

# Appunti del corso di Affidabilità e Diagnostica dei Sistemi Elettrici

Andrea Cavallini, Gian Carlo Montanari  
DIE-Università di Bologna  
viale Risorgimento 2, 40136 Bologna  
andrea.cavallini@mail.ing.unibo.it  
<http://limat.ing.unibo.it>

A.A 1999/2000

# Indice

<b>1</b>	<b>Calcolo delle probabilità</b>	<b>6</b>
1.1	Esperimento aleatorio . . . . .	6
1.2	Eventi e spazi rappresentativi . . . . .	6
1.3	Algebra degli eventi . . . . .	7
1.4	Probabilità . . . . .	9
1.5	Alcune conseguenze degli assiomi (1)-(3) . . . . .	10
1.5.1	Probabilità di $\emptyset$ . . . . .	10
1.5.2	Probabilità di $\bar{\mathcal{E}}$ . . . . .	10
1.5.3	Probabilità di $\mathcal{E} - \mathcal{F} = \mathcal{E}\bar{\mathcal{F}}$ . . . . .	10
1.5.4	Probabilità di $\mathcal{E} + \mathcal{F}$ . . . . .	11
1.6	Combinazioni di eventi equiprobabili: il campionamento . . . . .	13
1.6.1	Campionamento con reintroduzione . . . . .	17
1.6.2	Campionamento senza reintroduzione . . . . .	18
1.7	La legge dei grandi numeri . . . . .	19
<b>2</b>	<b>Indipendenza e dipendenza stocastica</b>	<b>20</b>
2.1	Probabilità condizionata . . . . .	20
2.2	Indipendenza stocastica . . . . .	21
2.2.1	Chain rule . . . . .	22
2.2.2	Esempi . . . . .	22
2.3	Numero di successi in esperimenti ripetuti . . . . .	24
2.4	Approssimazioni della distribuzione binomiale . . . . .	26
2.4.1	Il teorema di deMoivre Laplace . . . . .	26
2.4.2	Il teorema di Poisson . . . . .	27
2.5	Probabilità totale . . . . .	30
<b>3</b>	<b>Variabili aleatorie</b>	<b>32</b>
3.1	Il concetto di variabile aleatoria . . . . .	32
3.2	Eventi . . . . .	33
3.3	Funzioni di distribuzione . . . . .	34
3.4	Densità di probabilità . . . . .	37
3.4.1	Classificazione delle VA ed eventi delle VA continue . . . . .	37
3.4.2	Densità di probabilità . . . . .	39
3.4.3	VA discrete come caso particolare di VA continue . . . . .	39

3.4.4	VA miste . . . . .	40
3.5	Percentili . . . . .	40
3.6	Trasformazioni lineari . . . . .	41
3.7	Funzioni di uso comune . . . . .	42
3.7.1	Distribuzione normale (gaussiana) . . . . .	43
3.7.2	Distribuzione lognormale . . . . .	48
3.7.3	Distribuzione di Weibull . . . . .	50
3.7.4	Distribuzione esponenziale . . . . .	51
3.7.5	Distribuzione chi-quadro . . . . .	52
3.7.6	Legge di probabilità, di Student . . . . .	53
3.7.7	La distribuzione $F$ di Snedecor . . . . .	54
3.8	Distribuzioni condizionate . . . . .	54
3.9	Appendice 1: L'impulso di Dirac e derivata generalizzata . . . . .	57
3.9.1	Definizione . . . . .	57
3.9.2	Derivata di funzioni con discontinuità . . . . .	59
3.10	Appendice 2: Tavole della distribuzione normale . . . . .	60
3.11	Appendice 3: Tavole della distribuzione chi-quadro . . . . .	63
<b>4</b>	<b>Variabili aleatorie bivariate</b>	<b>65</b>
4.1	Eventi . . . . .	65
4.2	Distribuzione e densità di probabilità . . . . .	66
4.3	Distribuzioni marginali . . . . .	66
4.4	Variabili aleatorie congiuntamente normali . . . . .	69
4.5	Indipendenza stocastica . . . . .	69
4.6	Alcune funzioni di VA doppie . . . . .	70
4.6.1	Somma di due variabili aleatorie . . . . .	70
4.6.2	Differenza di due VA . . . . .	72
4.6.3	Massimo di due VA . . . . .	73
4.6.4	Minimo di due VA . . . . .	74
4.7	Distribuzioni condizionate . . . . .	75
4.7.1	Variabili aleatorie congiuntamente normali . . . . .	77
<b>5</b>	<b>Momenti di una variabile aleatoria</b>	<b>79</b>
5.1	Previsione di una variabile aleatoria . . . . .	79
5.1.1	Previsione di una sequenza di dati . . . . .	79
5.1.2	Comportamento asintotico: il valore atteso e media . . . . .	80
5.1.3	La probabilità come valore atteso . . . . .	82
5.1.4	Esistenza del valore atteso . . . . .	82
5.1.5	Linearità del valore atteso . . . . .	83
5.1.6	Altre misure di intensità . . . . .	83
5.2	Momenti del secondo ordine di VA univariate: varianza . . . . .	84
5.3	Il lemma di Tchebycheff . . . . .	86
5.4	Altre misure di dispersione . . . . .	88
5.5	Momenti di ordine superiore a 2 . . . . .	88

5.6	Momenti del secondo ordine di VA doppie: covarianza . . . . .	89
5.6.1	Trasformazioni lineari . . . . .	92
5.7	Il teorema del limite centrale . . . . .	94
5.8	Valore atteso e varianza condizionati . . . . .	96
<b>6</b>	<b>Affidabilità</b>	<b>99</b>
6.1	Generalità sul guasto . . . . .	99
6.2	Sistemi non riparabili . . . . .	101
6.2.1	Funzioni affidabilistiche empiriche . . . . .	103
6.2.2	Il tasso di guasto istantaneo . . . . .	104
6.2.3	Parametri affidabilistici . . . . .	106
6.3	Tasso di guasto per componenti elettronici . . . . .	107
6.4	Generalità, concetto di missione . . . . .	110
6.5	Il diagramma affidabilistico . . . . .	110
6.6	Strutture semplici . . . . .	112
6.6.1	Sistemi di tipo serie . . . . .	112
6.6.2	Sistemi di tipo parallelo (ridondanza) . . . . .	112
6.6.3	Combinazione di strutture tipo serie e parallelo . . . . .	113
6.6.4	Influenza del modo di guasto dei dispositivi . . . . .	113
6.7	Strutture complesse . . . . .	116
6.7.1	Il metodo della probabilità totale . . . . .	116
6.7.2	Il metodo dello spazio degli stati . . . . .	117
<b>7</b>	<b>Disponibilità</b>	<b>119</b>
7.1	Definizioni . . . . .	119
7.1.1	Analisi con le catene di Markov . . . . .	120
7.2	Analisi combinatoria . . . . .	121
7.2.1	Frequenza . . . . .	121
7.3	Analisi di sistemi serie/parallelo . . . . .	124
7.3.1	Sistemi con dispositivi a guasti indipendenti . . . . .	124
7.3.2	Sistemi con dispositivi a guasti dipendenti . . . . .	127
7.4	Ridondanza . . . . .	129
7.5	Analisi affidabilistica di un sistema di distribuzione radiale . . . . .	132
7.5.1	Considerazioni generali . . . . .	132
7.5.2	Criterio di guasto . . . . .	134
7.5.3	Sistema radiale semplice (1) . . . . .	135
7.5.4	Sistema radiale con selezione della linea di alimentazione esterna all'impianto. . . . .	138
7.5.5	Sistema radiale con selezione della linea di alimentazione alla sbarra di media tensione dell'utente. . . . .	141
7.5.6	Sistema radiale con selezione della linea di alimentazione al primario del trasformatore. . . . .	143
7.5.7	Sistema radiale con selezione della linea di alimentazione al secondario del trasformatore. . . . .	145

<b>8</b>	<b>Metodi empirici</b>	<b>147</b>
8.1	Stima empirica delle leggi di probabilità . . . . .	147
8.2	Percentili . . . . .	152
8.3	Carte probabilistiche . . . . .	152
8.4	Stima empirica di momenti e percentili . . . . .	154
8.4.1	Valore atteso . . . . .	154
8.4.2	Varianza . . . . .	155
8.4.3	Covarianza e correlazione empiriche . . . . .	155
8.4.4	Momenti . . . . .	156
<b>9</b>	<b>Stime puntuali</b>	<b>157</b>
9.1	Introduzione . . . . .	157
9.2	Proprietà degli stimatori . . . . .	157
9.3	Il metodo dei momenti . . . . .	161
9.4	Principio di massima verosimiglianza, ML . . . . .	163
9.4.1	Proprietà dello stimatore ML . . . . .	164
9.4.2	Stima ML della probabilità di un evento . . . . .	165
9.4.3	Stima dei parametri di una distribuzione normale . . . . .	166
9.4.4	Stima ML del tasso di guasto . . . . .	166
9.4.5	Stima dei parametri di una distribuzione di Weibull . . . . .	168
<b>10</b>	<b>Stime per intervalli</b>	<b>170</b>
10.1	Introduzione . . . . .	170
10.2	Quantità pivotali . . . . .	170
10.2.1	Il metodo della quantità pivotale . . . . .	173
10.3	Campionamento da una distribuzione normale . . . . .	174
10.3.1	Calcolo degli intervalli di confidenza per la media . . . . .	174
10.3.2	Varianza . . . . .	176
10.3.3	Rapporto di varianze . . . . .	176
10.4	Il metodo statistico . . . . .	177
10.5	Intervallo di confidenza per la probabilità . . . . .	179
10.5.1	Calcolo mediante l'approssimazione normale . . . . .	181
10.6	Intervallo di confidenza per $\lambda$ di una distribuzione esponenziale . . . . .	182
<b>11</b>	<b>Verifica delle ipotesi</b>	<b>185</b>
11.1	Introduzione . . . . .	185
11.2	Ipotesi parametriche . . . . .	185
11.2.1	Esempio di test per la media . . . . .	188
11.2.2	Ipotesi semplici e composte . . . . .	189
11.3	Test bidirezionali . . . . .	192
11.3.1	Intervalli di confidenza . . . . .	193
11.4	Test unidirezionali . . . . .	194
11.4.1	Intervalli di confidenza . . . . .	194
11.5	Test sulla media . . . . .	195

11.5.1	Test bidirezionale . . . . .	195
11.5.2	Test unidirezionali . . . . .	197
11.6	Test sulla varianza per distribuzioni normali . . . . .	198
11.7	Test sul rapporto delle varianze per distribuzioni normali . . . . .	198
11.7.1	Test bidirezionali . . . . .	198
11.7.2	Test unidirezionali . . . . .	199
11.8	Test su due medie . . . . .	199
11.8.1	Varianze identiche . . . . .	199
11.8.2	Varianze diverse . . . . .	200
11.9	Test bilaterali . . . . .	201
11.9.1	Test bilaterale sulla probabilità . . . . .	202
11.9.2	Test bilaterale su <i>MTBF</i> . . . . .	203
11.9.3	Test sequenziali . . . . .	205
11.10	Test non parametrici . . . . .	206
11.10.1	Test di adattamendo del chi quadrato . . . . .	206

## Capitolo 3

# Variabili aleatorie

### 3.1 Il concetto di variabile aleatoria

Una variabile aleatoria, VA, é una funzione definita fra lo spazio rappresentativo degli eventi di un esperimento aleatorio,  $S$ , ed uno degli usuali spazi numerici: lo spazio dei numeri interi,  $Z$ , lo spazio dei numeri reali,  $R$ , lo spazio dei numeri complessi,  $C$ . Una VA puó anche essere vista come una regola che permette di associare un numero ad ogni evento elementare in  $S$ .

A livello di notazione, se  $\zeta_i$  é uno dei possibili eventi elementari dell'esperimento, si puó scrivere:

$$x(\zeta_i) = x \tag{3.1}$$

evidenziando il fatto che la VA é una funzione che ad ogni evento elementare associa un numero. In generale, però, si indicherá la VA come:

$$\mathbf{x} = x$$

dove il grassetto sottintende che non ci troviamo di fronte ad un numero, ma ad una particolare realizzazione (o, in altre parole, risultato di un particolare esperimento) di tutti i possibili risultati di un esperimento aleatorio.

Per fare qualche esempio, si consideri l'esperimento relativo al lancio di una moneta. E' tipico quantificare il risultato di questo esperimento utilizzando la VA  $\mathbf{x}$  definita dalla seguente regola:

$$\mathbf{x} = \begin{cases} 0 & \text{se é testa} \\ 1 & \text{se é croce} \end{cases}$$

Le possibili realizzazioni della VA cosí definita in un esperimento ripetuto tre volte sono 000, 001, 010, ..., 111.

Si noti che la mappatura fra  $S$  e lo spazio numerico considerato puó non essere univoca. Ciò accade quando non si definisce la VA a partire da eventi elementari. Ad esempio, nell'esperimento del lancio del dado la VA  $\mathbf{x}$  definita come:

$$\mathbf{x} = \begin{cases} 0 & \text{se é un numero pari} \\ 1 & \text{se é un numero dispari} \end{cases}$$

non é una mappatura univoca dello spazio  $S$  (a 0 e a 1 corrispondono piú eventi elementari), mentre una mappatura univoca é, ad esempio:

$$\mathbf{x} = \begin{cases} 1 & \text{se é 1} \\ 2 & \text{se é 2} \\ \dots & \end{cases}$$

Spesso é necessario definire la VA come mappature fra  $S$  e prodotti cartesiani di spazi numerici (ad esempio  $\mathbb{R}^n$ ). Un esempio per chiarire ciò é dato dalla lettura di un barometro elettronico dove, simultaneamente, sia possibile leggere temperatura, umidità e pressione. Se consideriamo aleatorio lo stato meteorologico, allora la lettura é una VA da  $S$  (che definire non é semplice) ad  $\mathbb{R}^3$ . Tali variabili aleatorie, dette multivariate, possono essere rappresentate come n-uple o vettori colonna del tipo:

$$\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T \quad (3.2)$$

## 3.2 Eventi

L'evento della VA non é, in generale, l'evento dell'esperimento. Generalmente (con l'eccezione delle VA discrete che saranno discusse piú avanti), per evento di una variabile aleatoria si intende il fatto che la realizzazione della variabile aleatoria in un certo esperimento sia compresa in un dato intervallo numerico. L'evento di un esperimento può essere qualcosa che non é strettamente numerico (ad esempio, si verifica testa nel lancio di una moneta). Esiste tuttavia una corrispondenza fra gli eventi elementari dell'esperimento e gli eventi della variabile aleatoria. Tale corrispondenza é indotta dalla definizione di VA, cioè dall'equazione (3.1). Tale corrispondenza permette di associare ad eventi della variabile aleatoria una probabilità, note le probabilità degli eventi elementari dell'esperimento.

**Esempio 3.2.1** Per chiarire come sia possibile passare dalla probabilità di eventi elementari dell'esperimento alla probabilità di eventi della variabile aleatoria, si consideri la seguente VA definita in base al risultato del lancio di un dado:

$$x(f_n) = 10n$$

dove  $f_n$  é l'evento elementare per cui la faccia uscita ha  $n$  puntini. Allora i seguenti eventi di  $\mathbf{x}$  sono in corrispondenza biunivoca con i seguenti insiemi di eventi elementari di  $S$ :

- $\{\mathbf{x} \leq 35\} \leftrightarrow \{f_1, f_2, f_3\}$
- $\{\mathbf{x} \leq 5\} \leftrightarrow \{\emptyset\}$
- $\{20 \leq \mathbf{x} \leq 35\} \leftrightarrow \{f_2, f_3\}$
- $\{\mathbf{x} = 40\} \leftrightarrow \{f_4\}$
- $\{\mathbf{x} = 35\} \leftrightarrow \{\emptyset\}$

Poiché gli eventi  $\zeta_i$  sono elementari, quindi escludentisi a vicenda, la probabilità di un evento di una variabile aleatoria coincide con la somma delle probabilità degli eventi elementari dell'esperimento aleatorio con cui é in corrispondenza. Riferendosi al caso precedente:

$$\Pr(\{\mathbf{x} \leq 35\}) = \Pr(f_1) + \Pr(f_2) + \Pr(f_3) = 1/6 + 1/6 + 1/6 = 0.5$$

**Esempio 3.2.2** Si consideri una centrale elettrica costituita da 4 unità di potenza 20 MW. Se  $Q = 0.001$  é la probabilità di guasto (inaffidabilità) di una unità, assumendo che i guasti siano indipendenti, é possibile calcolare la probabilità dei valori assumibili dalla VA  $\mathbf{x}$  "massima potenza generabile". Tale VA assume i valori 0, 20, 40, 60 e 80 MW. La probabilità di trovare  $k$  unità guaste é data da:

$$\Pr(k \text{ unita' guaste}) = \binom{4}{k} Q^k (1 - Q)^{4-k}$$

quindi



$$\begin{aligned}
\Pr(\mathbf{x} = 0 \text{ MW}) &= \binom{4}{4} Q^4(1-Q)^0 = 1.0 \cdot 10^{-12} \\
\Pr(\mathbf{x} = 20 \text{ MW}) &= \binom{4}{3} Q^3(1-Q)^1 = 3.996 \cdot 10^{-9} \\
\Pr(\mathbf{x} = 40 \text{ MW}) &= \binom{4}{2} Q^2(1-Q)^2 = 5.988 \cdot 10^{-6} \\
\Pr(\mathbf{x} = 60 \text{ MW}) &= \binom{4}{1} Q^1(1-Q)^3 = 3.988 \cdot 10^{-3} \\
\Pr(\mathbf{x} = 80 \text{ MW}) &= \binom{4}{0} Q^0(1-Q)^4 = 9.96 \cdot 10^{-1}
\end{aligned}$$

Inoltre é:

$$\Pr(\mathbf{x} \leq 60 \text{ MW}) = 1.0 \cdot 10^{-12} + 3.996 \cdot 10^{-9} + 5.988 \cdot 10^{-6} + 3.988 \cdot 10^{-3} = 3.994 \cdot 10^{-3}$$

$$\Pr(\mathbf{x} \leq 80 \text{ MW}) = 1$$

$$\Pr(\mathbf{x} < 80 \text{ MW}) = 3.994 \cdot 10^{-3}$$

### 3.3 Funzioni di distribuzione

La funzione di distribuzione di una VA permette di specificare in modo completo quelle che sono le proprietà probabilistiche della VA considerata. Tale funzione é definita come:

$$F_x(a) = \Pr(\mathbf{x} \leq a) \quad (3.3)$$

Cioé la funzione di distribuzione della variabile aleatoria  $\mathbf{x}$  é quella funzione che specifica la probabilità di ogni evento del tipo  $\mathbf{x} \in [-\infty, a]$ , cioé che il valore della variabile aleatoria sia minore o uguale ad  $a$  in funzione dell'estremo  $a$  stesso. Si noti che il pedice  $x$  può essere omesso ove sia chiaro dal contesto che ci si riferisce alla distribuzione della VA  $\mathbf{x}$ .

Quando lo spazio degli eventi  $S$  (dell'esperimento) può essere messo in corrispondenza biunivoca con un sottoinsieme di dimensione finita degli interi, allora la funzione di distribuzione ha un andamento a gradinata. Le figure 3.1 e 3.2 mostrano le funzioni di distribuzione delle VA ottenute dal lancio di un dado utilizzando le due seguenti relazioni:

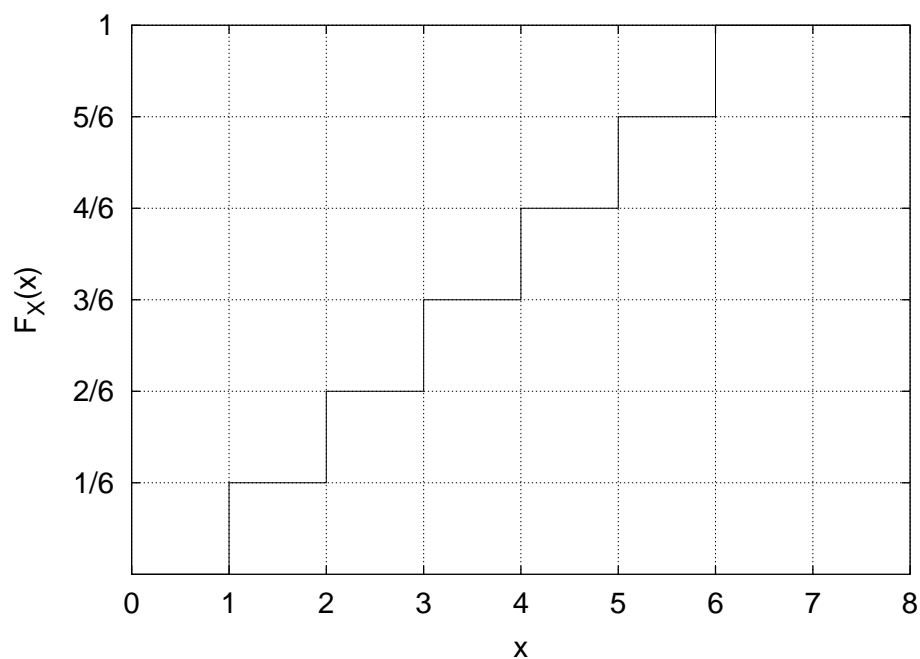
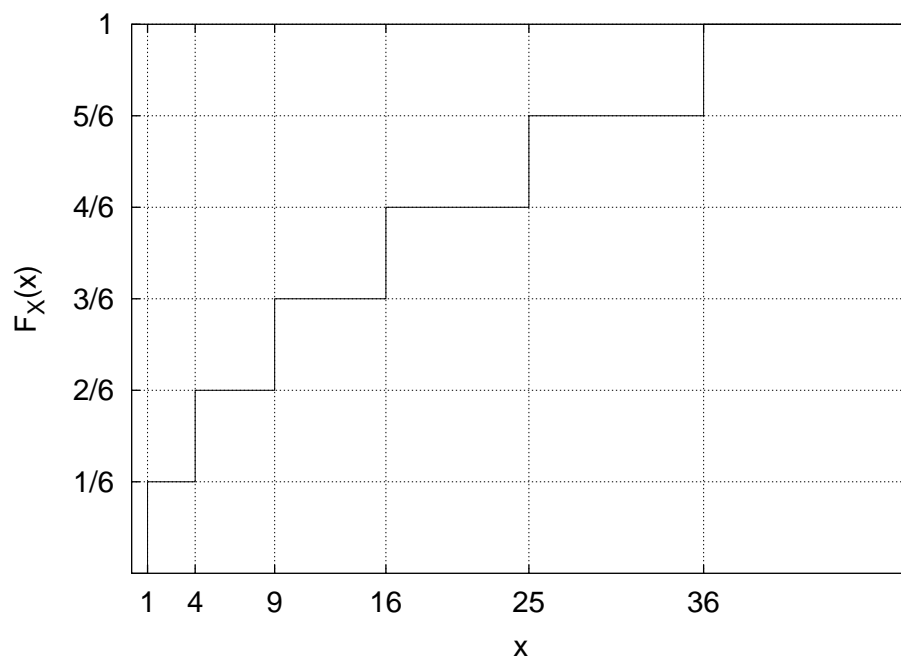
$$x(f_n) = n$$

$$x(f_n) = n^2$$

**Esempio 3.3.1** La centrale di generazione costituita da 4 unità discussa nell'esempio precedente ha potenza disponibile la cui distribuzione di probabilità é rappresentata nella figura 3.3

Per molte VA di interesse pratico (velocità di un vento in una certa regione, potenza domandata ad un nodo della rete elettrica, tempo intercorrente fra due guasti successivi di un dispositivo, ecc.) non presenta un andamento a scalinata (discontinuo) ma, piuttosto, un andamento continuo (quindi  $\lim_{x \rightarrow x_0^-} = \lim_{x \rightarrow x_0^+} = F(x_0)$ ) come quello rappresentato, ad esempio, in figura 3.4. Le conseguenze della continuità di  $F$  (o, viceversa, della presenza di discontinuità di tipo salto in  $F$ ) saranno chiarite in seguito.

Per la funzione di distribuzione valgono le seguenti proprietà:

Figura 3.1: Distribuzione della VA  $x(f_n) = n$  ottenuta dall'esperimento di lancio del dado.Figura 3.2: Distribuzione della VA  $x(f_n) = n^2$  ottenuta dall'esperimento di lancio del dado.

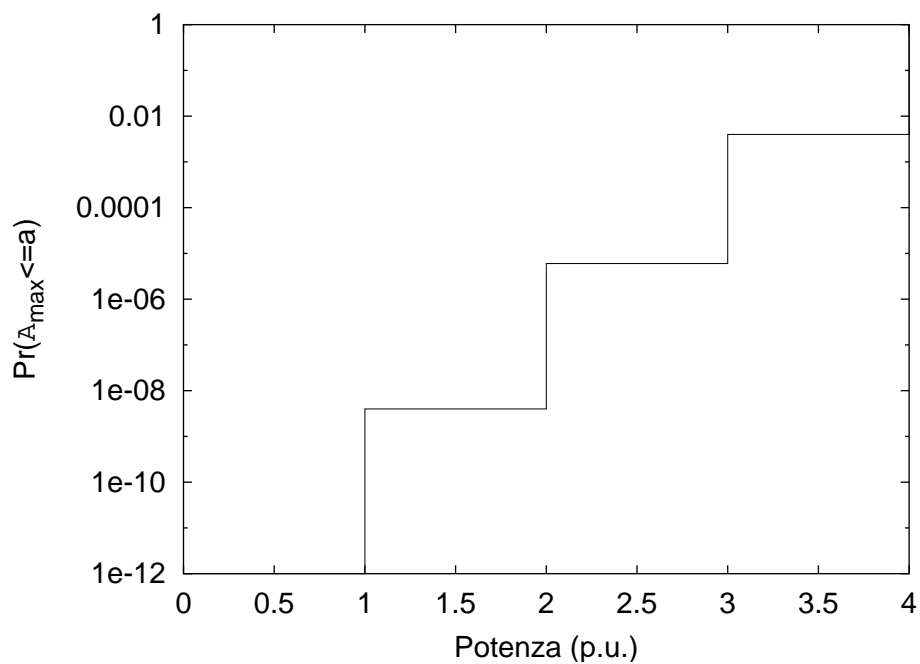


Figura 3.3: Distribuzione di probabilità della potenza massima generabile dall'impianto (potenza base  $A_{base} = 20$  MW) discusso negli esempi 3.2.2 e 3.3.1.

