



UNIVERSITA' DEGLI STUDI DI VERONA

# LABORATORIO DI PROBABILITA' E STATISTICA

Docente: Bruno Gobbi

Corso di laurea in Informatica e Bioinformatica

## 8 - VERIFICA DI IPOTESI

# LA VERIFICA DI IPOTESI

- ▶ Una delle applicazioni più comuni della statistica è quello della verifica di ipotesi.
- ▶ Si vuole verificare se una data ipotesi è vera oppure no, sulla base dei dati provenienti da un campione.
- ▶ Il caso più classico è la verifica di ipotesi sulla media, ovvero si vuole verificare l'ipotesi che la media di un dato campione sia uguale a un certo valore oppure no.

# LA VERIFICA DI IPOTESI

Dato un campione estratto da una variabile aleatoria normale  $X$  con media  $\mu_0$  e varianza  $\sigma^2$ , la verifica dell'ipotesi che la media del campione  $\bar{X}$  sia pari a  $\mu_0$  si effettua tramite la statistica test:

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}}$$

# LA VERIFICA DI IPOTESI

```
t.test(x, y=NULL, alternative="two.sided", mu=0, paired=FALSE,  
      var.equal=FALSE, conf.level=0.95)
```

- ▶ **x** è un vettore di dati
- ▶ **y** è un secondo vettore (opzionale) di dati, da specificare quando si vuole effettuare il confronto tra due campioni
- ▶ **alternative** specifica l'ipotesi alternativa. Può essere:
  - "two.sided" (opzione di default), quando l'ipotesi alternativa è che la media (o la differenza tra medie) sia diversa da quella fissata nell'ipotesi nulla
  - "less", quando è minore
  - "greater", quando è maggiore
- ▶ **mu** è l'ipotesi nulla sul valore della media (test su un campione), o sulla differenza tra medie (test su due campioni). Di default  $\mu = 0$ .
- ▶ **conf.level** è il livello di confidenza dell'intervallo, ovvero il complemento ad 1 della probabilità dell'errore di prima specie fissato, di default posto uguale a 0.95

Gli argomenti **paired** e **var.equal** verranno commentati successivamente.

# LA VERIFICA DI IPOTESI

ARGOMENTI TRATTATI NELLA LEZIONE DI OGGI:

1. VERIFICA DI IPOTESI SULLA MEDIA CON TEST BILATERALE (UGUALE O DIVERSO DALLA MEDIA)
2. INTERVALLO DI CONFIDENZA PER LA MEDIA
3. LA VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE (MAGGIORE O MINORE DELLA MEDIA)
4. LA VERIFICA DI IPOTESI SUL CONFRONTO FRA DUE MEDIE (PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE POPOLAZIONI UGUALI
5. LA VERIFICA DI IPOTESI SUL CONFRONTO FRA DUE MEDIE (PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE POPOLAZIONI DIVERSE
6. LA VERIFICA DI IPOTESI PER DATI APPAIATI

# 1 - VERIFICA DI IPOTESI SULLA MEDIA CON TEST BILATERALE

Sono stati rilevati i QI di 10 studenti di Informatica. Si vuole verificare l'ipotesi che gli studenti di Informatica abbiano un QI in media di 110. Il tutto a un livello di significatività dell'1% ovvero un livello di confidenza teorico del 99% (100-1).

Il test è bilaterale perché  $H_0$  è  $\mu=110$  e l'alternativa  $\mu \neq 110$

Ecco i dati rilevati dal campione di 10 studenti:

```
> qi=c(109, 99, 100, 120, 95, 123, 121, 121, 115, 109)
```

# 1 - VERIFICA DI IPOTESI SULLA MEDIA

**# FACCIAMO INNANZITUTTO UNA PRIMA RAPIDA ANALISI DELLE PRINCIPALI STATISTICHE:**

```
> summary(qi)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
95.0	102.2	112.0	111.2	120.8	123.0

