A Thinning Analogue of de Finetti's Theorem

Shannon Starr

University of Rochester

9 Jan 2007

Finite Exchangeability

$\mu_n \in \mathbf{M}_1(\Omega^n)$

$(X_1, \ldots, X_n)$ random, μ_n-distributed

$\mu_n \cdot \pi :=$ distribution of $(X_{\pi(1)}, \ldots, X_{\pi(n)})$

Finite Exchangeability

$\mu_n \in \mathbf{M}_1(\Omega^n)$

$(\mathsf{X}_1, \dots, \mathsf{X}_n)$ random, μ_n-distributed

$\mu_n \cdot \pi :=$ distribution of $(\mathsf{X}_{\pi(1)}, \dots, \mathsf{X}_{\pi(n)})$

μ_n is symmetric if $\mu_n \cdot \pi = \mu_n$ for all $\pi \in S_n$.

Let us call this "finite exchangeability".

Infinite Exchangeability

Assume $\Omega = \mathcal{X}$ a compact metric space

$\mathcal{X}^\infty := (x_1, x_2, \dots)$, all x_i in $\mathcal{X}$

S_∞ := bijections $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ s.t. $\#\{i : \pi(i) \neq i\} < \infty$

Infinite Exchangeability

Assume $\Omega = \mathcal{X}$ a compact metric space

$\mathcal{X}^\infty := (x_1, x_2, \dots)$, all x_i in $\mathcal{X}$

S_∞ := bijections $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ s.t. $\#\{i \, : \, \pi(i) \neq i\} < \infty$

$\mu_\infty \in \mathbf{M}_1(\mathcal{X}^\infty)$

$(\mathsf{X}_1, \mathsf{X}_2, \dots)$ random, μ_∞-distributed

$\mu_\infty \cdot \pi$:= distribution of $(\mathsf{X}_{\pi(1)}, \mathsf{X}_{\pi(2)}, \dots)$.

Infinite Exchangeability

Assume $\Omega = \mathcal{X}$ a compact metric space

$\mathcal{X}^\infty := (x_1, x_2, \dots)$, all x_i in $\mathcal{X}$

S_∞ := bijections $\pi : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ s.t. $\#\{i \, : \, \pi(i) \neq i\} < \infty$

$\mu_\infty \in \mathbf{M}_1(\mathcal{X}^\infty)$

$(\mathrm{X}_1, \mathrm{X}_2, \dots)$ random, μ_∞-distributed

$\mu_\infty \cdot \pi$:= distribution of $(\mathrm{X}_{\pi(1)}, \mathrm{X}_{\pi(2)}, \dots)$.

μ_∞ is infinitely exchangeable if $\mu_\infty \cdot \pi = \mu_\infty$ for all $\pi \in S_\infty$.

Specific Example

$\mathcal{X} = \{0, 1\}$

Bernoulli distribution: $\beta_p = p \cdot \delta_1 + (1 - p)\delta_0$

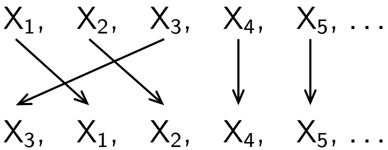
independent product measure: $\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots$

Specific Example

$\mathcal{X} = \{0,1\}$

Bernoulli distribution: $\beta_p = p \cdot \delta_1 + (1-p)\delta_0$

independent product measure: $\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots$



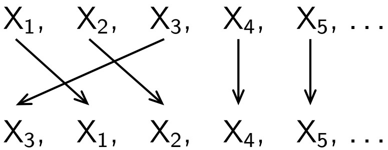
$(\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots) \cdot \pi = \beta_{p_{\pi(1)}} \otimes \beta_{p_{\pi(2)}} \otimes \cdots$

Specific Example

$\mathcal{X} = \{0, 1\}$

Bernoulli distribution: $\beta_p = p \cdot \delta_1 + (1 - p)\delta_0$

independent product measure: $\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots$



$$(\beta_{p_1} \otimes \beta_{p_2} \otimes \cdots) \cdot \pi = \beta_{p_{\pi(1)}} \otimes \beta_{p_{\pi(2)}} \otimes \cdots$$

So $\beta_p^{\otimes \infty}$ is exchangeable for all $p \in [0, 1]$.

More general example

Any mixture of i.i.d. product measure is exchangeable.

More general example

Any mixture of i.i.d. product measure is exchangeable.

Still consider $\mathcal{X} = \{0, 1\}$.

Suppose $P \in \mathbf{M}_1([0, 1])$, mixing measure.

Define

$$\mu_\infty(\cdot) = \int_0^1 P(dp)\, \beta_p^{\otimes\infty}(\cdot)$$

More general example

Any mixture of i.i.d. product measure is exchangeable.

Still consider $\mathcal{X} = \{0, 1\}$.

Suppose $P \in \mathbf{M}_1([0, 1])$, mixing measure.

Define

$$\mu_\infty(\cdot) = \int_0^1 P(dp)\, \beta_p^{\otimes\infty}(\cdot)$$

De Finetti's theorem says every infinitely exchangeable measure is a mixture of i.i.d. product measures.

