

Minicorso Controllo Statistico di Processo

di Andrea Saviano

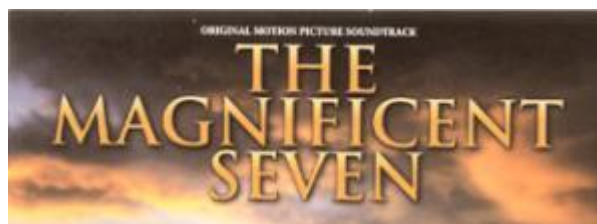
Parte 3

- I magnifici 7, premessa
- 7 strumenti per me posson bastare: variabili e attributi
- I diagrammi: cartesiani, a torta, a barre, a radar, a bolle e chi più ne ha, più ne metta
- I fogli di raccolta dati, quelli di riscontro e i diagrammi di concentrazione dei difetti
- Il raggruppamento in classi omogenee e i grafici di Pareto
- I diagrammi causa-effetto
- L'analisi per stratificazione
- I grafici a dispersione e l'analisi di correlazione e indipendenza
- L'analisi dei processi tramite diagrammi di flusso

Premessa



La storia si svolge in una (f)azienda di un paese di frontiera dove un consulente professionista, viene ingaggiato dal capo della (f)azienda perché ne protegga i prodotti dalle scorrerie dello scarto e dei suoi non conformi. Il professionista assolda così sette compari per difendere il processo di produzione dai non conformi.



I sette s'installano nella (f)azienda e, quando arriva lo scarto con i suoi non conformi, li affrontano e li mettono in fuga. Il giorno dopo l'intera (f)azienda festeggia l'avvenimento, ma ben presto tra gli operai si manifesta un sentimento di disagio per la presenza di questi strumenti della qualità. Appurato che il professionista costa troppo e che i non conformi sembrano ormai sconfitti, il padrone della (f)azienda decide di dare il benservito al consulente, anche perché (a suo avviso) gli operai sono troppo ignoranti per sprecare i soldi in formazione e addestramento.

Tuttavia, una volta andatosene il consulente, lo scarto ritorna insieme a tutti i non conformi. A questo punto il professionista prende il compito che gli era stato assegnato come una questione d'onore, ritorna indietro a difendere il processo di produzione insegnando agli operai come servirsi dei sette. Nel successivo scontro gli operai e i sette riusciranno nel loro intento, anche se a caro prezzo, perché alcuni dipendenti, grazie questa formazione superiore, si licenzieranno per andare a lavorare in altri settori più tecnici.

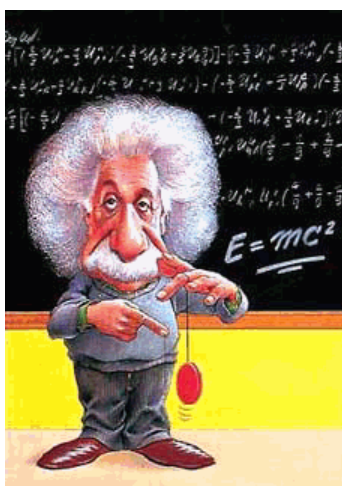
Dei superstiti, i più anziani gradualmente andranno in pensione, mentre i più giovani si appassioneranno alle nuove metodologie migliorando la qualità delle produzioni e abbassando i costi dovuti agli sprechi.



Acclamato dal padrone della (f)azienda, visto che ormai gli operai sanno affrontare i non conformi, il consulente saluterà tutti con un po' di tristezza negli occhi. Dopo essere salito sul suo cavallo, s'incamminerà per recarsi verso la linea dell'orizzonte, lì dove il sole tramonta. Quindi, oltre il confine fra cielo e terra, troverà un'altra (f)azienda, altri operai, altri non conformi da affrontare e sconfiggere.

The End

7 strumenti per me posson bastare: variabili e attributi



*La teoria è quando si sa tutto e niente funziona.
La pratica è quando tutto funziona e nessuno sa il perché.
Noi abbiamo messo insieme la teoria e la pratica:
non c'è niente che funzioni, e nessuno sa il perché!*¹

La logica alla base degli strumenti statistici della qualità è semplice e, per certi aspetti, persino banale:

- **accorciamento della catena decisionale**, le decisioni devono essere prese:
 - subito,
 - dove accadono gli eventi,
 - dalle persone che stanno conducendo l'attività in questione;

¹ Albert Einstein.

- **sintesi**, le informazioni devono essere divulgate/registrate in una forma:
 - regolamentata e condivisa,
 - facilmente sintetizzabile in indici o grafici,
 - chiara, semplice ed estremamente intuibile;
- **tracciabilità**, un processo deve poter essere analizzato:
 - per le attività che sono state svolte,
 - per i guasti/anomalie che si sono manifestati,
 - per le non conformità a cui ha dato luogo.

*Una grande quantità di piccole decisioni, prese rapidamente,
fa meglio il lavoro di una grande decisione, presa lentamente.*

La misura fa parte della vita quotidiana: usando strumenti di misura, si possono raccogliere dati, da utilizzare allo scopo di rappresentare quantitativamente e anche qualitativamente il mondo che ci circonda. Per questo occorre che ognuno di noi misuri diversi tipi di grandezze, progettando anche esperimenti di misura, per poi passare a descrivere i risultati, in un processo che deve portare a comprendere le differenze tra:

- la misura come procedimento pratico, aspetto tipico delle **scienze sperimentali**,
- la misura come teoria, aspetto tipico delle **scienze teoriche** (statistica e stocastica).

Di una caratteristica si può misurare:

- la sua quantità, nel qual caso si lavora per **variabili**;
- la sua presenza (o assenza), nel qual caso si lavora per **attributi**.

Questa prima selezione è importante, perché indirizza su quali strumenti preferire.

I sette strumenti che vengono proposti sono:

- i diagrammi;
- i fogli di raccolta dati,
- l'analisi di Pareto;
- i diagrammi causa-effetto;
- l'analisi per stratificazione;
- l'analisi per correlazione;
- i diagrammi di flusso.

Rispetto alla tradizione sono stati inseriti i diagrammi di flusso e tolte le carte di controllo sia perché si tratta di strumenti preliminari, volti a decidere quali carte di controllo usare, sia perché l'argomento carte di controllo è così complesso da necessitare una trattazione specifica.

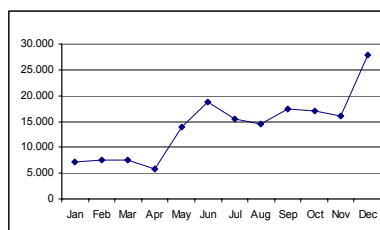
I diagrammi: cartesiani, a torta, a barre, a radar, a bolle e chi più ne ha, più ne metta

Il modo più semplice per sintetizzare i dati è rappresentato dai grafici e, in particolare, dai diagrammi.

Esiste una varietà sterminata di diagrammi, ma qui esamineremo solo le forme principali evidenziando i pregi che alcuni di essi rivestono in particolari applicazioni.

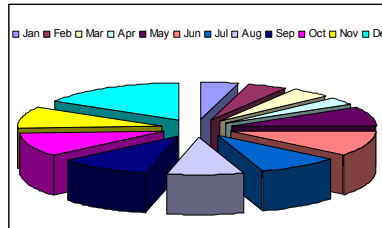
Qualsiasi foglio elettronico di calcolo attualmente rende disponibili i seguenti strumenti grafici:

- **line chart** (diagrammi cartesiani a due variabili), si tratta di diagrammi su cui tracciare dei punti la cui altezza in ordinata fornisce il valore e la cui posizione in ascissa fornisce l'ordine sequenziale/temporale, i valori così ottenuti possono essere collegati tramite segmenti o linee smussate che tentano di emulare per interpolazione l'andamento nel tempo della variabile o dell'attributo che si è esaminato;

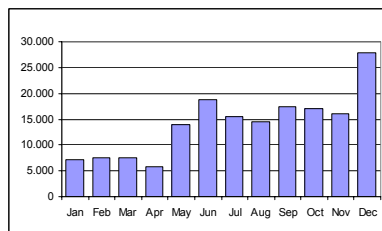


- **pie chart** (diagrammi a torta o circolare), si tratta di diagrammi che riportano in forma di torta l'aspetto qualitativo del controllo sotto forma di percentuale sintetizzando, ad esempio, come può

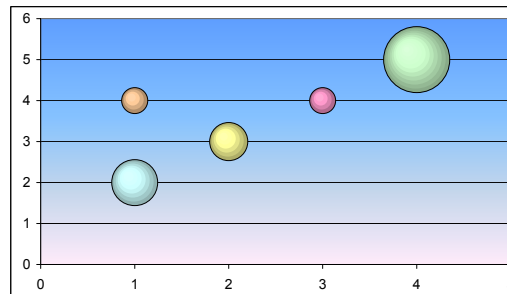
essersi suddiviso in classi di difettosità l'esito del controllo utilizzando cerchi di diametro differente a seconda della numerosità del campione esaminato;



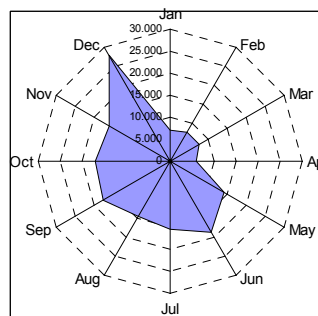
- **histogram & bar chart** (istogrammi o diagrammi a barre verticali/orizzontali), si tratta di un diagramma che distribuisce delle quantità finite in un controllo per attributi o rende leggibile un andamento di probabilità di un controllo per variabili associando i valori continui in classi d'appartenenza;



- **buble chart** (diagrammi a bolle), si tratta di un diagramma in grado di restituire informazioni quando si serve restituire l'andamento di due attributi che caratterizzano un medesimo elemento esaminato, fornendo al contempo una dimensione di numerosità.



