

TCVN

TIÊU CHUẨN QUỐC GIA

TCVN 10859:2015

ISO 3301:1975

Xuất bản lần 1

**GIẢI THÍCH DỮ LIỆU THỐNG KÊ - SO SÁNH HAI TRUNG
BÌNH TRONG TRƯỜNG HỢP QUAN TRẮC THEO CẶP**

*Statistical interpretation of data -
Comparison of two means in the case of paired observations*

HÀ NỘI - 2015

Lời nói đầu

TCVN 10859:2015 hoàn toàn tương đương với ISO 3301:1975;

TCVN 10859:2015 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.

Giải thích dữ liệu thống kê - So sánh hai trung bình trong trường hợp quan trắc theo cặp

Statistical interpretation of data –

Comparison of two means in the case of paired observations

1 Phạm vi áp dụng

Tiêu chuẩn này quy định phương pháp so sánh trung bình của tổng thể về chênh lệch giữa các cặp quan trắc với giá trị "không" hoặc giá trị bất kỳ được ấn định trước.

2 Thuật ngữ và định nghĩa

Thuật ngữ và định nghĩa sau đây được sử dụng cho tiêu chuẩn này.

2.1

quan trắc cặp (paired observation):

Hai quan trắc x_i và y_i của một tính chất hoặc đặc trưng nào đó được gọi là cặp nếu:

- Chúng được tạo thành từ cùng phần tử i trong tổng thể các phần tử nhưng dưới các điều kiện khác nhau (ví dụ: so sánh kết quả của hai phương pháp phân tích trên cùng một sản phẩm);
- Chúng được tạo thành từ hai phần tử riêng biệt giống nhau về tất cả các khía cạnh trừ khác biệt hệ thống là đối tượng kiểm nghiệm (ví dụ: so sánh năng suất của các mảnh ruộng liền kề được gieo hai giống hạt khác nhau).

Tuy nhiên, cần chú ý trong trường hợp thứ hai, hiệu quả kiểm nghiệm phụ thuộc vào hiệu lực của giả thuyết rằng không có khác biệt hệ thống nào khác giữa các quan trắc trong cùng một cặp ngoài khác biệt hệ thống được kiểm nghiệm.

3 Lĩnh vực áp dụng

Phương pháp này có thể áp dụng để xác định sự khác biệt giữa hai cách xử lý. Trong trường hợp

này, quan trắc x_i được tiến hành sau cách xử lý thứ nhất và y_i tiến hành sau xử lý thứ hai. Hai loạt kết quả quan trắc không độc lập với nhau vì mỗi kết quả x_i của loạt đầu tiên (xử lý thứ nhất) gắn với kết quả y_i của loạt thứ hai (xử lý thứ hai). Thuật ngữ "xử lý" cần được hiểu theo nghĩa rộng. Ví dụ: hai xử lý được so sánh có thể là hai phương pháp kiểm nghiệm, hai phương tiện đo hoặc hai phòng thử nghiệm nhằm phát hiện ra sai số hệ thống có thể có. Hai xử lý được thực hiện liên tiếp trên cùng một đối tượng thực nghiệm có thể ảnh hưởng đến nhau và giá trị thu được có thể phụ thuộc vào thứ tự thực hiện. Thiết kế thực nghiệm tốt cần cho phép loại bỏ độ chệch này. Ngoài ra, có thể chỉ áp dụng một xử lý và so sánh ảnh hưởng của nó với việc không xử lý; khi đó, mục đích của việc so sánh này là xác định ảnh hưởng của cách xử lý đó.

4 Điều kiện áp dụng

Phương pháp này có thể được áp dụng nếu thỏa mãn hai điều kiện sau:

- loạt các hiệu $d_i = x_i - y_i$ có thể được xem là loạt các cá thể ngẫu nhiên độc lập;
- phân bố các hiệu $d_i = x_i - y_i$ giữa các cặp quan trắc được giả định là phân bố chuẩn hoặc xấp xỉ chuẩn.

Nếu phân bố của các hiệu này sai lệch so với phân bố chuẩn thì phương pháp được mô tả vẫn có hiệu lực, với điều kiện cỡ mẫu đủ lớn; sai lệch so với phân bố chuẩn càng cao thì cỡ mẫu yêu cầu càng lớn. Tuy nhiên, ngay trong các trường hợp đặc biệt, cỡ mẫu 100 là đủ đáp ứng cho phần lớn các ứng dụng thực tế.

