Le curve normali di ordine *p* come distribuzioni di errori accidentali: una rassegna dei risultati e problemi aperti per il caso univariato e per quello multivariato[1]

The Normal Distributions of order p as Random Errors Distributions: a Review of present results and open Problems in the Univariate and Multivariate situation

Marcello Chiodi
Istituto di Statistica, Facoltà di Economia di PalermoViale delle Scienze, 90128 – Palermo – Italy; e-mail: chiodi@unipa.it)

Abstract: The normal distributions of order *p* (known also as exponential power distributions) are useful in the description of symmetrical random errors more general than the normal ones. After a brief sketch on the history of this family of curves, some results about the properties of M.L. estimators and tests are reviewed in the general situation of L_p-norm regression, as well as the problem of the estimation of *p*; we underline the importance of simulation studies, besides the analytical approaches, for the approximation of sampling distributions of tests and estimators. In the paper are also reported brief results concerning some multivariate extensions, even if these extensions have been used mainly to specify multivariate non-normal alternatives in testing normality or as departures from normality in studying the robustness of some multivariate tests, but only seldom to study real data sets of multivariate observations.

Parole chiave: Normal distributions of order *p*, exponential power distributions, L_p-norm regression, Monte Carlo studies, multivariate non normal distributions.

1. Introduzione.

Una legge di distribuzione di errori di osservazione più generale di quella normale fu introdotta da Subbotin (1923) in modo abbastanza naturale come distribuzione degli errori di osservazione di natura accidentale e_i , di misure dirette z_i di una quantità incognita m, con $z_i=e_i+m$, facendo solo due ipotesi:

1) la probabilità di un errore e dipende soltanto dalla grandezza dell'errore stesso e può essere espressa da una funzione $f(e)$ dotata di derivata prima continua in generale;
2) il valore più probabile di una quantità m, della quale siano note delle misure dirette z_i, non deve dipendere dall'unità di misura adottata (2° assioma di Schiaparelli);

Tali ipotesi sono più generali di quelle di Gauss che, nella *Theoria Motus corporum coelestium…*, ricavò la normale come legge di distribuzione degli errori di osservazione, ipotizzando che il valore più plausibile (di massima probabilità) da dedursi da un sistema di osservazioni dirette di egual precisione sia la media aritmetica "*se non con assoluta esattezza, almeno con molta approssimazione*" (Pizzetti, 1892). La

[1] Questa ricerca é stata realizzata utilizzando fondi MURST (ex 60%)

distribuzione normale scaturisce, fra gli altri modi, come legge di probabilitàdegli errori formati dall'accumularsi (additivo! di un gran numero di piccoli errori dovuti a cause indipendenti e con leggi di distribuzione arbitrarie.

In effetti la forma impiegata da Subbotin per la sua distribuzione generale di errori era:

$$j\,(u) = \frac{m\,h}{2\Gamma(1/m)}\exp[-h^m\,|u|^m] \qquad\qquad m,\, h > 0,\, u \in \hat{A}$$

Inoltre chiamò anche $1/m$ grado di precisione, dal momento che per $m \rightarrow 0$ si ha $j\,(0) \rightarrow 1$, mentre per $m \rightarrow \infty$ si ha una distribuzione uniforme nell' intervallo $(-h,+h)$.

In questo lavoro vengono richiamate brevemente le caratteristiche del modello delle curve normali di ordine *p*, ricavato da quello di Subbotin, con una rassegna dei principali risultati relativi alla stima dei parametri nella regressione di norma *p*, sulla stima di *p* e su alcune proprietà degli stimatori di massima verosimiglianza dei parametri, insieme con un cenno ad alcuni aspetti computazionali; nell'ultima sezione si fa cenno a problematiche relative alle estensioni multivariate di tali curve.

2. Le distribuzioni normali di ordine *p*.

La famiglia di curve di Subbotin, ripresa poi da vari Autori (Diananda,1949; Box, 1953, Turner, 1960), è citata spesso nella letteratura anglosassone come *exponential power distribution*, anche se con diverse parametrizzazioni: ad esempio quella di Box e Tiao (1973), che l'hanno utilizzata, piuttosto che come distribuzione generale di errori, nello studio dei modelli lineari per l'analisi della robustezza dell'inferenza bayesiana rispetto ad allontanamenti dall'ipotesi di normalità

$$f(z) = \frac{1}{2^{(b+3)/2}\,s\;\Gamma[(3+b)/2]}\exp\left(-\frac{1}{2}\left|\frac{z-m}{s}\right|^{2/[1+b]}\right) \qquad -1 \leq b \leq 1,\; s > 0,\, z \in \hat{A} \qquad (1)$$

in cui b è un parametro di non-normalità Già comunque Box (1953) aveva impiegato una famiglia simile di distribuzioni per studiare la robustezza del test *F* in funzione di allontanamenti dalla normalità

Un'altra parametrizzazione, introdotta da Vianelli (1963) e da Lunetta (1963), è quella che porta al modello delle *curve normali di ordine p (c.n.o.p)*, costituito da una famiglia di distribuzioni di variabile reale *Z*, con densitàdata da:

$$f(z) = \frac{1}{2\,s_p\,p^{1/p}\Gamma(1+1/p)}\exp\left(-\frac{|z-m|^p}{p\,s_p^p}\right) \qquad p > 0,\; s_p > 0\; z \in \hat{A} \qquad (2)$$

ove m è il vero valore di una quantità, i cui valori osservati z sono affetti da errori con dispersione s_p, scarto medio assoluto di ordine *p*, poiché si ha:

$$m = \mathrm{E}[Z] \qquad \text{e} \qquad s_p = \{\mathrm{E}|Z - m|^p\}^{1/p}.$$

