

SENSOMETRICS: ANÁLISE SENSORIAL DE BEBIDAS NUMA PERSPETIVA ESTATÍSTICA

Eduardo ESTEVES⁽¹⁾

Resumo

A utilização de pessoas como instrumentos de medida, desde «simples» consumidores a participar em estudos de mercado até provadores treinados ou peritos integrando painéis de provadores, é cada vez mais importante no desenvolvimento e controlo de qualidade de produtos alimentares e de bebidas.

O vinho é uma bebida fascinante e complexa, na qual uma grande multiplicidade de aromas e sabores podem ser identificados e dependem de vários fatores, das castas ao *terroir* ou ao processo de vinificação. A avaliação da sua qualidade é essencial e, tradicionalmente, realizada pelos enólogos, mas outros intervenientes na indústria têm a responsabilidade de avaliar a qualidade do vinho. Essa avaliação é, muitas vezes, realizada recorrendo à análise sensorial (viz. "ciência envolvida na avaliação das características organolépticas de um produto pelos sentidos"). Para além das condições ambientais e do "protocolo laboratorial", os testes sensoriais são, em termos de planeamento e análise (estatística) de resultados, como testes de hipóteses estatísticas/significância estatística cujo *modus operandi* está relacionado com o método científico.

Neste contributo, proponho a tradução e definição do conceito de Sensometria, apresento de forma elementar e sumária algumas das técnicas e métodos estatísticos mais relevantes e/ou emergentes em análise sensorial, com referência ao software utilizável e que facilita a tarefa (com enfoque na linguagem de programação e sistema de computação R), assim como exponho os aspetos importantes a considerar no planeamento e análise de resultados de provas/testes sensoriais que contribuem para a validade e generalidade das conclusões.

Palavras-chave: Análise sensorial, Estatística, Sensometria, *Software*.

1. INTRODUÇÃO

A utilização de pessoas como *instrumentos de medida*, sejam «simples» consumidores a participar em estudos de mercado ou provadores treinados ou peritos integrando painéis de provadores, é cada vez mais importante no desenvolvimento de produtos, designadamente alimentos e bebidas, e na inovação (BROCKHOFF, 2011). De facto, apesar da evolução tecnológica (e.g. *narizes e línguas eletrónicas*), alguns estímulos originados pelos alimentos e bebidas apenas são percebidos pelos sentidos enquanto outros são demasiado complexos para os instrumentos de análise!? Por outro

⁽¹⁾ Departamento de Engenharia Alimentar, Instituto Superior de Engenharia, Universidade do Algarve e CCMAR Centro de Ciências do Mar. DEA-ISE, Universidade do Algarve, Campus da Penha, 8005-139 Faro, Portugal

lado, em algumas situações, as metodologias (instrumentais) físico-químicas e microbiológicas não produzem resultados *utilizáveis* em “tempo útil” ou *consistentes* com as sensações provocadas.

O vinho é uma bebida fascinante e complexa, na qual uma grande multiplicidade de aromas e sabores podem ser identificados; alguns vinhos são caracterizados por complexidades e facetas que estão para além até das capacidades descritivas da linguagem (LANGSTAFF, 2010). A qualidade um vinho depende de vários fatores, e.g. as castas, o solo/*terroir* ou o processo de vinificação, e a sua avaliação é essencial e, tradicionalmente, realizada pelos enólogos, *sensorialmente*, que têm formação e experiência para elaborar os vinhos de acordo com requisitos do rótulo/marca e detetar vinhos defeituosos. Todavia, outros intervenientes na indústria têm a responsabilidade de avaliar a qualidade do vinho (LANGSTAFF, 2010).

A perceção sensorial dos alimentos e bebidas está obviamente relacionada com os sentidos (tópico muito interessante, mas que não se detalha neste artigo; cf. LAWLESS e HEYMANN, 2010 cap. 2, ou JACKSON, 2014 cap. 11). Por outro lado, a relação entre um estímulo físico e a resposta fisiológica/sensorial é um assunto relativamente complexo que, em bastantes aspetos, é ainda objeto de investigação no âmbito da psicofísica (cf. LAWLESS, 2013).

Entende-se por Análise Sensorial, a “disciplina da Ciência usada para evocar, medir, analisar e interpretar as reações às características dos alimentos e materiais tal como são percebidas pelos sentidos da visão, olfato, paladar, tato e audição” (STONE e SIDEL, 2004). Esquemáticamente, observe-se a Figura 1. A definição de análise sensorial na norma ISO 5492:2008² é: “ciência envolvida na avaliação das características organoléticas de um produto pelos sentidos”. Acrescento uma outra definição³: “disciplina científica que **aplica os princípios do planeamento experimental e da análise estatística** à utilização dos sentidos (da visão, olfato, paladar, tato e audição) para a avaliação de produtos de consumo”, destacando a parte a negrito que é o «mote» deste artigo.

Segundo BROCKHOFF (2011), “in food research, data is produced and used similar[ly] to the industrial use and academic environments specifically for sensory and consumer sciences exist world-wide. The development and application of statistics and data analysis in this area is called **sensometrics**”. Doutra forma, “**sensometrics** [is] the academic discipline that unites sensory perception [of food], or in other words how we experience food through our senses, with mathematics and statistics” (TOMASCO VARELA e NAES, 2016). Mais simplisticamente, “**sensometrics** is a branch of statistics that involves the use of statistical methods in sensory and consumer science” (CAMO,

² ISO 5492:2008 Sensory analysis – Vocabulary. International Organization for Standardization: Geneva, Switzerland.

³ https://pt.wikipedia.org/wiki/Análise_sensorial (acesso em 15/4/2016).

2016). Proponho, por analogia com a Econometria⁴ e a Quimiometria⁵, e seguindo alguns autores de língua portuguesa (SILVIA D. DUTCOSKY e ADILSON DOS ANJOS⁶, e ERIC B. FERREIRA⁷, *com. pess.*), que:

Sensometria é a disciplina da ciência relacionada com o desenvolvimento e a aplicação de métodos matemáticos e estatísticos a resultados de análise sensorial e ciência do consumo.

Na prática, se não for conduzida por um cientista, engenheiro ou tecnólogo alimentar, a Sensometria é realizada por alguém com formação em Estatística, Psicofísica/Psicologia experimental ou Química/Quimiometria.

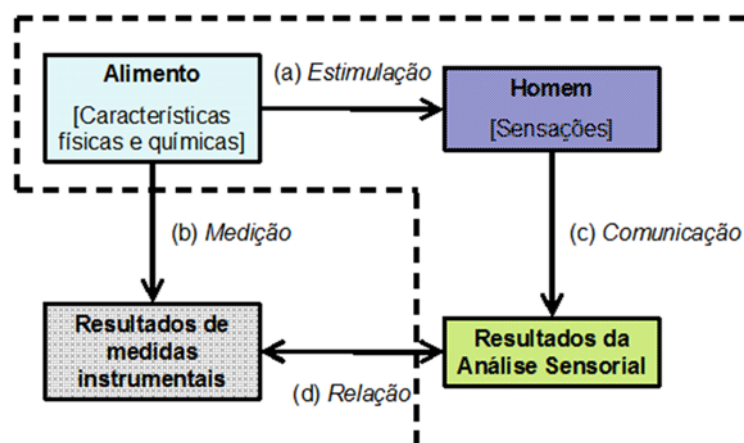


Figura 1 – Ilustração do conceito atual de Análise Sensorial cujas componentes estão delimitadas pela linha tracejada (ESTEVEZ, 2016).

Neste artigo, proponho a tradução e definição do conceito de Sensometria, apresento de forma elementar e sumária algumas das técnicas e métodos estatísticos mais relevantes e/ou emergentes em análise sensorial, com referência ao software utilizável e que facilita a tarefa (com enfoque na linguagem de programação e sistema de computação R), assim como exponho os aspetos importantes a considerar no planeamento e análise de resultados de provas/testes sensoriais que contribuem para a validade e generalidade das conclusões.

⁴ <https://pt.wikipedia.org/wiki/Econometria> (acesso em 15/4/2016).

⁵ <https://pt.wikipedia.org/wiki/Quimiometria> (acesso em 15/4/2016).

⁶ http://www.aboutsolution.com.br/upload/noticias/arg/2_foldersensometriacomr2016.pdf (acesso em 16/4/2016).

⁷ http://www.lce.esalq.usp.br/arquivos/seminarios/2015/resumo_sensometria.pdf (acesso em 16/4/2016).

2. TESTES SENSORIAIS: PAINÉIS DE PROVADORES E RESULTADOS DE PROVAS SENSORIAIS

O «equipamento» utilizado em análise sensorial é o painel de provadores⁸. Ao contrário dos instrumentos de medição, os provadores (humanos) podem ser afetados por fatores psicológicos e fisiológicos. A preparação e planeamento da análise sensorial devem assegurar a eliminação ou redução da influência desses fatores. Podem dividir-se os fatores que influenciam a análise sensorial de produtos alimentares em dois grupos: uns dependentes do indivíduo; e outros relacionados com o ambiente. Podem subdividir-se os fatores dependentes do indivíduo em capacidades “inatas” e adquiridas. As capacidades inatas são, por exemplo, a *atitude* (relacionada com a personalidade), a *motivação* (com raízes fisiológicas, mas também relacionada com o interesse pessoal), a *capacidade de adaptação*, os “erros psicológicos” (de habituação, de expectativa, de estímulo, de lógica, de tendência central, de contraste, de proximidade e de associação; cf. STONE e SIDEL, 2004) e *outras* (como por exemplo, a memória, a experiência, a concentração, a atenção, etc.). As capacidades adquiridas estão relacionadas com instruções específicas da análise sensorial e com o treino⁹ (cf. NP EN ISO 8586:2014¹⁰, ou DE VOS, 2010). Sendo certo que algumas podem ser despistadas *a priori*, outras podem/devem ser aferidas *a posteriori* (v. mais adiante). As condições (ambientais) do local/sala de provas e os equipamentos/utensílios são, também, muito importantes. No caso da avaliação sensorial de vinho, para além características das cabinas de prova e da iluminação, o número de amostras, os copos usados (“copo ISO”¹¹) ou modo de apresentação das amostras são muito relevantes, assim como o conjunto de atributos sensoriais a avaliar, a escala a usar e a ficha para registar os resultados das provas (JACKSON, 2014).

