5. Confronto tra dati appaiati (confronto tra medie e confronto tra mediane)

Questa volta l'esercizio consiste, una volta raccolti i dati, nel resistere alla tentazione di saltare direttamente alle conclusioni, calcolando la significatività del confronto tra dati appaiati e riportando un valore di probabilità p con la constatazione che la differenza tra medie è significativa (se $p < 0{,}05$) o non è significativa (se $p \geq 0{,}05$).

Invece di saltare direttamente alle conclusioni, proviamo invece a seguire alcune indicazioni di tipo metodologico, che trovate alla pagina http://www.bayes.it/html/statistica_e_laboratorio.html. Lo faremo utilizzando il mio programma di statistica Ministat, che potete scaricare liberamente dall'area di download (http://www.bayes.it/html/download.html). Se utilizzate Windows 7 dovete installare dal sito della Microsoft una patch: seguite le indicazioni che ho riportato nella Bacheca (http://www.bayes.it/html/bacheca.html) nel novembre 2010.

5.1. Il disegno sperimentale

Si vuole stabilire se l'enzima aspartato-amminotransferasi (AST) è stabile per 24 ore se conservato a temperatura di frigorifero. Per fare questo viene raccolto un certo numero di campioni di sangue, che sono immediatamente centrifugati. Per ciascun campione il siero viene suddiviso in due aliquote identiche, la prima della quali viene analizzata immediatamente, mentre la seconda, ben tappata, viene conservata in frigorifero a temperatura controllata compresa tra +2 °C e +8°C. L'indomani, trascorse 24 ore, i campioni sono nuovamente analizzati. Si ottengono i seguenti risultati (concentrazione dell'AST espressa in U/L):

Subito	Dopo_24_ore
20	20
45	48
24	25
17	17
52	57
42	40
101	106
67	71
174	180
327	300
24	27
652	631
433	440
18	17
22	20
30	25
62	60
95	86
19	24
495	515
37	42
476	430

Il disegno sperimentale adottato consente di avere una informazione molto pulita, con un rapporto segnale/rumore tra i migliori che si possano ottenere in un esperimento. Infatti i due insiemi di dati differiscono per un solo fattore, quello che lo sperimentatore ha deliberatamente introdotto: il fattore tempo. Ogni coppia di valori è stata ottenuta sullo stesso identico siero, quindi l'effetto osservato (differenza tra il primo e il secondo campione) deve essere stato causato esclusivamente dal fattore tempo, a meno dell'incertezza (rumore) legata all'errore di misura (le misure di concentrazione dell'AST sono comunque affette da un errore casuale, vedere ad esempio http://www.bayes.it/pdf/Appendice_B.pdf). In generale ci si riferisce alla variabile introdotta dallo sperimentatore come al "trattamento". L'espressione vale sia nel nostro caso, nel quale abbiamo sottoposto ad un trattamento specifico un campione di siero, sia nel caso di sperimentazioni cliniche, che prevedono di valutare per esempio il risultato del trattamento di pazienti con un farmaco A e con un farmaco B (anche se in questo caso il disegno sperimentale è necessariamente più complicato).

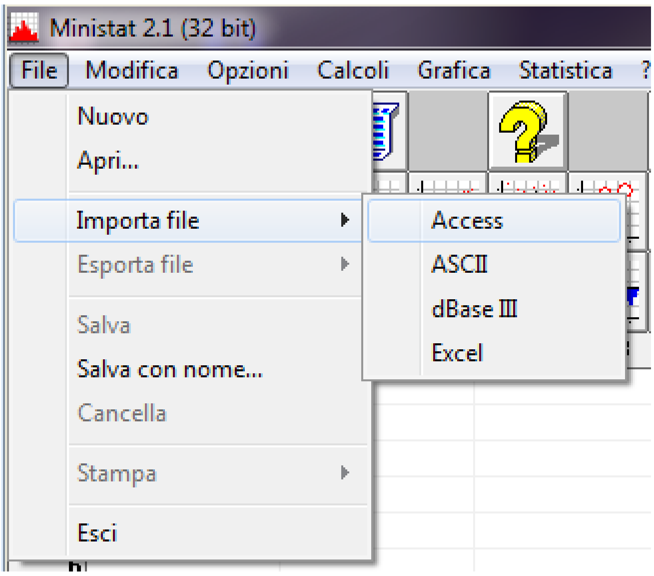
5.2. Analisi esplorativa dei dati

L'analisi esplorativa dei dati può fornire informazioni preliminari della massima importanza, e addirittura condizionare la scelta del test statistico da applicare. E' quindi un passo imprescindibile e propedeutico all'analisi statistica dei dati. La tabulazione dei dati e la rappresentazione grafica dei dati sono i due passaggi chiave nell'analisi esplorativa dei dati.

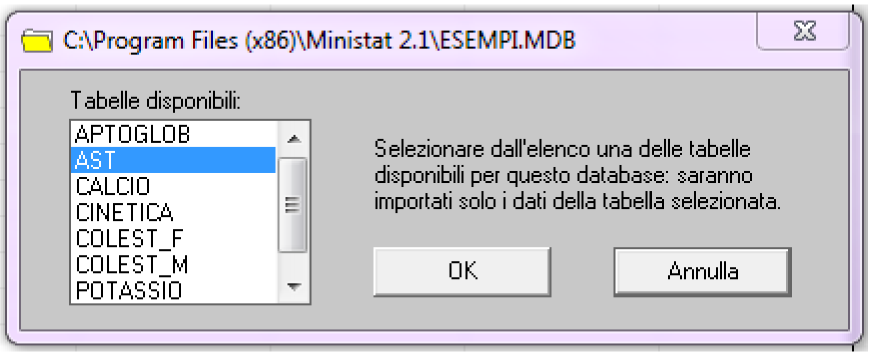
5.2.1. Tabulazione dei dati

La presentazione in forma tabellare consente una valutazione analitica dei dati ottenuti nell'esperimento. Permette di ordinarli in ordine crescente o decrescente, e quindi di valutare l'ampiezza della loro distribuzione. Permette di verificare se i dati sembrano seguire una distribuzione nota (per esempio una distribuzione gaussiana). Consente di individuare dati (sospetti "dati aberranti") che sembrano discostarsi eccessivamente, con una netta soluzione di continuo, dalla massa dei dati raccolti, e quindi di focalizzare l'attenzione sulle possibili cause di questo comportamento. Consente di effettuare sui dati trasformazioni o elaborazioni numeriche. Tutte operazioni che aiutano a fare emergere dai dati valore aggiunto in termini di informazione.

