

Politecnico di Torino
II Facoltà di Ingegneria - sede di Vercelli

Dispense per il corso di Metodi Matematici e Statistici

Per i corsi di laurea in Ingegneria Elettronica ed Informatica

Andrea Gamba Mariagrazia Graziano

27 maggio 2003

Copyright

Copyright (c) 2003 Andrea Gamba Mariagrazia Graziano.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.1 or any later version published by the Free Software Foundation; with the Invariant Sections being Copyright, with the Front-Cover Texts being *Dispense per il corso di Metodi Matematici e Statistici per i corsi di Laurea in Ingegneria Elettronica e Informatica*, and with no Back-Cover Texts. A copy of the license is included in the section entitled GNU Free Documentation License.

Indice

Copyright	i
Prefazione	iii
1 Variabili aleatorie e distribuzione di probabilità	1
1.1 Introduzione al concetto di distribuzione di probabilità	1
1.2 Valore atteso e deviazione standard	3
1.3 La delta di Dirac	8
1.4 Uso della delta di Dirac nella teoria delle probabilità	14
2 Funzione generatrice dei momenti e cammino aleatorio	16
2.1 Funzione generatrice dei momenti e trasformata di Laplace	16
2.2 Cumulanti e funzione generatrice dei cumulanti	18
2.3 Cammino aleatorio	19
3 La trasformata di Laplace	21
3.1 Le proprietà della Trasformata di Laplace	21
3.2 Cenni alla trasformata di Fourier	21
3.3 Il teorema dei residui	21
3.4 L'antitrasformata	21
4 GNU Free Documentation License	22

Prefazione

Gli appunti riportati nel seguito costituiscono una traccia ad uso degli studenti del corso di *Metodi matematici e statistici* dei corsi di Laurea in Ingegneria Elettronica ed Informatica della seconda facoltà del Politecnico di Torino. Il programma di tale corso offre argomenti che lo studente deve necessariamente affrontare in previsione di corsi successivi del suo piano di studi (come per esempio Teoria dei segnali ed Elettrotecnica). Poiché gli argomenti trattati in questo corso sono organizzati in modo differente rispetto ai corsi tradizionali (ci riferiamo qui ai corsi di Calcolo delle probabilità e di Complementi di matematica) si è voluto fornire uno spunto di materiale che sarebbe reperibile su testi diversi e che non risulterebbe in tal caso legato da un filo logico come invece si ritiene debbano presentarsi gli argomenti di un corso di laurea. La trattazione dunque non pretende di essere esaustiva nè formale ma si spera che possa comunque essere di aiuto allo studente del corso.

CAPITOLO 1

Variabili aleatorie e distribuzione di probabilità

1.1 Introduzione al concetto di distribuzione di probabilità

Esempio Supponiamo di fare un esperimento di questo tipo: prendo una vasca che contiene un fluido turbolento e metto una sonda a filo caldo che misura la velocità del flusso nella direzione trasversale al filo (principio di funzionamento: il flusso raffredda il filo e la sua resistenza cambia; vedi fig. 1.1).

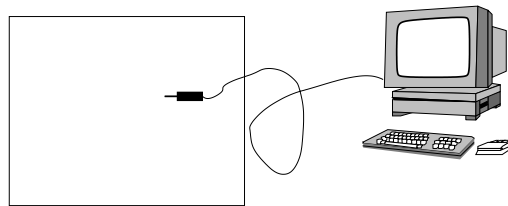


Figura 1.1: Misura della velocità in un fluido turbolento.

Misuro la componente di velocità v trasversale al filo.

Per costruire un istogramma divido l'asse delle velocità in tanti intervallini, o *classi* $I_i = (v_i, v_i + \Delta v)$ di ampiezza Δv fissata. Poi faccio N misure (p.es. una ogni secondo) e registro quante volte n_i la misura cade nell'intervallo $(v_i, v_i + \Delta v)$, come rappresentato in figura 1.2.

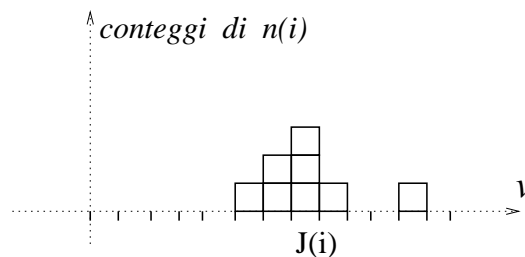


Figura 1.2: Costruzione dell'istogramma di misura.

Normalizzo tutto dividendo per N e trovo delle *frequenze*

$$f_i = \frac{n_i}{N}$$

(fig. 1.3) le quali hanno la seguente proprietà

$$\sum_i f_i = 1$$

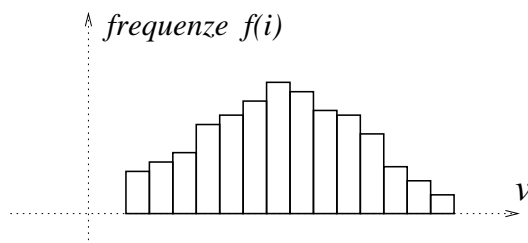


Figura 1.3: costruzione dell'istogramma di misura.