L'importanza di questa parametrizzazione risiede nell'avere introdotto esplicitamente s_p, come parametro di variabilitàdi ordine *p*. Le curve risultano unimodali, simmetriche e, per $p > 1$, campanulari: per ogni valore di p si ha una diversa distribuzione di errori. Come casi particolari si hanno la distribuzione di Laplace ($p = 1$), la normale ($p = 2$) e l'uniforme ($p \rightarrow \infty$) nell'intervallo $[-s_p, s_p]$. Quando $p \rightarrow 0^+$ si ottiene una distribuzione degenere che concentra tutta la massa nel punto $z=m$ e risulta nulla quasi ovunque, per cui in questo caso Z assume con probabilità 1 il valore m. Il parametro di forma p è legato a b della parametrizzazione di Box-Tiao nella (1) dalla relazione: $p=2/(1+b)$.

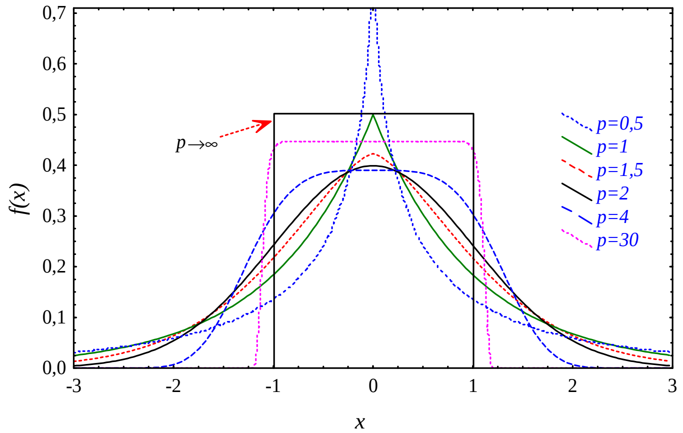

Figura 1 *Densità di curve normali di ordine p standardizzate, per alcuni valori di p.*

Il parametro di forma p è collegato con la kurtosi, in quanto si ha:

$$b_2 = \frac{m_4}{m_2^2} = \frac{\mathrm{E}(|Z - m|^4)}{\mathrm{E}^2(|Z - m|^2)} = \frac{G(1/p)G(5/p)}{[G(3/p)]^2}.$$

ed i valori dell'indice di kurtosi b_2 variano da $b_2 \rightarrow \infty$ per $p \rightarrow 0^+$ a 1,8 per la distribuzione uniforme ($p \rightarrow \infty$). Un indice di kurtosi generalizzato (Lunetta, 1963) é dato da:

$$b_p = \frac{m_{2p}}{m_p^2} = \frac{\mathrm{E}(|Z - m|^{2p})}{\mathrm{E}^2(|Z - m|^p)} = p+1,$$

mentre l'indice di forma I, utile per alcuni procedimenti empirici per la stima di p é:

$$I = \frac{m_1}{\sqrt{m_2}} = \frac{\mathrm{E}(|Z - m|)}{\sqrt{\mathrm{E}(|Z - m|^2)}} = \frac{G(2/p)}{\sqrt{G(1/p)G(3/p)}}.$$

I vari indici di kurtosi possono essere utili per la stima di p attraverso un campione di osservazioni.

2.1 Relazioni delle curve normali di ordine *p* con altre distribuzioni.

Sono numerosi i legami delle distribuzioni normali di ordine p con altre distribuzioni. Fra le trasformazioni elementari si ha che se Z é una variabile casuale distribuita secondo una c.n.o.p di densità data dalla (2), allora $(|Z - m|/s_p)^p$ segue una distribuzione Gamma con parametro di forma $1/p$ e parametro di scala $1/p$ (Lunetta, 1963). Questa proprietà sfruttata nella costruzione dei prontuari (Vianelli, Mineo,1978) é utile per il calcolo delle probabilità integrali della (2). Vianelli (1968) introduce le curve normali di ordine r per intervalli finiti, basate su trasformate della variabile casuale Beta. Bellavista

(1970) mostra che la c.n.o.*p*. è il limite di una particolare distribuzione ottenuta come generalizzazione della *t* di student; ancora Bellavista (1979) dimostra che la distribuzione normale di ordine *p* é la distribuzione limite di una particolare sequenza di variabili casuali di cui sia costante la media delle potenze *p*-esime.

Si può dimostrare che se X~*Gamma*$(1+1/p,\ 1/(p s_p^p))$, U~*Uniforme*[-1;1], allora $Z=XU$ é distribuita secondo una normale di ordine *p* (Johnson 1986).

La normale di ordine *p* può scaturire anche come mistura continua di scala di distribuzioni normali e distribuzioni stabili (West, 1987) per $1 \le p \le 2$; Choy e Smith (1997) utilizzano tale rappresentazione per studiare il comportamento della distribuzione a posteriori di un parametro di posizione di una normale in funzione di alcune scelte di distribuzioni a priori fra cui appunto la normale di ordine *p* per 1£*p*£2.

