

## TECNICHE DI STIMA - TEST delle Ipotesi

- determinare le ipotesi di lavoro 1
- verificare le ipotesi attraverso procedimenti sperimentali 2
- analizzare il rischio che si ha nel tenere valide le ipotesi determinate (analisi del rischio che deriva dalla validità delle ipotesi) 3

Noi continuiamo dei campioni di prodotto su cui verifichiamo l'ipotesi che le nostre caratteristiche siano migliori di quelle proposte dal mercato. Le ipotesi si riferiscono alla progettazione delle caratteristiche qualitative del nostro prodotto (la nostra è migliore delle preesistenti)

Verifico la validità attraverso l'analisi campionaria  $n \rightarrow \bar{x}, S$   
(la popolazione prodotta è rappresentata) dal valore di benchmark con cui mi confronto

ANALISI del RISCHIO  $\rightarrow$  gli stimatori servono a collegare la percezione di validità a un rischio

1 In ambito industriale si hanno due forme per esprimere ipotesi:

$(H_0)$  IPOTESI NULLA: non vale niente perché è espressa attraverso un'uguaglianza  $\mu = \mu_0$

$\rightarrow$  dal campione che estraggo la tendenza centrale della mia produzione ( $\mu$ ) è pari a un valore di riferimento ( $\mu_0$ )

Questa ipotesi comporta 2 tipi di errore:

- ERRORE di TIPO PRIMO: corrisponde alla probabilità che si ha nel momento in cui rifiuto  $H_0$  quando questa ipotesi è vera  
 $\rightarrow$  livello di rischio  $(\alpha) \rightarrow$  rifiuto di  $H_0$  che in realtà è vera
- ERRORE di TIPO SECONDO: corrisponde alla probabilità che si ha nel momento in cui accetto  $H_0$  quando però questa è falsa  
 $\rightarrow$  livello di rischio  $(\beta)$

$\alpha$  e  $\beta$  sono due probabilità che misurano il rischio:

- rifiuto  $H_0$  ma hai  $\alpha$  probabilità che sia vera  $\rightarrow$  POTENZA del TEST

- accetti  $H_0$  ma hai  $\beta$  probabilità che sia falsa  $\rightarrow$  SIGNIFICATIVITÀ

Se l'ipotesi nulla è a me favorevole l'errore peggiore sarà quello di tipo 2

Si rifiuta  $H_0$  in favore di ipotesi alternative  $(H_a)$  che sono di 3 tipi:

$[H_{a1}] \rightarrow \mu \neq \mu_0$  TEST BIDIREZIONALE: non mi interessa se è  $>$  o  $<$

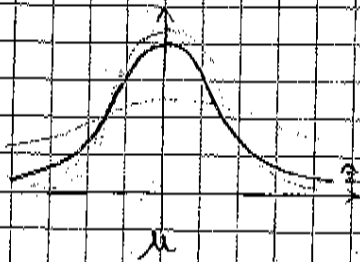
$[H_{a2}], [H_{a3}] \rightarrow \mu > \mu_0, \mu < \mu_0$  TEST UNIDIREZIONALE rifiuto  $H_0$  a favore di  $\mu > \mu_0$  o di  $\mu < \mu_0$

2 Quando devo verificare l'ipotesi sperimentalmente mi trovo davanti a una variabile  $y(\mu, \sigma)$  che non conosco

$\downarrow$   
prendo un campione  $(y_1, y_2, \dots, y_n)$  e calcolo  $\bar{y}$  e  $S$

$\rightarrow$  se voglio costruire la distribuzione  $t$  allora  $T = \frac{\bar{y} - \mu}{\frac{S}{\sqrt{n}}}$

per il teorema del limite centrale avrò:



$$z(y) = N(\mu, \sigma)$$

↳ parametri

alla variazione di quel parametro tendiamo a stringerli attorno a  $\mu$

Più  $n$  aumenta, più questa distribuzione  $t$  tende a  $z$ , che è la distribuzione normale della popolazione ( $n: t \rightarrow z$ )

GRADI DI LIBERTÀ ( $V$ )

$$V = n - 1$$

↳ all'aumentare del grado di libertà,  $t$  tende a  $z$

3. Supponiamo di conoscere la distribuzione  $t$  in base ai gradi di libertà ( $t$  tab), se il valore sperimentale, si allontana dalla tendenza centrale statisticamente siamo portati a rifiutare  $H_0$  a favore di un'alternativa ( $H_{a1}$  o  $H_{a2}$  o  $H_{a3}$ ) sulla base di un rischio ( $\alpha$ ) che è pari all'area che ho sulle code.

