TCVN **TIÊU CHUẨN QUỐC GIA**

TCVN 12740:2019

ISO/TR 13587:2012

Xuất bản lần 1

BA CÁCH TIẾP CẬN THỐNG KÊ ĐÁNH GIÁ VÀ GIẢI THÍCH ĐỘ KHÔNG ĐẢM BẢO ĐO

Three statistical approaches for the assessment and interpretation of measurement uncertainty

HÀ NỘI - 2019

Mục lục

Lời nói đầu

TCVN 12740:2019 hoàn toàn tương đương với ISO/TR 13587:2012.

TCVN 12740:2019 do Ban kỹ thuật tiêu chuẩn quốc gia TCVN/TC 69 *Ứng dụng các phương pháp thống kê* biên soạn, Tổng cục Tiêu chuẩn Đo lường Chất lượng đề nghị, Bộ Khoa học và Công nghệ công bố.

Lời giới thiệu

Việc chấp nhận TCVN 9595-3 [ISO/IEC Guide 98-3 (GUM)] đã dẫn đến sự thừa nhận ngày càng tăng về sự cần thiết phải bao gồm công bố về độ không đảm bảo trong kết quả đo. Việc công nhận phòng thí nghiệm dựa trên tiêu chuẩn như TCVN ISO/IEC 17025[2] đã đẩy nhanh quá trình này. Thừa nhận rằng công bố về độ không đảm bảo là cần thiết cho việc đưa ra quyết định hiệu quả, các nhà đo lường học trong các phòng thí nghiệm thuộc mọi loại hình, từ Viện Đo lường Quốc gia đến các phòng thí nghiệm hiệu chuẩn thương mại, đang nỗ lực rất nhiều trong việc xây dựng các đánh giá độ không đảm bảo thích hợp cho các loại phép đo khác nhau, sử dụng các phương pháp nêu trong GUM.

Những điểm mạnh của các qui trình phác thảo và phổ biến trong GUM là cách tiếp cận chuẩn hóa để đánh giá độ không đảm bảo, là sự thích ứng đối với các nguồn độ không đảm bảo được đánh giá thống kê (Loại A) hoặc phi thống kê (Loại B), là sự nhấn mạnh vào việc báo cáo tất cả các nguồn độ không đảm bảo cần xem xét. Cách tiếp cận chủ yếu về lan truyền độ không đảm bảo trong GUM, dựa trên phép xấp xỉ tuyến tính của hàm đo, thường dễ thực hiện và trong nhiều tình huống thực tế sẽ cho các kết quả tương tự với các kết quả thu được một cách chính thức hơn. Nói tóm lại, kể từ khi được chấp nhận áp dụng, GUM đã tạo ra một cuộc cách mạng trong đánh giá độ không đảm bảo.

Thực tế là sẽ luôn cần làm việc nhiều hơn để cải tiến việc đánh giá độ không đảm bảo trong các ứng dụng cụ thể, đồng thời mở rộng ra để bao trùm thêm các lĩnh vực khác. Trong số những công việc này, Ủy ban phối hợp về Hướng dẫn trong đo lường học (JCGM), chịu trách nhiệm về GUM kể từ năm 2000, đã hoàn thành Phần bổ sung 1 của GUM, tiêu đề là "Lan truyền các phân bố sử dụng phương pháp Monte Carlo" (gọi là GUMS1)[3]. JCGM còn xây dựng các phần bổ sung khác cho GUM về các chủ đề như mô hình hóa và các mô hình với số lượng đại lượng đầu ra bất kỳ.

Do cần áp dụng cho tập hợp các vấn đề đo lường rộng nhất có thể nên định nghĩa về độ không đảm bảo đo trong TCVN 6165:2009 (ISO/IEC Guide 99:2007)[4] là: "tham số không âm đặc trưng cho độ phân tán của các giá trị đại lượng được quy cho đại lượng đo, trên cơ sở thông tin đã sử dụng" không thể đưa nhiều hơn mức khái niệm tương đối một cách hợp lý. Kết quả là, việc xác định và hiểu về vai trò thích hợp của các đại lượng thống kê khác nhau trong đánh giá độ không đảm bảo, thậm chí cho các ứng dụng đo lường tương đối được thông hiểu, là một chủ đề được quan tâm đặc biệt với cả nhà thống kê và nhà đo lường.

Các nghiên cứu trước đây đã tiếp cận các chủ đề từ quan điểm đo lường học, một số tác giả tập trung vào mô tả đặc trưng thuộc tính thống kê của các quy trình nêu trong GUM. Tài liệu tham khảo [5] cho thấy các quy trình này không hoàn toàn phù hợp với giải thích theo Bayes hay giải thích theo tần suất. Tài liệu tham khảo [6] đề xuất một số sửa đổi nhỏ cho quy trình GUM mà trong một số tình huống sẽ đưa ra các kết quả thống nhất hơn với giải thích Bayes. Tài liệu tham khảo [7] thảo luận về quan hệ giữa các quy trình đánh giá độ không đảm bảo đề xuất trong GUMS1 và các kết quả của phân tích Bayes cho một lớp mô hình cụ thể. Tài liệu tham khảo [8] cũng thảo luận về các giải thích xác suất khác nhau có thể có về khoảng phủ và khuyến nghị xấp xỉ các phân bố hậu nghiệm cho lớp phân tích Bayes này bằng các phân bố xác suất từ họ phân bố Pearson.

Tài liệu tham khảo [9] so sánh cách tiếp cận tần suất ("quy ước") và Bayes để đánh giá độ không đảm bảo. Tuy nhiên, nghiên cứu chỉ giới hạn ở các hệ thống đo mà tất cả các nguồn độ không đảm bảo có thể đánh giá bằng cách sử dụng phương pháp Loại A. Ngược lại, các hệ thống đo có nguồn độ không đảm bảo được đánh giá bằng cả phương pháp Loại A và Loại B được đề cập trong tiêu chuẩn này và được minh họa bằng nhiều ví dụ, bao gồm một trong các ví dụ từ Phụ lục H của GUM.

Các nhà thống kê học đã từng rất chú trọng vào việc sử dụng các phương pháp đánh giá độ không đảm bảo có chứng minh hoặc giải thích xác suất. Thông qua công việc của họ, thường nằm ngoài lĩnh vực đo lường, nhiều cách tiếp cận khác nhau đối với suy luận thống kê liên quan đến đánh giá độ không đảm bảo đã được phát triển. Tiêu chuẩn này trình bày một số cách tiếp cận đánh giá độ không đảm bảo từ quan điểm thống kê và liên hệ chúng với các phương pháp hiện đang được sử dụng trong đo lường hoặc được phát triển trong cộng đồng đo lường. Các cách tiếp cận thống kê cụ thể trong đó các phương pháp khác nhau để đánh giá độ không đảm bảo sẽ được mô tả là cách tiếp cận tần suất, Bayes và cách tiếp cận dựa vào sự tin tưởng, được thảo luận thêm sau khi vạch ra các quy ước ký hiệu cần thiết để phân biệt các loại đại lượng khác nhau.

Ba cách tiếp cận thống kê để đánh giá và giải thích độ không đảm bảo đo

Three statistical approaches for the assessment and interpretation of measurement uncertainty

1 Phạm vi áp dụng

Tiêu chuẩn này đề cập đến ba cách tiếp cận thống kê để đánh giá và giải thích độ không đảm bảo đo: cách tiếp cận tần suất bao gồm khoảng không đảm bảo bootstrap, cách tiếp cận Bayes và cách tiếp cận lập luận dựa vào sự tin tưởng. Đặc điểm chung của ba cách tiếp cận này là giải thích xác suất phân định rõ ràng hoặc lý giải về khoảng độ không đảm bảo thu được. Đối với mỗi cách tiếp cận, phương pháp cơ bản được mô tả, các giả định cơ bản và việc giải thích xác suất về độ không đảm bảo thu được được thảo luận. Mỗi cách tiếp cận được minh họa bằng hai ví dụ, bao gồm ví dụ lấy từ TCVN 9595-3 (ISO/IEC Guide 98-3) [Độ không đảm bảo đo – Phần 3: Hướng dẫn trình bày độ không đảm bảo đo (GUM:1995)]. Ngoài ra, tiêu chuẩn này cũng thảo luận về mối quan hệ giữa các phương pháp đề xuất trong GUM phần bổ sung 1 và ba cách tiếp cận thống kê này.

2 Tài liệu viện dẫn

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

TCVN 8244-1:2010 (ISO 3534-1:2006), *Thống kê học – Từ vựng và ký hiệu – Phần 1: Thuật ngữ chung về thống kê và thuật ngữ dùng trong xác suất*

TCVN 8244-2:2010 (ISO 3534-2:2006), *Thống kê học – Từ vựng và ký hiệu – Phần 2: Thống kê ứng dụng*

TCVN 9595-3:2013 (ISO/IEC Guide 98-3:2008), *Độ không đảm bảo đo – Phần 3: Hướng dẫn trình bày độ không đảm bảo đo (GUM:1995)*

ISO/IEC Guide 98-3:2008/Suppl 1:2008, *Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) – Supplement 1: Propagation of distributions using a Monte Carlo method* [Độ không đảm bảo đo – Phần 3: Hướng dẫn trình bày độ không đảm bảo đo (GUM:1995) – Phần bổ sung 1: Lan truyền phân bố sử dụng phương pháp Monte Carlo]

3 Thuật ngữ và định nghĩa

Tiêu chuẩn này áp dụng các thuật ngữ và định nghĩa trong TCVN 8244-1 (ISO 3534-1), TCVN 8244-2 (ISO 3534-2) và các thuật ngữ, định nghĩa dưới đây.

3.1

Hàm phân bố thực nghiệm (empirical distribution function)

Hàm phân bố tích lũy thực nghiệm (empirical cumulative distribution function)

Hàm phân bố ấn định xác suất $1/n$ cho mỗi trong số n cá thể trong mẫu ngẫu nhiên, tức hàm phân bố thực nghiệm là một hàm bậc thang được xác định bởi

$$F_n(x) = \frac{|\{x_i \le x\}|}{n},$$

trong đó $\{x_1, ..., x_n\}$ là mẫu và $|A|$ là số phần tử trong tập A.

3.2

Phân tích độ nhạy Bayes (Bayesian sensitivity analysis)

Nghiên cứu ảnh hưởng của sự lựa chọn các phân bố tiên nghiệm cho các tham số của mô hình thống kê lên phân bố hậu nghiệm của đại lượng đo.

3.3

Thống kê đủ (sufficient statistic)

Hàm của mẫu ngẫu nhiên $X_1, ..., X_n$ từ hàm mật độ xác suất với tham số θ mà phân bố điều kiện của $X_1, ..., X_n$ cho hàm này không phụ thuộc vào θ.

CHÚ THÍCH: Thống kê đủ chứa toàn bộ thông tin về θ như $X_1, ..., X_n$.

3.4

Mô hình quan trắc (observation model)

Quan hệ toán học giữa tập hợp các phép đo (số chỉ), đại lượng đo và sai số đo ngẫu nhiên kèm theo.

3.5

Phương trình cấu trúc (structural equation)

Mô hình thống kê liên hệ biến ngẫu nhiên quan trắc được với tham số chưa biết và biến ngẫu nhiên không quan trắc được có phân bố đã biết và không phụ thuộc vào tham số chưa biết.

3.6

Phân bố Khi-bình phương không trung tâm (non-central chi-squared distribution)

Phân bố xác suất tổng quát hóa phân bố Khi-bình phương điển hình (hoặc trung tâm).

CHÚ THÍCH 1: Đối với k biến ngẫu nhiên độc lập, phân bố chuẩn X_i có trung bình μ_i và phương sai σ_i^2, biến ngẫu nhiên $X = \sum_{i=1}^{k} (X_i / \sigma_i)^2$ là phân bố Khi-bình phương không trung tâm. Phân bố Khi-bình phương không trung tâm có hai tham số: k, bậc tự do (tức là số lượng X_i) và λ, liên quan đến các trung bình của các biến ngẫu nhiên X_i bởi $\lambda = \sum_{i=1}^{k} (\mu_i / \sigma_i)^2$ và được gọi là tham số không trung tâm.

CHÚ THÍCH 2: Hàm mật độ xác suất tương ứng được biểu thị như tổng hợp các hàm mật độ xác suất χ^2 trung tâm cho bởi

$$g_X(\zeta) = \sum_{i=0}^{\infty} \frac{e^{-\lambda/2}(\lambda/2)^i}{i!} g_{Y_{k+2i}}(\zeta)$$

$$= \frac{e^{\frac{-(\zeta+\lambda)}{2}}}{2^{\frac{k}{2}}} \sum_{i=0}^{\infty} \frac{\zeta^{\frac{k}{2}+i-1}\lambda^i}{\Gamma\left(\frac{k}{2}+i\right)2^{2i} i!},$$

trong đó Y_q có phân bố như Khi-bình phương với q bậc tự do.

4 Ký hiệu (và từ viết tắt)

Trong 4.1.1 của GUM quy định các chữ cái La tinh được sử dụng để thể hiện các đại lượng vật lý cần xác định bằng phép đo (tức là các đại lượng đo theo thuật ngữ GUM) cũng như các biến ngẫu nhiên có thể lấy các giá trị quan trắc khác nhau của một đại lượng vật lý. Việc sử dụng cùng một ký hiệu, có các ý nghĩa khác nhau chỉ được chỉ ra theo bối cảnh, có thể khó giải thích và đôi khi dẫn đến sự không rõ ràng hay hiểu nhầm không cần thiết. Để giảm thiểu nguồn gây nhầm lẫn tiềm ẩn này, cách ký hiệu truyền thống hơn thường được sử dụng trong tài liệu thống kê được áp dụng trong tiêu chuẩn này. Trong cách ký hiệu này, các chữ cái Hy Lạp được sử dụng để thể hiện các tham số trong mô hình thống kê (ví dụ như đại lượng đo), có thể là biến ngẫu nhiên hoặc hằng số tùy thuộc vào cách tiếp cận thống kê được sử dụng và tính chất của mô hình đó. Chữ cái La tinh hoa được sử dụng để thể hiện biến ngẫu nhiên có thể lấy các giá trị khác nhau của đại lượng quan trắc được (ví dụ như giá trị đo được tiềm năng), còn chữ cái La tinh thường thể hiện giá trị quan trắc cụ thể của một đại lượng (ví dụ như giá trị đo được cụ thể). Vì có thể cần ký hiệu bổ sung để chỉ thị các khái niệm vật lý, toán học hoặc thống kê khác nên sẽ luôn có khả năng có sự không rõ ràng nhất định [1]. Trong những trường hợp này, bối cảnh sẽ làm rõ cách giải thích phù hợp.

5 Vấn đề cần giải quyết

5.1 Vấn đề quan tâm trong tiêu chuẩn này là mô hình đo trong đó μ_1 ... μ_p là các đại lượng đầu vào và θ là đại lượng đầu ra:

$$\theta = f(\mu_1, \ldots, \mu_p) \quad (1)$$

---

[1] Ví dụ, không phải tất cả các đại lượng trình bày bằng chữ cái Hy Lạp trong một mô hình thống kê đều là các tham số của mô hình đó. Một ví dụ phổ biến về loại đại lượng này là tập các đại lượng không quan trắc được đại diện cho các sai số đo ngẫu nhiên có trong hầu hết các mô hình thống kê (tức là, ε_i trong mô hình $Y_i = \mu + \varepsilon_i$).

trong đó f được biết là hàm đo. Hàm f được xác định theo toán học hoặc như một quy trình tính toán. Trong GUM (4.1, CHÚ THÍCH 1), một quan hệ hàm số giống hệt được đưa ra là

$$Y = f(X_1, \ldots, X_p) \quad (2)$$

có thể không dễ dàng phân biệt với nó hàm đo được đánh giá tại giá trị của các biến ngẫu nhiên tương ứng cho từng đầu vào quan trắc được.

Bằng cách sử dụng quy trình khuyến nghị trong GUM, p đại lượng chưa biết $\mu_1 \ldots \mu_p$, được ước lượng bằng các giá trị $x_1, \ldots, x_p$ thu được từ phép đo hoặc từ các nguồn khác. Độ không đảm bảo chuẩn kèm theo của chúng cũng thu được từ dữ liệu liên quan bằng các phương pháp thống kê hoặc từ hàm mật độ xác suất dựa trên kiến thức chuyên gia mô tả đặc trưng các biến. GUM (xem thêm 4.5 trong Tài liệu tham khảo [11]) khuyến nghị rằng cùng một mô hình đo liên hệ đại lượng đo θ với các đại lượng đầu vào $\mu_1, \ldots, \mu_p$ được dùng để tính y từ $x_1, \ldots, x_p$. Do đó, giá trị đo được (hoặc, theo thuật ngữ thống kê, là ước lượng) y của θ thu được là

$$y = f(x_1, \ldots, x_p) \quad (3)$$

tức là, Y được đánh giá bởi $y = f(x_1, \ldots, x_p)$, được lấy làm giá trị đo được của θ. Các ước lượng $y, x_1, \ldots, x_p$ tương ứng là các thể hiện của $Y, X_1, \ldots, X_p$.

5.2 Trong tiêu chuẩn này, ba cách tiếp cận thống kê được sử dụng để cung cấp (a) ước lượng tốt nhất y của θ, (b) độ không đảm bảo chuẩn kèm theo $u(y)$, và (c) khoảng tin cậy hoặc khoảng phủ đối với θ cho một xác suất phủ quy định (thường được lấy bằng 95 %).

5.3 Khi thảo luận về độ không đảm bảo chuẩn, sự phân biệt được đưa ra giữa độ không đảm bảo chuẩn đánh giá được gắn với ước lượng của các đại lượng khác nhau và giá trị lý thuyết tương ứng của chúng. Theo đó, các ký hiệu như σ_μ hoặc σ_x sẽ ký hiệu cho độ không đảm bảo chuẩn lý thuyết và các ký hiệu như S_x và s_x sẽ ký hiệu cho độ không đảm bảo chuẩn đánh giá, tương ứng với trước và sau khi được quan trắc.

6 Cách tiếp cận thống kê

6.1 Cách tiếp cận tần suất

6.1.1 Cách tiếp cận thống kê đầu tiên được xem xét, trong đó độ không đảm bảo có thể được đánh giá về mặt xác suất, là tiếp cận tần suất. Đôi khi, cách tiếp cận tần suất còn được gọi là "truyền thống" hay "quy ước". Tuy nhiên, do tính chất của độ không đảm bảo trong đo lường, các phương pháp quen thuộc này thường phải được điều chỉnh để có các khoảng không đảm bảo tần suất trong điều kiện thực tế.

6.1.2 Trong cách tiếp cận tần suất, các đại lượng đầu vào $\mu_1, \ldots, \mu_p$ trong mô hình đo (1) và đại lượng đầu ra θ được coi là các hằng số chưa biết. Khi đó, dữ liệu liên quan đến mỗi tham số đầu vào, μ_i, sẽ thu được và được dùng để ước lượng giá trị của θ dựa trên mô hình đo hoặc các mô hình thống kê tương ứng. Cuối cùng, khoảng tin cậy cho θ, đối với một mức tin cậy quy định, sẽ thu được bằng cách sử dụng một trong nhiều nguyên tắc hoặc phương pháp toán học, ví dụ, bình phương tối thiểu, hợp lý cực đại hoặc bootstrap.

6.1.3 Vì θ được coi như hằng số nên công bố xác suất kèm theo khoảng tin cậy cho θ không phải là công bố xác suất trực tiếp về giá trị của nó. Thay vào đó là công bố xác suất về tần suất mà quy trình sử dụng để thu được khoảng độ không đảm bảo cho đại lượng đo có thể bao gồm giá trị của θ khi sử dụng lặp lại. "Sử dụng lặp lại" có nghĩa là đánh giá độ không đảm bảo được lặp lại nhiều lần bằng cách sử dụng dữ liệu khác nhau lấy từ cùng các phân bố. Khoảng độ không đảm bảo tần suất truyền thống cung cấp một công bố xác suất về các thuộc tính dài hạn của quy trình sử dụng để thiết lập khoảng đó dưới một tập hợp các điều kiện cụ thể được giả định để áp dụng cho quá trình đo.

6.1.4 Mặt khác, trong hầu hết các bố trí đo lường thực tế, khoảng độ không đảm bảo là để tính đến độ không đảm bảo gắn với các ước lượng của đại lượng thu được bằng các giá trị đo được (dữ liệu quan trắc được) và cả độ không đảm bảo gắn với các ước lượng của đại lượng dựa trên kiến thức chuyên gia. Để thu được khoảng độ không đảm bảo tương tự với khoảng tin cậy, các đại lượng không dựa trên giá trị đo được được coi như biến ngẫu nhiên với phân bố xác suất cho giá trị của chúng trong khi những đại lượng có giá trị có thể ước lượng bằng dữ liệu thống kê được coi như các hằng số chưa biết.

6.1.5 Quy trình tần suất truyền thống để thiết lập khoảng tin cậy sau đó được sửa đổi để có được mức tin cậy quy định sau khi lấy trung bình các giá trị có thể có của các đại lượng đánh giá bằng ý kiến chuyên gia [5]. Những khoảng phủ sửa đổi này cung cấp công bố xác suất dài hạn về quy trình sử dụng để thu được khoảng với các phân bố xác suất cho đại lượng chưa được đo, cũng giống như khoảng tin cậy truyền thống đưa ra khi tất cả các tham số được coi là hằng số.