```
> sd(qi)
```

```
[1] 10.37947
```

# 1 - VERIFICA DI IPOTESI SULLA MEDIA

# VERIFICHIAMO ORA SE, SULLA BASE DI QUESTO CAMPIONE, SI PUO' RITENERE VERA L'IPOTESI SECONDO CUI IN MEDIA IL QI MEDIO DEGLI STUDENTI DI INFORMATICA E' UGUALE A 110 (QUINDI BEN SUPERIORE AL VALORE NORMALE DI 100), AL LIVELLO DI CONFIDENZA (CONF.LEVEL) DEL 99%

```
t.test(qi, alternative="two.sided", mu=110, conf.level=0.99)
```

**H0:  $\mu=110$  H1:  $\mu \neq 110$**

# 1 - VERIFICA DI IPOTESI SULLA MEDIA

**SE**      LIVELLO DI  
SIGNIFICATIVITA'  
ALPHA TEORICO      **<**      P-VALUE  
CALCOLATO

**→ SI ACCETTA L'IPOTESI NULLA H0  
DI UGUAGLIANZA**

**H0:  $\mu=110$       H1:  $\mu \neq 110$**

Il livello di significatività alpha teorico viene indicato nel testo dell'esercizio, mentre il p-value lo si ricava dalla funzione t.test

# 1 - VERIFICA DI IPOTESI SULLA MEDIA

```
> t.test(qi, alternative="two.sided", mu=110, conf.level=0.99)
```

One Sample t-test

data: qi

t = 0.3656, df = 9, p-value = 0.7231

alternative hypothesis: true mean is not equal to 110

99 percent confidence interval:

100.5331 121.8669

sample estimates:

mean of x

111.2

**# POICHE' IL LIVELLO DI SIGNIFICATIVITA' ALPHA TEORICO DELLO 0.01 (100-99=1%) E' INFERIORE AL P-VALUE, ACCETTO L'IPOTESI NULLA E CONSIDERO VALIDA L'IPOTESI CHE GLI STUDENTI DI INFORMATICA ABBIAMO UN QI SUPERIORE A QUELLO NORMALE.**

# 1 - VERIFICA DI IPOTESI SULLA MEDIA

**EXTRA - NON VERRA' CHIESTO ALL'ESAME**

# IN ALTERNATIVA SI PUO' VERIFICARE IL TEST CONFRONTANDO IL VALORE DELLA STATISTICA  $t$  CON I VALORI CRITICI, O SOGLIE, DELLA DISTRIBUZIONE  $t$  DI STUDENT, AL LIVELLO DI SIGNIFICATIVITA' PREFISSATO.

PER RICAVARE QUESTI VALORI SI USA LA FUNZIONE INVERSA DELLA  $T$  DI STUDENT, OSSIA  $qt(p, df)$ , DOVE  $p$  E' IL LIVELLO DI SIGNIFICATIVITA' TEORICO, NEL CASO DI IPOTESI SULLA MEDIA DA DIVIDERE PER DUE, MENTRE  $df$  SONO I GRADI DI LIBERTA' PARI A  $n-1$  (IN QUESTO CASO  $10-1=9$ )

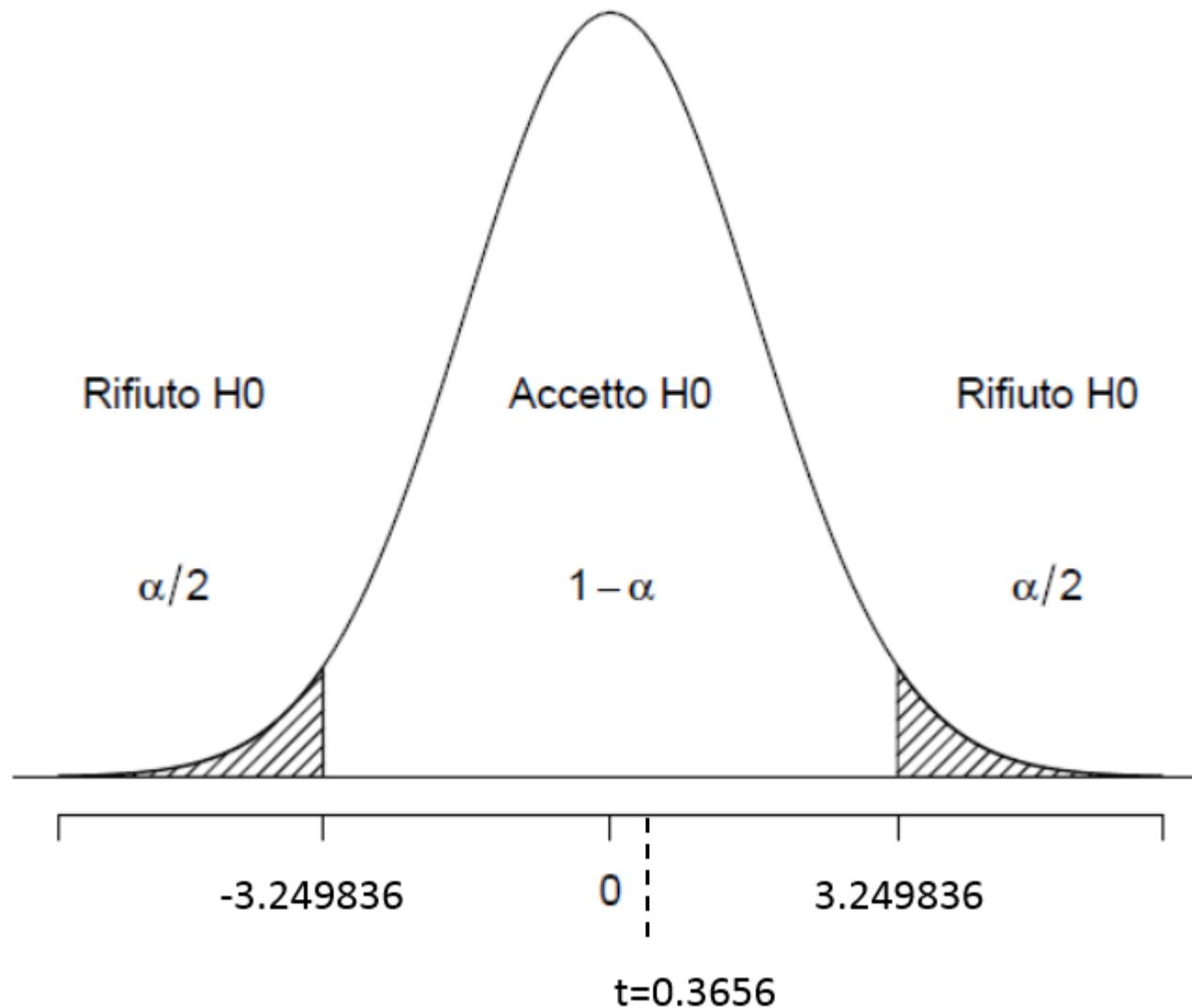
```
> qt(0.01/2, 9, lower.tail=TRUE)
```

