

Introduzione alla teoria dei piani sperimentali

Obiettivo:

Studiare l'effetto di fattori controllabili che influenzano la realizzazione di processi di varia natura (industriali, agricoli, ecc)

Modello : $Y = f(A, B, C, \dots) + E$

tipo di funzione f :

dal punto di vista statistico: *qualitativo o quantitativo*

dal punto di vista operativo: *controllabile, non controllabile*

La funzione è incognita: occorre formulare delle ipotesi sulla forma funzionale:

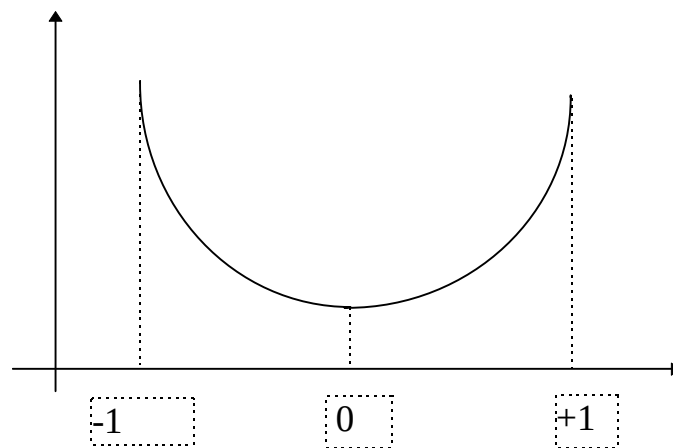
es.: di tipo lineare, polinomiale, ecc.

Ogni fattore qualitativo presenta un numero finito di modalità
(livelli fattoriali)

I fattori quantitativi possono assumere un numero infinito di valori: occorre “discretizzarli”
cioè scegliere dei livelli per ciascun fattore.

Per un modello espresso in forma lineare, è sufficiente individuare due livelli definiti come *minimo* e *massimo*.

Per una forma non lineare occorrono tre, o più, livelli



Obiettivo dell'esperimento:

- Determinare le variabili influenti sulla risposta del sistema.
- Stimare il modello
- Determinare il valore di un livello che fornisce un valore predefinito della risposta

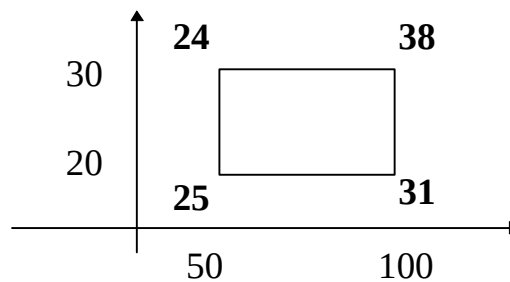
Esempio N°1

Consideriamo l'esperimento consistente nella misura del rendimento di un agente chimico sottoposto a diversi fattori ambientali (temperatura) e fisici (concentrazione)

Per la temperatura consideriamo un livello minimo di funzionamento del sistema di 50° ed un livello max di 100°

La concentrazione viene considerata ad un livello minimo di 20g/l ed un livello max di 30 g/l

Si effettuano quattro prove sperimentali, considerando le possibili combinazioni dei livelli



Matrice del piano sperimentale

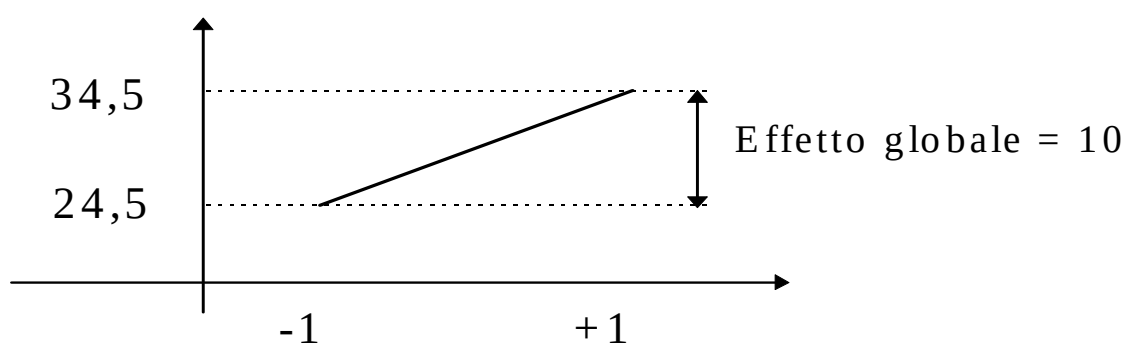
Prove	Livelli dei fattori		Risposta
	Temperatur	Concentrazion	
	a	e	
1	-1	-1	25
2	+1	-1	31
3	-1	+1	24
4	+1	+1	38

Media generale della risposta $Y = 29,5$

Effetto della *temperatura* sulla risposta:

$$(25 + 24)/2 = 24,5$$

$$(31 + 38)/2 = 34,5$$



$$\text{Effetto medio} = 10/2 = 5$$

L'effetto medio si calcola come prodotto tra la variabile temperatura e il vettore della risposta, diviso il numero di prove.

