

ALMA MATER STUDIORUM · UNIVERSITÀ DI
BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica, Curriculum applicativo

**IL METODO ENTROPICO NELLE
APPLICAZIONI DELLA
MECCANICA STATISTICA ALLE
SCIENZE SOCIALI**

Tesi di Laurea in Meccanica Statistica

Relatore:

**Chiar.mo Prof.
Pierluigi Contucci**

Correlatore:

**Dott.ssa
Micaela Fedele**

**Presentata da:
Ilaria Brunetti**

**Seconda Sessione
Anno Accademico 2010-2011**

Indice

Elenco delle Figure	iv
Elenco delle Tabelle	v
Introduzione	vi
1 La Matematica per gli studi socio-economici	1
1.1 Quando le scienze dure incontrano le scienze sociali	1
1.2 La Teoria dei Modelli a Scelta Discreta	4
1.2.1 La nascita dei Modelli a Scelta Discreta	4
1.2.2 Il modello Logit Multinomiale	6
1.2.3 I limiti della DCT	8
1.3 La Meccanica Statistica	9
1.3.1 I modelli della Meccanica Statistica	10
1.4 Il modello di McFadden in chiave fisica	12
2 Il modello di Curie-Weiss	16
2.1 Il modello	17
2.2 Limite termodinamico	19
2.2.1 Esistenza del limite termodinamico	19
2.2.2 Le proprietà di fattorizzazione	21
2.2.3 I bounds della funzione pressione	24
2.2.4 L'equazione di campo medio	29
2.2.5 La transizione di fase	34

2.2.6	Studio grafico dell'equazione di campo medio	36
3	Il modello di Curie-Weiss multipopolato	40
3.1	Il modello	40
3.2	Limite termodinamico	42
3.3	Esistenza del limite	42
3.3.1	Il metodo di Talagrand per il calcolo del limite	46
3.3.2	Equazione di campo medio	49
4	Il problema inverso	50
4.1	Problema inverso per il modello di Curie-Weiss	51
4.1.1	Modello monopolato	51
4.1.2	Modello multipopolato	53
4.2	La tecnica della massima verosimiglianza	54
5	L'entropia come strumento per la modellizzazione	57
5.1	L'entropia: in termodinamica e in teoria dell'informazione	58
5.2	Il Principio di Massima Entropia	61
5.3	Il Principio di Massima Entropia ed il modello di Curie-Weiss	64
5.3.1	Modello ad una popolazione	64
5.3.2	Modello multipopolato	66
5.4	Critiche al Principio della Massima Entropia	68
6	Dati e risultati sperimentali	70
6.1	Il problema	70
6.2	I dati	71
6.3	Il modello	74
6.3.1	Il problema inverso teorico	75
6.4	Il problema dell'inversione	76
7	Prospettive future	80
A	La distribuzione di Gumbel	83

B Teorema di esistenza del limite per i modelli di campo medio	85
Bibliografia	89
Bibliografia	89
Ringraziamenti	94

Elenco delle figure

2.1	Rappresentazione grafica dell'equazione di consistenza - $h \neq 0$	37
2.2	Rappresentazione grafica dell'equazione di consistenza - $h = 0$	38
2.3	(a) Superficie m in funzione di J e h - modello Curie-Weiss (b) Stessa superficie vista dall'alto	38
2.4	Grafico di m al variare di J per $h = 0$ - modello Curie-Weiss .	39

Elenco delle tabelle

6.1	Percentuale di adesione su primo invito dal 01/01/03 al 31/12/05 per i singoli comuni della provincia di Parma	73
-----	---	----

Introduzione

‘J’entends par physique sociale la science qui a pour objet propre l’étude des phénomènes sociaux, considérés dans le même esprit que les phénomènes astronomiques, physiques, chimiques et physiologiques, c’est-à-dire assujettis à des lois naturelles invariables, dont la découverte est le but spécial de ses recherches’

August Comte [10]

Le dinamiche che determinano il comportamento di un gruppo di individui sono oggetto di un interesse multidisciplinare, che negli ultimi decenni è notevolmente cresciuto e si è sviluppato coinvolgendo non solo gli studiosi delle materie socio-economiche, ma anche quelli di materie apparentemente distanti, ossia di Matematica e Fisica.

Le suddette scienze, infatti, possono aiutare a comprendere i meccanismi fondamentali alla base dei comportamenti di un gruppo di persone, nonché le relazioni tra motivazioni individuali e azioni collettive. Il nostro lavoro ha l’obiettivo di illustrare come Matematica e Fisica forniscano effettivamente un’utile chiave di lettura delle dinamiche di gruppo. Utilizzeremo, in particolare, i modelli della Meccanica Statistica nati per descrivere un insieme di particelle interagenti, al fine di descrivere quantitativamente il comportamento di una popolazione.

Il fisico francese Serge Galam iniziò ad occuparsi di queste problematiche comportamentali già durante gli anni Settanta, ma solo recentemente si sono superate l’iniziale ostilità della comunità scientifica e la diffidenza dei sociologi e si è affermata la Sociofisica, scienza che indaga i fenomeni sociali attraverso la teoria della Termodinamica e della Meccanica Statistica.

Il formalismo matematico proprio della Meccanica Statistica consente di rappresentare le transizioni di fase che avvengono in un sistema di particelle, quantificando gli effetti delle interazioni tra elementi. Parallelamente, in Sociologia si è compreso che le azioni di una persona non sono determinate unicamente da caratteristiche ed obiettivi soggettivi, ma anche da quella che è stata definita come ‘pressione sociale’, ovvero dalla tendenza a conformarsi agli altri, o, viceversa, a distinguersi, a seconda della situazione. In particolare, si è capito che la presenza di interazione tra individui consente di giustificare il verificarsi delle crisi all’interno di un sistema, degli ‘opinion shifts’ ed, in generale, di tutti quei cambiamenti drastici ed improvvisi che si osservano passando dalle intenzioni dei singoli al comportamento collettivo. L’economista Thomas Schelling, uno dei principali studiosi di queste dinamiche di gruppo, si è dedicato all’analisi di diverse problematiche sociali, tra le quali il razzismo e le strategie militari, focalizzando l’attenzione sull’effetto che le influenze reciproche hanno sul comportamento collettivo. In [41] mostrò, ad esempio, come le preferenze di un individuo riguardo alla composizione razziale del proprio vicinato potessero originare il fenomeno della segregazione.

L’analogia tra le discontinuità che si riscontrano nel comportamento di un sistema di particelle e di un gruppo di persone ha posto quindi le basi per una descrizione quantitativa dei fenomeni sociali. Nonostante le azioni umane siano guidate da un complesso insieme di forze, sicuramente più difficili da individuare e da precisare rispetto a quelle che intervengono sulle particelle, si pensò di approcciare lo studio dei fenomeni sociali con gli strumenti della Matematica e della Meccanica Statistica, al fine di rendere in qualche modo prevedibile il comportamento di un insieme di persone. Tra i diversi metodi utilizzati in Sociofisica, [19], [8], [9], [16], [34], [21], in questo lavoro seguiremo quello sviluppato in [11], in cui, per la prima volta, si introduce un nuovo modello di campo medio, il modello di Curie-Weiss a due popolazioni, per rappresentare le dinamiche relative al contatto tra due culture.

Chiaramente, attingere alla Fisica per gli studi sociali significa utilizzarne

modelli e metodi senza però considerare l'intero apparato teorico sottostante: come vedremo in seguito, si trascura il concetto di temperatura e non si dispone di un'unità di misura per l' 'Hamiltoniana sociale'.

Nonostante ciò, la distribuzione di Boltzmann-Gibbs che caratterizza i modelli della Meccanica Statistica sembrerebbe essere la 'misura naturale' per descrivere un sistema costituito da elementi interagenti, siano questi particelle o esseri umani.

La prima volta che comparve la misura di Boltzmann in studi di carattere sociale non fu nel lavoro di un fisico bensì di un economista, Daniel McFadden, uno dei massimi esponenti della Teoria dei Modelli a Scelta Discreta (DCT), teoria economica che si occupa di stabilire il legame tra caratteristiche del singolo e comportamento di gruppo, al fine specifico di prevedere la risposta di una popolazione posta di fronte a diverse alternative. Il modello che formulò prevedeva che l'errore legato alla funzione utilità avesse distribuzione logit multinomiale e, come vedremo in seguito, massimizzando l'utilità così definita, lo stato fondamentale che si ricava corrisponde proprio allo stato di Boltzmann-Gibbs.

I modelli della DCT non considerano, tuttavia, l'effetto delle influenze reciproche, in quanto i 'decision makers' vengono connotati esclusivamente da attributi individuali, ed il loro impiego si limita alla rappresentazione di quelle situazioni in cui le interazioni sono trascurabili. Nei primi anni Novanta, gli economisti Steven Durlauf e William Brock [8], [9], grazie alle teorie della fisica, integrarono questi modelli al fine di descrivere i sistemi interagenti ed osservarono come i modelli della DCT ed i modelli della Meccanica Statistica, non solo potessero essere utilizzati per affrontare gli stessi problemi, ma potessero essere tradotti nello stesso linguaggio.

Il fatto che la misura di Boltzmann-Gibbs emerga anche dal modello di McFadden, senza considerare la teoria fisica, risponde, almeno parzialmente, ai dubbi ed alle incertezze circa la possibilità di applicare modelli meccanico statistici ai comportamenti umani. Le persone, infatti, non sono particelle in un bagno a temperatura T , per cui le ipotesi fisiche necessarie alla

formulazione del modello potrebbero venire meno.

In questo lavoro ci siamo chiesti se esista un'altra strada possibile per giustificare l'uso della misura di Boltzmann-Gibbs per le Scienze Sociali. Mostriamo quindi un metodo diverso, che non implica alcuna ipotesi fisica, ma sfrutta il concetto di entropia, così come sviluppato dalla Teoria dell'Informazione. Si tratta del Principio della Massima Entropia, proposto per la prima volta da Jaynes alla fine degli anni Cinquanta, che consente, pur seguendo un percorso diverso, di giungere alla formulazione di modelli equivalenti a quelli della Meccanica Statistica.

Questo lavoro è organizzato come segue. Nel Primo Capitolo si mostrerà come la necessità di una teoria scientifica per le scienze sociali non sia un'esigenza recente ma abbia origine nel passato, ed in particolare come, grazie alla rivoluzione di Maxwell e Boltzmann, fondatori della Meccanica Statistica, si sia arrivati a concepire la Fisica Sociale allo stato attuale. Vedremo che non è stata solo la Sociologia ad attingere alle metodologie scientifiche, ma anche Fisica e Matematica hanno ricevuto contributi e stimoli importanti dalle scienze umane. Illustreremo i modelli a Scelta Discreta, soffermandoci in particolare sul modello Logit Multinomiale. Introduciamo i modelli della Meccanica Statistica e dimostreremo come anche il modello Logit Multinomiale possa essere pensato e riscritto nei termini della fisica.

Nel Secondo e nel Terzo capitolo presenteremo in dettaglio due modelli della Meccanica Statistica detti di campo medio: il modello di Curie-Weiss ed una sua estensione al caso multipopolato. Ne definiremo l'Hamiltoniana e le osservabili principali, e, per verificare che si tratta di modelli ben posti, dimostreremo l'esistenza del limite termodinamico della funzione pressione ad essi associata e ne calcoleremo il valore esatto. Sfrutteremo due diverse tecniche: individuando bound inferiore e superiore per la pressione per il modello monopopolato ed il metodo di Talagrand per il multipopolato.

Nel Quarto Capitolo analizzeremo il problema dell'inversione di tali modelli, cioè vedremo come sia possibile ricavare i parametri a partire da un insieme di dati.

Nel Quinto Capitolo affronteremo lo studio del Principio della Massima Entropia: daremo innanzitutto la definizione di entropia nell'ambito della termodinamica e in quello della Teoria dell'Informazione e spiegheremo esattamente in cosa consiste il Principio di Massima Entropia introdotto da Jaynes. Mostreremo quindi come questo possa essere utilizzato per elaborare modelli sostanzialmente equivalenti a quelli di Curie-Weiss, senza bisogno del supporto di ipotesi fisiche.

Il Sesto Capitolo sarà incentrato, infine, sul progetto di ricerca al quale sta attualmente lavorando un gruppo di docenti e ricercatori delle università di Bologna, Modena, Ferrara e Parma. Si tratta di uno studio sul comportamento delle donne della provincia di Parma coinvolte nella campagna di screening per il Pap test. Sebbene il problema sia ancora aperto ed in fase di studio, mostreremo quali ipotesi si sono sviluppate dall'analisi dei dati a disposizione, illustrando un possibile modello di campo medio pensato per rappresentare la situazione. Vedremo, infine quali sono le difficoltà e le problematiche rilevate e quali soluzioni sono state proposte.

Capitolo 1

La Matematica per gli studi socio-economici

Essentially, all models are wrong, but some are useful

George E.P.Box [7]

1.1 Quando le scienze dure incontrano le scienze sociali

Le scienze umane e le cosiddette scienze dure -*hard science*- ovvero quelle che si basano su dati sperimentali e applicano il metodo scientifico, quali Matematica, Fisica, Chimica o Biologia, vengono spesso considerate come discipline molto distanti, quasi in antitesi tra loro, con interessi, scopi e strumenti completamente diversi. Potrebbe quindi sembrare alquanto difficile stabilire una comunicazione tra loro, eppure non è così. Negli ultimi anni infatti, si è affermata la cosiddetta Sociofisica, ovvero lo studio della società nell'ottica della teoria fisica, in particolare grazie agli strumenti della meccanica statistica e della teoria del Caos. Il fisico francese Serge Galam è ritenuto uno dei fondatori di tale disciplina; *'Do humans behave like atoms?'* si chiede lo

scienziato in [21] e la provocatoria risposta che dà è chiaramente affermativa. Questo non implica, secondo l'autore, svilire la complessità dell'essere umano e ridurre la sua libertà, come molti potrebbero obiettare, ma, anzi, consente di aprire nuove ed interessanti prospettive sull'uomo.

In questo lavoro mostreremo che cosa significhi e come sia effettivamente possibile pensare ad un insieme di persone come ad un insieme di atomi e come questo approccio possa contribuire a comprendere e a prevedere alcuni comportamenti di una popolazione.

Le origini delle fruttuose interazioni tra Fisica, Matematica e Scienze Sociali non sono in realtà così recenti, ma vanno ricercate più lontano: non sono rari gli esempi di studiosi e scienziati che in passato si sono dedicati al tentativo di unificare questi mondi apparentemente così diversi.

La prima testimonianza significativa di tale sforzo risale al 1600, con il lavoro di Thomas Hobbes (1588-1679), [1] intellettuale inglese che, nonostante gli studi classici, nel 1629 venne a conoscenza dell'opera di Euclide '*Elementi di Geometria*' e rimase affascinato dal ragionamento deduttivo del matematico. Qualche anno più tardi, incontrò Galileo ed, influenzato anche dalle sue teorie, iniziò a concepire l'idea che si potessero studiare le leggi universali che regolano gli esseri umani e la società, così come la Fisica studiava le leggi universali che governano la natura. Nella sua opera principale, il *Leviathan* (1651), emergono chiaramente la visione meccanicistica del mondo che aveva acquisito ed il suo scopo, ovvero di attingere agli strumenti propri della scienza, quali il rigoroso ragionamento logico-deduttivo e la costruzione assiomatica, per affrontare gli studi sociali e politici. Sebbene dal punto di vista politico le argomentazioni e le conclusioni di Hobbes siano certamente discutibili, il suo lavoro resta storicamente e filosoficamente molto importante, in quanto costituisce un sorprendente antecedente all'attuale sviluppo della Fisica sociale.

William Petty (1623-1687), discepolo di Hobbes, nella sua opera '*Political Arithmetic*' portò avanti l'idea di fare della politica una scienza, sostenendo che la società potesse essere compresa solo misurandola e quantificandola, e

rendendo in tal modo verificabili i principi del governo [1]. Petty, rispetto a Hobbes, compì un ulteriore passo verso le scienze dure, in quanto lavorò sui ‘numeri sociali’, ovvero su un insieme di dati relativi ai tassi di morte della popolazione londinese, fornitigli dall’amico e collaboratore John Graunt (1620-1674), considerato uno dei fondatori della statistica moderna. Nel 1749 lo studio di questi ‘numeri sociali’ venne definito *Statistica*, ed iniziò ad attirare l’interesse dei matematici. Fu in seguito il filosofo francese Auguste Comte (1798-1857) a coniare il termine *Physique sociale*, Fisica sociale, per definire quella scienza che ha come oggetto di studio i fenomeni sociali considerati allo stesso modo dei fenomeni chimici, fisici o astronomici.

Un altro nome importante nella storia della Fisica sociale è senza dubbio Adolphe Quetelet (1796-1874) che, ispirato dai tentativi passati e grazie ai nuovi strumenti della statistica, elaborò delle leggi social, illustrate nelle sue opere di ‘meccanica sociale’, formalmente simili alle leggi della natura.

Nel corso della storia non sono state solo le discipline sociali ad attingere alle idee ed ai metodi della Fisica, ma è avvenuto anche l’inverso, come testimonia il lavoro di Maxwell (1831-1879). Per comprendere come fossero distribuite le velocità delle particelle rispetto alla media ed elaborare la sua teoria cinetica dei gas, egli fu infatti ispirato proprio delle leggi statistiche nate per lo studio della società. Nel 1866, il fisico scozzese formulò la nota distribuzione di Maxwell, generalizzando così le leggi della termodinamica precedentemente stabilite e fornendo una migliore spiegazione delle osservazioni sperimentali. Anche Ludwig Boltzmann (1844-1906), considerato il fondatore della termodinamica statistica, era venuto a contatto con la statistica sociale: *‘The molecules are like to many individuals, having the most various states of motion, and the properties of gases only remain unaltered because the number of these molecules which on average have a given state of motion is constant.’* [6]

La nuova scienza di Boltzmann si basa sull’idea che le caratteristiche macroscopiche di un sistema, quali energia, calore o pressione, siano comprensibili sono in chiave microscopica; poichè il numero di costituenti ele-

mentari di un sistema fisico è molto elevato, si necessita di descrizioni in termini probabilistici e le proprietà del sistema sono interpretabili in termini statistici. Non deve quindi stupire se negli ultimi anni la Fisica sta di nuovo guardando all'uomo, nel tentativo di capire gli schemi che regolano il comportamento di una popolazione: gli stessi fisici furono ispirati dagli studi sociologici. Attualmente, non si cercano più le leggi universali che regolano la società, ma piuttosto, si vuole comprendere la struttura organizzativa di una rete sociale per poter prevedere il comportamento di un gruppo di individui che agiscono nel proprio interesse.

Anche le discipline economiche si sono occupate di questo aspetto ed hanno apportato contributi importanti; in particolare, di grande rilievo è la Teoria dei Modelli a Scelta Discreta, che, come vedremo nel prossimo paragrafo, si pone il problema di comprendere il comportamento emergente di un insieme di individui di fronte ad una scelta. Di seguito illustreremo alcuni modelli della DCT e della meccanica statistica, mostrando innanzitutto come, sebbene partendo da teorie completamente diverse, possano non solo essere usati per scopi comuni, ma addirittura possano essere scritti nello stesso linguaggio della Fisica.

1.2 La Teoria dei Modelli a Scelta Discreta

1.2.1 La nascita dei Modelli a Scelta Discreta

Negli anni Settanta, lo studio dei sistemi di trasporto portò alla nascita e allo sviluppo della Teoria dei Modelli a Scelta Discreta, la *Discrete Choice Theory* (DCT) [3].

Si tratta di modelli statistici di tipo *'bottom-up'*, ovvero in cui si studiano i singoli individui all'interno di un gruppo per comprendere il comportamento emergente a livello globale. L'obiettivo principale della DCT non consiste semplicemente nel descrivere, quanto, piuttosto, nel prevedere il comportamento della popolazione in esame a partire dalla conoscenza dei singoli. A tale scopo, le persone vengono identificate sulla base di determinate caratter-

istiche individuali, opportunamente selezionate in modo da creare un legame tra attributi del singolo e scelta effettuata. Se, ad esempio, si volesse studiare quale mezzo di trasporto venga preferito per recarsi al lavoro in una certa popolazione, si potrebbe supporre che la scelta di ciascun individuo dipenda dall'impiego, dal sesso, dal reddito, dal luogo di residenza

Il contributo principale in questo campo fu quello di Daniel McFadden che, nel 1975, riuscì a stimare la percentuale della popolazione di San Francisco che avrebbe utilizzato il nuovo mezzo di trasporto pubblico della città, il BART (6.3% percentuale prevista, 6.2% percentuale effettiva misurata dopo l'avvento del BART).

Tale studio valse il premio Nobel per l'economia (2000) a McFadden e segnò il successo della DCT nelle scienze economiche e sociali [33].

Per i modelli della DCT, si assume che gli individui siano *'razionali'*, ovvero che agiscano in modo che la scelta effettuata apporti loro il maggior beneficio possibile. Matematicamente, questo si traduce nella massimizzazione di una funzione utilità, o funzione guadagno, che rappresenta il vantaggio che la persona trae dalla scelta. Per quanto una persona possa essere supposta *'razionale'*, affinché il modello sia realistico, è necessario considerare un certo grado di imprevedibilità, intrinseca all'essere umano. La funzione utilità che ne risulta assume così la seguente forma:

$$U = V + \epsilon \quad (1.1)$$

dove V è la parte deterministica, mentre ϵ è il contributo casuale che rispecchia l'irrazionalità legata alla persona.

Al fine di specificare il modello, è necessario definire i seguenti elementi:

- gli agenti, o *decision makers*, cioè gli individui sottoposti alla scelta;
- l'insieme delle opzioni possibili;
- gli attributi che caratterizzano gli individui e le alternative disponibili;
- le *decision rules*, ovvero le regole secondo cui avviene il processo di scelta.

Consideriamo il caso in cui i decision makers siano un gruppo finito di persone, che devono compiere una scelta tra un numero finito di alternative. Sia gli individui che le alternative sono caratterizzati da alcuni attributi, per cui, la parte deterministica di (1.1) può essere espressa come:

$$V = \sum_a \lambda_a x_a + \sum_a \alpha_a y_a$$

dove a è l'indice associato ad un individuo, le x_a e y_a rappresentano, rispettivamente, le caratteristiche di una data scelta e di un dato individuo, mentre i parametri λ_a e α_a sono i parametri che indicano il peso di ciascun attributo sulla scelta effettuata.

Per definire il modello è inoltre necessario assegnare ad ϵ una certa distribuzione, che generalmente appartiene alla classe delle distribuzioni di valore estremo. Nel prossimo paragrafo mostreremo il modello Multinomial Logit, che si ottiene scegliendo, in questa classe, la distribuzione di Gumbel (si veda l'appendice (A)).

1.2.2 Il modello Logit Multinomiale

Il modello Logit Multinomiale, introdotto dal matematico sociologo Duncan Luce nel 1959 [3] e sviluppato in seguito da McFadden, gode della proprietà di generare una soluzione in forma chiusa per la probabilità di una scelta, senza di fatto imporre restrizioni al modello. Come vedremo in seguito, inoltre, tale modello risulta particolarmente interessante per la Sociofisica, in quanto massimizzandone la funzione utilità si ricava lo stato fondamentale legato alla distribuzione di Boltzmann. Supponiamo di avere N individui che scelgono tra k alternative. Il beneficio che la scelta dell' i -esima alternativa apporta all' l -esimo individuo è indicata con:

$$U_i^{(l)} = V_i^{(l)} + \epsilon_i^{(l)}. \quad (1.2)$$

Al termine casuale ϵ della funzione utilità viene assegnata la distribuzione di Gumbel (si veda appendice A), la cui densità di probabilità è la seguente:

$$f(\epsilon) = \mu e^{-\eta(\epsilon-\eta)} \exp[-e^{-\eta(\epsilon-\eta)}], \quad (1.3)$$

dove i parametri sono $\eta \in \mathbb{R}$ e $\mu > 0$, detti rispettivamente parametro di locazione e di scala.

