The theory of quantum statistical comparison

where do we stand

Francesco Buscemi (Nagoya University)

IQSA 2022 Conference, Tropea
27 June 2022

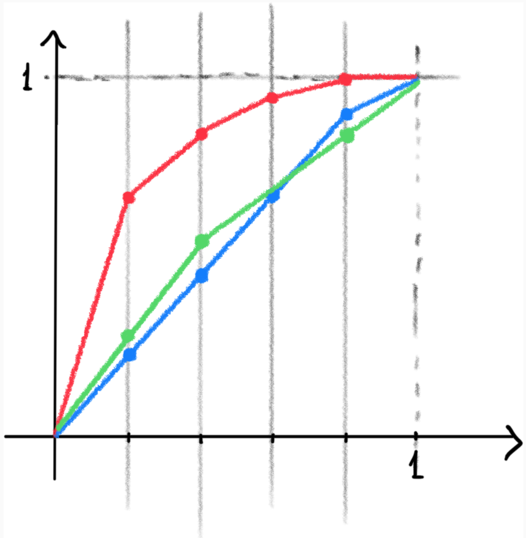
The precursor: majorization

Lorenz curves and majorization

- two probability distributions, $\boldsymbol{p} = (p_1, \ldots, p_n)$ and $\boldsymbol{q} = (q_1, \ldots, q_n)$
- truncated sums $P(k) = \sum_{i=1}^{k} p_i^{\downarrow}$ and $Q(k) = \sum_{i=1}^{k} q_i^{\downarrow}$, for all $k = 1, \ldots, n$
- $\boldsymbol{p}$ majorizes $\boldsymbol{q}$, i.e., $\boldsymbol{p} \succ \boldsymbol{q}$, whenever $P(k) \geqslant Q(k)$, for all k
- minimal element: uniform distribution $\boldsymbol{e} = n^{-1}(1, 1, \cdots, 1)$

Hardy–Littlewood–Pólya, 1934

$\boldsymbol{p} \succ \boldsymbol{q} \iff \boldsymbol{q} = M\boldsymbol{p}$, for some bistochastic matrix M.



$(x_k, y_k) = (k/n, P(k)), \quad 1 \leqslant k \leqslant n$

Blackwell's extension

Statistical experiments

Lucien Le Cam (1924-2000)

"The basic structures in the whole affair are systems that Blackwell called experiments, and transitions between them. An experiment is a mathematical abstraction intended to describe the basic feature of an observational process if that process is contemplated in advance of its implementation."

Lucien Le Cam (1984)

The formulation

Definition (Statistical models and decision problems)

$$\begin{tikzcd} \Omega \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{A} \arrow[d, squiggly] \\ \omega \arrow[r, "w(x|\omega)"'] & x \arrow[r] & a \end{tikzcd}$$

- *parameter* set $\Omega = \{\omega\}$, *sample* set $\mathcal{X} = \{x\}$, *action* set $\mathcal{A} = \{a\}$
- a **statistical model/experiment** is a triple $\mathbf{w} = \langle \Omega, \mathcal{X}, w(x|\omega) \rangle$
- a **statistical decision problem/game** is a triple $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$, where $c : \Omega \times \mathcal{A} \to \mathbb{R}$ is a payoff function

Playing statistical games with experiments

- the experiment/model is the resource: it is given
- the decision is the transition: it can be optimized

$$\begin{tikzcd} \Omega \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{A} \arrow[d, squiggly] \\ \omega \arrow[r, "w(x|\omega)"'] & x \arrow[r, "d(a|x)"'] & a \end{tikzcd}$$

Definition

The **(expected) maximin payoff** of a statistical model $\mathbf{w} = \langle \Omega, \mathcal{X}, w \rangle$ w.r.t. a decision problem $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$ is given by

$$c^*_{\mathbf{g}}(\mathbf{w}) \stackrel{\text{def}}{=} \max_{d(a|x)} \min_{\omega} \sum_{a,x} c(\omega, a) d(a|x) w(x|\omega) \; .$$