De Finetti's Theorem

Theorem

Suppose $\mu_\infty \in \mathbf{M}_1(\mathcal{X}^\infty)$ *is exchangeable. Then there is a unique probability measure* $P \in \mathbf{M}_1(\mathbf{M}_1(\mathcal{X}))$ *such that*

$$\mu_\infty(\cdot) = \int_{\mathbf{M}_1(\mathcal{X})} P(d\nu)\, \nu^{\otimes\infty}(\cdot)$$

De Finetti's Theorem

Theorem

Suppose $\mu_\infty \in \mathbf{M}_1(\mathcal{X}^\infty)$ *is exchangeable. Then there is a unique probability measure* $P \in \mathbf{M}_1(\mathbf{M}_1(\mathcal{X}))$ *such that*

$$\mu_\infty(\cdot) = \int_{\mathbf{M}_1(\mathcal{X})} P(d\nu)\, \nu^{\otimes\infty}(\cdot)$$

Note: This also implies the extreme points are i.i.d. product measures.

Thinning

The right action generalizes to maps other than permutations.

$\phi : [m] \to [n]$ and $\mu_n \in \mathbf{M}_1(\mathcal{X}^n) \rightsquigarrow \mu_n \cdot \phi \in \mathbf{M}_1(\mathcal{X}^m)$:

Thinning

The right action generalizes to maps other than permutations.

$\phi : [m] \to [n]$ and $\mu_n \in \mathbf{M}_1(\mathcal{X}^n) \rightsquigarrow \mu_n \cdot \phi \in \mathbf{M}_1(\mathcal{X}^m)$:

Let $(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ be μ_n-distributed.

$\mu_n \cdot \phi =$ distribution of $(\mathrm{X}_{\phi(1)}, \ldots, \mathrm{X}_{\phi(m)})$

Thinning

The right action generalizes to maps other than permutations.

$\phi : [m] \to [n]$ and $\mu_n \in \mathbf{M}_1(\mathcal{X}^n) \rightsquigarrow \mu_n \cdot \phi \in \mathbf{M}_1(\mathcal{X}^m)$:

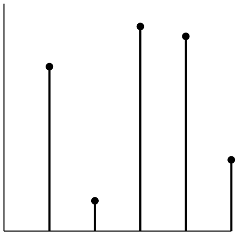
Let $(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ be μ_n-distributed.

$\mu_n \cdot \phi \; = \;$ distribution of $(\mathrm{X}_{\phi(1)}, \ldots, \mathrm{X}_{\phi(m)})$

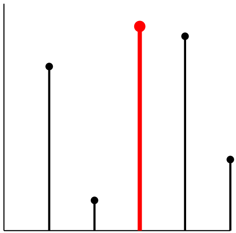
"Thinning" $=$ linear transformation $\Theta^n_{n-1} : \mathbf{M}_1(\mathcal{X}^n) \to \mathbf{M}_1(\mathcal{X}^{n-1})$,

$\phi_{n,k} : [n-1] \to [n]$:
$(\phi_{n,k}(1), \ldots, \phi_{n,k}(n-1)) = (1, \ldots, k-1, k+1, \ldots, n)$.

$$\mu_n \Theta^n_{n-1} \; = \; \frac{1}{n} \sum_{k=1}^{n} \mu_n \cdot \phi_{n,k} \, ,$$

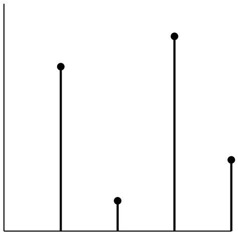


$(X_1, \ldots, X_n)$ random, μ_n-distributed.



$(X_1, \ldots, X_n)$ random, μ_n-distributed.

$k \in [n]$ uniform and independent of $(X_1, \ldots, X_n)$.



$(X_1, \dots, X_n)$ random, μ_n-distributed.

$k \in [n]$ uniform and independent of $(X_1, \dots, X_n)$.

$\mu_n \Theta^n_{n-1}$ is distribution of $(X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n)$.

Thinning-invariance

Sequence $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ is "thinning-invariant" if $\mu_n \Theta_{n-1}^n = \mu_{n-1}$ all $n > 1$

Thinning-invariance

Sequence $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ is "thinning-invariant" if $\mu_n \Theta^n_{n-1} = \mu_{n-1}$ all $n > 1$

Call $\boldsymbol{\mu}$ "exchangeable" if:
(1) each $\mu_n \in \mathbf{M}_1(\mathcal{X}^n)$ is finitely-exchangeable;
(2) $\boldsymbol{\mu}$ is thinning-invariant.

Thinning-invariance

Sequence $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ is "thinning-invariant" if $\mu_n \Theta^n_{n-1} = \mu_{n-1}$ all $n > 1$

Call $\boldsymbol{\mu}$ "exchangeable" if:
(1) each $\mu_n \in \mathbf{M}_1(\mathcal{X}^n)$ is finitely-exchangeable;
(2) $\boldsymbol{\mu}$ is thinning-invariant.

Because of finite exchangeability, $\mu_n \Theta^n_{n-1}$ is equal-in-distribution to the marginal distribution of $(\mathsf{X}_1, \dots, \mathsf{X}_{n-1})$, where $(\mathsf{X}_1, \dots, \mathsf{X}_n)$ is μ_n-distributed.