Altre relazioni, in particolare con la distribuzione beta, sono state sfruttate anche per la generazione efficiente di numeri casuali (Chiodi, 1986, 1995, 2000a).

3. Regressione di norma *p*.

Il modello trova applicazione (Mineo, 1989) in problemi di regressione, lineare e non, supponendo che nella (2) si abbia: $m=g(\mathbf{x}, \mathbf{b})$, ossia il valore atteso sia funzione di alcuni regressori noti $\mathbf{x}$ e di un vettore $\mathbf{b}$ di k parametri incogniti; la variabile osservabile Y é allora funzione dei regressori X_j e di una componente accidentale secondo la relazione:

$$Y_i = g(\mathbf{x}_i,\mathbf{b})+e_i$$

e le e_i sono distribuite indipendentemente secondo delle normali di ordine *p*, con uguali dispersioni s_p e speranze matematiche nulle. Con queste ipotesi le stime $\hat{\mathbf{b}}$ di massima verosimiglianza dei coefficienti di regressione $\mathbf{b}$, *se p é noto*, sono fornite dalle stime di minima norma *p*, che si ottengono minimizzando l'espressione:

$$\mathrm{R}(\mathbf{b},p) = \sum_{i=1}^{n} | y_i - g(\mathbf{x}_i,\mathbf{b})|^p.$$

Infatti, l'inferenza sui parametri $\mathbf{b}$, s_p^p, ed eventualmente anche *p*, sulla base di un campione di *n* osservazioni $(y_i,\mathbf{x}_i)$, $i=1,2,\ldots,n$ è basata sulla log-verosimiglianza:

$$l(\mathbf{b}\ \ , s_p^p, p; \mathbf{y}) = -\frac{\sum_{i=1}^{n} | y_i - g(\mathbf{x}_i,\mathbf{b})|^p}{p s_p^p} - \frac{n \log p}{p} - n \log\left(2G(1+1/p)\right) - n \log s_p$$

e dipende dai dati campionari solo attraverso la quantità $\sum_{i=1}^{n} | y_i - g(\mathbf{x}_i,\mathbf{b})|^p$, mentre la stima di massima verosimiglianza di s_p^p é data da:

$$\hat{s}_p^p = \sum_{i=1}^{n} | y_i - g(x_i, \hat{\mathbf{b}})|^p/n,$$

per cui si può eliminare l'influenza del parametro di disturbo s_p^p attraverso la verosimiglianza profilo per $\mathbf{b}$ e *p*:

$$l(b,\hat{s}_p^p(b,p),p;\mathbf{y}) = -\frac{n}{p}(1+\log p)-n \log \left(2G(1+1/p)\right) - n \log\left(\sum_{i=1}^{n} | y_i - g(\mathbf{x}_i,b)|^p\right)+n \log n$$

Nel caso in cui $g(x_i,b)=m$ (ossia le Y_i costituiscono un campione di n osservazioni i.i.d.), la stima di massima verosimiglianza del parametro di posizione $\hat{m}$ è, per $p\geq 1$, la soluzione dell'equazione:

$$\sum_{i=1}^{n} |y_i - \hat{m}|^{p-1} sign(y_i - \hat{m})=0,$$

che per $p>1$ è unica. Gli unici casi in cui esiste una soluzione esplicita per $\hat{m}$ si hanno per $p=2$, caso in cui $\hat{m}$ é la media aritmetica M, per $p\rightarrow\infty$, quando $\hat{m}$ è la semisomma dei valori estremi, e per $p=4$, caso in cui $\hat{m}$ è l'unica soluzione reale dell'equazione di terzo grado: $(\hat{m}-M)^3+3(\hat{m}-M)\hat{m}_2-\hat{m}_3=0$ essendo $\hat{m}_2$ e $\hat{m}_3$ i momenti centrali campionari secondo e terzo. Se $p=1$ ed n dispari $\hat{m}$ é la mediana, mentre con n pari $\hat{m}$ é un qualsiasi valore compreso fra i due valori centrali del campione. Quando $p<1$ il valore $\hat{m}$ che massimizza la verosimiglianza é uno degli n valori campionari y_i; per $p\rightarrow 0^+$, $\hat{m}$ tende alla moda della distribuzione.

Nel caso più generale della regressione lineare di norma p, con $g(\mathbf{x}_i,b)= \mathbf{x}_i,b$, la possibilità di far variare il parametro p con continuità consente una generalizzazione dei metodi alternativi ai minimi quadrati ordinari ($p=2$), ossia la regressione L_1 (minimizzazione degli scarti assoluti) o L_∞ (minimizzazione dello scarto massimo), ed anzi li inquadra in un corretto contesto parametrico; per $p<1$ la verosimiglianza non é unimodale (Zeckhauser, Thompson, 1970), ed il valore ottimo del vettore dei parametri si ottiene in corrispondenza di configurazioni di valori delle componenti di $\hat{b}$ che annullano almeno uno dei residui.

