BÀI GIẢI
XÁC SUẤT THỐNG KÊ
(GV: Trần Ngọc Hội – 2009)

CHƯƠNG 1
NHỮNG ĐỊNH LÝ CƠ BẢN TRONG LÝ THUYẾT XÁC SUẤT

Bài 1.1: Có ba khẩu súng I, II và III bắn độc lập vào một mục tiêu. Mỗi khẩu bắn 1 viên. Xác suất bắn trúng mục tiêu cuả ba khẩu I, II và III lần lượt là 0,7; 0,8 và 0,5. Tính xác suất để
a) có 1 khẩu bắn trúng.
b) có 2 khẩu bắn trúng.
c) có 3 khẩu bắn trúng.
d) ít nhất 1 khẩu bắn trúng.
e) khẩu thứ 2 bắn trúng biết rằng có 2 khẩu trúng.

Lời giải

Tóm tắt:

Khẩu súng	I	II	III
Xác suất trúng	0,7	0,8	0,5

Gọi A_j (j = 1, 2, 3) là biến cố khẩu thứ j bắn trúng. Khi đó A_1, A_2, A_3 độc lập và giả thiết cho ta:

$$P(A_1) = 0,7; P(\bar{A}_1) = 0,3;$$
$$P(A_2) = 0,8; P(\bar{A}_2) = 0,2;$$
$$P(A_3) = 0,5; P(\bar{A}_3) = 0,5.$$

a) Gọi A là biến cố có 1 khẩu trúng. Ta có
$$A = A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3$$
Vì các biến cố $A_1\bar{A}_2\bar{A}_3, \bar{A}_1A_2\bar{A}_3, \bar{A}_1\bar{A}_2A_3$ xung khắc từng đôi, nên theo công thức Cộng xác suất ta có
$$P(A) = P(A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3)$$
$$= P(A_1\bar{A}_2\bar{A}_3) + P(\bar{A}_1A_2\bar{A}_3) + P(\bar{A}_1\bar{A}_2A_3)$$
Vì các biến cố A_1, A_2, A_3 độc lập nên theo công thức Nhân xác suất ta có

$$P(A_1\bar{A}_2\bar{A}_3) = P(A_1)P(\bar{A}_2)P(\bar{A}_3) = 0,7.0,2.0,5 = 0,07;$$
$$P(\bar{A}_1A_2\bar{A}_3) = P(\bar{A}_1)P(A_2)P(\bar{A}_3) = 0,3.0,8.0,5 = 0,12;$$
$$P(\bar{A}_1\bar{A}_2A_3) = P(\bar{A}_1)P(\bar{A}_2)P(A_3) = 0,3.0,2.0,5 = 0,03.$$
Suy ra P(A) = 0,22.

b) Gọi B là biến cố có 2 khẩu trúng. Ta có
$$B = A_1A_2\bar{A}_3 + \bar{A}_1A_2A_3 + A_1\bar{A}_2A_3$$
Tính toán tương tự câu a) ta được P(B) = 0,47.

c) Gọi C là biến cố có 3 khẩu trúng. Ta có
$$C = A_1A_2A_3.$$
Tính toán tương tự câu a) ta được P(C) = 0,28.

d) Gọi D là biến cố có ít nhất 1 khẩu trúng. Ta có
$$D = A + B + C.$$
Chú ý rằng do A, B, C xung khắc từng đôi, nên theo công thức Cộng xác suất ta có:
$$P(D) = P(A) + P(B) + P(C) = 0,22 + 0,47 + 0,28 = 0,97.$$

e) Gỉa sử có 2 khẩu trúng. Khi đó biến cố B đã xảy ra. Do đó xác suất để khẩu thứ 2 trúng trong trường hợp này chính là xác suất có điều kiện $P(A_2/B)$.

Theo công thức Nhân xác suất ta có:
$$P(A_2B) = P(B)P(A_2/B)$$
Suy ra
$$P(A_2/B) = \frac{P(A_2B)}{P(B)}.$$

Mà $A_2B = A_1A_2\bar{A}_3 + \bar{A}_1A_2A_3$ nên lý luận tương tự như trên ta được
$$P(A_2B)=0,4$$
Suy ra $P(A_2/B)$ =0,851.

Bài 1.2: Có hai hộp I và II mỗi hộp chứa 10 bi, trong đó hộp I gồm 9 bi đỏ, 1 bi trắng; hộp II gồm 6 bi đỏ, 4 bi trắng. Lấy ngẫu nhiên từ mỗi hộp 2 bi.
a) Tính xác suất để được 4 bi đỏ.
b) Tính xác suất để được 2 bi đỏ và 2 bi trắng.
c) Tính xác suất để được 3 bi đỏ và 1 bi trắng.
d) Giả sử đã lấy được 3 bi đỏ và 1 bi trắng. Hãy tìm xác suất để bi trắng có được của hộp I.

Lời giải

Gọi A_i, B_i (i = 0, 1, 2) lần lượt là các biến cố có i bi đỏ và (2 - i) bi trắng có trong 2 bi được chọn ra từ hộp I, hộp II.

Khi đó

- A_0, A_1, A_2 xung khắc từng đôi và ta có:

$$P(A_0) = 0;$$
$$P(A_1) = \frac{C_9^1 C_1^1}{C_{10}^2} = \frac{9}{45};$$
$$P(A_2) = \frac{C_9^2 C_1^0}{C_{10}^2} = \frac{36}{45}.$$

- B_0, B_1, B_2 xung khắc từng đôi và ta có:

$$P(B_0) = \frac{C_6^0 C_4^2}{C_{10}^2} = \frac{6}{45};$$
$$P(B_1) = \frac{C_6^1 C_4^1}{C_{10}^2} = \frac{24}{45};$$
$$P(B_2) = \frac{C_6^2 C_4^0}{C_{10}^2} = \frac{15}{45}.$$

- A_i và B_j độc lập.
- Tổng số bi đỏ có trong 4 bi chọn ra phụ thuộc vào các biến cố A_i và B_j theo bảng sau:

	B_0	B_1	B_2
A_0	0	1	2
A_1	1	2	3
A_2	2	3	4

a) Gọi A là biến cố chọn được 4 bi đỏ. Ta có:

$$A = A_2 B_2.$$

Từ đây, do tính độc lập, Công thức nhân xác suất thứ nhất cho ta:

$$P(A) = P(A_2)P(B_2) = \frac{36}{45}.\frac{15}{45} = 0,2667.$$

b) Gọi B là biến cố chọn được 2 bi đỏ và 2 bi trắng. Ta có:

$$B = A_0B_2 + A_1B_1 + A_2B_0$$

Do tính xung khắc từng đôi của các biến cố A_0B_2, A_1B_1, A_2B_0, công thức Cộng xác suất cho ta:

$$P(B) = P(A_0B_2 + A_1B_1 + A_2B_0) = P(A_0B_2) + P(A_1B_1) + P(A_2B_0)$$

Từ đây, do tính độc lập, Công thức nhân xác suất thứ nhất cho ta:

$$P(B) = P(A_0)P(B_2) + P(A_1)P(B_1) + P(A_2)P(B_0) = 0,2133.$$

c) Gọi C là biến cố chọn được 3 bi đỏ và 1 bi trắng. Ta có:

$$C = A_1B_2 + A_2B_1.$$

Lý luận tương tự như trên ta được

$$P(C) = P(A_1)P(B_2) + P(A_2)P(B_1) = 0,4933.$$

d) Giả sử đã chọn được 3 bi đỏ và 1 bi trắng. Khi đó biến cố C đã xảy ra. Do đó xác suất để bi trắng có được thuộc hộp I trong trường hợp này chính là xác suất có điều kiện $P(A_1/C)$. Theo Công thức nhân xác suất, ta có

$$P(A_1C) = P(C)P(A_1/C).$$

Suy ra

$$P(A_1/C) = \frac{P(A_1C)}{P(C)}.$$

Mà $A_1C = A_1B_2$ nên

$$P(A_1C) = P(A_1B_2) = P(A_1)P(B_2) = \frac{9}{45}.\frac{15}{45} = 0,0667.$$

Do đó xác suất cần tìm là: $P(A_1/C) = 0,1352$.

Bài 1.3: Một lô hàng chứa 10 sản phẩm gồm 6 sản phẩm tốt và 4 sản phẩm xấu. Khách hàng kiểm tra bằng cách lấy ra từng sản phẩm cho đến khi nào được 3 sản phẩm tốt thì dừng lại.

a) Tính xác suất để khách hàng dừng lại ở lần kiểm tra thứ 3.
b) Tính xác suất để khách hàng dừng lại ở lần kiểm tra thứ 4.
b) Giả sử khách hàng đã dừng lại ở lần kiểm tra thứ 4. Tính xác suất để ở lần kiểm tra thứ 3 khách hàng gặp sản phẩm xấu.

Lời giải

Gọi T_i, X_i lần lượt là các biến cố chọn được sản phẩm tốt, xấu ở lần kiểm tra thứ i.

a) Gọi A là biến cố khách hàng dừng lại ở lần kiểm tra thứ 3. Ta có:

$$A = T_1T_2T_3.$$

Suy ra $P(A) = P(T_1T_2T_3) = P(T_1)\,P(T_2/T_1)\,P(T_3/\,T_1T_2)$
$= (6/10)(5/9)(4/8) = 0,1667.$

b) Gọi B là biến cố khách hàng dừng lại ở lần kiểm tra thứ 4. Ta có:

$$B = X_1T_2T_3T_4 + T_1X_2T_3T_4 + T_1T_2X_3T_4\,.$$

Suy ra

$P(B) = P(X_1T_2T_3T_4) + P(T_1X_2T_3T_4) + P(T_1T_2X_3T_4)$
$= P(X_1)\,P(T_2/X_1)\,P(T_3/X_1T_2)\,P(T_4/X_1T_2T_3)$
$+ P(T_1)\,P(X_2/T_1)\,P(T_3/T_1X_2)\,P(T_4/T_1X_2T_3)$
$+ P(T_1)\,P(T_2/T_1)\,P(X_3/\,T_1T_2)\,P(T_4/\,T_1T_2\,X_3)$
$= (4/10)(6/9)(5/8)(4/7) + (6/10)(4/9)(5/8)(4/7)+(6/10)(5/9)(4/8)(4/7)$
$= 3(4/10)(6/9)(5/8)(4/7) = 0,2857.$

c) Giả sử khách hàng đã dừng lại ở lần kiểm tra thứ 4. Khi đó biến cố B đã xảy ra. Do đó xác suất để ở lần kiểm tra thứ 3 khách hàng gặp sản phẩm xấu trong trường hợp này chính là xác suất có điều kiện $P(X_3/B)$.

Theo Công thức nhân xác suất , ta có

$$P(X_3B) = P(B)P(X_3/B)\ .$$

Suy ra

$$P(X_3/B) = \frac{P(X_3B)}{P(B)}\ .$$

Mà $X_3B = T_1T_2X_3T_4$ nên

$P(X_3B) = P(T_1T_2X_3T_4) = P(T_1)\,P(T_2/T_1)\,P(X_3/\,T_1T_2)\,P(T_4/\,T_1T_2\,X_3)$
$= (6/10)(5/9)(4/8)(4/7) = 0,0952.$

Suy ra $P(X_3/B) = 0,3333.$

Bài 1.4: Một hộp bi gồm 5 bi đỏ, 4 bi trắng và 3 bi xanh có cùng cỡ. Từ hộp ta rút ngẫu nhiên không hòan lại từng bi một cho đến khi được bi đỏ thì dừng lại. Tính xác suất để

a) được 2 bi trắng, 1 bi xanh và 1 bi đỏ.

b) không có bi trắng nào được rút ra.

Lời giải

Gọi D_i, T_i, X_i lần lượt là các biến cố chọn được bi đỏ, bi trắng, bi xanh ở lần rút thứ i.

a) Gọi A là biến cố rút được 2 bi trắng, 1 bi xanh và 1 bi đỏ. Ta có:

$$A \text{ xảy ra} \Leftrightarrow \text{Rút được} \left[\begin{array}{l} T - T - X - D \\ T - X - T - D \\ X - T - T - D \end{array}\right.$$

Suy ra

$$A = T_1T_2X_3D_4 + T_1X_2T_3D_4 + X_1T_2T_3D_4$$

Từ đây, do tính xung khắc từng đôi của các biến cố thành phần, ta có:

$$P(A) = P(T_1T_2X_3D_4)+ P(T_1X_2T_3D_4) + P(X_1T_2T_3D_4)$$

Theo Công thức Nhân xác suất, ta có

$P(T_1T_2X_3D_4) = P(T_1)P(T_2/T_1)P(X_3/T_1T_2)P(D_4/T_1T_2X_3)$
$= (4/12)(3/11)(3/10)(5/9) = 1/66;$

$P(T_1X_2T_3D_4) = P(T_1)P(X_2/T_1)P(T_3/T_1X_2)P(D_4/T_1X_2T_3)$
$= (4/12)(3/11)(3/10)(5/9) = 1/66;$

$P(X_1T_2T_3D_4) = P(X_1)P(T_2/X_1)P(T_3/X_1T_2)P(D_4/X_1T_2T_3)$
$= (3/12)(4/11)(3/10)(5/9) = 1/66.$

Suy ra $P(A) = 3/66 = 1/22 = 0,0455.$

b) Gọi B là biến cố không có bi trắng nào được rút ra. Ta có:

$$B \text{ xảy ra} \Leftrightarrow \text{Rút được} \left[\begin{array}{l} D \\ X - D \\ X - X - D \\ X - X - X - D \end{array}\right.$$