- **radar chart** (diagrammi radar), si tratta di un diagramma in grado di restituire informazioni quando si serve restituire l'andamento di più di due attributi che caratterizzano un medesimo elemento esaminato, dal centro partono degli assi su cui si pongono le scale rappresentative dei vari attributi, quindi con delle linee si collegano i valori adiacenti ottenendo una specie di ragnatela.



A questi grafici, che oserei definire "classici", se ne possono abbinare altri di più coreografici, come quelli per i cosiddetti "cruscotti aziendali".



Come preparare un grafico

I dati che si raccolgono durante una campagna di rilevazione sono usualmente dispersi. Per conoscere e comprendere il significato di questa dispersione, le informazioni non vanno solo raccolte, ma anche organizzate. Questo è il fine per cui si realizzano i diagrammi, cioè per fornire una visione globale ed organizzata dei dati campionati. Ovviamente l'aspetto grafico deve essere in grado di proporre un significato alle informazioni che tenta di rielaborare. L'obiettivo è quello di riorganizzare in modo chiaro ed univoco ciò che espresso in maniera "oscura" dalla tabella che li può contenere.

Per conseguire questo fine con efficacia, è necessario innanzitutto conoscere il campo di variabilità (*range*) della dispersione. Tale valore può essere identificato in maniera univoca dal massimo e dal minimo assunti dai dati oppure sotto forma di intervallo o sotto forma di escursione.

In base al campo di variabilità è poi possibile raggruppare i dati in classi omogenee, suddividendo tale intervallo in un numero limitato di sotto-intervalli (da un minimo di tre).

Il modo più semplice è quello di collegarli ai:

- quartili,
- decili,
- principali percentili.

Un diagramma è quindi un metodo grafico in grado di fornire a vista una rappresentazione della dispersione/addensamento dei dati.

Per costruire un diagramma occorre

- individuare la massima escursione dei dati;
- definire degli intervalli di suddivisione in modo da ottenere delle classi di eguale ampiezza ed evitare che il dato si disperda troppo togliendo al grafico la significatività;
- definire i limiti delle classi in modo da evitare che alcuni valori possano cadere a cavallo tra una classe e quella successiva;
- classificare i dati in tabelle di frequenza;
- tracciare i grafici corredandoli di:
 - tutte le leggende necessarie (ad esempio: rendendo disponibili: il valor medio, la deviazione standard o altri indici importanti),
 - una scala che permetta di realizzare una griglia di significatività;
- costruire la curva caratteristica operativa o l'andamento probabilistico di tipo normale di raffronto;
- evidenziare tramite segmenti verticali la posizione del valore obiettivo e il campo di tolleranza previsto.

Come analizzare un diagramma

Un grafico è essenzialmente uno strumento di sintesi dei dati che deve permettere un controllo "a vista" dei processi. Un occhio "allenato" può così porre particolare attenzione a:

- la forma (ad esempio: asimmetrie, curtosi, presenza di dati isolati, etc.);
- il posizionamento rispetto alle specifiche ovvero ai valori di controllo o obiettivo;
- il livello di dispersione o di raggruppamento;
- l'eventuale differente distribuzione al variare di alcuni parametri.

Conclusioni

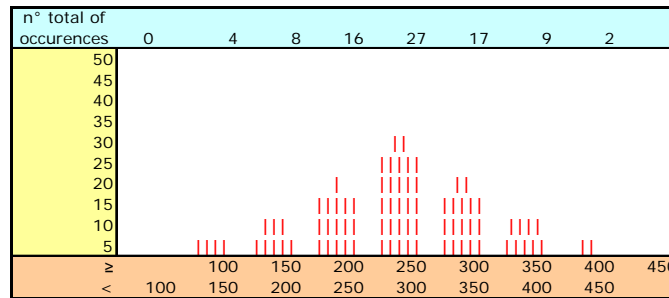
Per poter utilizzare correttamente un grafico (come strumento di sintesi ed analisi) è necessario avere una certa esperienza, sia per quanto concerne l'applicazione delle metodologie della statistica, sia per quanto concerne i più diffusi comportamenti anomali nella distribuzione di un campione rispetto all'andamento della popolazione o al modello probabilistico di riferimento.

Ne consegue che solo la pratica può rendere effettivamente valido l'utilizzo di questo strumento.

I fogli di raccolta dati, quelli di riscontro e i diagrammi di concentrazione dei difetti

La tradizione inserisce come strumento il “**foglio raccolta dati**”, tuttavia questo strumento ha dato luogo a varie forme applicative come:

- **data-sheet**, si tratta del più semplice esempio di carta di raccolta dati, in cui:
 - la prima riga viene riservata al computo dei totali,
 - l’ultima riga viene riservata alla definizione delle classi d’appartenenza,
 - nelle celle intermedie si segnano gli esiti dei controlli con delle barre (|) o delle crocette (×);



- **tally-sheet**, si tratta di una tabella in cui si segnano le ricorrenze, cioè il numero di volte che un predeterminato evento si verifica, nella prima colonna si pone un elenco di possibili difetti, mentre sulla prima riga si segnano i lotti d’appartenenza del campione esaminato, nelle restanti celle si pongono con la notazione $x_{i,j}/n_{i,j}$ i valori registrati, dove:
 - $x_{i,j}$ è il numero di elementi non conformi alla specifica i ,
 - $n_{i,j}$ è la dimensione del campione proveniente dal lotto j ;

	line 1						line 2																	
	1	2	3	4	5	6	Total	p	ln(p)	p*ln(p)	1	2	3	4	5	6	Total	p	ln(p)	p*ln(p)				
feature 1							0										0							
feature 2							0										0							
feature 3							0										0							
feature 4							0										0							
feature 5							0										0							
feature 6							0										0							
feature 7							0										0							
feature 8							0										0							
feature 9							0										0							
feature 10							0										0							
feature 11							0										0							
feature 12							0										0							
feature 13							0										0							
feature 14							0										0							
feature 15							0										0							
feature 16							0										0							
feature 17							0										0							
feature 18							0										0							
feature 19							0										0							
feature 20							0										0							
n° total of occurrences	0	0	0	0	0	0	0									0	0							
							H ^ˆ	0											H ^ˆ	0				
							exp(H ^ˆ)	1											exp(H ^ˆ)	1				
							H ^ˆ max												H ^ˆ max					
							J ^ˆ												J ^ˆ					

- **check-list**, si tratta di una tabella in cui lungo la prima colonna si segnano tutte le attività che devono essere svolte, tali attività non devono necessariamente essere consequenziali o propedeutiche una a quella successiva,

	drafting	inspection	approval	issue
task 1	x	x	x	x
task 2	x	x		
task 3	x	x	x	
task 4				
task 5	x	x	x	x
task 6	x	x	x	
task 7	x			
task 8	x	x		
task 9	x	x	x	
task 10				
task 11	x	x	x	x
task 12	x	x	x	x
task done	10	9	7	4

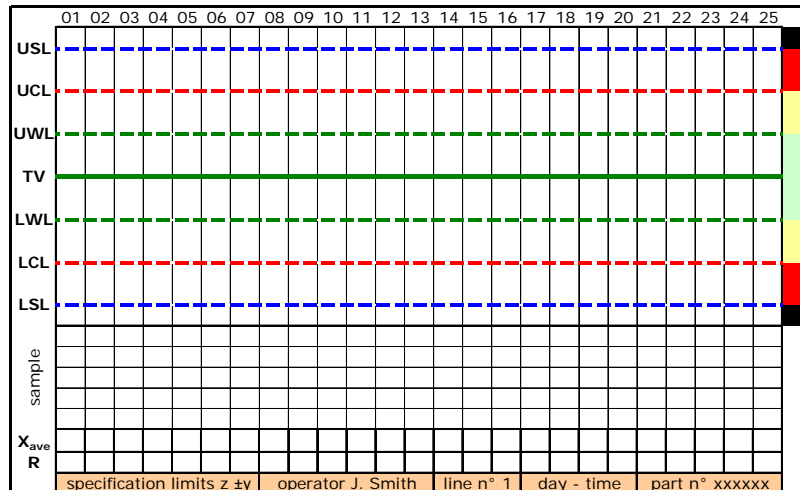
- **defect concentration diagram**, solitamente è un disegno in più viste del prodotto o dell'impianto dove vengono segnati il tipo di difetti riscontrato e la numerosità nel campione esaminato, tuttavia a volte può avere un aspetto tabellare.

