

# CHÍNH XÁC HÓA MỘT KHÁI NIỆM TRONG NGHIÊN CỨU ĐỊNH LƯỢNG

PHẠM XUÂN GIANG, NGUYỄN THỊ PHƯƠNG THẢO  
Khoa Quản trị kinh doanh, Trường Đại học Công nghiệp thành phố Hồ Chí Minh  
phamxuanguiang@iuh.edu.vn

**Tóm tắt.** Trong nghiên cứu định lượng, chỉ tiêu trung bình (mean) được dùng để mô tả ý kiến chung. Hiện nay có ba cách phân tổ để xếp loại giá trị mean nhằm đánh giá đối tượng nghiên cứu. Vì ba cách có sự khác nhau nên cùng một nguồn dữ liệu nhưng sử dụng cách khác nhau, nhiều khi lại cho ra kết quả phân tổ và xếp loại giá trị mean không giống nhau. Chính vì vậy, bài báo đã trình bày nội dung của ba cách, sau đó đối chiếu với cơ sở khoa học về phân tổ thống kê để khẳng định cách thứ ba là đúng. Hy vọng rằng, đây sẽ là cách chính thức được sử dụng rộng rãi trong nghiên cứu khoa học của giảng viên và sinh viên khối ngành kinh tế thuộc trường Đại học Công nghiệp thành phố Hồ Chí Minh, thay vì cách thứ nhất như hiện nay. Và sau cùng, cách phân tổ giá trị mean theo thang đo Likert 7 mức độ cũng được trình bày chi tiết trong bài báo.

**Từ khóa.** Nghiên cứu định lượng; Giá trị mean; Thang đo Likert; Phân tổ thống kê; Biến liên tục; Biến rời rạc; Tiêu thức; Biện chứng; Logic.

## ACCURACY OF A CONCEPT IN QUANTITATIVE RESEARCH

**Abstract.** In quantitative research, mean are used to describe opinions. There are currently three ways to classify mean to evaluate the research object. Because the three ways are different, the same data source but using different ways, sometimes results in disaggregation and classification of mean are not the same. Therefore, the paper presented the content of the three ways, then collated with the scientific basis of statistical clustering to assert the third way is right. Hopefully, this will be the most widely used way in scientific research of lecturers and economics students of the University Industry of Ho Chi Minh City, instead of the first way as currently. Finally, the method of class intervals of mean based on 7-Likert scale is also mentioned in the detailed paper.

**Key word.** Quantitative research; Mean; Likert scales; Statistical divisions; Continuous variables; Discrete variables; Criteria; Dialectic; Logic.

### 1. Mục tiêu và phương pháp nghiên cứu

Hiện nay đang tồn tại ba cách phân tổ để xếp loại giá trị trung bình (mean) trong các nghiên cứu định lượng có sử dụng thang đo Likert 5 mức độ. Vì có ba cách nên cùng một nguồn dữ liệu nhưng sử dụng cách khác nhau, nhiều khi lại cho ra kết quả phân tổ và xếp loại giá trị mean không giống nhau. Chính vì vậy, mục tiêu của bài báo là (1)Trình bày nội dung của ba cách trên, sau đó đối chiếu với cơ sở lý thuyết và thực tế về phân tổ thống kê để khẳng định cách nào đúng, nên áp dụng khi cần phải phân tổ, phân loại giá trị mean trong các nghiên cứu khoa học và (2)Trình bày cách phân tổ giá trị mean theo thang đo Likert 7 mức độ. Để đạt được hai mục tiêu trên, các phương pháp tổng hợp tài liệu, so sánh, đối chiếu, logic và biện chứng khoa học đã được sử dụng.

### 2. Cơ sở lý thuyết và thực tế

#### 2.1. Số trung bình (mean)

Số trung bình trong thống kê là chỉ tiêu biểu hiện mức độ chung nhất, điển hình nhất của một tổng thể bao gồm nhiều đơn vị cùng chất [1]. Chẳng hạn, chiều cao trung bình của các sinh viên trong cùng một lớp học là 1.67mét. Số trung bình có đặc điểm là biểu hiện được mức độ chung nhất, điển hình nhất và là chỉ tiêu đại diện dùng để so sánh các tổng thể không cùng quy mô. Với thí dụ trên, chúng ta chỉ có thể so sánh chiều cao của hai lớp thông qua chiều cao trung bình của từng lớp mà không thể so sánh tổng chiều cao hay chiều cao của hai em bất kỳ trong hai lớp. Bởi vì, tổng chiều cao của từng lớp phụ thuộc vào số (tức quy mô) sinh viên, còn hai em bất kỳ thì không mang tính đại diện chung.

