Atleticastudi, 3-4 1988, pp. 435-457

ATLETICASTUDI

ASPETTI TECNICO-STATISTICI DELLA POLIVALENZA E VERSATILITÀ ATLETICA DI TIPO QUALITATIVO IN ALCUNE DISCIPLINE DI ATLETICA LEGGERA
Prima parte

Un impiego sportivo dell'analisi di regressione e correlazione lineare

Otello Donzelli, *collaboratore del Centro Studi & Ricerche F.I.D.A.L.*

Indice

Premessa

Alcune considerazioni di carattere generale sulle carriere sportive di alcuni grandi campioni di decathlon del presente e del passato hanno suggerito all'autore di proporre un'analisi tecnico-statistica circa i particolari aspetti della polivalenza e versatilità atletica nell'ambito, normalmente poco trattato, del solo carattere qualitativo del fenomeno sportivo.

Pertanto, la figura del decathleta, che primeggia a buoni livelli assoluti in ognuna delle note 10 specialità, verrà sostituita nell'indagine da quella dell'atleta che si distingue in poche discipline da lui liberamente scelte e praticate, ma ai massimi livelli mondiali.

Le tecniche statistiche impiegate dall'autore nella ricerca — analisi di regressione e correlazione lineare con preliminare studio di alcuni aspetti della variabilità del fenomeno (deviazione-standard e coefficiente di variazione) — permettono di inquadrare il fenomeno sportivo in questione, attualmente in costante evoluzione, secondo schemi di indagine che, seppure non troppo consueti, risultano, specialmente alla luce degli importanti avvenimenti di questi ultimi tempi, di grande attualità ed interesse. Al fine di una più proficua osservazione del fenomeno sportivo ed in relazione alle specifiche situazioni tecniche ante e post-Olimpiade di Seul '88, la ricerca è stata predisposta in due parti cronologicamente collegate, nella seconda delle quali sarà analizzato il dettaglio delle variazioni tecniche che saranno espresse dalla manifestazioni olimpica.

1. Introduzione

I tempi e le misure delle prestazioni atletiche dei vincitori delle gare di atletica leggera della prima olimpiade moderna disputata in Grecia (Atene - 1896) rappresentano oggi, indubbiamente, poco più di una semplice curiosità, almeno dal punto di vista tecnico attuale. C'è da sottolineare, però, che lo storico avvenimento di quell'epoca non si esaurisce nello sbiadito ricordo delle imprese sportive, ormai tanto lontane dal nostro tempo, degli atleti che vi hanno preso parte, e neanche nella disamina più o meno critica dei valori tecnici espressi dalle varie competizioni disputate. Oggi come allora, una delle problematiche più interessanti che merita di essere esaminata è l'aspetto, sportivamente immutabile nel tempo, della capacità di un determinato atleta ad esprimersi, ai massimi livelli mondiali, in più specialità aventi caratteristiche diverse.

Gli atleti americani T. Burke, E. Clark, R. Garrett e l'australiano E. Flack, ognuno vincitore nella prima olimpiade moderna in due differenti specialità, rappresentano in assoluto, nella storia sportiva delle attuali discipline di atletica leggera, il primo esempio di predisposizione alla *polivalenza e versatilità atletica* nel senso indicato. Il dettaglio tecnico che riguarda, nel lungo arco di tempo delle olimpiadi finora disputate, quanto appena accennato, è riportato nella Tabella n. 1.

Peraltro, c'è da aggiungere che problematiche sportive di questo genere possono essere poste, ovviamente, anche a proposito di atleti detentori, nel passato come nel presente, di records mondiali, con la sola differenza cronologica dell'inizio del fenomeno, all'incirca collocabile, per la maggior parte delle specialità, intorno al 1912. In quest'ultimo caso, il numero ancora maggiore di atleti ci suggerisce di soprassedere ad una analoga elencazione; ciò, comunque, non sposta eccessivamente i termini della questione che, semmai, proprio per questo ne viene ad essere magnificata. Si tenga presente, inoltre, che l'esclusione dalla Tabella n. 1 dei dati tecnici relativi a titoli olimpici conseguiti nelle specialità collettive delle staffette 4×100 m. e 4×400 m. è dovuta solo al fatto pratico di non incorrere nella citazione ripetitiva di numerosi atleti vincitori di analoghe gare individuali.

Prima di procedere oltre, è intanto opportuna una puntualizzazione sul concetto di polivalenza e versatilità atletica.

TABELLA N. 1 - ATLETI VINCITORI DI OLIMPIADI IN GARE DIVERSE

Settore Maschile

N.	Atleta	Naz.	Olimpiade	Gare	Tempi o Misure
1	T. Burke	USA	1896-Atene	metri 100	12.0
			1896-Atene	metri 400	54.2
2	E. Flack	AUS	1896-Atene	metri 800	2:11.0
			1896-Atene	metri 1500	4:33.2
3	E. Clark	USA	1896-Atene	salto in alto	1.81
			1896-Atene	salto in lungo	6.34
4	R. Garrett	USA	1896-Atene	getto del peso	11.22
			1896-Atene	lancio del disco	29.15
5	J. Tewkesbury	USA	1900-Parigi	metri 200	22.2
			1900-Parigi	metri 400 ost.	57.6
6	J. Baxter	USA	1900-Parigi	salto in alto	1.90
			1900-Parigi	salto con asta	3.30
7	M. Prinstein	USA	1900-Parigi	salto triplo	14.47
			1904-St. Louis	salto triplo	14.325
			1904-St. Louis	salto in lungo	7.35
8	A. Hahn	USA	1904-St. Louis	metri 100	11.0
			1904-St. Louis	metri 200	21.6
9.	J. Lightbody	USA	1904-St. Louis	metri 800	1:56.0
			1904-St. Louis	metri 1500	4:05.4
10	H. Hillman	USA	1904-St. Louis	metri 400	49.4
			1904-St. Louis	metri 400 ost.	53.0
				(su ostacoli di cm. 76.2)	
11	M. Sheppard	USA	1908-Londra	metri 800	1:52.8
			1908-Londra	metri 1500	4:03.4
12	R. Craig	USA	1912-Stoccolma	metri 100	10.8
			1912-Stoccolma	metri 200	21.7
13	H. Kolehmainen	FIN	1912-Stoccolma	metri 5000	14:36.6
			1912-Stoccolma	metri 10000	31:20.8
			1920-Anversa	maratona	2h32:35.8
				(gara disputata sulla distanza di Km. 42.750)	
14	A. Hill	GBR	1920-Anversa	metri 800	1:53.4
			1920-Anversa	metri 1500	4:01.8
15	P. Nurmi	FIN	1920-Anversa	metri 10000	31:45.8
			1924-Parigi	metri 1500	3:53.6
			1928-Amsterdam	metri 10000	30:18.8
16	V. Ritola	FIN	1924-Parigi	metri 10000	30:23.2
			1924-Parigi	metri 3000 siepi	9:33.6
			1928-Amsterdam	metri 5000	14:38.0
17	P. Williams	CAN	1928-Amsterdam	metri 100	10.8
			1928-Amsterdam	metri 200	21.8
18	E. Tolan	USA	1932-Los Angeles	metri 100	10.3
			1932-Los Angeles	metri 200	21.2
19	J. Owens	USA	1936-Berlino	metri 100	10.3
			1936-Berlino	metri 200	20.7
			1936-Berlino	salto in lungo	8.06

20	H. Dillard	USA	1948-Londra	metri 100	10.3
			1952-Helsinki	metri 110 ost.	13.7
21	E. Zatopek	TCH	1948-Londra	metri 10000	29:59.6
			1952-Helsinki	metri 5000	14:06.6
			1952-Helsinki	metri 10000	29:17.0
			1952-Helsinki	maratona	2h23:03.2
22	B. Morrow	USA	1956-Melbourne	metri 100	10.5
			1956-Melbourne	metri 200	20.6
23	V. Kuts	URS	1956-Melbourne	metri 5000	13:39.6
			1956-Melbourne	metri 10000	28:45.6
24	P. Snell	NZL	1960-Roma	metri 800	1:46.3
			1964-Tokyo	metri 800	1:45.1
			1964-Tokyo	metri 1500	3:38.1
25	K. Keino	KEN	1968-Città d. Messico	metri 1500	3:34.9
			1972-Monaco d. Baviera	metri 3000 siepi	8:23.6
26	V. Borzov	URS	1972-Monaco d. Baviera	metri 100	10.14
			1972-Monaco d. Baviera	metri 200	20.00
27	L. Viren	FIN	1972-Monaco d. Baviera	metri 5000	13:26.4
			1972-Monaco d. Baviera	metri 10000	27:38.4
			1976-Montreal	metri 5000	13:24.8
			1976-Montreal	metri 10000	27.40.4
28	M. Yfter	ETI	1980-Mosca	metri 5000	13:21.0
			1980-Mosca	metri 10000	27:42.7
29	C. Lewis	USA	1984-Los Angeles	metri 100	9.99
			1984-Los Angeles	metri 200	19.80
			1984-Los Angeles	salto in lungo	8.54

