La progettazione degli esperimenti

Progettazione degli esperimenti

Quando si progetta un esperimento si considerano i seguenti parametri:

RISPOSTA

E' il "risultato", il parametro che si vuole riuscire a controllare/predire:
ad es. la resa di una reazione, la proprietà di un materiale, la riduzione di un prodotto secondario non voluto, la minimizzazione di un difetto del materiale, etc...

VARIABILI (più di una)

Sono i parametri che possono essere monitorati/variati durante il processo per ottenere la risposta.

Scopo della progettazione: individuare quali variabili (e loro eventuali interazioni) influenzano significativamente la risposta per **ottimizzare** il risultato che si vuole ottenere

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

1) **Progettazione degli esperimenti**

(*Design of Experiments o Experimental Design)*

Di estrema importanza, da applicare ove possibile

Impiega:

- ANOVA
- F-test
- t-test
- Diagrammi
- Superfici di risposta

RECUPERARE E LEGGERE
Riccardo Leardi
«Experimental design in chemistry: A tutorial»
Analytica Chimica Acta Volume 652, Issues 1-2, 2009, Pages 161-172

Progettazione degli esperimenti

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_Kx_K + b_{11}x_1^2 + b_{22}x_2^2 + \ldots + b_{KK}x_K^2 + b_{12}x_1x_2 + \ldots + \varepsilon$$

I Fattori x_1, x_2,...x_K possono essere modificati sistematicamente

La Risposta y è misurata e modellata

Perché progettare gli esperimenti

- *Screening* (per capire quali sono le variabili importanti nel determinare il valore di una risposta)
- *Saving time* (per risparmiare tempo)
- *Quantitative modelling* (per costruire un modello quantitativo dell'esperimento)
- *Optimisation* (per massimizzare rese di reazione, ottimizzare tempi, consumo di reagenti …)

Ottimizzazione sperimentale: approccio classico

L'approccio classico al disegno sperimentale è un approccio univariato:

prevede di valutare gli effetti di ogni variabile sulla risposta tenendo ferme ad un determinato valore costante le altre.

Questo approccio comporta una serie di **svantaggi**:

Grande dispendio di tempi e risorse per effettuare tanti esperimenti

Stima dell'effetto di una variabile condizionato dal valore (livello) a cui vengono bloccate le altre variabili

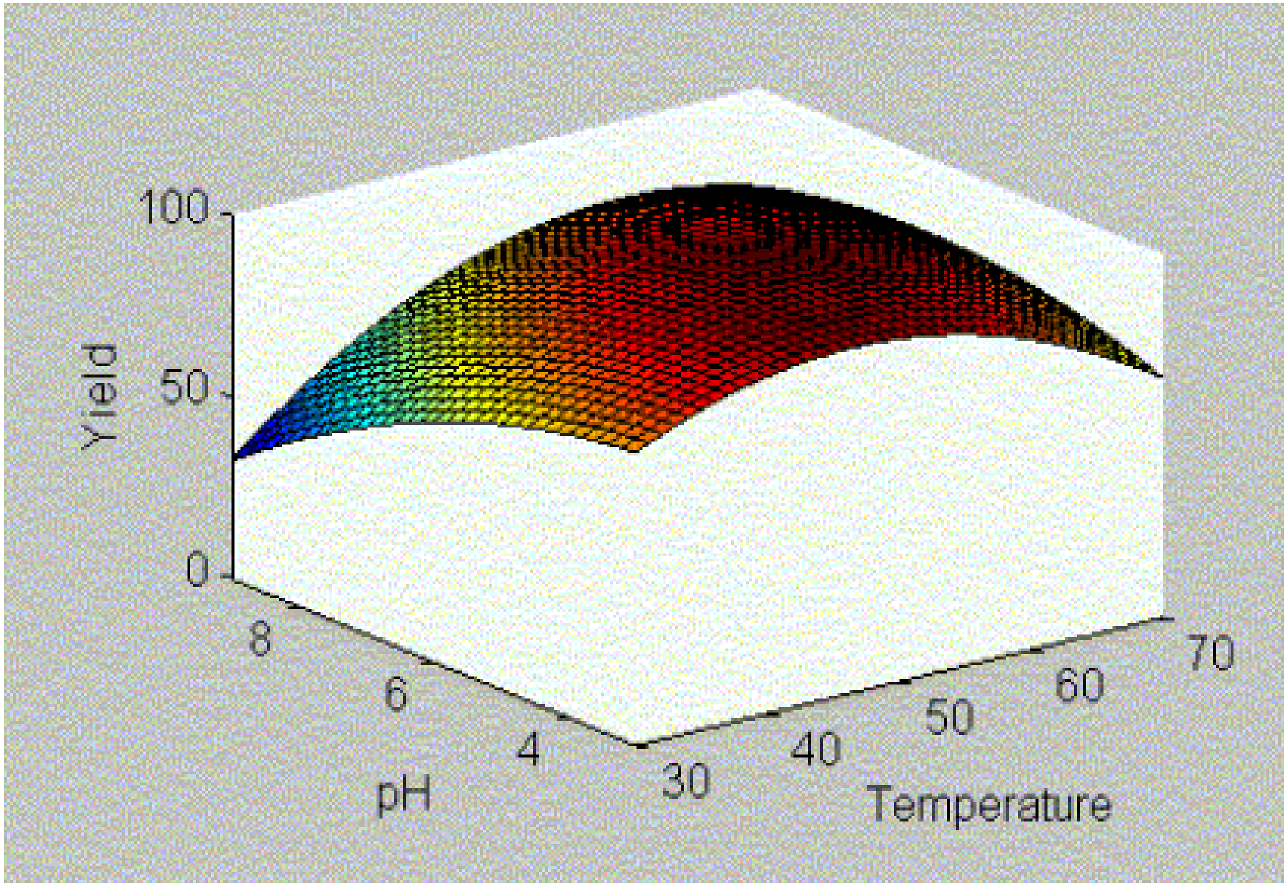
Impossibilità di rilevare interazioni tra variabili

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Perché progettare gli esperimenti?

Un problema : Ottimizzazione di una resa di reazione con pH e temperatura.

Possiamo trovare la combinazione di pH e temperatura che producono la resa migliore della reazione?

