

ĐẶNG HÙNG THẮNG

# BÀI TẬP THỐNG KÊ

TT TT-TV \* ĐHQGHN

519.2

ĐA-T

2010

V-G1



NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

ĐẶNG HÙNG THẮNG

# BÀI TẬP THỐNG KÊ

(Dùng cho các trường Đại học và Cao đẳng)

*(Tái bản lần thứ hai)*

NHÀ XUẤT BẢN GIÁO DỤC VIỆT NAM

**Công ty CP Sách Đại học - Dạy nghề – Nhà xuất bản Giáo dục Việt Nam giữ  
quyền công bố tác phẩm.**

---

**19 – 2010/CXB/233 – 2244/GD**

**Mã số : 7B723y0 – DAI**

## *Lời nói đầu*

---

Cuốn sách bài tập này bao gồm 256 bài tập về Thống kê toán học trong chương trình đại học đại cương nhằm trang bị cho người học những kiến thức cơ sở của môn học này. Các bài tập trong cuốn sách chủ yếu là những bài tập cơ bản, giúp bạn đọc rèn luyện kỹ năng toán học, kỹ năng vận dụng lý thuyết vào các tình huống thực tế, phát triển tư duy thống kê.

Cuốn sách bài tập này đi kèm với cuốn sách lý thuyết "Thống kê và ứng dụng" của cùng tác giả. Bởi vậy các bài toán cũng được sắp xếp thành 7 chương tương ứng với 7 chương trong cuốn sách lý thuyết đã nêu.

*Chương I : Thống kê mô tả*

*Chương II : Đại cương về lý thuyết xác suất*

*Chương III : Ước lượng tham số*

*Chương IV : Kiểm định giả thiết thống kê*

*Chương V : Bài toán so sánh*

*Chương VI : Bài toán so sánh mở rộng*

*Chương VII : Phân tích tương quan và hồi quy*

Phần A là các đề toán. Phần B là lời giải gồm hướng dẫn, đáp số hoặc bài giải chi tiết các bài tập đã cho tùy theo mức độ khó dễ của chúng. Chúng tôi hi vọng rằng cuốn sách bài tập này sẽ có ích cho tất cả những ai quan tâm tới việc học tập hay giảng dạy bộ môn Thống kê. Mong rằng việc xuất bản đầy đủ bộ sách lý thuyết

và bài tập Thống kê sẽ góp phần nâng cao chất lượng dạy và học bộ môn quan trọng này trong các trường đại học của nước ta.

Nhân đây chúng tôi xin bày tỏ lời cảm ơn tới Nhà xuất bản Giáo dục đã gợi ý và khuyến khích chúng tôi biên soạn bộ sách Lí thuyết và Bài tập Thống kê này.

Vì khả năng và thời gian có hạn, cuốn sách khó tránh khỏi thiếu sót. Chúng tôi rất mong nhận được sự góp ý phê bình của độc giả. Xin chân thành cảm ơn.

*Hà nội 2/2008*

**Đặng Hùng Thắng**

## *Phần A*

# **CÁC ĐỀ BÀI TẬP**

---

### *Chương 1*

## **THỐNG KÊ MÔ TẢ**

1. Cho dãy số liệu sau đây :

54, 75, 121, 142, 154, 159, 171, 189, 203, 211, 225, 247, 251, 259, 264, 278, 290, 305, 315, 322, 355, 367, 388, 450, 490.

Lập bảng phân bố tần số với 6 khoảng bắt đầu từ (49,5 – 124,5), độ rộng của khoảng là 75.

2. Cho dãy số liệu sau đây :

53, 47, 59, 66, 36, 69, 84, 77, 42, 57, 51, 60, 78, 63, 46, 63, 42, 55, 63, 48, 75, 60, 58, 80, 44, 59, 60, 75, 49, 63.

Lập bảng phân bố tần số với 10 khoảng bắt đầu từ khoảng (35,5 – 40,5), độ rộng của khoảng là 5.

3. Lập bảng phân bố tần suất đối với số liệu của bài tập 1.

4. Một nhóm 25 học sinh có số điểm thi môn ngoại ngữ như sau :

860, 940, 1120, 900, 840, 980, 1050, 1220, 860, 770, 1010, 870, 890, 910, 930, 1040, 1280, 1020, 970, 1330, 890, 980, 1260, 980, 760.

Lập bảng phân bố tần số cho các số liệu này. Sử dụng 6 khoảng chia.

5. Trọng lượng (tính bằng pound : 1 pound = 0,454 kg) của một nhóm 26 người tham gia câu lạc bộ sức khỏe là :

135, 175, 166, 148, 183, 206, 190, 128, 147, 156, 166, 174, 158, 196, 120, 165, 189, 174, 148, 225, 192, 177, 154, 140, 180, 172.

Lập bảng phân bố tần số với 7 khoảng chia.

6. Chiều dài của 20 cây gỗ được ghi lại như sau (đơn vị là mét) :

5,2 ; 5,6 ; 4,9 ; 5,3 ; 5,8 ; 4,8 ; 5,0 ; 5,2 ; 5,4 ; 4,8 ; 4,4 ; 5,1 ; 5,5 ; 4,9 ; 5,2 ; 5,7 ; 5,0 ; 5,3 ; 4,9 ; 4,8.

Lập bảng phân bố tần suất với 5 khoảng chia.

7. Vẽ tổ chức đồ tần số cho dãy số liệu sau đây :

54, 75, 121, 142, 154, 159, 171, 189, 203, 211, 225, 247, 251, 259, 264, 278, 290, 305, 315, 322, 355, 367, 388, 450, 490.

Hãy sử dụng 6 khoảng với độ dài là 75, bắt đầu từ (49,5 – 124,5).

8. Vẽ tổ chức đồ tần số cho dãy số liệu sau đây :

860, 940, 1120, 900, 840, 980, 1050, 1220, 860, 770, 1010, 870, 890, 910, 930, 1040, 1280, 1020, 970, 1330, 890, 980, 1260, 980, 760.

Sử dụng 7 khoảng với độ dài 82 bắt đầu từ (759,5 – 841,5).

9. Vẽ tổ chức đồ tần số cho bảng số liệu sau đây ghi lại kết quả thi của một lớp gồm 131 học sinh :

Khoảng	Tần số
0,5 – 20,5	14
20,5 – 30,5	9
30,5 – 40,5	15
40,5 – 45,5	11
45,5 – 50,5	18
50,5 – 55,5	14
55,5 – 60,5	10
60,5 – 70,5	16
70,5 – 100,5	24
Tổng số	131

Sử dụng tổ chức đồ này, hãy ước lượng số học sinh có điểm nằm trong đoạn [43 ; 54].

10. Vẽ tổ chức đồ tần số cho bảng phân bố tần số sau đây :

Khoảng	Tần số
49,5 – 79,5	18
79,5 – 109,5	24
109,5 – 129,5	23
129,5 – 149,5	33
149,5 – 179,5	27
Tổng số	125

11. Tuổi của dân cư trong một vùng A nào đó được thống kê trong bảng phân bố tần số dưới đây :

Khoảng	Tần số
0,5 – 9,5	440
9,5 – 19,5	480
19,5 – 34,5	630
34,5 – 54,5	440
54,5 – 79,5	150
Tổng số	2140

Hãy vẽ tổ chức đồ tần suất.

12. Tuổi của dân cư trong một vùng B nào đó được thống kê lại trong bảng phân bố tần số dưới đây :

Khoảng	Tần số
0,5 – 3,5	54
3,5 – 23,5	180
23,5 – 38,5	291
38,5 – 48,5	315
48,5 – 58,5	360
58,5 – 73,5	384
73,5 – 88,5	90
Tổng số	1074

Vẽ tổ chức đồ tần suất. So sánh cơ cấu dân cư của vùng B với cơ cấu dân cư vùng A (bài tập 11).



**13.** Tính giá trị trung bình, median (số trung vị) và mod của dãy số liệu sau đây :

53, 47, 59, 66, 36, 36, 69, 84, 77, 42, 57, 51, 60, 78, 63, 46, 63, 42, 55, 63, 48, 75, 60, 58, 80, 44, 59, 60, 75, 49, 63.

**14.** Tính giá trị trung bình, median (số trung vị) và mod của dãy số liệu sau đây :

860, 940, 1120, 900, 840, 1050, 1220, 860, 770, 1010, 870, 890, 910, 930, 1040, 1280, 1020, 970, 1330, 890, 980, 1260, 980, 760.

**15.** Cho bảng số liệu sau đây :

Khoảng	Tần số
0 – 10	7
10 – 20	20
20 – 30	15
30 – 50	20
50 – 70	10
70 – 100	9

Tìm giá trị trung bình, khoảng mod và median (số trung vị).

**16.** Cho bảng phân bố tần số sau :

Khoảng	Tần số
10 – 14	5
14 – 16	10
16 – 17	11
17 – 18	22
18 – 19	22
19 – 20	10
20 – 21	11
21 – 22	21
22 – 23	44
23 – 24	34
24 – 30	10

Tìm giá trị trung bình, khoảng mod và median (số trung vị).

17. Tính điểm thi trung bình của 61 sinh viên từ bảng thống kê sau đây :

Khoảng	Tần số
9,5 – 19,5	2
19,5 – 29,5	5
29,5 – 39,5	16
39,5 – 49,5	8
49,5 – 59,5	13
59,5 – 69,5	11
69,5 – 79,5	5
79,5 – 89,5	1

18. Tính tuổi trung bình của 228 người từ bảng thống kê sau đây :

Khoảng	Tần số
16 – 19	10
19 – 22	13
22 – 25	22
25 – 28	38
28 – 31	35
31 – 34	37
34 – 37	28
37 – 40	20
40 – 43	18
43 – 46	7

19. Giá tiền của 80 căn hộ trong vùng A nào đó được thống kê như sau (đơn vị là triệu) :

Khoảng	Tần số
79,5 – 84,5	5
84,5 – 89,5	10
89,5 – 94,5	15
94,5 – 99,5	26
99,5 – 104,5	13
104,5 – 109,5	7
109,5 – 114,5	4

Tìm median (số trung vị).

**20.** Cho bảng phân bố tần số sau đây :

Khoảng	Tần số
19,5 – 24,5	2
24,5 – 29,5	7
29,5 – 34,5	15
34,5 – 39,5	8
39,5 – 44,5	3
Tổng số	35

Tính giá trị trung bình và median (số trung vị).

**21.** Cho dãy số liệu sau

123 132 130 119 106 97  
 121 109 118 128 132 115 130  
 125 126 127 144 115 107 110  
 112 118 115 134 132 139 144  
 104 128 138 114 121 129 128  
 116 138 129 113 105 142 122  
 131 126 111 142

a) Lập bảng phân bố tần số với 6 khoảng.

b) Tìm giá trị trung bình.

c) Tìm median (số trung vị).

**22.** Tuổi của 70 nhân viên trong một cơ quan được ghi lại như sau :

Khoảng	Tần số
Dưới 20	7
20 – 30	13
30 – 40	26
40 – 50	15
50 – 60	6
Trên 60	3

Tìm median (số trung vị) và lớp mod.

**23. Cho dãy số liệu sau đây**

$x_i$	20	21	22	23	24	25
Tần số	5	8	11	10	5	1

Tìm median và giá trị trung bình của dãy số liệu này.

**24. Tỉ lệ tử vong của những người mắc bệnh A và người mắc bệnh B trong những năm 1983 – 1992 được thống kê trong bảng sau đây :**

Năm	Bệnh A (%)	Bệnh B (%)
1983	49,1	26,0
1984	36,2	20,0
1985	23,3	12,3
1986	21,7	11,7
1987	21,5	11,2
1988	21,1	10,8
1989	19,8	10,5
1990	19,9	10,6
1991	20,0	10,0
1992	19,6	9,6

a) Tính biên độ và độ lệch trung bình của hai nhóm số liệu.

b) Có sự khác nhau về mức độ phân tán giữa hai nhóm số liệu hay không ?

**25. Cho dãy số liệu sau :**

9,75    3,75    5,50    7,30    6,00  
7,30    7,25    4,25    3,30    6,60

Tính phương sai và độ lệch tiêu chuẩn mẫu.

**26. Tính giá trị trung bình và độ lệch tiêu chuẩn từ bảng số liệu sau :**

X	114	115	116	117	118	119
Tần số	21	57	111	78	45	18

Tính giá trị trung bình và độ lệch tiêu chuẩn của các số liệu cho dưới dạng bảng phân bố trung vị trong các bài 27, 28, 29, 30 và 31.

27.

Khoảng	Tần số
0 – 10	5
10 – 20	13
20 – 30	24
30 – 40	35
40 – 50	40
50 – 60	36
60 – 70	22
70 – 80	16
80 – 90	9

28.

Khoảng	Tần số
10 – 15	7
15 – 20	21
20 – 25	38
25 – 30	46
30 – 35	50
35 – 40	54
40 – 45	18
45 – 50	2

29.

Khoảng	Tần số
0,5 – 3,5	24
3,5 – 6,5	22
6,5 – 9,5	8
9,5 – 12,5	4
12,5 – 15,5	1
15,5 – 18,5	0
18,5 – 21,5	1

30.

Khoảng	Tần số
27,5 – 32,5	18
32,5 – 37,5	76
37,5 – 42,5	200
42,5 – 47,5	100
47,5 – 52,5	6

31.

Khoảng	Tần số
99,5 – 199,5	47
199,5 – 299,5	139
299,5 – 399,5	328
399,5 – 499,5	555
499,5 – 599,5	419
599,5 – 699,5	280
699,5 – 799,5	311
799,5 – 899,5	107

## Chương 2

# ĐẠI CƯƠNG VỀ LÝ THUYẾT XÁC SUẤT

**32.** Gieo đồng thời hai con xúc sắc. Tính xác suất để :

- a) Tổng số nốt xuất hiện trên hai con là 7 ;
- b) Tổng số nốt xuất hiện trên hai con là 8 ;
- c) Số nốt xuất hiện trên hai con hơn kém nhau 2.

**33.** Gieo đồng thời ba con xúc sắc. Tính xác suất để :

- a) Tổng số nốt xuất hiện của ba con là 8 ;
- b) Tổng số nốt xuất hiện của ba con là 11.

**34.** Một chiếc hộp đựng 6 quả cầu trắng, 4 quả cầu đỏ và 2 quả cầu đen. Chọn ngẫu nhiên 6 quả cầu. Tính xác suất để ta thu được 3 quả trắng, 2 quả đỏ và 1 quả đen.

**35.** Có 30 tấm thẻ đánh số từ 1 đến 30. Chọn ngẫu nhiên 10 tấm thẻ. Tính xác suất để :

- a) Tất cả 10 tấm thẻ đều mang số chẵn ;
- b) Có đúng 5 tấm thẻ mang số chia hết cho 3 ;
- c) Có 5 tấm thẻ mang số lẻ, 5 tấm thẻ mang số chẵn trong đó chỉ có 1 số chia hết cho 10.

**36.** Trong một hòm đựng 100 thẻ đánh số từ 1 đến 100. Chọn ngẫu nhiên 3 tấm thẻ rồi cộng các số ghi trên ba tấm thẻ ấy lại. Tính xác suất để ta thu được một số chia hết cho 3.

**37.** Chọn ngẫu nhiên một số trong 9 số 1, 2, ..., 9. Tính xác suất để sau  $n$  lần chọn ( $n > 1$ ) có hoàn lại thì tích các số thu được là một số chia hết cho 10.

**38.** Một công ti cần tuyển hai nhân viên. Có 6 người nộp đơn trong đó có 4 nữ và 2 nam. Khả năng được tuyển của mỗi người là như nhau.

a) Tính xác suất để cả hai nữ đều được chọn biết rằng ít nhất một nữ đã được chọn.

b) Giả sử Hoa là một trong 4 nữ. Tính xác suất để Hoa được chọn.

c) Tính xác suất để Hoa được chọn nếu biết rằng ít nhất 1 nữ đã được chọn.

**39.** Một đoàn tàu có 4 toa đỗ ở một sân ga. Có 4 hành khách từ sân ga lên tàu, mỗi người độc lập với nhau chọn ngẫu nhiên một toa. Tính xác suất để 1 toa có 3 người, 1 toa có 1 người và hai toa còn lại không có ai.

**40.** Chọn ngẫu nhiên một vé xổ số có 5 chữ số. Tính xác suất để số vé không có chữ số 1 hoặc không có chữ số 5.

**41.** Một đoàn tàu có 3 toa đỗ ở sân ga. Có 5 hành khách lên tàu. Mỗi hành khách độc lập với nhau chọn ngẫu nhiên một toa. Tính xác suất để mỗi toa đều có ít nhất một hành khách mới lên.

**42.** Một người bỏ ngẫu nhiên 4 lá thư vào 4 chiếc phong bì đã ghi địa chỉ. Tính xác suất để có ít nhất một lá thư bỏ đúng phong bì của nó.

**43.** Một sọt cam rất lớn được phân loại theo cách sau : Chọn ngẫu nhiên 20 quả cam để xem xét. Nếu trong đó không có quả cam nào hỏng thì sọt cam được xếp loại 1. Nếu trong đó có một hoặc hai quả hỏng thì sọt cam được xếp loại 2. Trong trường hợp còn lại (có từ ba quả hỏng trở lên) thì sọt cam được xếp loại 3.

Giả sử tỉ lệ cam hỏng trong sọt cam là 3%. Hãy tính xác suất để :

a) Sọt cam được xếp loại 1 ;

b) Sọt cam được xếp loại 2 ;

c) Sọt cam được xếp loại 3.



**44.** Một bài thi trắc nghiệm (*multiple choice test*) gồm 12 câu hỏi, mỗi câu hỏi cho 5 câu trả lời, trong đó chỉ có một câu đúng. Giả sử mỗi câu trả lời đúng được 4 điểm, mỗi câu trả lời sai bị trừ 1 điểm. Một học sinh kém làm bài bằng cách chọn hù họa một câu trả lời. Tính xác suất để :

- a) Anh ta được 13 điểm ;
- b) Anh ta bị điểm âm.

**45.** Một trò chơi bao gồm việc gieo đồng thời ba con xúc sắc. Người chơi thắng cuộc nếu có xuất hiện ít nhất 2 "lục". Tính xác suất để trong 5 ván chơi người chơi thắng ít nhất 3 ván.

**46.** Một người say rượu bước 8 bước. Mỗi bước anh ta tiến lên phía trước một nửa mét hoặc lùi lại phía sau một nửa mét với xác suất như nhau. Tính xác suất để sau 8 bước :

- a) Anh ta trở lại điểm xuất phát ;
- b) Anh ta cách điểm xuất phát hơn 2 mét.

**47.** Để chuẩn bị thi một sinh viên được cho hai tập câu hỏi để ôn tập. Mỗi tập gồm 10 câu. Giả sử trước khi thi anh ta học thuộc 9 câu của tập 1 và 8 câu của tập 2. Đề thi gồm 3 câu được thiết kế như sau : Chọn ngẫu nhiên một câu hỏi rồi từ đó chọn ngẫu nhiên hai câu. Câu thứ ba chọn ngẫu nhiên từ tập còn lại.

- a) Tính xác suất để sinh viên chỉ trả lời được một câu.
- b) Tính xác suất để sinh viên trả lời được hai câu.

**48.** Một cuộc thi có ba vòng. Vòng 1 lấy 90% thí sinh. Vòng 2 lấy 80% thí sinh vượt qua vòng 1 và vòng 3 lấy 90% thí sinh vượt qua vòng 2.

- a) Tính xác suất để một thí sinh lọt qua cả ba vòng thi.
- b) Tính xác suất để một thí sinh bị loại ở vòng 2 nếu biết rằng thí sinh đó bị loại.

49. Có hai chuồng thỏ. Chuồng thứ nhất có 5 con thỏ đen và 10 con thỏ trắng. Chuồng thứ hai có 3 con thỏ trắng và 7 con thỏ đen. Từ chuồng thứ hai ta bắt ngẫu nhiên một con thỏ cho vào chuồng thứ nhất, rồi sau đó bắt ngẫu nhiên một con thỏ từ chuồng thứ nhất ra thì được một thỏ trắng. Tính xác suất để con thỏ trắng này lúc đầu ở chuồng thứ nhất.

50. Có ba người đi săn. Họ cùng trông thấy một con nai và cùng nổ súng. Con nai bị trúng một viên đạn và chết. Nên chia phần theo tỉ lệ thế nào là công bằng ? Biết rằng xác suất bắn trúng của người thứ nhất là 0,3 ; của người thứ hai là 0,4 và của người thứ ba là 0,5.

51. Một mạng điện có 4 bóng đèn 1, 2, 3 và 4 với xác suất hỏng tương ứng là 0,1 ; 0,2 ; 0,3 và 0,4. Biết rằng có hai bóng bị hỏng. Tính xác suất để đó là hai bóng 1 và 2.

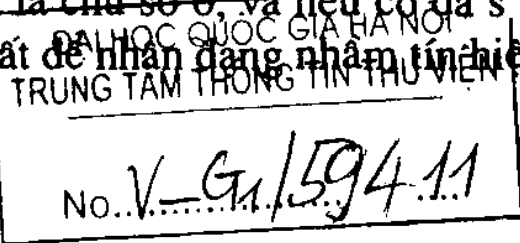
52. Biết rằng một người có nhóm máu AB có thể nhận máu của bất kì nhóm máu nào. Nếu người đó có nhóm máu còn lại (A hoặc B hoặc O) thì chỉ có thể nhận máu của người cùng nhóm máu với mình hoặc người có nhóm máu O.

Cho biết tỉ lệ người có nhóm máu O, A, B và AB tương ứng là 33,7% ; 37,5% ; 20,9% và 7,9%.

a) Chọn ngẫu nhiên một người cần tiếp máu và một người cho máu. Tính xác suất để sự truyền máu thực hiện được.

b) Chọn ngẫu nhiên một người cần tiếp máu và hai người cho máu. Tính xác suất để sự truyền máu thực hiện được.

53. Trên một kênh liên lạc, người ta truyền đi hai tín hiệu 0 và 1. Vì tiếng ồn gây nhiễu nên máy thu có thể nhận dạng nhầm tín hiệu với xác suất  $q = 0,2$ . Để tăng độ tin cậy, thay vì truyền chữ số 0 ta truyền một cụm 5 chữ số 0 (00000) và truyền chữ số 1 bằng một cụm 5 chữ số 1 (11111). Nếu khi nhận được một cụm 5 tín hiệu có đa số là 0 thì ta coi là chữ số 0, và nếu có đa số là 1 thì ta coi là chữ số 1. Tính xác suất để nhận dạng nhầm tín hiệu.



**54.** Một đại lí có 4 chiếc ô tô cho thuê. Số khách có nhu cầu thuê trong một ngày là một ĐLNN có phân bố Poisson với tham số  $\lambda = 2$ .

a) Tính số ô tô trung bình mà đại lí cho thuê trong một ngày.

b) Đại lí cần có ít nhất bao nhiêu chiếc xe ô tô để đáp ứng được nhu cầu với xác suất ít nhất là 0,98 ?

**55.** Số hoa mọc trong một chậu cây cảnh là một ĐLNN có phân bố Poisson với tham số  $\lambda = 3$ . Người ta chỉ đem bán các chậu cây với số hoa là 2, 3, 4 hoặc 5. .

a) Hỏi trong số các chậu cây đem bán có bao nhiêu phần trăm cây có 2 hoa ? 3 hoa ? 4 hoa ? 5 hoa ?

b) Tính số hoa trung bình của các chậu hoa đem bán.

**56.** Một nhà nghỉ có 1000 khách. Nhà ăn phục vụ bữa trưa làm hai đợt liên tiếp. Số chỗ ngồi của nhà ăn phải ít nhất là bao nhiêu để xác suất của biến cố : "không đủ chỗ cho người đến ăn" nhỏ hơn 1% ?

**57.** Cũng bài toán trên nhưng giả thiết nhà nghỉ có 500 cặp vợ chồng (mọi cặp vợ chồng luôn đi ăn cùng với nhau).

**58.** Một kì thi gồm 45 câu hỏi, với mỗi câu hỏi thí sinh cần chọn một trong 4 câu trả lời kèm theo, trong đó chỉ có duy nhất một câu trả lời đúng. Một sinh viên hoàn toàn không học gì khi đi thi chọn ngẫu nhiên một trong 4 câu trả lời đối với mỗi câu hỏi. Tính xác suất để :

a) Sinh viên đó trả lời đúng ít nhất 16 câu hỏi ;

b) Sinh viên đó trả lời đúng nhiều nhất 9 câu ;

c) Số câu trả lời đúng từ 8 đến 12.

**59.** Xác suất làm ra phế phẩm của một nhà máy là 0,02. Trong một lô hàng gồm 25000 sản phẩm hãy ước lượng số phế phẩm với

độ tin cậy 95%. Cụ thể là hãy tìm các hằng số  $a$ ,  $b$  để với xác suất 0,95 ta có thể tin chắc rằng số phế phẩm nằm trong khoảng  $(a, b)$ .

**60.** Một trường đại học có chỉ tiêu tuyển sinh là 300.

a) Giả sử có 325 người dự thi với xác suất thi đỗ của mỗi người là 90%. Tính xác suất để số người trúng tuyển nhiều hơn chỉ tiêu.

b) Cần cho phép tối đa bao nhiêu người đăng kí dự thi (giả sử xác suất đỗ của mỗi người vẫn là 90%) để khả năng số người trúng tuyển nhiều hơn chỉ tiêu chỉ xảy ra với xác suất không quá 1% ?

**61.** Thời gian làm việc của một linh kiện điện tử máy tính là một ĐLNN có trung bình là 250 (giờ) và độ lệch tiêu chuẩn là 250 (giờ).

a) Giả sử ta dự trữ 30 linh kiện. Tính xác suất để 30 linh kiện này đủ dùng trong ít nhất 1 năm (8760 giờ).

b) Phải dự trữ ít nhất là bao nhiêu linh kiện để với xác suất 0,99 ta có thể đảm bảo cho máy tính hoạt động ít nhất 1 năm ?

### Chương 3

## ƯỚC LƯỢNG THAM SỐ

**62.** Xây dựng khoảng tin cậy 95% cho tiền lương trung bình trong 1 giờ của công nhân ngành thép ở Mỹ, biết rằng khi khảo sát một mẫu gồm 50 công nhân ta thu được  $\bar{X} = 14,75$  (USD) với độ lệch tiêu chuẩn  $S = 3$  (USD).

**63.** Một mẫu ngẫu nhiên gồm 80 công nhân ở tỉnh A được hỏi về quãng đường họ đi từ nhà tới chỗ làm việc. Giá trị trung bình và độ lệch tiêu chuẩn của mẫu này tương ứng là 7,52 km và 1 km. Hãy tìm khoảng tin cậy 95% cho quãng đường trung bình đi làm của tất cả các công nhân của tỉnh A.

**64.** Để xác định chiều cao trung bình của các cây bạch đàn trong một khu rừng bạch đàn rất lớn, người ta chọn ngẫu nhiên 35 cây để đo. Kết quả thu được như sau :

Khoảng	Tần số
6,5 – 7	2
7 – 7,5	4
7,5 – 8	10
8 – 8,5	11
8,5 – 9	5
9 – 9,5	3

Tìm khoảng tin cậy 95% cho chiều cao trung bình của tất cả các cây bạch đàn.

**65.** Một trường đại học tiến hành một cuộc điều tra xem trung bình một sinh viên của trường tiêu hết bao nhiêu tiền gọi điện

thoại trong một học kì. Một mẫu ngẫu nhiên gồm 59 sinh viên được chọn và số tiền chi cho việc gọi điện thoại của họ như sau (đơn vị : nghìn) :

14	18	22	30	36	28	42	79
36	52	15	47	95	16	27	111
37	63	127	23	31	70	27	11
30	147	72	37	25	7	33	29
35	41	48	15	29	73	26	15
26	31	57	40	18	85	28	
32	22	37	60	41	35	26	20
58	33	23	35				

Hãy xây dựng khoảng tin cậy 95% cho số tiền gọi điện thoại trung bình  $\mu$  hàng tháng của một sinh viên.

66. Để xác định trọng lượng trung bình của các bao bột mì được đóng bằng máy tự động, người ta chọn ngẫu nhiên 15 bao và tính được  $\bar{x} = 39,8$  kg và  $s^2 = 0,144$ .

Tìm khoảng tin cậy 99% của trọng lượng trung bình  $\mu$  của các bao bột.

67. Hai nhóm chuột nuôi trong phòng thí nghiệm được cho ăn theo hai chế độ ăn khác nhau. Ở nhóm đối chứng cho ăn chế độ ăn bình thường, còn ở nhóm thí nghiệm cho ăn thêm một lượng coban. Sự tăng trọng của hai nhóm sau một thời gian được ghi lại như sau (mg) :

*Nhóm đối chứng :* 504 560 420 600 580 530  
490 580 470

*Nhóm thí nghiệm :* 580 692 700 621 640 561  
680 630

Tìm khoảng tin cậy 95% cho sự tăng trọng trung bình của mỗi nhóm.

**68.** Để khảo sát chiều cao trung bình  $\mu$  của thanh niên trong một vùng A nào đó, một mẫu ngẫu nhiên gồm 16 thanh niên được chọn. Chiều cao của các thanh niên này đo được như sau (đơn vị cm) :

172   173   173   174   174   175   176

166   166   167   165   173   171   170

171   170

Tìm khoảng tin cậy 95% cho  $\mu$ .

**69.** Một phương pháp điều trị bệnh mới đang được xem xét nghiệm thu. Một chỉ tiêu đánh giá tính hiệu quả của phương pháp là số ngày trung bình  $\mu$  từ lúc điều trị cho tới khi bệnh nhân khỏi bệnh. Một mẫu ngẫu nhiên gồm 11 bệnh nhân được theo dõi và số ngày điều trị cho tới khi khỏi bệnh được ghi lại như sau :

4, 4, 3, 8, 5, 6, 7, 12, 5, 3, 8

Tìm khoảng tin cậy 95% cho  $\mu$ .

**70.** Trong một cuộc khảo sát 64 khách hàng ở một tiệm ăn nhanh, thời gian đợi trung bình là 3 phút và độ lệch tiêu chuẩn là 1,5 phút.

Tìm khoảng tin cậy 98% cho thời gian đợi trung bình ở tiệm ăn này.

**71.** Trong một cuộc khảo sát thị trường của công ty sản xuất thuốc lá, 150 người nghiện thuốc lá được chọn ngẫu nhiên. Số điếu thuốc hút trung bình trong 1 tuần của nhóm người này là 97 với độ lệch tiêu chuẩn là 36. Tìm khoảng tin cậy 99% cho số điếu thuốc hút trung bình trong 1 tuần của tất cả những người nghiện thuốc.

**72.** Một cuộc khảo sát 50 em ở lứa tuổi là 6 cho thấy số giờ xem tivi trung bình trong một tuần của các em này là 38 giờ với độ lệch tiêu chuẩn là 6,4 giờ. Tìm khoảng tin cậy 99% cho thời gian xem tivi trung bình trong 1 tuần của các em nhỏ 6 tuổi.

73. Một công ty lớn muốn ước lượng trung bình một ngày một thư kí phải đánh máy bao nhiêu trang giấy. Khảo sát 50 thư kí được chọn ngẫu nhiên cho thấy số trang trung bình họ đánh máy trong một ngày là 32 với độ lệch tiêu chuẩn bằng 6. Tìm khoảng tin cậy 90% cho số trang đánh máy trung bình của một thư kí trong công ty.

74. Khảo sát một mẫu gồm 12 người cho thấy số lần họ đi xem phim trong 1 năm như sau :

14, 16, 17, 17, 24, 20, 32, 18, 29, 31, 15, 35.

Tìm khoảng tin cậy 95% cho số lần trung bình mà một người tới rạp chiếu bóng trong 1 năm.

75. Một nhà sưu tập tem khảo giá chiếc tem A trong 9 cửa hàng thì thấy giá trung bình của chiếc tem đó là 17 USD với độ lệch tiêu chuẩn là 3 USD. Tìm khoảng tin cậy 90% cho giá tiền trung bình của chiếc tem này trong tất cả các cửa hàng bán tem.

76. Chọn ngẫu nhiên 12 lớp trung học trong một thành phố A nào đó, ta tính được số học sinh trung bình trong một lớp là 28 với độ lệch tiêu chuẩn là 5. Tìm khoảng tin cậy 99% cho số học sinh trung bình trong một lớp của thành phố A đó.

77. Khảo sát 18 giám đốc các công ty ở Mỹ cho thấy lương trung bình hàng năm của họ là 275000 USD với độ lệch tiêu chuẩn là 62000 USD. Tìm khoảng tin cậy 90% cho mức lương trung bình hàng năm của các giám đốc công ty ở Mỹ.

78. Một người dự định mua một chiếc radio đi khảo giá loại radio này ở 8 cửa hàng. Giá bán ở 8 cửa hàng đó như sau (đơn vị USD) : 138, 149, 129, 135, 145, 125, 139, 142. Tìm khoảng tin cậy 90% cho giá của loại radio này.

79. Trong một tháng, ở một sân bay nọ có 80 chuyến bay hạ cánh. trong số đó có 74 chuyến hạ cánh đúng giờ. Tìm khoảng tin cậy 99% cho tỉ lệ chuyến bay hạ cánh đúng giờ tại sân bay đó.



**80.** Trong số 500 người mua xe máy ở một cửa hàng nọ có 325 người mua xe Honda. Tìm khoảng tin cậy 99% cho tỉ lệ người mua xe Honda trong số những người dùng xe máy.

**81.** Trong số 40 bệnh nhân dùng thuốc A có 32 người thấy có kết quả tốt.

a) Tìm khoảng tin cậy 95% cho tỉ lệ bệnh nhân dùng thuốc A có kết quả tốt.

b) Giả sử rằng trong số 400 bệnh nhân dùng thuốc A thì thấy 320 người có kết quả tốt. Tìm khoảng tin cậy 95% cho tỉ lệ bệnh nhân dùng thuốc A có kết quả tốt.

**82.** Trước ngày bầu cử tổng thống, một cuộc thăm dò dư luận đã được tiến hành. người ta chọn ngẫu nhiên 100 người để hỏi ý kiến thì 60 người nói rằng họ sẽ bỏ phiếu cho ông A. Tìm khoảng tin cậy cho tỉ lệ cử tri bỏ phiếu cho ông A với độ tin cậy 90%.

**83.** Trong một nhóm gồm 200 người có xe máy được chọn ngẫu nhiên, ta thấy có 162 người dùng xe phân khối lớn (từ 100 phân khối trở lên). Tìm khoảng tin cậy 95% cho tỉ lệ những người dùng xe phân khối lớn.

**84.** Trong một cuộc thăm dò ý kiến 100 khách hàng, người ta thấy 55 người nói rằng ưa thích mặt hàng A. Tìm khoảng tin cậy 90% cho tỉ lệ người tiêu dùng ưa thích mặt hàng A.

**85.** Cơ quan cảnh sát giao thông kiểm tra hệ thống phanh của 40 chiếc xe tải trên đường quốc lộ. Họ phát hiện ra 14 chiếc có phanh chưa đảm bảo an toàn.

a) Tìm khoảng tin cậy 98% cho tỉ lệ xe tải có phanh chưa an toàn.

b) Tìm khoảng tin cậy 98% có tỉ lệ xe tải có phanh tốt.

**86.** Chọn ngẫu nhiên 200 sinh viên ta thấy có 37% số em không có trong kí thức xá. Tìm khoảng tin cậy 90% cho tỉ lệ sinh viên ngoại trú.

87. Một cuộc điều tra trong 2074 gia đình trí thức cho thấy 373 gia đình có máy vi tính ở nhà. Tìm khoảng tin cậy 96% cho tỉ lệ những gia đình trí thức có máy vi tính ở nhà.

88. Ở cây ngọc trâm thì đặc tính lá phẳng hay lá nhăn là do một gen có hai alen, A trội và a lặn, quyết định. Các đồng hợp tử AA và dị hợp tử Aa có lá phẳng, còn đồng hợp tử aa có lá nhăn.

Trong số 560 cây có được khi lai hai dị hợp tử thì có 110 cây lá nhăn. Tìm khoảng tin cậy 95% có xác suất  $p$  để có cây lá nhăn khi lai hai dị hợp tử. Số liệu trên có phù hợp với lý thuyết của Mendel hay không (lý thuyết của Mendel cho rằng  $p = \frac{1}{4}$ ) ?

89. Trước ngày bầu cử, để biết tỉ lệ phần trăm  $p$  các cử tri đang còn do dự chưa biết bỏ phiếu cho ai, người ta hỏi ý kiến  $n$  cử tri được chọn một cách ngẫu nhiên.

Số  $n$  phải là bao nhiêu để khoảng tin cậy 95% có độ dài không vượt quá 0,04 ?

