

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA
SCUOLA DI ECONOMIA E STATISTICA

CORSO DI LAUREA IN
SCIENZE STATISTICHE ED ECONOMICHE



ANALISI DELL'ACCURATEZZA DI UNO STRUMENTO DI MISURA PER SAATI SPA

RELATORE: Dott. Tommaso Rigon

TESI DI LAUREA DI:
Alessandro Colello
MATRICOLA N. 851334

ANNO ACCADEMICO 2021/2022

A mia nonna Angela

Indice

1	Introduzione	1
1.1	I dati	2
1.2	Sintesi e descrizione dei dati	3
1.3	Inferenza su medie e varianze	5
2	Indicatore per la valutazione dell'accuratezza degli strumenti	11
2.1	Definizione dell'indicatore	12
2.2	Stima di massima verosimiglianza dell'indicatore	13
2.3	Stimatore MV e distribuzione asintotica	14
3	Inferenza sull'indicatore	17
3.1	Intervallo di confidenza alla Wald	17
3.2	Intervallo di confidenza basato sul rapporto di verosimiglianza . .	19
3.3	Test di Wald	19
3.3.1	Test di Wald unilaterale	21
3.4	Test del rapporto di verosimiglianza	22
3.5	Applicazione ai dati	23
4	Individuazione delle cause d'inaccuratezza delle misure	27
4.1	Il dataset	27
4.2	Descrizione generale della procedura di modellizzazione	28
4.3	Modello lineare per gli scarti dello strumento digitale con ricetta E	31
4.3.1	Specificazione e stima del modello d'interesse	32
4.3.2	Trattamento dell'eteroschedasticità	32
4.3.3	Stima del modello trasformato	33
5	Conclusioni	35
	Bibliografia	37

Capitolo 1

Introduzione

Saati SpA (Figura 1.1) è un'azienda multinazionale che opera nei settori della serigrafia, della filtrazione e della protezione balistica sviluppando, producendo e commercializzando prodotti chimici per uso industriale e tessuti tecnici avanzati e personalizzati. A maggio, ho avuto la possibilità di fare uno stage di 3 mesi presso lo stabilimento produttivo di Veniano. Tale *plant* è interamente dedicato alla filtrazione ed è diviso in area medica, alimentare e industriale a seconda del settore in cui le soluzioni prodotte sono impiegate. Nell'area industriale vengono prodotte strisce di tessuto, dette nastri, destinate all'industria automobilistica, alimentare, dell'acustica e degli elettrodomestici. Dai nastri ottenuti tramite taglio a caldo da una pezza, ovvero una striscia di tessuto tolta dal telaio senza orli o finiture e avvolta attorno ad una struttura cilindrica, vengono tagliate le estremità di inizio e di fine rotolo, dette tirelle. Esse vengono quindi misurate dagli operatori tramite un righello di precisione dotato di lenti d'ingrandimento che permettono di raggiungere una risoluzione di 0.1 mm. Successivamente, viene effettuato un secondo controllo della correttezza delle altezze dei nastri. Il laboratorio di controllo qualità, infatti, misura ogni tirella con un proiettore di profili con risoluzione di 0.001 mm. Per evitare possibili errori di misura tramite il righello, è stato introdotto uno strumento digitale per la misura dell'altezza dei nastri. Questo macchinario è dotato di 11 programmi, detti ricette, ognuno caratterizzato da parametri opportunamente impostati per diverse categorie di



Figura 1.1

tessuti. Per effettuare una nuova misurazione, bisogna selezionare la ricetta corretta. Successivamente, si ripone la tirella su un piano e la si schiaccia con una lastra di plexiglass. Facendo passare una fonte di luce attraverso il tessuto, il macchinario, grazie alla fotocamera da 20 MP di cui è dotato, rileva i bordi del nastro. Contando il numero di pixel che compongono l'immagine da un lato all'altro della tirella e moltiplicandolo per la dimensione del pixel ($5\text{ }\mu\text{m}$), si ottiene una prima misura dell'altezza del nastro. Ripetendo tale operazione in diversi punti del tessuto e calcolando la media delle misure rilevate, si ottiene la misura che viene poi mostrata a video. Di questa misura vengono considerate significative le prime due cifre decimali dopo il millimetro.

Il primo obiettivo di quest'elaborato è quello di fornire una procedura oggettiva per la valutazione dell'accuratezza dello strumento digitale. Con accuratezza si intende la capacità di uno strumento di fornire misure esatte, ovvero in linea con i valori ritenuti veri, e precise, ovvero coerenti con misurazioni successive. L'altezza effettiva di una tirella è assunta essere pari alla misura ottenuta tramite il proiettore di profili. Se si dovesse stabilire che lo strumento digitale è più accurato del righello di precisione, lo si potrebbe riprodurre per renderlo disponibile in più centri di lavoro. Inoltre, si potrebbe eliminare la fase di ricontrollo delle altezze dei nastri effettuata dal laboratorio di controllo qualità, riducendo il suo carico di lavoro. Il secondo obiettivo, invece, consiste nel determinare una procedura per l'identificazione di eventuali cause d'inesattezza o d'imprecisione delle misure dello strumento digitale al fine di aiutare i suoi produttori a renderlo più accurato.

1.1 I dati

I dati raccolti provengono da 28 ordini prodotti da fine marzo a fine luglio 2022. Le misure delle tirelle considerate sono state ottenute, dagli operatori, tramite il righello di precisione e lo strumento digitale e, dal laboratorio di controllo qualità, tramite il proiettore di profili. Il numero di tirelle varia da un minimo di 3 a un massimo di 22 per ogni ordine. Le misure prese dagli operatori sono state recuperate dai cartacei degli ordini di produzione mentre quelle prese dal laboratorio di controllo qualità da SAP, il sistema ERP aziendale. Sono state considerate 572 misure prese con il righello di precisione. Invece, le misure prese con lo strumento digitale sono in totale 372 e sono state ottenute utilizzando le sei ricette di uso più frequente, indicate con le lettere dalla A alla F, le cui

numerosità campionarie sono riportate nella Tabella 1.1. Dato che tirelle di

	Strumento digitale						Righello
	A	B	C	D	E	F	
Numerosità	18	50	95	13	138	28	578

Tabella 1.1: Numerosità campionarie delle misure degli strumenti

diversi ordini hanno potenzialmente altezze diverse, si è deciso di analizzare gli scostamenti delle misure dello strumento digitale e del righello di precisione dalle misure del proiettore di profili. Gli scarti così ottenuti rappresentano campioni tra loro indipendenti. Nonostante i campioni siano ottenuti osservando in parte le medesime unità statistiche (le tirelle), tra i campioni cambia la natura del fenomeno d'interesse (lo scarto). Gli scarti, infatti, sono influenzati dalle misure dello strumento digitale oppure da quelle del righello a seconda del campione. Conoscere la distribuzione degli scarti dello strumento digitale con una particolare ricetta non dà informazioni sulla distribuzione degli scarti del righello e viceversa. Per esempio, alti scarti dello strumento digitale su una specifica tirella non implicano a priori alti scarti del righello sulla stessa tirella. Queste considerazioni sono utili per giustificare l'indipendenza dei campioni, assunzione fondamentale per il prosieguo.

1.2 Sintesi e descrizione dei dati

Lo strumento ideale è caratterizzato dall'esattezza e dalla precisione. Mentre l'esattezza fa sì che le differenze tra le misure e i valori assunti veri siano nulli in media, la precisione fa sì che la varianza delle misure sia minima. La Tabella 1.2 riporta alcune statistiche di sintesi degli scostamenti delle misure degli strumenti dalle altezze reali dei nastri. Lo strumento digitale è apparentemente accurato quando si utilizzano le ricette A, C e D e solamente esatto quando si utilizza la ricetta E. Utilizzando le ricette B ed F, si ottengono, in media, misure sovrastimate rispettivamente di 0.1086 e 0.18 mm con deviazioni standard relativamente elevate, rispettivamente di 0.168 e 0.179 mm. Inoltre, i valori positivi del primo quartile delle distribuzioni degli scarti ricavati utilizzando queste due ricette sono indicativi del fatto che almeno nel 75% dei casi lo strumento digitale abbia sovrastimato le misure. Utilizzando il righello di precisione si ottengono, in media, misure non distorte ma con una deviazione standard relativamente elevata:

0.1668 mm. Infine, si nota che le misure più imprecise nei casi di sottostima e di sovrastima sono state entrambe rilevate con il righello: -0.621 mm e 0.9 mm.

	Strumento digitale						Righello
	A	B	C	D	E	F	
Minimo	-0.148	-0.31	-0.106	-0.072	-0.4266	-0.327	-0.621
1° quartile	-0.0618	0.009	-0.0235	-0.007	-0.0903	0.1544	-0.0623
Media	-0.0369	0.1086	0.0373	-0.0003	-0.0097	0.1804	0.0285
Mediana	-0.0355	0.1355	0.029	0.005	0.0033	0.2275	0.0115
3° quartile	-0.0155	0.1938	0.0825	0.018	0.0913	0.2715	0.1018
Massimo	0.125	0.63	0.308	0.058	0.466	0.58	0.9
Dev. std	0.0612	0.168	0.0774	0.0366	0.1637	0.179	0.1668

Tabella 1.2: Sintesi degli scarti dello strumento digitale e del righello di precisione (mm)

Approfondendo l'esplorazione dei dati, si visualizzano le distribuzioni di densità degli scarti degli strumenti tramite istogrammi e boxplot (Figura 1.2). Vengono anche riportati i valori osservati w_{oss} della statistica test e i p-value dei test di Shapiro-Wilk per la normalità degli scarti. Per la costruzione del test di Shapiro-Wilk si rimanda a [Shapiro & Wilk \(1965\)](#). Fissando il livello di significatività al 5%, c'è evidenza empirica che sostiene il rifiuto dell'ipotesi di normalità degli scarti dello strumento digitale con ricette C, E ed F e del righello di precisione. D'altra parte, gli scarti dello strumento con ricette A, B e D possono essere considerati normalmente distribuiti secondo il test di Shapiro-Wilk. Concentrando l'attenzione sugli istogrammi relativi agli scarti per i quali si è rifiutata l'ipotesi di normalità, si nota l'unimodalità delle distribuzioni di densità. Si visualizzano quindi i grafici delle probabilità normali (o *normal Q-Q plot*) degli scarti dello strumento digitale con ricette A, E ed F e del righello di precisione (Figura 1.4). Nelle Figure 1.4b, 1.4c e 1.4d si vede che le code delle distribuzioni di densità sono più pesanti di quelle della distribuzione normale. Nella Figura 1.4a si vede che solo la coda di destra è più pesante di quella della distribuzione normale. La coda sinistra è invece più leggera. In tutti i grafici, però, i quantili empirici centrali tendono a distribuirsi lungo la retta che collega i punti dei primi e dei terzi quartili. Le distribuzioni degli scarti degli strumenti rispecchiano quindi la distribuzione normale almeno tra il primo e il terzo quartile. Nonostante l'assunzione di normalità degli scarti dello strumento digitale con ricette A, E ed F e del righello di precisione sia errata, si procede come se fosse valida. Ciò semplifica l'analisi e non dovrebbe avere un forte impatto sui risultati finali per

le considerazioni fatte. Infine, guardando i boxplot, si nota la presenza di outlier in tutti i grafici. Si decide di non rimuovere questi valori dall'analisi per non perdere l'informazione che essi apportano.