```
[1] -3.249836
```

```
> qt(0.01/2, 9, lower.tail=FALSE)
```

```
[1] 3.249836
```

# 1 - VERIFICA DI IPOTESI SULLA MEDIA



# POICHE' IL VALORE DI  $t$  OSSERVATO (0.3656) RICADE NELL'INTERVALLO -3.249836 E +3.249836 SI ACCETTA L'IPOTESI NULLA.

## 2 - INTERVALLO DI CONFIDENZA PER LA MEDIA

NOTA: NELL'OUTPUT DELLA FUNZIONE t.test E' PRESENTE ANCHE L'INTERVALLO DI CONFIDENZA DELLA MEDIA, OSSIA AL 99% LA MEDIA DELLA POPOLAZIONE SI TROVERA' FRA I VALORI 100.5331 E 121.8669

data: qi

t = 0.3656, df = 9, p-value = 0.7231

alternative hypothesis: true mean is not equal to 110

**99 percent confidence interval:**

**100.5331 121.8669**

sample estimates:

mean of x

111.2

# ESERCIZIO SU VERIFICA DI IPOTESI SULLA MEDIA

IL NUMERO MEDIO DI ORE DI SONNO RILEVATE PER UN MESE SU UN CAMPIONE DI 10 PERSONE E' RISULTATO PARI AL SEGUENTE VETTORE:

sonno=c(7, 6.5, 7.5, 6, 5, 4.5, 8, 7, 7, 6.5)

VERIFICARE L'IPOTESI CHE IL NUMERO MEDIO DI ORE DI SONNO SIA PARI A 7 AL LIVELLO DI CONFIDENZA DEL 99%.

INDICARE ANCHE L'INTERVALLO DI CONFIDENZA PER LA MEDIA.

IPOTESI DA VALUTARE  
 $H_0: \mu=7$     $H_1: \mu \neq 7$

# ESERCIZIO SU VERIFICA DI IPOTESI SULLA MEDIA

