

```
. ****
. *** Esercizio 1 cap. 3 ***
. *** soluzione ***
. ****
.
. /* Per verificare quanto tempo impiega il poliestere a deteriorarsi in una
discarica,
> un ricercatore ha preso 10 strisce di poliestere e le ha interrate. Cinque
strisce
> scelte a caso sono state estratte dopo due settimane e le restanti dopo 3
mesi. Su
> ogni striscia è stata misurata la forza di rottura in libbre. Sono state
ottenute le
> statistiche seguenti.
>
> Gruppo          n    media    deviazione standard
> 1 (2 settimane)  5    123.8      4.60
> 2 (3 mesi)       5    116.4     16.09
>
> a) Costruire un intervallo di confidenza al 95% per la differenza delle medie
> assumendo la normalità e l'omoschedasticità.
> b) Sottoporre a test l'ipotesi di uguaglianza delle medie. */
.
. *IC e test t per l'uguaglianza delle medie
. *campioni INDIPENDENTI, stessa varianza
. ttesti 5 123.8 4.60 5 116.4 16.09
```

	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
x	5	123.8	2.057183	4.6	118.0883	129.5117
y	5	116.4	7.195667	16.09	96.42163	136.3784
combined	10	120.1	3.737338	11.8185	111.6456	128.5544
diff		7.4	7.483958		-9.858037	24.65804

Ha: diff < 0	Ha: diff != 0	Ha: diff > 0
Pr(T < t) = 0.8241	Pr(T > t) = 0.3517	Pr(T > t) = 0.1759

```
. /* I dati seguenti rappresentano i pesi di 20 uomini partecipanti a un
programma di
> perdita di peso, prima e dopo il programma.
>
> Prima 100.8 102.0 105.9 108.0 92.0 116.7 110.2 135.0 123.5 95.0 105.0 85.0
107.2 80.0 115.1 103.5 82.0 101.5 103.5 93.0
>
```

```

HW2.log
> Dopo 97.0 107.5 97.0 108.0 84.0 111.5 102.5 127.5 118.5 94.2 105.0 82.4 98.2
83.6 115.0 103.0 80.0 101.5 102.6 93.0
>
> Vogliamo sapere se in media il programma ha avuto effetto.
> Fare un test opportuno esplicitando le assunzioni. */
.
. *Siamo di fronte a due gruppi di campioni appaiati,
. *in quanto le misurazioni sono state effettuate sugli stessi individui prima
e dopo la dieta
.
. *inserire i dati (vedere pesi.csv)
. * y1 peso PRIMA della dieta
. * y2 peso DOPO la dieta
.
. import delimited "Y:\carla\modelli_triennale\esempi_Stata\pesi.csv",
delimiter(" ", collapse) clear
(2 vars, 20 obs)

.
. *costruire la variabile differenza
. gen diff=y1-y2

.
. summ diff

```

Variable	Obs	Mean	Std. Dev.	Min	Max
diff	20	2.645	4.116651	-5.5	9

```

.
. *Per le assunzioni e statistica test vedere slides 12-13 in Modelli2017_04 e
esempio in appaiati.do
. *Per verificare se c'è stata una riduzione del peso o se il peso è rimasto in
media sostanzialmente invariato (ipotesi H0),
. *dobbiamo effettuare un test t di student per campioni appaiati
.
. *usando diff
. ttest diff == 0

```

One-sample t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
diff	20	2.645	.9205111	4.116651	.7183482 4.571652

```

mean = mean(diff)
Ho: mean = 0
Ha: mean < 0
Pr(T < t) = 0.9951
t = 2.8734
degrees of freedom = 19
Ha: mean != 0
Pr(|T| > |t|) = 0.0097
Ha: mean > 0
Pr(T > t) = 0.0049

```

```

.
. *oppure usando i dati originali
. ttest y1 == y2

```

Paired t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
y1	20	103.245	3.02256	13.5173	96.91871 109.5713
y2	20	100.6	2.789388	12.47452	94.76174 106.4383
diff	20	2.645	.9205111	4.116651	.7183482 4.571652

```

mean(diff) = mean(y1 - y2)
Ho: mean(diff) = 0
Ha: mean(diff) < 0
t = 2.8734
degrees of freedom = 19
Ha: mean(diff) != 0
Ha: mean(diff) > 0