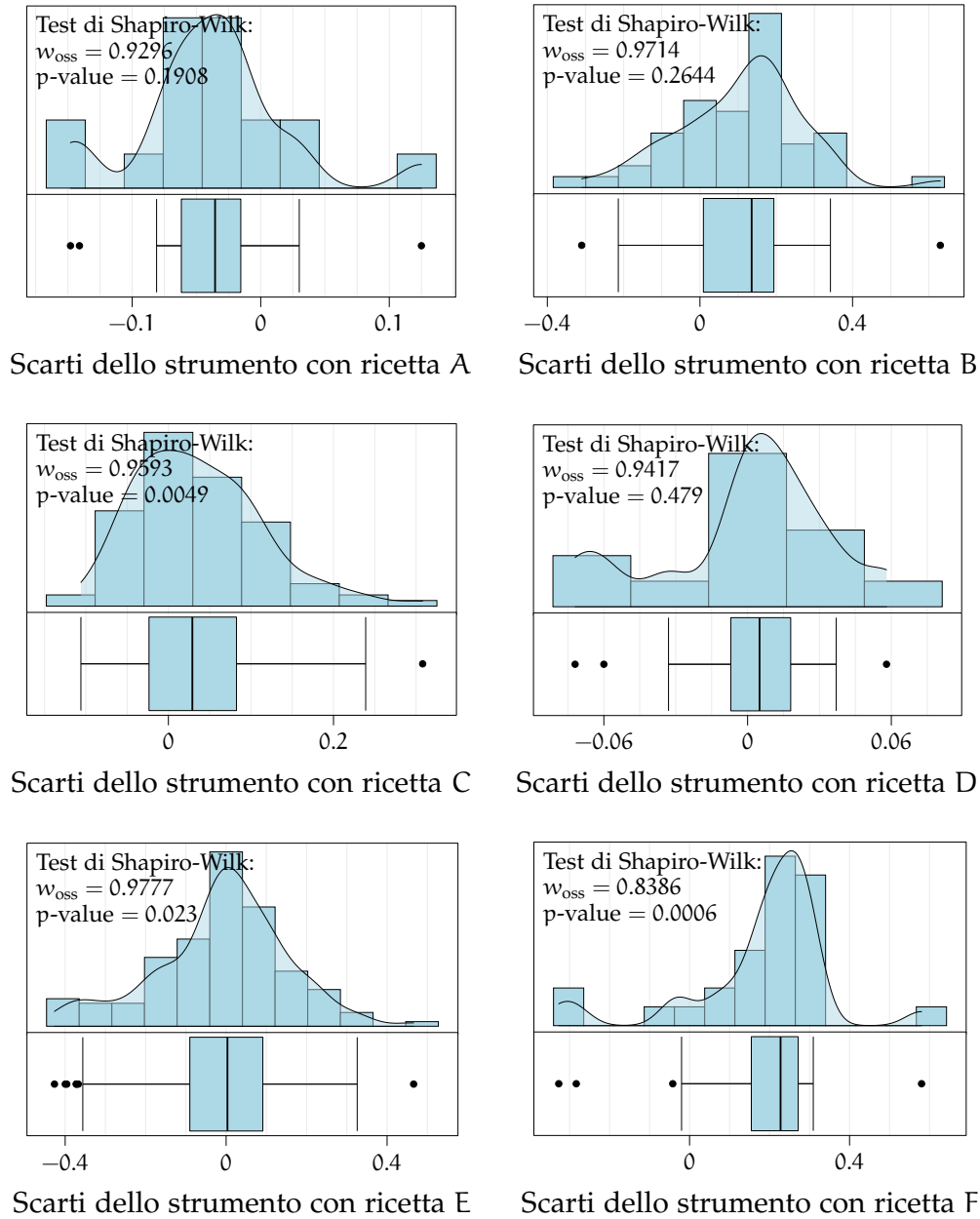


Figura 1.2: Istogrammi, boxplot e test di Shapiro–Wilk per le sei ricette

1.3 Inferenza su medie e varianze

In questa sezione vengono fornite le stime puntuali e intervallari dei parametri essenziali per la valutazione dell'esattezza e della precisione di uno strumento di misura: rispettivamente la media, indicata con μ_j , e la varianza, indicata con σ_j^2 .

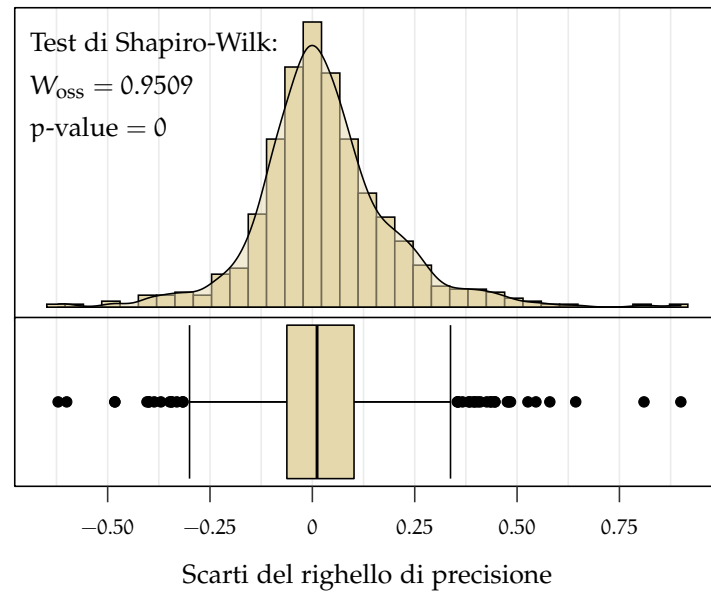


Figura 1.3: Istogramma, boxplot e test di Shapiro-Wilk per gli scarti del righello

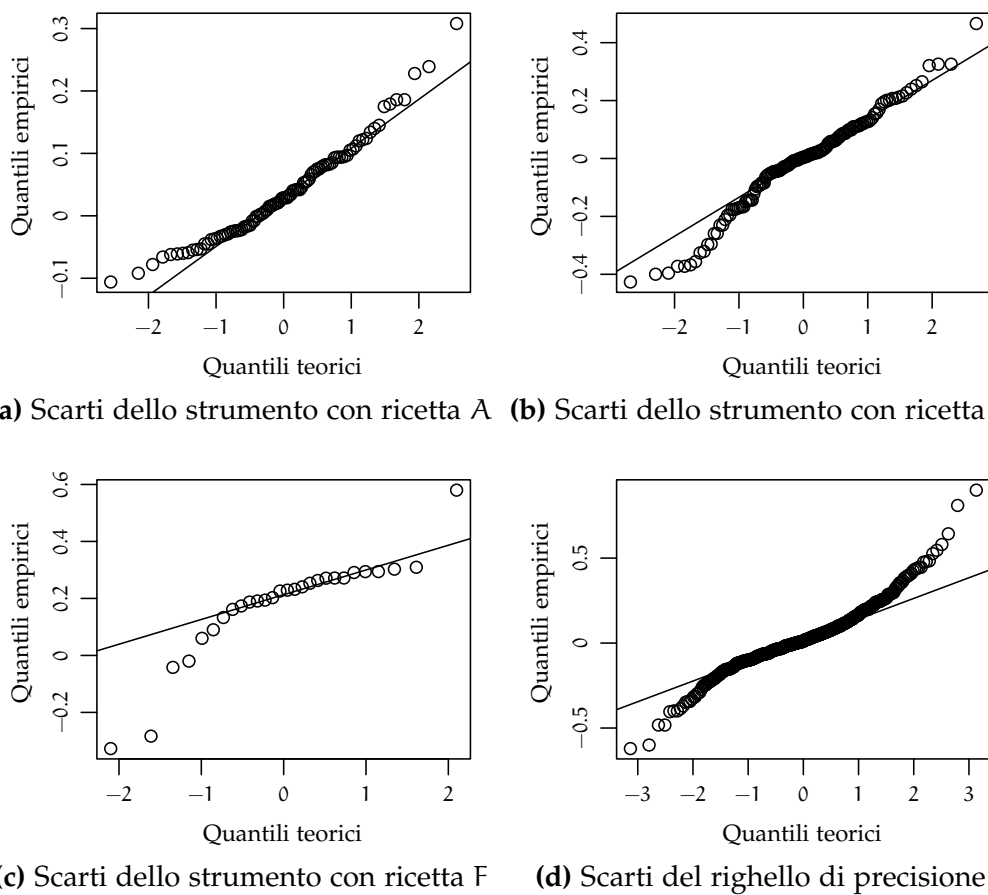


Figura 1.4: Normal Q-Q plot degli scarti dello strumento con ricette A, E ed F e del righello

L'indice j viene utilizzato per indicare lo strumento di misura, macchinario (M) o righello (R), utilizzato per ottenere gli scarti ai quali i parametri si riferiscono. Quando si indica un parametro relativo al macchinario si utilizza l'indice r per specificare la ricetta. Vengono inoltre riportati i risultati del test T ad un campione per la verifica dell'ipotesi di media nulla.

Nella Tabella 1.3, sono riportate le stime puntuali delle medie e delle varianze degli scarti dello strumento digitale e del righello di precisione. Gli stimatori utilizzati sono la media campionaria

$$\bar{X}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} X_{ij},$$

e la varianza campionaria corretta

$$S_j^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

dove X_{ij} rappresenta l' i -esimo scarto relativo allo strumento di misura j .

