



Construtivismo cognitivo e estatística bayesiana

Julio MICHAEL STERN



RESUMO

Neste artigo, algumas questões epistemológicas, ontológicas e sociológicas concernentes à significância estatística de hipóteses precisas na pesquisa científica são investigadas dentro do arcabouço fornecido pelo construtivismo cognitivo e do FBST – o teste completo de significância bayesiana. O arcabouço construtivista é contrastado com o da teoria da decisão e o do falsificacionismo, os cenários epistemológicos tradicionais para a estatística bayesiana ortodoxa e para a estatística frequentista.

PALAVRAS-CHAVE • Autopoiese. Auto-soluções. Construtivismo cognitivo. Epistemologia. Hipóteses precisas. Significância estatística. Teoria dos sistemas.

INTRODUÇÃO

Neste artigo, algumas questões epistemológicas, ontológicas e sociológicas, concernentes à significância estatística de hipóteses precisas no contexto científico, são investigadas no arcabouço fornecido pelo construtivismo cognitivo, ou teoria construtivista (TCon), como apresentada em Maturana & Varela (1980), Foerster (2003),¹ e Luhmann (1989, 1990a, 1990b, 1995). Várias conclusões deste estudo permanecem válidas, *mutatis mutandis*, em várias outras organizações e sistemas (cf. Bakken & Hernes, 2002; Christis, 2001; Mingers, 1995; Rasch, 1998).

O interesse do autor neste tópico de investigação provém de seu envolvimento no desenvolvimento do teste completo de significância bayesiana (FBST – *Full Bayesian Significance Test*), uma nova solução bayesiana para o problema estatístico de medir o suporte a hipóteses precisas, proposto em Pereira & Stern (1999). O problema de medir o suporte a hipóteses precisas coloca várias dificuldades conceituais e metodo-

¹ Os seguintes artigos (capítulos) desta antologia são de especial interesse: (1) On self-organizing systems and their environments, p. 1-19; (8) On constructing a reality, p. 211-27; (11) Objects: tokens for eigen-behaviors, p. 261-71; (15) For Niklas Luhmann: how recursive is communication? , p. 305-23.

lógicas para a análise estatística tradicional, tanto na abordagem frequentista (clássica), quanto na abordagem bayesiana. A solução dada pelo FBST tem vantagens significativas sobre as alternativas tradicionais, em relação a suas propriedades estatísticas e lógicas. Como essas propriedades já foram extensamente analisadas em outros artigos do autor, o foco deste artigo é dirigido exclusivamente a questões ontológicas e epistemológicas.

O FBST é totalmente compatível com a Teoria da Decisão (TDec) (cf. Madruga *et al.*, 2001), e esta fornece um arcabouço epistemológico forte e coerente para a estatística bayesiana ortodoxa. Todavia, as propriedades lógicas do FBST abrem possibilidades para o uso de arcabouços epistemológicos alternativos. Neste artigo, o arcabouço epistemológico da TCon é contraposto ao da TDec. O contraste é, todavia, limitado em escopo por nosso interesse em estatística, e conduzido de forma exploratória e não-exaustiva. O arcabouço epistemológico da TCon é também contraposto ao do falsificacionismo, o arcabouço epistemológico no qual a teoria estatística frequentista de teste de hipóteses é, geralmente, apresentada (cf. Boyd *et al.*, 1991; Popper 1972 [1934]).

Na seção 1, revisamos as noções fundamentais de autopoiese e auto-soluções em sistemas autopoieticos. Na seção 2, fazemos o mesmo com as noções de sistemas sociais e diferenciação funcional e, na seção 3, apresentamos uma visão construtivista da ciência. Na seção 4, o material apresentado nas seções 1, 2 e 3 é relacionado à significância estatística de hipóteses precisas em ciência, e essas conclusões são contrapostas às interpretações tradicionais, como as da TDec. Na seção 5, revisamos algumas análises sociológicas de fenômenos de “dediferenciação”. Na seção 6 e 7, apresentamos nossas conclusões finais.

As seções 1, 2, 3 e 5 visam tornar mais acessíveis conceitos já bem estabelecidos da TCon, que infelizmente não são amplamente conhecidos nas comunidades científica e estatística, as quais têm maior familiaridade com as interpretações tradicionais do falsificacionismo frequentista ou da TDec. Optamos por citar as fontes primárias sempre que foi possível fazê-lo (em nosso contexto) de forma clara e concisa, evitando assim rephrasing desnecessariamente esses conceitos. As contribuições nas seções 4, 6 e 7 referem-se essencialmente à análise do papel de métodos quantitativos concebidos para medir o suporte estatístico a hipóteses precisas.

I AUTOPOIESE E AUTO-SOLUÇÕES

O conceito de autopoiese tenta capturar uma característica essencial de organismos vivos (auto = própria; poiesis = produção). Seu propósito e definição são enunciados em Maturana & Varela (1980, p. 84 e 78-79):

Nosso propósito foi o de propor uma caracterização de seres vivos que explicasse a geração de todos os fenômenos que lhe são próprios. Fizemos isto designando autopoiese em um espaço físico como uma condição necessária e suficiente para um sistema ser considerado vivo. [...]

Um sistema autopoietico é organizado (definido como uma unidade) como uma rede de processos de produção (transformação e destruição) de componentes, componentes estes que:

- (i) pelas suas interações e transformações continuamente regeneram e realizam a rede de processos (relações) que os produzem, e
- (ii) constituem (a máquina) como uma unidade concreta no espaço em que eles (os componentes) existem pela especificação do domínio topológico de sua realização como uma rede.

Sistemas autopoieticos são sistemas dinâmicos fora de equilíbrio (dissipativos) que exibem estruturas (meta-)estáveis, cuja organização permanece invariante no (por longos períodos de) tempo, não obstante a freqüente substituição de seus componentes. Ademais, estes componentes são produzidos pelas próprias estruturas que regeneram. Por exemplo, a população de macro-moléculas de uma única célula pode ser renovada milhares de vezes durante sua vida (cf. Bertalanffy, 1969). A investigação desses processos regenerativos em redes de produção de sistemas autopoieticos leva à definição de domínio cognitivo, como enunciado em Maturana & Varela (1980, p.10):

A circularidade de sua organização continuamente os leva de volta aos mesmos estados internos (mesmos com respeito ao processo cíclico). Cada estado interno requer que certas condições (interações com o ambiente) sejam satisfeitas para prosseguir para o estado seguinte. Portanto, a organização circular implica uma predição de que uma interação que ocorreu no passado voltará a ocorrer novamente. Se tal não ocorre, o sistema mantém sua integridade (identidade em relação ao observador) e faz uma nova predição. Em um ambiente continuamente mutável, estas predições só serão bem sucedidas se o ambiente não se alterar em relação ao que é predito. Assim, as predições decorrentes da organização do sistema vivo não são predições de eventos particulares, mas de classes de interações. Cada interação é uma interação particular, mas cada predição é uma predição para uma classe de interações definida pelas características dos elementos que permitirão ao sistema vivo reter sua organização circular após a interação, e portanto interagir novamente. Isso torna sistemas vivos sistemas inferenciais, e seu domínio de interações um domínio cognitivo.

As características desses processos regenerativos circulares (cíclicos ou recursivos), e seus auto-estados (*eigen-states*, ou estados de equilíbrio, fixos, homeostáticos, invariantes, recorrentes, recursivos, etc.) são ainda investigados, em sistemas autopoieticos, concretos ou abstratos, em Foerster (2003) e Segal (2001):

O significado de recursão é o de novamente percorrer seu próprio caminho. Um de seus resultados é, sob certas condições, a existência de soluções que, ao reentrar no formalismo, produzem novamente as mesmas soluções. Estas são denominadas auto-valores, auto-funções, auto-comportamentos, etc., dependendo do domínio de aplicação da transformação – um domínio de números, de funções, de comportamentos, etc. (Segal, 2001, p. 145).

O conceito de auto-solução para um sistema autopoietico é uma noção-chave para distinguir objetos específicos em um domínio cognitivo. Von Foerster também identifica várias propriedades essenciais de auto-soluções que são básicas para as análises feitas neste artigo e suas conclusões:

Objetos são *tokens* [instâncias] para auto-comportamentos. *Tokens* representam algo distinto. Em troca de dinheiro (ele mesmo um *token* para ouro guardado por um governo, embora desafortunadamente não mais conversível), *tokens* são usados para se ter acesso a uma viagem de metrô, ou a um jogo em um equipamento de diversão. Na esfera cognitiva, objetos são nomes-*token* que damos aos nossos auto-comportamentos. Este é o *insight* construtivista do que acontece quando falamos sobre nossa experiência com objetos (Segal, 2001, p.127).

