ThS. NGUYỄN TRUNG ĐÔNG

Bài tập

XÁC SUẤT THỐNG KÊ

TP. HỒ CHÍ MINH - 2013

Chương 0. GIẢI TÍCH TỔ HỢP

0.1. Tóm tắt lý thuyết

0.1.1. Quy tắc đếm

Ta chỉ khảo sát tập hữu hạn: $X = \{x_1, x_2, ..., x_n\}$, X có n phần tử,

ký hiệu $|X| = n$.

0.1.2. Công thức cộng

Cho X, Y là hai tập hữu hạn và $X \cap Y = \varnothing$, ta có $|X \cup Y| = |X| + |Y|$

Tổng quát: Nếu cho k tập hữu hạn $X_1, X_2, ..., X_k$ sao cho $X_i \cap Y_j = \varnothing,\ i \neq j$,

ta có

$$\left|X_1 \cup X_2 \cup ... \cup X_k\right| = \left|X_1\right| + \left|X_2\right| + ... + \left|X_k\right|$$

0.1.3. Công thức nhân

Cho X, Y là hai tập hữu hạn, định nghĩa tập tích nhý sau

$X \times Y = \{(x, y) / x \in X \wedge y \in Y\}$, ta có $|X \times Y| = |X| \cdot |Y|$

Tổng quát: Nếu cho n tập hữu hạn $X_1, X_2, ..., X_k$, ta có

$$\left|X_1 \times X_2 \times ... \times X_k\right| = \left|X_1\right| \cdot \left|X_2\right| \cdot ... \cdot \left|X_k\right|$$

0.1.4. Quy tắc cộng

Giả sử một công việc có thể thực hiện một trong k phương pháp, trong đó

- Phương pháp 1 có n_1 cách thực hiện,
- Phương pháp 2 có n_2 cách thực hiện,…,
- Phương pháp k có n_k cách thực hiện,

và hai phương pháp khác nhau không có cách thực hiện chung.

Khi đó, ta có $n_1 + n_2 + ... + n_k$ cách thực hiện công việc.

0.1.5. Quy tắc nhân

Giả sử một công việc có thể thực hiện tuần tự theo k bước, trong đó

- Bước 1 có n_1 cách thực hiện,

- Bước 2 có n_2 cách thực hiện,…,
- Bước k có n_k cách thực hiện,

Khi đó, ta có $n_1 \times n_2 \times \ldots \times n_k$ cách thực hiện công việc.

0.1.6. Giải tích tổ hợp

a. Chỉnh hợp

Định nghĩa: Chỉnh hợp chập k từ n phần tử là một bộ có kể thứ tự gồm k phần tử khác nhau lấy từ n phần tử đã cho.

Số chỉnh hợp: Số chỉnh hợp chập k từ n phần tử, ký hiệu là : A_n^k

Công thức tính :

$$A_n^k = n(n-1)...(n-k+1) = \frac{n!}{(n-k)!}$$

b. Chỉnh hợp lặp

Định nghĩa: Chỉnh hợp lặp chập k từ n phần tử là một bộ có kể thứ tự gồm k phần tử không cần khác nhau lấy từ n phần tử đã cho.

Số chỉnh hợp lặp: Số chỉnh hợp lặp chập k từ n phần tử ký, hiệu là : $\widetilde{A_n^k}$

Công thức tính: $\widetilde{A_n^k} = n^k$

c. Hoán vị

Định nghĩa: Một hoán vị từ n phần tử là một bộ có kể thứ tự gồm n phần tử khác nhau đã cho.

Số hoán vị: Số hoán vị từ n phần tử, ký hiệu là P_n

Công thức tính:

$$P_n = n! = (n-1)(n-2)...(1)$$

d. Tổ hợp

Định nghĩa: Một tổ hợp chập k từ n phần tử là một tập con gồm k phần tử lấy từ n phần tử.

Số tổ hợp : Số tổ hợp chập k từ n phần tử ký hiệu là : C_n^k

Công thức tính:

$$C_n^k = \frac{n!}{k!(n-k)!}$$

e. Nhị thức Newton

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k$$

$$(1+x)^n = \sum_{k=0}^{n} C_n^k x^k$$

Bài tập mẫu

Bài 1. Đêm chung kết hoa khôi sinh viên thành phố có 12 thí sinh, chọn 3 thí sinh trao giải: Hoa khôi, Á khôi 1, Á khôi 2. Có bao nhiêu cách chọn ?

Giải

Nhận xét: thí sinh được trao giải, được chọn từ 12 thí sinh, và có thứ tự (A, B, C cùng được trao giải, nhưng trường hợp A là hoa khôi, khác trường hợp B là hoa khôi).

Suy ra mỗi cách chọn là một chỉnh hợp chập 3 từ 12 phần tử.

Vậy số cách chọn là: $A_{12}^3 = 12.11.10 = 1320$.

Bài 2. Giả sử có một vị thần có quyền phân phát ngày sinh cho con người, có bao nhiêu cách phân bố ngày sinh cho 10 em bé ra đời trong năm 1999 tại 1 khu tập thể của công nhân viên chức.

Giải

Nhận xét: Mỗi ngày sinh của một em bé là 1 trong 365 ngày của năm 1999, nên các ngày sinh có thể trùng nhau.

Suy ra mỗi cách phân bố 10 ngày sinh là một chỉnh hợp lặp chập 10 từ 365 phần tử.

Vậy số cách phân bố ngày sinh là: $\tilde{A}_{365}^{10} = 365^{10}$

Bài 3. có 3 bộ sách:

Toán cao cấp C : 6 tập,

Kinh tế quốc tế : 2 tập,

Xác suất thống kê : 3 tập,

Được đặt lên giá sách. Có bao nhiêu cách sắp:

a) Tuỳ ý;

b) Các tập sách được đặt theo từng bộ.

Giải

a) Nhận xét: 3 bộ sách có tất cả 11 tập; đặt lên giá sách, mỗi cách sắp là hoán vị của 11 phần tử.

Suy ra số cách sắp tuỳ ý: $P_{11} = 11!$

b) Nhận xét:

- Xem mỗi bộ sách là một phần tử.

  $\Rightarrow$ có $3!$ cách sắp xếp 3 phần tử này.

- Các cặp sách trong mỗi bộ sách xáo trộn với nhau.

  Toán cao cấp C : $6!$

  Kinh tế lượng : $2!$

  Xác suất thống kê : $3!$

Suy ra: số cách sắp xếp 3 bộ sách theo từng bộ là: 3!6!2!3!

Bài 4. Giải bóng đá ngoại hạng Anh có 20 đội bóng thi đấu vòng tròn, có bao nhiêu trận đấu được tổ chức nếu:

a) Thi đấu vòng tròn 1 lượt.

b) Thi đấu vòng tròn 2 lượt.

<u>Giải</u>

a) Nhận xét: Mỗi trận đấu ứng với việc chọn 2 đội chọn từ 20 đội. Suy ra mỗi trận đấu là một tổ hợp chập 2 từ 20 phần tử.

Số mỗi trận đấu được tổ chức là :

$$C_{20}^{2} = \frac{20!}{2!18!} = 190 \text{ trận}$$

b) Nhận xét: Mỗi trận đấu ứng với việc chọn 2 đội chọn từ 20 đội. (đội chủ, đội khách).

Suy ra mỗi trận đấu là một chỉnh hợp chập 2 từ 20 phần tử.

Vậy số trận đấu là : $A_{20}^{2} = \frac{20!}{18!} = 380 \text{ trận}$

Bài tập rèn luyện

Bài 1. Trong một lớp gồm 30 sinh viên, cần chọn ra ba sinh viên để làm lớp trưởng, lớp phó và thủ quỹ. Hỏi có tất cả bao nhiêu cách bầu chọn ?

Bài 2. Có bao nhiêu cách xếp 10 người ngồi thành hàng ngang sao cho A và B ngồi cạnh nhau.

Bài 3. Một hộp đựng 6 bi trắng và 4 bi đen.

a) Có tất cả bao nhiêu cách lấy ra 5 bi ?

b) Có bao nhiêu cách lấy ra 5 bi trong đó có 2 bi trắng ?

Bài 4. Trong một nhóm ứng viên gồm 7 nam và 3 nữ,

a) có bao nhiêu cách thành lập một ủy ban gồm 3 người ?

b) có bao nhiêu cách thành lập một ủy ban gồm 3 người trong đó có đúng 1 nữ ?

c) có bao nhiêu cách thành lập một ủy ban gồm 3 người trong đó có ít nhất 1 nữ ?

Bài 5. Cho các chữ số 0, 1, 2, 3, 4, 5. Hỏi từ các chữ số này:

a) Lập được bao nhiêu số có 4 chữ số khác nhau trong đó nhất thiết có mặt chữ số 5?

b) Lập được bao nhiêu số có 7 chữ số trong đó chữ số 5 có mặt đúng 3 lần còn các chữ số khác có mặt không quá một lần?

Chương 1. ĐẠI CƯƠNG VỀ XÁC SUẤT

1.1. Tóm tắt lý thuyết

1.1.1. Định nghĩa xác suất

Xét biến cố A với không gian mẫu Ω tương ứng, ta có định nghĩa cổ điển

$$P(A) = \frac{|A|}{|\Omega|},$$

trong đó $|A|$ và $|\Omega|$ lần lượt là số phần tử của A và của Ω và định nghĩa bằng tần suất

$$P(A) = \frac{\text{Số trường hợp thuận lợi cho A}}{\text{Số trường hợp xảy ra}}$$

1.1.2. Tính chất cơ bản của xác suất

a) $0 \le P(A) \le 1,\ P(\varnothing) = 0,\ \ P(\Omega) = 1$.

b) Công thức cộng: Cho họ biến cố $A_1, A_2, \ldots, A_n$ xung khắc với nhau từng đôi một (nghĩa là $A_iA_j = \varnothing,\ \text{khi}\ i \ne j$), ta có

$$P\left(A_1 + A_2 + \ldots + A_n\right) = P\left(A_1\right) + P\left(A_2\right) + \ldots + P\left(A_n\right).$$

c) Với A, B là hai biến cố bất kỳ, ta có

$$P\left(A + B\right) = P(A) + P(B) - P(AB).$$

d) $P(\bar{A}) = 1 - P(A)$

1.1.3. Xác suất có điều kiện

Xác suất để biến cố A xảy ra khi biết biến cố B đã xảy ra

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

với $P(B) > 0$, và ta có công thức nhân

$$P(AB) = P\left(A|B\right)P(B) = P\left(B|A\right)P(A).$$

Khi biến cố B xảy ra hay không xảy ra không ảnh hưởng đến việc biến cố A xảy ra hay không xảy ra, ta nói A, B là hai biến cố độc lập và khi đó

$$P(AB) = P(A)P(B).$$

Ta có *công thức nhân tổng quát,*

$$P(A_1A_2...A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)...P(A_n|A_1A_2...A_{n-1})$$

Khi A_1, A_2, …, A_n là họ các biến cố độc lập, nghĩa là một biến cố xảy ra hay không xảy ra không ảnh hưởng đến việc xảy ra một hay nhiều biến cố khác, nghĩa là với bất kỳ họ hữu hạn các biến cố A_{i_1}, A_{i_2}, …, A_{i_k}, ta có

$$P(A_{i_1}A_{i_2}...A_{i_k}) = P(A_{i_1})P(A_{i_2})...P(A_{i_k}).$$

1.1.4. Công thức xác suất toàn phần – công thức Bayes

Cho $B_1, B_2, ..., B_n$ là họ đầy đủ các biến cố, nghĩa là

i) $B_iB_j = \varnothing$

ii) $B_1 + B_2 + ... + B_n = \Omega$

với A là một biến cố bất kỳ, ta có

a) Công thức xác suất toàn phần

$$P(A) = P(A | B_1)P(B_1) + P(A | B_n)P(B_n) + ... + P(A | B_n)P(B_n)$$

b) Công thức Bayes

$$P(B_k | A) = \frac{P(A | B_k)P(B_k)}{P(A)}, \quad k = 1, 2, ..., n.$$

1.1.5. Dãy phép thử Bernoulli

Khi thực hiện n lần phép thử độc lập nhau và gọi X là số lần biến cố A xảy ra trong n lần thực hiện phép thử, thì biến cố $(X = k)$ chỉ trường hợp biến cố A xảy ra đúng k lần trong n lần thực hiện phép thử, ta có

$$P(X = k) = C_n^k p^k (1 - p)^{n-k}, \quad k = 0, 1, 2, .., n$$

với $p = P(A)$. Ta ký hiệu $X \sim B(n; p)$.

1.2. Bài tập mẫu

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

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

<u>Giải</u>

Ta có

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

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

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

nên

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

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 $P(\bar{A}B)$.

<u>Giải</u>

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

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

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

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

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

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

Tương tự, ta có

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

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.

<u>Giải</u>

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

$$P(A+B) = P(A) + P(B) - P(AB) = 0,09 + 0,12 - 0,07 = 0,14.$$

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

$$P(\overline{A}.\overline{B}) = P(\overline{A+B}) = 1 - P(A+B) = 1 - 0,14 = 0,86.$$

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

$$P(\overline{A} + \overline{B}) = P(\overline{AB}) = 1 - P(AB) = 1 - 0,07 = 0,93.$$

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

$$P(A.\overline{B}) = P(A) - P(AB) = 0,09 - 0,07 = 0,02.$$

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

$$P(\overline{A}.B) = P(B) - P(AB) = 0,12 - 0,07 = 0,05.$$

Bài 4. 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

$$P\left(A_1\middle|A\right)=\frac{30}{44}=0,682,$$
$$P\left(B_1\middle|A\right)=\frac{4}{44}=0,091,$$
$$P\left(C_1\middle|A\right)=\frac{10}{44}=0,227.$$

Bài 5. 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 ?

<u>Giải</u>

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

$$\begin{aligned}P(A)&=P(A_1+\bar{A}_1A_2+\bar{A}_1\bar{A}_2A_3)\\&=P(A_1)+P(\bar{A}_1)\cdot P(A_2\mid\bar{A}_1)+P(\bar{A}_1)\cdot P(\bar{A}_2\mid\bar{A}_1)\cdot P(A_3\mid\bar{A}_1\bar{A}_2)\\&=\frac{1}{10}+\frac{9}{10}\cdot\frac{1}{9}+\frac{9}{10}\cdot\frac{8}{9}\cdot\frac{1}{8}=\frac{3}{10}=0,3.\end{aligned}$$

Khi đã biết số cuối cùng là số lẻ thì khi đó các số để chọn quay chỉ ̣cn 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 6. 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 ?

<u>Giải</u>

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

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

Từ đó, suy ra xác suất nhận được bi đỏ

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

và xác suất nhận được bi trắng là

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

Bài 7. 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.

<u>Giải</u>

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

$$P(B|\bar{A}) = P(\bar{B}|\bar{A}) = 0,5,$$

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

$$P(B) = 0,3 + 0,34 = 0,64 \text{ và } P(\bar{B}) = 0,36.$$

a) Do công thức xác suất toàn phần,

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

ta suy ra

$$P(A) = \frac{P(B) - P(B|\bar{A})}{P(B|A) - P(B|\bar{A})} = \frac{0,64 - 0,5}{1 - 0,5} = 0,28.$$

b) Do công thức Bayes,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0,28}{0,64} = 0,4375.$$

Bài 8. 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

$$P(A) = 0,8\,;\ P(A|B) = 0,9\,;\ P(\overline{A}|\overline{B}) = 0,5\,.$$

a) Do công thức xác suất toàn phần,

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

mà $P(A|\overline{B}) = 1 - P(\overline{A}|\overline{B}) = 0,5$, nên xác suất để phép kiểm định là dương tính cho bởi

$$P(B) = \frac{P(A) - P(A|\overline{B})}{P(A|B) - P(A|\overline{B})} = \frac{0,8 - 0,5}{0,9 - 0,5} = 0,75\,.$$

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

$$P(AB + \overline{A}\overline{B}) = P(AB) + P(\overline{A}\overline{B}) = P(A|B)P(B) + P(\overline{A}|\overline{B})P(\overline{B}) = 0,7125.$$

Bài 9. 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

$$P(B_1) = \frac{60}{100} = 0,6\,;\ P(B_2) = \frac{40}{100} = 0,4\,;$$

$$P(A|B_1) = 0,9\,;\ P(A|B_2) = 0,85.$$

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 10. 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

$$P\left(A\middle|\overline{B}\right)=\frac{P\left(\overline{B}\middle|A\right)P\left(A\right)}{P\left(\overline{B}\middle|A\right)P\left(A\right)+P\left(\overline{B}\middle|\overline{A}\right)P\left(\overline{A}\right)}=\frac{0,4\times 0,3}{0,4\times 0,3+0,7\times 0,7}=0,1967.$$

Bài 11. 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.

<u>Giải</u>

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

$P(A_1)=0,1$, $P(A_2)=0,05$, 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

$$\begin{aligned}P(B)&=P\left(\overline{A}_1\overline{A}_2\overline{A}_3\right)=P\left(\overline{A}_1\right)P\left(\overline{A}_2\right)P\left(\overline{A}_3\right)\\&=0,9\times 0,95\times 0,85=0,7267.\end{aligned}$$

Bài 12. 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 $\geq 0,9$.

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

$$\begin{aligned} P\left(X \geq 1\right) &= 1 - P\left(X = 0\right) \\ &= 1 - C_3^0 (0,7)^0 (1 - 0,7)^{3-0} \\ &= 1 - (0,3)^3 = 0,973. \end{aligned}$$

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

$$\begin{aligned} P\left(X \geq 1\right) &= 1 - P\left(X = 0\right) \\ &= 1 - C_n^0 (0,7)^0 (1 - 0,7)^{n-0} \\ &= 1 - (0,3)^n. \end{aligned}$$

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 $\geq 0,9$.

Bài 13. 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 $0,95$.

<u>Giải</u>

Gọi X là số phế phẩm nhận được trong n sản phẩm lấy ra từ lô hàng. Ta có $X \sim B\left(n; 0,01\right)$. Khi đó xác suất để nhận được ít nhất một sản phẩm hỏng là

$$\begin{aligned} P\left(X \geq 1\right) &= 1 - P\left(X = 0\right) \\ &= 1 - C_n^0 (0,01)^0 (1 - 0,01)^{n-0} = 1 - (0,99)^n. \end{aligned}$$

Để tìm n sao cho xác suất nhận được ít nhất một sản phẩm hỏng lớn hơn $0,95$, nghĩa là $P(X \geq 1) > 0,95$, 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 14. 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.

<u>Giải</u>

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

$$P(X=0) = C_3^0(0,1)^0(1-0,1)^3 = (0,9)^3 = 0,729.$$

1.3. Bài tập rèn luyện

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 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 và người thứ hai bắn trật,

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,

C : có ít nhất 1 người bắn trúng.

Bài 3. Ba sinh viên A, B, C cùng thi môn xác suất thống kê. Xét các biến cố:

A : sinh viên A đậu,

WWW.VIETMATHS.COM

B : sinh viên B đậu,

C : sinh viên C đậu.

Hãy biểu diễn qua A, B, C các biến cố sau :

a) chỉ có A đậu,

b) A đậu và B rớt,

c) có ít nhất một người đậu,

d) cả 3 cùng đậu,

e) có ít nhất 2 người đậu,

f) chỉ có 2 người đậu,

g) không ai đậu,

h) không có quá 2 người đậu.

Bài 4. 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 viết 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 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) kế toán trưởng là nữ,

b) ít nhất 1 nữ.