	Strumento digitale						Righello
	A	B	C	D	E	F	
$\hat{\mu}$	-0.0369	0.1086	0.0373	-0.0003	-0.0097	0.1804	0.0285
$\hat{\sigma}^2$	0.0037	0.0282	0.006	0.0013	0.0268	0.032	0.0278

Tabella 1.3: Stime puntuali delle medie e delle varianze

Sfruttando l'assunzione di normalità è possibile fornire le stime intervallari per le medie e le varianze. L'intervallo di confidenza di livello $1 - \alpha$ per la media della popolazione j è definito nel seguente modo:

$$IC_{\mu_j}(\alpha) = \left(\bar{X}_j - t_{n_j-1, 1-\frac{\alpha}{2}} \frac{S_j}{\sqrt{n_j}}; \bar{X}_j + t_{n_j-1, 1-\frac{\alpha}{2}} \frac{S_j}{\sqrt{n_j}} \right)$$

dove $t_{n_j-1, 1-\frac{\alpha}{2}}$ indica il quantile di livello $1 - \frac{\alpha}{2}$ della distribuzione T di Student con $n_j - 1$ gradi di libertà. Invece, l'intervallo di confidenza di livello $1 - \alpha$ per la varianza della popolazione j è definito nel seguente modo:

$$IC_{\sigma_j^2}(\alpha) = \left(\frac{(n_j - 1) S_j^2}{\chi_{n_j-1, 1-\frac{\alpha}{2}}^2}; \frac{(n_j - 1) S_j^2}{\chi_{n_j-1, \frac{\alpha}{2}}^2} \right)$$

dove $\chi^2_{\frac{\alpha}{2}, n_j-1}$ e $\chi^2_{1-\frac{\alpha}{2}, n_j-1}$ indicano rispettivamente i quantili di livello $\frac{\alpha}{2}$ e $1 - \frac{\alpha}{2}$ della distribuzione chi-quadro con $n_j - 1$ gradi di libertà. Le stime intervallari al 95% delle medie e delle varianze vengono riportate nella Tabella 1.4. Utilizzando gli intervalli di confidenza per le medie è possibile sottoporre a verifica il seguente sistema di ipotesi:

$$\begin{cases} H_0: \mu_j = 0 \\ H_1: \mu_j \neq 0. \end{cases}$$

Infatti, dato che gli intervalli di confidenza al 95% degli scarti dello strumento digitale con ricette A, B, C ed F e del righello di precisione non includono lo zero, si rifiuta l'ipotesi che i suddetti scarti siano in media nulli. Lo strumento digitale con ricette A, B, C ed F e il righello di precisione hanno una distorsione significativa nelle misure delle altezze dei nastri. In particolare, lo strumento con ricetta A tende a sottostimare le misure mentre lo strumento con ricette B, C ed F e il righello tendono a sovrastimarle. È possibile correggere le distorsioni di nuove misure dello strumento digitale sottraendo la media degli scarti della ricetta utilizzata.

		IC per μ		IC per σ^2	
		Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
Strumento digitale	A	-0.0683	-0.0056	0.0021	0.0084
	B	0.0604	0.1568	0.0197	0.0438
	C	0.0214	0.0531	0.0046	0.0081
	D	-0.0233	0.0227	0.0007	0.0036
	E	-0.0374	0.0179	0.021	0.0345
	F	0.1097	0.251	0.02	0.0594
Righello		0.0149	0.0421	0.0249	0.0313

Tabella 1.4: Intervalli di confidenza al 95% per la media e la varianza

Si utilizza poi il test F per il confronto delle varianze dello strumento digitale con la varianza del righello di precisione. Il sistema di ipotesi che si sottopone a verifica è il seguente:

$$\begin{cases} H_0: \sigma_{M,r}^2 \geq \sigma_R^2 \\ H_0: \sigma_{M,r}^2 < \sigma_R^2, \end{cases}$$

dove r indica la ricetta dello strumento digitale. La Tabella 1.5 riporta i risultati della procedura di verifica di ipotesi sulle varianze. Fissando il livello di significatività al 5%, si rifiuta l'ipotesi nulla per lo strumento digitale con ricette A,

C e D. Ne consegue che le varianze degli scarti di queste ricette sono considerate maggiori della varianza degli scarti del righello.

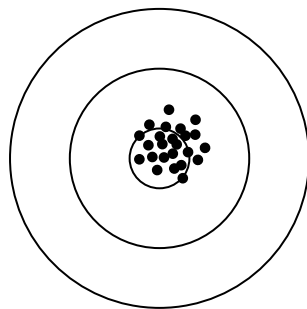
	Strumento digitale					
	A	B	C	D	E	F
Statistica test	0.1347	1.0151	0.2153	0.0481	0.9635	1.1524
Gradi di libertà	17, 577	49, 577	94, 577	12, 577	137, 577	27, 577
p-value	0	0.5519	0	0	0.4020	0.7271

Tabella 1.5: Test F per il confronto delle varianze dello strumento e del righello

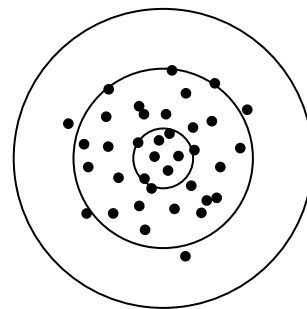
Capitolo 2

Indicatore per la valutazione dell'accuratezza degli strumenti

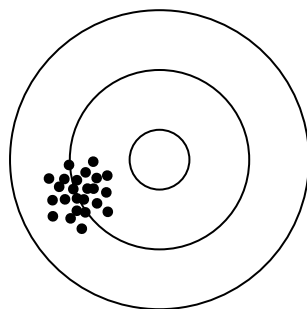
Per studiare l'accuratezza di uno strumento di misura bisogna analizzarne l'esattezza e la precisione. Così facendo, uno strumento può appartenere ad una delle quattro categorie illustrate nella Figura 2.1. La categoria più auspicabile



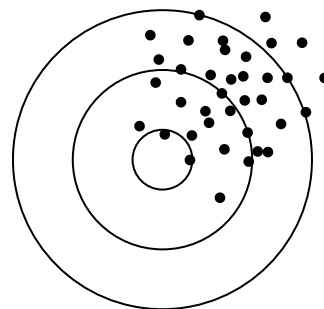
(a) Distorsione e varianza basse



(b) Distorsione bassa, varianza alta



(c) Distorsione alta, varianza bassa



(d) Distorsione e varianza alte

Figura 2.1: Descrizione grafica dell'accuratezza di uno strumento di misura

è quella rappresentata nella Figura 2.1a. Lo strumento che appartiene a tale categoria è definibile accurato in quanto esatto, per la bassa distorsione, e preciso, per la bassa varianza. La Figura 2.1d mostra invece la categoria peggiore. In

questo caso, lo strumento di misura non è né esatto né preciso. Categorie intermedie sono quelle rappresentate nelle Figure 2.1b e 2.1c.

Finora, l'esattezza e la precisione degli strumenti di misura sono state valutate in modo separato accettando l'impossibilità di determinare lo strumento più accurato. In questo capitolo, invece, si introduce un indicatore che permette di valutare l'accuratezza degli strumenti considerando congiuntamente distorsione e varianza e permettendone il confronto. I risultati riportati nelle sezioni successive valgono sia per gli scarti relativi al macchinario sia per gli scarti del righello.

2.1 Definizione dell'indicatore

Siccome la variabile casuale che modella la misurazione di uno strumento è centrata nella vera altezza della tirella misurata, la variabile aleatoria X_j , definita come lo scarto tra la misura e il vero valore, è centrata in zero. Il parametro ψ_j permette di valutare la bontà delle misure rilevate tramite lo strumento j ed è definito come l'errore quadratico medio (o *mean squared error*) degli scarti relativi al macchinario esaminato rispetto a zero. Si scrive quindi che

$$\begin{aligned}\psi_j &= \text{MSE}(X_j) = \\ &= \mathbb{E} \left[(X_j - 0)^2 \right],\end{aligned}$$

da cui segue che il parametro ψ_j coincide con il momento secondo della popolazione j , ovvero

$$\psi_j = \mathbb{E} \left(X_j^2 \right).$$

Ricordando che $\text{var}(X_j) = \mathbb{E} \left[(X_j - \mu_j)^2 \right] = \mathbb{E} \left(X_j^2 \right) - \mathbb{E} (X_j)^2$, si può riscrivere il parametro ψ_j come somma tra il quadrato della media μ_j e la varianza σ_j^2 , infatti

$$\begin{aligned}\psi_j &= \mathbb{E} \left(X_j^2 \right) = \\ &= \mathbb{E} (X_j)^2 + \mathbb{E} \left(X_j^2 \right) - \mathbb{E} (X_j)^2 = \\ &= \mu_j^2 + \sigma_j^2.\end{aligned}\tag{2.1}$$

Quest'ultima forma di ψ_j risulterà utile nelle prossime sezioni.

2.2 Stima di massima verosimiglianza dell'indicatore

Per ottenere la stima di massima verosimiglianza (SMV) dell'indicatore ψ_j si sfruttano le proprietà degli stimatori di massima verosimiglianza. È noto che la stima di massima verosimiglianza del vettore di parametri (μ_j, σ_j^2) di una popolazione normalmente distribuita è

$$(\hat{\mu}_j, \hat{\sigma}_j^2) = \left(\bar{x}_j, \frac{1}{n_j} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 \right)$$

dove \bar{x}_j indica la media campionaria relativa alla popolazione j . Per la dimostrazione si rimanda a [Pace & Salvan \(2001, pag. 137–138\)](#). Si considera ora la seguente trasformazione lineare:

$$\begin{aligned} \tau: \mathbb{R} \times \mathbb{R}^+ &\rightarrow \mathbb{R} \times \mathbb{R}^+ \\ (\mu_j, \sigma_j^2) &\mapsto (\mu_j, \mu_j^2 + \sigma_j^2). \end{aligned}$$

Tale trasformazione è

- iniettiva, perché il suo nucleo è l'insieme vuoto, cioè

$$\ker \tau = \left\{ (\mu_j, \sigma_j^2) \in \mathbb{R} \times \mathbb{R}^+ : \tau(\mu_j, \sigma_j^2) = 0 \right\} = \emptyset,$$

in quanto $\mu_j^2 + \sigma_j^2 > 0$ per ogni $\mu_j \in \mathbb{R}$ e $\sigma_j^2 \in \mathbb{R}^+$;

- suriettiva, perché la sua immagine coincide con il suo codominio, cioè

$$\text{Im } \tau = \left\{ \tau(\mu_j, \sigma_j^2) : (\mu_j, \sigma_j^2) \in \mathbb{R} \times \mathbb{R}^+ \right\} = \mathbb{R} \times \mathbb{R}^+.$$

Segue che la trasformazione τ è biunivoca ed è quindi applicabile la proprietà di equivarianza ([Azzalini \(2008, pag. 63\)](#)).

Proprietà 2.1 (Proprietà di equivarianza). Sia $\psi(\cdot)$ una funzione biunivoca dallo spazio Θ nello spazio Ψ . Allora la SMV di $\psi(\theta)$ è $\psi(\hat{\theta})$, se $\hat{\theta}$ è la SMV di θ relativa alla verosimiglianza $\mathcal{L}(\theta)$.

