Bài tập

XÁC SUẤT THỐNG KÊ

Chương 1

ĐẠI CƯƠNG VỀ XÁC SUẤT

A. BÀI TẬP MẪU

Bài 1. Cho A, B, C là ba biến cố. Chứng minh

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

Giải

Ta có

$$P(A \cup B \cup C) = P\left[(A \cup B) \cup C\right] = P(A \cup B) + P(C) - P[(A \cup B)C],$$

$$P(A \cup B) = P(A) + P(B) - P(AB),$$

$$P[(A \cup B)C] = P[AC \cup BC] = P(AC) + P(BC) - P(ABC)$$

nên

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

Bài 2. Cho $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{2}$ và $P(A + B) = \frac{3}{4}$.

Tính $P(AB)$, $P(\bar{A}\bar{B})$, $P(\bar{A} + \bar{B})$, $P(A\bar{B})$ và $P(\bar{A}B)$.

Giải

Do

$$P(A + B) = P(A) + P(B) - P(AB),$$

ta suy ra

$$P(AB) = P(A) + P(B) - P(A + B) = \frac{1}{12}.$$

Do $\bar{A}\bar{B} = \overline{A + B}$, nên

$$P\left(\bar{A}\bar{B}\right) = P\left(\overline{A + B}\right) = 1 - P(A + B) = \frac{1}{4}.$$

Tương tự, vì $\bar{A} + \bar{B} = \overline{AB}$ ta suy ra

$$P\left(\bar{A} + \bar{B}\right) = 1 - P(AB) = \frac{11}{12}.$$

Xuất phát từ đẳng thức $A = AB + A\bar{B}$ và vì AB, $A\bar{B}$ là các biến cố xung khắc, ta được $P(A) = P(AB) + P\left(A\bar{B}\right)$ và do đó

$$P\left(A\bar{B}\right) = P(A) - P(AB) = \frac{1}{4}.$$

Tương tự, ta có

$$P\left(\bar{A}B\right) = P(B) - P\left(AB\right) = \frac{5}{12}.$$

Bài 3. Tỷ lệ người mắc bệnh tim trong một vùng dân cư là 9%, mắc bệnh huyết áp là 12%, mắc cả hai bệnh là 7%. Chọn ngẫu nhiên một người trong vùng. Tính xác suất để người đó

a) Bị bệnh tim hay bị bệnh huyết áp.

b) Không bị bệnh tim cũng không bị bệnh huyết áp.

c) Không bị bệnh tim hay không bị bệnh huyết áp.

d) Bị bệnh tim nhưng không bị bệnh huyết áp.

e) Không bị bệnh tim nhưng bị bệnh huyết áp.

Giải

Xét các biến cố A : "nhận được người mắc bệnh tim",

B : "nhận được người mắc bệnh huyết áp",

Ta có $P(A) = 0.09$; $P(B) = 0.12$; $P(AB) = 0.07$.

a) Biến cố "nhận được người bị bệnh tim hay bị bệnh huyết áp" là A+B, với

$$\begin{aligned} P(A+B) &= P(A) + P(B) - P(AB) \\ &= 0.09 + 0.12 - 0.07 = 0.14. \end{aligned}$$

b) Biến cố "nhận được người không bị bệnh tim cũng không bị bệnh huyết áp" là $\overline{A}.\overline{B}$, với

$$\begin{aligned} P(\overline{A}.\overline{B}) &= P(\overline{A+B}) = 1 - P(A+B) \\ &= 1 - 0.14 = 0.86. \end{aligned}$$

c) Biến cố "nhận được người không bị bệnh tim hay không bị bệnh huyết áp" là $\overline{A} + \overline{B}$, với

$$\begin{aligned} P(\overline{A} + \overline{B}) &= P(\overline{AB}) = 1 - P(AB) \\ &= 1 - 0.07 = 0.93. \end{aligned}$$

d) Biến cố "nhận được người bị bệnh tim nhưng không bị bệnh huyết áp" là $A.\overline{B}$, với

$$\begin{aligned} P(A.\overline{B}) &= P(A) - P(AB) \\ &= 0.09 - 0.07 = 0.02. \end{aligned}$$

e) Biến cố "nhận được người không bị bệnh tim nhưng bị bệnh huyết áp" là $\overline{A}.B$, với

$$\begin{aligned} P(\overline{A}.B) &= P(B) - P(AB) \\ &= 0.12 - 0.07 = 0.05. \end{aligned}$$

Bài 4. Một hộp đựng 10 phiếu trong đó có 2 phiếu trúng thưởng. Có 10 người lần lượt rút thăm. Tính xác suất nhận được phần thưởng của mỗi người.

Giải

Gọi T_k $(k = 1, 2, \ldots, 10)$ là biến cố "người thứ k nhận được phiếu trúng thưởng". Ta có

$$P(T_1) = \frac{2}{10} = \frac{1}{5} = 0.2,$$

$$\begin{aligned} P(T_2) &= P(T_1) \cdot P\left(T_2 \middle| T_1\right) + P\left(\overline{T}_1\right) \cdot P\left(T_2 \middle| \overline{T}_1\right) \\ &= \frac{1}{5} \cdot \frac{1}{9} + \frac{4}{5} \cdot \frac{2}{9} = \frac{1}{5} = 0.2, \end{aligned}$$

$$P(T_3) = P\left(\overline{T}_1\right)P\left(T_2|\overline{T}_1\right)P\left(T_3|\overline{T}_1T_2\right) + P(T_1)P\left(\overline{T}_2|T_1\right)P\left(T_3|T_1\overline{T}_2\right)$$
$$+ P\left(\overline{T}_1\right)P\left(\overline{T}_2|\overline{T}_1\right)P\left(T_3|\overline{T}_1\overline{T}_2\right)$$
$$= \frac{4}{5}\cdot\frac{2}{9}\cdot\frac{1}{8} + \frac{1}{5}\cdot\frac{8}{9}\cdot\frac{1}{8} + \frac{4}{5}\cdot\frac{7}{9}\cdot\frac{2}{8} = \frac{1}{5} = 0.2,$$

...

$$P(T_{10}) = \frac{1}{5} = 0.2.$$

Bài 5. Một bài thi trắc nghiệm gồm 12 câu hỏi, mỗi câu có 5 câu trả lời, trong đó chỉ có một câu đúng. Giả sử mỗi câu trả lời đúng, thí sinh được 4 điểm, mỗi câu trả lời sai, thí sinh bị trừ 1 điểm. Một thí sinh làm bài bằng cách chọn ngẫu nhiên các câu trả lời. Tìm xác suất để

a) thí sinh được 13 điểm,

b) thí sinh bị điểm âm.

Giải

Gọi X là số câu trả lời đúng trong 12 câu hỏi được trả lời một cách ngẫu nhiên. Ta có $X \sim B\left(12;\frac{1}{5}\right)$.

Xét sự tương quan giữa số câu trả lời đúng và số điểm nhận được tương ứng, ta có

Số câu đúng (X)	Số điểm
0	–12
1	–7
2	–2
3	3
4	8
5	13
6	18
7	23
8	28
9	33
10	38
11	43
12	48

a) Biến cố "thí sinh được 13 điểm" chính là biến cố $X = 5$, với xác suất

$$P(X=5) = C_{12}^{5}(0.2)^5(1-0.2)^{12-5}$$
$$= \frac{12!}{5!\times(12-5)!}\cdot(0.2)^5\cdot(0.8)^7$$
$$= 0.0532$$

b) Biến cố "thí sinh bị điểm âm" chính là biến cố $X \le 2$, với xác suất

$$P(X\le 2) = P(X=0)+P(X=1)+P(X=2)$$
$$= C_{12}^{0}(0.2)^0\cdot(0.8)^{12} + C_{12}^{1}(0.2)^1\cdot(0.8)^{11} + C_{12}^{2}(0.2)^2\cdot(0.8)^{10}$$
$$= (0.8)^{12} + 12\cdot(0.2)\cdot(0.8)^{11} + 66\cdot(0.2)^2\cdot(0.8)^{10} = 0.558.$$

Bài 6. Theo dõi dự báo thời tiết trên đài truyền hình (nắng, sương mù, mưa) và so sánh với thời tiết thực tế xảy ra, ta có bảng thống kê sau

Dự báo / Thực tế	Nắng	Sương mù	Mưa
Nắng	30	5	5
Sương mù	4	20	2
Mưa	10	4	20

nghĩa là có 30 lần dự báo nắng, trời nắng, 4 lần dự báo nắng, trời sương mù; 10 lần dự báo nắng, trời mưa, v.v...

a) Tính xác suất dự báo trời nắng của đài truyền hình.

b) Tính xác suất dự báo của đài truyền hình là đúng thực tế.

c) Được tin dự báo là trời nắng. Tính xác suất để thực tế thì trời mưa ? trời sương mù ? trời nắng ?

Giải

Xét các biến cố A : "Đài truyền hình dự báo trời nắng", A_1 : "Thực tế trời nắng".

B : "Đài truyền hình dự báo trời sương mù", B_1 : "Thực tế trời sương mù".

C : "Đài truyền hình dự báo trời mưa", C_1 : "Thực tế trời mưa".

a) Do trong 100 lần theo dõi dự báo đài truyền hình, ta thấy có $30+4+10$ lần dự báo trời nắng nên xác suất dự báo trời nắng của đài truyền hình là

$$P(A)=\frac{30+4+10}{100}=0.44.$$

b) Do trong 100 lần theo dõi, ta thấy có $30+20+20$ dự báo của đài truyền hình đúng so với thực tế nên xác suất dự báo của đài truyền hình đúng so với thực tế là

$$\frac{30+20+20}{100}=0.7.$$

c) Do trong 44 lần đài truyền hình dự báo là trời nắng có 30 lần thực tế trời nắng, 4 lần thực tế trời sương mù và 10 lần thực tế trời mưa nên xác suất để thực tế thì trời mưa, trời sương mù, trời nắng lần lượt là

$$P\left(A_1|A\right)=\frac{30}{44}=0.682,$$
$$P\left(B_1|A\right)=\frac{4}{44}=0.091,$$
$$P\left(C_1|A\right)=\frac{10}{44}=0.227.$$

Bài 7. Bạn quên mất số cuối cùng trong số điện thoại cần gọi (số điện thoại gồm 6 chữ số) và bạn chọn số cuối cùng này một cách ngẫu nhiên. Tính xác suất để bạn gọi đúng số điện thoại này mà không phải thử quá 3 lần. Nếu biết số cuối cùng là số lẻ thì xác suất này là bao nhiêu ?

Giải

Gọi A_i là biến cố "gọi đúng ở lần thứ i", $i=1,2,3$. Ta có A_1 là biến cố "gọi đúng khi thử một lần", $\bar{A}_1A_2$ là biến cố "gọi đúng khi phải thử hai lần" và $\bar{A}_1\bar{A}_2A_3$ là biến cố "gọi đúng khi phải thử ba lần". Do đó biến cố "gọi đúng khi không phải thử quá ba lần là $A=A_1+\bar{A}_1A_2+\bar{A}_1\bar{A}_2A_3$ với

$$\begin{aligned}P(A) &= P(A_1 + \bar{A}_1A_2 + \bar{A}_1\bar{A}_2A_3)\\ &= P(A_1) + P(\bar{A}_1)\cdot P(A_2 \mid \bar{A}_1) + P(\bar{A}_1)\cdot P(\bar{A}_2 \mid \bar{A}_1)\cdot P(A_3 \mid \bar{A}_1\bar{A}_2)\\ &= \frac{1}{10} + \frac{9}{10}\cdot\frac{1}{9} + \frac{9}{10}\cdot\frac{8}{9}\cdot\frac{1}{8} = \frac{3}{10}.\end{aligned}$$

Khi đã biết số cuối cùng là số lẻ thì khi đó các số để chọn quay chỉ còn giới hạn lại trong 5 trường hợp (số lẻ) nên công thức trên trở thành

$$P(A) = \frac{1}{5} + \frac{4}{5}\cdot\frac{1}{4} + \frac{4}{5}\cdot\frac{3}{4}\cdot\frac{1}{3} = \frac{3}{5} = 0.6.$$

Bài 8. Một người bắn bia với xác suất bắn trúng là $p = 0.7$.

a) Bắn liên tiếp 3 phát. Tính xác suất có ít nhất 1 lần trúng bia.

b) Hỏi phải bắn ít nhất mấy lần để có xác suất ít nhất một lần trúng bia ≥ 0.9.

Giải

Gọi X là số viên đạn trúng bia trong 3 phát. Ta có $X \sim B(n;p)$, với $n = 3$ và $p = 0.7$.

a) Xác xuất có ít nhất một lần trúng bia khi bắn liên tiếp 3 phát là

$$\begin{aligned}P(X \geq 1) &= 1 - P(X = 0)\\ &= 1 - C_3^0(0.7)^0(1-0.7)^{3-0}\\ &= 1 - (0.3)^3 = 0.973.\end{aligned}$$

b) Gọi n là số lần bắn để xác suất ít nhất một lần trúng bia ≥ 0.9. Do $X \sim B(n;p)$ với $p = 0.7$, nên xác suất có ít nhất 1 lần trúng bia trong n phát là

$$\begin{aligned}P(X \geq 1) &= 1 - P(X = 0)\\ &= 1 - C_n^0(0.7)^0(1-0.7)^{n-0}\\ &= 1 - (0.3)^n.\end{aligned}$$

Để $P(X \geq 1) \geq 0.9$, ta giải bất phương trình

$$1 - (0.3)^n \geq 0.9,$$

hay tương đương

$$(0.3)^n \leq 0.1.$$

Lấy lôgarít hai vế của bất phương trình trên, ta được

$$n \times \ln(0.3) \leq \ln(0.1).$$

Do $\ln(0.3) < 0$, ta suy ra

$$n \geq \frac{\ln(0.1)}{\ln(0.3)} \approx 1.91.$$

Vậy, cần phải bắn ít nhất 2 phát đạn để xác suất có ít nhất 1 lần trúng bia ≥ 0.9.

Bài 9. Có hai hộp đựng bi :

- Hộp H_1 đựng 20 bi trong đó có 5 bi đỏ và 15 bi trắng,

- Hộp H_2 đựng 15 bi trong đó có 6 bi đỏ và 9 bi trắng.

Lấy một bi ở hộp H_1, bỏ vào hộp H_2, trộn đều rồi lấy ra một bi. Tính xác suất nhận được bi đỏ ? bi trắng ?

Giải

Xét các biến cố

A : "Bi nhận được từ hộp H_2 là bi đỏ",

B : "Bi từ hộp H_1 bỏ sang hộp H_2 là bi đỏ".

Do giả thuyết, ta có

$$P(B) = \frac{5}{20} = \frac{1}{4};\ P(A|B) = \frac{7}{16};\ P(A|\overline{B}) = \frac{6}{16} = \frac{3}{8}.$$

Từ đó, suy ra xác suất nhận được bi đỏ

$$P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B}) = \frac{25}{64},$$

và xác suất nhận được bi trắng là

$$P(\overline{A}) = 1 - P(A) = \frac{39}{64}.$$

Bài 10. Một cặp trẻ sinh đôi có thể do cùng một trứng (sinh đôi thật) hay do hai trứng khác nhau sinh ra (sinh đôi giả). Các cặp sinh đôi thật luôn luôn có cùng giới tính. Các cặp sinh đôi giả thì giới tính của mỗi đứa độc lập với nhau và có xác suất là 0.5. Thống kê cho thấy 34% cặp sinh đôi là trai; 30% cặp sinh đôi là gái và 36% cặp sinh đôi có giới tính khác nhau.

a) Tính tỷ lệ cặp sinh đôi thật.

b) Tìm tỷ lệ cặp sinh đôi thật trong số các cặp sinh đôi có cùng giới tính.

Giải

Xét các biến cố

A : "nhận được cặp sinh đôi thật",

B : "nhận được cặp sinh đôi có cùng giới tính".

Do các cặp sinh đôi thật luôn luôn có cùng giới tính nên

$$P(B|A) = 1,$$

với các cặp sinh đôi giả thì giới tính của mỗi đứa độc lập nhau và có xác suất là 0.5 nên

$$P(B|\overline{A}) = P(\overline{B}|\overline{A}) = 0.5,$$

và do thống kê trên các cặp sinh đôi nhận được thì

$$P(B) = 0.3 + 0.34 = 0.64 \text{ và } P(\overline{B}) = 0.36.$$

a) Do công thức xác suất toàn phần,

$$\begin{aligned} P(B) &= P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) \\ &= P(B|A)P(A) + P(B|\overline{A})\left[1 - P(A)\right] \\ &= P(B|\overline{A}) + \left[P(B|A) - P(B|\overline{A})\right]P(A), \end{aligned}$$

ta suy ra

$$P(A)=\frac{P(B)-P\left(B\middle|\bar{A}\right)}{P(B|A)-P\left(B\middle|\bar{A}\right)}=\frac{0.64-0.5}{1-0.5}=0.28.$$

b) Do công thức Bayes,

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}=\frac{0.28}{0.64}=0.4375.$$

Bài 11. Một trung tâm chẩn đoán bệnh dùng một phép kiểm định T. Xác suất để một người đến trung tâm mà có bệnh là 0.8. Xác suất để người khám có bệnh khi phép kiểm định dương tính là 0.9 và xác suất để người khám không có bệnh khi phép kiểm định âm tính là 0.5. Tính các xác suất

a) phép kiểm định là dương tính,

b) phép kiểm định cho kết quả đúng.

Giải

Xét các biến cố

A : "nhận được người có bệnh",

B : "nhận được người có kiểm định dương tính".

Do giả thiết, ta có

$$P(A)=0.8;\ P(A|B)=0.9;\ P\left(\bar{A}\middle|\bar{B}\right)=0.5.$$

a) Do công thức xác suất toàn phần,

$$\begin{aligned}P(A)&=P(A|B)P(B)+P\left(A\middle|\bar{B}\right)P\left(\bar{B}\right)\\&=P(A|B)P(B)+P\left(A\middle|\bar{B}\right)\left[1-P(B)\right]\\&=P\left(A\middle|\bar{B}\right)+\left[P(A|B)-P\left(A\middle|\bar{B}\right)\right]P(B),\end{aligned}$$

mà $P\left(A\middle|\bar{B}\right)=1-P\left(\bar{A}\middle|\bar{B}\right)=0.5$, nên xác suất để phép kiểm định là dương tính cho bởi

$$P(B)=\frac{P(A)-P\left(A\middle|\bar{B}\right)}{P(A|B)-P\left(A\middle|\bar{B}\right)}=\frac{0.8-0.5}{0.9-0.5}=0.75.$$

b) Xác suất để phép kiểm định cho kết quả đúng là

$$\begin{aligned}P\left(AB+\bar{A}\bar{B}\right)&=P(AB)+P\left(\bar{A}\bar{B}\right)\\&=P(A|B)P(B)+P\left(\bar{A}\middle|\bar{B}\right)P\left(\bar{B}\right)\\&=0.7125.\end{aligned}$$

Bài 12. Một thiết bị gồm 3 cụm chi tiết, mỗi cụm bị hỏng không ảnh hưởng gì đến các cụm khác và chỉ cần một cụm bị hỏng thì thiết bị ngừng hoạt động. Xác suất để cụm thứ nhất bị hỏng trong ngày là 0.1, cụm thứ hai là 0.05 và cụm thứ ba là 0.15. Tìm xác suất để thiết bị không ngừng hoạt động trong ngày.

Giải

Xét các biến cố

A_i : "Cụm chi tiết thứ i bị hỏng", với $i=1,2,3$, B : “thiết bị không ngừng hoạt động”.

Do giả thiết, ta có

$$P(A_1)=0.1,\ P(A_2)=0.05,\ \text{và}\ P(A_3)=0.15.$$

Do A_1, A_2 và A_3 là họ các biến cố độc lập nên xác suất để thiết bị không ngừng hoạt động là

$$\begin{aligned}P(B)&=P(\bar{A}_1\bar{A}_2\bar{A}_3)=P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)\\&=0.9\times0.95\times0.85=0.7267.\end{aligned}$$

Bài 13. Một phân xưởng có 5 máy. Xác suất để trong một ca, mỗi máy bị hỏng là 0.1. Tìm xác suất để trong một ca, có đúng 2 máy bị hỏng.

Giải

Gọi X là số máy bị hỏng của phân xưởng trong một ca. Do biến cố các máy bị hỏng độc lập nhau nên X thỏa lược đồ Bernoulli, nghĩa là $X\sim B(5;0.1)$.

Do đó, xác suất để trong một ca, có đúng 2 máy bị hỏng là

$$P(X=2)=C_5^2(0.1)^2(1-0.1)^{5-2}=C_5^2(0.1)^2(0.9)^3=0.0729.$$

Bài 14. Tính xác suất để gieo con xúc xắc 10 lần, mặt một nút xuất hiện không quá 3 lần.

Giải

Gọi X là số lần mặt một nút xuất hiện trong 10 lần thảy. Ta có $X\sim B(10;\frac{1}{6})$. Do đó, xác suất để mặt một nút xuất hiện không quá 3 lần là

$$\begin{aligned}P(X\le3)&=P(X=0)+P(X=1)+P(X=2)+P(X=3)\\&=C_{10}^0\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{10}+C_{10}^1\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^9+C_{10}^2\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^8+C_{10}^3\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7\\&=\left(\frac{5}{6}\right)^{10}+10\cdot\left(\frac{1}{6}\right)^1\cdot\left(\frac{5}{6}\right)^9+45\cdot\left(\frac{1}{6}\right)^2\cdot\left(\frac{5}{6}\right)^8+120\cdot\left(\frac{1}{6}\right)^3\cdot\left(\frac{5}{6}\right)^7\\&=0.857.\end{aligned}$$

Bài 15. Tỷ lệ phế phẩm của một lô hàng (lớn) là 1%. Từ lô hàng này, lấy ra n sản phẩm. Hỏi n ít nhất phải là bao nhiêu để xác suất nhận được ít nhất một phế phẩm lớn hơn 0.95.

Giải

Gọi X là số phế phẩm nhận được trong n sản phẩm lấy ra từ lô hàng. Ta có $X\sim B(n;0.01)$. Khi đó xác suất để nhận được ít nhất một sản phẩm hỏng là

$$\begin{aligned}P(X\ge1)&=1-P(X=0)\\&=1-C_n^0(0.01)^0(1-0.01)^{n-0}\\&=1-(0.99)^n.\end{aligned}$$

Để tìm n sao cho xác suất nhận được ít nhất một sản phẩm hỏng lớn hơn 0.95, nghĩa là $P(X\ge1)>0.95$, ta giải bất phương trình

$$1-(0.99)^n>0.95.$$

Từ đó, suy ra $n>298.073$. Vậy cần phải lấy ra ít nhất 299 sản phẩm để xác suất trong đó có ít nhất một sản phẩm hỏng lớn hơn 0.95.

Bài 16. Một người viết n lá thư và bỏ ngẫu nhiên n lá thư này vào trong n phong bì đã viết sẵn địa chỉ. Tìm xác suất sao cho có ít nhất một lá thư được bỏ vào đúng phong bì.

Giải

Gọi A_j là biến cố "lá thư thứ j đến đúng người nhận", $j = \overline{1,n}$ và gọi A là biến cố "có ít nhất một lá thư đến đúng người nhận". Ta có $A = \bigcup_{j=1}^{n} A_j$ và do công thức cộng tổng quát cho n biến cố

$$P(A) = P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j) - \sum_{i<j} P(A_i A_j) +$$
$$+ \sum_{i<j<k} P(A_i A_j A_k) - \ldots + (-1)^{n-1} P\left(\bigcap_{j=1}^{n} A_j\right)$$

Do $P\left(A_j\right) = \frac{1}{n}$, với mọi j,

$$P\left(A_i A_j\right) = P\left(A_i \middle| A_j\right) P\left(A_j\right) = \tfrac{1}{n-1} \cdot \tfrac{1}{n} = \tfrac{(n-2)!}{n!} ,$$

với mọi $i < j$,

$$P\left(A_i A_j A_k\right) = P\left(A_i \middle| A_j A_k\right) P\left(A_j \middle| A_k\right) P\left(A_k\right) = \tfrac{1}{n-2} \cdot \tfrac{1}{n-1} \cdot \tfrac{1}{n} = \tfrac{(n-3)!}{n!},$$

với mọi $i < j < k$, ..., ta suy ra

$$P(A) = n\frac{1}{n} - C_n^2 \frac{(n-2)!}{n!} + C_n^3 \frac{(n-3)!}{n!} - \ldots + (-1)^{n-1} \frac{1}{n!}$$
$$= \sum_{k=1}^{n} (-1)^{k-1} \frac{1}{k!} \approx 1 - e^{-1}$$

khi n đủ lớn.

Bài 17. Một dây chuyền lắp ráp nhận các chi tiết từ hai nhà máy khác nhau. Tỷ lệ chi tiết do nhà máy thứ nhất cung cấp là 60%, của nhà máy thứ hai là 40%. Tỷ lệ chính phẩm của nhà máy thứ nhất là 90%, của nhà máy thứ hai là 85%. Lấy ngẫu nhiên một chi tiết trên dây chuyền và thấy rằng nó tốt. Tìm xác suất để chi tiết đó do nhà máy thứ nhất sản xuất.

Giải

Xét các biến cố

A : "nhận được sản phẩm tốt",

B_i : "nhận được sản phẩm do nhà máy thứ i sản xuất",

với $i = 1, 2$.

Từ giả thuyết, ta có

$$P(B_1) = \frac{60}{100} = 0.6\,;\; P(B_2) = \frac{40}{100} = 0.4\,;$$

$$P\left(A \middle| B_1\right) = 0.9\,;\; P\left(A \middle| B_2\right) = 0.85\,.$$

Do B_1, B_2 tạo thành họ đầy đủ các biến cố nên từ công thức Bayes, ta được xác suất để chi tiết tốt nhận được trên dây chuyền là do nhà máy thứ nhất sản xuất

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2)} = 0.614.$$

Bài 18. Trong một vùng dân cư, cứ 100 người thì có 30 người hút thuốc lá. Biết tỷ lệ người bị viêm họng trong số người hút thuốc lá là 60%, trong số người không hút thuốc lá là 30%. Khám ngẫu nhiên một người và thấy người đó bị viêm họng. Tìm xác suất để người đó hút thuốc lá. Nếu người đó không bị viêm họng thì xác suất để người đó hút thuốc lá là bao nhiêu.

Giải

Khám ngẫu nhiên một người trong vùng dân cư, xét các biến cố

A : "nhận được người hút thuốc lá",

B : "nhận được người bị viêm họng".

Giả thiết cho

$$P(A) = 0.3;\ P(B|A) = 0.6 \text{ và } P(B|\bar{A}) = 0.3.$$

Do người đó đã bị viêm họng nên từ công thức Bayes, ta suy ra xác suất để người đó hút thuốc lá là

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$
$$= \frac{0.6 \times 0.3}{0.6 \times 0.3 + 0.3 \times 0.7} = 0.4615.$$

Khi người đó không bị viêm họng thì xác suất để anh ta hút thuốc lá là

$$P(A|\bar{B}) = \frac{P(\bar{B}|A)P(A)}{P(\bar{B}|A)P(A) + P(\bar{B}|\bar{A})P(\bar{A})}$$
$$= \frac{0.4 \times 0.3}{0.4 \times 0.3 + 0.7 \times 0.7} = 0.1967.$$

Bài 19. a) Cho A, B là hai biến cố độc lập. Chứng minh rằng $\bar{A}, B$; $A, \bar{B}$ và $\bar{A}, \bar{B}$ cũng là các cặp biến cố độc lập.

b) Cho $A_1, A_2, \ldots, A_n$ là n biến cố độc lập. Chứng minh rằng $\bar{A}_1, A_2, \ldots, A_n$ cũng là n biến cố độc lập. Suy ra rằng nếu xét n biến cố $B_1, B_2, \ldots, B_n$, với $B_i = A_i$ hay $B_i = \bar{A}_i$, thì $B_1, B_2, \ldots, B_n$, cũng là n biến cố độc lập.

Giải

Vì $B = \bar{A}B + AB$, $\bar{A}B$ và AB là các biến cố xung khắc nên công thức cộng cho

$$P(\bar{A}B) = P(B) - P(AB)$$
$$= P(B) - P(A)P(B) = [1 - P(A)]P(B)$$
$$= P(\bar{A})P(B),$$

và do đó $\bar{A}$ và B là hai biến cố độc lập. Tương tự

$$\begin{aligned}P\left(A\bar{B}\right) &= P(A) - P(AB)\\ &= P(A) - P(A)P(B) = \left[1 - P(B)\right]P(A)\\ &= P(A)P\left(\bar{B}\right),\end{aligned}$$

và

$$\begin{aligned}P\left(\bar{A}\bar{B}\right) &= P\left(\bar{A}\right) - P\left(\bar{A}B\right)\\ &= P\left(\bar{A}\right) - P\left(\bar{A}\right)P(B) = \left[1 - P(B)\right]P\left(\bar{A}\right)\\ &= P\left(\bar{A}\right)P\left(\bar{B}\right).\end{aligned}$$

Do đó, $A, \bar{B}$ và $\bar{A}, \bar{B}$ cũng là các cặp biến cố độc lập.

b) Để chứng minh rằng họ các biến cố $\bar{A}_1, A_2, ..., A_n$ là độc lập, ta lấy một họ con bất kỳ gồm k biến cố khác nhau của nó. Nếu họ con này không chứa biến cố $\bar{A}_1$, ta có thể viết nó dưới dạng A_{i_1}, A_{i_2}, ..., A_{i_k}, với $2 \le i_1 < i_2 < ... < i_k \le n$, và do đó nó là họ con của họ các biến cố độc lập $A_1, A_2, ..., A_n$. Suy ra

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P\left(A_{i_j}\right).$$

Nếu họ này chứa biến cố $\bar{A}_1$, nghĩa là nó có dạng $\bar{A}_{i_1}$, A_{i_2}, ..., A_{i_k}, với $i_1 = 1$, $2 \le i_2 < ... < i_k \le n$. Do giả thiết A_1 và $\bigcap_{j=2}^{k} A_{i_j}$ là hai biến cố độc lập nên từ câu a), ta được $\bar{A}_1$ và $\bigcap_{j=2}^{k} A_{i_j}$ cũng độc lập. Do đó

$$\begin{aligned}P\left(\bar{A}_1 \cap \left(\bigcap_{j=2}^{k} A_{i_j}\right)\right) &= P\left(\bar{A}_1\right)P\left(\bigcap_{j=2}^{k} A_{i_j}\right) = P\left(\bar{A}_1\right)\prod_{j=2}^{k} P\left(A_{i_j}\right)\\ &= \prod_{j=1}^{k} P\left(A_{i_j}\right).\end{aligned}$$

Tóm lại họ các biến cố $\bar{A}_1, A_2, ..., A_n$ là độc lập.

Để chứng minh rằng họ các biến cố $B_1, B_2, ..., B_n$, với $B_i = A_i$ hay $B_i = \bar{A}_i$, cũng là n biến cố độc lập, ta dùng quy nạp trên số k các biến cố $B_i = \bar{A}_i$, với $k \le n$.

Không mất tính tổng quát, ta có thể giả sử $B_i = \bar{A}_i$ với i thay đổi từ 1 đến k và $B_i = A_i$ khi $i > k$.

Trường hợp $k = 1$ đã được khảo sát trong phần đầu câu b).