6.1.6 Bảng 1 tổng hợp các giải thích về cách tiếp cận tần suất, Bayes và cách tiếp cận dựa vào sự tin tưởng để đánh giá độ không đảm bảo.

Bảng 1 – Giải thích các cách tiếp cận đánh giá độ không đảm bảo

Cách tiếp cận	Mô tả đặc trưng các đại lượng trong mô hình đo $\theta = f(\mu_1, \dots, \mu_p)$	Khoảng độ không đảm bảo đối với đại lượng đầu ra θ	Chú thích
Tần suất	θ và μ_i đều là các hằng số chưa biết	Tần suất xuất hiện dài hạn mà khoảng đó chứa θ	Cách tiếp cận tần suất truyền thống mở rộng để tích hợp các độ không đảm bảo không được đánh giá thống kê
Bayes	θ và μ_i là các biến ngẫu nhiên. Phân bố xác suất của chúng đại diện cho sự tin tưởng về giá trị của đại lượng đầu vào và đầu ra	Khoảng phủ chứa θ dựa trên phân bố hậu nghiệm đối với θ	Tính không đơn nhất có thể có của khoảng là do lựa chọn các giá trị tiên nghiệm
Tin tưởng	μ_i coi là biến ngẫu nhiên có phân bố thu được từ giả định về dữ liệu quan trắc được sử dụng để ước lượng μ_i và kiến thức chuyên gia về μ_i	Khoảng phủ chứa θ dựa trên phân bố dựa vào sự tin tưởng đối với θ	Tính không đơn nhất là do lựa chọn phương trình cấu trúc

6.2 Cách tiếp cận Bayes

Cách tiếp cận thứ hai được gọi là cách tiếp cận Bayes. Tên gọi này xuất phát từ định lý cơ bản là cơ sở của cách tiếp cận, đã được Reverend Thomas Bayes chứng minh vào giữa những năm 1700 [12]. Trong cách tiếp cận này, kiến thức về các đại lượng trong mô hình đo (1) ở Điều 5 được mô hình hóa như một tập các biến ngẫu nhiên tuân theo phân bố xác suất đồng thời đối với $\mu_1,\ldots,\mu_p$ và θ. Khi đó, định lý Bayes cho phép cập nhật các phân bố xác suất này dựa trên dữ liệu quan trắc được (cũng được mô hình hóa bằng các phân bố xác suất) và mối quan hệ của các tham số xác định bằng hàm f hoặc mô hình thống kê tương đương. Sau đó, phân bố xác suất thu được mô tả kiến thức về θ theo dữ liệu quan trắc được. Khoảng độ không đảm bảo chứa θ với xác suất quy định bất kỳ có thể thu được sau đó từ phân bố này. Vì kiến thức về các giá trị tham số được mô tả bởi phân bố xác suất nên các phương pháp Bayes đưa ra công bố xác suất trực tiếp về giá trị của θ và các tham số khác, bằng cách sử dụng định nghĩa xác suất như là thước đo của sự tin tưởng.

6.3 Cách tiếp cận dựa vào sự tin tưởng

6.3.1 Cách tiếp cận dựa vào sự tin tưởng được phát triển bởi R.A. Fisher [13] vào những năm 1930. Trong cách tiếp cận này, phân bố xác suất, gọi là phân bố dựa vào sự tin tưởng, đối với θ với điều kiện dữ liệu thu được dựa trên mối quan hệ giữa θ và μ_i; mô tả bởi f và các giả định phân bố về dữ liệu dùng để ước lượng μ_i. Khi có được, phân bố dựa vào sự tin tưởng cho θ có thể dùng để thu được khoảng độ không đảm bảo chứa θ với xác suất quy định bất kỳ.

6.3.2 Lập luận lý giải cho quá trình dùng để thu được phân bố dựa vào sự tin tưởng được minh họa bằng một ví dụ đơn giản. Giả định các giá trị lấy bởi đại lượng Y có thể được mô tả bằng công thức $Y = \mu + Z$, trong đó μ là đại lượng đo và Z là đại lượng đặc trưng bởi biến ngẫu nhiên phân bố chuẩn. Nếu y là giá trị thể hiện của Y tương ứng với z là giá trị thể hiện của Z thì khi đó $\mu = y - z$. Mặc dù Z không thể quan trắc được, kiến thức về phân bố từ đó z được tạo ra cho phép xác định một tập hợp các giá trị hợp lý của μ. Phân bố xác suất đối với Z có thể dùng để suy ra phân bố xác suất của μ. Quá trình biến mối quan hệ $\mu = y - z$ sang quan hệ $\mu = y - Z$ là những gì cấu thành lập luận dựa vào sự tin tưởng. Phân bố dựa vào sự tin tưởng đối với μ là phân bố xác suất cho biến ngẫu nhiên $y - Z$ với y cố định.

6.4 Thảo luận

Khi mô tả các phương pháp khác nhau để đánh giá độ không đảm bảo theo từng cách tiếp cận thống kê này, các giả định cơ bản của chúng, sự kết hợp các độ không đảm bảo thu được bằng đánh giá Loại A hoặc Loại B, và việc giải thích về xác suất của các đánh giá độ không đảm bảo thu được sẽ được thảo luận. Mô tả về cách các phương pháp sử dụng trong GUM liên quan đến kết quả tần suất, Bayes hoặc dựa vào sự tin tưởng cũng sẽ được đưa ra.

7 Các ví dụ

7.1 Khái quát

Hai ví dụ được đưa ra để minh họa các cách tiếp cận. Ví dụ 1 liên quan đến đại lượng vật lý được hiệu chính đối với nhiễu nền. Bảng 2 nêu các ký hiệu được sử dụng và 7.2 đến 7.4 xác định các biến thể của bài toán đánh giá này. Ví dụ 2 là hiệu chuẩn độ dài của can mẫu lấy từ Phụ lục H.1 của GUM. Do phức tạp hơn nên nó được xem xét trong Điều 11, sau khi ba phương pháp đánh giá độ không đảm bảo được thảo luận và minh họa bằng Ví dụ 1.

Trong các điều sau, ba cách tiếp cận sẽ được áp dụng cho các ví dụ này.

CHÚ THÍCH: Đơn vị của các đại lượng liên quan không được cho khi chúng không quan trọng cho ví dụ.

Bảng 2 – Ký hiệu dùng cho Ví dụ 1

Đại lượng	Ký hiệu
Đại lượng vật lý cần quan tâm (đại lượng đo)	θ
Đại lượng được phát hiện bởi phương pháp đo khi đo nền (nghĩa là giá trị kỳ vọng của B) (nhiễu nền)	β
Đại lượng được phát hiện bởi phương pháp đo khi đo đại lượng vật lý cần quan tâm (nghĩa là giá trị kỳ vọng của Y)	$\gamma = \theta + \beta$
Độ lệch chuẩn của phương pháp đo khi đo đại lượng vật lý cần quan tâm (nghĩa là độ lệch chuẩn của Y)	σ_Y
Độ lệch chuẩn của phương pháp đo khi đo nền (nghĩa là độ lệch chuẩn của B)	σ_B

7.2 Ví dụ 1a

Năm giá trị đo được thu được độc lập của tín hiệu cùng với nền được quan trắc. Mỗi giá trị đo được được giả định là thể hiện của biến ngẫu nhiên Y có phân bố Gauss với trung bình $\gamma = \theta + \beta$ và độ lệch chuẩn σ_Y. Các giá trị đo được y của tín hiệu cùng với nền là

3,738, 3,442, 2,994, 3,637, 3,874.

Dữ liệu này có trung bình mẫu $\bar{y} = 3,537$ và độ lệch chuẩn mẫu $s_y = 0,342$.

Tương tự, có được năm giá trị nền đo được thu được độc lập. Các giá trị đo được này được giả định là các thể hiện của biến ngẫu nhiên B có phân bố Gauss với trung bình β và độ lệch chuẩn σ_B. Các giá trị quan trắc b của nền là

1,410, 1,085, 1,306, 1,137, 1,200.

Vì có các giá trị đo được cho mỗi đại lượng là một nguồn độ không đảm bảo, nên Ví dụ 1a có giải thích thống kê đơn giản cho mỗi cách tiếp cận.

7.3 Ví dụ 1b

Ví dụ 1b giống với Ví dụ 1a ngoại trừ việc đánh giá nền dựa vào kiến thức chuyên gia hoặc kinh nghiệm trước đó hơn là vào dữ liệu thực nghiệm mới. Trong trường hợp này, nền β được tin là tuân theo phân bố đều (hoặc chữ nhật) với các điểm đầu 1,126 và 1,329. Vì áp dụng đánh giá chuyên gia nên độ không đảm bảo gắn với giá trị của nền sẽ thu được bằng đánh giá Loại B. Do đó, Ví dụ 1b có thể được coi là gần với tình huống đo thực tế hơn Ví dụ 1a.

7.4 Ví dụ 1c

Ví dụ 1c giống Ví dụ 1b ngoại trừ tín hiệu θ gần với nền hơn. Dữ liệu quan trắc được đối với tín hiệu cùng với nền trong trường hợp này là

1,340 1,078, 1,114, 1,256, 1,192.

Với tín hiệu chỉ cao hơn mức nền, Ví dụ 1c minh họa cách các ràng buộc vật lý có thể kết hợp như thế nào trong đánh giá độ không đảm bảo đối với từng cách tiếp cận.

8 Cách tiếp cận tần suất để đánh giá độ không đảm bảo

8.1 Phương pháp cơ bản

8.1.1 Trong bối cảnh tần suất, tham số là các hằng số chưa biết. Tuân thủ quy ước ký hiệu các biến ngẫu nhiên bằng chữ cái in hoa và giá trị quan trắc của các biến ngẫu nhiên bằng chữ cái thường, khoảng tin cậy có thể thu được từ đại lượng then chốt đối với θ, tức là hàm $W(Y, \theta)$ của dữ liệu Y (có thể là nhiều biến) và tham số θ, có phân bố xác suất không phụ thuộc tham số (với điều kiện có thể xác định được phân bố). Khi đó, khoảng tin cậy $100(1-\alpha)$ % đối với θ có thể được xác định bằng cách tính phân vị dưới và phân vị trên l_α và u_α để thỏa mãn $P_\theta(l_\alpha \le W(Y, \theta) \le u_\alpha) = 1-\alpha$.

8.1.2 Ví dụ, lấy $Y = (Y_1, \ldots, Y_n)$ là các biến ngẫu nhiên, phân bố là $N(\mu, \sigma^2)$, với biến ngẫu nhiên $\bar{Y} = \sum_{i=1}^{n} Y_i / n$. Nếu tham số quan tâm là μ thì đối với σ chưa biết, $Z = \dfrac{\bar{Y}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$

là đại lượng then chốt. Khoảng tin cậy tần suất cho μ là

$$\bar{Y} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \tag{4}$$

trong đó z_β là phân vị 100 β của phân bố chuẩn chuẩn hóa.

Nếu chưa biết σ thì có thể ước lượng bằng độ lệch chuẩn mẫu

$$S = \sqrt{\frac{\sum_{j=1}^{n}\left(Y_j - \bar{Y}\right)^2}{n-1}}$$

Khi đó, đại lượng then chốt (đúng) đối với μ sẽ thu được bằng cách thay σ trong khoảng (4) bằng S:

$$\frac{\bar{Y}-\mu}{S/\sqrt{n}} \sim t\,(n-1). \qquad (5)$$

Do đó, khoảng tin cậy 100(1 – α) % đối với μ dựa trên phân bố t-Student là

$$\bar{Y} \pm \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2},$$

trong đó $t_{n-1,\beta}$ là phân vị 100β của phân bố t với n – 1 bậc tự do.

8.1.3 Thay vì các đại lượng then chốt đúng, chỉ tồn tại trong các tình huống đơn giản, các đại lượng then chốt xấp xỉ thường được sử dụng trong các ứng dụng. Đối với các mẫu lớn, có thể vận dụng định lý giới hạn trung tâm để thu được các khoảng tin cậy xấp xỉ dựa trên phân bố chuẩn.

8.1.4 Các phương pháp khác để có được khoảng tin cậy (lấy nghịch đảo thống kê kiểm nghiệm, lập hàm phân bố tích lũy liên tục làm then chốt, sắp xếp các giá trị mẫu rời rạc theo thứ tự xác suất của chúng, v.v). được thảo luận trong Tài liệu tham khảo [14]. Một số trong các phương pháp này được đề cập trong Ví dụ 1. Phương pháp chuyên dùng máy tính, được gọi là bootstrap, cũng được sử dụng để thiết lập khoảng tin cậy cho các đại lượng then chốt có phân bố chưa biết. Quy trình bootstrap được trình bày ở 8.2.

8.1.5 Mặc dù không có được lý giải tần suất rõ ràng từ các xem xét khoa học cơ bản, nhưng có thể sử dụng các quy trình khuyến nghị trong GUM để có khoảng tin cậy xấp xỉ cho đại lượng đo. Những khoảng tin cậy này dựa trên đại lượng then chốt xấp xỉ phân bố t giả định thu được từ mô hình đo (1). Trong quy trình này, các đại lượng chưa biết μ_1, ..., μ_p được ước lượng bằng các giá trị x, ..., x_p thu được từ phép đo vật lý hoặc từ các nguồn khác. Một số giá trị x_i có thể là trung bình mẫu hoặc hàm dữ liệu khác dùng để ước lượng các đại lượng μ_i, i = 1, ..., m. Độ không đảm bảo chuẩn kèm theo $u(x_i)$ của chúng cũng được đánh giá từ dữ liệu bằng các phương pháp thống kê, thường sử dụng độ lệch chuẩn mẫu hoặc sử dụng quy trình dựa trên thứ hạng ổn định. Các phương pháp như vậy được gọi là đánh giá độ không đảm bảo Loại A. Bậc tự do ν_i kèm theo $u(x_i)$ được xác định từ cỡ mẫu dùng để ước lượng μ_i.

8.1.6 Vì các phép đo vật lý không phải lúc nào cũng có thể hoặc khả thi đối với một số μ_i, nên các ước lượng x_i của μ_i đối với một số i, ví dụ: $i = m + 1$, ..., p, thu được bằng các đánh giá chủ quan (hoặc có khả năng chủ quan) và sử dụng cùng với x_i, đối với i = 1,..., p, thu được từ đánh giá độ không đảm bảo Loại A. Do đó, loại thông tin phi thống kê được sử dụng để ước lượng μ_{m+1}, ..., μ_p, bằng cách sử dụng đánh giá độ không đảm bảo Loại B, bao gồm cả lập luận khoa học, quy định kỹ thuật của nhà sản xuất hoặc thông tin liên quan gián tiếp hoặc chưa quy định đầy đủ khác.

CHÚ THÍCH: Đôi khi độ không đảm bảo thu được bằng cả đánh giá độ không đảm bảo Loại A và Loại B.

8.1.7 GUM khuyến nghị nên sử dụng cùng một mô hình đo liên hệ đại lượng θ với các đại lượng đầu vào μ_1, ..., μ_p để tính y từ x_1, ..., x_p. Do đó, giá trị đo được (hoặc ước lượng) y của θ thu được là

$$y = f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_p),$$

đó là, Y được đánh giá, $y = f(x_1,\dots, x_p)$, được lấy là giá trị đo được của θ.

8.1.8 Trong GUM, định luật lan truyền độ không đảm bảo được sử dụng để đánh giá độ không đảm bảo chuẩn, $u(y)$, gắn với y. Độ không đảm bảo chuẩn $u(x_1),\dots, u(x_p)$ gắn với các giá trị $x = (x_1,\dots, x_p)$ được sử dụng trong khai triển chuỗi Taylor bậc 1 của hàm $f(x_1,\dots, x_p)$ tại $\mu_1, \dots, \mu_p$

$$f(x_1,\dots,x_p) \approx f(\mu_1, \dots, \mu_p) + \sum_{i=1}^{p} c_i(x_i - \mu_i) \tag{6}$$

Ký hiệu $\mu_1, \dots, \mu_p$ bằng μ, đạo hàm riêng

$$c_i = \frac{\partial f}{\partial \mu_i}\Big|_{\mu=x}$$

được gọi là hệ số độ nhạy. Áp dụng định luật lan truyền độ không đảm bảo trong GUM sẽ cho độ không đảm bảo chuẩn xấp xỉ gắn với y:

$$u(y) \approx \sqrt{\sum_{i=1}^{p} c_i^2 u^2(x_i) + 2\sum_{i<j} c_i c_j u(x_i, x_j)}, \tag{7}$$

trong đó $u(x_i, x_j)$ là hiệp phương sai giữa X_i và X_j.

8.1.9 Để đánh giá độ không đảm bảo chuẩn $u(y)$, GUM sử dụng bậc tự do hiệu dụng ν_{eff} tính từ công thức Welch-Satterthwaite,

$$\nu_{eff} = \frac{u^4(y)}{\sum_{i=1}^{p} \frac{c_i^4 u^4(x_i)}{\nu(x_i)}}. \tag{8}$$

CHÚ THÍCH: Tài liệu tham khảo [15] thảo luận về tính chất phản trực quan theo đó trong nghiên cứu liên phòng thí nghiệm khoảng tin cậy dựa trên phép xấp xỉ Welch-Satterthwaite đối với chênh lệch giữa các phòng thí nghiệm có thể ngắn hơn đối với một trong các thành phần của nó.

8.1.10 Cuối cùng, để thiết lập khoảng tin cậy cho θ, đại lượng then chốt xấp xỉ,

$$W(y,\theta) = \frac{y-\theta}{u(y)} \tag{9}$$

được sử dụng. Theo GUM,

$$W(Y, \theta) \sim t(\nu_{eff}), \tag{10}$$

đó là, $W(Y, \theta)$ là đại lượng then chốt xấp xỉ có phân bố t với ν_{eff} bậc tự do. Khoảng tin cậy 100(1 – α) %

$$y \pm u(y)\ t_{\nu_{eff},1-\alpha/2}, \tag{11}$$

đối với θ khi đó có thể được khuyến nghị là khoảng độ không đảm bảo 100(1 – α) %. Một nửa độ rộng $t_{\nu_{eff},1-\alpha/2}u(y)$ của khoảng này được gọi là độ không đảm bảo mở rộng gắn với y.

8.1.11 Khuyến nghị này thống nhất với thực hành thống kê chuẩn khi tất cả các độ không đảm bảo được xác định bằng đánh giá Loại A, trong trường hợp mà ước lượng thống kê được sử dụng phổ biến nhất đối với đại lượng đầu vào cụ thể μ là trung bình mẫu của n giá trị quan trắc. Phương pháp truyền thống dùng cho tổng hợp dữ liệu để thu được độ không đảm bảo chuẩn Loại A của ước lượng này là $S/\sqrt{n}$ với $n-1$ bậc tự do. Điều này dựa trên thực tế là $(n-1)S^2/\sigma^2$ có phân bố Khi-bình phương với $n-1$ bậc tự do. Phương pháp này áp dụng cho các thống kê tổng quát hơn có dạng $Y = G(X_1,\dots,X_p)$, trong đó các ước lượng, X_i, $i = 1,\dots,p$ tuân theo định lý giới hạn trung tâm. Thực sự trong tình huống này, độ lệch chuẩn của Y có thể tính xấp xỉ bằng biểu thức (7) với $u(x_i, x_j)$ thay bằng $\mathrm{Cov}(X_i,X_j)$.

Phương pháp GUM thể hiện trí tuệ tập thể của nhiều nhà khoa học đo lường nhưng bị giới hạn bởi các giả định về

- tính tuyến tính cục bộ của hàm f: lý tưởng là hệ số độ nhạy không nên biến động nhiều và không bị triệt tiêu;
- tính chuẩn của phân bố xác suất các ước lượng điểm $Y = f(X_1,\dots,X_p)$: có thể không duy trì ngay cả là xấp xỉ đối với các mẫu nhỏ;
- giá trị sử dụng của công thức Welch-Satterthwaite (8): có thể không áp dụng được khi các đại lượng đầu vào phụ thuộc lẫn nhau, đại lượng đầu vào không theo phân bố chuẩn và độ không đảm bảo chuẩn khác nhau (bậc tự do đối với các phân bố không liên quan đến luật Khi-bình phương rất khó để lý giải, chúng thực sự không được dùng trong lý thuyết thống kê).

8.1.12 Để phát triển biểu thức (7) trong bối cảnh tần suất, có thể sử dụng các khái niệm lý thuyết quyết định thống kê và giải thích phương sai (độ không đảm bảo chuẩn bình phương) $u^2(y)$ như là sai số bình phương trung bình của ước lượng thống kê của $f(x_1, x_2,\dots, x_p)$. Có thể thực hiện các bước này với điều kiện là các đại lượng có độ không đảm bảo được xác định bằng đánh giá Loại B, đó là $x_{m+1},\dots,x_p$, được loại bỏ bằng cách lấy tích phân trên phân bố của chúng. Xem Tài liệu tham khảo [5]. Nếu f "đủ gần để là tuyến tính", thì biểu thức (7) sẽ cho xấp xỉ bậc một của sai số bình phương trung bình.

