Distribuzioni

Vers. 1.1.0

Gianluca Mastrantonio

gianluca.mastrantonio@polito.it

Outline

1. Note introduttive

2. Distribuzioni
   - Normal Multivariata (Vettori Gaussiani)
   - Chi-quadro
   - T di Student
   - F di Fisher Snedecor

Alcune note sul corso

- Nel corso viene seguito il libro Casella Berger (con poche eccezioni, come la normale multivariata e la regressione multivariata) e ogni volta che finiremo un argomento metterò a quali capitoli faceva riferimento;
- Useremo il software R per fare degli esempi e chiarire con dataset alcuni concetti che useremo, ma non è chiesta la conoscenza di R per poter superare l'esame;
- Prima della lezione metterò a disposizione i codici R e le slides che useremo. Delle slides ci saranno due versioni, una pulita e una su cui scriverò durante la lezione;
- La scritta "chunk x" sulle slides indica che c'è un esempio nel file R allegato, al chunk x, con esempi o più dettagli;
- al fine di ogni sezione caricherò degli esercizi.

Notazione I

La notazione usata dagli statistici è un po' diversa da quella dei probabilisti (e meno precisa), e se alcune cose si possono determinare dal contesto non si indicano. Alcuni esempi

- Sia la funzione di probabilità (pmf) che di densità (pdf) si indicano spesso con $f()$, dove f è la pmf se l'argomento è discreto, o la pfm se continuo, a meno che non serva fare una distinzione;
- La pmf e pdf di una variabile X, dipendente da parametri $\boldsymbol{\theta}$, si può scrivere equivalentemente come

  $$f(X = x) \equiv f_X(x) \equiv f(x) \equiv f(x|\boldsymbol{\theta}) \equiv f(x; \boldsymbol{\theta}) \equiv f_{\boldsymbol{\theta}}(x)$$

  e varie altre permutazioni. L'importante è che dal contesto si capisca a cosa ci si riferisce;

Notazione II

- Se una variabile assume valori solo in un sottoinsieme di $\mathbb{R}$, per esempio $\mathbb{R}^+$, o di un'altro spazio, non c'è bisogno di usare la funzione indicatrice per dire dove la sua densità è zero: per esempio se $X \sim G(\alpha, \beta)$, la sua densità la si indica con
$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$
dove è implicito che vada calcolata solo su $\mathbb{R}^+$.
- Il dominio dei parametri viene in generale omesso e è quello in cui la densità esiste, se è diverso viene specificato;
- Alcune distribuzioni possono essere formalizzate in modi diversi da quello che avete visto con la Siri, per esempio la Gamma può essere anche scritta come $G(\alpha, \theta)$ con densità
$$\frac{\theta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\theta}}$$
altri esempi sono l'esponenziale, le geometrica etc.

Richiami Normale Univariata I

- Richiamiamo dei concetti delle normali univariate. Se $X \sim N(0,1)$ (distribuita come una normale standard), la sua densità è

$$f(x) = (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2}\right)$$

con $\mathrm{E}(X) = 0$ e $\mathrm{Var}(X) = 1$ e funzione caratteristica

$$\varphi_X(t) = \exp\left(-\frac{t^2}{2}\right),\ \forall t \in \mathbb{R}$$

- la variabile aleatorie $Y = \mu + \sigma X$, se $\sigma \neq 0$ è distribuita normalmente ($N(\mu, \sigma^2)$) con densità

$$f(y; \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

con $\mathrm{E}(X) = \mathrm{E}(\mu + \sigma X) = \mu$, $\mathrm{Var}(X) = \mathrm{Var}(\mu + \sigma X) = \sigma^2 \mathrm{Var}(X) = \sigma^2$ e funzione caratteristica

$$\varphi_Y(t) = \varphi_{\mu+\sigma X}(t) = \exp(it\mu)\,\varphi_X(\sigma t) = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)$$

Richiami Normale Univariata II

- Se $\sigma = 0$ la variabile Y è degenere;
- I parametri σ e $-\sigma$ danno luogo alla stessa distribuzione;
- Trasformazione lineare di una variabile normale è ancora normale. Se $Z = a + bY$ allora $Z \sim N(a + b\mu, b^2\sigma^2)$, visto che

  $$Z = a + bY = a + b(\mu + \sigma X) = a + b\mu + b\sigma X = \mu^* + \sigma^* X$$

  dove $\mu^* = a + b\mu$ e $\sigma^* = b\sigma$;

Potete vedere quache esempio in **"Chunk normale"**

Normale Multivariata I

Consideriamo una collezione di dimensione k di variabili aleatorie normali standard $X_i \sim N(0,1)$, con $i = 1, \dots, k$ ($X_i \overset{iid}{\sim} N(0,1)$) e indichiamo con $\mathbf{X} = (X_1, X_2, \dots, X_k)^T$. La densità congiunta è

$$f(\mathbf{x}) = \prod_{i=1}^{k} f(x_i) = \prod_{i=1}^{k} (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{x_i^2}{2}\right) = \prod_{i=1}^{k} (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{x_i^2}{2}\right) =$$

$$(2\pi)^{-\frac{k}{2}} \exp\left(-\frac{\sum_{i=1}^{k} x_i^2}{2}\right) = (2\pi)^{-\frac{k}{2}} \exp\left(-\frac{\mathbf{x}^T\mathbf{x}}{2}\right)$$