Settore femminile

1	F.E. Blankers-Koen	OLA	1948-Londra	metri 100	11.9
			1948-Londra	metri 200	24.4
			1948-Londra	metri 80 ost.	11.2
2	M.D.M. Ostermeyer	FRA	1948-Londra	getto del peso	13.75
				lancio del disco	41,92
3	M. Jackson	AUS	1952-Helsinki	metri 100	11.5
			1952-Helsinki	metri 200	23.7
4	B. Cuthbert	AUS	1956-Melbourne	metri 100	11.5
			1964-Tokyo	metri 400	52.0
5	W. Rudolph	USA	1960-Roma	metri 100	11.0
			1960-Roma	metri 200	24.0
6	I. Press	URS	1960-Roma	metri 80 ost.	10.8
			1960-Roma	pentathlon	p. 5246
7	T. Press	URS	1960-Roma	getto del peso	17.32
			1964-Tokyo	getto del peso	18.14
			1964-Tokyo	lancio del disco	57.27
8	R. Stecher	RDT	1972-Monaco d. Baviera	metri 100	11.07
			1972-Monaco d. Baviera	metri 200	22.40
9	T. Kazankina	URS	1976-Montreal	metri 800	1:54.9
			1980-Mosca	metri 1500	3:56.6
10	V. Brisco-Hooks	USA	1984-Los Angeles	metri 200	21.81
			1984-Los Angeles	metri 400	48.83

Le problematiche di questo particolare aspetto possono essere affrontate, nel campo delle attuali discipline di atletica leggera, secondo due distinti criteri, affini nella forma, ma fondamentalmente diversi nella sostanza. Il primo criterio è quello della polivalenza e versatilità intesa come capacità di un determinato atleta ad esprimersi in due o più specialità, non prestabilite, ma ai massimi livelli assoluti per ogni singola specialità, conformando pertanto l'analisi di tale problematica ad una interpretazione eminentemente di tipo qualitativo (criterio di verticalità, ovvero valore, delle prestazioni atletiche). Il secondo criterio, viceversa, è relativo alla capacità di un determinato atleta ad esprimersi in molte prestabilite specialità (quelle del decathlon, per l'appunto), ma non necessariamente ai massimi livelli assoluti per ogni singola specialità, conformando l'analisi, di conseguenza, ad una interpretazione di tipo quantitativo (criterio di orizzontalità, ovvero numerosità, del tipo delle prestazioni atletiche).

Facendo riferimento alla realtà sportiva attuale, relativamente a quanto appena accennato, è facile constatare, ad esempio, che la polivalenza e versatilità di atleti come gli americani Carl Lewis (metri 100 e 200 piani, salto in lungo e staffetta 4×100) e André Phillips (metri 200 e 400 piani, 110 e 400 ad ostacoli) od anche il tedesco federale Harald Schmid (metri 400 e 800 piani, 400 ad ostacoli) non è certamente confrontabile con quella di atleti come il britannico Daley Thompson, il tedesco federale Jürgen Hingsen ed il tedesco orientale Uwe Freimuth, specialisti della gara di decathlon. La citazione dei primi tre atleti, in proposito, non è senza motivo. Infatti, i propri records personali nelle specialità indicate permettono di inquadrare molto chiaramente i termini di ciò che deve intendersi per polivalenza e versatilità qualitativa di un atleta. La Tabella n. 2 riporta i dati tecnici delle migliori prestazioni degli atleti in questione, la cui collocazione è indubbiamente ai vertici mondiali assoluti delle rispettive specialità, sebbene in presenza di reciproche contrastanti caratteristiche tecniche. Questo, appunto, è l'aspetto dell'argomento che sarà trattato in questa sede, rimandando ad altra occasione l'analisi delle problematiche tecnico-statistiche relative alle interpretazioni prevalentemente quantitative delle gare di decathlon.

TABELLA N. 2 - RECORDS PERSONALI DI 3 RAPPRESENTATIVI ATLETI POLIVALENTI A LIVELLO MONDIALE

1) CARL LEWIS (USA) 1961

Metri 100 piani
9.93 30 Agosto 1987 Roma
2° atleta *all time* in parità con Calvin Smith (USA)

Metri 200 piani
19.75 19 Giugno 1983 Indianapolis
2° atleta *all time*

Salto in lungo
8.79 19 Giugno 1983 Indianapolis
2° atleta *all time*

IVª frazione staffetta 4×100 metri
37.83 11 Agosto 1984 Los Angeles
attuale record mondiale

2) ANDRÈ PHILLIPS (USA) 1959

Metri 200 piani
20.55 1 Agosto 1986 Los Angeles

Metri 400 piani
44.71 17 Maggio 1986 Westwood
25° atleta *all time*

Metri 110 ad ostacoli
13.25 28 Luglio 1985 Baton Rouge
9° atleta *all time*

Metri 400 ad ostacoli
47.51 5 Settembre 1986 Bruxelles
3° atleta *all time*

3) HARALD SCHMID (FRG) 1957

Metri 400 piani
44.92 11 Agosto 1979 Stoccarda
48° atleta *all time*

Metri 800 piani
1:44.84 15 Agosto 1979 Zurigo

Metri 400 ad ostacoli
47.48 8 Settembre 1982 Atene
2° atleta *all time*

Una osservazione attenta della Tabella n. 1 ci permette di valutare, sebbene molto sinteticamente, il significato tecnico delle performances di atleti del passato come J. Tewkesbury, H. Hillman, H. Kolehmainen, P. Nurmi, V. Ritola, E. Zatopek, K. Keino in campo maschile, e analogamente di atlete come F.E. Blankers-Koen, B. Cuthbert, I. Szewinska, L. Bragina, in campo femminile.

In effetti, le relative specialità praticate dagli atleti citati, evidenziando notoriamente un sensibile contrasto tecnico rispetto ad altre gare (ovviamente considerate in coppia), conferiscono alle misure delle prestazioni atletiche osservate connotati specifici di particolare valore ed interesse sportivo.

Per il presente e recente passato, peraltro, non si presentano molte difficoltà per le corrispondenti citazioni, considerando la grande notorietà e numerosità degli atleti che praticano, o hanno di recente praticato, due o più specialità ad alto livello.

In relazione a quanto appena detto, quindi, il fenomeno in questione sarà analizzato, contestualmente in due diverse specialità, mediante lo studio dei parametri tecnici relativi alle dieci migliori gare di alcuni atleti di indiscussa caratura mondiale. Per il settore femminile, alcune difficoltà nel reperimento dei dati statistici necessari ci hanno indotto, almeno per il momento, a limitare la scelta ad una sola atleta che, fortunatamente però, rappresenta, per motivi che saranno esposti adeguatamente quando sarà analizzata la relativa scheda tecnica, un caso molto significativo dello specifico fenomeno osservato. Una particolarità degna di nota è il fatto che alcune gare saranno analizzate più volte allo scopo di effettuare anche comparazioni indirette fra performances di differenti atleti in un medesimo tipo di gara. Infine, in considerazione del fatto che le analisi dei singoli atleti presentano caratteristiche formali ripetitive, è stato deciso di trattare in modo esauriente, con funzione di analisi-guida, le performances di un solo atleta e, di conseguenza, conformare a tale analisi-guida tutte le altre.

Per la scelta delle performances da inserire nell'analisi-guida, è stato fatto riferimento obiettivo alle risultanze della pregevole pubblicazione «ALL TIME WORLD BEST PERFORMANCES» compilata, per l'edizione 1987 A.T.F.S., a cura dell'Arch. Pino Mappa dell'Association of Track and Field Members.

In detta pubblicazione, in effetti, si può rilevare che l'italiano Pietro Mennea, da noi scelto per lo scopo, può vantare 15 delle 1070 migliori performances mai finora registrate sui 100 metri piani da parte di 192 atleti, con tempi compresi fra 10.01 (suo record personale) e 10.23 (sua 10ª migliore performance personale), e 85 delle 1020 migliori prestazioni mai finora registrate sui 200 metri piani da parte di 208 atleti, con tempi compresi fra 19.72 (attuale suo record mondiale dal 1979) e 20.60 (sua 10ª migliore performance personale). Come si può notare, con le 85 presenze nella predetta lista *all time* Pietro Mennea raggiunge ben l'8.31% del totale delle predette performances; salvo errore, non ci sembra esistano attualmente al mondo atleti che, in una qualsiasi specialità e nei termini indicati, siano in grado soltanto di avvicinare un simile exploit. Nelle due gare in questione, ad esempio, Carl Lewis e Calvin Smith possono vantare, sui 100 e 200 metri piani rispettivamente, solo 50/15 presenze il primo e 51/31 presenze il secondo, mentre il prodigioso canadese Ben Johnson, attuale recordman mondiale dei 100 metri piani (9.83, Roma 30 Agosto 1987) può vantare, sì, 34 presenze nella lista *all time* della gara più breve, ma è poi praticamente assente, anche se per propria scelta tecnica, in quella dei 200 metri piani (due sole presenze: 20.41 e 20.76, entrambe a Bruxelles, rispettivamente nel 1985 e 1987, se non andiamo errati).