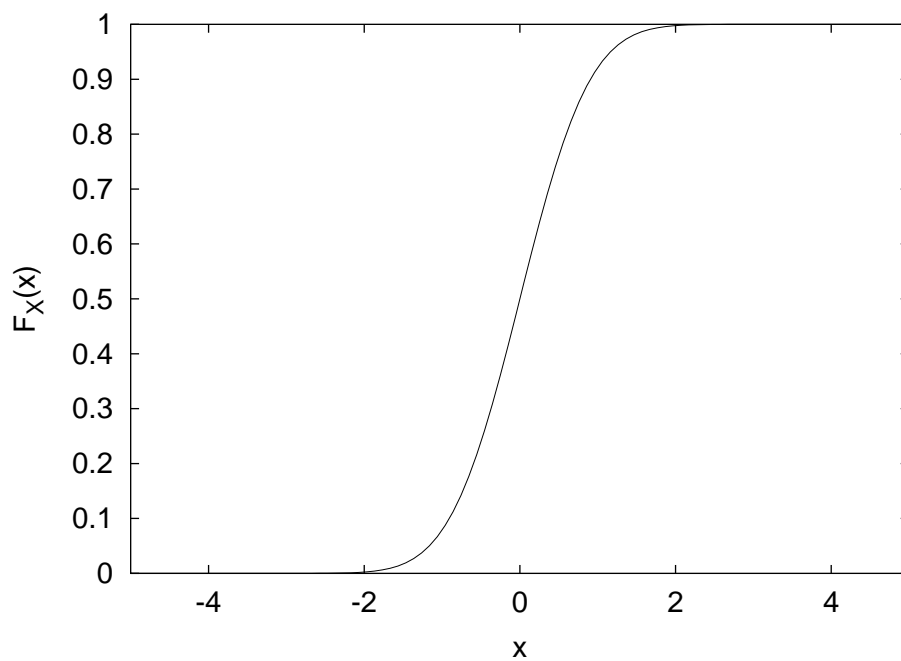
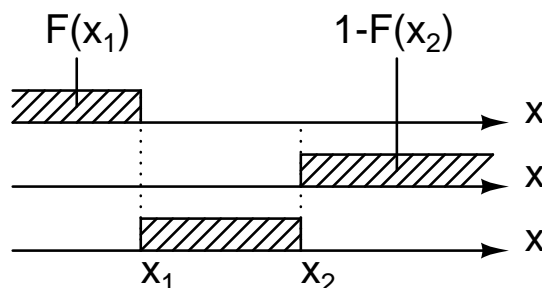


Figura 3.4: Distribuzione di una VA  $x$  continua.



$$\Pr(x_1 < \mathbf{x} < x_2) = 1 - (F(x_1) + (1 - F(x_2))) = F(x_2) - F(x_1)$$

Figura 3.5: Procedimento logico per il calcolo di  $\Pr(\mathbf{x} \in [x_1, x_2])$

- $0 \leq F \leq 1$  ( $F$  esprime una probabilitá)
- $F(-\infty) = 0$
- $F(+\infty) = 1$
- $F$  é monotona crescente. Infatti, se  $x_1 < x_2$ , allora  $F(x_2) = \Pr(\mathbf{x} \leq x_1) + \Pr(x_1 < \mathbf{x} \leq x_2)$
- se  $F(x_0) = 0$  allora  $F(\mathbf{x} < x_0) = 0$  (per la monotonia di  $F$ )
- $1 - F(a) = \Pr(\mathbf{x} > a)$ , in quanto l'evento  $\{\mathbf{x} > a\}$  é il complementare di  $\{\mathbf{x} \leq a\}$
- $\Pr(x_1 < \mathbf{x} \leq x_2) = F(x_2) - F(x_1)$ , infatti l'evento  $\{x_1 < \mathbf{x} \leq x_2\}$  é il complementare dell'unione  $\{\mathbf{x} \leq x_1\} \cup \{\mathbf{x} > x_2\}$ , come mostrato in figura 3.5

**Esempio 3.3.2** Una distribuzione uniforme é caratterizzata dal fatto che la probabilitá di osservare la VA  $\mathbf{u}$  in  $[0, u]$  dipende, esclusivamente da  $u$  ed é:

$$\Pr(\mathbf{u} \leq u) = u \rightarrow F(u) = u$$

## 3.4 Densitá di probabilitá

### 3.4.1 Classificazione delle VA ed eventi delle VA continue

Prima di introdurre il concetto di densitá di probabilitá é bene, a questo punto, distinguere vari classi di variabili aleatorie. Tali classi sono generate dalle proprietá della funzione di distribuzione. Si definiscono:

- continue quelle VA per cui  $F$  é continua,
- discrete quelle VA per cui  $F$  é costante a tratti con discontinuitá di tipo salto,
- miste quelle VA per cui  $F$  é generalmente continua, eccettuate alcune discontinuitá di tipo salto.

Le VA discrete e continue sono quelle piú comunemente usate in ingegneria. Le VA discrete nascono, come mostrato negli esempi precedenti, da processi di conta (ad esempio, il numero di fulmini caduti su una linea nel corso di un anno, il numero di centrali elettriche funzionanti, ecc.). Le VA continue hanno origine dalla misura di grandezze fisiche. Ad esempio, la lettura di un voltmetro, la tensione di scarica di un isolante, la richiesta di energia elettrica ad un nodo della rete AT, ecc. Sará mostrato in seguito che le VA discrete possono essere considerate casi particolari delle VA continue utilizzando la teoria matematica

delle distribuzioni. Le VA continue hanno, pertanto, un campo di applicazione molto esteso ed ha quindi la massima importanza studiarne le proprietà<sup>1</sup>.

Per una VA continua, non é, in generale, possibile assegnare all'evento  $\mathbf{x} = a$  una probabilità che non sia infinitesima, cioè diversa da 0. Infatti:

$$\Pr(\mathbf{x} = a) = \lim_{\delta \rightarrow 0} \Pr(\mathbf{x} \in [a, a + \delta]) = \lim_{\delta \rightarrow 0} F(a + \delta) - F(a) = 0$$

(l'ultima eguaglianza deriva dalla continuità di  $F$ , precedentemente postulata). Per una VA continua ha, pertanto, senso parlare di eventi come quegli insiemi numerici per cui non é nullo l'integrale di una funzione positiva ovunque sull'insieme. Sono eventi elementari gli intervalli:

- $\{x_1 \leq \mathbf{x} \leq x_2\}$
- $\{x_1 < \mathbf{x} \leq x_2\}$
- $\{x_1 \leq \mathbf{x} < x_2\}$
- $\{x_1 < \mathbf{x} < x_2\}$

e tutte le loro combinazioni (unioni e/o intersezioni) di tali eventi, fra cui gli intervalli infiniti:

- $\{\mathbf{x} \leq x\}$
- $\{\mathbf{x} < x\}$
- $\{\mathbf{x} \geq x\}$
- $\{\mathbf{x} > x\}$

Per quanto esposto in precedenza, la probabilità degli eventi

- $\{x_1 \leq \mathbf{x} \leq x_2\}$
- $\{x_1 < \mathbf{x} < x_2\}$

é identica. Infatti l'intervallo chiuso può essere considerato come l'intervallo aperto piú i due punti  $x_1$  ed  $x_2$ . Quindi, se é  $A$  l'intervallo aperto e con  $C$  l'intervallo chiuso e visto che non esiste intersezione fra  $x = x_1$ ,  $A$ ,  $x = x_2$ , allora:

$$\Pr(C) = \Pr(\mathbf{x} = x_1) + \Pr(A) + \Pr(\mathbf{x} = x_2) = \Pr(A)$$

In modo analogo é possibile dimostrare che tutte le distinzioni fatte in precedenza fra intervalli aperti, chiusi e semiaperti non generano differenze nella valutazione della probabilità degli eventi. Queste considerazioni sono vere quando la VA é continua. Nel caso di VA discrete o miste tali considerazioni non sono applicabili.

---

<sup>1</sup>E' chiaro che, per effetto dell'errore strumentale le letture di grandezze fisiche sono sempre arrotondate. Si potrebbe convenire quindi che le uniche VA aventi interesse pratico sono quelle discrete. Tuttavia, per quanto concerne la costruzione di modelli appropriati non può essere trascurato il fatto che tali grandezze sono continue.

### 3.4.2 Densità di probabilità

Se l'evento  $\Pr(\mathbf{x} = a)$  è sempre a probabilità nulla per una VA continua, al contrario, generalmente diverso da zero è, il rapporto incrementale:

$$\lim_{\delta \rightarrow 0} \frac{\Pr(a \leq \mathbf{x} \leq a + \delta)}{\delta}$$

Ove questo limite esista esso è noto come densità di probabilità e si indica come:

$$f_x(a) = \lim_{\delta \rightarrow 0} \frac{\Pr(a \leq \mathbf{x} \leq a + \delta)}{\delta} \quad (3.4)$$

La densità di probabilità della VA  $\mathbf{x}$  non definisce, pertanto, la probabilità di  $\mathbf{x} = a$  ma il limite a cui tende la probabilità di osservare la VA nell'intervallo  $[a, a + \delta]$ , espresso in rapporto a  $\delta$ , quando la larghezza dell'intervallo tende a 0. La funzione  $f_x(a)$  può pertanto assumere tutti i valori (positivi) fra 0 e infinito. Poiché:

$$\Pr(a \leq \mathbf{x} \leq a + \delta) = F_x(a + \delta) - F_x(a)$$

è evidente da (3.4) che sussiste

$$f_x(x) = \lim_{\delta \rightarrow 0} \frac{F_x(a + \delta) - F_x(a)}{\delta} = \frac{dF_x(x)}{dx} \quad (3.5)$$

e

$$F_x(x) = \int_{-\infty}^x f_x(s) ds \quad (3.6)$$

quindi:

$$\Pr(\mathbf{x} \in [x_1, x_2]) = F_x(x_2) - F_x(x_1) = \int_{x_1}^{x_2} f_x(s) ds$$

La probabilità di osservare la VA  $\mathbf{x}$  nell'intervallo considerato coincide, pertanto, con l'area sottesa da  $f_x$  fra i punti  $x_1$  ed  $x_2$ , come illustrato dalla figura 3.6

Si noti che, contrariamente alla distribuzione,  $F$ , che indica una probabilità, la densità  $f$  è dotata di unità di misura. Ad esempio, la densità di probabilità di scarica di un sistema isolante dovrebbe essere indicata in  $kV^{-1}$ . Nella pratica, l'unità di misura della densità di probabilità è raramente indicata.

### 3.4.3 VA discrete come caso particolare di VA continue

Abbiamo visto che per una VA continua l'evento  $\mathbf{x} = a$  ha probabilità nulla. Al contrario, eventi di questo tipo sono quelli che permettono, in generale, di specificare il comportamento di una VA discreta.

Nonostante queste considerevoli differenze, la funzione di distribuzione è definita in un modo che non dipende dalla natura della VA, discreta o continua. In un caso, la funzione di distribuzione è una funzione a gradinata, discontinua e non derivabile se non in senso generalizzato (la derivata non appartiene al dominio delle funzioni ma a quello delle distribuzioni), nell'altro caso essa sarà una funzione continua, derivabile nel senso delle funzioni.

Il concetto di derivata generalizzata permette di assegnare un risultato al calcolo della derivata per distribuzioni continue o affette da discontinuità di tipo salto. Per una VA discreta, con distribuzione "a gradinata", la densità di probabilità è caratterizzabile come una serie di impulsi di Dirac. Nel caso continuo si otterrà una densità di probabilità che appartiene al dominio delle funzioni (non necessariamente continua).

E' quindi possibile pensare ogni VA discreta come una VA continua caratterizzata da una densità di probabilità appartenente alla classe delle distribuzioni.

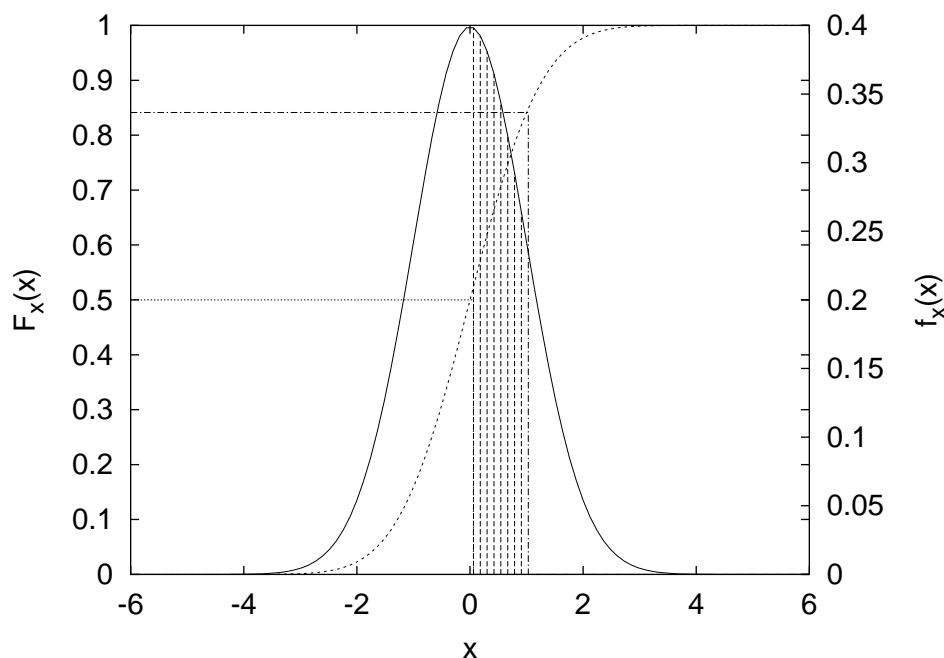


Figura 3.6: La probabilità di osservare  $\mathbf{x}$  nell'intervallo di estremi 0 e 1 è pari all'area della regione tratteggiata o alla differenza  $F_x(1) - F_x(0)$ .

#### 3.4.4 VA miste

### 3.5 Percentili

Si definisce il percentile a probabilità  $P$ ,  $x_P$ , della VA  $\mathbf{x}$  il numero per cui é:

$$\Pr(\mathbf{x} \leq x_P) = F_x(x_P) = P \quad (3.7)$$

Si noti che, conformement a ciò, il fatto che una VA  $\mathbf{x}$  sia distribuita secondo la distribuzione  $F$  può essere esplicitato anche come:

$$\Pr(\mathbf{x} \leq F^{-1}(P)) = P \quad (3.8)$$

**Esempio 3.5.1 (Il metodo di Monte Carlo)** Il metodo di Monte Carlo é largamente utilizzato nella generazione di numeri aleatori aventi distribuzione prefissata  $F$  mediante il calcolatore. Il metodo sfrutta la possibilità del calcolatore di generare numeri (pseudo)aleatori uniformemente distribuiti in  $[0, 1]$ . Normalmente, la routine che si occupa di queste operazioni é indicata come `rand()` (ad esempio, in C, in Matlab, ecc.).

Sia quindi  $\mathbf{u}$  la VA uniformemente distribuita in  $[0, 1]$ . Per generare una VA avente distribuzione  $F$  é sufficiente invertire  $F$  e calcolare  $\mathbf{x} = F^{-1}(\mathbf{u})$ . Poiché  $F$  é monotona, é facile provare che:

$$\Pr(\mathbf{x} \leq F^{-1}(u)) = \Pr(\mathbf{u} \leq u) = u$$

Il risultato ottenuto, cioè  $\Pr(\mathbf{x} \leq F^{-1}(u)) = u$  é formalmente identico a quanto formulato nell'equazione (3.8). Quindi, la VA  $\mathbf{x}$  ha distribuzione  $F$ , esattamente come si desiderava.

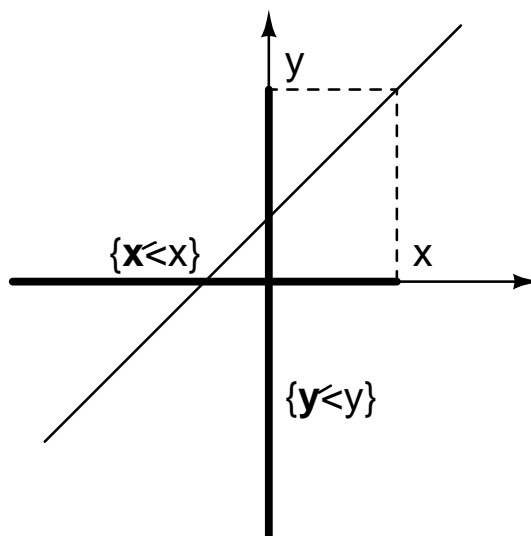


Figura 3.7: Trasformazione lineare di due variabili aleatorie. E' facile provare (utilizzando ad esempio l'approccio delle frequenze) che  $\Pr(\mathbf{x} \in ]-\infty, x]) = \Pr(\mathbf{y} \in ]-\infty, y])$  quando  $y = ax + b$ .

### 3.6 Trasformazioni lineari

Nelle applicazioni pratiche ed in statistica ha notevole interesse calcolare la distribuzione di VA scalate e traslate di opportuni coefficienti reali:

$$\mathbf{y} = a\mathbf{x} + b$$

Utilizzando il fatto che (si veda la figura 3.7):

$$\Pr(\mathbf{y} \leq y) = \Pr(a\mathbf{x} + b < y) = \Pr\left(\mathbf{x} \leq \frac{y-b}{a}\right)$$

si arriva a determinare la distribuzione di  $\mathbf{y}$  come:

$$F_y(y) = F_x\left(\frac{y-b}{a}\right) \quad (3.9)$$

e, derivando, si ottiene la densità:

$$f_y(y) = \frac{1}{|a|} f_x\left(\frac{y-b}{a}\right) \quad (3.10)$$

**Esempio 3.6.1** Un generatore reale con  $V_{OC} = -0.5$  V e conduttanza  $G = 10$  S é collegato ad una sorgente ideale di tensione. La sorgente ideale di tensione ha tensione fra i suoi morsetti distribuita uniformemente in  $[0, 1]$ . Il circuito é rappresentato in figura 3.8. Calcolare la distribuzione della corrente.

Applicando (3.9) e (3.10) si ottiene una densità costante, pari a 0.1, nell'intervallo  $[5, 15]$  A. La corrispondente distribuzione é nulla per valori di corrente inferiori a 5 A, cresce linearmente fino a 15 A, dove raggiunge valore unitario. Da 15 A in avanti rimane costantemente uguale ad 1.



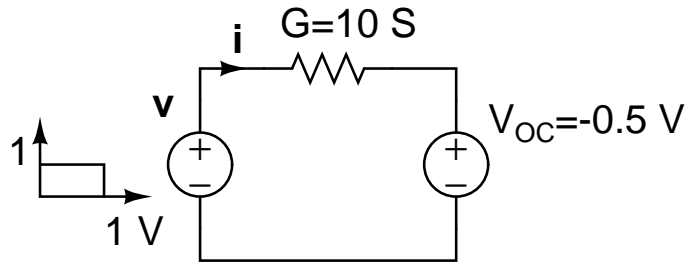


Figura 3.8: Circuito elettrico per il quale si intende calcolare la distribuzione della corrente.

### 3.7 Funzioni di uso comune

Per risolvere problemi inerenti alla teoria delle probabilità ed in statistica sono state introdotte famiglie parametriche di distribuzioni e funzioni di densità di probabilità. Tali famiglie sono parametriche in quanto dipendono dal valore generico della variabile aleatoria (che in seguito sarà indicato con  $t$ ) e da uno o più parametri che determinano la forma e la localizzazione della densità di probabilità sull'asse reale.

Si riportano, a titolo di esempio, alcune famiglie parametriche comunemente usati nei problemi di affidabilità e statistica (la funzione  $H(t)$  è la funzione di Heavyside, uguale a 0 per  $t < 0$ , a 1 per  $t \geq 0$ ).

- Uniforme:  $f(t) = \frac{1}{b-a}(H(a) - H(b))$ , ( $b > a$ )
- Esponenziale:  $f(t) = \lambda \exp(-\lambda t)H(t) \leftrightarrow F(t) = (1 - \exp(-\lambda t))H(t)$
- Raileigh:  $f(t) = \frac{t}{\sigma^2} \exp(-\frac{1}{2} \frac{t^2}{\sigma^2})H(t) \leftrightarrow F(t) = (1 - \exp(-\frac{t^2}{2\sigma^2}))H(t)$
- Normale:  $f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2} (\frac{t-\eta}{\sigma})^2)$
- LogNormale:  $f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \exp(-(\log(t) - \eta)^2 / (2\sigma^2))$
- Weibull:  $f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right]H(t) \leftrightarrow F(t) = \left(1 - \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right]\right)H(t)$
- Chiquadro:  $f(t) = \frac{1}{2^{\frac{n}{2}} \sigma^n \Gamma(n/2)} t^{\frac{n}{2}-1} \exp(-\frac{1}{2}(t/\sigma)^2)H(t)$
- Gamma:  $f(t) = \frac{\lambda}{(r-1)!} (\lambda t)^{r-1} \exp(-\lambda t)H(t)$
- Cauchy:  $f(t) = \frac{1}{\pi\beta \left(1 + \left(\frac{x-\alpha}{\beta}\right)^2\right)}$
- Student  $f(t) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$
- Maxwell  $f(t) = \frac{4}{\sqrt{\pi}} \frac{1}{\alpha^3} t^2 \exp(-\left(\frac{x}{\alpha}\right)^2)H(t)$
- Beta:  $f(t) = \frac{(n+1)!}{k!(n-k)!} t^k (1-t)^{n-k}$  con  $0 < t < 1$

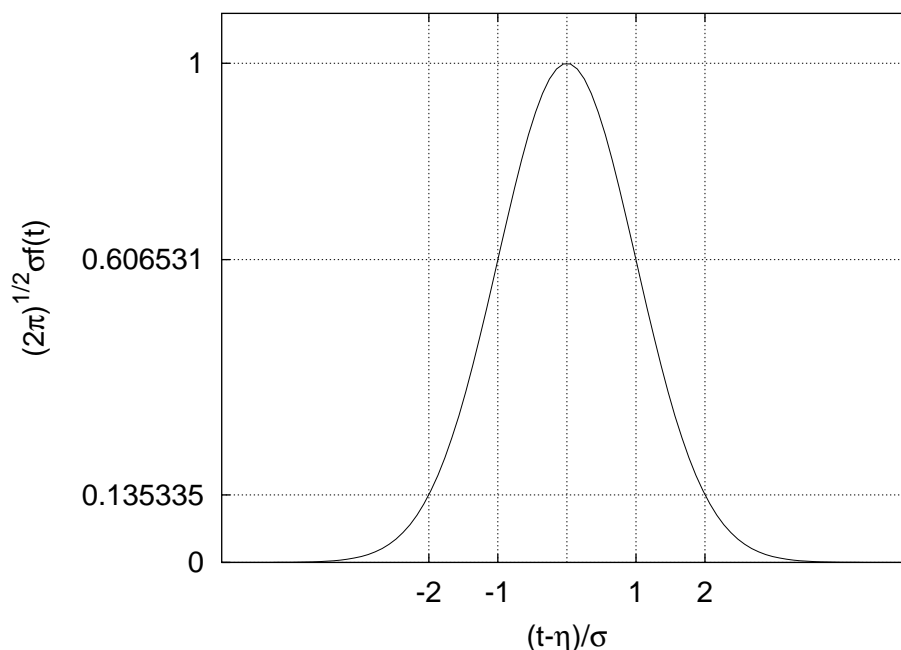


Figura 3.9: Andamento della densità di probabilità gaussiana.

### 3.7.1 Distribuzione normale (gaussiana)

Le variabili aleatorie di tipo gaussiano sono largamente impiegate per modellare fenomeni di tipo casuale che presentano una tendenza a raggrupparsi in modo simmetrico. La giustificazione teorica per l'uso della distribuzione normale in tante applicazioni risiede nel teorema del limite centrale, che sarà presentato successivamente.

Le variabili aleatorie gaussiane sono anche dette normali e sono descritte dalla densità di probabilità:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t-\eta}{\sigma}\right)^2\right)$$

Tale funzione è rappresentata in figura 3.9. Si noti che il parametro  $\eta$  definisce la posizione sulla retta dell'asse di simmetria della "campana". Il parametro  $\sigma$  ne controlla, invece, l'ampiezza. Ciò è posto in maggior rilievo dalla figura 3.10. Nella figura, si indica con  $N(\eta, \sigma)$  una distribuzione normale di parametri  $\eta$  e  $\sigma$ . Tale convenzione sarà utilizzata estesamente nel seguito. Sinteticamente, per affermare che una VA  $t$  segue la distribuzione  $N(\eta, \sigma)$  si indicherà  $t \sim N(\eta, \sigma)$ .

Si mostrerà ora una interessante proprietà della funzione gaussiana. Sia  $t$  una VA  $N(\eta, \sigma)$ . Si calcoli la distribuzione di:

$$u = \frac{t - \eta}{\sigma}$$

La trasformazione lineare in oggetto è riscrivibile come:

$$u = \frac{1}{\sigma}t - \frac{\eta}{\sigma}$$

Per applicare l'equazione (3.10) si deve porre  $a = \frac{1}{\sigma}$  e  $b = -\frac{\eta}{\sigma}$ . Pertanto, tale distribuzione può essere calcolata dalla distribuzione di partenza calcolata nel punto  $\sigma y + \eta$ . Utilizzando la densità di probabilità si ottiene:

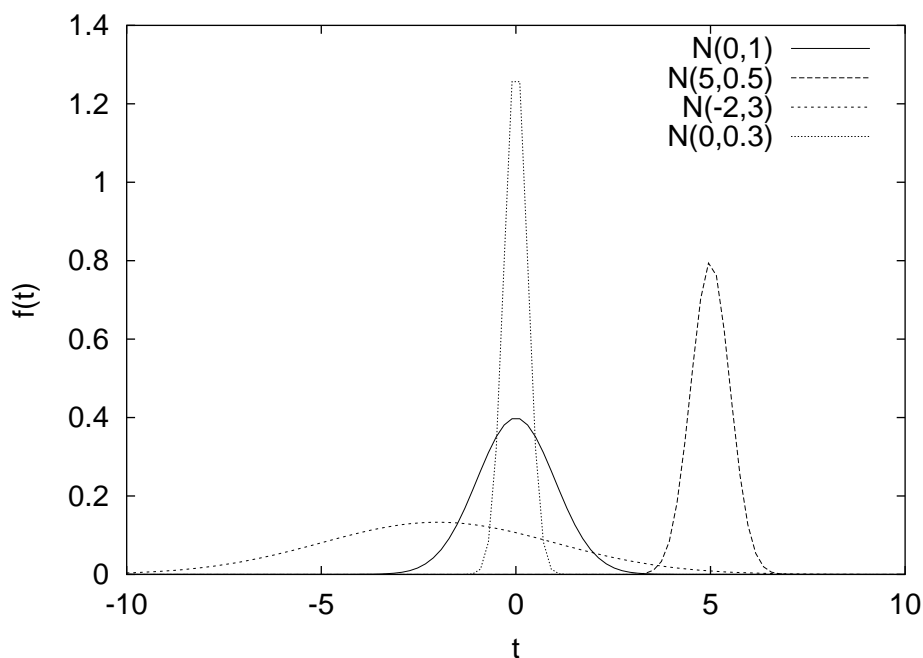


Figura 3.10: Andamento della densità di probabilità gaussiana per diversi valori dei parametri  $\eta$  e  $\sigma$

$$f_u(u) = \frac{1}{\sigma} f_x(\sigma u + \eta) \frac{1}{\sigma} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{(\sigma u + \eta) - \eta}{\sigma}\right)^2\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u^2\right)$$

Pertanto, la VA  $u$  ha densità di probabilità appartenente alla stessa famiglia di  $t$ . I parametri di  $u$ , invece di essere  $\eta$  e  $\sigma$ , sono 0 e 1. Si dice che  $u$  è una variabile normale normalizzata.

