

GIỚI THIỆU PHƯƠNG PHÁP PHÂN TÍCH BAYES

Phần 1: Diễn giải kết quả chẩn đoán

Nguyễn Văn Tuấn*

Tóm tắt: Phương pháp phân tích dữ liệu cổ điển và phổ biến dựa vào lý thuyết phản chứng, nhưng là một phương pháp có nhiều khiếm khuyết vốn đã được chỉ ra ngay từ khi phương pháp này ra đời vào khoảng 100 năm trước. Trong vài thập niên gần đây, phương pháp phân tích Bayes càng ngày càng phổ biến và đã đạt được nhiều thành công ngoạn mục trong di truyền học, nghiên cứu khoa học và nghiên cứu lâm sàng. Bài viết này giới thiệu ứng dụng phương pháp phân tích Bayes trong vài trường hợp lâm sàng.

Summary: The traditional frequentist methodology of data analysis relies on the concept of proof by contradiction. This methodology has several flaws and deficiencies which have been pointed out since its inception about 100 years ago. In recent decades, the Bayesian methodology has regained its prominence as an alternative approach to scientific inference and data analysis. The Bayesian methodology has proven a great success in genetics, scientific research, and clinical research. In this article, I introduce the basic concept of Bayesian inference via two cases of clinical diagnosis.

Hai trường hợp lâm sàng

Trường hợp 1. Nữ, 47 tuổi, tự phát hiện một u nhỏ bên vú trái và chị quan ngại đó là bướu ung thư nên đến gặp bác sĩ. Sau khi khám tổng quát và xem xét tiền sử gia đình, bác sĩ đề nghị chị đi chụp nhũ ảnh. Kết quả nhũ ảnh là dương tính. Chị muốn biết nguy cơ chị mắc bệnh ung thư vú là bao nhiêu. Bác sĩ nên trả lời như thế nào?

Trường hợp 2. Nam bệnh nhân, 60 tuổi, sức khỏe bình thường, BMI 23 kg/m², không có tiền căn gia đình với bệnh tiểu đường. Tuy nhiên kết quả xét nghiệm glucose trong máu là 127 mg/dL. Theo tiêu chuẩn do ADA đề nghị, ông được chẩn đoán là tiểu đường, nhưng ông không tin. Ông muốn biết nguy cơ thật mình mắc bệnh tiểu đường. Bác sĩ nên có lời khuyên nào cho bệnh nhân?

Có thể nói hai trường hợp trên rất tiêu biểu trong lâm sàng. Trước một xét nghiệm như nhũ ảnh với kết quả dương tính, và biết rằng phương pháp xét nghiệm có những hạn chế về độ chính xác, người bác sĩ cũng như bệnh nhân muốn biết khả năng mình mắc bệnh là cao hay thấp. Tương tự, đối với các kết quả xét nghiệm không có giá trị nhị phân mà là một dãy giá trị như nồng độ glucose trong máu, và kết quả xét nghiệm gần ngưỡng chẩn đoán bệnh, người thầy thuốc phân vân không biết nên ra mức bệnh hay

không mắc bệnh, bởi vì ai cũng biết rằng bất cứ xét nghiệm nào cũng không hoàn hảo. Hai trường hợp tiêu biểu trên cũng nói lên một đặc điểm của y học hiện đại: đó là tính bất định trong bất cứ đo lường nào, bất cứ xét nghiệm nào, và do đó bất cứ chẩn đoán nào. Chẳng thế mà “ông tổ” y khoa William Osler đã từng nói y khoa là một khoa học bất định và một nghệ thuật xác suất (medicine is a science of uncertainty and an art of probability).⁽¹⁾

Xử lý tình trạng bất định đòi hỏi đến khoa học. Xử lý xác suất cần phải có nghệ thuật. Trong vài năm gần đây, một phương pháp phân tích “mới” ra đời và đang dần trở thành phổ biến trong nghiên cứu khoa học và nghiên cứu lâm sàng có thể đáp ứng hai nhu cầu về khoa học và nghệ thuật. Nói là “mới” nhưng trong thực tế thì không mới, bởi vì cơ sở lý thuyết của phương pháp này đã ra đời từ thế kỷ 18. Đó là suy luận theo trường phái Bayes (Bayesian inference) do Thomas Bayes đề xuất vào năm 1763.⁽²⁾ Thomas Bayes là một linh mục, nhưng cũng là một nhà toán học tài tử. Tuy là “tài tử” nhưng di sản của ông để lại (chỉ một bài báo duy nhất) làm thay đổi cả thế giới khoa học, thay đổi cách suy nghĩ về sự bất định trong khoa học, và chỉ ra một phương pháp suy luận hoàn toàn logic. Ngày nay, phương pháp Bayes được ứng dụng trong hầu hết tất cả lĩnh vực khoa học, kể cả trong công nghệ thông tin (ứng dụng Bayes trong việc ngăn chặn những thư rác điện tử), tiên lượng kinh tế, phân tích các mối quan hệ xã hội, và lý giải qui trình suy nghĩ của con người. Ngày nay, suy luận theo trường phái Bayes được nhắc đến trên báo chí đại chúng chứ không chỉ trong báo khoa học. Những tờ báo lớn như *New York Times*, *Economist*, *Guardian*, v.v. đều thường xuyên nhắc đến phương pháp suy luận Bayes.

