Chương 2

Ước lượng và kiểm định giả thiết

Kiểm định giả thiết là một bài toán hay gặp trong thống kê. Phạm vi nghiên cứu khá rộng và về mặt lý thuyết có những vấn đề khá phức tạp nếu muốn giải quyết thật tỷ mỷ, chính xác. Trong chương này chỉ trình bầy một vài bài toán kiểm định giả thiết cụ thể liên quan đến các biến định lượng. Chương sau sẽ tiếp tục kiểm định giả thiết với biến định tính. Nhưng trước hết cần giới thiệu chung về giả thiết và đối thiết và hai loại sai lầm mắc phải khi kiểm định.

2.1. Giả thiết và đối thiết

Khi khảo sát một tổng thể (hoặc nhiều tổng thể) và xem xét một (hoặc nhiều) biến ngẫu nhiên có thể đưa ra một giả thiết nào đó liên quan đến phân phối của biến ngẫu nhiên hoặc nếu biết phân phối rồi thì đưa ra giả thiết về tham số của tổng thể. Để có thể đưa ra một kết luận thống kê nào đó đối với giả thiết thì phải chọn mẫu ngẫu nhiên, tính tham số mẫu, chọn mức ý nghĩa α sau đó đưa ra kết luận.

Bài toán kiểm định tham số Θ của phân phối có dạng $H_0 : \Theta = \Theta_o$ với Θ_o là một số đã cho nào đó. Kết luận thống kê có dạng: "chấp nhận H_0" hay "bác bỏ H_0". Nhưng nếu đặt vấn đề như vậy thì cách giải quyết hết sức khó, vì nếu không chấp nhận $H_0 : \Theta = \Theta_o$ thì điều đó có nghĩa là có thể chấp nhận một trong vô số Θ khác Θ_o, do đó thường đưa ra bài toán dưới dạng cụ thể hơn nữa: cho giả thiết H_0 và đối thiết H_1, khi kết luận thì hoặc chấp nhận H_0 hoặc bác bỏ $H_{0,}$ và trong trường hợp này, tuy không hoàn toàn tương đương, nhưng coi như chấp nhận đối thiết H_1.

Nếu chấp nhận H_0 trong lúc giả thiết đúng là H_1 thì mắc **sai lầm loại II** và xác suất mắc sai lầm này được gọi là rủi ro loại hai **β**. Ngược lại nếu bác bỏ H_0 trong lúc giả thiết đúng chính là H_0 thì mắc **sai lầm loại I** và xác suất mắc sai lầm đó gọi là rủi ro loại một α.

	Quyết định	
Giả thiết	Bác bỏ H_0	Chấp nhận H_0
H_0 đúng	**Sai lầm loại I (α)**	**Quyết định đúng**
H_0 sai	**Quyết định đúng**	**Sai lầm loại II (β)**

Như vậy trong bài toán kiểm định giả thiết luôn luôn có hai loại rủi ro, loại I và loại II, tuỳ vấn đề mà nhấn mạnh loại rủi ro nào. Thông thường người ta hay tập trung chú ý vào **sai lầm loại I** và khi kiểm định phải khống chế sao cho **rủi ro loại I** không vượt quá một mức α gọi là **mức ý nghĩa**.

Trước hết xem xét cụ thể bài toán kiểm định giả thiết H_0: $\Theta = \Theta_o$, đối thiết H_1: $\Theta = \Theta_1$ với Θ_1 là một giá trị khác Θ_o. Đây là bài toán kiểm định giả thiết đơn. Quy tắc kiểm định căn cứ vào hai giá trị cụ thể Θ_1 và Θ_o, vào mức ý nghĩa α và còn căn cứ vào cả sai lầm loại hai. Việc này về lý thuyết thống kê không gặp khó khăn gì.

Sau đó mở rộng quy tắc sang cho bài toán kiểm định giả thiết kép. H_1: $\Theta \neq \Theta_o$; $\Theta > \Theta_o$ hoặc $\Theta < \Theta_o$, việc mở rộng này có khó khăn nhưng các nhà nghiên cứu lý thuyết xác suất thống kê đã giải quyết được, do đó về sau khi kiểm định giả thiết H_0 : $\Theta = \Theta_o$ có thể chọn một trong 3 đối thiết H_1 sau:

H_1 : $\Theta \neq \Theta_o$ gọi là đối thiết hai phía

H_1 : $\Theta > \Theta_o$ gọi là đối thiết phải

H_1 : $\Theta < \Theta_o$ gọi là đối thiết trái

Hai đối thiết sau gọi là đối thiết một phía. Việc chọn đối thiết nào tuỳ thuộc vấn đề khảo sát cụ thể. Trong phạm vi tài liệu này đề cập chủ yếu đến đối thiết hai phía hay còn gọi là hai đuôi.

2.2. Ước lượng giá trị trung bình μ của biến phân phối chuẩn $N(\mu, \sigma^2)$.

2.2.1. Ước lượng μ khi biết phương sai σ^2

Dựa vào lý thuyết xác suất có thể đưa ra ước lượng giá trị trung bình quần thể (μ) theo các bước sau đây:

+ Chọn mẫu dung lượng n, tính trung bình cộng $\overline{x}$

+ Ở mức tin cậy P đã cho lấy α = 1- P, sau đó tìm giá trị tới hạn $z_{(\alpha/2)}$ trong bảng 1 (hàm $\Phi(z)$ tìm z sao cho $\Phi(z) = 1 - \alpha/2$)

+ Khoảng tin cậy đối xứng ở mức tin cậy P:

$$\overline{x} - z(\alpha/2)\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{x} + z(\alpha/2)\frac{\sigma}{\sqrt{n}}$$

Ví dụ 2.1: Khối lượng bao thức ăn gia súc phân phối chuẩn $N(\mu,\sigma^2)$ với σ = 1,5kg. Cân thử 25 bao được khối lượng trung bình $\overline{x}$ = 49kg. Hãy ước lượng kỳ vọng μ với mức tin cậy P = 0,95; z (0,025) = 1,96

$$49 - 1{,}96\,\frac{1{,}5}{\sqrt{25}} \le \mu \le 49 + 1{,}96\,\frac{1{,}5}{\sqrt{25}}$$

$$49 - 0{,}588 \le \mu \le 49 + 0{,}588$$

$$48{,}41\text{kg} \le \mu \le 49{,}59\text{kg}$$

2.2.2. Ước lượng μ khi không biết phương sai σ^2

Dựa vào phân phối Student có thể đưa ra ước lượng μ theo các bước sau đây:

+ Chọn mẫu dung lượng n, tính trung bình cộng $\bar{x}$ và độ lệch chuẩn s.
+ Ở mức tin cậy P lấy α = 1- P, tìm giá trị tới hạn t(α/2, n-1) trong bảng 2, cột α/2, dòng n-1
+ Khoảng tin cậy đối xứng ở mức tin cậy P:

$$\bar{x} - t(\alpha/2, n-1)\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t(\alpha/2, n-1)\frac{s}{\sqrt{n}}$$

Ví dụ 2.2: Cân 22 con gà được khối lượng trung bình $\bar{x}$ = 3,03kg; s = 0,0279 kg. Hãy ước lượng μ với mức tin cậy P = 0,98; α = 1- P = 0,02; α/2 = 0,01 t(0,01;21) = 2,518

$$3{,}03 - 2{,}518\sqrt{\frac{0{,}0279}{22}} \le \mu \le 3{,}03 + 2{,}518\sqrt{\frac{0{,}0279}{22}}$$

$$3{,}03 - 0{,}089 \le \mu \le 3{,}03 + 0{,}089$$

$$2{,}94\text{kg} \le \mu \le 3{,}12\text{ kg}$$

2.3. Kiểm định giá trị trung bình μ của biến phân phối chuẩn N(μ, σ^2).

2.3.1. Kiểm định giả thiết H_0: $\mu = \mu_0$ khi biết σ^2

Tiến hành kiểm định theo các các bước sau:

+ Chọn mẫu dung lượng n, tính trung bình cộng $\bar{x}$
+ Chọn mức ý nghĩa α
+ Tìm giá trị tới hạn z(α/2) nếu kiểm định 2 phía hoặc z(α) nếu kiểm định một phía
+ Tính giá trị thực nghiệm $Z_{TN} = \dfrac{(\bar{x} - \mu_0)}{\dfrac{\sigma}{\sqrt{n}}} = \dfrac{(\bar{x} - \mu_0)\sqrt{n}}{\sigma}$

So sánh Z_{TN} và z tới hạn để rút ra kết luận theo nguyên tắc sau:

Kết luận:

Với $H_1 : \mu \neq \mu_0$ (Kiểm định hai phía)
Nếu $|Z_{TN}|$ (giá trị tuyệt đối của Z_{TN}) nhỏ hơn hay bằng z(α/2) thì chấp nhận H_0 nếu ngược lại thì bác bỏ H_0, tức là chấp nhận H_1.

Với $H_1 : \mu > \mu_0$ (Kiểm định một phía)
Nếu Z_{TN} nhỏ hơn hay bằng giá trị tới hạn z(α) thì chấp nhận H_0, ngược lại thì chấp nhận H_1.

Với H_1: $\mu < \mu_0$ (Kiểm định một phía)
Nếu Z_{TN} lớn hơn hay bằng giá trị tới hạn - z(α) thì chấp nhận H_0, ngược lại thì chấp nhận H_1.

Ví dụ 2.3: Nuôi 100 con cừu theo một chế độ riêng. Mục đích của thí nghiệm là xem chế độ này có làm tăng khối lượng của cừu một năm tuổi hay không. Biết rằng 100 cừu này được lấy mẫu từ một quần thể có khối lượng trung bình một năm tuổi là 30 kg và phương sai là 25 kg². Giả thiết tăng trọng phân phối chuẩn N(μ,25), hãy kiểm định giả thiết H_0: μ = 30 đối thiết H_1: μ > 30 ở mức α= 0,05. Biết rằng khối lượng trung bình của 100 cừu thí nghiệm là 32 kg.

$$Z_{TN} = \frac{(32-30)\sqrt{100}}{5} = 4; \quad z(0,05) = 1,64$$

Kết luận: Vì $Z_{TN} > Z_{LT}$ nên giả thiết H_0 bị bác bỏ, như vậy tăng trọng trung bình không phải là 30 kg. Chế độ nuôi mới đã làm tăng khối lượng cừu một năm tuổi.

Ví dụ 2.4: Một mẫu cho trước gồm 100 bò sữa có sản lượng sữa một chu kỳ tiết sữa trung bình là 3850kg. Số bò này có xuất phát từ quần thể có giá trị trung bình là 4000kg và độ lệch chuẩn là 1000 hay không? Giả sử sản lượng sữa của quần thể tuân theo phân phối chuẩn N((μ,1000²). Hãy kiểm định giả thiết H_0: μ = 4000 đối thiết H_1: μ ≠ 4000 ở mức α= 0,05

$$Z_{TN} = \frac{(3850-4000)\sqrt{100}}{1000} = -1,5 \Rightarrow |Z_{TN}| = 1,5; \quad z(0,025) = 1,96$$

Kết luận: Chấp nhận H_0, số bò sữa nêu trên xuất phát từ một quần thể ban đầu có sản lượng sữa chu kỳ là 4000kg.

2.3.2. Kiểm định giả thiết H_0: $\mu = \mu_0$ khi không biết σ^2

Đây là trường hợp phổ biến khi kiểm định giá trị trung bình của phân phối chuẩn. Tiến hành các bước sau:

+ Lấy mẫu dung lượng n, tính $\bar{x}$ và s^2

+ Tính giá trị T thực nghiệm $T_{TN} = \dfrac{(\bar{x}-\mu_0)\sqrt{n}}{s}$

+ Tìm giá trị tới hạn t(α/2, n-1) với kiểm định 2 phía hoặc tìm t(α, n-1) nếu kiểm định 1 phía trong bảng 2.

Kết luận:

Với $H_1 : \mu \neq \mu_0$ (Kiểm định hai phía)
Nếu $|T_{TN}|$ (giá trị tuyệt đối của T_{tn}) nhỏ hơn hay bằng t(α/2,n-1) thì chấp nhận H_0 nếu ngược lại thì bác bỏ H_0, tức là chấp nhận H_1

Với $H_1 : \mu > \mu_0$ (Kiểm định một phía)
Nếu $T_{TN} \leq$ t(α, n-1) thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Với H_1: $\mu < \mu_0$ (Kiểm định một phía)
Nếu $T_{TN} \geq$ - t(α, n-1) thì chấp nhận H_0, ngược lại thì chấp nhận H_1.