Có ba loại số trung bình trong thống kê, đó là số bình quân cộng, bao gồm số bình quân cộng giản đơn và số bình quân cộng gia quyền. Chính số mean của thang đo Likert được tính theo số bình quân cộng giản đơn. Nó phản ánh ý kiến chung của những đơn vị được khảo sát khi đánh giá mức độ đạt được của một yếu tố hay của một biến quan sát có trong mô hình nghiên cứu. Ngoài số bình quân cộng còn có số bình quân điều hòa và số bình quân nhân. Việc áp dụng tính theo loại số bình quân nào là tùy thuộc vào nguồn dữ liệu và mục tiêu nghiên cứu.

## 2.2. Thang đo Likert

Có nhiều loại thang đo định lượng khác nhau, nhưng được sử dụng nhiều hơn cả là thang đo likert 5 (hoặc 7) mức độ. Theo đó, mức độ 1-hoàn toàn không đồng ý; 2-tương đối đồng ý; 3-bình thường; 4-đồng ý và 5-hoàn toàn đồng ý. Đây là loại thang đo nhằm mô tả thái độ, chính kiến của đối tượng được khảo sát về một vấn đề kinh tế, xã hội nào đó và được đặt tên theo tên của người đã đề xướng ra nó, là nhà khoa học xã hội người Mỹ-Rensis Likert (viết tắt Likert).

Thông thường một cuộc khảo sát được tiến hành trên nhiều đơn vị. Đơn vị có thể là người, doanh nghiệp, cơ quan hay địa phương,... Cùng một câu hỏi, nhưng tùy vào nhận thức và thái độ mà những người được hỏi có thể chọn mức độ (câu trả lời) không giống nhau. Bởi vậy, để mô tả ý kiến chung, cần phải tính mean, tiếp theo là phải phân tổ để xếp loại nó. Việc đếm số đơn vị (tức xác định tần số) và tính tỷ trọng (tức xác định tần suất) của từng tổ chiếm trong tổng số đơn vị cho thấy kết cấu của tổng thể hoặc của mẫu khảo sát.

## 2.3. Phân tổ thống kê

Phân tổ thống kê là căn cứ vào một (hay một số) tiêu thức nào đó để phân chia các đơn vị của tổng thể thành các tổ (hoặc tiểu tổ) có tính chất khác nhau [1]. Chẳng hạn phân tổ các sinh viên theo chiều cao thì tiêu thức phân tổ là chiều cao; phân tổ giá trị mean của sự hài lòng thì tiêu thức phân tổ chính là sự hài lòng, còn tổng thể nghiên cứu là tập hợp các đơn vị thuộc đối tượng khảo sát.

Phép duy vật biện chứng có 3 quy luật, trong đó có quy luật “Lượng biến, Chất biến”. Theo Wikipedia tiếng Việt thì quy luật này có nội dung là “*Biến đổi về lượng đến một mức nhất định sẽ dẫn đến sự biến đổi về chất, rồi trên nền tảng của chất mới lại bắt đầu biến đổi về lượng mới, lượng mới được tích lũy dần dần lại dẫn đến sự thay đổi về chất của sự vật*”. Triết học gọi trạng thái tích lũy về lượng dẫn đến sự thay đổi về chất là “bước nhảy” và cứ mỗi “bước nhảy” một tổ mới sẽ được hình thành. Điều này có nghĩa, các đơn vị trong cùng một tổ cơ bản là giống nhau về chất nhưng ở hai tổ thì phải có sự khác nhau. Đây chính là cơ sở, là lý thuyết nền cho phân loại nói chung và phân tổ trong khoa học thống kê nói riêng.

## 2.4. Cách phân tổ thống kê vận dụng trong phân tổ giá trị mean

Có hai cách phân tổ thống kê nếu căn cứ vào loại tiêu thức phân tổ. Đó là phân tổ theo tiêu thức chất lượng và phân tổ theo tiêu thức số lượng. Tiêu thức chất lượng (hay thuộc tính) là loại tiêu thức không được biểu hiện bằng con số cụ thể, như: giới tính, dân tộc, quê quán, loại hình doanh nghiệp,... [2][3]. Tiêu thức số lượng (hay khối lượng) là tiêu thức được biểu hiện bằng con số cụ thể, như: chiều cao, cân nặng, số lao động của doanh nghiệp,...[2][3]. Theo đó, việc phân tổ theo tiêu thức số lượng được chia thành hai trường hợp, tùy thuộc vào loại biến và đặc điểm biến động của nó. Nếu là biến liên tục (biến có thể nhận giá trị nguyên hoặc thập phân; như: tiền lương, lợi nhuận, chi phí, số diện tích,...) và biến thiên đều thì phải phân tổ đều. Khoảng cách tổ đều ( $h$ ) trong trường hợp này được tính theo công thức:

$$h = \frac{X_{\max} - X_{\min}}{k} \quad (2.1)$$

Với:  $X_{\max}$ ,  $X_{\min}$  là lượng biến (mức độ) lớn nhất và nhỏ nhất;  $k$  là số tổ định chia [2][3]

Mỗi tổ sẽ được hình thành từ hai giới hạn là giới hạn dưới và giới hạn trên, chênh lệch của hai giới hạn này được gọi là khoảng cách tổ. Ngoài ra trong phân tổ này, giới hạn trên của tổ đứng trước phải lấy đúng bằng giới hạn dưới của tổ đứng sau [2][3] nhằm thể hiện tính liên tục của biến số. Trường hợp có một lượng biến trùng với giới hạn trên của tổ trước và giới hạn dưới của tổ liền kề thì phải xếp đơn vị tương ứng vào tổ đứng sau [2][3]. Chẳng hạn, khi phân tổ các doanh nghiệp theo tình hình hoàn thành kế hoạch, có hai tổ liền kề là 80-100% và 100-120%. Nếu doanh nghiệp hoàn thành đúng 100% kế hoạch thì xếp doanh nghiệp này vào tổ thứ hai là tổ “hoàn thành và hoàn thành vượt mức kế hoạch”.

Ngược lại, với biến rời rạc (biến chỉ nhận các giá trị nguyên mà không lấy thập phân; như: số doanh nghiệp, số lao động, số bàn, ghế,...) thì các giới hạn này không được lấy trùng nhau mà phải cách nhau một đơn vị, nhằm thể hiện tính rời rạc của biến số. Chẳng hạn, phân tổ các doanh nghiệp trong 1 địa phương theo số lượng lao động thành 4 tổ: nhỏ hơn hoặc bằng 100 người; 101-200 người; 201-300 người và từ 301 người trở lên [4].

Trường hợp là biến liên tục nhưng biến thiên không đều thì phải phân tổ không đều và các giới hạn lấy như trong trường hợp biến thiên đều [2][3]. Liên hệ với thang đo Likert có 5 (hoặc 7) mức độ thì chắc chắn rằng, biến thiên từ mức độ 1 tuần tự đến mức độ 5 (hoặc 7) là biến thiên đều. Bởi vì, khi chọn 1, 2, 3, 4, 5, 6 hoặc 7 cho câu trả lời, người được khảo sát phải cân nhắc, so sánh lần lượt các mức độ với nhau. Các mức độ này hơn kém đều nhau và đúng bằng một bậc. Điều thú vị là giá trị mean của các câu trả lời thuộc biến liên tục vì nó có thể nhận giá trị nguyên hoặc thập phân. Chẳng hạn mean của mẫu này là 3.0, của mẫu kia có thể là 3.15,... Trong khi đó, 5 (hoặc 7) mức độ của thang đo Likert lại là biến rời rạc vì chúng chỉ nhận các giá trị nguyên từ 1 đến 5 (hoặc 7).

### 3. Ba cách phân tổ giá trị mean đang áp dụng

#### 3.1. Cách thứ nhất [5]

Theo cách này, giới hạn trên của tổ đứng trước và giới hạn dưới của tổ đứng sau trong hai tổ liền kề được tính theo công thức:

$$Mean = \frac{i + (i+1)}{2} \quad (3.1)$$

Với:  $i$  là mức độ thứ  $i$  ( $i=1...5$ )

Từ công thức (3.1), giới hạn trên của tổ 1 đồng thời là giới hạn dưới của tổ 2 sẽ là  $[1+(1+1)]/2=1.5$ ; tương tự, giới hạn trên của tổ 2 đồng thời là giới hạn dưới của tổ 3 sẽ là  $[2+(2+1)]/2=2.5$ ; giới hạn trên của tổ 3 đồng thời là giới hạn dưới của tổ 4 sẽ là  $[3+(3+1)]/2=3.5$ ; và cuối cùng, giới hạn trên của tổ 4 đồng thời là giới hạn dưới của tổ 5 sẽ là  $[4+(4+1)]/2=4.5$ . Vì thang đo có 5 mức độ nên giới hạn dưới của tổ thứ nhất phải là 1 và giới hạn trên của tổ thứ năm phải là 5. Từ đó giá trị mean của cuộc khảo sát được phân thành 5 tổ như sau:

Tổ 1: Từ 1 -> 1.5 - Rất thấp/ Tổ 2: Từ 1.5 -> 2.5 - Thấp/ Tổ 3: Từ 2.5 -> 3.5 - Trung bình/ Tổ 4: Từ 3.5 -> 4.5 - Cao/ Tổ 5: Từ 4.5 -> 5- Rất cao.

