



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Tecnica e Gestione dei sistemi industriali

Corso di Laurea in Ingegneria Gestionale

Tesi di Laurea

Il controllo statistico della qualità

tramite il software Minitab

Relatore
Ch.mo Prof. Luigi Salmaso

Laureando
Giorgio Belvedere

Anno Accademico 2016/2017



Sommario

Introduzione	8
Capitolo 1.....	9
Quality planning tools (Strumenti statistici per il controllo statistico di un processo produttivo)	9
1.1– Run chart	9
1.2 – Pareto chart	14
1.3 – Fishbone diagram (diagramma causa-effetto)	15
1.4 – Multi-Vari chart	17
1.5 – Symmetry plot	19
Capitolo 2.....	20
Measurement Systems Analysis (Analisi dei sistemi di misura)	20
2.1 – Gage R&R study crossed and nested	22
2.2 – Gage run chart	26
2.3 – Gage linearity and accuracy study	27
Capitolo 3.....	29
Variables control charts (Le carte di controllo)	29
3.1 – Che cosa sono e a cosa servono?	29
3.2 – Come inserire i dati nel foglio di lavoro.....	30
3.3 – Special causes (cause speciali).....	31
3.4 - Variables control charts for data in subgroups (per dati raggruppati in sottogruppi).....	32
3.4.1 – R chart.....	33
3.4.2 – S chart	34
3.4.3 – X medio chart.....	34
3.4.4 - X medio and R chart e X medio and S chart	36
3.4.5 - I-MR-R/S (Between/within) chart.....	37
3.5 Variables control charts for individual observations (per osservazioni individuali).....	37
3.5.1 - Moving Range (MR) chart.....	38
3.5.2 - Individual (I) chart;	40
3.5.3 - I-MR chart.....	40
3.6 - Variables control charts using subgroups combinations (usanti combinazioni di sottogruppi)	41



3.6.1 – EWMA chart.....	42
3.6.2 - Moving Average chart	43
3.6.3 – CUSUM chart	44
3.6.4 – Zone chart.....	45
3.7 - Variables control charts for Short Runs (per piccole quantità di dati).....	46
3.7.1 – Z-MR chart	46
Capitolo 4.....	47
Distribution analysis	47
4.1 – Come inserire i dati nel foglio di lavoro.....	48
4.2 – Distribution ID plot	50
4.2.1 – Distribution ID plot – Right censoring.....	50
4.2.2 - Distribution ID plot – Arbitrary censoring	51
4.3 – Distribution overview plot.....	54
4.3.1 - Distribution overview plot – Right censoring.....	55
4.3.2 - Distribution overview plot – Arbitrary censoring.....	56
4.4 – Parametric Distribution Analysis	57
4.4.1 - Parametric distribution analysis – Right censoring	58
4.4.2 – Parametric distribution analysis – Arbitrary censoring	58
4.5 – Nonparametric distribution analysis	65
Capitolo 5.....	68
Regression with life data	68
5.1 - Come inserire i dati nel foglio di lavoro	68
5.2 – Accelerated life testing.....	69
5.2.2 - Percentiles and survival probabilities.....	72
5.3 - Regression with life data	76
5.3.1 - Come inserire i dati nel foglio di lavoro.....	76
5.3.2 How to specify the model terms	78
5.3.3 Probability plots.....	78
5.3.4 Multiple degrees of freedom test	78
5.3.5 Estimating the model parameters.....	79



Università di Padova – Facoltà di Ingegneria
Giorgio Belvedere – matricola n. 1051159

Bibliografia.....	82
Sitografia.....	82



Lista delle tabelle

Tabella 1.0.1 Test cause speciali.....	10
Tabella 1.0.2 Struttura foglio di lavoro diagramma causa-effetto.....	16
Tabella 2.0.1 Crossed con metodo ANOVA.....	25
Tabella 2.0.2 Crossed con metodo ANOVA.....	25
Tabella 3.0.1 Tipologie di carte di controllo.....	30
Tabella 3.0.2 Dati dell'esempio 3.4.3.....	39
Tabella 3.0.3 Dati per esempi.....	41
Tabella 3.0.4 Dati per esempi.....	41
Tabella 4.0.1 I tipi dei tempi di guasto.....	48
Tabella 4.0.2 Struttura del foglio di lavoro con indicatori di censura.....	49
Tabella 4.0.3 Struttura del foglio di lavoro con indicatori di frequenza.....	50
Tabella 4.0.4 Tabella risultante di un Distribution ID plot.....	54
Tabella 4.0.5 Parameter Estimates.....	59
Tabella 4.0.6 Characteristics of distribution.....	60
Tabella 4.0.7 Table of percentiles.....	60
Tabella 4.0.8 Table of Survival Probabilities.....	61
Tabella 4.0.9 Table of percentiles dell'esempio 4.4 temperatura 80°C.....	64
Tabella 4.0.10 Table of percentiles dell'esempio 4.4 temperatura 100°C.....	64
Tabella 4.0.11 Table of survival probabilities dell'esempio 4.4 temperatura 80°C.....	64
Tabella 4.0.12 Table of survival probabilities dell'esempio 4.4 temperatura 100°C.....	64
Tabella 5.0.1 Regression table dell'esempio 5.1.....	74
Tabella 5.0.2 Table of percentiles dell'esempio 5.1.....	74
Tabella 5.0.3 Regression table esempio 5.2.....	80
Tabella 5.0.4 Table of percentiles esempio 5.2.....	80



Lista delle figure

Figura 1.0.1 Esempio di Mixture.....	11
Figura 1.0.2 Esempio di clusters.....	11
Figura 1.0.3 Esempio di oscillation.....	12
Figura 1.0.4 Esempio di trend.....	12
Figura 1.0.5 Finestra di dialogo Run chart.....	13
Figura 1.0.6 Run chart.....	14
Figura 1.0.7 Esempio di grafico di Pareto.....	15
Figura 1.0.8 Diagramma causa-effetto.....	17
Figura 1.0.9 Multi-vari chart dell'esempio 1.4.....	18
Figura 1.0.10 Esempio symmetry plot.....	19
Figura 2.0.1 Rappresentazione della varianza totale.....	21
Figura 2.0.2 Esempio Crossed e Nested.....	22
Figura 2.0.3 Grafico crossed con metodo ANOVA.....	26
Figura 2.0.4 Esempio di Gage run chart.....	27
Figura 2.0.5 Esempio di Gage linearity and accuracy study.....	28
Figura 3.0.1 Carta di controllo.....	29
Figura 3.0.2 Finestra di dialogo principale di una carta di controllo generica.....	31
Figura 3.0.3 Finestra "Test" di una generica carta di controllo.....	32
Figura 3.0.4 R chart.....	33
Figura 3.0.5 Finestra di dialogo principale X medio chart, esempio 3.4.2.....	35
Figura 3.0.6 X medio chart dell'esempio 3.4.2.....	36
Figura 3.0.7 Esempio di MR chart.....	39
Figura 3.0.8 Finestra Options EWMA.....	42
Figura 3.0.9 EWMA chart.....	43
Figura 3.0.10 Esempio di Zone chart.....	45
Figura 4.0.1 Esempio di finestra di dialogo principale di un Distribution ID plot.....	51
Figura 4.0.2 Probability plot del Distribution ID plot.....	53
Figura 4.0.3 Esempio di una finestra di dialogo principale di un Distribution Overview Plot.....	55
Figura 4.0.4 Distribution overview plot.....	57
Figura 4.0.5 Probability plot.....	61
Figura 4.0.6 Parametric hazard plot.....	62
Figura 4.0.7 Procedimento dell'esempio 4.4.....	63
Figura 4.0.8 Probability plot dell'esempio 4.4.....	65
Figura 4.0.9 Survival plot dell'esempio 4.4.....	65
Figura 5.0.1 Finestra di dialogo principale dell'opzione Estimate dell'Accelerated life testing.....	72
Figura 5.0.2 Probability plot dell'esempio 5.1.....	75
Figura 5.0.3 Relation plot dell'esempio 5.1.....	75
Figura 5.0.4 Esempi di how to specify the model terms.....	78



Figura 5.0.5 Probability plot esempio 5.2.....81



Introduzione

Lo scopo di questa tesi è mostrare come utilizzare il software statistico Minitab, attraverso i comandi principali, al fine di controllare i vari ambiti aziendali, con particolare attenzione al controllo della qualità.

Per far questo, in prima battuta, ho studiato e tradotto il manuale d'uso del software, per poi applicare quanto imparato ai dati raccolti nell'azienda di famiglia, così da poter fornire esempi utili e reali a chi vorrà consultare il mio lavoro.

La conoscenza e l'utilizzo di Minitab permettono di analizzare e valutare a fondo tutti gli aspetti operativi della propria azienda, così da diagnosticarne i problemi e consentire di proporre le soluzioni più opportune per il miglioramento degli standard qualitativi aziendali, così come un medico farebbe col proprio paziente.



Capitolo 1

Quality planning tools (Strumenti statistici per il controllo statistico di un processo produttivo)

I quality planning tools sono un insieme di strumenti, *“utili per raggiungere la stabilità del processo produttivo e per migliorare la produttività attraverso la riduzione della variabilità”* (Cfr. Cap.4 pag. 137-138, da “Controllo statistico della qualità”) dei parametri nominali del processo o dei pezzi prodotti. Raggiungere la stabilità del processo e ridurre la variabilità dei parametri nominali sono condizioni necessarie affinché un prodotto possa soddisfare le esigenze del produttore e le aspettative dei consumatori.

Tuttavia, l’efficacia di questi strumenti dipende da come la dirigenza aziendale li assimila ed applica: *“solo una loro applicazione continua e sistematica consente di farne parte integrante del modo di pensare alla programmazione della produzione e al miglioramento della qualità”* del prodotto. (Cfr. Cap.4 pag. 137-138, da “Controllo statistico della qualità”).

Minitab offre i seguenti cinque strumenti grafici per aiutare a esplorare e rintracciare i problemi nella qualità del processo produttivo:

- il **Run chart**: serve a rintracciare schemi ricorrenti nei dati di processo e ad effettuare due test per la verifica di comportamento non casuale dei dati;
- il **Pareto chart** (diagramma di Pareto): aiuta ad identificare i problemi più significativi delle fasi del processo produttivo o suscettibili di un più ampio margine di miglioramento;
- il **Fishbone diagram** (diagramma causa-effetto): aiuta ad organizzare e focalizzare le cause più rilevanti dei problemi del processo;
- il **Multi-vari chart**: mostra l’analisi della varianza dei dati per via grafica, consentendo una analisi immediata sulla distribuzione dei dati;
- il **Simmetry plot**: aiuta a verificare se i dati appartengono ad una distribuzione simmetrica.

1.1– Run chart

Il Run chart grafica tutte le osservazioni individuali di un insieme di dati in funzione del numero del rispettivo sottogruppo (o sottoinsieme) di appartenenza e traccia la linea orizzontale della mediana come riferimento; quando la dimensione campionaria dei sottogruppi è maggiore di uno, grafica (con un punto blu) la media o la mediana dei sottogruppi e le connette con una linea. Per vederne un esempio si veda la figura 1.0.6.



Il Run chart serve a rintracciare la presenza di schemi (pattern) non casuali presenti nei dati, i quali dovrebbero disporsi in modo casuale attorno al valor medio, e attua due test per verificare la loro presenza.

Test for randomness	Condition	Indicates
number of runs about the median	more runs observed than expected	mixed data from two population
	fewer runs observed than expected	clustering of data
number of runs up or down	more runs observed than expected	oscillation—data varies up and down rapidly
	fewer runs observed than expected	trending of data

Tabella 1.0.1 Test cause speciali

I due test per verificare la presenza di comportamenti non casuali nei dati rintracciano i seguenti comportamenti: trend, oscillazioni, mescolanze e raggruppamenti.

Questi comportamenti sono sintomo di «cause speciali», ovvero di cause derivanti dall'esterno del processo e che dovrebbero essere eliminate. Esse rendono il processo fuori controllo. Al contrario, un processo è sotto controllo quando nessuna «causa speciale», bensì solo cause comuni, influenzano il processo.

Vediamo nel dettaglio i due test.

Il primo test è basato sul numero totale di runs che si verificano sia sopra sia sotto la mediana. Un run, in questo caso, consiste in uno o più punti consecutivi dalla stessa parte della mediana.

Il test per il numero di runs attorno alla mediana è sensibile a due comportamenti non casuali:

- **mixtures:** un numero di runs osservato statisticamente più grande del numero di runs atteso supporta il mixing;
- **clustering:** un numero di runs osservato statisticamente minore del numero atteso di runs supporta l'alternativa di clustering.

Il risultato di questo test è riportato nelle ultime due righe del riquadro di sinistra delle figure 1.0.1 e 1.0.2.

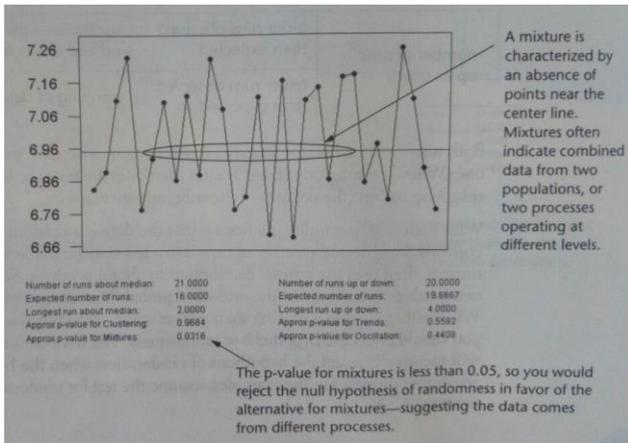


Figura 1.0.1 Esempio di Mixture

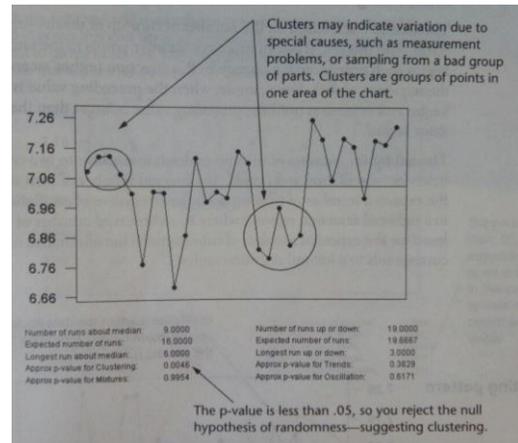


Figura 1.0.2 Esempio di clusters

Il secondo test è basato sul numero di runs che crescono o decrescono. Un run, in questo test, consiste in uno o più punti consecutivi nella medesima direzione (o crescente o discendente) nella sequenza dei dati. Per esempio, quando il precedente valore è più piccolo, un run up comincia e continua finché il valore precedente non è più grande del punto successivo, allora un run down avrà inizio.

Il test per il numero di runs crescenti o decrescenti è sensibile a due tipi di comportamento non casuale:

- **oscillation:** un numero di runs osservate statisticamente più grande del numero aspettato di corse supporta l'oscillation;
- **trends:** un numero di corse osservato statisticamente più piccolo del numero osservato di corse supporta l'alternativa trend.

Il risultato di questo test è riportato nelle ultime due righe del riquadro di destra della figura 1.0.3 e 1.0.4.

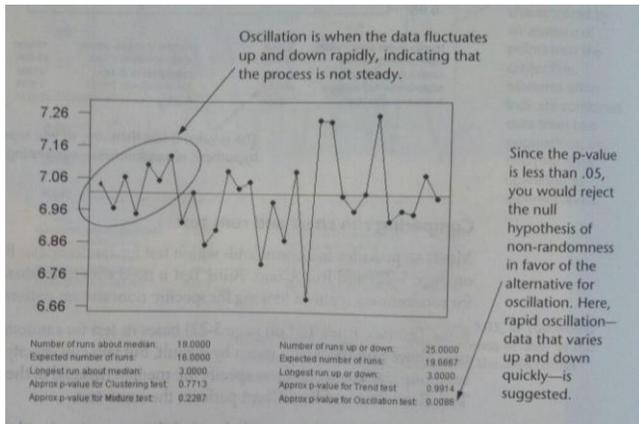


Figura 1.0.3 Esempio di oscillation

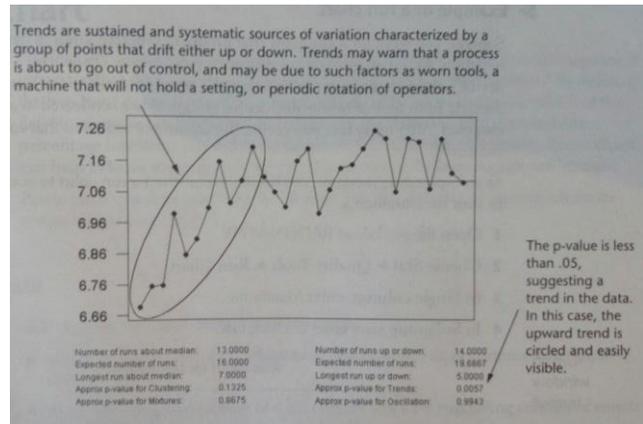


Figura 1.0.4 Esempio di trend

I dati da utilizzare col Run chart possono essere dati individuali oppure sottoinsiemi di dati.

Nel foglio di lavoro, i sottoinsiemi possono essere strutturati in:

- righe attraverso più colonne, se hanno numerosità campionaria costante;
- una singola colonna, se hanno numerosità campionaria variabile.

Quando si hanno sottoinsiemi di numerosità diversa, tutti i sottogruppi vanno inseriti in una singola colonna, per poi creare una seconda colonna (colonna di indicatori) indicante, per ogni dato (quindi per ogni riga), il sottoinsieme di appartenenza; nell'esempio precedente è la colonna "Subgroup".

In Minitab, per aprire la finestra di dialogo principale del Run chart, bisogna seguire la seguente procedura: Stat → Quality tools → Run chart. Così apparirà una finestra, in cui si potrà:

- quando i tutti i sottogruppi di dati o i dati individuali sono raggruppati in una unica colonna, inserire questa colonna in «Single column». In «subgroup size», inserire: nel caso di sottogruppi di uguale dimensione campionaria, la dimensione dei sottogruppi; nel caso di sottogruppi di dimensione variabile, la colonna di indicatori del sottogruppo di appartenenza dei dati; nel caso di osservazioni individuali, inserire 1;
- Quando i sottogruppi sono in righe, inserire la serie di colonne in «subgroups across rows of».



Esempio 1.1

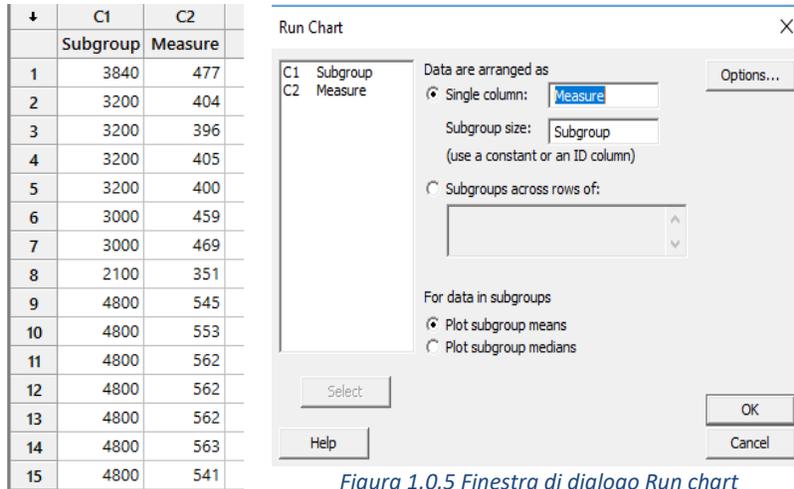


Figura 1.0.5 Finestra di dialogo Run chart

Si supponga di voler controllare le tempistiche del processo produttivo nel produrre cinque modelli diversi. Così si raccolgono, per ogni modello, i tempi (in secondi) macchina di lavorazione e li si inserisce nel foglio di lavoro sotto la colonna “Measure”. Dopo di che, per indicare il prodotto a cui fa riferimento ciascuna misura, si inserisce la colonna “Subgroup”.

Infine, avremo un foglio di lavoro organizzato come mostrato sopra.

Per creare il Run chart di queste misure apriamo la finestra di dialogo principale e la compiliamo come mostrato alla figura 1.0.5. Cliccando «OK», apparirà il grafico della figura 1.0.6.

Come si vede, il grafico presenta i diversi sottogruppi sull’asse delle ascisse, mentre lungo l’asse delle ordinate le misure dei tempi. Inoltre, i punti blu rappresentano le medie per ogni campione, mentre i punti grigi le singole osservazioni.

Dal grafico deduciamo che i tempi di lavorazione da modello a modello possono variare significativamente. Inoltre, a prima vista, i dati sembrano disporsi casualmente: quindi il processo sembra essere sotto controllo.

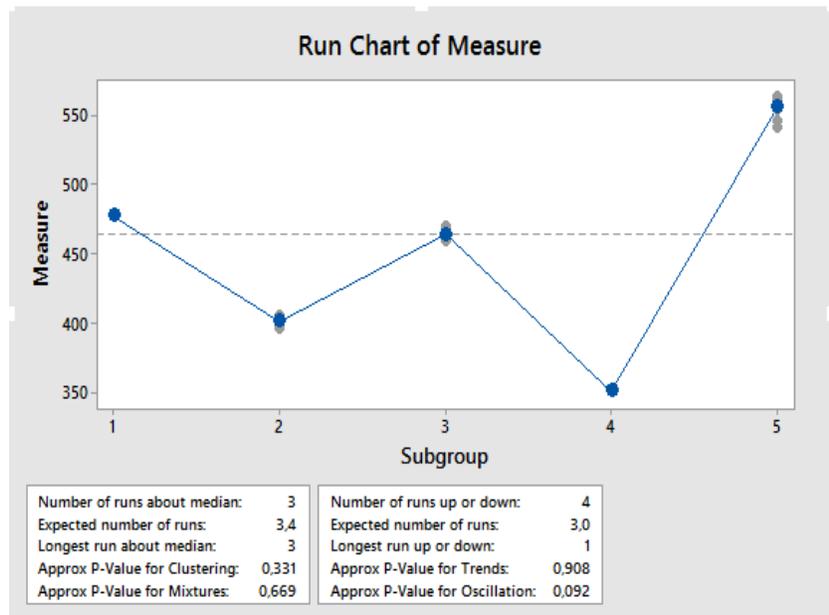


Figura 1.0.6 Run chart

1.2 – Pareto chart

Il Pareto chart (grafico di Pareto) è un grafico in cui lungo l'asse orizzontale vengono elencate delle categorie di interesse (come per esempio il tipo di difetti riscontrabili in un pezzo prodotto), mentre sull'asse verticale una scala di numeri reali o interi e una percentuale (come per esempio il numero di difetti al mese riscontrati nei pezzi prodotti).

Ad ogni categoria di interesse corrisponde un certo valore sull'asse verticale e ciò viene rappresentato tramite barre verticali, come per gli istogrammi. Quest'ultime vengono ordinate in modo decrescente; così il diagramma di Pareto aiuta a determinare quali, tra tutte le categorie di interesse, sono le più frequenti o le più rilevanti e quali lo sono meno. In aggiunta, sopra le barre, viene tracciata una curva percentuale cumulativa che mostra il contributo aggiuntivo dato da ogni ulteriore categoria.

Il diagramma di Pareto aiuta a individuare le categorie di interesse più rilevanti e quindi degne della maggior attenzione.

Per esempio, nella figura 1.0.7, siccome le categorie di interesse corrispondono a vari tipi di difetti, allora il Pareto chart mostra i difetti più frequenti e quindi quelli che vanno risolti o contenuti per primi.