3.1 Distribuzioni asintotiche

Le distribuzioni campionarie non sono in generale note per valori di n finiti, tuttavia per la regressione lineare di norma p con k regressori, le distribuzioni degli stimatori di massima verosimiglianza $\hat{b}$, $\hat{s}_p^p$, $\hat{p}$ per $p\geq 1$ sono asintoticamente normali, con matrice di varianze e covarianze asintotiche data da:

$$V(\hat{b}, \hat{s}_p^p, \hat{p}) = \begin{pmatrix} V(\hat{b}) & \mathbf{0}_{k,2} \\ \mathbf{0}_{2,k} & V(\hat{b}, \hat{s}_p^p, \hat{p}) \end{pmatrix},$$

con $\mathbf{0}_{a,b}$ una matrice di zeri con a righe e b colonne,

$V(\hat{b})=\dfrac{p^{2/p-1}\, s_p^2\, G(1+1/p)}{G(2-1/p)}\,(\mathbf{X}'\mathbf{X})^{-1}$, essendo $\mathbf{X}$ la matrice degli scarti dei k regressori e:

$$V(\hat{s}_p^p, \hat{p})=p\begin{pmatrix} (A_p^2+B_p)s_p^{2p} & p^2A_p s_p^p \\ p^2A_p s_p^p & p^3 \end{pmatrix}/ n[B_p-A_p^2(p-1)],$$

avendo posto per comodità $A_p = \log\left(p s_p^p\right) + y(1+1/p)$; $B_p = (1+1/p)\, y'(1+1/p) -1$; si veda anche Canal (1989) e Agrò (1995) per il caso univariato.

Asintoticamente quindi gli stimatori dei coefficienti di regressione sono indipendenti da quelli del parametro σ_p^p e di quello di struttura p, mentre risultano sempre correlati gli stimatori del parametro di forma e di quello di σ_p^p.

3.2 Proprietà degli stimatori e approssimazioni analitiche.

Per valori di p qualsiasi non esistono in generale stimatori sufficienti, tranne che per μ noto, caso in cui $\hat{\sigma}_p^p$ é sufficiente per σ_p^p (Cifarelli, Regazzini, 1974); qualche caso particolare si ha per valori di p pari (Box-Tiao,1973).

La superiorità, anche per piccoli campioni, degli stimatori di norma p rispetto ad un'ampia classe di stimatori robusti, è stata mostrata mediante tecniche Montecarlo da Mineo (1983,1986).

Per la determinazione analitica delle distribuzioni campionarie di stimatori e test nel caso in cui il parametro di struttura p sia noto, a parte alcuni (rari!) casi più semplici si può ricorrere ad approssimazioni asintotiche di varia natura (Barndorff-Nielsen e Cox, 1989; McCullagh, 1987), fra cui particolarmente utili quelle basate sulla determinazione dei punti di sella (Chiodi, 1994 e 2000b). Infatti per approssimare la distribuzione dello stimatore di massima verosimiglianza del parametro di posizione, o dei parametri di una relazione di regressione, si può ricorrere utilmente alle tecniche generali per stimatori di tipo M (Field, 1982; Lugannani, Rice, 1980; Ronchetti,1990).

Fra i risultati di base si ha che la quantità campionaria:

$$\hat{\sigma}_p^p(\mu) = \sum_{i=1}^{n} |Y_i - \mu|^p/n$$

si distribuisce secondo una Gamma con parametro di forma $c=n/p$ e parametro di scala $\lambda=n/p$ (Lunetta, 1966). Nel caso in cui μ non sia noto e si impieghi $\hat{\mu}$, la quantità

$\hat{\sigma}_p^p(\hat{\mu}) = \sum_{i=1}^{n} |Y_i-\hat{\mu}|^p/n$ per $p>1$ ha una distribuzione ben approssimabile da quella di una distribuzione Gamma con parametro di forma $c=n/p-1/2$ e parametro di scala $\lambda=n/p$ (Chiodi, 1988); il risultato si ottiene partendo dall'approssimazione: $E[n\hat{\sigma}_p^p(\hat{\mu})]\cong n-p/2$, e poi impiegando una distribuzione gamma con parametro di scala n/p, e parametro di forma corretto in modo di avere la stessa speranza matematica; in effetti questa stessa distribuzione approssimante si ottiene (Chiodi, 1994) ricavando l'approssimazione di punto di sella alla distribuzione congiunta di $\hat{\mu}$ e $\hat{\sigma}_p^p(\hat{\mu})$ ed integrandola rispetto a $\hat{\mu}$ mediante approssimazione di Laplace nel punto $\hat{\mu}=\mu$: si ottiene come approssimazione alla distribuzione marginale di $\hat{\sigma}_p^p(\hat{\mu})$ una distribuzione Gamma(n/p-½,n/p). Questa approssimazione ed altre simili relative al caso della regressione lineare possono essere impiegate per migliorare ad esempio le approssimazioni alle distribuzioni campionarie dei test basati su rapporti di verosimiglianza, almeno nel caso in cui p sia noto (Chiodi, 2000b).

3.3 Approssimazioni mediante tecniche Montecarlo: generazione di numeri casuali da una distribuzione normale di ordine *p*.