Os resultados de provas sensoriais *resultam* de apreciações, *juízos*, realizados por pessoas – por isso são de natureza, eminentemente, subjetiva – e a sua quantificação inclui-se no âmbito da psicofísica, *i.e.*, a ciência que estuda as relações entre os estímulos e as respostas sensoriais. Nestas relações, assumem-se existir duas *escalas*, a psicológica (*S*) e a física (*R*), cuja *correspondência* pode ser descrita por várias *leis*, ou modelos (matemáticos). Os quatro principais modelos são de Weber (1834?), de Fechner-Weber (ca. 1860), de Steven (ca. 1960) e de Thurstone (em 1927 e depois em 1959) (cf. MEILGAARD *et al.*, 2007, GACULA Jr. *et al.*, 2009).

⁸ NAES *et al.* (2010) distinguem os painéis de provadores treinados dos painéis de consumidores. Com os primeiros pretende-se, com um nº relativamente pequeno de provadores (10–15) e da forma mais objetiva possível, descrever as características ou distinguir/diferenciar os produtos alimentares, enquanto os segundos (que envolvem mais pessoas, 100–150) são utilizados para estudar a apreciação, aceitabilidade ou preferência das pessoas ou grupos de pessoas pelo(s) produto(s).

⁹ https://www.ivdp.pt/pt/docs/Instrucoes_sensorial_alabe.pdf (acesso em 12/5/2016).

¹⁰ NP EN ISO 8586:2014 Sensory analysis -- General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessor. International Organization for Standardization: Geneva, Switzerland.

¹¹ ISO 3591:1977 Sensory analysis -- Apparatus -- Wine-tasting glass. International Organization for Standardization: Geneva, Switzerland.

Os testes sensoriais são o *modus operandi* para “responder” aos objetivos estabelecidos em Análise Sensorial. Existem testes objetivos e subjetivos. Os primeiros providenciam dados sobre as características dos produtos e são executados por provadores (treinados). Estes testes sensoriais podem dividir-se em testes discriminatórios (para determinar se existem diferenças sensoriais entre amostras) e descritivos (para identificar a natureza e/ou a magnitude da diferença sensorial entre produtos). Os testes subjetivos, também designados testes afetivos ou testes a/de consumidores, providenciam dados sobre a aceitabilidade, agrado ou preferência de/por um produto e são levados a cabo por provadores não-treinados.

Para além dos procedimentos “laboratoriais”, práticos que têm a ver com as condições¹² e «utensílios» utilizados durante a prova, o planeamento e a análise dos resultados estão intimamente relacionados com a robustez e validade ou «generalidade» das conclusões. Testes sensoriais bem-sucedidos dependem de objetivos claros, de planeamento experimental robusto, da aplicação de técnicas estatísticas apropriadas, de boas-práticas éticas (no laboratório, e não só...) e da apresentação assertiva/informativa de resultados aos decisores (KEMP *et al.*, 2009; LAWLESS e HEYMANN, 2010).

Os testes sensoriais são, em termos de planeamento e análise (estatística) de resultados, como testes de hipóteses estatísticas/significância estatística cujo *modus operandi* tem «muito a ver» com o método científico. Resumidamente: estabelecidas inicialmente as hipóteses “de trabalho” (hipóteses nula e alternativa, H_0 e H_1) e definida a confiança c que se pretende nas conclusões (sabendo que $c=1-\alpha$), obtêm-se (das provas sensoriais) os resultados pertinentes e calcula-se uma estatística de teste (e.t.) adequada – que, de certa forma, resume a informação angariada com a experiência/prova sensorial – e, finalmente, “interpreta-se” essa e.t. “à luz” das hipóteses iniciais e em função de “certa e determinada” regra de decisão (muitas vezes definida como uma região de rejeição) para, finalmente, se concluir *a favor* de uma das hipóteses “de trabalho” (e.g. ESTEVES, 2011; VINING e KOWALSKI, 2011).

As apreciações dos atributos sensoriais dos produtos alimentares (na prática, os resultados das provas sensoriais) podem ser de vários tipos: nominais, ordinais, intervalares e absolutas (*cf.* ESTEVES, 2011; GACULA Jr. *et al.*, 2009; MEILGAARD *et al.*, 2007). O tipo de dados que se obtém das provas sensoriais, desde os testes mais simples e clássicos, e.g. teste triangular, aos mais elaborados, por ex. perfis sensoriais, *Quantitative Descriptive Analysis* (QDA®), ou *Spectrum*TM, que recorrem a painéis de provadores treinados que utilizam escalas para descrever e avaliar, qualitativamente e

¹² cf. OIV (2016). Document sur l’analyse sensorielle du vin. <http://www.oiv.int/public/medias/4291/document-sur-l-analyse-sensorielle-du-vin.pdf> (acesso em 11/4/16).

quantitativamente, os atributos sensoriais dos produtos alimentares e bebidas, condiciona, em certa medida, as metodologias de análise, nomeadamente as estatísticas de teste usadas e as conclusões que se podem elaborar.

Concordo com FINDLAY (2010) quando aquele autor escreve que “deve ficar claro desde o início que a Estatística é simplesmente uma ferramenta para melhorar a tomada de decisões sobre a qualidade”. Nas secções seguintes, abordo algumas aplicações estatísticas em testes discriminatórios e em testes descritivos.

3. APLICAÇÕES ESTATÍSTICAS EM TESTES DISCRIMINATÓRIOS

Os testes discriminatórios¹³, na prática, “have acquired individual names and a history of use” (MEILGAARD *et al.*, 2007), como por exemplo, os testes de comparação-par, duo-trio e triangular. Nestes testes, tanto o «protocolo laboratorial» como a análise de resultados consequente estão estabelecidos há bastante tempo. Os dados obtidos, do tipo categórico (de facto, binário: respostas corretas e incorretas), podem ser analisados com base na distribuição de probabilidades binomial ou, mais vulgarmente, em aproximações àquela distribuição, *viz.* qui-quadrado (χ^2) ou normal (z) (LAWLESS e HEYMANN, 2010; MEILGAARD *et al.* 2007; STONE e SIDEL, 2004): basicamente, se o número de respostas corretas (considerado por via da e.t.) for superior a um certo valor limite, rejeitamos a hipótese nula e concluímos que existe uma diferença perceptível entre os produtos.

Por exemplo, no caso do teste triangular, em que se pretende que o provador identifique a amostra diferente quando lhe são apresentadas em simultâneo três amostras codificadas, duas iguais e uma diferente¹⁴, calcula-se uma e.t. (χ^2) a partir do nº de respostas corretas (X_1) e incorretas (X_2):

$$\chi^2 = \frac{(|4X_1 - 2X_2| - 3)^2}{8n}$$

Se $\chi^2 > \chi^2_{[1-\alpha;1]}$, rejeita-se a H_0 com $(1-\alpha)\cdot 100\%$ de confiança e, por conseguinte, os produtos são percebidos como diferentes pelos provadores.

Apesar desta abordagem ser simples e direta, não corresponde de facto a uma teoria sobre o comportamento sensorial humano neste tipo de testes sensoriais. Por outro lado, esta perspetiva não

¹³ Podem subdividir-se naqueles se destinam a avaliar se existem, globalmente, diferenças sensoriais entre produtos¹³ (*overall difference tests*) e nos testes em que se pede aos provadores para avaliar uma “certa e determinada” característica sensorial (*attribute difference tests*).

¹⁴ O provador testa, pelo menos, uma das combinações possíveis dos dois produtos A e B: AAB, ABA, BAA, BBA, BAB e ABB.

dá respostas a algumas questões interessantes/importantes, como por exemplo: qual é a magnitude da diferença? Como posso garantir que o teste é suficientemente sensível e não *falhar* uma diferença importante (i.e., proteger relativamente ao erro tipo II)? Ou, pelo contrário, como garanto que um resultado (i.e., diferença) significativo é real, tem significado prático (i.e., proteger do erro tipo I)?

THURSTONE (1927, 1959) propôs uma teoria – *lei dos julgamentos comparativos* (ou *law of comparative judgements* no original em inglês) – com base no método dos estímulos constantes¹⁵ para descrever o comportamento face a uma escolha entre duas opções¹⁶. Assumindo que os provadores são capazes de tomar uma decisão relativamente às opções (ou estímulos) e que as respostas sensoriais aos estímulos têm distribuição normal (FIGURA 2a) (MEILGAARD *et al.*, 2007), o resultado da “aplicação” da Estatística àquela *lei* é dado pela equação:

$$\mu_A - \mu_B = Z_{AB} \sqrt{s_A^2 + s_B^2 - 2 \cdot r \cdot s_A s_B}$$

em que $\mu_A - \mu_B$ é a diferença entre intensidade (média) das sensações provocadas pelos estímulos A e B (que podem ser entendidos como dois produtos), Z_{AB} se determina a partir da proporção de vezes em que o estímulo A é considerado maior do que B (Figura 2c), r é a correlação entre A e B, e s_A e s_B são os desvios-padrão de A e B (Figura 2b). A diferença $\mu_A - \mu_B$ quando expressa em unidades de desvio-padrão, i.e. $\delta = (\mu_1 - \mu_2) / \sigma$, é designada distância discriminial δ de Thurstone. Aplicando a perspetiva thurstoniana aos testes triangulares e duo-trio, por exemplo, a relação entre as funções discriminiais correspondentes está ilustrada na Figura 2d.

Mais recentemente, na década de 1960, David M. Green e John A. Swets, aplicaram a teoria da deteção do sinal (*signal detection theory*¹⁷ ou SDT, no original em Inglês) no âmbito da psicofísica. Matematicamente relacionada com a *lei* de Thurstone, permite separar a sensibilidade “real” do observador (leia-se provador) do enviesamento da resposta (ou *response bias*) para avaliar a sensibilidade e desempenho humanos em situações de discriminação (LAWLESS e HEYMANN, 2010; VAN DER KELLEN *et al.*, 2008).

Num contexto experimental, a “matriz de deteção do sinal” ou “matriz estímulo-resposta”, ilustrada na Figura 3a, resume as possibilidades de resposta de um observador/provador (“Sim” ou “Não”) quando confrontado com a presença ou ausência de um estímulo (“Sinal” ou “Ruído”). Com

¹⁵ e.g. https://en.wikipedia.org/wiki/Psychophysics#Method_of_constant_stimuli (acesso em 2/5/2016).

¹⁶ Assume-se que as perceções têm uma componente probabilística conforme com a distribuição normal e que os provadores são capazes de tomar uma decisão.

¹⁷ https://en.wikipedia.org/wiki/Detection_theory (acesso em 2/5/2016).

base nesta experiência “elementar”, a SDT estabelece algumas condições: após muitos ensaios semelhantes, as sensações originadas pelo “sinal” e pelo “ruído” têm distribuição normal com igual variância; e o provador estabelece “naturalmente” um critério (valor-limite ou *cutoff*) estável, com base no qual decide se está perante um “sinal” ou “ruído”¹⁸ (Figura 3b). É, portanto, possível calcular a proporção de vezes em que acerta (“acerto”) e a proporção de vezes em que indica, erradamente, um “sinal” (“falso alarme”) (Figura 3c).

