CAPITOLO 4 – Controllo statistico della qualità dei prodotti e dei processi produttivi
Paragrafo 4.2.7 – ANOVA a 2 vie

4

Paragrafo 4.3. Metodi off-line: ANOVA a due vie con dati bilanciati

Esempio: produzione di film per shopper biodegradabili

Variabile risposta: spessore (in micron).

Fattore 1: temperatura di lavorazione

Livelli di fattore 1: Bassa: 190°C, Media: 192°C, Alta:194°C

Fattore 2: quantità (in percentuale) di una speciale sostanza che favorisce la biodegradabilità del film

Livelli di fattore 2: 2%, 3%

N. trattamenti: $6=2\times 3$

N. replicazioni per trattamento: n (**dati bilanciati**)

Totale n. prove: $6n$

I trattamenti

		Livelli del fattore colonna (temperatura)		
		bassa (livello 1)	media (livello 2)	alta (livello 3)
Livelli del fattore riga (% sostanza aggiunta)	2% (livello 1)			
	3% (livello 2)			

6 trattamenti

Ogni trattamento è identificato dalla coppia (i,j) dove:

i: indica il livello del fattore riga: % sostanza ($i=1,2$)

j: indica il livello del fattore colonna: livello temperatura ($j=1,2,3$)

Ipotesi di base (*n* replicazioni)

	bassa (col. 1)	media (col. 2)	alta (col. 3)	Marginale
2% (riga 1)	$N(\mu_{11};\sigma^2)$	$N(\mu_{12};\sigma^2)$	$N(\mu_{13};\sigma^2)$	$N(\mu_{1\bullet};\sigma^2)$
3% (riga 2)	$N(\mu_{21};\sigma^2)$	$N(\mu_{22};\sigma^2)$	$N(\mu_{23};\sigma^2)$	$N(\mu_{2\bullet};\sigma^2)$
Marginale	$N(\mu_{\bullet 1};\sigma^2)$	$N(\mu_{\bullet 2};\sigma^2)$	$N(\mu_{\bullet 3};\sigma^2)$	$N(\mu;\sigma^2)$

$$\mu_{i\cdot} = \frac{n\mu_{i1} + n\mu_{i2} + n\mu_{i3}}{3n} = \frac{\mu_{i1} + \mu_{i2} + \mu_{i3}}{3}, \quad i = 1,2$$

$$\mu_{\cdot j} = \frac{n\mu_{1j} + n\mu_{2j}}{2n} = \frac{\mu_{1j} + \mu_{2j}}{2}, \quad j = 1,2,3$$

$$\mu = \text{media generale}$$

ANOVA 2 vie: formalizzazione

Quantità sostanza aggiunta	Temperatura		
	bassa	medie	alta
2%	$Y_{111}, Y_{112}, \ldots, Y_{11n}$	$Y_{121}, Y_{122}, \ldots, Y_{12n}$	$Y_{131}, Y_{132}, \ldots, Y_{13n}$
3%	$Y_{211}, Y_{212}, \ldots, Y_{21n}$	$Y_{221}, Y_{222}, \ldots, Y_{22n}$	$Y_{231}, Y_{232}, \ldots, Y_{23n}$

Y_{ijk}: spessore dello shopper k prodotto col trattamento (i,j)

$$Y_{ijk} \sim N(\mu_{ij}; \sigma^2)$$

$$\Downarrow$$

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \qquad \varepsilon_{ijk} \sim N(0; \sigma^2)$$

$i=1,2;\ j=1,2,3;\ k=1,\ldots,n$

ANOVA a 2 vie: il modello generale

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$$

componente sistematica

componente accidentale

dove $\varepsilon_{ijk} = (Y_{ijk} - \mu_{ij}) \sim N(0;\sigma^2)$ è l'errore sperimentale

$i=1,..,a$ livello fattore riga

$j=1,...,b$ livello fattore colonna

Possibili configurazioni del modello (specificazioni di μ_{ij})

Esemplifichiamo mediante:

- **tabella delle medie** dei trattamenti
- relativa **rappresentazione grafica**
- formalizzazione del **modello**

le seguenti situazioni:

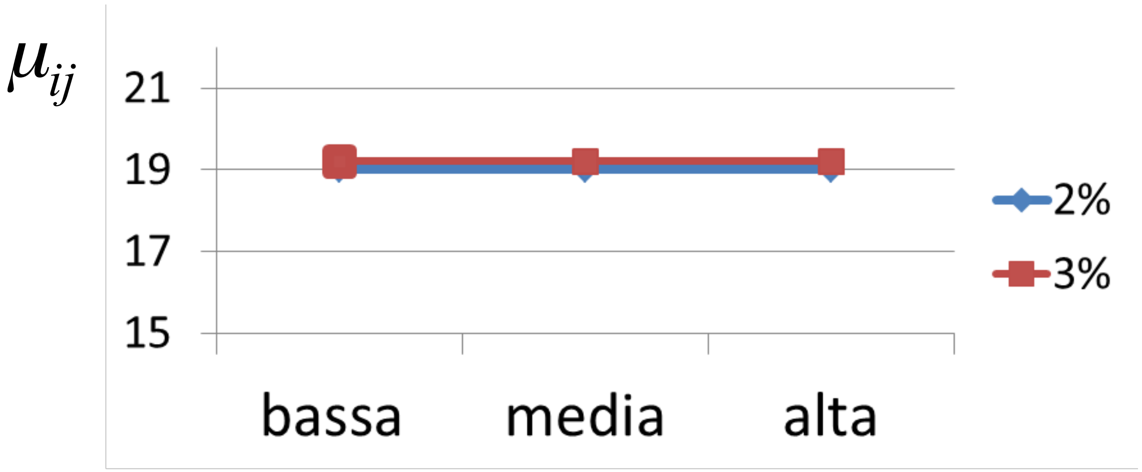
1. nessun fattore influenza lo spessore
2. solo effetto del fattore **riga** (quantità sostanza aggiunta)
3. solo effetto del fattore **colonna** (temperatura)
4. effetto **additivo** dei due fattori
5. effetto **additivo** e di **interazione** fra i due fattori

1. Nessun effetto dei due fattori

Tutte le medie sono uguali fra loro. Esempio: $\mu_{ij}=\mu=19$

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	19	19	19	19
3%	19	19	19	19
Marginale	19	19	19	19

Medie per ogni trattamento



PEARSON

1. Nessun effetto dei due fattori: $\mu_{ij} = \mu$

$$Y_{ijk} = \mu + \varepsilon_{ijk} \quad \Longrightarrow \quad Y_{ijk} - \mu = \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim N(0;\sigma^2)$$

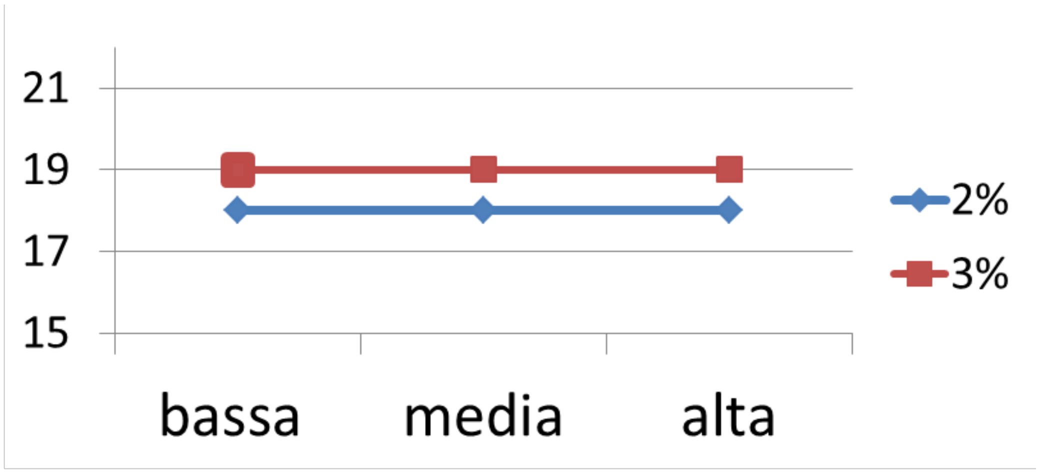
La variabilità dei valori della variabile risposta intorno alla media generale è dovuta solo a fattori accidentali (all'**errore sperimentale** ε_{ijk})

2. Solo effetto del fattore riga

Tutte le medie delle caselle di riga sono uguali: $\mu_{ij} = \mu_{i.}$

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	18	18	18	18
3%	19	19	19	19
Marginale	18.5	18.5	18.5	18.5

μ_{ij}



PEARSON

2. Solo effetto del fattore riga: $\mu_{ij} = \mu_{i\cdot}$

$$Y_{ijk} = \mu_{i\cdot} + \varepsilon_{ijk} \quad \Longrightarrow \quad Y_{ijk} - \mu = (\mu_{i\cdot} - \mu) + \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim N(0;\sigma^2)$$