Suy luận Bayes dựa vào định lý Bayes (Bayesian theorem). Có thể phát biểu định lý Bayes theo ngôn ngữ hàng ngày như sau: những gì chúng ta biết là tổng hợp những gì chúng ta *đã biết* cộng với *chứng cứ thực tế*. Có thể nói rằng định lý Bayes thể hiện cách suy nghĩ rất phổ biến của tất cả chúng ta: đó là chúng ta tiếp thu kiến thức theo kiểu tích lũy. Trong hai trường hợp trên, trước khi gửi bệnh nhân đi xét nghiệm, chúng ta đã biết được khả năng bệnh nhân mắc bệnh như thế nào (qua các thông tin về tỉ lệ

*Viện nghiên cứu y khoa Garvan Sydney, Australia

hiện hành trong cộng đồng), sau khi có kết quả xét nghiệm chúng ta có thêm chứng cứ thực tế, và hai thông tin này giúp cho chúng ta đánh giá lại khả năng mắc bệnh của bệnh nhân.

Định lý Bayes dĩ nhiên cũng có thể mô tả một cách đơn giản qua xác suất. Gọi H là bệnh trạng, và D là chứng cứ (có thể là kết quả xét nghiệm hay dữ liệu), Định lý Bayes phát biểu rằng xác suất H với điều kiện D xảy ra – ký hiệu $P(H | D)$ – là:

$$P(H | D) = \frac{P(D | H) \times P(H)}{P(D)} \quad [1]$$

trong đó

- $P(H)$ là khả năng mắc bệnh trước khi xét nghiệm; và
- $P(D | H)$ là xác suất kết quả dương tính với điều kiện có bệnh H ;
- $P(D)$ là độ phân bố của dữ liệu.

Nhìn qua định lý trên, chúng ta thấy suy luận Bayes có 3 thông tin. Thông tin thứ nhất là thông tin mà chúng ta muốn biết, thuật ngữ tiếng Anh gọi là *posterior information* – thông tin hậu định. Thông tin thứ hai là thông tin chúng ta đã biết, tiếng Anh là *prior information* – thông tin tiền định. Và, thông tin thứ ba là thông tin thực tế, thuật ngữ tiếng Anh là *likelihood*. Ở đây, “thông tin” có nghĩa là khả năng hay xác suất. Chúng ta muốn biết khả năng bệnh nhân thật sự mắc bệnh. Do đó, 3 yếu tố trên thường được gọi là posterior probability, prior probability, và likelihood, có thể thể hiện qua công thức chung như sau:

Xác suất hậu định = Xác suất tiền định + Dữ liệu thực tế

Chúng ta thử xét qua 3 thông tin trong công thức trên một cách chi tiết hơn như sau:

Thông tin tiền định

Trước khi thực hiện một công trình nghiên cứu, chúng ta đã có vài ý niệm về mức độ ảnh hưởng của một liệu pháp can thiệp. Trước khi có kết quả xét nghiệm, chúng ta thường đã biết khả năng một cá nhân mắc bệnh cao cỡ nào. Những thông tin chúng ta biết trước như thế được gọi là thông tin tiền định. Trong phân tích và suy luận Bayes, có thể nói thông tin tiền định đóng một vai trò quan trọng. Quan trọng là vì kết quả phân tích có thể thay đổi tùy theo cách chúng ta cung cấp thông tin tiền định. Quay lại hai trường hợp đặt ra trên đây, chúng ta thử xét qua thông tin tiền định.

Đối với trường hợp 1, thông tin tiền định là khả năng mắc bệnh. Khả năng này có thể là tỉ lệ hiện hành (prevalence) ung thư vú trong cộng đồng thuộc độ tuổi của cá nhân. Thông tin tiền định cũng

có thể là giá trị tiên lượng qua mô hình Gail.⁽³⁾ Chẳng hạn như đối với phụ nữ 47 tuổi, không có tiền sử ung thư vú, không có những yếu tố nguy cơ khác, thì số liệu dịch tễ học cho biết tỉ lệ hiện hành ung thư vú là khoảng 1%. Do đó, chúng ta có một thông tin tiền định: $P = 0,01$.