8.1.13 Thảo luận trong Ví dụ 1 cung cấp một quy trình tần suất thông thường khác để có được khoảng tin cậy.

8.2 Khoảng độ không đảm bảo bootstrap

8.2.1 Bootstrap là một chiến lược lấy mẫu lại [16] dùng cho ước lượng các tham số của phân bố như phương sai và xác định khoảng tin cậy cho các tham số khi dạng của phân bố cơ bản là chưa biết. Ý tưởng chính cho phương pháp bootstrap là mối quan hệ giữa phân bố xác suất tích lũy (CDF) F của Y và một mẫu từ F là tương tự với quan hệ giữa một CDF được ước lượng $\hat{F}$, nó có thể không phải là phân bố thực nghiệm tạo ra bởi mẫu đó, và mẫu thứ hai lấy từ $\hat{F}$. Khi không có sẵn F, việc lấy mẫu từ F không thể thực hiện được, nhưng máy tính hiện đại cho phép lấy một số lượng lớn từ $\hat{F}$. Vì vậy, ta sử dụng mẫu sơ cấp để hình thành xấp xỉ $\hat{F}$ của F rồi sau đó tính phân bố lấy mẫu của tham số ước lượng dựa trên $\hat{F}$. Tính toán này được thực hiện bằng cách lấy nhiều mẫu thứ cấp và tạo thành ước lượng (hoặc hàm của ước lượng) cho từng mẫu thứ cấp. Nếu $\hat{F}$ là xấp xỉ tốt của F thì H, phân bố lấy mẫu của ước lượng dựa trên $\hat{F}$, thường là xấp xỉ tốt cho phân bố lấy mẫu đối với ước lượng dựa trên F. H thường được gọi là phân bố bootstrap của tham số.

8.2.2 Có hai loại quy trình bootstrap hữu ích tương ứng với suy luận phi tham số và suy luận tham số. Bootstrap phi tham số dựa vào xem xét phân bố thực nghiệm $\hat{F}$ tạo ra bởi mẫu sơ cấp từ F. Trong thiết lập bootstrap tham số, phân bố xác suất F là thành phần của một họ tham số quy định và $\hat{F}$ thu được bằng cách ước lượng (các) tham số từ dữ liệu.

CHÚ THÍCH: Vì trong các vấn đề đo lường điển hình, tập hợp dữ liệu không đủ lớn để đảm bảo giá trị sử dụng của cách tiếp cận bootstrap phi tham số nên cách tiếp cận này không được xem xét ở đây.

8.2.3 Giả định chính sử dụng trong thiết lập khoảng tin cậy GUM là (10). Giả định này có thể không gần đúng ngay cả với các vấn đề đơn giản. Tuy nhiên, bootstrap cho phép có được khoảng tin cậy mà không liên quan đến việc đưa ra các giả định như (10). Một cách để có được các khoảng này là cách tiếp cận "bootstrap-t". Quy trình này tạo ra phân bố thực nghiệm cho đại lượng then chốt xấp xỉ $W(Y,\theta)$ [để thay thế phân bố t trong (10)]. Khi (10) là đúng, phân bố bootstrap- t sẽ tái tạo phân bố t. Khi đó, phân bố bootstrap- t thực nghiệm được sử dụng để thiết lập khoảng tin cậy theo cùng một cách mà phân bố t được sử dụng để xây dựng (11).

Về mối quan hệ giữa bootstrap và các phương pháp đề xuất trong GUMS1, xem 12.2.

8.2.4 Phác thảo về việc tạo ra mẫu bootstrap như sau đây. Giả định rằng x_1 và $u(x_1)$ là trung bình và độ lệch chuẩn đối với biến ngẫu nhiên X_1, biến này được giả định là theo phân bố xác suất trong một họ tham số xác định. Để minh họa, ở đây sử dụng phân bố Gauss:

a) x_1 và $u(x_1)$ là trung bình và độ lệch chuẩn ước lượng của biến ngẫu nhiên cỡ k theo phân bố Gauss.

b) Từ $N(x_1, u^2(x_1))$, tạo ra một mẫu có cỡ mẫu k, đó là, $\{x^*_{1,1},\ldots,x^*_{1,k}\}$.

c) Từ $\{x^*_{1,1},\ldots,x^*_{1,k}\}$, tính trung bình mẫu x^*_1 và sai số chuẩn mẫu $u(x^*_1)$.

$\{x^*_1, u(x^*_1)\}$ là mẫu bootstrap của X_1. Tương tự, đối với một số B cho trước, B mẫu bootstrap có thể được tạo ra cho biến bất kỳ.

8.2.5 Giống như GUM lấy $(x_i, u(x_i))$, với $i = 1,\ldots, p$), làm đầu vào để tạo ra y, $u(y)$ và $W(Y, \theta)$, các mẫu bootstrap $\{x^*_i, u(x^*_i)\}$, $i = 1,\ldots, p$ (xem 8.2.4) có thể được lấy làm đầu vào, để tạo ra y^*, $u(y^*)$, và

$$W^* = W(y^*, y) = \frac{y^* - y}{u(y^*)} \tag{12}$$

8.2.6 Để có được phân bố bootstrap cho $W(Y, \theta)$, với giá trị B lớn phù hợp, chẳng hạn 100 000, tạo ra B mẫu bootstrap $\{x^*_i(b), u(x^*_i(b))\}$, $i = 1,\ldots, p$, và với từng mẫu tính $W^*(b)$, $b = 1,\ldots, B$. Phân vị thứ 100 α của phân bố t bootstrap của $W(Y, \theta)$ khi đó được tính xấp xỉ bằng giá trị $\hat{t}_\alpha$ sao cho

$$\left|\left\{W^{*}(b) \le \hat{t}_{\alpha}\right\}\right| / B = \alpha,$$

trong đó |A| là số phần tử trong tập A. Cuối cùng, khoảng tin cậy bootstrap t 100(1 – α) % là

$$\left(y - \hat{t}_{1-\alpha/2} \cdot u(y) \quad y + \hat{t}_{\alpha/2} \cdot u(y)\right) \tag{13}$$

Phân vị Student-t là đối xứng qua 0, và kết quả là, (11) phải luôn đối xứng qua y. Ngược lại , phân vị bootstrap t sử dụng ở (13) có thể là bất đối xứng qua 0, dẫn đến khoảng độ không đảm bảo bất đối xứng qua y, điều này có thể cung cấp mô tả chính xác hơn về tình huống thực tế trong một số ứng dụng. Chi tiết về quá trình này trong việc thiết lập khoảng độ không đảm bảo 95 % được trình bày trong thuật toán dưới đây.

a) Đối với i = 1,..., p, sử dụng các phân bố đã cho của X_i, tạo B mẫu bootstrap $(x_i^*(1)), u(x_i^*(1))), \ldots, (x_i^*(B)), u(x_i^*(B)))$.

b) Đối với từng mẫu bootstrap $(x_i^*(b)), u(x_i^*(b)))$, i = 1,..., p, và b = 1,..., B, tính $y^*(b)$, $u(y^*(b))$ và $W^*(b) = (y^*(b) - y)/u(y^*(b))$ theo GUM.

c) Ước lượng phân vị thứ 100α của phân bố bootstrap t của $W(Y,\theta)$ bằng giá trị $\hat{t}_{\alpha}$.sao cho $\left|\left\{W^{*}(b) \le \hat{t}_{\alpha}\right\}\right| / B = \alpha$.

d) Khoảng tin cậy bootstrap t 95 % là $\left(y - \hat{t}_{0,975} \cdot u(y)\ y + \hat{t}_{0,025} \cdot u(y)\right)$.

8.2.7 Mẫu bootstrap có thể được sử dụng để thay cho $u(y)$ bằng cách ước lượng độ lệch chuẩn của Y, khi phép xấp xỉ Taylor (6) được coi là không thích hợp. Để làm được điều này, với i = 1,..., p và b = 1,..., B, chỉ các ước lượng đầu vào $x_i^*(b)$được tạo ra. Đối với mỗi mẫu bootstrap, $y^*(b) = f(x_1^*(b), \ldots, x_p^*(b))$được đánh giá. Ước lượng bootstrap của độ không đảm bảo chuẩn gắn với y là độ lệch chuẩn mẫu của B lần lặp:

$$u_c(y) = \sqrt{\sum_{b=1}^{B}\left[y^*(b) - y^*(\cdot)\right]^2 / (B-1)}, \qquad y^*(\cdot) = \sum_{b=1}^{B} y^*(b)/B.$$

8.2.8 Cuối cùng, khi phép xấp xỉ Taylor có thể không thích hợp và có sự bất đối xứng đáng kể trong phân bố cơ bản đối với Y, một bootstrap lồng của B_1 x B_2 mẫu bootstrap có thể được thực hiện để xây dựng khoảng bootstrap t sử dụng ước lượng độ lệch chuẩn bootstrap. B_1 mẫu bootstrap của các ước lượng đầu vào và y^* tương ứng được tạo ra. Đối với mỗi mẫu bootstrap, $u_c(y^*)$ được tính bằng B_2 mẫu bootstrap mức hai, và đánh giá

$$\frac{y^* - y}{u(y^*)}$$

Tập hợp B_1 các tỷ số này sau đó được dùng để ước lượng các phân vị của $W(Y, \theta)$, dẫn đến việc xây dựng khoảng bootstrap t như trong (13). Thuật toán để xây dựng khoảng độ không đảm bảo 95 % sử dụng bootstrap lồng được trình bày dưới đây.

a) Đối với i = 1,..., p, sử dụng phân bố của X_i, tạo ra B_1 mẫu bootstrap mức một $x_i^*(1),\ldots,x_i^*(B_1)$.

b) Đối với từng mẫu bootstrap mức một $x_i^*(b_1)$, i = 1,..., p, b_1 = 1,..., B_1 tính $y^*(b_1) = f(x_1^*(b_1),\ldots,x_p^*(b_1))$ và $W^*(b_1) = (y^*(b_1) - y)/u(y^*(b_1))$, trong đó $u(y^*(b_1)))$ được xác định bằng bootstrap mức hai sử dụng thuật toán sau đây:

1. Đối với i = 1,..., p, sử dụng phân bố của μ_i, tạo ra B_2 mẫu bootstrap mức hai $x_i^*(1),\ldots,x_i^*(B_2)$.

2. Đối với mỗi mẫu bootstrap mức hai, đánh giá $y^*(b_2) = f(x_1^*(b_2),\ldots,x_p^*(b_2))$

3. Tạo ước lượng bootstrap của độ không đảm bảo chuẩn của $y^*(b_1)$ như độ lệch chuẩn mẫu

$$u(y^*(b_1)) = \sqrt{\sum_{b_2=1}^{B_2}[y^*(b_2) - y^*(\cdot)]^2 / (B_2 - 1)}$$

của B_2 phép lặp, trong đó $y^*(\cdot) = \sum_{b_2=1}^{B_2} y^*(b_2)/B_2$.

c) Ước lượng phân vị thứ 100α của phân bố bootstrap t của $W(Y, \theta)$ bằng giá trị $\hat{t}_\alpha$ sao cho $|\{W^*(b_1) \le \hat{t}_\alpha\}|/B_1 = \alpha$.

d) Khoảng tin cậy "bootstrap-t lồng" 95 % là

$$\left(y - \hat{t}_{0,975} \cdot u(y)\ y + \hat{t}_{0,025} \cdot u(y)\right).$$

Mặc dù đây là cách tiếp cận tổng quát hơn, bootstrap lồng liên quan đến máy tính nhiều hơn và khó thực hiện hơn. Phương pháp bootstrap đơn giản hơn được chọn để phân tích tất cả các ví dụ.

8.3 Ví dụ 1

8.3.1 Khái quát

8.3.1.1 Như một minh họa, xét mô hình thống kê được cho trong Ví dụ 1 ở Điều 7 là

$Y_i = \theta + \beta + \varepsilon_i$, i = 1,..., n, (14)

trong đó θ là đại lượng đo, β thể hiện nền và ε_i là các sai số $N(0, \sigma^2)$ độc lập. Đối với giá trị β cố định, với y ký hiệu cho trung bình của dữ liệu, công thức đo cho mô hình này là $\theta = f(\beta, y) = y - \beta$.

8.3.1.2 Nếu nền, β, có phân bố đều trên khoảng $(a - d\ \ a + d)$, thì khoảng đối với θ rút ra bằng cách sử dụng GUM là

$$\overline{Y} - a \pm 2\sqrt{\frac{\sigma^2}{n} + \frac{d^2}{3}}.$$

Tài liệu tham khảo [5] đề cập đến các tính chất của các khoảng này và so sánh chúng với khoảng

$$\bar{Y} - a \pm \left[2\sqrt{\frac{\sigma^2}{n}} + d \right], \qquad (15)$$

được khuyến nghị bởi Eisenhart [17] và có thể được phát triển như dưới đây. Vì phân bố điều kiện của $\bar{Y}$ đối với β cho trước là phân bố Gauss, $N(\theta + \beta, \sigma^2/n)$

$$P\left(|\bar{Y} - \theta - \beta| \le \frac{2\sigma}{\sqrt{n}} \right) \ge 0{,}95,$$

trong khi

$$P(|\,a - \beta\,| \le d) = 1.$$

Từ đó kéo theo khoảng Eisenhart trong (15) vẫn bảo toàn, cụ thể là

$$P\left(|\bar{Y} - a - \theta| \le \frac{2\sigma}{\sqrt{n}} + d \right) \ge 0{,}95, \qquad (16)$$

8.3.1.3 Tuy nhiên, nếu $d > 12\sigma/\sqrt{n}$, thì khoảng khuyến nghị trong GUM chứa khoảng (15) chứng tỏ sự khác biệt giữa hai cách tiếp cận này.

8.3.1.4 Khoảng (15) có thể được điều chỉnh cho tỷ số phân bố $\sqrt{n}(\bar{Y} - a - \beta)/S$. Cũng có thể điều chỉnh khoảng này cho các phân bố khác của nền (hình tam giác, hình thang, v.v.). Còn có các phương pháp tần suất khác để xây dựng khoảng tin cậy trong tình huống này. Thực sự trong mô hình (14), $\bar{Y}$ tổng hợp tất cả các thông tin dữ liệu về θ (tức là, $\bar{Y}$ là thống kê đủ cho θ) với mật độ xác suất

$$\frac{\sqrt{n}}{2\sqrt{2\pi}\sigma d} \int_{a-d}^{a+d} e^{-0{,}5n(\bar{y}-\theta-\beta)^2/\sigma^2} d\beta$$

Dạng đặc biệt của phân bố này cho phép suy ra [14] các khoảng tin cậy khác (tất cả tập trung ở ước lượng hợp lý cực đại, $\bar{Y} - a$, nhưng khác nhau về độ dài).

8.3.2 Ví dụ 1a

Ví dụ đơn giản nêu trong Điều 7 tổng hợp các giá trị đo được trong mô hình (14) với $\bar{y}$ = 3,537 và $u(\bar{y})$ = 0,153. Giá trị nêu sau được thay cho $\sigma\sqrt{n}$ trong bất đẳng thức (16) và hệ số 2 cần được thay bằng phân vị của phân bố t với bậc tự do hiệu dụng 5,15. Trong Ví dụ 1a, nền β có thể được ước lượng từ các giá trị đo được được coi là lấy từ phân bố Gauss, dẫn đến $\bar{b}$ = 1,228 và $u(\bar{b})$ = 0,059. Ước lượng thu được của θ là $\bar{y} - \bar{b}$ = 2,309 với độ không đảm bảo chuẩn kèm theo $\sqrt{u^2(\bar{y}) + u^2(\bar{b})}$ = 0,164. Khoảng tin cậy GUM là

$$2{,}309 \pm 2{,}548 \times 0{,}164 = 2{,}309 \pm 0{,}417 = (1{,}892;\ \ 2{,}727).$$

Khoảng tin cậy bootstrap-*t* 100(1 - α) % theo (13) là $(2{,}309 - 0{,}164 \cdot \hat{t}_{1-\alpha/2};\ 2{,}309 + 0{,}164 \cdot \hat{t}_{\alpha/2})$, trong đó $\hat{t}_\beta$ là phân vị thứ 100β của W^* của (12).

Về lợi ích của người sử dụng ngôn ngữ R và WinBUGS, một số đoạn mã R [18] và đoạn WinBUGS [19] được dùng để minh họa một số khái niệm trong tiêu chuẩn này. Đối với Ví dụ 1a, chương trình R để tạo ra B = 10 000 thể hiện của W^* được cho dưới đây.

```
B = 10000
y.star = rnorm(B, mean=3.537, sd=0.153)
u.y.star = 0.153 * sqrt(rchisq(B,  df=4)/4),
b.star = rnorm(B, mean=1.228, sd=0.059)
u.b.star = 0.059 * sqrt(rchisq(B,  df=4)/4)
w.star = ((y.star-b.star) - 2.309)/sqrt(u.y.star^2+u.b.star^2)
```

Khoảng tin cậy bootstrap-t 95 % dựa trên các phân vị 0,025 và 0,975 của phân bố mô phỏng là

```
2.309 - quantile(w.star, c(0.975, 0.025))*0,164
## 1.895754 2.728817
```
[2)]

Tức là, khoảng tin cậy bootstrap-t 95 % được cho bởi (1,896; 2,729)[3)]

8.3.3 Ví dụ 1b

Khi không có dữ liệu thống kê của nền, thay vào đó β được giả định là có phân bố đều trong khoảng (1,126 1,329). Khi đó, khoảng tin cậy xấp xỉ rút ra từ việc sử dụng GUM là

$$3{,}537 - 1{,}228 \pm 2{,}533\sqrt{\frac{0{,}342^2}{5} + \frac{0{,}102^2}{3}} = 2{,}310 \pm 0{,}415 = (1{,}895 \quad 2{,}724).$$

Khoảng tin cậy Eisenhart rộng hơn, cụ thể là

$$3{,}537 - 1{,}228 \pm \left[2{,}776\frac{0{,}342}{\sqrt{5}} + 0{,}102\right] = 2{,}310 \pm 0{,}526 = (1{,}783 \quad 2{,}836).$$

Tương tự Ví dụ 1a, khoảng tin cậy bootstrap-t có thể xây dựng cho θ. Đối với ví dụ này, các ước lượng và độ không đảm bảo chuẩn kèm theo đối với y, β và θ có trị số giống với trong Ví dụ 1a, ngoại trừ β được xác định dựa trên kinh nghiệm hoặc ý kiến chuyên gia và độ không đảm bảo gắn với của nó thu được bằng đánh giá Loại B. Do đó, các thể hiện của W^* được tạo ra theo cách khác với trong Ví dụ 1a, đó là, chỉ tạo ra mẫu bootstrap b^* và độ không đảm bảo gắn với nó. Mẫu bootstrap b^* lúc này được tạo ra từ phân bố đều (1,126 1,329) đã biết với độ không đảm bảo chuẩn 0,059. Dòng lệnh R dùng cho việc tạo ra B = 10 000 thể hiện W^* như dưới đây.

[2)] Trong các ví dụ tính toán sử dụng R, WinBugs, hoặc các phần mềm khác, đầu ra được cho như báo cáo bởi định dạng chuẩn của phần mềm. Như chỉ ra rõ ràng bằng các giá trị của độ không đảm bảo báo cáo, không phải tất cả các chữ số trong đầu ra đều là chữ số có nghĩa. Cũng cần lưu ý là đầu ra chuẩn từ các phần mềm này sử dụng dấu chấm chứ không phải dấu phẩy làm chỉ thị thập phân.

[3)] Trong độ không đảm bảo mở rộng các giá trị làm tròn tương ứng đến ba chữ số có nghĩa. Chú thích: đối với phương pháp Monte Carlo, việc tính toán lại các ví dụ sẽ dẫn đến sai số ngẫu nhiên do mô phỏng.

```
B = 10000

y.star = rnorm(B, mean=3.537, sd=0.153)

u.y.star = 0.153 * sqrt(rchisq(B,  df=4)/4)

b.star = runif(B, min=1.126, max=1.329)

u.b.star = 0.059

w.star = ((y.star-b.star) - 2.309)/sqrt(u.y.star^2+u.b.star^2)
```

Khoảng tin cậy bootstrap-*t* 95 % dựa trên các phân vị 0,025 và 0,975 của phân bố tạo ra là

```
2.309 - quantile(w.star, c(0.975, 0.025))*0,164

## 1.918643 2.699749
```

Tức là, khoảng tin cậy bootstrap-t 95 % được cho bởi (1,919 2,700).

8.3.4 Ví dụ 1c

Vì $\bar{y}$ = 1,196, $s_{\bar{y}}$ = 0,047, cả hai khoảng đều có điểm cận dưới âm. Nếu trung bình θ đã biết là dương thì các điểm đầu mút này được thay bằng 0 dẫn đến khoảng khuyến nghị của GUM (0; 0,124) và khoảng Eisenhart (0 0,202).

Chương trình R dùng để tạo ra B = 10 000 thể hiện của W^* để tạo ra khoảng bootstrap giống như Ví dụ 1b với $\bar{y}$ = 1,196 và $u(\bar{y})$ = 0,047.

```
B = 10000

y.star = rnorm(B, mean=1.196, sd=0.047)

u.y.star = 0.047 * sqrt(rchisq(B,  df=4)/4)

b.star = runif(B, min=1.126, max=1.329)

u.b.star = 0.059

w.star = ((y.star-b.star) + 0.032)/sqrt(u.y.star^2+u.b.star^2)
```

Khoảng tin cậy bootstrap-*t* 95 % không cắt là

```
-0.032-quantile(w.star, c(0.975, 0.025))*0,075

## -0.1762648  0.1128422
```

Tức là, khoảng tin cậy bootstrap-*t* 95 % được cho bởi (-0,176; 0,113).

Vì θ đã biết là dương nên khoảng tin cậy bootstrap-*t* 95 % đã cắt đối với θ là (0; 0,113).