La variabilità dei valori della variabile risposta intorno alla media generale è dovuta a:

i)fattori accidentali (**errore sperimentale** ε_{ijk})

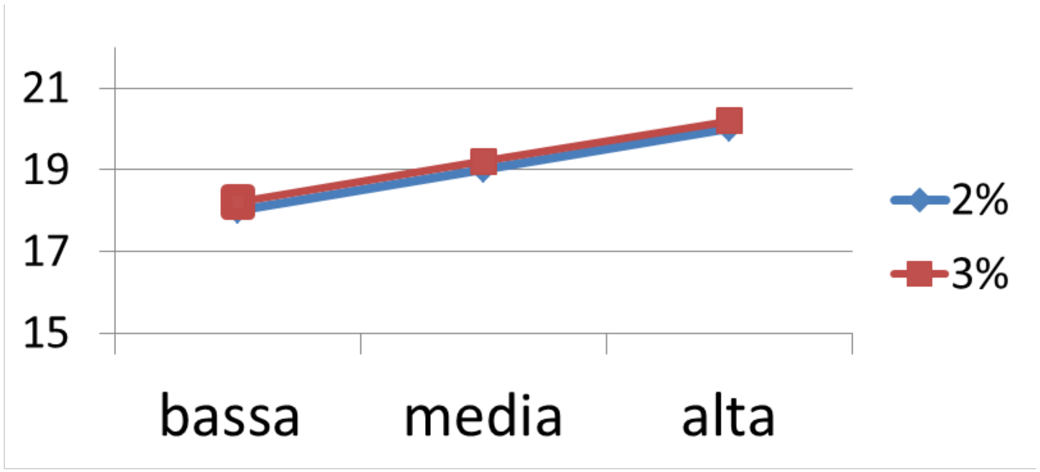
ii)effetto sistematico del fattore riga: $(\mu_{i\cdot} - \mu) \neq 0$ per almeno un livello i

3. Solo effetto del fattore colonna

Tutte le medie delle caselle di colonna sono uguali: $\mu_{ij} = \mu_{\cdot j}$

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	18	19	20	19
3%	18	19	20	19
Marginale	18	19	20	19

μ_{ij}



3. Solo effetto del fattore colonna: $\mu_{ij} = \mu_{\cdot j}$

$$Y_{ijk} = \mu_{\cdot j} + \varepsilon_{ijk} \quad \Longrightarrow \quad Y_{ijk} - \mu = (\mu_{\cdot j} - \mu) + \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim N(0;\sigma^2)$$

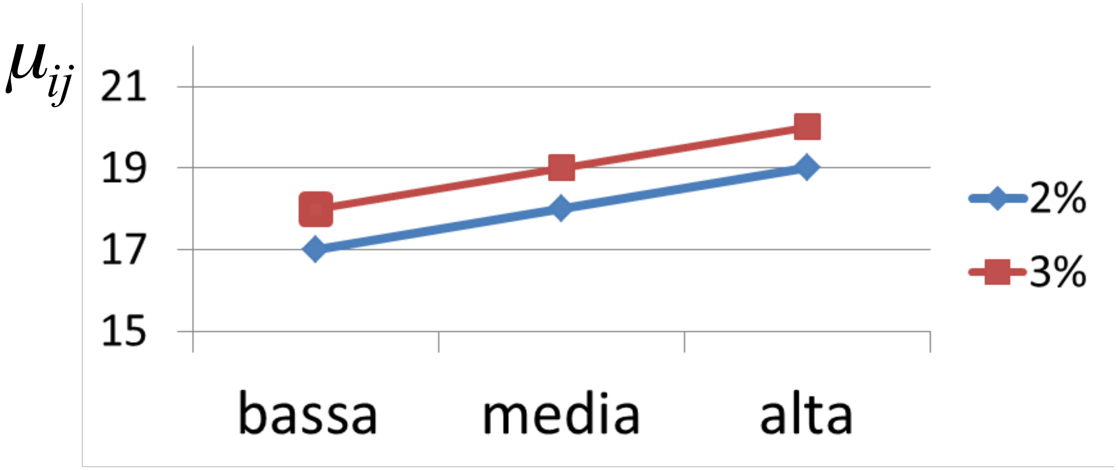
La variabilità dei valori della variabile risposta intorno alla media generale è dovuta a:

i)fattori accidentali (**errore sperimentale** ε_{ijk})

ii)effetto sistematico del fattore colonna: $(\mu_{\cdot j} - \mu) \neq 0$ per almeno un livello j

4. Effetto additivo dei due fattori

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	17	18	19	18
3%	18	19	20	19
Marginale	17.5	18.5	19.5	18.5



Effetto additivo (**spezzate parallele**): l'effetto di un fattore non dipende dal livello dell'altro fattore.

