COMPLEMENTI DI STATISTICA

Parte II

Fulvio Ricci

Dipartimento di Fisica, Università di Roma *La Sapienza*

INDICE

LA DISEGUAGLIANZA DI CHEBYCHEV

Sia X una variabile aleatoria avente valore medio μ e varianza σ^2. Sia poi K un arbitrario numero positivo. È possibile dimostrare che la probabilità di avere $|X-\mu| \geq K\sigma$ è sempre minore $1/K^2$.

$$P[|X-\mu| \geq K\sigma] \leq \frac{1}{K^2}$$

Questa è relazione nota come la diseguaglianza di Chebychev.
Dimostriamo allora tale diseguaglianza limitandoci al caso delle distribuzioni continue. Se $K < 1$ ciò è ovviamente verificato, essendo la probabilità $P \leq 1$. Dobbiamo quindi rivolgere la nostra attenzione ai valori che può assumere la variabile aleatoria X per cui $|X - \mu| \geq K\sigma$ cioè per tutti quei valori di x fuori dall'intervallo $\mu \pm K\sigma$ tali da verificare la condizione

$$(x-\mu)^2 \geq K^2\sigma^2$$

Calcoliamo la varianza σ^2

$$\sigma^2 = \int_{-\infty}^{\mu-K\sigma}(x-\mu)^2 f(x)dx + \int_{\mu-K\sigma}^{\mu+K\sigma}(x-\mu)^2 f(x)dx + \int_{\mu+K\sigma}^{+\infty}(x-\mu)^2 f(x)dx$$

Limitiamoci a considerare i valori di x tali che $(x-\mu)^2 \geq K^2\sigma^2$,

$$\sigma^2 \geq K^2\sigma^2 \left[\int_{-\infty}^{\mu-K\sigma} f(x)dx + \int_{\mu+K\sigma}^{+\infty} f(x)dx\right] + \int_{\mu-K\sigma}^{\mu+K\sigma}(x-\mu)^2 f(x)dx$$

Poichè

$$\int_{\mu-K\sigma}^{\mu+K\sigma}(x-\mu)^2 f(x)dx \geq 0$$

segue che

$$\sigma^2 \leq K^2\sigma^2 \underbrace{\left[\int_{-\infty}^{\mu-K\sigma} f(x)dx + \int_{\mu+K\sigma}^{+\infty} f(x)dx\right]}_{P[|X-\mu|\geq K\sigma}$$

$$P[|X-\mu| \geq K\sigma] \leq \frac{1}{K^2}$$

come volevasi dimostare.
La diseguaglianza di Chebychev fornisce un limite superiore alla probabilità che X scarti da μ K volte σ. In genere tale limite è molto più grande del valore effettivo di P che dipende dalla distribuzione. Si noti che tale limitazione al valore della probabilità è <u>indipendente dal tipo di distribuzione</u>. Ad esempio se poniamo $K = 2$

$$P[|X-\mu| \geq 2\ \sigma] \leq 0,25$$

(che implica $P[|X-\mu| < 2\ \sigma] \geq 0,75$). Considerando il caso della distribuzione normale, ci rendiamo conto immediatamente che, a parità di deviazione standard σ, la probabilità che X sia fuori dall'intervallo $\mu \pm K\sigma$ è minore tanto più grande è K. X deve assumere valori sempre più lontani da μ e che quindi cadono sempre più sulle code della curva di Gauss.
A parità di K, al diminuire della deviazione standard σ, la curva a campana di Gauss si restringe progressivamente, così che per avere un eguale valore di probabilità

delle code si devono considerare intervalli di ampiezza diversa. In altre parole se considero due distribuzioni gaussiane aventi $\sigma_2 < \sigma_1$, allora X risulta molto più probabilmente concentrata attorno al valore medio μ nel caso della distribuzione con σ_2.

<u>LA LEGGE DEI GRANDI NUMERI.</u>

Questa legge è connessa alla osservazione empirica (assunta dagli oggettivisti come definizione stessa di probabilità) che, detto N il numero di prove, la frequenza relativa di un evento tende alla probabilità. Questa osservazione, che vogliamo esprimere ora in termini matematici è anche nota come *legge empirica del Caso*.

Sia allora S_N una variabile aleatoria (ad esempio il numero di successi in N prove indipendenti); S_N esprime il possibile risultato di N tentativi a ciascuno dei quali associamo una variabile aleatoria X_i. Avremo allora

$$S_N = \sum_{i=1}^{N} X_i$$

X_i sono variabili aleatorie indipendenti aventi la stessa distribuzione di probabilità. Definiamo la frequenza relativa di successo in queste N prove nel modo seguente

$$\frac{S_N}{N} = frequenza\ relativa\ di\ successo$$

ed il suo valore aspettato è

$$E\left[\frac{S_N}{N}\right] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] = \mu$$

$$Var\left[\frac{S_N}{N}\right] = \frac{1}{N^2}\sum_{i=1}^{N} Var[X_i] = \frac{\sigma^2}{N}$$