9 Cách tiếp cận Bayes để đánh giá độ không đảm bảo

9.1 Phương pháp cơ bản

9.1.1 Trong đo lường, đại lượng đo và các biến đầu vào của mô hình (1) là các đại lượng vật lý có giá trị đại lượng cố định. Tuy nhiên, trong cách tiếp cận Bayes, các tham số tương ứng μ_i và θ được coi là các biến ngẫu nhiên theo nghĩa là phân bố xác suất của chúng tổng hợp hiểu biết về các đại lượng này.

9.1.2 Khuôn khổ Bayes sử dụng định nghĩa xác suất cho phép xác định các phân bố xác suất mà không cần dữ liệu vật lý, ví dụ, sử dụng quy định kỹ thuật của nhà sản xuất hoặc kiến thức chuyên môn khác. Tuy nhiên, trong các ứng dụng đo lường điển hình, có các giá trị (dữ liệu) đo được của các đại lượng vật lý có thể sử dụng để ước lượng một hoặc nhiều đại lượng đầu vào. Trong trường hợp này, có thể thu được hàm mật độ xác suất cho đại lượng bằng cách sử dụng định lý Bayes như dưới đây. Lấy $p(\mu_i)$ là hàm mật độ xác suất cho μ_i, như đã cho trước khi thu được dữ liệu vật lý. Hàm này được gọi là mật độ tiên nghiệm cho μ_i. Lấy Y ký hiệu cho biến ngẫu nhiên có thể hiện y (dữ liệu) tồn tại. Mật độ xác suất $p(y|\mu_i)$ của Y được gọi là mô hình thống kê. Trong khuôn khổ Bayes, vì μ_i là biến ngẫu nhiên nên ký hiệu $|$ thể hiện thực tế là mật độ xác suất của Y là có điều kiện đối với (hoặc phụ thuộc vào) μ_i. Đối với một thể hiện cụ thể $p(y|\mu_i)$ của Y, xem như hàm của μ_i được gọi là hàm hợp lý. Áp dụng định lý Bayes,

$$p(\mu_i|y)=\frac{p(y|\mu_i)p(\mu_i)}{\int p(y|\mu_i)p(\mu_i)d\mu_i} \quad (17)$$

là mật độ hâu nghiệm của μ_i tổng hợp kiến thức về μ_i sau khi dữ liệu y_i được quan trắc.

9.1.3 Khi không có hiểu biết trước đó về μ_i, thì sử dụng phân bố tiên nghiệm gọi là phi thông tin [20]. Trong trường hợp có thông tin trước đó, thì thể hiện bằng phân bố xác suất có thông tin. Đây là một trong các cơ chế, trong cách tiếp cận Bayes, nhằm bao gồm thông tin được sử dụng để thực hiện đánh giá độ không đảm bảo Loại B. Dạng của hàm hợp lý thường được chọn dựa trên hiểu biết về quá trình tạo ra dữ liệu.

9.1.4 Dạng của hàm hợp lý và các mật độ tiên nghiệm xác định hình dạng của mật độ hậu nghiệm. Điều quan trọng là chọn kỹ hàm hợp lý và mật độ tiên nghiệm và thực hiện phân tích độ nhạy các kết quả về những thay đổi hợp lý trong các phân bố này. Đối với các phân bố tiên nghiệm, điều này có thể nghĩa là so sánh các kết quả của việc sử dụng nhiều mật độ khác nhau. Kiểm nghiệm tính thích hợp của hàm hợp lý (mô hình thống kê mô tả dữ liệu đo) là một hình thức xác nhận giá trị mô hình [21], áp dụng giống nhau với các mô hình Bayes, tần suất và dựa vào sự tin tưởng.

9.1.5 Có thể giải thích định nghĩa về độ không đảm bảo đo nêu trong phần giới thiệu trong bối cảnh thống kê Bayes khi quy về phân bố xác suất hậu nghiệm đối với đại lượng đo θ, đó là, độ không đảm bảo chuẩn là độ lệch chuẩn của biến (đại lượng) ngẫu nhiên đặc trưng bởi phân bố xác suất này. Để thu được độ lệch chuẩn này, trước tiên cần tìm phân bố xác suất đồng thời của μ_i, rồi sau đó áp dụng công thức thay đổi của biến [14] để rút ra phân bố cho θ. Mô men của phân bố này có thể thu được đơn giản hơn như trình bày dưới đây. Đối với hàm $h(\theta)$, có giá trị mong muốn $E(h(\theta) = \int \ldots \int h(f(\mu_1, \ldots, \mu_p)) p(\mu_1, \ldots, \mu_p) d\mu_1 \ldots d\mu_p$. Phương sai tương ứng có thể thu được là $Var(\theta) = E(\theta^2) - [E(\theta)]^2$. Thông thường, tích phân cần thiết được thực hiện bằng các phương pháp Monte Carlo [20].

9.1.6 Khi μ_i là các biến ngẫu nhiên độc lập, phân bố xác suất đồng thời của chúng là tích của các phân bố riêng lẻ. Tuy nhiên, trong nhiều tình huống, μ_i không độc lập như khi phân bố xác suất của Y là hàm của μ_1 và μ_2, tức là, $p(y|\mu_1, \mu_2)$ là mô hình thống kê và $p(\mu_1, \mu_2) \neq p(\mu_1)\, p(\mu_2)$. Khi đó, mật độ hậu nghiệm đối với (μ_1, μ_2) thu được là

$$p(\mu_1, \mu_2 \mid y) = \frac{p(y \mid \mu_1, \mu_2) p(\mu_1, \mu_2)}{\int p(y \mid \mu_1, \mu_2) p(\mu_1, \mu_2) d\mu_1 d\mu_2}$$

9.1.7 Tình huống phổ biến dẫn đến sự phụ thuộc này là khi mô hình thống kê là hàm của θ cũng như của một số μ_i. Cả hai ví dụ xét ở đây đều thuộc loại này, minh họa điểm mà theo cách tiếp cận Bayes, bất cứ khi nào có sẵn dữ liệu đo, quá trình quy định phân bố xác suất liên quan đòi hỏi xác định thích hợp mô hình thống kê. Thực hiện việc này sẽ tự động dẫn đến các hàm hợp lý cần thiết cho ứng dụng định lý Bayes và các mật độ hậu nghiệm thích hợp. Quá trình này có thể được tổng hợp như dưới đây.

a) Xác định tất cả các dữ liệu đo liên quan đến các đại lượng vật lý quan tâm (tham số).

b) Quy định mô hình thống kê (còn gọi là mô hình quan trắc) liên hệ dữ liệu với các tham số, có thể là μ_i hoặc đôi khi là đại lượng đo θ.

c) Quy định các phân bố tiên nghiệm đối với tất cả các tham số liên quan.

d) Áp dụng định lý Bayes để có được các phân bố hậu nghiệm cho các tham số.

e) Tính trung bình hậu nghiệm và độ lệch chuẩn hậu nghiệm của đại lượng đo.

f) Tiến hành phân tích độ nhạy của các kết quả đối với những thay đổi hợp lý trong phân bố tiên nghiệm.

9.1.8 Khi thích hợp, có thể sử dụng phép xấp xỉ chuỗi Taylor và giả định tính chuẩn để tránh các tính toán số học. Cụ thể, có thể sử dụng khai triển chuỗi Taylor $f(\mu_1, \ldots, \mu_p)$ về các giá trị mong muốn của μ_i cùng với giả định tính chuẩn để công bố là $f(\mu_1, \ldots, \mu_p)$ xấp xỉ phân bố như $\nu(f(E(\mu_1), \ldots, E(\mu_p)), \omega^2)$, trong đó

$$\omega^2 = \sum_{i=1}^{p} c_i^2 \mathrm{Var}(\mu_i) + 2\sum_{i<j} c_i c_j \mathrm{Cov}(\mu_i, \mu_j).$$

$\mathrm{Cov}(\mu_i, \mu_j)$ ký hiệu cho hiệp phương sai của μ_i và μ_j, và c_i là đạo hàm riêng của θ đối với μ_i đánh giá tại các giá trị mong muốn của μ_i.

CHÚ THÍCH: Tương tự như công thức (6) và (7) được sử dụng ở 8.1.8, nhưng có sử dụng phép khai triển để tìm ước lượng phương sai của ước lượng θ chứ không phải cho θ.

9.2 Ví dụ 1

9.2.1 Khái quát

Quá trình này giờ được minh họa ở Ví dụ 1 trong Điều 7. Đại lượng đo trong ví dụ này được ký hiệu là θ. Mô hình đo được mô tả trong 8.3.1.1 là

$$\theta = \gamma - \beta \tag{18}$$

9.2.2 Ví dụ 1a

9.2.2.1 Có hai tập hợp dữ liệu liên quan: (i) năm giá trị đo được y_i, thu được độc lập, của tín hiệu cộng với nền, và (ii) năm giá trị đo được b_i, thu được độc lập, riêng của nền. Mỗi giá trị trong tập dữ liệu (i) được coi như thể hiện của biến ngẫu nhiên Y_i có phân bố Gauss với trung bình $\gamma = \theta + \beta$ và độ lệch chuẩn σ_Y, cũng tương tự đối với từng giá trị trong (ii) nhưng đối với biến ngẫu nhiên B_i, với trung bình β và độ lệch chuẩn σ_B. Do đó, mô hình thống kê đối với Y_i là

$$Y_i | \theta, \beta, \sigma_Y^2 \sim N(\theta + \beta, \sigma_Y^2),$$

và vì năm giá trị đo được là độc lập, nên

$$p(y_1, \ldots, y_5 | \theta, \beta, \sigma_Y) = \left(\frac{1}{\sigma_Y\sqrt{2\pi}}\right)^5 \exp\left\{-\frac{\sum_{i=1}^{5}(y_i - \theta - \beta)^2}{2\sigma_Y^2}\right\}.$$

9.2.2.2 Mô hình thống kê đối với B_i là

$$B_i | \beta, \sigma_B^2 \sim N(\beta, \sigma_B^2),$$

đó là,

$$p(b_1, \ldots, b_5 | \beta, \sigma_B) = \left(\frac{1}{\sigma_B\sqrt{2\pi}}\right)^5 \exp\left\{-\frac{\sum_{i=1}^{5}(b_i - \beta)^2}{2\sigma_B^2}\right\}.$$

9.2.2.3 Vì hai tập hợp quan trắc độc lập với nhau nên mô hình thống kê đối với Y và B là

$$p(y, b | \theta, \beta, \sigma_Y, \sigma_B) = p(b_1, \ldots, b_5 | \beta, \sigma_B)\, p(y_1, \ldots, y_5 | \theta, \beta, \sigma_Y).$$

9.2.2.4 Có bốn tham số, θ, β, σ_Y và σ_B, được ấn định cho phân bố tiên nghiệm. Trong ví dụ này, không có thêm thông tin về các tham số này ngoài việc chúng không âm, và do đó biến ngẫu nhiên sẽ được lấy là độc lập. Mong muốn rằng dạng của phân bố tiên nghiệm có ảnh hưởng ít nhất đến các kết quả phân tích. Ảnh hưởng này thu được từ việc sử dụng cái gọi là tiên nghiệm tham chiếu [20]. Đối với các tham số gắn với trung bình, tức là θ và β, mật độ này có thể tính xấp xỉ bằng

$\theta \sim$ Uniform $(0, c)$, $\beta \sim$ Uniform $(0, c)$

với một giá trị lớn cho c. Đối với tham số thang đo σ_Y và σ_B, mật độ tiên nghiệm tham chiếu

$p(\sigma_Y) = 1/\sigma_Y$, $p(\sigma_B) = 1/\sigma_B$

là không đúng, tức là, chúng không có tích phân bằng 1. Vì khía cạnh này có thể gây khó khăn trong tính toán số học nên mật độ đúng là

$\sigma_Y \sim$ Uniform $(0, c)$,

hoặc

$\sigma_Y \sim$ Gamma(c, c),

với các giá trị c lớn được sử dụng. Ký hiệu Gamma(ϕ_1, ϕ_2) thể hiện phân bố gamma với các tham số ϕ_1 và ϕ_2, tức là, đối với biến ngẫu nhiên X, mật độ xác suất này được cho bởi

$$p(x \mid \phi_1, \phi_2) = \frac{\phi_2^{\phi_1}}{\Gamma(\phi_1)} x^{\phi_1 - 1} e^{-x\phi_2}$$

Điều này hoàn thành quy định về các phân bố tiên nghiệm.

9.2.2.5 Áp dụng định lý Bayes dẫn đến mật độ hậu nghiệm đồng thời đối với θ, β, σ_Y và σ_B như sau:

$$p(\theta, \beta, \sigma_Y, \sigma_B | y, b) = \frac{p(y, b | \theta, \beta, \sigma_Y, \sigma_B) p(\theta) p(\beta) p(\sigma_Y) p(\sigma_B)}{\int p(y, b | \theta, \beta, \sigma_Y, \sigma_B) p(\theta) p(\beta) p(\sigma_Y) p(\sigma_B) d\theta\, d\beta\, d\sigma_Y\, d\sigma_B}.$$

Mật độ hậu nghiệm của đại lượng đo θ thu được bằng tích phân

$$p(\theta | y, b) = \int p(\theta, \beta, \sigma_Y, \sigma_B | y, b) d\beta\, d\sigma_Y\, d\sigma_B.$$

Phân bố hậu nghiệm này tổng hợp tất cả thông tin về θ sẵn có sau khi thu được các giá trị đo được. Kỳ vọng của phân bố này được lấy làm ước lượng của đại lượng vật lý và độ lệch chuẩn của phân bố này được sử dụng làm độ không đảm bảo chuẩn gắn với ước lượng này. Từ phân bố này sẽ dễ dàng thu được khoảng phủ cho đại lượng đo. Khoảng phủ này là khoảng các giá trị có thể có đối với θ với xác suất cố định. Trong thống kê Bayes, khoảng này được gọi là khoảng đáng tin cậy. Trong nhiều trường hợp, phương pháp số học được sử dụng để thực hiện các tích phân cần thiết khi áp dụng định lý Bayes. Một phương pháp khác có thể rút ra từ phân bố hậu nghiệm là phân bố Markov Chain Monte Carlo (MCMC) [22] sử dụng phần mềm WinBUGS [19]. Các dòng lệnh cho ví dụ này, với phân bố tiên nghiệm có c = 100, được cho như sau đây

```
Example1a{
theta~dunif(0,100)
beta~dunif(0,100)
gamma <- theta+beta
```

```
    sigma.Y~dunif(0,1)
    sigma.B~dunif(0,1)
    tau.Y <- 1/(sigma.Y*sigma.Y)
   tau.B <- 1/(sigma.B*sigma.B)
    for(i in 1:n){
    y[i]~dnorm(gamma,tau.Y)
    b[i]~dnorm(beta,tau.B)}
    }
```

Với dữ liệu cho trong 7.2, với n = 5, chương trình tạo ra trung bình hậu nghiệm của θ là 2,309 và độ lệch chuẩn hậu nghiệm là 0,247. Khoảng tin cậy 95 % đối với θ là (1,805, 2,815). Phân tích độ nhạy Bayes đối với những thay đổi về hình thức của bốn phân bố tiên nghiệm có thể được thực hiện bằng cách thay đổi giá trị của c (xem 9.2.2.4) và bằng cách thay thế các dòng

```
tau.Y~dgamma (1,0E-5,1,0E-5)
tau.B~dgamma 1,0E-5,1,0E-5)
```

cho bốn dòng

```
sigma. Y~dunif(0,1)
sigma. B~dunif(0,1)
tau. Y <- 1 / (sigma.Y*sigma.Y)
tau. B <- 1 / (sigma.B*sigma.B)
```

và so sánh các giá trị thu được của trung bình và độ lệch chuẩn hậu nghiệm. Các kết quả trong ví dụ này là ổn định với các thay đổi như vậy.

9.2.3 Ví dụ 1b

9.2.3.1 Thông tin về tham số nền β được cung cấp dưới dạng phân bố xác suất thu được bằng đánh giá độ không đảm bảo Loại B. Trong trường hợp này, mô hình quan trắc là chỉ về tập dữ liệu (i) trong Ví dụ 1a (9.2.2), đó là,

$$Y_i | \theta, \beta, \sigma_Y^2 \sim N(\theta + \beta, \sigma_Y^2).$$

9.2.3.2 Lúc này có ba tham số được ấn định cho phân bố tiên nghiệm. Đối với tham số nền β, mật độ tiên nghiệm dựa vào thông tin nêu trong Lời giới thiệu, đó là,

$$\beta \sim \text{Uniform } (1,126;\ 1,329).$$

Đối với θ và σ_Y,

$$\theta \sim \text{Uniform } (0, c),\ \sigma_Y \sim \text{Uniform } (0, c),$$

với giá trị lớn cho c.

Điều này hoàn thành các quy định kỹ thuật đối với phân bố tiên nghiệm.

9.2.3.3 Các dòng lệnh WinBUGS cho ví dụ này như dưới đây.

```
Example1b{
theta~dunif(0,100)
beta~dunif(1.126,1.329)
sigma.Y~dunif(0,1)
gamma <- theta+beta
tau.Y <- 1/(sigma.Y*sigma.Y) for (i in 1:n) {
y[i] ~dnorm (gamma,tau.Y)
}
```

Với dữ liệu nêu trong Lời giới thiệu, các dòng lệnh này tạo ra trung bình hậu nghiệm cho θ là 2,309 và độ lệch chuẩn hậu nghiệm đối với θ là 0,232. Khoảng tin cậy 95 % đối với θ là (1,832 2,788). Phân tích độ nhạy của các kết quả lại được thỏa mãn.

9.2.4 Ví dụ 1c

9.2.4.1 Khác biệt duy nhất với Ví dụ 1b là ở các giá trị đo được thực tế (lúc này gần với nền) và có thể sử dụng cùng mô hình và các dòng lệnh WinBUGS. Lúc này, trung bình hậu nghiệm cho θ là 0,069, độ lệch chuẩn hậu nghiệm cho θ là 0,067 và khoảng tin cậy 95 % cho θ là (0,000 0,188). Các kết quả này là ổn định với các thay đổi giá trị của c với các giá trị tiên nghiệm đều. Thay đổi hình thức của mật độ tiên nghiệm σ_Y từ đều thành Gamma dẫn đến trung bình hậu nghiệm là 0,058, độ lệch chuẩn hậu nghiệm là 0,052 và khoảng tin cậy 95 % là (0,000 0,150). Đây là thay đổi lớn hơn trong các ví dụ trước và được chỉ ra ở đây vì độ gần của dữ liệu với nền, dữ liệu không hoàn toàn là có thông tin về đại lượng đo. Cỡ của σ_Y (kiểm soát ở mức độ nào đó bằng phân bố tiên nghiệm vì chỉ có năm giá trị đo được để làm cơ sở cho ước lượng) ảnh hưởng đến mức độ thông tin của dữ liệu. Trong trường hợp như vậy, giải pháp thận trọng là sử dụng khoảng tin cậy dài hơn dựa trên phân bố đều. Cách tốt hơn là thu được nhiều giá trị đo được hơn. Hệ quả sẽ là giảm ảnh hưởng của mật độ tiên nghiệm của σ_Y lên các kết quả. (Một thực tế thú vị về khoảng tin cậy Bayes như được nêu ở đây có thể tìm trong Tài liệu tham khảo [23]. Trong đó tác giả cho thấy rằng trong các mô hình như Ví dụ 1, khoảng tin cậy Bayes 95 % dựa trên phân bố tiên nghiệm đều có độ phủ tần suất gần với 95 %, trong khi khoảng dựa trên tiên nghiệm gamma thường có độ phủ tần suất thấp hơn.).

9.2.5 Tóm tắt các ví dụ

Ví dụ 1a minh họa trường hợp giá trị đo được từ hai nguồn độc lập được sử dụng trong một đánh giá độ không đảm bảo. Ví dụ 1b thể hiện cách thức thông tin về nền được sử dụng để thực hiện đánh giá độ không đảm bảo Loại B có thể được đưa vào trong mô hình Bayes như thế nào. Ví dụ 1c minh họa sự dễ dàng mà một ràng buộc có thể được đưa vào mô hình Bayes, như ràng buộc tích cực ở đây với giá trị của đại lượng đo. Ví dụ cũng cho thấy việc lựa chọn phân bố tiên nghiệm thiếu thông tin có thể ảnh hưởng đến kết quả như thế nào.

10 Lập luận tin tưởng để đánh giá độ không đảm bảo

10.1 Phương pháp cơ bản

10.1.1 Đối với hàm đo (1), đánh giá độ không đảm bảo đối với đại lượng đo θ có thể dựa trên phân bố dựa vào sự tin tưởng đối với θ. Các ví dụ dưới đây minh họa cách thức để thu được phân bố dựa vào sự tin tưởng đối với các tham số quan tâm.