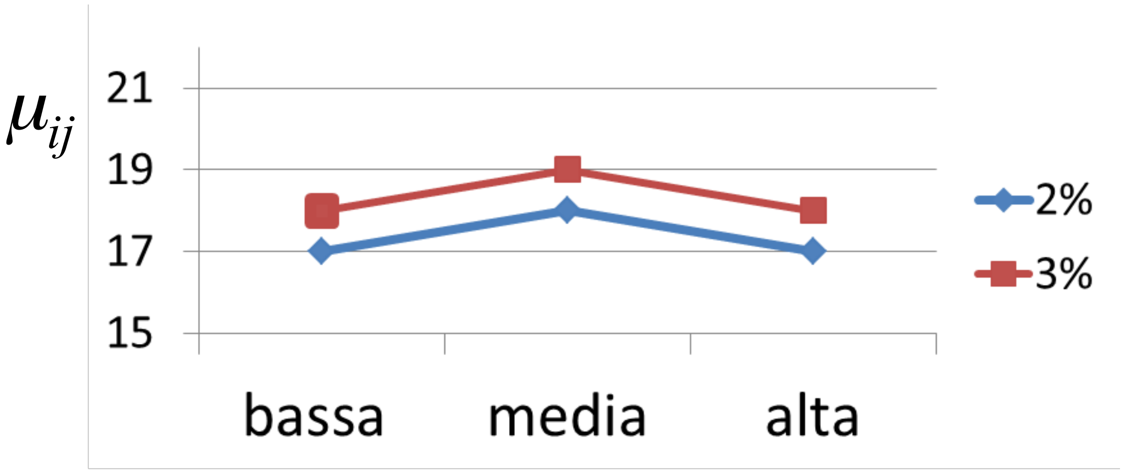
La temperatura fa aumentare lo spessore medio qualunque sia la quantità della sostanza aggiunta.

PEARSON

4. Ancora effetti additivi dei due fattori

Effetto additivo (spezzate parallele)

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	17	18	17	17.3
3%	18	19	18	18.3
Marginale	17.5	18.5	17.5	17.8



4. Effetto additivo dei due fattori

$$\mu_{ij} = \mu + (\mu_{i.} - \mu) + (\mu_{.j} - \mu)$$

media generale che cosa accade in assenza di effetti	**effetto principale fattore riga** differenza fra media di riga e media generale	**effetto principale fattore colonna** differenza fra media di colonna e media generale

$$Y_{ijk} - \mu = (\mu_{i.} - \mu) + (\mu_{.j} - \mu) + \varepsilon_{ijk} \qquad \varepsilon_{ijk} \sim N(0;\sigma^2)$$

La variabilità intorno alla media generale è dovuta a:

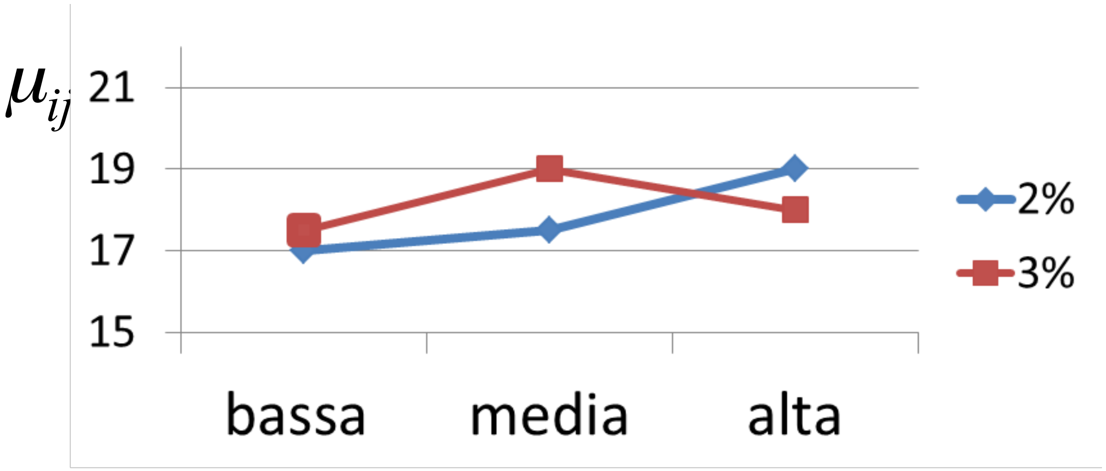
i)fattori accidentali (**errore sperimentale** ε_{ijk})

ii)effetto sistematico del fattore riga: $(\mu_{i.} - \mu) \neq 0$ per almeno un i

iii)effetto sistematico del fattore colonna: $(\mu_{.j} - \mu) \neq 0$ per almeno un j

