Chương 4

Một số phân phối xác suất thông dụng

4.1 Phân phối Bernoulli, nhị thức

Bài 4.1. Có 8000 sản phẩm trong đó có 2000 sản phẩm không đạt tiêu chuẩn kỹ thuật. Lấy ngẫu nhiên (không hoàn lại) 10 sản phẩm. Tính xác suất để trong 10 sản phẩm lấy ra có 2 sản phẩm không đạt tiêu chuẩn.

Đáp án. 0.282 ■

Hướng dẫn. Gọi X là số sản phẩm không đạt tiêu chuẩn trong 10 sản phẩm lấy ra.

Ta có, $X \sim B(10, \frac{2000}{8000}) = B(10, 0.25)$ ■

Bài 4.2. Khi tiêm truyền một loại huyết thanh, trung bình có một trường hợp phản ứng trên 1000 trường hợp. Dùng loại huyết thanh này tiêm cho 2000 người. Tính xác suất để

(a) có 3 trường hợp phản ứng,

(b) có nhiều nhất 3 trường hợp phản ứng,

(c) có nhiều hơn 3 trường hợp phản ứng.

Đáp án. (a) 0.18 (b) 0.86 (c) 0.14 ■

Bài 4.3. Giả sử tỷ lệ sinh con trai và con gái là bằng nhau và bằng $\frac{1}{2}$. Một gia đình có 4 người con. Tính xác suất để 4 đứa con đó gồm

(a) 2 trai và 2 gái.

(b) 1 trai và 3 gái.

(c) 4 trai.

Đáp án. (a) 0.375 (b) 0.25 (c) 0.0625 ■

Bài 4.4. Một nhà máy sản xuất với tỷ lệ phế phẩm là 7%.

(a) Quan sát ngẫu nhiên 10 sản phẩm. Tính xác suất để

i) có đúng một phế phẩm.

ii) có ít nhất một phế phẩm.

iii) có nhiều nhất một phế phẩm.

(b) Hỏi phải quan sát ít nhất bao nhiêu sản phẩm để xác suất nhận được ít nhất một phế phẩm ≥ 0.9

Đáp án. (a)-(i) 0.364 -(ii) 0.516 -(iii) 0.848 (b) 32 ∎

Bài 4.5. Tỷ lệ một loại bệnh bẩm sinh trong dân số là $p = 0.01$. Bệnh này cần sự chăm sóc đặc biệt lúc mới sinh. Một nhà bảo sinh thường có 20 ca sinh trong một tuần. Tính xác suất để

(a) không có trường hợp nào cần chăm sóc đặc biệt,

(b) có đúng một trường hợp cần chăm sóc đặc biệt,

(c) có nhiều hơn một trường hợp cần chăm sóc đặc biệt.

Tính bằng quy luật nhị thức rồi dùng quy luật Poisson để so sánh kết quả khi ta xấp xỉ phân phối nhị thức $B(n; p)$ bằng phân phối Poisson $P(np)$.

Đáp án. (a) 0.818 (b) 0.165 (c) 0.017 ∎

Bài 4.6. Tỷ lệ cử tri ủng hộ ứng cử viên A trong một cuộc bầu cử là 60%. Người ta hỏi ý kiến 20 cử tri được chọn một cách ngẫu nhiên. Gọi X là số người bỏ phiếu cho A trong 20 người đó.

(a) Tìm giá trị trung bình, độ lệch chuẩn và Mod của X.

(b) Tìm $P(X \leq 10)$

(c) Tìm $P(X > 12)$

(d) Tìm $P(X = 11)$

Đáp án. (a) 12; 2.191; 12 (b) 0.245 (c) 0.416 (d) 0.16 ∎

Bài 4.7. Giả sử tỷ lệ dân cư mắc bệnh A trong vùng là 10%. Chọn ngẫu nhiên 1 nhóm 400 người.

(a) Viết công thức tính xác suất để trong nhóm có nhiều nhất 50 người mắc bệnh A.

(b) Tính xấp xỉ xác suất đó bằng phân phối chuẩn.

Đáp án. (b) 0.953 ∎

Bài 4.8. Một chiếc máy bay muốn bay được thì phải có ít nhất một nửa số động cơ hoạt động. Nếu mỗi động cơ hoạt động, độc lập nhau, với xác suất 0.6, thì một máy bay có 4 động cơ có đáng tin cậy hơn một máy bay có 2 động cơ hay không? Giải thích?

Đáp án. không ∎

Hướng dẫn. Gọi X, Y lần lượt là số động cơ hoạt động trong 4 động cơ và trong 2 động cơ. So sánh $P(X \geq 2)$ và $P(Y \geq 1)$. ∎

Bài 4.9. Số lượng X các phân tử phát ra từ một nguồn phóng xạ nào đó trong 1 giờ là một biến ngẫu nhiên có phân phối Poisson với tham số $\lambda = \ln 5$. Hơn nữa, ta giả sử rằng sự phát xạ này độc lập nhau qua mỗi giờ.

(a) Tính xác suất có ít nhất 30 giờ, trong 168 giờ của một tuần nào đó, không có phân tử nào được phát ra.

(b) Sử dụng phân phối Poisson để tính xấp xỉ xác suất trong câu (a).

Đáp án. (a) 0.7549 (b) 0.7558 ∎

Bài 4.10. Một máy sản xuất ra sản phẩm loại A với xác suất 0.485. Tính xác suất sao có trong 200 sản phẩm do máy sản xuất ra có ít nhất 95 sản phẩm loại A.

Đáp án. 0.6368 ∎

Hướng dẫn. Sử dụng xấp xỉ chuẩn và hiệu chỉnh liên tục. ∎

Bài 4.11. Dựa vào số liệu trong quá khứ, ta ước lượng rằng 85% các sản phẩm của một máy sản xuất nào đó là thứ phẩm. Nếu máy này sản xuất 20 sản phẩm mỗi giờ, thì xác suất 8 hoặc 9 thứ phẩm được sản xuất trong mỗi khoảng thời gian 30 phút là bao nhiêu?

Đáp án. 0.6233 ∎

Bài 4.12. Mười mẫu có kích thước 10 được rút ra ngẫu nhiên và không hoàn lại từ các thùng chứa 100 sản phẩm, trong mỗi thùng có 2 phế phẩm. Một thùng sản phẩm được chấp nhận nếu có nhiều nhất một thứ phẩm được phát hiện trong mẫu tương ứng. Hỏi xác suất có ít hơn chín thùng được chấp nhận là bao nhiêu?

Đáp án. 0.0036 ∎

Bài 4.13. Xác suất để một sản phẩm được sản xuất bởi một máy nào đó phù hợp với các yêu cầu kĩ thuật là 0.95, độc lập với các sản phẩm khác. Ta tiến hành lấy ra các sản phẩm được sản xuất bởi máy này cho đến khi được sản phẩm đạt các yêu cầu kĩ thuật. Thí nghiệm ngẫu nhiên này được lặp lại trong 15 ngày liên tiếp (độc lập). Gọi X là số ngày, trong 15 ngày thí nghiệm, mà ta phải lấy ít nhất 2 sản phẩm để nhận được một sản phẩm phù hợp với các yêu cầu kĩ thuật.

(a) Tìm giá trị trung bình của X.

(b) Sử dụng phân phối Poisson để tính xấp xỉ xác suất có điều kiện $P(X = 2|X \geq 1)$.

Đáp án. (a) 0.75 (b) 0.2519 ∎

Bài 4.14. Xác suất trúng số là 1%. Mỗi tuần mua một vé số. Hỏi phải mua vé số liên tiếp trong tối thiểu bao nhiêu tuần để có không ít hơn 95% hy vọng trúng số ít nhất 1 lần.

Đáp án. 299 ∎

Bài 4.15. Trong trò chơi "bầu cua" có ba con xúc sắc, mỗi con có sáu mặt hình là: bầu, cua, hưu, nai, tôm và gà. Giả sử có hai người, một người chơi và một người làm cái. Nếu mỗi ván người chơi chỉ đặt ở một ô (một trong các hình: bầu, cua, hưu, nai, tôm và gà) sau khi chơi nhiều ván thì người nào sẽ thắng trong trò chơi này. Giả sử thêm mỗi ván người chơi đặt 1000 đ nếu thắng sẽ được 5000 đ, nếu thua sẽ mất 1000 đ. Hỏi trung bình mỗi ván người thắng sẽ thắng bao nhiêu?

Đáp án. 972.2222 ∎

Bài 4.16. Có ba lọ giống nhau: hai lọ loại I, mỗi lọ có 3 bi trắng và 7 bi đen; một lọ loại II có 4 bi trắng và 6 bi đen. Một trò chơi được đặt ra như sau: Mỗi ván, người chơi chọn ngẫu nhiên một lọ và lấy ra hai bi từ lọ đó. Nếu lấy được đúng hai bi trắng thì người chơi thắng, ngược lại người chơi thua.

(a) Người A chơi trò chơi này, tính xác suất người A thắng ở mỗi ván.

(b) Giả sử người A chơi 10 ván, tính số ván trung bình người chơi thắng được và số ván người A thắng tin chắc nhất.

(c) Người A phải chơi ít nhất bao nhiêu ván để xác suất thắng ít nhất một ván không dưới 0,99.

Đáp án. (a) 0.0889 (b) 0.889; 0 (c) 50 ∎

Bài 4.17 (*). Cho X và Y là hai đại lượng ngẫu nhiên độc lập.

(a) Giả sử $X \sim B(1, \frac{1}{5})$, $Y \sim B(2, \frac{1}{5})$. Lập bảng phân phối xác suất của $X + Y$ và kiểm tra rằng $X + Y \sim B(3, \frac{1}{5})$

(b) Giả sử $X \sim B(1, \frac{1}{2})$, $Y \sim B(2, \frac{1}{5})$. Tìm phân bố xác suất của $X + Y$. Chứng minh rằng $X + Y$ không có phân bố nhị thức.

(c) Giả sử $X \sim B(n_1, p_1)$, $Y \sim B(n_2, p_2)$. Chứng minh rằng $X + Y$ có phân phối nhị thức khi và chỉ khi $p_1 = p_2$.

Bài 4.18. Hai cầu thủ ném bóng vào rổ. Cầu thủ thứ nhất ném hai lần với xác suất trúng rổ của mỗi lần là 0.6. Cầu thủ thứ hai ném một lần với xác suất trúng rổ là 0.7. Gọi X là số lần trúng rổ của cả hai cầu thủ. Lập bảng phân phối xác suất của X, biết rằng kết quả của các lần ném rổ là độc lập với nhau.

Đáp án.

X	0	1	2	3
$\mathbb{P}$	0.048	0.256	0.444	0.252

∎

Bài 4.19. Bưu điện dùng một máy tự động đọc địa chỉ trên bì thư để phân loại từng khu vực gởi đi, máy có khả năng đọc được 5000 bì thư trong 1 phút. Khả năng đọc sai 1 địa chỉ trên bì thư là 0,04% (xem như việc đọc 5000 bì thư này là 5000 phép thử độc lập).

(a) Tính số bì thư trung bình mỗi phút máy đọc sai.

(b) Tính số bì thư tin chắc nhất trong mỗi phút máy đọc sai.

(c) Tính xác suất để trong một phút máy đọc sai ít nhất 3 bì thư.