10.1.2 Giả định $Y \sim N(\theta,1)$, trong đó θ là đại lượng đo, quá trình đo có phương sai đã biết bằng 1 và Y là biến ngẫu nhiên biểu thị các giá trị có thể được quan trắc. Ta có thể diễn tả quan hệ giữa giá trị đo được và quá trình sai số thực nghiệm ngẫu nhiên cần quan tâm bằng công thức

$$Y = \theta + E, \tag{19}$$

trong đó E là sai số ngẫu nhiên với phân bố $N(0,1)$. Mỗi giá trị đo được được gắn với một sai số thực nghiệm ngẫu nhiên cụ thể. Giả định thu được một giá trị đo được là 10. Sai số đo kèm theo được ký hiệu là e. Như vậy

$$10 = \theta + e.$$

Do đó, $\theta = 10 - e$. Nếu đã biết giá trị của e thì đại lượng đo sẽ được biết chính xác, nhưng giá trị của e là chưa biết. Tuy nhiên, thực tế là biết được phân bố từ đó e được tạo ra sẽ giúp xác định tập hợp giá trị của θ được coi là hợp lý. Ví dụ, giá trị $\theta = 2$ hợp lý như thế nào với đại lượng đo? Để điều này là đúng, cần có $e = 8$. Giá trị 8 là khó có thể có được từ phân bố $N(0,1)$. Do đó, kết luận là giá trị $\theta = 2$ là không thể có. Khả năng θ nằm giữa 10 và 12 là như thế nào? Đối với θ giữa 10 và 12, e cần bằng từ 0 đến 2 và xác suất cho điều này là $\Phi(2) - \Phi(0)$, trong đó $\Phi(z)$ là giá trị của phân bố Gauss chuẩn tích lũy tại z. Do đó, các xác suất gắn với E có thể được chuyển thành xác suất cho θ. Hiểu biết về θ, dựa trên giá trị đo được là 10, có thể được mô tả bằng phân bố của biến ngẫu nhiên $\tilde{\theta}$ có phân bố được cho bằng phân bố của $10 - E$. Tức là, $\hat{\theta} \sim N(10,1)$ hoặc phân bố dựa vào sự tin tưởng cho θ (tức là, phân bố cho $\tilde{\theta}$) là $N(0,1)$. Biến ngẫu nhiên $\tilde{\theta}$ còn được gọi là đại lượng dựa vào sự tin tưởng (FQ) cho θ. FQ này được liên hệ với cái gọi là đại lượng then chốt tổng quát hóa [24], [25] hoặc đại lượng then chốt tổng quát hóa dựa vào sự tin tưởng [26], [27] trong tài liệu.

10.1.3 Trong ví dụ ở trên, giả định hai phép đo được thực hiện. Lấy Y_1 và Y_2 là biến ngẫu nhiên ký hiệu cho các giá trị có thể có thu được cho hai phép đo, được biểu thị là

$$Y_1 = \theta + E_1,$$

$$Y_2 = \theta + E_2. \tag{20}$$

Giả định các giá trị đo được thực tế là 10 và 8. Khi đó, các công thức dưới đây liên hệ giá trị đo được, đại lượng đo và giá trị thể hiện của sai số thực nghiệm, là e_1 và e_2

$$10 = \theta + e_1,$$

$$8 = \theta + e_2.$$

Giá trị hợp lý cho θ được liên hệ với giá trị hợp lý của (e_1, e_2). Điều làm ví dụ này khác biệt với ví dụ trước là ở đây đã biết rằng $e_1 - e_2 = 2$. Vì vậy, tập hợp các giá trị có thể có cho (e_1, e_2) lúc này bị giới hạn bởi yêu cầu này. Biết rằng (e_1, e_2) là từ phân bố Gauss chuẩn hai chiều nhưng có ràng buộc là nằm trên đường $e_1 - e_2 = 2$. Vì vậy, xác suất có thể gắn cho θ là xác suất với $10 - e_1$ hoặc $8 - e_2$ biết rằng (e_1, e_2) là thể hiện từ phân bố Gauss chuẩn hai chiều với điều kiện bổ sung là $e_1 - e_2 = 2$. Do đó FQ $\widetilde{\theta}$ được xác định là có phân bố có điều kiện với phân bố điều kiện của $10 - E_1$ với $E_1 - E_2 = 2$. Đây chính là phân bố giống với phân bố có điều kiện của $8 - E_2$ với $E_1 - E_2 = 2$. Tính toán đơn giản cho thấy rằng phân bố của $\widetilde{\theta}$ là $N(\bar{y}, 1/2)$ trong đó $\bar{y} = (y_1 + y_2)/2 = (10 + 8)/2 = 9$.

10.1.4 Tổng quát hơn, đối với n phép đo độc lập thực hiện từ $N(\theta, \sigma^2)$,

$$Y_1 = \theta + \sigma E_1,$$

$$Y_2 = \theta + \sigma E_2,$$

$$\ldots,$$

$$Y_n = \theta + \sigma E_n, \tag{21}$$

trong đó $E_1, \ldots, E_n$ là độc lập, các biến ngẫu nhiên Gauss chuẩn. Phân bố dựa vào sự tin tưởng đồng thời (θ,σ) có thể thu được như sau đây. Sử dụng hai phương trình cấu trúc đầu tiên (hoặc bất kỳ) trong số n công thức ở trên để giải cho θ hoặc σ, ký hiệu bằng $\widetilde{\theta}$ và $\widetilde{\sigma}$, là hàm của y_1, y_2, E_1 và E_2. Phân bố dựa vào sự tin tưởng đồng thời (θ,σ) là phân bố đồng thời cho ($\widetilde{\theta}$, $\widetilde{\sigma}$) điều kiện là E_i áp đặt bởi phần con lại của công thức n - 2. Cụ thể, phân bố dựa vào sự tin tưởng cho θ là

$$\widetilde{\theta} = \bar{y} - \frac{s}{\sqrt{n}} T_{n-1}, \tag{22}$$

gọi là phân bố t dịch chuyển và tỷ lệ với $n - 1$ bậc tự do. Do vậy, $\bar{y}$ và s là giá trị thể hiện của trung bình mẫu $\bar{X}$ và độ lệch chuẩn mẫu S đối với n giá trị đo được, còn T_{n-1} là biến ngẫu nhiên có phân bố t với $n - 1$ bậc tự do.

10.1.5 Có một phương pháp thay thế, đơn giản hơn là chỉ phác thảo để rút ra phân bố dựa vào sự tin tưởng cho θ ở (22). Phương pháp này sẽ được minh họa trong các ví dụ sau.

10.1.6 Lập luận ở trên có thể được tổng quát hóa và phân bố dựa vào sự tin tưởng có thể được xây dựng cho các tham số mô hình trong các vấn đề phạm vi rộng hơn. Điểm bắt đầu cho quá trình này được gọi là phương trình cấu trúc [28]. Ký hiệu phương trình cấu trúc này là $Y = G(\beta,E)$. Đối với một phép đo, công thức (19) tạo thành phương trình cấu trúc. Đối với n phép đo, công thức (21) tạo thành các phương trình cấu trúc. Phương trình cấu trúc liên hệ các phép đo Y với tham số mô hình β và sai số quá trình E có phân bố đã biết. Ví dụ, đối với một phép đo, phân bố của E được biết đầy đủ. Đối với giá trị cố định bất kỳ của β, phân bố của E và phương trình cấu trúc $G(\cdot)$ xác định phân bố cho dữ liệu Y. Sau khi quan trắc dữ liệu Y, vai trò của dữ liệu và tham số có thể hoán đổi nhau. Cụ thể, giá trị của Y là cố định và phân bố của E, phương trình cấu trúc $G(\cdot)$ được sử dụng để suy ra phân bố cho β. Đó là những gì cấu thành nên lập luận dựa vào sự tin tưởng.

10.2 Ví dụ 1

10.2.1 Ví dụ 1a

10.2.1.1 Để minh họa, xét Ví dụ 1a mô tả ở Điều 7 trong đó đại lượng vật lý θ được ước lượng từ các giá trị đo được theo mô hình

$$Y_i = \theta + \beta + \varepsilon_i,\ i = 1, \ldots, n, \tag{23}$$

trong đó ε_i là các sai số đo độc lập với $\varepsilon_i \sim N(0,\sigma_Y^2)$. Còn β đại diện cho nền và có thể được ước lượng từ các phép đo theo mô hình

$$B_i = \beta + \delta_i,\quad i = 1, \ldots, n_b, \tag{24}$$

trong đó δ_i là các sai số đo độc lập với $\delta_i \sim N(0,\sigma_B^2)$. Giả định rằng ε_i và δ_i là độc lập. Từ (23) và (24) kéo theo $\bar{Y} - \bar{B}$ có phân bố Gauss với trung bình θ và phương sai $\sigma_Y^2/n + \sigma_B^2/n_b$, trong đó $\bar{Y}$ và $\bar{B}$ là trung bình của Y_i và B_i, tương ứng và có thể biểu thị bằng

$$\bar{Y} - \bar{B} = \theta + \sqrt{\frac{\sigma_Y^2}{n} + \frac{\sigma_B^2}{n_b}}\,Z, \tag{25}$$

trong đó Z là biến ngẫu nhiên Gauss chuẩn. Đây là phương trình cấu trúc $\bar{Y} - \bar{B}$ cho. Như vậy

$$W_y = \frac{(n-1)S_y^2}{\sigma_Y^2} \sim \chi^2(n-1)$$

và

$$W_b = \frac{(n_b-1)S_b^2}{\sigma_B^2} \sim \chi^2(n_b-1)$$

trong đó $\chi^2(\nu)$ ký hiệu cho phân bố Khi-bình phương với ν bậc tự do, S_y^2 và S_b^2 là phương sai mẫu của Y_i và B_i, tương ứng. Do đó

$$S_y^2 = \frac{\sigma_Y^2 W_y}{n-1} \tag{26}$$

là phương trình cấu trúc cho S_y^2 và

$$S_b^2 = \frac{\sigma_B^2 W_b}{n_b-1} \tag{27}$$

là phương trình cấu trúc cho S_b^2. Bằng việc giải ba phương trình cấu trúc ở trên cho θ, σ_r và σ_B, FQ cho θ thu được là

$$\tilde{\theta} = \bar{y} - \bar{b} - \sqrt{\frac{(n-1)s_y^2}{nW_y} + \frac{(n_b-1)s_b^2}{n_bW_b}}Z. \tag{28}$$

10.2.1.2 Khoảng dựa vào sự tin tưởng 1 – α đối với θ được cho bởi $\left(\tilde{\theta}_{\alpha/2} \quad \tilde{\theta}_{1-\alpha/2}\right)$, trong đó $\tilde{\theta}_\alpha$ là phân vị α của phân bố của $\tilde{\theta}$. Các phân vị này có thể được xác định về mặt thống kê trong những tình huống đơn giản. Tuy nhiên, cách thuận tiện nhất là tính xấp xỉ bằng phép Monte Carlo. Cách tiếp cận này liên quan đến việc tạo ra một số lượng lớn các thể hiện từ phân bố của $\tilde{\theta}$ và xác định các phân vị α/2 và 1–α/2 thực nghiệm. Các phân vị này được sử dụng như các ước lượng cho $\tilde{\theta}_{\alpha/2}$ và $\tilde{\theta}_{1-\alpha/2}$. Một thể hiện của $\tilde{\theta}$ có thể được tạo ra như dưới đây.

a) Tạo ra thể hiện của biến ngẫu nhiên Gauss chuẩn Z.

b) Tạo ra các thể hiện W_y và W_b của các biến ngẫu nhiên χ^2 độc lập với n –1 và n_b – 1 bậc tự do, tương ứng.

c) Tính $\tilde{\theta}$ như trong (28).

Đối với ví dụ này, $n = n_b = 5$, $\bar{y} = 3{,}537$, $s_y = 0{,}342$, $\bar{b} = 1{,}228$ và $s_b = 0{,}131$. Chương trình R dùng cho tạo ra 500 000 thể hiện của $\tilde{\theta}$ được cho dưới đây.

```
nrun = 500000
Z = rnorm(nrun)
W1 = rchisq(nrun, 4)
Wb = rchisq(nrun, 4)
theta = 3.537 - 1.228 - sqrt(4*0.342^2/(5*W1)+4*0.131^2/(5*Wb))*Z
```

Trung bình của phân bố mô phỏng là

```
mean(theta)
##2.308893
```

và khoảng dựa vào sự tin tưởng 95 % dựa trên các phân vị 0,025 và 0,975 của phân bố mô phỏng là

```
quantile(theta, c(0.025, 0.975))
## 2.5%  97.5 %
## 1.857814  2.760931
```

Tức là, khoảng dựa vào sự tin tưởng 95 % được cho bởi (1,858; 2,761).

10.2.2 Ví dụ 1b

10.2.2.1 Lúc này không có dữ liệu thống kê liên quan đến nền. Giả định rằng thông tin về β được quy định về phân bố xác suất cho β còn β và ε_i là độc lập. Ngoài ra, giả định phân bố xác suất cho β được biết đầy đủ, tức là, không liên quan đến tham số chưa biết bất kỳ nào.

10.2.2.2 Phương trình cấu trúc cho $\bar{Y}$ được cho bởi

$$\bar{Y} = \theta + \beta + \frac{\sigma}{\sqrt{n}} Z. \tag{29}$$

Cùng với phương trình cấu trúc cho S_y^2 trong (26), ta thu được FQ cho θ là

$$\tilde{\theta} = \bar{y} - \beta - \frac{s_y}{\sqrt{n}} \frac{Z}{\sqrt{W_y/(n-1)}}.$$

Vì $Z/\sqrt{W_y/(n-1)} = T_{n-1}$ là biến ngẫu nhiên có phân bố t với $n-1$ bậc tự do,

$$\tilde{\theta} = \bar{y} - \beta - \frac{s_y}{\sqrt{n}} T_{n-1}. \tag{30}$$

Một thể hiện của $\tilde{\theta}$ có thể được tạo ra như sau đây.

a) Tạo ra thể hiện của T_{n-1} của biến ngẫu nhiên phân bố t student với n –1 bậc tự do.

b) Tạo ra β theo phân bố của nó, độc lập với T_{n-1}.

c) Tính $\tilde{\theta}$ như trong (30).

Đối với ví dụ này, β được giả định có phân bố đều trong khoảng (1,126 1,329). 500 000 thể hiện của $\tilde{\theta}$ được tạo ra bằng

```
beta = runif(nrun, 1.126, 1.329)

theta = 3.537 - beta - 0.342/sqrt(5)*rt(nrun, 4)
```

Trung bình của phân bố mô phỏng là

```
mean(theta)

## 2.309454
```

và khoảng dựa vào sự tin tưởng 95 % dựa trên các phân vị 0,025 và 0,975 của phân bố mô phỏng là

```
quantile(theta, c(0.025, 0.975))

##2.5%    97.5%

##1.871685  2.745590
```

Tức là, khoảng dựa vào sự tin tưởng 95 % được cho bởi (1,872 2,746).

10.2.2.3 Khoảng dựa vào sự tin tưởng ở trên thống nhất với khoảng độ không đảm bảo thu được bằng cách sử dụng phương pháp đề xuất trong GUMS1.

10.2.3 Ví dụ 1c

10.2.3.1 Trường hợp của Ví dụ 1b áp dụng ngoại trừ $\bar{y}$ = 1,196 và s_y = 0,106. 500 000 thể hiện của $\tilde{\theta}$ được tạo ra bằng

```
theta = 1.196 - beta - 0.106/sqrt(5)*rt(nrun, 4)
```

Trung bình của phân bố mô phỏng là

```
mean(theta)

## -0.03158058
```

giá trị này nằm ngoài khoảng tham số cho θ. Số thể hiện nằm ngoài khoảng tham số có thể tìm được bằng

```
length((1:nrun) [theta < 0])

## 319168
```

Cách tiếp cận để xử lý ràng buộc tham số là cắt đuôi phân bố dựa vào sự tin tưởng đến khoảng tham số giới hạn. Đó là, sử dụng max($\tilde{\theta}$,0) để thu được thể hiện của phân bố dựa vào sự tin tưởng cho θ. Khoảng dựa vào sự tin tưởng 95 % được tính là

```
quantile(pmax(theta, 0), c(0.025, 0.975))

## 2.5%    97.5%

## 0.0000000  0.1361553
```

Tức là, khoảng dựa vào sự tin tưởng 95 % được cho bởi (0,000 0,136).

10.2.3.2 Phương pháp mô tả trong 10.2.1.1 và 10.2.2.2 có thể được tổng quát hóa thành mô hình thống kê tùy biến. Mô tả về việc thiết lập FQ được nêu trong Tài liệu tham khảo [29]. Phương pháp đơn giản hơn dùng cho các vấn đề phổ biến hơn trong đó có đủ các thống kê được nêu trong báo cáo kỹ thuật (Tài liệu tham khảo [30]) và được thảo luận thêm trong Tài liệu tham khảo [24] và [25]. Ở đây nhắc lại cho đảm bảo tính đầy đủ. Phương pháp bao gồm các bước sau:

a) Biểu thị từng thống kê đủ như hàm của một hoặc nhiều tham số và biến ngẫu nhiên có phân bố được biết đầy đủ, không có tham số nào chưa biết. Đó là, thu được phương trình cấu trúc cho mỗi thống kê đủ.

b) Trong từng phương trình cấu trúc, biểu thị mỗi tham số như hàm của thống kê đủ và biến ngẫu nhiên có phân bố đã biết hoàn toàn.

c) Thu được FQ cho từng tham số bằng cách thay thế các thống kê đủ bằng các giá trị quan trắc được tương ứng của chúng.

11 Ví dụ 2: hiệu chuẩn can mẫu

11.1 Khái quát

11.1.1 Ví dụ này, được lấy từ Phụ lục H.1 của GUM, liên quan đến việc xác định độ dài của can mẫu bằng cách so sánh với can mẫu giống hệt về danh nghĩa đã được hiệu chuẩn trước đó. Ký hiệu sử dụng trong GUM được tuân thủ chặt chẽ tuy nhiên đã được sửa đổi khi cần cho thống nhất với quy ước ký hiệu trong Điều 4 để phân biệt đại lượng đo với giá trị đo được. Bảng 2 liệt kê các đại lượng vật lý được sử dụng.

11.1.2 Sử dụng ký hiệu nêu trong Bảng 3 và dựa trên

- dòng thứ nhất của Công thức (H.2) trong GUM,
- quan hệ giữa $\alpha = \alpha_s + \delta_\alpha$ và $\theta_s = \bar{\theta} + \Delta - \delta_\theta$ xác định trong H.1.2, và
- suy luận rút ra từ định luật lan truyền độ không đảm bảo trong H.1.3.2 và H.1.3.4 của GUM.

11.1.3 Mô hình đo cho λ sử dụng trong phân tích GUM về ví dụ H.1 có thể được biểu thị bằng

$$\lambda = \frac{\lambda_s\left[1+\alpha_s\left(\bar{\theta}+\Delta-\delta_\theta\right)\right]+\delta_\lambda+\delta_{C_r}+\delta_{C_{nr}}}{1+\left(\alpha_s+\delta_\alpha\right)\left(\bar{\theta}+\Delta\right)}. \tag{31}$$

Công thức (31) là mô hình đo như được mô tả trong H.1.1 và dòng thứ nhất của Công thức (H.2) của GUM, chứ không phải phép xấp xỉ thực hiện đối với dòng thứ hai của Công thức (H.2) và sau đó được sử dụng trong suốt phần còn lại của H.1.

Bảng 3 – Ký hiệu dùng cho phân tích Ví dụ H.1 của GUM trong từng cách tiếp cận trong số ba cách tiếp cận thống kê. Biến ngẫu nhiên tương ứng với δ_λ được ký hiệu là $\overline{D}_\lambda$ và giá trị quan trắc ký hiệu là $\overline{d}_\lambda$

Đại lượng	Ký hiệu
Độ dài của can mẫu chưa biết ở 20 °C	λ
Độ dài của can mẫu chuẩn ở 20 °C	λ_s
Chênh lệch giữa các độ dài can mẫu ở nhiệt độ môi trường phòng thí nghiệm	δ_λ
Hiệu chính chênh lệch độ dài can mẫu để bù cho sai số so sánh ngẫu nhiên	δ_{C_r}
Hiệu chính chênh lệch độ dài can mẫu để bù cho sai số so sánh hệ thống	$\delta_{C_{nr}}$
Hệ số giãn nở nhiệt của can mẫu chuẩn	α_s
Chênh lệch hệ số giãn nở nhiệt của can mẫu chuẩn và can mẫu chưa biết	δ_s
Độ lệch trung bình của nhiệt độ bệ thử so với điều kiện tiêu chuẩn trong quá trình thu thập dữ liệu	$\bar{\theta}$
Biến thiên chu kỳ của nhiệt độ bệ thử so với nhiệt độ trung bình do kiểm soát nhiệt tĩnh	Δ
Chênh lệch nhiệt độ của can mẫu chuẩn và can mẫu chưa biết	δ_θ

11.1.4 Công thức (31) còn được biểu thị về mặt đại lượng vật lý dùng để xác định độ dài của can mẫu chứ không phải tổng hợp trước các ảnh hưởng do chênh lệch độ dài của hai can mẫu và chênh lệch nhiệt độ của bệ thử. Thực hành tốt là biểu thị mô hình đo theo tất cả các đại lượng cần thiết để xác định nó. Thực hành này giúp giảm thiểu sai lỗi có thể có khi xác định mối tương quan giữa các đại lượng vật lý khác nhau, ví dụ như θ và θ_s, α và α_s, như đề cập trong H.1.2, với giá trị cuối cùng có thể dựa trên cùng một dữ liệu.

11.1.5 Bảng 4 tóm tắt phần thông tin còn lại lấy từ phân tích Ví dụ H.1 của GUM cần thiết cho phân tích bằng các cách tiếp cận thống kê khác nhau được thảo luận và so sánh trong phần còn lại của điều này.