Suy ra

$$B = D_1 + X_1D_2 + X_1X_2D_3+ X_1X_2X_3\,D_4$$

Từ đây, do tính xung khắc từng đôi của các biến cố thành phần, ta có:

$$P(B) = P(D_1)+ P(X_1D_2) + P(X_1X_2D_3) + P(X_1X_2X_3\,D_4)$$

Theo Công thức Nhân xác suất, ta có

$P(B) = P(D_1) + P(X_1)P(D_2/X_1) + P(X_1)P(X_2/X_1)P(D_3/X_1X_2)$

$+ P(X_1)P(X_2/X_1)P(X_3/X_1X_2)P(D_4/X_1X_2X_3)$

$= 5/12+ (3/12)(5/11) + (3/12)(2/11)(5/10) + (3/12)(2/11)(1/10)(5/9)$

$= 5/9$

Bài 1.5: Sản phẩm X bán ra ở thị trường do một nhà máy gồm ba phân xưởng I, II và III sản xuất, trong đó phân xưởng I chiếm 30%; phân xưởng II chiếm 45% và phân xưởng III chiếm 25%. Tỉ lệ sản phẩm loại A do ba phân xưởng I, II và III sản xuất lần lượt là 70%, 50% và 90%.

a) Tính tỉ lệ sản phẩm lọai A nói chung do nhà máy sản xuất.

b) Chọn mua ngẫu nhiên một sản phẩm X ở thị trường. Giả sử đã mua được sản phẩm loại A. Theo bạn, sản phẩm ấy có khả năng do phân xưởng nào sản xuất ra nhiều nhất?

c) Chọn mua ngẫu nhiên 121 sản phẩm X (trong rất nhiều sản phẩm X) ở thị trường.

1) Tính xác suất để có 80 sản phẩm loại A.

2) Tính xác suất để có từ 80 đến 85 sản phẩm loại A.

Lời giải

Tóm tắt:

Phân xưởng	I	II	III
Tỉ lệ sản lượng	30%	45%	25%
Tỉ lệ loại A	70%	50%	90%

a) Để tính tỉ lệ sản phẩm loại A nói chung do nhà máy sản xuất ta chọn mua ngẫu nhiên một sản phẩm ở thị trường. Khi đó tỉ lệ sản phẩm loại A chính là xác suất để sản phẩm đó thuộc loại A.

Gọi B là biến cố sản phẩm chọn mua thuộc loại A.

A_1, A_2, A_3 lần lượt là các biến cố sản phẩm do phân xưởng I, II, III sản xuất. Khi đó A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi và

$P(A_1) = 30\% = 0{,}3$; $P(A_2) = 45\% = 0{,}45$; $P(A_3) = 25\% = 0{,}25$.

Theo công thức xác suất đầy đủ, ta có:

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + P(A_3)P(B/A_3)$$

Theo giả thiết,

$P(B/A_1) = 70\% = 0{,}7$; $P(B/A_2) = 50\% = 0{,}5$; $P(B/A_3) = 90\% = 0{,}9$.

Suy ra $P(B) = 0{,}66 = 66\%$. Vậy tỉ lệ sản phẩm loại A nói chung do nhà máy sản xuất là 66%.

b) Chọn mua ngẫu nhiên một sản phẩm X ở thị trường. Giả sử đã mua được sản phẩm loại A. Theo bạn, sản phẩm ấy có khả năng do phân xưởng nào sản xuất ra nhiều nhất?

Giả sử đã mua được sản phẩm loại A. Khi đó biến cố B đã xảy ra. Do đó, để biết sản phẩm loại A đó có khả năng do phân xưởng nào sản xuất ra nhiều nhất ta cần so sánh các xác suất có điều kiện $P(A_1/B)$, $P(A_2/B)$ và $P(A_3/B)$. Nếu $P(A_i/B)$ là lớn nhất thì sản phẩm ấy có khả năng do phân xưởng thứ i sản xuất ra là nhiều nhất. Theo công thức Bayes ta có:

$$P(A_1/B) = \frac{P(A_1)P(B/A_1)}{P(B)} = \frac{0{,}3.0{,}7}{0{,}66} = \frac{21}{66};$$

$$P(A_2/B) = \frac{P(A_2)P(B/A_2)}{P(B)} = \frac{0{,}45.0{,}5}{0{,}66} = \frac{22{,}5}{66};$$

$$P(A_3/B) = \frac{P(A_3)P(B/A_3)}{P(B)} = \frac{0{,}25.0{,}9}{0{,}66} = \frac{22{,}5}{66}.$$

Vì $P(A_2/B) = P(A_3/B) > P(A_1/B)$ nên sản phẩm loại A ấy có khả năng do phân xưởng II hoặc III sản xuất ra là nhiều nhất.

c) Chọn mua ngẫu nhiên 121 sản phẩm X (trong rất nhiều sản phẩm X) ở thị trường.

1) Tính xác suất để có 80 sản phẩm loại A.

2) Tính xác suất để có từ 80 đến 85 sản phẩm loại A.

Áp dụng công thức Bernoulli với $n = 121$, $p = 0{,}66$, ta có:

1) Xác suất để có 80 sản phẩm loại A là

$$P_{121}(80) = C_{121}^{80}p^{80}q^{41} = C_{121}^{80}(0{,}66)^{80}(0{,}34)^{41} = 0{,}076.$$

2) Xác suất để có từ 80 đến 85 sản phẩm loại A là

$$\sum_{k=80}^{85} P_{121}(k) = \sum_{k=80}^{85} C_{121}^{k}p^{k}q^{121-k} = \sum_{k=80}^{85} C_{121}^{k}(0{,}66)^{k}(0{,}34)^{121-k} = 0{,}3925.$$

Bài 1.6: Có ba cửa hàng I, II và III cùng kinh doanh sản phẩm Y. Tỉ lệ sản phẩm loại A trong ba cửa hàng I, II và III lần lượt là 70%, 75% và 50%. Một khách hàng chọn nhẫu nhiên một cửa hàng và từ đó mua một sản phẩm

a) Tính xác suất để khách hàng mua được sản phẩm loại A.
b) Giả sử đã mua được sản phẩm loại A. Theo bạn, khả năng người khách hàng ấy đã chọn cửa hàng nào là nhiều nhất?

Lời giải

Tóm tắt:

Cửa hàng	I	II	III
Tỉ lệ loại A	70%	75%	50%

Chọn nhẫu nhiên một cửa hàng và từ đó mua một sản phẩm.

a) Tính xác suất để khách hàng mua được sản phẩm loại A.