Per quanto riguarda il fenomeno sportivo nel suo insieme, abbiamo ritenuto che le 10 migliori performances di ogni atleta nelle due gare oggetto di analisi potessero offrire un quadro abbastanza preciso delle relative situazioni. Inoltre, tutti gli atleti scelti per l'indagine, ad esclusione degli americani Larry Myricks e André Phillips, e del tedesco federale Harald Schmid (i quali, comunque, restano di assoluto valore mondiale), hanno stabilito recentemente almeno un record mondiale, vigente o meno, ovvero hanno conseguito almeno una

TABELLA N. 3 - ATLETI E GARE

N.	Atleta		Anno di nascita	Naz.	Iª Gara		IIª Gara
Settore maschile							
1	Mennea	Pietro	1952	ITA	metri 100		metri 200
2	Lewis	Carl	1961	USA	metri 100		metri 200
3	Juantorena	Alberto	1950	CUB	metri 400		metri 800
4	Coe	Sebastian	1956	GBR	metri 800		metri 1500
5	Aouita	Said	1960	MAR	metri 1500		metri 50000
6	Cova	Alberto	1958	ITA	metri 5000		metri 10000
7	Mamede	Fernando	1951	POR	metri 5000		metri 10000
8	Phillips	André	1959	USA	metri 110 h		metri 400 h
9	Schmid	Harald	1957	FRG	metri 400		metri 400 h
10	Rono	Henry	1952	KEN	metri 5000		metri 3000 siepi
11	Lewis	Carl	1961	USA	metri 100		salto in lungo
12	Myricks	Larry	1956	USA	metri 200		salto in lungo
Settore Femminile							
13	Koch	Marita	1957	GDR	metri 200	—	metri 400

vittoria olimpica. Diciamo anche, per la precisione, che non vi sono stati particolari motivi sulla scelta di questi atleti, salvo il fatto della loro grande notorietà e dell'elevatissimo valore tecnico dell'insieme delle 10 migliori prestazioni atletiche nelle relative specialità. Numerosi altri atleti di pari valore e notorietà avrebbero fornito, certamente, un'alternativa ugualmente valida, ma, tant'è, gli atleti da noi scelti per l'indagine ci sembrano in grado di offrire, obiettivamente, la massima garanzia sul loro grado di rappresentatività del fenomeno sportivo preso in esame. L'elenco completo degli atleti e delle relative coppie di gare è riportato nella Tabella n. 3.

2. Analisi della polivalenza e versatilità qualitativa di alcuni atleti di livello mondiale

Premesso che l'analisi considerata viene condotta, per definizione, mediante lo studio delle 10 migliori performances di ogni atleta in due distinte gare, appare chiaro che occorre, in primis, indagare sulla dinamica del fenomeno sportivo nel suo insieme, nel senso che le predette performances siano in grado di stabilire termini comuni (ed omogenei) di riferimento che abbiano, per così dire, la capacità di esprimere la *legge* di interrelazione secondo la quale il fenomeno stesso si *muove*.

Come primo passo in questa direzione, sembra senz'altro conveniente affrontare il fenomeno su alcuni particolari aspetti della sua variabilità. La prima parte dell'analisi, pertanto, sarà dedicata allo studio della dinamica degli scarti fra tutte le singole performances e le relative medie (aritmetiche) di ognuna delle due gare di volta in volta considerate. Si tratterà, in sostanza, di arrivare alla definizione delle cosiddette *costanti caratteristiche* o *valori segnaletici* del fenomeno e quindi, mediante opportuna standardizzazione, al confronto fra serie diverse di performances, eventualmente anche nel caso di unità di misura non coincidenti.

Come già accennato in precedenza, il primo atleta oggetto di analisi sarà l'italiano Pietro Mennea, sulle cui relative performances nelle gare dei 100 e 200 metri piani sarà incentrata l'analisi-guida che farà da riferimento per tutte le altre, secondo l'ordine della predetta Tabella n. 3.

2.1 - Pietro Mennea (ITA) 1952
Metri 100 e 200 piani
Analisi-guida

2.1.1 Alcuni aspetti della variabilità

Per quanto riguarda il calcolo degli scarti, indispensabile per ricavare gli elementi essenziali del grado di variabilità del fenomeno nella forma richiesta, la soluzione più logica sarebbe quella di operare sullo *scostamento semplice medio* (media degli scarti - in assoluto - fra le singole performances di ognuna delle due gare e le relative medie aritmetiche). Il valore dello scostamento semplice medio è dato dalla formula:

$$s_m = \frac{\sum_{i=1}^{n} \left| (x_i - \bar{x}) \right|}{n}$$

dove: x_i = valore osservato
$\bar{x}$ = media aritmetica
n = numero delle osservazioni

la quale, però, implica il ricorso ai valori assoluti poiché, altrimenti, la somma algebrica di tutti gli scostamenti stessi risulterebbe nulla.

Per ovviare alla scomoda trattazione in termini di valori assoluti è preferibile ricorrere al quadrato degli scostamenti in questione e, successivamente, alla estrazione della radice quadrata *positiva* della sommatoria degli scostamenti così ricavati, secondo la sequenza:

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

formula della *varianza*, e

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

formula dello *scarto quadratico medio*, ovvero della *deviazione-standard*.

TABELLA N. 4

SCHEDA DELLE 10 MIGLIORI PRESTAZIONI DI PIETRO MENNEA (ITA) 1952 NELLE GARE DEI METRI 100 E 200

N.	Gara	Tempo	Data		Località
1	Metri 100	10.01	4 Settembre	1979	Città del Messico
	Metri 200	19.72	12 Settembre	1979	Città del Messico
2	Metri 100	10.15	4 Agosto	1979	Torino
	Metri 200	19.96	10 Settembre	1979	Cltà del Messico
3	Metri 100	10.15	23 Settembre	1979	Spalato
	Metri 200	19.96	17 Agosto	1980	Barletta
4	Metri 100	10.18	19 Settembre	1979	Bologna
	Metri 200	20.01	5 Agosto	1980	Roma
5	Metri 100	10.19	29 Agosto	1978	Praga
	Metri 200	20.03	20 Settembre	1980	Tokyo
6	Metri 100	10.19	24 Giugno	1980	Torino
	Metri 200	20.03	27 Settembre	1980	Pechino
7	Metri 100	10.20	12 Luglio	1975	Torino
	Metri 200	20.04	11 Settembre	1979	Cltà del Messico
8	Metri 100	10.22	18 Agosto	1979	Lignano S.
	Metri 200	20.05	22 Agosto	1980	Bruxelles
9	Metri 100	10.23	20 Agosto	1975	Zurigo
	Metri 200	20.07	13 Agosto	1980	Rovereto
10	Metri 100	10.23	1 Luglio	1978	Milano
	Metri 200	20.07	13 Ottobre	1984	Brindisi

La Tabella n. 5 riporta tutti i calcoli effettuati per ricavare i valori delle deviazioni-standard con il procedimento indicato poc'anzi.

I valori relativi alle 10 performances delle due gare di Pietro Mennea risultano:

a) 0.0612 secondi per i 100 metri
b) 0.0985 secondi per i 200 metri

che rapportati alla corrente prassi di cronometraggio delle gare corrispondono a 6.12 e 9.85 centesimi di secondo, rispettivamente.

Ciò evidenzia chiaramente una maggiore variabilità, in assoluto, delle prestazioni dell'atleta nelle gare dei 200 metri. Le gare in questione, però, pur avendo una comune unità di misura del tempo, presentano, pur sempre, caratteristiche tecniche e di grandezza differenti fra loro, le principali delle quali, come è noto, sono costituite rispettivamente dai tempi diversi di *ammortamento*, nell'economia delle due distinte gare, dell'handicap della partenza (tempo di reazione allo sparo dello «starter» e critici 10/15 metri iniziali della gara) e dalla diversa lunghezza di gara (e quindi diverso arco di tempo impiegato a percorrere la distanza). Conseguentemente, il confronto diretto fra i valori assoluti di dette deviazioni-standard potrebbe risultare non precisamente rispondente alla reale variabilità del fenomeno, che è poi ciò che ci interessa maggiormente.

In questo caso, allora, l'uso del coefficiente di variazione:

$$V = \frac{s}{\bar{x}},$$

dove: V = Coefficiente di variazione
s = deviazione-standard
$\bar{x}$ = media aritmetica delle 10 performances,

svincolato dall'unità di misura, sarà in grado di interpretare più correttamente la sostanza del fenomeno osservato.

