

TCVN

TIÊU CHUẨN QUỐC GIA

TCVN 12387:2018

ISO 6658:2017

Xuất bản lần 1

**PHÂN TÍCH CẢM QUAN –
PHƯƠNG PHÁP LUẬN – HƯỚNG DẪN CHUNG**

Sensory analysis – Methodology – General guidance

HÀ NỘI - 2018

Lời nói đầu

TCVN 12387:2018 hoàn toàn tương đương với ISO 6658:2017;

TCVN 12387:2018 do Ban kỹ thuật tiêu chuẩn quốc gia TCVN/TC/F13
Phương pháp phân tích và lấy mẫu biên soạn, Tổng cục Tiêu chuẩn
Đo lường Chất lượng thẩm định, Bộ Khoa học và Công nghệ công bố.

Lời giới thiệu

Tiêu chuẩn này bao gồm phần giới thiệu chung về phương pháp luận của phân tích cảm quan và phải được đọc trước khi thực hiện các quy trình thử chi tiết hơn được đề cập trong các tiêu chuẩn khác. Tiêu chuẩn này bao gồm các vấn đề chung của phương pháp luận và có nhiệm vụ sau:

- cung cấp những kiến thức cơ bản của các tính năng thiết yếu của phương pháp phân tích cảm quan cho người sử dụng các phép thử cụ thể;
- cung cấp chi tiết các yêu cầu chung, quy trình chung và diễn giải kết quả chung cho tất cả hoặc hầu hết các phép thử;
- cung cấp hướng dẫn đầy đủ các yêu cầu, quy trình và diễn giải kết quả cho các phép thử cụ thể để lựa chọn các quy trình thích hợp nhất nhằm giải quyết một vấn đề cụ thể.

Tiêu chuẩn bao gồm ba khía cạnh chính, được nêu trong Điều 4, Điều 5 và Điều 6.

Cần phải đọc Điều 4 "Các yêu cầu chung" trước tiên. Điều 5 "Phương pháp thử" mô tả tổng quát tất cả các phép thử chính. Điều 6 liên quan đến một số nguyên tắc chung về thu thập, phân tích dữ liệu cảm quan và cũng đề cập đến các nguyên tắc chung về xử lý thống kê kết quả.

Phân tích cảm quan – Phương pháp luận – Hướng dẫn chung

Sensory analysis – Methodology – General guidance

CẢNH BÁO – Tiêu chuẩn này không đề cập đến tất cả các vấn đề về an toàn, liên quan đến việc sử dụng tiêu chuẩn, nếu có. Người sử dụng có trách nhiệm thiết lập các thực hành về an toàn và sức khoẻ phù hợp và đảm bảo tuân thủ các quy định hiện hành.

1 Phạm vi áp dụng

Tiêu chuẩn này đưa ra hướng dẫn chung về việc sử dụng phép phân tích cảm quan. Tiêu chuẩn này mô tả các phép thử để kiểm tra thực phẩm và các sản phẩm khác bằng phân tích cảm quan và bao gồm một số thông tin chung về các kỹ thuật cần sử dụng nếu có yêu cầu phân tích thống kê các kết quả.

Nhìn chung, các phép thử này chỉ dành cho phân tích cảm quan một cách khách quan. Tuy nhiên, nếu phép thử được sử dụng để xác định sự ưa thích trong phép thử thị hiếu thì điều này cần được chỉ ra.

Phép thử thị hiếu nhằm mục đích xác định khả năng chấp nhận các sản phẩm và/hoặc xác định sự ưa thích giữa hai hoặc nhiều sản phẩm bởi số lượng người tiêu dùng quy định. Các phương pháp này có hiệu quả trong việc xác định sự ưa thích có thể cảm nhận được (sự khác biệt về mức độ ưa thích), hoặc không có sự ưa thích có thể cảm nhận được (phép thử cặp đôi giống nhau). Hướng dẫn chung đối với phép thử thị hiếu được nêu trong ISO 11136.

2 Tài liệu viện dẫn

Các tài liệu viện dẫn sau rất cần thiết cho việc áp dụng tiêu chuẩn này. Đối với các tài liệu viện dẫn ghi năm công bố thì áp dụng phiên bản được nêu. Đối với các tài liệu viện dẫn không ghi năm công bố thì áp dụng phiên bản mới nhất, bao gồm cả các sửa đổi, bổ sung (nếu có).

TCVN 11182 (ISO 5492), *Phân tích cảm quan – Thuật ngữ và định nghĩa*

3 Thuật ngữ và định nghĩa

Trong tiêu chuẩn này sử dụng các thuật ngữ và định nghĩa nêu trong TCVN 11182 (ISO 5492) cùng với thuật ngữ và định nghĩa sau:

3.1

Phân tích cảm quan (sensory analysis)

Khoa học liên quan đến việc đánh giá các thuộc tính cảm quan của một sản phẩm bằng các giác quan.

4 Yêu cầu chung

4.1 Thông tin cơ bản

Điều này đề cập đến các yêu cầu chung cho tất cả các tình huống gặp phải trong phân tích cảm quan. Thông tin cơ bản về các yêu cầu này như sau.

- a) Độ nhạy của con người đối với một kích thích không thể được tách rời với những trải nghiệm có được trước đó hoặc từ những kích thích cảm quan khác nhận được từ môi trường. Tuy nhiên, những ảnh hưởng phát sinh từ hai nguồn này có thể được kiểm soát và kết quả được chuẩn hóa.
- b) Sự dao động độ nhạy về cảm quan là điều tự nhiên trong bất kỳ nhóm người đánh giá nào và điều này là không thể tránh khỏi; điều này có thể do sự không nhất quán của từng cá nhân, sự khác biệt về tâm, sinh lý giữa các cá nhân. Tuy nhiên, thông qua huấn luyện, một nhóm như vậy có thể cho độ nhạy của các cá nhân có tính nhất quán cao. Việc công nhận các yếu tố này rất quan trọng trong phân tích kết quả.
- c) Độ chệch có hệ thống trong các thử nghiệm cảm quan liên quan đến độ nhạy của con người có thể làm cho dữ liệu bị hiểu nhầm và việc diễn giải không đúng có thể khó nhận biết. Các yếu tố có thể dẫn đến độ chệch cần được xác định và kiểm soát càng nhiều càng tốt bằng cách thiết kế các thử nghiệm và tiến hành thử thích hợp.
- d) Hiệu lực của các kết luận được rút ra từ các kết quả phụ thuộc vào phép thử đã sử dụng và cách thức tiến hành, bao gồm cả các câu hỏi được yêu cầu hỏi.

4.2 Mục đích thử nghiệm

Có ba mục đích chính, như sau:

- a) mục đích thử nghiệm là để phân nhóm, xếp hạng hoặc mô tả sản phẩm;
- b) khi mục đích nhằm phân biệt giữa hai hoặc nhiều sản phẩm trở lên; ở đây điều quan trọng là để phân biệt giữa nhu cầu cần biết:

- i) nếu có sự khác biệt,
 - ii) độ lớn của sự khác biệt này như thế nào,
 - iii) xu hướng (hoặc chất lượng) của sự khác biệt đó,
 - iv) ảnh hưởng của sự khác biệt đó, ví dụ: liên quan đến sự ưa thích, hoặc
 - v) nếu tất cả hoặc chỉ một phần dân số phát hiện được sự khác biệt;
- c) khi mục đích để chắc chắn tìm được tìm được sản phẩm đủ tương đồng để được sử dụng hoán đổi cho nhau.

Trong phân tích cảm quan, đối với một vấn đề nhất định đòi hỏi phải thảo luận thường xuyên hoặc suy nghĩ kỹ trước khi lựa chọn phép thử thích hợp. Đây là do cần phải làm rõ khái niệm ban đầu của vấn đề.

4.3 Lựa chọn phép thử

Việc lựa chọn phép thử thích hợp chủ yếu phụ thuộc vào bản chất của mục đích thử nghiệm, nhưng cũng cần phải tính đến các yếu tố liên quan đến sản phẩm, người đánh giá, môi trường thử nghiệm, mức độ chụm của phép phân tích yêu cầu và độ tin cậy thống kê trong kết luận. Cần xác định trước những hành động xảy ra dựa trên kết quả của phép thử.

Đối với mỗi phép thử, cách tiến hành được nêu trong Điều 5 đưa ra những hướng dẫn có liên quan đến phép thử. Có thể cần đến các phép thử sơ bộ để khẳng định khả năng áp dụng của phép thử.

Vì cảm giác mệt mỏi và ảnh hưởng của sự thích nghi, chỉ có một số lượng mẫu hạn chế có thể được thực hiện trong buổi đánh giá, điều này còn tùy thuộc vào bản chất của phép thử và loại sản phẩm. Một số ảnh hưởng này có thể được giảm thiểu bằng các thủ tục thanh vị và nghỉ ngơi giữa các lần thử mẫu.

Trong hầu hết các trường hợp, việc sử dụng các mẫu kiểm soát là cần thiết trong buổi đánh giá, việc sử dụng mẫu kiểm soát sẽ hạn chế số lượng mẫu cần được đánh giá.

Phải luôn xác định phương án thống kê trước khi bắt đầu phép thử. Điều này đặc biệt được khuyến cáo nếu số lượng mẫu cần đánh giá đòi hỏi nhiều buổi đánh giá. Chi tiết về các phương án thống kê phải được lựa chọn từ các tài liệu chuyên ngành. Bất kể phương pháp thử nào được sử dụng, phép thử tuân tự được mô tả trong TCVN 11186 (ISO 16820) cần được cân nhắc sử dụng để giữ số lượng mẫu hoặc số lượng người đánh giá tối thiểu.

4.4 Lựa chọn và huấn luyện người đánh giá

Hội đồng phân tích cảm quan chính xác là một "thiết bị đo lường", vì vậy mà các kết quả phân tích phụ thuộc vào các thành viên của hội đồng. Do đó, cần tiến hành tuyển chọn cẩn thận các thành viên hội

đồng và phải đầu tư thực sự cả về thời gian và tài chính. Sự hỗ trợ của quản lý trong tổ chức là cần thiết để có hiệu quả.

Đánh giá cảm quan có thể được thực hiện bởi ba kiểu người đánh giá: "người đánh giá cảm quan", "người đánh giá được chọn" hoặc "chuyên gia đánh giá cảm quan". Người đánh giá có thể là "người đánh giá chưa qua huấn luyện" là người không đáp ứng bất kỳ một tiêu chí cụ thể nào về tuyển chọn hoặc huấn luyện, hoặc những người đã tham gia một phần ở một số phép thử cảm quan ("người đánh giá ban đầu"). "Người đánh giá được lựa chọn" là người đánh giá đã được lựa chọn và huấn luyện về phép thử cảm quan cụ thể. "Chuyên gia đánh giá cảm quan" là người đánh giá được lựa chọn và huấn luyện các phương pháp phân tích cảm quan khác nhau và có thể thực hiện đánh giá cảm quan một cách nhất quán và lặp lại các sản phẩm trong một nhóm hoặc một số nhóm.

