TRƯỜNG ĐẠI HỌC QUY NHƠN

KHOA KINH TẾ - KẾ TOÁN

BỘ MÔN TOÁN KINH TẾ

…………………………………………………..

CAO TẤN BÌNH

BÀI GIẢNG
LÝ THUYẾT XÁC SUẤT VÀ THỐNG KÊ TOÁN

Bình Định, 2015

CHƯƠNG 1

BIẾN CỐ NGẪU NHIÊN VÀ XÁC SUẤT

Thuật ngữ xác suất đề cập đến việc nghiên cứu tính ngẫu nhiên và không chắc chắn. Ngôn ngữ của xác suất thường xuyên được sử dụng một cách không chính thức trong cả văn nói và văn viết. Chẳng hạn như " Rất có khả năng giá cổ phiếu A sẽ tăng trong phiên giao dịch tới", " Khoảng 50-50 cơ hội ông B sẽ tái đắc cử", "Hy vọng rằng ít nhất 10,000 vé hòa nhạc sẽ được bán ra". Chương 1 giới thiệu các khái niệm xác suất cơ bản, cho phép chúng ta diễn đạt các hiện tượng ngẫu nhiên bằng ngôn ngữ toán học một cách sáng sủa và logic. Các nghiên cứu về xác suất như một nhánh của toán học ra đời cách đây khoảng hơn 300 năm, bắt nguồn từ những câu hỏi liên quan đến các trò chơi ngẫu nhiên.

1.1 CÁC KHÁI NIỆM CƠ BẢN VỀ QUY TẮC ĐẾM

Quy tắc nhân

Một công việc được tiến hành qua k giai đoạn. Giai đoạn 1 có m_1 cách thực hiện, giai đoạn 2 có m_2 cách thực hiện, …, giai đoạn k có m_k cách thực hiện. Khi đó có $n = m_1.m_2...m_k$ cách thực hiện công việc đó.

Tổ hợp

$$C_n^k = \frac{n!}{k!(n-k)!} \text{ với } k,n \in \mathbb{N} \text{ và } 0 \le k \le n$$

Chỉnh hợp

$$A_n^k = \begin{cases} \dfrac{n!}{(n-k)!} \\ n^k \end{cases}$$

Hoán vị

$$P_n = \begin{cases} n! \\ \dfrac{n!}{n_1!n_2!...n_k!} \end{cases}$$

1.2 PHÉP THỬ VÀ CÁC LOẠI BIẾN CỐ

Phép thử

Khi quan sát một hiện tượng hay làm một thí nghiệm và chú ý đến kết quả của hiện tượng hay thí nghiệm đó, ta nói đã thực hiện một phép thử. Kết quả có thể xảy ra trong một phép thử gọi là các biến cố.

Ví dụ 1.2.1 Quan sát tình trạng hoạt động của một máy là một phép thử. Việc máy chạy tốt hay hỏng hóc là các biến cố.

Ví dụ 1.2.2 Tung một đồng xu là lam một phép thử. Mặt sấp xuất hiện hay mặt ngửa xuất hiện là các biến cố.

Ví dụ 1.2.3 Từ một lô sản phẩm gồm chính phẩm và phế phẩm, lấy ngẫu nhiên một sản phẩm. Khi đó việc lấy sản phẩm là phép thử, còn việc lấy được chính phẩm hay phế phẩm là các biến cố.

Các loại biến cố

Biến cố ngẫu nhiên: Là biến cố có thể xảy ra cũng có thể không xảy ra khi thực hiện một phép thử. Ta thường ký hiệu các biến cố ngẫu nhiên bởi các chữ cái in: A, B, C, D,…

Biến cố chắc chắn: Là biến cố nhất định sẽ xảy ra khi thực hiện một phép thử. Ta thường dùng ký hiệu U để biểu diễn biến cố chắc chắn.

Biến cố không thể có: Là biến cố không thể xảy ra khi thực hiện một phép thử. Ta thường dùng ký hiệu V để biểu diễn biến cố không thể có.

Ví dụ 1.2.4 Thực hiện phép thử tung một con xúc xắc. Khi đó biến cố ngẫu nhiên là A: mặt hai chấm xuất hiện, biến cố chắc chắn U là: mặt có số chấm nhỏ hơn 6 xuất hiện, và biến cố không thể có là: Mặt 7 chấm xuất hiện.

1.3 CÁC PHÉP TOÁN VỀ BIẾN CỐ

Tổng các biến cố

Biến cố E được gọi là tổng của hai biến cố A và B nếu E chỉ xảy ra khi có ít nhất một trong hai biến cố A hoặc B xảy ra. Ký hiệu $E = A + B$.

Biến cố A được gọi là tổng của n biến cố $A_1, A_2, ..., A_n$ nếu A chỉ xảy ra khi có ít nhất một trong n biến cố đó xảy ra. Ký hiệu $A = A_1 + A_2 + ... + A_n = \sum_{i=1}^{n} A_i$.

Tích các biến cố

Biến cố E được gọi là tích của hai biến cố A và B nếu E chỉ xảy ra khi cả hai biến cố A và B cùng đồng thời xảy ra. Ký hiệu $E = A.B$ hoặc $E = AB$.

Biến cố A được gọi là tích của n biến cố $A_1, A_2, ..., A_n$ nếu A chỉ xảy ra khi cả n biến cố này cùng đồng thời xảy ra. Ký hiệu $A = A_1 A_2 ... A_n = \prod_{i=1}^{n} A_i$.

Ví dụ 1.3.1 Có hai danh sách lớp, mỗi danh sách đều có cả nam và nữ. Chọn ngẫu nhiên một sinh viên từ mỗi danh sách lớp. Gọi A là biến cố: chọn được nữ ở danh sách thứ nhất, B là biến cố: chọn được nữ ở danh sách thứ hai, E là biến cố: chọn được ít nhất một nữ, F là biến cố: chọn được cả hai nữ. Khi đó ta được $E = A + B$, và $F = AB$.

Biến cố đối

Biến cố B được gọi là biến cố đối của biến cố A nếu $A + B = U$ và $AB = V$. Ký hiệu $B = \overline{A}$.

Ví dụ 1.3.2 Trở lại ví dụ 1.3.1, nếu ta đặt M là biến cố: chọn được cả hai nam thì M là biến cố đối của biến cố E.

Các tính chất

$$A + A = A$$

$$A + B = B + A$$

$$A + (B + C) = (A + B) + C$$

$$\overline{A + B} = \overline{A}.\overline{B}$$

$$AB = BA$$

$$A(BC) = (AB)C$$

$$A(B + C) = AB + AC$$

$(A+B)(A+C)=A+BC$

$\overline{AB}=\overline{A}+\overline{B}$

1.4 QUAN HỆ GIỮA CÁC BIẾN CỐ

Biến cố xung khắc nhau

Hai biến cố A và B được gọi là xung khắc nhau nếu $AB=V$.

Nhóm n biến cố c được gọi là xung khắc từng đôi nếu $A_iA_j=V,\forall 1\le i\ne j\le n$.

Nhóm (họ) đầy đủ các biến cố

Nhóm các biến cố $A_1,A_2,...,A_n$ được gọi là một họ đầy đủ nếu chúng thỏa mãn hai điều kiện sau đây

(i) $A_iA_j=V,\forall 1\le i\ne j\le n$

(ii) $A_1+A_2+...+A_n=\Omega$

Ví dụ 1.4.1 Thực hiện phép thử gieo một con xúc xắc và ký hiệu $A_i,i=1,...,6$ là các biến cố mặt i chấm xuất hiện trong phép thử này. Khi đó Nhóm các biến cố $A_1,A_2,...,A_6$ tạo thành một nhóm đầy đủ.

Nhận xét: Hai biến cố đối nhau tạo thành họ đây đủ.

1.5 XÁC SUẤT CỦA BIẾN CỐ

Xác suất của một biến cố là một con số đặc trưng khả năng khách quan xuất hiện biến cố đó khi thực hiện phép thử. Ta có các kiểu định nghĩ khác nhau về xác suất của biến cố dưới đây.

Định nghĩa cổ điển về xác suất

Giả sử A là một biến cố nào đó trong một phép thử. Khi đó xác suất (probability) của A là

$$P(A):=\frac{m}{n}$$

Trong đó m là số các trường hợp có thể xảy ra A (số trường hợp thuận lợi cho A), n là số tất cả các trường hợp có thể xảy ra khi phép thử được thực hiện (số trường hợp đồng khả năng của phép thử).

Ví dụ 1.5.1 Thực hiện phép thử tung một đồng xu. Gọi S là biến cố mặt sấp xuất hiện, khi đó ta có xác suất của S là $P(S)=1/2$.

Ví dụ 1.5.2 Gieo một con xúc xắc (cân đối đồng chất). Gọi $A_i, i=1,...,6$ là biến cố mặt i chấm xuất hiện, khi đó ta được $P(A_1)=P(A_2)=...=P(A_6)=1/6$.

Tính chất:

$$0 \le P(A) \le 1$$

$$P(U)=1$$

$$P(V)=0$$

Nhận xét: Xác suất của biến cố càng gần số 1 thì khả năng biến cố xảy ra càng cao, xác suất của biến cố càng gần số 0 thì khả năng biến cố xảy ra càng thấp.

Định nghĩa thống kê về xác suất

Xác suất xuất hiện biến cố A trong một phép thử là một số p không đổi mà tần suất xuất hiện biến cố đó trong n phép thử sẽ dao động rất ít xung quanh nó khi số phép thử tăng lên vô hạn. Như vậy về mặt thực tế với số phép thử đủ lớn ta có thể lấy $P(A) \approx f(A) = \frac{k}{n}$*, với n là số phép thử và k là số lần xuất hiện biến cố A trong phép thử.*

Ví dụ 1.5.3 Để nghiên cứu khả năng xuất hiện mặt sấp khi tung một đồng xu, người ta tiến hành tung một đồng xu nhiều lần và thu được kết quả sau đây:

Người làm thí nghiệm	Số lần tung n	Số lần được mặt sấp k	Tần xuất f(A)
Buffon	4040	2048	0,5069
Pearson	12000	6019	0,5016
Pearson	24000	12012	0,5005

Qua ví dụ này ta có thể nói rằng khả năng (xác suất) xuất hiện mặt sấp là 0,5.

Định nghĩa hình học về xác suất

Xác suất để điểm ngẫu nhiên rơi vào phần A của B là

$$P(A) = \frac{S_A}{S_B}$$

Trong đó S_A, S_B lần lượt là độ đo hình học (diện tích, thể tích,…) của A và B.

Ví dụ 1.5.4 Hai người hẹn gặp nhau tại một quán café trong khoảng thời gian từ 8h đến 9 giờ sáng, mỗi người có thể đến chỗ hẹn vào bất kỳ thời điểm nào trong khoảng thời gian này. Hôm đó cả hai cùng có việc bận đột xuất nên giao hẹn rằng người nào đến trước sẽ chờ người kia và thời gian chờ không quá 20 phút. Tính xác suất để hai người gặp được nhau.

Trả lời: Xác suất để hai người gặp được nhau là 5/9.

Định nghĩa tiên đề về xác suất

Cho $\mathfrak{M} = \{A \mid A \subset \Omega\}$, với Ω là không gian mẫu gồm các biến cố của phép thử. Nếu ánh xạ

$P : \mathfrak{M} \to [0,1]$

thỏa mãn các tính chất sau đây:

(i) $0 \le P(\Omega) \le 1, \forall A \in \mathfrak{M}$

(ii) $P(\Omega) = 1$

(iii) $P(A+B) = P(A) + P(B), \forall AB = V$

Khi đó với mọi biến cố $A \subset \Omega$, giá trị $P(A)$ được gọi là xác suất của biến cố A.

Ba tính chất trên gọi là hệ tiên đề Kolmogorov.

Xác suất có điều kiện

Xác suất của biến cố phụ thuộc những gì được biết về kết quả phép thử.

Cho A, B là hai biến cố với $P(B) > 0$ *(biến cố B đã xảy ra). Khi đó xác suất có điều kiện của biến cố A với biến cố B đã xảy ra được xác định như sau:*

$$P(A|B) := \frac{P(AB)}{P(B)}$$

Ví dụ 1.5.5 Gieo mọt con xúc xắc. Goi M là biến cố mặt một chấm xuất hiện, L là biến cố mặt lẻ chấm xuất hiện. Khi đó $P(M|L) = \frac{P(ML)}{P(L)} = 1/3$ và $P(L|M) = \frac{P(LM)}{P(M)} = 1$

Các tính chất về xác suất của biến cố

- Với A và B là hai biến cố cho trước bất kỳ, ta có $P(A+B) = P(A) + P(B) - P(AB)$

Ví dụ 1.5.6 Trong một kết quả khảo sát gồm 100 người nữ có 60 người thích loại nước hoa E, 70 người thích loại nước hoa F, và 50 người thích cả hai loại nước hoa trên. Chọn ngẫu nhiên một người. Tính xác suất để người này thích ít nhất một trong hai loại nước hoa.

- Nếu A và B là hai biến cố xung khắc nhau thì $P(A+B) = P(A) + P(B)$

Ví dụ 1.5.7 Trong một thùng có 10 chi tiết máy trong đó có 2 chi tiết bị hỏng. Tìm xác suất để khi lấy ngẫu nhiên ra 6 chi tiết thì có không quá 1 chi tiết bị hỏng.

- Với mọi biến cố A, ta có $P(A) = 1 - P(\overline{A})$
- Giả sử $A_1, A_2, ..., A_n$ là các biến cố, khi đó

$$P(A_1 + A_2 + ... + A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k) - \cdots + (-1)^{n-1} P(A_1 ... A_n)$$

- Nếu các A_i xung khắc nhau thì

$$P(A_1 + A_2 + ... + A_n) = \sum_{i=1}^{n} P(A_i)$$

- Giả sử $A_1, A_2, ..., A_n$ là các biến cố, khi đó

$$P(A_1 A_2 ... A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) ... P(A_n | A_1 A_2 ... A_{n-1})$$

Ví dụ 1.5.8 Trong một hộp có 12 bóng đèn, trong đó có 3 bóng hỏng. Lấy ngẫu nhiên có thứ tự không hoàn lại 3 bóng để dùng. Tính xác suất để:

a. Cả 3 bóng đều hỏng.
b. Cả 3 bóng đều không hỏng.
c. Có ít nhất 1 bóng không hỏng.
d. Chỉ có bóng thứ hai hỏng.

- Nếu các biến cố $A_1, A_2, ..., A_n$ độc lập từng đôi thì

$$P(A_1A_2...A_n) = P(A_1)P(A_2)...P(A_n)$$

Ví dụ 1.5.9 Xác suất vi trùng kháng mỗi loại thuốc T_1, T_2, T_3 lần lượt là 5%, 10% và 20%. Nếu dùng cả 3 loại để diệt vi trùng thì xác suất để vi trùng bị tiêu diệt là bao nhiêu? Giả sử tác dụng tiêu diệt vi trùng của 3 loại thuốc độc lập nhau.

- Công thức xác suất đầy đủ: Cho $A, H_1, H_2, ..., H_n$ là các biến cố, trong đó nhóm $H_1, H_2, ..., H_n$ là một họ đầy đủ và $P(H_i) > 0$. Khi đó xác suất của biến cố A là

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A \mid H_i)$$

Ví dụ 1.5.10 Có 3 hộp giống nhau. Hộp thứ nhất đựng 10 sản phẩm trong đó có 6 chính phẩm, hộp thứ hai đựng 15 sản phẩm trong đó có 10 chính phẩm, hộp thứ ba đựng 20 sản phẩm trong đó có 15 chính phẩm. Lấy ngẫu nhiên một hộp và từ đó lấy ngẫu nhiên một sản phẩm. Tìm xác suất để lấy được chính phẩm.

- Công thức Bayes: Cho $A, H_1, H_2, ..., H_n$ là các biến cố, trong đó nhóm $H_1, H_2, ..., H_n$ là một họ đầy đủ, $P(H_i) > 0$, và $P(A) > 0$. Khi đó xác suất có điều kiện của từng H_i với biến cố A đã xảy ra là

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{P(A)} = \frac{P(H_i)P(A \mid H_i)}{\sum_{i=1}^{n} P(H_i)P(A \mid H_i)}$$

Ví dụ 1.5.11 Trước khi đưa sản phẩm ra thị trường, người ta đã phỏng vấn ngẫu nhiên 200 khách hàng về sản phẩm đó và thấy có 34 người trả lời “sẽ mua”, 96 người trả lời “có thể sẽ mua”, và 70 người trả lời “không mua”. Kinh nghiệm cho thấy khách hàng thực sự mua sản phẩm tương ứng với những cách trả lời trên là 40%, 20% và 10%.

a. Hãy đánh giá thị trường tiềm năng của sản phẩm đó.
b. Trong số khách hàng thực sự mua sản phẩm thì có bao nhiêu phần trăm trả lời “sẽ mua”?

Ví dụ 1.5.12 Hai người cùng bắn vào một mục tiêu. Khả năng bắn trúng của từng người là 0,8 và 0,9. Tìm xác suất:

a. Chỉ có một người bắn trúng
b. Có người bắn trúng mục tiêu
c. Cả hai người bắn trượt

ĐS: a. 0,26 b. 0,98 c. 0,02

Ví dụ 1.5.13 Tín hiệu thông tin được phát ba lần với xác suất thu được cử mỗi lần là 0,4.

a. Tìm xác suất để nguồn thu nhận được thông tin đó
b. Nếu muốn xác suất thu được thông tin lên đến 0,9 thì phải phát ít nhất bao nhiêu lần

ĐS: a. 0,784 b. 5 lần

Ví dụ 1.5.14 Có hai xạ thủ loại I và tám xạ thủ loại II, xác suất bắn trúng đích của các loại xạ thủ theo thứ tự là 0,9 và 0,8.

a. Chọn ngẫu nhiên ra một xạ thủ và xạ thủ đó bắn một viên đạn. Tìm xác suất để viên đạn đó trúng đích.
b. Nếu chọn ra hai xạ thủ và mỗi người bắn một viên thì khả năng để cả hai viên đều trúng đích là bao nhiêu?

ĐS: a. 0,82 b. 0,67

Ví dụ 1.5.15 Tỷ lệ người dân nghiện thuốc lá ở một vùng là 30%. Biết rằng tỷ lệ người viêm họng trong số người nghiện thuốc là 60%, còn tỷ lệ người bị viêm họng trong số người không hút thuốc lá là 40%.

a. Chọn ngẫu nhiên một người, biết rằng người đó viêm họng. Tính xác suất người đó nghiện thuốc.
b. Nếu người đó không bị viêm họng, tính xác suất để người đó nghiện thuốc.

ĐS: a. 0,3913 b. 0,222

Ví dụ 1.5.16 Một người tham gia đấu thầu hai dự án. Khả năng trúng thầu dự án thứ nhất là 0,6. Nếu trúng thầu dự án thứ nhất thì khả năng trúng thầu dự án thứ hai tăng lên là 0,8; còn nếu không trúng thầu dự án thứ nhất thì khả năng trúng thầu dự án thứ hai chỉ còn là 0,2. Tìm xác suất để người đó:

a. Trúng thầu cả hai dự án.
b. Chỉ trúng thầu một dự án.
c. Trúng thầu ít nhất một dự án.

ĐS: a. 0,48 b. 0,2 c. 0,68

Ví dụ 1.5.17 Tỷ lệ phế phẩm của một công ty là 5%. Trước khi đưa ra thị trường người ta sử dụng một thiết bị kiểm tra chất lượng để loại phế phẩm. Thiết bị kiểm tra có độ chính xác đối với chính phẩm là 90%, còn đối với phế phẩm là 99%.

a. Tìm tỷ lệ phế phẩm trong sản phẩm của công ty trên thị trường.
b. Tìm tỷ lệ chính phẩm bị loại.
c. Tìm tỷ lệ sai sót của thiết bị kiểm tra đó.

ĐS: a. 0,00058 b. 0,6574 c. 0,0955

Ví dụ 1.5.18 Ba người chơi bóng rổ, mỗi người ném một quả. Xác suất ném trúng rổ của mỗi người lần lượt là 0,5; 0,6 và 0,7. Tính xác suất để:

a. Cả ba người đều ném trúng rổ.
b. Có ít nhất một người ném trúng rổ.
c. Có ít nhất một người ném không trúng rổ.
d. Có đúng hai người ném trúng rổ.
e. Người thứ ba ném không trúng rổ, biết rằng có hai người ném trúng rổ.