Giả sử họ $B_1, B_2, ..., B_n$, với $B_i = \bar{A}_i$ trong đó i thay đổi từ 1 đến k là họ các biến cố độc lập.

Xét họ $C_1, C_2, ..., C_n$ các biến cố với $C_i = \bar{A}_i$ khi i thay đổi từ 1 đến $k+1$, và $C_i = A_i$ với $i > k+1$. Do $C_i = B_i$ với $i \ne k+1$, hai họ $C_1, C_2, ..., C_n$ và $B_1, B_2, ..., B_n$ chỉ khác nhau đúng một phần tử là $C_{k+1} = \bar{A}_i \ne B_{k+1} = A_i$, và do đó, như trong trường hợp $k = 1$, $C_1, C_2, ..., C_n$ cũng là họ các biến cố độc lập.

Do đó, ta kết luận rằng họ các biến cố $B_1, B_2, ..., B_n$, với $B_i = A_i$ hay $B_i = \overline{A_i}$ cũng là n biến cố độc lập.

Bài 20. Hai nhà máy X, Y cùng sản xuất một loại sản phẩm. Xác suất nhận được sản phẩm hỏng ở nhà máy X là $p_X = 0.03$ và ở nhà máy Y là $p_Y = 0.05$.

a) Một người mua 3 sản phẩm ở nhà máy X. Tính xác suất có ít nhất một sản phẩm hỏng .

b) Nếu mua 3 sản phẩm ở nhà máy X và 2 sản phẩm ở nhà máy Y. Tính xác suất có ít nhất một sản phẩm hỏng .

Giải

Xét các biến cố

A : "nhận được sản phẩm hỏng của nhà máy X",

B : "nhận được sản phẩm hỏng của nhà máy Y".

Dựa theo giả thiết, ta có

$$P(A) = 0.03 \text{ và } P(B) = 0.05.$$

a) Gọi X là số sản phẩm hỏng trong 3 sản phẩm lấy ra từ nhà máy X. Ta có

$$X \sim B(n;p) \text{ với } n = 3 \text{ và } p = P(A) = 0.03.$$

Do đó, xác suất có ít nhất một sản phẩm hỏng là

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - C_3^0 (0.03)^0 (1 - 0.03)^3 = 0.087327. \end{aligned}$$

b) Gọi X là số sản phẩm hỏng trong 3 sản phẩm lấy ra từ nhà máy X và Y là số sản phẩm hỏng trong 2 sản phẩm lấy ra từ nhà máy Y, thì

$$X \sim B(n;p) \text{ với } n = 3,\ p = P(A) = 0.03\ ,$$

và

$$Y \sim B(n;p) \text{ với } n = 2,\ p = P(B) = 0.05.$$

Do "số sản phẩm hỏng nhận được từ nhà máy X" và "số sản phẩm hỏng nhận được từ nhà máy Y" là các biến cố độc lập và biến cố "nhận được ít nhất một sản phẩm hỏng trong 5 sản phẩm, 3 sản phẩm từ nhà máy X và 2 sản phẩm từ nhà máy Y", $X + Y \geq 1$, có biến cố đối lập là biến cố "$X = 0$ và $Y = 0$" nên xác suất để nhận ít nhất 1 sản phẩm hỏng khi mua 3 sản phẩm của nhà máy X và 2 sản phẩm của nhà máy Y là

$$\begin{aligned} P(X + Y \geq 1) &= 1 - P(X = 0; Y = 0) = 1 - P(X = 0)P(Y = 0) \\ &= 1 - (0.97)^3 (0.95)^2 = 0.1763. \end{aligned}$$

Bài 21. Trong một lô thuốc (rất nhiều) với xác suất nhận được thuốc hỏng là $p = 0.1$. Lấy ngẫu nhiên 3 lọ để kiểm tra. Tính xác suất để

a) cả 3 lọ đều hỏng,

b) có 2 lọ hỏng và 1 lọ tốt,

c) có 1 lọ hỏng và 2 lọ tốt,

d) cả 3 lọ đều tốt.

Giải

Gọi X là số lọ hỏng trong 3 lọ lấy ra để kiểm tra. Ta có $X \sim B(3;0.1)$. Do đó xác suất để

a) cả 3 lọ đều hỏng

$$P(X=3) = C_3^3(0.1)^3(1-0.1)^0 = (0.1)^3 = 0.001,$$

b) có hai lọ hỏng và một lọ tốt

$$P(X=2) = C_3^2(0.1)^2(0.9)^{3-2} = 3\times 0.01\times 0.9 = 0.027,$$

c) có một lọ hỏng và hai lọ tốt

$$P(X=1) = C_3^1(0.1)^1(0.9)^{3-1} = 3\times 0.1\times 0.81 = 0.243,$$

d) cả 3 lọ đều tốt

$$P(X=0) = C_3^0(0.1)^0(1-0.1)^3 = (0.9)^3 = 0.729.$$

B. BÀI TẬP

Bài toán về biểu diễn các biến cố.

Bài 1. Kiểm tra 3 sản phẩm. Gọi A_k là biến cố sản phẩm thứ k tốt. Hãy trình bày các cách biểu diễn qua A_k và qua giản đồ Venn các biến cố sau đây :

A : tất cả đều xấu,

B : có ít nhất một sản phẩm xấu,

C : có ít nhất một sản phẩm tốt,

D : không phải tất cả sản phẩm đều tốt,

E : có đúng một sản phẩm xấu,

F : có ít nhất 2 sản phẩm tốt.

Bài 2. Ba người, mỗi người bắn một phát. Gọi A_i là biến cố người thứ i bắn trúng. Hãy biểu diễn qua A_i các biến cố sau :

A : chỉ có người thứ nhất bắn trúng,

B : người thứ nhất bắn trúng còn người thứ hai bắn trật,

C : có ít nhất 1 người bắn trúng,

D : cả 3 người đều bắn trúng,

E : có ít nhất 2 người bắn trúng,

F : chỉ có 2 người bắn trúng,

G : không ai bắn trúng,

H : không có hơn 2 người bắn trúng,

I : người thứ nhất bắn trúng, hoặc người thứ hai và người thứ ba cùng bắn trúng,

K : người thứ nhất bắn trúng hay người thứ hai bắn trúng.

Bài 3. Quan sát 4 sinh viên làm bài thi. Kí hiệu $B_j (j = 1,2,3,4)$ là biến cố sinh viên j làm bài thi đạt yêu cầu. Hãy biểu diễn các biến cố sau đây

a) có đúng một sinh viên đạt yêu cầu,

b) có đúng 3 sinh viên đạt yêu cầu,

c) có ít nhất 1 sinh viên đạt yêu cầu,

d) không có sinh viên nào đạt yêu cầu.

Xác suất bằng định nghĩa.

Bài 4. Một hộp có 7 bi đỏ và 3 bi đen.

a) Lấy ngẫu nhiên 1 viên bi từ hộp ra để kiểm tra, tính xác suất nhận được bi đen.

b) Lấy ngẫu nhiên lần lượt có hoàn lại 2 bi. Tính xác suất để lấy được 2 bi đen.

c) Lấy ngẫu nhiên ra 2 viên bi từ hộp. Tính xác suất để lấy được 2 bi đen.

Đáp số : a) 0.3.
b) 0.09.
c) 0.067.

Bài 5. Một công ty liên doanh cần tuyển một kế toán trưởng, một trưởng phòng tiếp thị, có 40 người dự tuyển trong đó có 15 nữ. Tính xác suất trong 2 người được tuyển có:

a) ít nhất 1 nữ,

b) 1 nữ,

c) kế toán trưởng là nữ.

Đáp số : a) 0.616.
b) 0.481.
c) 0.75.

Bài 6. Mỗi sinh viên được thi tối đa 2 lần một môn thi. Xác suất để một sinh viên đậu môn xác suất thống kê ở lần thi thứ 1 là P_1, lần thi thứ 2 là P_2. Tính xác suất để sinh viên này vượt qua được môn xác suất thống kê.

Đáp số : $P_1 + (1 - P_1)P_2$.

Bài 7. Gieo đồng thời 2 con xúc xắc cân đối, đồng chất. Tính xác suất để tổng số nút xuất hiện là 6.

Đáp số : $\frac{5}{36} = 0.139$

Bài 8. Trước cổng trường đại học có 3 quán cơn bình dân chất lượng ngang nhau. Ba sinh viên A, B, C độc lập với nhau chọn ngẫu nhiên một quán cơm để ăn trưa. Tính xác suất để

a) 3 sinh viên vào cùng một quán.

b) 2 sinh viên vào cùng một quán, còn người kia thì vào quán khác.

Đáp số : a) $\frac{1}{9}$.
b) $\frac{2}{3}$.

Bài 9. Một lô hàng có 10 sản phẩm, trong đó có 7 sản phẩm tốt, 3 sản phẩm xấu. Lấy ngẫu nhiên từ lô hàng ra 4 sản phẩm. Tính xác suất để 4 sản phẩm lấy ra có 3 sản phẩm tốt.

Đáp số : 0.5.

Bài 10. Trong hộp có 4 bi trắng, 6 bi đỏ cùng kích cỡ. Rút hú họa 2 bi. Tính xác suất để trong đó có

a) hai viên bi trắng,

b) ít nhất một viên bi đỏ,

c) viên thứ 2 đỏ.

Đáp số : *a)* 0.133.
b) 0.867.
c) 0.867

Bài 11. Chọn lần lượt không hoàn lại 2 con domino từ bộ 28 con. Tính xác suất chọn được 2 con domino có thể sắp nối tiếp nhau.

Đáp số : 0.238.

Bài 12. Rút ngẫu nhiên từ bộ bài (gồm 52 lá) ra 9 quân bài. Tính xác suất sao cho trong 9 quần bài rút ra có

a) 3 con Át, 2 con 10, 2 con 2, 1 con K, 1 con J,

b) 3 con cơ, 1 con rô, 2 con bích, 3 con chuồn,

c) 5 con màu đỏ, 4 con màu đen,

d) 4 con chủ bài (4 con đồng chất nào đó; chất đó đã được xác định trước, chẳng hạn 4 con cơ).

Đáp số : *a)* 6.262×10^{-7}.
b) 0.02254.
c) 0.2673.
d) 0.448.

Công thức cộng – nhân – xác suất có điều kiện.

Bài 13. Trong 100 người phỏng vấn có 40 người thích dùng nước hoa A, 28 người thích dùng nước hoa B, 10 người thích dùng cả 2 loại A, B. Chọn ngẫu nhiên 1 người trong số 100 người trên. Tính xác suất người này :

a) thích dùng ít nhất 1 loại nước hoa trên,

b) không dùng loại nào cả.

Đáp số : *a)* 0.58.
b) 0.42.

Bài 14. Một cơ quan có 210 người, trong đó có 100 người ở gần cơ quan, 60 người trong 100 người là nữ, biết rằng số nữ chiếm gấp đôi số nam trong cơ quan.

Chọn ngẫu nhiên 1 người trong cơ quan. Tính xác suất :

a) người này là nam,

b) người này ở gần cơ quan,

c) người này phải trực đêm (người trực đêm phải ở gần cơ quan hoặc là nam).

Đáp số : *a)* $\frac{1}{3}$.
b) 0.4762.
c) 0.619.

Bài 15. Có 3 loại súng bề ngoài hoàn toàn giống nhau, với xác suất bắn trúng bia tương ứng là 0.6, 0.7, 0.8. Loại thứ I có 5 khẩu, loại thứ II có 3 khẩu, loại thứ III có 2 khẩu. Chọn ngẫu nhiên 1 khẩu và bắn vào bia. Tính xác suất bắn trúng bia.

Đáp số : 0.67.

Bài 16. Cho 3 biến cố A, B, C sao cho

$P(A) = 0{,}5$; $P(B) = 0{,}7$; $P(C) = 0{,}6$;

$P(AB) = 0{,}3$; $P(BC) = 0{,}4$; $P(AC) = 0{,}2$

và $P(ABC) = 0{,}1$.

a) Tìm xác suất để cả 3 biến cố A, B, C đều không xảy ra.

b) Tìm xác suất để có đúng 2 trong 3 biến cố đó xảy ra.

c) Tìm xác suất để chỉ có đúng 1 biến cố trong 3 biến cố đó xảy ra.

Đáp số :a) 0.
b) 0.6.
c) 0.3.

Bài 17. Cho A và B là 2 biến cố sao cho $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$, $P(AB) = \frac{1}{6}$. Hãy tính :

1) $P(A \cup B)$, 8) $P(A|B)$,

2) $P(\bar{A} \cup \bar{B})$, 9) $P(\bar{A}|B)$,

3) $P(\overline{A \cup B})$, 10) $P(AB|B)$,

4) $P(\overline{AB})$, 11) $P(A\bar{B}|B)$,

5) $P(A\bar{B})$, 12) $P(A\bar{B}|\bar{B})$,

6) $P(\bar{A}B)$, 13) $P(A \cup B|A\bar{B})$,

7) $P(\bar{A} \cup B)$, 14) $P(\bar{A}B|\bar{A} \cup B)$.

Đáp số : 1) $\frac{2}{3}$. *2)* $\frac{5}{6}$. *3)* $\frac{1}{3}$. *4)* $\frac{5}{6}$. | *5)* $\frac{1}{3}$. *6)* $\frac{1}{6}$. *7)* $\frac{2}{3}$. *8)* $\frac{1}{2}$. | *9)* $\frac{1}{2}$. *10)* $\frac{1}{2}$. *11)* 0. *12)* $\frac{1}{2}$. *13)* 1. | *14)* $\frac{1}{4}$.

Bài 18. Đội tuyển bóng bàn của Khoa Kinh Tế có 3 vận động viên, mỗi vận động viên thi đấu một trận. Xác suất thắng trận của các vận viên A, B, C lần lượt là : 0.7; 0.8; 0.9. Tính xác suất :

a) đội tuyển thắng ít nhất 1 trận,

b) đội tuyển thắng 2 trận,

c) C thua, biết rằng đội tuyển thắng 2 trận.

Đáp số : a) 0.994.
b) 0.398.
c) 0.0621.

Bài 19. Trong 1 khu phố, tỷ lệ người mắc bệnh tim là 6%; mắc bệnh phổi là 8% và mắc cả hai bệnh là 5%. Chọn ngẫu nhiên 1 người trong khu phố đó. Tính xác suất để người đó không mắc cả 2 bệnh tim và bệnh phổi.

Đáp số : 0.91.

Bài 20. Một người có 5 con gà mái, 2 con gà trống nhốt chung trong một cái lồng. Một người đến mua, người bán gà bắt ngẫu nhiên 1 con. Người mua chấp nhận con đó.

a) Tính xác suất để người đó mua được con gà mái.

Người thứ hai lại đến mua, người bán gà lại bắt ngẫu nhiên ra 1 con.

b) Tìm xác suất để người thứ hai mua được con gà trống.

c) Xác suất này sẽ bằng bao nhiêu nếu người bán gà quên mất rằng con gà bán cho người thứ nhất là gà trống hay gà mái.

Đáp số : a) 0.7143.

b) $\frac{1}{3} = 0.33$.

c) $\frac{2}{7} = 0.2857$

Bài 21. Hai công ty A, B cùng kinh doanh một mặt hàng. Xác suất để công ty A thua lỗ là 0,2; xác suất để công ty B thua lỗ là 0,4. Tuy nhiên trên thực tế, khả năng cả 2 công ty cùng thua lỗ là 0,1. Tìm xác suất để

a) có ít nhất một công ty làm ăn không thua lỗ,

b) chỉ có một công ty thua lỗ.

Đáp số : a) 0.9.

b) 0.4.

Bài 22. Một thủ quỹ có một chùm chìa khóa gồm 12 chiếc bề ngoài giống hệt nhau, trong đó có 4 chiếc mở được cửa chính của thư viện. Cô ta thử từng chìa một một cách ngẫu nhiên, chìa nào không trúng thì bỏ ra. Tìm xác suất để cô ta mở được cửa chính của thư viện ở lần mở thứ 5.

Đáp số : 0.0707.

Bài 23. Một chàng trai viết 4 lá thư cho 4 cô gái; nhưng vì đãng trí nên anh ta bỏ 4 lá thư vào 4 phong bì một cách ngẫu nhiên, dán kín rồi mới ghi địa chỉ gửi,

a) tính xác suất để không có cô nào nhận đúng thư viết cho mình,

b) tính xác suất để có ít nhất 1 cô nhận đúng thư của mình,

c) tổng quát hóa với n cô gái. Tính xác suất có ít nhất 1 cô nhận đúng thư. Xấp xỉ giá trị xác suất này khi cho $n \to \infty$.

Bài 24. Trong 1 lô hàng 10 sản phẩm có 2 sản phẩm xấu, chọn không hoàn lại để phát hiện ra 2 sản phẩm xấu, khi nào chọn được sản phẩm xấu thứ 2 thì dừng lại.

a) Tính xác suất dừng lại ở lần chọn thứ 4.

b) Biết rằng đã chọn được sản phẩm xấu ở lần chọn thứ nhất, tính xác suất dừng lại ở lần chọn thứ 4.

c) Nếu việc kiểm tra dừng lại ở lần chọn thứ 3, tính xác suất lần chọn đầu được sản phẩm xấu.

Đáp số : a) 0.067.

b) $\frac{1}{7} = 0.143$.

c) 0.044.

Bài 25. Đội tuyển bóng bàn Thành phố có 4 vận động viên A, B, C, D . Mỗi vận động viên thi đấu 1 trận, với xác suất thắng trận lần lượt là : 0.6, 0.7, 0.8, 0.9. Tính

a) xác suất đội tuyển thắng ít nhất 1 trận,

b) xác suất đội tuyển thắng 2 trận,

c) xác suất đội tuyển thắng 3 trận,

d) xác suất D thua, trong trường hợp đội tuyển thắng 3 trận.

Đáp số : a) 0.9976.
b) 0.2144.

Bài 26. 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. Tìm 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 2 hỏng.

Đáp số : a) 0.004545.
b) 0.3818.
c) 0.9954.
d) 0.1636.

Bài 27. Ở một cơ quan nọ có 3 chiếc ôtô. Khả năng có sự cố của mỗi xe ôtô lần lượt là 0.15 ; 0.20 ; 0.10.

a) Tìm khả năng 3 ôtô cùng bị hỏng.

b) Tìm khả năng có ít nhất 1 ôtô hoạt động tốt.

c) Tìm khả năng cả 3 ôtô cùng hoạt động được.

d) Tìm xác suất có không quá 2 ôtô bị hỏng.

Đáp số : a) 0.003, *b)* 0.997.
c) 0.612, *d)* 0.997.

Công thức xác suất đầy đủ – Công thức Bayès.

Bài 28. Một hộp có 15 quả bóng bàn, trong đó có 9 mới 6 cũ, lần đầu chọn ra 3 quả để sử dụng, sau đó bỏ vào lại, lần hai chọn ra 3 quả.

a) Tính xác suất 3 quả bóng chọn lần hai là 3 bóng mới.

b) Biết rằng lần hai chọn được 3 bóng mới, tính xác suất lần đầu chọn được 2 bóng mới.

Đáp số : a) 0.0025.
b) 0.4091.

Bài 29. Một nhà máy sản xuất bóng đèn, máy A sản xuất 25%, máy B: 35%, máy C: 40% số bóng đèn. Tỉ lệ sản phẩm hỏng của mỗi máy trên số sản phẩm do máy đó sản xuất lần lượt là 3%, 2%, 1%. Một người mua 1 bóng đèn do nhà máy sản xuất.

a) Tính xác suất để sản phẩm này do máy A sản xuất.

b) Tính xác suất để sản phẩm này tốt.

c) Biết rằng sản phẩm này là xấu. Tính xác suất để sản phẩm do máy C sản xuất.

Đáp số : a) 0.25.
b) 0.9815.
c) 0.22.

Bài 30. Có 8 bình đựng bi, trong đó có :

2 bình loại 1: mỗi bình đựng 6 bi trắng 3 bi đỏ,

3 bình loại 2: mỗi bình đựng 5 bi trắng 4 bi đỏ,

3 bình loại 3: mỗi bình đựng 2 bi trắng 7 bi đỏ.

Lấy ngẫu nhiên một bình và từ bình đó lấy ngẫu nhiên 1 bi.

a) Tính xác suất để bi lấy ra là bi trắng.

b) Biết rằng bi lấy ra là bi trắng. Tính xác suất để bình lấy ra là bình loại 3.

Đáp số : a) 0.458.
b) 0.182.

Bài 31. Một bộ đề thi có 20 câu hỏi. Sinh viên giỏi sẽ trả lời đúng hết cả 20 câu. Sinh viên khá trả lời đúng 15 câu. Sinh viên trung bình trả lời đúng 10 câu. Sinh viên kém trả lời đúng 5 câu. Tỷ lệ sinh viên giỏi, khá, trung bình và kém lần lượt là 10%, 20%, 30%, 40%.

Một sinh viên lên bắt thăm 3 câu từ 20 câu trên. Giám khảo thấy anh trả lời đúng cả 3 câu. Tính xác suất anh ta là sinh viên khá hoặc trung bình.

Đáp số : 0.5184.

Bài 32. Có 2 lô hàng cũ. Lô I có 10 cái tốt, 2 cái hỏng. Lô II có 12 cái tốt, 3 cái hỏng. Từ mỗi lô lấy ngẫu nhiên ra 1 cái. Tìm xác suất để :

a) nhận được 2 cái tốt,

b) nhận được 2 cái cùng chất lượng,

c) nếu lấy từ cùng 1 lô ra 2 cái thì nên lấy từ lô nào để được 2 cái tốt với khả năng cao hơn.

Đáp số : a) 0.67.
b) 0.7.
c) Lấy từ lô I.

Bài 33. Có 3 hộp bi; hộp một có 10 bi trong đó có 3 bi đỏ; hộp hai có 15 bi trong đó có 4 bi đỏ; hộp ba có 12 bi trong đó có 5 bi đỏ. Gieo một con xúc xắc. Nếu xuất hiện mặt 1 thì chọn hộp một, xuất hiện mặt hai thì chọn hộp 2, xuất hiện các mặt còn lại thì chọn hộp ba. Từ hộp được chọn, lấy ngẫu nhiên 1 bi

a) tính xác suất để được bi đỏ,

b) giả sử lấy được bi đỏ. Tính xác suất để bi đỏ này thuộc hộp hai.

Đáp số : a) 0.372.
b) 0.1194.

Bài 34. Có 2 hộp áo; hộp một có 10 áo trong đó có 1 phế phẩm; hộp hai có 8 áo trong đó có 2 phế phẩm. Lấy hú họa 1 áo từ hộp một bỏ sang hộp hai; sau đó từ hộp này chọn hú họa ra 2 áo. Tìm xác suất để cả 2 áo này đều là phế phẩm.

Đáp số : 0.033.

Bài 35. Có 3 xạ thủ cùng bắn vào một con mồi, mỗi người bắn 1 viên đạn, với xác suất bắn trúng lần lượt là 0,6; 0,7; 0,8. Biết rằng nếu trúng 1 phát đạn thì xác suất để con thú bị tiêu diệt là 0,5; trúng 2 phát thì xác suất để con thú bị tiêu diệt là 0,8; còn nếu trúng 3 phát đạn thì chắc chắn con thú bị tiêu diệt.

a) Tính xác suất con thú bị tiêu diệt.

b) Hãy tính xác suất con thú bị tiêu diệt do trúng 2 phát đạn.

Đáp số : a) 0.7916.
b) 0.3616.

Bài 36. Có 2 chuồng thỏ. Chuồng thứ nhất có 5 con thỏ đen và 10 con thỏ trắng. Chuồng thứ hai có 3 con thỏ trắng và 7 con thỏ đen. Từ chuồng thứ hai, bắt ngẫu nhiên 1 con thỏ cho vào chuồng một và sau đó lại bắt ngẫu nhiên một con thỏ ở chuồng một ra thì được 1 con thỏ trắng. Tính xác suất để con thỏ trắng này là của chuồng một.

Đáp số : 0.973.

Bài 37. Một chuồng gà có 9 con gà mái và 1 con gà trống. Chuồng gà kia có 1 con mái và 5 con trống. Từ mỗi chuồng lấy ngẫu nhiên 1 con đem bán. Các con gà còn lại được dồn vào chuồng thứ ba. Nếu ta lại bắt ngẫu nhiên 1 con gà nữa từ chuồng này ra thì xác suất để bắt được con gà trống là bao nhiêu ?

Đáp số : 0.362.

Bài 38. Hai nhà máy cùng xản suất 1 loại linh kiện điện tử. Năng suất nhà máy hai gấp 3 lần năng suất nhà máy một. Tỷ lệ hỏng của nhà máy một và hai lần lượt là 0,1% và 0,2%. Giả sử linh kiện bán ở Trung tâm chỉ do hai nhà máy này sản xuất. Mua 1 linh kiện ở Trung tâm.

a) Tính xác suất để linh kiện ấy hỏng.

b) Giả sử mua linh kiện và thấy linh kiện bị hỏng. Theo ý bạn thì linh kiện đó do nhà máy nào sản xuất.

Đáp số : a) 0.00025.
b) 0.857*, linh kiện do nhà máy 2 sản xuất.*

Bài 39. Biết rằng $p_1 = 0,04$ là xác suất để mỗi sản phẩm được sản xuất ra từ dây chuyền 1 là phế phẩm. Tương tự, đối với dây chuyền 2 thì xác suất đó là $p_2 = 0,03$, với dây chuyền 3 là $p_3 = 0,05$ và với dây chuyền 4 là $p_4 = 0,058$. Từ một lô gồm 8 sản phẩm của dây chuyền 1; 12 sản phẩm của dây chuyền 2; 10 sản phẩm của dây chuyền 3 và 5 sản phẩm của dây chuyền 4, lấy ngẫu nhiên ra 1 sản phẩm. Tính xác suất để nhận được sản phẩm xấu ? nhận được sản phẩm tốt ?

Đáp số : 0.042, 0.958.

Bài 40. Trên mặt bàn có 5 đồng xu, trong đó có 3 đồng xu xấp và 2 đồng xu ngửa. Gieo tiếp lên mặt bàn 2 đồng xu và sau đó khoanh ngẫu nhiên 4 đồng xu. Tính xác suất để trong 4 đồng xu này có 3 đồng xu xấp.

Đáp số : 0.343.

Bài 41. Có 3 cái thùng. Thùng 1 có 6 bi trắng, 4 bi đỏ; thùng 2 có 5 bi trắng, 5 bi đỏ và thùng 3 có 10 bi trắng. Giả sử người ta lấy ngẫu nhiên 2 bi từ thùng 1 bỏ vào thùng 2. Sau đó, lại lấy ngẫu nhiên 1 bi từ thùng 2 bỏ vào thùng 3 rồi từ thùng 3 lấy ngẫu nhiên ra 1 bi. Tìm xác suất để bi lấy ra là đỏ.

Đáp số : 0.4833.

Công thức Bernoulli

Bài 42. Một bác sĩ chữa khỏi bệnh A cho một người với xác suất là 95%. Giả sử có 10 người bị bệnh A đến chữa một cách độc lập nhau. Tính xác suất để

a) có 8 người khỏi bệnh,

b) có nhiều nhất 9 người khỏi bệnh.

Đáp số : a) 0.0746.
b) 0.4013.

Bài 43. Một cầu thủ đá thành công quả phạt 11m với xác suất 80%.

- Đá 4 thành công 2.

- Đá 6 thành công 3.

Công việc nào dễ thực hiện ?

Đáp số : Đá 4 quả dễ hơn.

Bài 44. Trong một thành phố có 70% dân cư thích xem bóng đá. Chọn ngẫu nhiên 10 người, tính xác suất có :

a) 5 người thích xem bóng đá,

b) ít nhất 2 người thích xem bóng đá.

Đáp số : a) 0.103.
b) 0.999856.

Bài 45. Một nhà toán học có xác suất giải được một bài toán khó là 0,9. Cho nhà toán học này 5 bài toán khó được chọn một cách ngẫu nhiên.

a) Tính xác suất để nhà toán học này giải được 3 bài.

b) Tính xác suất để nhà toán học này giải được ít nhất 1 bài.

c) Tính số bài có khả năng nhất mà nhà toán học này giải được.

Đáp số : a) 0.0729.
b) 0.99999.
c) 5.

Bài 46. Tỷ lệ mắc bệnh Basedow ở một vùng rừng núi nào đó là 7%. Trong đợt khám tuyển sức khoẻ để xuất cảnh, người ta khám cho 100 người. Tìm xác suất để

a) trong 100 người có 6 người bị Basedow,

b) trong 100 người có 95 người không bị Basedow,

c) trong 100 người có ít nhất một người bị Basedow.

Đáp số : a) 0.153, *b)* 0.1283.
c) 0.999295.

Bài 47. Một lô hàng với tỷ lệ phế phẩm là 5%. Cần phải lấy mẫu cỡ bao nhiêu sao cho xác suất để bị ít nhất một phế phẩm không bé hơn 0,95.

Đáp số : Cỡ mẫu lớn hơn hay bằng 59.

Bài 48. Hai đấu thủ A, B thi đấu cờ. Xác suất thắng của người A trong một ván là 0,6 (không có hòa). Trận đấu bao gồm 5 ván, người nào thắng một số ván lớn hơn là người thắng cuộc. Tính xác suất để người B thắng cuộc.

Đáp số : 0.31744.

Bài 49. Một máy sản xuất lần lượt từng sản phẩm. Xác suất sản xuất ra một phế phẩm của máy là 0,01.

a) Cho máy sản xuất 10 sản phẩm. Tính xác suất để có 2 phế phẩm.

b) Máy cần sản xuất ít nhất bao nhiêu sản phẩm để xác suất có ít nhất một chính phẩm trên 0,99.

Đáp số : a) 0.00415.
b) Cần sản xuất ít nhất 459 sản phẩm.

Chương 2

BIẾN SỐ NGẪU NHIÊN

A. BÀI TẬP MẪU

Bài 1. Có hai thùng thuốc A và B, trong đó :

- thùng A có 20 lọ gồm 2 lọ hỏng và 18 lọ tốt,
- thùng B có 20 lọ gồm 3 lọ hỏng và 17 lọ tốt.

a) Lấy ở mỗi thùng 1 lọ. Gọi X là số lọ hỏng trong hai lọ lấy ra. Tìm hàm mật độ của X.

b) Lấy ở thùng B ra 3 lọ. Gọi Y là số lọ hỏng trong 3 lọ lấy ra. Tìm hàm mật độ của Y.

