Fields Institute Distinguished Lecture

Series In Statistical Sciences

Probabilistic Phenomena in Mathematics and Science

Don Dawson

UBIQUITY OF PROBABILISTIC PHENOMENA

Certain basic statistical patterns and probabilistic structures arise in empirical data and theoretical models from completely different fields.

The Workshop on Current and Emerging Research Opportunities in Probability – 2002 identified the following areas

- ***Algorithms***
- ***Statistical Physics***
- ***Dynamical and physical systems***
- ***Complex networks***
- ***Mathematical finance, risk and dependency***
- ***Perception in artificial systems***
- ***Genetics and ecology***

Outline

FIRST CIRCLE OF IDEAS:

Probability laws and universality classes

SECOND CIRCLE OF IDEAS:

Markovian dynamics and conditional independence.

THIRD CIRCLE OF IDEAS:

Probabilistic modelling of reversible interacting systems

FOURTH CIRCLE OF IDEAS:

Spatially distributed nonreversible stochastic dynamics

FIFTH CIRCLE OF IDEAS:

Hierarchy, Genealogy and History

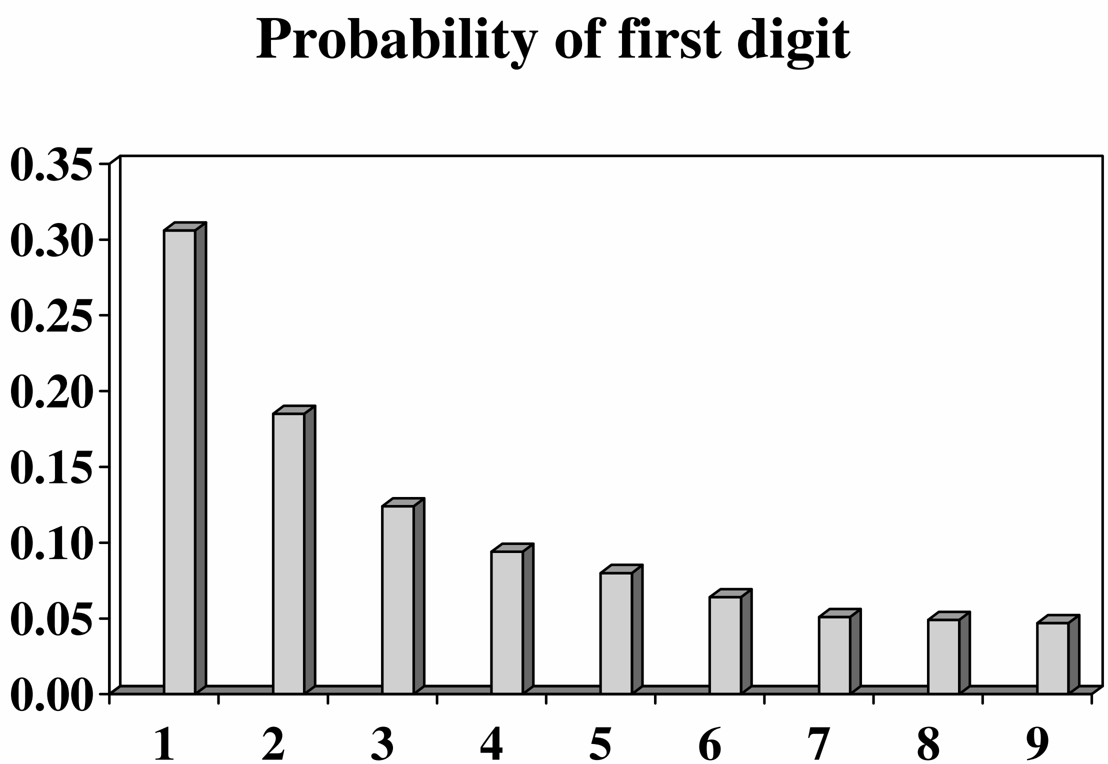
SIXTH CIRCLE OF IDEAS:

Universality classes of spatial and space-time structures

Empirical discovery of probability laws

Example:Benford's Law

- Simon Newcomb (1881) - tables of logarithms.
- Frank Benford (1938) - 20,229 data sets (molecular weights of chemical compounds, population sizes, etc.



Benford's distribution of first digit

$$\boxed{p = log_{10}(1 + 1/d)}$$

where p = the probability that the first significant digit is d.