Tiến hành phương pháp lựa chọn và huấn luyện phụ thuộc vào nhiệm vụ và phương pháp được sử dụng cho "người đánh giá được lựa chọn". Quy trình huấn luyện người đánh giá của phép thử mô tả khác với quy trình huấn luyện người đánh giá của phép thử phân biệt.

Quy trình chi tiết và phương pháp lựa chọn và huấn luyện người đánh giá được nêu trong TCVN 12389 (ISO 8586). Cần lưu ý rằng, những phương pháp này đôi khi chỉ đưa ra cách để chọn được ứng viên tốt nhất trong số những người có sẵn, hơn là để đáp ứng tiêu chí đã được xác định trước. Ngoài ra, việc lựa chọn người đánh giá về khả năng phân biệt, mô tả thực phẩm và các sản phẩm hoàn toàn khác so với việc lựa chọn người đánh giá dành cho các phép thử ưa thích. Đối với việc lựa chọn người đánh giá về khả năng phân biệt, mô tả đòi hỏi hội đồng có sự lựa chọn và huấn luyện, trong khi việc lựa chọn người đánh giá trong phép thử ưa thích chỉ đòi hỏi hội đồng đại diện cho một bộ phận người đánh giá cụ thể, ví dụ nhóm người tiêu dùng.

Nếu tiến hành quy trình lựa chọn thì một số tiêu chí quan trọng để chọn người đánh giá như sau:

- a) năng lực tổng thể để thực hiện nhiệm vụ cảm quan cụ thể;
- b) khả năng sẵn sàng tham gia giống như với một việc làm bình thường;
- c) động cơ (sẵn sàng và quan tâm);
- d) sức khoẻ tốt (bao gồm cả việc không bị dị ứng đặc biệt hoặc đang điều trị bằng thuốc), tình trạng răng tốt và điều kiện vệ sinh chung tốt.

Phải thường xuyên giám sát hoạt động của "người đánh giá được lựa chọn" và "chuyên gia đánh giá cảm quan" để đảm bảo rằng tiêu chí mà họ được lựa chọn ban đầu tiếp tục được đáp ứng (xem TCVN 11132 (ISO 22118)).

4.5 Vật liệu cần được thử nghiệm

Bản chất của sản phẩm cần được thử nghiệm quyết định quy trình tiến hành phép thử và có thể cũng ảnh hưởng đến dạng phép thử được yêu cầu nhằm thỏa mãn các mục đích thử nghiệm. Ví dụ, thực phẩm được tiêu thụ ở dạng nóng thì cần tính đến tốc độ nguội của sản phẩm, khả năng ảnh hưởng đến các thuộc tính cảm quan và sự thay đổi các thuộc tính cảm quan có thể xuất hiện trong khi giữ cho sản phẩm nóng trước khi thử nghiệm.

Các phương pháp chuẩn bị và trình bày mẫu phải phù hợp với sản phẩm và các vấn đề liên quan.

Ví dụ 1 Sản phẩm thường được tiêu thụ ở dạng nóng được chuẩn bị theo cách thông thường và được thử nóng; tuy nhiên, có thể việc sử dụng nhiệt độ cao trong một số trường hợp làm tăng sự dễ cảm nhận đối với một số hương vị được đánh giá.

Ví dụ 2 Sản phẩm thường được tiêu thụ ở dạng miếng rời thì không đồng hóa nhằm giữ các đặc trưng về cấu trúc. Tuy nhiên, đối với từng người đánh giá, cần thận trọng để đảm bảo sự đồng nhất tối đa giữa các mẫu con; sự đồng nhất này bao gồm tỷ lệ kích cỡ tương tự và sự đồng đều của thành phần.

Nguyên tắc chung về lấy mẫu sản phẩm (tuân thủ theo các tiêu chuẩn liên quan đến sản phẩm cần thử nghiệm) phải được áp dụng đối với các mẫu thử. Trong tất cả các trường hợp, phải có tài liệu chứng minh mã nhận biết mẫu hoặc số lô. Có thể rút ra kết luận đúng về một sản phẩm như đối với toàn bộ lô chỉ khi các mẫu được thử nghiệm là đại diện.

Đôi khi có thể sử dụng vật chứa cho các phép thử liên quan đến việc đánh giá các sản phẩm mà thử nghiệm trực tiếp là không khả thi (xem ISO 5497)⁽¹⁾, ví dụ các thành phần thực phẩm.

Phải quy định điều kiện ánh sáng khi đánh giá bề ngoài. Khi phép thử chỉ liên quan đến sự khác biệt về hương, ảnh hưởng của sự khác biệt về màu sắc có thể được che bớt một phần bằng cách sử dụng các điều kiện ánh sáng nhằm giảm thiểu sự khác biệt về màu sắc.

Phải chọn những vật chứa sao cho không ảnh hưởng đến phép thử hoặc sản phẩm. Chúng có thể bao gồm các vật chứa bằng gốm hoặc bằng thủy tinh có thể rửa được, hoặc các vật chứa bằng chất dẻo hoặc bằng giấy dùng một lần, nhưng không thôi nhiễm vật liệu hóa chất có thể gây hư hỏng sản phẩm. Đặc biệt, các vật chứa có thể rửa được phải được rửa bằng chất tẩy rửa không mùi, không gây hư hỏng và tráng được bằng nước, vật chứa bằng polyme và giấy, kể cả vật chứa cách nhiệt dùng cho mẫu nóng hoặc mẫu lạnh cũng không được có mùi và không gây hư hỏng sản phẩm.

Chất làm thanh vị có thể được người đánh giá sử dụng giữa các mẫu và giữa các buổi đánh giá, nhưng phải làm sạch vòm miệng cẩn thận để đảm bảo rằng chúng không ảnh hưởng đến hương vị của sản phẩm cần được đánh giá. Nước không ga, nước có ga và các thực phẩm ngọt (ví dụ: bánh quy giòn không muối) có thể được sử dụng giữa các mẫu và giữa các buổi đánh giá. Cần kiểm tra nguồn cung cấp nước để đảm bảo rằng chúng nhạt và an toàn để sử dụng. Đối với các mục đích đặc biệt, có thể sử dụng nước đã khử khoáng, nước cất bằng dụng cụ thủy tinh, nước muối có hàm lượng khoáng thấp, nước đã lọc carbon hoặc nước máy đun sôi, nhưng cần lưu ý rằng chúng có nhiều mùi hương khác nhau.

4.6 Phòng thử nghiệm

Phân tích cảm quan phải được thực hiện trong phòng thử nghiệm chuyên dụng [xem hướng dẫn của TCVN 12390 (ISO 8589)]. Mục đích là nhằm tạo cho từng người đánh giá một môi trường riêng biệt với sự phân tâm tối thiểu, để từng người đánh giá có thể nhanh chóng điều chỉnh về trạng thái tự nhiên của công việc mới. Trong suốt quá trình thử nghiệm, không cho phép các hoạt động khác, bao gồm cả chuẩn bị mẫu, vì có thể dẫn đến sai lệch kết quả. Phòng thử nghiệm phải có nhiệt độ phòng phù hợp và phải được thông gió với khí không mùi; cần hạn chế dòng không khí để tránh sự biến động nhiệt độ quá mức. Không được phép làm nhiễm bẩn môi trường của phòng thử nghiệm bởi mùi dai dẳng như thuốc lá hoặc mỹ phẩm.

Phải hạn chế âm thanh. Tiếng ồn xung quanh thấp thường được chấp nhận hơn tiếng ồn dao động. Chuyển trò sẽ gây mất tập trung hơn tiếng ồn xung quanh. Sự gián đoạn gây ra sự phân tâm lớn nhất.

Thông thường, nên kiểm soát cả màu sắc và cường độ ánh sáng, mặc dù ánh sáng màu hiếm khi che giấu được hoàn toàn sự khác biệt về trạng thái bề ngoài.

Các bề mặt của phòng thử nghiệm không được thấm nước và cần thiết kế với tiêu chuẩn vệ sinh cao. Kích cỡ của buồng thử nếm là rất quan trọng; trần nhà rất thấp và buồng rất hẹp có thể gây áp lực hoặc có thể gây ra cảm giác sợ hãi. Cần có chỗ ngồi phù hợp.

Nếu thực hiện thu thập dữ liệu trên máy vi tính thì cần thực hiện một cách an toàn, vệ sinh và không làm ảnh hưởng đến phán đoán cảm quan.

4.7 Lập kế hoạch và tiến hành thử nghiệm

Việc lập kế hoạch và vận hành phép thử được xác định bởi các mục đích của chương trình, phép thử được chọn và những ràng buộc thực tế liên quan đến việc sử dụng con người. Đặc biệt, điều quan trọng là nhận ra độ chệch có thể có trong phép thử đã chọn và tiến hành phép thử sao cho giảm thiểu ảnh hưởng của độ chệch. Độ chệch tiềm tàng có thể do cả tâm lý và sinh lý.

Độ chệch kết quả do tâm lý là nghiêm trọng nhất vì tương tác giữa người đánh giá làm ảnh hưởng đến phán đoán của nhau và cần được giảm thiểu bằng cách sử dụng buồng thử nghiệm đơn lẻ hoặc tách riêng những người đánh giá một cách thích hợp. Ngoài ra, cần quản lý chặt chẽ các hoạt động của những người đánh giá.

Cách thức và thứ tự trình bày các mẫu là những khía cạnh quan trọng của phép thử và có thể dẫn đến độ chệch do tâm lý. Ví dụ, các mẫu phải được mã hoá bằng số ba chữ số ngẫu nhiên và các mã cần thay đổi đối với từng mẫu thử. Thứ tự đánh giá cũng có thể là nguyên nhân gây độ chệch, thông thường, phải quy định thứ tự đánh giá. Với số lượng mẫu và người đánh giá ít, thứ tự đánh giá có thể được cân bằng sao cho có thể đánh giá số lần bằng nhau. Trong các thí nghiệm lớn hơn, thứ tự có thể được cân bằng hoặc ngẫu nhiên hóa.

Độ chệch do sinh lý thường liên quan đến bản chất của các mẫu thử. Đặc biệt, sự thích nghi với kích thích về mùi hương cụ thể có thể xuất hiện khi tiếp xúc nhiều lần với kích thích đó và sự mệt mỏi có thể xảy ra khi nhai những thực phẩm cứng. Cả hai yếu tố đều có thể dẫn đến giới hạn trên về số lượng mẫu được đánh giá trong một buổi thử. Việc nhà mẫu có thể được khuyến cáo với các hội đồng đã được huấn luyện nhưng có thể làm mất thông tin về các thuộc tính cảm quan cụ thể.

