BÀI 1 – XÁC SUẤT CỦA BIẾN CỐ

+ Định nghĩa xác suất cổ điển.	**+ Các bài toán tìm xác suất bằng định nghĩa.**

1.1 (4.t28) Một phép thử bao gồm tung 2 con xúc xắc, một con màu đỏ và một con màu xanh, rồi ghi lại số chấm xuất hiện trên mỗi con. Nếu x là số chấm xuất hiện trên con xúc xắc màu xanh và y là số chấm xuất hiện trên con xúc xắc màu đỏ, hãy mô tả không gian mẫu S bằng 2 cách:

a) Liệt kê tất cả các phần tử (x, y); b) Chỉ ra quy luật của các phần tử trong S.

1.2 (17.t28) Cho A, B, C là các biến cố liên quan đến không gian mẫu S. Sử dụng sơ đồ Venn, hãy bôi đen vùng tương ứng với các biến cố:

a) $\overline{(A\cap B)}$; b) $\overline{(A\cup B)}$; c) $(A\cap C)\cup B$.

1.3 (10.t46) Gieo đồng thời hai con xúc xắc. Tìm xác suất để nhận được:

a) Tổng số chấm là 8. b) Tổng số chấm lớn nhất là 5.

Giải:

a) + Gọi A là biến cố " Tổng số chấm hai con xúc xắc là 8".
+ Không gian mẫu $N = 6\times 6 = 36$.
+ Số khả năng thuận lợi cho biến cố A là $(2,6);(3,5);(4,4);(5,3);(6,2)$ nên $n = 5$.
+ Áp dụng công thức $P(A) = \dfrac{n}{N} = \dfrac{5}{36}$

b) + Gọi B là biến cố " Tổng số chấm hai con xúc xắc lớn nhất là 5".
+ Không gian mẫu $N = 6\times 6 = 36$.
+ Số khả năng thuận lợi cho biến cố B là
$(1,1);(1,2);(1,3);(1,4);(2,1);(2,2);(2,3);(3,2);(3,1);(4,1)$ nên $n = 10$.
+ Áp dụng công thức $P(A) = \dfrac{n}{N} = \dfrac{5}{18}$

1.4 (12.t46) Chọn ngẫu nhiên 3 quyển sách từ một giá sách gồm 5 quyển tiểu thuyết, 3 quyển thơ và một quyển từ điển. Tìm xác suất để:

a) Quyển từ điển được chọn.
b) Hai quyển tiểu thuyết và một quyển thơ được chọn.

Đáp án

a) $P(A) = \dfrac{n}{N} = \dfrac{C_8^2}{C_9^3} = \dfrac{1}{3}$.

b) $P(B) = \dfrac{n}{N} = \dfrac{C_5^2 C_3^1}{C_9^3} = \dfrac{5}{14}$.

1.5 (9.t46) Mỗi cuốn sách trong thư viện được mã hóa với 3 chữ cái đứng trước và 4 chữ số khác không đứng sau. Tìm xác suất để khi chọn ngẫu nhiên một cuốn sách ta được cuốn sách có mã hóa là chữ cái đầu tiên là một nguyên âm và chữ số cuối cùng là số chẵn. (Biết rằng tiếng Anh có 26 chữ cái với 5 nguyên âm).

Đáp án: $P(A) = \dfrac{n}{N} = \dfrac{5.26^2.9^3.4}{26^3.9^4} = \dfrac{10}{117}$.

1.6 (11.t46) Lấy lần lượt hai quân bài từ một cỗ bài theo phương thức không hoàn lại. Tính xác suất để cả hai quân bài đều lớn hơn 2 và nhỏ hơn 8.

Đáp án: $P(A) = \dfrac{n}{N} = \dfrac{20.19}{52.51}$.

1.7 (12.t63) Từ một hộp đựng 6 quả bóng đen và 4 quả bóng xanh, lần lượt lấy ra 3 quả bóng theo phương thức có hoàn lại. Tìm xác suất để:

a) Cả 3 quả bóng được lấy ra cùng màu. b) 3 quả bóng lấy ra có đủ cả 2 màu.

Giải:

a) $P(A) = \dfrac{n}{N} = \dfrac{6^3 + 4^3}{10^3}$.

b) $P(B) = 1 - P(A)$

BÀI 2 – CÁC CÔNG THỨC TÍNH XÁC SUẤT

+ Các định lý về phép toán xác suất.	**+ Công thức Xác suất đầy đủ, công thức Bayes.**

2.1.(5.t45): Xác suất để 1 ngành kinh doanh của Mỹ có trụ sở tại Munich là 0.7; xác suất để nó có trụ sở tại Brussels là 0.4 và xác suất để nó có trụ sở ở Munich hoặc Brussels hoặc cả hai là 0.8. Tính xác suất để ngành kinh doanh đó có trụ sở:

a) ở cả hai thành phố trên
b) Không ở thành phố nào trong hai thành phố trên.

Giải:

a) + Gọi A là biến cố " có trụ sở tại Munich", B là biến cố " có trụ sở tại Brussels", C là biến cố "ở cả hai thành phố trên"

+ Ta có $P(A)=0.7; P(B)=0.4; P(A\cup B)=0.8$ và $C=AB$

+ Áp dụng công thức $P(A\cup B)=P(A)+P(B)-P(AB)$

+ Thay số vào ta có $P(AB)=P(A)+P(B)-P(A\cup B)=0.3$

b) + Gọi D là biến cố " không ở thành phố nào trong 2 thành phố trên"

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

+ Áp dụng công thức De – morgan ta có $D=\overline{A}\overline{B}=\overline{A\cup B}$

+ Nên $P(D)=P(\overline{A\cup B})=1-P(A\cup B)=0.2$

2.2 (6.t45) Từ kinh nghiệm của mình, một người mua bán cổ phiếu tin rằng, với điều kiện kinh tế hiện nay, một khách hàng sẽ đầu tư vào trái phiếu miễn thuế với xác suất là 0,6, đầu tư vào chứng chỉ quỹ với xác suất là 0,3 và đầu tư vào cả hai loại trên với xác suất là 0,15. Tìm xác suất để tại thời điểm này một khách hàng sẽ:

a) Đầu tư vào trái phiếu miễn thuế hoặc chứng chỉ quỹ.
b) Không đầu tư vào trái phiếu miễn thuế cũng không đầu tư vào chứng chỉ quỹ.

Giải:

a) + Gọi A là biến cố " đầu tư vào trái phiếu"
B là biến cố " đầu tư vào chứng chỉ quỹ"
C là biến cố "đầu tư vào miễn thuế hoặc chứng chỉ quỹ"

+ Ta có $P(A)=0.6; P(B)=0.3; P(AB)=0.15$ và $C=A\cup B$

+ Áp dụng công thức $P(A\cup B)=P(A)+P(B)-P(AB)$

+ Thay số vào ta có $P(A\cup B)=0.75$

b) + Gọi D là biến cố " không đầu tư vào cả hai loại trên"

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

+ Áp dụng công thức De – morgan ta có $D=\overline{A}\overline{B}=\overline{A\cup B}$

+ Nên $P(D)=P(\overline{A\cup B})=1-P(A\cup B)=0.25$

2.3 (15.t46) Một lớp học có 100 sinh viên, trong đó có 54 sinh viên học toán, 69 sinh viên học lịch sử và 35 sinh viên học cả toán và lịch sử. Chọn ngẫu nhiên một sinh viên, tìm xác suất để:

a) Sinh viên đó học toán hoặc lịch sử.
b) Sinh viên đó không học cả hai môn.
c) Sinh viên đó học lịch sử nhưng không học toán.

Giải:

Gọi A là biến cố " Sinh viên đó học Toán"
B là biến cố " sinh viên học lịch sử"
C là biến cố " sinh viên đó học cả Toán và lịch sử"

a. D là biến cố " sinh viên học toán hoặc học lịch sử

+ Ta có $P(A)=0.54; P(B)=0.69$

$C=AB, D=A\cup B$

+ Áp dụng công thức $P(A\cup B)=P(A)+P(B)-P(AB)$

+ Thay số vào ta có $P(A\cup B)=0.54+0.69-0.35=0.88$

b. + Gọi E là biến cố " sinh viên đó không học cả hai môn"

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

+ Áp dụng công thức De – morgan ta có $E = \overline{A}\overline{B} = \overline{A \cup B}$

+ Nên $P(\text{E}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 0{,}12$

c. + Gọi F là biến cố " sinh viên đó học lịch sử nhưng không học Toán"

+ $E = \overline{A}B$

+ Áp dụng công thức $P(E) = P(\overline{A}B) = P(B) - P(AB) = 0.69 - 0.35 = 0.34$

2.4 (3.t55) Cho một mẫu ngẫu nhiên gồm 200 người đã trưởng thành, được phân loại theo giới tính và trình độ học vấn như sau:

Trình độ học vấn	**Nam**	**Nữ**	**Tổng**
Sơ cấp	38	45	**83**
Trung cấp	28	50	**78**
Cao đẳng	22	17	**39**
Tổng	**88**	**112**	

Chọn ngẫu nhiên một người từ nhóm này, tìm xác suất để:

a) Người được chọn là nam giới, biết rằng người đó có trình độ trung cấp.

b) Người được chọn không có trình độ cao đẳng, biết rằng người đó là nữ giới.

Giải:

Gọi A là biến cố " người đó là nam giới " $P(A) = 0.44$

B là biến cố " người đó có trình độ trung cấp" $P(B) = 0.39$

C là biến cố "người đó vừa là nam và có trình độ trung cấp" $P(C) = 0.14$

a. Gọi D là biến cố " người được chọn là nam giới, biết rằng ng đó có trình độ trung cấp"

+ Ta có $C = AB, D = A|B$

+ Áp dụng công thức $P(D) = P\left(A|B\right) = \dfrac{P(AB)}{P(B)}$

+ Thay số vào ta có $P(D) = \dfrac{0.14}{0.39}$

b. + Gọi F là biến cố " người được chọn có trình độ cao đẳng" $P(F) = 0.195$

Gọi H là biến cố " người được chọn là nữ giới" $P(H) = 0.56$

Gọi E là biến cố " Người được chọn không có trình độ cao đẳng, biết rằng người đó là nữ giới"

+ Ta có $E = \overline{F}|H$

+ Áp dụng công thức $P(E) = \dfrac{P(\overline{F}H)}{P(H)}$

+ Thay số vào ta có $P(E) = \dfrac{95/200}{0.56} = \dfrac{95}{112}$.