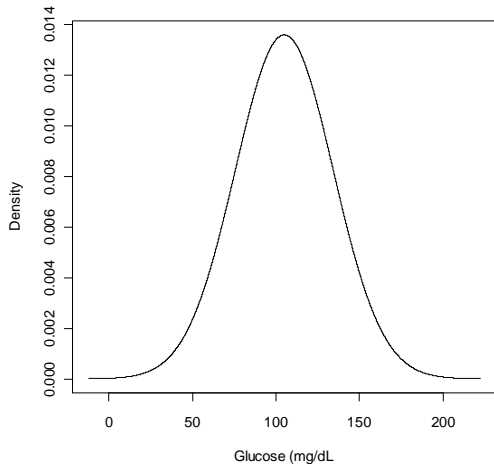
Thông tin tiền định trong trường hợp 1 còn là độ chính xác của phương pháp chụp nhũ ảnh. Thông thường, có hai chỉ số có thể sử dụng phản ánh độ chính xác của một phương pháp xét nghiệm: *độ nhạy* (sensitivity) và *độ đặc hiệu* (specificity). Thật ra, hai thuật ngữ tiếng Anh này không hẳn thích hợp, nhưng vì giới y khoa đã sử dụng quá lâu nên chúng ta tạm chấp nhận hai thuật ngữ đó. Tôi sẽ giải thích hai chỉ số này như sau:

• **Độ nhạy** là xác suất có kết quả dương tính nếu cá nhân thật sự mắc bệnh. Nói cách khác, độ nhạy trả lời câu hỏi: nếu 100 người mắc bệnh ung thư vú đều đi chụp nhũ ảnh thì có bao nhiêu người có kết quả dương tính. Nếu phương pháp nhũ ảnh hoàn toàn chính xác, chúng ta kỳ vọng tất cả 100 người sẽ có kết quả dương tính. Nhưng trong thực tế, không có phương pháp nào hoàn chỉnh, nên độ nhạy của nhũ ảnh thường khoảng 90% hoặc thấp hơn (nhưng chúng ta sẽ lạc quan với 90%).⁽⁴⁾

• **Độ đặc hiệu** là xác suất có kết quả âm tính nếu cá nhân thật sự không mắc bệnh. Thông thường độ đặc hiệu của nhũ ảnh khoảng 80%. Độ đặc hiệu 80% có thể hiểu như sau: nếu 100 người không mắc bệnh đều đi chụp nhũ ảnh thì sẽ có 80 người có kết quả âm tính. Nói cách khác, sẽ có 20 người có kết quả dương tính, và đây là trường hợp *dương tính giả* (false positive).

Đối với trường hợp 2, thông tin tiền định là sự phân bố về nồng độ glucose trong cộng đồng. Trong trường hợp 1, thông tin tiền định là tỉ lệ hiện hành mắc bệnh ung thư vú. Nhưng trong trường hợp 2, thông tin tiền định là sự phân bố glucose trong cộng đồng, nhất là ở những người cùng giới tính và cùng độ tuổi. Hai thông số để phản ánh một phân bố là số trung bình và độ lệch chuẩn (hoặc phương sai). Nghiên cứu của chúng tôi (chưa công bố) trong cộng đồng người Việt cho thấy ở nam 60 tuổi, nồng độ glucose trung bình là 105 mg/dL và phương sai là 860 mg/dL². Có thể hình dung sự phân bố của glucose trong cộng đồng qua biểu đồ 1.

Một thông tin tiền định khác là độ tin cậy của phương pháp đo lường glucose. Chúng ta biết rằng nồng độ glucose trong máu dao động trong mỗi cá nhân ngay trong điều kiện bình thường (không có can thiệp sinh học). Chẳng hạn như trường hợp



Biểu đồ 1: Phân bố nồng độ glucose trong cộng đồng nam giới 60 tuổi, trung bình 105 mg/dL và phương sai 860 mg/dL²

bệnh nhân được đo glucose 5 lần liên tiếp trong 5 ngày, và kết quả như sau (mg/dL): 127, 124, 125, 120, 126

Nếu chỉ dựa vào kết quả ngày thứ nhất và ngày thứ năm, cá nhân này được phân vào nhóm tiểu đường (vì nồng độ glucose cao hơn hay bằng 126 mg/dL). Nhưng nếu dựa vào kết quả ngày thứ hai, thứ ba và thứ tư thì cá nhân không được chẩn đoán tiểu đường. Đây là một tình trạng bất định rất phổ biến trong lâm sàng và chẩn đoán. Khác với sự bất định trong chẩn đoán ung thư vú (khi độ chính xác được định lượng bằng độ nhạy, độ đặc hiệu hay tỉ lệ dương tính giả), sự bất định trong các đo lường mang tính liên tục (continuous variable) như nồng độ glucose được định lượng bằng *hệ số tin cậy* mà thuật ngữ tiếng Anh là coefficient of reliability.