Đáp số: a) 0,75;b) 0,6154.

Bài 6. 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 7. 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) $\frac{1}{15}$.

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

Bài 8. 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 9. 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) 0,476; c) 0,619.*

Bài 10. 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 11. 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) $P(A+B)$, 8) $P(A|B)$,

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

3) $P(\overline{A + B})$, 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) $P(A + B|A\bar{B})$,

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

Bài 12. Đội tuyển cầu lông của Trường Đại học Tài chính - Marketing 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 : 0,9; 0,7; 0,8. 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) 0,317.

Bài 13. Một lớp học có 50 học sinh trong kỳ thi giỏi Toán và Văn, trong đó có 20 người giỏi Toán, 25 người giỏi Văn, 10 người giỏi cả Toán lẫn Văn. Chọn ngẫu nhiên 1 học sinh của lớp này. Tính xác suất để học sinh được chọn giỏi Toán hoặc Văn.

Đáp số: 0,7.

Bài 14. 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 15. 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 16. 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) $\frac{5}{7}$*; b)* $\frac{1}{3}$*; c)* $\frac{2}{7}$.

Bài 17. 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) chỉ có một công ty thua lỗ,

b) có ít nhất một công ty làm ăn không thua lỗ.

Đáp số: a) 0,44; b) 0,92.

Bài 18. 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 : 0,071.

Bài 19. Một lô hàng có 6 sản phẩm tốt, 4 sản phẩm xấu. Lấy ngẫu nhiên không hoàn lại từng sản phẩm cho đến khi lấy được 2 sản phẩm tốt thì ngừng,

a) Tính xác suất để ngừng lại ở lần lấy sản phẩm thứ 2,

b) Biết đã ngừng lại ở lần lấy sản phẩm thứ 4. Tính xác suất để lần lấy thứ nhất lấy được sản phẩm tốt.

Đáp a) $\frac{1}{3}$*; b)* $\frac{1}{14}$.

Bài 20. 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 21. 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) 0,0667; b) 0,0222; c) 0,0222.

Bài 22. 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 la : 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) 0,1496; c) 0,4404; d) 0,0763.

Bài 23. 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 24. 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 tốt.

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

Đáp số: a) 0,981;b) 0,22.

Bài 25. Trong một trạm cấp cứu bỏng : 80% bệnh nhân bỏng do nóng, 20% bỏng do hóa chất. Loại bỏng do nóng có 30% bi biến chứng, loại bỏng do hóa chất có 50% bị biến chứng.

a) Chọn ngẫu nhiên một bệnh án. Tính xác suất để gặp một bệnh án của bệnh nhân bị biến chứng.

b) Rút ngẫu nhiên được một bệnh án của một bệnh nhân bị biến chứng. Tính xác suất để bệnh án đó là của bệnh nhân bị biến chứng do nóng gây ra? do hóa chất gây ra?

Đáp số: a) 0,34; b) 0,71; 0,2942.

Bài 26. Một lô hạt giống được phân thành ba loại. Loại 1 chiếm 2/3 số hạt cả lô, loại 2 chiếm 1/4, còn lại là loại 3. Loại 1 có tỉ lệ nẩy mầm 80%, loại 2 có tỉ lệ nẩy mầm 60% và loại 3 có tỉ lệ nẩy mầm 40%. Hỏi tỉ lệ nẩy mầm chung của lô hạt giống là bao nhiêu ?

Đáp số: 0,72.

Bài 27. Hai nhà máy cùng sản 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) 0,175%; b) nhà máy 2.

Bài 28. 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 29. 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 30. 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: 0,36.

Bài 31. 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 ngẫu nhiên 1 áo từ hộp một bỏ sang hộp hai; sau đó từ hộp này chọn ngẫu nhiên ra 2 áo. Tìm xác suất để cả 2 áo này đều là phế phẩm.

Bài 32. Có 3 xạ thủ cùng bắn vào một con thú, 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) Giả sử con thú bị tiêu diệt. Tính xác suất nó bị trúng 2 phát đạn.

Đáp số: a) 0,7916; b) 0,4567.

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. 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) 0,0893; b) 0,4091.

Bài 35. 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: 0,044.

Công thức Bernoulli

Bài 36. 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) 0,389.

Bài 37. Một thiết bị có 10 chi tiết với độ tin cậy của mỗi chi tiết là 0,9. (Xác suất làm việc tốt trong khoảng thời gian nào đó).

Tính xác suất để trong khoảng thời gian ấy :

a) Có đúng một chi tiết làm việc tốt,

b) Có ít nhất 2 chi tiết làm việc tốt.

Đáp số: a) $9\cdot 10^{-9}$ *; b)* ≈ 1.

Bài 38. 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 thành công 2 dễ hơn.

Bài 39. 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) 0,99986.

Bài 40. 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.

Bài 41. Tỷ lệ mắc bệnh Basedow ở một vùng rừng núi nào đó là 70%. 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) 0,1528; b) 0,12826; c) 0,999.

Bài 42. 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: $n \geq 59$.

Bài 43. 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 44. 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) 0,0041; b) 2.

Chương 2. BIẾN SỐ NGẪU NHIÊN

2.1. Tóm tắt lý thuyết

2.1.1. Biến số ngẫu nhiên rời rạc

a) Bảng phân phối xác suất

Biến số ngẫu nhiên rời rạc X được xác định bằng *bảng phân phối xác suất*

X	x_1	x_2	…	x_n	…
P	p_1	p_2	…	p_n	…

trong đó $x_1 < x_2 < ... < x_n < ...$ là các giá trị nhận được bởi X và $p_i = P(X = x_i)$, với mọi i.

b) Hàm xác suất

Hàm số f(x) được gọi là hàm xác suất biến số ngẫu nhiên rời rạc X, nếu f(x) được xác định như sau:

$$f(x) = \begin{cases} p_i & \text{khi} \quad x = x_i \\ 0 & \text{khi} \quad x \neq x_i, \forall i \end{cases}$$

c) Hàm phân phối xác suất

Hàm số F(x) được gọi là hàm phân phối (xác suất) của biến số ngẫu nhiên rời rạc X, nếu F(x) được xác định như sau:

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

d) Trung bình và phương sai

Giá trị trung bình (kỳ vọng) của X cho bởi

$$\mu_X = E(X) = \sum_{x_i} x_i f(x_i) = \sum_i x_i p_i ,$$

và phương sai của X là

$$\sigma_X^2 = Var(X) = \sum_{x_i} (x_i - \mu_X)^2 f(x_i) = \sum_i (x_i - \mu_X)^2 p_i .$$

Căn bậc hai của phương sai gọi là *độ lệch chuẩn*,

$$\sigma_X = se(X) = \sqrt{\sigma_X^2} .$$

2.1.2. Biến số ngẫu nhiên liên tục

a) Hàm mật độ xác suất

Hàm số $f : \mathbb{R} \to \mathbb{R}$ được gọi là hàm mật độ (xác suất) biến số ngẫu nhiên liên tục X, nếu f(x) được xác định như sau:

$$P(a \le X \le b) = \int_a^b f(x)dx \text{, với } a, b \in \mathbb{R},\ a < b.$$

b) Hàm phân phối xác suất

Hàm số $F : \mathbb{R} \to \mathbb{R}$ được gọi là hàm phân phối (xác suất) của biến số ngẫu nhiên liên tục X, nếu F(x) được xác định như sau:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$

c) Trung bình và phương sai

Giá trị trung bình (kỳ vọng) của X cho bởi

$$\mu_X = E(X) = \int_{-\infty}^{+\infty} xf(x)dx,$$

và phương sai của X là

$$\sigma_X^2 = Var(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x)dx. \qquad \text{(Var : Variance)}$$

Căn bậc hai của phương sai gọi là *độ lệch chuẩn,*

$$\sigma_X = se(X) = \sqrt{\sigma_X^2}. \qquad \text{(Se : Standard error)}$$

2.2. Bài tập mẫu

Biến số ngẫu nhiên rời rạc

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.

<u>Giải</u>

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

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

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. Có hai lô sản phẩm.

- Lô 1: Có 8 chính phẩm và 2 phế phẩm
- Lô 2: Có 7 chính phẩm và 3 phế phẩm

Từ lô 1 lấy ngẫu nhiên 2 sản phẩm bỏ sang lô 2, sau đó từ lô 2 lấy ra 2 sản phẩm.

a) Tìm bảng phân phối xác suất của số chính phẩm được lấy ra

b) Tìm hàm phân phối xác suất của số chính phẩm được lấy ra

Giải

a) Xét các biến cố

A : "nhận được 2 phế phẩm từ lô 1",

B : "nhận được 1 chính phẩm và 1 phế phẩm từ lô 1",

C : "nhận được 2 chính phẩm từ lô 1",

và gọi X là số chính phẩm lấy ra từ lô 2. Ta có X lấy các giá trị 0, 1 và 2.

Nhận xét A, B, C là họ đầy đủ các biến cố

$$\begin{aligned}P(X=0)&=P(X=0\mid A)P(A)+P(X=0\mid B)P(B)+P(X=0\mid C)P(C)\\ &=\frac{C_5^2}{C_{12}^2}\cdot\frac{C_2^2}{C_{10}^2}+\frac{C_8^0C_4^2}{C_{12}^2}\cdot\frac{C_8^1C_2^1}{C_{10}^2}+\frac{C_3^2}{C_{12}^2}\cdot\frac{C_8^2}{C_{10}^2}\\ &=\frac{10}{66}\cdot\frac{1}{45}+\frac{6}{66}\cdot\frac{16}{45}+\frac{3}{66}\cdot\frac{28}{45}=\frac{190}{2970}\end{aligned}$$

$$\begin{aligned}P(X=1)&=P(X=1\mid A)P(A)+P(X=1\mid B)P(B)+P(X=1\mid C)P(C)\\ &=\frac{C_7^1C_5^1}{C_{12}^2}\cdot\frac{C_2^2}{C_{10}^2}+\frac{C_8^1C_4^1}{C_{12}^2}\cdot\frac{C_8^1C_2^1}{C_{10}^2}+\frac{C_9^1C_3^1}{C_{12}^2}\cdot\frac{C_8^2}{C_{10}^2}\\ &=\frac{35}{66}\cdot\frac{1}{45}+\frac{32}{66}\cdot\frac{16}{45}+\frac{27}{66}\cdot\frac{28}{45}=\frac{1303}{2970}\end{aligned}$$

$$P(X=2)=P(X=2\mid A)P(A)+P(X=2\mid B)P(B)+P(X=2\mid C)P(C)$$
$$=\frac{C_7^2}{C_{12}^2}\cdot\frac{C_2^2}{C_{10}^2}+\frac{C_8^2}{C_{12}^2}\cdot\frac{C_8^1C_2^1}{C_{10}^2}+\frac{C_9^2}{C_{12}^2}\cdot\frac{C_8^2}{C_{10}^2}$$
$$=\frac{21}{66}\cdot\frac{1}{45}+\frac{28}{66}\cdot\frac{16}{45}+\frac{36}{66}\cdot\frac{28}{45}=\frac{1477}{2970}$$

Từ đó, ta được bảng phân phối xác suất

X	0	1	2
P	$\frac{190}{2970}$	$\frac{1303}{2970}$	$\frac{1477}{2970}$

b) Tìm hàm phân phối xác suất của số chính phẩm được lấy ra

$$F(x)=\begin{cases}0 & \text{khi } x<0\\ \frac{190}{2970} & \text{khi } 0\le x<1\\ \frac{1493}{2970} & \text{khi } 1\le x<2\\ 1 & \text{khi } 2\le x\end{cases}$$

Bài 3. Thực hiện ba lần bắn bia với xác suất trúng bia tương ứng là 0,3; 0,4; 0,6. Tìm trung bình và phương sai của số lần bắn trúng bia.

<u>Giải</u>

Gọi A_i "biến cố bắn trúng bia lần thứ i"

Ta có: $P(A_1)=0,2;\ P(A_2)=0,3;\ P(A_3)=0,6$

Gọi X là số lần bắn trúng bia trong 3 lần bắn, X lấy các giá trị 0, 1, 2, 3. Chú ý rằng $A_1,\ A_2,\ A_3$ là các biến cố độc lập. Ta có

$$P(X=0)=P(\bar{A}_1\bar{A}_2\bar{A}_3)=P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)$$
$$=0,7\times0,6\times0,4=0,168,$$

$$P(X=1)=P(A_1\bar{A}_2\bar{A}_3+\bar{A}_1A_2\bar{A}_3+\bar{A}_1\bar{A}_2A_3)$$
$$=P(A_1\bar{A}_2\bar{A}_3)+P(\bar{A}_1A_2\bar{A}_3)+P(\bar{A}_1\bar{A}_2A_3)$$

$$= P(A_1)P(\bar{A}_2)P(\bar{A}_3) + P(\bar{A}_1)P(A_2)P(\bar{A}_3) + P(\bar{A}_1)P(\bar{A}_2)P(A_3)$$
$$= 0,3 \times 0,6 \times 0,4 + 0,7 \times 0,4 \times 0,4 + 0,7 \times 0,6 \times 0,6 = 0,436,$$

$$P(X = 2) = P(A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3)$$
$$= P(A_1A_2\bar{A}_3) + P(A_1\bar{A}_2A_{33}) + P(\bar{A}_1A_2A_3)$$
$$= P(A_1)P(A_2)P(\bar{A}_3) + P(A_1)P(\bar{A}_2)P(A_3) + P(\bar{A}_1)P(A_2)P(A_3)$$
$$= 0,3 \times 0,4 \times 0,4 + 0,3 \times 0,6 \times 0,6 + 0,7 \times 0,4 \times 0,6 = 0,324$$
$$P(X = 3) = P(A_1A_2A_3) = P(A_1)P(A_2)P(A_3) = 0,3 \times 0,4 \times 0,6 = 0,072.$$

Từ đó, ta được bảng phân phối xác suất

X	0	1	2	3
P	0,168	0,436	0,324	0,072

Trung bình $EX = 1,3$; phương sai $\sigma_X^2 = 0,69$.

Biến số ngẫu nhiên liên tục

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.

<u>Giải</u>

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

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

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

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

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

$$= 3,10^{-9}10^5\left(\frac{10^4}{3} - \frac{11376}{5}\right) = \frac{992}{3125} = 0,31744.$$

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

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

$$= \frac{P\left(X \geq 60\right)}{P\left(X \geq 50\right)} = \frac{0,31744}{0,5} = 0,63488,$$

với $P(X \geq 50)$ được tính như ở phần c và bằng 0,5.

Bài 5. Cho biến số ngẫu nhiên liên tục X có hàm mật độ xác suất

$$f(x) = \begin{cases} \dfrac{\sin x}{2} & \text{khi } x \in (0,\pi) \\ 0 & \text{khi } x \notin (0,\pi) \end{cases}$$

a) Tìm hàm phân phối xác suất của X.

b) Tìm $P\left(0 \leq X \leq \dfrac{\pi}{4}\right)$

c) Tìm trung bình và phương sai của X.

<u>Giải</u>

a) Tìm hàm phân phối xác suất của X.

Trường hợp 1. Nếu $x \leq 0$ thì $F(x) = \int_{-\infty}^{x} f(t)dt = 0$.

Trường hợp 2. Nếu $0 < x \le \pi$

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt$$
$$= \int_{0}^{x} \frac{\sin t}{2} dt = -\frac{\cos t}{2}\Big|_{0}^{x} = \frac{1-\cos x}{2}$$

Trường hợp 3. Nếu $\pi < x < +\infty$

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{\pi} f(t)dt + \int_{\pi}^{x} f(t)dt$$
$$= \int_{0}^{\pi} \frac{\sin t}{2} dt = -\frac{\cos t}{2}\Big|_{0}^{\pi} = 1$$

Vậy hàm phân phối xác suất của X là

$$F(x) = \begin{cases} 0 & \text{khi } x \le 0 \\ \dfrac{1-\cos x}{2} & \text{khi } 0 < x \le \pi \\ 1 & \text{khi } \pi < x \end{cases}$$

b) $P\left(0 \le X \le \dfrac{\pi}{4}\right) = \dfrac{1-\cos x}{2}\Big|_{0}^{\frac{\pi}{4}} = \dfrac{1}{2} - \dfrac{\sqrt{2}}{4}$

c) Trung bình và phương sai

Trung bình

$$EX = \int_{-\infty}^{+\infty} xf(x)dx = \frac{1}{2}\int_{0}^{\pi} x\sin xdx$$

Đặt

$$u = x \ \Rightarrow du = dx$$
$$dv = \sin xdx \Rightarrow v = -\cos x$$

$$EX = -\frac{1}{2}x\cos x\Big|_{0}^{\pi} + \frac{1}{2}\int_{0}^{\pi}\cos xdx$$
$$= -\frac{1}{2}x\cos x + \frac{1}{2}\sin x\Big|_{0}^{\pi} = \frac{\pi}{2}$$

Phương sai: $\text{Var}(X) = EX^2 - \left(EX\right)^2$

Với

$$EX^2 = \frac{1}{2}\int_0^{\pi} x^2 \sin x dx$$

Đặt

$$u = x^2 \Rightarrow du = 2xdx$$
$$dv = \sin x dx \Rightarrow v = -\cos x$$

$$EX^2 = -\frac{1}{2}x^2 \cos x\Big|_0^{\pi} + \int_0^{\pi} x\cos x dx = \frac{\pi^2}{2} + I$$

Tính $I = \int_0^{\pi} x\cos x dx$

Đặt

$$u = x \Rightarrow du = dx$$
$$dv = \cos x dx \Rightarrow v = \sin x$$

$$I = x\sin x\Big|_0^{\pi} - \int_0^{\pi} \sin x dx = x\sin x + \cos x\Big|_0^{\pi} = -2$$

Vậy $Var(X) = \frac{\pi^2}{2} - 2 - \left(\frac{\pi}{2}\right)^2 = \frac{\pi^2}{4} - 2$

2.3. Bài tập rèn luyện

Biến số ngẫu nhiên rời rạc

Bài 1. Cho X là một biến số 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) $a = 0,1$*; b)* $P\left[X \geq 5\right] = 0,2$; $P\left[X < 3\right] = 0,3$*; c)* $k = 3$.

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

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

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

Đáp số: a) $A = 1$; *b)* $A = 2$.

Bài 3. 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 biến số 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 4. 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. Tính E(X); $\mathrm{Var}(X)$ và Mod(X).

Đáp số: $EX = 2,2$; $\mathrm{Var}(X) = 0,54$; $\mathrm{Mod}(X) = 2$.

Bài 5. Một phân xưởng có ba máy M_1, M_2, M_3. Trong một giờ, mỗi máy sản xuất được 10 sản phẩm. Số sản phẩm không đạt tiêu chuẩn trong 10 sản phẩm của M_1, M_2, M_3 lần lượt là 1, 2, 1. Lấy ngẫu nhiên 1 sản phẩm từ 10 sản phẩm do mỗi máy sản xuất. Gọi X là số sản phẩm không đạt tiêu chuẩn trong 3 sản phẩm được lấy ra.

a) Lập bảng phân phối xác xuất của X.

b) Tìm E(X), Var(X), Mod(X).

c) Tính $P(X \leq 1)$.

Đáp số: a)

X	0	1	2	3
P	0,648	0,306	0,044	0,002

b) $\mathrm{EX} = 0,4;\ \mathrm{Var(X)} = 0,34; \mathrm{Mod(X)} = 0$; c) 0,954.