line position	01	02	03	04	05	06	07	08	09	10	Totale
defect 1	1						1			3	5
defect 2		1						1		2	2
defect 3			1			1	2		2	6	6
defect 4	2	2		1				1		1	7
defect 5				2				1		3	3
defect 6		1				1			3	5	5
defect 7	1	1					2			1	5
defect 8	3	1				1		1	1	7	7
defect 9				1			1			2	2
defect 10		2				2		1		5	5
defect 11	1							1	2	4	4
defect 12			3		1					1	5
defect 13		1				2		1	1	5	5
defect 14		3					1		1	6	6
defect 15	2				3					5	5
defect 16			2							2	2
defect 17	1	1				4				8	8
defect 18	1						1		2	4	4
defect 19		1		1						2	2
defect 20			1					1		2	2
n° total of occurences	12	10	11	4	5	11	8	8	12	9	

- **input change sheet**, si tratta di una tabella in cui:
 - sulla prima colonna sono riportati i principali parametri d'impostazione dell'impianto o della macchina,
 - sulla seconda colonna sono riportati i valori di set-up,
 - sulla prima riga viene segnato il giorno e l'ora della modifica;
 - sulla seconda riga viene segnato il motivo.

when?		
why?		
input 1		
input 2		
input 3		
input 4		
input 5		
input 6		
input 7		
input 8		

- **control chart**, si tratta di una tabella in cui:
 - sulla prima colonna sono collocati i principali limiti di controllo (esplicitandone i valori);
 - sulla prima riga sono collocati i provini appartenente al campione di controllo,
 - nella parte bassa sono registrati i valori misurati o calcolati e i campi di variabilità,
 - nella parte alta sono riportati i valori medi,
 - nella parte di destra possono essere riassunte le regole di sensibilità delle carte di controllo e cosa fare al presentarsi di una di queste anomalie.



Appare ovvio da questi esempi che esiste la più ampia libertà nel creare documenti di questo tipo, essi risultano utilissimi purché si tenga sempre presente:

- lo scopo per cui si creano;
- il livello di formazione di chi lo deve utilizzare;
- i tempi e le attrezzature necessarie per effettuare le misurazioni e le registrazioni;
- le informazioni che si vogliono ottenere;
- le attività che si chiede di svolgere in conseguenza dei dati registrati.

Come preparare una carta di raccolta dati

La base di qualsiasi controllo statistico risiede nella sintesi dei dati, che devono essere raccolti in maniera chiara, facile e univoca.

Relativamente ad un processo produttivo i dati possono essere suddivisi in:

- **dati per comprendere la situazione attuale**, in pratica i dati vengono raccolti per monitorare una caratteristica speciale di un prodotto;
- **dati per un'analisi specifica**, in pratica i dati vengono raccolti per sostenere o confutare un'ipotesi legata essenzialmente alla presenza di una correlazione causa – effetto;
- **dati per controllare il processo**, in pratica i dati vengono raccolti per monitorare un processo;
- **dati per la regolazione di parametri**, in pratica i dati vengono raccolti per la messa a punto del processo nella fase intermedia tra il set-up² e lo start-up³;
- **dati per accettare o rifiutare lotti**, in pratica i dati vengono raccolti per monitorare le materie prime e i semi-lavorati dei fornitori in fase di accettazione.

Per quanto riguarda la forma dei dati, essi possono essere di tipo:

- **continuo** (ad esempio: peso e spessore),
- **discreto** (ad esempio: numero di pezzi non conformi o i valori basati su attribuzione di punteggi).

Comunque stiano le cose, è basilare che l'azione di raccolta avvenga in modo tale da garantire una corretta analisi dei fatti. Questo si traduce nell'esigenza di effettuare sempre un corretto:

- campionamento,
- utilizzo dei metodi statistici.

Affinché la raccolta dei dati non si trasformi in una semplice archiviazione burocratica è necessario che la raccolta avvenga sotto il dettame di precise regole e con la registrazione di tutti gli elementi che possono poi rendere effettivamente utilizzabili le informazioni contenute nel campione, ad esempio, vanno raccolti anche dati relativi a:

- prodotto,
- data,

² Le lavorazioni vengono impostate secondo una scheda tecnica d'impianto o di macchina che, tuttavia, deve essere di volta in volta riadattata con successive micro-regolazioni.

³ Le lavorazioni hanno sempre una fase in cui si succedono delle micro-regolazioni per ottenere il giusto equilibrio tra tutte le caratteristiche del prodotto.

- operatore,
- strumentazione utilizzata,
- metodologia seguita,
- valore nominale atteso e tolleranza ammissibile.

Tutto questo perché nessuno in seguito dovrebbe effettuare inutili sforzi mentali per ricordarsi l'origine dei dati qualora li si utilizzasse parecchio tempo dopo per un eventuale confronto.

I dati, inoltre, devono essere registrati in modo estremamente chiaro e facile per le successive elaborazioni (ad esempio: database Access, foglio elettronico Excel o su file di testo TXT) in modo di poterli eventualmente importare anche su specifici applicativi (ad esempio: KaleidaGraph, Minitab, Statistica, SQCPack).

Ne deriva che, in accordo con la premessa che i dati devono essere riutilizzati e confrontati tra loro, è opportuno prevedere dei fogli di raccolta opportunamente pre-strutturati.

Tre sono le raccomandazioni da tener presenti:

- la raccolta delle informazioni numeriche deve essere effettuata per poter discutere un problema non sulla base di semplici illazioni, ma di dati certi che evidenziano i fatti osservati o contraddicono eventuali ipotesi alternative;
- l'obiettivo della raccolta delle informazioni (scopo, livello di responsabilità, modalità esecutive) deve essere chiaro a tutte le persone coinvolte;
- ritenere di aver identificato le informazioni da raccogliere non significa automaticamente ottenere il risultato atteso, questo a causa di problematiche che si sono sottostimate o sottovalutate;

tutto ciò implica che sono sempre necessari grande forza di volontà, ingegnosità e grande impiego di metodi scientifici.

Foglio non significa necessariamente carta

Innanzitutto il termine "foglio", dall'avvento dell'informatica, è diventato un vocabolo più virtuale che reale, nel senso che il sistema di raccolta e archiviazione dei dati può prevedere non solo un supporto cartaceo, ma anche l'esistenza di un'interfaccia grafica del tutto virtuale.

Seppur sembri banale, è utile ribadire che il principale scopo dei fogli di raccolta dati è quello di rendere agevole e possibilmente "non invasiva" la raccolta dei dati, cioè l'intervento da parte degli operatori a bordo macchina deve essere ridotto al minimo indispensabile⁴. Inoltre, i fogli di raccolta dati devono essere strutturati per consentire il fine per il quale si è deciso di effettuare un controllo o un'analisi di tipo statistico.

Al fine di semplificare e agevolare la raccolta dei dati è conveniente organizzare i valori di tipo continuo per classi, il numero minimo richiesto è 3, anche se per una corretta analisi sarebbe meglio prevederne almeno 10. Si raccomanda, invece, in direzione opposta di non superare la dozzina di classi.

Il foglio di raccolta dati deve prevedere adeguati campi per la conservazione delle cosiddette condizioni al contorno, cioè di quei parametri presenti durante il processo di rilevazione che potrebbero influenzare i dati o costituire un interessante parametro di confronto (ad esempio: il numero identificativo del prodotto e del lotto, l'operatore presente a bordo macchina, la configurazione impostata nei macchinari, le materie prime utilizzate e i relativi fornitori).

Separando poi il concetto di raccolta dati dal concetto di sintesi degli stessi, il modo più semplice di registrare le informazioni che s'intendono fare oggetto dell'analisi statistica è quello di ordinare in ascissa le classi e in ordinata la frequenza con cui si presenta il valore sotto esame.

Risulta poi utile prevedere la presenza dei valori limite d'accettazione e quello obiettivo per il dato sotto esame.

In altri casi può essere richiesta la registrazione di possibili configurazioni (ad esempio: il motivo per una non conformità). In questo caso, le classi non saranno altro che tutte le possibili configurazioni riunite però in macro-famiglie (ad esempio: colore errato, resistenza al di sotto del valore minimo).

Appare quindi evidente che non esiste e non potrà mai esistere un foglio di raccolta dati standard, ma che si dovranno generare di volta in volta delle interfacce adatte al fine per cui si effettua una campagna di raccolta dati.

Tra gli allegati sono stati inseriti alcuni esempi pratici di fogli di raccolta dati.

⁴ Da un lato la raccolta deve essere efficace (conseguire l'obiettivo) e dall'altro deve essere economica (ridurre al minimo indispensabile: i costi, il tempo, lo sforzo).

Conclusioni

Quando si progetta un foglio di raccolta dati è importante mettere prima su carta ogni informazione che possa poi essere utile in fase di diagnosi. Se il foglio di raccolta dati deve essere utilizzato anche per successive e differenti elaborazioni è importante che lo si realizzi in maniera adeguata agli scopi sono stati prefissati, questo prima di raccogliere materialmente i dati, anche per questo il lancio di un programma di rilevazioni “di prova” può essere utile per schiarirsi le idee in merito e validare l’impostazione di massima di questo strumento di sintesi ed analisi.