- Statistical derivation - T.P. Hill (1996)
  - Introduced idea of Scale invariant and Base invariant distributions

Discovery of the Central Limit Theorem

$\boxed{\text{The Normal (Gaussian) Law}}$

$$P(a < X < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

Abraham De Moivre 1733 - aproximation to Binomial - games of chance

Johann Carl Friedrich Gauss 1809 - Law of Errors - astronomical data

Adolphe Quetelet ~1830 "the average man"

Francis Galton ~1860 - biological data - inheritance

Surprising Places: Kac-Steinhaus Example

$$x_n(t) := \frac{\cos(\lambda_1 t) + \cos(\lambda_2 t) + \cdots + \cos(\lambda_n t)}{\sqrt{n}}$$

$\{\lambda_i\}$ are are linearly independent over the field of rationals.

$$L_n(a, b) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} 1_{(a,b)}(x_n(t))dt$$

$$\lim_{n \to \infty} L_n(a, b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

Q. WHY does the Gauss (normal) law arise?

"Even before the climax of our search for the meaning of independence was reached, it became abundantly clear why tout le monde was justified in believing in the loi des erreurs. It proved to be both un fait d'observation and une théorème de mathématiques."

- referring to his work with Steinhaus (1935-38)

Mark Kac - Enigmas of Chance

Basic Concepts

$$X_1, \ldots, X_N$$

zero mean independent identically distributed random variables.

Independence and the Characteristic Function

$$E(e^{i\theta S_N}) = E(e^{i\theta X_1} \cdots e^{i\theta X_N}) = E(e^{i\theta X_1}) \cdots E(e^{i\theta X_N})$$

where $\quad S_N := (X_1 + \ldots X_N)$

Scaling Limit

$$X_n(t) := \frac{1}{a_n} S_{\lfloor nt \rfloor}$$

$$\Psi_t^n(\theta) = E(e^{i\theta X_n(t)}) \to \Psi_t(\theta)$$

Properties of the Scaling Limit

$$\Psi_{t+s} = \Psi_t \Psi_s$$

$$= E(e^{i\theta(X_t - X_0)} e^{i\theta(X_{t+s} - X_t)})$$

$$P_{t+s} = P_t \bigstar P_s \quad \text{Convolution Semigroup}$$

- Infinite Divisibility $\boxed{\text{Khinchin-Levy Representation}}$
- Functional Fixed Point Equation

$$\psi_1(\theta) = F(\psi_1)(\theta)$$

$$F(\psi_1)(\theta) = \psi_1^2(\alpha\theta) \text{ for some } \alpha > 0.$$

Universality Classes
Fixed Points and their Domains of attraction

- The Gaussian Law - “Short Tails”

$$\Psi_t(\theta) = \exp\left(-\frac{\theta^2 t}{2}\right),$$

$$P_t(dx) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{|x|^2}{2t}} dx \qquad \text{in } \mathbb{R}^d$$

- The Stable Laws - “Long Tails”- “Noah Effect”

$$\Psi_t(\theta) = \exp(-t|\theta|^\alpha), \quad 0 < \alpha < 2$$

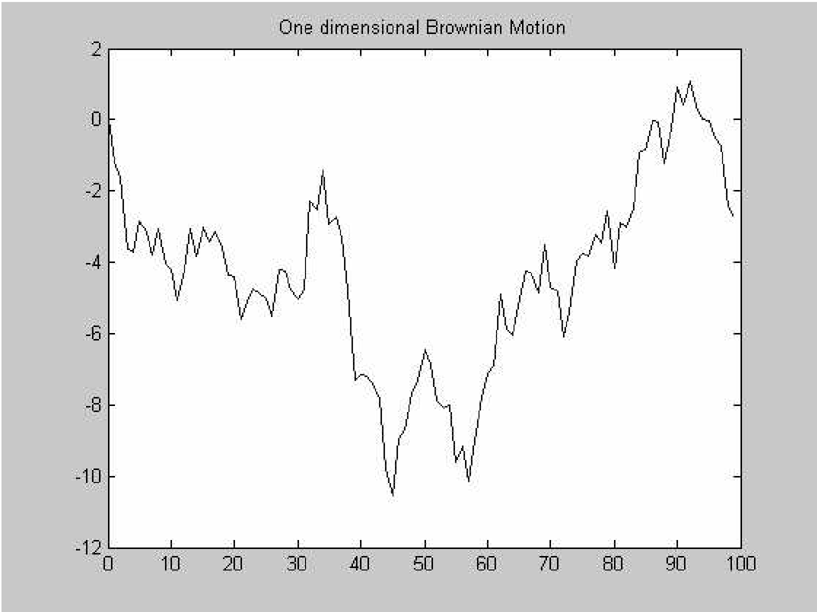
The Function Space Perspective - Brownian Motion