Sia ora ε una quantità positva arbitrariamente piccola; invocando la diseguaglianza di Chebychev, possiamo affermare che la probabilità che il modulo della differenza tra $\frac{S_N}{N}$ ed il suo valor medio μ sia maggiore od uguale ad ε, è

$$P\left[\left|\frac{S_N}{N}-\mu\right| \geq \varepsilon\right] \leq \frac{\sigma^2}{N\varepsilon^2}$$

avendo posto

$$\left(\varepsilon = K\frac{\sigma}{\sqrt{N}}\right)$$

In altri termini questo significa che la probabilità che $\frac{S_N}{N}$ si discosta da μ di ε, è tanto più piccola quante più è alto il numero di prove N,

$$\lim_{N\to\infty} P = 0$$

Nello stesso limite la frequenza relativa di successo tende al suo valore aspettato

$$\lim_{N\to\infty} \frac{S_N}{N} = \mu$$

DA UNA DISTRIBUZIONE DISCRETA AD UNA CONTINUA: IL CASO DELLA DISTRIBUZIONE BINOMIALE ED IL TEOREMA LIMITE DI DE MOIVRE - LAPLACE

Ricordiamo che la distribuzione binomiale esprime la probabilità P di ottenere γ successi in n prove quando p è la probabilità del successo in una singola prova

$$
\begin{array}{lll}
n & \Longrightarrow & \textit{numero di prove} \\
\gamma & \Longrightarrow & \textit{numero di successi}
\end{array}
$$

$$
p = 1 - q \begin{cases} p & \Longrightarrow \quad probabilita' \ di \ successo \ in \ una \ prova \\ q & \Longrightarrow \quad probabilita' \ di \ insuccesso \ in \ una \ prova \end{cases}
$$

$$P(\gamma, n, p) = \binom{n}{\gamma} p^{\gamma} q^{n-\gamma}$$

$$E[\gamma] = \bar{\gamma} = np$$

$$Var[\gamma] = \sigma_{\gamma}^{2} = np(1-p)$$

Osserviamo che la distribuzione binomiale nel caso $p = \frac{1}{2}$ è simmetrica attorno al valore medio $n/2 = \bar{\gamma}$.

$$P\left(\gamma, n, \frac{1}{2}\right) = P\left(n - \gamma, n, \frac{1}{2}\right)$$

Per $P \neq \frac{1}{2}$ questo non è vero. La distribuzione di Gauss che è definita per variabili aleatorie continue, è comunque simmetrica attorno al valore medio. Nonostante ciò vi è una importante connessione tra le due distribuzioni.

Infatti possiamo facilmente renderci conto che quando n è grande, per un fissato valore di p l'andamento della $P(\gamma, n, p)$ è approssimato dalla distribuzione di Gauss.

$$P(\gamma, n, p) \sim \frac{1}{\sigma\sqrt{2\ \pi}} e^{-(x-\mu)^2/2\ \sigma^2}$$

con

$$\mu = np \qquad\qquad \sigma = \sqrt{np(1-p)}$$

Ad esempio calcoliamo la probabilità che si ottenga 23 volte "Testa" nel lancio di moneta, avendo ripetuto il lancio 36 volte:

$$P(23, 36, \frac{1}{2} = \frac{36!}{23!13!}\left(\frac{1}{2}\right)^3 6 = 0.0336$$

Se facciamo il calcolo ipotizzando che il processo segua la distribuzione di Gauss con

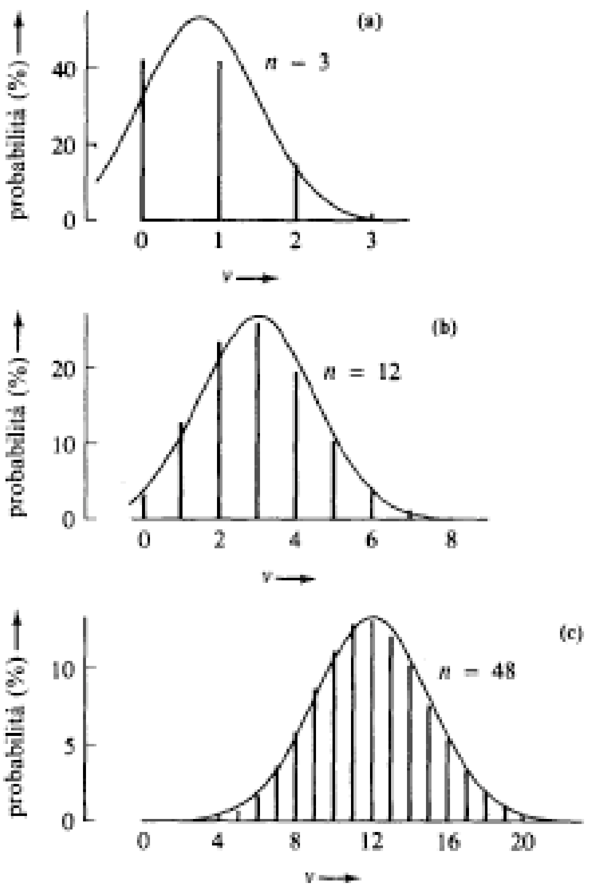
$$\sigma = \sqrt{\frac{36}{2} \cdot \frac{1}{2}} = 3 \qquad\qquad \bar{\gamma} = \mu = 36 \cdot \frac{1}{2} = 18$$

otterremo un valore molto prossimo al precedente

$$P(23) \simeq \frac{1}{3\sqrt{2\ \pi}} e^{-\frac{(23-18)^2}{2 \cdot 9}} = 0.0332$$

Questa coincidenza di numeri non è casuale. Infatti tali considerazioni qualitative possono essere espresse in forma matematicamente rigorosa dal teorema limite di De Moivre - Laplace, di cui qui ci limiteremo solo ad enunciarne il testo.