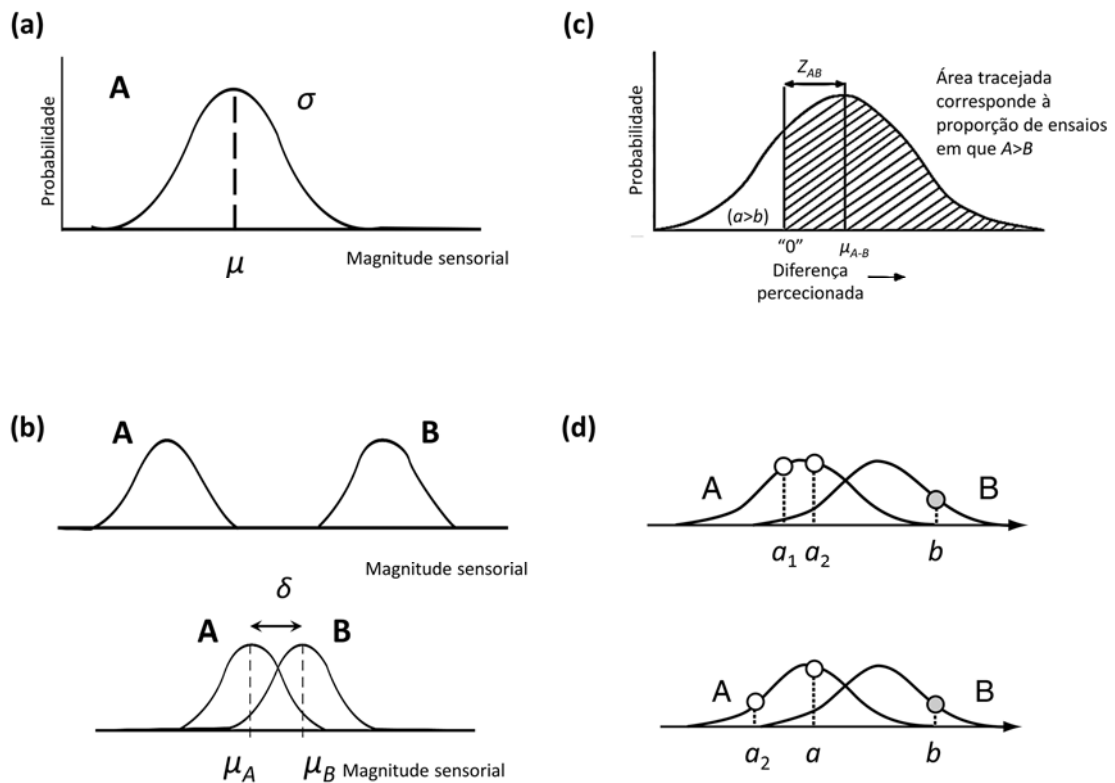


Figura 2 – Ilustrações (a) da distribuição (normal, com média μ e desvio-padrão σ) das respostas sensoriais ao estímulo A – função discriminacional, (b) das funções discriminacionais relativas aos estímulos A e B distintos e da diferença discriminacional δ entre dois estímulos A e B (de médias μ_A e μ_B) parcialmente distinguíveis, (c) da distribuição das diferenças $\mu_A - \mu_B$ e (d) das funções discriminacionais dos estímulos A e B no contexto dos testes triangulares (esq.) e duo-trio (dir.), no contexto da perspectiva thurstoniana dos testes discriminatórios (adaptado de MEILGAARD *et al.*, 2007, e de LAWLESS, 2013).

¹⁸ O provador desconhece a verdadeira natureza do estímulo, “sinal” ou “ruído”.

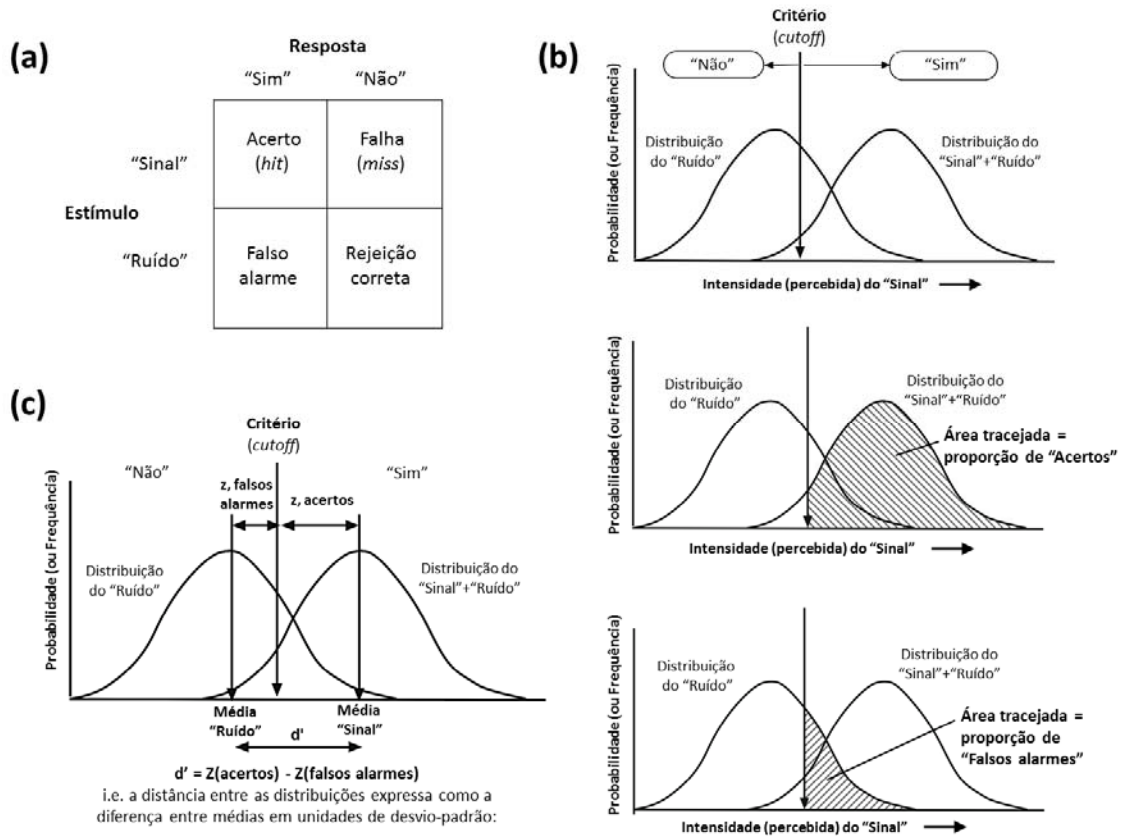


Figura 3 – Ilustração da aplicação da SDT aos testes (sensoriais) discriminatórios: (a) “matriz estímulo-resposta”; (b) estabelecimento de um critério (valor-limite ou *cutoff*) estável; e (c) diferença d' , frequentemente utilizada como estimativa da distância discriminial δ de Thurstone (adaptado de LAWLESS e HEYMANN, 2010).

Conjugando a “matriz estímulo-resposta” e as distribuições de probabilidades de “acertos” e “falsos alarmes”, é possível calcular a diferença d' , frequentemente utilizada como estimativa da distância discriminial δ de Thurstone (em que $\delta = (\mu_1 - \mu_2) / \sigma$) (GACULA Jr. *et al.*, 2009; LAELESS e HEYMANN, 2010):

$$d' = Z(\text{acertos}) - Z(\text{falsos alarmes}).$$

A significância de d' pode ser testada recorrendo, por exemplo, às tabelas publicadas em JI (2006).

O pacote *sensR* (CHRISTENSEN e BROCKHOFF, 2016) do R implementa a perspectiva thurstoniana à análise de resultados de testes (sensoriais) discriminatórios (incl. a possibilidade de calcular a potência dos testes), nomeadamente através da função `discrim()` (cf. BROCKHOFF e CHRISTENSEN, 2010; CHRISTENSEN, 2011).

4. APLICAÇÕES ESTATÍSTICAS EM TESTES DESCRITIVOS

Os testes descritivos utilizam-se para descrever detalhadamente a aparência, a textura, o aroma, e o sabor dos produtos (alimentares), i.e., obter o perfil sensorial, tanto no contexto da I&D como na indústria (MEILGAARD *et al.*, 2007). Os resultados (as cotações, ou *scores*, dos atributos) Y podem entender-se, *estatisticamente*, como (GACULA Jr. *et al.*, 2009):

$$Y = f(\text{aparência, textura, aroma, sabor}) + \text{erro}$$

Os testes descritivos são mais sofisticados e, por isso, mais complexos do que os testes discriminatórios. Para além de envolverem provadores treinados, por ex. seguindo as diretrizes descritas na NP EN ISO 8586:2014 ou por DE VOS (2010), os resultados das provas são, em termos estatísticos, do tipo intervalar, já que foram obtidos através da avaliação quantitativa dos atributos sensoriais dos produtos utilizando escalas.

Simplificando, o processo estímulo-resposta tem duas fases (Figura 4) (LAWLESS e HEYMANN, 2010): o estímulo dá origem a uma sensação por processos psicofísicos (i.e., experiência consciente); essa sensação é, então, traduzida (por processo de julgamento¹⁹) para uma resposta. O processo psicofísico pode ser modificado por processos fisiológicos como sejam a adaptação ou a ampliação/supressão. Por outro lado, o processo de julgamento pode ser alterado por processos cognitivos, como por exemplo, efeitos de contexto, utilização de números aleatórios, ou outros envidamentos da resposta (cf. KEMP *et al.*, 2009; MEILGAARD *et al.*, 2007). Estas duas fases são determinantes para a avaliação sensorial no contexto dos testes descritivos.



Figura 4 – Diagrama ilustrando os fatores envolvidos no processo estímulo-resposta (adaptado de LAWLESS e HEYMANN, 2010)

Nos testes descritivos as escalas utilizadas refletem (muitas vezes numericamente) um juízo acerca da perceção da intensidade de um certo atributo sensorial e são, usualmente, do tipo categorizada/estruturada ou, então, linear (esta última é, aliás, preconizada pelo método QDA[®]) (Figura 5).

¹⁹ Ato ou efeito de julgar ou de emitir um juízo (de valor); Exame, apreciação ou avaliação (Dicionário da Língua Portuguesa com Acordo Ortográfico [em linha]. Porto: Porto Editora, 2003-2016. <http://www.infope-dia.pt/dicionarios/lingua-portuguesa/julgamento> [acesso em 2016-04-16]).