2.6 (10.T55) Trong một cuộc trưng cầu dân ý, xác suất để người chồng tham gia bỏ phiếu là 0,21; xác suất để người vợ tham gia bỏ phiếu là 0,28; và xác suất để cả 2 cùng tham gia bỏ phiếu là 0,15. Tìm xác suất để:

a) Có ít nhất 1 người trong gia đình tham gia bỏ phiếu.

b) Người vợ sẽ tham gia bỏ phiếu, biết rằng chồng cô ta cũng tham gia bỏ phiếu.

c) Người chồng sẽ tham gia bỏ phiếu, biết rằng vợ anh ta không tham gia bỏ phiếu.

Giải:

Gọi A là biến cố " người chồng tham gia bỏ phiếu" $P(A) = 0.21$

B là biến cố " người vợ tham gia bỏ phiếu" $P(B) = 0.28$

C là biến cố "cả hai cùng tham gia bỏ phiếu" $P(C) = 0.15$

a. Gọi D là biến cố " có ít nhất 1 người trong gia đình tham gia bỏ phiếu"

+ Ta có $C = AB, D = A \cup B$

+ AP dụng công thức $P(A \cup B) = P(A) + P(B) - P(AB)$

+ Thay số vào ta có $P(D) = 0.21 + 0.28 - 0.15 = 0.34$

b. + Gọi E là biến cố " Người vợ sẽ tham gia bỏ phiếu biết người chồng tham gia bỏ phiếu"

+ Ta có $E = B|A$

+ AP dụng công thức $P(E) = P\left(B|A\right) = \dfrac{P(AB)}{P(A)}$

+ Thay số vào ta có $P(E) = \dfrac{0.15}{0.21}$.

c. + Gọi F là biến cố " Người chồng sẽ tham gia bỏ phiếu biết người vợ không tham gia bỏ phiếu"

+ Ta có $F = A|\overline{B}$

+ AP dụng công thức $P(F) = P\left(A|\overline{B}\right) = \dfrac{P(A\overline{B})}{P(\overline{B})}$

+ Thay số vào ta có $P(F) = P\left(A|\overline{B}\right) = \dfrac{P(A\overline{B})}{P(\overline{B})} = \dfrac{P(A) - P(AB)}{1 - P(B)}$.

2.7 (11.T55) Xác suất để 1 phương tiện có mang biển kiểm soát của Canada tới thăm quan khu hang động Luray là 0.12; xác suất để khách du lịch tới đó cắm trại là 0.28 và xác suất để khách du lịch tới khu hang động cắm trại có sử dụng phương tiện mang biển kiểm soát của Canada là 0.09. Tìm xác suất
a. Một khách du lịch tới khu hang động cắm trại biết rằng người đó sử dụng phương tiện mang biển kiêm soát Canada.
b. Một phương tiện mang biểm kiểm soát Canada tới khu hang động, biết rằng phương tiện đólà của khách du lịch đi cắm trại.
c. một phương tiện tới khu hang động không mang biểm kiểm soát Canada hoặc không phải là của khách du lịch đi cắm trại.

Giải:

Gọi A là biến cố " mang biển kiểm soát Canada" $P(A) = 0.12$

B là biến cố " Khách du lịch tới căm trại" $P(B) = 0.28$

C là biến cố "khách du lịch cắm trại và mang biển kiểm soát của Canada" $P(C) = 0.09$

a. Gọi D là biến cố " khách du lịch tới cắm trại biết mang biển kiểm soát Canada"

+ Ta có $C = AB, D = B|A$.

+ AP dụng công thức $P(B|A) = \dfrac{P(AB)}{P(A)}$

+ Thay số vào ta có $P(D) = \dfrac{0.09}{0.12} = 0.75$

b. + Gọi E là biến cố " mang biển kiểm soát Canada biết khách du lịch cắm trại"

+ Ta có $E = A|B$

+ AP dụng công thức $P(E) = P\left(A|B\right) = \dfrac{P(AB)}{P(\mathrm{B})}$

+ Thay số vào ta có $P(E) = \dfrac{0.09}{0.28}$.

c. + Gọi F là biến cố "không mang biểm kiểm soát Canada hoặc không phải là khách du lịch cắm trại"

+ Ta có $F = \overline{A} \cup \overline{B}$

+ Áp dụng công thức $\overline{A} \cup \overline{B} = \overline{AB} \Rightarrow P(F) = 1 - P(AB)$

+ Thay số vào ta có $P(F) = 0.91$.

2.8 (13.T55) Xác suất để một bác sĩ chuẩn đoán đúng bệnh là 0,7. Nếu bác sĩ chẩn đoán sai, xác suất để bệnh nhân bị chẩn đoán sai phát đơn kiện đòi bồi thường là 0.9. Tìm xác suất để bác sĩ chẩn đoán sai bệnh và bị bệnh nhân phát đơn kiện đòi bồi thường.

Giải:

+ Gọi A là biến cố " Bác sĩ chẩn đoán đúng bệnh" $P(A) = 0.7$

B là biến cố " Bệnh nhân phát đơn kiện" $P(B|\overline{A}) = 0.9$

C là biến cố “chẩn đoán sai bệnh và bị bệnh nhân phát đơn kiện đòi bồi thường”

+ $C = \overline{A}B$

+ Áp dụng công thức $P(\overline{A}B) = P(\overline{A})P(B|\overline{A})$.

+ Thay số vào ta có $P(\overline{A}B) = 0.3 \times 0.9 = 0.27$

2.11(2. T60) Để buộc mọi người lái xe đúng tốc độ quy định, cảnh sát giao thông đặt hệ thông rada bắn tốc độ ở 4 vị trí khác nhau trong thành phố. Tỉ lệ và thời gian hoạt động của mỗi hệ thống ra đa ở mỗi vị trí L_1, L_2, L_3, L_4 lần lượt là 40%, 20%, 20% và 30%. Một người lái xe quá tốc độ quy định phải đi qua các vị trí này với xác suất tương ứng là 0.2; 0.1; 0.5 và 0.2. Tìm xác suất để

a. anh ta nhận biên lai phạt.

b. Nếu người đó nhận biên lai phạt, tìm xác suất để anh ta đi qua hệ thống ra đa đặt ở vị trí L_2.

Giải:

Gọi L_i là biến cố “ người lái xe qua vị trí L_i”.

A là biến cố “ nhận biên lai phạt”

a) Từ giả thiết ta có $P(L_1) = 0.2, P(L_2) = 0.1, P(L_3) = 0.5, P(L_4) = 0.2$.

$P(A|L_1) = 0.4, P(A|L_2) = 0.2, P(A|L_3) = 0.2, P(A|L_4) = 0.3$.

Áp dụng công thức xác suất đầy đủ ta có

$P(A) = P(L_1)P(A|L_1) + P(L_2)P(A|L_2) + P(L_3)P(A|L_3) + P(L_4)P(A|L_4) = 0.26$.

b) C là biến cố “người đó nhận biên lai phạt biết anh ta đi qua vị trí L_2”.

$C = L_2|A$

Áp dụng công thức $P(C) = P(L_2|A) = \dfrac{P(L_2)P(A|L_2)}{P(A)}$.

Thay số vào ta có $P(C) = P(L_2|A) = \dfrac{0.02}{0.26}$

2.12 (8.t61) Một cửa hàng bán sơn Latex và Semigloss. Tỷ lệ khách hàng mua sơn Latex là 75%. Trong số các khách mua sơn Latex, có 60% khách hàng mua kèm chổi lăn sơn. Tỷ lệ khách hàng mua sơn Semigloss kèm chổi lăn sơn là 30%. Chọn ngẫu nhiên 1 khách hàng mua 1 thùng sơn kèm chổi lăn sơn, tính xác suất để khách hàng đó mua loại sơn Latex.

Giải:

Gọi L là biến cố “ Khách hàng mua sơn Latex” $P(L) = 0.75$

S là biến cố “ Khách hang mua sơn Semigloss” $P(S) = 0.25$

A là biến cố “ mua kèm chổi lăn sơn”.

B là biến cố “ khách hang mua sơn Latex, biết khách mua kèm chổi lăn sơn”

+ ta thấy $P(A|L) = 0.6, P(A|S) = 0.3$.

+ Áp dụng công thức xác suất đầy đủ ta có

$P(A) = P(L)P(A|L) + \mathrm{P(S)}P(A|S) = 0.75 \times 0.6 + 0.25 \times 0.3 = 0.525$

+ $B = L|A$. Áp dụng công thức ta có $P(L|A) = \dfrac{P(A|L)P(L)}{P(A)}$

Thay số vào ta có:

2.13 (10. T63) Trong một nhà tù liên bang có 2/3 số tù nhân dưới 25 tuổi. Biết rằng 3/5 số tù nhân là nam, 5/8 số tù nhân là nữ hoặc lớn hơn hoặc bằng 25 tuổi. Tìm xác suất để một tù nhân được chọn ngẫu nhiên từ nhà tù này là nữ và có độ tuổi lớn hơn hoặc bằng 25 tuổi.

Giải:

+ Gọi A là biến cố “ Tù nhân dưới 25 tuổi” $P(A) = \dfrac{2}{3}$.

B là biến cố " Tù nhân là nam giới" $P(B)=\dfrac{3}{5}$

C là biến cố " tù nhân là nữ và có độ tuổi lớn hơn hoặc bằng 25"

+ Theo giả thiết $P(\overline{B}\cup\overline{A})=\dfrac{5}{8}$, $C=\overline{B}\,\overline{A}$

+ Áp dụng công thức De – Morgan. $\overline{B}\cup\overline{A}=\overline{AB}\Rightarrow P(AB)=\dfrac{3}{8}$, $C=\overline{B}\,\overline{A}=\overline{A\cup B}$

+ Khi đó ta có $P(C)=P\left(\overline{A\cup B}\right)=1-P(A\cup B)=1-\left(P(A)+P(B)-P(AB)\right)=\dfrac{13}{120}$.

BÀI 3 – BIẾN NGẪU NHIÊN

+ Biến ngẫu nhiên	**+ Phân phối xác suất rời rạc**	**+ Phân phối liên tục.**

3.1 (3.T75) Giả sử W là bnn chỉ số lần xuất hiện mặt ngửa trừ đi số lần xuất hiện mặt sấp khi tung 1 đồng xu 3 lần. Liệt kê các phàn tử của không gian mẫu S khi tung đồng xu 3 lần ứng với mỗi điểm mẫu xác định giá trị w của W.