TABELLA N. 5

CALCOLO DELLA VARIANZA E DELLA DEVIAZIONE-STANDARD
PIETRO MENNEA (ITA) 1952

Gara dei metri 100 - Media delle 10 migliori performances: $\bar{x} = 10.175$

1 $(x_1 - \bar{x}^2) = (10.01 - 10.175)^2 = (-0.165)^2 = 0.027225$
2 $(x_2 - \bar{x}^2) = (10.15 - 10.175)^2 = (-0.025)^2 = 0.000625$
3 $(x_3 - \bar{x}^2) = (10.15 - 10.175)^2 = (-0.025)^2 = 0.000625$
4 $(x_4 - \bar{x}^2) = (10.18 - 10.175)^2 = (\ 0.005)^2 = 0.000025$
5 $(x_5 - \bar{x}^2) = (10.19 - 10.175)^2 = (\ 0.015)^2 = 0.000125$
6 $(x_6 - \bar{x}^2) = (10.19 - 10.175)^2 = (\ 0.015)^2 = 0.000125$
7 $(x_7 - \bar{x}^2) = (10.20 - 10.175)^2 = (\ 0.025)^2 = 0.000625$
8 $(x_8 - \bar{x}^2) = (10.22 - 10.175)^2 = (\ 0.045)^2 = 0.002025$
9 $(x_9 - \bar{x}^2) = (10.23 - 10.175)^2 = (\ 0.055)^2 = 0.003025$
10 $(x_{10} - \bar{x}^2) = (10.23 - 10.175)^2 = (\ 0.055)^2 = 0.003025$

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \qquad 0.03745$$

$$\text{Varianza} = s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} = \frac{0.03745}{10} = 0.003745$$

$$\text{Deviazione-standard} = s = \sqrt{0.003745} = 0.0611964051 = 0.0612$$

Gara dei metri 200 - Media delle 10 migliori performances: $\bar{y}$ = 19.994

$$
\begin{array}{rlll}
1 & (y_1 - \bar{y}^2) = (19.72 - 19.994)^2 & = (-0.274)^2 & = 0.075076 \\
2 & (y_2 - \bar{y}^2) = (19.96 - 19.994)^2 & = (-0.034)^2 & = 0.001156 \\
3 & (y_3 - \bar{y}^2) = (19.96 - 19.994)^2 & = (-0.034)^2 & = 0.001156 \\
4 & (y_4 - \bar{y}^2) = (20.01 - 19.994)^2 & = (0.016)^2 & = 0.000256 \\
5 & (y_5 - \bar{y}^2) = (20.03 - 19.994)^2 & = (0.036)^2 & = 0.001296 \\
6 & (y_6 - \bar{y}^2) = (20.03 - 19.993)^2 & = (0.036)^2 & = 0.001296 \\
7 & (y_7 - \bar{y}^2) = (20.04 - 19.994)^2 & = (0.046)^2 & = 0.002116 \\
8 & (y_8 - \bar{y}^2) = (20.05 - 19.994)^2 & = (0.056)^2 & = 0.003136 \\
9 & (y_9 - \bar{y}^2) = (20.07 - 19.994)^2 & = (0.076)^2 & = 0.005776 \\
10 & (y_{10} - \bar{y}^2) = (20.07 - 19.994)^2 & = (0.076)^2 & = 0.005776
\end{array}
$$

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \qquad 0.09704$$

$$\text{Varianza} = s^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n} = \frac{0.09704}{10} = 0.009704$$

$$\text{Deviazione-standard} = s = \sqrt{0.009704} = 0.0985088828 = 0.0985$$

Avremo allora:

V_1 = Coefficiente di variazione riferito alla gara dei 100 metri =

$$= \frac{s_1}{\bar{x}_1} = \frac{0.0612}{10.175} = 0.00601$$

V_2 = Coefficiente di variazione riferito alla gara dei 200 metri =

$$= \frac{s_2}{\bar{x}_2} = \frac{0.0985}{19.994} = 0.00493$$

I due valori, trasformati da assoluti a relativi con l'introduzione nel rapporto della media aritmetica ed espressi, pertanto, come numeri puri (cioè adimensionali), sono ora correttamente paragonabili reciprocamente. Ovviamente, detti valori possono essere espressi anche in termini percentuali. In questo caso avremo i seguenti valori:

a) 0.601% per la gara dei 100 metri
b) 0.493% per la gara dei 200 metri

i quali, naturalmente, non sono ora più riferiti alla variabilità assoluta — in secondi — del fenomeno, bensì alla percentuale della sua intensità media.

È opportuno, adesso, fare un passo indietro. Riprendendo in esame la Tabella n. 5 si può notare che la deviazione-standard così ricavata è una misura che necessita del calcolo degli scarti fra tutte le singole performances considerate e la loro media. Tale procedimento, che certamente illustra con molta proprietà il significato intrinseco del nesso esistente fra la varianza e la deviazione-standard, costringe, per il calcolo delle differenze successive e dei rispettivi quadrati, a continui arrotondamenti e, ovviamente, ad approssimazioni sempre più sensibili.

È lo sviluppo della stessa formula della deviazione-standard che ci suggerisce il modo per evitare lo scomodo calcolo degli scarti.

Infatti:

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

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

$$= \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2}{n}} ;$$

essendo però: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

e quindi: $\sum_{i=1}^{n} x_i = n\ \bar{x}$

avremo:

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

$$= \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - 2\,n\bar{x}^2 + n\ \bar{x}^2}{n}} =$$

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

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

Semplificando, infine, avremo la formula finale che permette di operare senza ricorrere al calcolo degli scarti:

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

Pertanto, la deviazione-standard può essere espressa indifferentemente con e senza calcolo degli scarti, secondo la seguente uguaglianza:

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

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

ovvero:

la media quadratica degli scarti delle singole performances dalla media è uguale alla radice quadrata della differenza tra la media dei quadrati delle singole performances ed il quadrato della media delle performances stesse.

Esaminiamo le due modalità di calcolo:

a) gara dei metri 100

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

$$= \sqrt{\frac{0.03745}{10}} =$$

$$= \sqrt{0.003745} = 0.0611964051$$

ovvero:

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

$$= \sqrt{\frac{1035.3439}{10} - 103.530625} = 0.0613595958$$

Come si può vedere, le approssimazioni effettuate nei due diversi tipi di operazione consentono l'uguaglianza dei valori solo fino al terzo decimale;
b) gara dei metri 200

$$s = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}} =$$

$$= \sqrt{\frac{0.09704}{10}} =$$

$$= \sqrt{0.009704} = 0.0985088828$$

ovvero:

$$s = \sqrt{\frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2} =$$

$$= \sqrt{\frac{3997.6974}{10} - 399.760036} = 0.0985088828$$

In questo caso, sebbene con diverso procedimento, si giunge al medesimo valore di varianza (al sesto decimale), per cui il conseguente valore della deviazione-standard è coincidente.

Per quanto finora detto e per la praticità che deriva dalla eliminazione del problema delle differenze successive nel calcolo eseguito in forma canonica, oltreché per la maggiore speditezza delle operazioni, i valori delle deviazioni-standard relativi alle performances degli altri atleti saranno ricavati, d'ora in poi, seguendo le modalità della 2ª formula, la quale, tra l'altro, prevede l'uso vantaggioso di alcuni parametri presenti nelle tabelle destinate alle successive analisi di regressione e correlazione.

Come prima impressione, possiamo intanto dire che i valori della deviazioni-standard e dei coefficienti di variazione sembrano veramente molto contenuti e già in grado di offrire una risposta adeguata circa le caratteristiche qualitative dell'adattamento dell'atleta alle due specialità praticate. Comunque, più ancora che la singola valutazione tecnico-statistica delle misure appena espresse dall'analisi di variabilità condotta nel senso indicato, sarà l'insieme dei confronti con i corrispondenti valori relativi alle coppie di gare degli altri atleti a dare una esauriente risposta alle problematiche in questione. Come è facile intuire, il pregio maggiore di detti confronti, che saranno effettuati dopo la completa acquisizione dei dati indispensabili, risiede nel fatto che la standardizzazione dei valori potrà permettere, naturalmente con le dovute puntualizzazioni, la comparazione del grado di polivalenza e versatilità specifica fra atleti specialisti di gare anche non affini tra loro.

2.1.2 Analisi di regressione lineare

La disposizione dei dati tecnici dell'atleta in un diagramma cartesiano di assi coordinati (Tabella n. 4 e Tavola n. 1), indica molto chiaramente l'esistenza di una *relazione* (nel senso di *associazione*) fra le due variabili considerate. Lo *scatter*, infatti, suggerisce che la relazione in questione può essere adeguatamente rappresentata da una funzione del tipo: $Y = a+bx$, cioè una retta. Il problema della determinazione dei parametri a e b è facilmente risolvibile con l'adozione del metodo dei minimi quadrati in un sistema di equazioni normali (cfr. «Atleticastudi», n. 4/5-Luglio/Ottobre 1986 e n. 4 - Luglio/Agosto 1987).

Per la natura del fenomeno sportivo osservato, occorre intanto dire che non si hanno motivi sufficienti per considerare analiticamente come *indipendente* una qualsiasi delle due variabili, per cui è consigliabile studiare il fenomeno di relazione tracciando due distinte *rette di regressione*, precisamente: $Y = a+bx$ e $X = a+by$, che in genere non coincidono; le rette stesse si incontreranno nel punto di baricentro dei dati, punto rappresentato nel diagramma dalla coppia di valori relativi alle medie (aritmetiche) $\bar{x}$ e $\bar{y}$ (Tavola n. 1).

Una prima valutazione dell'intensità della relazione è già possibile con l' osservazione dell'angolo α che misura la divergenza fra le rette di regressione (Tavola n. 2). Il grado di relazione intercorrente fra le due variabili sotto controllo, infatti, è inversamente proporzionale al grado di divergenza delle rette di regressione, ovvero, il che è la stessa cosa, è direttamente proporzionale al grado di convergenza delle stesse.