Thinning-invariance

Sequence $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ is "thinning-invariant" if $\mu_n \Theta^n_{n-1} = \mu_{n-1}$ all $n > 1$

Call $\boldsymbol{\mu}$ "exchangeable" if:
(1) each $\mu_n \in \mathbf{M}_1(\mathcal{X}^n)$ is finitely-exchangeable;
(2) $\boldsymbol{\mu}$ is thinning-invariant.

Because of finite exchangeability, $\mu_n \Theta^n_{n-1}$ is equal-in-distribution to the marginal distribution of $(\mathsf{X}_1, \dots, \mathsf{X}_{n-1})$, where $(\mathsf{X}_1, \dots, \mathsf{X}_n)$ is μ_n-distributed.

Kolmogorov's extension theorem $\Rightarrow \exists$ infinitely exchangeable $\mu_\infty \in \mathbf{M}_1(\mathcal{X}^\infty)$, such that μ_n is the marginal distribution of $(\mathsf{X}_1, \dots, \mathsf{X}_n)$, where $(\mathsf{X}_1, \mathsf{X}_2, \dots)$ is μ_∞-distributed.

Another version of de Finetti's Theorem

Theorem

Suppose $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ *is "exchangeable". Then there is a unique probability measure* $P \in \mathbf{M}_1(\mathbf{M}_1(\mathcal{X}))$ *such that, for each* $n > 0$,

$$\mu_n(\cdot) = \int_{\mathbf{M}_1(\mathcal{X})} P(d\nu)\, \nu^{\otimes n}(\cdot)$$

Main Question

There is a representation theorem for sequences $\boldsymbol{\mu}$ which are thinning-invariant, and such that each μ_n is finitely exchangeable.

Is there a representation theorem for sequences which are thinning-invariant but not necessarily finitely-exchangeable?

An example of thinning-invariance : Order Statistics

Define $\mathcal{I} := [0, 1]$.

$\lambda :=$ Borel version of standard Lebesgue measure on $\mathcal{I}$

Let $\mathcal{X} = \mathcal{I}$.

An example of thinning-invariance : Order Statistics

Define $\mathcal{I} := [0, 1]$.

$\lambda :=$ Borel version of standard Lebesgue measure on $\mathcal{I}$

Let $\mathcal{X} = \mathcal{I}$.

Let $\mathsf{T}_1, \dots, \mathsf{T}_n \in \mathcal{I}$ be i.i.d. λ-distributed r.v.'s.

Almost surely, there is a unique $\hat{\pi} \in S_n$ such that

$$\mathsf{T}_{\hat{\pi}(1)} < \dots < \mathsf{T}_{\hat{\pi}(n)} .$$

An example of thinning-invariance : Order Statistics

Define $\mathcal{I} := [0, 1]$.

$\lambda :=$ Borel version of standard Lebesgue measure on $\mathcal{I}$

Let $\mathcal{X} = \mathcal{I}$.

Let $\mathsf{T}_1, \dots, \mathsf{T}_n \in \mathcal{I}$ be i.i.d. λ-distributed r.v.'s.

Almost surely, there is a unique $\hat{\pi} \in S_n$ such that

$$\mathsf{T}_{\hat{\pi}(1)} < \dots < \mathsf{T}_{\hat{\pi}(n)} \,.$$

Let $\mu_n \in \mathbf{M}_1(\mathcal{I}^n)$ be the distribution of $(\mathsf{T}_{\hat{\pi}(1)}, \dots, \mathsf{T}_{\hat{\pi}(n)})$.

An example of thinning-invariance : Order Statistics

Define $\mathcal{I} := [0, 1]$.

$\lambda :=$ Borel version of standard Lebesgue measure on $\mathcal{I}$

Let $\mathcal{X} = \mathcal{I}$.

Let $\mathsf{T}_1, \dots, \mathsf{T}_n \in \mathcal{I}$ be i.i.d. λ-distributed r.v.'s.

Almost surely, there is a unique $\hat{\pi} \in S_n$ such that

$$\mathsf{T}_{\hat{\pi}(1)} < \dots < \mathsf{T}_{\hat{\pi}(n)} .$$

Let $\mu_n \in \mathbf{M}_1(\mathcal{I}^n)$ be the distribution of $(\mathsf{T}_{\hat{\pi}(1)}, \dots, \mathsf{T}_{\hat{\pi}(n)})$.

Then $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ is thinning-invariant, but not exchangeable.

A more general example

$\mathcal{I} = [0, 1]$ and $\lambda =$ Lebesgue measure.

A more general example

$\mathcal{I} = [0,1]$ and $\lambda =$ Lebesgue measure.

$\mathbf{M}_1^{\lambda}(\mathcal{X} \times \mathcal{I}) :=$ all $\alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I})$ such that marginal on $\mathcal{I}$ is λ

A more general example

$\mathcal{I} = [0, 1]$ and $\lambda =$ Lebesgue measure.

$\mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I}) :=$ all $\alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I})$ such that marginal on $\mathcal{I}$ is λ

Let $(\mathsf{X}_1, \mathsf{T}_1), \dots, (\mathsf{X}_n, \mathsf{T}_n) \in \mathcal{X} \times \mathcal{I}$ be i.i.d., α-distributed.