3.2 (5.t73) Tìm c để mỗi hàm số sau là hàm phân phối xác suất của biến ngẫu nhiên rời rạc X:

(a) $f(x)=c(x^2+4)$ với $x=0,1,2,3$ (b) $f(x)=cC_2^xC_3^{3-x}$ với $x=0,1,2$.

Giải:

a. Áp dụng tính chất $\sum_{x=0}^{3}f(x)=1$ ta có $4c+5c+8c+13c=1\Rightarrow c=\dfrac{1}{30}$

b. Áp dụng tính chất $\sum_{x=0}^{2}f(x)=1$ ta có $c+6c+3c=1\Rightarrow c=\dfrac{1}{10}$

3.3 (11.t76) Một kiện hàng gồm 7 chiếc tivi trong đó có 2 chiếc bị hỏng. Một khách sạn mua ngẫu nhiên 3 chiếc. Gọi X là số chiếc bị hỏng mà khách sạn đó mua phải, tìm phân phối xác suất của X.

Giải:

Ta có hàm phân phối xác suất $f(x)=P(X=x)=\dfrac{C_2^xC_5^{3-x}}{C_7^3}$

+ Bảng ppxs

x	0	1	2
$f(x)$	2/7	4/7	1/7

3.4 (12.t76) Một công ty đầu tư phát hành đợt trái phiếu có kì hạn biến đổi theo năm. Gọi T là kì hạn (tính theo năm) của một trái phiếu được chọn ngẫu nhiên. Biết T có hàm phân phối tích lũy như sau:

$$F(t)=\begin{cases}0 & ,\ t<1\\ 1/4 & ,\ 1\le t<3\\ 1/2 & ,\ 3\le t<5\\ 3/4 & ,\ 5\le t<7\\ 1 & ,\ t\ge 7\end{cases}.$$

Tính các xác suất sau bằng hai cách:

(a) $P(T=5)$. (b) $P(T>3)$. (c) $P(1{,}4<T<6)$.

Giải:

Cách 1: Bảng ppxs là

t	1	3	5	7
$f(t)$	$\dfrac{1}{4}$	$\dfrac{1}{4}$	$\dfrac{1}{4}$	$\dfrac{1}{4}$

Nên $P(T=5)=f(5)=\frac{1}{4}$

$P(T>3)=f(5)+f(7)=\frac{1}{2}$

$P(1.4<T6<)=f(3)+f(5)=\frac{1}{2}$

Cách 2: Áp dụng công thức

$P(T=5)=P(T\le 5)-P(T<5)=F(5)-F(3)=\frac{1}{4}$

$P(T>3)=1-P(T\le 3)=1-F(3)=\frac{1}{2}$

$P(1.4<T<6)=F(6)-F(1.4)-P(X=6)=\frac{1}{2}$

3.5 (13.t77) Phân phối xác suất của *X*, trong đó *X* là số lỗi trên 10m vải sợi tổng hợp trong một súc vải có độ rộng giống nhau, được cho bởi bảng sau:

x	0	1	2	3	4
$f(x)$	0,41	0,37	0,16	k	0,01

a) Tìm *k*.
b) Tìm hàm phân phối tích lũy của biến ngẫu nhiên *X*.

Giải:

a) k = 0,05.
b) Hàm phân phối tích lũy

$$F(x)=\begin{cases}0 & x<0\\ 0,41 & 0\le x<1\\ 0,78 & 1\le x<2\\ 0,94 & 2\le x<3\\ 0.99 & 3\le x<4\\ 1 & x>4\end{cases}$$

3.6 (25.t78) Một hộp chứa 4 đồng một đồng và 2 đồng năm xu. Chọn ngẫu nhiên 3 đồng tiền. Tìm phân phối xác suất của tổng số tiền *T* của 3 đồng tiền được chọn.

Giải:

+ Gọi T là bnn tổng số tiền được chọn. $T=20,25,30$

Bảng ppxs

T	20	25	30
$f(t)$	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

3.7 (22.t76) Rút ngẫu nhiên liên tiếp không hoàn lại 3 quân bài từ một bộ bài. Tìm phân phối xác suất của số quân bích rút được.

Giải:

+ Gọi X là bnn chỉ số quân bích rút được. $X=0,1,2,3$.

Ta có hàm phân phối xác suất $f(x)=P(X=x)=\frac{C_3^x A_{13}^x A_{39}^{3-x}}{A_{52}^3}$

+ Bảng ppxs

x	0	1	2	3

$f(x)$	$\frac{703}{1700}$	$\frac{741}{1700}$	$\frac{117}{850}$	$\frac{11}{850}$

3.8 (26.t76) Một hộp có 4 quả bóng đen và 2 quả bóng xanh. Lấy ngẫu nhiên lần lượt 3 quả bóng theo phương thức có hoàn lại. Lập bảng phân phối xác suất của số quả bóng xanh lấy được.

Giải:

+ Gọi X là bnn chỉ số qủa bóng xanh lấy được. $X = 0,1,2,3$

+ Ta có hàm phân phối xác suất $f(x) = P(X = x) = \dfrac{C_3^x 2^x 4^{3-x}}{6^3}$

+ Bảng ppxs

x	0	1	2	3
$f(x)$	$\frac{8}{27}$	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{1}{27}$

3.9 (9.t74) Tỷ lệ người trả lời các thư chào hàng qua đường bưu điện là một biến ngẫu nhiên liên tục X có hàm mật độ như sau:

$$f(x) = \begin{cases} \dfrac{2(x+2)}{5} & ; \quad 0 < x < 1 \\ 0 & ; \quad x \notin (0,1) \end{cases}.$$

a) Hãy chứng minh $P(0 < X < 1) = 1$.

b) Tìm xác suất để có từ 1/4 đến 1/2 số người được liên hệ trả lời các thư chào hàng nói trên.

Giải:

a) $P(0<X<1) = \int_{-\infty}^{+\infty} f(x)dx = \frac{2}{5}\int_0^1 (x+2)f(x)dx = \frac{2}{5}\left(\frac{x^2}{2} + 2x\right)\Big|_0^1 = 1$.

b) $P\left(\frac{1}{4} < X < \frac{1}{2}\right) = \frac{2}{5}\int_{\frac{1}{4}}^{\frac{1}{2}} (x+2)f(x)dx = \frac{2}{5}\left(\frac{x^2}{2} + 2x\right)\Big|_{\frac{1}{4}}^{\frac{1}{2}} = \frac{19}{80}$

3.10 (21.T78) Xét hàm mật độ

$$f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1 \\ 0, & x \notin (0,1) \end{cases}.$$

a) Tìm k.

b) Tìm $F(x)$ và sử dụng $F(x)$ để tính $P(0,3 < X < 0,6)$.

Giải:

a) $k = \frac{3}{2}$.

b) $F(x) = \int_{-\infty}^{+\infty} f(x)dx = \begin{cases} x^{\frac{3}{2}} & x \in (0,1) \\ 0 & x \notin (0,1) \end{cases}$. $P(0,3 < X < 0,6) = F(0,6) - F(0,3) = 0,3004$

3.11 (14.t75) Thời gian chờ (tính theo giờ) giữa 2 lần bắn liên tiếp của một thiết bị bắn tốc độ ô tô sử dụng công nghệ rada là một biến ngẫu nhiên liên tục có hàm phân phối tích lũy như sau:

$$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-8x}, & x > 0 \end{cases}.$$

Tìm xác suất để thời gian chờ ít hơn 12 phút bằng cách:

a) Sử dụng hàm mật độ tích lũy của X.

b) Sử dụng hàm mật độ xác suất của X.

Giải:

Đổi đơn vị $12' = 0,2h$.

$$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-8x}, & x > 0 \end{cases} \Rightarrow f(x) = \begin{cases} 0, & x \le 0 \\ 8e^{-8x}, & x > 0 \end{cases}$$

a) Áp dụng công thức $P(X<0.2)=\int_{-\infty}^{0,2} f(x)dx=\int_{0}^{0,2} 8e^{-8x}dx=-e^{-8x}\Big|_0^{0.2}=1-e^{-1.6}$.

b) $P(X<0.2)=F(0.2)-F(-\infty)=1-e^{-1.6}$.

BÀI 4 – BIẾN NGẪU NHIÊN HAI CHIỀU

4.1(1.t94) Xác định giá trị của c để các hàm số sau là phân phối xác suất đồng thời của các biến ngẫu nhiên X và Y:

a) $f(x,y)=cxy$ với $x=1,2,3$; $y=1,2,3$.

b) $f(x,y)=c|x-y|$ với $x=-2,0,2$; $y=-2,3$.

Giải:

a) $c=\frac{1}{36}$.

b) $c=\frac{1}{15}$.

4.2(2.t94) Cho phân phối xác suất đồng thời của X và Y là : $f(x,y)=\frac{x+y}{30}$, với $x = 0,1,2,3$; $y = 0,1,2$.

Tính:

(a) $P(X\le 2, Y=1)$ (b) $P(X>2, Y\le 1)$ (c) $P(X>Y)$ (d) $P(X+Y=4)$.

Giải

a) $P(X\le 2,Y=1)=f(0,1)+f(1,1)+f(2,1)=\frac{1}{5}$

b) $P(X>2,Y\le 1)=f(3,0)+f(3,1)=\frac{7}{30}$

c) $P(X>Y)=f(1,0)+f(2,0)+f(3,0)+f(2,1)+f(3,1)+f(3,2)=\frac{3}{5}$.

d) $P(X+Y=4)=f(2,2)+f(3,1)=\frac{4}{15}$.

4.3(3.t93) Từ một túi trái cây gồm 3 quả cam, 2 quả táo và 3 quả chuối, lấy ngẫu nhiên ra 4 quả. Gọi X là số quả cam, Y là số quả táo được lấy ra, tìm:

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

b) $P[(X,Y)\in A]$, trong đó A là miền $\{(x,y)|\, x+y\le 2\}$.