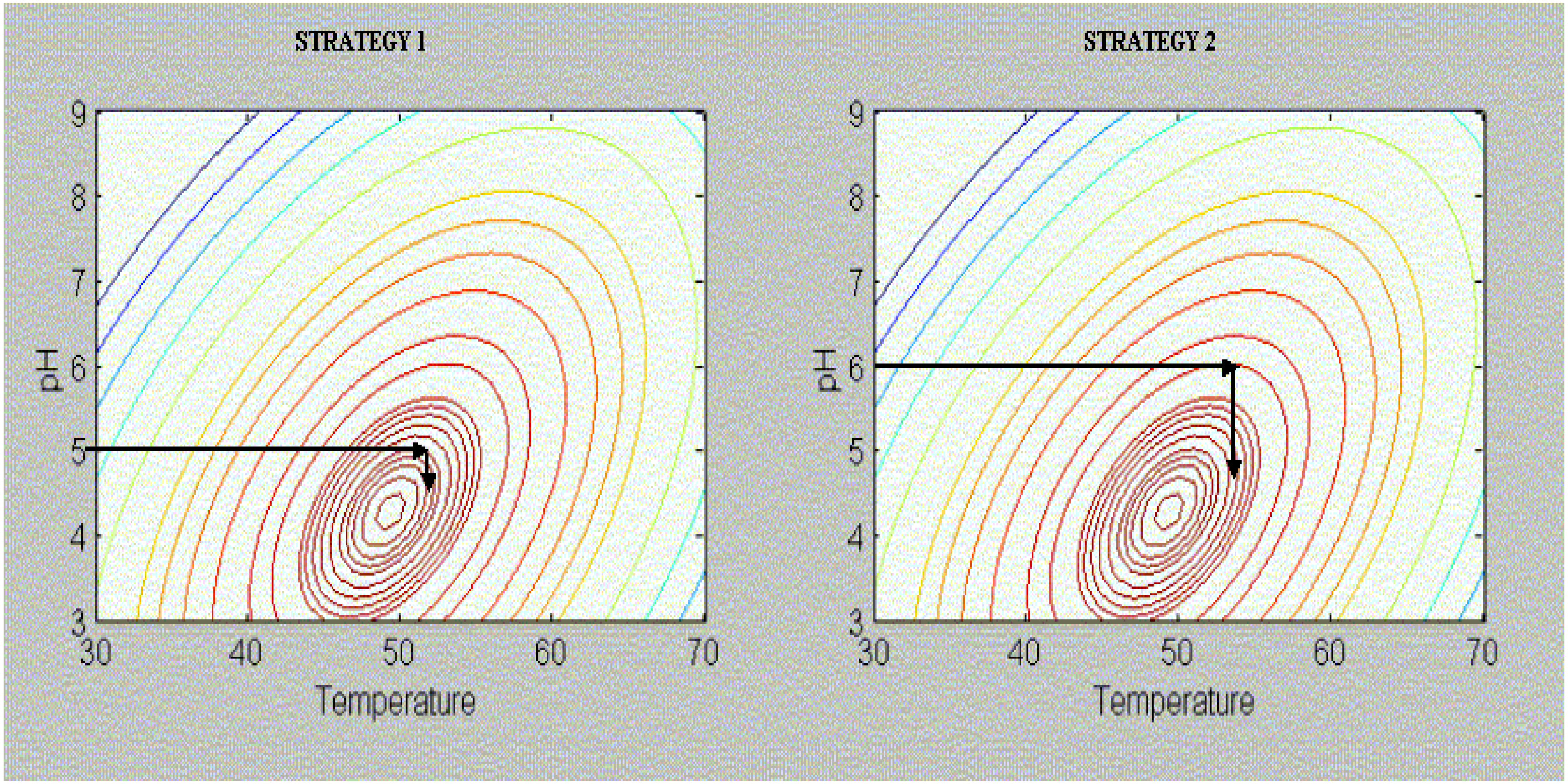


Progettazione degli esperimenti

La strategia di variare un fattore alla volta

(*One Variable At Time)*:

può mancare di cogliere l' "ottimo"



DIFFICOLTA’

Interazioni – la risposta per ciascun fattore non è indipendente

La temperatura ottimale a pH 5 è diversa da quella a pH 6.

Come affrontare il problema? Forza bruta?

- **Una griglia di esperimenti (*Grid search*). 10 pHs, 10 temperatures, 100 experiments.**

- **Si inizia con una griglia a maglia larga. Poi a maglia più stretta.**

Controindicazioni

- **Dispendioso in termini di tempo e denaro.**
- **Molti esperimenti vengono condotti in aree del dominio sperimentale che sono quasi sicuramente "non vicine" a un ottimo (quindi una perdita di tempo e denaro)**
- **Come stimare riproducibilità ed errore sperimentale? (Altri esperimenti, replica dei precedenti ?!?)**

Che facciamo?

Abbiamo bisogno di regole !

La progettazione formale degli esperimenti

Analytica Chimica Acta Volume 652, Issues 1-2,
12 October 2009, Pages 161-172
Experimental design in chemistry: A tutorial
Riccardo Leardi,
https://www.academia.edu/237865/Experimental_Design_in_Chemistry_a_Tutorial

Ottimizzazione sperimentale: approccio multivariato (multifattoriale)

L'approccio multivariato o multifattoriale:

prevede di valutare gli effetti delle variabili sulla risposta modificando più variabili contemporaneamente.

Questo approccio comporta una serie di **vantaggi**:

Riduzione del dispendio di tempi e risorse per effettuare gli esperimenti

Stima dell'effetto di tutte le variabili all'interno di diversi valori (livelli) di altre

Possibilità di rilevare interazioni tra variabili

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Disegno sperimentale (*Design of Experiments - DoE*)

Il disegno sperimentale prevede una pianificazione degli esperimenti da eseguire, identificando i seguenti parametri:

- ✓ la **RISPOSTA** (o variabile dipendente) che si vuole ottimizzare;
- ✓ i **FATTORI** (o variabili indipendenti) di cui si vuole osservare l'effetto sulla risposta.

Lo scopo del DoE è massimizzare le informazioni che sono ottenibili ***con un contenuto numero di esperimenti***.

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Disegno sperimentale (*Design of Experiments - DoE*)

La costruzione del disegno sperimentale prevede di:

- Individuare un <u>modello matematico</u> in base alla RISPOSTA e ai FATTORI scelti (di solito modelli di regressione multipla);
- Selezionare il <u>tipo</u> di disegno sperimentale;
- Selezionare dettagli rispetto alla <u>randomizzazione</u> e <u>ripetizione</u> degli esperimenti;
- Scegliere il metodo di <u>analisi dei dati</u> ottenuti dagli esperimenti.

I risultati consentiranno di <u>ottimizzare</u> il processo considerato, ad esempio:

- individuando condizioni adatte per ottenere una determinato valore della risposta;
- massimizzando o minimizzando la risposta;
- individuando la combinazione di parametri che riduce la variabilità della risposta.

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Fattori e risposta

I fattori e la risposta possono essere sia qualitativi che quantitativi.

Risposta quantitativa	**es. resa di una reazione, proprietà di un materiale, etc.**
Risposta qualitativa	**es. prodotto conforme o non conforme**
Fattori quantitativi	**es. temperatura, pressione, concentrazione, etc.**
Fattori qualitativi	**es. aggiunta o meno di catalizzatore, co-solvente, etc.**

In presenza di fattori quantitativi, in genere, essi vengono divisi in livelli (es. alto-basso, o più livelli) per consentire la costruzione del disegno sperimentale.

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Classificazione dei DoE

Uno dei modi di classificare i DoE è relativo agli scopi:

Esperimenti di Screening:

- Moderato numero di esperimenti;
- Identificazione dei fattori che danno l'effetto principale sulla variabilità della risposta;
- Modelli lineari;
- Fattori divisi in due livelli (es. alto-basso).