5. Effetto di interazione

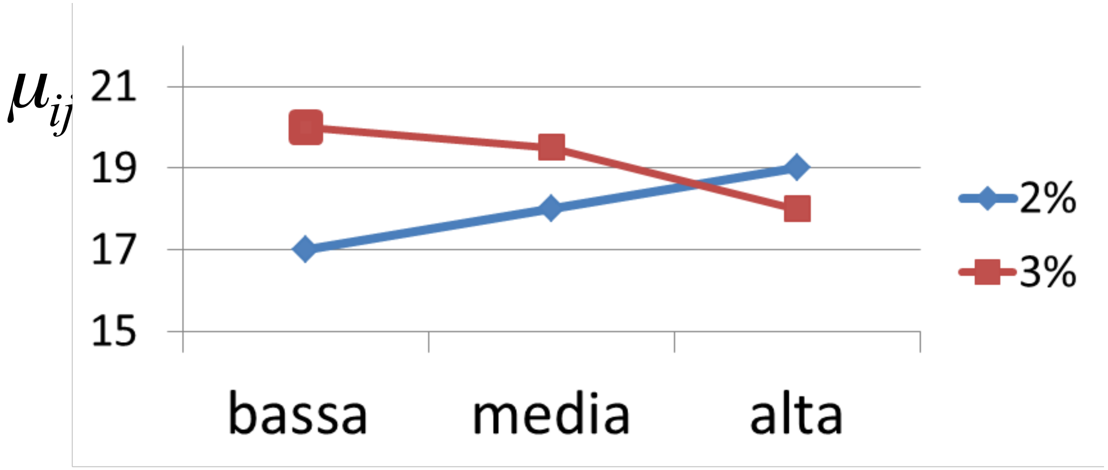
Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	17.0	17.5	19.0	17.8
3%	17.5	19.0	18.0	18.2
Marginale	17.3	18.3	18.5	18.0



Effetto di interazione (spezzate non parallele): l'effetto di un fattore dipende dal livello dell'altro fattore

5. Ancora su effetto di interazione

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	17.0	18.0	19.0	18.0
3%	20.0	19.5	18.0	19.2
Marginale	18.5	18.8	18.5	18.6



5. Il modello con interazione

$$\mu_{ij}=\mu + (\mu_{i.}-\mu)+(\mu_{.j}-\mu)+[(\mu_{ij}-\mu_{i.})-(\mu_{.j}-\mu)]$$

$$Y_{ijk}-\mu = \underbrace{(\mu_{i.}-\mu)+(\mu_{.j}-\mu)}_{\text{effetti additivi}}+\underbrace{[(\mu_{ij}-\mu_{i.})-(\mu_{.j}-\mu)] + \varepsilon_{ijk}}_{\text{interazione}}$$

$$\varepsilon_{ijk} \sim N(0;\sigma^2)$$

Forma equivalente per l'effetto di interazione: $[(\mu_{ij}-\mu_{.j})-(\mu_{i.}-\mu)]$

5. Interpretazione dell'effetto di interazione

	…	j	…	**Marginale**
:	…	…	…	…
i	…	μ_{ij}	…	$\mu_{i\cdot}$
:	…	…	…	…
Marginale	…	$\mu_{\cdot j}$	…	μ

$$(\mu_{ij}-\mu_{i.})-(\mu_{.j}-\mu)$$

se diversa da zero per almeno un *(i,j)* è presente interazione fra fattori

PEARSON

5. Effetto di interazione: esempio numerico

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	media	alta	
2%	17.0	18.0	19.0	18.0
3%	20.0	19.5	18.0	19.2
Marginale	18.5	18.8	18.5	18.6

Esempio per i=1, j=3

$$(\mu_{ij}-\mu_{i.}) - (\mu_{.j}-\mu) = (19-18)-(18.5-18.6) \neq 0$$

PEARSON

Riepilogo degli effetti nel modello completo

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \qquad \varepsilon_{ijk} \sim N(0;\sigma^2)$$

$$\mu_{ij} = \mu + (\mu_{i.} - \mu) + (\mu_{.j} - \mu) + [(\mu_{ij} - \mu_{i.}) - (\mu_{.j} - \mu)]$$

