

Test Non Parametrici Classici

Aldo Solari

Scuola di Alta Formazione Statistica
Inferenza Statistica in Biologia e Scienze Umane
Asti, 7 Luglio 2016



Outline

① Introduzione

② Due campioni indipendenti

③ Campioni appaiati



Verifica di ipotesi: ingredienti

Modello

Specificare eventuali assunzioni sulla distribuzione dei dati, etc.

Ipotesi nulla e alternativa

Specificare H_0 e il suo complementare H_1

Statistica test e distribuzione nulla

Specificare T e la sua distribuzione quando è vera H_0



Proprietà di un test

Controllo dell'Errore di I tipo

La probabilità di rifiutare H_0 quando è vera $\leq \alpha$

Se $= \alpha$ si dice che il test è *esatto*

Non distorsione

La probabilità di rifiutare H_0 quando è vera $\leq \alpha$

La probabilità di rifiutare H_0 quando è falsa $\geq \alpha$

Consistenza

La probabilità di rifiutare H_0 quando è falsa $\rightarrow 1$ quando $n \rightarrow \infty$



Outline

① Introduzione

② Due campioni indipendenti

③ Campioni appaiati



Due campioni indipendenti

<i>Dataset</i>	<i>Risposta</i>	<i>Test parametrico</i>	<i>Test non parametrico</i>
Cholesterol	Continua	t di Student Welch	Wilcoxon-Mann-Whitney Brunner e Munzel Kolmogorov-Smirnov
ECMO	Bernoulliana	Chi-quadrato	Esatto di Fisher



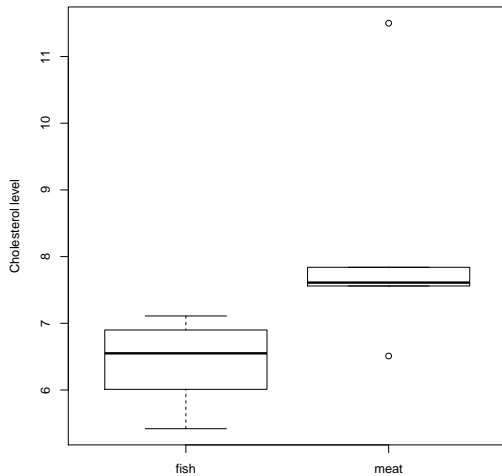
Dataset Cholesterol

- Stabilire se una dieta a base di pesce diminuisce il livello di colesterolo rispetto ad una dieta a base di carne (Ludbrook and Dudley, 1998)
- 12 soggetti sono stati assegnati casualmente alle due diete: 7 alla dieta a base di pesce, e 5 alla dieta a base di carne. Dopo un anno, è stato misurato il loro livello di colesterolo:

Fish	5.42	5.86	6.16	6.55	6.80	7.00	7.11
Meat	6.51	7.56	7.61	7.84	11.50		



Dataset Cholesterol



Modelli parametrici e non parametrici

Modello non parametrico

Due campioni indipendenti

X_1, \dots, X_{n_x} i.i.d. X con f.d.r. \mathcal{F}_X incognita

Y_1, \dots, Y_{n_y} i.i.d. Y con f.d.r. \mathcal{F}_Y incognita

Modello parametrico

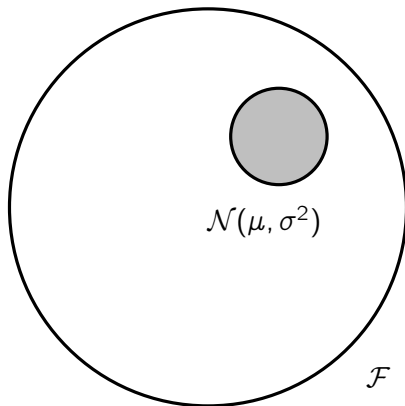
Due campioni indipendenti + Normalità + Omoschedasticità

X_1, \dots, X_{n_x} i.i.d. $X \sim \mathcal{N}(\mu_x, \sigma^2)$, μ_x e σ^2 incogniti

Y_1, \dots, Y_{n_y} i.i.d. $Y \sim \mathcal{N}(\mu_y, \sigma^2)$, μ_y e σ^2 incogniti



Modelli parametrici e non parametrici



Test t di Student per due campioni indipendenti

Modello parametrico

Due campioni indipendenti + Normalità + Omoschedasticità

X_1, \dots, X_{n_x} i.i.d. $X \sim \mathcal{N}(\mu_x, \sigma^2)$, μ_x e σ^2 incogniti

Y_1, \dots, Y_{n_y} i.i.d. $Y \sim \mathcal{N}(\mu_y, \sigma^2)$, μ_y e σ^2 incogniti

