

Introduzione agli L -momenti

Nelle procedure di analisi di frequenza, e di analisi di frequenza regionale, si adattano ai dati delle distribuzioni la cui forma si ritiene conosciuta a meno di un numero finito di parametri incogniti. I momenti campionari ordinari, in particolare media, scarto, skewness e kurtosis, sono spesso utilizzati per la stima dei parametri delle distribuzioni di probabilità. Hosking and Wallis [1997] suggeriscono invece di utilizzare, al posto dei momenti ordinari, gli L -momenti perché adatti a descrivere più distribuzioni, perché più robusti nella stima da campioni poco consistenti di dati in presenza di outliers e perché meno soggetti a distorsione nella stima. In questa appendice, tratta da Hosking and Wallis [1997], si definiscono gli L -momenti in maniera formale. Dopo una breve introduzione sui concetti di distribuzione di probabilità, di stimatori dei parametri e di momenti, vengono definiti gli L -momenti, si discutono alcune loro proprietà, le differenze rispetto ai momenti ordinari ed il loro utilizzo nella stima dei parametri delle distribuzioni (argomento approfondito in Capitolo 4).

3.1 Momenti

La forma di una distribuzione di probabilità può essere descritta dai *momenti della distribuzione*, che sono la *media* $\mu = E(X)$ e i *momenti di ordine superiore* $\mu_r = E(X - \mu)^r$, con $r = 2, 3, \dots$. La media definisce il baricentro della distribuzione (si veda l'Equazione ??). La dispersione della distribuzione intorno alla media può essere misurata con la *deviazione standard*

$$(3.1) \quad \sigma = \mu_2^{1/2} = \{E(X - \mu)^2\}^{1/2},$$

o con la varianza, $\sigma^2 = \text{var}(X)$. Il *coefficiente di variazione* (CV), $C_v = \sigma/\mu$, esprime la dispersione della distribuzione adimensionalizzata con la media. Spesso si utilizzano momenti adimensionalizzati di ordine superiore $\mu_r/\mu_2^{r/2}$, in particolare lo *skewness* (asimmetria)

$$(3.2) \quad \gamma = \mu_3/\mu_2^{3/2},$$

ed il *kurtosis*

$$(3.3) \quad \kappa = \mu_4 / \mu_2^2 .$$

Quantità analoghe a queste possono essere calcolate da un campione di dati x_1, x_2, \dots, x_n . La *media campionaria*

$$(3.4) \quad \bar{x} = n^{-1} \sum_{i=1}^n x_i ,$$

è lo stimatore naturale di μ . I momenti di ordine superiore

$$(3.5) \quad m_r = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^r ,$$

sono ragionevoli stimatori di μ_r , ma non sono indistorti. Al loro posto vengono spesso usati stimatori indistorti come, nel caso particolare di σ^2 , μ_3 e $\kappa_4 = \mu_4 - 3\mu_2^2$, sono, rispettivamente,

$$(3.6) \quad s^2 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2 ,$$

$$(3.7) \quad \tilde{m}_3 = \frac{n^2}{(n-1)(n-2)} m_3 ,$$

$$(3.8) \quad \tilde{k}_4 = \frac{n^2}{(n-2)(n-3)} \left\{ \left(\frac{n+1}{n-1} \right) m_4 - 3m_2^2 \right\} .$$

La deviazione standard campionaria, $s = \sqrt{s^2}$, è uno stimatore di σ ma non è indistorto. Gli stimatori campionari del CV, dello skewness e del kurtosis sono, rispettivamente,

$$(3.9) \quad \hat{C}_v = s/\bar{x} , \quad g = \tilde{m}_3/s^3 \quad k = \tilde{k}_4/s^4 + 3 .$$

Gli stimatori dei momenti hanno alcune proprietà indesiderabili. Ad esempio, gli stimatori g e k possono essere molto distorti in quanto sono caratterizzati da limiti algebrici che dipendono dalla lunghezza del campione; per un campione di n dati si ha che $|g| \leq n^{1/2}$ mentre $k \leq n+3$ [Hosking and Wallis, 1997]. Da ciò consegue che, se la distribuzione è particolarmente asimmetrica, potrebbe essere impossibile calcolare questa skewness da un campione di dimensioni fisse. L'inferenza di distribuzioni asimmetriche basata sui momenti campionari può essere estremamente poco affidabile. Per questo motivo in questo lavoro vengono utilizzate altre misure della forma delle distribuzioni, gli L -momenti di Hosking e Wallis.