Auto-valores apresentam-se como sendo ontologicamente discretos, estáveis, separáveis e componíveis, enquanto que, ontogeneticamente, surgem como equilíbrios que se determinam por processos circulares. Ontologicamente, auto-valores e objetos, e da mesma forma ontogeneticamente, comportamentos estáveis e a manifestação da compreensão que um sujeito tem de um objeto, não podem ser distinguidos (Foerster, 2003, p. 266).

A linha de argumentação desenvolvida neste estudo depende fortemente de duas propriedades qualitativas de auto-soluções referidas por von Foerster pelos termos “discreto” e “equilíbrio”. No que se segue, o significado destes qualificativos, como entendidos por von Foerster e aqui utilizados, será examinado:

(a) Discreto (ou preciso):

Há um ponto adicional que quero mencionar, um ponto importante. Dentre um contínuo de possibilidades, operações recursivas esculpem um conjunto preciso de soluções discretas. Auto-comportamentos geram entidades discretas e identificáveis. Produzir soluções discretas dentre uma variedade infinita tem consequências incrivelmente importantes. Nos permite começar a nomear objetos. Linguagem é a possibilidade de esculpir, dentre um número infinito de experiências possíveis, aquelas que permitem uma interação estável de nós mesmos conosco (Segal, 2001, p. 128).

É importante entender que, na seqüência, o termo “discreto”, utilizado por von Foerster para caracterizar auto-soluções em geral, pode ser substituído, dependendo do contexto específico, por termos como “de menor dimensão”, preciso, singular etc. Mesmo no caso familiar da álgebra linear, se definirmos um auto-vetor, correspondente a um auto-valor singular, c , de uma transformação linear, T , apenas pela propriedade essencial de invariância direcional, $T(x) = cx$, obteremos sub-variedades de dimensão 1, que neste caso são retas passando pela origem. Apenas se acrescentarmos à definição a usual (embora não essencial) condição de normalização, $\|x\| = 1$, é que obteremos auto-vetores discretos.

(b) Estável (ou equilíbrio):

Dada uma auto-solução estável de um operador, $Op()$, definida por uma equação de invariância ou equação de ponto fixo, $x_{inv} = Op(x_{inv})$, esta auto-solução pode ser encontrada (construída ou computada) como o limite, x_∞ , da seqüência $\{x_n\}$ quando $n \rightarrow \infty$, seqüência esta definida pela aplicação recursiva do operador, $x_{n+1} = Op(x_n)$. Em condições apropriadas (como por exemplo dentro de um limite de atração), a convergência do processo e sua auto-solução limite não dependem do ponto inicial, x_0 . No exemplo de álgebra linear, utilizando quase qualquer ponto inicial, a seqüência gerada pelo operador dado pela aplicação da transformação linear, T , seguida de uma normalização, $x_{n+1} = T(x_n)/\|T(x_n)\|$, converge para o auto-vetor unitário correspondente ao máximo auto-valor.

Nas próximas duas seções mostraremos, para a análise estatística no contexto científico, como a propriedade de precisão indica que muitas, e talvez as mais relevantes, hipóteses científicas são precisas, e como a propriedade de estabilidade indica que considerar essas hipóteses é razoável e natural. As consequências estatísticas dessas conclusões serão discutidas nas seções 4, 6 e 7. Antes disto, todavia, mais alguns conceitos da teoria construtivista (TCon) precisam ser introduzidos.

“Autopoiese” encontrou seu nome na obra de Maturana & Varela (1980), juntamente com uma formulação simples, poderosa e elegante, usando a moderna linguagem de teoria de sistemas. Não obstante, alguns conceitos básicos da teoria, como auto-organização e autonomia dos organismos vivos, têm uma longa história, que alguns autores remontam a Kant (cf. Kant, 1995 [1790], seção 65, p. 216), onde um “ser vivo” é definido da seguinte maneira:

[...] é necessário que as partes do mesmo se produzam umas às outras reciprocamente em conjunto, tanto segundo a sua forma como na sua ligação, e assim produzam um todo a partir da sua própria causalidade, cujo conceito por sua vez e inversamente (num ser que possuísse a causalidade adequada a um tal produto) poderia ser causa dele mesmo segundo um princípio, e em consequência a conexão das *causas eficientes* poderia ser ajuizada simultaneamente como *efeito mediante causas finais*.

Num tal produto da natureza cada uma das partes, assim como só existe *mediante* as restantes, também é pensada em *função das outras* e por causa do todo, isto é, como instrumento (órgão).

Para outros comentários históricos, vide Zelleny (1980).

2 DIFERENCIAÇÃO FUNCIONAL

Para poder dar respostas apropriadas às complexidades ambientais, sistemas autopoieticos podem ser hierarquicamente organizados como *sistemas autopoieticos de ordem superior*. Em Maturana & Varela (1980, p. 107, 109), essa noção é definida via o conceito de “acoplamento”:

Toda vez que a conduta de duas ou mais unidades é tal que existe um domínio no qual a conduta de cada unidade é função da conduta das demais, dizemos que elas estão acopladas nesse domínio. [...]

Um sistema autopoietico cuja autopoiese requer a autopoiese das unidades acopladas que o realizam é um sistema autopoietico de ordem superior.

Um exemplo típico é uma colméia, um sistema de terceira ordem, formado por abelhas, sistemas de segunda ordem, formadas pelo acoplamento de células, sistemas de primeira ordem.

O filósofo e sociólogo Niklas Luhmann aplicou esta noção para estudar sociedades humanas modernas e seus sistemas. A abstração básica de Luhmann é olhar para

um sistema social apenas em seu nível hierárquico mais alto, onde ele é visto como sistema autopoietico de redes de comunicação. Na terminologia de Luhmann, um evento de comunicação consiste da sua enunciação, a forma de transmissão; da sua informação, o conteúdo específico da mensagem; e do seu entendimento, a relação com eventos futuros na rede, como a ativação ou supressão de futuras comunicações.

Sistemas sociais usam comunicação como seu modo particular de (re)produção autopoietica. Seus elementos são comunicações que são recursivamente produzidas e reproduzidas por uma rede de comunicações, que não são unidades vivas, nem unidades conscientes, nem ações. Sua unidade requer uma síntese de três seleções, a saber: informação, enunciação e entendimento (incluindo desentendimento) (Luhmann, 1990b, p. 3).

Para Luhmann, a melhor estratégia de uma sociedade para lidar com crescentes complexidades é a mesma observada na maioria dos organismos, a saber, diferenciação. Organismos biológicos se diferenciam em sistemas especializados, como tecidos e órgãos, em formas de vida pluricelulares (sistemas não-autopoieticos ou alopoiéticos), ou indivíduos especializados, em colônias de insetos (sistemas autopoieticos). De fato, sociedades e organismos podem ser caracterizados pela forma com que se diferenciam em sistemas. Para Luhmann, sociedades modernas são caracterizadas por uma diferenciação vertical em sistemas funcionais autopoieticos, onde cada sistema é caracterizado pelo seu código, programa e mídia. Nesta caracterização, o código fornece uma referência bipolar para o sistema, daquilo que é positivo, aceito, favorecido ou válido, contra aquilo que é negativo, rejeitado, desfavorecido ou inválido. O programa fornece um contexto específico onde o código é aplicado, e a mídia é o espaço no qual o sistema opera.

Alguns exemplos-padrão de sistemas sociais são:

- Ciência: tendo como código verdadeiro/falso, trabalhando no programa de uma teoria científica, e utilizando uma mídia de artigos em anais e periódicos.
- Jurídico: tendo como código legal/ilegal, trabalhando em um programa definido por leis e regulamentos existentes, e utilizando uma mídia de documentos legais.
- Religioso: tendo como código bem/mal, trabalhando no programa definido por textos revelados e hermenêuticos, e utilizando como mídia o estudo, preces e boas ações.
- Economia: tendo como código ter/não-ter, trabalhando em um programa definido por cenários de planejamento econômico e métodos de precificação, e utilizando como mídia ativos financeiros.