```

```

                                HW2.log
Pr(T < t) = 0.9951              Pr(|T| > |t|) = 0.0097              Pr(T > t) = 0.0049

. *poichè il p-value è <0.01 concludiamo dicendo che la dieta è risultata
efficace
. *Pr(T < t) = 0.0049
.
.
. *****
. *** Esercizio 4 cap. 3 ***
. *** soluzione ***
. *****
.
. /* In un ospedale 485 pazienti sono stati sottoposti a due tipi di chirurgia,
ginecologica
> (Gruppo 1) e addominale (Gruppo 2), ed è stata registrata la presenza
> successiva di complicazioni postoperatorie. I dati sono i seguenti
>
>
> Operazione      Complicazioni
> Ginecologica    No Sì Totale % Complicazioni
> Addominale      235 5 240 2.1
> Totale          210 35 245 14.3
>
> Costruire un intervallo di confidenza per la differenza delle proporzioni.
> Sottoporre a test che non ci siano differenze tra i due gruppi. */
.
.
. *** I dati sono stati ottenuti con un esperimento controllato ***
. *** con assegnazione assimilabile alla randomizzazione completa ***
.
. *** La struttura dell'esperimento permette di ipotizzare che ***
. *** i dati siano due campioni indipendenti da due popolazioni di Bernoulli
***
.
.
. *Confronto di due proporzioni - abbiamo visto esempio in salk.do, formule in
slides modelli2015_06 (da pag.14 in poi)
.
. *** IC e test approssimato per la differenza tra proporzioni ***
. *** modello di Bernoulli, campioni indipendenti ***
.
. *indicare: n_1, r_1, n_2, r_2
. prtesti 240 5 245 35 , count

Two-sample test of proportions                                x: Number of obs = 240
                                                            y: Number of obs = 245
-----
|          |          Mean      Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
|          |          .0208333    .0092194          |          .0027637    .038903
| x         |          .1428571    .022356          |          .0990401    .1866741
| y         |          +-----+
| diff      |          -.1220238    .0241824          |          -.1694205    -.0746272
|          |          under Ho:    .0249833          |          -4.88    0.000
-----+-----
|          |          diff = prop(x) - prop(y)          |          z = -4.8842
|          |          Ho: diff = 0
|          |          Ha: diff < 0          |          Ha: diff != 0          |          Ha: diff > 0
|          |          Pr(Z < z) = 0.0000          |          Pr(|Z| > |z|) = 0.0000          |          Pr(Z > z) = 1.0000
.
. *poichè Pr(|Z| < |z|) = 0.0000
. *p-value<0.001 ==> differenza tra proporzioni significativamente diversa da
zero!
.
.
. *****

```

```

HW2.log
. *** soluzione esrcizi 11-14 capitolo 3 ***
. *****
.
. /* con riferimento agli esercizi 11, 12 e 13 calcolare il rischio relativo e
il rapporto delle quote.*/
.
. *Il calcolo fatto è appropriato per gli esrcizi 11 e 12, ma non per
l'esercizio 13 (studio caso-controllo), salvo sotto l'ipotesi che
. * la probabilità di sviluppare la malattia sia piccola
.
.
. *Nell'esercizio 11 i dati sono stati ottenuti con un esperimento controllato
con assegnazione randomizzata.
. *La struttura dell'esperimento permette di ipotizzare che i dati siano due
campioni indipendenti da due popolazioni di Bernoulli
. /*
>
> Gruppo      Infarto
> Gruppo      Si      No      Totale
> Placebo      189    10845    11034
> Aspirina      104    10933    11037
>
> */
.
. *Rilegge dati: Tabella con i dati vista sopra
. use "Y:\carla\modelli_triennale\esempi_Stata\ese11_ch3.dta", clear
.
. *tabella di frequenza
. table gruppo malato [weight=freq], row col
(frequency weights assumed)

-----
      Gruppo |      malato
              |      0      1      Total
-----+-----
Aspirina     | 10,933      104    11,037
Placebo      | 10,845      189    11,034
Total        | 21,778      293    22,071
-----

.
. *probabilità di malattia
. table gruppo [weight=freq], row col contents(mean malato)
(frequency weights assumed)

-----
      Gruppo | mean(malato)
-----+-----
Aspirina     |      .009423
Placebo      |      .017129
Total        |      .013275
-----

.
.
. *creo variabile gruppo come variabile numerica per usarla con tabodds
. gen group=0
.
.
. * Costruire un intervallo di confidenza per l'OR e dire quando si può
applicare approssimativamente al rischio relativo.

```

```

HW2.log
. *calcolo RR, OR e IC per odds ratio
. cs malato group [fw=freq], or

```

	group Exposed	Unexposed	Total
Cases	0	293	293
Noncases	0	21778	21778
Total	0	22071	22071
Risk	.	.0132753	.0132753
	Point estimate		[95% Conf. Interval]
Risk difference	.	.	.
Risk ratio	.	.	.
Attr. frac. ex.	.	.	.
Attr. frac. pop	.	.	.
Odds ratio	.	.	(Cornfield)

```

+-----+
chi2(1) = . Pr>chi2 = .

```

```

.
. *calcolo odds e IC per odds
. tabodds malato group [fw=freq]

```

group	cases	controls	odds	[95% Conf. Interval]	
0	293	21778	0.01345	0.01199	0.01510

```

Test of homogeneity (equal odds): chi2(0) = 0.00
Pr>chi2 = .

```

```

Score test for trend of odds: chi2(1) = .
Pr>chi2 = .

```

```

.
. *calcolo odds ratio e IC per OR
. tabodds malato group[fw=freq], or base(1)

```

group	Odds Ratio	chi2	P>chi2	[95% Conf. Interval]	
0	1.000000

```

Test of homogeneity (equal odds): chi2(0) = 0.00
Pr>chi2 = .

```

```

Score test for trend of odds: chi2(1) = .
Pr>chi2 = .

```

```

.
. *Fornire il p-value del test di uguaglianza delle due proporzioni usando come
parametro il logaritmo dell'OR.
. logit malato group [fw=freq]

```

```

note: group omitted because of collinearity
Iteration 0: log likelihood = -1557.3477
Iteration 1: log likelihood = -1557.3477

```

Logistic regression	Number of obs	=	22,071
	LR chi2(0)	=	-0.00
	Prob > chi2	=	.
Log likelihood = -1557.3477	Pseudo R2	=	-0.0000

HW2.log

malato	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
group	0	(omitted)				
_cons	-4.308483	.0588123	-73.26	0.000	-4.423753	-4.193213

```
. log close
  name: <unnamed>
  log: D:\Dati\corsi\modelli_triennale\Homework\HW2.log
  log type: text
  closed on: 14 Oct 2019, 09:20:03
```