Đáp án. (a) 2 (b) 2 (c) 0.323 ∎

Bài 4.20. Một bài thi trắc nghiệm gồm có 10 câu hỏi, mỗi câu có 4 phương án trả lời, trong đó chỉ có một phương án đúng. Giả sử mỗi câu trả lời đúng được 4 điểm và câu trả lời sai bị trừ 2 điểm. Một sinh viên kém làm bài bằng cách chọn ngẫu nhiên một phương án cho mỗi câu hỏi.

(a) Tính xác suất để học sinh này được 4 điểm.

(b) Tính xác suất để học sinh này bị điểm âm.

(c) Gọi X là số câu trả lời đúng, tính $E(X)$ và $Var(X)$.

(d) Tính số câu sinh viên này có khả năng trả lời đúng lớn nhất.

Đáp án. (a) 0.146 (b) 0.2503 (c) 2.5; 1.875 (d) 2 ∎

Bài 4.21. Các sản phẩm được sản xuất trong một dây chuyền. Để thực hiện kiểm tra chất lượng, mỗi giờ người ta rút ngẫu nhiên không hoàn lại 10 sản phẩm từ một hộp có 25 sản phẩm. Quá trình sản xuất được báo cáo là đạt yêu cầu nếu có không quá một sản phẩm là thứ phẩm.

(a) Nếu tất cả các hộp được kiểm tra đều chứa chính xác hai thứ phẩm, thì xác suất quá trình sản xuất được báo cáo đạt yêu cầu ít nhất 7 lần trong một ngày làm việc 8 giờ là bao nhiêu?

(b) Sử dụng phân phối Poisson để xấp xỉ xác suất được tính trong câu (a).

(c) Biết rằng lần kiểm tra chất lượng cuối cùng trong câu (a), quá trình sản xuất được báo cáo đạt yêu cầu. Hỏi xác suất mẫu 10 sản phẩm tương ứng không chứa thứ phẩm là bao nhiêu?

Đáp án. (a) 0.6572 (b) 0.6626 (c) 0.4118 ∎

Bài 4.22. Một công ty bảo hiểm có 20 nhân viên kinh doanh. Mỗi người, tại một thời điểm nào đó, có thể ở văn phòng hoặc đang trên đường giao dịch. Biết rằng nhân viên kinh doanh làm việc ở văn phòng vào lúc 14h30, vào một ngày làm việc trong tuần, với xác suất là 0.2, độc lập với các ngày làm việc khác và những nhân viên khác.

(a) Công ty muốn bố trí một số lượng ít nhất các bàn làm việc sao cho một nhân viên kinh doanh bất kì có thể tìm thấy một bàn trống để làm việc trong ít nhất 90% trường hợp. Tìm số lượng bàn ít nhất này.

(b) Tính số lượng bàn ít nhất trong phần (a) bằng cách sử dụng xấp xỉ Poisson.

(c) Một người phụ nữ đã gọi điện đến công ty vào lúc 14h30 vào 2 ngày làm việc cuối cùng trong tuần để nói chuyện với một nhân viên kinh doanh nào đó. Cho rằng cô ta không sắp xếp cuộc hẹn từ trước. Tìm xác suất cô ta phải gọi ít nhất hai lần nữa với giả sử rằng cô ta luôn gọi vào 14h30.

Đáp án. (a) 6 (b) 7 (c) 0.8 ■

4.2 Phân phối Poisson

Bài 4.23. Một trung tâm bưu điện nhận được trung bình 3 cuộc điện thoại trong mỗi phút. Tính xác suất để trung tâm này nhận được 1 cuộc, 2 cuộc, 3 cuộc gọi trong 1 phút, biết rằng số cuộc gọi trong một phút có phân phối Poisson.

Đáp án. 0.149; 0.224; 0.224 ■

Bài 4.24. Tính $P(X \geq 1 | X \leq 1)$ nếu $X \sim P(5)$

Đáp án. 5/6 ■

Bài 4.25 (*). Cho X, Y là các biến ngẫu nhiên độc lập, $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$

(a) Tính xác suất $P(X + Y = n)$

(b) Tính xác suất $P(X = k | X + Y = n)$

Bài 4.26. Một cửa hàng cho thuê xe ôtô nhận thấy rằng số người đến thuê xe ôtô vào ngày thứ bảy cuối tuần là một đại lượng ngẫu nhiên X có phân phối Poisson với tham số $\lambda = 2$. Giả sử cửa hàng có 4 chiếc ôtô.

(a) Tìm xác suất không phải tất cả 4 chiếc ôtô đều được thuê.

(b) Tìm xác suất tất cả 4 chiếc ôtô đều được thuê.

(c) Tìm xác suất cửa hàng không đáp ứng được yêu cầu.

(d) Trung bình có bao nhiêu ôtô được thuê.

(e) Cửa hàng cần có ít nhất bao nhiêu ôtô để xác suất không đáp ứng được nhu cầu thuê bé hơn 2%

Đáp án. (a) 0.857 (b) 0.143 (c) 0.053 (d) 2 (e) 5 ■

Bài 4.27. Một tổng đài bưu điện có các cuộc điện thoại gọi đến xuất hiện ngẫu nhiên, độc lập với nhau và có tốc độ trung bình 2 cuộc gọi trong 1 phút. Tìm xác suất để

(a) có đúng 5 cuộc điện thoại trong 2 phút,

(b) không có cuộc điện thoại nào trong khoảng thời gian 30 giây,

(c) có ít nhất 1 cuộc điện thoại trong khoảng thời gian 10 giây.

Đáp án. (a) 0.156 (b) 0.368 (c) 0.283 ∎

Bài 4.28. Các cuộc gọi điện đến tổng đài tuân theo phân phối Poisson với mức λ trên mỗi phút. Từ kinh nghiệm có được trong quá khứ, ta biết rằng xác suất nhận được chính xác một cuộc gọi trong một phút bằng ba lần xác suất không nhận được cuộc gọi nào trong cùng thời gian.

(a) Gọi X là số cuộc gọi nhận được trong mỗi phút. Tính xác suất $P(2 \leq X \leq 4)$.

(b) Ta xét 100 khoảng thời gian một phút liên tiếp và gọi U là số khoảng thời gian một phút không nhận được cuộc gọi điện nào. Tính $P(U \leq 1)$.

Đáp án. (a) 0.6161 (b) 0.0377 ∎

Hướng dẫn. $U \sim B(100, 0.0498)$ ∎

Bài 4.29. Tại một điểm bán vé máy bay, trung bình trong 10 phút có 4 người đến mua vé. Tính xác suất để:

(a) Trong 10 phút có 7 người đến mua vé.

(b) Trong 10 phút có không quá 3 người đến mua vé.

Đáp án. (a) 0.06 (b) 0.433 ∎

Bài 4.30. Các khách hàng đến quầy thu ngân, theo phân phối Poisson, với số lượng trung bình 5 người mỗi phút. Tính xác suất xuất hiện ít nhất 10 khách hàng trong khoảng thời gian 3 phút.

Đáp án. 0.9301 ∎

Bài 4.31. Số khách hàng đến quầy thu ngân tuân theo phân phối Poisson với tham số $\lambda = 1$ trong mỗi khoảng 2 phút. Tính xác suất thời gian đợi đến khi khách hàng tiếp theo xuất hiện (từ khách hàng trước đó) nhỏ hơn 10 phút.

Đáp án. 0.9933 ∎

Bài 4.32. Số lượng nho khô trong một cái bánh quy bất kì có phân phối Poisson với tham số λ. Hỏi giá trị λ là bao nhiêu nếu ta muốn xác suất có nhiều nhất hai bánh quy, trong một hộp có 20 bánh, không chứa nho khô là 0.925?

Đáp án. 2.9977 ∎

Bài 4.33. Một trạm cho thuê xe Taxi có 3 chiếc xe. Hàng ngày trạm phải nộp thuế 8 USD cho 1 chiếc xe (bất kể xe đó có được thuê hay không). Mỗi chiếc được cho thuê với giá 20USD. Giả sử số xe được yêu cầu cho thuê của trạm trong 1 ngày là đại lượng ngẫu nhiên có phân phối Poisson với $\mu = 2.8$.

(a) Tính số tiền trung bình trạm thu được trong một ngày.

(b) Giải bài toán trên trong trường hợp trạm có 4 chiếc xe.

(c) Theo bạn, trạm nên có 3 hay 4 chiếc xe?

Đáp án. (a) 32 USD (b) 24 USD (c) 3 ∎

Bài 4.34 (*). Ta có 10 máy sản xuất (độc lập nhau), mỗi máy sản xuất ra 2% thứ phẩm (không đạt chuẩn).

(a) Trung bình có bao nhiêu sản phẩm được sản xuất bởi máy đầu tiên trước khi nó tạo ra thứ phẩm đầu tiên?

(b) Ta lấy ngẫu nhiên một sản phẩm từ mỗi máy sản xuất. Hỏi xác suất nhiều nhất hai thứ phẩm trong 10 sản phẩm này là bao nhiêu?

(c) Làm lại câu (b) bằng cách sử dụng xấp xỉ Poisson.

(d) Phải lấy ra ít nhất bao nhiêu sản phẩm được sản xuất bởi máy đầu tiên để xác suất đạt được ít nhất một thứ phẩm không nhỏ hơn 1/2 (giả sử rằng các sản phẩm là độc lập với nhau)?

Đáp án. (a) 49 (b) 0.9991 (c) 0.9989 (d) 35 ∎

Bài 4.35. Số lỗi đánh máy trong một quyển sách 500 trang có phân phối Poisson với tham số $\lambda = 2$ mỗi trang, độc lập trên từng trang.

(a) Hỏi xác suất phải lấy ít nhất 10 trang, ngẫu nhiên và có hoàn lại, để đạt được 3 trang trong đó mỗi trang chứa ít nhất 2 lỗi là bao nhiêu?

(b) Giả sử rằng thật sự có 20 trang, trong 500 trang, mỗi trang chứa chính xác 5 lỗi.

   (i) Nếu 100 trang được lấy, ngẫu nhiên và không hoàn lại, thì xác suất nhiều nhất 5 trang chứa chính xác 5 lỗi mỗi trang là bao nhiêu?

   (ii) Ta xét 50 bản sao của quyển sách này. Nếu thí nghiệm ngẫu nhiên trong phần (i) được lặp lại cho mỗi bản sao, thì xác suất có chính xác 30 trong 50 bản sao mà mẫu lấy ra có nhiều nhất 5 trang với 5 lỗi mỗi trang là bao nhiêu?

Đáp án. (a) 0.0273 (b)-(i) 0.8083 -(ii) 0.000357 ∎

4.3 Phân phối chuẩn

Bài 4.36. Các kết quả của bài kiểm tra chỉ số thông minh (IQ) cho các học sinh của một trường tiểu học cho thấy điểm IQ của các học sinh này tuân theo phân phối chuẩn với các tham số là $\mu = 100$ và $\sigma^2 = 225$. Tỉ lệ học sinh có điểm IQ nhỏ hơn 91 hoặc lớn hơn 130 là bao nhiêu?

Đáp án. 0.2971 ∎

Bài 4.37. Giả sử chiều dài X (đơn vị tính m) của một nơi đỗ xe bất kì tuân theo phân phối chuẩn $N(\mu, 0.01\mu^2)$.