Hệ số tin cậy cũng giống như hệ số tương quan (coefficient of correlation). Hệ số tương quan đo lường mức độ tương quan giữa hai biến số. Để hiểu hệ số tin cậy, chúng ta cần đến khái niệm giá trị thật (true values) và giá trị quan sát (observed values hay measured values). Chẳng hạn trong trường hợp trên, chúng ta không biết nồng độ glucose thật của cá nhân đó là bao nhiêu, mà chỉ biết các giá trị đo được dao động trong khoảng 120 đến 127 nmg/dL. Tuy không biết giá trị thật là bao nhiêu, chúng ta có thể ước tính từ giá trị đo lường được, và ước lượng số (estimate) đơn giản nhất là giá trị trung bình. Hệ số tin cậy đo lường mức độ tương quan giữa giá trị quan sát được và giá trị thật. Hệ số tin cậy dao động từ 0 (hoàn toàn không tin cậy) đến 1 (tin cậy tuyệt đối).

Để ước tính hệ số tin cậy, người ta thường làm những nghiên cứu ngắn hạn. Trong đó, một nhóm

cá nhân được lấy máu và đo hơn 2 lần (thường là 3 lần). Từ dữ liệu đó, có thể ước tính phương sai (variance) nồng độ glucose. Phương sai này thực chất gồm có 2 thành phần: phương sai do dao động trung bình trong mỗi cá nhân (viết tắt là W) và phương sai do dao động giữa các cá nhân (B). Hệ số tin cậy (viết tắt là R) được tính bằng cách lấy B chia cho tổng phương sai:

$$R = \frac{B}{B + W}$$

Hệ số tin cậy và đặc biệt là phương sai W là những thông tin tiền định rất quan trọng trong việc đi đến một chẩn đoán chính xác về lâu về dài. Nghiên cứu trước đây của chúng tôi cho thấy phương sai W là 815 mg/dL², và hệ số tin cậy là 0.86.

Tóm lại, thông tin tiền định của hai trường hợp có thể tóm lược trong bảng số liệu sau đây:

Thông tin tiền định	Trường hợp 1	Trường hợp 2
Về tần số mắc bệnh hay phân bố trong cộng đồng	Tần số mắc bệnh là 1%, hay $P=0.01$	Phân bố glucose trong cộng đồng, với trung bình 105 mg/dL và phương sai 860 mg/dL ² .
Độ chính xác và tin cậy của phương pháp xét nghiệm	Độ nhạy: 0.90 Độ đặc hiệu: 0.80 Dương tính giả: 0.20	Hệ số tin cậy: $R=0.86$ Phương sai trung bình trong mỗi cá nhân $W = 815$

Dữ liệu thực tế

Dữ liệu thực tế, trong bối cảnh hai trường hợp trên, là kết quả xét nghiệm.

Đối với trường hợp 1 (ngghi ngờ ung thư vú), kết quả xét nghiệm đơn giản là dương tính.

Đối với trường hợp 2 (tiểu đường) thì dữ liệu thực tế có thể mô tả bằng luật phân phối. Chúng ta biết rằng nồng độ glucose tuân theo luật phân phối chuẩn (normal distribution). Luật phân phối chuẩn, như đề cập trên, được xác định bằng hai thông số: trung bình và phương sai. Luật phân phối có thể áp dụng để mô tả một biến số cho một nhóm đối tượng hay cho một cá nhân.

Trong trường hợp 2, chúng ta chỉ có một kết quả xét nghiệm glucose với giá trị 127 mg/dL. Nhưng chúng ta biết rằng nếu nồng độ glucose của cá nhân này được đo nhiều lần thì chắc chắn nồng độ không phải là 127 nhưng có thể dao động theo luật phân phối chuẩn. Do đó, chúng ta có thể phát biểu rằng đối với cá nhân 2, nồng độ glucose tuân theo luật phân phối chuẩn với trung bình 127 mg/dL và phương sai 815 mg/dL² (dựa vào thông tin tiền định). Các giá trị glucose của cá nhân này có thể thể hiện bằng biểu đồ 2.