Se ora eseguo moltissime misure e scelgo l'ampiezza Δv degli intervallini molto piccola, accade comunemente che le frequenze *discrete* f_i siano ben approssimate da una funzione continua $f(v_i)$ (si veda figura 1.4).

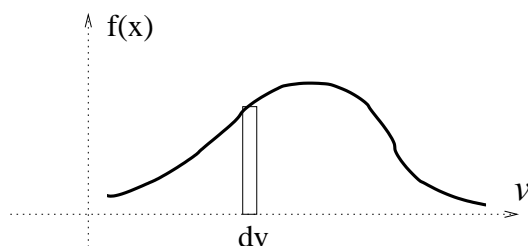


Figura 1.4: costruzione dell'istogramma di misura.

Precisamente, quando gli intervallini sono molto piccoli è chiaro che se li dimezzo anche la frequenza viene dimezzata, quindi

$$f_i \simeq f(v_i) \Delta v \quad (1.1)$$

Nel limite di moltissime misure e intervallini piccolissimi l'istogramma diventa una curva continua $f(v)$, la *distribuzione di probabilità*, o *densità di probabilità* (inglese: *probability distribution function*, abbreviata *pdf*).

Significato probabilistico: la probabilità di osservare una fluttuazione di velocità compresa tra i valori a e b viene identificata (nel limite ideale $\Delta v \rightarrow 0$, $N \rightarrow \infty$) con

$$P(a < v < b) = \sum_{a < v_i < b} f_i = \sum_{a < v_i < b} f(v_i) \Delta v \simeq \int_a^b f(v) dv = F(b) - F(a) \quad (1.2)$$

dove

$$F(v) = \int_{-\infty}^v f(v') dv'$$

è detta *distribuzione di probabilità cumulata* o *funzione di ripartizione*. Dalla definizione stessa (1.1) invece segue che

$$P(v_0 < v < v_0 + \Delta v) \simeq f(v) \Delta v \quad (1.3)$$

dal che si capisce perché la f viene detta *densità* di probabilità: perché non è una probabilità, ma lo diventa se la moltiplico per Δv . Conviene tenere presente le formule (1.2) e (1.3) perchè risultano utili nella soluzione di molti problemi.

Esempio con Matlab Facciamo un esempio di come si procederebbe con Matlab. Creiamo noi a mano una distribuzione di probabilità f gaussiana centrata in $m=10$ e avente deviazione standard $s=2$ (nella realtà la f potrebbe essere il risultato di un esperimento). Siccome il calcolatore non gestisce realmente funzioni continue scegliamo un'ampiezza finita $dv=.1$ delle classi:

```
m=10
s=2
dv=.1
v=-30:dv:30
f=exp(-(v-m).^2/(2*s^2))/sqrt(2*pi*s^2)
```

vediamo come è fatta la distribuzione di probabilità:

```
plot(v,f)
```

naturalmente se ridefiniamo m o s troviamo una d.pr. diversa:

```
m=-10
f=exp(-(v-m).^2/(2*s^2))/sqrt(2*pi*s^2)
plot(v,f)
s=4
f=exp(-(v-m).^2/(2*s^2))/sqrt(2*pi*s^2)
plot(v,f)
```

verifichiamo la correttezza della normalizzazione calcolando numericamente un'approssimazione di $\int_{-\infty}^{\infty} f(v) dv$:

```
sum(f*dv)
```

1.2 Valore atteso e deviazione standard

Ora viene naturale porsi domande come:

- “qual è la velocità media del fluido turbolento?”
- “qual è l'ampiezza delle fluttuazioni intorno al valor medio?”

Per quanto riguarda la velocità media, si capisce che dobbiamo calcolare una media pesata del tipo

$$\text{vel. media} \simeq \frac{1}{N} \sum_i n_i v_i = \sum_i f_i v_i \simeq \sum_i v_i f(v_i) \Delta v_i \simeq \int_0^{\infty} v f(v) dv$$

cioè ogni valore approssimato v_i lo contiamo il numero di volte che è stato osservato e normalizziamo tutto col numero totale di misure. La cosa interessante è che come conseguenza abbiamo una semplicissima espressione per calcolare il valore atteso. Infatti, immaginiamo che il nostro esperimento ripetuto molte volte ci abbia fornito una buona approssimazione della distribuzione di probabilità $f(v)$, per esempio sotto forma di un vettore di numeri f_i che sta nel nostro calcolatore. Allora per calcolare la velocità media basta usare l'espressione

$$\int v f(v) dv \tag{1.4}$$

ovvero in pratica si moltiplica la $f(v)$ per v e si integra.

Con Matlab dovremmo semplicemente scrivere

```
sum(v.*f*dv)
```

Valore atteso Nella letteratura matematica si preferisce parlare di *valore atteso* della velocità (o *valore di aspettazione*) piuttosto che di valore medio, per sottolineare la differenza che corre tra il valore medio effettivamente calcolato su un numero N di misure facendo la media pesata

$$\bar{v} = \frac{1}{N} \sum n_i v_i$$

(*media campionaria*) e quello “ideale” calcolato a partire da una data distribuzione di probabilità “matematica” $f(v)$ per mezzo della formula (1.4). Si usa anche appositamente un simbolo diverso da \bar{v} :

$$E(V) = \int v f(v) dv \quad (1.5)$$

dove la E sta per “expectation” (ingl.), “*é*xpectation” (fr.), “*ex*pectatio” (lat.) = attesa, aspettazione.