(a) Một người đàn ông sở hữu một chiếc xe hơi cao cấp có chiều dài lớn hơn 15% chiều dài trung bình của một chỗ đậu xe. Hỏi tỉ lệ chỗ đậu xe có thể sử dụng là bao nhiêu?

(b) Giả sử rằng $\mu = 4$. Hỏi chiều dài của xe là bao nhiêu nếu ta muốn chủ của nó có thể sử dụng 90% chỗ đậu xe?

Đáp án. (a) 0.0668 (b) ≈ 3.49 ∎

Bài 4.38. Đường kính của một chi tiết máy do một máy tiện tự động sản xuất có phân phối chuẩn với trung bình $\mu = 50$ mm và độ lệch chuẩn $\sigma = 0.05$ mm. Chi tiết máy được xem là đạt yêu cầu nếu đường kính không sai quá 0.1 mm.

(a) Tính tỷ lệ sản phẩm đạt yêu cầu.

(b) Lấy ngẫu nhiên 3 sản phẩm. Tính xác suất có ít nhất một sản phẩm đạt yêu cầu.

Đáp án. (a) 95.4% (b) 0.999 ∎

Bài 4.39. Trọng lượng X (tính bằng gam) một loại trái cây có phân phối chuẩn $N(\mu, \sigma^2)$, với $\mu = 500$ (gam) và $\sigma^2 = 16$ (gam^2). Trái cây thu hoạch được phân loại theo trọng lượng như sau:

(a) loại 1 : trên 505 gam,

(b) loại 2 : từ 495 đến 505 gam,

(c) loại 3 : dưới 495 gam.

Tính tỷ lệ mỗi loại.

Đáp án. (a) 0.106 (b) 0.788 (c) 0.106 ∎

Bài 4.40. Một công ty kinh doanh mặt hàng A dự định sẽ áp dụng một trong 2 phương án kinh doanh. Ký hiệu X_1 là lợi nhuận thu được khi áp dụng phương án thứ 1, X_2 là lợi nhuận thu được khi áp dụng phương án thứ 2. X_1, X_2 đều được tính theo đơn vị triệu đồng/ tháng) và $X_1 \sim N(140, 2500)$, $X_2 \sim N(200, 3600)$. Nếu biết rằng, để công ty tồn tại và phát triển thì lợi nhuận thu được từ mặt hàng kinh doanh A phải đạt ít nhất 80 triệu đồng/tháng. Hãy cho biết công ty nên áp dụng phương án nào để kinh doanh mặt hàng A? Vì sao?

Đáp án. phương án 2 ■

Bài 4.41. Nghiên cứu chiều cao của những người trưởng thành, người ta nhận thấy rằng chiều cao đó tuân theo quy luật phân bố chuẩn với trung bình là 175 cm và độ lệch tiêu chuẩn 4 cm. Hãy xác định:

(a) tỷ lệ người trưởng thành có tầm vóc trên 180 cm.

(b) tỷ lệ người trưởng thành có chiều cao từ 166 cm đến 177 cm.

(c) tìm h_0, nếu biết rằng 33% người trưởng thành có tầm vóc dưới mức h_0.

(d) giới hạn biến động chiều cao của 90% người trưởng thành xung quanh giá trị trung bình của nó.

Đáp án. (a) 0.106 (b) 0.68 (c) 173.24 (d) 6.6 ■

Bài 4.42. Ta quan tâm đến tuổi thọ X (theo năm) của một thiết bị. Từ kinh nghiệm trong quá khứ, ta ước lượng xác suất thiết bị loại này còn hoạt động tốt sau 9 năm là 0.1.

(a) Ta đưa ra mô hình sau cho hàm mật độ của X

$$f_X(x) = \frac{a}{(x+1)^b} \quad \text{với} \quad x \geq 0$$

trong đó $a > 0$ và $b > 1$. Tìm hai hằng số a, b.

(b) Nếu ta đưa ra một phân phối chuẩn với trung bình $\mu = 7$ cho X, thì giá trị tham số σ là bao nhiêu?

(c) Ta xét 10 thiết bị loại này một cách độc lập. Tính xác suất 8 hoặc 9 thiết bị loại này có tuổi đời hoạt động ít hơn 9 năm.

Đáp án. (a) 1; 2 (b) 1.5601 (c) 0.5811 ■

Bài 4.43. Entropy H của một biến ngẫu nhiên liên tục X được định nghĩa là $H = E[-\ln f_X(X)]$ với f_X là hàm mật độ xác suất của biến ngẫu nhiên X và ln là logarit tự nhiên. Tính entropy của biến ngẫu nhiên Gauss với trung bình 0 và phương sai $\sigma^2 = 2$.

Đáp án. ≈ 1.766 ■

Bài 4.44 (*). Một nhà sản xuất bán sản phẩm với một mức giá cố định s. Nhà sản xuất sẽ hoàn lại tiền cho khách hàng nếu khách hàng phát hiện trọng lượng sản phẩm nhỏ hơn trọng lượng cho trước w_0 và thu lại sản phẩm, có giá trị tái chế là $r(< s)$. Trọng lượng W tuân theo phân phối chuẩn với trung bình μ và phương sai σ^2. Một cài đặt thích hợp cho phép nhà sản xuất cố định giá trị μ bằng một giá trị mong muốn, nhưng không thể cố định giá trị σ. Chi phí sản xuất C là một hàm theo trọng lượng của sản phẩm: $C = \alpha + \beta W$, với α và β là các hằng số dương.

(a) Hãy xác định biểu thức cho lợi nhuận Z theo W.

(b) Chứng minh rằng lợi nhuận trung bình, $z(\mu)$, được xác định bởi

$$z(\mu) = s - \alpha - \beta\mu - (s - r)P[W < w_0]$$

Tìm giá trị μ_0 của μ làm cực đại $z(\mu)$.

Chương 6

Dữ liệu thống kê

Bài 6.1. Một nhà cổ sinh vật học đo chiều rộng (theo mm) của răng hàm trên cùng trong 36 mẫu vật của một loài động vật có vú đã tuyệt chủng *Acropithecus rigidus*. Các kết quả được ghi lại như sau:

6.1	5.7	6.0	6.5	6.0	5.7
6.1	5.8	5.9	6.1	6.2	6.0
6.3	6.2	6.1	6.2	6.0	5.7
6.2	5.8	5.7	6.3	6.2	5.7
6.2	6.1	5.9	6.5	5.4	6.7
5.9	6.1	5.9	5.9	6.1	6.1

(a) Xây dựng phân phối tần số và biểu diễn dưới dạng bảng và đồ thị tổ chức tần số (histogram).

(b) Mô tả hình dạng của phân phối.

Hướng dẫn. Chọn 7 khoảng và độ rộng mỗi khoảng là 0.2. ∎

Bài 6.2. Trong một nghiên cứu về bệnh tâm thần phân liệt, các nhà nghiên cứu đã đo hoạt động của enzyme monoamine oxidase (MAO) trong các tiểu cầu của 18 bệnh nhân. Các kết quả (là trung bình số hợp chất nmoles benzylaldehyde trong mỗi 108 tiểu cầu) được ghi lại như sau:

6.8	8.4	8.7	11.9	14.2	18.8
9.9	4.1	9.7	12.7	5.2	7.8
7.8	7.4	7.3	10.6	14.5	10.7

Xây dựng phân phối tần số và biểu diễn dưới dạng bảng và đồ thị tổ chức tần số (histogram).

Hướng dẫn. Chọn 8 khoảng và độ rộng mỗi khoảng là 2. ∎

Bài 6.3. Hãy tạo ra một mẫu kích thước 5 có trung bình là 20 và không phải tất cả các quan sát đều bằng nhau.

Bài 6.4. Hãy tạo ra một mẫu kích thước 5 có trung bình là 20 và trung vị mẫu là 15.

Bài 6.5. Sự tăng trọng của các con bò được đo mỗi chu kì 140 ngày. Trọng lượng tăng trung bình hằng ngày (lb/ngày) của 9 con bò có cùng chế độ ăn uống giống nhau như sau:

3.89	3.51	3.97	3.31	3.21
3.36	3.67	3.24	3.27	

(a) Xác định trung bình và trung vị.

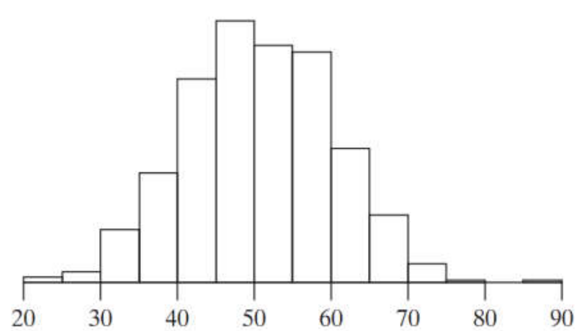
(b) Giả sử rằng một quan sát khác là 2.46 được thêm vào mẫu. Tìm trung bình và trung vị của 10 quan sát này.

Bài 6.6. Bảng kèm theo cung cấp lứa đẻ (số lợn con sống sau 21 ngày) của 36 con heo nái. Xác định trung vị lứa đẻ.

Hướng dẫn. Chú ý rằng có một con đẻ 5 lứa, nhưng có hai con đẻ 7 lứa, ba con đẻ 8 lứa,...

Số heo con	Tần số (số heo nái)
5	1
6	0
7	2
8	3
9	3
10	9
11	8
12	5
13	3
14	2
Tổng	36

Bài 6.7. Dưới đây là đồ thị histogram.

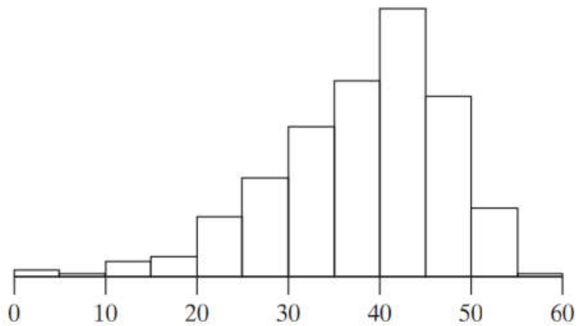


(a) Ước lượng trung vị của phân phối.

(b) Ước lượng trung bình của phân phối.

Bài 6.8. Xét đồ thị histogram từ bài tập 6.7. Bằng cách "đọc" đồ thị, hãy ước lượng phần trăm số quan sát nhỏ hơn 40. Phần trăm này gần nhất với 15%, 25%, 35%, hay 45%? (Chú ý: tần số không được cho trong đồ thị này, vì không cần phải tính số quan sát cho mỗi khoảng. Thay vào đó, phần trăm số quan sát nhỏ hơn 40 có thể được ước lượng bằng cách nhìn vào diện tích.)

Bài 6.9. Dưới đây là đồ thị histogram



(a) Ước lượng trung vị của phân phối.

(b) Ước lượng trung bình của phân phối.

Bài 6.10. Xét đồ thị histogram từ bài tập 6.9. Bằng cách "đọc" đồ thị, hãy ước lượng phần trăm số quan sát lớn hơn 45. Phần trăm này gần nhất với 15%, 25%, 35%, hay 45%? (Chú ý: tần số không được cho trong đồ thị này, vì không cần phải tính số quan sát cho mỗi khoảng. Thay vào đó, phần trăm số quan sát lớn hơn 45 có thể được ước lượng bằng cách nhìn vào diện tích.)