Gọi B là biến cố sản phẩm chọn mua thuộc loại A.
A_1, A_2, A_3 lần lượt là các biến cố chọn cửa hàng I, II, III. Khi đó A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi và

$$P(A_1) = P(A_2) = P(A_3) = 1/3.$$

Theo công thức xác suất đầy đủ, ta có:

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/ A_2)+ P(A_3)P(B/A_3)$$

Theo giả thiết,

$$P(B/A_1) = 70\% = 0{,}7;$$
$$P(B/A_2) = 75\% = 0{,}75;$$
$$P(B/A_3 = 50\% = 0{,}5.$$

Suy ra P(B) = 0,65 = 65%. Vậy xác suất để khách hàng mua được sản phẩm loại A là 65%.

b) Giả sử đã mua được sản phẩm loại A. Theo bạn, khả năng người khách hàng ấy đã chọn cửa hàng nào là nhiều nhất?

Giả sử đã mua được sản phẩm loại A. Khi đó biến cố B đã xảy ra. Do đó, để biết sản phẩm loại A đó có khả năng khách hàng ấy đã chọn cửa hàng nào là nhiều nhất ta cần so sánh các xác suất có điều kiện $P(A_1/B)$,

$P(A_2/B)$ và $P(A_3/B)$. Nếu $P(A_i/B)$ là lớn nhất thì cửa hàng thứ i có nhiều khả năng được chọn nhất.
Theo công thức Bayes ta có:

$$P(A_1/B) = \frac{P(A_1)P(B/A_1)}{P(B)} = \frac{(1/3).0,7}{0,65} = \frac{70}{195};$$
$$P(A_2/B) = \frac{P(A_2)P(B/A_2)}{P(B)} = \frac{(1/3).0,75}{0,65} = \frac{75}{195};$$
$$P(A_3/B) = \frac{P(A_3)P(B/A_3)}{P(B)} = \frac{(1/3).0,5}{0,65} = \frac{50}{195}.$$

Vì $P(A_2/B) > P(A_1/B) > P(A_3/B)$ nên cửa hàng II có nhiều khả năng được chọn nhất.

Bài 1.7: Có hai hộp I và II mỗi hộp chứa 12 bi, trong đó hộp I gồm 8 bi đỏ, 4 bi trắng; hộp II gồm 5 bi đỏ, 7 bi trắng. Lấy ngẫu nhiên từ hộp I ba bi rồi bỏ sang hộp II; sau đó lấy ngẫu nhiên từ hộp II bốn bi.

a) Tính xác suất để lấy được ba bi đỏ và một bi trắng từ hộp II.
b) Giả sử đã lấy được ba bi đỏ và một bi trắng từ hộp II. Tìm xác suất để trong ba bi lấy được từ hộp I có hai bi đỏ và một bi trắng.

Lời giải

Gọi A là biến cố chọn được 3 bi đỏ và 1 bi trắng từ hộp II.
A_i (i = 0, 1, 2, 3) là biến cố có i bi đỏ và (3-i) bi trắng có trong 3 bi chọn ra từ hộp I. Khi đó A_0, A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi và ta có:

$$P(A_0) = \frac{C_8^0 C_4^3}{C_{12}^3} = \frac{4}{220};$$
$$P(A_1) = \frac{C_8^1 C_4^2}{C_{12}^3} = \frac{48}{220};$$
$$P(A_2) = \frac{C_8^2 C_4^1}{C_{12}^3} = \frac{112}{220};$$
$$P(A_3) = \frac{C_8^3 C_4^0}{C_{12}^3} = \frac{56}{220}.$$

a) Tính xác suất để lấy được 3 bi đỏ và 1 bi trắng từ hộp II.

Theo công thức xác suất đầy đủ, ta có:

$$P(A)=P(A_0)P(A/A_0)+P(A_1)P(A/A_1)+P(A_2)P(A/A_2)+P(A_3)P(A/A_3)$$

Theo công thức tính xác suất lựa chọn, ta có

$$P(A/A_0)=\frac{C_5^3C_{10}^1}{C_{15}^4}=\frac{100}{1365};$$

$$P(A/A_1)=\frac{C_6^3C_9^1}{C_{15}^4}=\frac{180}{1365};$$

$$P(A/A_2)=\frac{C_7^3C_8^1}{C_{15}^4}=\frac{280}{1365};$$

$$P(A/A_3)=\frac{C_8^3C_7^1}{C_{15}^4}=\frac{392}{1365}.$$

Suy ra xác suất cần tìm là $P(A) = 0,2076$.

b) Giả sử đã lấy được 3 bi đỏ và 1 bi trắng từ hộp II. Tìm xác suất để trong 3 bi lấy được từ hộp I có 2 bi đỏ và 1 bi trắng.

Giả sử đã lấy được 3 bi đỏ và 1 bi trắng từ hộp II. Khi đó biến cố A đã xảy ra. Do dó xác suất để trong 3 bi lấy được từ hộp I có 2 bi đỏ và 1 bi trắng trong trường hợp này chính là xác suất có điều kiện $P(A_2/A)$. Áp dụng công thức Bayes, ta có:

$$P(A_2/A) = \frac{P(A_2)P(A/A_2)}{P(A)}=\frac{\frac{112}{220}\cdot\frac{280}{1365}}{0,2076}=0,5030.$$

Vậy xác suất cần tìm là $P(A_2/A) = 0,5030$.

Bài 1.8: Có ba hộp mỗi hộp đựng 5 viên bi trong đó hộp thứ nhất có 1 bi trắng, 4 bi đen; hộp thứ hai có 2 bi trắng, 3 bi đen; hộp thứ ba có 3 bi trắng, 2 bi đen.

a) Lấy ngẫu nhiên từ mỗi hộp một bi.
   1) Tính xác suất để được cả 3 bi trắng.
   2) Tính xác suất được 2 bi đen, 1 bi trắng.
   3) Giả sử trong 3 viên lấy ra có đúng 1 bi trắng.Tính xác suất để bi trắng đó là của hộp thứ nhất.

b) Chọn ngẫu nhiên một hộp rồi từ hộp đó lấy ngẫu nhiên ra 3 bi. Tính xác suất được cả 3 bi đen.

Lời giải

a) Gọi A_j (j = 1, 2, 3) là biến cố lấy được bi trắng từ hộp thứ j. Khi đó A_1, A_2, A_3 độc lập và

$$P(A_1)=\frac{1}{5};\ P(\bar{A}_1)=\frac{4}{5};$$

$$P(A_2)=\frac{2}{5};P(\bar{A}_2)=\frac{3}{5};$$

$$P(A_3)=\frac{3}{5};P(\bar{A}_3)=\frac{2}{5}.$$

1) Gọi A là biến cố lấy được cả 3 bi trắng. Ta có

$$A = A_1A_2A_3.$$

Suy ra $P(A) = P(A_1)\ P(A_2)\ P(A_3) = 0,048$.

2) Gọi B là biến cố lấy 2 bi đen, 1 bi trắng. Ta có

$$B = A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3$$

Suy ra $P(B) =0,464$.