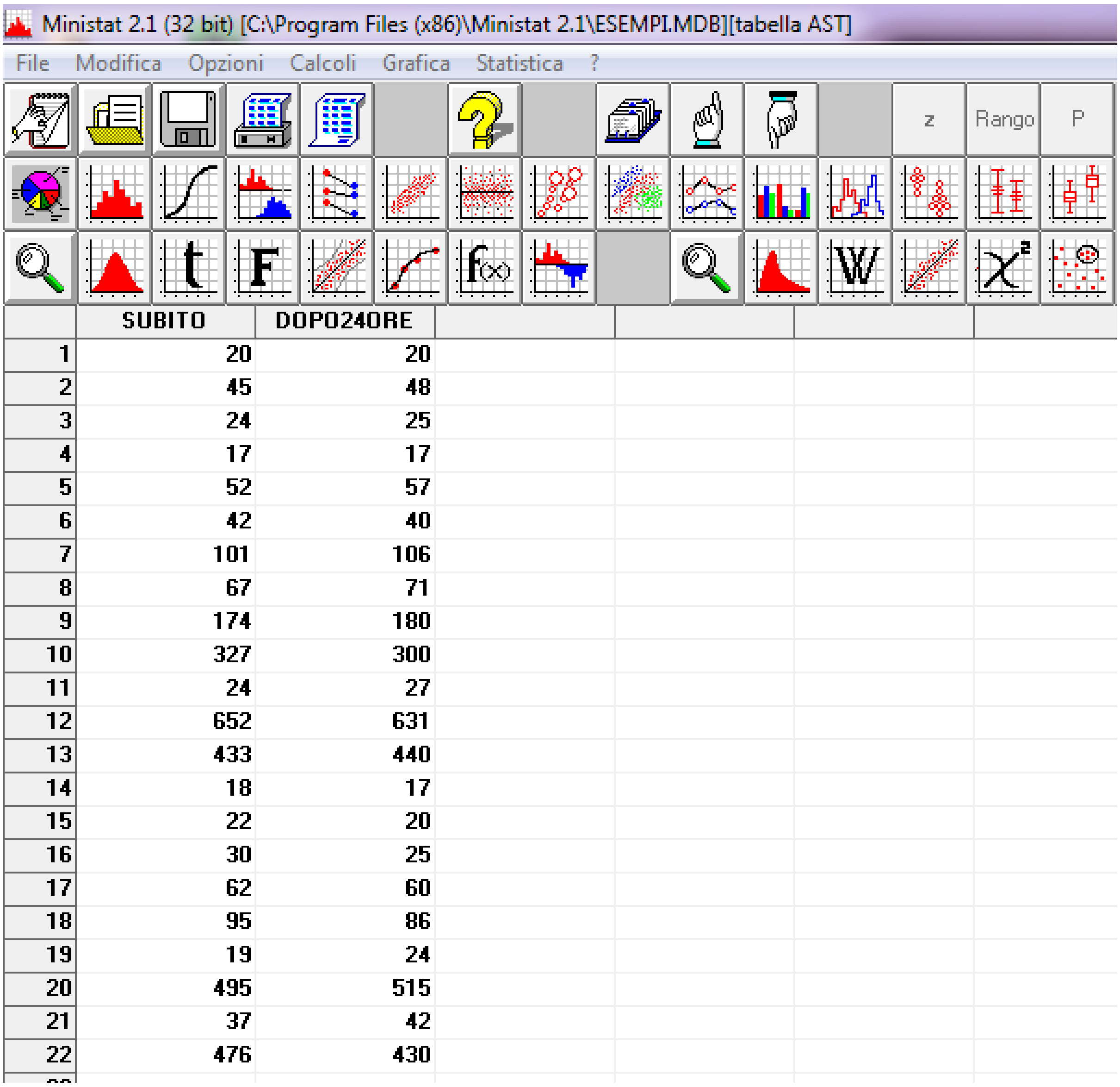
Lanciate Ministat e dal menù *File* selezionate *Importa File* e selezionate *Access*:



Il file ESEMPI.MDB contiene alcune tabelle di dati dimostrativi. Selezionate la tabella AST, che contiene esattamente i dati relativi alla concentrazione dell'AST riportati sopra, e premete OK