La Tabella n. 6 contiene tutte le elaborazioni che hanno portato alla definizione delle rette di regressione, la cui sistemazione nel diagramma cartesiano della Tavola n. 2 ci ha consentito di avere

TABELLA N. 6 - PIETRO MENNEA (ITA) 1952 - CALCOLO DEI PARAMETRI DELLE RETTE DI REGRESSIONE

N.	x	y	x^2	y^2	xy	valori $\hat{Y}$	teorici $\hat{X}$
1	10.01	19.72	100.2001	388.8784	193.3972	19.73	10.00
2	10.15	19.96	103.0225	398.4016	202.5940	19.95	10.15
3	10.15	19.96	103.0225	398.4016	202.5940	19.95	10.15
4	10.18	20.01	103.6324	400.4001	203.7018	20.00	10.18
5	10.19	20.03	103.8361	401.2009	204.1057	20.02	10.19
6	10.19	20.03	103.8361	401.2009	204.1057	20.02	10.19
7	10.20	20.04	104.0400	401.6016	204.4080	20.03	10.20
8	10.22	20.05	104.4484	402.0025	204.9110	20.06	10.21
9	10.23	20.07	104.6529	402.8049	205.3161	20.08	10.22
10	10.23	20.07	104.6529	402.8049	205.3161	20.08	10.22
	101.75	199.94	1035.3439	3997.6974	2034.4496		

$$a_{yx} = \frac{(199.94 \times 1035.3439) - (101.75 \times 2034.4496)}{(10 \times 1035.3439) - (101.75)^2} = \frac{1.412566}{0.3765} = 3.75183$$

$$b_{yx} = \frac{(10 \times 2034.449.6) - (101.75 \times 199.94)}{(10 \times 1035.3439) - (101.75)^2} = \frac{0.601}{0.3765} = 1.59628$$

Equazione della I^a retta di regressione: $\hat{Y}_i = 3.75183+1.59628\ x_i$

$$a_{xy} = \frac{(101.75 \times 3997.6974) - (199.94 \times 2034.4496)}{(10 \times 3997.6974) - (199.94)^2} = \frac{-2.142574}{0.9704} = -2.20793$$

$$b_{xy} = \frac{(10 \times 2034.4496) - (199.94 \times 101.75)}{(10 \times 3997.6974) - (199.94)^2} = \frac{0.601}{0.9704} = 0.61933$$

Equazione della II^a retta di regressione: $\hat{X}_i = -2.20793+0.61933\ y_i$

Tavola n. 1

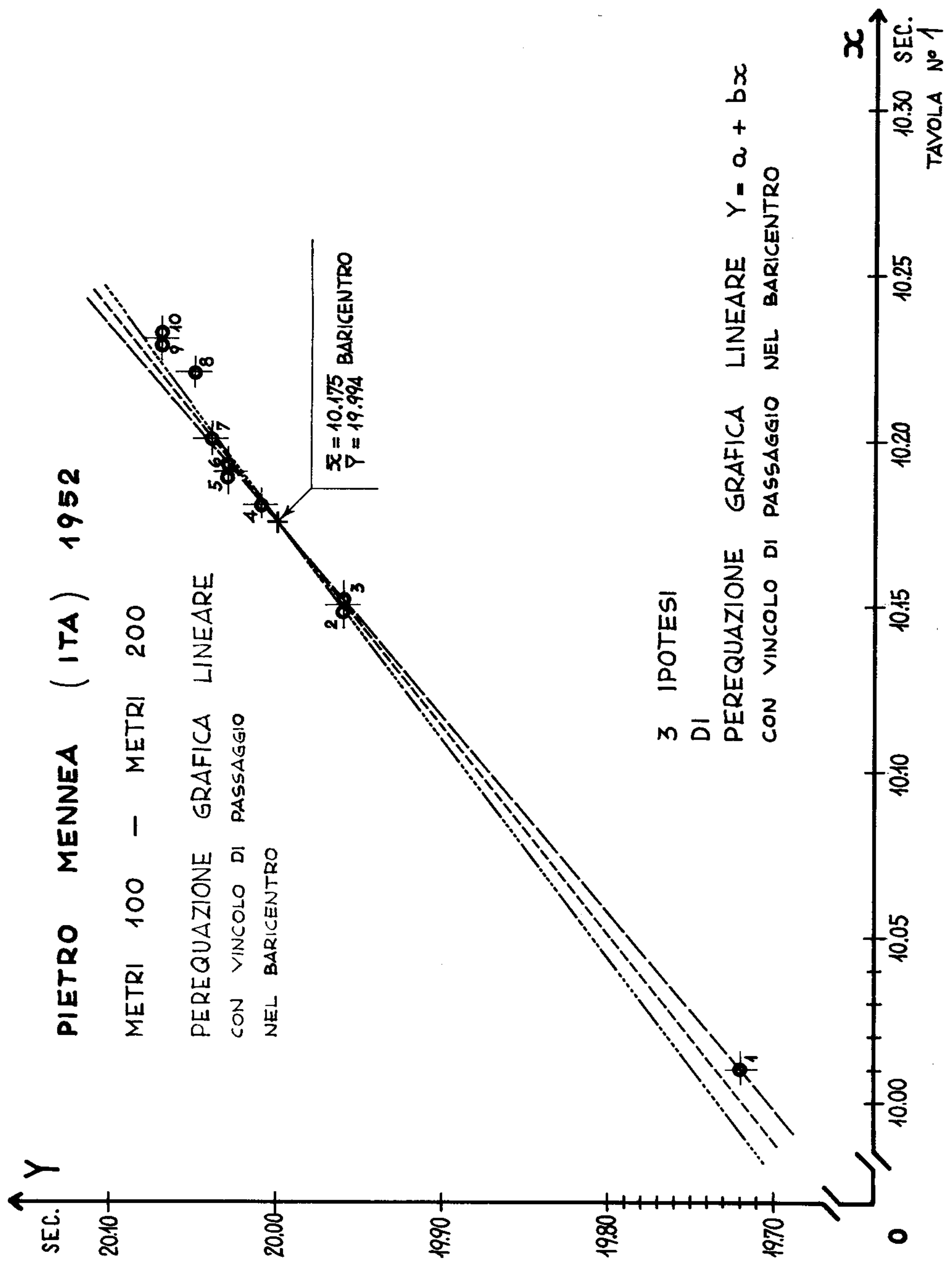
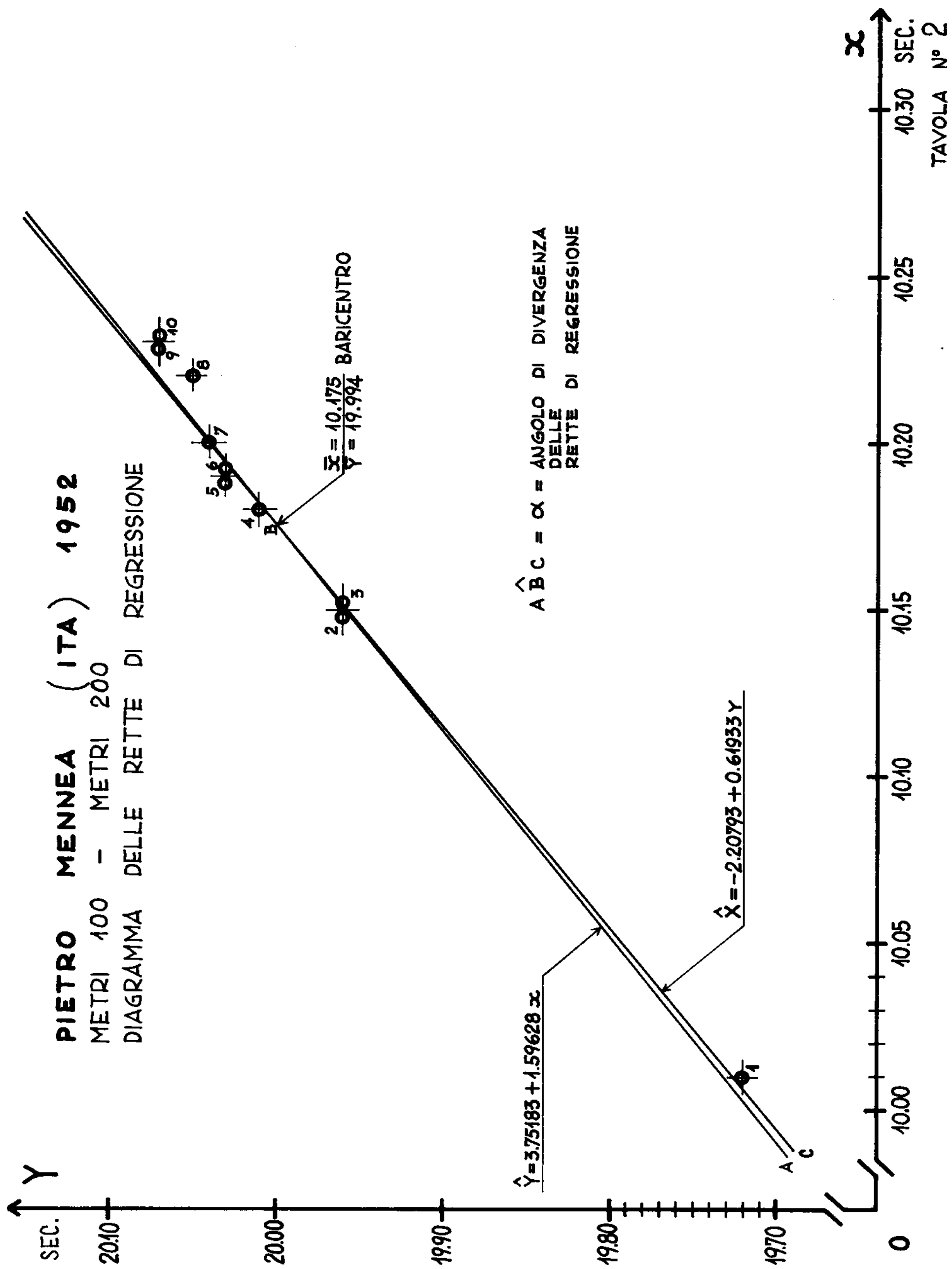

