

KIỂM TRA NGẪU NHIÊN HÀNG HÓA XUẤT NHẬP KHẨU

Võ Minh Trí⁽¹⁾, Ung Ngọc Quang⁽²⁾,

Dương Tôn Đảm⁽²⁾, Tô Anh Dũng⁽²⁾, Nguyễn Minh Hải⁽²⁾

⁽¹⁾Cục Hải Quan Thành phố Hồ Chí Minh, ⁽²⁾Trường Đại học Khoa học Tự nhiên – ĐHQG-HCM
(Bài nhận ngày 31 tháng 5 năm 2004, hoàn chỉnh sửa chữa ngày 01 tháng 7 năm 2004)

TÓM TẮT: Trong bài báo này, các tác giả ứng dụng phương pháp thống kê toán học vào bài toán kiểm tra ngẫu nhiên hàng xuất nhập khẩu trong Hải quan.

Từ khóa: Hàng đóng theo kiện, khai báo đúng, bác bỏ giả thuyết, chấp nhận giả thuyết.

1. Đặt vấn đề

Theo quy định hiện hành của Hải quan, hàng hóa xuất nhập khẩu qua các cửa khẩu Việt Nam phải kiểm tra theo các hình thức:

- Miễn kiểm tra
- Kiểm tra toàn bộ
- Kiểm tra ngẫu nhiên.(xem[1])

Trong bài báo này, ta chỉ xét cách thức kiểm tra ngẫu nhiên. Hiện nay, các tài liệu chính thức của Hải quan gọi cách kiểm tra này là kiểm tra xác suất. Cách này có nghĩa, từ lô hàng xuất nhập khẩu, người ta chọn ngẫu nhiên n kiện hàng để kiểm tra. Sau đó, trên cơ sở kiểm tra n kiện hàng này, người ta mới rút ra kết luận về toàn bộ lô hàng. Tuy nhiên theo cách thức kiểm tra như vậy, thì việc dùng danh từ “Kiểm tra ngẫu nhiên” sẽ đúng đắn hơn là dùng danh từ “Kiểm tra xác suất”. Vậy nên trong bài này, ta sẽ dùng danh từ “Kiểm tra ngẫu nhiên”.

Mặt khác, vì cách thức kiểm tra là ngẫu nhiên, nên ta có thể dùng phương pháp xác suất – thống kê để giải quyết bài toán này.

2. Phương hướng giải quyết

Trong thuật ngữ Hải quan, lô hàng được kiểm tra thường có các loại sau:

- Hàng xá (hàng không có bao bì, chia nhỏ khối lượng)
- Hàng tháo rời
- Hàng đóng theo kiện

Trong bài này, ta chỉ xét lô hàng đóng theo kiện. Nghĩa là trong lô hàng có nhiều kiện hàng và ta chỉ xét trường hợp các kiện hàng là đồng nhất (hàng hóa trong mỗi kiện giống hệt nhau) hoặc gần như đồng nhất. Các trường hợp khác sẽ được xét trong các bài báo tiếp theo.

Ngoài ra, hàng hóa được kiểm tra trên thực tế phải đúng so với khai báo của người khai thể hiện trên tờ khai của Hải quan, gồm các yếu tố sau:

- Tên hàng hóa
- Số lượng, khối lượng, trọng lượng hàng hóa
- Chất lượng, chủng loại hàng hóa
- Xuất xứ hàng hóa

Nếu phát hiện có sai ít nhất một trong 4 yếu tố trên, thì lô hàng được coi là “Khai báo sai”. Ngược lại, nếu cả 4 yếu tố trên đều không sai thì lô hàng được coi là “Khai báo đúng”.

Để kiểm tra, người ta lấy ngẫu nhiên n kiện hàng từ lô hàng (còn gọi là lấy mẫu ngẫu nhiên cỡ n). Theo qui định hiện hành của Hải quan thì tỷ lệ kiện hàng lấy ra để kiểm tra bằng 3%, 5% hoặc 10% số kiện hàng của toàn bộ lô hàng. (xem [1],[2],[3]).

Có nhiều bài toán có thể đặt ra xung quanh việc kiểm tra ngẫu nhiên. Nhưng bài báo này chỉ giải quyết một bài toán: Nếu số kiện hàng kiểm tra trên thực tế đều “Khai báo đúng” thì có thể kết luận rằng toàn bộ lô hàng là “Khai báo đúng” hay không?