90. Ta muốn xây dựng một khoảng tin cậy 95% cho trọng lượng trung bình của các gói đường đóng bằng máy tự động. Một mẫu điều tra sơ bộ cho ta  $\bar{x} = 11,8$  kg, với độ lệch tiêu chuẩn là  $s = 0,7$  kg. Hỏi cần phải lấy kích thước mẫu tối thiểu là bao nhiêu để đạt được sai số không vượt quá 0,2 kg ?

91. Một cuộc nghiên cứu được tiến hành nhằm xác định lượng trung bình các luật sư giỏi ở Mỹ dựa trên một mẫu điều tra. Hỏi cần lấy mẫu với kích thước tối thiểu là bao nhiêu để sai số không vượt quá 100 USD, với độ tin cậy ấn định là 95% ? Với độ tin cậy 99% thì kích thước mẫu phải là bao nhiêu ? Biết rằng độ lệch tiêu chuẩn của tập hợp chính là 1000 USD.

92. Người ta muốn xây dựng khoảng tin cậy 90% cho chiều cao trung bình  $\mu$  của người trưởng thành dựa trên một mẫu với kích thước  $n$ . Hãy xác định  $n$  để sai số giữa trung bình mẫu và trung

biên của tập hợp chính không quá 0,5 inches (1 inch = 2,54 cm). Biết rằng độ lệch tiêu chuẩn của tập hợp chính là 0,3 inches.

**93.** Người ta muốn xây dựng một khoảng tin cậy 95% cho vận tốc của ô tô trên đường cao tốc với sai số không vượt quá 2 dặm. Hỏi cần lấy kích thước mẫu tối thiểu là bao nhiêu để đạt được yêu cầu đề ra ? Một mẫu điều tra sơ bộ cho ta độ lệch tiêu chuẩn mẫu là  $s = 9$  dặm.

**94.** Một nhà nông học muốn ước lượng tỉ lệ nảy mầm của một loại hạt giống.

a) Với 1000 hạt đem gieo thì có 640 hạt nảy mầm. Tìm khoảng tin cậy 90% cho tỉ lệ nảy mầm  $p$ . Sai số ở đây là bao nhiêu ?

b) Nếu muốn có một khoảng tin cậy 90% với sai số không vượt quá 0,02 thì cần lấy mẫu với kích thước là bao nhiêu ?

**95.** Người ta muốn tìm khoảng tin cậy 90% cho điểm trung bình của học sinh trong kì thi tốt nghiệp với sai số không vượt quá 0,2. Kích thước mẫu cần lấy là bao nhiêu ? Giả sử rằng một mẫu điều tra sơ bộ đã cho ta độ lệch tiêu chuẩn  $s \approx 1,2$ .

**96.** Người ta muốn tìm khoảng tin cậy 95% cho tỉ lệ những gia đình có máy giặt với độ chính xác 0,04. Hỏi kích thước mẫu cần lấy là bao nhiêu ? Giả sử rằng một mẫu điều tra sơ bộ cho ta tần suất  $f = 0,72$ .

**97.** Một khách sạn lớn tiến hành một nghiên cứu để xác định tỉ lệ phần trăm các khách trọ với thời gian nhiều hơn 1 ngày. Người chủ khách sạn muốn đạt được độ tin cậy 95% và sai số không vượt quá 0,05. Anh ta ước lượng sơ bộ tỉ lệ này khoảng 30%. Hỏi cần lấy mẫu với kích thước bao nhiêu ?

**98.** Tương tự như bài toán 97 nhưng ở đây người chủ khách sạn không có trước một chút thông tin gì về tỉ lệ cần ước lượng.

## Chương 4

# KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ

**99.** Những thống kê trong năm ngoái cho thấy một người Mỹ đi du lịch ở châu Âu trong vòng 3 tuần sẽ chi hết trung bình 1010 USD cho việc mua sắm. Một cuộc nghiên cứu dự định tiến hành cuối năm nay để xác định xem có sự thay đổi gì trong việc tiêu tiền mua sắm khi du lịch của người Mỹ hay không.

a) Hãy phát biểu giả thiết  $H_0$  và đối thiết  $H_1$ .

b) Giả sử thống kê trên 50 khách du lịch cho thấy số tiền trung bình họ tiêu là 1090 USD, với độ lệch tiêu chuẩn là 300 USD. Với mức ý nghĩa  $\alpha = 0,01$ , hãy nhận định xem có sự thay đổi về sự tiêu tiền khi du lịch ở châu Âu của người Mỹ hay không.

c) Tính P – giá trị.

**100.** Một nhà nghiên cứu nhân chủng học muốn tìm hiểu xem chiều cao trung bình của thanh niên có thay đổi không so với mức 1,70m cách đây 10 năm. Một mẫu gồm 121 thanh niên được chọn và từ mẫu đó tính được  $\bar{x} = 1,72m$  với độ lệch tiêu chuẩn  $s = 0,02m$ . Với mức ý nghĩa  $\alpha = 0,05$ , ta có thể kết luận gì về sự thay đổi chiều cao trung bình của thanh niên ?

**101.** Một báo cáo của Sở Điện lực nói rằng trung bình một hộ hàng tháng phải trả 125 nghìn tiền điện với độ lệch tiêu chuẩn là 5 nghìn. Ta muốn kiểm định báo cáo nói trên bằng cách chọn ngẫu nhiên 400 hộ gia đình.

a) Phát biểu giả thiết  $H_0$  và đối thiết  $H_1$ .

b) Giả sử trung bình mẫu là 124 nghìn. Giả thiết  $H_0$  được chấp nhận hay bác bỏ ?

c) Tính P – giá trị.

**102.** Một nghiên cứu cho rằng trung bình một khách hàng vào siêu thị A tiêu hết 140 ngàn đồng. Ta muốn kiểm định khẳng định trên bằng cách chọn ngẫu nhiên 50 khách hàng và tính được số tiền trung bình họ tiêu là 154 nghìn với độ lệch tiêu chuẩn 62 nghìn. Phát biểu giả thiết  $H_0$  và đối thiết  $H_1$  và kiểm định giả thiết  $H_0$  với mức ý nghĩa 2%.

**103.** Một công ty có một loại máy tính có thể xử lý trung bình 1200 hóa đơn/giờ. Công ty được giới thiệu một loại máy tính mới. Loại máy này khi cho chạy kiểm tra thử trong 40 giờ cho thấy số hóa đơn xử lý trung bình trong 1 giờ là 1260 với độ lệch tiêu chuẩn là 125.

a) Phát biểu giả thiết  $H_0$  và đối thiết  $H_1$ .

b) Số liệu đã cho dẫn ta tới kết luận như thế nào ở mức ý nghĩa 5% ?

**104.** Theo thông báo của Lầu Năm góc, quân đội Mỹ bố trí trung bình 90 tên lửa ở mỗi căn cứ tên lửa. Một tổ chức hòa bình quốc tế dự định kiểm tra 49 căn cứ để xác định xem thông báo nói trên có đúng hay không. Đoàn kiểm tra sử dụng mức ý nghĩa 0,1.

a) Phát biểu giả thiết  $H_0$  và đối thiết  $H_1$ .

b) Giả sử đoàn kiểm tra tính được số tên lửa bố trí trung bình ở mỗi căn cứ là 92 với độ lệch tiêu chuẩn 9. Họ cần phải rút ra kết luận gì ?

c) Tính P – giá trị.

**105.** Theo quy định của chính phủ, nếu bình quân thu nhập đầu người của một huyện dưới 160 nghìn/tháng thì được trợ cấp thêm ngân sách. Huyện A có yêu cầu trợ cấp thêm ngân sách. Xem xét một mẫu gồm 75 gia đình ở huyện A cho thấy trung bình mẫu là 170 nghìn và độ lệch tiêu chuẩn mẫu là 30 nghìn. Huyện A có thể

biện luận rằng trung bình mẫu lớn hơn này là do ngẫu nhiên hay không ? Mức ý nghĩa sử dụng  $\alpha = 0,01$ .

**106.** Một nhà máy sản xuất sản phẩm ô tô tuyên bố rằng tuổi thọ trung bình một chiếc sản phẩm ô tô của nhà máy là 30000 dặm. Cơ quan giám định chất lượng nghi ngờ lời tuyên bố này đã kiểm tra ngẫu nhiên 100 chiếc sản phẩm và tìm được  $\bar{x} = 29000$  dặm và  $s = 5000$  dặm.

a) Với mức ý nghĩa  $\alpha = 5\%$  cơ quan giám định chất lượng cần rút ra kết luận gì ?

b) Cũng với câu hỏi trên nhưng với mức ý nghĩa  $\alpha = 2\%$ .

**107.** Một công ty sản xuất pin tuyên bố rằng pin của họ có tuổi thọ trung bình là 21,5 giờ. Cơ quan giám định chất lượng kiểm tra ngẫu nhiên 6 chiếc pin của công ty và thu được số liệu sau đây : tuổi thọ của 6 chiếc pin này là : 19, 18, 22, 20, 16, 25 (giờ). Với mức ý nghĩa 5% cơ quan giám định có kết luận gì ?

**108.** Trong năm học trước mức chi tiêu trung bình hàng tháng của một sinh viên là 420 nghìn. Trong năm học này với một mẫu ngẫu nhiên gồm 16 sinh viên ta tìm được trung bình mỗi tháng họ chi tiêu 442 nghìn với độ lệch tiêu chuẩn là 60 nghìn. Với mức ý nghĩa 5%, có thể kết luận rằng mức chi phí của sinh viên năm nay cao hơn năm ngoái hay không ?

**109.** Một bản thông báo nói rằng một người dân ở nước A trung bình mỗi năm đọc 10 cuốn sách. Trong một mẫu điều tra gồm 136 người được hỏi cho thấy trung bình 1 năm họ đọc 12 cuốn sách. Với mức ý nghĩa  $\alpha = 5\%$ , có nhận định gì về kết luận của bản thông báo nói trên ?

**110.** Một nhà sản xuất bánh ngọt tuyên bố rằng mỗi chiếc bánh của họ trung bình có chứa 88 calo. Với 36 chiếc bánh được kiểm tra cho thấy lượng calo trung bình trong mỗi chiếc bánh là 90 với

độ lệch tiêu chuẩn là 4. Với mức ý nghĩa 5%, kiểm định tính đúng đắn lời tuyên bố của nhà sản xuất.

**111.** Điểm trung bình trong kì thi tốt nghiệp của trường A năm ngoái là 9 (với thang điểm 20). Năm học này khi chọn ngẫu nhiên 400 em để xem xét, ta tính được trung bình của 400 em này là 9,2 với độ lệch tiêu chuẩn là 2,4. Với mức ý nghĩa 5%, có thể nhận định gì về điểm trung bình của học sinh năm nay ?

**112.** Năng suất lúa trung bình của giống lúa A là 43 tạ/ha. Trong vụ mùa năm nay khi bón thêm phân năng suất lúa trung bình tính trên 60 thửa ruộng là 46,2 tạ với độ lệch tiêu chuẩn  $s = 12$  tạ. Với mức ý nghĩa 5%, có thể kết luận được rằng do bón phân nên năng suất trung bình của vụ mùa năm nay là cao hơn 43 tạ/ha hay không ?

**113.** Trong một cửa hàng lớn có rất nhiều quầy hàng, mỗi nhân viên bán hàng trung bình mỗi ngày bán được 780 ngàn đồng. Trong một ngày khuyến mại, 80 nhân viên bán hàng mỗi người trung bình bán được 920 ngàn với độ lệch tiêu chuẩn 620 ngàn. Với mức ý nghĩa  $\alpha = 0,10$  có thể cho rằng trong ngày khuyến mại lượng hàng trung bình bán ra nhiều hơn ngày bình thường hay không ?

**114.** Một phương pháp ăn kiêng được quảng cáo rằng sẽ làm giảm trọng lượng trung bình là 45 pound (1 pound = 0,45 kg) trong vòng 6 tháng. Theo dõi 28 người theo chế độ ăn kiêng nói trên trong 6 tháng, ta thấy trung bình mỗi người giảm trọng lượng 35 pound với độ lệch tiêu chuẩn là 20 pound. Với mức ý nghĩa 0,01, hãy nhận định xem phương pháp ăn kiêng nói trên có quảng cáo quá sự thực hay không ?

**115.** Một loại dây cáp theo thiết kế được coi là có thể treo được vật nặng trung bình 1800 kg. Người ta đem thử một mẫu gồm 16

dây cáp nói trên và kết quả cho thấy vật nặng trung bình mà dây cáp treo được là 1740 kg, với độ lệch tiêu chuẩn là 60 kg. Sử dụng mức ý nghĩa 0,05, hãy đưa ra kết luận thích hợp trên cơ sở chứng cứ ở trên.

**116.** Một công ty lớn của Hoa Kỳ cho biết lương kỹ sư trung bình ở công ty là 48000 USD/năm. Một người dự định xin vào làm ở công ty này đã thăm dò 12 kỹ sư thì thấy lương trung bình của 12 người này là 45850 USD với độ lệch tiêu chuẩn 6300 USD. Với mức ý nghĩa 5%, người này nên rút ra kết luận gì ?

**117.** Theo quy định một vận động viên đẩy tạ cần đẩy xa trung bình mỗi lần là 43m. Huấn luyện viên thử kiểm tra anh ta 10 lần thì thấy khoảng cách anh ta đẩy được trung bình trong 10 lần này là 40,6m với độ lệch tiêu chuẩn là 3,8 m. Với mức ý nghĩa 5% huấn luyện viên có thể cho rằng anh ta đã đạt yêu cầu được không ? (Kết quả  $40,6m < 43m$  khi đó sẽ được xem là do ngẫu nhiên).

**118.** Một cửa hàng thông báo với công ty rằng trung bình mỗi ngày họ có 42 khách. Công ty nghi ngờ thông tin này đã kiểm tra ngẫu nhiên 22 ngày thì tính được số khách trung bình một ngày là 37,2 với độ lệch tiêu chuẩn là 7,4. Sử dụng mức ý nghĩa 2%, hãy nhận định xem sự nghi ngờ của công ty có cơ sở hay không ?

**119.** Công ty xe buýt nói rằng cứ trung bình 5 phút lại có một chuyến xe. Chọn ngẫu nhiên 8 thời điểm và ghi lại thời gian giữa hai chuyến xe buýt ta thu được số liệu sau đây :

5,3 ;    4,5 ;    4,8 ;    5,1 ;    4,3 ;    4,8 ;    4,9 ;    4,7

Với mức ý nghĩa 5%, nhận định xem công ty xe buýt nói có đúng không ?

**120.** Tỷ lệ không khỏi bệnh khi dùng thuốc A qua thống kê nhiều năm cho thấy là 16%. Với một loại thuốc B mới khi dùng cho 600 bệnh nhân thì có 70 người không khỏi bệnh. Với mức ý



nghe 5%, hãy so sánh tỉ lệ không khỏi bệnh của hai loại thuốc A và B. Tính P – giá trị.

**121.** Một tỉnh A báo cáo rằng tỉ lệ học sinh giỏi của tỉnh mình ít nhất là 20%. Thanh tra Bộ tiến hành xác minh báo cáo này. Họ chọn ngẫu nhiên 400 học sinh và thấy rằng trong đó có 60 em học sinh giỏi. Với mức ý nghĩa 1%, Bộ có cơ sở để bác bỏ báo cáo của tỉnh A không ?

**122.** Một đảng chính trị dự đoán rằng trong cuộc bầu cử tổng thống sắp tới ứng viên đảng mình sẽ giành được 45% số phiếu bầu. Chọn ngẫu nhiên 200 cử tri để thăm dò ý kiến cho thấy 80 người nói rằng họ sẽ bỏ phiếu cho ứng viên của đảng đó. Với mức ý nghĩa 5%, nhận định như thế nào về lời dự đoán của đảng đó ?

**123.** Biết rằng 18% số gia đình ở thành phố A có máy tính cá nhân. Trong số 80 gia đình có trẻ em đang đi học, thì thấy có 22 gia đình có máy tính cá nhân. Với mức ý nghĩa 2%, hãy nhận định xem có phải trong số các gia đình có trẻ em đang đi học, tỉ lệ gia đình có máy tính cá nhân cao hơn tỉ lệ chung hay không ?

**124.** Một công ty sản xuất bánh kẹo quảng cáo rằng ít nhất  $\frac{2}{3}$  số trẻ em thích ăn bánh kẹo của công ty. Trong một cuộc thăm dò ý kiến của 100 em, có 55 em nói rằng thích ăn bánh kẹo của công ty. Sử dụng mức ý nghĩa 5%, có nhận xét gì về lời quảng cáo của công ty ?

**125.** Một báo cáo nói rằng 62% số người lái xe trên đường là có bằng lái xe. Cảnh sát giao thông kiểm tra ngẫu nhiên 130 người phát hiện được 62 người không có bằng lái xe.

a) Phát biểu giả thiết  $H_0$  và đối thiết  $H_1$ .

b) Tính P – giá trị.

c) Giả thiết  $H_0$  bị bác bỏ ở mức ý nghĩa thấp nhất là bao nhiêu ?

**126.** Tỷ lệ học sinh tốt nghiệp phổ thông năm ngoái của tỉnh A là 88%. Trong kì thi năm nay trong 100 em được chọn ngẫu nhiên có 82 em thi đỗ. Với mức ý nghĩa 5%, có thể kết luận rằng tỉ lệ học sinh thi đỗ năm nay thấp hơn năm ngoái hay không ?

**127.** Một công ty quảng cáo rằng có 40% dân chúng ưa thích sản phẩm của công ty. Một cuộc điều tra 400 người tiêu dùng cho thấy có 125 người ưa thích sản phẩm của công ty. Sử dụng mức ý nghĩa 5%, kiểm định tính đúng đắn của lời quảng cáo của công ty.

**128.** Theo thống kê của Bộ Y tế, 12% dân cư trong tỉnh A mắc bệnh đau mắt hột. Khi kiểm tra 200 người ở tỉnh A đã phát hiện 21 người đau mắt hột. Số liệu này có khẳng định kết luận mà Bộ Y tế đưa ra không ? với mức ý nghĩa 1%.

**129.** Thống kê trong nhiều năm cho thấy 75% bệnh nhân chữa trị bằng thuốc A không thuyên giảm bệnh. Một loại thuốc mới B đang được thử nghiệm trên 120 người cho thấy có 82 người không thuyên giảm bệnh. Với mức ý nghĩa 5%, có thể khẳng định rằng thuốc B tốt hơn thuốc A được không ?

**130.** Một trung tâm cai nghiện ma túy công bố rằng nhiều nhất 22% số bệnh nhân của họ mắc nghiện trở lại trong vòng 2 năm. Theo dõi 35 bệnh nhân từ trung bình này trở về cho thấy có 10 người mắc nghiện trở lại. Với mức ý nghĩa 1%, hãy nhận định xem công bố của trung tâm có cao hơn so với thực tế hay không ?

**131.** Một Bộ nọ báo cáo rằng trong số nhân viên của họ có ít nhất 35% là nữ. Kiểm tra ngẫu nhiên danh sách 92 nhân viên cho thấy có 22 nữ. Sử dụng phương pháp P – giá trị, nhận định xem báo cáo của Bộ có đúng không ? với mức ý nghĩa  $\alpha = 0,025$ .

**132.** Gieo một con xúc sắc 600 lần. Số lần mặt trên của con xúc sắc ra 1, 2, 3, 4, 5, 6 được thống kê trong bảng sau đây :

1	2	3	4	5	6
106	92	97	105	88	112

Sử dụng mức ý nghĩa 5%, nhận định xem đây có phải là con xúc sắc được chế tạo cân đối đồng chất hay không ?

**133.** Một nhà tâm lí học quan tâm tới việc xác định xem trẻ em chậm phát triển về trí tuệ có tỏ ra ưa thích một loại màu nào đó hơn các màu khác hay không. Ông ta thử nghiệm trên 80 đứa trẻ chậm phát triển với 4 màu : xanh, vàng, da cam và nâu. Kết quả như sau :

Màu	Số em ưa thích
Nâu	25
Da cam	18
Vàng	19
Xanh	18
Tổng số	80

Với mức ý nghĩa 5%, nhận định xem trẻ em chậm phát triển về trí tuệ có thiên hướng ưa thích một màu nào đó hay không ?

**134.** Một bánh xe quay trong khi quay xổ số phải có xác suất dừng lại như nhau ở các chữ số 0, 1, 2, ..., 9. Để kiểm tra điều này, người ta quay thử 200 lần và thu được kết quả sau :

Chữ số	0	1	2	3	4	5
Tần số	10	23	15	24	21	23

Chữ số	6	7	8	9		
Tần số	19	18	25	22		

Với mức ý nghĩa 1%, hãy rút ra kết luận về chiếc bánh xe quay xổ số này.

**135.** Công ty bánh kẹo A cho biết rằng trong số kẹo sôcôla của công ty có 30% kẹo màu nâu, 20% màu xanh, 20% màu đỏ, 20%

màu vàng và 10% màu da cam. Mua một túi kẹo gồm 188 chiếc ta có 67 chiếc màu nâu, 24 chiếc màu xanh, 51 chiếc màu đỏ, 22 chiếc màu vàng và 24 chiếc màu da cam.

Với mức ý nghĩa 1%, kiểm định xem sự phân bố về màu kẹo sôcôla của công ty có đúng như đã nói hay không ?

**136.** Trong mùa thi đại học năm trước, tỉ lệ thí sinh đăng kí thi vào khối A là 64%, vào khối B là 10% và vào khối C là 26%.

Trong năm học này khi xem xét 360 hồ sơ đăng kí dự thi, ta thấy có 206 em đăng kí khối A, 37 em khối B và 117 em khối C. Sử dụng mức ý nghĩa 1%, kiểm định giả thiết  $H_0$  : "Không có sự thay đổi gì về nguyện vọng của thí sinh so với năm ngoái".

**137.** Theo thông báo của Bộ Tư pháp Mỹ, trong vòng 5 năm sau khi được trả tự do 20% phạm nhân không tái phạm, 38% tái phạm 1 lần, 18% tái phạm 2 lần, 13,5% tái phạm 3 lần và 10,5% tái phạm 4 lần trở lên.

Một nhà nghiên cứu xã hội học theo dõi hồ sơ 200 phạm nhân và thu được số liệu sau :

Số lần tái phạm của phạm nhân	Số phạm nhân
0	58
1	62
2	28
3	16
$\geq 4$	36
Tổng số	200

Với mức ý nghĩa 1%, số liệu thực nghiệm này có phù hợp với thông báo của tòa án hay không ?

**138.** Chi nhánh điện quận A có ghi lại các vụ báo cáo sửa chữa điện cho tuần sau :

Chủ nhật	Thứ hai	Thứ ba	Thứ tư	Thứ năm	Thứ sáu	Thứ bảy
22	12	15	8	14	27	35

Căn cứ trên số liệu thống kê này, với mức ý nghĩa 1%, hãy nhận định xem các sự cố về điện có phân bố đều giữa các ngày trong tuần hay không ?

**139.** Hai con xúc sắc được ném 360 lần, mỗi lần người ta tính tổng số nốt ghi trên mặt của hai con xúc sắc. Kết quả cho như sau :

Tổng	2	3	4	5	6	7	8	9	10	11	12
Tần số	8	15	26	42	50	65	48	44	32	22	8

Với mức ý nghĩa  $\alpha = 5\%$ , nhận định xem có phải hai con xúc sắc được chế tạo cân đối hay không ?

**140.** Thống kê 250 ngày nghỉ ốm của công nhân trong một xí nghiệp lớn ta có được số liệu sau :

Ngày nghỉ ốm	Số công nhân nghỉ
Thứ hai	57
Thứ ba	39
Thứ tư	37
Thứ năm	54
Thứ sáu	63
Tổng số	250

Với mức ý nghĩa 5%, hãy nhận định xem sự cố "ốm" của công nhân trong xí nghiệp có được phân bố đều giữa các ngày làm việc hay không ?

**141.** Trong dân gian lưu truyền một quan niệm rằng một loại thức ăn A nào đó làm tăng khả năng sinh con trai. Để kiểm chứng quán niệm này người ta cho một nhóm phụ nữ dùng thức ăn A rồi xem xét 80 trường hợp có 3 con trong thời gian dùng loại thức ăn A đó. Kết quả thống kê cho ở bảng dưới đây :

Số bé trai	3	2	1	0	Tổng số
Số phụ nữ	14	36	24	6	80

Với mức ý nghĩa 5% kiểm định xem loại thức ăn A có ảnh hưởng hay không đến việc sinh con trai ?

**142.** Một nhà máy sản xuất máy in A nói rằng số lỗi in trong một cuốn sách 300 trang là một ĐLNN có quy luật phân bố Poisson với tham số  $\lambda = 4,7$ . Kiểm tra ngẫu nhiên 50 cuốn sách có 300 trang in từ máy in A đó ta thu được số liệu sau :

Số lỗi	Số cuốn
0	1
1	1
2	8
3	6
4	13
5	10
6	4
7	5
8	2
$\geq 9$	0

Với mức ý nghĩa 5%, số liệu trên có xác nhận lời tuyên bố của các nhà sản xuất máy in A hay không ?

**143.** Dân cư trong một nước X có phân bố nhóm máu như sau : 45% nhóm máu O ; 40% nhóm máu A ; 10% nhóm máu B và 5% nhóm máu AB. Một mẫu gồm 200 người ở nước Y được kiểm tra nhóm máu và cho kết quả sau :

Nhóm máu	O	A	B	AB
Số người	80	72	24	24

Với mức ý nghĩa 5%, ta có thể kết luận được rằng dân cư của nước Y có phân bố nhóm máu khác với dân cư của nước X hay không ?

**144.** Kích thước của một chi tiết máy là một ĐLNN của phân bố chuẩn. Chọn ngẫu nhiên 30 chi tiết máy ta tìm được giá trị trung bình  $\bar{x} = 0,47$  với độ lệch tiêu chuẩn của tập hợp chính gồm kích thước của tất cả các chi tiết máy.

**145.** Đo đường kính của 12 chi tiết máy do một dây chuyền tự động sản xuất, nhân viên kiểm tra chất lượng tính được  $s = 0,3$ . Biết rằng nếu độ biến động của sản phẩm lớn hơn 0,2 thì dây chuyền sản xuất cần phải ngừng lại để điều chỉnh. Với mức ý nghĩa 5%, dây chuyền sản xuất có cần phải ngừng lại để điều chỉnh hay không ?

**146.** Trong một mẫu ngẫu nhiên 30 đĩa trẻ, thời gian trung bình xem T.V. trong một tuần là 48 giờ với độ lệch tiêu chuẩn là 12,4 giờ. Tìm khoảng tin cậy cho độ lệch tiêu chuẩn của thời gian xem T.V với tất cả các đĩa trẻ.

**147.** Ở một khách sạn A, người ta thấy rằng thời gian đợi phục vụ của một khách hàng là một ĐLNN có phân bố chuẩn với trung bình là 3,4 phút và độ lệch tiêu chuẩn là 2,75 phút. Sau một thời gian cải tiến cách phục vụ nhằm giảm độ lệch tiêu chuẩn, theo dõi 30 khách hàng cho thấy thời gian đợi trung bình là 3,6 phút với độ lệch tiêu chuẩn là 2,32 phút. Hãy kiểm định xem sau một thời gian cải tiến, độ lệch tiêu chuẩn của thời gian phục vụ đã giảm được chưa ? Sử dụng mức ý nghĩa 5%.

## Chương 5

### BÀI TOÁN SO SÁNH

**148.** Một cuộc nghiên cứu được tiến hành nhằm so sánh tuổi trung bình của người mẹ khi sinh đứa con cuối cùng giữa hai vùng A và B. Một mẫu ngẫu nhiên gồm 36 phụ nữ ở vùng A cho kết quả tuổi trung bình khi sinh con lần cuối là 32,9 với độ lệch tiêu chuẩn là 5,7 năm. Một mẫu ngẫu nhiên gồm 49 phụ nữ ở vùng B cho kết quả tuổi trung bình khi sinh con lần cuối là 29,6 với độ lệch tiêu chuẩn là 5,5 năm.

a) Với mức ý nghĩa 5%, nhà nghiên cứu cần rút ra kết luận gì ?

b) Tính P – giá trị. Giả thiết  $H_0$  bị bác bỏ ở mức thấp nhất là bao nhiêu ?

**149.** Người ta muốn so sánh tốc độ phát nổ của hai loại đầu đạn A và B. Mỗi loại được thử với 30 đầu đạn. Đối với 30 đầu đạn loại A ta có giá trị trung bình là 20,44 giây và độ lệch tiêu chuẩn là 2,96 giây. Đối với đầu đạn loại B, giá trị trung bình là 19,45 giây với độ lệch tiêu chuẩn là 2,13 giây. Ta cần rút ra kết luận như thế nào khi sử dụng mức ý nghĩa 1% ?

**150.** Người ta muốn so sánh chất lượng đào tạo tại hai cơ sở đào tạo A và B bằng cách căn cứ trên điểm trung bình tại kì thi quốc gia. Trong một mẫu gồm 100 thí sinh được đào tạo tại cơ sở A được chọn ngẫu nhiên, điểm trung bình của họ là 9,4 với độ lệch tiêu chuẩn 0,8 (thang điểm 10). Đối với mẫu gồm 80 thí sinh của cơ sở B, điểm trung bình là 9 với độ lệch tiêu chuẩn 1.



Với mức ý nghĩa 1%, hãy phát biểu bài toán so sánh điểm thi trung bình của hai cơ sở đào tạo A và B, trong đó được biết rằng cơ sở đào tạo A có cơ sở vật chất, đội ngũ giáo viên tốt hơn cơ sở B.

**151.** Một cuộc nghiên cứu được tiến hành để so sánh mức lương trung bình của phụ nữ và mức lương trung bình của nam giới trong một công ty lớn. Một mẫu gồm 100 phụ nữ có mức lương trung bình là 7,23 đôla/giờ với độ lệch tiêu chuẩn là 1,64 đôla. Một mẫu gồm 75 nam có mức lương trung bình là 8,06 đôla/giờ với độ lệch tiêu chuẩn là 1,85 đôla. Số liệu đã cho có chứng minh được rằng mức lương trung bình của phụ nữ trong công ty là thấp hơn nam giới hay không ? Sử dụng mức ý nghĩa 1%.

**152.** Cơ quan không gian Mỹ (NASA) đã kí hợp đồng với hai công ty A và B để sản xuất thử pin dùng cho vệ tinh viễn thông.

Dựa trên kết quả của các pin thử nghiệm NASA sẽ quyết định chọn công ty nào làm nhà cung cấp pin cho vệ sinh viễn thông. Công ty A cho sản xuất thử 10 chiếc, tuổi thọ trung bình của chúng là 4,8 năm và độ lệch tiêu chuẩn là 1,1 năm. Công ty B sản xuất thử được 12 chiếc với tuổi thọ trung bình là 4,3 năm và độ lệch tiêu chuẩn là 0,3 năm. Với mức ý nghĩa 1% kiểm định xem có sự khác nhau về tuổi thọ trung bình của hai loại pin hay không ?

**153.** Người ta ghi lại sản lượng lúa mì tính bằng tạ/ha của các mảnh ruộng đã bón lót 50 và 100 đơn vị đạm. Trên 11 thửa bón 100 đơn vị đạm cho ta năng suất như sau :

47,9	48,9	43,5	53,1	50,8	46,1
41,1	43,0	41,0	48,5		
47,7					

Trên 8 thửa bón 50 đơn vị đạm cho ta năng suất sau :

47,2      43,1      35,7      47,0

45,7      42,6      40,7      42,3

Với mức ý nghĩa 5%, có thể kết luận rằng bón lót 100 đơn vị đạm cho năng suất cao hơn bón lót 50 đơn vị đạm hay không ?

**154.** Một cuộc nghiên cứu được tiến hành để so sánh thời gian nằm viện trung bình của bệnh nhân nam điều trị bệnh A và thời gian nằm viện trung bình của bệnh nhân nữ điều trị bệnh A. Khảo sát một mẫu ngẫu nhiên 50 bệnh nhân nam và một mẫu ngẫu nhiên 40 bệnh nhân nữ điều trị bệnh A cho thấy : Đối với nhóm bệnh nhân nam, thời gian trung bình là 5,3 ngày với độ lệch tiêu chuẩn là 2,1 ngày. Đối với nhóm bệnh nhân nữ, thời gian trung bình là 6,2 ngày với độ lệch tiêu chuẩn là 1,8 ngày. Với mức  $\alpha = 5\%$  ta cần có kết luận như thế nào ?

**155.** Một nhà khoa học tiến hành một công trình nghiên cứu nhằm so sánh tác dụng của hai loại phân A và B đối với năng suất cây cà chua. Với 60 cây cà chua được bón phân A, nhà nghiên cứu thu được sản lượng trung bình 32,2 quả với độ lệch tiêu chuẩn 8,5. Với 72 cây cà chua được bón phân B anh ta thu được sản lượng trung bình 28,4 quả với độ lệch tiêu chuẩn 9,3. Với mức ý nghĩa 5%, nhà nghiên cứu cần rút ra kết luận gì ?

**156.** Một công ty vận tải lớn muốn so sánh số ô tô tải gặp trục trặc trong mùa đông với số ô tô tải gặp trục trặc trong mùa hè. Chọn ngẫu nhiên 35 ngày trong mùa đông cho ta số ô tô tải gặp trục trặc kỹ thuật trung bình trong một ngày là 16,6 với mức ý nghĩa 7,1. Chọn ngẫu nhiên 32 ngày trong mùa hè cho ta số ô tô tải gặp trục trặc kỹ thuật trung bình trong một ngày là 12,4 với độ lệch tiêu chuẩn 5,8. Với mức ý nghĩa  $\alpha = 2\%$ , dựa trên số liệu đó công ty vận tải đó cần có kết luận như thế nào ?

**157.** Để so sánh chiều cao trung bình của thanh niên nam của hai vùng dân cư A và B, người ta chọn ngẫu nhiên 10 thanh niên nam ở vùng A và 10 thanh niên nam ở vùng B. Số đo chiều cao của hai nhóm người này được cho như sau (đơn vị cm) :

*Vùng A :* 165   167   174   172   165

167   168   172   170   173

*Vùng B :* 172   170   167   169   171

167   173   165   163   174

Với mức ý nghĩa  $\alpha = 1\%$ , hãy so sánh chiều cao trung bình của thanh niên giữa hai vùng A và B.

**158.** Một tạp chí Y học vừa công bố một công trình nghiên cứu cho thấy : những đứa trẻ được nuôi bằng sữa mẹ có chỉ số thông minh (IQ) cao hơn những đứa trẻ được nuôi bằng bú chai. Để kiểm tra tính đúng đắn của thông báo này người ta chọn một mẫu 10 đứa trẻ nuôi bằng sữa mẹ và một mẫu 10 đứa trẻ bú chai và ghi lại các chỉ số IQ của chúng. Kết quả như sau :

*Nhóm bú mẹ :*    121   105   111   119   108   101

90    131   106   112

*Nhóm bú chai :*   102   110   107   98    99    103

86    117   113   87

Dựa trên số liệu này có thể kết luận rằng công trình nghiên cứu nói trên là đúng hay không ? Sử dụng mức ý nghĩa 5%.

**159.** Có hai lô chuột thí nghiệm tăng trọng với hai khẩu phần ăn khác nhau. Lô thứ nhất cho ăn khẩu phần ăn nhiều đậm. Lô thứ hai cho ăn với khẩu phần ăn ít đậm hơn. Sự tăng trọng của hai lô chuột sau một thời gian được ghi lại như sau (đơn vị mg) :

*Lô thứ nhất :*   123,   134,   146,   104,   119,   124,

161,   107,   83,    113,   129,   97

*Lô thứ hai :*    70,    118,   85,    107,   132,   94,    101.

Với mức ý nghĩa  $\alpha = 5\%$ , nhận định xem việc cho ăn nhiều đạm có tác dụng tăng trọng hay không ?

**160.** Đo thời gian cắt của một máy tiện loại mới với một máy tiện loại cũ 10 lần ta thu được kết quả sau đây (đơn vị là giây) :

Máy cũ : 58    58    56    38    70    38    42    75    68    67

Máy mới : 57    55    63    24    67    43    33    68    56    54

Với mức ý nghĩa 5%, có thể cho rằng máy mới tốt hơn máy cũ hay không ?

**161.** Để khảo sát ảnh hưởng của magiê (Mg) đến sự phát triển của một loại mốc, người ta so sánh 10 mẫu cấy một liều Mg là 5mg và 10 mẫu cấy một liều 10mg. Đối với mẫu thứ nhất  $(x_i)_{i=1}^{10}$

ta có  $\sum_{i=1}^{10} x_i = 65$ ,  $\sum_{i=1}^{10} x_i^2 = 621$ . Đối với mẫu thứ hai  $(y_i)_{i=1}^{10}$  ta có

$$\sum_{i=1}^{10} y_i = 245, \sum_{i=1}^{10} y_i^2 = 6503.$$

Có thể khẳng định rằng magiê có tác dụng tới sự tăng trưởng của mốc hay không ? Sử dụng mức ý nghĩa 5%.