Thông tin hậu định

Dựa vào hai thông tin tiên định và dữ liệu thực tế, chúng ta có thể trả lời câu hỏi đặt ra lúc ban đầu. Phương pháp để có thông tin hậu định là ứng dụng định lý Bayes như mô tả trong phần đầu.

Đối với trường hợp 1, cần nhắc lại rằng chúng ta muốn biết xác suất bệnh nhân mắc bệnh ung thư vú là bao nhiêu sau khi đã có kết quả dương tính từ nhũ ảnh. Để hiểu công thức 1 tôi sẽ giải thích bằng cách giả định rằng chúng ta có một quần thể gồm 1000 phụ nữ trong độ tuổi 47. Chúng ta biết rằng (qua thông tin tiên định) trong số 1000 người sẽ có 10 người mắc bệnh ung thư vú, và do đó 990 người không mắc bệnh.

Cũng qua thông tin tiên định, chúng ta biết rằng độ nhạy là 90%. Do đó, trong số 10 người mắc bệnh, sẽ có 9 người (10 x 0,90) có kết quả dương tính (và 1 người có kết quả âm tính).

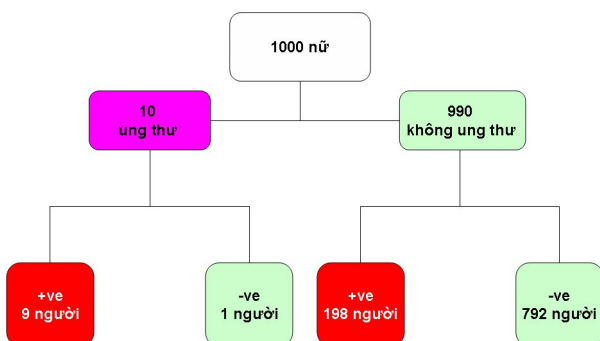
Ngoài ra, vì độ đặc hiệu là 80%. Do đó, trong số 990 người không mắc bệnh, sẽ có 792 người (990 x 0,80) có kết quả âm tính (và 198 người có kết quả dương tính - ở đây thật ra là dương tính giả).

Như vậy, tổng cộng chúng ta có 9 + 198 = 207 người có kết quả dương tính. Tuy nhiên, trong số này, chỉ có 9 người thật sự mắc bệnh ung thư vú, còn 198 người không mắc bệnh (vì do tỉ lệ dương tính giả). Do đó, xác suất mà người phụ nữ mắc bệnh ung thư vú với điều kiện có kết quả dương tính là: 9 / 207 = 4,3%.

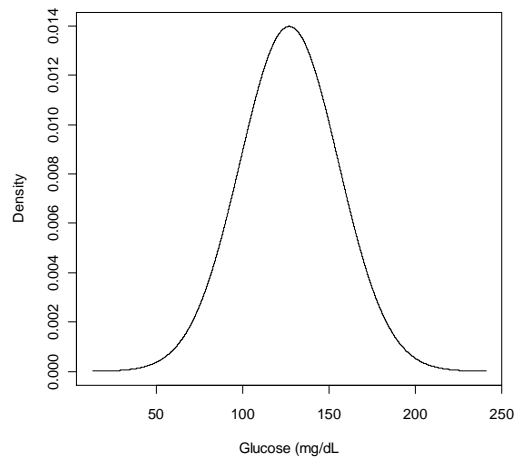
Thật ra, có thể ước tính dựa vào công thức (1) một cách nhanh gọn hơn. Gọi độ nhạy là $P(D | H)$; xác suất dương tính giả là $P(D | NoH)$, NoH là không mắc bệnh; và tỉ lệ hiện hành là $P(H)$, xác suất mắc bệnh với kết quả dương tính là:

$$P(H | D) = \frac{P(H) \times P(D | H)}{P(H) \times P(D | H) + [1 - P(H)] \times P(D | NoH)}$$

$$P(H | D) = \frac{0.01 \times 0.90}{0.01 \times 0.90 + 0.10 \times 0.2} = 0.043$$



Có lẽ nhiều người sẽ ngạc nhiên tại sao xác suất mắc bệnh thấp. Nhưng ở đây, cần nói thêm rằng