$$X = np \quad e \quad \sigma = \sqrt{np(1-p)}$$



Il teorema afferma che data una variabile aleatoria discreta S_n avente distribuzione binomiale, il limite per il numero di prove n tendente all'infinito della probabilità che la variabile aleatoria ridotta S_n^* sia compreso nell'intervallo dipendente da due particolari valori x_1 e x_2,

$$x_1 \leq S_n^* \leq x_2$$

è espresso in termini della funzione di ripartizione normale ridotta calcolata in x_1 e x_2

$$\lim_{n\to\infty} P[x_1 \leq S_n^* \leq x_2] = \frac{1}{\sqrt{2\ \pi}} \int_{x_1}^{x_2} e^{-\frac{x^{*2}}{2}} dx^*$$

Se utilizziamo variabili aleatorie non ridotte, avremo in modo analogo

$$\lim_{n\to\infty} P\left[np + x_1\sqrt{npq} \leq S_n \leq np + x_2\sqrt{npq}\right] = \frac{1}{\sqrt{2\ \pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx$$

IL TEOREMA DEL LIMITE CENTRALE

Il teorema del limite centrale asserisce che *la distribuzione associata ad una variabile aleatoria S_n, somma di n variabile aleatoria indipendenti X_i tende alla distribuzione normale al crescere di n, qualunque sia la distribuzione delle n variabili, purchè nessun termine della sommatoria sia predominante rispetto agli altri.*

In particolare affinchè sia valido questo teorema, è condizione <u>sufficiente</u> ma <u>non</u> <u>necessaria</u> che le variabili X_i abbiano tutte la stessa distribuzione di probabilità con valore medio μ e varianza σ^2.
Data la limitatezza della nostra esposizione, verifichiamo soltanto la condizione di sufficienza.

Poichè per ipotesi $S_n = \sum_i X_i$ è definita da variabili X_i aventi tutte la stessa distribuzione, possiamo scrivere

$$\mu = E[S_n] = \sum_{i=1}^{n} E[X_i] = nE[X_i]$$

e poichè le X_i sono indipendenti

$$\sigma^2 = \sum_{i=1}^{n} \sigma_i^2 = n\sigma_i^2$$

Per semplificare ulteriormente i calcoli, passiamo alla variabile ridotta

$$S_n^* = \frac{S_n - \mu}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i^*$$

dove

$$X_i{}^* = \frac{X_i - \mu_i}{\sigma_i}$$

con $i = 1, 2, \cdots n$. La funzione generatrice dei momenti $g(t)$ della variabile aleatoria S_n è calcolabile dalle funzioni generatrici $a(t)$, uguali per tutte le X_i, ricordando che, essendo le X_i variabili indipendenti, allora

$$g(\tau) = \prod_{i=1}^{n} A_i(\tau) \qquad\qquad per\ S_n = \sum_{i=1}^{n} X_i$$

Ricorrendo ad un ulteriore piccolo artificio matematico, esprimiamo il parametro τ della funzione generatrice tramite un' altra variabile t con $\tau = \frac{t}{\sqrt{n}}$. Avremo allora

$$g(t) = \left[a\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left\{E\left[e^{\frac{tX_i^*}{\sqrt{n}}}\right]\right\}^n$$

Sviluppando l'esponenziale in serie di Taylor, otteniamo

$$g(t) = \left[\int_{-\infty}^{+\infty}\left[1 + \frac{t}{\sqrt{n}} x_i^* + \cdots + \frac{1}{n!}\left(\frac{t}{\sqrt{n}}\right)^n x_i^{*n}\right] f(x_i^*) dx_i^*\right]^n$$

Introducendo i momenti di ordine k della distribuzione μ_k, riscriviamo in modo più compatto lo sviluppo

$$g(t) = \left[\sum_{k=0}^{\infty} \frac{\mu_k}{k!}\left(\frac{t}{\sqrt{n}}\right)^k\right]^n$$

Le X_i^* sono variabili ridotte, quindi $\mu_1 = 0, \quad \mu_2 = 1$

$$g(t) = \left(1 + \frac{t^2}{2\ n} + \cdots\right)^n$$

e

$$\lim_{n\to\infty} g(t) = \lim_{n\to\infty} \left(1 + \frac{t^2}{2\ n} + \cdots\right)^n = e^{t^2/2}$$

Questa é la *funzione generatrice della distribuzione normale ridotta* ed il teorema di unicità ci garantisce che la funzione di distribuzione avente tale funzione generatrice non può che essere la funzione di Gauss.

Nei casi che si presentano in pratica n assume un valore finito, pertanto la funzione di distribuzione della variabile aleatoria oggetto di studio è solamente prossima a quella di Gauss:

$$f(x) \cong \frac{1}{\sigma\sqrt{2\ \pi}} e^{-\frac{(x-\mu)^2}{2\ \sigma^2}}$$

essendo

$$\sigma^2 = \sum_{i=1}^{n} a_i \sigma_i^2$$

Questa approssimazione sarà tanto più buona quanto più n è grande. Supponiamo allora che x_i sia il <u>risultato di una misura</u> della variabile aleatoria X. $E[X] = \mu$ rappresenta il valor vero grandezza. Allora l'errore $\Delta X = X - \mu$ *è pensabile a sua volta come una variabile aleatoria* determinata da *tante cause indipendenti* e definibile quindi come somma di processi indipendenti ΔX_i

$$\Delta X = \sum_i \Delta X_i$$

Questa é la ragione per cui sovente si assume che **la distribuzione di probabilità degli errori casuali sia una Gaussiana a valor medio nullo**.