Data la presenza del termine casuale, massimizzare $U^{(l)}$ equivale a determinare la probabilità $p_{l,i} = P('l \text{ sceglie } i')$. Se l sceglie i , ciò significa che l'utilità che ne ricava è maggiore rispetto alle altre scelte, per cui:

$$p_{l,i} = P(U_i^{(l)} \geq \max_{j \neq i} U_j^{(l)}) = P(V_i^{(l)} + \epsilon_i^{(l)} \geq \max_{j \neq i} (V_j^{(l)} + \epsilon_j^{(l)}))$$

La distribuzione di Gumbel gode della seguente proprietà (A.3): il massimo di un insieme di variabili di Gumbel aventi lo stesso parametro di scala μ ha ancora distribuzione di Gumbel, con medesimo μ e parametro di locazione $\frac{1}{\mu} \ln \sum_{j=1}^J e^{\mu \eta_j}$. Ponendo quindi $U^* = \max_{j \neq i} (V_j^{(l)} + \epsilon_j^{(l)})$, si ha che:

$$U^* = {}^d \text{Gumbel} \left(\frac{1}{\mu} \ln \sum_{j \neq i} e^{\mu V_j^{(l)}}, \mu \right)$$

e definendo V^* come:

$$V^* = \frac{1}{\mu} \ln \sum_{j \neq i} e^{\mu V_j^{(l)}}$$

si ha $U^* = V^* + \epsilon^*$, con $\epsilon^* = {}^d \text{Gumbel}(0, \mu)$.

Sfruttando un'altra proprietà della distribuzione, (A.2) per cui la distribuzione di una variabile ottenuta come differenza di due variabili indipendenti con distribuzione di Gumbel con lo stesso parametro μ ha distribuzione logistica con densità di probabilità:

$$f(\epsilon) = \frac{\mu e^{-\mu(\eta_2 - \eta_1 - \epsilon)}}{(1 + e^{-\mu(\eta_2 - \eta_1 - \epsilon)})^2},$$

si ricava che la probabilità $p_{l,i}$ vale:

$$\begin{aligned}
 p_{l,i} &= P\left(V_i^{(l)} + \epsilon_i^{(l)} \geq V^* + \epsilon^*\right) = \frac{1}{1 + e^{\mu(V^* - V_i^{(l)})}} = \frac{e^{\mu V_i^{(l)}}}{e^{\mu V_i^{(l)}} + e^{\mu V^*}} \\
 &= \frac{e^{\mu V_i^{(l)}}}{\sum_{j=1}^k e^{\mu V_j^{(l)}}}
 \end{aligned}$$

Il parametro μ può essere incorporato ai gradi di libertà di $V_i^{(l)}$, per cui:

$$p_{l,i} = \frac{e^{V_i^{(l)}}}{\sum_{j=1}^k e^{V_j^{(l)}}} \quad (1.4)$$

1.2.3 I limiti della DCT

La DCT ha sicuramente rappresentato una grande svolta nello studio e nella modellizzazione del comportamento di una popolazione di individui. Il successo del lavoro di McFadden è stato largamente riconosciuto ed i modelli della DCT sono tuttora utilizzati, non solo nell'analisi della domanda di trasporti [20], [35], ma anche in altri settori, come ad esempio in studi di tipo medico [13], [24], [38] o sociale [39].

I modelli della DCT, tuttavia, non considerano in alcun modo le interazioni che esistono tra individui, e le conseguenti ripercussioni che queste hanno sulle scelte effettuate.

Si pensi ad un esempio molto semplice, quale può essere la scelta di un ristorante. Certamente contano molto i gusti personali, la posizione del ristorante o il budget a disposizione. Tuttavia, è indubbio che il 'passaparola', il consiglio di un amico, o il sapere che un ristorante è più frequentato rispetto agli altri, pesano spesso in maniera determinante sulla scelta finale.

Questi limiti della Teoria dei Modelli a Scelta Discreta emersero già durante gli anni '80, in concomitanza con la diffusione di numerosi studi di psicologia sociale che evidenziavano il peso che queste interazioni tra individui possono avere. Di grande rilievo, in questo campo, furono gli studi di Schelling [42] e Granovetter [25] i quali riuscirono a spiegare quei cambiamenti

drastici che si osservano nei comportamenti di gruppo, proprio considerando le interazioni.

Vedremo quindi come si iniziò a pensare a nuovi modelli, integrando il modello di McFadden con delle quantità che rappresentassero gli effetti delle influenze reciproche tra individui [16], [8],[8] e consentissero quindi di spiegare i comportamenti collettivi emergenti.

Il modelli della DCT restano tuttora validi, ma limitatamente a quelle situazioni in cui l'effetto delle intarazioni tra individui è trascurabile. Negli altri casi, è invece fondamentale capire come quantificare e pesare opportunamente le influenze reciproche all'interno di una popolazione.

1.3 La Meccanica Statistica

L'avvento della Meccanica Statistica segnò un profondo cambiamento nel mondo della Fisica, non solo per il diverso approccio e per i nuovi strumenti matematici utilizzati, ma anche per le ripercussioni che questa ebbe sulla percezione del mondo e dell'uomo.

Un importante stimolo per l'introduzione della statistica nella termodinamica si ebbe nel 1738, con la pubblicazione del testo di Bernoulli 'Idrodinamica', che pose le basi per la teoria cinetica dei gas. Per la prima volta, i gas vennero descritti come un insieme di molecole che si muovono in tutte le direzioni, il cui impatto su una superficie dà origine alla pressione, ed il cui movimento genera il calore che noi percepiamo. Anche Clausius si occupò della diffusione delle molecole, ma fu poi Maxwell, nel 1859, a studiare la distribuzione delle velocità medie delle particelle; questa fu la prima legge statistica in Fisica e prese così il nome di distribuzione di Maxwell. Boltzmann proseguì gli studi da lui avviati, ed è tutt'ora considerato il fondatore della Meccanica Statistica. Nel 1875 formulò la relazione tra entropia e numero di microstati possibili per un certo macrostato di un gas e riformulò in termini probabilistici il secondo principio della termodinamica.

Il passaggio dal determinismo newtoniano alle scienze statistiche e la mod-

ellizzazione delle interazioni tra elementi di un sistema fisico, insieme agli studi sociali per la comprensione degli *opinion shifts* in un gruppo di individui, sono gli elementi che hanno reso possibile lo sviluppo della Sociofisica allo stato attuale. Si vide infatti l'analogia tra questi drastici 'cambi di opinione' e le transizioni di fase dovute all'interazione tra particelle. Nel 1974 il matematico Hans Föllmer sviluppò un modello economico per descrivere il comportamento di un insieme di agenti, basandosi proprio su un modello di Ising per gli atomi dei magneti, in cui quindi si consideravano le interazioni tra elementi [19]

1.3.1 I modelli della Meccanica Statistica

Dato un sistema di particelle, gli elementi che caratterizzano i modelli di Meccanica Statistica sono:

- la funzione Hamiltoniana $H_N(\bar{x})$, che rappresenta l'energia legata ad una configurazione \bar{x} ;
- la funzione di distribuzione dello stato \bar{x} , detta **misura di Boltzmann-Gibbs**:

$$P(\bar{x}) = \frac{e^{-\beta H_N(\bar{x})}}{\int_{\bar{x}} e^{-\beta H_N(\bar{x})}}, \quad (1.5)$$

dove $\beta = \frac{1}{kT}$, con k costante di Boltzmann e T temperatura assoluta. Questo parametro è di grande importanza per i modelli fisici, in quanto fornisce il fattore di scala dell'Hamiltoniana, legando temperatura ed energia. Poichè $\beta \geq 0$ si ha che le configurazioni più probabili sono quelle con energia più bassa.

Vediamo più in dettaglio un esempio di modello meccanico-statistico. Consideriamo un sistema costituito da N particelle; ad ogni particella i si associa il valore del suo spin σ_i , che rappresenta il suo momento angolare intrinseco e quindi la direzione di rotazione. Per semplicità, ci limitiamo a considerare

un modello dicotomico, per cui $\sigma_i = \pm 1$. Supponiamo che su ciascuna particella i , agisca un campo esterno, la cui intensità si indica con h_i ; si considera inoltre, un effetto dovuto alla presenza degli altri spin. Ci limitiamo alle interazioni tra coppie, per cui, data la coppia di particelle (i, j) , l'intensità dell'interazione è indicata con J_{ij} . Se J_{ij} è positiva, significa che le due particelle tendono ad allinearsi, mentre quando J_{ij} è negativa queste hanno la tendenza ad opporsi.

La funzione energia associata ad una certa configurazione di spin è data dalla seguente Hamiltoniana:

$$H_N = - \sum_{i,j=1}^N J_{ij} \sigma_i \sigma_j - \sum_{i=1}^N h_i \sigma_i. \quad (1.6)$$

Per i modelli sociali, il parametro β che compare in (1.5) viene incluso incluso nei gradi di libertà legati ai parametri J_{ij} e h_i perchè, in questo caso, il concetto di temperatura perde di significato; la distribuzione di Boltzmann-Gibbs risulta essere:

$$P(\sigma) = \frac{e^{-H_N(\sigma)}}{\sum_{\sigma \in \Omega_N} e^{-H_N(\sigma)}} \quad (1.7)$$

dove $\Omega_N = \{-1, 1\}^N$ è l'insieme di tutte le configurazioni possibili.

Nel momento in cui si assegna l'Hamiltoniana e si definisce la distribuzione di Gibbs, una funzione molto importante legata al modello è la funzione pressione, p_N :

$$p_N = \frac{1}{N} \ln \sum_{\sigma} e^{-H_N(\sigma)} = \frac{1}{N} \ln Z_N \quad (1.8)$$

dove Z_N è la **funzione di partizione**, $Z_N = \sum_{\sigma} e^{-H_N(\sigma)}$.

Vediamo come sia possibile pensare a questo modello di particelle interagenti come ad un modello per gli studi comportamentali, ovvero come da un insieme di N particelle si possa passare ad un insieme di persone. Innanzitutto si suppone che gli individui affrontino una scelta di tipo binomiale, ovvero siano sottoposti ad una *'yes or no question'*, per cui le alternative possibili

sono solo due (*binomial choice model*). In questo conteso σ_i rappresenta la scelta dell' i -esimo individuo e vi si assegna un valore nel modo seguente:

$$\sigma_i = \begin{cases} +1 & \text{se } i \text{ dice SI} \\ -1 & \text{se } i \text{ dice NO} \end{cases} \quad (1.9)$$

Il campo esterno h_i si interpreta come l'effetto che hanno, sulla scelta della persona, le proprie inclinazioni, dovute alla cultura, alla storia ed al carattere personali; le interazioni tra coppie di particelle, J_{ij} , non diventano altro che le influenze reciproche. L'hamiltoniana del sistema, rappresenterà così, non l'energia, quanto piuttosto una 'funzione costo', che gli individui tendono a minimizzare.

Prima di passare ad illustrare i modelli di campo medio, vediamo come anche il modello di scelta discreta precedentemente illustrato possa essere interpretato in chiave fisica, e descritto anch'esso in termini di Hamiltoniana.

1.4 Il modello di McFadden in chiave fisica

Il modello di McFadden può essere riletto e tradotto nel linguaggio della Meccanica Statistica come un modello di gas perfetti non interagenti.

Dal punto di vista della Fisica, determinare il massimo della funzione guadagno equivale a minimizzare la funzione energia, determinando lo stato fondamentale, (*ground state*).

Per i gas non interagenti la funzione energia si esprime come:

$$H_N(\sigma) = - \sum_{i=1}^N h_i \sigma_i \quad (1.10)$$

dove la quantità h_i indica il campo esterno agente sull' i -esima particella.

Dimostriamo matematicamente l'effettiva equivalenza tra il modello di gas perfetti e il modello Logit Multinomiale illustrato. Consideriamo sempre N individui, ciascuno dei quali sceglie tra k alternative; la scelta della m -esima alternativa si rappresenta con l' m -esimo vettore della base canonica di \mathbb{R}^k :

$$\sigma_m = e_m = (0, \dots, 1, \dots, 0)$$

La configurazione del sistema non è altro che l'insieme delle scelte degli N individui, ovvero:

$$\sigma = \{\sigma^{(1)}, \dots, \sigma^{(N)}\}$$

dove preso un individuo l , si avrà $\sigma^{(l)} = \sigma_m$ per un qualche $m \in 1, \dots, k$. Ciascuna alternativa apporta un certo guadagno all'individuo l ; il vettore composto da tali quantità sarà quindi:

$$v^{(l)} = (V_1^{(l)}, \dots, V_k^{(l)}),$$

dove, $V_i^{(l)}$ è la parte deterministica di (1.2), ovvero il beneficio che l trae dalla scelta dell'alternativa i . Con queste notazioni, l'Hamiltoniana del modello diventa:

$$H_N(\sigma) = - \sum_{l=1}^N v^{(l)} \cdot \sigma^{(l)}.$$

La funzione pressione ad essa associata è:

$$p_N = \ln \sum_{\sigma} e^{-H_N(\sigma)},$$

funzione generatrice dei momenti rispetto alla distribuzione di Gibbs:

$$p(\sigma) = \frac{e^{-H_N(\sigma)}}{\sum_{\sigma} e^{-H_N(\sigma)}}$$

Derivando la funzione pressione rispetto a $V_i^{(l)}$, si ricava la probabilità che l scelga i . Vogliamo dimostrare che l'espressione che si ottiene in questo modo è la medesima ricavata dal modello Logit Multinomiale. Innanzitutto si esplicita p_N :

$$\begin{aligned}
p_N &= \ln \sum_{\sigma} e^{-H_N(\sigma)} = \ln \sum_{\sigma} e^{\sum_{l=1}^N v^{(l)} \cdot \sigma^{(l)}} \\
&= \ln \left(\sum_{\sigma^{(1)}} e^{v^{(1)} \cdot \sigma^{(1)}} \dots \sum_{\sigma^{(N)}} e^{v^{(N)} \cdot \sigma^{(N)}} \right) \\
&= \ln \prod_{l=1}^N \sum_{i=1}^k e^{V_i^{(l)}} = \sum_{l=1}^N \ln \sum_{i=1}^k e^{V_i^{(l)}} \quad (1.11)
\end{aligned}$$

Derivando tale espressione si ha:

$$p_{i,l} = \frac{\partial p_N}{\partial V_i^{(l)}} = \frac{e^{V_i^{(l)}}}{\sum_{j=1}^k e^{V_j^{(l)}}},$$

che coincide esattamente con (1.4), da cui l'equivalenza dei due modelli. Abbiamo quindi dimostrato che, pur partendo da ipotesi diverse ed utilizzando diversi strumenti e linguaggi, la distribuzione che si ricava dal modello Logit Multinomiale della DCT coincide con la misura di Boltzmann-Gibbs.

Vediamo, inoltre, come sia possibile interpretare il campo esterno h nei termini della DCT. Poichè in DCT ad ogni agente i si associano delle caratteristiche, si definisce un vettore $a_i = \{a_i^{(1)}, a_i^{(2)}, \dots, a_i^{(k)}\}$, che le rappresenta. Per semplicità, si considera il caso in cui anche le $a_i^{(j)}$ sono di tipo binario, ad esempio:

$$a_i^{(1)} = \begin{cases} 1 & \text{se } i \text{ è donna} \\ 0 & \text{se } i \text{ è uomo} \end{cases} \quad a_i^{(2)} = \begin{cases} 1 & \text{se } i \text{ è italiano} \\ 0 & \text{se } i \text{ è straniero} \end{cases} \quad a_i^{(3)} = \dots$$

Si può così esprimere h_i come:

$$h_i = h^{(0)} + \sum_{j=1}^h h^{(j)} a_i^{(j)}. \quad (1.12)$$

I parametri del modello sono le componenti del vettore $\mathbf{h} = (h^{(0)}, h^{(1)}, \dots, h^{(k)})$, dove $h^{(0)}$ è comune a tutti gli individui, mentre le altre componenti sono pesate con le caratteristiche specifiche di ciascuno. Va notato come gli elementi

di \mathbf{h} non dipendano da i , bensì dalla partizione della popolazione generata dal raggruppamento secondo i k attributi considerati.

Capitolo 2

Il modello di Curie-Weiss

Nel 1907 il fisico francese Pierre Weiss [51] introdusse e sviluppò la teoria del campo molecolare, che costituisce il primo studio moderno dei fenomeni magnetici, visti come fenomeni cooperativi. Riuscì così a spiegare le osservazioni sperimentali raccolte da Pierre Curie sul comportamento dei ferromagneti, cioè i materiali capace di ritenere un certo grado di magnetizzazione dopo essere stati esposti ad un campo magnetico.

Nel 1895 Pierre Curie [12] aveva infatti notato l'esistenza di una temperatura critica al di sopra della quale un ferromagnete, perde bruscamente la magnetizzazione. Rappresentando graficamente l'andamento della magnetizzazione in funzione della temperatura si può osservare chiaramente un punto di irregolarità, che dipende dal materiale in esame. La temperatura corrispondente alla singolarità viene detta *temperatura di Curie* per un dato materiale.

Weiss trovò l'origine di questo fenomeno considerando le interazioni tra gli atomi del ferromagnete. Egli assunse che gli spin interagissero tra loro attraverso un campo molecolare proporzionale alla media della magnetizzazione (da qui **teoria di campo medio**). Le interazioni J_{ij} che compaiono nell'Hamiltoniana (1.6), che si riporta per chiarezza:

$$H_N = - \sum_{i,j=1}^N J_{ij} \sigma_i \sigma_j - \sum_{i=1}^N h_i \sigma_i$$

vengono sostituite con un valore uniforme $J/2N$, e la simmetria che ne consegue consente di passare da un problema multy-body ad uno one-body.

Assumere che tutte le particelle interagiscano tra loro con la medesima intensità è un'assunzione molto forte, che se da un lato consente di ridurre drasticamente i parametri del modello, dall'altro comporta inevitabilmente una perdita di accuratezza. Tuttavia, per i modelli sociofisici, questa risulta essere un'ipotesi sensata ed il modello di Curie-Weiss costituisce un buon compromesso tra semplicità e precisione.

Negli anni '90, William Brock e Steven Durlauf approfondirono gli studi di Föllmer e svilupparono dei nuovi modelli, apportando l'idea delle interazioni tra agenti ai modelli della DCT [8] ed unificando in un quadro coerente modelli DCT e modelli meccanico statistici [9] per studiare problemi di natura socio-economica.

2.1 Il modello

Come nel caso interagente generale, si considera un insieme di N individui, indicizzati da i , $i = 1, \dots, N$, che affrontano una scelta di tipo binario, rappresentata da $\sigma_i \in \{-1, 1\}$.

Si pone $h_i = h$ e $J_{ij} = J/2N$, $\forall i = 1, \dots, N$, ovvero si assume che il campo esterno e le interazioni siano di tipo uniforme; il parametro J viene detto *costante di accoppiamento*.

L'hamiltoniana (1.6) assume così la seguente forma semplificata:

$$H_N = -\frac{J}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i \quad (2.1)$$

Si definisce anche qui la funzione di partizione canonica $Z_N(J, h)$:

$$Z_N(J, h) = \sum_{\sigma \in \Omega_N} e^{-H_N(\sigma)}$$

dove, come per l'Hamiltoniana generale, β è 'assorbito' dagli altri parametri J e h . La probabilità di una certa configurazione $\sigma = (\sigma_1, \dots, \sigma_N)$ è data

dalla misura di Boltzmann-Gibbs definita in (1.7), per cui:

$$P_{N,J,h}(\sigma) = \frac{e^{-H_N(\sigma)}}{Z_N(J, h)}$$

L'osservabile principale del modello di Curie-Weiss è la **magnetizzazione** di una configurazione, interpretabile come la 'scelta media' degli individui:

$$m_N(\sigma) \equiv \frac{1}{N} \sum_{i=1}^N \sigma_i \quad (2.2)$$

L'Hamiltoniana (2.1) può essere espressa in funzione di $m_N(\sigma)$:

$$H_N(\sigma) = -N \left(\frac{J}{2} m_N(\sigma)^2 + h m_N(\sigma) \right). \quad (2.3)$$

Assegnata l'Hamiltoniana, per qualsiasi osservabile del modello $f(\sigma)$, è possibile definirne il valor medio rispetto alla distribuzione di Gibbs definita in (1.7), detto **stato di Gibbs**:

$$\omega_N(f) \equiv \frac{\sum_{\sigma} f(\sigma) e^{-H_N(\sigma)}}{\sum_{\sigma} e^{-H_N(\sigma)}} = \frac{\sum_{\sigma} f(\sigma) e^{-H_N(\sigma)}}{Z_N(J, h)} \quad (2.4)$$

Saremo quindi interessati a determinare lo stato di Gibbs della magnetizzazione. A tale scopo si introduce la funzione **pressione**, le cui derivate rispetto al campo esterno h , forniscono i momenti della magnetizzazione; si definisce come:

$$p_N(J, h) = \frac{1}{N} \ln Z_N(J, h). \quad (2.5)$$

Derivando (2.5) rispetto al campo esterno h si ottiene:

$$\frac{\partial p_N(J, h)}{\partial h} = \frac{\sum_{\sigma} m_N(\sigma) e^{-H_N(\sigma)}}{Z_N(J, h)} = \omega_N(m_N(\sigma)) \quad (2.6)$$

La derivata seconda della pressione (2.5) rispetto ad h , è detta **suscettività**; vedremo in seguito che questa è un'altra osservabile fondamentale per la risoluzione del problema inverso:

$$\chi = \frac{\partial^2 p_N(J, h)}{\partial h^2} \quad (2.7)$$

Sviluppando la (2.7) si ricava:

$$\begin{aligned}
 \chi &= \frac{\partial \omega_N(m_N)}{\partial h} = \frac{\sum_{\sigma} m_N(\sigma) e^{-H(\sigma)}}{Z_N(J, h)} \\
 &= \frac{\sum_{\sigma} m_N^2(\sigma) e^{-H(\sigma)} Z_N(J, h) - \sum_{\sigma} m_N(\sigma) Z_N(J, h) e^{-H(\sigma)}}{Z_N(J, h)^2} \\
 &= N[\omega_N(m_N^2(\sigma)) - \omega_N(m_N(\sigma))^2] = N \text{Var}(\sigma_i)
 \end{aligned} \tag{2.8}$$

2.2 Limite termodinamico

Affinchè il modello sia ben definito, è necessario che l'Hamiltoniana sia una quantità intensiva rispetto al numero di spin, cioè che l'energia non dipenda dalle dimensioni del sistema, ma soltanto dalla natura e dalle condizioni nelle quali esso si trova.

A tale scopo si mostra che esiste finito il limite per N che tende all'infinito della funzione pressione. Tale problema, inizialmente trascurato, venne studiato in maniera rigorosa a partire dagli anni '60 e ricevette nuovi stimoli in seguito al modello introdotto da Sherrington e Kirkpatrick nel 1975 [43].

Nei prossimi paragrafi si dimostra dapprima l'esistenza del limite termodinamico della pressione e poi, una volta mostrata l'importante proprietà di fattorizzazione delle funzioni di correlazione, se ne calcola il valore esatto. Si illustra, infine, la soluzione del modello.

2.2.1 Esistenza del limite termodinamico

Esistono diversi metodi per dimostrare l'esistenza del limite termodinamico per il modello di Curie-Weiss. E' possibile, ad esempio, sfruttare la convessità dell'Hamiltoniana per mostrare la subadditività in N della pressione del sistema e quindi l'esistenza del limite, come mostrato in [2]. Tuttavia l'applicabilità di tale metodo è limitata: non è valida nei casi in cui l'Hamiltoniana non è convessa, come vedremo, ad esempio per il modello di Curie-Weiss multipopolato. Illustriamo quindi un suo raffinamento, detta tecnica di interpolazione, che sfrutta un teorema di esistenza per i modelli di campo

medio enunciato in [5], per cui l'esistenza della pressione per particella per grandi volumi è assicurata da una condizione di monotonia dell'Hamiltoniana. Questo metodo ha permesso a Guerra, [26], di dimostrare l'esistenza del limite termodinamico per il modello di Sherrington-Kirkpatrick.

Proposizione 2.1. *Data un'Hamiltoniana (2.1) di Curie-Weiss e la relativa pressione p_N , fissati i parametri J e h , esiste una funzione $p(J, h)$ tale per cui:*

$$\lim_{N \rightarrow \infty} p_N = p(J, h)$$

Dimostrazione. Dal teorema dimostrato in appendice (B) si ha che, data un'Hamiltoniana H_N tale per cui:

$$\omega_N(H_N) \geq \omega_N(H_{N_1}) + \omega_N(H_{N_2}), \quad (2.9)$$

per ogni partizione di $\{1, \dots, N\}$ in $\{1, \dots, N_1\}$ e $\{N_1+1, \dots, N_2\}$, il modello ammette limite termodinamico.