Per il fattore *Concentrazione*:

$$\text{Effetto globale} = [(24+38)/2] - [(25 +31)/2] = 31 - 28 = 3$$

$$\text{Effetto medio} = \frac{1}{4} \langle c, y \rangle = 1,5$$

Ipotesi sulla forma funzionale:

Modello lineare additivo

$$Y = \textit{media} + \textit{temp} + \textit{conc} + \textit{err}$$

$$\hat{Y} = 29,5 + 5[-1 +1] + 1,5 [-1 +1]$$

Con semplici calcoli

$$\hat{Y} = [23 \ 33 \ 26 \ 36]$$

i residui sono pari a $Y - \hat{Y}$:

$$\text{err} = [+2 \ -2 \ -2 \ +2]$$

$$\begin{bmatrix} 25 \\ 31 \\ 24 \\ 38 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & +1 & -1 \\ 1 & -1 & +1 \\ 1 & +1 & +1 \end{bmatrix} \cdot \begin{bmatrix} 29,5 \\ +5 \\ +1,5 \end{bmatrix} + \begin{bmatrix} +2 \\ -2 \\ -2 \\ +2 \end{bmatrix}$$

Si ritrova il modello lineare

$$Y = XB + E$$

La matrice del piano è ortogonale

I coefficienti sono le stime ottenute col metodo dei minimi quadrati

$$B = (X'X)^{-1}X'y$$

$$X'X = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$B = \frac{1}{4}X'y$$

La varianza delle stime è

$$\sigma^2 = \text{Var}(E) = \frac{16}{n-2} = 8$$

Ritornando ai valori reali del processo:

$$\hat{Y} = 29,5 + 5 \left(\frac{\text{Temp} - 75}{100 - 50} \right) + 1,5 \left(\frac{\text{Conc} - 25}{30 - 20} \right)$$

$$\hat{Y} = 29,5 + \frac{5}{25} \text{Temp} - \frac{5 \cdot 75}{25} + \frac{1,5}{5} \text{Conc} - \frac{1,5 \cdot 25}{5}$$

$$\hat{Y} = 29,5 + 0,2\text{Temp} - 15 + 0,3\text{Conc} - 7,5$$

$$\hat{Y} = 7 + 0,2\text{Temp} + 0,3\text{Conc}$$

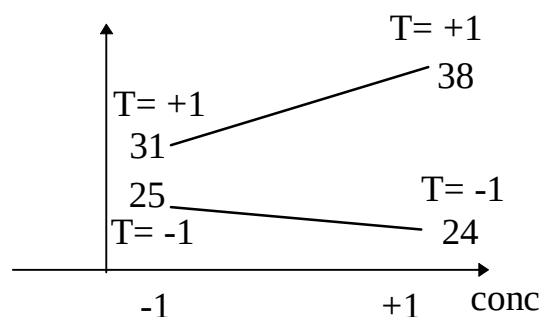
Il modello precedente non è completo poiché non si è considerato

l'effetto dell'interazione tra i due fattori

Temperatura e Concentrazione

L'effetto della temperatura dipende dal valore assunto dalla
concentrazione

Verifica grafica



Per C = -1 l'effetto di T è +6; effetto medio = 3

Per C = +1 l'effetto di T è +14; effetto medio = 7

In assenza di interazione, l'effetto di T doveva essere costante:
verifica grafica: le rette devono essere parallele

Calcolo dell'interazione:

$$\frac{7 - 3}{2} = 2 = \frac{1}{2} \left(\frac{38 - 24}{2} - \frac{31 - 25}{2} \right) = \frac{1}{4} (25 - 31 - 24 + 38)$$

semi differenza degli effetti medi d'interazione.

In termini matriciali si effettua il prodotto scalare tra i vettori
colonna relativi ai fattori temperatura e concentrazione nella
matrice del piano sperimentale:

$$\begin{pmatrix} -1 \\ +1 \\ -1 \\ +1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ +1 \\ +1 \end{pmatrix} = \begin{pmatrix} +1 \\ -1 \\ -1 \\ +1 \end{pmatrix}$$

Matrice completa degli effetti per il piano 2^2

$$X = \begin{bmatrix} I & T & C & TC \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

La matrice X è ancora una matrice ortogonale

Questa proprietà permette di stimare l'interazione senza influenzare la stima degli effetti principali

Nel modello precedente l'effetto interazione influiva sui residui del modello

Nel modello completo, le risposte osservate vengono ricostruite esattamente dal modello.

Il modello si dice saturo: il numero di parametri da stimare è uguale al numero di prove sperimentali.

Nota: Non è possibile stimare la $\text{var}(\text{err})$ con il numero di esperimenti effettuati.

Il modello completo è

$$\hat{Y} = 29,5 + 5 \frac{T - 75}{25} + 1,5 \frac{C - 25}{5} + 2 \left(\frac{T - 75}{25} \right) \cdot \left(\frac{C - 25}{5} \right)$$

In assenza di residui non è possibile effettuare test di significatività sull'effetto interazione.

Occorre ripetere o duplicare le prove sperimentali

I piani completi permettono di stimare tutte le interazioni tra i fattori.

È possibile ridurre il numero di esperimenti se non si ha interesse agli effetti d'interazione

Problemi:

Le stime sono significative ?

La forma funzionale del modello è corretta ?