LAWLESS e HEYMANN (2010, cap. 7), STONE e SIDEL (2004, cap 3) e MEILGAARD *et al.* (2007, cap. 5) são referências para este tópico assim como a norma ISO 4121:2003²⁰. Os resultados que se obtêm são do tipo «intervalar» e, por isso, são passíveis de ser analisados com técnicas estatísticas ditas paramétricas (v. seguir).

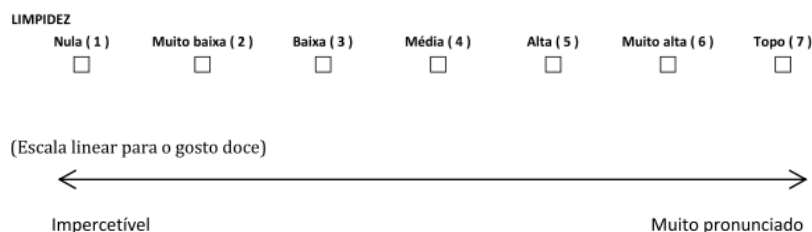


Figura 5 – Ilustração de tipos de escalas que se podem usar em Análise Sensorial.

Habitualmente, nos testes descritivos os provadores avaliam, utilizando um conjunto pré-definido de termos descritivos (ou descritores)²¹, os atributos sensoriais dos produtos alimentares, incl. bebidas. Entre os métodos descritivos mais “famosos” destaco: *Flavour profile*[®] desenvolvido na Arthur D. Little Inc. nos anos 1940 (qualitativo), *Texture profile*[®] desenvolvido no General Food Research Center na década de 1960, *Quantitative Descriptive Analysis* ou QDA[®] (desenvolvido por Herbert Stone na Tragon Corp. nos anos 1970), *Spectrum*[™] (desenvolvido por Gail V. Civille na década de 1990), e *Generic Descriptive Analysis* (*sensu* LAWLESS e HEYMANN, 2010²²) (todos quantitativos). Na literatura internacional anglófona, JACKSON (2014) e LANGSTAFF (2010) descrevem a ficha de Davis (*Modified Davis Score Card for Wine Grading* ou *Davis 20-point scorecard*) como a mais utilizada: um conjunto de características sensoriais é avaliado (por provadores treinados/peritos) com pontos (de mérito) de 0 a 4 que perfazem um total de 20; os intervalos 1–8, 9–12, 13–16 e 17–20 definem as categorias *unacceptable/spoiled*, *below standard*, *standard*, e *superior*, respetivamente. Pelo contrário, PEYNAD e BLOUIN (2013, p. 88) descrevem vários tipos de degustação/análise sensorial de vinhos conforme os objetivos, e.g. para engarrafamento, para inspeção,

²⁰ ISO 4121:2003. Sensory analysis -- Guidelines for the use of quantitative response scales. International Organization for Standardization: Geneva, Switzerland.

²¹ O painel de provadores desenvolve e consensualiza de antemão os descritores e o vocabulário a utilizar no teste sensorial. Noutros testes descritivos, e.g. *free-choice profiling*, os provadores utilizam termos próprios para definir e avaliar os atributos sensoriais (cf. apresentação do Prof. Luís Cunha neste Fórum).

²² As técnicas QDA[®] e o Spectrum[™] têm sido adaptados e utilizados de variados modos, pelo que, apesar de das semelhanças e de se seguirem as diretrizes, não se podem utilizar – com propriedade – os nomes registados (LAWLESS e HEYMANN, 2010).

para concurso, ou para I&D, e que decorrem em locais diferentes ou envolvem “tipos” e número de provadores distintos.

Genericamente, os testes descritivos envolvem 3 etapas (LAWLESS e HEYMANN, 2010): treinar os provadores (relativamente aos descritores e ao modo de avaliação, escalas usadas), determinar a coerência e reprodutibilidade dos provadores durante o treino, e, finalmente, avaliar (e comparar) os produtos. Nesta última parte, concentro-me na análise (estatística) dos resultados da avaliação sensorial através de testes descritivos.

A estrutura dos resultados de testes descritivos pode ser ilustrada como na Figura 6. Os K atributos sensoriais de J amostras (de outros tantos produtos) são avaliados por um painel de I provadores, i.e., $K \times J \times I$ observações. Se, eventualmente, houve oportunidade de repetir provas em mais do que uma sessão S , então os resultados “complicam-se” (no “bom sentido”).

A análise de variância (ANOVA) é a técnica estatística de eleição para muitos autores para analisar os resultados de testes descritivos – e perfis sensoriais – de vários produtos. Neste contexto, a ANOVA permite, numa abordagem atributo-a-atributo, avaliar o desempenho do painel de provadores, por um lado, e identificar os atributos que contribuem para distinguir os produtos em teste, por outro.

Na ANOVA, a variabilidade total dos resultados é subdividida por componentes devidas à influência dos fatores em estudo (por ex., produtos, provadores ou sessões) e ao erro aleatório, a partir das quais se estimam e comparam as respetivas variâncias, σ^2 . A tabela da ANOVA resume a estrutura dos dados assim como a influência relativa das fontes de variação (componentes).

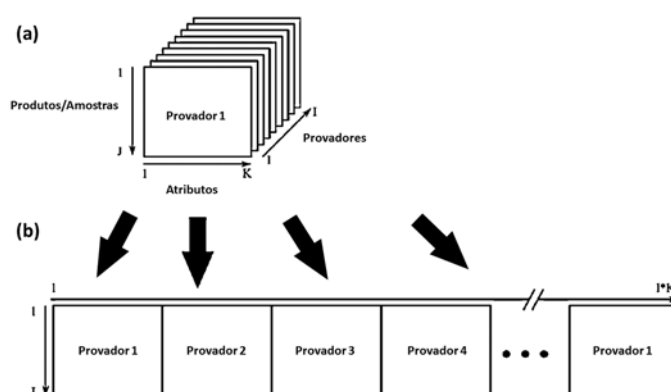


Figura 6 – Diagrama que ilustra (a) a estrutura de dados que se obtém em análise sensorial descritiva. Cada provador, origina resultados (*scores*) para um determinado número de atributos (k) num conjunto de produtos (j). O processo de *unfolding* (b), através do qual os dados são reorganizados, está na base do método Tucker-1 (que serve para elaborar os *Manhattan plots*, v. adiante) (adaptado de NAES *et al.*, 2010)

Para estudar a variabilidade de um dado atributo sensorial Y (variável dependente) entre produtos, avaliado por provadores diferentes em sessões de prova distintas (variáveis independentes ou fatores), o modelo (de efeitos mistos) na ANOVA é (NAES *et al.*, 2010):

$$Y_{ij_s} \sim \mu + \alpha_j + \beta_i + \gamma_s + \alpha\beta_{ij} + \alpha\gamma_{js} + \beta\gamma_{is} + \varepsilon_{ij_s}$$

em que,

α_j é o coeficiente associado efeito do Produto j e indica se os produtos são percebidos como diferentes pelos provadores (no que diz respeito àquele atributo) – na análise do desempenho de um painel corresponde à capacidade discriminatória (*discrimination ability*); por outro lado, permite selecionar quais os atributos que contribuem significativamente para a diferenciação dos produtos e é, neste contexto, uma técnica preparatória de análise estatística subsequente (por ex. PCA; v. seguir);

β_i é o coeficiente associado ao efeito do Provador i , e permite perceber se os provadores estão a usar a escala de forma similar;

γ_s é o coeficiente associado ao efeito da Sessão s , e indica se a escala é usada de forma consistente de sessão para sessão;

$\alpha\beta_{ij}$ é o coeficiente associado à interação Produto-Provador ij , e indica se existe consenso entre provadores na avaliação do atributo sensorial – em termos de avaliação de desempenho do painel corresponde à concordância (*agreement*) entre provadores;

$\alpha\gamma_{js}$ é o coeficiente relativo ao efeito da interação Produto-Sessão js , e indica se, de sessão para sessão, os produtos são percebidos de forma similar – corresponde à repetibilidade em termos de avaliação do desempenho do painel;

$\beta\gamma_{is}$ é o coeficiente associado à interação Provador-Sessão is , e tem a ver com a utilização diferenciada da escala entre sessões de prova; e

ε_{ij_s} refere-se ao erro.

Neste modelo, os erros assumem-se independentes e normalmente distribuídos (com média 0 e variância constante, σ^2). Mais ainda, o efeito *Produto* é considerado *fixo*, i.e., o *interesse* da análise está nos produtos testados *de facto*, enquanto os efeitos *Provador* e *Sessão* são considerados como *aleatórios*, ou seja, não *interessam aqueles* provadores, mas sim o efeito dos provadores *em geral* – daí a designação modelo de efeitos mistos utilizada acima.

Se se verificar que o efeito *Produto* é significativo para um dado atributo (i.e., existem diferenças entre produtos relativamente ao atributo sensorial em estudo), é possível esclarecer as diferenças entre os produtos utilizando testes *post-hoc* de comparações múltiplas, nomeadamente testes de Duncan, de Newman-Keuls, de Tukey, LSD de Fisher, de Dunnett ou de Sheffé (cf. STONE e SIDEL, 2004, ou MEILGAARD *et al.*, 2007). A escolha do teste depende do contexto da comparação pretendida e permanece como motivo de debate e polémica em estatística.

A repetição da ANOVA, atributo-a-atributo, é muito facilitada pela utilização de software, por ex. a função `panelperf()` integrada na *library* *SensoMineR* (LÊ e HUSSON, 2008; HUSSON e LÊ, 2009) do software R (R CORE TEAM, 2016) permite utilizar o modelo acima num contexto de avaliação do desempenho do painel de provadores (Figura 7) de modo similar àquele que segue o software *PanelCheck* (<http://www.panelcheck.com/>), enquanto a função `decat` (da mesma *library* do R) possibilita a análise com enfoque no efeito dos Produtos (i.e., permite perceber quais os atributos que *definem* os produtos e elaborar o respetivo perfil sensorial).