Phương pháp toán học dùng để giải quyết bài toán này là một nhánh của Thống kê toán học có tên gọi là Lý thuyết kiểm định giả thuyết thống kê (có thể xem, chẳng hạn, các tài liệu [4],[5],[6],[7]).

3. Lời giải bài toán Kiểm định giả thuyết thống kê

Trong mục này, ta sẽ mô hình hóa bài toán kiểm tra ngẫu nhiên hàng hóa xuất nhập khẩu nói trên, thành bài toán kiểm định giả thuyết thống kê về tham số tỷ lệ.

Như đã biết, bài toán kiểm định giả thuyết thống kê là một bài toán lớn và quan trọng trong thống kê toán học. Người ta đã chứng minh được sự tồn tại lời giải tối ưu cho một số bài toán kiểm định giả thuyết về tham số. Ta sẽ sử dụng lời giải tối ưu đó vào việc giải quyết bài toán kiểm tra ngẫu nhiên đã nêu trong mục 2.

3.1 Định nghĩa: Giả sử ta phải kiểm tra một lô hàng gồm nhiều kiện hàng như đã nêu ở mục 2.

Đặt p = "Tỷ lệ kiện hàng "Khai báo đúng" trong toàn bộ lô hàng". Lúc đó p là một tham số thống kê chưa biết và được gọi là tham số tỷ lệ. Rõ ràng $p \in [0,1]$.

3.2 Định nghĩa: Chọn n kiện hàng từ lô hàng để kiểm tra .

Đặt S_n = " Số kiện hàng "Khai báo đúng" trong n kiện hàng lấy ra từ lô hàng ". Rõ ràng S_n là một đại lượng ngẫu nhiên rời rạc, nhận các giá trị : 0, 1, 2, ..., n . Hơn nữa có thể coi các lần kiểm tra là độc lập với nhau. Lúc đó S_n là đại lượng ngẫu nhiên nhị thức và ký hiệu: $S_n \sim B(n, p)$.

Vì tỷ lệ p chưa biết, nên ta tìm cách xác định nó bằng cách đưa ra giả thuyết rằng p bằng một tỷ lệ p_0 đã biết. Do đó ta có bài toán kiểm định giả thuyết về tham số tỷ lệ p được viết dưới dạng sau:

3.3 Định nghĩa:
$$\begin{cases} H : p = p_0 \\ K : p > p_0 \end{cases}$$

Trong đó H là giả thuyết , K là đối thuyết , $p_0 \in (0,1)$ đã biết và mức ý nghĩa α cho trước .

3.4 Lời giải tối ưu:

Như đã biết (xem, chẳng hạn, [5], [6], [7]), bài toán kiểm định giả thuyết về tham số tỷ lệ p đã có lời giải tối ưu. Lời giải ấy có dạng:

Bác bỏ H : $Z > C$ và chấp nhận H : $Z \leq C$

Trong đó $Z = \frac{\sqrt{n}(\hat{p} - p_0)}{\sqrt{p_0 \cdot q_0}}$, với $\hat{p} = \frac{S_n}{n}$ gọi là tỷ lệ mẫu, $q_0 = 1 - p_0$, còn hằng số C được xác

định bởi phương trình $\Phi(c) = 1 - \alpha$ và tra bảng chuẩn $N(0,1)$. Hơn nữa cỡ mẫu n phải thỏa điều kiện:

$$\begin{cases} n \cdot p_0 \geq 5 \text{ (*)} \\ n \cdot q_0 \geq 5 \end{cases}$$

Chú ý rằng bài toán với cỡ mẫu n thỏa điều kiện (*) được gọi là bài toán với cỡ mẫu lớn .

3.5 Nhận xét: Lời giải tối ưu nói trên chưa áp dụng chưa áp dụng được cho bài toán của ta. Để giải quyết nó, cần chú ý rằng, theo cách đặt bài toán, ta chỉ xét trường hợp: Tất cả n kiện hàng lấy ra kiểm tra đều "Khai báo đúng" (vì chỉ cần một kiện hàng "Khai báo sai" thì ta phải kiểm tra toàn bộ lô hàng và khi đó tính ngẫu nhiên sẽ không còn nữa).