Ví dụ 1.5.19 Trung tâm cứu nạn quốc gia nhận được tin báo là có một máy bay bị rơi. Theo đánh giá thì khả năng máy bay bị rơi ở vùng núi, vùng biển và vùng đồng bằng tương ứng là 0,6; 0,3 và 0,1. Khả năng tìm thấy máy bay rơi ở những nơi đó tương ứng là 0,2; 0,6 và 0,9.

a. Đầu tiên người ta cử ngay một đội tìm kiếm đến vùng núi và không tìm thấy máy bay rơi. Vậy khả năng máy bay rơi ở các vùng nói trên bằng bao nhiêu?

b. Người ta cử tiếp ba đội tìm kiếm khác đến tìm ở cả ba nơi và vẫn không thấy. Vậy khả năng máy bay rơi ở các vùng nói trên là bao nhiêu?

ĐS: a. 0,545; 0,341 và 0,114   b. 0,747; 0,234 và 0,019

Ví dụ 1.5.20 Trong một cơ quan điều tra, người ta dùng máy dó tìm tội phạm. Kinh nghiệm cho biết cứ 10 người bị tình nghi thì 7 người là tội phạm. Máy báo đúng người có tội với xác suất 0,85, máy báo sai người vô tội với xác suất 0,1. Một người được máy phân tích, hãy tính xác suất:

a. Người này là tội phạm.
b. Máy báo người này là tội phạm.
c. Người này thực sự có tội, biết rằng máy đã báo có tội.
d. Máy báo đúng.

ĐS: a. 0,7   b. 0,625   c. 0,952   d. 0,865

CHƯƠNG 2

BIẾN NGẪU NHIÊN VÀ QUY LUẬT PHÂN PHỐI XÁC SUẤT

2.1 ĐỊNH NGHĨA VÀ PHÂN LOẠI BIẾN NGẪU NHIÊN

Định nghĩa

Một biến số được gọi là biến ngẫu nhiên nếu trong kết quả của phép thử nó sẽ nhận một và chỉ một trong các giá trị có thể có của nó tùy thuộc vào sự tác động của các nhân tố ngẫu nhiên.

Ta thường sử dụng các chữ cái in: X, Y, Z, T, ... để ký hiệu cho biến ngẫu nhiên, và sử dụng các các chữ cái thường x, y, z, t, ... để ký hiệu giá trị của X, Y, Z, T, ...tương ứng.

Ví dụ 2.1.1 Gieo một con xúc xắc. Gọi X là số chấm xuất hiện, khi đó X là biến ngẫu nhiên và các kết quả có thể có của X là $X(\Omega)=\{1,2,3,4,5,6\}$.

Ví dụ 2.1.2 Giả sử có một xạ thủ bắn không hạn chế số lần vào một mục tiêu đến khi trúng mục tiêu thì dừng. Nếu gọi Y là số lần xạ thủ bắn trượt thì Y là một biến ngẫu nhiên với các giá trị nó có thể nhận được là $Y(\Omega)=\{0,1,2,3,\ldots\}$.

Ví dụ 2.1.3 Chọn ngẫu nhiên một người và đo chiều cao. Nếu gọi Z là chiều cao của người này thì Z là một biến ngẫu nhiên và giá trị có thể nhận được của Z là một đoạn trong trục số thực R.

Phân loại biến ngẫu nhiên

Chúng ta chia biến ngẫu nhiên thành hai loại: biến ngẫu nhiên rời rạc và biến ngẫu nhiên liên tục.

Biến ngẫu nhiên rời rạc là biến ngẫu nhiên mà các giá trị có thể có của nó lập nên một tập hữu hạn hoặc đếm được, có nghĩa là ta có thể liệt kê được tất cả các giá trị của nó.

Biến ngẫu nhiên liên tục là biến ngẫu nhiên mà các giá trị có thể có của nó lấp đầy một khoảng trên trục số, có nghĩa là ta không thể kiệt kê được tất cả các giá trị có thể xảy ra của nó.

Ví dụ 2.1.4 Các biến ngẫu nhiên X và Y trong các Ví dụ 2.1.1 và Ví dụ 2.1.2 là rời rạc, và biến ngẫu nhiên Z trong Ví dụ 2.1.3 là liên tục.

2.2 QUY LUẬT PHÂN PHỐI XÁC SUẤT CỦA BIẾN NGẪU NHIÊN

Định nghĩa

Quy luật phân phối xác suất của biến ngẫu nhiên là sự tương ứng giữa các giá trị có thể có của nó và các xác suất tương ứng với các giá trị đó.

Bảng phân phối xác suất

Chỉ sử dụng để mô tả quy luật phân phối xác suất của biến ngẫu nhiên rời rạc. Giả sử biến ngẫu nhiên rời rạc X với $X(\Omega)=\{x_1,x_2,...,x_n,...\}$ và $p_i=P(X=x_i)$. Khi đó bảng phân phối xác suất của X có dạng:

X	x_1	x_2	...	x_i	...	x_n	...
P	p_1	p_2	...	p_i	...	p_n	...

trong đó

$$\begin{cases} 0\le p_i\le 1, \forall i \\ \sum_{i\ge 1} p_i = 1 \end{cases}$$

Ví dụ 2.2.1 Trong phép thử gieo một con xúc xắc, ta có bảng phân xác suất của biến ngẫu nhiên X " số chấm xuất hiện" là

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

Ví dụ 2.2.2 Giả sử xác suất để xạ thủ bắn trúng mục tiêu là 0,8. Xạ thủ phát từng viên đạn để bắn cho đến khi trúng mục tiêu. Khi đó quy luật phân phối xác suất của biến ngẫu nhiên Y "số viên đạn được phát" là

Y	1	2	...	k	...
P	0,8	0,2.0,8	...	$(0,2)^{k-1}.0,8$	...

Hàm phân bố xác suất

Giả sử X là biến ngẫu nhiên bất kỳ (rời rạc hoặc liên tục), x là một số thực nào đó.

Hàm phân bố xác suất F(x) của X được định nghĩa là

$$F(x) = P(X < x)$$

Nhận xét:

- Khi X là biến ngẫu nhiên rời rạc, ta có

$$F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$$

- Khi X là biến ngẫu nhiên rời rạc với các giá trị $x_1, x_2, ..., x_n$, ta có

$$F(x) = \begin{cases} 0 & , x \le x_1 \\ ... & \\ p_1 + ... + p_i & , x_i < x \le x_{i+1} \\ ... & \\ 1, & , x > x_n \end{cases}$$

Ví dụ 2.2.3 Sử dụng nhận xét trên ta được hàm phân bố của biến ngẫu nhiên X trong Ví dụ 2.2.1 là

$$F(x) = \begin{cases} 0 & , x \le 1 \\ 1/6 & , 1 < x \le 2 \\ 2/6 & , 2 < x \le 3 \\ 3/6 & , 3 < x \le 4 \\ 4/6 & , 4 < x \le 5 \\ 5/6 & , 5 < x \le 6 \\ 1, & , x > 6 \end{cases}$$

Các tính chất:

- $0 \le F(x) \le 1, \forall x$
- $F(-\infty) = 0$
- $F(+\infty) = 1$
- X là biến ngẫu nhiên liên tục $\Rightarrow$ F(x) là hàm liên tục
- F(x) là hàm số không giảm và $P(a \le X < b) = F(b) - F(a)$
- X là biến ngẫu nhiên liên tục $\Rightarrow P(X = a) = 0$
- X là biến ngẫu nhiên liên tục $\Rightarrow P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$

Ví dụ 2.3.4 Cho biến ngẫu nhiên X có hàm phân bố xác suất:

$$F(x) = \begin{cases} 0 & ,x \le -1 \\ \frac{3}{4}x + \frac{3}{4} & ,-1 < x \le 1/3 \\ 1 & ,x > 1/3 \end{cases}$$

Khi đó $P(0 \le X < 1/3) = F(1/3) - F(0) = 1 - 3/4 = 1/4$

Hàm mật độ xác suất

Giả sử F(x) là hàm phân bố xác suất của biến ngẫu nhiên liên tục X.

Hàm mật độ của X được định nghĩa như sau: $f(x) := F'(x)$

Các tính chất:

- $f(x) \ge 0, \forall x$
- $P(a < X < b) = \int_a^b f(x)dx$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$
- $F(x) = \int_{-\infty}^{x} f(x)dx$

Chú ý: Điều kiện đủ để hàm số f(x) là hàm mật độ xác suất của một biến ngẫu nhiên liên tục nào đó là

$$\begin{cases} f(x) \geq 0, \forall x \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases}$$

Ví dụ 2.3.5 Cho biến ngẫu nhiên liên tục có hàm phân bố

$$F(x) = \begin{cases} 0 & ,x \leq 0 \\ \text{mx}^2 & ,0 < x \leq 1 \\ 1 & ,x > 1 \end{cases}$$

a. Hàm mật độ của X là

$$f(x) = F'(x) = \begin{cases} 0 & ,x \notin (0,1] \\ 2\text{mx}, & x \in (0,1] \end{cases}$$

b. Tham số m được xác định từ đẳng thức

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \Leftrightarrow \int_0^1 2mxdx = 1 \Rightarrow m = 1$$

Hoặc m được xác định từ tính liên tục của hàm F(x) tại $x = 1$.

Ví dụ 2.3.6 Tuổi thọ của một loại sản phẩm là niến ngẫu nhiên liên tục X với hàm mật độ

$$f(x) = \begin{cases} \dfrac{m}{x^2}, & x \geq 400 \\ 0 & ,x < 400 \end{cases}$$

a. Tìm tham số m.

b. Tính xác suất để lấy ngẫu nhiên 1 sản phẩm thì tuổi thọ của nó kéo dài ít nhất là 600 giờ.

2.3 CÁC THAM SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN

Kỳ vọng toán

Kỳ vọng toán của biến ngẫu nhiên X, ký hiệu E(X), được xác định như sau:

$$E(X) := \begin{cases} \sum_{i\geq 1} x_i p_i, \text{ nếu } X \text{ là bnn rời rạc với } X(\Omega) = \{x_1, x_2, ..., x_n, ...\} \\ \int_{-\infty}^{+\infty} xf(x)dx, \text{ nếu } X \text{ là bnn liên tục} \end{cases}$$

Kỳ vọng toán của một biến ngẫu nhiên phản ánh giá trị trung tâm của phân phối xác suất của niến ngẫu nhiên đó.

Ví dụ 2.3.1 Nghiên cứu về lương hưu của 400 công nhân ngành may với số liệu:

Lương (triệu đồng/năm)	8	9	10	11	12	13
Số công nhân	16	60	160	100	40	24

a. Tính E(X).

b. Tính tổng thu nhập của 400 công nhân trong một năm và so sánh bình quân thu nhập của 1 công nhân.

Các tính chất:

- $E(C) = C$, với C là hằng số
- $E(CX) = CE(X)$
- $E(X+Y) = E(X) + E(Y)$
- $E(aX+bY) = aE(X) + bE(Y)$
- $E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$
- Nếu X, Y là hai biến ngẫu nhiên độc lập thì $E(XY) = E(X)E(Y)$
- Nếu nhóm các biến ngẫu nhiên $X_1, X_2, ..., X_n$ độc lập thì $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

Ví dụ 2.3.2 Một dự án xây dựng được một viện thiết kế V soạn thảo cho hai bên A và B xét duyệt độc lập. Xác suất để A và B chấp nhận dự án khi xét duyệt là 0,7 và 0,8. Nếu chấp nhận dự án thì A phải trả cho V 4 triệu đồng, ngược lại thì phải trả 1 triệu đồng. Với B, nếu chấp nhận dự án thì phải trả cho V 10 triệu đồng, ngược lại thì phải trả 3 triệu đồng. Cho biết cho phí cho thiết kế là 10 triệu đồng và thuế 10% doanh thu. Hỏi viện V có nên nhận thiết kế theo thỏa thuận như trên hay không?

Ví dụ 2.3.3 Thống kê số khách trên một ô tô bus tại một tuyến giao thông thu được các số liệu sau:

Số khách trên một chuyến	20	25	30	35	40
Tần suất	0,2	0,3	0,15	0,1	0,25

a. Tìm kỳ vọng toán của số khách đi mỗi chuyến.
b. Giả sử chi phí cho mỗi chuyến xe là 200 ngàn đồng không phụ thuộc vào số khách đi trên xe thì để công ty xe bus có thể thu được lãi bình quân cho mỗi chuyến xe là 100 ngàn đồng thì phải quy định giá vé là bao nhiêu?

ĐS: 10,17

Ví dụ 2.3.4 Một người đi từ nhà đến cơ quan phải qua ba ngã tư, xác suất để người đó gặp đèn đỏ ở các ngã tư tương ứng là 0,2; 0,4 và 0,5. Hỏi thời gian trung bình phải ngừng trên đường là bao nhiêu, biết rằng mỗi khi gặp đèn đỏ người đó phải đợi khoảng 3 phút.

ĐS: 3,3 phút

Ví dụ 2.3.5 Nhu cầu hàng năm về loại hàng A là biến ngẫu nhiên liên tục X có hàm mật độ xác suất như sau (đơn vị: ngàn sản phẩm).

$$f(x)=\begin{cases} k(30-x), x\in(0,30) \\ 0 \qquad\qquad , x\notin(0,30) \end{cases}$$

a. Tìm k.
b. Tìm xác suất để nhu cầu về loại hàng đó không vượt quá 12000 sản phẩm trong một năm.
c. Tìm nhu cầu trung bình hàng năm về loại hàng đó.

ĐS: a. k=1/450 b. 0,64 c. 10

Ví dụ 2.3.6 Thời gian xếp hàng chờ mua hàng của khách là biến ngẫu nhiên liên tục với hàm phân bố xác suất như sau (đơn vị: phút).

$$F(x)=\begin{cases}0 & ,\ x\le 0\\ \mathrm{ax}^3-3x^2+2x & ,0<x\le 1\\ 1 & ,x>1\end{cases}$$

a. Tìm hệ số a.
b. Tìm thời gian xếp hàng trung bình.
c. Tìm xác suất để trong ba người xếp hàng thì có không quá hai người phải chờ quá 0,5 phút.

ĐS: a. 2 b. 0,5 c. 0,875

Phương sai và độ lệch chuẩn

Phương sai của một biến ngẫu X, ký hiệu Var(X), được xác định như sau:

$$Var(X):=E\left[X-E(X)\right]^2=\begin{cases}\sum\limits_{i\ge 1}\left[x_i-E(X)\right]^2 p_i,\ \textit{nếu } X \textit{ là bnn rời rạc với } X(\Omega)=\left\{x_1,x_2,...,x_n,...\right\}\\ \int\limits_{-\infty}^{+\infty}\left[x-E(X)\right]^2 f(x)dx,\ \textit{nếu } X \textit{ là bnn liên tục}\end{cases}$$

Phương sai của biến ngẫu nhiên phản ánh mức độ phân tán của các giá trị của biến ngẫu nhiên đó xung quanh giá trị kỳ vọng toán của nó.

Trong thực hành, ta thường dùng công thức sau đây của Var:

$$\mathrm{Var}(X):=E(X)^2-\left[E(X)\right]^2==\begin{cases}\sum\limits_{i\ge 1}x_i^2 p_i-\left[E(X)\right]^2,\ \text{nếu } X \text{ là bnn rời rạc với } X(\Omega)=\left\{x_1,x_2,...,x_n,...\right\}\\ \int\limits_{-\infty}^{+\infty}x^2 f(x)dx-\left[E(X)\right]^2,\ \text{nếu } X \text{ là bnn liên tục}\end{cases}$$

Ví dụ 2.3.7 Tỷ lệ mắc một loại bệnh trong một vùng dân cư là biến ngẫu nhiên liên tục có hàm mật độ:

$$f(x)=\begin{cases}\dfrac{1}{20}, x\in(0,30)\\ 0\ \ , x\notin(0,30)\end{cases}$$

a. Tính $P\left(|X-10|>2,5\right)$.

b. Tính tỷ lệ mắc bệnh trung bình và phương sai.

ĐS: a. 0,75 b. 15 và 33,3

Các tính chất:

- $Var(C)=0$, với C là hằng số
- $Var(C+X)=Var(X)$
- $Var(CX)=C^2Var(X)$
- Nếu X, Y là hai biến ngẫu nhiên độc lập thì $Var(X\pm Y)=Var(X)+Var(Y)$
- Nếu nhóm các biến ngẫu nhiên $X_1,X_2,...,X_n$ độc lập thì $Var\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}Var(X_i)$

Độ lệch chuẩn của biến ngẫu nhiên X, ký hiệu $\sigma(X)$, được xác định là: $\sigma(X)=\sqrt{\mathrm{Var(X)}}$.

Trung vị

Giá trị trung vị của biến ngẫu nhiên X, ký hiệu Med(X), được xác định như sau:

- *Nếu X là biến ngẫu nhiên rời rạc, $Med(X)=x_i$ nếu x_i thỏa mãn điều kiện*
$$F(x_i)\le 0,5<F(x_{i+1})$$
trong đó F(x) là hàm phân bố xác suất của X.
- *Nếu X là biến ngẫu nhiên liên tục, trung vị Med(X) của X sẽ là giá trị thỏa mãn điều kiện*
$$\int_{-\infty}^{Med(X)} f(x)dx=0,5$$
trong đó f(x) là hàm mật độ xác suất của X.

Giá trị trung vị của một biến ngẫu nhiên chia phân phối của biến ngẫu nhiên đó thành hai phần bằng nhau.

Nói chung, giá trị trung vị của biến ngẫu nhiên không duy nhất.

Mốt

Giá trị mốt của biến ngẫu nhiên X, ký hiệu Mod(X), được xác định như sau:

- *Nếu X là biến ngẫu nhiên rời rạc, Mod(X) là giá trị x_i của X có p_i cực đại.*

- *Nếu X là biến ngẫu nhiên liên tục, Mod(X) là giá trị x tại đó hàm mật độ f(x) đạt cực đại.*

Nói chung, giá trị mốt của biến ngẫu nhiên không duy nhất.

Ví dụ 2.3.8 Một thiết bị gồm 3 bộ phận hoạt động độc lập với nhau. Xác suất trong thời gian t các bộ phận bị hỏng tương ứng là 0,4; 0,2 và 0,3.

a. Tìm quy luật phân phối xác suất của số bộ phận bị hỏng X.
b. Thiết lập hàm phân bố xác suất của X.
c. Tính xác suất trong thời gian t có không quá hai bộ phận bị hỏng.
d. Tìm mốt và trung vị của X.

ĐS: c. 0,976 d. 1

Hệ số biến thiên

Hệ số biến thiên của biến ngẫu nhiên X, ký hiệu CV(X), được xác định như sau:

$$CV(X) = \frac{\sigma(X)}{|E(X)|}.100 \quad (\%)$$

Hệ số biến thiên được dùng để đo mức độ thuần nhất của một phân phối (hệ số biến thiên có giá trị càng bé thì tính thuần nhất của số liệu càng lớn), và nó cũng được sử dụng để so sánh mức độ phân tán của hai phân phối mà kỳ vọng toán và độ lệch chuẩn của chúng khác nhau.

Hệ số tương quan

Hệ số tương quan của hai biến ngẫu nhiên X và Y, ký hiệu r(X,Y), được xác định như sau:

$$r(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)Var(Y)}}$$

Hệ số tương quan đánh giá mức độ liên hệ của X và Y

- $r(X,Y) = 0$: X và Y không tương quan nhau
- $0 < r(X,Y) < 1$: X và Y tương quan thuận
- $r(X,Y) = 1$: X và Y tương quan tuyến tính thuận

- $-1 < r(X,Y) < 0$: X và Y tương quan nghịch
- $r(X,Y) = -1$: X và Y tương quan tuyến tính nghịch

Giá trị tới hạn

Giá trị tới hạn mức α của biến ngẫu nhiên X, ký hiệu x_α, là giá trị thỏa mãn:

$$P(X > x_\alpha) = \alpha$$

CHƯƠNG 3

MỘT SỐ QUY LUẬT PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

3.1 QUY LUẬT PHÂN PHỐI NHỊ THỨC

Định nghĩa

Biến ngẫu nhiên rời rạc X với $X(\Omega)=\{0,1,\ldots,n\}$ *được gọi là phân phối theo quy luật nhị thức với tham số n và p nếu tồn tại số* $0<p<1$ *sao cho:*

$$P(X=k)=C_n^k p^k(1-p)^{n-k}, k\in X(\Omega)$$

Ký hiệu $X\sim B(n,p)$.