- Independent increments

$$P_{x_0}(B_{t_1} \in dx_1, \cdots, B_{t_n} \in dx_n)$$
$$= p_{t_1}(x_1 - x_0) p_{t_2 - t_1}(x_2 - x_1) \ldots p_{t_n - t_{n-1}}(x_n - x_{n-1})$$

- Law of Brownian Motion = Wiener Measure (1923):

$$P_{x_0} = \text{ measure on } C([0, \infty)).$$



- **Donsker's Invariance Principle (1952)**

$$X_n(t) := \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}$$

$$X_n(t) \Longrightarrow B(t)$$

$B(t) =$ **Brownian motion**

- **Self-similarity**

$$B(ct) = \sqrt{c} B(t)$$

- $t \to B_t$ **Almost surely non-differentiable, positive quadratic variation.**

Brownian motion is an incredibly rich mathematical object and serves as a key building block of stochastic analysis. Mark Yor:

"I was extremely astonished at the number of very natural questions which have escaped attention until very recently."

One of the questions he raises is "*To understand better the ubiquity of Brownian motion in a great number of probabilistic problems* "

SECOND CIRCLE: Dynamics and Conditional Independence

Markov property Conditional independence of past and future

- **Markov Semigroup - Brownian Motion**

$$T_t f(x) = E_x(f(B_t)) = \int f(y) P_t(x-y) dy$$

$$T_{t+s} = T_t T_s$$

$$u(t,x) := T_t f(x)$$

$$\frac{\partial u(t,x)}{\partial t} = \frac{1}{2} \Delta u(t,x)$$

The Impact of Ito Calculus

$$dF(B(t)) = F'(B(t))dt + \frac{1}{2}F''(B(t))dB(t)$$

Brownian Motion as building block: Stochastic Differential Equations.

Geometric Brownian Motion:

(multiplicative random effects, multiplicative CLT)

$$dx(t) = \mu x(t)dt + \sigma x(t)dw(t)$$

$$x(t) = e^{(\mu - \frac{\sigma^2}{2})t + \sigma w(t)}$$

- Black-Scholes formula for option pricing
- $\sigma =$ volatility, volatility modelling

Brownian Motion with Killing - Feynman Kac change of probability measure

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u - Vu$$
$$u(0) = f \geq 0$$

$$T_t f(x) = E_x\left[\exp(-\int_0^t V(x_s)ds)f(x_t)\right]$$

Mutation-Selection, nonlinear Genetic algorithm:

$$T_t f(x) = \frac{1}{Z}E_x\left[\exp(-\int_0^t V(x_s)ds)f(x_t)\right]$$
$$Z = \text{ normalizing constant}$$

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u - Vu + u\int V(x)u(x)dx$$

Gradient motion in a potential V

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2}\Delta u - \nabla u . \nabla V$$
$$dx(t) = -\nabla V(x(t))dt + \sigma dw(t) \qquad \text{Ito SDE}$$

Gibbs Distribution: $\lim_{t\to\infty} T_t f(x) = \frac{1}{Z}\int f(x)\exp(-\frac{1}{\sigma^2}V(x))dx$

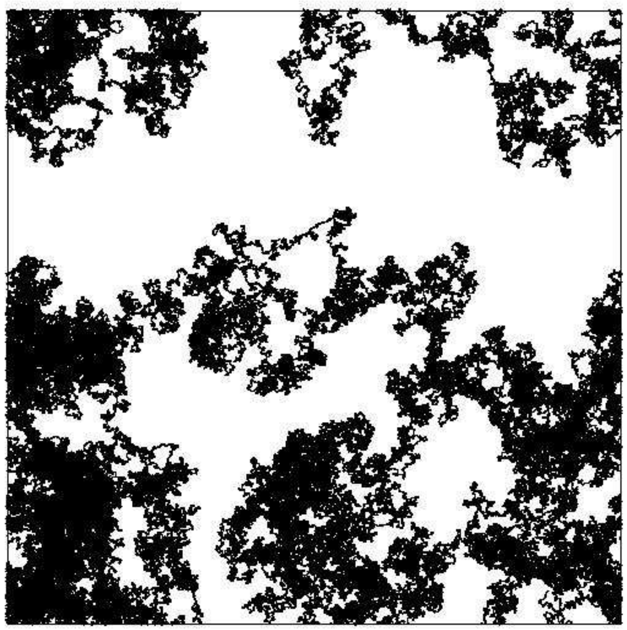
Example: Ornstein-Uhlenbeck

$$V(x) = \frac{1}{2}\gamma x^2$$
$$dx(t) = -\gamma x(t)dt + \sigma dw(t)$$

→ Method of simulated annealing.