3.2 L -momenti delle distribuzioni di probabilità

Gli L -momenti sono un sistema alternativo di descrivere la forma delle distribuzioni di probabilità. Storicamente essi nascono come modifica dei *momenti pesati in probabilità* di Greenwood et al. [1979]. I momenti pesati in probabilità di una variabile casuale X con distribuzione di frequenza cumulata $F(\cdot)$ sono le quantità

$$(3.10) \quad M_{p,r,s} = E[X^p \{F(X)\}^r \{1 - F(X)\}^s] .$$

Particolarmente utili sono i momenti pesati in probabilità $\alpha_r = M_{1,0,r}$ e $\beta_r = M_{1,r,0}$. Per una distribuzione caratterizzata da una funzione dei quantili $x(u)$, dalle Equazioni ?? e 3.10 si ottiene

$$(3.11) \quad \alpha_r = \int_0^1 x(u)(1-u)^r du, \quad \beta_r = \int_0^1 x(u)u^r du .$$

Queste equazioni possono essere paragonate alla definizione dei momenti ordinari, che può essere scritta anche come

$$(3.12) \quad E(X^r) = \int_0^1 \{x(u)\}^r du .$$

Mentre i momenti ordinari considerano successive elevazioni di potenza della funzione dei quantili $x(u)$, i momenti pesati in probabilità considerano successive elevazioni di potenza di u oppure $1-u$ e possono essere visti come integrali di $x(u)$ pesati con i polinomi u^r oppure $(1-u)^r$.

I momenti pesati in probabilità α_r e β_r sono stati usati in letteratura come base di metodi per la stima dei parametri delle distribuzioni di probabilità ma sono difficilmente interpretabili come misure di scala e forma di queste. Queste informazioni sono contenute in certe combinazioni lineari dei momenti pesati in probabilità. Ad esempio, multipli di $\alpha_0 - 2\alpha_1$ o $2\beta_1 - \beta_0$ sono stime dei parametri di scala delle distribuzioni, mentre lo skewness può essere misurato da $6\beta_2 - 6\beta_1 + \beta_0$. Queste combinazioni lineari derivano naturalmente dall'integrazione di $x(u)$ pesata non con i polinomi u^r o $(1-u)^r$, ma con un set di polinomi ortogonali.

Si definiscano *polinomi di Legendre sfasati* (perché definiti nell'intervallo $[0, 1]$ invece che nell'intervallo $[-1, +1]$) i polinomi $P_r^*(u)$, con $r = 0, 1, 2, \dots$, che godono delle seguenti proprietà:

- (i) $P_r^*(u)$ è un polinomio di grado r in u ;
- (ii) $P_r^*(1) = 1$;
- (iii) $\int_0^1 P_r^*(u)P_s^*(u)du = 0$ se $r \neq s$ (condizione di ortogonalità).

I polinomi di Legendre sfasati hanno forma esplicita

$$(3.13) \quad P_r^*(u) = \sum_{k=0}^r p_{r,k}^* u^k ,$$

dove

$$(3.14) \quad p_{r,k}^* = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} = \frac{(-1)^{r-k} (r+k)!}{(k!)^2 (r-k)!} .$$