Giải

a) Gọi X là số quả cam, Y là số quả táo được lấy ra được $X=\{0,1,2,3\}$, $Y=\{0,1,2\}$

+ Hàm phân phối xác suất là $f(x,y)=P(X=x,Y=y)=\frac{C_3^x C_2^y C_3^{4-x-y}}{C_8^4}$

b) $A=\{(x,y)|\, x+y\le 2\}=\{(0,1),(0,2),(1,0),(1,1),(2,0)\}$. $P[(X,Y)\in A]=\frac{1}{2}$

4.4(13.t94) Giả sử một biến ngẫu nhiên hai chiều có bảng phân phối xác suất đồng thời như sau:

f(x,y)	*X*	1	2	3
	1	0,05	0,05	0,1
Y	2	0,05	0,1	0,35
	3	0	0,2	0,1

Tìm:

a) Phân phối biên duyên của X.

b) Phân phối biên duyên của Y.

c) $P(Y=3|X=2)$.

Giải

a) Phân phối biên duyên của X là

x	1	2	3
$g(x)$	0.1	0.35	0.55

b) Phân phối biên duyên của Y

y	1	2	3
$h(y)$	0.2	0.5	0.3

c) $P(Y=3|X=2)$

y	1	2	3
$f(y\|2)$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$

$$P(Y=3|X=2)=f(y=3|x=2)=\frac{4}{7}$$

4.5(4.t95) Một cửa hàng rượu tổ chức bán rượu tại quầy và trong các tủ trưng bày. Chọn ngẫu nhiên 1 ngày, gọi X và Y lần lượt là tỷ lệ thời gian hoạt động của quầy rượu và tủ rượu. Biết hàm mật độ đồng thời của X và Y là :

$$f(x,y)=\begin{cases}\frac{2}{3}(x+y) & , 0\le x\le 1,\quad 0\le y\le 1,\\ 0 & , \text{tại các điểm khác.}\end{cases}$$

Tìm : (a) Hàm mật độ biên duyên của X. (b) Hàm mật độ biên duyên của Y.
(c) Xác suất để thời gian hoạt động của quầy rượu nhỏ hơn một nửa ngày.

Giải Gọi X và Y lần lượt là tỷ lệ thời gian hoạt động của quầy rượu và tủ rượu.

a) Hàm mật độ biên duyên của X **:**

$$g(x)=\begin{cases}\frac{2}{3}x+\frac{1}{3} & , 0\le x\le 1,\\ 0 & , \text{tại các điểm khác.}\end{cases}$$

b) Hàm mật độ biên duyên của Y **:**

$$h(y)=\begin{cases}y+\frac{1}{3} & , 0\le y\le 1,\\ 0 & , \text{tại các điểm khác.}\end{cases}$$

c) $P(0<X<0.5)=\int_0^{0.5} g(x)dx=0,25$

4.6(6.t95) Giả sử X và Y là tuổi thọ (tính theo năm) của hai bộ phận trong một hệ thống điện tử. Biết hàm mật độ đồng thời của các biến ngẫu nhiên này là:

$$f(x,y)=\begin{cases}e^{-(x+y)} & , x>0,\quad y>0\\ 0 & , \text{tại các điểm khác.}\end{cases}$$

Tính $P(0<X<1|Y=2)$.

Giải: Hàm mật độ biên duyên của Y **:** $h(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^{+\infty} \mathrm{e}^{-(x+y)}\,dx=e^{-y}$.

$$f(x|y)=\frac{f(x,y)}{h(y)}=\frac{e^{-(x+y)}}{e^{-y}}=e^{-x}.\ P(0<X<1|Y=2)=\int_0^1 f(x|y)\,dx=\int_0^1 e^{-x}dx=1-\frac{1}{e}$$

4.7(5.t93) Một công ty kẹo phân phối các hộp kẹo sôcôla tổng hợp với các loại nhân kem, nhân bơ cứng và nhân rượu. Giả sử trọng lượng của mỗi hộp là 1kg, nhưng trọng lượng của từng loại nhân kem, nhân bơ cứng và nhân rượu ở mỗi hộp là khác nhau. Chọn ngẫu nhiên một hộp, gọi X và Y lần lượt là trọng lượng của kẹo sôcôla nhân kem và kẹo sôcôla nhân bơ cứng trong hộp đó. Biết hàm mật độ đồng thời của các biến ngẫu nhiên này là:

$$f(x,y)=\begin{cases}24xy & , 0\le x\le 1,\quad 0\le y\le 1,\quad x+y\le 1.\\ 0 & , \text{ tại các điểm khác.}\end{cases}$$

a) Tìm xác suất để trong một hộp được chọn có lượng sôcôla nhân rượu lớn hơn 1/2 trọng lượng của hộp.
b) Tìm hàm mật độ biên duyên của trọng lượng sôcôla nhân kem.
c) Tìm xác suất để trọng lượng của kẹo sôcôla nhân bơ cứng trong một hộp ít hơn 1/8 kg biết rằng trọng lượng của kẹo sôcôla nhân kem trong hộp đó là 3/4 kg.

Giải: Gọi X và Y lần lượt là trọng lượng của kẹo sôcôla nhân kem và kẹo sôcôla nhân bơ cứng trong hộp

a) xác suất để trong một hộp được chọn có lượng sôcôla nhân rượu lớn hơn 1/2 trọng lượng của hộp, tức là trọng lượng của nhân kem và nhân bơ cứng nhỏ hơn ½.

$$P\left(x+y\le\frac{1}{2}\right)=\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}-x}24xydydx=\frac{1}{16}$$

b) Hàm mật độ biên duyên của X $g(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\int_0^{1-x}24xydy=12x(1-x)^2$.

c) xác suất để trọng lượng của kẹo sôcôla nhân bơ cứng trong một hộp ít hơn 1/8 kg biết rằng trọng lượng của kẹo sôcôla nhân kem trong hộp đó là 3/4 kg là $P\left(y\le\frac{1}{8}\middle|x=\frac{3}{4}\right)$.

$$f\left(y|x\right)=\frac{f(x,y)}{g(x)}=\frac{24xy}{12x(1-x)^2}=\frac{2y}{(1-x)^2}$$

$$P\left(y\le\frac{1}{8}\middle|x=\frac{3}{4}\right)=\int_0^{\frac{1}{8}}\frac{2y}{(1-0.75)^2}dy=\frac{1}{4}$$

BÀI 5 – CÁC SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN MỘT CHIỀU

5.1(2.t107) Phân phối xác suất của biến ngẫu nhiên rời rạc X là

$$f(x)=C_3^x\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x},\quad x=0,1,2,3.$$

Hãy tìm trung bình của X.

Giải: $\mu=\sum xf(x)=0.75$.

5.2(5.t107) Cho biến ngẫu nhiên X có bảng phân phối xác suất như sau

X	0	1	2	3	4
$f(x)$	0,41	0,37	0,16	0,05	0,01

Hãy tìm số lỗi trung bình trên mỗi 10 mét của loại sợi vải nói trên.

Giải: $\mu=\sum xf(x)=0.88$.

5.3(7.t107) Bằng cách đầu tư vào cổ phần đã xác định, một người có thể kiếm được 4000USD một năm với xác suất 0,3 hoặc lỗ 1000USD với xác suất 0,7. Người này hy vọng sẽ kiếm được bao nhiêu trong một năm?

Giải: $\mu=\sum xf(x)=500$.

5.4 (12.t108) Nếu lợi nhuận của một nhà kinh doanh ô tô (tính theo đơn vị 5000USD), thu được từ mỗi chiếc xe ô tô là một biến ngẫu nhiên X với hàm mật độ như sau

$$f(x)=\begin{cases}2(1-x), & x\in(0;1)\\ 0, & x\notin(0;1)\end{cases}.$$

Tìm lợi nhuận trung bình từ 1 chiếc ô tô

Giải: Lợi nhuận trung bình là $1666,6667$ (đô la)

5.5 (15.t108) Hàm mật độ của biến ngẫu nhiên liên tục *X*, biểu thị tổng số giờ (theo đơn vị 100 giờ) mà một gia đình sử dụng máy hút bụi trong một năm, được cho ở Bài tập 3.9, như sau

$$f(x)=\begin{cases} x, & x\in(0;1) \\ 2-x, & x\in[1;2) \\ 0, & x\notin(0;1)\cup[1;2) \end{cases}.$$

Hãy tìm số giờ sử dụng máy hút bụi trung bình của một gia đình trong một năm?

Giải: **Số** giờ dùng máy hút bụi trung bình là 100(giờ)

5.6 (17.t109) Gọi X là biến ngẫu nhiên với phân phối xác suất như sau:

x	-3	6	9
$f(x)$	1/6	1/2	1/3

Hãy tìm kỳ vọng của biến ngẫu nhiên $g(X)=(2X+1)^2$.

Giải: $\mu_{g(x)}=209$.

5.7 (20.t109) Biến ngẫu nhiên liên tục X có hàm mật độ là

$$f(x)=\begin{cases} e^{-x}, & \text{khi } x>0 \\ 0, & \text{khi } x\le 0 \end{cases}.$$

Hãy tìm kỳ vọng của biến ngẫu nhiên $g(X)=e^{\frac{2X}{3}}$.

Giải: $\mu_{g(X)}=3$

5.8 (2.t117) Cho X là biến ngẫu nhiên với phân phối xác suất như sau:

X	-2	3	5
$f(x)$	0,3	0,2	0,5

Hãy tìm độ lệch chuẩn của *X*.

Giải: $\sigma=\sqrt{E(X^2)-\mu^2}=3,0414$

5.9 (6.t117) Tỷ lệ người trả lời các thư chào hàng qua đường bưu điện là một biến ngẫu nhiên liên tục X có hàm mật độ như sau:

$$f(x)=\begin{cases} \dfrac{2(x+2)}{5} & ;\ 0<x<1 \\ 0 & ;\ x\notin(0,1) \end{cases}.$$

Hãy tìm phương sai của *X*.

Giải: $\sigma^2=0.0822$

5.10 (11.t117) Thời gian, tính theo đơn vị phút, để một chiếc máy bay nhận được giấy phép cất cánh tại một sân bay nào đó là biến ngẫu nhiên $Y=3X-2$, trong đó X là biến ngẫu nhiên có hàm mật độ

$$f(x)=\begin{cases} \dfrac{1}{4}e^{\frac{-x}{4}}, & x>0 \\ 0, & x\le 0 \end{cases}.$$

Hãy tìm giá trị trung bình và phương sai của biến ngẫu nhiên *Y*.