3) Giả sử trong 3 viên lấy ra có đúng 1 bi trắng. Khi đó biến cố B đã xảy ra. Do đó xác suất để bi trắng đó là của hộp thứ nhất trong trường hợp này chính là xác suất có điều kiện $P(A_1/B)$. Theo công thức Nhân xác suất ta có:

$$P(A_1B) = P(B)P(A_1/B)$$

Suy ra

$$P(A_1/B) = \frac{P(A_1B)}{P(B)}.$$

Mà $A_1B = A_1\bar{A}_2\bar{A}_3$ nên lý luận tương tự như trên ta được $P(A_1B) = 0,048$.

Suy ra

$$P(A_1/B) =0,1034 .$$

b) Chọn ngẫu nhiên một hộp rồi từ hộp đó lấy ngẫu nhiên ra 3 bi. Tính xác suất được cả 3 bi đen.

Gọi A là biến cố lấy được cả 3 bi đen.

A_1, A_2, A_3 lần lượt là các biến cố chọn được hộp I, II, III. Khi đó A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi và

$$P(A_1) = P(A_2) = P(A_3) = 1/3.$$

Theo công thức xác suất đầy đủ, ta có:

$$P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/ A_2)+ P(A_3)P(A/A_3)$$

Theo công thức xác suất lựa chọn, ta có:

$$P(A/A_1) = \frac{C_1^0 C_4^3}{C_5^3} = \frac{4}{10}; P(A/A_2) = \frac{C_2^0 C_3^3}{C_5^3} = \frac{1}{10}; P(A/A_3) = 0.$$

Suy ra P(A) = 0,1667.

Bài 1.9: Có 20 hộp sản phẩm cùng loại, mỗi hộp chứa rất nhiều sản phẩm, trong đó có 10 hộp của xí nghiệp I, 6 hộp của xí nghiệp II và 4 hộp của xí nghiệp III. Tỉ lệ sản phẩm tốt của các xí nghiệp lần lượt là 50%, 65% và 75%. Lấy ngẫu nhiên ra một hộp và chọn ngẫu nhiên ra 3 sản phẩm từ hộp đó.

a) Tính xác suất để trong 3 sản phẩm chọn ra có đúng 2 sản phẩm tốt.

b) Giả sử trong 3 sản phẩm chọn ra có đúng 2 sản phẩm tốt. Tính xác suất để 2 sản phẩm tốt đó của xí nghiệp I.

Lời giải

Gọi A là biến cố trong 3 sản phẩm chọn ra có đúng 2 sản phẩm tốt.

A_j (j = 1, 2, 3) là biến cố chọn được hộp của xí nghiệp thứ j.

Khi đó A_1, A_2, A_3 là một đầy đủ, xung khắc từng đôi và ta có:

$$P(A_1) = \frac{C_{10}^1}{C_{20}^1} = \frac{10}{20};$$

$$P(A_2) = \frac{C_6^1}{C_{20}^1} = \frac{6}{20};$$

$$P(A_3) = \frac{C_4^1}{C_{20}^1} = \frac{4}{20}.$$

Mặt khác, từ giả thiết, theo công thức Bernoulli, ta có

$$P(A/A_1) = C_3^2(0,5)^2(1-0,5) = 0,375$$

$$P(A/A_2) = C_3^2(0,65)^2(1-0,65) = 0,443625$$

$$P(A/A_3) = C_3^2(0,75)^2(1-0,25) = 0,421875$$

Theo công thức xác suất đầy đủ, ta có

$P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3)$

$= (10/20).0,375 + (6/20).\ 0,443625 + (4/20).\ 0,421875 = 0,4050.$

b) Giả sử trong 3 sản phẩm chọn ra có đúng 2 sản phẩm tốt. Khi đó, biến cố A đã xảy ra. Do đó, xác suất để 2 sản phẩm tốt đó của xí nghiệp I chính là xác suất có điều kiện $P(A_1/A)$.

Áp dụng Công thức Bayes và sử dụng kết quả vừa tìm được ở câu a) ta có

$$P(A_1/A) = \frac{P(A_1)P(A/A_1)}{P(A)} = \frac{(10/20).0,375}{0,4050} = 0,4630.$$

Bài 1.10: Có 10 sinh viên đi thi, trong đó có 3 thuộc loại giỏi, 4 khá và 3 trung bình. Trong số 20 câu hỏi thi qui định thì sinh viên loại giỏi trả lời được tất cả, sinh viên khá trả lời được 16 câu còn sinh viên trung bình được 10 câu. Gọi ngẫu nhiên một sinh viên và phát một phiếu thi gồm 4 câu hỏi thì anh ta trả lời được cả 4 câu hỏi. Tính xác suất để sinh viên đó thuộc loại khá.

Lời giải

Tóm tắt:

Xếp loại sinh viên	Giỏi	Khá	Trung bình
Số lượng	3	4	3
Số câu trả lời được/20	20	16	10

Gọi A là biến cố sinh viên trả lời được cả 3 câu hỏi.

A_1, A_2, A_3 lần lượt là các biến cố sinh viên thuộc loại Giỏi, Khá; Trung bình.

Yêu cầu của bài toán là tính xác suất có điều kiện $P(A_2/A)$.

Các biến cố A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi, và ta có:

$P(A_1) = 3/10$; $P(A_2) = 4/10$; $P(A_3) = 3/10$.

Theo công thức Bayes, ta có

$$P(A_2/A) = \frac{P(A_2)P(A/A_2)}{P(A)}.$$

Mặt khác, theo công thức xác suất đầy đủ, ta có

$P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3)$.

Theo công thức tính xác suất lựa chọn, ta có:

$$P(A/A_1) = \frac{C_{20}^4}{C_{20}^4} = 1;$$

$$P(A/A_2) = \frac{C_{16}^4 C_4^0}{C_{20}^4} = \frac{1820}{4845};$$

$$P(A/A_3) = \frac{C_{10}^4 C_{10}^0}{C_{20}^4} = \frac{210}{4845}.$$

Suy ra $P(A_2/A) = 0{,}3243$.

Bài 1.11: Có hai hộp I và II, trong đó hộp I chứa 10 bi trắng và 8 bi đen; hộp II chứa 8 bi trắng và 6 bi đen. Từ mỗi hộp rút ngẫu nhiên 2 bi bỏ đi, sau đó bỏ tất cả các bi còn lại của hai hộp vào hộp III (rỗng). Lấy ngẫu nhiên 2 bi từ hộp III. Tính xác suất để trong 2 bi lấy từ hộp III có 1 trắng, 1 đen.

Lời giải

Gọi A là biến cố bi lấy được 1 trắng, 1 đen.

A_j (j = 0, 1, 2, 3, 4) là biến cố có j bi trắng và (4-j) bi đen có trong 4 bi bỏ đi (từ cả hai hộp I và II). Khi đó A_0, A_1, A_2, A_3, A_4 là một hệ đầy đủ, xung khắc từng đôi.