Gli L -momenti di una variabile casuale X con funzione dei quantili $x(u)$ sono definiti come

$$(3.15) \quad \lambda_r = \int_0^1 x(u) P_{r-1}^*(u) du .$$

In termini di momenti pesati in probabilità, gli L -momenti sono dati da

$$(3.16) \quad \begin{aligned} \lambda_1 &= \alpha_0 & &= \beta_0 , \\ \lambda_2 &= \alpha_0 - 2\alpha_1 & &= 2\beta_1 - \beta_0 , \\ \lambda_3 &= \alpha_0 - 6\alpha_1 + 6\alpha_2 & &= 6\beta_2 - 6\beta_1 + \beta_0 , \\ \lambda_4 &= \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 & &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 , \end{aligned}$$

e, in generale,

$$(3.17) \quad \lambda_{r+1} = (-1)^r \sum_{k=0}^r p_{r,k}^* \alpha_k = \sum_{k=0}^r p_{r,k}^* \beta_k .$$

È conveniente definire le versioni adimensionali degli L -momenti, cosa che si può ottenere dividendo gli L -momenti di ordine superiore per la misura di scala λ_2 . Si ottengono così i *rapporti degli L -momenti*

$$(3.18) \quad \tau_r = \lambda_r / \lambda_2 , \quad r = 3, 4, \dots$$

che misurano la forma di una distribuzione indipendentemente dalla scala. Si definisce inoltre il coefficiente di L -variazione (L -CV) come

$$(3.19) \quad \tau = \lambda_2 / \lambda_1 ,$$

quantità analoga al coefficiente di variazione CV.

3.3 Proprietà degli L -momenti

Gli L -momenti λ_1 e λ_2 , l' L -CV τ e i rapporti degli L -momenti τ_3 e τ_4 sono le quantità che Hosking and Wallis [1997] consigliano di utilizzare per descrivere le distribuzioni di probabilità.

Le loro più importanti proprietà sono:

- **Esistenza.** Se esiste la media della distribuzione, allora esistono tutti i suoi L -momenti.
- **Unicità.** Se esiste la media della distribuzione, allora gli L -momenti definiscono tale distribuzione in maniera univoca, ovvero non esistono due distribuzioni diverse con gli stessi L -momenti.

- **Terminologia.** Gli L -momenti (ed i rapporti degli L -momenti) che si utilizzano hanno un determinato significato, paragonabile a quello dei momenti campionari: λ_1 è la L -posizione (o la media) della distribuzione; λ_2 è l' L -scala; τ è l' L -CV; τ_3 è l' L -skewness; τ_4 è l' L -kurtosis.
- **Limiti algebrici.** λ_1 può assumere qualsiasi valore; $\lambda_2 \geq 0$; per una distribuzione che assume solo valori positivi $0 \leq \tau < 1$; i rapporti degli L -momenti soddisfano l'uguaglianza $|\tau_r| < 1$ per ogni $r \geq 3$. Limiti più precisi possono essere trovati per ogni τ_r : ad esempio, dato τ_3 , allora $(5\tau_3^2 - 1)/4 \leq \tau_4 < 1$ e, per distribuzioni che assumono solo valori positivi, dato τ si ha che $2\tau - 1 \leq \tau_3 < 1$.
- **Trasformazioni lineari.** Siano X e Y due variabili casuali con L -momenti λ_r e λ_r^* rispettivamente, e si supponga che $Y = aX + b$. Allora $\lambda_1^* = a\lambda_1 + b$; $\lambda_2^* = |a|\lambda_2$; $\tau_r^* = (\text{sign}(a))^r \tau_r$ per $r \geq 3$.
- **Simmetria.** Sia X una variabile casuale simmetrica con media μ , ossia $\Pr[X \geq \mu + x] = \Pr[X \leq \mu - x]$ per ogni x . Allora tutti i rapporti degli L -momenti di ordine dispari valgono 0, ovvero $\tau_r = 0$ se $r = 3, 5, 7, \dots$.

Gli L -momenti sono stati calcolati per molte distribuzioni (si veda l'Appendice ??). La distribuzione che gioca un ruolo centrale nella teoria degli L -momenti, analoga alla distribuzione Normale nella teoria dei momenti ordinari, è la distribuzione uniforme. Si può dimostrare che tutti gli L -momenti λ_r e rapporti degli L -momenti τ_r di ordine superiore (con $r \geq 3$) valgono zero per la distribuzione uniforme. La distribuzione Normale, per il fatto che è simmetrica, presenta gli L -momenti di ordine dispari nulli, ma quelli di ordine pari non sono particolarmente semplici: ad esempio $\tau_4 \approx 0.123$. La distribuzione esponenziale, invece, ha dei rapporti degli L -momenti particolarmente semplici: $\tau_3 = 1/3$, $\tau_4 = 1/6$.