Ví dụ 3.1.1 Giả sử ta có lược đồ Bernoulli, tức là tiến hành *n* phép thử độc lập, trong mỗi phép thử chỉ có hai trường hợp, hoặc biến cố *A* xuất hiện, hoặc biến cố *A* không xuất hiện, xác suất xuất hiện biến cố *A* trong mỗi phép thử đều bằng p. Nếu gọi X là biến ngẫu nhiên: số lần xuất hiện biến cố *A* trong *n* phép thử độc lập nói trên thì X phân phối theo quy luật nhị thức với tham số *n* và *p*.

Các tham số đặc trưng

- $E(X)=np$
- $Var(X)=np(1-p)$
- $np+p-1\le Mod(X)\le np+p$

Ví dụ 3.1.2 Một nữ công nhân quản lý 12 náy dệt. Xác suất để mỗi máy trong khoảng thời gian t cần đến sự chăm sóc của nữ công nhân là 1/3. Tính xác suất:

a. Trong khoảng thời gian t có 4 máy cần đến sự chăm sóc của nữ công nhân.
b. Trong khoảng thời gian t có từ 3 đến 6 máy cần đến sự chăm sóc của nữ công nhân.

ĐS: a. 0,238 b. 0,751

Ví dụ 3.1.3 Việc sản xuất ra các sản phẩm được tiến hành độc lập. Hỏi phải sản xuất mỗi đợt bao nhiêu sản phẩm để trung bình có được 10 sản phẩm đạt tiêu chuẩn, biết rằng xác suất được sản phẩm đạt tiêu chuẩn là 0,8.

Ví dụ 3.1.4 Một kho hàng chuyên cung cấp hàng cho 12 cửa hàng. Xác suất để mỗi cửa hàng đặt hàng cho kho đó trong ngày là 0,3. Tìm số đơn đặt hàng có khả năng nhiều nhất cho một ngày và xác suất tương ứng với nó.

ĐS: 3 và 0,2397

Ví dụ 3.1.5 Trên một chuyến bay người ta dùng loại máy bay ATR 72 có 72 chỗ ngồi. Thực tế cho thấy đến giờ chót vẫn có khách bỏ chuyến bay. Để tận dụng hết chỗ ngồi bằng cách bán thêm vé dự phòng, người ta đã thống kê 20 chuyến bay và thu được các số liệu sau:

Số khách bỏ chuyến bay	2	3	4	5	6	7	8	9	10	11
Số chuyến tương ứng	1	4	0	4	2	5	1	1	0	2

Hãy ước lượng xác suất để trong một chuyến bay nào đó có:

a. Một hành khách bỏ chuyến bay đó.
b. Hai hành khách bỏ chuyến bay đó.
c. Tìm số hành khách bỏ chuyến trung bình ở mỗi chuyến bay.

Ví dụ 3.1.6 Xác suất để máy bị hỏng trong một ngày hoạt động là 0,01. Mỗi lần máy hỏng chi phí sửa chữa khoảng một triệu đồng. Vậy có nên ký một hợp đồng bảo dưỡng thường xuyên với chi phí là 120 ngàn đồng một tháng để giảm xác suất máy hỏng xuống còn một nửa hay không? Biết rằng một năm máy hoạt động 300 ngày.

ĐS: Nên ký hợp đồng bảo dưỡng

3.2 QUY LUẬT PHÂN PHỐI SIÊU BỘI

Định nghĩa

Biến ngẫu nhiên rời rạc X với $X(\Omega)=\{0,1,...,n\}$ *được gọi là phân phối theo quy luật siêu bội với tham số n và N nếu tồn tại số M và N sao cho:*

$$P(X=k)=\frac{C_M^k.C_{N-M}^{n-k}}{C_N^n},k\in X(\Omega)$$

Ký hiệu $X \sim H(N,n)$.

Ví dụ 3.2.1 Giả sử trong bình có *N* quả cầu trong đó có *M* quả cầu trắng và *N* - *M* quả cầu đen. Lấy ngẫu nhiên lần lượt ra *n* quả cầu theo phương thức không hoàn lại. Nếu gọi X là biến ngẫu nhiên: số lần lấy được quả cầu trắng trong *n* lần như trên thì *X* phân phối theo quy luật siêu bội với tham số *n* và *N*.

Nhận xét: Ta có

$$\lim_{\frac{n}{N}\to 0}\frac{C_M^k.C_{N-M}^{n-k}}{C_N^n}=C_n^k.\left(\frac{M}{N}\right)^k.\left(1-\frac{M}{N}\right)^{n-k}$$

Do đó khi số n là rất bé so với số N thì phân phối siêu bội thực tế không khác biệt so với phân phối nhị thức, điều đó có nghĩa là khi tỷ số $\frac{n}{N}$ rất nhỏ ($n<0,1N$) thì phương pháp lấy không hoàn lại và có hoàn lại xem như không khác nhau. Do đó có thể lấy theo phương thức không hoàn lại nhưng vẫn có thể tính toán như trong trường hợp lấy có hoàn lại cho đơn giản.

Các tham số đặc trưng

- $E(X)=np$
- $Var(X)=np(1-p).\frac{N-n}{N-1}$

với $p=\frac{M}{N}$, $\frac{N-n}{N-1}$ được gọi là hệ số điều chỉnh.

Ví dụ 3.2.2 Để thanh toán một triệu đồng tiền hàng, một khách hàng gian lận đã xếp lẫn 5 tờ 50 ngàn tiền giả với 15 tờ tiền thật. Chủ cửa hàng rút ngẫu nhiên 3 tờ giấy bạc đem đi kiểm tra và giao hẹn nếu phát hiện có bạc giả thì cứ mỗi tờ giả khách hàng phải đền hai tờ thật. Tìm số tiền mà khách có thể phải trả.

ĐS: 75 ngàn

3.3 QUY LUẬT PHÂN PHỐI POISSON

Định nghĩa

Biến ngẫu nhiên rời rạc X với $X(\Omega)=\{0,1,...,n,...\}$ *được gọi là phân phối theo quy luật poisson với tham số* λ *nếu tồn tại số* $\lambda>0$ *sao cho:*

$$P(X=k)=\frac{\lambda^k}{k!}.e^{-\lambda},k\in X(\Omega)$$

Ký hiệu $X\sim P(\lambda)$.

Nhận xét: Giả sử $X\sim B(n,p)$. Khi n rất lớn ($n\geq 20$) và p rất bé ($p\leq 0,1$), và $np=\lambda$ không đổi, ta có xấp xỉ sau đây:

$$P(X=k)=C_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}.e^{-\lambda}$$

Các tham số đặc trưng

- $E(X)=\lambda$
- $Var(X)=\lambda$
- $\lambda-1\leq Mod(X)\leq\lambda$

Ví dụ 3.3.1 Số khách hàng vào một cửa hàng bách hóa trong một giờ là biến ngẫu nhiên tuân theo quy luật Poisson với mật độ (số khách trung bình) là 8 khách hàng trong một giờ. Tìm xác suất để trong một giờ nào đó có hơn 4 khách vào.

ĐS: 0,9

Ví dụ 3.3.2 Cứ 5000 con cá biển đánh bắt được thì có một con bị nhiễm khuẩn có hại cho sức khỏe con người. Tìm xác suất để trong một lô cá gồm 1800 con mới được đánh bắt về, có không quá hai con bị nhiễm khuẩn.

3.4 QUY LUẬT PHÂN PHỐI CHUẨN

Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là phân phối theo quy luật chuẩn với các tham số μ *và* σ^2 *nếu hàm mật độ của X có dạng*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Kí hiệu $X \sim N\left(\mu,\sigma^2\right)$.

Các tham số đặc trưng

- $E(X) = \mu$
- $Var(X) = \sigma^2$
- $Mod(X) = Med(X) = \mu$

Nhận xét:

- Chuẩn hóa biến ngẫu nhiên $X \sim N\left(\mu,\sigma^2\right)$: Ký hiệu $U = \frac{X-\mu}{\sigma}$, khi đó $U \sim N(0,1)$ được gọi là phân phối theo quy luật chuẩn hóa, và hàm mật độ của U là

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

- Giả sử $X \sim N\left(\mu,\sigma^2\right)$, khi đó xác suất để giá trị của X rơi vào khoảng *(a,b)* là

$$P(a < X < b) = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right), \text{ với } \phi(u) = \frac{1}{\sqrt{2\pi}}\int_0^u e^{-\frac{t^2}{2}} dt$$

  Chú ý rằng $\phi(u) \approx \phi(5) = 0{,}5$ với mọi $u > 5$.

- $X \sim N\left(\mu,\sigma^2\right) \Rightarrow P\left(\left|X-\mu\right| < \varepsilon\right) = 2\phi\left(\frac{\varepsilon}{\sigma}\right)$

- Giả sử các biến ngẫu nhiên $X_1,...,X_n$ độc lập lẫn nhau và $X_i \sim N\left(\mu_i,\sigma_i^2\right), i=\overline{1,n}$.

  Khi đó

  $$X=\sum_{i=1}^{n} X_i \sim N\left(E(X)=\sum_{i=1}^{n}\mu_i, Var(X)=\sum_{i=1}^{n}\sigma_i^2\right)$$

- Giả sử các biến ngẫu nhiên $X_1,...,X_n$ độc lập lẫn nhau, có cùng phân phối và $n \geq 30$.

  Khi đó

  $$X=\sum_{i=1}^{n} X_i \approx N\left(E(X)=\sum_{i=1}^{n}E\left(X_i\right), Var(X)=\sum_{i=1}^{n}Var\left(X_i\right)\right)$$

- Giả sử $X \sim B(n,p)$. Nếu $n>5$ và $\left|\sqrt{\frac{p}{1-p}}-\sqrt{\frac{1-p}{p}}\right|<0,3\sqrt{n}$ thì

  $$X \sim N\left(\mu=np, \sigma^2=np(1-p)\right)$$

  và

  $$P(X=K) \approx \frac{1}{\sqrt{np(1-p)}}\varphi\left(\frac{k-np}{\sqrt{np(1-p)}}\right),\ P(k_1 \leq X \leq k_2) \approx \phi\left(\frac{k_2-np}{\sqrt{np(1-p)}}\right)-\phi\left(\frac{k_1-np}{\sqrt{np(1-p)}}\right)$$

- $X \sim P(\lambda), \lambda>20 \Rightarrow X \approx N\left(\mu=\lambda, \sigma^2=\lambda\right)$
- Quy tắc *k*-xích ma:

  $X \sim N\left(\mu,\sigma^2\right) \Rightarrow P\left(|X-\mu|<k\sigma\right)=2\phi(k)$

  K=2: $P\left(|X-\mu|<2\sigma\right)=2\phi(2)=0,9544$ (quy tắc 2-xích ma)

  K=3: $P\left(|X-\mu|<3\sigma\right)=2\phi(3)=0,9974$ (quy tắc 3-xích ma)
- Nếu ta chưa biết quy luật phân phối của X nhưng nó thỏa mãn quy tắc 2-xích ma hoặc 3-xích ma thì có thể xem X có phân phối chuẩn.

Ví dụ 3.4.1 Tìm xác suất để biến ngẫu nhiên phân phối chuẩn hóa nhận giá trị:

a. Trong khoảng (-2,33;2,33).
b. Trong khoảng (-2;1).
c. Lớn hơn 3,02.
d. Nhỏ hơn 2,5.

Ví dụ 3.4.2 Việc tiêu dùng điện hàng tháng của các hộ gia đình ở Hà Nội là biến ngẫu nhiên phân phối chuẩn với trung bình là 200 KWh và độ lệch chuẩn là 40 KWh. Tìm xác suất để chọn ngẫu nhiên một hộ gia đình thì hộ đó:

a. Có mức tiêu dùng điện hàng tháng trên 250 KWh.
b. Có mức tiêu dùng điện hàng tháng dưới 180 KWh.

Ví dụ 3.4.3 Kích thước chi tiết là biến ngẫu nhiên phân phối chuẩn với kích thước trung bình là 50 cm. Kích thước thực tế của các chi tiết không nhỏ hơn 32 cm và không lớn hơn 68 cm. Tìm xác suất để lấy ngẫu nhiên một chi tiết có kích thước:

a. Lớn hơn 55 cm.
b. Nhỏ hơn 40 cm.

ĐS: a. 0,0823 b. 0,0027

Ví dụ 3.4.4 Tiến hành kiểm tra chất lượng 900 chi tiết. Xác suất được chi tiết đạt tiêu chuẩn là 0,9. Hãy tìm với xác suất 0,9544, xem số chi tiết đạt tiêu chuẩn nằm trong khoảng nào xung quanh số chi tiết đạt tiêu chuẩn trung bình?

ĐS: (792;828)

Ví dụ 3.4.5 Tiến hành kiểm tra chất lượng 900 chi tiết. Xác suất được chi tiết đạt tiêu chuẩn là 0,9. Hãy tìm với xác suất 0,9544, xem số chi tiết đạt tiêu chuẩn nằm trong khoảng nào xung quanh số chi tiết đạt tiêu chuẩn trung bình?

ĐS: (792;828)

Ví dụ 3.4.6 Tuổi thọ của một loại thiết bị điện là biến ngẫu nhiên phân phối chuẩn với trung bình là 1500 giờ và độ lệch chuẩn là 150 giờ. Nếu thiết bị hỏng trước 1200 giờ thì nhà máy phải bảo hành miễn phí.

a. Tìm tỷ lệ sản phẩm phải bảo hành.
b. Phải quy định thời gian bảo hành là bao nhiêu để tỷ lệ sản phẩm phải bảo hành chỉ còn 1%

ĐS: a. 0,0228 b. 1150,5

Ví dụ 3.4.7 Khi xâm nhập một thị trường mới, doanh nghiệp chỉ sự kiến được rằng doanh số hàng tháng có thể đạt được phân phối tuân theo quy luật chuẩn. Khả năng đạt được doanh số trên 40 triệu là 0,2 và dưới 25 triệu là 0,1. Tìm xác suất để doanh nghiệp đạt được doanh số ít nhất là 32 triệu/tháng.

ĐS: 0,6141

Ví dụ 3.4.8 Xác suất để sản phẩm sau khi sản xuất không được kiểm tra chất lượng bằng 0,2. Tìm xác suất để trong 400 sản phẩm được sản xuất ra có:

a. 80 sản phẩm không được kiểm tra chất lượng.
b. Có từ 70 đến 100 sản phẩm không được kiểm tra chất lượng.

ĐS: a. 0,04986 b. 0,8882

Ví dụ 3.4.9 Một nhà sản xuất cần mua một loại gioăng cao su có độ dày từ 0,118 cm đến 0,122 cm. Có hai cửa hàng cùng bán loại gioăng này với độ dày là các biến ngẫu nhiên phân phối chuẩn với các đặc trưng được cho trong bảng sau:

	Độ dày trung bình	Độ lệch chuẩn	Giá bán
Cửa hàng A	0,12	0,001	3 USD/hộp/1000 cái
Cửa hàng B	0,12	0,0015	2,6 USD/hộp/1000 cái

Hỏi nhà sản xuất nên mua gioăng của cửa hàng nào?

ĐS: Cửa hàng A

Ví dụ 3.4.10 Lãi suất đầu tư vào hai thị trường A và B là các biến ngẫu nhiên độc lập và cùng phân phối chuẩn với trung bình là 10% và 9%; độ lệch chuẩn là 4% và 3%.

a. Muốn có lãi suất trên 8% thì nên chọn phương án nào trong các phương án sau:

   Phương án 1: Đầu tư toàn bộ vào A.

   Phương án 2: Đầu tư toàn bộ vào B.

   Phương án 3: Chia đều vốn vào hai thị trường.

b. Nếu muốn rủi ro về lãi suất là nhỏ nhất thì nên đầu tư như thế nào?

ĐS: a. Nên chọn phương án 3 b. Nên chọn phương án 3

3.5 QUY LUẬT PHÂN PHỐI CHI BÌNH PHƯƠNG

Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là phân phối theo quy luật Chi bình phương với n bậc tự do nếu hàm mật độ của nó có dạng

$$f(x)=\begin{cases}0 & ,x\le 0\\ \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}e^{-\frac{x}{2}}.x^{\frac{x}{2}-1} & ,x>0\end{cases}$$

Ký hiệu $X\sim\chi^2(n)$, trong đó hàm Gamma $\Gamma(x)=\int_0^{+\infty}t^{x-1}e^{-t}dt$.

Các tham số đặc trưng

- $E(X)=n$
- $Var(X)=2n$

Nhận xét:

- $X\sim N(0,1)\Rightarrow X^2\sim\chi^2(1)$
- Giả sử $X_1\sim N(0,1),...,X_k\sim N(0,1)$ và các $X_1,...,X_k$ độc lập lẫn nhau. Khi đó $X=X_1^2+...+X_k^2\sim\chi^2(k)$
- Giả sử $X_1\sim\chi^2(n_1),...,X_k\sim\chi^2(n_k)$ và các $X_1,...,X_k$ độc lập lẫn nhau. Khi đó $X=X_1+...+X_k\sim\chi^2(n_1+...+n_k)$

3.6 QUY LUẬT PHÂN PHỐI STUDENT

Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là phân phối theo quy luật Student với n bậc tự do nếu hàm mật độ của nó có dạng

$$f(x)=\frac{\Gamma\left(\dfrac{n}{2}\right)}{\sqrt{(n-1)\pi}.\Gamma\left(\dfrac{n-1}{2}\right)}\left(1+\frac{x^2}{n-1}\right)^{-\frac{n}{2}}\quad,\forall x$$

Ký hiệu $X \sim T(n)$, trong đó hàm Gamma $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$.

Các tham số đặc trưng

- $E(X) = 0$
- $Var(X) = \dfrac{n}{n-2}$

Nhận xét:

Giả sử $X \sim N(0,1), Y \sim \chi^2(n)$ và X, Y độc lập nhau . Khi đó $\dfrac{X}{\sqrt{\dfrac{Y}{n}}} \sim T(n)$.

3.7 QUY LUẬT PHÂN PHỐI FISHER-SNEDECOR

Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là phân phối theo quy luật Fisher-Snedecor với n_1 và n_2 bậc tự do nếu hàm mật độ của nó có dạng

$$f(x) = \begin{cases} 0 & ,x \le 0 \\ C.\dfrac{x^{\frac{n_1-n_2}{2}}}{(n_1x+n_2)^{\frac{n_1-n_2}{2}}} & ,x > 0 \end{cases}$$

Ký hiệu $X \sim F(n_1, n_2)$, trong đó $C = \dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right).n_1^{\frac{n_1}{2}}.n_2^{\frac{n_2}{2}}}{\Gamma\left(\dfrac{n_1}{2}\right).\Gamma\left(\dfrac{n_2}{2}\right)}$ và hàm Gamma $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$.

Các tham số đặc trưng

- $E(X) = \dfrac{n_2}{n_2-2}$
- $Var(X) = \dfrac{2n_2^2\left(n_1+n_2^2-2\right)}{n_1\left(n_2-1\right)^2\left(n_2-4\right)}$

Nhận xét:

- $X \sim \chi^2(n_1), Y \sim \chi^2(n_2) \Rightarrow \dfrac{\frac{X}{n_1}}{\frac{Y}{n_2}} \sim F(n_1, n_2)$
- $X \sim T(n) \Rightarrow X^2 \sim F(1, n)$

3.8 QUY LUẬT PHÂN PHỐI ĐỀU

Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là phân phối theo quy luật đều trong khoảng (a,b) nếu hàm mật độ của nó có dạng

$$f(x) = \begin{cases} \dfrac{1}{b-a} & , x \in (a,b) \\ 0 & , x \notin (a,b) \end{cases}$$

Ký hiệu $X \sim U(a,b)$.

Các tham số đặc trưng

- $E(X) = \dfrac{a+b}{2}$
- $Var(X) = \dfrac{(b-a)^2}{12}$

3.9 QUY LUẬT PHÂN PHỐI LŨY THỪA

Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là phân phối theo quy luật lũy thừa nếu hàm mật độ của nó có dạng

$$f(x) = \begin{cases} \lambda . e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases} \quad \text{(hằng số } \lambda > 0\text{)}$$

Ký hiệu $X \sim E(\lambda)$.