è facile da verificare che il vettore $\mathrm{E}(\mathbf{X})$ ha elemento i-esimo $[\mathrm{E}(\mathbf{X})]_i = 0$ e matrice di varianza e covarianza $\mathrm{Var}(\mathbf{X})$ con element ij pari a

$$[\mathrm{Var}(\mathbf{X})]_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Normale Multivariata II

la funzione caratteristica è

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \prod_{i=1}^{k} \varphi_{X_i}(\mathbf{t}) = \prod_{i=1}^{k} \exp\left(-\frac{t_i^2}{2}\right) = \exp\left(-\frac{\mathbf{t}^T\mathbf{t}}{2}\right)$$

Normale multivariate standard (e vettori Gaussiani standard)

Un vettore aleatorio $\mathbf{X}$ a valori in $\mathbb{R}^k$, assolutamente continuo, si dice vettore Gaussiano standard, e si scrive $\mathbf{X} \sim N_k(\mathbf{0}, \mathbf{I}_k) \equiv N(\mathbf{0}, \mathbf{I}_k)$ se ha densità

$$f(\mathbf{x}) = (2\pi)^{-\frac{k}{2}} \exp\left(-\frac{\mathbf{x}^T\mathbf{x}}{2}\right) \forall \mathbf{x} \in \mathbb{R}^k.$$

Un vettore Gaussiano standard si dice essere distribuito come una Normale multivariata standard o, equivalentemente, come una Gaussiana multivariata standard.

Come con il caso univariato, si può generalizzare

Normale Multivariata III

Normale multivariata - Caso generale

Definiamo un vettore $\boldsymbol{\mu} \in \mathbb{R}^n$ e una matrice $\mathbf{A} \in m_{n \times k}(\mathbb{R})$, allora il vettore aleatorio $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{X}$ è un vettore Gaussiano di dimensione n con parametri $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, dove $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$, e si scrive come

$$\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Per costruzione $\boldsymbol{\Sigma} \in m_{n \times n}(\mathbb{R})$.

Indichiamo con μ_i l'i-esimo elemento di $\boldsymbol{\mu}$, a_{ij} il valore in posizione ij di $\mathbf{A}$, e con σ_{ij} il valore in posizione ij di $\boldsymbol{\Sigma}$ se $j \neq j$, mentre l'elemento in posizione ii si indica con σ_i^2).

Normale Multivariata IV

Abbiamo che

$$Y_i = \sum_{h=1}^{k} a_{ih} X_h + \mu_i$$

$$\mathrm{E}(Y_i) = \sum_{h=1}^{k} a_{ih} \mathrm{E}(X_h) + \mu_i = \mu_i$$

$$\mathrm{Var}(Y_i) = \sum_{h=1}^{k} a_{ih}^2 \mathrm{Var}(X_h) = \sum_{h=1}^{k} a_{ih}^2 = [\mathbf{A}\mathbf{A}^T]_{ii} = \sigma_i^2$$

$$\mathrm{Cov}(Y_i, Y_j) = \mathrm{Cov}\left(\sum_{h=1}^{k} a_{ih} X_h + \mu_i, \sum_{l=1}^{k} a_{jl} X_l + \mu_j\right) = \sum_{h=1}^{k} \sum_{l=1}^{k} a_{ih} a_{jl} \mathrm{Cov}(X_h, X_l) =$$

$$\sum_{h=1}^{k} a_{ih} a_{jh} = [\mathbf{A}\mathbf{A}^T]_{ij} = \sigma_{ij}$$

Quindi il vettore $\boldsymbol{\mu}$ è il vettore medie e $\boldsymbol{\Sigma}$ la matrice di varianza e covarianza.

Normale Multivariata V

Gli stessi risultati si possono ottenere utilizzando le regole per il calcolo di medie e varianze di combinazioni lineari di vettori

$$\mathrm{E}(\mathbf{Y}) = \boldsymbol{\mu} + \mathbf{A}\mathrm{E}(\mathbf{X}) = \boldsymbol{\mu}$$

$$\mathrm{Var}(\mathbf{Y}) = \mathbf{A}\mathrm{Var}(\mathbf{X})\mathbf{A}^T = \mathbf{A}\mathbf{I}_k\mathbf{A}^T = \boldsymbol{\Sigma}$$

la funzione caratteristica di $\mathbf{Y}$ è

$$\varphi_{\mathbf{Y}}(\mathbf{t}) = \varphi_{\boldsymbol{\mu}+\mathbf{AX}}(\mathbf{t}) = \exp\left(i\mathbf{t}^T\boldsymbol{\mu}\right)\varphi_{\mathbf{X}}(\mathbf{A}^T\mathbf{t}) = \exp\left(i\mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}|\mathbf{A}^T\mathbf{t}|^2\right)$$

dove possiamo scrivere

$$|\mathbf{A}^T\mathbf{t}|^2 = (\mathbf{A}^T\mathbf{t})^T(\mathbf{A}^T\mathbf{t}) = \mathbf{t}^T\mathbf{A}\mathbf{A}^T\mathbf{t} = \mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}$$