A more general example

$\mathcal{I} = [0,1]$ and $\lambda =$ Lebesgue measure.

$\mathbf{M}_1^{\lambda}(\mathcal{X} \times \mathcal{I}) :=$ all $\alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I})$ such that marginal on $\mathcal{I}$ is λ

Let $(\mathsf{X}_1, \mathsf{T}_1), \dots, (\mathsf{X}_n, \mathsf{T}_n) \in \mathcal{X} \times \mathcal{I}$ be i.i.d., α-distributed.

Almost surely, $\exists\ !\ \hat{\pi} \in S_n$ s.t. $\mathsf{T}_{\hat{\pi}(1)} < \dots < \mathsf{T}_{\hat{\pi}(n)}$

A more general example

$\mathcal{I} = [0,1]$ and $\lambda =$ Lebesgue measure.

$\mathbf{M}_1^{\lambda}(\mathcal{X} \times \mathcal{I}) :=$ all $\alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I})$ such that marginal on $\mathcal{I}$ is λ

Let $(\mathsf{X}_1, \mathsf{T}_1), \dots, (\mathsf{X}_n, \mathsf{T}_n) \in \mathcal{X} \times \mathcal{I}$ be i.i.d., α-distributed.

Almost surely, $\exists\ !\ \hat{\pi} \in S_n$ s.t. $\mathsf{T}_{\hat{\pi}(1)} < \dots < \mathsf{T}_{\hat{\pi}(n)}$

Define $\mathcal{M}_n(\alpha) \in \mathbf{M}_1(\mathcal{X}^n)$ distribution of $(\mathsf{X}_{\hat{\pi}(1)}, \dots, \mathsf{X}_{\hat{\pi}(n)})$

A more general example

$\mathcal{I} = [0,1]$ and $\lambda =$ Lebesgue measure.

$\mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I}) :=$ all $\alpha \in \mathbf{M}_1(\mathcal{X} \times \mathcal{I})$ such that marginal on $\mathcal{I}$ is λ

Let $(\mathsf{X}_1, \mathsf{T}_1), \dots, (\mathsf{X}_n, \mathsf{T}_n) \in \mathcal{X} \times \mathcal{I}$ be i.i.d., α-distributed.

Almost surely, $\exists\ !\ \hat{\pi} \in S_n$ s.t. $\mathsf{T}_{\hat{\pi}(1)} < \dots < \mathsf{T}_{\hat{\pi}(n)}$

Define $\mathcal{M}_n(\alpha) \in \mathbf{M}_1(\mathcal{X}^n)$ distribution of $(\mathsf{X}_{\hat{\pi}(1)}, \dots, \mathsf{X}_{\hat{\pi}(n)})$

Define $\boldsymbol{\mathcal{M}}(\alpha) = (\mathcal{M}_1(\alpha), \mathcal{M}_2(\alpha), \dots)$.

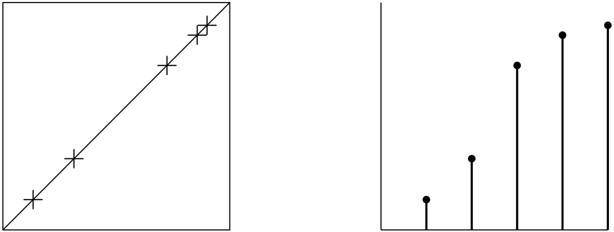
$\boldsymbol{\mathcal{M}}(\alpha)$ thinning-invariant for every $\alpha \in \mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I})$

E.g., recovering order statistics

If $\mathcal{X} = \mathcal{I}$ and α is uniform measure on diagonal $\{(t,t)\,:\,t \in \mathcal{I}\}$,

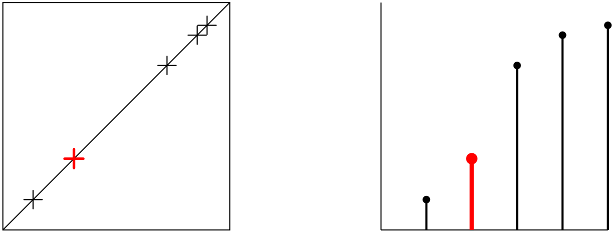
E.g., recovering order statistics

If $\mathcal{X} = \mathcal{I}$ and α is uniform measure on diagonal $\{(t,t) \, : \, t \in \mathcal{I}\}$,



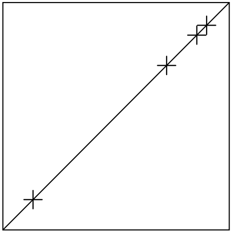
E.g., recovering order statistics

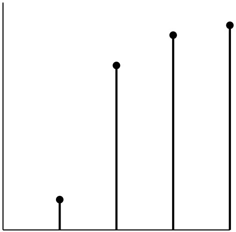
If $\mathcal{X} = \mathcal{I}$ and α is uniform measure on diagonal $\{(t,t) : t \in \mathcal{I}\}$,



E.g., recovering order statistics

If $\mathcal{X} = \mathcal{I}$ and α is uniform measure on diagonal $\{(t,t) : t \in \mathcal{I}\}$,





Main Result

Theorem

Suppose $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ *is thinning-invariant. Then there is a unique probability measure* $P \in \mathbf{M}_1(\mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I}))$ *such that, for each* $n > 0$,

$$\mu_n(\cdot) = \int_{\mathbf{M}_1^\lambda(\mathcal{X}\times\mathcal{I})} P(d\alpha)\, \mathcal{M}_n(\alpha)(\cdot)$$

Proof deferred ...