Esperimenti con Superfici di Risposta:

- Elevato numero di esperimenti;
- Identificazione di superfici di risposta;
- Identificazione di interazione tra fattori;
- Fattori divisi in tre livelli (es. alto-medio-basso).

Fonte: https://www.youtube.com/watch?v=RKhXKPV4JE0

Design di tipo fattoriale

***Full factorial design*:**

- Prevede l'utilizzo di due livelli per ogni fattore;
- Prevede di effettuare 2^k esperimenti, con k = numero di fattori;
- Quindi, ad es. per 3 fattori $2^3 = 8$ esperimenti; per 8 fattori $2^8 = 256$ esperimenti.

***Fractional factorial design*:**

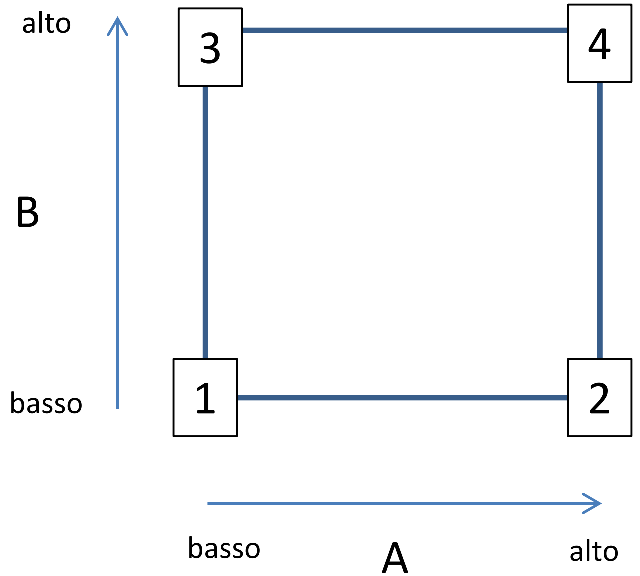
- Prevede l'utilizzo di due livelli per ogni fattore;
- Prevede di effettuare esperimenti su una sola parte delle combinazioni previste dal *Full factorial design*;
- ½ *factorial* = 2^{k-1} esperimenti (es. $2^{8-1} = 128$ esperimenti)
- ¼ *factorial* = 2^{k-2} esperimenti (es. $2^{8-2} = 64$ esperimenti)
- ⅛ *factorial* = 2^{k-3} esperimenti (es. $2^{8-3} = 32$ esperimenti)

Fonte: https://www.youtube.com/watch?v=RKhXKPV4JE0

Full factorial: esempio a 2 fattori

Fattori: A, B
Livelli: alto, basso codificati come +1 e -1 (es. alto e basso)
N° esperimenti: $2^2 = 4$

Exp	A	B
1	- 1	- 1
2	+ 1	- 1
3	- 1	+ 1
4	+ 1	+ 1

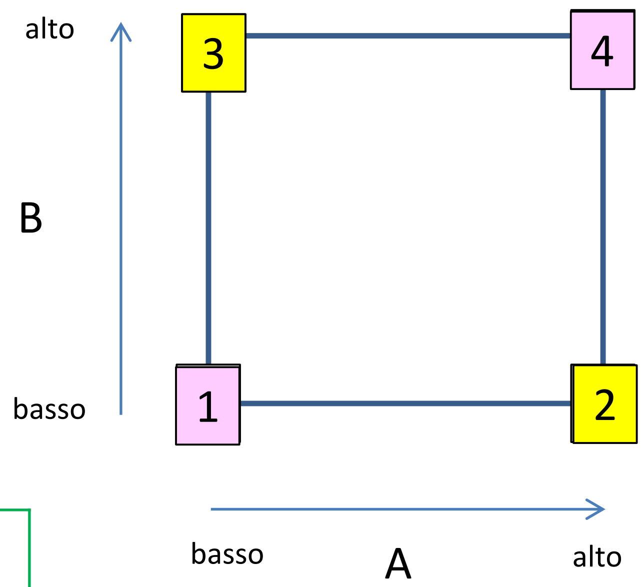


Fractional factorial: esempio a 2 fattori

Fattori: A, B
Livelli: alto, basso codificati come +1 e -1 (es. alto e basso)
N° esperimenti: $2^{2-1} = 2$

Come si sceglie? Il sottoinsiemi di esperimenti deve prevedere che entrambi i livelli di ogni fattore siano rappresentati almeno una volta (il controllo viene effettuato osservando la tabella per colonna).

Exp	A	B
1 (full)	- 1	- 1
2 (full)	+ 1	- 1
3 (full)	- 1	+ 1
4 (full)	+ 1	+ 1



Exp	A	B
1	+ 1	- 1
2	- 1	+ 1

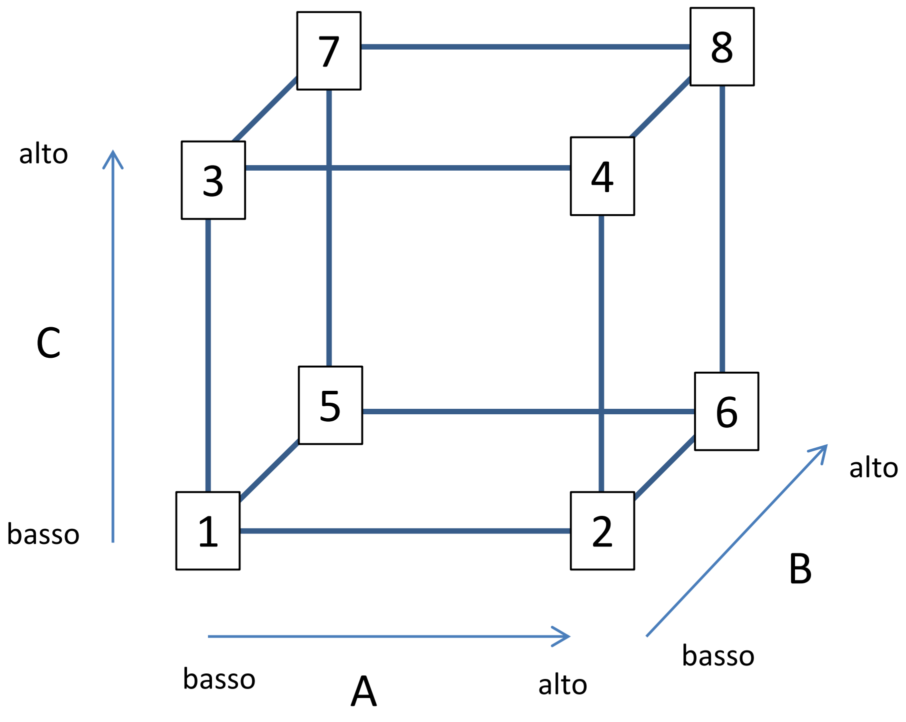
oppure

Exp	A	B
1	- 1	- 1
2	+ 1	+ 1

Full factorial: esempio a 3 fattori

Fattori: A, B, C
Livelli: alto, basso codificati come +1 e -1 (es. alto e basso)
N° esperimenti: $2^3 = 8$

