

INTERVALLO DI CONFIDENZA PER LA DIFFERENZA TRA DUE PROPORZIONI

$$(\rho_1 - \rho_2) \pm Z_{\alpha/2} \sqrt{\frac{\rho_1(1-\rho_1)}{n_1} + \frac{\rho_2(1-\rho_2)}{n_2}} \quad (10.6)$$

oppure:

$$\begin{aligned} (\rho_1 - \rho_2) \pm Z_{\alpha/2} \sqrt{\frac{\rho_1(1-\rho_1)}{n_1} + \frac{\rho_2(1-\rho_2)}{n_2}} &\leq (\pi_1 - \pi_2) \\ &\leq (\rho_1 - \rho_2) + Z_{\alpha/2} \sqrt{\frac{\rho_1(1-\rho_1)}{n_1} + \frac{\rho_2(1-\rho_2)}{n_2}} \end{aligned}$$

Ritornando all'esempio della soddisfazione dei clienti per i due hotel nell'isola dei tropici, utilizzando i risultati del test Z per il confronto tra le proporzioni:

$$p_1 = \frac{X_1}{n_1} = \frac{163}{227} = 0.7181 \quad p_2 = \frac{X_2}{n_2} = \frac{154}{262} = 0.5878$$

e l'Equazione (10.6), si ottiene:

$$\begin{aligned} & (0.7181 - 0.5878) \pm (1.96) \sqrt{\frac{0.7181(1-0.7181)}{227} + \frac{0.5878(1-0.5878)}{262}} \\ & 0.1303 \pm (1.96)(0.0426) \\ & 0.1303 \pm 0.0835 \\ & 0.0468 \leq (\pi_1 - \pi_2) \leq 0.2138 \end{aligned}$$

In definitiva, con un livello di confidenza del 95%, la differenza tra le proporzioni di clienti che ritornerebbero in uno dei due hotel è compresa tra 0.0468 e 0.2138. Dato che la differenza è sempre positiva, risulta che la proporzione di clienti che ritornerebbero presso l'hotel 1 è maggiore di quella riscontrata per l'hotel 2.

Esercizi al Paragrafo 10.3

Apprendere i concetti di base

- 10.25** Siano $n_1 = 100$, $X_1 = 50$, $n_2 = 100$, $X_2 = 30$.
- Con un livello di significatività di 0,05, verificare se vi è una differenza significativa tra le due proporzioni.
 - Costruire un intervallo di confidenza al 95% per stimare la differenza tra le proporzioni.
- 10.26** Siano $n_1 = 100$, $X_1 = 45$, $n_2 = 50$, $X_2 = 25$.
- Con un livello di significatività di 0.01, verificare se vi è una differenza significativa tra le due proporzioni.
 - Costruire un intervallo di confidenza al 99% per stimare la differenza tra le proporzioni

Applicazioni

- 10.27** Un campione di 500 soggetti è stato selezionato casualmente per una ricerca di mercato. È emerso che 136 maschi su 240 amano fare shopping e così 224 femmine su 260.
- Si vuole determinare se, con un livello di significatività di 0.01, vi è una significativa differenza tra la proporzione di maschi e femmine che amano fare shopping.
 - Calcolare il *p-value* e interpretarne il significato.
 - Costruire un intervallo di confidenza al 99% per la differenza tra le due proporzioni e interpretarne il significato.

d. Quale sarebbe stata la risposta ai punti (a) e (c) se il numero di maschi che ha dichiarato di amare lo shopping fosse stato di 206 anziché di 136?

10.28 Uno studio condotto da alcuni ricercatori del MIT (Massachusetts Institute of Technology) aveva lo scopo di verificare se per alcune persone, anche quando si tratta di placebo, i medicinali più costosi appaiono essere più efficaci di quelli economici. Un campione di 82 volontari è stato suddiviso in due gruppi. Tutti i volontari erano stati informati che si stava effettuando un test su un nuovo antidolorifico. Al primo gruppo è stato comunicato che il nuovo farmaco aveva un costo \$2.50 a pillola, mentre al secondo è stato indicato un prezzo di soli \$0.1 a pillola. In realtà le pillole in questione non sono nient'altro che zucchero. Per testare l'efficacia dell'antidolorifico, a tutti i volontari veniva data una leggera scossa elettrica che poi veniva ripetuta dopo l'assunzione del farmaco. Dei 41 soggetti che avevano assunto le pillole più costose, 35 avevano dichiarato di aver sentito meno dolore per la scossa successiva all'assunzione del farmaco. Degli altri 41 volontari che avevano assunto le pillole meno costose erano stati 25 ad aver dichiarato di aver sentito meno dolore alla seconda scossa (fonte: R. Rubin, Placebo Study Tests "Costlier is Better" Notion, *usatoday.com*, 5 marzo, 2008).