Il foglio di raccolta dati non deve necessariamente essere compilato da un’unica per-sona, tuttavia è utile affidare tale incarico in maniera univoca, delegando eventual-mente la responsabilità di fornire informazioni provenienti da aree funzionali diffe-renti ai relativi responsabili.

Una volta validata l’impostazione del foglio di raccolta dati e la metodologia di acquisizione e registrazione delle informazioni, è necessario far diventare il tutto una procedura interna codificata, in modo da assicurare la ripetitività del processo di acquisizione nelle successive campagne di campionamento secondo una metodica condivisa.

È infine utile evidenziare il nome anglosassone di questo strumento: *check sheet*, perché ci ricorda che il simbolo di spunta è quello che dovrebbe risultare dominante nella registrazione del dato grezzo, questo significa che:

- tutte le informazioni legate al contorno devono essere già presenti,
- possibilmente non devono essere segnati i valori (cifre), ma con una spunta la classe d’appartenenza,
- l’esame deve essere fatta direttamente dal responsabile del processo che ne deve ottenere immediatamente delle informazioni a vista.

Il raggruppamento in classi omogenee e i grafici di Pareto



Sotto il nome di analisi di Pareto non c’è la semplice costruzione di un istogramma, ma l’idea di raggruppare, per quanto possibile, i dati in classi omogenee, in modo da poterli confrontare per categorie uniformi.

L’analisi di Pareto si basa quindi sulla costruzione di un particolare istogramma cause-effetti in cui i dati sono organizzati in classi omogenee e raggruppati in ordine (dalle cause più frequenti a quelle meno). Ne deriva un grafico a barre che mostra le cause in funzione della loro rilevanza, evidenziando la cosiddetta **legge universale delle priorità**.

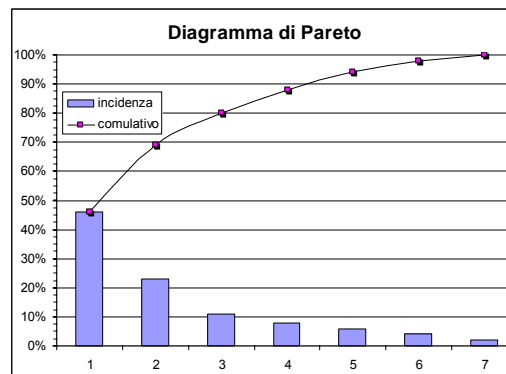
La maggior parte degli effetti è dovuta ad un numero ristretto di cause.

Come preparare un diagramma di Pareto

Per costruire un diagramma di Pareto occorre:

- effettuare un'analisi preliminare di correlazione cause-effetto (magari tramite un diagramma a spina di pesce) in modo da definire le voci che vanno classificate;
- decidere l'arco temporale su cui effettuare le osservazioni e l'eventuale sotto suddivisione sempre in termini temporali al fine di vedere l'effetto di eventuali azioni correttive;
- raccogliere i dati e riassumerli in una tabella di frequenza completando le differenti elaborazioni;
- impostare il grafico evidenziando sull'asse verticale la frequenza di ogni voce di classificazione e ponendo ad intervalli uguali le varie voci di classificazione;
- costruire le barre partendo da quella più alta (maggior frequenza) a sinistra e procedendo quindi verso destra;
- riunire le voci minori in un'unica barra sotto un'unica voce **altre cause**;
- tracciare eventualmente un diagramma a torta evidenziando le frequenze sotto forma percentuale;
- completare il grafico con tutte le informazioni necessarie a comprendere l'oggetto dell'analisi (ad esempio: titolo, arco temporale relativo all'analisi, chi ha effettuato lo studio, condizioni nelle quali si è svolta la campagna di rilevazione, ipotesi, tesi, diagnosi).

Utilizzo dei diagrammi di Pareto



I diagrammi di Pareto sono un utile strumento per:

- ottenere la collaborazione di tutti;
- coinvolgere gli altri;
- scegliere e conseguire obiettivi concreti.

Questo strumento è molto utile per ottenere il coinvolgimento di tutti, perché può sottoporre in maniera evidente e semplice le problematiche aziendali inoltre l'esperienza mostra che è più facile ridurre a metà una barra molto alta piuttosto che ridurre a zero una barra molto bassa.

Viceversa si richiede un'analisi approfondita per comprendere il limite oltre il quale non è più possibile ridurre, in termini di ritorno economico, le cause di non conformità.

I diagrammi di Pareto possono inoltre essere realizzati secondo una differente logica, non quella numerica delle frequenze ma quella dei costi evidenziando le cause che producono in termini più onerosi per l'azienda il medesimo effetto, questo perché difetti riscontrati in maniera molto frequente possono comportare a volte perdite piuttosto modeste in termini di denaro e viceversa.

Un ulteriore passo in avanti può essere fatto utilizzando degli opportuni pesi che introducono il concetto di gravità delle conseguenze.

Conclusioni

Dato che in ogni azienda è necessario conseguire risultati con una quantità spesso limitata di:

- capacità,
- tempo,
- risorse umane,

è importante poter comprendere come e dove concentrare gli sforzi, seguendo sempre la logica che gli obiettivi devono essere remunerativi ovvero economicamente conseguibili.

Questo strumento inoltre è molto valido per verificare i fornitori e le loro materie prime e in generale confrontare input omogenei o analoghi rispetto ai relativi output.

In sintesi lo scopo del diagramma di Pareto è quello di identificare pochi problemi vitali, ma molto triviali. Per cui è utile avere poche barre che coprano all'incirca l'80% dei problemi e che indichino con chiarezza in quale direzione muoversi.

È utile anche costruire un cumulativo per verificare se il diagramma di Pareto funziona bene, tale curva deve essere ripida a sinistra evidenziando come siano le prime categorie a contribuire principalmente ai problemi di non conformità.

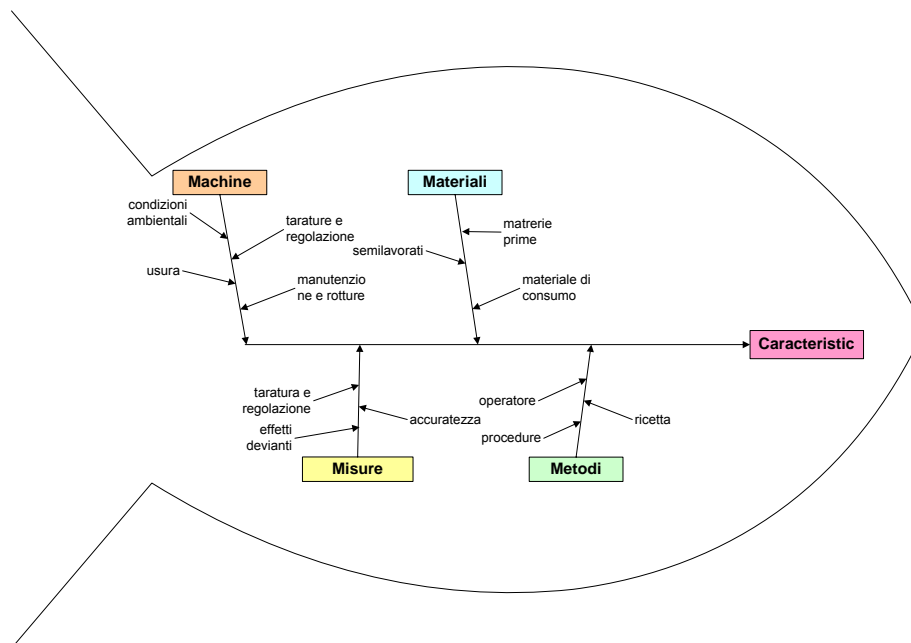
Se la curva cumulativa dovesse rivelarsi poco inclinata all'inizio, questo significherebbe che tutte le barre sono all'incirca equivalenti e che il contributo delle varie cause è pressoché analogo. Questo comporterebbe che nessuna causa sarebbe prevalente e, in questo anomalo caso, il diagramma di Pareto si rivelerebbe poco utile. Tuttavia la causa di ciò potrebbe essere connessa con un'eccessiva suddivisione delle cause, che non permette di evidenziare l'esistenza di una netta supremazia, per cui in questo caso può tornare utile riconsiderare il raggruppamento utilizzato per costruire il diagramma di Pareto, individuando una metodologia che permetta di ottenere aree più grandi e, conseguentemente, più rappresentative.

I diagrammi causa-effetto

Al fine di raggruppare e relazionare tra loro cause ed effetti, così da poter effettuare una successiva analisi, è stato concepito uno strumento tanto semplice quanto efficace: il diagramma a lisca di pesce (*fishbone diagram*). Con riferimento alla forma "classica" definita da Ishikawa, ogni difetto che sia stato individuato dovrebbe essere analizzato per le sue possibili cause e per i suoi indesiderati effetti. Le cause, in un approccio generico, possono essere suddivise in quattro macro-aree:

- **metodi**;
- **materiali**;
- **macchine**, impianti e attrezzature;
- **misurazioni** o metodi d'ispezione;

in modo da sintetizzare al massimo i dati qualitativi o semantici in possesso (tali rappresentazioni per questo motivo vengono anche detti diagrammi a 4M).