Exp	A	B	C
1	- 1	- 1	- 1
2	+ 1	-1	- 1
3	- 1	- 1	+ 1
4	+ 1	- 1	+ 1
5	- 1	+ 1	- 1
6	+ 1	+ 1	- 1
7	- 1	+ 1	+ 1
8	+ 1	+ 1	+ 1



Fractional factorial: esempio a 3 fattori

Fattori: A, B, C
Livelli: alto, basso codificati come +1 e -1 (es. alto e basso)
N° esperimenti per ½ *factorial* $2^{3-1} = 4$

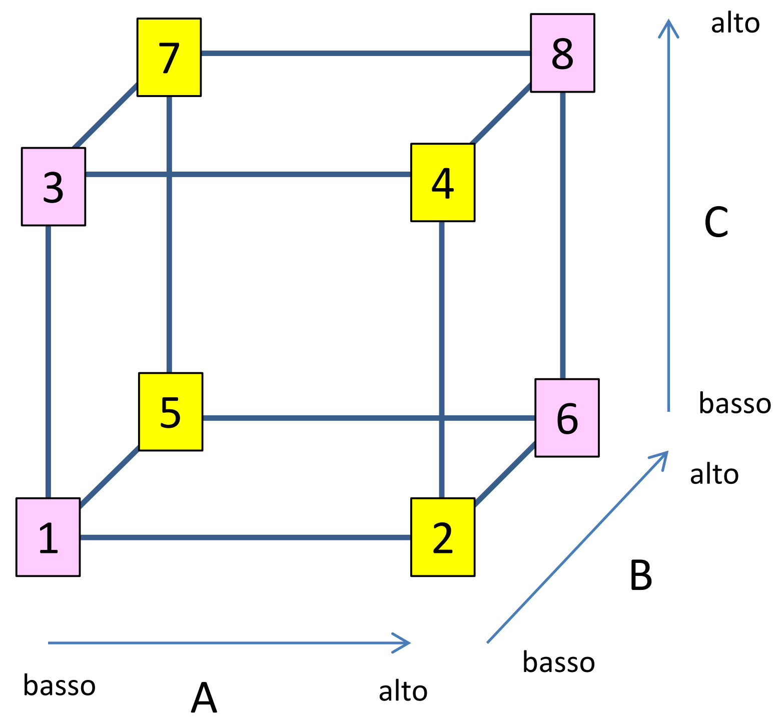
Come si sceglie? Il sottoinsiemi di esperimenti deve prevedere che entrambi i livelli di ogni fattore siano rappresentati due volte.

Exp	A	B	C
1(full)	- 1	- 1	- 1
2(full)	+ 1	- 1	- 1
3(full)	- 1	- 1	+ 1
4(full)	+ 1	- 1	+ 1
5(full)	- 1	+ 1	- 1
6(full)	+ 1	+ 1	- 1
7(full)	- 1	+ 1	+ 1
8(full)	+ 1	+ 1	+ 1

Exp	A	B	C
1	- 1	- 1	- 1
2	- 1	- 1	+ 1
3	+ 1	+ 1	- 1
4	+ 1	+ 1	+ 1

oppure

Exp	A	B	C
1	+ 1	- 1	- 1
2	+ 1	- 1	+ 1
3	- 1	+ 1	- 1
4	- 1	+ 1	+ 1



Qualità del disegno sperimentale

Ortogonalità (*orthogonality*) e disegno bilanciato (*balanced design*) sono due caratteristiche desiderabili che il disegno sperimentale dovrebbe avere.

Exp	A	B	C	*
1	+ 1	- 1	- 1	1
2	+ 1	- 1	+ 1	-1
3	- 1	+ 1	- 1	1
4	- 1	+ 1	+ 1	-1
	0	**0**	**0**	**0**

Balanced Design: la somma dei valori di colonna per ogni fattore è uguale a zero. Cioè i livelli del fattore sono rappresentati in eguale numero negli esperimenti

Orthogonality: due o più vettori sono ortogonali se la somma dei prodotti dei loro elementi è uguale a zero.

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Qualità del disegno sperimentale (2)

Randomizzazione: è necessario eseguire gli esperimenti in ordine più possibile casuale, ad esempio:

Exp	A	B	C
1	- 1	- 1	- 1
2	- 1	- 1	+ 1
3	+ 1	+ 1	- 1
4	+ 1	+ 1	+ 1

Exp	A	B	C
1	- 1	- 1	- 1
2	+ 1	+ 1	- 1
3	- 1	- 1	+ 1
4	+ 1	+ 1	+ 1

Meglio questo

Prove al punto centrale

Il disegno potrà essere migliorato addizionando punti centrali (solitamente 3, non randomizzati: inizio, metà e fine dell'esperimento). Punto centrale: tutti i fattori = 0 (cioè, nella realtà, il valore medio tra i livelli).

Replicazione

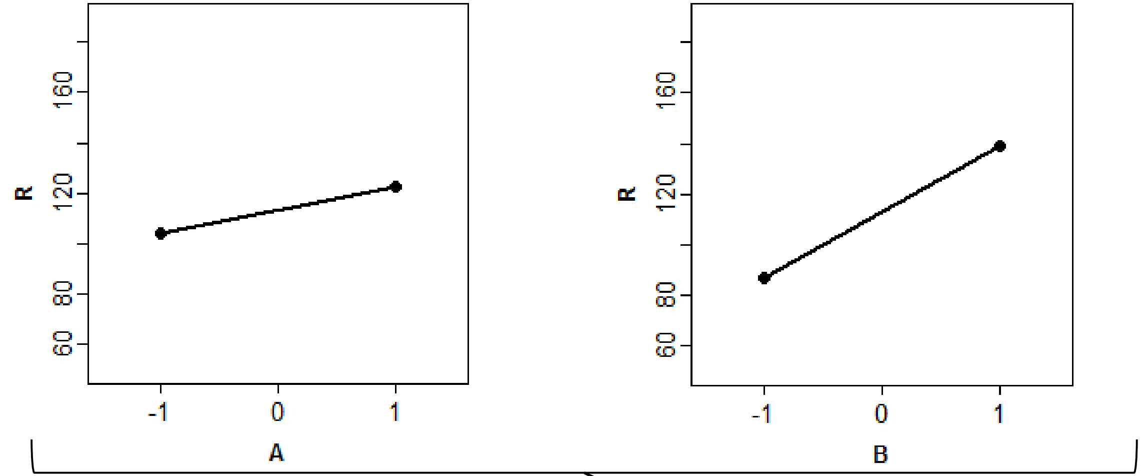
Fornisce informazioni sulla variabilità delle risposte: eseguire il disegno più di una volta facilita l'analisi dei dati perché per ogni prova si ha media e dispersione delle risposte.