Đói và no có thể ảnh hưởng đến hiệu suất của người đánh giá và nếu hội đồng được tổ chức quá thường xuyên thì hiệu suất có thể kém đi. Nếu có thể, người đánh giá được yêu cầu kiểm chế không hút thuốc và ăn uống vật, ví dụ uống cà phê, trong 1 h trước khi thử nghiệm. Người đánh giá không được mang bất kỳ mùi lạ nào vào buổi đánh giá, ví dụ mùi thuốc hoặc mùi mỹ phẩm vì chúng có thể ảnh hưởng đến độ nhạy của những người đánh giá khác.

Thời gian trong ngày thực hiện phép thử là rất quan trọng. Lịch trình cần tính đến các bữa ăn theo phong tục địa phương vì hiệu suất thường được coi là tối ưu vào giữa buổi sáng và giữa buổi chiều. Người đánh giá bị rối loạn cảm xúc, bị cảm lạnh và các bệnh khác không được tham gia thử nghiệm cho đến khi hồi phục sức khỏe.

Việc đối chiếu các kết quả bao gồm ba khía cạnh:

- kiểm tra xem tất cả dữ liệu đã được ghi lại chính xác hay chưa, bằng máy vi tính hoặc bằng tay;
- xác minh xem mọi thông tin liên quan bổ sung có thể hỗ trợ hoặc gây nghi ngờ về việc diễn giải kết quả đã được ghi chép lại hay chưa;
- kiểm tra xem người đánh giá có động lực để tiếp tục tham gia hay không nếu có kế hoạch thử nghiệm tiếp.

5 Phương pháp thử

5.1 Yêu cầu chung

Các phép thử được sử dụng phổ biến nhất được chia thành ba nhóm:

- a) các phép thử phân biệt được sử dụng để xác định xác suất khác nhau hoặc giống nhau giữa các sản phẩm;
- b) các phép thử sử dụng thang đo và phân nhóm sử dụng để ước tính thứ tự hoặc quy mô của sự khác biệt hoặc phân nhóm hoặc phân loại để định vị mẫu;
- c) các phép thử mô tả được sử dụng để mô tả đặc điểm, cả định tính, định lượng và các thuộc tính cảm quan cụ thể của mẫu (xem 5.4).

Phần lớn các phép thử này với một số sửa đổi có thể được sử dụng cho phép thử người tiêu dùng [27].

TCVN 12387:2018

Về số lượng người đánh giá, tham khảo các tiêu chuẩn tương ứng, có tính đến các rủi ro α hoặc β tùy thuộc vào mục đích thử nghiệm. Ngoài ra, phép thử tuân tự [xem TCVN 11186 (ISO 16820)] có thể cho phép đưa ra quyết định với ít phép thử hơn so với yêu cầu của phương pháp truyền thống sử dụng một số đánh giá đã được xác định trước.

5.2 Phép thử phân biệt

5.2.1 Yêu cầu chung

Trước khi thực hiện phép thử phân biệt, phải xác định mục đích của phép thử:

- để chỉ ra có tồn tại sự khác biệt có nghĩa giữa hai sản phẩm;
- để chỉ ra có tồn tại sự giống nhau có nghĩa giữa hai sản phẩm.

Trong trường hợp thứ nhất, sử dụng phép thử khác biệt. Trong trường hợp thứ hai, sử dụng phép thử tương đồng.

Các phép thử sau đây thường được sử dụng để xác định khả năng khác nhau hoặc giống nhau giữa các mẫu:

- a) phép thử so sánh cặp đôi (xem 5.2.2);
- b) phép thử tam giác (xem 5.2.3);
- c) phép thử hai-ba (xem 5.2.4);
- d) phép thử chọn hai trong năm (xem 5.2.5);
- e) phép thử "A - không A" (xem 5.2.6);
- f) phép thử bốn mẫu (xem 5.2.7).

Đối với tất cả các phép thử này, có những cách phân tích kết quả khác nhau.

5.2.2 Phép thử so sánh cặp đôi

5.2.2.1 Định nghĩa

Đây là phép thử trong đó các mẫu được trình bày theo cặp để so sánh và phát hiện sự khác nhau hoặc sự giống nhau về cảm quan trên cơ sở một số tiêu chí xác định.

Xem TCVN 4831 (ISO 5495) để biết chi tiết.

5.2.2.2 Ứng dụng

Phép thử so sánh cặp đôi được khuyến cáo:

- a) để xác định nếu tồn tại sự khác biệt có thể có thể cảm nhận được (mục đích của sự khác biệt) trong thuộc tính cụ thể (ví dụ vị ngọt) hoặc để xác định liệu có tồn tại sự giống nhau trong thuộc tính đó,
- b) để lựa chọn, huấn luyện và giám sát hoạt động của người đánh giá,
- c) để so sánh hai sản phẩm tốt nhất là trong khuôn khổ phép thử người tiêu dùng.

Ưu điểm của phép thử trên so với các phép thử phân biệt khác là sự đơn giản và cảm giác ít mệt mỏi hơn.

Nhược điểm của phương pháp so sánh cặp đôi là khi số lượng mẫu cần so sánh tăng lên, dẫn đến số lượng các phép thử so sánh tăng nhanh đến mức không khả thi.

5.2.2.3 Cách tiến hành

Người đánh giá nhận một bộ hai mẫu (nghĩa là một cặp mẫu). Họ chỉ định mẫu mà họ coi là mạnh nhất (Intense) đối với thuộc tính đang được xem xét, ngay cả khi lựa chọn này chỉ dựa trên phán đoán. Một trong hai mẫu (của cặp mẫu) có thể là mẫu đối chứng. Đếm số lần mỗi mẫu được chọn và mức ý nghĩa được xác định bằng cách tham chiếu với bảng thống kê.

Trước khi tiến hành thử nghiệm, cần xác định,

- a) mục đích của phép thử là gì (phép thử khác biệt hay phép thử tương đồng),
- b) độ nhạy thích hợp nhất là bao nhiêu, và
- c) thử nghiệm thống kê cần thực hiện là:
 - i) một chiều (nghĩa là người thực hiện phép thử mong đợi một chiều hướng khác biệt cụ thể và giả thuyết thay thế tương ứng với sự tồn tại sự khác biệt theo hướng đó), hoặc
 - ii) hai chiều (nghĩa là người thực hiện phép thử không có chiều hướng khác biệt mong đợi nào và giả thuyết thay thế tương ứng với hướng khác biệt có thể ở bất cứ chiều nào).

Không được kết hợp các câu hỏi về sự khác biệt và sự ưa thích: đối với những câu hỏi này, tiêu chí tuyển chọn đối với mỗi hội đồng là khác nhau.

5.2.2.4 Phân tích kết quả

Xem 6.2.2.

TCVN 12387:2018

5.2.3 Phép thử tam giác

5.2.3.1 Định nghĩa

Đây là phép thử phân biệt liên quan đến ba mẫu đã được mã hóa, hai trong số ba mẫu đó là đồng nhất, được trình bày đồng thời. Người đánh giá được yêu cầu chọn mẫu khác biệt với hai mẫu còn lại.

Xem TCVN 11184 (ISO 4120) để biết chi tiết.

5.2.3.2 Ứng dụng

Phép thử tam giác được khuyến cáo

- a) khi chưa biết được bản chất của sự khác biệt, và
- b) để lựa chọn và huấn luyện người đánh giá.

Không được dùng phép thử để xác định sự ưa thích. Một số bất lợi của phép thử tam giác đó là:

- không kinh tế khi đánh giá số lượng mẫu lớn,
- có thể bị ảnh hưởng nhiều hơn bởi sự thích nghi về cảm quan so với phép thử so sánh cặp đôi sử dụng các mẫu có các thuộc tính cảm quan mạnh,
- nếu biết được bản chất của sự khác biệt thì phép thử ít này hiệu quả hơn so với một số phép thử khác, và
- phương pháp này chỉ có thể áp dụng được nếu sản phẩm càng đồng nhất càng tốt.

5.2.3.3 Cách tiến hành

Mỗi lần trình bày với người đánh giá một bộ gồm ba mẫu đã được mã hoá, trong đó có hai mẫu giống nhau và được yêu cầu chọn mẫu khác.

Các mẫu cần được trình bày với các số lần như nhau, thành hai bộ có ba hoán vị khác nhau, đó là:

BAA ABA AAB

ABB BAB BBA

5.2.3.4 Phân tích kết quả

Xem 6.2.3.

5.2.4 Phép thử hai-ba

5.2.4.1 Định nghĩa

Đây là phép thử phân biệt, trong đó mẫu chuẩn được trình bày trước tiên. Sau đó đến hai mẫu, một trong hai mẫu giống hệt mẫu chuẩn và yêu cầu người đánh giá nhận ra mẫu đó.

Xem TCVN 11185 (ISO 10399) để biết chi tiết.

5.2.4.2 Ứng dụng

Phép thử hai-ba này được sử dụng để xác định nếu giữa mẫu đã cho và mẫu chuẩn khác nhau hoặc giống nhau về cảm quan. Phép thử này đặc biệt thích hợp khi người đánh giá biết rõ về mẫu chuẩn, ví dụ một mẫu sản xuất thường xuyên.

Nếu có những cảm giác mạnh sau thử thì phép thử này kém phù hợp hơn so với phép thử so sánh cặp đôi (5.2.2) hoặc phép thử "A - không A" (5.2.6).

5.2.4.3 Cách tiến hành

Đầu tiên người đánh giá giới thiệu với mẫu chuẩn đã được nhận dạng. Tiếp theo là hai mẫu được mã hóa, một trong hai mẫu giống hệt với mẫu chuẩn. Người đánh giá được yêu cầu nhận dạng mẫu này.

5.2.4.4 Kết quả phân tích

Xem 6.2.4.

5.2.5 Phép thử chọn hai trong năm

5.2.5.1 Định nghĩa

Đây là phép thử nghiệm phân biệt liên quan đến năm mẫu được mã hóa, hai trong số năm mẫu là một loại và ba mẫu kia là của loại khác. Người đánh giá được yêu cầu chia mẫu thành hai nhóm.

5.2.5.2 Ứng dụng

Phép thử chọn hai trong năm được khuyến cáo để thiết lập một sự khác biệt một cách kinh tế hơn so với các phép thử khác (phương pháp này hiệu quả hơn về mặt thống kê).

Nhược điểm của phép thử này tương tự như phép thử tam giác (5.2.3). Phép thử bị ảnh hưởng nhiều hơn do sự mệt mỏi cảm giác và việc ghi nhớ nhưng có năng lực thống kê cao hơn. Phép thử này được sử dụng chủ yếu trong các ứng dụng thị giác, thính giác hoặc xúc giác.