Bài 6. Tỷ lệ khách hàng phản ứng tích cực đối với một chiến dịch quảng cáo là biến số ngẫu nhiên có bảng phân phối xác suất như sau :

X (%)	0	10	20	30	40	50
P	0,1	0,2	0,35	0,2	0,1	0,05

a) Tìm tỷ lệ khách hàng bình quân phản ứng tích cực đối với chiến dịch quảng cáo đó.

b) Tìm xác suất để có trên 20% khách hàng phản ứng tích cực đối với chiến dịch quảng cáo.

Đáp số: a) 21,5%; b) 0,35.

Bài 7. Qua theo dõi trong nhiều năm kết hợp với sự đánh giá của các chuyên gia tài chính thì lãi suất đầu tư vào một công ty là biến số ngẫu nhiên có bảng phân phối xác suất như sau:

X (%)	9	10	11	12	13	14	15
P	0,05	0,15	0,3	0,2	0,15	0,1	0,05

a) Tính xác suất để khi đầu tư vào công ty đó thì sẽ đạt được lãi suất ít nhất là 12%.

b) Tính lãi suất kỳ vọng khi đầu tư vào công ty đó.

c) Mức độ rủi ro khi đầu tư vào công ty đó có thể đánh giá bằng cách nào?

Đáp số a) 0,5; b) $\mathrm{EX} = 11,75$ *; c)* $\sigma_X^2 = 2,2875$.

Bài 8. Cho biến số ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau:

X	1	4	8
P	0,3	0,1	0,6

Tính $P\left(\left|X - E(X)\right| < 4\right)$.

Đáp số: 0,7.

Bài 9. Lợi nhuận X thu được khi đầu tư 50 triệu đồng vào một dự án có bảng phân phối xác suất như sau (đơn vị : triệu đồng).

X	-2	-1	0	1	2	3
P	0,1	0,1	0,2	0,2	0,3	0,1

a) Tìm mức lợi nhuận có khả năng nhiều nhất khi đầu tư vào dự án đó.

b) Việc đầu vào dự án này có hiệu quả hay không? Tại sao?

c) Làm thế nào để đo được mức độ rủi ro của vụ đầu tư này? Hãy tìm mức độ rủi ro đó.

Đáp số: a) $\mathrm{Mod}(X) = 2$*; b)* $EX = 0,8$*;* $\sigma_X^2 = 3,24$.

Bài 10. Tại một cửa hàng bán xe máy Honda người ta thống kê được số xe máy bán ra hàng tuần (X) với bảng phân phối xác suất như sau :

X	0	1	2	3	4	5	6	7	8	9	10	11
P	0,05	0,12	0,17	0,08	0,12	0,2	0,07	0,02	0,07	0,02	0,03	0,05

a) Tìm số xe trung bình bán được mỗi tuần.

b) Tìm phương sai và độ lệch chuẩn của số xe bán được mỗi tuần và giải thích ý nghĩa của kết quả nhận được.

Đáp số: a) 4,33; b) $\sigma_X^2 = 8,3411; \sigma_X = 2,89$.

Bài 11. Sản phẩm nhà máy được đóng thành từng hộp, mỗi hộp có 10 sản phẩm. Gọi X là số sản phẩm loại một có trong hộp. Cho biết X có bảng phân phối xác suất như sau:

X	6	7
P	0,7	0,3

Khách hàng chọn cách kiểm tra để mua hàng như sau : Từ mỗi hộp lấy ngẫu nhiên ra 3 sản phẩm để kiểm tra, nếu thấy có ít nhất 2 sản phẩm loại một thì mua hộp đó. Lấy ngẫu nhiên 3 hộp để kiểm tra. Tính xác suất để có 2 hộp được mua.

Đáp số: 0,381.

Biến số ngẫu nhiên liên tục

Bài 12. Tuổi thọ của một loại bóng đèn nào đó là một biến số ngẫu nhiên liên tục 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 hỏng trước khi nó được 1 năm tuổi.

Đáp số: a) $k = \dfrac{3}{64}$*; b) 0,0508.*

Bài 13. Khối lượng của một con vịt 6 tháng tuổi là một biến số 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, tính

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

(ii) tỷ lệ vịt chậm lớn, biết vịt 6 tháng tuổi chậm lớn là vịt có khối lượng nhỏ hơn 2Kg,

(iii) hàm phân phối xác suất của X.

Đáp số: a) $k = \dfrac{3}{20}$*; b)* $EX = 2,4$*;* $P\left(X < 2\right) = 0,2$

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

Bài 14. 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 của X.

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

Đáp số: a) $a = 0,5;\ F(x) = \begin{cases} 0 & \text{khi } x \le -\frac{\pi}{2} \\ \frac{\sin x + 1}{2} & \text{khi } -\frac{\pi}{2} < x \le \frac{\pi}{2} \\ 1 & \text{khi } x > \frac{\pi}{2} \end{cases}$

b) 0,1465.

Bài 15. Cho X là biến số ngẫu nhiên liên tục có hàm phân phối xác suất sau:

$$F(x) = \frac{1}{4} + \frac{1}{\pi}\arctan x$$

a) Tính $P(0 < X < 1)$.

b) Tìm hàm mật độ xác suất của X.

Đáp số: a) 0,25; b) $f(x) = \frac{1}{\pi(1 + x^2)}$.

Bài 16. Cho X là biến số ngẫu nhiên liên tục có hàm phân phối xác suất sau:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan\frac{x}{2}$$

Tìm giá trị x_1 thỏa mãn điều kiện: $P(X > x_1) = \frac{1}{4}$.

Đáp số: $x_1 = 2$.

Bài 17. Thời gian xếp hàng chờ mua hàng của khách là biến số ngẫu nhiên liên tục X với hàm phân phối xác suất như sau:

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

a) Tìm hệ số a.

b) Tìm thời gian trung bình.

c)Tìm xác suất để trong 3 người xếp hàng thì có không quá 2 người phải chờ quá 0,5 phút.

Đáp số: a) $a = 2$*; b)* $EX = 0,5$*; c) 0,875.*

Bài 18. Tỷ lệ mắc một loại bệnh trong một vùng dân cư là biến số ngẫu nhiên liên tục X có hàm mật độ như sau:

$$f(x) = \begin{cases} \dfrac{1}{20} & \text{khi } x \in [5,25] \\ 0 & \text{khi } x \notin [5,25] \end{cases}$$

a) Tính $P(|X-10| > 2,5)$.

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

Đáp số: a) 0,75; b) $EX = 15$; $\sigma_X^2 = 33,3$.

Bài 19. Tuổi thọ (tính theo giờ) của một trò chơi điện tử bấm tay là một biến số ngẫu nhiên liên tục có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} ke^{-\frac{x}{100}} & \text{khi } x \ge 0 \\ 0 & \text{khi } x < 0 \end{cases}$$

a) Tìm hằng số k.

b) Tính xác suất tuổi thọ của trò chơi này nằm trong khoảng từ 50 đến 150 giờ.

c) Tính xác suất tuổi thọ của trò chơi này ít hơn 100 giờ.

Đáp số: a) $k = 0,01$*; b) 0,239; c) 0,632.*

Bài 20. Cho biến số ngẫu nhiên liên tục X có hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} x^m \dfrac{e^{-x}}{m!} & \text{khi } x \ge 0 \\ 0 & \text{khi } x < 0 \end{cases}$$

Tính trung bình và phương sai của X.

Đáp số : $EX = \sigma_X^2 = m+1$.

Chương 3. PHÂN PHỐI XÁC SUẤT

1. Phân phối nhị thức $B(n;p)$

1.1. Định nghĩa

Biến số ngẫu nhiên rời rạc X được gọi là phân phối nhị thức, ký hiệu $X \sim B(n;p)$ nếu hàm xác suất của X có dạng sau

$$f(x) = \begin{cases} C_n^x p^x (1-p)^{n-x} & \text{khi } x = 0,1,2,\dots,n \\ 0 & \text{khi } x \neq 0,1,2,\dots,n \end{cases}$$

Công thức xác suất

$$P\left(X = k\right) = C_n^k p^k (1-p)^{n-k}, \text{ với } k = 0,1,2,\dots,n$$

1.2. Mệnh đề:

Cho $X \sim B(n;p)$, ta có

i) Trung bình: $\mu_X = np$,

ii) Phương sai: $\sigma_X^2 = np(1-p)$,

iii) Giá trị tin chắc: $Mod(X) = k_0$ thỏa $np - q \le k_0 \le np - q + 1$, với $q = 1 - p$.

2. Phân phối siêu bội $H(N,K,n)$

2.1. Định nghĩa

Biến số ngẫu nhiên rời rạc X được gọi là phân phối siêu bội, ký hiệu $X \sim H(N,K,n)$ nếu hàm xác suất của X có dạng sau

$$f(x) = \begin{cases} \dfrac{C_K^x C_{N-K}^{n-x}}{C_N^n} & \text{khi } x \in \left[\max\{0, n-N+K\}, \min\{n,K\}\right] \\ 0 & \text{khi } x \notin \left[\max\{0, n-N+K\}, \min\{n,K\}\right] \end{cases}$$

Công thức xác suất

$$P\left(X = k\right) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}, \text{ với } \max\{0, n-N+K\} \le k \le \min\{n,K\}$$

2.2. Mệnh đề:

Cho $H(N,K,n)$, ta có

i) Trung bình: $\mu_X = np$,

ii) Phương sai: $\sigma_X^2 = np(1-p)\left(\dfrac{N-n}{N-1}\right)$.

3. Phân phối Poisson $P(\mu)$

3.1. Định nghĩa

Biến số ngẫu nhiên rời rạc X được gọi là phân phối poisson, ký hiệu $X \sim P(\mu)$ nếu hàm xác suất của X có dạng sau

$$f(x) = \begin{cases} e^{-\mu}\dfrac{\mu^x}{x!} & \text{khi } x = 0,1,2,\ldots,n \\ 0 & \text{khi } x \neq 0,1,2,\ldots,n \end{cases}$$

Công thức xác suất

$$P\left(X = k\right) = e^{-\mu}\frac{\mu^k}{k!}, \text{ với } k = 0,1,2,\ldots,n$$

3.2. Mệnh đề:

Cho $X \sim P(\mu)$, ta có

i) Trung bình: $\mu_X = \mu$,

ii) Phương sai: $\sigma_X^2 = \mu$,

iii) Độ lệch chuẩn: $\sigma_X = \sqrt{\mu}$.

3.3. Chú ý: Nếu $X \sim B(n,p)$, trong đó p đủ nhỏ và n đủ lớn thì X được xem như có phân phối Poisson $X \sim P(\mu)$, với $\mu = np$.

Bằng cách viết

$$\begin{aligned} C_n^k p^k (1-p)^{n-k} &= \frac{n(n-1)\ldots(n-k+1)}{k!} p^k (1-p)^{n-k} \\ &= \frac{1}{k!} \cdot \frac{n(n-1)\ldots(n-k+1)}{n^k} \left(np\right)^k (1-p)^{n-k} \end{aligned}$$

và với $\mu = np$ không đổi, khi $n \to \infty$, ta có $p \to 0$ và

$$\lim_{n\to\infty} \frac{n(n-1)\ldots(n-k+1)}{n^k} = 1$$

$$\lim_{n\to\infty}(1-p)^{n-k} = \lim_{p\to 0}(1-p)^{\frac{\mu}{p}-k} = e^{-\mu}$$

Vậy

$$C_n^k p^k (1-p)^{n-k} \approx e^{-\mu} \frac{\mu^k}{k!}$$

4. Phân phối Chuẩn $N\left(\mu,\sigma^2\right)$

4.1. Định nghĩa

Biến số ngẫu nhiên liên tục X được gọi là phân phối chuẩn, ký hiệu $X \sim N\left(\mu,\sigma^2\right)$ nếu hàm mật độ của X có dạng sau

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty,$$

Công thức xác suất

$$P\left(a \le X \le b\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \text{ với } a, b \in \mathbb{R},\ a \le b$$

4.2. Mệnh đề:

Cho $X \sim N\left(\mu,\sigma^2\right)$, ta có

i) Trung bình: $\mu_X = \mu$,

ii) Phương sai: $\sigma_X^2 = \sigma^2$,

iii) Độ lệch chuẩn: $\sigma_X = \sigma$.

4.3. Chú ý:

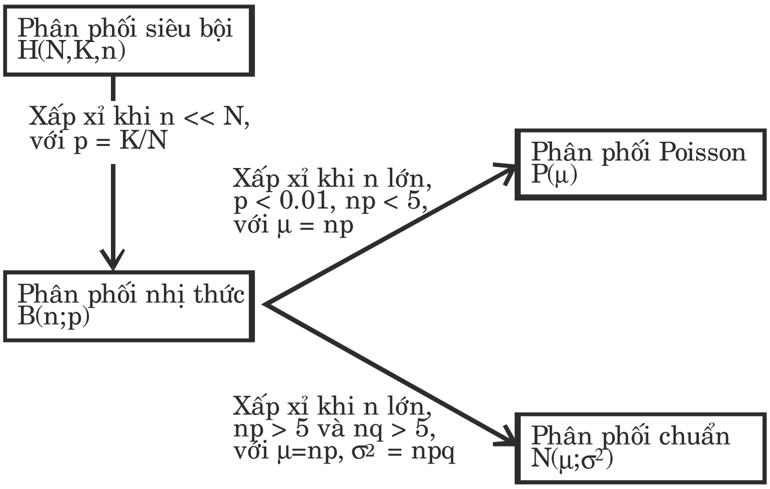
i) Nếu $X \sim N\left(\mu,\sigma^2\right)$ thì đặt $Y = \dfrac{X-\mu}{\sigma}$, ta có $Y \sim N(0,1)$ (Y: phân phối Gauss hay là phân phối chuẩn tắc). Do đó, với $a, b \in \mathbb{R},\ a \le b$, ta có

$$P\left(a \le X \le b\right) = P\left(\frac{a-\mu}{\sigma} \le Y \le \frac{b-\mu}{\sigma}\right) = \varphi\left(\frac{b-\mu}{\sigma}\right) - \varphi\left(\frac{a-\mu}{\sigma}\right)$$

với $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^x e^{-\frac{t^2}{2}} dx$

ii) Phân phối chuẩn dùng để khảo sát các hiện tượng bình thường. Cụ thể, nếu $X \sim B\left(n;p\right)$ với tích np lớn thì ta xấp xỉ phân phối nhị thức $B\left(n;p\right)$ bằng phân phối chuẩn $N\left(\mu;\sigma^2\right)$, với $\mu = np,\ \sigma^2 = npq$.

iii) Sự liên hệ giữa các phân phối nhị thức, siêu bội, Poisson và chuẩn được cho trong sơ đồ sau:



5. Phân phối Gamma và phân phối chi bình phương

Định nghĩa: hàm Gamma $\Gamma : \left(0, \infty\right) \to \mathbb{R}$

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$$

5.1. Định nghĩa

Biến số ngẫu nhiên liên tục X được gọi là phân phối Gamma, ký hiệu $X \sim \Gamma\left(\alpha, \beta\right)$, với $\alpha, \beta > 0$, nếu hàm mật độ của X có dạng sau

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} & \text{khi } x > 0 \\ 0 & \text{khi } x \le 0 \end{cases},$$

5.2. Mệnh đề

Cho $X \sim \Gamma(\alpha, \beta)$, ta có

i) Trung bình: $\mu_X = \alpha\beta$,

ii) Phương sai: $\sigma_X^2 = \alpha\beta^2$.

5.3. Định nghĩa

Biến số ngẫu nhiên liên tục X được gọi là phân phối chi bình phương, ký hiệu $X \sim \chi^2(r)$, nếu hàm mật độ của X có dạng sau

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

nghĩa là $X \sim \Gamma\left(\dfrac{r}{2},2\right)$

5.4. Mệnh đề

Cho $X \sim \chi^2(r)$, ta có

i) Trung bình: $\mu_X = r$,

ii) Phương sai: $\sigma_X^2 = 2r$.

6. Phân phối Student $St(n)$

6.1. Định nghĩa

Cho X là biến số ngẫu nhiên liên tục có phân phối Gauss, $X \sim N(0,1)$; Y là biến số ngẫu nhiên liên tục có phân phối Chi bình phương với n bậc tự do, $Y \sim \chi^2(n)$ và X, Y là hai biến số độc lập.

Đặt

$$T=\frac{X}{\sqrt{Y/n}}$$

thì T có phân phối Student với n bậc tự do, $T \sim St(n)$.

6.2. Mệnh đề

Cho $T \sim St(n)$, ta có

i) Trung bình: $\mu_T = 0$,

ii) Phương sai: $\sigma_T^2 = \dfrac{n}{n-2}$.

6.3. Chú ý : Nếu $X \sim St(n)$, với $n \geq 30$, thì $X \sim N(0,1)$.

7. Phân phối Fisher $F(n,m)$

7.1. Định nghĩa

Cho X, Y là hai biến số ngẫu nhiên liên tục có phân phối Chi bình phương, $X \sim \chi^2(n)$, $Y \sim \chi^2(m)$ và X, Y là hai biến số độc lập.

Đặt

$$F = \frac{X/n}{Y/m}$$

thì F có phân phối Fisher với n, m bậc tự do, $F \sim F(n,m)$.

7.2. Mệnh đề

Cho $F \sim F(n,m)$, ta có

i) Trung bình: $\mu_F = \dfrac{m}{m-2}$,

ii) Phương sai: $\sigma_F^2 = \dfrac{2m^2(n+m^2-2)}{n(m-2)^2(m-4)}$.

3.1. 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.

<u>Giải</u>

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 $\geq 0,9$.

<u>Giải</u>

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

$$\begin{aligned} P(X=1) &= C_{10}^{1}(0,07)^{1}(1-0,07)^{10-1} \\ &= 10\cdot 0,07\cdot (0,93)^{9} = 0,3643. \end{aligned}$$

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 $\geq 0,9$. 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 tính

$$1-(0,93)^{n} \geq 0,9.$$

Giải bất phương trính trên, ta nhận được giá trị $n \geq 31,73$. Vậy phải quan sát ít nhất 32 sản phẩm.

Bài 3. Trong số 20 công nhân của một công ty có 12 người có tay nghề khá. Tìm xác suất để kiểm tra ngẫu nhiên tay nghề của 5 công nhân thì có ít nhất 3 người có tay nghề khá.

Giải

Gọi X là số công nhân có tay nghề trong 5 công nhân kiểm tra thì $X \sim H(20,12,5)$.

Xác suất để có ít nhất 3 công nhân có tay nghề là

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5)$$
$$= \frac{C_{12}^3 C_8^2}{C_{20}^5} + \frac{C_{12}^4 C_8^1}{C_{20}^5} + \frac{C_{12}^5 C_8^0}{C_{20}^5} = \frac{10912}{15504} = 0,704.$$

Bài 4. Để thanh toán 1 triệu đồng tiền hàng, một khách hàng gian lận đã xếp lẫn 5 tờ 50 ngàn đồng tiền giả với 15 tờ tiền thật. Chủ cửa hàng rút ngẫu nhiên 3 tờ giấy bạc đem đi kiểm tra và giao hẹn nếu phát hiện có bạc giả thì cứ mỗi tờ giả khách hàng phải đền hai tờ thật. Tìm số tiền phạt mà khách có thể phải trả.

Giải

Gọi X là số tờ giả trong 3 tờ rút ra thì $X \sim H(20,5,3)$.