Teorema - Chiusura rispetto a trasformazioni lineari

La famiglia dei vettori gaussiani è chiusa rispetto a trasformazione lineari

Dimostrazione:
Sia $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{X}$ con $\mathbf{X} \sim N_k(\mathbf{0}, \mathbf{I}_k)$, con $\boldsymbol{\mu} \in \mathbb{R}^n$ e dove $\mathbf{A} \in m_{n \times k}(\mathbb{R})$, allora $\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \mathbf{A}\mathbf{A}^t)$. Introduciamo $\mathbf{b} \in \mathbb{R}^q$ e $\mathbf{B} \in m_{q \times n}(\mathbb{R})$ e definiamo la variabile $\mathbf{Z} = \mathbf{b} + \mathbf{B}\mathbf{Y}$ abbiamo che

$$\mathbf{Z} = \mathbf{b} + \mathbf{B}\mathbf{Y} = \mathbf{b} + \mathbf{B}(\boldsymbol{\mu} + \mathbf{A}\mathbf{X}) = (\mathbf{b} + \mathbf{B}\boldsymbol{\mu}) + (\mathbf{B}\mathbf{A})\mathbf{X}$$

dove $(\mathbf{b} + \mathbf{B}\boldsymbol{\mu}) \in \mathbb{R}^q$ e $(\mathbf{B}\mathbf{A}) \in m_{q \times k}(\mathbb{R})$ e quindi

$$\mathbf{Z} \sim N_q(\mathbf{b} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\mathbf{A}\mathbf{A}^T\mathbf{B}^T)$$

$\square$

Normale Multivariata VII

Proposizione - Proprietà delle Matrice di covarianza

Se $\mathbf{\Sigma}$ è una matrice di covarianza di un vettore aleatorio $\mathbf{Y}$ a valori in $\mathbb{R}^n$, allora (i) è simmetrica e (ii) semi-definita positiva

Dimostrazione:

(i) Simmetrica:

$$\sigma_{ij} = [\mathbf{\Sigma}]_{ij} = \mathrm{Cov}(Y_i, Y_j) = \mathrm{Cov}(Y_j, Y_i) = [\mathbf{\Sigma}]_{ji} = \sigma_{ji}$$

(ii) Semi-definita positiva: prendiamo un vettore $\mathbf{u} \in \mathbb{R}^n$, abbiamo che

$$\mathbf{u}^T \mathbf{\Sigma} \mathbf{u} = \mathbf{u}^T \mathrm{Cov}(\mathbf{Y}) \mathbf{u} = \mathbf{u}^T \mathrm{E}\left((\mathbf{Y} - \mathrm{E}(\mathbf{Y})) (\mathbf{Y} - \mathrm{E}(\mathbf{Y}))^T \right) \mathbf{u} =$$

$$\mathrm{E}\left(\mathbf{u}^T (\mathbf{Y} - \mathrm{E}(\mathbf{Y})) (\mathbf{Y} - \mathrm{E}(\mathbf{Y}))^T \mathbf{u} \right) = \mathrm{E}\left(\left(\mathbf{u}^T \mathbf{Y} - \mathbf{u}^T \mathrm{E}(\mathbf{Y}) \right) \left(\mathbf{u}^T \mathbf{Y} - \mathbf{u}^T \mathrm{E}(\mathbf{Y}) \right)^T \right)$$

Normale Multivariata VIII

dove $\mathbf{u}^T\mathbf{Y}$ è uno scalare che ha valore atteso pari a $\mathbf{u}^T\mathrm{E}(\mathbf{Y})$ e quindi

$$\mathbf{u}^T\mathbf{\Sigma}\mathbf{u} = \mathrm{Var}(\mathbf{u}^T\mathbf{Y}) \geq 0$$

□

Proposizione - Radice principale matrice simmetrica e semi-definita positiva

Se $\mathbf{\Sigma}$ è una matrice simmetrica e semi-definita positiva di dimensione $n \times n$, allora esiste uan matrice $\mathbf{A}$ simmetrica e semi-definita positiva di dimensione $n \times n$, tale per cui vale

$$\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^T = \mathbf{A}^2$$

e $\mathbf{A}$ viene chiamata la radice principale di $\mathbf{\Sigma}$

Se la matrice $\mathbf{\Sigma}$ è invetibile, lo è anche la sua radice principale.