**162.** Người ta muốn so sánh trọng lượng trung bình bộ não của người trên 50 tuổi và bộ não của người dưới 50 tuổi. Kết quả đo trên một mẫu 127 người trên 50 tuổi cho ta các số liệu  $(x_i)_{i=1}^{127}$  với  $\sum x_i = 13$ ,  $\sum x_i^2 = 313$ . Số liệu đo trên một mẫu 265 người dưới 50 tuổi cho ta 265 số liệu  $(y_i)$  với  $\sum y_i = 55$ ,  $\sum y_i^2 = 767$ . Với mức ý nghĩa  $\alpha = 5\%$ , ta có thể coi rằng trọng lượng trung bình bộ não của người trên 50 tuổi và dưới 50 tuổi là như nhau hay không ?

**163.** Người ta tiến hành một cuộc khảo sát về giá cả trong hai siêu thị lớn trong thành phố 12 mặt hàng được chọn ngẫu nhiên và giá bán của chúng tại hai cửa hiệu được ghi lại như sau :

Mặt hàng	Cửa hàng A	Cửa hàng B
1	0,89	0,95
2	0,59	0,55
3	1,29	1,49
4	1,50	1,69
5	2,49	2,39
6	0,65	0,79
7	2,25	2,39
8	0,99	0,99
9	1,99	1,79
10	0,5	0,59
11	1,99	2,19
12	1,79	1,99

Với mức ý nghĩa  $\alpha = 2\%$ , hãy kiểm định xem có sự khác nhau về giá bán ở hai cửa hiệu hay không ?

**164.** Để xác định hiệu quả của chế độ ăn kiêng trong việc làm giảm trọng lượng cơ thể, 10 người tình nguyện đã được chọn để thử nghiệm chế độ ăn kiêng trong 6 tuần lễ. Kết quả được ghi lại trong bảng dưới đây :

Người thử nghiệm	Trước ăn kiêng	Sau ăn kiêng
1	190	185
2	202	197
3	177	185
4	160	152
5	225	205
6	180	184
7	196	185
8	208	200
9	185	187
10	177	170

(Đơn vị đo ở đây là pound (1 pound = 0,454 kg). Hãy kiểm định xem chế độ ăn kiêng có giúp làm giảm trọng lượng hay không ? Mức ý nghĩa sử dụng là 5%.

**165.** Trước và sau dịp Tết giá của mặt hàng A tại 8 cửa hiệu trong thành phố như sau :

Cửa hiệu	Trước Tết	Sau Tết
1	95	98
2	109	105
3	99	99
4	98	99
5	105	109
6	99	105
7	109	115
8	102	110

Với mức ý nghĩa 5%, kiểm định xem có phải có một khuynh hướng tăng giá sau Tết đối với mặt hàng A hay không ?

**166.** Người ta chọn ngẫu nhiên 7 sinh viên của một trường đại học và ghi lại điểm thi trung bình của họ ở năm thứ nhất và năm thứ hai. Kết quả được cho như sau :

Sinh viên	Năm I	Năm II
1	7,5	7,25
2	7,78	7,62
3	9,15	9,34
4	9,62	9,45
5	6,15	6,55
6	4,84	4,58
7	8,61	8,7

Với mức ý nghĩa 5%, hãy nhận định xem có sự khác nhau về điểm trung bình giữa năm I và năm II hay không ?

**167.** Để nghiên cứu ảnh hưởng của côban lên sự phát triển của thỏ, người ta làm thí nghiệm với hai nhóm thỏ. Nhóm thứ nhất (nhóm đối chứng) không cho ăn côban. Nhóm thứ hai (nhóm thí nghiệm hàng ngày cho ăn thêm vào thức ăn một lượng côban dưới dạng dung dịch. Sự tăng trọng (sau 1 tháng) của hai nhóm được cho như sau :

*Nhóm thứ nhất :* 504, 560, 420, 600, 580, 530, 490, 480, 470

*Nhóm thứ hai :* 580, 692, 700, 621, 640, 561, 680, 630

Với mức ý nghĩa 1%, có kết luận gì về ảnh hưởng của côban lên sự tăng trọng ?

**168.** Để so sánh chiều cao của đàn ông trưởng hành ở vùng A và vùng B ta lấy một nhóm 7 người đàn ông vùng A được chọn ngẫu nhiên và 6 người đàn ông vùng B được chọn ngẫu nhiên. Chiều cao của hai nhóm này như sau (đơn vị cm) :

*Vùng A :* 170 178 160 180 160 160 168

*Vùng B :* 180 175 162 170 168 171

Với mức  $\alpha = 5\%$ , hỏi có sự khác nhau về chiều cao trung bình của đàn ông tỉnh A và tỉnh B hay không ?

**169.** Hai giống lúa A và B được trồng trên 7 thửa ruộng thí nghiệm cạnh nhau ở 7 trạm thực nghiệm nông nghiệp. Sản lượng của các giống lúa A và B được ghi lại như sau :

Trạm	Giống A	Giống B
1	7,7	8,26
2	9,0	7,22
3	9,4	8,43
4	7,4	5,57
5	10,9	6,35
6	8,0	8,00
7		9,13

Với mức ý nghĩa 5%, có sự khác nhau về năng suất giữa giống lúa A và giống lúa B hay không ?

170. Sản lượng của hai giống lúa A và B được trồng cùng ở một trạm thực nghiệm trong 6 năm cho kết quả sau :

Giống A : 31,6 24,2 24,8 29,1 29,9 31

Giống B : 31,1 24,0 24,6 28,6 29,1 30,1

Với mức ý nghĩa 5%, có sự khác nhau về năng suất giữa các giống lúa A và B hay không ?

171. Một người lái xe thường xuyên đi lại giữa hai địa điểm A và B. Có hai con đường X và Y nối A với B. Chọn ngẫu nhiên 10 ngày đi trên đường Y, ghi lại số thời gian đi, anh ta có số liệu sau (thời gian đo bằng phút) :

Đường X: 34 28 46 42 56 85 48 25 37 49

Đường Y : 43 49 41 55 39 45 65 50 47 51

Sử dụng tiêu chuẩn hạng Mann – Whitney, so sánh thời gian trung bình đi giữa hai con đường X và Y của anh ta (sử dụng mức ý nghĩa 5%).

172. Một số nhà nghiên cứu cho rằng các bệnh nhân nam thường nằm viện lâu hơn các bệnh nhân nữ. Để kiểm tra tính đúng đắn của giả thiết này, người ta chọn ngẫu nhiên 9 hồ sơ bệnh nhân nam và 8 hồ sơ bệnh nhân nữ của một bệnh viện và ghi lại thời gian nằm viện của những người này. Số liệu được cho như sau (đơn vị thời gian là ngày) :

Nam	Nữ
13	11
15	14
9	10
18	8
11	16
20	9
24	17
22	21
25	



Với mức ý nghĩa 5%, có kết luận gì về thời gian nằm viện trung bình của các bệnh nhân nam và nữ (sử dụng tiêu chuẩn hạng Mann – Whitney).

**173.** Một mẫu gồm 8 công nhân trong xí nghiệp A và 10 công nhân trong xí nghiệp B được chọn ngẫu nhiên để xem xét xem liệu có sự khác nhau về thời gian nghỉ phép trong năm của các công nhân hai xí nghiệp hay không. Các số liệu được ghi lại như sau :

*Xí nghiệp A :* 11, 13, 14, 15, 8, 10, 5, 6

*Xí nghiệp B :* 8, 9, 10, 16, 17, 1, 7, 2, 5, 4

Sử dụng tiêu chuẩn hạng Mann – Whitney, hãy kiểm định giả thiết rằng không có sự khác nhau về thời gian nghỉ phép trong năm của công nhân hai xí nghiệp A và B. Mức ý nghĩa 10%.

**174.** Tuổi thọ của các bóng đèn (tính bằng giờ) do hai nhà máy A và B sản xuất được thử nghiệm trên 8 bóng của nhà máy A và 8 bóng của nhà máy B. Kết quả như sau :

*Nhà máy A :* 372, 283, 712, 849, 623, 382, 427, 821

*Nhà máy B :* 645, 682, 913, 742, 691, 689, 842, 751

Với mức ý nghĩa 5%, ta có thể kết luận rằng bóng đèn do các nhà máy A và B sản xuất có khác nhau về chất lượng hay không ? (Sử dụng tiêu chuẩn hạng Mann – Whitney).

**175.** Người ta muốn so sánh phân bố chiều cao của các trẻ em 6 tuổi ở các nước phát triển và các nước đang phát triển. Ta có hai mẫu số liệu sau đây (đơn vị chiều cao là inch ; 1 inch = 2,54cm).

*Các nước phát triển :*      30,2   34,6   37,8   40,8   43,4  
   45,9   39,2   40,4   40

*Các nước đang phát triển :*      29,4   33,8   37,5   40,7  
   43,4   31,2   32,1   33,4

Với mức ý nghĩa 5%, có thể kết luận rằng hai phân bố đó khác nhau hay không ? (Sử dụng tiêu chuẩn hạng Mann – Whitney).

**176.** Một nhà tâm lí học đã đưa ra một phương pháp giáo dục nhằm làm tăng tính tự tin của các em hay nhút nhát. Phương pháp này thử nghiệm trên 13 em cho kết quả sau

Trẻ em	Trước thí nghiệm	Sau thí nghiệm
1	Nhút nhát	Hơi tự tin
2	Bình thường	Tự tin
3	Nhút nhát	Tự tin
4	Bình thường	Tự tin
5	Bình thường	Nhút nhát
6	Nhút nhát	Khá tự tin
7	Bình thường	Tự tin
8	Nhút nhát	Tự tin
9	Nhút nhát	Hơi tự tin
10	Nhút nhát	Hơi tự tin
11	Tự tin	Nhút nhát
12	Nhút nhát	Tự tin
13	Nhút nhát	Nhút nhát

Với mức ý nghĩa 10%, chúng ta có thể kết luận rằng phương pháp của nhà tâm lí học đó có tác dụng làm tăng mức tự tin của trẻ em hay không ?

**177.** Một nhà tâm lí học khảo sát tâm trạng "lạc quan" của các cặp vợ chồng già. Ông ta đưa ra một thang điểm từ 0 đến 100 trong đó 0 điểm là điểm hoàn toàn không lạc quan, còn điểm 100 là cực kì lạc quan. Kết quả khảo sát 20 cặp vợ chồng cho số liệu sau đây :

Chồng	70	85	73	75	65
Vợ	65	41	45	80	84

Chồng	50	80	71	80	51
Vợ	50	71	52	42	78

Chồng	72	76	79	65	59
Vợ	62	38	80	65	54

Chồng	72	84	90	56	57
Vợ	67	87	90	38	43

Với mức ý nghĩa 5%, sử dụng tiêu chuẩn dấu, hãy kiểm định xem liệu có sự khác nhau giữa vợ và chồng về tâm trạng lạc quan hay không ?

**178.** Một thầy giáo dạy Toán cho rằng việc cho học sinh ôn tập một vài buổi trước khi thi có tác dụng tốt tới kết quả học tập của các em. Một mẫu gồm 21 học sinh được chọn để theo dõi điểm thi của các em trước và sau khi ôn tập. Kết quả cho ở bảng sau đây :

Học sinh	Điểm thi trước khi ôn tập	Điểm thi sau khi ôn tập
1	22	21
2	26	29
3	17	15
4	20	20
5	28	26
6	31	32
7	23	25
8	13	14
9	19	19
10	25	27
11	28	27
12	24	25
13	27	27
14	18	20
15	20	23
16	14	16
17	24	26
18	15	20
19	19	20
20	18	17
21	27	19

Với mức ý nghĩa 5%, có thể kết luận rằng sau khi được ôn tập kết quả thi của học sinh tốt hơn hay không ? Sử dụng tiêu chuẩn dấu với mức ý nghĩa 5%.

**179.** Kiểm tra số biên lai phạt vi cảnh mà hai cảnh sát giao thông A và B xuất ra trong 10 ngày chọn ngẫu nhiên, ta có số liệu sau :

*Cảnh sát A :* 32, 14, 26, 37, 45, 58, 28, 32, 36, 25, 30

*Cảnh sát B :* 44, 37, 24, 33, 27, 41, 29, 25, 34

Sử dụng tiêu chuẩn hạng Mann – Whitney, hãy so sánh phân bố số biên lai phạt vi cảnh của hai cảnh sát A và B. Sử dụng mức ý nghĩa 5%.

**180.** Người quản lí một nhà hàng muốn so sánh số khách trung bình mà hai tiếp viên A, B của nhà hàng phục vụ trong mỗi ngày. Anh ta thu được các con số thống kê sau :

*Tiếp viên A :* 42, 36, 58, 27, 48, 85, 38, 44, 62

*Tiếp viên B :* 53, 48, 65, 41, 57, 49, 74, 49, 56

Sử dụng tiêu chuẩn Mann – Whitney, nhận định xem số khách trung bình mà A và B phục vụ mỗi ngày là như nhau hay khác nhau ?

**181.** Một nữ giáo sư bị chỉ trích là có xu hướng thiên vị các sinh viên nam. Để kiểm tra tính đúng đắn của lời chỉ trích này, ông chủ nhiệm khoa đã chọn một số bài thi của sinh viên nam và sinh viên nữ cùng học vị nữ giáo sư trên để so sánh. Kết quả thu được như sau (điểm tối đa mỗi bài là 100) :

*Sinh viên nam :* 75, 86, 77, 72, 89, 94,

81, 83, 77, 73, 86, 90, 90, 97.

*Sinh viên nữ :* 77, 83, 72, 67, 84, 91, 82, 73, 65, 72, 70,

72, 65, 88.

Sử dụng tiêu chuẩn hạng Mann – Whitney, với mức ý nghĩa 5% có thể cho rằng điểm trung bình của sinh viên nam là khác điểm trung bình của sinh viên nữ hay không ?

**182.** Một phương pháp để đo lường "sức khỏe" của một nền kinh tế là theo dõi số công việc cần người làm được quảng cáo trên các báo. Trong một nghiên cứu nhằm so sánh tình trạng kinh tế của hai năm 1996 và 1997 ở một nước, người ta đã đếm số trang quảng cáo việc tìm người trên mặt báo mỗi ngày. Số liệu thu được ghi lại trong 9 ngày chọn ngẫu nhiên của tháng 7 năm 1996 và tháng 7 năm 1997 như sau :

*Tháng 7/1996 :* 12, 10, 25, 18, 14, 18, 20, 24, 20

*Tháng 7/1997 :* 21, 16, 26, 17, 20, 28, 19, 20, 18

Với mức ý nghĩa 2% hãy so sánh số trang quảng cáo việc tìm người trung bình trong mỗi ngày của hai năm 1996 và 1997.

**183.** Có hai nhóm công nhân, một nhóm toàn nam, một nhóm gồm toàn nữ, được đào tạo để làm cùng một nghề. Thời gian để hoàn thành công việc được giao (tính bằng ngày) của mỗi công nhân được ghi lại như sau :

*Công nhân nam :* 48, 42, 55, 62, 60, 74, 53, 34, 40, 66,  
72, 65, 48, 52, 73, 79.

*Công nhân nữ :* 60, 54, 50, 51, 53, 65, 41, 45, 50, 61,  
63, 37, 43, 48, 46, 49.

Sử dụng tiêu chuẩn hạng Mann – Whitney hãy so sánh năng suất lao động của công nhân nữ và công nhân nam. Mức ý nghĩa 5%.

**184.** Để so sánh năng suất lao động trước khi nghỉ phép của công nhân, một mẫu gồm 22 công nhân được lựa chọn để theo dõi năng suất lao động của họ trước khi nghỉ và sau khi nghỉ 1 tuần. Kết quả ghi lại như sau :

Số thứ tự	Năng suất trước khi nghỉ	Năng suất sau khi nghỉ
1	83	79
2	85	87
3	75	70
4	91	93
5	80	85
6	75	75
7	90	80
8	65	71
9	78	80
10	85	88
11	83	82
12	75	71
13	78	75
14	80	85
15	82	86
16	88	85
17	85	82
18	80	87
19	78	78
20	81	84
21	70	85
22	80	81

a) Dùng tiêu chuẩn dấu, với mức ý nghĩa  $\alpha = 5\%$ , hãy nhận định xem năng suất lao động của công nhân trước và sau khi nghỉ phép có thay đổi hay không ?

b) Cũng câu hỏi trên nhưng hãy sử dụng tiêu chuẩn hạng có dấu Wilconxon.

**185.** Có ý kiến cho rằng trong hai anh em trai, người em luôn cao hơn người anh. Một mẫu ngẫu nhiên gồm 25 cặp anh em trai được chọn. Chiều cao của người anh (X) và của người em (Y) của 25 cặp này như sau (đơn vị cm) :

X	170	169	167	168	166
Y	175	172	167	166	163

X	165	165	164	164	165
Y	166	164	167	163	167

X	166	166	169	168	168
Y	168	164	170	172	171

X	166	168	168	169	169
Y	170	167	165	166	171

X	164	170	169	169	166
Y	163	168	166	167	166

Với mức ý nghĩa  $\alpha = 5\%$ , sử dụng tiêu chuẩn hạng có dấu hãy nhận định xem ý kiến nêu ra có đúng không ?

**186.** Sử dụng số liệu sau đây và tiêu chuẩn hạng có dấu Wilconxon, hãy xác định xem liệu có sự khác nhau về hiệu quả của hai phương pháp A, B khi tác động lên cùng một đối tượng hay không ? Mức ý nghĩa 5%.

Đối tượng	Phương pháp A	Phương pháp B
1	68	32
2	55	55
3	49	55
4	40	75
5	20	50
6	18	25
7	30	23
8	70	49
9	52	62
10	50	41

**187.** Chọn ngẫu nhiên 8 cặp vợ chồng và ghi lại tuổi của họ ta có số liệu sau :

Số thứ tự	Tuổi chồng	Tuổi vợ
1	24	21
2	25	23
3	27	19
4	22	24
5	20	18
6	30	25
7	18	17
8	21	22

Sử dụng tiêu chuẩn hạng có dấu Wilconxon, hãy nhận định xem tuổi trung bình của chồng và vợ như nhau hay khác nhau ? Mức sử dụng ý nghĩa 10%.

**188.** Để đánh giá hiệu quả của một chiến dịch chống hút thuốc lá, các nhà nghiên cứu chọn ra 12 đối tượng hút thuốc lá để tuyên truyền. Số điếu thuốc mà các đối tượng hút trong 1 tuần trước khi tuyên truyền và sau khi tuyên truyền được ghi lại dưới đây :

	Đối tượng											
	1	2	3	4	5	6	7	8	9	10	11	12
Trước	23	23	21	24	21	27	22	19	19	23	26	27
Sau	23	12	21	15	19	15	17	21	18	14	18	19

Sử dụng tiêu chuẩn dấu để kiểm định xem chiến dịch tuyên truyền chống hút thuốc lá có hiệu quả hay không ? Mức ý nghĩa  $\alpha = 5\%$ .

**189.** Cũng với bài toán 188 nhưng yêu cầu sử dụng tiêu chuẩn hạng có dấu Wilconxon.

**190.** Số trận mưa rào ở hai vùng A và B trong 25 tháng được ghi lại như sau :



Vùng A	10	11	7	8	8
Vùng B	7	5	4	8	9

Vùng A	7	4	5	8	7
Vùng B	3	2	8	1	2

Vùng A	11	12	11	7	5
Vùng B	8	10	7	2	6

Vùng A	6	7	9	12	10
Vùng B	5	9	3	8	7

Vùng A	10	9	8	5	6
Vùng B	8	7	4	2	7

a) Sử dụng tiêu chuẩn dấu, với mức ý nghĩa 1%, liệu có thể cho rằng vùng A hay mưa hơn vùng B hay không ?

b) Cũng với câu hỏi trên nhưng hãy sử dụng tiêu chuẩn hạng có dấu của Wilconxon.

**191.** Trong một cuộc thăm dò dư luận tiến hành trước ngày bầu cử, 42 trong số 100 cử tri nam được hỏi cho biết sẽ bỏ phiếu cho ứng viên A. Trong khi đó 92 trong số 200 cử tri nữ được hỏi cho biết sẽ bỏ phiếu cho ứng viên A. Với mức ý nghĩa 5%, kiểm định xem tỉ lệ cử tri nam bầu cho ông A và tỉ lệ cử tri nữ bầu cho ông A có như nhau hay không ?

**192.** Công ty nước giải khát Côca – Cola đang nghiên cứu việc đưa vào một công thức mới để cải tiến sản phẩm của mình. Với công thức cũ, khi cho 500 người dùng thử thì có 120 người tỏ ra ưa

thích nó. Với công thức mới, khi cho 1000 người khác dùng thử thì có 300 người tỏ ra ưa thích nó.

Hãy kiểm định xem liệu công thức mới đưa vào có làm tăng tỉ lệ những người ưa thích uống Côca – Cola hay không ?

Sử dụng mức ý nghĩa 2%.

**193.** Một công ty dược phẩm đang tiến hành thử nghiệm một loại thuốc mới chống lại sự rụng tóc của những người hói đầu. Công ty chọn ngẫu nhiên 150 người để thí nghiệm. Trước khi cho đối tượng dùng thuốc họ đếm số tóc mọc trên một vùng đầu. Sau khi dùng thuốc 1 tháng các đối tượng quay trở lại và được đếm lại số tóc mọc trong vùng đầu ấy. Kết quả cho thấy có 116 đối tượng mọc nhiều tóc hơn, 22 đối tượng có ít tóc hơn và 12 đối tượng có số tóc không thay đổi.

Dựa trên số liệu này, với mức ý nghĩa 2%, có thể kết luận rằng thuốc chống hói đầu có tác dụng hay không ?

**194.** Hai giáo sư A và B cùng dạy một môn ở hai trường đại học lớn. Trong số 400 sinh viên theo học giáo sư A có 80 sinh viên thi trượt. Trong số 500 sinh viên theo học giáo sư B có 125 sinh viên thi trượt. Với mức ý nghĩa  $\alpha = 5\%$  hãy kiểm định xem có sự khác nhau hay không về tỉ lệ trượt của các sinh viên theo học giáo sư A và giáo sư B ?

**195.** Trong một cuộc thăm dò trước bầu cử, 40 trong số 100 cử tri được hỏi cho biết rằng họ sẽ bỏ phiếu cho ông A. Một tuần sau, một cuộc thăm dò dư luận khác lại được tiến hành cho thấy có 68 trên tổng số 150 được hỏi ủng hộ ông A. Hỏi rằng tỉ lệ cử tri bỏ phiếu cho ông A có thay đổi không ? Mức ý nghĩa  $\alpha = 5\%$ .

**196.** Một cuộc nghiên cứu được tiến hành nhằm so sánh tỉ lệ học sinh bỏ học trước khi tốt nghiệp ở hai vùng dân cư A và B. Ở vùng A trong số 600 học sinh được theo dõi có 90 học sinh bỏ học

trước khi tốt nghiệp. Ở vùng B, trong số 400 học sinh được theo dõi có 48 em bỏ học trước khi tốt nghiệp. Với mức ý nghĩa 2%, có thể cho rằng hai tỉ lệ học sinh bỏ học ở vùng A và vùng B khác nhau hay không ?

**197.** Tạp chí Washington Post tiến hành một cuộc thăm dò dư luận sau cái chết của một ngôi sao bóng rổ. Có 729 người trong số 1656 người được hỏi nói rằng việc các vận động viên sử dụng các thuốc kích thích bất hợp pháp là một vấn đề nổi cộm trong thể thao. Một năm trước đó chỉ có 301 trong số 1432 người được hỏi là có ý kiến như vậy. Với mức ý nghĩa  $\alpha = 1\%$ , nhận định xem phải chăng đã có một sự thay đổi có ý nghĩa trong quan điểm của công chúng về vấn đề sử dụng chất kích thích ?

**198.** Trong mùa bóng vừa rồi, cầu thủ tiền đạo A ghi được 8 bàn trong 37 lần sút bóng. Cầu thủ B ghi được 7 bàn trong 38 lần sút bóng. Hỏi có sự khác nhau về tỉ lệ ghi bàn của hai cầu thủ A và B hay không ? Mức ý nghĩa  $\alpha = 5\%$ .

## Chương 6

### BÀI TOÁN SO SÁNH MỞ RỘNG

**199.** So sánh tác dụng của 6 mẫu thuốc thử nghiệm trên 6 lô chuột, kết quả thu được như sau :

Mẫu thuốc	1	2	3	4	5	6	Tổng
Số sống	79	82	77	83	76	81	478
Số chết	21	18	23	17	24	19	122
Tổng số	100	100	100	100	100	100	600

Với mức ý nghĩa 5% hãy kiểm định giả thiết :

$H_0$  : "Tỉ lệ chết trong 6 mẫu thuốc như nhau"

$H_1$  : "Tỉ lệ chết trong 6 mẫu thuốc khác nhau".

**200.** Có 4 thầy giáo A, B, C, D cùng dạy một giáo trình Thống kê. Ban chủ nhiệm khoa muốn tìm hiểu chất lượng dạy của 4 thầy này nên đã làm một cuộc khảo sát. Kết quả thu được như sau :

Thầy Kết quả thi	A	B	C	D	Tổng
Đạt	60	75	150	125	410
Không đạt	40	75	50	75	240
Tổng	100	150	200	200	650

Với mức ý nghĩa  $\alpha = 1\%$ , có thể cho rằng tỉ lệ đỗ của các học sinh đã học các thầy A, B, C, D là như nhau được không ?

**201.** Một hãng sản xuất ô tô tiến hành một nghiên cứu nhằm xác định xem có sự khác nhau giữa đàn ông và đàn bà trong việc chọn mua một trong 3 loại ô tô của hãng hay không. Kết quả thu được như sau :

Loại ô tô	A	B	C
Phụ nữ	70	80	150
Nam giới	40	60	100

Với mức ý nghĩa 5%, ta cần rút ra kết luận gì ?

**202.** Ở thành phố X người ta tiến hành một cuộc khảo sát để xác định phương tiện giao thông mà các công nhân viên chức sử dụng khi đến chỗ làm việc. Việc điều tra được tiến hành trên hai nhóm đàn ông và đàn bà. Kết quả thu được như sau :

Loại phương tiện	Xe máy	Ô tô buýt	Xe đạp
Phụ nữ	25	100	125
Nam giới	75	120	205

Với mức ý nghĩa  $\alpha = 1\%$ , hãy nhận định xem có sự khác nhau về cơ cấu sử dụng phương tiện giao thông khi đến chỗ làm việc giữa hai nhóm phụ nữ và nam giới hay không ?

**203.** Một xí nghiệp may sản xuất áo khoác với 4 màu : đỏ, xanh, vàng và tím than. Số khách hàng nam và nữ mua áo khoác với các màu nói trên được thống kê trong bảng dưới đây :

Màu áo	Đỏ	Xanh	Vàng	Tím than
Nữ	62	34	71	42
Nam	125	223	52	54

Với mức ý nghĩa  $\alpha = 1\%$ , hãy so sánh tỉ lệ khách hàng nam và nữ ưa chuộng một trong các màu sắc nói trên.

204. Trong một thí nghiệm nghiên cứu ung thư, 162 con chuột được ghép khối u và chia thành 4 nhóm. Mỗi nhóm được chiếu xạ với cường độ khác nhau. Bảng sau đây ghi lại kết quả thí nghiệm.

Mức \ Kết quả	1	2	3	4
Hết khối u	10	32	37	32
Không hết	32	9	2	8

a) Với mức ý nghĩa  $\alpha = 1\%$ , hãy so sánh tỉ lệ chữa khỏi ung thư của 4 nhóm nói trên.

b) Với mức ý nghĩa 5%, hãy so sánh tỉ lệ chữa khỏi ung thư của 3 nhóm 2, 3 và 4.

204. Phân bố giới tính của trẻ em sinh ra ở Thụy Điển trong năm 1938 được thống kê như sau :

Tháng 1 :	3743 bé trai,	3537 bé gái
Tháng 2 :	2550 bé trai,	3407 bé gái
Tháng 3 :	4017 bé trai,	3866 bé gái
Tháng 4 :	4173 bé trai,	3711 bé gái
Tháng 5 :	4117 bé trai,	3775 bé gái
Tháng 6 :	3944 bé trai,	3665 bé gái
Tháng 7 :	3964 bé trai,	3621 bé gái
Tháng 8 :	3797 bé trai,	3596 bé gái
Tháng 9 :	3712 bé trai,	3491 bé gái
Tháng 10 :	3812 bé trai,	3391 bé gái
Tháng 11 :	3392 bé trai,	3160 bé gái
Tháng 12 :	3761 bé trai,	3371 bé gái

Hãy kiểm định giả thiết cho rằng xác suất sinh con trai là như nhau trong tất cả các tháng. Mức ý nghĩa là 5%.

**205.** Bảng sau đây cho ta phân bố thu nhập của hai nhóm tuổi : Nhóm từ 40 đến 50 tuổi và nhóm từ 50 đến 60 tuổi trong tất cả các viên chức ở Thụy Điển năm 1930 (Đơn vị là 1000 curon) :

Nhóm tuổi	Thu nhập					
	0 – 1	1 – 2	2 – 3	3 – 4	4 – 6	≥ 6
Từ 40 – 50	7831	26740	35572	20009	11527	6919
Từ 50 – 60	7858	20685	24186	12280	6776	4222

Với mức ý nghĩa 1%, nhận định xem liệu có sự khác nhau về phân bố thu nhập giữa hai nhóm tuổi này không ?

**206.** Bảng sau đây cho ta phân bố về màu tóc của 196 đàn ông và 226 đàn bà :

Màu tóc	Đàn ông	Đàn bà
Đen	56	32
Vàng nhạt	37	66
Nâu	84	90
Đỏ	19	38

a) Hãy tính các tần số lí thuyết.

b) Với mức 1%, nhận định xem phân bố màu tóc có như nhau trong hai giới nam và nữ hay không ?

**207.** Bảng sau đây cho ta phân bố thu nhập của hai nhóm tuổi : Nhóm từ 40 – 50 tuổi và nhóm từ 50 – 60 tuổi trong số các công nhân lành nghề ở Thụy Điển vào năm 1930.

Nhóm tuổi	Thu nhập					
	0 – 1	1 – 2	2 – 3	3 – 4	4 – 6	≥ 6
Từ 40 – 50	71	430	1072	1609	1178	158
Từ 50 – 60	54	324	894	1202	963	112

Với mức ý nghĩa 5%, hãy nhận định xem liệu có sự khác nhau về phân bố thu nhập giữa hai nhóm tuổi này hay không ?

**208.** Người ta tiến hành một cuộc điều tra xã hội học ở 5 thành phố A, B, C, D, E. Những người tham gia phỏng vấn được yêu cầu diễn tả mức độ thỏa mãn của mình đối với thành phố mà họ đang sống. Kết quả thu được cho như sau :

Thành phố	Mức độ thỏa mãn		
	Rất thỏa mãn	Tương đối thỏa mãn	Không thỏa mãn
A	220	121	63
B	130	207	75
C	84	54	24
D	156	95	43
E	122	164	73

Với mức ý nghĩa 5%, kiểm định xem mức độ thỏa mãn cuộc sống có phân bố như nhau trong 5 thành phố nói trên hay không ?

**209.** Một nhà nghiên cứu muốn so sánh số băng trên vỏ của ba loài ốc sên rừng I, II và III. Quan sát 169 con ốc sên loại I, 100 con ốc sên loại II và 374 con ốc sên loại II và 374 con ốc sên loại III, ông ta ghi được số liệu sau :

Số băng trên vỏ	Loài			Tổng số
	I	II	III	
0	49	31	126	206
1 hoặc 2	33	20	56	109
3 hoặc 4	52	20	83	155
5 trở lên	35	29	109	173
Tổng số	169	100	374	643

Với mức ý nghĩa 5%, hãy nhận định xem số băng trên vỏ của ba loài ốc sên này có phân bố giống nhau hay không ?



**210.** Đài truyền hình Việt Nam muốn thăm dò ý kiến khán giả về thời lượng phát sóng phim truyện Việt Nam hàng tuần. Phiếu thăm dò đặt ra 4 mức :

$A_1$  : Tăng thời lượng phát sóng

$A_2$  : Giữ như cũ

$A_3$  : Giảm

$A_4$  ; Không có ý kiến

Đài truyền hình tổ chức thăm dò ý kiến ở ba nhóm xã hội khác nhau : Công nhân, Nông dân, Trí thức. Kết quả cuộc thăm dò được cho trong bảng sau đây :

<div>Tầng lớp</div> <div>Ý kiến</div>	Công nhân	Nông dân	Trí thức
Tăng	100	300	20
Như cũ	200	400	30
Giảm	50	80	5
Không ý kiến	30	70	5

Với mức ý nghĩa 5%, hãy nhận định xem có sự khác nhau giữa các tầng lớp xã hội nói trên xung quanh vấn đề thời lượng phát sóng của phim truyện hay không ?

**211.** Bệnh đau mắt hột được chia làm 4 thời kì  $T_1$ ,  $T_2$ ,  $T_3$  và  $T_4$ . Kết quả kiểm tra mắt hột ở ba tỉnh A, B, C cho ta số liệu sau đây :

Địa phương	Mức độ đau mắt hột			
	$T_1$	$T_2$	$T_3$	$T_4$
A	47	189	807	1768
B	53	746	1387	946
C	16	228	438	115

Với mức ý nghĩa 1%, hãy nhận định xem tình hình đau mắt hột ở ba tỉnh A, B, C giống nhau hay khác nhau.

**212.** Bảng sau đây cho số liệu người chết vì bệnh ung thư ở ba nước A, B, C. Người chết về bệnh ung thư được phân loại theo cơ quan bị ung thư.

Tên nước	Cơ quan bị ung thư			
	Ruột	Ngực	Dạ dày	Các bộ phận khác
A	11	15	3	41
B	5	3	22	30
C	5	7	3	15

a) Hãy tính tần số lí thuyết của bảng số liệu trên.

b) Có thể áp dụng tiêu chuẩn  $\chi^2$  để phân tích bảng số liệu nói trên được không ?

c) Với mức ý nghĩa 1%, hãy so sánh phân bố tỉ lệ chết về ung thư của ba nước A, B, C.

**213.** Để so sánh tác dụng của hai loại vaccin A và B, người ta tiêm chủng cho 348 trẻ em được chia làm hai nhóm. Một nhóm tiêm vaccin A, nhóm kia tiêm vaccin B. Kết quả về phản ứng được cho trong bảng sau :

Vaccin	Mức độ phản ứng		
	Nhẹ	Trung bình	Mạnh
A	12	156	9
B	29	135	7

a) Hãy tính các tần số lí thuyết.

b) Với mức ý nghĩa 5% hãy so sánh hai loại vaccin A và B trên khía cạnh gây phản ứng.

**214.** Một nông trường nuôi bò ba giống bò sữa A, B, C. Lượng sữa của các con bò này được thống kê trong bảng dưới đây :

Loại bò	Lượng sữa		
	Ít	Trung bình	Nhiều
A	92	37	46
B	53	15	19
C	75	19	12

Với mức ý nghĩa  $\alpha = 0,025$ , hãy nhận định xem có phải ba giống bò này là thuần nhất như nhau về phương diện sản lượng sữa hay không ?

**215.** Điểm thi của 12 sinh viên theo học các giáo sư A, B, C, được cho trong bảng dưới đây (thang điểm 100) :

Giáo sư A	Giáo sư B	Giáo sư C
79	71	82
86	77	68
94	81	70
89	83	76

Với mức ý nghĩa 5%, hãy nhận định xem liệu điểm thi trung bình của tất cả các sinh viên theo học các giáo sư A, B, C có giống nhau hay không ?

**216.** Sản lượng (tính bằng tạ/ha) của 4 giống lúa cao sản A, B, C, D trồng trên các thửa ruộng thí nghiệm được cho ở bảng sau đây :

Giống lúa			
A	B	C	D
56	61	58	68
64	66	60	74
67	52	65	59
61	48	49	54
70	47	75	66
	56		64

Với mức ý nghĩa 5%, nhận định xem năng suất của 4 giống lúa này như nhau hay khác nhau.

**217.** Để so sánh ba loại thuốc bổ A, B, C ta có 3 nhóm súc vật I, II, III chọn tương đối thuần nhất và được nuôi dưỡng trong những điều kiện như nhau. Nhóm thứ I có 6 con, nhóm thứ II có 8 con và nhóm thứ III có 7 con. Qua một thời gian theo dõi sự tăng trọng  $x_{ij}$  của các súc vật này ta tính được :

$$T_1 = \sum x_{j1} = 61$$

$$T_2 = \sum x_{j2} = 126$$

$$T_3 = \sum x_{j3} = 25$$

và 
$$Q = \sum \sum x_{ij}^2 = 2854$$

Với mức ý nghĩa 1%, hãy nhận định xem ba loại thuốc bổ nói trên có tác dụng giống nhau hay khác nhau.