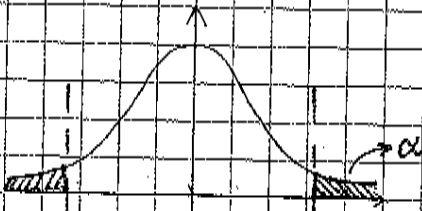


TABELLE  
( $t_{\alpha, V}$ )

Ci sono delle tabelle che in base ai gradi di libertà e del livello di rischio che voglio prendere quando rifiuto  $H_0$  mi determinano la REGIONE CRITICA

Se sperimentalmente vado a finire nella regione critica allora vuol dire che sono andata molto lontano da  $\mu$  e che sono portata a rifiutare  $H_0$   
↳ più ci allontaniamo da  $\mu$  più il rischio diminuisce

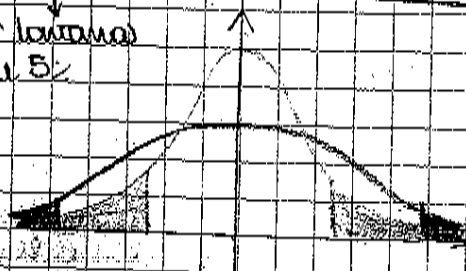
$$t^* = \frac{y - \mu_0}{\frac{s}{\sqrt{n}}}$$

↳ se  $t^*$  è nella regione critica siamo portati a rifiutare

↳ se il rischio aumenta, la regione critica si stringe

EX  $t_{5\%, 10}$   
↳ più lontano dal 5%

$t_{5\%, 100}$  ↳ più spaziosa



Supponiamo di voler verificare  $H_{a1} \Rightarrow \mu \neq \mu_0$

qual è il valore di  $z$  che lascia alle code un valore del 5%?

$$z_{5\%} = 1.96$$

$$\mu \neq 30 \quad 99.73\%$$

$$\mu \neq 1.96\sigma \quad 95\%$$

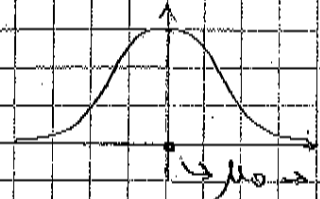
RISPOSTA:  $t^*$  non è nella regione critica quindi non possiamo dire nulla ma possiamo scegliere un livello di rischio tale per cui sono più probabili l'ipotesi critica a favore di un'altrea.

In sede di design review il livello di rischio che doveva essere standard è del 25% perché è il livello più vicino per cui  $t^*$  si trova nella regione critica e la produzione sarà dunque più performante di 25m.

(passo test delle ipotesi)

9.10.11

Il TEST delle IPOTESI serve a verificare qual è il livello di rischio quando mi confronto con una determinata prestazione o parametro fimo

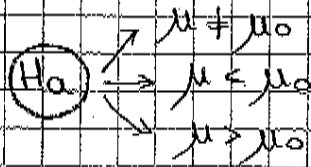


Ci confrontiamo con un valore di riferimento  $\mu_0$ , attraverso le statistiche a nostra disposizione  $\bar{x}$  e  $S$

1. definisco le ipotesi:  $(H_0) \Rightarrow \mu = \mu_0$

2. effettuo analisi sperimentale e calcolo  $\bar{x}$  e  $S$

→ se la distanza dei valori da  $\mu$  è grande, posso pensare di rifiutare questa ipotesi a favore di un'alternativa (sono talmente lontano da pensare che con un certo rischio sono optare per un'ipotesi alternativa.)