Biểu đồ 2: Phân bố nồng độ glucose của một cá nhân 60 tuổi, trung bình 127 mg/dL và phương sai 815 mg/dL²

phần lớn bác sĩ hiểu lầm rằng độ nhạy là xác suất mắc bệnh, nên nhiều bác sĩ cho rằng xác suất mắc bệnh là 90%. Nhưng cách diễn giải đó sai. Sai lầm là vì có sự nhầm lẫn giữa hai xác suất. Gọi H là bệnh ung thư và D là kết quả xét nghiệm dương tính, độ nhạy được định nghĩa là: $P(D | H)$

Còn xác suất mắc bệnh với điều kiện có kết quả dương tính là: $P(H | D)$

Tuy cách viết chỉ khác nhau về thứ tự của H và D , nhưng ý nghĩa thì rất khác nhau! Độ nhạy phản ánh độ chính xác của phương pháp nhũ ảnh chứ không nói lên xác suất mắc bệnh ung thư vú.

Trong trường hợp 2, câu hỏi về thông tin hậu định có phần khác hơn trường hợp 1. Chúng ta muốn biết với kết quả xét nghiệm glucose trong máu là 127 mg/dL, xác suất mà cá nhân mắc bệnh tiểu đường là bao nhiêu. Vì 126 mg/dL là ngưỡng để chẩn đoán tiểu đường, nên câu hỏi có thể viết bằng ngôn ngữ xác suất như sau:

$$P(\text{glucose} > 126) = ?$$

Để trả lời câu hỏi đó, chúng ta cần xác định luật phân phối của glucose cho cá nhân. Gọi số trung bình và độ lệch chuẩn của glucose trong cộng đồng lần lượt là m_{prior} và s_{prior} . Tương tự, gọi số trung bình và độ lệch chuẩn của glucose của cá nhân là m_{data} và s_{data} . Qua vài phép tính, có thể chứng minh rằng glucose trung bình và độ lệch chuẩn của cá nhân đó là:

$$m_{\text{posterior}} = \frac{\frac{m_{\text{prior}}}{s_{\text{prior}}^2} + \frac{m_{\text{data}}}{s_{\text{data}}^2}}{\frac{1}{s_{\text{prior}}^2} + \frac{1}{s_{\text{data}}^2}}$$

và

$$S_{posterior} = \frac{1}{\sqrt{\left(\frac{1}{S_{prior}^2} + \frac{1}{S_{data}^2}\right)}}$$

Quay lại trường hợp 2, xin nhắc lại, chúng ta biết rằng trong cộng đồng, nồng độ glucose tuân theo luật phân phối chuẩn với trung bình 105 mg/dL và phương sai 860 mg/dL². Chúng ta cũng biết nồng độ glucose của cá nhân tuân theo luật phân phối chuẩn với trung bình 127 mg/dL và phương sai 815 mg/dL². Từ hai thông tin này, chúng ta có thể xác định phân phối glucose của cá nhân đó sau khi điều chỉnh với thông tin tiên định.

$$m_{posterior} = \frac{\frac{105}{860} + \frac{127}{815}}{\frac{1}{860} + \frac{1}{815}} = 116$$

$$S_{posterior} = \frac{1}{\sqrt{\left(\frac{1}{860} + \frac{1}{815}\right)}} = 20.5$$

Nói cách khác, nồng độ glucose trung bình về lâu về dài (hay nồng độ thật) của cá nhân là 116 mg/dL, nhưng có thể dao động trong khoảng 75.8 đến 156.2 mg/dL với xác suất 95% (tức 116–1.96×20.5 đến 116+1.96×20.5). Có thể hình dung phân bố glucose của cá nhân này như Biểu đồ 3.

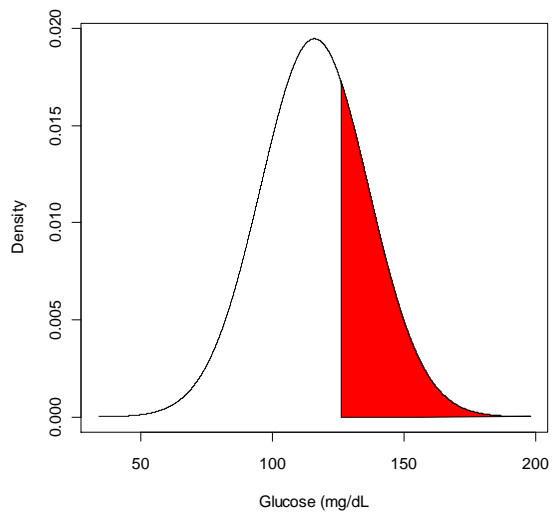
Nhưng chúng ta muốn biết P(glucose > 126). Bởi vì nồng độ glucose tuân theo luật phân phối chuẩn với trung bình 116 mg/dL và độ lệch chuẩn 20,5 mg/dL, nên có thể mô tả bằng công thức sau đây:

$$f(\text{glucose}) = \frac{1}{20.5\sqrt{2\pi}} e^{-\frac{(\text{glucose} - 116)^2}{2 \times (20.5)^2}}$$