- a. Formulare l'ipotesi nulla e quella alternativa per verificare l'affermazione dei ricercatori ovvero che le persone credono che i farmaci più costosi siano più efficaci di quelli economici.
- b. Testare le ipotesi formulate nel punto (a) con un livello di significatività di 0.05.
- c. I risultati del test al punto (b) sono in linea con l'ipotesi dei ricercatori?

10.29 Il senso comune ritiene che le donne siano più avverse al rischio rispetto agli uomini. Alcuni studiosi hanno misurato la propensione al rischio attraverso i comportamenti che i soggetti mostrano nel mercato azionario. A un campione di donne e uomini è stato chiesto di scegliere tra due tipi di alternative di investimento: la prima (alternativa 1) connessa a una quantità di rischio più alta della seconda (alternativa 2). È emerso che 368 donne su 965 e 554 uomini su 1066 hanno optato per l'alternativa 1, ovvero per lo scenario più rischioso.

- a. Formulare l'ipotesi nulla e quella alternativa per verificare se le donne sono più avverse al rischio rispetto agli uomini, sulla base dell'esperimento condotto.
- b. Testare le ipotesi formulate nel punto (a) con un livello di significatività di 0.05.
- c. Calcolare il *p-value* e interpretarne il significato.

10.30 Da un'indagine è emerso che alcune caratteristiche socio-demografiche influenzano la percezione che gli italiani hanno della pressione fiscale. In generale è emerso che il 60% di essi ritiene che la pressione fiscale sia troppo elevata mentre il 40% l'ha definita giusta. Un gruppo di ricercatori vuole verificare se il livello di reddito influisca su tale percezione. Un campione di 1005 intervistati è stato suddiviso in due gruppi: il primo ha un reddito annuo inferiore a 50000 euro e il secondo superiore a 50000 euro annui. La seguente tabella di contingenza riporta i risultati dell'indagine sui due gruppi.

Pressione fiscale	Reddito annuo		Totale
	Minore di 50 000 euro	Maggiore di 50 000 euro	
Giusta	225	180	405
Troppo elevata	280	320	600
Totale	505	500	1005

a. Vi è differenza sulla percezione della pressione fiscale tra i due gruppi di reddito ($\alpha = 0.05$)?

b. Calcolare il *p-value* e interpretarne il significato.

10.31 I responsabili del controllo di qualità di una ditta che produce microprocessori hanno effettuato un'indagine per individuare le cause della bassa qualità di alcuni microprocessori. L'indagine si basa su un campione casuale di 450 microprocessori. La seguente tabella di contingenza riporta i dati incrociati per le due variabili: "Vi sono delle impurità sul wafer di silicio?" e "Qualità del microprocessore":

Presenza di impurità	Qualità del microprocessore		Totale
	Alta	Bassa	
Sì	14	36	50
No	320	80	400
Totale	334	116	450

Fonte: S.W. Hall, Analysis of Defectivity of Semiconductor Wafers by Contingency Table, *Proceedings Institute of Environmental Sciences*, Vol. 1, 1994, pp. 177-183.

- a. Vi è evidenza empirica che vi sia una significativa differenza nella qualità (alta o bassa) del microprocessore in base alla presenza di impurità sul wafer di silicio ($\alpha = 0.05$)?
- b. Calcolare il *p-value* e interpretarne il significato.
- c. Costruire un intervallo di confidenza al 95% per la differenza tra le due proporzioni e interpretarne il significato.
- d. Quali conclusioni dovrebbero trarre i responsabili del controllo di qualità dai risultati emersi?