La VA normale normalizzata ha importanza fondamentale in quanto la densità e la distribuzione di ogni variabile aleatoria normale  $t \sim N(\eta, \sigma)$  può essere ottenuta attraverso quella della VA normale normalizzata:

$$f_t(t) = \frac{1}{\sigma} f_u\left(\frac{t - \eta}{\sigma}\right) \quad (3.11)$$

$$F_t(t) = F_u\left(\frac{t - \eta}{\sigma}\right) \quad (3.12)$$

Si noti che, se  $t \sim N(\eta, \sigma)$ , allora:

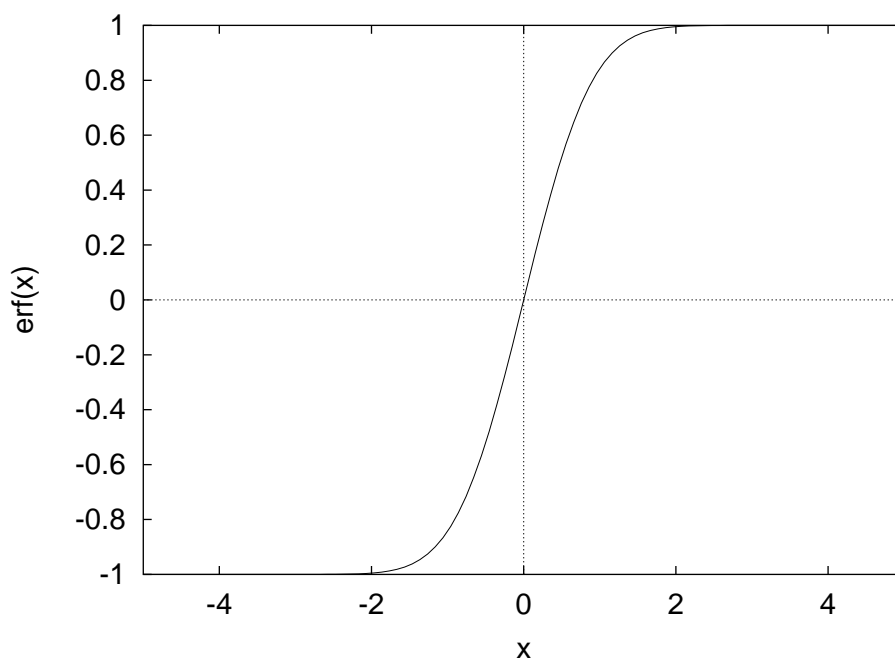
$$(t + a) \sim N(\eta + a, \sigma) \quad (3.13)$$

$$(at) \sim N(a\eta, a\sigma) \quad (3.14)$$

La densità di probabilità di  $u$ , variabile aleatoria normale normalizzata, è:

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

La distribuzione, non calcolabile analiticamente, viene indicata come:

Figura 3.11: Andamento della *error function*.

$$F_u(u) = \Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u \exp\left(-\frac{1}{2}s^2\right) ds$$

Per la simmetria della funzione gaussiana é:

$$\Phi(-u) = 1 - \Phi(u)$$

La funzione  $\Phi(u)$  é tabulata in tutti i libri di teoria delle probabilitá e statistica. Due forme possibili di tavola possono essere fornite, una che mostra il valore di  $\Phi(u)$  per determinati valori di  $u$  (calcolo diretto) l'altra che mostra  $u$  corrispondenti a determinati livelli di probabilitá (calcolo inverso). In Appendice 2 sono riportate tabelle di entrambi i tipi che permettono di risolvere la maggioranza dei problemi pratici. A livello computazionale é disponibile, in molti pacchetti di calcolo, una implementazione numerica della *error function*, cioé la funzione definita come:

$$\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-s^2} ds$$

La funzione  $\operatorname{erf}()$ , rappresentata in figura 3.11, permette di calcolare  $\Phi(u)$  servendosi della seguente relazione:

$$\Phi(u) = \frac{1 + \operatorname{erf}(u/\sqrt{2})}{2} \quad (3.15)$$

La relazione (3.15) puó essere ottenuta a partire da semplici passaggi analitici. Infatti, per  $u > 0$  é:

$$\frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{1}{2}t^2\right) dt = \frac{\sqrt{2}}{\sqrt{2\pi}} \int_0^x \exp\left(-\left(\frac{t}{\sqrt{2}}\right)^2\right) d\frac{t}{\sqrt{2}} =$$

$$= \frac{1}{\sqrt{\pi}} \int_0^{\frac{x}{\sqrt{2}}} \exp(-t^2) dt = \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)$$

Sfruttando il fatto che  $\Pr(\mathbf{u} \leq 0) = \Phi(0) = 1/2$  si perviene immediatamente a (3.15). Tale relazione vale anche per valori di  $u$  negativi.

Dal punto di vista pratico, se  $\mathbf{t} \sim N(\eta, \sigma)$ , allora:

$$\Pr(\mathbf{t} \in [a, b]) = \Phi\left(\frac{b-\eta}{\sigma}\right) - \Phi\left(\frac{a-\eta}{\sigma}\right) \quad (3.16)$$

Viceversa, se  $\Pr(\mathbf{u} \in [c, d]) = P$ , allora

$$\Pr\left(\mathbf{t} \in [\sigma c + \eta, \sigma d + \eta]\right) = P \quad (3.17)$$

### Stima intervallare

Il problema della stima intervallare può essere

- diretto: consiste nel calcolo della probabilità  $\Pr(\mathbf{x} \in [a, b])$  con  $a$  e  $b$  dati;
- indiretto, cioè data  $P$ , calcolare  $a$  e  $b$  tali che  $\Pr(\mathbf{x} \in [a, b]) = P$ ;

Si noti che il problema indiretto non ha soluzione unica, in generale (salvo i casi in cui  $a$  oppure  $b$  sono dati). Si richiede pertanto, per determinare univocamente gli estremi dell'intervallo, una condizione addizionale. Generalmente, tale condizione è che  $a = -b$ . Il significato di questo vincolo sarà chiarito nel corso della trattazione della stima statistica per intervalli.

Il problema diretto è di semplice soluzione. Si supponga, inizialmente, che la VA sia normale normalizzata. In questo caso è sufficiente utilizzare le tabelle riportate in Appendice 2. Supponiamo, ad esempio, che si intenda calcolare la probabilità con cui  $u$  appartiene all'intervallo  $[1, 2]$ . Dalle tabelle 3.1 e 3.2 osserviamo che  $u = 2$  ha una probabilità (cumulata)  $\Phi(2) = 0.977250$ . Similmente 1 ha probabilità  $\Phi(1) = 0.841345$ . Quindi la probabilità cercata è  $0.977250 - 0.841345 = 0.13591$ . Si supponga ora che l'intervallo sia  $[-1, 2]$ . In questo caso, occorre osservare che, per la simmetria della densità di probabilità gaussiana,

$$F(-a) = 1 - F(a) \quad (3.18)$$

La probabilità cercata è pertanto,  $0.977250 - (1 - 0.841345) = 0.81859$ . Si noti che quando l'intervallo è del tipo  $[-a, a]$ , allora la probabilità cercata è:

$$F(a) - F(-a) = F(a) - (1 - F(a)) = 2F(a) - 1 \quad (3.19)$$

Dunque, per l'intervallo  $[-1, 1]$ , ad esempio, la probabilità cercata vale:  $2 \cdot 0.841345 - 1 = 0.68269$ .

Quando la VA considerata non è normalizzata, allora è necessario utilizzare (3.9). Ciò consiste, in pratica, nel normalizzare gli estremi dell'intervallo ed utilizzare il procedimento descritto per la VA normale normalizzata. Se  $[a, b]$  è l'intervallo richiesto,  $\eta$  e  $\sigma$  sono la media e la deviazione standard della VA, allora gli estremi dell'intervallo diventano:

$$\left[\frac{a-\eta}{\sigma}, \frac{b-\eta}{\sigma}\right]$$

e la probabilità cercata è:

$$\Pr(\mathbf{x} \in [a, b]) = \Phi\left(\frac{b-\eta}{\sigma}\right) - \Phi\left(\frac{a-\eta}{\sigma}\right) \quad (3.20)$$

Ad esempio, se si vuole calcolare la probabilità di osservare la VA nell'intervallo  $[\eta - \sigma, \eta + \sigma]$ , tale probabilità coincide con quella di osservare una VA normale normalizzata nell'intervallo  $[-1, 1]$  e vale, come nel caso precedente, 0.68269.

Il problema inverso, sotto il vincolo  $a = -b$  si risolve in modo abbastanza agevole utilizzando (3.19). Sia infatti  $P = 0.68$ . Allora è necessario cercare  $a$  in modo che  $2F(a) - 1 = 0.68$ , cioè  $F(a) = (1 + 0.68)/2 = 0.84$ . Dalle tabelle 3.3 e 3.4 si ottiene immediatamente il percentile desiderato, 0.994458, approssimativamente uguale ad 1.

**Esempio 3.7.1 (Controllo statistico di qualità)** Si consideri una produzione industriale di resistori aventi un valore nominale di  $R_0 = 1000 \Omega$ . Nessuna macchina è in grado di produrre resistori tutti uguali ed aventi il valore nominale specificato. Ciò può essere dovuto a disomogeneità del materiale, fluttuazioni della tensione di alimentazione della macchina, variazioni di temperatura, errore umano, ecc. Normalmente si considera che una linea di produzione industriale produca dispositivi il cui parametro caratteristico (la resistenza, nell'esempio proposto) è dato da una variabile aleatoria (errore di produzione)  $\Delta R$  che si somma al valore nominale. Il parametro stesso è, pertanto, aleatorio:

$$R = R_0 + \Delta R$$

Normalmente, l'errore di produzione,  $\Delta R$ , si considera che segua una distribuzione normale di parametri  $\eta$  e  $\sigma$ . In una produzione ben fatta  $\eta = 0$  e  $\sigma$  ha il valore minimo possibile, compatibile con la tecnologia di produzione, caratteristiche della macchina, ecc.. La richiesta  $\eta = 0$  equivale ad affermare che la produzione è priva di errori sistematici. La richiesta che  $\sigma$  sia il minimo compatibile con la tecnologia produttiva significa che la variabilità di produzione è la minore possibile. In questa analisi si considererà sempre verificata questa seconda richiesta, e si valuteranno gli effetti di una variazione di  $\eta$ , cioè dell'errore sistematico di produzione. Quando  $\eta$  è diverso da zero (positivo o negativo) la linea si dice fuori controllo.

La tolleranza di produzione sia  $\pm 10 \Omega$ ,  $\sigma = 2.5 \Omega$ . Pertanto, resistori con resistenze minori di  $990 \Omega$  e superiori a  $1100 \Omega$  sono scartate. Si calcoli quale frazione di dispositivi sono scartati (difettosità della produzione) nel caso che  $\eta = 0 \Omega$  e nel caso  $\eta = 1 \Omega$ .

Per risolvere il problema dato è sufficiente valutare, nel primo caso, la probabilità che una VA  $N(1000, 2.5)$  appartenga all'intervallo  $[990, 1010]$ . Il complementare di tale probabilità rappresenta il risultato cercato. Pertanto:

$$P_A = 1 - \left( \left( \frac{1010 - 1000}{2.5} \right) - \left( \frac{990 - 1000}{2.5} \right) \right) \approx 6 \cdot 10^{-5}$$

Quindi, su 100000 pezzi prodotti ne saranno scartati 6. Nel secondo caso, visto che la resistenza dei resistori è una VA con distribuzione  $N(1001, 2.5)$ , applicando lo stesso procedimento, si ottiene:

$$P_B = 1 - \left( \left( \frac{1010 - 1001}{2.5} \right) - \left( \frac{990 - 1001}{2.5} \right) \right) \approx 2 \cdot 10^{-4}$$

e, ogni 10000 pezzi prodotti, 2 saranno difettosi. La difettosità, nel secondo caso, è il 333% di quella del primo caso.

Si convenga ora di volere stabilire se è il caso, o meno, di interrompere la produzione per eliminare eventuali errori sistematici mediante controlli sul valore delle resistenze dei resistori prodotti. Per fare ciò si può procedere in due modi:

- (a) se un resistore prodotto esce dall'intervallo  $R_0 \pm 3\sigma$  si interrompe la produzione e si verifica il comportamento della linea.
- (b) se un lotto di  $N$  resistori ha un valore medio di resistenza che esce dall'intervallo  $R_0 \pm 3\sigma/\sqrt{N}$  si interrompe la produzione e si verifica il comportamento della linea.

Si consideri il procedimento (a). Se  $\eta = 0$  (la linea non deve essere controllata) per una frazione

$$P = 1 - \left( \left( \frac{(1000 + 7.5) - 1000}{2.5} \right) - \left( \frac{(1000 - 7.5) - 1000}{2.5} \right) \right) \approx 2.7 \cdot 10^{-4}$$

dei resistori prodotti il sistema di controllo segnalerá che é necessario controllare la linea quando, al contrario, tutto funziona perfettamente. Se  $\eta = 1$ , allora in una frazione

$$P = 1 - \left( \left( \frac{(1001 + 7.5) - 1000}{2.5} \right) - \left( \frac{(1001 - 7.5) - 1000}{2.5} \right) \right) \approx 5 \cdot 10^{-4}$$

dei resistori prodotti il sistema di controllo segnalerá che, come in effetti é, é necessario controllare la linea. In media, però,  $10000/5=2000$  pezzi saranno prodotti con la linea fuori controllo prima che il sistema segnali l'anomalia nella produzione. Pertanto  $P_B 2000 = 2 \cdot 10^{-4} 2000 = 0.4$  pezzi saranno prodotti in modo non conforme (fuori dalla tolleranza).

Nel caso (b), é possibile dimostrare che se i resistori hanno resistenza che segue una legge  $N(R_0 + \eta, \sigma)$ , allora il valore medio su  $N$  resistori é una variabile aleatoria che segue una legge  $N(R_0 + \eta, \sigma_N)$  dove  $\sigma_N = \sigma/\sqrt{N}$ . Si convenga, pertanto di segnalare che é necessario controllare la linea se la media su  $N$  resistori esce dall'intervallo  $R_0 \pm 3\sigma/\sqrt{N}$ . Nel caso  $\eta = 0$  ed  $N = 100$  é immediato provare che si avrá una frazione di falsi allarmi pari a:

$$P = 1 - \left( \left( \frac{(1000 + 7.5/\sqrt{100}) - 1000}{2.5/\sqrt{100}} \right) - \left( \frac{(1000 - 7.5/\sqrt{100}) - 1000}{2.5/\sqrt{100}} \right) \right) \approx 2.7 \cdot 10^{-4}$$

identica al caso precedente. Tuttavia, se  $\eta = 1$ , allora per una frazione

$$P = 1 - \left( \left( \frac{(1001 + 7.5/\sqrt{100}) - 1000}{2.5/\sqrt{100}} \right) - \left( \frac{(1001 - 7.5/\sqrt{100}) - 1000}{2.5/\sqrt{100}} \right) \right) \approx 0.84$$

dei lotti di 100 dispositivi prodotti si avrá una segnalazione che la linea deve essere controllata. Mediamente, si avrá un segnale di anomalia ogni  $100/84 = 1.2$  lotti, per un totale di 120 pezzi prodotti con la linea fuori controllo. Il rischio di produrre pezzi fuori dai limiti di tolleranza prima che si abbia un segnale di allarme é, in questo caso  $120 \cdot 2 \cdot 10^{-4} = 0.02$  pezzi. Il rischio si é pertanto ridotto, rispetto al caso precedente (0.4) al 5% con un notevole miglioramento della difettosità della catena di produzione.

### 3.7.2 Distribuzione lognormale

La distribuzione lognormale appartiene alla famiglia delle distribuzioni ai valori estremi. Le distribuzioni appartenenti a questa famiglia sono caratterizzate da una forte variabilità dei valori assumibili dalla variabile aleatoria. In genere tali distribuzioni presentano lunghe "code" in corrispondenza dei percentili superiori. In particolare, la lognormale é stata utilizzata per modellare i tempi di riparazione di dispositivi elettrici e la velocità dei venti all'interno degli uragani.

L'espressione analitica della densità di probabilità lognormale é:

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \exp(-(\log(t) - \eta)^2 / (2\sigma^2)) \quad (3.21)$$

La figura 3.12 mostra il grafico della distribuzione lognormale ottenuta per  $\eta = 0$  e  $\sigma = 1$ . La distribuzione lognormale é completamente caratterizzata dai parametri  $\eta$  e  $\sigma$ .

La distribuzione lognormale puó essere ricavata tramite la distribuzione di una VA normale normalizzata. Per provare ciò é sufficiente calcolare l'integrale fra 0 e  $x$  della densità di probabilità:

$$F(t) = \int_0^t \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log x - \eta}{\sigma}\right)^2\right) dx =$$

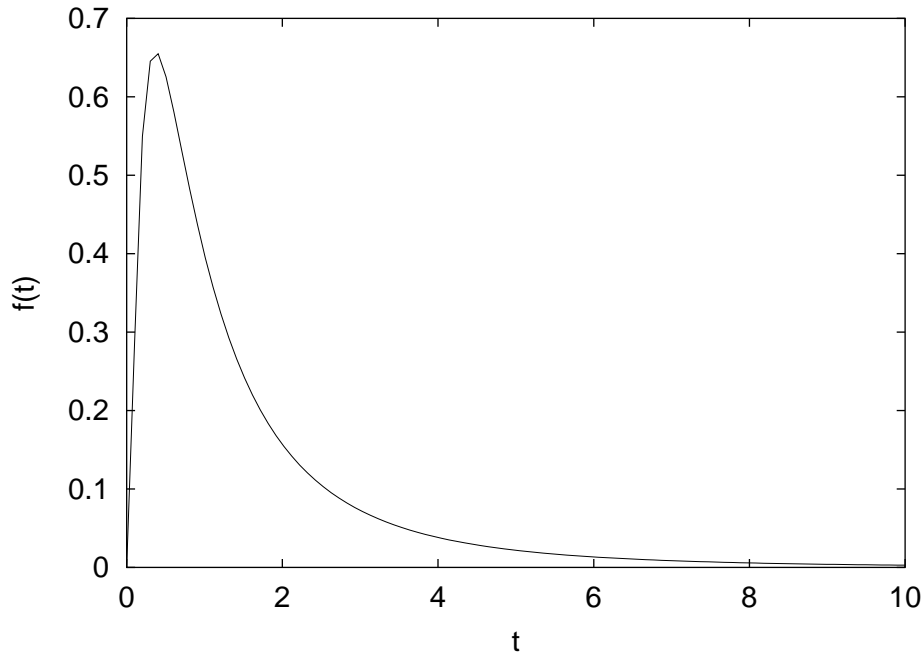


Figura 3.12: Grafico della distribuzione lognormale.

$$\begin{aligned}
 &= \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}\left(\frac{\log x - \eta}{\sigma}\right)^2\right) \frac{dx}{x\sigma} = \\
 &= \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}\left(\frac{\log x - \eta}{\sigma}\right)^2\right) d\frac{\log x}{\sigma} = \\
 &= \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{1}{2}\left(\frac{\log x - \eta}{\sigma}\right)^2\right) d\frac{\log x - \eta}{\sigma} = \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\log t - \eta}{\sigma}} \exp\left(-\frac{1}{2}y^2\right) dy
 \end{aligned} \tag{3.22}$$

Quindi é possibile scrivere:

$$F(t) = \Phi\left(\frac{\log t - \eta}{\sigma}\right) = \Phi\left(\frac{\lambda \log t}{\sigma}\right) \tag{3.23}$$

dove l'ultima equazione é stata ottenuta ponendo:

$$\eta = \log \lambda \tag{3.24}$$

E' inoltre possibile dimostrare che se  $\mathbf{x} = \log \mathbf{t}$  con  $\mathbf{t}$  distribuita in modo lognormale, allora  $\mathbf{x}$  é distribuita in modo normale ed ha parametri  $\eta$  e  $\sigma$ . Infatti, per la monotonia della funzione logaritmo (si veda la figura 3.13) é:

$$\Pr(\mathbf{t} \leq a) = \Pr(\mathbf{x} \leq \log a) \rightarrow F_t(a) = F_x(\log a) \tag{3.25}$$

Poiché la distribuzione della VA  $\mathbf{t}$  é calcolabile utilizzando la distribuzione  $\Phi$  di una variabile aleatoria normale normalizzata (si veda (3.23)), é possibile scrivere:



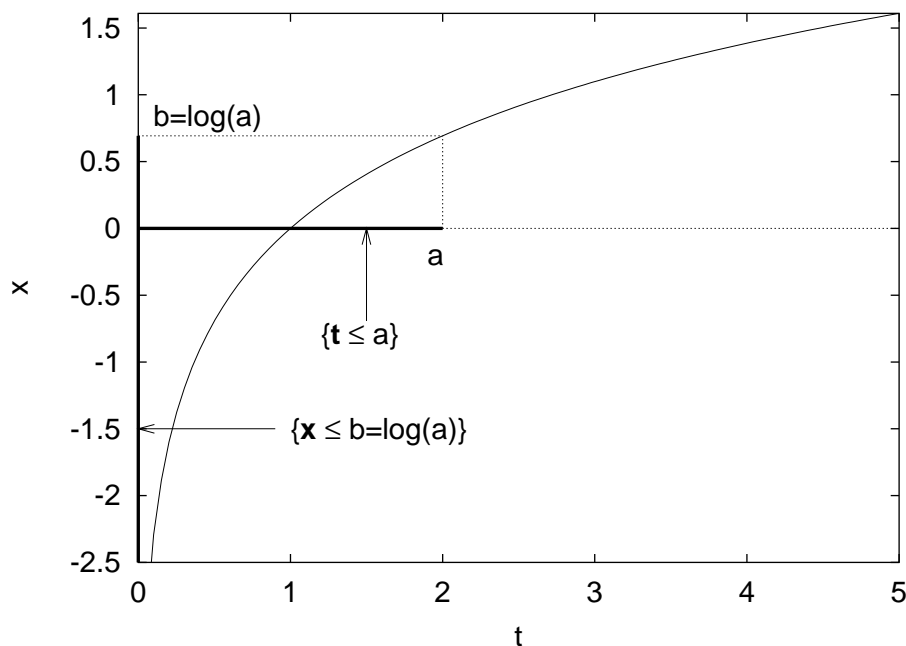


Figura 3.13: Eventi delle VA  $x$  e  $t$  se  $x = \log t$ . Per effetto dell'andamento monotono del logaritmo, gli eventi indicati in figura hanno medesima probabilità.

$$F_x(\log a) = \Phi\left(\frac{\log a - \eta}{\sigma}\right) \quad (3.26)$$

Infine, ponendo  $b = \log a$  si ottiene il risultato cercato:

$$F_x(b) = \Phi\left(\frac{b - \eta}{\sigma}\right) \quad (3.27)$$

che é equivalente a scrivere:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\left(\frac{x - \eta}{\sigma}\right)^2\right) \quad (3.28)$$

Cioé la  $x$  é, come era stato premesso, normale di parametri  $\eta$  e  $\sigma$ .

### 3.7.3 Distribuzione di Weibull

La funzione di Weibull appartiene, come la lognormale, alla famiglia di distribuzioni ai valori estremi. Questa distribuzione é comunemente utilizzata per modellare i fenomeni di guasto in materiali dielettrici solidi sottoposti a livelli di campo costanti.

La distribuzione di Weibull é:

$$F(t) = (1 - \exp(-\left(\frac{t}{\alpha}\right)^\beta))H(t) \quad (3.29)$$

con densità

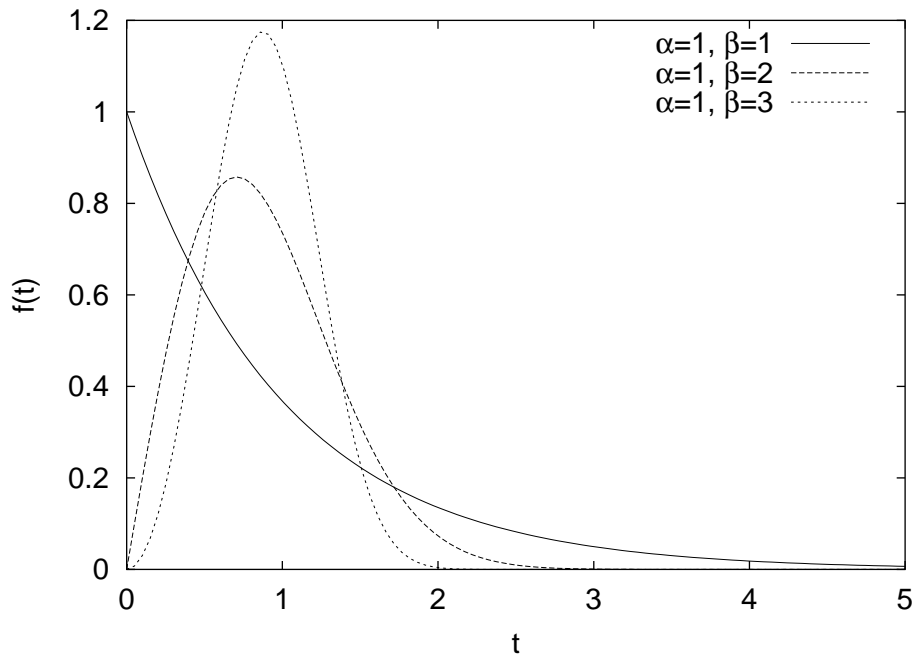


Figura 3.14: Grafico della distribuzione di Weibull per  $\alpha = 1$  e  $\beta = 1$ ,  $\beta = 2$  e  $\beta = 3$ .

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{t}{\alpha}\right)^\beta\right) H(t) \quad (3.30)$$

I parametri  $\alpha$  e  $\beta$  caratterizzano completamente la funzione. Il parametro  $\alpha$  corrisponde al percentile con probabilità del 63.2% ed è detto, pertanto, parametro di scala. Il parametro  $\beta$  caratterizza la forma della distribuzione ed è detto, appunto, parametro di forma (si veda la figura 3.14).

I momenti della distribuzione di Weibull (raramente utilizzati in pratica) sono dati da:

$$\eta = \frac{\alpha}{\beta} \Gamma(1/\beta) \quad (3.31)$$

$$\sigma^2 = \eta^2 \left( 2\beta \frac{\Gamma(2/\beta)}{\Gamma^2(1/\beta)} - 1 \right) \quad (3.32)$$

In pratica, si utilizza anche la distribuzione di Weibull a tre parametri:

$$F(t) = \left( 1 - \exp\left(-\left(\frac{t-\gamma}{\alpha}\right)^\beta\right) \right) H(t-\gamma) \quad (3.33)$$

con  $\gamma$  parametro di localizzazione.

### 3.7.4 Distribuzione esponenziale

La distribuzione esponenziale è un caso particolare della distribuzione di Weibull, ottenuta con  $\beta = 1$ . Il parametro  $\alpha$  viene sostituito, generalmente, dal reciproco  $\lambda = \frac{1}{\alpha}$  e la distribuzione esponenziale viene pertanto scritta come:

$$F(t) = 1 - \exp(-\lambda t) \quad (3.34)$$

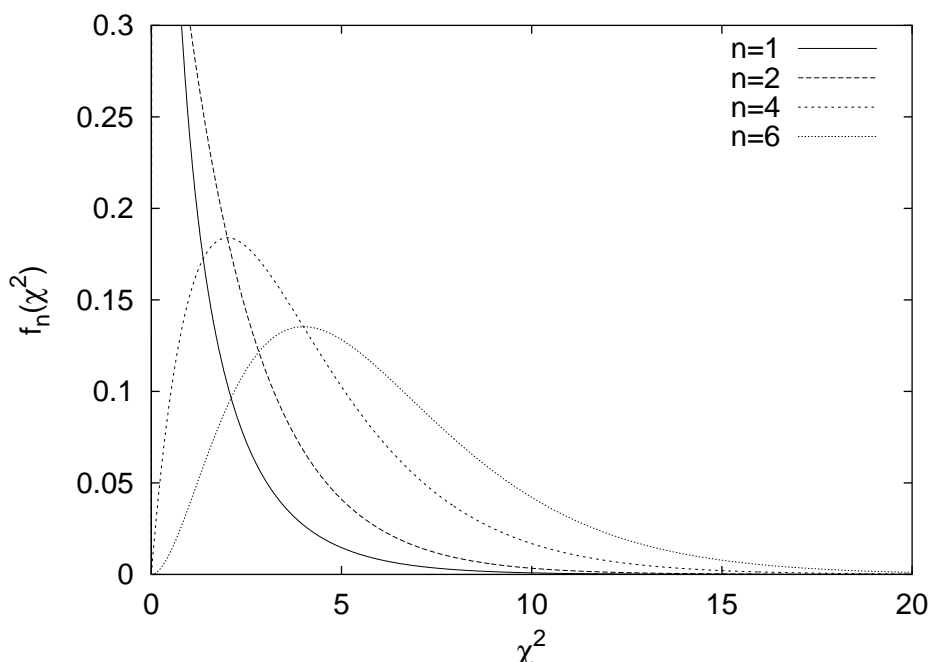


Figura 3.15: Grafico della distribuzione di chi quadro in corrispondenza a diversi gradi di libert .

con densit :

$$f(t) = \lambda \exp(-\lambda t) \quad (3.35)$$

La distribuzione esponenziale trova numerosissime applicazioni nel campo degli studi affidabilistici in quanto si postula frequentemente che i tempi al guasto dei dispositivi seguano una legge esponenziale. Il motivo di ci  sar  chiarito in seguito.