Bài 6.11. Ở đây là 18 độ đo về hoạt động MAO được trình bày trong bài tập 6.2:

6.8	8.4	8.7	11.9	14.2	18.8
9.9	4.1	9.7	12.7	5.2	7.8
7.8	7.4	7.3	10.6	14.5	10.7

(a) Xác định trung vị và tứ phân vị.

(b) Xác định khoảng tứ phân vị.

(c) Tìm giá trị outlier nhỏ nhất trong tập dữ liệu.

(d) Xây dựng một đồ thị boxplot của dữ liệu.

Bài 6.12. Trong một nghiên cứu về sản lượng sữa ở cừu (dùng để làm pho mát), một nhà nghiên cứu đã đo sản lượng sữa trong 3 tháng của 11 con cừu cái. Các sản lượng (lít) được cho như sau:

56.5	89.8	110.1	65.6	63.7	82.6
75.1	91.5	102.9	44.4	108.1	

(a) Xác định trung vị và tứ phân vị.

(b) Xác định khoảng tứ phân vị.

(c) Xây dựng đồ thị boxplot của dữ liệu.

Bài 6.13. Tính độ lệch chuẩn của các mẫu giả sau:

(a) 16, 13, 18, 13

(b) 38, 30, 34, 38, 35

(c) 1, -1, 5, -1

(d) 4, 6, -1, 4, 2

Bài 6.14. (a) Tạo một mẫu giả có kích thước 5 có các độ lệch $y_i - \bar{y}$ là $-3, -1, 0, 2, 2$.

(b) Tính độ lệch chuẩn của mẫu vừa tạo.

(c) Có phải mọi người đều nhận cùng kết quả trong câu (b)? Tại sao?

Bài 6.15. Bốn lô đất, mỗi lô 346 feet vuông, được trồng cùng một giống ("Beau") của lúa mì. Sản lượng lô (lb) như sau:

$$35.1, 30.6, 36.9, 29.8$$

(a) Tính trung bình và độ lệch chuẩn.

(b) Tính hệ số biến thiên.

Bài 6.16. Dopamine là một chất hóa học đóng một vai trò trong việc truyền tải các tín hiệu trong não. Một dược sĩ đo lượng dopamine trong não của bảy con chuột. Mức độ dopamine (nmoles/g) như sau:

$$6.8, 5.3, 6.0, 5.9, 6.8, 7.4, 6.2$$

(a) Tính trung bình và độ lệch chuẩn.

(b) Xác định trung vị và khoảng tứ phân vị.

(c) Tính hệ số biến thiên.

(d) Thay quan trắc 7.4 bằng 10.4 và lặp lại câu (a) và (b). Các độ đo nào thay đổi và không thay đổi?

Bài 6.17. Trong một nghiên cứu về loài thằn lằn *Sceloporus occidentalis*, các nhà sinh học đã đo khoảng cách (m) chạy trong hai phút của 15 con. Kết quả (được liệt kê theo thứ tự tăng) như sau:

18.4	22.2	24.5	26.4	27.5	28.7	30.6	32.9
32.9	34.0	34.8	37.5	42.1	45.5	45.5	

(a) Xác định tứ phân vị và khoảng tứ phân vị.

(b) Xác định miền giá trị.

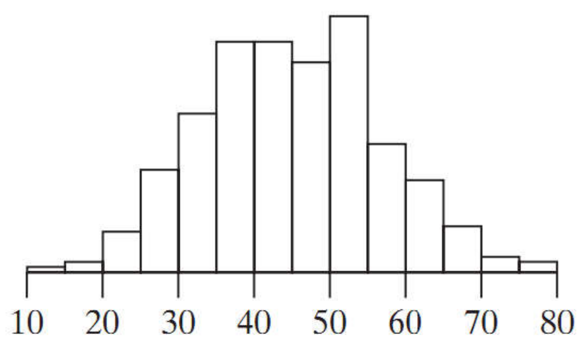
Bài 6.18. Sử dụng dữ liệu khoảng cách chạy trong bài 6.17. Trung bình mẫu là 32.23 m và độ lệch chuẩn mẫu (s) 8.07 m. Phần trăm số quan sát nằm trong

(a) 1 s so với trung bình?

(b) 2 s so với trung bình?

Bài 6.19. So sánh các kết quả của bài tập 6.18 với các dự đoán của quy tắc thực nghiệm.

Bài 6.20. Dưới đây là histogram. Hãy ước lượng trung bình và độ lệch chuẩn của phân phối



Bài 6.21. Một nhà sinh học đo độ pH nào đó ở 24 con ếch; các giá trị tiêu biểu là

$$7.43, 7.16, 7.51, \ldots$$

Cô ta tính được trung bình là 7.373 và độ lệch chuẩn 0.129 cho các độ đo pH gốc này. Tiếp theo, cô ta biến đổi dữ liệu bằng cách trừ 7 mỗi quan trắc và sau đó nhân với 100. Ví dụ, 7.43 được biến đổi thành 43. Dữ liệu được biến đổi là

$$43, 16, 51, \ldots$$

Trung bình và độ lệch chuẩn của dữ liệu được biến đổi là gì?

Bài 6.22. Một nhà nghiên cứu đo mức tăng trung bình hằng ngày (theo kg/ngày) của 20 con bò thịt; các giá trị tiêu biểu là

$$1, 39, 1.57, 1.44, \ldots$$

Trung bình của dữ liệu là 1.461 và độ lệch chuẩn là 0.178.

(a) Tính trung bình và độ lệch chuẩn theo lb/ngày. (*Gợi ý*: 1kg = 2.20 lb.)

(b) Tính hệ số biến thiên khi dữ liệu được trình bày (i) theo kg/ngày; (ii) theo lb/ngày.

Bài 6.23. Một mẫu bốn sinh viên có chiều cao như sau (theo cm): 180, 182, 179, 176. Giả sử sinh viên thứ năm được thêm vào mẫu. Chiều cao sinh viên này phải là bao nhiêu để chiều cao trung bình của mẫu bằng 181?

Bài 6.24. Một nhà thực vật học trồng 15 cây hồ tiêu trong một nhà kính. Sau 21 ngày, cô ta đo chiều dài thân cây (cm) của mỗi cây, và đạt được các giá trị sau:

12.4	12.2	13.4
10.9	12.2	12.1
11.8	13.5	12.0
14.1	12.7	13.2
12.6	11.9	13.1

(a) Xây dựng đồ thị stem - leaf cho các dữ liệu này, và đánh dấu vị trí của các tứ phân vị.

(b) Tính khoảng tứ phân vị.

Bài 6.25. Trong một nghiên cứu về hành vi của ruồi giấm *Drosophila melanogaster*, một nhà sinh vật học xác định, đối với mỗi con ruồi, tổng thời gian rỉa lông trong một thời đoạn quan sát sáu phút. Các số liệu sau đây là thời gian rỉa lông (giây) của 20 con ruồi:

34	24	10	16	52
76	33	31	46	24
18	26	57	32	25
48	22	48	29	19

(a) Xác định trung vị và tứ phân vị.

(b) Xác định khoảng tứ phân vị.

(c) Xây dựng đồ thị boxplot cho dữ liệu trên.

Bài 6.26. Một độ đo sức khỏe thể chất là sự hấp thu oxy tối đa, đó là tốc độ tối đa mà một người có thể tiêu thụ oxy. Một bài kiểm tra bằng máy chạy bộ đã được sử dụng để xác định sự hấp thu oxy tối đa của chín sinh viên nữ trước và sau khi tham gia vào một chương trình 10 tuần tập luyện sức khỏe. Bảng kèm theo cho thấy các độ đo trước và sau và sự thay đổi (sau−trước); tất cả các giá trị là theo ml O_2 mỗi mm mỗi kg thể trọng.

	TỐC ĐỘ HẤP THU OXY TỐI ĐA		
NGƯỜI THAM GIA	TRƯỚC	SAU	THAY ĐỔI
1	48.6	38.8	-9.8
2	38.0	40.7	2.7
3	31.2	32.0	0.8
4	45.5	45.4	-0.1
5	41.7	43.2	1.5
6	41.8	45.3	3.5
7	37.9	38.9	1.0
8	39.2	43.5	4.3
9	47.2	45.0	-2.2

Các tính toán sau được thực hiện lên biến *thay đổi* trong sự hấp thu oxy tối đa (cột bên tay phải).

(a) Tính trung bình và độ lệch chuẩn.

(b) Xác định trung vị.

(c) Loại bỏ người tham gia 1 khỏi dữ liệu và lặp lại phần (a) và (b). Độ đo mô tả nào không thay đổi? thay đổi?

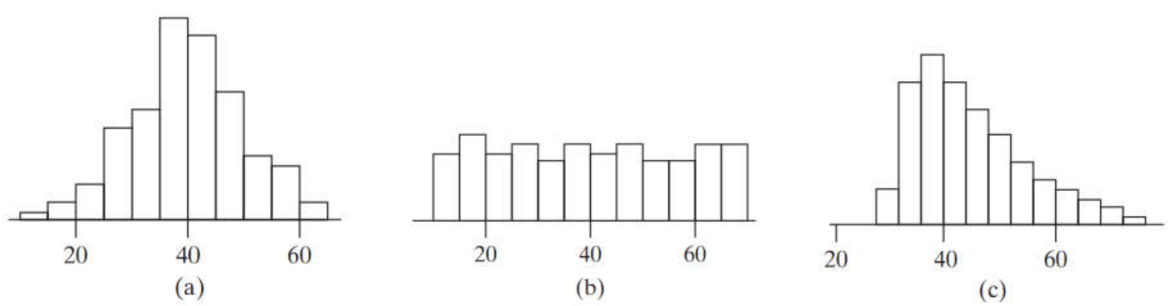
Bài 6.27. Một nhà giải phẫu học thú y khảo sát sự sắp xếp không gian của các tế bào thần kinh trong ruột của một con ngựa. Ông đã lấy ra một khối mô từ thành ruột, cắt khối thành nhiều phần bằng nhau, và đếm số lượng tế bào thần kinh trong mỗi 23 phần được chọn ngẫu nhiên. Các kết quả đếm được như sau:

35	19	33	34	17	26	16	40
28	30	23	12	27	33	22	31
28	28	35	23	23	19	29	

(a) Xác định trung vị, tứ phân vị và khoảng tứ phân vị.

(b) Xây dựng đồ thị boxplot cho dữ liệu.

Bài 6.28. Các đồ thị histogram (a), (b), và (c) sau trình bày ba phân phối.