Ví dụ 2.5: Thời gian mang thai của bò phân phối chuẩn $N(285,\sigma^2)$. Theo dõi thời gian mang thai (ngày) của 6 bò được các số liệu

307 293 293 283 294 297

Kiểm định giả thiết H_0: $\mu = 285$ ngày đối thiết H_1: $\mu \neq 285$ ngày

Tính $\bar{x} = \dfrac{(307+293+293+283+294+297)}{6} = \dfrac{1767}{6} = 294{,}5$

$$s^2 = \frac{307^2 + 293^2 + ... + 294^2 + 297^2) - \dfrac{1767^2}{6}}{5} = 59{,}9;\ s = \sqrt{59{,}9} = 7{,}7395 \approx 7{,}74$$

$$T_{TN} = \frac{(294{,}5 - 285)}{7{,}74} \times \sqrt{6} = \frac{9{,}5}{3{,}16} = 3{,}007;\ \ t(0{,}025;5) = 2{,}571$$

Kết luận: Vì $|T_{TN}| = 3{,}007 > t(0{,}025;5)$ nên bác bỏ H_0 như vậy thời gian mang thai không phải 285 ngày

Ví dụ 2.6: Trong điều kiện chăn nuôi bình thường, lượng sữa trung bình của một con bò là 19 kg / ngày. Trong một đợt hạn, người ta theo dõi 25 con bò và được lượng sữa trung bình 17,5 kg/ ngày, độ lệch chuẩn s = 2,5 kg. Giả thiết lượng sữa phân phối chuẩn, hãy kiểm định giả thiết H_0: $\mu = 19$ với đối thiết $\mu < 19$ ở mức $\alpha = 0{,}05$.

$$T_{TN} = = \frac{(17{,}5 - 19)\sqrt{25}}{2{,}5} \text{ - } 3\ ; \qquad t(0{,}05;24) = 1{,}711$$

Kết luận: $T_{TN} < -1{,}711$ nên giả thiết H0 bị bác bỏ, như vậy sản lượng sữa trung bình không còn là 19 kg / ngày nữa mà thấp hơn.

2.4. Kiểm định hai giá trị trung bình của hai biến phân phối chuẩn

Giả sử chúng ta có hai tổng thể và theo dõi một biến định lượng X nào đó, ví dụ khối lượng sau 6 tháng nuôi của hai đàn gà, năng suất của hai giống lúa, năng suất của một giống ngô khi bón theo hai công thức phân bón khác nhau, sản lượng một loại quả khi trồng theo hai khoảng cách hàng . . .

Chúng ta gọi biến X trên tổng thể thứ nhất là X_1 (phân phối chuẩn $N(\mu_1,\sigma_1^2)$) và biến X trên tổng thể thứ hai là X_2 (phân phối chuẩn $N(\mu_2,\sigma_2^2)$). Để so sánh μ_1 và μ_2 chúng ta phải chọn mẫu. Có hai cách chọn mẫu: **Chọn mẫu theo cặp** và **chọn mẫu độc lập**.

2.4.1. Chọn mẫu theo cặp

Từ tổng thể thứ nhất ta chọn một mẫu n cá thể được các giá trị $x_1, x_2, . . . ,x_n$, từ tổng thể thứ hai chọn một mẫu cũng gồm n cá thể được $y_1, y_2, . . ., y_n$.

Giữa hai mẫu này có mối quan hệ cặp, tức là có n cặp (x_i, y_i) (i = 1, n). Các cặp này hình thành do khi chọn mẫu ta đã dùng những quan hệ cặp như quan hệ gia đình (vợ chồng, anh em, thí dụ chọn n tổ chim sau đó bắt chim đực vào mẫu đại diện cho tổng thể chim đực, bắt chim cái vào mẫu đại diện cho tổng thể chim cái), quan hệ trước sau (thí dụ cá thể được đo một chỉ số trước khi dùng thuốc và số liệu này đại diện cho tổng thể trước khi dùng thuốc, một thời gian sau khi dùng thuốc lại đo lại chỉ số và số liệu này đại diện cho tổng thể sau khi dùng thuốc), cũng có khi các cặp này là các cặp số liệu do chúng ta bố trí thí nghiệm theo cặp: chọn 2 ô ruộng, một ô ruộng(hay một chuồng) bố trí giống thử nghiệm, một ô ruộng (một chuồng) bố trí giống đối chứng.

Viết lại số liệu dưới dạng hai cột hay hai hàng rồi tính hiệu số $d_i = y_i - x_i$

X_1	x_1	x_2	...	x_n
X_2	y_1	y_2	...	y_n
d	d_1	d_2	...	d_n

Tiếp theo tính giá trị trung bình $\bar{d}$ và độ lệch chuẩn s_d

Giả thiết H_0: $\mu_2 = \mu_1$ đối thiết H_1: $\mu_2 \neq \mu_1$ được chuyển thành H_0: $\mu_d = 0$ đối thiết H_1: $\mu_d \neq 0$ (tương tự H_1: $\mu_2 > \mu_1$ chuyển thành H_1: $\mu_d > 0$ và H_1: $\mu_2 < \mu_1$ chuyển thành H_1: $\mu_d < 0$).

Ở mức ý nghĩa α việc kiểm định gồm các bước sau:

+ Tính giá trị thực nghiệm $T_{TN} = \dfrac{\bar{d}\sqrt{n}}{s_d}$

+ Tìm giá trị tới hạn $t(\alpha/2, n-1)$ nếu kiểm định 2 phía hoặc $t(\alpha, n-1)$ nếu kiểm định một phía bảng 2

Kết luận:

+ Kiểm định hai phía H_1: $\boldsymbol{\mu_2 \neq \mu_1}$
Nếu $|T_{TN}| \leq t(\alpha/2, n-1)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\boldsymbol{\mu_2 > \mu_1}$
Nếu $T_{TN} \leq t(\alpha, n-1)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\boldsymbol{\mu_2 < \mu_1}$
Nếu $T_{TN} \geq - t(\alpha, n-1)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Ví dụ 2.7: Tăng trọng (pound) của 10 cặp bê sinh đôi giống hệt nhau dưới hai chế độ chăm sóc khác nhau (A và B). Bê trong từng cặp được bắt thăm ngẫu nhiên về một trong hai cách chăm sóc. Giả thiết tăng trọng có phân phối chuẩn. Hãy kiểm định giả thiết H_0: Tăng trọng trung bình ở hai cách chăm sóc như nhau, đối thiết H_1: Tăng trọng trung bình khác nhau ở hai cách chăm sóc với mức ý nghĩa $\alpha = 0,05$. Số liệu thu được như sau:

Cặp sinh đôi	1	2	3	4	5	6	7	8	9	10
Tăng trọng ở cách A	43	39	39	42	46	43	38	44	51	43
Tăng trọng ở cách B	37	35	34	41	39	37	35	40	48	36
Chênh lệch (d)	6	4	5	1	7	6	3	4	3	7

n = 10; $\bar{d}$ = 4,6; s_d = 1,955; $T_{TN} = \dfrac{4,6\sqrt{10}}{1,955} = 7,44$; t(0,025;9) = 2,262

Kết luận: Bác bỏ giả thiết H_0, chấp nhận H_1: “Tăng trọng trung bình ở hai cách chăm sóc là khác nhau”.

Ví dụ 2.8: Có 15 trại phối hợp tham gia thử nghiệm khẩu phần ăn bình thường (A) và khẩu phần ăn có bổ sung đồng (B). Mỗi trại lấy 2 khu nuôi lợn tương tự về mọi mặt sau đó chỉ định ngẫu nhiên một khu ăn khẩu phần A, một khu ăn khẩu phần B. Tăng trọng trung bình (kg/ngày) của một con lợn được trình bày ở bảng dưới. Kiểm định giả thiết H_0: "Hai khẩu phần A và B cho kết quả tăng trọng trung bình như nhau" với đối thiết H_1: "Khẩu phần có bổ sung đồng cho tăng trọng trung bình cao hơn"

	Khẩu phần			Khẩu phần			Khẩu phần	
Trại	A (x_i)	B (y_i)	Trại	A (x_i)	B (y_i)	Trại	A (x_i)	B (y_i)
1	0,42	0,53	6	0,50	0,52	11	0,50	0,51
2	0,53	0,47	7	0,44	0,44	12	0,54	0,54
3	0,48	0,56	8	0,45	0,46	13	0,46	0,50
4	0,50	0,59	9	0,30	0,43	14	0,48	0,50
5	0,42	0,47	10	0,52	0,57	15	0,53	0,59

Giá trị trung bình $\bar{d}$ = 0,0407; độ lệch chuẩn s_d = 0,0489

$T_{TN} = \frac{0{,}0407}{0{,}0489} \times \sqrt{15} = 3{,}22$; t(0,05;14) = 1,761

Kết luận: Vì $T_{TN} > t$ nên bác bỏ H_0, chấp nhận H_1. Như vậy khẩu phần bổ sung đồng cho tăng trọng trung bình cao hơn khẩu phần ăn thường.

2.4.2. Chọn mẫu độc lập

Từ hai tổng thể chọn ra hai mẫu độc lập, dung lượng có thể bằng nhau hoặc khác nhau. Tính các tham số thống kê $\bar{x}_1$; s_1^2 của mẫu thứ nhất; $\bar{x}_2$; s_2^2 của mẫu thứ hai. Để kiểm định giả thiết H_0: $\mu_2 = \mu_1$ với các đối thiết H_1 ở mức ý nghĩa α ta chia ra 3 trường hợp:

2.4.2.1. Biết phương sai σ_1^2 và σ_2^2

+ Tính Z thực nghiệm $$Z_{TN} = \frac{(\bar{x}_2 - \bar{x}_1)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

+ Tìm giá trị tới hạn $z(\alpha/2)$ nếu kiểm định 2 phía hoặc $z(\alpha)$ nếu kiểm định một phía trong bảng 1

Kết luận:

+ Kiểm định hai phía H_1: $\mu_2 \neq \mu_1$
Nếu $|Z_{TN}| \leq z(\alpha/2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\mu_2 > \mu_1$
Nếu $Z_{TN} \leq z(\alpha)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\mu_2 < \mu_1$
Nếu $Z_{TN} \geq - z(\alpha)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Ví dụ 2.9: Chiều dài cá trong 2 ao phân phối chuẩn với độ lệch chuẩn σ_1 = 2cm và σ_2 = 2,2cm. Lấy mẫu 100 con của ao thứ nhất được giá trị trung bình $\bar{x}_1$ = 8 cm; lấy mẫu 120 con của ao thứ hai được giá trị trung bình $\bar{x}_2$ = 8,5 cm. Hãy kiểm định giả thiết H_0: $\mu_1 = \mu_2$ với đối thiết H_1: $\mu_1 \neq \mu_2$ ở mức ý nghĩa α=0,05

$$Z_{TN} = \frac{(8,5-8)}{\sqrt{\frac{2^2}{100}+\frac{2,2^2}{120}}} = 1,764 \quad ; \quad z(0,025) = 1,96$$

Vì $|Z_{TN}|$ = 1,764 < 1,96 nên chấp nhận H_0: "Chiều dài cá trung bình trong 2 ao như nhau".

2.4.2.2. Không biết phương sai σ_1^2 và σ_2^2 mẫu lớn($n_1 \geq 30$, $n_2 \geq$ 30).

+ Tính giá trị thực nghiệm $Z_{TN} = \dfrac{(\bar{x}_2 - \bar{x}_1)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$

+ Tìm giá trị tới hạn $z(\alpha/2)$ nếu kiểm định 2 phía hoặc $z(\alpha)$ nếu kiểm định một phía trong bảng 1

Kết luận:

+ Kiểm định hai phía H_1: $\mu_2 \neq \mu_1$
Nếu $|Z_{TN}| \leq z(\alpha/2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\mu_2 > \mu_1$
Nếu $Z_{TN} \leq z(\alpha)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\mu_2 < \mu_1$
Nếu $Z_{TN} \geq -\ z(\alpha)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Ví dụ 2.10: Để đánh giá tăng trọng của lợn ở hai chế độ ăn khác nhau. Khối lượng sau 4 tháng ở hai chế độ nuôi có các số liệu sau. Ở chế độ thứ nhất, tiến hành thí nghiệm 64 con (n_1 = 64) được giá trị trung bình $\bar{x}_1$ = 73,2 kg biết σ_1 = 10,9 kg; tương tự với chế độ thứ 2 ta có n_2 = 68; $\bar{x}_2$ = 76,6; σ_2 = 11,4 kg. Giả thiết khối lượng phân phối chuẩn $N(\mu_1,\sigma_1^2)$ và $N(\mu_2,\sigma_2^2)$). Kiểm định giả thiết H_0: $\mu_2 = \mu_1$ với đối thiết H_1: $\mu_2 > \mu_1$

$$Z_{TN} = \frac{76,6-73,2}{\sqrt{\frac{10,9^2}{64}+\frac{11,4^2}{68}}} = 1,75; \qquad z(0,05) = 1,645$$

Kết luận:

$Z_{TN} > z(0,05)$ vì vậy chấp nhận H_1: "chế độ ăn thứ hai cho kết quả trung bình cao hơn chế độ ăn thứ nhất".