Di conseguenza, la stima di massima verosimiglianza del vettore di parametri $(\mu_j, \mu_j^2 + \sigma_j^2) = \tau(\mu_j, \sigma_j^2)$ è

$$(\hat{\mu}_j, \hat{\mu}_j^2 + \hat{\sigma}_j^2) = \tau(\hat{\mu}_j, \hat{\sigma}_j^2) = \left(\bar{x}_j, \bar{x}_j^2 + \frac{1}{n_j} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 \right).$$

Ricordare l'Equazione 2.1 permette di determinare la stima di massima verosimiglianza per ψ_j , infatti, sapendo che $\frac{1}{n_j} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij}^2 - \bar{x}_j^2$, si trova

$$\hat{\psi}_j = \hat{\mu}_j^2 + \hat{\sigma}_j^2 = \bar{x}_j^2 + \frac{1}{n_j} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij}^2.$$

2.3 Stimatore MV e distribuzione asintotica

Una volta determinata la stima di massima verosimiglianza $(\hat{\mu}_j, \hat{\psi}_j)$ per il vettore di parametri (μ_j, ψ_j) , se ne deriva lo stimatore sostituendo il valore osservato x_{ij} con la relativa variabile aleatoria X_{ij} con $i = 1, 2, \dots, n_j$. Lo stimatore di massima verosimiglianza per (μ_j, ψ_j) è dunque

$$(\bar{X}_j, \Psi_j) = \left(\frac{1}{n_j} \sum_{i=1}^{n_j} X_{ij}, \frac{1}{n_j} \sum_{i=1}^{n_j} X_{ij}^2 \right).$$

La distribuzione di tale stimatore è asintotica ed è garantita dal seguente teorema.

Teorema 2.1 (Metodo delta). *Sia $\hat{\theta}_n$ lo stimatore di massima verosimiglianza di θ e $\tau = \tau(\theta)$ una funzione derivabile. Allora:*

$$\left[\mathcal{J}_\tau(\theta) \mathcal{J}_n(\theta)^{-1} \mathcal{J}_\tau(\theta)' \right]^{-\frac{1}{2}} (\hat{\tau}_n - \tau) \xrightarrow{d} \mathcal{N}(0, I_n)$$

dove $\mathcal{J}_\tau(\theta)$ indica la matrice jacobiana relativa alla trasformazione τ , $\mathcal{J}_n(\theta)$ indica l'informazione attesa di Fisher relativa a θ e I_n indica la matrice identità di ordine n .

Siccome la trasformazione lineare τ tale che a (μ_j, σ_j^2) associa $(\mu_j, \mu_j^2 + \sigma_j^2) = (\mu_j, \psi_j)$ è derivabile in tutto il suo dominio, lo stimatore $\mathbf{T}_j = (\bar{X}_j, \Psi_j)$ si distribuisce come una normale di media $E(\mathbf{T}_j) \rightarrow (\mu_j, \psi_j)$ e matrice di varianze/covarianze $\text{cov}(\mathbf{T}_j) \rightarrow \mathcal{J}_\tau(\mu_j, \sigma_j^2) \mathcal{J}_{n_j}(\mu_j, \sigma_j^2)^{-1} \mathcal{J}_\tau(\mu_j, \sigma_j^2)'$ con $n_j \rightarrow +\infty$,

cioè

$$\mathbf{T}_j = \begin{bmatrix} \bar{X}_j \\ \Psi_j \end{bmatrix} \xrightarrow{d} \mathcal{N}_2 \left(\begin{bmatrix} \mu_j \\ \psi_j \end{bmatrix}, \mathcal{J}_\tau(\mu_j, \sigma_j^2) \mathcal{J}_{n_j}(\mu_j, \sigma_j^2)^{-1} \mathcal{J}_\tau(\mu_j, \sigma_j^2)' \right).$$

Siccome la matrice jacobiana relativa alla trasformazione lineare τ è

$$\begin{aligned} \mathcal{J}_\tau(\mu_j, \sigma_j^2) &= \begin{bmatrix} \frac{\partial}{\partial \mu_j} \mu_j & \frac{\partial}{\partial \sigma_j^2} \mu_j \\ \frac{\partial}{\partial \mu_j} (\mu_j^2 + \sigma_j^2) & \frac{\partial}{\partial \sigma_j^2} (\mu_j^2 + \sigma_j^2) \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0 \\ 2\mu_j & 1 \end{bmatrix} \end{aligned}$$

e l'informazione attesa di Fisher relativa a (μ_j, σ_j^2) è

$$\mathcal{J}_{n_j}(\mu_j, \sigma_j^2) = \begin{bmatrix} \frac{n_j}{\sigma_j^2} & 0 \\ 0 & \frac{n_j}{2\sigma_j^4} \end{bmatrix},$$

la matrice di varianze/covarianze di \mathbf{T}_j risulta essere

$$\begin{aligned} \text{cov}(\mathbf{T}_j) &\rightarrow \mathcal{J}_\tau(\mu_j, \sigma_j^2) \mathcal{J}_{n_j}(\mu_j, \sigma_j^2)^{-1} \mathcal{J}_\tau(\mu_j, \sigma_j^2)' = \\ &= \begin{bmatrix} 1 & 0 \\ 2\mu_j & 1 \end{bmatrix} \begin{bmatrix} \frac{\sigma_j^2}{n_j} & 0 \\ 0 & 2\frac{\sigma_j^4}{n_j} \end{bmatrix} \begin{bmatrix} 1 & 2\mu_j \\ 0 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} \frac{\sigma_j^2}{n_j} & 2\mu_j \frac{\sigma_j^2}{n_j} \\ 2\mu_j \frac{\sigma_j^2}{n_j} & 2\frac{\sigma_j^4}{n_j} \end{bmatrix} \end{aligned}$$

che, se riscritta in funzione dei parametri μ_j e ψ_j , diventa

$$\text{cov}(\mathbf{T}_j) \rightarrow \begin{bmatrix} \frac{\psi_j - \mu_j^2}{n_j} & 2\mu_j \frac{\psi_j - \mu_j^2}{n_j} \\ 2\mu_j \frac{\psi_j - \mu_j^2}{n_j} & 2\frac{(\psi_j - \mu_j^2)^2}{n_j} \end{bmatrix}$$

con n_j che diverge a $+\infty$. In conclusione, lo stimatore di massima verosimiglianza Ψ_j del parametro ignoto ψ_j si distribuisce in modo asintotico come una normale di media ψ_j e varianza $\frac{2}{n_j} (\psi_j - \mu_j^2)^2$, ovvero

$$\Psi_j \xrightarrow{d} \mathcal{N} \left(\psi_j, \frac{2}{n_j} (\psi_j - \mu_j^2)^2 \right). \quad (2.2)$$

Capitolo 3

Inferenza sull'indicatore

In questo capitolo viene trattato il tema dell'inferenza sull'indicatore ψ_j fornendo le procedure per la costruzione di intervalli di confidenza e la verifica di ipotesi basate sul metodo di Wald e sul rapporto di verosimiglianza. L'ipotesi che si vuole sottoporre a verifica è quella di uguaglianza tra l'indicatore degli scarti dello strumento digitale con una specifica ricetta e quello degli scarti del righello di precisione, rispettivamente $\psi_{M,r}$ con r che indica la ricetta utilizzata e ψ_R . Il sistema di ipotesi che si considera è il seguente:

$$\begin{cases} H_0: \psi_{M,r} = \psi_R \\ H_1: \psi_{M,r} \neq \psi_R \end{cases} \quad (3.1)$$

3.1 Intervallo di confidenza alla Wald

Per determinare l'intervallo di confidenza alla Wald di livello $1 - \alpha$ per il parametro ψ_j è necessario conoscere il seguente teorema.

Teorema 3.1 (Teorema di Slutsky). *Siano $\{X_n\}$ e $\{Y_n\}$ due successioni di variabili aleatorie tali che $X_n \xrightarrow{d} X$ e $Y_n \xrightarrow{p} c$ con $c \in \mathbb{R}$, allora valgono i seguenti risultati:*

$$X_n Y_n \xrightarrow{d} cX \quad e \quad X_n + Y_n \xrightarrow{d} X + c.$$

Ricordando la distribuzione asintotica dello stimatore Ψ_j (2.2) e sfruttando il teorema appena enunciato, si trova che

$$Z = \frac{\Psi_j - \psi_j}{\left(\Psi_j - \bar{X}_j^2\right) \sqrt{\frac{2}{n_j}}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Infatti, se $X_{n_j} = \frac{\Psi_j - \psi_j}{(\Psi_j - \bar{X}_j^2) \sqrt{\frac{2}{n_j}}}$ e $Y_{n_j} = \frac{\psi_j - \mu_j^2}{\Psi_j - \bar{X}_j^2}$, si ha che

$$\underbrace{X_{n_j}}_{\xrightarrow{d} \mathcal{N}(0,1)} \underbrace{Y_{n_j}}_{\xrightarrow{p} 1} = Z \xrightarrow{d} \mathcal{N}(0,1)$$

in quanto \bar{X}_j e Ψ_j sono stimatori consistenti rispettivamente per μ_j e ψ_j . La statistica Z è quindi una quantità asintoticamente pivotale ed è utilizzabile per la costruzione dell'intervallo di confidenza d'interesse. Fissato α , si ricavano i valori $v_1, v_2 \in \mathbb{R}$ tali che

$$\mathbb{P}(v_1 < Z < v_2) = 1 - \alpha.$$

Conoscendo la distribuzione asintotica di Z , i valori v_1 e v_2 sono immediatamente determinabili. Infatti, tali valori coincidono rispettivamente con i quantili di ordini $\frac{\alpha}{2}$ e $1 - \frac{\alpha}{2}$ della distribuzione normale standard, ovvero

$$v_1 = z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}} \quad \text{e} \quad v_2 = z_{1-\frac{\alpha}{2}}.$$

Invertendo la quantità pivotale Z rispetto al parametro ψ_j si trovano gli estremi dell'intervallo di confidenza:

$$\begin{aligned} -z_{1-\frac{\alpha}{2}} &< Z < z_{1-\frac{\alpha}{2}}, \\ -z_{1-\frac{\alpha}{2}} &< \frac{\Psi_j - \psi_j}{(\Psi_j - \bar{X}_j^2) \sqrt{\frac{2}{n_j}}} < z_{1-\frac{\alpha}{2}}, \\ \Psi_j - z_{1-\frac{\alpha}{2}} (\Psi_j - \bar{X}_j^2) \sqrt{\frac{2}{n_j}} &< \psi_j < \Psi_j + z_{1-\frac{\alpha}{2}} (\Psi_j - \bar{X}_j^2) \sqrt{\frac{2}{n_j}}. \end{aligned}$$

L'intervallo di confidenza aleatorio a livello $1 - \alpha$ per ψ_j è

$$\text{IC}_{\psi_j}(\alpha) : \Psi_j \mp z_{1-\frac{\alpha}{2}} (\Psi_j - \bar{X}_j^2) \sqrt{\frac{2}{n_j}}.$$

Essendo i limiti dell'intervallo appena trovati quantità aleatorie, è possibile affermare che la probabilità che il vero parametro ψ_j sia compreso tra limite inferiore e limite superiore è proprio pari a $1 - \alpha$. La stima dell'intervallo di

confidenza è ottenuta sostituendo alle quantità aleatorie le relative stime:

$$IC_{\psi_j}(\alpha) : \hat{\psi}_j \mp z_{1-\frac{\alpha}{2}} \left(\hat{\psi}_j - \hat{\mu}_j^2 \right) \sqrt{\frac{2}{n_j}}.$$

3.2 Intervallo di confidenza basato sul rapporto di verosimiglianza

Si enuncia il teorema che garantisce la distribuzione asintotica di una trasformazione del rapporto di verosimiglianza.