Normale Multivariata IX

Possiamo determinare le densità del vettore $\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Proposizione - Densità di $\mathbf{Y}$

Se $\boldsymbol{\Sigma}$ è invertibile, allora $\mathbf{Y}$ è assolutamente continua e ha densità

$$(2\pi)^{-\frac{n}{2}} \det(\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp\left(-\frac{(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y}-\boldsymbol{\mu})}{2}\right)$$

Dimostrazione:

Definiamo $\mathbf{A}$ come la radice principale di $\boldsymbol{\Sigma}$, quindi invertibile e sia

$$\mathbf{X} = \mathbf{A}^{-1}(\mathbf{Y} - \boldsymbol{\mu}) = \mathbf{A}^{-1}\mathbf{Y} - \mathbf{A}^{-1}\boldsymbol{\mu}$$

e quindi

$$\mathbf{X} \sim N_n(\mathbf{A}^{-1}\boldsymbol{\mu} - \mathbf{A}^{-1}\boldsymbol{\mu}, \mathbf{A}^{-1}\boldsymbol{\Sigma}(\mathbf{A}^{-1})^T)$$

ma $\mathbf{A}^{-1}\boldsymbol{\mu} - \mathbf{A}^{-1}\boldsymbol{\mu} = \mathbf{0}$ e $\mathbf{A}^{-1}\boldsymbol{\Sigma}(\mathbf{A}^{-1})^T = \mathbf{A}^{-1}\mathbf{A}\mathbf{A}^T(\mathbf{A}^{-1})^T = \mathbf{I}_n$

Normale Multivariata X

Per costruzione abbiamo $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{X}$, quindi $\mathbf{Y}$ è assolutamente continuo ($\boldsymbol{\mu} + \mathbf{A}\mathbf{X}$ è un diffeomorfismo) e usando il metodo dello Jacobiano abbiamo che la densità è

$$f_{\mathbf{Y}}(y) = \frac{f_{\mathbf{X}}(\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}))}{\det(\mathbf{A})} = \frac{(2\pi)^{-\frac{n}{2}} \exp\left(-\frac{(\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}))^T(\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}))}{2}\right)}{\det(\mathbf{A})}$$

e visto che

$$(\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}))^T(\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu})) = (\mathbf{y}-\boldsymbol{\mu})^T(\mathbf{A}^{-1})^T\mathbf{A}^{-1}(\mathbf{A}^{-1}(\mathbf{y}-\boldsymbol{\mu}) =$$
$$(\mathbf{Y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu})$$

e $\det(\boldsymbol{\Sigma}) = \det(\mathbf{A})\det(\mathbf{A}) = \det(\mathbf{A})^2$ allora

$$f_{\mathbf{Y}}(y) = (2\pi)^{-\frac{n}{2}}\det(\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp\left(-\frac{(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{2}\right)$$

$\square$

Normale Multivariata XI

Osservazioni

- Vale anche il viceversa, cioè se $\mathbf{\Sigma}$ non è invertibile, allora $\mathbf{Y}$ non è assolutamente continua.
- se $\mathbf{\Sigma}$ non è invertibile allora $\mathbf{Y}$ è degenere.

Normale Multivariata XII

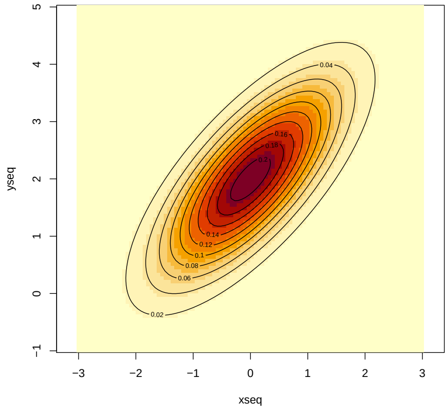
Esempio di una normale bivariata con media

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

e matrice di covarianza

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1.2 \end{pmatrix}$$

Il plot è stata fatto con il "**Chunk normale multivariata**".



Esercizio 1

Sia $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}$ trovare la legge di

$\mathbf{Z} = \begin{pmatrix} x_1 - x_2 \\ x_1 + x_3 \end{pmatrix}$

Soluzione:

Possiamo scrivere $\mathbf{Z} = \mathbf{A}\mathbf{X} + \mathbf{b}$, con $\mathbf{b} = \mathbf{0}$ e $\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, quindi $\mathbf{Z}$ è un vettore gaussiano e $\mathbf{Z} \sim N_2(\mathbf{m}, \mathbf{V})$, dove $\mathbf{m} = \mathbf{0}$ e $\mathbf{V} = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Un'alternativa è calcolare direttamente i valori attesi e varianze e covarianze. $\square$

Esercizio 2

Sia $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ $\mathbf{X}$ è un vettore gaussiano? è assolutamente continuo?

Soluzione:
la matrice $\boldsymbol{\Sigma}$ è simmetrica e definita positiva, visto che il suo determinante è 1, quindi $\mathbf{X}$ è un vettore Gaussiano. Visto che $\boldsymbol{\Sigma}$ è invertibile (è definita positiva), allora $\mathbf{X}$ è assolutamente continua. $\square$

Esercizio 3

Sia $f(\mathbf{x}) = (2\pi)^{-1}3^{-\frac{1}{2}} \exp\left(-\frac{x_1^2+x_1x_2+x_2^2}{3}\right)$, $\forall \mathbf{x} \in \mathbb{R}^2$, la funzione di densità di $\mathbf{X}$. La variabile $\mathbf{X}$ è $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$? e che valore assumono i parametri?