Như vậy, ở đây ta có $S_n = n$. Do đó tỷ lệ mẫu $\hat{p} = \frac{S_n}{n} = 1$.

$$\text{Suy ra : } Z = \frac{\sqrt{n}(\hat{p} - p_0)}{\sqrt{p_0 \cdot q_0}} = \frac{\sqrt{n}(1 - p_0)}{\sqrt{p_0 \cdot q_0}} = \frac{\sqrt{n} \cdot q_0}{\sqrt{p_0 \cdot q_0}} = \frac{\sqrt{n \cdot q_0}}{\sqrt{p_0}}$$

Điều này có nghĩa Z chỉ phụ thuộc vào n, p_0, q_0 và thỏa điều kiện:
$$\begin{cases} n.p_0 \geq 5 \\ n.q_0 \geq 5 \end{cases} (*)$$

Vậy nên nếu cho n và p_0 những giá trị cụ thể thì ta sẽ biết khi nào thì việc kết luận toàn bộ lô hàng "Khai báo đúng" là nên chấp nhận, khi nào không nên chấp nhận.

Để thấy rõ điều này, ta xét 2 thí dụ cụ thể sau.

3.6 Thí dụ: Giả sử từ lô hàng ta lấy cỡ mẫu $n = 100$. Cho tỷ lệ giả thuyết $p_0 = 0.90$. Ấn định mức ý nghĩa $\alpha = 0.01$. Lúc đó bài toán kiểm định giả thuyết về tham số tỷ lệ p (chưa biết) và lời giải tối ưu của nó như sau: Bài toán:

$$\begin{cases} H : p = 0.90 \\ K : p > 0.90 \end{cases}$$

Lời giải tối ưu: Bác bỏ $H: Z > C$ và chấp nhận $H: Z \leq C$.

Để có thể dùng lời giải tối ưu thì các số n, p_0, q_0 phải thỏa điều kiện:
$$\begin{cases} n.p_0 \geq 5 \\ n.q_0 \geq 5 \end{cases} (*)$$

Với $n = 100; p_0 = 0.90; q_0 = 0.10$ thì điều kiện (*) được thỏa, nên ta có quyền dùng lời giải tối ưu.

$$\text{Tính toán ta được: } Z = \frac{\sqrt{n}(\hat{p} - p_0)}{\sqrt{p_0 \cdot q_0}} = \frac{\sqrt{n}q_0}{\sqrt{p_0}} = \frac{\sqrt{100 \cdot 0.10}}{\sqrt{0.90}} = 3.33.$$

Mặt khác, với $\alpha = 0.01$, tra bảng chuẩn $N(0,1)$, ta có: $C = 2.33$. Suy ra $Z > C$, nên ta bác bỏ H và do đó chấp nhận K . Tức là ta khẳng định rằng $p > 0.90$. Điều này có nghĩa, khi kiểm tra một mẫu gồm 100 kiện hàng đều "Khai báo đúng" thì việc khẳng định rằng tỷ lệ kiện hàng "Khai báo đúng" trong toàn bộ lô hàng sẽ cao hơn 90%. Nghĩa là tỷ lệ khai báo đúng của toàn bộ lô hàng khá cao.

3.7 Thí dụ: Xét trường hợp $n = 49, p_0 = 0.90, \alpha = 0.01$, ta có:

$$\text{Bài toán: } \begin{cases} H : p = 0.90 \\ K : p > 0.90 \end{cases}$$

Lời giải tối ưu: Bác bỏ $H: Z > C$ và chấp nhận $H: Z \leq C$.

$$\text{Trước hết, ta xét điều kiện: } \begin{cases} n.p_0 \geq 5 \\ n.q_0 \geq 5 \end{cases} (*)$$

Lúc đó $n.p_0 = 49 \cdot 0.90 = 44.1 > 5$, nhưng $n.q_0 = 49 \cdot 0.10 = 4.9 < 5$. Vậy điều kiện (*) không thỏa, nên ta không thể dùng lời giải tối ưu.