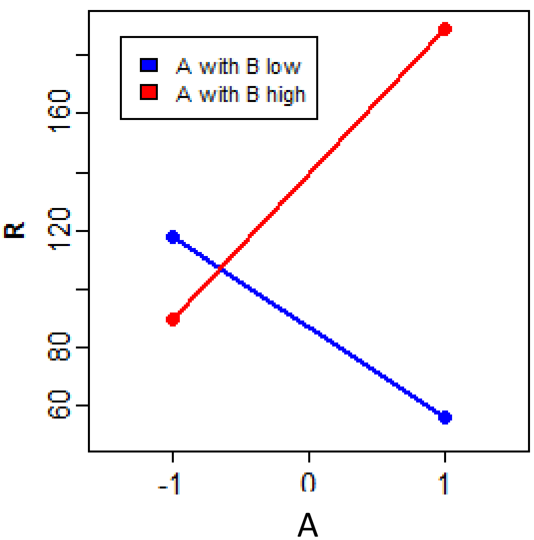
Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Interazioni

Una interazione avviene quando il livello di uno o più fattori determina l'effetto di uno o più altri fattori sulla risposta. Esempio:

A	B	R
1	1	185
1	1	193
1	-1	46
1	-1	66
-1	1	96
-1	1	83
-1	-1	132
-1	-1	104





Ma se si considera l'effetto di A sulla risposta, separatamente, a seconda dei livelli di B si ottiene questo grafico.

Quindi per massimizzare la risposta, di fatto, se B è elevato, A deve essere elevato. Mentre, se si deve minimizzare la risposta, A deve essere elevato e B basso.

Osservando i grafici pare che per massimizzare la risposta, A deve essere elevato e anche B deve essere elevato.

L'interazione è presente se le rette rappresentate non risultano parallele

Fonte: https://www.youtube.com/watch?v=RKhXKPV4JE0

Fattori confondenti *(Confunding factors)*

Nel disegno sperimentale, quando la moltiplicazione di due fattori può dare lo stesso effetto di un terzo fattore, essi si definiscono confondenti.

Exp	A	B	C	A*B
1	- 1	+ 1	- 1	- 1
2	- 1	- 1	+ 1	+ 1
3	+ 1	+ 1	+ 1	+ 1
4	+ 1	- 1	- 1	- 1

In questo esempio l'effetto del fattore C sulla risposta può essere confuso con l'effetto del prodotto tra i fattori A e B (ovvero della loro interazione) sulla risposta (o viceversa).

Questa condizione avviene più facilmente quando il disegno sperimentale è *fractional factorial*.

Fonte: https://www.youtube.com/watch?v=RKhXKPV4JE0

Blocking factors

Molto spesso quando si effettuano degli esperimenti esistono delle variabili "indesiderate" di cui però bisogna tener conto e che devono il meno possibile condizionare i risultati ottenuti dal disegno sperimentale, cioè non devono impedire l'identificazione dei fattori di interesse che contribuiscono alla variabilità della risposta.

Esempi di queste variabili possono essere:
il giorno in cui viene effettuato l'esperimento, il lotto del materiale su cui avviene l'esperimento, il cambio di operatore, l'aggiornamento della taratura di uno strumento, etc..

Per considerare l'effetto di queste variabili così dette *blocking variables*, bisogna considerarle come fattori da aggiungere al disegno sperimentale.
In questo caso, dai dati ottenuti, deve risultare che questi fattori di *blocking*, non portano a differenze significative nella risposta.

Fonte: https://www.youtube.com/watch?v=RKhXKPV4JE0

Altri tipi di disegni sperimentali

La scelta del disegno sperimentale è fatta sulla base dei risultati desiderati. Altri esempi sono:

Completely randomized designs

Randomized block designs
- Latin squares
- Graeco-Latin squares
- Hyper-Graeco-Latin squares

Plackett-Burman designs

Response surface (second-order) designs
- Central composite designs
- Box-Behnken designs
- Response surface designs

Three-level full factorial designs

Three-level, mixed level and fractional factorial designs

In R:

agricolae package

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Modelli matematici

I modelli matematici principalmente utilizzati sono modelli lineari di primo o secondo ordine, con la considerazione o meno di interazioni tra fattori.

Modello lineare di primo ordine senza interazioni:

$$\hat{R} = \hat{\beta}_0 + \hat{\beta}_1 * F_{1} + \hat{\beta}_2 * F_2 + \ldots + \hat{\beta}_k * F_k$$

Modello lineare di primo ordine con interazioni:

$$\hat{R} = \hat{\beta}_0 + \hat{\beta}_1 * F_{1} + \hat{\beta}_2 * F_2 + \hat{\beta}_{12} * F_1 * F_2 + \ldots + \hat{\beta}_{k,k-1} * F_{k-1} * F_k + \hat{\beta}_k * F_k$$

Modello lineare di secondo ordine senza interazioni:

$$\hat{R} = \hat{\beta}_0 + \hat{\beta}_1 * F_{1} + \hat{\beta}_2 * F_2 + \hat{\beta}_{11} * F_1^2 + \hat{\beta}_{22} * F_2^2 + \ldots + \hat{\beta}_k * F_k^2$$

Modello lineare di secondo ordine con interazioni:

$$\hat{R} = \hat{\beta}_0 + \hat{\beta}_1 * F_{1} + \hat{\beta}_2 * F_2 + \hat{\beta}_{12} * F_1 * F_2 + \hat{\beta}_{123} * F_1 * F_2 * F_3 + \hat{\beta}_{11} * F_1^2 + \ldots + \hat{\beta}_k * F_k^2$$

Fonte: http://users.unimi.it/gazzalab/wordpress/wp-content/uploads/2011/12/15-Design-of-Experiment-DOE.pdf

Analisi dei dati

Dopo aver proceduto a:

Exp	A	B	C
1	- 1	- 1	- 1
2	+ 1	-1	- 1
3	- 1	- 1	+ 1
4	+ 1	- 1	+ 1
5	- 1	+ 1	- 1
6	+ 1	+ 1	- 1
7	- 1	+ 1	+ 1
8	+ 1	+ 1	+ 1

- scelta del disegno sperimentale, ad es.:
- scelta del modello, ad es.: $\hat{R} = \hat{\beta}_0 + \hat{\beta}_1 * F_1 + \hat{\beta}_2 * F_2 + \ldots + \hat{\beta}_k * F_k$
- effettuazione gli esperimenti,

Si ottiene una tabella dei risultati del tipo:

Exp	A	B	C	R (risposta)
1	- 1	- 1	- 1	35
2	+ 1	-1	- 1	87
3	- 1	- 1	+ 1	29
4	+ 1	- 1	+ 1	76.5
5	- 1	+ 1	- 1	34
6	+ 1	+ 1	- 1	55.2
7	- 1	+ 1	+ 1	64
8	+ 1	+ 1	+ 1	92.3

COME INDIVIDUO I FATTORI CHE INFLUISCONO SULLA VARIABILITA' DELLA RISPOSTA?