### 3.7.5 Distribuzione chi-quadro

La distribuzione chi-quadro ( $\chi^2$ )   comunemente utilizzata nel campo dell'analisi statistica dei dati per la verifica delle ipotesi sulla distribuzione di una variabile aleatoria.

La distribuzione di chi-quadro descriva la legge di probabilit  di una variabile aleatoria ottenuta come:

$$\chi^2 = \sum_{i=1}^n \mathbf{u}_i^2 \quad (3.36)$$

essendo  $\mathbf{u}_i$  ( $i = 1, \dots, n$ ) variabili aleatorie normali normalizzate fra loro indipendenti. Il numero di variabili normali normalizzate,  $n$ ,   detto numero di gradi di libert , e definisce completamente la distribuzione di chi quadro. L'espressione analitica di tale distribuzione  :

$$f(t) = \frac{1}{2^{\frac{n}{2}} \sigma^n \Gamma(n/2)} t^{\frac{(n-1)}{2}} \exp\left(-\frac{1}{2}(t/\sigma)^2\right) H(t) \quad (3.37)$$

La funzione  $\Gamma$  (gamma)   definita come:

$$\Gamma(\alpha) = \int_0^{\infty} s^{\alpha-1} \exp(-s) ds \quad (3.38)$$

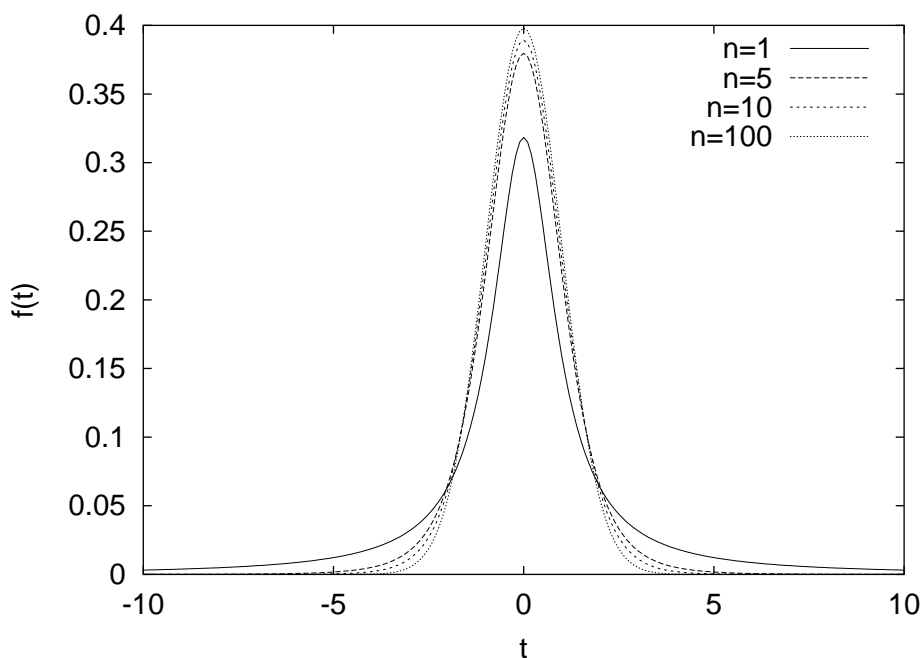


Figura 3.16: Andamento della densità di probabilità della legge  $t$  di Student per diversi valori del numero di gradi di libertà

e gode delle seguenti proprietà:

$$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$$

e se  $\alpha$  è un intero

$$\Gamma(\alpha + 1) = \alpha!$$

I valori della distribuzione di chi quadro sono normalmente forniti in forma tabulare.

### 3.7.6 Legge di probabilità , di Student

Sia  $z$  una variabile aleatoria avente distribuzione normale normalizzata ( $N(0,1)$ ),  $\chi$  una VA avente distribuzione di chi quadro con  $n$  gradi di libertà . Si dimostra che se  $z$  e  $\chi$  sono indipendenti, allora la variabile aleatoria  $t$ :

$$t = \frac{z}{\sqrt{\chi/n}} = \sqrt{n} \frac{z}{\sqrt{\chi}} \quad (3.39)$$

segue la distribuzione  $t$  di Student ad  $n$  gradi di libertà :

$$f(t) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{(n+1)}{2}} \quad (3.40)$$

L'andamento della distribuzione di Student è riportato in figura 3.16. La distribuzione di Student per 1 grado di libertà è coincidente con la distribuzione di Cauchy. Al tendere dei gradi di libertà all'infinito la distribuzione tende alla normale normalizzata.

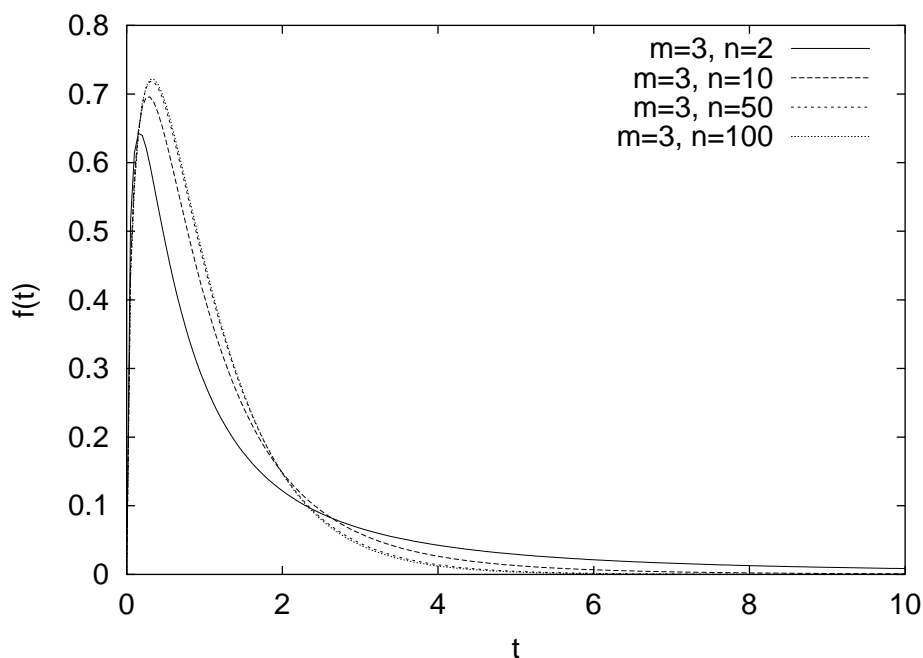


Figura 3.17: Andamento della densità di probabilità della legge  $F$  di Snedecor per valori crescenti del parametro  $n$ .

### 3.7.7 La distribuzione $F$ di Snedecor

Date due variabili aleatorie,  $\mathbf{x}$  e  $\mathbf{y}$  che seguono la distribuzione di chi quadro con  $n$  ed  $m$  gradi di libertà rispettivamente, allora il rapporto:

$$t = \frac{\mathbf{x}/n}{\mathbf{y}/m} \quad (3.41)$$

segue la legge  $F$  di Snedecor ad  $(n, m)$  gradi di libertà. Tale legge ha densità:

$$f(t) = \frac{t^{n/2-2}}{\sqrt{(1+nt/m)^{m+n}}} \quad (3.42)$$

La distribuzione di Snedecor è indicata con  $F(n, m)$ , il suo percentile a probabilità  $u$  con  $F_u(n, m)$ . Alcune rappresentazioni della densità di probabilità della legge di Snedecor sono riportate nelle figure 3.17 e 3.18

## 3.8 Distribuzioni condizionate

La probabilità condizionata di una VA è definita in modo identico a quanto fatto per un evento di un esperimento aleatorio. Detto  $\mathcal{E}$  l'evento della VA, cioè  $\mathcal{E} = \{\mathbf{x} \in U\}$ , dove  $U$  è un insieme di numeri reali misurabile in senso probabilistico, allora, se  $\mathcal{F}$  è l'evento condizionante:

$$\Pr(\{\mathbf{x} \in U\} | \mathcal{F}) = \frac{\Pr(\{\mathbf{x} \in U\} \cap \mathcal{F})}{\Pr(\mathcal{F})}$$

In particolare è definibile la distribuzione condizionata:

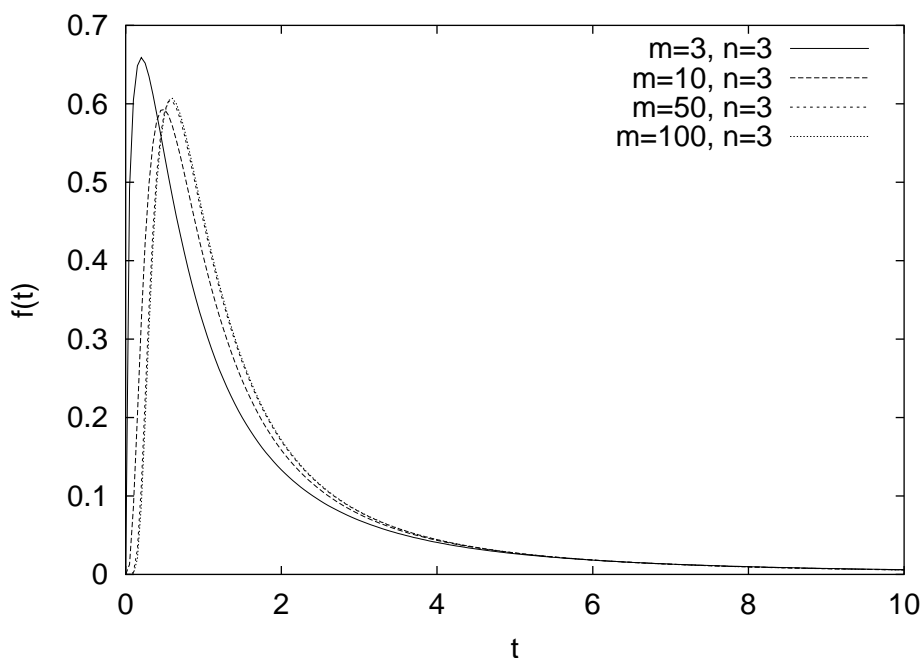


Figura 3.18: Andamento della densità di probabilità della legge  $F$  di Snedecor per valori crescenti del parametro  $m$ .

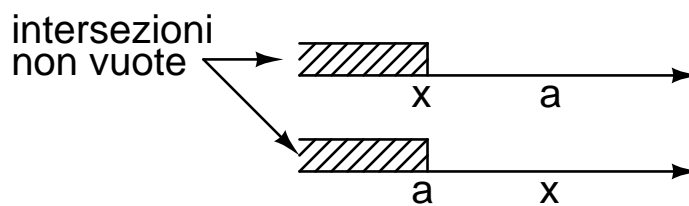


Figura 3.19: Calcolo della probabilità dell'evento  $\{\mathbf{x} \leq x\}\{\mathbf{x} \leq a\}$ .

$$F(x | \mathcal{F}) = \frac{\Pr(\{\mathbf{x} \leq x\}\{\mathcal{F}\})}{\Pr(\mathcal{F})}$$

e la densità condizionata:

$$f(x | \mathcal{F}) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x | \mathcal{F}) - F(x | \mathcal{F})}{\Delta x}$$

Tali definizioni hanno carattere generale. Per scopi pratici é tuttavia necessario specificare la natura dell'evento  $\mathcal{F}$ . In particolare, considerare i casi in cui  $\mathcal{F}$  é un evento della VA  $\mathbf{x}$  stessa.

Quando  $\mathcal{E} = \{\mathbf{x} \leq x\}$  e  $\mathcal{F} = \{\mathbf{x} \leq a\}$  allora é facile verificare (si veda figura 3.19) che:

$$\Pr(\{\mathbf{x} \leq x\}\{\mathbf{x} \leq a\}) = \begin{cases} \Pr(\{\mathbf{x} \leq x\}) = F(x) & \text{se } x \leq a \\ \Pr(\{\mathbf{x} \leq a\}) = F(a) & \text{se } x > a \end{cases}$$

Quindi

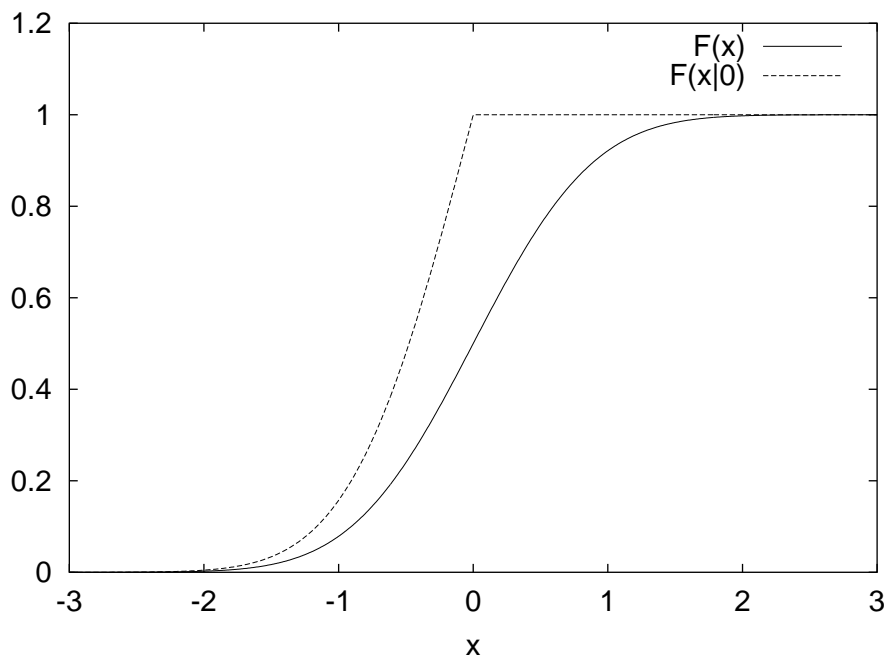


Figura 3.20: Distribuzione di una VA  $x \sim N(0,1)$  condizionata da  $x \leq 0$ .

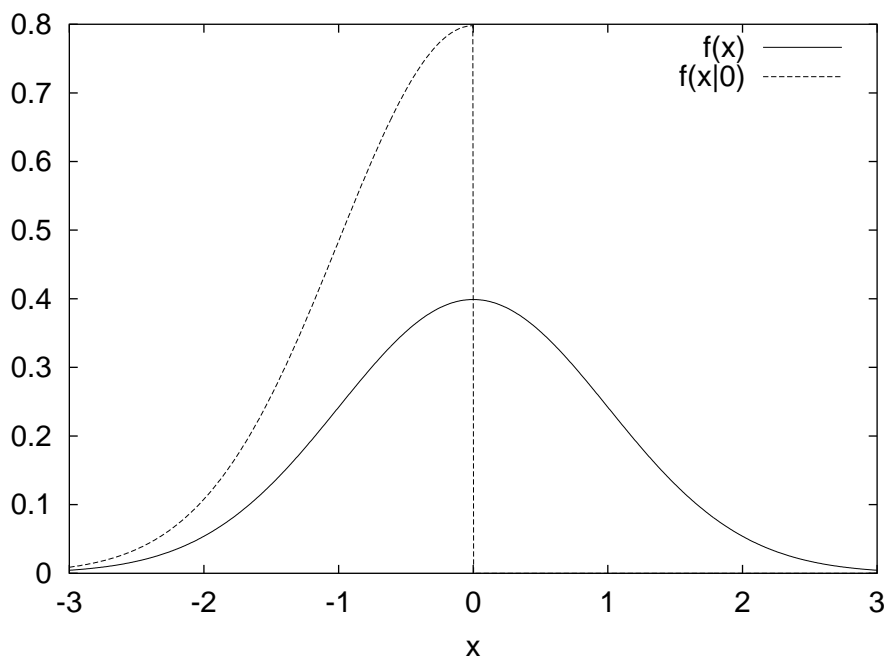


Figura 3.21: Densità di probabilità di una VA  $x \sim N(0,1)$  condizionata da  $x \leq 0$ .

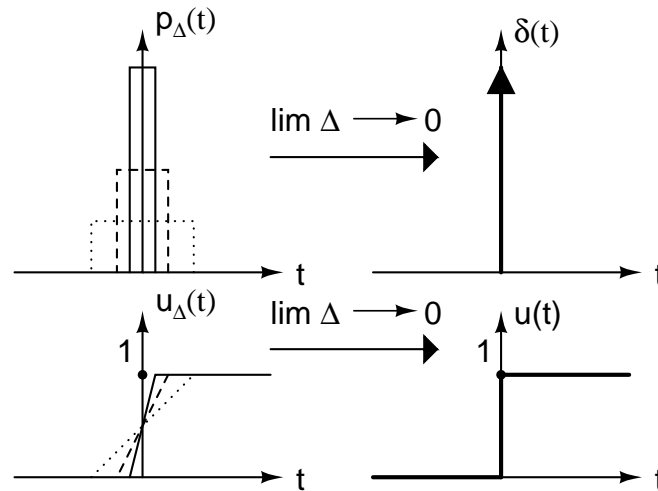


Figura 3.22: Definizione dell'impulso come limite della derivata di funzioni continue.

$$F(x | a) = \begin{cases} F(x)/F(a) & \text{se } x \leq a \\ 1 & \text{se } x > a \end{cases}$$

(si osservi la notazione semplificata  $F(x | a) = F(x | \{x \leq a\})$ ). La figura 3.20 mostra l'esempio di una VA gaussiana  $x$  condizionata da  $x \leq 0$ . E' importante osservare che la curva presenta un punto a cuspide (discontinuita' della derivata) in  $x = a$ , pur essendo continua in tal punto.

Per quanto riguarda la densita' di probabilita', poiche'

$$\Pr(\{x \in [x, x + \Delta x]\} \{x \leq a\}) = \begin{cases} \Pr(\{x \in [x, x + \Delta x]\}) & \text{se } x \leq a \\ 0 & \text{se } x > a \end{cases}$$

Quindi, calcolando il limite del rapporto incrementale per  $\Delta x$  tendente a zero, si ottiene la densita' di probabilita' condizionata dall'evento  $x > a$ :

$$f(x | a) = \lim_{\Delta x \rightarrow 0} \frac{\Pr(\{x \in [x, x + \Delta x]\} \{x \leq a\})}{\Delta x} = \begin{cases} f(x)/F(a) & \text{se } x \leq a \\ 0 & \text{se } x > a \end{cases}$$

La figura 3.21 mostra la densita' relativa alle distribuzioni di 3.20. Si noti la discontinuita' nel punto  $x = a$ , corrispondente al punto di cuspide nella distribuzione.

### 3.9 Appendice 1: L'impulso di Dirac e derivata generalizzata

L'impulso di Dirac si introduce per rappresentare eventi di durata estremamente limitata ma, comunque, in grado di fornire energia ad un sistema. Sono particolarmente utili per definire la "derivata" (da intendersi opportunamente) di funzioni affette da discontinuita' di tipo salto.

#### 3.9.1 Definizione

Uno dei possibili modi (non l'unico!) di definire l'impulso di Dirac parte dal "limite" (nel senso opportuno) di funzioni positive aventi supporto limitato ed integrale (sul supporto) unitario, indipendente dall'ampiezza del supporto. Una di tali funzioni e' la derivata del "gradino di durata  $\Delta$ ". Tale derivata, per costruzione di  $u_\Delta(t)$ , e' un rettangolo di durata  $\Delta$ . L'ampiezza del rettangolo e' pari alla derivata del tratto inclinato in  $u_\Delta(t)$ , quindi e'  $1/\Delta$ . Se ne deduce che, per un arbitrario valore di  $\Delta$  oltre alla funzione



gradino di ampiezza  $\Delta$ , possiamo definire la funzione "impulso di ampiezza  $\Delta$ ",  $p_\Delta(t)$ , caratterizzato da una durata temporale  $\Delta$ , ampiezza  $1/\Delta$ , quindi integrale unitario, indipendentemente da  $\Delta$ .

**Definizione 3.9.1 (Impulso di Dirac)** *Si definisce (non e' l'unico modo possibile) impulso unitario il limite:*

$$\delta(t) = \lim_{\Delta \rightarrow 0} p_\Delta(t) = \begin{cases} 0 & \text{per } t < 0 \\ +\infty & \text{per } t = 0 \\ 0 & \text{per } t > 0 \end{cases} \quad (3.43)$$

Evidentemente, per stabilire se la successione di funzioni  $p_\Delta$  converge verso l'impulso di Dirac (che non e' una funzione) bisogna definire un criterio di convergenza. Criterio di convergenza avente senso in uno spazio piu' ampio che l'ordinario spazio delle funzioni. L'approfondimento di tale concetto sara' oggetto del corso di Complementi di Matematica del terzo anno e richiede la conoscenza del concetto di distribuzione. Per il momento, limitiamoci ad osservare che  $p_\Delta(t)$  converge, in qualche modo, verso l'impulso di Dirac.

A volte si dice che un impulso ha ampiezza  $a$ . Tale notazione indica che la funzione generalizzata che stiamo considerando puo' essere pensata, ad esempio, come il limite a cui converge la funzione  $ap_\Delta(t)$ , che, evidentemente, ha area pari ad  $a$  per ogni valore di  $\Delta$ .

Per definizione di impulso di Dirac e di gradino unitario si pone, (anche se non ha molto senso, matematicamente):

$$\frac{du(t)}{dt} = \delta(t) \quad (3.44)$$

### Proprieta' di selezione

Studiamo l'integrale:

$$\int_{-\infty}^{\infty} f(t)\delta(t)dt \quad (3.45)$$

dove  $f$  e' una funzione continua. Utilizzando la definizione data di impulso, scriviamo l'integrale come il limite:

$$\lim_{\Delta \rightarrow 0} \int_{-\infty}^{\infty} f(t)p_\Delta(t)dt = \frac{1}{\Delta} \lim_{\Delta \rightarrow 0} \int_{-\Delta/2}^{+\Delta/2} f(t)dt \quad (3.46)$$

Al tendere di  $\Delta$  a zero il termine entro integrale tende, se la funzione e' continua, a:

$$f(0)\Delta \quad (3.47)$$

quindi:

$$\lim_{\Delta \rightarrow 0} \int_{-\infty}^{\infty} f(t)p_\Delta(t)dt = f(0) \quad (3.48)$$

Quindi e' possibile affermare che per l'impulso di Dirac vale la proprieta' di selezione:

$$\int_{-\infty}^{\infty} f(t)\delta(t)dt = f(0) \quad (3.49)$$

ed anche:

$$\int_{-\infty}^{\infty} f(t)\delta(t - \tau)d\tau = f(t) \quad (3.50)$$

Tale proprietà gode di una interessante interpretazione. Infatti, in analogia agli spazi vettoriali a dimensione finita,  $n$ , dove il generico vettore,  $\mathbf{x}$ , può essere scritto come combinazione lineare dei vettori componenti la base,  $\zeta_i = 1, \dots, n$ , mediante le componenti  $x_i$ :

$$\mathbf{x} = \sum_{i=1}^n x_i \zeta_i$$

così una funzione  $f(t)$  può essere rappresentata in una base ad infinite dimensioni (l'indice discreto  $i$  è sostituito da un indice continuo  $\tau$ , la sommatoria con un integrale), in cui i vettori della base sono gli impulsi di Dirac e le componenti sono  $f(t)$

### 3.9.2 Derivata di funzioni con discontinuità

Se una funzione  $f(t)$  è costituita da una funzione continua  $g(t)$  e da una serie di discontinuità modellabili come gradini di ampiezza e tempo di applicazione opportuno, allora, per quanto visto in precedenza:

$$\frac{df(t)}{dt} = \frac{dg(t)}{dt} + \sum_k a_k \delta(t - t_k) \quad (3.51)$$

(anche in questo caso, “derivata” è intesa in senso opportuno, facendo riferimento alle funzioni generalizzate).

## 3.10 Appendice 2: Tavole della distribuzione normale

Tabella 3.1: Problema diretto: probabilità  $\Phi(u)$  per valori di  $u$  prefissati

$u$	0.000000	0.010000	0.020000	0.030000	0.040000
0.000000	0.500000	0.503989	0.507978	0.511966	0.515953
0.100000	0.539828	0.543795	0.547758	0.551717	0.555670
0.200000	0.579260	0.583166	0.587064	0.590954	0.594835
0.300000	0.617911	0.621720	0.625516	0.629300	0.633072
0.400000	0.655422	0.659097	0.662757	0.666402	0.670031
0.500000	0.691462	0.694974	0.698468	0.701944	0.705401
0.600000	0.725747	0.729069	0.732371	0.735653	0.738914
0.700000	0.758036	0.761148	0.764238	0.767305	0.770350
0.800000	0.788145	0.791030	0.793892	0.796731	0.799546
0.900000	0.815940	0.818589	0.821214	0.823814	0.826391
1.000000	0.841345	0.843752	0.846136	0.848495	0.850830
1.100000	0.864334	0.866500	0.868643	0.870762	0.872857
1.200000	0.884930	0.886861	0.888768	0.890651	0.892512
1.300000	0.903200	0.904902	0.906582	0.908241	0.909877
1.400000	0.919243	0.920730	0.922196	0.923641	0.925066
1.500000	0.933193	0.934478	0.935745	0.936992	0.938220
1.600000	0.945201	0.946301	0.947384	0.948449	0.949497
1.700000	0.955435	0.956367	0.957284	0.958185	0.959071
1.800000	0.964070	0.964852	0.965620	0.966375	0.967116
1.900000	0.971283	0.971933	0.972571	0.973197	0.973810
2.000000	0.977250	0.977784	0.978308	0.978822	0.979325
2.100000	0.982136	0.982571	0.982997	0.983414	0.983823
2.200000	0.986097	0.986447	0.986791	0.987126	0.987455
2.300000	0.989276	0.989556	0.989830	0.990097	0.990358
2.400000	0.991802	0.992024	0.992240	0.992451	0.992656
2.500000	0.993790	0.993963	0.994132	0.994297	0.994457
2.600000	0.995339	0.995473	0.995604	0.995731	0.995855
2.700000	0.996533	0.996636	0.996736	0.996833	0.996928
2.800000	0.997445	0.997523	0.997599	0.997673	0.997744
2.900000	0.998134	0.998193	0.998250	0.998305	0.998359

Tabella 3.2: Problema diretto: probabilità  $\Phi(u)$  per valori di  $u$  prefissati

$u$	0.050000	0.060000	0.070000	0.080000	0.090000
0.000000	0.519939	0.523922	0.527903	0.531881	0.535856
0.100000	0.559618	0.563559	0.567495	0.571424	0.575345
0.200000	0.598706	0.602568	0.606420	0.610261	0.614092
0.300000	0.636831	0.640576	0.644309	0.648027	0.651732
0.400000	0.673645	0.677242	0.680822	0.684386	0.687933
0.500000	0.708840	0.712260	0.715661	0.719043	0.722405
0.600000	0.742154	0.745373	0.748571	0.751748	0.754903
0.700000	0.773373	0.776373	0.779350	0.782305	0.785236
0.800000	0.802337	0.805105	0.807850	0.810570	0.813267
0.900000	0.828944	0.831472	0.833977	0.836457	0.838913
1.000000	0.853141	0.855428	0.857690	0.859929	0.862143
1.100000	0.874928	0.876976	0.879000	0.881000	0.882977
1.200000	0.894350	0.896165	0.897958	0.899727	0.901475
1.300000	0.911492	0.913085	0.914657	0.916207	0.917736
1.400000	0.926471	0.927855	0.929219	0.930563	0.931888
1.500000	0.939429	0.940620	0.941792	0.942947	0.944083
1.600000	0.950529	0.951543	0.952540	0.953521	0.954486
1.700000	0.959941	0.960796	0.961636	0.962462	0.963273
1.800000	0.967843	0.968557	0.969258	0.969946	0.970621
1.900000	0.974412	0.975002	0.975581	0.976148	0.976705
2.000000	0.979818	0.980301	0.980774	0.981237	0.981691
2.100000	0.984222	0.984614	0.984997	0.985371	0.985738
2.200000	0.987776	0.988089	0.988396	0.988696	0.988989
2.300000	0.990613	0.990863	0.991106	0.991344	0.991576
2.400000	0.992857	0.993053	0.993244	0.993431	0.993613
2.500000	0.994614	0.994766	0.994915	0.995060	0.995201
2.600000	0.995975	0.996093	0.996207	0.996319	0.996427
2.700000	0.997020	0.997110	0.997197	0.997282	0.997365
2.800000	0.997814	0.997882	0.997948	0.998012	0.998074
2.900000	0.998411	0.998462	0.998511	0.998559	0.998605