2.4.2.3. Không biết phương sai σ_1^2 và σ_2^2, mẫu bé (ít nhất một trong 2 số n_1, n_2 <30)

Đây là một bài toán còn rất nhiều vướng mắc về mặt lý thuyết do đó chúng ta chỉ trình bầy trường hợp có thêm giả thiết phụ : $\sigma_1^2 = \sigma_2^2$

+ Tính phương sai chung: $s^2_c = \dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

+ Tính $T_{TN} = \dfrac{(\bar{x}_2-\bar{x}_1)}{\sqrt{s_c^2(\frac{1}{n_1}+\frac{1}{n_2})}}$

+ Tìm giá trị tới hạn $t(\alpha/2, n_1 + n_2 - 2)$ với kiểm định 2 phía hoặc $t(\alpha, n_1 + n_2 - 2)$ nếu kiểm định một phía

Kết luận:

+ Kiểm định hai phía H_1: $\mu_2 \neq \mu_1$
Nếu $|T_{TN}| \leq t(\alpha/2, n_1+n_2-2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\mu_2 > \mu_1$
Nếu $T_{TN} \leq t(\alpha, n_1+n_2-2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

+ Kiểm định một phía H_1: $\mu_2 < \mu_1$
Nếu $T_{TN} \geq -t(\alpha, n_1+n_2-2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Ví dụ 2.11: Để so sánh khối lượng của 2 giống bò, chọn ngẫu nhiên 12 bò của giống thứ nhất và 15 bò của giống thứ 2. Khối lượng (kg) của từng bò được xác định và thu được các tham số thống kê sau: $n_1 = 12$; $\bar{x}_1 = 196,2$kg; $s_1 = 10,62$ kg; $n_2 = 15$; $\bar{x}_2 = 153,70$kg; $s_2 = 12,30$kg. Kiểm định giả thiết H_0: Hai giống bò có khối lượng trung bình như nhau với đối thiết H_1: Giống bò thứ nhất có khối lượng trung bình lớn hơn giống bò thứ hai. Giả sử khối lượng của 2 giống bò có phân phối chuẩn và hai phương sai bằng nhau với mức ý nghĩa $\alpha = 0,05$.

$$s_c^2 = \frac{(11\times 10,62^2 + 14\times 12,30^2)}{11+14} = 134,33$$

$$T_{TN} = \frac{(196,2-153,7)}{\sqrt{134,33\times\left(\frac{1}{12}+\frac{1}{15}\right)}} = \frac{42,5}{4,489} = 9,46\text{ ; } t(0,05,25) = 1,708$$

Kết luận: Ở mức ý nghĩa $\alpha = 0,05$ vì $T_{TN} > t$ nên bác bỏ H_0. Như vậy giống thứ nhất có khối lượng trung bình cao hơn giống thứ hai.

Ví dụ 2.12 : Hai giống gà có khối lượng phân phối chuẩn, lấy mẫu 10 gà đối với giống thứ nhất và 16 gà của giống thứ 2. Các tham số về khối lượng 45 ngày tuổi của 2 mẫu nêu trên như sau:

Với mẫu thứ nhất $n_1 = 10$; $\bar{x}_1 = 2,8$kg; $s_1^2 = 0,1111$ kg² với mẫu thứ hai $n_2 = 16$; $\bar{x}_2 = 2,35$kg; $s_2^2 = 0,0667$kg². Kiểm định giả thiết H_0: Hai giống gà có khối lượng trung bình như nhau với đối thiết H_1: Hai giống gà có khối lượng trung bình khác nhau. Mức ý nghĩa $\alpha = 0,05$.

$$s_c^2 = \frac{9 \times 0{,}1111 + 15 \times 0{,}0667}{9+15} = \frac{1{,}99995}{24} = 0{,}83331$$

T_{TN} = = -3,866 ➔ $|T_{TN}|$ = 3,866; t(0,025;24) = 2,064

Kết luận: Bác bỏ H_0, như vậy hai giống gà có khối lượng trung bình khác nhau.

2.5. Ước lượng và kiểm định xác suất

Trường hợp tổng thể có 2 loại cá thể A và A', loại A chiếm tỷ lệ p và A' chiếm tỷ lệ q = 1-p. Sau khi chọn mẫu có thể dùng phân phối chuẩn để tính gần đúng phân phối nhị thức, từ đó suy ra công thức ước lượng p.

2.5.1. Ước lượng xác suất p

Khi dung lượng mẫu lớn (n ≥ 30 nhưng thực tế tốt nhất là trên 100) và p không bé quá, cũng không lớn quá (np > 5, nq > 5). Từ mẫu có dung lượng n, tính số cá thể loại A được tần số m và tần suất f = m/ n với mức tin cậy P có khoảng tin cậy đối xứng sau:

$$f - z(\alpha/2)\sqrt{\frac{f(1-f)}{n}} \le p \le f + z(\alpha/2)\sqrt{\frac{f(1-f)}{n}}$$

Ví dụ 2.13: Để biết tỷ lệ trứng nở p của một loại trứng; cho vào máy ấp 100 quả, kết quả có 80 quả nở.

f = 80 / 100 = 0,8 ở mức tin cậy P = 0,95 thì α = 0,05 và z(0,025) = 1,96. Ta có thể tính được khoảng tin cậy như sau:

$$0{,}8 - 1{,}96\sqrt{\frac{0{,}8 \times 0{,}2}{100}} \le p \le 0{,}8 + 1{,}96\sqrt{\frac{0{,}8 \times 0{,}2}{100}}$$

$$0{,}8 - 0{,}0784 \le p \le 0{,}8 + 0{,}0784 \Leftrightarrow 0{,}72 \le p \le 0{,}88$$

2.5.2. Kiểm định giả thiết H_0: p = p_0

Khi dung lượng mẫu lớn (n ≥ 30 nhưng thực tế thấy tốt nhất là trên 100) và p không bé quá, cũng không lớn quá (np > 5, nq > 5). Từ mẫu có dung lượng n, tính số cá thể loại A được tần số m và tần suất f = m / n. Ở mức ý nghĩa α tính z(α/2) với kiểm định 2 phía hoặc z(α) nếu kiểm định một phía.