4. Kết luận

4.1 Qua các thí dụ trên, ta không phải lúc nào lời giải tối ưu cũng dùng được. Còn nếu dùng được thì chưa chắc tỷ lệ "Khai báo đúng" đáng được chấp nhận. Vì vậy ta phải lập một bảng số tỷ mỷ để nhân viên Hải quan khi nhìn vào bảng ấy sẽ biết ngay lô hàng nào có tỷ lệ "Khai báo đúng" khá cao, còn lô hàng nào có tỷ lệ "Khai báo đúng" chưa cao. Độc giả quan tâm tới các bảng số liệu này, có thể tham khảo theo địa chỉ email: minhhaikhtn@yahoo.com.

4.2 Việc loại trừ các yếu tố chủ quan mà chỉ xét sự việc dưới góc độ toán học thuần túy sẽ giúp cho người quản lý thấy rõ bản chất của quá trình kiểm tra hàng hóa và giúp người thực hiện kiểm tra định hướng phương pháp kiểm tra một lô hàng trên thực tế.

4.3 Để tiện cho việc theo dõi của độc giả, chúng tôi sẽ trình bày 2 biểu đồ về tỷ lệ chấp nhận được của các kiện hàng trong phần phụ lục dưới đây.

Lời cảm ơn: Các tác giả cảm ơn Giáo sư Nguyễn Bác Văn và các ông Đặng Đông Triều, Lê Thanh Nhân đã đóng góp nhiều ý kiến bổ ích cho bài báo.

ON THE PROBLEM OF RANDOMLY CHECKING IMPORT – EXPORT GOODS

Vo Minh Tri⁽¹⁾, Ung Ngoc Quang⁽²⁾,

Duong Ton Dam⁽²⁾, To Anh Dung⁽²⁾, Nguyen Minh Hai⁽²⁾

⁽¹⁾Customs of Ho Chi Minh City, ⁽²⁾University of Natural Sciences – VNU-HCM

ABSTRACT: *In this article, we apply the mathematical statistics method to the problem of randomly checking import – export goods in the customs.*

Key words: *The goods are baled, right declaration, hypothesis reject, hypothesis accept.*

TÀI LIỆU THAM KHẢO

- [1]. Tờ khai hàng hóa nhập khẩu, Bản lưu Hải quan, Cục Hải quan Tp. HCM, 2002 .
- [2]. Nghiên cứu Hải quan, số 1-2 / 2003, Tổng cục Hải quan.
- [3]. Nghiên cứu Hải quan, số 11 – 12 / 2003, Tổng cục Hải quan.
- [4]. V.K. Rohatgi, A.K. Ehsanes Saleh, *An Introduction to Probability and Statistics*, John Wiley, New York, 2001.
- [5]. I.S.Milton, J.C.Arnold, *An Introduction to Probability and Statistics*, Mc Graw – Hill, New York, 1995.
- [6]. Đặng Hùng Thắng, *Thống kê và Ứng dụng*, NXB Giáo dục, 1999.
- [7]. Nguyễn Bác Văn, *Xác suất và Xử lý số liệu thống kê*, NXB Giáo dục, 1996.