Tabella 3.3: Problema inverso: percentili a livello di probabilità prefissato

Probabilità	0.500000	0.600000	0.700000	0.800000	0.900000
0.000000	0.000000	0.253347	0.524401	0.841621	1.281552
0.010000	0.025069	0.279319	0.553385	0.877896	1.340755
0.020000	0.050154	0.305481	0.582842	0.915365	1.405072
0.030000	0.075270	0.331853	0.612813	0.954165	1.475791
0.040000	0.100434	0.358459	0.643345	0.994458	1.554774
0.050000	0.125661	0.385320	0.674490	1.036433	1.644853
0.060000	0.150969	0.412463	0.706303	1.080319	1.750683
0.070000	0.176374	0.439913	0.738847	1.126391	1.880794
0.080000	0.201893	0.467699	0.772193	1.174987	2.053749
0.090000	0.227545	0.495850	0.806421	1.226528	2.326348

Tabella 3.4: Problema inverso: percentili a livello di probabilità prefissato

Probabilità	0.950000	0.960000	0.970000	0.980000	0.990000
0.000000	1.644853	1.750683	1.880794	2.053749	2.326348
0.001000	1.654628	1.762406	1.895698	2.074855	2.365618
0.002000	1.664562	1.774382	1.911036	2.096927	2.408916
0.003000	1.674664	1.786613	1.926837	2.120072	2.457263
0.004000	1.684940	1.799118	1.943134	2.144411	2.512144
0.005000	1.695397	1.811911	1.959964	2.170090	2.575829
0.006000	1.706042	1.825007	1.977368	2.197286	2.652070
0.007000	1.716885	1.838424	1.995393	2.226212	2.747781
0.008000	1.727932	1.852180	2.014091	2.257129	2.878162
0.009000	1.739195	1.866296	2.033520	2.290368	3.090232

## 3.11 Appendice 3: Tavole della distribuzione chi-quadro

n/P	0.001	0.005	0.025	0.05	0.1	0.25	0.5
1	0.000002	0.000039	0.000982	0.003932	0.015791	0.101531	0.454936
2	0.002001	0.010025	0.050636	0.102587	0.210721	0.575364	1.386294
3	0.024298	0.071722	0.215795	0.351846	0.584374	1.212533	2.365974
4	0.090804	0.206989	0.484419	0.710723	1.063623	1.922558	3.356694
5	0.210213	0.411742	0.831212	1.145476	1.610308	2.674603	4.351460
6	0.381067	0.675727	1.237344	1.635383	2.204131	3.454599	5.348121
7	0.598494	0.989256	1.689869	2.167350	2.833107	4.254852	6.345811
8	0.857105	1.344413	2.179731	2.732637	3.489539	5.070640	7.344121
9	1.151950	1.734933	2.700389	3.325113	4.168159	5.898826	8.342833
10	1.478743	2.155856	3.246973	3.940299	4.865182	6.737201	9.341818
11	1.833853	2.603222	3.815748	4.574813	5.577785	7.584143	10.340998
12	2.214209	3.073824	4.403789	5.226029	6.303796	8.438419	11.340322
13	2.617218	3.565035	5.008751	5.891864	7.041505	9.299066	12.339756
14	3.040673	4.074675	5.628726	6.570631	7.789534	10.165314	13.339274
15	3.482684	4.600916	6.262138	7.260944	8.546756	11.036538	14.338860
16	3.941628	5.142205	6.907664	7.961646	9.312236	11.912220	15.338499
17	4.416093	5.697217	7.564186	8.671760	10.085186	12.791926	16.338182
18	4.904849	6.264805	8.230746	9.390455	10.864936	13.675290	17.337902
19	5.406816	6.843971	8.906516	10.117013	11.650910	14.561997	18.337653
20	5.921041	7.433844	9.590777	10.850811	12.442609	15.451774	19.337429
21	6.446677	8.033653	10.282898	11.591305	13.239598	16.344384	20.337228
22	6.982968	8.642716	10.982321	12.338015	14.041493	17.239619	21.337045
23	7.529240	9.260425	11.688552	13.090514	14.847956	18.137297	22.336878
24	8.084882	9.886234	12.401150	13.848425	15.658684	19.037253	23.336726
25	8.649344	10.519652	13.119720	14.611408	16.473408	19.939341	24.336587
26	9.222127	11.160237	13.843905	15.379157	17.291885	20.843431	25.336458
27	9.802777	11.807587	14.573383	16.151396	18.113896	21.749405	26.336339
28	10.390879	12.461336	15.307861	16.927875	18.939242	22.657156	27.336229
29	10.986053	13.121149	16.047072	17.708366	19.767744	23.566586	28.336127
30	11.587951	13.786720	16.790772	18.492661	20.599235	24.477608	29.336032
40	17.916427	20.706535	24.433039	26.509303	29.050523	33.660295	39.335345
50	24.673905	27.990749	32.357364	34.764252	37.688648	42.942084	49.334937
60	31.738342	35.534491	40.481748	43.187958	46.458888	52.293817	59.334666
70	39.036377	43.275180	48.757565	51.739278	55.328940	61.698330	69.334474
80	46.519876	51.171932	57.153173	60.391478	64.277844	71.144509	79.334330
90	54.155244	59.196304	65.646618	69.126030	73.291090	80.624665	89.334218
100	61.917939	67.327563	74.221927	77.929465	82.358136	90.133220	99.334129

n/P	0.5	0.75000	0.90000	0.95000	0.97500	0.99500	0.99900
1	0.454936	1.323304	2.705543	3.841459	5.023886	7.879439	10.827566
2	1.386294	2.772589	4.605170	5.991465	7.377759	10.596635	13.815511
3	2.365974	4.108345	6.251389	7.814728	9.348404	12.838156	16.266236
4	3.356694	5.385269	7.779440	9.487729	11.143287	14.860259	18.466827
5	4.351460	6.625680	9.236357	11.070498	12.832502	16.749602	20.515006
6	5.348121	7.840804	10.644641	12.591587	14.449375	18.547584	22.457744
7	6.345811	9.037148	12.017037	14.067140	16.012764	20.277740	24.321886
8	7.344121	10.218855	13.361566	15.507313	17.534546	21.954955	26.124482
9	8.342833	11.388751	14.683657	16.918978	19.022768	23.589351	27.877165
10	9.341818	12.548861	15.987179	18.307038	20.483177	25.188180	29.588298
11	10.340998	13.700693	17.275009	19.675138	21.920049	26.756849	31.264134
12	11.340322	14.845404	18.549348	21.026070	23.336664	28.299519	32.909490
13	12.339756	15.983906	19.811929	22.362032	24.735605	29.819471	34.528179
14	13.339274	17.116934	21.064144	23.684791	26.118948	31.319350	36.123274
15	14.338860	18.245086	22.307130	24.995790	27.488393	32.801321	37.697298
16	15.338499	19.368860	23.541829	26.296228	28.845351	34.267187	39.252355
17	16.338182	20.488676	24.769035	27.587112	30.191009	35.718466	40.790217
18	17.337902	21.604890	25.989423	28.869299	31.526378	37.156451	42.312396
19	18.337653	22.717807	27.203571	30.143527	32.852327	38.582257	43.820196
20	19.337429	23.827692	28.411981	31.410433	34.169607	39.996846	45.314747
21	20.337228	24.934777	29.615089	32.670573	35.478876	41.401065	46.797038
22	21.337045	26.039265	30.813282	33.924438	36.780712	42.795655	48.267942
23	22.336878	27.141336	32.006900	35.172462	38.075627	44.181275	49.728232
24	23.336726	28.241150	33.196244	36.415029	39.364077	45.558512	51.178598
25	24.336587	29.338850	34.381587	37.652484	40.646469	46.927890	52.619656
26	25.336458	30.434565	35.563171	38.885139	41.923170	48.289882	54.051962
27	26.336339	31.528412	36.741217	40.113272	43.194511	49.644915	55.476020
28	27.336229	32.620494	37.915923	41.337138	44.460792	50.993376	56.892285
29	28.336127	33.710909	39.087470	42.556968	45.722286	52.335618	58.301173
30	29.336032	34.799743	40.256024	43.772972	46.979242	53.671962	59.703064
40	39.335345	45.616014	51.805057	55.758479	59.341707	66.765962	73.401958
50	49.334937	56.333605	63.167121	67.504807	71.420195	79.489978	86.660815
60	59.334666	66.981461	74.397006	79.081944	83.297675	91.951698	99.607233
70	69.334474	77.576655	85.527043	90.531225	95.023184	104.214899	112.316932
80	79.334330	88.130258	96.578204	101.879474	106.628568	116.321057	124.839224
90	89.334218	98.649932	107.565009	113.145270	118.135893	128.298944	137.208354
100	99.334129	109.141241	118.498004	124.342113	129.561197	140.169489	149.449253

## Capitolo 4

# Variabili aleatorie bivariate

### 4.1 Eventi

E' spesso utile potere trattare variabili aleatorie che siano una mappatura da  $S$  ad un prodotto cartesiano dell'insieme dei reali. In altre parole, il risultato degli esperimenti aleatori considerati in questa sede é una  $n$ -upla o un vettore di  $n$  componenti. Tale tipo di VA ha grande importanza in quanto si utilizza nella costruzione di modelli probabilistici relativi a misure effettuate simultaneamente, a misure tratte da esperimenti ripetuti, a serie storiche, ecc. In questa sezione saranno trattate le VA bivariate o doppie, che sono il caso piú semplice delle VA multiple o multivariate.

Le VA multivariate sono generalmente indicizzate, cioé descritte come un vettore (colonna) di VA:

$$\mathbf{z} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T$$

Quando ci si limita al caso di variabili doppie é tuttavia consuetudine semplificare la trattazione utilizzando due VA non indicizzate  $\mathbf{x}$  e  $\mathbf{y}$ .

Analogamente a quanto visto nel caso univariato, sono considerati eventi non tanto i singoli punti del piano quanto prodotti cartesiani di eventi delle due variabili aleatorie considerati singolarmente. Ad esempio, se

$$\{\mathbf{x} \in U_x\} = \{x_1 \leq \mathbf{x} \leq x_2\}$$

$$\{\mathbf{y} \in U_y\} = \{y_1 \leq \mathbf{y} \leq y_2\}$$

il corrispondente evento della VA doppia  $\mathbf{z} = (\mathbf{x}, \mathbf{y})$  é il prodotto cartesiano

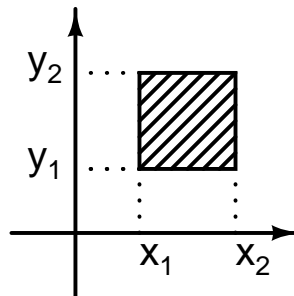


Figura 4.1: Evento elementare di una VA doppia



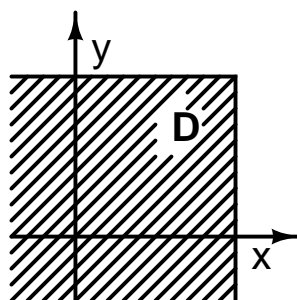


Figura 4.2: Evento di una VA doppia

$$\{\mathbf{z} \in U_{xy} = U_x \otimes U_y\} = \{\{x_1 \leq \mathbf{x} \leq x_2\}\{y_1 \leq \mathbf{y} \leq y_2\}\}$$

cioè il rettangolo di figura 4.1. In modo analogo a quanto visto nel caso della linea reale, in cui possono essere definiti semi-intervalli infiniti, semipiani infiniti come,  $D$ , di figura 4.2 possono essere eventi della variabile aleatoria doppia.

## 4.2 Distribuzione e densità di probabilità

A partire dalle definizioni relative agli eventi di una variabile aleatoria doppia è possibile definire la distribuzione della VA doppia:

$$F_{xy}(x, y) = \Pr(D) = \Pr(\{\mathbf{x} \leq x\}\{\mathbf{y} \leq y\}) \quad (4.1)$$

e la densità di probabilità:

$$\begin{aligned} f_{xy}(x, y) &= \\ &= \lim_{(\Delta x \rightarrow 0, \Delta y \rightarrow 0)} \frac{F_{xy}(x + \Delta x, y + \Delta y) - F_{xy}(x, y + \Delta y) - F_{xy}(x + \Delta x, y) + F_{xy}(x, y)}{\Delta x \Delta y} = \\ &= \frac{\partial^2 F_{xy}(x, y)}{\partial x, \partial y} \end{aligned}$$

Il termine a numeratore nel limite, ha il consueto significato, cioè rappresenta la probabilità di osservare la VA nel rettangolo infinitesimo di spigoli  $(x, y)$  ed  $(x + \Delta x, y + \Delta y)$  (si veda figura 4.2), cioè un evento elementare infinitesimo. Quindi, come nel caso univariato, la densità di probabilità non è una probabilità vera e propria ma il rapporto fra la misura probabilistica di un evento infinitesimo e la misura ordinaria (area) dell'evento. Esattamente come nel caso univariato, la densità può assumere valori compresi fra 0 e infinito e possiede unità di misura.

Il legame fra densità e distribuzione si ottiene integrando (4.2), cioè:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(s, p) dp ds \quad (4.2)$$

## 4.3 Distribuzioni marginali

Si supponga di conoscere la distribuzione di probabilità congiunta di una VA doppia (o di due VA  $\mathbf{x}$  e  $\mathbf{y}$ ), cioè

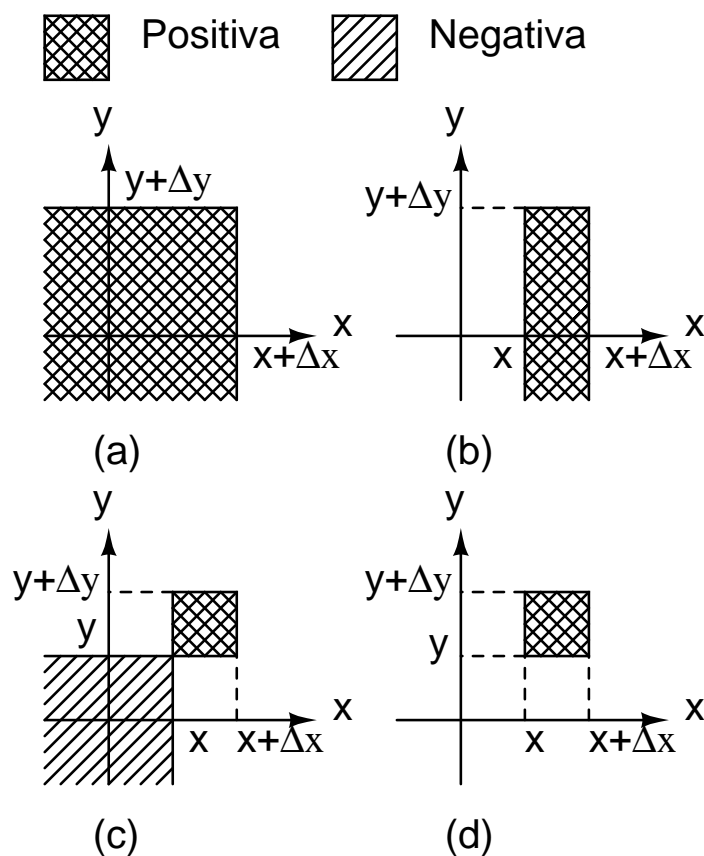


Figura 4.3: Calcolo della probabilità di un evento elementare per una VA doppia. In (a) è riportata la probabilità (legata all'area tratteggiata) di  $\{\{x \leq x + \Delta x\}\{y \leq y + \Delta y\}\}$ ; in (b) quella dell'evento  $\{\{x \leq x \leq x + \Delta x\}\{y \leq y + \Delta y\}\}$ ; in (c) quella dell'evento  $\{\{x \leq x \leq x + \Delta x\}\{y \leq y + \Delta y\} - \{x \leq x + \Delta x\}\{y \leq y\}\}$ ; in (d) quella dell'evento  $\{\{x \leq x \leq x + \Delta x\}\{y \leq y \leq y + \Delta y\}\}$ . Le probabilità di tali eventi si calcolano come: (a)  $F(x + \Delta x, y + \Delta y)$ , (b)  $F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y)$ , (c)  $F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y)$ , (d)  $F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y)$ .

$$F(x, y) = \Pr(\{\mathbf{x} \leq x\}\{\mathbf{y} \leq y\})$$

dove il pedice  $xy$  é stato omesso per semplicitá di notazione.

Il problema qui affrontato é quello di determinare la distribuzione di una sola delle due VA, ad esempio  $F_x(x)$  (oppure  $F_y(y)$ ). Tale distribuzione viene indicata come distribuzione marginale della  $\mathbf{x}$ , per indicare che la  $\mathbf{x}$  é una delle componenti di una VA bivariata o, piú in generale, multivariata.

Per risolvere il problema si consideri il significato di:

$$F(x, \infty) = \Pr(\{\mathbf{x} \leq x\}\{\mathbf{y} \leq \infty\})$$

Poiché l'evento  $\{\mathbf{y} \leq \infty\}$  avviene certamente, allora é possibile affermare che:

$$F(x, \infty) = F_x(x) \quad (4.3)$$

La correttezza di (4.3) é evidente se ci si riferisce alla definizione (4.1). Si può altresí osservare che, utilizzando l'approccio delle frequenze, la probabilità dell'evento  $\{\mathbf{x} \leq x\}\{\mathbf{y} \leq \infty\}$  é quel numero a cui tende la frequenza relativa delle osservazioni del medesimo evento rispetto al numero delle osservazioni complessive, cioè il numero di osservazioni per cui é simultaneamente verificato  $\{\mathbf{x} \leq x\}$  e  $\{\mathbf{y} \leq \infty\}$ . Ma poiché la seconda condizione é sempre vera, allora la frequenza relativa di  $\{\mathbf{x} \leq x\}\{\mathbf{y} \leq \infty\}$  coincide con quella di  $\{\mathbf{x} \leq x\}$ .

In modo analogo si osserva che:

$$F(\infty, y) = F_y(y) \quad (4.4)$$

Utilizzando l'equazione (4.2), si può esprimere la distribuzione marginale di  $\mathbf{x}$  come:

$$F_x(x) = \int_{-\infty}^x \left( \int_{-\infty}^{\infty} f(s, q) ds \right) dq$$

Analogamente:

$$F_x(x + \Delta x) = \int_{-\infty}^{x + \Delta x} \left( \int_{-\infty}^{\infty} f(s, q) ds \right) dq$$

Quindi, in base alla definizione di densitá di probabilità e scambiando l'ordine degli integrali, si ottiene:

$$f_x(x) = \lim_{\Delta x \rightarrow 0} \int_{-\infty}^{\infty} \left( \frac{1}{\Delta x} \int_x^{x + \Delta x} f(s, q) ds \right) dq$$

Il termine entro parentesi tende, per il teorema del valore medio a  $f(x, q)$ , quindi:

$$f_x(x) = \int_{-\infty}^{\infty} f(x, q) dq \quad (4.5)$$

Analogamente

$$f_y(y) = \int_{-\infty}^{\infty} f(s, y) ds \quad (4.6)$$

Si noti che la conoscenza della distribuzione congiunta permette di ricavare le distribuzioni marginali mentre il procedimento inverso non é in generale possibile in quanto le distribuzioni marginali non contengono l'informazione relativa al comportamento congiunto delle due VA. L'unico caso in cui il procedimento inverso é possibile é quando le due VA sono indipendenti, come sará mostrato nella prossima sezione.

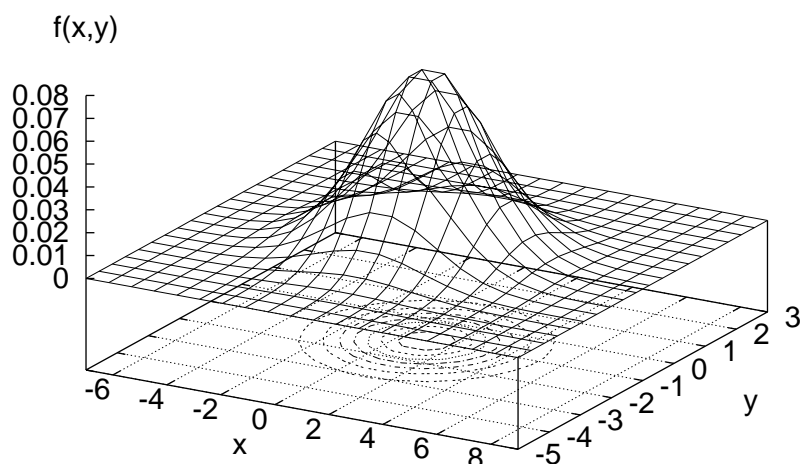


Figura 4.4: Densità di probabilità di una VA doppia congiuntamente normale avente parametri:  $\eta_x = 1$ ,  $\eta_y = -1$ ,  $\sigma_x = 2$ ,  $\sigma_y = 1$ ,  $r = 0$ .

## 4.4 Variabili aleatorie congiuntamente normali

Una coppia di VA si dice congiuntamente normale se la densità di probabilità congiunta è descritta dalla funzione:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \exp\left(-\frac{1}{2(1-r^2)}\left(\frac{(x-\eta_x)^2}{\sigma_x^2} - 2r\frac{(x-\eta_x)(y-\eta_y)}{\sigma_x\sigma_y} + \frac{(y-\eta_y)^2}{\sigma_y^2}\right)\right) \quad (4.7)$$

Si dimostra che le densità marginali delle VA  $\mathbf{x}$  e  $\mathbf{y}$  sono normali, descritte dalle funzioni:

$$f_x(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\eta_x}{\sigma_x}\right)^2\right) \quad (4.8)$$

$$f_y(y) = \frac{1}{\sigma_y\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\eta_y}{\sigma_y}\right)^2\right) \quad (4.9)$$

Il parametro  $r$  che compare nell'espressione della densità congiunta e non nelle densità marginali descrive l'accoppiamento fra le due VA  $\mathbf{x}$  ed  $\mathbf{y}$ . Esso è noto come coefficiente di correlazione. Il significato di tutti i parametri in (4.7) sarà chiarito nel capitolo relativo ai momenti di una variabile aleatoria. Al momento si osservi che il punto  $(\eta_x, \eta_y)$  del piano  $(x, y)$  coincide con il massimo di  $f(x, y)$ , che  $\sigma_x > \sigma_y$  produce un maggiore allargamento della campana nella direzione  $x$  (più chiaro se si osserva la figura 4.4) ed infine che le curve di livello sono cerchi per  $r = 0$  ellissi inclinate di 45 gradi quando  $r = 1$ .

## 4.5 Indipendenza stocastica

Quando le VA  $\mathbf{x}$  e  $\mathbf{y}$  sono indipendenti, allora la densità di probabilità assume una forma particolarmente semplice. Con riferimento alla notazione precedente ed utilizzando rettangoli aventi spigoli infinitesimi:

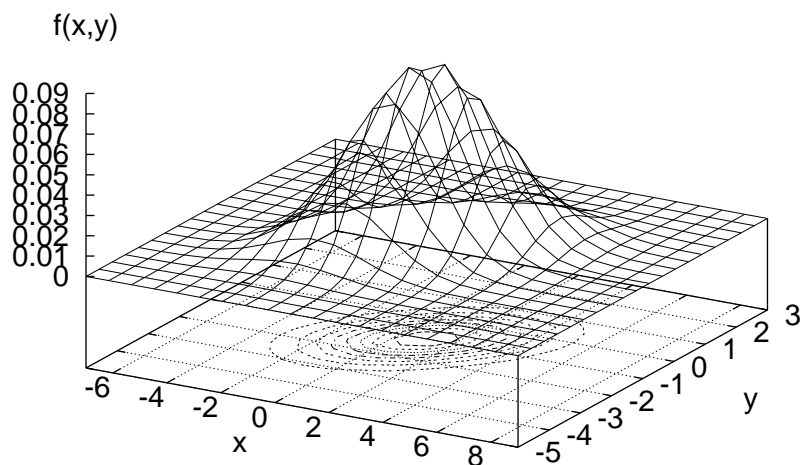


Figura 4.5: Densità di probabilità di una VA doppia congiuntamente normale avente parametri:  $\eta_x = 1$ ,  $\eta_y = -1$ ,  $\sigma_x = 2$ ,  $\sigma_y = 1$ ,  $r = 0.5$ .

$$f_{xy}(x, y) dx dy = \Pr((x, y) \in U_{xy}) = \Pr(\{x \in U_x\} \{y \in U_y\})$$

cioè il verificarsi dell'evento  $U_{xy}$  della VA doppia corrisponde al verificarsi contemporaneo (intersezione) dei due eventi  $U_x$  ed  $U_y$ . Se le due variabili sono indipendenti allora sappiamo che la probabilità dell'intersezione di due eventi è il prodotto delle probabilità, quindi:

$$f_{xy}(x, y) dx dy = f_x(x) dx f_y(y) dy$$

quindi:

$$f_{xy}(x, y) = f_x(x) f_y(y) \quad (4.10)$$

## 4.6 Alcune funzioni di VA doppie

### 4.6.1 Somma di due variabili aleatorie

Nella pratica è frequentemente richiesto il calcolo della somma di due variabili aleatorie. Si mostra qui la procedura per il calcolo della distribuzione di:

$$z = x + y$$

Si osservi che:

$$\Pr(z \leq z) = \Pr((x, y) \in D)$$

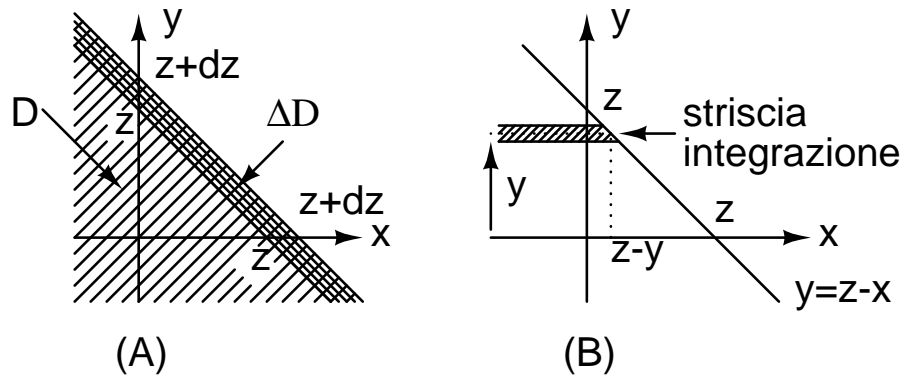


Figura 4.6: Calcolo della probabilità della somma di due variabili aleatorie.

essendo  $D$  il dominio tratteggiato in figura 4.6 (A). Dunque, utilizzando per il calcolo dell'integrale multiplo le strisce descritte nella figura 4.6 (B) si ottiene:

$$F_z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f(x, y) dx dy \quad (4.11)$$

Il calcolo di tale integrale dipende dalla densità congiunta delle due variabili aleatorie che si intende sommare. Per quanto concerne la densità di probabilità di  $z$ , utilizzando un ragionamento identico a quanto fatto per pervenire alla formulazione di (4.11), si ottiene:

$$\Pr(z \in [z, z + \Delta z]) = \int_{-\infty}^{\infty} \int_{-\infty}^{z+\Delta z-y} f(x, y) dx dy - \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f(x, y) dx dy =$$

(tale probabilità coincide con la misura probabilistica della striscia  $\Delta D$ ). Sfruttando l'additività dell'integrale ed il teorema del valore medio è possibile riscrivere

$$\Pr(z \in [z, z + \Delta z]) = \int_{-\infty}^{\infty} \int_{z-y}^{z+\Delta z-y} f(x, y) dx dy \approx \Delta z \int_{-\infty}^{\infty} f(z-y, y) dy$$

Infine, utilizzando la definizione di densità di probabilità:

$$f_z(z) = \lim_{\Delta z \rightarrow 0} \frac{\Pr(z \in [z, z + \Delta z])}{\Delta z}$$

si ottiene:

$$f_z(z) = \int_{-\infty}^{\infty} f(z-y, y) dy \quad (4.12)$$

Un caso particolare di applicazione di (4.12) è quando le due variabili aleatorie sono indipendenti, quindi  $f(x, y) = f_x(x)f_y(y)$ . Se  $x$  e  $y$  sono indipendenti si ha:

$$f_z(z) = \int_{-\infty}^{\infty} f_x(z-y)f_y(y) dy \quad (4.13)$$

Cioè la densità di probabilità della somma è la convoluzione delle densità di probabilità delle due variabili aleatorie  $x$  e  $y$ .