```
> t.test(sonno, mu=7, alternative="two.sided",  
conf.level=0.99)
```

One Sample t-test

data: sonno

t = -1.4639, df = 9, p-value = 0.1773

alternative hypothesis: true mean is not equal to 7

99 percent confidence interval:

5.38997 7.61003

sample estimates:

mean of x

6.5

# ESERCIZIO SU VERIFICA DI IPOTESI SULLA MEDIA

```
> t.test(sonno, mu=7, alternative="two.sided", conf.level=0.99)
```

One Sample t-test

data: sonno

t = -1.4639, df = 9, **p-value = 0.1773**

alternative hypothesis: true mean is not equal to 7

99 percent confidence interval:

5.38997 7.61003

sample estimates:

mean of x

6.5

**# IL LIVELLO DI SIGNIFICATIVITA' TEORICO ALPHA E' PARI A 1%  
(100%-99%, OVVERO 0.01)**

**IL P-VALUE CALCOLATO (CHE E' UGUALE A 0.1773),**

**POICHE' IL LIVELLO DI SIGNIFICATIVITA' < P-VALUE CALCOLATO**

**→ SI ACCETTA L'IPOTESI NULLA**

# ESERCIZIO SU VERIFICA DI IPOTESI SULLA MEDIA

```
> t.test(sonno, mu=7, alternative="two.sided", conf.level=0.99)
```

One Sample t-test

data: sonno

t = -1.4639, df = 9, p-value = 0.1773

alternative hypothesis: true mean is not equal to 7

**99 percent confidence interval:**

**5.38997 7.61003**

sample estimates:

mean of x

6.5

**# NOTA: NELL'OUTPUT DELLA FUNZIONE t.test E' PRESENTE ANCHE L'INTERVALLO DI CONFIDENZA DELLA MEDIA, OSSIA AL 99% LA MEDIA DELLA POPOLAZIONE SI TROVERA' FRA I VALORI 5.38997 E 7.61003**

# 3 - VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

NEL CASO IN CUI VOLESSIMO NON PIU' VALUTARE SE LA MEDIA E' UGUALE O MENO AD UN CERTO VALORE, MA VOLESSIMO VERIFICARE SE LA MEDIA E' MAGGIORE O MINORE DI UN CERTO VALORE, ALLORA USIAMO LO STESSO TEST PRECEDENTE, MA UTILIZZANDO LE OPZIONI alternative="less" O alternative="greater" ALL'INTERNO DEL t.test

# 3 - VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

VOGLIAMO VERIFICARE L'IPOTESI CHE IL QI MEDIO DEL CAMPIONE DI 10 STUDENTI DI INFORMATICA PRECEDENTE SIA MAGGIORE DI 120. QUINDI NON SEMPLICEMENTE CHE SIA DIVERSA, MA CHE SIA MAGGIORE DI 120.

IL LIVELLO DI SIGNIFICATIVITA' E' SEMPRE ALL'1%.

$$H_0: \mu=120 \quad H_1: \mu>120$$

# 3 - VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

$H_0: \mu=120$     $H_1: \mu>120$

```
> t.test(qi, mu=120, conf.level=0.99, alternative="greater")
```

# 3 - VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