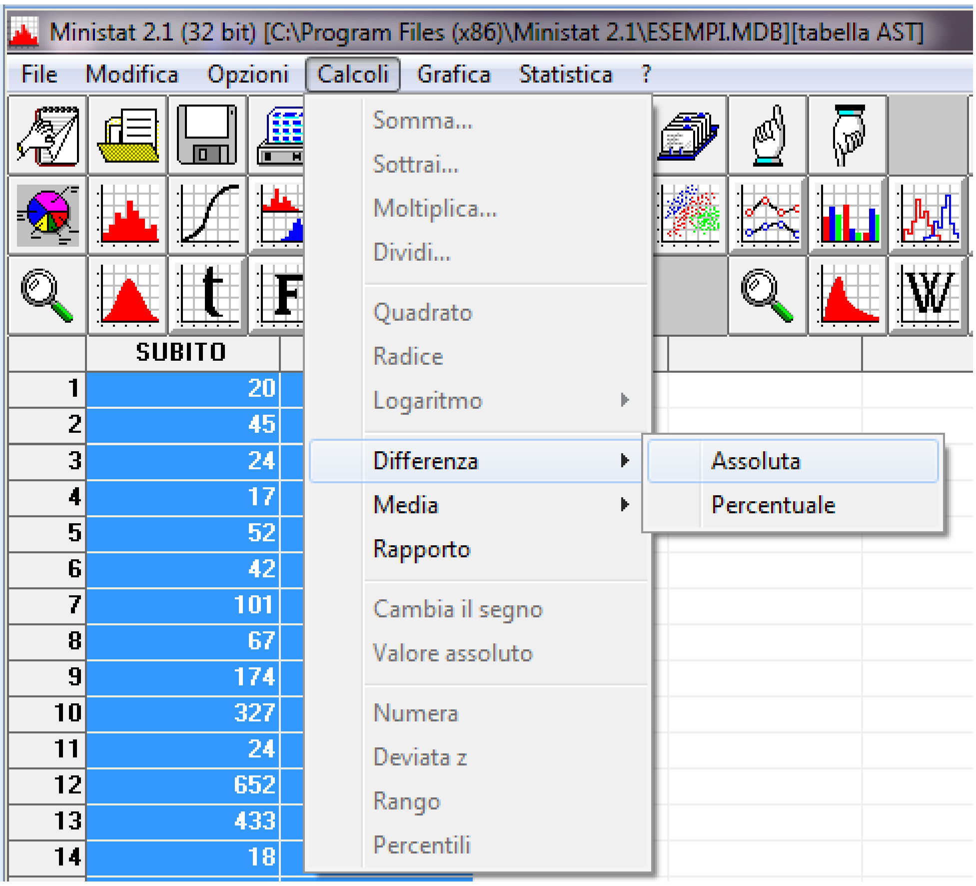


A questo punto avrete caricato i dati nella griglia dati di Ministat

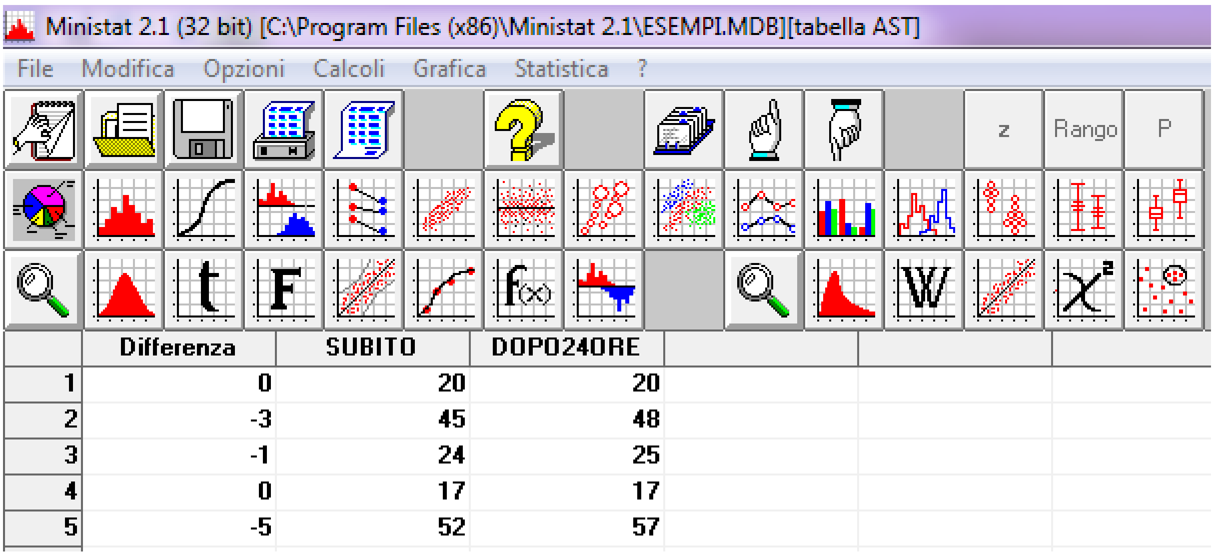


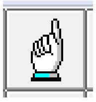
	SUBITO	DOPO24ORE
1	20	20
2	45	48
3	24	25
4	17	17
5	52	57
6	42	40
7	101	106
8	67	71
9	174	180
10	327	300
11	24	27
12	652	631
13	433	440
14	18	17
15	22	20
16	30	25
17	62	60
18	95	86
19	19	24
20	495	515
21	37	42
22	476	430

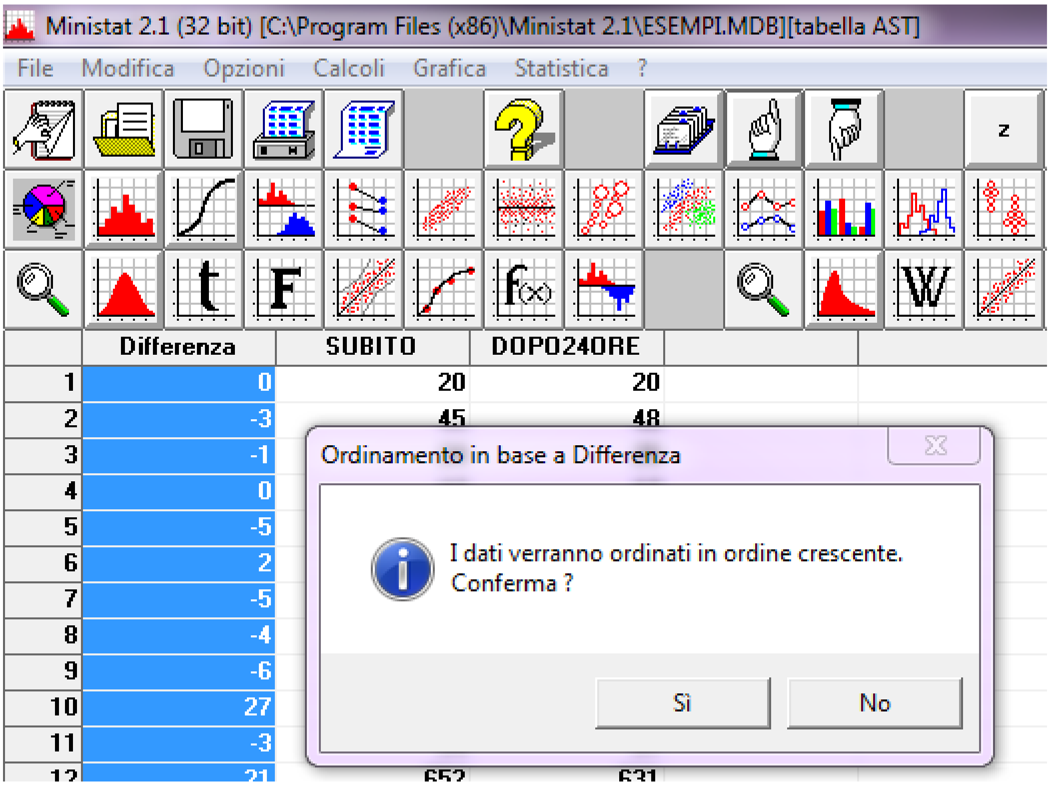
Iniziamo ora ad elaborare i dati. Come prima cosa fate click sulla colonna SUBITO e, senza rilasciare il tasto del mouse, spostatevi sulla colonna DOPO24ORE: in questo modo le avrete selezionate entrambe. Dal menù *Calcoli* selezionate *Differenza* e selezionate *Assoluta*



Rispondete in modo affermativo alle due successive domande. Nella griglia di Ministat compare ora una nuova colonna, che contiene la differenza, con il segno, tra il valore di concentrazione dell'AST misurato subito e quello misurato dopo 24 ore



Fate ora click sulla colonna *Differenza* per selezionarla, quindi fate click sull’icona per ordinare i dati in ordine crescente 



Già da questa semplice tabulazione dei dati possiamo ricavare alcune importanti informazioni.

	Differenza	SUBITO	DOPO24ORE
1	-20	495	515
2	-7	433	440
3	-6	174	180
4	-5	19	24
5	-5	101	106
6	-5	37	42
7	-5	52	57
8	-4	67	71
9	-3	45	48
10	-3	24	27
11	-1	24	25
12	0	17	17
13	0	20	20
14	1	18	17
15	2	62	60
16	2	22	20
17	2	42	40
18	5	30	25
19	9	95	86
20	21	652	631
21	27	327	300
22	46	476	430

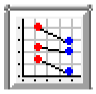
Risultano infatti subito evidenti alcune cose:
→ le differenze (SUBITO – DOPO24ORE) includono sia valori negativi sia valori positivi;
→ i valori negativi e i valori positivi sono distribuiti in modo abbastanza simmetrico;
→ a concentrazioni elevate di AST corrispondono differenze elevate tra le coppie di valori;
→ a concentrazioni basse di AST corrispondono differenze ridotte tra le coppie di valori.

L'impressione che si ricava dalla tabulazione dei dati è che sembra quindi esistere una proporzionalità tra valore di concentrazione dell'AST e differenza tra le coppie di valori, con un bilanciamento "quasi esatto" tra valori negativi e valori positivi.