Các kết quả xuất ra của máy tính đi kèm về trung bình, trung vị, và độ lệch chuẩn của ba phân phối, cộng với trung bình, trung vị, và độ lệch chuẩn cho phân phối thứ tư. Chọn đồ thị histogram phù hợp với các thống kê. Giải thích.

```
1.Count    100           2.Count    100
  Mean     41.3522         Mean     39.6761
  Median   39.5585         Median   39.5377
  StdDev   13.0136         StdDev   10.0476

3.Count    100           4.Count    100
  Mean     37.7522         Mean     39.6493
  Median   39.5585         Median   39.5448
  StdDev   13.0136         StdDev   17.5126
```

Bài 6.29. Các bác sĩ đo nồng độ canxi (nM) trong các mẫu máu của 38 người khỏe mạnh. Các dữ liệu được liệt kê như sau

95	110	135	120	88	125
112	100	130	107	86	130
122	122	127	107	107	107
88	126	125	112	78	115
78	102	103	93	88	110
104	122	112	80	121	126
90	96				

Tính các độ đo hướng tâm và sự biến thiên. Mô tả hình dạng của phân phối và các đặc điểm bất thường trong dữ liệu.

Bài 6.30. Số liệu về chiều cao của các sinh viên nữ (Đơn vị: inch) trong một lớp học như sau:

62	64	66	67	65	68	61	65	67	65	64	63	67
68	64	66	68	69	65	67	62	66	68	67	66	65
69	65	70	65	67	68	65	63	64	67	67		

(a) Tính chiều cao trung bình, phương sai và độ lệch tiêu chuẩn.

(b) Vẽ đồ thị Stem – Leaf cho chiều cao. Nhận xét.

(c) Tìm 3 phân vị q_1, q_2, q_3. Vẽ đồ thị boxplot cho chiều cao.

Hướng dẫn. (b) Để đồ thị Stem – Leaf phản ánh đúng phân phối số liệu chiều cao, ta sẽ chia stem là số trước dấu thập phân và leaf là số sau dấu thập phân, ví dụ $61 = 61.0$ có stem $= 61$ và leaf $= 0$. ■

Bài 6.31. Cho bộ dữ liệu sau:

4.2	4.7	4.7	5.0	3.8	3.6	3.0	5.1	3.1	3.8
4.8	4.0	5.2	4.3	2.8	2.0	2.8	3.3	4.8	5.0

(a) Tính trung bình mẫu, phương sai mẫu và độ lệch tiêu chuẩn.

(b) Tính hệ số biến thiên và miền giá trị mẫu.

Đáp án. (a) 4; 0.866; 0.931 ■

Bài 6.32. Cho bộ dữ liệu sau:

43	47	51	48	52	50	46	49
45	52	46	51	44	49	46	51
49	45	44	50	48	50	49	50

(a) Tính trung bình mẫu, phương sai mẫu và độ lệch tiêu chuẩn.

(b) Vẽ đồ thị Stem–Leaf cho dữ liệu trên.

(c) Xác định khoảng tứ phân vị (IQR).

Đáp án. (a) 48.125; 7.245; 2.692 ■

Bài 6.33. Xét biểu thức $y = \sum_{i=1}^{n}(x_i - a)^2$. Với a nào thì y đạt giá trị nhỏ nhất?

Đáp án. $\overline{x}$ ■

Bài 6.34. Xét $y_i = a + bx_i$, $i = 1, \ldots, n$ và a, b là các hằng số khác 0. Hãy tìm mối liên hệ giữa $\overline{x}$ và $\overline{y}$, s_x và s_y.

Đáp án. $\overline{y} = a + b\overline{x}$; $s_y = |b|s_x$ ■

Bài 6.35 (*). Giả sử ta có mẫu cỡ n gồm các giá trị quan trắc $x_1, x_2, \ldots, x_n$ và đã tính được trung bình mẫu $\overline{x}_n$ và phương sai mẫu s_n^2. Quan trắc thêm giá trị thứ $(n+1)$ là x_{n+1}, gọi $\overline{x}_{n+1}$ và s_{n+1}^2 lần lượt là trung bình mẫu và phương sai mẫu ứng với mẫu có $(n+1)$ quan trắc.

(a) Tính $\overline{x}_{n+1}$ theo $\overline{x}_n$ và x_{n+1}.

(b) Chứng tỏ rằng

$$ns_{n+1}^2 = (n-1)s_n^2 + \frac{n(x_{n+1} - \overline{x}_n)^2}{n+1}$$

Bài 6.36. Từ bảng các số ngẫu nhiên người ta lấy ra 150 số. Các số đó được phân thành 10 khoảng như sau:

x_i	1–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100
n_i	16	15	19	13	14	19	14	11	13	16

(a) Xác định trung bình mẫu và phương sai mẫu.

(b) Vẽ đồ thị tần số (histogram) cho dữ liệu trên.

Đáp án. (a) 48.97; 834.9 ■

Bài 6.37. Khảo sát thu nhập của công nhân ở một công ty, cho bởi bảng sau (đơn vị ngàn đồng).

Thu nhập	$[500, 600]$	$[600, 700]$	$[700, 800]$	$[800, 900]$	$[900, 1000]$	$[1000, 1100]$	$[1100, 1200]$
Số người	2	10	15	30	25	14	4

Xác định thu nhập trung bình, độ lệch chuẩn.

Đáp án. 874; 136.4 ■

Bài 6.38. Quan sát thời gian cần thiết để sản xuất một chi tiết máy, ta thu được số liệu cho bảng sau:

Khoảng thời gian (phút)	Số lần quan sát
20-25	2
25-30	14
30-35	26
35-40	32
40-45	14
45-50	8
50-55	4

(a) Tính trung bình mẫu $\overline{x}$, phương sai mẫu s^2.

(b) Vẽ đồ thị tần số (histogram) cho dữ liệu trên.

Đáp án. 36.6; 45.14 ■

Bài 6.39. Đo độ dài của một loại trục xe, ta có kết quả

Nhóm	18.4-18.6	18.6-18.8	18.8-19	19-19.2	19.2-19.4	19.4-19.6	19.6-19.8
n_i	1	4	20	41	19	8	4

Hãy tính độ dài trung bình và phương sai mẫu.

Đáp án. 19.133; 0.054 ■

Chương 7

Ước lượng tham số thống kê

7.1 Ước lượng trung bình tổng thể

Bài 7.1. Trên tập mẫu gồm 100 số liệu, người ta tính được $\overline{x} = 0.1$ $s = 0.014$. Xác định khoảng tin cậy 95% cho giá trị trung bình thật.

Đáp án. $(0.0973, 0.1027)$ ■

Bài 7.2. Chọn ngẫu nhiên 36 công nhân của xí nghiệp thì thấy lương trung bình là 380 ngàn đ/tháng. Giả sử lương công nhân tuân theo phân phối chuẩn với $\sigma = 14$ ngàn đồng. Với độ tin cậy 95%, hãy ước lượng mức lương trung bình của công nhân trong toàn xí nghiệp.

Đáp án. $(375.423, 384.573)$ ngàn đ/tháng ■

Bài 7.3. Đo sức bền chịu lực của một loại ống thí nghiệm, người ta thu được bộ số liệu sau

$$4500, 6500, 5200, 4800, 4900, 5125, 6200, 5375$$

Từ kinh nghiệm nghề nghiệp, người ta cũng biết rằng sức bền đó có phân phối chuẩn với độ lệch chuẩn $\sigma = 300$. Hãy xây dựng khoảng tin cậy 90% cho sức bền trung bình của loại ống trên.

Đáp án. $(5149.991, 5500.009)$ ■

Bài 7.4. Sản lượng mỗi ngày của một phân xưởng là biến ngẫu nhiên tuân theo luật chuẩn. Kết quả thống kê của 9 ngày cho ta:

$$27, 26, 21, 28, 25, 30, 26, 23, 26$$

Hãy xác định các khoảng tin cậy 95% cho sản lượng trung bình.

Đáp án. $(23.755, 27.805)$ ■

Bài 7.5. Quan sát chiều cao X (cm) của một số người, ta ghi nhận

x (cm)	140-145	145-150	150-155	155-160	160-165	165-170
Số người	1	3	7	9	5	2

(a) Tính $\overline{x}$ và s^2

(b) Ước lượng μ ở độ tin cậy 0.95

Đáp án. (a) 156.2; 37.68 (b) $(153.77, 158.63)$ ∎

Bài 7.6. Điểm trung bình môn toán của 100 thí sinh dự thi vào trường A là 5 với độ lệch chuẩn là 2.5.

(a) Ước lượng điểm trung bình môn toán của toàn thể thí sinh với độ tin cậy là 95%.

(b) Với sai số ước lượng điểm trung bình ở câu a) là 0.25 điểm, hãy xác định độ tin cậy.

Đáp án. (a) $(4.51, 5.49)$ (b) 68.26% ∎

Bài 7.7. Tuổi thọ của một loại bóng đèn được biết theo quy luật chuẩn với độ lệch chuẩn 100 giờ.

(a) Chọn ngẫu nhiên 100 bóng đèn để thử nghiệm, thấy mỗi bóng tuổi thọ trung bình là 1000 giờ. Hãy ước lượng tuổi thọ trung bình của bóng đèn xí nghiệp A sản xuất với độ tin cậy là 95%.

(b) Với dung sai của ước lượng tuổi thọ trung bình là 15 giờ, hãy xác định độ tin cậy.

(c) Để dung sai của ước lượng tuổi thọ trung bình không quá 25 giờ với độ tin cậy là 95% thì cần phải thử nghiệm ít nhất bao nhiêu bóng.

Đáp án. (a) $(980.4, 1019.6)$ (b) 86.64% (c) 62 ∎

Bài 7.8. Khối lượng các bao bột mì tại một cửa hàng lương thực tuân theo phân phối chuẩn. Kiểm tra 20 bao, thấy khối lượng trung bình của mỗi bao bột mì là 48kg, và phương sai mẫu $s^2 = (0.5\ kg)^2$.

(a) Với độ tin cậy 95% hãy ước lượng khối lượng trung bình của một bao bột mì thuộc cửa hàng.

(b) Với dung sai của ước lượng ở câu a) là 0.284 kg, hãy xác định độ tin cậy.

(c) Để dung sai của ước lượng ở câu a) không quá 160 g với độ tin cậy là 95%, cần phải kiểm tra ít nhất bao nhiêu bao?

Đáp án. (a) $(47.766, 48.234)$ (b) 0.98 (c) 38 ∎

Bài 7.9. Đo đường kính của một chi tiết máy do một máy tiện tự động sản xuất, ta ghi nhận được số liệu như sau:

x	12.00	12.05	12.10	12.15	12.20	12.25	12.30	12.35	12.40
n	2	3	7	9	10	8	6	5	3

với n chỉ số trường hợp tính theo từng giá trị của X (mm).

(a) Tính trung bình mẫu $\overline{x}$ và độ lệch chuẩn s của mẫu.

(b) Ước lượng đường kính trung bình μ ở độ tin cậy 0.95.

(c) Nếu muốn sai số ước lượng không quá $\varepsilon = 0.02$ mm ở độ tin cậy 0.95 thì phải quan sát ít nhất mấy trường hợp.

Đáp án. (a) 12.21; 0.103 (b) (12.18, 12.24) (c) 102 ∎

Bài 7.10. Người ta đo ion $Na+$ trên một số người và ghi nhận lại được kết quả như sau

$$129, 132, 140, 141, 138, 143, 133, 137, 140, 143, 138, 140$$

(a) Tính trung bình mẫu $\overline{x}$ và phương sai mẫu s^2.

(b) Ước lượng trung bình μ của tổng thể ở độ tin cậy 0.95.