```
> t.test(qi, mu=120, conf.level=0.99, alternative="greater")
```

One Sample t-test

data: qi

t = -2.6811, df = 9, p-value = 0.9874

alternative hypothesis: true mean is greater than 120

99 percent confidence interval:

101.9393      Inf  
sample estimates:

mean of x

111.2

**# IN QUESTO CASO IL VALORE ALPHA DI SIGNIFICATIVITA' TEORICA E' 0.01 (1%),  
MENTRE IL p-value CALCOLATO E' 0.9874. POICHE' L'ALPHA E' MINORE DEL p-  
value, SI ACCETTA L'IPOTESI NULLA E SI RIFIUTA L'IPOTESI ALTERNATIVA CHE  
PREVEDEVA CHE IL QI MEDIO DEGLI STUDENTI DI INFORMATICA FOSSE  
SUPERIORE A 120, CON UN LIVELLO DI CONFIDENZA DEL 99%.**

# ES. VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

RIPRENDENDO I DATI DELL'ESERCIZIO PRECEDENTE SULLE ORE DI SONNO, VALUTARE L'IPOTESI CHE IL NUMERO MEDIO DI ORE DI SONNO SIA SUPERIORE A 7 A UN LIVELLO DI CONFIDENZA DEL 95%.

sonno=c(7, 6.5, 7.5, 6, 5, 4.5, 8, 7, 7, 6.5)

# ES. VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

$H_0: \mu=7$   $H_1: \mu>7$

```
t.test(sonno, mu=7, conf.level=0.95, alternative="greater")
```

# ES. VERIFICA DI IPOTESI SULLA MEDIA CON TEST UNILATERALE

```
> t.test(sonno, mu=7, conf.level=0.95, alternative="greater")
```

One Sample t-test

data: sonno

t = -1.4639, df = 9, p-value = 0.9114

alternative hypothesis: true mean is greater than 7

95 percent confidence interval:

5.873873      Inf  
sample estimates:

mean of x

6.5

**# POICHE' IL LIVELLO DI SIGNIFICATIVITA' TEORICO ALPHA E' PARI A 5% (100%-95%), OVVERO 0.05, MENTRE IL P-VALUE CALCOLATO E' UGUALE A 0.9114, SI ACCETTA L'IPOTESI NULLA CHE LA MEDIA SIA 7 E SI RIFIUTA L'IPOTESI ALTERNATIVA CHE QUESTA SIA MAGGIORE DI 7.**

# 4 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

LA VERIFICA DI IPOTESI PUO' ESSERE CONDOTTA ANCHE PER  
CONFRONTARE LE MEDIE DI DUE CAMPIONI INDIPENDENTI.

SI DISTINGUONO DUE SOTTOCASI:

- ▶ VARIANZE DELLE POPOLAZIONI UGUALI
- ▶ VARIANZE DELLE POPOLAZIONI DIVERSE

VEDIAMO UN ESEMPIO DEL PRIMO CASO, QUANDO LE  
VARIANZE DELLE POPOLAZIONI OGGETTO DI STUDIO SONO  
UGUALI.

# 4 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

I POTIZZIAMO DI AVER RILEVATO IL QI DI 5 STUDENTI E DI 5  
STUDENTESSE DI BIOINFORMATICA E DI VOLER VERIFICARE L'IPOTESI  
CHE IL QI DELLE FEMMINE SIA DIVERSO DA QUELLO DEI MASCHI. SI  
IPOTIZZA CHE LA VARIANZA DELLE DUE POPOLAZIONI SIA UGUALE. IL  
LIVELLO DI CONFIDENZA E' DEL 99%

**# COSTRUIAMO I DUE CAMPIONI IN QUESTO MODO:**

qi.m = c(105, 99, 115, 130, 80)

qi.f = c(103, 104, 115, 118, 106)

$H_0: M(QI_f) = M(QI_m)$        $H_1: M(QI_f) \neq M(QI_m)$

# 4 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

$$H_0: M(QI_f) = M(QI_m) \quad H_1: M(QI_f) \neq M(QI_m)$$

```
t.test(qi.f, qi.m, var.equal=TRUE, conf.level=0.99)
```

# 4 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