PIETRO MENNEA (ITA) 1952
METRI 100 – METRI 200
PEREQUAZIONE GRAFICA LINEARE
CON VINCOLO DI PASSAGGIO
NEL BARICENTRO
3 IPOTESI
DI
PEREQUAZIONE GRAFICA LINEARE Y = a + bx
CON VINCOLO DI PASSAGGIO NEL BARICENTRO
X̄ = 10.175
Ȳ = 19.994
BARICENTRO
Y
SEC.
20.10
20.00
19.90
19.80
19.70
0
X
SEC.
10.00
10.05
10.10
10.15
10.20
10.25
10.30
TAVOLA N° 1

Tavola n. 2


PIETRO MENNEA (ITA) 1952
METRI 100 - METRI 200
DIAGRAMMA DELLE RETTE DI REGRESSIONE
$\bar{x} = 10.175$
$\bar{y} = 19.994$ BARICENTRO
$A\hat{B}C = \alpha =$ ANGOLO DI DIVERGENZA DELLE RETTE DI REGRESSIONE
$\hat{Y} = 3.75183 + 1.59628\,x$
$\hat{X} = -2.20793 + 0.61933\,Y$
TAVOLA N° 2
SEC.
10.00
10.05
10.10
10.15
10.20
10.25
10.30
X
Y
19.70
19.80
19.90
20.00
20.10
0

un'idea abbastanza precisa del tipo di connessione dei dati delle variabili. Le equazioni delle rette:

$$\hat{Y}_i = 3.75183+1.59628\ x_i$$
$$\hat{X}_i = -2.20793+0.61933\ y_i$$

ci forniscono nelle colonne 6 e 7 della medesima Tabella n. 6 anche i valori teorici delle variabili che, come è facile constatare, rispondono egregiamente alla realtà effettiva del caso Mennea.

Le analisi di regressione, come è noto, sono state affrontate per la prima volta dallo statistico e naturalista inglese Francis Galton (1822-1911) nello studio di alcuni caratteri di ereditarietà umana. Precisamente, il Galton aveva osservato che la statura dei figli tendeva a variare nella stessa direzione dei propri genitori, ma con minore intensità, nel senso che i figli di coppie la cui statura era superiore (inferiore) al valore medio della popolazione erano anch'essi più alti (meno alti), ma il loro scarto medio dalla media era minore di quello dei genitori; pertanto, i figli tendevano a *regredire* verso la media.

Nel nostro caso sportivo, i valori delle variabili esplicative sono naturalmente riferiti alle performances degli atleti nelle relative coppie di gare prese in esame, per cui sarà possibile, analogamente, stabilire eventuali nessi di interrelazione fra le variabili stesse. L'analisi di regressione, però, permette solo una descrizione sintetica, per quanto abbastanza soddisfacente, del *tipo* di relazione. Il problema della sua *misura*, come vedremo in seguito, sarà risolto dall'impiego dei coefficienti di correlazione e determinazione.

Giova aggiungere, peraltro, che l'anasisi della regressione implica, *ma non dimostra*, l'esistenza di un nesso di causalità fra la variabile indipendente, x, e la variabile dipendente, y. Di contro, l'analisi della correlazione, che esclude qualsiasi implicazione di nessi causali, esprime *soltanto* la misura del grado di connessione rilevata nell'analisi di regressione. In effetti, nulla esclude che le due variabili considerate possano essere più o meno fortemente correlate non autonomamente, ma in conseguenza dell'azione o influenza esercitata su entrambe da parte di una ulteriore variabile. In questo senso, l'analisi di correlazione si dimostra strumento statistico molto meno potente, seppure complementare per eccellenza, dell'analisi di regressione, la quale, nel nostrò caso sportivo specifico, pur dovendo necessariamente escludere rapporti causali di tipo strutturale (funzionale) fra le due gare in esame, è tuttavia in grado di *codificare*, per via sperimentale, la *legge di associazione* delle performances dell'atleta.

È noto che uno dei metodi più efficaci per stimare i parametri di una certa funzione che interpreti il senso di un insieme di osservazioni empiriche è quello dei minimi quadrati, di cui si è già parlato in precedenti occasioni nelle quali, però, non si era accennato ad un particolare problema: su quale tipo di scarto fra valori effettivi e teorici si deve operare, una volta posto il vincolo del minimo della somma dei quadrati degli scarti in questione? Esaminiamo il diagramma dimostrativo della Tavola n. 2/bis approntata per lo scopo, visto che quello della Tavola n. 2 relativo alle performances di Mennea non consente, per il ridottissimo grado di dispersione delle prestazioni dell'atleta rispetto ai valori teorici delle rette di regressione, una chiara rappresentazione grafica dei tipi di scarto. Come si può notare, esistono tre diverse interpretazioni analitiche dello scarto fra valore effettivo e valore teorico, e precisamente:

a) scarto verticale (segmento $\overline{AB}$)

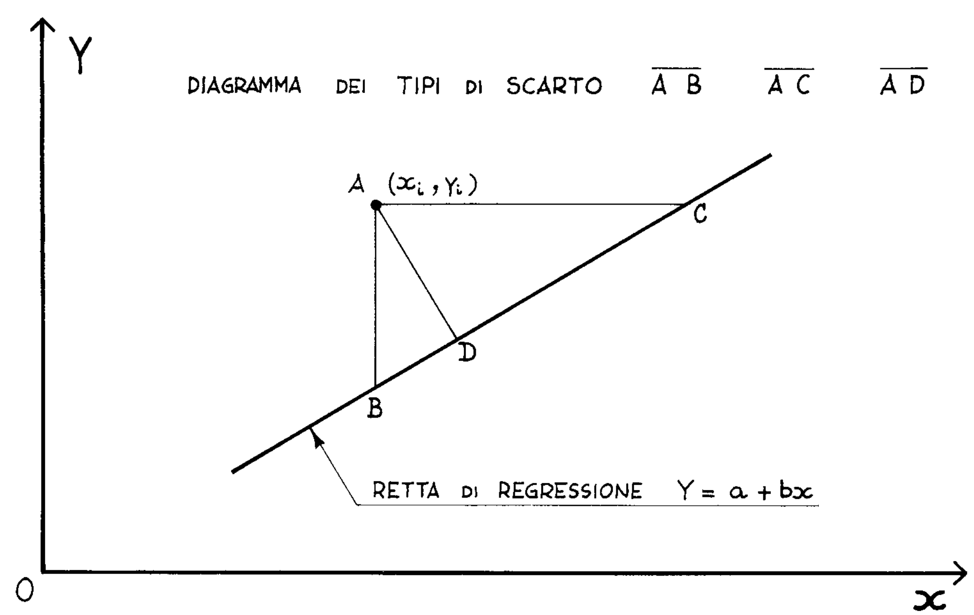
b) scarto orizzontale (segmento $\overline{AC}$)

c) scarto perpendicolare alla retta di regressione (segmento $\overline{AD}$).

Se viene ritenuta indipendente la variabile x (relativa, nel nostro caso, alla gara dei 100 metri), gli scarti di cui necessita minimizzare la somma dei quadrati sono evidentemente quelli verticali del tipo $\overline{AB}$ (per la precisione, si tenga presente che la teoria della regressione nasce da questa ipotesi); in questo caso si è in presenza di regressione di Y su X. Nel caso venga invece ritenuta indipendente la variabile y (relativa alla gara dei 200 metri), gli scarti in questione sono quelli orizzonali del tipo $\overline{AC}$; in questo caso, allora, si è in presenza di regressione di X su Y. Infine, se non si hanno motivi sufficienti per considerare indipendente una delle due variabili (come già accennato, è il nostro caso specifico) il fenomeno può essere analizzato tramite la definizione di due distinte rette di regressione, procedimento che, nella fattispecie, permette di evitare il calcolo degli scarti ortogonali del tipo $\overline{AD}$, che risulta molto laborioso dal punto di vista analitico ed anche criticabile da quello teorico. Per la natura dei fenomeni sportivi presi in esame e per alcuni vantaggi pratici operativi, collegati alla successiva analisi di correlazione (calcolo dei coefficienti di determinazione in funzione dei parametri angolari delle due rette di regressione), abbiamo considerato conveniente tale soluzione, anche se l'analisi di regressione secondo il metodo di Wald-Bartlett avrebbe comportato la definizione di una sola retta di regressione, ma senza il vincolo della minimizzazione del quadrato degli scarti, pur comprendendo quello del passaggio della retta nel punto di baricentro relativo alle due medie aritmetiche $\bar{x}$ e $\bar{y}$.