TCVN 12387:2018

5.2.5.3 Cách tiến hành

Mỗi lần trình bày với người đánh giá một bộ gồm năm mẫu đã được mã hóa và được cho biết hai mẫu là của một loại và ba mẫu kia là của loại khác. Người đánh giá được yêu cầu chia mẫu thành hai nhóm

Khi số lượng người đánh giá nhỏ hơn 20, thứ tự trình bày phải được chọn ngẫu nhiên từ 20 hoán vị khác nhau sau đây:

AAABB BBBA AABAB BBABA ABAAB BABBA BAAAB ABBA AABBA BBAAB ABABA BABAB
BAABA ABBAB ABBAA BAABB BABAA ABABB BBAAB AABBB

5.2.5.4 Phân tích kết quả

Xem 6.2.5.

5.2.6 Phép thử "A – không A"

5.2.6.1 Định nghĩa

Đây là phép thử trong đó một loạt mẫu, có thể là "A" hoặc "không A", được trình bày để người đánh giá thực hiện đánh giá sau khi họ đã học được cách nhận diện mẫu "A". Người đánh giá được yêu cầu chỉ ra mẫu nào là "A".

Xem ISO 8588 để biết chi tiết.

5.2.6.2 Ứng dụng

Phép thử này là phép thử phân biệt, có thể được sử dụng để đánh giá các mẫu có bề ngoài khác nhau hoặc để lại cảm giác dai dẳng sau đó.

Phép thử này đặc biệt hữu ích khi không thể thu được các mẫu lặp lại giống nhau.

5.2.6.3 Cách tiến hành

Mỗi thời điểm, trình bày một mẫu với người đánh giá. Người đánh giá trước tiên được trình bày với mẫu chuẩn "A" một vài lần, cho đến khi họ có thể nhận ra chúng. Người đánh giá được nhận ngẫu nhiên một số mẫu, một trong số các mẫu đó là có thể là "A" hoặc "không A", và họ phải xác định chúng là gì. Cần cho phép một khoảng thời gian có thể đánh giá được (ví dụ từ 2 min đến 5 min) giữa thời gian nhận mẫu và chỉ kiểm tra một vài mẫu trong một buổi đánh giá.

5.2.6.4 Phân tích kết quả

Xem 6.2.6.

5.2.7 Phép thử bốn mẫu

5.2.7.1 Định nghĩa

Đây là phép thử phân biệt liên quan đến bốn mẫu đã được mã hóa, hai mẫu trong số đó là một loại và hai mẫu còn lại là loại khác. Người đánh giá được yêu cầu nhóm hai bộ mẫu.

Xem ASTM E3009-15.

5.2.7.2 Ứng dụng

Phép thử này có thể được sử dụng để xác định liệu có tồn tại sự khác biệt về cảm giác có thể nhận thấy giữa các mẫu của hai sản phẩm hoặc để ước tính độ lớn của sự khác biệt có thể cảm nhận được, ví dụ khi có sự thay đổi về thành phần, quá trình chế biến, bao gói, xử lý hoặc bảo quản.

5.2.7.3 Cách tiến hành

Người đánh giá nhận được một bộ gồm bốn mẫu đã được mã hóa, hai mẫu từ một nhóm và hai mẫu từ nhóm khác và được yêu cầu lựa chọn các mẫu thành hai nhóm của hai mẫu dựa trên sự giống nhau bằng cách sử dụng phương pháp được quy định hoặc không được quy định.

Các mẫu cần được thể hiện với các số lần như nhau, thành một trong hai bộ có ba hoán vị khác nhau, đó là:

AABB ABAB ABAB

BBAA BABA BAAB

5.2.7.4 Phân tích kết quả

Xem 6.2.7.

5.3 Sử dụng thang đo và nhóm

5.3.1 Xem xét chung

Xem TCVN 5090 (ISO 4121) để biết chi tiết.

Các phương pháp đo lường trong phân tích cảm quan có thể tìm cách quyết định các nhóm, loại hoặc cấp mà các mẫu được phân bổ. Phương pháp này cũng có thể biểu thị số lượng ước tính về độ lớn các thuộc tính cảm nhận được của mẫu hoặc sự khác nhau giữa các mẫu.

Không sử dụng trực tiếp thang đo cảm nhận để suy ra số lượng và thang đo tương ứng với các giá trị đã được ghi lại. Do đó, với cùng một phương pháp thu được các số (thang đo cảm nhận) có thể dẫn

TCVN 12387:2018

đến các giá trị của thang đo chỉ là thứ tự (khoảng không bằng nhau) hoặc thang đo khoảng (khoảng bằng nhau). Với thang đo thứ bậc, kích cỡ của sự khác biệt giữa hai giá trị không thể được giả định để phản ánh sự khác biệt giữa các cường độ cảm nhận được. Cũng không thể giả định tỉ số của hai giá trị phản ánh tỉ số của các cường độ cảm nhận được. Với thang đo khoảng, các giá trị bằng số lớn hơn tương ứng với cường độ cảm nhận lớn hơn (hoặc mức độ hài lòng) và kích cỡ của sự khác biệt giữa hai giá trị phản ánh kích cỡ của sự khác biệt trong cường độ cảm nhận đặc tính cần đo. Tuy nhiên, giá trị bằng zero (0) có thể không thể hiện là không có toàn bộ đặc tính có và tỷ số của hai giá trị không thể được giả định để phản ánh tỷ số của cường độ cảm nhận.

Lựa chọn thang đo cảm nhận phụ thuộc vào các mục đích nghiên cứu và các sản phẩm cần phân tích. Trong bất kỳ trường hợp cụ thể nào, có thể chọn một vài thang đo tốt như nhau. Bất kể thang đo cảm nhận nào được chấp nhận, chúng phải dễ sử dụng, dễ phân biệt, không bị sai lệch và dễ hiểu đối với người đánh giá [xem TCVN 5090 (ISO 4121)].

5.3.2 Diễn giải kết quả

5.3.2.1 Chất lượng của các phép đo thu được

Bất kể thang đo cảm nhận nào thì chất lượng của các phép đo phụ thuộc vào cách mà chúng thu được. Các khía cạnh được xem xét là:

- mức huấn luyện của người đánh giá [xem TCVN 12389 (ISO 8586)], và
- phương pháp trình bày các mẫu (xem 4.5 và 4.7).

5.3.2.2 Năng lực thống kê và diễn giải

Phân tích thống kê bị ảnh hưởng bởi bản chất của thang đo (thứ bậc, khoảng hoặc tỷ lệ) chứ không phải do thang đo cảm nhận đã sử dụng.

Kết quả đo được trên thang đo thứ bậc được phân tích tốt nhất bằng cách sử dụng phương pháp phi tham số, ví dụ phép thử Wilcoxon trong trường hợp hai mẫu như nhau hoặc phép thử Friedman với nhiều hơn hai mẫu. Các phép đo trên thang đo khoảng hoặc thang đo tỷ số có thể được phân tích bằng phép thử tham số, ví dụ phân tích phương sai, nếu có thể giả định phân bố chuẩn các số dư.

Các phép thử tham số thường có hiệu lực hơn so với phép thử phi tham số. Nghĩa là, nếu tồn tại sự khác biệt thì phép thử tham số có khả năng chứng minh tốt hơn. Ngược lại, các phép thử phi tham số có tính chất "thô" (robust) lớn hơn các phép thử tham số; nghĩa là chúng ít bị ảnh hưởng bởi các bất thường trong dữ liệu.

5.3.2.3 Các dạng phép đo

Cần phân biệt các loại kiểu đo lường sau đây:

- a) phân loại (xem 5.3.3);
- b) phân cấp (xem 5.3.4);
- c) xếp hạng (xem 5.3.5);
- d) xếp loại và cho điểm (xem 5.3.6).

Thuật ngữ "phép đo" là chung và đề cập đến tất cả những điều này.

5.3.3 Phân loại

5.3.3.1 Yêu cầu chung

Việc phân loại liên quan đến phương pháp phân nhóm mẫu (theo cách vật lý hoặc theo nhận nhận biết chúng) thành các nhóm đã được xác định trước.

5.3.3.2 Ứng dụng

Việc phân loại có thể ứng dụng khi muốn phân bổ các mẫu vào nhóm phù hợp nhất trong số các nhóm được sắp xếp không theo thứ tự nào cả. Ví dụ, cá có thể được chọn lọc theo loài, hoặc các mẫu có thể được phân bổ vào các hạng phù hợp với kiểu khuyết tật của chúng. Nguyên tắc là mỗi mẫu được phân bổ theo nhóm là điển hình nhất. Nếu các số được sử dụng để đại diện cho nhóm thì phép đo được gọi là thang đo danh nghĩa, với các số chỉ được sử dụng như tên nhãn. Đối với những con số như vậy thông tin không phải là thứ tự hay độ lớn.

5.3.4 Phân cấp

5.3.4.1 Yêu cầu chung

Phân cấp là sự phân loại sản phẩm theo chất lượng, dựa trên một hoặc nhiều thuộc tính cảm quan, thường dựa trên các khuyết tật. Người đánh giá thường là các chuyên gia hoặc người đánh giá đã được huấn luyện. Phương pháp này yêu cầu phải lựa chọn trước các thuộc tính và định nghĩa về cấp.

Việc phân cấp liên quan đến phương pháp chọn lọc mẫu tạo thành nhóm cấu thành thang đo thứ bậc về chất lượng, ví dụ cấp I (không có mùi lỗi/hương lỗi), cấp II (có mùi lỗi/hương lỗi nhẹ).

5.3.4.2 Ứng dụng

Việc phân cấp được áp dụng khi muốn phân bổ các mẫu vào một số nhóm phù hợp nhất được giả định tương ứng với chất lượng. Ví dụ, cá có thể được chọn lọc theo độ tươi hoặc mẫu có thể được phân bổ vào các nhóm theo mức độ nghiêm trọng của các khuyết tật. Nguyên tắc là mỗi mẫu được phân bổ vào nhóm điển hình nhất. Nếu các số được sử dụng đại diện cho nhóm thì phép đo được gọi là thang đo thứ bậc. Những số như vậy chỉ là thứ tự thông tin.

5.3.5 Xếp hạng

5.3.5.1 Yêu cầu chung

Xem TCVN 11183 (ISO 8587) để biết chi tiết.

CHÚ THÍCH So sánh với 5.3.6.

Xếp hạng là phương pháp phân loại trong đó một loạt các mẫu được đặt theo thứ tự cường độ hoặc mức độ của một số thuộc tính cụ thể. Phương pháp này cho phép tìm ra sự khác biệt nếu có, nhưng không nhằm đánh giá độ lớn của sự khác biệt giữa các mẫu. Nếu vị trí trong thứ tự được ghi lại là con số thì thang đo kết quả là thang đo thứ bậc.

5.3.5.2 Ứng dụng

Xếp hạng có thể được sử dụng như là biện pháp nhanh để mô tả mẫu với số lượng nhỏ (khoảng sáu mẫu) trên các thuộc tính có thể tạo ra cảm giác một mùi (ví dụ: hương vị) hoặc trên mẫu với số lượng lớn (lên đến 20 mẫu) trong trường hợp không có cảm giác một mùi (ví dụ ở bề ngoài).