```
> t.test(qi.f, qi.m, var.equal=TRUE, conf.level=0.99)
```

Two Sample t-test

data: qi.f and qi.m

t = 0.3838, df = 8, p-value = 0.7111 **# I DF SONO (N1-1)+(N2-1)=N-2**

alternative hypothesis: true difference in means is not equal to 0

99 percent confidence interval:

-26.32502 33.12502

sample estimates:

mean of x mean of y

109.2 105.8

**# POICHE' L'ALPHA TEORICO (OSSIA IL LIVELLO DI SIGNIFICATIVITA') E' 0.01 ED E' INFERIORE AL P-VALUE CALCOLATO DI 0.7111, SI ACCETTA L'IPOTESI NULLA E SI CONSIDERANO UGUALI I QI DEI MASCHI E DELLE FEMMINE AD UN LIVELLO DI CONFIDENZA DEL 99%.**

# ES. CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

I POTIZIAMO DI AVER RILEVATO IL NUMERO DI KM PERCORSI CON UN LITRO DI BENZINA DA 5 AUDI A3 E 5 BMW SERIE1. SI IPOTIZZA CHE LA VARIANZA DELLE DUE POPOLAZIONI SIA UGUALE. IL LIVELLO DI CONFIDENZA E' DEL 95%. VERIFICARE SE SI PUO' DIRE CHE IN MEDIA L'AUDI E LA BMW CONSUMANO UGUALE.

**# COSTRUIAMO I DUE CAMPIONI IN QUESTO MODO:**

km.audi = c(20, 21, 15, 16, 17)

km.bmw = c(18, 13, 14, 19, 17)

$$H_0: M(KM_{\text{audi}}) = M(KM_{\text{bmw}}) \quad H_1: M(KM_{\text{audi}}) \neq M(KM_{\text{bmw}})$$

# ES. CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

$$H_0: M(KM_{\text{audi}}) = M(KM_{\text{bmw}})$$

$$H_1: M(KM_{\text{audi}}) \neq M(KM_{\text{bmw}})$$

```
t.test(km.audi, km.bmw, var.equal=TRUE,  
conf.level=0.95)
```

# ES. CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI UGUALI

```
> t.test(km.audi, km.bmw, var.equal=TRUE, conf.level=0.95)
```

Two Sample t-test

data: km.audi and km.bmw

t = 0.9774, df = 8, p-value = 0.357

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.175091 5.375091

sample estimates:

mean of x mean of y

17.8 16.2

**# POICHE' L'ALPHA TEORICO (OSSIA IL LIVELLO DI SIGNIFICATIVITA') E' 0.05 ED E' INFERIORE AL p-value CALCOLATO DI 0.357, SI ACCETTA L'IPOTESI NULLA E SI CONSIDERANO UGUALI I CONSUMI MEDI DI AUDI E BMW AD UN LIVELLO DI CONFIDENZA DEL 95%.**

# 5 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

RIPRENDIAMO L'ESERCIZIO PRECEDENTE SUL QI DEI 5 STUDENTI  
MASCHI E 5 FEMMINE E IPOTIZZIAMO STAVOLTA CHE LE VARIANZE  
DELLE POPOLAZIONI SIANO DIVERSE. IL LIVELLO DI CONFIDENZA E'  
DEL 99%

**# RIPRENDIAMO I DUE CAMPIONI IN QUESTO MODO:**

qi.m = c(105, 99, 115, 130, 80)

qi.f = c(103, 104, 115, 118, 106)

$H_0: M(QI_f) = M(QI_m)$        $H_1: M(QI_f) \neq M(QI_m)$

# 5 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

$$H_0: M(QI_f) = M(QI_m) \quad H_1: M(QI_f) \neq M(QI_m)$$

```
t.test(qi.f, qi.m, var.equal=FALSE, conf.level=0.99)
```

# 5 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

$$H_0: M(QI_f) = M(QI_m) \quad H_1: M(QI_f) \neq M(QI_m)$$

```
t.test(qi.f, qi.m, var.equal=FALSE, conf.level=0.99)
```

# 5 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