Comunque, dell'equazione di regressione lineare ricavata secondo il metodo di Wald-Bartlett ce ne serviremo molto probabilmente in futuro, stante la sua flessibilità e praticità di impiego. Tuttavia, non sarà inutile sottolineare che l'adozione del metodo dei minimi quadrati è normalmente preferibile in quanto, comportando la penalizzazione delle *deviazioni* più grandi rispetto a quelle più piccole, può consentire una ottimale definizione della retta di regressione.

Tavola n. 2/bis



2.1.3 Analisi di correlazione

Abbiamo già visto nel diagramma cartesiano della Tavola n. 2 che le 10 coppie ordinate delle performances conseguite da Mennea nelle due gare in esame si situano nelle estreme vicinanze delle rette di regressione, avallando, pertanto, la scelta del tipo di regressione. Nel diagramma, anzi, si può osservare che le due rette in questione sono quasi al limite della sovrapposizione, significando con ciò un elevatissimo grado di relazione (in senso associativo, s'intende) fra le due variabili.

Sorge naturale, a questo punto, l'esigenza della *costruzione* di un indice analitico che dia rigorosamente la misura del tipo di relazione già riscontrato valido a livello di regressione. Osserviamo, in proposito, la Tavola n. 3: nel diagramma rappresentato è stata effettuata una doppia traslazione di assi, assegnando come nuova origine il punto di coordinate ($\bar{x}$; $\bar{y}$), ovvero il punto di baricentro delle performances dell'atleta; conseguentemente, le variabili saranno ora espresse in termini di scarto dalle relative medie, precisamente $\hat{x}_i = (x_i - \bar{x})$; $\hat{y}_i = (y_i - \bar{y})$, e le performances in questione risulteranno distribuite secondo la suddivisione in quadranti operata dal nuovo sistema di assi cartesiani. Sarà ora interessante constatare quanto emerge dalla seguente classificazione delle coordinate dei punti nei quattro quadranti in relazione al loro segno algebrico:

	Segno algebrico $\hat{x}_i$	$\hat{y}_i$
I° quadrante	+	+
II° quadrante	−	+
III° quadrante	−	−
IV° quadrante	+	−

I prodotti $\hat{x}_i\,\hat{y}_i$ risulteranno quindi di segno positivo per i punti del I° e III° quadrante, e negativo per quelli del II° e IV°. Conseguentemente, la distribuzione dei punti nei quadranti potrà essere *descritta* dalla somma algebrica:

$$\sum_{i=1}^{n} \hat{x}_i\,\hat{y}_i\,.$$

Con la prevalenza delle osservazioni nel I° e III° quadrante la predetta somma algebrica risulterà positiva, per cui i punti si troveranno più o meno allineati lungo una retta passante per detti quadranti, il che significa chiaramente presenza di correlazione positiva. Nel caso di

$$\sum_{i=1}^{n} \hat{x}_i\,\hat{y}_i < 0$$

la correlazione sarà evidentemente negativa, per la prevalenza dei punti nel II° e IV° quadrante. Il caso di

$$\sum_{i=1}^{n} \hat{x}_i\,\hat{y}_i = 0,$$

invece, implica equa distribuzione dei punti nei quattro quadranti, per cui con valori di $\hat{x}_1$ (ovvero: $\hat{y}_i$) sempre maggiori o sempre minori, la $\hat{y}_i$ (ovvero: $\hat{x}_i$) tenderà alla stazionarietà, con ciò escludendo l'ipotesi di correlazione lineare fra le variabili (ma non necessariamente quella curvilinea).

A questo punto, potrebbe sembrare opportuno assumere il parametro

$$\sum_{i=1}^{n} \hat{x}_i\,\hat{y}_i$$

come misura della correlazione fra le due variabili; c'è da osservare, però, che tale misura è direttamente influenzata dalla numerosità delle osservazioni in quanto, ferma restando la misura della correlazione, un eventuale raddoppio delle frequenze, in costanza degli stessi dati, moltiplicherebbe automaticamente per 2 anche il parametro

$$\sum_{i=1}^{n} \hat{x}_i\,\hat{y}_i\,.$$

La media aritmetica dei prodotti degli scarti permette di superare l'inconveniente. Pertanto si avrà:

$$\frac{\sum_{i=1}^{n} \hat{x}_i\,\hat{y}_i}{n}\,,$$

Tavola n. 3

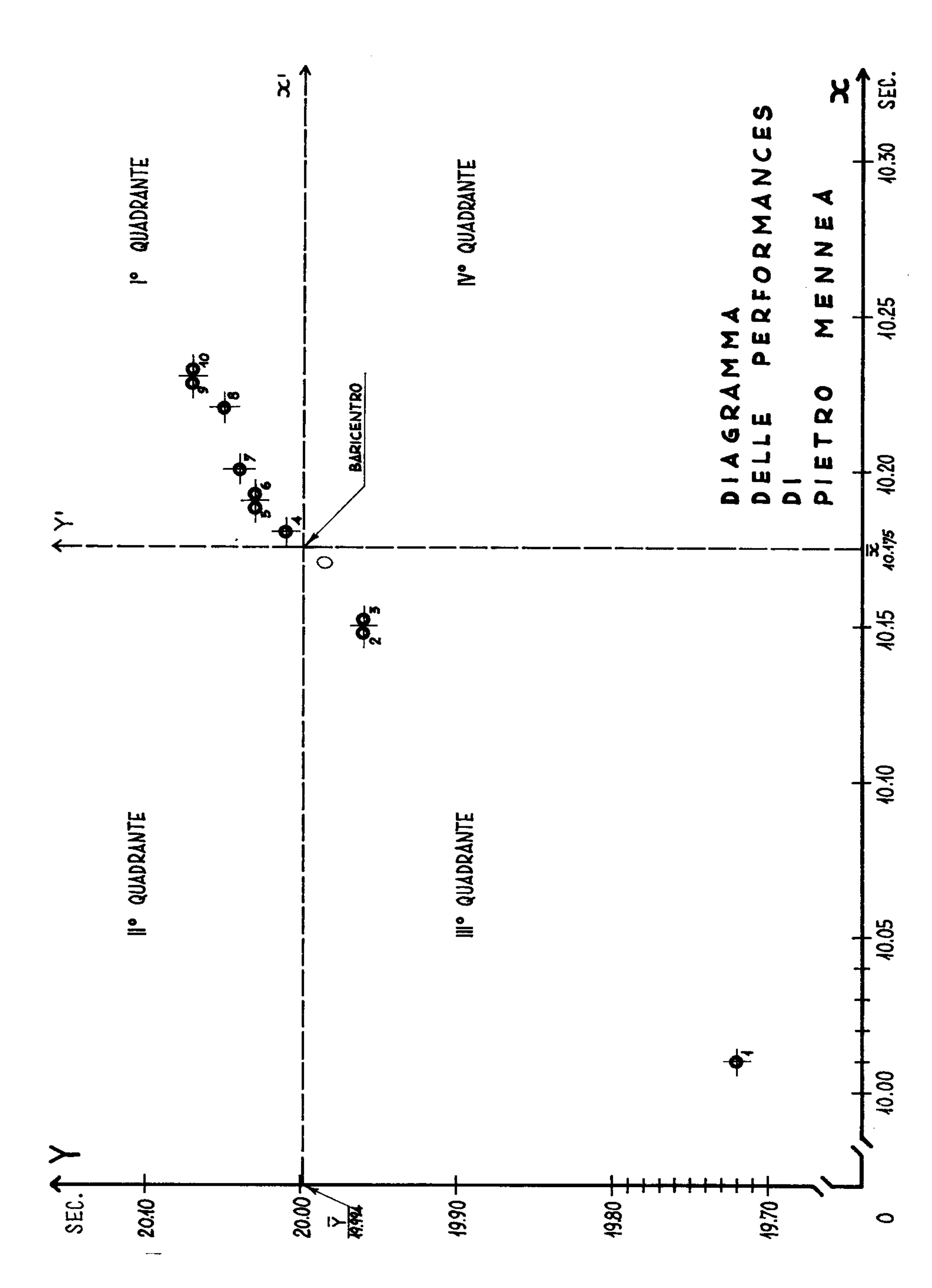


DIAGRAMMA
DELLE PERFORMANCES
DI
PIETRO MENNEA
I° QUADRANTE
II° QUADRANTE
III° QUADRANTE
IV° QUADRANTE
BARICENTRO
X
X'
Y
Y'
O
SEC.
10.00
10.05
10.10
10.15
10.20
10.25
10.30
10.175
20.10
20.00
19.90
19.80
19.70
19.994
0

che è il valore di *covarianza* delle due variabili. In considerazione che anche la covarianza è assoggettata all'unità di misura (con i relativi inconvenienti), il problema viene superato definitivamente con il rapporto covarianza/prodotto delle deviazioni-standard.