STONE e SIDEL (2004) propõem a representação das médias das avaliações dos atributos, i.e., perfil sensorial, em gráficos de coordenadas polares (Figura 8) para a “comunicação” dos resultados relativos aos produtos.

| | Product | Panelist | Session | Product:Panelist | Product:Session | Panelist:Session |
|----------|-----------|-----------|----------|------------------|-----------------|------------------|
| Heady | 4.308e-31 | 1.703e-15 | 0.6443 | 1.438e-07 | 0.3529 | 0.1002 |
| Greedy | 1.147e-30 | 6.125e-05 | 0.4906 | 2.958e-07 | 0.6149 | 0.0009707 |
| Vanilla | 9.031e-23 | 8.553e-10 | 0.2305 | 7.708e-07 | 0.5682 | 0.02814 |
| Spicy | 1.996e-22 | 1.648e-12 | 0.8911 | 0.0000576 | 0.2128 | 0.1516 |
| Floral | 1.076e-19 | 4.325e-16 | 0.02687 | 0.0003138 | 0.0375 | 0.1845 |
| Fruity | 2.769e-13 | 3.704e-09 | 0.9368 | 8.842e-05 | 0.9033 | 0.7593 |
| Oriental | 2.682e-12 | 5.35e-18 | 0.1423 | 1.474e-12 | 0.9806 | 0.0009239 |
| Wrapping | 4.459e-11 | 3.472e-17 | 0.007783 | 1.795e-07 | 0.4035 | 0.4648 |
| Green | 2.586e-09 | 8.403e-09 | 0.06641 | 0.0001503 | 0.8727 | 0.7347 |
| Woody | 1.623e-06 | 6.315e-11 | 0.3825 | 0.03666 | 0.5552 | 0.05019 |
| Marine | 0.01519 | 2.843e-05 | 0.1156 | 0.5438 | 0.01695 | 2.458e-06 |
| Citrus | 0.08308 | 1.799e-09 | 0.1557 | 2.213e-06 | 0.5643 | 0.5178 |

Figura 7 – Avaliação do desempenho do painel (através da função `panelperf()`) ilustrada numa tabela (obtida com a função `coltable()`) na qual os efeitos significativos da ANOVA (com modelo de efeitos mistos), atributo-a-atributo (Heady, Greedy, Vanilla, etc.), estão coloridos a cinzento (de LÊ e WORCH, 2015).

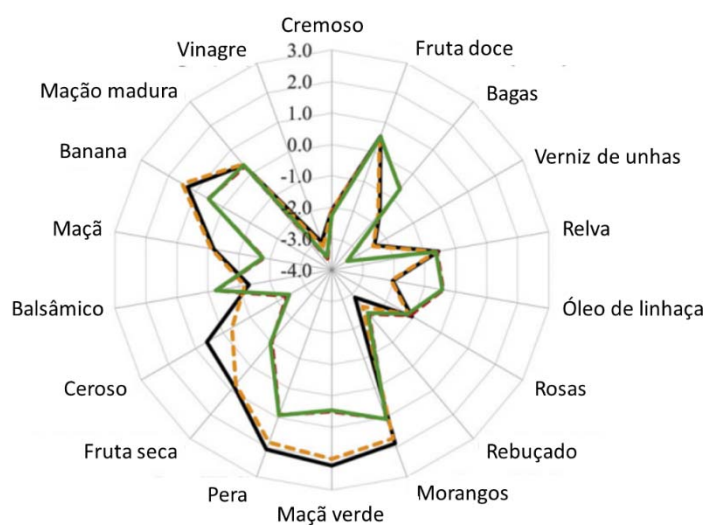


Figura 8 – Perfis sensoriais de vinhos Chardonnay (adaptado de <http://wine.landfood.ubc.ca/files/2013/03/FoodResIntl2012.pdf>).

Os resultados de testes descritivos são inerentemente multivariados (ou multivariáveis) porque se avaliam várias características nos diversos produtos (BROKHOFF, 2011). Do conjunto de técnicas estatísticas multivariadas (ou multivariáveis) disponíveis, a análise de componentes principais (*principal component analysis* ou PCA, em inglês) é a metodologia habitualmente utilizada para simplificar e descrever as relações entre múltiplas variáveis (neste caso, os atributos sensoriais/descriptores) e em vários indivíduos/objetos (na análise sensorial, os produtos) (ANDERSON, 2003). É uma das técnicas estatísticas elementares em Quimiometria (VARMUZA e FILSMOSER, 2009).

A PCA é usualmente aplicada à matriz produtos-atributos X (na prática aos perfis sensoriais), na qual os dados representam médias por provador e sessão. Com a PCA, transformam-se as variáveis originais (i.e., os atributos sensoriais), que se supõem correlacionadas, num conjunto menor de “novas” *variáveis* (latentes, não-observadas) independentes, designados componentes principais (PC) p , que ainda assim representam a maior parte da informação presente nas variáveis originais (LAWLESS e HEYMANN, 2010; MEILGAARD *et al.*, 2007; NAES *et al.*, 2010). É, nestes termos, um método de redução de complexidade dos dados (MAROCO, 2003). As componentes correspondem a combinações lineares das m variáveis, que maximizam a variância total das variáveis; por exemplo para o PC1:

$$p_1 = p_{11}x_1 + p_{12}x_2 + \dots + p_{1m}x_m$$

os coeficientes p_{ij} , designados pesos (ou *factor loadings*), «medem» a importância relativa das variáveis originais nesse PC. Como se fossem coeficientes de correlação, tomam valores entre -1 e +1 e quanto mais próximo desses “limites” maior a influência no valor desse componente principal. As componentes, extraídas sucessivamente, explicam cada vez menor proporção da variância dos dados e são ortogonais (ou não-correlacionadas) entre si. A extração dessas componentes (cuja *matemática* não pormenorizamos aqui, mas está associada à estimação dos pesos de modo a cumprir certas condições; cf. ANDERSON, 2003; HUSSON *et al.*, 2011; NAES *et al.*, 2010) baseia-se, usualmente, na matriz de covariâncias²³ ou de correlações²⁴ das variáveis (obtida a partir dos dados *centrados* e, muitas vezes, *standardizados*).

O número de componentes a reter (no limite, tantas quantas as variáveis) é decidido usando critérios relativamente concretos, nomeadamente se a variância (explicada) pela componente (ou *eigenvalue*) é maior do que 1 (critério de Kaiser) e a análise do *scree plot*, para além do bom senso e interpretabilidade (LAWLESS e HEYMANN, 2010).

Os indivíduos/objetos (i.e., os produtos) tomam valores, chamados *object scores*, nestas novas *variáveis* (nas PC) como acontecia nas variáveis originais. Esses valores (de facto, as coordenadas resultantes das projeções ortogonais dos objetos sobre (o vetor) componente; cf. Fig. 3.2 de VARMUZA e FILSMOZER, 2009) T obtêm-se a partir dos dados originais X e da matriz P que inclui as componentes extraídas (Figura 9).

Os resultados da PCA são apresentados graficamente através de *loadings plots* (para as variáveis/atributos) e de *scores plot* (para os objetos/produtos), em separado ou em conjunto (num *biplot*). Estas representações, entre outras derivadas de outros métodos estatísticos multivariados (por ex., escalonamento multidimensional ou *multidimensional scaling*), são classificadas como *perceptual mapping*²⁵ por GACULA Jr *et al.* (2009) e LAWLESS (2013). No fundo são, nas palavras de LAWLESS (2013), um exercício de redução (da complexidade) de dados e constituem um tópico ativo de investigação em Sensometria.

É possível representar, junto com os *loadings* e *scores plots*, outras variáveis (de interesse), categóricas ou contínuas, para além daquelas utilizadas para a extração das componentes na PCA, *variáveis suplementares ou informativas* (GERRARD e JOHNSON, 2015; HUSSON *et al.* 2011; LÊ e WORCH,

²³ Quando as variáveis *centradas* foram medidas na mesma escala e *interessam* as diferenças nas variâncias entre variáveis.

²⁴ Quando as variáveis (*centradas* e *standardizadas*) foram medidas em escalas diferentes e não *interessam* as diferenças nas variâncias entre variáveis.

²⁵ Se os resultados têm a ver com preferência e derivam de testes afetivos/hedónicos, os *mapas* designam-se *preference maps* (LAWLESS, 2013).

2015), por forma a completar a interpretação dos resultados (Figura 10a,b). O pacote FactoMineR (HUSSON *et al.*, 2016; LÊ *et al.*, 2008) para o R permite esta abordagem de forma bastante simples (por via dos parâmetros `quali.sup` e `quantif.sup` da função `PCA()`).

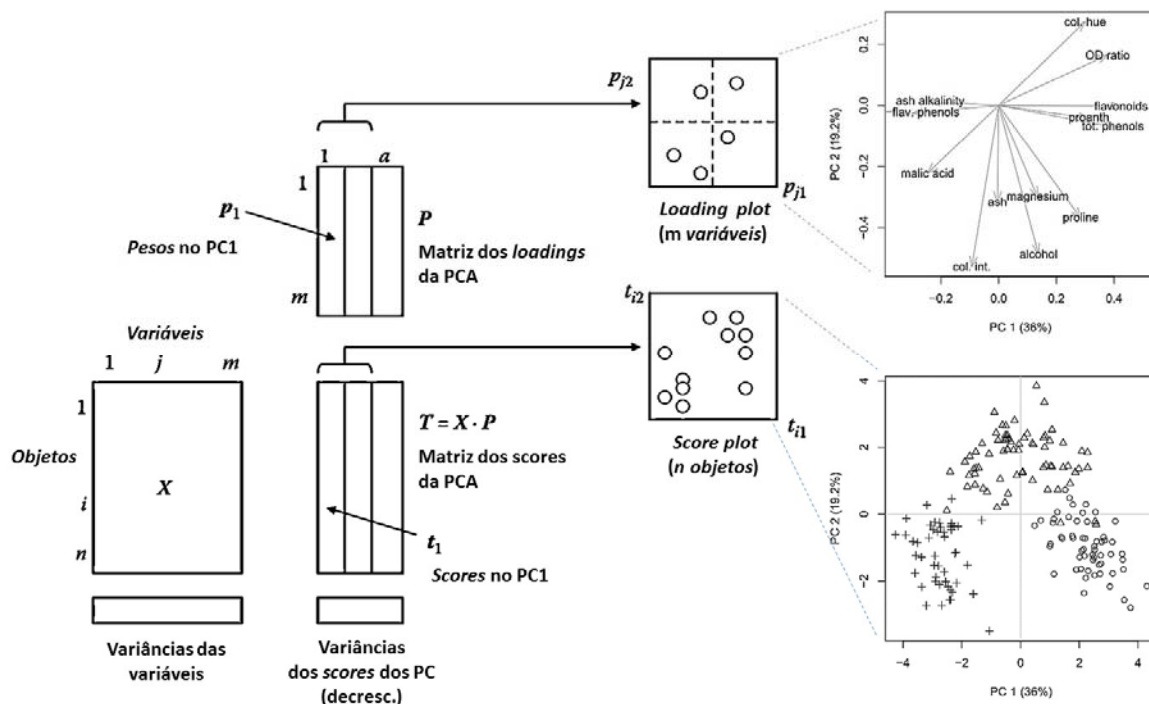


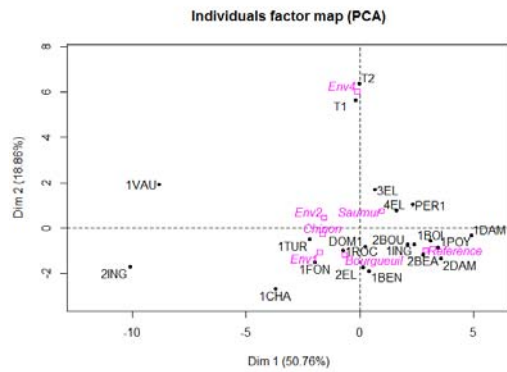
Figura 9 – Diagrama da PCA (adaptado de VARMUZA e FILSMOZER, 2009, e de WEHRENS, 2011).