I punti di forza di questo strumento d'analisi visiva sono:

- riuscire a focalizzare più persone su un obiettivo comune;
- localizzare rapidamente le reali cause di una non conformità;
- stabilire nuove procedure operative e punti di controllo per monitorare il processo;
- il creare l'abitudine a realizzare una traccia per discutere un problema;
- conservare una traccia dei ragionamenti effettuati realizzando un archivio delle esperienze;

- fornire il supporto primario per la realizzazione dei fogli di raccolta dati.

Come costruire un diagramma a spina di pesce

Per costruire un diagramma cause – effetto occorre:

- individuare il problema che si vuole analizzare;
- individuare le persone da intervistare per ottenere le informazioni necessarie al tracciamento del diagramma organizzando eventualmente un team;
- disegnare il blocco effetto sul lato a destra (si tratta di un riquadro con il nome del problema) e tracciare una riga orizzontale con una freccia terminale che conduce al blocco dell'effetto;
- elencare le possibili cause individuate individuando delle macro-aree di pertinenza (ad esempio le 4M);
- ordinare le cause rispetto alla barra orizzontale da quelle che sembrano avere uno specifico collegamento con l'effetto o un maggior impatto sull'effetto;
- collegare le cause ad ulteriori eventuali sotto-cause andando nel dettaglio ottenendo la costruzione di un diagramma ad albero (evitando di approfondire la cosa oltre il terzo sotto-livello, altrimenti il diagramma perde in chiarezza);
- individuare delle possibili azioni correttive e misurarne l'efficacia.

Appare evidente come sia importante non solo elencare le cause, ma anche poterle classificare e correlare tra loro.

Se in apparenza il diagramma pare costruirsi ponendo in varia successione dei perché, in realtà permette di correlare cause e sotto-cause in uno schema che ne permette la sintesi visiva.

Ora il difetto principale risiede nel fatto che le persone costruiscono diagrammi differenti a seconda del ruolo che rivestono, per questo è utile che il medesimo schema cause – effetto (cioè realizzato per la medesima non conformità) venga realizzato da tutte le persone coinvolte nel team separatamente per poi originare un documento unico di sintesi.

Il diagramma cause – effetto per processi

Il diagramma cause – effetto può complicarsi quando sulla linea orizzontale si vadano a porre i blocchi relativi alle varie fasi del processo produttivo, in modo da poter analizzare in dettaglio dove le varie cause possano entrare in azione, così da monitorare a valle di ogni singola fase la presenza o meno della non conformità.

Questo tipo d'approccio, più complesso, è però quello che permette rapidamente di individuare le principali effettive cause per le non conformità.

Conclusioni

L'utilizzo dei diagrammi cause – effetto è essenzialmente connesso al fatto di imporre una metodologia di raccolta di dati semantici al fine di sintetizzare ed organizzare le informazioni in essi contenute.

Se è innegabile il suo valore educativo è tuttavia utile sottolineare che si tratta di uno strumento che deve servire all'interno della gestione di un problema da parte di un team, poiché ruoli differenti forniscono informazioni (cause) differenti.

Inoltre, i diagrammi cause – effetto sono degli strumenti “vivi” in cui vanno eliminate le cause che non risultino effettivamente collegate all'effetto (semplici illazioni) ed è solo attraverso questa azione che i diagrammi cause – effetto diventano veramente efficaci.

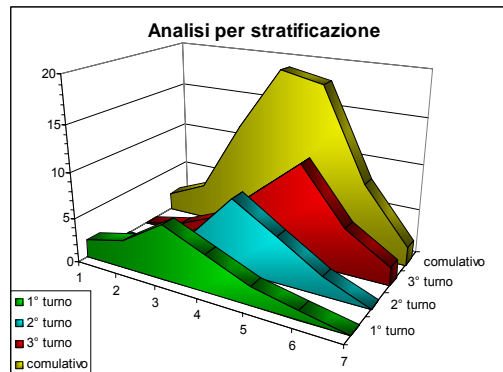
Se è vero che una delle 4M è rappresentata dalla materia prima, non una cosa positiva addossare sempre tutte le colpe ad facile alibi esterno.

L'analisi per stratificazione

Durante la raccolta, i dati sono spesso caratterizzati da fattori differenti (condizioni al contorno) che poi influiscono sul modo che questi hanno di distribuirsi.

Se le classi tendono ad omologare valori simili, classificandoli in gruppo di frequenza, la stratificazione effettua un approccio analogo, raggruppando distribuzioni e dispersioni caratterizzati dal medesimo stato di parametri fondamentali del processo, in modo da confrontare situazioni in cui alcune variabili risultino bloccate ed altre libere.

Lo scopo della stratificazione è quello d'individuare le **vere cause** che influiscono sulla qualità oppure di **pesare adeguatamente le cause tra loro** tramite un'opera di confronto dei valori di distribuzione e di dispersione prima e dopo (cioè relativamente alla popolazione e al campione).



Come preparare un'analisi per stratificazione

Per effettuare un'analisi per stratificazione occorre:

- determinare la caratteristica speciale che s'intende porre sotto analisi;
- prendere in esame la distribuzione e la dispersione che caratterizza la popolazione (solitamente tramite istogramma);
- individuare i possibili fattori di stratificazione (attraverso l'analisi di Pareto);
- classificare i dati in gruppi relativamente ai fattori individuati (attraverso l'introduzione di tale parametro all'interno delle informazioni che caratterizzano il foglio di raccolta dati);
- confrontare le distribuzioni così ottenute tra loro e rispetto alla distribuzione ottenuta per l'intera popolazione.

La stratificazione può avvenire ponendo sotto controllo parametri come:

- tempo (suddividendo la produzione in periodi differenti dell'anno);
- operatori (suddivisi per esperienza, turno, età, etc.);
- macchinari ed attrezzature (suddivisi per tecnologia, tipologia, modello, etc.);
- processo o metodo (individuando differenti condizioni operative o ambientali);
- materiale (fornitore, composizione, conservazione);
- processo di controllo (operatore, attrezzature per i test, metodologie applica).

Conclusione

Al fine di rendere la stratificazione uno strumento d'analisi efficace è necessario:

- cercare molti tipi differenti di cause, di cui il foglio di raccolta dati deve tener conto;
- raccogliere dati, tenendo opportunamente sotto controllo i parametri (o condizioni al contorno);
- filtrare i dati relativamente all'operatore che ha effettuato le operazioni o le misurazioni.

Ora, poiché è spesso difficile rintracciare quali cause contribuiscono maggiormente alla non conformità è opportuno raggruppare le informazioni secondo delle macro-aree al fine di poi realizzare indagini più accurate con cognizione di causa, così da diminuire i tempi e i costi dell'esperimento, ne consegue una stretta connessione con i diagrammi cause – effetto.

I grafici a dispersione e l'analisi di correlazione e indipendenza

Di fronte ad un fenomeno quale, ad esempio, una non conformità è spesso utile poter verificare se esista:

- una qualsivoglia legge tra alcune possibili cause e l'effetto;
- una particolare correlazione tra loro di alcune delle cause;

ed è proprio per questo fine che si utilizza lo strumento dei grafici di dispersione e dell'analisi di correlazione.

È da tenere ben presente che non è sempre possibile creare una legge "universale" (un modello matematico in cui l'insieme di definizione delle variabili indipendenti è rappresentato da tutti i valori possibili per tali variabili), a volte risulta più semplice e più pratico osculare l'andamento per tratti con leggi di tipo empirico.

Come preparare un'analisi di correlazione

Per costruire un diagramma di correlazione è necessario:

- numerare i campioni e i relativi provini;
- raccogliere i dati su colonne;
- tracciare gli assi e riportare i dati da correlare sotto forma di punti (nel piano, nello spazio tridimensionale o in quello iper-geometrico),
- se due o più punti cadono nella medesima posizione allora si disegna un correlato numero di cerchi concentrici di diametro unitario, conservando così l'informazione di molteplicità che riveste quel dato.

Al termine di queste operazioni è facile intuire visivamente l'eventuale tendenza dei punti a distribuirsi o addensarsi secondo forme conosciute (ad esempio: la forma del nugolo di punti, a seconda dei casi, sarà: una retta, una conica, una sinusoidale, etc.) oppure in modo puramente casuale (ad esempio: una nube informe).

Finalità dell'utilizzo dei diagrammi di correlazione

L'esperienza guidata da un opportuno strumento d'analisi può permettere d'individuare la presenza di una o più leggi di correlazione tra un fenomeno e le sue cause. Ciò significa che non è necessario formulare ambiziose teorie e complicati modelli matematici per descrivere un qualsiasi fenomeno, per descrivere la correlazione tra cause ed effetti all'interno di un determinato ambito è sufficiente soddisfare una serie limitata di requisiti:

- aver effettuato un'analisi dimensionale al fine di rendere l'equazione omogenea dal punto di vista delle grandezze in gioco;
- aver individuato una serie di relazioni elementari esprimibili attraverso dei monomi o dei polinomi semplici che rappresentino le equazioni di base che legano il fenomeno ad alcune cause;
- aver effettuato un'analisi per stratificazione in modo da poter comprendere come costruire una legge sotto forma di sequenza d'equazioni di base.