Ipotesi nulla e alternativa

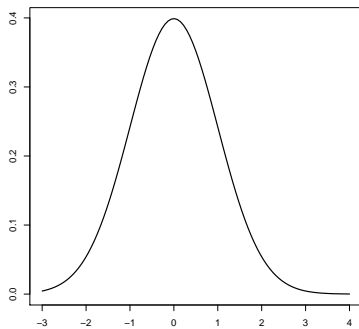
$H_0 : \Delta \leq 0$ vs $H_1 : \Delta > 0$ dove $\Delta = \mathbb{E}(Y) - \mathbb{E}(X) = \mu_y - \mu_x$

Statistica test e distribuzione nulla

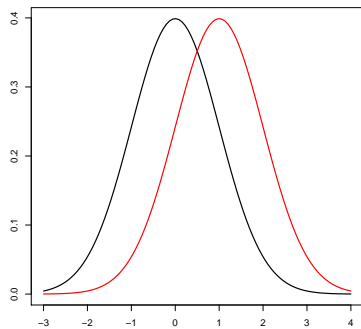
$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{S^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}} \stackrel{\Delta=0}{\sim} \mathcal{T}_{n_x+n_y-2}, \text{ dove } S^2 = \frac{(n_x-1)S_x^2 + (n_y-1)S_y^2}{n_x+n_y-2}$$



H0



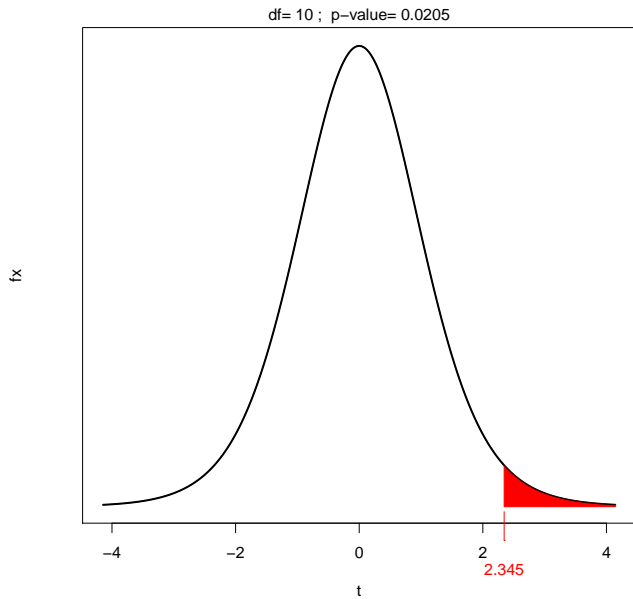
H1



```
> t.test(meat,fish,  
         alternative = "greater",  
         var.equal = TRUE)
```

```
Two Sample t-test with fish and meat  
t = 2.3451, df = 10, p-value = 0.02049  
alternative hypothesis:  
true difference in means is greater than 0  
95 percent confidence interval:  
 0.4064741      Inf  
sample estimates:  
mean of y mean of x  
 8.204000  6.414286
```





Alternativa bilaterale

Alternativa unilaterale destra

$$H_0^+ : \Delta \leq 0 \text{ vs } H_1^+ : \Delta > 0$$

Alternativa unilaterale sinistra

$$H_0^- : \Delta \geq 0 \text{ vs } H_1^- : \Delta < 0$$

Alternativa bilaterale

$$H_0 = H_0^- \cap H_0^+ : \Delta = 0$$

$$H_1 = H_1^- \cup H_1^+ : \Delta \neq 0$$

p -value per alternativa bilaterale

$$p = 2 \cdot \min(p^-, p^+) = 2 \cdot \min(0.02049, 0.9795) = 0.04098$$



La potenza è nulla senza controllo

- $X \sim \text{Esponenziale}(1)$, $n_x = 5$
- $Y \sim \text{Esponenziale}(1)$, $n_y = 15$
- Quindi è vera H_0 , ovvero le due distribuzioni sono identiche
- In teoria, il test t dovrebbe controllare l'errore di I tipo al livello $\alpha = 5\%$ (ma le distribuzioni non sono Normali...)
- In realtà, la probabilità di commettere un errore di I tipo è $\approx 10\%$



Il problema di Behrens-Fisher

- $X \sim \mathcal{N}(0, 1)$, $n_x = 120$
- $Y \sim \mathcal{N}(0, 9)$, $n_y = 40$
- Quindi è vera H_0 , ovvero le due medie sono identiche
- In teoria, il test t dovrebbe controllare l'errore di I tipo ad $\alpha = 5\%$ (ma le distribuzioni non sono omoschedatiche...)
- In realtà, la probabilità di commettere un errore di I tipo è $\approx 14\%$