Xác suất trên chính là diện tích dưới đường biểu diễn của hàm số glucose cho cá nhân đó. Theo đó:

$$P(\text{glucose} > 126) = \int_{126}^{\infty} f(\text{glucose}) = 0,312$$

Nói cách khác, mặc dù nồng độ glucose của cá nhân này nằm trong ngưỡng chẩn đoán tiểu đường (127 mg/dL), nhưng vì biết được độ tin cậy của phương pháp xét nghiệm và thông tin tiên định trong cộng đồng, xác suất mà bệnh nhân có nồng độ glucose trên 126 mg/dL thật ra chỉ 31,2%, chưa đủ thuyết phục để chẩn đoán cá nhân này mắc bệnh tiểu đường. Chi tiết về lí thuyết và cách ứng dụng có thể xem qua vài nghiên cứu trước của người viết bài này^(5,6) và đồng nghiệp khác.⁽⁷⁻¹⁰⁾



Biểu đồ 3: Phân bố nồng độ glucose của một cá nhân 60 tuổi, trung bình 116 mg/dL và độ lệch chuẩn 20.5 mg/dL. Diện tích dưới đường biểu diễn (màu đỏ) là xác suất cá nhân có nồng độ glucose cao hơn 126 mg/dL (tức có thể chẩn đoán tiểu đường)

Suy luận Bayes

Phương pháp phân tích thống kê đóng vai trò rất quan trọng trong nghiên cứu y khoa, đặc biệt là nghiên cứu lâm sàng. Không có thống kê, nghiên cứu lâm sàng chỉ là những con số vô hồn và không có giá trị khoa học. Trong thời gian khoảng 100 năm qua, khoa học thống kê đã phát triển được rất nhiều phương pháp phân tích có thể ứng dụng cho rất nhiều tình huống khác nhau trong nghiên cứu y học. Những phương pháp này dựa vào triết lý phản nghiệm (falsificationism), mà theo đó nhà nghiên cứu tiến hành 3 bước.

- Bước thứ nhất, phát biểu giả thuyết vô hiệu H_0 ;
- Bước thứ hai, thu thập dữ liệu D ; và
- Bước thứ ba, ước tính xác suất D xảy ra nếu H_0 là thật: $P(D | H_0)$.

Đó là qui trình chứng minh đảo ngược (proof by contradiction) hay còn gọi là phản chứng. Qui trình này có thể áp dụng trong toán học, nhưng khi ứng dụng vào y học thì trở nên ... vô duyên. Trong y khoa chúng ta muốn biết với dữ liệu thu thập được, xác suất giả thuyết H_0 thật là bao nhiêu, chứ không ai lại đặt câu hỏi nếu giả thuyết H_0 là đúng thì xác suất D xảy ra là bao nhiêu! Tương tự, không ai ngó ngàng hỏi nếu tôi mắc bệnh thì xác suất kết quả xét nghiệm dương tính là bao nhiêu (vì nếu mắc bệnh thì phải điều trị, chứ không hỏi như thế). Người ta hỏi: nếu kết quả xét nghiệm là dương tính, khả năng tôi mắc bệnh là bao

nhieu, để lên kế hoạch đối phó. Phương pháp phân tích cổ điển không thể trả lời câu hỏi này (mà chỉ trả lời câu hỏi đảo ngược).

Trong y khoa, bất định là qui luật chứ không phải là điều bất bình thường. Để xử lý sự bất định, chúng ta cần một phương pháp phân tích và suy luận logic. Về mặt đánh giá bằng chứng trong điều kiện bất định, chỉ có phương pháp Bayes là phương pháp logic nhất. Có thể nói rằng phương pháp suy luận Bayes không phải xa lạ gì với bác sĩ lâm sàng. Theo Bayes, kiến thức chúng ta tiếp thu qua kiến thức đã biết và dữ liệu thực tế. Tương tự, trong chẩn đoán bệnh, xác suất mắc bệnh tùy thuộc vào thông tin liên quan đến cá nhân và kết quả xét nghiệm. Nếu hai cá nhân có những yếu tố nguy cơ giống nhau, nhưng cá nhân có kết quả dương tính sẽ có khả năng mắc bệnh cao hơn so với cá nhân có kết quả âm tính. Nhìn theo góc độ này, suy luận Bayes chính là một phương pháp cá nhân hóa trong định lượng lâm sàng và chẩn đoán bệnh. Do đó, phương pháp suy luận Bayes rất thích hợp cho các nhà lâm sàng.