Comparison of statistical models

Definition (Information Preorder)

Given two statistical models $\mathbf{w} = \langle \Omega, \mathcal{X}, w \rangle$ and $\mathbf{w}' = \langle \Omega, \mathcal{Y}, w' \rangle$ on the same parameter set but possibly different sample sets, we say that $\mathbf{w}$ is (always) more informative than $\mathbf{w}'$, and write

$$\mathbf{w} \succ \mathbf{w}' \; ,$$

if and only if

$$c^*_{\mathbf{g}}(\mathbf{w}) \geqslant c^*_{\mathbf{g}}(\mathbf{w}') \; ,$$

for all decision problems $\mathbf{g} = \langle \Omega, \mathcal{A}, c \rangle$.

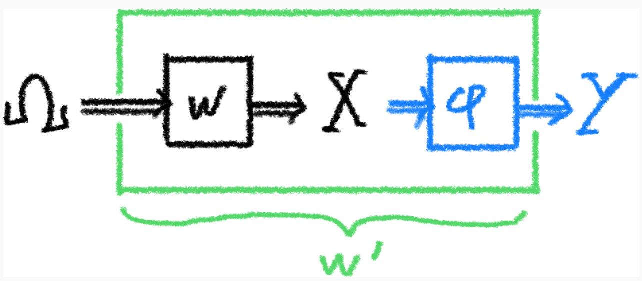
Can we visualize the information preorder more concretely?

Information preorder = statistical sufficiency

Theorem (Blackwell, 1953)

Given two statistical experiments $\mathbf{w} = \langle \Omega, \mathcal{X}, w \rangle$ *and* $\mathbf{w}' = \langle \Omega, \mathcal{Y}, w' \rangle$, *the following are equivalent:*

1. $\mathbf{w} \succ \mathbf{w}'$;
2. $\exists$ *cond. prob. dist.* $\varphi(y|x)$ *such that* $w'(y|\omega) = \sum_x \varphi(y|x) w(x|\omega)$ *for all* y *and* ω.



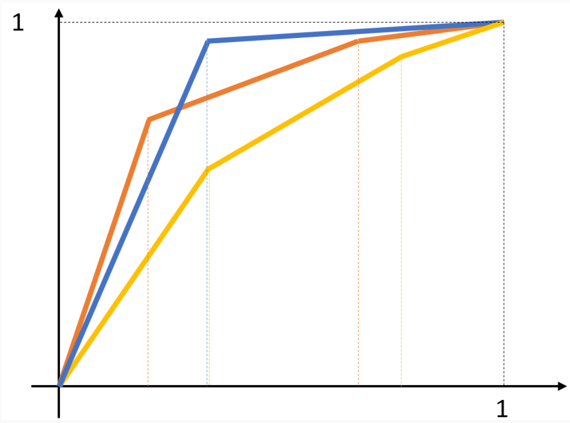
David Blackwell (1919-2010)

The case of dichotomies (a.k.a. relative majorization)

- for $\Omega = \{1, 2\}$, we compare two dichotomies, i.e., two pairs of probability distributions $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ and $(\boldsymbol{q}_1, \boldsymbol{q}_2)$, of dimension m and n, respectively
- relabel entries such that ratios p_1^i / p_2^i and q_1^j / q_2^j are nonincreasing
- construct the truncated sums $P_\omega(k) = \sum_{i=1}^k p_\omega^i$ and $Q_\omega(k) = \sum_{j=1}^k q_\omega^j$
- $(\boldsymbol{p}_1, \boldsymbol{p}_2) \succ (\boldsymbol{q}_1, \boldsymbol{q}_2)$ iff the relative Lorenz curve of the former is never below that of the latter

Blackwell, 1953

$(\boldsymbol{p}_1, \boldsymbol{p}_2) \succ (\boldsymbol{q}_1, \boldsymbol{q}_2) \iff \boldsymbol{q}_\omega = M\boldsymbol{p}_\omega$, for some stochastic matrix M.