Test di Wilcoxon-Mann-Whitney

Modello non parametrico

Due campioni indipendenti + *location shift*

X_1, \dots, X_{n_x} i.i.d. X con f.d.r. $\mathcal{F}_x(z)$

Y_1, \dots, Y_{n_y} i.i.d. Y con f.d.r. $\mathcal{F}_y(z) = \mathcal{F}_x(z - \Delta)$

Ipotesi nulla e alternativa

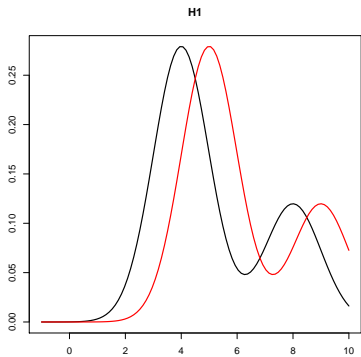
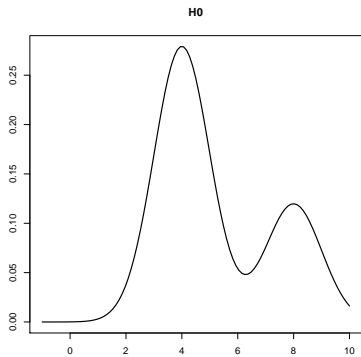
$H_0 : \Delta \leq 0$ vs $H_1 : \Delta > 0$ dove $\Delta = \mathbb{E}(Y) - \mathbb{E}(X)$

Statistica test di Wilcoxon (1945)

$$W = \sum_{i=n_x+1}^{n_x+n_y} R_i$$

dove $R_1, \dots, R_{n_x+n_y}$ sono i ranghi di $\{X_1, \dots, X_{n_x}, Y_1, \dots, Y_{n_y}\}$





Ranghi

i		R_i	
1	x_1	5.42	1
2	x_2	5.86	2
3	x_3	6.16	3
4	x_4	6.55	5
5	x_5	6.80	6
6	x_6	7.00	7
7	x_7	7.11	8
8	y_1	6.51	4
9	y_2	7.56	9
10	y_3	7.61	10
11	y_4	7.84	11
12	y_5	11.50	12

$$W = \sum_{i=8}^{12} R_i = 4 + 9 + 10 + 11 + 12 = 46$$



Statistica test di Mann-Whitney (1947)

$$MW = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} I\{Y_j > X_i\} \quad \text{dove } I\{\cdot\} \text{ è la funzione indicatrice}$$

Relazione con la statistica di Wilcoxon

$$MW = W - n_y(n_y + 1)/2 = 46 - 5(5 + 1)/2 = 31$$

Interpretazione

$$\hat{\theta} = \frac{MW}{n_x n_y} = \frac{31}{7 \cdot 5} = 0.886 \text{ stima di } \theta = \mathbb{P}(Y > X)$$

Probabilità di ottenere un valore più alto se pesco una osservazione dal campione Y rispetto ad una dal campione X



Distribuzione nulla di W

Se è vera $\mathcal{F}_x = \mathcal{F}_y \Leftrightarrow \Delta = 0$

Media

$$\mathbb{E}_0(W) = \frac{n_y(n_x + n_y + 1)}{2}$$

Varianza

$$\mathbb{V}\text{ar}_0(W) = \frac{n_x n_y (n_x + n_y + 1)}{12}$$

Distribuzione nulla (approssimata)

$$Z = \frac{W - \mathbb{E}_0(W)}{\sqrt{\mathbb{V}\text{ar}_0(W)}} \approx \mathcal{N}(0, 1)$$

Correzione per la continuità

$$Z^c = \frac{W - \mathbb{E}_0(W) - 1/2}{\sqrt{\mathbb{V}\text{ar}_0(W)}} \approx \mathcal{N}(0, 1)$$



```
> wilcox.test(meat, fish,  
              alternative="greater",  
              exact=FALSE,  
              correct=FALSE)
```

Wilcoxon rank sum test with fish and meat

W = 31, p-value = 0.01418

alternative hypothesis:

true location shift is greater than 0

```
> wilcox.test(meat, fish,  
              alternative="greater",  
              exact=FALSE,  
              correct=TRUE)
```