Aqueles autores também apresentam a possibilidade de, por um lado, desenhar elipses de confiança²⁶ em torno dos produtos no *scores plot* e, dessa forma, esclarecer as diferenças entre os produtos e, por outro, elucidar a variação dos efeitos das variáveis (i.e., dos atributos sensoriais) representando simultaneamente as componentes principais e os “dados” no *loadings plot* (Figura 10c). As elipses no *scores plot* resultam de um processo de reamostragem (ou *bootstrapping*²⁷) dos dados que permite obter painéis *virtuais* de provedores a partir dos quais se realiza a PCA e obtêm as elipses. Os “dados” originados pelo *bootstrapping* são representados juntamente conjunto as variáveis no *loadings plot* e permitem perceber a *estabilidade* da variável (Figura 10d). O pacote SensoMineR (HUSSON e LÊ, 2009; LÊ e HUSSON, 2008) inclui a função `panellipse()` que permite esta análise.

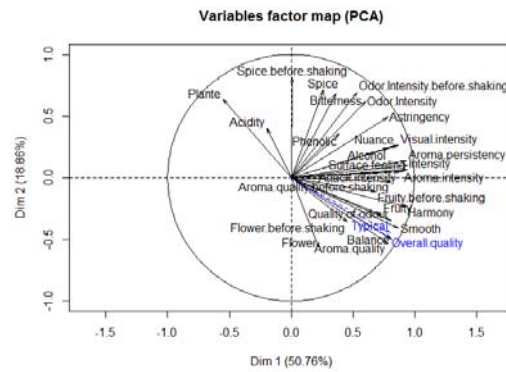
²⁶ Estão para dados relativos a duas variáveis (bidimensionais) como os intervalos de confiança estão para os dados relativos a uma única variável (unidimensionais).

²⁷ [https://pt.wikipedia.org/wiki/Bootstrapping_\(estatística\)](https://pt.wikipedia.org/wiki/Bootstrapping_(estatística)) ou [https://en.wikipedia.org/wiki/Bootstrapping_\(statistics\)](https://en.wikipedia.org/wiki/Bootstrapping_(statistics))

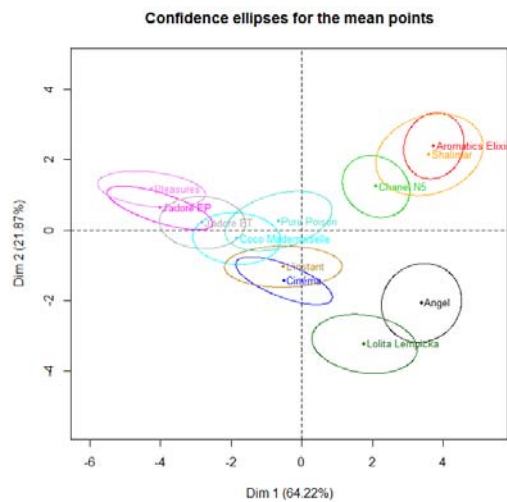
(a)



(b)



(c)



(d)

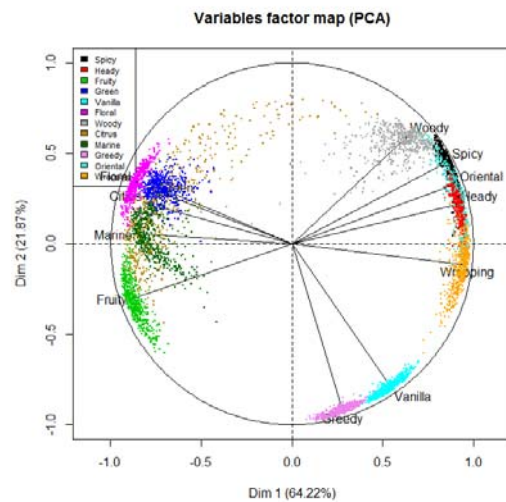


Figura 10 – Representações gráficas resultantes de PCA. *Scores plot* (a) e *loadings plot* (b) relativos às duas primeiras componentes principais (Dim 1 e Dim 2) extraídas dos dados da análise sensorial de uma série de atributos (plante, acidity, spice.before.shaking, etc.) de vinhos (1VAU, 2ING, T2, T1, etc.) incluindo variáveis suplementares (Env4, Env2, Saumur, etc., à esq. e Typical e Overall.quality, à dir.). *Scores plot* (c) com elipses de confiança para facilitar a discriminação dos produtos e *loadings plot* (d) incluindo (as projeções) dos dados originados por reamostragem (dados da análise sensorial de um conjunto de fragrâncias).

Se, porventura, as características sensoriais de produtos alimentares, por ex. bebidas, foram avaliadas utilizando os termos dos próprios provadores, como acontece no caso do *free-choice profiling*, a Análise Generalizada de Procrustes²⁸ (ou GPA na sigla em Inglês) desenvolvida por GOWER (1975) permite *consensualizar* os resultados das matrizes de dados individuais através de uma série de transformações (i.e., translação, rotação, reflexão, e reescalonamento isotópico) (Figura 11). LÊ e

²⁸ *Generalized Procrustes Analysis* (ou GPA) que deve o nome ao mito grego de Procusto que esticava ou amputava os convidados para que se conformassem com o tamanho da sua cama de ferro.

WORCH (2015) advogam e apresentam uma metodologia de análise estatística alternativa, a *multiple factor analysis* (MFA), para analisar os resultados e que permite considerar variáveis quantitativas e/ou qualitativas.

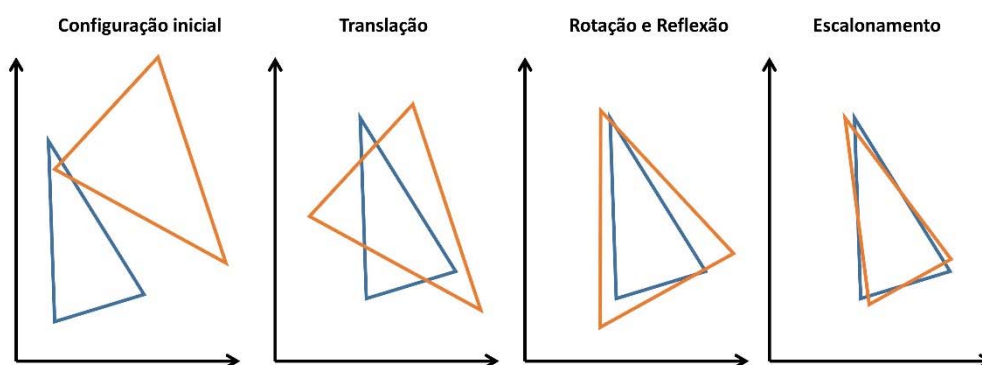


Figura 11 – Representação das etapas da GPA (adaptada de SensorySociety.org²⁹).

Finalmente, é possível relacionar os resultados, mais objetivos digamos assim, obtidos com recurso a um painel de provadores treinados com aqueles que resultam de provas afetivas/hedónicas com consumidores – designada *external preference mapping* (GACULA Jr. *et al.*, 2009; LAWLESS, 2013). No caso de apenas se analisarem os dados obtidos com base em consumidores, então estaremos a falar de *internal preference mapping*. A análise dos resultados recorre à PCA (v. acima) com os produtos (pelo menos 6) como *objetos* (linhas) e as cotações (hedónicas) dos consumidores como *variáveis* (colunas) (LAWLESS e HEYMANN, 2010). MEULLENET *et al.* (2007, cap. 3) providenciam uma descrição não-técnica de *preference mapping*. VARELA e ARES (2012) reviram e DEHLHOLM *et al.* (2012) compararam um conjunto de metodologias relativamente recentes para a caracterização sensorial de produtos alimentares que são mais rápidas e flexíveis, e que recorrem a provadores semi-treinados ou até consumidores, e que providenciam *mapas sensoriais* muito aproximados àqueles que se obtêm através de análise descritiva “clássica”, e.g. *check-all-that-apply* (CATA) *questions*, *flash profiling*, *projective mapping* ou *Napping*[®]. TUORILA e MONTELEONE (2009) apresentam e refletem sobre os avanços recentes na análise sensorial de produtos alimentares e bebidas, nomeadamente o vinho.

²⁹ <http://www.sensorysociety.org/knowledge/sspwiki/Pages/Generalized%20Procrustes%20Analysis.aspx>

5. QUALIDADE DOS DADOS

A utilização correta de métodos estatísticos adequados permite extrair o máximo de informação dos dados experimentais (e.g. provas sensoriais) – que se admitem de «boa» qualidade, i.e. são relativamente precisos e resultam de amostras aleatórias de tamanho adequado ou em número suficiente – e elaborar conclusões válidas. Duas características dos dados (obtidos em análise sensorial) são (muito) importantes para a aplicação de testes estatísticos (de inferência): a «normalidade», ou seja, os dados distribuem-se de acordo com a curva normal; e a homocedasticidade, i.e., a variação (de facto, a variância) dos dados é uniforme. O primeiro pressuposto, a «normalidade» dos dados, é importante para testes estatísticos paramétricos como sejam a análise de variância (ANOVA), análise de correlação e regressão (simples e múltipla), ou testes de t. É possível testar este pressuposto graficamente, usando por ex. gráficos de probabilidades normais (normal probability plots), ou através de testes estatísticos, por ex. teste de Kolmogorov-Smirnov/Lilliefors ou de Shapiro-Wilk. Quanto ao segundo, a homogeneidade das variâncias, é um importante pressuposto nos métodos em que se comparam médias (por ex. ANOVA) ou naqueles em que modela (matematicamente) a relação entre os dados (e.g. regressão). Pode ser verificada graficamente ou por via de testes estatísticos como sejam, por ex., o teste de Levene ou de Bartlett. Em análise sensorial é relativamente comum que os dados não se conformem com distribuição normal e/ou as variâncias são heterogéneas. A transformação dos dados, por via de funções matemáticas e.g. $\ln(x)$ ou $x^{1/2}$, pode resolver aqueles *problemas*; em alternativa, podem utilizar-se métodos não-paramétricos (cf. GRANATO *et al.* 2014; RAYNER *et al.*, 2005).