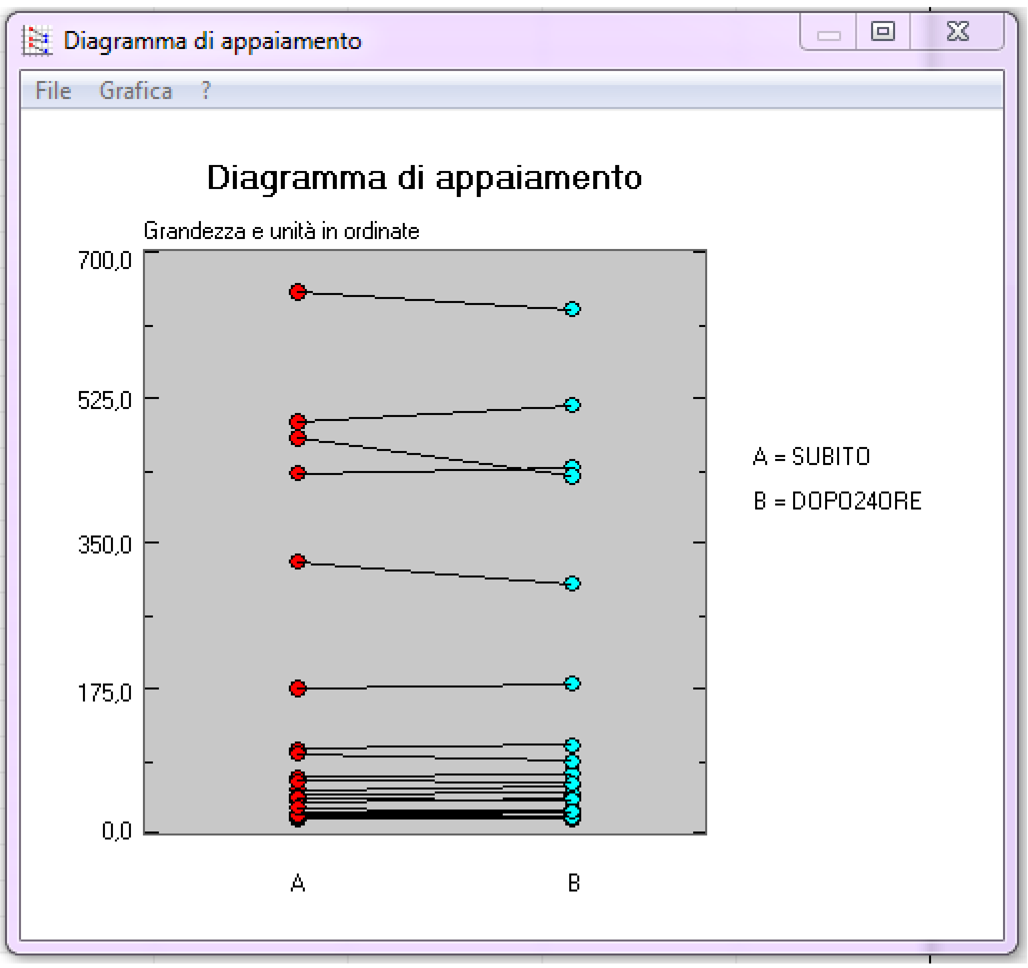
5.2.2. Rappresentazione grafica dei dati

Se la presentazione in forma tabellare consente una valutazione analitica dei dati ottenuti nell'esperimento, la rappresentazione grafica consente di riassumerli in modo sintetico. E i due momenti dell'analisi esplorativa dei dati, tabulazione e rappresentazione grafica, diventano in questo modo complementari.

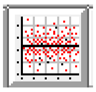
Selezionate le colonne SUBITO e DOPO24ORE e fate click sull'icona del *Diagramma di appaiamento*



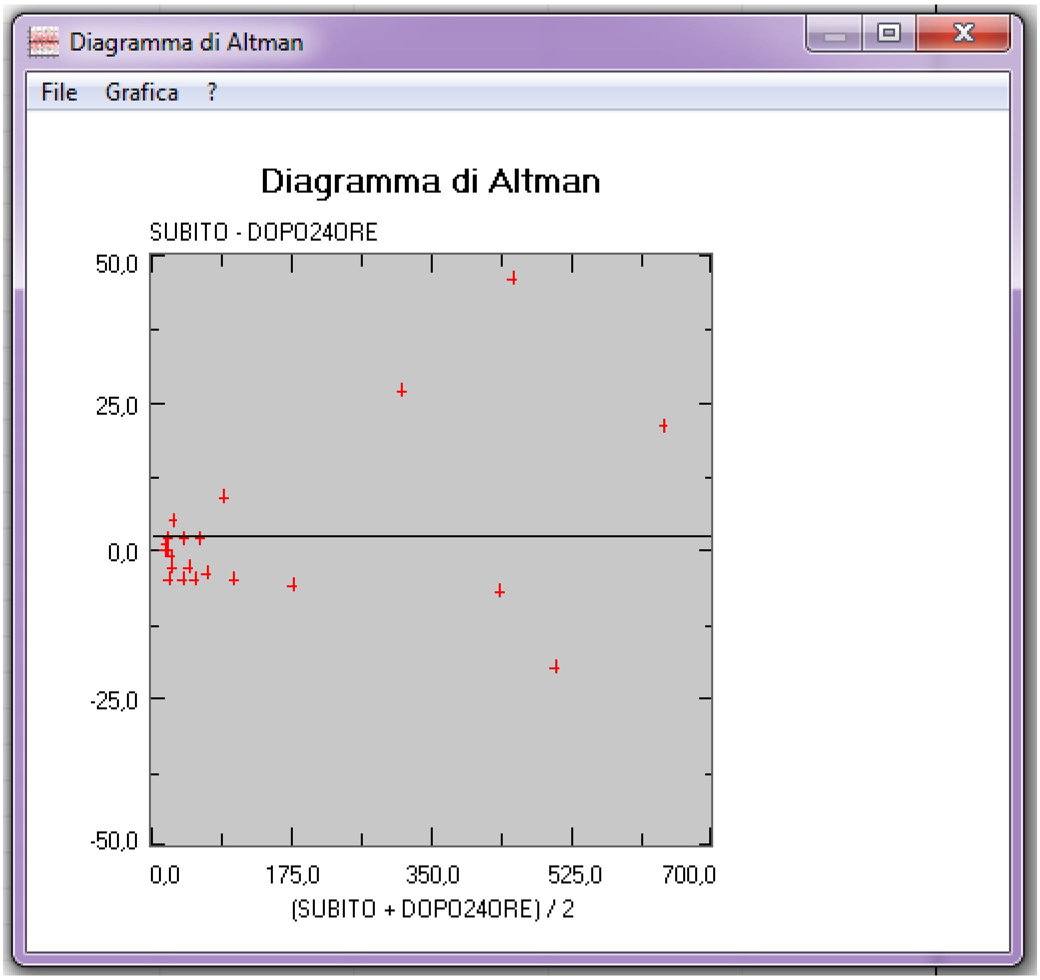
Abbiamo ora una rappresentazione grafica dei dati che mette in corrispondenza per ogni coppia di dati il valore iniziale (SUBITO) e il valore finale (DOPO24ORE)



L'impressione che si ricava da questa semplice rappresentazione grafica conferma quanto ricavato dalla tabulazione dei dati: sembra che non vi sia una sistematicità nelle differenze osservate.

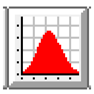
Chiudete ora la finestra del *Diagramma di appaiamento* e fate click sull'icona del *Diagramma di Altman* 

Anche questa rappresentazione grafica sembra confermare la mancanza di sistematicità nelle differenze osservate tra le coppie di valori



5.3. Analisi statistica dei dati

Siamo finalmente arrivati all'analisi statistica dei dati. Per capirne il senso è necessario notare che nel nostro caso tutta l'informazione ricavabile dai dati, rappresentati da coppie di valori ottenuti sullo stesso campione a meno di una differenze specificamente introdotta dal disegno sperimentale, è contenuta in una sola variabile, la variabile *Differenza* (SUBITO – DOPO24ORE) che abbiamo calcolato a partire dai dati originali (vedere **5.2.1. Tabulazione dei dati**).

D'ora in poi ci concentreremo su questa sola variabile. Innanzitutto calcoliamo con Ministat le statistiche elementari della variabile *Differenza*, selezionando la relativa colonna e facendo click sull'icona delle *Statistiche parametriche* 

Le cose incominciano a farsi interessanti se dal menù *Statistica* selezionate *Tabella dei percentili*

Statistiche elementari parametriche

File Modifica Grafica Statistica ?