Mean-Field Models

Let's start by (re)considering mean-field models

E.g., Curie-Weiss model. $\mathcal{X} = \{+1, -1\}$, $H_N : \mathcal{X}^N \to \mathbb{R}$,

$$\frac{H_N(\sigma_1, \dots, \sigma_n)}{N} = -\frac{J}{2} \cdot \frac{\sum_{1 \leq j < k \leq N} \sigma_j \sigma_k}{\binom{N}{2}} - h \cdot \frac{\sum_{k=1}^{N} \sigma_k}{N}$$

for some fixed $J, h \in \mathbb{R}$.

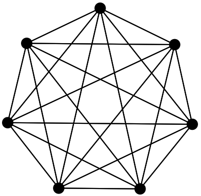
Mean-Field Models

Let's start by (re)considering mean-field models

E.g., Curie-Weiss model. $\mathcal{X} = \{+1, -1\}$, $H_N : \mathcal{X}^N \to \mathbb{R}$,

$$\frac{H_N(\sigma_1, \dots, \sigma_n)}{N} = -\frac{J}{2} \cdot \frac{\sum_{1 \leq j < k \leq N} \sigma_j \sigma_k}{\binom{N}{2}} - h \cdot \frac{\sum_{k=1}^N \sigma_k}{N}$$

for some fixed $J, h \in \mathbb{R}$.
Ising model on a complete graph K_N



Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathscr{C}(\mathcal{X}^N) \to \mathscr{C}(\mathcal{X}^N)$,

$$\big(\pi \cdot f_N\big)(\sigma_1, \dots, \sigma_N) \ := \ f_N(\sigma_{\pi(1)}, \dots, \sigma_{\pi(N)})\,.$$

Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathscr{C}(\mathcal{X}^N) \to \mathscr{C}(\mathcal{X}^N)$,

$$\big(\pi \cdot f_N\big)(\sigma_1, \ldots, \sigma_N) \ := \ f_N(\sigma_{\pi(1)}, \ldots, \sigma_{\pi(N)})\,.$$

Define $\Theta_{N-1}^N : \mathscr{C}(\mathcal{X}^{N-1}) \to \mathscr{C}(\mathcal{X}^N)$ by

$$\big(\Theta_{N-1}^N f_{N-1}\big)(\sigma_1, \ldots, \sigma_N) \ := \ \frac{1}{N} \sum_{k=1}^N f_{N-1}(\sigma_{\phi_{N,k}(1)}, \ldots, \sigma_{\phi_{N,k}(N-1)})$$

Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathscr{C}(\mathcal{X}^N) \to \mathscr{C}(\mathcal{X}^N)$,

$$(\pi \cdot f_N)(\sigma_1, \dots, \sigma_N) := f_N(\sigma_{\pi(1)}, \dots, \sigma_{\pi(N)}) .$$

Define $\Theta_{N-1}^N : \mathscr{C}(\mathcal{X}^{N-1}) \to \mathscr{C}(\mathcal{X}^N)$ by

$$(\Theta_{N-1}^N f_{N-1})(\sigma_1, \dots, \sigma_N) := \frac{1}{N} \sum_{k=1}^N f_{N-1}(\sigma_{\phi_{N,k}(1)}, \dots, \sigma_{\phi_{N,k}(N-1)})$$

Note: $\langle \mu_N \cdot \pi, f_N \rangle = \langle \mu_N, \pi \cdot f_N \rangle$,

$\langle \mu_N \Theta_{N-1}^N, f_{N-1} \rangle = \langle \mu_N, \Theta_{N-1}^N f_{N-1} \rangle$

where $\langle \mu_N, f_N \rangle := \mathbf{E}^{\mu_N}[f_N]$.

Symmetry properties

For $\pi \in S_N$ define left-action $\pi : \mathscr{C}(\mathcal{X}^N) \to \mathscr{C}(\mathcal{X}^N)$,

$$(\pi \cdot f_N)(\sigma_1, \dots, \sigma_N) := f_N(\sigma_{\pi(1)}, \dots, \sigma_{\pi(N)}) .$$

Define $\Theta_{N-1}^N : \mathscr{C}(\mathcal{X}^{N-1}) \to \mathscr{C}(\mathcal{X}^N)$ by

$$(\Theta_{N-1}^N f_{N-1})(\sigma_1, \dots, \sigma_N) := \frac{1}{N} \sum_{k=1}^N f_{N-1}(\sigma_{\phi_{N,k}(1)}, \dots, \sigma_{\phi_{N,k}(N-1)})$$

Note: $\langle \mu_N \cdot \pi, f_N \rangle = \langle \mu_N, \pi \cdot f_N \rangle$,
$\langle \mu_N \Theta_{N-1}^N, f_{N-1} \rangle = \langle \mu_N, \Theta_{N-1}^N f_{N-1} \rangle$
where $\langle \mu_N, f_N \rangle := \mathbf{E}^{\mu_N}[f_N]$.