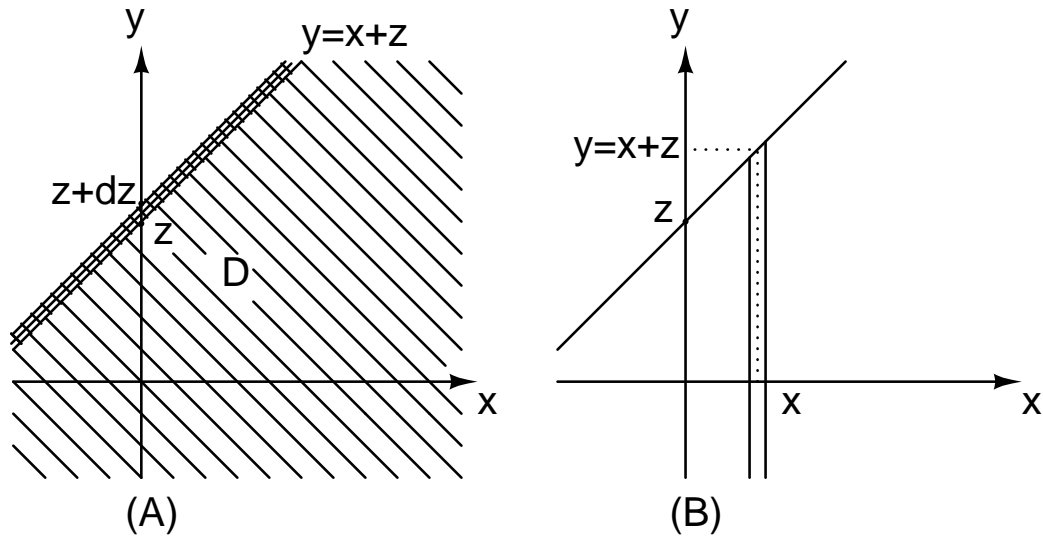


Figura 4.7: Calcolo della probabilità della differenza di due variabili aleatorie.

**Esempio 4.6.1 (Somma di due VA con distribuzione esponenziale)** Calcolare la densità di probabilità della VA  $z$ , somma di due VA esponenziali con parametri  $\lambda_1$  e  $\lambda_2$ , rispettivamente.

La trasformata di Laplace di una densità esponenziale è data da:

$$(L)(\lambda \exp(-\lambda t)) = \frac{\lambda}{s + \lambda}$$

Poiché la trasformata della convoluzione è il prodotto delle trasformate, la trasformata della densità di probabilità della somma è data da:

$$(L)(f_z(z)) = \frac{\lambda_1 \lambda_2}{(s + \lambda_1)(s + \lambda_2)}$$

Mediante l'espansione in fratti semplici ed anti trasformazioni si ottiene:

$$f_z(z) = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left( e^{-\lambda_1 z} - e^{-\lambda_2 z} \right)$$

se  $\lambda_1 \neq \lambda_2$ , altrimenti:

$$f_z(z) = \lambda^2 z e^{-\lambda z} \tag{4.14}$$

In quest'ultimo caso si dice che la VA  $z$  segue la legge di Rayleigh.

### 4.6.2 Differenza di due VA

Il calcolo della distribuzione di probabilità della differenza di VA si effettua in modo analogo a quanto fatto per la somma. Infatti, la distribuzione di probabilità è pari alla misura probabilistica del dominio  $D$  di figura 4.7 (A).

$$\Pr(z = y - x \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z+x} f(x, y) dy dx$$

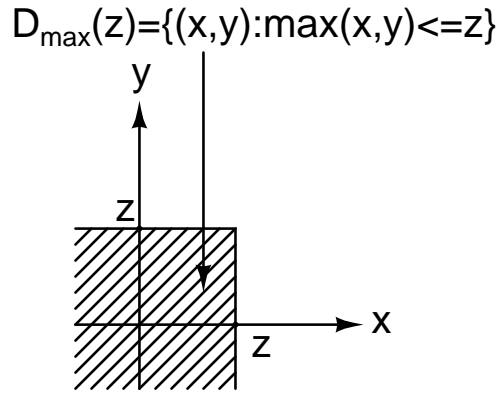


Figura 4.8: Calcolo della probabilità del massimo di due variabili aleatorie.

L'integrazione è stata fatta seguendo strisce infinitesime simili a quella mostrata in figura 4.7 (B). Per calcolare la densità di probabilità si calcola la probabilità che la VA appartenga alla striscia infinitesima di figura 4.7 (A)

$$\begin{aligned} \Pr(z \in [z, z + \Delta z]) &= \int_{-\infty}^{\infty} \int_{-\infty}^{z+\Delta z+x} f(x, y) dy dx - \int_{-\infty}^{\infty} \int_{-\infty}^{z+x} f(x, y) dy dx = \\ &= \int_{-\infty}^{\infty} \int_z^{z+\Delta z+x} f(x, y) dy dx = \\ &\approx \Delta z \int_{-\infty}^{\infty} f(x, x+z) dx \end{aligned}$$

da cui:

$$f_z(z) = \lim_{\Delta z \rightarrow 0} \frac{\Pr(z \in [z, z + \Delta z])}{\Delta z} = \int_{-\infty}^{\infty} f(x, x+z) dx \quad (4.15)$$

Nel caso  $x$  e  $y$  siano indipendenti, vale:

$$f_z(z) = \int_{-\infty}^{\infty} f_x(x) f_y(x+z) dx = \int_{-\infty}^{\infty} f_x(y-z) f_y(y) dy \quad (4.16)$$

L'ultima eguaglianza si ottiene mediante un cambio di variabile di integrazione:  $x = y - z$ . Si noti che tale equazione poteva essere ottenuta direttamente da (4.13) cambiando il segno della quantità in  $f_x$ . Ciò è conseguenza del fatto che la differenza fra due VA può essere pensata come una somma, in cui la variabile aleatoria sottratta deve essere preventivamente moltiplicata per il fattore  $-1$ . Utilizzando (3.10) si verifica immediatamente che  $f_{-x}(x) = f_x(-x)$ .

### 4.6.3 Massimo di due VA

Per calcolare la distribuzione del massimo di due variabili aleatorie,  $z = \max(x, y)$  è necessario valutare la probabilità:

$$\Pr(z = \max(x, y) \leq z) = F_z(z)$$

È immediato verificare che tutte le coppie  $(x, y)$  all'interno del dominio  $D_{\max}(z)$  (area tratteggiata) in figura 4.8 soddisfano la condizione  $\max(x, y) \leq z$ . Pertanto è:



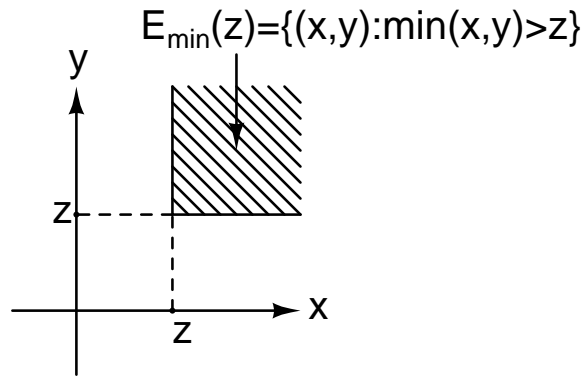


Figura 4.9: Calcolo della probabilità del minimo di due variabili aleatorie.

$$F_z(z) = \int \int_{D_{\max}(z)} f(x, y) dx dy \quad (4.17)$$

Quando le VA  $x$  e  $y$  sono indipendenti, tale relazione diventa, semplicemente:

$$F_z(z) = F_x(z)F_y(z) \quad (4.18)$$

Ciò può essere provato scrivendo  $f(x, y) = f(x)f(y)$  ed integrando (4.17). Oppure, più semplicemente, scrivendo:

$$\Pr(z = \max(\mathbf{x}, \mathbf{y}) \leq z) = \Pr(\{\mathbf{x} \leq z\}\{\mathbf{y} \leq z\}) = F_x(z)F_y(z)$$

**Esempio 4.6.2 (Tempo di guasto di un sistema parallelo)** Si supponga che due dispositivi siano posti in parallelo. Si verifica un guasto del sistema quando entrambi i dispositivi non sono più in grado di fornire il servizio a loro richiesto. Sia  $t_F$  il tempo di guasto del sistema,  $t_1$  e  $t_2$  i tempi di guasto dei due dispositivi. Il tempo di guasto del sistema coincide con il massimo dei valori assunti  $t_1$  e  $t_2$ . Pertanto, in generale è:

$$F_{t_F}(t) = \Pr(t_F \leq t) = \int \int_{D_{\max}(t)} f(t_1, t_2) dt_1 dt_2$$

Nel caso i guasti siano due VA indipendenti, allora:

$$F_{t_F}(t) = F_{t_1}(t)F_{t_2}(t)$$

#### 4.6.4 Minimo di due VA

In modo analogo a quanto fatto per il massimo di due VA, si vuole calcolare la probabilità di  $z = \min(\mathbf{x}, \mathbf{y})$ . Si comincia con il calcolo della probabilità:

$$\Pr(z = \min(\mathbf{x}, \mathbf{y}) > z) = R_z(z)$$

Dalla figura 4.9 ci si può convincere che tale probabilità può essere calcolata come la probabilità che la VA doppia  $(\mathbf{x}, \mathbf{y})$  abbia valori compresi nel dominio  $E_{\min}(z)$ :

$$R_z(z) = \int \int_{E_{\min}(z)} f(x, y) dx dy \quad (4.19)$$

quindi

$$F_z(z) = 1 - \int \int_{E_{\min}(z)} f(x, y) dx dy \quad (4.20)$$

Se  $\mathbf{x}$  e  $\mathbf{y}$  sono indipendenti, allora, posto  $R_x(z) = \Pr(\mathbf{x} > z)$  e  $R_y(z) = \Pr(\mathbf{y} > z)$ , è immediato scrivere (4.19) e (4.20) come::

$$R_z(z) = R_x(z)R_y(z) \quad (4.21)$$

$$F_z(z) = 1 - R_x(z)R_y(z) \quad (4.22)$$

**Esempio 4.6.3 (Tempo di guasto di un sistema serie)** Si supponga che due dispositivi siano posti in serie. Si verifica un guasto del sistema quando uno dei due dispositivi non è più in grado di fornire il servizio richiesto. Sia  $\mathbf{t}_F$  il tempo di guasto del sistema,  $\mathbf{t}_1$  e  $\mathbf{t}_2$  i tempi di guasto dei due dispositivi. Il tempo di guasto del sistema coincide con il minimo dei tempi di guasto dei dispositivi. In generale vale:

$$F_{\mathbf{t}_F}(t) = \Pr(\mathbf{t}_F \leq t) = 1 - \int \int_{E_{\min}(t)} f(t_1, t_2) dt_1 dt_2$$

Nel caso i guasti siano due VA indipendenti, allora, utilizzando (4.22):

$$F_{\mathbf{t}_F}(t) = 1 - R_{\mathbf{t}_1}(t)R_{\mathbf{t}_2}(t)$$

## 4.7 Distribuzioni condizionate

Si supponga di misurare, ad una data ora, la potenza attiva  $x_k$  e quella reattiva  $y_k$  a tutti i nodi di una rete elettrica di trasmissione (i simboli non sono conformi a quelli utilizzati normalmente nella letteratura tecnica per non generare fraintendimenti fra probabilità e potenza attiva). Poiché la prova è stata condotta su tutti i nodi della rete i valori misurati descrivono la distribuzione di probabilità della potenza attiva  $f_x$  e della potenza reattiva  $f_y$ .

Si supponga ora che, dall'intero insieme di dati (popolazione), si estrarrebbero solo quelli per cui la potenza attiva assume valori elevati (campione). Potenza attiva e reattiva sono generalmente dipendenti quindi, per alti valori di potenza attiva domandata, ci si attende in generale che la potenza reattiva sia abbastanza elevata. Sulla base di queste considerazioni, cioè di un campione così definito, la distribuzione della potenza reattiva per il sottoinsieme di misure considerato non coinciderebbe con quella originale (popolazione) ma tenderebbe a concentrarsi verso i più alti valori assunti dalla potenza reattiva, come mostra, qualitativamente la figura 4.10 (Nota: in questo caso il campione non è rappresentato dalla popolazione).

Si fornirà ora una impostazione matematica che permette di determinare le funzioni probabilistiche di una VA  $\mathbf{y}$  noto  $x$ , il valore assunto dalla VA  $\mathbf{x}$  che si supporrà accoppiata con  $\mathbf{y}$ . Per fare ciò si osservi la figura 4.11. La probabilità di osservare  $\mathbf{y}$  in  $[y, y + dy]$  se  $\mathbf{x}$  è in  $[x, x + dx]$  coincide con la probabilità di osservare  $(\mathbf{x}, \mathbf{y})$  in  $[x, x + dx] \otimes [y, y + dy]$  (rettangolo infinitesimo  $D$ ) divisa per la probabilità di osservare  $\mathbf{x}$  in  $[x, x + dx]$ . Poiché:

$$\Pr((\mathbf{x}, \mathbf{y}) \in [x, x + dx] \otimes [y, y + dy]) = f(x, y) dx dy$$

$$\Pr(\mathbf{x} \in [x, x + dx]) = f_x(x) dx$$

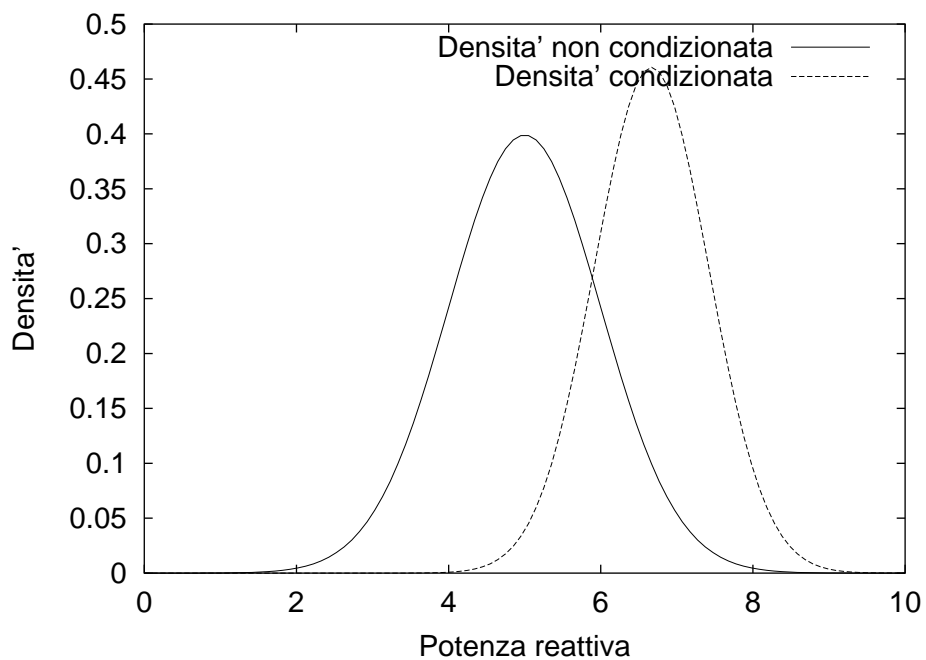


Figura 4.10: Rappresentazione qualitativa della rivalutazione della probabilità della potenza reattiva descritta nell'esempio.

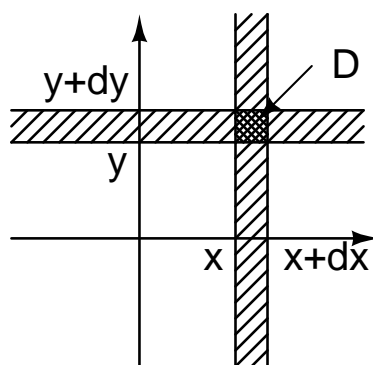


Figura 4.11: Regioni infinitesime utilizzate per il calcolo della densità di probabilità condizionata di  $y$  per specifici valori  $x$  di  $x$ .

sfruttando la definizione di probabilità condizionata ( $\Pr(\mathcal{E} | \mathcal{F}) = \Pr(\mathcal{E}\mathcal{F})/\Pr(\mathcal{F})$ ) si scrive:

$$\Pr(\mathbf{y} \in [y, y + dy] | \mathbf{x} \in [x, x + dx]) = \frac{f(x, y)dxdy}{f_x(x)dx} = \frac{f(x, y)dy}{f_x(x)} \quad (4.23)$$

Da quest'ultima relazione, definita  $f(y | x)$  come:

$$f(y | x)dy = \Pr(\mathbf{y} \in [y, y + dy] | \mathbf{x} \in [x, x + dx])$$

si ottiene:

$$f(y | x) = \frac{f(x, y)}{f_x(x)} \quad (4.24)$$

Un'altra formulazione utile dal punto di vista pratico che permette di calcolare una probabilità congiunta di variabili aleatorie note due probabilità univariate (una di questa condizionata) è:

$$f(x, y) = f(y | x)f_x(x) \quad (4.25)$$

Poiché:

$$f_y(y) = \int_{-\infty}^{\infty} f(x, y)dx$$

la distribuzione marginale di  $\mathbf{y}$  può essere ottenuta come:

$$f_y(y) = \int_{-\infty}^{\infty} f(y | x)f_x(x)dx \quad (4.26)$$

Questa relazione esprime, in forma continua, il teorema della probabilità totale (confronta equazione (2.14)).

Si noti infine che, spesso, ha interesse conoscere la distribuzione di probabilità condizionata. Per fare ciò basta integrare la densità condizionata (si veda (4.23) da cui si evince che  $f(y | x) = \Pr(\mathbf{y} \in [y, y + dy] | \mathbf{x} \in [x, x + dx])/dy$ ) ottenendo:

$$\Pr(\mathbf{y} \leq y | \mathbf{x} \in [x, x + dx]) = \int_{-\infty}^y f(y | x)dy$$

#### 4.7.1 Variabili aleatorie congiuntamente normali

Sia:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \exp\left(-\frac{1}{2(1-r^2)}\left((x/\sigma_x)^2 - 2r\frac{xy}{\sigma_x\sigma_y} + (y/\sigma_y)^2\right)\right)$$

la densità di probabilità congiunta di due VA normali aventi  $\eta_x = \eta_y = 0$ .

Poiché la densità marginale di  $\mathbf{x}$  è (si veda (4.8))

$$f_x(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x/\sigma_x)^2\right)$$

allora, dopo alcuni passaggi algebrici, è possibile provare che:

$$f(y | x) = \frac{f(x, y)}{f_x(x)} = \frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-r^2}} \exp\left(-\frac{1}{2}\left(\frac{y - r\sigma_y x/\sigma_x}{\sigma_y\sqrt{1-r^2}}\right)^2\right)$$

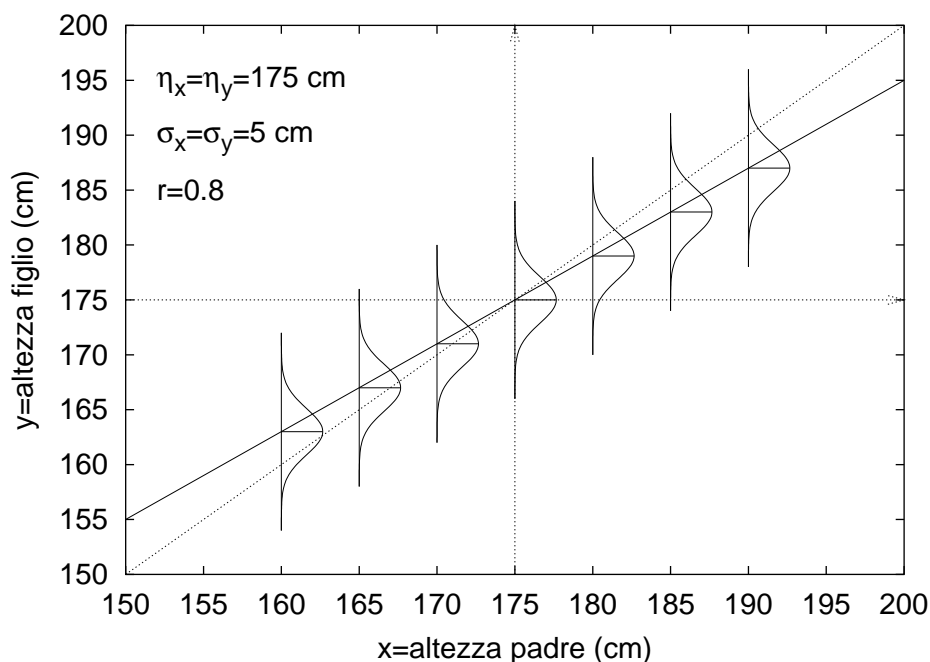


Figura 4.12: Densità di probabilità condizionate  $f(y | x)$  relative all'esempio 4.7.1

Cioè la VA  $y$ , noto il valore  $x$  di  $x$  è una variabile normale di parametri:

$$\eta_{y|x} = \left( r \frac{\sigma_y}{\sigma_x} \right) x \quad (4.27)$$

$$\sigma_{y|x} = \sigma_y \sqrt{1 - r^2} \quad (4.28)$$

Quando i parametri  $\eta_x$  ed  $\eta_y$  non siano entrambi uguali a 0, allora (4.27) può essere riscritta come:

$$\eta_{y|x} = \eta_y + r(x - \eta_x) \frac{\sigma_y}{\sigma_x} \quad (4.29)$$

**Esempio 4.7.1 (Altezza di genitori e figli (legge di Galton))** Tracciare la densità di probabilità condizionate  $f(y | x)$  se  $x$  e  $y$  sono due VA congiuntamente normali aventi parametri  $\eta_x = \eta_y = 175$ ,  $\sigma_x = \sigma_y = 5$ ,  $r = 0.8$ .

La densità condizionata è normale, avente parametri:

$$\eta_{y|x} = 175 + 0.8(x - 175)$$

$$\sigma_{y|x} = \sigma_y \sqrt{1 - r^2} = 3$$

Alcune delle densità di probabilità (per  $x = 165, x = 170, \dots, x = 190$  sono mostrate nella figura 4.12). Tali curve, specificamente, rappresentano la densità di probabilità dell'altezza dei figli (VA  $y$ ) calcolate sapendo l'altezza media dei genitori (VA  $x$ ). Tali curve formano la base della cosiddetta legge di regressione verso la media (attribuibile al genetista Galton), che sarà discussa nell'esempio 5.8.1

## Capitolo 5

# Momenti di una variabile aleatoria

Nella pratica non é quasi mai necessario specificare completamente la distribuzione di una variabile aleatoria. In genere é sufficiente localizzare una regione della linea dei numeri reali in cui tendono a concentrarsi maggiormente i risultati sperimentali. Per questo motivo, analogo a quello per cui nella geometria delle masse vengono definiti il baricentro ed il momento di inerzia, nella teoria delle probabilità vengono introdotte misure di intensità, valore atteso, moda e mediana e misure di dispersione, come la deviazione standard. I primi tipi di misura ci dicono, approssimativamente, quale srá il punto dell'asse dei numeri reali in cui, piú frequentemente, si osserveranno i valori della VA. Le misure di dispersione ci informano, invece, su quanto é larga la regione su cui si può considerare significativamente diversa da zero la probabilità di osservare i valori della variabile aleatoria.

### 5.1 Previsione di una variabile aleatoria

#### 5.1.1 Previsione di una sequenza di dati

Si supponga di avere  $n$  palle in una urna. Su ciascuna di queste palle sia stampigliato un numero  $x_k$ . La sequenza di tali numeri  $\{x_k, k = 1, \dots, n\}$  non deve essere necessariamente costituita da interi. Non é nemmeno richiesto che lo stesso numero non compaia su due distinte palle. Si proceda ad estrarre senza reintroduzione le palle dall'urna e si cerchi di fissare un numero  $\eta$  tale che la differenza fra il numero letto sulla palla ed  $\eta$  sia minima su tutte le  $n$  successive estrazioni. Per risolvere il problema si deve innanzitutto definire cosa é la differenza. Normalmente la differenza  $d_k$  é definita come  $d_k = (x_k - \eta)^2$ . Il vantaggio di tale scelta é legato a considerazioni di natura analitica ( $\eta$  é determinabile in modo semplice), e fisica ( $d_k$  é quadratica, la somma dei valori assunti da  $d_k$  é assimilabile ad un valore efficace). E' immediato provare che:

$$\eta = \frac{1}{n} \sum_{k=1}^n x_k \quad (5.1)$$

é quel valore di  $\eta$  per cui la funzione costo di predizione (o energia dell'errore di predizione):

$$W(\eta) = \sum_{k=1}^n (x_k - \eta)^2 \quad (5.2)$$

é minima. Infatti, azzerando la derivata di  $W$  rispetto ad  $\eta$  si ha:

$$\frac{\partial W}{\partial \eta} = 2 \left( \sum_{k=1}^n (x_k - \eta) \right) (-1) = 0$$

Si ottiene l'equazione (5.1), che esprime la media aritmetica della sequenza.

Quindi, supponendo di svuotare l'urna, calcolare  $\eta$  ed eseguire l'esperimento proposto, allora  $\eta$  sarà la migliore predizione del  $k$ -esimo numero  $x_k$  in uscita dall'urna, migliore nel senso che l'energia dell'errore (o l'errore quadratico medio  $W/n$ ) è minima.

È importante notare che scegliendo altre nozioni di distanza fra predittore e numero  $x_k$  si arriverebbe a soluzioni radicalmente diverse. Si provi a pensare che accade se (a)  $W$  fosse definita come  $\sum_k |x_k - \eta|$  o (b) come il massimo delle differenze assolute  $|x_k - \eta|$ . Queste misure dell'errore di predizione sono altrettanto ragionevoli di quella quadratica. Tuttavia hanno il grosso inconveniente di non condurre ad una soluzione analitica del problema proposto. Inoltre, per sua natura, l'errore quadratico tende a enfatizzare (quadraticamente) le differenze  $x_k - \eta$  di più alto valore assoluto, mentre le altre nozioni di distanza, specialmente (a) tendono a non pesare in modo adeguato forti errori di predizione.

### 5.1.2 Comportamento asintotico: il valore atteso e media

È stato stabilito che il valore medio è il migliore predittore di una sequenza numerica i cui valori sono noti ma di cui non è noto l'ordine di presentazione. Si cercherà ora di estendere i concetti visti in precedenza alle variabili aleatorie. Sia, pertanto,  $\mathbf{x}$  una variabile aleatoria discreta di cui si possono osservare i valori  $v_1, \dots, v_M$ . Della VA  $\mathbf{x}$  siano disponibili  $n$  misure,  $x_1, \dots, x_N$  (dove  $x_k = v_j$ ) (si noti che tali misure possono fornire più valori uguali,  $v_j$ ). Il calcolo della media aritmetica può essere condotto in due modi distinti, il primo sommando tutti i valori osservati della VA:

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k \quad (5.3)$$

Il secondo osservando che se  $n_i$  è il numero di volte in cui la VA ha assunto il valore  $v_i$ , allora:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^M v_i n_i = \sum_{i=1}^M v_i \frac{n_i}{n} \quad (5.4)$$

Le equazioni (5.3) e (5.4) sono espressioni della media aritmetica dei valori assunto da una variabile aleatoria presi singolarmente (eq. (5.3)) oppure raggruppati in base al loro valore ((5.4)). L'equazione (5.4) è anche l'espressione della cosiddetta media pesata. Infatti, si supponga di avere misurato i valori della VA  $x_k$ , con  $k = 1, \dots, n$ . È ora possibile raggruppare tutti quegli esperimenti che abbiano fornito risultati identici,  $v_i$ . Sia  $n_i$  il numero di volte in cui la VA ha assunto il valore  $v_i$ . Evidentemente è  $\sum_i n_i = n$ . È immediato verificare che il valore medio è dato dalla (5.4).