$$(x_k, y_k) = (P_2(k), P_1(k)), \quad 1 \leqslant k \leqslant n$$

Quantum extensions

Quantum statistical decision theory (Holevo, 1973)

classical case	**quantum case**
• decision problems $\mathbf{g} = \langle \Omega, \mathcal{A}, c\rangle$	• decision problems $\mathbf{g} = \langle \Omega, \mathcal{A}, c\rangle$
• models $\mathbf{w} = \langle \Omega, \mathcal{X}, \{w(x\|\omega)\}\rangle$	• quantum models $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{\rho_S^\omega\}\rangle$
• decisions $d(a\|x)$	• POVMs $\{P_S^a : a \in \mathcal{A}\}$
• $c_{\mathbf{g}}^*(\mathbf{w}) = \max_{d(a\|x)} \min_\omega \cdots$	• $c_{\mathbf{g}}^*(\mathcal{E}) = \max_{\{P_S^a\}} \min_\omega \sum_a c(\omega, a) \operatorname{Tr}[\rho_S^\omega \; P_S^a]$

Quantum statistical morphisms (FB, CMP 2012)

Definition (Tests)

Given a quantum statistical model $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{\rho_S^\omega\}\rangle$, a family of operators $\{Z_S^a\}$ is said to be an $\mathcal{E}$-test if and only if there exists a POVM $\{P_S^a\}$ such that

$$\operatorname{Tr}[\rho_S^\omega \; Z_S^a] = \operatorname{Tr}[\rho_S^\omega \; P_S^a] \;, \quad \forall \omega, \forall a \;.$$

Definition (Morphisms)

Given two quantum statistical models $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{\rho_S^\omega\}\rangle$ and $\mathcal{E}' = \langle \Omega, \mathcal{H}_{S'}, \{\sigma_{S'}^\omega\}\rangle$, a linear map $\mathcal{M} : \mathsf{L}(\mathcal{H}_S) \to \mathsf{L}(\mathcal{H}_{S'})$ is said to be an $\mathcal{E} \to \mathcal{E}'$ quantum statistical morphism iff

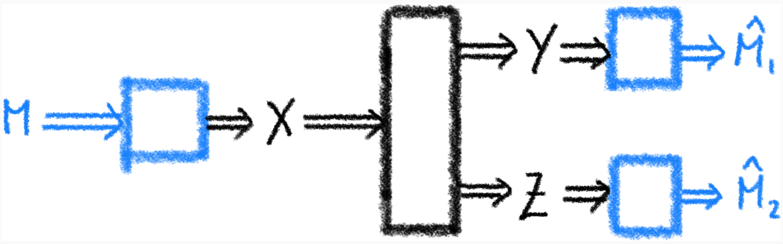
1. $\mathcal{M}$ is trace-preserving;
2. $\mathcal{M}(\rho_A^\omega) = \sigma_{S'}^\omega$, for all $\omega \in \Omega$;
3. the trace-dual map $\mathcal{M}^\dagger : \mathsf{L}(\mathcal{H}_{S'}) \to \mathsf{L}(\mathcal{H}_S)$ maps $\mathcal{E}'$-tests into $\mathcal{E}$-tests.

Quantum statistical comparison (FB, CMP 2012)