Antes de encerrar esta seção, introduzimos uma noção relacionada à quebra de autopoiese: “dediferenciação” (*Entdifferenzierung*) é a degradação da coerência interna do sistema pela adulteração, corrupção, ou dissolução de suas relações autopoieticas próprias. Uma forma de dediferenciação (quer em sistemas biológicos, quer em sistemas sociais) é a penetração do sistema por agentes externos que tentam utilizar os recursos do sistema de uma forma que não é compatível com a autonomia desse sistema. Na concepção de Luhmann da sociedade moderna, cada sistema pode perceber eventos em outros sistemas (ser cognitivamente aberto), mas deve manter sua própria diferenciação (ser operacionalmente fechado).

Autopoiese requer que o bom funcionamento das sociedades modernas dependa criticamente da manutenção da autonomia operacional de cada um de seus (sub)sistemas (Luhmann, 1989, p. 109).

3 AUTO-SOLUÇÕES E HIPÓTESES CIENTÍFICAS

A interpretação do conhecimento científico como auto-solução de um processo de pesquisa é parte da abordagem construtivista para a epistemologia. A figura 1 apresenta a estrutura idealizada e a dinâmica da produção de conhecimento. Este diagrama representa, do lado experimental (coluna da esquerda) as operações de campo ou laboratório de uma ciência empírica, onde experimentos são desenhados e construídos, efeitos observáveis são gerados e medidos, e o banco de dados experimental é montado. Do lado teórico (coluna da direita), o diagrama representa o trabalho teórico de análise estatística, interpretação e (assim se espera) entendimento de acordo com padrões aceitos. Caso necessário, novas hipóteses (incluindo teorias inteiras) são formuladas, motivando o desenho de novos experimentos. Teoria e experimento constituem um sistema cíclico de dupla retro-alimentação, deixando claro que o desenho de novos experimentos é guiado pela teoria existente e suas interpretações, que por sua vez precisa ser constantemente checada, adaptada, e modificada para dar conta dos resultados experimentais. O sistema como um todo constitui uma unidade autopoietica (cf. Krohn & Küppers, 1990):

- Do lado experimental do diagrama, a tarefa da estatística é o de fazer enunciados probabilísticos sobre a ocorrência de eventos pertinentes, isto é, descrever distribuições de probabilidade para o que, onde, quando, ou quais eventos podem ocorrer. Se os eventos ocorrerão no futuro, essas descrições são denominadas predições, como freqüentemente ocorre em ciências naturais. Também é possível (mais freqüentemente em ciências sociais) lidar com observações relacionadas a eventos passados, que podem ou não ser experimentalmente gerados ou repetidos, impondo limitações em relação à qualidade ou quantidade dos dados disponíveis. Mesmo assim, o hábito de denominar esse tipo de enunciado de “preditivo” será mantido.
- Do lado teórico do diagrama, o papel da estatística é o de medir o suporte estatístico a hipóteses, isto é, o de medir, quantitativamente, a plausibilidade ou possibilidade de hipóteses, dentro do arcabouço teórico em que estas são formuladas, em função dos dados observados. Do material apresentado até agora, fica claro que, nesse papel, a função principal da estatística é a de medir o suporte estatístico a hipóteses precisas, pois o fato dessas hipóteses serem precisas (ou discretas, de menor dimensão etc.) é uma característica essencial de auto-soluções.

Examinemos agora quão bem os paradigmas estatísticos tradicionais e, em contraste, o teste completo de significância bayesiana (FBST), são capazes de dar conta dessa tarefa dual. Para examinar esta questão, o primeiro passo é distinguir que tipos de enunciados estatísticos podem ser feitos. Faremos uso de três categorias de enunciados: freqüentistas, epistêmicos e bayesianos.

Enunciados probabilísticos *freqüentistas* são feitos exclusivamente com base na freqüência de ocorrência de um evento, em uma seqüência (potencialmente) infinita de observações geradas por uma variável aleatória.

Enunciados probabilísticos *epistêmicos* são feitos com base no status epistêmico (grau de crença, possibilidade, validade, verossimilhança) de um evento dentre os possíveis resultados gerados por uma variável aleatória. Essa geração pode ser atual ou potencial (pode ter sido realizada ou não), pode ser observada (ou observável) ou não, pode ser repetida um número infinito ou finito de vezes, ou mesmo acontecer apenas uma única vez.

Enunciados probabilísticos *bayesianos* são enunciados probabilísticos epistêmicos gerados pela (na prática sempre finita) aplicação recursiva da fórmula de Bayes:

$$p_n(\theta) \propto p_{n-1}(\theta)p(x_n|\theta).$$

Em modelos padrão, o parâmetro θ , uma variável aleatória não observada, e a amostra x , uma variável aleatória observada, são relacionadas pela sua distribuição conjunta de probabilidade, $p(x, \theta)$. A distribuição *a priori*, $p_0(\theta)$, é o ponto de partida para a operação bayesiana de recursão. Ela representa a informação inicialmente disponível sobre θ . Em particular, a distribuição *a priori* pode representar nenhuma informação disponível, como em uma distribuição uniforme ou uma distribuição obtida via o princípio de máxima entropia (cf. Dugdale, 1996; Kapur, 1989). A distribuição *a posteriori*, $p_n(\theta)$, representa a informação disponível sobre o parâmetro após o n -ésimo passo de aprendizado, no qual a fórmula de Bayes é usada para incorporar a informação na observação x_n . Dada a natureza recursiva do procedimento, a distribuição *a posteriori* de um dado passo é usada como a distribuição *a priori* do passo seguinte.

A estatística freqüentista exige dogmaticamente que todas os enunciados probabilísticos sejam freqüentistas. Portanto, qualquer enunciado probabilístico formulado diretamente no espaço paramétrico é categoricamente proibido. Hipóteses científicas são enunciados epistêmicos sobre os parâmetros de modelos estatísticos. Assim, a estatística freqüentista não pode fazer nenhum enunciado direto sobre a significância estatística (veracidade) de hipóteses. Estritamente falando, ela pode apenas fazer enunciados no lado experimental do diagrama. A maneira freqüentista de lidar com questões, no lado teórico do diagrama, é imergi-las de alguma forma no lado experimental. Uma forma de fazer isso é usando uma construção em que todo o processo de aquisição de dados é visto como um único resultado em uma seqüência infinita de um metaprocessos estocástico imaginário, e então fazer um enunciado freqüentista, no meta processo, sobre a freqüência de resultados insatisfatórios em relação a alguma medida de incompatibilidade entre o banco de dados observado com a hipótese. Este é o clássico (e muitas vezes esquecido) raciocínio usado quando enunciamos um p -valor.

Um p -valor é a probabilidade de obter uma amostra que seja mais extrema (incompatível com a hipótese) que a amostra que obtivemos. Assim, devemos sempre falar do p -valor de um banco de dados (nunca da hipótese). A confusão conceitual e a frustração resultante (para a maioria dos cientistas) com esse tipo de raciocínio convoluto são capturadas na maravilhosa paródia dos diálogos de Galileu em Rouanet *et al.* (1998).

Ao enunciar um p -valor, devemos também especificar que critério será usado para definir aquilo que entendemos como “mais extremo”, isto é, para ordenar o espaço amostral, e geralmente há vários critérios possíveis para tanto.