Giải

a) Xét các biến cố

A : "nhận được lọ hỏng từ thùng A",

B : "nhận được lọ hỏng từ thùng B",

và gọi X là số lọ hỏng trong hai lọ lấy ra. Ta có X lấy các giá trị 0, 1 và 2. Chú ý rằng A, B là các biến cố độc lập. Ta có

$$P(X=0) = P(\bar{A}\bar{B}) = P(\bar{A})P(\bar{B}) = \frac{18}{20}\cdot\frac{17}{20} = \frac{306}{400} = 0.765,$$

$$\begin{aligned} P(X=1) &= P(A\bar{B} + \bar{A}B) = P(A)P(\bar{B}) + P(\bar{A})P(B) \\ &= \frac{2}{20}\cdot\frac{17}{20} + \frac{18}{20}\cdot\frac{3}{20} = \frac{88}{400} = 0.22, \end{aligned}$$

$$P(X=2) = P(AB) = P(A)P(B) = \frac{2}{20}\cdot\frac{3}{20} = \frac{6}{400} = 0.015.$$

Từ đó, ta được bảng phân phối xác suất

X	0	1	2
P	0.765	0.22	0.015

và hàm mật độ của X

$$f(x) = \begin{cases} 0.765 & \text{khi } x = 0 \\ 0.22 & \text{khi } x = 1 \\ 0.015 & \text{khi } x = 2 \\ 0 & \text{khi } x \neq 0, 1, 2 \end{cases}$$

b) Gọi Y là số lọ hỏng trong 3 lọ lấy ra từ thùng B. Ta có $Y \sim H(20,3,3)$, nghĩa là

$$P(Y=k) = \frac{C_3^k C_{17}^{3-k}}{C_{20}^3}$$

và ta nhận được bảng phân phối xác suất

Y	0	1	2	3
P	0.596	0.358	0.045	0.001

cũng như hàm mật độ của Y

$$f(x)=\begin{cases}0.596 & \text{khi } x=0\\ 0.358 & \text{khi } x=1\\ 0.045 & \text{khi } x=2\\ 0.001 & \text{khi } x=3\\ 0 & \text{khi } x\neq 0,\ 1,\ 2,\ 3\end{cases}$$

Bài 2. Một xạ thủ bắn bia với xác suất bắn trúng bia là $p=0.6$. Có 5 viên đạn được bắn lần lượt và xạ thủ dừng bắn khi hết đạn hay ngay khi có một viên đạn trúng bia. Gọi X là số lần bắn. Tìm hàm mật độ của X. Tính trung bình μ và phương sai σ^2.

Giải

Xét các biến cố T_i : "bắn trúng bia ở lần bắn thứ i", với $i=1,2,3,4,5$. Gọi X số lần bắn, ta có $X=1,\ 2,\ 3,\ 4,\ 5$ và

$$P(X=1)=P(T_1)=0.6,$$

$$P(X=2)=P\left(\overline{T}_1T_2\right)=P\left(\overline{T}_1\right)P(T_2)=0.4\times 0.6,$$

$$P(X=3)=P\left(\overline{T}_1\overline{T}_2T_3\right)=P\left(\overline{T}_1\right)P\left(\overline{T}_2\right)P(T_3)$$
$$=(0.4)^2\times 0.6,$$

$$P(X=4)=P\left(\overline{T}_1\overline{T}_2\overline{T}_3T_4\right)=P\left(\overline{T}_1\right)P\left(\overline{T}_2\right)P\left(\overline{T}_3\right)P(T_4)$$
$$=(0.4)^3\times 0.6,$$

$$P(X=5)=P\left(\overline{T}_1\overline{T}_2\overline{T}_3\overline{T}_4\right)=P\left(\overline{T}_1\right)P\left(\overline{T}_2\right)P\left(\overline{T}_3\right)P\left(\overline{T}_4\right)$$
$$=(0.4)^4.$$

Từ đó, ta được bảng phân phối xác suất

X	1	2	3	4	5
P	0.6	0.24	0.096	0.0384	0.0256

và hàm mật độ xác suất của X

$$f(x)=\begin{cases}0.6 & \text{khi} & x=1\\ 0.24 & \text{khi} & x=2\\ 0.096 & \text{khi} & x=3\\ 0.0384 & \text{khi} & x=4\\ 0.0256 & \text{khi} & x=5\\ 0 & \text{khi} & x\neq 0,\ 1,\ 2,\ 3,4,5\end{cases}$$

Ta có trung bình của X

$$\mu_X=\sum_i x_i f(x_i)=1\times 0.6+2\times 0.24+\ldots+5\times 0.0256$$
$$=1.6496,$$

và phương sai là

$$\sigma_X^2=E\left(X^2\right)-\mu_X^2=\left(\sum_x x^2 f(x)\right)-\mu_X^2$$
$$=1^2\times 0.6+2^2\times 0.24+\ldots+5^2\times 0.0256-(1.6496)^2$$
$$=0.95722.$$

Bài 3. Một thùng đựng 10 lọ thuốc trong đó có 1 lọ hỏng. Ta kiểm tra từng lọ (không hoàn lại) cho tới khi phát hiện được lọ hỏng thì dừng. Gọi X là số lần kiểm tra. Tìm hàm mật độ của X. Tính trung bình μ và phương sai σ^2.

Giải

Xét các biến cố T_k : "lấy được lọ hỏng ở lần lấy thứ k", $k = 1, 2, ..., 10$. Gọi X là số lần kiểm tra. Ta có, $X = 1, 2, ..., 10$. Hơn nữa, gọi Y_k là biến cố "không lấy được lọ hỏng trong k lần lấy đầu tiên", với $k = 1, 2, ..., 10$. Ta được

$(X = k) = Y_{k-1}T_k$ và $Y_k = Y_{k-1}\overline{T}_k$.

$P(X = 1) = P(T_1) = \frac{1}{10}$; $P(Y_1) = P(\overline{T}_1) = \frac{9}{10}$;

$P(X = 2) = P(Y_1T_2) = P(T_2|Y_1)P(Y_1) = \frac{1}{9}.\frac{9}{10} = \frac{1}{10}$;

$P(Y_2) = P(Y_1\overline{T}_2) = P(\overline{T}_2|Y_1)P(Y_1) = \frac{8}{9}.\frac{9}{10} = \frac{8}{10}$;

$P(X = 3) = P(Y_2T_3) = P(T_3|Y_2)P(Y_2) = \frac{1}{8}.\frac{8}{10} = \frac{1}{10}$;

$P(Y_3) = P(Y_2\overline{T}_3) = P(\overline{T}_3|Y_2)P(Y_2) = \frac{7}{8}.\frac{8}{10} = \frac{7}{10}$;

$P(X = 4) = P(Y_3T_4) = P(T_4|Y_3)P(Y_3) = \frac{1}{7}.\frac{7}{10} = \frac{1}{10}$;

...

Tương tự, ta có $P(X = k) = \frac{1}{10}$, với mọi $k = 1, 2, ..., 10$.

Từ đó, ta được bảng phân phối xác suất

X	1	2	3	4	5	6	7	8	9	10
P	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

và hàm mật độ xác suất của X

$$f(x) = \begin{cases} \frac{1}{10} & \text{khi } x \in \{1, 2, 3, ..., 10\} \\ 0 & \text{khi } x \notin \{1, 2, 3, ..., 10\} \end{cases}$$

Suy ra trung bình và phương sai của X

$$\mu_X = (1 + 2 + .. + 10)\frac{1}{10} = 5.5.$$

$$\sigma_X^2 = E(X^2) - \mu_X^2 = (1^2 + 2^2 + .. + 10^2)\frac{1}{10} - (5.5)^2 = 8.25.$$

Bài 4. Gọi X là tuổi thọ của con người. Một công trình nghiên cứu cho biết hàm mật độ của X là

$$f(x) = \begin{cases} cx^2(100 - x)^2 & \text{khi } \ 0 \le x \le 100 \\ 0 & \text{khi } \ x < 0 \text{ hay } x > 100 \end{cases}$$

a) Xác định hằng số c.

b) Tính trung bình và phương sai của X.

c) Tính xác suất của một người có tuổi thọ ≥ 60.

d) Tính xác suất của một người có tuổi thọ ≥ 60, biết rằng người đó hiện nay đã 50 tuổi.

Giải

a) Để $f(x)$ là hàm mật độ, ta cần

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

mà

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{0}^{100} cx^2(100-x)^2 dx = c\left(10^4\frac{x^3}{3} - 2.10^2\frac{x^4}{4} + \frac{x^5}{5}\right)\Bigg|_0^{10^2},$$

nên ta được phương trình

$$c\left(10^4\frac{x^3}{3} - 2.10^2\frac{x^4}{4} + \frac{x^5}{5}\right)\Bigg|_0^{10^2} = 1.$$

Giải phương trình này, ta được $c = 3.10^{-9}$.

b) Ta có trung bình

$$\mu_X = E(X) = \int_{-\infty}^{+\infty} xf(x)dx = c\int_0^{100} x^3(100-x)^2 dx$$

$$= c\int_0^{100}(10^4x^3 - 2.10^2x^4 + x^5)dx$$

$$= c\left(10^4\frac{x^4}{4} - 2.10^2\frac{x^5}{5} + \frac{x^6}{6}\right)\Bigg|_0^{10^2} = 50,$$

và phương sai

$$\sigma_X^2 = E(X^2) - \mu_X^2 = \int_{-\infty}^{+\infty} x^2f(x)dx - 50^2 = c\int_0^{100} x^4(100-x)^2 dx - 2500$$

$$= c\int_0^{100}(10^4x^4 - 2.10^2x^5 + x^6)dx - 2500$$

$$= c\left(10^4\frac{x^5}{5} - 2.10^2\frac{x^6}{6} + \frac{x^7}{7}\right)\Bigg|_0^{10^2} - 2500$$

$$= 3.10^{-9}\left(\frac{10^{14}}{105}\right) - 2500 = \frac{10^5}{35} - 2500 = \frac{2500}{7}.$$

c) Xác suất của một người có tuổi thọ ≥ 60 là

$$P(X \ge 60) = \int_{60}^{+\infty} f(x)dx = \int_{60}^{100} cx^2(100-x)^2\,dx$$

$$= c\int_{60}^{100} (10^4x^2 - 2.10^2x^3 + x^4)dx$$

$$= c\left(10^4\frac{x^3}{3} - 2.10^2\frac{x^4}{4} + \frac{x^5}{5}\right)\Bigg|_{60}^{10^2}$$

$$= c\left[\left(\frac{10^{10}}{30}\right) - 10^5\left(100.\frac{216}{3} - 20.\frac{1296}{4} + \frac{7776}{5}\right)\right]$$

$$= 3.10^{-9}10^5\left(\frac{10^4}{3} - \frac{11376}{5}\right) = \frac{992}{3125} = 0.31744.$$

d) Để tính xác suất của một người có tuổi thọ ≥ 60, khi biết người đó đã 50 tuổi, ta tính xác suất có điều kiện

$$P\left(X \ge 60 | X \ge 50\right) = \frac{P\left((X \ge 60)(X \ge 50)\right)}{P(X \ge 50)}$$

$$= \frac{P(X \ge 60)}{P(X \ge 50)} = \frac{0.31744}{0.5} = 0.63548,$$

với $P(X \ge 50)$ được tính như ở phần c và bằng 0.5.

Bài 5. Cho biến số ngẫu nhiên X có hàm mật độ

$$f(x) = \begin{cases} \frac{1}{\lambda}e^{-\frac{x}{\lambda}} & \text{khi} \quad x > 0 \\ 0 & \text{khi} \quad x \le 0 \end{cases} \qquad \text{với } \lambda > 0$$

a) Tính trung bình μ phương sai σ^2.

b) Tìm hàm đặc trưng M(t). Dùng hàm đặc trưng, tính lại trung bình μ và phương sai σ^2.

Giải

a) Ta có

$$\mu = E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \frac{1}{\lambda}\int_0^{+\infty} xe^{-\frac{x}{\lambda}}dx, \qquad (1)$$

Dùng công thức tích phân từng phần, với $u = x$, $dv = e^{-x/\lambda}dx$, ta được $du = dx$, $v = -\lambda e^{-x/\lambda}$ và biểu thức (1) cho

$$\mu = -\lambda xe^{-\frac{x}{\lambda}}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-\frac{x}{\lambda}}dx$$

$$= -\lambda e^{-\frac{x}{\lambda}}\Big|_0^{+\infty} = \lambda$$

Phương sai σ^2 cho bởi

$$\sigma^2 = E\left(X^2\right) - \mu_X^2,$$

với $E\left(X^2\right)=\int_{-\infty}^{+\infty} x^2 f(x)dx=\frac{1}{\lambda}\int_{0}^{+\infty} x^2 e^{-\frac{x}{\lambda}}dx$.

Cũng do công thức tích phân từng phần, ta có

$$E\left(X^2\right)=-x^2e^{-\frac{x}{\lambda}}\Big|_0^{+\infty}+2\int_0^{+\infty} xe^{-\frac{x}{\lambda}}dx=2\lambda\cdot\frac{1}{\lambda}\int_0^{+\infty} xe^{-\frac{x}{\lambda}}dx$$
$$=2\lambda^2.$$

Từ đó suy ra $\sigma^2=E\left(X^2\right)-\mu_X^2=2\lambda^2-\lambda^2=\lambda^2$.

b) Hàm đặc trưng $M(t)$ của biến số ngẫu nhiên X cho bởi

$$M(t)=E\left(e^{tX}\right)=\int_{-\infty}^{+\infty} e^{tx}f(x)dx=\frac{1}{\lambda}\int_0^{+\infty} e^{\left(t-\frac{1}{\lambda}\right)x}dx$$
$$=\frac{1}{\lambda t-1}e^{\left(t-\frac{1}{\lambda}\right)x}\Bigg|_0^{+\infty}=\frac{1}{1-\lambda t}.$$

Với hàm đặc trưng $M(t)$ này, ta nhận được trở lại giá trị trung bình

$$\mu=M'(0)=\frac{\lambda}{\left[1-\lambda\times(0)\right]^2}=\lambda,$$

và phương sai

$$\sigma^2=M''(0)-\left[M'(0)\right]^2=\frac{2\lambda^2}{\left[1-\lambda(0)\right]^3}-\left(\frac{\lambda}{\left[1-\lambda\times(0)\right]^2}\right)^2$$
$$=2\lambda^2-\lambda^2=\lambda^2.$$

Bài 6. Cho vectơ ngẫu nhiên có bảng phân phối xác suất

Y / X	1	2	3
0	0.1	0.2	0.1
1	0.2	0.2	0.2

a) Tìm các hàm mật độ thành phần $f_X(x), f_Y(y)$.

b) Tìm các trung bình μ_X, μ_Y, các phương sai σ_X^2, σ_Y^2 và hệ số tương quan $\rho(X,Y)$.

Giải

a) Hàm mật độ thành phần $f_X(x)$ cho bởi

$$f_X(0)=P(X=0)=P\left(X=0|Y=1\right)+P\left(X=0|Y=2\right)+P\left(X=0|Y=3\right)$$
$$=0.1+0.2+0.1=0.4,$$
$$f_X(1)=P(X=1)=P\left(X=1|Y=1\right)+P\left(X=1|Y=2\right)+P\left(X=1|Y=3\right)$$
$$=0.2+0.2+0.2=0.6,$$

và $f_X(x)=0$ khi $x\neq 0,1$.

Tương tự, hàm mật độ thành phần $f_Y(y)$ cho bởi

$$f_Y(1)=P(Y=1)=P\left(Y=1|X=0\right)+P\left(Y=1|X=1\right)$$
$$=0.1+0.2=0.3,$$

$$f_Y(2) = P(Y=2) = P\left(Y=2 \middle| X=0\right) + P\left(Y=2 \middle| X=1\right)$$
$$= 0.2 + 0.2 = 0.4,$$

$$f_Y(3) = P(Y=3) = P\left(Y=3 \middle| X=0\right) + P\left(Y=3 \middle| X=1\right)$$
$$= 0.1 + 0.2 = 0.3.$$

và $f_Y(y) = 0$, với $y \neq 1, 2, 3$.

b) Từ các hàm mật độ, ta suy ra

$$\mu_X = \sum_x x f_X(x) = 0 \cdot 0.4 + 1 \cdot 0.6 = 0.6,$$

$$\mu_Y = \sum_y y f_Y(y) = 1 \cdot 0.3 + 2 \cdot 0.4 + 3 \cdot 0.3 = 2,$$

$$\sigma_X^2 = \sum_x x^2 f_X(x) - \mu_X^2 = 0^2 \cdot 0.4 + 1^2 \cdot 0.4 - (0.4)^2 = 0.24,$$

$$\sigma_Y^2 = \sum_y y^2 f_y(y) - \mu_y^2 = 1^2 \cdot 0.3 + 2^2 \cdot 0.4 + 3^2 \cdot 0.3 - 2^2 = 0.6.$$

Do

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X.\sigma_Y} = \frac{E(XY) - \mu_X.\mu_Y}{\sigma_X.\sigma_Y},$$

và

$$E(XY) = \sum_{x,y} xyf(x,y) = 0.2 + 0.4 + 0.6 = 1.2,$$

ta suy ra

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X.\sigma_Y} = \frac{E(XY) - \mu_X.\mu_Y}{\sigma_X.\sigma_Y} = \frac{1.2 - 0.6 \times 2}{\sqrt{0.24 \times 0.6}} = 0.$$

Bài 7. Cho vectơ ngẫu nhiên có hàm mật độ

$$f(x,y) = \begin{cases} c(x+y)^2 & \text{khi} \quad (x,y) \in [0,1] \times [0,1] \\ 0 & \text{khi} \quad (x.y) \notin [0,1] \times [0,1] \end{cases}$$

a) Tìm các hàm mật độ thành phần $f_X(x), f_Y(y)$.

b) Tìm các trung bình μ_X, μ_Y, các phương sai σ_X^2, σ_Y^2 và hệ số tương quan $\rho(X,Y)$.

Giải

Trước hết, ta cần xác định hằng số c. Do tính chất hàm mật độ, ta có

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1.$$

Mà

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = c\int_0^1\int_0^1 (x+y)^2 dxdy = \frac{c}{3}\left(x^3 + \frac{3}{2}x^2 + x\right)\Bigg|_0^1 = \frac{7c}{6},$$

nên ta suy ra $c = \frac{6}{7}$. Khi đó, các hàm mật độ thành phần cho bởi

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^1 c(x+y)^2dy=\left(\frac{c}{3}(x+y)^3\right)\Big|_0^1$$
$$=\frac{2}{7}\left(3x^2+3x+1\right),$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^1 c(x+y)^2dx=\left(\frac{c}{3}(x+y)^3\right)\Big|_0^1$$
$$=\frac{2}{7}\left(3y^2+3y+1\right).$$

b) Từ hàm mật độ $f(x,y)$, ta suy ra

$$\mu_X=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)dxdy=\frac{6}{7}\int_0^1\int_0^1 x(x+y)^2dxdy=\frac{2}{7}\int_0^1\left(x(x+y)^3\right)\Big|_0^1 dx$$
$$=\frac{2}{7}\int_0^1\left(3x^3+3x^2+x\right)dx=\frac{2}{7}\left(\frac{3}{4}x^4+x^3+\frac{x^2}{2}\right)\Big|_0^1=\frac{9}{14},$$

$$\mu_Y=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)dxdy=\frac{6}{7}\int_0^1\int_0^1 y(x+y)^2dxdy$$
$$=\frac{2}{7}\int_0^1\left(y(x+y)^3\right)\Big|_0^1 dy=\frac{2}{7}\int_0^1\left(3y^3+3y^2+y\right)dy$$
$$=\frac{2}{7}\left(\frac{3}{4}y^4+y^3+\frac{y^2}{2}\right)\Big|_0^1=\frac{9}{14},$$

$$\sigma_X^2=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_X)^2 f(x,y)dxdy$$
$$=\frac{1}{686}\left(\frac{588}{5}x^5-\frac{168}{4}x^4-\frac{317}{3}x^3-\frac{9}{2}x^2+81x\right)\Big|_0^1=\frac{199}{2940},$$

$$\sigma_Y^2=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-\mu_Y)^2 f(x,y)dxdy=\frac{199}{2940},$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy=\frac{6}{7}\int_0^1\int_0^1 xy(x+y)^2dxdy$$
$$=\frac{6}{7}\int_0^1\left(x^3\frac{1}{2}+2x^2\frac{1}{3}+x\frac{1}{4}\right)dx$$
$$=\frac{6}{7}\left(\frac{x^4}{8}+\frac{2x^3}{9}+\frac{x^2}{8}\right)\Big|_0^1=\frac{17}{42},$$

$$\rho(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}=\frac{E(XY)-\mu_X\mu_Y}{\sigma_X\sigma_Y}$$
$$=\frac{\frac{17}{42}-\frac{9}{14}\cdot\frac{9}{14}}{\sqrt{\frac{199}{2940}\cdot\frac{199}{2940}}}=\frac{-\frac{5}{588}}{\frac{199}{2940}}=-\frac{25}{199}=-0.127.$$

Bài 8. Cho vectơ ngẫu nhiên $V=(X,Y)$, với X, Y độc lập. Giả sử X, Y có trung bình μ_X, μ_Y và phương sai σ_X^2, σ_Y^2
Đặt $Z=\alpha X+\beta Y$. Chứng minh rằng

a) $\mu_Z=\alpha\mu_X+\beta\mu_Y$,

b) $\sigma_Z^2=\alpha^2\sigma_X^2+\beta^2\sigma_Y^2$.

Giải

Ta chứng minh cho trường hợp X, Y là các biến số ngẫu nhiên rời rạc. Trường hợp X và Y là các biến số ngẫu nhiên liên tục được chứng minh tương tự.

a) Gọi $f(x,y)$ là hàm mật độ (đồng thời) của V. Ta có

$$\begin{aligned}E(Z)=E(\alpha X+\beta Y)&=\sum_{x,y}(\alpha x+\beta y)f(x,y)\\&=\alpha\sum_{x,y}xf(x,y)+\beta\sum_{x,y}yf(x,y)=E(X)+E(Y),\end{aligned}$$

nghĩa là $\mu_Z=\mu_X+\mu_Y$.

b) Do định nghĩa,

$$\begin{aligned}\sigma_Z^2&=\sum_z(z-\mu_Z)^2f(x,y)=\sum_{x,y}\left((\alpha x+\beta y)-(\alpha\mu_X+\beta\mu_Y)\right)^2f(x,y)\\&=\sum_{x,y}\left[\alpha(x-\mu_X)+\beta(y-\mu_Y)\right]^2f(x,y)\\&=\alpha^2\sum_{x,y}(x-\mu_X)^2f(x,y)+\beta^2\sum_{x,y}(y-\mu_Y)^2f(x,y)+\\&\qquad+2\alpha\beta\sum_{x,y}(x-\mu_X)(y-\mu_Y)f(x,y)).\end{aligned}$$

Mà X và Y độc lập nên

$$\sum_{x,y}(x-\mu_X)(y-\mu_Y)f(x,y))=E\left((X-\mu_X)(Y-\mu_Y)\right)=0,$$

và do đó

$$\begin{aligned}\sigma_Z^2&=\alpha^2\sum_{x,y}(x-\mu_X)^2f(x,y)+\beta^2\sum_{x,y}(y-\mu_Y)^2f(x,y)\\&=\alpha^2\sigma_X^2+\beta^2\sigma_Y^2.\end{aligned}$$

Bài 9. Cho vectơ ngẫu nhiên $V=(X,Y)$. Đặt $Z=X+Y$. Chứng minh rằng

$$\mu_Z=\mu_X+\mu_Y \text{ và } \sigma_Z^2=\sigma_X^2+2\rho(X,Y)\sigma_X\sigma_Y+\sigma_Y^2.$$

Suy ra rằng, nếu X và Y không tương quan, nghĩa là $\rho(X,Y)=0$, thì $\sigma_Z^2=\sigma_X^2+\sigma_Y^2$.

Giải

Tương tự bài 8, ta chứng minh cho trường hợp X, Y là biến số ngẫu nhiên rời rạc.

$$\mu_Z=\sum_z zf(x,y)=\sum_{x,y}xf(x,y)+\sum_{x,y}yf(x,y)=\mu_x+\mu_y,$$

$$\sigma_Z^2 = \sum_z (z-\mu_Z)^2 f(x,y) = \sum_z (x-\mu_x+y-\mu_y)^2 f(x,y)$$
$$= \sum_{x,y}(x-\mu_X)^2 f(x,y) + \sum_{x,y}(y-\mu_y)^2 f(x,y) +$$
$$+2\sum_{x,y}(x-\mu_X)(y-\mu_y) f(x,y)$$
$$= \sigma_X^2+\sigma_Y^2+2\mathrm{cov}(X,Y) = \sigma_X^2+\sigma_Y^2+2\rho(X,Y)\sigma_X\sigma_Y.$$

Khi X và Y không tương quan, thì $\rho(X,Y)=0$ và do đó

$\sigma_Z^2 = \sigma_X^2+\sigma_Y^2$.

Bài 10. Cho vectơ ngẫu nhiên $V=(X,Y)$ có bảng phân phối xác suất

X \ Y	0	1
–1	1/3	0
0	0	1/3
1	1/3	0

a) Tính trung bình và phương sai của X và Y.

b) Tính hệ số tương tương quan $\rho(X,Y)$.

c) X và Y có độc lập không ?

Giải

a) Ta có các hàm mật độ thành phần

$$f_X(x)=\begin{cases}1/3 & \text{khi} \quad x=-1,0,1\\ 0 & \text{khi} \quad x\neq -1,0,1\end{cases}$$

$$f_Y(y)=\begin{cases}2/3 & \text{khi} \quad y=0\\ 1/3 & \text{khi} \quad y=1\\ 0 & \text{khi} \quad y\neq 0,1\end{cases}$$

Từ đó suy ra

$$\mu_X = \sum_x x f_X(x) = -1\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}=0,$$

$$\mu_Y = \sum_y y f_Y(y) = 0\cdot\frac{2}{3}+1\cdot\frac{1}{3}=\frac{1}{3},$$

$$\sigma_X^2 = \sum_x x^2 f_X(x)-\mu_X^2 = (-1)^2\cdot\frac{1}{3}+0^2\cdot\frac{1}{3}+1^2\cdot\frac{1}{3}-0^2=\frac{2}{3},$$

$$\sigma_Y^2 = \sum_y y^2 f_y(y)-\mu_y^2 = 0^2\cdot\frac{2}{3}+1^2\cdot\frac{1}{3}-\left(\frac{1}{3}\right)^2=\frac{2}{9}.$$

b) Do

$$E(XY)=\sum_{x,y} xyf(x,y)=0,$$

ta suy ra

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X.\sigma_Y} = \frac{E(XY) - \mu_X.\mu_Y}{\sigma_X.\sigma_Y} = 0.$$

c) Với hàm mật độ đồng thời $f(x,y)$, ta có

$$f(0,0) = 0 \neq f_X(0)f_Y(0) = \frac{2}{9}.$$

Do đó, X và Y không độc lập.

Bài 11. Chứng minh rằng nếu vectơ ngẫu nhiên $V = (X,Y)$ có X, Y độc lập, thì $\rho(X,Y) = 0$.

Giải

Vì X, Y độc lập nên $E(XY) = E(X).E(Y)$. Từ đó suy ra

$$\begin{aligned}\rho(X,Y) &= \frac{\operatorname{cov}(X,Y)}{\sigma_X.\sigma_Y} = \frac{E(XY) - \mu_X.\mu_Y}{\sigma_X.\sigma_Y} \\ &= \frac{E(X).E(Y) - E(X).E(Y)}{\sigma_X.\sigma_Y} = 0.\end{aligned}$$

Bài 12. Chứng minh rằng với mọi vectơ ngẫu nhiên $V = (X,Y)$, ta có hệ số tương quan $\rho(X,Y)$ thỏa $-1 \leq \rho(X,Y) \leq 1$.

Giải

Ta chứng minh cho trường hợp X, Y là các biến số ngẫu nhiên rời rạc. Trường hợp biến số ngẫu nhiên liên tục được chứng minh tương tự. Với $f(x,y)$ chỉ hàm mật độ (đồng thời) của $V = (X,Y)$, ta có

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X.\sigma_Y} = \frac{E\big((X-\mu_X)(Y-\mu_Y)\big)}{\sigma_X.\sigma_Y},$$

và

$$\begin{aligned}E\big((X-\mu_X)(Y-\mu_Y)\big) &= \sum_{x,y}(X-\mu_X)(Y-\mu_Y)f(x,y) \\ &= \sum_{x,y}(X-\mu_X)\sqrt{f(x,y)}(Y-\mu_Y)\sqrt{f(x,y)}.\end{aligned}$$

nên từ bất đẳng thức Cauchy Schwarz, ta có

$$\begin{aligned}E\big((X-\mu_X)(Y-\mu_Y)\big) &\leq \sqrt{\sum_{x,y}(X-\mu_X)^2 f(x,y)}\sqrt{\sum_{x,y}(Y-\mu_Y)^2 f(x,y)} \\ &= \sigma_X\sigma_Y.\end{aligned}$$

Do đó

$$|\rho(X,Y)| = \frac{\left|E\big((X-\mu_X)(Y-\mu_Y)\big)\right|}{\sigma_X\sigma_Y} \leq 1,$$

nghĩa là $-1 \leq \rho(X,Y) \leq 1$.

B. BÀI TẬP

Xác định biến ngẫu nhiên.

Bài 1. Xác suất chữa khỏi bệnh A của 1 bác sĩ là 0,8.

a) Lập bảng phân phối xác suất của số người được chữa khỏi bệnh trong 1 nhóm bệnh nhân gồm 5 người do bác sĩ đó điều trị.

b) Gọi X là số bệnh nhân chữa khỏi bệnh. Tìm hàm phân phối xác suất của X.

Đáp số : a)

X	*0*	*1*	*2*	*3*	*4*	*5*
P	*0.00032*	*0.0064*	*0.0512*	*0.2048*	*0.4096*	*0.32768*

b)

$$F(x) = \begin{cases} 0 & \text{khi} \quad x < 0 \\ 0.00032 & \text{khi} \quad 0 \le x < 1 \\ 0.00672 & \text{khi} \quad 1 \le x < 2 \\ 0.05792 & \text{khi} \quad 2 \le x < 3 \\ 0.26272 & \text{khi} \quad 3 \le x < 4 \\ 0.67232 & \text{khi} \quad 4 \le x < 5 \\ 1 & \text{khi} \quad x \ge 5 \end{cases}$$

Bài 2. Có 2 cái hộp. Hộp một chứa 10 bi gồm 3 bi đỏ và 7 bi đen. Hộp hai chứa 5 bi gồm 2 bi đỏ và 3 bi đen. Lấy ngẫu nhiên 1 bi từ hộp một bỏ vào hộp hai; rồi từ hộp hai lấy ngẫu nhiên 1 bi.

a) Tính xác suất để bi lấy ra từ hộp hai là bi đỏ.

b) Lập bảng phân phối xác suất cho số bi đỏ có trong hộp hai sau khi bỏ vào 1 bi lấy từ hộp một.

Đáp số : a) 0.383.

b)

X	*2*	*3*
P	$\frac{7}{10}$	$\frac{3}{10}$

Bài 3. Một thiết bị gồm 3 bộ phận hoạt động độc lập với nhau, xác suất trong khoảng thời gian t các bộ phận hỏng tương ứng bằng 0.2; 0.3; 0.25. Gọi X là số bộ phận bị hỏng trong khoảng thời gian t.

a) Lập bảng phân phối xác suất của X.

b) Viết biểu thức hàm phân phối của X.

c) Tính $P(0 < X \le 4)$ theo hai cách.

Đáp số : a)

X	*0*	*1*	*2*	*3*
P	*0.42*	*0.425*	*0.14*	*0.015*

b)

$$F(x) = \begin{cases} 0 & \text{khi} \quad x < 0 \\ 0.42 & \text{khi} \quad 0 \le x < 1 \\ 0.845 & \text{khi} \quad 1 \le x < 2 \\ 0.985 & \text{khi} \quad 2 \le x < 3 \\ 1 & \text{khi} \quad x \ge 3 \end{cases}$$

c) 0.58.