Soluzione:

Dobbiamo avere che $\det(\boldsymbol{\Sigma}) = 3$, e inoltre

$$-\frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2} = -\frac{x_1^2+x_1x_2+x_2^2}{3}.$$

Siccome non ci sono termini con gradi < 2, allora $\boldsymbol{\mu} = \mathbf{0}$, cioè

$$(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = (x_1-\mu_1)^2[\boldsymbol{\Sigma}^{-1}]_{11}+(x_2-\mu_2)^2[\boldsymbol{\Sigma}^{-1}]_{22}+2(x_1-\mu_1)(x_2-\mu_2)[\boldsymbol{\Sigma}^{-1}]_{12}$$

e tutti i prodotti $x_1\mu_1$, $x_2\mu_2$, $x_2\mu_1$ e $x_1\mu_2$ si devono eliminare. Abbiamo anche che

$$(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = 2\frac{x_1^2+x_1x_2+x_2^2}{3}$$

e quindi $\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$ il cui determinante è $\frac{1}{3}$ e quindi quello di $\boldsymbol{\Sigma}$ è 3. La sua inverse è $\boldsymbol{\Sigma} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ $\square$

Esercizio 3

Sia $\mathbf{X} \sim N_2(\mathbf{0}, \mathbf{I}_2)$ e definiamo $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{AX}$, con $\boldsymbol{\mu} = \mathbf{0}$ e $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$.
Determinare la natura di $\mathbf{Y}$.

Soluzione:

Normale Multivariata XVII

La matrice di covarianza di $\mathbf{Y}$ è $\boldsymbol{\Sigma} = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}$ il cui determinante è 0, quindi non invertibile e $\mathbf{Y}$ è degenere.
Si può anche vedere dal valore della correlazione tra Y_1 e Y_2, che è $\frac{[\boldsymbol{\Sigma}]_{12}}{\sqrt{[\boldsymbol{\Sigma}]_{11}[\boldsymbol{\Sigma}]_{22}}} = \frac{4}{\sqrt{16}} = 1$, quindi perfetta correlazione.
Un'altro metodo è vedere come $Y_1 = X_1 + X_2$ e $Y_2 = 2X_1 + 2X_2 = 2Y_1$, e $P(Y_1 = 2Y_2) = P(\mathbf{Y} \in r)$, con $r = \{\mathbf{y} \in \mathbb{R}^2 : y_2 = 2y_1\}$ $\square$

Normale Multivariata XVIII

Proposizione - Distribuzione Marginale

Se $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, allora $\forall\ i = 1, \ldots, n\ X_i \sim N(\mu_i, \sigma_i^2)$

Soluzione:

Possiamo usare la funzione caratteristica della normale multivariata

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left(i\mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right)$$

e, introducendo il vettore $\mathbf{e}_i$ di lunghezza n, che assume valore 0 per tutti gli elementi tranne l'i-esimo, in cui è uno. Allora la funzione caratteristica della i-esima marginale è

$$\varphi_{X_i}(t) = \varphi_{\mathbf{X}}(\mathbf{e}_i t) = \exp\left(i(\mathbf{e}_i t)^T\boldsymbol{\mu} - \frac{1}{2}(\mathbf{e}_i t)^T\boldsymbol{\Sigma}(\mathbf{e}_i t)\right)$$

e visto che $(\mathbf{e}_i t)^T\boldsymbol{\mu} = t\mu_i$ e $(\mathbf{e}_i t)^T\boldsymbol{\Sigma}(\mathbf{e}_i t) = t^2\sigma_i^2$, allora

$$\varphi_{X_i}(t) = \exp\left(it\mu_i - \frac{1}{2}t^2\sigma_i^2\right)$$

che è la funzione caratteristica di una normale univariata di media μ_i e varianza σ_i^2. $\square$

fate attenzione che il viceversa non vale, cioè se $X_1, \ldots, X_n$ sono marginalmente normali, il vettore $\mathbf{X}$ potrebbe non essere normale

Proposizione - Indipendenza

Consideriamo una v.a. $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dividiamola in blocchi come $\mathbf{X} = \begin{pmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{pmatrix}$ dove $\mathbf{X}_A \in \mathbb{R}^{n_A}$ e $\mathbf{X}_B \in \mathbb{R}^{n_B}$, e assumiamo $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{pmatrix}$ e $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{AB}^T & \boldsymbol{\Sigma}_B \end{pmatrix}$ dove $\boldsymbol{\mu}_A$ e $\boldsymbol{\mu}_B$ sono le medie di $\mathbf{X}_A$ e $\mathbf{X}_B$ rispettivamente, mentre $\boldsymbol{\Sigma}_A$ è la matrice di covarianza di $\mathbf{X}_A$, $\boldsymbol{\Sigma}_B$ è la matrice di covarianza di $\mathbf{X}_B$, mentre $\boldsymbol{\Sigma}_{AB}$ la matrice di covarianza tra A e B. Allora $\mathbf{X}_A$ è indipendente da $\mathbf{X}_B \iff \boldsymbol{\Sigma}_{AB}$ è una matrice di zeri.