(c) Nếu muốn sai số ước lượng trung bình không quá $\varepsilon = 1$ với độ tin cậy 0.95 thì phải quan sát mẫu gồm ít nhất mấy người?

Đáp án. (a) 137.83; 19.42 (b) (135.01, 140.63) (c) 75 ∎

Bài 7.11. Quan sát tuổi thọ x (giờ) của một số bóng đèn do xí nghiệp A sản xuất, ta ghi nhận

x	1000	1100	1200	1300	1400	1500	1600	1700	1800
n	10	14	16	17	18	16	16	12	9

với n chỉ số trường hợp theo từng giá trị của x.

(a) Tính trung bình mẫu $\overline{x}$ và độ lệch chuẩn mẫu s.

(b) Ước lượng tuổi thọ trung bình của bóng đèn ở độ tin cậy 0.95.

(c) Nếu muốn sai số ước lượng không quá $\varepsilon = 30$ giờ với độ tin cậy 0.99 thì phải quan sát mẫu gồm ít nhất mấy bóng đèn?

Đáp án. (a) 1391.41; 234.45 (b) (1350.79, 1432.03) (c) 235 ∎

Bài 7.12. Chiều dài của một loại sản phẩm được xuất khẩu hàng loạt là biến ngẫu nhiên phân phối chuẩn với $\mu = 100$ mm và $\sigma^2 = 4^2$ mm^2. Kiểm tra ngẫu nhiên 25 sản phẩm. Khả năng chiều dài trung bình của số sản phẩm kiểm tra nằm trong khoảng từ 98mm đến 101mm là bao nhiêu?

Đáp án. 88.82% ∎

7.2 Ước lượng tỉ lệ tổng thể

Bài 7.13. Trước bầu cử, người ta phỏng vấn ngẫu nhiên 2000 cử tri thì thấy có 1380 người ủng hộ một ứng cử viên K. Với độ tin cậy 95%, hỏi ứng cử viên đó thu được tối thiểu bao nhiêu phần trăm phiếu bầu?

Đáp án. 66.97% ■

Bài 7.14. Để ước lượng xác suất mắc bệnh gan với độ tin cậy 90% và sai số không vượt quá 2% thì cần phải khám ít nhất bao nhiêu người, biết rằng tỷ lệ mắc bệnh gan thực nghiệm đã cho bằng 0,9.

Đáp án. 613 ■

Bài 7.15. Giả sử quan sát 100 người thấy có 20 người bị bệnh sốt xuất huyết. Hãy ước lượng tỷ lệ bệnh sốt xuất huyết ở độ tin cậy 97%. Nếu muốn sai số ước lượng không quá 3% ở độ tin cậy 95% thì phải quan sát ít nhất bao nhiêu người?

Đáp án. $(0.1132, 0.2868)$; 683 ■

Bài 7.16. Một loại thuốc mới đem điều trị cho 50 người bị bệnh B, kết quả có 40 người khỏi bệnh.

(a) Ước lượng tỷ lệ khỏi bệnh p nếu dùng thuốc đó điều trị với độ tin cậy 0.95 và 0.99.

(b) Nếu muốn sai số ước lượng không quá 0.02 ở độ tin cậy 0.95 thì phải quan sát ít nhất mấy trường hợp?

Đáp án. (a) $(0.69, 0.91)$; $(0.65, 0.946)$ (b) 1537 ■

Bài 7.17. Ta muốn ước lượng tỷ lệ viên thuốc bị sức mẻ p trong một lô thuốc lớn.

(a) Nếu muốn sai số ước lượng không quá 0.01 với độ tin cậy 0.95 thì phải quan sát ít nhất mấy viên?

(b) Quan sát ngẫu nhiên 200 viên, thấy có 18 viên bị sứt mẻ. Hãy ước lượng p ở độ tin cậy 0.95.

(c) Khi đó, nếu muốn sai số ước lượng không quá 0.01 với độ tin cậy 0.95 thì phải quan sát ít nhất mấy viên?

Đáp án. (a) 9604 (b) $(0.051, 0.13)$ (c) 3147 ■

Bài 7.18. Muốn biết trong ao có bao nhiêu cá, người ta bắt lên 2000 con, đánh dấu xong lại thả xuống hồ. Sau một thời gian, người ta bắt lên 500 con và thấy có 20 con cá có đánh dấu của lần bắt trước. Dựa vào kết quả đó hãy ước lượng số cá có trong hồ với độ tin cậy 95%.

Đáp án. $(34965.03, 877719.3)$ ■

Bài 7.19. Để có thể dự đoán được số lượng chim thường nghỉ tại vườn nhà mình, người chủ bắt 89 con, đem đeo khoen cho chúng rồi thả đi. Sau một thời gian, ông bắt ngẫu nhiên được 120 con và thấy có 7 con có đeo khoen. Hãy dự đoán số chim giúp ông chủ vườn ở độ tin cậy 99%.

Đáp án. $(785.1688, 27396.59)$ ■

7.3 Tổng hợp

Bài 7.20. Cân thử 100 quả cam, ta có bộ số liệu sau:

Khối lượng (g)	32	33	34	35	36	37	38	39	40
Số quả	2	3	15	26	28	6	8	8	4

(a) Hãy ước lượng khối lượng trung bình các quả cam ở độ tin cậy 95%.

(b) Cam có khối lượng dưới 34 g được coi là cam loại 2. Tìm khoảng ước lượng cho tỷ lệ loại 2 với độ tin cậy 90%.

Đáp án. (a) $(35.539, 36.241)$ (b) $(0.014, 0.086)$ ■

Bài 7.21. Đem cân một số trái cây vừa thu hoạch, ta được kết quả sau:

X (gam)	200-210	210-220	220-230	230-240	240-250
Số trái	12	17	20	18	15

(a) Tìm khoảng ước lượng của trọng lượng trung bình μ của trái cây với độ tin cậy 0.95 và 0.99.

(b) Nếu muốn sai số ước lượng không quá $\varepsilon = 2$ *gam* ở độ tin cậy 99% thì phải quan sát ít nhất bao nhiêu trái?

(c) Trái cây có khối lượng $X \geq 230$ *gam* được xếp vào loại A. Hãy tìm khoảng ước lượng cho tỷ lệ p của trái cây loại A ở độ tin cậy 0.95 và 0.99. Nếu muốn sai số ước lượng không quá 0.04 ở độ tin cậy 0.99 thì phải quan sát ít nhất mấy trường hợp?

Đáp án. (a) (222.98, 228.72); (222.08, 229.63) (b) 293 (c) (0.2963, 0.5085); (0.2627,0.5421); 1001 ■

Chương 8

Kiểm định giả thuyết thống kê

8.1 So sánh kì vọng với một số cho trước

Bài 8.1. Giám đốc một xí nghiệp cho biết lương trung bình của 1 công nhân thuộc xí nghiệp là 380 ngàn đ/tháng. Chọn ngẫu nhiên 36 công nhân thấy lương trung bình là 350 ngàn đ/tháng, với độ lệch chuẩn $s = 40$. Lời báo cáo của giám đốc có tin cậy được không, với mức có ý nghĩa là $\alpha = 5\%$.

Hướng dẫn. Ta cần kiểm định các giả thuyết

$$\begin{cases} H_0 : \mu = 380 \\ H_1 : \mu \neq 380 \end{cases}$$

Đây là trường hợp $n = 36 \geq 30$ và σ^2 chưa biết, nên ta dùng

$$\begin{aligned} z &= \frac{\sqrt{n}(\overline{x} - \mu)}{s} \\ &= \frac{\sqrt{36}(350 - 380)}{40} \\ &= -4.5 \end{aligned}$$

Ta thấy $|z| > z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$. Do đó ta bác bỏ giả thuyết H_0. Nghĩa là lời báo cáo của giám đốc không đáng tin cậy. ∎

Bài 8.2. Trong thập niên 80, trọng lượng trung bình của thanh niên là 48 kg. Nay để xác định lại trọng lượng ấy, người ta chọn ngẫu nhiên 100 thanh niên đo trọng lượng trung bình là 50 kg và phương sai mẫu $s^2 = (10\ kg)^2$. Thử xem trọng lượng thanh niên hiện nay phải chăng có thay đổi, với mức có ý nghĩa là 1%?

Đáp án. $z = 2$. Trọng lượng thanh niên hiện nay không thay đổi so với trước kia. ∎

Bài 8.3. Một cửa hàng thực phẩm nhận thấy thời gian vừa qua trung bình một khách hàng mua 25 ngàn đồng thực phẩm trong ngày. Nay cửa hàng chọn ngẫu nhiên 15 khách hàng thấy trung bình một khách hàng mua 24 ngàn đồng trong ngày và phương sai mẫu là $s^2 = (2 \text{ ngàn đồng})^2$.
Với mức ý nghĩa là 5%, kiểm định xem có phải sức mua của khách hàng hiện nay thực sự giảm sút hay không. Biết rằng sức mua của khách hàng có phân phối chuẩn.

Đáp án. $t = -1.9365$. Sức mua của khách hàng hiện nay thực sự giảm sút. ∎

Bài 8.4. Đối với người Việt Nam, lượng huyết sắc tố trung bình là 138.3 g/l. Khám cho 80 công nhân ở nhà máy có tiếp xúc hoá chất, thấy huyết sắc tố trung bình $\overline{x} = 120\ g/l$; $s = 15\ g/l$. Từ kết quả trên, có thể kết luận lượng huyết sắc tố trung bình của công nhân nhà máy hoá chất này thấp hơn mức chung hay không? Kết luận với $\alpha = 0.05$.

Đáp án. $z = -10.912$. Lượng huyết tố trung bình của công nhân nhà máy thấp hơn mức chung. ∎

Bài 8.5. Trong điều kiện chăn nuôi bình thường, lượng sữa trung bình của 1 con bò là 14 kg/ngày. Nghi ngờ điều kiện chăn nuôi kém đi làm cho lượng sữa giảm xuống, người ta điều tra ngẫu nhiên 25 con và tính được lượng sữa trung bình của 1 con trong 1 ngày là 12.5 và độ lệch chuẩn $s = 2.5$. Với mức ý nghĩa $\alpha = 0.05$. hãy kết luận điều nghi ngờ nói trên. Giả thiết lượng sữa bò là 1 biến ngẫu nhiên chuẩn.

Đáp án. $t = -3$. Điều kiện chăn nuôi kém đi làm cho lượng sữa giảm xuống. ∎

Bài 8.6. Tiền lương trung bình của công nhân trước đây là 400 ngàn đ/tháng. Để xét xem tiền lương hiện nay so với mức trước đây thế nào, người ta điều tra 100 công nhân và tính được $\overline{x} = 404.8$ ngàn đ/tháng và $s = 20$ ngàn đ/tháng. Với $\alpha = 1\%$

(a) Nếu lập giả thiết 2 phía và giả thiết 1 phía thì kết quả kiểm định như thế nào?

(b) Giống câu a, với $\overline{x} = 406$ ngàn đ/tháng và $s = 20$ ngàn đ/tháng.

Bài 8.7. Một máy đóng gói các sản phẩm có khối lượng 1 kg. Nghi ngờ máy hoạt động không bình thường, người ta chọn ra một mẫu ngẫu nhiên gồm 100 sản phẩm thì thấy như sau:

Khối lượng	0.95	0.97	0.99	1.01	1.03	1.05
Số gói	9	31	40	15	3	2

Với mức ý nghĩa 0.05, hãy kết luận về nghi ngờ trên.