Các tham số đặc trưng

- $E(X) = \frac{1}{\lambda}$
- $Var(X) = \frac{1}{\lambda^2}$

Nhận xét:

- Hàm phân bố $F(x) = 1 - e^{-\lambda x}$.
- $P\left(a < X < b\right) = e^{-a\lambda} - e^{-b\lambda}$

CHƯƠNG 4

ƯỚC LƯỢNG THAM SỐ CỦA BIẾN NGẪU NHIÊN

4.1 ĐỊNH NGHĨA THỐNG KÊ TOÁN

Để nghiên cứu biến ngẫu nhiên gốc X trong tổng thể theo phương pháp chọn mẫu, người ta tiến hành n quan sát độc lập về biến ngẫu nhiên đó. Gọi $X_i, i = 1,2,...,n$ là quan sát thứ i về biến ngẫu nhiên X. khi đó $W = (X_1,...,X_n)$ được gọi là mẫu ngẫu nhiên và hàm $G = f(X_1,...,X_n)$ được gọi là thống kê của X. Như vậy thống kê toán là một biến ngẫu nhiên, và khi mẫu ngẫu nhiên nhận một giá trị cụ thể $w = (x_1,...,x_n)$ thì G cũng nhận một giá trị cụ thể $g = f(x_1,...,x_n)$.

4.2 PHƯƠNG PHÁP ƯỚC LƯỢNG BẰNG KHOẢNG TIN CẬY

Khi kích thước mẫu nhỏ thì việc áp dụng phương pháp ước lượng điểm để ước lượng tham số θ của biến ngẫu nhiên gốc X trở nên không hiệu quả. Khi đó người ta sử dụng phương pháp ước lượng bằng khoảng tin cậy như sau: Từ mẫu ngẫu nhiên gốc, xây dựng mẫu

$W = (X_1,...,X_n)$, thống kê $G = f(X_1,...,X_n)$, $G_1 = f_1(X_1,...,X_n)$ và $G_2 = f_2(X_1,...,X_n)$ sao cho với xác suất $1-\alpha$ cho trước, tham số θ sẽ rơi vào khoảng (G_1, G_2). Nếu với xác suất $1-\alpha$ cho trước thỏa mãn điều kiện $P(G_1 < \theta < G_2) = 1-\alpha$ thì $1-\alpha$: độ tin cậy của ước lượng, $I = G_1 - G_2$: độ dài khoảng tin cậy, (G_1, G_2): khoảng tin cậy.

Ước lượng tham số $\mu = E(X)$ của $X \sim N(\mu, \sigma^2)$

Trường hợp 1: Đã biết σ^2

Lập mẫu $W = (X_1,...,X_n)$ và xét thống kê $G = U = \dfrac{\overline{X} - E(\overline{X})}{Var(\overline{X})} = \dfrac{\overline{X} - \mu}{\sigma}.\sqrt{n} \sim N(0,1)$.

Với độ tin cậy $1-\alpha$, tồn tại α_1, α_2 sao cho $\alpha_1 + \alpha_2 = \alpha$ và $u_{1-\alpha_1}, u_{\alpha_2}$ sao cho

$$P\left(U < u_{1-\alpha_1}\right) = \alpha_1, P\left(U > u_{\alpha 2}\right) = \alpha_2$$

Khi đó

$$P\left(u_{1-\alpha_1} < U < u_{\alpha_2}\right) = 1-\alpha \Leftrightarrow P\left(-u_{\alpha_1} < U < u_{\alpha_2}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\overline{X} - \frac{\sigma}{\sqrt{n}} u_{\alpha_2} < \mu < \overline{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha_1}\right) = 1-\alpha$$

Đối với mẫu cụ thể $w = \left(x_1, ..., x_n\right)$ ta có

$$\overline{x} - \frac{\sigma}{\sqrt{n}} u_{\alpha_2} < \mu < \overline{x} + \frac{\sigma}{\sqrt{n}} u_{\alpha_1}$$

Các khoảng tin cậy của μ:

- Khoảng tin cậy bên phải: $\left(\overline{x} - \frac{\sigma}{\sqrt{n}} u_\alpha, +\infty\right)$
- Khoảng tin cậy bên trái: $\left(-\infty, \overline{x} + \frac{\sigma}{\sqrt{n}} u_\alpha\right)$
- Khoảng tin cậy đối xứng: $\left(\overline{x} - \frac{\sigma}{\sqrt{n}} u_{\alpha/2}, \overline{x} + \frac{\sigma}{\sqrt{n}} u_{\alpha/2}\right)$

Trong trường hợp này $\varepsilon = \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$ được gọi là độ chính xác của ước lượng, và khi đó kích thước mẫu tối thiểu để $\varepsilon \le \varepsilon_0$ cho trước là $n \ge n_0 = \left[\frac{\sigma^2}{\varepsilon_0^2} . u_{\alpha/2}^2\right] + 1$.

Trường hợp 2: Chưa biết σ^2 và $n \ge 30$

Xét thống kê $G = U = \frac{\overline{X} - \mu}{S} . \sqrt{n} \sim N\left(0,1\right)$, lập luận tương tự như trên ta có các khoảng tin cậy là:

- Khoảng tin cậy bên phải:

$$\left(\overline{x} - \frac{s}{\sqrt{n}} u_\alpha, +\infty\right)$$

- Khoảng tin cậy bên trái:

$$\left(-\infty, \overline{x} + \frac{s}{\sqrt{n}} u_\alpha\right)$$

- Khoảng tin cậy đối xứng:

$$\left(\overline{x}-\frac{s}{\sqrt{n}}u_{\alpha/2},\overline{x}+\frac{s}{\sqrt{n}}u_{\alpha/2}\right)$$

Trường hợp 3: Chưa biết σ^2 và $n<30$

Xét thống kê $G=U=\frac{\overline{X}-\mu}{S}.\sqrt{n}\sim T(n-1)$, lập luận tương tự như trên ta có các khoảng tin cậy là:

- Khoảng tin cậy bên phải:

$$\left(\overline{x}-\frac{s}{\sqrt{n}}t_{\alpha}^{(n-1)},+\infty\right)$$

- Khoảng tin cậy bên trái:

$$\left(-\infty,\overline{x}+\frac{s}{\sqrt{n}}t_{\alpha}^{(n-1)}\right)$$

- Khoảng tin cậy đối xứng:

$$\left(\overline{x}-\frac{s}{\sqrt{n}}t_{\alpha/2}^{(n-1)},\overline{x}+\frac{s}{\sqrt{n}}t_{\alpha/2}^{(n-1)}\right)$$

Ví dụ 4.2.1 Với độ tin cậy 95%, hãy ước lượng xăng hao phí trung bình cho một ô tô chạy từ A đến B nếu chạy thử 30 lần trên đoạn dường này, người ta ghi nhận được lượng xăng hao phí như sau:

Lượng xăng hao phí (lít)	Số lần tương ứng
9,6 - 9,8	3
9,8 - 10,0	5
10,0 - 10,2	10
10,2 - 10,4	8
10,4 - 10,6	4

	n=30

Biết rằng lượng xăng hao phí là biến ngẫu nhiên tuân theo quy luật chuẩn.

ĐS: (10,1238;10,1428)

Ví dụ 4.2.2 Có số liệu về trọng lượng của một loại trứng gà như ở bảng dưới đây. Bằng khoảng tin cậy đối xứng hãy ước lượng trọng lượng trung bình của loại trứng gà này với độ tin cậy 0,95. Giả thiết trọng lượng trứng gà là biến ngẫu nhiên tuân theo quy luật chuẩn.

Trọng lượng trứng (gam)	Số quả trứng tương ứng
25 – 30	2
30 – 35	3
35 – 40	10
40 – 45	8
45 – 50	2
	n=25

ĐS: (36,35;40,65)

Ước lượng tham số *p* của $X \sim B(n,p)$

Giả sử $X \sim B(n,p)$ và $f = \frac{X}{n}$. Nếu $n > 5$ và $\left|\sqrt{\frac{p}{1-p}} - \sqrt{\frac{1-p}{p}}\right| < 0,3\sqrt{n}$ thì

$$f \approx N\left(p, \frac{p(1-p)}{n}\right)$$

Xét thống kê $G = U = \frac{f - E(f)}{Var(f)} = \frac{f - p}{\sqrt{p(1-p)}}.\sqrt{n} \sim N(0,1)$, khi đó khoảng tin cậy của p đối với một mẫu cụ thể nào đấy là: (p_1, p_2)

Với

$$p_1 = \frac{nf + u_{\alpha/2}^2 + u_{\alpha/2}.\sqrt{nf(1-f) + \frac{1}{4}u_{\alpha/2}^2}}{n + u_{\alpha/2}^2}$$

$$p_2 = \frac{nf + u_{\alpha/2}^2 - u_{\alpha/2}.\sqrt{nf(1-f) + \frac{1}{4}u_{\alpha/2}^2}}{n + u_{\alpha/2}^2}$$

Nhận xét: Khi $n \geq 100$ thì $G = U = \frac{f - p}{\sqrt{f(1-f)}}.\sqrt{n} \sim N(0,1)$ và các khoảng tin cậy của *p* là:

- Khoảng tin cậy bên phải:

$$\left(\overline{x} - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_\alpha, +\infty\right)$$

- Khoảng tin cậy bên trái:

$$\left(-\infty, \overline{x} + \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_\alpha\right)$$

- Khoảng tin cậy đối xứng:

$$\left(\overline{x} - \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{\alpha/2}, \overline{x} + \frac{\sqrt{f(1-f)}}{\sqrt{n}}u_{\alpha/2}\right)$$

Kích thước mẫu tối thiểu để $\varepsilon \leq \varepsilon_0$ cho trước là $n \geq n_0 = \left[\frac{f(1-f)}{\varepsilon_0^2}.u_{\alpha/2}^2\right] + 1$.

Ví dụ 4.2.3 Trọng lượng X một loại chi tiết là một biến ngẫu nhiên tuân theo quy luật chuẩn với độ lệch chuẩn 1,2 kg. Phải chọn ít nhất bao nhiêu chi tiết để điều tra, nếu muốn độ chính xác của ước lượng là 0,3 và độ tin cậy của ước lượng là 0,95.

ĐS: 62

Ví dụ 4.2.4 Hãy ước lượng tỷ lệ chính phẩm của một nhà máy bằng khoảng tin cậy đối xứng với độ tin cậy 0,95 biết rằng kiểm tra 100 sản phẩm của nhà máy thì thấy có 10 phế phẩm.

ĐS: (0,8412;0,9588)

Ví dụ 4.2.5 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 có 60 người nói rằng họ sẽ bỏ phiếu cho ông A. tìm khoảng tin cậy cho tỷ lệ cử tri bỏ phiếu cho ông A với độ tin cậy 90%.

ĐS: (0,52;0,68)

Ví dụ 4.2.6 Một công ty lớn muốn ước lượng trung bình một ngày thư ký phải đánh máy bao nhiêu trang giấy. Một mẫu gồm 50 thư ký được chọn ngẫu nhiên cho thấy số trang trung bình mà họ đánh máy là 32 với độ lệch tiêu chuẩn là 6. Tìm khoảng tin cậy 99% cho số trang trung bình mà một thư ký của công ty đánh máy trong một ngày.

ĐS: (30,61;33,39)

Ví dụ 4.2.7 Cơ quan cảnh sát giao thông kiểm tra hệ thống phanh của 40 chiếc xe tải trên quốc lộ. Họ phát hiện 14 chiếc có phanh chưa đảm bảo an toàn.

a. Tìm khoảng tin cậy 95% cho tỷ lệ xe tải có phanh chưa an toàn.

b. Tìm khoảng tin cậy 98% cho tỷ lệ xe tải có phanh tốt.

ĐS: a. (17,5%;52,5%)   b. (47,5%;82,5%)

Ước lượng tham số σ^2 của $X \sim N\left(\mu,\sigma^2\right)$

Trường hợp 1: Đã biết μ

Xét thống kê $G=\dfrac{nS^{*2}}{\sigma^2}\sim\chi^2\left(n\right)$, từ điều kiện $P\left(\chi^2_{1-\alpha_1}(n)<G<\chi^2_{\alpha_2}(n)\right)=1-\alpha$ ta được

$$P\left(\frac{nS^{*2}}{\chi^2_{\alpha_2}(n)}<\sigma^2<\frac{nS^{*2}}{\chi^2_{1-\alpha_1}(n)}\right)=1-\alpha$$

Khi đó ta có các khoảng ước lượng là:

- $\left(\dfrac{nS^{*2}}{\chi^2_{\alpha/2}(n)}, \dfrac{nS^{*2}}{\chi^2_{1-\alpha/2}(n)}\right)$
- $\left(\dfrac{nS^{*2}}{\chi^2_{\alpha}(n)}, +\infty\right)$
- $\left(0, \dfrac{nS^{*2}}{\chi^2_{1-\alpha}(n)}\right)$

Trường hợp 1: Chưa biết μ

Xét thống kê $G = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, Khi đó ta có các khoảng ước lượng là:

- $\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$
- $\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}, +\infty\right)$
- $\left(0, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}\right)$

Ví dụ 4.2.7 Để nghiên cứu độ ổn định của một máy gia công, người ta lấy ngẫu nhiên 25 chi tiết do máy đó gia công, đem đo và thu được các kích thước sau:

24,1	27,2	26,7	23,6	26,4
25,8	27,3	23,2	26,9	27,1
22,7	26,9	24,8	24,0	23,4
24,5	26,1	25,9	25,4	22,9
26,4	25,4	23,3	23,0	24,3

Với độ tin cậy 95% hãy ước lượng độ phân tán của kích thước các chi tiết do máy đó gia công. Biết kích thước chi tiết được gia công là biến ngẫu nhiên phân phối chuẩn.

CHƯƠNG 5

KIỂM ĐỊNH GIẢ THUYẾT THỐNG KÊ

5.1 KHÁI NIỆM CƠ BẢN

Giả thuyết thống kê?

Giả thuyết thống kê là giả thuyết về dạng phân phối xác suất của biến ngẫu nhiên, về các tham số đặc trưng của biến ngẫu nhiên, hoặc về tính độc lập của các biến ngẫu nhiên.

Ký hiệu H_0: giả thuyết gốc, H_1: giả thuyết đối của H_0. Vì các giả thuyết thống kê có thể đúng hoặc có thể sai nên cần kiểm định nó. Việc kiểm định này được gọi là kiểm định giả thuyết thống kê.

Tiêu chuẩn kiểm định giả thuyết thống kê

Giả sử $W = (X_1, ..., X_n)$ là mẫu ngẫu nhiên của X và thống kê $G = f(X_1, ..., X_n, \theta_0)$, với θ_0 là tham số liên quan đến giả thuyết cần kiểm định.

Nếu H_0 đúng thì quy luật phân phối xác suất của G được xác định, khi đó G được gọi là tiêu chuẩn kiểm định.

Miền bác bỏ giả thuyết thống kê

Giả sử G là tiêu chuẩn kiểm định. Với mức ý nghĩa $\alpha > 0$ khá bé cho trước, luôn tồn tại W_α sao cho

$$P\left(G \in W_\alpha \middle| H_0\right) = \alpha$$

Khi đó W_α được gọi là miền bác bỏ giả thuyết H_0 với mức ý nghĩa α.

Giá trị quan sát của tiêu chuẩn kiểm định

Giả sử $w = (x_1, ..., x_n)$ là mẫu cụ thể của mẫu ngẫu nhiên $W = (X_1, ..., X_n)$ của biến ngẫu nhiên X. Khi đó $G_{qs} = f(x_1, ..., x_n, \theta_\alpha)$ được gọi là giá trị quan sát của G.

- Nếu $G_{qs} \in W_\alpha$ thì bác bỏ giả thuyết H_0, thừa nhận giả thuyết đối H_1
- Nếu $G_{qs} \notin W_\alpha$ thì chưa có cơ sở để bác bỏ H_0, trến thực tế vẫn thừa nhận H_0.

Sai lầm loại 1, sai lầm loại 2

Sai lầm loại 1: Bác bỏ H_0 trong khi nó đúng, xác suất mắc sai lầm loại 1 đúng bằng α .

Sai lầm loại 2: Thừa nhận H_0 trong khi nó sai, xác suất mắc sai lầm loại 2 là

$$\beta = P\left(G \notin W_\alpha \middle| H_1\right)$$

và $1-\beta = P\left(G \in W_\alpha \middle| H_1\right)$ được gọi là lực kiểm định.

5.2 KIỂM ĐỊNH THAM SỐ

Kiểm định giả thuyết về giá trị trung bình của tổng thể chung

Trường hợp một tham số

Giả sử lượng biến của tiêu thức X trong tổng thể chung tuân theo quy luật phân phối chuẩn $N(\mu, \sigma^2)$. Ta chưa biết μ nhưng nếu có cơ sở để cho rằng nó bằng μ_0, ta đưa ra giả thuyết $H_0 : \mu = \mu_0$.

Trường hợp đã biết σ^2*:*Khi H_0 đúng thì thống kê $G = \dfrac{\overline{X} - \mu_0}{\sigma}.\sqrt{n}$ tuân theo quy luật $N(0,1)$

Nếu giả thuyết đối $H_1 : \mu > \mu_0$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} > u_\alpha$$

thì ta bác bỏ H_0, trong đó u_α là giá trị tới hạn của phân phối chuẩn $N(0,1)$ và α mức ý nghĩa cho trước. Ngược lại, nếu

$$G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} \le u_\alpha$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Nếu giả thuyết đối $H_1: \mu < \mu_0$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} < -u_\alpha$$

thì ta bác bỏ H_0. Ngược lại, nếu

$$G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} \ge u_\alpha$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Nếu giả thuyết đối $H_1: \mu \neq \mu_0$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} > u_{\alpha/2} \text{ hoặc } G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} < -u_{\alpha/2}$$

thì ta bác bỏ. Ngược lại, nếu

$$-u_{\alpha/2} \le G_{qs} = \frac{\overline{x} - \mu_0}{\sigma}.\sqrt{n} \le u_{\alpha/2}$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ngoài ra nếu chưa cho trước mức ý nghĩa α thì ta có thể sử dụng giá trị xác suất $P-value$ để kiểm định giả thuyết H_0, trong đó $P-value$ được xác định phụ thuộc vào H_1. Cụ thể, $H_1: \mu > \mu_0$, $P-value = P(G > G_{qs})$; $H_1: \mu < \mu_0$, $P-value = P(G < G_{qs})$;

và $H_1: \mu \neq \mu_0$, $P-value = 2P(G > |G_{qs}|)$. Sau đó ta so sánh $P-value$ với các giá trị sau đây: $P-value < 0,001$: bác bỏ H_0 và $P-value > 0,1$: chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận nó. Nếu cho trước mức ý nghĩa α thì chúng ta sử dụng $P-value$

để so sánh với α : nếu $P - value < \alpha$ thì bác bỏ H_0; ngược lại ta nói chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận nó.

Trường hợp chưa biết σ^2: Khi H_0 đúng thì thống kê $G = \dfrac{\overline{X} - \mu_0}{S}.\sqrt{n}$ tuân theo quy luật Student với $n-1$ bậc tự do $T(n-1)$, trong đó

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{k}\left(X_i - \overline{X}\right)^2.n_i} \;.$$

Tuy nhiên khi $n \geq 30$, $G = \dfrac{\overline{X} - \mu_0}{S}.\sqrt{n}$ tuân theo quy luật $N(0,1)$, do đó chúng ta làm tương tự như trường hợp đã biết σ^2 bằng cách thay σ bởi S (S trở thành s đối với mẫu cụ thể). Khi $n < 30$, $G = \dfrac{\overline{X} - \mu_0}{S}.\sqrt{n}$ vẫn tuân theo quy luật $T(n-1)$. Khi đó chúng ta kiểm định H_0 bằng cách so sánh giá trị quan sát $G_{qs} = \dfrac{\overline{x} - \mu_0}{s}.\sqrt{n}$ với giá trị tới hạn của $T(n-1)$, do đó chúng ta làm tương tự như trường hợp đã biết σ^2 bằng cách thay u_α bởi $t_\alpha^{(n-1)}$, $u_{\alpha/2}$ bởi $t_{\alpha/2}^{(n-1)}$.