Giải:
$$\mu_{3X-2}=3\mu_X-2=10.$$
$$\sigma^2_{3X-2}=9\sigma^2_X=144$$

BÀI 6 – CÁC SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN HAI CHIỀU

6.1(8.t128) Giả sử X, Y là hai biến ngẫu nhiên độc lập với phân phối xác suất đồng thời như sau

$f(x,y)$		x:	2	4
	1		0,10	0,15
Y	3		0,20	0,30
	5		0,10	0,15

Hãy tìm (a) $E(2X-3Y)$; (b) $E(XY)$.

Giải:

a) $E(2X-3Y)=-2.6$

b) $E(XY)=9,6$

6.2(3.t93) Từ một túi trái cây gồm 3 quả cam, 2 quả táo và 3 quả chuối, lấy ngẫu nhiên ra 4 quả. Gọi X là số quả cam, Y là số quả táo được lấy ra, tìm :

a) $E(X-Y)$.

b) Tính covaricance của X và Y.

Giải:

+ $E(X-Y)=0,5$

+ $\sigma_{XY}=-\dfrac{3}{14}$

6.3(1.t97_ĐHH) Cho biến ngẫu nhiên 2 chiều với bảng phân phối xác suất đồng thời như sau:

$f(x,y)$	x	20	40	60
	10	3λ	λ	0
y	20	2λ	4λ	2λ
	30	λ	2λ	5λ

(a) Tìm λ. (b) Lập bảng phân phối của X và Y. (c) Tính hệ số tương quan của X và Y.

Giải :

a) $\lambda=\dfrac{1}{20}$.

b)

$f(x,y)$	x	20	40	60
	10	0,15	0,05	0
y	20	0,1	0,2	0,1
	30	0,05	0,1	0,25

c) Hệ số tương quan $\rho_{XY}=0,5646$.

6.4(21.tr127) Nếu hàm mật độ đồng thời của X và Y được cho như sau :

$$f(x,y)=\begin{cases}\dfrac{2}{7}(x+2y) & ; \quad (x,y)\in(0,1)\times(1,2)\\ 0 & ; \quad (x,y)\notin(0,1)\times(1,2)\end{cases}.$$

(a) Tìm kỳ vọng của $\dfrac{X}{Y^3}+X^2Y$. (b) Tìm covaricance của X và Y.

Giải:

a) kỳ vọng của $\dfrac{X}{Y^3}+X^2Y$ là $E\left(\dfrac{X}{Y^3}+X^2Y\right)=0,6587$

b) $\sigma_{XY}=-\dfrac{293}{441}$

6.5 (2.t130) Tìm covariance và hệ số tương quan của các biến ngẫu nhiên X và Y biết hàm mật độ đồng thời như sau:

$$f(x,y)=\begin{cases}(x+y) & ; \quad (x,y)\in(0,1)\times(0,1)\\ 0 & ; \quad (x,y)\notin(0,1)\times(0,1)\end{cases}.$$

Giải :

+ $\sigma_{XY}=-\dfrac{1}{144}$

Hệ số tương quan là $\rho_{XY}=\dfrac{\sigma_{XY}}{\sigma_X\sigma_X}=-\dfrac{1}{11}$

BÀI 7 – MỘT SỐ PHÂN PHỐI THƯỜNG GẶP

7.1 (4.t141) Tại một thành phố nào đó, 75% các vụ trộm cắp là do muốn có tiền mua ma túy. Hãy tìm xác suất để trong 5 vụ trộm cắp liên tiếp tại thành phố này,

a) Có đúng 2 vụ là do muốn có tiền mua ma túy.
b) Nhiều nhất là 3 vụ do muốn có tiền mua ma túy.

a) Xác xuất để có đúng 2 vụ do muốn có tiền mua ma túy là $P(X=2)=C_5^2 0,75^2 0,25^3=0,0879$

b) Có nhiều nhất 3 vụ $P(X\le 3)=\sum_{x=0}^{3}b(\mathrm{x},5,0,75)=0,3672$

7.2 (7.141) Một bác sĩ có uy tín tuyên bố rằng 70% trong tổng số người mắc ung thư phổi là những người nghiện thuốc lá. Nếu khẳng định của ông ta là đúng,

a) Tìm xác suất để 10 bệnh nhân mắc ung thư phổi nhập viện gần đây có dưới một nửa là người nghiện thuốc lá.
b) Tìm xác suất để 20 bệnh nhân mắc ung thư phổi nhập viện gần đây có dưới một nửa là người nghiện thuốc lá.

Giải:

a) Xác xuất để có dưới 1 nửa là người nghiện là $P(X<5)=\sum_{x=0}^{4}b(x;10;0.7)=0,0474$

b) Đây là phân phối nhị thức với $P(X<10)=\sum_{x=0}^{9}b(x,20;0,7)=0,0171$

7.3 (9.141) Khi chảy thử 1 loại lốp xe tải nào đó trên vùng địa hình gồ ghề, người ta thấy rằng 25% trong số các xe tải không hoàn thành bài chạy thử do bị nổ lốp. Trong số 15 chiếc chạy thử tiếp Theo, hãy tìm xác suất để

a) Có từ 3 đến 6 chiếc bị nổ lốp.
b) có ít hơn 4 xe bị nổ lốp.
c) nhiều hơn 5 xe bị nổ lốp.

Giải:

a) Xác suất có từ 3 đến 6 chiếc bị nổ lốp là $P(3\le X\le 6)=0.7073$

b) Xác suất có ít hơn 4 xe bị nổ lốp là $P(X\le 3)=\sum_{x=0}^{3}b(x,15;0.25)=0.4613$

c) Xác suất có nhiều hơn 5 xe bị nổ lốp $P(X>5)=1-P(X\le 5)=0.1484$

7.5 (2.t174) Tìm giá trị của z nếu diện tích của phần nằm bên dưới đường cong chuẩn và
(a) ở bên trái của z là 0.1131; (b) ở bên phải của z là 0.3622;
(c) ở giữa 0 và z (với $z>0$), là 0.4838; (d) giữa –z và z (với $z>0$) là 0.9500.

Giải:

a) $P(Z<z)=0.1131\Rightarrow z=-1.21$
b) $P(Z>z)=0.3622$. $P(Z>z)=1-P(Z<z)=0.3622\Rightarrow P(Z<z)=0.6378\Rightarrow z=0.35$
c) $P(0<Z<z)=0.4838$. $P(0<Z<z)=P(Z<z)-P(Z<0)\Rightarrow P(Z<z)=0.9838$. Nên $z=2.14$.
d) $P(-\mathrm{z}<Z<z)=0.9500\Rightarrow 1-2\alpha=0.95\Rightarrow \alpha=0.025$. Nên $P(Z<z)=1-\alpha=0.975\Rightarrow z=1.96$.

7.5 (5.t174) Cho biến ngẫu nhiên X tuân theo phân phối chuẩn với trung bình 18 và độ lệch tiêu chuẩn bằng 2.5, tìm

(a) $P(X<15)$; (b) giá trị của k thỏa mãn $P(X<k)=0.2236$;
(c) giá trị của k thỏa mãn $P(X>k)=0.1814$; (d) $P(17<X<21)$.

Giải:

X là biến ngẫu nhiên tuân theo pp chuẩn với $\mu=18, \sigma=2.5$. Đặt $Z=\frac{X-\mu}{\sigma}$.

a) Với $x=15 \Rightarrow z=-1.2$, ta có $P(X<15)=P(Z<-1.2)=0{,}1151$.

b) $P(X<k)=P(Z<\frac{k-18}{2.5})=0.2236 \Rightarrow \frac{k-18}{2.5}=-0{,}76 \Rightarrow k=16{,}1$.

c) $P(X>k)=1-P(X<k)=0.1814 \Rightarrow P(X<k)=0{,}8186 \Rightarrow z=0.91 \Rightarrow k=20{,}275$.

d) Với $x=17 \Rightarrow z=-0{,}4; x=21 \Rightarrow z=1{,}2$.
Nên $P(17<X<21)=P(-0{,}4<Z<1.2)=P(Z<1{,}2)-P(Z<-0{,}4)=0{,}8849-0{,}3446=0{,}5403$

7.6 (11.t174) Một luật sư đi làm hàng ngày từ nhà tới cơ quan, thời gian trung bình cho một lần đi là 24 phút, với độ lệch chuẩn là 3.8 phút. Giả sử thời gian của mỗi lần đi có phân phối chuẩn.

a) Hỏi với xác suất bằng bao nhiêu thì một lần đi mất ít nhất 1/2 h?
b) Nếu cơ quan anh ta mở cửa vào lúc 9:00 sáng và anh ta rời nhà lúc 8:45 hằng ngày thì số ngày anh ta đi muộn chiếm bao nhiêu phần trăm?
c) Nếu anh ta rời nhà vào lúc 8:35 và tại cơ quan anh ta cà phê chỉ được phục vụ từ 8:50 đến 9:00 sáng thì xác suất để một lần nào đó anh ta không được phục vụ cà phê là bao nhiêu?
d) Tìm xác suất để có 2 trong 3 hành trình tiếp theo anh ta đi hết ít nhất 1/2 h.

Giải

Đặt $Z=\frac{X-\mu}{\sigma}$

a) $P(X>30)=1-P(X<30)=1-P(Z<1.58)=1-0{,}9428=0{,}0572$.

b) $P(X>15)=1-P(X<15)=1-P(Z<-2.37)=1-0{,}0089=0.9911$.

c) Xác xuất để anh ta được phục vụ cà phê là $P(X<25)=P(Z<0{,}26)=P(Z<0{,}26)=0{,}6026$
Vậy xs để anh ta không được phục vụ cà phê là 1 – 0,6026 =0,3974.

d) $P(X=2)=b(3,2,0{,}0572)=C_3^2(0{,}0572)^2(1-0{,}0572)=0{,}0093$.

7.8 (19.t174) Chỉ số IQ của 600 người nộp đơn xin học ở một trường đại học có phân phối xấp xỉ chuẩn với trung bình là 115 và độ lệch tiêu chuẩn là 12. Nếu trường đại học yêu cầu chỉ số IQ phải đạt ít nhất là 95, thì có bao nhiêu sinh viên sẽ bị loại trong đợt tuyển chọn hồ sơ của họ bởi tiêu chí trên?

Giải: Đặt $Z=\frac{X-\mu}{\sigma}$. $P(X<95)=P(Z<-1{,}67)=0{,}0475$.

Số sinh viên bị loại là 600x0,0475=28.5. Số sv bị loại là 29 em.