Avremo allora, in forma compatta, la formula:

r = coefficiente di correlazione lineare

$= \frac{s_{xy}}{s_x\, s_y}$ dove il numeratore rappresenta la quantità:

$$\frac{1}{n} \sum_{i=1}^{n} \hat{x}_i\, \hat{y}_i \quad \text{ovvero:}$$

$$\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

cioè la covarianza fra le x_i e le y_i, ed il denominatore la quantità:

$$\sqrt{\left(\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2\right)\left(\frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2\right)}$$

ovvero: radice quadrata del prodotto delle deviazioni-standard secondo la già citata 2ª formula abbreviata.

Pertanto, la formula operativa del coefficiente di correlazione ricavata dalla poc'anzi esposta analisi dei segni algebrici dei punti nello scatter a quadranti sarà data da:

$$\frac{\frac{1}{n} \sum_{i=1}^{n} \hat{x}_i\, \hat{y}_i}{\sqrt{\left(\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2\right)\left(\frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2\right)}},$$

cioè: r =

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2\right)\left(\frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2\right)}}.$$

Chiarito ora, però, il significato di detta analisi ci serviremo, per il calcolo relativo alle performances di Mennea, del più pratico coefficiente di correlazione ricavato in funzione dei parametri angolari delle due rette di regressione (per la cui dimostrazione di uguaglianza si rimanda ad un testo di statistica di base).

La formula in questione è data da: $r = b_x\, b_y$. In effetti, la dimostrazione di cui si è detto porta alla conclusione che:

$$b_x = r\frac{s_x}{s_y} \qquad \text{e} \qquad b_y = r\frac{s_y}{s_x}\ ;$$

per cui sarà:

$$r^2\, \frac{s_x}{s_y}\, \frac{s_y}{s_x} = b_x\, b_y\,, \quad \text{e quindi:}$$

$r^2 = b_x\, b_y$, da cui, appunto, discende:

$$|\, r\,| = \sqrt{b_x\, b_y}.$$

Il coefficiente di correlazione lineare così calcolato, pertanto, assume anche un significato geometrico, in quanto *r* rappresenta la media geometrica in valore assoluto, dei parametri angolari (o gradienti) delle due rette di regressione, con la conseguente conferma che un indice del grado di associazione tra le due variabili in questione può essere posto in diretta relazione all'angolo di divergenza fra le rette stesse.

Dalle due rette di regressione relative alle performances di Mennea ricaviamo

i due parametri angolari per ottenere ora il coefficiente di correlazione lineare cercato. Detti parametri risultano:
1.59628 per la I^a^ retta di regressione
0.61933 per la II^a^ retta di regressione. Si avrà, allora:

$|r|$ = coefficiente di correlazione lineare =

$$= \sqrt{1.59628 \times 0.61933} =$$

$$= \sqrt{0.9886240924} = \quad 0.9942957771$$

$$\cong 0.994 = 99.4\%.$$

In termini sportivi, possiamo finalmente affermare che la misura dell'intensità della relazione lineare (secondo il metodo dei minimi quadrati), come adattamento delle rette di regressione al fenomeno espresso dalle performances dell'atleta, è pari al 99.4%.

La misura della parte di variabilità che viene invece sottratta dal procedimento di regressione, e che pertanto stima *indirettamente* la frazione di variabilità da attribuire all'associazione delle due variabili, nel senso della reciproca sistematica connessione delle performances nelle due gare, è data dal quadrato del coefficiente di correlazione *r*, il quale esprime la percentuale della varianza assorbita rispetto a quella totale emersa dal processo perequatorio. Nella nostra fattispecie, l'influenza reciproca delle corrispettive performances dell'atleta nelle due gare dei metri 100 e 200 piani è descritta, come abbiamo già visto nel calcolo del coefficiente di correlazione, da:
r^2 = coefficiente di determinazione = 0.989 = 98.9%.

Il valore del coefficiente di determinazione r^2 ci permette, quindi, di ricavare un ovvio complemento a 1, cioè $(1 - r^2)$, chiamato coefficiente di alienazione (ovvero di non determinazione), il quale, conseguentemente, misura l'influenza che la presenza e la variazione di altri fattori, diversi da *x* (ovvero: *y*), esercitano su *y* (ovvero: *x*).

Il valore in questione, pertanto, è dato da:
C_a = coefficiente di alienazione =
$= (1 - r^2) = 1 - 0.989 = 0.011 = 1.1\%$,

In riferimento al caso Mennea, si potrà allora affermare che l'1.1% rappresenta l'incidenza percentuale di fattori perturbatori non considerati dal processo di perequazione lineare, nel senso che la specifica analisi di regressione ha preso in esame *solo* e *soltanto* l'influenza dicotomica performance gara 100 metri/performance gara 200 metri.

C'è da aggiungere, peraltro, che è possibile una risposta, pertinente al fenomeno, circa la natura di detti fattori perturbatori del rapporto (equilibrio di corrispondenza) instauratosi fra le due gare praticate dall'atleta, naturalmente nell'ambito delle relative 10 migliori performances. La risposta in questione è a livello generale ed è da porsi in relazione all'insieme di tutti gli eventuali fattori potenzialmente in grado di influire nella proporzione indicata, dei quali proviamo a dare una possibile elencazione, ovviamente incompleta e senza ordine d'importanza, ma certamente significativa:
1) condizione climatica ed ambientale (temperatura, percentuale di umidità, vento, pioggia, altitudine sul livello del mare, ecc.)
2) grado d'importanza della manifestazione atletica
3) livello della condizione di forma dell'atleta
4) tipo del fondo della pista
5) grado di refrattarietà (reattività) alle condizioni psico-fisiche e caratteriali contrarie dell'atleta (stato di tensione nervosa pre-agonistica, degrado del rapporto di efficienza *allenamento/risultato tecnico,* mancanza di serenità personale e/o familiare, postumi di infortuni, ecc.)
6) applicazione non corretta del regime dietetico
ai quali, peraltro, vanno aggiunti numerosi altri fattori di cui non necessita, in questa sede, la descrizione.

È d'obbligo precisare, però, che la presente indicativa elencazione di possibili fattori perturbatori deve, naturalmente, essere interpretata in senso parziale (purtroppo non quantificabile), nell'accezione specifica di parte complementare rispetto a quella — preponderante — che rientra, per così dire, nel normale *corpo* delle componenti della prestazione atletica.

La seguente Tabella n. 7, che riassume i dati essenziali relativi al fenomeno sportivo indagato, è stata predisposta per consentire, con quelle corrispondenti agli altri atleti, una sintetica ma efficace comparazione tecnica e statistica fra tutte le prestazioni atletiche considerate nell'indagine.

TABELLA N. 7 - CALCOLO DEI PARAMETRI - PIETRO MENNEA (ITA) 1952 - METRI 100/METRI 200

1 - DEVIAZIONE-STANDARD

a) Metri 100:

$$s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{1035.3439}{10} - (10.175)^2} = \sqrt{103.53439 - 103.530625} =$$

$$= \sqrt{0.003745} = 0.061196405 \cong 0.0612 \text{ secondi} = 6.12 \text{ centesimi di secondo}$$

b) Metri 200:

$$s = \sqrt{\frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2} = \sqrt{\frac{3997.6974}{10} - (19.994)^2} = \sqrt{399.76974 - 399.760036} =$$

$$= \sqrt{0.009704} = 0.0985088828 \cong 0.0985 \text{ secondi} = 9.85 \text{ centesimi di secondo}$$

2 - COEFFICIENTE DI VARIAZIONE

a) Metri 100:

$$V = \frac{0.0612}{10.175} \cong 0.0060 = 0.60\% \text{ dell'intensità media del fenomeno}$$

b) Metri 200:

$$V = \frac{0.0985}{19.994} \cong 0.0049 = 0.49\% \text{ dell'intensità media del fenomeno}$$

3 - COEFFICIENTE DI DETERMINAZIONE

$$r^2 = b_{yx} \times b_{xy} = 1.59628 \times 0.61933 = 0.9886240924 \cong 0.989 = 98.9\%$$

4 - COEFFICIENTE DI CORRELAZIONE

$$r = \sqrt{0.9886240924} = 0.9942957771 \cong 0.994 = 99.4\%$$

5 - COEFFICIENTE DI ALIENAZIONE

$$C_a = (1 - r^2) = 1 - 0.989 = 0.011 = 1.1\%$$

Equazioni di regressione: $\hat{Y}_i = 3.75183 + 1.59628\, x_i$

$\hat{X}_i = -2.20793 + 0.61933\, y_i$

Con l'analisi-guida condotta sulle performances di Pietro Mennea si conclude la prima parte della ricerca. Nel numero successivo di Atleticastudi, saranno trattate le problematiche relative agli altri atleti scelti per l'indagine, precisando che — qualora i risultati di Seoul '88 ne modificassero la sostanza — accanto all'analisi originaria ante-olimpiade, sarà proposta, per gli atleti interessati, quella relativa alla nuova situazione tecnica creatasi nel corso della manifestazione olimpica.

(continua)