Ví dụ 5.2.1 Một công ty dự định mở một của hàng siêu thị tại một khu dân cư A. Để đánh giá khả năng mua hàng của nhân dân trong khu dân cư này, giám đốc công ty đã cho điều tra thu nhập bình quân hàng tháng của 100 hộ được chọn một cách ngẫu nhiên trong khu và thu được bảng số liệu sau:

Thu nhập bình quân (ngàn đồng/người/tháng)	150	200	250	300	350
Số hộ	10	15	20	30	10

Theo bộ phận tiếp thị thì siêu thị sẽ hoạt động không có hiệu quả tại khu vực này nếu thu nhập bình quân hàng tháng của các hộ chỉ đạt mức 250 ngàn đồng. Vậy qua kết quả điều tra trên, công ty có nên quyết định mở siêu thị tại khu dân cư A này hay không? (Yêu cầu kết luận với xác suất tin cậy 95% và biết rằng thu nhập bình quân hàng tháng của các hộ trong khu vực này tuân theo quy luật chuẩn).

Ví dụ 5.2.2 Một mẫu kích thước n=25 được rút ra từ tổng thể phân phối chuẩn với phương sai là 64. Nếu trung bình mẫu tìm được là 55,4 thì với mức ý nghĩa 0,05 hãy tiến hành một thủ tục kiểm định giả thuyết là trung bình tổng thể bằng 52.

ĐS: U_{qs}=2,125

Ví dụ 5.2.3 Thông qua một mẫu gồm 100 gia đình, người ta thu được kết quả là chi tiêu trung bình hàng tháng của các gia đình đó là 2,455 triệu đồng với độ lệch chuẩn là 0,3 triệu. Với mức ý nghĩa 5% có thể cho rằng chi tiêu trung bình hàng tháng của các gia đình ít hơn 2,5 triệu hay không. Giả thiết mức chi tiêu phân phối chuẩn.

ĐS: Không bác bỏ giả thuyết

Ví dụ 5.2.4 Bột mì được đóng bao bằng máy tự động có trọng lượng đóng bao theo quy định là 16 kg và độ lệch chuẩn là 1,2 kg. Lấy ngẫu nhiên 25 bao bột để kiểm tra tìm được trọng lượng đóng gói trung bình là 16,5 kg.

a. Với mức ý nghĩa 5% có cần dừng máy để điều chỉnh hay không.
b. Tìm xác suất mắc sai lầm loại 2 nếu giá trị thực của trọng lượng đóng gói trung bình là 15,5 và 16,6.
c. Nếu muốn kiểm định vấn đề trên với xác suất mắc sai lầm loại 1 là 0,05 và xác suất mắc sai lầm loại 2 là 0,02 trong điều kiện trọng lượng đóng gói thực của bột mì sai lệch so với trọng lượng quy định không quá 1 kg thì phải lấy ra bao nhiêu bao bột mì để kiểm tra. Giả thiết trọng lượng đóng bao phân phối chuẩn.

ĐS: a. Cần dừng máy   b. 0,4522 và 0,2946   c. 24

Ví dụ 5.2.5 Thời gian trước số tiền gửi tiết kiệm bằng ngoại tệ trung bình của mỗi khách hàng là 1000 USD. Để đánh giá xem hiện nay xu hướng này còn giữa nguyên hay không người ta kiểm tra ngẫu nhiên 64 sổ tiết kiệm và tìm được số tiền gửi trung bình là 990 USD, độ lệch chuẩn là 100 USD.

a. Với mức ý nghĩa 0,05 hãy kiểm định xem số tiền gửi trung bình có thay đổi hay không.
b. Hãy kiểm định vấn đề trên bằng cách sử dụng P-value.

c. Tìm xác suất mắc sai lầm loại 2 nếu số tiền tiết kiệm trung bình của mỗi khách hàng thực sự bằng 1050 USD.

d. Tìm lực kiểm định nếu $\mu = 980$ USD.

e. Nếu đòi hỏi xác suất mắc sai lầm loại 1 và loại 2 đều bằng 0,05 và sai lệch giữa giá trị trung bình thực sự với giá trị giả thuyết không quá 30 USD thì phải điều tra một mẫu có kích thước tối thiểu là bao nhiêu?

ĐS: a. T_{qs}=-0,8 b. P-value=0,21 c. $\beta = 0,0207$ d. 0,3594 e. 144

Ví dụ 5.2.6 Một loại cây nào đó trong điều kiện bình thường có chiều cao trung bình là 11 inches. Người ta muốn thử xem một nguyên tố vi lượng A có ảnh hưởng tới chiều cao của cây không. Trong một vườn thí nghiệm trồng 48 cây này có bón thêm nguyên tố vi lượng A, ta tính được chiều cao trung bình là 10,3 với độ lệch tiêu chuẩn là 2,3. Sử dụng phương pháp P-value, hãy kết luận xem nguyên tố vi lượng A có ảnh hưởng đến chiều cao của cây hay không với mức ý nghĩa 5%?

ĐS: Phân vi lượng A có ảnh hưởng tới chiều cao của cây

Trường hợp hai tham số

Trường hợp hai mẫu độc lập: Giả sử có hai tổng thể chung, tổng thể thứ nhất có lượng biến của tiêu thức X_1 tuân theo quy luật phân phối chuẩn $N\left(\mu_1, \sigma_1^2\right)$ và tổng thể thứ hai có lượng biến của tiêu thức X_2 tuân theo quy luật phân phối chuẩn $N\left(\mu_2, \sigma_2^2\right)$. Nếu μ_1 và μ_2 chưa biết nhưng có cơ sở để giả định rằng giá trị của chúng bằng nhau ta có giả thuyết gốc là $H_0 : \mu_1 = \mu_2$.

Trường hợp đã biết phương sai của hai tổng thể chung σ_1^2 và σ_2^2: Từ hai tổng thể trên có thể rút ra hai mẫu độc lập $W_1 = \left(X_{11}, X_{12}, ..., X_{1n_1}\right)$ và $W_2 = \left(X_{21}, X_{22}, ..., X_{2n_2}\right)$.

Khi H_0 đúng thì thống kê

$$G = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \text{ trở thành } G = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

và tuân theo quy luật $N(0,1)$.

Nếu giả thuyết đối $H_1 : \mu_1 > \mu_2$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_\alpha$$

thì ta bác bỏ H_0, trong đó u_α là giá trị tới hạn của phân phối chuẩn $N\left(0,1\right)$ và α mức ý nghĩa cho trước. Ngược lại, nếu

$$G_{qs} = \frac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \le u_\alpha$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Nếu giả thuyết đối $H_1 : \mu_1 < \mu_2$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -u_\alpha$$

thì ta bác bỏ H_0. Ngược lại, nếu $G_{qs} = \frac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \ge -u_\alpha$ thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Nếu giả thuyết đối $H_1 : \mu_1 \ne \mu_2$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_{\alpha/2} \text{ hoặc } G_{qs} = \frac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -u_{\alpha/2}$$

thì ta bác bỏ H_0. Ngược lại, nếu $-u_{\alpha/2} \le G_{qs} = \dfrac{\left(\overline{x}_1 - \overline{x}_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \le u_{\alpha/2}$ thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.2.7 Tại một xí nghiệp người ta xây dựng hai phương án gia công cùng một loại chi tiết. Để đánh giá xem chi phí trung bình về nguyên liệu theo hai phương án đó có khác nhau hay không người ta tiến hành sản xuất thử và thu được kết quả sau:

Phương án1	2.5	3.2	3.5	3.8	3.5	
Phương án 2	2.0	2.7	2.5	2.9	2.3	2.6

Với mức ý nghĩa $\alpha = 0.05$, hãy kết luận về vấn đề trên biết rằng chi phí nguyên vật liệu theo cả hai phương án gia công đều là các biến ngẫu nhiên tuân theo quy luật chuẩn với $\sigma_1^2 = \sigma_2^2 = 0,16$.

Trường hợp chưa biết phương sai của hai tổng thể chung σ_1^2 và σ_2^2 và không thể cho rằng chúng bằng nhau:

Nếu có thể điều tra được từ tổng thể hai mẫu độc lập kích thước n_1 và n_2 thì khi H_0 đúng, thống kê

$$G = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \text{ trở thành } G = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$$

và tuân theo quy tuân theo quy luật Student với k bậc tự do $T(k)$, trong đó S_1^2 và S_2^2 lần lượt là phương sai của mẫu thứ nhất và thứ hai, còn

$$k = \frac{(n_1 - 1)(n_2 - 1)}{(n_1 - 1)(1 - C)^2 + (n_2 - 1)C^2} \text{ với } C = \frac{\dfrac{S_1^2}{n_1}}{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}.$$

Khi đó chúng ta làm tương tự như trường hợp đã biết phương sai của hai tổng thể chung σ_1^2 và σ_2^2 bằng cách thay σ_1^2 bởi S_1^2, σ_2^2 bởi S_2^2, u_α bởi $t_\alpha^{(k)}$, $u_{\alpha/2}$ bởi $t_{\alpha/2}^{(k)}$.

Ví dụ 5.2.8 Có ý kiến cho rằng chiều cao của nam thanh niên sống ở khu vực thành thị cao hơn chiều cao của nam thanh niên sống ở khu vực nông thôn, người ta tiến hành chọn ra 10 nam thanh niên sống ở khu vực thành thị và 12 nam thanh niên sống ở khu vực nông thôn để đo chiều cao và thu được kết quả như sau (cm):

Thành thị	168	171	165	169	168	173	165	162	167	169		
Nông thôn	162	168	174	164	165	166	160	163	165	167	167	163

Với mức ý nghĩa $\alpha = 0.05$, hãy đưa ra kết luận về nhận định trên, biết rằng chiều cao của nam thanh niên sống ở khu vực thành thị và sống ở khu vực nông thôn là biến ngẫu nhiên tuân theo quy luật chuẩn.

Trường hợp hai mẫu điều tra phụ thuộc theo từng cặp: Trong phần đầu, khi kiểm định về hai kỳ vọng toán ta luôn giả định rằng hai mẫu ngẫu nhiên được chọn ra từ các tổng thể nghiên cứu là độc lập nhau. Trong thực tế có nhiều trường hợp hai mẫu ngẫu nhiên được lấy ra từ cùng một tổng thể thì khả năng chúng phụ thuộc nhau rất cao. Khi đó chúng ta sử dụng phương pháp kiểm định như sau: Từ hai tổng thể, lấy ra hai mẫu ngẫu nhiên kích thước n, ký hiệu $W_1 = \left(X_{11}, X_{12}, ..., X_{1n}\right)$ và $W_2 = \left(X_{21}, X_{22}, ..., X_{2n}\right)$. Lập biến ngẫu nhiên D với các giá trị có thể có của nó là $D_1, ..., D_n$ với $D_i = X_{1i} - X_{2i}, \left(i = \overline{1, n}\right)$. Khi đó kỳ vọng toán của D là $\mu_D = E(X_1 - X_2) = \mu_1 - \mu_2$ và việc kiểm định giả thuyết $H_0 : \mu_1 = \mu_2$ trở thành kiểm định

$H_0 : \mu_D = 0$. Xét tiêu chuẩn kiểm định

$$G = \frac{\overline{D}}{S_D}.\sqrt{n} \text{ với } \overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \text{ và } S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(D_i - \overline{D}\right)^2.$$

Ta có G tuân theo quy luật phân phối Student với $n-1$ bậc tự do. Khi đó ta thực hiện các thủ tục kiểm định tiếp theo như trong phần kiểm định một kỳ vọng của tổng thể đã biết.

Ví dụ 5.2.9 Theo dõi doanh số bán hàng của một công ty (triệu đồng) trong 15 ngày đầu tháng ba và 15 ngày đầu tháng năm và thu được kết quả sau:

Ngày	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tháng 3	7.6	10.2	9.5	1.3	3.0	6.3	5.3	6.2	2.2	4.8	11.3	12.1	6.9	7.6	8.4
Tháng 5	7.3	9.1	8.4	1.5	2.7	5.0	4.9	5.3	2.0	4.2	11.0	11.0	6.1	6.7	7.5

Nếu giả thiết rằng doanh số hàng ngày phân phối chuẩn thì với mức ý nghĩa $\alpha = 0.05$ có thể cho rằng doanh số bán trung bình hàng ngày trong tháng 5 giảm sút so với tháng 3 hay không?

Trường hợp kiểm định trung bình của nhiều tổng thể chung (phân tích phương sai)

Phân tích phương sai một nhân tố: dùng để phân tích ảnh hưởng của một yếu tố nguyên nhân đến một yếu tố đang nghiên cứu. Giả sử có k tổng thể nghiên cứu trong đó các biến ngẫu nhiên $X_i, (i = \overline{1,k})$ tuân theo quy luật phân phối chuẩn $N\left(\mu_i, \sigma_i^2\right)$. Các phương sai μ_i chưa biết nhưng nếu có cơ sở để giả định rằng tất cả chúng bằng nhau, khi đó ta có thể đưa ra giả thuyết gốc $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$. Để kiểm định giả thuyêt này, từ các tổng thể chung lấy ra k mẫu ngẫu nhiên, độc lập nhau, với kích thước tương ứng là $n_1, \ldots, n_k$ và các số liệu được trình bày dưới dạng bảng sau:

Mức độ					
1	2	…	I	…	K
x_{11}	x_{21}	…	x_{i1}	…	x_{k1}
x_{12}	x_{22}	…	x_{i2}	…	x_{k2}
…	…	…	…	…	…
x_{1j}	x_{2j}	…	x_{ij}	…	x_{kj}
…	…	…	…	…	…
x_{1n_1}	x_{2n_2}	…	x_{in_i}	…	x_{kn_k}

Tính các trung bình:

Trung bình của các mẫu $\overline{x}_i = \dfrac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$

Trung bình chung $\overline{x} = \dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}}{n_i} = \dfrac{\sum_{i=1}^{k} \overline{x}_i . n_i}{\sum_{i=1}^{k} n_i}$

Tính các tổng bình phương độ lệch:

Tổng bình phương chung: $SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij} - \overline{x}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^2 - n\left(\overline{x}\right)^2$

Tổng bình phương do ảnh hưởng của nhân tố: $SSF = \sum_{i=1}^{k}\left(\overline{x}_i - \overline{x}\right)^2 .n_i$

Tổng bình phương do sai số: $SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij} - \overline{x}_i\right)^2$

Nhận thấy $SST = SSF + SSE$

Tính các phương sai tương ứng:

Phương sai do ảnh hưởng nhân tố: $MSF = \dfrac{SSF}{k-1}$

Phương sai do sai số: $MSE = \dfrac{SSE}{n-k}$

Khi giả thuyết H_0 đúng thì thống kê $F = \dfrac{MSF}{MSE}$ tuân theo quy luật phân phối Fisher-Snedecor với bậc tự do là $(k-1, n-k)$. Nếu $F > f_{\alpha}^{(k-1,n-k)}$ thì bác bỏ giả thuyết H_0.

Ví dụ 5.2.10 Năng suất của một giống ngô đêm trồng ở ba vùng A, B, C có thực sự khác nhau không? Hãy cho kết luận với mức ý nghĩa 1% trên cơ sở bảng số liệu sau:

Số thứ tự quan sát		1	2	3	4	5	6	7
Vùng	A	30.56	32.66	34.78	35.5	36.63	40.2	42.28
	B	43.44	47.51	53.8	42.23			
	C	31.36	36.2	36.38				

Kiểm định giả thuyết về phương sai của tổng thể chung

Trường hợp một tham số

Giả sử trong tổng thể biến ngẫu nhiên gốc X tuân theo quy luật phân phối chuẩn $N(\mu, \sigma^2)$. Ta chưa biết σ^2 nhưng nếu có cơ sở để cho rằng nó bằng σ_0^2 ta đưa ra giả thuyết $H_0 : \sigma^2 = \sigma_0^2$. Khi H_0 đúng thì thống kê $G = \frac{(n-1)S^2}{\sigma_0^2}$ tuân theo quy luật $\chi^2(n-1)$. Nếu giả thuyết đối $H_1 : \mu > \mu_0$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} > \chi_\alpha^2(n-1)$$

thì ta bác bỏ H_0, trong đó $\chi_\alpha^2(n-1)$ là giá trị tới hạn của phân phối $\chi^2(n-1)$ và α mức ý nghĩa cho trước. Ngược lại, nếu

$$G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} \le \chi_\alpha^2(n-1)$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét. Nếu giả thuyết đối $H_1 : \mu < \mu_0$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} < \chi_{1-\alpha}^2(n-1)$$

thì ta bác bỏ H_0. Ngược lại, nếu

$$G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} \ge \chi_{1-\alpha}^2(n-1)$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét. Nếu giả thuyết đối $H_1 : \mu \neq \mu_0$ và giá trị quan sát đối với mẫu cụ thể

$$G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{\alpha/2}(n-1) \text{ hoặc } G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{1-\alpha/2}(n-1)$$

thì ta bác bỏ H_0. Ngược lại, nếu

$$\chi^2_{1-\alpha/2}(n-1) \leq G_{qs} = \frac{(n-1)s^2}{\sigma_0^2} \leq \chi^2_{\alpha/2}(n-1)$$

thì ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.2.11 Để kiểm tra độ chính xác của một máy, người ta đo ngẫu nhiên kích thước của 15 chi tiết do máy đó sản xuất và tính được $s^2 = 14,6$. Với mức ý nghĩa 1%, hãy kết luận máy có hoạt động bình thường không? Biết rằng kích thước chi tiết là biến ngẫu nhiên phân phối chuẩn có dung sai theo thiết kế là $\sigma^2 = 12$.

Trường hợp hai tham số

Giả sử có hai tổng thể chung, tổng thể thứ nhất có lượng biến của tiêu thức X_1 tuân theo quy luật phân phối chuẩn $N\left(\mu_1, \sigma_1^2\right)$ và tổng thể thứ hai có lượng biến của tiêu thức X_2 tuân theo quy luật phân phối chuẩn $N\left(\mu_2, \sigma_2^2\right)$. Nếu σ_1^2 và σ_2^2 chưa biết nhưng có cơ sở để giả định rằng giá trị của chúng bằng nhau ta có giả thuyết gốc là $H_0 : \sigma_1^2 = \sigma_2^2$. Từ hai tổng thể trên có thể rút ra hai mẫu độc lập $W_1 = \left(X_{11}, X_{12}, ..., X_{1n_1}\right)$ và $W_2 = \left(X_{21}, X_{22}, ..., X_{2n_2}\right)$.

Khi H_0 đúng thì thống kê $G = \dfrac{S_1^2.\sigma_2^2}{S_2^2.\sigma_1^2}$ trở thành $G = \dfrac{S_1^2}{S_2^2}$ và tuân theo quy luật Fisher-Snedecor $F\left(n_1 - 1, n_2 - 1\right)$. Với mức ý nghĩa α cho trước, nếu giả thuyết đối $H_1 : \sigma_1^2 > \sigma_2^2$ và giá trị quan sát đối với mẫu cụ thể $G_{qs} = \dfrac{s_1^2}{s_2^2} > f_\alpha^{(n_1-1, n_2-1)}$ thì ta bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫuđang xét. Nếu giả thuyết đối $H_1:\sigma_1^2<\sigma_2^2$ và giá trị quan sát đối với mẫu cụ thể $G_{qs}=\frac{s_1^2}{s_2^2}<f_{1-\alpha}^{(n_1-1,n_2-1)}$ thì ta bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét. Nếu giả thuyết đối $H_1:\sigma_1^2\neq\sigma_2^2$ và giá trị quan sát đối với mẫu cụ thể $G_{qs}=\frac{s_1^2}{s_2^2}>f_{\alpha/2}^{(n_1-1,n_2-1)}$ hoặc $G_{qs}=\frac{s_1^2}{s_2^2}<f_{1-\alpha/2}^{(n_1-1,n_2-1)}$ thì ta bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.2.12 Một số nhà kinh tế cho rằng độ phân tán của thị phần trong các công ty hoạt động có cạnh tranh về giá cả cao hơn trong các công ty độc quyền. Để kết luận về điều đó người ta tiến hành điều tra thị phần của một công ty có cạnh tranh về giá cả trong 4 năm và tìm được phương sai mẫu là 85.576, đồng thời điều tra thị phần của một công ty độc quyền trong 7 năm tìm được phương sai mẫu là 13.78. Với mức ý nghĩa 5%, hãy kết luận về ý kiến trên. Cho biết rằng thị phần của các công ty là biến ngẫu nhiên phân phối chuẩn.