FUNZIONI DI VARIABILI ALEATORIE

Siano date due variabili aleatorie X ed Y legate da una relazione funzionale nota

$$y = h(X)$$

vogliamo vedere come sia possibile ricavare la funzione di distribuzione della Y, nota la funzione di distribuzione della X.

Distinguiamo il caso in cui la funzione $h(X)$ è monotona. Ciò implica che quando X assume il valore x_i in corrispondenza Y assumerà il valore $y_i = h(x_i)$. Nel caso di relazione $h(x)$ monotona crescente, diremo allora che la probabilità che $Y \leq y_i$

$$P[Y \leq y_i] = P[X \leq x_i]$$

e, nel caso di relazione $h(x)$ monotona decrescente

$$P[Y \leq y_i] = P[X \geq x_i]$$

Usando allora il linguaggio delle funzioni di ripartizione, detta $G(y)$ la funzione di ripartizione della Y e $F(x)$, l'analoga per la X.

$$h(x) \; monotona \; crescente \rightarrow G(y) = F(x)$$

$$h(x) \; monotona \; decrescente \rightarrow G(y) = 1 - F(x)$$

Differenziando ambo i membri ricaviamo la relazione che lega le densità di probabilità

$$g(y) = f(x)\Big|\frac{dx}{dy}\Big|$$

dove il modulo tiene conto dei due possibili casi.

Nel caso in cui la funzione $h(x) = y$ non sia monotona, allora un dato valore y della variabile aleatoria Y corrisponderà a più valori assunti dalla variabile indipendente.

In tal caso la probabilità che Y assuma un certo valore y nell'intervallo $y\ ;\ y+dy$ é pari alla somma delle probabilità che x cada negli intervalli $x_j\ ;\ x_j+dx, \quad x_s\ ;\ x_s+dx, \quad x_K\ ;\ x_K+dx$.

Quindi in sintesi avremo:

$$g(y) = \{f(x_j(y)) + f[x_s(y)] + f[x_K(y)]\}\left|\frac{dx}{dy}\right|$$

ESEMPIO: sia X distribuito normalmente ed inoltre sia $y = x^2$. Applichiamo la relazione sopra scritta.

$$g(y) = \{f[x(y)] + f[-x(y)]\}\,\frac{dx(y)}{dy}$$

dove $x = \pm\sqrt{y}$.
Quindi si ha

$$|\frac{dx}{dy}| = \frac{d}{dy}\sqrt{y} = \frac{1}{2\sqrt{y}}$$

per cui

$$g(y) = \{f(\sqrt{y}) + f(-\sqrt{y})\}\,\frac{1}{2\sqrt{y}}$$

sostituendo, otteniamo

$$g(y) = \frac{1}{2\sqrt{\pi y}}e^{-\frac{y}{2}}$$

Nel caso in cui si abbia una relazione tra variabili aleatorie X, Y, Z del tipo $z = f(x, y)$ e sia $p(x, y)$ la probabilità congiunta delle X e Y si dimostra che

$$h(x, z) = p(x, y)\Big|\frac{\partial y}{\partial z}\Big|$$

e quindi la probabilità marginale per la variabile $z, q(z)$

$$q(z) = \int_{-\infty}^{+\infty} p(x, y)\Big|\frac{\partial y}{\partial z}\Big|dx$$

ESEMPIO IMPORTANTE: Siano X e Y variabili aleatorie indipendenti e connesse tra loro dalla relazione $Z = Y/X$. Possiamo quindi scrivere che

$$dz = \frac{1}{x}dy$$

e

$$h(x, z) = f(x)g(y)|x| = f(x)g(xz)|x|$$

ne segue che

$$q(z) = \int_{-\infty}^{+\infty} |x|f(x)g(xz)dx$$

Supponiamo poi

$$f(x) = e^{-x} \quad per \quad x \geq 0$$

$$g(y) = e^{-y} \quad per \quad y \geq 0$$

In tal caso allora otteniamo

$$q(z) = \int_{-0}^{+\infty} xe^{-x}e^{-xz}dx = (1+z)^{-2} \quad per \quad z \geq 0$$

Applichiamo in un caso particolarissimo quanto dedotto finora.

Consideriamo la funzione cumulativa di una variabile aleatoria X la cui densità di probabilità sia $f(x)$

$$F(x) = \int_{-0}^{+\infty} f(x')dx'$$

$F(x)$ è notoriamente una funzione monotona crescente compresa tra 0 e 1 ed è quindi in corrispondenza biunivoca con la x.

Se trattiamo a sua volta la funzione $F(X)$ come variabile aleatoria, o più esattamente come una funzione di variabile aleatoria

$$Y = F(X)$$

allora possiamo dedurre la funzione densità di probabilità dell Y

$$g(y) = g(F(x)) = f(x)\frac{dx}{dy} = f(x)\left(\frac{dy}{dx}\right)^{-1} = f(x)\frac{1}{f(x)}$$

Concludiamo che la funzione di distribuzione della cumulativa F è <u>uniforme</u>:

$$g(y) = 1$$

Questo risultato ha una importante ricaduta nello sviluppo di metodi per la generazione casuale di numeri secondo un'assegnata distribuzione statistica. É di basilare importanza per sviluppare i programmi al calcolatore per la simulazione di fenomeni casuali complessi (Metodo di Monte Carlo).