Un modo conveniente per rappresentare gli L -momenti di diverse distribuzioni è il *diagramma dei rapporti degli L -momenti*, esemplificato in Figura 3.1. Questo diagramma mostra gli L -momenti in un grafico i cui assi sono l' L -skewness e l' L -kurtosis. Una distribuzione a due parametri, caratterizzata da un parametro di posizione ed uno di scala, viene rappresentata sul diagramma da un punto. Infatti se due distribuzioni differiscono solo nei parametri di posizione e di scala, allora sono distribuzioni di due variabili casuali X e $Y = aX + b$ con $a > 0$, per cui, dato la proprietà delle trasformazioni lineari degli L -momenti ($\tau_r^* = (\text{sign}(a))^r \tau_r$), hanno gli stessi L -skewness ed L -kurtosis. Una distribuzione a tre parametri, invece, dal momento che è caratterizzata dai parametri di posizione, scala e forma, viene rappresentata sul diagramma da

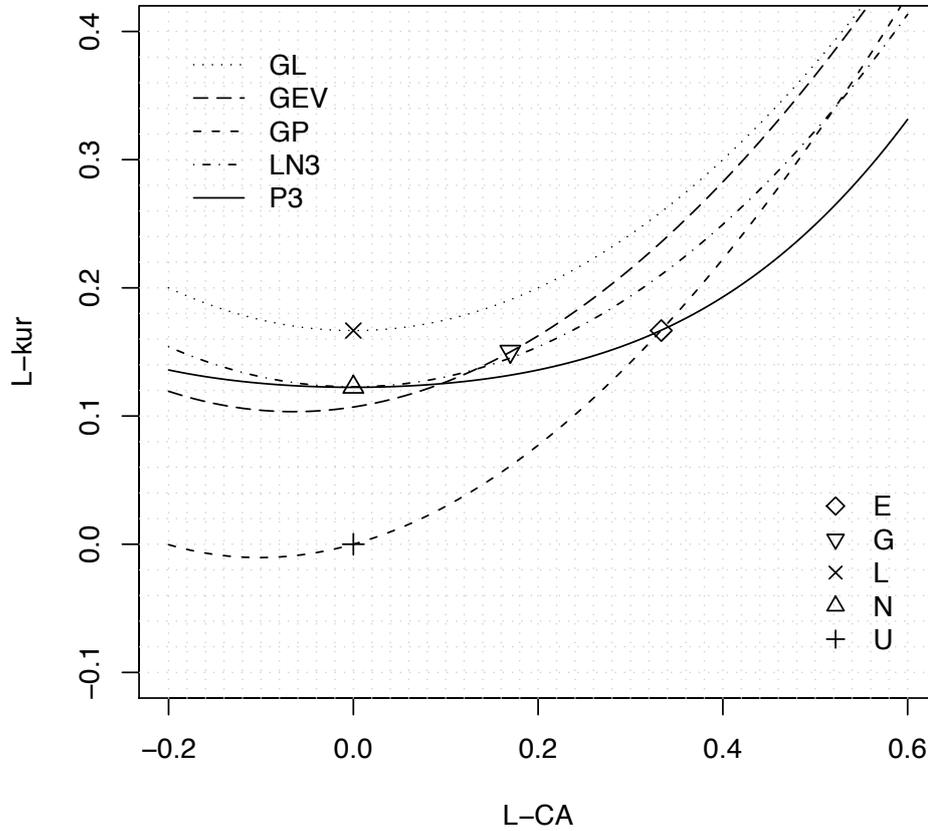


Figura 3.1: Diagramma dei rapporti degli L -momenti. Le distribuzioni a due ed a tre parametri sono riportate come punti e come linee rispettivamente. Le distribuzioni a due parametri sono: esponenziale (E), Gumbel (G), lognormale (L), normale (N) ed uniforme (U); quelle a tre parametri sono: logistica generalizzata (GL), generalizzata del valore estremo (GEV), Pareto generalizzata (GP), lognormale a 3 parametri (LN3) e Pearson tipo III (P3).

una linea, i cui punti corrispondono a differenti valori del parametro di forma. Distribuzioni con più di un parametro di forma generalmente ricoprono aree bidimensionali sul diagramma.