Ministat 2.1 (32 bit) 24/03/2011 21:01:06

Tabella dei percentili parametrici (*inizio*)

File : C:\Program Files (x86)\Ministat 2.1\ESEMPI.MDB
Tabella : AST
Variabile : Differenza

		Limiti di confidenza al 90%		Valore
Percentile	Valore	inferiore	superiore	non parametrico comparativo
2,5	-24,4153	-32,71764	-16,11295	###
5,0	-20,11885	-27,55643	-12,68126	-18,05001
10,0	-15,1677	-21,69321	-8,64218	-6,7
15,0	-11,81237	-17,79129	-5,83345	-5,82
20,0	-9,15266	-14,75269	-3,55263	-5,36
25,0	-6,87486	-12,19725	-1,55248	-4,9
30,0	-4,82893	-9,94497	0,28711	-4,44
35,0	-2,93304	-7,89886	2,03278	-3,96667
40,0	-1,13262	-5,99574	3,7305	-3,2
45,0	0,61324	-4,18989	5,41637	-1,86667
50,0	2,31818	-2,46541	7,10177	-0,66667
55,0	4,02312	-0,78	8,82625	0,1
60,0	5,76898	0,90586	10,63211	0,86667
65,0	7,5694	2,60358	12,53522	1,475
70,0	9,4653	4,34926	14,58134	2,15
75,0	11,51123	6,18884	16,83361	3,875
80,0	13,78903	8,189	19,38906	6,6
85,0	16,44874	10,46982	22,42765	15,6
90,0	19,80406	13,27855	26,32957	25,2
95,0	24,75521	17,31762	32,1928	43,15001
97,5	29,05166	20,74932	37,35401	###

Tabella dei percentili parametrici (*fine*)

Come potete constatare il valore del 50° percentile parametrico, che è banalmente la media, è 2,31818. Nella terza e nella quarta colonna della tabella compaiono i limiti di confidenza della media (e di ciascuno degli altri percentili che compaiono nella tabella). I limiti di confidenza al 90% della media sono -2,46541 (il limite inferiore) e 7,10117 (il limite superiore). Vediamo come utilizzare questa informazione ricordandoci il senso dell'esperimento che è stato fatto

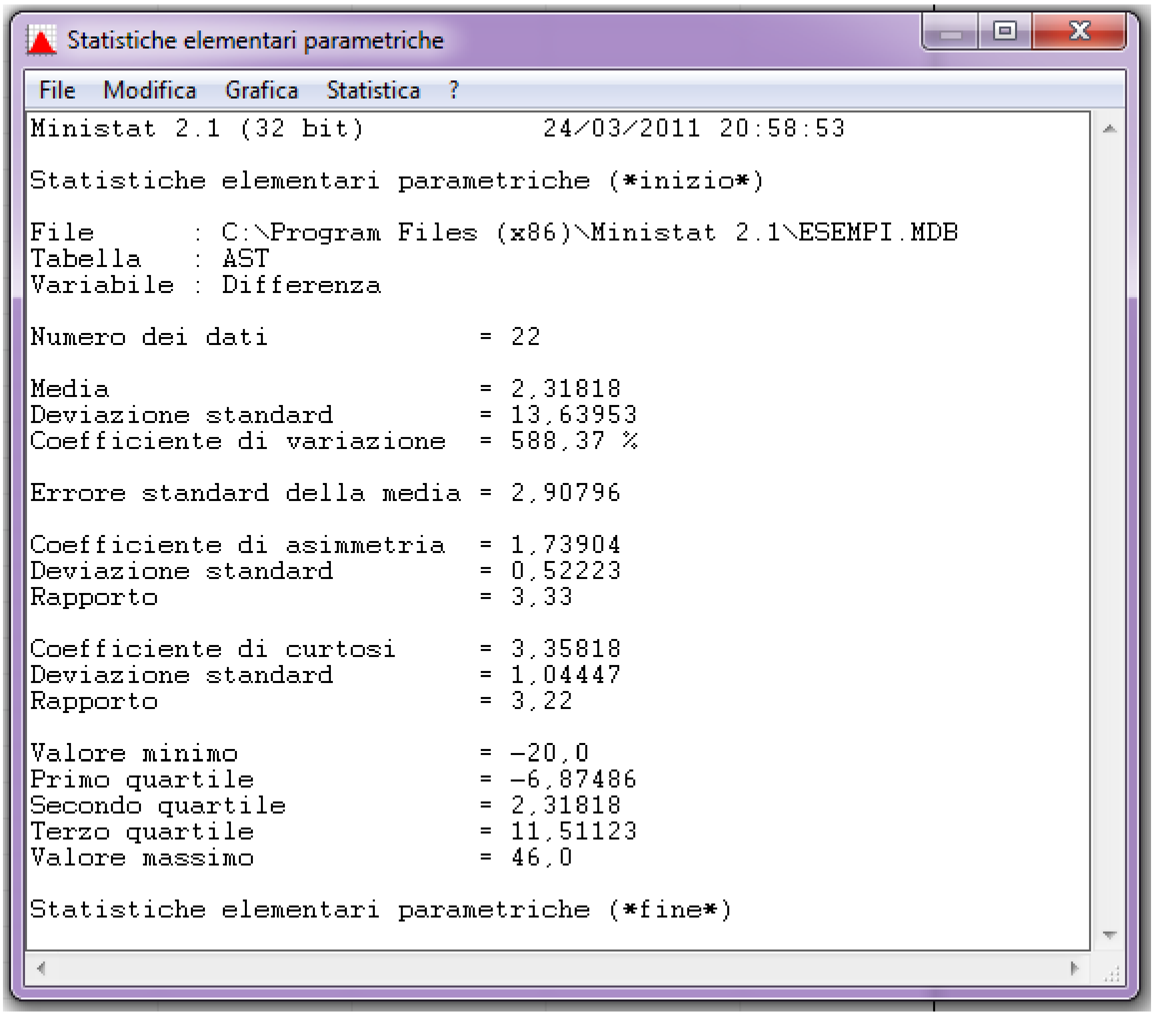
Abbiamo raccolto un certo numero di campioni di sangue, che sono stati immediatamente centrifugati. Per ciascun campione il siero è stato suddiviso in due aliquote identiche, la prima della quali è stata analizzata immediatamente, mentre la seconda, ben tappata, è stata conservata in frigorifero a temperatura controllata compresa tra +2 °C e +8°C. L'indomani, trascorse 24 ore, i campioni sono nuovamente analizzati. La domanda era: l'AST è stabile in un siero conservato per 24 ore in frigorifero a temperatura controllata compresa tra +2 °C e +8°C?

In statistica per rispondere ad una domanda bisogna fare un'ipotesi. Potremmo fare l'ipotesi che l'AST aumenta mediamente di 5 U/L, potremmo fare l'ipotesi che diminuisce mediamente di 10

U/L, potremmo fare l'ipotesi che aumenta mediamente di 13 U/L, e così facendo apriremmo la strada ad infinite ipotesi, tra le quali è impossibile individuare la più plausibile. Per evitare la paralisi indotta da questo approccio, possiamo fare l'ipotesi che l'AST non cambia dopo 24 ore. Sembra un trucco banale, ma in realtà è il solo modo per trarci d'impaccio. E funziona. Vediamo come.