Soluzione:

(i) $\Longrightarrow$ è ovvio

(ii) per $\Longleftarrow$ definiamo $\mathbf{t} = \begin{pmatrix} \mathbf{t}_A \\ \mathbf{t}_B \end{pmatrix}$ con $\mathbf{t}_A$ e $\mathbf{t}_B$ di lunghezza rispettivamente n_A e n_B, e prendiamo la funzione caratteristica

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left(i\mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right) = \exp\left(i(\mathbf{t}_A^T\boldsymbol{\mu}_A + \mathbf{t}_B^T\boldsymbol{\mu}_B) + -\frac{1}{2}(\mathbf{t}_A^T\boldsymbol{\Sigma}_A\mathbf{t}_A + \mathbf{t}_B^T\boldsymbol{\Sigma}_B\mathbf{t}_B)\right)$$

che è uguale a

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left(i\mathbf{t}_A^T\boldsymbol{\mu}_A + -\frac{1}{2}\mathbf{t}_A^T\boldsymbol{\Sigma}_A\mathbf{t}_A\right)\exp\left(i\mathbf{t}_B^T\boldsymbol{\mu}_B + -\frac{1}{2}\mathbf{t}_B^T\boldsymbol{\Sigma}_B\mathbf{t}_B\right) = \varphi_{\mathbf{X}_A}(\mathbf{t}_A)\varphi_{\mathbf{X}_B}(\mathbf{t}_B)$$

e quindi $\mathbf{X}_A \perp\!\!\!\perp \mathbf{X}_B$ $\square$

Si può generalizzare il risultato a più di due blocchi, anche univariati.

Proposizione - Condizionamento

Consideriamo una v.a. $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Dividiamola in blocchi cone $\mathbf{X} = \begin{pmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{pmatrix}$ dove $\mathbf{X}_A \in \mathbb{R}^{n_A}$ e $\mathbf{X}_B \in \mathbb{R}^{n_B}$, e assumiamo $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{pmatrix}$ e $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{BA} & \boldsymbol{\Sigma}_B \end{pmatrix}$ dove $\boldsymbol{\mu}_A$ e $\boldsymbol{\mu}_B$ sono le medie di $\mathbf{X}_A$ e $\mathbf{X}_B$ rispettivamente, mentre $\boldsymbol{\Sigma}_A$ è la matrice di covarianza di $\mathbf{X}_A$, $\boldsymbol{\Sigma}_B$ è la matrice di covarianza di $\mathbf{X}_B$, mentre $\boldsymbol{\Sigma}_{AB}$ la matrice di covarianza tra A e B.

Allora la distribuzione di $\mathbf{X}_A | \mathbf{X}_B = \mathbf{x}_B$ è ancora normale multivariata con

$$\boldsymbol{\mu}_{A|B} = \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B)$$

$$\boldsymbol{\Sigma}_{A|B} = \boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{BA}$$

Soluzione:
Definiamo la seguente v.a. $\mathbf{Z} = \mathbf{X}_A + \mathbf{A}\mathbf{X}_B$, con $\mathbf{A} = -\mathbf{\Sigma}_{AB}\mathbf{\Sigma}_B^{-1}$ in modo tale che $\mathbf{Z} \perp\!\!\!\perp \mathbf{X}_B$:

$$\mathrm{Cov}(\mathbf{Z}, \mathbf{X}_B) = \mathrm{Cov}(\mathbf{X}_A, \mathbf{X}_B) + \mathrm{Cov}(\mathbf{A}\mathbf{X}_B, \mathbf{X}_B) = \mathbf{\Sigma}_{AB} - \mathbf{\Sigma}_{AB}\mathbf{\Sigma}_B^{-1}\mathbf{\Sigma}_B = \mathbf{0}$$

dove in questo caso $\mathbf{0}$ è una matrice di zeri e siccome $(\mathbf{Z}, \mathbf{X}_B)$ sono un vettore gaussiano sono anche indipendenti.

Visto che

$$\mathbf{X}_A = \mathbf{Z} - \mathbf{A}\mathbf{X}_B$$

abbiamo che $\mathbf{X}_A|\mathbf{X}_B = \mathbf{x}_B \sim \mathbf{Z} - \mathbf{A}\mathbf{x}_B$ è gaussiano con

$$\mathrm{E}(\mathbf{X}_A|\mathbf{x}_B) = \mathrm{E}(\mathbf{Z} - \mathbf{A}\mathbf{X}_B|\mathbf{x}_B) = \mathrm{E}(\mathbf{Z}|\mathbf{x}_B) - \mathrm{E}(\mathbf{A}\mathbf{X}_B|\mathbf{x}_B) = \mathrm{E}(\mathbf{Z}) - \mathbf{A}\mathbf{x}_B$$