- $(\mu_{i.} - \mu) = 0$ per ogni *i*, non c'è effetto del fattore riga
- $(\mu_{.j} - \mu) = 0$ per ogni *j*, non c'è effetto del fattore colonna
- $[(\mu_{ij} - \mu_{i.}) - (\mu_{.j} - \mu)] = 0$ per ogni *(i,j)*, non c'è interazione

PEARSON

Effetto additivo e di interazione: conseguenze

Solo effetto additivo (**spezzate parallele**): l'effetto di un fattore sulla variabile risposta non dipende dal livello dell'altro fattore.
→ I due fattori possono essere gestiti (es. calibrati, scelti) indipendentemente l'uno dall'altro.

Anche effetto di interazione (**spezzate non parallele**): l'effetto di un fattore dipende dal livello dell'altro fattore.
→ Nel fissare il livello di un fattore devo tener conto anche del livello dell'altro fattore.

Test di ipotesi per studiare gli effetti dei fattori

Test sull'**assenza** di effetto del fattore riga

H$_0$: $(\mu_{i.}-\mu)=0$ per ogni i

H$_1$: $(\mu_{i.}-\mu)\neq 0$ per almeno un i

Test sull'**assenza** di effetto del fattore colonna

H$_0$: $(\mu_{.j}-\mu)=0$ per ogni j

H$_1$: $(\mu_{.j}-\mu)\neq 0$ per almeno un j

Test sull'**assenza** di interazione

H$_0$: $[(\mu_{ij}-\mu_{i.})-(\mu_{.j}-\mu)]=0$ per ogni (i,j)

H$_1$: $[(\mu_{ij}-\mu_{i.})-(\mu_{.j}-\mu)]\neq 0$ per almeno un trattamento (i,j)

La scomposizione della devianza (1)

Quantità sostanza aggiunta	Temperatura			Marginale
	bassa	medie	alta	
2%	$\bar{y}_{11}$	$\bar{y}_{12}$	$\bar{y}_{13}$	$\bar{y}_{1\cdot}$
3%	$\bar{y}_{21}$	$\bar{y}_{22}$	$\bar{y}_{23}$	$\bar{y}_{2\cdot}$
Marginale	$\bar{y}_{\cdot 1}$	$\bar{y}_{\cdot 2}$	$\bar{y}_{\cdot 3}$	$\bar{y}$

Tabella delle medie osservate

Scomposizione da modello

$$\overbrace{(Y_{ijk}-\mu) = (\mu_{i.}-\mu) + (\mu_{.j}-\mu) + [(\mu_{ij}-\mu_{i.})-(\mu_{.j}-\mu)] + \varepsilon_{ijk}}$$

$$\underbrace{(y_{ijk}-\bar{y}) = (\bar{y}_{i}\cdot-\bar{y}) + (\bar{y}\cdot_{j}-\bar{y}) + [(\bar{y}_{ij}-\bar{y}_{i}\cdot)-(\bar{y}\cdot_{j}-\bar{y})] + (y_{ijk}-\bar{y}_{ij})}$$

Scomposizione dei valori osservati

PEARSON

La scomposizione della devianza (2)

i) si elevano al quadrato i due membri

$$(y_{ijk} - \bar{y})^2 = \Big((\bar{y}_{i\cdot} - \bar{y}) + (\bar{y}_{\cdot j} - \bar{y}) + [(\bar{y}_{ij} - \bar{y}_{i\cdot}) - (\bar{y}_{\cdot j} - \bar{y})] + (y_{ijk} - \bar{y}_{ij})\Big)^2$$

ii) si sommano rispetto a i,j,k

$$\sum_{i,j,k}(y_{ijk} - \bar{y})^2 = \sum_{i,j,k}\Big((\bar{y}_{i\cdot} - \bar{y}) + (\bar{y}_{\cdot j} - \bar{y}) + [(\bar{y}_{ij} - \bar{y}_{i\cdot}) - (\bar{y}_{\cdot j} - \bar{y})] + (y_{ijk} - \bar{y}_{ij})\Big)^2$$

iii) per le proprietà delle medie, si ottiene (v. lucido seguente)

La scomposizione della devianza (3)

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y})^2$$

sst: devianza totale

$$=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{i}\cdot-\bar{y})^2$$

ssbr: devianza fra le medie di riga (between)