Theo công thức xác suất đầy đủ, ta có

$$P(A) = P(A_0)P(A/A_0) + P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3) + P(A_4)P(A/A_4).$$

trong đó

$P(A/A_0) = \dfrac{C_{18}^1 C_{10}^1}{C_{28}^2} = \dfrac{10}{21}$ (Vì khi A_0 đã xảy ra thì trong hộp III có 28 bi gồm 18 trắng , 10 đen).

Tương tự,

$$P(A/A_1) = \frac{C_{17}^1 C_{11}^1}{C_{28}^2} = \frac{187}{378}; P(A/A_2) = \frac{C_{16}^1 C_{12}^1}{C_{28}^2} = \frac{32}{63};$$

$$P(A/A_3) = \frac{C_{15}^1 C_{13}^1}{C_{28}^2} = \frac{65}{126}; P(A/A_4) = \frac{C_{14}^1 C_{14}^1}{C_{28}^2} = \frac{14}{27}.$$

Bây giờ ta tính $P(A_0)$; $P(A_1)$; $P(A_2)$; $P(A_3)$; $P(A_4)$.

Gọi B_i , C_i (i = 0, 1, 2) lần lượt là các biến cố có i bi trắng và (2 - i) bi đen có trong 2 bi được chọn ra từ hộp I, hộp II. Khi đó

- B_0, B_1, B_2 xung khắc và ta có:

$$P(B_0) = \frac{C_{10}^0 C_8^2}{C_{18}^2} = \frac{28}{153}; P(B_1) = \frac{C_{10}^1 C_8^1}{C_{18}^2} = \frac{80}{153}; P(B_2) = \frac{C_{10}^2 C_8^0}{C_{18}^2} = \frac{5}{17}.$$

- C_0, C_1, C_2 xung khắc và ta có:

$$P(C_0) = \frac{C_8^0 C_6^2}{C_{14}^2} = \frac{15}{91}; P(C_1) = \frac{C_8^1 C_6^1}{C_{14}^2} = \frac{48}{91}; P(C_2) = \frac{C_8^2 C_6^0}{C_{14}^2} = \frac{28}{91}.$$

- B_i và C_j độc lập.

- Tổng số bi trắng có trong 4 bi chọn ra phụ thuộc vào các biến cố B_i và C_j theo bảng sau:

	C_0	C_1	C_2
B_0	0	1	2
B_1	1	2	3
B_2	2	3	4

$A_0 = B_0C_0 \Rightarrow P(A_0) = P(B_0)P(C_0) = 20/663.$

$A_1 = B_0C_1 + B_1C_0 \Rightarrow P(A_1) = P(B_0)P(C_1) + P(B_1)P(C_0) = 848/4641.$

$A_2 = B_0C_2 + B_1C_1 + B_2C_0 \Rightarrow P(A_2) = P(B_0)P(C_2)+P(B_1)P(C_1)+P(B_2)P(C_0) =757/1989.$

$A_3 = B_1C_2 + B_2C_1 \Rightarrow P(A_3) = P(B_1)P(C_2)+P(B_2)P(C_1) = 4400/13923.$

$A_4 = B_2C_2 \Rightarrow P(A_4) = P(B_2)P(C_2) = 20/221.$

Từ đó suy ra $P(A) = 0{,}5080$.

Bài 1.12: Có hai hộp cùng cỡ. Hộp thứ nhất chứa 4 bi trắng 6 bi xanh, hộp thứ hai chứa 5 bi trắng và 7 bi xanh. Chọn ngẫu nhiên một hộp rồi từ hộp đó lấy ra 2 bi thì được 2 bi trắng. Tính xác suất để viên bi tiếp theo cũng lấy từ hộp trên ra lại là bi trắng.

Lời giải

Gọi A_1 là biến cố 2 bi lấy đầu tiên là bi trắng.

A_2 là biến cố bi lấy lần sau là bi trắng.

Bài tóan yêu cầu tính $P(A_2/A_1)$.

Theo công thức nhân xác suất, ta có $P(A_1A_2) = P(A_1)\,P(A_2/A_1)$. Suy ra

$$P(A_2/A_1) = \frac{P(A_1A_2)}{P(A_1)}.$$

Bây giờ ta tính các xác suất $P(A_1)$ và $P(A_1A_2)$.

Gọi B_1, B_2 lần lượt là các biến cố chọn được hộp I, hộp II. Khi đó B_1, B_2 là một hệ đầy đủ, xung khắc từng đôi và ta có: $P(B_1) = P(B_2) = 0{,}5$.

Theo công thức xác suất đầy đủ, ta có

$$P(A_1) = P(B_1)\,P(A_1/B_1) + P(B_2)\,P(A_1/B_2)$$

Mà

$$P(A_1/B_1)=\frac{C_4^2C_6^0}{C_{10}^2}=\frac{6}{45};$$

$$P(A_1/B_2)=\frac{C_5^2C_7^0}{C_{12}^2}=\frac{10}{66}.$$

nên $P(A_1) = 47/330$.

Theo công thức xác suất đầy đủ, ta có

$$P(A_1A_2) = P(B_1)\,P(A_1A_2/B_1) + P(B_2)\,P(A_1A_2/B_2).$$

Mà

$$P(A_1A_2/B_1)=P(A_1/B_1)P(A_2/A_1B_1)=\frac{6}{45}\frac{2}{8}=\frac{1}{30};$$

$$P(A_1A_2/B_2)=P(A_1/B_2)P(A_2/A_1B_2)=\frac{10}{66}\frac{3}{10}=\frac{1}{22}.$$

nên $P(A_1A_2) = 13/330$. Suy ra xác suất cần tìm là $P(A_2/A_1) = 13/47 = 0,2766$.

Bài 1.13: Một lô hàng gồm a sản phẩm loại I và b sản phẩm loại II được đóng gới để gửi cho khách hàng. Nơi nhận kiểm tra lại thấy thất lạc 1 sản phẩm. Chọn ngẫu nhiên ra 1 sản phẩm thì thấy đó là sản phẩm loại I. Tính xác suất để sản phẩm thất lạc cũng thuộc loại I.

Lời giải

Gọi A là biến cố sản phẩm được chọn ra thuộc loại I.

A_1, A_2 lần lượt là các biến cố sản phẩm thất lạc thuộc loại I, loại II.

Yêu cầu của bài toán là tính xác suất có điều kiện $P(A_1/A)$.