dove $\mathrm{E}(\mathbf{Z}) = \mathrm{E}(\mathbf{Z}|\mathbf{x}_B)$ perchè c'è indipendenza. Allora

$$\mathrm{E}(\mathbf{X}_A|\mathbf{x}_B) = \boldsymbol{\mu}_A - \mathbf{A}(\mathbf{x}_B - \boldsymbol{\mu}_B) = \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B)$$

$$\begin{aligned}\mathrm{Var}(\mathbf{X}_A|\mathbf{x}_B) =& \mathrm{Var}(\mathbf{Z} - \mathbf{A}\mathbf{X}_B|\mathbf{x}_B) = \mathrm{Var}(\mathbf{Z}) = \mathrm{Var}(\mathbf{X}_A + \mathbf{A}\mathbf{X}_B) = \\ & \mathrm{Var}(\mathbf{X}_A) + \mathbf{A}\mathrm{Var}(\mathbf{X}_B)\mathbf{A}^T + \mathbf{A}\mathrm{Cov}(\mathbf{X}_B, \mathbf{X}_A) + \mathrm{Cov}(\mathbf{X}_A, \mathbf{X}_B)\mathbf{A}^T = \\ & \boldsymbol{\Sigma}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_B\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{BA} - 2\boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{BA} = \\ & \boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{BA}\end{aligned}$$

$\square$

Normale Multivariata XXIV

Si possono riarrangiare gli elementi di $\mathbf{X}$ per dimostrare che per qualsiasi partizione di $\mathbf{X}$, in $\mathbf{X}_A$ e $\mathbf{X}_B$, la distribuzione di $\mathbf{X}_A|\mathbf{X}_B$ è normale con media e varianza

$$\begin{aligned}
\boldsymbol{\mu}_{A|B} &= \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B) \\
\boldsymbol{\Sigma}_{A|B} &= \boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{BA}
\end{aligned}$$

Alcuni esempi di distribuzione conditionate della normale multivariata si possono vedere su "**Chunk normale multivariata**".

Normale Multivariata XXV

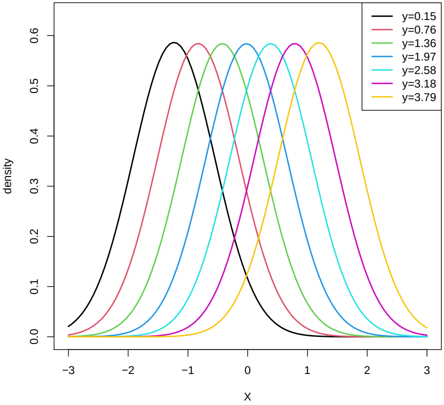
Esempi della distribuzione di $X|Y = y$, con $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

e matrice di covarianza

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1.2 \end{pmatrix}$$

per vari valori di y.
Il plot è stata fatto con il "**Chunk normale multivariata**".



Chi-quadro I

Chi-quadro

Una variabile aleatoria $\mathbf{Y} \in \mathbb{R}^+$ assolutamente continua si dice essere distribuita come una chi-quadro con $n \in \mathbb{N}$ (zero escluso) gradi di libertà (gdl), $Y \sim \chi^2(n)$, se

$$f(y) = 2^{-\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)^{-1} y^{\frac{n}{2}-1} \exp\left(-\frac{y}{2}\right)$$

Abbiamo che $\mathrm{E}(Y) = n$ e $\mathrm{Var}(Y) = 2n$.

Alcune interessanti proprietà

- Se $X_1, \ldots X_n$ sono iid normali standard, allora $Q = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$ - la chi-quadro viene definita partendo da questa relazione;

Chi-quadro II

- La somma di due chi-quadro indipendenti di n_1 e n_2 gdl è una chi quadro di $n = n_1 + n_2$ gradi di libertà. La dimostrazione è molto semplice visto che

$$\sum_{i=1}^{n} X_i^2 \sim \chi^2(n) = \sum_{i=1}^{n_1} X_i^2 + \sum_{i=1}^{n_2} X_{i+n_1}^2$$

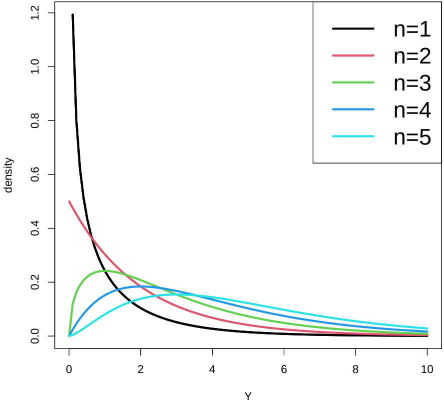
  e $Q_1 = \sum_{i=1}^{n_1} X_i^2 \sim \chi^2(n_1)$ e $Q_2 = \sum_{i=1}^{n_2} X_{i+n_1}^2 \sim \chi^2(n_2)$, dove per costruzione $Q_1 \perp\!\!\!\perp Q_2$;
- la $\chi^2(n)$ è una $G(n/2, 2)$ (fate attenzione a come definite il secondo parametro della gamma, altrimenti è $G(n/2, 1/2)$), visto che una $G(a, b)$ ha densità