5 Trình bày chính thức các tính toán

Vấn đề nghiên cứu	
Điều kiện thực nghiệm.....	
Dữ liệu thống kê Cỡ mẫu: $n =$ Tổng các giá trị quan trắc: $\sum x_i = \quad \sum y_i =$ Tổng các hiệu: $\sum d_i =$ Tổng bình phương các hiệu: $\sum d_i^2 =$ Giá trị đã cho: $d_0 =$ Bậc tự do: $v = n - 1$ Mức ý nghĩa được chọn: $\alpha =$	Tính toán $\bar{d} = \frac{1}{n} (\sum x_i - \sum y_i)$ $= \frac{1}{n} \sum d_i =$ $s_d^2 = \frac{1}{n-1} \left[\sum d_i^2 - \frac{1}{n} (\sum d_i)^2 \right] =$ $\sigma_d^* = \sqrt{s_d^2} =$ $A_1 = [t_{1-\alpha}(v)/\sqrt{n}] \sigma_d^* =$ $A_2 = [t_{1-\alpha/2}(v)/\sqrt{n}] \sigma_d^* =$
Kết quả Trường hợp hai phía: Giả thuyết rằng trung bình tổng thể của các hiệu bằng d_0 (giả thuyết không) bị bác bỏ nếu: $ \bar{d} - d_0 > A_2$ Trường hợp một phía: a) Giả thuyết rằng trung bình tổng thể của các hiệu lớn hơn hoặc bằng d_0 (giả thuyết không) bị bác bỏ nếu: $\bar{d} < d_0 - A_1$ b) Giả thuyết rằng trung bình tổng thể của các hiệu nhỏ hơn hoặc bằng d_0 (giả thuyết không) bị bác bỏ nếu: $\bar{d} > d_0 + A_1$	

CHÚ THÍCH: $t_{1-\alpha}(v)$ là phân vị $1 - \alpha$ của biến t của phân bố Student với v bậc tự do. Các giá trị của $t_{1-\alpha}(v)/\sqrt{n}$ được cho trong Bảng 1.

Bảng 1 – Giá trị của tỉ số $t_{1-\alpha}(v)/\sqrt{n}$ đối với $v = n - 1$

$v = n - 1$	Trường hợp hai phía		Trường hợp một phía	
	$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,995}}{\sqrt{n}}$	$\frac{t_{0,95}}{\sqrt{n}}$	$\frac{t_{0,99}}{\sqrt{n}}$
1	8,985	45,013	4,465	22,501
2	2,434	5,730	1,686	4,021
3	1,591	2,920	1,177	2,270
4	1,242	2,059	0,953	1,676
5	1,049	1,646	0,823	1,374
6	0,925	1,401	0,734	1,188
7	0,836	1,237	0,670	1,060
8	0,769	1,118	0,620	0,966
9	0,715	1,028	0,580	0,897
10	0,672	0,956	0,546	0,833
11	0,635	0,897	0,518	0,785
12	0,604	0,847	0,494	0,744
13	0,577	0,806	0,473	0,708
14	0,554	0,769	0,455	0,678
15	0,533	0,737	0,438	0,651
16	0,514	0,708	0,423	0,626
17	0,497	0,683	0,410	0,605
18	0,482	0,660	0,398	0,586
19	0,468	0,640	0,387	0,568
20	0,455	0,621	0,376	0,552
21	0,443	0,604	0,367	0,537
22	0,432	0,588	0,358	0,523
23	0,422	0,573	0,350	0,510
24	0,413	0,559	0,342	0,498
25	0,404	0,547	0,335	0,487
26	0,396	0,535	0,328	0,477
27	0,388	0,524	0,322	0,467
28	0,380	0,513	0,316	0,458
29	0,373	0,503	0,310	0,449
30	0,367	0,494	0,305	0,441
40	0,316	0,422	0,263	0,378
50	0,281	0,375	0,235	0,337
60	0,256	0,341	0,214	0,306
70	0,237	0,314	0,198	0,283
80	0,221	0,293	0,185	0,264
90	0,208	0,276	0,174	0,248
100	0,197	0,261	0,165	0,235
200	0,139	0,183	0,117	0,165
500	0,088	0,116	0,074	0,104
∞	0	0	0	0

VÍ DỤ: Bảng dữ liệu dưới đây được thu thập trong quá trình nghiên cứu để xác định xem liệu tỷ lệ trung bình mài mòn trực gây ra do kim loại làm ổ trục khác nhau trong động cơ đốt trong có khác nhau giữa các kim loại hay không.