Si intende dire che se i valori casuali di una grandezza osservata V sono distribuiti secondo la curva di frequenze $f(v)$ ed effettuiamo un gran numero di misure, *possiamo ragionevolmente aspettarci* che il valor medio osservato \bar{v} sia un’approssimazione (tanto migliore quanto più grande è il numero di misure) del valore atteso ideale $E(V)$.

Nel loro amore per la precisione i matematici scelgono anche di indicare in modo diverso la v a seconda che si intenda la *variabile aleatoria* osservata V o i valori concreti v che può assumere.

Riassumendo: si dice *variabile aleatoria continua* V descritta dalla distribuzione di probabilità continua $f(v)$ una grandezza osservabile i cui valori v vengono osservati con probabilità

$$P(v < V < v + \Delta v) \simeq f(v) \Delta v$$

Il *valore atteso* della variabile aleatoria V è dato dalla formula (1.5).

Esempio: la distribuzione gaussiana Nell’esempio della turbolenza sviluppata (alti numeri di Reynolds) si trova che le componenti trasverse della velocità hanno con buona approssimazione una distribuzione gaussiana

$$f(v) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(v-v_0)^2}{2\sigma^2}\right] = \frac{1}{\sigma} \varphi\left(\frac{v-v_0}{\sigma}\right) \quad (1.6)$$

Dove con $\varphi(x)$ intendiamo la gaussiana standard

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

opportunamente normalizzata in modo che si abbia

$$\int_{-\infty}^{+\infty} \varphi(x) dx = 1$$

Ricordiamo anche la notazione

$$\Phi(x) = \int_{-\infty}^x \varphi(x') dx'$$

per la distribuzione cumulata.

Ricordiamo ora che v_0 nell’espressione 1.6 è il massimo di una distribuzione simmetrica, dunque ci aspettiamo che v_0 sia il valore atteso definito in (1.5). Questa è anche l’indicazione che ci viene dal

calcolo fatto con Matlab.

Verifichiamo questa affermazione formalmente, cioè con un calcolo algebrico:

$$\begin{aligned}
 E(V) &= \int_{-\infty}^{\infty} v f(v) \, dv \\
 &= \int_{-\infty}^{\infty} v \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(v-v_0)^2}{2\sigma^2}\right] \, dv \\
 &= \int_{-\infty}^{\infty} (v_0 + \sigma u) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{u^2}{2}\right] \sigma \, du \\
 &= v_0 \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] \, du}_{=1} + \sigma \underbrace{\int_{-\infty}^{\infty} u \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] \, du}_{=0} \\
 &= v_0
 \end{aligned}$$

dove abbiamo usato nella seconda riga la sostituzione

$$u = \frac{v - v_0}{\sigma}, \quad v = v_0 + \sigma u, \quad du = \frac{dv}{\sigma}$$

Il primo integrale è 1 perché la distribuzione di probabilità è normalizzata, il secondo è 0 perché integrale di una funzione dispari (fig. 1.5).

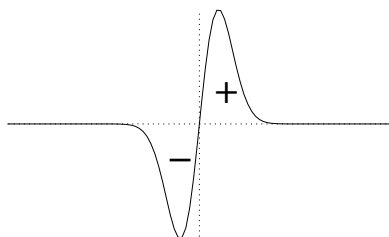


Figura 1.5: Integrale di una funzione dispari.

Il conto si poteva fare anche senza scrivere esplicitamente la $f(v)$, ma usando solo cambiamenti di variabili e le proprietà della gaussiana standard φ , infatti:

$$\begin{aligned}
 &\int_{-\infty}^{+\infty} v \frac{1}{\sigma} \varphi\left(\frac{v-v_0}{\sigma}\right) \, dv \\
 &= \int_{-\infty}^{+\infty} (v_0 + \sigma u) \frac{1}{\sigma} \varphi(u) \, du \\
 &= v_0 \underbrace{\int_{-\infty}^{+\infty} \varphi(u) \, du}_{=1} + \sigma \underbrace{\int_{-\infty}^{+\infty} u \varphi(u) \, du}_{=0}
 \end{aligned}$$

Deviazione standard La σ che compare nell'espressione della funzione gaussiana dovrebbe avere il significato di dispersione attorno al valor medio. Come possiamo definire in generale (cioè per una distribuzione di probabilità qualsiasi) l'idea di dispersione attorno al valor medio?

Potremmo scegliere di calcolare lo scarto medio dal valore atteso:

$$\begin{aligned}
 & \frac{1}{N} \sum_i n_i [v_i - E(V)] \\
 &= \sum_i f_i [v_i - E(V)] \\
 &\simeq \int [v - E(V)] f(v) \, dv \\
 &= \underbrace{\int v f(v) \, dv}_{=E(V)} - E(V) \underbrace{\int f(v) \, dv}_{=1} \\
 &= 0
 \end{aligned}$$

Proviamo anche con Matlab nel caso della distribuzione gaussiana:

```
sum((v-m).*f*dv)
```

Il problema è che la somma che sto considerando ha termini negativi che bilanciano esattamente i termini positivi, quindi la prima idea per calcolare la dispersione attorno al valor medio fallisce.