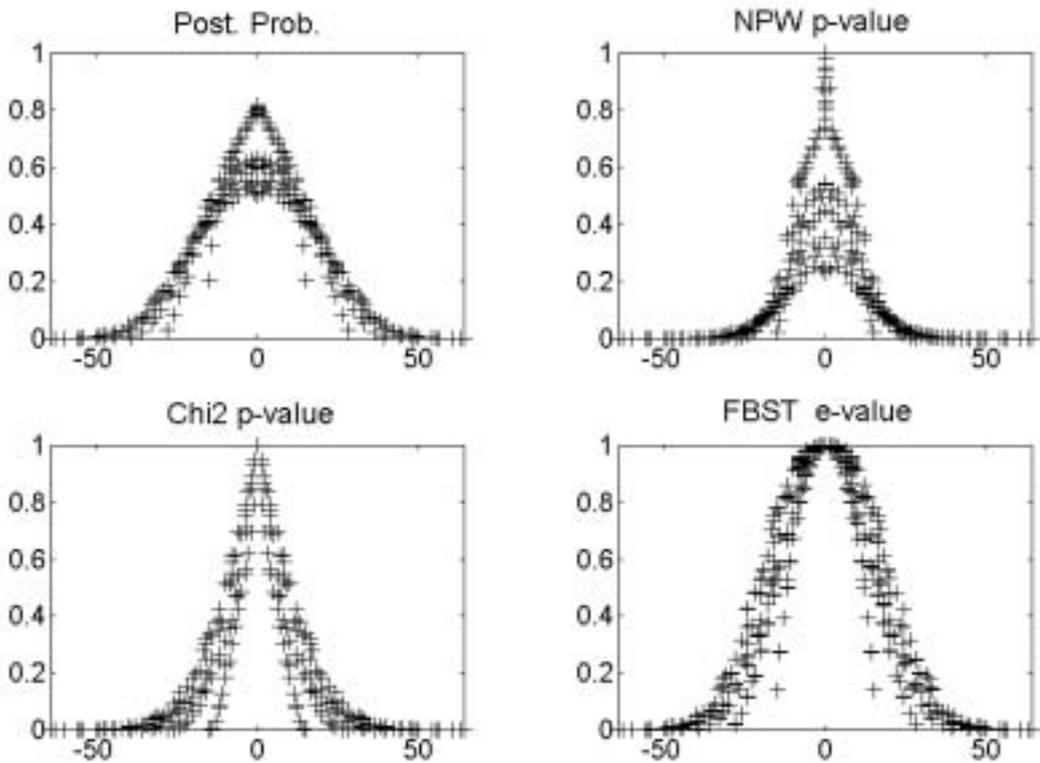


Figura 2. Hipótese de independência, $n=16$.

A figura 2 compara quatro estatísticas, a saber, probabilidades *a posteriori* bayesianas ortodoxas, *p*-valores de Neyman-Pearson-Wald (NPW), *p*-valores obtidos por aproximação chi-quadrado, e os *e*-valores dados pelo FBST em favor de H . Neste exemplo, H é a hipótese de independência em uma tabela de contingência 2×2 , para um tamanho de amostra $n = 16$. O eixo horizontal mostra a estatística de “assimetria diagonal” (diferença entre os produtos das diagonais), que pode ser vista como um estimador para uma versão não-normalizada de ρ , a correlação de Pearson. Para explicações sobre os detalhes técnicos vide Irony & Pereira (1995), Irony, Pereira & Tiwari (2000), Irony, Lauretto, Pereira & Stern (2002), Stern & Zacks (2002), Madruga *et al.* (2003) e Lauretto *et al.* (2003).

Amostras que são “perfeitamente compatíveis” com a hipótese, isto é, sem nenhuma assimetria, estão no centro do eixo horizontal, com amostras crescentemente assimétricas para os lados. A curva envoltória dos *e*-valores produzidos pelo FBST, a ser comentada mais adiante nesta seção, é suave (diferenciável), e portanto nivelada no seu máximo, onde alcança o valor 1.

Em contraste, a curva envoltória dos p -valores tem a forma de uma cúspide, isto é, uma curva pontiaguda, que é quebrada (não diferenciável) no seu máximo, onde também atinge o valor 1. A agudeza da cúspide aumenta com o aumento do tamanho da amostra. No caso dos p -valores de NPW, vemos também, no topo da cúspide, uma “escada” ou “agulha” com várias amostras, todas sem assimetria, mas tendo diferentes probabilidades de ocorrência, “competindo” pelo maior p -valor. Esse é um efeito colateral típico da estratégia de escamotear a pergunta sobre a significância de H , que pede como resposta uma probabilidade no espaço paramétrico, para uma pergunta, condicional à veracidade de H , sobre a probabilidade de obtenção da amostra observada, que tem como resposta uma probabilidade no espaço amostral.

Esta análise qualitativa da metodologia de p -valores nos dá um *insight* sobre alguns (ab)usos da expressão “aumente a amostra para rejeitar”. Nas palavras de I. J. Good (1983, p. 135):

Muitas vezes o estatístico não se dá ao trabalho de esclarecer se a hipótese nula é realmente precisa ou apenas aproximadamente precisa. [...] Não é pois surpreendente que muitos fisherianos (e popperianos) digam que não se pode ter (muita) evidência em favor da hipótese nula, mas apenas refutá-la.

Na estatística bayesiana nos é permitido fazer tanto enunciados no espaço paramétrico, como também (é claro) no espaço amostral. Assim, parece que a estatística bayesiana é a ferramenta certa para o trabalho, e de fato ela o é! No entanto, devemos primeiramente examinar o papel da Teoria da Decisão (TDec) na estatística bayesiana ortodoxa. Desde o trabalho pioneiro de de Finetti, Savage e vários outros, a estatística bayesiana ortodoxa desenvolveu sólidos fundamentos baseados na TDec, onde muitas questões puderam ser analisadas e resolvidas com sucesso.

Essas fundações podem ser estratificadas em duas camadas:

- Na primeira camada, a TDec fornece um sistema de coerência para o uso de enunciados estatísticos (cf. Finetti, 1974, 1981, 1991, 1993). Nesse sentido, o uso que o FBST faz da teoria de probabilidades é completamente compatível com a TDec, como demonstrado por Madruga *et al.* (2001).
- Na segunda camada, a TDec fornece um arcabouço epistemológico para a interpretação de procedimentos estatísticos. As propriedades lógicas do FBST abrem a possibilidade de utilizar e de beneficiar-se de estruturas epistemológicas alternativas, como a teoria construtivista (TCon), de modo que a TDec não tenha que ser “o único instrumento disponível”.

Afirmamos que, no caso específico de procedimentos estatísticos para medir o suporte a (significância estatística de) hipóteses científicas precisas, a TCon fornece um arcabouço epistemológico mais adequado que a TDec. Este ponto é tão importante quanto sutil. Para entendê-lo, lembremo-nos primeiramente do paradigma ortodoxo, como enunciado concisamente em Dubins & Savage (1965). Em uma segunda citação, de Savage (1954), vemos que hipóteses precisas, mesmo se importantes, fazem pouco sentido nesse paradigma (cf. também Levi, 1974; Maher *et al.*, 1993).

Problemas de apostas, nos quais a distribuição de várias quantidades são importantes na descrição da fortuna do jogador, parecem abarcar a totalidade da estatística teórica de acordo com uma visão (que pode ser chamada de visão bayesiana baseada na teoria da decisão) do assunto.

[...] Desde o ponto de vista da estatística baseada em teoria da decisão, o jogador neste problema é a pessoa que deve, em última instância, agir de uma de duas maneiras (os dois palpites), uma das quais seria apropriada sob a hipótese (H_0), e a outra sob sua negação (H_1).

[...] Muitos problemas, dos quais este é uma instância, são grosso modo do seguinte tipo: a opinião de uma pessoa sobre um parâmetro desconhecido é descrita por uma distribuição de probabilidade; sendo-lhe permitido sucessivamente comprar fragmentos de informação sobre o parâmetro, a preços que podem depender (talvez aleatoriamente) dos próprios parâmetros desconhecidos, até que ela finalmente escolhe uma ação terminal pela qual ela recebe um prêmio que depende da ação e dos parâmetros (Dubins & Savage, 1965, seção 12.8, p. 229-30).

Volto-me agora para um tópico diferente e, ao menos para mim, delicado, em conexão com aplicações da teoria de teste (de hipóteses). Muita atenção é dada na literatura estatística àquilo que é pertinente ao teste de hipóteses, nas quais a hipótese nula é tal que esta não seria realmente aceita por ninguém. [...] Hipóteses extremas (precisas) como eu as chamarei.

[...] A inaceitabilidade de hipóteses nulas extremas (precisas) é perfeitamente bem conhecida; ela está intimamente relacionada à conhecida máxima que reza que a ciência falseia, mas nunca comprova hipóteses.

O papel de hipóteses extremas (precisas) em ciência e outras atividades estatísticas parece ser importante mas obscura. Em particular eu, como todo praticante de estatística, muitas vezes “testo” hipóteses extremas (precisas), embora não possa dar uma análise muito satisfatória do processo, nem dizer claramente como ele se relaciona a testes como definidos neste capítulo e em outras discussões da teoria (Savage, 1954, seção 16.3, p. 254).