Teorema 3.2 (Teorema di Wilks). *Sia $\mathbf{x} = (x_1, \dots, x_n)$ un campione casuale estratto da $X \sim f(\mathbf{x}; \boldsymbol{\theta})$ con $\boldsymbol{\theta} \in \mathbb{R}^k$. Sia $f(\mathbf{x}; \boldsymbol{\theta})$ la funzione di densità congiunta di \mathbf{x} . Si supponga che $f(\mathbf{x}; \boldsymbol{\theta})$ soddisfi le condizioni di regolarità. Sia $H_0: \theta_1 = \theta_1^0, \dots, \theta_r = \theta_r^0$ con $1 \leq r \leq k$. Allora*

$$-2 \ln \Lambda \xrightarrow[H_0]{d} \chi_r^2$$

dove Λ indica il rapporto di verosimiglianza.

Dunque, la quantità pivotale che permette di ottenere intervalli di confidenza basati sul rapporto di verosimiglianza a livello $1 - \alpha$ per il parametro ψ_j è la seguente:

$$W(\psi_j) = -2 \log \frac{\mathcal{L}(\mu_j, \psi_{0j})}{\mathcal{L}(\mu_j, \psi_j)} \xrightarrow{d} \chi_1^2.$$

Fissato il livello di confidenza $1 - \alpha$, la regione di confidenza di livello $100(1 - \alpha)\%$ per ψ_j coincide con l'insieme dei valori $\psi_j \in \mathbb{R}^+$ tali che $W(\psi_j)$ sia minore del quantile di ordine $1 - \alpha$ della distribuzione chi-quadro con un grado di libertà, cioè

$$IC_{\psi_j}(\alpha) = \left\{ \psi_j \in \mathbb{R}^+ : W(\psi_j) < \chi_{1,1-\alpha}^2 \right\}.$$

3.3 Test di Wald

Siccome le variabili aleatorie $X_{i,M,r}$ e $X_{l,R}$ sono stocasticamente indipendenti tra di loro, cioè $X_{i,M,r} \perp\!\!\!\perp X_{l,R}$ con $i = 1, 2, \dots, n_{M,r}$, $l = 1, 2, \dots, n_R$, anche gli stimatori $\Psi_{M,r}$ e Ψ_R sono indipendenti tra di loro, essendo funzioni note rispettivamente delle variabili $X_{i,M,r}$ con $i = 1, 2, \dots, n_{M,r}$ e $X_{l,R}$ con $l = 1, 2, \dots, n_R$. Per determinare la statistica test bisogna considerare la variabile aleatoria data dalla differenza tra gli stimatori $\Psi_{M,r}$ e Ψ_R . Asintoticamente, la media di tale variabile

è

$$\begin{aligned}
 \lim_{\substack{n_{M,r} \rightarrow \infty \\ n_R \rightarrow \infty}} \mathbb{E}(\Psi_{M,r} - \Psi_R) &\stackrel{\text{lin.}}{=} \lim_{\substack{n_{M,r} \rightarrow \infty \\ n_R \rightarrow \infty}} [\mathbb{E}(\Psi_{M,r}) - \mathbb{E}(\Psi_R)] = \\
 &= \lim_{n_{M,r} \rightarrow \infty} \mathbb{E}(\Psi_{M,r}) - \lim_{n_R \rightarrow \infty} \mathbb{E}(\Psi_R) = \\
 &= \psi_{M,r} - \psi_R
 \end{aligned}$$

mentre la varianza è

$$\begin{aligned}
 \lim_{\substack{n_{M,r} \rightarrow \infty \\ n_R \rightarrow \infty}} \text{var}(\Psi_{M,r} - \Psi_R) &= \lim_{\substack{n_{M,r} \rightarrow \infty \\ n_R \rightarrow \infty}} [\text{var}(\Psi_{M,r}) + \text{var}(\Psi_R) - 2 \text{cov}(\Psi_{M,r}, \Psi_R)] \stackrel{||}{=} \\
 &= \lim_{\substack{n_{M,r} \rightarrow \infty \\ n_R \rightarrow \infty}} [\text{var}(\Psi_{M,r}) + \text{var}(\Psi_R)] = \\
 &= \lim_{n_{M,r} \rightarrow \infty} \text{var}(\Psi_{M,r}) + \lim_{n_R \rightarrow \infty} \text{var}(\Psi_R) = \\
 &= \frac{2}{n_{M,r}} \left(\psi_{M,r} - \mu_{M,r}^2 \right)^2 + \frac{2}{n_R} \left(\psi_R - \mu_R^2 \right)^2.
 \end{aligned}$$

Essendo $\Psi_{M,r}$ e Ψ_R asintoticamente normali (2.2), anche la variabile $\Psi_{M,r} - \Psi_R$ segue una distribuzione normale asintotica, cioè

$$\Psi_{M,r} - \Psi_R \xrightarrow{d} \mathcal{N} \left(\psi_{M,r} - \psi_R, \frac{2}{n_{M,r}} \left(\psi_{M,r} - \mu_{M,r}^2 \right)^2 + \frac{2}{n_R} \left(\psi_R - \mu_R^2 \right)^2 \right).$$

Sotto l'ipotesi nulla di uguaglianza tra $\psi_{M,r}$ e ψ_R , si ha che

$$\Psi_{M,r} - \Psi_R \xrightarrow[H_0]{d} \mathcal{N} \left(0, \frac{2}{n_{M,r}} \left(\psi_{M,r} - \mu_{M,r}^2 \right)^2 + \frac{2}{n_R} \left(\psi_R - \mu_R^2 \right)^2 \right).$$

Standardizzando la variabile $\Psi_{M,r} - \Psi_R$ si trova che

$$\frac{\Psi_{M,r} - \Psi_R}{\sqrt{\frac{2}{n_{M,r}} \left(\psi_{M,r} - \mu_{M,r}^2 \right)^2 + \frac{2}{n_R} \left(\psi_R - \mu_R^2 \right)^2}} \xrightarrow[H_0]{d} \mathcal{N}(0, 1).$$

Per il teorema di Slutsky (3.1), se $X_{n_{M,r}, n_R} = \frac{\Psi_{M,r} - \Psi_R}{\sqrt{\frac{2}{n_{M,r}} (\psi_{M,r} - \mu_{M,r}^2)^2 + \frac{2}{n_R} (\psi_R - \mu_R^2)^2}}$ e

$$Y_{n_{M,r}, n_R} = \sqrt{\frac{\frac{1}{n_{M,r}} (\psi_{M,r} - \mu_{M,r}^2)^2 + \frac{1}{n_R} (\psi_R - \mu_R^2)^2}{\frac{1}{n_{M,r}} (\Psi_{M,r} - \bar{X}_{M,r}^2)^2 + \frac{1}{n_R} (\Psi_R - \bar{X}_R^2)^2}}, \text{ si ha che}$$

$$\underbrace{X_{n_{M,r}, n_R}}_{\xrightarrow[H_0]{d} \mathcal{N}(0,1)} \underbrace{Y_{n_{M,r}, n_R}}_{\xrightarrow{p} 1} = Z \xrightarrow[H_0]{d} \mathcal{N}(0, 1)$$

in quanto \bar{X}_j e Ψ_j sono stimatori consistenti rispettivamente per μ_j e ψ_j con $j = M, R$. Dunque, la statistica test d'interesse è

$$Z = \frac{\Psi_{M,r} - \Psi_R}{\sqrt{\frac{2}{n_{M,r}} (\Psi_{M,r} - \bar{X}_{M,r}^2)^2 + \frac{2}{n_R} (\Psi_R - \bar{X}_R^2)^2}} \xrightarrow[H_0]{d} \mathcal{N}(0, 1). \quad (3.2)$$

Fissato il livello di significatività α , si rifiuta l'ipotesi nulla H_0 se il valore osservato della statistica test

$$z_{\text{oss}} = \frac{\hat{\psi}_{M,r} - \hat{\psi}_R}{\sqrt{\frac{2}{n_{M,r}} (\hat{\psi}_{M,r} - \hat{\mu}_{M,r}^2)^2 + \frac{2}{n_R} (\hat{\psi}_R - \hat{\mu}_R^2)^2}}$$

è, in valore assoluto, maggiore del quantile di livello $1 - \frac{\alpha}{2}$ della distribuzione normale standard,

$$|z_{\text{oss}}| > z_{1-\frac{\alpha}{2}},$$

oppure se il p-value è minore del livello di significatività,

$$\text{p-value} = \mathbb{P}_{H_0} (|Z| > |z_{\text{oss}}|) < \alpha.$$

3.3.1 Test di Wald unilaterale

Il test di Wald unilaterale considerato è quello che sottopone a verifica il seguente sistema di ipotesi:

$$\begin{cases} H_0: \psi_{M,r} \geq \psi_R \\ H_1: \psi_{M,r} < \psi_R. \end{cases} \quad (3.3)$$

Sotto ipotesi nulla, il righeello di precisione è accurato almeno quanto il macchinario con una specifica ricetta. Ci si auspica che ci sia abbastanza evidenza empirica per considerare il test significativo e poter affermare che il macchinario con

una specifica ricetta sia più accurato del righello. Utilizzando la statistica test Z dell'Equazione 3.2, si rifiuta l'ipotesi nulla se il valore osservato della statistica test z_{oss} è minore del quantile di ordine $1 - \alpha$ della distribuzione normale standard $z_{1-\alpha}$, ovvero se

$$z_{\text{oss}} < z_{1-\alpha},$$

oppure se il p-value è minore del livello di significatività α scelto, ovvero se

$$\text{p-value} = \mathbb{P}_{H_0} (Z < z_{\text{oss}}) = \Phi (z_{\text{oss}}) < \alpha$$

dove Φ indica la funzione di ripartizione della distribuzione normale standard.