E' quindi sufficiente verificare che, per il modello di Curie-Weiss (2.1), vale la (2.9). Innanzitutto si definiscono le magnetizzazioni parziali:

$$m_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} \sigma_i \quad m_2 = \frac{1}{N_2} \sum_{i=N_1+1}^{N_2} \sigma_i$$

tali per cui $m = \frac{N_1}{N} m_1 + \frac{N_2}{N} m_2$. Si ha:

$$\begin{aligned} H_N - H_{N_1} - H_{N_2} &= -N \left(\frac{J}{2} m^2 + h m \right) + N_1 \left(\frac{J}{2} m_1^2 + h m_1 \right) + N_2 \left(\frac{J}{2} m_2^2 + h m_2 \right) \\ &= -N \left[\frac{J}{2} \left(m^2 - \frac{N_1}{N} m_1^2 - \frac{N_2}{N} m_2^2 \right) + h \left(m - \frac{N_1}{N} m_1 - \frac{N_2}{N} m_2 \right) \right] \\ &= -N \frac{J}{2} \left(m^2 - \frac{N_1}{N} m_1^2 - \frac{N_2}{N} m_2^2 \right). \end{aligned}$$

Dalla convessità della funzione $f(x) = x^2$ si ha quindi:

$$H_N \geq H_{N_1} + H_{N_2}$$

e poichè tale disuguaglianza vale per ogni configurazione σ , la (2.9) è soddisfatta, e la proposizione è dimostrata. □

2.2.2 Le proprietà di fattorizzazione

Prima di calcolare la soluzione esatta del limite termodinamico, si dimostra che le funzioni di correlazione fattorizzano completamente al limite per quasi ogni coppia di parametri. Tale proprietà risulterà essere fondamentale, in quanto implica che tutte le proprietà termodinamiche del sistema in esame possono essere descritte dalla magnetizzazione.

Proposizione 2.2. *Per quasi ogni scelta del parametro h , vale:*

$$\lim_{N \rightarrow \infty} (\omega_N(m^2) - \omega_N(m)^2) = 0$$

Dimostrazione. Si calcolano le derivate prima e seconda della funzione pressione, che si ricorda essere: $p_N = \frac{1}{N} \ln Z_N = \frac{1}{N} \ln \sum_{\sigma} e^{-H_N(\sigma)}$:

$$\frac{\partial p_n}{\partial h} = \frac{1}{N} \sum_{\sigma} N m(\sigma) \frac{e^{-H(\sigma)}}{Z_N} = \omega_N(m)$$

$$\frac{\partial^2 p_n}{\partial h^2} = \omega_N(m^2) - \omega_N(m)^2$$

Sfruttando queste equazioni si può maggiorare l'integrale seguente:

$$\begin{aligned} \left| \int_{h^{(1)}}^{h^{(2)}} (\omega_N(m^2) - \omega_N(m)^2) dh \right| &= \frac{1}{N} \left| \int_{h^{(1)}}^{h^{(2)}} \frac{\partial^2 p_n}{\partial h^2} dh \right| = \frac{1}{N} \left| \frac{\partial p_n}{\partial h} \right|_{h^{(1)}}^{h^{(2)}} \\ &\leq \frac{1}{N} (|\omega_N(m)|_{h^{(2)}} + |\omega_N(m)|_{h^{(1)}}) \\ &= O\left(\frac{1}{N}\right) \end{aligned} \quad (2.10)$$

Poichè la pressione $p = \lim_{N \rightarrow \infty} p_N$ è una funzione convessa, si ha che $\frac{\partial p_n}{\partial h}$ e $\frac{\partial p_n}{\partial J}$ ammettono quasi ovunque limite termodinamico ben definito, ed essendo:

$$\omega_N(m) = \frac{\partial p_n}{\partial h} \quad \omega_N(m^2) = \frac{\partial p_n}{\partial J},$$

la disuguaglianza in (2.10):

$$\lim_{N \rightarrow \infty} (\omega_N(m^2) - \omega_N(m)^2) = 0 \quad q.o. \text{ in } h \quad (2.11)$$

□

L'equazione (2.11) ha un significato molto importante, in quanto indica che la magnetizzazione $m(\sigma)$ è una quantità **automediante** (*self-averaging quantity*), ovvero le cui fluttuazioni si annullano al limite termodinamico. Da questa proprietà, sfruttando la seguente disuguaglianza di Cauchy-Schwartz, si ricava la proposizione (2.4)

Proposizione 2.3 (Disuguaglianza di Cauchy-Schwartz). *Siano X ed Y due variabili aleatorie definite in uno spazio di probabilità finito, tali per cui $P(X_i) = P(Y_i) = p_i$. Si ha che:*

$$\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$$

Dimostrazione. Siano:

$$\mu_X = \mathbb{E}(X) = \sum_i X_i p_i \quad \mu_Y = \mathbb{E}(Y) = \sum_i Y_i p_i$$

$$\sigma_X^2 = \text{Var}(X) \quad \sigma_Y^2 = \text{Var}(Y)$$

Si definiscono le nuove variabili aleatorie \bar{X} e \bar{Y} :

$$\bar{X} = \frac{X - \mu_X}{\sigma_X} \quad \bar{Y} = \frac{Y - \mu_Y}{\sigma_Y},$$

ed essendo $\{\bar{X}_i\}$ e $\{\bar{Y}_i\}$ vettori di modulo 1 si ha:

$$|\mathbb{E}(\bar{X}\bar{Y})| = \left| \sum_i \bar{X}_i \bar{Y}_i p_i \right| = \left| \sum_i \bar{X}_i p_i^{1/2} \bar{Y}_i p_i^{1/2} \right| \leq 1.$$

Andando a sostituire con X ed Y si ha il risultato.

□

Proposizione 2.4. *Dato un intero k , per quasi ogni scelta di h vale:*

$$\lim_{N \rightarrow \infty} (\omega_N(m^k) - \omega_N(m)^k) = 0$$

Dimostrazione. la dimostrazione segue direttamente dalla disuguaglianza di Cauchy-Schwartz, ponendo $X = m^{k-1}$ e $Y = m$:

$$|\omega_N(m^{k-1}m) - \omega_N(m^{k-1})\omega_N(m)| \leq \sqrt{\text{Var}_N(m^{k-1})\text{Var}_N(m)} \quad (2.12)$$

Poichè m è una quantità automediante, $\text{Var}_N(m)$ si annulla al limite, ed essendo m^{k-1} una quantità limitata:

$$\lim_{N \rightarrow \infty} (\omega_N(m^k) - \omega_N(m)^{k-1}\omega_N(m)) = 0$$

da cui, per induzione, la proposizione è dimostrata. \square

La proposizione seguente mostra come la proprietà di fattorizzazione della funzione di correlazione degli spin possa essere ricavata dalla fattorizzazione della magnetizzazione; di conseguenza le proprietà termodinamiche del sistema possono essere espresse in funzione della magnetizzazione.

Proposizione 2.5. *Dati due spin σ_i e σ_j , con $i \neq j$, e per q.o. h vale:*

$$\lim_{N \rightarrow \infty} (\omega_N(\sigma_i\sigma_j) - \omega_N(\sigma_i)\omega_N(\sigma_j)) = 0$$

Dimostrazione. Poichè stiamo considerando un modello di campo medio, si ha la proprietà di invarianza rispetto alle traslazioni della misura di Gibbs, per cui:

$$\omega_N(m) = \omega_N\left(\frac{1}{N} \sum_{i=1}^N \sigma_i\right) = \omega_N(\sigma_i) \quad (2.13)$$

$$\begin{aligned} \omega_N(m^2) &= \omega_N\left(\frac{1}{N^2} \sum_{i,j=1}^N \sigma_i\sigma_j\right) = \omega_N\left(\frac{1}{N^2} \sum_{i \neq j=1}^N \sigma_i\sigma_j\right) + \omega_N\left(\frac{1}{N^2} \sum_{i=j=1}^N \sigma_i\sigma_j\right) \\ &= \frac{N-1}{N} \omega_N(\sigma_1\sigma_2) + \frac{1}{N} \end{aligned} \quad (2.14)$$

Da quest'ultima e da (2.11) si ottiene:

$$\lim_{N \rightarrow \infty} \omega_N(\sigma_i\sigma_j) - \omega_N(\sigma_i)\omega_N(\sigma_j) = 0 \quad (2.15)$$

per q.o. h, e per ogni coppia (i, j) con $i \neq j$. \square

In modo del tutto analogo si mostra la proprietà di fattorizzazione per k spins, con $k > 2$. Come già detto, tale proprietà è valida quasi ovunque in h ; resta escluso il caso $h = 0$, con $J > 1$, in cui, tuttavia, si può ottenere una versione debole della regola di fattorizzazione, illustrata nella proposizione seguente.

Proposizione 2.6. *Dati quattro spin distinti $\sigma_i, \sigma_j, \sigma_k, \sigma_l$, per quasi ogni scelta di J , vale:*

$$\lim_{N \rightarrow \infty} \omega_N(\sigma_i \sigma_j \sigma_k \sigma_l) - \omega_N(\sigma_i \sigma_j) \omega_N(\sigma_k \sigma_l) = 0$$

Dimostrazione. Si dimostra facilmente che anche l'energia per particella è una quantità automediante. Questa, infatti, si ricava come derivata della pressione rispetto a J ; da questo e dalla proposizione (2.4) si dimostra la proposizione. \square

A questo punto, mostrate le proprietà di fattorizzazione della magnetizzazione e delle correlazioni, si può passare al calcolo della soluzione esatta del modello, individuando i bounds superiore ed inferiore della pressione al limite termodinamico.

2.2.3 I bounds della funzione pressione

Bound superiore

Per ricavare il bound superiore della funzione pressione al limite termodinamico si utilizza una tecnica di 'disaccoppiamento', per cui lo spazio delle configurazioni viene partizionato in sottoinsiemi di stati con la stessa magnetizzazione, come in [15], [27], [28].

Proposizione 2.7. *Data un'Hamiltoniana di Curie-Weiss (2.1) e la relativa pressione $p_N = \frac{1}{N} \ln Z$, fissati i parametri J e h , vale la seguente disuguaglianza:*

$$\lim_{N \rightarrow \infty} \sup p_N \leq \sup_{\bar{m}} p_U,$$

dove

$$p_U(\bar{m}) = \ln 2 - \frac{J}{2}\bar{m}^2 + \ln \cosh(J\bar{m} + h) \quad \bar{m} \in [-1, 1]$$

Dimostrazione. Lo spettro dei possibili valori di m è il seguente:

$$S_{m_N} = \left\{-1, -1 + \frac{2}{N}, \dots, 1 - \frac{2}{N}, 1\right\}$$

La funzione di partizione Z_N può quindi essere riscritta come:

$$Z_N = \sum_{\sigma} \exp \left[N \left(\frac{J}{2} m^2 + hm \right) \right] = \sum_{\sigma} \sum_{\bar{m} \in S_{m_N}} \delta_{m, \bar{m}} \exp \left[N \left(\frac{J}{2} m^2 + hm \right) \right] \quad (2.16)$$

dove $\delta_{x,y}$ è la delta di Kronecker, per cui vale: $\sum_{\bar{m} \in S_{m_N}} \delta_{m, \bar{m}}$.

Utilizzando la seguente relazione

$$(m - \bar{m})^2 = 0 \quad \rightarrow \quad m^2 = 2m\bar{m} - \bar{m}^2$$

si può sostituire m^2 linearizzando il termine quadratico, in quanto \bar{m} non dipende dalla configurazione di spin. Maggiorando poi la delta di Kronecker con il valore 1 si ricava:

$$\begin{aligned} Z_N &= \sum_{\sigma} \sum_{\bar{m} \in S_m} \delta_{m, \bar{m}} \exp N \left(\frac{J}{2} (2m\bar{m} - \bar{m}^2) + hm \right) \\ &\leq \sum_{\sigma} \sum_{\bar{m} \in S_m} \exp N \left(\frac{J}{2} (2m\bar{m} - \bar{m}^2) + hm \right) \end{aligned}$$

Scambiando le due sommatorie ed eseguendo la sommatoria in σ , che fattorizza grazie alla linearità delle interazioni rispetto a m , si ottiene:

$$Z_N \leq \sum_{\bar{m} \in S_m} G(\bar{m})$$

dove

$$G(\bar{m}) \equiv e^{-N \frac{1}{2} J \bar{m}^2} 2^N (\cosh(J\bar{m} + h))^N \quad (2.17)$$

Poichè lo spettro delle magnetizzazioni ha cardinalità finita $|S_{m_N}| = N + 1$, si ha:

$$Z_N \leq (N + 1) \sup_{\bar{m}} G \quad (2.18)$$

Da questa disuguaglianza si ottiene così il bound superiore per la pressione $p_N = \frac{1}{N} \ln Z_N$:

$$p_N \leq \frac{1}{N} \ln[(N + 1) \sup_{\bar{m}} G] = \frac{1}{N} \ln(N + 1) + \frac{1}{N} \sup_{\bar{m}} \ln G. \quad (2.19)$$

Definendo quindi la funzione p_U , indipendente da N :

$$p_U(\bar{m}) = \frac{1}{N} \ln G = \ln 2 - \frac{J}{2} \bar{m}^2 + \ln \cosh(J\bar{m} + h)$$

e ricordando che $\lim_{N \rightarrow \infty} \frac{1}{N} \ln(N + 1) = 0$, al limite termodinamico si ottiene:

$$\lim_{N \rightarrow \infty} \sup p_N \leq \sup_{\bar{m}} p_U(\bar{m}) \quad (2.20)$$

□

Bound inferiore dalla convessità dell'Hamiltoniana

Proposizione 2.8. *Data un'Hamiltoniana di Curie-Weiss (2.1) e la relativa pressione $p_N = \frac{1}{N} \ln Z$, fissati i parametri J e h , vale la seguente disuguaglianza:*

$$p_N \geq \sup_{-1 \leq \bar{m} \leq 1} p_L$$

dove

$$p_L(\bar{m}) = -\frac{J}{2} \bar{m}^2 + \ln 2 + \ln \cosh(J\bar{m} + h)$$

Dimostrazione. L'Hamiltoniana per particella può essere riscritta come funzione della magnetizzazione $m(\sigma)$:

$$\frac{H(\sigma)}{N} = -\frac{J}{2} m^2 - hm.$$

Preso un qualsiasi numero $\bar{m} \in [-1, 1]$, si ha che

$$(m - \bar{m})^2 \geq 0 \Rightarrow m^2 \geq 2m\bar{m} - \bar{m}^2,$$

da cui si ottiene:

$$\begin{aligned} p_N &= \frac{1}{N} \ln Z_N = \frac{1}{N} \ln \sum_{\sigma} \exp N\left(\frac{J}{2}m^2 + hm\right) \\ &\geq \frac{1}{N} \ln \sum_{\sigma} \exp N\left(Jm\bar{m} - \frac{J}{2}\bar{m}^2 + hm\right) \\ &= \frac{1}{N} \ln \left(\exp -\frac{NJ}{2}\bar{m}^2 \sum_{\sigma} \exp N(J\bar{m}m + hm) \right) \\ &= -\frac{J}{2}\bar{m}^2 + \frac{1}{N} \ln (2^N \cosh(J\bar{m} + h)^N) = -\frac{J}{2}\bar{m}^2 + \ln 2 + \ln \cosh(J\bar{m} + h) \end{aligned}$$

da cui

$$p_N \geq \sup_{-1 \leq \bar{m} \leq 1} p_L$$

dove

$$p_L = -\frac{J}{2}\bar{m}^2 + \ln 2 + \ln \cosh(J\bar{m} + h)$$

□

Bound inferiore dal principio variazionale

Il limite inferiore per la pressione al limite termodinamico può essere ottenuto anche attraverso un'altra tecnica, che sfrutta il principio variazionale entropico di Gibbs, [40].

Sostanzialmente l'idea è quella di minorare l'Hamiltoniana (2.1) con un'Hamiltoniana non interagente, restringendo così il problema variazionale ad una misura prodotto con un unico grado di libertà.

Si considera quindi:

$$\tilde{H} = -r \sum_{i=1}^N \sigma_i,$$

il cui stato di Gibbs corrispondente di un'osservabile $f(\sigma)$ è il seguente:

$$\tilde{\omega}(f) = \frac{\sum_{\sigma} f(\sigma) e^{-\tilde{H}(\sigma)}}{\sum_{\sigma} e^{-\tilde{H}(\sigma)}}$$

Il principio variazionale [40] può essere espresso nella seguente forma:

Proposizione 2.9 (Principio variazionale). *Siano date un'Hamiltoniana H e la sua funzione di partizione $Z = \sum_{\sigma} e^{-H}$. Presa un'arbitraria Hamiltoniana \tilde{H} con relativa funzione di partizione \tilde{Z} , vale la seguente disuguaglianza:*

$$\ln Z \geq \ln \tilde{Z} - \tilde{\omega}(H) + \tilde{\omega}(\tilde{H}). \quad (2.21)$$

Data l'Hamiltoniana (2.1) del modello di Curie-Weiss, con funzione pressione p_N , vale la seguente disuguaglianza:

$$\liminf_{N \rightarrow \infty} p_N \geq \sup_{\bar{m}} p'_L \quad (2.22)$$

con

$$p'_L \equiv \frac{J}{2} \bar{m}^2 + h \bar{m} - \frac{1 + \bar{m}}{2} \ln(1 + \bar{m}) - \frac{1 - \bar{m}}{2} \ln(1 - \bar{m}) + \ln 2 \quad (2.23)$$

con $\bar{m} \in [-1, 1]$

Dimostrazione. la disuguaglianza (2.21) segue direttamente dalla disuguaglianza di Jensen:

$$e^{\tilde{\omega}(-H+\tilde{H})} \leq \tilde{\omega}(e^{-H+\tilde{H}}). \quad (2.24)$$

Data l'Hamiltoniana (2.1), lo stato $\tilde{\omega}(H)$ è pari a:

$$\begin{aligned} \tilde{\omega}(H) &= -\frac{J}{2N} \sum_{i,j} \tilde{\omega}(\sigma_i, \sigma_j) - h \sum_i \tilde{\omega}(\sigma_i) \\ &= -N \left(1 - \frac{1}{N}\right) \frac{J}{2} (\tanh r)^2 - N \frac{J}{2} - Nh \tanh r \end{aligned} \quad (2.25)$$

dove (2.25) deriva dal calcolo dei momenti di un sistema non interagente.

In maniera analoga si calcola lo stato $\tilde{\omega}$ di \tilde{H} :

$$\tilde{\omega}(\tilde{H}) = -Nr \tanh r$$

e la rispettiva funzione di partizione \tilde{Z}_N :

$$\tilde{Z}_N = \sum_{\sigma} e^{-\tilde{H}(\sigma)} = 2^N (\cosh r)^N,$$

da cui:

$$\tilde{p}_N = \frac{1}{N} \ln \tilde{Z}_N = \ln 2 + \ln \cosh r.$$

Applicando la (2.21) si ha:

$$\begin{aligned} p_N &\geq \frac{1}{N} \left(\ln \tilde{Z}_N - \tilde{\omega}(H) + \omega(\tilde{H}) \right) = \\ &= \ln 2 + \ln \cosh r + \frac{J}{2} (\tanh r)^2 + h \tanh r - r \tanh r + \frac{J}{2} N - \frac{J}{N} (\tanh r)^2 \end{aligned} \quad (2.26)$$

Prendendo il limite inferiore sugli N e l'estremo superiore in r del membro di sinistra, con un cambio di variabili $\bar{m} = \tanh r$ si ottiene:

$$\begin{aligned} p'_L(\bar{m}) &= \frac{J}{2} \bar{m}^2 + h \bar{m} - \frac{1 + \bar{m}}{2} \ln \left(\frac{1 + \bar{m}}{2} \right) - \frac{1 - \bar{m}}{2} \ln \left(\frac{1 - \bar{m}}{2} \right) \\ &= \frac{J}{2} \bar{m}^2 + h \bar{m} - \frac{1 + \bar{m}}{2} \ln(1 + \bar{m}) - \frac{1 - \bar{m}}{2} \ln(1 - \bar{m}) + \ln 2 \end{aligned}$$

□

2.2.4 L'equazione di campo medio

Abbiamo mostrato che i limiti della funzione pressione, al limite termodinamico, sono dati dall'estremo superiore rispetto ad \bar{m} delle seguenti funzioni:

- $p_U(\bar{m}) = p_L(\bar{m}) = \ln 2 - \frac{J}{2} \bar{m}^2 + \ln \cosh(J\bar{m} + h);$
- $p'_L = \frac{J}{2} \bar{m}^2 + h \bar{m} - \frac{1 + \bar{m}}{2} \ln(1 + \bar{m}) - \frac{1 - \bar{m}}{2} \ln(1 - \bar{m}) + \ln 2$

Essendo $p_U = p_L$, l'estremo superiore di tale funzione coincide con il limite termodinamico della pressione e fornisce la soluzione esatta del modello; è necessario verificare che tale limite coincida effettivamente con quello dato da p'_L .

Per il modello di Curie-Weiss ad una popolazione, l'equivalenza dei due bounds può essere provata utilizzando la trasformata di Legendre (definita in (2.27)).

Proposizione 2.10 (Trasformata di Legendre). *La funzione*

$$f^*(y) = \frac{1}{J} \left(\frac{1+y}{2} \ln \frac{1+y}{2} + \frac{1-y}{2} \ln \frac{1-y}{2} - yh \right)$$

è la trasformata di Legendre della funzione:

$$f(x) = \frac{1}{J} \ln 2 \cosh(Jx + h)$$

Dimostrazione. La trasformata di Legendre di una generica funzione f si definisce come:

$$f^*(y) = \sup_x (xy - f(x)) \quad (2.27)$$

Poichè la funzione f considerata è convessa, il suo estremo si può individuare annullando la derivata, ovvero imponendo:

$$\frac{df}{dx} = y - \tanh(Jx + h) = 0$$

da cui

$$Jx = \operatorname{arctanh} y - h.$$

Sostituendo in (2.27) si ottiene:

$$\begin{aligned}
f^*(x) &= y \frac{1}{J} (\operatorname{arctanh} y - h) - \frac{1}{J} \ln 2 \cosh(\operatorname{arctanh} y - h + h) \\
&= \frac{y}{J} \operatorname{arctanh} y - \frac{yh}{J} - \frac{1}{J} \ln 2 \cosh \operatorname{arctanh} y \\
&= \frac{y}{2J} \ln \frac{1+y}{1-y} - \frac{yh}{J} - \frac{1}{J} \ln \left(\exp\left\{\frac{1}{2} \ln \frac{1+y}{1-y}\right\} + \exp\left\{\frac{1}{2} \ln \frac{1-y}{1+y}\right\} \right) \\
&= \frac{y}{2J} \ln \frac{1+y}{1-y} - \frac{yh}{J} - \frac{1}{J} \ln \left(\frac{1+y+1-y}{\sqrt{1-y^2}} \right) \\
&= \frac{y}{2J} \ln \frac{1+y}{1-y} - \frac{yh}{J} - \frac{1}{J} \ln \left(\frac{2}{\sqrt{1-y^2}} \right) \\
&= \frac{1}{J} \left(\frac{1+y}{2} \ln(1+y) + \frac{1-y}{2} \ln(1-y) - yh - \ln 2 \right) \\
&= \frac{1}{J} \left(\frac{1+y}{2} \ln \frac{1+y}{2} + \frac{1-y}{2} \ln \frac{1-y}{2} - yh \right)
\end{aligned}$$

□

In maniera del tutto analoga si può dimostrare che la funzione

$$g^*(x) = \frac{1}{2}x^2$$

è la trasformata di Legendre di:

$$g(x) = -\frac{1}{2}x^2.$$

Di conseguenza, si ha che

$$\begin{aligned}
p_U(\bar{m}) &= p_L(\bar{m}) = J(f(\bar{m}) - g(\bar{m})), \\
p'L(\bar{m}) &= J(g^*(\bar{m}) - f^*(\bar{m}))
\end{aligned} \tag{2.28}$$

Utilizzando la seguente proposizione si ha quindi l'equivalenza dei bounds.

Proposizione 2.11. *Siano f e g due funzioni convesse e siano f^* e g^* le rispettive trasformate di Legendre. Vale:*

$$\sup_x (f(x) - g(x)) = \sup_y (g^*(y) - f^*(y))$$

Dimostrazione. Si veda [18] o [50]. □

I risultati finora mostrati si possono riassumere con il seguente teorema:

Teorema 2.12. *Data l'Hamiltoniana di Curie-Weiss (2.1) e la relativa pressione per particella p_N , fissati i parametri J e h , il limite termodinamico:*

$$\lim_{N \rightarrow \infty} p_N = p$$

esiste. Tale limite può essere espresso equivalentemente come:

$$(i) \quad p = \sup_{\bar{m}} p_U(\bar{m}) = \sup_{\bar{m}} p_L(\bar{m});$$

$$(ii) \quad p = \sup_{\bar{m}} p'_L(\bar{m}).$$

Quando le funzioni $p_U = p_L$ e p'_L ammettono un unico punto di massimo, allora tale massimo coincide proprio con il limite termodinamico della magnetizzazione. Se i punti di massimo sono più di uno, allora si ha transizione di fase e ciascun argomento degli estremi fornisce uno stato puro per m .