Como claramente vimos, no arcabouço da TDec falamos sobre chances de aposta para “conseguir” a hipótese em um “jogo” acontecendo no espaço paramétrico. Mas hipóteses precisas são conjuntos com medida (de Lebesgue) nula, portanto nossas chances de aposta devem ser nulas, isto é, hipóteses precisas são (quase) certamente falsas. Se aceitarmos a visão construtivista de que hipóteses se relacionam à identificação de auto-soluções, e que auto-soluções são ontologicamente precisas, teremos um paradoxo!

Nestas considerações não deve ser surpreendente que muitos frequentistas e a maioria dos bayesianos ortodoxos considerem hipóteses precisas como sendo no máximo aproximações grosseiras, utilizadas quando o cientista é incapaz de especificar corretamente intervalos de erro, funções de perda, ganho, custo ou utilidade etc., ou então as consideram pura e simplesmente “meras tolices”. Nas palavras de Williams (2001, p. 234):

Significância bayesiana de hipóteses precisas: um apelo à sanidade. [...] Me surpreende portanto que alguns bayesianos queiram agora dar probabilidade *a priori* não zero de que uma hipótese precisa seja exatamente verdadeira, para obter resultados que parecem suportar fortemente hipóteses nulas que qualquer frequentista definitivamente rejeitaria. (Claro que ter como consequência este tipo de resultado é o óbvio ululante.)

Mas não importa quantas vezes o estatístico repreenda os cientistas por seu desleixo e incompetência, eles continuam formulando hipóteses precisas, como se fossem por elas magneticamente atraídos. Da perspectiva construtivista, com o auxílio da metodologia do FBST, eles estão apenas fazendo o que é correto!

A estatística baseada na TDec também desenvolveu métodos para lidar com hipóteses precisas, geralmente colocando tenebrosos avisos de perigo (*caveat emptor*) para quem queira utilizá-los. Os mais conhecidos destes métodos são os testes de Jeffreys, baseados em fatores de Bayes, obtidos colocando uma massa de probabilidade *a priori* positiva sobre a hipótese precisa. Esta massa *a priori* positiva deveria funcionar como um sistema de *handicap*, desenhado para equilibrar as chances iniciais e tornar o jogo uma competição “justa”. Disto tudo obtemos apenas mais paradoxos, como o bem documentado paradoxo de Lindley (ao contrário de sua contraparte frequentista, este é um efeito do tipo “aumente a amostra para aceitar”) (cf. Shafer, 1982).

O FBST foi especialmente desenhado para fornecer um valor de evidência, o *e*-valor, dando suporte a uma hipótese precisa, H . Esta função de suporte, $Ev(H, p_n)$, é baseada na medida de probabilidade *a posteriori* de um conjunto denominado o conjunto tangente, $T(H, p_n)$, que é um conjunto de medida não nula; portanto, nada de pa-

radoxos oriundos de probabilidades nulas (cf. Pereira & Stern, 1999; Madruga *et al.*, 2003). Não obstante o e -valor $Ev(H, p_n)$ ser uma probabilidade no espaço paramétrico, ele é também uma função de suporte possibilística. O termo “possibilístico” tem uma conotação técnica específica, indicando que $Ev()$ obedece a uma estrutura lógica e algébrica bem definida (cf. Darwiche & Ginsberg, 1992; Darwiche, 1993; Stern, 2003). Ademais, $Ev()$ tem muitas propriedades desejáveis ou necessárias a uma função de suporte estatístico a hipóteses precisas, tais como:

- (1) Fornecer uma função de suporte à (ou medida de significância estatística da) hipótese em teste, idealmente uma medida de *probabilidade no espaço paramétrico* original ou natural do problema.
- (2) Ter uma definição intrinsecamente geométrica, independente de qualquer aspecto não geométrico, como a particular parametrização da (sub-variedade representando a) hipótese sendo testada, ou o particular sistema de coordenadas escolhido para o espaço paramétrico, isto é, ser um procedimento *invariante*.
- (3) Fornecer uma função de suporte *suave*, isto é, contínua e diferenciável, nos parâmetros da hipótese e nas estatísticas da amostra, dentro de condições apropriadas de regularidade do modelo.
- (4) Obedecer ao *princípio da verossimilhança*, isto é, a informação obtida das observações deve ser representada pela, e apenas pela, função de verossimilhança.
- (5) Não requerer qualquer artifício *ad hoc*, como dar probabilidades positivas a conjuntos de medida nula, ou estabelecer razões de crença iniciais arbitrárias entre hipóteses.
- (6) Fornecer um função de suporte *possibilística*.
- (7) Ser um procedimento *exato*, isto é, não utilizar no cálculo do e -valor qualquer aproximação assintótica.
- (8) Fornecer um teste *consistente* para uma dada hipótese precisa.
- (9) Fornecer operações de *composicionalidade* para modelos complexos.
- (10) Permitir a incorporação de experiência prévia ou opiniões de especialistas via *distribuições a priori*.

Para uma explicação detalhada e cuidadosa da definição do FBST, sua implementação computacional, propriedades lógicas e estatísticas, e várias aplicações já desenvolvidas, o leitor pode consultar alguns dos artigos do autor na bibliografia.

5 DEGRADAÇÃO SEMÂNTICA

Nesta seção revisamos algumas análises construtivistas de fenômenos de dediferenciação em sistemas sociais. Se as conclusões das últimas seções estiverem corretas, então é surpreendente o quanto a TDec, por vezes com uma interpretação pseudo-econômica muito estreita, foi mal utilizada em análise estatística. As dificuldades de testar hipóteses nos paradigmas estatísticos tradicionais são extensivamente discutidos e documentados na literatura (cf. por exemplo Harlow *et al.*, 1997). Esperamos que o material nesta seção possa ajudar a entender essas dificuldades como sintomas de problemas com raízes muito mais profundas. De forma alguma este autor foi o primeiro a apontar o perigo de análises feitas pela cega transplantação de categorias entre sistemas heterogêneos. Em particular, quanto ao abuso de análises econômicas:

Neste sentido, não faz sentido falar de custos não econômicos. Esta é apenas uma forma metafórica de falar que transfere a especificidade do modo econômico de pensar indiscriminadamente para outros sistemas sociais (Luhmann, 1989, p. 164).

Para uma análise sociológica desse fenômeno no contexto da ciência, ver as seguintes citações:

[...] ciências de maior *status* podem, mais ou menos agressivamente, colonizar áreas de menor *status* em uma tentativa de reduzi-las a seus próprios Primeiros Princípios. Para a física de partículas, tudo são quarks e as quatro forças. Para a neurofisiologia, consciência é a saída agregada de redes neurais comportamentais. Para a sociobiologia, filosofia é feita por formigas e ratos com cérebros super-desenvolvidos, que tagarelam baboseira metafísica de acordo com reflexos adquiridos. Em resumo, cadeias reducionistas críveis e bem sucedidas usualmente vão de cima para baixo na hierarquia de prestígio das disciplinas (Fuchs, 1996, p. 310).

Isso pode explicar a popularidade de dar um “entendimento econômico” a processos em áreas funcionalmente distintas, mesmo que (ou talvez justamente porque) esta semântica é freqüentemente escondida pela teoria e métodos da estatística baseados em análises da teoria da decisão. Isto também pode explicar porque algumas áreas, como ecologia, sociologia ou psicologia são (ou foram) muito mais suscetíveis a sofrer esse tipo de dediferenciação por degradação semântica, que outras, como a física (DiMaggio & Powell, 1991, p. 63).

Uma vez que as forças levando à degradação sistêmica estão claramente expostas, esperamos que se entendam os seguintes corolários dos famosos imperativos (ético e estético) de von Foerster:

- Imperativo teórico: Preserve a autopoiese sistêmica e a integridade semântica, pois de-diferenciação é a própria in-sanidade.
- Imperativo operacional: Escolha a ferramenta certa para cada tarefa: “se você só tem um martelo, tudo fica parecendo prego”.

6 HIPÓTESES PRECISAS CONCORRENTES

Nesta seção examinaremos o conceito de hipóteses precisas concorrentes (ou competitivas). Este conceito tem muitas variantes, mas a idéia básica é que um bom cientista nunca deve testar uma única hipótese precisa, pois isto seria uma luta desleal da pobre hipótese precisa enfrentando sozinha o resto do mundo. Ao contrário, o bom cientista deveria sempre confrontar uma hipótese precisa com uma hipótese precisa adversária, tornando o jogo uma competição justa.