3.4 Test del rapporto di verosimiglianza

Per determinare la statistica test d'interesse, bisogna ottenere la funzione di verosimiglianza congiunta delle due popolazioni considerate. Grazie all'ipotesi d'indipendenza tra le popolazioni stesse, la funzione di densità congiunta dei campioni casuali osservati $\mathbf{x}_{M,r}$ e \mathbf{x}_R è ottenuta immediatamente moltiplicando le funzioni di densità marginali, ovvero

$$f(\mathbf{x}_{M,r}, \mathbf{x}_R; \boldsymbol{\mu}, \boldsymbol{\psi}) = \prod_{j=M,R} f(\mathbf{x}_j; \boldsymbol{\mu}_j, \boldsymbol{\psi}_j)$$

dove $\boldsymbol{\mu} = (\boldsymbol{\mu}_{M,r}, \boldsymbol{\mu}_R)$ e $\boldsymbol{\psi} = (\boldsymbol{\psi}_{M,r}, \boldsymbol{\psi}_R)$. Di conseguenza, la funzione di verosimiglianza coincide con il prodotto delle funzioni di verosimiglianza marginali, cioè

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\psi}; \mathbf{x}_{M,r}, \mathbf{x}_R) = \prod_{j=M,R} \mathcal{L}(\boldsymbol{\mu}_j, \boldsymbol{\psi}_j; \mathbf{x}_j)$$

e, sotto ipotesi nulla, diventa

$$\mathcal{L}_{H_0}(\boldsymbol{\mu}_0, \boldsymbol{\psi}_0; \mathbf{x}_{M,r}, \mathbf{x}_R) = \prod_{j=M,R} \mathcal{L}(\boldsymbol{\mu}_{0j}, \boldsymbol{\psi}_{0j}; \mathbf{x}_j).$$

La distribuzione asintotica della statistica test è garantita dal teorema 3.2. Se $\Lambda = \frac{\mathcal{L}_{H_0}(\boldsymbol{\mu}_0, \boldsymbol{\psi}_0; \mathbf{x}_{M,r}, \mathbf{x}_R)}{\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\psi}; \mathbf{x}_{M,r}, \mathbf{x}_R)}$, allora la statistica test è

$$W = -2 \log \Lambda \xrightarrow{d} \chi_1^2.$$

La distribuzione di W è una chi-quadro con un grado di libertà perché, sotto ipotesi nulla (3.1), si passa da un modello con spazio parametrico di dimensione 4, $\mathbb{R}^2 \times (\mathbb{R}^+)^2$, ad uno con spazio parametrico di dimensione 3, $\mathbb{R}^2 \times \mathbb{R}^+$. Il valore osservato della statistica test è

$$w_{\text{oss}} = -2 \log \lambda = -2 \log \frac{\mathcal{L}_{H_0}(\hat{\mu}_0, \hat{\psi}_0; \mathbf{x}_{M,r}, \mathbf{x}_R)}{\mathcal{L}(\hat{\mu}, \hat{\psi}; \mathbf{x}_{M,r}, \mathbf{x}_R)}$$

dove $\hat{\mu} = (\hat{\mu}_{M,r}, \hat{\mu}_R)$ con $\hat{\mu}_{M,r} = \frac{1}{n_{M,r}} \sum_{i=1}^{n_{M,r}} x_{i,M,r}$ e $\hat{\mu}_R = \frac{1}{n_R} \sum_{i=1}^{n_R} x_{i,R}$ mentre $\hat{\psi} = (\hat{\psi}_{M,r}, \hat{\psi}_R)$ con $\hat{\psi}_{M,r} = \frac{1}{n_{M,r}} \sum_{i=1}^{n_{M,r}} x_{i,M,r}^2$ e $\hat{\psi}_R = \frac{1}{n_R} \sum_{i=1}^{n_R} x_{i,R}^2$. Per determinare le stime $\hat{\mu}_0$ e $\hat{\psi}_0$ bisogna risolvere il seguente problema di ottimizzazione:

$$(\hat{\mu}_0, \hat{\psi}_0) = \arg \max_{\mu_0, \psi_0} \mathcal{L}_{H_0}(\mu_0, \psi_0; \mathbf{x}_{M,r}, \mathbf{x}_R)$$

Purtroppo, tale problema non ammette una soluzione analitica e si è quindi costretti a procedere per via numerica. Fatto ciò, è possibile calcolare il valore λ e quindi anche il valore osservato della statistica test.

Fissato il livello di significatività α , si rifiuta l'ipotesi nulla H_0 se il valore osservato della statistica test è maggiore del quantile di livello $1 - \alpha$ della distribuzione chi-quadro con un grado di libertà, cioè se

$$w_{\text{oss}} > \chi_{1,1-\alpha}^2$$

oppure se il p-value è minore del livello di significatività, cioè se

$$\text{p-value} = \mathbb{P}_{H_0}(W > w_{\text{oss}}) < \alpha.$$

3.5 Applicazione ai dati

Nella Tabella 3.1 sono riportate le stime di massima verosimiglianza dell'indicatore per gli scarti dello strumento digitale e del righello di precisione. La Figura 3.1

	Strumento digitale						Righello
	A	B	C	D	E	F	
$\hat{\psi}$	0.0049	0.0395	0.0073	0.0012	0.0267	0.0634	0.0286

Tabella 3.1: Stime puntuali dell'indicatore

è una visualizzazione degli strumenti di misura nello spazio bidimensionale

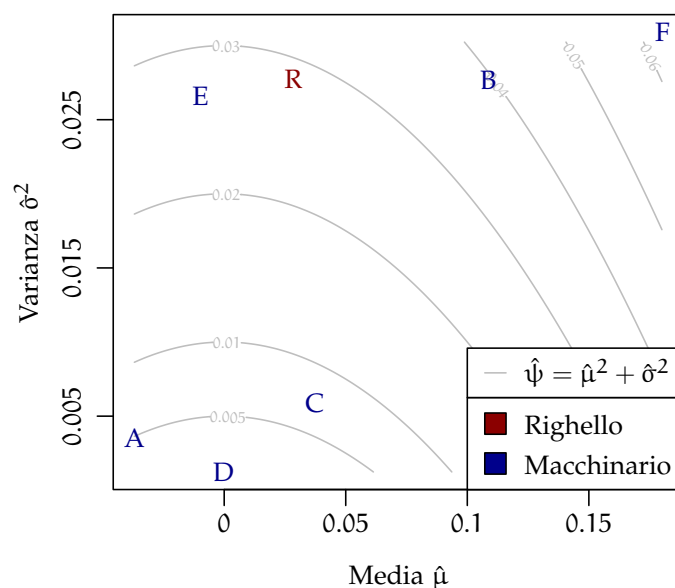


Figura 3.1: Visualizzazione degli strumenti nello spazio della media e della varianza

identificato dalle medie e dalle varianze degli scarti. Osservando tale grafico si può notare la presenza di tre gruppi: uno composto dalle ricette A, D e C con medie e varianze degli scarti relativamente piccole, uno composto dalla ricetta E e dal righello con medie basse e varianze alte e, infine, uno composto dalle ricette B ed F con medie basse e varianze alte. Questi gruppi rispecchiano tre delle quattro situazioni illustrate nella Figura 2.1. La Tabella 3.2 riporta le stime intervallari basate sul metodo di Wald e sul rapporto di verosimiglianza per gli indicatori dei diversi strumenti di misura.

		IC alla Wald		IC basato sul RV	
		Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
Strumento digitale	A	0.0026	0.0072	0.0033	0.0087
	B	0.0286	0.0503	0.0309	0.0539
	C	0.0056	0.009	0.0059	0.0094
	D	0.0003	0.0022	0.0006	0.0030
	E	0.0204	0.033	0.0213	0.0341
	F	0.0472	0.0796	0.0516	0.0874
Righello		0.0254	0.0318	0.0256	0.0320

Tabella 3.2: Intervalli di confidenza al 95% per l'indicatore

La Tabella 3.3 riporta invece i p-value del test di Wald e del test del rapporto di verosimiglianza per il confronto bilaterale degli indicatori dello strumento digitale con l'indicatore del righello di precisione. Fissato il livello di significatività al 5%, il due test portano agli stessi risultati anche nel caso di piccoli campioni:

solo per il macchinario con ricette B ed E non si rifiuta l'ipotesi di uguaglianza tra ciascun indicatore dello strumento digitale e l'indicatore del righello. Gli indicatori del macchinario con ricette A, C, D e F sono quindi statisticamente diversi dall'indicatore del righello.

	Strumento digitale					
	A	B	C	D	E	F
Test di Wald	0	0.0592	0	0	0.6008	0
Test RV	0	0.0941	0	0	0.6138	0.0004

Tabella 3.3: p-value dei test per il confronto tra lo strumento e il righello

Il criterio per poter considerare uno strumento più o meno accurato dell'altro consiste nel test di Wald per il confronto unilaterale degli indicatori. I risultati di tale test sono riportati nella Tabella 3.4. A tutti i livelli di significatività usualmente scelti si rifiuta l'ipotesi nulla $H_0: \psi_{M,r} \geq \psi_R$ solo per le ricette A, C e D. Di conseguenza, secondo la procedura scelta, il macchinario con ricette A, C e D è più accurato del righello di precisione. Invece, per le ricette B, E ed F, non c'è evidenza empirica sufficiente per rifiutare la suddetta ipotesi nulla. Il macchinario con ricette B, E ed F viene quindi considerato accurato al massimo come il righello. Però, siccome gli indicatori del macchinario con ricette B ed E sono statisticamente uguali all'indicatore del righello di precisione al 5%, solo lo strumento digitale con ricetta F viene ritenuto meno accurato del righello.

	Strumento digitale					
	A	B	C	D	E	F
Statistica test	-11.7535	1.8867	-11.5203	-16.0505	-0.5233	4.1406
p-value	0	0.9704	0	0	0.3004	1

Tabella 3.4: Test di Wald unilaterale

Capitolo 4

Individuazione delle cause d'inaccuratezza delle misure

Questo capitolo è dedicato alla ricerca di cause di inaccuratezza delle misure dello strumento digitale. Per raggiungere questo obiettivo sono state rilevate alcune caratteristiche delle tirelle misurate del macchinario. Tramite la stima di modelli di regressione lineare si studiano le relazioni tra l'imprecisione delle misure e le variabili considerate. Sulla base del coefficiente stimato, si determina la significatività della relativa covariata. Determinare quali sono le fonti d'imprecisione delle misure permette di apportare i giusti cambiamenti nei parametri delle diverse ricette per aumentarne l'accuratezza. Nelle sezioni successive vengono riportate una descrizione generale dei modelli utilizzati e l'applicazione di tale procedura agli scarti ottenuti dal macchinario utilizzando la ricetta E, quella per cui si ha il maggior numero di osservazioni.