Tính $Z_{TN} = \dfrac{f - p_0}{\sqrt{\dfrac{p_o(1-p_0)}{n}}}$

Kết luận:

Với đối thiết hai phía H_1: $p \ne p_0$
Nếu $|Z_{TN}| \le z(\alpha/2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Với đối thiết một phía H_1: $p > p_0$
Nếu $Z_{TN} \le z(\alpha)$ thì chấp nhận H0, ngược lại thì chấp nhận H_1

Với đối thiết một phía H_1**:** $p < p_0$
Nếu $Z_{TN} \ge -z(\alpha)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Ví dụ 2.14: Ấp 100 quả trứng có 82 quả nở. Kiểm định giả thiết H_0: tỷ lệ nở p = 0,80, đối thiết H_1: $p \neq 0,8$ với $\alpha = 0,05$.

n = 100; m = 82; f = 82/100 = 0,82;

$$Z_{TN} = \frac{0,82 - 0,80}{\sqrt{\frac{0,8.0,2}{100}}} = 0,5 \text{ ; } z(0,025) = 1,96$$

Kết luận: Chấp nhận H_0: "Tỷ lệ ấp nở là 0,80".

2.5.3. Kiểm định giả thiết H_0: $p_2 = p_1$

Khi dung lượng cả 2 mẫu đều lớn ($n_1 > 100$, $n_2 > 100$) và các p_i không bé quá (hoặc lớn quá) có thể kiểm định như sau (ở mức ý nghĩa α)

Tính các tần suất: $f_1 = \frac{m_1}{n_1}$; $f_2 = \frac{m_2}{n_2}$

Tính tần suất chung: $f = \frac{m_1 + m_2}{n_1 + n_2}$

Tìm giá trị tới hạn z(α/2) nếu kiểm định 2 phía hoặc z(α) nếu kiểm định một phía

Tính giá trị thực nghiệm: $Z_{TN} = \frac{f_2 - f_1}{\sqrt{f(1-f)(\frac{1}{n_1} + \frac{1}{n_2})}}$

Kết luận:

Với đối thiết hai phía H_1: $p_2 \neq p_1$
Nếu $|Z_{TN}| \leq z(\alpha/2)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Với đối thiết một phía H_1: $p_2 > p_1$
Nếu $Z_{TN} \leq z(\alpha)$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Với đối thiết một phía H_1: $p_2 < p_1$
Nếu $Z_{TN} \geq -z(\alpha)$ thì chấp nhận H_0, ngược lại thì thì chấp nhận H_1

Ví dụ 2.15: Dùng thuốc A điều trị cho 200 bệnh nhân thấy 150 người khỏi bệnh. Tương tự với thuốc B đối với 100 bệnh nhân thì 72 người khỏi bệnh. Hãy kiểm định giả thiết H_0: Tỷ lệ khỏi bệnh của hai thuốc như nhau với đối thiết H_1: tỷ lệ khỏi bệnh của hai thuốc khác nhau với mức ý nghĩa α = 0,05.

$n_1 = 200$; $m_1 = 150$; $f_1 = 150/200 = 0,75$; $n_2 = 100$; $m_2 = 72$; $f_2 = 72/100 = 0,72$;

$$f = \frac{150 + 72}{200 + 100} = 0,74$$

$$Z_{TN} = \frac{0,72 - 0,75}{\sqrt{0,74 \times 0,26 \times (\frac{1}{200} + \frac{1}{100})}} = -0,5584 \Rightarrow |Z_{TN}| = 0,5584; z(0,025) = 1,96.$$

Kết luận: Chấp nhận H_0; tức là tỷ lệ khỏi bệnh ở 2 loại thuốc là như nhau.

2.6. Phân tích phương sai

Mở rộng bài toán so sánh hai trung bình của hai tổng thể ở mục trên khi có nhiều hơn 2 trung bình chúng ta có bài toán phân tích phương sai một nhân tố. Thí dụ có a tổng thể, để khảo sát các biến $X_1, X_2, \dots, X_a$ trên các tổng thể đó chúng ta lấy ở mỗi tổng thể một mẫu các quan sát độc lập:

Mẫu 1	$x_{11}, x_{12}, \dots, x_{1r1}$
Mẫu 2	$x_{21}, x_{22}, \dots, x_{2r2}$
….	….
Mẫu a	$x_{a1}, x_{a2}, \dots, x_{2ra}$

Tất cả có $n = \Sigma r_i$ quan sát. Viết lại các quan sát $x_{i\,j}$ dưới dạng

$$x_{i\,j} = \mu_i + e_{i\,j} \qquad e_{i\,j} \text{ gọi là sai số hay phần dư} \qquad (2.1)$$

Giả thiết các biến X_i độc lập, phân phối chuẩn $N(\mu_i, \sigma^2)$, các quan sát trong mẫu độc lập.Từ giả thiết trên có thể nêu cụ thể 3 giả thiết sau đối với các sai số $e_{i\,j}$

a- Các biến $e_{i\,j}$ độc lập với nhau

b- Các biến $e_{i\,j}$ phân phối chuẩn với kỳ vọng bằng 0

c- Các biến $e_{i\,j}$ có phương sai bằng nhau (σ^2)

Bài toán phân tích phương sai một nhân tố chính là bài toán kiểm định giả thiết H_0: "Các trung bình μ_i bằng nhau" với đối thiết H_1: "Có ít nhất một cặp trung bình khác nhau".

