Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2017-2018

STATISTICA – 24.01.18 - II PROVA PARZIALE

Modalità B -- *TRACCIA DI SOLUZIONE*

(A) ai fini della valutazione verranno considerate **solo** le risposte riportate dallo studente **negli appositi riquadri** bianchi
(B) nello svolgimento del compito si utilizzino almeno tre cifre decimali.

COGNOME………………………….NOME…..……………………..MATR………………..

ESERCIZIO 1 (8 punti)

Su un campione di 22 notebook tra quelli in commercio si rileva il peso in Kg. I 22 pesi rilevati hanno una media pari a 1.91 ed una varianza pari a 0.09. Si assume distribuzione normale per il peso dei notebook.
a) Si stabilisca a livello 0.05 se il peso medio dei notebook in commercio è inferiore a 2 Kg, usando il metodo del p-value.
b) Si determini l'intervallo di confidenza al 90% per il peso medio dei notebook in commercio.
c) Si dica se la lunghezza dell'intervallo di confidenza per il peso medio dei notebook in commercio diminuisce o aumenta al diminuire del livello di confidenza (a parità dei dati campionari).

a) Le ipotesi sono: $H_0: \mu = 2$ e $H_1: \mu < 2$. Il p-value del test è:

$$p - value = P(T \leq t_{oss}) = P\left(T \leq \frac{1.91 - 2}{\sqrt{\frac{0.09}{22}}}\right) = P(T \leq -1.407) = 1 - P(T \leq 1.407),$$

con la probabilità calcolata supponendo vera l'ipotesi nulla, nel qual caso la statistica test *T* ha distribuzione T di Student con 21 gradi di libertà. Dalle tavole si rileva che il valore della probabilità sopra indicata (cioè il p-value) è compreso tra 0.05 e 0.1, per cui è superiore al livello fissato 0.05. Quindi non si rifiuta l'ipotesi nulla a livello 0.05; non c'è dunque evidenza per stabilire, a questo livello, che il peso medio dei notebook in commercio è inferiore a 2 Kg.

b) L'intervallo è

$$(\bar{x} - t^{21}_{\frac{\alpha}{2}} s/\sqrt{n}, \bar{x} + t^{21}_{\frac{\alpha}{2}} s/\sqrt{n}) = (1.91 - 1.721\frac{\sqrt{0.09}}{\sqrt{22}}, 1.91 + 1.721\frac{\sqrt{0.09}}{\sqrt{22}}) = (1.799, 2.020)$$

c) Al diminuire del livello di confidenza (a parità di dati campionari) la lunghezza dell'intervallo diminuisce; ovvero, al diminuire del livello di fiducia sul fatto che la media appartenga all'intervallo aumenta la precisione (ovvero diminuisce la lunghezza) dell'intervallo.

ESERCIZIO 2 (4 punti)

Il gestore di una sala cinematografica vuole stabilire se è opportuno introdurre degli sconti sui biglietti nei giorni feriali per incrementare gli incassi. Su un primo campione di 20 giorni feriali in cui non vengono praticati gli sconti rileva un incasso medio giornaliero pari a 2800 euro ed una deviazione standard pari a 410 euro. Su un secondo campione di 10 giorni feriali, in cui vengono praticati gli sconti, l'incasso medio giornaliero è pari a 3000 euro e la deviazione standard corrispondente 405 euro. Si assumono distribuzioni normali con varianze uguali per gli incassi giornalieri nei giorni feriali in cui sono praticati gli sconti e per quelli nei giorni feriali in cui non lo sono. Si stabilisca, a livello 0.01, se l'introduzione degli sconti sui biglietti incrementa gli incassi medi giornalieri dei giorni feriali. Si giustifichi la risposta, riportando ipotesi nulla ed alternativa e regione di rifiuto del test.