PHỤ LỤC

Các biểu đồ dùng trong kiểm tra ngẫu nhiên hàng hóa xuất nhập khẩu

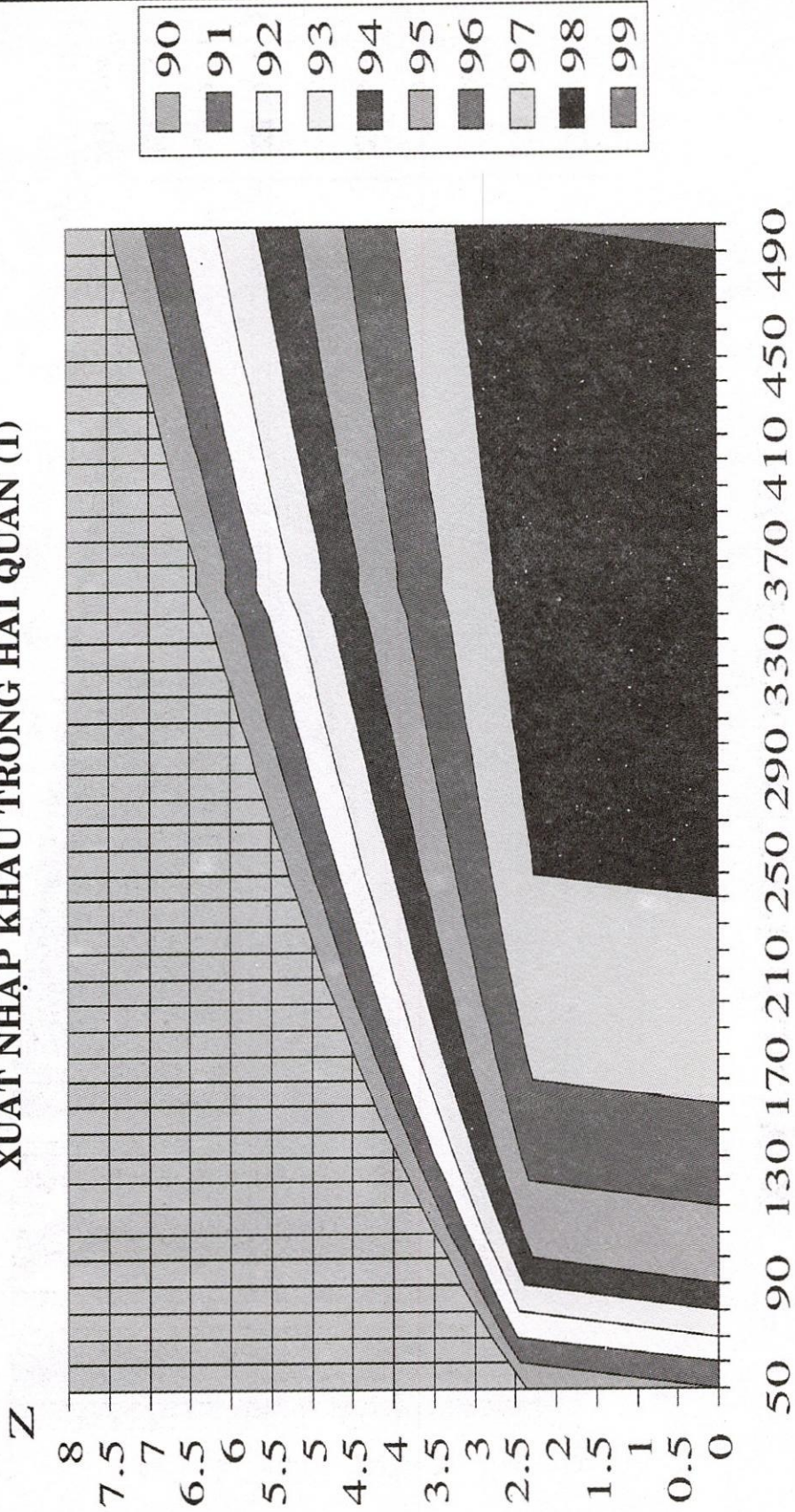
1. Chú thích biểu đồ 1:

- Trục hoành: ký hiệu là n , chỉ số lượng kiện hàng trong mẫu.
- Trục tung: ký hiệu là Z , chỉ giá trị của công thức: $Z = \frac{\sqrt{n}(\hat{p} - p_0)}{\sqrt{p_0 \cdot q_0}}$
- Các màu trong bảng ứng với các tỷ lệ $p_0 = 0.90; 0.91 \dots 0.99$

2. Chú thích biểu đồ 2:

- Trục hoành: ký hiệu là n , chỉ số lượng kiện hàng trong mẫu.
- Trục tung: ký hiệu là p_0 chỉ tỷ lệ giả thiết kiện hàng “Khai báo đúng”
- Trục đứng: ký hiệu là Z , chỉ giá trị của công thức: $Z = \frac{\sqrt{n}(\hat{p} - p_0)}{\sqrt{p_0 \cdot q_0}}$
- Các màu trong bảng ứng với các tỷ lệ $p_0 = 0.90; 0.91 \dots 0.99$

**KIỂM TRA NGẪU NHIÊN HÀNG HÓA
XUẤT NHẬP KHẨU TRONG HẢI QUAN (1)**



KIỂM TRA NGẪU NHIÊN HÀNG HÓA XUẤT NHẬP KHẨU TRONG HẢI QUAN (2)