Como eu considero refutação e corroboração como sendo ambos critérios válidos para esta demarcação, é conveniente usar outro termo, checabilidade, para abarcar ambos os processos. Eu considero a checabilidade como uma medida de quão científica é uma teoria, onde checar deve ser tomado em ambos os sentidos, positivo e negativo, confirmando ou refutando.

[...] Se por “a verdade da mecânica newtoniana” nós entendêssemos que esta é aproximadamente verdadeira, em algum sentido apropriadamente bem definido, poderíamos obter uma forte evidência de que ela seja verdadeira. Se, no entanto, entendêssemos que ela é exatamente verdadeira, então é como se já a tivéssemos refutado.

[...] Eu penso que a probabilidade inicial é positiva para qualquer teoria científica consistente que tenha conseqüências verificáveis em um sentido probabilístico. Nenhuma contradição pode ser deduzida desta afirmação, pois o número de teorias enunciáveis é no máximo contavelmente finito (enumerável).

[...] É muito difícil decidir sobre o valor numérico de probabilidades, mas não é tão difícil julgar quocientes de probabilidades iniciais subjetivas para duas teorias por comparação de suas complexidades. Esta é uma razão pela qual a história da ciência é cientificamente importante (Good, 1983, p. 167, 135, 126).

O argumento das hipóteses concorrentes não contradiz diretamente o arcabouço epistemológico apresentado neste artigo, e pode até ser apropriado em certas circunstâncias. Este argumento pode também minorar ou remediar parcialmente os paradoxos apontados, nas seções precedentes, nos testes de hipóteses precisas dos paradigmas tradicionais da estatística, freqüentista ou bayesiano ortodoxo. No entanto, o autor não acredita que ter sempre hipóteses precisas concorrentes seja nem uma condição necessária para a boa prática científica, nem uma descrição acurada da história da ciência.

Examinemos rapidamente o primeiro incidente do tumultuado desmoronamento da mecânica newtoniana (apenas para ficar com o exemplo usado por Good). Este incidente foi o experimento de Michelson sobre o efeito do “vento etéreo” sobre a velocidade da luz (cf. Michelson & Morley, 1887; Lorentz *et al.*, 1952; Jaffe, 1960). Na verdade, Michelson não encontrou tal efeito, isto é, ele descobriu que a velocidade da luz é constante, invariante com a velocidade relativa do observador. Este resultado, uma contradição na mecânica newtoniana, é facilmente explicado pela teoria da relatividade restrita de Einstein. A diferença fundamental entre as duas teorias são seus grupos de simetria ou invariância: o grupo de Galileu para a mecânica newtoniana, e o grupo de Lorentz para a relatividade restrita. Um resultado fundamental da física, o teorema de Noether, garante que, para qualquer grupo contínuo de simetria em uma teoria física, deve haver uma quantidade invariante ou lei de conservação (cf. Noether, 1971 [1918]; Doncel *et al.*, 1987; Gruber *et al.*, 1980-98; French, 1968; Landau & Lifschitz, 1966). Leis de conservação são hipóteses precisas por excelência, ideais para verificação experimental. Assim, parece que estamos exatamente na situação de hipóteses precisas concorrentes, e assim estamos, hoje, em uma perspectiva histórica muito distante. Mas esta é uma análise *post-mortem* da mecânica newtoniana. Ao tempo do experimento não havia nenhuma teoria concorrente. Ao invés de confirmar um efeito esperado, Michelson encontrou, para grande espanto seu e de todo mundo, um efeito nulo, dentro da precisão do experimento.

Experimentos complexos, como o de Michelson, requerem uma análise cuidadosa de erros experimentais, identificando todas as fontes significantes de ruído e flutuações de medida. Este tipo de análise é usual em física experimental, e motiva um breve comentário sobre uma fonte secundária de críticas a hipóteses precisas. No passado era freqüentemente necessário trabalhar com modelos estatísticos super simplificados. Esta situação era muitas vezes imposta por limitações, como a falta de modelos melhores e mais realísticos, ou a indisponibilidade dos algoritmos numéricos necessários, ou o custo computacional de utilizá-los. Sob estas limitações, era comum o uso de modelos estatísticos minimalistas ou de técnicas de aproximação, mesmo quando não recomendadas. Estes modelos ou técnicas foram instrumentais para fornecer ferramentas viáveis de análise estatística, mas tornaram muito difícil trabalhar (ou ti-

nam baixa performance) com sistemas complexos, poucas observações, bancos de dados muito grandes etc. A necessidade de trabalhar com sistemas complexos e outras situações difíceis que requerem o uso de modelos e técnicas estatísticas sofisticadas é muito comum (e freqüentemente inescapável) em áreas de pesquisa lidando com sistemas complexos, como biologia, medicina, ciências sociais, psicologia, e muitas outras áreas que compartilham a misteriosa denominação de “ciências *soft*”. Como um colega uma vez me disse, “parece que a física ficou com todos os problemas fáceis...”.

Se há uma área na qual as técnicas computacionais da estatística bayesiana fizeram contribuições dramáticas nas últimas décadas, esta é a análise de modelos complexos. O desenvolvimento de técnicas estatísticas computacionais avançadas como MCMC (Monte Carlo em cadeias de Markov), redes neurais e bayesianas, modelos de campos aleatórios, e muitas outras, nos dão a esperança que a maioria das dificuldades advindas do uso de modelos super simplificados possam agora ser superadas. Hoje em dia, a boa prática estatística requer que todas as influências estatisticamente relevantes sejam incorporadas ao modelo, e raramente há uma desculpa aceitável para não proceder dessa maneira (cf. Pereira & Stern, 2001).

7 COMENTÁRIOS FINAIS

Devemos novamente frisar que a maior parte do material apresentado nas seções 1, 2, 3 e 5 não é novo na Teoria Construtivista (TCon). Infelizmente a TCon tem tido pouco impacto em estatística, ou por vezes provocado uma reação hostil (ainda que pouco informada). Uma possível explicação para essa situação pode ser encontrada no desenvolvimento histórico da TCon. A reação construtivista ao dogmatismo objetivista do realismo metafísico originou uma retórica inflamada cuja intenção era esclarecer e explicitar quão frágeis e ingênuas eram as fundações desse realismo simplório. Esta retórica foi muito bem sucedida, rapidamente influenciando as idéias daqueles interessados nas áreas de história e filosofia da ciência (ou despertando-os de seu sono dogmático) e rapidamente disseminando-se também para outras áreas. Infelizmente, essa mesma retórica pode, em uma leitura superficial, fazer com que a TCon seja percebida como hostil ou intrinsecamente incompatível com o uso de métodos estatísticos e quantitativos, ou tendendo a formas extremadas de subjetivismo. Na TCon (ou idealismo) nem afirmamos ter acesso à “coisa em si” (*Ding an sich*) no ambiente externo, como o fazem as formas dogmáticas do realismo, nem nos rendemos ao solipsismo, como o fazem as formas céticas do subjetivismo, incluindo formas extremas da interpretação subjetivista em probabilidade e estatística (cf. Caygill, 1995; Finetti, 1974, seções 1.11, 7.5.7). De fato, o papel das restrições externas impostas pelo am-

biente, conjuntamente com as relações autopoieticas internas do sistema, é guiar a convergência do processo de aprendizado para auto-soluções precisas, sendo estas, ao fim e ao cabo, os últimos (e neste sentido reais) objetos do conhecimento. Como dito por Luhmann:

[...] o construtivismo mantém nada além da inacessibilidade do mundo externo “em si mesmo” e o fechamento do conhecimento – sem se render jamais à velha dúvida cética ou solipsista de questionar a própria existência do mundo exterior (Luhmann, 1990a, p. 65).

[...] ao menos na teoria dos sistemas, eles (enunciados) se referem ao mundo real. Portanto o conceito de sistema se refere a algo que em realidade é um sistema e que portanto tem a responsabilidade de testar seus enunciados contra a realidade (Luhmann, 1995, p. 12).

[...] teorias subjetivistas e objetivistas do conhecimento devem ambas ser substituídas por distinções da forma sistema / ambiente, que então tornam a distinção sujeito / objeto irrelevante (Luhmann, 1990a, p. 66).