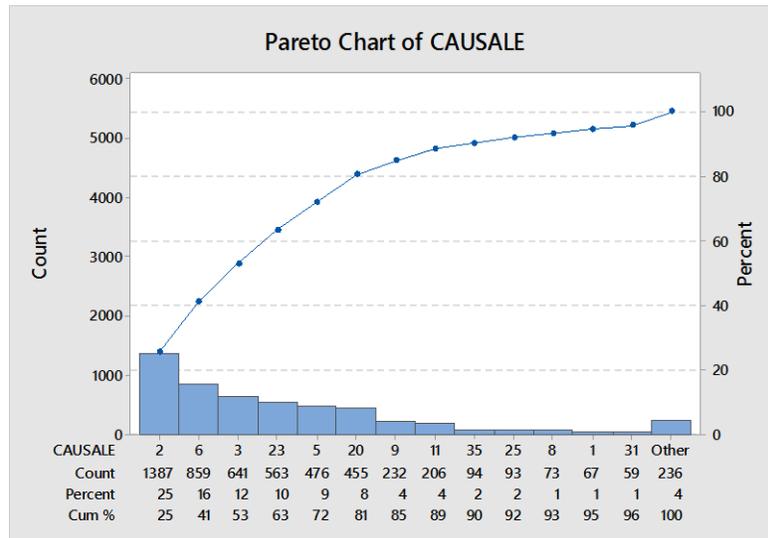


Figura 1.0.7 Esempio di grafico di Pareto

Nel foglio di lavoro, i dati possono essere raggruppati in due modi:

- tutti in una colonna;
- in due colonne: una contenente il nome della categoria di interesse e l'altra contenente il corrispettivo conteggio.

Per aprire la finestra di dialogo principale di un diagramma di Pareto, bisogna seguire la seguente procedura: Stat → Quality tools → Pareto chart.

Nella finestra di dialogo si potrà:

- se si ha una colonna di dati, inserire tale colonna in «Chart defects data in»;
- se si ha una colonna con i nomi delle categorie di interesse e una colonna coi conteggi: spuntare «chart defects table»;
- in «labels in» inserire la colonna coi nomi dei difetti; in «frequencies in», inserire la colonna coi conteggi.

1.3 – Fishbone diagram (diagramma causa-effetto)

Questo diagramma aiuta ad organizzare le idee sulle possibili cause di un problema e ad analizzare le possibili relazioni tra le cause stesse.

Dalla figura 1.0.8, si può notare come i «rami» dell'albero siano associati a delle possibili cause generali del problema, mentre su ogni ramo sono presenti alcune cause specifiche del problema. All'apice destro della linea rossa è mostrato il nome del problema principale.



Prima di aprire la finestra di dialogo del diagramma, bisogna disporre i dati in modo tale che nella prima casella di ogni colonna (quella non numerata) si indichi una causa generale e nelle celle sottostanti le cause specifiche di quella causa generale, come nella tabella 1.0.2.

↓	C1-T	C2-T	C3-T	C4-T	C5-T	C6-T
	Measurements	Materials	Personnel	Environment	Methods	Machines
1	Micrometers	Alloys	Shifts	Moisture %	Brake	Sockets
2	Microscopes	Lubricants	Supervisors	Condensation	Engager	Bits
3	Inspectors	Suppliers	Training		Angle	Lathes
4			Operators			Speed
5						

Tabella 1.0.2 Struttura foglio di lavoro diagramma causa-effetto

Per aprire la finestra di dialogo di questo diagramma bisogna seguire la seguente procedura: Stat → Quality tools → Cause-and-effect...

Fatto ciò apparirà una finestra in cui si potranno selezionare le cause generali elencate a sinistra e inserirle nelle celle bianche sotto il nome «Causes». Automaticamente anche la colonna «Labels» verrà completata dalle cause specifiche prima inserite nel foglio di lavoro.

Fatto ciò si potrà inserire:

- in «Effect» il nome del problema;
- in «Title» intitolare a piacimento il grafico e, volendo, spuntare le due opzioni sottostanti.

Esempio 1.3

Nell' esempio 1.2 del Pareto chart, si era identificata, come causa principale di difetto, la causale 2, ovvero la scheggiatura sul bordo. Così, si è deciso di investigarne le cause.

Dopo aver consultato gli operatori del reparto molatura, si sono identificate quattro principali cause di scheggiatura: il tipo di mola, la quantità d'acqua per raffreddare la zona di contatto vetro-mola, il vetro stesso, la velocità di asportazione.



Ognuna di queste cause generali ha delle sotto cause, indicate nel diagramma dell'esempio.

1.4 – Multi-Vari chart

Questo è un grafico multi-variabile che traccia, tramite punti, il valore di una variabile dipendente in funzione di quattro o meno variabili indipendenti, dette fattori.

I multi-vari chart sono un modo grafico per rappresentare l'analisi della varianza dei dati. Questi grafici possono anche essere usati nelle fasi iniziali dell'analisi.

Questo grafico richiede che tutte le combinazioni tra i livelli dei quattro fattori vengano esplorate attraverso i dati raccolti. Per esempio, considerando un grafico multi-variabile a due variabili, ogni valore assunto da ciascuna delle due variabili va combinato con tutti gli altri valori assunti dall'altra variabile. Ciò può rendere ardua l'analisi.

Nel foglio di lavoro, i dati vanno disposti in modo tale che ogni riga contenga i dati di una singola osservazione, mentre:

- in una colonna saranno riportati i dati numerici della risposta o della variabile dipendente;
- e in massimo altre quattro colonne saranno riportati i dati delle variabili indipendenti, che possono essere testo, numeriche, data o tempo e i loro sotto-livelli.

Fatto questo, si potrà aprire la finestra di dialogo principale del Multi-Vari chart seguendo la procedura: Stat → Quality tools → Multi-Vari chart.

In questa finestra si dovrà:

- in «Response», inserire la colonna delle misurazioni della variabile dipendente;
- in ciascuno dei quattro «Factor», la colonna di una certa variabile indipendente.

Fatto ciò si clicca «OK» ed apparirà un grafico simile a quello dell'esempio successivo.

Esempio 1.4

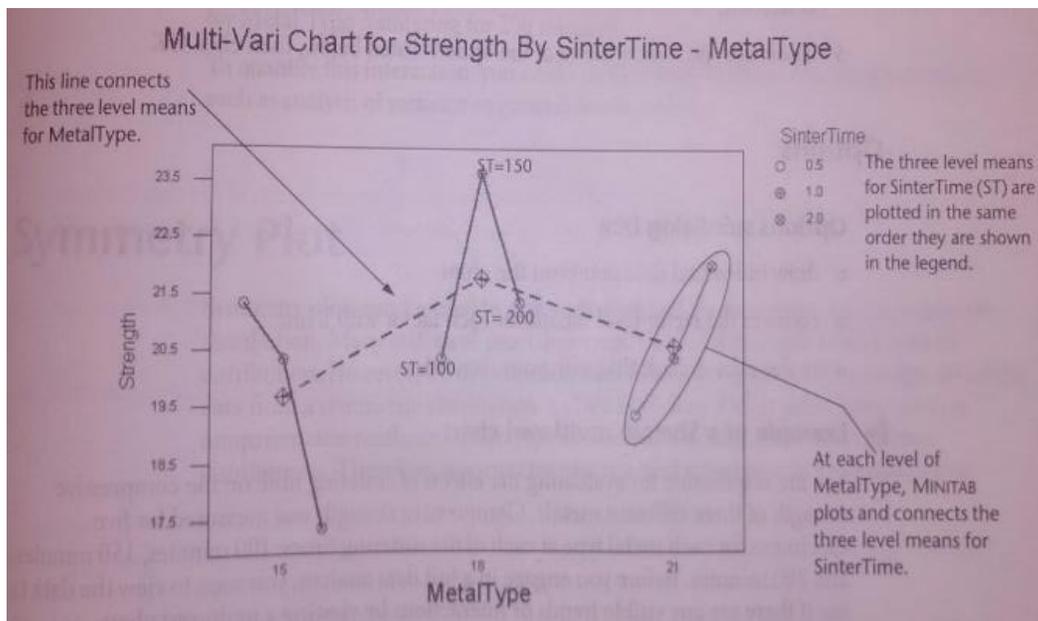


Figura 1.0.9 Multi-vari chart dell'esempio 1.4

Supponiamo di voler studiare la resistenza a compressione di tre tipi di metallo (15, 18 e 21), tramite tre provini per ciascun tipo, dopo aver sottoposto ciascun provino a tre tempi di sinterizzazione diversi ma uguali tra i tipi di metallo.

In questo caso, abbiamo due fattori: tipo di metallo («Metal type») e tempo di sinterizzazione («ST»). Ciascun fattore ha solo tre valori (livelli) che può assumere: Metal type può essere 15, 18 o 21; ST può essere 0,5, 1 e 2.

Mostrando graficamente la resistenza a compressione del metallo in funzione dei due fattori, otterremo il grafico di figura 1.0.9, il quale mostra come varia la resistenza del metallo variando tipo di metallo e tempo di sinterizzazione a cui è stato sottoposto.



1.5 – Symmetry plot

Questo strumento è utilizzato per verificare se un insieme di dati proviene da una distribuzione simmetrica o asimmetrica. Verificare ciò è importante perché per molte procedure è sufficiente che i dati siano distribuiti simmetricamente, anche se non sono proprio distribuiti normalmente.

I dati analizzabili da questo grafico devono essere di tipo numerico.

Per aprire la finestra di dialogo principale di questo grafico bisogna seguire la seguente procedura: Stat → Quality tools → Symmetry Plot. Fatto ciò in «Variables» va inserita la colonna contenente i dati da analizzare. Poi, cliccando “OK”, comparirà un grafico simile alla figura 1.0.10, che si legge notando come i dati si dispongono rispetto alla linea rossa di riferimento e guardando l’istogramma a destra del grafico.

Vi sono tre possibilità:

- se i dati si dispongono approssimativamente lungo la linea rossa, allora i dati avranno distribuzione approssimativamente simmetrica;
- se i dati divergono al di sopra della linea rossa, allora saranno distribuiti asimmetricamente verso destra;
- se i dati divergono al di sotto della linea rossa, allora si distribuiranno asimmetricamente verso sinistra, come nel grafico esemplificativo.

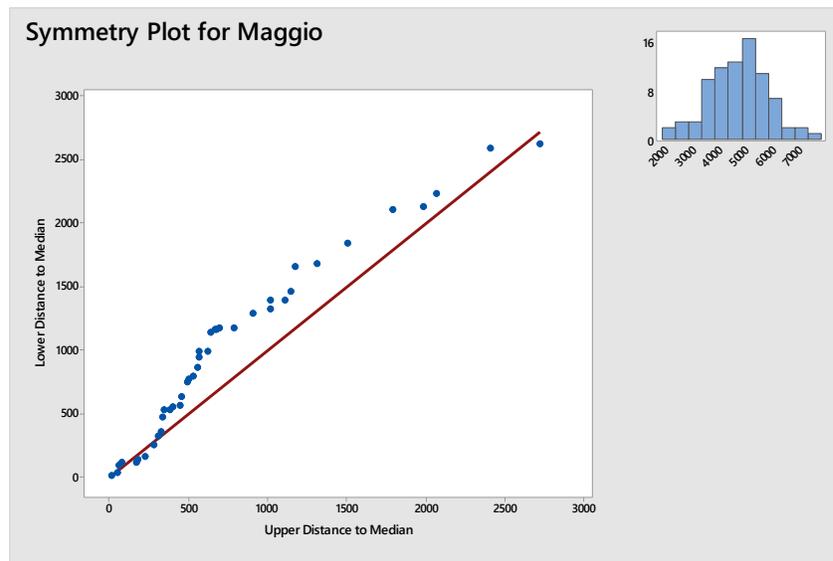


Figura 1.0.10 Esempio symmetry plot



Capitolo 2

Measurement Systems Analysis (Analisi dei sistemi di misura)

Determinare la capacità del sistema di misurazione è importante al fine di perseguire il miglioramento del processo, quindi della qualità. In generale, in qualunque operazione che coinvolga misurazioni, parte della variabilità osservata sarà imputabile agli elementi o agli oggetti che si misurano e parte deriverà dal sistema di misurazione che si è utilizzato. Tale sistema consisterà di uno strumento (gauge) e spesso anche di un operatore che utilizza lo strumento.

Lo scopo della maggior parte degli studi sulla capacità dei sistemi di misurazione riguardano:

1. determinare quanto della variabilità osservata totale è dovuta allo strumento;
2. isolare le componenti di variabilità del sistema di misurazione;
3. valutare se il mezzo di misurazione è adatto all'utilizzo che ne si sta facendo.

In ogni problema legato a procedure di misura, parte della variabilità dei risultati può essere dovuta alla variabilità propria dei prodotti ma anche alla variabilità derivante da strumenti di misura. Si ha quindi:

$$\sigma_{totale}^2 = \sigma_{prodotto}^2 + \sigma_{strumento}^2$$

Dove σ_{totale}^2 è la variabilità osservata, $\sigma_{prodotto}^2$ è la componente di varianza dovuta al prodotto e $\sigma_{strumento}^2$ è la componente legata agli errori di misura.

Dunque è necessario valutare sia l'incertezza dovuta allo strumento di misura che l'incertezza legata alle differenze tra i pezzi misurati; così Minitab offre vari comandi che aiutano a riconoscere le cause della variabilità:

- Gage R&R Crossed, Gage R&R Nested e Gage Run Chart esaminano la precisione dello strumento di misura ma anche quanto della variazione totale dei dati sia dovuta a differenze tra i pezzi;
- Gage Linearity and Accuracy esamina la linearità e l'accuratezza del misuratore.

Gli errori degli strumenti di misura possono essere classificati come segue:

- l'accuratezza, la quale si riferisce all'attitudine a misurare correttamente, in media, il vero valore (vedi Gage linearity and accuracy study). Essa è suddivisa in:
 - linearità, che è una misura di quanto la dimensione del misurato influisce sull'accuratezza dello strumento di misura. Questa è la differenza tra il valore dell'accuratezza osservata e del range atteso delle misurazioni;



- accuratezza, che è una misura del pregiudizio dello strumento di misura. Questa è la differenza tra la media osservata delle misurazioni e un valore di riferimento;
- stabilità, che è una misura di quanto accuratamente lo strumento opera nel tempo. Questa è la variazione totale ottenuta con uno strumento particolare, sullo stesso pezzo, quando si misura la stessa caratteristica nel tempo;
- la precisione, la quale è una misura della variabilità inerente il procedimento di misura (vedi Gage R&R study). Questa può essere suddivisa in due specifiche componenti dell'errore di misurazione dello strumento, che sono:
 - ripetibilità, che è la variazione dovuta allo strumento di misura. Questa è la variazione osservata quando lo stesso operatore misura lo stesso pezzo ripetutamente con lo stesso strumento;
 - riproducibilità, che è la variabilità dovuta alla diversità di operatori che utilizzano lo stesso strumento di misura. È la variazione osservata quando diversi operatori misurano lo stesso pezzo utilizzando lo stesso strumento.

Possiamo scrivere: $\sigma_{strumento}^2 = \sigma_{ripetibilità}^2 + \sigma_{riproducibilità}^2$.

Infine,

$$\sigma_{totale}^2 = \sigma_{prodotto}^2 + \sigma_{ripetibilità}^2 + \sigma_{riproducibilità}^2$$

graficamente potremmo rappresentare la varianza totale così:

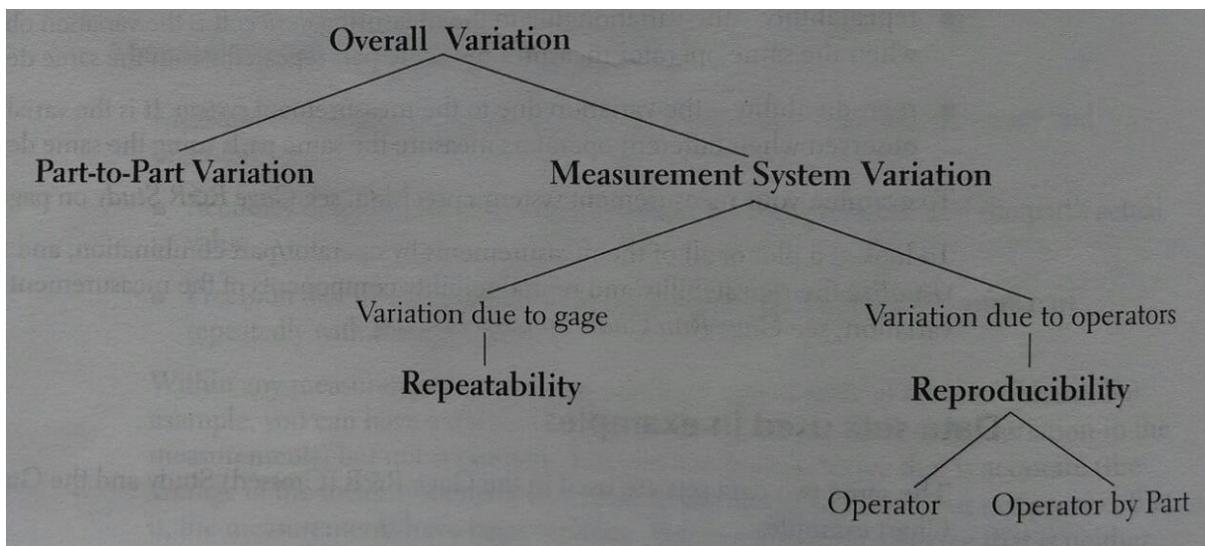


Figura 2.0.1 Rappresentazione della varianza totale

2.1 – Gage R&R study crossed and nested

Lo studio Gage R&R determina quanto della variabilità osservata nel processo è dovuta all'incertezza dello strumento di misura e a differenze tra i pezzi misurati.

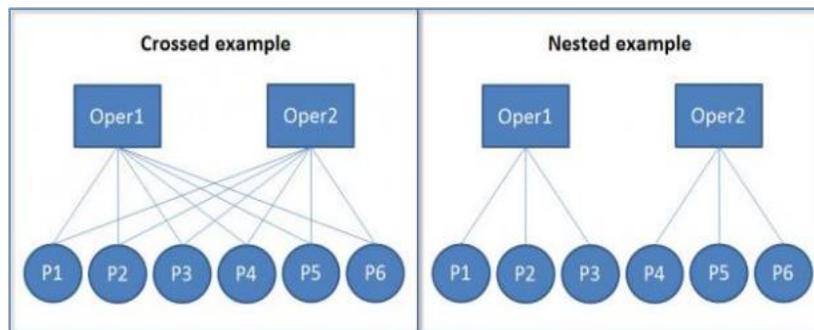


Figura 2.0.2 Esempio Crossed e Nested

Minitab permette di eseguire due test al riguardo:

- **Gage R&R study crossed:** va impiegato quando ogni operatore misura ciascun pezzo, si veda l'esempio crossed della figura 2.0.2;
- **Gage R&R study nested:** va impiegato quando ciascuna parte è misurata una sola volta da un operatore, come nei crash test (destructive testing), dove i pezzi, per essere testati, vengono rotti; si veda l'esempio nested della figura 2.0.2;

Per valutare la riproducibilità e la ripetibilità nei due studi precedenti, possiamo usare due metodi:

- X medio and R, che separa la variazione complessiva in: variazione da pezzo a pezzo; ripetibilità; riproducibilità;
- ANOVA, consente un passo ulteriore perché separa la variazione complessiva in: variazione da pezzo a pezzo; ripetibilità; riproducibilità: Operatore; Operatore by part.

Gage R&R study crossed permette di scegliere tra il metodo X medio and R e il metodo ANOVA, mentre il Gage R&R study nested usa solo il metodo ANOVA. In ogni caso, il metodo ANOVA è più accurato dell'X medio and R metodo.

Se si usa il destructive testing e se è possibile assumere che i pezzi appartenenti ad una singola partita siano abbastanza simili da poter esser considerati tutti perfettamente uguali, allora:

- se tutti gli operatori misurano dei pezzi da ciascuna partita (o lotto), si userà Gage R&R study crossed;



- se ciascuna partita viene misurata solo da un operatore (ma non necessariamente sempre dallo stesso), si userà il Gage R&R study nested. Infatti, ogniqualvolta gli operatori misurano singoli componenti o partite o lotti, si ha il nested.

Sia nel caso di Gage R&R study crossed che nested, i dati vanno inseriti nel foglio di lavoro in modo tale che per ciascuna riga vi sia: il nome o il numero dell'operatore, il numero o il testo indicante il tipo di pezzo, il valore della misura.

Gli studi Gage R&R richiedono che la quantità di dati sia bilanciata. Per esempio: se tre operatori misurano tre volte ciascuno i due pezzi diversi che devono lavorare, allora dovremo collezionare (3*3) dati per pezzo, in totale ((3*3)*2). Se anche per un solo operatore non riuscissimo a raccogliere 3 misure di un pezzo, allora, nella Session Window, Minitab restituirebbe l'errore: “* ERROR * Design is not balanced; execution aborted”.

Per aprire la finestra di dialogo principale del Gage R&R study crossed, bisogna seguire la seguente procedura: Stat → Quality tools → Gage study → Gage R&R study (crossed). Dopo di che:

- in «**Part numbers**» inserire la colonna indicante il tipo di pezzo («Part Num» nell'esempio);
- in «**Measurement data**» inserire la colonna con le misurazioni («Measure» nell'esempio);
- in «**Operators**» inserire la colonna indicante il nome o il numero degli operatori. Completare questo campo è facoltativo; in “**Method of analysis**” scegliere se usare l'ANOVA o Xbar and R come metodo di analisi dei dati.

Per aprire la finestra di dialogo principale del Gage R&R study nested, bisogna seguire la seguente procedura: Stat → Quality tools → Gage study → Gage R&R study (nested). Dopo di che:

- in «**Part numbers**» inserire la colonna indicante il tipo di pezzo («Part Num» nell'esempio);
- in «**Measurement data**» inserire la colonna con le misurazioni («Measure» nell'esempio);
- in «**Operators**» inserire la colonna indicante il nome o il numero degli operatori.

Per entrambi i Gage R&R study, una volta dato l'«OK» alle finestre di dialogo principali, sulla session window, appariranno alcune tabelle, che sono:

- **ANOVA table:** compare solo nel caso si utilizzi il metodo ANOVA e mostra l'output dell'analisi della varianza per i risultati adattati (fitted effects). Nel caso in cui l'ultimo valore della colonna P, ovvero il p-value per «Num Part * Operator», sia <0,25 allora MINITAB adatta (fit) l'intero modello e il metodo ANOVA è più adeguato rispetto all' X and R method nell'analisi di questi dati;
- **Number of distinct categories:** è il numero di categorie diverse nei dati delle misure che il sistema di misura può riconoscere. Quando questo valore è:



- minore di 2, allora i tipi di pezzi non possono essere distinti l'un l'altro e dunque serve uno strumento di misura più preciso;
 - uguale a 2, allora i dati possono essere divisi in due categorie, una coi dati maggiori e l'altra coi dati minori;
 - uguale a 3, allora i dati possono essere suddivisi in tre categorie: bassa, medio e alta;
 - uguale a 4 o maggiore, allora significa che il sistema di misura è adeguato alle misurazioni svolte.
- **Gage R&R:** mostra varie colonne indicanti:
 - VarComp che è il contributo alla varianza totale dato da ciascun fattore;
 - %Contribution (of VarComp) che è la VarComp in percentuale rispetto alla varianza totale, «Total Variation».;
 - StdDev che è la deviazione standard per ciascun componente della varianza totale;
 - StudyVar è la deviazione standard moltiplicata per 5,15 e stima la larghezza necessaria di un intervallo per contenere il 99% di tutte le misure.
 - %StudyVar che è la percentuale della varianza dello studio per ogni componente.

N.B. *Tutte queste osservazioni valgono sia per il Gage R&R study crossed che nested.*

Dopo aver cliccato «OK» sulle finestre di dialogo principale, oltre all'output della session window, comparirà anche una finestra contenente i seguenti grafici (dalla colonna a sinistra a quella a destra e dall'alto verso il basso):

- **components of variation:** è una visualizzazione della tabella Gage R&R nella session window e confronta la «%Contribution» con la «%Study Var» per i seguenti quattro fattori: Gage R&R; Ripetibilità; Riproducibilità; la varianza pezzo-pezzo. Nell'esempio a lato, guardando le colonne blu, possiamo notare che la maggior parte della variazione nei dati è dovuta al sistema di misura;
- **R chart by operator:** mostra la variabilità delle misurazioni per ciascun operatore, così da poter confrontare tra loro i vari operatori e con la media. Ciò aiuta a verificare se tutti gli operatori hanno la media delle loro misurazioni entro i limiti di controllo;
- **X bar chart by operator:** mostra le misurazioni in relazione alla media complessiva per ogni operatore, così si possono confrontare gli operatori tra di loro e alla media.

Nel caso in cui:

- la maggior parte dei punti sia all'interno dei limiti di controllo, allora significa che la variazione nei dati è dovuta principalmente al sistema di misura;
- la maggior parte dei punti è fuori dai limiti di controllo, allora la variazione è dovuta principalmente a differenze tra i tipi di pezzo. Nell'esempio, vi sono forti differenze soprattutto tra il primo e gli altri operatori e la variabilità dei dati è dovuta principalmente al sistema di misura.