The Mystery of Dimension

- **Two dimensional Brownian motion**



- Low dimensions
- Intermediate dimensions
- High dimensions

Dynamics of Eigenvalues and Wigner's Semicircle Law

Real symmetric $n \times n$ random matrices: $M\ (t)$ - the upper triangle entries are independent OU.

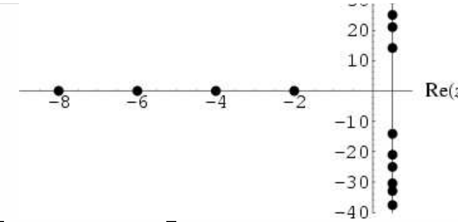
Eigenvalues $\lambda_j(t)$:

$$d\lambda_j = \left(\frac{1}{n} \sum_{i \neq j} \frac{dt}{\lambda_j - \lambda_i} - \lambda_j \right) dt + \sqrt{\frac{1}{n}} dw_j$$

Semicircle Law: Empirical distribution of eigenvalues converges to

$$\lim_{n \to \infty} \frac{1}{n} N_n(\lambda) = \frac{2}{\pi} \int_{-1}^{\lambda} \left(1 - u^2\right)^{\frac{1}{2}} du, \quad |\lambda| \leq 1$$

Universality: $n \times n$ real symmetric matrices with $E(\xi_{ij}^2) = \frac{1}{4},\ i < j,\ E(\xi_{ii}^2) \leq const$ + subgaussian decay at infinity