Wilcoxon rank sum test with continuity correction with fish

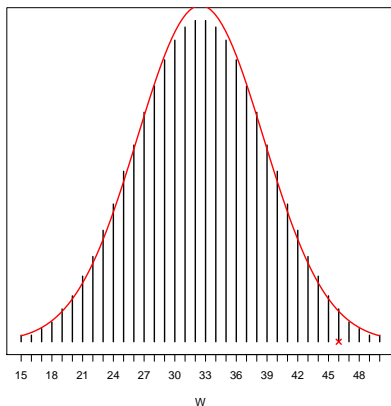
W = 31, p-value = 0.01738

alternative hypothesis:

true location shift is greater than 0



Distribuzione nulla di W



$$\mathbb{P}(W_0 \geq w) = \mathbb{P}(W_0 \geq 46) = (5 + 3 + 2 + 1 + 1)/792 = 0.01515$$

$$\mathbb{P}(\mathcal{N}(0, 1) \geq z) = \mathbb{P}(\mathcal{N}(0, 1) \geq 2.1924) = 0.01418$$

$$\mathbb{P}(\mathcal{N}(0, 1) \geq z^c) = \mathbb{P}(\mathcal{N}(0, 1) \geq 2.1112) = 0.01738$$



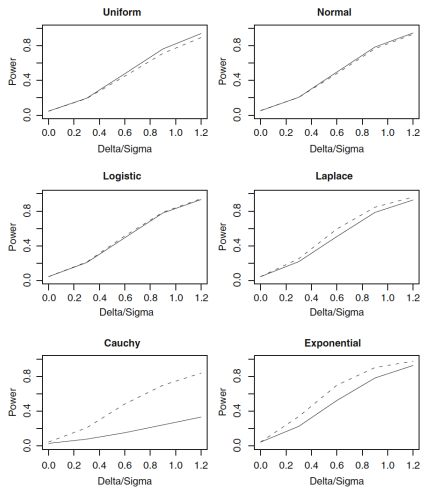
Invarianza del test di WMW rispetto a trasformazioni monotone dei dati

Test	dati originali p -value	dati trasformati in logaritmo p -value
t di Student	0.02049	0.01451
WMW	0.01418	0.01418
WMW ^c	0.01738	0.01738

Inoltre, WMW è robusto rispetto a valori anomali (*outliers*)



Potenza: t vs WMW



$$n_x = n_y = 15$$



Test di Kolmogorov-Smirnov

Modello non parametrico

Due campioni indipendenti

X_1, \dots, X_{n_x} i.i.d. X con f.d.r. \mathcal{F}_x

Y_1, \dots, Y_{n_y} i.i.d. Y con f.d.r. \mathcal{F}_y

Ipotesi nulla e alternativa

$H_0 : \mathcal{F}_x(z) = \mathcal{F}_y(z) \forall z$ vs $H_1 : \mathcal{F}_x(z) \neq \mathcal{F}_y(z)$ per almeno un z

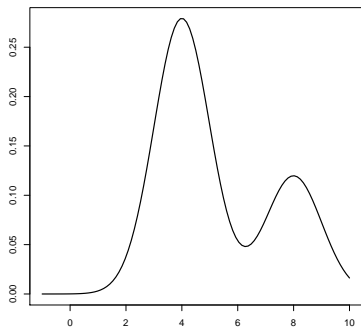
Statistica test di Kolmogorov-Smirnov

$$KS = \max_z |\hat{\mathcal{F}}_x(z) - \hat{\mathcal{F}}_y(z)|$$

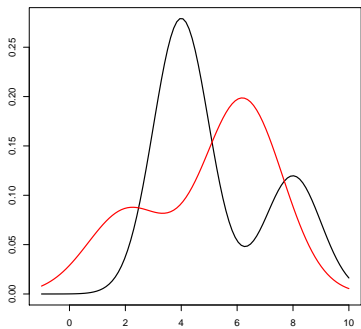
dove $\hat{\mathcal{F}}_x$ e $\hat{\mathcal{F}}_y$ sono le f.d.r. empiriche



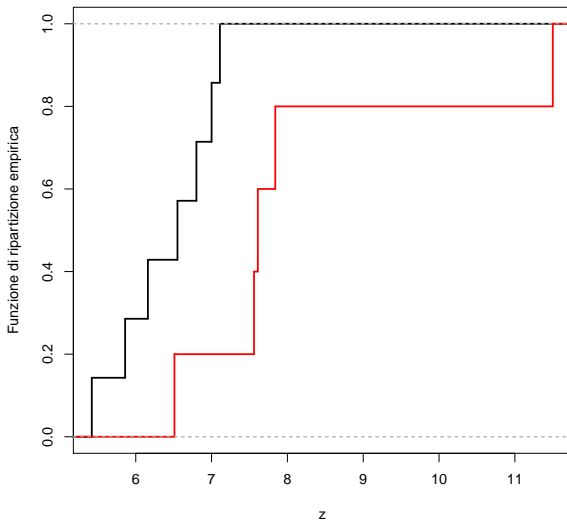
H0



H1



Funzione di ripartizione empirica



```
> ks.test(fish, meat)
```

Two-sample Kolmogorov-Smirnov test with fish and meat
 $D = 0.8$, $p\text{-value} = 0.0303$
alternative hypothesis: two-sided

```
> ks.test(fish, meat,  
          alternative="greater")
```