- **Measure by Num Part:** permette di notare se vi è differenza tra i tipi di pezzi. Dall'esempio possiamo notare che tra i due tipi di pezzi vi è un poco di differenza;
- **Measure by Operator:** mostra le misurazioni in base all'operatore; si possono anche comparare le misurazioni medie degli operatori;
- **Num Part*Operator Interaction:** permette di visualizzare il p-value per Oper*Part, 0,001 nell'esempio, indicante una significativa relazione tra tipo di pezzo e operatore.

N.B. Il grafico «NumPart*Operator Interaction» compare solo nel Gage R&R study crossed e tutte le osservazioni fatte in queste slide valgono per entrambi i tipi di Gage R&R study.

Esempio 2.1

Vediamo nelle tabelle 2.0.1 e 2.0.2 un esempio di risultato dell'analisi per il Gage R&R study crossed con il metodo ANOVA, col quale si è studiata la differenza nei pezzi misurati tra due operatori.

Gage R&R Study - ANOVA Method					
Two-Way ANOVA Table With Interaction					
Source	DF	SS	MS	F	P
Pezzo	1	8493,4	8493,39	47,9702	0,020
Operatore	2	716,3	358,17	2,0229	0,331
Pezzo * Operatore	2	354,1	177,06	0,7085	0,512
Repeatability	12	2998,7	249,89		
Total	17	12562,5			

α to remove interaction term = 0,05

Two-Way ANOVA Table Without Interaction					
Source	DF	SS	MS	F	P
Pezzo	1	8493,4	8493,39	35,4654	0,000
Operatore	2	716,3	358,17	1,4956	0,258
Repeatability	14	3352,8	239,48		
Total	17	12562,5			

Tabella 2.0.1 Crossed con metodo ANOVA

Gage R&R			
Source	VarComp	%Contribution (of VarComp)	
Total Gage R&R	259,26	22,04	
Repeatability	239,48	20,36	
Reproducibility	19,78	1,68	
Operatore	19,78	1,68	
Part-To-Part	917,10	77,96	
Total Variation	1176,37	100,00	

Source	StdDev (SD)	Study Var (6 * SD)	%Study Var (%SV)
Total Gage R&R	16,1017	96,610	46,95
Repeatability	15,4753	92,852	45,12
Reproducibility	4,4475	26,685	12,97
Operatore	4,4475	26,685	12,97
Part-To-Part	30,2837	181,702	88,30
Total Variation	34,2982	205,789	100,00

Number of Distinct Categories = 2

Tabella 2.0.2 Crossed con metodo ANOVA

L'ultimo valore della colonna «P» nella tabella «Two-Way ANOVA table with interaction» (0,512) è >0,25, dunque MINITAB adatta (fit) il modello senza le interazioni e usa il modello ridotto per definire le statistiche Gage R&R, infatti genera l'ulteriore tabella «Two-Way ANOVA table without interaction», dove la riga «Operator*Part» viene omessa e la omette pure nella tabella «Gage R&R». Nel caso in cui il valore fosse stato <0,25, allora il metodo ANOVA sarebbe stata la scelta migliore perché più accurato.

Oltre a ciò, nella tabella «Gage R&R», alla colonna «%Contribution (of VarComp)», possiamo notare che il 77,96% della variazione complessiva è dovuta alla differenza da pezzo a pezzo (riga «Part-to-Part»). Dunque il sistema di misura è sotto controllo.

All'ultima riga, il «Number of distinct categories» è pari a 2. Ciò indica che il mio sistema di misura non è adatto alle misurazioni svolte e andrebbe impiegato uno strumento più preciso.

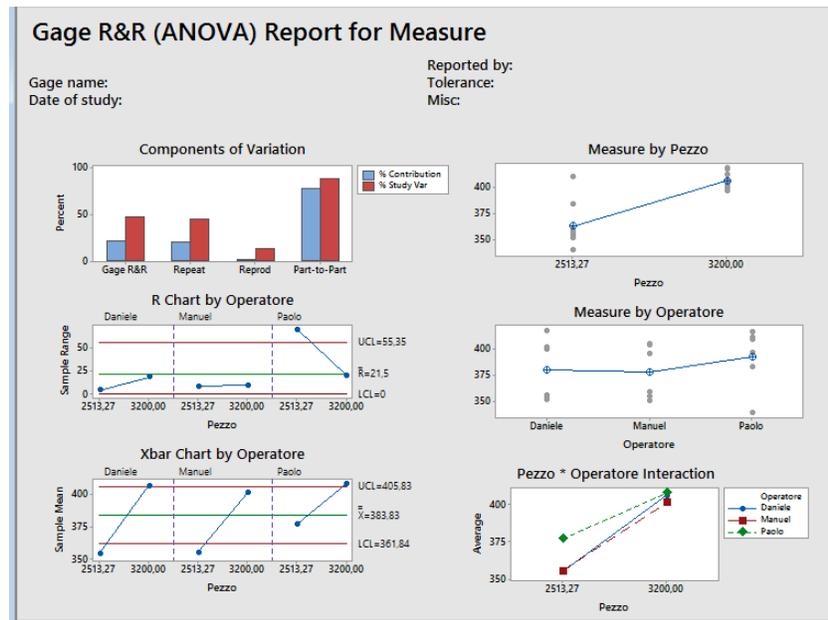


Figura 2.0.3 Grafico crossed con metodo ANOVA

Alla figura 2.0.3 è riportata anche la rappresentazione grafica delle tabelle 2.0.1 e 2.0.2. Dal grafico «Measure by Pezzo», notiamo che vi è una significativa differenza tra le misure dei pezzi, dato che la linea non è orizzontale. Invece, guardando «Measure by Operatore», notiamo che non vi sono grandi differenze tra gli operatori, soprattutto tra Daniele e Manuel.

Nel grafico «Xbar Chart by Operatore», molti dei punti cadono fuori dai limiti di controllo e ciò indica che gran parte della variazione complessiva delle misure è dovuta a differenze tra i pezzi.

2.2 – Gage run chart

Il Gage run chart mostra tutte le osservazioni in funzione dell'operatore e del tipo di pezzo. Si può utilizzare questo grafico per valutare velocemente eventuali differenze nelle misurazioni tra diversi operatori e pezzi.

Per produrre il grafico tutti i pezzi devono essere misurati da ciascun operatore, come per il crossed.

Un processo sotto controllo dovrebbe dare un grafico in cui le osservazioni sono sparpagliate in modo casuale attorno alla media; in caso contrario i punti seguirebbero un certo schema e dunque il processo non sarebbe sotto controllo.



Prima di aprire la finestra di dialogo per tracciare questo grafico, bisogna inserire i dati come per i Gage R&R study, ovvero: ogni riga deve contenere un numero o un testo indicante il tipo di pezzo misurato; il nome dell'operatore che ha fatto la misura; il valore della misurazione.

Inseriti i dati, per aprire la finestra di dialogo principale del Gage run chart, bisogna seguire la seguente procedura: Stat → Quality tools → Gage study → Gage run chart. A questo punto, bisognerà:

- in «Part numbers», inserire la colonna indicante i tipi di pezzi;
- in «Operators», inserire la colonna indicante gli operatori;
- in «Measurement data», inserire la colonna con le misurazioni.

Inoltre, si possono inserire dei numeri di prova e un valore storico della media («Historical mean»), sennò Minitab calcolerà il valore medio dai dati.

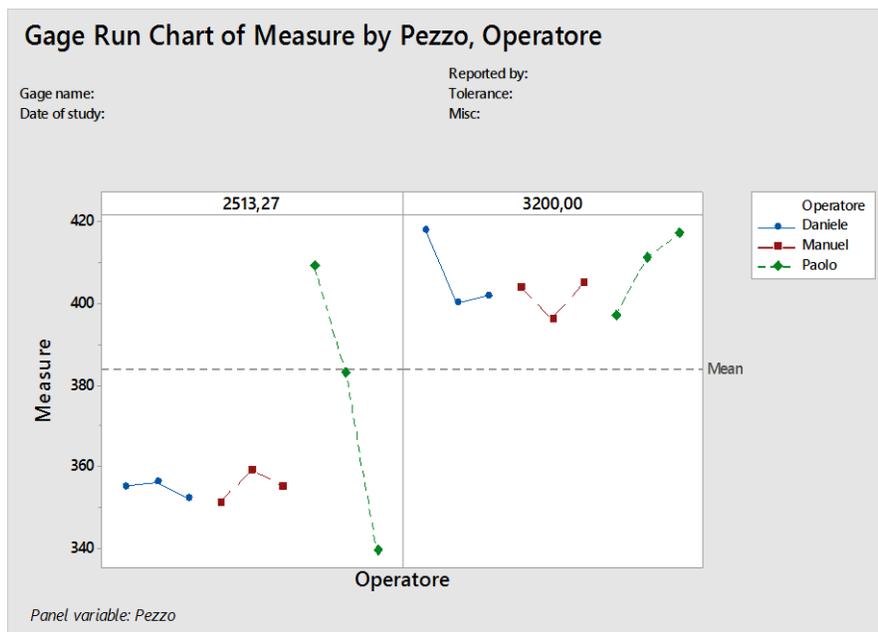


Figura 2.0.4 Esempio di Gage run chart

2.3 – Gage linearity and accuracy study

Il gage linearity study dice quanto accurate siano le misurazioni rispetto al range aspettato delle misurazioni. Questo risponde alla domanda: “Il mio strumento ha la stessa accuratezza per tutti le dimensioni degli oggetti misurati?”.

Il gage accuracy study ed esamina la differenza tra la media delle osservazioni fatte e un valore della media di riferimento, come un valore storico. Questo risponde alla domanda: “Quanto accurato è il mio strumento rispetto al valore di riferimento o di progetto?”.



Per aprire la finestra di dialogo di questi due studi, bisogna seguire la seguente procedura: Stat → Quality tools → Gage study → Gage linearity and bias Study. In questa finestra bisognerà:

- in «Part numbers» inserire la colonna contenente il tipo di pezzo;
- in «Reference value» inserire la colonna dei valori di riferimento;
- in «Measurement data» inserire i valori delle misurazioni.

Applicando un gage linearity and bias study, si ottiene un grafico simile a quello in figura 2.0.5, dal quale si possono conoscere:

- la variazione nelle misurazioni dovuta alla linearità dello strumento, che è la pendenza (slope) della retta. Più la pendenza della retta è bassa, meno le dimensioni dell'oggetto misurato influiscono sull'accuratezza dello strumento. Nella figura 2.0.5 la pendenza è pari a 0,074;
- la variazione nelle misurazioni dovuta all'accuratezza dello strumento (bias). Nell'esempio è pari, in media, a 1,667.

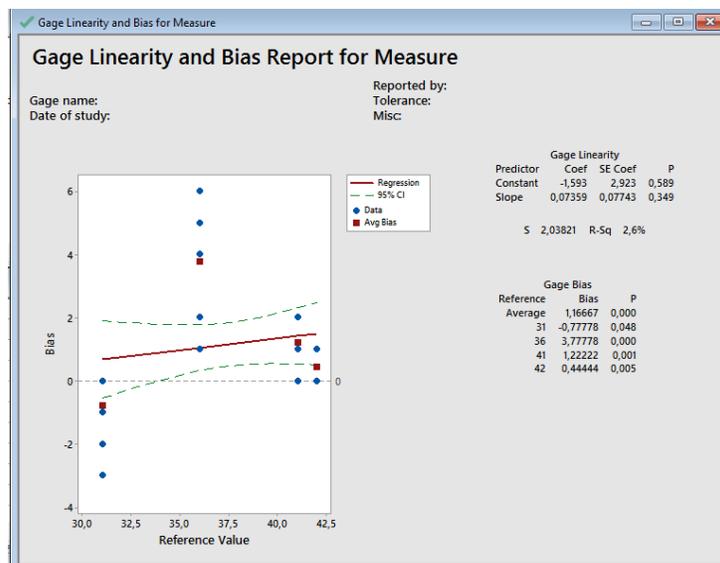


Figura 2.0.5 Esempio di Gage linearity and accuracy study



Capitolo 3

Variables control charts (Le carte di controllo)

3.1 – Che cosa sono e a cosa servono?

Le carte di controllo sono grafici che aiutano a tenere sotto controllo dei parametri statistici (come la media e la variabilità) dei pezzi prodotti (o del processo produttivo) al trascorrere del tempo. Le carte riportano:

- nell'asse delle ascisse il tempo o il numero del campione raccolto;
- nell'asse delle ordinate il valore di un parametro statistico per ciascun campione (come media campionaria, range dei dati, ...).
- la linea centrale (CL), che rappresenta il valor medio della qualità, che dovrebbe corrispondere al valore desiderato se il processo è sotto controllo;
- il limite di controllo superiore (UCL) e il limite di controllo inferiore (LCL). Di base, questi limiti sono limiti 3σ , ovvero distanti dalla media 3 volte la deviazione standard (σ).

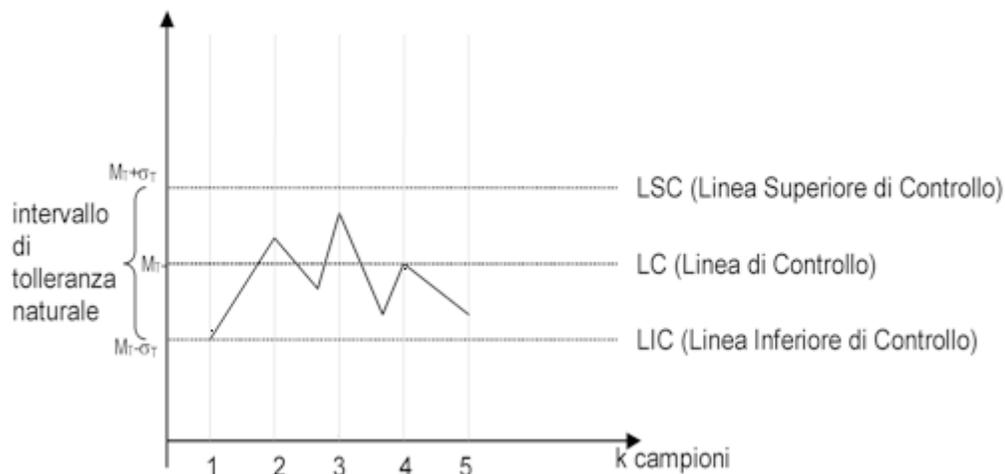


Figura 3.0.1 Carta di controllo

Lo scopo delle carte di controllo è studiare uno o più gruppi di dati normalmente distribuiti, che rappresentano caratteristiche quantitative misurate (quali la lunghezza, la temperatura, il tempo ...) di un pezzo prodotto o di un processo produttivo. Ne consegue che le carte di controllo sono utili per:

- controllare o stimare la media e la variabilità delle caratteristiche dei pezzi prodotti o del processo produttivo, così da monitorare quest'ultimo;
- trovare dati anomali e quindi errori nel processo produttivo (come per il Run chart).



Le carte di controllo possono essere di vario tipo, come mostrato nella tabella sottostante.

	To plot this...	Use this chart
For data in subgroups:	subgroup means, \bar{X}	\bar{X}
	subgroup ranges, r	R
	subgroup standard deviations, s	S
	\bar{X} and r on same screen	\bar{X} and R
	\bar{X} and s on same screen	\bar{X} and S
For individual observations:	individual observations	Individuals
	moving ranges	Moving Range
	individual observations and moving ranges on same screen	I-MR
For subgroup combinations:	exponentially weighted moving averages	EWMA
	moving averages	Moving Average
	cumulative sums	CUSUM
	individual observations or subgroup means according to their distance from the center line	Zone
For short runs:	standardized individual observations and moving ranges from short run processes	Z-MR

Tabella 3.0.1 Tipologie di carte di controllo

3.2 – Come inserire i dati nel foglio di lavoro

Nel foglio di lavoro, i dati per le carte di controllo possono essere raggruppati:

- in righe. Questo modo di raggruppare va usato nel caso in cui tutti i campioni abbiano ugual numerosità campionaria;
- tutti in una unica colonna. Questo modo di raggruppare va usato nel caso in cui si abbiano sottogruppi (o campioni) di numerosità campionaria differente tra loro; può essere, però, applicata anche nel caso di numerosità campionaria (anche 1) costante tra i campioni.

I dati da inserire nelle carte di controllo dovrebbero provenire da una distribuzione normale. Nel caso in cui i dati provengano da una distribuzione fortemente asimmetrica, allora bisognerà attuare una «Box-Cox transformation» per normalizzarla. Questo comando si trova pigiando il pulsante “Options” di una qualsiasi carta di controllo generica, vedi figura 3.0.2.

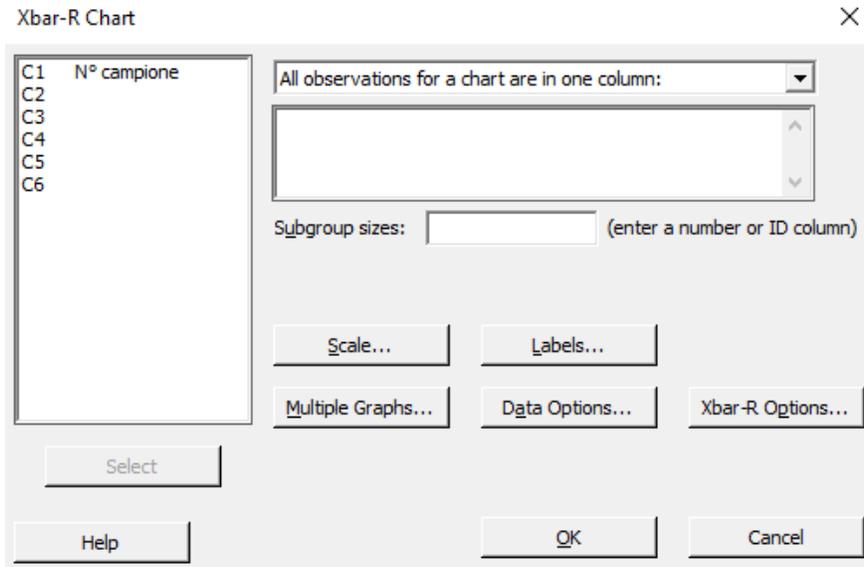


Figura 3.0.2 Finestra di dialogo principale di una carta di controllo generica

Nella finestra di dialogo di una carta di controllo generica (vedi figura 3.0.2), per l’inserimento dei dati si dovrà:

1. nel caso in cui tutti i dati sia raggruppati in una colonna, nel menu a tendina indicare “All observations for a chart are in one column:”; nel caso in cui tutti i dati siano raggruppati in più righe, nel menu a tendina indicare «Observations for a subgroup are in one row of columns:»;
2. nella finestra vuota sotto il menu a tendina, indicare la/e colonna/e in cui sono raggruppati i dati. Per far ciò si posiziona il cursore, con un clic sinistro, nella finestra vuota per poi selezionare con un doppio clic sinistro le colonne desiderate dall’elenco a sinistra;
3. solo nel caso di dati raggruppati tutti in una colonna, in «Subgroup size»:
 - nel caso in cui i dati siano raggruppati in sottogruppi di numerosità campionaria variabile, indicare la colonna che indica il sottogruppo di appartenenza per ogni dato;
 - oppure, nel caso in cui tutti i campioni avessero pari numerosità campionaria, indicare la dimensione di ogni campo.

In certi casi, questo campo non è presente, come per esempio nella carta di controllo Individuals in cui si inseriscono solo dati individuali.

3.3 – Special causes (cause speciali)

Come detto, le carte di controllo servono anche a rintracciare dati anomali nel processo. Questi dati anomali sono provocati dalle cosiddette “cause speciali”; quando queste si verificano, si dice che il processo è fuori controllo e va riportato ad un corretto funzionamento.



Riconoscere la presenza delle cause speciali è molto importante, e ciò si può fare:

- per via visiva, perché: uno o più punti del grafico cadono fuori dai limiti di controllo; più punti si dispongono in modo non casuale;
- tramite Minitab utilizzando appositi test. Per attuarli bisognerà aprire la finestra di dialogo principale della carta utilizzata, cliccare su “Options”, così si aprirà una seconda finestra in cui cliccare “Test”. Fatto ciò apparirà un elenco, vedi figura 3.0.3, in cui si potranno spuntare i test voluti.

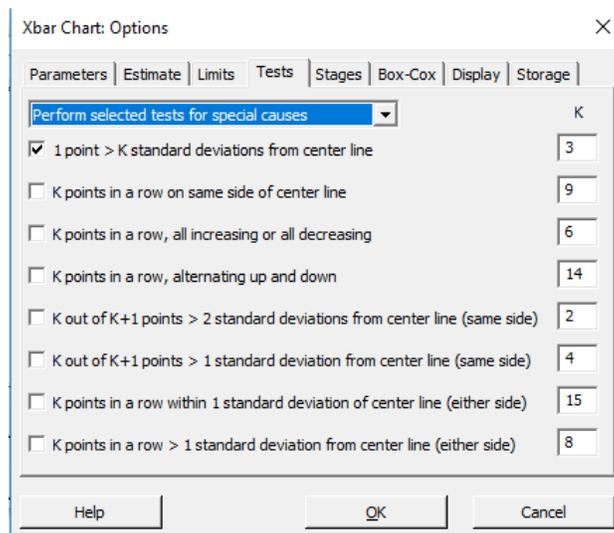


Figura 3.0.3 Finestra "Test" di una generica carta di controllo

3.4 - Variables control charts for data in subgroups (per dati raggruppati in sottogruppi)

Nel caso in cui i dati siano raggruppati in sottogruppi, si useranno le seguenti carte di controllo:

- R chart;
- S chart;
- X medio chart (X bar...);
- X medio and R chart;
- X medio and S chart;
- I-MR-R/S (Between/within) chart.

Tutte queste carte, tranne l’X medio chart, richiedono che la numerosità campionaria dei sottogruppi sia almeno pari a 2. Inoltre, i sottogruppi possono essere di numerosità campionaria diversa.



3.4.1 – R chart

Questa è usata per valutare la variabilità nei sottogruppi, ovvero la variabilità del processo in un dato istante, e per rintracciare la presenza di cause speciali.

Minitab stima la variazione del processo (σ) in base alla media dei range dei sottogruppi.

Essa si applica a campioni di numerosità pari o inferiore a 5 e prima della costruzione dell'X medio chart.

Per aprire la finestra di dialogo principale dell'R chart, bisognerà seguire la seguente procedura: Stat → Control Chart → Variables Charts for Subgroups → R. A questo punto, bisognerà indicare come sono raggruppati i dati e indicarne le colonne interessate e si potrà:

- modificare il layout della carta di controllo tramite i pulsanti "Scale..." e "Labels...";
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- cliccare su «R options...».