ANOVA = ANalysis Of VAriance

Fonte: https://www.youtube.com/watch?v=RKhXKPV4JE0

F-value

Serve a confrontare la varianza di <u>due</u> serie di dati (con unità di misura omogenea).

$$F = \frac{\sigma_a^2}{\sigma_b^2}$$

n_a = n° osservazioni di a

n_b = n° osservazioni di b

$$\sigma_a^2 > \sigma_b^2$$

Si pone al numeratore sempre la varanza più alta

$$F_{critical} = F_{\alpha,d.f.(a),d.f.(b)}$$

Si confronta con i valori critici della *F-distribution* (*one-tailed* test)

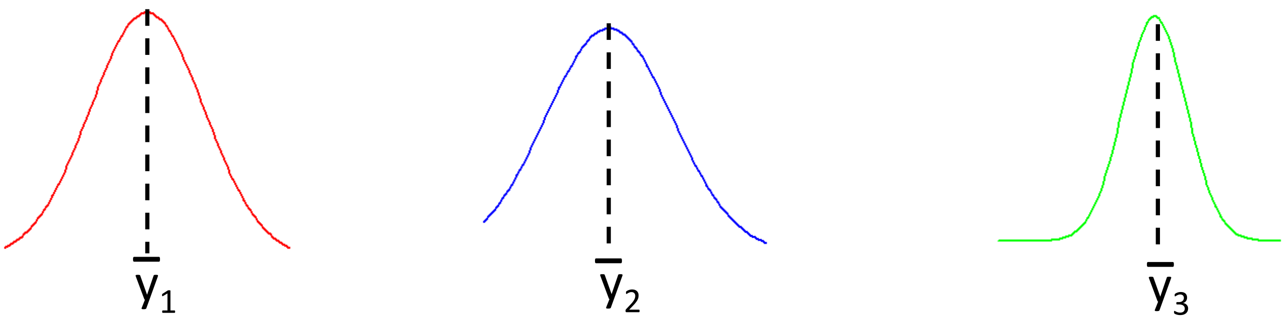
$$\text{se } F > F_{critical}$$

allora la differenza tra le varianze della serie "a" e "b" è statisticamente significativa

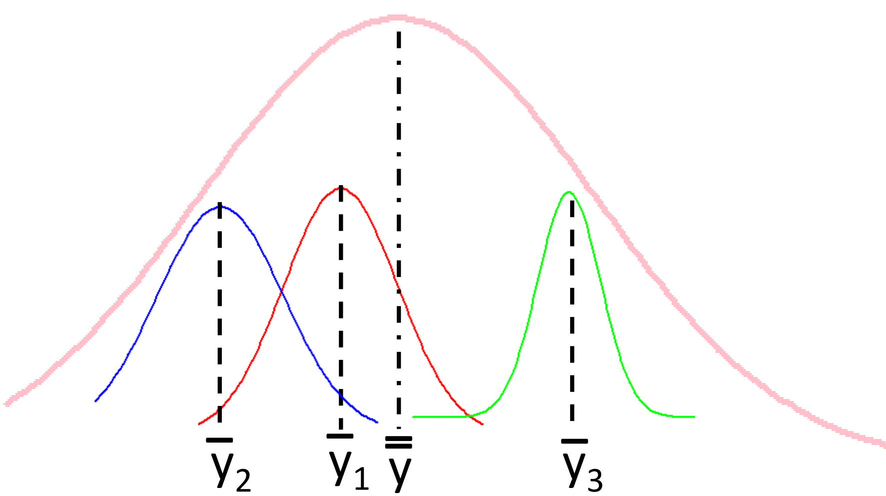
Fonte: https://www.bcfoltz.com/blog/stats-101/

ANOVA: ANalysis Of VAriance

Serve a confrontare la varianza di <u>più di due serie</u> di dati (con unità di misura omogenea).



Ipotesi nulla $H_0 : \mu_1 = \mu_2 = \mu_3$

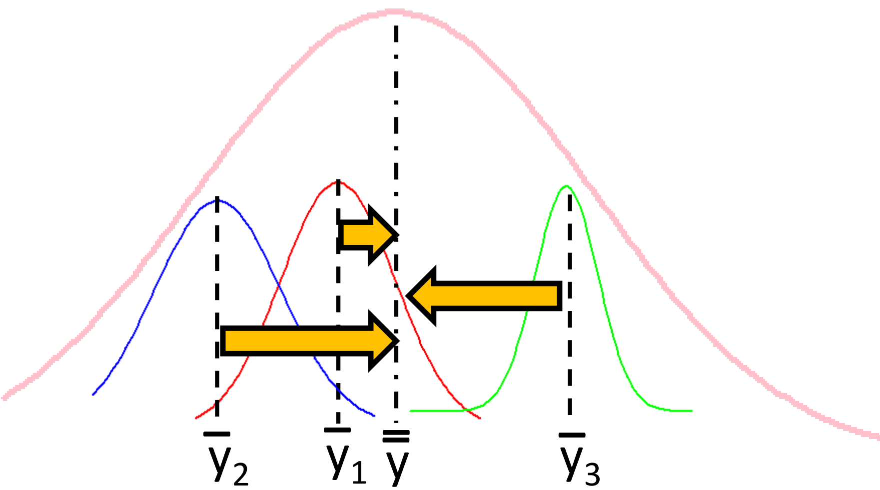


Le medie delle serie di dati (con le loro varianze) appartengono alla stessa popolazione (di media $\bar{\bar{y}}$)

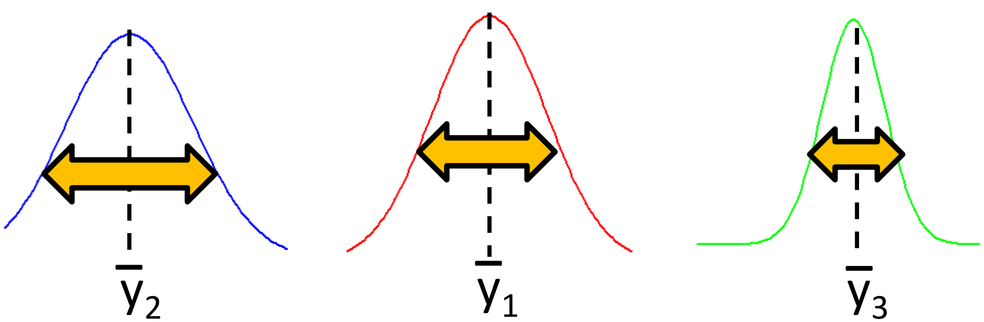
Fonte: https://www.bcfoltz.com/blog/stats-101/

ANOVA: ANalysis Of VAriance (2)

In ANOVA si confrontano due tipi di varianze:

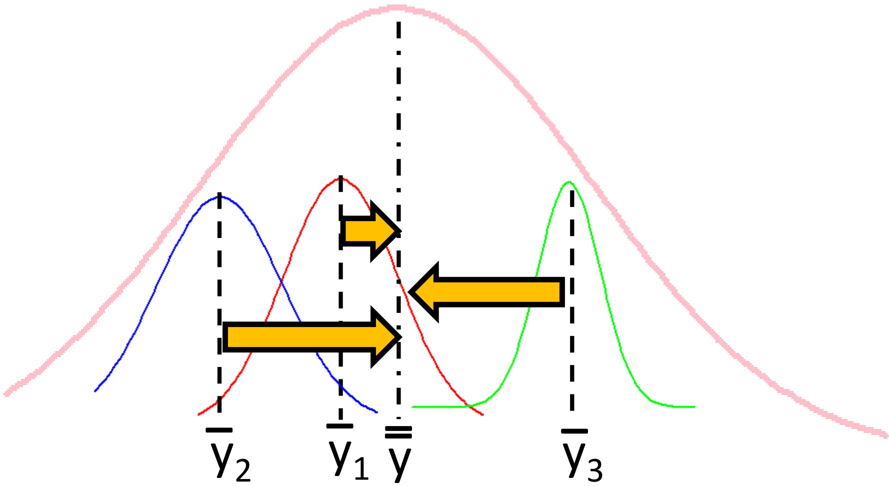


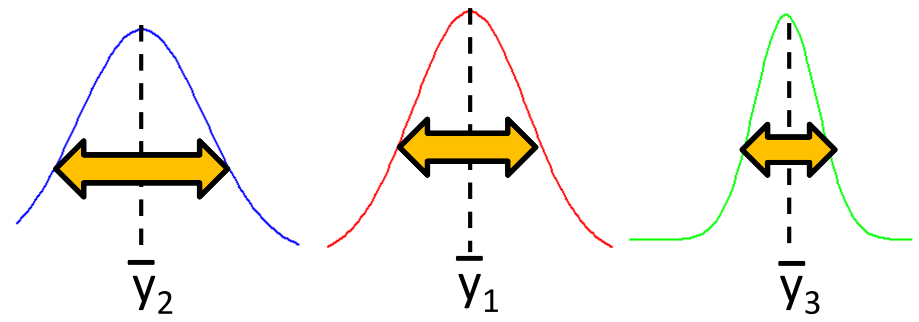
Variabilità tra le medie (***BETWEEN*** o *among*). Si misura la distanza delle medie delle diverse serie, dalla media complessiva (*overall*),



Variabilità interne (***WITHIN*** o *around*) delle serie.

ANOVA: ANalysis Of VAriance (3)



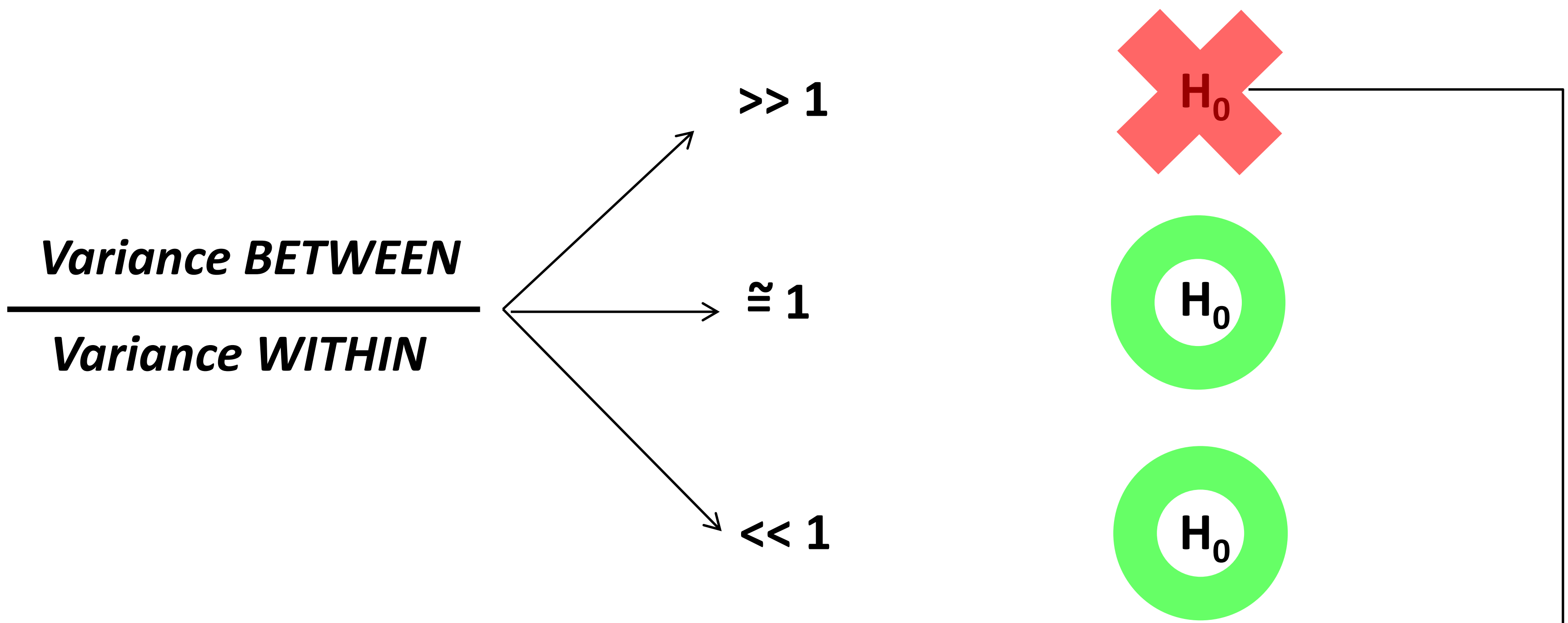


TOTAL Variance = Variance BETWEEN + Variance WITHIN

Fonte: https://www.bcfoltz.com/blog/stats-101/

ANOVA: ANalysis Of VAriance (4)

TOTAL Variance = Variance BETWEEN + Variance WITHIN



Le distribuzioni sono strette e distinte, almeno una è nettamente distinta dalle altre

Fonte: https://www.bcfoltz.com/blog/stats-101/

Assunzioni su ANOVA

L'utilizzo di ANOVA prevede che alcune assunzioni siano soddisfatte:

- ✓ La variabile dipendente (risposta) deve essere continua;
- ✓ La variabile dipendente (risposta) deve avere una distribuzione normale;
- ✓ Omogeneità delle varianze (omoschedasticità).

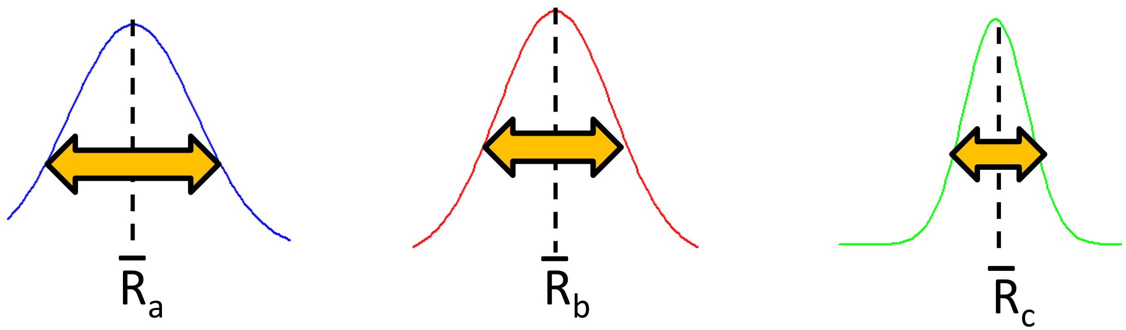
Fonte: http://www.sthda.com/english/wiki/compare-multiple-sample-variances-in-r; http://www.sthda.com/english/wiki/normality-test-in-r

One-way or single factor ANOVA

Serve a comparare l'influenza di più livelli di un fattore su una risposta.