Trường hợp nhiều hơn hai tham số

Giả sử có k tổng thể nghiên cứu trong đó các biến ngẫu nhiên $X_i,(i=\overline{1,k})$ tuân theo quy luật phân phối chuẩn $N\left(\mu_i,\sigma_i^2\right)$. Các phương sai σ_i^2 chưa biết nhưng nếu có cơ sở để giả định rằng tất cả chúng bằng nhau, khi đó ta có thể đưa ra giả thuyết gốc $H_0:\sigma_1^2=\sigma_2^2=\cdots=\sigma_k^2$. Để kiểm định giả thuyêt này, từ các tổng thể chung lấy ra k mẫu ngẫu nhiên, độc lập nhau, với kích thước tương ứng là $n_1,\ldots,n_k$.

Trường hợp 1: Nếu các kích thước mẫu $n_1,\ldots,n_k$ khác nhau, ta tiến hành các thủ tục sau đây để kiểm định H_0. Tính

$$\overline{S^2}=\frac{\sum_{i=1}^{k}\left(n_i-1\right)S_i^2}{\sum_{i=1}^{k}\left(n_i-1\right)}.$$

Khi H_0 đúng thì thống kê Bartlett

$$G = \frac{2.303\left[h.\lg \overline{S^2} - \sum_{i=1}^{n} h_i.\lg S_i^2\right]}{1 + \frac{1}{3(k-1)}\left[\sum_{i=1}^{k}\frac{1}{h_i} - \frac{1}{h}\right]} \text{ với } h = \sum_{i=1}^{k} h_i \text{ và } h_i = n_i - 1$$

tuân theo quy luật phân phối khi bình phương $\chi^2\left(k-1\right)$, $\forall n_i > 3$. Với mức ý nghĩa α cho trước, nếu $G_{qs} > \chi_\alpha^2(k-1)$ thì ta bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.2.13 Từ bốn mẫu độc lập với kích thước 10, 13, 15 và 16 được lấy ra từ các tổng thể phân phối chuẩn và tìm được các phương sai mẫu tương ứng là 0.25, 0.4, 0.36, 0.46. Với mức ý nghĩa 5%, hãy kiểm định xem phương sai của các tổng thể có bằng nhau hay không?

Trường hợp 2: Nếu các kích thước mẫu $n_1, \ldots, n_k$ bằng nhau, việc kiểm định thông qua tiêu chuẩn kiểm định Bartlett không còn đáng tin cậy nữa. Khi đó người ta sử dụng tiêu chuẩn kiểm định Cochran $G = \frac{s_{\max}^2}{s_1^2 + s_2^2 \cdots + s_k^2}$ để kiểm định H_0. Với mức ý nghĩa α cho trước, nếu $G = \frac{s_{\max}^2}{s_1^2 + s_2^2 \cdots + s_k^2} > g_\alpha^{(n-1,k)}$ với $n = n_1 = \cdots = n_k$ thì ta bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấ p nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.2.14 Từ bốn mẫu độc lập kích thước như nhau và bằng 17 được lấy ra từ các tổng thể phân phối chuẩn và tìm được các phương sai mẫu tương ứng là 0.26, 0.36, 0.40 và 0.42. Với mức ý nghĩa 5%, hãy kiểm định sự bằng nhau của các phương sai của tổng thể.

Kiểm định giả thuyết về tham số p

Trường hợp một tham số

Giả sử lượng biến của tiêu thức X trong tổng thể chung tuân theoquy luật phân phối không-một. Ta chưa biết p nhưng nếu có cơ sở để cho rằng nó bằng p_0, ta đưa ra giả thuyết $H_0 : p = p_0$.

Nếu n và p thỏa mãn điều kiện

$$n > 5, \left|\sqrt{\frac{p}{1-p}} - \sqrt{\frac{1-p}{p}}\right| < 0.3\sqrt{n}$$

thì khi H_0 đúng, thống kê $G = \frac{(F - p_0)}{\sqrt{p_0(1-p_0)}}.\sqrt{n}$ tuân theo quy luật chuẩn hóa $N(0,1)$. Khi đó ta quyết định bác bỏ hay thừa nhận H_0 phụ thuộc vào việc so sánh giá trị quan sát $G_{qs} = \frac{(f - p_0)}{\sqrt{p_0(1-p_0)}}.\sqrt{n}$ với giá trị tới hạn của $N(0,1)$ dựa vào giả thuyết đối H_1 tương tự như đã làm cho việc kiểm định giả thuyết về một kỳ vọng toán của tổng thể chung.

Ví dụ 5.2.15 Tỷ lệ khách hàng tiêu dùng một loại sản phẩm tại một địa phương là 60%. Sau một chiến dịch quảng cáo người ta muốn đánh giá xem chiến dịch quảng cáo này có thật sự mang lại hiệu quả hay không. Để làm điều đó người ta phỏng vấn ngẫu nhiên 400 khách hàng thì thấy có 250 người tiêu dùng loại sản phẩm nói trên. Với mức ý nghĩa $\alpha = 0.05$, hãy kết luận về hiệu quả của chiến dịch quảng cáo đó.

Ví dụ 5.2.16 Lô hàng đủ tiêu chuẩn xuất khẩu nếu tỷ lệ phế phẩm không vượt quá 3%. Kiểm tra ngẫu nhiên 400 sản phẩm cua lô hàng này thấy có 14 phế phẩm. Với mức độ tin cậy 95% có cho phép lô hàng xuất khẩu được không.

ĐS: U_{qs}=0,586

Ví dụ 5.2.17 Một báo cáo nói rằng 18% gia đình ở thành phố A có máy tính cá nhân ở nhà. Để kiểm tra, người ta chọn ngẫu nhiên 80 gia đình trong thành phố có trẻ em đang đi học và thấy rằng có 22 gia đình có máy tính. Với mức ý nghĩa 0,02 hãy kiểm định xem liệu trong các gia đình có trẻ em đang đi học, tỷ lệ gia đình có máy tính có cao hơn tỷ lệ chung hay không?

ĐS: Tỷ lệ gia đình có máy tính cao hơn tỷ lệ chung

Ví dụ 5.2.18 Cơ quan cảnh sát giao thông cho rằng 62% số người lái xe trên đường là có bằng lái. Kiểm tra ngẫu nhiên 130 người lái xe, cảnh sát giao thông thấy chỉ có 68 người có bằng lái xe. Số liệu này có chứng tỏ tỷ lệ người có bằng lái xe thấp hơn 62% hay không (sử dụng phương pháp P-value với mức ý nghĩa 2%)?

ĐS: Tỷ lệ người có bằng lái xe thực tế thấp hơn 62%

Ví dụ 5.2.19 Một tỉnh báo cáo rằng tỷ lệ học sinh đỗ tốt nghiệp của họ là 88%. Một mẫu ngẫu nhiên gồm 100 em được chọn cho thấy trong đó chỉ có 82 em đỗ. Với mức ý nghĩa 5%, kiểm định xem phải chăng báo cáo của tỉnh về tỷ lệ đỗ 88% là cao hơn sự thật?

Trường hợp hai tham số

Giả sử có hai tổng thể chung, tổng thể thứ nhất có lượng biến của tiêu thức X_1 tuân theo quy luật phân phối không-một $A(p_1)$ và tổng thể thứ hai có lượng biến của tiêu thức X_2 tuân theo quy luật phân phối không-một $A(p_2)$. Nếu p_1 và p_2 chưa biết nhưng có cơ sở để giả định rằng giá trị của chúng bằng nhau ta có giả thuyết gốc là $H_0 : p_1 = p_2$. Từ hai tổng thể trên có thể rút ra hai mẫu độc lập $W_1 = \left(X_{11}, X_{12}, ..., X_{1n_1}\right)$ và $W_2 = \left(X_{21}, X_{22}, ..., X_{2n_2}\right)$. Khi H_0 đúng thì thống kê

$$G = \frac{(F_1 - F_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \text{ trở thành } G = \frac{F_1 - F_2}{\sqrt{p(1-p)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

và tuân theo quy luật $N(0,1)$ khi $n_1 \geq 30, n_2 \geq 30$. Thông thường p chưa biết nên nó được

thay bởi $\overline{f} = \dfrac{n_1 f_1 + n_2 f_2}{n_1 + n_2}$. Do đó việc quyết định bác bỏ hay thừa nhận H_0 phụ thuộc vào việc so sánh giá trị quan sát $G_{qs} = \dfrac{(f_1 - f_2)}{\sqrt{\overline{f}\left(1-\overline{f}\right)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ với giá trị tới hạn của $N(0,1)$ dựa vào giả thuyết đối H_1 tương tự như đã làm cho việc kiểm định giả thuyết về hai kỳ vọng toán của tổng thể chung.

Ví dụ 5.2.20 Hiện tượng học sinh bỏ học là vấn đề đang được đặc biệt quan tâm, nhất là ở các vùng nông thôn. Tại hai trường trung học ở hai vùng nông thôn M và N trong năm học 2011-2012 có các số liệu sau:

Trường	Số học sinh	Số học sinh bỏ học
M	1900	175
N	2600	325

Với mức ý nghĩa $\alpha = 0.05$, có thể cho rằng tình trạng bỏ học ở trường N là nghiêm trọng hơn ở trường M hay không?

Trường hợp nhiều hơn hai tham số

Giả sử có k tổng thể nghiên cứu trong đó các biến ngẫu nhiên $X_i, (i = \overline{1,k})$ tuân theo quy luật phân không-một $A(p_i)$. Các tỷ lệ p_i chưa biết nhưng nếu có cơ sở để giả định rằng tất cả chúng bằng nhau, khi đó ta có thể đưa ra giả thuyết gốc $H_0 : p_1 = p_2 = \cdots = p_k = p$. Để kiểm định giả thuyêt này, từ các tổng thể chung lấy ra k mẫu ngẫu nhiên, độc lập nhau, với kích thước tương ứng là $n_1, \ldots, n_k$ và các số liệu được trình bày dưới dạng bảng sau:

Mẫu	Số lần xuất hiện biến cố	Số lần không xuât hiện biến cố	Tổng số n_i
1	X_{11}	X_{12}	n_1
…	…	…	…
I	X_{i1}	X_{i2}	n_i
…	…	…	…
K	X_{k1}	X_{k2}	n_k
Tổng số m_j	m_1	m_2	$\sum n_i = n$

Nếu $n_i > 50, \forall i$ thì thống kê toán

$$U_i = \frac{X_{i1} - n_i p_i}{\sqrt{n_i p_i \left(1 - p_i\right)}} \sim N\left(0,1\right),\ i = \overline{1,k}$$

và thống kê toán $G = \sum_{i=1}^{k} U_i^2 \sim \chi^2\left(k\right)$. Khi H_0 đúng thì thống kê $G = \sum_{i=1}^{k} \frac{\left(X_{i1} - n_i p\right)^2}{n_i p\left(1 - p\right)}$.

Thông thường p chưa biết nên nó được thay bởi xấp xỉ của nó $\overline{f} = \frac{x_{11} + x_{21} + \ldots x_{k1}}{n_1 + n_2 + \cdots + n_k}$ và

$$G = \sum_{i=1}^{k} \frac{\left(X_{i1} - n_i p\right)^2}{n_i p\left(1 - p\right)} \text{ trở thành } G = \sum_{i=1}^{k} \frac{\left(X_{i1} - n_i \overline{f}\right)^2}{n_i \overline{p}\left(1 - \overline{f}\right)} = n\left[\sum_{i=1}^{k}\sum_{j=1}^{2} \frac{X_{ij}^2}{n_i m_j} - 1\right] \sim \chi^2\left(k-1\right).$$

Với mức ý nghĩa α cho trước, nếu

$$G_{qs} = n\left[\sum_{i=1}^{k}\sum_{j=1}^{2} \frac{x_{ij}^2}{n_i m_j} - 1\right] > \chi_\alpha^2\left(k-1\right)$$

thì bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Lưu ý: Để sử dụng tiêu chuẩn kiểm định trên thì số lần xuất hiện biến cố trong mỗi mẫu ít nhất là 5 lần.

Ví dụ 5.2.21 Trước khi đưa ra thị trường một loại sản phẩm với màu sơn mới, người ta muốn xem xét các lứa tuổi khác nhau phản ứng như thế nào với màu sắc sản phẩm đó. Một cuộc điều tra theo những nhóm lứa tuổi khác nhau về phản ứng của họ đối với màu sắc của sản phẩm mà doanh nghiệp sản xuất thu được kết quả như sau:

	Thích	Không thích	Tổng
Dưới 25 tuổi	104	21	125
Từ 25 đến 35	122	28	150
Từ 36 đến 45	68	32	100
Từ 46 trở lên	41	34	75
Tổng	335	115	450

Với mức ý nghĩa $\alpha = 0.05$, có thể cho rằng tỷ lệ những người thích màu sắc đó của sản phẩm là như nhau đối với mọi lứa tuổi hay không?

5.3 KIỂM ĐỊNH PHI THAM SỐ

Kiểm định giả thuyết về tính độc lập của hai dấu hiệu định tính

Giả sử cần nghiên cứu đồng thời hai dấu hiệu định tính A và B trên cùng một tổng thể. Dấu hiệu A có các phạm trù là $A_1, A_2, \ldots, A_k$ và dấu hiệu B có các phạm trù là $B_1, B_2, \ldots, B_l$. Nếu có cơ sở để giả thuyết rằng A và B độc lập nhau, ta đưa ra cặp giả thuyết thống kê như sau: H_0: A và B độc lập nhau, H_1: A và B không độc lập nhau. Để kiểm định giả thuyết trên ta lập mẫu kích thước n và trình bày các số liệu mẫu dưới dạng bảng như sau:

	B_1	B_2	...	B_j	...	B_l	Tổng số n_i
A_1	n_{11}	n_{12}	...	n_{1j}	...	n_{1l}	n_1
A_2	n_{21}	n_{22}	...	n_{2j}	...	n_{2l}	n_2
...	...	...	...	...	...	...	...
A_i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{il}	n_i
...	...	...	...	...	...	...	...
A_k	n_{k1}	n_{k2}	...	n_{kj}	...	n_{kl}	n_k
Tổng số m_j	m_1	n_{k2}	...	m_j	...	m_l	$\sum = n$

Trong đó n_{ij} là tần số ứng với các phần tử đồng thời mang dấu hiệu A_i và B_j. Với n khá lớn thì theo định nghĩa thống kê về xác suất ta có:

$$P\left(A_iB_j\right) \approx \frac{n_{ij}}{n},\ P\left(A_i\right) \approx \frac{n_i}{n} \text{ và } P\left(B_j\right) \approx \frac{m_j}{n}.$$

Nếu H_0 đúng có nghĩa là A và B độc lập nhau. Khi đó

$$P\left(A_iB_j\right) = P\left(A_i\right)P\left(B_j\right) \Leftrightarrow \frac{n_{ij}}{n} = \frac{n_i}{n}.\frac{m_j}{n}$$

và thống kê

$$G = n\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{\left(\frac{n_{\text{ij}}}{n}-\frac{n_i}{n}.\frac{m_j}{n}\right)}{\frac{n_i}{n}.\frac{m_j}{n}} = n\left[\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n_{\text{ij}}}{n_i.m_j}-1\right] \sim \chi^2\left\{(k-1)(l-1)\right\}.$$

Với mức ý nghĩa α cho trước và nếu

$$G_{qs} = n\left[\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n_{\text{ij}}}{n_i.m_j}-1\right] > \chi^2_\alpha\left\{(k-1)(l-1)\right\}$$

thì bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.3.1 Để lập kế hoạch sản xuất, một công tuy đã làm một cuộc điều tra về sở thích của khách hàng về ba loại mẫu hàng khác nhau là A, B, C của cùng một loại hàng và thu được kết quả như sau:

	A	B	C	Tổng
Thích	43	30	42	115
Không thích	35	53	39	127
Không có ý kiến	22	17	19	58
Tổng	100	100	100	300

Có hay không sự phân biệt về sở thích của khách hàng đối với ba loại mẫu nói trên? Cho kết luận với mức ý nghĩa 5%.

Kiểm định giả thuyết về quy luật phân phối xác suất của biến ngẫu nhiên:

Giả sử chưa biết quy luật phân phối xác suất của biến ngẫu nhiên gốc X trong tổng thể nghiên cứu nhưng nếu có cơ sở để giả định rằng X tuân theo quy luật phân phối A nào đó. Khi đó ta đưa ra cặp giả thuyết thống kê sau:

H_0: X tuân theo quy luật phân phối xác suất A

H_1: X không tuân theo quy luật phân phối xác suất A

Trường hợp 1: X là biến ngẫu nhiên rời rạc.

Từ tổng thể, thực hiện một mẫu kích thước nvới bảng phân phối thực nghiệm như sau:

Giá trị X có thể nhận	x_1	x_2	...	x_k
Số lần xuất hiện x_i	m_1	m_2	...	m_k

Với $\sum_{i=1}^{k} m_i = n$. Nếu H_0 đúng có nghĩa là X tuân theo quy luật phân phối xác suất A thì về mặt lý thuyết ta tính được các giá trị xác suất tương ứng với x_i bằng $p_i = P(X = x_i), i = \overline{1,k}$. Khi đó tần số lý thuyết là $n_i = np_i, i = \overline{1,k}$. Xét tiêu chuẩn kiểm định

$$G = \sum_{i=1}^{k} \frac{(m_i - n_i)^2}{n_i} \sim \chi^2(k - r - 1)$$

trong đó r là tham số cần ước lượng của quy luật cần kiểm định bằng phương pháp ước lượng điểm hợp lý tối đa. Với mức ý nghĩa α cho trước, nếu

$$G_{qs} = \sum_{i=1}^{k} \frac{(m_i - n_i)^2}{n_i} > \chi^2_\alpha(k - r - 1)$$

thì bác bỏ H_0. Ngược lại, ta chưa có cơ sở để bác bỏ H_0, tạm thời vẫn chấp nhận giả thuyết này đối với mẫu đang xét.

Ví dụ 5.3.2 Số con của 2000 phụ nữ thủ đô dưới 25 tuổi với số liệu như sau:

Số con X	0	1	2	3	4
Số phụ nữ	1090	650	220	30	10

Với mức ý nghĩa $\alpha = 0.05$, có thể xem X tuân theo quy luật phân phối Poisson được không?

Trường hợp 2: X là biến ngẫu nhiên liên tục.

Từ tổng thể, thực hiện một mẫu kích thước nvới bảng phân phối thực nghiệm dưới dạng ghép lớp như sau:

$x_{i-1} - x_i$	$x_0 - x_1$	$x_1 - x_2$	...	$x_{k-1} - x_k$
Số lần xuất hiện	m_1	m_2	...	m_k

Nếu H_0 đúng thì về mặt lý thuyết ta tính được $p_i = P(x_{i-1} < X < x_i), i = \overline{1,k}$ và tần số lý thuyết là $n_i = np_i, i = \overline{1,k}$. Các bước kiểm định tiếp theo ta tiến hành tương tự như trường hợp X là biến rời rạc ở trên. Xét trường hợp cụ thể đối với việc kiểm định giả thuyết về dạng phân phối chuẩn. Khi đó

$$p_i = P(x_{i-1} < X < x_i) = \Phi_0\left(\frac{x_i - \mu}{\sigma}\right) - \Phi_0\left(\frac{x_{i-1} - \mu}{\sigma}\right).$$

Thông thường ta chưa biết μ và σ nên chúng được thay bằng các ước lượng điểm hợp lý tối đa tương ứng là $\overline{x}$ và $\sqrt{ms}$ với

$$ms = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \overline{x}\right)^2 .m_i$$

Khi đó

$$p_i \approx \Phi_0\left(\frac{x_i - \overline{x}}{\sqrt{ms}}\right) - \Phi_0\left(\frac{x_{i-1} - \overline{x}}{\sqrt{ms}}\right).$$

Ví dụ 5.3.3 Gặt 200 thửa ruộng của một vùng thu được các số liệu sau:

Năng suất (tạ/ha)	Số thửa ruộng tương ứng
4-6	15
6-8	26
8-10	25
10-12	30
12-14	26
14-16	21
16-18	24
18-20	20
20-22	13
	n=200

Với mức ý nghĩa $\alpha = 0.05$, có thể xem năng suất lúa của vùng đó tuân theo quy luật phân

phối chuẩn hay không?