Bài 4. Mỗi cầu thủ có 3 quả bóng. Hai cầu thủ lần lượt ném bóng vào rổ cho đến khi có người ném trúng hoặc hết bóng thì ngưng. Biết xác suất ném trúng của cầu thủ thứ nhất là 0,7, của cầu thủ thứ hai là 0,8 và cầu thủ 1 ném trước.

a) Gọi X_i là số lần cầu thủ thứ i ném. Lập bảng phân phối xác suất của X_1 và X_2.

b) Gọi Y_i là số lần cầu thủ thứ i ném trúng. Lập bảng phân phối xác suất của Y_1 và Y_2.

Đáp số : a)

X_1	1	2	3
P	0.94	0.0564	0.0036

X_2	0	1	2	3
P	0.7	0.282	0.01692	0.00108

b)

Y_1	0	1
P	0.25548	0.74452

Y_2	0	1
P	0.744736	0.255264

Tham số đặc trưng của biến ngẫu nhiên.

Bài 5. Tung một đồng xu xấp ngửa 2 lần độc lập. Gọi X là số lần được mặt xấp.

a) Lập bảng phân phối xác suất cho X.

b) Tính xác suất có ít nhất một lần được mặt xấp.

c) Tính kỳ vọng, phương sai.

d) Tính Mod[X], Me[X].

e) Tính hệ số bất đối xứng, hệ số nhọn.

Đáp số : a)

X	0	1	2
P	0.25	0.5	0.25

b) 0.75.

c) $\mu_X = 1$, $\sigma_X^2 = 0.5$.

d) $\text{Mod}[X] = 1$, $\text{Me}[X] = 1$.

e) $\gamma_1(X) = 0$, $\gamma_2(X) = 8$.

Bài 6. Gọi X là số lần mặt nhất xuất hiện sau ba lần tung một con xúc xắc.

a) Lập bảng phân phối xác suất của X.

b) Tính xác suất có ít nhất một lần được mặt nhất.

c) Tính xác suất có tối đa hai lần mặt nhất.

d) Tính μ_X, σ_X^2.

Đáp số : a)

X	0	1	2	3
P	0.579	0.347	0.069	0.005

b) 0.421.

c) 0.995.

d) $\mu_X = 0.5$, $\sigma_X^2 = 0.417$.

Bài 7. Có 3 xạ thủ cùng bắn vào một mục tiêu, mỗi người bắn 1 viên, trong cùng một số điều kiện nhất định. Xác suất để mỗi xạ thủ bắn trúng mục tiêu lần lượt là 0,6; 0,7; 0,9. Gọi X là số viên đạn trúng mục tiêu. Hãy lập bảng phân phối xác suất của X. Tính trung bình (μ_X), phương sai (σ_X^2) và Mod[X].

Đáp số :

X	*0*	*1*	*2*	*3*
P	*0.012*	*0.154*	*0.456*	*0.378*

$\mu_X = 2.2$, $\sigma_X^2 = 0.54$, $\text{Mod}[X] = 2$.

Bài 8. Một phân xưởng có ba máy M_1, M_2, M_3. Trong một giờ, mỗi máy sản xuất được 10 sản phẩm, trong đó số sản phẩm không đạt tiêu chuẩn của M_1, M_2, M_3 lần lượt là 1, 2, 1. Lấy ngẫu nhiên từ mỗi máy một sản phẩm. Gọi X là số sản phẩm không đạt tiêu chuẩn trong 3 sản phẩm được lấy ra.

a) Lập bảng phân phối sản xuất của X.

b) Tìm μ_X, σ_X^2, Mod[X].

c) Tính $P(X \le 1)$.

Đáp số : a)

X	*0*	*1*	*2*	*3*
P	*0.648*	*0.306*	*0.044*	*0.002*

b) $\mu_X = 0.4$, $\sigma_X^2 = 0.34$, $\text{Mod}[X] = 0$.

c) 0.954.

Bài 9. Xét trò chơi, tung một con xúc xắc ba lần: nếu cả ba lần được 6 nút thì lĩnh 6 ngàn đ, nếu hai lần 6 nút thì lĩnh 4 ngàn đ, một lần 6 nút thì lĩnh 2 ngàn đ, và nếu không có 6 nút thì không lĩnh gì hết. Mỗi lần chơi phải đóng A ngàn đ. Hỏi :

a) A là bao nhiêu thì người chơi về lâu về dài huề vốn (gọi là trò chơi công bằng),

b) A là bao nhiêu thì trung bình mỗi lần người chơi mất 1 ngàn đ.

Đáp số : a) $A = 1000$.

b) $A = 2000$.

Bài 10. Một nhà đầu tư có 3 dự án. Gọi X_i(i=1, 2, 3) là số tiền thu được khi thực hiện dự án thứ i (giá trị âm chỉ số tiền bị thua lỗ). X_i là đại lượng ngẫu nhiên. Qua nghiên cứu, giả sử có số liệu như sau : (Đơn vị tính : 10 triệu đồng)

X_1	-20	30	60
P	0.3	0.2	0.5

X_2	-20	-10	100
P	0.4	0.2	0.4

X_3	-25	-30	80
P	0.2	0.3	0.5

Theo anh (chị), ta nên chọn dự án nào ?

Đáp số : Nên chọn dự án 1.

Bài 11. Cho X là một đại lượng ngẫu nhiên có phân phối xác suất như sau

X	0	1	2	3	4	5	6	7
P_X	0	a	2a	2a	3a	a^2	$2a^2$	$7a^2 + a$

a) Xác định a.

b) Tính $P[X \geq 5]$, $P[X < 3]$.

c) Tính k nhỏ nhất sao cho $P[X \leq k] \geq \frac{1}{2}$.

Đáp số : a) $a = \frac{1}{10}$.

b) $P[X \geq 5] = 0.2$, $P[X < 3] = 0.3$.

c) $k = 3$.

Bài 12. Cho hàm mật độ của biến ngẫu nhiên X có dạng

a) $f(x) = \begin{cases} Ax & \text{khi} \quad x \in [0,1] \\ 0 & \text{khi} \quad x \notin [0,1] \end{cases}$

b) $f(x) = \begin{cases} A\sin x & \text{khi} \quad x \in [0,\pi] \\ 0 & \text{khi} \quad x \notin [0,\pi] \end{cases}$

c) $f(x) = \begin{cases} A\cos \pi x & \text{khi} \quad x \in [0,\frac{1}{2}] \\ 0 & \text{khi} \quad x \notin [0,\frac{1}{2}] \end{cases}$

d) $f(x) = \begin{cases} A\frac{1}{x^4} & \text{khi} \quad x \geq 1 \\ 0 & \text{khi} \quad x < 1 \end{cases}$

Hãy xác định A. Tìm hàm phân phối xác suất của X. Tính μ_X, σ_X^2, nếu có.

Đáp số : a) $A = 2$, $\mu_X = \frac{2}{3}$, $\sigma_X^2 = 0.055$,

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

b) $A = 0.5$, $\mu_X = \frac{\pi}{2}$, $\sigma_X^2 = \frac{\pi^2}{4} - 2$,

$$F(x) = \begin{cases} \frac{1}{2}(1-\cos x) & \text{khi} \quad 0 \leq x \leq \pi \\ 0 & \text{khi} \quad x < 0 \\ 1 & \text{khi} \quad x > \pi \end{cases}.$$

c) $A = \pi$, $\mu_X = \frac{1}{2} - \frac{1}{\pi}$, $\sigma_X^2 = \frac{\pi - 3}{\pi^2}$,

$$F(x) = \begin{cases} \sin(\pi x) & \text{khi} \quad 0 \leq x \leq \frac{1}{2} \\ 0 & \text{khi} \quad x < 0 \\ 1 & \text{khi} \quad x > \frac{1}{2} \end{cases}.$$

d) $A = 3,\ \mu_X = \dfrac{3}{2},\ \sigma_X^2 = \dfrac{3}{4},$

$$F(x) = \begin{cases} 1 - \dfrac{1}{x^3} & \text{khi} \quad x \ge 1 \\ 0 & \text{khi} \quad x < 1 \end{cases}.$$

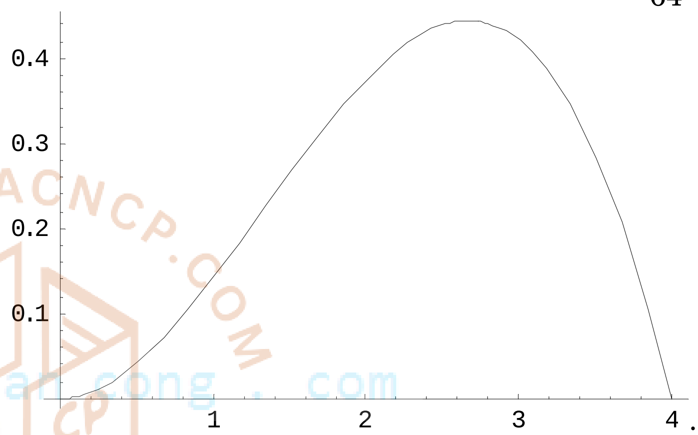
Bài 13. Tuổi thọ của một loại bóng đèn nào đó là 1 biến ngẫu nhiên X (đơn vị năm) với hàm mật độ như sau

$$f(x) = \begin{cases} kx^2(4-x) & \text{khi} \quad 0 \le x \le 4 \\ 0 & \text{khi} \quad x \notin [0,4] \end{cases}$$

a) Tìm k và vẽ đồ thị f(x).

b) Tìm xác suất để bóng đèn cháy trước khi nó được 1 năm tuổi.

Đáp số : a) $k = \dfrac{3}{64}$,



b) 0.0508.

Bài 14. Trọng lượng của một con vịt 6 tháng tuổi là 1 biến ngẫu nhiên X (đơn vị tính là Kg) có hàm mật độ

$$f(x) = \begin{cases} k(x^2-1) & \text{khi} \quad 1 \le x \le 3 \\ 0 & \text{khi} \quad x \notin [1,3] \end{cases}$$

a) Tìm k.

b) Với k tìm được, tìm

(i) trọng lượng trung bình của vịt 6 tháng tuổi,

(ii) hàm phân phối xác suất của X,

(iii) tỷ lệ vịt chậm lớn, biết vịt 6 tháng tuổi chậm lớn là vịt có trọng lượng nhỏ hơn 2Kg.

Đáp số : a) $k = \dfrac{3}{20}$.

b) (i) $\mu_X = 2.4\,kg$.

(ii)

$$F(x) = \begin{cases} \dfrac{x^3 - 3x + 2}{20} & \text{khi} \quad 1 \le x \le 3 \\ 0 & \text{khi} \quad x < 1 \\ 1 & \text{khi} \quad x > 3 \end{cases}.$$

(iii) 0.2.

Bài 15. Cho hàm mật độ của biến ngẫu nhiên X có dạng

$$f(x) = \begin{cases} a\cos x & \text{khi} \quad x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ 0 & \text{khi} \quad x \notin \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \end{cases}$$

a) Tìm a và xác định hàm phân phối xác suất F(x) của X.

b) Tính xác suất để X nhận giá trị trong khoảng $\left(\frac{\pi}{4}, \pi\right)$.

Đáp số : a) $a = \frac{1}{2}$,

$$F(x) = \begin{cases} \frac{\sin x + 1}{2} & \text{khi} \quad -\frac{\pi}{2} \le x \le \frac{\pi}{2} \\ 0 & \text{khi} \quad x < -\frac{\pi}{2} \\ 1 & \text{khi} \quad x > \frac{\pi}{2} \end{cases}.$$

b) 0.1465.

Bài 16. Cho biến ngẫu nhiên liên tục X có hàm phân phối

$$F(x) = \begin{cases} 0 & \text{khi} \quad x < -\frac{\pi}{2}, \\ a + b\sin x & \text{khi} \quad -\frac{\pi}{2} \le x \le \frac{\pi}{2}, \\ 1 & \text{khi} \quad x > \frac{\pi}{2} \end{cases}$$

với a, b là hằng số.

a) Tìm a và b.

b) Với a và b tìm được ở câu a), tính hàm mật độ f(x) của X; $\text{Mod}[x]$; $\text{Me}[x]$; $P\left[X > \frac{\pi}{4}\right]$.

Đáp số : a) $a = \frac{1}{2}$, $b = \frac{1}{2}$.

b) $\text{Mod}[x] = 0$, $\text{Me}[x] = 0$, $P\left[X > \frac{\pi}{4}\right] = 0.1465$,

$$f(x) = \begin{cases} \frac{1}{2}\cos x & \text{khi} \quad x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ 0 & \text{khi} \quad x \notin \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \end{cases}$$

Vectơ ngẫu nhiên.

Bài 17. Số trẻ em sinh ra trong một tuần ở một làng A nào đó là một đại lượng ngẫu nhiên có phân bố xác suất là

X	0	1	2	3
P	0,4	0,3	0,2	0,1

Số người chết trong một tuần ở làng A là một đại lượng ngẫu nhiên Y có phân bố xác suất là

Y	0	1	2	3	4
P	0,1	0,3	0,4	0,15	0,05

Giả sử rằng X và Y độc lập.

a) Tìm phân phối xác suất đồng thời của X và Y.

b) Tính P(X > Y).

Đáp số : a)

Y / X	0	1	2	3	4
0	0.04	0.12	0.16	0.06	0.02
1	0.03	0.09	0.12	0.045	0.015
2	0.02	0.06	0.08	0.03	0.01
3	0.01	0.03	0.04	0.015	0.005

b) 0.19.

Bài 18. Cho bảng phân phối xác suất đồng thời của X, Y như sau :

Y / X	4	5
1	0,1	0,06
2	0,3	0,18
3	0,2	0,16

a) Lập bảng phân phối xác suất thành phần của X và Y.

b) Lập bảng phân phối xác suất có điều kiện của X và Y.

c) Tính covariance và hệ số tương quan của X và Y.

Đáp số : a)

X	1	2	3
P_X	0.16	0.48	0.36

Y	4	5
P_Y	0.6	0.4

b)

Y / X	4	5
1	0.17	0.15
2	0.5	0.45
3	0.33	0.4

X / Y	1	2	3
4	0.625	0.625	0.56
5	0.375	0.375	0.44

c) $\text{cov}(X, Y) = 0.02$, $\rho(X, Y) = 0.059$.

Tham số đặc trưng của biến ngẫu nhiên.

Bài 19. Các đại lượng ngẫu nhiên X và Y có bảng phân phối xác suất đồng thời như sau

Y / X	1	2	3
1	0,12	0,15	0,03
2	0,28	0,35	0,07

a) Chứng minh rằng X và Y độc lập.

b) Lập bảng phân phối xác suất của Z = XY. Từ đó tính E(Z) và kiểm tra rằng $E(Z) = E(X)E(Y)$.

Đáp số : b)

Z	1	2	3	4	6
P	0.12	0.43	0.03	0.35	0.07

$E(Z) = 2.89,\ E(X) = 1.7,\ E(Y) = 1.7.$

Bài 20. Cho X, Y là hai đại lượng ngẫu nhiên có phân bố xác suất đồng thời như sau

X \ Y	-1	1
-1	$1/6$	$1/4$
0	$1/6$	$1/8$
1	$1/6$	$1/8$

Hãy tính E(X), E(Y), cov(X,Y) và $\rho(X,Y)$.

Đáp số : $\mu_X = -\frac{1}{8},\ \mu_Y = 0,\ \text{cov}(X,Y) = -0.125,\ \rho(X,Y) = -0.1502.$

Bài 21. Cho X,Y là hai đại lượng ngẫu nhiên có phân bố xác suất đồng thời như sau

X \ Y	-1	0	1
-1	$4/15$	$1/15$	$4/15$
0	$1/15$	$2/15$	$1/15$
1	0	$2/15$	0

a) Tìm μ_X, μ_Y, cov(X,Y) và $\rho(X,Y)$.

b) X và Y có độc lập không ?

Đáp số : a) $\mu_X = -0.467,\ \mu_Y = 0,\ \text{cov}(X,Y) = 0,\ \rho(X,Y) = 0.$

b) X và Y độc lập.

Bài 22. Có hai hộp, mỗi hộp đựng 6 bi. Trong hộp một có : 1 bi mang số 1, 2 bi mang số 2, 3 bi mang số 3. Trong hộp hai có : 2 bi mang số 1, 3 bi mang số 2, 1 bi mang số 3. Rút từ mỗi hộp 1 bi. Gọi X là số ghi trên bi rút ra từ hộp một, Y là số ghi trên bi rút ra từ hộp hai.

a) Hãy lập bảng phân phối xác suất đồng thời của $V = (X, Y)$.

b) Bảng phân phối xác suất lề của X , Y.

c) Kỳ vọng, phương sai của X , Y.

d) Hiệp phương sai, hệ số tương quan.

Đáp số : a)

X \ Y	1	2	3
1	$2/36$	$3/36$	$1/36$

2	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{2}{36}$
3	$\frac{6}{36}$	$\frac{9}{36}$	$\frac{3}{36}$

b)

X	*1*	*2*	*3*
P_X	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$

Y	*1*	*2*	*3*
P_Y	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

c) $\mu_X = 2.33$, $\mu_Y = 1.83$, $\sigma_X^2 = 0.555$, $\sigma_Y^2 = 0.472$.

d) $\text{cov}(X, Y) = 0.0139$, $\rho(X, Y) = 0.027$.

Bài 23. Tung ba lần độc lập một con xúc xắc. Gọi X là số lần mặt chẵn xuất hiện và Y là số lần mặt lẻ xuất hiện.

a) Lập bảng phân phối xác suất của X và Y.

b) Tính hệ số tương quan $\rho(X, Y)$. Nhận xét?

Đáp số : a)

X	*0*	*1*	*2*	*3*
P_X	*0.125*	*0.375*	*0.375*	*0.125*

Y	*0*	*1*	*2*	*3*
P_Y	*0.125*	*0.375*	*0.375*	*0.125*

b) $\rho(X, Y) = -1$, *X và Y phụ thuộc chặt, nghịch biến.*

PHÂN PHỐI XÁC SUẤT

A. BÀI TẬP MẪU

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

a) 2 trai và 2 gái,

b) 1 trai và 3 gái,

c) 4 trai.

Giải

Gọi X là số con trai trong một gia đình có 4 con thì $X \sim B(4; 0.5)$.

a) Xác suất để có hai trai và hai gái trong bốn đứa con là

$$P(X = 2) = C_4^2 (0.5)^2 (0.5)^2 = \frac{3}{8}$$
$$= 0.375.$$

b) Xác suất để có một con trai trong số bốn đứa con là

$$P(X = 1) = C_4^1 (0.5)^1 (0.5)^3 = \frac{1}{4}$$
$$= 0.25.$$

c) Xác suất để cả bốn đều là trai

$$P(X = 4) = C_4^4 (0.5)^4 (0.5)^0 = \frac{1}{16}$$
$$= 0.0625.$$

Bài 2. Một nhà máy sản xuất với tỷ lệ phế phẩm là 7%

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

i) có đúng một phế phẩm,

ii) có ít nhất một phế phẩm,

iii) có nhiều nhất một phế phẩm.

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

Giải

a) Gọi X là số phế phẩm nhận được trong 10 sản phẩm thì $X \sim B(10; 0.07)$.

i) Xác suất để có đúng 1 phế phẩm trong 10 sản phẩm là

$$P(X = 1) = C_{10}^1 (0.07)^1 (1 - 0.07)^{10-1}$$
$$= 10 \cdot 0.07 \cdot (0.93)^9 = 0.3643.$$

ii) Xác suất để có ít nhất một phế phẩm là

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - C_{10}^{0}(0.07)^{0}(0.93)^{10} = 1 - (0.93)^{10} \\ &= 0.516. \end{aligned}$$

iii) Và xác suất để có nhiều nhất một phế phẩm là

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) \\ &= C_{10}^{0}(0.07)^{0}(0.93)^{10} + C_{10}^{1}(0.07)^{1}(0.93)^{9} \\ &= 0.8483. \end{aligned}$$

b) Gọi n là số sản phẩm quan sát để xác suất nhận được ít nhất một phế phẩm ≥ 0.9. Với biến số X chỉ số phế phẩm nhận được trong n lần quan sát này thì $X \sim B(n; 0.07)$. Do

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - C_{n}^{0}(0.07)^{0}(0.93)^{n} \\ &= 1 - (0.93)^{n}. \end{aligned}$$

Từ $P(X \geq 1) \geq 0.9$, ta được bất phương trình

$$1 - (0.93)^{n} \geq 0.9.$$

Giải bất phương trình trên, ta nhận được giá trị $n \geq 31.73$. Vậy phải quan sát ít nhất 32 sản phẩm.

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

Giải

Gọi X là số cuộc gọi nhận được trong 1 phút thì X có phân phối Poisson với trung bình 3, nghĩa là $X \sim P(3)$.

Xác suất để trung tâm bưu điện nhận được 1 cuộc, 2 cuộc và 3 cuộc gọi trong 1 phút lần lượt là

$$P(X = 1) = e^{-3}\frac{3^{1}}{1!} = 0.1494,$$

$$P(X = 2) = e^{-3}\frac{3^{2}}{2!} = 0.224,$$

và $$P(X = 3) = e^{-3}\frac{3^{3}}{3!} = 0.224.$$

Bài 4. Khi tiêm truyền một loại huyết thanh, trung bình có một trường hợp phản ứng trên 1000 trường hợp. Dùng loại huyết thanh này tiêm cho 2000 người. Tính xác suất để

a) có 3 trường hợp phản ứng,

b) có nhiều nhất 3 trường hợp phản ứng,

c) có nhiều hơn 3 trường hợp phản ứng.

Giải

Do xác suất để một người bị phản ứng với loại huyết thanh này là $\frac{1}{1000}$ nên với X chỉ số người bị phản ứng với loại huyết thanh này trong 2000 người thì $X \sim B(2000; 0.001)$.

Vì $p = 0.001 < 0.01$ và $np = 2 < 5$ nên phân phối nhị thức có thể xấp xỉ bằng phân phối Poisson, nghĩa là

$$X \sim P(2000 \times 0.001) = P(2).$$

a) Vậy, xác suất để có ba trường hợp phản ứng trong 1000 trường hợp là

$$P(X = 3) = e^{-2}\frac{2^3}{3!} = \frac{4}{3}e^{-2} = 0.18.$$

b) Xác suất có nhiều nhất 3 trường hợp phản ứng trong 1000 trường hợp là

$$\begin{aligned} P(X \le 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= e^{-2} + 2e^{-2} + 2e^{-2} + \frac{4}{3}e^{-2} \\ &= \frac{19}{3}e^{-2} = 0.86. \end{aligned}$$

c) Và xác xuất có nhiều hơn 3 trường hợp phản ứng là

$$\begin{aligned} P(X > 3) &= 1 - P(X \le 3) \\ &= 1 - \frac{19}{3}e^{-2} = 0.14. \end{aligned}$$

Bài 5. Tỷ lệ một loại bệnh bẩm sinh trong dân số là $p = 0.01$. Bệnh này cần sự chăm sóc đặc biệt lúc mới sinh. Một nhà bảo sinh thường có 20 ca sinh trong một tuần. Tính xác suất để

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

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

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

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

Giải

Gọi X là số trường hợp cần chăm sóc đặc biệt trong 20 ca sinh. Ta có $X \sim B(20;0.01)$.

a) Xác suất để không có trường hợp nào cần chăm sóc đặc biệt là

$$\begin{aligned} P(X = 0) &= C_{20}^0 (0.01)^0 (1 - 0.01)^{20} \\ &= (0.99)^{20} = 0.8179. \end{aligned}$$

b) Xác suất để có đúng một trường hợp cần chăm sóc đặc biệt là

$$\begin{aligned} P(X = 1) &= C_{20}^1 (0.01)^1 (1 - 0.01)^{20-1} \\ &= 20 \cdot (0.01) \cdot (0.99)^{19} = 0.1652. \end{aligned}$$

c) Xác suất có nhiều hơn một trường hợp cần chăm sóc đặc biệt là

$$\begin{aligned} P(X > 1) &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - (0.8179 + 0.1652) = 0.0168. \end{aligned}$$

Khi xấp xỉ phân phối nhị thức bằng phân phối Poisson, nghĩa là $X \sim P(20 \cdot 0.01) = P(0.2)$, ta nhận được

$$P(X = 0) = e^{-0.2} = 0.8187,$$

$$P(X=1) = e^{-0.2}\frac{(0.2)^1}{1!} = 0.1637,$$

và

$$\begin{aligned} P(X>1) &= 1-[P(X=0)+P(X=1)] \\ &= 1-(0.8187+0.1637) = 0.01755. \end{aligned}$$

Kết luận : Với cỡ mẫu 20 và tỷ lệ bệnh $p = 0.01$ thì kết quả của hai loại phân phối này xấp xỉ như nhau.

Bài 6. Cho vectơ ngẫu nhiên $V = (X, Y)$, với X, Y độc lập, $X \sim P(\mu_X)$ và $Y \sim P(\mu_Y)$.

a) Tính xác suất $P(X+Y=n)$,

b) Tính xác suất $P(X=k \mid X+Y=n)$.

Giải

a) Ta có

$$P(X+Y=n) = \sum_{k=0}^{n} P(X=k; Y=n-k).$$

Do X, Y độc lập, $X \sim P(\mu_X)$ và $Y \sim P(\mu_Y)$, nên

$$\begin{aligned} P(X=k; Y=n-k) &= P(X=k)P(Y=n-k) \\ &= e^{-\mu_X}\frac{\mu_X^k}{k!} e^{-\mu_Y}\frac{\mu_Y^{n-k}}{(n-k)!}. \end{aligned}$$

Từ đó suy ra

$$\begin{aligned} P(X+Y=n) &= \sum_{k=0}^{n} e^{-\mu_X}\frac{\mu_X^k}{k!} e^{-\mu_Y}\frac{\mu_Y^{n-k}}{(n-k)!} = \\ &= e^{-(\mu_X+\mu_Y)}\sum_{k=0}^{n}\frac{\mu_X^k}{k!}\frac{\mu_Y^{n-k}}{(n-k)!} = \frac{e^{-(\mu_X+\mu_Y)}}{n!}\sum_{k=0}^{n} C_n^k \mu_X^k \mu_Y^{n-k} \\ &= e^{-(\mu_X+\mu_Y)}\frac{(\mu_X+\mu_Y)^n}{n!}. \end{aligned}$$

b) Từ công thức xác suất có điều kiện

$$\begin{aligned} P(X=k \mid X+Y=n) &= \frac{P(X=k; X+Y=n)}{P(X+Y=n)} = \frac{P(X=k; Y=n-k)}{P(X+Y=n)} \\ &= \frac{P(X=k)P(Y=n-k)}{P(X+Y=n)}, \end{aligned}$$

ta được

$$\begin{aligned} P(X=k \mid X+Y=n) &= \frac{e^{-\mu_X}\dfrac{\mu_X^k}{k!}\cdot e^{-\mu_Y}\dfrac{\mu_Y^{n-k}}{(n-k)!}}{e^{-(\mu_X+\mu_Y)}\dfrac{(\mu_X+\mu_Y)^n}{n!}} \\ &= C_n^k\left(\frac{\mu_Y}{\mu_X+\mu_Y}\right)^{n-k}\cdot\left(\frac{\mu_X}{\mu_X+\mu_Y}\right)^k. \end{aligned}$$

Bài 7. Đường kính của một chi tiết máy do một máy tiện tự động sản xuất có phân phối chuẩn với trung bình $\mu = 50$ mm và độ lệch chuẩn $\sigma = 0.05$ mm. Chi tiết máy được xem là đạt yêu cầu nếu đường kính không sai quá 0.1mm.

a) Tính tỷ lệ sản phẩm đạt yêu cầu.

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

Giải

Gọi X là đường kính của chi tiết máy thì $X \sim N(\mu;\sigma^2)$, với $\mu = 50$ mm và $\sigma = 0.05$ mm.

a) Xét biến cố A : "nhận được sản phẩm đạt yêu cầu", ta có

$$P(A) = P(49.9 \le X \le 50.1).$$

Mặt khác, nếu ta đặt $Y = \dfrac{X-\mu}{\sigma} = \dfrac{X-50}{0.05}$, thì $Y \sim N(0;1)$. Do đó

$$\begin{aligned} P(49.9 \le X \le 50.1) &= P\left(\frac{49.9-50}{0.05} \le \frac{X-50}{0.05} \le \frac{50.1-50}{0.05}\right) \\ &= P(-2 \le Y \le 2) = \varphi(2) - \varphi(-2) = 2\varphi(2) \\ &= 0.9544. \end{aligned}$$

Vậy xác suất để nhận được sản phẩm đạt yêu cầu là 95.44%.

b) Gọi X là số sản phẩm đạt yêu cầu trong 3 sản phẩm lấy ra thì $X \sim B(3;0.9544)$.

Suy ra xác suất để có ít nhất một sản phẩm đạt yêu cầu là

$$\begin{aligned} P(X \ge 1) &= 1 - P(X = 0) \\ &= 1 - C_3^0 (0.9544)^0 (1-0.9544)^3 \\ &= 1 - (0.0456)^3 = 0.9999. \end{aligned}$$

Bài 8. Trọng lượng X (tính bằng gam) một loại trái cây có phân phối chuẩn $N(\mu;\sigma^2)$, với $\mu = 500(\text{gam})$ và $\sigma^2 = 16(\text{gam}^2)$. Trái cây thu hoạch được phân loại theo trọng lượng như sau :

a) loại 1 : trên 505 gam,

b) loại 2 : từ 495 đến 505 gam,

c) loại 3 : dưới 495 gam.

Tính tỷ lệ mỗi loại.

Giải

Gọi X là trọng lượng trái cây thì $X \sim N(\mu;\sigma^2) = N(500;4^2)$. Với $Y = \dfrac{X-500}{4}$ thì $Y \sim N(0;1)$. Do đó

a) Tỷ lệ trái cây loại 1 là

$$\begin{aligned} P(X > 505) &= P\left(\frac{X-500}{4} > \frac{505-500}{4}\right) \\ &= P(Y > 1.25) = \varphi(+\infty) - \varphi(1.25) = 0.5 - \varphi(1.25) \\ &= 0.10565. \end{aligned}$$

b) Tỷ lệ trái cây loại 2 là

$$P(495 \le X \le 505) = P\left(\frac{495-500}{4} \le \frac{X-500}{4} \le \frac{505-500}{4}\right)$$
$$= P(-1.25 \le Y \le 1.25) = 0.7887.$$

c) Và tỷ lệ của loại 3 là

$$P(X < 495) = P(\frac{X-500}{4} < \frac{495-500}{4})$$
$$= P(Y < -1.25) = \varphi(-1.25) - \varphi(-\infty)$$
$$= -\varphi(1.25) + 0.5 = 0.10565.$$

Vậy, trái cây thu hoạch được có khoảng 11% loại 1, 78% loại 2 và 11% loại 3.

B. BÀI TẬP.

Bài 1. Có 8000 sản phẩm trong đó có 2000 sản phẩm không đạt tiêu chuẩn kỹ thuật. Lấy ngẫu nhiên (không hoàn lại) 10 sản phẩm. Tính xác suất để trong 10 sản phẩm lấy ra có 2 sản phẩm không đạt tiêu chuẩn.

Đáp số : 0.282.

Bài 2. Đường kính của một loại chi tiết do một máy sản xuất có phân phối chuẩn, kỳ vọng 20mm, phương sai $(0,2\text{mm})^2$. Lấy ngẫu nhiên 1 chi tiết máy. Tính xác suất để

a) có đường kính trong khoảng 19,9mm đến 20,3mm,

b) có đường kính sai khác với kỳ vọng không quá 0,3mm.

Đáp số : a) 0.6247.
b) 0.8664.