Two-sample Kolmogorov-Smirnov test with fish and meat
 $D^+ = 0.8$, $p\text{-value} = 0.02391$
alternative hypothesis: the CDF of x lies above that of y



Test di Welch

Modello parametrico

Due campioni indipendenti + Normalità

X_1, \dots, X_{n_x} i.i.d. $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$, μ_x e σ_x^2 incogniti
 Y_1, \dots, Y_{n_y} i.i.d. $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$, μ_y e σ_y^2 incogniti

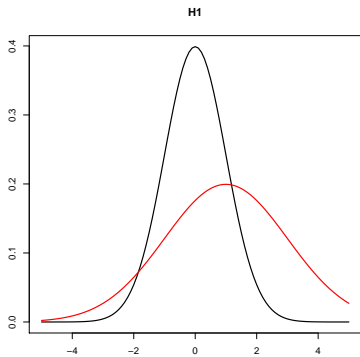
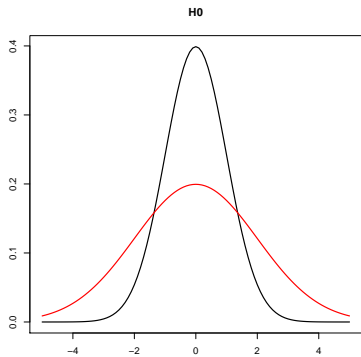
Ipotesi nulla e alternativa

$H_0 : \Delta \leq 0$ vs $H_1 : \Delta > 0$ dove $\Delta = \mathbb{E}(Y) - \mathbb{E}(X)$

Statistica test e distribuzione nulla

$$T_W = \frac{\bar{Y} - \bar{X}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} \stackrel{\Delta=0}{\sim} \mathcal{T}_g$$






```
> t.test(meat, fish,  
         alternative="greater",  
         var.equal = FALSE)
```

```
Welch Two Sample t-test with meat and fish  
t = 2.0165, df = 4.618, p-value = 0.05225  
alternative hypothesis:  
true difference in means is greater than 0  
95 percent confidence interval:  
-0.03188968      Inf  
sample estimates:  
mean of x mean of y  
8.204000  6.414286
```



Test di Brunner e Munzel

Modello non parametrico

Due campioni indipendenti

X_1, \dots, X_{n_x} i.i.d. X con f.d.r. \mathcal{F}_x

Y_1, \dots, Y_{n_y} i.i.d. Y con f.d.r. \mathcal{F}_y

Ipotesi nulla e alternativa

$H_0 : \theta \leq 1/2$ vs $H_1 : \theta > 1/2$ dove $\theta = \mathbb{P}(Y > X)$

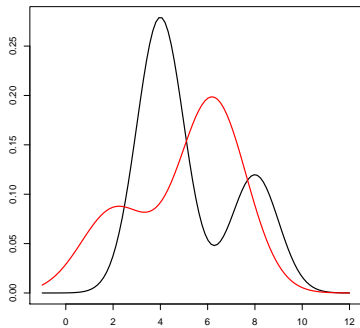
Notare che $\mathcal{F}_x = \mathcal{F}_y \Rightarrow \theta = 1/2$ ma non viceversa

Statistica test e distribuzione nulla

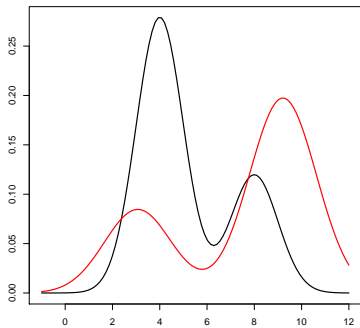
$$BM = \frac{\bar{R}_Y - \bar{R}_X}{\sqrt{\hat{V}^2}} \stackrel{\theta=1/2}{\approx} \mathcal{N}(0, 1)$$



H0



H1



```
> npar.t.test(colesterolo ~ dieta, data = dati,
               alternative="greater",
               method = "t.app")
```

- Alternative Hypothesis:

True relative effect p is greater than $1/2$

- Confidence level: 95 %

- Method = Brunner - Munzel - T - Approx with 5.01 DF

#-----Interpretation-----

$p(a,b) > 1/2$: b tends to be larger than a

#-----

	Effect	Estimator	Lower	Upper	T	p.Value
1	$p(\text{fish}, \text{meat})$	0.886	0.642	1	3.182	0.012



Risultati

Test	dati originali p -value	dati trasformati in logaritmo p -value
t di Student	0.02049	0.01451
WMW	0.01418	0.01418
WMW ^c	0.01738	0.01738
KS	0.02391	0.02391
Welch	0.05225	0.03513
BM	0.01200	0.01200



Dataset ECMO

- Verificare se la tecnica di ossigenazione extracorporea a membrana riduce la mortalità dei neonati con ipertensione polmonare rispetto alla terapia tradizionale
- 39 neonati con ipertensione polmonare sono stati assegnati casualmente al trattamento ExtraCorporeal Membrane Oxygenation (ECMO) oppure alla terapia medica convenzionale conventional (CMT)

ECMO : 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
CMT : 0 1 1 0 1 1 0 0 0 0

dove 1 = Sopravvissuto e 0 = Deceduto



Dataset ECMO

- Frequenze osservate:

	Deceduti	Sopravvissuti	
ECMO	1	28	29
CMT	4	6	10
	5	34	39

- Frequenze attese sotto l'ipotesi $p_x = p_y$:

	Deceduti	Sopravvissuti	
ECMO	3.72	25.28	29
CMT	1.28	8.72	10
	5	34	39



```
> chisq.test(ecmo, correct=F)
```

Pearson's Chi-squared test with ecmo

X-squared = 8.8885, df = 1, p-value = 0.00287

Warning message:

In `chisq.test(ecmo, correct = F)` :

Chi-squared approximation may be incorrect



Risposta bernoulliana

X_1, \dots, X_{n_x} i.i.d. $X \in \{0, 1\} \sim \text{Bernoulli}(p_x)$

Y_1, \dots, Y_{n_y} i.i.d. $Y \in \{0, 1\} \sim \text{Bernoulli}(p_y)$

$S_x = \sum_{i=1}^{n_x} X_i \sim \text{Binomiale}(n_x, p_x)$

$S_y = \sum_{j=1}^{n_y} Y_j \sim \text{Binomiale}(n_y, p_y)$

	Successi	Insuccessi	
campione X	S_x	$n_x - S_x$	n_x
campione Y	S_y	$n_y - S_y$	n_y
	$S = S_x + S_y$	$n_x + n_y - S$	$n_x + n_y$

Ipotesi nulla e alternativa

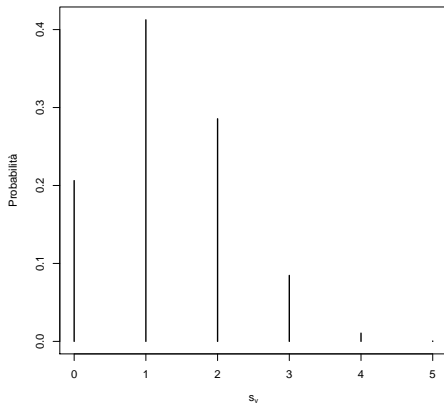
$H_0 : p_y \leq p_x$ vs $H_1 : p_y > p_x$



Test esatto di Fisher

$$\text{Se } p_x = p_y, \text{ allora } \mathbb{P}(S_y = s_y | S = s) = \frac{\binom{n_x}{s-s_y} \binom{n_y}{s_y}}{\binom{n_x+n_y}{s}}$$

ovvero $S_y|S$ segue una distribuzione Ipergeometrica



Fisher's Exact Test for Count Data with ecmo

p-value = 0.01102

alternative hypothesis: true odds ratio is greater than 1

95 percent confidence interval:

1.833681 Inf

sample estimates:

odds ratio

16.78571



Outline

- 1 Introduzione
- 2 Due campioni indipendenti
- 3 Campioni appaiati**



Campioni appaiati

<i>Dataset</i>	<i>Risposta</i>	<i>Test parametrico</i>	<i>Test non parametrico</i>
SAR	Continua	t di Student	Ranghi con segno di Wilcoxon
AHR	Bernoulliana		McNemar