```
t.test(qi.f, qi.m, var.equal=FALSE, conf.level=0.99)
```

## Welch Two Sample t-test

data: qi.f and qi.m

t = 0.3838, df = 5.061, p-value = 0.7167

alternative hypothesis: true difference in means is not equal to 0

99 percent confidence interval:

-32.09857 38.89857

sample estimates:

mean of x mean of y

109.2 105.8

**# POICHE' L'ALPHA TEORICO (OSSIA IL LIVELLO DI SIGNIFICATIVITA') E' 0.01 ED E' INFERIORE AL p-value CALCOLATO DI 0.7167, SI ACCETTA L'IPOTESI NULLA E SI CONSIDERANO UGUALI I QI DEI MASCHI E DELLE FEMMINE AD UN LIVELLO DI CONFIDENZA DEL 99%.**

# 5 - CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

- ▶ RISPETTO AL CASO PRECEDENTE CON LE VARIANZE UGUALI, “R” UTILIZZA UNA FORMULA PER IL CALCOLO DELLA STATISTICA  $t$  DIVERSA RISPETTO AGLI ALTRI CASI, OSSIA IL **Welch Two Sample t-test**
- ▶ I GRADI DI LIBERTA' VENGONO CALCOLATI IN MODO PARTICOLARE, IN QUESTO CASO 5.061 E QUESTO E' DOVUTO AL MAGGIOR GRADO DI INCERTEZZA CHE SI HA QUANDO LE VARIANZE DELLE POPOLAZIONI NON SONO UGUALI.
- ▶ LA LETTURA DEI RISULTATI E' COMUNQUE SEMPRE LA STESSA
- ▶ IL P-VALUE E' LEGGERMENTE PIU' ELEVATO (DA 0.7111 A 0.7167)

# ES. CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

I POTIZZIAMO DI AVER RILEVATO IL PESO DI 5 STUDENTI DI INFORMATICA E 5 DI SCIENZE MOTORIE. SI IPOTIZZA CHE LA VARIANZA DELLE DUE POPOLAZIONI SIA DIVERSA. IL LIVELLO DI CONFIDENZA E' DEL 99%. VERIFICARE L'IPOTESI CHE LE MEDIE DELLE DUE POPOLAZIONI SIANO UGUALI.

peso.inf = c(64, 61, 120, 105, 58)

peso.mot = c(78, 72, 84, 77, 79)

$$H_0: M(P_{\text{inf}}) = M(P_{\text{mot}})$$

$$H_1: M(P_{\text{inf}}) \neq M(P_{\text{mot}})$$

# ES. CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

$$H_0: M(P_{\text{inf}}) = M(P_{\text{mot}})$$

$$H_1: M(P_{\text{inf}}) \neq M(P_{\text{mot}})$$

```
t.test(peso.inf, peso.mot, var.equal=FALSE, conf.level=0.99)
```

# ES. CONFRONTO FRA DUE MEDIE

(PER CAMPIONI INDIPENDENTI) CON VARIANZE DELLE  
POPOLAZIONI DIVERSE

```
> t.test(peso.inf, peso.mot, var.equal=FALSE, conf.level=0.99)
```

Welch Two Sample t-test

data: peso.inf and peso.mot

t = 0.2766, df = 4.179, p-value = 0.7952

alternative hypothesis: true difference in means is not equal to 0

99 percent confidence interval:

-54.60384 61.80384

sample estimates:

mean of x mean of y

81.6 78.0

**# POICHE' L'ALPHA TEORICO (OSSIA IL LIVELLO DI SIGNIFICATIVITA') E' 0.01 ED E' INFERIORE AL p-value CALCOLATO DI 0.7952, SI ACCETTA L'IPOTESI NULLA E SI CONSIDERANO UGUALI I PESI DEGLI STUDENTI DI INFORMATICA E DI SCIENZE MOTORIE AD UN LIVELLO DI CONFIDENZA DEL 99%.**

## 6 - VERIFICA DI IPOTESI PER DATI APPAIATI

A VOLTE I DATI VENGONO PRESENTATI COME COPPIE DI UNA SITUAZIONE PRIMA E DOPO UN DATO TRATTAMENTO.

AD ES. SI VUOLE VALUTARE L'EFFETTO DIMAGRANTE DI UNA NUOVA DIETA SU UN GRUPPO DI 12 PERSONE, I CUI PESI, PRIMA E DOPO LA CURA, SONO I SEGUENTI:

> prima=c(80, 76, 92, 87, 79, 86, 98, 92, 90, 89, 85, 88)

> dopo=c(76, 75, 86, 82, 78, 86, 95, 90, 88, 85, 84, 86)

**H0:  $\text{diff}(\text{prima-dopo}) \leq 0$     H1:  $\text{diff}(\text{prima-dopo}) > 0$**

## 6 - VERIFICA DI IPOTESI PER DATI APPAIATI

**H0:  $\text{diff}(\text{prima-dopo}) \leq 0$     H1:  $\text{diff}(\text{prima-dopo}) > 0$**