Đáp án. $z = -6.9204$. Máy hoạt động không bình thường. ∎

Bài 8.8. Trọng lượng trung bình khi xuất chuồng ở một trại chăn nuôi trước là 3.3 kg/con. Năm nay người ta sử dụng một loại thức ăn mới, cân thử 15 con khi xuất chuồng ta được các số liệu như sau:

$$3.25, 2.50, 4.00, 3.75, 3.80, 3.90, 4.02, 3.60, 3.80, 3.20, 3.82, 3.40, 3.75, 4.00, 3.50$$

Giả thiết trọng lượng gà là đại lượng ngẫu nhiên phân phối theo quy luật chuẩn.

(a) Với mức ý nghĩa $\alpha = 0.05$. Hãy cho kết luận về tác dụng của loại thức ăn này?

(b) Nếu trại chăn nuôi báo cáo trọng lượng trung bình khi xuất chuồng là 3.5 kg/con thì có chấp nhận được không? ($\alpha = 0.05$).

Đáp án. (a) $t = 3.0534$. Thức ăn mới này làm thay đổi trọng lượng gà.

(b) $t = 1.1409$. Trại chăn nuôi báo cáo trọng lượng trung bình khi xuất chuồng là chấp nhận được. ∎

Bài 8.9. Đo cholesterol (đơn vị mg%) cho một nhóm người, ta ghi nhận lại được

Chol.	150 –160	160 - 170	170 - 180	180 - 190	190 - 200	200 - 210
Số người	3	9	11	3	2	1

Cho rằng độ cholesterol tuân theo phân phối chuẩn.

(a) Tính trung bình mẫu $\overline{x}$ và phương sai mẫu s^2.

(b) Tìm khoảng ước lượng cho trung bình cholesterol trong dân số ở độ tin cậy 0.95.

(c) Có tài liệu cho biết lượng cholesterol trung bình là $\mu_0 = 175\ mg\%$. Giá trị này có phù hợp với mẫu quan sát không? (kết luận với $\alpha = 0.05$).

Đáp án. (a) 173.2759; 143.3498 (b) $(168.7226, 177.8292)$ (c) $t = -0.7755$. Giá trị mẫu phù hợp với tài liệu. ∎

Bài 8.10. Quan sát số hoa hồng bán ra trong một ngày của một cửa hàng bán hoa sau một thời gian, người ta ghi được số liệu sau:

Số hoa hồng (đoá)	12	13	15	16	17	18	19
Số ngày	3	2	7	7	3	2	1

Giả thiết rằng số hoa bán ra trong ngày có phân phối chuẩn.

(a) Tìm trung bình mẫu $\overline{x}$, phương sai mẫu s^2.

(b) Sau khi tính toán, ông chủ cửa hàng nói rằng nếu trung bình một ngày không bán được 15 đoá hoa thì chẳng thà đóng cửa còn hơn. Dựa vào số liệu trên, anh (chị) hãy kết luận giúp ông chủ cửa hàng xem có nên tiếp tục bán hay không ở mức ý nghĩa $\alpha = 0.05$.

(c) Giả sử những ngày bán được từ 13 đến 17 đoá hồng là những ngày “bình thường”. Hãy ước lượng tỉ lệ của những ngày bình thường của cửa hàng ở độ tin cậy 90%.

Đáp án. (a) 15.4; 3.5 (b) $t = 1.069$. Ông chủ cửa hàng nên tiếp tục bán. (c) $(0.6191, 0.9009)$ ∎

Bài 8.11. Một xí nghiệp đúc một số rất lớn các sản phẩm bằng thép với số khuyết tật trung bình ở mỗi sản phẩm là 3. Người ta cải tiến cách sản xuất và kiểm tra 36 sản phẩm. Kết quả như sau:

Số khuyết tật trên sản phẩm	0	1	2	3	4	5	6
Số sản phẩm tương ứng	7	4	5	7	6	6	1

Giả sử số khuyết tật của các sản phẩm có phân phối chuẩn.

(a) Hãy ước lượng số khuyết tật trung bình ở mỗi sản phẩm sau khi cải tiến, với độ tin cậy 90%.

(b) Hãy cho kết luận về hiệu quả của việc cải tiến sản xuất ở mức ý nghĩa 0.05.

Đáp án. (a) $(2.1333, 3.1445)$ (b) $z = -1.1785$. Cải tiến không hiệu quả. ■

Bài 8.12. Đánh giá tác dụng của một chế độ ăn bồi dưỡng mà dấu hiệu quan sát là số hồng cầu. Người ta đếm số hồng cầu của 20 người trước và sau khi ăn bồi dưỡng:

x_i	32	40	38	42	41	35	36	47	50	30
y_i	40	45	42	50	52	43	48	45	55	34
x_i	38	45	43	36	50	38	42	41	45	44
y_i	32	54	58	30	60	35	50	48	40	50

Với mức ý nghĩa $\alpha = 0.05$, có thể kết luận gì về tác dụng của chế độ ăn bồi dưỡng này?

Đáp án. $t = 3.0386$. Chế độ ăn bồi dưỡng làm thay đổi hồng cầu. ■

Hướng dẫn. Đặt $Z = Y - X$ để chỉ số lượng hồng cầu thay đổi sau khi ăn bồi dưỡng. ■

Bài 8.13. Giả sử ta muốn xác định xem hiệu quả của chế độ ăn kiêng đối với việc giảm trọng lượng như thế nào. 20 người quá béo đã thực hiện chế độ ăn kiêng. Trọng lượng của từng người trước khi ăn kiêng (X kg) và sau khi ăn kiêng (Y kg) được cho như sau:

X	80	78	85	70	90	78	92	88	75	75
Y	75	77	80	70	84	74	85	82	80	65
X	63	72	89	76	77	71	83	78	82	90
Y	62	71	83	72	82	71	79	76	83	81

Kiểm tra xem chế độ ăn kiêng có tác dụng làm thay đổi trọng lượng hay không ($\alpha = 0.05$).

Đáp án. $t = -3.3002$. Chế độ ăn kiêng có tác dụng làm thay đổi trọng lượng. ■

8.2 So sánh hai kì vọng

Bài 8.14. Một nhà phát triển sản phẩm quan tâm đến việc giảm thời gian khô của sơn. Vì vậy hai công thức sơn được đem thử nghiệm. Công thức 1 là công thức có các thành phần chuẩn và công thức 2 có thêm một thành phần làm khô mới được cho rằng sẽ làm giảm thời gian khô của sơn. Từ các thí nghiệm người ta thấy rằng $\sigma_1 = \sigma_2 = 8$ phút. 10 đồ vật được sơn với công thức 1 và 10 đồ vật khác được sơn với công thức 2. Thời gian khô trung bình của từng mẫu là $\overline{x_1} = 121$ phút và $\overline{x_2} = 112$ phút. Nhà phát triển sản phẩm có thể rút ra kết luận gì về ảnh hưởng của thành phần làm khô mới? Với mức ý nghĩa 5%.

Hướng dẫn. Ta cần kiểm định các giả thuyết

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

Ta tính được

$$\begin{aligned} z &= \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \\ &= \frac{121 - 112}{\sqrt{\frac{8^2}{10} + \frac{8^2}{10}}} \\ &= 2.5156 \end{aligned}$$

Ta thấy $z > z_{1-\alpha} = z_{0.95} = 1.65$. Do đó, ta bác bỏ giả thuyết H_0 nghĩa là thành phần làm khô mới làm giảm thời gian khô. ∎

Bài 8.15. Tốc độ cháy của hai loại chất nổ lỏng được dùng làm nhiên liệu trong tàu vũ trụ được nghiên cứu. Người ta biết rằng độ lệch chuẩn của tốc độ cháy của hai loại nhiên liệu bằng nhau và bằng 3 cm/s. Hai mẫu ngẫu nhiên kích thước $n_1 = 20$ và $n_2 = 20$ được thử nghiệm; trung bình mẫu tốc độ cháy là $\overline{x_1} = 18\ cm/s$ và $\overline{x_2} = 24\ cm/s$. Với mức ý nghĩa $\alpha = 0.05$ hãy kiểm định giả thuyết hai loại chất nổ lỏng này có cùng tốc độ đốt cháy.

Đáp án. $z = -6.3246$. Hai loại chất nổ lỏng này có tốc độ đốt cháy khác nhau. ∎

Bài 8.16. Theo dõi giá cổ phiếu của 2 công ty A và B trong vòng 31 ngày người ta tính được các giá trị sau

	$\overline{x}$	s
Công ty A	37.58	1.50
Công ty B	38.24	2.20

Giả thiết rằng giá cổ phiếu của hai công ty A và B là hai biến ngẫu nhiên phân phối theo quy luật chuẩn. Hãy cho biết ý nghĩa kì vọng của các biến ngẫu nhiên nói trên? Hãy cho biết có sự khác biệt thực sự về giá cổ phiếu trung bình của hai công ty A và B không? Với mức ý nghĩa $\alpha = 5\%$

Đáp án. $t = -1.3801$. Giá cổ phiếu trung bình của hai công ty A và B bằng nhau. ∎

Bài 8.17. Hàm lượng đường trong máu của công nhân sau 5 giờ làm việc với máy siêu cao tần đã đo được ở hai thời điểm trước và sau 5 giờ làm việc. Ta có kết quả sau:

Trước:	$n_1 = 50$	$\overline{x} = 60\ mg\%$	$s_x = 7$
Sau:	$n_2 = 40$	$\overline{y} = 52\ mg\%$	$s_y = 9.2$

Với mức ý nghĩa $\alpha = 0.05$, có thể khẳng định hàm lượng đường trong máu sau 5 giờ làm việc đã giảm đi hay không?

Đáp án. $t = 4.6851$. Hàm lượng đường trong máu sau 5 giờ làm việc đã giảm đi. ∎

Bài 8.18. Trồng cùng một giống lúa trên hai thửa ruộng như nhau và bón hai loại phân khác nhau. Đến ngày thu hoạch ta có kết quả như sau:

- Thửa thứ nhất lấy mẫu 1000 bông lúa thấy số hạt trung bình của mỗi bông là $\overline{x} = 70$ hạt và $s_x = 10$.

- Thửa thứ hai lấy mẫu 500 bông thấy số hạt trung bình mỗi bông là $\overline{y} = 72$ hạt và $s_y = 20$.

Hỏi sự khác nhau giữa X và Y là ngẫu nhiên hay bản chất, với $\alpha = 0.05$?

Đáp án. $t = -2.5824$. Sự khác nhau giữa X và Y là do bản chất. ∎

Bài 8.19. Để so sánh trọng lượng trung bình của trẻ sơ sinh ở thành thị và nông thôn, người ta thử cân trọng lượng của 10000 cháu và thu được kết quả sau đây:

Vùng	Số cháu được cân	Trọng lượng trung bình	Độ lệch chuẩn mẫu
Nông thôn	8000	3.0 kg	0.3 kg
Thành thị	2000	3.2 kg	0.2 kg

Với mức ý nghĩa $\alpha = 0.05$ có thể coi trọng lượng trung bình của trẻ sơ sinh ở thành thị cao hơn ở nông thôn hay không? (Giả thiết trọng lượng trẻ sơ sinh là biến ngẫu nhiên chuẩn).