Mi costruisco una  $T^* = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$

$y \sim N(\mu, \sigma^2)$        $z = \frac{y - \mu}{\sigma} \rightarrow$  cambio di variabile

$X \sim N(0, 1)$       [ $\mu_0 - \mu = 0$ ]

di questa si conosce tutto, se  $t^*$  è vicino non scarto  $H_0$ , se  $t^*$  è molto lontano dal riferimento allora si può dire che scarto  $H_0$  a favore di un'ipotesi alternativa

$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$

da una misura più corretta rispetto a  $\frac{1}{n}$   
 GdL (Gradi di libertà)  
 $\downarrow$   
 $= n - 1$

varianza VAR       $\left[ \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right] \rightarrow$  devianza DEV       $VAR = \frac{DEV}{GdL}$

Conosco tutto anche di una serie di distribuzioni t che variano al secondo dei gradi di libertà ( $\downarrow$ ):

$$t = \frac{\mu - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

più aumentiamo i gradi di libertà (più le curve si avvicinano alla distribuzione normale)

$$n \rightarrow N \rightarrow N(0, 1)$$

- ↳ bassi  $\rightarrow$  curve spaziate
- ↳ alti  $\rightarrow$  sempre più vicino alla normale standardizzata

$\alpha$  è il rischio che devo sopportare nel momento in cui rifiuto  $H_0$

$$\left[ \begin{array}{c} t^{cr} \\ -t^{cr} \end{array} \right] \rightarrow \text{delimita la REGIONE CRITICA}$$

per rifiutare  $H_0$  devo ragionare sul fatto che  $t^*$  debba finire il più lontano possibile, sulla coda della regione critica

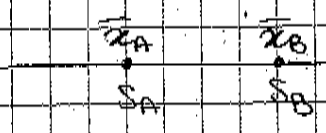
- se  $t^* > t^{cr}$ : sono portato a rifiutare  $H_0$  in favore di  $H_1$  ( $\mu > \mu_0$ )
- se  $t^* < -t^{cr}$ : sono portato a rifiutare  $H_0$  in favore di  $H_1$  ( $\mu < \mu_0$ )
- se  $t^* \notin$  regione critica: statisticamente non posso dire nulla, ma sapendo che  $t^{cr}$  dipende da  $\nu$  e da  $\alpha$  posso vedere se, facendo variare  $\alpha$  ( $\nu$  è costante) riesco a far cadere  $t^*$  nella regione critica  $\rightarrow$  devo aumentare  $\alpha$  per poter dire qualcosa statisticamente

### ANOVA - Analysis Of Variance

L'ANOVA è la prosecuzione della tecnica del test delle ipotesi ed è dedicata al confronto fra due o più opzioni tecnologiche, rispetto alla prestazione in esame

Ex: opzioni tecnologiche  $\rightarrow$  benzina A, B o C  
prestazioni  $\rightarrow$  km/l

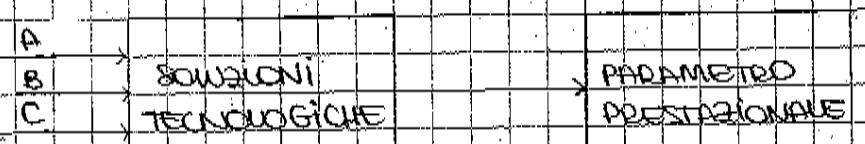
L'analisi delle varianze mi consente di prevedere quale opzione tecnologica massimizza la prestazione oggetto di studio. Per scegliere l'opzione migliore NON si usa la  $\mu$  perché, sia campionando: la media campionaria cade nell'intervallo della media della popolazione ma si distribuisce secondo una normale.