$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y\cdot_{j}-\bar{y})^2$$

ssbc: devianza fra le medie di colonna (between)

$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left((\bar{y}_{ij}-\bar{y}_{i}\cdot)+(y\cdot_{j}-\bar{y})\right)^2$$

sster: devianza di interazione

$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{ijk}-\bar{y}_{ij})^2$$

ssw: devianza residua (interna o within)

$$sst=\underbrace{(ssbr+ssbc+sster)}_{sstrat}+ssw$$

PEARSON

La scomposizione della devianza (4)
Tavola dell'ANOVA – Modello con interazione

	Devianza	Gradi di libertà	Varianza
Fra medie di riga	$ssbr$	$a-1$	$msbr=ssbr/(a-1)$
Fra medie di colonna	$ssbc$	$b-1$	$msbc=ssbc/(b-1)$
Interazione	$sster$	$(a-1)(b-1)$	$mster=sster/((a-1)(b-1))$
Totale trattamenti	$sstrat$	$ab-1$	$mstrat=sstrat/(ab-1)$
Residua (within)	ssw	$nab-ab$	$msw=ssw/(nab-ab)$
Totale	sst	$nab-1$	

Stima di σ^2

$$sst=\underbrace{(ssbr+ssbc+sster)}_{}+\underbrace{ssw}_{}$$

sstrat: tutti gli effetti dei trattamenti. Contiene eventuali effetti sistematici e (sempre) effetti accidentali (errore sperimentale)

Solo effetti accidentali (errore sperimentale)

Test delle ipotesi (1/4)

A) Test sull'**assenza** di effetto del fattore riga

H_0: $(\mu_{i.}-\mu)=0$ per ogni i

H_1: $(\mu_{i.}-\mu)\neq 0$ per almeno un i

$$F_{oss} = \frac{msbr}{msw} = \frac{ssbr/(a-1)}{ssw/(nab-ab)}$$

È la determinazione di una F di Fisher con $(a\text{-}1),(nab\text{-}ab)$ *gradi di libertà.*

si rifiuta H_0 se $F_{oss} > F_\alpha$ dove $\alpha = P(F_{(a-1),(nab-ab)} > F_\alpha)$

ovvero

si rifiuta H_0 se $p\text{-}level < \alpha$ dove $p\text{-}level = P(F_{(a-1),(nab-ab)} > F_{oss})$

Test delle ipotesi (2/4)

B) Test sull'**assenza** di effetto del fattore colonna

H_0: $(\mu_{.j}-\mu)=0$ per ogni j

H_1: $(\mu_{.j}-\mu)\neq 0$ per almeno un j

$$F_{oss} = \frac{msbc}{msw} = \frac{ssbc/(b-1)}{ssw/(nab-ab)}$$

È la determinazione di una F di Fisher con $(b\text{-}1),(nab\text{-}ab)$ *gradi di libertà.*

si rifiuta H_0 se $F_{oss} > F_\alpha$ dove $\alpha = P(F_{(b-1),(nab-ab)} > F_\alpha)$

ovvero

si rifiuta H_0 se $p\text{-}level < \alpha$ dove $p\text{-}level = P(F_{(b-1),(nab-ab)} > F_{oss})$

Test delle ipotesi (3/4)

C) Test sull'**assenza** di interazione

H_0: $[(\mu_{ij}-\mu_{i\cdot})-(\mu_{\cdot j}-\mu)]=0$ per ogni (i,j)

H_1: $[(\mu_{ij}-\mu_{i\cdot})-(\mu_{\cdot j}-\mu)]\neq 0$ per almeno un trattamento (i,j)

$$F_{oss}=\frac{mster}{msw}=\frac{sster/((a-1)(b-1))}{ssw/(nab-ab)}$$

È la determinazione di una F di Fisher con $(a-1)(b-1),(nab-ab)$ *g.d.l.*

si rifiuta H_0 se $F_{oss}>F_\alpha$ dove $\alpha=P(F_{((a-1)(b-1)),(nab-ab)}>F_\alpha)$

ovvero

si rifiuta H_0 se $p\text{-}level<\alpha$ dove $p\text{-}level=P(F_{((a-1)(b-1)),(nab-ab)}>F_{oss})$

PEARSON

Test delle ipotesi sull'assenza di effetti dei trattamenti (4/4)