Un'idea alternativa sarebbe di calcolare

$$\frac{1}{N} \sum_i n_i |v_i - E(V)|$$

e questa fornirebbe senz'altro una misura non negativa della dispersione, che darebbe 0 se tutti i valori v_i fossero esattamente uguali a $E(V)$ (e in questo caso V non sarebbe aleatoria) e un valore positivo tanto più grande quanto più i valori v osservati fossero dispersi attorno a $E(V)$.

Tuttavia il valore assoluto ha proprietà analitiche "brutte" (per esempio non si può derivare). Una scelta migliore è calcolare lo *scarto quadratico medio* o *varianza*:

$$\begin{aligned}
 & \frac{1}{N} \sum_i n_i (v_i - E(V))^2 \\
 &= \sum_i f_i (v_i - E(V))^2 \\
 &\simeq \int (v - E(V))^2 f(v) \, dv
 \end{aligned}$$

che è sicuramente positiva (nulla se la variabile non è aleatoria) e dà comunque una misura della dispersione dei valori v_i .

Più formalmente, possiamo definire

$$\text{Var}(V) = E \left[(V - E(V))^2 \right]$$

Non è difficile vedere che si può anche scrivere

$$\text{Var}(V) = E(V^2) - E(V)^2 \tag{1.7}$$

infatti:

$$\begin{aligned}
 \int (v - E(V))^2 f(v) \, dv &= \int (v^2 - 2E(V)v + E(V)^2) f(v) \, dv \\
 &= E(V^2) - 2E(V) \underbrace{\int v f(v) \, dv}_{=E(V)} + E(V)^2 \underbrace{\int f(v) \, dv}_{=1}
 \end{aligned}$$

Formule di questo tipo discendono dal fatto che il valore atteso è sostanzialmente una somma (o un integrale) e come tale è lineare: per esempio, l'integrale della somma diventa la somma degli integrali, e i fattori costanti come $E(V)$ vengono portati fuori dall'integrale.

Un'altra osservazione è la seguente: in un primo tempo abbiamo calcolato il valore atteso di V ; ora, per calcolare la dispersione attorno al valore atteso ci troviamo a dover calcolare il valore atteso V^2 ; non è difficile capire che è possibile calcolare il valore atteso di qualsiasi funzione di $g(V)$, per esempio $g_n(V) = V^n$:

$$\frac{1}{N} \sum_i n_i g(v_i) = \sum_i f_i g(v_i) \simeq \int g(v) f(v) dv \quad (1.8)$$

di qui la "ricetta": se voglio calcolare il valore atteso di una qualsiasi funzione $g(V)$ di una variabile aleatoria V basta calcolare

$$E(g(V)) = \int g(v) f(v) dv \quad (1.9)$$

In particolare se $g_n(V) = V^n$ i numeri che si ottengono si chiamano *momenti* della distribuzione di probabilità f :

$$\begin{aligned} E(V) &= \int v f(v) dv \\ E(V^2) &= \int v^2 f(v) dv \\ E(V^3) &= \int v^3 f(v) dv \\ \dots &\dots \dots \end{aligned}$$

Esempio A che cosa serve calcolare il valore atteso di *funzioni* di una variabile aleatoria? Consideriamo l'esempio del flusso turbolento visto prima. L'energia cinetica di un volumetto di fluido di massa m è

$$T = T(V) = m \frac{v^2}{2} = m \frac{v_x^2 + v_y^2 + v_z^2}{2}$$

Se la turbolenza è isotropa, ci aspettiamo che la componente v_x sia una variabile aleatoria che ha esattamente la stessa distribuzione di probabilità di qualsiasi altra componente della velocità, come per esempio della componente trasversa v_{\perp} che misuriamo con la sonda.

Allora, se vogliamo conoscere l'energia cinetica media di un volumetto di fluido il calcolo che dobbiamo fare è

$$\begin{aligned} E(T(V)) &= E\left(m \frac{V_x^2 + V_y^2 + V_z^2}{2}\right) \\ &= \frac{m}{2} E(V_x^2) + \frac{m}{2} E(V_y^2) + \frac{m}{2} E(V_z^2) \\ &= \frac{3}{2} m E(V_{\perp}^2) \\ &= \frac{3}{2} m \int v^2 f(v) dv \end{aligned}$$

dunque l'energia cinetica media del volumetto è proporzionale al secondo momento della distribuzione di probabilità $f(v)$.

Nel caso della gaussiana, per calcolarne la varianza, dovremmo calcolare $E(V^2)$, ovvero il momento secondo, e dovremmo quindi utilizzarlo nella relazione 1.7.

La risoluzione degli integrali per il calcolo dei momenti per funzioni di distribuzione qualsiasi può essere complessa, soprattutto se si vogliono calcolare i momenti di ordine successivo. Esiste uno strumento che utile dal punto di vista pratico per il calcolo dei momenti, che tuttavia tratteremo

più avanti (si veda il capitolo 2 a pagina 16). Nel seguito del capitolo introdurremo ancora alcune proprietà e funzioni particolari che completano questa sezione e che potremo utilizzare anche nelle sezioni successive.