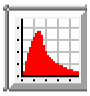
La differenza SUBITO – DOPO24ORE ha una media uguale a 2,31818. I suoi limiti di confidenza, che descrivono l'incertezza che caratterizza la media calcolata, vanno da -2,46541 (il limite inferiore) a 7,10117 (il limite superiore). Qualsiasi valore compreso tra questi due limiti è altrettanto plausibile. I due valori appartenenti a qualsiasi coppia di valori compresi tra questi due limiti risultano tra loro altrettanto plausibili. E due valori appartenenti a qualsiasi coppia di valori compresi tra questi due limiti risultano pertanto tra loro indistinguibili. Questo significa che 0 (la differenza media ipotizzata) e 2,31818 (la differenza media osservata) essendo entrambi compresi tra -2,46541 e 7,10117 sono altrettanto plausibili e tra loro indistinguibili: tecnicamente si dice che la differenza media osservata di 2,31818 non differisce significativamente da 0 (zero).

In realtà un amico pignolo potrebbe farci notare che abbiamo utilizzato per le nostre conclusioni la media, che fa parte delle statistiche (statistiche parametriche) che basano la loro validità sull'assunto che i dati siano distribuiti in modo gaussiano, senza verificare questo assunto.

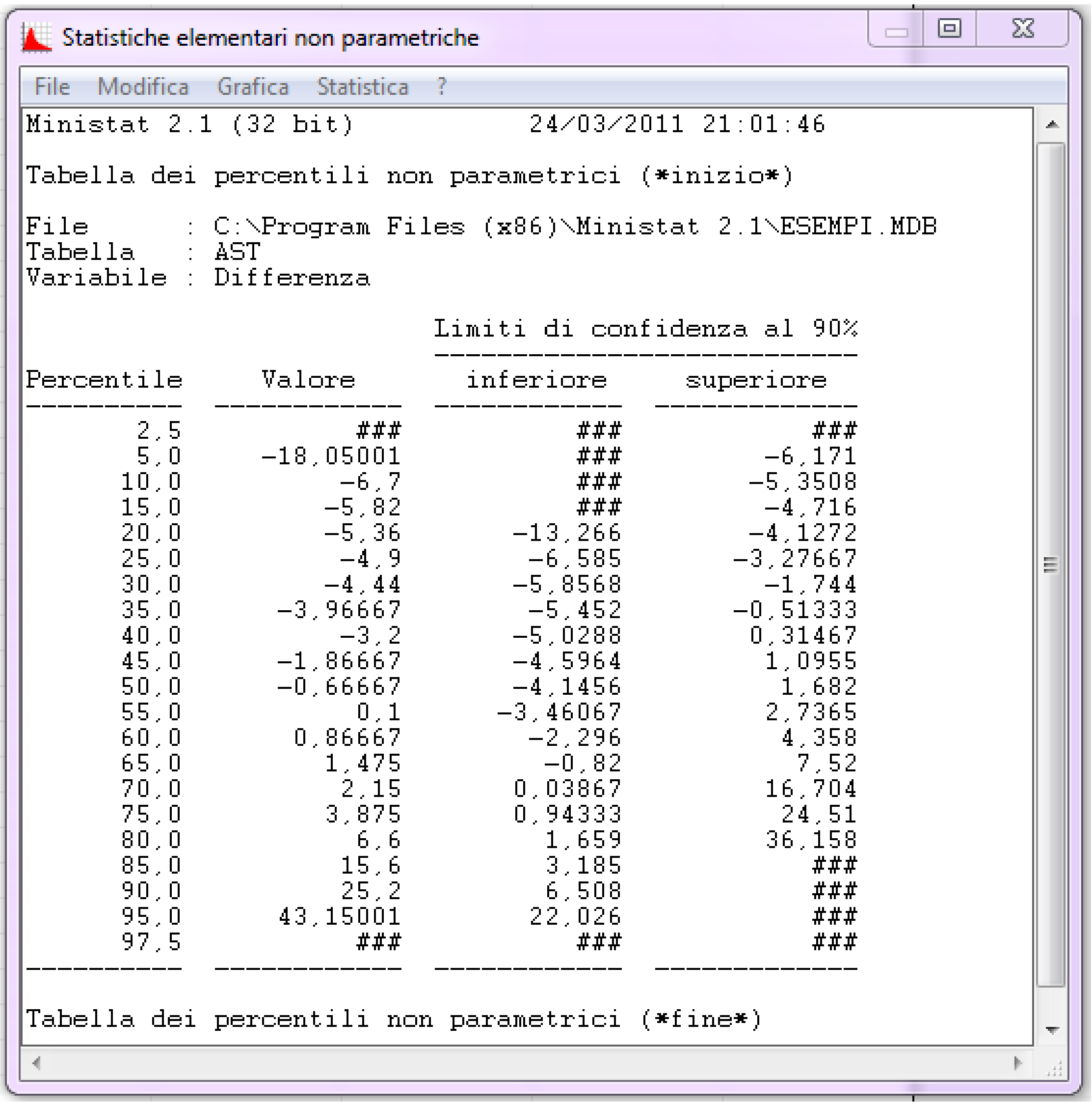



```
Ministat 2.1 (32 bit)          24/03/2011 20:58:53

Statistiche elementari parametriche (*inizio*)

File       : C:\Program Files (x86)\Ministat 2.1\ESEMPI.MDB
Tabella    : AST
Variabile  : Differenza

Numero dei dati              = 22

Media                        = 2,31818
Deviazione standard          = 13,63953
Coefficiente di variazione   = 588,37 %

Errore standard della media  = 2,90796

Coefficiente di asimmetria   = 1,73904
Deviazione standard          = 0,52223
Rapporto                     = 3,33

Coefficiente di curtosi      = 3,35818
Deviazione standard          = 1,04447
Rapporto                     = 3,22

Valore minimo                = -20,0
Primo quartile               = -6,87486
Secondo quartile             = 2,31818
Terzo quartile               = 11,51123
Valore massimo               = 46,0

Statistiche elementari parametriche (*fine*)
```

Possiamo rispondere all'osservazione, assolutamente corretta, selezionando il menù *Statistica* e selezionando *Asimmetria e Curtosi*. Come vedete abbiamo molti dati, ma quelli più importanti sono rappresentati dalla voce *Rapporto*, che compare sia nel *Coefficiente di asimmetria* sia nel *Coefficiente di curtosi*. Se tale rapporto supera 2,7 il coefficiente in questione indica che la distribuzione si discosta eccessivamente dalla gaussiana. In questi casi l'utilizzo delle statistiche parametriche è sconsigliato, e dobbiamo utilizzare statistiche non parametriche. Per dirla più semplicemente nel nostro caso dobbiamo utilizzare la mediana invece della media.

Fate ora click sull'icona che rappresenta una distribuzione non gaussiana 

Dal menù *Statistica* selezionate la *Tabella dei percentili*, che vi apparirà così



Ministat 2.1 (32 bit) 24/03/2011 21:01:46

Tabella dei percentili non parametrici (*inizio*)

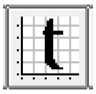
File : C:\Program Files (x86)\Ministat 2.1\ESEMPI.MDB
Tabella : AST
Variabile : Differenza

Percentile	Valore	Limiti di confidenza al 90% inferiore	superiore
2,5	###	###	###
5,0	-18,05001	###	-6,171
10,0	-6,7	###	-5,3508
15,0	-5,82	###	-4,716
20,0	-5,36	-13,266	-4,1272
25,0	-4,9	-6,585	-3,27667
30,0	-4,44	-5,8568	-1,744
35,0	-3,96667	-5,452	-0,51333
40,0	-3,2	-5,0288	0,31467
45,0	-1,86667	-4,5964	1,0955
50,0	-0,66667	-4,1456	1,682
55,0	0,1	-3,46067	2,7365
60,0	0,86667	-2,296	4,358
65,0	1,475	-0,82	7,52
70,0	2,15	0,03867	16,704
75,0	3,875	0,94333	24,51
80,0	6,6	1,659	36,158
85,0	15,6	3,185	###
90,0	25,2	6,508	###
95,0	43,15001	22,026	###
97,5	###	###	###