**218.** Người ta tiến hành đo kích thước hộp sọ của ba nhóm người. Nhóm thứ nhất có 83 hộp sọ, nhóm thứ hai có 51 hộp sọ và nhóm thứ ba có 8 hộp sọ. Số liệu tính được là

$$T_1 = \sum x_{j1} = 11277$$

$$T_2 = \sum x_{j2} = 7049$$

$$T_3 = \sum x_{j3} = 1102$$

và 
$$\sum x_{ij}^2 \frac{T^2}{n} = 4616,64$$

Với mức ý nghĩa 5%, so sánh kích thước trung bình của hộp sọ của ba nhóm người nói trên.

**219.** Tuổi thọ bóng đèn điện tử được sản xuất từ 4 nhà máy A, B, C, D được kiểm tra và cho kết quả sau đây :

Nhà máy			
A	B	C	D
1600	1580	1460	1510
1610	1640	1550	1520
1650	1640	1600	1530
1680	1700	1620	1570
1700	1750	1640	1600
1700		1660	1680
1800		1740	
		1820	

Với mức ý nghĩa 5%, hãy so sánh chất lượng bóng đèn của ba nhà máy nói trên.

**220.** Để nghiên cứu mối quan hệ giữa thời điểm mắc bệnh và mức độ bệnh, người ta chia 38 con chuột thí nghiệm ra làm 4 nhóm, mỗi nhóm chuột cho tiêm virus ở những thời điểm khác nhau. Tất cả các con chuột sau một thời gian đều chết. Thời gian sống của chúng từ lúc tiêm cho tới lúc chết được ghi lại ở bảng sau (đơn vị thời gian là giờ).

Thời điểm tiêm			
8 giờ	16 giờ	20 giờ	24 giờ
3,33	3,29	3,33	3,37
3,40	3,36	3,36	3,53
3,43	3,40	3,46	3,76
3,49	3,43	3,52	4,10
3,63	3,68	3,61	4,13
3,76	3,82	3,71	4,22
3,85	3,87	3,87	4,28
3,97	4,20	4,09	4,55
4,40	4,49	4,18	4,69
4,24			5

Với mức ý nghĩa 5%, nhận định xem khả năng chống bệnh của chuột có liên quan tới thời điểm mà chuột bị tiêm hay không ?

**221.** Hãy tiến hành phân tích phương sai đối với bảng số liệu sau đây :

*Mẫu thứ nhất :* 22, 19, 13, 19, 23, 15, 16, 18, 20, 20

*Mẫu thứ hai :* 27, 25, 22, 27, 19, 23, 21, 28, 23, 25, 27

*Mẫu thứ ba :* 20, 18, 21, 21, 16, 17, 20, 18, 17, 19, 18

*Mẫu thứ tư :* 18, 16, 24, 19, 22, 22, 24

**222.** Một nhà tâm lí học muốn nghiên cứu ảnh hưởng của quê quán đối với thời gian (tính bằng phút) để một sinh viên có thể trả lời một câu đố. 4 nhóm sinh viên từ 4 vùng khác nhau (nội thành, ngoại thành, thị trấn, nông thôn) được tiến hành khảo sát và kết quả thu được như sau :

*Nội thành :* 16,5 5,2 12,1 14,3

*Ngoại thành :* 10,9 5,2 10,8 8,9 16,1

*Thị trấn :* 18,6 8,1 6,4

*Nông thôn :* 14,2 24,5 14,8 24,9 5,1

Hãy tiến hành phân tích phương sai bảng số liệu nói trên

**223.** Có 4 báo cáo viên A, B, C, D đi báo cáo về cùng một chủ đề ở các địa điểm khác nhau. Sau đây là thời gian (tính bằng phút) mà mỗi báo cáo viên đó sử dụng trong 5 buổi báo cáo :

A	B	C	D
25	35	30	28
29	20	27	32
30	20	18	33
42	17	19	35
35	30	26	24

Với mức ý nghĩa 5%, hãy so sánh thời gian báo cáo trung bình của 4 báo cáo viên nói trên.

**224.** Hãy tiến hành phân tích phương sai cho bảng số liệu dưới đây :

I	II	III	IV
12	19	34	19
16	20	29	21
12	18	31	17
14	9	19	24
26	22	26	
	19		

## Chương 7

# PHÂN TÍCH TƯƠNG QUAN VÀ HỒI QUY

**225.** Một nhà khoa học quan tâm tới vấn đề về mối tương quan giữa tuổi và mạch đập của người phụ nữ. Trong một mẫu quan sát gồm 5 phụ nữ được chọn một cách ngẫu nhiên, ông ta thu được số liệu sau :

Tuổi (X)	Mạch đập (Y)
23	210
39	185
19	220
44	164
51	123

Tính hệ số tương quan của hai biến X và Y.

**226.** Tính hệ số tương quan của hai biến X, Y căn cứ trên mẫu quan sát dưới đây :

X	Y	X	Y
80	2,4	92	3
85	2,8	82	2,5
88	2,3	75	2,3
90	3,1	78	2,8
95	3,7	85	3,1

**227.** Tính hệ số tương quan của các tập số liệu sau :

a)

X	2	5	7	10	11
Y	10	20	35	50	65



b)

X	5	8	10	12	15	22
Y	50	42	40	35	33	28

**228.** Một công ty nhỏ quan tâm tới việc phân tích hiệu quả của việc quảng cáo. Trong thời gian 5 tháng công ty thu được kết quả sau :

X	5	8	10	15	22
Y	6	15	20	30	39

trong đó X là số tiền chi vào quảng cáo (đơn vị là triệu), còn Y là tổng doanh thu (đơn vị là chục triệu). Hãy xác định hệ số tương quan.

**229.** Một trường đại học thu thập các số liệu về số chứng chỉ mà một sinh viên theo học và số giờ học ở nhà của anh ta trong một tuần. Kết quả cho thấy như sau :

X	20	25	30	50	20	23
Y	12	13	12	15	16	16

trong đó X là số giờ học, Y là số chứng chỉ. Hãy tính hệ số tương quan. Ở mức ý nghĩa  $\alpha = 5\%$ , có sự tương quan tuyến tính giữa hai biến đó không ?

**230.** Một cuộc nghiên cứu được tiến hành ở Mỹ nhằm xác định mối quan hệ giữa chiều cao của một người và cỡ giày của họ. Nhà nghiên cứu đã thu được số liệu sau :

X	66	63	67	71	62
Y	9	7	8,5	10	6

X	65	72	68	60	66
Y	8,5	12	10,5	5,5	8

trong đó X là chiều cao (đơn vị đo : inch) còn Y là cỡ giày.

Hãy tính hệ số tương quan giữa X và Y.

231. Cho mẫu quan sát sau đây của cặp ĐLNN (X, Y) :

X	8	8	12	12	20	20
Y	82	78	65	50	60	47

X	24	24	8	8	8	12
Y	52	41	87	58	70	62

X	12	12	20	20	20	24
Y	55	52	44	66	41	57

X	24	24	8	12	20	24
Y	50	47	65	49	57	63

Hãy tính hệ số tương quan, hệ số xác định và tỉ số tương quan của Y đối với X.

232. Với mức ý nghĩa  $\alpha = 5\%$ , hãy kiểm định xem liệu có tương quan phi tuyến giữa Y và X hay không ? trong đó (X, Y) là mẫu quan sát đã nêu ở bài tập 231.

233. Trên 100 thửa ruộng thí nghiệm kích thước như nhau, người ta bón phân với số lượng khác nhau đã nghiên cứu mối quan hệ giữa lượng phân bón và sản lượng. Kí hiệu X là lượng phân, còn Y là sản lượng. Kết quả được cho ở bảng dưới đây :

X \ Y	1	2	3	4	5
14	10	8	7	6	9
15		12		8	
16			28		
17					
18					12

a) Tính sản lượng trung bình của mỗi mức phân và sản lượng trung bình chung.

b) Tính hệ số tương quan, hệ số xác định và tỉ số tương quan.

c) Lập phương trình hồi quy tuyến tính của Y đối với X.

**234.** Một nhà xã hội học tuyên bố rằng kết quả học tập của sinh viên (do bằng điểm thi tốt nghiệp) không có liên quan tới thu nhập của gia đình họ. Với một mẫu ngẫu nhiên gồm 20 sinh viên người ta tính được hệ số tương quan là 0,4. Với mức ý nghĩa 1% hãy kiểm định giả thiết : "Điểm thi tốt nghiệp và thu nhập của gia đình một sinh viên là không tương quan".

**235.** Khảo sát 20 trận đấu bóng đá cho thấy hệ số tương quan giữa số lần sút bóng vào khung thành đối phương và số bàn thắng là 0,21. Với mức ý nghĩa 5%, kiểm định giả thiết  $H_0$  : "Số lần sút bóng và số bàn thắng không tương quan".

**236.** Từ 13 cặp quan sát  $(x_i, y_i)$  sau đây về cặp ĐLNN  $(X, Y)$  hãy kiểm định giả thiết :

$H_0$  : "X và Y không tương quan với mức ý nghĩa 5%".

X	123	358,9	111,1	118
Y	4,15	4,10	3,74	3,89

X	265	390	67	24
Y	3,75	3,10	3,13	2,97

X	156,9	106	86	85,5
Y	2,96	2,92	2,33	2,26

X	241,3			
Y	2,03			

**237.** Tính tỉ số tương quan của Y đối với X và hệ số xác định từ 10 cặp quan sát sau đây của cặp ĐLNN (X, Y) :

X	15	25	10	15	20	10
Y	13	22	6	17	21	10

X	20	25	30	30		
Y	25	18	14	10		

Có nhận xét gì về mối quan hệ giữa X và Y ?

**238.** Tính tỉ số tương quan của Y đối với X và hệ số xác định từ mẫu quan sát sau đây của cặp ĐLNN (X, Y) :

X	50	130	170	270	90
Y	15	115	215	335	95

X	210	50	130	270	240
Y	295	55	155	295	315

X	170	210	90	210	90
Y	175	275	75	255	115

X	240	50	240	270	130
Y	255	35	275	315	135

X	270	90	240	130	170
Y	355	135	295	175	235

X	50	170	195		
Y	75	210	235		

Có kết luận gì về mối tương quan giữa X và Y (tương quan phi tuyến hay tuyến tính) ?

**239.** Tính tỉ số tương quan của Y đối với X và hệ số xác định từ mẫu quan sát sau đây của cặp ĐLNN (X, Y) :

X	50	130	210	240	90
Y	75	235	255	195	115

X	210	240	50	240	170
Y	295	215	35	235	295

X	270	210	50	240	90
Y	135	315	55	175	135

X	90	270	130	50	270
Y	175	115	255	15	75

X	130	170	210	270	170
Y	295	355	275	95	335

X	170	90	130		
Y	315	155	275		

**240.** Ở các cây ngọc trâm, lá có hai dạng "lá phẳng hoặc lá nhân". Hoa cũng có hai dạng "Hoa bình thường hoặc hoa hoàng hậu". Quan sát một mẫu gồm 560 cây ngọc trâm ta thu được kết quả sau :

<div style="text-align: center;">Hoa Lá</div>	Bình thường	Hoàng hậu	Tổng số
Thẳng	328	122	450
Nhân	77	33	110
Tổng số	405	155	560

Có thể chấp nhận rằng hai đặc tính về hoa và lá nói trên là độc lập hay không ?

**241.** Một con ốc sên rừng có thể có màu vỏ là vàng hoặc hồng. Số vạch trên vỏ của nó có thể là 0, 1, 2, 3, 4, 5. Xem xét một mẫu ngẫu nhiên gồm 169 con ốc sên rừng ta thu được số liệu sau đây :

<div>Số vạch Màu vỏ</div>	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>	Tổng số
Vàng	35	19	36	25	115
Hồng	14	14	16	10	54
Tổng số	49	33	52	35	169

Với mức ý nghĩa 5%, kiểm định xem hai đặc tính "màu vỏ" và "số vạch trên vỏ" có mối tương quan nào đó hay không ?

**242.** Giả sử X và Y tương ứng là số đo huyết áp và trọng lượng của trẻ em 14 tuổi. (Đơn vị đo trọng lượng ở đây là pound, 1 pound = 0,454 kg). Một mẫu gồm 200 trẻ em được đo huyết áp và trọng lượng cho ta số liệu sau đây :

<div>Huyết áp Trọng lượng</div>	Nhỏ hơn 99	Từ 99 tới 100	Từ 110 đến 120	Từ 120 trở lên
Nhỏ hơn 102	10	20	11	5
Lớn hơn hay bằng 102	6	48	50	50

Với mức ý nghĩa 1%, hãy nhận định xem huyết áp và trọng lượng có độc lập với nhau hay không ?

**243.** Một nhà xã hội học muốn tìm hiểu mối quan hệ giữa các dạng tội phạm (hình sự và không hình sự) với tuổi tác của phạm nhân. Chọn ngẫu nhiên 100 phạm nhân trong hồ sơ của tòa án, ông ta thu được số liệu sau :

Loại tội phạm	Tuổi		
	dưới 25	Từ 25 đến 50	Trên 50
Hình sự	15	30	10
Không hình sự	5	30	10

Với mức ý nghĩa 5%, căn cứ trên số liệu nói trên có thể kết luận rằng có mối tương quan giữa tuổi và loại hình tội phạm hay không ?

**244.** Một nhà xã hội học chọn một mẫu ngẫu nhiên gồm 500 người và trao cho mỗi người một bản thăm dò ý kiến với các câu hỏi sau đây :

- Ông (bà) có đi nhà thờ hay không ? (Trả lời : có hoặc không).
- Mức độ thành kiến chủng tộc của ông bà tự đánh giá thế nào ? (Trả lời : rất cao, đôi khi hoặc không chút nào).

Kết quả tổng hợp lại được cho ở bảng sau đây :

Đi nhà thờ	Mức độ thành kiến chủng tộc		
	Cao	Đôi khi	Không
Có	70	160	170
Không	20	50	30

Với mức ý nghĩa 5%, hãy nhận định xem có mối tương quan giữa việc có đi nhà thờ và vấn đề thành kiến chủng tộc hay không ?

**245.** Một cuộc thăm dò được tiến hành bởi viện nghiên cứu xã hội nổi tiếng Gallup (Mỹ) để nghiên cứu mối quan hệ giữa nghề nghiệp của một người với quan niệm của anh ta về tiêu chuẩn đạo đức và tính nghề nghiệp. Khảo sát một mẫu ngẫu nhiên gồm 380 người cho ta số liệu thống kê sau đây :

Nghề nghiệp	Quan điểm về tiêu chuẩn đạo đức, nghề nghiệp		
	Cao	Trung bình	Thấp
Bác sĩ	53	35	10
Luật sư	24	43	27
Nhà kinh doanh	18	55	20
Nhà chính trị	14	43	38

Với mức ý nghĩa 5%, viện nghiên cứu Gallup cần rút ra kết luận gì ?

**246.** Bảng sau đây cho ta về số liệu màu tóc của 422 người châu Âu :

Màu tóc	Giới tính	
	Nam	Nữ
Đen	56	32
Hung	37	66
Nâu	84	90
Vàng	19	38

Với mức ý nghĩa 1%, nhận định xem liệu có mối quan hệ giữa màu tóc và giới tính hay không ?

**247.** Một nhóm gồm 105 nhà doanh nghiệp Mỹ được phân loại căn cứ theo thu nhập hàng năm và tuổi của họ để nghiên cứu mối tương quan giữa tuổi và mức thu nhập. Kết quả phân loại được cho ở bảng dưới đây :

Tuổi	Thu nhập hàng năm		
	Dưới 10 vạn USD	Từ 10 vạn đến 40 vạn USD	Trên 40 vạn USD
Dưới 40	6	9	5
Từ 40 đến 54	18	19	8
Trên 54	11	12	17

Với mức ý nghĩa 1%, hãy kiểm định giả thiết rằng tuổi và mức thu nhập là độc lập với nhau.

**248.** Các số liệu về số trang của một cuốn sách (X) và giá bán của nó (Y) được cho trong bảng sau đây :

Tên sách	X	Y (nghìn đồng)
A	400	44
B	600	47
C	500	48
D	600	48
E	400	43
F	500	46



a) Hãy tìm đường thẳng hồi quy của Y theo X căn cứ trên số liệu nói trên.

b) Hãy tính sai số tiêu chuẩn  $s_{Y.X}$ .

**249.** Một mẫu gồm 7 sinh viên được chọn để nghiên cứu mối quan hệ giữa điểm thi vào đại học (X) và điểm thi ở kì thi cuối năm thứ nhất (Y) (thang điểm : 5).

X	2	1	3	3	4	4	4
Y	2.5	2	2	3	3,5	4	4

a) Tính hệ số tương quan.

b) Tìm phương trình hồi quy tuyến tính của Y theo X.

c) Nếu điểm thi vào đại học của một sinh viên là 3,3 thì dự đoán số điểm thi cuối năm thứ nhất của anh ta là bao nhiêu ?

**250.** Bảng sau đây cho ta tuổi (tính theo năm) và chiều cao thân răng của 15 con lừ. Tuổi được kí hiệu là X và chiều cao thân răng được kí hiệu là Y.

X	0,06	0,08	0,08	0,18
Y	5,05	4,95	5,10	4,85

X	0,28	0,36	0,56	0,53
Y	4,40	5,25	4,40	5,15

X	1,02	1,23	1,77	1,94
Y	4,3	4,15	3,35	3,40

X	2,14	2,44	2,44	
Y	3,05	2,65	3,00	

a) Tìm phương trình đường hồi quy của Y đối với X.

b) Tính sai số tiêu chuẩn  $s_{Y.X}$ .

c) Kiểm định giả thiết  $H_0$  : "Hệ số hồi quy lí thuyết bằng 0" với mức ý nghĩa 5%.

**251.** Đo đường kính  $X$  và chiều cao  $Y$  của 20 cây ta thu được số liệu sau :

X	2,3	2,5	2,6	3,1	3,4
Y	7	8	4	4	6

X	3,7	7,3	3,9	4	4,1
Y	6	14	12	8	5

X	4,1	4,2	4,4	4,7	5,1
Y	7	8	7	9	10

X	5,5	5,8	6,2	6,9	6,9
Y	13	7	11	11	16

- Tính đường hồi quy của  $Y$  đối với  $X$ .
- Tính sai số tiêu chuẩn  $s_{Y.X}$ .
- Kiểm định giả thiết cho rằng có hồi quy tuyến tính của  $Y$  theo  $X$ , với mức ý nghĩa 1%.

**252.** Bảng sau đây cho ta một mẫu quan sát gồm 11 quan sát  $(x_i, y_i)$  từ tập hợp chính gồm các giá trị của cặp ĐLNN  $(X, Y)$  :

X	0,9	1,22	1,32	0,77	1,3
Y	-0,3	0,1	0,7	-0,28	-0,25

X	1,2	1,32	0,95	1,45	1,3	1,2
Y	0,02	0,37	-0,70	0,55	0,35	0,32

- Tìm đường hồi quy của  $Y$  đối với  $X$ .
- Tính sai số tiêu chuẩn của đường hồi quy.
- Với mức ý nghĩa 1%, hãy kiểm định giả thiết cho rằng có hồi quy tuyến tính của  $Y$  đối với  $X$ .

**253.** Bảng sau đây ghi lại kỉ lục thế giới về chạy 1 dặm trong thế kỉ 20 :

Năm	Thời gian chạy
1911	4 phút 15,4 giây
1913	4 phút 14,6 giây
1915	4 phút 12,6 giây
1923	4 phút 10,4 giây
1931	4 phút 9,2 giây
1933	4 phút 7,6 giây
1934	4 phút 6,8 giây
1937	4 phút 6,4 giây
1942	4 phút 6,2 giây
1942	4 phút 4,6 giây
1943	4 phút 2,6 giây
1944	4 phút 1,6 giây
1945	4 phút 1,4 giây
1954	3 phút 59,4 giây
1954	3 phút 58 giây
1957	3 phút 57,2 giây
1958	3 phút 54,5 giây
1962	3 phút 54,4 giây
1964	3 phút 54,1 giây
1965	3 phút 53,6 giây
1966	3 phút 51,3 giây
1967	3 phút 51,1 giây
1975	3 phút 49,4 giây
1979	3 phút 49 giây

1980	3 phút 48,9 giây
1981	3 phút 48,8 giây
1981	3 phút 48,7 giây
1981	3 phút 47,6 giây
1985	3 phút 46,5 giây

- Tính hệ số tương quan giữa X và Y (đơn vị của Y là giây).
- Tìm đường thẳng hồi quy giữa Y và X.
- Tìm sai số tiêu chuẩn của đường thẳng hồi quy. Kiểm định giả thiết : "Có mối tương quan giữa Y và X".
- Dựa trên phương trình hồi quy hãy ước lượng kỉ lục thế giới năm 2050.
- Ước lượng năm mà kỉ lục thế giới sẽ là 3 phút 30 giây.
- Dựa trên phương trình hồi quy hãy ước lượng kỉ lục thế giới năm 2500. Kết quả này có hợp lý hay không ?

**254.** Người ta cho 5 bệnh nhân ở các lứa tuổi khác nhau uống cùng một loại thuốc và theo dõi thời gian phân hủy hoàn toàn loại thuốc đó trong cơ thể mỗi người. Kết quả ghi lại như sau :

Tuổi (X) (tính bằng năm)	Thời gian phân hủy thuốc (phút) (Y)
30	15
25	28
65	30
50	22
40	24

- Tính đường hồi quy của Y đối với X.
- Nêu ý nghĩa của các hệ số hồi quy.
- Dự báo thời gian phân hủy thuốc của một bệnh nhân 42 tuổi.

**255.** Cho bảng số liệu sau đây :

X	73	82	90	60	51	40
Y	51	63	67	35	23	20

- Tìm đường hồi quy tuyến tính của Y đối với X.
- Tính sai số tiêu chuẩn  $s_{Y.X}$ .
- Dự báo giá trị trung bình của Y khi  $X = 80$ .
- Tìm khoảng tin cậy 95% cho giá trị trung bình của Y khi  $X = 80$ .
- Tìm khoảng tin cậy cho giá trị Y khi  $X = 80$ .

**256.** Một nhà nghiên cứu tội phạm học quan tâm tới việc dự báo tuổi của một phạm nhân khi bị tổng giam cầm cứ theo tuổi của anh ta khi lần đầu tiên phạm pháp. Nhà nghiên cứu đã thu thập được số liệu sau đây :

Tuổi lần đầu phạm pháp X	Tuổi khi bị bắt giam Y
11	21
17	20
13	20
12	19
15	18
10	23
12	20

- Tìm đường hồi quy tuyến tính của Y theo X.
- Tìm sai số tiêu chuẩn  $s_{Y.X}$ .
- Ước lượng tuổi bị tổng giam của một phạm nhân khi biết rằng anh ta lần đầu phạm pháp ở tuổi 14. Tìm khoảng tin cậy 95% cho giá trị dự báo này.

## Phần B

# LỜI GIẢI VÀ ĐÁP SỐ

---

### Chương 1

## THỐNG KÊ MÔ TẢ

1.

Khoảng	Tần số
49,5 – 124,5	3
124,5 – 199,5	5
199,5 – 274,5	7
274,5 – 349,5	5
349,5 – 424,5	3
424,5 – 499,5	2
Tổng số	25

2.

Khoảng	Tần số
35,5 – 40,5	1
40,5 – 45,5	3
45,5 – 50,5	4
50,5 – 55,5	3
55,5 – 60,5	7
60,5 – 65,5	4
65,5 – 70,5	2
70,5 – 75,5	2
75,5 – 80,5	3
80,5 – 85,5	1
Tổng số	30

3.

Khoảng	Tần số
49,5 – 124,5	0,12
124,5 – 199,5	0,20
199,5 – 274,5	0,28
274,5 – 349,5	0,20
349,5 – 424,5	0,12
424,5 – 499,5	0,08
Tổng số	1

4.

Khoảng	Tần số
749,5 – 849,5	3
849,5 – 949,5	9
949,5 – 1049,5	7
1049,5 – 1149,5	2
1149,5 – 1249,5	1
1249,5 – 1349,5	3
Tổng số	25

5.

Khoảng	Tần số
119,5 – 134,5	2
134,5 – 149,5	5
149,5 – 164,5	3
164,5 – 179,5	8
179,5 – 194,5	5
194,5 – 209,5	2
209,5 – 225,5	1
Tổng số	26

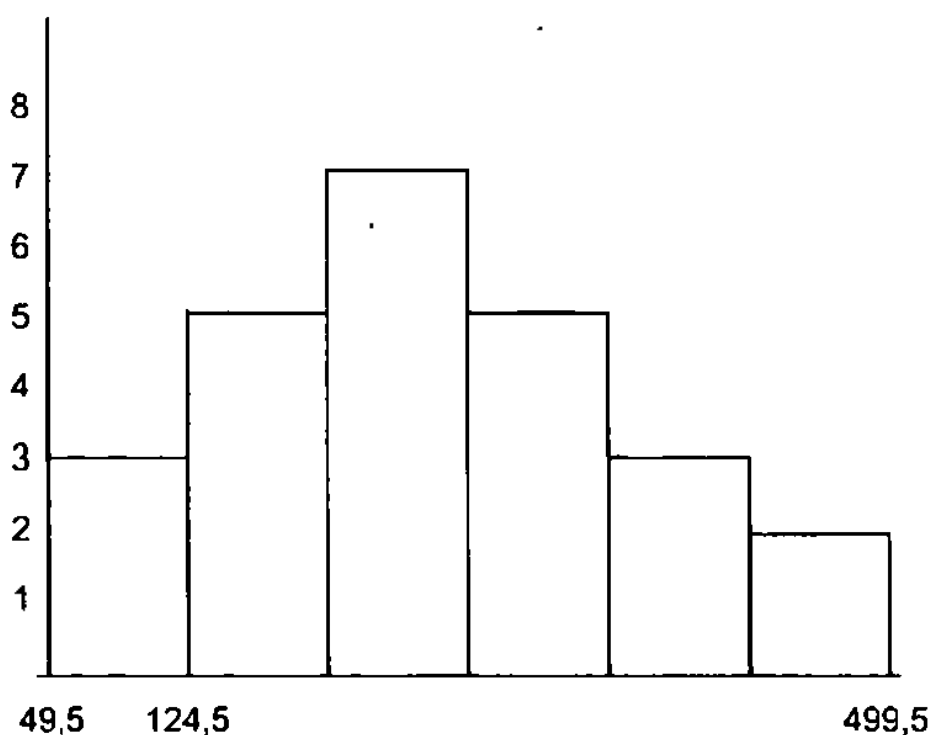
6.

Khoảng	Tần số
4,35 – 4,65	0,05
4,65 – 4,95	0,30
4,95 – 5,25	0,30
5,25 – 5,55	0,20
5,55 – 5,85	0,15
Tổng số	1,00

7. Bảng phân bố tần số như sau :

Khoảng	Tần số
49 – 124,5	3
124,5 – 199,5	5
199,5 – 274,5	7
274,5 – 349,5	5
349,5 – 424,5	3
424,5 – 499,5	2
Tổng số	25

Các khoảng có độ dài bằng nhau nên chiều cao của mỗi hình chữ nhật bằng tần số của khoảng dựng trên đó.

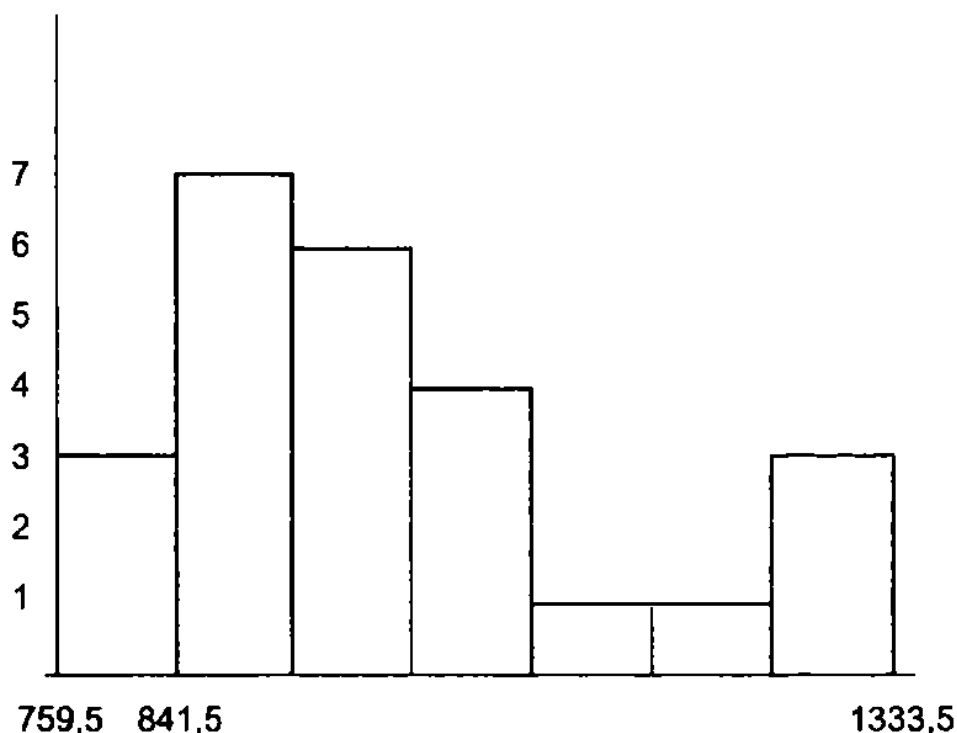




### 8. Bảng phân bố tần số như sau

Khoảng	Tần số
759,5 – 841,5	3
841,5 – 923,5	7
923,55 – 1005,5	6
1005,5 – 1087,5	4
1087,5 – 1169,5	1
1169,5 – 1251,5	1
1251,5 – 1333,5	3
Tổng số	25

Các khoảng có độ dài bằng nhau nên chiều cao mỗi hình chữ nhật bằng tần số của khoảng dựng trên đó.



9. Chiều cao của các hình chữ nhật trong tổ chức đồ này tỉ lệ giữa tần số của khoảng đó với độ dài của khoảng. Từ đó chiều cao của các hình chữ nhật sẽ là :

0,7   0,9   1,5   2,2   3,6   2,8  
2,0   1,6   0,8

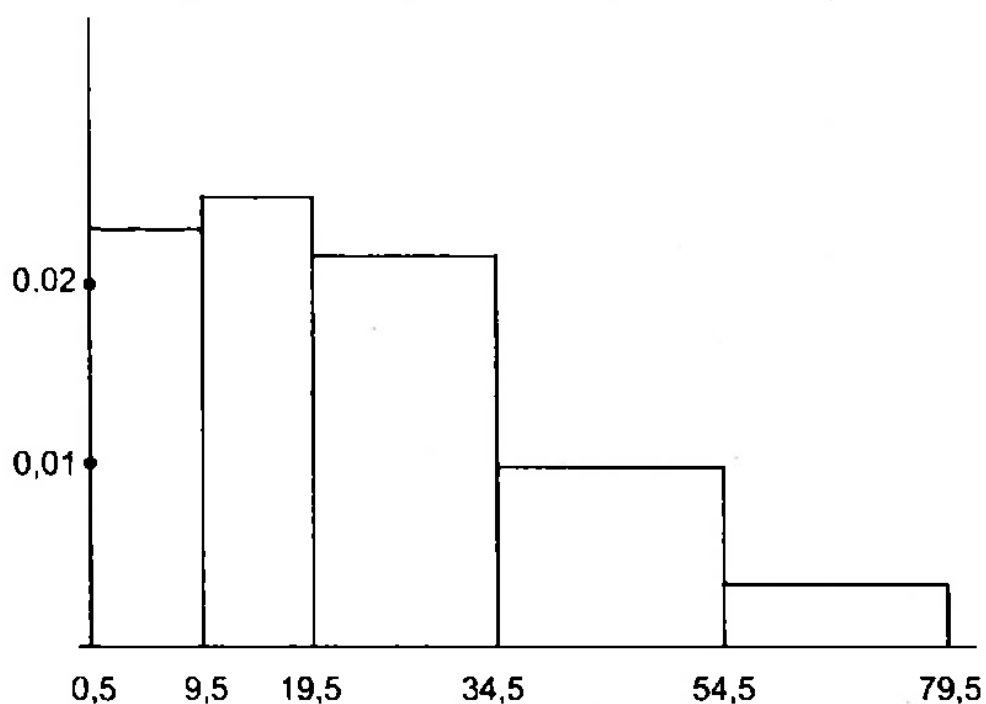
số thí sinh là 35 hoặc 36.

10. Chiều cao của các hình chữ nhật là

0,6   0,8   1,15   1,65   0,9

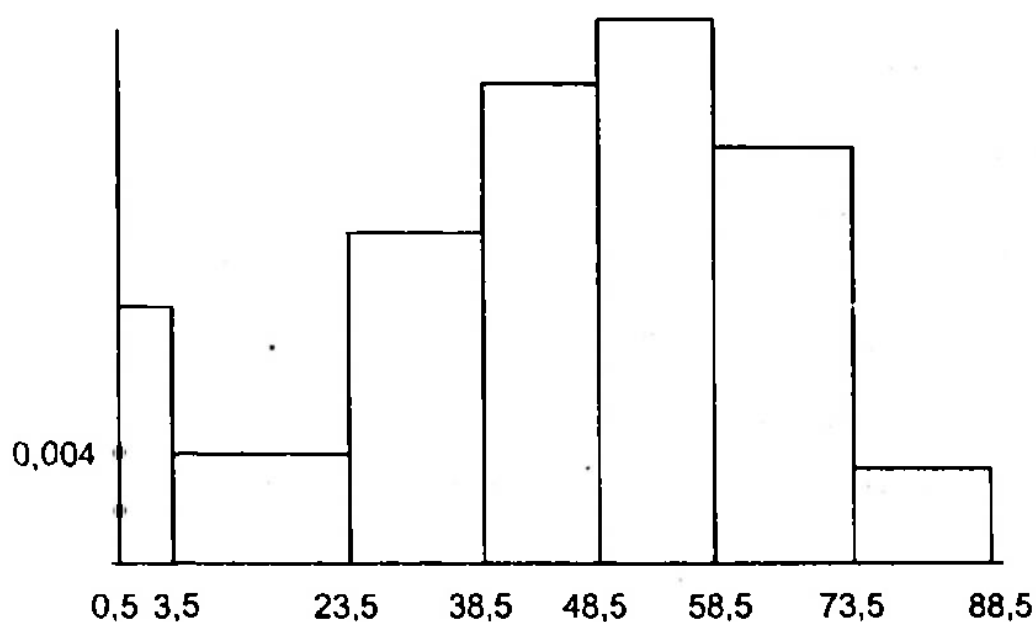
11. Chiều cao của các hình chữ nhật là

0,021   0,022   0,02   0,01   0,003



12. Chiều cao của các hình chữ nhật lần lượt là

0,008   0,005   0,012   0,019  
0,002   0,015   0,004



Nhìn vào tổ chức đồ tần suất dân cư của hai vùng A và B ta thấy ở vùng A dân cư chủ yếu là thanh niên, còn ở vùng B có nhiều người già.

$$\begin{aligned} 13. \quad x &= 59,4 \\ \text{mêđian } \mu &= 59,5 \\ \text{mod} &= 63 \end{aligned}$$

$$\begin{aligned} 14. \quad x &= 986,4 \\ \text{mêđian } \mu &= 970 \\ \text{mod} &= 980. \end{aligned}$$

$$15. \quad \bar{x} = 35,5$$

Khoảng mod là (10 – 20) và (30 – 50)

Số trung vị (mêđian) là 29. Cách tìm số trung vị như sau. Ta có

$$7 + 20 = 27 < \frac{81}{2} = 40,5$$

và  $7 + 20 + 15 = 42 > 40,5$

Vậy khoảng trung vị là (20 – 30).

Tần số của khoảng này  $f = 15$ , độ rộng là  $r = 10$ . Vậy số trung vị (mêđian) là

$$\mu = 20 + \left( \frac{40,5 - 27}{15} \right) \cdot 10 = 29$$

$$16. \quad \bar{x} = 20,57.$$

Khoảng mod là (22 – 23).

$$\text{Ta có } \frac{n}{2} = \frac{200}{2} = 100 \text{ và } 5 + \dots + 11 = 91 < 100$$

còn  $5 + \dots + 11 + 21 > 100$

do đó khoảng trung vị là (21 – 22).