Em análise sensorial e ciência do consumo, a obtenção de informação que permita identificar quais os potenciais fatores que, influenciando os atributos sensoriais dos produtos, determinam a preferência dos consumidores está intimamente ligada com uma área da estatística, o planeamento experimental (cf. MONTGOMERY, 2013). STONE e SIDEL (2004) examinam os aspetos mais relevantes para o planeamento experimental e análise dos resultados de provas sensoriais, dos quais destaco aqui: a fiabilidade (*reliability* no original em Inglês) e a validade; a replicação; (in)dependência dos resultados; e os riscos envolvidos no processo de decisão.

A fiabilidade é a medida em que os sujeitos fornecem respostas semelhantes para produtos similares em ocasiões repetidas enquanto a validade tem a ver com consistência entre resultados e os factos (a “realidade”). A replicação, i.e., a obtenção de réplicas dos resultados para cada provador, é essencial até para avaliar a consistência dos provadores (STONE e SIDEL, 2004). NAES *et al.* (2010) consideram que para que os dados sejam válidos é importante que os provadores estejam «calibrados» e usem os atributos sensoriais do mesmo modo (cf. AMERINE *et al.*, 1965; MEILGAARD *et*

al., 2007 cap. 4; STONE e SIDEL, 2004 cap. 4; LAWLESS e HEYMANN, 2010 cap. 3). As principais diferenças individuais com implicações na análise sensorial são (NAES *et al.*, 2010):

1. Utilização da escala: diferenças na média e variabilidade/amplitude dos resultados que têm a ver com o modo distinto como os provadores utilizam a escala;
2. Concordância/Reprodutibilidade: discordância entre provadores na avaliação de um atributo num dado produto;
3. Repetibilidade: diferentes níveis de precisão, i.e., diferenças na avaliação entre réplicas independentes;
4. Discriminação: diferenças na capacidade para diferenciar produtos.

O conceito (estatístico, literal) de independência (dos resultados) não é realístico (mais que não seja porque muitas vezes se exige aos provadores contacto/consumo prévio do(s) produtos em teste). Todavia, a apresentação equilibrada (*sensu* balanced) das amostras, nomeadamente a apresentação monádica sequencial, minimiza o efeito da dependência. De qualquer forma, não é completamente clara a influência da (inter)dependência dos resultados nos resultados e conclusões (elaboradas, por exemplo, a partir de testes de t ou ANOVA). Finalmente, as conclusões que se obtêm não estão isentas de erro (em virtude dos pressupostos e técnicas usadas, nomeadamente a utilização de amostras e respetivas estatísticas). Distinguem-se dois tipos de erro na inferência estatística: de tipo I e de tipo II (Quadro 1). Enquanto o erro de tipo I é controlado diretamente (através da definição *a priori* de α), para um dado n , α é inversamente proporcional a β . Para reduzir ambos os erros, é necessário aumentar n . A quantidade $(1 - \beta)$ é designada “poder estatístico” do teste.

Quadro 1 – Erros na inferência estatística.

| Conclusões elaboradas a partir do teste de hipóteses (Amostra) | Alternativas verdadeiras (População/"Realidade") | |
|--|--|---------------------------------------|
| | H ₀ | H ₁ |
| H ₀ | Conclusão correta | Erro Tipo II Probabilidade β |
| H ₁ | Erro Tipo I Probabilidade α | Conclusão correta |

O controlo da qualidade dos dados pode seguir, com a flexibilidade necessária, fluxograma proposto por TOMIC *et al.* (2007, 2009) e reproduzido por NAES *et al.* (2010, fig. 3.2). Mais uma vez, a relativa complexidade de algumas técnicas e a repetitividade de outras podem ser contornadas usando software, por exemplo PanelCheck (www.panelcheck.com) – que «segue» o fluxograma

referido acima – ou as funções integradas no SensoMineR para R – por ex. `panelperf()`. A análise de variância (ANOVA) é o método de eleição (e não só para a avaliar o desempenho do painel como dos provadores) em conjunto com a análise de componentes principais (PCA na sigla em inglês). Aliás das várias metodologias, eminentemente gráficas, derivadas da ANOVA ou da PCA e que se podem utilizar neste contexto, destaco aqui os *p*MSE plots* (que derivam de resultado da ANOVA) e os *Manhattan plots* (que se obtêm a partir dos resultados de PCA). Complementarmente, na norma internacional ISO 11132:2012³⁰, estabelece-se a metodologia para a monitorização do desempenho de um painel sensorial (de índole quantitativa).

Os *p*MSE plots* obtêm-se a partir dos valores-p (probabilidade de significância) associados ao efeito *Produto* (β_{ij}) e das somas dos quadrados dos erros (MSE, em Inglês) de ANOVAS «simples» realizadas para todas as combinações provador-atributo *ik* e cujo modelo dos efeitos é (NAES *et al.*, 2010):

$$y_{ijk} = \mu_{ir} + \beta_{ijk} + \varepsilon_{ijk}$$

em que β_{ij} é o efeito *Produto* (e cujo valor-p se utiliza) e ε_{ijk} é o erro (para o qual se estima a MSE). Os valores-p são representados no eixo das ordenadas e as estimativas de MSE no eixo das abcissas (Figura 12a). Cada ponto corresponde a uma combinação provador-atributo, num total de $i*k$ (por ex., para 9 provadores e 20 atributos, o *p*MSE plot* inclui 180 pontos)³¹. Provadores cujos resultados estejam próximos da origem do gráfico são considerados «melhores».

Os *Manhattan plot* podem utilizar-se para visualizar as variâncias explicadas para combinação provador-atributo sensorial obtidas através do método Tucker-1 (cf. Figura 6) ou da PCA aplicada aos dados de cada atributo individualmente (Figura 12b).

A análise do desempenho dos provadores, quer sejam experientes, treinados, peritos ou «simples» consumidores, é essencial, mas não é suficiente (HYLDIG, 2010) para avaliar o desempenho de um painel de provadores. É, por isso, importante que os painéis de provadores participem em ensaios interlaboratoriais (aliás como acontece com laboratórios [acreditados] de análises químicas e/ou microbiológicas) e está a acontecer na ALABE. Existe uma panóplia de normas aplicáveis, e.g. ISO 17025:2005³², ISO/IEC 17043:2010³³ e EA-4/09³⁴ (HYLDIG, 2010).

³⁰ ISO 11132:2012 Sensory analysis -- Methodology -- Guidelines for monitoring the performance of a quantitative sensory panel. International Organization for Standardization: Geneva, Switzerland.

³¹ De acordo com NAES *et al.* (2010), esta abordagem requiere 2-3 repetições de avaliações/réplicas por combinação provador-atributo.

³² cf. <http://www.ipac.pt/docs/publicdocs/regras/OGC001.pdf>.

³³ ISO/IEC 17043:2010 Conformity assessment -- General requirements for proficiency testing. International Organization for Standardization: Geneva, Switzerland.

³⁴ v. <http://www.european-accreditation.org/publication/ea-4-09-g-rev01-july-2003>

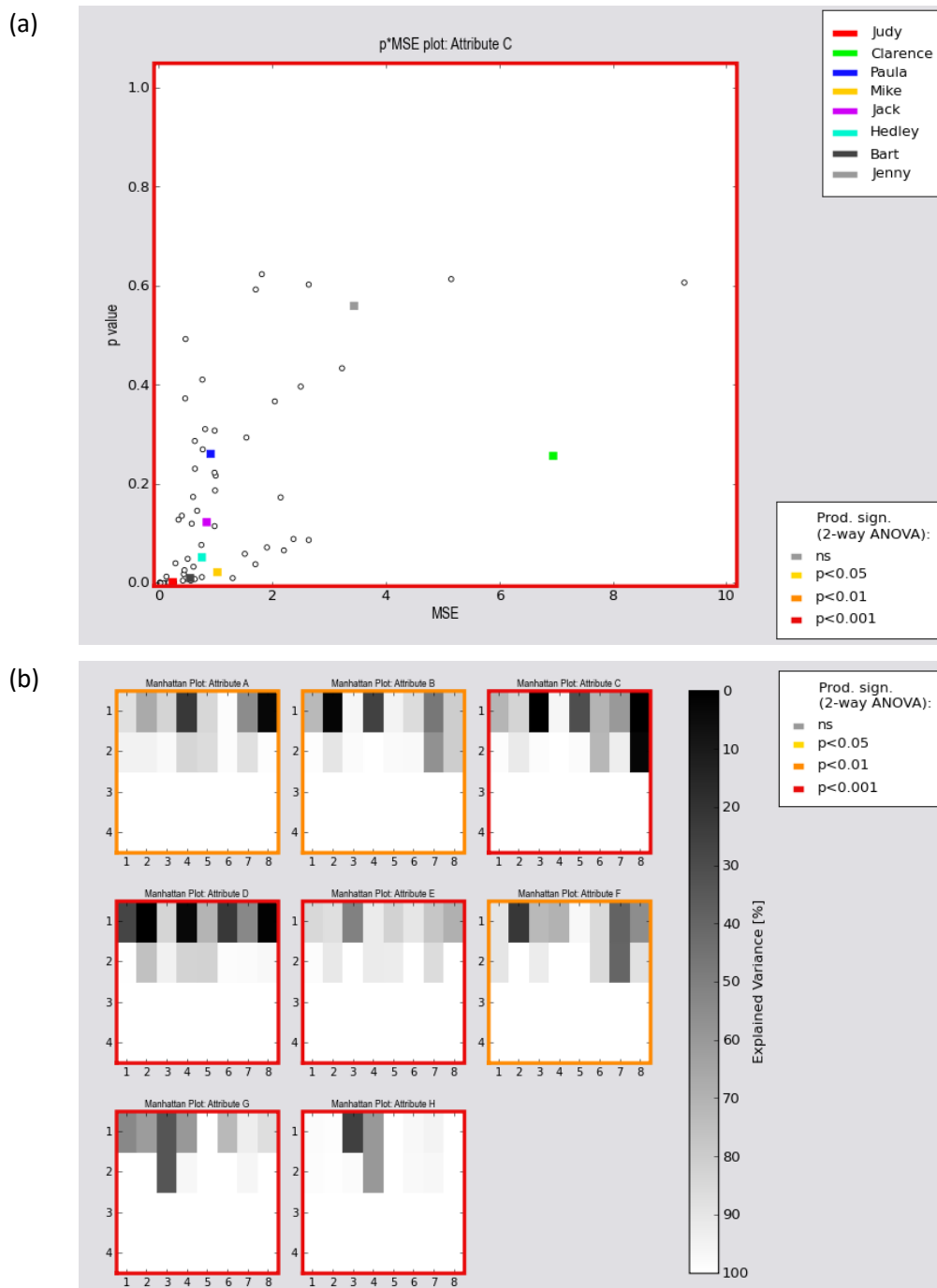


Figura – (a) p *MSE plot para um atributo C relativo a 8 provedores (Judy, Clarence, Paula, etc.) e (b) *Manhattan plots* relativo a 8 atributos (A–H), 8 provedores (1–8 no eixo XX) e 2 PC (eixo YY). A cor da margem dos gráficos, em (a) e em (b), reflete a significância do efeito *Produto* numa ANOVA com fatores *Produto* e *Provedores*. (<http://www.panelcheck.com/Home/softwarefeatures>).