Nếu gọi μ là trung bình của các μ_i thì có thể viết (2.1) lại như sau:

$$x_{i\,j} = \mu + a_i + e_{i\,j} \qquad (2.2)$$

với $a_i = \mu_i - \mu$; $\Sigma a_i = 0$

Giả thiết H_0 bây giờ là : "Các a_i đều bằng 0" còn H_1 là "Không phải tất cả các a_i đều bằng 0".

Để phân tích phương sai chúng ta gọi các trung bình cộng của các mẫu quan sát là $\bar{x}_i$. Nếu giả thiết H_0 đúng thì các X_i có cùng phân phối $N(\mu,\sigma^2)$ và có thể coi các mẫu quan sát nói trên được lấy ra từ cùng một tổng thể.

Gọi $\bar{x}$ là trung bình chung của tất cả các mẫu.

Tính tổng bình phương tất cả các sai số (gọi là tổng bình phương toàn bộ SS_{TO})

$$SS_{TO} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i}x_{ij}^2 - n\bar{x}^2$$

Đem tổng bình phương này chia cho (n - 1) được một ước lượng của σ^2.

SS_{TO}/σ^2 phân phối χ^2 với df_{TO} = (n - 1) bậc tự do.

Đối với mỗi mẫu quan sát chúng ta tính tổng bình phương sai số trong mẫu (mà nếu đem chia cho bậc tự do tương ứng (n_i - 1) thì được một ước lượng của σ^2) sau đó gộp lại thành tổng bình phương do sai số SS_E (Giống như cách đã làm khi đi tìm phương sai chung s^2_c trong trường hợp mẫu bé và hai phương sai bằng nhau ở mục 2.4.2.3)

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$$

Đem SS_E chia cho n - a được một ước lượng của σ^2

SS_E / σ^2 phân phối χ^2 với $df_E = (n - a)$ bậc tự do.

Có thể chứng minh hệ thức sau:

$$\sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij}-\bar{x})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2 + \sum_{i=1}^{a}\sum_{j=1}^{n_i}(\bar{x}_i-\bar{x})^2$$

Tổng thứ ba gọi là tổng bình phương do nhân tố SS_A.

Nếu x_{ij} phân phối chuẩn $N(\mu_i, \sigma^2)$ thì các trung bình cộng $\bar{x}_{i.}$ phân phối chuẩn $N(\mu_i, \sigma^2/n_i)$. Từ đó suy ra nếu đem SS_A chia cho (a - 1) thì được ước lượng của σ^2.

Tổng SS_A/σ^2 phân phối χ^2 với $df_A = (a-1)$ bậc tự do.

Như vậy chúng ta đã tách tổng bình phương toàn bộ ra hai tổng:

$$SS_{TO} = SS_A + SS_E$$

Đồng thời bậc tự do toàn bộ cũng tách thành 2 bậc tự do:

$$df_{TO} = df_A + df_E$$

Mỗi tổng bình phương chia cho bậc tự do tương ứng sẽ cho một ước lượng của phương sai σ^2 và mỗi tổng sau khi chia cho σ^2 sẽ phân phối χ^2 với số bậc tự do tương ứng.

Bây giờ xét tỷ số MS_A / MS_E với $MS_A = SS_A / df_A$ và $MS_E = SS_E / df_E$

Dựa trên lý thuyết về phân phối Khi bình phương (χ^2) và phân phối F có kết luận sau: MS_A / MS_E phân phối Fisher- Snederco (F). Từ đó có cách kiểm định sau đây đối với giả thiết H_0 (đối thiết H_1):

+ Tính giá trị thực nghiệm $F_{TN} = MS_A / MS_E$
+ Tìm giá trị tới hạn $F_{(\alpha,dfA,dfE)}$
+ Nếu $F_{TN} \le F_{(\alpha,dfA,dfE)}$ thì chấp nhận H_0, ngược lại thì chấp nhận H_1

Toàn bộ quy trình phân tích phương sai được tóm tắt trong bảng phân tích phương sai sau:

Nguồn biến động	Bậc tự do	Tổng bình phương	Trung bình bình phương	F_{TN}	F tới hạn
Nhân tố	df_A = a-1	SS_A	$MS_A = SS_A/df_A$	MS_A/MS_E	$F_{(\alpha,dfA,dfE)}$
Sai số ngẫu nhiên	df_E = n-a	SS_E	$MS_E = SS_E/df_E$		
Tổng biến động	df_{TO} = n-1	SS_{TO}			

Để thuận tiện thường kẻ bảng chứa dữ liệu và tính theo thứ tự sau:

+ Tính dung lượng n_i, tổng hàng TH_i, trung bình $\bar{x}_{i.}$, TH^2_i / n_i
+ Tổng các dung lượng $n = \Sigma n_i$, tổng tất cả các x_{ij} $ST = \Sigma\Sigma x_{ij}$
+ Số điều chỉnh $G = ST^2 / n$
+ $SS_{TO} = \Sigma\Sigma x^2_{ij} - G$ bậc tự do $df_{TO} = n - 1$
+ $SS_A = \Sigma TH^2_i / n_i - G$ bậc tự do $df_A = a - 1$
+ $SS_E = SS_{TO} - SS_A$ bậc tự do $df_E = df_{TO} - df_A = n - a$
+ Tính các trung bình $MS_A = SS_A / df_A$ và $MS_E = SS_E / df_E$
+ Tính $F_{TN} = MS_A / MS_E$
+ Tìm giá trị $F_{(\alpha,dfA,dfE)}$
+ So sánh F_{TN} với $F_{(\alpha,dfA,dfE)}$.

Ví dụ 2.16: Khối lượng (kg) của 20 lợn 90 ngày tuổi được nuôi ở 5 chế độ khác nhau từ lúc cai sữa 21 ngày tuổi. Biết rằng 20 lợn được chọn đồng đều nhau vào thời điểm cai sữa và bố trí ngẫu nhiên về một trong 5 công thức thí nghiệm. Số liệu được trình bày trong bảng dưới. Giả thiết khối lượng tuân theo phân phối chuẩn. Kiểm định giả thiết H_0: Khối lượng trung bình của lợn 90 ngày tuổi ở 5 chế độ chăm sóc bằng nhau với đối thiết H1: Khối lượng trung bình của lợn 90 ngày tuổi ở 5 chế độ chăm sóc không bằng nhau. Mức ý nghĩa α = 0,05.