Cách phân tổ này có điểm phù hợp là giới hạn trên và giới hạn dưới của hai tổ liền kề trùng nhau. Điểm chưa phù hợp là phân tổ không đều nên khoảng cách tổ không bằng nhau. Cụ thể tổ 1 và tổ 5 có khoảng cách tổ là 0.5 khác với tổ 2, 3 và 4 có khoảng cách tổ là 1. Trong khi đó, như đã phân tích, thì giá trị mean thuộc biến liên tục và có biến thiên đều. Thêm nữa, cơ sở khoa học cho việc hình thành công thức (3.1) là chưa rõ ràng.

#### 3.2. Cách thứ hai [6]

Theo cách này, khoảng cách tổ của từng tổ được tính theo công thức (2.1). Cụ thể với thang đo Likert 5 mức độ thì khoảng cách tổ =  $(5 - 1) / 5 = 0.8$

Từ đó giá trị mean của cuộc khảo sát được xếp thành 5 tổ như sau:

Tổ 1: Từ 1 -> 1.8 - Rất thấp/ Tổ 2: Từ 1.81 -> 2.6 - Thấp/ Tổ 3: Từ 2.61 -> 3.4 - Trung bình/ Tổ 4: Từ 3.41 -> 4.2 - Cao/ Tổ 5: Từ 4.21 -> 5- Rất cao.

Cách phân tổ này có điểm phù hợp là xác định khoảng cách tổ theo phân tổ đều, phù hợp với sự biến thiên của giá trị mean. Nhưng có điểm chưa phù hợp là giới hạn trên và giới hạn dưới của hai tổ liền kề không lấy trùng nhau, vì vậy, khoảng cách tổ đã không bằng nhau. Cụ thể: tổ thứ nhất có khoảng cách tổ là 0.8, bốn tổ còn lại có khoảng cách tổ là 0.79.

#### 3.3. Cách thứ ba [6]

Gần giống với cách thứ hai, nhưng các giới hạn trên và giới hạn dưới của hai tổ liền kề lấy trùng nhau. Cụ thể:

Tổ 1: Từ 1 -> 1.8 - Rất thấp/ Tổ 2: Từ 1.8 -> 2.6 - Thấp/ Tổ 3: Từ 2.6 -> 3.4 - Trung bình/ Tổ 4: Từ 3.4 -> 4.2 - Cao/ Tổ 5: Từ 4.2 -> 5- Rất cao.

Cách phân tổ này đã thể hiện được tính liên tục của giá trị mean, vì các giới hạn trên và dưới của hai tổ liền kề được lấy trùng nhau và giữ nguyên được khoảng cách tổ đều (là 0.8).

### 4. Kết quả nghiên cứu và thảo luận

Cách thứ hai và thứ ba được nhiều người ghi nguồn trích dẫn từ sách “*Phân tích dữ liệu nghiên cứu với SPSS*” [6]. Tuy vậy, trong cuốn sách này và cả nhiều sách khác, như: “*Giáo trình nguyên lý thống kê*” [2], “*Thống kê ứng dụng*” [3],... các tác giả chỉ hướng dẫn cách phân tổ thống kê nói chung mà không đưa ra thí dụ cụ thể về cách phân tổ, phân loại giá trị mean. Bởi vậy, do cách hiểu không thống nhất nên có người tiến hành phân tổ giá trị mean theo cách thứ hai, lại có ít người theo cách thứ ba (hai giới hạn trùng nhau).