Esempio 3.4.1

↓	C1-T	C2	C3	C4
	Operatore	Ordine	p (mm)	t (s)
1	m	161111	4100,00	359
2	Missing	*	4100,00	334
3	Missing	*	4100,00	334
4	Missing	*	4100,00	334
5	m	170030	3470,57	420
6	m	161086	2513,27	351
7	Missing	*	2513,27	359
8	Missing	*	2513,27	355
9	Missing	*	2513,27	356
10	Missing	*	2513,27	352
11	Missing	*	2513,27	354
12	Missing	*	2513,27	352
13	m	161113	3141,59	420
14	Missing	*	3141,59	419
15	Missing	*	3141,59	419
16	Missing	*	3141,59	419
17	Missing	*	3141,59	419
18	m	161086	2513,27	374
19	Missing	*	2513,27	379
20	Missing	*	2513,27	382
21	Missing	*	2513,27	380
22	m	170071	2984,51	411
23	Missing	*	2984,51	405
24	Missing	*	2984,51	405
25	Missing	*	2984,51	385

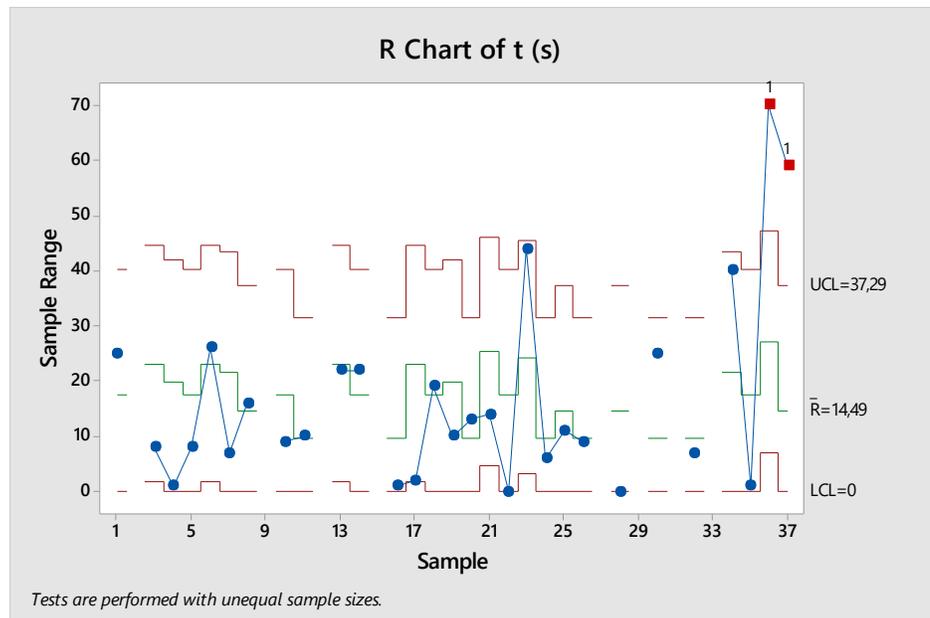


Figura 3.0.4 R chart

Supponendo di essere gli addetti alla gestione del personale e di voler valutare la variabilità dei tempi di lavorazione dell'operatore Manuel (m), si raccoglieranno vari tempi di lavorazione per diversi tipi di



prodotti per poi riportarli nel foglio di lavoro come mostrato sopra, e si aprirà la finestra di dialogo principale dell'R chart, quindi la si compilerà.

Dando l'«OK» alla finestra di dialogo principale dell'R chart, comparirà il grafico di figura 3.0.4.

Il limite superiore e quello inferiore e la linea del range medio non sono continue perché i vari campioni hanno numerosità campionaria differente.

Dal grafico notiamo che solo due punti escono dai limiti di controllo (sintomo di causa speciale), infatti, nella Session Window, Minitab mostra le seguenti scritte:

- “TEST 1. One point more than 3,00 standard deviations from center line.”;
- “Test Failed at points: 36; 37”.

Dato che i due punti fuori controllo sono proprio gli ultimi due, bisognerà verificare che il processo non stia andando fuori controllo.

3.4.2 – S chart

S chart è una carta di controllo per le deviazioni standard dei sottogruppi. Si può utilizzare questa carta per monitorare la variazione di processo e individuare la presenza di cause speciali.

Le S chart sono tipicamente usate con campioni di numerosità campionaria variabile e maggiore di 5, mentre gli R chart vengono usati per campioni di numerosità campionaria pari o inferiore a 5.

Per aprire la finestra di dialogo dell'S chart, bisogna seguire la seguente procedura: Stat → Control Chart → Variables Charts for Subgroups → S. A questo punto, bisognerà indicare come sono raggruppati i dati e indicarne le colonne interessate, come mostrato alla sezione 3.2. Inoltre, si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri, simili a quelli dell'R chart, pigiando il pulsante «S Options...».

3.4.3 – X medio chart

X medio chart è una carta di controllo per le medie dei sottogruppi o per osservazioni individuali. Lo si può utilizzare per monitorare il livello del processo e rintracciare la presenza di cause speciali.

Questa carta di controllo calcola la variazione del processo (σ) usando la pooled standard deviation o sulla media delle deviazioni standard di tutti i sottogruppi o su un valore storico. Nel caso in cui l'X medio chart rappresenti osservazioni individuali, allora Minitab stima σ calcolando la media dei range mobili (moving range) per poi dividerla per una costante.



X medio chart confronta le medie aritmetiche dei singoli sottogruppi e permette di monitorare o individuare:

- la media aritmetica tra i sottogruppi (il livello medio di attività del processo monitorato);
- le variazioni indesiderate del valor medio del processo (cause speciali) e quindi i problemi nella produzione;
- la misura della variabilità tra i campioni, ovvero la variabilità del processo nel tempo.

Per aprire la finestra di dialogo principale dell’X medio chart, bisogna seguire la seguente procedura: Stat → Control Chart → Variables Charts for Subgroups → Xbar. Così ci apparirà una finestra in cui dovremo indicare come e dove sono stati inseriti. Inoltre, si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri, pigiando il pulsante «Xbar Options...».

Esempio 3.4.2

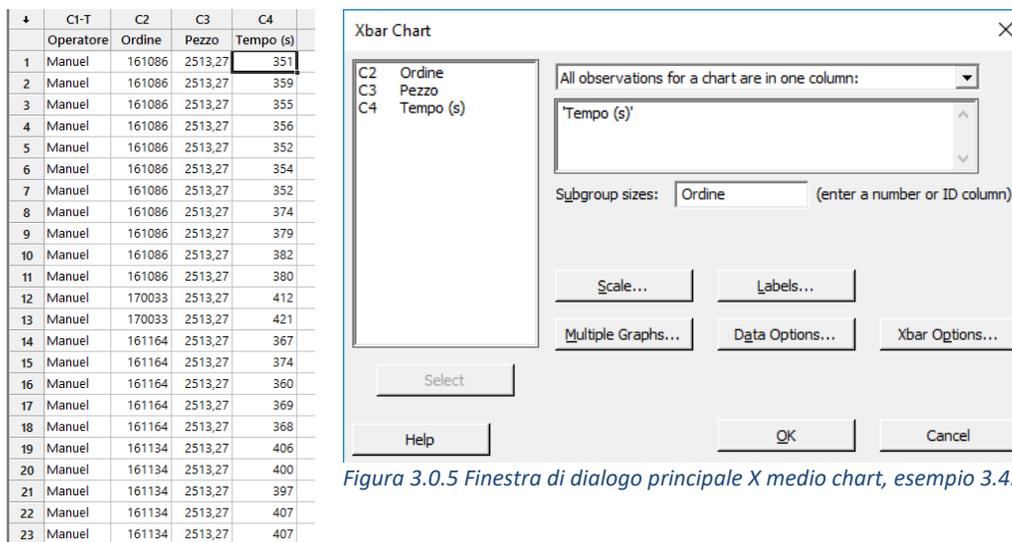


Figura 3.0.5 Finestra di dialogo principale X medio chart, esempio 3.4.2

Si supponga di essere addetti alla gestione del personale e di voler monitorare le prestazioni di un operatore (Manuel) nel produrre il pezzo 2513,27.

Per far ciò, a campione, sono stati raccolti i tempi di lavorazione (t(s)) e inseriti nel foglio di lavoro di MINITAB, come mostrato in alto. Si noti che è presente pure la colonna «Ordine», la quale indica il sottogruppo di appartenenza di ogni misura. Poi si è aperta la finestra di dialogo principale della carta di controllo X bar e completata come mostrato in figura 3.0.5.

Cliccato «OK» sulla finestra di dialogo principale, comparirà il grafico sottostante, in cui «Sample» sta per «Ordine», mentre «Sample Mean» sta per media di «Misura» per ciascun campione.

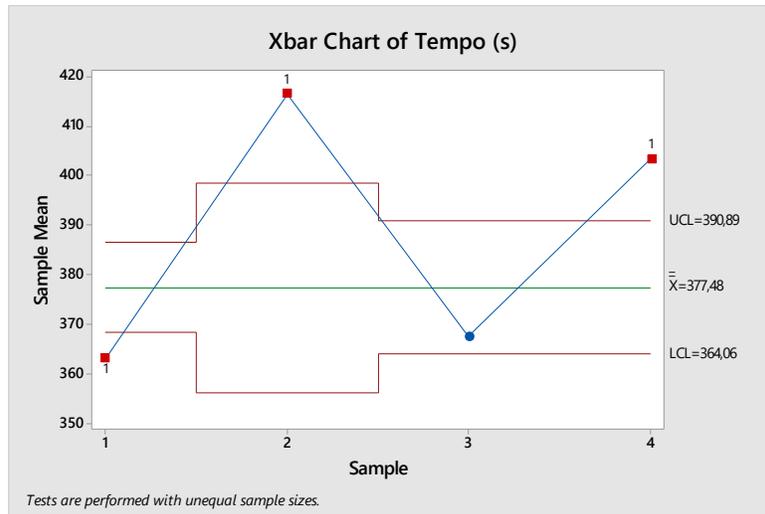


Figura 3.0.6 X medio chart dell'esempio 3.4.2

Dal grafico possiamo notare che, per questo tipo di pezzo, i tempi di lavorazione di Manuel sono molto variabili. Ciò non è un bene se si vuole standardizzare il processo produttivo. Ne segue che bisognerà convincere l'operatore, formandolo, ad essere più costante.

Nella Session Window, MINITAB mostra le scritte:

- TEST 1. One point more than 3,00 standard deviations from center line.
- Test Failed at points: 1; 2; 4

Ciò significa che vi è la presenza di cause speciali.

3.4.4 - X medio and R chart e X medio and S chart

In questo tipo di carte i grafici X medio chart e R chart (o S chart) vengono raccolti in una unica finestra. Visualizzando i grafici contemporaneamente, è possibile monitorare sia il livello del processo sia la variazione del processo, così da poter anche rintracciare la presenza di cause speciali.

Tutte le caratteristiche e le osservazioni finali fatte precedentemente sui singoli grafici rimangono invariate.

L'X medio accoppiato con l'R chart è usato per campioni con numerosità campionaria inferiore o pari a 5, mentre l'X medio accoppiato con l'S chart per i campioni con numerosità campionaria maggiore di 5.

Per aprire la finestra di dialogo principale dell'X medio and R chart o dell'X medio and S chart bisogna seguire la seguente procedura: Stat → Control Chart → Variables charts for subgroups → X and R / X and S...



Essendo questi due grafici un mix dei tre visti precedentemente, tutte le varie opzioni e considerazioni sui grafici si ripetono.

3.4.5 - I-MR-R/S (Between/within) chart

Questa carta di controllo è impiegata nei casi in cui sia rilevante sia la variazione tra i sottogruppi che tra i dati di ciascun sottogruppo, ovvero all'interno dei sottogruppi.

L'I-MR-R/S chart crea tre distinte valutazioni della variazione nel processo:

- **individuals (I) chart:** grafica la media di ogni sottogruppo in una carta di controllo individuale (individuals control chart) invece che in un X bar chart. I limiti di controllo sono calcolati tramite il range mobile tra medie consecutive, ciò permette di non considerare la variazione all'interno dei sottogruppi nella stima dei limiti di controllo;
- **moving range (MR) chart:** grafica la media dei sottogruppi col range mobile per eliminare, dalla variazione, la componente interna ai sottogruppi;
- **R/S chart:** quest'ultimo grafico monitora la variabilità dei dati all'interno dei sottogruppi e può essere a scelta un R chart o un S chart;

Per aprire la finestra di dialogo di un I-MR-R/S chart bisognerà seguire la seguente procedura: Stat → Control Chart → Variables Charts for Subgroups → I-MR-R/S (Between/Within)...

Così facendo apparirà una finestra in cui si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri, pigiando il pulsante «I-MR-R/S Options...».

3.5 Variables control charts for individual observations (per osservazioni individuali)

Queste sono impiegate nel caso in cui la dimensione campionaria di tutti i campioni sia uno. Si potrebbero raccogliere dati individuali nel caso in cui:

- i dati non possano essere raggruppati;
- sia controllata ogni singola unità prodotta;
- il tasso di produzione sia molto basso;
- la deviazione standard sia piccola rispetto al globale.

I grafici adatti a questo tipo di carte sono i seguenti:

- individual (I) chart;



- moving Range (MR) chart, che va tracciato prima dell'I chart;
- I-MR chart.

In questo tipo di grafici i dati devono essere sempre raggruppati sotto una unica colonna, perché sono dati individuali.

3.5.1 - Moving Range (MR) chart

Questa carta di controllo grafica la variabilità del processo tramite l'utilizzo di range mobili calcolati a partire da sottogruppi artificiali, non realmente presenti, dato che stiamo trattando dati individuali.

Tramite la seguente procedura si potrà aprire la finestra di dialogo del Moving Range chart: Stat → Control Chart → Variables Charts for Individuals → Moving Ranges. Inoltre, si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri, pigiando il pulsante «MR Options...».

Esempio 3.5.1

Supponiamo di essere gli addetti al controllo produzione e di voler monitorare come varia la differenza tra il tempo macchina e il tempo complessivo necessario all'operatore per: caricare e scaricare il pezzo dalla macchina; attrezzare la macchina per la lavorazione.

A tale scopo, raccogliamo i dati dal campo per poi inserirli nel foglio di lavoro di Minitab, ottenendo la tabella 3.0.2, in cui: $t(s)$ è il tempo macchina; $T(s)$ è il tempo totale impiegato dall'operatore a lavorare il pezzo; $T-t$ pz è la differenza dei due tempi precedenti divisa per il numero di pezzi lavorati.



↓	C1-T	C2	C3	C4	C5	C6	C7	C8	C9
	Operatore	Ordine	p (mm)	t (s)	Media t(s)	T(s) consuntivo	n° pezzi	T-t (min)	T-t al pz
1	m	161111	4100,00	359	340,25	4200	8	25	3
2	p	161285	2133,14	273	273,00	3000	5	27	5
3	m	170030	3470,57	420	420,00	13500	29	22	1
4	m	161086	2513,27	351	354,14	7800	19	18	1
5	m	161113	3141,59	420	419,20	18300	33	74	2
6	m	161086	2513,27	374	378,75	9900	24	14	1
7	d	161161	1256,64	192	207,57	7800	32	19	1
8	p	161271	1875,53	282	280,36	13800	20	137	7
9	d	170042	2513,27	355	354,33	11700	36	-18	-0
10	d	170078	1627,34	247	261,67	4500	12	23	2
11	m	170071	2984,51	411	402,71	6000	11	26	2
12	m	170078	3173,01	416	421,50	6900	13	24	2
13	m	161084	3840,00	477	477,00	3900	6	17	3
14	m	161066	3200,00	404	401,25	3900	8	12	1
15	m	56465	3000,00	459	464,00	1800	2	15	7
16	m	56478	2100,00	351	351,00	1500	2	13	7
17	m	161244	4800,00	545	555,43	11400	14	60	4
18	d	170015	2800,00	398	404,33	9900	22	17	1
19	m	161066	2513,27	379	379,29	3900	8	14	2
20	p	161118	1727,87	323	322,50	2700	4	24	6
21	m	161144	3141,59	409	423,25	4200	6	28	5
22	m	161134	2513,27	406	403,40	6900	13	28	2
23	m	161164	3769,91	484	477,50	3900	6	17	3

Tabella 3.0.2 Dati dell'esempio 3.4.3

A questo punto, basterà aprire la finestra di dialogo principale dell'I chart e compilarla. Cliccando «OK» alla slide di dialogo principale, comparirà la figura 3.0.7.

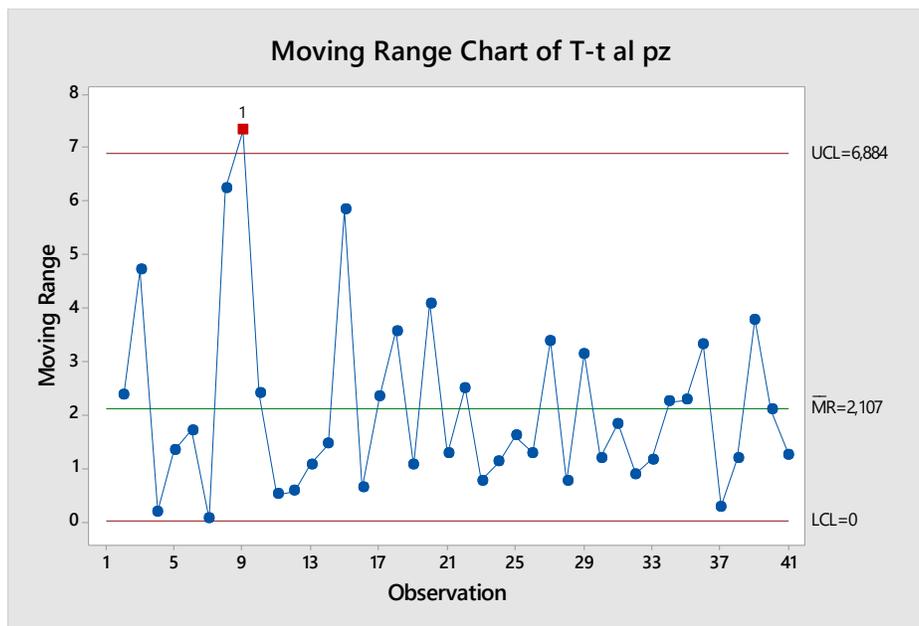


Figura 3.0.7 Esempio di MR chart

Come si nota, una osservazione è uscita dai limiti di controllo e il tempo impiegato dall'operatore per attrezzare la macchina per la lavorazione, caricare e scaricare il pezzo era molto variabile fino al 21°



campione ma poi è andato via via stabilizzandosi. Sarebbe opportuno ridurre questa variabilità il più possibile, per rendere il processo produttivo più standardizzato e organizzare meglio la produzione.

Inoltre, Minitab ha mostrato nella Session Window:

- TEST 1. One point more than 3,00 standard deviations from center line.
- Test Failed at points: 9

evidenziando la presenza di cause speciali.

3.5.2 - Individual (I) chart;

Questo grafico le osservazioni individuali e si utilizza per:

- tracciare la media del processo;
- individuare cause speciali nel processo.

Questo grafico è uguale al grafico X medio chart con l'unica differenza che, invece di rappresentare i valori medi dei sottogruppi, rappresenta direttamente il valore della singola misurazione.

Per aprire la finestra di dialogo dell'I chart, si dovrà seguire la seguente procedura: Stat → Control Chart → Variables Charts for Individuals → Individuals. Inoltre, si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri, pigiando il pulsante «I Chart Options...».

3.5.3 - I-MR chart

Questo raggruppa un Individuals chart e un Moving Range chart in una unica finestra.

Per aprire la finestra di dialogo con questa carta di controllo bisogna seguire la seguente procedura: Stat → Control Chart → Variables Charts for Individuals → I-MR. Inoltre, si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri pigiando il pulsante «I-MR Chart Options...».



3.6 - Variables control charts using subgroups combinations (usanti combinazioni di sottogruppi)

Le carte di controllo che lavorano combinazioni di sottogruppi sono:

- **EWMA chart:** grafico di medie mobili esponenzialmente pesate;
- **Moving Average chart:** grafico di medie mobili non pesate;
- **CUSUM chart:** grafico di somme cumulate delle deviazioni da un valore nominale;
- **Zone chart:** grafico che assegna un peso a ciascun punto in base alla sua distanza dalla linea centrale e grafica il punteggio cumulativo.

Queste carte di controllo possono essere applicate su campioni di numerosità campionaria variabile e maggiore o pari a 1. Queste sono tipicamente usate per valutare il livello del processo.

Per facilitare la costruzione delle carte di controllo, consideriamo l'esempio delle tabelle 3.0.3 e 3.0.4 con 30 misurazioni (che sono disposte tutte lungo la stessa colonna C2) per applicarlo alle carte di controllo per combinazioni di sottogruppi.

↓	C1	C2
	campione	misurazione
1	1	9,45
2	2	7,99
3	3	9,29
4	4	11,66
5	5	12,16
6	6	10,18
7	7	8,04
8	8	11,46
9	9	9,20
10	10	10,34
11	11	9,03
12	12	11,47
13	13	10,51
14	14	9,40
15	15	10,08
16	16	9,37
17	17	10,62
18	18	10,31
19	19	8,52
20	20	10,84
21	21	10,90
22	22	9,33
23	23	12,29

Tabella 3.0.3 Dati per esempi

↓	C1	C2
	campione	misurazione
10	10	10,34
11	11	9,03
12	12	11,47
13	13	10,51
14	14	9,40
15	15	10,08
16	16	9,37
17	17	10,62
18	18	10,31
19	19	8,52
20	20	10,84
21	21	10,90
22	22	9,33
23	23	12,29
24	24	11,50
25	25	10,60
26	26	11,08
27	27	10,38
28	28	11,62
29	29	11,31
30	30	10,52

Tabella 3.0.4 Dati per esempi



3.6.1 – EWMA chart

Le EWMA sono carte di controllo a medie mobili pesate esponenzialmente. Ogni punto delle EWMA conserva informazioni dalle osservazioni o dai sottogruppi precedenti. Ciò è utile per rintracciare qualsiasi tipo di spostamento del processo da un parametro di riferimento o obiettivo.

Le EWMA sono insensibili all'assunzione di normalità dei dati e quindi sono ideali nel controllo delle osservazioni individuali, ma possono essere applicate anche a sottogruppi perché da essi ne calcolerà la media.

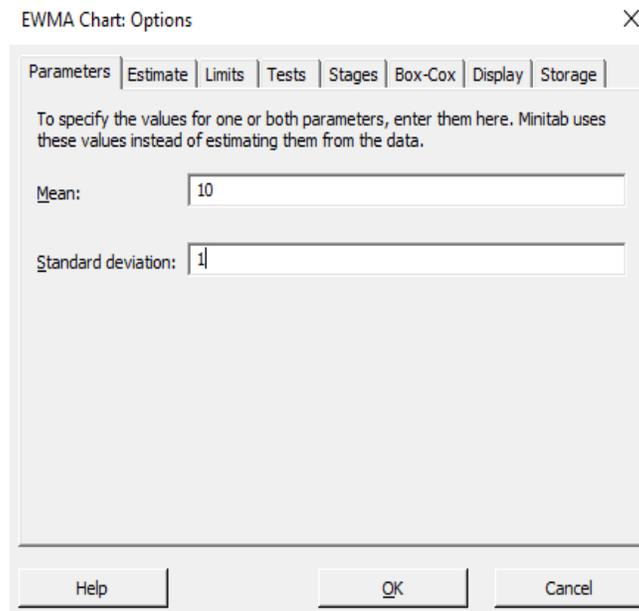


Figura 3.0.8 Finestra Options EWMA

Per aprire la loro finestra di dialogo principale bisogna seguire la seguente procedura: Stat → Control Chart → Time-weighted charts → EWMA...

Così apparirà la finestra di dialogo principale, vedi figura 3.0.8, in cui si potrà:

- specificare il valore del peso (weight of EWMA), compreso tra 0 e 1, che il programma dovrà usare per pesare i dati. Di base è preimpostato a 0,2;
- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;

Cliccando sul pulsante «EWMA options...» comparirà una finestra della figura 3.0.8. I valori 10 e 1, rispettivamente per media e deviazione standard, sono stati inseriti arbitrariamente ai fini dell'esempio successivo; di default, il programma non specifica nessun valore. Inoltre, in questa finestra si potranno personalizzare varie opzioni.



Esempio 3.6.1

Vogliamo verificare se, con un valore di riferimento per la media di 10 e per la deviazione standard di 1, le misurazioni dei campioni raccolti rientrano tra i limiti di controllo e dunque se il sistema è sotto controllo.

Per questo disponiamo i dati nel foglio di lavoro come mostrato alle tabelle 3.0.3 e 3.0.4 e apriamo la finestra di dialogo principale di queste carte compilandola. Infine, cliccando «OK», comparirà il grafico mostrato nella figura 3.0.9.

Dal grafico possiamo notare che tutti i punti ricadono tra l'ULC e l'LCL, mentre le misurazioni raccolte dal 22° campione in poi cominciano a mostrare la presenza di una causa speciale che può essere o una deriva o lo spostamento del valore di riferimento. Possiamo concludere che il processo è sotto controllo ma che è bene indagare l'origine della causa speciale, dato che i punti blu sono molto vicini al limite di controllo superiore.