Công thức	Khối lượng (kg)					n_i	TH_i	TH_i^2/n_i	$\bar{x}_{i.}$
A	32,2	34,9	29,7			3	96,8	3123,413	32,27
B	28,4	28,0	22,8	28,5	29,4	5	137,1	3759,282	27,42
C	28,8	29,5	23,1	20,1		4	101,5	2575,563	25,38
D	41,5	36,3	31,7	31,0	38,2	5	178,7	6386,738	35,74
E	33,0	26,0	30,6			3	89,6	2676,053	29,87
Tổng						20	603,7	18521,0492	

n = 20 ST = 603,7 $\Sigma TH^2_i / n_i$ = 18521,0492

Số điều chỉnh G = $603,7^2$ / 20 = 18222,6845

Tổng các bình phương $\Sigma\Sigma\, x^2_{i\,j}$ = 18727,6900

SS_{TO} = 18727,69 - 18222,68 = 505,0055; bậc tự do df_{TO} = 20 -1 = 19

SS_A = 18521,0492 - 18222,6845 = 298,3647; bậc tự do df_A = 5 - 1 = 4

SS_E = 505,0055 - 298,3647 = 206,6408; bậc tự do df_E = 19 - 4 = 15

MS_A = 298,3647 / 4 = 74,5912; MS_E = 206,6408 / 15 = 13,7761

F_{TN} = 74,5912/ 13,7761 = 5,4145 F(0,05;4;15) = 3,056

Có thể tổng hợp các kết quả thu được theo bảng phân tích phương sai (ANOVA) sau:

Nguồn biến động	Bậc tự do	Tổng bình phương	Trung bình bình phương	F_{TN}	F tới hạn
Công thức	4	298,3647	74,5912	3,056	$F_{(0,05;4;15)}$ = 3,056
Sai số ngẫu nhiên	15	206,6408	13,7761		
Tổng biến động	19	505,0055			

Kết luận: Bác bỏ H_0, như vậy là bác bỏ giả thiết "Khối lượng trung bình của lợn 90 ngày tuổi ở 5 chế độ chăm sóc bằng nhau".

Sau khi có kết luận như trên thì vấn đề đặt ra là phải so sánh 5 trung bình của 5 lô để tìm ra các trung bình nào bằng nhau, các trung bình nào khác nhau. Vấn đề này sẽ được trình bầy kỹ ở phần sau.

Qua cách làm như trên chúng ta thấy để kiểm định giả thiết H_0: "Các trung bình bằng nhau" với đối thiết H_1: "Có ít nhất một cặp trung bình khác nhau" phải tìm cách tách tổng bình phương toàn bộ SS_{TO} thành các tổng bình phương SS_A và SS_E căn cứ vào 2 nguồn biến động của số liệu: biến động do sự khác nhau giữa các mẫu và biến động do sự khác nhau giữa các số liệu trong cùng một mẫu. Đồng thời phải tách bậc tự do toàn bộ df_{TO} thành các bậc tự do df_A và df_E tương ứng với các tổng SS_A, SS_E. Từ đó có tên phân tích phương sai.

Trong phần sau khi có nhiều nguồn biến động thì phải tách SS_{TO} thành nhiều tổng ứng với các nguồn biến động và tách bậc tự do df_{TO} thành nhiều bậc tự do, sau đó kiểm định các giả thiết tương ứng với các nguồn biến động nhờ phân phối Fisher- Snederco.

2.7. Bài tập

2.7.1

Tăng trọng trung bình (gram/ngày) của 36 lợn nuôi vỗ béo giống Landrace được rút ngẫu nhiên từ một trại chăn nuôi. Số liệu thu được như sau:

577	596	594	612	600	584	618	627	588	601	606	559	615	607	608	591	565	586
621	623	598	602	581	631	570	595	603	605	616	574	578	600	596	619	636	589

Cán bộ kỹ thuật trại cho rằng tăng trọng trung bình của toàn đàn lợn trong trại là 607 gram/ngày. Theo anh chị kết luận đó đúng hay sai, vì sao?

2.7.2

Anh chị hãy kiểm tra kết luận với bài tập tương tự như 2.7.1, biết rằng độ lệch chuẩn của tính trạng này ở Landrace là 24 gram/ngày.

2.7.3

Tỷ lệ thụ thai bằng thụ tinh nhân tạo từ tinh trùng của 2 bò đực giống được xác định trên nhóm bò cái gồm 50 con; 18 nhóm bò cái sử dụng tinh trùng của bò đực A và 16 đối với bò đực B. Tỷ lệ thụ thai (%) thu được như sau:

Bò đực A	74,2	62,1	57,7	71,7	62,0	76,1	70,6	68,3	68,4	79,8
	71,1	70,9	65,5	61,2	60,8	73,9	51,9	63,7		
Bò đực B	49,6	49,2	53,2	56,5	69,1	54,2	80,7	62,7	71,5	67,5
	64,6	75,4	79,6	59,8	68,8	60,2				

Hãy cho biết tỷ lệ thụ thai của 2 bò đực nêu trên.

2.7.4

Nồng độ fructoza (mg%) trong tinh dịch bò trước và sau khi ủ được xác định trên 12 mẫu tinh bò đực; các giá trị thu được như sau:

Mẫu số	1	2	3	4	5	6	7	8	9	10	11	12
Trước khi ủ	116	190	570	375	236	505	120	322	429	102	167	299
Sau khi ủ	30	58	100	48	58	153	54	66	67	34	69	82

Kết luận về nồng độ fructoza trong tinh dịch bò trước và sau khi ủ.

2.7.5

Một thí nghiệm được tiến hành nhằm nghiên cứu ảnh hưởng của progesterone lên chu kỳ động dục của cừu Merino. Sử dụng 4 liều khác nhau (0, 10, 25 và 40 mg/ngày) tiêm dưới da liên tục trong 4 ngày tính từ ngày động dục. Chu kỳ động dục (ngày) của 8 cừu trong mỗi nhóm thu được như sau:

Liều 0 mg/ngày	18	14	18	18	18	18	18	19
Liều 10 mg/ngày	15	14	17	14	12	13	12	13
Liều 25 mg/ngày	11	13	11	11	12	11	11	12
Liều 40 mg/ngày	9	10	12	10	11	11	10	11

Cho biết ảnh hưởng của progesterone lên chu kỳ động dục ở cừu Merino.