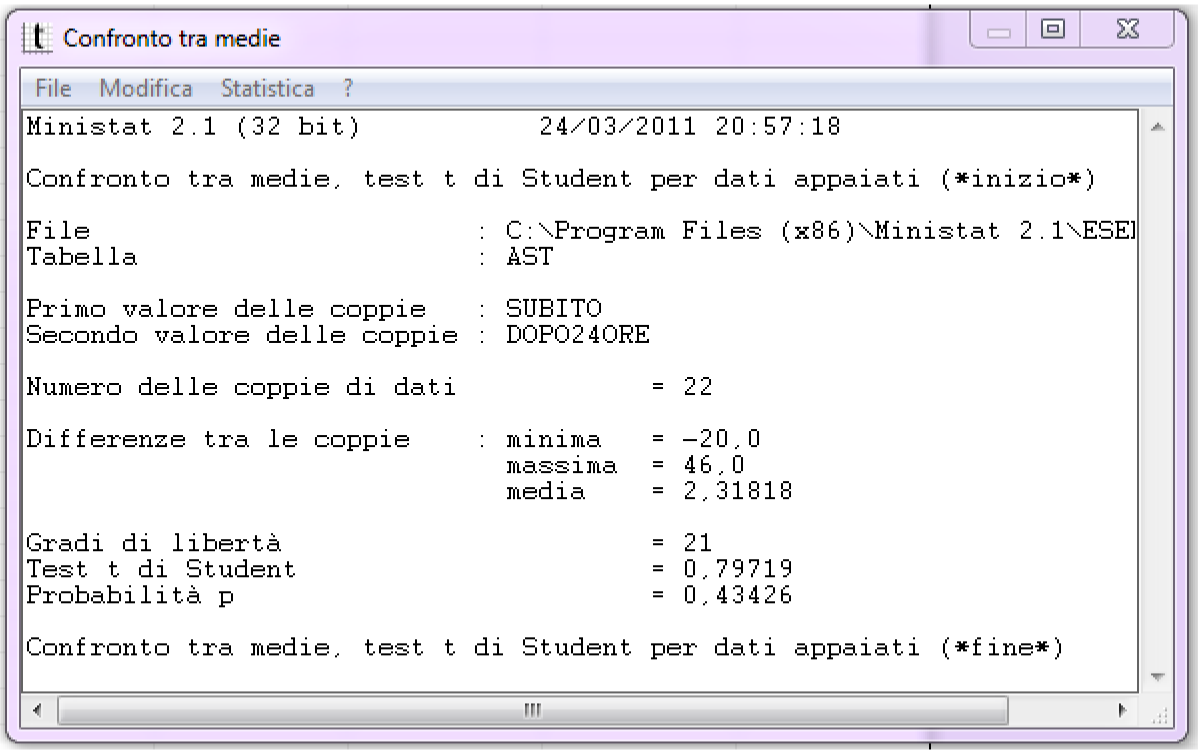
Tabella dei percentili non parametrici (*fine*)

Senza ripetere tutte le considerazioni fatte in precedenza, possiamo verificare rapidamente che 0 (la mediana delle differenze ipotizzata) e -0,66667 (la mediana delle differenze osservata), dati i limiti di confidenza di quest'ultima che vanno da -4,1456 a 1,682 sono tra loro indistinguibili: <u>tecnicamente si dice che la mediana delle differenze osservata di -0,66667 non differisce significativamente da 0 (zero)</u>.

Ci manca un ultimo punto. In effetti sappiamo che è disponibile una forma del test t di Student che consente di confrontare tra di loro le medie di dati appaiati. Potete utilizzarla anche con Ministat, se fate click sull'icona



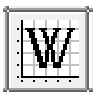
Selezionate quindi il menù *Statistica* e la voce *Test t di Student per dati appaiati*



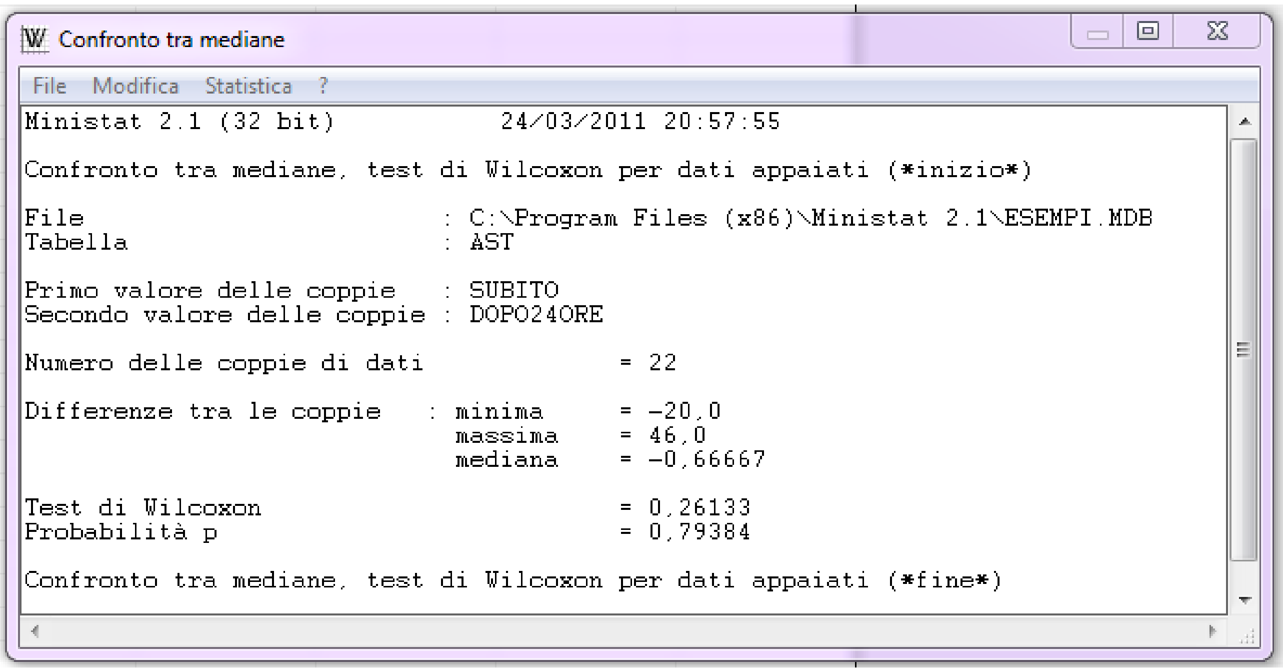
Come vedete nell'ultima riga compare già direttamente il valore di probabilità *p* di osservare per caso una differenza tra le medie pari a 2,31818. In statistica si è soliti accettare una conclusione quando il rischio di sbagliare è basso, in genere inferiore al 5% ($p < 0{,}05$). Qui la probabilità che la differenza tra le medie osservata sia dovuta al caso è elevata, pari al 43% (valore esatto 43,426%). Poiché nessuno accetterebbe il rischio di attribuire a un fatto reale (cioè a un reale cambiamento della concentrazione dell'AST nell'arco delle 24 ore) una differenza che ha una probabilità così elevata di essere dovuto al caso, diciamo che la differenza tra le medie osservata di 2,31818 non differisce significativamente da 0 (zero).