L'analisi di regressione, il metodo dell'analisi dimensionale e la metodologia delle curve oscuratrici diventano quindi gli strumenti per poter ricavare in modo semplice dei modelli matematici in grado di ottenere informazioni previsionali sull'andamento di un fenomeno in output, una volta che ne siano noti i fattori in input.

Più in generale, l'obiettivo è quello di ricavare una funzione continua (seppur a tratti) in grado di descrivere il fenomeno, correlando tra loro le variabili in gioco.

Le principali curve di regressione cartesiane e le formule d'interpolazione

Se si considera l'equazione della retta nel piano essa è definita dalla forma generica:

$$y = a_0 + a_1 \cdot x$$

Si nota la presenza di due costanti:

- $a_0=q$ → intercetta (valore in cui la retta interseca l'asse delle ordinate).
- $a_1=m$ → parametro angolare (valore trigonometrico della tangente avente come angolo α quello formato tra l'asse delle ascisse e la retta);

Per ottenere il valore di queste costanti sono sufficienti due generici punti $P_1=(x_1, y_1)$ e $P_2=(x_2, y_2)$ che devono rendere vero il sistema:

$$\begin{cases} y_1 = a_1 \cdot x_1 + a_0 \\ y_2 = a_1 \cdot x_2 + a_0 \end{cases} \Rightarrow \begin{cases} a_1 = \frac{y_1 - y_2}{x_1 - x_2} \\ a_0 = y_1 - \frac{y_1 - y_2}{x_1 - x_2} \cdot x_1 \end{cases}$$

Diverso è però il caso di dover definire una retta che deve soddisfare al meglio la distribuzione di più punti che non giacciono su una medesima retta. In questo caso si parla di retta interpolante. La caratteristica peculiare di questa retta è di rendere minima la somma delle distanze da tutti i punti che la retta deve interpolare.

Solitamente per distanza di un punto ad una retta si è abituati alla formula:

$$d = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

ma nel caso dell'interpolazione si parla solitamente in termini di **distanza verticale**:

$$d = y(x_0) - y_0$$

Dove $y(x)$ è la funzione della retta interpolatrice.

Se si considerano tutti i punti, la precedente equazione si trasforma in:

$$D = d^2 = \sum_{i=1}^n d_i^2 = \sum_{i=1}^n (a_0 + a_1 \cdot x_i - y_i)^2$$

Come è noto D è minimo quando le derivate parziali rispetto a_0 e a_1 valgono zero, allora:

$$\begin{cases} \frac{\partial D}{\partial a_0} = 2 \cdot \sum_{i=1}^n (a_0 + a_1 \cdot x_i - y_i) = 0 \\ \frac{\partial D}{\partial a_1} = 2 \cdot \sum_{i=1}^n (a_0 + a_1 \cdot x_i - y_i) \cdot x_i = 0 \end{cases}$$

Ne derivano così le equazioni normali della **retta dei minimi quadrati** può essere espressa nella forma:

$$\begin{cases} \sum_{i=1}^n y_i = a_1 \cdot \sum_{i=1}^n x_i + a_0 \cdot n \\ \sum_{i=1}^n (y_i \cdot x_i) = a_1 \cdot \sum_{i=1}^n x_i^2 + a_0 \cdot \sum_{i=1}^n x_i \end{cases} \Rightarrow \begin{cases} a_1 = \frac{n \cdot \sum_{i=1}^n (y_i \cdot x_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \\ a_0 = \frac{\sum_{i=1}^n y_i \cdot \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n (y_i \cdot x_i)}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \end{cases}$$

Posto:

$$\begin{aligned} X &= \sum_{i=1}^n x_i \\ Y &= \sum_{i=1}^n y_i \\ XY &= \sum_{i=1}^n (y_i \cdot x_i) \\ XX &= \sum_{i=1}^n x_i^2 \end{aligned}$$

Si ha il sistema:

$$\begin{cases} a_0 \cdot n + a_1 \cdot X = Y \\ a_0 \cdot X + a_1 \cdot XX = XY \end{cases}$$

La soluzione è facilmente ottenibile utilizzando la regola di Cramer:

$$\begin{vmatrix} n & X \\ X & XX \end{vmatrix} \cdot \begin{vmatrix} a_0 \\ a_1 \end{vmatrix} = \begin{vmatrix} Y \\ XY \end{vmatrix}$$

Supponiamo di voler verificare se esiste una correlazione lineare tra la statura e il peso degli studenti universitari, analizzando un campione e suddividendolo in classi si ottiene uno specchietto di valori x_i e y_i da cui poi ottenere i coefficienti intercetta e parametro angolare della retta interpolatrice.

x_i	y_i	x_i^2	$x_i \cdot y_i$	$(x_i - m_x)^2$	$(y_i - m_y)^2$	$(x_i - m_x) \cdot (y_i - m_y)$		
150	40	22.500	6.000	400	484	-20	-22	440
160	59	25.600	9.440	100	9	-10	-3	30
170	60	28.900	10.200	0	4	0	-2	0
180	71	32.400	12.780	100	81	10	9	90
190	80	36.100	15.200	400	324	20	18	360
850	310	145.500	53.620	1.000	902			920
170	62			200,0	180,4			184
medie				varianze				COV

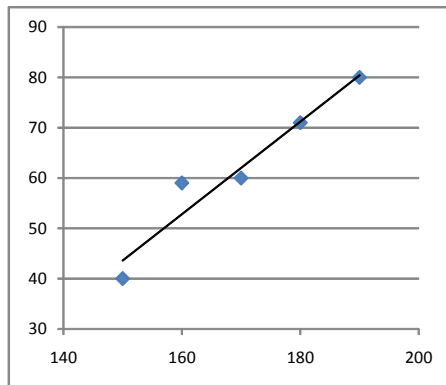
Inserire qui le due equazioni

$$1 > 5 a_0 + 850 a_1 = 310$$

$$2 > 850 a_0 + 145.500 a_1 = 53.620$$

Det della matrice dei coefficienti	
5	850
850	145500

n= 5
 $a_0 = -94,40$
 $a_1 = 0,920$
 $\sigma_x = 14,14$
 $\sigma_y = 13,43$
 $\sigma_{xy} = 184,00$
 $r = 96,87\%$
 $y(178) = 69,36$



$$a_0 = \frac{\begin{vmatrix} 310 & 850 \\ 53620 & 145500 \end{vmatrix}}{\begin{vmatrix} 5 & 850 \\ 850 & 145500 \end{vmatrix}} = \frac{-472000}{5000} = -94,400$$

$$a_1 = \frac{\begin{vmatrix} 5 & 310 \\ 850 & 53620 \end{vmatrix}}{\begin{vmatrix} 5 & 850 \\ 850 & 145500 \end{vmatrix}} = \frac{4600}{5000} = 0,920$$

Nell'esempio si estrapola il peso atteso per uno studente universitario alto 178 cm (69,36 kg).
 Se si considera l'equazione della parabola nel piano essa è definita dalla forma generica:

$$y = a_0 + a_1 \cdot x + a_2 \cdot x^2$$

In questo caso si itera il precedente modello ponendo 3 condizioni:

$$\begin{cases} \sum_{i=1}^n y_i = a_2 \cdot \sum_{i=1}^n x_i^2 + a_1 \cdot \sum_{i=1}^n x_i + a_0 \cdot n \\ \sum_{i=1}^n (y_i \cdot x_i) = a_2 \cdot \sum_{i=1}^n x_i^3 + a_1 \cdot \sum_{i=1}^n x_i^2 + a_0 \cdot \sum_{i=1}^n x_i \\ \sum_{i=1}^n (y_i \cdot x_i^2) = a_2 \cdot \sum_{i=1}^n x_i^4 + a_1 \cdot \sum_{i=1}^n x_i^3 + a_0 \cdot \sum_{i=1}^n x_i^2 \end{cases}$$

La medesima regola vale se si considera una legge cubica del tipo:

$$y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$

Quindi questo approccio è valido e generalizzabile per tutte le curve generiche del tipo:

$$y = \sum_{i=0}^n a_i \cdot x^i$$

Consideriamo ora due funzioni generiche $g(x_i)$ e $f(x_i)$ che potrebbero essere, ad esempio, due funzioni trigonometriche:

$$y(x_i) = a_0 \cdot g(x_i) + a_1 \cdot f(x_i)$$

da cui:

$$D = d^2 = \sum_{i=1}^n d_i^2 = \sum_{i=1}^n [a_0 \cdot g(x_i) + a_1 \cdot f(x_i) - y_i]^2$$