Tần số của khoảng này là  $f = 21$ , độ rộng  $r = 1$ . Vậy

$$\mu = 21 + \frac{100 - 91}{21} = 21,4285$$

17.  $\bar{x} = 48,11.$

18.  $\bar{x} = 30,8$

19.  $\mu = 94,5 + \frac{\frac{80}{2} - 30}{26} \cdot 5 = 96,423$

20. Điểm giữa các khoảng là : 22, 27, 32, 37, 42.

Vậy 
$$\bar{x} = \frac{2(22) + 7(27) + 15(32) + \dots + 3(42)}{35}$$
$$= \frac{1135}{35} = 32,43$$

Số trung vị là

$$\mu = 29,5 + \left( \frac{17,5 - 9}{15} \right) 5 = 32,33$$

21. Bảng phân bố tần số là

Khoảng	Tần số	$x_i$
89,5 – 99,5	1	95
99,5 – 109,5	5	105
109,5 – 119,55	12	115
119,5 – 129,5	13	125
129,5 – 139,5	10	135
139,5 – 149,5	4	145
Tổng số	45	

$$\bar{x} = \frac{5555}{45} = 123,44$$

Số trung vị :

$$\mu = 119,5 + \frac{22,5 - 18}{13} \cdot 10 = 122,96$$

**22. Số trung vị :**

$$\mu = 30 + \left( \frac{\frac{70}{2} - 20}{26} \right) \cdot 10 = 35,77$$

Lớp mod là (30 – 40).

**23. Ta lập bảng phân bố tần số như sau :**

Khoảng	Tần số
19,5 – 20,5	5
20,5 – 21,5	8
21,5 – 22,5	11
22,5 – 23,5	10
23,5 – 24,5	5
24,5 – 25,5	1
Tổng số	40

Ta có  $\mu = 21,5 + \left( \frac{20 - 13}{11} \right) = 22,14$

$$\bar{x} = \frac{885}{40} = 22,125$$

**24. Đối với nhóm A**

Biên độ :  $49,1 - 19,6 = 29,5$ .

Trung bình :  $\bar{x} = 25,22$ .

Độ lệch trung bình :

$$\frac{23,88 + 10,98 + \dots + 5,62}{10} = \frac{69,72}{10} = 6,972.$$

**Đối với nhóm B**

Biên độ :  $26,0 - 9,6 = 16,4$

Giá trị trung bình :

$$\bar{x} = \frac{132,7}{10} = 13,27.$$

Độ lệch trung bình :

$$\frac{17,23 + 6,73 + \dots + 3,13}{10} = \frac{38,92}{10} = 3,892.$$

b) Nhóm A có mức độ phân tán lớn hơn nhóm B.

25.  $\bar{x} = 6,1 ;$

$$s^2 = \frac{34,92}{9} = 3,88 ;$$

$$s = \sqrt{3,88} = 1,97.$$

26.  $\bar{x} = 116,3727 ;$

$$s = 1,249.$$

27.  $\bar{x} = 46,1 ;$

$$s = 19,2.$$

28.  $\bar{x} = 30,02 ;$

$$s = 7,872.$$

29.  $\bar{x} = 5.$

$$s = 3,5365.$$

30.  $\bar{x} = 40 ;$

$$s^2 = 17 ;$$

$$s = 4,12.$$

31.  $\bar{x} = 522,4 ;$

$$s = 168,4.$$

## Chương 2

### ĐẠI CƯƠNG VỀ LÝ THUYẾT XÁC SUẤT

32. a) Các trường hợp thuận lợi là :

$$(1,6) ; (2,5) ; (3,4) ; (4,3) ; (5,2) ; (6,1).$$

$$\text{Vậy } p = \frac{1}{6}$$

b) Các trường hợp thuận lợi là  $(2,6) ; (3,5) ; (4,4) ; (5,3) ; (6,2)$ .

$$\text{Vậy } p = 5/36.$$

$$\text{c) } p = \frac{2}{9}$$

33. a) Các tổ hợp có tổng bằng 8 là  $(2, 3, 3) ; (2, 2, 4) ; (1, 1, 6) ; (1, 2, 5)$  và  $(1, 3, 4)$ .

Từ đó số trường hợp thuận lợi là  $3 + 3 + 3 + 6 + 6 = 21$ .

Để thấy số trường hợp có thể là  $6^3 = 216$ .

$$\text{Vậy } p = \frac{21}{216} \approx 0,097.$$

b) Các tổ hợp có tổng bằng 11 là  $(1, 4, 6) (1, 5, 5) (2, 3, 6), (2, 4, 5), (3, 4, 4), (3, 3, 5)$ .

Từ đó số trường hợp thuận lợi là :

$$6 + 3 + 6 + 6 + 3 + 3 = 27.$$

$$\text{Do đó } p = \frac{27}{216} = 0,125.$$

34. Số trường hợp có thể  $C_{12}^6$ .

Số trường hợp thuận lợi  $C_6^3 \cdot C_4^2 \cdot C_2^1$ .

$$\text{Vậy } p = \frac{20}{77}$$

35. a) Trong 30 số có 15 số chẵn do đó :

$$p = \frac{C_{15}^{10}}{C_{30}^{10}} = 0,00009$$

b) Trong 30 số từ 1 đến 30 có đúng 10 số chia hết cho 3. Vậy

$$p = \frac{C_{10}^5 \cdot C_{20}^5}{C_{30}^{10}} = 0,130.$$

c) Trong 30 số có 15 số lẻ, 15 số chẵn, trong đó chỉ có 3 số chia hết cho 10 (là 10, 20, 30). Vậy

$$p = \frac{C_{15}^5 \cdot C_{12}^4 \cdot C_3^1}{C_{30}^{10}} \approx 0,1484.$$

36. Dễ thấy rằng tổng của ba số nguyên dương  $a + b + c$  chia hết cho 3 khi và chỉ khi hoặc cả ba số đều chia hết cho 3 hoặc số dư của các số khi chia cho 3 khác nhau (không có số nào đồng dư với nhau theo mod 3). Có 33 số chia 3 dư 1, 33 số chia hết cho 3 và 34 số chia 3 dư 2. Vậy có

$$C_{33}^3 + C_{34}^3 + C_{33}^3 = 16896$$

khả năng rút được ba tấm thẻ chia 3 có cùng số dư và có

$$C_{33}^1 C_{33}^1 C_{34}^1 = 37026$$



khả năng rút được ba tấm thẻ với các số trên đó không đồng dư với nhau (mod 3). Vậy

$$p = \frac{16896 + 37026}{C_{100}^3} \approx 0,3334$$

**37.** Số trường hợp có thể là  $9^n$ .

Ta hãy tính xác suất của biến cố đối C : "Tích các số thu được chia hết cho 10".

Gọi A là biến cố "trong n số rút được không có số chẵn" ;

B : "Trong n số rút được không có số 5".

Ta có  $C = A \cup B$ .

Vậy  $P(C) = P(A) + P(B) - P(AB)$ .

Dễ thấy  $P(A) = \left(\frac{5}{9}\right)^n$

$$P(B) = \left(\frac{8}{9}\right)^n$$

$$P(AB) = \left(\frac{4}{9}\right)^n$$

Thành thử  $P(C) = \frac{5^n + 8^n - 4^n}{9^n}$  và do đó xác suất cần tìm là

$$p = 1 - P(C) = 1 - \frac{5^n + 8^n - 4^n}{9^n}$$

**38.** Gọi A là biến cố : "Cả hai nữ được chọn" ;

B là biến cố : "Có ít nhất một nữ được chọn".

Ta có

$$P(A/B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{6/15}{14/15} = \frac{3}{7}$$

b) Có 5 khả năng Hoa được chọn. Vậy xác suất là  $\frac{5}{15} = \frac{1}{3}$ .

$$c) P = \frac{5/15}{14/15} = \frac{5}{14}.$$

39. Số cách chọn 1 toa có 3 người, 1 toa có 1 người và hai toa trống là  $A_4^2 = 12$ . Do đó

$$P = \frac{A_4^2 \cdot C_4^3}{4^4} = \frac{3}{16}$$

40. Gọi A là biến cố : "Vé có chữ số 1" ;

B là biến cố : "Vé có chữ số 5".

$$\text{Ta có } P(\bar{A}) = \left(\frac{9}{10}\right)^5 ; P(\bar{B}) = \left(\frac{9}{10}\right)^5 ;$$

$$P(\overline{AB}) = \left(\frac{8}{10}\right)^5$$

Xác suất cần tìm là

$$\begin{aligned} P(\bar{A} \cup \bar{B}) &= P(\bar{A}) + P(\bar{B}) - P(\overline{AB}) \\ &= 2\left(\frac{9}{10}\right)^5 - \left(\frac{8}{10}\right)^5 \end{aligned}$$

41. Giả sử ba toa tàu được kí hiệu là A, B, C. Gọi A, B, C tương ứng là các biến cố : "Toa A không có hành khách mới lên" ; "Toa B không có hành khách mới lên" và "Toa C không có hành khách mới lên".

Biến cố đối của biến cố đang xét là  $A \cup B \cup C$  :

"Có ít nhất một toa không có hành khách mới lên".

$$\begin{aligned} \text{Ta có } P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(AB) - P(BC) - P(CA) + P(ABC) \end{aligned}$$

Dễ thấy

$$P(A) = \left(\frac{2}{3}\right)^5 ; P(B) = \left(\frac{2}{3}\right)^5 ; P(C) = \left(\frac{2}{3}\right)^5$$

$$P(AB) = P(BC) = P(CA) = \left(\frac{1}{3}\right)^5 ; P(ABC) = 0.$$

Thay vào ta được

$$P(A \cup B \cup C) = 3\left(\frac{2}{3}\right)^5 - 3\left(\frac{1}{3}\right)^5 = \frac{31}{81}.$$

$$\text{Do đó xác suất cần tìm là } 1 - \frac{31}{81} = \frac{50}{81}.$$

**42.** Kí hiệu 4 lá thư đó là A, B, C, D.

Gọi A là biến cố : "Lá thư A bỏ đúng địa chỉ" ; B là biến cố : "Lá thư B bỏ đúng địa chỉ" ; C là biến cố : "Lá thư C bỏ đúng địa chỉ" và D là biến cố : "Lá thư D bỏ đúng địa chỉ".

Ta phải tìm

$$\begin{aligned} P(A \cup B \cup C \cup D) &= P(A) + P(B) + P(C) + P(D) - P(AB) \\ &\quad - P(AC) - P(AD) - P(BC) - P(BD) - P(CD) + P(ABC) \\ &\quad + P(ACD) + P(BCD) - P(ABCD). \end{aligned}$$

$$\text{Dễ thấy } P(A) = P(B) = P(C) = P(D) = \frac{3!}{4!} = \frac{1}{4} ;$$

$$P(AB) = P(AC) = \dots P(CD) = \frac{2!}{4!} = \frac{1}{12} ;$$

$$P(ABC) = \dots = P(BCD) = \frac{1}{4!} = \frac{1}{24} ;$$

$$P(ABCD) = \frac{1}{24}.$$

Thay vào ta được

$$P(A \cup B \cup C \cup D) = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} = \frac{15}{24}.$$

43. a)  $p = (0,97)^{20} = 0,54338$

b)  $p = 20(0,03).(0,97)^{19} + 190.(0,03)^2(0,97)$   
 $= 0,3364 + 0,0988 = 0,4352$

c)  $p = 1 - 0,54338 - 0,4352 = 0,021.$

44. a) Anh ta được 13 điểm trong trường hợp trả lời đúng 5 câu và trả lời sai 7 câu. Vậy xác suất là

$$C_{12}^5 \left(\frac{1}{5}\right)^5 \cdot \left(\frac{4}{5}\right)^7 = 0,0532.$$

b) Anh ta bị điểm âm khi trả lời đúng ít hơn 3 câu. Thật vậy gọi  $x$  là số câu trả lời đúng. Số câu trả lời sai là  $12 - x$ . Bất phương trình  $4x < 12 - x$  xảy ra khi  $x = 0, 1, 2$ .

Vậy  $P\{\text{nhận điểm âm}\} =$

$$= C_{12}^0 \left(\frac{4}{5}\right)^{12} + C_{12}^1 \left(\frac{1}{5}\right) \left(\frac{4}{5}\right)^{11} + C_{12}^2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{10}$$
$$= 0,0687 + 0,2062 + 0,2835 = 0,5583.$$

45. Xác suất thắng trong 1 ván là

$$C_3^2 \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right) + \left(\frac{1}{6}\right)^3 = \frac{2}{27} \approx 0,071.$$

Vậy xác suất để thắng ít nhất 3 ván là :

$$C_5^3 \left(\frac{2}{27}\right)^3 \left(\frac{25}{27}\right)^2 + C_5^4 \left(\frac{2}{27}\right)^4 \left(\frac{25}{27}\right) + \left(\frac{2}{27}\right)^5 =$$
$$= \frac{52032}{14348907} \approx 0,003$$

46. a) Anh ta trở lại điểm xuất phát khi tiến 4 bước và lùi 4 bước. Vậy

$$p = C_8^4 \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^4 = \frac{70}{256}$$

b) Anh ta cách điểm xuất phát hơn 2m nếu số bước tiến là 8 hoặc 7 hoặc 1 hoặc 0. Vậy :

$$p = C_8^8 \left(\frac{1}{2}\right)^8 + C_8^7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right) + C_8^1 \left(\frac{1}{2}\right)^8 + \left(\frac{1}{2}\right)^8 = \frac{18}{256}$$

47. a) Gọi  $B_1$  là biến cố : "chọn được tập 1", và  $B_2$  là biến cố : "chọn được tập 2"

A là biến cố cần tìm... Ta có

$$\begin{aligned} P(A) &= P(B_1) P(A/B_1) + P(B_2) P(A/B_2) \\ &= \frac{1}{2} P(A/B_1) + \frac{1}{2} P(A/B_2). \end{aligned}$$

Tính  $P(A/B_1)$

Khi  $B_1$  xảy ra thì 2 câu được chọn từ tập thứ nhất mà anh ta đã thuộc 9 câu. Vậy

$$P(A/B_1) = \frac{C_9^1 \cdot C_1^1}{C_{10}^2} \cdot \frac{2}{10} = \frac{36}{900}$$

(Câu thuộc duy nhất của anh ta phải ở tập thứ nhất).

Tính  $P(A/B_2)$

Có thể xảy ra các khả năng sau :

+ Không thuộc hai câu của tập thứ hai, thuộc câu tập thứ nhất.

Xác suất là :

$$\frac{C_2^2}{C_{10}^2} \cdot \frac{9}{10} = \frac{18}{900}$$

+ Không thuộc 1 câu của tập 2 và 1 câu của tập 1, thuộc 1 câu tập 2. Xác suất là

$$\frac{C_8^1 \cdot C_2^1}{C_{10}^2} \cdot \frac{1}{10} = \frac{32}{900}$$

Từ đó  $P(A/B_2) = \frac{50}{900}$

Thành thử  $P(A) = \left( \frac{36}{900} + \frac{50}{900} \right) \frac{1}{2} = \frac{43}{900} = 0,04$

b) Gọi B là biến cố cần tìm. Ta có

$$P(B) = P(B_1)P(B/B_1) + P(B_2)P(B/B_2).$$

Tính  $P(B/B_1)$

Có hai khả năng :

+ Thuộc hai câu của tập 1, không thuộc câu nào của tập 2. Xác suất là

$$\frac{C_9^2}{C_{10}^2} \cdot \frac{2}{10} = \frac{144}{900}$$

+ Thuộc một câu của tập 1 và 1 câu của tập 2. Xác suất là

$$\frac{C_9^1 \cdot C_1^1}{C_{10}^2} \cdot \frac{8}{10} = \frac{144}{900}$$

Vậy  $P(B/B_1) = \frac{288}{900}$

Tính  $P(B/B_2)$

Có hai khả năng :

+ Thuộc hai câu tập 2, không thuộc câu nào ở tập 1. Xác suất là :

$$\frac{C_8^2}{C_{10}^2} \cdot \frac{1}{10} = \frac{56}{900}$$

+ Thuộc 1 câu tập 2, 1 câu tập 1. Xác suất là :

$$\frac{C_8^1 C_2^1}{C_{10}^2} \cdot \frac{9}{10} = \frac{288}{900}$$

Vậy  $P(B/B_2) = \frac{344}{900}$

Từ đó  $P(B) = \left( \frac{344 + 288}{900} \right) \frac{1}{2} = \frac{316}{900} = 0,351.$

48. a)  $P = (0,9)(0,8)(0,9) = 0,648.$

b)  $P \{ \text{trượt ở vòng 2} \} = (0,9)(0,2) = 0,18.$

Vậy xác suất để thí sinh trượt ở vòng 2 nếu biết rằng thí sinh đó trượt là

$$P = \frac{P\{\text{trượt ở vòng 2}\}}{P\{\text{trượt}\}} = \frac{0,18}{0,352} = 0,511.$$

49. Kí hiệu  $E_1$  : "Từ chuông 2 bắt được thỏ trắng" ;

$E_2$  : "Từ chuông 2 bắt được thỏ đen" ;

$A$  : "Bắt được thỏ trắng ở lần bắt sau" ;

$B$  : "Bắt được thỏ trắng của chuông 1 ở lần bắt sau".

Ta có

$$\begin{aligned} P(A) &= P(E_1)P(A/E_1) + P(E_2)P(A/E_2) = \\ &= \frac{3}{10} \cdot \frac{11}{16} + \frac{7}{10} \cdot \frac{10}{16} = \frac{103}{160} ; \end{aligned}$$

$$\begin{aligned} P(B) &= P(E_1)P(B/E_1) + P(E_2)P(B/E_2) = \\ &= \frac{3}{10} \cdot \frac{10}{16} + \frac{7}{10} \cdot \frac{10}{16} = \frac{100}{160} . \end{aligned}$$

Vậy  $P(B/A) = \frac{P(AB)}{P(A)} = \frac{P(B)}{P(A)} = \frac{100}{103}$

50. Gọi  $A_1$  là biến cố : "Con nai trúng đạn của người thứ nhất" ;  
 $B$  là biến cố : "Con nai bị trúng 1 viên đạn".

$$\text{Ta có } P(B) = (0,3)(0,6)(0,5) + (0,7)(0,6)(0,5) + (0,7)(0,4)(0,5) \\ = 0,44 ;$$

$$P(A_1) = (0,3)(0,6)(0,5) = 0,09.$$

$$\text{Vậy : } P(A_1/B) = \frac{9}{44}$$

Tương tự gọi  $A_i$  là biến cố : "Con nai trúng đạn của người thứ  $i$ "  
 $(i = 2,3)$ , thì

$$P(A_2/B) = \frac{14}{44} ; P(A_3/B) = \frac{21}{44}$$

Thành thử nên chia thu hoạch theo tỉ lệ

$$9 : 14 : 21$$

(Đa số nghĩ rằng nên chia theo tỉ lệ

$$3 : 4 : 5)$$

51. Gọi  $A$  là biến cố : "Bóng 1 và bóng 2 bị hỏng" ;  $B$  là biến cố : "Có đúng hai bóng hỏng".

Ta có

$$P(A) = (0,1)(0,2)(0,7)(0,6) = 0,0084 ;$$

$$P(B) = (0,1)(0,2)(0,7)(0,6) + (0,1)(0,8)(0,3)(0,6) + \\ + (0,1)(0,8)(0,7)(0,4) + (0,9)(0,2)(0,3)(0,6) + \\ + (0,9)(0,2)(0,7)(0,4) + (0,9)(0,8)(0,3)(0,4) = \\ = 0,2144.$$

$$\text{Vậy } P(A/B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{0,0084}{0,2144} = 0,039$$



**52.** a) Kí hiệu O, A, B, AB tương ứng là các biến cố "Người cần máu có nhóm máu là O, A, B và AB".

Gọi H là biến cố : "Sự truyền máu thực hiện được".

Theo công thức xác suất đầy đủ ta có

$$P(H) = P(O)P(H/O) + P(A)P(H/A) + \\ + P(B)P(H/B) + P(AB)P(H/AB).$$

Theo dữ kiện đầu bài

$$P(O) = 0,337 ; P(A) = 0,375 ; P(B) = 0,209 ; P(AB) = 0,079 ;$$

$$P(H/O) = P(O) = 0,337 ;$$

$$P(H/A) = P(O) + P(A) = 0,712 ;$$

$$P(H/B) = P(O) + P(B) = 0,576 ;$$

$$P(H/AB) = 1.$$

Thay vào ta tìm được  $P(H) = 0,5737$ .

b) Gọi E là biến cố : "Sự truyền máu không thực hiện được"

Ta có :

$$P(E/O) = [1 - P(O)]^2 = 0,663^2 ;$$

$$P(E/A) = [1 - P(O) - P(A)]^2 = 0,288^2 ;$$

$$P(E/B) = [1 - P(O) - P(B)]^2 = 0,454^2 ;$$

$$P(E/AB) = 0.$$

Áp dụng công thức xác suất đầy đủ ta được

$$P(E) = 0,223.$$

Vậy xác suất để truyền máu được là

$$1 - P(E) = 0,7777.$$

$$\mathbf{53.} \quad p = q^5 + 5q^4p + 10q^3p^2 = 0,05792.$$

54. a) Gọi  $\psi$  là số ô tô cho thuê,  $X$  là số khách đến thuê. Ta có

$$P\{\psi = 0\} = P\{X = 0\} = 0,1353 ;$$

$$P\{\psi = 1\} = P\{X = 1\} = 0,2707 ;$$

$$P\{\psi = 2\} = P\{X = 2\} = 0,2707 ;$$

$$P\{\psi = 3\} = P\{X = 3\} = 0,1804 ;$$

$$P\{\psi = 4\} = P\{X \geq 4\} = 0,1429.$$

Từ đó  $E\psi = 1,925$ .

b) Gọi  $n$  là số ô tô mà đại lí cần có.

Ta cần có :  $P\{X \leq n\} > 0,98$ .

Tra bảng ta thấy  $P\{X \leq 4\} = 0,9473 ;$

$$P\{X \leq 5\} = 0,9834.$$

Vậy  $n = 5$ .

55. a) Gọi  $X$  là số hoa trên một chậu cây.

Theo giả thiết  $X$  có phân bố Poisson với tham số  $\lambda = 3$ .

Vậy  $P\{X = 2\} = 0,2240 ;$

$$P\{X = 3\} = 0,2240 ;$$

$$P\{X = 4\} = 0,1680 ;$$

$$P\{X = 5\} = 0,1008 ;$$

và  $P\{2 \leq X \leq 5\} = 0,7169.$

Gọi  $Y$  là số hoa của chậu cây đem bán.

Ta có  $P\{Y = 2\} = \frac{P\{X = 2\}}{P\{2 \leq X \leq 5\}} = 0,3125 ;$

Tương tự  $P\{Y = 3\} = 0,3125 ;$

$$P\{Y = 4\} = 0,2344 ;$$

$$P\{Y = 5\} = 0,1406.$$

b)  $EY = 3,203$ .

**56.** Giả sử  $X$  là số người chọn ăn ở đợt 1.

Khi đó  $1000 - X$  là số người chọn. Gọi  $k$  là số chỗ ngồi trong nhà ăn. Ta phải chọn  $k$  nhỏ nhất để

$$P\{X < k, 1000 - X < k\} \geq 0,99$$

$$\Leftrightarrow P\{1000 - k < X < k\} \geq 0,99$$

$X$  có phân bố xấp xỉ chuẩn với kì vọng  $\mu = 500$ , độ lệch tiêu chuẩn  $\sigma = \sqrt{250}$ . Vậy ta phải có :

$$\Phi\left(\frac{k - 500}{\sqrt{250}}\right) - \Phi\left(\frac{500 - k}{\sqrt{250}}\right) \geq 0,99$$

$$\Leftrightarrow 2\Phi\left(\frac{k - 500}{\sqrt{250}}\right) \geq 0,99$$

$$\Leftrightarrow \Phi\left(\frac{k - 500}{\sqrt{250}}\right) \geq \Phi(2,58)$$

$$\Leftrightarrow k \geq 500 + 2,58 \sqrt{250} = 540,79.$$

Vậy  $k = 541$ .

**57.** Gọi  $X$  là số cặp vợ chồng chọn ăn ở đợt 1. Khi đó  $500 - X$  là số cặp chọn ăn ở đợt 2. Ta phải chọn  $k$  nhỏ nhất để

$$P\{2X < k, 100 - 2X < k\} \geq 0,99$$

$$\Leftrightarrow P\{500 - \frac{k}{2} < X < \frac{k}{2}\} \geq 0,99.$$

$X$  có phân bố nhị thức,  $X \approx B\left(500, \frac{1}{2}\right)$ . Dùng xấp xỉ chuẩn :  $X$

có phân bố xấp xỉ chuẩn với kì vọng  $\mu = 250$ , độ lệch tiêu chuẩn  $\sigma = \sqrt{125}$ . Vậy ta phải có

$$\Phi\left(\frac{\frac{k}{2} - 250}{\sqrt{125}}\right) - \Phi\left(\frac{250 - \frac{k}{2}}{\sqrt{125}}\right) \geq 0,99$$

$$\Leftrightarrow \frac{k}{2} \geq 250 + 2,58 \sqrt{125}$$

$$\Leftrightarrow k \geq 557,69.$$

Vậy  $k = 558$ .

58. Gọi  $X$  là số câu trả lời đúng của sinh viên. Ta có  $X \approx B(45, 1)$ . Dùng xấp xỉ chuẩn, ta được

$$\begin{aligned} \text{a) } P\{X \geq 16\} &= P\{\bar{X} > 15,5\} = P\left\{Z > \frac{15,5 - 11,25}{2,904}\right\} = \\ &= 1 - \Phi(1,463) = 0,0717; \end{aligned}$$

$$\text{b) } P\{X < 10\} = P\{\bar{X} \leq 9,5\} = 0,2737.$$

$$\text{c) } P\{8 \leq X \leq 12\} = P\{7,5 < \bar{X} < 12,5\} = 0,5681.$$

59. Ta chọn  $\varepsilon$  sao cho

$$P\left\{\left|\frac{X - 500}{22,13}\right| < \varepsilon\right\} = 0,95$$

$$\Leftrightarrow 2\Phi(\varepsilon) = 1,95 \Leftrightarrow \varepsilon = 1,96.$$

$$\text{Vậy } P\{500 - (1,96)(22,13) < X < 500 + (1,96)(22,13)\} = 0,95$$

$$\Leftrightarrow P\{456,62 < X < 543,37\} = 0,95.$$

Vậy ta kết luận : Số phế phẩm nằm trong khoảng từ 457 đến 543. Kết luận này có độ tin cậy 95%.

60. a) Gọi  $X$  là số người trúng tuyển :

$X \sim B(325 ; 0,9)$  ;  $X$  có phân bố xấp xỉ chuẩn với  $\mu = 292,5$  ;  $\sigma = 5,4$ . Vậy :

$$\begin{aligned}
 P\{X \leq 300\} &= P\{\bar{X} < 300,5\} \\
 &= \Phi\left(\frac{8}{5,4}\right) = \Phi(1,48) = 0,9306.
 \end{aligned}$$

b) Giả sử  $n$  là số người đăng kí. Khi đó  $X$  có phân bố xấp xỉ chuẩn với  $\mu = 0,9n$  và  $\sigma = 0,3\sqrt{n}$ .

$$\begin{aligned}
 \text{Vậy : } P\{X \leq 300\} &= P\{\bar{X} < 300,5\} \\
 &= \Phi\left(\frac{300,5 - 0,9n}{0,3\sqrt{n}}\right) \geq 0,99 = \Phi(2,33)
 \end{aligned}$$

$$\Leftrightarrow 300,5 - 0,9n \geq (0,3)(2,33)\sqrt{n}$$

Đặt  $\sqrt{n} = u$ . Ta cần có

$$0,9u^2 + 0,699u - 300,5 \leq 0$$

$$\Rightarrow u \leq 17,88 \Leftrightarrow n \leq 319,99.$$

Vậy cho phép tối đa là 319 người đăng kí dự thi.

**61. a)** Gọi  $X_i$  là thời gian làm việc cho tới khi hỏng của linh kiện thứ  $i$ . Khi đó  $S = \sum_{i=1}^{30} X_i$  là thời gian duy trì hoạt động của máy

với 30 linh kiện dự trữ. Theo yêu cầu ta phải tính  $P\{S > 8760\}$ . Sử dụng định lí giới hạn trung tâm, ta thấy  $S$  có phân bố xấp xỉ chuẩn với

$$\mu = 250 \times 30 = 7500$$

và độ lệch tiêu chuẩn

$$\sigma = \sqrt{(30)(250)^2} = 1369,3$$

$$\text{Vậy } P\{S > 8760\} = 1 - \Phi\left(\frac{8760 - 7500}{1369,3}\right) =$$

$$1 - \Phi(0,92) = 0,1788,$$

b) Gọi  $n$  là số linh kiện cần dự trữ và  $S = \sum_{i=1}^n X_i$ . Theo yêu cầu

ta phải tìm  $n$  để  $P\{S > 8760\} \geq 0,99$ .

Sử dụng định lí giới hạn trung tâm, ta thấy  $S$  có phân bố xấp xỉ chuẩn với  $\mu = 250n$  và độ lệch tiêu chuẩn  $\sigma = 250\sqrt{n}$ .

$$\text{Vậy } P\{S > 8760\} \geq 0,99$$

$$\Leftrightarrow 1 - \Phi\left(\frac{8760 - 250n}{250\sqrt{n}}\right) \geq 0,99 = \Phi(2,33)$$

$$\Leftrightarrow 250n - 8760 \geq 250(2,33)\sqrt{n}$$

$$\Leftrightarrow n - 2,33\sqrt{n} - 35,04 \geq 0$$

$$\Leftrightarrow \sqrt{n} \geq 7,195 \Leftrightarrow n \geq 51,76$$

Vậy  $n = 52$ .

### Chương 3

## ƯỚC LƯỢNG THAM SỐ

62.  $14,65 \pm 0,82$ .

63.  $7,52 \pm 0,29$

64. Từ mẫu ta tính được  $\bar{X} = 8,06$  và  $s = 0,64$ .

Vậy khoảng tin cậy 95% là

$$8,06 \pm 1,96 \frac{0,64}{\sqrt{35}} = 8,06 \pm 0,21$$

65. Từ các số liệu trên ta có :

$$n = 59 ; \bar{x} = 41,05 ;$$

$$s = 27,99$$

Do đó  $\frac{s}{\sqrt{n}} = \frac{27,99}{\sqrt{59}} = 3,64$ .

Khoảng tin cậy 95% cho  $\mu$  là :

$$x \pm 1,96(3,64) = 41,05 \pm 7,13$$

hay

$$33,92 \leq \mu \leq 48,18.$$

66. Ta có

$$\alpha = 1 - 0,99 = 0,01 ;$$

$$\frac{\alpha}{2} = 0,005$$

Tra bảng phân bố Student với 14 bậc tự do ta tìm được :

$t = t_{0,005} = 2,997$ . Vậy khoảng tin cậy 95% cho  $\mu$  là :

$$39,8 \pm 2,997 \frac{0,379}{\sqrt{15}}$$

hay  $39,5023 \leq \mu \leq 40,0977$

67.  $\bar{x} = 526$  ;  $s_x = 61,03$  ;

$\bar{y} = 638$  ;  $s_y = 53,42$ .

Từ đó khoảng tin cậy cho  $\mu_x$  là :  $526 \pm 45,982$  và khoảng tin cậy  $\mu_y$  là :  $638 \pm 42,96$ .

68. Từ các số liệu trên ta tính được :

$\bar{x} = 171$  ;  $s = 3,4254$  ;

$\alpha = 0,05$  ;  $\frac{\alpha}{2} = 0,025$ .

Tra bảng phân bố Student với 15 bậc tự do ta tìm được

$t = t_{0,025} = 2,131$ .

Vậy khoảng tin cậy 95% của  $\mu$  là

$$171 \pm 2,131 \frac{3,4254}{\sqrt{16}} = 171 \pm 1,885$$

hay  $169,115 \leq \mu \leq 172,885$ .

69.  $4,095 \leq \mu \leq 7,723$ .

70.  $[2,563 ; 3,437]$ .

71.  $81,1 \leq \mu \leq 110,9$ .

72.  $35,65 \leq \mu \leq 40,35$ .

73.  $30,61 \leq \mu \leq 33,39$ .



74.  $\bar{x} = 22,333 ; s = 7,5116$

$$\alpha = 0,05 \Rightarrow t = t_{\alpha/2} = t_{0,025} = 2,201$$

(11 bậc tự do)

Vậy khoảng tin cậy 95% là

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 22,33 \pm 2,201 \cdot \frac{7,51}{\sqrt{12}}$$

$$\Rightarrow 17,56 \leq \mu \leq 27,10.$$

75.  $15,14 \leq \mu \leq 18,86.$

76.  $23,53 \leq \mu \leq 32,17.$

77.  $[242190 ; 307810].$

78.  $132,4 \leq \mu \leq 143,1.$

79.  $f = \frac{74}{80} = 0,925$

Khoảng tin cậy 99% là :

$$0,925 \pm 2,58 \frac{\sqrt{(0,925)(0,075)}}{80} = 0,925 \pm 0,076.$$

Vậy :  $0,849 \leq p \leq 1.$

80.  $f = \frac{325}{500} = 0,65.$

Khoảng tin cậy 95% là :

$$\begin{aligned} 0,65 \pm 2,58 \frac{\sqrt{(0,65)(0,35)}}{500} \\ = 0,65 \pm 2,58 (0,0213) = 0,65 \pm 0,06. \end{aligned}$$

Vậy  $0,59 \leq p \leq 0,71.$

81. a)  $f = \frac{32}{40} = 0,8.$

khoảng tin cậy 95% là :

$$0,8 \pm 1,96 \frac{\sqrt{(0,8)(0,2)}}{40} = 0,8 \pm (1,96)(0,0632) = \\ = 0,8 \pm 0,124.$$

Vậy  $0,676 \leq p \leq 0,924$

b)  $f = \frac{320}{400} = 0,8$

Khoảng tin cậy 95% là

$$0,8 \pm 1,96 \frac{\sqrt{(0,8)(0,2)}}{400} = 0,8 \pm (1,96)(0,02) \\ = 0,8 \pm 0,039.$$

Vậy :  $0,761 \leq p \leq 0,839$

82.  $f = \frac{60}{100} = 0,6.$

Khoảng tin cậy 90% là :

$$0,60 \pm 1,64 \frac{\sqrt{(0,6)(0,4)}}{100} = 0,6 \pm 0,08.$$

Vậy :  $0,52 \leq p \leq 0,68.$

83.  $f = \frac{162}{200} = 0,81.$

Khoảng tin cậy 95% là :

$$0,81 \pm 1,96 \frac{\sqrt{(0,81)(0,19)}}{200} = 0,81 \pm (1,96)(0,028) = \\ = 0,81 \pm 0,055.$$

84.  $0,468 \leq p \leq 0,632.$

85. a)  $0,175 \leq p \leq 0,525$ .

b)  $0,475 \leq p \leq 0,825$ .

86.  $0,314 \leq p \leq 0,426$ .

87.  $0,163 \leq p \leq 0,197$ .

88.  $f = \frac{110}{60} = 0,1964$

Khoảng tin cậy 95% là :

$$0,1964 \pm 1,96 \frac{\sqrt{(0,1964)(0,8036)}}{560}$$

hay  $0,163 \leq p \leq 0,23$ .

Vì  $p = \frac{1}{4} = 0,25$  không nằm trong khoảng này, nên số liệu tỏ ra không phù hợp với lí thuyết của Mendel.

89.  $n \geq \frac{1,96^2}{4.(2.10^{-2})^2} = \frac{1,96^2}{16(10^{-4})} \approx 2400$ .

90.  $n \geq \left[ \frac{(1,96)(0,7)}{0,2} \right]^2 = 47,06$ . Vậy  $n = 48$ .

91. Với độ tin cậy 95% thì :

$$n \geq \left[ \frac{(1,96)(1000)}{100} \right]^2 = 385$$

Vậy  $n = 385$ .

Với độ tin cậy 99% thì :

$$n \geq \left[ \frac{(2,58)(1000)}{100} \right]^2 = 666$$

Vậy  $n = 666$ .

92. Ta có 
$$n \geq \left[ \frac{(1,64).3}{0,5} \right]^2 = 96,826.$$

Vậy kích thước mẫu tối thiểu là  $n = 97$ .

93. Ta có 
$$n \geq \left[ \frac{(1,96).9}{2} \right]^2 = 77,79.$$
 Vậy kích thước mẫu tối thiểu là  $n = 78$ .

94. a) Ta có 
$$f = \frac{640}{1000} = 0,64.$$

Khoảng tin cậy 90% là :

$$0,64 \pm 1,64 \frac{\sqrt{(0,64)(0,36)}}{1000} = 0,64 \pm (1,64)(0,0152) = 0,64 \pm 0,0249.$$

Sai số ở đây là 0,0249.

b) 
$$n \geq \frac{(1,64)^2 (0,64)(0,36)}{(0,02)^2} = 1549,2.$$
 Vậy  $n = 1550$ .