Table A.1 Cumulative Binomial Probabilities

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

a. $n = 5$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	.951	.774	.590	.328	.237	.168	.078	.031	.010	.002	.001	.000	.000	.000	.000
	1	.999	.977	.919	.737	.633	.528	.337	.188	.087	.031	.016	.007	.000	.000	.000
x	2	1.000	.999	.991	.942	.896	.837	.683	.500	.317	.163	.104	.058	.009	.001	.000
	3	1.000	1.000	1.000	.993	.984	.969	.913	.812	.663	.472	.367	.263	.081	.023	.001
	4	1.000	1.000	1.000	1.000	.999	.998	.990	.969	.922	.832	.763	.672	.410	.226	.049

b. $n = 10$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	.904	.599	.349	.107	.056	.028	.006	.001	.000	.000	.000	.000	.000	.000	.000
	1	.996	.914	.736	.376	.244	.149	.046	.011	.002	.000	.000	.000	.000	.000	.000
	2	1.000	.988	.930	.678	.526	.383	.167	.055	.012	.002	.000	.000	.000	.000	.000
	3	1.000	.999	.987	.879	.776	.650	.382	.172	.055	.011	.004	.001	.000	.000	.000
	4	1.000	1.000	.998	.967	.922	.850	.633	.377	.166	.047	.020	.006	.000	.000	.000
x	5	1.000	1.000	1.000	.994	.980	.953	.834	.623	.367	.150	.078	.033	.002	.000	.000
	6	1.000	1.000	1.000	.999	.996	.989	.945	.828	.618	.350	.224	.121	.013	.001	.000
	7	1.000	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.474	.322	.070	.012	.000
	8	1.000	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.756	.624	.264	.086	.004
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.944	.893	.651	.401	.096

c. $n = 15$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
	2	1.000	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
	3	1.000	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
	4	1.000	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
	5	1.000	1.000	.998	.939	.852	.722	.403	.151	.034	.004	.001	.000	.000	.000	.000
	6	1.000	1.000	1.000	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
x	7	1.000	1.000	1.000	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
	8	1.000	1.000	1.000	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
	9	1.000	1.000	1.000	1.000	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000
	10	1.000	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.314	.164	.013	.001	.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.539	.352	.056	.005	.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.764	.602	.184	.036	.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.920	.833	.451	.171	.010
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.987	.965	.794	.537	.140

(continued)

Table A.1 Cumulative Binomial Probabilities *(cont.)*

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

d. $n = 20$

x	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.818	.358	.122	.012	.003	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.983	.736	.392	.069	.024	.008	.001	.000	.000	.000	.000	.000	.000	.000	.000
2	.999	.925	.677	.206	.091	.035	.004	.000	.000	.000	.000	.000	.000	.000	.000
3	1.000	.984	.867	.411	.225	.107	.016	.001	.000	.000	.000	.000	.000	.000	.000
4	1.000	.997	.957	.630	.415	.238	.051	.006	.000	.000	.000	.000	.000	.000	.000
5	1.000	1.000	.989	.804	.617	.416	.126	.021	.002	.000	.000	.000	.000	.000	.000
6	1.000	1.000	.998	.913	.786	.608	.250	.058	.006	.000	.000	.000	.000	.000	.000
7	1.000	1.000	1.000	.968	.898	.772	.416	.132	.021	.001	.000	.000	.000	.000	.000
8	1.000	1.000	1.000	.990	.959	.887	.596	.252	.057	.005	.001	.000	.000	.000	.000
9	1.000	1.000	1.000	.997	.986	.952	.755	.412	.128	.017	.004	.001	.000	.000	.000
10	1.000	1.000	1.000	.999	.996	.983	.872	.588	.245	.048	.014	.003	.000	.000	.000
11	1.000	1.000	1.000	1.000	.999	.995	.943	.748	.404	.113	.041	.010	.000	.000	.000
12	1.000	1.000	1.000	1.000	1.000	.999	.979	.868	.584	.228	.102	.032	.000	.000	.000
13	1.000	1.000	1.000	1.000	1.000	1.000	.994	.942	.750	.392	.214	.087	.002	.000	.000
14	1.000	1.000	1.000	1.000	1.000	1.000	.998	.979	.874	.584	.383	.196	.011	.000	.000
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.994	.949	.762	.585	.370	.043	.003	.000
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.984	.893	.775	.589	.133	.016	.000
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.965	.909	.794	.323	.075	.001
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.992	.976	.931	.608	.264	.017
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.997	.988	.878	.642	.182

(continued)

Table A.1 Cumulative Binomial Probabilities *(cont.)*

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

e. $n = 25$

		p														
		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.778	.277	.072	.004	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.974	.642	.271	.027	.007	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000
	2	.998	.873	.537	.098	.032	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000
	3	1.000	.966	.764	.234	.096	.033	.002	.000	.000	.000	.000	.000	.000	.000	.000
	4	1.000	.993	.902	.421	.214	.090	.009	.000	.000	.000	.000	.000	.000	.000	.000
	5	1.000	.999	.967	.617	.378	.193	.029	.002	.000	.000	.000	.000	.000	.000	.000
	6	1.000	1.000	.991	.780	.561	.341	.074	.007	.000	.000	.000	.000	.000	.000	.000
	7	1.000	1.000	.998	.891	.727	.512	.154	.022	.001	.000	.000	.000	.000	.000	.000
	8	1.000	1.000	1.000	.953	.851	.677	.274	.054	.004	.000	.000	.000	.000	.000	.000
	9	1.000	1.000	1.000	.983	.929	.811	.425	.115	.013	.000	.000	.000	.000	.000	.000
	10	1.000	1.000	1.000	.994	.970	.902	.586	.212	.034	.002	.000	.000	.000	.000	.000
	11	1.000	1.000	1.000	.998	.980	.956	.732	.345	.078	.006	.001	.000	.000	.000	.000
x	12	1.000	1.000	1.000	1.000	.997	.983	.846	.500	.154	.017	.003	.000	.000	.000	.000
	13	1.000	1.000	1.000	1.000	.999	.994	.922	.655	.268	.044	.020	.002	.000	.000	.000
	14	1.000	1.000	1.000	1.000	1.000	.998	.966	.788	.414	.098	.030	.006	.000	.000	.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	.987	.885	.575	.189	.071	.017	.000	.000	.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	.996	.946	.726	.323	.149	.047	.000	.000	.000
	17	1.000	1.000	1.000	1.000	1.000	1.000	.999	.978	.846	.488	.273	.109	.002	.000	.000
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.993	.926	.659	.439	.220	.009	.000	.000
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.971	.807	.622	.383	.033	.001	.000
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.910	.786	.579	.098	.007	.000
	21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.967	.904	.766	.236	.034	.000
	22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.968	.902	.463	.127	.002
	23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.993	.973	.729	.358	.026
	24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.996	.928	.723	.222

Table A.2 Cumulative Poisson Probabilities

$$F(x; \mu) = \sum_{y=0}^{x} \frac{e^{-\mu}\mu^y}{y!}$$

		μ									
		.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
	0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368
	1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736
	2	1.000	.999	.996	.992	.986	.977	.966	.953	.937	.920
x	3		1.000	1.000	.999	.998	.997	.994	.991	.987	.981
	4				1.000	1.000	1.000	.999	.999	.998	.996
	5							1.000	1.000	1.000	.999
	6										1.000

(continued)

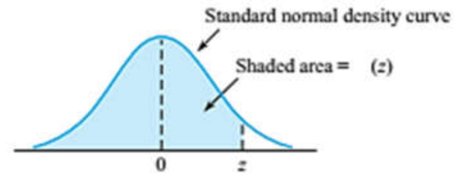
Table A.2 Cumulative Poisson Probabilities *(cont.)*

$$F(x;\mu)=\sum_{y=0}^{x}\frac{e^{-\mu}\mu^{y}}{y!}$$

	μ										
x	**2.0**	**3.0**	**4.0**	**5.0**	**6.0**	**7.0**	**8.0**	**9.0**	**10.0**	**15.0**	**20.0**
0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
6	.995	.966	.889	.762	.606	.450	.313	.207	.130	.008	.000
7	.999	.988	.949	.867	.744	.599	.453	.324	.220	.018	.001
8	1.000	.996	.979	.932	.847	.729	.593	.456	.333	.037	.002
9		.999	.992	.968	.916	.830	.717	.587	.458	.070	.005
10		1.000	.997	.986	.957	.901	.816	.706	.583	.118	.011
11			.999	.995	.980	.947	.888	.803	.697	.185	.021
12			1.000	.998	.991	.973	.936	.876	.792	.268	.039
13				.999	.996	.987	.966	.926	.864	.363	.066
14				1.000	.999	.994	.983	.959	.917	.466	.105
15					.999	.998	.992	.978	.951	.568	.157
16					1.000	.999	.996	.989	.973	.664	.221
17						1.000	.998	.995	.986	.749	.297
18							.999	.998	.993	.819	.381
19							1.000	.999	.997	.875	.470
20								1.000	.998	.917	.559
21									.999	.947	.644
22									1.000	.967	.721
23										.981	.787
24										.989	.843
25										.994	.888
26										.997	.922
27										.998	.948
28										.999	.966
29										1.000	.978
30											.987
31											.992
32											.995
33											.997
34											.999
35											.999
36											1.000

Table A.3 Standard Normal Curve Areas

(z) = $P(Z \leq z)$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0038
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3482
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

(continued)

Table A.3 Standard Normal Curve Areas *(cont.)* $\Phi(z) = P(Z \leq z)$

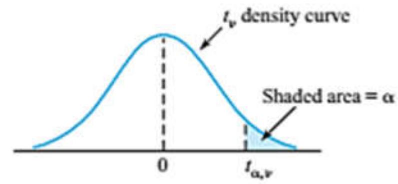
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table A.4 The Incomplete Gamma Function

$$F(x;\alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y}\, dy$$

x \ α	1	2	3	4	5	6	7	8	9	10
1	.632	.264	.080	.019	.004	.001	.000	.000	.000	.000
2	.865	.594	.323	.143	.053	.017	.005	.001	.000	.000
3	.950	.801	.577	.353	.185	.084	.034	.012	.004	.001
4	.982	.908	.762	.567	.371	.215	.111	.051	.021	.008
5	.993	.960	.875	.735	.560	.384	.238	.133	.068	.032
6	.998	.983	.938	.849	.715	.554	.394	.256	.153	.084
7	.999	.993	.970	.918	.827	.699	.550	.401	.271	.170
8	1.000	.997	.986	.958	.900	.809	.687	.547	.407	.283
9		.999	.994	.979	.945	.884	.793	.676	.544	.413
10		1.000	.997	.990	.971	.933	.870	.780	.667	.542
11			.999	.995	.985	.962	.921	.857	.768	.659
12			1.000	.998	.992	.980	.954	.911	.845	.758
13				.999	.996	.989	.974	.946	.900	.834
14				1.000	.998	.994	.986	.968	.938	.891
15					.999	.997	.992	.982	.963	.930

Table A.5 Critical Values for t Distributions

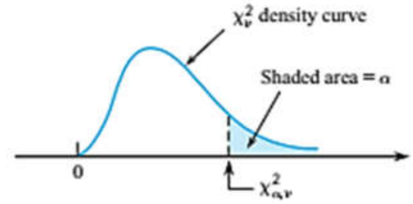


	α						
ν	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table A.6 Tolerance Critical Values for Normal Population Distributions

		Two-sided Intervals						One-sided Intervals					
Confidence Level		95%			99%			95%			99%		
% of Population Captured		≥ 90%	≥ 95%	≥ 99%	≥ 90%	≥ 95%	≥ 99%	≥ 90%	≥ 95%	≥ 99%	≥ 90%	≥ 95%	≥ 99%
	2	32.019	37.674	48.430	160.193	188.491	242.300	20.581	26.260	37.094	103.029	131.426	185.617
	3	8.380	9.916	12.861	18.930	22.401	29.055	6.156	7.656	10.553	13.995	17.370	23.896
	4	5.369	6.370	8.299	9.398	11.150	14.527	4.162	5.144	7.042	7.380	9.083	12.387
	5	4.275	5.079	6.634	6.612	7.855	10.260	3.407	4.203	5.741	5.362	6.578	8.939
	6	3.712	4.414	5.775	5.337	6.345	8.301	3.006	3.708	5.062	4.411	5.406	7.335
	7	3.369	4.007	5.248	4.613	5.488	7.187	2.756	3.400	4.642	3.859	4.728	6.412
	8	3.136	3.732	4.891	4.147	4.936	6.468	2.582	3.187	4.354	3.497	4.285	5.812
	9	2.967	3.532	4.631	3.822	4.550	5.966	2.454	3.031	4.143	3.241	3.972	5.389
	10	2.839	3.379	4.433	3.582	4.265	5.594	2.355	2.911	3.981	3.048	3.738	5.074
	11	2.737	3.259	4.277	3.397	4.045	5.308	2.275	2.815	3.852	2.898	3.556	4.829
	12	2.655	3.162	4.150	3.250	3.870	5.079	2.210	2.736	3.747	2.777	3.410	4.633
	13	2.587	3.081	4.044	3.130	3.727	4.893	2.155	2.671	3.659	2.677	3.290	4.472
	14	2.529	3.012	3.955	3.029	3.608	4.737	2.109	2.615	3.585	2.593	3.189	4.337
	15	2.480	2.954	3.878	2.945	3.507	4.605	2.068	2.566	3.520	2.522	3.102	4.222
	16	2.437	2.903	3.812	2.872	3.421	4.492	2.033	2.524	3.464	2.460	3.028	4.123
Sample Size n	17	2.400	2.858	3.754	2.808	3.345	4.393	2.002	2.486	3.414	2.405	2.963	4.037
	18	2.366	2.819	3.702	2.753	3.279	4.307	1.974	2.453	3.370	2.357	2.905	3.960
	19	2.337	2.784	3.656	2.703	3.221	4.230	1.949	2.423	3.331	2.314	2.854	3.892
	20	2.310	2.752	3.615	2.659	3.168	4.161	1.926	2.396	3.295	2.276	2.808	3.832
	25	2.208	2.631	3.457	2.494	2.972	3.904	1.838	2.292	3.158	2.129	2.633	3.601
	30	2.140	2.549	3.350	2.385	2.841	3.733	1.777	2.220	3.064	2.030	2.516	3.447
	35	2.090	2.490	3.272	2.306	2.748	3.611	1.732	2.167	2.995	1.957	2.430	3.334
	40	2.052	2.445	3.213	2.247	2.677	3.518	1.697	2.126	2.941	1.902	2.364	3.249
	45	2.021	2.408	3.165	2.200	2.621	3.444	1.669	2.092	2.898	1.857	2.312	3.180
	50	1.996	2.379	3.126	2.162	2.576	3.385	1.646	2.065	2.863	1.821	2.269	3.125
	60	1.958	2.333	3.066	2.103	2.506	3.293	1.609	2.022	2.807	1.764	2.202	3.038
	70	1.929	2.299	3.021	2.060	2.454	3.225	1.581	1.990	2.765	1.722	2.153	2.974
	80	1.907	2.272	2.986	2.026	2.414	3.173	1.559	1.965	2.733	1.688	2.114	2.924
	90	1.889	2.251	2.958	1.999	2.382	3.130	1.542	1.944	2.706	1.661	2.082	2.883
	100	1.874	2.233	2.934	1.977	2.355	3.096	1.527	1.927	2.684	1.639	2.056	2.850
	150	1.825	2.175	2.859	1.905	2.270	2.983	1.478	1.870	2.611	1.566	1.971	2.741
	200	1.798	2.143	2.816	1.865	2.222	2.921	1.450	1.837	2.570	1.524	1.923	2.679
	250	1.780	2.121	2.788	1.839	2.191	2.880	1.431	1.815	2.542	1.496	1.891	2.638
	300	1.767	2.106	2.767	1.820	2.169	2.850	1.417	1.800	2.522	1.476	1.868	2.608
	∞	1.645	1.960	2.576	1.645	1.960	2.576	1.282	1.645	2.326	1.282	1.645	2.326

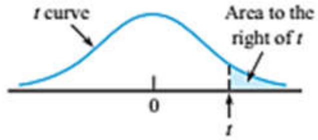
Table A.7 Critical Values for Chi-Squared Distributions



ν \ α	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

For $\nu > 40$, $\chi^2_{\alpha,\nu} \approx \nu\left(1 - \frac{2}{9\nu} + z_\alpha\sqrt{\frac{2}{9\nu}}\right)^3$

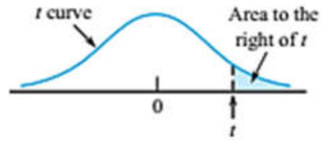
Table A.8 *t* Curve Tail Areas



t \ ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0.0	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
0.1	.468	.465	.463	.463	.462.	.462	.462	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461
0.2	.437	.430	.427	.426	.425	.424	.424	.423	.423	.423	.423	.422	.422	.422	.422	.422	.422	.422
0.3	.407	.396	.392	.390	.388	.387	.386	.386	.386	.385	.385	.385	.384	.384	.384	.384	.384	.384
0.4	.379	.364	.358	.355	.353	.352	.351	.350	.349	.349	.348	.348	.348	.347	.347	.347	.347	.347
0.5	.352	.333	.326	.322	.319	.317	.316	.315	.315	.314	.313	.313	.313	.312	.312	.312	.312	.312
0.6	.328	.305	.295	.290	.287	.285	.284	.283	.282	.281	.280	.280	.279	.279	.279	.278	.278	.278
0.7	.306	.278	.267	.261	.258	.255	.253	.252	.251	.250	.249	.249	.248	.247	.247	.247	.247	.246
0.8	.285	.254	.241	.234	.230	.227	.225	.223	.222	.221	.220	.220	.219	.218	.218	.218	.217	.217
0.9	.267	.232	.217	.210	.205	.201	.199	.197	.196	.195	.194	.193	.192	.191	.191	.191	.190	.190
1.0	.250	.211	.196	.187	.182	.178	.175	.173	.172	.170	.169	.169	.168	.167	.167	.166	.166	.165
1.1	.235	.193	.176	.167	.162	.157	.154	.152	.150	.149	.147	.146	.146	.144	.144	.144	.143	.143
1.2	.221	.177	.158	.148	.142	.138	.135	.132	.130	.129	.128	.127	.126	.124	.124	.124	.123	.123
1.3	.209	.162	.142	.132	.125	.121	.117	.115	.113	.111	.110	.109	.108	.107	.107	.106	.105	.105
1.4	.197	.148	.128	.117	.110	.106	.102	.100	.098	.096	.095	.093	.092	.091	.091	.090	.090	.089
1.5	.187	.136	.115	.104	.097	.092	.089	.086	.084	.082	.081	.080	.079	.077	.077	.077	.076	.075
1.6	.178	.125	.104	.092	.085	.080	.077	.074	.072	.070	.069	.068	.067	.065	.065	.065	.064	.064
1.7	.169	.116	.094	.082	.075	.070	.065	.064	.062	.060	.059	.057	.056	.055	.055	.054	.054	.053
1.8	.161	.107	.085	.073	.066	.061	.057	.055	.053	.051	.050	.049	.048	.046	.046	.045	.045	.044
1.9	.154	.099	.077	.065	.058	.053	.050	.047	.045	.043	.042	.041	.040	.038	.038	.038	.037	.037
2.0	.148	.092	.070	.058	.051	.046	.043	.040	.038	.037	.035	.034	.033	.032	.032	.031	.031	.030
2.1	.141	.085	.063	.052	.045	.040	.037	.034	.033	.031	.030	.029	.028	.027	.027	.026	.025	.025
2.2	.136	.079	.058	.046	.040	.035	.032	.029	.028	.026	.025	.024	.023	.022	.022	.021	.021	.021
2.3	.131	.074	.052	.041	.035	.031	.027	.025	.023	.022	.021	.020	.019	.018	.018	.018	.017	.017
2.4	.126	.069	.048	.037	.031	.027	.024	.022	.020	.019	.018	.017	.016	.015	.015	.014	.014	.014
2.5	.121	.065	.044	.033	.027	.023	.020	.018	.017	.016	.015	.014	.013	.012	.012	.012	.011	.011
2.6	.117	.061	.040	.030	.024	.020	.018	.016	.014	.013	.012	.012	.011	.010	.010	.010	.009	.009
2.7	.113	.057	.037	.027	.021	.018	.015	.014	.012	.011	.010	.010	.009	.008	.008	.008	.008	.007
2.8	.109	.054	.034	.024	.019	.016	.013	.012	.010	.009	.009	.008	.008	.007	.007	.006	.006	.006
2.9	.106	.051	.031	.022	.017	.014	.011	.010	.009	.008	.007	.007	.006	.005	.005	.005	.005	.005
3.0	.102	.048	.029	.020	.015	.012	.010	.009	.007	.007	.006	.006	.005	.004	.004	.004	.004	.004
3.1	.099	.045	.027	.018	.013	.011	.009	.007	.006	.006	.005	.005	.004	.004	.004	.003	.003	.003
3.2	.096	.043	.025	.016	.012	.009	.008	.006	.005	.005	.004	.004	.003	.003	.003	.003	.003	.002
3.3	.094	.040	.023	.015	.011	.008	.007	.005	.005	.004	.004	.003	.003	.002	.002	.002	.002	.002
3.4	.091	.038	.021	.014	.010	.007	.006	.005	.004	.003	.003	.003	.002	.002	.002	.002	.002	.002
3.5	.089	.036	.020	.012	.009	.006	.005	.004	.003	.003	.002	.002	.002	.002	.002	.001	.001	.001
3.6	.086	.035	.018	.011	.008	.006	.004	.004	.003	.002	.002	.002	.002	.001	.001	.001	.001	.001
3.7	.084	.033	.017	.010	.007	.005	.004	.003	.002	.002	.002	.002	.001	.001	.001	.001	.001	.001
3.8	.082	.031	.016	.010	.006	.004	.003	.003	.002	.002	.001	.001	.001	.001	.001	.001	.001	.001
3.9	.080	.030	.015	.009	.006	.004	.003	.002	.002	.001	.001	.001	.001	.001	.001	.001	.001	.001
4.0	.078	.029	.014	.008	.005	.004	.003	.002	.002	.001	.001	.001	.001	.001	.001	.001	.000	.000