- let $\mathcal{E} = \langle \Omega, \mathcal{H}_S, \{\rho_S^\omega\} \rangle$ and $\mathcal{E}' = \left\langle \Omega, \mathcal{H}_{S'}, \{\sigma_{S'}^\omega\} \right\rangle$ be given
- information ordering: $\mathcal{E} \succ \mathcal{E}'$ iff $c^*_{\mathbf{g}}(\mathcal{E}) \geqslant c^*_{\mathbf{g}}(\mathcal{E}')$ for all $\mathbf{g}$
- *complete* information ordering: $\mathcal{E} \ggcurly \mathcal{E}'$ iff $\mathcal{E} \otimes \mathcal{F} \succ \mathcal{E}' \otimes \mathcal{F}$ for all ancillary models $\mathcal{F} = \left\langle \Theta, \mathcal{H}_A, \{\tau_A^\theta\} \right\rangle$
- **Theorem 1/3**: $\mathcal{E} \succ \mathcal{E}'$ iff there exists a *quantum statistical morphism* $\mathcal{M} : \mathsf{L}(\mathcal{H}_S) \to \mathsf{L}(\mathcal{H}_{S'})$ such that $\mathcal{M}(\rho_S^\omega) = \sigma_{S'}^\omega$, for all $\omega \in \Omega$
- **Theorem 2/3**: $\mathcal{E} \ggcurly \mathcal{E}'$ iff there exists a *completely positive trace-preserving linear map* $\mathcal{N} : \mathsf{L}(\mathcal{H}_S) \to \mathsf{L}(\mathcal{H}_{S'})$ such that $\mathcal{N}(\rho_S^\omega) = \sigma_{S'}^\omega$, for all $\omega \in \Omega$
- **Theorem 3/3**: if $\mathcal{E}'$ is *commutative*, that is, if $[\sigma_{S'}^{\omega_1}, \sigma_{S'}^{\omega_2}] = 0$ for all $\omega_1, \omega_2 \in \Omega$, then $\mathcal{E} \ggcurly \mathcal{E}'$ iff $\mathcal{E} \succ \mathcal{E}'$

Applications in information theory

Classical broadcast channels



How to capture the idea that Y carries more information than Z?

(i) (stochastically) degradable: $\exists$ channel $Y \to Z$

(ii) less noisy: for all M, $H(M|Y) \leqslant H(M|Z)$

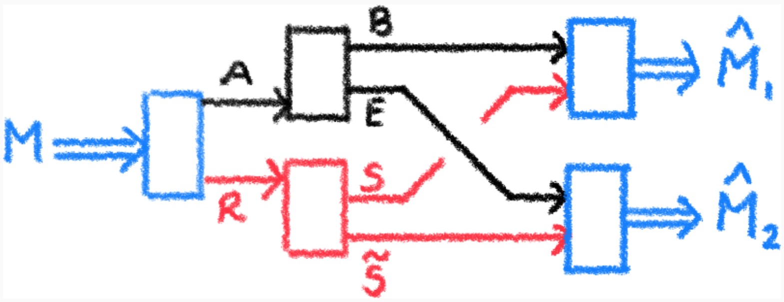
(iii) less ambiguous: for all M, $\max \mathbf{Pr}\{\hat{M}_1 = M\} \geqslant \max \mathbf{Pr}\{\hat{M}_2 = M\}$

(iv) less ambiguous (reformulation): for all M, $H_{\min}(M|Y) \leqslant H_{\min}(M|Z)$

Theorem (Körner–Marton, 1977; FB, 2016)

less noisy $\underset{\not\Rightarrow}{\Longleftarrow}$ *degradable* $\Longleftrightarrow$ *less ambiguous*

Quantum broadcast channels



(i) (CPTP) degradable: $\exists$ channel $B \to E$

(ii) *completely* less noisy: for all M and all symmetric side-channels $R \to S\tilde{S}$, $H(M|BS) \leqslant H(M|E\tilde{S})$

(iii) *completely* less ambiguous: for all M and all symmetric side-channels $R \to S\tilde{S}$, $H_{\min}(M|BS) \leqslant H_{\min}(M|E\tilde{S})$

Theorem (FB–Datta–Strelchuk, 2014)

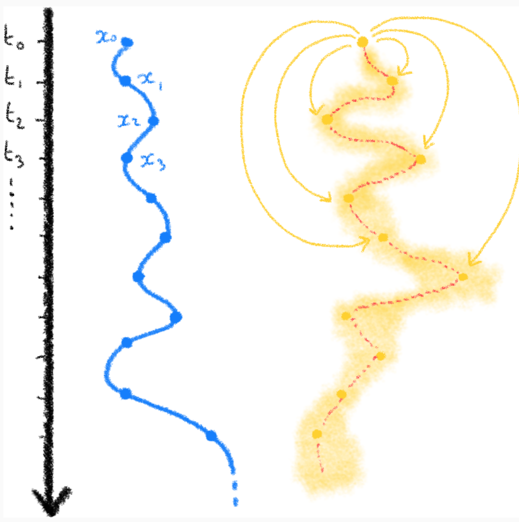
completely less noisy $\underset{\not\Rightarrow}{\Longleftarrow}$ *degradable* $\Longleftrightarrow$ *completely less ambiguous*