Bảng phân phối xác suất của X

X	0	1	2	3
P	$\frac{455}{1140}$	$\frac{525}{1140}$	$\frac{150}{1140}$	$\frac{10}{1140}$

Gọi Y là số tiền bị phạt, ta có $Y = 100X$

Y	0	100	200	300
P	$\frac{455}{1140}$	$\frac{525}{1140}$	$\frac{150}{1140}$	$\frac{10}{1140}$

Trung bình số tiền bị phạt: $EY = 75$ ngàn đồng.

Bài 5. 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 6. 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

$$X \sim P(2000 \cdot 0,001) = P(2).$$

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

$$\begin{aligned}P(X \le 3) &= P(X=0)+P(X=1)+P(X=2)+P(X=3)\\ &= e^{-2}+2e^{-2}+e^{-2}+\frac{4}{3}e^{-2}=\frac{16}{3}e^{-2}=0,722.\end{aligned}$$

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

$$\begin{aligned}P(X>3) &= 1-P(X\le 3)\\ &= 1-\frac{16}{3}e^{-2}=0,278.\end{aligned}$$

Bài 7. 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 $P(np)$.

<u>Giải</u>

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

$$\begin{aligned}P(X=0) &= C_{20}^{0}\left(0,01\right)^{0}\left(1-0,01\right)^{20}\\ &= \left(0,99\right)^{20}=0,8179.\end{aligned}$$

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

$$\begin{aligned}P(X=1) &= C_{20}^{1}\left(0,01\right)^{1}\left(1-0,01\right)^{20-1}\\ &= 20\cdot\left(0,01\right)\cdot\left(0,99\right)^{19}=0,1652.\end{aligned}$$

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

$$\begin{aligned}P(X>1) &= 1-\left[P(X=0)+P(X=1)\right]\\ &= 1-(0,8179+0,1652)=0,0168.\end{aligned}$$

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

$$\begin{aligned} P(X>1) &= 1 - \left[P(X=0) + P(X=1)\right] \\ &= 1 - (0,8187 + 0,1637) = 0,01755. \end{aligned}$$

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 8. 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

$$\begin{aligned} P(49,9 \le X \le 50,1) &= P\left(\frac{49,9-50}{0,05} \le \frac{X-50}{0,05} \le \frac{50,1-50}{0,05}\right) \\ &= P(-2 \le X \le 2) = \varphi(2) - \varphi(-2) = 2\varphi(2) = 0,9544. \end{aligned}$$

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

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - C_3^0 (0,9544)^0 (1 - 0,9544)^3 \\ &= 1 - (0,0456)^3 = 0,9999. \end{aligned}$$

Bài 9. 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

$$\begin{aligned} P(X > 505) &= P\left(\frac{X-500}{4} > \frac{505-500}{4}\right) \\ &= P(X > 1,25) = \varphi(+\infty) - \varphi(1,25) = 0,5 - \varphi(1,25) = 0,10565. \end{aligned}$$

b) Tỷ lệ trái cây loại 2 là

$$\begin{aligned} P(495 \leq X \leq 505) &= P\left(\frac{495-500}{4} \leq \frac{X-500}{4} \leq \frac{505-500}{4}\right) \\ &= P(-1,25 \leq Y \leq 1,25) = 0,7887. \end{aligned}$$

c) Và tỷ lệ của loại 3 là

$$\begin{aligned} P(X < 495) &= P(\frac{X-500}{4} < \frac{495-500}{4}) \\ &= P(Y < -1,25) = \varphi(-1,25) - \varphi(-\infty) \\ &= -\varphi(1,25) + 0,5 = 0,10565. \end{aligned}$$

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

Bài 10. Một máy sản xuất ra sản phẩm loại A với xác suất 0,485. Tính xác suất trong 200 sản phẩm do máy sản xuất ra có ít nhất 95 sản phẩm loại A.

Giải

Gọi X là số sản phẩm loại A trong 200 do máy sản xuất ra. Ta có $X \sim B(200; 0,485)$

Ta có $np = 200 \cdot 0,485 = 97 > 5$ và $nq = 200 \cdot 0,515 = 103 > 5$

$$\sigma_X^2 = npq = 200 \cdot 0,485 \cdot 0,515 = 49,955 \approx (7,07)^2.$$

Ta xấp xỉ phân phối nhị thức bằng phân phối chuẩn, nghĩa là $X \sim N\left(97,(7,07)^2\right)$. Với $Y = \dfrac{X-97}{7,07}$ thì $Y \sim N(0;1)$. Do đó xác suất có ít nhất 95 sản phẩm loại A

$$\begin{aligned} P(X \ge 95) &= P(95 \le X \le 200) \\ &= P\left(\frac{95-97}{7,07} \le \frac{X-97}{7,07} \le \frac{200-97}{7,07}\right) \\ &= P(-0,28 \le Y \le 14,57) \\ &= \varphi(14,57) - \varphi(-0,28) = 0,5 + 0,1103 = 0,6103. \end{aligned}$$

3.2. Bài tập rèn luyện

Bài 1. 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 2. 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 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 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. Biết rằng số cuộc gọi trong một khoảng thời gian cố định có phân phối Poisson. 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 5. 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) $\mathrm{EX} = 5$; $\mathrm{Mod}(\mathrm{X}) = 5$.

Bài 6. 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 7. 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 8. 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 9. Giả sử xác suất trúng số là 1%. Mỗi tuần mua một vé số. Hỏi phải mua vé số liên tiếp trong tối thiểu bao nhiêu tuần để có không ít hơn 95% hy vọng trúng số ít nhất 1 lần.

Đáp số: 296.

Bài 10. 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 11. 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: b) 0,9525.

Bài 12. 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.

1) Lập bảng phân phối xác suất của số sản phẩm tốt trong 3 sản phẩm lấy ra.

2) 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

a) có ít nhất 80 kiện hàng được mua,

b) có ít nhất 60 kiện được mua.

Đáp số: a)

X	0	1	2	3
P	0	0,066	0,467	0,467

b) 0,0038.

Bài 13. 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 ?

Đáp số: a) 20,76 ; b) ; c) .

Bài 14. 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 15. Trong hệ thống tỷ giá hối đoái thả nổi, sự biến động của tỷ giá hối đoái chịu sự tác động của nhiều nhân tố và có thể xem như là biến ngẫu nhiên có phân phối chuẩn. Giả sử ở giai đoạn nào đó tỷ giá của USD với VND có trung bình là 18000đ và độ lệch chuẩn là 800đ. Tìm xác suất để trong một ngày nào đó.

a) Tỷ giá sẽ cao hơn 19000đ,

b) Tỷ giá sẽ thấp hơn 17500đ,

c) Tỷ giá nằm trong khoảng từ 17500đ đến 19500.

Đáp số: a) 0,1056; b) 0,2643; c) 0,7342.

Bài 16. Khối lượng của một gói đường (đóng bằng máy tự động) có phân phối chuẩn. Trong 1000 gói đường có 70 gói có khối lượng lớn hơn 1015. Hãy ước lượng xem có bao nhiêu gói đường có khối lượng ít hơn 1008g. Biết rằng khối lượng trung bình của 1000 gói đường là 1012g.

Đáp số: 24,4 gói.

Bài 17. Lãi suất (%) đầu tư vào một dự án năm 2000 được coi như 1 đại lượng ngẫu nhiên có 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: 0,9987.

Bài 18. 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: $P(X_1 \geq 80) = 0,8849 < P(X_2 \geq 80) = 0,9772$, *chọn phương án thứ 2.*

Bài 19. Độ dài của một chi tiết máy được tiện ra có phân phối chuẩn $N(\mu\,cm; (0,2cm)^2)$. 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 đạt yêu cầu.

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

c) Nếu sản phẩm tốt mà bị loại trong kiểm tra thì mắc phải sai lầm loại 1, nếu sản phẩm không đạt mà được nhận thì mắc phải sai lầm loại 2. giả sử khả năng mắc phải sai lầm loại 1, loại 2 lần lượt là 0,1 và 0,2. Tính xác suất để trong 3 lần kiểm tra hoàn toàn không nhầm lẫn.

Đáp số a) 0,8664; b) 0,9512; c) 0,697.

Bài 20. Khối lượng của 1 loại trái cây có quy luật phân phối chuẩn với khối lượng trung bình là 250g, độ lệch chuẩn về khối lượng là 5g.

a) Một người lấy 1 trái từ trong sọt trái cây ra. Tính xác suất người này lấy được trái loại 1 (trái loại 1 là trái có khối 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) 0,0228; b) 0,02.

Bài 21. Có hai thị trường A và B, lãi suất của cổ phiếu trên hai thị trường này là các đại lượng ngẫu nhiên có 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

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

b) Để tránh rủi ro thì nên đầu tư vào cổ phiếu trên cả hai thị trường theo tỷ lệ như thế nào?

Đáp số: a) nên đầu tư vào cổ phiếu trên thị trường loại A.

b) 74% vào thị trường A còn lại là thị trường B.

Bài 22. 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 chiều cao trên 180cm,

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

c) Giá trị h_0, nếu biết rằng 33% người trưởng thành có chiều cao 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 23. 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) $\left(7,7.10^{-3}\right)^2$.

Bài 24. Khối lượng X của một loại trái cây ở nông trường được biết có kỳ vọng 250gr và phương sai $81\left(\text{gr}\right)^2$. Trái cây được đóng thành sọt, mỗi sọt 100 trái. Mỗi sọt được gọi là loại A nếu khối lượng không dưới 25kg. Kiểm tra ngẫu nhiên 100 sọt. Tính xác suất :

a) có nhiều nhất 60 sọt loại A,

b) ít nhất 45 sọt loại A.

Đáp số: a) 0,0228; b) 0,1587.

Bài 25. Việc kiểm tra các viên bi được tiến hành như sau: nếu viên bi không lọt qua lỗ có đường kính d_1 song lọt qua lỗ có đường kính d_2 thì viên bi được coi là đạt tiêu chuẩn, nếu không thì viên bi bị loại. Biết đường kính các viên bi sản xuất ra là biến ngẫu nhiên có phân phối chuẩn với trung bình là $\frac{d_1 + d_2}{2}$ và độ lệch chuẩn là $\frac{d_2 - d_1}{4}$. Tìm xác suất để viên bi bị loại.

Đáp số: 0,0456.

Chương 4. MẪU VÀ ƯỚC LƯỢNG THAM SỐ

4.1. Tóm tắt lý thuyết

Đối với số liệu tổng thể có N phần tử, $X_1, X_2, ..., X_N$, người ta quan tâm đến các tham số sau:

- Trung bình tổng thể: $\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$,

- Phương sai tổng thể: $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \mu\right)^2$,

- Tỷ lệ tổng thể: $p = \frac{K}{N}$, trong đó K là các X_i thỏa một tính chất nào đó.

Khi không có số liệu cho toàn bộ tổng thể, người ta dùng một mẫu $X_1, X_2, ..., X_n$, trích ta từ tổng thể, và từ đó ta có các tham số tính toán trên mẫu, người ta tìm cách ước lượng các tham số tổng thể. Có hai loại ước lượng:

4.1.1. Ước lượng điểm

Ước lượng điểm tốt nhất của trung bình tổng thể, μ, là trung bình mẫu:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

Ước lượng điểm tốt nhất của phương sai tổng thể, σ^2, là phương sai mẫu có hiệu chỉnh:

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

Ước lượng điểm tốt nhất của tỷ lệ tổng thể, p, là tỷ lệ mẫu:

$$f = \frac{k}{n},$$

trong đó k là X_i thỏa tính chất tương ứng trên mẫu.

4.1.2. Ước lượng khoảng

a) Ước lượng trung bình tổng thể μ khi biết phương sai tổng thể $\sigma^2 = \sigma_0^2$:

Ta dùng thống kê

$$Z = \frac{\left(\bar{X} - \mu\right)\sqrt{n}}{\sigma_0} \sim N\left(0,1\right)$$

Trong thống kê này, $\bar{X}$ (trung bình mẫu), n (cỡ mẫu) đã biết, σ_0 (độ lệch chuẩn tổng thể) cho trước. Với độ tin cậy γ cho trước, ta có $C = Z_{\frac{\gamma}{2}}$

Khoảng ước lượng của trung bình tổng thể μ, được ký hiệu

$$\mu = \bar{X} \pm C\frac{\sigma_0}{\sqrt{n}} \text{ hay } \mu \in \left[\bar{X} - C\frac{\sigma_0}{\sqrt{n}}, \bar{X} + C\frac{\sigma_0}{\sqrt{n}}\right]$$

b) Ước lượng trung bình tổng thể μ khi chưa biết phương sai tổng thể σ^2:

Ta dùng thống kê

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

Trong thống kê này, $\bar{X}$ (trung bình mẫu), n (cỡ mẫu), S_X (độ lệch chuẩn mẫu có hiệu chỉnh) đã biết. Với độ tin cậy γ cho trước, ta có $\alpha = 1 - \gamma$ suy ra $C = t_{\alpha}^{n-1}$

Khoảng ước lượng của trung bình tổng thể μ, được ký hiệu

$$\mu = \bar{X} \pm C\frac{S_X}{\sqrt{n}} \text{ hay } \mu \in \left[\bar{X} - C\frac{S_X}{\sqrt{n}}, \bar{X} + C\frac{S_X}{\sqrt{n}}\right]$$

c) Ước lượng phương sai tổng thể σ^2 khi biết trung bình tổng thể $\mu = \mu_0$:

Ta dùng thống kê

$$Y = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \mu_0\right)^2 \sim \chi^2(n).$$

Trong thống kê này, X_i (số liệu của mẫu), n (cỡ mẫu) đã biết, μ_0 (trung bình tổng thể) cho trước. Với độ tin cậy γ cho trước, ta chọn khoảng tin cậy cho Y : $\left[a, b\right]$

Với $a = \chi^2_{1-\frac{\alpha}{2}}(n),\ b = \chi^2_{\frac{\alpha}{2}}(n)$

Khoảng ước lượng của phương sai tổng thể σ^2

$$\sigma^2 \in \left[\frac{1}{b}\sum_{i=1}^{n}\left(X_i-\mu_0\right)^2; \frac{1}{a}\sum_{i=1}^{n}\left(X_i-\mu_0\right)^2\right]$$

d) Ước lượng phương sai tổng thể σ^2 khi chưa biết trung bình tổng thể μ :

Ta dùng thống kê

$$Y=\frac{(n-1)S_X^2}{\sigma^2}\sim\chi^2(n-1).$$

Trong thống kê này, S_X^2 (phương sai mẫu có hiệu chỉnh), n (cỡ mẫu) đã biết, với độ tin cậy γ cho trước, ta chọn khoảng tin cậy cho Y : $\left[a,b\right]$, với

$a=\chi^2_{1-\frac{\alpha}{2}}(n-1),\ b=\chi^2_{\frac{\alpha}{2}}(n-1)$

Khoảng ước lượng của phương sai tổng thể σ^2

$$\sigma^2 \in \left[\frac{(n-1)S_X^2}{b},\frac{(n-1)S_X^2}{a}\right]$$

e) Ước lượng tỷ lệ tổng thể p :

Ta dùng thống kê

$$T=\frac{\left(f-p\right)\sqrt{n}}{\sqrt{f(1-f)}}\sim St(n-1).$$

Trong thống kê này, f (tỷ lệ mẫu), n (cỡ mẫu) đã biết. Với độ tin cậy γ cho trước, ta có

$\alpha=1-\gamma$ suy ra $C=t_{\alpha}^{n-1}$

Khoảng ước lượng của tỷ lệ tổng thể p, được ký hiệu

$$p=f\pm C\sqrt{\frac{f(1-f)}{n}} \text{ hay } p\in\left[f-C\sqrt{\frac{f(1-f)}{n}};f+C\sqrt{\frac{f(1-f)}{n}}\right]$$

4.2. Bài tập mẫu

Bài 1. Phân tích Vitamin của 17 mẫu, ta được $\overline{X}=20mg$. Biết rằng lượng Vitamin có phân phối chuẩn $N\left(\mu;\sigma^2\right)$ với $\sigma=3,98mg$.

a) Hãy ước lượng lượng Vitamin trung bình với độ tin cậy 95%.

b) Nếu muốn sai số ước lượng không quá 1mg ở độ tin cậy 95% thì phải quan sát ít nhất mấy trường hợp.

Giải

a) Để ước lượng trung bình tổng thể μ khi biết phương sai tổng thể ta dùng thống kê

$$Z = \frac{\left(\overline{X} - \mu\right)\sqrt{n}}{\sigma_0} \sim N(0,1).$$

Với số liệu mẫu, ta có

$$Z = \frac{\left(20 - \mu\right)\sqrt{17}}{3,98} \sim N(0,1).$$

Ở độ tin cậy $\gamma = 0,95$, ta tìm được $C = 1,96$. Do đó ước lượng lượng Vitamin trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{\sigma_0}{\sqrt{n}} = 20 \pm 1,96\frac{3,98}{\sqrt{17}},$$

và ta nhận được khoảng ước lượng $\left[18,11; 21,89\right]$.

b) Ta có sai số ước lượng là $C\frac{\sigma_0}{\sqrt{n}}$ nên nếu muốn sai số ước lượng không quá 1mg, ta phải có $C\frac{\sigma_0}{\sqrt{n}} \leq 1$

Với độ tin cậy 0,95, thì $C = 1,96$ và ta nhận được bất phương trình

$$n \geq \left(C\frac{\sigma_0}{1}\right)^2 = \left(1,96\frac{3,98}{1}\right)^2 = 60,85.$$

Vậy phải phân tích ít nhất 61 trường hợp.

Bài 2. Doanh số của một cửa hàng là biến số ngẫu nhiên có phân phối chuẩn với độ lệch chuẩn là 2 triệu/tháng. Điều tra ngẫu nhiên doanh số của 600 cửa hàng có quy mô tương tự nhau tìm được doanh số trung bình là 8,5 triệu. Với độ tin cậy 95% hãy ước lượng doanh số trung bình của các cửa hàng thuộc quy mô đó.

Giải

Để ước lượng trung bình tổng thể μ khi biết phương sai tổng thể ta dùng thống kê

$$Z = \frac{\left(\overline{X} - \mu\right)\sqrt{n}}{\sigma_0} \sim N(0,1).$$

Với số liệu mẫu, ta có

$$Z = \frac{\left(8,5 - \mu\right)\sqrt{600}}{2} \sim N(0,1).$$

Ở độ tin cậy $\gamma = 0,95$, ta tìm được $C = 1,96$. Do đó ước lượng doanh số trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{\sigma_0}{\sqrt{n}} = 8,5 \pm 1,96\frac{2}{\sqrt{600}},$$

và ta nhận được khoảng ước lượng $\left[8,34; 8,66\right]$.

Bài 3. Đo chiều sâu của biển bằng một loại dụng cụ có sai số chuẩn tuân theo quy luật chuẩn với phương sai bằng 400 (m^2). Phải đo bao nhiêu lần độc lập với nhau để kết quả có sai số không quá 15 m với độ tin cậy 95%.

Giải

Ta có sai số ước lượng là $C\frac{\sigma_0}{\sqrt{n}}$ nên nếu muốn sai số ước lượng không quá 15 m, ta phải có $C\frac{\sigma_0}{\sqrt{n}} \le 15$

Với độ tin cậy 0,95, thì $C = 1,96$ và ta nhận được bất phương trình

$$n \ge \left(C\frac{\sigma_0}{15}\right)^2 = \left(1,96\frac{20}{15}\right)^2 = 6,83.$$

Vậy phải đo ít nhất 7 lần.