$\forall \pi \in S_N$, $\pi \cdot H_N = H_N$. Finite-exchangeability.

$\forall N > 2$, $\Theta_{N-1}^N H_{N-1} = [(N-1)/N] H_N$. Thinning-invariance.

Thermodynamic Quantities

Let $\nu_0 \in \mathbf{M}_1(\mathcal{X})$ be *a-priori* measure, e.g. uniform.

Partition function

$$Z_N(\beta) = \int_{\mathcal{X}^N} e^{-\beta H_N} \, d\nu_0^{\otimes N}$$

Free energy

$$F_N(\beta) = -\frac{1}{\beta} \log(Z_N(\beta))$$

"Pressure"

$$p_N(\beta) = \frac{1}{N} \log(Z_N(\beta)) \,.$$

Boltzmann-Gibbs measure $\mu^*_{N,\beta} \in \mathbf{M}_1(\mathcal{X}^N)$ s.t.

$$\frac{d\mu^*_{N,\beta}}{d\nu_0^{\otimes N}} = Z_N(\beta)^{-1} e^{-\beta H_N} \,.$$

Gibbs Variational Principle

For $\mu_N \in \mathbf{M}_1(\mathcal{X}^N)$, define

$$\mathcal{G}_N(\mu_N;\beta) := \frac{1}{N}\left(S_N(\mu_N|\nu_0^{\otimes N}) - \beta \mathbf{E}^{\mu_N}[H_N]\right)$$

where $S_N(\mu_N|\nu_N)$ is relative entropy

$$S_N(\mu_N|\nu_N) = -\int_{\mathcal{X}^N} \log\left(\frac{d\mu_N}{d\nu_N}\right)\, d\mu_N$$

and $\mathbf{E}^{\mu_N}[\cdot]$ is expectation.

Gibbs Variational Principle

For $\mu_N \in \mathbf{M}_1(\mathcal{X}^N)$, define

$$\mathcal{G}_N(\mu_N;\beta) := \frac{1}{N}\left(S_N(\mu_N|\nu_0^{\otimes N}) - \beta \mathbf{E}^{\mu_N}[H_N]\right)$$

where $S_N(\mu_N|\nu_N)$ is relative entropy

$$S_N(\mu_N|\nu_N) = -\int_{\mathcal{X}^N} \log\left(\frac{d\mu_N}{d\nu_N}\right)\, d\mu_N$$

and $\mathbf{E}^{\mu_N}[\cdot]$ is expectation.

$$p_N(\beta) = \max_{\mu_N} \mathcal{G}_N(\mu_N;\beta)$$

and the unique argmax is Boltzmann-Gibbs distribution $\mu^*_{N,\beta}$

Consequences of symmetry

$$p_N(\beta) = \max_{\mu_N} \mathcal{G}_N(\mu_N;\beta) = \mathcal{G}_N(\mu^*_{N,\beta};\beta)\,,$$

$$\frac{d\mu^*_{N,\beta}}{d\nu_0^{\otimes N}}(\cdot) = Z_N(\beta)^{-1}\, e^{-\beta H_N(\cdot)}$$

Consequences of symmetry

$$p_N(\beta) = \max_{\mu_N} \mathcal{G}_N(\mu_N; \beta) = \mathcal{G}_N(\mu^*_{N,\beta}; \beta)\,,$$

$$\frac{d\mu^*_{N,\beta}}{d\nu_0^{\otimes N}}(\cdot) = Z_N(\beta)^{-1}\, e^{-\beta H_N(\cdot)}$$

Since $\pi \cdot H_N = H_N$, know $\mu^*_{N,\beta} = \mu^*_{N,\beta} \cdot \pi$.

So, defining $\mathbf{M}_1^{\mathrm{ex}}(\mathcal{X}^N)$ the finite exchangeable measures,

$$p_N(\beta) = \max_{\mu_N \in \mathbf{M}_1^{\mathrm{ex}}(\mathcal{X}^N)} \mathcal{G}_N(\mu_N; \beta)\,.$$

$\mathcal{G}_N$ for product measures $\subset \mathbf{M}_1^{\mathrm{ex}}(\mathcal{X}^N)$

If we further restrict to i.i.d. product measures, we get a simpler formula.

$\mathcal{G}_N$ for product measures $\subset \mathbf{M}_1^{\text{ex}}(\mathcal{X}^N)$

If we further restrict to i.i.d. product measures, we get a simpler formula.

Suppose $\nu \in \mathbf{M}_1(\mathcal{X})$ and $\mu_n = \nu^{\otimes N}$. Then

$$S_N(\mu_N|\nu_0^{\otimes N}) \;=\; S_N(\nu^{\otimes N}|\nu_0^{\otimes N}) \;=\; N \cdot S_1(\nu|\nu_0)\,.$$

$\mathcal{G}_N$ for product measures $\subset \mathbf{M}_1^{\text{ex}}(\mathcal{X}^N)$

If we further restrict to i.i.d. product measures, we get a simpler formula.