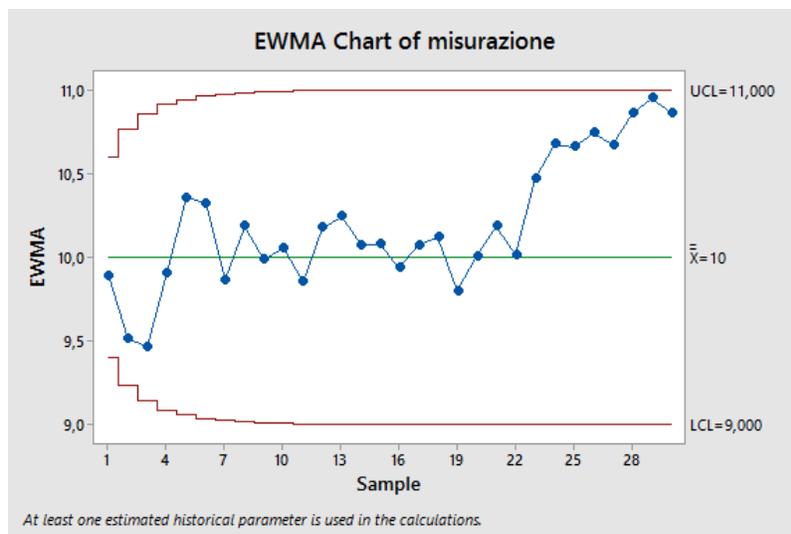


Figura 3.0.9 EWMA chart

3.6.2 - Moving Average chart

È un grafico di medie mobili calcolate a partire da sottogruppi artificiali creati da osservazioni consecutive. In ogni caso, le osservazioni possono essere sia individuali che medie di un sottogruppo. Solitamente si preferisce usare l'EWMA chart al moving average chart, perché quest'ultimo non pesa i dati.



La procedura per aprire la finestra di dialogo del moving average chart è la seguente: Stat → Control Chart → Time-weighted charts → Moving average. Così apparirà una finestra di dialogo, in cui si potrà:

- specificare il valore della lunghezza delle medie mobili (length of MA), di base è 3;
- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare tutta una serie di parametri pigiando il pulsante «MA Options...».

3.6.3 – CUSUM chart

Un cumulative sum chart (CUSUM) rappresenta graficamente i valori $C_i = \sum_{j=1}^i (A_j - \mu_0)$ (μ_0 = valore di riferimento per la media) dell'i-esimo campione su n campioni totali. I C_i sono le somme cumulate degli scostamenti del valore A dell'i-esimo campione dal parametro di riferimento per la media. Siccome le CUSUM possono analizzare sia osservazioni individuali che le medie dei valori di vari sottogruppi, il parametro A può essere un valore individuale o la media dell'i-esimo sottogruppo.

Sia le EWMA che le CUSUM sono ottime carte di controllo per verificare se un certo parametro dei campioni analizzati si sta discostando anche di poco da un certo valore di riferimento.

Le carte CUSUM possono essere di due tipi:

- una carta contenente due one-sided CUSUM (è la CUSUM di base). Questa carta di controllo è costituita da due grafici unilaterali: Upper CUSUM che permette di verificare l'esistenza di spostamenti oltre l'UCL del processo; Lower CUSUM che permette di verificare l'esistenza di spostamenti oltre LCL del processo;
- una carta contenente una CUSUM two-sided. Piuttosto che i limiti di controllo a 3σ , questo grafico usa una V-mask, che sono due limiti di controllo che si incrociano formando una V;

Per aprire la finestra di dialogo principale delle CUSUM, bisogna seguire la seguente procedura: Stat → Control Chart → Time-weighted charts → CUSUM. A questo punto bisogna selezionare le o la colonna coi dati e inserire la dimensione campionaria dei sottogruppi. Inoltre, è possibile:

- modificare il layout della carta tramite i pulsanti «Scale...» e «Labels...»;
- specificare un «Target» diverso da 0,0. Questo è importante perché il target rappresenta il parametro di riferimento, μ_0 ;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;



In “CUSUM options...” e sotto la voce «Plan/Type», una volta spuntato il flag «One-sided (LCL, UCL)» si genera la carta CUSUM unilaterale e c'è la possibilità:

- sotto la voce «Type of CUSUM», di spuntare il flag «Use FIR», che inizializza la carta CUSUM unilaterale. Ciò serve nel caso in cui il processo è fuori controllo all'inizio del suo studio;
- sotto la voce «CUSUM Plan», specificare i valori di h e k. Di base, h e k sono impostati, rispettivamente, a 4 e 0,5.

In “CUSUM options...” e sotto la voce «Plan/Type», per generare una carta CUSUM V-mask, bisognerà spuntare il flag «Two-sided (V-mask)». Volendo, c'è anche la possibilità di:

1. specificare il sottogruppo a cui centrare l'apice della maschera a V;
2. specificare i valori di h e k sotto «CUSUM Plan». Di base, h e k sono impostati, rispettivamente, a 4 e 0,5. Nelle carte bilaterali: H è parte dell'equazione usata per calcolare metà della larghezza della V-mask; k definisce l'inclinazione dei bracci della maschera a V.

3.6.4 – Zone chart

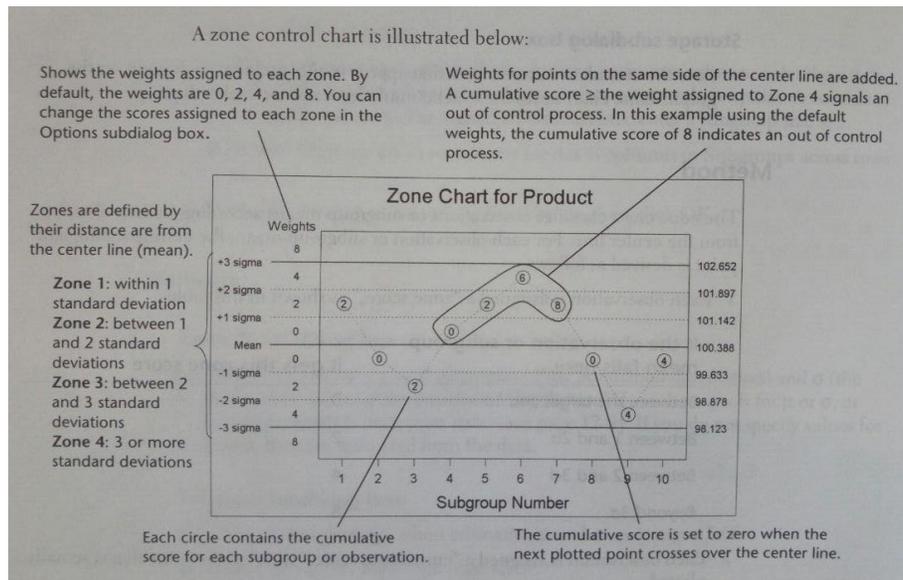


Figura 3.0.10 Esempio di Zone chart

Esso è un incrocio tra l'X medio chart (or Individuals chart) e il CUSUM. Esso mostra, per ogni dato raccolto o per medie di sottogruppi, graficamente un punteggio cumulativo (dalle osservazioni precedenti) basato su zone a 1, 2 e 3 sigma dalla linea centrale. Ne segue che il grafico è suddiviso in 4 zone sia sopra che sotto la linea di riferimento della media.



Grazie a ciò questa carta è facile da utilizzare: un punto è fuori controllo quando il suo punteggio (cumulativo) è maggiore o uguale 8 e non serve rintracciare altri schemi nei dati, dunque non vi sono «Tests» per cause speciali.

Questo può essere utilizzato sia con medie di sottogruppi di dati che con dati individuali.

La procedura per aprire la finestra di dialogo principale del Zone chart è la seguente: Stat → Control Chart → Variables charts for subgroup → Zone. Così facendo apparirà la finestra a lato, in cui andranno compilati i vari campi. Successivamente si potrà:

- modificare il layout della carta tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;
- modificare le opzioni del grafico tramite il pulsante «Zone Options...».

3.7 - Variables control charts for Short Runs (per piccole quantità di dati)

I grafici visti precedentemente sono affidabili nella stima dei parametri di un processo, come la media e la deviazione standard, solo nel caso in cui si abbiano a disposizione grandi quantità di dati.

Quando la quantità di dati è limitata come, per esempio, nel caso in cui si debba studiare un processo produttivo o un macchinario che produce pochi pezzi ma per molti modelli diversi o come nel caso in cui si debbano studiare i parametri di pezzi prodotti in piccole quantità, allora bisognerà usare la carta di controllo Z-MR.

3.7.1 – Z-MR chart

In questa carta vi è un grafico per le osservazioni individuali standardizzate (Z) e uno per le medie mobili (MR).

Per aprire la sua finestra di dialogo, bisogna seguire la seguente procedura: Stat → Control charts → Variables Charts for Individuals → Z-MR. Dopo di che, bisognerà indicare come sono raggruppati di dati e indicarne le colonne interessate. Inoltre, si potrà:

- modificare il layout della carta di controllo tramite i pulsanti «Scale...» e «Labels...»;
- includere o escludere dati dallo studio tramite il pulsante «Data options...»;



Capitolo 4

Distribution analysis

Minitab mette a disposizione comandi che consentono di descrivere alcune caratteristiche del ciclo di vita di un prodotto, di un componente, di una persona o di un organismo. Per esempio, si potrebbe essere interessati a stimare:

- per quanto tempo un certo componente di una macchina funzionerà correttamente (non si romperà) sotto certe condizioni;
- oppure per quanto tempo un paziente sopravvivrà dopo un intervento chirurgico importante.

Lo scopo di quest'analisi è stimare la distribuzione del tempo di guasto di un prodotto. Infatti, raccolti i dati sul ciclo di vita si potranno usare i comandi presenti in questo capitolo per selezionare la distribuzione che meglio si adatta a loro, per poi stimare la funzione che meglio rappresenta la distribuzione. Le distribuzioni di dati possono essere:

- **parametriche**, quando è possibile assumere che i dati seguono una distribuzione parametrica;
- **non parametriche**, quando non è possibile l'assunzione del punto precedente. In questo caso, Minitab fornisce stime non parametriche per le stesse funzioni.

Per conoscere quale comando di analisi delle distribuzioni usare, bisognerà considerare due aspetti:

1. se posso o meno assumere una distribuzione parametrica per i miei dati. In questo caso si sceglieranno i comandi di «parametric distribution analysis»; in caso contrario, si sceglieranno i comandi di «nonparametric distribution analysis». Fatto ciò, si dovrà scegliere:
2. il tipo di censura applicata ai dati (vedi tabella 4.0.1). Se si sceglie la censura a destra o se si conosce il tempo di guasto esatto del componente, allora si usano i comandi «right-censoring»; se i tempi di guasto sono stati censurati secondo gli altri criteri o un loro assortimento, allora si usano i comandi «arbitrary-censoring».

Tutti questi comandi possono essere usati indipendentemente dal tipo di dati censurati (censored data) che ho a disposizione.

I comandi di analisi parametrica dei dati includono:

- la «Parametric Distribution Analysis», che esegue l'analisi completa e adatta una delle otto distribuzioni parametriche (Weibull, extreme value, exponential, normal, lognormal base e, lognormal base 10, logistic e loglogistic) ai dati e usa questa distribuzione per stimare percentuali e probabilità di sopravvivenza (survival probabilities), e tracciare grafici di sopravvivenza (survival) e del tasso di guasto e probability plots per i tuoi dati;



- i grafici «Distribution ID Plot» che tracciano fino a quattro probability plots, a partire dalle distribuzioni scelte tra le seguenti otto: Weibull, extreme value, exponential, normal, logonormal base e, logonormal base 10, logistic e loglogistic. I quattro grafici aiuteranno a scegliere la distribuzione che, tra queste, si adatta al meglio ai tuoi dati;
- «Distribution Overview Plot», spesso usati prima dell’analisi completa per scegliere una distribuzione o vedere informazioni generali, tracciano un probability plot, una funzione di densità di probabilità, un grafico di sopravvivenza (survival plot), e un grafico del tasso di guasto (hazard plot). Questi ti aiuteranno a valutare l’adattamento della distribuzione scelta precedentemente e mostreranno dei grafici riassuntivi sui tuoi dati.

I comandi di analisi non parametrica delle distribuzioni includono:

- la «Nonparametric Distribution analysis» che fornisce stime non parametriche su probabilità di sopravvivenza, stime del tasso di guasto, altre stime dipendenti dal tipo di tecnica non parametrica scelta, e tracciano i grafici del tasso di guasto e di sopravvivenza;
- «Distribution overview plot» (uncensored/right-censored data only) che traccia un grafico di sopravvivenza Kaplan-Meier o Actuarial e un grafico del tasso di guasto;
- «Distribution Overview plot» (uncensored/arbitrarily censored data only) che traccia un grafico di sopravvivenza Turnbull or Actuarial e un tasso di guasto.

4.1 – Come inserire i dati nel foglio di lavoro

I dati da raccogliere in questo capitolo sono i tempi di guasto individuali dei componenti che possono essere di vario tipo come mostrato dalla tabella 4.0.1. In base al tipo di censura applicata al tempo di guasto, dovrò impostare il foglio di lavoro diversamente.

Type of observation	Description	Example
Exact failure time	You know <i>exactly</i> when the failure occurred.	The fan failed at exactly 500 days.
Right censored	You only know that the failure occurred <i>after</i> a particular time.	The fan had not yet failed at 500 days.
Left censored	You only know that the failure occurred <i>before</i> a particular time.	The fan failed sometime before 500 days.
Interval censored	You only know that the failure occurred <i>between</i> two particular times.	The fan failed sometime between 475 and 500 days.

Tabella 4.0.1 I tipi dei tempi di guasto



I tempi di guasto censurati a destra possono essere:

- **singly**: significa che tutti i pezzi sono stati testati per lo stesso ammontare di tempo e che tutti gli esatti tempi di guasto si sono verificati prima del tempo di censura. Questo tipo di dati sono i più comuni in studi controllati. Questo tipo di dato può essere di due tipi:
 - **time censored**: lo studio continua per un certo periodo di tempo e tutte le unità ancora in funzione alla fine di questo periodo saranno censurate a destra;
 - **failed censored**: lo studio prosegue finché non si verifica un certo numero di guasti.
- **multiply censored**: significa che i pezzi testati sono stati censurati in momenti differenti, dunque per alcuni pezzi si conosce l'esatto tempo di guasto, mentre per altri si ha solo il tempo di censura. Questo tipo di dati sono più comuni, dato che i componenti possono entrare il servizio in tempi differenti.

Il foglio di calcolo va impostato in base al tipo di censura applicata ai dati. Oltre alle colonne indicanti la durata della prova:

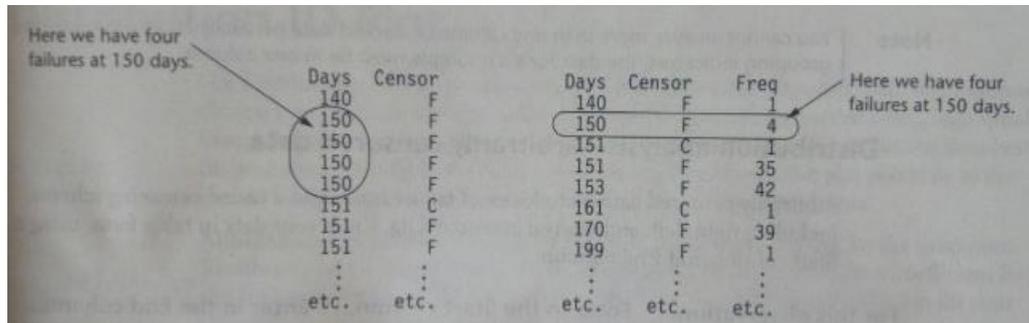
- **nel caso di dati singly time censored**, quindi per usare un tempo di guasto costante che definisce la censura, bisognerà aggiungere una colonna indicante i tempi di guasto per ciascun campione. Più tardi si specificherà il tempo di guasto al quale cominciare la censura;
- **nel caso di dati singly failed censored**, quindi per specificare un numero di fallimenti oltre il quale si attiva la censura, bisognerà inserire una colonna di tempi di guasto per ciascun campione. Poi, si specificherà il numero di fallimenti oltre il quale cominciare la censura.
- **nel caso di dati singly o multiply censored**, per usare le colonne con gli indicatori di censura per definire il momento di censura, bisognerà aggiungere una colonna di tempi di guasto e una che indica gli indicatori di censura. Questi indicatori possono essere numeri o testo.

La colonna dei dati e la colonna della censura associata, devono essere della stessa lunghezza. In basso vi è un esempio.

	Months	Censor	
This column contains failure times for engine windings in a turbine assembly.	50	F	This column contains the corresponding censoring indicators: an F designates an actual failure time; a C designates a unit that was removed from the test, and was thus censored.
	60	F	
	53	F	
	40	F	
	51	F	
	99	C	
	35	F	
	55	F	
	:	:	
	etc.	etc.	

Tabella 4.0.2 Struttura del foglio di lavoro con indicatori di censura

È possibile strutturare il foglio di lavoro tale che ogni riga corrisponda ad una singola osservazione, oppure tramite l'aggiunta di una colonna delle frequenze, far sì che ogni riga contenga un tipo specifico di osservazione a cui corrisponde una cella dedicata alla frequenza di quella osservazione, come nella tabella 4.0.3. Questa seconda metodologia è utile nel caso in cui si abbia un gran numero di dati con tempi di guasto e di censura comuni.



Days	Censor	Days	Censor	Freq
140	F	140	F	1
150	F	150	F	4
150	F	151	C	1
150	F	151	F	35
150	F	153	F	42
151	C	161	C	1
151	F	170	F	39
151	F	199	F	1
...
etc.	etc.	etc.	etc.	etc.

Tabella 4.0.3 Struttura del foglio di lavoro con indicatori di frequenza

4.2 – Distribution ID plot

Si usa questo grafico per rappresentare fino a quattro diversi probability plot, i quali aiutano a determinare quale tra le diverse distribuzioni si adatta al meglio ai dati, i quali sono tempi di guasto individuali, forse censurati. Solitamente ciò è fatto notando quanto i punti cadono in prossimità della curva della distribuzione, in particolare i punti lungo le code della distribuzione.

Inoltre, Minitab fornisce anche il test Anderson-Darling per aiutare a capire se una certa distribuzione descrive bene i dati. Più il valore di Anderson-Darling è basso, più la distribuzione calza bene i dati.

Questo tipo di grafico accetta vari tipi di dati:

- Distribution ID Plot – right censoring accetta esatti tempi di guasto e dati censurati a destra;
- Distribution ID plot – arbitrary censoring accetta esatti tempi di guasto, dati censurati a destra, a sinistra e in un intervallo.

4.2.1 – Distribution ID plot – Right censoring

Per fare questo grafico right – censoring, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Distribution ID Plot – Right Censoring.

In questo modo comparirà una finestra in cui:

- in «variables», inserire la colonna coi tempi di guasto;
- in caso si avesse la colonna con le frequenze, inserirla in «Frequency columns»;
- se tutti i campioni sono raggruppati in una colonna, spuntare «by variable» e inserire la colonna con gli indicatori dei gruppi.

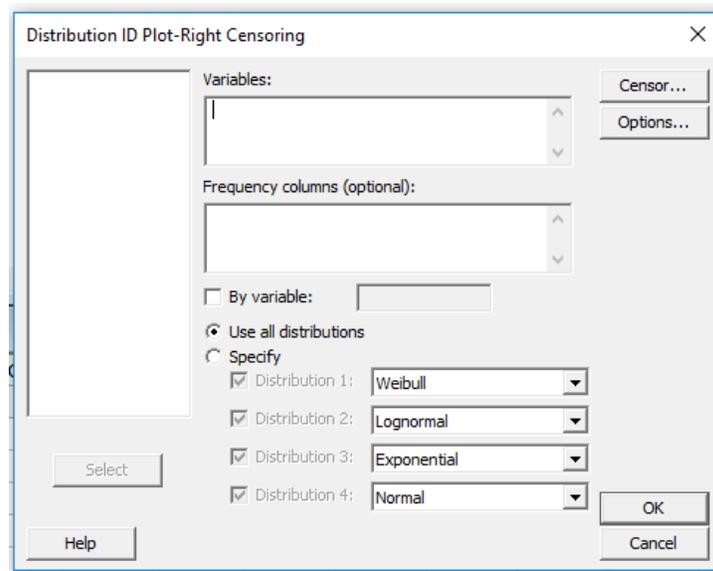


Figura 4.0.1 Esempio di finestra di dialogo principale di un Distribution ID plot

Nel caso in cui i tempi di guasto siano tutti esatti (ovvero, non ci siano tempi censurati) allora si può già pigiare «OK». In caso contrario, bisognerà anche:

- se si sono raccolti anche dati censurati, cliccare «censor» (alla finestra di dialogo principale, figura 4.0.1), così apparirà una finestra. Nel caso in cui ci sia una colonna di indicatori di censura, inserirla in «Use censoring columns»; la prima colonna di questi indicatori è appaiata con la prima colonna di dati, mentre la seconda colonna di indicatori con la seconda colonna di dati;
- nel caso in cui si abbiano dati censurati in base ad un tempo di guasto, selezionare «Time censor at:» e indicare il tempo voluto;
- nel caso in cui si abbiano dati censurati in base al numero di fallimenti, selezionare «failure censor at» e indicare il numero desiderato;
- infine cliccare OK.

4.2.2 - Distribution ID plot – Arbitrary censoring

Per realizzare questo tipo di grafico bisognerà seguire la seguente procedura: Stat → Reliability/Survival → Distribution ID Plot – Arbitrary Cens. Per completare la finestra di dialogo bisogna:

- in «Start variables», inserire la colonna dei tempi di inizio prova;
- in «End variables», inserire la colonna dei tempi di fine prova;
- in caso di colonne di frequenza, inserirle in «Frequency columns»;



- se tutti i dati sono raggruppati in una colonna e loro sottogruppi appartengono a campioni differenti, allora spuntare «by variable» e inserire la colonna con gli indicatori dei campioni;
- è possibile scegliere tra quattro distribuzioni per trovare quella che più si adatta ai dati;
- pigiando su «options», apparirà una finestra.

Per entrambi i Distribution ID plot, dopo aver compilato tutti i campi, cliccando «OK» il programma darà come risultato:

- statistiche di buon adattamento dei dati alla distribuzione scelta;
- tabella delle percentuali (Table of percents) e i loro percentili, errore standard e intervalli di confidenza al 95%;
- tabella dei tempi medi di guasto e dei loro errori standard e gli intervalli di confidenza al 95%;
- quattro probability plot per le distribuzioni: Weibull, logonormal base e, logonormal base 10, exponential e normal.

Minitab fornisce due statistiche di buon adattamento dei dati alla distribuzione scelta:

- l'**Anderson-Darling** per il metodo della massima probabilità (maximum likelihood) e per il metodo least square estimation. Un valore basso della statistica Anderson-Darling indica che la distribuzione calza meglio i dati;
- il **coefficiente di correlazione Pearson** per il metodo last squares estimation. Questo indicatore assumerà un valore compreso tra 0 e 1 e più tende a 1, più la distribuzione calza al meglio i dati.

Esempio 4.2

Si ipotizzi di essere l'addetto aziendale alla manutenzione dei macchinari e di prendere nota, ad ogni guasto, del relativo tempo intercorso tra questi in un apposito registro (ovvero del tempo di guasto o time to failure).

Tra i diversi macchinari, ipotizziamo di scegliere quello a maggior frequenza di guasto e, analizzando il registro dei tempi di guasto, si crei una tabella come quella sottostante.

Prima di poter studiare approfonditamente il tempo di guasto del macchinario, innanzitutto, vogliamo farci un'idea di come questi tempi si distribuiscano, ovvero quale è la distribuzione che meglio segue questi dati.

Così, decidiamo di usare il «Distribution ID Plot», nello specifico il right censoring, perché abbiamo solo tempi esatti di guasto.