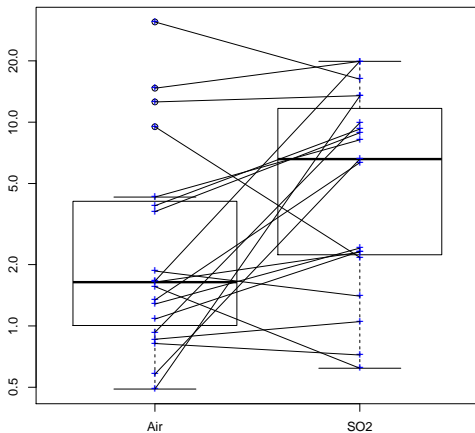
Dataset SAR

- Stabilire se la resistenza delle vie aeree (*specific airways resistance*) è maggiore quando si respira diossido di zolfo (SO₂) oppure quando si respira aria normale (Air)
- Per 19 soggetti è stata misurata la SAR in uno stato di riposo e dopo 5 minuti di esercizio.
Per ciascun soggetto, la differenza SAR(esercizio) - SAR(riposo) è stata valutata in due condizioni: respirando aria 'normale' e diossido di zolfo

	1	2	3	4	5	6	7	8	9	10
Air	0.82	0.86	1.86	1.64	12.57	1.56	1.28	1.08	4.29	1.34
SO2	0.72	1.05	1.40	2.30	13.49	0.62	2.41	2.32	8.19	6.33
	11	12	13	14	15	16	17	18	19	
Air	14.68	3.64	3.89	0.58	9.50	0.93	0.49	31.04	1.66	
SO2	19.88	8.87	9.25	6.59	2.17	9.93	13.44	16.25	19.89	

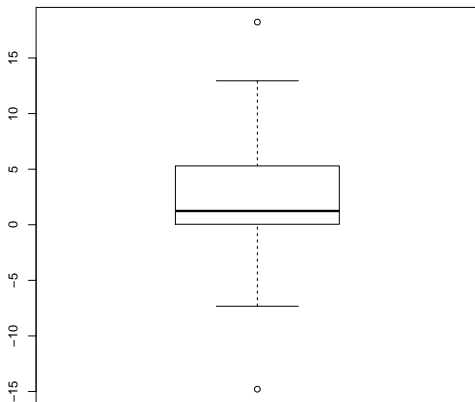


Dataset SAR



Differenza SO2 - Air

	1	2	3	4	5	6	7	8	9	10
Dif	-0.10	0.19	-0.46	0.66	0.92	-0.94	1.13	1.24	3.90	4.99
	11	12	13	14	15	16	17	18	19	
Dif	5.20	5.23	5.36	6.01	-7.33	9.00	12.95	-14.79	18.23	



Test t di Student per campioni appaiati

Modello parametrico

Campioni dipendenti + Normalità

$(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d. (X, Y)

$D_1 = Y_1 - X_1, \dots, D_n = Y_n - X_n$ i.i.d. $D \sim \mathcal{N}(\Delta, \sigma^2)$

Ipotesi nulla e alternativa

$H_0 : \Delta \leq 0$ vs $H_1 : \Delta > 0$ con $\Delta = \mathbb{E}(D) = \mathbb{E}(Y) - \mathbb{E}(X)$

Statistica test e distribuzione nulla

$$T = \frac{\bar{D}}{\sqrt{\frac{S^2}{n}}} \stackrel{\Delta=0}{\sim} \mathcal{T}_{n-1}$$



Test dei ranghi con segno di Wilcoxon

Modello non parametrico

Campioni dipendenti + simmetria rispetto a Δ

$(X_1, Y_1), \dots, (X_n, Y_n)$ indipendenti

$D_1 = Y_1 - X_1, \dots, D_n = Y_n - X_n$ indipendenti e

$D_i - \Delta \stackrel{d}{=} -(D_i - \Delta)$, distribuzione simmetrica rispetto a Δ

Ipotesi nulla e alternativa

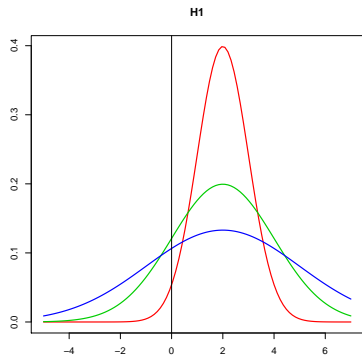
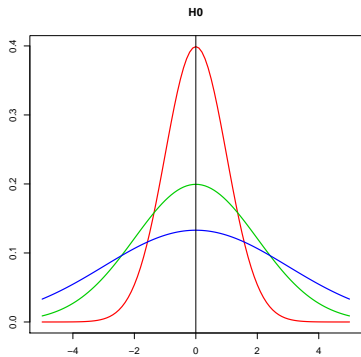
$H_0 : \Delta \leq 0$ vs $H_1 : \Delta > 0$ dove $\Delta = \mathbb{E}(D_i) = \mathbb{E}(Y_i) - \mathbb{E}(X_i)$