Bài 4. Đo đường kính X(mm) 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

a) Tính trung bình mẫu và độ lệch chuẩn của mẫu có hiệu chỉnh.

b) Ước lượng đường kính trung bình tổng thể μ ở độ tin cậy 95%.

c) Nếu muốn sai số ước lượng không quá 0,02mm ở độ tin cậy 95% thì phải quan sát ít nhất mấy trường hợp.

<u>Giải</u>

a) Ta được cỡ mẫu $n = 53,$ trung bình mẫu $\overline{X} = 12,21,$ độ lệch chuẩn mẫu có hiệu chỉnh $S_X = 0,103$

b) Để ước lượng trung bình tổng thể μ khi chưa biết phương sai tổng thể ta dùng thống kê

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

Với số liệu mẫu, ta có

$$T = \frac{\left(12,21 - \mu\right)\sqrt{53}}{0,103} \sim St(52) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,95$, ta tìm được $C = 1,96$. Do đó ước lượng đường kính trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{S_X}{\sqrt{n}} = 12,21 \pm 1,96\frac{0,103}{\sqrt{53}},$$

và ta nhận được khoảng ước lượng $\left[12,18; 12,24\right]$.

c) Ta có sai số ước lượng là $C\frac{S_X}{\sqrt{n}}$ nên nếu muốn sai số ước lượng không quá 0,02mm, ta phải có $C\frac{S_X}{\sqrt{n}} \le 0,02$

Với độ tin cậy 0.95, thì $C = 1,96$ và ta nhận được bất phương trình

$$n \ge \left(C\frac{S_X}{0,02}\right)^2 = \left(1,96\frac{0,103}{0,02}\right)^2 = 101,89.$$

Vậy phải quan sát ít nhất 102 trường hợp.

Bài 5. Cân ngẫu nhiên 45 con heo 3 tháng tuổi trong một trại chăn nuôi, ta được kết quả sau

X_i	35	37	39	41	43	45	47
n_i	2	6	10	11	8	5	3

Giả sử khối lượng X (kg) tuân theo quy luật phân phối chuẩn.

a) Hãy ước lượng khoảng cho khối lượng trung bình các con heo 3 tháng tuổi trong trại trên với độ tin cậy 95%.

b) Heo có khối lượng $\geq 38\text{kg}$ là heo đạt tiêu chuẩn. Hãy tìm ước lượng tỷ lệ heo đạt chuẩn với độ tin cậy 90%.

<u>Giải</u>

Với số liệu, ta có : cỡ mẫu $n = 45$, trung bình $\overline{X} = 40,96$, phương sai $S_X^2 = 9,73$.

a) Để ước lượng trung bình tổng thể μ khi chưa biết phương sai tổng thể ta dùng thống kê

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

Với số liệu mẫu, ta có

$$T = \frac{\left(40,96 - \mu\right)\sqrt{45}}{3,12} \sim St(44) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,95$, ta tìm được $C = 1,96$. Do đó ước lượng khối lượng trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{S_X}{\sqrt{n}} = 40,96 \pm 1,96\frac{3,12}{\sqrt{45}},$$

và ta nhận được khoảng ước lượng $\left[40,5\text{kg}; 41,87\text{kg}\right]$.

b) Tỷ lệ heo đạt tiêu chuẩn là $f = \frac{10+11+8+5+3}{45} \approx 0,8222$.

Để ước tỷ lệ tổng thể p, ta dùng thống kê

$$T = \frac{\left(f - p\right)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1).$$

Với số liệu mẫu, ta có

$$T = \frac{\left(0,8222 - p\right)\sqrt{45}}{\sqrt{0,8222(1-0,8222)}} \sim St(44) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,9$, ta tìm được $C = 1,64$. Do đó ước lượng tỷ lệ p cho bởi

$$p = f \pm C \cdot \sqrt{\frac{f(1-f)}{n}} = 0,8222 \pm 1,64\sqrt{\frac{0,8222(1-0,8222)}{45}},$$

và ta nhận được khoảng ước lượng $\left[0,7282; 0,9162\right]$

Bài 6. 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 ?

<u>Giải</u>

a) Từ các số liệu nhận được của mẫu, ta có

$n = 12$, $\overline{X} = 137,83$, $S_X^2 = 19,42$, và $S_X = 4,41$.

b) Để ước lượng trung bình tổng thể μ, ta dùng thống kê

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

Với số liệu mẫu, ta có

$$T = \frac{\left(137,83 - \mu\right)\sqrt{12}}{4,41} \sim St\left(11\right).$$

Với độ tin cậy $\gamma = 0,95$, ta có $C = t_{0.05}^{11} = 2,201$. Do đó ước lượng trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{S_X}{\sqrt{n}} = 137,83 \pm 2,201\frac{4,41}{\sqrt{12}},$$

và ta nhận được khoảng ước lượng $[135,01;140,63]$.

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

$$a \equiv \chi^2_{1-\frac{\alpha}{2}}(11) = 3,816, \text{ và } b \equiv \chi^2_{\frac{\alpha}{2}}(11) = 21,925.$$

Do đó

$$3,816 \le \frac{(n-1)S_X^2}{\sigma^2} \le 21,925,$$

và ta nhận được bất đẳng thức

$$\frac{11\times(19,42)}{21,925} \le \sigma^2 \le \frac{11\times(19,42)}{3,816}$$

Từ đó suy ra ước lượng cho phương sai tổng thể là $[9,74;55,98]$.

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 7. Khảo sát chỉ tiêu X của một loại sản phẩm, người ta quan sát một mẫu và có kết quả sau

X(cm)	11 - 15	15 - 19	19 – 23	23 - 27	27 - 31	31 - 35	35 - 39
Số sản phẩm	8	9	20	16	16	13	18

Giả sử X có phân phối chuẩn. Hãy ước lượng phương sai của X với độ tin cậy 95% trong các trương hợp sau:

a) Biết trung bình tổng thể của X là 25 cm.
b) Chưa biết giá trị trung bình của X.

Giải

a) Để ước lượng phương sai tổng thể khi biết trung bình tổng thể $\mu_0 = 25$, ta dùng thống kê

$$Y = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \mu_0\right)^2 \sim \chi^2(n).$$

Với độ tin cậy $\gamma = 0,95$, ta chọn khoảng tin cậy cho Y : $\left[a, b\right]$

Với $a = \chi^2_{0.975}(100) = 77,929,\ b = \chi^2_{0.025}(100) = 129,561$

Ta lập bảng

$X - \mu_0$	-12	-8	-4	0	4	8	12
N	8	9	20	16	16	13	18

Từ đó ta tìm được cỡ mẫu $n = 100,\ \sum n_i\left(X_i - \mu_0\right)^2 = 5728$

Khoảng ước lượng của phương sai tổng thể σ^2

$$\sigma^2 \in \left[\frac{1}{b}\sum_{i=1}^{n} n_i\left(X_i-\mu_0\right)^2, \frac{1}{a}\sum_{i=1}^{n} n_i\left(X_i-\mu_0\right)^2\right] = \left[44,211;73,503\right]$$

b) 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

$$Y = \frac{99\times\left(55,991\right)}{\sigma^2} \sim \chi^2\left(99\right).$$

Với độ tin cậy $\gamma = 0,95$, ta tìm được a và b sao cho

Từ bảng phân phối xác suất của phân phối Chi-Bình phương, ta tìm được

$$a = \chi^2_{0.975}(99) = 77,929,\ b = \chi^2_{0.025}(99) = 129,561.$$

Khoảng ước lượng của phương sai tổng thể σ^2

$$\sigma^2 \in \left[\frac{(n-1)S_X^2}{b}, \frac{(n-1)S_X^2}{a}\right] = \left[42,784;71,130\right]$$

Bài 8. Trong kho có 10000 hộp thịt. Kiểm tra ngẫu nhiên 100 hộp thấy có 5 hộp bị hỏng. Với độ tin cậy 95%, tính xem trong kho có khoảng bao nhiêu hộp bị hỏng.

<u>Giải</u>

Gọi N là số hộp thịt bị hỏng ở trong kho

Tỷ lệ (tổng thể) những hộp bị hỏng: $p = \dfrac{N}{10000}$.

Tỷ lệ (mẫu) những hộp bị hỏng: $f = \dfrac{5}{100} = 0,05$.

Để ước tỷ lệ tổng thể p, ta dùng thống kê

$$T = \frac{\left(f-p\right)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1).$$

Với số liệu mẫu, ta có

$$T = \frac{\left(0,05 - p\right)\sqrt{100}}{\sqrt{0,05(1 - 0,05)}} \sim St(99) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,95$, ta tìm được $C = 1,96$. Do đó ước lượng tỷ lệ p cho bởi

$$p = f \pm C \cdot \sqrt{\frac{f(1-f)}{n}} = 0,05 \pm 1,96\sqrt{\frac{0,05(1-0,05)}{100}},$$

và ta nhận được khoảng ước lượng $p \in \left[0,00728; 0,09272\right]$.

Vậy số hộp bị hỏng nằm trong khoảng $N \in \left[73; 927\right]$.

Bài 9. Trong kho có 1000 sản phẩm của xí nghiệp A sản xuất bỏ lẫn với nhiều sản phẩm của xí nghiệp B sản xuất. Lấy ngẫu nhiên từ kho ra 100 sản phẩm thì thấy có 9 sản phẩm của xí nghiệp A sản xuất. Với độ tin cậy 90% hãy ước lượng số sản phẩm của xí nghiệp B sản xuất có ở trong kho.

<u>Giải</u>

Gọi N là số sản phẩm của xí nghiệp B sản xuất có ở trong kho

Tỷ lệ (tổng thể) sản phẩm của xí nghiệp A sản xuất có ở trong kho: $p = \dfrac{1000}{1000 + N}$.

Tỷ lệ (mẫu) sản phẩm của xí nghiệp A sản xuất có ở trong kho: $f = \dfrac{9}{100} = 0,09$.

Để ước tỷ lệ tổng thể p, ta dùng thống kê

$$T = \frac{\left(f - p\right)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1).$$

Với số liệu mẫu, ta có

$$T = \frac{\left(0,09 - p\right)\sqrt{100}}{\sqrt{0,09(1 - 0,09)}} \sim St(99) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,9$, ta tìm được $C = 1,64$. Do đó ước lượng tỷ lệ p cho bởi

$$p = f \pm C \cdot \sqrt{\frac{f(1-f)}{n}} = 0,09 \pm 1,64\sqrt{\frac{0,09(1-0,09)}{100}},$$

và ta nhận được khoảng ước lượng $p \in \left[0,0431; 0,1369\right]$.

Vậy số hộp bị hỏng nằm trong khoảng $N \in [6305; 22202]$.

Bài 10. Số liệu thống kê về doanh số bán hàng của một siêu thị trong 7 tháng qua là :

Doanh số (triệu đ/ngày)	20 - 30	30 - 35	35 - 40	40 - 45	45 - 50	50- 55	55 – 60	60 - 70
Số ngày	10	25	30	40	38	30	15	8

a) Ước lượng doanh số bán trung bình trong một ngày của siêu thị này với độ tin cậy 95%.

b) Những ngày có doanh số bán trên 50 triệu là những ngày đắt hàng. Hãy ước lượng số ngày bán đắt hàng ở siêu thị này trong một năm (360 ngày) với độ tin cậy 99%.

Giải

Với số liệu, ta có : cỡ mẫu $n = 196$, trung bình $\overline{X} = 44,133$, phương sai $S_X = 9,382$.

a) Để ước lượng trung bình tổng thể μ khi chưa biết phương sai tổng thể ta dùng thống kê

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

Với số liệu mẫu, ta có

$$T = \frac{\left(44,133 - \mu\right)\sqrt{196}}{9,382} \sim St(195) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,95$, ta tìm được $C = 1,96$. Do đó ước lượng khối lượng trung bình μ cho bởi

$$\mu = \overline{X} \pm C\frac{S_X}{\sqrt{n}} = 44,133 \pm 1,96\frac{9,382}{\sqrt{196}},$$

và ta nhận được khoảng ước lượng $[42,82; 45,45]$.

b) Tỷ lệ những ngày bán đắt hàng trong năm là $f = \frac{30+15+8}{196} \approx 0,27041$.

Để ước tỷ lệ tổng thể p, ta dùng thống kê

$$T = \frac{\left(f - p\right)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1).$$

Với số liệu mẫu, ta có

$$T = \frac{(0,27041 - p)\sqrt{196}}{\sqrt{0,27041(1 - 0,27041)}} \sim St(195) \equiv N(0,1).$$

Ở độ tin cậy $\gamma = 0,99$, ta tìm được $C = 2,58$. Do đó ước lượng tỷ lệ p cho bởi

$$p = f \pm C \cdot \sqrt{\frac{f(1-f)}{n}} = 0,27041 \pm 2,58\sqrt{\frac{0,27041(1 - 0,27041)}{196}},$$

và ta nhận được khoảng ước lượng $[0,1130; 0,4278]$.

Gọi N là số ngày bán đắt hàng trong năm

Tỷ lệ (tổng thể) những ngày bán đắt hàng trong năm: $p = \frac{N}{360}$.

Vậy số ngày bán đắt hàng trong năm : $N \in [41;154]$.

4.3. Bài tập rèn luyện

Bài 1. 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.

Bài 2. 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 điểm độ dài trung bình và phương sai của trục xe.

Đáp số: $\overline{X} = 19,133$; $S_X^2 = 0,539$.

Bài 3. Đo sức bền chịu lực của một loại ống thí nghiệm, người ta thu được bộ số liệu sau

4500	6500	5200	4800	4900	5125	6200	5375

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

Đáp số: $\mu \in [5151; 5499]$.

Bài 4. 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: 67%.

Bài 5. a) Muốn ước lượng tỷ lệ bệnh sốt xuất huyết ở Tp. Hồ Chí Minh với sai số không quá 3% ở độ tin cậy 95% thì phải quan sát ít nhất bao nhiêu người ?

b) 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 Tp. Hồ Chí Minh ở độ tin cậy 97%. Nếu muốn sai số ước lượng không quá 3% ở độ tin cậy 95% thì phải quan sát ít nhất bao nhiêu người ?

Đáp số: a) 1068 người; b) $p \in [0,1132; 0,2868]$*; 683 người.*

Bài 6. 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 người.

Bài 7. 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: $\left[34865; 87719\right]$.

Bài 8. 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: $\left[784; 27812\right]$.

Bài 9. 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 ước lượng sản lượng trung bình và phương sai mỗi ngày, với độ tin cậy 95%.

Đáp số: $\mu \in \left[23,75; 27,81\right]$; $\sigma^2 \in \left[3,166; 25,468\right]$.

Bài 10. Trên tập mẫu gồm 100 số liệu, người ta tính được $\overline{X} = 0,1;\ S_X = 0,014$. Hãy ước lượng giá trị trung bình tổng thể, với độ tin cậy 95%.

Đáp số: $\mu \in \left[0,0973; 0,103\right]$.

Bài 11. 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 tỷ lệ cam loại 2 với độ tin cậy 90% .

Đáp số: a) $\mu \in \left[35,539; 36,241\right]$*; b)* $p \in \left[0,0143; 0,0857\right]$.

Bài 12. 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: 0,8828.

Bài 13. 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 luật chuẩn với $\sigma = 14$ ngàn đồng. Với độ tin cậy 95%, hãy ước lượng mức lương trung bình của công nhân trong toàn xí nghiệp.

Đáp số: $\mu \in [375,427; 384,573]$.

Bài 14. Điểm trung bình môn toán của 100 thí sinh dự thi vào ĐHKT là 5 với độ lệch chuẩn mẫu đã điều chỉnh $S_X = 2,5$.

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

b) Với sai số 0,25 điểm. Hãy xác định độ tin cậy.

Đáp số: a) $\mu \in [4,51; 5,49]$*; b) 68,26%.*

Bài 15. 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 là 15 giờ. Hãy xác định độ tin cậy.

c) Với độ chính xác là 25 giờ và độ tin cậy là 95% thì cần thử nghiệm bao nghiêu bóng.

Đáp số: a) $\mu \in [980,4; 1019,6]$*; b) 86,64%; c) 62.*

Bài 16. Khối lượng các bao bột mì tại một cửa hàng lương thực theo quy luật chuẩn. Kiểm tra 20 bao, thấy khối lượng trung bình của mỗi bao bột mì là 48kg, và phương sai mẫu có điều 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 là 0,26kg. Hãy xác định độ tin cậy.

c) Với độ chính xác là 160g và độ tin cậy là 95%, tính cỡ mẫu.

Đáp số: a) $\mu \in [47,766; 48,234]$*; b) 97%; c) 43.*

Bài 17. 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 $\varepsilon = 3\%$, hãy xác định độ tin cậy.

Đáp số: a) $p \in [0,051; 0,169]$*; b) 66,3%.*

Bài 18. 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% và độ chính xác 1% thì cần kiểm tra bao nhiêu sọt.

d) Muốn ước lượng tỷ lệ trái cây không đạt tiêu chuẩn với độ tin cậy 99% thì độ chính xác đạt được là bao nhiêu?

Đáp số: a) $p \in [0,082; 0,098]$*; b) 78,5%; c) 0,012; d) 55 sọt.*

Bài 19. 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
Diện tích (ha)	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 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 20. 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 đường kính trung bình của những chi tiết đạt tiêu chuẩn với độ tin cậy 95%.

b) Ước lượng tỷ lệ chi tiết đạt tiêu chuẩn với độ tin cậy 95%.

c) Muốn ước lượng đường kính trung bình của chi tiết đạt tiêu chuẩn muốn độ chính xác đạt 0,02mm và khi ước lượng tỷ lệ chi tiết đạt tiêu chuẩn muốn độ chính xác là 5%, với cùng độ tin cậy là 99% thì cần đo thêm bao nhiêu chi tiết nữa.

Đáp số: a) $\mu \in [19,986; 20,008]$*; b)* $p \in [0,733; 0,887]$*; c) 310.*

Bài 21. 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 $\overline{X} = 0,47\text{cm}$ và $S_x = 0,032$ cm. Tìm khoảng tin cậy cho phương sai và trung bình chuẩn 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 22. Lấy 28 mẫu xi măng của một nhà máy sản xuất 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) $\mu \in [11,64; 12,64]$; $\sigma^2 \in [1,156; 3,427]$.

Chương 5. KIỂM ĐỊNH GIẢ THUYẾT

5.1. Tóm tắt lý thuyết.

5.1.1.Kiểm định tham số.

Quan sát mẫu X_1, X_2, ..., X_n độc lập và có cùng phân phối chuẩn $N(\mu;\sigma^2)$, ta có một số bài toán kiểm định tham số sau :

5.1.1.1. So sánh trung bình tổng thể μ với một số μ_0 cho trước.

Nếu biết phương sai tổng thể $\sigma^2 = \sigma_0^2$, ta có mô hình

a) $\begin{cases} H : \mu = \mu_0 \\ \overline{H} : \mu \neq \mu_0 \end{cases}$

Khi H đúng, ta có

$$Z = \frac{(\overline{X} - \mu_0)\sqrt{n}}{\sigma_0} \sim N(0;1)$$

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1 - \alpha$, cho trước, ta tìm được khoảng tin cậy $[-C, C]$ cho Z.

c) Do $\overline{X}$, n có được từ mẫu, μ_0, σ_0 cho trước nên ta tính được giá trị cụ thể của Z.

d) So sánh Z với khoảng tin cậy :

$|Z| > C$: bác bỏ H,

$|Z| \leq C$: chấp nhận H.