In genere i calcolatori destinati al calcolo scientifico dispongono di algoritmi capaci di generare numeri (pseudo) casuali con distribuzione uniforme tra 0 e 1, *cioè con la stessa distribuzione di ogni funzione cumulativa.*

Indichiamo con r un valore estratto dal calcolatore

$$r = \int_{-0}^{+\infty} f(x')dx'$$

Se la dipendenza funzionale F é nota, allora la si puó invertire, cioè si puó calcolare $F^{-1}(r) = x$. Si otterrà cosí il valore di x, cioé una realizzazione della variabile aleatoria X distribuita secondo la funzione densitá di probabilitá $f(x)$.

<u>NUOVI TIPI DI DISTRIBUZIONE :</u>
<u>LA DISTRIBUZIONE MULTINOMIALE</u>

La distribuzione multinomiale è una generalizzazione della distribuzione binomiale. Nel caso della distribuzione binomiale i risultati di un esperimento si raggruppano in due classi; quando tale raggruppamento é basato su r classi, allora si parla di distribuzione multinomiale.

$A_1 \quad A_2 \quad \cdots \quad A_r$ <u>siano le classi</u>

$p_1 \quad p_2 \quad \cdots \quad p_r$ <u>probabilità associate</u>

$$\sum_{i=1}^{r} p_i = 1$$

$n_1 \quad n_2 \quad \cdots \quad n_r$ <u>numero di successi della relativa classe</u>

ovvero il numero di volte in cui il risultato dell'esperimento cade nella relativa classe. La probabilità che si realizzi una particolare sequenza di valori $n_1 \quad n_2 \quad \cdots \quad n_r$ distribuiti nella classe $A_1 \cdots n_r$ in A_r é

$$P_{n_1 \cdots n_2, p_1 \cdots p_2} = \underbrace{\frac{n!}{n_1! \cdots n_r!}} \quad \underbrace{p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}}$$

ESEMPIO: calcoliamo la probabilità che in 36 lanci di un dato a 6 facce, i 6 possibili risultati escano tutti 6 volte per ciascuno.

$$P = \frac{36!}{(6!)^6} \left(\frac{1}{6}^{36} \right) = \frac{3.72 \cdot 10^{41}}{1.393 \cdot 10^{17}} 9.695 \cdot 10^{-29}$$

poichè $\quad n = 36 \quad n_i = 6\ Vi \quad p_i = \frac{1}{6}\ Vi$ da cui

$$P = 2.589 \cdot^{-4}$$

La distribuzione multinomiale è utilizzata per caratterizzare la probabilità di ottenere l' istogramma costruito sul totale di n *eventi, distribuiti in* r *classi, e* n_i $(i = 1, \cdots, r)$ *eventi per classe.*

LA DISTRIBUZIONE χ^2 (O DI PEARSON)

Siano

a) $X_1, X_2, \cdots X_\gamma$, variabili casuali indipendenti distribuite tutte normalmente

b) $m_1, m_2, \cdots m_\gamma$ i rispettivi valori medi

c) $\sigma_1, \sigma_2, \cdots \sigma_\gamma$ le rispettive varianze

Costruiamo allora γ variabili <u>normali, ridotte</u>, quindi per definizione con media nulla e varianza unitaria

$$X_k^* = \frac{X_k - m_k}{\sigma_k}, \qquad k = 1, \cdots, \gamma$$

e definiamo la variabile ${\chi_\gamma}^2$ <u>chi quadro</u>

$${\chi_\gamma}^2 = \sum_{K=1}^{\gamma} X_k^{*2}$$

La funzione di distribuzione di questa variabile si dimostra essere la seguente:

$$f(\chi^2) = \frac{1}{2^{\gamma/2}\Gamma\left(\frac{\gamma}{2}\right)} e^{-\frac{\chi^2}{2}} (\chi^2)^{\frac{\gamma}{2}-1}$$

dove Γ rappresenta la funzione definita

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$$

Essa è chiamata funzione Gamma (o funzione fattoriale) ed ha le seguenti proprietà:

1) $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ in generale $\left(\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\right)$

2) per α intero $\Gamma(\alpha) = (\alpha - 1)!$ $\quad (\Gamma(1) = 1)$

3) per α semintero $\Gamma(\alpha) = (\alpha - 1)(\alpha - 2) \cdots \frac{3}{2} \cdot \frac{1}{2}\sqrt{\pi}$

La funzione generatrice è ottenibile come prodotto delle funzioni generatrici di X_k^{*2}:

$$g(t) = E[e^{\chi^2 t}] \Longrightarrow \prod_{i=1}^{\gamma} g_{x_K^2}(t)$$

dove

$$g_{x_K^2}(t) = \frac{1}{\sqrt{2\ \pi}} \int_{-\infty}^{+\infty} e^{x_K^2}(t) e^{-x_K^2/2} dx_K$$

$$g_{\chi^2}(t) = (1 - 2\ t)^{-\gamma/2}$$

- per $\gamma = 1$ abbiamo che $\lim_{\chi^2 \to 0} f(\chi^2) = \infty$

Per $\gamma \geq 2$ la funzione $f(\chi^2)$ ha un massimo nel punto $\chi^2_{MAX} = \gamma - 2$ <u>valore Modale</u>.