Bài 3. Một máy dệt có 4000 ống sợi. Xác suất để mỗi ống sợi bị đứt trong 1 phút là 0,0005. Tính xác suất để trong 1 phút

a) có 3 ống sợi bị đứt,

b) có ít nhất 2 ống sợi bị đứt.

Đáp số : a) 0.18.
b) 0.595.

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

a) không phải tất cả 4 chiếc ôtô đều được thuê,

b) tất cả 4 chiếc ôtô đều được thuê,

c) cửa hàng không đáp ứng được yêu cầu,

d) trung bình có bao nhiêu ôtô được thuê,

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

Đáp số : a) 0.857.
b) 0.1429.
c) 0.0527.
d) 2.
e) 5.

Bài 5. Một tổng đài bưu điện có các cuộc điện thoại gọi đến xuất hiện ngẫu nhiên, độc lập với nhau và có tốc độ trung bình 2 cuộc gọi trong 1 phút. Tìm xác suất để

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

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

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

Đáp số : a) 0.1563.
b) 0.3679.
c) 0.284.

Bài 6. Tỷ lệ cử tri ủng hộ ứng cử viên A trong một cuộc bầu cử là 60%. Người ta hỏi ý kiến 20 cử tri được chọn một cách ngẫu nhiên. Gọi X là số người bỏ phiếu cho A trong 20 người đó.

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

b) Tìm $P(X \le 10)$.

c) Tìm $P(X > 12)$.

d) Tìm $P(X = 11)$.

Đáp số : a) $\mu_X = 12$, $\sigma_X = 2.191$, $\text{Mod}[X] = 12$.
b) 0.245.
c) 0.416.
d) 0.16.

Bài 7. Xác suất để một máy sản xuất ra phế phẩm là 0.02.

a) Tính xác suất để trong 10 sản phẩm do máy sản xuất có không quá 1 phế phẩm.

b) Một ngày máy sản xuất được 250 sản phẩm. Tìm số phế phẩm trung bình và số phế phẩm tin chắc nhất của máy đó trong một ngày.

Đáp số : a) 0.98.
b) Số phế phẩm trung bình = 5, số phế phẩm tin chắc nhất = 5.

Bài 8. Một máy sản xuất ra sản phẩm loại A với xác suất 0.485. Tính xác suất sao có trong 200 sản phẩm do máy sản xuất ra có ít nhất 95 sản phẩm loại A.

Đáp số : 0.6103.

Bài 9. Xác suất để một máy sản xuất ra sản phẩm loại A là 0.25. Tính xác suất để trong 80 sản phẩm do máy sản xuất ra có từ 25 đến 30 sản phẩm loại A.

Đáp số : 0.0936.

Bài 10. Gieo 100 hạt giống của một loại nông sản. Xác suất nảy mầm của mỗi hạt là 0.8. Tính xác suất để có ít nhất 90 hạt nảy mầm.

Đáp số : 0.0062.

Bài 11. Một sọt cam có 10 trái trong đó có 4 trái hư. Lấy ngẫu nhiên ra 3 trái.

a) Tính xác suất lấy được 3 trái hư.

b) Tính xác suất lấy được 1 trái hư

c) Tính xác suất lấy được ít nhất 1 trái hư.

d) Tính xác suất lấy được nhiều nhất 2 trái hư.

Đáp số : a) 0.033.
b) 0.5.
c) 0.83.
d) 0.967.

Bài 12. Giả sử tỷ lệ dân cư mắc bệnh A trong vùng là 10%. Chọn ngẫu nhiên 1 nhóm 400 người.

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

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

Đáp số : a) 0.9564.
b) 0.9525.

Bài 13. Một nhà xã hội học cho rằng 12% số dân của thành phố ưa thích một bộ phim A mới chiếu trên tivi. Để khẳng định dự đoán này, ông ta chọn một mẫu ngẫu nhiên gồm 500 người để hỏi ý kiến và thấy 75 người trả lời ưa thích bộ phim đó. Tính xác suất để trong một mẫu ngẫu nhiên gồm 500 người, số người ưa thích bộ phim ít nhất là 75 nếu giả thuyết p = 12% là đúng.

Đáp số : a) 0.0233.
b) 0.9525.

Bài 14. Cho X và Y là hai đại lượng ngẫu nhiên độc lập.

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

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

Đáp số : a)

X+Y	*0*	*1*	*2*	*3*
P	$\frac{64}{125}$	$\frac{48}{125}$	$\frac{12}{125}$	$\frac{1}{125}$

b)

Z	*0*	*1*	*2*	*3*
P	$\frac{16}{50}$	$\frac{24}{50}$	$\frac{9}{50}$	$\frac{1}{50}$

Bài 15. Xác suất để một con gà đẻ trong ngày là 0,6. Nuôi 5 con.

1) Tính xác suất để trong một ngày :

a) không con nào đẻ,

b) cả 5 con đẻ,

c) có ít nhất 1 con đẻ,

d) có ít nhất 2 con đẻ.

2) Nếu muốn mỗi ngày có trung bình 100 trứng thì phải nuôi bao nhiêu con gà.

Đáp số : 1) a) 0.01024, *b)* 0.07776, *c)* 0.98976, *d)* 0.91296.
2) 167 con.

Bài 16. Sản phẩm sau khi hoàn tất được đóng thành kiện, mỗi kiện gồm 10 sản phẩm với tỷ lệ thứ phẩm là 20%. Trước khi mua hàng, khách hàng muốn kiểm tra bằng cách từ mỗi kiện chọn ngẫu nhiên 3 sản phẩm.

a) Tìm luật phân phối xác suất của số sản phẩm tốt trong 3 sản phẩm lấy ra.

b) Nếu cả 3 sản phẩm được lấy ra đều là sản phẩm tốt thì khách hàng sẽ đồng ý mua kiện hàng đó. Tính xác suất để khi kiểm tra 100 kiện có ít nhất 60 kiện được mua.

Đáp số : a) Gọi X là số sản phẩm tốt trong 3 sản phẩm lấy ra, $X \sim H(10;8;3)$,

X	*0*	*1*	*2*	*3*
P	*0*	*0.066*	*0.467*	*0.467*

b) 0.0038.

Bài 17. Xác suất trúng số là 1%. Mỗi tuần mua một vé số. Hỏi phải mua vé số liên tiếp trong tối thiểu bao nhiêu tuần để có không ít hơn 95% hy vọng trúng số ít nhất 1 lần. $(\text{cho } \lg 99 = 1,9956;\ \lg 5 = 0,6990)$

Đáp số : 296 *tuần.*

Bài 18. Bưu điện dùng một máy tự động đọc địa chỉ trên bì thư để phân loại từng khu vực gởi đi, máy có khả năng đọc được 5000 bì thư trong 1 phút. Khả năng đọc sai 1 địa chỉ trên bì thư là 0,04% (xem như việc đọc 5000 bì thư này là 5000 phép thử độc lập).

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

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

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

Đáp số : a) 2.
b) 2.
c) 0.3233.

Bài 19. Xác suất để một máy sản xuất ra một phế phẩm là 0.001. Tính xác suất để trong 4000 sản phẩm do máy này sản xuất ra có không quá 5 phế phẩm.

Đáp số : 0.7851.

Bài 20. Tại một điểm bán vé máy bay, trung bình trong 10 phút có 4 người đến mua vé. Tính xác suất để:

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

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

Đáp số : a) 0.0596.
b) 0.4335.

Bài 21. Lãi suất (%) đầu tư vào một dự án năm 2000 được coi như 1 đại lượng ngẫu nhiên phân phối theo quy luật chuẩn. Theo đánh giá của uỷ ban đầu tư thì lãi suất cao hơn 20% có xác suất 0,1587, và lãi suất cao hơn 25% có xác suất là 0,0228. Vậy khả năng đầu tư mà không bị thua lỗ là bao nhiêu?.

Đáp số : 0.5.

Bài 22. Độ dài của một chi tiết máy được tiện ra có phân phối chuẩn $N(\mu\,\text{cm};(0,2\text{cm})^2)$. Sản phẩm coi là đạt nếu độ dài sai lệch so với độ dài trung bình không quá 0,3cm.

a) Tính xác suất chọn ngẫu nhiên 1 sản phẩm thì được sản phẩm yêu cầu.

b) Chọn ngẫu nhiên 3 sản phẩm. Tính xác suất có ít nhất 2 sản phẩm đạt yêu cầu .

Đáp số : a) 0.8664.
b) 0.9512.

Bài 23. Trọng lượng của 1 loại trái cây có quy luật phân phối chuẩn với trọng lượng trung bình là 250g, độ lệch chuẩn về trọng lượng là 5g. Một người lấy 1 trái từ trong sọt trái cây ra.

a) Tính xác suất người này lấy được trái loại 1 (trái loại 1 là trái có trọng lượng > 260g).

b) Nếu lấy được trái loại 1 thì người này sẽ mua sọt đó. Người này kiểm tra 100 sọt, tính xác suất mua được 6 sọt.

Đáp số : a) 0.1587.
b) 0.0029.

Bài 24. Một công ty kinh doanh mặt hàng A dự định sẽ áp dụng một trong 2 phương án kinh doanh. Ký hiệu X_1 là lợi nhuận thu được khi áp dụng phương án thứ 1, X_2 là lợi nhuận thu được khi áp dụng phương án thứ 2. X_1, X_2 đều được tính theo đơn vị triệu đồng/ tháng) và $X_1 \sim N(140, 2500)$, $X_2 \sim N(200, 3600)$. Nếu biết rằng, để công ty tồn tại và phát triển thì lợi nhuận thu được từ mặt hàng kinh doanh A phải đạt ít nhất 80 triệu đồng/tháng. Hãy cho biết công ty nên áp dụng phương án nào để kinh doanh mặt hàng A? Vì sao?.

Đáp số : $P(X_1 \geq 80) = 0.8849$, $P(X_2 \geq 80) = 0.9772$, *nên ta chọn phương án thứ 2.*

Bài 25. Có hai thị trường A và B, lãi suất của cổ phiếu trên hai thị trường này là các biến ngẫu nhiên phân phối chuẩn, độc lập với nhau, có kỳ vọng và phương sai được cho trong bảng dưới đây:

	Trung bình	Phương sai
Thị trường A	19%	36
Thị trường B	22%	100

Nếu mục đích là đạt lãi suất tối thiểu bằng 10% thì nên đầu tư vào loại cổ phiếu nào?

Đáp số : Nên đầu tư vào loại cổ phiếu trên thị trường A.

Bài 26. Nghiên cứu chiều cao của những người trưởng thành, người ta nhận thấy rằng chiều cao đó tuân theo quy luật phân bố chuẩn với trung bình là 175cm và độ lệch tiêu chuẩn 4cm. Hãy xác định :

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

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

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

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

Đáp số : a) 0.1056.
b) 0.6793.
c) 173.24.
d) 6.6.

Bài 27. Chiều dài của chi tiết được gia công trên máy tự động là biến ngẫu nhiên tuân theo quy luật phân phối chuẩn với độ lệch tiêu chuẩn là 0.01mm. Chi tiết được coi là đạt tiêu chuẩn nếu kích thước thực tế của nó sai lệch so với kích thước trung bình không vượt quá 0.02mm.

a) Tìm tỷ lệ chi tiết không đạt tiêu chuẩn.

b) Xác định độ đồng đều (phương sai) cần thiết của sản phẩm để tỷ lệ chi tiết không đạt tiêu chuẩn chỉ còn 1%.

Đáp số : a) 0.9544.
b) 0.03^2.

Bài 28. Trọng lượng X của một loại trái cây ở nông trường được biết có kỳ vọng 250gr và phương sai $81(gr)^2$. Trái cây được đóng thành sọt, mỗi sọt 100 trái. Mỗi sọt được gọi là loại A nếu trọng lượng không dưới 25kg. Kiểm tra ngẫu nhiên 100 sọt. Tính xác suất :

a) có nhiều nhất 30 sọt loại A,

b) ít nhất 10 sọt loại A.

Đáp số : a) 0.8413.

b) 0.9987.

Bài 29. Một trạm cho thuê xe Taxi có 3 chiếc xe. Hàng ngày trạm phải nộp thuế 8USD cho 1 chiếc xe (bất kể xe đó có được thuê hay không). Mỗi chiếc được cho thuê với giá 20USD. Giả sử số xe được yêu cầu cho thuê của trạm trong 1 ngày là đại lượng ngẫu nhiên có phân phối Poisson với $\mu = 2,8$.

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

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

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

MẪU THỐNG KÊ & ƯỚC LƯỢNG THAM SỐ

A. BÀI TẬP MẪU

Bài 1. Đo lượng cholesterol (đơn vị mg%) cho một số người, ta được

X(mg%)	150-160	160-170	170-180	180-190	190-200	200-210
Số người	2	4	5	6	4	3

a) Tính trung bình mẫu $\overline{X}$ và độ lệch chuẩn S_X.

b) Một mẫu thứ nhì Y có 30 người cho $\overline{Y} = 180$ mg%, $S_Y = 16$ mg%.

Nhập hai mẫu lại, tính trung bình và độ lệch chuẩn của mẫu nhập.

Giải

a) Tính trung bình mẫu μ_X và độ lệch chuẩn S_X.

X(mg%)	155	165	175	185	195	205
Số người	2	4	5	6	4	3

$\overline{X} = \mu_X = 181.25$, $S_X = 14.98$, $n = 24$.

b) Một mẫu thứ nhì Y có 30 người cho $\mu_Y = 180$mg%, $S_Y = 16$mg%. Nhập hai mẫu lại và gọi Z là mẫu nhập. Ta có cỡ mẫu nhập là $N = 24 + 30 = 54$ người và trung bình của mẫu nhập là

$$\mu_Z = \frac{24\mu_X + 30\mu_Y}{54} = \frac{24 \times 181.25 + 30 \times 180}{54} = 180.55.$$

Để tính độ lệch chuẩn mẫu nhập, ta dùng công thức

$$S_Z^2 = \frac{1}{N-1}\left(\sum Z^2 - N\overline{Z}^2\right),$$

trong đó $\overline{Z} = \mu_Z = 180.55$ và $\sum Z^2 = \sum X^2 + \sum Y^2$.

Mặt khác, từ mẫu của X, ta có

$$\sum X^2 = 793598,71.$$

Với mẫu Y, do

$$S_Y^2 = \frac{1}{n_Y - 1}\left(\sum Y^2 - n_Y\overline{Y}^2\right),$$

ta suy ra

$$\sum Y^2 = (n_Y - 1)S_Y^2 + n_Y\overline{Y}^2 = 29 \times (16)^2 + 30 \times (180)^2 = 979424.$$

Do đó

$$S_Z^2 = \frac{1}{53}\left((793598,71 + 979424) - 54 \times (180.55)^2\right) = 239.89,$$

nên $S_Z = 15,48$ mg%.

Bài 2. Có 3 mẫu quan sát sức nặng con người, kết quả ghi nhận

	Lần quan sát	Trung bình	Độ lệch
Mẫu 1	70	55kg	8.30kg
Mẫu 2	75	57kg	8.60kg
Mẫu 3	90	54kg	8.50kg

Nhập chung 3 mẫu lại, tính trung bình và độ lệch mẫu nhập.

Dựa vào mẫu nhập để ước lượng trung bình của tổng thể ở độ tin cậy 95% và 99%.

Giải

Gọi X_1, X_2 và X_3 lần lượt là biến số ngẫu nhiên cho bởi các mẫu 1, 2 và 3. Ta có

Mẫu 1 có cỡ mẫu $n_1 = 70$, trung bình $\bar{X}_1 = 55$, độ lệch $S_{X_1} = 8.3$,

Mẫu 2 có cỡ mẫu $n_2 = 75$, trung bình $\bar{X}_2 = 57$, độ lệch $S_{X_2} = 8.6$,

Mẫu 3 có cỡ mẫu $n_3 = 90$, trung bình $\bar{X}_3 = 54$, độ lệch $S_{X_3} = 8.5$.

Từ đó, với X chỉ mẫu nhập, ta có cỡ mẫu $n = n_1 + n_2 + n_3 = 235$, trung bình

$$\begin{aligned}\bar{X} &= \frac{1}{n}\sum X = \frac{1}{n}\left(\sum X_1 + \sum X_2 + \sum X_3\right)\\ &= \frac{1}{n}\left(n_1\bar{X}_1 + n_2\bar{X}_2 + n_3\bar{X}_3\right)\\ &= \frac{70\times 55 + 75\times 57 + 90\times 54}{235} = 55.25.\end{aligned}$$

Để tính độ lệch cho X, ta dùng công thức

$$S_X^2 = \frac{1}{n-1}\left(\sum X^2 - n\bar{X}^2\right),$$

trong đó

$$\sum X^2 = \sum X_1^2 + \sum X_2^2 + \sum X_3^2.$$

Mà

$$\sum X_i^2 = (n_i - 1)S_{X_i}^2 + n_i\bar{X}_i^2,$$

với $i = 1, 2, 3$, nên ta có

$$\sum X_1^2 = 69\times(8.3)^2 + 70\times 55^2 = 216503.41,$$

$$\sum X_2^2 = 74\times(8.6)^2 + 75\times 57^2 = 249148.04,$$

$$\sum X_3^2 = 89\times(8.5)^2 + 90\times 54^2 = 268870.25.$$

Từ đó suy ra

$$\begin{aligned}S_X^2 &= \frac{(216503.41 + 249148.04 + 268870.25) - 235\times(55.25)^2}{234}\\ &= 73.37.\end{aligned}$$

Độ lệch : $S_X = 8.56$ kg .

Gọi μ là trung bình tổng thể cần ước lượng. Ta có

$$T = \frac{(\overline{X} - \mu)\sqrt{n}}{S_X} \sim St(n-1),$$

nghĩa là

$$T = \frac{(55.25 - \mu)\sqrt{235}}{8.56} \sim St(234) \equiv N(0;1).$$

Với độ tin cậy $\gamma = 0.95$, ta được $C = 1.96$, và $\mu = 55.25 \pm 1.96 \times \frac{8.56}{\sqrt{235}}$. Nên ta có khoảng ước lượng cho μ là $[54.364; 56.136]$

Với độ tin cậy $\gamma = 0.99$, ta tìm được $C = 2.58$, và $\mu = 55.25 \pm 2.58 \times \frac{8.56}{\sqrt{235}}$, nên ta tìm được khoảng ước lượng cho μ là $[53.81; 56.69]$.

Bài 3. Đo đường kính của một chi tiết máy do một máy tiện tự động sản xuất, ta ghi nhận được số liệu như sau:

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

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

a) Tính trung bình mẫu $\overline{X}$ và độ lệch chuẩn S_X của mẫu.

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

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

Giải

a) Ta được cỡ mẫu $n = 53$, trung bình $\overline{X} = 12.21$, độ lệch $S_X = 0.103$.

b) Ta dùng thống kê

$$T = \frac{(\overline{X} - \mu)\sqrt{n}}{S_X} \sim St(n-1).$$

Với số liệu mẫu, ta có

$$T = \frac{(12.21 - \mu)\sqrt{53}}{0.103} \sim St(52) \equiv N(0;1).$$

Ở độ tin cậy $\gamma = 0.95$, ta tìm được $C = 1.96$. Do đó ước lượng đường kính trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{S_X}{\sqrt{n}} = 12.21 \pm 1.96 \times \frac{0.103}{\sqrt{53}},$$

và ta nhận được khoảng ước lượng $[12.18; 12.24]$.

c) Do sai số của ước lượng là $C\frac{S_X}{\sqrt{n}}$ nên nếu muốn sai số ước lượng không quá $\varepsilon = 0.02mm$, ta phải có

$$C\frac{S_X}{\sqrt{n}} \le \varepsilon.$$

Với độ tin cậy 0.95, thì $C = 1.96$ và ta nhận được bất phương trình

$$n \ge \left(C\frac{S_X}{\varepsilon}\right)^2 = \left(1.96\frac{0.103}{0.02}\right)^2 = 101.89.$$

Vậy phải quan sát ít nhất 102 trường hợp.

Bài 4. Đem cân một số trái cây vừa thu hoạch, ta được kết quả sau

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

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

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

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

Giải

a) Từ số liệu của mẫu, ta có

$$n = 82,\ \overline{X} = 225.854,\ S_X = 13.259.$$

Để tìm khoảng ước lượng của trung bình tổng thể μ khi chưa biết phương sai tổng thể, ta dùng thống kê

$$T = \frac{(\overline{X} - \mu)\sqrt{n}}{S_X} \sim St(n-1).$$

Nhận được từ bộ số liệu của mẫu, ta có

$$T = \frac{(225.854 - \mu)\sqrt{82}}{13.259} \sim St(81) \equiv N(0;1).$$

Với độ tin cậy $\gamma = 0.95$, ta nhận được $C = 1.96$. Nên ước lượng của trọng lượng trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{S_X}{\sqrt{n}} = 225.854 \pm 1.96 \times \frac{13.259}{\sqrt{82}}.$$

Ta có khoảng ước lượng là $[222.98; 228.72]$.

Tương tự, với độ tin cậy $\gamma = 0.99$, ta tìm được $C = 2.58$. Từ đó ta suy ra khoảng ước lượng là $[222.08; 229.63]$.

b) Do sai số của ước lượng là $C\frac{S_X}{\sqrt{n}}$ nên nếu muốn sai số ước lượng không quá $\varepsilon = 2$ gam, ta phải có

$$C\frac{S_X}{\sqrt{n}} \le \varepsilon.$$

Ở độ tin cậy 99% thì $C = 2.58$, ta có bất phương trình

$$n \geq \left(C\frac{S_X}{\varepsilon}\right)^2 = \left(2.58\frac{13.259}{2}\right)^2 = 292.551.$$

Như vậy, cần phải quan sát ít nhất 293 trái.

c) Từ bộ số liệu, ta có tần số trái cây loại A là

$$f = \frac{18+15}{82} = 0.4024.$$

Để ước lượng tỷ lệ p trái cây loại A của tổng thể, ta dùng thống kê

$$T = \frac{(f-p)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1),$$

nghĩa là

$$T = \frac{(0.4024-p)\sqrt{82}}{\sqrt{0.4024\times 0.5976}} \sim St(81) \equiv N(0;1).$$

Với độ tin cậy $\gamma = 0.95$, ta tìm được $C = 1.96$ và do đó

$$p = 0.4024 \pm 1.96\times\sqrt{\frac{0.4024\times 0.5976}{82}}.$$

Ta được khoảng ước lượng cho tỷ lệ p của trái cây loại A $[0.296; 0.509]$.

Ta có sai số của ước lượng là

$$\varepsilon = C\sqrt{\frac{f(1-f)}{n}}.$$

Với độ tin cậy $\gamma = 0.99$, ta được $C = 2.58$ nên với dữ liệu cho, để sai số ước lượng không quá $\varepsilon = 0.04$ ở độ tin cậy 0.99, ta nhận được bất phương trình

$$2.58\times\sqrt{\frac{0.4024\times 0.5976}{n}} \leq 0.04.$$

Từ đó suy ra $n \geq 1000.43$, nghĩa là phải quan sát ít nhất 1001 trường hợp.

Bài 5. Người ta đo ion Na^+ trên một số người và ghi nhận lại được kết quả như sau

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

a) Tính trung bình mẫu $\overline{X}$ và phương sai mẫu S_X^2.

b) Ước lượng trung bình μ và phương sai σ^2 của tổng thể ở độ tin cậy 0.95.

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

Giải

a) Từ các số liệu nhận được của mẫu, ta có

$$n = 12, \overline{X} = 137.83, \ S_X^2 = 19.42, \text{ và } S_X = 4.41.$$

b) Để ước lượng trung bình μ, ta dùng thống kê

$$T = \frac{(\overline{X}-\mu)\sqrt{n}}{S_X} \sim St(n-1),$$

nghĩa là

$$T = \frac{(137.83 - \mu)\sqrt{12}}{4.41} \sim St(11).$$

Với độ tin cậy $\gamma = 0.95$, ta có $C = t_{0.05}^{11} = 2.201$. Ta dễ dàng tìm được khoảng ước lượng cho trung bình μ là $[135.01; 140.63]$.

Để ước lượng phương sai tổng thể khi chưa biết trung bình của tổng thể, ta dùng thống kê

$$Y = \frac{(n-1)S_X^2}{\sigma^2} \sim \chi^2(n-1),$$

nghĩa là

$$Y = \frac{11 \times (19.42)}{\sigma^2} \sim \chi^2(11).$$

Với độ tin cậy $\gamma = 0.95$, ta tìm được a và b sao cho

$$P(Y \le a) = P(Y \ge b) = \frac{1-\gamma}{2}.$$

Từ bảng phân phối xác suất của phân phối Chi-Bình phương, ta tìm được

$$a \equiv \chi^2_{\frac{1+\gamma}{2}} = 3.816, \text{ và } b \equiv \chi^2_{\frac{1-\gamma}{2}} = 21.92.$$

Do đó

$$3.816 \le \frac{(n-1)S_X^2}{\sigma^2} \le 21.92,$$

và ta nhận được bất đẳng thức

$$\frac{11 \times (19.42)}{21.92} \le \sigma^2 \le \frac{11 \times (19.42)}{3.816}$$

Từ đó suy ra ước lượng cho phương sai tổng thể là $[9.76; 56.1]$.

c) Sai số của ước lượng trung bình cho bởi $C\frac{S_X}{\sqrt{n}}$, nên để sai số này không quá $\varepsilon = 1$, ta giải bất phương trình

$$C\frac{S_X}{\sqrt{n}} \le \varepsilon = 1.$$

Suy ra

$$n \ge \left(C\frac{S_X}{\varepsilon}\right)^2 = \left(2.201\frac{4.41}{1}\right)^2 = 94.2.$$

Vậy phải quan sát ít nhất 95 người.

Bài 6. Quan sát tuổi thọ X (giờ) của một số bóng đèn do xí nghiệp A sản xuất, ta ghi nhận

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

với N chỉ số trường hợp theo từng giá trị của X.

a) Tính trung bình mẫu $\bar{X}$ và độ lệch chuẩn mẫu S_X.

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

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

Giải

a) Từ bảng số liệu của mẫu, ta có

$$n = 128, \ \bar{X} = 1391.41 \text{ và } S_X = 234.45.$$

b) Để ước lượng trung bình tuổi thọ của bóng đèn, ta dùng thống kê

$$T = \frac{(\bar{X} - \mu)\sqrt{n}}{S_X} \sim St(n-1).$$

Dựa theo bộ số liệu của thống kê, ta nhận được

$$T = \frac{(1391.41 - \mu)\sqrt{128}}{234.45} \sim St(127) \equiv N(0;1).$$

Với độ tin cậy $\gamma = 0.95$, ta tìm được $C = 1.96$ và do đó ước lượng của trung bình cho bởi

$$\mu = \bar{X} \pm C\frac{S_X}{\sqrt{n}} = 1391.41 \pm 1.96 \times \frac{234.45}{\sqrt{128}}.$$

Ta được khoảng ước lượng cho tuổi thọ trung bình của bóng đèn là $[1350.79; 1432.03]$.

c) Sai số của ước lượng cho bởi $C\frac{S_X}{\sqrt{n}}$, nên để sai số không quá 30 giờ, ta có

$$C\frac{S_X}{\sqrt{n}} = 1.96 \times \frac{234.45}{\sqrt{n}} \le 30,$$

Giải bất phương trình trên, ta tìm được $n = 234.63$. Vậy, phải quan sát ít nhất 235 bóng đèn.

Bài 7. Ta muốn ước lượng tỷ lệ viên thuốc bị sức mẻ p trong một lô thuốc lớn.

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

b) Quan sát ngẫu nhiên 200 viên, thấy có 18 viên bị sứt mẻ. Hãy ước lượng p ở độ tin cậy 0.95. Khi đó, nếu muốn sai số ước lượng không quá 0.01 với độ tin cậy 0.95 thì phải quan sát ít nhất mấy viên ?

Giải

a) Để ước lượng tỷ lệ thuốc, p, bị sứt mẻ trong tổng thể, ta dùng thống kê

$$T = \frac{(f-p)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1),$$

nghĩa là

$$p = f \pm C\sqrt{\tfrac{f(1-f)}{n}}.$$

Hơn nữa, từ nhận xét rằng hàm số $y = x(1-x)$ đạt giá trị lớn nhất là $\frac{1}{4}$ tại điểm $x = \frac{1}{2}$, ta có thể viết

$$\frac{C^2 f(1-f)}{\varepsilon^2} \le \frac{\frac{1}{4}C^2}{\varepsilon^2} = \frac{C^2}{4\varepsilon^2}.$$

Do đó, ta có thể chọn cỡ mẫu

$$n \ge \frac{C^2}{4\varepsilon^2}.$$

Với độ tin cậy $\gamma = 0.95$, ta tìm được $C = 1.96$ và do đó

$$n \ge \frac{C^2}{4\varepsilon^2} = \left(\frac{1.96}{2\cdot 0.01}\right)^2 = 9604.$$

Vậy phải quan sát ít nhất 9604 viên thuốc.

b) Theo giả thuyết, ta có tần số của thuốc bị sứt mẻ là $f = \frac{18}{200} = 0.09$. Với độ tin cậy $\gamma = 0.95$, ta có

$$p = f \pm C\sqrt{\frac{f(1-f)}{n}} = 0.09 \pm 1.96\sqrt{\frac{0.09\cdot 0.91)}{200}}.$$

Và ta có khoảng ước lượng p với độ tin cậy 0.95 là $[0.051; 0.13]$

Nếu muốn sai số ước lượng không quá 0.01, ta có bất phương trình $C\sqrt{\frac{f(1-f)}{n}} \le \varepsilon$. Suy ra

$$n \ge (1.96)^2 \frac{0.09(1-0.09)}{(0.01)^2} = 3146.27.$$

Vậy phải quan sát ít nhất 3147 viên thuốc.

Bài 8. Quan sát chiều cao X (cm) của một số người, ta ghi nhận

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

a) Tính $\overline{X}$ và S_X^2.

b) Ước lượng μ và σ^2 ở độ tin cậy 0.95.

Giải

a) Ta có bảng tần số xuất hiện số liệu như sau

X(cm)	142.5	147.5	152.5	157.5	162.5	167.5
Số người	1	3	7	9	5	2

Do đó, ta nhận được

$$n = 27, \overline{X} = 156.2, \quad S_X^2 = 37.68, \text{ và } S_X = 6.14.$$

b) Để ước lượng trung bình μ, ta dùng thống kê

$$T = \frac{(\overline{X}-\mu)\sqrt{n}}{S_X} \sim St(n-1),$$

nghĩa là

$$T = \frac{(156.2 - \mu)\sqrt{27}}{6.14} \sim St(26).$$

Với độ tin cậy $\gamma = 0.95$, ta có $C = t_{0.05}^{26} = 2.056$. Và do đó ước lượng của trung bình cho bởi

$$\mu = \bar{X} \pm C\frac{S_X}{\sqrt{n}} = 156.2 \pm 2.056 \times \frac{6.14}{\sqrt{27}}.$$

Ta tìm được khoảng ước lượng cho trung bình μ là $[153.77;158.63]$.