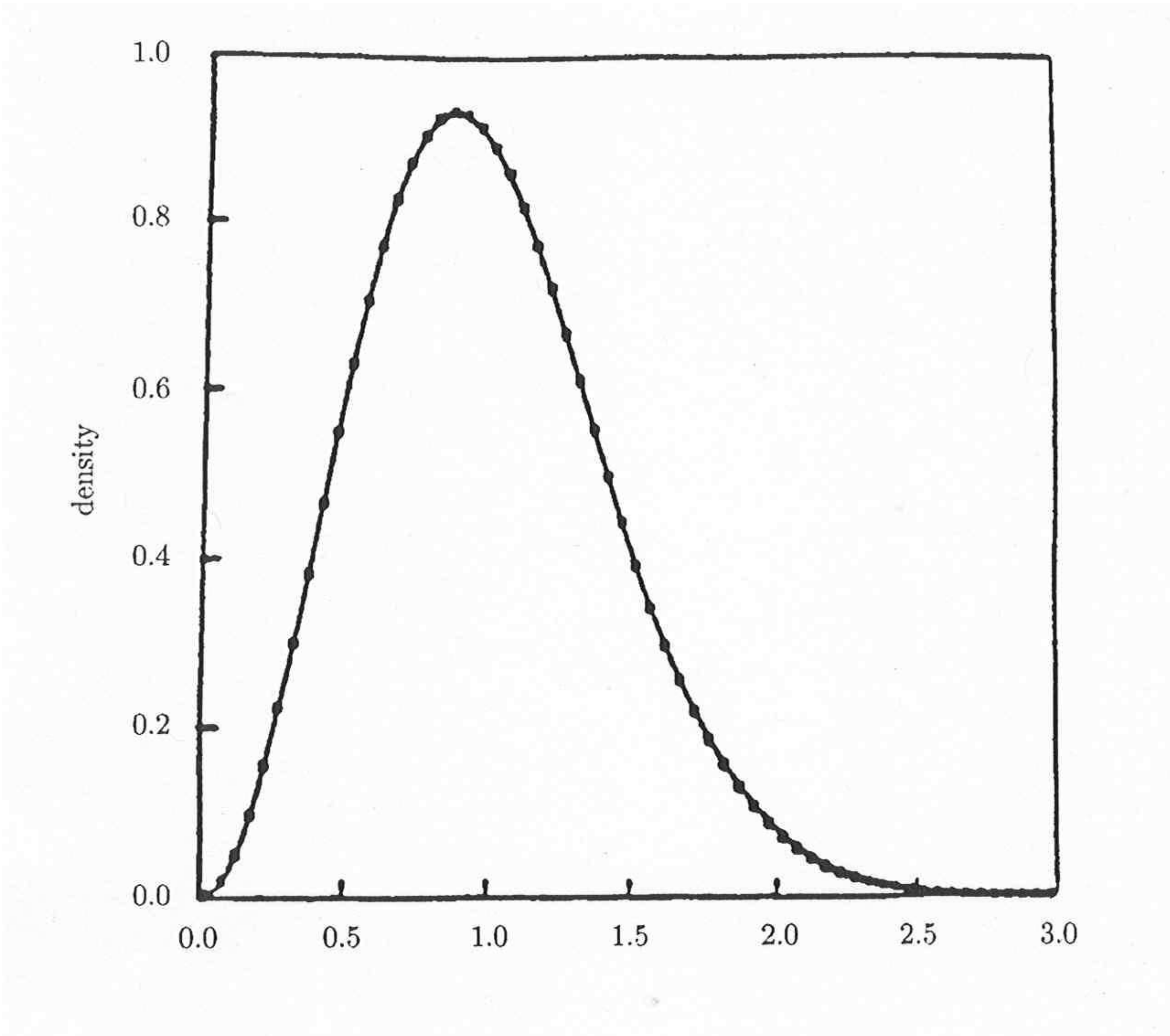


Gaussian unitary ensemble:

Real and complex off-diagonal entries are independent Gaussian.

- **"GUE-Hypothesis":** (Dyson – Montgomery) **the spacings of N successive zeros of the Riemann zeta function and eigenvalues of $N \times N$ Hermitian matrices have the same statistical properties in the $N \to \infty$ limit.**
- **Persi Diaconis: "Statistical test of hypothesis" – Andrew Odlyzko's computation of the 10^{20}th zero of the Riemann zeta function and 175 million of it neighbours.**
- **Universality Rudnick and Sarnak: universality of the behavior of correlations between successive zeros for a class of L-functions.**

Nearest neighbor spacing among 70 million zeros beyond the 10^{20}th zero of the zeta function compared to the Gaussian Unitary Ensemble



⇒ **THIRD CIRCLE:**

Probabilistic Modelling of Interacting Systems

Gibbs Random Fields and Ising Model

Configurations on cube of side K, C_K in $\mathbb{Z}^d$. $E = \{\pm 1\}^{C_K}$.