Calcoliamo i punti stazionari μ di $p_U = p_L$ e p'_L imponendo l'annullamento delle derivate:

$$\frac{\partial p_U(\mu)}{\partial \mu} = -J\mu + \frac{1}{\cosh(J\mu + h)} \sinh(J\mu + h)J = 0$$

$$\Rightarrow \mu = \tanh(J\mu + h)$$

$$\frac{\partial p'_L(\mu)}{\partial \mu} = J\mu + h - \frac{1}{2} \ln(1 + \mu) + \frac{1}{2} \ln(1 - \mu) = J\mu + h - \frac{1}{2} \ln \left(\frac{1 + \mu}{1 - \mu} \right) = 0$$

E dalla relazione $\frac{1}{2} \ln \left(\frac{1+x}{1-x} \right) = \tanh^{-1}(x)$, si ottiene l'espressione di punto fisso precedente.

Riassumendo, i punti stazionari delle funzioni bounds della pressione soddisfano la seguente espressione, detta **equazione di campo medio** o **equazione di consistenza**:

$$\mu = \tanh(J\mu + h) \tag{2.29}$$

Proposizione 2.13. *Siano J e h tali per cui $p_U = p_L$ ammette un unico punto di massimo in μ^* . Allora vale:*

$$\lim_{N \rightarrow \infty} \omega_N(m) = \lim_{N \rightarrow \infty} \omega_N(\sigma_i)$$

Dimostrazione. Dalla definizione della funzione pressione $p_N(J, h)$, per N finito si ha:

$$\frac{\partial p_N}{\partial h} = \omega_N(m_N)$$

Poichè p_N è una serie convergente di funzioni convesse, anche la funzione limite è convessa e differenziabile quasi ovunque. Di conseguenza vale:

$$\lim_{N \rightarrow \infty} \omega_N(m) = \lim_{N \rightarrow \infty} \frac{\partial p_N}{\partial h} = \frac{\partial \sup_{\mu} p_L}{\partial h}$$

quando l'ultima derivata esiste.

Quindi, ponendo $p(J, h, \mu^*(J, h)) \equiv \lim_{N \rightarrow \infty} p_N$, si ha:

$$\frac{\partial \sup_{\mu} p_L}{\partial h} = \frac{\partial p(J, h, \mu^*(J, h))}{\partial h} = -J \frac{\partial \mu^*}{\partial h} + \tanh(J\mu^* + h) + J \frac{\partial \mu^*}{\partial h} \tanh(J\mu^* + h),$$

e, utilizzando l'equazione di campo medio (2.29), si ricava:

$$\frac{\partial \sup_{\mu} p_L}{\partial h} = \mu^*,$$

da cui il risultato. □

Lo stesso risultato può essere provato per la funzione p'_L .
Siano:

$$\omega(m) = \lim_{N \rightarrow \infty} \omega_N(m)$$

$$\omega(\sigma_i) = \lim_{N \rightarrow \infty} \omega_N(\sigma_i)$$

Come conseguenza della proposizione (2.13) si può riscrivere p'_L come:

$$p'_L(\mu^*) = S - U,$$

dove S è l'entropia termodinamica:

$$S = -\frac{1 + \omega(\sigma_i)}{2} \ln \left(\frac{1 + \omega(\sigma_i)}{2} \right) - \frac{1 - \omega(\sigma_i)}{2} \ln \left(\frac{1 - \omega(\sigma_i)}{2} \right),$$

mentre U è l'energia interna:

$$U = \frac{J}{2} \omega(m)^2 + h\omega(m).$$

2.2.5 La transizione di fase

Per una descrizione completa del modello di Curie-Weiss, è interessante studiare la sua transizione di fase, che può essere individuata cercando i punti di non-analiticità del limite termodinamico della pressione per $h \in \mathbb{R}$ e $J \in \mathbb{R}^+$.

Quando $h \neq 0$ non si osserva alcun punto di irregolarità, per cui non si ha alcuna transizione di fase. La situazione cambia nel caso $h = 0$: dal teorema del passaggio al limite sotto segno di derivata si ha:

$$\lim_{N \rightarrow \infty} p_N(J, 0) = \begin{cases} 0 & J \leq 1 \\ \frac{J}{2} \mu_0^2 + \ln \cosh(J\mu_0) & J > 1 \end{cases}$$

Quando $J \rightarrow 1^+$ la magnetizzazione μ_0 , e di conseguenza il limite della pressione, sono continui per ogni valore di J . Si cerca quindi la discontinuità differenziando il limite rispetto a J . Si ha:

$$\begin{aligned} \frac{\partial}{\partial J} \left(\lim_{N \rightarrow \infty} p_N(J, h) \right) &= -\frac{J}{2} \mu^2 - J\mu \frac{\partial \mu}{\partial J} + \tanh(J\mu + h) \left(\mu + J \frac{\partial \mu}{\partial J} \right) \\ &= \frac{1}{2} \mu^2. \end{aligned}$$

Quando il campo esterno è nullo si ha quindi:

$$\frac{\partial}{\partial J} \left(\lim_{N \rightarrow \infty} p_N(J, 0) \right) = \begin{cases} 0 & J \leq 1 \\ \frac{1}{2} \mu_0^2 & J > 1. \end{cases}$$

La funzione è ancora continua in J . Differenziando ulteriormente, si ha che la derivata seconda del limite della pressione vale:

$$\frac{\partial^2}{\partial J^2} \left(\lim_{N \rightarrow \infty} p_N(J, h) \right) = \mu \frac{\partial \mu}{\partial J}.$$

Poichè si può riscrivere

$$\mu \frac{\partial \mu}{\partial J} = \frac{1}{2} \frac{\partial \mu^2}{\partial J}, \quad (2.30)$$

nel caso $h = 0$ si ottiene:

$$\frac{\partial^2}{\partial J^2} \left(\lim_{N \rightarrow \infty} p_N(J, 0) \right) = \begin{cases} 0 & \text{when } J \leq 1 \\ \frac{1}{2} \frac{d\mu_0^2}{dJ} & \text{when } J > 1. \end{cases} \quad (2.31)$$

Utilizzando l'equazione di campo medio (2.29), si può espandere la tangente iperbolica, ricavando:

$$\mu_0 = J\mu_0 - \frac{(J\mu_0)^3}{3} + O(\mu_0^5) \quad J \rightarrow 1^+. \quad (2.32)$$

Poichè μ_0 non si annulla per $J > 1$, si può dividere per $J\mu_0$ l'equazione (2.32). Si ricava:

$$\frac{1}{J} = 1 - \frac{(J\mu_0)^2}{3} + O(\mu_0^4) \quad J \rightarrow 1^+.$$

Da cui:

$$\mu_0 \sim \left(\frac{3}{(J)^2} \left(1 - \frac{1}{J} \right) \right)^{\frac{1}{2}} \sim \left(3 \left(1 - \frac{1}{J} \right) \right)^{\frac{1}{2}} \quad J \rightarrow 1^+ \quad (2.33)$$

e quindi, sostituendo quest'espressione in (2.31), il caso $J > 1$ può essere approssimato come:

$$\frac{1}{2} \frac{d\mu_0^2}{dJ} \sim \frac{1}{2} \frac{d}{dJ} \left(3 \left(1 - \frac{1}{J} \right) \right) = \frac{3}{2J^2} \quad J \rightarrow 1^+. \quad (2.34)$$

da cui segue la discontinuità del limite termodinamico (2.31). Ciò significa che si ha una transizione di fase del secondo ordine per $h = 0$ e $J = 1$. Nel paragrafo successivo, in cui si rappresenta graficamente l'equazione di campo medio e si visualizza la magnetizzazione come superficie in funzione di J e h , si può osservare questa transizione di fase come discontinuità del grafico (2.3).

Studiamo ora il comportamento della suscettività $\chi = \frac{\partial^2 p_N}{\partial h^2}(J, h)$ nel caso $h = 0$, $J \rightarrow 1^+$. Al limite termodinamica vale:

$$\chi = \frac{\partial \mu}{\partial h} = \frac{1 - \mu^2}{1 - J(1 - \mu^2)},$$

per cui, sfruttando lo sviluppo (2.33) si ricava che, quando il campo esterno è nullo, al limite termodinamico vale:

$$\begin{aligned} \chi &= \frac{1 - \mu_0^2}{1 - J(1 - \mu_0^2)} \sim \frac{1 - (3(1 - \frac{1}{J}))}{1 - J(1 - (3(1 - \frac{1}{J})))} \\ &= \frac{-2 + \frac{3}{J}}{2J - 2} = \frac{3 - 2J}{J(2J - 2)}. \end{aligned}$$

Si ha quindi:

$$\chi \rightarrow +\infty \quad J \rightarrow 1^+ \quad (2.35)$$

2.2.6 Studio grafico dell'equazione di campo medio

Abbiamo visto che quando il punto di massimo della funzione $p_U = p_L$ o di p'_L è unico, esso coincide con il limite termodinamico della magnetizzazione, mentre se sono in numero superiore ad uno si ha transizione di fase ed i punti di massimo corrispondono ad uno stato puro.

Le soluzioni dell'equazione di campo medio possono essere visualizzate graficamente come punti di intersezione tra la bisettrice $y = \mu$ e la tangente iperbolica $y = \tanh(J\mu + h)$. A seconda dei valori assunti dai parametri del

modello, J ed h , si possono stabilire quante soluzioni ammette (2.29). Si possono distinguere i seguenti casi:

- $h \neq 0$ e $J > 0$: la soluzione μ_h è unica, diversa da zero, di segno concorde al campo esterno h e coincide con l'unico massimo di p_U ;

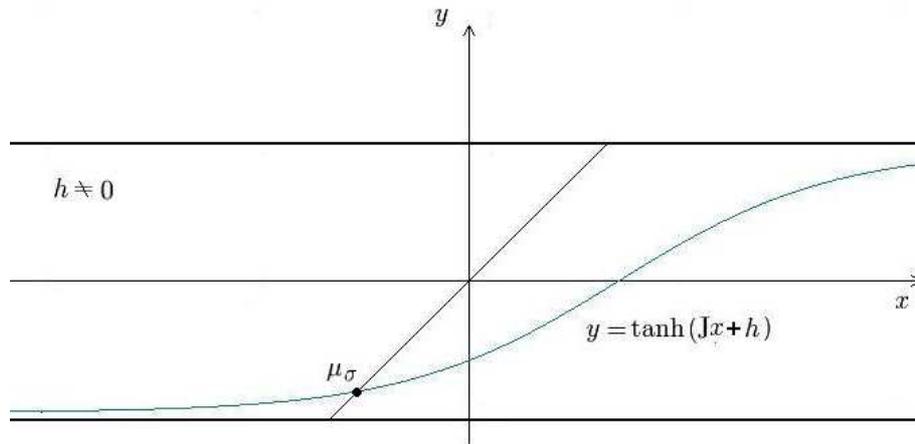


Figura 2.1: Rappresentazione grafica dell'equazione di consistenza - $h \neq 0$

- $h = 0$
 - $0 \leq J \leq 1$: la soluzione è unica, uguale a zero, unico massimo di p_U ;
 - $J > 1$: le soluzioni sono tre, zero e $\pm\mu_0$, dove questi ultimi sono i massimi di p_U .

In figura (2.3) si rappresenta, invece, l'equazione di campo medio come superficie, dove la magnetizzazione è funzione dei parametri J e h . Si può osservare qui e in figura (2.4), che per $h = 0$ e $J > 1$, la superficie si 'sdoppia'. Essendo, al limite, $\chi = \frac{\partial \mu}{\partial h}$, possiamo visualizzare la suscettività come la tangente al grafico in (2.4): per $h = 0$, $J \rightarrow 1^+$, in accordo con quanto calcolato in (2.35), per cui $\chi = +\infty$, si vede che la tangente tende alla retta verticale $J = 1$.

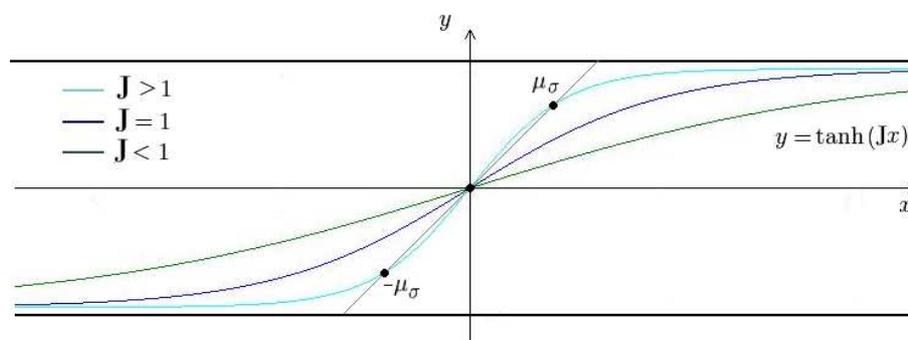


Figura 2.2: Rappresentazione grafica dell'equazione di consistenza - $h = 0$

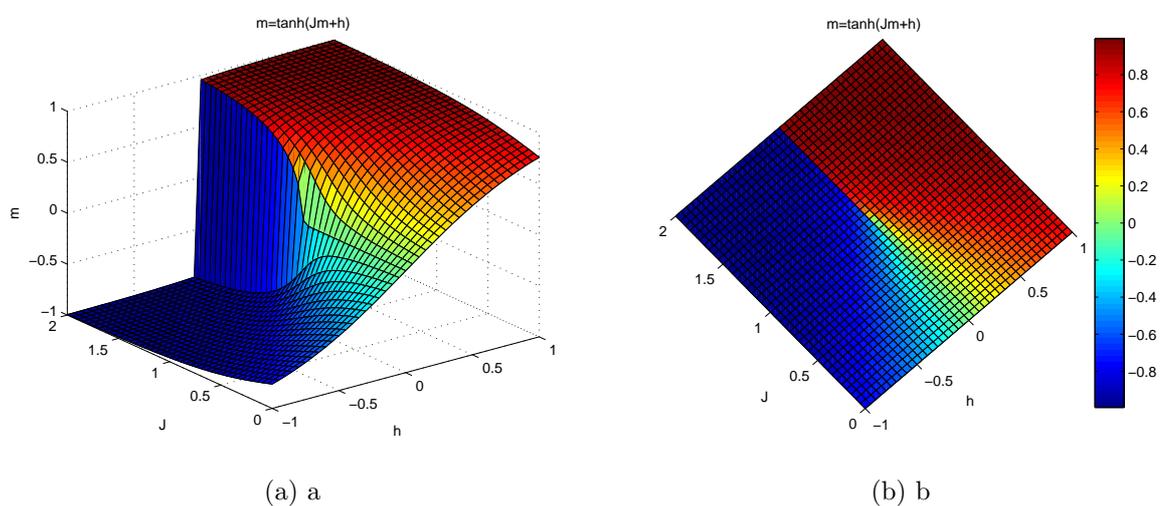


Figura 2.3: (a) Superficie m in funzione di J e h - modello Curie-Weiss (b) Stessa superficie vista dall'alto

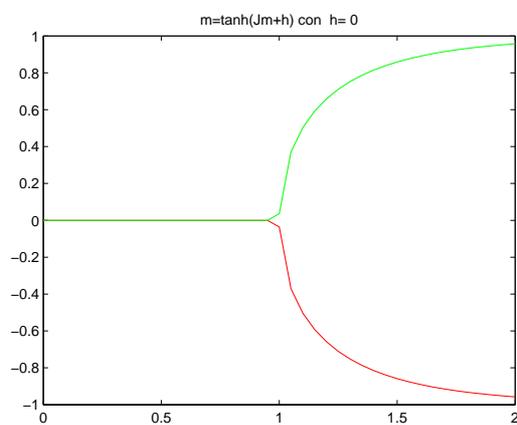


Figura 2.4: Grafico di m al variare di J per $h = 0$ - modello Curie-Weiss

Capitolo 3

Il modello di Curie-Weiss multipopolato

3.1 Il modello

Negli studi di tipo socio-economico, una volta osservati i diversi attributi di un insieme di individui, si può pensare di suddividere l'insieme in sottogruppi che riuniscano soggetti con caratteristiche molto vicine. Ad esempio, conoscendo gli anni di nascita dei membri di una popolazione, potrebbe essere conveniente raggrupparli per fasce d'età, e supporre che gli individui appartenenti ad una stessa fascia abbiano la tendenza a volersi imitare.

Il modello di Curie-Weiss può essere quindi esteso ed utilizzato per descrivere questo caso detto 'multipopolato'.

Si consideri una popolazione di N individui, partizionata in n sottogruppi P_1, \dots, P_n , a due a due disgiunti, ciascuno di cardinalità $|P_s| = N_s$ $s = 1, \dots, n$, tale per cui $\sum_{s=1}^n N_s = N$. Gli elementi del gruppo interagiscono tra di loro secondo l'Hamiltoniana:

$$H_N = -\frac{1}{2N} \sum_{i,j=1}^N J_{ij} \sigma_i \sigma_j - \sum_{i=1}^N h_i \sigma_i$$

La partizione è scelta in modo da poter supporre che all'interno di un

gruppo gli individui siano omogenei rispetto ad un certo attributo. Di conseguenza si assume che tutti gli elementi di un dato gruppo l interagiscano con tutti quelli appartenenti al gruppo s con la medesima intensità. E' possibile utilizzare la seguente matrice a blocchi per illustrare come variano i parametri J_{ij} :

$$\begin{array}{c}
 \left. \begin{array}{l} N_1 \\ N_2 \\ \vdots \\ N_n \end{array} \right\} \left(\begin{array}{c|c|c|c}
 \overbrace{\mathbf{J}_{11}}^{N_1} & \overbrace{\mathbf{J}_{12}}^{N_2} & \cdots & \overbrace{\mathbf{J}_{1n}}^{N_n} \\
 \hline
 \mathbf{J}_{12} & \mathbf{J}_{22} & & \\
 \hline
 \vdots & & & \\
 \hline
 \mathbf{J}_{1n} & \mathbf{J}_{2n} & \cdots & \mathbf{J}_{nn}
 \end{array} \right)
 \end{array}$$

dove ciascun blocco \mathbf{J}_{ls} ha elementi costanti e tutti uguali tra loro. In maniera del tutto analoga, per il campo esterno si ha:

$$\left. \begin{array}{l} N_1 \\ N_2 \\ \vdots \\ N_n \end{array} \right\} \left(\begin{array}{c} \mathbf{h}_1 \\ \hline \mathbf{h}_2 \\ \hline \vdots \\ \hline \mathbf{h}_n \end{array} \right)$$

Indicando $m_l(\sigma)$ la magnetizzazione del sottogruppo P_l e con $\alpha_l = N_l/N$ la sua dimensione relativa, e definendo la funzione g :

$$g(x_1, \dots, x_n) = \frac{1}{2} \sum_{l,s=1}^n \alpha_l \alpha_s J_{ls} x_l(\sigma) x_s(\sigma) + \sum_{l=1}^n \alpha_l h_l x_l(\sigma). \quad (3.1)$$

l'Hamiltoniana per il modello di Curie-Weiss a n popolazioni può essere espressa come:

$$H_N(\sigma) = -Ng(m_1(\sigma), \dots, m_n(\sigma)) \quad (3.2)$$

Definendo la matrice simmetrica \mathbf{J} ed il vettore \mathbf{h} come:

$$\mathbf{J} = \begin{pmatrix} J_{1,1} & J_{1,2} & \dots & J_{1,n} \\ J_{1,2} & J_{2,2} & \dots & J_{2,n} \\ \vdots & \vdots & & \vdots \\ J_{1,n} & J_{2,n} & \dots & J_{n,n} \end{pmatrix} \quad \mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}, \quad (3.3)$$

la funzione g può essere espressa in forma compatta come:

$$g(\mathbf{x}) = \frac{1}{2} \langle \tilde{\mathbf{J}}, \mathbf{x} \rangle + \langle \tilde{\mathbf{h}}, \mathbf{x} \rangle, \quad (3.4)$$

dove:

$$\begin{aligned} \tilde{\mathbf{J}} &= \mathbf{D}_\alpha \mathbf{D}_\alpha \mathbf{J} \mathbf{D}_\alpha \mathbf{D}_\alpha, \\ \tilde{\mathbf{h}} &= \mathbf{D}_\alpha \mathbf{D}_\alpha \mathbf{h}, \\ \mathbf{D}_\alpha &= \text{diag}\{\sqrt{\alpha_1}, \dots, \sqrt{\alpha_n}\} \end{aligned}$$

Nel prossimo paragrafo vedremo, dimostrando l'esistenza del limite termodinamico, che tale modello è ben posto.

3.2 Limite termodinamico

3.3 Esistenza del limite

Anche in questo caso si prova l'esistenza del limite termodinamico della pressione senza calcolarne la soluzione esatta, grazie al teorema per i modelli di campo medio riportato in appendice (B); questa tecnica è molto utile in quanto può essere applicata anche a quei modelli per cui non si dispone dell'espressione esatta della funzione pressione.

Per il modello di Curie-Weiss multipopolato vale la seguente:

Proposizione 3.1. *Data un'Hamiltoniana del modello di Curie-Weiss multipopolato (3.2) e la relativa pressione p_N , esiste una funzione p dei parametri $(\{J_{ij}\}, \{h_i\})$, tale per cui:*

$$\lim_{N \rightarrow \infty} p_N = p$$

Per dimostrare la proposizione si definisce innanzitutto un'Hamiltoniana di prova \tilde{H}_N , per la quale vale il teorema (B.1); si mostra poi che la funzione pressione ad essa relativa, \tilde{p}_N tende alla funzione pressione dell'Hamiltoniana del nostro modello multipopolato (3.2); per semplicità di calcolo si considera l'Hamiltoniana di prova e l'Hamiltoniana (3.2) con campo esterno nullo, per cui $h_l = 0$ per $l = 1, \dots, n$.

Si sceglie come Hamiltoniana di prova la seguente:

$$\tilde{H}_N = \sum_{l=1}^n \tilde{H}_N^{(l)} + \sum_{l < s} \tilde{H}_N^{(ls)}$$

dove:

$$\tilde{H}_N^{(l)} = \alpha_l J_l \frac{1}{2\alpha_l N - 1} \sum_{i \neq j \in P_l} \xi_i \xi_j \quad l = 1, \dots, n$$

$$\tilde{H}_N^{(ls)} = \frac{1}{N} J_{ls} \sum_{\substack{i \in P_l \\ j \in P_s}} \xi_i \xi_j \quad l, s = 1, \dots, n \quad l < s$$

Con il seguente lemma si dimostra che esiste il limite termodinamico per la funzione pressione \tilde{p}_N relativa all'Hamiltoniana di prova.