BÀI 9 – ƯỚC LƯỢNG CHO GIÁ TRỊ TRUNG BÌNH

9.1

a. (4.t278) Một công ty sản xuất các bóng đèn có tuổi thọ tuân theo phân phối chuẩn với độ lệch chuẩn 40 giờ. Nếu một mẫu 30 bóng có tuổi thọ trung bình là 780 giờ

a) hãy xác định khoảng tin cậy 96% đối với kỳ vọng tổng thể của tất cả các bóng điện do công ty này sản xuất .

b) mẫu cần lớn bao nhiêu nếu chúng ta mong muốn rằng với độ tin cậy 96%, trung bình mẫu sẽ sai khác với giá trị trung bình tổng thể không quá 10 giờ.

Giải:

a) Đây là bài toán ước lượng 1 trung bình đã biết phương sai. $765.0289<\mu<794.9711$

b) $n \geq 68$

9.2 (6.t279) Chiều cao của một mẫu ngẫu nhiên 50 sinh viên đại học cho thấy có giá trị trung bình 174.5cm và độ lệch chuẩn 6.9cm.

a) Xác định khoảng tin cậy 98% cho chiều cao trung bình của tất cả sinh viên đó;
b) Chúng ta có thể khẳng định điều gì với độ tin cậy 98% về sai số nếu chúng ta ước lượng chiều cao trung bình của tất cả các sinh viên là 174.5cm

Giải:

a) Đây là bài toán ước lượng 1 trung bình chưa biết σ cỡ mẫu lớn. $172.2264<\mu<176.7736$

b) Sai số giữa chiều cao trung bình mẫu và chiêu cao trung bình tổng thể là $e = z_{\alpha/2}\dfrac{s}{\sqrt{n}} = 2,2736$

9.3(10.279) Một chuyên gia về hiệu quả muốn xác định thời gian trung bình cần để khoan 3 lỗ trên 1 khóa kim loại. Anh ấy cần cỡ mẫu lớn như thế nào để tin cậy 95% rằng số trung bình mẫu của anh ấy nằm trong khoảng 15s của trung bình chân thực. Giả thiết trong nghiên cứu trước đó $\sigma = 40s$.
Giải:
Đây là bài toán xác định cỡ mẫu đã biết phương sai nên dùng bảng tra A3.
Thay số vào ta có $n \geq \left(z_{\alpha/2}\dfrac{\sigma}{e}\right)^2 = 27.3180$. Chọn n =28.

9.4 (13.t273) Một máy sản xuất các mảnh kim loại có hình trụ. Một mẫu các mảnh được lấy ra và các đường kính là 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01 và 1.03cm. Xác định khoảng tin cậy 99% đối với đường kính trung bình của các mảnh được sản xuất ra, giả thiết đường kính có phân phối xấp xỉ chuẩn.
Giải:
Đây là bài toán ước lượng 1 trung bình, chưa biết phương sai, cỡ mẫu nhỏ $0.9781 < \mu < 1.0331$

9.5 (15.t273) Một mẫu ngẫu nhiên 12 chốt nghiền được lấy trong một nghiên cứu về độ cứng Rockwell của đầu trên chốt. Các lần đo được tiến hành lần lượt cho 12 chốt, cho giá trị trung bình 48.50 với độ lệch chuẩn mẫu là 1.5. Giả thiết các giá trị đo có phân phối chuẩn, xác định một khoảng tin cậy 90% cho độ cứng Rockwell trung bình.
Giải:
Đây là bài toán ước lượng 1 trung bình, chưa biết phương sai, cỡ mẫu nhỏ $47.7223 < \mu < 49.2777$

BÀI 10 – ƯỚC LƯỢNG HIỆU HAI TRUNG BÌNH VÀ TỶ LỆ

10.1 (1.t290) Một mẫu ngẫu nhiên kích thước n_1=25 được lấy từ một tổng thể có phân phối chuẩn có độ lệch chuẩn σ_1=5 và giá trị trung bình $\overline{x}_1 = 80$. Một mẫu ngẫu nhiên thứ hai kích thước $n_2 = 6$ được lấy từ một tổng thể khác cũng có phân phối chuẩn với độ lệch chuẩn σ_2=3, và giá trị trung bình $\overline{x}_2 = 75$. Xác định khoảng tin cậy 94% cho $\mu_1 - \mu_2$.
Giải:
Đây là bài toán UL hiệu hai TB, đã biết phương sai. $2.0259 < \mu_1 - \mu_2 < 7.9741$

10.2 (2.290) Hai loại ren được so sánh về cường độ. Tiến hành kiểm tra 50 mẫu của mỗi loại trong cùng điều kiện. Nhãn hang A có cường độ kéo căng trung bình 78.3kg với độ lệch chuẩn 5.6kg, trong khi đó nhã hang B có cường độ kéo căng trung bình 87.2kg với độ lệch chuẩn 6.3kg. Xác định khoảng tin cậy 95% cho hiệu số của các giá trị trung bình.
Giải
Đây là bài toán UL hiệu hai TB, chưa biết phương sai, cỡ mẫu lớn. $-11.2364 < \mu_1 - \mu_2 < -6.5636$

10.3 (4.t290) Trong một phản ứng hóa học, hai chất xúc tác được so sánh về tác động lên hiệu suất của quá trình phản ứng. Một mẫu 12 phản ứng được sử dụng chất xúc tác 1 và một mẫu 10 phản ứng được sử dụng chất xúc tác 2. 12 mẫu sử dụng chất xúc tác 1 cho khối lượng bình quân 85 với độ lệch chuẩn mẫu là 4 và khối lượng bình quân cho mẫu thứ hai là 81 với độ lệch chuẩn là 5. Xác định khoảng tin cậy 90% cho hiệu trung bình hai tổng thể, giả thiết các tổng thể có phân phối xấp xỉ chuẩn với phương sai bằng nhau.
Giải

$n_1 = 12; \overline{x}_1 = 85; s_1 = 4$; $n_2 = 10; \overline{x}_2 = 81; s_1 = 5$; $s_p = \sqrt{\dfrac{(n_1-1)s_1 + (n_2-1)s_2}{n_1+n_2-2}} = 3,2558$

Đây là bài toán ước lượng hiệu hai TB, chưa biết ps, cỡ mẫu nhỏ, ps bằng nhau
Áp dụng công thức

$$\overline{x}_1-\overline{x}_2-t_{\alpha/2}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le \mu_1-\mu_2\le \overline{x}_1-\overline{x}_2+t_{\alpha/2}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

với $1-\alpha=0,90\Rightarrow t_{\alpha/2}=1,725$

Thay số vào ta có $1,5953\le \mu_1-\mu_2\le 6,4047$

10.4 (**7. t285**) Dữ liệu sau đây, được ghi nhận theo ngày, thể hiện khoảng thời gian hồi phục đối với các bệnh nhân được điều trị ngẫu nhiên bằng một trong hai loại thuốc để điều trị nhiễm trùng:

Loại thuốc 1	**Loại thuốc 2**
$n_1=14$	$n_2=16$
$\overline{x_1}=17$	$\overline{x_2}=19$
$s_1^2=1,5$	$s_2^2=1,8$

Xác định khoảng tin cậy 99% cho hiệu số $\mu_1-\mu_2$ về hiệu thời gian hồi phục trung bình cho hai loại thuốc, giả thiết các tổng thể có phân phối chuẩn với phương sai bằng nhau.

10.5 (1.t299)

a) Một mẫu ngẫu nhiên 200 cử tri được lựa chọn và 114 được xác định ủng hộ một ứng viên. Xác định khoảng tin cậy 96% cho tỷ lệ cử tri ủng hộ cho ứng viên đó.

b) Cần một mẫu lớn như thế nào nếu chúng ta muốn với độ tin cậy 96%, tỷ lệ mẫu sẽ sai khác so với tỷ lệ tổng thể không quá 0,02.

Giải

$n=200; x=114; \hat{p}=0.57$

Đây là bài toán ước lượng 1 tỉ lệ. Áp dụng công thức $\hat{p}-z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\le p\le \hat{p}+z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

Với $1-\alpha=0,96\Rightarrow 1-\frac{\alpha}{2}=0.98\Rightarrow z_{\alpha/2}=2.05$

Thay số vào ta có $0,4982\le p\le 0,6417$

b. Áp dụng công thức $n\ge \frac{z^2{}_{\alpha/2}\hat{p}\hat{q}}{e^2}$. Thay số vào ta có n ~2576

10.6 (15.t294) Một nhà nghiên cứu gen quan tâm đến tỷ lệ nam giới và nữ giới trong tổng thể bị rối loạn tiểu cầu. Trong một mẫu ngẫu nhiên gồm 1000 nam giới thì có 250 người được xác định bị rối loạn tiểu cầu, đối với nữ, tỷ lệ này là 275 người trong số 1000 nữ giới được kiểm tra. Tìm khoảng tin cậy 95% cho sự khác nhau giữa tỷ lệ nam giới và nữ mắc bệnh này.

BÀI 11 – KIỂM ĐỊNH CHO GIÁ TRỊ TRUNG BÌNH

+ Bài toán kiểm định giả thuyết
+ Kiểm định giả thuyết về một kỳ vọng và hiệu hai kỳ vọng

11.1(1.351) Một hãng sản suất bóng đèn, có tuổi thọ trung bình của bóng là xấp xỉ phân phối chuẩn với kỳ vọng 800 giờ và độ lệch chuẩn 40 giờ. Kiểm định giả thuyết $\mu=800$ giờ với đối thuyết $\mu\ne 800$ giờ nếu một mẫu ngẫu nhiên gồm 30 bóng có tuổi thọ trung bình là 778 giờ, mức ý nghĩa 0,04.

Giải:

+ Giả thuyết: $H_0:\mu=800$ Đối thuyết: $H_1:\mu\ne 800$.

+ Đây là bài toán kiểm định 1 trung bình đã biết phương sai

+ Chỉ tiêu kiểm định $z=\dfrac{\overline{x}-\mu}{\sigma/\sqrt{n}}=\dfrac{778-800}{40/\sqrt{30}}=-3,0125$

+ Với $\alpha=0.04\Rightarrow z_{\alpha/2}=2.05$. Miền bác bỏ $H_0=D:(-\infty;-2.05)\cup(2.05;+\infty)$

+ $z\in D$. Bác bỏ H_0 chấp nhận H_1

+ Kết luận Tuổi thọ trung bình của bóng đèn là khác 800h.