11.1.6 Mô tả ví dụ trong GUM chỉ ra rằng chỉ có một đại lượng, δ_λ, có giá trị được ước lượng bằng cách sử dụng phân tích dữ liệu thống kê. Phân bố của trung bình các giá trị đo được, đưa ra ước lượng của δ_λ, được lấy là Gauss (chuẩn) với giá trị kỳ vọng phụ thuộc vào độ dài của can mẫu và các đại lượng vật lý khác mô tả trong Bảng 3 và Bảng 4.

11.1.7 Giá trị và độ không đảm bảo chuẩn gắn với các ước lượng của tất cả các đại lượng khác thu được bằng đánh giá Loại B. Tuy nhiên, do các đại lượng δ_α và δ_θ theo phân bố chữ nhật chứ không phải Gauss, nên không có phương pháp thống kê được thừa nhận rộng rãi nào để giải thích cho bậc tự do trong hai trường hợp này. Kết quả là, bậc tự do đã cho sẽ không được sử dụng cho các đại lượng này.

Bảng 4 – Tổng hợp thông tin từ phân tích Ví dụ H.1 của GUM cần thiết cho tái phân tích

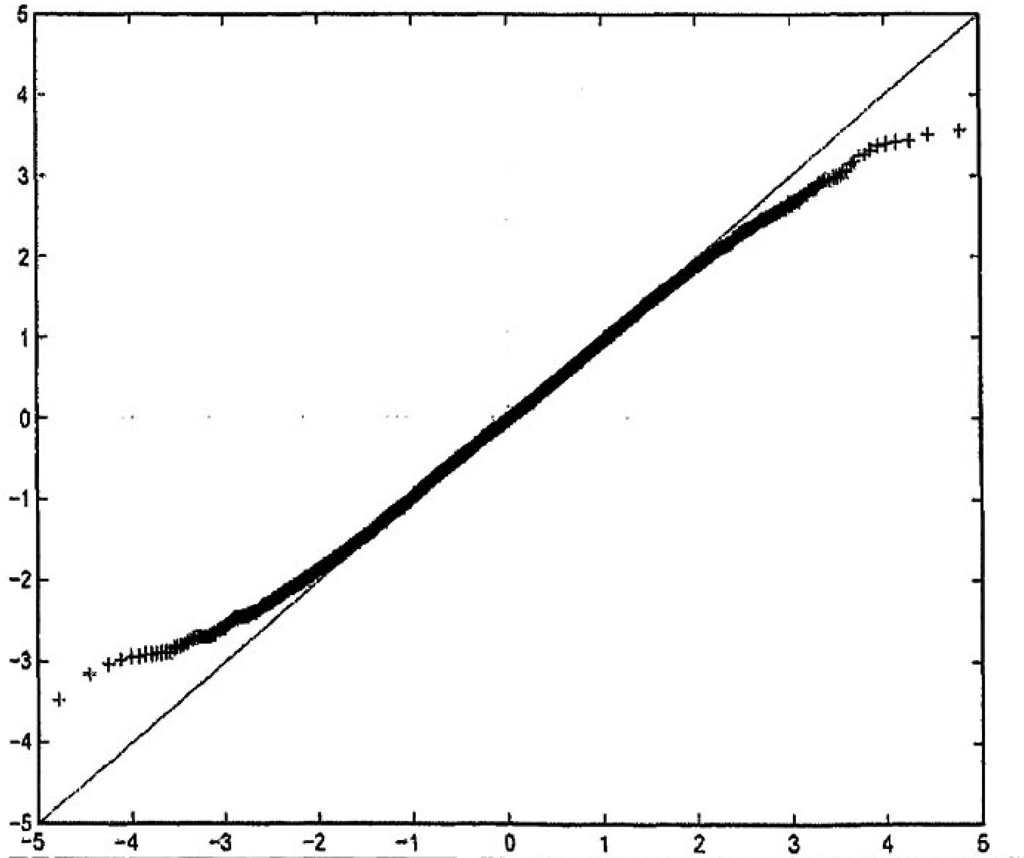
Đại lượng	Giá trị	Độ không đảm bảo chuẩn	Bậc tự do	Loại đánh giá độ không đảm bảo	Mô tả đặc trưng phân bố
λ_s	50 000 623 nm	25 nm	18	B	Gauss
$\bar{d}_\lambda$	215 nm	5,8 nm	24	A	Gauss
δ_{C_r}	0 nm	3,9 nm	5		Gauss
$\delta_{C_{nr}}$	0 nm	6,7 nm	8	B	Gauss
α_s	$11,5\times10^{-6}$ °C^{-1}	$1,2\times10^{-6}$ °C^{-1}		B	Hình chữ nhật
δ_α	0 °C^{-1}	$0,58\times10^{-6}$ °C^{-1}	50	B	Hình chữ nhật
$\bar{\theta}$	-0,1 °C^{-1}	0,2 °C		B	Không quy định
Δ	0 °C	0,35 °C		B	Arcsin
δ_θ	0 °C	0,29 °C	2	B	Hình chữ nhật

11.2 Cách tiếp cận tần suất

11.2.1 Trong ví dụ này, hệ số độ nhạy $c_{\alpha_s} = c_{\theta_s}$ triệt tiêu và số hạng thứ hai được đưa vào Công thức (6) và (7) mặc dù chỉ một trong hai là khác 0 đáng kể (GUM, p71).

11.2.1 Giá trị y = 50 000 838 nm của đại lượng đo, gọi là, độ dài của can mẫu được hiệu chuẩn, và độ không đảm bảo chuẩn kèm theo $u(y)$ = 34 nm được trả lại bằng đánh giá trong GUM sử dụng số hạng cấp hai. Như đã đề cập ở 8.1.12, độ không đảm bảo gắn với các ước lượng này được tính xấp xỉ bằng sai số bậc hai cận biên nếu các tham số λ, θ_s được lấy xấp xỉ theo các phân bố chuẩn, còn λ_s, θ_s và δ_θ được lấy tích phân theo phân bố đều của chúng.

11.2.2 Các kết quả này được khẳng định bằng lan truyền phân bố áp dụng bởi phương pháp Monte Carlo như trong GUMS1, đưa ra một đáp án rất sát. Ngoài ra, phép xấp xỉ bằng phân bố t (10) có vẻ là hợp lý. Hình 1 thể hiện phân vị thực nghiệm vẽ theo phân vị phân bố t khi bậc tự do được ước lượng theo công thức (8). Các kết quả khác từ phương pháp Monte Carlo được báo cáo trong Tài liệu tham khảo [31].



Hình 1 – Phân vị thực nghiệm với phân vị phân bố t trong Ví dụ 2

11.2.3 Thiết lập khoảng bootstrap cho ví dụ này, λ được ước lượng là 50 000 838 nm với độ không đảm bảo chuẩn đồng thời u = 31,7 nm,. Từ công thức (13), khoảng tin cậy bootstrap-t 100(1 – α) % là

$(50\,000\,838 - \hat{t}_{1-\alpha/2} \cdot 31{,}7 \;\; 500\,00838 + \hat{t}_{1-\alpha/2} \cdot 31{,}7)$, trong đó $\hat{t}_\beta$ là phân vị thứ 100β của W^* của (12). Các dòng lệnh R dùng để tạo B = 10 000 thể hiện của W^* như dưới đây.

```
B = 10000
x.star = cbind(
    rnorm(B, mean=50000623, sd=25),
    rnorm(B, mean=215, sd=5.8),
    rnorm(B, mean=0, sd=3.9),
    rnorm(B, mean=0, sd=6.7),
    runif(B, min=0.0000095, max=0.0000135),
    runif(B, min=-0.000001, max=0.000001),
    runif(B, min=-0.45, max=0.25),
    rbeta(B, 0.5,  0.5) -0.5,
    runif(B, min=-0.05, max=0.05))
  u.star = cbind(
    25 * sqrt(rchisq(B, df=18)/18),
    5.8 * sqrt(rchisq(B, df=24)/24,
    3.9 * sqrt(rchisq(B, df=5)/5,
    6.7 * 0.000001 2,
    0.00000058 * sqrt(rchisq(B, df=50)/50,
    0.2,
    0.35,
    0.029 * sqrt(rchisq(B, df=2)/2),
 x.name = c("L.s","D.lambda",Dc.r","Dc.s","A.std","D.alpha",
    "T.bar","T.cv","D.theta")
 f = expression((L.s*(1+A.std.*(T.bar+T.cv-D.theta))+
    D.lambda+Dc.r+Dc.s)/ (1+A.std+D.alpha)*(T.bar+T.cv))
 star = delta(f,x.star,u.star, x.name)
 w.star = (star$y-50000838)/star$uc
```

Hàm `R delta` được xác định như dưới đây.

```
delta = function(meq,x,u,namevec) {
    for(i in 1:ncol(x)) assign(namevec[i], x[,i])
```

```
    c = attr(eval(deriv(meq,namevec)),"gradient")
  list(y=eval(meq),uc=sqrt(apply((c*u)^2,1,sum)))
  )
```

Hàm này chấp nhận biểu thức `R` hợp lý `meq`, hàm đo, có tên tham số được cho là `namevec`, và ma trận giá trị đầu vào `x`, một cột của `x` chứa các giá trị lặp bootstrap cho mỗi đại lượng trong `meq`. Hàm sử dụng hàm `R` `deriv` để đánh giá hàm đo và thu được đạo hàm bậc 1 `c` ("`gradient`" của `meq` tại `x`) đối với tất cả các tham số trong `namevec` đánh giá tại giá trị đầu vào cho bởi `x`. Cuối cùng, hàm trả lại biểu thức đánh giá (`y`) và độ không đảm bảo gắn với tính được sử dụng phép xấp xỉ Taylor bậc nhất thông thường. Trong các lệnh bootstrap ở trên, hàm delta được áp dụng cho hàm đo xác định là `f` trong các lệnh.

Khoảng tin cậy bootstrap-*t* 95 % dựa trên các phân vị 0,025 và 0,975 của phân bố mô phỏng là

```
50000838-quantile(w.star,c(0.975,0.025))*31.70511
## 50000777  50000899
```

Khoảng này, (50000777 50000899) nm, hầu như ngắn hơn 10 % so với khoảng thu được bằng việc áp dụng GUM. Biểu hiện chung này, độ rộng của khoảng bootstrap ngắn hơn của khoảng rút ra từ đánh giá độ không đảm bảo dựa trên phép xấp xỉ Taylor bậc nhất, được thảo luận thêm trong Tài liệu tham khảo [32].

11.3 Cách tiếp cận Bayes

11.3.1 Mô hình quan trắc liên hệ dữ liệu với các tham số có thể được quy định.

11.3.2 GUM có thể được giải thích là nêu rõ rằng giá trị kỳ vọng $E(D_\lambda)$ của phép đo bằng δ_λ, trong đó

$$\delta_\lambda = \lambda(1 + (\delta_\alpha + \alpha_s)(\bar{\theta} + \Delta)) - \lambda_s(1 + (\bar{\theta} + \Delta) - \delta_\theta)\alpha_s). \tag{32}$$

Giá trị kỳ vọng của phép đo là hàm của véctơ tham số $\gamma = (\lambda, \bar{\theta}, \Delta, \delta_\alpha, \alpha_s, \delta_\theta)$ và đại lượng đo λ. GUM đưa ra hai thành phần bổ sung của độ không đảm bảo bao gồm hàm so sánh xác định bằng đánh giá Loại B. Do đó, có độ không đảm bảo về giá trị kỳ vọng của phép đo khác nhau bằng δ_λ. Tương tự với GUM, hai thành phần này có thể cộng lại để có được độ không đảm bảo là 7,8 nm, với 12 bậc tự do sử dụng công thức Welch-Satterthwaite. Mô hình thống kê hai bước dưới đây kết hợp các thông tin sẵn có:

$$\bar{D}_\lambda | \delta_{\lambda_*} \sim N(\delta_{\lambda_*}, \sigma^2_{D_\lambda}/5) \tag{33}$$

$$\delta_{\lambda_*} | \delta_\lambda \sim \delta_\lambda + 7.8 \cdot T_{12}.$$

11.3.3 Cho trong ví dụ là độ không đảm bảo gắn với chênh lệch đo được s_{d_λ} thu được bằng đánh giá Loại A, cung cấp ước lượng của σ_{D_λ}. Từ lý thuyết xác suất cơ bản, đối với mẫu cỡ *n* từ phân bố Gauss có phương sai đã biết σ^2,

$$\frac{(n-1)}{\sigma^2}S^2 \sim \chi^2(n-1)$$

Vì $\chi^2(n-1)$ cũng là mật độ Gamma $(\frac{n-1}{2}, \frac{1}{2})$,

$$S_{d_\lambda}^2 | \sigma_{D_\lambda} \sim \text{Gamma}\left(\frac{(25-1)}{2}, \frac{1}{2\sigma_{D_\lambda}^2}\right).$$

11.3.4 Có 8 tham số trong mô hình thống kê, bao gồm đại lượng đo λ. Để tìm phân bố hậu nghiệm cho λ, trước tiên phân bố tiên nghiệm đồng thời của 8 tham số được quy định. Theo lý thuyết, các biến ngẫu nhiên này có thể được coi là độc lập và vì vậy phân bố đồng thời của chúng là tích các phân bố tiên nghiệm riêng lẻ. Đối với các thành phần của véctơ tham số γ, thông tin sử dụng để thực hiện đánh giá độ không đảm bảo Loại B có thể được hiểu là mật độ tiên nghiệm thông tin như sau:

$$\lambda_S \sim N(50000623, 625), \tag{34}$$

$\delta_\alpha \sim$ Uniform $(-1 \times 10^{-6}, 1 \times 10^{-6})$,

$\bar{\theta} \sim N(-0{,}1, 0{,}1681)$,

$\Delta \sim$ Beta$(0{,}5, 0{,}5) - 0{,}5$,

$\alpha_s \sim$ Uniform $(9{,}5 \times 10^{-6}, 13{,}5 \times 10^{-6})$,

$\delta_\theta \sim$ Uniform $(-0{,}05, 0{,}05)$.

Vì vậy, phân bố tiên nghiệm đồng thời là

$$p(\gamma) = p(\lambda_s)\, p(\delta_\alpha) p(\bar{\theta})\, p(\Delta) p(\alpha_s)\, p(\delta_\theta).$$

Phân bố tiên nghiệm của đại lượng đo λ và σ_{D_λ} là cần thiết để hoàn thiện quy định trước. Trong ví dụ này, không có thông tin thêm về hai tham số này ngoài việc chúng không âm. Như trong Ví dụ 1, các tham số được ấn định là tiên nghiệm tham chiếu [20]. Đối với λ, mật độ tham chiếu xấp xỉ là

$$\lambda \sim \text{Uniform}(0, c) \tag{35}$$

với giá trị lớn cho c. Tương tự, đối với σ_{D_λ},

$$\sigma_{D_\lambda} \sim \text{Uniform}(0, c) \tag{36}$$

hoặc Gamma(c, c). Điều này hoàn thành quy định phân bố tiên nghiệm.

Lưu ý là hai phân bố tiên nghiệm tham chiếu, mà phân tích độ nhạy chỉ ra là có ít tác động đến kết quả, là phân bố duy nhất không sử dụng theo một cách thức nhất định bằng cách tiếp cận tần suất hoặc cách tiếp cận dựa vào sự tin tưởng.

11.3.5 Áp dụng định lý Bayes dẫn đến mật độ hậu nghiệm đồng thời dưới đây của $[\lambda, \gamma, \sigma_{D_\lambda}]$:

$$p\left(\lambda,\gamma,\sigma_{D_\lambda}\middle|\bar{d}_\lambda,s_{d_\lambda}\right)=\frac{p\left(\bar{d}_\lambda\middle|\delta_\lambda\right)p\left(\delta_\lambda\middle|\delta_\lambda\right)p(\gamma)p(\lambda)p\left(\sigma_{D_\lambda}\right)}{\int p\left(\bar{d}_\lambda\middle|\delta_\lambda\right)p\left(\delta_\lambda\middle|\delta_\lambda\right)p(\gamma)p(\lambda)d\gamma\, d\lambda\, d\sigma_{D_\lambda}}.$$

Khi đó, mật độ hậu nghiệm của λ thu được bằng cách lấy tích phân là

$$p\left(\lambda\middle|\bar{d}_\lambda,s_{d_\lambda}\right)=\int p\left(\lambda,\gamma,\sigma_{D_\lambda}\middle|\bar{d}_\lambda,s_{d_\lambda}\right)d\gamma\, d\sigma_{D_\lambda}.$$

Phân bố hậu nghiệm này tổng hợp tất cả thông tin về λ có sẵn thu được sau các phép đo. Các dòng lệnh WinBUGS cho ví dụ này như sau:

```
Example2 {
   n<-25
   df<-(n-1)/2
   lambda~dnorm(0.1,0E-18)
   delta.a~dunif(-0.000001, 0.000001)
   alpha~dunif(0.0000095,0.0000135)
   theta~dnorm(-0.1,5.94)
   ddelt~dbeta(0.5,0.5)
   delta<-ddelt-0.5
   delta.t~dunif(-0.05,0.05)
   lambda.s~dnorm(50000623, 0.0016)
   sigma.D~dunif(0.20)
   tau.D<-1/(sigma.D*sigma.D)
   delta.l<-lambda*(1+(delta.a+alpha)*(theta+delta))
        -lambda.s*(1+((theta+delta)-delta.t)*alpha)
        delta.l.r~dt(delta.l, 0.0164,12)
   msg<-5*tau.D
   dbar~dnorm(delta.l.r,msg)
   pg<-tau.D/2
   ssq<-(n-1)*s.y*s.y
   ssq~ dgamma(df,pg)
   }
```

Đối với $\bar{d}_\lambda$= 215 và s_Y = 13, dòng lệnh WinBUGS này thu được trung bình hậu nghiệm của λ là 50 000 837 nm, với độ lệch chuẩn hậu nghiệm là 34 nm. Khoảng tin cậy 95 % là (50 000 768 nm 50 000 908 nm). Các kết quả này hầu như giống với trong GUM.

11.3.6 Trong lời giải ở đây, mô hình đo đối với λ, tức là công thức (31), không bao giờ được sử dụng, vì vậy tránh được công việc khó khăn và không cần thiết là xác định các phân bố của các tham số khác nhau có liên quan như thế nào. Như trong Ví dụ 1 với hai tham số, cách tiếp cận nêu ở đây dẫn đến phân bố hậu nghiệm đồng thời thích hợp cho cả 8 tham số.

11.3.7 Xét lời giải gần đúng cho ví dụ này dựa trên phép xấp xỉ bằng chuỗi Taylor. Trong cách giải của GUM, công thức (31) được tính xấp xỉ là

$$\delta_\lambda = \lambda - \lambda_s(1 - (\delta_\alpha(\bar{\theta} + \Delta) + \alpha_s\delta_\theta)).$$

Định nghĩa tham số $\eta = \lambda - \delta_\lambda$. Sử dụng phép xấp xỉ chuỗi Taylor, mật độ xác suất của η có thể được tính xấp xỉ bằng Gauss là $\eta \sim N$(50 000 623 nm, 911 nm^2). Để đơn giản, cũng tính xấp xỉ σ_{D_λ} bằng s_{d_λ}. Khi đó, mô hình thống kê trở thành

$$\tilde{D}_\lambda|\delta_\lambda \sim N\left(|\delta_\lambda, \frac{(13)^2}{5}\right),$$

$$|\delta_\lambda|\delta_\lambda \sim N\left(\lambda - \eta (7,8)^2\right),$$

$$\lambda \sim N(0, c),$$

$$\eta \sim N(50\,000\,623, 911{,}47),$$

Đối với mô hình này, mật độ hậu nghiệm của λ có thể thu được theo phân tích [33]. Ta có được

$$\lambda \sim N\left(\bar{d} + 50\,000\,623, \frac{(13)^2}{5} + (7{,}8)^2 + 911{,}47\right)$$

Vì $\bar{d}_\lambda$ = 215 nm, ta có trung bình hậu nghiệm của λ là 50 000 838 nm với độ lệch chuẩn hậu nghiệm 31 nm, giá trị này cũng gần với giá trị trong GUM.