O autor espera ter mostrado que a TCon não apenas fornece uma visão balanceada e eficiente dos aspectos teóricos e experimentais da pesquisa científica, mas também que ela está apta (ou mesmo é a mais apta) a fornecer os fundamentos epistemológicos necessários ao uso dos métodos quantitativos da análise estatística requeridos na prática científica. Deve ser ainda enfatizada, conforme a interpretação do autor, a importância de medir o suporte estatístico a hipóteses precisas. Nesse sentido, o autor acredita que, devido a suas características lógicas e estatísticas, o FBST é a ferramenta certa para esta tarefa. Essa perspectiva abre assim várias áreas para pesquisa futura, dentre as quais, mencionamos duas a seguir.

7.1 TEOREMAS DE NOETHER E DE FINETTI

A primeira área para pesquisa futura refere-se a algumas similaridades entre teoremas de Noether em física, e teoremas de de Finetti em estatística. Teoremas de Noether fornecem quantidades físicas invariantes ou leis de conservação a partir de grupos de simetria de uma teoria física, e leis de conservação são hipóteses precisas por excelência. De forma similar, teoremas de de Finetti fornecem distribuições invariantes a partir de grupos de simetria do modelo estatístico. Essas distribuições invariantes podem,

por seu turno, fornecer protótipos de hipóteses precisas em muitas áreas de aplicação. A física, via teoremas de Noether, tem seu poderoso aparato para lidar com os importantíssimos assuntos de invariância e simetria. A estatística, via os teoremas de de Finetti, poderia fornecer esse tipo de aparato a outras áreas, mesmo em situações que não estão naturalmente imersas em um formalismo matemático tão pesado (cf. Feller, 1971, cap.7; Diaconis & Freeman, 1987; Diaconis, 1988; Eaton, 1989; Ressel, 1987).

7.2 COMPOSICIONALIDADE

A segunda área para pesquisa futura refere-se a uma das propriedades de auto-soluções mencionadas por von Foerster que não foi diretamente explorada neste artigo, a saber, que auto-soluções são “componíveis” (cf. Borges & Stern, 2005). Propriedades de composicionalidade referem-se à relação entre a credibilidade de uma hipótese complexa, H , e a de suas hipóteses constituintes elementares, H^j , $j=1, \dots, k$. Questões de composicionalidade têm um papel central em filosofia analítica.

De acordo com Wittgenstein (1993 [1921], seções 2.0201, 5, 5.32, p. 139, 203, 217):

Todo enunciado sobre complexos pode-se decompor em um enunciado sobre as partes constituintes [...]

A proposição é uma função de verdade das proposições elementares.

Todas as funções de verdade são resultados da aplicação sucessiva de um número finito de operações de verdade às proposições elementares.

Questões de composicionalidade surgem também em contextos bem mais concretos, como na engenharia de confiabilidade, conforme expresso na seguinte citação de Birnbaum *et al.* (1961, seção 1.4):

Um dos propósitos principais de teoria matemática de confiabilidade é o de desenvolver formas de avaliar a confiabilidade de uma estrutura quando a confiabilidade de seus componentes é conhecida. O presente estudo abordará esse tipo de desenvolvimento matemático. Com este propósito, será necessário rephrasing nossos conceitos intuitivos de estrutura, componente, confiabilidade etc. em uma linguagem mais formal, enunciar cuidadosamente nossos pressupostos e introduzir um aparato matemático apropriado.

Em Luhmann (1989, p.79), encontramos um comentário sobre a evolução da ciência que indica diretamente a importância dessa propriedade:

Após o sistema (científico) ter trabalhado por vários séculos sob essas condições, ficou claro para onde ele se dirigia. Isso é algo que idealização, matematização, abstração etc. não descrevem adequadamente. Isso se refere à crescente capacidade de decomposição e recombinação, uma nova formulação de conhecimento como produto de análise e síntese [...] (assim) revelando um enorme potencial de recombinação.

Na visão do autor, a composição (ou re-combinação) de conhecimento científico e seu uso, tão relevante no desenvolvimento tecnológico e em engenharia, pode nos dar uma perspectiva diferente (talvez uma abordagem *bottom-up*, ao invés da abordagem *top-down* deste artigo) da importância de hipóteses precisas na prática de ciência e tecnologia. O lado formal deste problema é tratado em Borges & Stern (2005). O estudo das implicações desta questão poderá também fornecer algum *insight* sobre as formas válidas de interação da ciência com outros sistemas sociais ou, na terminologia de Luhmann, como a ciência pode (ou deve) “ressoar” na sociedade humana. ☉

AGRADECIMENTOS. O autor se beneficiou do suporte da FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo, CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico, BIOINFO, Programa de Bio-Informática, e DCC, Departamento de Computação, do IME-USP, Instituto de Matemática e Estatística da Universidade de São Paulo, e ainda do Departamento de Ciências Matemáticas da SUNY-Binghamton, State University of New York, USA. O autor é grato a muitas pessoas por úteis e interessantes discussões, muito especialmente Wagner Borges, Carlos Humes, Joseph Kadane, Luis Gustavo Esteves, Marcelo Lauretto, Fabio Nakano, Carlos Alberto de Bragança Pereira, Rafael Bassi Stern, Sergio Wechsler e Shelemyahu Zacks. O autor recebeu diversos comentários e sugestões interessantes sobre uma versão preliminar deste artigo (Stern, 2005), apresentado na FIS-2005, a Terceira Conferência sobre Fundamentos de Ciência da Informação, em Paris, em julho de 2005.

Julio MICHAEL STERN

Instituto de Matemática e Estatística da

Universidade de São Paulo.

jstern@ime.usp.br

ABSTRACT

In this paper, some epistemological, ontological, and sociological issues are investigated, concerning the statistical significance of sharp hypotheses in scientific research, within the framework offered by cognitive constructivism and FBST – the full Bayesian significance test. The constructivist framework is contrasted to those furnished by decision theory and falsificationism, the traditional epistemological approaches to orthodox Bayesian statistics and to frequentist statistics.

KEYWORDS • Autopoiesis. Eigen-solutions. Cognitive constructivism. Epistemology. Sharp hypotheses. Statistical significance. Systems theory.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAKKEN, T. & HERNES, T. *Autopoietic organization theory: drawing on Niklas Luhmann's social systems perspective*. Copenhagen: Copenhagen Business School, 2002.
- BERTALANFFY, L. VON. *General system theory*. New York: George Braziller, 1969.
- BIRNBAUM, Z. W.; ESARY, J. D. & SAUNDERS, S. C. Multicomponent systems and structures, and their reliability. *Technometrics*, 3, p. 55-77, 1961.
- BORGES, W. & STERN, J. M. On the truth value of complex hypotheses. *CIMCA-2005: International conference on computational intelligence for modelling control and automation*. Los Alamitos: IEEE, 2005.
- BOYD, R.; CASPER, P. & TROUT, J. D. *The philosophy of science*. Cambridge: MIT Press, 1991.
- CAYGILL, H. *A Kant dictionary*. Oxford: Blackwell, 1995.
- CHRISTIS, J. Luhmann's theory of knowledge: beyond realism and constructivism? *Soziale Systeme*, 7, p. 328-49, 2001.
- DARWICHE, A. Y. *A symbolic generalization of probability theory*. Stanford: Stanford University Press, 1993.
- DARWICHE, A. Y. & GINSBERG, M. L. A symbolic generalization of probability theory. *Proceedings of the 10th national conference on artificial intelligence*, 1992. p. 622-7. (AAAI-92.)
- DIACONIS, P. *Group representation in probability and statistics*, 1988. (Institute of Mathematical Statistics Lecture Notes 11.)
- DIACONIS, P. & FREEMAN, D. A dozen de Finetti style results in search of a theory. *Annales de l'Institut Henri Poincaré, Probabilité et Statistique*, 23, p. 397-423, 1987.
- DI MAGGIO, P. & POWELL, W. W. *New institutionalism in organizational analysis*. Chicago: Chicago University Press, 1991.
- DONCEL, M. G.; HERMANN, A.; MICHEL, L. & PAIS, A. *Symmetries in physics (1600-1980)*. Barcelona: Seminari d'Història des les Ciències, Universitat Autònoma de Barcelona, 1987.
- DUBINS, L. E. & SAVAGE, L. J. *How to gamble if you must: inequalities for stochastic processes*. New York: McGraw-Hill, 1965.
- DUGDALE, J. S. *Entropy and its physical meaning*. London: Taylor & Francis, 1996.
- EATON, M. L. *Group invariance applications in statistics*. Hayward: Institute of Mathematical Statistics and American Statistical Association, 1989.
- FELLER, W. *An introduction to probability theory and its applications*. New York: Wiley, 1971.
- FINETTI, B. de. *Theory of probability*. London: Wiley, 1974. 2 v.
- _____. *Scritti*. Padova: Cedam, 1981. v. 1.
- _____. *Scritti*. Padova: Cedam, 1991. v. 2.
- _____. *Probabilità e induzione*. Bologna: Clueb, 1993.
- FOERSTER, H. VON. *Understanding understanding: essays on cybernetics and cognition*. New York: Springer, 2003.
- FRENCH, A. P. *Special relativity*. New York: Chapman & Hall, 1968.
- FUCHS, S. The new wars of truth: conflicts over science studies as differential modes of observation. *Social Science Information*, 35, p. 307-26, 1996.
- GOOD, I. J. *Good thinking*. Minneapolis: University of Minnesota Press, 1983.
- GRUBER, B. et al. (Ed.). *Symmetries in science*. New York: Plenum, 1980-98. 9 v.
- HARLOW, L. L.; MULAİK, S. A. & STEIGER, J. H. *What if there were no significance tests?* London: Lawrence Erlbaum, 1997.
- HAYAKAWA, Y.; IRONY, T. Z. & XIE, M. (Org.). *Systems and Bayesian reliability*. Singapore: World Scientific, 2002.