Nếu không có một thông tin trước gì về  $p$  thì ta phải lấy

$$n \geq \frac{(1,64)^2}{4.(0,02)^2} = 1681, \text{ tức là } n = 1681.$$

95.  $n = 97$ .

96.  $n = 485$ .

97. 
$$n \geq \frac{(1,96)^2 (0,3)(0,7)}{(0,05)^2} = 322,69.$$
 Vậy kích thước mẫu tối thiểu là  $n = 323$ .

98. 
$$n \geq \frac{(1,96)^2}{4.(0,05)^2} = 384,16.$$

Vậy  $n = 385$ .

## Chương 4

### KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ

99. a)  $H_0 : \mu = 1010 ;$

$$H_1 : \mu \neq 1010.$$

b)  $T = \frac{(1090 - 1010)\sqrt{50}}{300} = 1,89.$

Với mức ý nghĩa  $\alpha = 0,01$  thì  $c = 2,58.$

Vì  $|T| < c$  nên ta không có đủ cơ sở để kết luận là có sự thay đổi.

c) P – giá trị là :  $2(0,55 - 0,4706) = 0,0588.$

100.  $H_0 : \mu = 1,70 \text{ (m)} ;$

$$H_1 : \mu \neq 1,70 \text{ (m)} ;$$

$$T = \frac{(1,72 - 1,70)\sqrt{121}}{0,02} = 11 ;$$

$$c = 1,96.$$

$|T| > c$  : Ta bác bỏ  $H_0$  và kết luận : Chiều cao trung bình của thanh niên đã thay đổi so với cách đây 10 năm.

101.  $H_0 : \mu = 125 ;$

$$H_1 : \mu \neq 124 ;$$

$$T = \frac{124 - 125}{5} \sqrt{400} = -4.$$

Vì  $|T| = 4 > c = 2,58$ , nên ta bác bỏ  $H_0$  : Số tiền điện trung bình hàng tháng của một hộ khác 125 nghìn một cách có ý nghĩa (mức ý nghĩa 1%).

c) P – giá trị là

$$2(0,5 - 0,49997) = 0,00006$$

**102.**  $H_0 : \mu = 140 ;$

$$H_1 : \mu \neq 140 ;$$

$$T = \frac{(154 - 140)\sqrt{50}}{62} = 1,59.$$

Hàng số  $c = z_{0,01} = 2,33$ .

Vì  $|T| < c$  nên ta chưa có cơ sở để bác bỏ  $H_0$ . Kết luận : "Thông báo của nhóm nghiên cứu đó chấp nhận được".

**103.** Vì ta tin một cách tiên quyết rằng loại máy tính đời mới sẽ không thể tồi hơn máy tính đời cũ, do đó ta có bài toán kiểm định một phía :

$$H_0 : \mu = 1200 ;$$

$$H_1 : \mu > 1200 ;$$

$$T = \frac{(1260 - 1200)\sqrt{40}}{215} = 1,76 ;$$

$$c = z_{0,05} = 1,64 < T.$$

Số liệu đã cho dẫn tới bác bỏ  $H_0$ . Kết luận : "Loại máy mới tốt hơn loại máy cũ một cách có ý nghĩa với mức ý nghĩa 5%".

**104. a)**  $H_0 : \mu = 90 ;$

$$H_1 : \mu > 90.$$

$$b) \quad T = \frac{(92 - 90)\sqrt{49}}{9} = 1,56.$$

Hàng số c là :  $c = z_{0,1} = 1,28$ .

$T > c$ , vậy ta bác bỏ  $H_0$ . Đoàn thanh tra kết luận là mật độ tên lửa tại các căn cứ của Mỹ dày đặc hơn là đã công bố

c) P – giá trị là :

$$P\{X \geq 92\} = 1 - \Phi(1,56) = 1 - 0,9406 = 0,0594$$

$$105. \quad H_0 : \mu = 160 ;$$

$$H_1 : \mu > 160 ;$$

$$T = \frac{(170 - 160)\sqrt{75}}{30} = 2,89.$$

Hàng số c =  $z_{0,99} = 2,33$ .

Ta có  $T > c$  ; vậy ta bác bỏ  $H_0$ . Kết luận : "Thu nhập bình quân đầu người của huyện A lớn hơn 160 nghìn một cách có ý nghĩa (mức ý nghĩa 1%)".

$$106. \quad H_0 : \mu = 30000 ;$$

$$H_1 : \mu < 30000.$$

$$\text{Ta có} \quad T = \frac{(30000 - 25000)\sqrt{100}}{5000} = 2.$$

a) Với mức ý nghĩa  $\alpha = 5\%$ , ta có  $c = 1,64$ .

Ta bác bỏ  $H_0$  ở mức ý nghĩa 5% : "Tuổi thọ trung bình của lớp ô tô nhỏ hơn 30000 dặm một cách có ý nghĩa với mức ý nghĩa 5%".

b) Với mức ý nghĩa  $\alpha = 2\%$  thì  $c = z_{0,02} = 2,06$ .

Ta chưa có cơ sở bác bỏ  $H_0$  ở mức ý nghĩa này.

**107.** Ta cần kiểm định giả thiết ;

$$H_0 : \mu = 21,5 ;$$

$$H_1 : \mu < 21,5 ;$$

$$\bar{x} = 20, s = 10$$

$$T = \frac{(21,5 - 20)\sqrt{6}}{\sqrt{10}} = 1,16.$$

Tra bảng phân bố Student với bậc tự do  $k = n - 1 = 5$ , ta tìm được :

$$c = t_{0,05} = 2,015.$$

Vì  $T < c$  nên ta không có cơ sở bác bỏ  $H_0$  : "Trung bình mẫu  $x = 20 < 21,5$  có thể chỉ là do ngẫu nhiên".

**108.** Giả thiết  $H_0 : \mu = 420$

với đối thiết  $H_1 : \mu > 420$ .

Ta có

$$T = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(442 - 420)\sqrt{16}}{60} = \frac{22}{15} = 1,47.$$

Tra bảng phân bố Student với 15 bậc tự do, ta tìm được :

$$c = t_{0,05} = 1,753 < T.$$

Do đó ta chưa có cơ sở để kết luận rằng mức chi phí của sinh viên năm nay cao hơn năm ngoái.

**109.**  $T = 1,33 ;$

$$c = 1,64.$$

Chấp nhận  $H_0 : \mu = 10$ , bác bỏ  $H_1 : \mu > 10$ .

**110.**  $H_0 : \mu = 88 ;$

$$H_1 : \mu > 88 ;$$



$$T = 2,99 ; c = 1,69.$$

Dẫn tới bác bỏ  $H_0$ .

**Kết luận :** Mỗi chiếc bánh trung bình chứa nhiều hơn 88 calo.

$$111. \quad H_0 : \mu = 9 ;$$

$$H_1 : \mu \neq 9 ;$$

$$T = 1,67 ; c = 1,96.$$

Chấp nhận  $H_0$ .

$$112. \quad H_0 : \mu = 43 ;$$

$$H_1 : \mu > 43 ;$$

$$T = 2,07 ; c = 1,64.$$

Bác bỏ  $H_0$  và kết luận : "Năng suất trung bình vụ mùa này cao hơn 43 tạ/ha một cách có ý nghĩa".

$$113. \quad H_0 : \mu = 780 ;$$

$$H_1 : \mu > 780 ;$$

$$T = 2,03 ; c = z_{0,10} = 1,28.$$

Bác bỏ  $H_0$  và kết luận : "Trong ngày khuyến mại lượng hàng được bán nhiều hơn".

$$114. \text{ Giả thiết } H_0 : \mu = 45 ;$$

$$\text{Đối thiết } H_1 : \mu < 45 ;$$

$$T = \frac{(54 - 35)\sqrt{28}}{20} = 2,646.$$

Tra bảng phân bố Student với  $n - 1 = 28 - 1 = 27$  bậc tự do, ta tìm được  $c = t_{0,01} = 2,473$ . Vì  $T > c$  nên ta bác bỏ  $H_0$ .

**Kết luận :** Lời quảng cáo nói trên là "phóng đại".

$$115. \quad H_0 : \mu = 1800 ;$$

$$H_1 : \mu < 1800 ;$$

$$T = \frac{(1800 - 1740)\sqrt{16}}{60} = 4.$$

Tra bảng phân bố Student với  $n - 1 = 16 - 1 = 15$  bậc tự do, ta tìm được  $c = t_{0,05} = 1,753$ . Vậy ta bác bỏ  $H_0$  và kết luận : "Loại dây cáp nói trên chưa đạt yêu cầu thiết kế".

$$116. \quad H_0 : \mu = 48000 ;$$

$$H_1 : \mu < 48000 ;$$

$$T = \frac{(48000 - 45850)\sqrt{12}}{3600} = 1,18.$$

Tra bảng phân bố Student với  $n - 1 = 12 - 1 = 11$  bậc tự do, ta tìm được  $c = 2,201$ .

Vậy có thể cho rằng mức lương trung bình của công ty là 48000.

$$117. \quad H_0 : \mu = 43 ;$$

$$H_1 : \mu < 43 ;$$

$$T = \frac{(43 - 40,6)\sqrt{10}}{3,8} = 2$$

Tra bảng phân bố Student với  $n - 1 = 9$  bậc tự do, ta tìm được  $c = t_{0,05} = 1,833$ . Vì  $T > c$  nên bác bỏ  $H_0$  : "Kết quả của vận động viên thấp hơn mức yêu cầu một cách có ý nghĩa".

$$118. \quad H_0 : \mu = 42 ;$$

$$H_1 : \mu \neq 42 ;$$

$$T = \frac{(42 - 37,2)\sqrt{22}}{7,4} = 3,04$$

Tra bảng phân bố Student với  $n - 1 = 21$  bậc tự do, ta tìm được  $c = t_{0,01} = 2,518$ . Ta bác bỏ  $H_0$  : "Sự nghi ngờ của công ty là có cơ sở".

$$119. \quad H_0 : \mu = 5 ;$$

$$H_1 : \mu \neq 5 ;$$

$$T = 1,875.$$

Tra bảng phân bố Student với  $n - 1 = 7$  bậc tự do, ta tìm được  $c = t_{0,025} = 2,365$ .

Ta chấp nhận  $H_0$ .

$$120. \quad H_0 : p = 0,16 ;$$

$$H_1 : p \neq 0,16 ;$$

$$f = \frac{70}{600} \approx 0,116 ;$$

$$T = \frac{(0,116 - 0,16) \sqrt{600}}{\sqrt{(0,16)(0,84)}} = - 2,9 ;$$

Hàng số  $c = 1,96$ .

Vì  $|T| > c$  nên ta bác bỏ  $H_0$  : "Tỉ lệ không khỏi lệch khi dùng thuốc A và thuốc B khác nhau một cách có ý nghĩa ở mức 5%".  
P – giá trị là 0,0038.

121. Gọi  $p$  là tỉ lệ học sinh giỏi của tỉnh :

$$H_0 : p = 0,2 ;$$

$$H_1 : p < 0,2 ;$$

$$f = \frac{60}{400} = 0,15 ;$$

$$T = \frac{(0,20 - 0,15) \sqrt{400}}{\sqrt{(0,2)(0,8)}} = 2,5.$$

Với mức ý nghĩa 1% thì  $c = z_{0,01} = 2,33$ .  $T > c$  nên ta bác bỏ  $H_0$  : "Bộ có cơ sở để cho rằng tỉ lệ học sinh giỏi của tỉnh A thấp hơn 20%".

122.  $H_0 : p = 0,45 ;$

$H_1 : p = 0,45 ;$

$f = \frac{80}{200} = 0,4 ;$

$$T = \frac{(0,4 - 0,45)\sqrt{200}}{\sqrt{(0,45)(0,55)}} = -1,43 ;$$

Hàng số  $c = 1,96$ .

$|T| < c$  do đó ta không có cơ sở bác bỏ  $H_0$  : "Dự đoán của đảng đó có thể đúng".

123. Ta cần lựa chọn giữa hai giả thiết

$H_0 : p = 0,18 ;$

$H_1 : p > 0,18 ;$

$f = \frac{22}{80} = 0,275 ;$

$$T = \frac{(0,275 - 0,18)\sqrt{80}}{\sqrt{(0,18)(0,82)}} = 2,21 ; c = z_{0,02} = 2,05.$$

$T > c$  nên ta bác bỏ  $H_0$  : Trong các gia đình có trẻ em đi học thì tỉ lệ gia đình có máy tính cao hơn tỉ lệ chung một cách có ý nghĩa.

124.  $H_0 : p = \frac{2}{3} ;$

$H_1 : p < \frac{2}{3} ;$

$$f = \frac{55}{100} = 0,55 ;$$

$$T = \frac{\left(\frac{2}{3} - 0,55\right)\sqrt{100}}{\sqrt{\frac{2}{3} \cdot \frac{1}{3}}} = 2,49 ;$$

$$c = z_{0,05} = 1,64.$$

Vì  $T > c$  nên ta bác bỏ  $H_0$ . Ta kết luận rằng lời quảng cáo của công ty là quá sự thật.

**125.** Ta cần lựa chọn giữa hai giả thiết :

$$H_0 : p = 0,62 ;$$

$$H_1 : p < 0,62 ;$$

$$f = \frac{68}{130} = 0,523.$$

$f$  có phân bố xấp xỉ chuẩn với kì vọng 0,62 và độ lệch tiêu chuẩn

$$\sigma = \frac{\sqrt{(0,62)(0,38)}}{130} = 0,043.$$

$$\begin{aligned} \text{Vậy } P - \text{ giá trị là } P\{f \leq 0,523\} &= \Phi\left(\frac{0,523 - 0,62}{0,043}\right) = \\ &= \Phi(-2,26) = 0,0113. \end{aligned}$$

Giả thiết  $H_0$  sẽ bị bác bỏ và giả thiết  $H_1$  được chấp nhận ở mức ý nghĩa thấp nhất là 0,0113. Chẳng hạn,  $H_0$  bị bác bỏ ở mức 2%.

$$\mathbf{126.} \quad H_0 : p = 0,88 ;$$

$$H_1 : p < 0,88 ;$$

$$T = \frac{(0,88 - 0,82)\sqrt{100}}{\sqrt{(0,88)(0,12)}} = 1,875 ; c = 1,64.$$

Ta bác bỏ  $H_0$  và kết luận :

Tỉ lệ học sinh thi đỗ năm nay thấp hơn năm ngoái.

127.  $H_0 : p = 0,4 ;$

$$H_1 : p < 0,4 ;$$

$$f = \frac{125}{400} = 0,3125 ;$$

$$T = \frac{(0,4 - 0,3125)\sqrt{400}}{\sqrt{(0,4)(0,6)}} = 3,65 ; c = 1,64.$$

Ta bác bỏ  $H_0$  và kết luận : "Công ty quảng cáo quá sự thực".

128.  $H_0 : p = 0,12 ;$

$$H_1 : p \neq 0,12 ;$$

$$f = \frac{21}{200} = 0,105 ;$$

$$T = \frac{(0,105 - 0,12)\sqrt{200}}{\sqrt{(0,12)(0,88)}} = -0,65 ;$$

$$c = 2,58.$$

Chấp nhận  $H_0$  và kết luận : "Báo cáo của Bộ Y tế được chấp nhận là đúng".

129.  $H_0 : p = 0,75 ;$

$$H_1 : p < 0,75 ;$$

$$f = \frac{82}{120} = 0,6833 ;$$

$$T = \frac{(0,75 - 0,6833)\sqrt{120}}{\sqrt{(0,75)(0,25)}} = 1,75 ;$$

$$c = 1,64.$$

Bác bỏ  $H_0$  và kết luận : "Thuốc B tốt hơn thuốc A".

130.  $H_0 : p = 0,22 ;$

$$H_1 : p > 0,22 ;$$

$$f = \frac{10}{35} = 0,2857 ;$$

$$T = \frac{(0,2857 - 0,22)\sqrt{35}}{\sqrt{(0,22)(0,78)}} = 1 ; c = 1,64.$$

Chấp nhận  $H_0$ .

131.  $H_0 : p = 0,35 ;$

$$H_1 : p \neq 0,35 ;$$

$$f = \frac{22}{92} = 0,239 ;$$

$$|T| = 2,2 ;$$

P – giá trị = 0,0139.

P – giá trị bé hơn mức ý nghĩa  $\alpha = 0,025$ .

Vậy ta bác bỏ  $H_0$ .

132. Ta tính test thống kê  $\chi^2$  :

$$T = \frac{(106 - 100)^2}{100} + \dots + \frac{(112 - 100)^2}{100} = 4,42$$

Tra bảng phân bố  $\chi^2$  với  $6 - 1 = 5$  bậc tự do, ta tìm được  $c = \chi_{0,05}^2 = 11,070$ .

Vì  $T < c$  nên ta không có cơ sở bác bỏ giả thiết  $H_0$  : "Con xúc sắc được chế tạo cân đối, đồng chất".

**133.**  $H_0$  : "4 màu được ưa thích như nhau" ;

$H_1$  : "4 màu không được ưa thích như nhau" ;

$$T = \frac{(25 - 20)^2}{20} + \frac{(18 - 20)^2}{20} + \frac{(19 - 20)^2}{20} + \frac{(18 - 20)^2}{20}$$

$$= 1,70 ;$$

$$c = \chi_{0,05}^2 = 7,815 \text{ với 3 bậc tự do.}$$

Vì  $T < c$  nên ta không có cơ sở bác bỏ  $H_0$ .

**134.** Các tần số lí thuyết là  $\frac{200}{10} = 20$ . Từ đó

$$T = \frac{(10 - 20)^2}{20} + \dots + \frac{(22 - 20)^2}{20} = 9,70.$$

Hàng số  $c = \chi_{0,01}^2 = 21,666$  (9 bậc tự do). Ta không có cơ sở bác bỏ  $H_0$  : "Số liệu đã có chứng tỏ chiếc bánh xe quay xổ số chấp nhận được".

**135.** Các tần số lí thuyết là

Nâu : 56,4 ;

Xanh : 37,6 ;

Đỏ : 37,6 ;

Vàng : 37,6 ;

Da cam : 18,8.

$$\text{Vậy } T = \frac{(67 - 56,4)^2}{56,4} + \dots + \frac{(24 - 18,8)^2}{18,8} = 19,598$$



Hàng số  $c = \chi_{0,01}^2 = 13,277$  (4 bậc tự do). Vì  $T > c$  nên ta bác bỏ giả thiết  $H_0$  : "Phân bố về màu kẹo sôcôla không đúng như đã công bố".

**136.** Các tần số lí thuyết là :

Khối A : 230,4 ;

Khối B : 36 ;

Khối C : 93,6.

Từ đó  $T = 8,46$ .

Hàng số  $c = \chi_{0,01}^2 = 9,21$  (2 bậc tự do).

Chấp nhận giả thiết  $H_0$  : "Không có sự thay đổi về nguyện vọng đăng kí vào đại học so với năm ngoái".

**137.** Các tần số lí thuyết là

0 lần : 40 ;

1 lần : 76 ;

2 lần : 36 ;

3 lần : 27 ;

4 lần trở lên : 21.

Từ đó  $T = \frac{324}{40} + \frac{196}{76} + \frac{64}{36} + \frac{121}{27} + \frac{225}{21} = 27,653$

Hàng số  $c = \chi_{0,01}^2 = 13,227$  (4 bậc tự do).

Vì  $T > c$  nên ta bác bỏ  $H_0$  : "Số liệu thực nghiệm không phù hợp với công bố của Bộ Tư pháp".

**138.**  $T = 28,42$  ;

$c = 16,812$ .

Bác bỏ  $H_0$  : "Các sự cố về điện phân bố không đều giữa các ngày trong tuần".

$$139. \quad T = 3,91 ; \\ c = 19,675.$$

Chấp nhận  $H_0$  : "Có thể coi rằng con xúc sắc được chế tạo cân đối".

140. Giả thiết  $H_0$  : "Các sự cố "ốm" được phân bố đều trong các ngày làm việc".

Tần số lí thuyết đều bằng 50.

$$T = 10,48.$$

Tra bảng  $\chi^2$  với 4 bậc tự do, mức 5%, ta tìm được  $c = 9,488$ .

Bác bỏ  $H_0$  và kết luận : "Có những ngày nào đó công nhân hay nghỉ ốm hơn".

141. Giả thiết  $H_0$  : "Thức ăn A không có ảnh hưởng tới việc sinh con trai". Nếu  $H_0$  đúng thì số bé trai trong một gia đình có 3 con là một ĐLNN có quy luật nhị thức với tham số  $n = 3$ ,  $p = \frac{1}{2}$ .

Giả sử  $B_k$  là biến cố : "Trong 3 trẻ em có  $k$  đứa trẻ là con trai".

Ta có dưới giả thiết  $H_0$  :

$$P(B_0) = \frac{1}{8} ; P(B_1) = \frac{3}{8} ; P(B_2) = \frac{3}{8} ; P(B_3) = \frac{1}{8}.$$

Vậy các TSLT là : 10, 30, 30 và 10.

Số bé trai	3	2	1	0
TSQS	14	36	24	6
TSLT	10	30	30	10

Từ đó  $T = 1,3$ .

Với 3 bậc tự do, mức 5% ta tra bảng phân bố  $\chi^2$ , tìm được  
 $c = 7,815$ .

Vì  $T < c$  nên dẫn đến chấp nhận  $H_0$ .

**142.** Gọi  $X$  là số lỗi trong 300 trang in.

Xét các biến cố sau

$$B_1 = \{X \leq 2\}$$

$$B_2 = \{X = 3\}$$

$$B_3 = \{X = 4\}$$

$$B_4 = \{X = 5\}$$

$$B_5 = \{X = 6 \text{ hoặc } 7\}$$

$$B_6 = \{X = 8\}$$

Nếu  $H_0$  đúng thì  $X$  có phân bố Poisson với tham số  $\lambda = 4,7$ .

Khi đó :  $P(B_1) = 0,152$  ;  $P(B_2) = 0,157$  ;

$P(B_3) = 0,185$  ;  $P(B_4) = 0,174$  ;

$P(B_5) = 0,228$  ;  $P(B_6) = 0,104$ .

Các TSQS và TSLT được cho trong bảng dưới đây

	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	$B_6$
TSQS	10	6	13	10	9	2
TSLT	7,6	7,85	0,25	8,5	11,4	5,2

Tính toán cho ta  $T = 5,385$ .

Tra bảng  $\chi^2$  với  $6 - 1 = 5$  bậc tự do, ta tìm được

$$c = \chi_{0,05}^2 = 11,070.$$

Vì  $T < c$  nên ta chấp nhận  $H_0$ .

143.  $T = 22,3$  ;

$c = 7,815$ .

Vì  $T > c$  nên dẫn tới bác bỏ  $H_0$ . Kết luận : "Dân cư của nước Y có phân bố nhóm máu khác với dân cư của nước X".

144. Ta có  $n = 30$  ;

$s = 0,032$  ;

do đó  $s^2 = (0,032)^2 = 0,001024$ .

Với  $\beta = 0,95$  thì  $\frac{\alpha}{2} = \frac{1 - \beta}{2} = 0,025$  và  $1 - \frac{\alpha}{2} = 0,975$ .

Tra bảng phân bố  $\chi^2$  với 29 bậc tự do ta tìm được :

$\chi_{0,025}^2 = 45,772$  ;

$\chi_{0,975}^2 = 16,047$

Vậy khoảng tin cậy 95% cho  $\sigma^2$  là

$$\left[ \frac{29(0,001024)}{45,772}, \frac{29(0,001024)}{16,047} \right]$$

hay  $[0,000649, 0,001851]$ .

Từ đó khoảng tin cậy cho  $\sigma$  sẽ là

$[\sqrt{0,000649}, \sqrt{0,001851}]$

hay  $[0,025 ; 0,043]$ .

145.  $H_0 : \sigma^2 = (0,2)^2 = 0,04$  ;

$H_1 : \sigma^2 > 0,04$ .

Tra bảng phân bố  $\chi^2$  ta tìm được

$c = \chi_{0,05}^2 = 19,98$

(số bậc tự do là  $n - 1 = 12 - 1 = 11$ ).

$$T = \frac{(12 - 1)(0,09)}{0,04} = 24,75.$$

$T < c$  do đó ta bác bỏ giả thiết  $H_0$ . Kết luận : "Dây chuyền sản xuất cần phải điều chỉnh lại vì độ biến động người đã lớn hơn mức cho phép".

**146.**  $9,87 \leq \sigma \leq 16,66$ .

**147.**  $H_0 : \sigma^2 = (2,75)^2 = 7,5625 ;$

$$H_1 : \sigma^2 < 7,5625 ;$$

$$T = \frac{29(2,32)^2}{(2,75)^2} = 20,63.$$

Tra bảng phân bố  $\chi^2$  với 29 bậc tự do ta tìm được :

$$c = \chi_{0,95}^2 = 17,708$$

Ta có  $T > c$ , do đó ta chấp nhận  $H_0$  và kết luận : "Sự cải tiến chưa làm giảm độ lệch tiêu chuẩn một cách có ý nghĩa".

## Chương 5

### BÀI TOÁN SO SÁNH

148. a) Gọi  $\mu_1$ ,  $\mu_2$  tương ứng là tuổi trung bình các bà mẹ sinh con lần cuối ở vùng A và vùng B.

Ta có bài toán kiểm định giả thiết sau

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 \neq \mu_2 ;$$

$$T = \frac{32,9 - 29,6}{\sqrt{\frac{(5,7)^2}{36} + \frac{(5,5)^2}{49}}} = \frac{3,3}{1,23} = 2,68.$$

Hàng số c ứng với  $\alpha = 5\%$  là 1,96. Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$  và kết luận : "Có sự khác nhau giữa  $\mu_1$  và  $\mu_2$ ".

b) P – giá trị là  $2(0,037) = 0,074$ .

149.  $H_0 : \mu_1 = \mu_2 ;$

$$H_1 : \mu_1 \neq \mu_2 ;$$

$$T = \frac{20,44 - 19,45}{\sqrt{\frac{(2,96)^2}{30} + \frac{(2,13)^2}{30}}} = \frac{0,99}{0,666} = 1,49.$$

Hàng số c là 2,58.

Vì  $T < c$  nên ta chấp nhận  $H_0$  và kết luận : "Không có sự khác nhau giữa tuổi trung bình của các bà mẹ sinh con lần cuối ở hai vùng A và B".

**150.** Gọi  $\mu_1, \mu_2$  tương ứng là điểm thi trung bình của học viên của cơ sở A và B. Do điều kiện bài toán ta tin chắc rằng  $\mu_1 \leq \mu_2$ . Vậy ta cần lựa chọn giữa hai giả thiết :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 > \mu_2 ;$$

$$\text{Ta có } T = \frac{9,4 - 9}{\sqrt{\frac{0,8^2}{100} + \frac{1^2}{80}}} = 2,91.$$

Hàng số c là 2,33.

Ta có  $T > c$ , vậy giả thiết  $H_0$  cần phải bác bỏ ; kết luận : "Điểm thi trung bình của cơ sở A thực sự là cao hơn cơ sở B'.

**151.** Kí hiệu  $\mu_1$  và  $\mu_2$  tương ứng là mức lương trung bình của nữ và nam. Ta có bài toán kiểm định

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 < \mu_2 ;$$

$$T = \frac{8,06 - 7,23}{\sqrt{\frac{1,85^2}{75} + \frac{1,64^2}{100}}} = 3,07$$

Với mức ý nghĩa  $\alpha = 0,05$ , ta có hàng số c = 2,33.

Vì  $T > c$  nên dẫn tới bác bỏ  $H_0$ . Kết luận : "Mức lương trung bình của phụ nữ thấp hơn mức lương trung bình của nam giới trong công ty này".

**152.** Gọi  $\mu_1, \mu_2$  tương ứng là tuổi thọ trung bình của pin do công ty A và công ty B sản xuất ;

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 \neq \mu_2 ;$$

$$s^2 = \frac{(10-1)(1,1)^2 + (12-1)(0,9)^2}{10+12-2} = \frac{19,8}{20} = 0,99 ;$$

$$T = \frac{4,8 - 4,3}{\sqrt{0,99 \left( \frac{1}{10} + \frac{1}{12} \right)}} = \frac{0,5}{0,426} = 1,174.$$

Với mức ý nghĩa 1%, tra bảng phân bố Student với 20 bậc tự do, ta tìm được  $c = t_{0,005} = 2,845$ .

Vì  $|T| < c$  nên ta chấp nhận giả thiết  $H_0$ .

*Kết luận* : "Tuổi thọ trung bình của hai loại pin như nhau".

**153.** Gọi  $\mu_1$  là sản lượng trung bình khi bón lót 100 đơn vị đạm và  $\mu_2$  là sản lượng trung bình khi bón lót 50 đơn vị đạm. Ta thừa nhận rằng phương sai của hai sản lượng bằng nhau (giả thiết này tỏ ra hợp lí vì đối với cùng một loại cây, sự chăm sóc khác nhau không làm thay đổi phương sai của sản lượng).

Giả thiết  $H_0 : \mu_1 = \mu_2 ;$

Đối thiết  $H_1 : \mu_1 > \mu_2.$

Tính toán cho ta

$$x = 46,54 ;$$

$$y = 43,35 ;$$

$$s = 3,84 ;$$

$$T = \frac{46,54 - 43,85}{3,84 \sqrt{\frac{1}{8} + \frac{1}{11}}} = 1,69.$$



Tra bảng phân bố Student với  $8 + 11 - 2 = 17$  bậc tự do, ta tìm được

$$c = t_{0,05} = 1,74.$$

Vì  $T < c$  nên ta quyết định giữ giả thiết  $H_0$ . Kết luận : "Chưa có cơ sở để cho rằng bón 100 đơn vị đạm tốt hơn 50 đơn vị đạm".

$$154. \quad T = \frac{6,2 - 5,3}{\sqrt{\frac{1,8^2}{40} + \frac{2,1^2}{50}}} = 2,2$$

Hằng số  $c$  ứng với  $\alpha = 5\%$  là 1,96. Vì  $|T| > c$  nên ta quyết định bác bỏ  $H_0$  : "Thời gian nằm viện trung bình của bệnh nhân nam và nữ là khác nhau".

155. Giả sử  $\mu_A, \mu_B$  tương ứng là năng suất trung bình của cây cà chua khi bón các loại phân A, B. Ta có bài toán kiểm định hai phía

$$H_0 : \mu_A = \mu_B ; H_1 : \mu_A \neq \mu_B ;$$

$$T = \frac{32,2 - 28,4}{\sqrt{\frac{8,5^2}{60} + \frac{9,3^2}{72}}} = 2,45.$$

Với  $\alpha = 5\%$  ta có  $c = 1,96$ .

Vì  $|T| > c$  nên ta bác bỏ  $H_0$ . Ta kết luận : "Tác dụng của hai loại phân A và B là khác nhau ở mức ý nghĩa 5%".

156. Giả sử  $\mu_1$  và  $\mu_2$  tương ứng là số ô tô tải trung bình gặp trục trặc kĩ thuật trong một ngày mùa đông và một ngày mùa hè. Ta có bài toán kiểm định giả thiết :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 \neq \mu_2 ;$$

$$T = \frac{16,6 - 12,4}{\sqrt{\frac{7,1^2}{35} + \frac{5,8^2}{32}}} = 2,66.$$

Hàng số  $c = 2,33$  ứng với  $\alpha = 2\%$ .

Vì  $|T| > c$  nên ta bác bỏ  $H_0$  và kết luận có sự khác nhau giữa một ngày mùa đông và một ngày mùa hè đối với vấn đề trục trặc kĩ thuật của ô tô.

**157.** Giả sử  $\mu_1$  và  $\mu_2$  tương ứng là chiều cao trung bình của thanh niên vùng A và vùng B. Ta có bài toán kiểm định giả thiết :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 \neq \mu_2 ;$$

Ta có  $\bar{x} = 169,5 ;$

$$\bar{y} = 169,1 ;$$

$$s^2 = 11,1 ; s = 3,33 ;$$

$$T = \frac{169,5 - 169,1}{3,33 \sqrt{\frac{1}{10} + \frac{1}{10}}} = 0,268 ;$$

$c = 2,88$  là phân vị mức 0,005 của phân bố Student với 18 bậc tự do.

Vì  $|T| < c$  nên ta kết luận :

"Không có sự khác nhau giữa chiều cao trung bình của thanh niên hai vùng A và B".

**158.**  $\bar{x} = 110,4 ;$

$$\bar{y} = 101,2 ;$$

$$s = 11,21 ;$$

$$T = \frac{110,4 - 101,2}{11,21 \sqrt{\frac{1}{10} + \frac{1}{10}}} = 1,82.$$

Với bài toán kiểm định một phía :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 > \mu_2 ;$$

ở đó  $\mu_1, \mu_2$  tương ứng là chỉ số IQ trung bình của trẻ bú mẹ và trẻ bú chai. Tra bảng phân bố Student với  $10 + 10 - 2 = 18$  bậc tự do, ta tìm được

$$c = t_{0,05} = 1,734.$$

Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Các trẻ bú mẹ có chỉ số IQ cao hơn một cách có ý nghĩa".

**159.** Ta có bài toán kiểm định giả thiết một phía :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 > \mu_2 ;$$

trong đó  $\mu_1, \mu_2$  tương ứng là độ tăng trọng trung bình ứng với khẩu phần ăn nhiều đạm và ít đạm.

$$\text{Ta có } \bar{x} = 120 ;$$

$$\bar{y} = 101 ;$$

$$s^2 = 446,12 ;$$

$$s = 21,12 ;$$

$$T = \frac{120 - 101}{21,12 \sqrt{\frac{1}{12} + \frac{1}{7}}} = 1,89.$$

Với mức ý nghĩa  $\alpha = 5\%$ , tra bảng phân bố Student với  $12 + 7 - 2 = 17$  bậc tự do, ta tìm được

$$c = t_{0,05} = 1,74.$$

Vì  $T > c$  nên ta kết luận rằng khẩu phần ăn nhiều đạm có tác dụng tăng trọng một cách có ý nghĩa.

**160.** Kiểm định một phía :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 > \mu_2 ;$$

$$T = 0,786 ;$$

$$c = t_{0,05} = 1,734 \text{ (bậc tự do} = 18).$$

Vì  $T < c$  nên ta giữ giả thiết  $H_0$  : "Chưa có cơ sở để cho rằng loại máy mới cắt nhanh hơn máy cũ".

$$\mathbf{161. Ta có } \bar{x} - \bar{y} = \frac{245 - 65}{10} = 18 ;$$

$$\begin{aligned} \sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2 &= \sum x_i^2 - \frac{(\sum x)^2}{n} + \sum y_i^2 - \frac{(\sum y)^2}{n} = \\ &= 621 - \frac{65^2}{10} + 6503 - \frac{245^2}{10} = 699. \end{aligned}$$

$$\text{Vậy } s^2 = \frac{699}{18} = 38,83 ;$$

$$s = 6,23 ;$$

$$T = \frac{18}{6,23 \sqrt{\frac{1}{10} + \frac{1}{10}}} = 6,47.$$

Với mức ý nghĩa  $\alpha = 5\%$ , tra bảng phân bố Student với 18 bậc tự do, ta có  $c = t_{0,05} = 1,73$ .

Ta có  $T > c$  nên  $H_0$  bị bác bỏ. Kết luận : "Việc bón thêm magiê là có tác dụng".

162. Ta có  $\bar{x} = \frac{13}{127} = 0,102$  ;

$$\bar{y} = \frac{55}{265} = 0,208 ;$$

$$s_x^2 = \frac{1}{126} \left( 313 - \frac{13^2}{127} \right) = 2,47 ;$$

$$s_y^2 = \frac{1}{264} \left( 767 - \frac{55^2}{265} \right) = 2,86 ;$$

$$T = \frac{0,208 - 0,102}{\sqrt{\frac{2,47}{127} + \frac{2,86}{265}}} = \frac{0,106}{0,173} = 0,612$$

Ta có  $c = 1,96$ .

Vì  $T < c$  nên ta quyết định giữ giả thiết  $H_0$ . Kết luận : "Không có sự khác nhau giữa trọng lượng não của người trên 50 tuổi và người dưới 50 tuổi".

163. Ta lập bảng các giá trị của hiệu số  $d = x_A - x_B$ .

d :	-0,6	0,4	-0,20	-0,19	0,10
	-0,14	0	0,20	-0,14	-0,09
	-0,20	-0,20			

Từ đó ta có  $\bar{d} = -0,073$  ;

$$s_d = 0,133.$$

Ta có bài toán kiểm định :

$$H_0 : \mu_d = 0$$

$$H_1 : \mu_d \neq 0 ;$$

$$T = \frac{(-0,073)\sqrt{12}}{0,133} = -1,921.$$

Với mức ý nghĩa 2%, tra bảng phân bố Student với  $n - 1 = 12 - 1 = 11$  bậc tự do, ta có  $c = t_{0,01} = 2,718$ .