Lemma 3.2. *Esiste una funzione \tilde{p} tale che:*

$$\lim_{N \rightarrow \infty} \tilde{p}_N = \tilde{p}$$

Dimostrazione. Dalla definizione di $\tilde{H}_N^{(l)}$ e dall'invarianza dello stato di Boltzmann-Gibbs rispetto alle permutazioni di spin, si ha:

$$\begin{aligned}\omega_N(\tilde{H}_N^{(l)}) &= \omega_N \left(\alpha_l J_{ll} \frac{1}{2\alpha_l N - 1} \sum_{i \neq j \in P_l} \xi_i \xi_j \right) = \alpha_l J_{ll} \frac{(\alpha_l N - 1)\alpha_l N}{2\alpha_l N - 1} \omega(\xi_i \xi_j) \\ &= \frac{N}{2} \alpha_l^2 J_{ll} \omega(\xi_i \xi_j)\end{aligned}$$

dove $l = 1, \dots, n$. Analogamente per $\tilde{H}_N^{(ls)}$:

$$\omega_N(\tilde{H}_N^{(ls)}) = N \alpha_l \alpha_s J_{ls} \omega(\xi_i \xi_j)$$

$l = 1, \dots, n-1$ $s = 2, \dots, n$, $l < s$. Di conseguenza, data una qualsiasi partizione di $\{1, \dots, N\}$ in $\{1, \dots, N_1\}$ e $\{N_1 + 1, \dots, N_2\}$ con $N_1 + N_2 = N$, vale l'uguaglianza:

$$\omega_N(\tilde{H}_N) = \omega_N(\tilde{H}_{N_1} + \tilde{H}_{N_2})$$

per cui l'ipotesi del teorema (B.1) è verificata ed il lemma è dimostrato. \square

Mostriamo ora, attraverso i seguenti lemmi che la differenza tra H_N e \tilde{H}_N è termodinamicamente trascurabile e che le rispettive funzioni pressione coincidono al limite. Per praticità si riscrive l'Hamiltoniana H_N in (3.2) nella seguente forma:

$$H_N = \sum_{l=1}^n H_N^{(l)} + \sum_{l < s} H_N^{(ls)}$$

dove:

$$\begin{aligned}H_N^{(l)} &= \frac{1}{2N} J_{ll} \sum_{i, j \in P_l} \xi_i \xi_j \quad l = 1, \dots, n \\ H_N^{(ls)} &= \frac{1}{N} J_{ls} \sum_{\substack{i \in P_l \\ j \in P_s}} \xi_i \xi_j \quad l, s = 1, \dots, n \quad l < s\end{aligned}$$

Lemma 3.3. *Si ha:*

$$\lim_{n \rightarrow \infty} \frac{H_N}{N} = \lim_{n \rightarrow \infty} \frac{\tilde{H}_N}{N}$$

o, equivalentemente:

$$H_N = \tilde{H}_N + O(1)$$

Dimostrazione.

$$\begin{aligned} H_N^{(l)} &= \frac{1}{2N} J_{ll} \sum_{i,j \in P_l} \xi_i \xi_j = \frac{N_l - 1}{2N} J_{ll} \frac{1}{N_l - 1} \sum_{i \neq j \in P_l} \xi_i \xi_j + \frac{1}{2N} J_{ll} \sum_{i \in P_l} \xi_i \xi_i \\ &= \frac{\alpha_l N - 1}{2\alpha_l N} \alpha_l J_{ll} \frac{1}{\alpha_l N - 1} \sum_{i \neq j \in P_l} \xi_i \xi_j + \frac{\alpha_l J_{ll}}{2} \\ &= \alpha_l J_{ll} \frac{1}{\alpha_l N - 1} \sum_{i \neq j \in P_l} \xi_i \xi_j - \alpha_l J_{ll} \frac{1}{\alpha_l N (\alpha_l N - 1)} \sum_{i \neq j \in P_l} \xi_i \xi_j + \alpha_l J_{ll} \quad (3.5) \end{aligned}$$

da cui:

$$H_N^{(l)} = \tilde{H}_N^{(l)} + O(1).$$

In maniera del tutto analoga per i termini $H_N^{(ls)}$ in relazione a $\tilde{H}_N^{(ls)}$, da cui:

$$H_N = \tilde{H}_N + O(1)$$

□

Lemma 3.4. *Siano p_N la funzione pressione relativa a H_N , $h_N(\sigma) = \frac{H_N(\sigma)}{N}$ e analogamente per \tilde{p}_N e \tilde{h}_N . Sia*

$$k_N = \|h_N - \tilde{h}_N\| = \sup_{\sigma \in \{-1,1\}^N} \{|h_N(\sigma) - \tilde{h}_N(\sigma)|\} < \infty.$$

Allora vale:

$$|p_N - \tilde{p}_N| \leq \|h_N - \tilde{h}_N\|.$$

Dimostrazione.

$$\begin{aligned} p_N - \tilde{p}_N &= \frac{1}{N} \ln Z_N - \frac{1}{N} \ln \tilde{Z}_N = \frac{1}{N} \ln \frac{Z_N}{\tilde{Z}_N} \\ &= \frac{1}{N} \ln \frac{\sum_{\sigma} e^{-H_N(\sigma)}}{\sum_{\sigma} e^{-\tilde{H}_N(\sigma)}} \leq \frac{1}{N} \ln \frac{\sum_{\sigma} e^{-H_N(\sigma)}}{\sum_{\sigma} e^{-N(h_N(\sigma) + k_N)}} = \\ &= \frac{1}{N} \ln \frac{\sum_{\sigma} e^{-H_N(\sigma)}}{e^{-Nk_N} \sum_{\sigma} e^{-Nh_N(\sigma)}} = \frac{1}{N} \ln e^{Nk_N} = k_N = \|h_N - \tilde{h}_N\|. \end{aligned}$$

dove la disuguaglianza segue dalla definizione di k_N e dalla monotonia dell'esponenziale e del logaritmo. Allo stesso modo per $\tilde{p}_N - p_N$. \square

A questo punto possiamo dimostrare la proposizione (3.1):

Dimostrazione della proposizione (3.1). L'esistenza del limite termodinamico segue direttamente dai tre lemmi: il primo garantisce l'esistenza \tilde{p}_N , mentre dai seguenti si ha:

$$\lim_{N \rightarrow \infty} |p_N - \tilde{p}_N| \leq \lim_{N \rightarrow \infty} \|h_N - \tilde{h}_N\| = 0,$$

da cui il risultato. \square

3.3.1 Il metodo di Talagrand per il calcolo del limite

Come per il modello di Curie-Weiss, anche per il modello multipopolato è possibile ricavare il valore del limite termodinamico calcolando i bounds della funzione pressione ed il loro valore al limite termodinamico. Tuttavia, per utilizzare tale metodo, è necessario supporre che la matrice ridotta delle interazioni \mathbf{J} sia definita positiva, il che è restrittivo, in quanto equivale ad affermare che tutte le interazioni tra le diverse popolazioni sono di tipo imitativo.

In questo paragrafo si mostra quindi una tecnica diversa, introdotta già da Talagrand in [46] per il modello ad una popolazione, che presenta il vantaggio di essere valida in generale. Questa consiste nel calcolo del numero di configurazioni con la stessa magnetizzazione. Da tale stima, si ottengono i bounds inferiore e superiore per la funzione di partizione, da cui il valore del limite per $N \rightarrow \infty$.

Sia σ_l la configurazione di spin del sottogruppo P_l . Si definisce la quantità:

$$A_{\mu_l} = \text{card} \left\{ \sigma_l \in \Omega_{N_l} \mid m_l(\sigma) = \mu_l \right\} \quad (3.6)$$

che indica il numero di configurazioni σ_l possibili, con magnetizzazione μ_l .

Utilizzando la funzione g espressa in (3.4) e (3.6), la funzione di partizione può essere espressa come:

$$Z_N(\mathbf{J}, \mathbf{h}) = \sum_{\mu} \prod_{l=1}^n A_{\mu_l} \exp \left(N \left(\frac{1}{2} \langle \tilde{\mathbf{J}} \mu, \mu \rangle + \langle \tilde{\mathbf{h}}, \mu \rangle \right) \right) \quad (3.7)$$

dove la somma si estende su tutti i possibili valori del vettore (m_1, \dots, m_n) .

Lemma 3.5. *Dato l'insieme $\Omega_{N_l} = \{-1, 1\}^{N_l}$ di tutte le possibili configurazioni s_l . Sia A_{μ_l} la cardinalità definita in (3.6). Vale la seguente disuguaglianza:*

$$\frac{1}{C} \frac{2^{N_l}}{\sqrt{N_l}} \exp(-N_l \mathcal{I}(\mu_l)) \leq A_{\mu_l} \leq 2^{N_l} \exp(-N_l \mathcal{I}(\mu_l)) \quad (3.8)$$

con C costante e \mathcal{I}

$$\mathcal{I}(x) = \frac{1}{2} \left((1+x) \ln(1+x) + (1-x) \ln(1-x) \right) \quad (3.9)$$

Dimostrazione. Poichè $m_l(\sigma) = \mu_l$, ciò significa che in σ_l compare $N_l(1 + \mu_l)/2$ volte 1 e $N_l(1 - \mu_l)/2$ volte -1 . Di conseguenza:

$$A_{\mu_l} = \binom{N_l}{\frac{N_l(1+\mu_l)}{2}}$$

Dalla formula di Stirling, $n! \sim n^n e^{-n} \sqrt{2\pi n}$, si ottiene:

$$\begin{aligned} A_{\mu_l} &\geq \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{N_l(1-\mu_l^2)}} \frac{N_l^{N_l}}{\left(\frac{N_l(1+\mu_l)}{2}\right)^{N_l(1+\mu_l)/2} \left(\frac{N_l(1-\mu_l)}{2}\right)^{N_l(1-\mu_l)/2}} \\ &\geq \frac{1}{C} \frac{2^{N_l}}{\sqrt{N_l}} \frac{1}{(1+\mu_l)^{N_l(1+\mu_l)/2} (1-\mu_l)^{N_l(1-\mu_l)/2}} \\ &= \frac{1}{C} \frac{2^{N_l}}{\sqrt{N_l}} \exp(-N_l \mathcal{I}(\mu_l)) \end{aligned} \quad (3.10)$$

ovvero si ha un bound inferiore di A_{μ_l} . A questo punto si vuole ricavare anche il bound superiore. A tal fine, si suppone che gli spin siano tra loro *indipendenti*, in modo che tutte le configurazioni σ_l abbiano la stessa probabilità, da cui:

$$A_{\mu_l} = 2^{N_l} P\left\{m_l(\sigma) = \mu_l\right\} \leq 2^{N_l} P\left\{m_l(\sigma) \geq \mu_l\right\}$$

dove, per definizione di magnetizzazione:

$$P\left\{m_l(\sigma) \geq \mu_l\right\} = P\left\{S_l(\sigma) \geq \mu_l N_l\right\}. \quad (3.11)$$

dato $\lambda > 0$, dalla disuguaglianza di Chebyshev, si migliora la probabilità (3.11):

$$\begin{aligned} P\left\{S_l(\sigma) \geq \mu_l N_l\right\} &\leq e^{-\lambda \mu_l N_l} \prod_{i=1}^{N_l} \omega_N(\exp(\lambda \sigma_i)) \\ &= \exp(N_l(-\lambda \mu_l + \ln \cosh \lambda)) \\ &\leq \min_{\lambda} \{\exp(N_l(-\lambda \mu_l + \ln \cosh \lambda))\} \end{aligned} \quad (3.12)$$

Se $|\mu_l| < 1$, l'esponente in (3.12) è minimizzato da:

$$\lambda = \tanh^{-1}(\mu_l) = \frac{1}{2} \ln \left(\frac{1 + \mu_l}{1 - \mu_l} \right) \quad (3.13)$$

Poichè $1/(\cosh^2 y) = 1 - \tanh^2 y$, si ha:

$$\ln \cosh \lambda = -\frac{1}{2} \ln(1 - \mu_l^2) \quad (3.14)$$

Quindi, da (3.13) e (3.14)

$$\min_{\lambda} \{\exp(N_l(-\lambda \mu_l + \ln \cosh \lambda))\} = \exp(-N_l \mathcal{I}(\mu_l)).$$

Siamo quindi arrivati a migliorare A_{μ_l} :

$$A_{\mu_l} \leq 2^{N_l} \exp(-N_l \mathcal{I}(\mu_l)). \quad (3.15)$$

Da (3.10) e (3.15), si ricava quindi la disuguaglianza in (3.8). \square

Dal lemma di Talagrand 3.5 si ricavano i bounds per la funzione di partizione:

$$\frac{2^N}{C} \prod_{l=1}^n \frac{1}{\sqrt{N_l}} \exp \left(N \max_{\boldsymbol{\mu}} \bar{f}(\boldsymbol{\mu}) \right) \leq Z_N(\mathbf{J}, \mathbf{h}) \leq 2^N \prod_{l=1}^n (N_l+1) \exp \left(N \max_{\boldsymbol{\mu}} \bar{f}(\boldsymbol{\mu}) \right)$$

dove la funzione \bar{f} è definita come:

$$\bar{f}(x_1, \dots, x_n) \equiv \frac{1}{2} \sum_{l,s=1}^n \alpha_l \alpha_s J_{ls} x_l x_s + \sum_{l=1}^n \alpha_l h_l x_l - \sum_{l=1}^n \alpha_l \mathcal{I}(x_l). \quad (3.16)$$

mentre \mathcal{I} è la funzione definita in (3.9). Per la pressione si ricava quindi:

$$\begin{aligned} \ln 2 - \frac{1}{N} \left(\ln C + \frac{1}{2} \sum_{l=1}^n \ln N_l \right) + \max_{\boldsymbol{\mu}} \bar{f}(\boldsymbol{\mu}) &\leq p_N(\mathbf{J}, \mathbf{h}) \\ &\leq \frac{1}{N} \left(\sum_{l=1}^n \ln(N_l + 1) \right) + \max_{\boldsymbol{\mu}} \bar{f}(\boldsymbol{\mu}) + \ln 2. \end{aligned}$$

3.3.2 Equazione di campo medio

Il limite termodinamico della pressione si ottiene dalla massimizzazione della funzione \bar{f} definita in (3.16). Imponendo l'annullamento delle derivate di \bar{f} rispetto a x_1, \dots, x_n si ottengono le n equazioni di campo medio:

$$\begin{cases} x_1 = \tanh \left(\sum_{l=1}^n \alpha_l J_{1l} x_l + h_1 \right) \\ x_2 = \tanh \left(\sum_{l=1}^n \alpha_l J_{2l} x_l + h_2 \right) \\ \vdots \\ x_n = \tanh \left(\sum_{l=1}^n \alpha_l J_{ln} x_l + h_n \right). \end{cases} \quad (3.17)$$

Capitolo 4

Il problema inverso

Nel momento in cui si definisce l'Hamiltoniana del sistema, è possibile determinare i momenti delle osservabili del modello, tramite calcoli analitici per i problemi particolarmente semplici, o, nei casi più complessi, attraverso metodi numerici quali quello di Monte Carlo. Assegnati i parametri, è quindi possibile ricavare i valori delle magnetizzazioni e delle correlazioni tra spin.

Tuttavia, quando si cerca di modellizzare una situazione reale, di fatto si procede in maniera opposta. Supponiamo di disporre di un certo numero di dati, e di voler determinare le equazioni che meglio rappresentano il problema. In tal caso saranno i dati a fornire una stima della magnetizzazione e si cercano i valori da assegnare a parametri in modo tale che l'Hamiltoniana fitti al meglio i dati reali. Ciò significa risolvere il **problema inverso**, per cui a partire dai momenti delle osservabili del modello si ricavano i valori dei parametri.

Sono state studiate diverse tecniche per affrontare il problema: si vedano [36] e [37]; in questo lavoro si utilizza il metodo per i modelli di campo medio utilizzato da Tanaka in [47] e, successivamente, da Hertz in [36] e in [37] per risolvere il problema inverso per il modello di Ising.

4.1 Problema inverso per il modello di Curie-Weiss

4.1.1 Modello monopopolato

Consideriamo il modello di Curie-Weiss ad una popolazione; abbiamo visto che l'Hamiltoniana del sistema è:

$$H_N = -\frac{J}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j - \sum_{i,j=1}^N h_i \sigma_i$$

e la probabilità di una configurazione è data dalla distribuzione di Boltzmann-Gibbs:

$$P_N(\sigma) = \frac{e^{-H_N(\sigma)}}{\sum_{\sigma} e^{-H_N(\sigma)}}.$$

Dalla proposizione (2.13), vale: $\omega_N(m_N(\sigma)) = \omega_N(\sigma_i)$. Supponiamo che μ sia l'unica soluzione dell'equazione di campo medio (2.29)

$$\mu = \tanh(J\mu + h),$$

punto di massimo di $p_U = p_L$; si ha quindi $h \neq 0$ e $J > 0$ oppure $h=0$ e $J < 1$.

Abbiamo dimostrato che al limite termodinamico vale:

$$\lim_{N \rightarrow \infty} \omega_N(m_N(\sigma)) = \mu \quad (4.1)$$

Quest'uguaglianza non è sufficiente per risolvere il problema inverso per il modello di Curie-Weiss: oltre al valor medio della magnetizzazione, è necessario conoscerne anche il momento secondo, la suscettività (2.7) $\chi = \frac{\partial^2 p_N}{\partial h^2}$. Al limite termodinamico si ha:

$$\chi = \frac{\partial \mu}{\partial h}$$

per cui, derivando ambo i membri dell'equazione (4.1) si ricava:

$$\lim_{N \rightarrow \infty} \frac{\partial}{\partial h} \omega_N(m_N(\sigma)) = \chi \quad (4.2)$$

Svolgendo la derivata al membro di sinistra, come visto in (2.8), si ha:

$$\frac{\partial}{\partial h} \omega_N(m_N(\sigma)) = N (\omega_N(m_N^2(\sigma)) - \omega_N(m_N(\sigma))^2)$$

Derivando rispetto ad h l'equazione di campo medio (2.29) si ha:

$$\frac{\partial \mu}{\partial h} = \frac{\partial}{\partial h} (\operatorname{tgh}(J\mu + h)) = (1 - \operatorname{tgh}^2(J\mu + h)) \left(J \frac{\partial \mu}{\partial h} + 1 \right)$$

$$\Rightarrow \frac{\partial \mu}{\partial h} = \frac{1 - \operatorname{tgh}^2(J\mu + h)}{1 - J(1 - \operatorname{tgh}^2(J\mu + h))}.$$

La suscettività può quindi essere scritta come:

$$\chi = \frac{1 - \mu^2}{1 - J(1 - \mu^2)}$$

Utilizzando queste relazioni, l'equazione (4.2) diventa:

$$\lim_{N \rightarrow \infty} N (\omega_N(m_N^2(\sigma)) - \omega_N(m_N(\sigma))^2) = \lim_{N \rightarrow \infty} \frac{1 - \omega_N(m_N(\sigma))^2}{1 - J(1 - \omega_N(m_N(\sigma))^2)}$$

da cui si ricava il parametro J in funzione del valor medio e della varianza della magnetizzazione:

$$J = \frac{1}{1 - \omega_N(m_N(\sigma))^2} - \frac{1}{N (\omega_N(m_N^2(\sigma)) - \omega_N(m_N(\sigma))^2)} \quad (4.3)$$

Una volta che si è calcolato J , diventa possibile ricavare h invertendo l'equazione di consistenza (2.29):

$$h = \tanh^{-1}(\omega_N(m_N(\sigma))) - J\omega_N(m_N(\sigma)) \quad (4.4)$$

risolvendo così il problema inverso nel caso in cui f ammette un unico punto di massimo.

Se vale $h = 0$ e $J > 1$, i massimi della funzione f sono due, $\pm\mu_0$ e l'equazione (4.1) non è valida, se non a livello locale. E' quindi necessario restringere il dominio della magnetizzazione: $m_N(\sigma) \in (\pm\mu_0 - \epsilon, \pm\mu_0 + \epsilon)$ con $\epsilon > 0$, in modo che in tale dominio valga:

$$\lim_{N \rightarrow \infty} \omega_N(m_N(\sigma)) = \pm\mu_0$$

4.1.2 Modello multipopolato

Consideriamo l'Hamiltoniana del modello di Curie-Weiss multipopolato:

$$H = -\frac{N}{2} \sum_{l,s} \alpha_l \alpha_s J_{l,s} m_l(\sigma) m_s(\sigma) - \sum_{l=1}^k \alpha_l h_l m_l(\sigma)$$

Si procede in maniera analoga al caso precedente, sapendo che, al limite termodinamico, localmente, si ha:

$$\omega_N(m_l(\sigma)) = \mu_l \quad l = 1, \dots, k$$

dove $\mu = (\mu_1, \mu_2, \dots, \mu_k)$ è il punto di massimo della funzione

$$f(x_1, \dots, x_k) = -\frac{1}{2} \sum_{l,s} \alpha_l \alpha_s J_{l,s} x_l x_s - \sum_{l=1}^k \alpha_l \ln(\cosh(\sum_{s=1}^n \alpha_s J_{l,s} x_s + h_l))$$

per cui soddisfa le equazioni di campo medio del modello.

Si ha inoltre, al limite:

$$\frac{\partial}{\partial h_s} \omega_N(m_l(\sigma)) = \chi_{l,s}$$

dove $\chi_{l,s}$ è un elemento della matrice di suscettività. Vale:

$$\chi_{l,s} = \frac{\partial \mu_l}{\partial h_s} = \frac{\partial \tanh(h_l + \sum_{p=1}^k \alpha_p J_{l,p} \mu_p)}{\partial h_j} = (1 - \mu_l^2) \left(\delta_{l,s} + \sum_{p=1}^k \alpha_p J_{l,p} \chi_{p,s} \right)$$

La matrice di suscettività è quindi $\chi = P(\mathbb{I} + J J D_\alpha D_\alpha \chi)$, dove la matrice P ha elementi $P_{i,j} = \delta_{i,j}(1 - \mu_i^2)$ e $(D_\alpha)_{i,j} = \delta_{i,j}(\sqrt{\alpha_i})$.

Derivando anche $\omega_N(m_l(\sigma))$:

$$\frac{\partial}{\partial h_s} \omega_N(m_l(\sigma)) = N_s (\omega_N(m_l(\sigma) m_s(\sigma)) - \omega_N(m_l(\sigma)) \omega_N(m_s(\sigma)))$$

Si ricava quindi la matrice delle interazioni:

$$J = (P^{-1} - \chi^{-1}) D_\alpha^{-1} D_\alpha^{-1} \quad (4.5)$$

e di conseguenza:

$$h_l = \tanh^{-1}(\mu_l) - \sum_{s=1}^k \alpha_s J_{l,s} \mu_s \quad l = 1, \dots, k \quad (4.6)$$

4.2 La tecnica della massima verosimiglianza

Uno strumento possibile per passare da un insieme di dati alla stima effettiva dei parametri della distribuzione di Boltzmann-Gibbs è il *principio di massima verosimiglianza*, che si basa sull'idea di utilizzare una certa realizzazione campionaria per determinare i parametri incogniti della distribuzione da cui questa è stata estratta.

Supponiamo di disporre di un campione di dati di dimensione n , $\{x_i\}_{i=1}^n$, estratti da una distribuzione con funzione di densità dipendente da un parametro θ . Se il parametro θ è incognito, la probabilità $P(\{x_i\}_{i=1}^n | \theta)$ di osservare la realizzazione $\{x_i\}_{i=1}^n$ condizionata al parametro, può essere pensata come funzione del parametro. Si definisce così la funzione di verosimiglianza:

$$L(\theta | \{x_i\}_{i=1}^n) = P(\{x_i\}_{i=1}^n | \theta).$$

Il principio di massima verosimiglianza afferma che per determinare θ è necessario massimizzare la funzione di verosimiglianza rispetto a $\theta \in \Theta$, ovvero si cerca il valore del parametro che, con maggior probabilità, può aver prodotto i dati osservati.

Vediamo quindi come applicare tale principio per invertire i modelli di Curie-Weiss illustrati. Supponiamo di disporre di M configurazioni di spin $\sigma^{(1)}, \dots, \sigma^{(M)}$, **indipendenti ed identicamente distribuite** secondo la misura di Boltzmann-Gibbs. La funzione di massima verosimiglianza risulta essere:

$$\begin{aligned} L(J, h) &= P_{N,J,h}(\sigma^{(1)}, \dots, \sigma^{(M)}) = \prod_{m=1}^M P_{N,J,h}(\sigma^{(m)}) \\ &= \prod_{m=1}^M \frac{e^{-H_N(\sigma^{(m)})}}{\sum_{\sigma} e^{-H_N(\sigma^{(m)})}} \end{aligned} \quad (4.7)$$

Poichè una funzione ed il suo logaritmo assumono i massimi negli stessi punti, per praticità si massimizza $\ln L$:

$$\ln L(J, h) = \sum_{m=1}^M (-H_N(\sigma^{(m)}) - \ln Z_N(J, h)) \quad (4.8)$$

rispetto ai parametri del modello.