Soffermiamoci ora sulla formula (1.9). Come abbiamo visto nell'esempio, molti calcoli algebrici si svolgono semplicemente tenendo conto di alcune proprietà generali del valore atteso, come la linearità. In effetti si è detto che il valore atteso altro non è, filosoficamente, che una somma (o un integrale), e quindi ha le stesse proprietà dell'integrale, in primo luogo la linearità:

$$\begin{aligned} E(\alpha g(V) + \beta h(V)) &= \int (\alpha g(v) + \beta h(v)) f(v) dv \\ &= \alpha \int g(v) f(v) dv + \beta \int h(v) f(v) dv \\ &= \alpha E(g(V)) + \beta E(h(V)) \end{aligned} \quad (1.10)$$

In più, il fatto che $f(v)$ è una distribuzione di probabilità assicura anche che

$$E(1) = \int f(v) dv = 1$$

L'espressione (1.7) per la varianza, ad esempio, si può ricavare usando solo queste proprietà algebriche del valore atteso:

$$\begin{aligned} \text{Var}(V) &= E[(V - E(V))^2] \\ &= E[V^2 - 2E(V)V + E(V)^2] \\ &= E(V^2) - 2E(V)E(V) + E(V^2)E(1) \\ &= E(V^2) - E(V)^2 \end{aligned}$$

1.3 La delta di Dirac

Ritorniamo ancora alla formula (1.9):

$$E(g(V)) = \int g(v) f(v) dv$$

la versione discreta sarebbe

$$\sum_i g_i f_i \Delta v = \Delta v \cdot \sum_i g_i f_i$$

a parte il fattore Δv questa espressione ha la forma di un prodotto scalare tra i due vettori g_i e f_i , entrambi con moltissime componenti. In effetti questo è precisamente il modo come calcolerei il valore atteso $E(g(V))$ usando Matlab:

```
g=v.^2
sum(g.*f*dv)
g*f'*dv
```

Per questi motivi diciamo che l'espressione

$$(g, f) = \int g(v) f(v) dv$$

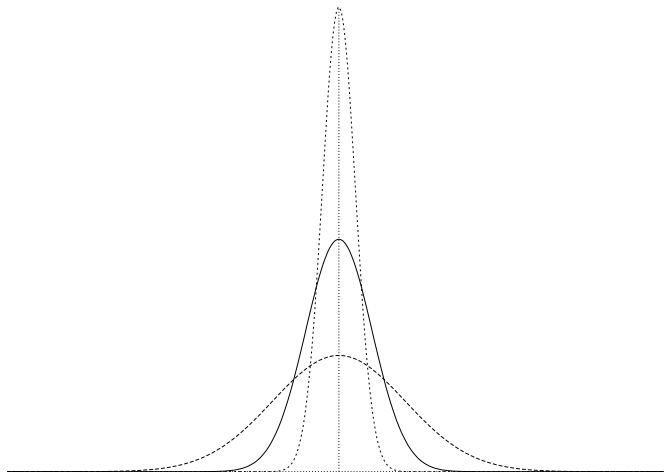


Figura 1.6: Distribuzioni gaussiane centrate.

è un *prodotto scalare* su uno spazio di funzioni. Dunque il valore atteso di $g(V)$, dove g è una funzione di una variabile aleatoria V distribuita secondo la legge $f(v)$, o semplicemente il prodotto scalare delle funzioni g ed f .

Se decidiamo di tenere fissa f , allora per ogni funzione $g(V)$ posso calcolare il valore atteso

$$E(g(V)) = \int g(v) f(v) dv$$

quindi ho un sistema per associare un numero ad ogni funzione $g(v)$:

$$g(v) \mapsto E(g(V)) = \int g(v) f(v) dv \quad (1.11)$$

ho in un certo senso una “funzione di funzione”. I matematici per brevità chiamano *funzionali* queste “funzioni di funzione”

La cosa interessante è che ogni scelta di una distribuzione di probabilità diversa f fornisce un funzionale E_f diverso, dove metto l'indice f per intendere che calcolo il valore atteso rispetto alla distribuzione di probabilità f :

$$E_f(g(V)) = \int g(v) f(v) dv$$

quindi la distribuzione f può essere vista in due modi diversi:

- come funzione $f(v)$
- come funzionale $g(v) \mapsto E_f(g(V))$

Perché queste osservazioni sono importanti? Perché i matematici hanno scoperto che lo spazio dei funzionali contiene quello delle funzioni (ad ogni $f(v)$ corrisponde un funzionale $g \mapsto E_f(g(V))$), ma che lo spazio dei funzionali è *più grande*.

Facciamo un esempio. Consideriamo le distribuzioni gaussiane centrate (fig. 1.6)

$$\varphi_\sigma(v) \equiv \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{v^2}{2\sigma^2}\right]$$

Man mano che σ viene scelta più piccola la curva diventa sempre più stretta ($\sim \sigma$) e alta ($\sim 1/\sigma$), in modo che l'area sottostante risulti sempre invariata e uguale a 1.