Vediamo ora di calcolare il <u>valore medio</u>

$$E[\chi^2] = \sum_{K=1}^{\gamma} E\left[\left(\frac{\chi_k - m_K}{\sigma_K}\right)^2\right] = \gamma E\left[\left(\frac{\chi_k - m_K}{\sigma_K}\right)^2\right] = \gamma$$

e la <u>varianza</u>

$$\left[\frac{d^2}{dt^2} G_{\chi^2}(t)\right]_{t=0} = \sigma^2$$

dove $G_{\chi^2}(t) = g_{\chi^2}(t) e^{-\gamma t}$

$$\left.\frac{d^2G}{dt^2}\right|_{t=0} = \sigma^2 = 2\ \gamma$$

Vediamo poi di ricavare alcune proprietà fondamentali delle variabili χ^2.

1) Additività delle variabili χ^2.

Consideriamo due variabili ${\chi_{\gamma_1}}^2$ e ${\chi_{\gamma_2}}^2$ tra loro indipendenti. La loro somma è ancora una variabile ${\chi_\gamma}^2$ con $\gamma = \gamma_1 + \gamma_2$ gradi di libertà.

$${\chi_\gamma}^2 = \chi_{\gamma_1} + {\chi_{\gamma_2}}^2 \quad \longrightarrow \quad con\ \ \gamma_1 + \gamma_2 = \gamma$$

Questa proprietà segue dal fatto che

$$g_{x^2}(t) = g_{x_1^2}(t) \cdot g_{x_2^2}(t) = (1-2\,t)^{-\gamma_1/2} \cdot (1-2\,t)^{-\gamma_1/2} = (1-2\,t)^{-(\gamma_1+\gamma_2)/2}$$

UNA PARTICOLARE VARIABILE χ^2.

Siano $X_1, \cdots, X_n$ variabili aleatorie indipendenti normali aventi tutte lo stesso valore aspettato $E[X_i] = \mu$ e la stessa varianza $\sigma^2(X_i) = \sigma^2$. Consideriamo la variabile

$$\bar{X} = \frac{\sum_i X_i}{n}$$

che ha distribuzione normale con valore aspettato μ e varianza

$$\sigma_{\bar{X}}{}^2 = \frac{\sum_i \sigma_i^2}{n^2} = \frac{\sigma^2}{n}$$

Consideriamo inoltre una funzione di queste n variabili aleatorie (lo scarto quadratico) così definita:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 = \frac{\sigma^2}{n-1} Z$$

Z é una nuova variabile il cui comportamento statistico è descritto da una distribuzione di tipo χ^2. Dimostreremo ció in modo rigoroso e metteremo in evidenza il fatto che il numero di gradi di libertà associati a tale variabile è $n-1$. Procediamo applicando un semplice artificio matematico,

$$s^2 = \sum_1^n \frac{(X_i - \bar{X})^2}{n-1} = \frac{\sigma^2}{n-1} \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$$

Sommando e sottraendo in parentesi il valore aspettato μ e sviluppando avremo:

$$s^2 = \frac{\sigma^2}{n-1}\sum_{i=1}^{n}\left((\frac{X_i-\mu}{\sigma})^2 - (\frac{\bar{X}-\mu}{\sigma})^2 \right)$$

tale eguaglianza può essere riscritta come

$$(n-1)\frac{s^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 - \left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

Il primo termine a destra dell'eguaglianza è immediatamente riconoscibile come una variabile χ^2 con n gradi di libertá ed il secondo un χ^2 ad un solo grado di libertá. Per la proprietá additiva delle variabili χ^2 concludiamo che i gradi di libertà della $(n-1)\frac{s^2}{\sigma^2}$ sono quindi $n-1$.

Sulla base della dipendenza di s^2 da Z, possiamo poi dedurre esplicitamente la funzione di distribuzione di s^2:

$$g(s^2) = f(Z)\Big|\frac{dZ}{ds^2}\Big| = f(Z(s^2))\Big|\frac{dZ(s^2)}{ds^2}\Big|$$

e poichè

$$Z = \frac{n-1}{\sigma^2}s^2 \quad \longrightarrow \quad \frac{dZ}{ds^2} = \frac{n-1}{\sigma^2}$$

allora avremo

$$g(s^2) = \left(\frac{n-1}{2\ \sigma^2}\right)^{\frac{n-1}{2}} \frac{(s^2)^{\frac{n-3}{2}}}{2^{\frac{s-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} e^{-\frac{n-1}{2}\frac{}{\sigma^2}s^2}$$

Concludiamo riportando il valore aspettato e la varianza di s^2:

$$E[s^2] = E\left[\frac{\sigma^2}{n-1}Z\right] = \frac{\sigma^2}{n-1}E[Z] = \sigma^2$$

$$Var[s^2] = Var\left[\frac{\sigma^2}{n-1}Z\right] = \frac{\sigma^4}{(n-1)^2}Var[Z] = \frac{2\ \sigma^4}{(n-1)}$$

<u>LA DISTRIBUZIONE DEL T DI STUDENT</u>

Sia X^* una variabile aleatoria <u>normale ridotta</u>

$$E[X^*] = 0 \qquad\qquad Var[X^*] = 1$$

Sia U^2 una variabile aleatoria indipendente da X^* con distribuzione χ^2 a γ gradi di libertà.