11.2 (2.351) Từ một mẫu ngẫu nhiên về khối lượng của 64 túi bắp rang bơ, tính được trung bình là 5,23 ounce và độ lệch chuẩn 0,24 ounce. Kiểm định giả thuyết rằng $\mu=5.5$ với đối thuyết $\mu<5.5$ theo mức ý nghĩa 0.05.

Giải:

+ Giả thuyết: $H_0:\mu=5.5$ Đối thuyết: $H_1:\mu<5.5$.

+ Đây là bài toán kiểm định 1 trung bình, 1 phía chưa biết phương sai, cỡ mẫu lớn.

+ Chỉ tiêu kiểm định $z=\dfrac{\overline{x}-\mu}{s/\sqrt{n}}=\dfrac{5.23-5.5}{0.24/\sqrt{64}}=-9$

+ Với $\alpha=0.05\Rightarrow z_{\alpha}=1.645$. Miền bác bỏ $H_0=D:(-\infty;-1.645)$

+ $z\in D$. Bác bỏ H_0 chấp nhận H_1

+ Kết luận khối lượng trung bình của túi bắp rang bơ là khác 5.5

11.2(6.351) Theo tạp chí của hiệp hội Tim mạch Mỹ, nghiên cứu cho thấy những người mắc bệnh cao huyết áp tham gia tập luyện, sẽ giảm đáng kể huyết áp của họ. Nếu từ mẫu ngẫu nhiên 225 người tham gia, tính được thời gian tập luyện trung bình là 8.5h một tuần và độ lệch chuẩn là 2.5h, thì ta có thể kết luận là thời gian tập luyện là trên 8h một tuần không? Với mức ý nghĩa 0.05

Giải:

+ Giả thuyết: $H_0:\mu=8$ Đối thuyết: $H_1:\mu>8$.

+ Đây là bài toán kiểm định 1 trung bình, 1 phía chưa biết phương sai, cỡ mẫu lớn.

+ Chỉ tiêu kiểm định $z=\dfrac{\overline{x}-\mu}{s/\sqrt{n}}=\dfrac{8.5-8}{2.5/\sqrt{225}}=3$

+ Với $\alpha=0.05\Rightarrow z_{\alpha}=1.645$. Miền bác bỏ $H_0=D:(1,645;+\infty)$

+ $z\in D$. Bác bỏ H_0 chấp nhận H_1

+ Kết luận thời gian tập luyện trung bình là trên 8h.

11.3(7.351) Kiểm định giả thuyết rằng thể tích của các hộp đựng loại dầu nhờn nào đó là 10 lít, nếu từ mẫu ngẫu nhiên gồm 10 hộp ta có các thể tích là:

10,2 9,7 10,1 10,3 10,1 9,8 9,9 10,4 10,3 9,8.

Sử dụng mức ý nghĩa 0,01 và giả thiết phân phối của thể tích là chuẩn.

Giải:

+ Giả thuyết: $H_0:\mu=10$ Đối thuyết: $H_1:\mu\neq 10$.

+ Đây là bài toán kiểm định 1 trung bình, 2 phía chưa biết phương sai, cỡ mẫu nhỏ.

+ Chỉ tiêu kiểm định $t=\dfrac{\overline{x}-\mu}{s/\sqrt{n}}=\dfrac{10.06-10}{0.2459/\sqrt{10}}=0.7716$

+ Với $\alpha=0.01\Rightarrow t_{\alpha/2}=3.25$. Miền bác bỏ $H_0=D:(-\infty;-3.25)\cup(3.25;+\infty)$

+ $z\notin D$. Chấp nhận H_0.

+ Kết luận thể tích trung bình là 10l.

(8.351) Cơ thể người cần khoảng 220mlg nattri 1 ngày. Khảo sát 20 loại suất ăn của 1 hãng nga ta thấy lượng nattri trung bình trong đó là 244mlg, với độ lệch chuẩn là 24.5. Với mức ý nghĩa 0.05 có thể nói lượng nattri trung bình trong các suất ăn của hang này lớn hơn 220 được không?

11.4 (11.351) Các thí nghiệm cho thấy, thời gian để học sinh phổ thông làm một bài kiểm tra là một biến ngẫu nhiên có phân phối chuẩn với kỳ vọng là 35 phút. Nghiên cứ trên một mẫu gồm 20 học sinh, người ta thấy thời gian trung bình để các em hoàn thành bài thi là 33,1 phút với độ lệch 4,3 phút. Với mức ý nghĩa 0,025 hãy kiểm định giả thuyết $\mu = 35$ phút với đối thuyết $\mu < 35$ phút.

Giải

dữ liệu mẫu $n = 20; \bar{x} = 33.1; s = 4.3$

Đây là bài toán kiểm định 1 phía, chưa biết phương sai, cỡ mẫu nhỏ

Ho: $\mu = 35$; H1: $\mu \neq 35$

Chỉ tiêu kiểm định $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ = -1,9761

với $\alpha = 0,025$, $\nu = 19$ ta có $t_{0,025,19} = 2,093$

Miền bác bỏ Ho là $D = (-\infty; -t_{0,025,19})$

chấp nhận H0

thời gian để học sinh phổ thông làm một bài kiểm tra là 35 phút

11.5 (12.351) Một mẫu ngẫu nhiên cỡ $n_1 = 25$, lấy từ phân phối chuẩn với $\sigma_1 = 5,2$ có trung bình mẫu $\bar{x}_1 = 81$. Một mẫu khác cỡ $n_2 = 36$, lấy từ phân phối chuẩn với $\sigma_2 = 3,4$ có trung bình mẫu $\bar{x}_2 = 76$. Kiểm định giả thuyết $\mu_1 = \mu_2$ với đối thuyết $\mu_1 < \mu_2$ và mức ý nghĩa 0.05.

(13.351) Một nhà sản xuất khẳng định sức căng trung bình của loại dây A hơn sức căng trung binhg của loại dây B ít nhất là 12kg. Để kiểm tra người ta lấy 50 mẫu dây mỗi loại được thử với cùng điều kiện. Kết quả cho thấy, loại dây A có sức căng trung bình là 86.7kg và độ lệch chuẩn là 6.28kg. Loại dậy B có sức căng trung bình là 77.8kg và độ lệch chuẩn 5.61kg. Hãy kiểm định khẳng định của nhà sản xuất với mức ý nghĩa 0.05.

11.6 (18) Một hãng sản xuất xe hơi muốn xác định xem, nên dùng loại lốp A hay B cho loại xe mới của họ. Họ thực hiện thí nghiệm với 12 chiếc lốp mỗi loại, và ghi lại số km đi được đến khi phải thay lốp. Kết quả như sau:

Loại A: $\bar{x}_1 = 37900$ km ; $s_1 = 5100$ km.

Loại B: $\bar{x}_2 = 39800$ km ; $s_1 = 5900$ km.

Kiểm định giả thuyết $\mu_1 = \mu_2$ với đối thuyết $\mu_1 > \mu_2$ và mức ý nghĩa 0.05. Giả sử các phân phối đều là phân phối chuẩn, với phương sai bằng nhau.

11.7 (22.354) Một mẫu gồm 32 phụ nữ đang có thai vào giai đoạn 3 tháng cuối của thai kỳ, có độ tuổi từ 15 đến 32, được chia làm hai nhóm: Có hút thuốc và Không hút thuốc. Người ta đo nồng độ axit huyết tương ascorbic (mg/ml) trong máu của họ, khi họ chưa ăn sáng, được số liệu sau:

Hút thuốc:	0,48	0,71	0,98	0,68	1,18	1,36	0,78	1,64				
Không hút:	0,97	0,72	1,00	0,81	0,62	1,32	1,24	0,99	0,90	0,74	1,24	0,88
	0,94	1,16	0,86	0,85	0,58	0,57	0,64	0,98	1,09	0,92	0,78	1,18

Hãy kiểm định xem có sự sai khác giữa nồng độ ascorbic trung bình của hai nhóm hay không, với mức ý nghĩa 0.05? (Giả sử các số liệu tuân theo phân phối chuẩn với phương sai khác nhau)

11.8(24.354) Năm mẫu quặng sắt, mỗi mẫu được chia thành hai phần, rồi lần lượt được xác định hàm lượng sắt bằng hai cách là dùng tia X và dùng phân tích hóa học, kết quả thu được là

	Số thứ tự mẫu				
Cách phân tích	**1**	**2**	**3**	**4**	**5**
Tia X	2,0	2,0	2,3	2,1	2,4
Phân tích hóa học	2,2	1,9	2,5	2,3	2,4

Giả sử các số liệu ở mỗi cách phân tích đều tuân theo phân phối chuẩn. Hãy kiểm định rằng hai phương pháp cho kết quả giống nhau, với mức ý nghĩa 0,05?

BÀI 12 – KIỂM ĐỊNH CHO TỶ LỆ

+ Kiểm định 1 tỷ lệ và kiểm định hiệu 2 tỷ lệ

12.1 Một chuyên gia marketing của công ty sản xuất mì ống tin rằng, 40% người thích mì ống hơn lasagna (một loại món ăn). Qua phỏng vấn 20 người, thì có 9 người thích mì ống hơn. Có thể kết luận gì về khẳng định của chuyên gia, với mức ý nghĩa 0,05?

12.2 Giả sử rằng trước đây có 40% người trưởng thành ủng hộ án tử hình. Có thể tin được hay không rằng tỷ lệ người ủng hộ án tử hình ngày nay đã tăng lên, nếu trong mẫu ngẫu nhiên gồm 15 người thì có 8 người đồng ý? Sử dụng mức ý nghĩa 0,05.

12.3 (5) Một công ty xăng dầu khẳng định 1/5 số nhà trong thành phố có máy sưởi sử dụng dầu. Có thể nghi ngờ khẳng định này không, nếu trong một mẫu ngẫu nhiên gồm 1000 ngôi nhà, thì có 136 ngôi nhà được sưởi bằng dầu? Dùng mức ý nghĩa 0,01.

12.4 (**6**) Tại một trường cao đẳng nào đó, người ta ước tính rằng có đúng 25% sinh viên tới trường bằng xe đạp. Khảo sát một mẫu gồm 90 sinh viên, thấy có 28 bạn tới trường bằng xe đạp. Kiểm định giả thuyết $p = 0,25$ với đối thuyết $p > 0,25$ và mức ý nghĩa 0.05.