Vì  $|T| = 1,921 < 2,718$  nên ta không có cơ sở bác bỏ  $H_0$ .

Kết luận : "Giá bán ở hai cửa hiệu như nhau".

**164.**  $\bar{d} = 5 ; s_d = 8,01 ;$

$$T = \frac{\bar{d}\sqrt{n}}{s_d} = \frac{5\sqrt{10}}{8,01} = 1,97 ;$$

$$c = t_{0,05} = 1,833 \text{ với } 9 \text{ bậc tự do.}$$

Ta bác bỏ  $H_0$  và kết luận : "Chế độ ăn kiêng có tác dụng làm giảm trọng lượng".

**165.**  $T = 1,897.$

Ta có bài toán kiểm định một phía :

$$H_0 : \mu_d = 0$$

$$H_1 : \mu_d > 0.$$

$$\begin{aligned} \text{Hàng số } c &= t_{0,05} \text{ (với } 7 \text{ bậc tự do)} \\ &= 1,895. \end{aligned}$$

Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Có sự tăng giá đối với mặt hàng A".

**166.**  $T = 0,56.$

Hàng số  $c = t_{\frac{\alpha}{2}} = t_{0,025} = 2,447$  với 6 bậc tự do. (Ở đây ta sử dụng bài toán kiểm định một phía  $H_0 : \mu_d = 0 ; H_1 : \mu_d > 0$ ).

Vì  $T < c$  nên ta giữ giả thiết  $H_0$ .

Kết luận : "Không có sự khác nhau giữa điểm trung bình của một sinh viên ở năm thứ nhất và năm thứ hai".

**167.** Gọi  $\mu_1$  và  $\mu_2$  theo thứ tự là mức tăng trọng trung bình của hai nhóm thỏ. Ta có bài toán kiểm định hai phía :

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 \neq \mu_2.$$

Ta có  $\bar{x}_1 = 514,89$

$$\bar{x}_2 = 638 ;$$

$$s^2 = 3118,8 ;$$

$$s = 55,84 ;$$

$$T = \frac{638 - 526}{55,84 \sqrt{\frac{1}{8} + \frac{1}{9}}} = \frac{112}{27,13} = 4,63$$

Tra bảng phân bố Student với 15 bậc tự do, ta tìm được

$$c = t_{\frac{\alpha}{2}} = t_{0,005} = 2,95.$$

Vì  $|T| > c$  nên ta bác bỏ  $H_0$ . Kết luận : "Côban có ảnh hưởng tới sự tăng trọng".

**168.** Gọi  $\mu_A$ ,  $\mu_B$  tương ứng là chiều cao trung bình của đàn ông tỉnh A và tỉnh B.

Ta có bài toán kiểm định hai phía :

$$H_0 : \mu_A = \mu_B ;$$

$$H_1 : \mu_A \neq \mu_B.$$

Ta có  $\bar{x}_A = 168$  ;

$$\bar{x}_B = 171 ;$$

$$s = 7,62 ;$$

$$T = \frac{168 - 171}{7,62 \sqrt{\frac{1}{6} + \frac{1}{7}}} = \frac{-3}{4,2} = -0,71.$$

Tra bảng phân bố Student với 11 bậc tự do, ta tìm được

$$c = t_{\frac{\alpha}{2}} = t_{0,025} = 2,2$$

Vì  $|T| = 0,71 < c$ , nên ta không có cơ sở bác bỏ  $H_0$ . Kết luận :  
"Không có sự khác nhau về chiều cao trung bình của đàn ông hai  
tử A, B".

**169.** Hiệu số  $d = x_A - x_B$  như sau :

$$-0,56 \quad 1,78 \quad 0,97 \quad 1,83 \quad 1,05 \quad 2,9 \quad -1,13$$

$$\bar{d} = 0,98 ; s_d = 1,402 ;$$

$$T = \frac{0,98\sqrt{7}}{1,402} = 1,85.$$

Ta có bài toán điểm định :

$$H_0 : \mu_d = 0$$

$$H_1 : \mu_d \neq 0.$$

Tra bảng phân bố Student với 6 bậc tự do, ta tìm được

$$c = t_{0,025} = 2,447.$$

Vì  $|T| < c$  nên ta chấp nhận  $H_0$ .

Kết luận : "Năng suất hai giống lúa A và B như nhau".

**170.**  $\bar{d} = 0,52 ;$

$$s_d = 0,2939 ;$$



$$T = \frac{0,52\sqrt{6}}{0,2939} = 4,33 ;$$

$$c = t_{0,025} = 2,57 \text{ (với 5 bậc tự do).}$$

Vì  $|T| > c$  nên ta bác bỏ  $H_0$ .

Kết luận : "Năng suất hai giống lúa là khác nhau".

**171.** Tổng các dạng của đường X là :

$$R = 1 + 2 + 3 + 4 + 7 + 10 + 12 + 13,5 + 18 + 20 = 90,5 ;$$

$$\mu_R = \frac{10(10 + 10 + 1)}{2} = 105 ;$$

$$\sigma_R^2 = \frac{(10)(10)(10 + 10 + 1)}{12} = 175 ;$$

$$T = \frac{90,5 - 105}{\sqrt{175}} = - 1,1.$$

Với mức ý nghĩa  $\alpha = 5\%$  hằng số  $c$  là 1,96.

Ta có  $|T| < 1,96$ . Vậy giả thiết  $H_0$  được chấp nhận. Kết luận : "Thời gian trung bình đi trên đường X và đường Y của người lái xe đó như nhau".

**172.** Hạng của các bệnh nhân nữ và nam được cho như sau :

Nam	
Thời gian	Hạng
13	7
15	9
9	2,5
18	12
11	5,5
20	13
24	16
22	15
25	17
Tổng	97

Nữ	
Thời gian	Hạng
11	5,5
14	8
10	4
8	1
16	10
9	2,5
17	11
21	14
Tổng	56

$R = 56$  (tổng hạng của nữ) ;

$$\mu_R = \frac{8(9 + 8 + 1)}{2} = 72 ;$$

$$\sigma_R^2 = \frac{9.8(9 + 8 + 1)}{12} = 108 ;$$

$$T = \frac{56 - 72}{\sqrt{108}} = -1,54.$$

Với mức ý nghĩa  $\alpha = 5\%$ , hằng số  $c$  là 1,96.

Vì  $|T| < c$  nên giả thiết  $H_0$  được chấp nhận. Ta chưa có đủ cơ sở để kết luận rằng bệnh nhân nam nằm viện lâu hơn bệnh nhân nữ.

173.  $H_0$  : không có sự khác nhau ;

$H_1$  : có sự khác nhau

Xí nghiệp A		Xí nghiệp B	
Số ngày nghỉ	Hạng	Số ngày nghỉ	Hạng
5	4,5	1	1
6	6	2	2
8	8,5	4	3
10	11,5	5	4,5
11	13	7	7
13	14	8	8,5
14	15	9	10
15	16	10	11,5
Tổng	88,5	16	17
		17	18
		Tổng	82,5

$R =$  tổng hạng của A = 88,5 ;

$$\mu_R = \frac{8(8 + 10 + 1)}{2} = 76$$

$$\sigma_R^2 = \frac{8(10)(8 + 10 + 1)}{12} = 126,66 ;$$

$$T = \frac{88,5 - 76}{\sqrt{126,66}} = 1,11.$$

Với mức ý nghĩa  $\alpha = 10\%$ , hằng số c là  $c = z_{\alpha/2} = 1,645$

Vì  $|T| < c$  nên ta chấp nhận  $H_0$ .

**174.**  $H_0$  : Không có sự khác nhau về chất lượng

$H_1$  : Có sự khác nhau về chất lượng

Nhà máy A	
Tuổi thọ	Hạng
372	2
283	1
712	10
849	15
623	5
382	3
427	4
821	13
Tổng	53

Nhà máy B	
Tuổi thọ	Hạng
645	6
682	7
913	16
742	11
691	9
689	8
842	14
751	12
Tổng	83

$R = \text{tổng hạng của A} = 53 ;$

$$\mu_r = \frac{8(8 + 8 + 1)}{2} = 68 ;$$

$$\sigma_R^2 = \frac{8.8(8 + 8 + 1)}{12} = 90,666 ;$$

$$T = \frac{53 - 68}{\sqrt{90,666}} = \frac{-15}{9,5} = -1,58.$$

Với  $\alpha = 5\%$  thì  $c = 1,96$ .

Vì  $|T| = 1,58 < c$  nên ta không có cơ sở bác bỏ  $H_0$ .

175.

Nước phát triển (A)		Nước đang phát triển (B)	
Chiều cao	Hạng	Chiều cao	Hạng
30,2	2	29,4	1
34,6	7	33,8	6
37,8	9	37,5	8
40,8	14	40,7	13
43,4	15,5	43,4	15,5
45,9	17	31,2	3
39,2	10	32,1	4
40,4	12	33,4	5
40	11		
Tổng	97,5	Tổng	55,5

$$R = \text{tổng hạng A} = 97,5 ;$$

$$\mu_R = \frac{9(9 + 8 + 1)}{2} = 81 ;$$

$$\sigma_R^2 = \frac{9.8(9 + 8 + 1)}{12} = 108 ;$$

$$T = \frac{97,5 - 81}{\sqrt{108}} = \frac{16,5}{10,39} = 1,59$$

Hàng số  $c = 1,96 > T$ . Do đó ta không có cơ sở bác bỏ  $H_0$ . Các số liệu thu được chưa cho phép ta kết luận rằng phân bố chiều cao của trẻ em 6 tuổi ở các nước phát triển và đang phát triển là khác nhau.

**176.** Dấu + chứng tỏ phương pháp giáo dục thành công, trẻ em sau khi được giáo dục sẽ tự tin hơn trước. Dấu – chứng tỏ phương pháp giáo dục thất bại.

Ta có kết quả sau

Trẻ em	Dấu
1	+
2	+
3	+
4	+
5	-
6	+
7	+
8	+
9	+
10	+
11	-
12	+
13	0

Như vậy số dấu khác không là  $\tilde{n} = 12$ . Số dấu cộng là  $n^+ = 10$ .

$$T = \frac{n^+ - 0,5\tilde{n}}{0,5\sqrt{\tilde{n}}} = \frac{10 - \frac{12}{2}}{\frac{1}{2}\sqrt{12}} = \frac{8}{\sqrt{12}} = 2,3$$

Với mức ý nghĩa 5%, hằng số  $c = 1,28$  (kiểm định 1 phía).

Ta bác bỏ  $H_0$  và kết luận rằng phương pháp giáo dục có làm tăng tính tự tin của trẻ em.

177.  $H_0 : p = 0,5 ;$

$H_1 : p \neq 0,5 ;$

$\tilde{n} = 17$  vì có 3 hiệu số bằng 0 ;

$n^+ = 12 ;$

$$T = \frac{12 - (17)(0,5)}{(0,5)\sqrt{17}} = \frac{3,5}{2,06} = 1,69.$$

Hàng số  $c$  ứng với mức  $\alpha = 5\%$  là 1,96. Vì  $T < c$  nên ta chấp nhận  $H_0$ . Kết luận : "Không có sự khác nhau giữa vợ và chồng về tâm trạng lạc quan".

**178.** Kí hiệu  $p$  là tỉ lệ học sinh có điểm khi sau ôn tập cao hơn điểm thi trước ôn tập. Ta có bài toán kiểm định một phía :

$$H_0 : p = 0,5 ;$$

$$H_1 : p > 0,5.$$

Gọi  $d$  là hiệu số giữa điểm thi sau và điểm thi trước. Có cả thấy 18 giá trị  $d \neq 0$  với 13 dấu +. Vậy

$$\bar{n} = 18 ; n^+ = 13 ;$$

$$T = \frac{13 - 18(0,5)}{0,5\sqrt{18}} = \frac{4}{(0,5)(4,24)} = 1,886.$$

Với mức ý nghĩa  $\alpha = 0,05$  ta có

$$c = z_{0,05} = 1,64.$$

$$T > c \text{ điều này dẫn tới bác bỏ } H_0.$$

Kết luận : "Việc cho học sinh ôn tập có tác dụng".

**179.**  $R = 100$ ,  $T = 0,38$ . Chấp nhận  $H_0$ .

Kết luận : "Số biên lai phạt của hai cảnh sát có phân bố như nhau".

**180.**  $R = 71,5 ;$

$$T = 1,27 ; c = 1,96 ; T < c.$$

Chấp nhận  $H_0$  : "Số khách trung bình mà mỗi tiếp viên A và B phục vụ trong một ngày là như nhau".

**181.** Gọi  $\mu_1$  và  $\mu_2$  là điểm trung bình của sinh viên nam và sinh viên nữ. Ta có bài toán kiểm định giả thiết

$$H_0 : \mu_1 = \mu_2 ;$$

$$H_1 : \mu_1 \neq \mu_2.$$

Sử dụng tiêu chuẩn hạng Mann – Whitney ta tìm được :

$$R = 253,3 ;$$

$$\mu_R = \frac{1}{2} (14)(14 + 14 + 1) = 203 ;$$

$$\sigma_R^2 = \frac{(14)(14)(14 + 14 + 1)}{12} = 473,66 ;$$

$$\sigma_R = 21,76 ;$$

$$T = \frac{253,3 - 203}{21,76} = 2,297.$$

Vì  $T > c = 1,96$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Điểm trung bình của sinh viên nam và sinh viên nữ khác nhau".

**182.**  $R = 74 ;$

$$\mu_R = 85,5 ;$$

$$\sigma_R^2 = \frac{(9)(9)(9 + 9 + 1)}{12} = 128,25 ;$$

$$\sigma_R = 11,32 ;$$

$$T = \frac{74 - 85,5}{11,32} = - 1,015.$$

Hàng số c ứng với  $\alpha = 2\%$  là

$$c = z_{0,01} = 2,33.$$

Vì  $|T| < c$  nên ta chấp nhận  $H_0$ .

**183.**  $R = 223,5$

$$\mu_R = \frac{16(16 + 16 + 1)}{2} = 264 ;$$

$$\sigma_R^2 = \frac{(16)(16)(16 + 16 + 1)}{12} = 704 ; \sigma_R = 26,53 ;$$

$$T = \frac{223,5 - 264}{26,53} = 1,526$$

$$T < c = 1,96$$

Ta chấp nhận  $H_0$ . Kết luận : "Năng suất lao động của công nhân nam và công nhân nữ như nhau".

**184. a)**  $\bar{n} = 20$  (vì có 2 hiệu số bằng 0) ;

$$n^+ = 12 ;$$

$$T = \frac{12 - 20(0,5)}{(0,5)\sqrt{20}} = \frac{4}{4,47} = 0,894.$$

Với mức  $\alpha = 5\%$  thì  $c = 1,96$ . Ta có  $T < c$  nên chưa có cơ sở bác bỏ  $H_0$ .

**b)**  $R^+ = 126,5 ;$

$$\bar{n} = 20 ;$$

$$T = \frac{126,5 - \frac{20(20 + 1)}{4}}{\sqrt{\frac{20(20 + 1)(40 + 1)}{24}}} = \frac{21,5}{26,78} = 0,803.$$

Ta chấp nhận  $H_0$ .

**185.** Từ số liệu trên ta tính được  $R^+ = 122$  và  $\bar{n} = 23$  (vì có hai số liệu có hiệu số bằng 0).

$$T = \frac{122 - \frac{23(23 + 1)}{4}}{\sqrt{\frac{23(23 + 1)(46 + 1)}{24}}} = \frac{-16}{32,88} = -0,4866.$$



Vì  $|T| = 0,4866 < 1,96 = c$  nên ta chấp nhận  $H_0$ . Kết luận :  
"Chiều cao trung bình của hai anh em trai là như nhau".

**186.** Ta có bảng sau đây :

Đối tượng	Hiệu $d = A - B$	Hạng của $ d $	Hạng dương	Hạng âm
1	- 36	9		9
2	0			
3	+6	1	1	
4	+35	8	8	
5	+30	7	7	
6	+7	2,5	2,5	
7	-7	2,5		2,5
8	-21	6		6
9	+10	5	5	
10	-9	4		4

$$R^+ = 23,5 ;$$

$$R^- = 21,5 ;$$

$$R = \min(23,5 ; 21,5) = 21,5 ;$$

$$\tilde{n} = 9 ;$$

$$T = \frac{21,5 - 9(9 + 1)/4}{\sqrt{\frac{9(9 + 1)(18 + 1)}{24}}} = \frac{1}{8,44} = 0,118.$$

Ta có  $|T| < c$ , do đó ta không có cơ sở bác bỏ  $H_0$ .

**187.** Ta có bảng sau đây :

Số thứ tự	Hiệu số d	Hạng của  d	Hạng dương
1	3	6	6
2	2	4	4
3	8	8	8
4	-2	4	
5	2	4	4
6	5	7	7
7	1	1,5	1,5
8	-1	1,5	
			30,5

Ta có  $R^+ = 30,5$  ;

$$\tilde{n} = 8 ;$$

$$T = \frac{30,5 - 8(8 + 1)/4}{\sqrt{\frac{8(8 + 1)(16 + 1)}{24}}} = \frac{12,5}{7,14} = 1,75$$

Với mức ý nghĩa  $\alpha = 10\%$  thì

$$c = z_{\frac{\alpha}{2}} = 1,65$$

Vì  $|T| > c$  nên ta bác bỏ  $H_0$ . Kết luận : "Tuổi trung bình của vợ và chồng là khác nhau".

**188.** Ta có :

$$\tilde{n} = 10 ;$$

$$n^+ = 1 ;$$

$$T = \frac{n^+ - 0,5\tilde{n}}{0,5\sqrt{\tilde{n}}} = \frac{1 - 5}{0,5\sqrt{10}} = \frac{-8}{\sqrt{10}} = -2,531.$$

Với mức ý nghĩa  $\alpha = 5\%$  và bài toán kiểm định ở đây là 1 phía nên

$$c = z_{\alpha} = 1,65.$$

Vì  $|T| > 1,65$  nên ta quyết định bác bỏ  $H_0$  : "Các số liệu đã chứng tỏ số điều thuốc hút trung bình trong 1 tuần của các đối tượng được tuyên truyền đã giảm một cách có ý nghĩa".

**189.** Ta có bảng sau đây :

Hiệu số d	Hạng	$R^+$	$R^-$
0			
-11	9		9
0			
-9	7,5		7,5
-2	2,5		2,5
-12	10		10
-5	4		4
+2	2,5	2,5	
-1	1		1
-9	7,5		7,5
-8	5,5		5,5
-8	5,5		5,5
		2,5	52,5

Ta có  $R^+ = 2,5$  ;

$$\tilde{n} = 10 ;$$

$$T = \frac{2,5 - 10(10 + 1)/4}{\sqrt{\frac{10(10 + 1)(20 + 1)}{24}}} = \frac{-25}{9,81} = -2,548$$

Với mức ý nghĩa  $\alpha = 5\%$  thì  $c = 1,96$ .

Ta bác bỏ  $H_0$  vì  $|T| > c$ .

190. a)  $\bar{n} = 24$  ;

$$n^+ = 19 ;$$

$$T = \frac{19 - 12}{0,5\sqrt{24}} = \frac{7}{\sqrt{6}} = 2,85$$

Với mức ý nghĩa 1% và bài toán kiểm định một phía  $c = 2,33$ .

Vì  $T > c$  nên ta bác bỏ  $H_0$  và kết luận rằng : "Vùng A có mưa nhiều hơn vùng B".

b) Ta có

$$\bar{n} = 24 ;$$

$$R^+ = 273 ; R^- = 27 ;$$

$$T = \frac{273 - \frac{(24)(24 + 1)}{4}}{\sqrt{\frac{24(24 + 1)(48 + 1)}{24}}} = \frac{123}{35} = 3,51.$$

Với mức ý nghĩa 1%,  $c = 2,58$ .

Ta có  $|T| > c$  do đó giả thiết  $H_0$  bị bác bỏ.

191. Gọi  $p_1$  và  $p_2$  là tỉ lệ cử tri nam và nữ bỏ phiếu cho ông A.  
Giả thiết

$$H_0 : p_1 = p_2$$

với đối thiết  $H_1 : p_1 \neq p_2$ .

Các số liệu mẫu cho ta

$$n_1 = 100 ; k_1 = 42 ; f_1 = \frac{42}{100} = 0,42 ;$$

$$n_2 = 200 ; k_2 = 92 ; f_2 = \frac{92}{200} = 0,46.$$

$$\text{Tần suất chung } f = \frac{42 + 92}{100 + 200} = \frac{134}{300} = 0,447.$$

$$T = \frac{0,42 - 0,46}{\sqrt{(0,447)(0,553)\left(\frac{1}{100} + \frac{1}{200}\right)}} = -0,66.$$

Với mức ý nghĩa  $\alpha = 5\%$  ta có  $c = 1,96$ . Vì  $|T| = 0,66 < c$  nên ta chấp nhận  $H_0$ . Kết luận : "Tỉ lệ cử tri nam bầu cho A và tỉ lệ cử tri nữ bầu cho A không có sự khác biệt".

**192.** Gọi  $p_1$  là tỉ lệ những người ưa thích Côca – Cola với công thức mới,  $p_2$  là tỉ lệ những người ưa thích Côca – Cola với công thức cũ. Ta có bài toán kiểm định giả thiết :

$$H_0 : p_1 = p_2 ;$$

$$H_1 : p_1 > p_2.$$

Các số liệu cho ta

$$n_1 = 1000, k_1 = 300, f_1 = 0,3 ;$$

$$n_2 = 500, k_2 = 120, f_2 = 0,24.$$

$$\text{Tần suất chung là : } f = \frac{420}{1500} = 0,28.$$

$$T = \frac{0,3 - 0,24}{\sqrt{(0,28)(0,72)(0,003)}} = 2,4.$$

Với mức ý nghĩa  $\alpha = 2\%$  và bài toán kiểm định một phía ta có :

$$c = z_{0,02} = 2,06.$$

Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Tỉ lệ những người ưa thích Côca – Cola với công thức mới cao hơn tỉ lệ những người ưa thích Côca – Cola với công thức cũ".

**193.**  $T = 8$ . Bác bỏ  $H_0$ . Kết luận : "Thuốc chống hói đầu tỏ ra có tác dụng".

**194.** Các số liệu cho ta :

$$n_1 = 400, k_1 = 80, f_1 = \frac{80}{400} = 0,2 ;$$

$$n_2 = 500, k_2 = 125, f_2 = \frac{125}{500} = 0,25 ;$$

$$f = \frac{80 + 125}{400 + 500} = \frac{205}{900} = 0,227 ;$$

$$T = \frac{0,25 - 0,20}{\sqrt{(0,227)(0,773)\left(\frac{1}{400} + \frac{1}{500}\right)}} = 1,79.$$

Hàng số  $c = 1,96$ .

Vì  $|T| < c$  nên ta chấp nhận  $H_0$ . Kết luận : "Tỉ lệ thi trượt của hai nhóm thí sinh là như nhau".

**195.**  $n_1 = 100, k_1 = 40, f_1 = 0,4 ;$

$$n_2 = 150 ; k_2 = 68 ; f_2 = 0,453 ;$$

$$f = \frac{40 + 68}{100 + 150} = \frac{108}{250} = 0,432 ;$$

$$T = \frac{0,4 - 0,453}{\sqrt{(0,432)(0,568)\left(\frac{1}{100} + \frac{1}{150}\right)}} = \frac{-0,053}{0,064} = -0,828.$$

Hàng số  $c = 1,96$ .

Vì  $|T| < c$  nên ta chấp nhận  $H_0$ .

$$196. \quad n_1 = 600, k_1 = 90, f_1 = 0,15 ;$$

$$n_2 = 400, k_2 = 48, f_2 = 0,12 ;$$

$$f = \frac{90 + 48}{600 + 400} = 0,13 ;$$

$$T = \frac{0,15 - 0,12}{\sqrt{(0,13)(0,87)\left(\frac{1}{600} + \frac{1}{400}\right)}} = 1,36$$

Với  $\alpha = 2\%$  thì  $c = 2,33$ . Ta có  $T < c$ , thành thử giả thiết "tỉ lệ bỏ học như nhau" được chấp nhận.

$$197. \quad n_1 = 1656, k_1 = 729, f_1 = 0,44 ;$$

$$n_2 = 1432, k_2 = 301, f_2 = 0,21 ;$$

$$f = \frac{301 + 729}{1656 + 1432} = 0,33 ;$$

$$T = \frac{0,44 - 0,21}{\sqrt{(0,333)(0,67)\left(\frac{1}{1656} + \frac{1}{1432}\right)}} = \frac{0,23}{0,017} = 13,6.$$

Với  $\alpha = 1\%$  thì  $c = 2,56$ .

Ta có  $T < c$  thành thử  $H_0$  bị bác bỏ. Kết luận : "Có sự thay đổi trong quan điểm của công chúng".

$$198. \quad T = \frac{\frac{8}{37} - \frac{7}{38}}{\sqrt{(0,2)(0,8)\left(\frac{1}{37} + \frac{1}{38}\right)}} = \frac{0,032}{0,092} = 0,35.$$

Giả thiết  $H_0$  được chấp nhận.

## Chương 6

### BÀI TOÁN SO SÁNH MỞ RỘNG

$$\begin{aligned} 199. T &= \frac{600^2}{(478)(122)} \left[ \frac{79^2}{100} + \frac{82^2}{100} + \frac{77^2}{100} + \frac{83^2}{100} + \frac{76^2}{100} + \frac{81^2}{100} \right] \\ &\quad - \frac{(600)(478)}{122} = 2353,24 - 2350,81 = 2,42. \end{aligned}$$

Với mức ý nghĩa  $\alpha = 5\%$ , tra bảng phân bố  $\chi^2$  với 5 bậc tự do, ta tìm được  $c = \chi_{0,05}^2 = 11,07$ .

Vì  $T < c$  nên ta chấp nhận  $H_0$ .

200.

$$\begin{aligned} T &= \frac{(650)^2}{(410)(240)} \left[ \frac{60^2}{100} + \frac{75^2}{150} + \frac{150^2}{200} + \frac{125^2}{200} \right] - \frac{(650)(410)}{240} \\ &= 1134,07 - 1110,41 = 23,63. \end{aligned}$$

Với mức ý nghĩa  $\alpha = 1\%$ , tra bảng  $\chi^2$  với  $4 - 1 = 3$  bậc tự do, ta tìm được  $c = \chi_{0,01}^2 = 11,343$ . Vì  $T > c$  nên ta bác bỏ  $H_0$  và kết luận : "Kết quả thi của học sinh các thầy A, B, C, D là khác nhau".

201.  $T = 1,08$ . Tra bảng phân bố  $\chi^2$  với bậc tự do  $3 - 1 = 2$  ở mức  $\alpha = 5\%$ , ta tìm được  $c = 5,991$ . Vì  $T < c$  nên ta chấp nhận  $H_0$ .

Kết luận : "Không có sự khác nhau giữa đàn ông và đàn bà trong thị hiếu tiêu dùng xe ô tô của công ty".



$$202. T = \frac{650^2}{(250)(400)} \left[ \frac{25^2}{100} + \frac{100^2}{220} + \frac{125^2}{330} \right] - \frac{(650)(250)}{400} =$$

$$= 418,49 - 406,25 = 12,24.$$

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với  $3 - 1 = 2$  bậc tự do, ta tìm được  $c = 9,21$ . Vì  $T > c$  nên ta bác bỏ  $H_0$ . Kết luận : "Có sự khác nhau giữa hai nhóm phụ nữ và đàn ông trong việc sử dụng phương tiện giao thông đi làm hàng ngày".

$$202. T = 85,838.$$

Tra bảng phân bố  $\chi^2$  với 3 bậc tự do ở mức  $\alpha = 1\%$ , ta tìm được  $c = 11,343$ .

Vì  $T > c$  nên ta bác bỏ  $H_0$ .

Kết luận : "Thị hiếu về màu sắc áo khoác của nam và nữ là khác nhau".

203. a)  $T = 5,68$ . Tra bảng phân bố  $\chi^2$  với 3 bậc tự do ở mức ý nghĩa 1%, ta tìm được  $c = 11,343$ .

Vì  $T > c$  nên ta bác bỏ  $H_0$ .

b)  $T = 5,07$ . Tra bảng phân bố  $\chi^2$  với 2 bậc tự do ở mức ý nghĩa 5% ta tìm được  $c = 5,99$ .

Vì  $T < c$  nên ta chấp nhận  $H_0$ . Như vậy tỉ lệ chữa khỏi khối u của 3 nhóm 2, 3, 4 là như nhau. Sự khác nhau của 4 nhóm 1, 2, 3, 4 là do sự có mặt của nhóm 1 trong khi so sánh.

204. Ở đây ta có 12 tập hợp chính và ta cần so sánh phân bố sinh con trai con gái của 12 tập hợp chính đó.

$$T = 14,986.$$

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với 10 bậc tự do, ta tìm được  $c = 18,3$ .

Vì  $T < c$  nên ta chấp nhận giả thiết đã nêu.

**205.**  $T = 840,62$ .

Với mức ý nghĩa 1% tra bảng phân bố  $\chi^2$  với 5 bậc tự do, ta tìm được  $c = 11,07$ .

Vì  $T > c$  nên ta bác bỏ  $H_0$ .

Kết luận : "Phân bố thu nhập của hai nhóm tuổi là khác nhau".

**206.** Các tần số lí thuyết như sau :

Màu tóc	Đàn ông	Đàn bà
Đen	40,9	47,1
Vàng nhạt	47,8	55,2
Nâu	80,8	93,2
Đỏ	26,5	30,5

$$\begin{aligned} \text{b) } T &= \frac{(56 - 40,9)^2}{40,9} + \frac{(32 - 47,1)^2}{47,1} + \dots + \\ &+ \dots + \frac{(19 - 26,5)^2}{26,5} + \frac{(38 - 30,5)^2}{30,5} = 19,17 \end{aligned}$$

Với mức ý nghĩa 1% tra bảng phân bố  $\chi^2$  với 3 bậc tự do, ta tìm được  $c = 11,345$ . Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$  và kết luận : "Số liệu cho thấy có sự khác nhau về phân bố màu tóc giữa hai giới tính nam và nữ".

**207.**  $T = 4,27$ .

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với 5 bậc tự do, ta tìm được  $c = 11,07$ .

Vì  $T < c$  nên ta chấp nhận  $H_0$  và kết luận : "Phân bố thu nhập trong hai nhóm tuổi không khác nhau".

**208.** Ta tính tổng các dòng và các cột :

				Tổng số
	220	121	63	404
	130	207	75	412
	84	54	24	162
	156	95	43	294
	122	164	73	359
Tổng số	772	641	278	1631

$$T = 1631 \left[ \frac{220^2}{(772)(404)} + \frac{121^2}{(641)(404)} + \frac{63^2}{(278)(404)} + \dots + \frac{73^2}{(278)(359)} - 1 \right] \\ \approx 75,72.$$

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với  $(5 - 1)(3 - 1) = 8$  bậc tự do, ta tìm được  $c = 15,07$ .

Vì  $T > c$  nên ta bác bỏ  $H_0$  và kết luận : "Phân bố về mức độ thỏa mãn cuộc sống của 5 thành phố là khác nhau".

**209.** Ta có

$$T = 643 \left[ \frac{49^2}{(169)(206)} + \frac{31^2}{(100)(206)} + \frac{126^2}{(374)(206)} + \dots + \frac{109^2}{(374)(173)} - 1 \right] \\ = 10,4$$

Tra bảng phân bố  $\chi^2$  với  $(3 - 1)(4 - 1) = 6$  bậc tự do, ta tìm được  $c = \chi_{0,05}^2 = 12,595$ . Giá trị này lớn hơn  $T$ . Vậy ta chấp nhận  $H_0$  và kết luận : "Số băng trên vỏ có phân bố như nhau đối với cả 3 loại ốc sên rừng này".

**210.** Tần suất lí thuyết của ô "trí thức không ý kiến" là  $\frac{(60)(105)}{1290} = 4,88$  bé hơn 5 do đó điều kiện cho phép áp dụng tiêu chuẩn "khi bình phương" không được thỏa mãn.

Để khắc phục khó khăn này, một cách hợp lí là ghép cột cuối cùng "trí thức" với cột "công nhân : vì trí thức có lẽ gần với công nhân hơn là nông dân. Ta có bảng số liệu dạng sau đây :

Ý kiến \ Tầng lớp	Công nhân, trí thức	Nông dân	Tổng số
Tăng	120	300	420
Như cũ	230	400	630
Giảm	55	80	135
Không ý kiến	35	70	105
Tổng số	440	850	1290

Sử dụng công thức ta tìm được :

$$T \approx 1290 \left[ \frac{120^2}{(440)(420)} + \frac{300^2}{(850)(420)} + \dots + \frac{35^2}{(440)(105)} + \frac{70^2}{(850)(105)} - 1 \right]$$

$$= 10,059.$$

Tra bảng phân bố  $\chi^2$  ở mức 5% với bậc tự do là

$$(2 - 1)(4 - 1) = 3, \text{ ta tìm được } \chi_{0,05}^2 = 7,815.$$

Số này bé hơn 10,059. Vậy ta kết luận : "Có sự khác nhau về ý kiến giữa hai nhóm : Nông dân và Công nhân, Trí thức (sống ở thành thị)".

**211. T = 1075.**

Tra bảng phân bố  $\chi^2$  với bậc  $(4 - 1)(3 - 2) = 6$  ở mức 1% ta tìm được  $\chi_{0,01}^2 = 16,812$ .

Vì  $T > 16,812$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Có sự khác nhau về tình hình đau mắt hột ở ba tỉnh A, B, C".

**212. a)** Tần số lí thuyết của bảng số liệu trên như sau :

Tên nước	Cơ quan bị ung thư			
	Ruột	Ngực	Dạ dày	Các bộ phận khác
A	9,19	10,91	12,25	37,63
B	7,88	9,38	10,50	32,25
C	3,94	4,69	5,25	16,13

b) Chưa áp dụng được vì có ô tần suất lí thuyết là  $3,94 < 5$ .

c) Ta cần ghép hai cột đầu tiên lại.

Khi đó  $T = 29,07$ .

Tra bảng phân bố  $\chi^2$  với bậc tự do là  $(3 - 1)(3 - 1) = 4$ , và mức ý nghĩa 1%, ta tìm được  $\chi_{0,01}^2 = 13,277$ .

Vì  $T > 13,277$  nên ta bác bỏ  $H_0$ . Kết luận : "Phân bố về bệnh ung thư dẫn đến tử vong là khác nhau giữa ba nước".

**213. a)** Các tần số lí thuyết như sau :

Vắc xin	Mức độ phản ứng		
	Nhẹ	Trung bình	Mạnh
A	20,9	148,6	8,1
B	20,1	143,0	7,9

b)  $T = 8,81$

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với  $(3 - 1)(2 - 1) = 2$  bậc tự do, ta tìm được :  $\chi_{0,05}^2 = 5,991$ . Vì  $T > c$  nên ta bác bỏ  $H_0$  và kết luận : "Hai loại vắc xin A và B khác nhau".

**214.**  $T = 11,389$ .

Với mức ý nghĩa 0,025 tra bảng phân bố  $\chi^2$  với  $(3 - 1)(3 - 1) = 4$  bậc tự do, ta tìm được  $c = \chi_{0,025}^2 = 11,143$

Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Ba giống A, B, C không như nhau về phương diện cung cấp sữa".

**215.** Kết quả tính toán cho ta bảng ANOVA như sau

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	354,67	2	177,34	4,96
Sai số	322	9	35,78	
Tổng	676,67	11		

Với mức ý nghĩa  $\alpha = 5\%$  tra bảng phân bố Fisher với bậc tự do (2,9), ta tìm được hằng số

$$c = f_{0,05} = 4,26$$

Vì  $F > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Điểm thi trung bình của các sinh viên theo học các giáo sư A, B, C là khác nhau ở mức ý nghĩa 5%".

**216.** Kết quả tính toán cho ta bảng ANOVA dưới đây

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	310,6	3	103,5	1,85
Sai số	1007,2	18	56	
Tổng	1317,8	21		

Tra bảng phân bố Fisher với bậc tự do (3,18) và mức ý nghĩa 5%, ta tìm được  $c = f_{0,05} = 3,16$ .

Vì  $F < c$  nên ta chấp nhận giả thiết  $H_0$ . Kết luận : "Năng suất trung bình của bốn giống lúa A, B, C, D như nhau".