devo confrontarli per capire quale mi porterebbe una prestazione migliore

$\downarrow$   
maggiore complessità

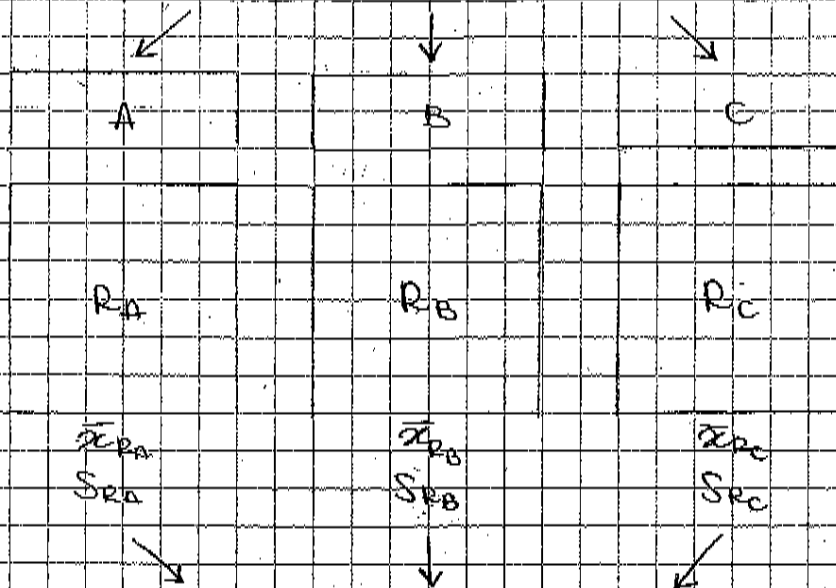
Si confrontano  $\bar{x}_A, s_A$  e  $\bar{x}_B, s_B \rightarrow$  due analisi campionarie



più opzioni tecnologiche (trattamenti o livelli)  $\rightarrow \bar{x}_A, s_A, \bar{x}_B, s_B, \bar{x}_C, s_C$

L'ANOVA risponde alla domanda: "quale delle opzioni tecnologiche impatta maggiormente (= più favorevolmente) sul parametro prestazionale?"

FRETTAZIONE DA MISURARE



POPOLAZIONI  
 $U = 1, 2, \dots, p$

PER OGNI OPZIONE  
 HO  $n$  REPLICAZIONI  
 $D_i = 1, 2, \dots, n$

da tutte le repliche POSSO AVERE  
 $\bar{x}$  (la media delle medie)

L'ipotesi nulla ( $H_0$ ) è:  $\mu_A = \mu_B = \mu_C$

( $H_a$ ) è solo una:  $\mu_A \neq \mu_B \neq \mu_C$  → ANOVA non consente altre ipotesi alternative, ( $> <$ )

se  $\bar{x}_{R_A} > \bar{x}_{R_B}$ : l'opzione A ha un impatto sul parametro prestazionale maggiore di quello dell'opzione B

stimatore di  $\mu_A$

VARIANZA = DEVIANZA / GRADI DI LIBERTA'

Questa tecnica prevede 3 tipologie di VARIANZA:

TOTALE ( $VAR_{TOT}$ ): mette in relazione tutti gli  $x_{ij}$  rispetto alla media delle medie  $\bar{x}$

TRA ( $VAR_{TRA}$ ): tra le opzioni; mette in relazione le medie  $\bar{x}_j$  e la media delle medie  $\bar{x}$

ENTRO ( $VAR_{ENTRO}$ ): mette in relazione tutti gli  $x_{ij}$  rispetto alle medie delle opzioni  $\bar{x}_j$

$DEV_{ENTRO} = DEV_{TOT} - DEV_{TRA}$

$DEV_{TOT} = DEV_{ENTRO} + DEV_{TRA}$

DEVIANZA

TOTALE ( $DEV_{TOT}$ ): è, in tutto, numero di repliche - 1 ( $GL = n - 1$ )

$DEV_{TOT} = \sum_{j=1}^p \sum_{i=1}^n (x_{ij} - \bar{x})^2$

TRA (DEV<sub>TRA</sub>) →  $\Gamma_{Gall} = p-1$

$$DEV_{TRA} = \sum_{j=1}^p (n_j) (\bar{x}_j - \bar{x})^2$$

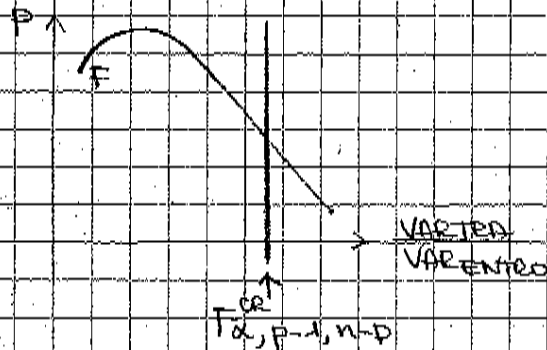
non è detto che per ogni opzione si abbia lo stesso numero di repliche