Nếu chưa biết phương sai tổng thể σ^2, ta dùng mô hình

a) $\begin{cases} H : \mu = \mu_0 \\ \overline{H} : \mu \neq \mu_0 \end{cases}$

Khi H đúng, ta có

$$T \equiv \frac{\left(\overline{X} - \mu_0\right)\sqrt{n}}{S_X} \sim St\left(n-1\right).$$

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1-\alpha$, cho trước, ta tìm được khoảng tin cậy $[-C, C]$ cho T.

c) Do $\overline{X}$, S_X, n có được từ mẫu, μ_0 cho trước nên ta tính được giá trị cụ thể của T.

d) So sánh T với khoảng tin cậy :

$|T| > C$: bác bỏ H,

$|T| \le C$: chấp nhận H.

5.1.1.2. So sánh tỷ lệ tổng thể, p, với một số p_0 cho trước.

Ta dùng mô hình

a) $\begin{cases} H : p = p_0 \\ \overline{H} : p \neq p_0 \end{cases}$

Khi H đúng, ta có

$$Z = \frac{(f - p_0)\sqrt{n}}{\sqrt{p_0 q_0}} \sim N(0;1),$$

với $q_0 = 1 - p_0$ và f là tỷ lệ của mẫu.

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1-\alpha$, cho trước, ta tìm được khoảng tin cậy $[-C, C]$ cho Z.

c) Do f, n có được từ mẫu, p_0 cho trước nên ta tính được giá trị cụ thể của Z.

d) So sánh Z với khoảng tin cậy :

$|Z| > C$: bác bỏ H,

$|Z| \le C$: chấp nhận H.

5.1.1.3. So sánh hai trung bình μ_X và μ_Y của hai tổng thể.

Để so sánh trung bình của hai tổng thể thỏa phân phối chuẩn, ta giả sử chúng có cùng phương sai σ^2 chưa biết và dựa vào hai mẫu quan sát độc lập lấy từ hai tổng thể này,

$$X_1,\ X_2, \ldots,\ X_n;\ X_i \sim N\left(\mu_X;\sigma^2\right);$$

$$Y_1,\ Y_2, \ldots,\ Y_m;\ Y_i \sim N\left(\mu_Y;\sigma^2\right).$$

Ta so sánh hai trung bình μ_X và μ_Y bằng cách dùng mô hình

a) $\begin{cases} H: \mu_X = \mu_Y \\ \overline{H}: \mu_X \neq \mu_Y \end{cases}$

Khi H đúng, ta có

$$T = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim St(n + m - 2),$$

với $S^2 = \dfrac{(n-1)S_X^2 + (m-1)S_Y^2}{n + m - 2}$.

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1 - \alpha$, cho trước, ta tìm được khoảng tin cậy $[-C, C]$ cho T.

c) Từ hai mẫu, ta tính được các giá trị S_X^2, S_Y^2, S^2, n, m, $\overline{X}$, $\overline{Y}$ và suy ra giá trị cụ thể của T.

d) So sánh T với khoảng tin cậy :

$|T| > C$: bác bỏ H,

$|T| \leq C$: chấp nhận H.

5.1.1.4. So sánh hai tỷ lệ p_X và p_Y của hai tổng thể.

Để so sánh tỷ lệ của hai tổng thể, ta cũng dựa vào các tỷ lệ lấy ra từ hai mẫu quan sát độc lập từ hai tổng thể này,

$$X_1,\ X_2, \ldots,\ X_n;$$

$Y_1,\ Y_2,\ ...,\ Y_m;$

trong đó X_i, Y_j chỉ lấy các giá trị là 0 hay 1. Khi đó, $f_X = \frac{1}{n}\sum_{i=1}^{n} X_i$ là tỷ lệ (tần suất) của mẫu X và $f_Y = \frac{1}{m}\sum_{j=1}^{m} Y_j$ là tỷ lệ (tần suất) của mẫu Y.

Ta so sánh hai tỷ lệ p_X và p_Y bằng cách dùng mô hình

a) $\begin{cases} H: p_X = p_Y \\ \overline{H}: p_X \neq p_Y \end{cases}$

Khi H đúng, ta có

$$T = \frac{f_X - f_Y}{\sqrt{pq\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim St(n+m-2),$$

với $p = \frac{nf_X + mf_Y}{n+m}$, $q = 1 - p$.

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1 - \alpha$, cho trước, ta tìm được khoảng tin cậy $[-C, C]$ cho T.

c) Từ hai mẫu, ta tính được các giá trị f_X, f_Y, p, q, n, m và suy ra giá trị cụ thể của T.

d) So sánh T với khoảng tin cậy :

$|T| > C$: bác bỏ H,

$|T| \leq C$: chấp nhận H.

5.1.1.5. So sánh hai phương sai σ_X^2 và σ_Y^2 của hai tổng thể.

Để so sánh phương sai của hai tổng thể thỏa phân phối chuẩn, ta dựa vào hai mẫu quan sát độc lập lấy từ hai tổng thể này,

$X_1,\ X_2,\ ...,\ X_n;\ X_i \sim N(\mu_X; \sigma_X^2);$

$Y_1,\ Y_2,\ ...,\ Y_m;\ Y_i \sim N(\mu_Y; \sigma_Y^2).$

Từ các mẫu, ta tính được các phương sai mẫu S_X^2, S_Y^2 và không mất tính tổng quát, ta có thể giả sử $S_X \geq S_Y$. Ta so sánh hai phương sai σ_X^2 và σ_Y^2 bằng cách dùng mô hình

a) $\begin{cases} H : \sigma_X^2 = \sigma_Y^2 \\ \bar{H} : \sigma_X^2 > \sigma_Y^2 \end{cases}$

Khi H đúng, ta có

$$F = \frac{S_X^2}{S_Y^2} \sim F(n-1, m-1).$$

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1 - \alpha$, cho trước, ta tìm được khoảng tin cậy $[0, C]$ cho F.

c) Từ hai mẫu, ta tính được các giá trị S_X^2, S_Y^2 và suy ra giá trị cụ thể của F.

d) So sánh F với khoảng tin cậy :

$F > C$: bác bỏ H,

$F \leq C$: chấp nhận H.

5.1.2. Kiểm định phi tham số.

Trong trường hợp tổng thể được chia thành r phạm trù khác nhau. Khi đó, ứng với mỗi một mẫu, ta được một bộ số liệu gồm r số hạng mà ta gọi là bộ số liệu quan sát

$$N_1,\ N_2,\ ...,\ N_r.$$

Khi đó, các phép kiểm định phi tham số nhằm mục đích so sánh trực tiếp các bộ số liệu như vậy với nhau hay so sánh chúng với một bộ số liệu lý thuyết nào đó.

5.1.2.1. So sánh bộ số liệu quan sát với bộ số liệu lý thuyết.

Trong trường hợp này, với một bộ số liệu quan sát,

$$N_1,\ N_2,\ ...,\ N_r,$$

ta cần so sánh nó với bộ số liệu lý thuyết

$$N'_1,\ N'_2,\ ...,\ N'_r,$$

trong đó bộ số liệu lý thuyết này được tính theo quy luật phân phối các phạm trù trong tổng thể cho trước.

Khi đó, Định lý Pearson cho ta mô hình kiểm định sau

a) $\begin{cases} H: \text{Các bộ số liệu quan sát và lý thuyết giống nhau} \\ \bar{H}: \text{Các bộ số liệu quan sát và lý thuyết khác nhau} \end{cases}$

Với

$$Q = \sum_{i=1}^{r} \frac{(N_i - N'_i)^2}{N'_i} = \frac{(N_1 - N'_1)^2}{N'_1} + \frac{(N_2 - N'_2)^2}{N'_2} + \ldots + \frac{(N_r - N'_r)^2}{N'_r},$$

thì phân phối xác suất của Q cho bởi :

$Q \sim \chi^2(r-1)$, nếu không có tham số nào cần ước lượng trong quá trình tính các số liệu lý thuyết.

$Q \sim \chi^2(r-k-1)$, nếu có k tham số cần ước lượng trong quá trình tính các số liệu lý thuyết. Chẳng hạn, nếu ta cần ước lượng trung bình tổng thể μ, thì $Q \sim \chi^2(r-2)$; nếu ta cần ước lượng cả trung bình lẫn phương sai tổng thể $(\mu;\sigma^2)$ thì $Q \sim \chi^2(r-3)$.

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1-\alpha$, cho trước, ta tìm được khoảng tin cậy $[0, C]$ cho Q.

c) Tính giá trị cụ thể của Q.

d) So sánh Q với khoảng tin cậy :

$Q > C$: bác bỏ H,

$Q \leq C$: chấp nhận H.

5.1.2.2. So sánh các bộ số liệu quan sát với nhau.

Trong trường hợp này, ta so sánh k bộ số liệu quan sát

$N_{1,1}$, $N_{1,2}$, ..., $N_{1,r}$,

$N_{2,1}$, $N_{2,2}$, ..., $N_{2,r}$,

$N_{k,1}$, $N_{k,2}$, ..., $N_{k,r}$.

với nhau mà người ta còn gọi là so sánh các số liệu trong một bảng :

	P. trù 1	P. trù 2	...	P. trù r
Bộ 1	$N_{1,1}$	$N_{1,2}$	...	$N_{1,r}$
Bộ 2	$N_{2,1}$	$N_{2,2}$	...	$N_{2,r}$
...				
Bộ k	$N_{k,1}$	$N_{k,2}$	...	$N_{k,r}$.

Người ta quy bài toán so sánh k bộ số liệu này với nhau (mỗi bộ số liệu gồm r phạm trù) bằng cách coi nó như là một bộ số liệu gồm $k \times r$ phạm trù và so sánh nó với một bộ số liệu lý thuyết. Và như vậy, ta chuyển về bài toán đã khảo sát trong phần 5.1.2.1.

Muốn vậy, ta thành lập các tổng hàng và tổng cột như sau

	P. trù 1	P. trù 2	...	P. trù r	Σ
Bộ 1	$N_{1,1}$	$N_{1,2}$	...	$N_{1,r}$	$\mathbf{H}_1$
Bộ 2	$N_{2,1}$	$N_{2,2}$	...	$N_{2,r}$	$\mathbf{H}_2$
...	...	...	...	...	...
Bộ k	$N_{k,1}$	$N_{k,2}$	...	$N_{k,r}$.	$\mathbf{H}_k$
Σ	$\mathbf{C}_1$	$\mathbf{C}_2$	...	$\mathbf{C}_r$	$\mathbf{N}$

trong đó

$\mathbf{H}_i = \sum_{j=1}^{r} \mathbf{N}_{i,j}$ là tổng số số liệu quan sát của bộ thứ i (hàng thứ i), với $i = 1, 2, ..., k$;

$\mathbf{C}_j = \sum_{i=1}^{k} \mathbf{N}_{i,j}$ là tổng số số liệu quan sát của phạm trù thứ j (cột thứ j), với $j = 1, 2, ..., r$,

và

$\mathbf{N} = \sum_{i=1}^{k} \mathbf{H}_i = \sum_{i=1}^{k}\sum_{j=1}^{r} \mathbf{N}_{i,j} = \sum_{j=1}^{r} \mathbf{C}_j = \sum_{j=1}^{r}\sum_{i=1}^{k} \mathbf{N}_{i,j}$ là tổng số toàn bộ số liệu quan sát.

Khi đó, nếu k bộ số liệu này là như nhau, thì

tỷ lệ của các phạm trù 1 trong tổng thể là $p_1 = \frac{\mathbf{C}_1}{\mathbf{N}}$,

tỷ lệ của các phạm trù 2 trong tổng thể là $p_2 = \frac{\mathbf{C}_2}{\mathbf{N}}$,

...

tỷ lệ của các phạm trù r trong tổng thể là $p_r = \frac{\mathbf{C}_r}{\mathbf{N}}$.

Từ đó, ta xây dựng được bộ số liệu lý thuyết gồm $k \times r$ số hạng liệt kê trong bảng như sau

	P. trù 1	P. trù 2	...	P. trù r
Bộ 1	$N'_{1,1}$	$N'_{1,2}$	...	$N'_{1,r}$
Bộ 2	$N'_{2,1}$	$N'_{2,2}$	...	$N'_{2,r}$
...				
Bộ k	$N'_{k,1}$	$N'_{k,2}$	...	$N'_{k,r}$.

trong đó

$$N'_{i.j} = p_j \times \mathbf{H}_i = \frac{\mathbf{C}_j \times \mathbf{H}_i}{\mathbf{N}}.$$

Chú ý rằng khi đó, ta dùng thống kê

$$Q = \sum_{i=1}^{k}\sum_{j=1}^{r} \frac{\left(\mathbf{N}_{i,j} - \mathbf{N}'_{i,j}\right)^2}{\mathbf{N}'_{i,j}}$$

và vì trong quá trình thành lập bộ số liệu lý thuyết, ta đã ước lượng $k + r - 2$ số hạng nên ta có mô hình kiểm định sau

a) $\begin{cases} H: & \text{Các bộ số liệu quan sát là giống nhau} \\ \overline{H}: & \end{cases}$

Các bộ số liệu quan sát là khác nhau

Nếu H đúng thì

$$Q \sim \chi^2\left(k \times r - (k + r - 2) - 1\right) \equiv \chi^2\left((k-1)\times(r-1)\right).$$

b) Với nguy cơ sai lầm α, hay độ tin cậy $\gamma = 1-\alpha$, cho trước, ta tìm được khoảng tin cậy $[0, C]$ cho Q.

c) Tính giá trị cụ thể của Q.

d) So sánh Q với khoảng tin cậy :

$Q > C$: bác bỏ H,

$Q \le C$: chấp nhận H.

Chú ý :

i) Công thức tính Q nêu trên còn có thể viết lại thành (xem chứng minh trong phần phụ lục)

$$Q = \mathbf{N} \times \left(\sum_{i=1}^{k} \sum_{j=1}^{r} \frac{\left(\mathbf{N}_{i,j}\right)^2}{\mathbf{H}_i \times \mathbf{C}_j} - 1 \right).$$

ii) Trường hợp ta so sánh hai bảng số liệu nhưng lại so sánh từng cặp số liệu với nhau. Chẳng hạn, với hai bảng số liệu

X_1	X_2	...	X_n
Y_1	Y_2	...	Y_n

ta không phải ta so sánh hai bộ số liệu với nhau mà là so sánh sự sai khác của từng cặp số liệu xem nó có ý nghĩa không. Do vậy, ta xét hiệu số

$$D_i = X_i - Y_i, \text{ với } i = 1, 2, \ldots, n.$$

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", và ta nhận được bài toán so sánh trung bình (của một bộ số liệu) với một số (số 0) khảo sát trong phần 5.1.1.

Tóm lại, ta có mô hình cho bài toán so sánh hai bộ số liệu từng cặp này như sau

a) Lập hiệu số từng cặp của hai bộ số liệu $D = X - Y$.

b) $\begin{cases} H: & \text{Hai bộ số liệu giống nhau từng cặp} \\ \bar{H}: & \text{Hai bộ số liệu không giống nhau từng cặp} \end{cases}$

và ta nhận được giả thiết tương đương

$$\begin{cases} H: \mu_D = 0 \\ \bar{H}: \mu_D \neq 0 \end{cases}$$

5.2. Bài tập mẫu

Kiểm định tham số

Bài 1. 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 $\bar{X}$ và phương sai S_X^2.

b) Tìm khoảng ước lượng cho trung bình cholesterol trong dân số μ_X, ở độ tin cậy 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 $\alpha = 0,05$).

<u>Giải</u>

a) Để tính trung bình $\bar{X}$ 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

$$S_Y^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} n_i Y_i^2 - n\overline{Y}^2\right) = \frac{164 - 29\left(-0,3448\right)^2}{28} = 5,734,$$

do $S_Y^2 = \left(\frac{1}{5}\right)^2 S_X^2$, ta có

$$S_X^2 = 25 S_Y^2 = 39,7971\text{kg}^2,$$

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

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

Với độ tin cậy $\gamma = 0,95$, từ bảng phân phối Student, ta tìm được $C = t_{0,05}^{28} = 2,048$ sao cho

$$P\left(-2,048 \le T \le 2,048\right) = 0,95,$$

thay $T = \dfrac{\left(\overline{X} - \mu\right)\sqrt{n}}{S_X}$, ta được

$$P\left(-2,048 \le T = \frac{\left(\overline{X} - \mu\right)\sqrt{n}}{S_X} \le 2,048\right) = 0,95.$$

Do đó, ta tìm được khoảng ước lượng cho trung bình tổng thể μ_X là

$$\left[\overline{X} - 2,048\frac{S_X}{\sqrt{n}}; \overline{X} + 2,048\frac{S_X}{\sqrt{n}}\right] = \left[168,73; 177,83\right].$$

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

$$T = \frac{\left(173,28 - 175\right)\sqrt{29}}{11,97} = -0,774.$$

Vì $|T| \le C$, nên ta chấp nhận H, nghĩa là giá trị này phù hợp với mẫu quan sát.

Bài 2. 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 5%, hãy kết luận về nghi ngờ trên.

<u>Giải</u>

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

$$T = \frac{\left(\overline{X} - \mu_0\right)\sqrt{n}}{S_X} \sim St(99) \equiv N(0;1).$$

Với mức ý nghĩa $\alpha = 0,05$, ta có $C = 1,96$. Với số liệu trên, ta được

$$T = \frac{\left(0,9856 - 1\right)\sqrt{100}}{0,021} = -6,86.$$

Vì $|T| > C$, nên ta bác bỏ H, nghĩa là máy hoạt động không bình thường

Bài 3. 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 5%.

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

$$\begin{cases} H : \mu = \mu_0 = 15 \\ \overline{H} : \mu < \mu_0 = 15 \end{cases}$$

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 $\alpha = 0,05 \to 2\alpha = 0,1$ thì $C = t_{0,1}^{24} = 1,711$.

Vì $|T| \le C$, 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

$$T = \frac{\left(f - p\right)\sqrt{n}}{\sqrt{f(1-f)}} \sim St(n-1) = St(24).$$

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 $\left[0,614; 0,906\right]$.

Bài 4. Quan sát sức nặng của bé trai (X) và bé gái (Y) lúc sơ sinh (đơn vị gam), ta có kết quả

Khối 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.

Khối 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

$$F = \frac{S_X^2}{S_Y^2} \sim F(n,m) \equiv F\left(24,27\right).$$

Với mức ý nghĩa $\alpha = 0,05$ thì $C = f_{0,05}(24,27) = 1,89$.

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

$$S^2 = \frac{\left(n-1\right)S_X^2 + \left(m-1\right)S_Y^2}{n+m-2} = \frac{24 \cdot 40266,67 + 27 \cdot 37407,41}{25+28-2} = 38752,94$$

$S = 196,86$;

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

$$T = \frac{\left(\overline{Z} - \mu\right)\sqrt{n}}{S_Z} = \frac{\left(3515,1 - \mu\right)\sqrt{53}}{206,98} \sim St(52) \equiv N(0;1).$$

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 $\left[3459,38; 3570,82\right]$.

Bài 5. 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 5%.

<u>Giải</u>

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

$$T = \frac{\left(\overline{X} - \mu\right)\sqrt{n}}{S_X} \sim St(n-1) = St(35) \equiv N(0;1).$$

Độ tin cậy $\gamma = 0,9$ thì $C = 1,645$. Do đó ta có khoảng ước lượng

$$\left[\overline{X} - C.\frac{S_X}{\sqrt{n}}; \overline{X} + C.\frac{S_X}{\sqrt{n}}\right] = \left[2,136; 3,144\right].$$

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.