$$f(y) = b^{-a}\Gamma(a)^{-1} y^{a-1} \exp\left(-\frac{y}{b}\right)$$

- Se $X_1, \ldots X_n$ sono indipendenti e $X_i \sim N(\mu_i, 1)$, allora $Q = \sum_{i=1}^{n} X_i^2$ è distribuita come una chi-quadro non centrata con n gdl e parametro di non centralità $\lambda = \sum_{i=1}^{n} \mu_i^2$ e si scrive $Q \sim \chi_\lambda^2(n)$.
  La media del chi-quadro non centrato è $n + \lambda$ e la varianza $2(n + 2\lambda)$;
- Se $X \sim U(0, 1)$ allora $-2\log(X) \sim \chi^2(2)$.

Chi-quadro III

Alcune densità della chi-quadro per differenti valori dei gradi di libertà ("**chunk chi-quadro**")



T di Student I

T di Student

Una variabile aleatoria $\mathbf{Y} \in \mathbb{R}$ assolutamente continua si dice essere distribuita come una T di Student con $n \in \mathbb{R}^+$ gradi di libertà (gdl), $Y \sim t(n)$, se

$$f(y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{y^2}{n}\right)^{-\frac{n+1}{2}}$$

La T di Student ha media 0 se $n > 0$, e varianza $\frac{n}{n-2}$ se $n > 2$, è infinita se $1 < n \leq 2$, altrimenti è indefinita. La distribuzione è simmetrica e unimodale.

- La distribuzione T di Student può essere definita come la distribuzione della variabile aleatoria

$$Y = \frac{Z}{\sqrt{\frac{X}{n}}}$$

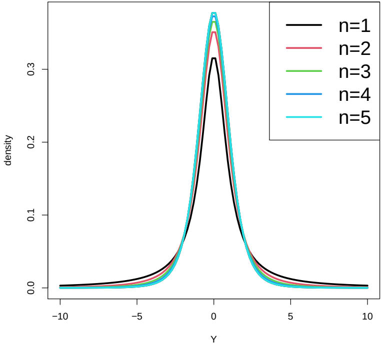
dove $Z \sim N(0,1)$, $X \sim \chi^2(n)$, e Z e X sono indipendenti.

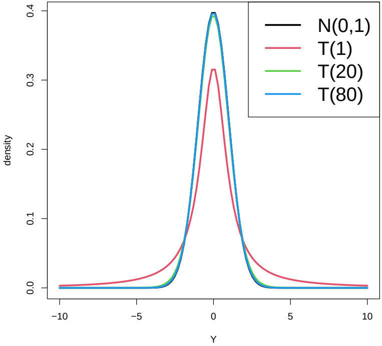
T di Student II

- Una generalizzazione si può ottenere se $Z \sim N(\mu, 1)$, e in questo caso Y è distribuita come una T non centrata, con parametro di non centralità μ: $Y \sim t_\mu(n)$.
- Con n abbastanza grande (più di 80), la T di Student può essere approssimata da una normale standard.

T di Student III

Alcune densità della T di Student per differenti valori dei gradi di libertà (sinistra) e confronto con normale standard (destra) (“**chunk T di Student**”).





F di Fisher Snedecor

Una variabile aleatoria $\mathbf{Y} \in \mathbb{R}^+$ assolutamente continua si dice essere distribuita come una F di Fisher Snedecor con $n_1 \in \mathbb{R}^+$ e $n_2 \in \mathbb{R}^+$ gradi di libertà, $Y \sim F(n_1, n_2)$, se

$$f(y) = \frac{\sqrt{\frac{(n_1 y)^{n_1} n_2^{n_2}}{(n_1 y + n_2)^{n_1+n_2}}}}{yB\left(\frac{n_1}{2}, \frac{n_2}{2}\right)}$$

La F di Fisher Snedecor ha media $\frac{n_2}{n_2-1}$ se $n_2 > 0$, altrimenti è indefinita, e varianza $\frac{2n_2^2(n_1+n_2-1)}{n_1(n_2-2)^2(n_2-4)}$ se $n_2 > 4$, altrimenti è indefinita.

- La distribuzione F di Fisher Snedecor può essere definita come la distribuzione della variabile aleatoria

$$Y = \frac{X_1/n_1}{X_2/n_2}$$

dove $X_1 \sim \chi^2(n_1)$, $X_2 \sim \chi^2(n_2)$, e X_1 e X_2 sono indipendenti.

F di Fisher Snedecor II

- se $X \sim t(n)$, allora $X^2 \sim F(1, n)$, e $X^{-2} \sim F(n, 1)$;
- se $X \sim B(n_1/2, n_2/2)$, allora $\frac{n_2 X}{n_1(1-X)} \sim F(n_1, n_2)$.

F di Fisher Snedecor III

Alcune densità della F di Fisher Snedecor per differenti valori dei gradi di libertà ("**chunk F di Fisher Snedecor**").