5.3.5.3 Cách tiến hành

Cần đảm bảo rằng người đánh giá hiểu và đồng ý về thuộc tính hoặc tiêu chí mà mẫu cần xếp hạng. Mỗi người đánh giá độc lập kiểm tra các mẫu đã được mã hoá theo thứ tự chỉ định và xếp hạng sơ bộ. Người đánh giá phải xem lại việc xếp hạng này bằng cách kiểm tra lại các mẫu và điều chỉnh, nếu cần, bằng cách xếp hạng thứ tự.

5.3.6 Xếp loại và cho điểm

5.3.6.1 Yêu cầu chung

Xếp loại là phương pháp phân loại trong đó mỗi mẫu được phân bổ với một số điểm theo thang đo thứ bậc. Nhiều mẫu có thể được phân bổ với cùng một điểm thang đo. Thang đo có thể dạng số, dạng chữ, hình ảnh hoặc kết hợp giữa chúng. Có thể là thang đo liên tục hoặc rời rạc và đơn cực hoặc lưỡng cực [xem TCVN 5090 (ISO 4121)]. Nếu thang đo dạng số thì quy trình thường được gọi là "cho điểm". Điều này giúp cho người đánh giá có một số mẫu làm chuẩn để nhận biết các điểm cụ thể trên thang đo.

Mặc dù cả xếp hạng và xếp loại đều chỉ là thang đo thứ bậc, nhưng chúng không tương đương nhau. Xếp hạng bố trí các mẫu theo thứ tự và kết quả chỉ dẫn đến xếp hạng mẫu. Còn xếp loại đưa ra cách ước tính thứ bậc về mức độ của các thuộc tính hoặc sự ưa thích vì sử dụng cùng một thang đo thứ bậc không kể các mẫu cụ thể đang được đánh giá. Vì vậy, xếp loại là thích hợp hơn nếu kết quả từ một bộ mẫu cần được so sánh với kết quả của các bộ mẫu khác. Nhưng vì xếp hạng khuyến khích

người đánh giá sử dụng mọi sự khác biệt giữa các mẫu mà họ cảm nhận được, nên sự khác biệt nhỏ giữa các mẫu phát hiện được sẽ được xếp loại giống nhau.

5.3.6.2 Ứng dụng

Xếp loại có thể được sử dụng để đánh giá cường độ của một hoặc nhiều thuộc tính hoặc mức độ thích (với người tiêu dùng) mẫu hoặc mức độ chênh lệch giữa mẫu thử và mẫu chuẩn.

Cho điểm thường được áp dụng trong kiểm soát chất lượng được tiến hành thường xuyên trên mẫu với số lượng lớn và/có một chút áp lực về thời gian và/hoặc số lượng chuyên gia đánh giá có sẵn bị hạn chế. Cho điểm được sử dụng để đánh giá mẫu trong thương mại quốc tế, ví dụ các sản phẩm sữa [xem TCVN 10565-3(ISO 22935-3)].

5.3.6.3 Cách tiến hành

Việc phân loại phải được những người đánh giá xác định và hiểu rõ. Mỗi người đánh giá kiểm tra độc lập từng mẫu theo thứ tự và cho từng điểm trên thang đo.

5.4 Phép thử mô tả

5.4.1 Yêu cầu chung

Các phép thử này có thể được áp dụng cho một hoặc nhiều mẫu để xác định đặc điểm, cả về chất lượng và số lượng của một hoặc nhiều thuộc tính cảm quan. Các phép thử thường được sử dụng nhất đó là:

- a) profile cảm quan định tính;
- b) profile cảm quan định lượng;
- c) profile đồng thuận
- d) profile lựa chọn tự do;
- e) profile mô tả nhanh;
- f) profile mô tả độ chênh so với mẫu chuẩn;
- g) các cảm nhận nổi trội theo thời gian (TDS).

Đối với tất cả các phép thử này, có nhiều phương pháp phân tích kết quả khác nhau. ISO 13299 mô tả các phương pháp được sử dụng để thiết lập profile cảm quan một cách chi tiết.

5.4.2 Profile cảm quan định tính

5.4.2.1 Định nghĩa

Đây là phép thử trong đó các thuộc tính cảm quan của sản phẩm được mô tả bởi hội đồng đã được huấn luyện nhưng không có giá trị cường độ đo.

5.4.2.2 Ứng dụng

Phép thử có thể được sử dụng:

- để nhận biết và mô tả các thuộc tính của một mẫu hoặc của các mẫu cụ thể, và
- để thiết lập trình tự trong đó các thuộc tính này được cảm nhận.

Phép thử này nên được sử dụng trong việc mô tả sự khác biệt đã được thiết lập trước đó. Phép thử này hữu ích cho việc huấn luyện người đánh giá, hoặc để xây dựng bộ từ vựng ban đầu cho các kỹ thuật mô tả sâu hơn được nêu dưới đây.

5.4.2.3 Cách tiến hành

Phép thử này nên được sử dụng cho một hoặc nhiều mẫu. Khi có nhiều mẫu được trình bày trong buổi đánh giá thì thứ tự trình bày các mẫu sẽ có ảnh hưởng. Cần ngẫu nhiên mã hóa mẫu cho mỗi người đánh giá.

Mẫu được đánh giá độc lập bởi từng người đánh giá và kết quả được ghi lại. Có thể cung cấp danh mục kiểm tra thuộc tính. Người đánh giá cảm quan sau đó có thể có cuộc thảo luận được kiểm soát bởi người phụ trách hội đồng.

5.4.2.4 Diễn giải kết quả

Các kết quả phải được so sánh để tạo ra danh sách các thuật ngữ mô tả áp dụng cho mẫu, dựa trên tần suất sử dụng từng từ mô tả. Khuyến cáo thảo luận mở vào lúc người đánh giá kết luận.

5.4.3 Profile mô tả định lượng

5.4.3.1 Định nghĩa

Đây là phép thử trong đó tất cả các thuộc tính cảm quan của sản phẩm được mô tả và giá trị cường độ của chúng được định lượng bởi hội đồng đã được huấn luyện.

Profile mô tả định lượng thu được là của hội đồng cụ thể và nhóm sản phẩm cụ thể. Chúng không thể được diễn giải bởi các nhóm khác nếu không có chuẩn đối chứng.

Có một số phương pháp để thiết lập profile cảm quan mô tả định lượng, trong đó một số kỹ thuật đã đăng ký nhãn hiệu.

5.4.3.2 Ứng dụng

Phép thử này được khuyến cáo:

- để sử dụng trong việc phát triển sản phẩm mới,
- để thiết lập bản chất của sự khác biệt giữa các sản phẩm,
- để sử dụng trong kiểm soát chất lượng,
- để cung cấp dữ liệu cảm quan nhằm liên kết với dữ liệu thiết bị và dữ liệu người tiêu dùng.

5.4.3.3 Người đánh giá

Cần phải có hội đồng gồm những người đánh giá hoặc gồm những chuyên gia đánh giá cảm quan, được huấn luyện đặc biệt về phương pháp và trên nhóm sản phẩm được thử nghiệm. Xem TCVN 12389 (ISO 8586) và ISO 13299 để biết chi tiết liên quan đến việc lựa chọn người đánh giá và quy trình huấn luyện phép thử mô tả.

Theo từng trường hợp:

- người phụ trách hội đồng phải hướng dẫn việc huấn luyện,
- người phụ trách hội đồng phải kiểm soát cuộc thảo luận và thiết lập sự đồng thuận.

5.4.3.4 Cách tiến hành

Tiến hành thử nghiệm (hoặc huấn luyện) sơ bộ với một loạt các sản phẩm cần thử nghiệm, để thiết lập các đặc tính cảm quan quan trọng trong việc mô tả đặc điểm và cách phân biệt chúng. Kết quả của các thử nghiệm này được sử dụng để xây dựng bộ các thuật ngữ mô tả được sử dụng và để thiết lập các quy trình thử nghiệm nhằm trình bày và kiểm tra mẫu. Hội đồng sau đó được huấn luyện về phương pháp luận và đặc biệt là việc sử dụng thuật ngữ. Ở giai đoạn này nên có bộ vật liệu chuẩn là các hợp chất tinh khiết hoặc các sản phẩm tự nhiên tạo ra các điểm mùi hoặc hương vị đặc biệt hoặc có cấu trúc đặc biệt hoặc có những đặc tính quan sát đặc biệt.

Trong buổi đánh giá, người đánh giá kiểm tra các mẫu dựa trên các thuật ngữ, cho điểm từng thuộc tính có mặt trên thang đo cường độ.

Thông thường phải lưu ý thứ tự mà các yếu tố cảm nhận được, bao gồm cả hậu vị và để cho điểm tổng thể của mùi và hương.

5.4.3.5 Diễn giải kết quả

Kết quả bao gồm điểm cường độ đối với từng thuộc tính. Có hai phương pháp cơ bản để xử lý dữ liệu.

Trong các phương pháp profile đồng thuận, ngay sau khi người đánh giá hoàn thành công việc, người phụ trách hội đồng lập bảng kết quả và bắt đầu một cuộc thảo luận để giải quyết sự khác biệt. Theo kết quả thảo luận và sau khi kiểm tra lại mẫu, nếu cần, người phụ trách hội đồng đưa ra quyết định của nhóm dựa trên profile.

Trong các phương pháp profile mô tả khác, không có thảo luận và profile thu được là dãy điểm trung bình được mỗi người đánh giá ấn định cho từng từ mô tả.

Kết quả trung bình có thể được so sánh về mặt thống kê, ví dụ bằng cách sử dụng phân tích phương sai. Ngoài ra, đối với tất cả các phương pháp phân tích mô tả, còn có kỹ thuật phân tích đa biến.

5.4.4 Profile đồng thuận

5.4.4.1 Định nghĩa

Đây là phương pháp mà người đánh giá thực hiện đánh giá đơn lẻ các mẫu và sau đó chia sẻ quan điểm cá nhân của họ để đạt được sự đồng thuận về các thuộc tính khác nhau, thứ tự xuất hiện và cường độ của chúng.

5.4.4.2 Ứng dụng

Phép thử này đặc biệt được khuyến cáo sử dụng trong việc phát triển các sản phẩm mới.

5.4.4.3 Cách tiến hành

Người đánh giá thực hiện đánh giá đơn lẻ mỗi mẫu một lần và ghi lại các thuộc tính, cường độ của các thuộc tính và thứ tự xuất hiện của chúng. Kết thúc việc đánh giá sản phẩm bằng nhóm các kết quả được thu được. Người phụ trách hội đồng tổ chức một cuộc thảo luận chung để đưa ra một profile đồng thuận về các thuộc tính, cường độ và mọi khía cạnh khác (ví dụ thứ tự của các thuộc tính). Mẫu chuẩn có thể được cung cấp để làm phong phú cuộc thảo luận. Một khi hội đồng đạt được sự đồng thuận thì sẽ tạo ra được "profile" cuối cùng. Khi không có sự đồng thuận, người phụ trách hội đồng có thể quyết định xem xét sản phẩm vào buổi đánh giá mới hoặc ghi lại sự khác nhau giữa những người đánh giá. Quá trình này được lặp lại cho đến khi tất cả các sản phẩm được đánh giá.