Applications in open quantum systems dynamics

Discrete-time stochastic processes

Formulation of the problem:

- for $i \in \mathbb{N}$, let x_i index the state of a system at time $t = t_i$
- given the system's initial state at time $t = t_0$, the process is fully predicted by the conditional distribution $p(x_N, \ldots, x_1 | x_0)$
- if the system evolving is quantum, we only have a quantum dynamical mapping $\left\{\mathcal{N}^{(i)}_{Q_0 \to Q_i}\right\}_{i \geqslant 1}$
- the process is divisible if there exist channels $\mathcal{D}^{(i)}$ such that $\mathcal{N}^{(i+1)} = \mathcal{D}^{(i)} \circ \mathcal{N}^{(i)}$ for all $i \geqslant 1$
- **problem**: to provide a *fully information-theoretic characterization* of divisibility

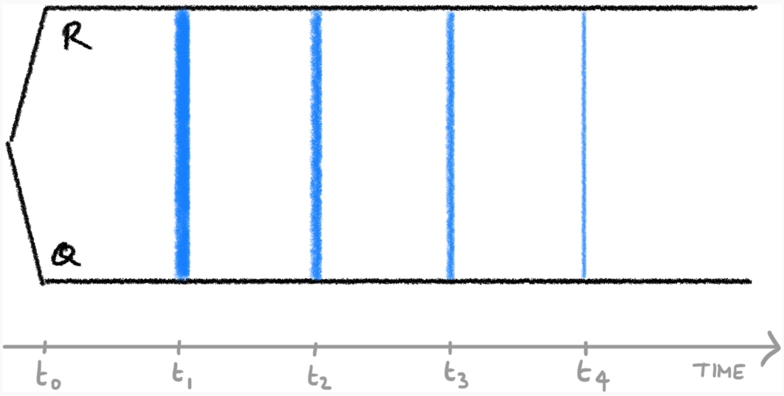


Divisibility as locality of information flow

Theorem (FB–Datta, 2016; FB, 2018)

Given an initial open quantum system Q_0, a quantum dynamical mapping $\left\{\mathcal{N}^{(i)}_{Q_0\to Q_i}\right\}_{i\geqslant 1}$ is divisible if and only if, for any initial state ω_{RQ_0},

$$H_{\min}(R|Q_1) \leqslant H_{\min}(R|Q_2) \leqslant \cdots \leqslant H_{\min}(R|Q_N)\ .$$



Thermalization as relative majorization

Basic idea (FB, arXiv:1505.00535)

Thermal accessibility $\rho \to \sigma$ can be characterized as the statistical comparison between quantum dichotomies (ρ, γ) and (σ, γ), for γ thermal state

Two main problems:

- for dimension larger than 2 and $[\sigma, \gamma] \neq 0$, we need a complete (i.e., extended) comparison
- moreover, Gibbs-preserving channels can create coherence between energy levels, while a truly thermal operation should not

Conclusions

Conclusions

- the theory of statistical comparison studies morphisms (preorders) of one "statistical system" X into another "statistical system" Y
- equivalent conditions are given in terms of (finitely or infinitely many) monotones, e.g., $f_i(X) \geqslant f_i(Y)$
- such monotones quantify the resources at stake in the operational framework at hand, e.g.
  - the expected maximin payoff in decision problems for experiments
  - the information asymmetry for broadcast channels
  - the non-divisibility for open systems dynamics
  - the joint time-energy information for quantum thermodynamics

Thank you