Consideriamo innanzitutto il modello di Curie-Weiss ad una popolazione. Le derivate cercate saranno:

$$\begin{aligned} \frac{\partial L(J, h)}{\partial h} &= \sum_{m=1}^M \left(N m_N(\sigma^{(m)}) - N \frac{\sum_{\sigma} e^{-H_N(\sigma)} m_N(\sigma)}{\sum_{\sigma} e^{-H_N(\sigma^{(m)})}} \right) \\ &= N \sum_{m=1}^M (m_N(\sigma^{(m)}) - \omega_N(m_N(\sigma))) ; \end{aligned}$$

$$\begin{aligned} \frac{\partial L(J, h)}{\partial J} &= \sum_{m=1}^M \left(\frac{N}{2} m_N^2(\sigma^{(m)}) - \frac{N}{2} \frac{\sum_{\sigma} e^{-H_N(\sigma)} m_N^2(\sigma)}{\sum_{\sigma} e^{-H_N(\sigma^{(m)})}} \right) \\ &= \frac{N}{2} \sum_{m=1}^M (m_N^2(\sigma^{(m)}) - \omega_N(m_N^2(\sigma))) . \end{aligned}$$

Si impone che queste si annullino, in modo da individuare i punti di massimo cercati. Si ottiene:

$$\begin{cases} \omega_N(m_N(\sigma)) = \frac{1}{M} \sum_{m=1}^M m_N(\sigma^{(m)}) \\ \omega_N(m_N^2(\sigma)) = \frac{1}{M} \sum_{m=1}^M m_N^2(\sigma^{(m)}) \end{cases} \quad (4.9)$$

Per il modello di Curie-Weiss a k popolazioni, con Hamiltoniana (3.2), il calcolo delle derivate è il seguente:

$$\begin{aligned} \frac{\partial L(J, h)}{\partial h_l} &= N_l \sum_{m=1}^M (m_l(\sigma^{(m)}) - \frac{\sum_{\sigma} e^{-H_N(\sigma)} m_l(\sigma)}{\sum_{\sigma} e^{-H_N(\sigma^{(m)})}}) \\ &= N_l \sum_{m=1}^M (m_l(\sigma^{(m)}) - \omega_N(m_l(\sigma))) ; \end{aligned} \quad (4.10)$$

$$\begin{aligned} \frac{\partial L(J, h)}{\partial J_{l,s}} &= \frac{N\alpha_l\alpha_s}{2} \sum_{m=1}^M (m_l(\sigma^{(m)})m_s(\sigma^{(m)}) - \frac{\sum_{\sigma} e^{-H_N(\sigma)} m_l(\sigma)m_s(\sigma)}{\sum_{\sigma} e^{-H_N(\sigma)}}) \\ &= \frac{N\alpha_l\alpha_s}{2} \sum_{m=1}^M (m_l(\sigma^{(m)})m_s(\sigma^{(m)}) - \omega_N(m_l(\sigma)m_s(\sigma))). \end{aligned}$$

Imponendo l'annullamento di tali derivate si ottengono le seguenti equazioni:

$$\begin{cases} \omega_N(m_l(\sigma)) = \frac{1}{M} \sum_{m=1}^M m_l(\sigma^{(m)}) & l = 1, \dots, n \\ \omega_N(m_l(\sigma)m_s(\sigma)) = \frac{1}{M} \sum_{m=1}^M m_l(\sigma^{(m)})m_s(\sigma^{(m)}) & l, s = 1, \dots, n \end{cases} \quad (4.11)$$

Dalle relazioni (4.9) e (4.11) si ricavano quindi le stime necessarie per esprimere i parametri cercati, ovvero le quantità da sostituire rispettivamente in (4.3) e (4.4) e in (4.5) e (4.6) per il modello di Curie-Weiss ed il modello multipopolato.

Capitolo 5

L'entropia come strumento per la modellizzazione

L'obiezione che può sorgere spontanea nel momento in cui si utilizza un'Hamiltoniana per modellizzare una popolazione è che gli esseri umani non sono particelle in un bagno a temperatura T . Mostreremo quindi come, senza disporre di alcuna ipotesi di carattere fisico e percorrendo una strada molto diversa, si possa ricavare un modello del tutto analogo ai modelli di Curie-Weiss. Nel capitolo precedente abbiamo visto come, per tali modelli, risolvendo il problema inverso con il metodo della massima verosimiglianza, si ottengono le equazioni (4.9) e (4.11). Utilizzando il metodo della Massima Entropia sviluppato da Jaynes, invece, l'imporre queste uguaglianze costituisce il punto di partenza, che, insieme alla massimizzazione dell'entropia, consente di determinare la probabilità di una configurazione, come verrà illustrato in seguito.

Nei paragrafi successivi vedremo innanzitutto cosa si intende per entropia, in che cosa consiste il Principio della Massima Entropia e come da questo sia possibile ottenere un modello corrispondente a quello di Curie-Weiss.

5.1 L'entropia: in termodinamica e in teoria dell'informazione

I propose to name the quantity S the entropy of the system, after the Greek word [τροπή], the transformation. I have deliberately chosen the word entropy to be as similar as possible to the word energy: the two quantities to be named by these words are so closely related in physical significance that a certain similarity in their names appears to be appropriate

R. Clausius [31]

My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.'

C. Shannon [44]

La parola entropia apparve per la prima volta nella seconda metà dell'Ottocento, nell'ambito della termodinamica, precisamente nel *'Trattato sulla teoria meccanica del calore'*, (*'Abhandlungen über die mechanische Wärmetheorie'*), di Rudolf Clausius. [31] Ben presto il concetto si sviluppò e si diffuse in altri settori della Fisica e non solo: venne ridefinita e studiata nell'ambito della meccanica statistica e, con un diverso significato, in teoria dell'informazione; viene, inoltre, largamente utilizzata in diversi campi, tra cui quello delle scienze sociali.

In termodinamica classica, l'entropia è una funzione di stato, indicata con S , che quantifica l'indisponibilità di un sistema a produrre lavoro. Si può definire in forma variazionale come:

$$\Delta S = \frac{Q}{T} \quad (5.1)$$

dove Q è la quantità di calore assorbito o ceduto in maniera reversibile e isoterma dal sistema a temperatura T .

In questo contesto, quindi, l'entropia indica quali sono i processi che possono avvenire in maniera spontanea: l'evoluzione di un sistema avviene sempre nella direzione dell'entropia crescente. Da qui il legame intrinseco con il secondo principio della termodinamica, che può essere riformulato come segue: *'In un sistema isolato l'entropia è una funzione non decrescente nel tempo'*.

Nel 1870, con lo sviluppo della meccanica statistica, il fisico J.W. Gibbs diede un nuovo significato all'entropia, legato alle possibili disposizioni molecolari di un sistema di particelle. L'entropia di Gibbs è definita come:

$$S = -k_B \sum_i p_i \ln p_i \quad (5.2)$$

dove k_B è la costante di Boltzmann e p_i è la probabilità che il sistema si trovi nell' i -esimo microstato; l'insieme di tutti i microstati costituisce l'insieme micro-canonico. Massimizzando la funzione S si ottiene lo stato di equilibrio del sistema. Questa può essere considerata come la definizione fondamentale di entropia, in quanto tutte le altre espressioni di S possono essere ricavate da (5.2) ma non viceversa.

Successivamente Boltzmann rielaborò il concetto di Gibb's, definendo l'entropia come la misura del numero di microstati possibili di un sistema, date le sue proprietà termodinamiche macroscopiche; ad uno stesso macrostato, infatti, possono corrispondere un alto numero di microstati. Si definisce quindi l'entropia di Boltzmann come:

$$S = k_B \ln \Omega \quad (5.3)$$

dove Ω è il numero di microstati che corrispondono ad un macrostato. L'entropia di Gibbs corrisponde a (5.3) quando tutti i microstati sono equiproba-

bili. Tale configurazione è quella con entropia massima e quindi, dal secondo principio, quella di equilibrio.

Vediamo, infine, il significato che l'entropia assume in teoria dell'informazione. Nel 1948, nell'articolo 'A mathematical theory of communication' Claude Shannon introdusse il concetto di *entropia dell'informazione* [44], mostrando come fosse possibile quantificare l'informazione contenuta in un messaggio emesso da una sorgente.

Egli trascurò completamente il contenuto semantico, pensando alla quantità di informazione esclusivamente in termini probabilistici. Per fare ciò, Shannon rappresentò una sorgente di informazione discreta come un processo stocastico che ad ogni istante di tempo n genera un carattere preso da un dato alfabeto finito. Si ha quindi un insieme di eventi possibili con probabilità p_1, p_2, \dots, p_P , dove $p_i = P(X = a_i)$ è la probabilità che il carattere generato sia a_i , con $a_i \in A$ (alfabeto finito). A questo punto, l'informazione viene quantificata attraverso una funzione che misura l'incertezza di X , l'entropia appunto, definita come:

$$H(p_1, p_2, \dots, p_P) \equiv -K \sum_{i=1}^P p_i \log p_i. \quad (5.4)$$

dove K è una costante positiva arbitraria, che dipende dalla base del logaritmo, ovvero dall'unità di misura scelta per H . Generalmente H si misura in bits e quindi si ha base 2 per il logaritmo.

Shannon giunse a tale definizione imponendo che H fosse una funzione delle probabilità e che:

- H fosse continua in $\{p_i\}$;
- dati degli eventi equiprobabili, ovvero $p_i = 1/P$ per ogni $i = 1, \dots, P$, H risultasse funzione di P monotona crescente;
- suddividendo l'insieme p_1, p_2, \dots, p_P in $A = p_1, \dots, p_k$ e $B = p_{k+1}, \dots, p_P$, dati $w_A = \sum_{i=1}^k p_i$ e $w_B = \sum_{i=k+1}^P p_i$, si verificasse $H(p_1, p_2, \dots, p_P) = H(w_A, w_B) + w_A H(p_1/w_A, \dots, p_k/w_A) + w_B H(p_{k+1}/w_B, \dots, p_P/w_B)$.

In questo caso, l'entropia misura quindi la quantità di incertezza o informazione presente in un segnale aleatorio.

A partire dal 1957, il fisico Jaynes si dedicò a dimostrare il nesso tra il concetto fisico di entropia e quello della teoria dell'informazione [29], sviluppando l'importante, seppur ancora molto dibattuto, Principio di Massima Entropia, illustrato nel seguente paragrafo.

5.2 Il Principio di Massima Entropia

Jaynes si servì degli strumenti della teoria dell'informazione per affrontare i problemi della Meccanica Statistica e, attraverso il Principio di Massima Entropia mostrò come fosse possibile determinare le distribuzioni di probabilità di una configurazione, a partire da informazioni parziali. L'idea base è quella di sfruttare le informazioni a disposizione ed imporre che la distribuzione cercata sia quella che massimizza l'entropia di Shannon, in quanto misura dell'incertezza e della quantità di informazione.

Grazie a questo approccio, Jaynes mostrò come fosse possibile ottenere le distribuzioni canoniche in maniera molto semplice e senza, di fatto, fare assunzioni di tipo fisico.

Il suo metodo non fu esente da critiche [49], [32], come illustreremo nell'ultimo paragrafo, ma merita sicuramente di essere indagato ed approfondito in quanto i risultati ottenuti sembrerebbero confermare la validità dei modelli di Curie-Weiss per gli studi di tipo sociale.

Vediamo in dettaglio in cosa consiste applicare il Principio di Massima Entropia. Si consideri un sistema descritto da un insieme di variabili di stato: $\{x_1, x_2, \dots, x_N\} \equiv \mathbf{x}$, per cui ciascuna possibile configurazione ha una certa probabilità di essere osservata. La probabilità di uno stato, in questo caso, non può essere pensata in senso frequentista ma va piuttosto intesa come la nostra conoscenza del sistema. Poichè spesso si ha a che fare con sistemi di dimensione molto elevata, con N molto grandi, è conveniente studiare la distribuzione di opportune funzioni degli stati.

Si definiscono perciò delle quantità legate alla configurazione, $f_1(\mathbf{x})$, $f_2(\mathbf{x})$, \dots , $f_K(\mathbf{x})$, che riassumono alcune proprietà di un sistema e i cui valori medi $\langle f_\nu(\mathbf{x}) \rangle_{exp}$ possono essere calcolati. Si cerca quindi la distribuzione $P(x)$ tale per cui i valori medi delle K funzioni considerate, $\langle f_\nu(\mathbf{x}) \rangle_{exp}$ osservati sperimentalmente coincidano con i loro valori attesi $\langle f_\nu(\mathbf{x}) \rangle_P$ rispetto alla distribuzione.

Si considera la seguente espressione della funzione entropia:

$$S[P] = - \sum_{\mathbf{x}} P(\mathbf{x}) \ln P(\mathbf{x}), \quad (5.5)$$

e si risolve quindi il problema di massimizzazione vincolata, dove i vincoli sono dati dalle informazioni parziali a disposizione, per cui si impone:

$$\langle f_\nu(\mathbf{x}) \rangle_{exp} = \langle f_\nu(\mathbf{x}) \rangle_P \equiv \sum_{\mathbf{x}} P(\mathbf{x}) f_\nu(\mathbf{x}) \quad (5.6)$$

con $\nu = 0, \dots, K$. Poichè la distribuzione di probabilità deve essere normalizzata, si utilizza la seguente tecnica: si sceglie $f_0(\mathbf{x}) = 1$ e si impone che sia uguale al valore sperimentale 1.

Per risolvere il problema si utilizzano i moltiplicatori di Lagrange, per cui si introducono i K parametri $\{\lambda_\mu\}$ e la **funzione di entropia generalizzata**:

$$\mathcal{S}[P; \{\lambda_\mu\}] = S[P] - \sum_{\mu=0}^K \lambda_\mu [\langle f_\mu(\mathbf{x}) \rangle_P - \langle f_\mu(\mathbf{x}) \rangle_{exp}]. \quad (5.7)$$

Si ottimizza quindi $\mathcal{S}[P; \{\lambda_\mu\}]$ rispetto alla probabilità $P(\mathbf{x})$ e rispetto ai parametri, imponendo:

1.

$$\begin{aligned} 0 &= \frac{\partial \mathcal{S}[P; \{\lambda_\mu\}]}{\partial P(\mathbf{x})} = \frac{\partial S[P]}{\partial P(\mathbf{x})} - \sum_{\mu=0}^K \lambda_\mu \frac{\partial \langle f_\mu(\mathbf{x}) \rangle_P}{\partial P(\mathbf{x})} = \\ &= -\ln P(\mathbf{x}) - 1 - \sum_{\mu=0}^K \lambda_\mu f_\mu(\mathbf{x}), \end{aligned}$$

da cui si ricava:

$$P_{me}(\mathbf{x}) = \frac{1}{Z(\{\lambda_\nu\})} \exp \left[- \sum_{\mu=1}^K \lambda_\mu f_\mu(\mathbf{x}) \right] \quad (5.8)$$

con $Z_{me}(\{\lambda_\nu\}) = \sum_{\mathbf{x}} \exp(-\lambda_0 - 1)$. Normalizzare rispetto a λ_0 equivale a normalizzare la distribuzione per cui si può scrivere esplicitamente:

$$Z_{me}(\{\lambda_\nu\}) = \sum_{\mathbf{x}} \exp \left[- \sum_{\mu=1}^K \lambda_\mu f_\mu(\mathbf{x}) \right] \quad (5.9)$$

2.

$$0 = \frac{\partial \mathcal{S}[P; \{\lambda_\mu\}]}{\partial \lambda_\mu} = \langle f_\mu(\mathbf{x}) \rangle_P - \langle f_\mu(\mathbf{x}) \rangle_{exp} \quad (5.10)$$

da cui si vede che massimizzare l'entropia generalizzata rispetto ai parametri $\{\lambda_\nu\}$ equivale ad imporre che le medie delle funzioni considerate misurate sperimentalmente coincidano con i valori predetti dalla distribuzione. Scrivendo esplicitamente $\langle f_\mu(\mathbf{x}) \rangle_P$ e sostituendo in essa l'espressione di $P(\mathbf{x})$ trovata in (5.8), si ha:

$$\langle f_\mu(\mathbf{x}) \rangle_{exp} = \frac{1}{Z(\{\lambda_\nu\})} \sum_{\mathbf{x}} f_\mu(\mathbf{x}) \exp \left[- \sum_{\mu=1}^K \lambda_\mu f_\mu(\mathbf{x}) \right] \quad (5.11)$$

Sostituendo in (5.7) la distribuzione $P_{me}(\mathbf{x})$, si ottiene la seguente espressione per l'entropia generalizzata:

$$\mathcal{S}[P; \{\lambda_\mu\}] = \ln Z(\{\lambda_\nu\}) + \sum_{\mu=0}^K \lambda_\mu \langle f_\mu(\mathbf{x}) \rangle_{exp}, \quad (5.12)$$

Tale espressione coincide con il logaritmo della probabilità che il modello generi i dati osservati, ovvero il logaritmo della verosimiglianza.

Il lavoro di Bialeck et al. [4] mostra come questo procedimento possa essere utilizzato in pratica per ricavare la distribuzione delle velocità di volo all'interno di uno stormo di uccelli, a partire solo dalla conoscenza della matrice delle correlazioni di queste velocità.

5.3 Il Principio di Massima Entropia ed il modello di Curie-Weiss

Si vuole ora indagare il nesso tra il Principio della Massima Entropia ed il modello di Curie-Weiss, mostrando come, attraverso strumenti molto diversi si arrivi a risultati analoghi. In particolare, si vedrà che, anche con l'approccio di Jaynes, si determina una distribuzione di probabilità per la configurazione di particelle che ha la medesima forma della distribuzione di Boltzmann-Gibbs assunta per il modello di Curie-Weiss.

5.3.1 Modello ad una popolazione

Studiamo un insieme di spin $\sigma = \{\sigma_1, \dots, \sigma_N\}$, supponendo di disporre di M configurazioni indipendenti $\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(M)}$. Supponiamo inoltre di conoscere, come funzioni della configurazione, la magnetizzazione ed il suo quadrato:

$$m_N(\sigma) = \frac{1}{N} \sum_{i=1}^N \sigma_i,$$

$$m_N(\sigma)^2 = \frac{1}{N^2} \sum_{i,j=1}^N \sigma_i \sigma_j,$$

per cui si impone che il valor medio di tali quantità misurato sperimentalmente coincida con il valor medio della distribuzione di probabilità che genera la configurazione, ovvero:

$$\begin{cases} \langle m_N(\sigma) \rangle_{exp} = \langle m_N(\sigma) \rangle_P \\ \langle m_N^2(\sigma) \rangle_{exp} = \langle m_N^2(\sigma) \rangle_P \end{cases} \quad (5.13)$$

dove:

$$\langle m_N(\sigma) \rangle_{exp} \equiv \frac{1}{M} \sum_{m=1}^M m_N(\sigma^{(m)}) \quad \langle m_N(\sigma) \rangle_P \equiv \sum_{\sigma} m_N(\sigma) P(\sigma)$$

$$\langle m_N^2(\sigma) \rangle_{exp} \equiv \frac{1}{M} \sum_{m=1}^M m_N^2(\sigma^{(m)}) \quad \langle m_N^2(\sigma) \rangle_P \equiv \sum_{\sigma} m_N^2(\sigma) P(\sigma)$$

Si impongono quindi due vincoli, e altrettanti saranno i moltiplicatori di Lagrange da introdurre. Si osserva che queste equazioni, che costituiscono il punto di partenza per il utilizzare il metodo della Massima Entropia, corrispondono al punto di arrivo per il metodo della massima verosimiglianza. Rispetto alla notazione precedente si pone:

$$\begin{cases} f_1(\mathbf{x}) = m_N(\sigma) \\ f_2(\mathbf{x}) = m_N^2(\sigma) \end{cases} \quad \begin{cases} \lambda_1(\mathbf{x}) = -Nh \\ \lambda_2(\mathbf{x}) = -\frac{NJ}{2} \end{cases} \quad (5.14)$$

Con queste ipotesi, l'equazione dell'entropia generalizzata (5.7) risulta essere:

$$\begin{aligned} \mathcal{S}[P; \{\lambda_\mu\}] = & S[P] - \lambda_0 f_0 + \frac{NJ}{2} [\langle m_N^2(\sigma) \rangle_P - \langle m_N^2(\sigma) \rangle_{exp}] \\ & + Nh [\langle m_N(\sigma) \rangle_P - \langle m_N(\sigma) \rangle_{exp}] \end{aligned} \quad (5.15)$$

Imponendo la massimizzazione rispetto a $P(\sigma)$ si ha:

$$\begin{aligned} 0 = \frac{\partial \mathcal{S}[P; \{J, h\}]}{\partial P(\sigma)} &= \frac{\partial S[P]}{\partial P(\sigma)} - \lambda_0 \frac{\partial \langle f_0(\sigma) \rangle_P}{\partial P(\sigma)} + \frac{NJ}{2} \frac{\partial \langle m_N^2(\sigma) \rangle_P}{\partial P(\sigma)} \\ &+ Nh \frac{\partial \langle m_N(\sigma) \rangle}{\partial P(\sigma)} \\ &= -\ln P(\sigma) - 1 - \lambda_0 f_0 + J m_N^2 + h m_N \end{aligned}$$

da cui:

$$P_{me}(\sigma) = \frac{1}{Z(J, h)} \exp \left[\frac{NJ}{2} m_N^2(\sigma) + N h m_N \right] \quad (5.16)$$

dove, massimizzando rispetto a λ_0 per la normalizzazione della probabilità, si ha:

$$Z_{me}(J, h) = \sum_{\sigma} \exp \left[\frac{NJ}{2} m_N^2(\sigma) + Nh m_N(\sigma) \right] \quad (5.17)$$

Applicando il Principio dell'Massima Entropia, imponendo i vincoli opportuni, abbiamo quindi ritrovato la probabilità di Boltzmann-Gibbs, in quanto (5.23) può essere scritta come:

$$P_{me}(\sigma) = \frac{1}{Z_{me}(J, h)} \exp[-H(\sigma)] = \frac{e^{-H(\sigma)}}{\sum_{\sigma} e^{-H(\sigma)}} \quad (5.18)$$

dove $H(\sigma)$ è l'Hamiltoniana del modello di Curie-Weiss ad una popolazione.

Infine, sostituendo la distribuzione $P_{me}(\sigma)$ in (5.15), si ottiene:

$$\begin{aligned} \mathcal{S}[P_{me}; (J, h)] &= \ln Z_{me}(J, h) - \frac{NJ}{2} \langle m_N^2(\sigma) \rangle_{exp} - Nh \langle m_N(\sigma) \rangle_{exp} \\ &= \ln \left(\sum_{\sigma} e^{-H(\sigma)} \right) - \langle H(\sigma) \rangle_{exp}, \end{aligned} \quad (5.19)$$

L'entropia generalizzata (5.19) così espressa coincide esattamente con N volte il logaritmo della funzione di massima verosimiglianza (4.8), ovvero: $\mathcal{S}[P_{me}; (J, h)] = N \ln L(J, h)$.

5.3.2 Modello multipopolato

Consideriamo ora il modello di Curie-Weiss multipopolato, in cui la popolazione di N individui viene suddivisa in n sottogruppi distinti, ciascuno con una magnetizzazione $m_l(\sigma)$, con $l = 1, \dots, n$. In questo caso, i vincoli che si impongono per la ricerca della distribuzione $P(\sigma)$ sono i seguenti:

$$\begin{cases} \langle m_l(\sigma) \rangle_{exp} = \langle m_l(\sigma) \rangle_P & l = 1, \dots, n \\ \langle m_l(\sigma) m_s(\sigma) \rangle_{exp} = \langle m_l(\sigma) m_s(\sigma) \rangle_P & l, s = 1, \dots, n \end{cases} \quad (5.20)$$

per cui, per ciascun gruppo, come in (5.13) si fissa la media delle magnetizzazioni, mentre il secondo vincolo è sulle correlazioni. In totale si avranno quindi $n + \frac{n^2}{2} = n(\frac{n}{2} + 1)$ vincoli. Si introducono i seguenti moltiplicatori di Lagrange: $J_{l,s}$ per le correlazioni e h_l per le magnetizzazioni, $l, s = 1 \dots n$.