Để ước lượng phương sai tổng thể khi chưa biết trung bình của tổng thể, ta dùng thống kê

$$Y = \frac{(n-1)S_X^2}{\sigma^2} \sim \chi^2(n-1),$$

nghĩa là

$$Y = \frac{26 \times (37.68)}{\sigma^2} \sim \chi^2(26).$$

Với độ tin cậy $\gamma = 0.95$, ta tìm được a và b sao cho

$$P(Y \le a) = P(Y \ge b) = \frac{1-\gamma}{2}.$$

Từ bảng phân phối xác suất của phân phối Chi-Bình phương, ta tìm được

$$a \equiv \chi^2_{\frac{1+\gamma}{2}} = 13.844, \text{ và } b \equiv \chi^2_{\frac{1-\gamma}{2}} = 41.923.$$

Do đó

$$3.816 \le \frac{(n-1)S_X^2}{\sigma^2} \le 21.925,$$

và ta nhận được bất đẳng thức

$$\frac{26 \times (37.68)}{41.923} \le \sigma^2 \le \frac{26 \times (37.68)}{13.844}.$$

Từ đó suy ra ước lượng cho phương sai tổng thể là $[23.38;70.80]$.

Bài 9. Một loại thuốc mới đem điều trị cho 50 người bị bệnh B, kết quả có 40 người khỏi bệnh.

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

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

Giải

a) Theo giả thuyết, ta có tần số khỏi bệnh là $f = \frac{40}{50} = 0.8$.

Để ước lượng tỷ lệ khỏi bệnh p của tổng thể, ta dùng thống kê

$$T = \frac{(f-p)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1),$$

nghĩa là

$$T = \frac{(0.8-p)\sqrt{50}}{\sqrt{0.8 \times 0.2}} \sim St(49) \equiv N(0;1).$$

Với độ tin cậy $\gamma = 0.95$, ta tìm được $C = 1.96$ và do đó

$$p = 0.8 \pm 1.96 \times \sqrt{\frac{0.8 \times 0.2}{50}}.$$

Ta được khoảng ước lượng cho tỷ lệ khỏi bệnh p là $[0.69; 0.91]$.

Với độ tin cậy $\gamma = 0.99$, ta tìm được $C = 2.58$. Tương tự, ta có khoảng ước lượng cho tỷ lệ khỏi bệnh với độ tin cậy 99% là $[0.65; 0.946]$.

b) Ta có sai số của ước lượng là

$$\varepsilon = C\sqrt{\frac{f(1-f)}{n}}.$$

Với độ tin cậy $\gamma = 0.95$, ta được $C = 1.96$ nên với dữ liệu cho, để sai số ước lượng không quá $\varepsilon = 0.02$, ta nhận được bất phương trình

$$1.96 \times \sqrt{\frac{0.8 \times 0.2}{n}} \le 0.02.$$

Từ đó suy ra $n \ge 1536.64$, nghĩa là phải quan sát ít nhất 1537 trường hợp.

Bài 10. Một loại bệnh có tỷ lệ tử vong là 0.01. Muốn chứng tỏ một loại thuốc có hiệu nghiệm (nghĩa là hạ thấp được tỷ lệ tử vong nhỏ hơn 0.005) ở độ tin cậy 0.95 thì phải thử thuốc đó trên ít nhất bao nhiêu người?

Giải

Theo giả thuyết, tần số tử vong là $f = 0.01$.

Để ước lượng tỷ lệ khỏi bệnh p của tổng thể, ta dùng thống kê

$$T = \frac{(f-p)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1),$$

nghĩa là

$$T = \frac{(0.01-p)\sqrt{n}}{\sqrt{0.01 \times 0.99}} \sim St(n-1).$$

Và ta có sai số của ước lượng này là $\varepsilon = C\sqrt{\frac{f(1-f)}{n}}$.

Với $\varepsilon = 0.005$, độ tin cậy $\gamma = 0.95$, ta được $C = 1.96$. Để hạ thấp độ tử vong nhỏ hơn 0.005, ta giải bất phương trình $C\sqrt{\frac{f(1-f)}{n}} \le \varepsilon$. Suy ra

$$n \ge (1.96)^2 \frac{0.01(1-0.01)}{(0.005)^2} = 1521.27.$$

Vậy phải thử thuốc ít nhất 1522 người.

B. BÀI TẬP

Bài 1. Đo lượng huyết tương của 8 người mạnh khoẻ, ta có

2,86	3,37	2,75	2,62	3,50	3,25	3,12	3,15

Hãy xác định các đặc trưng mẫu.

Đáp số : $n = 8$, $\overline{X} = 3.0775$ *và* $S_X^2 = 0.096$.

Bài 2. Quan sát thời gian cần thiết để sản xuất một chi tiết máy, ta thu được số liệu cho bảng sau :

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

Tính trung bình mẫu $\overline{X}$, phương sai mẫu có hiệu chỉnh S_X^2.

Đáp số : $n = 100$, $\overline{X} = 36.6$ *và* $S_X^2 = 45.14$.

Bài 3. Đo độ dài của một loại trục xe, ta có kết quả

Nhóm	18,4-18,6	18,6-18,8	18,8 -19	19 -19,2	19.2-19.4	19,4-19,6	19,6-19,8
n_i	1	4	20	41	19	8	4

Hãy ước lượng độ dài trung bình và phương sai.

Đáp số : $n = 97$, $\overline{X} = 19.133$ *và* $S_X^2 = 0.054$.

Bài 4. Đo sức bền chịu lực của một loại ống thí nghiệm, người ta thu được bộ số liệu sau

4500	6500	5200	4800	4900	5125	6200	5375

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

Đáp số : $[5151; 5499]$.

Bài 5. Trước bầu cử, người ta phỏng vấn ngẫu nhiên 2000 cử tri thì thấy có 1380 người ủng hộ một ứng cử viên K. Với độ tin cậy 95%, hỏi ứng cử viên đó thu được tối thiểu bao nhiêu phần trăm phiếu bầu ?

Đáp số : 66.97%.

Bài 6. Giả sử quan sát 100 người thấy có 20 người bị bệnh sốt xuất huyết. Hãy ước lượng tỷ lệ bệnh sốt xuất huyết ở độ tin cậy $\gamma = 97\%$. Nếu muốn sai số ước lượng không quá 3% ở độ tin cậy 95% thì phải quan sát ít nhất bao nhiêu người ?

Đáp số : 683.

Bài 7. Để ước lượng xác suất mắc bệnh gan với độ tin cậy 90% và sai số không vượt quá 2% thì cần phải khám ít nhất bao nhiêu người, biết rằng tỷ lệ mắc bệnh gan thực nghiệm đã cho bằng 0,9.

Đáp số : 606.

Bài 8. Muốn biết trong ao có bao nhiêu cá, người ta bắt lên 2000 con, đánh dấu xong lại thả xuống hồ. Sau một thời gian, người ta bắt lên 500 con và thấy có 20 con cá có đánh dấu của lần bắt trước. Dựa vào kết quả đó, Hãy ước lượng số cá có trong hồ với độ tin cậy 95%.

Đáp số : Số cá trong hồ trong khoảng $[34965; 87719]$.

Bài 9. Để có thể dự đoán được số lượng chim thường nghỉ tại vườn nhà mình, người chủ bắt 89 con, đem đeo khoen cho chúng rồi thả đi. Sau một thời gian, ông bắt ngẫu nhiên được 120 con và thấy có 7 con có đeo khoen. Hãy dự đoán số chim giúp ông chủ vườn ở độ tin cậy 99%.

Đáp số : Số chim trong khoảng $[784; 28400]$.

Bài 10. Trên tập mẫu gồm 100 số liệu, người ta tính được $\overline{X} = 0,1; \sigma_{n-1} = 0,014$. Xác định khoảng tin cậy 95% cho giá trị trung bình thật.

Đáp số : $[0.0973; 0.103]$.

Bài 11. Sản lượng mỗi ngày của một phân xưởng là biến ngẫu nhiên tuân theo luật chuẩn. Kết quả thống kê của 9 ngày cho ta :

27	26	21	28	25	30	26	23	26

Hãy xác định các khoảng tin cậy 95% cho sản lượng trung bình và cho phương sai tương ứng.

Đáp số : $\mu \in [23.755; 27.81]$, $\sigma^2 \in [3.17; 25.48]$.

Bài 12. Cân thử 100 quả cam, ta có bộ số liệu sau :

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

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

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

Đáp số : a) $\mu \in [35.539; 36.241]$.

b) $p \in [0.0143; 0.0857]$.

Bài 13. Chiều dài của một loại sản phẩm được xuất khẩu hàng loạt là biến ngẫu nhiên phân phối chuẩn với $\mu = 100mm$ và $\sigma^2 = 4^2 mm^2$. Kiểm tra ngẫu nhiên 25 sản phẩm. Khả năng chiều dài trung bình của số sản phẩm kiểm tra nằm trong khoảng từ 98mm đến 101mm là bao nhiêu.

Đáp số : 88.82%.

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

Đáp số : $\mu \in [375.427; 384.573]$.

Bài 15. Điểm trung bình môn toán của 100 thí sinh dự thi vào KKT là 5 với độ lệch chuẩn là 2.5.

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

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

Đáp số : a) $\mu \in [4.51; 5.49]$.

b) 68.26%.

Bài 16. Tuổi thọ của một loại bóng đèn được biết theo quy luật chuẩn với độ lệch chuẩn 100 giờ.

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

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

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

Đáp số : a) $\mu \in [980.4; 1019.6]$.

b) 86.64%.

c) 62.

Bài 17. Khối lượng các bao bột mì tại một cửa hàng lương thực tuân theo phân phối chuẩn. Kiểm tra 20 bao, thấy khối lượng trung bình của mỗi bao bột mì là 48kg, và phương sai mẫu có hiệu chỉnh là $S_X^2 = (0.5\text{kg})^2$.

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

b) Với độ chính xác của ước lượng ở câu a) là 0.26kg, hãy xác định độ tin cậy.

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

Đáp số : a) $\mu \in [47.766; 48.234]$.

b) 97%.

c) 43.

Bài 18. Để ước lượng tỷ lệ sản phẩm xấu của một kho đồ hộp, người ta kiểm tra ngẫu nhiên 100 hộp thấy có 11 hộp xấu.

a) Ước lượng tỷ lệ sản phẩm xấu của kho đồ hộp với độ tin cậy 94%.

b) Với sai số cho phép của sai số ước lượng $\varepsilon = 3\%$, hãy xác định độ tin cậy.

Đáp số : a) $p \in [0.051; 0.169]$.

b) 66.3%.

Bài 19. Lô trái cây của một chủ cửa hàng được đóng thành sọt mỗi sọt 100 trái. Kiểm tra 50 sọt thấy có 450 trái không đạt tiêu chuẩn.

a) Ước lượng tỷ lệ trái cây không đạt tiêu chuẩn của lô hàng với độ tin cậy 95%.

b) Muốn ước lượng tỷ lệ trái cây không đạt tiêu chuẩn với độ chính xác 0.5%, độ tin cậy đạt được là bao nhiêu.

c) Muốn ước lượng tỷ lệ trái cây không đạt tiêu chuẩn với độ tin cậy 99,7% thì độ chính xác đạt được là bao nhiêu?

d) Muốn ước lượng tỷ lệ trái cây không đạt tiêu chuẩn với độ tin cậy 99% và độ chính xác không quá 1% thì cần phải kiểm tra ít nhất bao nhiêu sọt.

Đáp số : a) $p \in [0.082; 0.098]$.

b) 78.5%, *c)* 1.24%.

d) 55.

Bài 20. Điều tra năng suất lúa trên diện tích 100 hec ta trồng lúa của một vùng, ta thu được bảng số liệu sau :

Năng suất (tạ/ha)	41	44	45	46	48	52	54
Số ha có năng suất tương ứng	10	20	30	15	10	10	5

a) Hãy ước lượng năng suất lúa trung bình của vùng đó với độ tin cậy 95%?

b) Những thửa ruộng có năng suất từ 48tạ/ha trở lên được xem là những thửa có năng suất cao. Hãy ước lượng tỉ lệ diện tích có năng suất cao trong vùng với độ tin cậy 97%.

Đáp số : a) $\mu \in [45.353; 46.647]$.

b) $p \in [0.156; 0.344]$.

Bài 21. Đo đường kính của 100 chi tiết do một máy sản suất kết quả cho ở bảng sau :

Đường kính(mm)	Số chi tiết
19,80 – 19,85	3
19,85 – 19,90	5
19,90 – 19,95	16
19,95 – 20,00	28
20,00 – 20,05	23
20,05 – 20,10	14
20,10 – 20,15	7
20,15 – 20,20	4

Quy định những chi tiết có đường kính 19,9mm đến 20,1mm là những chi tiết đạt tiêu chuẩn.

a) Ước lượng tỷ lệ chi tiết đạt tiêu chuẩn với độ tin cậy 95%.

b) Ước lượng đường kính trung bình của những chi tiết đạt tiêu chuẩn với độ tin cậy 95%.

Đáp số : a) $p \in [0.733; 0.887]$.

b) $\mu \in [19.986; 20.008]$.

Bài 22. Kích thước của một chi tiết máy là một đại lượng ngẫu nhiên có phân phối chuẩn. Trong một mẫu gồm 30 chi tiết máy được kiểm tra, ta tính được $\bar{X} = 0.47\text{cm}$ và $S_X = 0.032\,\text{cm}$. Tìm khoảng tin cậy cho phương sai và trung bình của kích thước của toàn bộ các chi tiết máy với độ tin cậy 95%.

Đáp số : $\mu \in [0.482; 0.458]$, $\sigma^2 \in [0.00065; 0.00185]$.

Bài 23. Lấy 28 mẫu xi măng của một nhà máy sản suất xi măng để kiểm tra. Kết quả kiểm tra về sức chịu lực R (kg/cm^2) như sau:

10.0	13.0	13.7	11.5	11.0	13.5	12.2
13.0	10.0	11.0	13.5	11.5	13.0	12.2
13.5	10.0	10.0	11.5	13.0	13.7	14.0
13.0	13.7	13.0	11.5	10.0	11.0	13.0

Với độ tin cậy 95% hãy ước lượng:

a) Sức chịu lực trung bình của xi măng do nhà máy sản suất,

b) Phương sai của sức chịu lực.

Đáp số : a) $\mu \in [11.62; 12.67]$.

b) $\sigma^2 \in [1.156; 3.427]$.

KIỂM ĐỊNH GIẢ THUYẾT

A. BÀI TẬP MẪU

Bài 1. Có một lô hàng mà người giao hàng cho biết tỷ lệ hỏng 0.10; thứ phẩm 0.30; đạt 0.40; tốt 0.20. Ta kiểm tra một số trường hợp thấy có 25 sản phẩm hỏng; 50 thứ phẩm; 50 sản phẩm đạt; 25 sản phẩm tốt. Hỏi rằng lời người giao hàng nói có đúng không ? (kết luận với $\alpha = 5\%$)

Giải

Ta có bài toán kiểm định

$$\text{a)}\quad \begin{cases} H: & \text{Người giao hàng nói đúng} \\ \overline{H}: & \text{Người giao hàng nói không đúng} \end{cases}$$

Ta có bảng phân phối tần số quan sát

Hỏng	Thứ phẩm	đạt	tốt
25	50	50	25

Nếu H đúng, thì trên tổng số 150 sản phẩm kiểm tra, ta được bảng tần số lý thuyết

Hỏng	Thứ phẩm	đạt	tốt
$0.1 \times 150 = 15$	$0.3 \times 150 = 45$	$0.4 \times 150 = 60$	$0.2 \times 150 = 30$

và khi đó

$$Q = \sum_{i=1}^{4} \frac{(N_i - N_i')^2}{N_i'} \sim \chi^2(3) \qquad (1)$$

với N_i là số liệu quan sát và N_i' là số liệu lý thuyết.

b) Với nguy cơ sai lầm $\alpha = 0.05$, ta được $C = \chi^2_{0.05}(3) = 7.815$

c) Thế các số liệu quan sát và lý thuyết vào biểu thức (1), ta nhận được $Q = 9.7222$.

d) Ta có $Q = 9.7222 > C = 7.815$. Do đó, ta từ chối H, nghĩa là người giao hàng nói không đúng.

Bài 2. Đo cholesterol (đơn vị mg%) cho một nhóm người, ta ghi nhận lại được

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

a) Tính trung bình $\overline{X}$ và phương sai S_X^2

b) Tìm khoảng ước lượng cho trung bình cholesterol trong dân số μ_X, ở độ tin cậy $\gamma = 0.95$

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

Giải

a) Để tính trung bình $\overline{X}$ và phương sai S_X^2, ta lập bảng

Lớp	Tần số	X_i	$Y_i = \frac{X_i - 175}{5}$	$n_i Y_i$	$n_i Y_i^2$
150-160	3	155	-4	-12	48
160-170	9	165	-2	-18	36
170-180	11	175	0	0	0
180-190	3	185	2	6	12
190-200	2	195	4	8	32
200-210	1	205	6	6	36
Tổng cộng	29			-10	164

Từ đó, suy ra

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{k} n_i Y_i = -\frac{10}{29} = -0.3448.$$

Do $\overline{Y} = \frac{\overline{X}-175}{5}$, ta suy ra $\overline{X} = 175 + 5\overline{Y} = 173.276$.

Ngoài ra

$$S_Y^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} n_i Y_i^2 - n\overline{Y}^2\right) = \frac{164 - 29(-0.3448)^2}{28} = 5.734,$$

do $S_Y^2 = \left(\frac{1}{5}\right)^2 S_X^2$, ta có

$$S_X^2 = 25\sigma_Y^2 = 39.7971\text{kg}^2,$$

do đó $S_X = 6.31\text{kg}$. Ta có trung bình mẫu

$$\overline{X} = \frac{1}{29}\sum_{i=1}^{6} n_i x_i = 173.28,$$

và phương sai mẫu là

$$S_X^2 = \frac{1}{28}\left(\sum_{i=1}^{6} n_i x_i^2 - 29\overline{X}^2\right) = 143.35.$$

b) Để ước lượng trung bình tổng thể khi chưa biết phương sai tổng thể, ta dùng thống kê

$$T = \frac{(\overline{X}-\mu)\sqrt{n}}{S_X} \sim St(n-1).$$

Với độ tin cậy $\gamma = 0.95$, từ bảng phân phối Student, ta tìm được $C = t_{0.05}^{28} = 2.048$ sao cho

$$P(-2.048 \le T \le 2.048) = 0.95,$$

thay $T = \dfrac{(\overline{X}-\mu)\sqrt{n}}{S_X}$, ta được

$$P\left(-2.048 \le T = \frac{(\overline{X}-\mu)\sqrt{n}}{S_X} \le 2.048\right) = 0.95.$$

Do đó, ta tìm được khoảng ước lượng cho trung bình tổng thể μ_X là

$$\left[\overline{X} - 2.048\frac{S_X}{\sqrt{n}}; \overline{X} + 2.048\frac{S_X}{\sqrt{n}}\right] = [168.73; 177.83].$$

c) Để so sánh trung bình tổng thể mà ta ước lượng với $\mu_0 = 175$ mg%, ta xét bài toán kiểm định

$$\begin{cases} H: & \text{Giá trị mẫu phù hợp tài liệu} \\ \overline{H}: & \text{Giá trị mẫu không phù hợp tài liệu} \end{cases}$$

Nếu H đúng, nghĩa là $\mu = \mu_0 = 175$, thì

$$T = \frac{\left(\overline{X} - \mu_0\right)\sqrt{n}}{S_x} \sim St(28).$$

Với $\alpha = 0.05$, ta tìm được $C = t_{0.05}^{28} = 2.048$. Từ số liệu của mẫu, ta có

$$T = \frac{(173.28 - 175)\sqrt{29}}{11.97} = -0.774.$$

Vì $|T| \le C$, nên ta chấp nhận H, nghĩa là giá trị mẫu phù hợp với tài liệu.

Bài 3. Quan sát sức nặng của bé trai (X) và bé gái (Y) lúc sơ sinh (đơn vị gam), ta có kết quả

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

a) Tính $\overline{X}$, $\overline{Y}$, S_X^2, S_Y^2.

b) So sánh các phương sai σ_X^2, σ_Y^2 (kết luận với $\alpha = 5\%$).

c) So sánh các trung bình μ_X, μ_Y (kết luận với 5%).

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

Giải

Với số liệu chọn là điểm giữa và với n_1 là số bé trai quan sát, n_2 số bé gái quan sát.

Trọng lượng	3100	3300	3500	3700	3900
n_1	1	3	8	10	3
n_2	2	10	10	5	1
Tổng số	3	13	18	15	4

a) Từ số liệu của mẫu, ta có

$$\overline{X} = 3588,\ \overline{Y} = 3450,$$

$$S_X^2 = \frac{1}{24}\left(\sum_{i=1}^{5} n_{1i}x_i^2 - 24\overline{X}^2\right) = 40266.67,$$

$$S_Y^2 = \frac{1}{27}\left(\sum_{i=1}^{5} n_{2i}y_i^2 - 24\overline{Y}^2\right) = 37407.41.$$

b) Ta xét bài toán kiểm định

$$\begin{cases} H: \sigma_X^2 = \sigma_X^2 \\ \overline{H}: \sigma_X^2 < \sigma_Y^2 \end{cases}$$

Nếu H đúng thì

$$T = \frac{f_X - f_Y}{\sqrt{pq\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim St\left(n + m - 2\right) \equiv St\left(24, 27\right).$$

Với mức ý nghĩa $\alpha = 0.05$ thì $C = f_{0.05} = 1.89$.

Với số liệu ở câu a) , ta có

$$F = \frac{S_X^2}{S_Y^2} = \frac{40266.67}{37407.41} = 1.076.$$

Vì $|F| \le C$, nên ta chấp nhận H, nghĩa là $\sigma_X^2 = \sigma_Y^2$.

c) Ta có bài toán kiểm định

$$\begin{cases} H: \mu_X = \mu_Y \\ \overline{H}: \mu_X \neq \mu_Y \end{cases}$$

Nếu H đúng thì

$$T = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim St(n + m - 2) = St(25 + 28 - 2) = St(51) \equiv N(0;1),$$

trong đó

$$S^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n + m - 2} = \frac{24 \cdot 40266.67 + 27 \cdot 37407.41}{25 + 28 - 2}$$
$$= 38752.94$$

$S = 196.86$;

Với mức ý nghĩa $\alpha = 0.05$ thì $C = Z_{0.475} = 1.96$.

Với số liệu ở câu a), ta có

$$T = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{3588 - 3450}{196.86 \cdot \sqrt{\frac{1}{25} + \frac{1}{28}}} = 2.581.$$

Vì $|T| > C$, nên ta bác bỏ H, nghĩa là $\mu_X \neq \mu_Y$.

d) Nhập hai mẫu lại. Gọi Z là mẫu nhập. Từ bảng số liệu, ta có

$$\overline{Z} = \frac{1}{53}\left(3100 \cdot 3 + 3300 \cdot 13 + 3500 \cdot 18 + 3700 \cdot 15 + 3900 \cdot 4\right)$$
$$= 3515.1,$$

$S_{XY}^2 = 42844.7$,

$S_Z = 206.98$.

Để tìm khoảng tin cậy cho trung bình mẫu nhập Z, ta dùng thống kê

$$T = \frac{\left(\overline{Z} - \mu\right)\sqrt{n}}{S_Z} = \frac{(3515.1 - \mu)\sqrt{53}}{206.98} \sim St(52) \equiv N(0;1).$$

Với độ tin cậy $\gamma = 0.95$ thì $C = 1.96$, ta suy ra

$$\mu_Z = 3515.1 \pm 1.96\frac{206.98}{\sqrt{53}},$$

nghĩa là ta được khoảng ước lượng trung bình của mẫu nhập $[3459.38; 3570.82]$.

Bài 4. Một máy đóng gói các sản phẩm có khối lượng 1kg. Nghi ngờ máy hoạt động không bình thường, người ta chọn ra một mẫu ngẫu nhiên gồm 100 sản phẩm thì thấy như sau :

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

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

Giải

Từ số liệu của mẫu, ta có

Trung bình mẫu : $\overline{X} = 0.9856$,

Phương sai mẫu : $S_X^2 = 0.000433$,

Độ lệch chuẩn mẫu : $S_X = 0.021$,

Cỡ mẫu : $n = 100$.

Xét giả thuyết H : "máy hoạt động bình thường", nghĩa là ta có bài toán kiểm định

$$\begin{cases} H: \mu = 1 \\ \overline{H}: \mu \neq 1 \end{cases}$$

Nếu H đúng thì

$$T = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S_X} \sim St(99) \equiv N(0;1).$$

Với mức ý nghĩa $\alpha = 0.05$, ta có $C = 1.96$. Với số liệu trên, ta được

$$T = \frac{(0.9856 - 1)\sqrt{100}}{0.021} = -6.86.$$

Vì $|T| > C$, nên ta bác bỏ H, nghĩa là máy hoạt động không bình thường

Bài 5. Quan sát số hoa hồng bán ra trong một ngày của một cửa hàng bán hoa sau một thời gian, người ta ghi được số liệu sau :

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

a) Tìm ước lượng điểm của số hoa hồng trung bình bán được trong một ngày.

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

c) Giả sử những ngày bán được từ 13 đến 17 đoá hồng là những ngày "bình thường". Hãy ước lượng tỉ lệ của những ngày bình thường của cửa hàng ở độ tin cậy 90%. (Giả thiết rằng số hoa bán ra trong ngày có phân phối chuẩn).

Giải

a) Trung bình $\overline{X} = 15.4$, $S_X = 1.871$, $n = 25$.

b) Xét giả thiết H : "nên bán tiếp", ta có bài toán kiểm định

$$\begin{cases} H: \mu = 15 \\ \overline{H}: \mu \neq 15 \end{cases}$$

Nếu H đúng thì

$$T = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S_X} \sim St(24).$$

Từ số liệu của câu a, ta có

$$T = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S_X} = \frac{(15.4 - 15).\sqrt{25}}{1.871} = 1.07.$$

Với mức ý nghĩa $\alpha = 0.05$ thì $C = t_{0.05}^{24} = 2.064$.

Vì $|T| \leq C$, nên ta chấp nhận H, nghĩa là ông chủ nên tiếp tục bán.

c) Ta có tỷ lệ mẫu những ngày bình thường là

$$f = \frac{2+7+7+3}{25} = 0.76.$$

Để ước lượng tỷ lệ tổng thể, ta dùng thống kê

$$Z = \frac{(f-p)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1) = St(24).$$

Với độ tin cậy 90%, ta được $C = 1.711$. Suy ra

$$p = f \pm C.\sqrt{\frac{f(1-f)}{n}} = 0.76 \pm 1.711\sqrt{\frac{0.76(1-0.76)}{25}} = 0.76 \pm 0.146.$$

Vậy, khoảng ước lượng là $[0.614; 0.906]$.

Bài 6. Một xí nghiệp đúc một số rất lớn các sản phẩm bằng thép với số khuyết tật trung bình ở mỗi sản phẩm là 3. Người ta cải tiến cách sản xuất và kiểm tra 36 sản phẩm. Kết quả như sau :

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

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

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

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

Giải

a) Ta có

$$\overline{X} = 2.64,\ S_X^2 = 3.38,\ S_x = 1.838,\ n = 36.$$

Để ước lượng số khuyết tật trung bình, ta dùng thống kê

$$T = \frac{(\overline{X} - \mu)\sqrt{n}}{S_X} \sim St(n-1) = St(35) \equiv N(0;1).$$

Độ tin cậy $\gamma = 0.9$ thì $C = 1.645$. Do đó ta có khoảng ước lượng

$$\left[\overline{X} - C.\frac{S_X}{\sqrt{n}}; \overline{X} + C.\frac{S_X}{\sqrt{n}}\right] = [2.136; 3.144].$$

b) Xét giả thiết H : "cải tiến không hiệu quả", ta có bài toán kiểm định

$$\begin{cases} H : \mu_0 = 3 \\ \overline{H} : \mu_0 \neq 3 \end{cases}$$

Nếu H đúng thì

$$T = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S_X} \sim St(35) \sim N(0;1).$$

Với mức ý nghĩa $\alpha = 0.05$ thì $C = 1.96$, và

$$T = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S_X} = \frac{(2.64 - 3).\sqrt{36}}{1.838} = -1.175.$$

Vì $|T| \leq C$, chấp nhận H, nghĩa là cải tiến không hiệu quả.

Bài 7. Quan sát ngẫu nhiên một số trường hợp trong 3 lô thuốc (rất nhiều), ta ghi nhận được

	Tốt	Tạm dùng	Hỏng
Lô A	125	52	23
Lô B	117	61	22
Lô C	178	97	25

Hỏi rằng chất lượng của 3 lô thuốc có như nhau không ? Kết luận với $\alpha = 0.05$.

Giải

Ta có

	Tốt	Tạm dùng	Hỏng	Tổng
Lô A	125	52	23	***200***
Lô B	117	61	22	***200***
Lô C	178	97	25	***300***
Tổng	***420***	***210***	***70***	**700**

Xét bài toán kiểm định

$$\begin{cases} H: & \text{Chất lượng 3 lô thuốc như nhau} \\ \overline{H}: & \text{Chất lượng 3 lô thuốc khác nhau} \end{cases}$$

Nếu H đúng thì

$$P(\text{tốt}) = \frac{125 + 117 + 178}{700} = 0.6,$$

$$P(\text{tạm}) = \frac{52 + 61 + 97}{700} = 0.3, \text{ và}$$

$$P(\text{hỏng}) = \frac{23 + 22 + 25}{700} = 0.1.$$

Khi đó, bảng phân phối tần số lý thuyết phải là

	Tốt	Tạm dùng	Hỏng
Lô A	$0.6 \times 200 = 120$	$0.3 \times 200 = 60$	$0.1 \times 200 = 20$
Lô B	$0.6 \times 200 = 120$	$0.3 \times 200 = 60$	$0.1 \times 200 = 20$
Lô C	$0.6 \times 300 = 180$	$0.3 \times 300 = 90$	$0.1 \times 300 = 30$

Độ khác biệt giữa quan sát và lý thuyết là

$$Q = \sum \frac{(N - N')^2}{N'} = \frac{(125 - 120)^2}{120} + \frac{64}{60} + \frac{9}{20} +$$
$$+ \frac{9}{120} + \frac{1}{60} + \frac{4}{20} + \frac{4}{180} + \frac{49}{90} + \frac{25}{30} = 3.42.$$

Nếu H đúng thì $Q \sim \chi^2(3-1)(3-1) = \chi^2(4)$, với mức ý nghĩa $\alpha = 0.05$, ta có $C = \chi^2_{0.05} = 9.488$.

Vì $|Q| \le C$, ta chấp nhận H, nghĩa là 3 lô thuốc như nhau.

Chú ý : Ta có thể thành lập trực tiếp bảng phân phối tần số lý thuyết như sau

	Tốt	Tạm dùng	Hỏng	Tổng
Lô A	$\frac{420 \times 200}{700}$	$\frac{210 \times 200}{700}$	$\frac{70 \times 200}{700}$	***200***
Lô B	$\frac{420 \times 200}{700}$	$\frac{210 \times 200}{700}$	$\frac{70 \times 200}{700}$	***200***
Lô C	$\frac{420 \times 300}{700}$	$\frac{210 \times 300}{700}$	$\frac{70 \times 300}{700}$	***300***
Tổng	***420***	***210***	***70***	**700**

Hơn nữa, ta có thể dùng trực tiếp công thức

$$Q = \mathbf{700}\left(\frac{(125)^2}{\mathbf{200} \times 420} + \frac{(52)^2}{200 \times 210} + \ldots + \frac{(25)^2}{70 \times 300} - 1\right).$$

Bài 8. Trong một công ty, người ta chọn ngẫu nhiên 1000 công nhân và theo dõi số ngày nghỉ của họ trong một năm. Kết quả thu được :

Giới tính / Số ngày nghỉ	Nữ	nam
0 – 5	300	500
5 – 20	80	70
> 20	20	30

Với mức ý nghĩa 0.01, hãy kiểm định giả thiết cho rằng sự nghỉ việc không phụ thuộc vào giới tính.