Đáp án. $t = -28.2885$. Trọng lượng trung bình của trẻ sơ sinh ở thành thị cao hơn ở nông thôn. ∎

Bài 8.20. Để so sánh năng lực học toán và vật lý của học sinh, người ta kiểm tra ngẫu nhiên 8 em bằng hai bài toán và vật lý. Kết quả cho bởi bảng dưới đây (X là điểm toán, Y là điểm lý):

X	15	20	16	22	24	18	20	14
Y	15	22	14	25	19	20	24	16

Giả sử X và Y đều có phân phối chuẩn. Hãy so sánh điểm trung bình giữa X và Y, mức ý nghĩa 5%.

Đáp án. $t = -0.3913$. Điểm trung bình của X và Y là như nhau. ∎

Bài 8.21. Hai máy được sử dụng để rót nước vào các bình. Người ta lấy mẫu ngẫu nhiên 10 bình do máy thứ nhất và 10 bình do máy thứ hai thì được kết quả sau:

Máy 1	16.03	16.01	16.04	15.96	16.05	15.98	16.05	16.02	16.02	15.99
Máy 2	16.02	16.03	15.97	16.04	15.96	16.02	16.01	16.01	15.99	16.00

Với mức ý nghĩa $\alpha = 0.05$ có thể nói rằng hai máy rót nước vào bình như nhau không?

Đáp án. $t = 0.7986$. Hai máy rót nước vào bình như nhau. ∎

Bài 8.22. Để nghiên cứu ảnh hưởng của một loại thuốc, người ta cho 10 bệnh nhân uống thuốc. Lần khác họ cũng cho bệnh nhân uống thuốc nhưng là thuốc giả. Kết quả thí nghiệm thu được như sau:

Bệnh nhân	1	2	3	4	5	6	7	8	9	10
Số giờ ngủ có thuốc	6.1	7.0	8.2	7.6	6.5	8.4	6.9	6.7	7.4	5.8
Số giờ ngủ với thuốc giả	5.2	7.9	3.9	4.7	5.3	5.4	4.2	6.1	3.8	6.3

Giả sử số giờ ngủ của bệnh nhân tuân theo phân phối chuẩn. Với mức ý nghĩa 5%, hãy kết luận về ảnh hưởng của loại thuốc trên.

Đáp án. $t = 3.7134$. Loại thuốc trên ảnh hưởng đến số giờ ngủ của bệnh nhân. ∎

Bài 8.23. Quan sát sức nặng của bé trai (X) và bé gái (Y) lúc sơ sinh (đơn vị gam), ta có kết quả

Trọng lượng	3000-3200	3200-3400	3400-3600	3600-3800	3800-4000
Số bé trai	1	3	8	10	3
Số bé gái	2	10	10	5	1

(a) Tính $\overline{x}$, $\overline{y}$, s_x^2, s_y^2.

(b) So sánh các kì vọng μ_X, μ_Y (kết luận với $\alpha = 5\%$).

(c) Nhập hai mẫu lại. Tính trung bình và độ lệch chuẩn của mẫu nhập. Dùng mẫu nhập để ước lượng sức nặng trung bình của trẻ sơ sinh ở độ tin cậy 95%.

Đáp án. (a) 3588; 3450; 40266.67; 37407.41

(b) $t = 2.5476$. Trọng lượng bé trai và bé gái lúc sơ sinh khác nhau.

(c) 3515.094; 206.9896; $(3459.367, 3570.821)$ ∎

8.3 So sánh tỉ lệ với một số cho trước

Bài 8.24. Một máy sản xuất tự động với tỷ lệ chính phẩm là 98%. Sau một thời gian hoạt động, người ta nghi ngờ tỷ lệ trên đã bị giảm. Kiểm tra ngẫu nhiên 500 sản phẩm thấy có 28 phế phẩm, với $\alpha = 0.05$ hãy kiểm tra xem chất lượng làm việc của máy có còn được như trước hay không?

Hướng dẫn. Gọi p là tỉ lệ chính phẩm của máy sản xuất tự động sau một thời gian hoạt động.
Ta cần kiểm định các giả thuyết

$$\begin{cases} H_0 : p = 0.98 \\ H_1 : p < 0.98 \end{cases}$$

Ta có $n = 500$, $f = \dfrac{500-28}{500} = 0.944$, $nf = 472 \geq 5$ và $n(1-f) = 28 \geq 5$.
Do đó, ta dùng

$$\begin{aligned} z &= \frac{\sqrt{n}(f-p)}{\sqrt{pq}} \\ &= \frac{\sqrt{500}(0.944-0.98)}{\sqrt{0.98 \times 0.02}} \\ &= -5.7499 \end{aligned}$$

Ta thấy $z < z_\alpha = z_{0.05} = -z_{0.95} = -1.65$. Do đó ta bác bỏ giả thuyết H_0. Nghĩa là chất lượng làm việc của máy không còn tốt như trước. ∎

Bài 8.25. Trong một vùng dân cư có 18 bé trai và 28 bé gái mắc bệnh B. Hỏi rằng tỷ lệ nhiễm bệnh của bé trai và bé gái có như nhau không? (kết luận với $\alpha = 0.05$ và giả sử rằng số lượng bé trai và bé gái trong vùng tương đương nhau, và rất nhiều).

Đáp án. $z = -1.4745$. Tỉ lệ mắc bệnh của bé trai và bé gái là như nhau. ∎

Bài 8.26. Đo huyết sắc tố cho 50 công nhân nông trường thấy có 60% ở mức dưới 110 g/l. Số liệu chung của khu vực này là 30% ở mức dưới 110 g/l. Với mức ý nghĩa $\alpha = 0.05$, có thể kết luận công nhân nông trường có tỷ lệ huyết sắc tố dưới 110 g/l cao hơn mức chung hay không?

Đáp án. $z = 4.6291$. Công nhân nông trường có tỷ lệ huyết sắc tố dưới 110 g/l cao hơn mức chung. ∎

Bài 8.27. Theo một nguồn tin thì tỉ lệ hộ dân thích xem dân ca trên Tivi là 80%. Thăm dò 36 hộ dân thấy có 25 hộ thích xem dân ca. Với mức có ý nghĩa là 5%. Kiểm định xem nguồn tin này có đáng tin cậy không?

Đáp án. $z = -1.584$. Nguồn tin này đáng tin cậy. ∎

Bài 8.28. Một máy sản suất tự động, lúc đầu tỷ lệ sản phẩm loại A là 20%. Sau khi áp dụng một phương pháp cải tiến sản xuất mới, người ta lấy 40 mẫu, mỗi mẫu gồm 10 sản phẩm đề kiểm tra. Kết quả kiểm tra cho ở bảng sau:

Số sản phẩm loại A trong mẫu	1	2	3	4	5	6	7	8	9	10
Số mẫu	2	0	4	6	8	10	4	5	1	0

Với mức ý nghĩa 5%. Hãy cho kết luận về phương pháp sản suất này.

Đáp án. $z = 16.875$. Phương pháp cải tiến sản xuất mới thay đổi tỉ lệ sản phẩm loại A. ∎

Bài 8.29. Tỷ lệ phế phẩm của một nhà máy trước đây là 5%. Năm nay nhà máy áp dụng một biện pháp kỹ thuật mới. Để nghiên cứu tác dụng của biện pháp kỹ thuật mới, người ta lấy một mẫu gồm 800 sản phẩm để kiểm tra và thấy có 24 phế phẩm.

(a) Với $\alpha = 0.01$. Hãy cho kết luận về biện pháp kỹ thuật mới này?

(b) Nếu nhà máy báo cáo tỷ lệ phế phẩm sau khi áp dụng biện pháp kỹ thuật mới là 2% thì có chấp nhận được không? ($\alpha = 0.01$).

Đáp án. (a) $z = -2.5955$. Biện pháp kĩ thuật mới làm thay đổi tỉ lệ phế phẩm.

(b) $z = 2.0203$. Nhà máy báo cáo tỷ lệ phế phẩm là chấp nhận được. ∎

8.4 So sánh hai tỉ lệ

Bài 8.30. Trong 90 người dùng DDT để ngừa bệnh ngoài da thì có 10 người nhiễm bệnh; trong 100 người không dùng DDT thì có 26 người mắc bệnh. Hỏi rằng DDT có tác dụng ngừa bệnh ngoài da không? (kết luận với $\alpha = 0.05$)

Hướng dẫn. Gọi

p_1 : tỉ lệ người mắc bệnh dùng DDT
p_2 : tỉ lệ người mắc bệnh không dùng DDT

Ta cần kiểm định các giả thuyết

$$\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 < p_2 \end{cases}$$

Ta có

$$\begin{aligned} n_1 &= 90 \geq 30 \\ n_2 &= 100 \geq 30 \\ f_1 &= \frac{10}{90} = 0.1111 \\ f_2 &= \frac{26}{100} = 0.26 \\ \hat{p} &= \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2} \\ &= \frac{10 + 26}{90 + 100} = 0.1895 \end{aligned}$$

Ta tính được

$$\begin{aligned} z &= \frac{f_1 - f_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ &= \frac{0.1111 - 0.26}{\sqrt{0.1895(1 - 0.1895)\left(\frac{1}{90} + \frac{1}{100}\right)}} \\ &= -2.6149 \end{aligned}$$

Ta thấy $z < z_\alpha = z_{0.05} = -z_{0.95} = -1.65$. Do đó ta bác bỏ giả thuyết H_0. Nghĩa là DDT có tác dụng ngừa bệnh ngoài da. ■

Bài 8.31. Người ta điều tra 250 người ở xã A thấy có 140 nữ và điều tra 160 người ở xã B thấy có 80 nữ. Hãy so sánh tỉ lệ nữ ở hai xã với mức ý nghĩa 5%.

Đáp án. $z = 1.1885$. Tỉ lệ nữ ở hai xã bằng nhau. ■

Bài 8.32. Áp dụng hai phương pháp gieo hạt. Theo phương pháp A gieo 180 hạt thì có 150 hạt nảy mầm; theo phương pháp B gieo 256 hạt thì thấy có 160 hạt nảy mầm. Hãy so sánh hiệu quả của hai phương pháp với mức ý nghĩa $\alpha = 5\%$.

Đáp án. $z = 4.7241$. Hiệu quả của hai phương pháp khác nhau ở mức ý nghĩa 5%. ■

Bài 8.33. Theo dõi trọng lượng của một số trẻ sơ sinh tại một số nhà hộ sinh thành phố và nông thôn, người ta thấy rằng trong số 150 trẻ sơ sinh ở thành phố có 100 cháu nặng hơn 3000 gam, và trong 200 trẻ sơ sinh ở nông thôn có 98 cháu nặng hơn 3000 gam. Từ kết quả đó hãy so sánh tỉ lệ trẻ sơ sinh có trọng lượng trên 3000 gam ở thành phố và nông thôn với mức ý nghĩa 5%.

Dáp án. $z = 3.3005$. Tỉ lệ trẻ sơ sinh có trọng lượng trên 3000 gam ở thành phố và nông thôn khác nhau ở mức ý nghĩa 5%. ■