Bảng 2 – Mài mòn trực sau thời gian làm việc cho trước theo 0,00001 in

Trục i	Mài mòn với		Hiệu $d_i = x_i - y_i$
	Hợp kim đồng chì x_i	Kim loại trắng y_i	
1	3,5	1,5	2,0
2	2,0	1,3	0,7
3	4,7	4,5	0,2
4	2,8	2,5	0,3
5	6,5	4,5	2,0
6	2,2	1,7	0,5
7	2,5	1,8	0,7
8	5,8	3,3	2,5
9	4,2	2,3	1,9
Tổng	34,2	23,4	10,8

Đặc tính kỹ thuật	
Dữ liệu thống kê Cỡ mẫu: $n = 9$ Tổng các giá trị quan trắc: $\sum x_i = 34,2$ $\sum y_i = 23,4$ Tổng các hiệu: $\sum d_i = 10,8$ Tổng bình phương các hiệu: $\sum d_i^2 = 19,22$ Giá trị đã cho: $d_0 = 0$ Bậc tự do: $v = 8$ Mức ý nghĩa được chọn: $\alpha = 0,01$	Công thức $\bar{d} = \frac{1}{9}(34,2 - 23,4) = 1,2$ $s_d^2 = \frac{1}{8}19,22 - \frac{10,8^2}{9} = 0,7825$ $\sigma_d^* = \sqrt{0,7825} = 0,8846$ $t_{0,995}/\sqrt{9} = 1,118$ $A_2 = 1,118 \times 0,8846 = 0,99$
Kết quả So sánh trung bình tổng thể với giá trị 0 cho trước: Trường hợp hai phía: $ \bar{d} - d_0 = 1,2 > 0,99$ Giả thuyết về sự bằng nhau của tỉ lệ mài mòn trực do hai kim loại bị bác bỏ ở mức 1 %.	

6 Sai lầm loại hai

Xác suất bác bỏ giả thuyết không khi nó đúng nhiều nhất là bằng mức ý nghĩa α . Việc bác bỏ giả thuyết không khi nó đúng được gọi là sai lầm loại một và do đó việc chọn α sẽ giới hạn rủi ro sai lầm như vậy.

Mặt khác, có khả năng phạm phải sai lầm loại hai, đó là, chấp nhận giả thuyết không khi nó sai. Xác suất $1 - \beta$ của việc bác bỏ giả thuyết không khi nó sai được gọi là hiệu lực của phép kiểm nghiệm; do đó xác suất sai lầm loại hai là β .

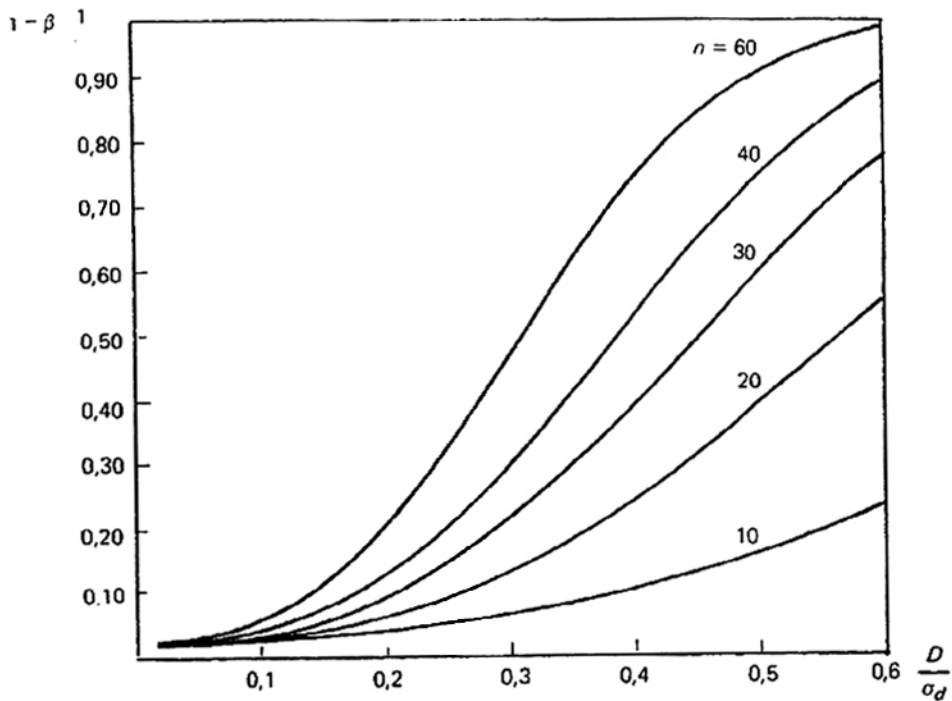
Với mẫu n và sai lầm loại một cho trước, các xác suất nói trên không chỉ phụ thuộc vào trung bình thực D của các hiệu quan trắc được $d_i = x_i - y_i$ mà ta có thể đặt ra các đối giả thuyết khác mà còn phụ thuộc vào độ lệch chuẩn σ_d của các hiệu này. Độ lệch chuẩn này nói chung là chưa biết và nếu n nhỏ thì mẫu sẽ chỉ cung cấp một ước lượng kém.