Le ipotesi sono: $H_0: \mu_X = \mu_Y$ e $H_1: \mu_X < \mu_Y$, dove $\mu_X e\ \mu_Y$ sono gli incassi medi giornalieri quando non vengono effettuati gli sconti e quando vengono effettuati rispettivamente. La regione di rifiuto del test è

$$R_{0.05}: T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_p^2}{n_X} + \frac{s_p^2}{n_Y}}} < -t^{28}_{0.01} = -2.467.$$

Usando i dati campionari si ottiene

$$s_p^2 = \frac{410^2 * 19 + 405^2 * 9}{20 + 10 - 2} = 166790.2$$

e quindi

$$T_{oss} = \frac{2800 - 3000}{\sqrt{\frac{166790.2}{20} + \frac{166790.2}{10}}} = -1.264.$$

Essendo il valore osservato della statistica test maggiore di -2.467, non si rifiuta l'ipotesi nulla a livello 0.01, per cui non c'è evidenza, a questo livello, del fatto che l'introduzione degli sconti incrementi gli incassi medi.

<u>ESERCIZIO 3 (8 punti)</u>

Si vuole stimare la proporzione di consegne effettuate da un corriere che arrivano a destinazione in ritardo rispetto alla data stabilita. Si considera un campione di 150 consegne; di queste, 20 arrivano a destinazione in ritardo.

a) Si determini una stima puntuale della proporzione di consegne che arrivano a destinazione in ritardo.

b) Si determini lo standard error della stima calcolata al punto precedente.

c) Si determini un intervallo di confidenza al 99% per la proporzione di consegne che arrivano a destinazione in ritardo.

a) La stima puntuale è $\hat{p} = \bar{x} = \frac{20}{150} = 0.133$.

b) Lo standard error è SE= $\sqrt{\frac{\bar{x}\,(1-\bar{x})}{n}} = \sqrt{\frac{0.133*0.867}{150}} = 0.028$.

c) L'intervallo è

$$(\bar{x} - 2.576 * SE, \bar{x} + 2.576 * SE) = (0.133 - 2.576 * 0.028, 0.133 + 2.576 * 0.028) = (0.061{,}0.205).$$

<u>ESERCIZIO 4 (7 punti)</u>

Si vuole analizzare, mediante un modello di regressione lineare, la dipendenza del tasso di riempimento dell'aereo sui voli effettuati su una certa tratta dal prezzo medio al quale è venduto il biglietto in classe economica e dalla congiuntura economica nel periodo in cui è effettuato il volo, misurata attraverso un indicatore numerico da 0 (che indica congiuntura economica molto negativa) fino a 100 (congiuntura economica molto positiva). Sulla base di un campione di voli, in ciascuno dei quali si rilevano le tre variabili (tasso di riempimento, prezzo medio del biglietto e congiuntura economica), si ottiene il seguente output excel:

ANOVA

	gdl	*SQ*	*MQ*	*F*	*Sig. F*
Regressione	2	1708.0237	854.0119	28.2929	3.2175E-34
Residuo	77	2324.2245	30.1847		
Total	79	4032.2482			

	Coefficienti	*Errore Standard*	*Stat t*	*Sig. P-value*
Intercetta	66.449	12.044	5.5172	3.2102E-07
Prezzo medio biglietto	-0.0523	0.0184	-2.8424	0.0047232
Congiuntura economica	0.0181	0.0021	8.6190	3.4536E-34

a) E' significativo, in questo modello, l'effetto della congiuntura economica sul tasso di riempimento? Si giustifichi la risposta precisando gli elementi fondamentali della procedura adottata.

b) Se la risposta alla domanda precedente è affermativa, si descriva l'effetto.

c) Si preveda il tasso di riempimento di un aereo per un volo con prezzo medio del biglietto pari a 200 effettuato in un periodo di congiuntura economica con valore dell'indicatore pari a 60.

a) Sì, l'effetto è significativo in quanto il p-value del test di significatività della variabile è 3.4536*10^(-34), quindi minore di ogni livello standard. Quindi si rifiuta l'ipotesi nulla $H_0: \beta_{congiuntura\ economica} = 0$ e, quindi, si ritiene che tale coefficiente sia diverso da 0, per cui la congiuntura economica ha effetto sul tasso di riempimento.

b) La stima del coefficiente $\beta_{congiuntura\ economica}$ è 0.0181; quindi, ad un incremento unitario del livello di congiuntura economica, corrisponde un incremento del tasso medio di riempimento pari a 0.0181, fissato il prezzo medio del biglietto.

c) La previsione è: 66.449-0.0523*200+0.0181*60=57.075.