12.5 (**10)**Trong một mẫu ngẫu nhiên gồm 200 phụ nữ trưởng thành sống ở thành thị, có 20 người mắc ung thư vú. Con số này là 10 trên 150 phụ nữ sống ở nông thôn được chọn ngẫu nhiên. Liệu có thể kết luận với mức ý nghĩa 0,06 rằng, tỷ lệ mắc bệnh ung thư vú ở phụ nữ thành thị là cao hơn phụ nữ nông thôn không?

Các bài toán ôn tập Kiểm định

12.6 Một nhà di truyền học quan tâm tới tỷ lệ nam và nữ trong dân số bị mắc chứng rối loạn máu. Trong mẫu ngẫu nhiên gồm 100 nam giới, có 31 người mắc chứng này; và trong 100 nữ giới có 24 người mắc. Có thể kết luận với mức ý nghĩa 0,01 rằng, tỷ lệ nam giới mắc chứng rối loạn máu cao hơn với tỷ lệ nữ giới mắc chứng này không?

12.7 Một nghiên cứu của Khoa Giáo dục thể chất, trường Đại học Virginia nhằm xác định xem sau 8 tuần luyện tập, lượng cholesterol của những người tham gia luyện tập có thực sự giảm không. Một nhóm 15 người tham gia luyện tập 2 lần một tuần. Một nhóm khác gồm 18 người với độ tuổi tương tự, không tham gia luyện tập. Sau 8 tuần, lượng cholesterol được ghi lại như sau:

Nhóm luyện tập: 129 131 154 172 115 126 175 191 122 238 159
156 176 175 126

Nhóm không luyện tập: 151 132 196 195 188 198 187 168 115
165 137 208 133 217 191 193 140 146

Ta có thể kết luận, với mức ý nghĩa 5% rằng, lượng cholesterol thực sự giảm sau khi thực hiện chương trình luyện tập không? Biết rằng các số liệu tuân theo phân phối xấp xỉ chuẩn với phương sai bằng nhau.

12.8 Một nghiên cứu được thực hiện bởi Trung tâm Thủy lợi và được phân tích bởi Trung tâm Thống kê, thuộc Đại học Virginia, nhằm so sánh hai thiết bị xử lý nước thải. Thiết bị A được đặt ở vùng dân cư có thu nhập trung bình dưới 22000$/năm. Thiết bị B được đặt ở vùng dân cư có thu nhập trung bình trên 60000$/năm. Lượng nước thải được xử lý bởi mỗi thiết bị (đơn vị: nghìn ga-lông/ ngày) được đo trong 10 ngày như sau:

Thiết bị A: 21 19 20 23 22 28 32 19 13 18

Thiết bị B: 20 39 24 33 30 28 30 22 33 24

Với mức ý nghĩa 5%, có thể kết luận rằng lượng nước thải trung bình được xử lý ở vùng có thu nhập cao là khác với vùng có thu nhập thấp không? Biết rằng các số liệu tuân theo phân phối xấp xỉ chuẩn với phương sai bằng nhau.

BÀI 13 – HỒI QUY TUYẾN TÍNH

+ Hồi quy và tương quan tuyến tính, ước lượng bình phương tối thiểu
+ Các tính chất của các ước lượng BPTT

13.1 Tại Viện Bách khoa và Đại học Tiểu bang Virginia đã tiến hành một cuộc nghiên cứu xác định liệu các số đo sức mạnh cánh tay tĩnh *X* có ảnh hưởng đến lực nâng động lực *Y* của một người hay không. Có 25 người đã tham gia thử nghiệm sức mạnh cánh tay của họ và sau đó được yêu cầu thực hiện thử nghiệm nâng cử tạ trong đó cử tạ được nâng qua đầu bằng lực nâng động lực. Số liệu như sau:

Đối tượng	***X***	***Y***	**Đối tượng**	***X***	***Y***
1	17,3	71,7	14	29,6	78,3
2	19,3	48,3	15	29,9	60,0
3	19,5	88,3	16	29,9	71,7
4	19,7	75,0	17	30,3	85,0
5	22,9	91,7	18	31,3	85,0
6	23,1	100,0	19	36,0	88,3
7	26,4	73,3	20	39,5	100,0
8	26,8	65,0	21	40,4	100,0
9	27,6	75,0	22	44,3	100,0
10	28,1	88,3	23	44,6	91,7
11	28,2	68,3	24	50,4	100,0
12	28,7	96,7	25	55,9	71,7
13	29,0	76,7			

a) Ước lượng α và β cho đường hồi quy tuyến tính $\mu_{Y|x} = \alpha + \beta x$.

b) Tìm một ước lượng điểm của $\mu_{Y|30}$.

c) Tính ước lượng không chệch s^2 của σ^2.

13.2 Trong một loại mẫu thử kim loại xác định, ứng suất chuẩn *X* trên một mẫu thử có liên quan về mặt chức năng đến sức bền cắt *Y*. Dưới đây là tập hợp số liệu thử nghiệm đã được mã hóa theo 2 biến số:

Ứng suất chuẩn *X*	**Sức bền cắt *Y***
26,8	26,5
25,4	27,3
28,9	24,2
23,6	27,1
27,7	23,6
23,9	25,9
24,7	26,3
28,1	22,5
26,9	21,7
27,4	21,4
22,6	25,8
25,6	24,9

a) Ước lượng α và β cho đường hồi quy tuyến tính $\mu_{Y|x} = \alpha + \beta x$.

b) Dự đoán sức bền cắt trung bình cho ứng suất chuẩn $x =$ 24,5 kg/cm^2.

c) Tìm ước lượng không chệch s^2 của σ^2.

13.3 Số liệu sau đây được thu thập để xác định mối quan hệ giữa áp suất và chỉ số thang độ tương ứng cho mục đích hiệu chỉnh.

Áp suất *X* (Pound/inch^2)	**Chỉ số thang độ *Y***
10	13
10	18
10	16
10	15
10	20
50	86
50	90
50	88
50	88
50	92

a) Tìm phương trình đường hồi quy quy tuyến tính thực nghiệm của *Y* theo *X*.
b) Mục đích hiệu chỉnh trong ứng dụng này là nhằm tính áp suất từ một chỉ số thang độ đã quan sát. Hãy tính áp suất cho chỉ số thang độ 54, sử dụng $\hat{x} = (54 - a)/b$.
c) Tính s^2.

13.4 Số lượng hợp chất hóa học *Y* hòa tan trong 100g nước tại các nhiệt độ biến thiên *X*, được ghi lại như sau:

***X* (^{o}C)**	***Y* (gram)**		
0	8	6	8
15	12	10	14
30	25	31	24
45	31	33	28
60	44	39	42
75	48	51	44

a) Tìm phương trình đường hồi quy tuyến tính thực nghiệm biểu diễn mối quan hệ của lượng chất hóa học hòa tan trong 100g nước theo nhiệt độ.
b) Dự đoán lượng hóa chất sẽ hòa tan trong 100g nước ở nhiệt độ 50^{o}C.
c) Tính s^2.

13.5 Một cuộc nghiên cứu về lượng mưa và lượng hạt ô nhiễm không khí thải ra đã cho các số liệu sau:

Lượng mưa hàng ngày *X* (0,01cm)	**Lượng hạt ô nhiễm thải ra *Y* (mcg/cum)**
4,3	126
4,5	121
5,9	116
5,6	118
6,1	114
5,2	118
3,8	132
2,1	141
7,5	108

a) Tìm phương trình đường hồi quy tuyến tính thực nghiệm để dự đoán trước lượng hạt ô nhiễm thoát ra từ lượng mưa hàng ngày.

b) Tính lượng hạt ô nhiễm thoát ra khi lượng mưa hàng ngày là $x = 4,8$ đơn vị.

c) Tính s^2.

BÀI 14 – BÀI TOÁN DỰ BÁO

+ Khoảng tin cậy cho các hệ số đường hồi quy
+ Bài toán dự báo, khoảng dự báo

14.1 Sử dụng các kết quả đã có trong các bài tập 13.1 đến 13.5 để tìm:

a) Khoảng tin cậy 99% cho α.

b) Khoảng tin cậy 99% cho β.

14.2 Sử dụng các kết quả đã có trong các bài tập 13.2 để tìm:

a) Khoảng tin cậy 95% cho độ bền cắt trung bình khi ứng suất chuẩn $x_0 = 24,5$.

b) Khoảng dự báo 95% cho độ bền cắt y_0 khi $x_0 = 24,5$.

14.3 Sử dụng các kết quả đã có trong các bài tập 13.4 để tìm:

a) Khoảng tin cậy 99% cho lượng hóa chất trung bình sẽ hòa tan trong 100g nước tại nhiệt độ 50^{o}C.

b) Khoảng dự báo 99% cho lượng hóa chất sẽ hòa tan trong 100 g nước tại nhiệt độ 50^{o}C.

14.4 Số liệu về độ tuổi (năm) và năng suất lao động (số sản phẩm/ngày) của một nhóm công nhân công ty may Việt Tiến được cho bởi bảng sau:

Tuổi *X*	20	22	25	30	35	40	42	45	48
Năng suất lao động *Y*	7	6	8	6	5	5	7	4	4

a) Tìm đường hồi quy tuyến tính thực nghiệm của *Y* theo *X*.

b) Tìm khoảng tin cậy cho năng suất lao động trung bình của công nhân 32 tuổi, với độ tin cậy 95%.

c) Tìm khoảng dự đoán cho năng suất lao động của công nhân 32 tuổi, với độ tin cậy như trên.

d) Tìm hệ số tương quan *r*.

14.5 Số liệu sau có được từ một cuộc nghiên cứu về mối liên hệ giữa trọng lượng và kích thước ngực của trẻ sơ sinh lúc sinh ra:

Trọng lượng (*kg*)	2,75	2,15	4,41	5,52	3,21	4,32	2,31	4,30	3,71
Kích thước ngực (*cm*)	29,5	26,3	32,2	36,5	27,2	27,7	28,3	30,3	28,7

a) Viết phương trình đường hồi quy tuyến tính thực nghiệm biểu diễn trọng lượng theo kích thước ngực của trẻ sơ sinh lúc sinh ra.

b) Tìm khoảng tin cậy 95% cho trọng lượng trung bình của trẻ sơ sinh khi kích thước ngực là 35*cm.*

c) Tìm khoảng dự đoán 95% cho trọng lượng trẻ sơ sinh khi kích thước ngực là 35*cm.*

d) Tìm hệ số tương quan *r*.