5.4.4.4 Người đánh giá

Hội đồng cần được huấn luyện để xác định các chỉ tiêu và có thể ra quyết định về các thuộc tính và cường độ khi họ nếm.

5.4.4.5 Diễn giải kết quả

Đối với mỗi thuộc tính, kết quả phải bao gồm một điểm duy nhất (điểm đã được đồng thuận). Có thể một người đánh giá không đồng ý với nhóm thì điều này phải được đề cập đến.

5.4.5 Profile lựa chọn tự do

5.4.5.1 Định nghĩa

Đây là phương pháp mô tả mà mỗi người đánh giá chọn riêng danh sách các thuộc tính cảm quan của mình và định lượng cường độ theo thang đo đã được xác định trước hoặc thang đo tự chọn.

5.4.5.2 Ứng dụng

Phương pháp hữu ích để tạo ra các thuộc tính, bản đồ cảm nhận về không gian sản phẩm và để nghiên cứu thị trường.

5.4.5.3 Người đánh giá

Người đánh giá có thể phải ngửi, nếm, chạm, nghe và có từ vựng toàn diện để mô tả các sản phẩm. Không cần huấn luyện cụ thể.

5.4.5.4 Cách tiến hành

Người đánh giá thu thập các thuộc tính riêng lẻ và chọn một trong hai thang đo mà họ muốn hoặc thang đo đã được xác định trước. Trong buổi đánh giá thứ hai, người đánh giá phải sử dụng các thuộc tính của mình một cách nhất quán cho mỗi sản phẩm mà họ phải so sánh và thực hiện đánh giá cường độ của sản phẩm theo các điều kiện chuẩn.

5.4.5.5 Diễn giải kết quả

Dữ liệu thường được diễn giải bằng phép phân tích đa chiều thích hợp (ví dụ: phân tích Procrustes khái quát) để tạo bản đồ cảm nhận đồng thuận về kích thước quan trọng và để đánh giá mức độ thống nhất dữ liệu của mỗi người đánh giá với bản đồ đồng thuận có nguồn gốc thống kê.

5.4.6 Profile mô tả nhanh

5.4.6.1 Định nghĩa

Đây là phương pháp mô tả trong đó người đánh giá đơn lẻ tạo ra các thuộc tính cho một tập hợp các sản phẩm được trình bày đồng thời và so sánh các sản phẩm bằng cách xếp hạng cường độ của thuộc tính.

CHÚ THÍCH Đây là biến mô tả tự do được phân biệt bằng cách sử dụng xếp hạng hạng chứ không phải là xếp loại.

5.4.6.2 Ứng dụng

Phương pháp này hữu ích để tạo ra các thuộc tính, bản đồ cảm nhận không gian sản phẩm và để nghiên cứu thị trường.

5.4.6.3 Người đánh giá

Người đánh giá có thể phải ngửi, nếm, chạm, nghe và có từ vựng toàn diện để mô tả sản phẩm. Không cần huấn luyện cụ thể.

5.4.6.4 Cách tiến hành

Trước khi thử nghiệm, cần hướng dẫn cho người đánh giá về phương pháp, cách mô tả và phân biệt giữa các sản phẩm.

Tất cả các sản phẩm phải được trình bày đồng thời cho mỗi người đánh giá. Người đánh giá thu thập các thuộc tính tại các buổi đánh giá độc lập và chọn một trong hai thang đo họ muốn hoặc thang đo đã được xác định trước. Sau đó, người đánh giá sử dụng các thuộc tính của mình cho mỗi sản phẩm họ phải so sánh và thực hiện đánh giá cường độ của sản phẩm ở các điều kiện chuẩn.

5.4.6.5 Diễn giải kết quả

Dữ liệu thường được diễn giải bằng phép phân tích đa chiều phù hợp như phân tích Procrustes khái quát. Kết quả luôn ở dạng bản đồ.

5.4.7 Profile mô tả độ chệch so với mẫu chuẩn

5.4.7.1 Định nghĩa

Đây là phương pháp mô tả trong đó người đánh giá so sánh hai sản phẩm được trình bày đồng thời với nhau về mỗi thuộc tính của một danh sách chung và cho điểm mức độ khác biệt.

5.4.7.2 Ứng dụng

Phương pháp này đặc biệt phù hợp khi có mẫu chuẩn rõ ràng hoặc khi khó phân biệt các mẫu thử. Phép thử này chủ yếu được sử dụng để kiểm soát chất lượng và đảm bảo chất lượng.

5.4.7.3 Người đánh giá

Thông thường, người đánh giá được lựa chọn và huấn luyện các sản phẩm cần được thử nghiệm như trong profile mô tả định lượng.

5.4.7.4 Cách tiến hành

Người đánh giá nhận được một bộ gồm hai mẫu. Các mẫu được người đánh giá so sánh trực tiếp với mẫu chuẩn và cho điểm theo mẫu chuẩn (mẫu chuẩn là điểm trung gian trên thang đo) hoặc, các mẫu được so sánh bằng cách tính toán sự khác nhau giữa các điểm được đưa ra cho mẫu chuẩn và các mẫu thử.

5.4.7.5 Diễn giải kết quả

Tiến hành phân tích dữ liệu dựa trên sự khác biệt giữa các mẫu thử và mẫu chuẩn.

5.4.8 Phép thử TDS

5.4.8.1 Định nghĩa

Đây là phương pháp mô tả trong đó người đánh giá liên tục chỉ ra từ danh sách các thuộc tính chung tính chất cảm quan chiếm ưu thế và đôi khi đánh giá chúng theo thời gian trong khi sản phẩm đang được đánh giá.

5.4.8.2 Ứng dụng

Phương pháp này đặc biệt phù hợp khi có đặc tính có khả năng kéo dài (ví dụ sản phẩm có chất tạo ngọt) hoặc đối với những sản phẩm phức tạp, khi profile mô tả tính không đủ để diễn giải sự khác biệt có thể cảm nhận được.

5.4.8.3 Người đánh giá

Thông thường, việc huấn luyện người đánh giá phải được đặc biệt định hướng để nhận biết các thuộc tính cảm quan khác nhau nhằm cải thiện việc lựa chọn các cảm giác chủ đạo.

5.4.8.4 Cách tiến hành

Người đánh giá nhận được một danh sách thuộc tính và được yêu cầu xác định xem cảm giác nào chiếm ưu thế trong khi sản phẩm đang được đánh giá. Vì là tùy chọn, nên việc chọn cường độ của thuộc tính chủ đạo có thể thực hiện bằng cách cho điểm.

5.4.8.5 Diễn giải kết quả

Dữ liệu phải bao gồm tỷ lệ của từng thuộc tính được chọn là ưu thế tại mỗi thời điểm. Dữ liệu thường được chuyển đổi thành đường chuẩn, trong đó thời gian là trục x.

6 Phân tích kết quả

6.1 Yêu cầu chung

Điều này đưa ra chỉ dẫn chung các phương pháp thích hợp được sử dụng để phân tích thống kê các kết quả của phép thử cảm quan. Các chi tiết khác về các phép thử cụ thể có thể tìm thấy trong các tiêu chuẩn thích hợp được liệt kê trong Thư mục tài liệu tham khảo. Các thuật ngữ thống kê chữ in đậm được diễn giải trong Phụ lục A, phù hợp với TCVN 9597-1 (ISO 3534-1), TCVN 8244-2 (ISO 3534-2) và ISO 3534-3.

6.2 Phép thử phân biệt

6.2.1 Yêu cầu chung

Mục đích của phép thử phân biệt được mô tả trong 5.2 là để xác định xem có sự khác biệt nào có thể nhận ra giữa hai sản phẩm A và B hay không (hoặc ưa thích sản phẩm nào hơn trong hai sản phẩm). Phép phân tích dựa trên kết quả thử nghiệm của người đánh giá trong từng nhóm cụ thể, ví dụ những người thích A, những người thích B, hoặc những người lựa chọn đúng mẫu khác.

Tiêu chuẩn này đề cập chi tiết từng phương pháp cũng như mô tả cách sử dụng để đảm bảo sự giống nhau khi được yêu cầu.

Thử nghiệm về sự giống nhau sử dụng các phép thử khác biệt, dùng các phương pháp thống kê khác nhau với thử nghiệm tính tương đương.

Tuy nhiên, mức độ có ý nghĩa của rủi ro α và β là như nhau đối với hai dạng phép thử. Rủi ro α là rủi ro bác bỏ không đúng giả thuyết không và rủi ro β là rủi ro chấp nhận không đúng giả thuyết không.

Trong thực tế, việc xây dựng giả thuyết không và việc tính các rủi ro α - và β - thay đổi theo bản chất của phép thử phân biệt: 2-afc, hai-ba, tam giác, v.v....

Đối với các phép thử phân biệt, cần chọn mức rủi ro α để thử sự khác biệt và chọn mức rủi ro β để thử tính giống nhau.

Khoảng tin cậy về tỷ lệ người thử (proportion of discriminators) có thể phân biệt các mẫu là tính được, nếu có yêu cầu.

6.2.2 Phép thử số sánh cặp đôi [xem TCVN 4831 (ISO 5495)]

6.2.2.1 Diễn giải thống kê

Có hai hình thức có thể có ở phép thử này. Hình thức thứ nhất liên quan đến việc phát hiện và xác định sự khác biệt cụ thể giữa hai sản phẩm; hình thức thứ hai liên quan đến việc ưu tiên cho một trong hai sản phẩm.

Trong cả hai trường hợp, giả thuyết không là không thể phân biệt giữa hai sản phẩm (kể cả cường độ và thứ tự ưu tiên). Về mặt định lượng, giả thuyết không là có xác suất bằng nhau (1/2) trong đó người đánh giá chọn ngẫu nhiên từ thùng chứa mẫu A hoặc mẫu B.

Việc diễn giải các kết quả dựa vào số lượng người tham gia chọn A (hoặc B) có cường độ lớn hơn hoặc được ưu tiên hơn, phụ thuộc vào giả thuyết thay thế trái ngược với giả thuyết không. Phụ thuộc vào bản chất của giả thuyết thay thế được quy định trước khi thực hiện phép thử, phép thử hai phía hoặc phép thử một phía.

6.2.2.2 Phép thử hai phía

Phép thử hai phía là phép thử trong đó chỉ đơn giản muốn tìm hiểu xem có sự khác biệt về cường độ giữa hai sản phẩm hay không (phép thử cường độ), hoặc nếu một trong hai sản phẩm được thích hơn (phép thử ưa thích). Giả thuyết thay thế được viết là $PA \neq PB$ (nghĩa là $PA > PB$ hoặc $PA < PB$).