Trong khi đó với cách thứ nhất, các tác giả đã trình bày cụ thể phương pháp phân tổ giá trị mean theo thang đo Likert 5 mức độ. Nội dung của ba cách trên được thể hiện trong bảng:

Bảng Tổng hợp ba cách phân tổ giá trị mean

Tổ	Phân loại giá trị mean	Cách thứ nhất	Cách thứ hai	Cách thứ ba
1.	Rất thấp	1.0 – 1.5	1.0 – 1.8	1.0 – 1.8
2.	Thấp	1.5 – 2.5	1.81 – 2.6	1.8 – 2.6
3.	Trung bình	2.5 – 3.5	2.61 – 3.4	2.6 – 3.4
4.	Cao	3.5 – 4.5	3.41 – 4.2	3.4 – 4.2
5.	Rất cao	4.5 – 5.0	4.21 – 5.0	4.2 – 5.0

*Nguồn: Tổng hợp*

Áp dụng ba cách phân tổ giá trị mean trên đây, đôi khi có kết quả không nhất quán. Chẳng hạn giá trị mean bằng 1.6 với cách thứ nhất phải xếp vào tổ 2 là tổ thấp nhưng với cách thứ hai lại phải xếp vào tổ 1 là tổ rất thấp. Tương tự, giá trị mean đúng bằng 2.6 với cách thứ hai phải xếp vào tổ 2 là tổ thấp còn với cách thứ ba lại phải xếp vào tổ thứ ba là tổ trung bình.

Đối chiếu với mục 2- Cơ sở lý thuyết và thực tế - thì giá trị mean là biến liên tục, biến thiên đều, nên phân tổ theo cách thứ ba là cách đúng, duy nhất. Cách phân tổ này phù hợp với cơ sở khoa học của phân tổ thống kê và giải quyết được tình trạng không thống nhất của hai cách phân tổ trước. Hy vọng rằng, cách thứ ba sẽ được sử dụng rộng rãi trong phân tổ, phân loại giá trị mean khi muốn thể hiện kết quả nghiên cứu thay vì cách thứ nhất như hiện nay trong khối ngành Kinh tế của trường Đại học Công nghiệp thành phố Hồ Chí Minh.

Với thang đo Likert có 7 mức độ, cách phân tổ tương tự như đối với 5 mức độ. Tuy nhiên, khoảng cách tổ đều (h) sẽ là  $(7-1)/5=1.2$  thay vì 0.8. Cụ thể: Tổ 1: Từ 1 -> 2.2 - Rất thấp/ Tổ 2: Từ 2.2 -> 3.4 - Thấp/ Tổ 3: Từ 3.4 -> 4.6 - Trung bình/ Tổ 4: Từ 4.6 -> 5.8 - Cao/ Tổ 5: Từ 5.8 -> 7 - Rất cao. Trong trường hợp này, tác giả vẫn phân loại mean của thang đo 7 mức độ thành 5 tổ (chứ không phải 7 tổ) là để đảm bảo được tính khái quát và tính đặc trưng chung, điển hình vốn có của chỉ tiêu trung bình. Sẽ là không đúng nguyên tắc phân tổ của khoa học thống kê, khi tổ được chia quá nhỏ, quá vụn vặt và giữa các tổ đã có sự khác nhau về lượng mà lại không có sự khác nhau về chất một cách rõ ràng.

Cuối cùng xin lưu ý rằng, trong phần cơ sở lý thuyết và thực tế, tác giả chỉ mới đề cập đến cách phân tổ thống kê có liên quan tới việc phân tổ giá trị mean. Những cách phân tổ khác chưa được trình bày và phân tích vì không cần thiết cho bài báo./.

## TÀI LIỆU THAM KHẢO

- [1] Patrick W.S & et al (2002). *A Course in Business Statistics*, Prentice Hall, Inc.
- [2] Trần Phước & Cộng sự (2011). *Giáo trình Nguyên lý thống kê*. Trường ĐH Công nghiệp Thành phố Hồ Chí Minh.
- [3] Hà Văn Sơn (2004). *Giáo trình Lý thuyết thống kê ứng dụng trong Quản trị và Kinh tế*. Nhà xuất bản Thống kê.
- [4] Nguyễn Minh Tuấn & Cộng sự (2013). *Giáo trình Thống kê kinh doanh*. Nhà xuất bản Đại học Quốc gia Thành phố Hồ Chí Minh.
- [5] Nguyễn Minh Tuấn & Cộng sự (2008). *Giáo trình Nghiên cứu Marketing*. Nhà xuất bản Đại học Quốc gia Thành phố Hồ Chí Minh.
- [6] Hoàng Trọng và Chu Nguyễn Mộng Ngọc (2008). *Phân tích dữ liệu nghiên cứu với SPSS, Tập 1*. Nhà xuất bản Hồng Đức.

*Ngày nhận bài: 21/02/2020*

*Ngày chấp nhận đăng: 07/04/2020*