Quanto più la distribuzione è stretta, tanto meno fluttuazioni ci sono attorno al valor medio, cioè la variabile V è sempre meno aleatoria. Nel limite in cui $\sigma \rightarrow 0$ diremo che la variabile V è *deterministica* e che non ci sono fluttuazioni

Tuttavia, il problema è che il limite delle funzioni φ_σ per $\sigma \rightarrow 0$ *non è una funzione*:

$$\lim_{\sigma \rightarrow 0} \varphi_\sigma(v) = \begin{cases} 0 & \text{se } v \neq 0 \\ +\infty & \text{altrimenti} \end{cases}$$

Il fatto che il limite non esista ci disturba, perchè fisicamente ci interessa considerare il caso di variabili deterministiche, i cui valori non possano fluttuare. Ci interessa anche considerare il caso di variabili aleatorie che possono assumere solo certi valori interi, come per esempio accade con il punteggio che si ottiene lanciando un dado. In entrambi questi casi vorremmo poter utilizzare il $\lim_{\sigma \rightarrow 0} \varphi_\sigma(v)$ per descrivere variabili aleatorie la cui distribuzione di probabilità è concentrata al di sopra di un punto o di un certo numero di punti discreti.

La chiave per uscire da questo problema è che il limite $\lim_{\sigma \rightarrow 0} \varphi_\sigma(v)$ non esiste se consideriamo le φ_σ come funzioni, ma esiste se le consideriamo come funzionali.

Infatti, se $g(v)$ è una qualsiasi funzione continua limitata, è intuitivo che

$$\lim_{\sigma \rightarrow 0} E_{\varphi_\sigma}(g(V)) = g(0)$$

perchè la funzione $\varphi_\sigma(v)$ è tutta concentrata nell'intorno di $v = 0$.

Dimostriamo questa proprietà rigorosamente. $g(v)$ limitata vuol dire che $|g(v)| \leq \text{Sup } |g(v)| = M < \infty$. Allora

$$\begin{aligned} E_{\varphi_\sigma}(g(V)) &= \int_{-\infty}^{+\infty} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv \\ &= \int_{-\infty}^{-1} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv + \int_{-1}^{+1} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv + \int_{+1}^{+\infty} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv \end{aligned}$$

Ora i due integrali estremi tendono a 0 e quello intermedio tende a $g(0)$:

$$\begin{aligned} \left| \int_{-\infty}^{-1} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv \right| &\leq \int_{-\infty}^{-1} |g(v)| \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv \\ &\leq M \int_{-\infty}^{-1} \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv = M \int_{-\infty}^{-1/\sigma} \varphi(u) du \\ &= M \Phi\left(-\frac{1}{\sigma}\right) \rightarrow 0 \text{ per } \sigma \rightarrow 0 \end{aligned}$$

perchè $\Phi(-\infty) = 0$. Analogamente per l'ultimo integrale. Quello intermedio invece

$$\begin{aligned} \int_{-1}^{+1} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv &= \int_{-1}^{+1} g(v) \frac{1}{\sigma} \varphi\left(\frac{v}{\sigma}\right) dv = \int_{-1/\sigma}^{1/\sigma} g(\sigma u) \varphi(u) du \\ \rightarrow \int_{-\infty}^{+\infty} g(0) \varphi(u) du &= g(0) \int_{-\infty}^{+\infty} \varphi(u) du = g(0) \text{ per } \sigma \rightarrow 0 \end{aligned}$$

Il limite finale nella prima riga si può dimostrare rigorosamente ricorrendo al teorema di convergenza limitata di Lebesgue¹.

¹Teorema di convergenza limitata: se

$$|f_n(x)| \leq g(x) \quad \text{e} \quad \int g(x) dx < \infty$$

Riassumendo: se identifico le funzioni gaussiane $\varphi_\sigma(v)$ con i corrispondenti funzionali

$$g(v) \mapsto \int g(v) \varphi_\sigma(v) dv$$

allora il limite

$$\lim_{\sigma \rightarrow 0} \varphi_\sigma$$

esiste ed è il funzionale che associa ad ogni funzione $g(v)$ il valore che questa funzione assume in $v = 0$:

$$g(v) \mapsto g(0)$$

questo funzionale è stato introdotto dall'ingegnere elettrico Heaviside e usato dal fisico Dirac per rappresentare una carica elettrica concentrata in un punto. Comunemente è detto δ di Dirac.

Si usa anche scrivere

$$\lim_{\sigma \rightarrow 0} \varphi_\sigma(v) = \delta(v) \quad (1.12)$$

e anche

$$\int g(v) \delta(v) dv = g(0) \quad (1.13)$$

ma bisogna tener presente che si tratta solo di utili convenzioni di scrittura: la δ che compare in (1.12) non è una funzione e il simbolo di integrale in (1.13) non è realmente un integrale. A rigor di termini $\delta(v)$ è solo il funzionale che associa a una qualsiasi funzione $g(v)$ continua e limitata il suo valore in $v = 0$. Precisamente, è un funzionale lineare continuo su uno spazio di funzioni opportuno (che qui non precisiamo). Un funzionale lineare continuo su un opportuno spazio di funzione è anche detto una *distribuzione* o funzione generalizzata (da non confondere con una distribuzione di probabilità). Le distribuzioni sono una generalizzazione del concetto di funzione che permette di fare operazioni che altrimenti sarebbero impossibili.