Ta thấy A_1, A_2 là một hệ đầy đủ, xung khắc từng đôi và

$$P(A_1)=\frac{C_a^1C_b^0}{C_{a+b}^1}=\frac{a}{a+b};\quad P(A_2)=\frac{C_a^0C_b^1}{C_{a+b}^1}=\frac{b}{a+b}.$$

Theo công thức Bayes, ta có

$$P(A_1/A)=\frac{P(A_1)P(A/A_1)}{P(A)}=\frac{P(A_1)P(A/A_1)}{P(A_1)P(A/A_1)+P(A_2)P(A/A_2)}$$

Mà

$$P(A/A_1)=\frac{C_{a-1}^1C_b^0}{C_{a+b-1}^1}=\frac{a-1}{a+b-1};\quad P(A/A_2)=\frac{C_a^1C_{b-1}^0}{C_{a+b-1}^1}=\frac{a}{a+b-1}.$$

nên

$$P(A_1/A)=\frac{\frac{a}{a+b}.\frac{a-1}{a+b-1}}{\frac{a}{a+b}.\frac{a-1}{a+b-1}+\frac{b}{a+b}.\frac{a}{a+b-1}}=\frac{a-1}{a+b-1}$$

Bài 1.14: Có 3 hộp phấn, trong đó hộp I chứa 15 viên tốt và 5 viên xấu, hộp II chứa 10 viên tốt và 4 viên xấu, hộp III chứa 20 viên tốt và 10 viên xấu. Ta gieo một con xúc xắc cân đối. Nếu thấy xuất hiện mặt 1 chấm thì ta chọn hộp I; nếu xuất hiện mặt 2 hoặc 3 chấm thì chọn hộp II, còn xuất hiện các mặt còn lại thì chọn hộp III. Từ hộp được chọn lấy ngẫu nhiên ra 4 viên phấn. Tìm xác suất để lấy được ít nhất 2 viên tốt.

Lời giải

Gọi A là biến cố chọn được ít nhất 2 viên phấn tốt.

A_j (j =1,2, 3) là biến cố chọn được hộp thứ j. Khi đó A_1, A_2, A_3 là hệ đầy đủ, xung khắc từng đôi và ta có:

- A_1 xảy ra khi và chỉ khi thảy con xúc xắc, xuất hiện mặt 1 chấm, do đó $P(A_1) = 1/6$.
- Tương tự, $P(A_2) = 2/6$; $P(A_3) = 3/6$.

Theo công thức xác suất đầy đủ, ta có

$$P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3).$$

Từ giả thiết ta có:

$$P(A/A_1)=\frac{C_{15}^2C_5^2}{C_{20}^4}+\frac{C_{15}^3C_5^1}{C_{20}^4}+\frac{C_{15}^4C_5^0}{C_{20}^4}=\frac{4690}{4845};$$

$$P(A/A_2)=\frac{C_{10}^2C_4^2}{C_{14}^4}+\frac{C_{10}^3C_4^1}{C_{14}^4}+\frac{C_{10}^4C_4^0}{C_{14}^4}=\frac{960}{1001};$$

$$P(A/A_3)=\frac{C_{20}^2C_{10}^2}{C_{30}^4}+\frac{C_{20}^3C_{10}^1}{C_{30}^4}+\frac{C_{20}^4C_{10}^0}{C_{30}^4}=\frac{24795}{27405}.$$

Suy ra $P(A) = 0,9334$.

Bài 1.15: Có hai kiện hàng I và II. Kiện thứ nhất chứa 10 sản phẩm, trong đó có 8 sản phẩm loại A. Kiện thứ hai chứa 20 sản phẩm, trong đó có 4 sản phẩm loại A. Lấy từ mỗi kiện 2 sản phẩm. Sau đó, trong 4 sản phẩm thu được chọn ngẫu nhiên 2 sản phẩm. Tính xác suất để trong 2 sản phẩm chọn ra sau cùng có đúng 1 sản phẩm loại A.

Lời giải

Gọi C là biến cố trong 2 sản phẩm chọn ra sau cùng có đúng 1 sản phẩm loại A.

A_j (j = 0, 1, 2, 3, 4) là biến cố có j sản phẩm loại A và (4-j) sản phẩm loại B có trong 4 sản phẩm lấy từ hai kiện I và II. Khi đó A_0, A_1, A_2, A_3, A_4 là một hệ đầy đủ, xung khắc từng đôi. Theo công thức xác suất đầy đủ, ta có

$$P(C) = P(A_0)P(C/A_0) + P(A_1)P(C/A_1) + P(A_2)P(C/A_2) + P(A_3)P(C/A_3) + P(A_4)P(C/A_4).$$

Ta có:

$$P(C/A_0) = 0;$$
$$P(C/A_1) = \frac{C_1^1C_3^1}{C_4^2} = \frac{3}{6}$$
$$P(C/A_2) = \frac{C_2^1C_2^1}{C_4^2} = \frac{4}{6}$$
$$P(C/A_3) = \frac{C_3^1C_1^1}{C_4^2} = \frac{3}{6}$$
$$P(C/A_4) = 0.$$

Bây giờ ta tính $P(A_1)$; $P(A_2)$; $P(A_3)$.

Gọi B_i, C_i (i = 0, 1, 2) lần lượt là các biến cố có i sp A và (2 - i) sp B có trong 2 sp được chọn ra từ kiện I, kiện II. Khi đó

- B_0, B_1, B_2 xung khắc từng đôi và ta có:

$$P(B_0) = \frac{C_8^0C_2^2}{C_{10}^2} = \frac{1}{45};$$
$$P(B_1) = \frac{C_8^1C_2^1}{C_{10}^2} = \frac{16}{45};$$
$$P(B_2) = \frac{C_8^2C_2^0}{C_{10}^2} = \frac{28}{45}.$$

- C_0, C_1, C_2 xung khắc từng đôi và ta có:

$$P(C_0) = \frac{C_4^0C_{16}^2}{C_{20}^2} = \frac{120}{190};$$
$$P(C_1) = \frac{C_4^1C_{16}^1}{C_{20}^2} = \frac{64}{190};$$
$$P(C_2) = \frac{C_4^2C_{16}^0}{C_{20}^2} = \frac{6}{190};$$

- B_i và C_j độc lập.

- Tổng số sp A có trong 4 sp chọn ra phụ thuộc vào các biến cố B_i và C_j theo bảng sau:

	C_0	C_1	C_2
B_0	0	1	2
B_1	1	2	3
B_2	2	3	4

Ta có:

$A_1 = B_0C_1 + B_1C_0$.
$A_2 = B_0C_2 + B_1C_1 + B_2C_0$.
$A_3 = B_1C_2 + B_2C_1$.

Từ đây, nhờ các công thức cộng và nhân xác suất ta tính được:

$P(A_1) = 0,2320$; $P(A_2) = 0,5135$; $P(A_3) = 0,2208$.

Suy ra xác suất cần tìm là $P(C) = 0,5687$.

Bài 1.16: Một xạ thủ bắn 10 viên đạn vào một mục tiêu. Xác suất để 1 viên đạn bắn ra trúng mục tiêu là 0,8 . Biết rằng: Nếu có 10 viên trúng thì mục tiêu chắc chắn bị diệt. Nếu có từ 2 đến 9 viên trúng thì mục tiêu bị diệt vơi xác suất 80%. Nếu có 1 viên trúng thì mục tiêu bị diệt với xác suất 20%.

a) Tính xác suất để mục tiêu bị diệt.

b) Giả sử mục tiêu đã bị diệt. Tính xác suất có 10 viên trúng.