Suppose $\nu \in \mathbf{M}_1(\mathcal{X})$ and $\mu_n = \nu^{\otimes N}$. Then

$$S_N(\mu_N|\nu_0^{\otimes N}) \;=\; S_N(\nu^{\otimes N}|\nu_0^{\otimes N}) \;=\; N \cdot S_1(\nu|\nu_0)\,.$$

Also,

$$\mathbf{E}^{\nu^{\otimes N}}[H_N] \;=\; N \cdot \varphi(\nu)\,,$$

where

$$\varphi(\nu) \;=\; -\frac{J}{2}\big(\mathbf{E}^{\nu}[\sigma]\big)^2 - h\mathbf{E}^{\nu}[\sigma]\,.$$

$\mathcal{G}_N$ for product measures $\subset \mathbf{M}_1^{\mathrm{ex}}(\mathcal{X}^N)$

If we further restrict to i.i.d. product measures, we get a simpler formula.

Suppose $\nu \in \mathbf{M}_1(\mathcal{X})$ and $\mu_n = \nu^{\otimes N}$. Then

$$S_N(\mu_N|\nu_0^{\otimes N}) \;=\; S_N(\nu^{\otimes N}|\nu_0^{\otimes N}) \;=\; N \cdot S_1(\nu|\nu_0)\,.$$

Also,

$$\mathbf{E}^{\nu^{\otimes N}}[H_N] \;=\; N \cdot \varphi(\nu)\,,$$

where

$$\varphi(\nu) \;=\; -\frac{J}{2}\big(\mathbf{E}^{\nu}[\sigma]\big)^2 - h\mathbf{E}^{\nu}[\sigma]\,.$$

So $\mathcal{G}_N(\nu^{\otimes N};\beta) \;=\; S_1(\nu|\nu_0) - \beta\varphi(\nu)$.

Thermodynamic Limit : One method of solution
Fannes, Spohn and Verbeure, *J. Math. Phys.* (1980)

"Gap equation"

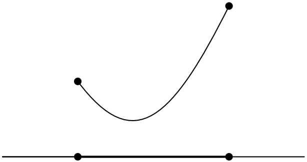
$$p(\beta) := \lim_{N\to\infty} p_N(\beta) = \max_{\nu\in\mathbf{M}_1(\mathcal{X})} \left(S_1(\nu|\nu_0) - \varphi(\nu)\right).$$

Thermodynamic Limit : One method of solution
Fannes, Spohn and Verbeure, *J. Math. Phys.* (1980)

"Gap equation"

$$p(\beta) := \lim_{N\to\infty} p_N(\beta) = \max_{\nu\in\mathbf{M}_1(\mathcal{X})} \left(S_1(\nu|\nu_0) - \varphi(\nu)\right).$$

Heuristic idea: Given a convex function on a simplex, it attains its maximum value at an extreme point.



Thermodynamic Limit : One method of solution Fannes, Spohn and Verbeure, *J. Math. Phys.* (1980)

"Gap equation"

$$p(\beta) := \lim_{N\to\infty} p_N(\beta) = \max_{\nu\in\mathbf{M}_1(\mathcal{X})} \big(S_1(\nu|\nu_0) - \varphi(\nu)\big)\,.$$

Heuristic idea: Given a convex function on a simplex, it attains its maximum value at an extreme point.

For any finite N, $\mathcal{G}_N(\cdot;\beta)$ is concave not convex, so cannot restrict to extreme points.

Thermodynamic Limit : One method of solution Fannes, Spohn and Verbeure, *J. Math. Phys.* (1980)

"Gap equation"

$$p(\beta) := \lim_{N\to\infty} p_N(\beta) = \max_{\nu\in\mathbf{M}_1(\mathcal{X})} \left(S_1(\nu|\nu_0) - \varphi(\nu)\right).$$

Heuristic idea: Given a convex function on a simplex, it attains its maximum value at an extreme point.

For any finite N, $\mathcal{G}_N(\cdot;\beta)$ is concave not convex, so cannot restrict to extreme points.

But the nonlinear part, entropy, is "almost convex".

$N \to \infty$ limit of $\mathcal{G}_N(\cdot;\beta)$ is linear/affine, when restricted to infinitely exchangeable measures

and extreme points are product measures.

Asymmetric Mean-Field Models

Consider a sequence of Hamiltonians $H_N : \mathcal{X}^N \to \mathbb{R}$, for $N \geq n$, s.t.

$$\frac{1}{N-1}\Theta^N_{N-1} H_{N-1} = \frac{1}{N} H_N$$

for all $N > n$. But do not assume finite exchangeability.

Asymmetric Mean-Field Models

Consider a sequence of Hamiltonians $H_N : \mathcal{X}^N \to \mathbb{R}$, for $N \geq n$, s.t.

$$\frac{1}{N-1}\Theta^N_{N-1} H_{N-1} = \frac{1}{N} H_N$$

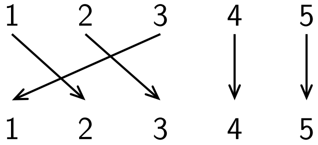
for all $N > n$. But do not assume finite exchangeability.