Analogamente se avessi fatto il limite delle gaussiane traslate

$$\varphi_{v_0, \sigma}(v) \equiv \frac{1}{\sigma} \varphi\left(\frac{v - v_0}{\sigma}\right)$$

identificate con i funzionali

$$g(v) \mapsto E_{\varphi_{v_0, \sigma}}(g(v)) \equiv \int_{-\infty}^{\infty} g(v) \frac{1}{\sigma} \varphi\left(\frac{v - v_0}{\sigma}\right) dv$$

avrei trovato il funzionale che dalla funzione $g(v)$ estrae il suo valore in $v = v_0$. Questo funzionale si indica col simbolo

$$\delta(v - v_0)$$

Per fare conti con la δ di Dirac basta quasi solamente ricordare le proprietà

$$\int g(v) \delta(v - v_0) dv = g(v_0) \quad (1.14)$$

dalla quale discende anche la “normalizzazione”

$$\int \delta(v - v_0) dv = 1$$

“ereditata” dalle gaussiane $\varphi_{v_0, \sigma}(v)$.

e se esiste il limite

$$\lim_{n \rightarrow \infty} f_n(x)$$

allora esiste

$$\lim_{n \rightarrow \infty} \int f_n(x) dx = \int \lim_{n \rightarrow \infty} f_n(x) dx$$

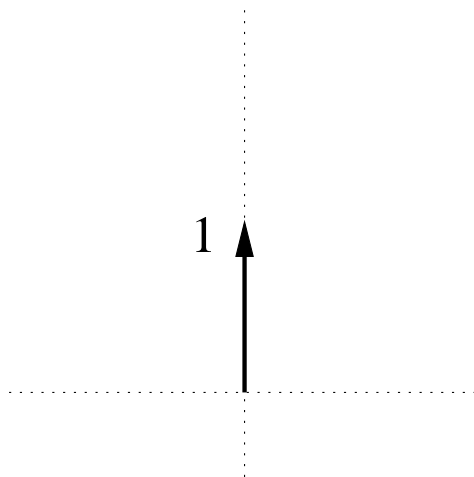


Figura 1.7: Delta di Dirac (impulso unitario).

L'impulso unitario $\delta(v)$ viene rappresentato comunemente come una freccina verticale unitaria al di sopra di $v = 0$ (fig.1.7).

L'introduzione della δ di Dirac è conveniente da vari punti di vista:

- da un punto di vista fisico, perché rappresenta una distribuzione di carica o di massa o di probabilità concentrata in un punto;
- da un punto di vista pratico: ogni δ di Dirac, applicando la formula (1.14), si “mangia” un integrale
- da un punto di vista matematico: è un'estensione del concetto di funzione sulla quale si possono fare operazioni che altrimenti non sarebbero possibili; questo è simile a quanto si fa quando si introducono i numeri complessi: si estende il campo reale per poter fare operazioni che altrimenti non si potrebbero fare.

Vediamo un esempio di operazioni che normalmente non si potrebbero fare. La funzione di distribuzione cumulata

$$\int_{-\infty}^v \frac{1}{\sigma} \varphi\left(\frac{w-v_0}{\sigma}\right) dw = \int_{-\infty}^{\frac{v-v_0}{\sigma}} \varphi(u) du = \Phi\left(\frac{v-v_0}{\sigma}\right)$$

è una primitiva della distribuzione di probabilità gaussiana

$$\frac{1}{\sigma} \varphi\left(\frac{v-v_0}{\sigma}\right) = \frac{d}{dv} \Phi\left(\frac{v-v_0}{\sigma}\right) \quad (1.15)$$

Cercando il limite $\sigma \rightarrow 0$ la funzione di distribuzione cumulata tende a un limite ben preciso (si veda figura 1.8):

$$\lim_{\sigma \rightarrow 0} \Phi\left(\frac{v-v_0}{\sigma}\right) = \lim_{\sigma \rightarrow 0} \int_{-\infty}^v \frac{1}{\sigma} \varphi\left(\frac{w-v_0}{\sigma}\right) dw = \begin{cases} 0 & \text{se } v < v_0 \\ 1/2 & \text{se } v = v_0 \\ 1 & \text{se } v > v_0 \end{cases}$$

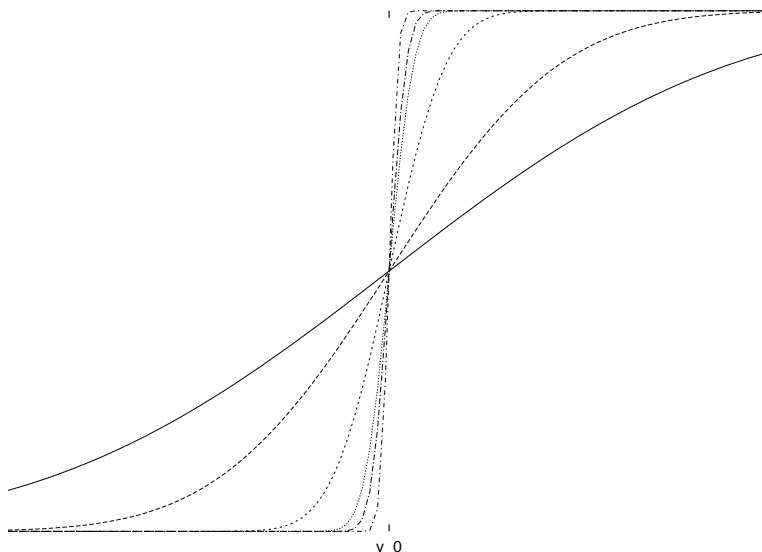


Figura 1.8: Limite per $\sigma \rightarrow 0$ della funzione di distribuzione cumulata.

cioè alla traslata del gradino unitario

$$\lim_{\sigma \rightarrow 0} \Phi \left(\frac{v - v_0}{\sigma} \right) = u(v - v_0)$$

dove

$$u(v) = \begin{cases} 0 & \text{se } v < 0 \\ 1/2 & \text{se } v = 0 \\ 1 & \text{se } v > 0 \end{cases}$$

rappresentato in figura 1.9. (il gradino unitario è noto anche come *funzione di Heaviside* e altre comuni notazioni sono $H(v)$ e $\theta(v)$; dal punto di vista delle teorie delle distribuzioni il fatto che $u(0) = 1/2$ è inessenziale, ma da altri punti di vista può essere importante).