Giả thuyết không bị bác bỏ nếu khi thử nghiệm sự khác biệt, số lượt chọn cho một mẫu lớn hơn hoặc ít nhất bằng số nêu trong các bảng thống kê TCVN 4831 (ISO 5495), tương ứng với số lượng người đánh giá và mức rủi ro được chọn để thử nghiệm) và khi thử nghiệm tính giống nhau, nếu số lượt chọn một mẫu nhỏ hơn hoặc bằng số nêu trong các bảng thống kê [TCVN 4831 (ISO 5495), tương ứng với số lượng người đánh giá và mức rủi ro được chọn cho phép thử].

Nếu trong trường hợp, kết luận có sự khác biệt đáng kể giữa hai sản phẩm và nếu đa số lượt chọn sản phẩm A thì kết luận rằng, đối với các đặc tính cần xác định, A có cường độ lớn hơn B đáng kể (hoặc được ưa thích hơn, nếu dựa trên số lượt chọn người đánh giá).

6.2.2.3 Phép thử một phía (một chiều)

Phép thử một phía là phép thử với yêu cầu phát hiện nếu một trong những sản phẩm được chọn đặc biệt (ví dụ A) có cường độ lớn hơn những sản phẩm khác; giả thuyết thay thế là $PA > 1/2$. Phép thử một phía chỉ phù hợp nếu mọi kết quả theo hướng ngược lại sẽ không diễn giải được như phản ánh thực tế nhưng chỉ đơn thuần là kết quả ngẫu nhiên mà không nghi ngờ về giả thuyết không.

Giả thuyết không bị bác bỏ nếu số lượt chọn một mẫu lớn hơn hoặc ít nhất bằng số lượt chọn nêu trong các bảng thống kê [TCVN 4831 (ISO 5495), tương ứng với số lượng người đánh giá và mức rủi ro được chọn cho phép thử]. Nếu trong trường hợp này, kết luận A có ưu thế hơn B (về cường độ) được hội đồng công nhận là có ý nghĩa.

Không có kết luận nào được đưa ra khi số lượng tối đa câu trả lời đúng dưới n/2.

6.2.3 Phép thử tam giác [xem TCVN 11184 (ISO 4120)]

Giả thuyết không là không thể phân biệt các sản phẩm. Trong trường hợp này, xác suất P mẫu khác với hai sản phẩm khác tương đương với $P_0 = 1/3$. Trong các thuật ngữ về thống kê, giả thuyết không H_0 được biểu thị bằng $P_0 = 1/3$.

TCVN 12387:2018

Người giám sát phép thử muốn biết liệu có thể phân biệt giữa hai sản phẩm hay không, vì vậy sẽ bác bỏ giả thuyết không để có lợi cho giả thuyết thay thế $P > 1/3$. Số lượng câu trả lời đúng phải được so sánh với dữ liệu được trình bày trong bảng của TCVN 11184 (ISO 4120) và mức rủi ro α phải được chọn khi thử nghiệm sự khác biệt và mức rủi ro β phải được chọn cho phép thử tương tự. Tính được khoảng tin cậy về tỷ lệ người đánh giá có thể phân biệt được các mẫu, nếu được yêu cầu.

Khi tiến hành phép thử tương tự, không đưa ra kết luận đối với số lượng tối đa câu trả lời đúng dưới $n/3$.

6.2.4 Phép thử hai-ba [xem TCVN 11185 (ISO 10399)]

Giả thuyết không là không thể phân biệt các sản phẩm. Trong trường hợp này, xác suất nhận biết mẫu giống hệt với mẫu chuẩn là $P_0 = 1/2$. Trong các thuật ngữ về thống kê, giả thuyết không H_0 được biểu thị bằng $P_0 = 1/2$.

Số lượng câu trả lời đúng phải được so sánh với dữ liệu được trình bày trong bảng của TCVN 11185 (ISO 10399) và mức rủi ro được chọn để thử nghiệm, cần chọn mức rủi ro α khi thử nghiệm sự khác biệt và mức rủi ro β khi thử nghiệm sự giống nhau.

6.2.5 Phép thử chọn hai trong năm

Giả thuyết không là $P_0 = 1/10$. Phép thử là một phía và giả thuyết thay thế là $P > 1/10$.

6.2.6 Phép thử "A- không A" (xem ISO 8588)

Số lượng câu trả lời "A" và số câu trả lời "không A" được tính riêng cho toàn bộ các mẫu được các nhà phân tích cảm quan biết là "A" và các mẫu được biết "không A", được nêu trong bảng 2 x 2. Phép thử độc lập chi-bình phương [χ^2] hoặc phép thử chuẩn Fisher có thể được sử dụng để xác định xem tỷ lệ câu trả lời "A" và "không A" có khác nhau đối với hai loại mẫu.

Phép thử một phía, với giả thuyết không là hai tỷ lệ là bằng nhau và giả thuyết thay thế đó là tỷ lệ câu trả lời "A" nhiều hơn các mẫu được biết là "A".

6.2.7 Phép thử terad

Giả thuyết không là không thể phân biệt các sản phẩm. Trong trường hợp phép thử bốn mẫu không quy định, xác suất P dự đoán chính xác là 1/3; xác suất dự đoán chính xác bằng 1/6 trong trường hợp phép thử bốn mẫu được quy định.

Để biết thêm chi tiết, xem ASTM E3009-15.

6.2.8 Xử lý câu trả lời "không có sự khác biệt" trong các phép thử phân biệt

Trong các phép thử phân biệt, có thể có câu trả lời "không có sự khác biệt". Tuy nhiên, điều này có thể được quy định là không được phép (kỹ thuật lựa chọn bắt buộc). Làm sao sử dụng câu trả lời của người đánh giá nếu họ trả lời "không có sự khác biệt", nhưng những phán đoán tuy vậy vẫn chính xác. Điều bất lợi ở đây là chúng có thể gây trở ngại cho người đánh giá trung thực muốn ghi "không có sự khác biệt".

Nếu kết quả "không có sự khác biệt" là được phép thì số lượng người đánh giá báo cáo "không có sự khác biệt" được ghi lại và phép phân tích thống kê chỉ sử dụng các kết quả từ những người báo cáo có sự khác biệt. Kết luận được biểu thị liên quan đến người đánh giá khi họ biểu thị sự ưa thích hoặc báo cáo về sự khác biệt.

6.2.9 Ảnh hưởng có hệ thống

Phải tiến hành kiểm tra để xem nếu có bất kỳ sự ảnh hưởng có hệ thống nào trong mối quan hệ, ví dụ,

- theo thứ tự thử nghiệm đối với người đánh giá đơn lẻ, và
- theo trình tự trình bày của một số người đánh giá.

6.2.10 Phép thử tuần tự [xem TCVN 11186 (ISO 16820)]

Trong phép thử tuần tự, kết quả tích lũy từ phép thử phân biệt được kiểm tra liên tục trong quá trình thử nghiệm. Phép thử sẽ dừng ngay khi có quyết định. Phép thử này thường được quyết định sau khi yêu cầu ít thử nghiệm hơn bằng cách sử dụng cỡ mẫu đã được xác định trước.

6.3 Phép thử sử dụng thang đo và phân nhóm

6.3.1 Yêu cầu chung

Xem TCVN 5090 (ISO 4121) để biết chi tiết.

Việc lựa chọn phương pháp thống kê để phân tích cảm quan sử dụng bất kỳ phép thử nào được nêu trong 5.3 tùy thuộc vào mục đích của phép thử và số lượng sản phẩm cần thử nghiệm. Các điều dưới đây cung cấp thông tin về các phương pháp thống kê được sử dụng. Để biết thêm chi tiết trong từng ngữ cảnh cụ thể của từng phép thử, cần phải tham khảo sách thống kê có liên quan hoặc tư vấn từ nhà thống kê.

6.3.2 Phân loại

Các kết quả thu được đối với một loại sản phẩm có thể được tổng hợp theo tần suất đối với mỗi hạng. Sau đó có thể sử dụng phép thử khi bình phương (χ^2) để so sánh sự phân bố của hai hoặc nhiều loại sản phẩm trong các hạng khác nhau, nghĩa là để thử giả thuyết không về sự phân bố là như nhau được so sánh với giả thuyết thay thế là khác nhau.

6.3.3 Phân cấp [xem TCVN 5090 (ISO 4121)]

Dữ liệu có thể được tổng hợp để phân loại. Kết quả có thể được tổng hợp cách khác bằng cách lấy kết quả ở giữa và các sản phẩm có thể được so sánh thống kê bằng các phép thử tổng hạng (rank sum test), mặc dù có thể cần phải điều chỉnh để cho phép xếp hạng với số lượng lớn. Nếu dữ liệu từ mẫu được sử dụng làm cơ sở để quyết định đối với số lượng lớn ("lò hàng"), xem TCVN 7790 (ISO 2859 (tất cả các phần) và TCVN 8243 (ISO 3951) (tất cả các phần) để có được các đặc tính của phương án lấy mẫu thích hợp.

6.3.4 Xếp hạng (xem TCVN 11183 (ISO 8587))

Khi các mẫu được một số người đánh giá xếp hạng như trong 5.3.5, có thể tiến hành các phép thử thống kê để xác định xem các mẫu có sự khác biệt đáng kể nào hay không (phép thử tổng hạng). Cũng có thể tiến hành thử nghiệm để xác định xem mẫu cụ thể nào có mức xếp hạng cao hơn hoặc thấp hơn đáng kể so với các mẫu khác.

Điểm xếp hạng có thể được nhóm lại thành các nhóm đồng nhất.

6.3.5 Xếp loại

Đối với việc xếp loại trên thang rời rạc với số điểm nhỏ, kết quả cho một mẫu có thể được xử lý như đối với cách phân loại. Dữ liệu liên tục, hoặc dữ liệu rời rạc với số điểm lớn có thể được nhóm lại và tập hợp theo tần suất trong từng khoảng.

Khi đánh giá nhiều mẫu, sử dụng phương pháp phi tham số để so sánh sự phân bố thu được.

Nếu dữ liệu đáp ứng các điều kiện cho điểm hoặc là chúng vẫn giữ nguyên hoặc sau đó chuyển đổi thì có thể sử dụng phương pháp trong 5.3.6.

6.3.6 Cho điểm

Các kết quả thu được đối với một mẫu có thể được tổng hợp theo kết quả ở giữa hoặc trung bình (trung bình số học) với một số phương pháp phân tán (ví dụ, độ lệch cấp hoặc độ lệch chuẩn).

Nếu chỉ có hai mẫu có liên quan và giả thuyết phân bố chuẩn các điểm là hợp lý thì có thể sử dụng phép thử t (xem ISO 2854). Nếu điểm thu được từ hơn hai mẫu thì tiến hành quy trình chuẩn là phân tích phương sai.