Giải

Ta có bài toán kiểm định

$$\begin{cases} H: & \text{Sự nghỉ việc không phụ thuộc vào giới tính} \\ \overline{H}: & \text{Sự nghỉ việc phụ thuộc vào giới tính} \end{cases}$$

Nếu H đúng thì

$$P(\text{nữ}) = \frac{300 + 80 + 20}{1000} = 0.4,$$

$$P(\text{nam}) = \frac{500 + 70 + 30}{1000} = 0.6.$$

Khi đó, bảng phân phối tần số lý thuyết là

Giới tính / Số ngày nghỉ	Nữ	nam
0 – 5	320	480
5 – 20	60	150
> 20	20	50

Độ khác biệt giữa quan sát và lý thuyết là

$$Q = \sum \frac{(N - N')^2}{N'} = \frac{400}{320} + \frac{400}{480} + \frac{400}{60} + \frac{400}{90} = 13.194.$$

Nếu H đúng thì $Q \sim \chi^2(3-1)(2-1) = \chi^2(2)$. Với mức ý nghĩa $\alpha = 0.01$, ta có $C = \chi^2_{0.01} = 9.21$.

Vì $|Q| > C$, nên ta bác bỏ H, nghĩa là sự nghỉ việc phụ thuộc vào giới tính.

Bài 9. Nghiên cứu ảnh hưởng của hoàn cảnh gia đình đối với tình hình phạm tội của trẻ em vị thành niên, người ta thu được.

Hoàn cảnh gia đình / Tình trạng phạm tội	Bố hoặc mẹ đã chết	Bố mẹ ly hôn	Còn cả bố mẹ
Không phạm tội	20	25	13
Phạm tội	29	43	18

Với mức ý nghĩa $\alpha = 0.05$, có thể kết luận là hoàn cảnh gia đình của trẻ em độc lập với tình trạng phạm tội hay không.

Giải

Gọi X : Bố hoặc mẹ đã chết, Y : Bố mẹ ly hôn, Z : còn cả bố mẹ.

Ta có bài toán kiểm định

$$\begin{cases} H: & \text{Hoàn cảnh gia đình độc lập với tình trạng xã hội} \\ \overline{H}: & \text{Hoàn cảnh gia đình không độc lập với tình trạng xã hội} \end{cases}$$

Nếu H đúng thì

$$P(X) = \frac{20+29}{148} = 0.331,$$

$$P(Y) = \frac{25+43}{148} = 0.459,$$

$$P(Z) = \frac{13+18}{148} = 0.209.$$

Ta có bảng phân phối tần số lý thuyết như sau

Hoàn cảnh gia đình / Tình trạng phạm tội	X	Y	Z
Không phạm tội	19	27	12
Phạm tội	30	41	19

Độ khác biệt giữa quan sát và lý thuyết là

$$Q = \sum \frac{(N - N')^2}{N'} = \frac{1}{19} + \frac{4}{27} + \frac{1}{12} + \frac{1}{30} + \frac{4}{41} + \frac{1}{19}$$
$$= 0.468.$$

Nếu H đúng thì $Q \sim \chi^2(2-1)(3-1) = \chi^2(2)$, với mức ý nghĩa $\alpha = 0.05$, ta có $C = \chi^2_{0.05} = 5.991$.

Vì $|Q| \le C$, nên ta chấp nhận H, nghĩa là hoàn cảnh gia đình độc lập với tình trạng phạm tội.

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

Giải

Đặt

p_1 : Tỷ lệ người mắc bệnh dùng DDT

p_2 : Tỷ lệ người mắc bệnh không dùng DDT

Ta có bài toán kiểm định

$$\begin{cases} H: p_1 = p_2 \\ \overline{H}: p_1 \neq p_2 \end{cases}$$

Vì $f_1 = \dfrac{10}{90} = \dfrac{1}{9}$, và $f_2 = \dfrac{26}{100} = 0.26$, nên ta có

$$p = \frac{nf_1 + mf_2}{n+m} = \frac{90.\frac{1}{9} + 100.\frac{26}{100}}{90+100} = 0.1895 .$$

Nếu H đúng thì

$$T = \frac{f_1 - f_2}{\sqrt{pq\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim St(n+m-2) = St(188) \equiv N(0;1).$$

Với mức ý nghĩa $\alpha = 0.05$ thì $C = 1.96$, và do đó

$$T = \frac{f_1 - f_2}{\sqrt{p(1-p)\left(\frac{1}{n} + \frac{1}{m}\right)}} = \frac{\frac{1}{9} - 0.26}{\sqrt{0.1895 \cdot 0.8105\left(\frac{1}{90} + \frac{1}{100}\right)}} = -2.616 .$$

Vì $|T| > C$, ta bác bỏ H, nghĩa là người dùng DDT có tác dụng ngừa bệnh ngoài da

Bài 11. Trong một vùng dân cư có 18 bé trai và 28 bé gái mắc bệnh B. Hỏi rằng tỷ lệ nhiễm bệnh của bé trai và bé gái có như nhau không ? (kết luận với $\alpha = 0.05$ và giả sử rằng số lượng bé trai và bé gái trong vùng tương đương nhau, và rất nhiều).

Giải

Ta có bài toán kiểm định

$$\begin{cases} H: p = \frac{1}{2} \\ \overline{H}: p \neq \frac{1}{2} \end{cases}$$

Nếu H đúng thì $Z = \dfrac{(f-p)\sqrt{n}}{\sqrt{pq}} \sim N(0;1)$.

Vì $f = \dfrac{18}{18+28} = 0.391$ nên ta có

$$Z = \frac{(0.391 - 0.5)\sqrt{46}}{\sqrt{0.5 \cdot 0.5}} = -1.48 .$$

Với mức ý nghĩa $\alpha = 0.05$, ta tìm được $C = 1.96$.

Vì $|Z| \le C$, nên ta chấp nhận H, nghĩa là tỷ lệ mắc bệnh B của bé trai và bé gái là như nhau.

Bài 12. Thống kê số tai nạn lao động tại 2 xí nghiệp, ta có các số liệu sau :

Xí nghiệp	Số công nhân	Số tai nạn lao động
A	200	20
B	800	120

Với mức ý nghĩa $\alpha = 0.05$, hãy kết luận xem chất lượng công tác bảo vệ an toàn lao động tại 2 xí nghiệp trên có khác nhau không ?

Giải

Ta có bài toán kiểm định

$$\begin{cases} H: & \text{Chất lượng bảo vệ an toàn của hai xí nghiệp như nhau} \\ \overline{H}: & \text{Chất lượng bảo vệ an toàn của hai xí nghiệp khác nhau} \end{cases}$$

Nếu H đúng thì

$$P(\text{Công nhân}) = \frac{200+800}{1140} = 0.8772$$

$$P(\text{tai nạn}) = \frac{20+120}{1140} = 0.123.$$

Ta có bảng phân phối tần số lý thuyết

Xí nghiệp	Số công nhân	Số tai nạn lao động
A	193	27
B	807	113

Độ khác biệt giữa quan sát và lý thuyết là

$$Q = \sum \frac{(|N - N'| - 0.5)^2}{N'} =$$

$$= \frac{42.25}{193} + \frac{42.25}{27} + \frac{42.25}{807} + \frac{42.25}{113} = 2.21.$$

Nếu H đúng thì $Q \sim \chi^2(2-1)(2-1) = \chi^2(1)$, với mức ý nghĩa $\alpha = 0.05$, ta có $C = \chi^2_{0.05} = 3.841$.

Vì $|Q| \le C$, nên ta chấp nhận H, nghĩa là chất lượng bảo vệ an toàn lao động của hai xí nghiệp là như nhau.

Bài 13. Đối với người Việt Nam, lượng huyết sắc tố trung bình là 138.3g/l. Khám cho 80 công nhân ở nhà máy có tiếp xúc hoá chất, thấy huyết sắc tố trung bình là 120g/l; $S = 15\text{g/l}$. Từ kết quả trên, có thể kết luận lượng huyết sắc tố trung bình của công nhân nhà máy hoá chất này thấp hơn mức chung hay không ? Kết luận với $\alpha = 0.05$

Giải

Ta có bài toán kiểm định

$$\begin{cases} H: \mu = 138.3 \\ \overline{H}: \mu \ne 138.3 \end{cases}$$

Theo giả thiết, ta có $\overline{X} = 120$, $S_X = 15$, và $n = 80$.

Nếu H đúng thì

$$T \equiv \frac{\left(\overline{X}-\mu_0\right)\sqrt{n}}{S_X} \sim St(n-1) = St(79) \equiv N(0;1).$$

Từ số liệu của mẫu ta tìm được giá trị của T là

$$T = \frac{\left(\overline{X}-138.3\right)\sqrt{n}}{S_X} = \frac{(120-138.3)\sqrt{80}}{15} = -10.91.$$

Với mức ý nghĩa $\alpha = 0.05$, ta có $C = 1.96$.

Vì $|T| > C$, nên ta bác bỏ H, nghĩa là lượng huyết tố trung bình của công nhân nhà máy thấp hơn mức chung.

Bài 14. Hàm lượng đường trong máu của công nhân sau 5 giờ làm việc với máy siêu cao tần đã đo được ở hai thời điểm trước và sau 5 giờ làm việc. Ta có kết quả sau :

Trước $n_1 = 50$, thì $\overline{X} = 60mg\%$, $S_X = 7$.

Sau $n_2 = 40$, thì $\overline{Y} = 52mg\%$, $S_Y = 9.2$.

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

Giải

Ta có bài toán kiểm định

$$\begin{cases} H: \mu_X = \mu_Y \\ \overline{H}: \mu_X \neq \mu_Y \end{cases}$$

Theo giả thiết, ta có

$$S^2 = \frac{(n_1-1)S_X^2 + (n_2-1)S_Y^2}{n_1+n_2-2} = \frac{49\cdot 49 + 39\cdot 84.64}{50+40-2}$$
$$= 64.795,$$

do đó $S = 8.05$.

Nếu H đúng thì

$$T = \frac{\overline{X}-\overline{Y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim St(n_1+n_2-2) = St(89) \equiv N(0;1).$$

Từ số liệu của hai mẫu, ta tính được giá trị của T là

$$T = \frac{\overline{X}-\overline{Y}}{S\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} = \frac{60-52}{8.05\sqrt{\dfrac{1}{50}+\dfrac{1}{40}}} = 4.68$$

Với mức ý nghĩa $\alpha = 0.05$, ta có $C = 1.96$.

Vì $|T| > C$: nên ta bác bỏ H, nghĩa là hàm lượng đường trong máu sau 5 giờ làm việc đã giảm đi.

Bài 15. Đánh giá tác dụng của một chế độ ăn bồi dưỡng mà dấu hiệu quan sát là số hồng cầu. Người ta đếm số hồng cầu của 20 người trước và sau khi ăn bồi dưỡng :

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

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

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

Giải

Đây là trường hợp dãy số liệu từng cặp. Ta không thể căn cứ trên tác dụng trung bình của từng dãy số để so sánh mà ta phải căn cứ trên sự thay đổi từng cá thể. Đặt $d = Y - X$ để chỉ số lượng gia tăng bồi bổ. Ta có bảng hiệu số $d_i = X_i - Y_i$ với $i = 1, 2, \ldots, 20$ như sau

d	8	5	4	8	11	8	12	-2	5	4

-6	9	15	-6	10	-3	8	7	-5	-4

Từ bảng trên, ta tính được

$$\overline{d} = 6.9,\ S_d = 4.28,\ \text{và}\ n = 20.$$

Khi đó, giả thiết H : "Hai bộ số liệu giống nhau từng cặp"

được thay bằng

H : "trung bình của bộ số liệu D_i bằng 0".

Ta có bài toán kiểm định

$$\begin{cases} H: \mu_d = 0 \\ \overline{H}: \mu_d \neq 0 \end{cases}$$

Nếu H đúng thì

$$T \equiv \frac{(\overline{d} - 0)\sqrt{n}}{S_d} \sim St(n-1) = St(19).$$

Từ đó, ta tìm được giá trị của T là

$$T = \frac{(\overline{d} - 0)\sqrt{n}}{S_d} = \frac{(6.9 - 0)\sqrt{20}}{4.28} = 7.21.$$

Với mức ý nghĩa $\alpha = 0.05$ thì $C = t_{0.05}^{19} = 2.093$.

Vì $|T| > C$, nên ta bác bỏ H, nghĩa là chế độ thức ăn bồi dưỡng làm thay đổi hồng cầu.

Bài 16. Trong đợt thi đua, phân xưởng A báo cáo chất lượng sản phẩm làm ra như sau : có 85% loại 1; 10% loại 2 và 5% loại 3. Ban thi đua đã lấy ngẫu nhiên từ lô sản phẩm chưa phân loại của phân xưởng A ra 100 sản phẩm, thấy có 80 loại 1, 13 loại 2 và 7 loại 3. Với mức ý nghĩa $\alpha = 0.01$, có thể kết luận gì về báo cáo của phân xưởng A ?

Giải

Bảng số liệu quan sát của phân xưởng A

	Loại 1	Loại 2	Loại 3
Sản phẩm	80	13	7
Tỉ lệ	0.85	0.1	0.05

Ta có bảng phân phối tần số lý thuyết

	Loại 1	Loại 2	Loại 3
Sản phẩm	85	10	5
Tỉ lệ	0.85	0.1	0.05

Độ khác biệt giữa quan sát và lý thuyết là

$$Q = \sum \frac{(N - N')^2}{N'} = \frac{(80-85)^2}{85} + \frac{(13-10)^2}{10} + \frac{(7-5)^2}{5} = 1.99.$$

Và ta có bài toán kiểm định

$$\begin{cases} H: & \text{Dự báo của phân xưởng A là đúng} \\ \overline{H}: & \text{Dự báo của phân xưởng A là không đúng} \end{cases}$$

Nếu H đúng thì $Q \sim \chi^2(2)$.

Với mức ý nghĩa $\alpha = 0.01$ thì $C = \chi^2_{0.01}(2) = 9.21$.

Vì $Q \le C$, nên ta chấp nhận H, nghĩa là dự báo của phân xưởng A là đúng.

B. BÀI TẬP

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

Đáp số : $T = -3$, *bác bỏ giả thiết.*

Bài 2. Một máy sản xuất tự động với tỷ lệ chính phẩm 98%. Sau một thời gian hoạt động, người ta nghi ngờ tỷ lệ trên đã bị giảm. Kiểm tra ngẫu nhiên 500 sản phẩm thấy có 28 phế phẩm, với $\alpha = 0.05$ hãy kiểm tra xem chất lượng làm việc của máy có còn được như trước hay không?

Đáp số : $Z = -5.75$, *bác bỏ giả thiết.*

Bài 3. Trồng cùng một giống lúa trên hai thửa ruộng như nhau và bón hai loại phân khác nhau. Đến ngày thu hoạch ta có kết quả như sau : Thửa thứ nhất lấy mẫu 1000 bông lúa thấy số hạt trung bình của mỗi bông $\overline{X} = 70$ hạt và $S_X = 10$. Thửa thứ hai lấy mẫu 500 bông thấy số hạt trung bình mỗi bông $\overline{Y} = 72$ hạt và $S_Y = 20$. Hỏi sự khác nhau giữa $\overline{X}$ và $\overline{Y}$ là ngẫu nhiên hay bản chất, với $\alpha = 0.05$?

Đáp số : $T = -2.58$, *bác bỏ giả thiết.*

Bài 4. Để so sánh trọng lượng trung bình của trẻ sơ sinh ở thành thị và nông thôn, người ta thử cân trọng lượng của 10000 cháu và thu được kết quả sau đây :

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

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

Đáp số : $T = -28.28$, *bác bỏ giả thiết.*

Bài 5. Số con của 2000 phụ nữ thủ đô dưới 25 tuổi cho ở bảng 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 luật Poisson hay không ?

Đáp số : $Q = 8.01$, *bác bỏ giả thiết.*

Bài 6. Giả sử ta muốn xác định xem hiệu quả của chế độ ăn kiêng đối với việc giảm trọng lượng như thế nào. 20 người quá béo đã thực hiện chế độ ăn kiêng. Trọng lượng của từng người trước khi ăn kiêng (Xkg) và sau khi ăn kiêng (Ykg) được cho như sau:

X	80	78	85	70	90	78	92	88	75
Y	75	77	80	70	84	74	85	82	80

X	75	63	72	89	76	77	71	83	78	82	90
Y	65	62	71	83	72	82	71	79	76	83	81

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

Đáp số : $T = -3.39$, *bác bỏ giả thiết.*

Bài 7. Dùng 3 phương án xử lý hạt giống kết quả cho như sau :

Kết quả	Phương án I	Phương án II	Phương án III
Số hạt mọc	360	603	490
Số hạt không mọc	40	97	180

Theo bảng số liệu ở trên, các phương án xử lý có tác dụng như nhau đối với tỷ lệ nảy mầm hay không ($\alpha = 0.05$) ?

Đáp số : $Q = 61.52$, *bác bỏ giả thiết.*

Bài 8. Theo dõi sự phụ thuộc giữa màu mắt và màu tóc ở 124 phụ nữ ở một nước Châu Âu ta có kết quả sau :

Màu tóc / Màu mắt	Vàng nâu	Nâu	Đen	Vàng hoe
Xanh	25	9	3	7
Xám	13	17	10	7
Nâu mực	7	13	8	5

Với $\alpha = 0.05$, hãy kiểm tra giả thiết cho rằng màu của tóc và màu của mắt độc lập với nhau.

Đáp số : $Q = 15.07$, *bác bỏ giả thiết.*

Bài 9. Để xác định thời vụ phun thuốc diệt sâu có lợi nhất, tổ bảo vệ cây trồng đã theo dõi các lứa sâu trong từng thời kỳ và đếm số sâu non mới nở bắt được. Kết quả ghi ở bảng sau

Thời kỳ theo dõi	Tháng 1	Tháng 2	Tháng 3	Tháng 4	Tháng 5
Số sâu non mới nở bắt được	62	28	70	75	15
Tổng số sâu non bắt được	488	392	280	515	185

Tỷ lệ sâu non mới nở trong các thời kỳ quan sát khác nhau có ý nghĩa hay không ($\alpha = 0.05$) ?

Đáp số : $Q = 50.83$, *bác bỏ giả thiết.*

Bài 10. Đo huyết sắc tố cho 50 công nhân nông trường thấy có 60% ở mức dưới 110g/l. Số liệu chung của khu vực này là 30% ở mức dưới 110g/l. Với mức ý nghĩa $\alpha = 0.05$, có thể kết luận công nhân nông trường có tỷ lệ huyết sắc tố dưới 110g/l cao hơn mức chung hay không ?

Đáp số : $Z = 4.63$, *bác bỏ giả thiết.*

Bài 11. Hàm lượng đường trong máu của công nhân sau 3 giờ làm việc với máy siêu cao tần đã được đo ở 2 thời điểm trước và sau 3 giờ làm việc. Ta có kết quả sau :

Trước : $n_1 = 50$: $\overline{X} = 60$mg%; $S_1 = 7$

Sau : $n_2 = 40$: $\overline{Y} = 52$mg%; $S_2 = 9.2$

Với mức ý nghĩa $\alpha = 0.05$, có thể khẳng định hàm lượng đường trong máu sau 3 giờ làm việc đã giảm đi hay không ?

Đáp số : $T = 4.69$, *bác bỏ giả thiết.*

Bài 12. Gọi X là số người tới một trạm điện thoại trong thời gian 3 phút. Theo dõi 50 khoảng thời gian như vậy ta có các số liệu sau:

Số người đến (X)	0	1	2	3	4	5	6
Số khoảng xảy ra	8	15	12	9	4	1	1

Với mức ý nghĩa $\alpha = 0.05$ có thể kết luận X tuân theo luật phân phối Poisson hay không ?

Đáp số : $Q = 1.06$, *chấp nhận giả thiết.*

Bài 13. Một nhà máy có 3 phân xưởng cùng sản xuất một loại sản phẩm. Chất lượng sản phẩm được chia thành 3 loại. Kiểm tra, phân loại ngẫu nhiên một số sản phẩm từ lô sản phẩm của 3 phân xưởng ta có số liệu sau :

Phân xưởng / Chất lượng	PX I	PX II	PX III
Loại I	70	80	60
Loại II	25	20	15
Loại III	5	10	5

Với $\alpha = 0.05$ có thể kết luận chất lượng sản phẩm phụ thuộc vào nơi làm ra chúng hay không ?

Đáp số : $Q = 2.8$, *chấp nhận giả thiết.*

Bài 14. Giám đốc một xí nghiệp cho biết lương trung bình của 1 công nhân thuộc xí nghiệp là 380 ngàn đ/tháng. Chọn ngẫu nhiên 36 công nhân thấy lương trung bình là 350 ngàn đ/tháng, với độ lệch chuẩn $S = 40$. Lời báo cáo của giám đốc có tin cậy được không, với mức có ý nghĩa là $\alpha = 5\%$.

Đáp số : $T = -4.5$, *bác bỏ giả thiết.*

Bài 15. Gieo đồng thời 2 đồng tiền 50 lần. Tần số xuất hiện số mặt xấp được cho như sau :

Số mặt xấp	0	1	2
Tần số xuất hiện	10	28	12

Với mức ý nghĩa $\alpha = 0.05$ có thể kết luận 2 đồng tiền là cân đối và đồng chất hay không ?

Đáp số : $Q = 0.88$, *chấp nhận giả thiết.*

Bài 16. Trong thập niên 80, trọng lượng trung bình của thanh niên là 48kg. Nay để xác định lại trọng lượng ấy, người ta chọn ngẫu nhiên 100 thanh niên đo trọng lượng trung bình là 50kg và phương sai mẫu điều chỉnh $S^2 = (10\text{kg})^2$. Thử xem trọng lượng thanh niên hiện nay phải chăng có thay đổi, với mức có ý nghĩa là 1% ?

Đáp số : $T = 2$, *chấp nhận giả thiết.*

Bài 17. Một cửa hàng thực phẩm nhận thấy thời gian vừa qua trung bình một khách hàng mua 25 ngàn đồng thực phẩm trong ngày. Nay cửa hàng chọn ngẫu nhiên 15 khách hàng thấy trung bình một khách hàng mau 24 ngàn đồng trong ngày và phương sai mẫu điều chỉnh là $S^2 = (2 \text{ ngàn đồng})^2$.

Với mức ý nghĩa là 5%, thử xem có phải sức mua của khách hàng hiện nay có thực sự giảm sút.

Đáp số : $T = -1.94$*, chấp nhận giả thiết.*

Bài 18. Theo một nguồn tin thì tỉ lệ hộ dân thích xem dân ca trên Tivi là 80%. Thăm dò 36 hộ dân thấy có 25 hộ thích xem dân ca. Với mức có ý nghĩa là 5%. Kiểm định xem nguồn tin này có đáng tin cậy không ?

Đáp số : $Z = -1.58$*, chấp nhận giả thiết.*

Bài 19. Một máy sản suất tự động, lúc đầu tỷ lệ sản phẩm loại A là 20%. Sau khi áp dụng một phương pháp cải tiến sản xuất mới, người ta lấy 40 mẫu, mỗi mẫu gồm 10 sản phẩm để kiểm tra. Kết quả kiểm tra cho ở bảng sau :

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

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

Đáp số : $Z = 16.875$*, bác bỏ giả thiết.*

Bài 20. Trọng lượng trung bình khi xuất chuồng ở một trại chăn nuôi trước là 3,3 kg/con. Năm nay người ta sử dụng một loại thức ăn mới, cân thử 15 con khi xuất chuồng ta được các số liệu như sau:

3,25 ; 2,50 ; 4,00 ; 3,75 ; 3,80 ; 3,90 ; 4,02 ; 3,60 ;3,80 ; 3,20 ; 3,82 ; 3,40 ; 3,75 ; 4,00 ; 3,50

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

a) Với mức ý nghĩa $\alpha = 0,05$. Hãy cho kết luận về tác dụng của loại thức ăn này ?

b) Nếu trại chăn nuôi báo cáo trọng lượng trung bình khi xuất chuồng là 3,5 kg/con thì có chấp nhận được không ? $(\alpha = 5\%)$.

Đáp số : $T = 3.06$*, bác bỏ giả thiết.*

Bài 21. Tỷ lệ phế phẩm của một nhà máy trước đây là 5%. Năm nay nhà máy áp dụng một biện pháp kỹ thuật mới. Để nghiên cứu tác dụng của biện pháp kỹ thuật mới, người ta lấy một mẫu gồm 800 sản phẩm để kiểm tra và thấy có 24 phế phẩm.

a) Với $\alpha = 0,01$. Hãy cho kết luận về biện pháp kỹ thuật mới này ?

b) Nếu nhà máy báo cáo tỷ lệ phế phẩm sau khi áp dụng biện pháp kỹ thuật mới là 2% thì có chấp nhận được không ? $(\text{với } \alpha = 0,01)$.

Đáp số : a) $Z = -2.6$*, bác bỏ giả thiết.*
b) $T = 2.02$*, chấp nhận giả thiết.*

Bài 22. Tiền lương trung bình của công nhân trước đây là 400 ngàn đ/tháng. Để xét xem tiền lương hiện nay so với mức trước đây thế nào, người ta điều tra 100 công nhân và tính được $\overline{X} = 404.8$ ngàn đ/tháng và $S = 20$ ngàn đ/tháng. Với $\alpha = 1\%$

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

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

Đáp số : a) $T = 2.4$*, chấp nhận giả thiết.*

b) $T = 3$*, bác bỏ giả thiết.*

Bài 23. Sản phẩm được sản xuất ra trên một dây chuyền tự động được đóng gói một cách ngẫu nhiên theo qui cách : 3 sản phẩm/hộp. Tiến hành kiểm tra 200 hộp ta được kết quả

Số sp loại I có trong hộp	0	1	2	3
Số hộp	6	14	110	70

Với $\alpha = 2\%$, có thể xem số sản phẩm loại I có trong hộp là đại lượng ngẫu nhiên có quy luật phân phối nhị thức không ?

Đáp số : $Q = 18.88$*, bác bỏ giả thiết.*

Bài 24. Một nhà máy sản xuất máy in nói rằng số lỗi in trong 1 cuốn sách dày 300 trang của máy in là 1 đại lượng ngẫu nhiên có quy luật phân phối Poisson với tham số $\lambda = 4.7$. Kiểm tra 300 trang sách in của 50 máy in cùng loại, ta được

Số lỗi	0	1	2	3	4	5	6	7	8	≥ 9
Số máy	1	1	8	6	13	10	5	5	1	0

Với mức ý nghĩa 1%, hỏi lời tuyên bố của nhà sản xuất có đúng không ?

Đáp số : $Q = 2.406$*, chấp nhận giả thiết.*

Bài 25. Kiểm tra 200 thùng một loại đồ hộp, người ta thu được số liệu sau

Số hộp bị hỏng/thùng	0	1	2	3	4
Số thùng	116	56	22	4	2

Với $\alpha = 5\%$, số hộp bị hỏng của một thùng có là biến ngẫu nhiên tuân theo qui luật Poisson ?

Đáp số : $Q = 2.393$*, chấp nhận giả thiết.*

Bài 26. Số tai nạn giao thông xảy ra mỗi ngày ở 1 thành phố quan sát được

Số tai nạn	0	1	2	3	4	5	6	7	8
Số ngày	10	32	46	35	20	9	2	1	1

Với mức ý nghĩa 1%, xét xem số tai nạn giao thông có quy luật Poisson ?

Đáp số : $Q = 2.311$*, chấp nhận giả thiết.*

Bài 27. Năng suất lúa thử nghiệm trên 100 lô đất cho kết quả

Năng suất(tấn/ha)	8-9	9-10	10-11	11-12	12-13	13-14	14-15
Số trường hợp	8	15	21	23	16	9	8

Với mức ý nghĩa 1%, xét xem năng suất lúa có tuân theo quy luật phân phối chuẩn không ?

Đáp số : $Q = 4.4$*, chấp nhận giả thiết.*

Bài 28. Gieo 1 con xúc xắc 600 lần. Số lần xuất hiện các mặt 1, 2, 3, 4, 5, 6 được cho trong bảng sau

Số nút	1	2	3	4	5	6
Số lần xuất hiện	106	92	97	105	88	112

Với mức ý nghĩa 5%, có thể xem con xúc xắc được chế tạo cân đối, đồng chất không ?

Đáp số : $Q = 4.2$*, chấp nhận giả thiết.*

PHỤ LỤC

1. Dùng các bảng phân phối xác suất

Các bảng phân phối xác suất quan trọng gồm phân phối Gauss, Chi-Bình phương, Student và Fisher. Các giá trị xác suất đặc biệt của chúng được tính sẵn và liệt kê thành bảng như sau

1.1. Phân phối Gauss $N(0,1)$.

Với $X \sim N(0,1)$, ta có hai bài toán xác suất quan trọng :

- tìm $P(a \le X \le b)$, với $a, b \in \mathbb{R}$, $a \le b$ cho trước,

- tìm giá trị C sao cho $P(-C \le X \le C) = P(|X| \le C) = \gamma$, với γ cho trước.

1.1.1. Tìm $P(a \le X \le b)$.

Do $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$ là hàm mật độ của X nên từ tính chất của tích phân, ta có

$$\begin{aligned} P(a \le X \le b) &= \int_a^b f(x)dx = \int_0^b f(x)dx - \int_0^a f(x)dx \\ &\equiv \varphi(b) - \varphi(a), \end{aligned}$$

trong đó $\varphi(x) = \int_0^x f(t)dt \equiv \dfrac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ được gọi là hàm Laplace.

Các giá trị của hàm Laplace được tính sẵn và liệt kê thành bảng gọi là bảng phân phối Gauss. Ngoài ra, vì φ là hàm lẻ, $\varphi(-x) = -\varphi(x)$, $\forall x \in \mathbb{R}$, nên người ta chỉ cần liệt kê các giá trị của $\varphi(x)$ với $x > 0$.

Bảng phân phối Gauss gồm 400 giá trị của $\varphi(x)$, với x thay đổi từ 0.00, 0.01, 0.02, ..., 3.99 và được bố trí như sau

- Các hàng trong bảng, trừ hàng đầu, được đánh số từ 0.0, 0.1, đến 3.9.

- Các cột trong bảng, trừ cột đầu, được đánh số từ 0.00, 0.01 tới 0.09.

Khi đó, ứng với mỗi giá trị x trong khoảng từ 0.00 đến 3.99 với hai số lẻ thập phân dạng $x = a.bc$, giá trị $\varphi(x)$ nằm ở hàng đánh số a.b và cột đánh số 0.0c. Chẳng hạn, với $x = 1.52$, $\varphi(x) \equiv \varphi(1.52)$ nằm ở hàng 1.5, cột 0.02, nghĩa là $\varphi(1.52) = 0.4357$.

	...	...	0.01	***0.02***	0.03	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
1.4	...	...	0.4207	0.4222	0.4236	...	...
1.5	...	...	0.4345	**0.4357**	0.4370	...	...
1.6	...	...	0.4463	0.4474	0.4484	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Ngoài ra, vì $\varphi(+\infty) \equiv \dfrac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-t^2/2} dt = 0.5$ nên $\varphi(x) = 0.5$, với $4 \le x \le +\infty$.