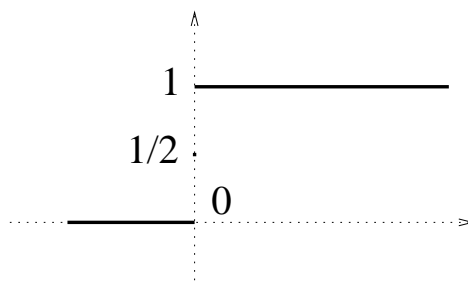


Figura 1.9: Funzione gradino o funzione di Heaviside.

Passando al limite $\sigma \rightarrow 0$ nel senso delle distribuzioni in (1.15) troviamo

$$\frac{d}{dv} u(v - v_0) = \delta(v - v_0) \tag{1.16}$$

cioè a dire che la derivata del gradino unitario è la δ . Oppure

$$u(v - v_0) = \lim_{\sigma \rightarrow 0} \int_{-\infty}^v \frac{1}{\sigma} \varphi \left(\frac{w - v_0}{\sigma} \right) dw = \int_{-\infty}^v \delta(w - v_0) dw$$

Questo mostra che mediante le distribuzioni è possibile calcolare la derivata di funzioni che non solo non sono derivabili, ma sono addirittura discontinue, come il gradino!

La formula (1.16) può essere dimostrata in due modi: uno quello appena visto, e un altro con l'uso di una funzione di prova $g(v)$:

$$\begin{aligned} & \int_{-\infty}^{\infty} g(v) \frac{d}{dv} u(v - v_0) dv \\ &= - \int_{-\infty}^{\infty} \frac{d}{dv} g(v) u(v - v_0) dv \\ &= - \int_{v_0}^{\infty} \frac{d}{dv} g(v) dv \\ &= g(v_0) \end{aligned}$$

1.4 Uso della delta di Dirac nella teoria delle probabilità

Abbiamo visto che la funzione di distribuzione di probabilità di una variabile aleatoria ha un significato probabilistico, ovvero:

$$P(v < V < v + \Delta v) \approx f(v)\Delta v$$

è una densità di probabilità e

$$\text{per } \Delta v \rightarrow 0 \quad f(v) = \frac{d}{dv} P(v < V)$$

dove

$$F(v) = P(V < v) \tag{1.17}$$

è la distribuzione di probabilità cumulata:

$$\begin{aligned} P(V < v) &\approx \sum_{v_i < v} P(v_i < V < v_i + \Delta v) \\ &\approx \sum_{v_i < v} f(v_i)\Delta v \approx \int_{-\infty}^{+\infty} f(v)dv \\ &= F(v) \end{aligned}$$

La distribuzione di probabilità cumulata può anche essere scritta come il valore atteso della funzione complementare del gradino (figura 1.10) :

$$1 - u(V - v)$$

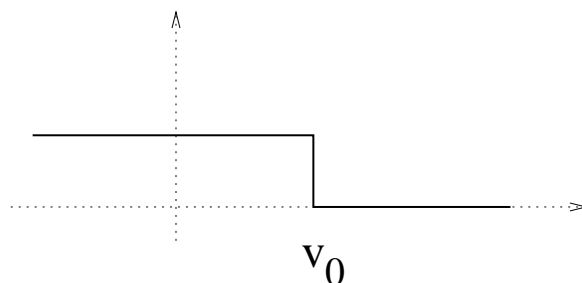


Figura 1.10: Funzione complementare del gradino traslata in v_0 .

$$E(1 - u(V - v)) = \int_{-\infty}^{+\infty} [1 - u(v' - v)]f(v')dv' = \int_{-\infty}^v f(v)dv = F(v_0) = P(V < v)$$

Dunque che cosa ci aspettiamo che sia la densità di probabilità? È intuitivo attendersi che sia il valore atteso della delta², e in effetti:

$$E(\delta(V - v)) = \int_{-\infty}^{+\infty} \delta(v' - v)f(v')dv = f(v)$$

O anche altrimenti:

$$\frac{d}{dv}P(V < v) = \frac{d}{dv}E(1 - u(V - v)) = E\left[\frac{d}{dv}\{1 - u(V - v)\}\right] = E(\delta(V - v)) = f(v) \quad (1.18)$$

²Il valore atteso del gradino è invece la distribuzione di probabilità complementare:

$$E(u(V - v)) = \int_{-\infty}^{+\infty} [u(v' - v)]f(v')dv' = \int_v^{+\infty} f(v')dv' = \int_{-\infty}^{+\infty} f(v')dv' - \int_{-\infty}^v f(v')dv' = 1 - F(v) = P(V > v)$$