contro la variabile Price (variabile di risposta) per tutti i panini e separatamente per i due tipi:

TUTTI I PANINI	
PANINI DI TIPO BURGER	
PANINI DI TIPO CHICKEN	

d) Rappresentare sui diagrammi di dispersione del punto a le tre rette di regressione

e) Che tipo di informazioni può desumere il gestore del fast-food dall'analisi dell'andamento delle tre rette di regressione?

--

f) Valutare la bontà di adattamento dei tre modelli di regressione stimati:

TUTTI I PANINI	
PANINI DI TIPO BURGER	
PANINI DI TIPO CHICKEN	

g) Valutare la significatività dei tre modelli di regressione (usando il test sul coefficiente angolare):

TUTTI I PANINI	
PANINI DI TIPO BURGER	

PANINI DI TIPO CHICKEN	
-------------------------------	--

h) Quali ipotesi sono necessarie per svolgere il punto precedente?

--

i) Quali conclusioni può desumere il gestore del fast-food dall'esame dei risultati dei test?

--

l) Valutare la significatività dei tre modelli di regressione (usando il test sull' R^2):

TUTTI I PANINI	
PANINI DI TIPO BURGER	
PANINI DI TIPO CHICKEN	

ESERCIZIO 2

Il nuovo gestore del fast-food ritiene, dall'esame dell'incartamento lasciategli dal vecchio gestore, che il contenuto medio di grassi (variabile FAT) dei panini sia pari a 20. In seguito all'esame degli ordini degli ultimi lotti di ingredienti e da un sommario esame del menu ha il dubbio che tale valore sottostimi il reale contenuto di grassi dei panini serviti.

a) Servendosi del campione riportato nella tabella 1, procedere ad un test di ipotesi ($\alpha = 0.1$) sapendo che dalle osservazioni passate è possibile ritenere che la variabile FAT segue una distribuzione normale con varianza pari a 121:

--

b) Calcolare le probabilità dell'errore di I e II specie e la potenza del test per le seguenti ipotesi alternative:

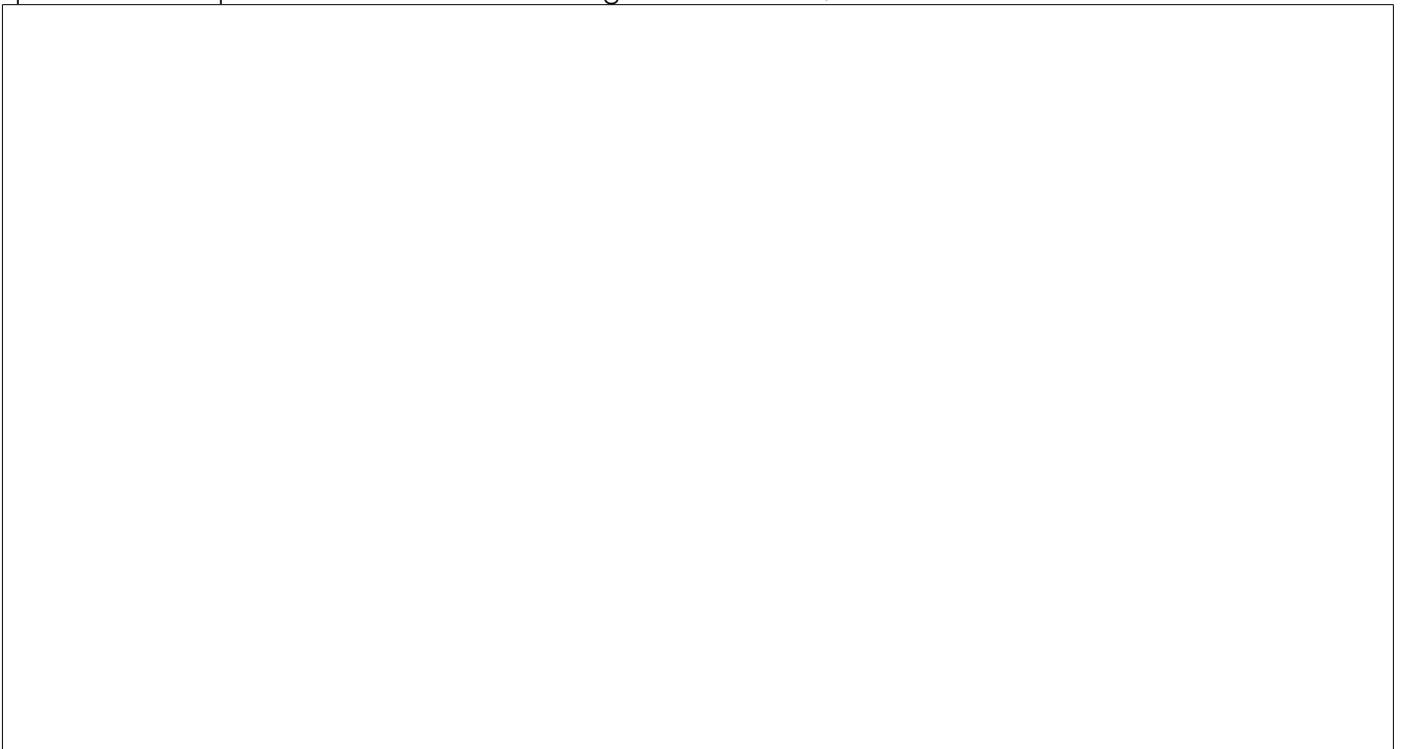
IPOTESI ALTERNATIVA	a	b	1-b
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			

c) Disegnare la curva di potenza del test:



ESERCIZIO 3

Un produttore di lampade a incandescenza informa i suoi clienti che la qualità dei suoi prodotti non è uniforme, e che ogni lampadina può essere indipendentemente di qualità A, B, C, D o E con probabilità del 15%, 25%, 35%, 20% e 5% rispettivamente. Uno dei clienti, acquistando grossi volumi di merce, ha tuttavia l'impressione di ricevere troppi pezzi di qualità E (la peggiore) e decide quindi di verificare l'affermazione del produttore investendo tempo e denaro per stabilire il livello qualitativo di 30 lampade. Il campione risulta in particolare composto da 3 lampade di qualità A, da 6 di qualità B, 9 di qualità C, 7 di qualità D e 5 di qualità E. Usando un livello di significatività del 5% cosa si decide?



ESERCIZIO 4

Si supponga che il numero di incidenti settimanali sul lavoro nella ditta AutoCar SpA in un periodo di 30 settimane sia stato il seguente:

8	0	0	1	3	4	0	2	12	5	1	8	0	2	0
1	9	3	4	5	3	3	4	7	4	0	1	2	1	2

Si verifichi l'ipotesi che la distribuzione del numero di incidenti settimanali sia del tipo di Poisson.

SUGGERIMENTO: usare la media campionaria per stimare il parametro della distribuzione di Poisson rispetto a cui effettuare il test di adattamento

ESERCIZIO 5

Indicare se le seguenti affermazioni sono vere o false:

- | | | |
|---|----------------------------|-----------------------------|
| - Gli stimatori dei minimi quadrati per i coefficienti della retta di regressione sono non distorti | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Le ipotesi alla base del modello classico di regressione richiedono, tra l'altro, la normalità degli errori | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - La regressione è un metodo che permette di studiare l'indipendenza assoluta tra due variabili X ed Y | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - La regressione è un metodo di analisi simmetrico | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Il coefficiente di determinazione lineare è un indice che varia tra -1 e +1 | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Il coefficiente di determinazione lineare vale -1 se i punti sono perfettamente allineati su una retta con pendenza crescente | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Il coefficiente di determinazione lineare vale -1 se i punti sono perfettamente allineati su una retta con pendenza decrescente | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Verificare la significatività del modello di regressione consiste nel valutare l'ipotesi che l'intercetta sia significativamente diversa da 0 | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Il coefficiente di determinazione lineare vale 1 solo se i punti sono perfettamente allineati su una retta | <input type="radio"/> Vero | <input type="radio"/> Falso |
| - Il coefficiente di determinazione lineare vale 1 solo se il coefficiente di correlazione lineare vale +1 | <input type="radio"/> Vero | <input type="radio"/> Fa |

- Il test di indipendenza mira a verificare se il coefficiente chi-quadro calcolato sul campione osservato sia significativamente diverso da 0 Vero Falso
- In un test sulla significatività del coefficiente chi-quadro, l'ipotesi nulla afferma che vi sia dipendenza tra le due variabili Vero Falso

ESERCIZIO 6

La seguente tabella riporta la ripartizione di un campione di 374 abitanti della nazione RAINBOW in base al colore dei capelli e al genere:

		GENERE		TOTALE
		Femminile	Maschile	
COLORE DEI CAPELLI	Biondi	27	32	59
	Castani	84	68	152
	Rossi	12	29	41
	Neri	57	65	122
TOTALE		180	194	374

Sapendo che l'indice chi-quadro risulta pari a 9.163, sottoporre a verifica l'ipotesi di indipendenza tra COLORE DEI CAPELLI e GENERE,

a) usando un livello di significatività $\alpha = 0.1$:

b) usando un livello di significatività $\alpha = 0.05$:

c) usando un livello di significatività $\alpha = 0.01$:

ESERCIZIO 7

I ricercatori di un'importante società finanziaria che opera prevalentemente on-line vogliono determinare se le scadenze delle obbligazioni industriali con rating AAA e di quelle con rating CCC hanno varianze diverse. Su un campione di 17 obbligazioni con rating AAA si è osservata una devianza campionaria pari a 1963.6 su un campione indipendente di 11 obbligazioni con rating CCC si è osservata una devianza campionaria pari a 80.2.

a) Si proceda alla verifica di ipotesi usando un livello di significatività pari a 0.02.

b) E' necessario formulare ipotesi per procedere al test richiesto al punto a? (motivare brevemente la risposta)