(continued)

Table A.8 *t* Curve Tail Areas *(cont.)*



t \ ν	19	20	21	22	23	24	25	26	27	28	29	30	35	40	60	120	$\infty (= z)$
0.0	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
0.1	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.460	.460	.460	.460	.460
0.2	.422	.422	.422	.422	.422	.422	.422	.422	.421	.421	.421	.421	.421	.421	.421	.421	.421
0.3	.384	.384	.384	.383	.383	.383	.383	.383	.383	.383	.383	.383	.383	.383	.383	.382	.382
0.4	.347	.347	.347	.347	.346	.346	.346	.346	.346	.346	.346	.346	.346	.346	.345	.345	.345
0.5	.311	.311	.311	.311	.311	.311	.311	.311	.311	.310	.310	.310	.310	.310	.309	.309	.309
0.6	.278	.278	.278	.277	.277	.277	.277	.277	.277	.277	.277	.277	.276	.276	.275	.275	.274
0.7	.246	.246	.246	.246	.245	.245	.245	.245	.245	.245	.245	.245	.244	.244	.243	.243	.242
0.8	.217	.217	.216	.216	.216	.216	.216	.215	.215	.215	.215	.215	.215	.214	.213	.213	.212
0.9	.190	.189	.189	.189	.189	.189	.188	.188	.188	.188	.188	.188	.187	.187	.186	.185	.184
1.0	.165	.165	.164	.164	.164	.164	.163	.163	.163	.163	.163	.163	.162	.162	.161	.160	.159
1.1	.143	.142	.142	.142	.141	.141	.141	.141	.141	.140	.140	.140	.139	.139	.138	.137	.136
1.2	.122	.122	.122	.121	.121	.121	.121	.120	.120	.120	.120	.120	.119	.119	.117	.116	.115
1.3	.105	.104	.104	.104	.103	.103	.103	.103	.102	.102	.102	.102	.101	.101	.099	.098	.097
1.4	.089	.089	.088	.088	.087	.087	.087	.087	.086	.086	.086	.086	.085	.085	.083	.082	.081
1.5	.075	.075	.074	.074	.074	.073	.073	.073	.073	.072	.072	.072	.071	.071	.069	.068	.067
1.6	.063	.063	.062	.062	.062	.061	.061	.061	.061	.060	.060	.060	.059	.059	.057	.056	.055
1.7	.053	.052	.052	.052	.051	.051	.051	.051	.050	.050	.050	.050	.049	.048	.047	.046	.045
1.8	.044	.043	.043	.043	.042	.042	.042	.042	.042	.041	.041	.041	.040	.040	.038	.037	.036
1.9	.036	.036	.036	.035	.035	.035	.035	.034	.034	.034	.034	.034	.033	.032	.031	.030	.029
2.0	.030	.030	.029	.029	.029	.028	.028	.028	.028	.028	.027	.027	.027	.026	.025	.024	.023
2.1	.025	.024	.024	.024	.023	.023	.023	.023	.023	.022	.022	.022	.022	.021	.020	.019	.018
2.2	.020	.020	.020	.019	.019	.019	.019	.018	.018	.018	.018	.018	.017	.017	.016	.015	.014
2.3	.016	.016	.016	.016	.015	.015	.015	.015	.015	.015	.014	.014	.014	.013	.012	.012	.011
2.4	.013	.013	.013	.013	.012	.012	.012	.012	.012	.012	.012	.011	.011	.011	.010	.009	.008
2.5	.011	.011	.010	.010	.010	.010	.010	.010	.009	.009	.009	.009	.009	.008	.008	.007	.006
2.6	.009	.009	.008	.008	.008	.008	.008	.008	.007	.007	.007	.007	.007	.007	.006	.005	.005
2.7	.007	.007	.007	.007	.006	.006	.006	.006	.006	.006	.006	.006	.005	.005	.004	.004	.003
2.8	.006	.006	.005	.005	.005	.005	.005	.005	.005	.005	.005	.004	.004	.004	.003	.003	.003
2.9	.005	.004	.004	.004	.004	.004	.004	.004	.004	.004	.004	.003	.003	.003	.003	.002	.002
3.0	.004	.004	.003	.003	.003	.003	.003	.003	.003	.003	.003	.003	.002	.002	.002	.002	.001
3.1	.003	.003	.003	.003	.003	.002	.002	.002	.002	.002	.002	.002	.002	.002	.001	.001	.001
3.2	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002	.001	.001	.001	.001	.001
3.3	.002	.002	.002	.002	.002	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000
3.4	.002	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000
3.5	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000	.000
3.6	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000	.000	.000	.000
3.7	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
3.8	.001	.001	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3.9	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
4.0	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Table A.9 Critical Values for *F* Distributions

ν_2 = denominator df	α	ν_1 = numerator df 1	2	3	4	5	6	7	8	9
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
	.010	4052.20	4999.50	5403.40	5624.60	5763.60	5859.00	5928.40	5981.10	6022.50
	.001	405,284	500,000	540,379	562,500	576,405	585,937	592,873	598,144	602,284
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.001	25.41	18.49	15.83	14.39	13.48	12.86	12.40	12.05	11.77
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.001	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.001	21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.001	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48

(continued)

Table A.9 Critical Values for *F* Distributions *(cont.)*

ν_1 = numerator df										
10	**12**	**15**	**20**	**25**	**30**	**40**	**50**	**60**	**120**	**1000**
60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.69	62.79	63.06	63.30
241.88	243.91	245.95	248.01	249.26	250.10	251.14	251.77	252.20	253.25	254.19
6055.80	6106.30	6157.30	6208.70	6239.80	6260.60	6286.80	6302.50	6313.00	6339.40	6362.70
605,621	610,668	615,764	620,908	624,017	626,099	628,712	630,285	631,337	633,972	636,301
9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.47	9.48	9.49
19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48	19.48	19.49	19.49
99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.48	99.49	99.50
999.40	999.42	999.43	999.45	999.46	999.47	999.47	999.48	999.48	999.49	999.50
5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15	5.15	5.14	5.13
8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.58	8.57	8.55	8.53
27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.35	26.32	26.22	26.14
129.25	128.32	127.37	126.42	125.84	125.45	124.96	124.66	124.47	123.97	123.53
3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.80	3.79	3.78	3.76
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.70	5.69	5.66	5.63
14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.69	13.65	13.56	13.47
48.05	47.41	46.76	46.10	45.70	45.43	45.09	44.88	44.75	44.40	44.09
3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.15	3.14	3.12	3.11
4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.44	4.43	4.40	4.37
10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.24	9.20	9.11	9.03
26.92	26.42	25.91	25.39	25.08	24.87	24.60	24.44	24.33	24.06	23.82
2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.77	2.76	2.74	2.72
4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.75	3.74	3.70	3.67
7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.09	7.06	6.97	6.89
18.41	17.99	17.56	17.12	16.85	16.67	16.44	16.31	16.21	15.98	15.77
2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.52	2.51	2.49	2.47
3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.32	3.30	3.27	3.23
6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.86	5.82	5.74	5.66
14.08	13.71	13.32	12.93	12.69	12.53	12.33	12.20	12.12	11.91	11.72
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.35	2.34	2.32	2.30
3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.02	3.01	2.97	2.93
5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.07	5.03	4.95	4.87
11.54	11.19	10.84	10.48	10.26	10.11	9.92	9.80	9.73	9.53	9.36
2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.22	2.21	2.18	2.16
3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.80	2.79	2.75	2.71
5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.52	4.48	4.40	4.32
9.89	9.57	9.24	8.90	8.69	8.55	8.37	8.26	8.19	8.00	7.84
2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.12	2.11	2.08	2.06
2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.64	2.62	2.58	2.54
4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.12	4.08	4.00	3.92
8.75	8.45	8.13	7.80	7.60	7.47	7.30	7.19	7.12	6.94	6.78
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.04	2.03	2.00	1.98
2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.51	2.49	2.45	2.41
4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.81	3.78	3.69	3.61
7.92	7.63	7.32	7.01	6.81	6.68	6.52	6.42	6.35	6.18	6.02
2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.97	1.96	1.93	1.91
2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.40	2.38	2.34	2.30
4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.57	3.54	3.45	3.37
7.29	7.00	6.71	6.40	6.22	6.09	5.93	5.83	5.76	5.59	5.44

(continued)

Table A.9 Critical Values for *F* Distributions *(cont.)*

		ν_1 = numerator df								
ν_2 = denominator df	α	1	2	3	4	5	6	7	8	9
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.001	17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.001	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.001	16.12	10.97	9.01	7.94	7.27	6.80	6.46	6.19	5.98
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.010	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68
	.001	15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.001	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.001	15.08	10.16	8.28	7.27	6.62	6.18	5.85	5.59	5.39
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.001	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24
21	.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.001	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.001	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.001	14.20	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.89
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.001	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80

(continued)

Table A.9 Critical Values for *F* Distributions *(cont.)*

ν_1 = numerator df										
10	12	15	20	25	30	40	50	60	120	1000
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.92	1.90	1.88	1.85
2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.31	2.30	2.25	2.21
4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.38	3.34	3.25	3.18
6.80	6.52	6.23	5.93	5.75	5.63	5.47	5.37	5.30	5.14	4.99
2.10	2.05	2.01	1.96	1.93	1.91	1.89	1.87	1.86	1.83	1.80
2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.24	2.22	2.18	2.14
3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.22	3.18	3.09	3.02
6.40	6.13	5.85	5.56	5.38	5.25	5.10	5.00	4.94	4.77	4.62
2.06	2.02	1.97	1.92	1.89	1.87	1.85	1.83	1.82	1.79	1.76
2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.18	2.16	2.11	2.07
3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.08	3.05	2.96	2.88
6.08	5.81	5.54	5.25	5.07	4.95	4.80	4.70	4.64	4.47	4.33
2.03	1.99	1.94	1.89	1.86	1.84	1.81	1.79	1.78	1.75	1.72
2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.12	2.11	2.06	2.02
3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.97	2.93	2.84	2.76
5.81	5.55	5.27	4.99	4.82	4.70	4.54	4.45	4.39	4.23	4.08
2.00	1.96	1.91	1.86	1.83	1.81	1.78	1.76	1.75	1.72	1.69
2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.08	2.06	2.01	1.97
3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.87	2.83	2.75	2.66
5.58	5.32	5.05	4.78	4.60	4.48	4.33	4.24	4.18	4.02	3.87
1.98	1.93	1.89	1.84	1.80	1.78	1.75	1.74	1.72	1.69	1.66
2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.04	2.02	1.97	1.92
3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.78	2.75	2.66	2.58
5.39	5.13	4.87	4.59	4.42	4.30	4.15	4.06	4.00	3.84	3.69
1.96	1.91	1.86	1.81	1.78	1.76	1.73	1.71	1.70	1.67	1.64
2.38	2.31	2.23	2.16	2.11	2.07	2.03	2.00	1.98	1.93	1.88
3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.71	2.67	2.58	2.50
5.22	4.97	4.70	4.43	4.26	4.14	3.99	3.90	3.84	3.68	3.53
1.94	1.89	1.84	1.79	1.76	1.74	1.71	1.69	1.68	1.64	1.61
2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.97	1.95	1.90	1.85
3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.64	2.61	2.52	2.43
5.08	4.82	4.56	4.29	4.12	4.00	3.86	3.77	3.70	3.54	3.40
1.92	1.87	1.83	1.78	1.74	1.72	1.69	1.67	1.66	1.62	1.59
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.94	1.92	1.87	1.82
3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.58	2.55	2.46	2.37
4.95	4.70	4.44	4.17	4.00	3.88	3.74	3.64	3.58	3.42	3.28
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.60	1.57
2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.91	1.89	1.84	1.79
3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.53	2.50	2.40	2.32
4.83	4.58	4.33	4.06	3.89	3.78	3.63	3.54	3.48	3.32	3.17
1.89	1.84	1.80	1.74	1.71	1.69	1.66	1.64	1.62	1.59	1.55
2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.88	1.86	1.81	1.76
3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.48	2.45	2.35	2.27
4.73	4.48	4.23	3.96	3.79	3.68	3.53	3.44	3.38	3.22	3.08
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.57	1.54
2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.86	1.84	1.79	1.74
3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.44	2.40	2.31	2.22
4.64	4.39	4.14	3.87	3.71	3.59	3.45	3.36	3.29	3.14	2.99

(continued)

Table A.9 Critical Values for *F* Distributions *(cont.)*

ν_2 = denominator df	α	ν_1 = numerator df 1	2	3	4	5	6	7	8	9
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.001	13.88	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.71
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.001	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.001	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.001	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.001	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.001	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.001	12.61	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02
50	.100	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76
	.050	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	.010	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78
	.001	12.22	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.001	11.97	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
100	.100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69
	.050	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97
	.010	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
	.001	11.50	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.44
200	.100	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66
	.050	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93
	.010	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50
	.001	11.15	7.15	5.63	4.81	4.29	3.92	3.65	3.43	3.26
1000	.100	2.71	2.31	2.09	1.95	1.85	1.78	1.72	1.68	1.64
	.050	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89
	.010	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43
	.001	10.89	6.96	5.46	4.65	4.14	3.78	3.51	3.30	3.13

(continued)

Table A.9 Critical Values for *F* Distributions *(cont.)*

ν_1 = numerator df										
10	**12**	**15**	**20**	**25**	**30**	**40**	**50**	**60**	**120**	**1000**
1.87	1.82	1.77	1.72	1.68	1.66	1.63	1.61	1.59	1.56	1.52
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.84	1.82	1.77	1.72
3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.40	2.36	2.27	2.18
4.56	4.31	4.06	3.79	3.63	3.52	3.37	3.28	3.22	3.06	2.91
1.86	1.81	1.76	1.71	1.67	1.65	1.61	1.59	1.58	1.54	1.51
2.22	2.15	2.07	1.99	1.94	1.90	1.85	1.82	1.80	1.75	1.70
3.09	2.96	2.81	2.66	2.57	2.50	2.42	2.36	2.33	2.23	2.14
4.48	4.24	3.99	3.72	3.56	3.44	3.30	3.21	3.15	2.99	2.84
1.85	1.80	1.75	1.70	1.66	1.64	1.60	1.58	1.57	1.53	1.50
2.20	2.13	2.06	1.97	1.92	1.88	1.84	1.81	1.79	1.73	1.68
3.06	2.93	2.78	2.63	2.54	2.47	2.38	2.33	2.29	2.20	2.11
4.41	4.17	3.92	3.66	3.49	3.38	3.23	3.14	3.08	2.92	2.78
1.84	1.79	1.74	1.69	1.65	1.63	1.59	1.57	1.56	1.52	1.48
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.79	1.77	1.71	1.66
3.03	2.90	2.75	2.60	2.51	2.44	2.35	2.30	2.26	2.17	2.08
4.35	4.11	3.86	3.60	3.43	3.32	3.18	3.09	3.02	2.86	2.72
1.83	1.78	1.73	1.68	1.64	1.62	1.58	1.56	1.55	1.51	1.47
2.18	2.10	2.03	1.94	1.89	1.85	1.81	1.77	1.75	1.70	1.65
3.00	2.87	2.73	2.57	2.48	2.41	2.33	2.27	2.23	2.14	2.05
4.29	4.05	3.80	3.54	3.38	3.27	3.12	3.03	2.97	2.81	2.66
1.82	1.77	1.72	1.67	1.63	1.61	1.57	1.55	1.54	1.50	1.46
2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.76	1.74	1.68	1.63
2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.25	2.21	2.11	2.02
4.24	4.00	3.75	3.49	3.33	3.22	3.07	2.98	2.92	2.76	2.61
1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.42	1.38
2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.66	1.64	1.58	1.52
2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.06	2.02	1.92	1.82
3.87	3.64	3.40	3.14	2.98	2.87	2.73	2.64	2.57	2.41	2.25
1.73	1.68	1.63	1.57	1.53	1.50	1.46	1.44	1.42	1.38	1.33
2.03	1.95	1.87	1.78	1.73	1.69	1.63	1.60	1.58	1.51	1.45
2.70	2.56	2.42	2.27	2.17	2.10	2.01	1.95	1.91	1.80	1.70
3.67	3.44	3.20	2.95	2.79	2.68	2.53	2.44	2.38	2.21	2.05
1.71	1.66	1.60	1.54	1.50	1.48	1.44	1.41	1.40	1.35	1.30
1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.56	1.53	1.47	1.40
2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.88	1.84	1.73	1.62
3.54	3.32	3.08	2.83	2.67	2.55	2.41	2.32	2.25	2.08	1.92
1.66	1.61	1.56	1.49	1.45	1.42	1.38	1.35	1.34	1.28	1.22
1.93	1.85	1.77	1.68	1.62	1.57	1.52	1.48	1.45	1.38	1.30
2.50	2.37	2.22	2.07	1.97	1.89	1.80	1.74	1.69	1.57	1.45
3.30	3.07	2.84	2.59	2.43	2.32	2.17	2.08	2.01	1.83	1.64
1.63	1.58	1.52	1.46	1.41	1.38	1.34	1.31	1.29	1.23	1.16
1.88	1.80	1.72	1.62	1.56	1.52	1.46	1.41	1.39	1.30	1.21
2.41	2.27	2.13	1.97	1.87	1.79	1.69	1.63	1.58	1.45	1.30
3.12	2.90	2.67	2.42	2.26	2.15	2.00	1.90	1.83	1.64	1.43
1.61	1.55	1.49	1.43	1.38	1.35	1.30	1.27	1.25	1.18	1.08
1.84	1.76	1.68	1.58	1.52	1.47	1.41	1.36	1.33	1.24	1.11
2.34	2.20	2.06	1.90	1.79	1.72	1.61	1.54	1.50	1.35	1.16
2.99	2.77	2.54	2.30	2.14	2.02	1.87	1.77	1.69	1.49	1.22

Table A.10 Critical Values for Studentized Range Distributions

		m										
ν	α	2	3	4	5	6	7	8	9	10	11	12
5	.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32
	.01	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70
6	.05	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79
	.01	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.48
7	.05	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43
	.01	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71
8	.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18
	.01	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03	8.18
9	.05	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98
	.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65	7.78
10	.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83
	.01	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.49
11	.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71
	.01	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25
12	.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61
	.01	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06
13	.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53
	.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90
14	.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46
	.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77
15	.05	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40
	.01	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66
16	.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35
	.01	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56
17	.05	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21	5.31
	.01	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48
18	.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27
	.01	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41
19	.05	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23
	.01	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34
20	.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11	5.20
	.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.28
24	.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10
	.01	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11
30	.05	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92	5.00
	.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93
40	.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82	4.90
	.01	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69	5.76
60	.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81
	.01	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60
120	.05	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71
	.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37	5.44
	.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62
	.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29