**217.** Tổng BP do nhân tố :

$$\frac{61^2}{6} + \frac{126^2}{8} + \frac{25^2}{7} - \frac{(61 + 126 + 25)^2}{21} = 553,76$$

Tổng BP do sai số :

$$2854 - \left( \frac{61^2}{6} + \frac{126^2}{8} + \frac{25^2}{7} \right) = 160,05$$

Từ đó bảng ANOVA thu được là

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	553,76	2	276,88	31,14
Sai số	160,05	18	8,89	
Tổng	713,81	20		

Tra bảng phân bố Fisher với bậc tự do (2,18) và mức ý nghĩa 1%, ta tìm được  $c = f_{0,01} = 10,39$ .

Vì  $F > c$  nên ta bác bỏ  $H_0$  và kết luận : "Ba loại thuốc đó khác nhau".

**218.** Kết quả tính toán cho ta bảng ANOVA sau đây.

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	238.59	2	119.29	3,79
Sai số	4378,05	139	31,5	
Tổng	4616,64	141		

Tra bảng phân bố Fisher với bậc tự do (2,139) với mức ý nghĩa 5% ta tìm được

$$c = f_{0,05} = 3$$

Giả thiết  $H_0$  bị bác bỏ và ta kết luận : "Kích thước trung bình các hộp sọ của 3 nhóm người này khác nhau".

**219.** Kết quả tính toán cho ta bảng ANOVA sau đây :

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	44,361	3	14,787	2,15
Sai số	151,351	22	6,88	
Tổng	195,712	25		

Tra bảng phân bố Fisher với bậc tự do (3,22) ta tìm được

$$c = f_{0,05} = 3,05$$

Vì  $F < c$  nên ta chưa có cơ sở bác bỏ  $H_0$ . Kết luận : "Chất lượng bóng đèn của 4 nhà máy A, B, C, D như nhau".

**220.** Bảng ANOVA được cho như sau :

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	1,552	3	0,517	3,31
Sai số	5,312	34	0,156	
Tổng	6,864	37		

Với mức ý nghĩa 5% tra bảng phân bố Fisher với bậc tự do (3,34), ta tìm được  $c = f_{0,05} = 2,85$ .

Vì  $f > c$  nên ta bác bỏ  $H_0$  và kết luận : "Khả năng kháng virus của chuột phụ thuộc vào thời điểm bị tiêm".

**221.** Bảng ANOVA như sau :

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	234,93	3	78,31	10,68
Sai số	256,66	35	7,33	
Tổng	491,59	38		



**222.**

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	117,04	3	39,01	1,03
Sai số	493,80	13	37,985	
Tổng	610,84	16		

**223.**

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	260,55	3	86,85	2,33
Sai số	595,20	16	37,20	
Tổng	855,75	19		

Với mức ý nghĩa 5% tra bảng phân bố Fisher với bậc tự do (3,16), ta tìm được  $c = f_{0,05} = 3,24$ .

Ta có  $F < c$  thành thử giả thiết  $H_0$  được chấp nhận : "Thời gian báo cáo trung bình của bốn báo cáo viên A, B, C, D là như nhau".

**224.**

Nguồn	Tổng BP	Bậc tự do	Trung bình BP	Tỉ số F
Nhân tố	410,17	3	136,72	5,52
Sai số	396,38	16	24,77	
Tổng	806,55	19		

## Chương 7

### PHÂN TÍCH TƯƠNG QUAN VÀ HỒI QUY

225.  $\sum x = 176$  ;  
 $\sum y = 902$  ;  
 $\sum xy = 29714$  ;  $\sum x^2 = 6948$  ;  $\sum y^2 = 168750$  ;  
 $n\sum xy - (\sum x)(\sum y) = 10182$  ;  
 $n\sum x^2 - (\sum x)^2 = 3764$  ;  
 $n\sum y^2 - (\sum y)^2 = 30146$ .

Từ đó  $r = - \frac{12882}{\sqrt{3764}\sqrt{30146}} = -0,956$ .

226. Ta có  $\sum x = 850$  ;  $\sum y = 29$  ;  
 $\sum xy = 2486,3$  ;  $\sum x^2 = 72617$  ;  
 $\sum y^2 = 85,78$ .

Từ đó

$$\begin{aligned} n\sum xy - (\sum x)(\sum y) &= 213 ; \\ n\sum x^2 - (\sum x)^2 &= 3670 ; \\ n\sum y^2 - (\sum y)^2 &= 16,8 ; \\ r &= \frac{213}{\sqrt{3670}\sqrt{16,8}} = 0,858. \end{aligned}$$

227. a)  $r = 0,982$ .

b)  $r = -0,951$ .

228.  $r = 0,985$ .

229.  $r = 0,102$ .

$$T = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{(0,102)\sqrt{4}}{\sqrt{1-0,102^2}} = \frac{0,204}{0,995} = 0,205.$$

Tra bảng phân bố Student với bậc tự do  $6 - 2 = 4$ , ta tìm được

$$c = t_{0,025} = 2,776.$$

Vì  $|T| < c$  nên ta chấp nhận  $H_0$  và kết luận : "Không có mối tương quan tuyến tính giữa X và Y".

230.  $r = 0,951$ .

231. Trước hết ta cần trình bày các số liệu trên dưới dạng bảng tương quan sau đây :

$\begin{array}{c} \backslash \\ X \\ / \\ Y \end{array}$	8	12	20	24	
	82	65	60	52	
	78	50	47	41	
	87	62	44	57	
	58	55	66	50	
	70	52	41	47	
	65	49	57	63	
$n_i$	6	6	6	6	$n = 24$
$T_i$	440	333	315	310	$T = 1398$

Ta có :  $\sum x = 6(8) + 6(12) + 6(20) + 6(24) = 384$  ;

$$\sum y = T = 1398 ;$$

$$\sum x^2 = 6(8)^2 + 6(12)^2 + 6(20)^2 + 6(24)^2 = 7140 ;$$

$$\sum y^2 = 82^2 + 78^2 + \dots + 47^2 + 63^2 = 84908 ;$$

$$\sum xy = 8(440) + 12(333) + 20(315) + 24(310) = 21256$$

Vậy  $n\sum xy - (\sum x)(\sum y) = -26688$  ;

$$\sqrt{n\sum x^2 - (\sum x)^2} = 151,789 ;$$

$$\sqrt{n\sum y^2 - (\sum y)^2} = 288,77$$

$$\text{Hệ số tương quan } r = \frac{-26688}{(151,789)(288,77)} = -0,6089$$

Hệ số xác định  $r^2 = 0,37$ .

Để tính tỉ số tương quan ta có :

$$SST = \sum y^2 - \frac{T^2}{n} = 3474,5 ;$$

$$SSF = \sum \frac{T_i^2}{n_i} - \frac{T^2}{n} = 1868,83.$$

Từ đó tỉ số tương quan là

$$\eta^2 = \frac{SSF}{SST} = 0,5378$$

$$\begin{aligned} 232. \text{ Ta có } F &= \frac{(\widehat{\eta^2} - r^2)(n - k)}{(1 - \widehat{\eta^2})(k - 2)} = \\ &= \frac{(0,5378 - 0,37)(24 - 4)}{(1 - 0,5378)(4 - 2)} = 3,63. \end{aligned}$$

Tra bảng phân bố Fisher với bậc tự do (2,20) ở mức 5%, ta tìm được  $c = f_{0,05} = 3,49$ .

Vì  $F > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Có tồn tại mối tương quan phi tuyến của Y đối với X".

**233. a)** Sản lượng trung bình của mức phân X = x kí hiệu là  $\bar{y}_X$ .  
Ta có kết quả sau :

Với  $X = 1 \rightarrow \bar{y}_X = 14$  ;

$X = 2 \rightarrow \bar{y}_X = 14,6$  ;

$X = 3 \rightarrow \bar{y}_X = 15,8$  ;

$X = 4 \rightarrow \bar{y}_X = 16,57$  ;

$X = 5 \rightarrow \bar{y}_X = 17,57$ .

Sản lượng trung bình chung của 100 thửa là  $y = 15,86$ .

b) Ta lập bảng tương quan

$\begin{matrix} Y \\ \backslash \\ X \end{matrix}$	1	2	3	4	5
14	10	8			
15		12	7		
16			28	6	
17				8	9
18					12
$n_i$	10	20	35	14	21
$T_i$	140	292	553	232	369

$T = \sum T_i = 1586$  ;

$\sum x = 316$  ;

$\sum y = \sum T_i = 1586$  ;

$$\sum x^2 = 1154 ;$$

$$\sum y^2 = 25308 ;$$

$$\sum xy = 5156 ;$$

$$n\sum xy - (\sum x)(\sum y) = 14424 ;$$

$$\sqrt{n\sum x^2 - (\sum x)^2} = 124,67 ;$$

$$\sqrt{n\sum y^2 - (\sum y)^2} = 124,11.$$

Từ đó hệ số tương quan :

$$r = \frac{14424}{(124,67)(124,11)} = 0,932$$

Hệ số xác định  $r^2 = 0,868$ .

Để tính tỉ số tương quan ta có

$$SST = \sum y^2 + \frac{T^2}{n} = 25308 - \frac{1586^2}{100} = 154,04 ;$$

$$SSF = \sum \frac{T_i^2}{n_i} - \frac{T^2}{n} = 135,069.$$

Từ đó tỉ số tương quan :

$$\widehat{\eta^2} = \frac{SSF}{SST} = \frac{135,069}{154,04} = 0,876$$

c) Phương trình hồi quy tuyến tính của Y theo X là

$$Y = 0,93x + 12,93.$$

**234.**  $T = 1,85$ .

Tra bảng phân bố Student với bậc tự do  $20 - 2 = 18$ , ta tìm được :

$$c = t_{0,005} = 2,861.$$

Vì  $T < c$  nên ta chấp nhận giả thiết  $H_0$  : "Không có sự tương quan tuyến tính giữa điểm thi tốt nghiệp của một sinh viên với mức thu nhập gia đình của họ".

**235.**  $T = 0,91$ .

Tra bảng phân bố Student với bậc tự do là  $20 - 2 = 18$ , mức ý nghĩa 5%, ta tìm được

$$c = t_{0,025} = 2,101.$$

Vì  $T < c$  nên ta chấp nhận giả thiết. Kết luận : "Số lần sút bóng và số bàn thắng không tương quan".

**236.**  $r = 0,222$  ;

$$T = 0,755.$$

Tra bảng phân bố Student với bậc tự do  $13 - 2 = 11$ , mức ý nghĩa 5%, ta tìm được  $c = t_{0,025} = 2,201$ .

Ta có  $T < c$  nên giả thiết  $H_0$  được chấp nhận.

**237.**  $\widehat{\eta^2} = 0,88$  ;  
 $r^2 = 0,1$ .

X và Y có mối tương quan phi tuyến khá chặt chẽ, nhưng mối tương quan tuyến tính thì yếu.

**238.**  $\widehat{\eta^2} = 0,947$  ;  
 $r^2 = 0,94$ .

X và Y có tương quan tuyến tính chặt chẽ.

**239.**  $\widehat{\eta^2} = 0,947$  ;  
 $r^2 = 0,08$ .

X và Y có tương quan phi tuyến chặt chẽ, nhưng không có tương quan tuyến tính.

$$240. \text{ Ta có } T = \frac{560 \begin{vmatrix} 328 & 122 \\ 77 & 33 \end{vmatrix}}{(450)(110)(405)(155)} = 0,368$$

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với một bậc tự do, ta tìm được  $c = \chi_{0.05}^2 = 3,841$ .

$T < c$  thành thử ta chấp nhận giả thiết  $H_0$  và kết luận : "Hai đặc tính về hoa và lá đã nêu là độc lập".

241. Sử dụng test  $\chi^2$  ta có :

$$T = 169 \left\{ \frac{35^2}{(49)(115)} + \frac{19^2}{(33)(115)} + \frac{36^2}{(52)(115)} + \frac{25^2}{(35)(115)} + \right. \\ \left. + \frac{14^2}{(49)(54)} + \frac{14^2}{(33)(54)} + \frac{16^2}{(52)(54)} + \frac{10^2}{(35)(54)} - 1 \right\} \\ = 2,13.$$

Với mức ý nghĩa  $\alpha = 5\%$  tra bảng phân bố  $\chi^2$  với

$(2 - 1)(4 - 1) = 3$  bậc tự do, ta tìm được  $c = \chi_{0.05}^2 = 7,81$ .

Ta thấy  $T < c$  thành thử giả thiết  $H_0$  được chấp nhận.

*Kết luận* : "Màu vỏ và số lượng vạch trên vỏ của một con ốc sên rừng không có mối tương quan với nhau".

242. Sử dụng tiêu chuẩn  $\chi^2$  ta có :

$$T = 200 \left\{ \frac{10^2}{(16)(46)} + \frac{20^2}{(68)(46)} + \frac{11^2}{(61)(46)} + \frac{5^2}{(55)(46)} + \right. \\ \left. + \frac{6^2}{(16)(154)} + \frac{48^2}{(68)(154)} + \frac{50^2}{(61)(154)} + \frac{50^2}{(55)(154)} - 1 \right\} \\ = 22,53.$$



Với mức ý nghĩa 1% tra bảng phân bố  $\chi^2$  với bậc tự do  $(2 - 1)(4 - 1) = 3$ , ta tìm được

$$c = \chi_{0,01}^2 = 11,345$$

Vì  $T > c$  nên ta bác bỏ giả thiết  $H_0$  và đi đến kết luận : "Giữa huyết áp và trọng lượng của trẻ 14 tuổi có sự phụ thuộc lẫn nhau".

**243.**  $T = 4,03$ .

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với  $(2 - 1)(3 - 1) = 2$  bậc tự do, ta tìm được

$$c = \chi_{0,05}^2 = 5,99.$$

Vì  $T < c$  nên ta chấp nhận  $H_0$  và kết luận : "Tuổi và loại hình phạm tội của một phạm nhân độc lập với nhau".

**244.**  $T = 5,308$ .

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với  $(2 - 1)(3 - 1) = 2$  bậc tự do, ta tìm được

$$c = \chi_{0,05}^2 = 55,99$$

Vì  $T < c$  nên ta chấp nhận  $H_0$  và kết luận : "Không có mối tương quan giữa việc đi nhà thờ và vấn đề thành kiến chủng tộc".

**245.**  $T = 55,5$ .

Với mức ý nghĩa 5% tra bảng phân bố  $\chi^2$  với  $(4 - 1)(3 - 1) = 6$  bậc tự do, ta tìm được

$$c = \chi_{0,05}^2 = 12,592$$

Vì  $T > c$  nên ta bác bỏ giả thiết  $H_0$  và kết luận : "Có mối liên hệ giữa nghề nghiệp của một người với quan niệm của anh ta về tiêu chuẩn đạo đức và tính nghề nghiệp".

**246.**  $T = 19,14$ .

Với mức ý nghĩa 1% tra bảng phân bố  $\chi^2$  với  $(4 - 1)(2 - 1) = 3$  bậc tự do, ta tìm được

$$c = \chi_{0,01}^2 = 11,343.$$

Vì  $T > c$  nên ta quyết định bác bỏ  $H_0$ . Kết luận : "Có tồn tại mối quan hệ phụ thuộc giữa giới tính và màu tóc".

**247.**  $T = 6,65$ .

Với mức ý nghĩa 1% tra bảng phân bố  $\chi^2$  với  $(3 - 1)(3 - 1) = 4$  bậc tự do, ta tìm được

$$c = \chi_{0,01}^2 = 13,277.$$

Vì  $T < c$  nên ta chấp nhận  $H_0$ . Kết luận : "Giữa tuổi và mức thu nhập của các nhà doanh nghiệp Mỹ không có mối tương quan".

**248. a)** Ta có  $\sum xy = 138800$  ;

$$\sum x = 3000$$
 ;

$$\sum y = 276$$
 ;

$$\sum x^2 = 1540000$$
 ;

$$\sum y^2 = 12718.$$

$$\begin{aligned} \text{Từ đó } a &= \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} \\ &= \frac{6(138800) - (3000)(276)}{6(1540000) - (3000)^2} = \frac{4800}{240000} = 0,02 \\ b &= \frac{\sum y - a \sum x}{n} = \frac{276 - (0,02)(3000)}{6} = 36. \end{aligned}$$

Vậy đường thẳng hồi quy là

$$y = 0,02x + 36$$

b) Ta có  $s_{Y.X}^2 = \frac{\sum y^2 - a \sum xy - b \sum y}{n - 2} = 1,5.$

Vậy  $s_{Y.X} = \sqrt{1,5} = 1,22.$

**249.** a)  $r = 0,84.$

b)  $\hat{y} = 0,75x + 0,61.$

c)  $\hat{y} = 3,09$

**250.** a)  $y = -0,934x + 5,162.$

b)  $s_{Y.X} = 0,367.$

c)  $|T| = 10,26.$

Tra bảng phân bố Student với bậc tự do  $15 - 2 = 13$ , ta tìm được phân vị  $c = t_{0,025} = 2,16$ . Ta có  $|T| > c$  thành thử giả thiết : "Chiều cao thân răng và tuổi độc lập với nhau" bị bác bỏ.

**251.** a)  $y = 1,677x + 1,045.$

b)  $s_{Y.X} = 2,22.$

c)  $F = 23,26.$

Tra bảng phân bố Fisher với bậc tự do (1,18) ở mức ý nghĩa 1%, ta tìm được  $c = 8,29$ . Vì  $F > c$  nên ta đi đến kết luận : "Có tồn tại mối tương quan tuyến tính giữa Y và X".

**252.** a)  $y = 1,548x - 1,739.$

b)  $s_{Y.X} = 0,29.$

c)  $F = 12,55.$

Tra bảng phân bố Fisher với bậc tự do (1,9) ở mức ý nghĩa 1%, ta tìm được  $c = 10,6$ .

Vì  $F > c$  nên ta kết luận : "Có mối tương quan tuyến tính giữa Y và X".

253. a)  $r = -0,989$ .

b)  $y = -0,4x + 1020,2$ .

c)  $s_{Y.X} = 1,31$ .

$|T| = 40$  lớn hơn hàng số  $c = 2,045$ .

Vậy ta kết luận : "Có mối tương quan giữa X và Y".

d) 3 phút 20,22 giây.

e) Năm 2025.

f) Kỉ lục thế giới năm 2500 là 20,22 giây!! Kết quả này là phi lí. Điều đó cho thấy sử dụng phương trình hồi quy để dự báo chỉ cho kết quả hợp lí trong một lân cận của năm 2000.

254.  $\sum x = 210$  ;

$\sum y = 119$  ;

$\sum xy = 5160$  ;

$\sum x^2 = 9850$ .

$$a = \frac{5(5160) - (210)(119)}{5(9850) - (210)^2} = 0,16 ;$$

$$b = \frac{119 - 0,16(210)}{5} = 17,08.$$

Vậy phương trình đường thẳng hồi quy là

$$y = 0,16x + 17,08.$$

b) Như vậy cứ tăng thêm 10 tuổi thì thời gian phân hủy thuốc sẽ tăng thêm 1,6 phút. Ngoài ra thời gian phân hủy thuốc ít nhất là 17,08 phút.

c)  $y = 0,16(42) + 17,08 = 23,8$ .

Với bệnh nhân 42 tuổi thì thời gian phân hủy thuốc trung bình là 23,8 phút.

255. a)  $y = 1,04x - 25,7$

b)  $\sum y^2 = 13213$  ;

$\sum xy = 18992$  ;

$\sum y = 259$ .

Vậy  $s_{Y.X} = \sqrt{\frac{13213 - (-25,7)(259) - (1,04)(18992)}{6 - 2}} = 5,42$ .

c) Giá trị trung bình được dự báo là

$$1,04(80) - 25,7 = 57,5.$$

d) Khoảng tin cậy 95% cho giá trị trung bình của Y khi X = 80

là :  $57,5 \pm 2,776(3,6) \sqrt{\frac{1}{6} + \frac{(80 - 66)^2}{27,954 - \frac{(396)^2}{6}}} = 57,5 \pm 5,24$ .

e) Khoảng tin cậy 95% cho giá trị của Y khi X = 80 là

$$57,5 \pm 2,776(3,6) \sqrt{1 + \frac{1}{6} + \frac{(80 - 66)^2}{27,954 - \frac{(396)^2}{6}}} = 57,5 \pm 11,28.$$

256.  $\sum x = 90$  ;

$\sum y = 141$  ;

$\sum xy = 1799$  ;

$\sum x^2 = 1192$  ;

$\sum y^2 = 2855$  ;

$$a = \frac{7(1799) - (90)(141)}{7(1192) - 90^2} = -\frac{97}{244} = -0,4 ;$$

$$b = \frac{141 - (-0,4)(90)}{7} = 25,29.$$

Vậy phương trình của đường hồi quy là

$$y = -0,4x + 25,29.$$

$$b) \quad s_{Y.X} = \sqrt{\frac{2855 - 25,29(141) - (-0,4)(1799)}{7 - 2}} = 1,32.$$

$$d) \quad \hat{y} = -0,4(14) + 25,29 = 19,69.$$

Khoảng tin cậy 95% cho giá trị dự báo là

$$19,69 \pm (2571)(132) \sqrt{1 + \frac{1}{7} + \frac{(14 - 12,86)^2}{1192 - \frac{90^2}{7}}}$$

$$= 19,69 \pm 3,69$$

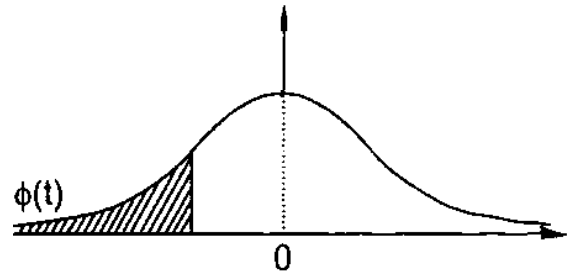
tức là khoảng từ 16 tuổi đến 23,38 tuổi.

## PHỤ LỤC

*Bảng 1. Hàm phân số chuẩn*

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}x^2} dx$$

(t từ -3,9 đến 0)



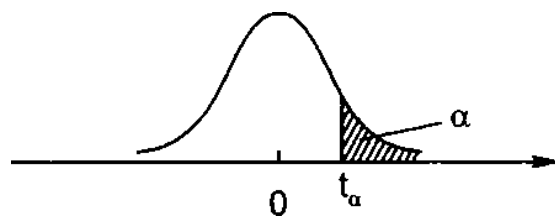
t	0	1	2	3	4	5	6	7	8	9
-0,0	0.5000	5960	4920	4880	4840	4801	4761	4721	4681	4641
1	4602	4562	4522	4483	4443	4404	4364	4325	4286	4247
2	4207	4168	4129	4090	4052	4013	3974	3936	3897	3859
3	3821	3783	3745	3707	3669	3632	3594	3557	3520	3483
4	3446	3409	3372	3336	3300	3264	3228	3192	3156	3121
-0,5	0.3085	3050	3015	3981	2946	2912	2877	2843	2810	2776
6	2743	2709	2676	2643	2611	2578	2546	2515	2483	2451
7	2420	2389	2358	2327	2297	2266	2236	2206	2177	2148
8	2119	2090	2061	2033	2005	1977	1949	1922	1894	1867
9	1841	1814	1788	1762	1736	1711	1685	1660	1635	1611
-1,0	0.1578	1562	1539	1515	1492	1469	1446	1423	1401	1379
1	1357	1335	1314	1292	1291	1251	1230	1210	1190	1170
2	1151	1131	1112	1093	1075	1056	1038	1020	1003	0985
3	0968	0951	0934	0918	0901	0885	0869	0853	0838	0823
4	0808	0793	0778	0764	0749	0735	0721	0708	0694	0681
-1,5	0.0668	0655	0643	0630	0618	0606	0594	0582	0571	0559
6	0548	0537	0526	0516	0505	0495	0485	0475	0465	0455
7	0446	0436	0427	0418	0409	0401	0392	0384	0375	0367
8	0359	0351	0344	0336	0329	0322	0314	0317	0301	0294
9	0288	0281	0274	0268	0262	0256	0250	0244	0239	0233
-2,0	0.0227	0222	0217	0212	2020	0202	0197	0192	0188	0183
1	0179	0174	0170	0166	0162	0158	0154	0150	0146	0143
2	0139	0136	0132	0129	0125	0122	0119	0116	0113	0110
3	0107	0104	0102	0099	0096	0094	0091	0089	0087	0084
4	0082	0080	0078	0075	0073	0071	0069	0068	0066	0064
-2,5	0.0062	0060	0059	0057	0055	0054	0052	0051	0049	0048
6	0047	0045	0044	0043	0041	0040	0039	0038	0037	0036
7	0035	0034	0033	0032	0031	0030	0029	0028	0027	0026
8	0026	0025	0024	0023	0023	0022	0021	0021	0020	0019
9	0019	0018	0018	0017	0016	0016	0015	0015	0014	0014
t	-3,0	-3,1	-3,2	-3,3	-3,4	-3,5	-3,6	-3,7	-3,8	-3,9
F(t)	0.0013	0010	0007	0005	0003	0002	0002	0001	0001	0000

Bảng 1. Hàm phân bố chuẩn  $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}x^2} dx$   
(t từ 0 đến + 3,9)

t	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0.5	0.6915	6950	6985	7019	7054	7088	7123	7156	7190	7224
6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
7	7580	7611	7642	7673	7703	7734	7764	7794	7823	7852
8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8132
9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1.0	0.8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1	8643	8665	8686	8708	8709	8649	8770	8790	8810	8830
2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015
3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177
4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319
1.5	0.9332	9345	9357	9370	9382	9394	9406	9418	9429	9441
6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545
7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633
8	9641	9649	9656	9664	9671	9678	9686	9683	9699	9706
9	9712	9719	9726	9732	9738	9744	9750	9756	9761	9767
2.0	0.9773	9778	9783	9788	9793	9798	9803	9808	9812	9817
1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857
2	9861	9864	9868	9871	9875	9878	9881	9884	9887	9890
3	9893	9896	9898	9901	9904	9906	9909	9911	9913	9916
4	9918	9920	9922	9925	9927	9929	9931	9932	9934	9936
2.5	0.9938	9940	9941	9943	9945	9946	9948	9949	9951	9952
6	9953	9955	9956	9957	9959	9960	9961	9962	9963	9964
7	9965	9969	9967	9968	9969	9970	9971	9972	9973	9974
8	9974	9975	9976	9977	9977	9978	9979	9979	9980	9981
9	9981	9982	9982	9983	9984	9984	9985	9985	9986	9986
t	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
F(t)	0.9987	9990	9993	9995	9996	9997	9998	9999	9999	9999

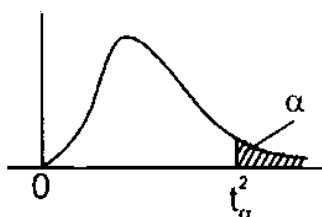


Bảng II. Phân vị  $t_{\alpha}$  của phân bố Student



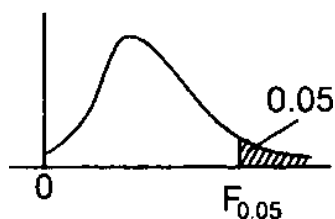
Bậc tự do	$t_{0,5}$	$t_{0,25}$	$t_{0,1}$	$t_{0,05}$	dl
1	6.314	12.706	31.821	63.657	1
2	2.920	4.303	6.965	9.925	2
3	2.353	3.182	4.541	5.841	3
4	2.132	2.776	3.747	4.604	4
5	2.015	2.571	3.365	4.032	5
6	1.943	2.447	3.143	3.707	6
7	1.895	2.365	2.998	3.499	7
8	1.860	2.306	2.896	3.355	8
9	1.833	2.262	2.821	3.250	9
10	1.812	2.228	2.764	3.169	10
11	1.796	2.201	2.718	3.106	11
12	1.782	2.179	2.681	3.055	12
13	1.771	2.160	2.650	3.012	13
14	1.761	2.145	2.624	2.977	14
15	1.753	2.131	2.606	2.947	15
16	1.746	2.120	2.583	2.921	16
17	1.740	2.110	2.567	2.898	17
18	1.734	2.101	2.552	2.878	18
19	1.729	2.093	2.539	2.861	19
20	1.725	2.086	2.528	2.845	20
21	1.721	2.080	2.518	2.831	21
22	1.717	2.074	2.508	2.819	22
23	1.714	2.069	2.500	2.807	23
24	1.711	2.064	2.492	2.797	24
25	1.708	2.060	2.485	2.787	25
26	1.796	2.056	2.479	2.779	26
27	1.703	2.052	2.473	2.771	27
28	1.701	2.048	2.467	2.763	28
29	1.699	2.045	2.462	2.756	29
inf	1.645	1.960	2.326	2.576	inf

Bảng III. Phân vị  $\chi^2_\alpha$  của phân bố "Khi bình phương"



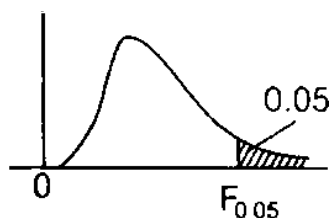
Bậc tự do	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$	df
1	.000	.000	.001	.004	3.841	5.024	6.635	7.879	1
2	.010	.020	.051	.103	5.991	7.378	9.210	10.597	2
3	.072	.115	.216	.352	7.815	9.348	11.345	12.838	3
4	.207	.297	.484	.711	9.488	11.143	13.277	14.860	4
5	.412	.554	.831	1.145	11.070	12.832	15.086	16.750	5
6	.676	.872	1.237	1.635	12.592	14.449	16.812	18.548	6
7	.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.982	22.362	24.736	27.688	28.819	13
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319	14
15	5.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	19
20	7.343	8.260	9.591	10.851	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	32.671	39.479	38.932	41.401	21
22	8.543	9.542	10.982	12.388	33.924	36.781	30.289	42.796	22
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558	24
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	25
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645	27
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	28
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336	29
30	13.787	14.93	16.791	18.493	43.773	46.979	50.892	53.672	30

Bảng IVa. Phân vị  $F_\alpha$  của phân bố Fisher mức  $\alpha = 5\%$



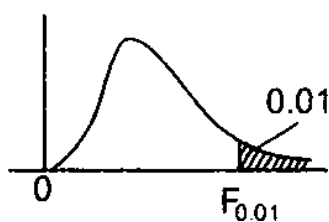
Bậc tự do của mẫu số	Bậc tự do của tử số									
	1	2	3	4	5	6	7	8	9	10
1	161	200	216	225	230	234	237	239	241	242
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.69	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91
$\infty$	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Bảng IVa (tiếp tục)



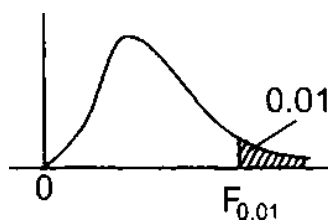
Bậc tự do của mẫu số	Bậc tự do của tử số								
	12	15	20	24	30	40	60	12	$\infty$
1	244	246	248	249	250	251	252	3.70	3.67
2	19.4	19.4	19.4	19.5	19.5	19.5	19.5	3.27	3.23
3	8.74	8.70	8.66	8.64	8.62	8.59	8.57	2.97	2.93
4	5.91	5.86	5.80	5.77	5.75	5.72	5.69	2.75	2.71
5	4.68	4.62	4.56	4.53	4.50	4.46	4.43	2.58	2.54
6	4.00	3.94	3.87	3.84	3.81	3.77	3.74	2.45	2.40
7	3.57	3.51	3.44	3.41	3.38	3.34	3.30	2.34	2.30
8	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.25	2.21
9	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.18	2.13
10	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.11	2.07
11	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.06	2.00
12	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.01	1.96
13	2.60	2.53	2.46	2.42	2.38	2.34	2.30	1.97	1.92
14	2.53	2.46	2.39	2.35	2.31	2.27	2.22	1.93	1.88
15	2.48	2.40	2.33	2.29	2.25	2.20	2.16	1.90	1.84
16	2.42	2.35	2.28	2.24	2.19	2.15	2.11	1.87	1.81
17	2.38	2.31	2.23	2.19	2.15	2.10	2.06	1.84	1.78
18	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.81	1.75
19	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.79	1.73
20	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.77	1.70
21	2.25	2.18	2.10	2.05	2.01	1.96	1.91	1.74	1.68
22	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.71	1.65
23	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.68	1.62
24	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.66	1.60
25	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.64	1.58
30	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.58	1.52
40	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.51	1.45
60	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.43	1.37
120	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.29
$\infty$	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Bảng IVb. Phân vị  $F_\alpha$  của phân bố Fisher mức  $\alpha = 1\%$



Bậc tự do của mẫu số	Bậc tự do của tử số									
	1	2	3	4	5	6	7	8	9	10
1	4.052	5.0000	5.403	5.625	5.764	5.859	5.928	5.982	6.023	6.056
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	10.6	8.02	6.99	6.42	5.64	5.80	5.61	5.47	5.35	5.26
10	10.0	7.56	6.55	5.99	5.32	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.06	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	4.86	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.70	4.62	4.44	4.30	4.19	4.10
14	8.86	6.51	5.56	5.04	4.56	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.44	4.32	4.14	4.00	3.89	3.80
16	8.53	6.23	5.29	4.77	4.34	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.19	4.67	4.25	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.17	4.01	3.84	3.71	3.60	3.51
19	8.19	5.93	5.01	4.50	4.10	3.94	3.77	3.63	3.52	3.43
20	8.10	5.85	4.94	4.43	4.04	3.87	3.70	3.56	3.46	3.37
21	8.02	5.78	4.87	4.37	3.99	3.81	3.64	3.51	3.40	3.31
22	7.95	5.72	4.82	4.31	3.94	3.76	3.59	3.45	3.35	3.26
23	7.88	5.66	4.76	4.26	3.90	3.71	3.54	3.41	3.30	3.21
24	7.82	5.61	4.72	4.22	3.86	3.67	3.50	3.36	3.26	3.17
25	7.77	5.57	4.68	4.18	3.70	3.63	3.46	3.32	3.22	3.13
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
$\infty$	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

Bảng IVb (tiếp tục)



Bậc tự do của mẫu số	Bậc tự do của tử số								
	12	15	20	24	30	40	60	120	$\infty$
1	6.106	6.157	6.209	6.235	6.261	6.287	6.313	6.339	6.366
2	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	14.4	14.2	14.2	13.9	13.8	13.7	13.7	13.6	13.5
5	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	4.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.34	2.10	2.03	1.95	1.86	1.76	1.66	1.53	1.38
$\infty$	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

## TÀI LIỆU THAM KHẢO

1. W. Mandenhall, *Mathematical Statistics with Applications*, Boston, 1990.
2. S.P. *Contemporary Statistics*, Mc Graw 1994.
3. D. A. Aczel *Statistics*, Irwin 1996.
4. W. Mandenhall and T.A. Sincich, *A Second Course in Statistics : Regression Analysis*, Prentice Hall 1996.
5. J.T. Nc Clave *Statistics*, Prentice Hall 1997.
6. R.F. Beal *Statistics for Economists. An Introduction*, Chicago, 1971
7. A. Agresti, *Statistical Methods for Social Sciences*, San Francisco 1986.
8. T.T. Soong, *Fundamentals of Probability and Statistics for engineers*, John Wiley 2004.
9. Đặng Hùng Thắng, *Mở đầu về lý thuyết xác suất và ứng dụng* (tái bản lần ba) Nhà xuất bản giáo dục 2008.
10. Đào Hữu Hồ, *Xác suất Thống kê*, NXB Đại học Quốc gia 2000.
11. Nguyễn Cao Văn, Trần Thái Ninh, *Lý thuyết xác suất và Thống kê toán*, NXB Khoa học Kỹ thuật 1996.

# MỤC LỤC

	<i>Trang</i>
<i>Lời nói đầu</i> .....	3
<b>Phần A – CÁC ĐỀ BÀI TẬP</b>	<b>5</b>
Chương 1. Thống kê mô tả.....	5
Chương 2. Đại cương về lí thuyết xác suất.....	14
Chương 3. Ước lượng tham số .....	20
Chương 4. Kiểm định giả thiết thống kê .....	27
Chương 5. Bài toán so sánh.....	39
Chương 6. Bài toán so sánh mở rộng.....	59
Chương 7. Phân tích tương quan và hồi quy.....	71
<b>Phần B – LỜI GIẢI VÀ ĐÁP SỐ</b>	<b>85</b>
Chương 1. Thống kê mô tả.....	85
Chương 2. Đại cương về lí thuyết xác suất.....	94
Chương 3. Ước lượng tham số .....	110
Chương 4. Kiểm định giả thiết thống kê .....	116
Chương 5. Bài toán so sánh.....	133
Chương 6. Bài toán so sánh mở rộng.....	159
Chương 7. Phân tích tương quan và hồi quy.....	169
<b>Phụ lục</b>	<b>182</b>
Bảng I. Bảng phân bố chuẩn .....	182
Bảng II. Bảng phân bố Student .....	184
Bảng III. Bảng phân bố "khi bình phương" .....	185
Bảng IV. Bảng phân bố Fisher .....	186
Tài liệu tham khảo .....	190