Kiểm định phi tham số

Bài 6. 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 $\alpha = 5\%$)

<u>Giải</u>

Ta có bài toán kiểm định

a) $\begin{cases} H: & \text{Người giao hàng nói đúng} \\ \overline{H}: & \text{Người giao hàng nói không đúng} \end{cases}$

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
$0,1\times150=15$	$0,3\times150=45$	$0,4\times150=60$	$0,2\times150=30$

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 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 mức ý nghĩa 5%.

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	$\frac{420\times200}{700}$	$\frac{210\times200}{700}$	$\frac{70\times200}{700}$	***200***
Lô B	$\frac{420\times200}{700}$	$\frac{210\times200}{700}$	$\frac{70\times200}{700}$	***200***
Lô C	$\frac{420\times300}{700}$	$\frac{210\times300}{700}$	$\frac{70\times300}{700}$	***300***
Tổng	***420***	***210***	***70***	**700**

Hơn nữa, ta có thể dùng trực tiếp công thức

$$Q = \mathbf{700}\left(\frac{(125)^2}{\mathbf{200}\times420} + \frac{(52)^2}{200\times210} + \ldots + \frac{(25)^2}{70\times300} - 1\right).$$

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 1%, hãy kiểm định giả thiết cho rằng sự nghỉ việc không phụ thuộc vào giới tính.

<u>Giải</u>

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

Gọi A là nữ

$$P(A) = \frac{300+80+20}{1000} = 0,4\,,$$

$$P(\bar{A}) = \frac{500+70+30}{1000} = 0,6\,.$$

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 5%, có thể kết luận là hoàn cảnh gia đình của trẻ em độc lập với tình trạng phạm tội hay không.

<u>Giải</u>

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 $|Q| \leq C$, 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 $f_1 = \frac{10}{90} = \frac{1}{9}$, và $f_2 = \frac{26}{100} = 0,26$, 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

$$T = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim St(n+m-2) = St(188) \equiv N(0;1).$$

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 cý 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 ý nghĩa 5% 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

$$\begin{cases} H:\ p = \dfrac{1}{2} \\ \overline{H}:\ p \neq \dfrac{1}{2} \end{cases}$$

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

$$Z = \dfrac{(0,391-0,5)\sqrt{46}}{\sqrt{0,5\cdot 0,5}} = -1,48.$$

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 5%, hãy kết luận xem chất lượng công tác bảo vệ an toàn lao động tại 2 xí nghiệp trên có khác nhau không ?

<u>Giải</u>

Ta có bài toán kiểm định

$\begin{cases} H: \\ \overline{H}: \end{cases}$ Chất lượng bảo vệ an toàn của hai xí nghiệp như nhau

Chất lượng bảo vệ an toàn của hai xí nghiệp khác nhau

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 nước ngoài, 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 mức ý nghĩa 5%.

<u>Giải</u>

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

$$T \equiv \frac{\left(\overline{X} - \mu_0\right)\sqrt{n}}{S_X} \sim St\left(n-1\right) = St\left(79\right) \equiv N\left(0;1\right).$$

Từ số liệu của mẫu ta tìm được giá trị của T là

$$T = \frac{\left(\overline{X} - 138,3\right)\sqrt{n}}{S_X} = \frac{\left(120 - 138,3\right)\sqrt{80}}{15} = -10,91.$$

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 ?

<u>Giải</u>

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

$$S^2 = \frac{\left(n_1 - 1\right)S_X^2 + \left(n_2 - 1\right)S_Y^2}{n_1 + n_2 - 2} = \frac{49 \cdot 49 + 39 \cdot 84.64}{50 + 40 - 2} = 64,795,$$

do đó $S = 8,05$.

Nếu H đúng thì

$$T = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim St\left(n_1 + n_2 - 2\right) = St\left(89\right) \equiv N\left(0;1\right).$$

Từ số liệu của hai mẫu, ta tính được giá trị của T là

$$T = \frac{\overline{X} - \overline{Y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{60 - 52}{8.05\sqrt{\frac{1}{50} + \frac{1}{40}}} = 4,68$$

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 ?

<u>Giải</u>

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 $i = 1, 2, ..., 20$ 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

$$T \equiv \frac{\left(\bar{d} - 0\right)\sqrt{n}}{S_d} \sim St\left(n-1\right) = St\left(19\right).$$

Từ đó, ta tìm được giá trị của T là

$$T = \frac{(\bar{d} - 0)\sqrt{n}}{S_d} = \frac{(6,9 - 0)\sqrt{20}}{4,28} = 7,21.$$

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 nghĩa1%, có thể kết luận gì về báo cáo của phân xưởng A ?

<u>Giải</u>

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.

5.3. Bài tập rèn luyện

Kiểm định tham số

Bài 1. Giám đốc một xí nghiệp cho biết lương trung bình của 1 công nhân thuộc xí nghiệp là 760 ngàn đ/tháng. Chọn ngẫu nhiên 36 công nhân thấy lương trung bình là 700 ngàn đ/tháng, với độ lệch chuẩn $\sigma = 80$. Lời báo cáo của giám đốc có tin cậy được không, với mức có ý nghĩa là 5%.

Đáp số: $Z = -4,5$, *bác bỏ.*

Bài 2. Khối lượng các bao gạo là biến ngẫu nhiên có phân phối chuẩn $N(50;\ 0,01)$. Có nhiều ý kiến khách hàng phản ánh là khối lượng bị thiếu. Một nhóm thanh tra đã cân ngẫu nhiên 25 bao gạo trong kho, kết quả như sau :

Khối lượng bao gạo (kg)	48-48,5	48,5-49	49-49,5	49,5-50	50-50,5
Số bao	2	5	10	6	2

Hãy xem ý kiến khách hàng có đúng không bằng cách kiểm tra giả thiết $\mu = 50$ và đối thiết $\mu < 50$, $\alpha = 0,05$.

Đáp số: $Z = -36,5$, *bác bỏ.*

Bài 3. 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 2,5. Với mức ý nghĩa 5%. Hãy kết luận điều nghi ngờ nói trên. Giả thiết lượng sữa bò là 1 biến ngẫu nhiên có phân phối chuẩn.

Đáp số: $T = -3$, *bác bỏ.*

Bài 4. 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; và độ lệch chuẩn 15g/l. Từ kết quả trên có thể kết luận lượng huyết sắc tố trung bình của công nhân nhà máy này thấp hơn mức chung hay không, với mức ý nghĩa 5%.

Đáp số: $T = -10,912$, *bác bỏ.*

Bài 5. Một cửa hàng thực phẩm nhận thấy thời gian vừa qua trung bình một khách hàng mua 25 ngàn đồng thực phẩm trong ngày. Nay cửa hàng chọn ngẫu nhiên 15 khách hàng thấy trung bình một khách hàng mua 24 ngàn đồng trong ngày và phương sai mẫu hiệu 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 hay không.

Đáp số: $T = -1,94$, *chấp nhận.*

Bài 6. Điều tra một mẫu gồm 100 gia đình ở vùng nông thôn người ta thu được kết quả về chi tiêu trung bình hàng tháng của các gia đình đó là 3,455 triệu đồng với độ lệch chuẩn là 0,3 triệu đồng. Với mức ý nghĩa 5% có thể cho rằng chi tiêu trung bình hàng tháng của các gia đình ít hơn 3,5 triệu hay không. Giả thiết mức chi tiêu có phân phối chuẩn.

Đáp số: $T = 1,5$, *chấp nhận giả thuyết.*

Bài 7. Khối 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 khối lượng gà là đại lượng ngẫu nhiên có phân phối theo quy luật chuẩn.

a) Với mức ý nghĩa 5%. Hãy cho kết luận về tác dụng của loại thức ăn này ?

b) Nếu trại chăn nuôi báo cáo khối lượng trung bình khi xuất chuồng là 3,5 kg/con thì có chấp nhận được không, với mức ý nghĩa 5%.

Đáp số: a) $T = 3,06$*; b)* $T = 1,15$.

Bài 8. 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 mức ý nghĩa 5%, hãy kiểm tra xem chất lượng làm việc của máy có còn được như trước hay không?

Đáp số: $Z = -5,75$, *bác bỏ.*

Bài 9. 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: $Z = -1,583$, *chấp nhận.*

Bài 10. 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 mức ý nghĩa 1%. 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, với mức ý nghĩa 5%.

Đáp số: a) $Z = -2,6$, *bác bỏ; b)* $Z = 2,02$, *chấp nhận.*

Bài 11. Nếu máy đóng bao hoạt động bình thường thì khối lượng của một loại sản phẩm sản xuất ra là đại lượng ngẫu nhiên có phân phối theo qui luật chuẩn với độ lệch chuẩn $\sigma = 0,2g$. Kiểm tra khối lượng của 1 số sản phẩm do máy sản xuất, ta được kết quả :

60; 60,2; 70; 60,8; 50,6 ;50,8; 50,9; 60,1; 50,3; 60,5; 60,1; 60,2; 60,3; 50,8; 60; 70

a) Với mức ý nghĩa 5%, hãy cho biết máy đóng bao hoạt động có bình thường hay không ?

b) Ước lượng khối lượng trung bình của loại sản phẩm này hiện nay với độ tin cậy 95%.

Đáp số: a) ; b) .

Bài 12. 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 mức ý nghĩa 5%?

Đáp số: $T = -2,58$, *bác bỏ.*

Bài 13. Để so sánh khối lượng trung bình của trẻ sơ sinh ở thành thị và nông thôn, người ta thử cân khối lượng của 10000 cháu và thu được kết quả sau đây :

Vùng	Số cháu được cân	Khối 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 5%, có thể coi khối lượng trung bình của trẻ sơ sinh ở thành thị cao hơn ở nông thôn hay không? (Giả thiết khối lượng trẻ sơ sinh là biến ngẫu nhiên có phân phối chuẩn).

Đáp số: $T = -28,28$, *bác bỏ giả thuyết.*

Bài 14. 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 :	$n_1 = 50$:	$\overline{X} = 60mg\%$;	$S_1 = 7mg\%$
Sau :	$n_2 = 40$:	$\overline{Y} = 52mg\%$;	$S_2 = 9,2mg\%$

Với mức ý nghĩa 5%, 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.*

Bài 15. Trong thập niên 80, khối lượng trung bình của thanh niên là 48kg. Nay để xác định lại khối lượng ấy, người ta chọn ngẫu nhiên 100 thanh niên đo khối lượng trung bình là 50kg và phương sai mẫu hiệu chỉnh $S^2 = (10\text{kg})^2$.

a) Thử xem khối lượng thanh niên hiện nay phải chăng có thay đổi, với mức có ý nghĩa là 1%.

b) Nếu khối lượng thực tế của người thanh niên là $a_1 = 51\text{kg}$ thì xác suất mắc sai lầm loại 2 là bao nhiêu.

c) Nếu muốn xác suất mắc sai lầm loại 1 là 1% và xác suất mắc sai lầm loại 2 không vượt quá 5% thì phải đo khối lượng của bao nhiêu thanh niên nếu khối lượng trung bình thực tế của thanh niên hiện nay không vượt quá 52kg.

d) Nếu muốn xác suất mắc sai lầm loại 1 là 1% và xác suất mắc sai lầm loại 2 không vượt quá 5% thì phải đo khối lượng của bao nhiêu thanh niên nếu khối lượng trung bình thực tế của thanh niên hiện nay trong khoảng (44 ; 52)kg.

Đáp số:

Kiểm định phi tham số

Bài 16. Cùng một loại hạt giống đem xử lý theo 2 phương án khác nhau. Kết quả quan sát chiều cao cây con của mỗi phương án được cho dưới đây

Phương án I	39,2	29	28,5	33,5	41,7	37,2
	37,3	27,7	23,4	33,4	29,2	35,6
Phương án II	20,8	33,8	28,6	23,4	22,7	30,9
	31,0	27,4	19,5	29,6	23,2	18,7
	20,7	17,6	29,4	27,7	25,5	14,5

Hãy dùng tiêu chuẩn phi tham số để kiểm tra xem 2 phương án xử lý có ảnh hưởng đến sinh trưởng chiều cao cây con hay không, với mức ý nghĩa 5%.

Đáp số:

Bài 17. Giả sử ta muốn xác định xem hiệu quả của chế độ ăn kiêng đối với việc giảm khối lượng như thế nào. 20 người quá béo đã thực hiện chế độ ăn kiêng. Khối lượng của từng người trước khi ăn kiêng (X kg) và sau khi ăn kiêng (Y kg) được cho như sau:

X	80	78	85	70	90	78	92	88	75		
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

Dùng tiêu chuẩn phi tham số kiểm tra xem chế độ ăn kiêng có tác dụng làm giảm khối lượng hay không, với mức ý nghĩa 5%.

Đáp số: $T = -3,39$, *bác bỏ giả thuyết.*

Bài 18. 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

Các phương án xử lý có tác dụng như nhau đối với tỷ lệ nảy mầm hay không, với mức ý nghĩa 5%.

Đáp số: $Q = 61,52$, *bácbỏ giả thuyết.*

Bài 19. 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 mức ý nghĩa 5%, hãy kiểm tra giả thiết cho rằng màu của tóc và màu của mắt độc lập với nhau.

Đáp số: $Q = 15,07$, *bácbỏ giả thuyết.*

Bài 20. 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, với mức ý nghĩa 5%.

Đáp số: $Q = 50,83$, *bácbỏ giả thuyết.*

Bài 21. 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 mức ý nghĩa 5%, 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 thuyết.*

Bài 22. Một máy sản suất tự động, lúc đầu tỷ lệ sản phẩm loại A là 20%. Sau khi áp dụng một phương pháp sản suất mới, người ta lấy 40 mẫu, mỗi mẫu gồm 10 sản phẩm đề kiểm tra. Kết quả kiểm tra cho ở bảng sau :

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

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

Đáp số: $Z = 16,875$, *bácbỏ giả thuyết.*

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 mức ý nghĩa 5%, 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: $Q = 18,88$, *bácbỏ giả thuyết.*

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 $\mu = 4,7$. 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 thuyết.*

Bài 25. 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 5% có thể xem X tuân theo luật Poisson hay không?

Đáp số: $Q = 8,01$, *bácbỏ giả thuyết.*

Bài 26. 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 mức ý nghĩa 5%, chứng tỏ rằng số hộp bị hỏng của một thùng là biến ngẫu nhiên tuân theo qui luật Poisson?

Đáp số: $Q = 2,393$, *chấp nhận giả thuyết.*

Bài 27. 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 thuyết.*

Bài 28. Năng suất lúa (X) 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 10%, xét xem X có phân phối chuẩn không?

Đáp số: $Q = 4,4$, *chấp nhận giả thuyết.*

MỘT SỐ ĐỀ THAM KHẢO

Đề 1

Câu 1. Một nhà máy có 3 phân xưởng A, B, C tương ứng làm ra 25%, 35% và 40% tổng sản phẩm của nhà máy. Giả sử xác suất làm ra một sản phẩm hỏng của các phân xưởng A, B và C lần lượt là 0,015; 0,025 và 0,035.

a) Tính xác suất để chọn ngẫu nhiên một sản phẩm của nhà máy thì nhận phải sản phẩm hỏng.

b) Biết rằng nhận phải sản phẩm hỏng của nhà máy. Theo bạn sản phẩm hỏng đó do phân xưởng nào sản xuất.

Câu 2. Một nhà máy sản xuất 100000 sản phẩm trong đó có 30000 sản phẩm loại A. KCS đến kiểm tra và lấy ngẫu nhiên không hoàn lại 500 sản phẩm ra thử. Hãy tính xác suất để số sản phẩm loại A mà KCS phát hiện ra có

a) Đúng 150 sản phẩm,

b) Từ 145 đến 155,

c) Ít hơn 151.

Câu 3. Đem cân một số trái cây vừa thu hoạch, ta được kết quả sau :

Khối lượng (g)	205	215	225	235	245	255	265	275
Số trái	8	14	16	23	16	10	8	5

Giả sử khối lượng của trái cây có phân phối chuẩn.

a) Tìm khoảng ước lượng của khối lượng trung bình của trái cây với độ tin cậy là 90%.

b) Trái cây có khối lượng lớn hơn 250 gam được gọi là trái cây loại một. Tìm khoảng ước lượng cho tỷ lệ này với độ tin cậy 99%.

c) Nếu muốn sai số ước lượng của khối lượng trung bình không vượt quá 2 gam ở độ tin cậy là 95% thì phải cân thêm ít nhất bao nhiêu trái nữa?

d) Theo tài liệu cho biết khối lượng trung bình của trái cây là 240 gam. Hãy cho biết bảng số liệu trên có phù hợp với tài liệu này không? Kết luận với mức ý nghĩa là 5%.

Đề 2

Câu 1. Giả sử thị trường xe ở Việt Nam do ba nước Nhật, Trung Quốc và Việt Nam cung cấp. Tỷ lệ xe hỏng của ba nước lần lượt là 1%, 20%, 10%. Biết rằng xe của Trung Quốc chiếm $\frac{2}{3}$ thị trường, xe của Nhật chiếm $\frac{1}{4}$ thị trường còn lại là xe của Việt Nam. Chọn ngẫu nhiên một xe trên thị trường.

a) Tìm xác suất để xe được chọn là xe hỏng.

b) Nếu chọn được xe hỏng. Theo bạn xe đó là do nước nào cung cấp.

Câu 2. Sản phẩm sau khi hoàn tất được đóng thành kiện, mỗi kiện gồm 10 sản phẩm với tỷ lệ chính phẩm là 80%. Trước khi mua hàng, khách hàng muốn kiểm tra bằng cách từ mỗi kiện chọn ngẫu nhiên 3 sản phẩm.

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

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

Câu 3. Điều tra năng suất lúa trên diện tích 100 hecta trồng lúa của một vùng, ta thu được bảng số liệu sau :

Năng suất (tạ/ha)	41	44	45	46	48	52	54
Diện tích (ha)	10	20	30	15	10	10	5

Giả sử năng suất lúa có phân phối chuẩn

a) Hãy ước lượng năng suất lúa trung bình của vùng đó với độ tin cậy 95%.

b) Nếu muốn sai số ước lượng của năng suất lúa trung bình không vượt quá 0,5 tạ/ha, với độ tin cậy 99% thì cần điều tra thêm ít nhất bao nhiêu hecta lúa nữa.

c) Những thửa ruộng có năng suất từ 48 tạ/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 90%.

d) Có tài liệu cho biết năng suất lúa trung bình là 47 tạ/ha. Giá trị này có phù hợp với mẫu quan sát không? (kết luận với mức ý nghĩa 5%).

Đề 3

Câu 1. Số liệu thống kê về doanh số bán (triệu đồng/ngày) của một siêu thị như sau :

Doanh số	Số ngày	Doanh số	Số ngày
20 – 40	5	80 - 90	15
40 – 50	10	90 - 100	10
50 – 60	20	100 - 110	8
60 – 70	25	110 - 130	3
70 – 80	25		

a) Những ngày có doanh số bán hàng trên 90 triệu đồng là những ngày bán đắt hàng. Hãy ước lượng tỷ lệ những ngày bán đắt hàng ở siêu thị này với độ tin cậy 95%.

b) Ước lượng doanh số bán hàng trung bình của một ngày ở siêu thị với độ tin cậy 90%, giả sử doanh số bán hàng của những ngày bán là đại lượng ngẫu nhiên có phân phối chuẩn.

c) Nếu muốn sai số ước lượng trung bình của một ngày bán hàng ở siêu thị không vượt quá 3 triệu đồng/ngày, ở độ tin cậy 99% thì cần quan sát thêm ít nhất bao nhiêu ngày nữa.

d) Trước đây doanh số bán hàng trung bình là 65 triệu đồng/ngày. Số liệu ở trên được thu thập sau khi siêu thị áp dụng phương pháp bán hàng mới. Hãy cho nhận xét về phương pháp bán hàng mới này với mức ý nghĩa 5%.