**Esempio 5.1.1** Sono state eseguite misure della tensione di scarica di un sistema isolante in aria. I risultati, in kV, sono pari a 50.9, 50.9, 50.9, 51.3, 52.1, 52.1, 53.2, 53.2, 53.2, 55.7. Calcolare il valore medio della tensione di scarica.

Evidentemente:

$$\bar{x} = \frac{50.9 + 50.9 + 50.9 + 51.3 + 52.1 + 52.1 + 53.2 + 53.2 + 53.2 + 55.7}{10} = 52.35 \text{ kV}$$

Alternativamente

$$\bar{x} = \frac{3(50.9) + 1(51.3) + 2(52.1) + 3(53.2) + 1(55.7)}{3 + 1 + 2 + 3 + 1} = 52.35 \text{ kV}$$

Al tendere di  $n$  all'infinito, in accordo con la definizione classica di probabilità, il rapporto  $n_i/n$  tende alla probabilità dell'evento  $\mathbf{x} = v_i$ . Per una VA discreta si definisce quindi il valore atteso:

$$\eta = \mathbf{E}[\mathbf{x}] = \sum_{i=1}^M v_i \Pr(\mathbf{x} = v_i) \quad (5.5)$$

La media aritmetica é pertanto una approssimazione (stima) del valore atteso della variabile aleatoria. Tanti piú campioni sono disponibili tanto piú la stima é precisa.

Il numero  $\eta = \mathbf{E}[\mathbf{x}]$  é il migliore predittore (nel senso menzionato in precedenza) di  $\mathbf{x}$ . Cioé, se l'esperimento di predizione potesse essere ripetuto infinite volte, il migliore predittore di una sequenza infinitamente lunga sarebbe  $\eta$ . Si noti che per sequenze non infinitamente lunghe il migliore predittore sarebbe la media aritmetica della sequenza, che tuttavia non può che essere calcolata *a posteriori*. Il valore atteso é invece qualcosa che appartiene alla conoscenza *a priori*, cioè nella teoria delle probabilità viene normalmente postulata la distribuzione della VA e, da questa, si può calcolare il valore atteso.

Per una VA continua si procede nel modo seguente. Si discretizza l'asse reale in tanti sub-intervalli di ampiezza  $\Delta$ :  $U_i = [i\Delta, (i+1)\Delta]$  dove  $-\infty < i < \infty$  per coprire l'intero asse reale. Se  $\mathbf{x} \in U_i$  allora si costruisce una nuova VA  $\mathbf{x}_\Delta$  che assume il valore  $v_i = (i+0.5)\Delta$ . In questo modo si é generato una VA discreta il cui valore atteso é calcolabile come

$$\mathbf{E}[\mathbf{x}_\Delta] = \sum_i v_i \Pr(\mathbf{x}_\Delta = v_i)$$

Poiché  $\Pr(\mathbf{x}_\Delta = v_i)$  é uguale alla probabilità che  $\mathbf{x}$  appartenga all'intervallo  $U_i$ :

$$\mathbf{E}[\mathbf{x}_\Delta] = \sum_i v_i (F_x(i\Delta) - F_x((i+1)\Delta)) \quad (5.6)$$

Facendo tendere a zero l'ampiezza dei sub-intervalli  $U_i$  la differenza  $F_x(i\Delta) - F_x((i+1)\Delta)$  può essere approssimata dal differenziale di  $F_x$  calcolato in  $v_i$  e la sommatoria può essere sostituita con un integrale. In questo modo, se esiste il limite per  $\Delta$  tendente a zero, si arriva a definire il valore atteso di una VA continua come:

$$\mathbf{E}[\mathbf{x}] = \int_{-\infty}^{\infty} x dF_x(x) = \int_{-\infty}^{\infty} x f_x(x) dx \quad (5.7)$$

La stima del valore atteso di una VA continua può essere ottenuta, analogamente a quanto fatto per la VA discreta, sfruttando l'equazione (5.3), cioè la media aritmetica dei risultati.

Si noti che, in base a (5.7), se la distribuzione presenta un asse di simmetria in  $x = \zeta$ , allora il valore atteso sarà localizzato in corrispondenza dell'asse di simmetria, cioè  $\eta = \zeta$ . Infatti:

$$\int x f(x) dx = \int (x - \zeta + \zeta) f(x) dx = \int (x - \zeta) f(x) dx + \zeta \int f(x) dx = \zeta$$

dove l'ultimo passaggio é stato fatto sfruttando la simmetria di  $f$  rispetto a  $x = \zeta$ , quindi  $\int (x - \zeta) f(x) dx = 0$ , e sfruttando il fatto che  $\int f(x) dx = 1$ .

**Esempio 5.1.2 (Valore atteso di una VA gaussiana)** Per la simmetria della funzione:

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

il valore atteso di una  $\mathbf{x}$  distribuita secondo una legge gaussiana di parametri  $\eta$  e  $\sigma$  coincide con il parametro  $\eta$  della legge normale.

---



Con un ragionamenti formalmente analoghi a quelli fatti per formulare (5.7), si dimostra che se:

$$\mathbf{y} = g(\mathbf{x})$$

allora:

$$\mathbf{E}[\mathbf{y}] = \int_{-\infty}^{\infty} g(x) dF_x(x) = \int_{-\infty}^{\infty} g(x) f_x(x) dx \quad (5.8)$$

Infine, quando le VA siano doppie e  $\mathbf{z}$  sia la VA definita come  $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$ , allora il valore atteso di  $\mathbf{z}$  si calcola come:

$$\mathbf{E}[\mathbf{z}] = \mathbf{E}[g(\mathbf{x}, \mathbf{y})] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy \quad (5.9)$$

### 5.1.3 La probabilità come valore atteso

Sia  $\mathcal{E}$  un evento avente probabilità  $p$ . Si costruisca la VA:

$$\mathbf{x} = l(\mathcal{E})$$

dove  $l$  é una funzione pari a 1 se si verifica  $\mathcal{E}$ , 0 altrimenti. Il valore atteso di  $\mathbf{x}$  é:

$$\mathbf{E}[\mathbf{x}] = 1(p) + 0(1 - p) = p$$

Quindi la probabilità dell'evento  $\mathcal{E}$  coincide con il valore atteso di  $\mathbf{x} = l(\mathcal{E})$ .

### 5.1.4 Esistenza del valore atteso

Non sempre é detto che il valore atteso esista. Condizione necessaria perché ciò si verifichi é che siano convergenti la serie

$$\sum_{i=1}^{\infty} |v_i| \Pr(\mathbf{x} = v_i)$$

e l'integrale:

$$\int_{-\infty}^{\infty} |x| f_x(x) dx$$

Ovvero che esistano i valori attesi del valore assoluto della VA considerata. Una variabile aleatoria distribuita secondo la densità di Cauchy:

$$f_x(x) = \frac{1}{\pi(1+x^2)}$$

non ammette il valore atteso in quanto l'integrale

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)} dx$$

ha come risultato  $\infty$ .

### 5.1.5 Linearit  del valore atteso

Poich  il valore atteso   definito in base a somme ed integrali gode della propriet  di linearit . Si dimostra (in modo analogo a quanto fatto in precedenza per pervenire a (5.7)) che vale:

$$E\left[\sum_k \alpha_k \mathbf{x}_k\right] = \sum_k \alpha_k E[\mathbf{x}_k] \quad (5.10)$$

Valgono i due casi particolari:

$$E[a\mathbf{x}] = aE[\mathbf{x}] \quad (5.11)$$

$$E[a + \mathbf{x}] = a + E[\mathbf{x}] \quad (5.12)$$

### 5.1.6 Altre misure di intensit 

Altre misure di intensit  sono la moda:

$$m = \max_x f(x)$$

e la mediana, cio  il percentile al 50% della distribuzione della VA:

$$\text{mediana}(\mathbf{x}) = x_{0.5} \rightarrow F_x(x_{0.5}) = 0.5$$

Se una distribuzione presenta un asse di simmetria e esiste il valore atteso, allora   immediato verificare che valore atteso e mediana coincidono. Se poi, come accade nella distribuzione normale, il massimo della densit  cade in corrispondenza dell'asse di simmetria, allora le tre misure sono coincidenti.

Si osservi che la moda   il previsore che ha la massima probabilit  di essere prossimo alla realizzazione della VA. Tuttavia, la moda accetta che, anche se con bassa probabilit , ci possano essere forti errori di previsione.

E' facile provare che la mediana   il previsore che minimizza lo scostamento assoluto medio fra la realizzazione ed il previsore. Infatti, detto  $\zeta$  il previsore (per non ingenerare fraintendimenti con il valore atteso) il valore atteso dell'errore di previsione  :

$$E[|x - \zeta|] = \int_{-\infty}^{\infty} |x - \zeta| f(x) dx$$

Per derivare rispetto al previsore ( $\zeta$ ) si osservi che

$$\frac{d|x - \zeta|}{d\zeta} = \lim_{h \rightarrow 0} \frac{|x - \zeta - h| - |x - \zeta|}{h} = \begin{cases} -1 & \text{se } x - \zeta \geq 0 \\ 1 & \text{se } x - \zeta < 0 \end{cases}$$

pertanto:

$$\frac{dW}{d\zeta} = - \int_{\zeta}^{\infty} f(x) dx + \int_{-\infty}^{\zeta} f(x) dx$$

La derivata si annulla nel punto in cui  $\zeta$  rende i due integrali uguali in valore assoluto, cio  per  $\zeta = x_{0.5}$

**Esempio 5.1.3 (Previsori di una VA distribuita secondo la legge esponenziale)** I previsori di una VA  $\mathbf{x}$  che segue la legge esponenziale di parametro  $\lambda$  sono:

$$m = 0$$

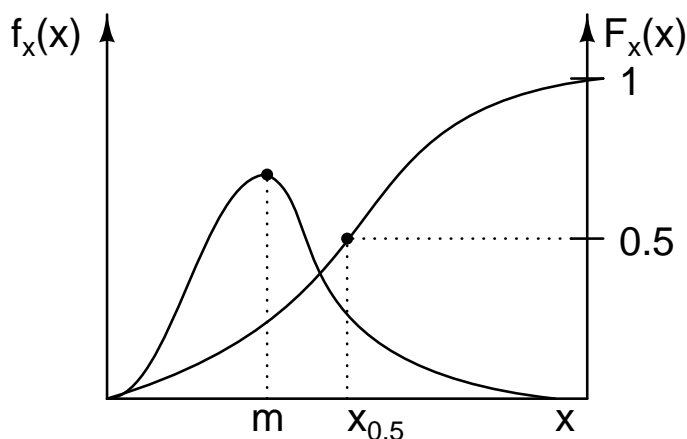


Figura 5.1: Significato di moda e varianza per una distribuzione di probabilità

$$\eta = \int_0^{\infty} x \lambda \exp(-\lambda x) dx = 1/\lambda$$

$$x_{0.5} = -\log(0.5)/\lambda = 0.693/\lambda$$

## 5.2 Momenti del secondo ordine di VA univariate: varianza

La varianza altro non é che il valore atteso della variabile aleatoria

$$y = (x - E[x])^2$$

cioé del quadrato della VA  $x$  centrata, cioè a cui é stato tolto il valore atteso (quindi a valore atteso nullo). Per quanto indicato in precedenza, la varianza altro non é che il valore quadratico medio dell'errore di previsione di  $x$  posto che sia noto  $E[x]$ .

A livello di notazione si indica:

$$\sigma_x^2 = E[y]$$

Quindi, per variabili aleatorie discrete é:

$$\sigma_x^2 = \sum_i (v_i - E[x])^2 \Pr(x = v_i) \quad (5.13)$$

Per variabili aleatorie continue, tenendo conto della definizione di valore atteso (eq. (5.7)) si ottiene:

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - E[x])^2 f_x(x) dx \quad (5.14)$$

é immediato verificare che  $\sigma^2 \geq 0$ . Il caso  $\sigma^2 = 0$  si raggiunge solo quando la VA non si discosta mai dal valore atteso, ossia quando la variabile considerata é perfettamente deterministica.

Applicando la proprietá di linearitá del valore atteso si dimostra inoltre che:

$$\sigma^2(a + \mathbf{x}) = \sigma_x^2 \quad (5.15)$$

$$\sigma^2(a\mathbf{x}) = a^2\sigma_x^2 \quad (5.16)$$

ed inoltre

$$\begin{aligned} \sigma_x^2 &= \mathbf{E}[(\mathbf{x} - \mathbf{E}[\mathbf{x}])^2] = \\ &= \mathbf{E}[\mathbf{x}^2 - 2\mathbf{x}\mathbf{E}[\mathbf{x}] + \mathbf{E}[\mathbf{x}]^2] = \\ &= \mathbf{E}[\mathbf{x}^2] - 2\mathbf{E}[\mathbf{x}]\mathbf{E}[\mathbf{E}[\mathbf{x}]] + \mathbf{E}[\mathbf{E}[\mathbf{x}]^2] = \\ &= \mathbf{E}[\mathbf{x}^2] - \mathbf{E}[\mathbf{x}]^2 \end{aligned}$$

(si noti che  $\mathbf{E}[\mathbf{E}[\mathbf{x}]] = \mathbf{E}[\mathbf{x}]$  in quanto  $\mathbf{E}[\mathbf{x}]$  non é una variabile aleatoria). Quest'ultima espressione permette di ricavare una stima per la varianza. Infatti, ricordando che il valore atteso di una variabile aleatoria é il limite a cui tende il valore medio delle osservazioni quando si eseguono infiniti esperimenti:

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{k=1}^n x_k^2 - \left( \frac{1}{n} \sum_{k=1}^n x_k \right)^2 = \frac{1}{n} \left( \sum_{k=1}^n x_k^2 - n\bar{x}^2 \right)$$

Quando sia possibile raggruppare i dati (cioé  $n_i$  risultati hanno il valore  $v_i$ ), allora é:

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_i (v_i n_i)^2 - \left( \frac{1}{n} \sum_i v_i n_i \right)^2$$

Ragionando in termini elettrici e supponendo che la VA  $\mathbf{x}$  sia una corrente, allora la varianza rappresenta quella quota parte di energia dissipata sui cavi della rete per effetto della variabilitá (rispetto al valore atteso della corrente) della corrente stessa.

**Esempio 5.2.1 (Media e varianza della legge esponenziale)** Calcolare media e varianza di  $\mathbf{t} \sim \lambda \exp(-\lambda t)$ .

Il valore atteso é dato da:

$$\mathbf{E}[\mathbf{t}] = \int_0^\infty t \lambda t \exp(-\lambda t) dt = \frac{1}{\lambda} \int_0^\infty s \exp(-s) ds$$

L'integrale si può calcolare per parti:

$$\int_0^\infty s \exp(-s) ds = \int_0^\infty s \frac{d(-\exp(-s))}{ds} ds = \left[ -s \exp(-s) \right]_0^\infty + \int_0^\infty \exp(-s) ds = 1$$

Quindi:

$$\mathbf{E}[\mathbf{t}] = \frac{1}{\lambda} \quad (5.17)$$

Per calcolare la varianza si può utilizzare la relazione:

$$\sigma_t^2 = \mathbf{E}[\mathbf{t}^2] - \mathbf{E}[\mathbf{t}]^2$$

$$\mathbf{E}[\mathbf{t}^2] = \int_0^\infty t^2 \lambda t \exp(-\lambda t) dt = \frac{1}{\lambda^2} \int_0^\infty s^2 \exp(-s) ds$$

L'integrale, come prima, si può calcolare per parti:

$$\int_0^{\infty} s^2 \exp(-s) ds = \int_0^{\infty} s^2 \frac{d(-\exp(-s))}{ds} ds = \left[ -s^2 \exp(-s) \right]_0^{\infty} + 2 \int_0^{\infty} s \exp(-s) ds = 2$$

Quindi:

$$\sigma_t^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \quad (5.18)$$

**Esempio 5.2.2 (Valore atteso della legge di Rayleigh)** Calcolare il valore atteso di  $t \sim \lambda^2 t \exp(-\lambda t)$

$$\mathbf{E}[t] = \int_0^{\infty} \lambda^2 t^2 \exp(-\lambda t) = \frac{1}{\lambda} \int_0^{\infty} s^2 \exp(-s) ds = \frac{2}{\lambda}$$

Si noti che la legge di Rayleigh è stata derivata calcolando la densità di probabilità della somma di due VA indipendenti che seguono la medesima legge esponenziale,  $t_1$  e  $t_2$  (si veda esempio 4.6.1. Era pertanto più rapido calcolare il valore atteso come:

$$\mathbf{E}[t] = \mathbf{E}[t_1 + t_2] = \mathbf{E}[t_1] + \mathbf{E}[t_2]$$

Procedendo per induzione, se si considera la somma di  $n$  variabili aleatorie indipendenti identicamente distribuite secondo la legge esponenziale si ottiene la VA con densità di probabilità:

$$f(t) = \lambda^n \frac{t^{(n-1)}}{(n-1)!} \exp(-\lambda t)$$

Tale densità di probabilità è nota come densità di Erlang, caso particolare della distribuzione Gamma. Il valore atteso di una VA che segua la legge di Erlang si calcola semplicemente come:

$$\mathbf{E}[t] = \mathbf{E}[t_1 + \dots + t_n] = \mathbf{E}[t_1] + \dots + \mathbf{E}[t_n] = \frac{n}{\lambda}$$

### 5.3 Il lemma di Tchebycheff

Il lemma di Tchebycheff fornisce una approssimazione per eccesso della probabilità di osservare la VA al di fuori di un intervallo centrato intorno al valore atteso. Tale valutazione è espressa da:

$$\Pr(|\mathbf{x} - \mathbf{E}[\mathbf{x}]| > \epsilon) \leq \frac{\sigma_x^2}{\epsilon^2} \quad (5.19)$$

Dove questa approssimazione ha senso se e solo se:

$$\epsilon > \sigma_x$$

(in caso contrario si potrebbero trovare valori di probabilità superiori ad 1).

**Esempio 5.3.1** Si confronti la probabilità di trovare una variabile aleatoria gaussiana in un intervallo di ampiezza  $k\sigma$  con la stima ottenuta utilizzando il lemma di Tchebycheff.

$\epsilon$	VA	VA normale
$\sigma$	$\geq 0$	0.6826
$2\sigma$	$\geq 0.75$	0.9545
$3\sigma$	$\geq 0.88$	0.9973
$4\sigma$	$\geq 0.93$	0.99993
$5\sigma$	$\geq 0.96$	0.9999994

Dalla tabella si evince che per  $k > 4$  il lemma di Tchebycheff permette di stimare la probabilità vera (calcolata come  $\Phi(k) - \Phi(-k)$ ) con un buon grado di approssimazione.

**Esempio 5.3.2 (Stima del valore medio di una VA)** Supponiamo che si stiano facendo misure di una VA. Tali misure potrebbero essere ridotte ad una sola rilevazione. La domanda che ci si pone è la seguente: la misura fatta quanto approssima il valore medio della VA?

Definiamo la deviazione relativa:

$$\mathbf{d}_r = \frac{\mathbf{x} - \mathbf{E}[\mathbf{x}]}{\mathbf{E}[\mathbf{x}]}$$

Il lemma di Tchebycheff per la deviazione relativa è applicabile tenendo conto del fatto che la deviazione relativa avrà una varianza pari a (ricordare che  $\sigma_x^2(a\mathbf{x}) = a^2\sigma_x^2$ ):

$$\frac{\sigma_x^2}{\mathbf{E}[\mathbf{x}]^2}$$

Quindi:

$$\Pr\left(|\mathbf{d}_r| \leq \epsilon\right) \geq 1 - \frac{1}{\epsilon^2} \frac{\sigma_x^2}{\mathbf{E}[\mathbf{x}]^2} + P_\epsilon$$

Si calcola immediatamente che tale probabilità è superiore a 0.95 quando  $\epsilon \geq 4.5\sigma_x / |\mathbf{E}[\mathbf{x}]|$ , superiore a 0.99 se  $\epsilon \geq 10\sigma_x / |\mathbf{E}[\mathbf{x}]|$ . Il rapporto:

$$SNR = (\text{Signal to Noise Ratio}) = \frac{\mathbf{E}[\mathbf{x}]}{\sigma_x}$$

noto come rapporto segnale/rumore, è fondamentale per decidere se una singola misura di un evento aleatorio può essere considerata una buona approssimazione del valore atteso dell'evento. Infatti, data  $P_\epsilon$ , risulta:

$$\epsilon = \frac{1}{SNR\sqrt{1-P}}$$

Per una VA  $\mathbf{x}$  distribuita secondo la legge esponenziale di parametro  $\lambda$  è:

$$\mathbf{E}[\mathbf{x}] = 1/\lambda$$

$$\sigma^2 = 1/\lambda^2$$

Pertanto

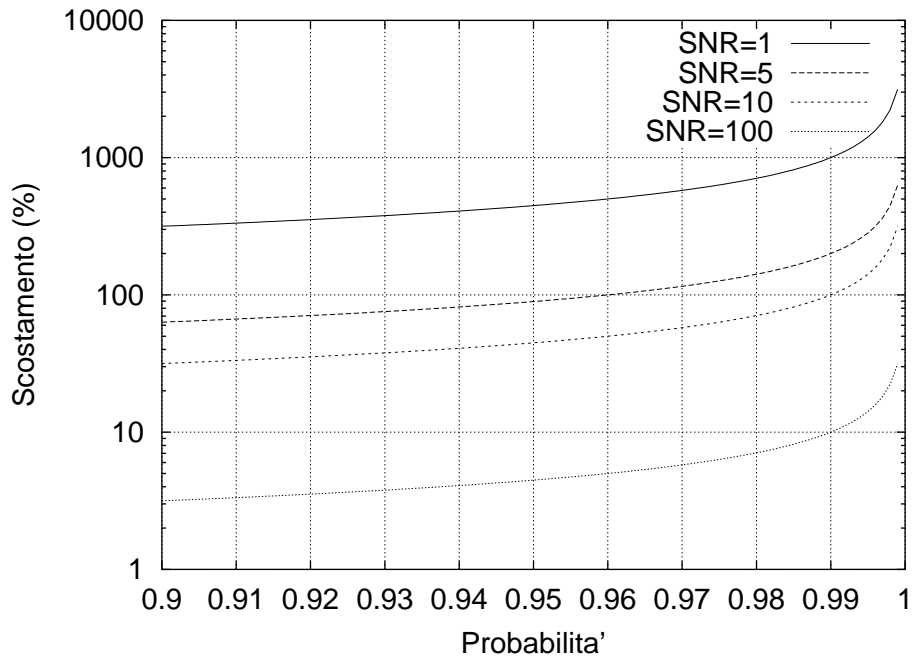


Figura 5.2: Ampiezza dell'intervallo (espressa in percentuale della media) all'interno del quale una VA avente rapporto segnale/rumore specificato (1, 5, 10 e 100) si trova con una probabilità eccedente il valore indicato in ascissa.

$$SNR = \frac{\mathbf{E}[x]^2}{\sigma^2} = \frac{1/\lambda^2}{1/\lambda^2} = 1$$

Una singola misura non é sufficiente per fornire una buona approssimazione del valore medio.

## 5.4 Altre misure di dispersione

Altre misure di dispersione sono:

**Campo di variazione**, cioè la differenza fra il massimo ed il minimo dei valori osservati.

**Scarto medio assoluto**, definito come:

$$\text{scarto medio assoluto} = \frac{1}{n} \sum_{k=1}^n |x_k - \mathbf{E}[x]| \quad (5.20)$$

(si noti che se non esistesse il valore assoluto, allora la media degli scarti rispetto al valore medio sarebbe nulla).

**Deviazione standard** E' la radice della varianza:

$$\text{deviazione standard} = \sqrt{\sigma_2^2} \quad (5.21)$$

## 5.5 Momenti di ordine superiore a 2

I momenti di ordine  $q$  sono definiti come:

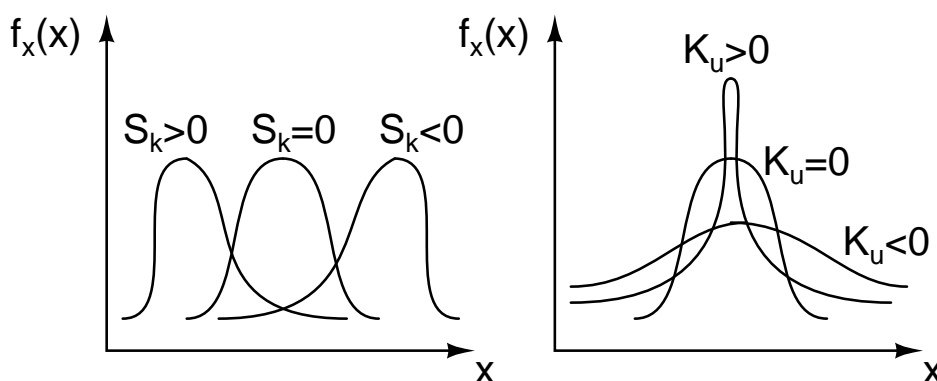


Figura 5.3: Influenza di skewness e kurtosis sulla forma della densità di probabilità.

$$m_q(x) = \int_{-\infty}^{\infty} (x - \mathbf{E}[x])^q f_x(x) dx \quad (5.22)$$

Fra i momenti di ordine superiore a 2 sono frequentemente utilizzati i momenti del terzo ordine, cioè la skewness ( $S_k$ ), e quelli del quarto ordine, cioè la kurtosis ( $K_u$ ). Tali momenti forniscono una descrizione della forma della densità di probabilità (si veda la figura 5.3).

Per applicazioni particolari sono utilizzati anche momenti frazionari e momenti negativi.

## 5.6 Momenti del secondo ordine di VA doppie: covarianza

Sfruttando l'approccio delle frequenze è immediato provare che, in generale, il valore atteso del prodotto di VA non coincide con il prodotto dei valori attesi

$$\mathbf{E}[xy] \neq \mathbf{E}[x]\mathbf{E}[y]$$

infatti, in generale, vale:

$$\sum_i x_i y_i \neq \sum_i x_i \sum_j y_j$$

L'eccezione è quella di VA indipendenti:

$$\begin{aligned} \mathbf{E}[xy] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_x(x) f_y(y) dx dy = \\ &= \int_{-\infty}^{\infty} x f_x(x) dx \int_{-\infty}^{\infty} y f_y(y) dy = \mathbf{E}[x]\mathbf{E}[y] \end{aligned} \quad (5.23)$$

Per descrivere il comportamento accoppiato di variabili aleatorie mediante momenti del secondo ordine viene, pertanto, definita la covarianza:

$$C_{xy} = \mathbf{E}[(x - \mathbf{E}[x])(y - \mathbf{E}[y])] \quad (5.24)$$

La covarianza altro non è che il valore atteso della VA:



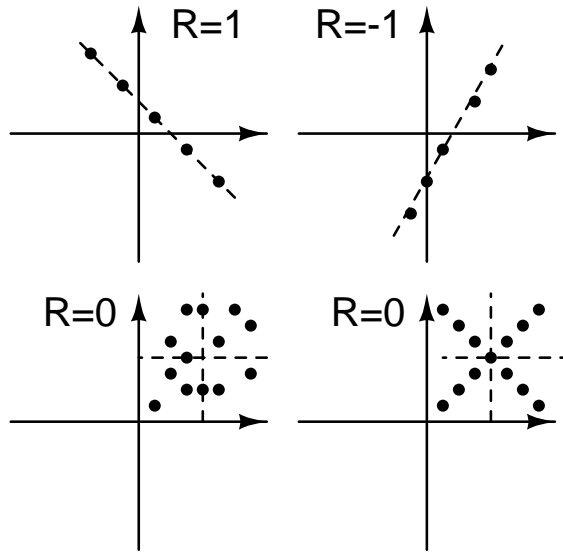


Figura 5.4: Configurazioni delle variabili aleatorie  $x$  e  $y$  per alcuni valori notevoli del coefficiente di correlazione.

$$z = (x - E[x])(y - E[y])$$

Facendo uso della proprietà di linearità di  $E[\cdot]$  si arriva ad una definizione alternativa:

$$C_{xy} = E[xy] - E[x]E[y] \quad (5.25)$$

Infatti:

$$\begin{aligned} C_{xy} &= E[xy - xE[y] - yE[x] + E[x]E[y]] \\ &= E[xy] - 2E[x]E[y] + E[x]E[y] \end{aligned} \quad (5.26)$$

Si noti che la varianza di una VA  $x$  può essere interpretata come la covarianza di  $x$  se  $y = x$ .