11.4 Cách tiếp cận dựa vào sự tin tưởng

11.4.1 Ví dụ này được dùng để minh họa cách tiếp cận lập luận tin tưởng trong một ứng dụng phức tạp hơn. Hàm đo được cho trong (31). Dựa trên thông tin cung cấp trong GUM, các giả định sau được đưa ra:

a) giá trị ước lượng của λ_s (tức là giá trị cho trong giấy chứng nhận hiệu chuẩn), ký hiệu là l_s, bằng 50 000 623 nm. Độ không đảm bảo chuẩn kèm theo ước lượng là 25 nm với 18 bậc tự do. Với giả định tính chuẩn, đại lượng dựa vào sự tin tưởng (FQ) cho λ_s là

$$\tilde{\lambda}_s = 50000623 - 25T_{18}. \qquad (37)$$

Biểu thức (37) thu được từ phân bố (22) với $\bar{y}$ = 50 000 623 nm, $u(\bar{y})$ = 25 nm và 18 bậc tự do đi kèm với $u(\bar{y})$.

b) Mỗi giá trị lặp đo được được coi là rút ra từ phân bố chuẩn có trung bình δ_λ và độ lệch chuẩn σ_{δ_λ}. Trung bình quan trắc được của năm giá trị đo được là $\bar{d}_\lambda$= 215 nm. Từ thực nghiệm riêng, σ_{δ_λ} được ước lượng là 13 nm với 24 bậc tự do. Do đó, $u(\bar{d}_\lambda)$ = $13/\sqrt{5}$. Vì vậy, FQ cho δ_λ là

$$\tilde{\delta}_\lambda = 215 - 13T_{24}/\sqrt{5} \tag{38}$$

Cũng dựa trên giấy chứng nhận hiệu chuẩn cho thiết bị so sánh, ước lượng của δ_{C_r} bằng 0 với độ không đảm bảo 3,9 nm (5 bậc tự do) và ước lượng của $\delta_{C_{nr}}$ bằng 0 với độ không đảm bảo 6,7 nm (8 bậc tự do). Ngoài ra, sai số so sánh có thể giả định là độc lập với sai số lặp. Do đó,

$$\tilde{\delta}_{C_r} = 3{,}9\ T_5, \tag{39}$$

và

$$\tilde{\delta}_{C_{nr}} = 6{,}7\ T_8, \tag{40}$$

Sự không phụ thuộc lẫn nhau giữa các biến ngẫu nhiên là kết quả của giả định GUM về quá trình đo.

c) Lấy $\bar{\theta}$ là sai lệch giữa nhiệt độ trung bình của bệ thử với giá trị danh nghĩa 20 °C. Ước lượng của $\bar{\theta}$ là -0,1 °C với độ lệch chuẩn 0,2 °C. Vì GUM không đưa thêm thông tin liên quan đến độ lệch chuẩn này nên bậc tự do vô hạn được giả định và $\bar{\theta}$ cũng được rút ra từ phân bố Gauss. Do đó

$$\tilde{\bar{\theta}} = -0{,}1 - 0{,}2Z, \tag{41}$$

trong đó Z là mẫu lấy từ biến ngẫu nhiên Gauss chuẩn, độc lập với tất cả các biến ngẫu nhiên khác.

d) FQ cho Δ có hàm mật độ xác suất cho bởi

$$g(x) = \frac{2}{\pi\sqrt{1-4x^2}}, \quad -0{,}5\ °\text{C} < x < 0{,}5\ °\text{C}. \tag{42}$$

Để lấy ngẫu nhiên từ phân bố arcsin (42), quan trắc thấy nếu U_1 là biến ngẫu nhiên Uniform (0,1), $-\cos(\pi U_1)/2$ có phân bố arcsin như yêu cầu. Vì vậy, FQ cho Δ có thể được lấy là

$$\tilde{\Delta} = -\cos(\pi U_1)/2. \tag{43}$$

e) FQ cho δ_α là

$$\tilde{\delta}_\alpha = U_2, \tag{44}$$

trong đó U_2 là biến ngẫu nhiên phân bố đều trong khoảng $\pm 1 \times 10^{-6}$ °C^{-1}.

f) FQ cho cho δ_θ là

$$\tilde{\delta}_\theta = U_3, \tag{45}$$

trong đó U_3 là biến ngẫu nhiên phân bố đều trong khoảng ±0,05 °C

g) FQ cho cho α_s là

$$\tilde{\alpha}_s = 11{,}5 \times 10^{-6} + U_4, \tag{46}$$

trong đó U_4 là biến ngẫu nhiên phân bố đều trong khoảng ±2 x 10^{-6} °C^{-1}.

11.4.2 Thay thế các đại lượng dựa vào sự tin tưởng trong các biểu thức từ (37) đến (46) vào biểu thức (31), thu được đại lượng dựa vào sự tin tưởng cho λ. Tính xấp xỉ phân bố cho $\tilde{\lambda}$ được cung cấp bằng cách sử dụng 500 000 phép thử Monte Carlo. Trung bình và độ lệch chuẩn của phân bố xấp xỉ này tương ứng là 50 000 838 nm và 35 nm. Khoảng dựa vào sự tin tưởng 95 % cho λ được cho bởi các phân vị 0,025 và 0,975 của phân bố này, tức là (50 000 768 nm 50 000 907nm). Các dòng lệnh R để tạo ra 500 000 thể hiện của $\tilde{\lambda}$ như dưới đây.

```
nrun = 500000
lambda.s = 50000623 - 25 * rt(nrun, 18)
delta.lambda = 215 - 13/sqrt(5) * rt(nrun, 24)
delta.cr = 3.9*rt(nrun, 5)
delta.cnr = 6.7*rt(nrun, 8)
theta = rnorm(nrun, -0.1, 0.2)
delta = (-cos(pi*runif(nrun))/2)
delta.alpha=runif(nrun,-10^(-6), 10^(-6))
delta.theta=runif(nrun, -0.05, 0.05)
alpha.s=runif(nrun, (11.5-2)*10^(-6), (11.5+2)*10^(-6))
lambda=(lambda.s*(1 + alpha.s*(theta + delta - delta.theta))+delta.lambda
        +delta.cr + delta.cnr)/(1 + (alpha.s + delta.aplha)*(theta+delta))
```

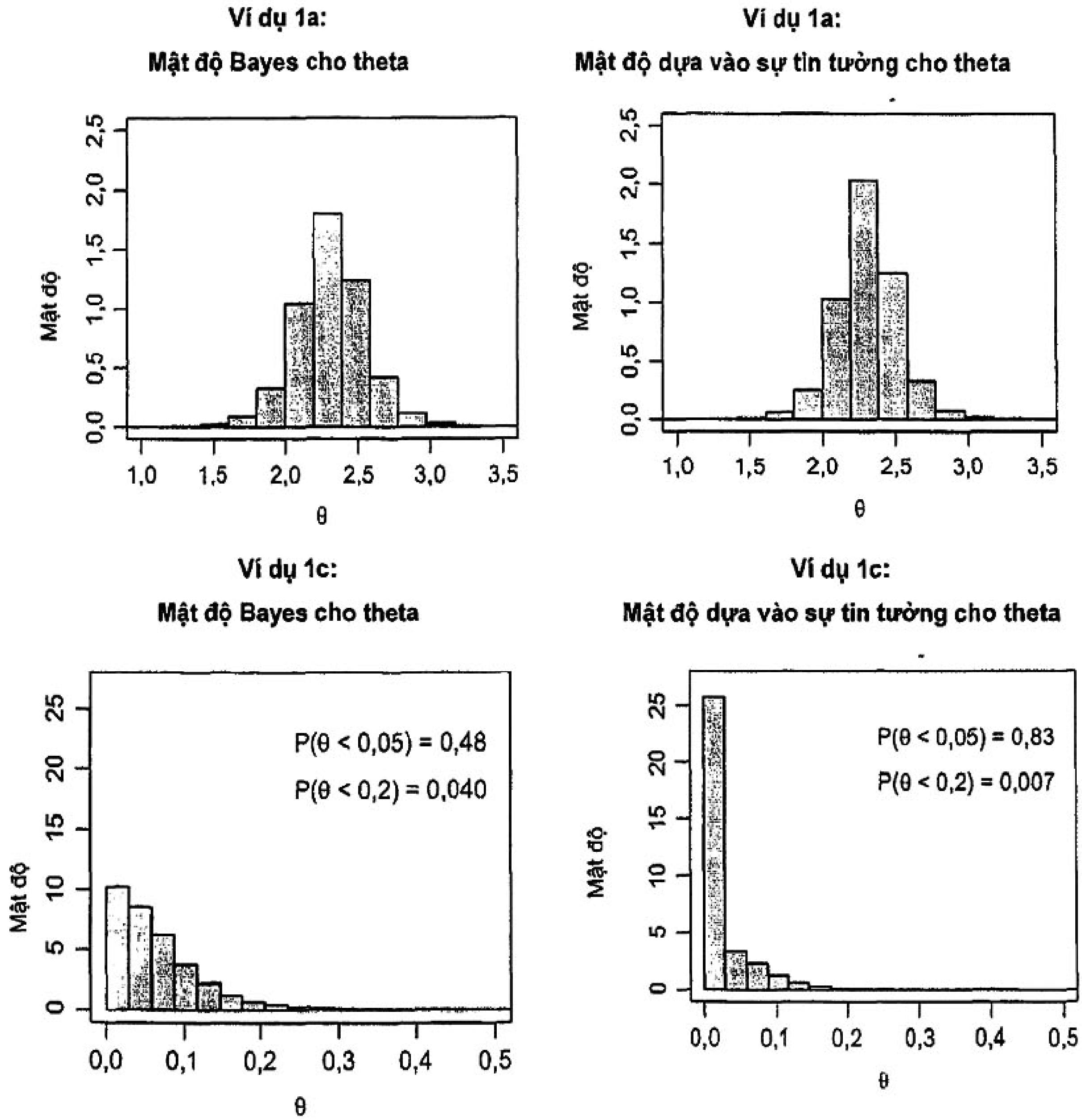
12 Thảo luận

12.1 So sánh các đánh giá độ không đảm bảo sử dụng ba cách tiếp cận thống kê

12.1.1 Bảng 4 tổng hợp các kết quả trong Ví dụ 1. Lời giải bootstrap tần suất, Bayes và dựa vào sự tin tưởng cho Ví dụ 1 a và Ví dụ 1b cũng rất giống như vậy. Lời giải bootstrap và GUM tạo ra các khoảng ngắn hơn một chút trong cả Ví dụ 1a và Ví dụ 1b. Có khác biệt đáng kể hơn trong cách giải ở Ví dụ 1c. Ở đây lời giải Bayes dựa trên mật độ tiên nghiệm đều tạo ra khoảng dài hơn đáng kể so với hầu hết các phương pháp khác; chỉ có khoảng Eisenhart bảo toàn là rộng hơn.

Bảng 4 – Khoảng độ không đảm bảo mở rộng cho ba cách tiếp cận thống kê ở Ví dụ 1

	GUM	Eisenhart	Bootstrap	Bayes	Tin tưởng
Ví dụ 1a	(1,89 2,73)	(1,89 2,73)	(1,83 2,66)	(1,81 2,82)	(1,86 2,76)
Ví dụ 1b	(1,90 2,72)	(1,78 2,84)	(1,86 2,64)	(1,83 2,79)	(1,87 2,75)
Ví dụ 1c	(0,00 0,12)	(0,00 0,20)	(0,00 0,11)	(0,00 0,19)	(0,00 0,14)



Hình 2 – So sánh các mật độ xấp xỉ Bayes và dựa vào sự tin tưởng đối với Ví dụ 1a và 1c

12.1.2 Vì cách tiếp cận Bayes và cách tiếp cận dựa vào sự tin tưởng tạo ra các phân bố xác suất cho đại lượng đo, θ, kết quả của chúng ở Ví dụ 1a và Ví dụ 1c được so sánh thêm trong Hình 2, bên cạnh so sánh khoảng độ không đảm bảo mở rộng ở Bảng 4. Kết quả ở Ví dụ 1b không được thể hiện vì không thể phân biệt bằng mắt thường với các kết quả ở Ví dụ 1a. Từ biểu đồ ở Hình 2, thấy rõ là phân bố xác suất hậu nghiệm Bayes cho θ và phân bố dựa vào sự tin tưởng cho θ là giống nhau khi tín hiệu vượt quá mức nền đáng kể. Tuy nhiên, khi tín hiệu gần với nền, hai phân bố có đặc trưng rất khác biệt do các phương pháp khác nhau trong kết hợp các ràng buộc vật lý vốn có của vấn đề.

12.1.3 Trong bối cảnh tần suất, đại lượng đo θ và đại lượng đầu vào $\mu_1,\ldots,\mu_p$ trong mô hình đo (1) đều được giả định là hằng số cố định chưa biết. Cách tiếp cận này có vẻ hoàn toàn hợp lý nếu đại lượng đo đại diện cho hằng số vật lý và các nghiên cứu trước đó không đưa ra phân bố tiên nghiệm (tham khảo) hoặc phương trình cấu trúc thích hợp cho nó. Phương pháp tần suất được ưa chuộng bởi các nhà thống kê là những người không tin rằng tất cả các tham số phải được mô hình như biến ngẫu nhiên, mặc dù phương pháp này thường xử lý độ không đảm bảo thu được bằng phương pháp đánh giá Loại B bằng cách ấn định cho chúng một phân bố xác suất và lấy tích phân theo phân bố này. Về mặt này, tương tự như cách tiếp cận Bayes, trong đó tất cả các tham số được đặc trưng bởi phân bố xác suất nhưng cần ít giả định về phân bố hơn.

12.1.4 Bootstrap là phương pháp thống kê được thiết lập tốt có thể thay thế các khoảng tin cây gần đúng phức tạp và thường không chính xác bằng các mô phỏng máy tính. Có nhiều chương trình bootstrap khác nhau được phát triển để thiết lập khoảng tin cậy trong các điều kiện khác nhau. Khoảng bootstrap-t tham số giới thiệu trong tiêu chuẩn này là lựa chọn tự nhiên như một cải tiến cho khoảng student-t của GUM. Ưu điểm của bootstrap là tính đơn giản – rất dễ dàng áp dụng bootstrap để rút ra khoảng tin cây như chứng minh trong các ví dụ.

12.1.5 Các ví dụ cho thấy rằng đánh giá độ không đảm bảo Bayes sử dụng mô hình thống kê là đơn giản về mặt khái niệm và có thể áp dụng cho các vấn đề đo phức tạp mà không làm thay đổi gì tới phương pháp cơ bản. Những ảnh hưởng hệ thống không thể ước lượng từ các giá trị đo được (tức là, không có hàm của các quan trắc có giá trị kỳ vọng bằng với ảnh hưởng hệ thống) và đối với chúng, thông tin được sử dụng để thực hiện đánh giá độ không đảm bảo loại B có thể dễ dàng đưa vào mô hình Bayes. Việc tính toán phân bố hậu nghiệm có thể được thực hiện bằng cách sử dụng phương pháp MCMC, thường sử dụng các phần mềm hiện có. Như đã thấy, không cần có các đối số tiệm cận để chứng minh công bố xác suất vì các mẫu nhỏ và lớn có chung một lý giải xác suất.

12.1.6 Có một số nhược điểm của phương pháp Bayes được mô tả ở đây. Quan trọng nhất, để sử dụng chúng các phân bố tiên nghiệm được quy định cho tất cả các tham số trong mô hình đo, bao gồm cả đại lượng đo. Mặc dù trong đo lường phân bố tiên nghiệm tham chiếu thường có sẵn dưới dạng đánh giá độ không đảm bảo Loại B, thường là trường hợp một hoặc hai tham số sẽ cần được ấn định phân bố tiên nghiệm không rõ ràng (phi thông tin) vì thiếu hiểu biết trước. Phân bố như vậy không phải là duy nhất và, như chứng minh trong Ví dụ 1c, chúng có thể ảnh hưởng đến kết quả. Do vậy, nên thực hiện phân tích độ nhạy để đánh giá mức độ lớn của các ảnh hưởng này. Ảnh hưởng lớn nảy sinh từ quy định về phân bố tiên nghiệm phi thông tin đòi hỏi nghiên cứu thêm về hệ thống đo. Nói chung, sự có mặt của những ảnh hưởng này có nghĩa là không có sẵn thông tin đầy đủ về đại lượng đo trong dữ liệu và do đó phân bố tiên nghiệm có ảnh hưởng đáng kể đến kết quả. Trong một số trường hợp, vấn đề này có thể được giải quyết bằng cách tăng số lượng giá trị đo lặp, hoặc thay đổi cách thức thu thập dữ liệu, ví dụ bằng cách cải thiện giải pháp. Trong các tình huống khác, có thể là mô hình toán học được sử dụng có quá nhiều tham số mà không biết thông tin thực trước đó về chúng, do đó mô hình cần được đơn giản hóa.

12.1.7 Khi có thông tin tiên nghiệm quan trọng về đại lượng đo, có thể đưa vào một cách đơn giản và cập nhật hiệu quả thông qua định lý Bayes. Ngoài ra, độ nhạy với dạng tiên nghiệm, không chỉ đối với đại lượng đo mà còn với độ lệch chuẩn của hàm hợp lý, có thể là một chỉ thị tốt là có vấn đề với hệ thống đo. Khi đó, điều này có thể được nghiên cứu và hiệu chính.

12.1.8 Lập luận tin tưởng cung cấp khuôn khổ cho việc kết hợp phân bố với tham số quan tâm. Các kết quả nghiên cứu gần đây [26] cho thấy lập luận tin tưởng là phương pháp thống kê hợp lệ với các đặc trưng hoạt động tốt nói chung. Ví dụ được sử dụng đã chứng minh rằng cách tiếp cận tin cậy có thể kết hợp dễ dàng và tự nhiên thông tin về độ không đảm bảo vào mô hình đo, và thu được ước lượng của đại lượng đo, độ lệch chuẩn kèm theo của nó bằng cách lan truyền các phân bố thống kê thành phần.

Trong cách tiếp cận dựa vào sự tin tưởng không có nhu cầu lan truyền độ không đảm bảo dựa trên khai triển chuỗi Taylor hoặc phép xấp xỉ Welch-Satterthwaite.

12.1.9 Có một vấn đề về tính phi đơn nhất trong phân bố dựa vào sự tin tưởng do việc lựa chọn dạng cụ thể của phương trình cấu trúc. Tuy nhiên, điều quan trọng cần lưu ý là trong nhiều ứng dụng thực tế, quá trình vật lý nhờ đó dữ liệu được tạo ra là đã biết. Trong trường hợp này, phương trình cấu trúc cần được chọn để phản ánh quá trình này, từ đó loại trừ được vấn đề phi-đơn nhất. Trong đo lường, khi một đại lượng đo chưa biết được đo bằng một quá trình đã biết nào đó, ta biết rằng sai số ngẫu nhiên ảnh hưởng đến phép đo theo một cách thức cụ thể. Các giá trị kết quả đo được được biểu thị trong mối quan hệ với mô hình đo kết hợp các đại lượng đo được và sai số ở dạng các đại lượng ảnh hưởng. Mô hình này có thể được lấy làm phương trình cấu trúc.

12.2 Quan hệ giữa các phương pháp đề xuất trong GUM phần bổ sung 1 (GUMS1) và ba cách tiếp cận thống kê

12.2.1 GUMS1[3] tạo ra các mẫu lấy ngẫu nhiên từ phân bố xác suất cho đại lượng đầu ra Y trong mô hình đo "mô tả hiểu biết về đại lượng đó, dựa trên kiến thức về đại lượng đầu vào, như mô tả bởi các PDF ấn định cho chúng" (trang vii của Tài liệu tham khảo [3]). Trong cùng trang, GUMS1 nêu rõ "PDF đối với đại lượng không được hiểu là mật độ tần suất". Cuối cùng, trong trang 7, GUMS1 định nghĩa đại lượng đầu ra Y là đại lượng đo. Do đó, kết quả của phân tích GUMS1, ví dụ như trung bình và độ lệch chuẩn của lấy mẫu Monte Carlo, là các ước lượng đặc trưng của mật độ xác suất cho đại lượng đo đó. Theo đó, so sánh trực tiếp giữa kết quả từ GUMS1 và từ phương pháp dựa vào sự tin tưởng cũng như phương pháp Bayes truyền thống, nếu có thể. Khoảng độ không đảm bảo GUMS1 có thể được nghiên cứu về thuộc tính phủ tần suất nhưng không nên giải thích là khoảng tin cậy tần suất thông thường.

12.2.2 Như chỉ ra trong 9.1.1 và 9.1.2, các phương pháp Bayes truyền thống dựa trên mô hình thống kê bao gồm hiểu biết trước đó về đại lượng đo. Công bố này không thực sự của GUMS1 vì phương pháp này dựa trên mô hình đo, trong đó đại lượng đo là đại lượng đầu ra. (Điều này cũng được nêu ở trang 2 và trang 8 của Tài liệu tham khảo [3]). Do đó, mọi so sánh trực tiếp các kết quả từ phương pháp Bayes truyền thống với phương pháp của GUMS1 đều giới hạn ở trường hợp không có hiểu biết trước đó về đại lượng đo.

12.2.3 Tài liệu tham khảo [34] trình bày so sánh đối với một vấn đề đo có thể áp dụng được cụ thể nhưng rộng rãi. Trong tài liệu này, đại lượng đo μ là hàm của α và β, tức là, mô hình đo là $\mu = f(\alpha,\beta)$. Tham số α có thể được ước lượng từ dữ liệu vì nó là trung bình của biến ngẫu nhiên Gauss X, biến có một tập hợp các dữ liệu quan trắc. Không có sẵn dữ liệu để ước lượng tham số β nhưng phân bố tin cậy đã cho. Phân tích GUMS1 (xem 6.4.9.2 của Tài liệu tham khảo [3]) ấn định phân bố student-t lấy tỷ lệ và dịch chuyển cho α rồi sau đó lan truyền các phân bố cho α và β bằng cách sử dụng hàm f. Trong tài liệu tham khảo [34] cho thấy rằng phân tích này tương đương với tính toán Bayes của mật độ xác suất cho hàm $f(\alpha,\beta)$, trong đó hai tham số được lấy là độc lập, hàm hợp lý cho X là Gauss với trung bình α, phân bố tiên nghiệm đều không phù hợp được sử dụng cho α, và mật độ cho β được cho bằng phân bố tin cậy. Chú ý là mật độ tiên nghiệm cho μ không được sử dụng ở đây.

12.2.4 Lúc này giả định rằng tồn tại hàm g sao cho $\alpha = g(\mu,\beta)$. Phân tích Bayes truyền thống sử dụng hàm hợp lý Gauss cho X với trung bình $g(\mu,\beta)$ và phân bố tiên nghiệm cho μ và β. Trường hợp không có thêm thông tin về đại lượng đo, phân bố đều không phù hợp có thể được sử dụng cho μ nhưng cũng có những lựa chọn khác. Phân bố tin cậy của β là lựa chọn tự nhiên cho phân bố tiên nghiệm. Thông thường, μ và β có thể được lấy là các biến ngẫu nhiên độc lập. Lưu ý là đối với mô hình này, mật độ tiên nghiệm cho α không được sử dụng.