4.1 Il dataset

Il dataset che viene analizzato è composto dalle variabili presenti nella Tabella 4.1. Di ogni tirella per la quale è stato calcolato lo scarto del macchinario sono stati rilevati i valori teorici dell'altezza, dello spessore, del numero di trame per cm e dell'apertura della maglia teorici. Si prendono in considerazione queste variabili perché sono tutte legate al grado di trasparenza del tessuto, aspetto fondamentale per il metodo di stima delle misure dello strumento digitale.

Variabile	Descrizione
altezza	altezza teorica della tirella in mm
spessore	spessore teorico della tirella in μm
fili.per.cm	numero teorico di trame per cm
apertura.maglia	apertura teorica delle maglie in μm
scarto	scostamenti delle misure dalle vere altezze
ricetta	ricetta usata per le misure: A, ..., F

Tabella 4.1: Descrizione delle variabili del dataset

4.2 Descrizione generale della procedura di modellizzazione

Si stima un modello di regressione lineare per ogni ricetta. Indicati con \mathbf{Y} il vettore aleatorio degli scarti dello strumento digitale con ricetta considerata e con \mathbf{X} la matrice che ha elementi pari a 1 sulla prima colonna e le n osservazioni delle $k - 1$ covariate sulle restanti colonne, si definisce il modello di regressione lineare

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

dove $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_k]$ è il vettore dei coefficienti di regressione e $\boldsymbol{\epsilon}$ è il vettore di errori classici. Per loro definizione, gli errori classici ϵ_i con $i = 1, 2, \dots, n$ hanno valore atteso nullo $\mathbb{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ e sono omoschedastici e indipendenti tra di loro $\text{var}(\boldsymbol{\epsilon}) = \sigma_\epsilon^2 \mathbf{I}_n$ con σ_ϵ^2 costante positiva. In forma scalare, il modello d'interesse è

$$Y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i \text{ con } i = 1, 2, \dots, n.$$

Il vettore $\boldsymbol{\beta}$ viene stimato attraverso lo stimatore OLS (*ordinary least squares*) così definito:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

Tale stimatore è non distorto per $\boldsymbol{\beta}$ ed efficiente con matrice di varianze/covarianze $\text{var}(\hat{\boldsymbol{\beta}}) = \sigma_\epsilon^2 (\mathbf{X}'\mathbf{X})^{-1}$.

Nel caso in cui non dovesse valere l'assunzione di omoschedasticità sui dati utilizzati, si ha che $\text{var}(\boldsymbol{\epsilon}) = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2)$ con $\sigma_i^2 > 0$ per $i = 1, 2, \dots, n$. In tal caso, l'eteroschedasticità degli errori viene descritta tramite modelli. I modelli che vengono considerati in seguito sono quello di Breusch-Pagan e quello di Harvey. Il modello di Breusch-Pagan prevede che la varianza dell' i -esimo errore

$\text{var}(\epsilon_i) = \sigma_i^2$ sia funzione lineare delle covariate del modello d'interesse, ovvero

$$\sigma_i^2 = \gamma_1 + \gamma_2 x_{i2} + \cdots + \gamma_k x_{ik} \text{ con } i = 1, 2, \dots, n.$$

Il modello di Harvey prevede invece che il logaritmo della varianza dell' i -esimo errore sia funzione lineare delle covariate del modello d'interesse, ovvero

$$\log \sigma_i^2 = \gamma_1 + \gamma_2 x_{i2} + \cdots + \gamma_k x_{ik} \text{ con } i = 1, 2, \dots, n.$$

In entrambi i modelli, i parametri $\gamma_0, \gamma_1, \dots, \gamma_k$ sono ignoti e devono essere stimati.

Si noti che modellizzare le varianze degli errori equivale a modellizzare le varianze della variabile dipendente Y , infatti

$$\text{var}(Y_i) = \text{var}(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i) = \text{var}(\epsilon_i).$$

La stima di tali modelli è effettuata considerando come stime delle varianze degli errori i quadrati dei rispettivi residui calcolati sul modello d'interesse. Si ha quindi che

$$\hat{\sigma}_i^2 = \hat{\epsilon}_i^2 \text{ con } i = 1, 2, \dots, n$$

dove $\hat{\epsilon}_i$ indica l' i -esimo residuo del modello d'interesse. Si stimano quindi le seguenti regressioni ausiliarie:

$$\hat{\epsilon}_i^2 = \gamma_1 + \gamma_2 x_{i2} + \cdots + \gamma_k x_{ik} + \eta_i \text{ con } i = 1, 2, \dots, n$$

e

$$\log \hat{\epsilon}_i^2 = \gamma_1 + \gamma_2 x_{i2} + \cdots + \gamma_k x_{ik} + \eta_i \text{ con } i = 1, 2, \dots, n$$

dove η_i con $i = 1, 2, \dots, n$ è un errore classico.

Per verificare che i modelli considerati siano opportuni per spiegare l'eteroschedasticità degli errori del modello d'interesse, si utilizza, un test di omoschedasticità comune per ognuno di essi. Tale test sottopone a verifica l'ipotesi nulla

$$H_0: \gamma_2 = \cdots = \gamma_k = 0.$$

La statistica test utilizzata per saggiare H_0 è la seguente:

$$nR^2 \xrightarrow[H_0]{d} \chi_{k-1}^2$$

dove R^2 indica il coefficiente di determinazione lineare relativo al modello di eteroschedasticità considerato. Valori elevati della statistica test sono critici per l'ipotesi nulla per cui si rifiuta H_0 se il p-value è minore del livello di significatività α scelto.

Scelto il miglior modello di eteroschedasticità per i dati analizzati, si ottengono i valori fittati $\widehat{\sigma}_i^2$ con $i = 1, 2, \dots, n$ e si determina la stima della matrice di varianze/covarianze degli errori $\text{vâr}(\epsilon) = \text{diag}(\widehat{\sigma}_1^2, \widehat{\sigma}_2^2, \dots, \widehat{\sigma}_n^2)$. Fatto ciò si calcola la matrice

$$\hat{L} = \text{diag} \left(\frac{1}{\sqrt{\widehat{\sigma}_1^2}}, \frac{1}{\sqrt{\widehat{\sigma}_2^2}}, \dots, \frac{1}{\sqrt{\widehat{\sigma}_n^2}} \right)$$

che è tale che $\hat{L}'\hat{L} = \text{vâr}(\epsilon)$ e $\hat{L}[\text{vâr}(\epsilon)]^{-1}\hat{L}' = I_n$. Ottenuta \hat{L} , si determina il modello d'interesse trasformato:

$$\underbrace{\hat{L}'Y}_{=Y^*} = \underbrace{\hat{L}'X}_{=X^*} \beta + \underbrace{\hat{L}'\epsilon}_{=\epsilon^*}$$

$$Y^* = X^* \beta + \epsilon^*.$$

In forma scalare, il modello d'interesse trasformato risulta essere

$$\underbrace{\frac{Y_i}{\sqrt{\widehat{\sigma}_i^2}}}_{=Y_i^*} = \beta_1 \underbrace{\frac{1}{\sqrt{\widehat{\sigma}_i^2}}}_{=1_i^*} + \beta_2 \underbrace{\frac{x_{i2}}{\sqrt{\widehat{\sigma}_i^2}}}_{=x_{i2}^*} + \dots + \beta_k \underbrace{\frac{x_{ik}}{\sqrt{\widehat{\sigma}_i^2}}}_{=x_{ik}^*} + \underbrace{\frac{\epsilon_i}{\sqrt{\widehat{\sigma}_i^2}}}_{=\epsilon_i^*},$$

$$Y_i^* = \beta_1 1_i^* + \beta_2 x_{i2}^* + \dots + \beta_k x_{ik}^* + \epsilon_i^* \text{ con } i = 1, 2, \dots, n.$$

Lo stimatore OLS sul modello trasformato coincide con lo stimatore FGLS (*feasible generalized least squares*):

$$\begin{aligned} \hat{\beta} &= (X^{*'}X^*)^{-1} X^{*'}Y^* = \\ &= (X'\hat{L}\hat{L}'X)^{-1} X'\hat{L}\hat{L}'Y = \\ &= (X'\hat{\Omega}X)^{-1} X'\hat{\Omega}Y = \\ &= \hat{\beta}_{\text{FGLS}}. \end{aligned}$$

Questo stimatore risulta essere consistente per β se il modello è corretto.

4.3 Modello lineare per gli scarti dello strumento digitale con ricetta E

Analizzando la matrice di correlazione della Figura 4.2, si decide di inserire come covariate nel modello di regressione lineare solo le variabili *altezza* e *fili.per.cm*. La Tabella 4.1 permette di studiare graficamente la relazione degli scarti con

	scarti	altezza	spessore	fili.per.cm	apertura.maglia
scarti	1	-0.08	-0.17	0.21	-0.19
altezza	-0.08	1	0.83	-0.56	0.62
spessore	-0.17	0.83	1	-0.91	0.92
fili.per.cm	0.21	-0.56	-0.91	1	-0.97
apertura.maglia	-0.19	0.62	0.92	-0.97	1

Tabella 4.2: Matrice di correlazione

l'altezza della tirella misurata e il numero di fili per cm da cui è composta. I grafici suggeriscono che l'irrelevanza di tali variabili nello spiegare gli scarti: all'aumentare dell'altezza o del numero di fili per cm, il livello medio degli scarti sembra costante. A cambiare, però, è la varianza degli scarti per diversi livelli di altezza e di fili per cm. Si nota infatti che, all'aumentare dell'altezza e al diminuire del numero di fili per cm, aumenta la variabilità degli scarti intorno al livello medio.