La covarianza normalizzata, nota come coefficiente di correlazione è definita come:

$$R_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \quad (5.27)$$

Utilizzando l'approccio delle frequenze è possibile dimostrare che:

$$-1 \leq R_{xy} \leq 1$$

Si postpone la dimostrazione di questa proprietà per analizzare quali VA diano origine ad alcuni valori notevoli del coefficiente di correlazione. Questi casi notevoli sono illustrati schematicamente nella figura 5.4. I casi  $R = -1$  ed  $R = 1$  si può dimostrare che hanno origine se esiste una relazione lineare fra le variabili aleatorie. Sia infatti:

$$y = a + bx$$

Da  $\mathbf{x}$  ed  $\mathbf{y}$  si possono ricavare le VA “centrate”, cioè private del valore medio:

$$\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{E}[\mathbf{x}]$$

$$\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{E}[\mathbf{y}] = (a + b\mathbf{x}) - (a + b\mathbf{E}[\mathbf{x}]) = b\tilde{\mathbf{x}}$$

E' immediato verificare che la covarianza é:

$$C_{xy} = \mathbf{E}[\tilde{\mathbf{x}}b\tilde{\mathbf{x}}] = b\mathbf{E}[(\mathbf{x} - \mathbf{E}[\mathbf{x}])^2] = b\sigma_x^2$$

Il termine a denominatore in (5.27) diventa:

$$|b| \sigma_x^2$$

(infatti:  $\sigma^2(b\mathbf{x}) = b^2\sigma_x^2$  e, pertanto  $\sigma_x\sigma_y = \sigma_x\sqrt{b^2\sigma_x^2} = |b| \sigma_x^2$ )

Ne deriva che:

$$R_{xy} = \text{sign}(b) = \pm 1$$

Il caso  $R_{xy} = 0$  può essere ottenuto quando esiste indipendenza stocastica. Infatti, poiché sotto l'ipotesi di indipendenza stocastica é  $\mathbf{E}[\mathbf{xy}] = \mathbf{E}[\mathbf{x}]\mathbf{E}[\mathbf{y}]$ , é possibile provare che  $C_{xy} = 0$  e quindi  $R_{xy} = 0$ . Tuttavia, correlazioni nulle possono essere ottenute anche nel caso di configurazioni di tipo quasi deterministico delle variabili aleatorie. Unica notevole eccezione, se le VA sono gaussiane, allora  $R_{xy} = 0$  implica anche indipendenza stocastica, come può essere facilmente verificato.

Per dimostrare che  $R \in [-1, 1]$  si supponga che siano noti i valori attesi delle VA  $\mathbf{x}$  e  $\mathbf{y}$ . Sia ora  $X_n$  un vettore di  $n$  osservazioni di  $\mathbf{x}$  a cui é stato sottratto  $\mathbf{E}[\mathbf{x}]$ ,  $Y_n$  un analogo vettore per  $\mathbf{y}$ :

$$X_n = (x_1, x_2, \dots, x_n)^T$$

$$Y_n = (y_1, y_2, \dots, y_n)^T$$

Le misure delle variabili aleatorie  $\mathbf{x}$  ed  $\mathbf{y}$  debbono essere state fatte simultaneamente per preservare le informazioni sul comportamento congiunto. E' facile, utilizzando le proprietà del prodotto riga per colonna, mostrare che:

$$\sigma_x^2 = \lim_{n \rightarrow \infty} \frac{1}{n} X_n^T X_n$$

$$\sigma_y^2 = \lim_{n \rightarrow \infty} \frac{1}{n} Y_n^T Y_n$$

$$C_{xy} = \lim_{n \rightarrow \infty} \frac{1}{n} X_n^T Y_n$$

Dunque:

$$R_{xy} = \lim_{n \rightarrow \infty} \frac{X_n^T Y_n}{\sqrt{X_n^T X_n} \sqrt{Y_n^T Y_n}} = \frac{\langle X_n, Y_n \rangle}{\|X_n\| \|Y_n\|}$$

Cioé il coefficiente di correlazione é il limite a cui tende, per  $n$  che tende all'infinito, il coseno dell'angolo formato dai vettori  $X_n$  ed  $Y_n$ . tale coseno deve essere compreso fra -1 e 1, e ciò prova quanto affermato in precedenza. Il coefficiente é massimo in valore assoluto quando noto  $\mathbf{x} = x$  allora  $\mathbf{y} = \pm x$ , nullo quando le due VA sono indipendenti.

### 5.6.1 Trasformazioni lineari

Si consideri:

$$z = ax + by + c$$

Supponendo di conoscere i momenti del primo e secondo ordine di  $\mathbf{x}$  e  $\mathbf{y}$ , si desidera calcolare quelli di  $z$ . Utilizzando la proprietà di linearità di  $\mathbf{E}[\cdot]$ , risulta immediatamente che:

$$\mathbf{E}[z] = a\mathbf{E}[x] + b\mathbf{E}[y] + c$$

Per calcolare la varianza di  $z$  calcoliamo  $\mathbf{E}[z^2]$ . Infatti:

$$\sigma_z^2 = \mathbf{E}[z^2] - \mathbf{E}[z]^2$$

Quindi, poiché:

$$z^2 = a^2x^2 + b^2y^2 + c^2 + 2abxy + 2axc + 2bxc$$

$$\mathbf{E}[z^2] = a^2\mathbf{E}[x^2] + b^2\mathbf{E}[y^2] + c^2 + 2ab\mathbf{E}[x]\mathbf{E}[y] + 2a\mathbf{E}[x]c + 2b\mathbf{E}[y]c$$

applicando la proprietà di linearità di  $\mathbf{E}[\cdot]$  si arriva a scrivere:

$$\sigma_z^2 = \mathbf{E}[z^2] - \mathbf{E}[z]^2 = a^2(\mathbf{E}[x^2] - \mathbf{E}[x]^2) + b^2(\mathbf{E}[y^2] - \mathbf{E}[y]^2) + 2ab(\mathbf{E}[xy] - \mathbf{E}[x]\mathbf{E}[y])$$

che, in sintesi, porge:

$$\sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2abC_{xy}$$

Quando le VA  $\mathbf{x}$  e  $\mathbf{y}$  sono indipendenti, allora  $C_{xy} = 0$  e la varianza della combinazione lineare delle due variabili aleatorie divente, semplicemente:

$$\sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 \quad (5.28)$$

**Esempio 5.6.1** Si calcoli media e varianza della media aritmetica:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k$$

Le variabili aleatorie  $\mathbf{x}_k$  sono identicamente distribuite ed indipendenti con  $\mathbf{E}[\mathbf{x}_k] = \eta \forall k$  e  $\sigma^2(\mathbf{x}_k) = \sigma^2 \forall k$ .

$$\mathbf{E}[\bar{\mathbf{x}}] = \frac{1}{n} \sum_{k=1}^n \mathbf{E}[\mathbf{x}_k] = \eta \quad (5.29)$$

$$\sigma^2(\bar{\mathbf{x}}) = \frac{1}{n} \sum_{k=1}^n \sigma^2(\mathbf{x}_k) = \frac{\sigma^2}{n} \quad (5.30)$$

Si noti che:

$$SNR(\bar{\mathbf{x}}) = \frac{SNR(\mathbf{x})}{\sqrt{n}}$$

pertanto la stima del valore atteso della VA  $\mathbf{x}$ , essendo uguale a quello di  $\bar{\mathbf{x}}$ , può essere fatto in modo accurato utilizzando la media aritmetica. Al crescere della numerosità del campione,  $n$ , la stima diventa via via più accurata.

**Esempio 5.6.2 (Propagazione degli errori di misura)** Sia  $\mathbf{x}$  una VA con valore atteso  $x_0$  e varianza  $\sigma_x^2$ . Analogamente,  $\mathbf{y}$  sia una VA con valore atteso  $y_0$  e varianza  $\sigma_y^2$ . Tali variabili aleatorie possono essere interpretate come le misure di due grandezze fisiche di valore  $x_0$  ed  $y_0$ . Le misure sono realizzate mediante strumenti affetti da errori casuali (incertezza, varianza non nulla) ma non affetti da errore sistematico (il valore atteso coincide con il valore vero). Si calcoli l'incertezza (varianza) della grandezza fisica ottenuta indirettamente mediante la funzione  $g$ :

$$z = g(\mathbf{x}, \mathbf{y})$$

Espandendo  $g$  in serie di Taylor si ottiene ( $P_0 + (x_0, y_0)$ ):

$$\begin{aligned} z &= f(x_0, y_0) + \left( \frac{\partial g(x, y)}{\partial x} \right)_{P=P_0} (\mathbf{x} - x_0) + \left( \frac{\partial g(x, y)}{\partial y} \right)_{P=P_0} (\mathbf{y} - y_0) + \dots \\ &\approx z_0 + \left( \frac{\partial g(x, y)}{\partial x} \right)_{P=P_0} (\mathbf{x} - x_0) + \left( \frac{\partial g(x, y)}{\partial y} \right)_{P=P_0} (\mathbf{y} - y_0) \end{aligned}$$

La varianza di  $z$  è:

$$\sigma_z^2 \approx \left( \frac{\partial g(x, y)}{\partial x} \right)_{P=P_0}^2 \sigma_x^2 + \left( \frac{\partial g(x, y)}{\partial y} \right)_{P=P_0}^2 \sigma_y^2 \quad (5.31)$$

Quando

$$z = \mathbf{x} + \mathbf{y}$$

allora:

$$\left( \frac{\partial g(x, y)}{\partial x} \right)_{P=P_0} = \left( \frac{\partial g(x, y)}{\partial y} \right)_{P=P_0} = 1$$

quindi

$$\sigma_z^2 \approx \sigma_x^2 + \sigma_y^2 \quad (5.32)$$

Analogamente, se

$$z = \mathbf{x} - \mathbf{y}$$

si ottiene:

$$\sigma_z^2 \approx \sigma_x^2 + \sigma_y^2 \quad (5.33)$$

Si consideri ora il prodotto:

$$z = \mathbf{x}\mathbf{y}$$

Le derivate parziali sono:

$$\left( \frac{\partial g(x, y)}{\partial x} \right)_{P=P_0} = y_0$$

$$\left( \frac{\partial g(x, y)}{\partial y} \right)_{P=P_0} = x_0$$

La varianza (normalizzata rispetto a  $z_0^2$ ) é:

$$\frac{\sigma_z^2}{z_0^2} \approx \left( \frac{\sigma_x}{x_0} \right)^2 + \left( \frac{\sigma_y}{y_0} \right)^2 \quad (5.34)$$

Per quanto riguarda il quoziente

$$z = \mathbf{x}/\mathbf{y}$$

Le derivate parziali sono:

$$\left( \frac{\partial g(x, y)}{\partial x} \right)_{P=P_0} = \frac{1}{y_0}$$

$$\left( \frac{\partial g(x, y)}{\partial y} \right)_{P=P_0} = \frac{x_0}{y_0^2}$$

La varianza normalizzata rispetto al valore  $z_0^2$  é identica a quella del prodotto:

$$\frac{\sigma_z^2}{z_0^2} = \frac{\frac{1}{y_0^2} \sigma_x^2 + \frac{x_0^2}{y_0^4} \sigma_y^2}{x_0^2/y_0^2} \approx \left( \frac{\sigma_x}{x_0} \right)^2 + \left( \frac{\sigma_y}{y_0} \right)^2 \quad (5.35)$$

## 5.7 Il teorema del limite centrale

Il teorema del limite centrale (TLC) asserisce che date  $n$  variabili aleatorie indipendenti,  $\mathbf{x}_k$ ,  $k = 1, \dots, n$ , la sommatoria:

$$\mathbf{y} = \sum_{k=1}^n \alpha_k \mathbf{x}_k$$

converge per  $n \rightarrow \infty$  ad una distribuzione normale:

$$\lim_{n \rightarrow \infty} f_Y(y) = \frac{1}{\sigma_y \sqrt{2\pi}} \exp \left( -\frac{1}{2} \left( \frac{y - \eta_y}{\sigma_y} \right)^2 \right)$$

Normalmente il TLC si applica per  $n \geq 30$  anche se, per densità di probabilità delle VA  $\mathbf{x}_k$  particolarmente regolari, allora anche  $n = 5$  può essere un numero per cui l'approssimazione guassiana é accettabile.

I parametri della distribuzione possono essere calcolati utilizzando le espressioni relative al calcolo dei momenti di una VA ottenuta come trasformazione lineare di due VA, presentate in precedenza. In particolare, il valore atteso sarà

$$\eta_y = \sum_{k=1}^n \alpha_k \mathbf{E}[\mathbf{x}_k]$$

mentre la varianza, tenendo conto che le VA sono indipendenti quindi vale (5.28), é data da:

$$\sigma_y^2 = \sum_{k=1}^n \alpha_k^2 \sigma_{x_k}^2$$

Il teorema del limite centrale ha fondamentale importanza nella modellazione probabilistica di sistemi stocastici. Infatti, ovunque una grandezza misurata possa essere pensata come la somma di piccoli contributi, numerosi ed indipendenti, allora si può ipotizzare ragionevolmente che questa grandezza segua la distribuzione gaussiana. Gli esempi di applicazione di tale principio sono numerosi. Ad esempio: (1) la deviazione di un parametro di un dispositivo prodotto su scala industriale, effetto cumulativo di tante piccole imperfezioni nella catena produttiva, (2) la potenza in un nodo di una rete elettrica, somma delle potenze domandate dai singoli utenti che agiscono indipendentemente l'uno dall'altro, (3) le deviazioni rispetto al valore medio del valore di un'azienda quotata in borsa, ecc..

E' importante, tuttavia, non abusare del teorema del limite centrale pretendendo che qualsiasi VA debba seguire una distribuzione gaussiana. Esistono numerosi controesempi. Ad esempio, i tempi al guasto di un dispositivo riparabile sono spesso descritti da una distribuzione esponenziale, i tempi di scarica di un materiale isolante solido seguono una distribuzione di Weibull, per un materiale gassoso la distribuzione lognormale, ecc: è solo dalla somma di VA indipendenti che tende a generarsi una distribuzione normale. Inoltre, particolari variabili aleatorie hanno somma che, anche nelle condizioni del teorema del limite centrale, non si distribuisce secondo la legge gaussiana.

Un interessante corollario del teorema del limite centrale è il seguente. Sia:

$$\mathbf{y} = \prod_{k=1}^n \alpha_k \mathbf{x}_k$$

(le VA  $\mathbf{x}_k$  siano indipendenti). Poiché:

$$\log(\mathbf{y}) = \sum_{k=1}^n \log(\alpha_k \mathbf{x}_k)$$

tende, per il teorema del limite centrale, ad una distribuzione normale di parametri  $\eta$  e  $\sigma$ , allora la VA  $\mathbf{y}$  tende alla corrispondente distribuzione lognormale di parametri  $\eta$  e  $\sigma$ . Questo corollario è noto come teorema del limite centrale per i prodotti.

**Esempio 5.7.1 (Distribuzione della media aritmetica)** Sia  $\mathbf{x}$  una VA di valore atteso  $\eta$  e deviazione standard  $\sigma$ . Calcolare la probabilità di trovare la media aritmetica di  $\mathbf{x}$  calcolata su  $n$  campioni in un intervallo  $[\eta - k\sigma, \eta + k\sigma]$  centrato attorno al valore atteso di  $\mathbf{x}$ .

Si premette che l'osservazione di  $n$  realizzazioni di  $\mathbf{x}$  può essere pensata come un esperimento multiplo consistente di  $n$  singoli esperimenti consecutivi, in ogni esperimento si osserva una realizzazione di  $\mathbf{x}$ . Alternativamente, si può pensare ad  $n$  esperimenti simultanei. In ognuno di questi esperimenti si ottiene una realizzazione di una VA  $\mathbf{x}_i$  ( $i = 1, \dots, n$ ). Le variabili aleatorie  $\mathbf{x}_i$  sono stocasticamente indipendenti e seguono tutte la legge di probabilità di  $\mathbf{x}$ . In questo esempio sarà utilizzata questa seconda interpretazione.

E' stato mostrato (si veda l'esempio 5.6.1) che la media aritmetica di  $\mathbf{x}$  calcolata su un campione di numerosità  $n$  è una VA di valore atteso  $\eta$  e deviazione standard  $\sigma/\sqrt{n}$ . Pertanto, ipotizzando che si possa applicare il teorema del limite centrale, la media aritmetica, essendo la somma di  $n$  VA indipendenti, tende a distribuirsi secondo una legge normale. I parametri di tale legge sono, ovviamente,  $\eta$  e  $\sigma/\sqrt{n}$ . La probabilità di trovare la media aritmetica nell'intervallo specificato è data da:

$$P(k, n) = \Phi(k\sqrt{n}) - \Phi(-k\sqrt{n})$$

Alcune curve di  $P(k, n)$  sono riportate in figura 5.5.

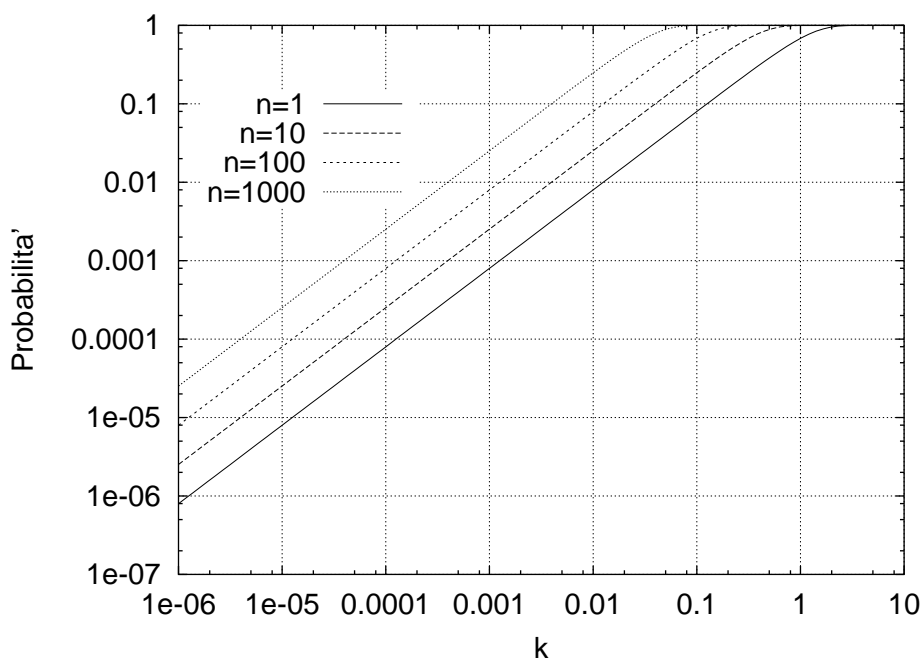


Figura 5.5: Grafico di  $P(k, n)$  relativo all'esempio 5.7.1. Il significato di  $k$  è legato all'ampiezza dell'intervallo:  $[\eta - k\sigma, \eta + k\sigma]$ .

## 5.8 Valore atteso e varianza condizionati

Il valore atteso condizionato di una VA è definito in modo analogo al valore atteso, sfruttando la densità di probabilità condizionata:

$$\mathbf{E}[\mathbf{y} | \mathcal{F}] = \int_{-\infty}^{\infty} y f(y | \mathcal{F}) dy$$

Se  $\mathcal{F}$  è l'evento  $\mathbf{x} \in [x, x + dx]$ , allora si può definire il valore atteso condizionato come:

$$\eta_{y|x} = \mathbf{E}[\mathbf{y} | x] = \int_{-\infty}^{\infty} y f(y | x) dy \quad (5.36)$$

ed in modo analogo la varianza condizionata:

$$\sigma_{y|x}^2 = \mathbf{E}[(\mathbf{y} - \eta_{y|x})^2 | x] = \int_{-\infty}^{\infty} (y - \eta_{y|x})^2 f(y | x) dy \quad (5.37)$$

La funzione:

$$\phi(x) = \mathbf{E}[\mathbf{y} | x]$$

è nota come *funzione di regressione* di  $\mathbf{y}$  su  $\mathbf{x}$ . Il nome di regressione deriva dagli studi del genetista Galton che analizzò la relazione fra altezza dei genitori ed altezza dei figli. Si veda l'esempio 5.8.1 per i dettagli.

Per spiegare il significato del valore atteso condizionato si pensi a quanto detto per il valore atteso. Sia infatti  $\eta$  un previsore costante per  $\mathbf{y}$ . Scelta come funzione costo la somma quadratica degli errori di

predizione, il valore atteso di  $\mathbf{y}$  é quel valore di  $\eta$  che minimizza il valore atteso del costo di predizione  $W = \mathbf{E}[(\mathbf{y} - \eta)^2]$ . Infatti:

$$W = \int_{-\infty}^{\infty} (y - \eta)^2 f(y) dy \quad (5.38)$$

Per calcolare il minimo di  $W$  si derivi rispetto a  $\eta$ :

$$\frac{dW}{d\eta} = \int_{-\infty}^{\infty} 2(y - \eta)f(y)dy = 0$$

Quindi:

$$\eta = \int_{-\infty}^{\infty} yf(y)dy \quad (5.39)$$

cióé il migliore stimatore di  $\mathbf{y}$  é esattamente il valore atteso di  $\mathbf{y}$ .

Si supponga ora di volere stimare  $\mathbf{y}$  mediante una funzione della VA  $\mathbf{x}$ :  $\eta(\mathbf{x})$ . Per fare ciò si scriva la funzione errore di predizione quadratico come nel modo precedente:

$$W = \mathbf{E}[(\mathbf{y} - \eta(\mathbf{x}))^2]$$

dove l'espressione mostra che  $W$  é una funzione di due variabili aleatorie. Per calcolarne il valore medio bisogna procedere in modo analogo a quanto fatto per scrivere le equazioni (5.38) e (5.39). In questo caso, poiché si é in presenza di due variabili aleatorie, allora é necessario utilizzare (5.9):

$$W = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \eta(x))^2 f(x, y) dx dy$$

La densità congiunta può essere riscritta utilizzando la probabilità condizionata ed evidenziando i termini che dipendono da  $x$  solamente come:

$$W = \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} (y - \eta(x))^2 f(y | x) dy dx$$

Gli integrandi in quest'ultima equazione sono entrambi positivi, quindi  $W$  é minima se l'integrale

$$\int_{-\infty}^{\infty} (y - \eta(x))^2 f(y | x) dy$$

é minimo per tutti i valori possibili di  $x$ . Per quanto mostrato in (5.39) e tenendo conto che  $x$  é una costante nell'integrale precedente, ne deriva che l'integrale precedente é minimizzato se:

$$\eta(x) = \mathbf{E}[\mathbf{y} | x]$$

Quindi, data una  $\mathbf{y}$  ed una  $\mathbf{x}$  la funzione di regressione  $\phi(x)$  costituisce il migliore predittore di  $\mathbf{y}$  noto il fatto che la VA  $\mathbf{x}$  ha assunto il valore  $x$ .

**Esempio 5.8.1 (Regressione di Galton)** E' stato mostrato che per VA congiuntamente normali il valore atteso di  $\mathbf{y}$ , condizionato dalla conoscenza di  $\mathbf{x} = x$  é dato da:

$$\eta_{y|x} = \eta_y + r(x - \eta_x) \frac{\sigma_y}{\sigma_x}$$

Il genetista Galton ipotizzó che l'altezza media dei genitori (VA  $\mathbf{x}$ ) e quella dei figli (VA  $\mathbf{y}$ ) potessero essere modellate come due VA congiuntamente gaussiane con  $\eta_x = \eta_y$ ,  $\sigma_x = \sigma_y$  e coefficiente di correlazione  $r$



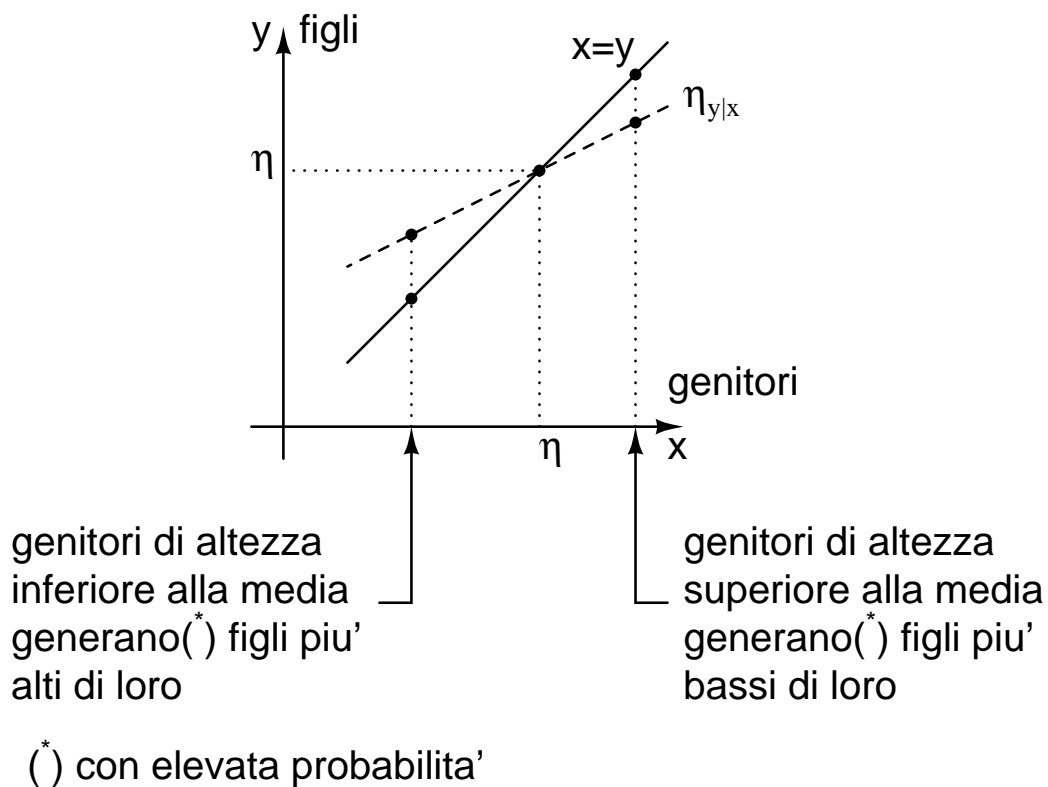


Figura 5.6: Miglior stima dell'altezza dei figli nota l'altezza media dei genitori

positivo. Pertanto, l'altezza attesa dei figli condizionata dalla conoscenza dell'altezza dei genitori sono esprimibili come:

$$\eta_{y|x} = \eta + r(x - \eta)$$

La funzione  $\eta_{y|x}$  è una retta passante per il punto  $(\eta, \eta)$  ed avente coefficiente angolare pari ad  $r$ , quindi minore di 1. Come mostra la figura 5.6 ciò porta ad affermare che genitori di altezza inferiore alla media genereranno (con alta probabilità) figli più alti di loro. Al contrario, genitori più alti della media genereranno (con alta probabilità) figli più bassi di loro.