Statistica test

$$W = \sum_{i=1}^n I\{D_i > 0\} \tilde{R}_i$$

dove $\tilde{R}_1, \dots, \tilde{R}_n$ sono i ranghi di $\{|D_1|, \dots, |D_n|\}$





Test dei ranghi con segno di Wilcoxon

i	1	2	3	4	5	6	7	8	9	10
D_i	-0.10	0.19	-0.46	0.66	0.92	-0.94	1.13	1.24	3.90	4.99
$ D_i $	0.10	0.19	0.46	0.66	0.92	0.94	1.13	1.24	3.90	4.99
\tilde{R}_i	1	2	3	4	5	6	7	8	9	10
$I\{D_i > 0\}$	0	1	0	1	1	0	1	1	1	1
i	11	12	13	14	15	16	17	18	19	
D_i	5.20	5.23	5.36	6.01	-7.33	9.00	12.95	-14.79	18.23	
$ D_i $	5.20	5.23	5.36	6.01	7.33	9.00	12.95	14.79	18.23	
\tilde{R}_i	11	12	13	14	15	16	17	18	19	
$I\{D_i > 0\}$	1	1	1	1	0	1	1	0	1	

$$W = 2 + 4 + 5 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 + 16 + 17 + 19 = 147$$



Distribuzione nulla di W

Se è vera $\Delta = 0 \Leftrightarrow D_i \stackrel{d}{=} -D_i$

Media

$$\mathbb{E}_0(W) = \frac{n(n+1)}{4}$$

Varianza

$$\mathbb{V}\text{ar}_0(W) = \frac{n(n+1)(2n+1)}{24}$$

Distribuzione nulla (approssimata)

$$Z = \frac{W - \mathbb{E}_0(W)}{\sqrt{\mathbb{V}\text{ar}_0(W)}} \approx \mathcal{N}(0, 1)$$

Correzione per la continuità

$$Z^c = \frac{W - \mathbb{E}_0(W) - 1/2}{\sqrt{\mathbb{V}\text{ar}_0(W)}} \approx \mathcal{N}(0, 1)$$



```
> wilcox.test(SO2, Air,  
              pair=TRUE,  
              exact=FALSE,  
              correct=FALSE,  
              alternative="greater")
```

Wilcoxon signed rank test with SO2 and Air
V = 147, p-value = 0.01819
alternative hypothesis:
true location shift is greater than 0



Dataset AHR

- Si vuole valutare se a seguito di un trapianto di cellule staminali (SCT) aumenta la presenza di iperreattività bronchiale (*Airway hyper-responsiveness*)
- Per 21 bambini è stata misurata la presenza o assenza di AHR prima e dopo un trapianto di cellule staminali

Prima SCT	Dopo SCT		
	AHR	No AHR	
AHR	1	1	2
No AHR	7	12	19
	8	13	21



Test di McNemar

Modello parametrico

Campioni dipendenti

$(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d. (X, Y)
con $X \sim \text{Bernoulli}(p_x)$ e $Y \sim \text{Bernoulli}(p_y)$

Ipotesi nulla e alternativa

$$H_0 : p_y \leq p_x \text{ vs } H_1 : p_y > p_x$$

Statistica test e distribuzione nulla (approssimata)

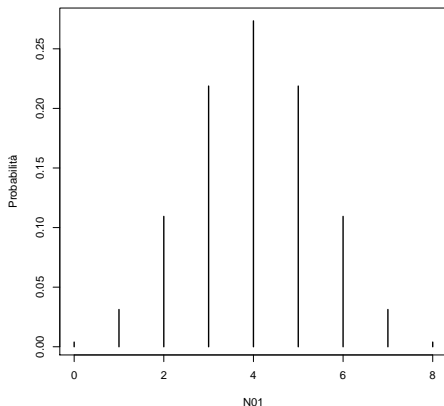
$$N = \frac{N_{01} - N_{10}}{\sqrt{N_{01} + N_{10}}} \stackrel{p_x=p_y}{\approx} \mathcal{N}(0, 1) \text{ dove}$$

$$N_{01} = \sum_{i=1}^n I\{X_i = 0, Y_i = 1\} \text{ e } N_{10} = \sum_{i=1}^n I\{X_i = 1, Y_i = 0\}$$



Test di McNemar esatto

$$N_{01} | N_{01} + N_{10} \stackrel{p_x = p_y}{\sim} \text{Binomiale}(N_{01} + N_{10}, 1/2)$$



$$\mathbb{P}(\text{Binomiale}(8, 1/2) \geq 7) = 0.03125 + 0.00391 = 0.03516$$