Definiamo la variabile aleatoria

$$T = \sqrt{\nu}\frac{X^*}{U}$$

variabile di Student (pseudonimo di Gosset) con $-\infty < t < \infty$

La variabile T è anch'essa aleatoria e la sua funzione densità di probabilitá si dimostra essere pari a

$$f(t) = \frac{\Gamma\left[\frac{1}{2}(\gamma+1)\right]}{\sqrt{\pi\gamma}\,\Gamma\left(\frac{\gamma}{2}\right)}\left(1+\frac{t^2}{\gamma}\right)^{-\left(\frac{\gamma+1}{2}\right)}$$

Infatti poniamo $Y = \sqrt{\gamma}X^*$

$$E[Y] = 0 \qquad\qquad Var[Y] = \gamma Var[Y^*] = \gamma$$

$$g(y) = \frac{1}{\sqrt{2\,\pi\gamma}}e^{-\frac{y^2}{2\,\gamma}}$$

e poniamo $U = \sqrt{\chi^2} \rightarrow 2\,udu = d\chi^2$ allora

$$f(U) = \frac{(U^2)^{\frac{\gamma}{2}-1}e^{-\frac{u^2}{2}}}{2^{\gamma/2}\Gamma\left(\frac{\gamma}{2}\right)}2\,u = 2\,u^{\gamma-1}e^{-\frac{u^2}{2}}a$$

Poichè allora u e y sono indipendenti e sono legate alla variabile t dalla relazione $tu = y$

$$q(t) = \int_0^\infty uf(u)g(tu)du = \frac{2\ a}{\sqrt{2\ \pi\gamma}} \int_0^\infty u^\gamma exp\left[-\frac{u^2}{2}\left(1+\frac{t^2}{\gamma}\right)\right] du$$

sostituendo

$$\omega = \frac{u^2}{2}\left(1+\frac{t^2}{\gamma}\right)$$

da cui si ricava

$$d\omega = \sqrt{2\ \omega(1+t^2/\gamma} du$$

per cui

$$q(t) = \frac{2^{\gamma/2} a}{\sqrt{\pi\gamma}}\left(1+\frac{t^2}{\gamma}\right)^{-\frac{\gamma+1}{2}} \underbrace{\int_0^\infty \omega^{1/2(\gamma-1)} du}_{\Gamma\left(\frac{\gamma=1}{2}\right)}$$

da cui

$$q(t) = \frac{\Gamma\left[\frac{1}{2}(\gamma+1)\right]}{\sqrt{\gamma}\ \Gamma\left(\frac{1}{2}\gamma\right)} \frac{1}{\left(1+\frac{t^2}{\gamma}\right)^{\frac{\gamma+1}{2}}}$$

Vediamo ora di riassumere alcune proprietà della distribuzione T. Osserviamo innanzitutto che il valore aspettato e la varianza della distribuzione sono

$$E[t] = 0\ \ \forall\ \gamma \qquad Var[t] = \begin{cases} \frac{\gamma}{\gamma-2} & \text{per } \gamma > 2 \\ \infty & \text{per } \gamma \leq 2 \end{cases}$$

Per $t = 0$, la funzione di distribuzione ha un massimo, $q(0) = q_{MAX}$, il cui valore è una funzione di γ e risulta contenuto nell'intervallo

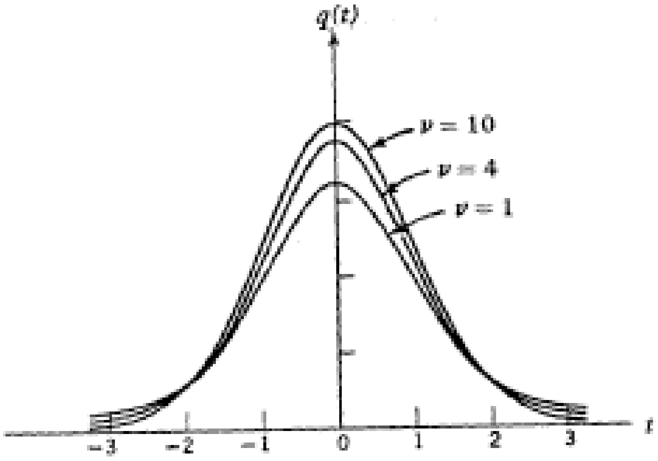
$$\frac{1}{\pi} \leq q_{MAX} \leq \frac{1}{\sqrt{2\ \pi}}$$

Si noti infine che, nel caso in cui $\gamma = 1$, la distribuzione T prende il nome di distribuzione di Cauchy:

$$q(t) = \frac{1}{\pi} \frac{1}{1 + t^2}$$

Per $\gamma = \infty$ la distribuzione T tende alla distribuzione di Gauss

$$q(t) = \frac{1}{\sqrt{2\ \pi}} e^{-t^2/2}$$



.

UNA PARTICOLARE VARIABILE T

.

Consideriamo la variabile $\bar{X}$, la media aritmetica di n variabili casuali indipendenti con distribuzione normale. Tutte le variabili che concorrono a definire $\bar{X}$, abbiano valore aspettato μ e varianza σ^2. Si consideri poi la variabile S^2

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

che ha distribuzione χ^2.

Poichè si dimostra che $\bar{X}$ ed S^2 sono indipendenti, è possibile costruire la variabile

$$t = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{n-1}{\sigma^2}s^2}}\sqrt{n-1}$$

cioè

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Si è così ottenuta una variabile che dipende dal valore aspettato μ ma non più dalla varianza σ. Questa variabile viene utilizzata per stimare l'intervallo fiduciario associato alla media, cioè l'intervallo $(\bar{x} - \Delta\bar{x}, \bar{x} + \Delta\bar{x})$ entro cui cade, con probabilità prefissata (sulla base del valore assunto dalla variabile T), il valore aspettato μ .