3.4 L -momenti campionari

Gli L -momenti sono stati definiti per una distribuzione di probabilità, ma nella pratica devono essere stimati a partire da campioni finiti. La loro stima è basata su un campione di lunghezza n , ordinato in senso crescente: $x_{1:n} \leq x_{2:n} \leq \dots \leq x_{n:n}$. È conveniente cominciare dalla definizione dello stimatore del momento pesato in probabilità β_r . Uno stimatore indistorto di β_r è

$$(3.20) \quad b_r = n^{-1} \binom{n-1}{r}^{-1} \sum_{j=r+1}^n \binom{j-1}{r} x_{j:n} ,$$

ovvero

$$(3.21) \quad b_r = n^{-1} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n} .$$

Analogamente alle Equazioni 3.16 e 3.17, gli L -momenti campionari sono definiti come

$$(3.22) \quad \begin{aligned} l_1 &= b_0 , \\ l_2 &= 2b_1 - b_0 , \\ l_3 &= 6b_2 - 6b_1 + b_0 , \\ l_4 &= 20b_3 - 30b_2 + 12b_1 - b_0 , \end{aligned}$$

e, in generale,

$$(3.23) \quad l_{r+1} = \sum_{k=0}^r p_{r,k}^* b_k , \quad r = 0, 1, \dots, n-1 ,$$

dove i coefficienti $p_{r,k}^*$ sono definiti dall'Equazione 3.14. L' L -momento campionario l_r è uno stimatore indistorto di λ_r .

Analogamente alle Equazioni 3.18 e 3.19, i rapporti degli L -momenti campionari sono definiti con

$$(3.24) \quad t_r = l_r / l_2 , \quad r = 3, 4, \dots ,$$

e l' L -CV campionario con

$$(3.25) \quad t = l_2 / l_1 .$$

Questi sono gli stimatori naturali di τ_r e di τ rispettivamente, ma non sono indistorti. Ad ogni modo la loro distorsione è molto piccola per campioni di lunghezza moderata o grande. Si può dimostrare che il bias relativo asintotico di t_3 per una distribuzione Gumbel vale $0.19n^{-1}$, e che vale $0.03n^{-1}$ per t_4 di una distribuzione Normale (n è la lunghezza del campione). Hosking and Wallis [1997] hanno calcolato il bias per campioni piccoli estratti da diverse

distribuzioni con metodi di simulazione Monte-Carlo. Ad esempio, si è visto come generalmente il bias di t sia trascurabile per campioni con $n \geq 20$. Per quanto riguarda i bias di t_3 e t_4 , essi sono certamente piccoli in confronto alle differenze negli L -momenti delle diverse famiglie di distribuzioni. Ciò significa che ci si può aspettare che l'identificazione di un tipo di distribuzione generatrice a partire dagli L -momenti di un campione sia comunque indistorta, a prescindere dalla distorsione degli stimatori. Ad ogni modo sia l' L -skewness che l' L -kurtosis sono molto meno distorti dello skewness e del kurtosis ordinari.