L'ampia disponibilità di software e la velocità di calcolo consentono di approssimare le distribuzioni campionarie di stimatori e test con tecniche di simulazione, almeno per ottenere delle soluzioni orientative. In alcuni miei lavori, (Chiodi, 1986,1995), ho mostrato che si possono ottenere numeri casuali X dalla distribuzione normale standardizzata di ordine p, per valori di $p > 1$, o generalizzando la formula di Box-Muller (1958), oppure (Chiodi, 1995) con una generalizzazione della trasformazione polare di Marsaglia (Marsaglia, Bray, 1964), che porta ad una codifica molto compatta, o con tecniche di compressione che portano ad algoritmi molto veloci, ma con codifica molto lunga; un algoritmo basato su una generalizzazione della tecnica del rapporto di numeri uniformi è fornito da Barabesi (1993). Eventualmente si può ricorrere alla proprietà che le potenze p-esime di variabili normali di ordine p si distribuiscono secondo una Gamma, per la quale esistono molte librerie efficienti di generazione di numeri casuali.

La disponibilità di software per la generazione di numeri pseudo-casuali da una distribuzione normale di ordine p rende più agevole lo studio delle distribuzioni campionarie di stimatori e test, mediante l'impostazione di studi di simulazione con metodo Montecarlo, in particolare per piccoli campioni, specie nei casi in cui p vada stimato dai dati campionari, o comunque per i casi in cui non siano noti risultati analitici anche approssimati.

3.4 Aspetti computazionali.

Gli algoritmi ed il corrispondente software numerico per trattare i problemi di regressione lineare di minima norma p, almeno con p noto e per insiemi di dati ben strutturati e non troppo numerosi, sono facilmente disponibili, o comunque implementabili a partire da buone routines di ottimizzazione numerica, presenti in molti packages. Per i problemi di regressione anche non lineare di norma p (Gonin,Money, 1989) esistono degli algoritmi fondati ad esempio su tecniche di minimi quadrati ponderati reiterati, riformulando opportunamente il problema di minimo. Tuttavia per alcuni problemi numerici sono sempre possibili nuovi algoritmi, specie per grandi quantitativi di dati: Poertnoy, Koenker (1997) propongono un miglioramento delle tecniche numeriche classiche basate sul metodo del simplesso per la risoluzione dei problemi di regressione di norma L_1, che porta a dei tempi di calcolo vicini a quelli dei minimi quadrati ordinari per problemi con migliaia di dati e poche variabili. Inoltre le soluzioni ai problemi L_1, L_2 e L_∞ possono essere usate per ricavare dei punti iniziali per il problema generale di norma L_p.

Fra gli altri aspetti di natura computazionale, va detto che in effetti buona parte dell'algebra relativa al modello delle curve normali di ordine p, e della regressione di norma p, può ricavarsi agevolmente mediante software di calcolo simbolico.

3.5 L' exponential power distribution nell'inferenza bayesiana

Nell'inferenza bayesiana il modello delle curve normali di ordine p, di solito nella parametrizzazione di Box e Tiao (*exponential power distribution*), é stato impiegato per l'analisi della robustezza dell'inferenza rispetto ad allontanamenti dalla normale; evidentemente l'inferenza é legata alle scelte della distribuzione a priori dei parametri. Recentemente é stato impiegato ad esempio per specificare particolari distribuzioni a priori non normali(Choy,Smith, 1997) nell'inferenza su parametri di posizione in modelli normali; Achcar J.De Araujo Pereira (1999) hanno impiegato una mistura di distribuzioni normali di ordine p in presenza di covariabili; Walker (1999) definisce una Uniform Power Distribution come approssimazione della Exponential power Distribution, per fare in modo che risulti comoda l'implementazione di metodi MCMC, impostando delle opportune distribuzioni a priori per i parametri.

In generale comunque in un ottica bayesiana l'impiego di una famiglia di distribuzioni di errori più generali di quella normale, parametrizzate da p, può essere utile per la validazione dei modelli e la sua influenza può essere eliminata integrando la distribuzione a posteriori dei parametri rispetto al parametro p.

4. Stima del parametro di struttura *p*.

Un problema fondamentale é in ogni caso, nell'analisi di dati reali, quello della stima del parametro p. Tale parametro in effetti svolge nella definizione del modello il ruolo di parametro di struttura, che caratterizza un particolare processo di generazione degli errori di osservazione, mentre quando occorre stimare i parametri b (o m) e s_p^p, p è solo un parametro di disturbo, poiché potrebbe non essere interessante in sé la stima di p, ma semplicemente la sua influenza sugli stimatori degli altri parametri del modello.

Una buona soluzione, valida anche per campioni piccoli (n=10), nel caso in cui $g(\boldsymbol{x}_i,\boldsymbol{b})=m$ o $g(\boldsymbol{x}_i,\boldsymbol{b})= b_0 + b_1\, x_i$ é fornita da A.M. Mineo (1994,1999). La tecnica di stima è basata su un algoritmo a due stadi in cui iterativamente vengono valutati $\hat{p}$ e le stime degli altri parametri: p è stimato risolvendo numericamente un'equazione nella quale si eguaglia il valore teorico dell'indice I con quello empirico; gli altri parametri vengono stimati dai corrispondenti stimatori di massima verosimiglianza in funzione del valore di $\hat{p}$ corrente; il procedimento viene iterato fino alla convergenza. La tecnica prevede accorgimenti particolari per valori molto elevati di $\hat{p}$ o per valori di $\hat{p}$ inferiori ad uno. La bontà dell'approccio è stata mostrata dall'Autore mediante piani di simulazione di tipo Montecarlo.