↓	C1-D	C2	C
	Data	tff (gg)	
1	05/10/2012	12,00	
2	19/10/2012	9,00	
3	30/10/2012	6,00	
4	31/10/2012	1,00	
5	19/11/2012	12,00	
6	24/11/2012	4,00	
7	27/11/2012	1,00	
8	29/11/2012	1,00	
9	12/12/2012	8,00	
10	25/02/2013	35,00	
11	29/04/2013	40,00	
12	10/05/2013	8,00	
13	24/06/2013	30,00	
14	02/08/2013	28,00	
15	28/08/2013	2,00	
16	29/08/2013	1,00	
17	06/09/2013	5,00	
18	10/09/2013	1,00	
19	04/11/2013	38,00	
20	19/11/2013	10,00	
21	11/12/2013	15,00	
22	18/12/2013	4,00	
23	24/02/2014	36,00	
24	11/03/2014	20,00	
25	07/04/2014	18,00	
26	07/04/2014	18,00	
27	24/04/2014	12,00	
28	13/06/2014	35,00	
29	03/07/2014	13,00	
30	07/07/2014	1,00	
31	11/07/2014	3,00	
32	28/08/2014	18,00	
33	01/09/2014	1,00	
34	09/09/2014	5,00	
35	12/09/2014	2,00	
36	18/09/2014	3,00	
37	10/11/2014	37,00	
38	13/11/2014	2,00	
39	19/11/2014	3,00	
40	25/11/2014	3,00	
41	05/12/2014	7,00	
42	15/01/2015	16,00	
43	26/01/2015	6,00	
44	20/03/2015	38,00	
45	14/04/2015	16,00	
46	21/04/2015	4,00	
47	15/06/2015	38,00	
48	08/09/2015	50,00	
49	16/11/2015	46,00	
50	19/11/2015	2,00	
51	16/11/2015	46,00	
52	18/12/2015	20,00	
53	14/01/2016	5,00	
54	17/02/2016	18,00	
55	19/02/2016	1,00	
56	29/04/2016	54,00	
57	30/05/2016	20,00	
58	24/06/2016	18,00	
59	13/10/2016	31,00	
60	18/10/2016	4,00	
61	20/10/2016	1,00	
62	10/11/2016	14,00	
63	30/11/2016	13,00	
64	13/01/2017	3,82	

Ne segue che inseriremo nella finestra «Variables» la colonna «tff (gg)» e cliccheremo «OK». A questo punto, compariranno più probability plot (figura 4.0.2), in cui potremo comparare la bontà delle varie distribuzioni (proposte da Minitab) all’adattamento ai dati.

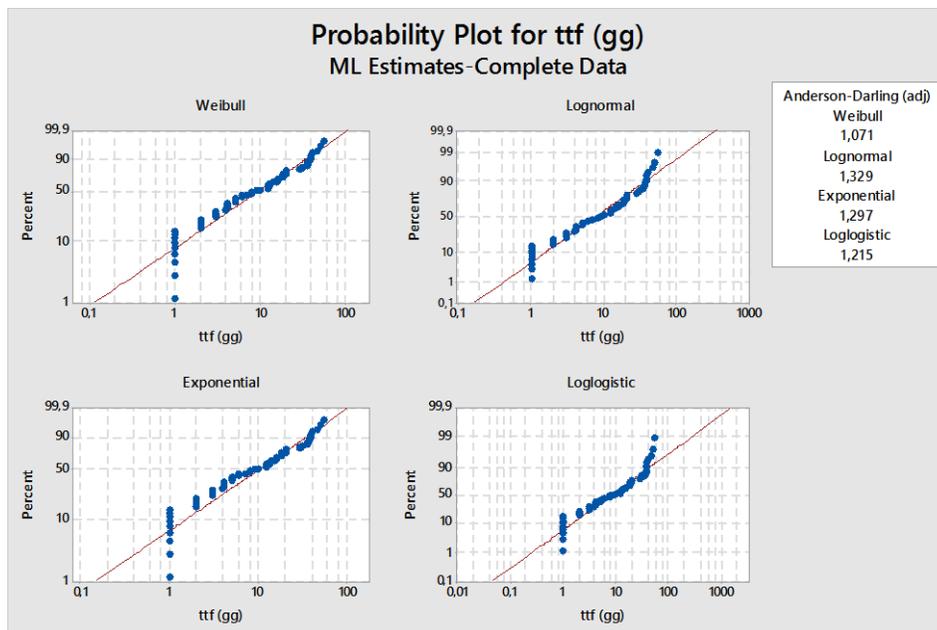


Figura 4.0.2 Probability plot del Distribution ID plot



Per scegliere quella più adatta al nostro esempio, basterà comparare i valori «Anderson-Darling» delle varie distribuzioni e scegliere quella col valore più basso (vedi figura 4.0.2).

Oltre a fornire la finestra 4.0.2, Minitab fornisce nella «Session Window» una tabella da cui non solo si potranno confrontare i valori «Anderson-Darling» ma anche avere informazioni sul tempo di guasto per ogni distribuzione, come quella alla tabella 4.0.4. Per esempio, per la distribuzione Weibull, il tempo di guasto, in media, si aggira attorno ai 14,65 giorni per il macchinario analizzato.

Distribution	Mean	Standard Error	95% Normal CI (CI = Intervallo di confidenza)	
			Lower	Upper
Weibull	14,65	1,95	11,28	19,02
Logonormal	17,01	3,57	11,27	25,68
Exponential	14,65	1,86	11,42	18,80

Tabella 4.0.4 Tabella risultante di un Distribution ID plot

4.3 – Distribution overview plot

Si usa questo grafico per generare un layout di grafici che permetteranno di osservare i tempi di guasto di un componente in vari modi. Questo tipo di comando può essere usato sia per i dati parametrici che non parametrici.

Nel caso si utilizzi il Distribution Overview Plot per dati parametrici, Minitab fornirà:

- un probability plot (per una distribuzione prescelta, magari utilizzando il Distribution ID Plot);
- un diagramma di affidabilità (reliability plot);
- una funzione di densità di probabilità e un diagramma di rischio (hazard plot).

Nel caso si utilizzi il Distribution Overview Plot per dati non parametrici, Minitab fornirà:

- nel caso di dati censurati a destra, un Kaplan-Meier diagramma di sopravvivenza (o affidabilità) e un diagramma di rischio o un Acturial diagramma di sopravvivenza e un diagramma di rischio;
- nel caso di dati censurati arbitrariamente, un Turnbull diagramma di sopravvivenza o un Acturial diagramma di sopravvivenza e un diagramma di rischio. In ogni caso, queste funzioni sono tutte tipici modi di descrivere la distribuzione dei tempi di guasto.



I dati utilizzati in questo comando sono tempi di guasto individuali, alcuni dei quali possono essere censurati.

Come detto, Distribution Overview Plot accetta sia tempi di guasto precisi che censurati, così, prima di iniziare questa analisi, occorrerà disporre i dati correttamente nel foglio di lavoro.

4.3.1 - Distribution overview plot – Right censoring

Per aprire la finestra di dialogo principale (figura 4.0.3) che permette di usare questo tipo di Distribution Overview Plot, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Distribution Overview Plot – Right.

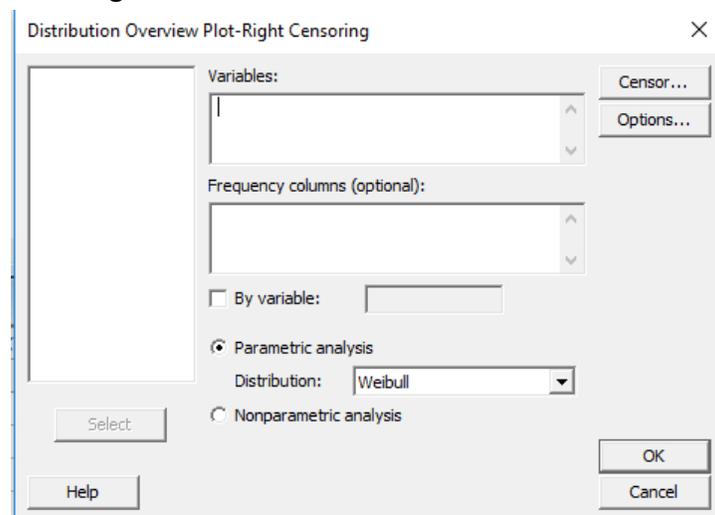


Figura 4.0.3 Esempio di una finestra di dialogo principale di un Distribution Overview Plot

In questo modo, apparirà una finestra dove:

- in «**Variables**» si inserisce la colonna (o fino a dieci colonne) con i tempi di guasto;
- in «**Frequency**» la o le colonne con le frequenze (opzionale);
- se tutti i dati sono incolonnati in una unica colonna, anche se appartenenti a gruppi (a campioni) diversi, allora spuntare «**By variable**» e inserire la colonna con tutti i dati;
- scegliere se tracciare una distribuzione non parametrica o parametrica. In questo secondo caso, scegliere la distribuzione più adatta; ciò può essere fatto tramite il Distribution ID Plot.

Nel caso in cui i dati siano tutti esatti tempi di guasto, allora si può già cliccare su «OK», in caso contrario cliccare su «**Censor...**», così apparirà una finestra in cui:

- in «**Use censoring columns:**» inserire le colonne con gli indicatori di censura, ricordando che la prima di queste colonne sarà accoppiata con la prima colonna di dati e così via. Volendo puoi inserire il valore al quale vuoi effettuare la censura in «**Censoring value**»;
- per i dati censurati in base al tempo, indicare il tempo di censura in «**Time censor at:**»;



- per i dati censurati in base al numero di guasti, allora inserire in «**Failure censor at:**» il numero dei guasti oltre il quale censurare;
- infine cliccare «OK».

4.3.2 - Distribution overview plot – Arbitrary censoring

Per aprire la finestra di dialogo che permette di usare questo tipo di Distribution Overview Plot, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Distribution Overview Plot – Arbitrarily Censoring. In questo modo, apparirà una finestra in cui:

- in «**Start Variables**» si inserirà la colonna (o fino a dieci colonne) con i tempi di inizio esperimento;
- in «**End Variables**» si inserirà la o le colonne con i tempi di fine esperimento;
- nel caso si avesse la colonna delle frequenze, allora bisognerà inserirla in «**Frequency columns**»;
- se tutti i dati sono incolonnati in una unica colonna, anche se appartenenti a gruppi (a campioni) diversi, allora spuntare «**By variable**» e inserire la colonna con tutti i dati;
- scegliere se tracciare una distribuzione non parametrica o parametrica. In questo secondo caso, scegliere la distribuzione più adatta; ciò può essere fatto tramite il **Distribution ID Plot**.

Esempio 4.3

Dall'esempio 4.2, abbiamo capito che la Weibull è la distribuzione che meglio si adatta ai dati. A questo punto, si vuole calcolare anche la funzione di affidabilità del macchinario e il suo tasso di guasto.

Dato che tutti i tempi di guasto sono esatti e non censurati, allora si userà il comando Distribution Overview Plot – right censoring. Così, dopo aver compilato la finestra di dialogo principale e dato l'«OK», si otterrà il grafico alla figura 4.0.4.

Questa mostra:

- la funzione del tasso di guasto (o di rischio) (Hazard function), solitamente indicata con $\lambda(t)$, definito come: «il rapporto tra il numero di oggetti guasti dopo un tempo t e il numero di quelli controllati.»;
- la funzione di sopravvivenza (survival function), solitamente indicata con $R(t)$ che rappresenta la garanzia che il nostro macchinario (o un componente o un organismo) dà di funzionare, dopo un certo tempo che lavora;
- un probability plot che mostra quanto bene la distribuzione scelta si adatta ai dati;
- un probability density function.

Sovrapponendo il cursore del mouse alla curva di ciascun grafico, comparirà una tabella dettagliata. Per esempio, per la funzione di sopravvivenza, possiamo notare che, dopo solo 9,5 giorni di lavoro, la garanzia di funzionamento del nostro macchinario è solo del 50,9157%. Ciò indica che il macchinario non è ulteriormente affidabile, per esempio perché vecchio; quindi sarebbe utile pensare ad una sua sostituzione, nel caso in cui questo sia un macchinario rilevante per il ciclo produttivo.

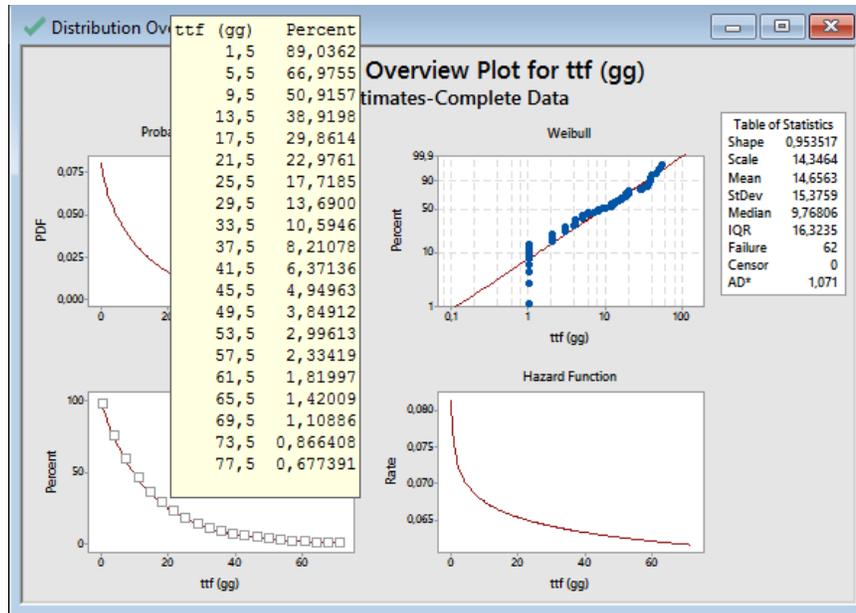


Figura 4.0.4 Distribution overview plot

4.4 – Parametric Distribution Analysis

Questo comando va utilizzato per:

- trovare la distribuzione che meglio si adatta ai tempi di guasto raccolti. Come risultato Minitab mostrerà due tabelle che descrivono la distribuzione: Parameter Estimates e Characteristics of Distribution;
- stimare percentili e probabilità di sopravvivenza. Come risultato Minitab mostrerà, rispettivamente, la Table of percentiles e la Table of survival probabilities;
- valutare se la distribuzione è appropriata. Come risultato Minitab mostrerà un probability plot;
- tracciare il diagramma di sopravvivenza, di rischio e un probability plot;

I dati raccolti sono tempi di guasto individuali, alcuni dei quali possono essere censurati.

Con la parametric distribution analysis, si possono inserire fino ad un massimo di dieci campioni di dati appartenenti a diversi componenti o macchinari od organismi. Così, si potranno confrontare i risultati ottenuti per i vari componenti e scegliere, per esempio, quello più affidabile o quello col tasso di guasto più basso.



In base al tipo di dati che si sta analizzando bisogna utilizzare tipi diversi di Parametric Distribution Analysis. Se si sono raccolti esatti tempi di guasto e tempi di guasto censurati a destra, allora bisogna utilizzare la Parametric Distribution Analysis – right censoring; se si sono raccolti tempi di guasto censurati arbitrariamente, allora bisogna utilizzare la Parametric Distribution Analysis – arbitrary censoring.

4.4.1 - Parametric distribution analysis – Right censoring

Per aprire la finestra di dialogo che permette di usare questo tipo di Parametric Distributio Analysis – right censoring, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Parametric Distributio Analysis – Right censoring. Per completare la finestra appena aperta:

- in «**Variables**» si inserisce la colonna (o fino a dieci colonne) con i tempi di guasto;
- in «**Frequency**» la o le colonne con le frequenze (opzionale);
- se tutti i dati sono incolonnati in una unica colonna, anche se appartenenti a gruppi (a campioni) diversi, allora spuntare «**By variable**» e inserire la colonna con tutti i dati;
- scegliere se tracciare una distribuzione non parametrica o parametrica. In questo secondo caso, scegliere la distribuzione più adatta; ciò può essere fatto tramite il **Distribution ID Plot**.

Nel caso in cui i dati siano tutti esatti tempi di guasto, allora si può già cliccare su «OK», in caso contrario cliccare su «**Censor...**», così apparirà una finestra in cui:

- in «**Use censoring columns:**» inserire le colonne con gli indicatori di censura, ricordando che la prima di queste colonne sarà accoppiata con la prima colonna di dati e così via. Volendo puoi inserire il valore al quale vuoi effettuare la censura in «**Censoring value**»;
- per i dati censurati in base al tempo, indicare il tempo di censura in «**Time censor at:**»;
- per i dati censurati in base al numero di guasti, allora inserire in «**Failure censor at:**» il numero dei guasti oltre il quale censurare;
- infine cliccare «OK».

4.4.2 – Parametric distribution analysis – Arbitrary censoring

Per aprire la finestra di dialogo che permette di usare questo tipo di Parametric Distributio Analysis – arbitrary censoring, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Parametric Distributio Analysis – arbitrary censoring. In questo modo, apparirà la finestra da completare come segue:

- in «**Start Variables**» si inserisce la colonna (o fino a dieci colonne) con i tempi di inizio esperimento;



- in «**End Variables**» si inserisce la o le colonne con i tempi di fine esperimento;
- nel caso si avesse la colonna delle frequenze, allora bisogna inserirla in «**Frequency columns**». Ciò è opzionale;
- se tutti i dati sono incolonnati in una unica colonna, anche se appartenenti a gruppi (a campioni) diversi, allora spuntare «**By variable**» e inserire la colonna con tutti i dati.

I risultati di base delle Parametric Distribution Analysis sono:

- l'informazione di censura;
- i parametri stimati e i loro errori standard e intervalli di confidenza al 95%;
- log-likelihood e Anderson-Darling goodness-of-fit statistic;
- caratteristiche delle distribuzioni e i loro errori standard e gli intervalli di confidenza al 95%;
- una tabella dei percentili e i loro errori standard e intervalli di confidenza al 95%;
- probability plot.

In entrambe le parametric distribution analysis, si hanno a disposizione otto distribuzioni (Weibull; extreme value; exponential; normal; lognormal base e; lognormal base 10; logistic; loglogistic distribution) tra le quali trovare quella che meglio si adatta ai dati del tempo di guasto.

Scelta la distribuzione ed avviata l'analisi, sulla session window compariranno due tabelle che descrivono la distribuzione scelta e quanto bene si adatta ai dati raccolti:

- nella tabella «**Parameter Estimates**», vedi tabella 4.0.5, viene mostrato the maximum likelihood o il least squares stime dei parametri della distribuzione, i loro errori standard e un intervallo di confidenza al 95%, statistiche (Anderson-Darling e log-likelihood) di buon adattamento ai dati della distribuzione scelta;

Estimation Method: Maximum Likelihood				
Distribution: Weibull				
Parameter Estimates				
Parameter	Estimate	Standard Error	95.0% Normal CI	
			Lower	Upper
Shape	2.3175	0.3127	1.7790	3.0191
Scale	73.344	5.203	63.824	84.286

Log-Likelihood = -186.128

Goodness-of-Fit
Anderson-Darling = 67.6366

Tabella 4.0.5 Parameter Estimates

- nella tabella «**Characteristics of Distribution**», vedi tabella 4.0.6, vengono mostrate misure comuni del centro e della diffusione della distribuzione con un intervallo di confidenza al 95%. La media e la deviazione standard non sono resistenti a lunghi tempi di vita, mentre lo sono la mediana, il primo e il terzo percentile e il range interquartile.



Characteristics of Distribution				
	Estimate	Standard Error	95.0% Normal CI	
			Lower	Upper
Mean (MTTF)	64.9829	4.6102	56.5472	74.6771
Standard Deviation	29.7597	4.1463	22.6481	39.1043
Median	62.6158	4.6251	54.1763	72.3700
First Quartile(Q1)	42.8439	4.3240	35.1546	52.2151
Third Quartile(Q3)	84.4457	6.2186	73.0962	97.5575
Interquartile Range(IQR)	41.6018	5.5878	31.9730	54.1305

Tabella 4.0.6 Characteristics of distribution

Oltre alle tabelle precedenti, Minitab mostra anche i percentili.

Nel caso in cui si dovesse rispondere a domande simili alle seguenti:

- in che momento metà dei campioni si romperanno?
- quanto tempo trascorrerà prima che il 10% dei componenti si rompa?

per rispondere bisogna usare i percentili.

I comandi della Parametric Distribution Analysis automaticamente mostrano i percentili 1-10, 20, 30, 40, 50, 60, 70, 80, 90-99.

Table of Percentiles					
	Percent	Percentile	Standard Error	95.0% Normal CI	
				Lower	Upper
At about 10 months, 1% of the windings failed.	1	10.0765	2.7845	5.8626	17.3193
	2	13.6193	3.2316	8.5543	21.6834
	3	16.2590	3.4890	10.6767	24.7601
	4	18.4489	3.6635	12.5009	27.2270

Tabella 4.0.7 Table of percentiles

Nella tabella 4.0.7 dei percentili sono stati studiati i tempi di guasto (in mesi) di un componente e, per esempio, dalla prima riga, si può notare che dopo 10 mesi (seconda colonna, «Percentile») l'1% (prima colonna, «Percent») dei componenti si sarà guastato. Nelle altre tre colonne viene mostrato l'errore standard sul percentile e i limiti inferiore e superiore di un intervallo di confidenza al 95% per la stima del percentile.

Per aiutare a verificare se una particolare distribuzione si adatta bene ai dati, Minitab fornisce anche un probability plot (figura 4.0.5). Questo grafico consiste di:

- **plot points**, che rappresentano la proporzione di fallimenti fino ad un certo tempo;
- **fitted line**, rappresentazione grafica dei percentili;
- un intervallo di confidenza al 95% per la fitted line.

Siccome i plot points non dipendono da nessuna distribuzione, prima di essere trasformati, questi sono uguali per tutti i probability plot. Comunque, la fitted line varia in base alla distribuzione parametrica scelta. Così si può usare il probability plot per verificare se una distribuzione calza i tuoi

dati. In generale, più i punti si sovrappongono alla fitted line, più la distribuzione descrive al meglio i dati raccolti.

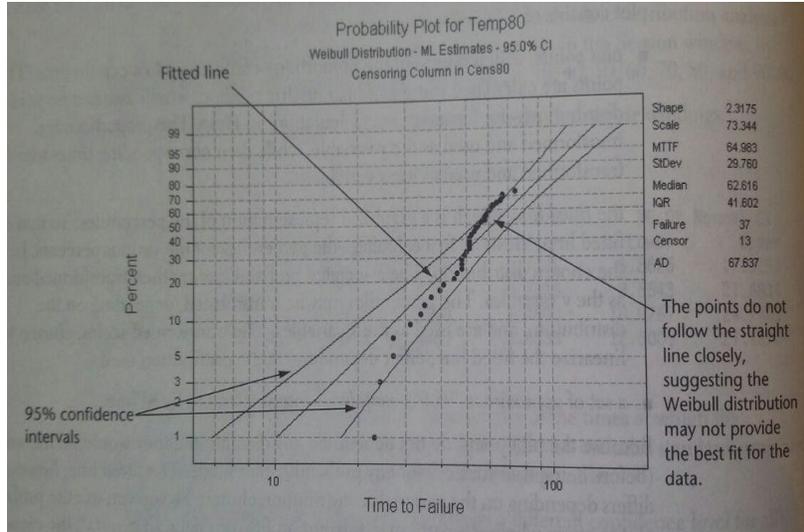


Figura 4.0.5 Probability plot

Minitab fornisce anche una tabella e un grafico circa la probabilità di sopravvivenza.

Nel caso in cui si dovesse rispondere a domande simili alle seguenti:

- qual è la probabilità che un componente superi senza guasti un certo periodo di tempo?
- qual è la probabilità che un paziente malato sia ancora vivo dopo cinque anni dalla somministrazione delle medicine?

Per rispondere, dobbiamo trovare le probabilità di sopravvivenza, che sono stime della proporzione di unità non guaste dopo un certo tempo.

Quando richiedi le probabilità di sopravvivenza nella sottofinestra di dialogo «Estimate», queste verranno mostrate nella Session Window, come nella tabella 4.0.8.

Table of Survival Probabilities			
40.76% of the engine windings last past 70 months.	Time	Probability	95.0% Normal CI
	70.0000	0.4076	Lower Upper
			0.2894 0.5222

Tabella 4.0.8 Table of Survival Probabilities

I grafici di sopravvivenza o affidabilità mostrano come varia la probabilità di sopravvivenza rispetto al tempo. Ogni grafico rappresenta la proporzione di unità ancora funzionanti dopo il tempo t . La curva di sopravvivenza è circondata dalle altre due curve dell'intervallo di confidenza al 95%, che forniscono un valore ragionevole per la vera funzione di sopravvivenza.



Per tracciare il grafico parametrico di sopravvivenza, nella finestra di dialogo principale delle “Parametric distribution analysis” cliccare su «Graphs...», così comparirà una finestra in cui bisogna spuntare «Survival Plot».

Il grafico di rischio (o del tasso di guasto o Hazard plot) mostra il tasso di guasto istantaneo in ogni momento.

Spesso, il grafico del tasso di rischio è:

- alto all’inizio del grafico (tempo iniziale). Questa zona viene chiamata fase della mortalità infantile;
- basso nel mezzo. Questa zona è chiamata fase della vita normale;
- nuovamente alto a fine grafico (tempi alti). Questa zona è chiamata fase di usura.

In questo modo, il grafico richiama la forma di una vasca da bagno.