L'equazione dell'entropia generalizzata (5.7) risulta essere:

$$\begin{aligned} \mathcal{S}[P; \{J_{l,s}\}, \{h_l\}] &= S[P] - \lambda_0 f_0 \\ &+ \frac{N}{2} \sum_{l,s=1}^n J_{l,s} [\langle m_l(\sigma) m_s(\sigma) \rangle_P - \langle m_l(\sigma) m_s(\sigma) \rangle_{exp}] \\ &+ N \sum_{l=1}^n h_l [\langle m_l(\sigma) \rangle_P - \langle m_l(\sigma) \rangle_{exp}] \end{aligned} \quad (5.21)$$

Imponendo la massimizzazione rispetto a $P(\sigma)$ si ha:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{S}[P; \{J_{l,s}\}, \{h_l\}]}{\partial P(\sigma)} = \frac{\partial S[P]}{\partial P(\sigma)} - \lambda_0 \frac{\partial \langle f_0(\sigma) \rangle_P}{\partial P(\sigma)} \\ &+ \frac{N}{2} \sum_{l,s=1}^n J_{l,s} \frac{\partial \langle m_l(\sigma) m_s(\sigma) \rangle_P}{\partial P(\sigma)} + N \sum_{l=1}^n h_l \frac{\partial \langle m_l(\sigma) \rangle}{\partial P(\sigma)} \\ &= -\ln P(\sigma) - 1 - \lambda_0 f_0 + \frac{N}{2} \sum_{l,s=1}^n J_{l,s} m_l(\sigma) m_s(\sigma) + N \sum_{l=1}^n h_l m_l(\sigma) \end{aligned} \quad (5.22)$$

da cui:

$$P_{me}(\sigma) = \frac{1}{Z(\{J_{l,s}\}, \{h_l\})} \exp \left[\frac{N}{2} \sum_{l,s=1}^n J_{l,s} m_l(\sigma) m_s(\sigma) + N \sum_{l=1}^n h_l m_l(\sigma) \right] \quad (5.23)$$

dove, massimizzando rispetto a λ_0 per la normalizzazione della probabilità, si ha:

$$Z_{me}(\{J_{l,s}\}, \{h_l\}) = \sum_{\sigma} \exp \left[\frac{N}{2} \sum_{l,s=1}^n J_{l,s} m_l(\sigma) m_s(\sigma) + N \sum_{l=1}^n h_l m_l(\sigma) \right] \quad (5.24)$$

Sostituendo la distribuzione $P_{me}(\sigma)$ in (5.15), si ottiene:

$$\begin{aligned} \mathcal{S}[P_{me}; (J, h)] &= \ln Z_{me}(J, h) - \frac{N}{2} \sum_{l,s=1}^n J_{l,s} \langle m_l(\sigma) m_s(\sigma) \rangle_{exp} \\ &- N \sum_{l=1}^n h_l \langle m_l(\sigma) \rangle_{exp} \end{aligned}$$

5.4 Critiche al Principio della Massima Entropia

Abbiamo visto come, grazie alla tecnica sviluppata da Jaynes, sia possibile ricavare la probabilità di una configurazione attraverso pochi e semplici calcoli, sfruttando solo le informazioni parziali a disposizione e la funzione entropia, senza necessitare di alcuna ipotesi di tipo fisico. Nonostante i vantaggi che presenta, la validità del Principio di Massima Entropia è controversa.

La prima obiezione che può sorgere è legata alla natura della funzione entropia: abbiamo visto che non si tratta una funzione intrinseca del sistema, ma dipende dalle variabili e, di conseguenza, utilizzando variabili diverse anche la distribuzione che si ottiene con questo metodo cambia. Jaynes ovviò a questo problema proponendo di massimizzare l'entropia relativa, che assume la seguente forma:

$$\tilde{S} = - \sum_x p(x) \ln \left[\frac{p(x)}{q(x)} \right],$$

dove $q(x)$ è una distribuzione nota, generalmente quella uniforme. Se da un lato è vero che \tilde{S} non dipende da x , non esistono regole per la scelta di $q(x)$, e quindi si ricade, sostanzialmente, nello stesso problema della scelta delle variabili da utilizzare.

Le critiche più profonde colpiscono l'idea stessa alla base del Principio e lo scopo ultimo di Jaynes. Il Principio di Massima Entropia venne presentato come un'estensione del Principio di ragion sufficiente di Laplace [49] secondo il quale, in assenza di informazioni su un insieme di eventi, questi possono essere assunti come equiprobabili. Allo stesso modo, il Principio di Massima Entropia fornisce la distribuzione '*maximally noncommittal with respect to missing information*' [29]; in [4], Bialeck et.al si afferma addirittura che, il fatto di assumere che la distribuzione cercata sia quella che massimizza l'entropia, non vada vista come un'assunzione legata al modello, ma piuttosto come l'assenza di assunzioni.

Si tratta di affermazioni sicuramente molto forti; per poterle giustificare è necessario inoltre interpretare la probabilità non in senso frequentista, come

generalmente si fa in Fisica, bensì soggettivista, per cui non la si intende come proprietà del mondo reale ma piuttosto come rappresentazione del nostro stato di conoscenza del mondo. La visione soggettivista è però oggetto di discussione anche nell'ambito della statistica; l'approccio di Jaynes viene inoltre criticato in quanto molti sostengono che non si possa trarre conoscenza dall'ignoranza. Significative, a tal proposito, le parole di Ellis in [17]: *“Mere ignorance is no ground for any inference whatsoever. ‘Ex nihilo nihil.’ It cannot be that because we are ignorant of the matter we know something about it”*.

L'obiettivo di Jaynes era quello di ridurre la Meccanica Statistica, da disciplina della Fisica a mera tecnica di inferenza [49], il che può difficilmente essere accettato all'interno della comunità scientifica.

Capitolo 6

Dati e risultati sperimentali

6.1 Il problema

Abbiamo visto diversi approcci possibili per modellizzare le situazioni in cui un gruppo di individui si trova ad affrontare una scelta. Vediamo ora come questi metodi possano essere utilizzati per indagare un problema reale, per cui a partire da un insieme di dati si cerca il modello più adatto a descrivere la situazione. Presentiamo quindi il progetto di ricerca sul quale sta attualmente lavorando un gruppo di fisici e matematici delle Università di Parma, Modena, Ferrara e Bologna. Si tratta di uno studio del comportamento delle donne della provincia di Parma di fronte all'esame citologico per la prevenzione dei tumori del collo e della cervice dell'utero, il Pap test. L'Azienda Sanitaria di Parma a partire dal 1998 ha attivato una campagna di screening per aumentare le adesioni al test, inviando a tutte le donne della provincia, in età compresa tra i 25 ed i 65 anni, un invito a partecipare al test. Nel caso in cui una donna invitata non si presenti, la partecipazione viene sollecitata con un secondo invito, e questa procedura si ripete ogni tre anni. Sebbene tale campagna abbia innalzato significativamente la percentuale delle adesioni, come vedremo in seguito, non è bastata per raggiungere la quota sperata. Lo scopo di questo progetto di ricerca consiste innanzitutto nella formulazione di un modello efficace per rappresentare l'atteggiamento delle donne di fronte

alla scelta di effettuare o meno il Pap test, considerando sia il peso della cultura individuale di ciascuna, sia il peso delle interazioni tra esse. Attraverso gli strumenti della meccanica statistica, in particolare utilizzando il modello di Curie-Weiss, si vogliono comprendere più a fondo le dinamiche di un comportamento collettivo, quantificando, così, l'effettiva incidenza degli inviti. E' importante capire cosa influenzi maggiormente la scelta di una donna perchè questo consentirebbe di individuare la strategia migliore per aumentare le percentuali di adesione ed attuare le politiche economiche e sociali che ottimizzino i costi sanitari e i benefici della campagna.

6.2 I dati

Grazie al lavoro del dottor Luigi Lombardozzi, responsabile del centro screening 'Vasari' di Parma per la diagnosi precoce dei tumori del collo dell'utero, a partire dal 2003 sono state registrate informazioni molto dettagliate relative al Pap test nel periodo di screening. Sono stati raccolti anche dati relativi alla fase pre-screening, in particolare dal 1991 al 1998; ad eccezione dei primi quattro anni della campagna (1998-2002), si dispone così di un quadro completo della situazione. Essendo i dati numerosi e complessi, prima ancora di poter pensare ad un modello, è stato necessario uno studio statistico molto attento ed approfondito, in modo da capire quali considerazioni fossero effettivamente possibili. Vediamo quindi come sono strutturate queste informazioni.

Ogni donna della provincia di Parma è stata identificata da un codice; per ciascuna di esse si hanno:

- dati di carattere generale, quali data di nascita, comune di residenza e di domicilio;
- dati di carattere medico, come, ad esempio, eventuale stato di gravidanza, di menopausa o di allattamento, numero di figli e di aborti, uso di pillola contraccettiva;

- dati relativi all'invito e al test, ovvero tipo di invito (primo, secondo, follow up, accesso spontaneo, . . .), data dell'invio e dell' eventuale esecuzione, sede dello screening, eventuale esito.

Sebbene all'origine si pensava di poter considerare anche gli accessi spontanei ed i secondi inviti, i numeri troppo bassi di risposte hanno suggerito di concentrarsi solo sull'analisi delle risposte al primo invito.

Un altro approccio che è stato scartato in seguito all'analisi dei dati è stato quello della suddivisione geografica per comuni di residenza. Considerando singolarmente i comuni della provincia di Parma, si rischiava infatti di avere delle anomalie dovute a numeri molto bassi: come si può osservare nella tabella (6.1), oltre alla percentuale di adesione, è necessario considerare anche il numero degli inviti mandati. Se guardiamo Compiano, ad esempio, la percentuale del 33% si riferisce a una sola adesione su un totale di 3 inviti spediti, il che, chiaramente, non ha consistenza statistica.

Poichè si conosce anche il centro in cui una donna esegue il pap-test, si era pensato anche alla suddivisione in base al centro di screening, ma anche qui i dati hanno suggerito di abbandonare l'idea.

Un'altra ipotesi possibile è che le donne agiscano in maniera differente secondo la loro età: potrebbe essere molto diverso il ruolo che hanno la cultura e le esperienze personali e quello delle influenze reciproche tra amiche, familiari e colleghe, per una ragazza ventenne piuttosto che per una donna sessantenne. E' quindi sembrato molto ragionevole considerare una suddivisione per età, anche se la cosa non è così ovvia. Innanzitutto è necessario capire se la suddivisione vada effettuata secondo gli anni di nascita, in modo da seguire le stesse donne nel tempo, o piuttosto se fissare delle 'finestre' di fasce d'età, in cui 'passano' le diverse donne. In entrambi i casi, inoltre, si deve poi scegliere quanti e quali gruppi individuare, in maniera tale che il comportamento all'interno di un gruppo sia sufficientemente omogeneo.

Tali suddivisioni vengono pensate al fine di sviluppare un adeguato modello di Curie-Weiss multipopolato per descrivere il comportamento delle donne di fronte al Pap-Test; tuttavia, poichè, come abbiamo visto, si tratta di un

Comune	Test eseguiti	Primi inviti	Percentuale di adesione su primo invito
Albaneto	12	26	46%
Bardi	1	10	10%
Bedonia	12	19	63%
Berceto	9	28	32%
Bore	5	11	45,40%
Borgo Val di Taro	11	23	47,80%
Busseto	347	757	45,80%
Calestano	27	73	37%
Collecchio	572	1127	50%
Colorno	324	735	44%
Compiano	1	3	33%
Corniglio	17	49	34,70%
Felino	456	1062	43%
Fidenza	1386	3215	43%
Fontanellato	317	802	40%
Fontevivo	204	463	44%
Fornoro di Taro	31	74	42%
Laughirano	99	192	51%
Lesignano de Bagni	114	276	41%
Medesano	418	882	47%
Mezzani	127	315	40%
Monchio delle Corti	16	33	48%
Montechiarugolo	578	1359	42,50%
Neviano degli Arduini	36	111	32,40%
Noceto	599	1378	43,50%
Palanzano	13	28	46,40%
Parma	9119	19852	46%
Pellegrino Parmense	14	27	52%
Polesine Parmense	63	141	44,70%
Roccabinca	156	426	36,60%
Sala Baganza	261	679	38,40%
Salsomaggiore Terme	805	2242	38,40%
San Secondo Parmense	265	646	41%
Sissa	191	478	40%
Solignano	7	19	36,80%
Soragna	209	561	37%
Sorbolo	566	1018	55,60%
Terenzo	17	31	54,80%
Tornolo	1	4	25%
Torrile	365	856	42,60%
Traversetolo	353	712	49,60%
Trecasali	159	410	39%
Valmozzola	6	13	46%
Varano dei Melegani	23	46	50%
Varsi	4	12	33%
Zibello	119	241	49,40%

Tabella 6.1: Percentuale di adesione su primo invito dal 01/01/03 al 31/12/05 per i singoli comuni della provincia di Parma

problema molto complesso, si è deciso di iniziare a lavorare con un modello monopolato, da raffinare e sviluppare in futuro.

6.3 Il modello

Presentiamo un possibile modello per rappresentare i dati relativi alla campagna di screening: sebbene il problema sia ancora in fase di studio e non sia ancora stato stabilito quale sarà il modello definitivo, è comunque utile illustrare le idee finora sviluppate, in modo da mostrare come un'Hamiltoniana possa essere effettivamente interpretata in termini sociologici.

Per modellizzare il comportamento delle donne si considera la situazione in due diversi momenti, ovvero periodo pre-screening (1991-1998), ed periodo in cui la campagna di screening è attiva (2003-2010). L'idea è quella di utilizzare un modello di Curie-Weiss, in cui l'Hamiltoniana costituisce la funzione utilità relativa al prendere parte al Pap test, e in cui l'azione del campo esterno corrisponde al sollecito dell'invito. In particolare, in fase pre-screening si avrà campo esterno nullo e l'unico contributo all'Hamiltoniana sarà dato dalle interazioni. Quando la campagna di screening è avviata, si suppone che ciò che prima era dato solo dalle interazioni, si trasformi in cultura individuale, e ad essa di aggiunga l'azione dell'invito.

Vediamo in dettaglio il modello che ne risulta.

Si indica con $t = 0$ la fase precedente allo screening, durante la quale si assume campo esterno nullo $h = 0$. L'Hamiltoniana in t_0 sarà quindi:

$$H_0(\sigma) = -\frac{J_0}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j \quad (6.1)$$

L'equazione di campo medio che si ottiene da (6.1) è la seguente:

$$m_0 = \tanh(J_0 m_0) \quad (6.2)$$

dove la magnetizzazione rappresenta la scelta media. Da (6.2):

$$J_0 = \frac{\operatorname{arctanh}(m_0)}{m_0} \quad (6.3)$$

In $t = 1$ la campagna di screening è attiva. Si suppone che a questo punto il fattore imitativo si trasformi in cultura, ovvero:

$$h_0 = \tanh^{-1}(m_0) = J_0 m_0$$

Dalla *conservazione della scala delle energie*, si ha che l'energia della fase pre-screening diventa il fattore di scala in $t=1$.

L'Hamiltoniana che ne deriva è la seguente:

$$H_1(\sigma) = -\frac{JJ_0}{2N} \sum_{i,j} \sigma_i \sigma_j - J_0 h \sum_{i=1}^N \sigma_i + h_0 \sum_{i=1}^N \sigma_i \quad (6.4)$$

dove il primo fattore rappresenta le interazioni, il secondo l'azione del capo esterno dovuta all'invito ed infine si ha la quantità che rappresenta la cultura acquisita. Da (6.4) si ricava la seguente equazione di campo medio:

$$m(J, h) = \tanh[J_0(Jm + h - m_0)] = \tanh[J_0(Jm + h) - h_0] \quad (6.5)$$

6.3.1 Il problema inverso teorico

Abbiamo visto, nel capitolo (4), il procedimento per invertire un modello di Curie-Weiss; si segue quindi lo stesso procedimento, per cui differenziando entrambi i membri dell'equazione (6.5) rispetto al campo esterno h si ha:

$$\chi = (1 - m^2(J, h))(J_0 J \chi + J_0)$$

e, raccogliendo χ :

$$\chi = \frac{J_0(1 - m^2(J, h))}{1 - J_0 J(1 - m^2(J, h))} \quad (6.6)$$

Invertendo tale equazione si ricava il parametro J ; andando a sostituire il valore così ottenuto in (6.5) si trova h :

$$\begin{cases} J = \frac{1}{J_0} \frac{1}{1 - m^2} - \frac{1}{\chi} \\ h = \frac{1}{J_0} \operatorname{arctanh}(m) - Jm + m_0 \end{cases} \quad (6.7)$$

I valori di m e χ sono ottenuti, al limite termodinamico, dal valor medio e dalla varianza della magnetizzazione $m_N(\sigma)$:

$$\begin{cases} \lim_{N \rightarrow \infty} \omega(m_N(\sigma)) = m(J, h) \\ \lim_{N \rightarrow \infty} \frac{\partial}{\partial h} \omega(m_N(\sigma)) = \chi \end{cases} \quad (6.8)$$

dove:

$$\begin{aligned} \frac{\partial}{\partial h} \omega(m_N(\sigma)) &= \frac{\partial}{\partial h} \left(\frac{\sum_{\sigma} m_N(\sigma) \exp(-H_N(\sigma))}{\sum_{\sigma} \exp(-H_N(\sigma))} \right) \\ &= J_0 N \frac{\sum_{\sigma} m_N^2(\sigma) \exp(-H_N(\sigma))}{\sum_{\sigma} \exp(-H_N(\sigma))} \\ &\quad - J_0 N \left(\frac{\sum_{\sigma} m_N(\sigma) \exp(-H_N(\sigma))}{\sum_{\sigma} \exp(-H_N(\sigma))} \right)^2 \\ &= J_0 N \left(\omega(m_N^2(\sigma)) - \omega^2(m_N(\sigma)) \right). \end{aligned} \quad (6.9)$$

6.4 Il problema dell'inversione

Per risolvere il problema inverso utilizzando i dati sperimentali a disposizione, bisogna stimare le quantità necessarie per esplicitare i parametri del modello. Dall'equazione (6.9) si evince che, per il modello sviluppato per rappresentare i dati relativi al programma di screening, come per il modello di Curie-Weiss, queste quantità sono $\omega_N(m_N(\sigma))$ e $\omega_N(m_N^2(\sigma))$.

Nel capitolo (4), per ricavare gli stimatori di $\omega_N(m_N(\sigma))$ e $\omega_N(m_N^2(\sigma))$, avevamo utilizzato il principio della massima verosimiglianza, da cui si ottengono

le stime (4.9). Nel caso in esame, però, tale metodo non può essere applicato in quanto non si dispone di un certo numero di configurazioni indipendenti in quanto:

- il periodo temporale che ricoprono (2002-2010) è troppo breve: se si pensasse ciascun anno come una configurazione indipendente, il numero di configurazioni ottenute sarebbe troppo basso ai fini statistici;
- se anche si disponesse di più anni o si considerassero periodi più brevi, come, ad esempio, trimestri o quadrimestri, l'ipotesi di indipendenza sarebbe troppo forte: non è plausibile che il comportamento di una donna sia totalmente scorrelato dal suo stesso comportamento nel passato;
- anche ammettendo l'indipendenza tra i diversi intervalli temporali scelti, i parametri J e h potrebbero cambiare nel tempo;

E' quindi necessario ricorrere ad una tecnica diversa. Si potrebbe pensare di sfruttare il metodo di Jaynes, massimizzando l'entropia rispetto ai parametri J ed h , il che equivale, di fatto, ad utilizzare il principio di massima verosimiglianza con un'unica configurazione. Consideriamo, per semplicità, il modello di Curie-Weiss, la cui Hamiltoniana abbiamo visto essere la seguente:

$$H_N(\sigma) = -\frac{J}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i \quad (6.10)$$

con relativa funzione pressione $p_N = \frac{1}{N} \ln Z_N$. Utilizzando la distribuzione di Boltzmann-Gibbs per rappresentare la probabilità, si ha che il logaritmo della funzione di verosimiglianza è dato da:

$$\ln P(\sigma) = -\ln Z_N + \frac{J}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j + h \sum_{i=1}^N \sigma_i \quad (6.11)$$

dove Z_N è la funzione di partizione. Poichè si ha $Z_N = Np_N$, da quanto visto per la funzione pressione, al limite termodinamico si ha:

$$\ln Z_N = N \sup_{\mu} p(\mu) \quad (6.12)$$

$$= N \left(-\frac{J}{2} \bar{\mu}^2 + \ln \cosh(J\bar{\mu} + h) \right) \quad (6.13)$$

dove $\bar{\mu}$ è appunto il valore in cui $p(\mu)$ raggiunge il sup e che quindi verifica l'equazione di campo medio $\bar{\mu} = \tanh(J\bar{\mu} + h)$. Applichiamo il principio della massima verosimiglianza, per cui si calcolano le derivate parziali rispetto a h e J di (6.11):

$$\begin{aligned} \frac{\partial \ln P(\sigma)}{\partial h} &= -N \left(-J\bar{\mu} \frac{\partial \bar{\mu}}{\partial h} + \tanh(J\bar{\mu} + h) \left(J \frac{\partial \bar{\mu}}{\partial h} + 1 \right) \right) + \sum_{i=1}^N \sigma_i \\ &= -N\bar{\mu} + \sum_{i=1}^N \sigma_i \\ \frac{\partial \ln P(\sigma)}{\partial J} &= -N \left(-\frac{\bar{\mu}^2}{2} - J\bar{\mu} \frac{\partial \bar{\mu}}{\partial J} + \tanh(J\bar{\mu} + h) \left(\bar{\mu} + J \frac{\partial \bar{\mu}}{\partial J} \right) \right) + \frac{1}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j \\ &= -N \left(\frac{\bar{\mu}^2}{2} \right) + \frac{1}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j \end{aligned} \quad (6.14)$$

Imponendo che le due derivate si annullino, otteniamo:

$$\begin{cases} \bar{\mu} = \frac{1}{N} \sum_{i=1}^N \sigma_i \\ \bar{\mu}^2 = \frac{1}{N^2} \sum_{i,j=1}^N \sigma_i \sigma_j = \left(\frac{1}{N} \sum_{i=1}^N \sigma_i \right)^2 \end{cases} \quad (6.15)$$

Utilizzando la massima verosimiglianza su un'unica configurazione, il problema non è risolvibile, in quanto otteniamo un unico stimatore per calcolare due parametri. Si sta quindi studiando un metodo alternativo, che sfrutta il concetto di debole dipendenza tra elementi di un sistema, per cui al limite termodinamico questi sono tra loro indipendenti. Avendo dati su un numero molto elevato di donne possiamo quindi supporre che siano tra loro indipendenti ed identicamente distribuite. La fotografia di cui disponiamo può così essere suddivisa in sottogruppi casuali di una data dimensione, e

questi possono essere considerati tra loro indipendenti; si creano così dei sottocampioni che consentono di ottenere le stime cercate. Va sottolineato che questi sottogruppi non costituiscono una partizione dell'insieme, una volta estratto un sottogruppo gli elementi vengono 'reimbussolati'. Se si considerasse una partizione, infatti, per avere un alto numero di campioni, questi andrebbero scelti di dimensione molto piccola. Con il nostro metodo, che definiamo *sottocampionamento* (*subsampling*), possiamo invece creare i sottocampioni di dimensione e in numero arbitrari. Il problema sarà quindi determinare come scegliere queste quantità.

Prima di poter applicare il sottocampionamento per invertire il nostro modello, lo si testa su delle configurazioni ottenute utilizzando il metodo di Monte Carlo e quindi parametri noti. Consideriamo, per semplicità, l'Hamiltoniana di Curie-Weiss (2.1) e fissiamo dei valori arbitrari per i parametri J ed h . Utilizzando il metodo di Monte Carlo si generano delle configurazioni di spin, che scegliamo di dimensione $N = 10000$. A questo punto si effettua il sottocampionamento e, con i dati creati, si risolve il problema inverso, verificando se i parametri così ottenuti corrispondono ai parametri di partenza. Dai test effettuati si è capito che per ottenere un buon match risulta essere di fondamentale importanza la dimensione scelta per i sottocampioni: esiste cioè una dimensione ottimale per la ricostruzione di J ed h , che varia al variare dei parametri. Si sta quindi cercando di capire come scegliere la lunghezza dei sottocampioni quando si opera sui dati reali, giacchè, in tal caso J e h sono le incognite da determinare. Si pensa di poter individuare, mediante considerazioni ragionate di tipo sociologico, una regione possibile per i parametri, sufficientemente piccola, in modo tale da sfruttare il metodo Monte Carlo per determinare la dimensione ottimale dei sottocampioni corrispondente a tale ragione e quindi scegliere la dimensione ottimale dei sottocampioni per i dati reali.

Capitolo 7

Prospettive future

Il nostro lavoro ha affrontato lo studio e l'analisi di un nuovo campo di ricerca, la Sociofisica, la cui evoluzione è tuttora *in itinere* e in cui si incontrano materie tradizionalmente in antitesi tra loro, quali le Scienze esatte e le Scienze Sociali.

Sebbene la Sociologia non sia una scienza galileiana, ovvero non possa essere indagata attraverso esperimenti controllati, ripetibili ed indipendenti, abbiamo visto come sia possibile avvalersi dei modelli della Meccanica Statistica per descrivere sistemi biologici, [36], [37], [4], sistemi economici, [19] e per affrontare diverse questioni sociali, dalla sanità ai sistemi di voto [22], fornendone un'originale chiave di lettura.

E' interessante osservare come vi sia un interscambio bidirezionale fra le discipline citate, che si completano e si integrano vicendevolmente per mezzo degli strumenti utilizzati ed attraverso i risultati raggiunti. In particolare, disporre di modelli quantitativi per studiare e comprendere i meccanismi del comportamento umano consente di valutare la rilevanza dei diversi fattori che concorrono a determinare le scelte di un gruppo di individui, mettendo in luce quanto rilevino le caratteristiche personali e quanto, invece, le influenze reciproche. L'ausilio della Fisica permette quindi alla Sociologia di definire gli schemi alla base delle dinamiche sociali, con lo scopo di rendere comprensibili ed, in futuro, parzialmente prevedibili i comportamenti collet-

tivi.