Un altro approccio, basato sulla verosimiglianza, è stato proposto da Agrò (1995) per il caso $g(\boldsymbol{x}_i,\boldsymbol{b})=m$. Si tratta sempre di un procedimento a due stadi che tuttavia è adatto essenzialmente per campioni medio-grandi (da n>50); successivamente (Agrò,1999) ha proposto una modifica della verosimiglianza basata sull'impostazione di Cox e Reid (1987), che prevede una riparametrizzazione in modo da avere stimatori asintoticamente non correlati, ottenendo risultati migliori (almeno per n>30).

I metodi basati interamente sulla verosimiglianza in effetti comportano comunque dei problemi dovuti al fatto che, almeno per campioni piccoli, la verosimiglianza può risultare non unimodale, nel caso in cui occorra stimare p insieme agli altri parametri.

Tuttavia nelle applicazioni può essere utile, per fare inferenza sui parametri, valutare la verosimiglianza profilo (3) per diversi valori di p, per vedere la sensibilità di $\hat{m}$ al variare di p. Eventualmente i contorni della verosimiglianza profilo rispetto ai due parametri p e m possono essere rappresentati in un grafico bidimensionale per valutare simultaneamente la plausibilità di particolari valori di p e m o per costruire regioni di confidenza approssimate.

5. Estensioni multivariate.

L'estensione al caso multivariato di distribuzioni non-normali a componenti non indipendenti è sempre ardua, perché le possibilità di estensione di sistemi di curve univariate non normali al caso multivariato possono essere di diversa natura e condurre a risultati differenti, mentre comunque si estenda la distribuzione normale univariata al caso multivariato, si arriva sempre alla stessa forma multivariata (Fang et al., 1989); ad esempio:

- dalla densità o dalla funzione caratteristica, sostituendo ad un quadrato una forma quadratica;
- se **X'a** é normale per qualsiasi **a**, allora **X** é normale multivariato
- come distribuzione di $\mathbf{X}=m+\mathbf{AY}$ (con **Y** a componenti indipendenti)
- da distribuzioni condizionate normali e omoscedastiche con funzioni di regressione lineari.

Nell'ambito degli allontanamenti non simmetrici dalla multinormalità, una famiglia di distribuzioni flessibile è costituita dal sistema di curve normali multivariate asimmetriche di Azzalini e Dalla Valle (1996), (Azzalini, Capitanio 1999).

Per quanto riguarda le estensioni multivariate della famiglia di distribuzioni normali di ordine p, queste non possono essere ottenute in unica forma, come avviene invece per la normale multivariata: una particolare estensione, per il caso bivariato è quella di Taguchi (1978), con distribuzioni marginali normali di ordine p, e caratterizzata da un particolare indice di correlazione r_p di norma p: tuttavia, con distribuzioni parziali in generale asimmetriche, ed in particolare con contorni non convessi, risulta a mio avviso di scarsa rilevanza pratica.

Lunetta (1967) ha ricavato un'estensione della c.n.o.p a due variabili, introducendo nel modello un parametro che misura l'interdipendenza fra le due componenti casuali, e che costituisce quindi un particolare indice generalizzato di correlazione; la densità di tale distribuzione é data da:

$$f(x,y)=k_p \exp[-h(p, r)\,(|x|^p-2r|xy|^{p/2} Sign(xy)+|y|^p)]$$

Le distribuzioni parziali risultano asimmetriche, tranne che nei casi p=2 oppure r=0; in quest'ultimo caso si ottiene una distribuzione di due variabili normali di ordine p indipendenti.

Un'estensione al caso bivariato della normale di ordine p é dovuta a De Simoni (1968) con densità

$$f(x,y)= D \exp\{-(A\,x^2 - 2\,B\,xy + C\,y^2)^{p/2}/p\}$$

in cui A, B, C e D sono delle costanti.

Altre generalizzazioni si trovano con opportune estensioni ad esempio del sistema multivariato di Johnson, oppure come caso particolare del sistema di curve simmetriche di Kotz, (Fang et al., 1989, Krzanowski,Marriott, 1994) di densità

$$f(\mathbf{y})= C\,|\mathrm{S}|^{-1/2}\,[(\mathbf{y}-m)'\mathrm{S}^{-1}(\mathbf{y}-m)]^{N-1}\exp\{-r[(\mathbf{y}-m)'\mathrm{S}^{-1}(\mathbf{y}-m)]^{s}\}, \qquad (4)$$

che, se N=1 e quando $s\neq1$, rappresenta degli allontanamenti simmetrici a contorni ellissoidali dalla normale multivariata (ponendo $s=p/2$).

Altre possibili estensioni scaturiscono dal sistema di distribuzioni multivariate, ottenuto come estensione delle rappresentazioni di distribuzioni univariate di Knintchine, in cui le singole componenti X_i siano esprimibili come $X_i=Z_i\,U_i$ che possono portare (Johnson, 1986, cap. 8) a diversi tipi di distribuzioni con differenti caratterizzazioni della interdipendenza fra le componenti X_i, se si considerano componenti di Z_i ed U_i non indipendenti e in modo tale che le X_i siano normali di ordine p. In qualche modo collegabili a questo approccio sono i risultati di Gupta e Song (1997) sulla caratterizzazione delle cosiddette distribuzioni sferiche di norma L_p.