Come è noto D è minimo quando le derivate parziali rispetto a_0 e a_1 valgono zero, allora:

$$\begin{cases} \frac{\partial D}{\partial a_0} = 2 \cdot \sum_{i=1}^n [a_0 \cdot g(x_i) + a_1 \cdot g(x_i) \cdot f(x_i) - g(x_i) \cdot y_i] = 0 \\ \frac{\partial D}{\partial a_1} = 2 \cdot \sum_{i=1}^n [a_0 \cdot g(x_i) \cdot f(x_i) + a_1 \cdot f(x_i) - f(x_i) \cdot y_i] = 0 \end{cases}$$

da cui:

$$\begin{cases} a_0 \cdot \sum_{i=1}^n g(x_i) + a_1 \cdot \sum_{i=1}^n g(x_i) \cdot f(x_i) = \sum_{i=1}^n g(x_i) \cdot y_i \\ a_0 \cdot \sum_{i=1}^n g(x_i) \cdot f(x_i) + a_1 \cdot \sum_{i=1}^n f(x_i) = \sum_{i=1}^n f(x_i) \cdot y_i \end{cases}$$

Consideriamo ora delle curve generiche del tipo:

$$y = \exp\left(\sum_{i=0}^n a_i \cdot x^i\right) = e^{\sum_{i=0}^n a_i \cdot x^i}$$

e

$$y = \ln\left(\sum_{i=0}^n a_i \cdot x^i\right)$$

Si nota che entrambe le forme sono facilmente riportabili a quelle polinomiali semplici già esaminate se le si trasformano in:

$$z = \ln(y) = \sum_{i=0}^n a_i \cdot x^i$$

e

$$z = e^y = \sum_{i=0}^n a_i \cdot x^i$$

Quindi la soluzione basata su un sistema di equazioni lineari è estendibile anche a queste forme d'interpolazione.

Sviluppo in serie di Fourier

Particolare rilievo ove si sia di fronte ad un processo ciclico è l'interpolazione mediante sviluppo in serie di Fourier, ovvero attraverso la composizione di funzioni trigonometriche.

$$y = a_0 + \sum_{j=0}^n [a_j \cdot \cos(x_j) + b_j \cdot \sin(x_j)]$$

in cui gli elementi sono così definiti:

$$a_0 = \frac{1}{2} \cdot \frac{\sum_{i=1}^n y_i}{n}$$
$$a_j = 2 \cdot \sum_{i=1}^n [y_i \cdot \cos(j \cdot x_i)]$$
$$b_j = 2 \cdot \sum_{i=1}^n [y_i \cdot \sin(j \cdot x_i)]$$

Utilizzando l'identità di Eulero:

$$z_j = \frac{2}{n} \cdot \sum_{k=0}^n [y_k \cdot \exp(i \cdot j \cdot x_k)] \Leftrightarrow i = \sqrt{-1}$$

La migliore interpolante

Per determinare quale possa essere la miglior funzione interpolante esiste il coefficiente di determinazione. Determinati i seguenti valori:

- **deviazione totale:**

$$dev.tot. = \sum_{i=1}^n [y_i - m_y]^2$$

- **deviazione spiegata** o di regressione:

$$dev.spie. = \sum_{i=1}^n [y(x_i) - m_y]^2$$

- **deviazione residua:**

$$dev.res. = \sum_{i=1}^n [y(x_i) - y_i]^2$$

dove:

$$dev.tot. = dev.res. + dev.spie.$$

Si definisce **indice di determinazione** il seguente rapporto:

$$r^2 = \frac{dev.spie.}{dev.tot.}$$

Tale coefficiente può variare tra 0 e 1 e più il suo valore s'avvicina all'unità e più è buona la curva d'interpolazione utilizzata.

Esiste un secondo coefficiente, detto **indice di correlazione lineare** che è definito come segue:

$$r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$

Attenzione, la correlazione lineare non è l'unica possibile, si pensi ad esempio ai fenomeni ciclici che oscillano lungo un valore centrale e che, quindi, fornirebbero $r=0$.

Il metodo dell'analisi dimensionale

L'analisi dimensionale è uno strumento concettuale che riutilizza per comprendere le grandezze fisiche coinvolte in un fenomeno.

Questa metodologia è utilizzata per:

- **verificare la plausibilità** delle equazioni;
- **formulare ipotesi ragionevoli** su dei fenomeni complessi;
- **progettare degli esperimenti** al fine di sviluppare delle teorie descrittive di un fenomeno.

Se si considera la seguente simbologia:

- **[M]** → massa;
- **[L]** → lunghezza;
- **[T]** → tempo;

si può facilmente comprendere come tutte le unità di misura siano riconducibili ad un numero limitato di unità di misura fondamentali di riferimento, cosicché ogni grandezza potrebbe essere espressa nella forma di un monomio:

$$[G] = \prod_{i=1}^n [A]^{x_i}$$

cioè di prodotto tra unità di misura fondamentali. Ne consegue che due monomi (grandezze) possono essere sommati solo se sono omogenei. Queste premesse permettono di utilizzare il metodo dell'analisi dimensionale come valido strumento di verifica per giudicare la correttezza formale di una relazione tra più fattori.

Ne consegue che gli argomenti di funzioni:

- esponenziali;
- logaritmiche;
- trigonometriche;

devono necessariamente essere valori adimensionali.

Progredendo ulteriormente in avanti, i fenomeni che avvengono nello spazio 3D (a tre dimensioni) richiedono un approccio vettoriale allo spazio geometrico, ampliando il concetto di lunghezza in quello di componenti:

$$\vec{L} = (L_x, L_y, L_z)$$

In conclusione si perviene al fatto che le equazioni fisiche devono essere indipendenti dalle unità di misura impiegate per misurare le variabili.

Si perviene quindi all'enunciato del **teorema π** (o di Buckingham): scelta una funzione $y(x_i)=0$ descritta da n variabili, il fenomeno così osservabile può essere studiato mediante una funzione $f(\pi_j)=0$ dove il numero m di parametri adimensionali π_j è pari a $n-q$, dove q è il numero delle unità fondamentali che descrivono il fenomeno, questo indipendentemente dalla forma matematica assunta dalla funzione.

Ciò consente di arrivare ad una formulazione matematica del fenomeno per via empirica, verificando le grandezze in gioco e ricavando i parametri adimensionali per via sperimentali.

Consideriamo le seguenti variabili di un fenomeno della fluidodinamica:

- forza: $F=[M][L][T]^{-2}$
- densità: $\rho=[M][L]^{-3}$
- velocità: $v=[L][T]^{-1}$
- lunghezza caratteristica: $l=[L]$
- velocità del suono: $a=[L][T]^{-1}$
- coefficiente di viscosità dinamica: $\mu=[M][L]^{-1}[T]^{-1}$

si hanno sei variabili e tre grandezze fondamentali, di conseguenza ci saranno tre prodotti Π . Scegliendo quale nuova terna di riferimento ρ, v, l , i prodotti Π sono:

$$\Pi_1 = f_1(F, \rho, v, l)$$

$$\Pi_2 = f_2(\mu, \rho, v, l)$$

$$\Pi_3 = f_3(a, \rho, v, l)$$

Consideriamo il seguente prodotto:

$$\Pi_1 = F \cdot \rho^a \cdot v^b \cdot l^c$$

Il teorema π assicura che i prodotti su riportati possono essere adimensionalizzati.

$$\frac{([M][L][T]^{-2})^1 ([M][L]^{-3})^a ([L][T]^{-1})^b ([L])^c}{[M]^{1+a} [L]^{1-3a+b+c} [T]^{-2-b}}$$

ora, gli esponenti devono risultare tutti nulli, per cui:

$$\begin{cases} 1 + a = 0 \\ 1 + 3a + b + c = 0 \\ -2 - b = 0 \end{cases}$$

Utilizzando lo schema per tutti e tre i prodotti si ottiene:

$$\begin{aligned} \Pi_1 &= \frac{F}{\rho \cdot v \cdot l} \\ \Pi_2 &= \frac{\rho \cdot v \cdot l}{\pi} \\ \Pi_3 &= \frac{v}{a} \end{aligned}$$

Che sono rispettivamente:

- il **coefficiente di portanza** (o di resistenza);
- il **numero di Reynolds**;
- il **numero di Mach**.

Questi parametri possono essere convenientemente utilizzati per studiare il moto dei fluidi e, più in generale, i fenomeni della fluidodinamica.

L'interpolazione spline

L'interpolazione spline è un particolare metodo d'interpolazione che, a differenza delle formule di regressione, che utilizza un concetto di "distanza minima", si serve nei successivi intervalli di una sequenza di polinomi di grado piccolo, scegliendoli in modo che due polinomi successivi si saldino in **modo liscio**, cioè osservando la continuità di qualche derivata.

Una tipica **equazione spline** è:

$$y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$

Con tale equazione, ad esempio, è possibile partire dal punto

$$y(0) = 0 \Rightarrow a_0 = 0$$

con una tangente orizzontale, cioè:

$$y'(0) = 0 \Rightarrow a_1 = 0$$

Se si considerano tutti i punti, la precedente equazione si trasforma in:

$$D = d^2 = \sum_{i=1}^n d_i^2 = \sum_{i=1}^n (a_2 \cdot x_2^2 + a_3 \cdot x_3^3 - y_i)^2$$

D è minimo quando le derivate parziali rispetto a_2 e a_3 valgono zero, allora:

$$\begin{cases} \frac{\partial D}{\partial a_2} = 2 \cdot \sum_{i=1}^n x_2^2 \cdot (a_2 \cdot x_2^2 + a_3 \cdot x_3^3 - y_i) = 0 \\ \frac{\partial D}{\partial a_3} = 2 \cdot \sum_{i=1}^n x_3^3 \cdot (a_2 \cdot x_2^2 + a_3 \cdot x_3^3 - y_i) = 0 \end{cases}$$

L'ultimo punto x_n viene così a definire tre condizioni per la successiva equazione spline:

- $y(x_n)$, valore che la funzione ha nell'ultimo punto;
- $y'(x_n)$, valore che la tangente alla funzione ha nell'ultimo punto;
- $y''(x_n)$, valore che del raggio di curvatura che la funzione ha nell'ultimo punto.