Kết quả là không thể đặt ra giới hạn trên cho xác suất sai lầm loại hai.

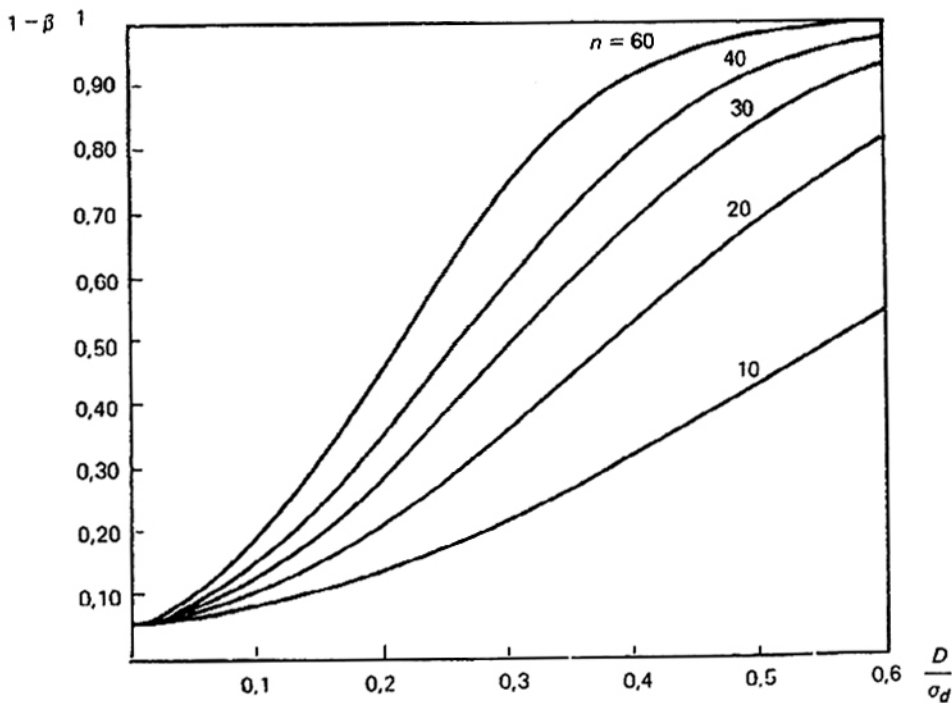
Tuy nhiên, trong các biểu đồ dưới đây mối quan hệ được chỉ ra giữa hiệu lực của phép kiểm nghiệm, $1 - \beta$, và trung bình tổng thể thực tế chia cho độ lệch chuẩn tương ứng, D/σ_d , với phép kiểm nghiệm một phía của giả thuyết $H_0: D \leq 0$, với các giá trị n khác nhau và mức ý nghĩa tương ứng là 0,05 và 0,01.

Từ các biểu đồ này có thể rút ra các kết luận sau:

- 1) Hiệu lực của phép kiểm nghiệm được xác định duy nhất bằng trung bình thực của các hiệu, được đo theo đơn vị của độ lệch chuẩn, theo mức ý nghĩa α và cỡ mẫu.
- 2) Hàm hiệu lực là hàm tăng ngặt của hiệu trung bình thực.
Hàm cũng tăng ngặt với cỡ mẫu và mức ý nghĩa α , với điều kiện $D > 0$ và α khác 0 và khác 1.
- 3) Với mức ý nghĩa 0,05 và cỡ mẫu 50, hiệu lực đạt được ít nhất là 0,95 khi hiệu trung bình thực lớn hơn một nửa độ lệch chuẩn của các hiệu. Khi $n = 20$ sẽ đạt được hiệu lực này với $D/\sigma_d \geq 0,78$.



Hình 1 – Hiệu lực của phép kiểm nghiệm Student một mẫu (một phía), $\alpha = 0,01$



Hình 2 – Hiệu lực của phép kiểm nghiệm Student một mẫu (một phía), $\alpha = 0,05$

CHÚ THÍCH: Các biểu đồ trên dựa trên tài liệu của D.B. OWEN, *Sổ tay bảng thống kê*, Addison Wesley.