- IRONY, T. Z. & PEREIRA, C. A. B. Bayesian hypothesis test: using surface integrals to distribute prior information among the hypotheses. *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo*, 2, 27-46, 1995.
- IRONY, T. Z.; PEREIRA, C. A. B. & TIWARI, R. C. Analysis of opinion swing: comparison of two correlated proportions. *The American Statistician*, 54, 57-62, 2000.
- IRONY, T. Z.; LAURETTO, M.; PEREIRA, C. A. B. & STERN, J. M. A Weibull wearout test: full Bayesian approach. In: HAYAKAWA, Y.; IRONY, T. Z. & XIE, M. (Org.). *Systems and Bayesian reliability*. Singapore: World Scientific, 2002. p. 287-300.
- JAFFE, B. *Michelson and the speed of light*. New York: Anchor, 1960.
- KANT, I. *Crítica da faculdade do juízo*. Tradução V. Rohden & A. Marques. Rio de Janeiro: Forense Universitária, 1995 [1790].
- KAPUR, J. N. *Maximum entropy models in science and engineering*. New York: Wiley, 1989.
- KROHN, W. & KÜPPERS, G. The selforganization of science: outline of a theoretical model. In: KROHN, W.; KÜPPERS, G. & NOWOTNY, H. (Org.). *Selforganization: portrait of a scientific revolution*. Dordrecht: Kluwer, 1990. p. 208-22.
- LANDAU, L. D. & LIFSHITZ, E. M. *Cours de physique théorique*. Moscou: MIR, 1966.
- LAURETTO, M.; PEREIRA, C. A. B.; STERN, J. M. & ZACKS, S. Comparing parameters of two bivariate normal distributions using the invariant full Bayesian significance test. *Brazilian Journal of Probability and Statistics*, 17, p. 147-68, 2003.
- LEVI, I. *Gambling with truth: an essay on induction and the aims of science*. Cambridge: The MIT Press, 1974.
- LORENTZ, H. A.; EINSTEIN, A.; MINKOWSKY, H. & WEYL, H. *The principle of relativity: a collection of original memoirs on the special and general theory of relativity*. New York: Dover, 1952.
- LUHMANN, N. *Ecological communication*. Chicago: Chicago University Press, 1989.
- _____. The cognitive program of constructivism and a reality that remains unknown. In: KROHN, W.; KÜPPERS, G. & NOWOTNY, H. (Org.). *Selforganization: portrait of a scientific revolution*. Dordrecht: Kluwer, 1990a. p. 64-86.
- _____. *Essays on self-reference*. New York: Columbia University Press, 1990b.
- _____. *Social systems*. Stanford: Stanford University Press, 1995.
- MADRUGA, M. R.; ESTEVES, L. G. & WECHSLER, S. On the bayesianity of Pereira-Stern tests. *Test*, 10, p. 291-9, 2001.
- MADRUGA, M. R.; PEREIRA, C. A. B. & STERN, J. Bayesian evidence test for precise hypotheses. *Journal of Statistical Planning and Inference*, 117, p. 185-98, 2003.
- MAHER, P.; SKYRMS, B.; ADAMS, E. W.; BINMORE, K.; BUTTERFIELD, J.; DIACONIS, P. & HARPER, W. L. *Betting on theories*. Cambridge: Cambridge University Press, 1993.
- MATURANA, H. R. & VARELA, F. J. *Autopoiesis and cognition: the realization of the living*. Dordrecht: Reidel, 1980.
- MICHELSON, A. A. & MORLEY, E. W. On the relative motion of the earth and the luminiferous ether. *American Journal of Physics*, 34, p. 333-45, 1887.
- MINGERS, J. *Self-producing systems: implications and applications of autopoiesis*. New York: Plenum, 1995.
- NACHBIN, L. *The Haar integral*. Princeton: Van Nostrand, 1965.
- NOETHER, E. Invariant variation problems. Tradução M. A. Tavel. *Transport Theory and Statistical Physics*, 1, p. 183-207, 1971 [1918].
- PEREIRA, C. A. B. & STERN, J. M. Evidence and credibility: full Bayesian significance test for precise hypotheses. *Entropy Journal*, 1, p. 69-80, 1999.
- _____ & _____. Model selection: full Bayesian approach. *Environmetrics*, 12, p. 559-68, 2001.
- PETITJEAN, M. (Org.). *Foundations of Information Science*. Basel: MDPI, 2005. v. 61.

- POPPER, K. R. *A lógica da pesquisa científica*. Tradução L. Hegenberg & O. S. da Mota. São Paulo: Cultrix, 1972 [1934].
- _____. *Conjecturas e refutações*. Tradução S. Bath. Brasília: Editora UnB, 1974 [1963].
- RASCH, W. Luhmann's Widerlegung des Idealismus: constructivism as a two-front war. *Soziale Systeme*, 4, p. 151-61, 1998.
- RESSEL, P. A Very general de Finetti type theorem. In: VIERTL, R. *Probability and Bayesian statistics*. New York: Plenum, 1987. p. 403-14.
- ROUANET, H.; BERNARD, J. M.; BERT, M. C.; LECOUTRE, B.; LECOUTRE, M. P. & LE ROUX, B. *New ways in statistical methodology: from significance tests to Bayesian inference*. Berna: Peter Lang, 1998.
- SAVAGE, L. J. *The foundations of statistics*. New York: Dover, 1954.
- SEGAL, L. *The dream of reality: Heinz von Foerster's constructivism*. New York: Springer, 2001.
- SHAFER, G. Lindley's paradox. *Journal of the American Statistical Association*, 77, p. 325-51, 1982.
- STERN, J. M. Significance tests, belief calculi, and burden of proof in legal and scientific discourse. *Frontiers in Artificial Intelligence and its Applications*, 101, p. 139-47, 2003. (Laptec'03.)
- _____. Cognitive constructivism, eigen-solutions, and sharp statistical hypotheses. In: PETITJEAN, M. (Org.). *Foundations of information science*. Basel: MDPI, 2005. v. 61. p. 1-23.
- STERN, J. M. & ZACKS, S. Testing independence of Poisson variates under the Holgate bivariate distribution: the power of a new evidence test. *Statistical and Probability Letters*, 60, p. 313-20, 2002.
- VIERTL, R. *Probability and Bayesian statistics*. New York: Plenum, 1987.
- WILLIAMS, D. *Weighing the odds*. Cambridge: Cambridge University Press, 2001.
- WITTGENSTEIN, L. *Tractatus logico-philosophicus*. Tradução L. H. L. dos Santos. São Paulo: Edusp, 1993 [1921].
- ZELLENY, M. *Autopoiesis, dissipative structures, and spontaneous social orders*. Washington: American Association for the Advancement of Science, 1980.