H0: assenza di effetti del fattore riga, del fattore colonna e di interazione

$$F_{oss} = \frac{mstrat}{msw} = \frac{sstrat/(ab-1)(b-1)}{ssw/(nab-ab)}$$

Esempio numerico

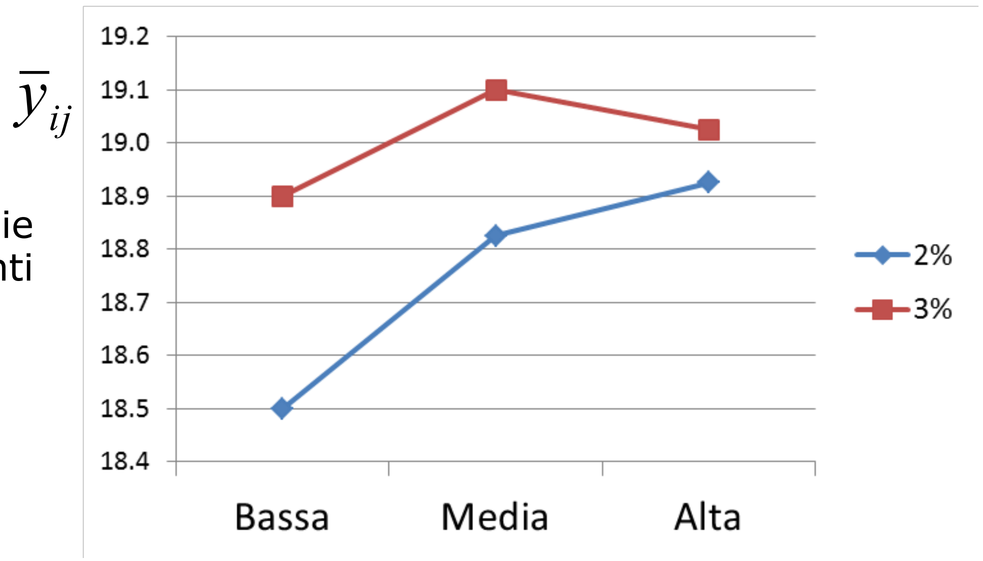
Dati

Fattore A sostanza aggiunta (%)	Fattore B (temperatura)		
	Bassa	Media	Alta
	18.4	18.8	19.0
2%	18.5	18.7	18.9
	18.5	18.9	18.9
	18.6	18.9	18.9
	18.8	19.0	19.1
3%	18.9	19.1	19.2
	19.0	19.2	19.0
	18.9	19.1	18.8

Tabella delle medie osservate

Fattore A sostanza aggiunta (%)	Fattore B (temperatura)			Marginale
	Bassa	Media	Alta	
2%	18.500	18.825	18.925	18.750
3%	18.900	19.100	19.025	19.008
Marginale	18.700	18.963	18.975	18.879

Grafico delle medie osservate dei trattamenti



PEARSON

Tavola dell'ANOVA (effetti principali e interazione)

Effetti	Devianza	gdl	Varianza	F oss.	*p-value*
Fattore riga (% sostanza aggiunta)	0.4004	1	0.40042	39.493	0.000006
Fattore colonna (temperatura)	0.3858	2	0.19292	19.027	0.000036
Interazione fra fattori	0.0908	2	0.04542	4.479	0.026371
Totale trattamenti	0.8771	5	0.17542	17.301	0.000003
Devianza interna	0.1825	18	0.01014		
Devianza totale	1.0596	23			

Stima di σ^2

Posto α=0.05, si conclude che è presente sia l'effetto del fattore riga, sia quello del fattore colonna sia l'effetto di interazione.

PEARSON

Modello a soli effetti additivi: scomposizione della devianza

	Devianza	Gradi di libertà	Varianza
Fra medie di riga	$ssbr$	$a-1$	$ssr/(a-1)$
Fra medie di colonna	$ssbc$	$b-1$	$ssc/(b-1)$
Residua	$ssres$	$nab-a-b+1$	$ssres/(nab-a-b+1)$
Totale	sst	$nab-1$	$sst/(nab-1)$

$$sst = \underbrace{(ssbr + ssbc)}_{} + \underbrace{ssres}_{}$$

Eventuali effetti sistematici e (sempre) effetti accidentali (errore sperimentale)

$ssres = sster + ssw$

Solo effetti accidentali (errore sperimentale)

PEARSON

Che cosa abbiamo imparato

- Che cosa si intende per trattamento con 2 fattori sperimentali
- Gli effetti principali dei fattori (effetti additivi)
- L'effetto di interazione
- Specificazione del modello con effetti additivi ed interazione e verifica degli effetti
- Grafico per indagare sulla presenza di interazione
- Tavola dell'ANOVA