6. CONSIDERAÇÕES FINAIS

Neste artigo tentei, seguindo a *lógica* de BROCKHOFF (2011), apresentar uma “outra face” da análise sensorial nomeadamente aquela que diz respeito ao desenvolvimento e aplicações de métodos estatísticos à análise de dados, a *Sensometria*.

Para além das questões práticas, laboratoriais essenciais para a realização das provas sensoriais, a proficiência com algumas técnicas estatísticas – principalmente em termos de análise e interpretação dos resultados – é cada vez mais importante dada a crescente complexidade dos *problemas*. Ainda que a análise dos resultados de alguns/bastantes testes sensoriais seja possível de realizar usando funções e ferramentas disponíveis em folhas de cálculo «correntes», por ex. Microsoft Excel™; noutros casos, a exaustividade e/ou complexidade dos cálculos «exige» a utilização de software de análise estatística (SPSS®, minitab®, Unscrambler® X, R, etc.) ou software «dedicado» à análise de resultados no contexto da análise sensorial (PanelCheck/ConsumerCheck³⁵, XLSTAT-Sensory³⁶ para MS Excel®, FIZZ by Biosystemes³⁷, Compusense³⁸, SensCheck™ Sensory Analysis System by Aroxa³⁹, RedJade⁴⁰, sensR/SensoMineR/FactoMineR para R). Neste artigo (e na comunicação complementar), menciono e apresento algumas aplicações usando principalmente o R, por ser uma linguagem de programação e um sistema (gratuito) que integra um conjunto de ferramentas computacionais que permitem a manipulação e análise de dados, o cálculo numérico e a produção de gráficos de qualidade, cujas qualidades matemático-estatísticas e computacionais são unanimemente reconhecidas (v. <https://www.r-project.org/>).

REFERÊNCIAS BIBLIOGRÁFICAS

- Amerine, M., Pangborn, R., Roessler, E. (1965). *Principles of sensory evaluation of food*. New York: Academic Press.
- Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd Edition. Chichester, UK: Wiley-Interscience.
- Bi, J. (2006). *Sensory Discrimination Tests and Measurements: Sensometrics in Sensory Evaluation*. Chichester, UK: John Wiley & Sons, Ltd.
- Brockhoff, P. (2011). Sensometrics. *International Encyclopedia of Statistical Science*: 1302-1305.
- Brockhoff, P., Christensen, R. (2010). Thurstonian models for sensory discrimination tests as generalized linear models. *Food Quality and Preference*, 21(3): 330-338.
- CAMO, (2016). *Sensometrics. Design of Experiments (DoE). Chemometrics*. [online] Camo.com. Available at: http://www.camo.com/rt/Products/Multivariate/application_consumer.html [Acesso em 15/4/2016].
- Christensen, R.H.B. (2011). Statistical methodology for sensory discrimination tests and its implementation in sensR. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.192.6348>.
- Christensen, R.H.B., Brockhoff, P.B. (2016). sensR - An R-package for sensory discrimination. R package version 1.4-7. <http://www.cran.r-project.org/package=sensR/>.

³⁵ <http://www.panelcheck.com/>

³⁶ <https://www.xlstat.com/en/solutions/sensory>

³⁷ <http://www.biosystemes.com/fizz.php>

³⁸ <https://www.compusense.com/en/>

³⁹ <http://www.aroxa.com/about-sensory-software/>

⁴⁰ <http://www.redjade.net/>

- De Vos, E. (2010). Selection and management of staff for sensory quality control. In Kilcast, D. (Ed.). *Sensory Analysis for Food and Beverage Quality Control*. Boca Raton: CRC Press.
- Dehlholm, C., Brockhoff, P., Meinert, L., Aaslyng, M.; Bredie, W. (2012). Rapid descriptive sensory methods – Comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference*, 26(2): 267-277.
- Esteves, E. (2011). Statistical Analysis in Food Science. In Cruz, R.M.S. (Ed.). *Practical food and research*. N.Y.: Nova Science Publishers.
- Esteves, E. (2016). *Apontamentos de Análise Sensorial*. Faro: Departamento de Engenharia Alimentar, Instituto Superior de Engenharia, Universidade do Algarve.
- Findlay C. (2010). Statistical approaches to sensory quality control. In Kilcast, D. (Ed.) *Sensory Analysis for Food and Beverage Quality Control*. Boca Raton: CRC Press.
- Gacula Jr., M., Singh, J., Bi, J., Altan, S. (2009). *Statistical methods in food and consumer research*. Amsterdam: Elsevier/Academic Press.
- Gerrard, P., Johnson, R. (2015). *Mastering scientific computing with R*. Birmingham, UK: Packt Publishing Ltd.
- Gower, J.C. (1975). Generalized procrustes analysis. *Psychometrika*, 40: 33-51.
- Granato, D., de Araújo Calado, V., Jarvis, B. (2014). Observations on the use of statistical methods in Food Science and Technology. *Food Research International*, 55: 137-149.
- Husson, F., Josse J., Lê, S., Mazet, J. (2016). FactoMineR: Factor Analysis and Data Mining with R. R package version 1.04, URL <http://CRAN.R-project.org/package=FactoMineR>.
- Husson, F., Lê, S. (2009). SensoMineR: Sensory data analysis with R. R package version 1.10. <http://CRAN.R-project.org/package=SensoMineR>
- Husson, F., Lê, S., Pagès, J. (2011). *Exploratory multivariate analysis by example using R*. Boca Raton: CRC Press.
- Hyldig, G. (2010) Proficiency testing of sensory panels. In Kilcast, D. (Ed.). *Sensory Analysis for Food and Beverage Quality Control*. Boca Raton: CRC Press.
- Jackson, R. (2014). *Wine Science. Principles and Applications*. 4th ed. USA: Academic Press.
- Kemp, S., Hollowood, T., Hort, J. (2009). *Sensory evaluation*. Chichester, U.K.: Ames, Iowa.
- Langstaff, S.A. (2010). Sensory quality control in the wine industry. In Kilcast, D. (Ed.). *Sensory Analysis for Food and Beverage Quality Control*. Boca Raton: CRC Press.
- Lawless, H., Heymann, H. (2010). *Sensory evaluation of food*. New York: Springer.
- Lawless, H.T. (2013). *Quantitative Sensory Analysis. Psychophysics, Models and Intelligent Design*. Chichester, West Sussex, UK: Wiley-Blackwell/John Wiley & Sons, Ltd.
- Lê, S., Husson, F. (2008). SensoMineR: a package for sensory data analysis. *Journal of Sensory Studies*, 23 (1): 14-25.
- Lê, S., Josse, J., Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1): 1-18.
- Lê, S., Worch, T. (2015). *Analyzing sensory data with R*. Boca Raton: CRC Press.
- Maroco, J. (2003) *Análise estatística. Com utilização do SPSS*. 2ª edição, Lisboa; Edições Sílabo.
- Meilgaard, M., Carr, B., Civille, G. (2007). *Sensory Evaluation Techniques*, Fourth Edition. Hoboken: CRC Press.

- Meullenet, J., Xiong, R., Findlay, C. (2007). *Multivariate and probalistic analyses of sensory science problems*. Oxford: Blackwell.
- Montgomery, D. (2013). *Design and analysis of experiments*. Hoboken, NJ: John Wiley & Sons, Inc.
- Næs, T., Brockhoff, P., Tomić, O. (2010). *Statistics for sensory and consumer science*. Chichester, West Sussex: Wiley.
- Peynaud, E., Blouin, J. (2013). *Le goût du vin: Le grand livre de la degustation*. 5e éd., Paris: Dunod.
- R Core Team (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing (<https://www.R-project.org>)
- Rayner, J., Best, D., Brockhoff, P., Rayner, G. (2005). *Nonparametrics for sensory science. A more informative approach*. Ames, Iowa: Blackwell Publ.
- Stone, H., Sidel, J. (2004). *Sensory evaluation practices*. USA: Elsevier Academic Press.
- Thurstone, L.L. (1927). A law of comparative judgement. *Psychological Review*, 34, 273-286.
- Thurstone, L.L. (1959). *The Measurement of Values*. Chicago: The University of Chicago Press.
- Tomasco Varela, P., Naes, T. (2016). *Sensometrics*. [online] Nofima. Available at: <http://nofima.no/en/forskningssomrade/consumer-and-sensory-sciences/sensometrics/> [acesso em 15/4/2016].
- Tomic, O., Luciano, G., Nilsen, A., Hyldig, G., Lorensen, K., Næs, T. (2010). Analysing sensory panel performance in proficiency tests using the PanelCheck software. *European Food Research and Technology*, 230: 497-511
- Tomic, O., Nilsen, A. N., Martens, M., Næs, T. (2007). Visualization of sensory profiling data for performance monitoring. *LWT - Food Science and Technology*, 40: 262-269
- Tuorila, H., Monteleone, E. (2009). Sensory food science in the changing society: Opportunities, needs, and challenges. *Trends in Food Science & Technology*, 20(2): 54-62.
- Van der Kellen, D., Nunes, L., Garcia-Marques, L. (2008). Sensibilidade e bom senso: Princípios fundamentais da teoria de detecção de sinal na investigação em Psicologia. *Laboratório de Psicologia*, 6(1): 75-91
- Varela, P., Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International*, 48(2): 893-908.
- Varmuza, K., Filzmoser, P. (2009). *Introduction to multivariate statistical analysis in chemometrics*. Boca Raton: CRC Press.
- Vining, G., Kowalski, S. (2011). *Statistical methods for engineers*. Boston, Mass.: Cengage Learning.
- Wehrens, R. (2011). *Chemometrics with R*. Heidelberg: Springer.