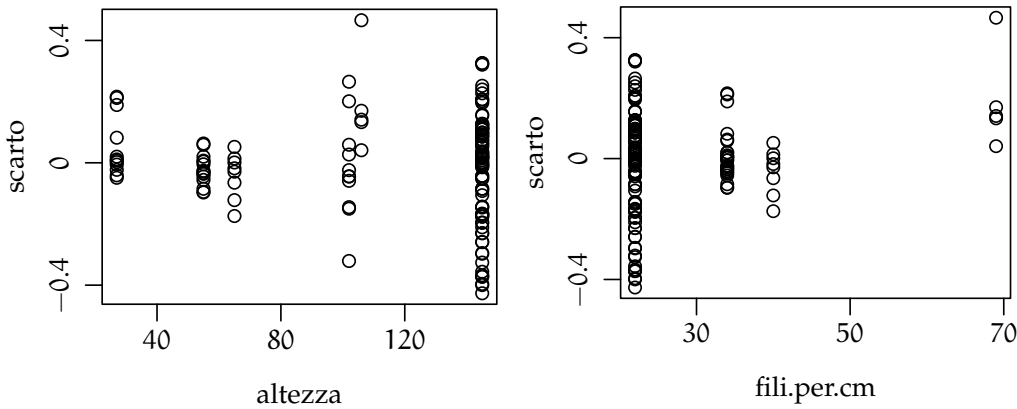


Figura 4.1: Grafici di dispersione scarto–altezza e scarto–fili.per.cm

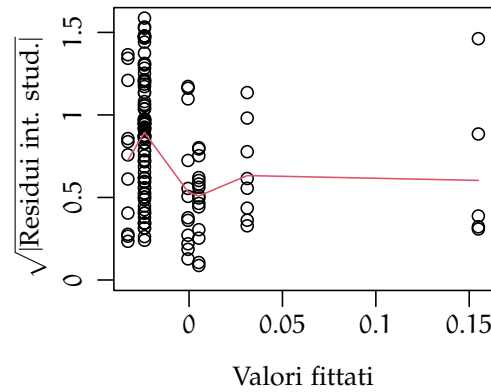


Figura 4.2: Scale-location plot del modello d'interesse

4.3.1 Specificazione e stima del modello d'interesse

Per le considerazioni fatte in precedenza, il modello d'interesse per gli scarti dello strumento digitale con ricetta E è il seguente:

$$\text{scarto}_i = \beta_0 + \beta_1 \text{altezza}_i + \beta_2 \text{fili.per.cm}_i + \epsilon_i \text{ con } i = 1, 2, \dots, n.$$

La stima del modello è riportata nella Tabella 4.3. Tali stime, tuttavia, non

	Stima	Std. error	Statistica test	p-value
(intercetta)	-0.1413	0.0807	-1.7513	0.0822
altezza	0.0002	0.0004	0.5282	0.5982
fili.per.cm	0.004	0.0016	2.4094	0.0173

Tabella 4.3: Stima del modello d'interesse con test di significatività dei coefficienti

sono affidabili perché, visualizzando lo *scale-location plot*, ovvero un grafico di dispersione con i valori fittati sull'asse delle ascisse e con la radice quadrata del valore assoluto dei residui internamente studentizzati sull'asse delle ordinate, si diagnostica un possibile problema dell'eteroschedasticità degli errori (Figura 4.2). Si tenta quindi di descrivere l'eteroschedasticità tramite i modelli di Breusch-Pagan e di Harvey.

4.3.2 Trattamento dell'eteroschedasticità

Per trattare l'eteroschedasticità degli errori si stimano le seguenti regressioni ausiliarie rispettivamente relative al modello di Breusch-Pagan e di Harvey:

$$\hat{\sigma}_i^2 = \hat{e}_i^2 = \gamma_0 + \gamma_1 \text{altezza}_i + \gamma_2 \text{fili.per.cm}_i + \eta_i$$

e

$$\log \hat{\sigma}_i^2 = \log \hat{e}_i^2 = \gamma_0 + \gamma_1 \text{altezza}_i + \gamma_2 \text{fili.per.cm}_i + \eta_i.$$

Stimati i due modelli, si calcolano i relativi test di omoschedasticità. I risultati sono riportati nella Tabella 4.4. Entrambi i test portano al rifiuto dell'ipotesi di

	Statistica test	Gradi di libertà	p-value
Test di Breusch-Pagan	11.1553	2	0.0038
Test di Harvey	25.2994	2	0

Tabella 4.4: Risultati dei test di Breusch-Pagan e di Harvey

omoschedasticità degli errori. Segue che sia il modello di Breusch-Pagan sia quello di Harvey sono compatibili con l'eteroschedasticità degli errori del modello d'interesse. Siccome le statistiche test dei due modelli considerati hanno la stessa distribuzione, si sceglie il modello che ha maggiore evidenza empirica contro l'ipotesi nulla. Si decide quindi di accettare come valido il modello di Harvey in quanto riporta il p-value minore.

La stima del modello scelto è riportata nella Tabella 4.5. Solo l'altezza

	Stima	Std. error	Statistica test	p-value
(intercetta)	-7.272	1.0769	-6.7529	0
altezza	0.0215	0.0052	4.1145	0.0001
fili.per.cm	-0.0154	0.022	-0.7004	0.4849

Tabella 4.5: Stima del modello di Harvey con test di significatività dei coefficienti

delle tirelle misurate, oltre all'intercetta, è significativa per spiegare la varianza degli scarti ed è quindi causa d'imprecisione delle misure. A parità del resto, all'aumentare di un mm dell'altezza della tirella misurata, la varianza degli scarti aumenta del 2.17%.

4.3.3 Stima del modello trasformato

Ottenuti i valori fittati $\hat{\sigma}_i^2$ con $i = 1, 2, \dots, n$ dal modello di Harvey stimato in precedenza, si stima il modello trasformato

$$\underbrace{\frac{\text{scarto}_i}{\sqrt{\hat{\sigma}_i^2}}}_{=\text{scarto}_i^*} = \beta_0 \underbrace{\frac{1}{\sqrt{\hat{\sigma}_i^2}}}_{=\text{intercetta}_i^*} + \beta_1 \underbrace{\frac{\text{altezza}_i}{\sqrt{\hat{\sigma}_i^2}}}_{=\text{altezza}_i^*} + \beta_2 \underbrace{\frac{\text{fili.per.cm}_i}{\sqrt{\hat{\sigma}_i^2}}}_{=\text{fili.per.cm}_i^*} + \underbrace{\frac{\epsilon_i}{\sqrt{\hat{\sigma}_i^2}}}_{=\epsilon_i^*},$$

$$\text{scarto}_i^* = \beta_0 \text{intercetta}_i^* + \beta_1 \text{altezza}_i^* + \beta_2 \text{fili.per.cm}_i^* + \epsilon_i^*.$$

Ispezionando la presenza di punti critici, si individua una tirella di altezza teorica pari a 106 mm con 69 fili per cm sulla quale il macchinario con ricetta E ha ottenuto uno scarto di 0.466 mm. Tale punto viene considerato un outlier al 5% perché ha un p-value con correzione di Bonferroni pari a 0.0475 circa. Quest'osservazione è ritenuta essere anche un punto di leva perché ha un valore di leva maggiore della media dei valori di leva. Si decide quindi di ristimare il modello trasformato non considerando questo punto. Così facendo non si ottengono indicazioni grafiche che portano a pensare che le assunzioni sottostanti al modello non siano valide. Osservando la Tabella 4.6 si nota che tutti i coefficienti non sono significativi al 5%. Non c'è evidenza empirica che supporta l'ipotesi iniziale che altezza e numero di fili per cm influenzino in modo statisticamente significativo gli scarti relativi allo strumento digitale con ricetta E.

	Stima	Std. error	Statistica test	p-value
intercetta*	-0.0328	0.0474	-0.6931	0.4894
altezza*	-0.0003	0.0002	-1.3187	0.1895
fili.per.cm*	0.0019	0.0011	1.6209	0.1074

Tabella 4.6: Stima del modello trasformato con test di significatività dei coefficienti

Capitolo 5

Conclusioni

Questa analisi è nata dalla necessità aziendale di misurare l'affidabilità dello strumento di misura digitale introdotto nell'area industriale per sostituire il righello di precisione. Gli obiettivi dell'elaborato sono stati due: il primo, creare una procedura statistica per la valutazione dello strumento digitale confrontandolo con il righello di precisione, strumento finora adottato; il secondo, invece, modellizzare gli scarti dello strumento digitale in modo da individuare le cause d'inaccuratezza delle misure. Sfruttando i principi di inferenza statistica si è definito l'indicatore che ha permesso di raggiungere il primo obiettivo. Sulla base dell'errore quadratico medio degli scarti, utilizzando i test di Wald e i test basati sul rapporto di verosimiglianza, è stato determinato un criterio per l'ordinamento degli strumenti di misura. Accettando la violazione dell'ipotesi di normalità degli scarti della maggior parte delle ricette dello strumento digitale e del righello di precisione secondo il test di Shapiro-Wilk, si è determinata la classifica degli strumenti in termini di accuratezza riportata nella Tabella 5.1. La ricetta F è

Posizione	Strumento di misura
1	D
2	A
3	C
4	E, Righello, B
5	F

Tabella 5.1: Classifica degli strumenti di misura dal più accurato al meno accurato

critica e necessita di una revisione. Il riesame dei parametri che caratterizzano questa ricetta potrà essere supportato dalla procedura di modellizzazione che ha permesso di raggiungere il secondo obiettivo dell'elaborato. Stimando modelli di regressione lineare per gli scarti dello strumento digitale con una ricetta scelta e

per le stime delle loro varianze, è possibile diagnosticare problemi d'inesattezza e/o d'imprecisione delle misure evidenziandone le cause. Per motivi di scarsità di dati, si è deciso di applicare tale procedura solamente agli scarti dello strumento digitale con ricetta E. Questo programma non ha mostrato una distorsione delle misure importante ma è stato rilevato un incremento significativo della variabilità delle stesse all'aumentare dell'altezza delle tirelle misurate. Ampliando i campioni sarà possibile ripetere l'analisi auspicando risultati più precisi ed estendere la stima dei modelli di regressione a tutte le ricette del macchinario con particolare attenzione alla ricetta F. Una volta che lo strumento digitale sarà più accurato del righello di precisione, lo si potrà riprodurre e renderlo disponibile in più centri di lavoro per poter eliminare la fase di ricontrollo delle altezze dei nastri effettuata dal laboratorio di controllo qualità.

Ulteriori sviluppi dello studio potrebbero consistere nella ricerca di differenze di accuratezza dello strumento digitale con una specifica ricetta per diverse tipologie di tessuti. Innanzitutto, si utilizzerebbero tecniche di *cluster analysis* per il raggruppamento di tirelle omogenee. Successivamente, si confronterebbe l'accuratezza dello strumento digitale sui diversi gruppi sfruttando l'indicatore ψ definito in precedenza. L'eventuale significatività della differenza di esattezza o di precisione per diverse gruppi di tessuti suggerirebbe la creazione di una nuova ricetta per la categoria associata a maggiore inaccuratezza.

Bibliografia

AZZALINI, A. (2008). *Inferenza Statistica*. Springer Verlag.

PACE, L. & SALVAN, A. (2001). *Introduzione alla statistica II*. Cedam.

SHAPIRO, S. S. & WILK, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611.