Fra i possibili impieghi delle estensioni multivariate delle distribuzioni normali di ordine *p* si hanno sinteticamente:

- descrizione di insiemi di dati multivariati;
- descrizione di errori accidentali di osservazione per misurazioni multiple ripetute;
- in modelli di regressione (lineare e non lineare), per la descrizione di errori non normali e in qualche modo correlati;
- per specificare particolari distribuzioni alternative alla normale multivariata in test di multinormalità;
- per incudere la normale multivariata in una classe più ampia di distribuzioni allo scopo di saggiare la robustezza di test univariati o multivariati, costruiti sull'ipotesi di multinormalità, rispetto ad allontanamenti dall'assunzione di multinormalità
- come distribuzione a priori in particolari problemi di inferenza bayesiana.

In letteratura invece alcune delle estensioni multivariate delle curve normali di ordine *p*, sono prevalentemente impiegate per specificare distribuzioni alternative in test di multinormalità (Naito, 1998), o per valutare la potenza di un test per specifiche alternative multivariate non normali; ad esempio Kuwana e Kariya (1991) impiegano una particolare forma della (4) (distribuzione ellittica di potenza in cui la non-normalità è riassunta in un solo parametro) come distribuzione alternativa in test di multinormalità

Meno utilizzate (a nostra conoscenza) sono invece tali estensioni multivariate per la descrizione di insiemi di dati multivariati, o per la descrizione di errori accidentali di osservazioni di misurazioni multiple ripetute o come distribuzione di errori generali in modelli di regressione lineare e non lineare.

Riferimenti bibliografici

Achcar J. A. De Araujo Pereira (1999) Use of exponential power distributions for mixture models in the presence of covariates. *Journal of Applied Statistics*, 26, 6, 669-679.

Agrò, G. (1995) Maximum likelihood estimation for the exponential power distribution parameters. *Comm.Stat. Simulation and Computation* 24, 523-536

Agrò, G. (1999) Ortogonalità dei parametri e verosimiglianza profilo condizionata: il caso della distribuzione normale di ordine *p*. *Annali della Facoltà di Economia di Palermo, Area Statistico-Matematica*, 9-18.

Azzalini A., Capitanio A. (1999) Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society* B, 61,3, 579-602.

Azzalini A., Dalla Valle A. (1996) The multivariate skew normal distribution. *Biometrika*, 83, 715-726.

Barabesi L., (1993), Optimized ratio-of-uniform method for generating exponential power variates, *Statistica Applicata, 5, 2, 149-155.*

Bellavista Vianelli L. (1970): Di una generalizzazione della distribuzione di Student. *Annali della Facoltà di Economia e Commercio di Palermo*, 24,2, 211-223.

Bellavista Vianelli L. (1979): The normal distribution off order *r* and some new functions arising from these laws. *Quaderno n.2 Annali della Facoltà di Economia e Commercio di Palermo-Contributi degli istituti di statistica e matematica per la ricerca operativa*, 2.

Barndorff-Nielsen O.E., Cox D.R.. (1989), *Asymptotic Techniques for Use in Statistics*. Chapman and Hall, London.

Box G.E.P., Muller M. E. (1958), A note on the generation of random normal deviates, *Annals of Mathematical Statistics*, 29, 610-611.

Box G.E.P., Tiao G.C. (1973), *Bayesian inference in statistical analysis*. Addison-

Wesley Ed.; Reading, Massachusetts.
Box G. E. P. (1953) A note on regions for test of kurtosis. *Biometrika*, 40, 465-468.
Canal L., (1989) Stimatori di Massima verosimiglianza dei parametric della distribuzione lognormale generalizzata e loro distribuzioni di probabilità *Quaderni di Statistica e Matematica Applicata alle Scienze Economico-Sociali. Facoltà di Ec. e Comm. Università di Trento*, 11,3-4 pag.27-36.
Chiodi M. (1988),Sulle distribuzioni di campionamento delle stime di massima verosimiglianza dei parametri delle curve normali di ordine *p*. *Pubbl. a cura dell'Istituto di Statistica di Palermo* il 21-12-1988 ai sensi dell' Art.1 D.L. 31-8-1945,n.660.
Chiodi M. (1986), Procedures for generating pseudo-random numbers from a normal distribution of order *p* (*p*>1), *Statistica Applicata*, 1, 7-26.
Chiodi M. (1994) Approssimazioni saddlepoint alle distribuzioni campionarie degli stimatori di massima verosimiglianza dei parametri delle curve normali di ordine *p* per piccoli campioni. *Atti della XXXVII Riunione Scientifica della SIS*. Ed. C.I.S.U. Roma; 1, 139-146.
Chiodi M. (1995), *Generation of pseudo random variates from a normal distribution of order p.* Statistica Applicata, 7,4, 401-416
Chiodi M. (2000 a) *Tecniche di Simulazione in Statistica.* Volume in corso di pubblicazione nella Collana del Dipartimento di Matematica e Statistica dell'Universitàdegli Studi di Napoli Federico II-Serie didattica.
Chiodi M.(2000 b), *Some results on likelihood-based inference in norm-p regression.* lavoro in corso di stesura.
Choy S. T. B., Smith A. F. M. (1997) On Robust Analysis of a Normal Location Parameter. *Journal of the Royal Statistical Society B*, 59, 2, 463-474
Cifarelli D.M., Regazzini E. (1974) Sulla caratterizzazione di una famiglia di distribuzioni basata sulla efficienza di una funzione campionaria. *Rend. Sem. Mat. Univers. Politecn. Torino*, 33, 299-311.
Cox D.R. Reid N. (1987) Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, 49, 1-39
De Simoni S. (1968), Su una estensione dello schema delle curve normali di ordine «r». alle variabili doppie. *Statistica*, 37. 447-474.
Diananda P.H. (1949), "Note on some proprieties of maximum likelihood estimates". *Proceedings of the Cambridge Philosophical Society*, 45, 536-544.
Fang K.T, Kotz S., Ng K.W. (1989) *Simmetric Multivariate and Related Distributions*. Chapman & Hall, Londra/New York.
Field C.A. (1982), Small sample asymptotic expansion for multivariate M-estimates. *Ann. Statist.*, 10, 672-689.
Gupta A.K. Song D. (1997) Characterization of *p*-generalized normality. *Journal of multivariate analysis*, 60, 61-71.
Johnson, *M. E.*, (1987), *Multivariate statistical simulation*, Wiley ed., New York.
Krzanowski W.J., Marriott F.H.C. (1994), *Multivariate Analsysis (Part I).* Edward Arnold, London
Kuwana Y. Kariya T. (1991) LBI tests for multivariate normality in exponential power distributions, *Journal of multivariate analysis*, 39, 117-134
Lugannani R., Rice S. (1980), "Saddlepoint approximation for the distribution of the sum of independent random variables". *Adv. Appl. Prob.*,12, 475-490.
Lunetta G., (1963), Di una generalizzazione dello schema della curva normale. *Annali della Facoltà di Economia e Commercio di Palermo*, 17, 2, 235-244.
Lunetta G. (1966) Di alcune distribuzioni deducibili da una generalizzazione dello schema della curva normale. *Annali della Facoltà di Economia e Commercio di Palermo*, 20, 1, 117-143.
Lunetta G. (1967): Su una estensione del coefficiente di correlazione lineare. *Annali della Facoltà di Economia e Commercio di Palermo*, 21,2, 187-201
Marsaglia, G., Bray, T. A., 1964, A convenient method for generating normal variables.