Tóm lại, khi $X \sim N(0,1)$ thì

$$P(a \le X \le b) = \varphi(b) - \varphi(a), \qquad (1)$$

với $\varphi(x)$ được tính như sau

- Khi $0 \le x \le 3.99$, giá trị $\varphi(x)$ được tìm thấy trong bảng,
- Khi $4 \le x \le +\infty$, $\varphi(x) = 0.5$,
- Khi $x < 0$, ta dùng công thức $\varphi(-x) = -\varphi(x)$.

Chú ý :

Công thức (1) vẫn đúng cho trường hợp $a = -\infty$ và / hay $b = +\infty$. Chẳng hạn

$$P(-\infty < X \le b) = \varphi(b) - \varphi(-\infty) = \varphi(b) + 0.5,$$

vì $\varphi(-\infty) = -\varphi(+\infty) = -0.5$,

$$P(a \le X < +\infty) = \varphi(+\infty) - \varphi(a) = 0.5 - \varphi(a).$$

Hơn nữa, do tính chất của tích phân, ta có

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b).$$

1.1.2. Tìm C sao cho $P(-C \le X \le C) = P(|X| \le C) = \gamma$, với γ cho trước.

Xuất phát từ đẳng thức $\gamma = P(-C \le X \le C) = 2\varphi(C)$, ta được $\varphi(C) = \frac{\gamma}{2}$.

Do đó, ứng với giá trị γ cho trước, tính $\varphi(C) = \frac{\gamma}{2}$ và tìm vị trí của số hạng này trong bảng. Bấy giờ, C chính là tổng của *số chỉ hàng* và *số chỉ cột*. Chẳng hạn, với $\gamma = 0.95$, $\varphi(C) = \frac{\gamma}{2} = 0.475$. Tra bảng, ta thấy giá trị 0.475 nằm ở hàng 1.9, cột 0.06, điều này có nghĩa là $\varphi(1.96) = 0.475$. Do đó $C = 1.96$.

	...	...	0.05	***0.06***	0.07	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
1.8	...	...	0.4678	0.4686	0.4693	...	...
1.9	...	...	0.4744	**0.4750**	0.4756	...	...
2.0	...	...	0.4798	0.4803	0.4808	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Chú ý :

Ta còn gặp bài toán tìm C sao cho $P(|X| > C) = \alpha$, với α cho trước. Khi đó

$$P(|X| \le C) = 1 - P(|X| > C) = 1 - \alpha = \gamma$$

và ta nhận trở lại bài toán vừa khảo sát. Thông thường, γ và α lần lượt được gọi là *độ tin cậy* và *nguy cơ sai lầm*.

1.2. Phân phối Student $St(n)$.

Do phân phối Student thường chỉ dùng trong các bài toán thống kê nên với $T \sim St(n)$, người ta chỉ có các nhu cầu

- tìm C sao cho $P(|T| \le C) = \gamma$,

- tìm C sao cho $P(|T| > C) = \alpha$.

Để làm được điều này, người ta tính sẵn $P(|T| > C) = \alpha$, $T \sim St(n)$, với một số giá trị của (nguy cơ sai lầm) α và (độ tự do) n và liệt kê trong bảng gọi là bảng phân phối Student.

Cụ thể, các hàng của bảng, trừ hàng 1, được đánh số theo độ tự do n, các cột của bảng, trừ cột 1, được đánh số theo (nguy cơ sai lầm) α. Khi đó, nội dung trong bảng ứng với hàng và cột nhận được chính là giá trị C cần tìm.

Ví dụ : Với $T \sim St(10)$, để tìm C sao cho $P(|X| > C) = 0.05$, nội dung bảng ứng với hàng 10, cột 0.05 cho giá trị $C = 2.228$.

	...	...	0.04	***0.05***	0.06	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
9	...	...	2.398	2.262	2.150	...	...
10	...	...	2.359	**2.228**	2.120	...	...
11	...	...	2.328	2.201	2.096	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Trường hợp tìm C sao cho $P(|T| \le C) = \gamma$ được khảo sát giống như trường hợp phân phối Gauss

$$P(|T| > C) = 1 - P(|T| \le C) = 1 - \gamma.$$

Chú ý :

Khi $T \sim St(n)$, với $n \ge 30$, ta có thể

- Dùng bảng phân phối Student với độ tự do $n = \infty$ (hàng cuối), hay

- Xấp xỉ phân phối Student bằng phân phối Gauss, nghĩa là $X \sim N(0,1)$.

1.3. Phân phối Chi-Bình phương.

Tương tự phân phối Student, phân phối Chi-Bình phương được dùng trong thống kê và ta gặp hai bài toán sau

- Tìm $a, b \in \mathbb{R}$ sao cho $P(X < a) = P(X > b) = \frac{\alpha}{2}$, với (nguy cơ sai lầm) α cho trước (bài toán ước lượng),

- Tìm $C \in \mathbb{R}$ sao cho $P(X > C) = \alpha$ (bài toán kiểm định).

Do $P(X < a) = 1 - P(X \ge a) = 1 - P(X > a)$ nên để giải các bài toán này, người ta tính sẵn một số giá trị x sao cho $P(X \ge x) = \alpha$, $X \sim \chi^2(n)$, tương ứng với các giá trị của α và n cho trước, và được liệt kê thành bảng. Các hàng, trừ hàng 1, được đánh số theo bậc tự do n, các cột, trừ cột 1, được đánh số theo các giá trị của α. Giá trị trong bảng tương ứng với hàng và cột tìm được chính là giá trị x cần tìm. Chẳng hạn, với $X \sim \chi^2(10)$, để tìm x sao cho $P(X \ge x) = 0.025$, tra bảng ứng với hàng 10, cột 0.025, ta được $x = 20.483$.

	...	...	0.02	***0.025***	0.03	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
9	...	...	19.679	19.023	18.480	...	...
10	...	...	21.161	**20.483**	19.922	...	...
11	...	...	22.618	21.920	21.342	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

1.4. Phân phối Fisher $F(n,m)$.

Tương tự như phân phối chi-bình phương, với $X \sim F(n,m)$, các giá trị x sao cho $P(X \geq x) = \alpha$ được tính sẵn với một số α, m, n cho trước.

Cụ thể, người ta chỉ xét hai giá trị của α là 0.05 và 0.01, và các giá trị của x được liệt kê thành hai bảng : bảng 1 ứng với $\alpha = 0.05$ và bảng 2 ứng với $\alpha = 0.01$. Trong mỗi bảng, hàng 1 liệt kê các giá trị của n. Cột 1 liệt kê các giá trị của m và giá trị trong bảng là giá trị x cần tìm tương ứng.

Chẳng hạn, nếu $X \sim F(5,10)$, để tìm x sao cho $P(X \geq x) = 0.05$, ta tra bảng 1, hàng 10, cột 5 và nhận được giá trị của x là 3.33

	...	...	4	***5***	6	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
9	...	...	3.63	3.48	3.37	...	...
10	...	...	3.48	**3.33**	3.22	...	...
11	...	...	3.36	3.2	3.09	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Để tìm x sao cho $P(X \geq x) = 0.01$, tra bảng 2, hàng 10, cột 5, ta được giá trị của x là 5.64

	...	...	4	***5***	6	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
9	...	...	6.42	6.06	5.80	...	...
10	...	...	5.99	**5.64**	5.39	...	...
11	...	...	5.67	5.32	5.07	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

2. Chứng minh đẳng thức

$$Q = \sum_{i=1}^{r}\sum_{j=1}^{k}\frac{\left(\mathbf{N}_{i,j} - \mathbf{N}'_{i,j}\right)^2}{\mathbf{N}'_{i,j}} = \mathbf{N} \times \left(\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{\left(\mathbf{N}_{i,j}\right)^2}{\mathbf{H}_i \times \mathbf{C}_j} - 1\right)$$

Do $N'_{i,j} = \dfrac{\mathbf{H}_i \times \mathbf{C}_j}{\mathbf{N}}$ nên

$$\frac{\left(N_{i,j}-N'_{i,j}\right)^2}{N'_{i,j}} = \frac{\left(N_{i,j}-\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}\right)^2}{\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}} = \frac{N_{i,j}^2-2N_{i,j}\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}+\left(\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}\right)^2}{\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}}$$

$$= \mathbf{N}\frac{N_{i,j}^2}{\mathbf{H}_i\times\mathbf{C}_j}-2N_{i,j}+\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}$$

Vì vậy

$$Q=\mathbf{N}\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{N_{i,j}^2}{\mathbf{H}_i\times\mathbf{C}_j}-2\sum_{i=1}^{r}\sum_{j=1}^{k}N_{i,j}+\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}},$$

trong đó

$$\sum_{i=1}^{r}\sum_{j=1}^{k}N_{i,j}=\mathbf{N};$$

$$\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{\mathbf{H}_i\times\mathbf{C}_j}{\mathbf{N}}=\frac{1}{\mathbf{N}}\sum_{i=1}^{r}\mathbf{H}_i\sum_{j=1}^{k}\mathbf{C}_j=\frac{\mathbf{N}^2}{\mathbf{N}}=\mathbf{N}.$$

Từ đó suy ra

$$Q=\mathbf{N}\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{N_{i,j}^2}{\mathbf{H}_i\times\mathbf{C}_j}-\mathbf{N}$$

$$=\mathbf{N}\left(\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{N_{i,j}^2}{\mathbf{H}_i\times\mathbf{C}_j}-1\right).$$

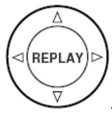
3. Hướng dẫn dùng máy tính

Trong phần này, ta chỉ khảo sát các phép tính thống kê trên ba loại máy : FX 500A, FX 500MS, 570MS và FX 500ES, 570ES.

3.1. Các ký hiệu, ghi chú

- Các phím bấm trên máy được ký hiệu bởi các biểu tượng được đóng khung, ví dụ : [0], [1], [2], [3], ..., [+], [−], [×], [÷], [SHIFT], [MODE], [sin], [cos], ...

- Các phím mũi tên như : [▶], [▲], [▼], [◀] là các phím trên nút Replay



.

- Chuỗi các ký hiệu biểu tượng như " [2] → [3] → [+] → [1] → [7] → [=] " nghĩa là bấm các phím [2], [3], [+], [1], [7], [=] trên máy theo thứ tự từ trái sang phải.

3.2. Các bước trong tính toán thống kê với các loại máy FX–500A, FX–500MS, FX–570MS

Để nhận được các kết quả trong tính toán thống kê một biến, ta thực hiện tuần tự các bước sau :

Bước 1 : Vào chế độ thống kê (SD).

Bước 2 : Xóa dữ liệu thống kê cũ.

Bước 3 : Nhập số liệu thống kê mới.

Bước 4 : Khai thác kết quả từ số liệu thống kê vừa nhập.

Bước 5 : Thoát khỏi chế độ thống kê.

Cụ thể, ta có

3.2.1. Bước 1 : Vào chế độ thống kê

Máy FX–500A : [MODE] → [•]

Máy FX–500MS : [MODE] → [2]

Máy FX–570MS : [MODE] → [1]

3.2.2. Bước 2 : Xóa số liệu thống kê cũ

Trong chế độ SD, ta thực hiện như sau :

Máy FX–500A : [MODE] → [0] → [MODE] → [•]

Máy FX–500MS và FX–570MS : [SHIFT] → [MODE] → [1] → [=] → [AC]

3.2.3. Bước 3 : Nhập số liệu thống kê mới

Số liệu không có tần số : Chẳng hạn để nhập dãy số liệu của X

X	1	2	3	4	5

Máy FX–500A : [1] → [M+] → [2] → [M+] → [3] → [M+] → [4] → [M+]
→ [5] → [M+] → [AC]

Máy FX–500MS và FX–570MS : [1] → [M+] → [2] → [M+] → [3]
→ [M+] → [4] → [M+] → [5] → [M+] → [AC]

Số liệu có tần số : Chẳng hạn để nhập dãy số liệu của X

X	1	2	3	4	5
Tần số	3	2	4	5	2

Máy FX–500A : [1] → [×] → [3] → [M+]
→ [2] → [×] → [2] → [M+]
→ [3] → [×] → [4] → [M+]
→ [4] → [×] → [5] → [M+]
→ [5] → [×] → [2] → [M+]
→ [AC]

Máy FX–500MS và FX–570MS : [1] → [SHIFT] → [,] → [3] → [M+]
→ [2] → [SHIFT] → [,] → [2] → [M+]
→ [3] → [SHIFT] → [,] → [4] → [M+]
→ [4] → [SHIFT] → [,] → [5] → [M+]
→ [5] → [SHIFT] → [,] → [2] → [M+]
→ [AC]

3.2.4. Bước 4 : Khai thác kết quả

Với $\overline{X}$ chỉ trung bình mẫu; σ_n chỉ phương sai mẫu chưa hiệu chỉnh; σ_{n-1} chỉ phương sai mẫu có hiệu chỉnh; $\sum x_i$ chỉ tổng số liệu mẫu; $\sum x_i^2$ chỉ tổng bình phương các số liệu mẫu và n chỉ cỡ mẫu. Ta có

Máy FX–500A :

$\overline{X}$: [SHIFT] → [7]

σ_n : SHIFT → 8

σ_{n-1} : SHIFT → 9

$\sum x_i$: SHIFT → 5

$\sum x_i^2$: SHIFT → 4

n : SHIFT → 6

Máy FX–500MS và FX–570MS :

$\overline{X}$: SHIFT → 2 → 1 → =

σ_n : SHIFT → 2 → 2 → =

σ_{n-1} : SHIFT → 2 → 3 → =

$\sum x_i$: SHIFT → 1 → 2 → =

$\sum x_i^2$: SHIFT → 1 → 1 → =

n : SHIFT → 1 → 3 → =

3.2.5. Bước 5 : Thoát khỏi chế độ thống kê

Máy FX–500A : MODE → 0

Máy FX–500MS : MODE → 1

Máy FX–570MS : MODE → 1

3.2. Các bước trong tính toán thống kê với các loại máy FX–500ES, FX–570ES

Để nhận được các kết quả trong tính toán thống kê một biến, ta thực hiện tuần tự các bước sau :

Bước 1 : Vào chế độ thống kê (STAT).

Bước 2 : Vào chế độ chỉnh sửa dữ liệu.

Bước 2.1 : Xóa dữ liệu thống kê cũ.

Bước 2.2 : Nhập số liệu thống kê mới.

Bước 3 : Khai thác kết quả từ số liệu thống kê vừa nhập.

Bước 4 : Thoát khỏi chế độ thống kê.

Chú ý rằng chỉ có sự khác biệt giữa hai loại máy này trong bước 1. Tất cả các bước còn lại là như nhau. Cụ thể, ta có

3.3.1. Bước 1 : Vào chế độ thống kê

Máy FX–500ES : MODE → 2 → 1 → AC

Máy FX–570ES : MODE → 3 → 1 → AC

3.3.2. Bước 2 : Vào chế độ chỉnh sửa dữ liệu

Trong chế độ STAT, ta thực hiện như sau :

SHIFT → 1 → 2

3.3.2.1 Bước 2.1 : Xóa số liệu thống kê cũ

SHIFT → 1 → 3 → 2

3.3.2.2 Bước 2.2 : Nhập số liệu thống kê mới

Số liệu không có tần số : Chẳng hạn để nhập dãy số liệu của X

X	1	2	3	4	5

1 → = → 2 → = → 3 → = → 4

→ = → 5 → = → AC

Số liệu có tần số : Chẳng hạn để nhập dãy số liệu của X

X	1	2	3	4	5
Tần số	3	2	4	5	2

Nếu trên màn hình không có cột Freq (cột để nhập tần số) thì bấm

SHIFT → MODE → ▼ → 4 → 1

Nhập dữ liệu : 1 → = → 2 → = → 3 → = → 4

→ = → 5 → = →

Nhập tần số : ▼ → ▶ → 3 → = → 2 → = → 4

→ = → 5 → = → 2 → = → AC

3.3.3. Bước 3 : Khai thác kết quả

Với $\overline{X}$ chỉ trung bình mẫu; σ_n chỉ phương sai mẫu chưa hiệu chỉnh; σ_{n-1} chỉ phương sai mẫu có hiệu chỉnh; $\sum x_i$ chỉ tổng số liệu mẫu; $\sum x_i^2$ chỉ tổng bình phương các số liệu mẫu và n chỉ cỡ mẫu. Ta có

$\overline{X}$: SHIFT → 1 → 5 → 2 → =

σ_n : SHIFT → 1 → 5 → 3 → =

σ_{n-1} : SHIFT → 1 → 5 → 4 → =

n : SHIFT → 1 → 5 → 1 → =

$\sum x_i$: SHIFT → 1 → 4 → 2 → =

$\sum x_i^2$: SHIFT → 1 → 4 → 1 → =

3.3.4. Bước 4 : Thoát khỏi chế độ thống kê

MODE → 1

PHÂN PHỐI POISSON

$$P(X \le x) = \sum_{k \le x} e^{-\mu} \frac{\mu^k}{k!} \text{ với } X \sim P(\mu)$$

Hàng 1 : các giá trị của μ.

Cột 1 : Các giá trị của x.

Nội dung bảng : Các giá trị $P(X \le x)$.

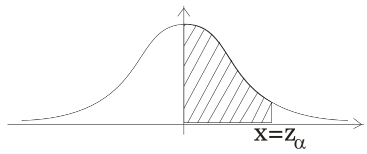
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

	1	2	3	4	5
0	0.3679	0.1353	0.0498	0.0183	0.0067
1	0.7358	0.4060	0.1991	0.0916	0.0404
2	0.9197	0.6767	0.4232	0.2381	0.1247
3	0.9810	0.8571	0.6472	0.4335	0.2650
4	0.9963	0.9473	0.8153	0.6288	0.4405
5	0.9994	0.9834	0.9161	0.7851	0.6160
6	0.9999	0.9955	0.9665	0.8893	0.7622
7	1.0000	0.9989	0.9881	0.9489	0.8666
8		0.9998	0.9962	0.9786	0.9319
9		1.0000	0.9989	0.9919	0.9682
10			0.9997	0.9972	0.9863
11			0.9999	0.9991	0.9945
12			1.0000	0.9997	0.9980
13				0.9999	0.9993
14				1.0000	0.9998
15					0.9999
16					1.0000

PHÂN PHỐI GAUSS

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt = P(0 \le X \le x) \equiv \alpha \,,$$

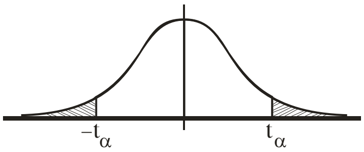
với $X \sim N(0;1)$, $x \equiv z_\alpha$.



	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

PHÂN PHỐI STUDENT

$P(|T| \geq t_\alpha) = \alpha$ với $T \sim St(n)$



Cột 1 : giá trị độ tự do n.

Hàng 1 : Giá trị nguy cơ sai lầm α

Nội dung bảng : Giá trị t_α tương ứng với n và α

	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.15	0.2
1	63.656	31.821	21.205	15.894	12.706	10.579	9.058	7.916	7.026	6.314	4.165	3.078
2	9.925	6.965	5.643	4.849	4.303	3.896	3.578	3.320	3.104	2.920	2.282	1.886
3	5.841	4.541	3.896	3.482	3.182	2.951	2.763	2.605	2.471	2.353	1.924	1.638
4	4.604	3.747	3.298	2.999	2.776	2.601	2.456	2.333	2.226	2.132	1.778	1.533
5	4.032	3.365	3.003	2.757	2.571	2.422	2.297	2.191	2.098	2.015	1.699	1.476
6	3.707	3.143	2.829	2.612	2.447	2.313	2.201	2.104	2.019	1.943	1.650	1.440
7	3.499	2.998	2.715	2.517	2.365	2.241	2.136	2.046	1.966	1.895	1.617	1.415
8	3.355	2.896	2.634	2.449	2.306	2.189	2.090	2.004	1.928	1.860	1.592	1.397
9	3.250	2.821	2.574	2.398	2.262	2.150	2.055	1.973	1.899	1.833	1.574	1.383
10	3.169	2.764	2.527	2.359	2.228	2.120	2.028	1.948	1.877	1.812	1.559	1.372
11	3.106	2.718	2.491	2.328	2.201	2.096	2.007	1.928	1.859	1.796	1.548	1.363
12	3.055	2.681	2.461	2.303	2.179	2.076	1.989	1.912	1.844	1.782	1.538	1.356
13	3.012	2.650	2.436	2.282	2.160	2.060	1.974	1.899	1.832	1.771	1.530	1.350
14	2.977	2.624	2.415	2.264	2.145	2.046	1.962	1.887	1.821	1.761	1.523	1.345
15	2.947	2.602	2.397	2.249	2.131	2.034	1.951	1.878	1.812	1.753	1.517	1.341
16	2.921	2.583	2.382	2.235	2.120	2.024	1.942	1.869	1.805	1.746	1.512	1.337
17	2.898	2.567	2.368	2.224	2.110	2.015	1.934	1.862	1.798	1.740	1.508	1.333
18	2.878	2.552	2.356	2.214	2.101	2.007	1.926	1.855	1.792	1.734	1.504	1.330
19	2.861	2.539	2.346	2.205	2.093	2.000	1.920	1.850	1.786	1.729	1.500	1.328
20	2.845	2.528	2.336	2.197	2.086	1.994	1.914	1.844	1.782	1.725	1.497	1.325
21	2.831	2.518	2.328	2.189	2.080	1.988	1.909	1.840	1.777	1.721	1.494	1.323
22	2.819	2.508	2.320	2.183	2.074	1.983	1.905	1.835	1.773	1.717	1.492	1.321
23	2.807	2.500	2.313	2.177	2.069	1.978	1.900	1.832	1.770	1.714	1.489	1.319
24	2.797	2.492	2.307	2.172	2.064	1.974	1.896	1.828	1.767	1.711	1.487	1.318
25	2.787	2.485	2.301	2.167	2.060	1.970	1.893	1.825	1.764	1.708	1.485	1.316
26	2.779	2.479	2.296	2.162	2.056	1.967	1.890	1.822	1.761	1.706	1.483	1.315
27	2.771	2.473	2.291	2.158	2.052	1.963	1.887	1.819	1.758	1.703	1.482	1.314
28	2.763	2.467	2.286	2.154	2.048	1.960	1.884	1.817	1.756	1.701	1.480	1.313
29	2.756	2.462	2.282	2.150	2.045	1.957	1.881	1.814	1.754	1.699	1.479	1.311
∞	2.576	2.326	2.170	2.054	1.960	1.881	1.812	1.751	1.695	1.645	1.440	1.282

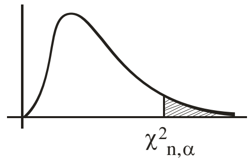
PHÂN PHỐI CHI – BÌNH PHƯƠNG

$P\left(X \geq \chi^2_{n,\alpha}\right) = \alpha$ khi $X \sim \chi^2(n)$

Hàng 1 : Giá trị của α

Cột 1 : Giá trị độ tự do n.

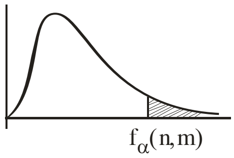
Nội dung bảng : Giá trị $\chi^2_{n,\alpha}$.



	0.005	**0.01**	**0.015**	**0.02**	**0.025**	**0.03**	**0.05**	**0.95**	**0.975**	**0.98**	**0.99**	**0.995**
1	7.879	6.635	5.916	5.412	5.024	4.709	3.841	0.004	0.001	0.001	0.000	0.000
2	10.597	9.210	8.399	7.824	7.378	7.013	5.991	0.103	0.051	0.040	0.020	0.010
3	12.838	11.345	10.465	9.837	9.348	8.947	7.815	0.352	0.216	0.185	0.115	0.072
4	14.860	13.277	12.339	11.668	11.143	10.712	9.488	0.711	0.484	0.429	0.297	0.207
5	16.750	15.086	14.098	13.388	12.832	12.375	11.070	1.145	0.831	0.752	0.554	0.412
6	18.548	16.812	15.777	15.033	14.449	13.968	12.592	1.635	1.237	1.134	0.872	0.676
7	20.278	18.475	17.398	16.622	16.013	15.509	14.067	2.167	1.690	1.564	1.239	0.989
8	21.955	20.090	18.974	18.168	17.535	17.011	15.507	2.733	2.180	2.032	1.647	1.344
9	23.589	21.666	20.512	19.679	19.023	18.480	16.919	3.325	2.700	2.532	2.088	1.735
10	25.188	23.209	22.021	21.161	20.483	19.922	18.307	3.940	3.247	3.059	2.558	2.156
11	26.757	24.725	23.503	22.618	21.920	21.342	19.675	4.575	3.816	3.609	3.053	2.603
12	28.300	26.217	24.963	24.054	23.337	22.742	21.026	5.226	4.404	4.178	3.571	3.074
13	29.819	27.688	26.403	25.471	24.736	24.125	22.362	5.892	5.009	4.765	4.107	3.565
14	31.319	29.141	27.827	26.873	26.119	25.493	23.685	6.571	5.629	5.368	4.660	4.075
15	32.801	30.578	29.235	28.259	27.488	26.848	24.996	7.261	6.262	5.985	5.229	4.601
16	34.267	32.000	30.629	29.633	28.845	28.191	26.296	7.962	6.908	6.614	5.812	5.142
17	35.718	33.409	32.011	30.995	30.191	29.523	27.587	8.672	7.564	7.255	6.408	5.697
18	37.156	34.805	33.382	32.346	31.526	30.845	28.869	9.390	8.231	7.906	7.015	6.265
19	38.582	36.191	34.742	33.687	32.852	32.158	30.144	10.117	8.907	8.567	7.633	6.844
20	39.997	37.566	36.093	35.020	34.170	33.462	31.410	10.851	9.591	9.237	8.260	7.434
21	41.401	38.932	37.434	36.343	35.479	34.759	32.671	11.591	10.283	9.915	8.897	8.034
22	42.796	40.289	38.768	37.659	36.781	36.049	33.924	12.338	10.982	10.600	9.542	8.643
23	44.181	41.638	40.094	38.968	38.076	37.332	35.172	13.091	11.689	11.293	10.196	9.260
24	45.558	42.980	41.413	40.270	39.364	38.609	36.415	13.848	12.401	11.992	10.856	9.886
25	46.928	44.314	42.725	41.566	40.646	39.880	37.652	14.611	13.120	12.697	11.524	10.520
26	48.290	45.642	44.031	42.856	41.923	41.146	38.885	15.379	13.844	13.409	12.198	11.160
27	49.645	46.963	45.331	44.140	43.195	42.407	40.113	16.151	14.573	14.125	12.878	11.808
28	50.994	48.278	46.626	45.419	44.461	43.662	41.337	16.928	15.308	14.847	13.565	12.461
29	52.335	49.588	47.915	46.693	45.722	44.913	42.557	17.708	16.047	15.574	14.256	13.121
30	53.672	50.892	49.199	47.962	46.979	46.160	43.773	18.493	16.791	16.306	14.953	13.787
35	60.275	57.342	55.553	54.244	53.203	52.335	49.802	22.465	20.569	20.027	18.509	17.192
40	66.766	63.691	61.812	60.436	59.342	58.428	55.758	26.509	24.433	23.838	22.164	20.707
45	73.166	69.957	67.994	66.555	65.410	64.454	61.656	30.612	28.366	27.720	25.901	24.311
50	79.490	76.154	74.111	72.613	71.420	70.423	67.505	34.764	32.357	31.664	29.707	27.991
55	85.749	82.292	80.173	78.619	77.380	76.345	73.311	38.958	36.398	35.659	33.571	31.735
60	91.952	88.379	86.188	84.580	83.298	82.225	79.082	43.188	40.482	39.699	37.485	35.534
65	98.105	94.422	92.161	90.501	89.177	88.069	84.821	47.450	44.603	43.779	41.444	39.383
70	104.215	100.425	98.098	96.387	95.023	93.881	90.531	51.739	48.758	47.893	45.442	43.275
75	110.285	106.393	104.001	102.243	100.839	99.665	96.217	56.054	52.942	52.039	49.475	47.206
80	116.321	112.329	109.874	108.069	106.629	105.422	101.879	60.391	57.153	56.213	53.540	51.172
85	122.324	118.236	115.720	113.871	112.393	111.156	107.522	64.749	61.389	60.412	57.634	55.170
90	128.299	124.116	121.542	119.648	118.136	116.869	113.145	69.126	65.647	64.635	61.754	59.196
95	134.247	129.973	127.341	125.405	123.858	122.562	118.752	73.520	69.925	68.879	65.898	63.250
100	140.170	135.807	133.120	131.142	129.561	128.237	124.342	77.929	74.222	73.142	70.065	67.328

PHÂN PHỐI FISHER

$$P\left(X \geq f_\alpha(n,m)\right) = \alpha \text{ khi } X \sim F(n,m)$$



Hàng 1 : Giá trị của độ tự do (tử số) n.

Cột 1 : Giá trị độ tự do (mẫu số) m.

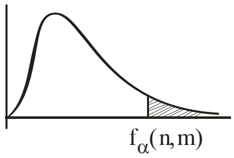
Nội dung bảng : Giá trị $f_\alpha(n,m)$.

Bảng 1 : $\alpha = 0.05$

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.51	19	19.16	19.25	19.3	19.33	19.35	19.37	19.38	19.4	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.5
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.7	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6	5.96	5.91	5.86	5.8	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.5	4.46	4.43	4.4	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1	4.06	4	3.94	3.87	3.84	3.81	3.77	3.74	3.7	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.3	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.9	2.86	2.83	2.79	2.75	2.71
10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.7	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.4
12	4.75	3.89	3.49	3.26	3.11	3	2.91	2.85	2.8	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.3
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.6	2.53	2.46	2.42	2.38	2.34	2.3	2.25	2.21
14	4.6	3.74	3.34	3.11	2.96	2.85	2.76	2.7	2.65	2.6	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.9	2.79	2.71	2.64	2.59	2.54	2.48	2.4	2.33	2.29	2.25	2.2	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.2	2.96	2.81	2.7	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.1	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.9	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.1	2.87	2.71	2.6	2.51	2.45	2.39	2.35	2.28	2.2	2.12	2.08	2.04	1.99	1.95	1.9	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.1	2.05	2.01	1.96	1.92	1.87	1.81
22	4.3	3.44	3.05	2.82	2.66	2.55	2.46	2.4	2.34	2.3	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.8	2.64	2.53	2.44	2.37	2.32	2.27	2.2	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.4	3.01	2.78	2.62	2.51	2.42	2.36	2.3	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.6	2.49	2.4	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4	3.15	2.76	2.53	2.37	2.25	2.17	2.1	2.04	1.99	1.92	1.84	1.75	1.7	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.5	1.43	1.35	1.25
∞	3.84	3	2.6	2.37	2.21	2.1	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1

PHÂN PHỐI FISHER

$$P\left(X \geq f_{\alpha}(n,m)\right) = \alpha \text{ khi } X \sim F(n,m)$$



Hàng 1 : Giá trị của độ tự do (tử số) n.

Cột 1 : Giá trị độ tự do (mẫu số) m.

Nội dung bảng : Giá trị $f_{\alpha}(n,m)$.

Bảng 2 : $\alpha = 0.01$

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6107	6157	6209	6234	6260	6286	6313	6340	6366
2	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.48	99.48	99.49	99.5
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.1
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.5
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.6
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.8
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.6
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00