ENTRO (DEV<sub>ENTRO</sub>) →  $\Gamma_{Gall} = n-p$

$$DEV_{ENTRO} = \sum_{j=1}^p \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$VAR_{TRA} = \frac{DEV_{TRA}}{p-1}$$

$$VAR_{ENTRO} = \frac{DEV_{ENTRO}}{n-p}$$



Ho una DISTRIBUZIONE di FISHER (F) che descrive la probabilità  $P$  che ho nell'acceptare  $H_0$ , più il valore in fondo più posso pensare di scartare  $H_0$  in favore di  $H_a$ , più avendo il mio rischio  $\alpha$

$$F_{\alpha, p-1, n-p}$$

→ il valore critico si trova con le tabelle

posso confrontare il mio valore critico di  $F$  con il mio  $F^*$  determinato sperimentalmente:

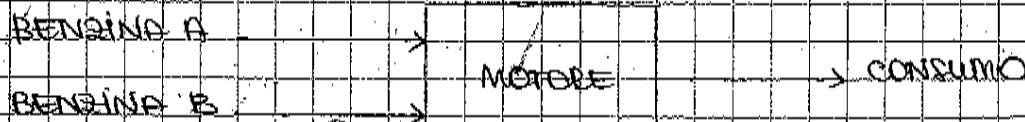
$$F^* = \frac{VAR_{TRA}}{VAR_{ENTRO}}$$

se  $F^*$  è nella regione critica allora statisticamente dico che con questo rischio  $\alpha$  rifiuto  $H_0$  in favore di  $H_a$ .

Se questo non succede, l'unico parametro che posso modificare è  $\alpha$ :

→ più  $\alpha$  aumenta, più il valore di  $F_{critico}$  diminuisce

### ESERCIZIO ANOVA



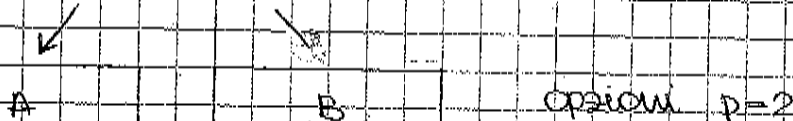
OBBIETTIVO: Calcolare  $F^* = \frac{VAR_{TRA}}{VAR_{ENTRO}}$

$$H_0 \Rightarrow \mu_A = \mu_B$$

$$H_a \Rightarrow \mu_A \neq \mu_B$$

$$\alpha = 5\%$$

CONSUMO



	A	B	OPZIONI p=2
$R_A$	20 23 23	16 17 24	$R_B$

REPPLICAZIONI n=6

$\bar{x}_{RA} = 22$        $\bar{x}_{RB} = 19$

$\bar{x} = \frac{22+19}{2} = 20,5$

$\rightarrow VAR_{TOT} = \frac{DEV_{TOT}}{n-1} = \frac{DEV_{TOT}}{5}$

$DEV_{TOT} = \sum_{j=1}^p \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = (20 - 20,5)^2 + (23 - 20,5)^2 + (23 - 20,5)^2 + (16 - 20,5)^2 + (17 - 20,5)^2 + (24 - 20,5)^2 = 57,5$

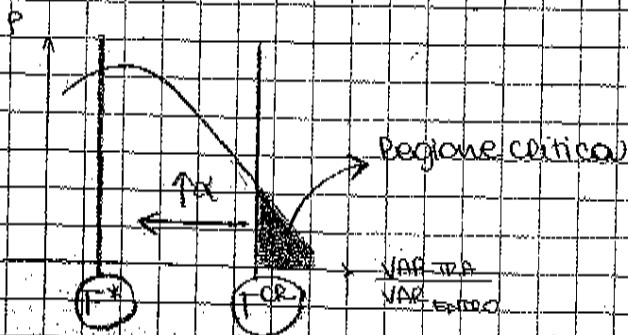
$\rightarrow VAR_{TRA} = \frac{DEV_{TRA}}{p-1} = \frac{DEV_{TRA}}{1}$