N.B. I grafici di rischio di MINITAB non assomigliano ad una vasca da bagno; Infatti, i fallimenti in zone differenti della curva a vasca da bagno sono causate da diverse modalità di rottura, mentre Minitab stima la distribuzione del tempo di fallimento causata solo da una modalità di rottura. In figura 4.0.6 vi è mostrato un esempio.

Per tracciare un grafico parametrico del tasso di rischio, bisogna cliccare su «Graphs...» nella finestra di dialogo principale delle “Parametric distribution analysis” e spuntare «Hazard Plot».

Minitab usa il metodo di stima della massima somiglianza o il metodo least squares per stimare i parametri della distribuzione.

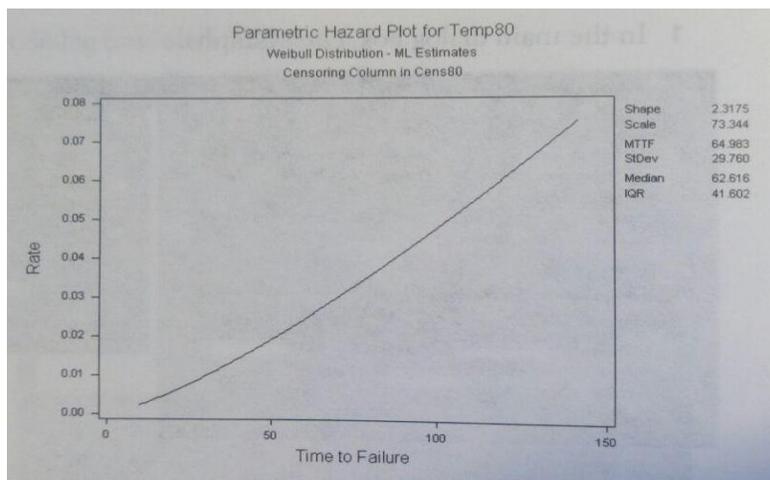


Figura 4.0.6 Parametric hazard plot



Altrimenti, se si preferisce, si potranno usare propri parametri. In questo caso, non verrà eseguita nessuna stima; tutti i risultati, come i percentili, sono basati sui parametri che sono stati inseriti. Si potrà scegliere di stimare i parametri o col metodo della massima somiglianza (maximum likelihood) o con quello della least squares.

Esempio 4.4

Esempio di parametric distribution analysis con tempi esatti di guasto o censurati a destra.

Si supponga di lavorare per una azienda che produce componenti elettrici, i quali si rompono troppo frequentemente alle alte temperature, così si decide di trovare il tempo di rottura di questi componenti ad 80°C e 100°C.

Per ogni temperatura si cerca:

- dopo quanto tempo di utilizzo il 10% dei componenti si sarà guastato;
- la proporzione dei componenti che dopo 70 mesi sarà ancora funzionante.

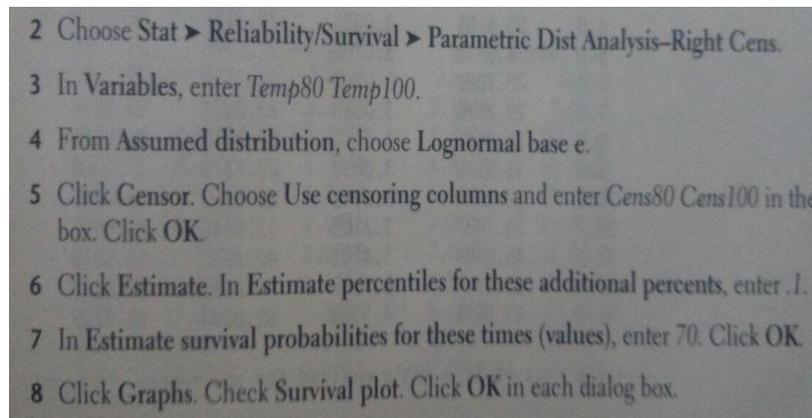


Figura 4.0.7 Procedimento dell'esempio 4.4

Inoltre, si vogliono tracciare due grafici: un probability plot, per verificare che la distribuzione logonormal base e si adatta bene ai dati raccolti; un grafico di sopravvivenza.

Per far ciò, bisognerà: raccogliere un primo campione di tempi di guasto (in mesi) su 50 componenti per la temperatura di 80°C e un secondo campione di tempi di guasto su 40 componenti per la temperatura di 100°C; seguire la procedura elencata in figura 4.0.7.

In questo modo, compariranno varie tabelle di cui considereremo solo la tabella «Table of percentiles» e la tabella «Table of Survival Probabilities», sia per la temperatura di 80°C che di 100°C. Inoltre, compariranno un probability plot e un survival plot. Il tutto è mostrato alla pagina seguente.



Temperatura 80°C:

Table of Percentiles				
Percent	Percentile	Standard Error	95.0% Normal CI	
			Lower	Upper
0.1	13.3317	2.5156	9.2103	19.2975
1.0	19.3281	2.8375	14.4953	25.7722
2.0	22.0674	2.9256	17.0178	28.6154
3.0	24.0034	2.9726	18.8304	30.5975
4.0	25.5709	3.0036	20.3126	32.1906
5.0	26.9212	3.0262	21.5978	33.5566
6.0	28.1265	3.0440	22.7506	34.7727
7.0	29.2276	3.0588	23.8074	35.8819
8.0	30.2501	3.0717	24.7910	36.9113
9.0	31.2110	3.0833	25.7170	37.8788
10.0	32.1225	3.0941	26.5962	38.7970
20.0	39.7837	3.2100	33.9646	46.5999
30.0	46.4184	3.4101	40.1936	53.6073
40.0	52.9573	3.7567	46.0833	60.8568
50.0	59.8995	4.3109	52.0192	68.9735

Tabella 4.0.9 Table of percentiles dell'esempio 4.4 temperatura 80°C

Temperatura 100°C:

Table of Percentiles				
Percent	Percentile	Standard Error	95.0% Normal CI	
			Lower	Upper
0.1	3.9350	1.1729	2.1940	7.0577
1.0	6.8776	1.6170	4.3383	10.9034
2.0	8.3941	1.7942	5.5212	12.7619
3.0	9.5253	1.9111	6.4283	14.1144
4.0	10.4756	2.0015	7.2036	15.2338
5.0	11.3181	2.0766	7.8995	16.2162
6.0	12.0884	2.1419	8.5418	17.1076
7.0	12.8069	2.2003	9.1453	17.9343
8.0	13.4863	2.2538	9.7195	18.7129
9.0	14.1354	2.3034	10.2707	19.4544

Tabella 4.0.10 Table of percentiles dell'esempio 4.4 temperatura 100°C

Dato che nella slide precedente abbiamo notato come a 80°C il 10% dei componenti si romperà dopo 19,3281 mesi, mentre a 100°C il 10% dei componenti si guasterà dopo 6.8776 mesi, possiamo concludere che l'aumento di temperatura aumenterà il numero di componenti che si guastano.

Inoltre, abbiamo le due tabelle:

Temperatura 80°C:

Table of Survival Probabilities			
Time	Probability	95.0% Normal CI	
		Lower	Upper
70.0000	0.3743	0.2631	0.4971

Tabella 4.0.11 Table of survival probabilities dell'esempio 4.4 temperatura 80°C

Temperatura 100°C:

Table of Survival Probabilities			
Time	Probability	95.0% Normal CI	
		Lower	Upper
70.0000	0.1982	0.1072	0.3248

Tabella 4.0.12 Table of survival probabilities dell'esempio 4.4 temperatura 100°C

Possiamo, inoltre, notare che la probabilità di sopravvivenza dei componenti sottoposti alla temperatura di 80°C (37,43%) è maggiore rispetto a quella dei componenti sottoposti a 100°C (19,82%).

Infine abbiamo i due seguenti grafici:

Probability plot

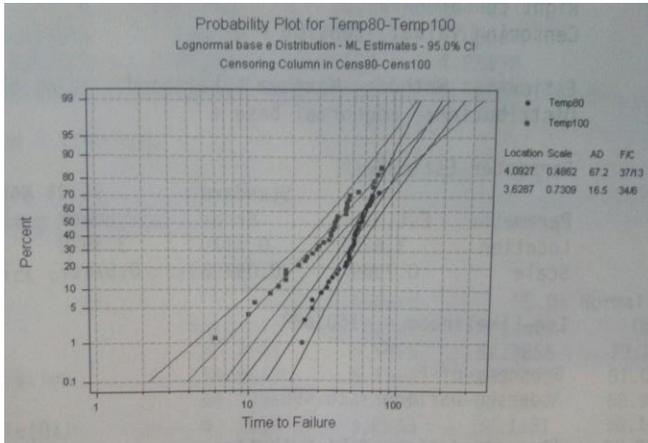


Figura 4.0.8 Probability plot dell'esempio 4.4

Survival plot

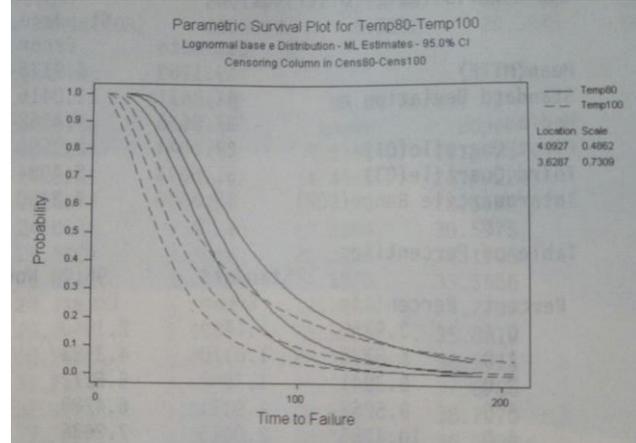


Figura 4.0.9 Survival plot dell'esempio 4.4

4.5 – Nonparametric distribution analysis

Quando nessuna distribuzione si adatta ai dati, si useranno i comandi della nonparametric distribution analysis per stimare:

- le probabilità di sopravvivenza;
- il tasso di rischio;
- tracciare il grafico di rischio e della sopravvivenza.

I dati da raccogliere sono tempi di guasto individuali, alcuni dei quali possono essere censurati. Per esempio, si possono collezionare tempi di guasto di unità lavoranti ad una certa temperatura, o a temperature diverse o sotto vari tipi di stress.

Nel caso in cui i dati siano:

- tempi esatti di guasto o censurati a destra, allora si potrà stimare tramite il metodo Kaplan-Maier o Acturial;
- censurati secondo vari schemi, allora si potrà stimare tramite il metodo Acturial o Turnbull.

Si possono analizzare fino a dieci campioni contemporaneamente. Quando si inseriscono più di un campione, Minitab stimerà le varie funzioni separatamente. Tutti le funzioni vengono mostrate nello stesso grafico ma con differenti colori e simboli, per aiutare a compararli.

La nonparametric distribution analysis accetta vari tipi di dati:

- la **nonparametric distribution analysis – right censoring** accetta tempi esatti di guasto o censurati a destra;



- la **nonparametric distribution analysis – arbitrary censoring** accetta tutti i tipi di tempo di guasto.

Per aprire la finestra di dialogo che permette di fare una nonparametric distribution analysis per tempi esatti di guasto o censurati a destra, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Distribution analysis (Right censoring) → Nonparametric distribution analysis.

Aperta la finestra sarà possibile:

- in «**Variables**» inserire la colonna coi tempi di guasto. Si possono inserire fino a dieci colonne, ciascuna per un campione differente;
- in caso di colonne di frequenza, inserirle in «**Frequency columns**»;
- se tutti i campioni sono raggruppati in una unica colonna, spuntare «**by variable**» e inserire la colonna con gli indicatori che indicano il campione di appartenenza.

Nel caso in cui tutti i tempi di guasto raccolti siano tempi esatti di guasto, allora si può cliccare su «OK». In caso contrario cliccare su «Censor», così apparirà una finestra in cui si potrà:

- se si è creata una colonna con gli indicatori di censura, spuntare «**Censoring columns**» e inserire la colonna contenente gli indicatori di censura. La prima colonna di indicatori di censura è accoppiata con la prima colonna dei tempi di guasto e così via. Volendo, si può inserire in «**Censoring value**» il valore oltre il quale censurare. Se non si inserisce nessun valore, Minitab userà il valore più piccolo della colonna degli indicatori di censura;
- per dati censurati in base al tempo, spuntare «**Time censor at**» e inserire il valore indicante il tempo oltre il quale censurare;
- per dati censurati in base al numero di pezzi guastatisi, spuntare «**Failure censor at**» e inserire il numero di guasti oltre il quale censurare;
- infine, cliccare su «OK».

Per aprire la finestra di dialogo che permette di fare una non parametric distribution analysis per tempi censurati in vario modo, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Distribution analysis (arbitrary censoring) → Nonparametric distribution analysis.

Aperta la finestra sarà possibile:

- In «**start variables**», inserendo la colonna dei tempi di partenza, si possono inserire fino a dieci colonne;
- in «**end variables**», inserendo la colonna coi tempi di fine, si possono inserire fino a dieci colonne;
- nel caso si avessero colonne indicanti le frequenze, si inserirà in «**frequency columns**» la colonna con le frequenze;
- se tutti i campioni sono raggruppati in una unica colonna, si spunterà «**by variable**» e indicare la colonna con gli indicatori di appartenenza ai camion;



- infine, cliccare «OK».

Entrambe le nonparametric distribution analysis forniscono risultati differenti a seconda del tipo di dato da analizzare. Se i tempi di guasto sono del tipo:

- censurati a destra o tempi esatti (right censoring). In questo caso, dall'analisi si ottengono:
 - informazioni sulla censura;
 - caratteristiche della variabile, che includono la media, l'errore standard e intervalli di confidenza al 95%, la mediana e il range interquartile Q1 e Q3;
 - stime delle probabilità di sopravvivenza tramite il metodo Kaplan-Meier e i loro: errori standard e intervalli di confidenza al 95%.
- censurati in vari modo (arbitrary censoring). In questo caso, dall'analisi si ottengono:
 - informazioni sulla censura;
 - stime sulla probabilità di guasto e i loro errori standard tramite il metodo Turnbull;
 - stime della probabilità di sopravvivenza e i loro errori standard e intervalli di confidenza al 95% tramite il metodo Turnbull.

Anche con la nonparametric distribution analysis, è possibile calcolare le probabilità di sopravvivenza, che stimano la proporzione di unità ancora funzionanti in funzione del passare del tempo.

Per tracciare un grafico non parametrico di sopravvivenza, bisogna spuntare «survival plot» nella finestra di dialogo che appare cliccando in «Graphs», nella finestra di dialogo principale.

Per tracciare un grafico del tasso di guasto:

- nel caso della **nonparametric distribution analysis – right censoring**, clicca «Graphs» nella finestra di dialogo principale e, in essa, spunta l'opzione «Hazard plot»;
- nel caso della **nonparametric distribution analysis – arbitrary censoring**, cliccare «Estimates», scegliere «Acturial» e cliccare «OK». Cliccare su «Graphs» e spuntare l'opzione «hazard plot». Infine cliccare «OK».



Capitolo 5

Regression with life data

I comandi di Minitab relativi alla regression with life data consentono di studiare la relazione tra una o più variabili indipendenti e il tempo di guasto di un componente, di un prodotto, di una persona o di un organismo in generale.

Lo scopo della regression with life data è quello di fornire un modello che predica il tempo di guasto. Da queste previsioni si potrà stimare anche l'affidabilità dell'oggetto o dell'organismo studiato.

Le regression with life data sono di due tipi:

- l'**accelerate life testing**, che esegue una semplice regressione con una sola variabile indipendente (predictor), che è usata per descrivere il tempo di guasto di prodotti altamente affidabili. La variabile indipendente è una variabile accelerata, ovvero i suoi livelli di intensità eccedono quelli normalmente trovati nella pratica. I dati ottenuti sotto condizioni estreme verranno poi utilizzati per studiare situazioni a condizioni normali di utilizzo. Per far questo, bisogna avere un buon modello della relazione tra la variabile indipendente e il tempo di guasto del componente o dell'organismo;
- la **regression with life data** esegue una regressione con una o più variabili indipendenti. Il modello può includere factors, covariates, interactions e nested terms. Questo modello aiuta a capire come vari factors e covariates influenzino la vita di un componente o un prodotto.

Si può scegliere tra una delle seguenti otto distribuzioni, per capire quale è la miglior rappresentazione dei propri dati: Weibull, extreme value, exponential, normal, lognormal base e, lognormal base 10, logistic e loglogistic.

5.1 - Come inserire i dati nel foglio di lavoro

La struttura base del foglio di lavoro per la regression with life data consiste in tre colonne, anche se possono essercene più di tre. Le tre colonne sono:

- la «**response variabile**» (variabile dipendente), che è il tempo di guasto;
- la «**censoring indicators**», ovvero la colonna contenente gli indicatori del tipo di censura applicata ai dati;
- la «**predictor variables**» (lo o le variabili indipendenti):
 - per l'accelerating life testing, inserire una colonna per la variabile indipendente contenente vari livelli di intensità per questa variabile;



- per la regression with life data, inserire una colonna per ogni variabile indipendente. Queste variabili indipendenti possono essere considerate come factors o covariates nel modello.

Ogni colonna va strutturata in modo tale che ogni sua riga contenga una singola osservazione o osservazioni uniche con una corrispondente colonna per la loro frequenza d'osservazione.

I comandi descritti in questo capitolo richiedono che i dati sul tempo di guasto siano tempi individuali e possono censurati allo stesso modo della distribution analysis (vedi tabella 4.0.1).

Nel caso in cui i tempi di guasto siano esatti o censurati a destra, bisognerà inserire nel foglio di lavoro due colonne: una unica per i tempi di guasto e una colonna corrispondente per gli indicatori di censura.

5.2 – Accelerated life testing

L'accelerated life testing si usa per studiare la relazione tra il tempo di guasto e una variabile indipendente che lo può influenzare, quindi l'accelerated life testing richiede che si conosca la relazione tra la variabile accelerata e il tempo di guasto.

Dato che in situazioni normali un componente o un organismo potrebbero impiegare molto tempo prima di guastarsi, vengono effettuati gli accelerated tests così da risparmiare tempo e denaro. In questi test, per accelerare il processo di guasto del componente o dell'organismo, l'intensità della variabile indipendente eccede le normali condizioni. Per questo la variabile è chiamata accelerata.

Di seguito sono indicati i passaggi per effettuare l'accelerated life testing:

- impostare i livelli della variabile accelerata;
- prendere nota del tempo di guasto esatto o censurato;
- eseguire l'accelerated life testing analysis, chiedendo a Minitab di risalire al comportamento del componente sotto condizioni normali di utilizzo.

Per la variabile accelerata (la variabile indipendente), è possibile richiedere la trasformazione: Arrhenius; inverse temperature; $\log(e)$; $\log(10)$; lineare (impostata di base). Di base Minitab assume che la relazione sia lineare.

Una relazione logaritmica è usata per modellare la vita di prodotti lavoranti sotto sforzo costante. Solitamente, la relazione logaritmica è usata in combinazione con una log-based time distribution. Applicazioni comuni della trasformazione logaritmica includono l'isolamento elettrico e la fatica nei metalli.



Nell'accelerated life testing, se i dati sono tempi esatti di guasto o censurati a destra o se sono tempi esatti di guasto o censurati arbitrariamente influisce sul come si esegue l'analisi dei dati.

Per eseguire l'accelerated life testing con tempi di guasto esatti o censurati a destra, bisognerà seguire la seguente procedura: Stat → Reliability/Survival → Accelerated Life Testing. Così apparirà una finestra in cui si potrà:

- in «**variables/start variables**», inserire le colonne coi tempi di guasto. Si possono inserire fino a dieci colonne;
- se si hanno colonne che indicano la frequenza delle osservazioni, inserirle in «**Freq. Columns**»;
- in «**Accelerating variable**», inserire la colonna della variabile indipendente.

Se tra i dati non ci sono tempi censurati, allora cliccare «OK», altrimenti si clicchi «Censor», così apparirà una finestra in cui si potrà:

- in «**use censoring columns**», inserire le colonne con gli indicatori di censura. La prima colonna con gli indicatori di censura è accoppiata con la prima colonna contenente i dati e così via. Di base, Minitab usa il valore più basso presente nella colonna degli indicatori di censura per indicare una osservazione censurata. Per un valore di censura alternativo, inserire tale valore in «**Censoring value**»;
- infine, cliccare «OK».

Per eseguire l'accelerated life testing con tempi di guasto esatti o arbitrariamente censurati, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Accelerated Life Testing.

Così apparirà una finestra cui bisognerà spuntare «Responses are uncens/arbitrarily censored data».

Poi, nella finestra si potrà:

- in «**variables/start variables**», inserire le colonne coi tempi di partenza. Si possono inserire fino a dieci colonne;
- in «**end variables**», inserire le colonne coi tempi di fine. Si possono inserire fino a dieci colonne;
- se si hanno colonne che indicano la frequenza delle osservazioni, inserirle in «**Freq. Columns**»;
- in «**Accelerating variable**», inserire la colonna della variabile indipendente.
- infine, cliccare «OK».

Da entrambi questi test si ricaverà una tabella di regressione, un grafico di relazione e un probability plot per ogni livello della variabile accelerata basata sul modello adattato. Vediamoli nel dettaglio.

La regression table mostra:

- I coefficienti stimati per il modello di regressione e i loro:
 - errori standard;
 - z-value e p-value. Il Z-test verifica che il coefficiente sia significativamente diverso da zero; in altre parole, è una variabile indipendente rilevante?;



- un intervallo di confidenza al 95%;
- il parametro di forma (Weibull or exponential) o parametro di scala (altre distribuzioni), una misura della variabilità complessiva e il suo: errore standard; l'intervallo di confidenza al 95%;
- il log-likelihood;
- statistiche Anderson-Darling di buon adattamento per ogni livello della variabile accelerata basata sul modello adattato.

Dai coefficienti di questa tabella sarà possibile ottenere il modello di regressione, il quale stima i percentili della distribuzione dei tempi di guasto:

$$Y_p = \beta_0 + \beta_1 X + \sigma \varepsilon_p,$$

Dove:

- Y_p = pth percentile della distribuzione del tempo di guasto o tempo di guasto o il logaritmo del tempo di guasto;
- β_0 = y-intercetta (costante);
- β_1 = coefficiente di regressione;
- X = variabile indipendente (può essere trasformata);
- σ = parametro di scala;
- ε_p = pth percentile della distribuzione dell'errore.

In base alla distribuzione, $Y_p = \text{tempo di guasto o } \log(\text{tempo di guasto})$:

- per le distribuzioni Weibull, exponential, logonormal base e, logonormal base 10 e le distribuzioni loglogistic, $Y_p = \log(\text{tempo di guasto})$;
- per le distribuzioni normal, extreme value e logistic, $Y_p = \text{tempo di guasto}$;

Quanto $Y_p = \log(\text{tempo di guasto})$, Minitab prende l'antilogaritmo per mostrare i percentili nella scala originale.

Anche il valore di ε_p dipende dalla distribuzione scelta.

Il relation plot mostra i tempi di fallimento in funzione della variabile accelerata mostrando graficamente i percentili per ogni livello della variabile accelerata. Di base, le linee sono tracciate al decimo, cinquantesimo e novantesimo percentile. Il cinquantesimo percentile è una buona stima della durata della vita di un componente sotto le condizioni date.

Arbitrariamente si possono specificare fino a dieci percentili per tracciare e mostrare i tempi di guasto sul grafico.

Il probability plot è creato per ogni livello della variabile accelerata basata sul modello adattato (line) e il modello non parametrico (points).



Di base, il probability plot include i parametri di forma e scala (distribuzioni Weibull e exponential) oppure i parametri di posizione e scala (altre distribuzioni). Questi parametri sono basati sul modello adattato. Questo grafico assume che le osservazioni per ogni livello della variabile accelerata abbiano tutte la stessa forma (distribuzioni Weibull o exponential) e scala (altre distribuzioni).

Il probability plot include anche la statistica Anderson-Darling, che è una misura di buon adattamento.

5.2.2 - Percentiles and survival probabilities

Quando si esegue un accelerated life testing, si sottopongono i componenti a livelli superiori alla norma di una variabile accelerata, per superare le normali condizioni di utilizzo e accelerare il processo di guasto. Ma spesso il dato che si cerca riguarda il comportamento delle unità sotto condizioni normali di utilizzo.

Nella finestra di dialogo dell'opzione «Estimate», si può richiedere a Minitab di estrapolare informazioni su una situazione di normale utilizzo a partire da una situazione accelerata tramite il campo “enter new predictor values”.