12.2.5 GUMS1 và phân tích Bayes truyền thống sử dụng các tham số hóa khác nhau của cùng một mô hình thống kê. Sự không biết về đại lượng đo được biểu thị khác nhau trong mỗi phép tham số hóa. Mô hình sử dụng bởi GUMS1 không thực hiện trực tiếp việc này bằng mật độ cho μ mà thay vào đó sử dụng phân bố tiên nghiệm phi thông tin, không đúng cho α. Như chỉ ra trong Tài liệu tham khảo [34], đây là một giả định khác biệt. Phân tích Bayes truyền thống thường sử dụng phân bố tiên nghiệm phi thông tin cho bản thân đại lượng đo μ. Tài liệu tham khảo [34] cho thấy rằng hai phân tích đưa ra các phân bố xác suất giống nhau cho đại lượng đo khi phân bố tiên nghiệm đều không đúng cho μ được sử dụng trong phân tích Bayes, và hàm f là tuyến tính. Đối với hàm phi tuyến tính, phân bố xác suất cho μ rút ra trong hai phép tham số hóa là không giống nhau. Điều quan trọng là lưu ý rằng nếu phân bố tiên nghiệm phi thông tin cho α được chuyển thành phân bố tiên nghiệm cho μ thì khi đó phân tích Bayes truyền thống tương ứng cho cùng kết quả như GUMS1 đối với hàm bất kỳ.

12.2.6 Như đã nêu ở trên, dựa vào mô hình đo, GUMS1 thu được PDF cho đại lượng đo bằng cách lan truyền PDF cho đại lượng đầu vào. PDF thu được mô tả hiểu biết về đại lượng đo với điều kiện dữ liệu quan trắc được và các giả định đưa ra khi ấn định PDF chung cho đầu vào. Trong nhiều mô hình tiêu chuẩn với giá trị đo được được coi là lấy từ phân bố chuẩn một chiều, khoảng độ không đảm bảo thu được bằng cách sử dụng GUMS1 và phương pháp dựa vào sự tin tưởng rất giống nhau, nếu không nói là giống hệt. Trở lại mô hình đo trong Ví dụ 1a, đó là

$$\theta = \gamma - \beta$$

với $Y_i \sim N(\gamma, \sigma_Y^2)$, $i = 1,...,5$ và $B_j \sim N(\beta, \sigma_B^2)$, $j = 1,...,5$. Dựa vào hướng dẫn trong GUMS1, phân bố t tỷ lệ và dịch chuyển được ấn định là PDF cho γ và PDF cho β. Đặc biệt, PDF cho γ có cùng phân bố với biến ngẫu nhiên

$$\bar{y} - \frac{s_y}{\sqrt{5}} T_4^{(1)},$$

trong đó $T_4^{(1)}$ là biến ngẫu nhiên Student-t với 4 bậc tự do, và PDF cho β có cùng phân bố với biến ngẫu nhiên

$$\bar{b} - \frac{s_b}{\sqrt{5}} T_4^{(2)},$$

trong đó $T_4^{(2)}$ là biến ngẫu nhiên Student-t với 4 bậc tự do độc lập với $T_4^{(1)}$. Kết quả là, PDF cho đại lượng đo θ có thể thu được từ phân bố của

$$\bar{y} - \bar{b} - \frac{s_y}{\sqrt{5}} T_4^{(1)} - \frac{s_b}{\sqrt{5}} T_4^{(2)}.$$

Dòng lệnh R dùng để tạo ra 500 000 thể hiện của phân bố ở trên được cho dưới đây.

```
nrun = 500000
T1=rt(nrun, 4)
T2=rt(nrun, 4)
theta=3.537 - 1.228 - 0.342/sqrt(5)*T1 + 0.131/sqrt(5)*T2
```

Khoảng độ không đảm bảo 95 % dựa trên các phân vị 0,025 và 0,975 của PDF gần đúng là

```
quantile(theta, c(0.025,  0.975))
##2.5%  97.5%
##1.853703  2.763999
```

về cơ bản giống với khoảng dựa vào sự tin tưởng đối với ví dụ này đã nêu trước đây. Tương tự, cách tiếp cận GUMS1 và cách tiếp cận dựa vào sự tin tưởng tạo ra cùng khoảng độ không đảm bảo cho các vấn đề trong Ví dụ 1b và Ví dụ 1c.

12.2.7 Trong nhiều trường hợp khác, phương pháp GUMS1 và dựa vào sự tin tưởng tạo ra các kết quả khác nhau. Trường hợp "cực trị" có thể thấy trong vấn đề mô tả ở Tài liệu tham khảo [35]. Trong ví dụ trình bày ở Tài liệu tham khảo [35], đại lượng đo là độ lớn của đại lượng giá trị phức.

$$\Gamma = \Gamma_1 + i\Gamma_2.$$

Tức là, đại lượng đo là

$$|\Gamma| = \sqrt{\Gamma_1^2 + \Gamma_2^2}.$$

Giả định $X_1 \sim N(\Gamma_1,\sigma^2)$ và $X_2 \sim N(\Gamma_2,\sigma^2)$ với σ đã biết, GUMS1 ấn định $N(x_1,\sigma^2)$ cho PDF đối với Γ_1 và $N(x_2,\sigma^2)$ cho PDF đối với Γ_2. Kết quả là PDF cho $|\Gamma|$ dựa trên GUMS1 là phân bố của biến ngẫu nhiên

$$\sqrt{(x_1 - \sigma Z_1)^2 + (x_2 - \sigma Z_2)^2}, \qquad (43)$$

trong đó Z_1 và Z_2 là biến ngẫu nhiên chuẩn chuẩn hóa độc lập. Tài liệu tham khảo [35] cho thấy rằng khoảng GUMS1 cho $|\Gamma|$ có hiệu năng tần suất không thỏa mãn (xác suất phủ không đủ) khi $|\Gamma|$ nhỏ so với σ. Điều này là vì biến ngẫu nhiên trong biểu thức (43) dương và do đó ranh giới dưới của khoảng độ không đảm bảo cho $|\Gamma|$ cũng sẽ dương, có thể không phủ $|\Gamma|$ khi $|\Gamma|$ tiến đến 0.

12.2.8 Lời giải dựa vào sự tin tưởng đối với bài toán có thể rút ra dựa trên thực tế là $(X_1^2 + X_2^2)/\sigma^2$ được phân bố χ^2 không trung tâm với 2 bậc tự do và tham số không trung tâm $\lambda = |\Gamma|^2/\sigma^2$. Tính chất phân bố này có thể được sử dụng để xây dựng phương trình cấu trúc liên hệ thống kê có thể quan trắc $(X_1^2 + X_2^2)/\sigma^2$ với λ, trong đó chứa tham số $|\Gamma|$ quan tâm. Dựa vào phương trình cấu trúc này, khoảng dựa vào sự tin tưởng cho $|\Gamma|$ có thể được xây dựng. Tài liệu tham khảo [36] chỉ ra rằng khoảng dựa vào sự tin tưởng duy trì độ phủ tần suất danh nghĩa trong mọi trường hợp.

13 Tổng kết

13.1 Trong tiêu chuẩn này thảo luận ba cách tiếp cận để thiết lập khoảng độ không đảm bảo có giải thích rõ ràng về mặt xác suất. Ngược lại, nhiều công việc khác trong lĩnh vực này tập trung vào đánh giá các tính chất thống kê của các quy trình hiện được sử dụng phổ biến trong cộng đồng đo lường. Một trong những mục tiêu trong việc tiếp cận nghiên cứu các phương pháp đánh giá độ không đảm bảo từ ưu điểm này là cố gắng để hiểu rõ hơn về các phương pháp hiện tại và nhấn mạnh những lựa chọn mới cũng có thể hữu ích.

13.2 Như đề cập trong Tài liệu tham khảo [9], khoảng độ không đảm bảo thu được theo các cách tiếp cận khác nhau sẽ thường giống nhau về số. Tuy nhiên, ngay cả trong trường hợp đó thì việc giải thích cũng khác nhau.

13.3 Khoảng độ không đảm bảo tần suất đưa ra công bố xác suất về hiệu năng lâu dài của quy trình cụ thể để thiết lập khoảng độ không đảm bảo trong quá trình sử dụng lặp lại ở các điều kiện giống nhau. Do đó, công bố xác suất không trực tiếp về giá trị của đại lượng đo mà về quan hệ lâu dài giữa quy trình nhờ đó khoảng được thiết lập và đại lượng đo. Khi đã thu được giá trị đo và khoảng độ không đảm bảo tần suất được tính, kết quả sẽ không còn mang tính ngẫu nhiên nữa. Mặc dù không biết được giá trị của đại lượng đo thu được trong khoảng cụ thể bất kỳ nhưng những khoảng này sẽ chứa giá trị của đại lượng với một xác suất xác định. Không giống như khoảng tin cậy truyền thống chỉ dựa trên dữ liệu thống kê, khoảng độ không đảm bảo tần suất thường được thiết lập sao cho mức tin cậy mong muốn có được bằng trung bình sau khi lấy tích phân theo các phân bố xác suất của các đại lượng thu được bằng đánh giá độ không đảm bảo Loại B.

13.4 Mặt khác, khoảng độ không đảm bảo Bayes và dựa vào sự tin tưởng đều dựa trên phân bố xác suất mô tả trực tiếp hiểu biết về giá trị của đại lượng đo. Các phương pháp sử dụng để thu được hai loại khoảng này là khác nhau, nhưng kết quả tương tự nhau về khía cạnh này của giải thích. Các kết quả Bayes thu được bằng cách kết hợp các phân bố xác suất cho từng tham số quy định trước khi phân tích dữ liệu với mô hình xác suất mô tả độ biến động dữ liệu sử dụng định lý Bayes. Các phân bố hậu nghiệm thu được cho từng tham số phản ánh xác suất của các giá trị tham số với thông tin và dữ liệu trước. Kết quả dựa vào sự tin tưởng thu được bằng cách nghịch đảo mô hình xác suất cho dữ liệu được cung cấp các tham số để thu được phân bố cho các giá trị tham số được cung cấp dữ liệu.

13.5 Nếu kết quả bằng số luôn giống nhau thì mỗi cách giải thích có thể áp dụng (ít nhất là gần đúng) cho mọi khoảng độ không đảm bảo. Tuy nhiên, như chứng tỏ trong các ví dụ trong tiêu chuẩn này, các kết quả bằng số có thể khác nhau đáng kể trong một số trường hợp, mặc dù mỗi kết quả có thể được chứng minh về mặt xác suất và chúng có cùng mức ý nghĩa (thường là 95 %). Cũng có thể thấy những khác biệt khác. Ví dụ, nếu một trong các nguồn độ không đảm bảo chi phối trong một ứng dụng cụ thể tương ứng với đại lượng có phân bố đối xứng, khoảng độ không đảm bảo thu được bằng cách sử dụng cách tiếp cận Bayes hoặc cách tiếp cận dựa vào sự tin tưởng phản ánh sự bất đối xứng trong khi khoảng tin cậy xấp xỉ thu được bằng cách sử dụng quy trình trong GUM sẽ tạo ra khoảng độ không đảm bảo đối xứng (và có thể dài hơn mức cần thiết về một phía). Kết quả tần suất dựa trên các nguyên tắc thống kê khác có thể khớp với các kết quả theo Bayes hoặc dựa vào sự tin tưởng trong một số trường hợp, nhưng các phương pháp khác nhau nhìn chung sẽ không thống nhất vì mỗi cách tiếp cận cuối cùng dựa trên một tập hợp các giả định toán học và tiêu chí khác nhau.

13.6 Sự tồn tại các cách tiếp cận khác nhau để đánh giá độ không đảm bảo không phải luôn thống nhất có thể được coi là sự phức tạp. Tuy nhiên, tốt hơn là nhìn nhận nó như một cơ hội. Chỉ bằng cách liên tục làm việc với nhau để đánh giá các đặc điểm của các cách tiếp cận khác nhau thì các phương pháp đánh giá độ không đảm bảo cuối cùng thu được mới đáp ứng tất cả những nhu cầu vô cùng lớn của khoa học và kinh tế: các phương pháp được áp dụng thực tế, tận dụng hiệu quả các nguồn lực, và áp dụng cho nhiều loại phép đo, cả cũ và mới, và có ý nghĩa rõ ràng.

Thư mục tài liệu tham khảo

[1] TCVN 9595-3:2013 (ISO/IEC Guide 98-3:2008), Độ không đảm bảo đo – Phần 3: Hướng dẫn trình bày độ không đảm bảo đo (GUM:1995)

[2] TCVN ISO/IEC 17025:2007[4] (ISO/IEC 17025:2005), Yêu cầu chung về năng lực của phòng thí nghiệm và hiệu chuẩn

[3] TCVN 6165:2009 (ISO/IEC Guide 99:2007), Từ vựng quốc tế về đo lường – Khái niệm, thuật ngữ chung và cơ bản (VIM)

[4] ISO/IEC Guide 98-3:2008/Suppl 1:2008, Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) – Supplement 1: Propagation of distributions using Monte Carlo method (Độ không đảm bảo đo – Phần 3: Hướng dẫn trình bày độ không đảm bảo đo (GUM:1995) – Phần bổ sung 1: Lan truyền phân bố sử dụng phương pháp Monte Carlo)

[5] GLESER, L.J. Assessing uncertainty in measurement. Statistical Science, 13:277-290, 1998 (Đánh giá độ không đảm bảo đo. Khoa học thống kê)

[6] KACKER, R. and JONES, A. On use of Bayesian statistics to make the guide to the expression of uncertainty in measurement consistent. Metrologia, 40:235-248, 2003 (Sử dụng thống kê Bayes để đưa ra hướng dẫn với cách diễn tả độ không đảm bảo trong đo lường phù hợp)

[7] ELSTER, C.W., WOGER, W. and COX, M.G. Draft GUM Supplement 1 and Bayesian analysis. Metrologia, 44:L31-L32, 2007

[8] WILLINK, R. A procedure for the evaluation of measurement uncertainty based on moments. Metrologia, 42:329-343, 2005 (Quy trình đánh giá độ không đảm bảo đo dựa trên các mô men)

[9] LIRA, I. and WOGER, W. Comparison between the conventional and Bayesian approaches to evaluate measurement data. Metrologia, 43:S249-S259, 2006 (So sánh giữa cách tiếp cận quy ước và cách tiếp cận Bayes để đánh giá dữ liệu đo)

[10] GUTHRIE, W.F., LIU, H.K., RUKHIN, A.L., TOMAN, B., WANG, C.M. and ZHANG, N.F. Three statistical paradigms for the assessment and interpretation of measurement uncertainty. *Data Modeling for Metrology and Testing in Measurement Sciences*, edit. Pavese, F. and Forbes, A. B. Birkhauser, Boston, 2008 (Mô hình hóa dữ liệu cho đo lường và kiểm nghiệm trong khoa học đo lường)

[11] KIRKUP, L. and FRENKEL, B. *An introduction to Uncertainty in Measurement using the GUM*. Cambridge University Press, Cambridge UK, 2006 (Giới thiệu độ không đảm bảo đo bằng cách sử dụng GUM)

[4] Tiêu chuẩn này hiện đã bị hủy bỏ và thay thế bằng TCVN ISO/IEC 17025:2017 (ISO/IEC 17025:2017).

[12] BAYES, T. An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370-418, 1764. (facsimile available at http://www.stat.ucla.edu/history/essay.pdf)

[13] FISHER, R.A. Inverse probability. Proc. Comb. Philos. Soc., 26:528-535, 1930

[14] CASELLA, G. and BERGER, R. *Statistical Inference*. Duxbury, MA, 2 edition, 2002 (Suy luận thống kê)

[15] GUTHRIE, W.F. Should (T1-T2) have larger uncertainty than T1? *Proceedings of the 8th International Conference on Temperature: Its Measurements and Control*, 2:887-892, 2002. http://www.itl.nist.gov./div898/pubs/author/guthrie/guthrie-2002-01.pdf) (Tiến trình của Hội thảo quốc tế lần thứ 8 về nhiệt độ: Phép đo và kiểm soát)

[16] EFRON, B. and TIBSHIRANI, R.J. An Introduction to the Bootstrap. *Monographs of Statistics and Applied Probability*, volume 57. Chapman and Hall, 1993 (Tài liệu chuyên khảo về thống kê và xác suất ứng dụng)

[17] EISENHART, C. Expression of the uncertainties of final measurements results. *NBS special publication*, NIST, Gaithersburg, MD, 1983 (Thể hiện độ không đảm bảo của kết quả đo cuối cùng. Ấn phẩm đặc biệt của NBS)

[18] R Development Core Team. R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, 2003. ISBN 3-900051-00-3, http://www.R-project.org (Ngôn ngữ và môi trường cho tính toán thống kê)

[19] LUNN, D.J. THOMAS, A., BEST, N. and SPIEGELHALTER, D. WinBUGS – a Bayesian modeling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10:325-337, 2000 (Khuôn khổ mô hình hóa Bayes: khái niệm, cấu trúc và mở rộng. Thống kê và tính toán)

[20] BERNARDO, J.M. and SMITH, A.F.M. Bayesian Theory. John Wiley and Sons Ltd. 1994

[21] HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag, New York, 2001 (Yếu tố của việc học thống kê: khai thác dữ liệu, suy luận và dự đoán)

[22] GELMAN, A., ARLIN, J.B., STERN, H.S., and RUBIN, D.B. *Bayesian Data Analysis*. Chapman and Hall, 1995 (Phân tích dữ liệu Bayes)

[23] BROWNE, W.J. and DRAPER, D.A comparison of Bayesian and likelihood-based methods for fitting multi-level models. Bayesian Analysis, 1:473-514,2006 (So sánh phương pháp Bayes và phương pháp dựa trên hàm hợp lý để khớp với các mô hình nhiều mức)

[24] WANG. C.M. and IYER, H.K. Propagation of uncertainties in measurements using generalized inference. Metrologia, 42:145-153, 2005 (Phổ biến độ không đảm bảo đo bằng cách sử dụng suy luận tổng quát)

[25] WANG. C.M. and IYER, H.K. A generalized confidence interval for a measurand in the presence of Type-A and Type-B uncertainties. *Measurements*, 39:856-863, 2006 (Khoảng tin cậy tổng quát cho thước đo trong độ không đảm bảo Loại A và Loại B)

[26] HANNIG, J., IYER, H.K. and PATTERSON, P.L. Fiducial generalized confidence intervals. *Journal of the American Statistical Association*, 101:254-269, 2006 (Khoảng tin cậy tổng quát dựa trên sự tin tưởng)

[27] WANG. C.M. and IYER, H.K. Uncertainty analysis for vector measurands using fiducial inference. Metrologia, 43:486-494, 2006 (Phân tích độ không đảm bảo đối với thước đo véc tơ bằng cách sử dụng suy luận tin tưởng)

[28] FRASER, D.A.S. *The Structure of Inference*. New York: Krieger, 1968 (Cấu trúc suy luận)

[29] HANNIG, J. On fiducial inference – the good, the bad and the ugly. Technical Report 2006/3, Department of Statistics, Colorado State University, Fort Collins, CO, 2006. URL http://www.stat.colostate.edu/research/2006_3.pdf

[30] IYER, H.K. and PATTERSON, P.L. A recipe for constructing generalized pivot quantities and generalized confidence intervals. Technical Report 2002/10, Department of Statistics, Colorado State University, Fort Collins, CO, 2002. URL http://www.stat.colostate.edu/research/2002_10.pdf (Công thức để xây dựng đại lượng then chốt tổng quát và khoảng tin cậy tổng quát)

[31] RUKHIN, A.L. and SEDRANSK, N. Statistics in metrology: international key comparisons and interlaboratory studies. *Journal of Data Science*, 7:393-412, 2007 (Thống kê trong đo lường: nghiên cứu so sánh và nghiên cứu liên phòng quốc tế chính)

[32] EFRON, B. Six questions raised by the bootstrap. In: *Exploring the Limits of Bootstrap* (R.LePage and L. Billard, editors) pages 99-126. Wiley, NY, 1992 (Khai thác giới hạn của Bootstrap)

[33] LINDLEY, D. and SMITH, A.F.M. Bayes estimates for the linear model, JRSS B., 34:1-41, 1972 (ước lượng Bayes cho mô hình tuyến tính)

[34] ELSTER, C. and TOMAN, B. Bayesian uncertainty analysis under prior ignorance of the measurand versus analysis using the Supplement 1 to the Guide: a comparison, Metrologia, 46:261-266, 2009 (Phân tích độ không đảm bảo Bayes theo sự thiếu hiểu biết trước của phép đo so với phân tích bằng cách sử dụng Bổ sung 1 cho hướng dẫn)

[35] HALL, B.D. Evaluating methods of calculating measurement uncertainty, *Metrologia*, 45:L5-L8, 2008 (Phương pháp đánh giá để tính độ không đảm bảo đo)

[36] WANG, C.M. and IYER, H.K. Fiducial intervals for the magnitude of an complex-valued quantity, *Metrologia*, 46:1 81-86, 2009 (Khoảng tin cậy cho độ lớn của đại lượng có giá trị phức tạp).