Similar result to Fannes,Spohn and Verbeure's holds, adapted to simplex of thinning-invariant measures

$$p(\beta) = \max_{\alpha \in \mathbf{M}_1^{\lambda}(\mathcal{X} \times \mathcal{I})} \left(S(\alpha | \nu_0 \otimes \lambda) - \beta \mathbf{E}^{\mathcal{M}_n(\alpha)}[H_n/n]\right)$$

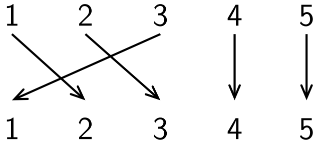
Example : Mallows Model

An inversion of $\pi \in S_n$ is a pair $i < j$ s.t. $\pi(i) > \pi(j)$.



Example : Mallows Model

An inversion of $\pi \in S_n$ is a pair $i < j$ s.t. $\pi(i) > \pi(j)$.



For $q > 0$ Mallows model of random permutations

$$\mathbf{P}_{n,q}(\{\pi\}) = Z_n(q)^{-1}\, q^{\#\mathrm{inv}(\pi)}\,,$$

$\pi \in S_n$.

Hamiltonian

$\mathcal{X} = \mathcal{I} = [0,1]$; $\nu_0 = \lambda$;

$$H_N(x_1,\dots,x_N) \;=\; N \cdot \frac{\sum_{1\leq i<j\leq N} 1_{x_i > x_j}}{\binom{N}{2}}\,.$$

Hamiltonian

$\mathcal{X} = \mathcal{I} = [0,1]$; $\nu_0 = \lambda$;

$$H_N(x_1,\dots,x_N) \;=\; N \cdot \frac{\sum_{1\leq i<j\leq N} 1_{x_i>x_j}}{\binom{N}{2}}\,.$$

Boltzmann-Gibbs measure $\mu^*_{N,\beta} \in \mathbf{M}_1(\mathcal{X}^N)$ concentrates on subset s.t. all $x_1,\dots,x_N$ distinct.

Hamiltonian

$\mathcal{X} = \mathcal{I} = [0, 1]$; $\nu_0 = \lambda$;

$$H_N(x_1, \dots, x_N) \;=\; N \cdot \frac{\sum_{1 \leq i < j \leq N} 1_{x_i > x_j}}{\binom{N}{2}}\,.$$

Boltzmann-Gibbs measure $\mu^*_{N,\beta} \in \mathbf{M}_1(\mathcal{X}^N)$ concentrates on subset s.t. all $x_1, \dots, x_N$ distinct.

Marginal on the ordering permutation $\hat{\pi}$ s.t.

$$x_{\hat{\pi}(1)} \;<\; \dots \;<\; x_{\hat{\pi}(N)}\,,$$

is $\mathbf{P}_{N,q}$ for $q = e^{-\beta/N}$. Weak asymmetry.

Solving the "Gap" equation

$\alpha \in \mathbf{M}_1^\lambda(\mathcal{X} \times \mathcal{I})$,

$$\alpha(dx \otimes dt) \;=\; f(t,x)\, dx\, dt\,,$$

satisfies, both marginals equal to λ and

$$\frac{\partial^2}{\partial t \partial x} \log\big(f(t,x)\big) \;=\; 2\beta f(t,x)\,.$$

Hyperbolic 2D Liouville's equation. Well-known PDE. Integrable. Arises when looking for 2D metric of constant negative curvature.

Solving the "Gap" equation

$\alpha \in \mathbf{M}_1^{\lambda}(\mathcal{X} \times \mathcal{I})$,

$$\alpha(dx \otimes dt) = f(t,x)\, dx\, dt\,,$$

satisfies, both marginals equal to λ and

$$\frac{\partial^2}{\partial t \partial x} \log\left(f(t,x)\right) = 2\beta f(t,x)\,.$$

Hyperbolic 2D Liouville's equation. Well-known PDE. Integrable. Arises when looking for 2D metric of constant negative curvature.

Unique solution with these boundary conditions.

Conclusion for Mallows model

Let $\hat{\pi} \in S_N$ be distributed by Mallows model $\mathbf{P}_{N,q}$ for $q = e^{-\beta/N}$.

Conclusion for Mallows model

Let $\hat{\pi} \in S_N$ be distributed by Mallows model $\mathbf{P}_{N,q}$ for $q = e^{-\beta/N}$.

Define random "empirical" measure

$$\hat{\alpha}_{N,\beta} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\left(\frac{i}{N}, \frac{\hat{\pi}(i)}{N}\right)} .$$

Conclusion for Mallows model

Let $\hat{\pi} \in S_N$ be distributed by Mallows model $\mathbf{P}_{N,q}$ for $q = e^{-\beta/N}$.

Define random "empirical" measure

$$\hat{\alpha}_{N,\beta} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\left(\frac{i}{N}, \frac{\hat{\pi}(i)}{N}\right)} .$$

Then in weak topology, $\hat{\alpha}_{N,\beta}(dx \otimes dt) \Rightarrow f_\beta(t,x)\, dx\, dt$, where:

$$f_\beta(t,x) = \frac{2\beta \sinh(\beta/2)}{\left[2e^{\beta/4} \cosh\left(\frac{\beta}{2}[x-t]\right) - 2e^{-\beta/4} \cosh\left(\frac{\beta}{2}[x+t-1]\right)\right]^2}$$