LA DISTRIBUZIONE F

Per introdurre la funzione di distribuzione F (o di Fisher) seguiremo uno sviluppo logico analogo a quello presentato nel paragrafo precedente, relativo alla variabile t.
Definiamo una nuova variabile aleatoria F,

$$F = \frac{\gamma U}{\mu V} = \frac{U/\mu}{V/\gamma}$$

dove U e V sono variabili indipendenti che seguono distribuzione χ^2 rispettivamente con μ e γ gradi di libertà.
Notiamo che U e V sono entrambe a distribuzione χ^2, quindi sono variabili non negative. Ne segue che $F \geq 0$
Poniamo $X = U$ ed $Y = U/\mu$ e ricaviamo subito la distribuzionne della Y a partire da quella della X:

$$g(y) = f(x) \cdot \mu = \frac{\mu x^{\frac{\mu}{2}-1} e^{-x/2}}{2^{\mu/2}\Gamma\left(\frac{\mu}{2}\right)} = g(y) = a y^{\frac{\mu}{2}-1} e^{-\frac{1}{2}\mu y}$$

avendo posto

$$a = \frac{\mu^{\mu/2}}{2^{\mu/2}\Gamma\left(\frac{\mu}{2}\right)}$$

Analogamente trattiamo la variabile V/γ

$$f(y) = \frac{\gamma x^{\frac{\gamma}{2}-1} e^{-x/2}}{2^{\gamma/2}\Gamma(\gamma/2)} = b e^{-\frac{1}{2}\gamma y} \cdot y^{\frac{\gamma}{2}-1}$$

avendo posto

$$b = \frac{\gamma^{\gamma/2}}{2^{\frac{\gamma}{2}}\Gamma\left(\frac{\gamma}{2}\right)}$$

Quindi per la variabile F adoperiamo la solita relazione (dedotta dall'esempio della variabile $Z = Y/X$ ed otteniamo

$$q(F) = \int_0^\infty y f(y) g(yF) dy$$

ovvero

$$q(F) = \int_0^\infty y b y^{\frac{\gamma}{2}-1} e^{-\frac{1}{2}\gamma y} a (yF)^{\frac{\mu}{2}-1} e^{-\frac{1}{2}\mu y F} dy =$$

$$= abF^{\frac{\mu}{2}-1} \int_0^\infty y^{\frac{1}{2}(\mu+\gamma-2)} e^{-\frac{y}{2}(\mu F+\gamma)} dy$$

Poniamo

$$\omega = \frac{y}{2}(\mu F + \gamma)$$

da cui consegue che

$$dy = \frac{2\ d\omega}{\mu F + \gamma}$$

La $q(F)$ diviene allora

$$q(F) = \frac{abF^{\frac{\mu}{2}-1} \cdot 2^{\frac{1}{2}(\mu+\gamma)}}{(\mu F + \gamma)^{\frac{1}{2}(\mu+\gamma)}} \underbrace{\int_0^\infty \omega^{\frac{1}{2}(\mu+\gamma)-1} e^{-\omega} dw}_{\Gamma\left(\frac{\mu+\gamma}{2}\right)}$$

e quindi, definendo

$$c = \mu^{\frac{\mu}{2}} \gamma^{\frac{\gamma}{2}} \frac{\Gamma\left(\frac{\mu+\gamma}{2}\right)}{\Gamma\left(\frac{\mu}{2}\right)\Gamma\left(\frac{\gamma}{2}\right)}$$

troviamo l'espressione esplicita della funzione densitá di probabilità della F

$$q(F) = cF^{\frac{1}{2}(\mu-2}(\mu F + \gamma)^{-\frac{1}{2}(\mu+\gamma)}$$

Elenchiamo ora alcune proprietà fondamentali di tale distribuzione.

a) $E[F] = \frac{\gamma}{\gamma-2}$ tale definizione è valida per $\gamma > 2$

b) $Var[F] = \frac{2\ \gamma^2(\mu+\gamma-2)}{\mu(\gamma-2)^2(\gamma-4)}$ tale definizione è valida per $\gamma > 4$.
Ciò segue dal fatto che poichè $F \geq 0$ sarà anche
$E[F] \geq 0$ e $Var[F] \geq 0$

c) la distribuzione $q(F)$ tende alla distribuzione di Gauss se μ e/o γ tendono all'infinito.

La distribuzione F è importante per l'analisi della varianza. Infatti se consideriamo come variabile aleatoria la somma degli scarti dalle rispettive medie di due serie di misure

$$\sum_i S_i'^2 = \sum_{i=1}^{\mu} \frac{(X_i - \bar{X})^2}{\mu} e \sum_{j=1}^{\gamma} S_j''^2 = \sum_i \frac{(Y_i - \bar{Y})}{\gamma}$$

il rapporto di due scarti quadratici

$$F_0 \equiv \frac{\sum_{i=1}^{\mu} S_i'^2}{\sum_{j=1}^{\gamma} S_j''^1}$$

è riconoscibile essere una variabile aleatoria che fluttua con densità di probabilità $q(F)$. Questo segue dal fatto che $S_j^{'}$ e $S_j^{''}$ sono a distribuzione normale e quindi i due scarti quadratici $\sum_{i=1}^{\mu} S_i^{'2}$ e $\sum_{i=1}^{\mu} S_i^{''2}$ sono a distribuzione χ^2 rispettivamente con μ e γ gradi di libertà.