Supponiamo di avere un **fattore F** di cui si considerano **3 livelli**: a, b, c.

Si raggruppano le risposte in base al livello del fattore che è stato usato nell'esperimento.



Domanda: c'è differenza significativa nella risposta R rispetto ai diversi livelli di F?

N.B: i livelli vengono anche chiamati gruppi o colonne o trattamenti

Fonte: https://www.bcfoltz.com/blog/stats-101/

Sum of Squares (SS)

Partendo dal concetto:

TOTAL Variance = Variance BETWEEN + Variance WITHIN

Si considera:

SST = SSG + SSE

$$SST = \sum_{i=1}^{n} (R_i - \bar{\bar{R}})^2$$

$$SSG = \sum_{j=1}^{g} n_j (\bar{R}_j - \bar{\bar{R}})^2$$

$$SSE = \sum_{j=1}^{g} \sum_{i=1}^{n_g} (R_{ji} - \bar{R}_j)^2 = \sum_{j=1}^{g} (n_j - 1)\sigma_j^2$$

n = numero esperimenti

g = numero gruppi/livelli

Fonte: https://www.bcfoltz.com/blog/stats-101/

Mean Squared (MS) e F-value

$$MSG = \frac{SSG}{d.f._{(G)}}$$

Media dei quadrati relativi ai gruppi - ***between***
(*mean squared of groups*)

$$d.f._{(G)} = g - 1$$

$$MSE = \frac{SSE}{d.f._{(E)}}$$

Media dei quadrati ***within***
(*mean squared error*)

$$d.f._{(E)} = n - g$$

$$F = \frac{MSG}{MSE}$$

$$F_{critical} = F_{\alpha, d.f.(G), d.f.(E)}$$

se $F > F_{critical}$ $\cancel{\mathbf{H_0}}$

ANOVA table

Alla fine delle elaborazioni si ottiene la così detta tabella ANOVA:

	SS	d.f.	MS	F-value
Gruppi	SSG	g - 1	MSG	F
Errore	SSE	n - g	MSE	
Totale	SST	n - 1		

Two-way or two factor ANOVA (senza repliche)

Serve a comparare l'influenza di più livelli di due fattori su una risposta.

Exp	F_1
1	a
2	b
3	c
4	d
5	a
6	b
7	c
8	d

Esempio one-way

Exp	F_1	F_2
1	a	α
2	b	β
3	c	α
4	d	β
5	a	α
6	b	β
7	c	α
8	d	β

Esempio two-way

Fonte: https://www.bcfoltz.com/blog/stats-101/

Two-way or two factor ANOVA (senza repliche) (2)

La varianza totale in questo caso viene spiegata da:

$$\mathbf{SST = SSG_{(F1)} + SSG_{(F2)} + SSE}$$

Quindi, in questo caso bisogna verificare due F-value contro $F_{critical}$:

$$F_{(F1)} = \frac{MSG_{(F1)}}{MSE} \qquad F_{(F2)} = \frac{MSG_{(F2)}}{MSE}$$

Se uno dei due fattori è un *blocking factor* si auspica che per esso, es. F_2, $F_{(F2)} < F_{critical}$, cioè che l'ipotesi H_0 sia confermata, ovvero che non influisca significativamente sulla variabilità della risposta.

Contemporaneamente si auspica che F_1 sia il *main factor*, cioè che $F_{(F1)} > F_{critical}$

Fonte: https://www.bcfoltz.com/blog/stats-101/

Three-way ... multi-way ANOVA

Aggiungendo fattori la complessità aumenta, ma comunque si deve focalizzare l'attenzione sugli F-value.

I software statistici, come R, effettuano anche questi calcoli più complessi.

Attenzione!!! Se si considerano anche le interazioni tra fattori i risultati da interpretare saranno numerosi!

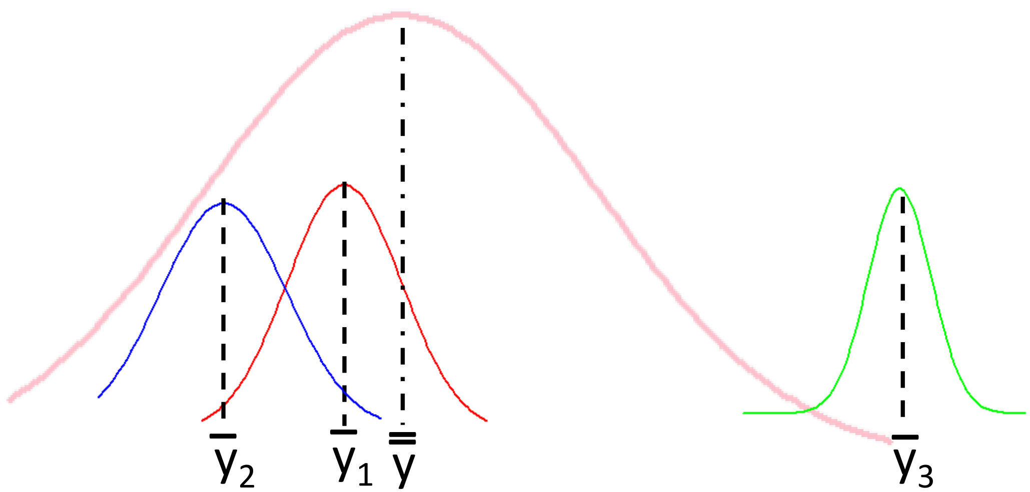
Se si aggiungono anche le repliche aumenta la complessità

	SS	**d.f.**	**MS**	**F-value**
F_1				
F_2				
F_3				
$F_1 * F_2$				
$F_1 * F_3$				
$F_2 * F_3$				
$F_1 * F_2 * F_3$				

Per aggiungere ancora complessità... se si vuole considerare l'effetto di più fattori su più di una risposta → **MANOVA (Multivariate ANalysis Of VAriance)**

Post-hoc tests

L'informazione che dà ANOVA, se $F_{(Fi)} > F_{critical}$, è che all'interno di quel fattore uno o più livelli appartengono a una popolazione diversa (ovviamente in caso di 3 o più livelli), ma quale? Ci sono alcuni test che consentono di capire quali sono i livelli "distanti".



Tukey's HSD (Honest Significant Difference) test

Student Newman Keuls (SNK)

Dunnett's test
(solo per trattamenti vs. controllo)