```
t.test(prima, dopo, alternative = "greater", paired = TRUE,  
conf.level=0.95)
```

**# SI INDICANO DUE DATABASE DA CONFRONTARE (prima E dopo) E SI SPECIFICA L'ARGOMENTO paired=TRUE PER INDICARE CHE SI TRATTA DI UN CONFRONTO FRA DUE BASE DATI**

# 6 - VERIFICA DI IPOTESI PER DATI APPAIATI

```
t.test(prima, dopo, alternative = "greater", paired = T, conf.level=0.95)
```

Paired t-test

data: prima and dopo

t = 4.8849, df = 11, p-value = 0.0002415

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

1.633598      Inf

sample estimates:

mean of the differences

2.583333

**# POICHE' L'ALPHA 0.05 E' SUPERIORE AL P-VALUE DI 0.0002415, SI RIFIUTA L'IPOTESI NULLA E SI ACCETTA L'IPOTESI ALTERNATIVA CHE LA DIFFERENZA FRA IL PRIMA E IL DOPO SIA MAGGIORE DI 0 E QUINDI CHE CI SIA DIFFERENZA > 0 FRA I PESI DEI PAZIENTI PRIMA E DOPO LA CURA.**

# ES. VERIFICA DI IPOTESI PER DATI APPAIATI

SI VUOLE VALUTARE SE UN SONNIFERO ABBIA EFFICACIA O MENO SU UN GRUPPO DI 5 PAZIENTI.

IL NUMERO DI ORE DI SONNO PRIMA E DOPO IL TRATTAMENTO E' STATO IL SEGUENTE:

prima.sonn=c(6, 6, 6.5, 5, 3)

dopo.sonn=c(6, 7, 8, 6.5, 6)

$H_0: \text{diff}(\text{prima.sonn}-\text{dopo.sonn}) \leq 0$        $H_1: \text{diff}(\text{prima.sonn}-\text{dopo.sonn}) > 0$

# ES. VERIFICA DI IPOTESI PER DATI APPAIATI

**H0:  $\text{diff}(\text{prima.sonn}-\text{dopo.sonn}) \leq 0$       H1:  $\text{diff}(\text{prima.sonn}-\text{dopo.sonn}) > 0$**

```
t.test(prima.sonn, dopo.sonn, alternative = "greater", paired = T, conf.level=0.95)
```

# ES. VERIFICA DI IPOTESI PER DATI APPAIATI

```
t.test(prima.sonn, dopo.sonn, alternative = "greater", paired = T,  
conf.level=0.95)
```

Paired t-test

```
data: prima.sonn and dopo.sonn
```

```
t = -2.888, df = 4, p-value = 0.9777
```

```
alternative hypothesis: true difference in means is greater than 0
```

```
95 percent confidence interval:
```

```
-2.433451      Inf
```

```
sample estimates:
```

```
mean of the differences
```

```
-1.4
```

**# POICHE' L'ALPHA 0.05 E' NETTAMENTE INFERIORE AL P-VALUE DI 0.0002415, SI ACCETTA L'IPOTESI NULLA, OVVERO CHE LA DIFFERENZA (prima.sonn-dopo.sonn)<=0 E QUINDI CHE IL NUMERO DI ORE DI SONNO DOPO IL SONNIFERO E' PIU' ALTO DI QUELLO PRECEDENTE ALL'ASSUNZIONE.**