In realtà il nostro amico pignolo potrebbe farci notare che abbiamo utilizzato il test t di Student per dati appaiati, che fa parte delle statistiche che basano la loro validità sull'assunto che i dati siano distribuiti in modo gaussiano, quando sappiamo che questo non è vero nel nostro caso. A questa osservazione, assolutamente corretta, possiamo rispondere utilizzando il test di Wilcoxon per dati appaiati, l'equivalente non parametrico del test t di Student per dati appaiati.

Il test di Wilcoxon non si basa su assunti distribuzionali e quindi fornisce risultati attendibili anche quando i dati non sono distribuiti in modo gaussiano. Potete vederne i risultati facendo click sull'icona



Con il test di Wilcoxon la probabilità *p* di osservare per caso una differenza tra le mediane uguale a -0,66667 è elevata, pari al 79% (valore esatto). Poiché nessuno accetterebbe il rischio di attribuire a un fatto reale (cioè un reale cambiamento della concentrazione dell'AST nell'arco delle 24 ore) una differenza che ha una probabilità così elevata di essere dovuto al caso, diciamo che <u>la differenza tra le mediane osservata di -0,66667 non differisce significativamente da 0 (zero)</u>.


Confronto tra mediane

File Modifica Statistica ?

```
Ministat 2.1 (32 bit)          24/03/2011 20:57:55

Confronto tra mediane, test di Wilcoxon per dati appaiati (*inizio*)

File                        : C:\Program Files (x86)\Ministat 2.1\ESEMPI.MDB
Tabella                     : AST

Primo valore delle coppie   : SUBITO
Secondo valore delle coppie : DOPO24ORE

Numero delle coppie di dati            = 22

Differenze tra le coppie   : minima    = -20,0
                             massima   = 46,0
                             mediana   = -0,66667

Test di Wilcoxon                       = 0,26133
Probabilità p                          = 0,79384

Confronto tra mediane, test di Wilcoxon per dati appaiati (*fine*)
```

5.4. Conclusioni

Lo scopo di questo esercizio era dimostrare come un problema apparentemente banale possa prestarsi a molte considerazioni.

Per rispondere alla domanda *"l'AST è stabile in un siero conservato per 24 ore in frigorifero?"* abbiamo cercato di seguire un percorso logico, che è iniziato con la tabulazione dei dati, è passato dalla loro rappresentazione grafica, per concludersi con l'impiego dei test statistici appropriati.

Abbiamo visto come nel caso di dati appaiati l'informazione originale possa essere conservata in un nuova variabile derivata dalle coppie di dati originali. Abbiamo visto che i limiti di confidenza della nuova variabile possono essere utilizzati per costruire un test di significatività, che fornisce gli stessi risultati di un test specificamente studiato per questi casi (test t di Student per dati appaiati). Abbiamo visto che è sempre necessario, prima di applicare un test statistico, verificare come sono distribuiti i dati: se i dati sono distribuiti in modo gaussiano possono essere utilizzati test parametrici, mentre se i dati non sono distribuiti in modo gaussiano è necessario ricorrere a test non parametrici (per semplicità qui non viene considerata la possibilità di applicare ai dati trasformazioni numeriche in grado di ricondurre distribuzioni non gaussiane a distribuzioni gaussiane). E infine abbiamo visto come esista un equivalente del test t di Student per dati appaiati, il test di Wilcoxon per dati appaiati, che deve essere utilizzato appunto quando l'assunto di gaussianità viene disatteso.

Appendice 1. Test t di Student per dati appaiati

Il test t di Student per dati appaiati è un caso particolare del confronto fra medie. Viene utilizzato quando un disegno sperimentale appropriato garantisce l'omogeneità fra i due campioni posti a confronto, che differiranno fra di loro (nel migliore dei casi) solamente per il fattore che lo sperimentatore ha deliberatamente introdotto. Il metodo è tratto da Snedecor (*Snedecor GW, Cochran WG. Statistical methods. VII Edition. Ames: The Iowa State University Press, 1980:83-89*).

Nel confronto tra medie mediante il test *t* di Student per dati appaiati la differenza fra le coppie di osservazioni diventa la variabile in esame. Dato un numero *n* di dati x_i , y_i sia d_i la differenza (presa con il segno) fra il valore del primo e il valore del secondo elemento della coppia, ovvero

$$d_i = x_i - y_i$$

Allora la differenza media $\bar{d}$ e la varianza s_d^2 delle differenza sono calcolate rispettivamente come

$$\bar{d} = \Sigma\, d_i / n$$
$$s_d^2 = \Sigma\,(d_i - \bar{d})^2 / (n - 1)$$

Si ricorda che anche in questo caso è possibile semplificare il calcolo della varianza s_d^2 ricordando che

$$\Sigma\,(d_i - \bar{d})^2 = \Sigma\, d_i^2 - (\Sigma\, d_i)^2 / n$$

Il valore del test t di Student è calcolato come rapporto fra la differenza media osservata e il suo errore standard, cioè

$$t = \bar{d} / \sqrt{(s_d^2 / n)}$$

con $n - 1$ gradi di libertà.

Il valore di *p* corrispondente alla statistica *t* rappresenta la probabilità di osservare per caso una differenza della grandezza di quella effettivamente osservata: se tale probabilità è sufficientemente piccola, si conclude per una significatività della differenza tra le medie.

Appendice 2. Test di Wilcoxon per dati appaiati

Il test di Wilcoxon per dati appaiati è l'equivalente non parametrico del test t di Student per dati appaiati, e va utilizzato in luogo di questo quando i dati non siano distribuiti in modo gaussiano. La soluzione qui utilizzata per il calcolo della significatività è sufficientemente accurata per n > 16. Il metodo è tratto da Snedecor (*Snedecor GW, Cochran WG. Statistical methods. VII Edition. Ames: The Iowa State University Press, 1980:141-3*).

Per i calcoli si procede in questo modo :
→ determinare le differenze (con il segno) fra le *n* coppie di valori;
→ stabilire, per ciascuna differenza, il numero di posizione nella lista delle differenze ordinate (questa volta ignorando il segno) in ordine numerico crescente: la più piccola differenza osservata avrà numero di posizione 1, e via dicendo;
→ quando due o più differenze sono uguali, assegnare a ciascuna di esse la media dei numeri di posizione che esse dovrebbero avere; così, per esempio, se la quinta e la sesta differenza sono uguali, assegnare come numero di posizione nella lista il valore 5.5 a entrambe;
→ riassegnare il segno ai numeri di posizione nella lista (se la differenza avente quel numero di posizione era negativa, assegnare il segno meno, se era positiva assegnare il segno più);
→ calcolare il totale per i numeri di posizione con segno negativo e per i numeri di posizione con segno positivo, e chiamare *T* il più piccolo di questi due totali;
→ calcolare la deviata normale standardizzata *Z* come

$$Z = (\mu - T - 0{,}5) / s$$

essendo

$$\mu = n \cdot (n + 1) / 4$$
$$s = \sqrt{((2n + 1) \cdot \mu / 6)}$$

La statistica *Z* così calcolata corrisponde a sottoporre al test la mediana delle differenze.

Il valore di *p* corrispondente alla statistica *Z* rappresenta la probabilità di osservare per caso una differenza della grandezza di quella effettivamente osservata: se tale probabilità è sufficientemente piccola, si conclude per una significatività della differenza fra le mediane. Questa soluzione è sufficientemente accurata per n > 16.