Nếu việc phân bố điểm số từ từng mẫu có vẻ không chuẩn, nên sử dụng phương pháp phân bố tự do.

6.4 Phép phân tích hoặc phép thử mô tả

Xem ISO 13299 để biết chi tiết.

7 Báo cáo thử nghiệm

Báo cáo thử nghiệm phải ghi rõ các thông tin sau:

- mục đích của phép thử;
- thông tin về sản phẩm, bao gồm mô tả mẫu (nguồn gốc, tính đại diện, phương pháp bảo quản và bao gói, v.v...);
- phương pháp chuẩn bị mẫu;
- phương pháp luận đã sử dụng: viện dẫn tiêu chuẩn này hoặc mô tả đầy đủ phép thử chưa được chuẩn hóa;
- ngày thử nghiệm và thời gian thử;
- điều kiện vật chất để đạt được phép thử (địa điểm, phương pháp và thứ tự trình bày mẫu, câu hỏi, phương pháp ghi câu trả lời);
- đặc điểm của nhóm người đánh giá:
 - hiệu quả;
 - phương pháp tuyển chọn;
 - trình độ và/hoặc tính đại diện;
- kết quả:
 - diễn giải kết quả;
 - mô hình thống kê đã sử dụng;
 - cách thức và điều kiện truy cập dữ liệu thô;
- kết luận;
- các khuyến cáo, nếu khách hàng yêu cầu;
- tên và thông tin liên lạc của người chịu trách nhiệm thực hiện phân tích cảm quan.

Khi có liên quan, phải đề cập bất kỳ thông tin bổ sung nào được đưa ra trong TCVN ISO/IEC 17025 (ISO/IEC 17025) và EA-4/09.

Phụ lục A

(Tham khảo)

Thuật ngữ về thống kê

A.1 Kiểm tra giả thuyết – Giả thuyết không

Trong phân tích cảm quan, phép thử được thực hiện để đánh giá một giả thuyết. Giả thuyết này cần được nêu rõ ràng và không có sự mơ hồ trước khi thực hiện phép thử. Giả thuyết cần được xây dựng để có thể chuyển được sang ngôn ngữ thống kê như giả thuyết không H_0 .

Giả thuyết không thông thường là giả thuyết cho rằng không có sự khác biệt về cường độ của một tính chất cảm quan (hoặc không có sự ưa thích cho một trong số những sản phẩm đó) giữa các sản phẩm.

A.2 Giả thuyết thay thế

Giả thuyết thay thế là giả thuyết đã nêu rõ ràng nó sẽ được chấp nhận nếu giả thuyết không bị bác bỏ. Nếu giả thuyết không H_0 là $P = P_0$ thì giả thuyết H_1 thay thế có thể là hai phía ($P \neq P_0$) hoặc một phía (ví dụ $P > P_0$). Ví dụ về cả hai loại được nêu trong 6.2.2.

A.3 Ý nghĩa và mức ý nghĩa

Khi phân tích kết quả của phép thử, có hai kết luận có thể xảy ra:

- Giả thuyết không không bị bác bỏ;
- Giả thuyết không bị bác bỏ.

Vì mọi phép thử được thực hiện bởi một một lượng hạn chế các chuyên gia đánh giá, nên kết luận bác bỏ giả thuyết không (ủng hộ cho giả thuyết thay thế) liên quan đến rủi ro. Mức ý nghĩa là xác suất (hoặc giá trị xác suất tối đa) việc bác bỏ giả thuyết không khi giả thuyết không là đúng. Đây được gọi là "rủi ro alpha".

Sự logic của phép thử nghiệm ý nghĩa thống kê kinh điển đòi hỏi quyết định thực hiện trước rủi ro alpha có thể chấp nhận được. Thông thường, giá trị được gán trước cho mức ý nghĩa là $\alpha = 0,05$ (5 %) hoặc $\alpha = 0,01$ (1 %). Hầu hết các bảng thống kê được sử dụng để diễn giải kết quả thử nghiệm bao gồm hai mức ý nghĩa này. Điều quan trọng cần lưu ý là giả thuyết không có thể bị bác bỏ "ở mức 5 %" và "không bị bác bỏ" ở mức 1 %".

Nếu giả thuyết không bị bác bỏ "ở mức 1 %" thì nó cũng bị bác bỏ đáng kể "ở mức 5 %". Điều này diễn giải tại sao việc biểu thị "có ý nghĩa" đối với mức 5 % và "rất có ý nghĩa" đối với mức 1 % đôi khi vẫn được sử dụng.

A.4 Kết luận sai lầm: Loại II - Hiệu quả

Nếu phép thử không dẫn đến việc bác bỏ giả thuyết không, điều này không chứng minh được giả thuyết không là đúng. Điều này chỉ có nghĩa là, từ các thông tin bị hạn chế có sẵn (thử với n người đánh giá), không có đủ bằng chứng để bác bỏ giả thuyết không (ở mức ý nghĩa đã chọn). Số lượng thông tin càng lớn (lớn hơn n người đánh giá) thì càng có lý khi bác bỏ giả thuyết không khi bị sai; hiệu quả của việc phép thử tăng lên cùng với số lượng người đánh giá tham gia phép thử. Ví dụ, trong trường hợp phép thử ưa thích (6.2.2) được thực hiện với 20 người đánh giá, giả thuyết không $P_0 = 1/2$ có thể không bị bác bỏ (kết luận là không có sự ưa thích có ý nghĩa đối với một trong hai sản phẩm), trong khi đó, với cùng tỷ lệ lựa chọn hai mẫu này, nếu phép thử được thực hiện với 100 người đánh giá thì sự ưa thích có ý nghĩa đối với một sản phẩm có thể chứng minh được.

Sai lầm loại II (phụ thuộc vào mức ý nghĩa được chọn) là xác suất (ký hiệu β) do nhầm lẫn không bác bỏ giả thuyết không khi giả thuyết thay thế (giả thuyết đối ngẫu) là thực sự đúng.

Nếu giả thuyết không và giả thuyết thay thế có thể được xác định bằng các giá trị của tham số như trong các phép thử phân biệt (phép thử so sánh cặp đôi, phép thử tam giác, phép thử hai-ba...) thì sai lầm loại II có thể được tính như theo hàm của tham số này. Đối với phép thử mà giả thuyết không và giả thuyết thay thế có thể không được xác định bằng cách sử dụng các giá trị của tham số (phép thử đánh giá, phân loại), thường không thể tính sai lầm loại II.

Đối với các mục đích khác nhau, cây quyết định có thể hữu ích trong việc xác định lựa chọn phương pháp và kết quả. Ví dụ được nêu trong Tài liệu tham khảo [28].

Thư mục tài liệu tham khảo

- [1] ISO 2854, *Statistical interpretation of data – Techniques of estimation and tests relating to means and variances*
- [2] TCVN 7790 (ISO 2859) (tất cả các phần), *Quy trình lấy mẫu để kiểm tra định tính*
- [3] TCVN 9597-1 (ISO 3534-1), *Phương pháp thống kê – Hướng dẫn đánh giá sự phù hợp với yêu cầu quy định – Phần 1: Nguyên tắc chung*
- [4] TCVN 8244-2 (ISO 3534-2), *Từ vựng và ký hiệu – Phần 2: Thống kê ứng dụng.*
- [5] TCVN 8244-3 (ISO 3534-3), *Thống kê học – Từ vựng và ký hiệu – Phần 3: Thiết kế thực nghiệm.*
- [6] TCVN 8243 (ISO 3951) (tất cả các phần), *Quy trình lấy mẫu để kiểm tra định lượng*
- [7] ISO 3972, *Sensory analysis – Methodology – Method of investigating sensitivity of taste*
- [8] TCVN 11184 (ISO 4120) , *Phân tích cảm quan – Phương pháp luận – Phép thử tam giác.*
- [9] TCVN 5090 (ISO 4121) *Phân tích cảm quan – Hướng dẫn sử dụng các thang đo định lượng đặc trưng*
- [10] TCVN 4831 (ISO 5495) *Phân tích cảm quan – Phương pháp luận – Phép thử so sánh cặp đôi*
- [11] ISO 5497, *Sensory analysis – Methodology – Guidelines for the preparation of samples for which direct sensory analysis is not feasible*
- [12] TCVN 12389 (ISO 8586), *Phân tích cảm quan – Hướng dẫn chung để lựa chọn, huấn luyện và giám sát người đánh giá được lựa chọn và chuyên gia đánh giá cảm quan*
- [13] TCVN 11183 (ISO 8587) *Phân tích cảm quan – Phương pháp luận – Xếp hạng*
- [14] ISO 8588, *Sensory analysis – Methodology – "A" - "not A" test*
- [15] TCVN 12390 (ISO 8589), *Phân tích cảm quan – Hướng dẫn chung về thiết kế phòng thử*
- [16] TCVN 11185 (ISO 10399) *Phân tích cảm quan – Phương pháp luận – Phép thử hai-ba*
- [17] ISO 11132, *Sensory analysis – Methodology – Guidelines for monitoring the performance of a quantitative sensory panel*

- [18] ISO 11136, *Sensory analysis – Methodology – General guidance for conducting hedonic tests with consumers in a controlled area*
- [19] ISO 13299, *Sensory analysis – Methodology – General guidance for establishing a sensory profile*
- [20] TCVN 12388-1 (ISO 13300-1), *Phân tích cảm quan – Hướng dẫn chung đối với nhân viên phòng đánh giá cảm quan – Phần 1: Trách nhiệm của nhân viên*
- [21] TCVN 12388-2 (ISO 13300-2), *Phân tích cảm quan – Hướng dẫn chung đối với nhân viên phòng đánh giá cảm quan – Phần 2: Tuyển chọn và huấn luyện người phụ trách hội đồng*
- [22] TCVN 11186 (ISO 16820), *Phân tích cảm quan – Phương pháp luận – Phép thử tuần tự*
- [23] TCVN 10565-3:2015 (ISO 22935-3:2009), *Sữa và sản phẩm sữa – Phân tích cảm quan – Phần 3: Hướng dẫn về phương pháp đánh giá sự phù hợp của các chỉ tiêu cảm quan với các quy định của sản phẩm bằng phương pháp cho điểm*
- [24] TCVN ISO/IEC 17025 (ISO/IEC 17025), *Yêu cầu chung về năng lực của phòng thử nghiệm và hiệu chuẩn*
- [25] ASTM E3009-15, *Standard Test Method for Sensory Analysis – Tetrad Test*
- [26] EA-4/09 Accreditation for sensory testing laboratories
- [27] VARELAP., ARESG., eds. *Novel techniques in sensory Characterization and consumer profiling*. CRC. Press, 2014
- [28] LAWLESS H.T., HEYMANN H. *Sensory evaluation of food. Principles and practices*. Springer.
-