Câu 2. Có 20 kiện hàng, mỗi kiện có 10 sản phẩm. Trong số đó có 8 kiện loại 1, mỗi kiện có 2 phế phẩm; 7 kiện loại 2, mỗi kiện có 3 phế phẩm; 5 kiện loại 3, mỗi kiện có 5 phế phẩm. Lấy ngẫu nhiên một kiện, rồi từ kiện đó lấy ngẫu nhiên một sản phẩm.

a) Tính xác suất sản phẩm lấy ra là phế phẩm.

b) Biết sản phẩm lấy ra là phế phẩm. Theo bạn sản phẩm đó thuộc kiện loại nào".

Câu 3. Độ dài của một chi tiết máy được tiện ra được biết là có phân phối chuẩn, với độ dài trung bình là 1,2 cm và độ lệch chuẩn về độ dài là 0,001 cm.

a) Sản phẩm tiện ra được xem là sản phẩm loại một nếu độ dài lớn hơn 1,202 cm. Tính xác suất để chọn ngẫu nhiên một sản phẩm thì được sản phẩm loại một.

b) Tính xác suất sản phẩm được chọn ra có độ dài từ 1,198 cm đến 1,202 cm.

c) Nếu chọn được sản phẩm loại một thì sẽ mua sản phẩm đó. Chọn ngẫu nhiên 10 sản phẩm, tính xác suất để mua được 3 sản phẩm loại một.

Đề 4

Câu 1. Trước khi đưa sản phẩm ra thị trường người ta đã phỏng vấn ngẫu nhiên 200 khách hàng về sản phẩm đó và thấy có 34 người trả lời "sẽ mua", 96 người trả lời "có thể mua" và 70 người trả lời "không mua". Kinh nghiệm cho thấy tỉ lệ khách hàng thực sự sẽ mua sản phẩm tương ứng với cách trả lời trên là 40%, 20% và 1%.

a) Tính tỉ lệ khách hàng thực sự mua sản phẩm đó.

b) Trong số khách hàng thực sự mua sản phẩm thì có bao nhiêu phần trăm trả lời "sẽ mua".

Câu 2. Khối 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 khối lượng như sau :

i) loại 1 : trên 505 gam,

ii) loại 2 : từ 495 đến 505 gam,

iii) loại 3 : dưới 495 gam.

Tính tỷ lệ mỗi loại.

Câu 3. Tiến hành điều tra số gạo bán ra hằng ngày ở một cửa hàng có kết quả sau :

Số gạo bán ra (kg)	120	130	150	160	180	190	210	220
Số ngày	2	9	12	25	30	20	13	4

Giả sử số gạo bán được trong ngày theo luật phân phối chuẩn

a) Những ngày bán được trên 200 kg là những ngày "cao điểm". Hãy ước lượng tiền bán được trong những ngày "cao điểm", biết rằng giá gạo trung bình là 10000 đồng/kg, với độ tin cậy 99%.

b) Hãy ước lượng tỉ lệ ngày "cao điểm" với độ tin cậy 90%.

c) Để ước lượng tỉ lệ ngày "cao điểm" với độ chính xác 5% thì độ tin cậy là bao nhiêu?

d) Chủ cửa hàng cho rằng nếu trung bình mỗi ngày bán ra không quá 150 kg thì tốt hơn là nghỉ bán. Từ số liệu trên, với mức ý nghĩa 5% cửa hàng nên quyết định thế nào?

Đề 5

Câu 1. Một lô hạt giống được phân thành ba loại. Loại 1 chiếm 2/3 số hạt cả lô, loại 2 chiếm 1/4, còn lại là loại 3. Loại 1 có tỉ lệ nảy mầm 80%, loại 2 có tỉ lệ nảy mầm 60% và loại 3 có tỉ lệ nảy mầm 40%. Hỏi tỉ lệ nảy mầm chung của lô hạt giống là bao nhiêu ?

Câu 2. 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.

Câu 3. Khảo sát về thu nhập của một số người của công ty, người ta thu được bảng số liệu sau (thu nhập triệu đồng/năm) :

Thu nhập	8-12	12-14	14-16	16-20	20-30

Số người	8	12	20	45	15

Giả sử thu nhập có phân phối chuẩn

a) Hãy ước lượng thu nhập trung bình của một người trong công ty với độ tin cậy 90%.

b) Những người có thu nhập trên 20 triệu đồng/năm là những người có thu nhập cao. Hăy ước lượng số người có thu nhập cao ở công ty với độ tin cậy 99%, biết rằng tổng số người làm việc trong công ty là 2000 người

c) Nếu muốn dùng mẫu trên để ước lượng thu nhập trung bình của một người ở công ty với độ chính xác 100 ngàn đồng/năm thì độ tin cậy là bao nhiêu?

d) Nếu công ty báo cáo mức thu nhập bình quân của một người trong công ty là 1,5 triệu đồng/tháng thì có chấp nhận được không? Kết luận với mức ư nghĩa 5% .

Đề 6

Câu 1. Một nhà máy X có ba phân xưởng khác nhau. Tỷ lệ phế phẩm của ba phân xưởng lần lượt là 1%, 5%, 10%. Biết rằng tỷ lệ sản phẩm của phân xưởng một và hai là như nhau và bằng một nửa của phân xưởng ba. Lấy ngẫu nhiên một sản phẩm của nhà máy.

a) Tìm xác suất để sản phẩm lấy ra là phế phẩm.

b) Nếu lấy ra phải phế phẩm. Tìm xác suất để phế phẩm đó là của phân xưởng một, của phân xưởng hai, của phân xưởng ba.

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

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

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

Câu 3. Để tìm hiểu lượng mủ X (g) mỗi cây cao su cho ta trong một ngày, ghi nhận 100 cây ta có kết quả sau :

X (g)	200-210	210-220	220-230	230-240	240-250	250-260	260-270

Số cây	2	8	14	30	25	12	9

Giả sử lượng mủ X có phân phối chuẩn

a) Hãy ước lượng lượng mủ trung bình của mỗi cây cao su với độ tin cậy 90%.

b) Nếu muốn ước lượng lượng mủ trung bình của mỗi cây cao su với độ tin cậy 99% và độ chính xác 3g thì cần phải điều tra thêm bao nhiêu cây nữa?

c) Nếu ước lượng lượng mủ trung bình của mỗi cây cao su với độ chính xác 2,5g thì sẽ đạt được độ tin cậy là bao nhiêu?

d) Một tài liệu thống kê cũ cho rằng lượng mủ trung bình của mỗi cây cao su là 235 gam. Hãy cho kết luận về tài liệu đó với mức ý nghĩa 5%.

Đề 7

Câu 1. Hai cửa hàng X và Y cung cấp các đĩa mềm máy tính cho một trung tâm tin học với tỷ lệ 3/2. Tỷ lệ đĩa bị lỗi của các cửa hàng tương ứng là 1% và 2%. Một sinh viên đến thực tập tại trung tâm chọn ngẫu nhiên một hộp đĩa gồm 20 chiếc và từ đó rút ngẫu nhiên một đĩa.

1) Tính xác suất để sinh viên đó rút phải đĩa bị lỗi.

2) Sau khi khởi động máy, sinh viên đó nhận thấy quả thật đĩa bị lỗi. Tính xác suất để đĩa này thuộc cửa hàng X.

Câu 2. Sản phẩm nhà máy được đóng thành từng hộp, mỗi hộp có 10 sản phẩm. Gọi X là số sản phẩm loại một có trong hộp. Cho biết X có bảng phân phối xác suất như sau:

X	6	7
P	0,7	0,3

Khách hàng chọn cách kiểm tra để mua hàng như sau : Từ mỗi hộp lấy ngẫu nhiên ra 3 sản phẩm để kiểm tra, nếu thấy có 2 sản phẩm loại một thì mua hộp đó. Lấy ngẫu nhiên 3 hộp để kiểm tra. Tính xác suất để có 2 hộp được mua.

Câu 3. Kết quả quan sát về hàm lượng vitamin của một loại trái cây, ta có số liệu sau :

Hàm lượng (%)	6-7	7-8	8-9	9-10	10-11	11-12
Số trái	5	10	20	35	25	5

Giả sử hàm lượng vitamin có phân phối chuẩn.

a) Hãy ước lượng hàm lượng vitamin trung bình trong một trái, với độ tin cậy 95%.

b) Những trái có hàm lượng vitamin trên 10% là trái loại I. Hãy ước lượng tỷ lệ trái loại I, với độ tin cậy 99%.

c) Nếu muốn sai số ước lượng của tỷ lệ trái loại I không quá 0.05, với độ tin cậy 95% thì cần quan sát thêm ít nhất bao nhiêu trái nữa.

Câu 4. Một công ty điện thoại nói rằng sẽ lắp đặt điện thoại cho khách hàng trong thành phố chậm nhất là 30 ngày tính từ lúc yêu cầu. Kiểm tra ngẫu nhiên 30 khách hàng thấy thời gian trung bình lắp đặt điện thoại là 34,5 ngày với độ lệch chuẩn mẫu là 3,3 ngày. Với mức ý nghĩa 1%, hãy cho biết có chấp nhận lời tuyên bố trên hay không?

Đề 8

Câu 1. Kiện hàng 1 có 5 sản phẩm loại A, 1 sản phẩm loại B. Kiện hàng 2 có 2 sản phẩm loại A, 4 sản phẩm loại B. Từ mỗi kiện ta chọn ngẫu nhiên ra 1 sản phẩm đem giao cho khách hàng. Sau đó các sản phẩm còn lại được dồn chung vào kiện hàng 3 (đang trống).

a) Nếu ta lấy ngẫu nhiên 1 sản phẩm từ kiện hàng 3 thì xác suất để chọn sản phẩm loại B là bao nhiêu?

b) Nếu ta lấy ngẫu nhiên 2 sản phẩm từ kiện 3, hãy tính xác suất để có ít nhất 1 sản phẩm loại B trong 2 sản phẩm được chọn.

Câu 2. Cho đại lượng ngẫu nhiên X có bảng phân phối xác suất như sau:

X	$0,6$	4	x_3
P	$0,3$	$0,5$	p_3

Cho biết $EX = 8$. Tính phương sai của X.

Câu 3. Đường kính của 100 chi tiết do một máy sản xuất kết quả cho trong 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

Giả sử đường kính của chi tiết máy có phân phối chuẩn. Quy định những chi tiết có đường kính từ 19,9 mm đến 20,1 mm là những chi tiết đạt tiêu chuẩn.

a) Hãy ước lượng đường kính trung bình của những chi tiết đạt tiêu chuẩn với độ tin cậy 95%.

b) Hãy ước lượng tỷ lệ chi tiết đạt tiêu chuẩn với độ tin cậy 99%.

Câu 4. Tỉ lệ phế phẩm của một xí nghiệp sản xuất là 5%. Nhằm giảm bớt tỉ lệ phế phẩm người ta đã cải tiến kỹ thuật. Sau khi cải tiến kỹ thuật, kiểm tra 400 sản phẩm thấy có 18 phế phẩm. Với mức ý nghĩa 5%, hãy cho kết luận về hiệu quả của việc cải tiến kỹ thuật.

Đề 9

Câu 1. Có 3 hộp bề ngoài giống hệt nhau. Các hộp chứa lần lượt 10, 15, 20 chính phẩm mỗi hộp đều chứa 5 phế phẩm. Lấy ngẫu nhiên một hộp rồi từ hộp đó lấy ngẫu nhiên 2 sản phẩm.

1) Tính xác suất lấy được cả hai phế phẩm.

2) Kiểm tra thì thấy cả hai sản phẩm lấy ra đúng là phế phẩm. Tính xác suất để hai sản phẩm đó thuộc hộp một.

Câu 2. Có 3 hộp, mỗi hộp có 35 sản phẩm.

Hộp 1 có 3 sản phẩm không đạt chất lượng.

Hộp 2 có 6 sản phẩm không đạt chất lượng.

Hộp 3 có 1 sản phẩm không đạt chất lượng.

Lấy ngẫu nhiên một hộp, rồi từ hộp đó lấy ngẫu nhiên ra 2 sản phẩm. Gọi X là số sản phẩm không đạt chất lượng trong 2 sản phẩm lấy ra.

1. Lập bảng phân phối xác suất cho X.
2. Tính trung bình và phương sai của X.

Câu 3. Khảo sát về thu nhập X (triệu đồng / tháng) của một số người, ta có bảng số liệu như sau:

Thu nhập	0 – 4	4 – 8	8 – 12	12 – 16	16 – 20	20 – 24
Số người	8	12	20	30	16	10

Giả sử thu nhập có phân phối chuẩn.

1. Hãy ước lượng thu nhập trung bình của một người trong một tháng với độ tin cậy 95%.
2. Những người có mức thu nhập từ 16 triệu đồng / tháng trở lên được gọi là những người có mức thu nhập "cao". Hãy ước lượng tỉ lệ những người có mức thu nhập "cao" với độ tin cậy 99%.
3. Nếu dùng mẫu này để ước lượng tỉ lệ những người có mức thu nhập "cao" với sai số không quá 5% thì độ tin cậy là bao nhiêu ?

Câu 4. Một công ty sản xuất bóng đèn đã quảng cáo rằng bóng đèn loại 75W của họ đốt sáng trung bình 800 giờ trước khi hỏng. Tổ chức những người tiêu dùng cần phải quyết định xem có phạt tiền liên quan đến chiến dịch quảng cáo của công ty hay không. Vì thế họ quyết định lấy ngẫu nhiên và kiểm tra 100 bóng đèn khiếu kiện. Với thí nghiệm này, 100 bóng đèn đốt sáng trung bình là 745,5 giờ trước khi cháy với độ lệch chuẩn là 238 giờ. Với mức ý nghĩa 5%, hãy cho biết công ty quảng cáo như vậy có bị phạt tiền hay không?

Đề 10

Câu 1. Cho hai bình. Bình thứ nhất chứa 3 bi trắng, 2 bi đỏ. Bình thứ hai chứa 2 bi trắng, 3 bi đỏ. Lấy ngẫu nhiên 2 bi từ bình thứ nhất bỏ sang bình thứ hai, rồi sau đó lấy ngẫu nhiên 2 bi từ bình thứ hai ra ngoài. Tính xác suất để hai bi này là 1 trắng và 1 đỏ.

Câu 2. Một nhà máy sản xuất 100000 sản phẩm trong đó có 30000 sản phẩm loại A. KCS đến kiểm tra và lấy ngẫu nhiên không hoàn lại 500 sản phẩm ra thử. Hãy tính xác suất để số sản phẩm loại A mà KCS phát hiện ra có

1. Đúng 150 sản phẩm,

2. Từ 145 đến 155.

Câu 3. Một đợt xổ số có 10% vé trúng thưởng. Hỏi phải mua ít nhất bao nhiêu vé để xác suất có ít nhất 1 vé trúng thưởng không nhỏ hơn 90%.

Câu 4. Để tìm hiểu lượng mủ X (g) mỗi cây cao su cho ta trong một ngày, ghi nhận được kết quả sau :

X (g)	200-210	210-220	220-230	230-240	240-250	250-260	260-270
Số cây	2	8	14	30	25	12	9

Giả sử lượng mủ X có phân phối chuẩn

1. Hãy ước lượng lượng mủ trung bình của mỗi cây cao su với độ tin cậy 95%.

2. Nếu muốn ước lượng lượng mủ trung bình của mỗi cây cao su với độ tin cậy 99% và độ chính xác 3g thì cần phải điều tra thêm bao nhiêu cây nữa?

3. Nếu ước lượng lượng mủ trung bình của mỗi cây cao su với độ chính xác 2,5g thì sẽ đạt được độ tin cậy là bao nhiêu?

Câu 5. Một lô hàng có 6000 sản phẩm. Chọn ngẫu nhiên 400 sản phẩm từ lô hàng để kiểm tra thì thấy có 360 sản phẩm loại A.

1. Hãy ước lượng số sản phẩm loại A có trong lô hàng với độ tin cậy 95%.

2. Nếu cho rằng số sản phẩm loại A có trong lô hàng là 5500 thì có chấp nhận được không? (với mức ý nghĩa 5%)

Đề 11

Câu 1. 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.

Câu 2. Để thanh toán 1 triệu đồng tiền hàng, một khách hàng gian lận đã xếp lẫn 5 tờ 50 ngàn đồng tiền giả với 15 tờ tiền thật. Chủ cửa hàng rút ngẫu nhiên 3 tờ giấy bạc đem đi kiểm tra và giao hẹn nếu phát hiện có bạc giả thì cứ mỗi tờ giả khách hàng phải đền hai tờ thật. Tìm số tiền phạt mà khách có thể phải trả.

Câu 3. Cân ngẫu nhiên 45 con heo 3 tháng tuổi trong một trại chăn nuôi, ta được kết quả sau

X_i	35	37	39	41	43	45	47
n_i	2	6	10	11	8	5	3

Giả sử khối lượng X (kg) tuân theo quy luật phân phối chuẩn.

a) Hãy ước lượng khoảng cho khối lượng trung bình các con heo 3 tháng tuổi trong trại trên với độ tin cậy 95%.

b) Heo có khối lượng $\geq 38\text{kg}$ là heo đạt tiêu chuẩn. Hãy tìm ước lượng tỷ lệ heo đạt chuẩn với độ tin cậy 90%.

Câu 4. 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.

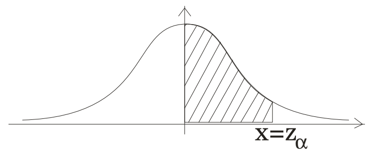
a) Với mức ý nghĩa 1%. Hãy cho kết luận về biện pháp kỹ thuật mới này ?

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, với mức ý nghĩa 5%.

PHÂN PHỐI GAUSS

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt = P\left(0 \le X \le x\right) \equiv \alpha,$$

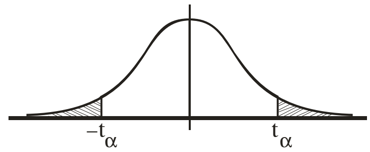
với $X \sim N\left(0;1\right)$, $x \equiv z_\alpha$.



	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

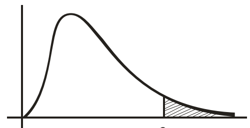
PHÂN PHỐI CHI – BÌNH PHƯƠNG

$P\left(X \geq \chi^2_{n,\alpha}\right) = \alpha$ khi $X \sim \chi^2(n)$

Hàng 1 : Giá trị của α

Cột 1 : Giá trị độ tự do n.

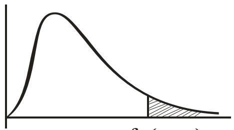
Nội dung bảng : Giá trị $\chi^2_{n,\alpha}$.


$\chi^2_{n,\alpha}$

	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.

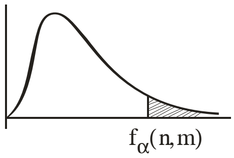
Nội dung bảng : Giá trị $f_{\alpha}(n,m)$.

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