Più in generale, il concetto è quello di una generica curva nel piano cartesiano con il concetto:

$$\delta_i = y(x_i) - y_i$$

In pratica si considerano le equazioni interpolatrici nella loro forma omogenea:

$$y(x) - y = 0 \Rightarrow \sum [y(x_i) - y_i]^2 = D$$

Ne deriva che anche un arco di conica può essere utilizzata come equazione spline di raccordo, essendo:

$$a_{1,1} \cdot x^2 + a_{2,2} \cdot y^2 + 2 \cdot a_{1,2} \cdot x \cdot y + 2 \cdot a_{1,3} \cdot x + 2 \cdot a_{2,3} \cdot y + a_{3,3} = 0$$

Si possono assegnare le tre condizioni ai due estremi per ottenere il valore delle costanti.

Come si è visto, sviluppando il quadrato del binomio $y(x_i) - x_i$ si ottengono una serie di somme di sommatorie che possono essere facilmente tabellate mediante foglio elettronico:

Y	X	X ²	...
y_1	x_1	x_1^2	...
y_2	x_2	x_2^2	...
...
y_n	x_n	x_n^2	...
$\sum y_i$	$\sum x_i$	$\sum x_i^2$...

I valori così ottenuti altro non sono che le costanti $b_{j,k}$ di una serie di m equazioni lineari a m variabili a_j facilmente risolvibili tramite un foglio elettronico sfruttando il calcolo matriciale e la **matrice completa associata** [B].

$$[B] = [B1][B2] = \begin{bmatrix} b_{1,1} & b_{2,1} & \cdots & b_{m,1} \\ b_{1,2} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b_{1,m} & \cdots & \cdots & b_{m,m} \end{bmatrix} \begin{bmatrix} b_{m+1,1} \\ \cdots \\ \cdots \\ b_{m+1,1} \end{bmatrix}$$

Dove l'ultima colonna rappresenta il vettore [B2] dei termini noti delle equazioni lineari di primo grado a più variabili.

$$[B1] \cdot [A] = -[B2] \Rightarrow [A] = -[B2] \cdot \text{inv}[B1]$$

L'unica condizione che deve essere rispettata è che:

$$\det[B1] \neq 0$$

la quale assicura l'invertibilità della matrice quadrata [B1].

Appare semplice la determinazione delle costanti utilizzandole come variabili di un sistema di equazioni lineari all'interno di un foglio di calcolo.

In pratica per estrapolare delle formule da una rilevazione di dati, si tende a realizzare correlazioni lineari a tratti (rami) per poi congiungere questi tratti con degli archi (gomiti) mantenendo la continuità della funzione e delle sue derivate prima (tangente) e seconda (curvatura).

Conclusioni

Al fine di non sprecare energie mentali utili altrove, è utile realizzare la ricerca dell'esistenza di una correlazione dei dati solo se l'analisi per stratificazione ha permesso di rilevare distribuzioni e dispersioni analoghe, inoltre la stratificazione può essere utile per verificare se la dispersione dei dati avvenga a sua volta in maniera stratificata.

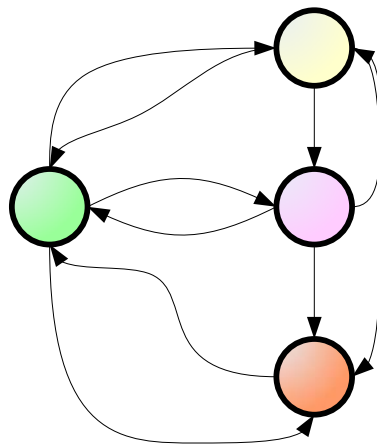
Come descritto, l'esistenza di una correlazione può avvenire in maniera univoca solo in un determinato intervallo differenziandosi così in varie leggi per i diversi intervalli. Tutto ciò si sintetizza nell'affermazione che è sempre possibile forzare empiricamente l'individuazione di una correlazione tra i principali parametri o misurazioni di un processo, ma che per arrivare a ciò è spesso necessario l'impiego di sofisticate tecniche di analisi, con il rischio per gli inesperti non solo di non trarre più alcuna conclusione, ma di trarre conclusioni errate.

L'analisi dei processi tramite diagrammi di flusso

Capita spesso (troppo spesso) che se si chiede a qualcuno di tracciare i flussi delle attività e delle informazioni inerenti ad un processo di cui lui è il responsabile, quello che s'ottiene è una specie d'intricata ragnatela.



Se chi coordina le attività si dimostra confuso su come queste si svolgano, è molto probabile che non ci si trova di fronte all'**uomo ragno** (supereroe “buono”), ma all'**uomo rogna** (supereroe “cattivo”), nel senso che è verosimile presumere che il processo che questi gestisce sia caratterizzato da mole criticità (rogne, per l'appunto).



Quasi sicuramente, chiunque abbia provato a chiedere di mettere su carta i flussi di attività e di informazioni inerenti alla gestione dell'**ordine** ha poi ricevuto qualcosa che dimostrava un certo **disordine**, con una serie di frecce proiettate in tutte le direzioni che al confronto, quelle dei persiani alle Termopili erano una miseria⁵.

Come preparare i diagrammi di flusso

Per preparare in maniera coerente dei diagrammi di flusso o flussogrammi è opportuno innanzitutto comprendere il significato della simbologia a blocchi.

Ogni blocco rappresenta un'istruzione, il fatto che la forma ne strutturi il significato permette una maggiore leggibilità delle attività svolte in un processo.

Ogni processo è caratterizzato da un blocco d'inizio dal quale obbligatoriamente si deve partire:



e da uno di fine a cui obbligatoriamente si deve giungere:

⁵ È celebre un piccolo aneddoto: alcuni disertori dell'esercito persiano (per lo più Greci arruolati con la forza) avevano dichiarato che i Persiani erano così tanti da oscurare il sole con le loro frecce; gli spartani risposero «*Bene, allora combatteremo nell'ombra*».



Ogni processo è svolto da attività elementari che rielaborano quanto giunge in ingresso:



In ogni processo ci sono dei momenti in cui la scelta dipende dall'esito sì/no di un controllo:



Appare evidente che non possono allontanarsi da questo blocco più frecce con il medesimo esito positivo/negativo. Una decisione è sempre univoca se non lo è si è trascurato l'esistenza di un predicato (cioè manca la domanda a cui si deve dare risposta). Infine, in un processo si esegue una scrittura/lettura di informazioni prima di svolgere le attività:



Nonostante i blocchi codificati siano molti di più, questi cinque possono bastare e avanzare per descrivere dei processi elementari. Molti processi danno luogo a un "semilavorato". In pratica questo qualcosa per essere completato deve ricorrere ad altri processi, come accade ad esempio quando s'invisano delle richieste ad un'altra funzione per ottenerne risposte. Tutto ciò può essere semplificato come il ricorso ad un processo esterno:



in questo modo, all'interno di questo semplice blocco, si può celare un altro flussogramma come fosse un allegato o un rimando. Infine si può introdurre un settimo simbolo, il nodo:



che si utilizza per far confluire più flussi affluenti in un unico emissario qualora non vi sia alcuna attività.

Conclusioni

Quanto proposto è la forma più elementare di rappresentazione di un processo. Esistono poi forme molto più sofisticate e connesse anche alle cosiddette "mappe mentali", ma non è obiettivo di queste dispense addentrarsi in argomenti così ampi. Lo scopo principale di questi elementari diagrammi è, infatti, quello di rendere leggibile in maniera semplice cosa accade in azienda, evidenziando le cosiddette "zone grigie", quegli ambiti caratterizzati ad grande fragilità che poi ciclicamente innescano le "rogne".

Le aziende sono generalmente strutturate per funzioni separate ("verticali") mentre i flussi di lavoro sono interfunzionali ("orizzontali"); questo genera vuoti o sovrapposizioni di responsabilità che peggiorano l'efficienza e l'efficacia complessiva. Un'analisi "per processi" permette di avere una corretta visione delle modalità di funzionamento dell'azienda.

È poi utile suddividere i processi in:

- **processi chiave:** che aggiungono valore al cliente;
- **processi di supporto:** che assicurano la coerenza dei processi chiave.

Utilizzare un approccio per processi permette quindi di:

- una chiara visione delle attività da svolgere per trasformare un input in output;
- un miglior controllo sui prodotti, i servizi e i controlli;
- una migliore gestione delle interrelazioni funzionali, che permette di:
 - evidenziare errori sistemici,
 - allineare velocemente le persone agli obiettivi.