SIAM rev. 6, 260-264.
McCullagh, (1987) *Tensor Methods in Statistics*. Chapman and Hall, London.
Mineo A. (1978), La stima dei parametri delle curve normali di ordine r per intervalli finiti e di quelle per intervalli infiniti, *Studi in onore di Paolo Fortunati,* C.L.E.U.B. *ed., Bologna*, 1, 565-577.
Mineo A., (1983), La miglior combinazione delle osservazioni per la stima dei parametri di intensità e di dispersione, *Atti del Convegno della S.I.S. di Trieste*, 1, 463-487.
Mineo A. (1986), La migliore combinazione delle osservazioni per la stima dei parametri di intensità e di dispersione. *Studi in onore di Silvio Vianelli: Istituto di Statistica, Palermo;*1, 517-548.
Mineo A. (1989) The norm-*p* estimation of location, scale and simple linear regression parameters, *Lecture notes in Statistics, Statistical Modeling Proceedings*, Trento, 222-223.
Mineo A. M. (1994) Un nuovo metodo di stima di *p* per una corretta valutazione dei parametri di intensità e di scala di una curva normale di ordine *p*. *Atti della XXXVII Riunione Scientifica della SIS*. Ed. C.I.S.U. Roma; 1, 147-154.
Mineo A. M. (1999) stima dei parametri di regressione lineare semplice quando gli errori seguono una distribuzione normale di ordine *p* (*p* incognito). *Annali della Facoltà di Economia di Palermo, Area Statistico-Matematica (1988-95)* pag.
Naito K. (1998), Approximation of the power of kurtosis test for multinormality, *Journal of multivariate analysis*, 65, 166-180.
Pizzetti P. (1892) *I fondamenti matematici per la critica dei risultati sperimentali.* Cappelli editore, Ristampa del 1963.
Poertnoy S. Koenker R. (1997) The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute-Error Estimators. *Statistical Science*, 12, 4, 279-300
Ronchetti E. (1990), Small sample asymptotics and bootstrap. *Quaderni di Statistica e Econometria*, Centro di Specializzazione e Ricerche Economiche Agrarie per il Mezzogiorno, Portici (NA), 12, 19-38
Subbotin M. T.(1923) On the law of frequency of errors, *Matematicheskii Sbornik*, 31, 296-301.
Taguchi, T. (1978) On a generalization of gaussian distribution. *Ann.Inst.Statist. Math*..30,A, 211-242
Turner M. E (1960), On Heuristic Estimation Methods. *Biometrics*, 16,2, 299-302
Vianelli S. (1963), La misura della variabilitàcondizionata in uno schema generale delle curve normali di frequenza. *Statistica, 33. 447-474.*
Vianelli S. (1968): Sulle curve normali di ordine r per intervalli finiti delle variabili statistiche. *Annali della Facoltà di Economia e Commercio di Palermo* 22,2, 241-257.
Vianelli S., Mineo A. (1978), *Prontuari delle probabilità integrali delle curve normali di ordine p*. Edigraphica Sud Europa, Palermo.
Walker S. (1999) The uniform power distribution, *Journal of Applied Statistics,* 26, 4, 509-517.
West M. (1987) On scale mixtures of normal distributions. *Biometrika* 74,3, 646-648
Zeckhauser R. Thompson M. (1970) Linear regression with non-normal error terms. *The Review of Economics and Statistics*, 52,3, 280-286.