$DEV_{TRA} = \sum_{j=1}^p n_j (\bar{x}_j - \bar{x})^2 = 3(22 - 20,5)^2 + 3(19 - 20,5)^2 = 13,5$

$\rightarrow DEV_{ENTRO} = DEV_{TOT} - DEV_{TRA} = 57,5 - 13,5 = 44$

$F^*$	$\frac{DEV_{TRA}}{GdL} = \frac{13,5}{1}$	$\frac{13,5}{11} = 1,227$
	$\frac{DEV_{ENTRO}}{GdL} = \frac{44}{4}$	

$F_{52, 1, 4}^{CR} = 7,709 \rightarrow F_{critica}$



$F^*$  non cade nella regione critica, non posso far nulla se non determinare quell'  $F_{crit}$  tale per cui  $F^*$  cade nella regione critica

↓  
determinare un nuovo livello di rischio  $\alpha$

Se posso dire che  $H_0$  è rifiutabile:  $\mu_A \neq \mu_B \rightarrow$  scelgo l'opzione con media maggiore:

$\bar{x}_{RA} > \bar{x}_{RB}$

# ESERCITAZIONE ANOVA

$$\alpha = 1\%$$

DUREZZA

A	B
2,7	2,2
2,8	2,1
2,9	2,2
2,5	2,3
2,6	2,1
2,7	2,2
2,8	2,3
	2,6

$$\bar{x}_{PA} = 2,714$$

$$\bar{x}_{PB} = 2,25$$

$$\bar{x} = 2,48$$

$$\rightarrow DEV_{TOT} = \sum_{j=1}^p \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = 1,096$$

$$\rightarrow DEV_{TRA} = \sum_{j=1}^p n_j (\bar{x}_j - \bar{x})^2 = 7(2,714 - 2,48)^2 + 8(2,25 - 2,48)^2 = 0,806$$

$$\rightarrow DEV_{entro} = 1,096 - 0,806 = 0,29$$

$$F^* = \frac{0,806}{\frac{1}{13}} = 36,13$$

$$F_{1\%, 1, 13}^{CR} = 9,074$$

$$F^* > F^{CR}$$

$F^*$  cade nella regione critica, rifiutiamo  $H_0$ .  $\bar{x}_{PA} > \bar{x}_{PB}$ : l'opzione A è la migliore

Il fatto che  $F^*$  si sposti lontano dipende da quanto sono lontano le medie dalla media delle medie ( $VAR_{TRA}$ )



## Esercitazione ANOVA (ANALYSIS OF VARIANCE)

Si stanno provando due nuove resine (nel seguito resina A e resina B) per la realizzazione del guscio di una nuova pallina da golf con prestazioni migliori rispetto a quelle già in commercio. Scopo dei test è verificare le proprietà di durezza fornite al guscio della pallina dalle due diverse resine. Cioè capire qual è la miglior resina secondo il criterio: la migliore è quella che dà alla pallina una maggiore resistenza.

Poiché le prove sono di carattere distruttivo in quanto il test di carico si basa sulla determinazione della pressione di rottura del guscio, il team di progetto ha commissionato al laboratorio la realizzazione di 8 esemplari di pallina con la resina A e 8 con la resina B.

La quantità di resina di tipo A a disposizione del laboratorio ha consentito la realizzazione di solo 7 esemplari di palline anziché di 8.

Tutte le palline sono state sottoposte al test di carico.

La seguente tabella riporta i risultati del test.

Test di carico	Resina A Durezza (x10 bar)	Resina B Durezza (x10 bar)
1	2,7	2,2
2	2,8	2,1
3	2,9	2,2
4	2,5	2,3
5	2,6	2,1
6	2,7	2,2
7	2,8	2,3
8	---	2,6

Il team di lavoro ha avuto l'incarico di verificare le prestazioni delle due resine assumendo un livello di rischio pari all'1%. Livelli di rischio maggiori dovranno essere validati in sede Design Review.

Effettuare l'Analisi del Rischio sulla base dei risultati ottenuti.