3.5 Momenti e L -momenti

Sia i momenti che gli L -momenti sono misure di posizione, scala e forma delle distribuzioni di probabilità. Il parametro di L -posizione λ_1 è la media e, ovviamente, equivale al momento del primo ordine μ . Per quanto riguarda il parametro di L -scala λ_2 , rispetto alla deviazione standard σ vale la relazione $\sigma \geq \sqrt{3}\lambda_2$. Entrambe le misure valutano le differenze tra i valori estratti casualmente da una distribuzione, ma σ dà più peso alle differenze più marcate. La stessa relazione esiste tra CV e L -CV, i cui stimatori soddisfano la disuguaglianza

$$(3.26) \quad \hat{C}_v \geq \left(\frac{3n}{n+1} \right)^{1/2} t.$$

Una relazione di questo genere non può essere scritta per le misure di skewness τ_3 e γ , dal momento che la situazione è molto diversa a seconda della distribuzione considerata. Per quanto riguarda il kurtosis, τ_4 è una misura simile a κ e, come quest'ultima, difficilmente interpretabile (spessore delle code, per alcune distribuzioni). Tuttavia τ_4 dà meno peso alle code estreme della distribuzione, rispetto a κ . Come già accennato sia l' L -skewness che l' L -kurtosis sono molto meno distorti dello skewness e del kurtosis ordinari.

Inoltre gli L -momenti godono della proprietà di esistere alla sola condizione di esistenza della media della distribuzione, il che include anche casi in cui i momenti ordinari non esistono. Ad esempio, per una distribuzione GEV (Appendice ??) i momenti di ordine 3 e 4 non esistono quando il parametro k di forma della distribuzione è inferiore a $-1/3$ e $-1/4$ rispettivamente. Per questi valori di k i rapporti degli L -momenti assumono valori moderati quali $\tau_3 = 0.403$ e $\tau_4 = 0.241$ rispettivamente (e campioni che presentano L -momenti campionari così elevati sono frequenti nell'analisi di dati come velocità del vento o portate di piena).

Un'altro vantaggio dei rapporti degli L -momenti è quello di essere contenuti nell'intervallo $(-1, 1)$, mentre i rapporti dei momenti possono assumere valori arbitrariamente grandi o negativi. Questa proprietà permette di dare un'interpretazione più semplice ai valori di τ_r .

I limiti algebrici dei momenti campionari sono stati menzionati nel Paragrafo 3.1. I rapporti degli L -momenti non sono soggetti a tali restrizioni e i loro stimatori campionari possono assumere qualsiasi valore raggiungibile da quelli teorici.

La differenza principale tra momenti e L -momenti è che i primi danno un peso maggiore alle code estreme delle distribuzioni. Questo può essere visto semplicemente confrontando le Equazioni 3.12 e 3.15. Al crescere di r , il peso assegnato alla coda della distribuzione, $u \approx 1$, cresce come $\{x(u)\}^r$ nell'Equazione 3.12 ma come u^r nell'Equazione 3.15. Per molte distribuzioni $x(u)$ cresce molto più velocemente di u all'avvicinarsi di quest'ultima ad 1; per distribuzioni non limitate superiormente, ovviamente, $x(u) \rightarrow \infty$ se $u \rightarrow 1$. Anche i momenti campionari, di conseguenza, sono più affetti dalle osservazioni estreme degli L -momenti corrispondenti.

3.6 Stima dei parametri mediante gli L -momenti

Un problema che ci si pone comunemente in statistica è la stima, a partire da un campione casuale di n dati, della distribuzione di probabilità la cui specificazione coinvolge un numero finito, p , di parametri incogniti. Analogamente all'usuale metodo dei momenti, il *metodo degli L -momenti* ottiene una stima dei parametri sostituendo i primi p L -momenti campionari ai corrispondenti L -momenti della distribuzione. Ciò richiede di conoscere l'espressione dei parametri in termini degli L -momenti, il che è riportato nell'Appendice ?? per molte distribuzioni standard dell'idrologia.

Le distribuzioni esatte degli stimatori dei parametri ottenuti con il metodo degli L -momenti sono generalmente difficili da derivare, ma approssimazioni per campioni caratterizzati da n elevato possono essere ottenute dalla teoria asintotica. Per molte distribuzioni si è riscontrato che tali stimatori sono asintoticamente distribuiti secondo una distribuzione Normale [Hosking and Wallis, 1997], e si sono derivati gli errori standard e gli intervalli di confidenza.

Sempre Hosking and Wallis [1997] sostengono che spesso, con campioni di lunghezza piccola o moderata, il metodo degli L -momenti è più efficiente di quello della massima verosimiglianza.