Tóm lại, suy luận Bayes là một phương pháp suy luận dựa vào thông tin chúng ta đã thu thập trước đây cộng với dữ liệu thực tế để có được tri thức mới hoàn chỉnh hơn. Có thể nói đó là một qui trình suy luận tích lũy. Tri thức khoa học là tri thức được tích lũy theo thời gian, và phương pháp Bayes cung cấp cho chúng ta một phương tiện rất có ích cho sự phát triển khoa học.

Quay lại hai trường hợp nêu ra trong phần đầu bài viết. Đối với trường hợp 1, chúng ta có thể trả lời rằng kết quả nhũ ảnh là dương tính, và đó là một kết quả đáng quan tâm, nhưng cần phải diễn giải trong điều kiện bất định của phương pháp chụp nhũ ảnh. Cứ 100 người có kết quả dương tính, thì có khoảng 4 người thật sự bị ung thư vú, nhưng bác sĩ không thể chắc chắn rằng cá nhân này sẽ mắc bệnh. Xác suất 4% có lẽ thấp, nhưng vẫn thể hiện cao gấp 4 lần so với một người phụ nữ “trung bình” cùng tuổi. Thật ra, trong thực tế, kết quả nhũ ảnh chỉ là một bước (có thể quan trọng) trong qui trình chẩn đoán và xét nghiệm để xác định khả năng mắc bệnh.

Đối với trường hợp 2, mặc dù kết quả xét nghiệm cho thấy nồng độ glucose trong máu là 127 (tức ngưỡng tiểu đường), nhưng vì phương pháp xét

nghiệm có độ tin cậy chưa cao, nên kết quả tự nó cũng mang tính bất định. Sau khi xem xét độ dao động trong mỗi cá nhân và kết hợp với thông tin trong cộng đồng, có thể nói rằng khả năng cá nhân này thật sự bị tiểu đường chỉ 31%, thấp hơn ngưỡng 95% khá xa. Trước kết quả và đánh giá này, có lẽ một xét nghiệm khác trong một thời điểm khác sẽ cung cấp thông tin chắc chắn hơn.

Phương pháp suy luận và phân tích dựa vào trường phái Bayes đã và đang trở nên phổ biến trong nghiên cứu y khoa. Ngày nay, hầu như bất cứ lĩnh vực nghiên cứu nào – nhất là chẩn đoán – đều cần đến phương pháp Bayes. Trong tình trạng bất định đo lường, ngay cả với những phương pháp đo lường có độ chính xác cao, suy luận Bayes chứng tỏ là một phương tiện rất có ích. Hi vọng rằng bài này đã cung cấp cho bạn đọc vài ý niệm và cách ứng dụng phương pháp suy luận Bayes trong chẩn đoán lâm sàng.

Tài liệu tham khảo

- (1) Osler W. Trích trong Clinical ethics: a practical approach to ethical decisions in clinical medicine (trang 17) của Albert R. Jonsen và cs. Nhà xuất bản McGraw-Hill Professional 2002.
- (2) Bayes, Thomas; Price, Mr. (1763). An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London* 53 (0): 370–418. Có thể xem bản gốc của bài báo quan trọng này tại địa chỉ: <http://www.stat.rice.edu/~blairc/seminar/Files/danTalk.pdf>
- (3) Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Chairer C, Mulvihill JJ, Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989; 81:1879-1886.
- (4) Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249-267). Cambridge, England: Cambridge University Press.
- (5) Nguyen TV, et al. Within-subject variability and analytic imprecision of insulinlike growth factor axis and collagen markers: implications for clinical diagnosis and doping tests. *Clin Chem* 2008;54:1268–76.
- (6) Nguyen TV, Pocock NA, Eisman JA. On the interpretation of bone mineral density measurements and its change. Special Article: *J Clin Densitometry* 2000; 3:107-19.
- (7) Dunson DB. Commentary: Practical Advantages of Bayesian Analysis of Epidemiologic Data. *Am J Epidemiol* 2001;153:1222–6.
- (8) Spiegelhalter DJ, Myles JP, Jones DR, et al. An introduction to Bayesian methods in health technology assessment. *BMJ* 1999;319:508–12.
- (9) Lilford RJ, Braunholtz D. Who's afraid of Thomas Bayes? *J Epidemiol Community Health* 2000;54:731–9.
- (10) Berry DA. Bayesian clinical trials. *Nature Reviews Drug Discovery*. 2006;5(1):27-36.