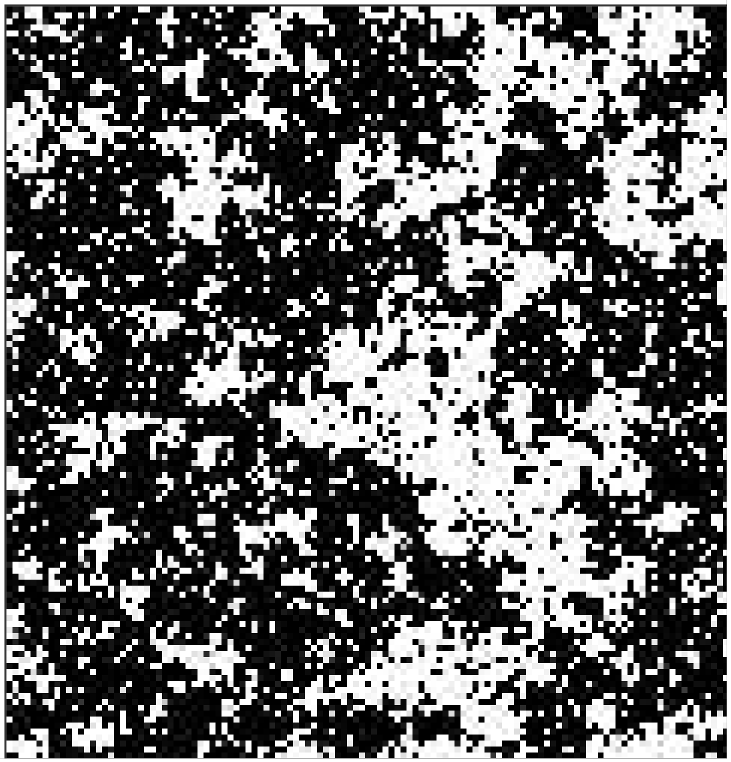
Energy

$$H(\underline{x}) = -\sum_{|i-j|=1} J_{ij} x_i x_j$$

Gibbs Distribution

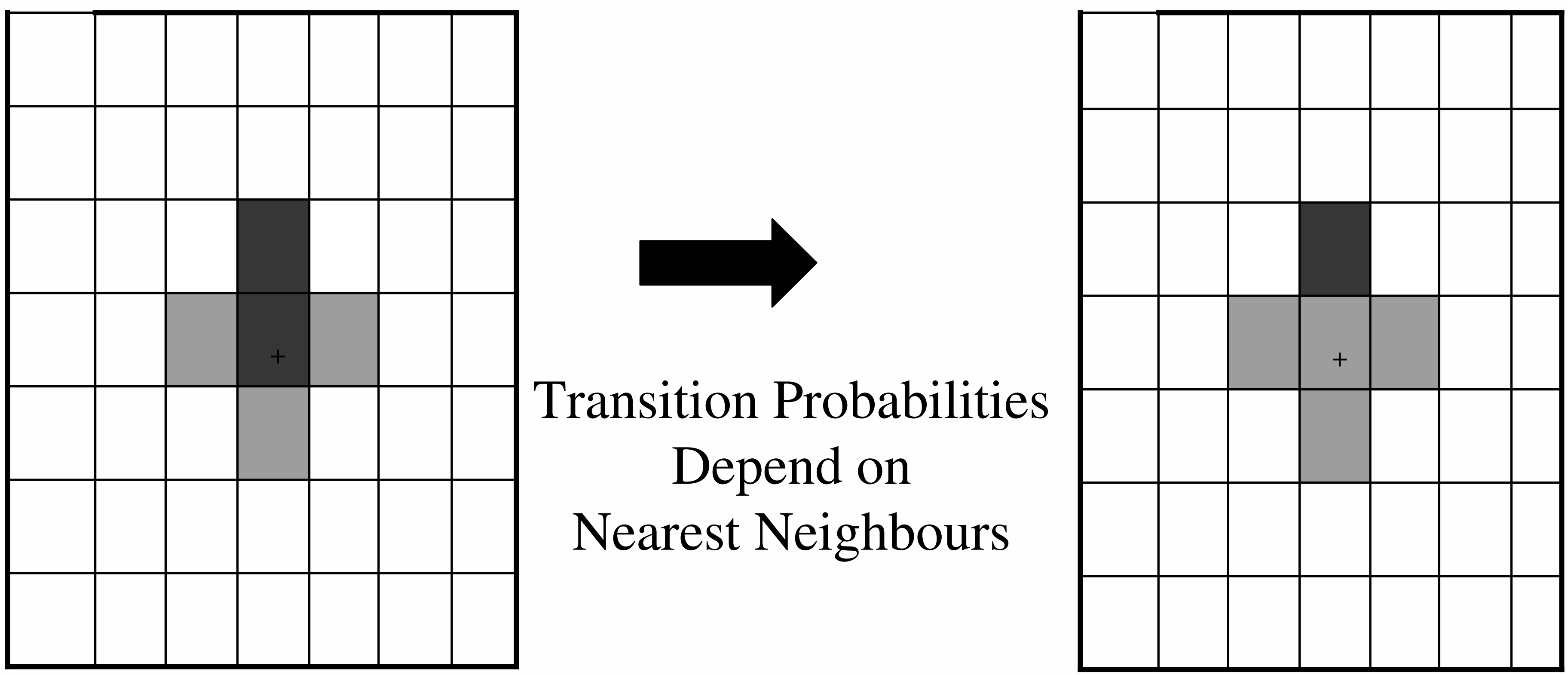
$$P(\underline{x}) = \frac{1}{Z} e^{-\beta H(\underline{x})}$$

$\beta =$ **inverse temperature**



Spitzer-Dobrushin

- Markov Random Field (conditional independence)
- Reversible equilibrium, detailed balance.
- Glauber Markovian dynamics: spin-flip



From Statistical Physics to Statistics:

- **Phase Transitions, Critical temperature.**
- **Critical fluctuations $d < 4$ vs. $d > 4$.**
- **Spin Glass Models, disordered medium: random $J_{i,j}$, frustration.**
- **Markov chain Monte Carlo in statistics and optimization**
  - **Find a Markov chain whose stationary distribution coincides with the desired distribution**
    - **the Gibbs sampler inspired by Glauber dynamics**
  - **MCMC has many important applications, for example, in Bayesian networks - used e.g. in machine learning.**