D'altra parte, anche la Fisica si è evoluta grazie all'apporto delle Scienze Sociali: le nuove tematiche e le nuove sfide proposte hanno stimolato un importante sviluppo per la Meccanica Statistica. Il lavoro di A.O. Sousa, K. Malarz e S. Galam [45] mostra, ad esempio, come grazie ad alcune simulazioni Monte Carlo eseguite per uno studio sociofisico, si siano scoperti nuovi risultati per un particolare modello di Ising bidimensionale.

Di fatto, la Sociofisica si serve degli strumenti della Meccanica Statistica pur trascurando molti aspetti della teoria fisica: gli esseri umani non sono particelle immerse in un bagno a temperatura T e non si ha un fattore di scala equivalente per l' 'Hamiltoniana sociale'. Quest'ultima, inoltre, contrariamente alla funzione energia della fisica, misurata in Joule, non ha unità di misura. Eppure, nonostante ciò, abbiamo visto come la distribuzione di Boltzmann-Gibbs, alla base dei modelli meccanico statistici, sembrerebbe essere la misura più adatta anche per i modelli sociali; per giustificare tale affermazione abbiamo illustrato come questa stessa distribuzione caratterizzi anche il modello Logit Multinomiale formulato in DCT e come emerga utilizzando un approccio innovativo, il Principio di Massima Entropia, che consente di ricavare dei modelli sostanzialmente analoghi a quelli della Meccanica Statistica, senza bisogno di formulare alcuna ipotesi fisica. L'obiettivo di Jaynes era quello di sostituire, con il suo principio, l'articolato sistema di argomentazioni fisiche connesse alla Meccanica Statistica, riducendo quest'ultima da branca della Fisica a tecnica di inferenza statistica. Pur non condividendo tali idee riteniamo che il suo metodo meriti di essere approfondito e che possa costituire un valido strumento di supporto per la formulazione di modelli nell'ambito della Sociofisica.

Abbiamo visto che non si tratta della prima volta nella storia in cui le scienze umane attingono alle scienze dure o viceversa, ma sicuramente, rispetto ai tentativi passati, la prospettiva attuale è assolutamente nuova: non si cercano più leggi universali che definiscano la natura umana o equazioni che pretendano di cogliere l'essenza dell'uomo, celando una visione determinis-

tica della realtà, ma si vuole, piuttosto, comprendere la struttura alla base di una rete sociale, servendosi della scienza come di *'a servant guide, not a dictator'*, [1]. Studiando la società attraverso l'ottica della Fisica non si intende privare l'essere umano della sua personalità, della sua irrazionalità e del suo libero arbitrio, ma si è osservato che esistono alcuni meccanismi fondamentali che emergono naturalmente in una rete di elementi interconnessi, indipendentemente dalla natura stessa della rete. Vi sono, cioè, alcune dinamiche che si ripetono quando si ha un insieme di unità interagenti, siano queste esseri umani o particelle.

Sicuramente la Sociofisica è una scienza 'giovane', in piena evoluzione e molta strada è ancora da percorrere: abbiamo visto come spesso si faccia ricorso a modelli di campo medio, che implicano assunzioni e semplificazioni forti, per cui tutti gli individui interagiscono con tutti gli elementi appartenenti allo stesso gruppo con la medesima intensità. Questi andranno quindi complicati in modo da cogliere anche gli aspetti più complessi; si potrà, ad esempio, pensare di elaborare dei modelli ispirati a quello di Sherrington e Kirkpatrick [43], in cui cioè le interazioni non hanno intensità costante ma sono distribuiti secondo una gaussiana.

Attraverso il progetto di ricerca sulla campagna di screening attuata nella provincia di Parma, abbiamo mostrato cosa significhi lavorare con un insieme di dati empirici per sviluppare un modello. Nonostante le difficoltà incontrate, lo studio è molto promettente e si pensa che presto si avranno tutti gli strumenti necessari per risolvere il problema inverso ed individuare il modello più appropriato per rappresentare i dati. Questo consentirà di capire quali siano i fattori che influenzano maggiormente la scelta di una donna riguardo al prendere parte alla campagna di screening e di conseguenza si potrà suggerire la strategia migliore per innalzare le percentuali di adesione. Questi studi sono animati, quindi, dalla speranza che, utilizzando i metodi delle scienze dure, si potranno attuare le politiche sociali più efficaci.

Appendice A

La distribuzione di Gumbel

La distribuzione di Gumbel appartiene alla classe delle distribuzioni di valore estremo, ovvero quelle distribuzioni che descrivono il comportamento del massimo (o del minimo) di una sequenza di variabili indipendenti ed identicamente distribuite.

Sia ϵ una variabile con distribuzione di Gumbel; allora si ha la seguente **funzione caratteristica**:

$$F(\epsilon) = \exp[-e^{-\mu(\epsilon-\mu)}], \quad \mu > 0 \quad (\text{A.1})$$

e **densità di probabilità**:

$$f(\epsilon) = \mu e^{-\eta(\epsilon-\eta)} \exp[-e^{-\eta(\epsilon-\eta)}]$$

dove η è il '*location parameter*', mentre μ è un parametro di scala positivo.

La distribuzione di Gumbel gode delle seguenti proprietà:

1. La media è $\mathbb{E}(\epsilon) = \eta + \frac{\gamma}{\eta}$, mentre la varianza è $\text{Var}(\epsilon) = \frac{\pi^2}{6\eta^2}$, dove η è la costante di Eulero-Mascheroni (~ 0.577);
2. la moda è η .
3. Se ϵ ha distribuzione di Gumbel con parametri (η, μ) e V e $\alpha > 0$ sono delle costanti arbitrarie, allora $\alpha\eta + V$ ha distribuzione di Gumbel di parametri $(\alpha\eta + V, \mu/\alpha)$.

4. Se ϵ_1 e ϵ_2 sono due variabili indipendenti con distribuzione di Gumbel, rispettivamente di parametri (η_1, μ_1) e (η_2, μ_2) , allora $\epsilon = \epsilon_1 - \epsilon_2$ ha distribuzione logistica, con funzione di ripartizione e densità di probabilità rispettivamente:

$$\begin{aligned} F(\epsilon) &= \frac{1}{1 + e^{-\mu(\eta_2 - \eta_1 - \epsilon)}} \\ f(\epsilon) &= \frac{\mu e^{-\mu(\eta_2 - \eta_1 - \epsilon)}}{(1 + e^{-\mu(\eta_2 - \eta_1 - \epsilon)})^2}. \end{aligned} \quad (\text{A.2})$$

5. Date k variabili indipendenti con distribuzione di Gumbel, si ha:

$$\max_{i=1, \dots, k} =^d \text{Gumbel} \left(\frac{1}{\mu} \ln \sum_{i=1}^k e^{\mu \eta_i} \right). \quad (\text{A.3})$$

Le prime due proprietà derivano semplicemente dalla definizione della distribuzione. La proprietà (3) afferma che la distribuzione di Gumbel si conserva rispetto a trasformazioni lineari; la (4) mostra la distribuzione di una variabile ottenuta come differenza di due variabili indipendenti con distribuzione di Gumbel che hanno lo stesso parametro di scala μ . L'ultima proprietà, infine, afferma che il massimo tra un insieme di variabili di Gumbel con lo stesso parametro di scala ha ancora distribuzione di Gumbel, con medesimo μ , e il cui parametro di locazione $\frac{1}{\mu} \ln \sum_{j=1}^J e^{\mu \eta_j}$ corrisponde al valore di aspettazione del massimo meno la costante (γ/μ) .

Appendice B

Teorema di esistenza del limite per i modelli di campo medio

Si dimostra in questa sezione il teorema di Bianchi, Contucci e Giardinà, che assicura l'esistenza del limite termodinamico della funzione pressione per i modelli di campo medio con Hamiltoniana che soddisfa una condizione di superadditività per lo stato di Boltzmann-Gibbs. Si consideri la Hamiltoniana di campo medio:

$$H_N = -Ng(m_N) \tag{B.1}$$

dove $g : [-1, 1] \rightarrow \mathbb{R}$ è una data funzione limitata della magnetizzazione $m_N(\sigma)$. A (B.1) si associano la funzione di partizione Z_N :

$$Z_N = \sum_{\sigma \in \Omega_N} e^{-\beta H_N(\sigma)},$$

la pressione:

$$p_N = \frac{1}{N} \ln Z_N$$

e, per ogni osservabile del sistema $f(\sigma)$, lo stato di Boltzmann-Gibbs:

$$\omega_N(f) = \frac{\sum_{\sigma \in \Omega_N} f e^{-\beta H_N(\sigma)}}{Z_N}.$$

Vale il seguente teorema:

Teorema B.1. Sia $H_N(\sigma)$ l'Hamiltoniana di campo medio (B.1). Se, per ogni partizione dell'insieme $\{1, \dots, N\}$ in $\{1, \dots, N_1\}$ e $\{N_1+1, \dots, N_2\}$ con $N = N_1 + N_2$ e

$$H_{N_1} = H_{N_1}(\sigma_1, \dots, \sigma_{N_1}) \quad H_{N_2} = H_{N_2}(\sigma_{N_1}, \dots, \sigma_N),$$

si verifica che

$$\omega_N(H_N) \geq \omega_N(H_{N_1}) + \omega_N(H_{N_2}) \quad (\text{B.2})$$

allora esiste il limite termodinamico della funzione pressione:

$$\lim_{N \rightarrow \infty} p_N = \inf_N p_N = p$$

Prima di dimostrare il teorema si definisce l'Hamiltoniana interpolante e si dimostrano tre lemmi.

Definizione B.2. Si dice Hamiltoniana interpolante la seguente funzione del parametro $t \in [0, 1]$:

$$H_N(t) = tH_N + (1-t)[H_{N_1} + H_{N_2}] \quad (\text{B.3})$$

con funzione di partizione $Z_N(t)$, pressione $p_N(t)$ e stato di Gibbs $\omega_N(H_N(t))$.

Lemma B.3. Sia H_N l'Hamiltoniana di campo medio e $H_N(t)$ la sua interpolante. Se

$$\frac{d}{dt} p_N \leq 0$$

per ogni $t \in [0, 1]$, allora si ha che:

$$p_N \leq \frac{N_1}{N} p_{N_1} + \frac{N_2}{N} p_{N_2}$$

per ogni decomposizione $N = N_1 + N_2$.

Dimostrazione. Segue direttamente dal teorema fondamentale del calcolo e dalla definizione (B.3), per cui:

$$\begin{aligned} Z_N(1) &= Z_N, & p_N(1) &= p_N, \\ Z_N(0) &= Z_{N_1} Z_{N_2}, & p_N(0) &= \frac{N_1}{N} p_{N_1} + \frac{N_2}{N} p_{N_2}, \end{aligned}$$

□

Lemma B.4. *Calcolando la derivata rispetto a t di $p_N(t)$ si ottiene:*

$$\begin{aligned} p'_N(t) &= \frac{d}{dt} \frac{1}{N} \log Z_N(t) = -\frac{\beta}{N} \sum_{\sigma \in \Omega} [H_N - H_{N_1} - H_{N_2}] \frac{e^{-\beta H_N(t)}}{Z_N(t)} \\ &= -\frac{\beta}{N} \omega_{N,t}[H_N - H_{N_1} - H_{N_2}] \end{aligned} \quad (\text{B.4})$$

Lemma B.5. *La derivata seconda di $p_N(t)$ è positiva:*

$$p''_N(t) = \frac{d^2}{dt^2} p_N(t) \geq 0$$

Dimostrazione. Il risultato si ottiene semplicemente dal calcolo diretto della derivata seconda $p''_N(t)$:

$$\begin{aligned} p''_N(t) &= \frac{d}{dt} \left(\frac{\beta}{N} \omega_{N,t}[H_N - H_{N_1} - H_{N_2}] \right) \\ &= \frac{\beta^2}{N} (\omega_{N,t}[(H_N - H_{N_1} - H_{N_2})^2] - \omega_{N,t}^2[(H_N - H_{N_1} - H_{N_2})]). \end{aligned} \quad (\text{B.5})$$

Dalla disuguaglianza di Jensen applicata alla funzione $x \rightarrow x^2$, segue che $p''_N(t) \geq 0$. \square

Dimostrazione del teorema B.1. Dal lemma (B.4) si ha che l'ipotesi (B.2)

$$\omega_N(H_N) \geq \omega_N(H_{N_1}) + \omega_N(H_{N_2})$$

equivale alla condizione:

$$p_N(1) \leq 0.$$

D'altra parte, dal lemma (B.5), $p'_N(t)$ è una funzione crescente in t , e poichè si ha $p'_N(1) \leq 0$, allora:

$$p'_N(t) \leq 0, \quad \forall t \in [0, 1]$$

Dal lemma (B.3), p_N è subadditiva, per cui:

$$\lim_{N \rightarrow \infty} p_N = \inf_N p_N$$

L'esistenza del limite segue dalla limitatezza della funzione g . Per cui, definendo K il massimo della funzione $g(x)$ in $[-1, 1]$, si ottiene:

$$p_N = \frac{1}{N} \ln \sum_{\sigma \in \Omega_N} e^{\beta N g(m_N)} \geq \frac{1}{N} \ln e^{\beta N K} = \beta K$$

□

Bibliografia

- [1] Ball P., *Critical Mass*, Arrow books, 2004
- [2] Barra A., The Mean Field Ising Model through Interpolating Techniques, *J.Stat. Phys.* **145**: 234-261, 2008
- [3] Ben-Akiva M., Lerman S.R., Discrete Choice Analysis, *The MIT Press*, 1985
- [4] Bialek W., Cavagna A., Giardinà I., Mora T., Silvestri E., Viale M., Walczak M., Statistical Mechanics for Natural Flock of Birds, *arXiv:1107.0604v1*, 2011
- [5] Bianchi A., Contucci P., Giardinà C., Thermodynamic limit for mean field spin models, *Math. Phys. E J* **9**, 2004
- [6] Boltzman L., Weitere Studien über das Wärmegleichgewicht unter Gasmolek^ül, 1872. Cited in Bart J.A., *Wissenschaftliche Abhandlungen*, ed. F. Hasenöhrl, vol1, 317., Leipzig
- [7] Box G.E.P., co-authored with Norman R. Draper, *Empirical Model-Building and Response Surfaces*, p. 424, 1987.
- [8] Brock W., Durlauf S., Interaction Based Models, *Handbook of Econometrics*, vol.5, 3297-3380 (2001).
- [9] Brock W., Durlauf S., Discrete Choice with Social Interaction, *Review of Economic Studies*, 68: 235-260 (2001).

-
- [10] Comte A., *Opuscules de philosophie sociale*, cited in Enciclopedia Larousse
- [11] Contucci P., Ghirlanda S., Modelling Society with Statistical Mechanics: an Application to Cultural Contact and Immigration, *arXiv:physics/0606062v1*, 2006
- [12] Curie P., Propriété ferromagnétique des corps à diverse températures, *Ann. de Chim. et de Phys.*, 7^e série, **V**: 289, 1895
- [13] Dam L., Hol L., de Bekker-Grob E.W., Steyerberg E.W., Kuipers E.J., Habbema J.D.F., Essink-Bot M.L., van Leerdman M.E., What individual preferences for colorectal cancer screening programmes? A discrete choice experiment, *Eur J Cancer* **46**, 150-159, 2010
- [14] De Condorcet J.A.N., (1785) Essay sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix, Quoted in P. Ball, *Critical Mass*, Arrow books, 2004
- [15] De Sanctis L., Structural Approaches to Spi Glasses and optimization Problems, Ph.D. Thesis, Department of Mathematics, Princeton University, 2005
- [16] Durlauf S., How Can Statistical Mechanics Contribute to Social Science?, *Proceeding of the National Academy of Science of the USA*, vol. 96, No. 19, 10582-10584, 1999
- [17] Ellis R.S., Remarks on an Alleged proof of the Method of Least Squares, contained in a late number of *Edinburgh Review*, 1850, in W. Walton (ed), *Mathematical and other Writings of R.L. Ellis*, Cambridge University Press, 53-61, 1863.
- [18] Ellis R. S., *Large Deviations and Statistical Mechanics*, Springer, 1985
- [19] Föllmer H, Random Economies with Many Interacting Agents, *J. Math. Econ.* **1**: 51-62, 1973

-
- [20] Fox J., Daly A.J., Gunn H., *Review of RAND Europes Transport Demand Model System*, RAND, 2003
- [21] Galam S., Sociophysics: a Personal Testimony, *Physica A: Statistical and Theoretical Physics*, 336(1-2) 49-55 (2004).
- [22] Galam S., Sociophysics: A Review of Galam Models, *Internat. J. Modern Physics C*, **19**(4) 409-440, 2008
- [23] Gallo I., Contucci P., Bipartite Mean Field Spin System. Existence and Solution, *MPEJ* **14**, 2008
- [24] Gerard K., Shanahan M., Louviere J., Using Stated Preference Discrete Choice Modelling to Inform Health Care Decision-making: a Pilot Study of Breast Screening Participation, *Applied Economics* **35**(9): 1073-1085, 2003
- [25] Granovetter M., Threshold model of collective behaviour, *Am. J. Sociol.* **83**: 1420-1443, 1978
- [26] Guerra F., Toninelli F. L., The Thermodynamic Limit in Mean Field Spin Glasses Models, *Communications in Mathematical Physics* **230**, 2002
- [27] Guerra F., Spin Glasses, *cond-mat/0507581*, 2006
- [28] Guerra F., Mathematical Aspects of Mean Field Spin Glass Theory, *cond-mat/0410435*, 2005
- [29] Jaynes E.T., Information Theory and Statistical Mechanics, *Phys Rev Lett* **106**, 620-630, 1957.
- [30] Jaynes E.T., Papers on Probability, Statistics and Statistical Physics, *R. Rosenkrantz (ed) (Dordrecht:Reidel)*, 1981
- [31] Laidler, Keith J. (1995). **The Physical World of Chemistry**, Oxford University Press, 104, 105

-
- [32] Lavis D.A., Milligan P.J., The Work of E.T. Jaynes on Probability, Statistics and Statistical Mechanics, *British Journal for the Philosophy of Science* **36**, 193-2010
- [33] McFadden D., Economic Choices, *The American Economic Review* **91**: 351-378, 2001
- [34] Michard Q., Bouchaud J.P., Theory of Collective Opinion Shifts: from Smooth Trends to Abrupt Swings, *arXiv: 0504.079v1*, 2005
- [35] Paag H., Daly A.J., Rohr C., Predicting Use of the Copenhagen Harbour Tunnel, in david hensher, *Travel Behaviour Research: the Leading Edge*, Pergamon, 2001
- [36] Roudy Y., Tyrcha J., Hertz J., The Ising Model for neural data: Model Quality and Approximate Methods for Extracting Functional Connectivity, *arXiv: 0902.2885v1*, 2009
- [37] Roudy Y., Aurell E., Hertz J., Statistical physics of pairwise probability models, *arXiv: 0905.110v1*, 2009
- [38] Ryan M., Gerard K., Using Discrete Choice Experiments to Value health Economics and Health Policy, *Applied Health Economics and Health Policy*, **2**(1): 55-64, 2003
- [39] Ryan M., Netten A., Skatn D., Smith P., Using Discrete Choice Experiments to Estimate a Preference-based Measure of Outcome. An Application to Social Care for Older People, *Journal of Health Economics*, **25**(5): 927-944, 2006.
- [40] Ruelle D., *Statistical Mechanics: rigorous results*, Addison Wesley, 188, 1989.
- [41] Schelling T., Dynamic Models of Segregation, *Journal of Mathematical Sociology*, **1**, 143-186, 1971

-
- [42] Schelling T., *Micromotives and Macrobehaviour*, W W Norton & Co Ltd, 1978
- [43] Sherrington D., Kirkpatrick S., Solvable Model of a Spin Glass, *Phys. Rev. Lett*, **35** 1792, 1975
- [44] Shannon C., A Mathematical Theory of Communication, *Bell Sym. Tech. J.* **27**, 379-423 & 623-656, 1948
- [45] Sousa A.O., Malarz K., Galam S., Reshuffling Spins with Short Range Interactions: When Sociophysics Produces Physical results, *arXiv:cond-mat/0503702v1*, 2005
- [46] Talagrand M., *Spin Glasses: a Challenge for Mathematicians. Cavity and Mean Field Models*, Springer Verlag, 2003
- [47] Tanaka T., *Men-field Theory of Boltzmann Machine Learning*, Physical Review E **58**, no. 2, 2302-2310, 1998.
- [48] Thompson C.J., *Classical Equilibrium Statistical Mechanics*, Clarendon Press, 1988
- [49] Uffink J., Can the Maximum Entropy Principle Be Explained as a Consistency Requirement?, *Studies in History and Philosophy of Modern Physics B* **26**, 223, 1995
- [50] Van Hemmen J.L., Van Enter A.C.D., Canisius J., On a Classical Spin Glass Model, *Z Phys B* **50**: 311-336, 1983
- [51] Weiss P., L'hypothese du champ moléculaire e la propriété ferromagnétique, *J. di Phys.*, 4^e série **VI**: 661, 1907

Ringraziamenti

Arrivata a questo importante traguardo, desidero esprimere la mia riconoscenza verso le persone che mi sono state vicine in questi anni e mi hanno aiutato, in diversi modi, a raggiungere questo importante traguardo.

Ringrazio innanzitutto chi, concretamente, ha contribuito alla realizzazione di questo lavoro: il *Prof. Pierluigi Contucci*, per i suoi preziosi insegnamenti e per avermi spronato anche nei miei momenti di indecisione, e la *Dott.ssa Micaela Fedele* per la sua disponibilità e per tutto il tempo e l'attenzione che mi ha dedicato.

Un grande, grandissimo grazie va alla mia famiglia, in particolare ai miei *genitori* per il loro profondo amore e per avermi dato la libertà di seguire la mia strada, equipaggiandomi di tutti gli strumenti necessari per poterla affrontare al meglio e grazie alla mia splendida *sorellina* per il suo dotto aiuto nella stesura di questa tesi ma soprattutto perchè il nostro indissolubile legame è, per me, fonte di quella forza e quella sicurezza che mi guidano nel mio cammino.

Vorrei esprimere tutta la mia gratitudine allo staff dello *Scacco Matto* perchè, accogliendomi come 'infiltrata speciale' nella loro brigata, mi ha fatto vivere un'esperienza indimenticabile, mi ha dato importanti lezioni di cucina e di vita, mi ha permesso di conoscere delle persone meravigliose e mi ha trasmesso tutta la serenità e la grinta necessarie per realizzare i miei studi.

Come non ringraziare poi i colleghi che hanno affrontato con me questo percorso: grazie a *Sara*, compagna di studi, di serate, di viaggi, con la quale ho condiviso dubbi ed incertezze ed ho superato molti degli scogli di questi

due anni; grazie a *Candia*, collega ed amica speciale, per tutto ciò che ha fatto per me e per essermi sempre vicina, grazie ad *Alberto* piccolo grande amico...perchè tanto ci prendiamo in giro quanto ci vogliamo bene; grazie a *Rita* per non essersi spaventata al nostro primo incontro e per le belle serate insieme e grazie a *Michele* per essere stato sempre così disponibile e pronto a darmi una mano.

Vorrei poi dire un grazie di cuore a tutti gli amici che ho la fortuna di aver avuto accanto in questi anni, che mi hanno fatto ridere, divertire, incuriosire, crescere, maturare, arrabbiare, commuovere, inorgoglire, e molto altro e senza i quali non sarei quella che sono. In particolare, grazie a *Martina* per i tanti, indimenticabili ricordi, perchè se non ci fosse non sarei in grado di inventarla e perchè che mondo sarebbe senza Martina non lo so, ma so che non mi piacerebbe; grazie a *Marta* per tutto ciò che abbiamo condiviso: serate, risate, vacanze, incontri, scontri e confidenze e perchè la nostra amicizia costituisce uno dei punti fermi della mia vita; grazie a *Raffaella* perchè, nonostante tutto, nei momenti importanti della mia vita c'è sempre, pronta a gioire per i miei successi, a consolarmi nei momenti brutti e...sgridarmi per le mie marachelle; grazie a *Mirko* per i suoi saggi consigli e per essermi stato così vicino in questo momento particolare; grazie a *Sabina* per avermi sempre stimolato a guardare più a fondo dentro ed intorno a me e per avermi insegnato che "le uniche cose che sbagli sono quelle che non provi a fare".

Un grazie, infine, a *Bologna* per tutti i momenti indicabili ed i tanti ricordi che conserverò sempre con me, dovunque la vita mi porterà.