Per stimare i percentili e le probabilità di sopravvivenza, si aprirà la finestra di dialogo dell'opzione «Estimate» (vedi figura 5.0.1).

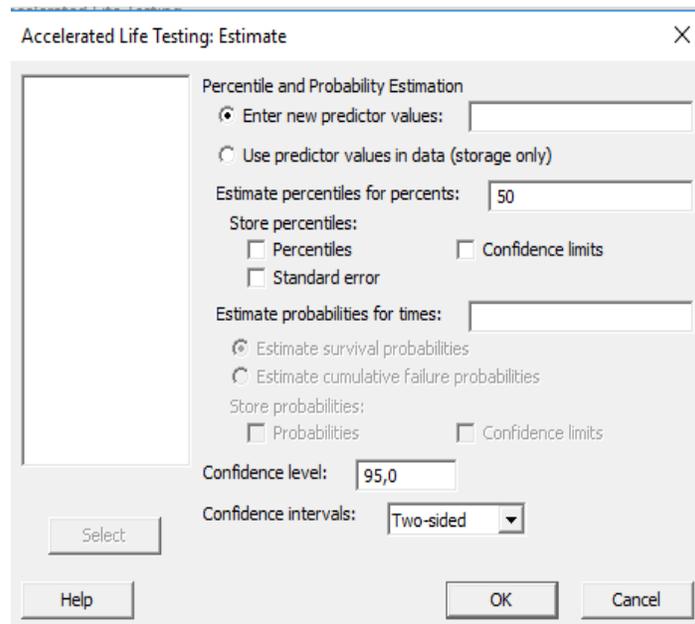


Figura 5.0.1 Finestra di dialogo principale dell'opzione Estimate dell'Accelerated life testing



In essa si potrà:

- in «**enter new predictor values**» inserire un nuovo valore o una colonna di nuovi valori. Spesso si inserisce la variabile designata, o condizioni normali di utilizzo, per le unità;
- stimare i percentili, inserendo le percentuali in «**estimate percentiles for percents**». Di base Minitab stima il cinquantesimo percentile. Se si vuole analizzare il completo ciclo vita di un prodotto in funzione di una certa variabile indipendente, si inserirà 10, 50 e 90. Così Minitab stimerà quanto tempo passi prima che il 10%, il 50% e il 90% delle unità si rompa;
- per stimare le probabilità di sopravvivenza, si inseriranno i tempi in «**estimate survival probabilities for times**». Per esempio, quando si inseriscono 70 (ore), Minitab stimerà la probabilità, per ogni variabile indipendente, che le unità siano ancora funzionanti dopo 70 (ore).

Qualche volta si vorranno stimare i percentili o le probabilità di sopravvivenza per i livelli della variabile accelerata usata nello studio. Per far ciò, nella finestra a lato, spunta «**Use predictor values in data (storage only)**». Minitab mostrerà i risultati nel foglio di lavoro.

Esempio 5.1

Si immagini di voler studiare come un isolante per motori elettrici, sottoposto ad una certa temperatura, si deteriora al passare del tempo. Normalmente, i motori lavorano tra gli 80 e i 100°C. Per non sprecare tempo e soldi si decide di studiare il deterioramento dell'isolante tramite l'accelerated life testing.

Innanzitutto, per accelerare il deterioramento dell'isolante, si raccoglieranno i tempi di guasto di motori lavoranti a temperature più alte del normale, come 110°C, 130°C, 150°C e 170°C. Da questi dati, sarà possibile estrapolare i dati di deterioramento dell'isolante tra gli 80°C e i 100°C.

Per vedere quanto bene il modello si adatta, si tratterà un probability plot basato sui residui standardizzati.

Per vedere se il modello si adatta ai dati trovati, allora bisognerà seguire la seguente procedura:

1. seguire Stat → Reliability/Survival → accelerated life testing;
2. in «**Variables/Stat variables**», inserire i tempi di fallimento e in «**Accelerating variable**» inserire le temperature superiori alla norma;
3. in «**Relationship**» scegliere «**Arrhenius**»;
4. in «**Censor**» e in «**Use censoring columns**» inserire la colonna con gli indicatori di censura;
5. in «**Graphs**» e in «**Enter design value to include on plot**», inserire il valore di riferimento della variabile accelerata;
6. in «**Estimate**» e in «**Enter new predictor values**» inserire i nuovi valori della variabile indipendente;



7. infine, cliccare «OK».

Fatto ciò, compariranno varie tabelle e due grafici, come mostrato di seguito:

Regression with Life Data: FailureT versus Temp

Response Variable: FailureT

Censoring Information Count
Uncensored value 66
Right censored value 14
Censoring value: Censor = C

Estimation Method: Maximum Likelihood
Distribution: Weibull
Transformation on accelerating variable: Arrhenius

Regression Table		Standard			95.0% Normal CI	
Predictor	Coef	Error	Z	P	Lower	Upper
Intercept	-15.1874	0.9862	-15.40	0.000	-17.1203	-13.2546
Temp	0.83072	0.03504	23.71	0.000	0.76204	0.89940
Shape	2.8246	0.2570			2.3633	3.3760

Log-Likelihood = -564.693

Tabella 5.0.1 Regression table dell'esempio 5.1

Anderson-Darling (adjusted) Goodness-of-Fit

At each accelerating level

Level	Fitted Model
110	*
130	*
150	*
170	*

Table of Percentiles			Standard	95.0% Normal CI	
Percent	Temp	Percentile	Error	Lower	Upper
50	80.0000	159584.5	27446.85	113918.2	223557.0
50	100.0000	36948.57	4216.511	29543.36	46209.94

Tabella 5.0.2 Table of percentiles dell'esempio 5.1

Dalla tabella di regressione (tabella 5.0.1) si ottengono i coefficienti per il modello di regressione. In questo esempio, si è applicata una distribuzione Weibull, così, questo modello, descrive la relazione tra temperatura e tempo di guasto dell'isolamento secondo l'equazione:

$$\log_e(\text{failure time}) = -15,1874 + 0,83072(\text{ArrTemp}) + 2,8246\varepsilon_p$$

dove $\text{ArrTemp} = \frac{11604,83}{\text{Temp}+273,16}$ e Temp è la temperatura in °C rilevata sul campo.

La tabella dei percentili (tabella 5.0.2) mostra i cinquantiesimi percentili per le temperature inserite (80 e 100°C). Il cinquantiesimo percentile è una buona stima del tempo di guasto dell'isolamento. A 80°C l'isolamento resisterà per circa 159.548,5 ore; mentre a 100°C l'isolamento durerà per circa 36.948,57 ore.

Col relation plot (figura 5.0.3) si può guardare alla distribuzione dei tempi di guasto per ogni temperatura, nell'esempio il decimo, cinquantiesimo e novantesimo percentile.

Il probability plot (figura 5.0.2) basato sul modello adattato aiuta a determinare se la distribuzione, la trasformazione e l'assunzione di uguale forma ad ogni livello della variabile accelerata sono appropriate.

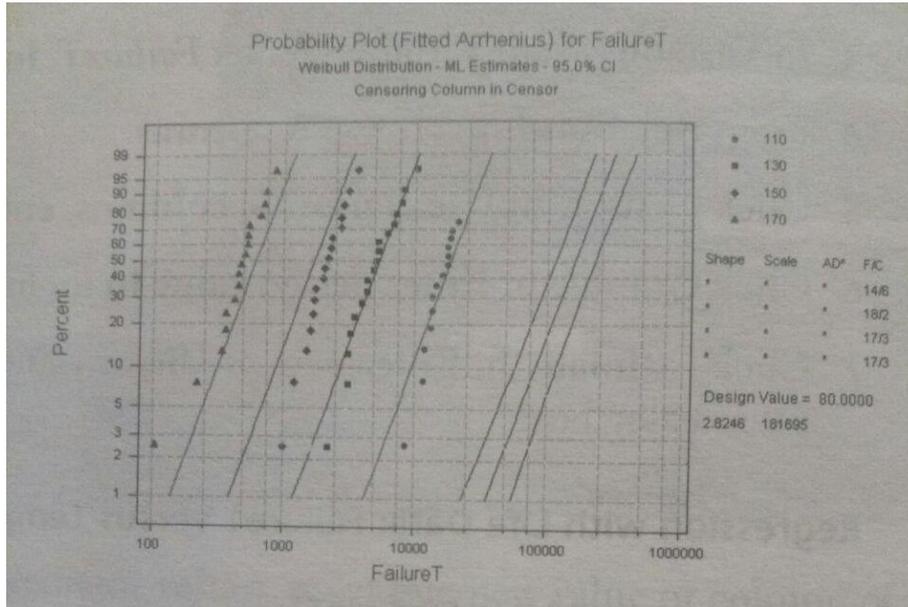


Figura 5.0.2 Probability plot dell'esempio 5.1

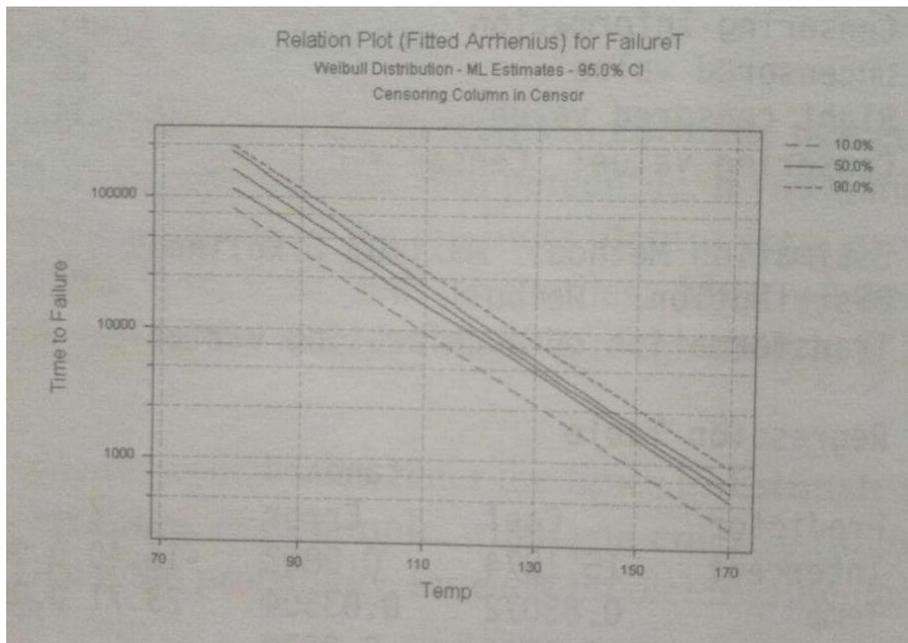


Figura 5.0.3 Relation plot dell'esempio 5.1



5.3 - Regression with life data

La regression with life data si usa per vedere se una o più variabili indipendenti influenzano il tempo di guasto di un prodotto, eventualmente così da definire un modello che predica i tempi di guasto. Questo modello usa la variabile indipendente per spiegare cambiamenti nella variabile dipendente (il tempo di guasto), per esempio perché alcuni prodotti si guastano velocemente, mentre altri non si guastano per lungo tempo. Il modello può includere factors, covariates, interactions e nested terms.

Regression with life data si distingue dalla regressione di Minitab per il fatto che accetta dati censurati e usa distribuzioni differenti.

Per fare una regression with life data, si ha bisogno dei seguenti dati:

- la variabile dipendente, ovvero i tempi di guasto;
- model terms, che consiste di qualsiasi numero di variabili indipendenti e quando appropriato, varie interazioni tra le variabili indipendenti. Alcuni dei termini possono essere factors.

5.3.1 - Come inserire i dati nel foglio di lavoro

Si inseriscano tre tipi di colonne nel foglio di lavoro:

- quella della variabile dipendente (tempi di guasto);
- quella degli indicatori di censura per la variabile dipendente;
- quella della variabile indipendente, che possono essere fattori (categorical variables) o covariates (continuous variables).

A meno che non si specifichi una variabile indipendente come un factor, la variabile indipendente è considerata una covariate.

Nel modello, i termini possono essere creati da queste variabili indipendenti e trattate come factors, covariates, interactions o nested terms.

Per eseguire una regression with uncensored/right censored data, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Regression with life data. Così comparirà la finestra a lato, in cui si potrà:

- in «**variables/start variables**», inserire fino a dieci colonne di tempi di guasto (da dieci differenti campioni);
- il «**Freq. Columns**» inserire le colonne di frequenza, se si hanno;
- in «**Model**», inserire i model terms. Se alcuni di queste variabili indipendenti sono factors, inseriscili anche in «**Factors**».



Se non si hanno dati censurati, allora si potrà cliccare «OK», altrimenti, nel caso si avessero dati censurati, allora cliccare «**Censor**», così comparirà una finestra in cui si potrà:

- in «**use censoring columns**», inserire la colonna con gli indicatori di censura. La prima colonna di indicatori di censura è accoppiata con la prima colonna di dati e così via. Di base Minitab usa il valore più basso nella colonna degli indicatori di censura per indicare un valore censurato. Per usare un altro valore, inserire quel valore in «Censoring value».
- Infine, cliccare «OK».

Per eseguire una regression with uncensored/arbitrary censored data, bisogna seguire la seguente procedura: Stat → Reliability/Survival → Regression with life data. Così comparirà una finestra in cui bisognerà spuntare l'opzione «responses are uncensored/arbitrarily censored data». Nella finestra si potrà:

- in «**Variables/start variables**», inserire fino a dieci colonne di tempi d'inizio (di dieci differenti campioni);
- in «**end variables**», inserire fino a dieci colonne di tempi di arrivo (di dieci differenti campioni);
- se si hanno colonne di frequenza, inserirle in «**Freq. Columns**»;
- in «**Model**», inserire i model terms. Se alcuni di queste variabili indipendenti sono factors, inseriscili anche in «**Factors**»;
- infine clicca «OK».

Nella finestra di dialogo della regression with life data si potrà scegliere una tra le otto più comuni distribuzioni della lifetime di un componente per la distribuzione errore. Le otto distribuzioni sono: Weibull (di base), extreme value, exponential, normal, lognormal base e, lognormal base 10, logistic e loglogistic.

I risultati di base consistono in una tabella di regressione che mostra:

- i coefficienti stimati per il modello di regressione e i loro:
 - ✓ errori standard;
 - ✓ z-value e p-value. Lo z-test verifica che i coefficienti siano significativamente diversi da zero; in altre parole, è una variabile indipendente rilevante?
 - ✓ intervallo di confidenza al 95%.
- il parametro di forma (Weibull o exponential) o il parametro di scala (altre distribuzioni), una misura della variabilità complessiva, e il suo:
 - ✓ errore standard;
 - ✓ intervallo di confidenza al 95%.
- Il log-likelihood.



5.3.2 How to specify the model terms

Si possono adattare modelli con:

- fino a 9 factors e fino a 50 covariates;
- crossed o nested factors;
- covariates che sono crossed l'un l'altro o con factors, o nested con factors.

In figura 5.0.3 ci sono alcuni esempi. A è un factor e X un covariate.

Model terms	
A X A*X	fits a full model with a covariate crossed with a factor
A X	an alternative way to specify the previous model
A X X*X	fits a model with a covariate crossed with itself making a squared term
A X(A)	fits a model with a covariate nested within a factor

Figura 5.0.4 Esempi di how to specify the model terms

5.3.3 Probability plots

Il comando per la regression with life data traccia i probability plot per residui standardizzati e Cox-Snell. Puoi usare questi grafici per verificare se una particolare distribuzione si adatta ai tuoi dati. In generale, più i punti cadono lungo la linea del probability plot, migliore è l'adattamento del modello ai dati.

MINITAB fornisce una misura del buon adattamento ai dati da parte del modello: la statistica Anderson-Darling. La statistica Anderson-Darling è utile per confrontare l'adattamento dei dati a diverse distribuzioni. Un basso valore della statistica Anderson-Darling indica che la distribuzione si adatta bene ai dati.

5.3.4 Multiple degrees of freedom test

Quando si ha un termine con più di un grado di libertà, si può richiedere un test a più gradi di libertà. Questa procedura verifica se il termine è significativo o meno. In altre parole: almeno uno dei coefficienti associati con questo termine è significativamente diverso da zero?

Per effettuare un multiple degrees of freedom test, bisogna cliccare l'opzione «Results» nella finestra di dialogo principale della regression with life data. Spuntando l'opzione «In addition, list of factor level values, tests for terms with more than 1 degree of freedom», e infine cliccando «OK», si effettuerà il test voluto.



5.3.5 Estimating the model parameters

Minitab usa un algoritmo Newton-Raphson modificato per stimare i parametri del modello. Se si preferisce, si possono inserire dei propri parametri. In questo caso, nessuna stima verrà eseguita; tutti i risultati - come i percentili - sono basati su questi parametri.

Quanto MINITAB stima i parametri dai dati, opzionalmente si può:

- inserire valori di partenza per l'algoritmo;
- cambiare il massimo numero di iterazioni per raggiungere la convergenza. Minitab ottiene le maximum likelihood estimates attraverso un processo iterativo. Se il massimo numero di iterazioni è raggiunto prima della convergenza, il processo termina;
- stimare altri coefficienti del modello mentre il parametro di forma (Weibull) o il parametro di scala (altre distribuzioni) rimangono fissati ad uno specifico valore.

Perché è possibile inserire dei valori di partenza per l'algoritmo? La maximum likelihood solution può non convergere se le stime iniziali non sono nelle vicinanze della vera soluzione, così si potrebbe voler specificare quale si pensa sia un buon valore di partenza per le stime dei parametri.

In tutti i casi, inserisce una colonna con inserimenti che corrisponde ai termini del modello nell'ordine in cui li si è inseriti nel Model box. Coi modelli complicati, trova l'ordine di inserimento per le stime di partenza guardando alla tabella di regressione tra i risultati.

Esempio 5.2

Si immagini di voler studiare come un isolante per motori elettrici si deteriora al passare del tempo, se sottoposto ad una certa temperatura. Si vuole sapere se è possibile predire i tempi di guasto per l'isolante in base all'impianto in cui è stato prodotto e alla temperatura alla quale deve lavorare. Si sa che tra la temperatura di lavoro e il tempo di guasto vi è la relazione Arrhenius.

Innanzitutto, si devono raccogliere, per ogni impianto, i dati sull'isolante per le temperature 110°C, 130°C, 150°C e 170°C. Solitamente gli isolanti lavorano tra la temperatura di 80°C e di 100°C, dunque dai dati raccolti si vorrà dedurre il comportamento dell'isolante a queste temperature più basse.

Per vedere quanto bene il modello si adatta, si tratterà un probability plot basato sui residui standardizzati.

Per vedere se il modello si adatta ai dati trovati, allora bisognerà seguire la seguente procedura:

1. Seguire Stat → Reliability/Survival → regression with life data;
2. In «Variables/Stat variables, inserire i tempi di fallimento;



3. In “Model” inserire le temperature Arrhenius ($ArrTemp = \frac{11604,83}{Temp+273,16}$, con Temp le temperature in °C raccolte). In “Factors” inserire gli impianti.
4. In «Censor» e in «Use censoring columns» inserire la colonna con gli indicatori di censura;
5. In «Estimate» e in «Enter new predictor values» inserire i nuovi valori della variabile indipendente;
6. In «Graphs», spuntare «Probability plot for standardized residuals».
7. Infine, cliccare «OK».

Fatto ciò, compariranno varie tabelle e un grafico, il tutto è mostrato alle ultime due figure.

Regression Table						
Predictor	Coef	Standard Error	Z	P	95.0% Normal CI	
					Lower	Upper
Intercept	-15.3411	0.9508	-16.13	0.000	-17.2047	-13.4775
ArrTemp	0.83925	0.03397	24.71	0.000	0.77267	0.90584
Plant 2	-0.18077	0.08457	-2.14	0.033	-0.34652	-0.01501
Shape	2.9431	0.2707			2.4577	3.5244

Log-Likelihood = -562.525

Anderson-Darling (adjusted) Goodness-of-Fit

Standardized Residuals = 0.5078

Tabella 5.0.3 Regression table esempio 5.2

Table of Percentiles						
Percent	Predictor Row Number	Percentile	Standard Error	95.0% Normal CI		
				Lower	Upper	
50	1	182093.6	32466.16	128389.8	258260.9	
50	2	151980.8	25286.65	109689.6	210577.6	
50	3	41530.38	5163.756	32548.44	52990.94	
50	4	34662.51	3913.866	27781.00	43248.61	

Tabella 5.0.4 Table of percentiles esempio 5.2

Dalla regression table (tabella 5.0.3) si ottengono i coefficienti per il modello di regressione. Per la distribuzione Weibull, di seguito ci sono le equazioni che descrivono la relazione tra la temperatura e il tempo di guasto per l'isolante per l'impianto 1 e 2, rispettivamente:

- $\log_e(\text{failure time}) = -15,3411 + 0,83925(ArrTemp) + 2,9431\varepsilon_p$;
- $\log_e(\text{failure time}) = -15,5219 + 0,83925(ArrTemp) + 2,9431\varepsilon_p$,

con $ArrTemp = \frac{11604,83}{Temp+273,16}$.



La table of percentiles (tabella 5.0.4) mostra i cinquantiesimi percentili per le combinazioni di temperature e impianti inseriti.

- per i motori lavoranti alla temperatura di 80°C, l'isolamento prodotto nell'impianto uno dura circa 182.093,6 ore; mentre l'isolamento prodotto nell'impianto due dura circa 151.980,8 ore;
- per i motori lavoranti alla temperatura di 100°C, l'isolamento prodotto nell'impianto uno dura circa 41.530,38 ore; mentre l'isolamento prodotto nell'impianto due dura circa 34.662,51 ore.

Dai bassi valori dei p-value, si può dedurre che gli impianti sono significativamente differenti al valore di $\alpha=0,05$, e la temperatura è una variabile indipendente rilevante.

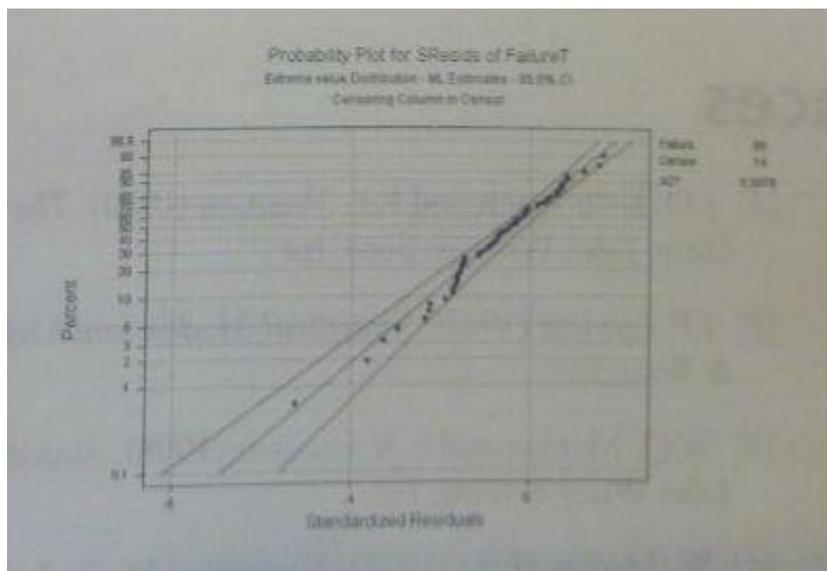


Figura 5.0.5 Probability plot esempio 5.2

Il probability plot per i residui standardizzati aiuterà a determinare se la distribuzione, la trasformazione, e la forma uguale (Weibull o exponential) o il parametro di scala (altre distribuzioni) sono assunzioni adeguate. Dalla figura 5.0.4 possiamo notare che la retta segue bene i punti e dunque possiamo assumere che il modello sia adeguato ai dati raccolti.



Bibliografia

- “User’s Guide 2: Data Analysis and Quality Tools” di “MINITAB Statistical Software”;
- “Controllo statistico della qualità” di “Douglas C. Montgomery”, Seconda edizione, McGraw-Hill;
- “Statistica” di “David M. Levine, Timothy C. Krehbiel e Mark L. Berenson”, Quinta edizione, Pearson;

Sitografia

- [http://support.minitab.com/en-us/minitab/17/topic-library/quality-tools/quality-tools/run-chart-basics/;](http://support.minitab.com/en-us/minitab/17/topic-library/quality-tools/quality-tools/run-chart-basics/)