Lời giải

Tóm tắt:

- Số viên bắn ra: 10 viên.
- Xác suất trúng của mỗi viên: 0,8.

Số viên trúng	1	2-9	10
Xác suất mục tiêu bị diệt	20%	80%	100%

a) Gọi A là biến cố mục tiêu bị diệt.

A_0, A_1, A_2, A_3 lần lượt là các biến cố có 0; 1; 2-9; 10 viên trúng. Khi đó, A_0, A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi và giả thiết cho ta:

$$P(A/A_0) = 0;\ P(A/A_1) = 20\% = 0{,}2;$$
$$P(A/A_2) = 80\% = 0{,}8;\ P(A/A_3) = 100\% = 1.$$

Theo công thức xác suất đầy đủ, ta có:

$$P(A) = P(A_0)P(A/A_0) + P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3).$$

Theo công thức Bernoulli với n =10; p = 0,8, q = 0,2, ta có

$$P(A_0) = q^{10} = (0,2)^{10};$$
$$P(A_1) = C_{10}^1 pq^9 = 10(0,8)(0,2)^9;$$
$$P(A_3) = p^{10} = (0,8)^{10};$$
$$P(A_2) = 1 - P(A_0) - P(A_1) - P(A_3) = 1 - (0,2)^{10} - 10(0,8)(0,2)^9 - (0,8)^{10}.$$

Suy ra $P(A) = 0{,}8215$.

b) Giả sử mục tiêu đã bị diệt. Khi đó biến cố A đã xảy ra. Do đó xác suất có 10 viên trúng trong trường hợp này chính là xác suất có điều kiện $P(A_3/A)$.

Theo công thức Bayes, ta có:

$$P(A_3/A) = \frac{P(A_3)P(A/A_3)}{P(A)}$$

Từ đây ta tính được $P(A_3/A) = 0{,}1307$.

Bài 1.17: Một máy sản xuất sản phẩm với tỉ lệ sản phẩm loại A là 60%. Một lô hàng gồm 10 sản phẩm với tỉ lệ sản phẩm loại A là 60%. Cho máy sản xuất 2 sản phẩm và từ lô hàng lấy ra 3 sản phẩm.

a) Tính xác suất để số sản phẩm loại A có trong 2 sản phẩm do máy sản xuất bằng số sản phẩm loại A có trong 3 sản phẩm được lấy ra từ lô hàng.

b) Giả sử trong 5 sản phẩm thu được có 2 sản phẩm loại A. Tính xác suất để 2 sản phẩm loại A đó đều do máy sản xuất.

Lời giải

Gọi A_j (j = 0, 1, 2) là các biến cố có j sản phẩm loại A và (2-j) sản phẩm không thuộc loại A có trong 2 sản phẩm do máy sản xuất.

Gọi B_j (j = 0, 1, 2, 3) là các biến cố có j sản phẩm loại A và (3-j) sản phẩm không thuộc loại A có trong 3 sản phẩm lấy từ lô hàng.

Khi đó

- A_0, A_1, A_2 xung khắc từng đôi và theo công thức Bernoulli với n = 2; p = 0,6; q = 0,4 ta có:

$$P(A_0) = C_2^0 p^0 q^2 = (0,4)^2 = 0,16;$$
$$P(A_1) = C_2^1 p^1 q^1 = 2(0,6)(0,4) = 0,48;$$
$$P(A_2) = C_2^2 p^2 q^0 = (0,6)^2 = 0,36.$$

- B_0, B_1, B_2, B_3 xung khắc từng đôi và theo công thức tính xác suất lựa chọn với N = 10, N_A = 6, n= 3 ta có (vì lô hàng gồm 10 sản phẩm với tỉ lệ sản phẩm loại A là 60%, nghĩa là lô hàng gồm 6 sản phẩm loại A và 4 sản phẩm không thuộc loại A):

$$P(B_0) = \frac{C_6^0 C_4^3}{C_{10}^3} = \frac{4}{120};$$
$$P(B_1) = \frac{C_6^1 C_4^2}{C_{10}^3} = \frac{36}{120};$$
$$P(B_2) = \frac{C_6^2 C_4^1}{C_{10}^3} = \frac{60}{120};$$
$$P(B_3) = \frac{C_6^3 C_4^0}{C_{10}^3} = \frac{20}{120}.$$

- A_i và B_j độc lập.

a) Gọi C là biến cố số sản phẩm loại A có trong 2 sản phẩm do máy sản xuất bằng số sản phẩm loại A có trong 2 sản phẩm được lấy ra từ lô hàng. Ta có:

$$C = A_0B_0 + A_1B_1 + A_2B_2.$$

Từ đây, do tính xung khắc và độc lập, các công thức cộng và nhân xác suất cho ta:

$$P(C) = P(A_0)P(B_0) + P(A_1)P(B_1) + P(A_2)P(B_2) = 0{,}3293.$$

b) Gọi D là biến cố có 2 sản phẩm loại A trong 5 sản phẩm có được.
Giả sử trong 5 sản phẩm trên có 2 sản phẩm loại A. Khi đó biến cố D đã xảy ra. Do đó, xác suất để 2 sản phẩm loại A đó đều do máy sản xuất chính là xác suất có điều kiện $P(A_2/D)$.

Theo công thức nhân xác suất ta có:

$$P(A_2/D) = \frac{P(A_2D)}{P(D)}.$$

Nhận xét rằng tổng số sản phẩm loại A có trong 5 sản phẩm thu được phụ thuộc vào các biến cố A_i và B_j theo bảng sau:

	B_0	B_1	B_2	B_3
A_0	0	1	2	3
A_1	1	2	3	4
A_2	2	3	4	5

Suy ra

$$D = A_0B_2 + A_1B_1 + A_2B_0 \quad \text{và} \quad A_2D = A_2B_0.$$

Từ đây, ta tính được $P(D) = 0,236$; $P(A_2D) = 0,012$. Suy ra xác suất cần tìm là

$$P(A_2/D) = 0,0508.$$

Bài 1.18: Có hai lô hàng, mỗi lô chứa 60% sản phẩm tốt, trong đó lô I chứa 15 sản phẩm, lô II chứa rất nhiều sản phẩm. Từ lô II lấy ra 3 sản phẩm bỏ vào lô I, sau đó từ lô I lấy ra 2 sản phẩm.

a) Tính xác suất lấy được 1sp tốt, 1sp xấu từ lô I.

b) Tính xác suất lấy được 1sp tốt, 1sp xấu từ lô I, trong đó sp tốt có trong lô I từ trước.

c) Giả sử đã lấy được 1sp tốt, 1sp xấu từ lô I. Tính xác suất đã lấy được 2sp tốt, 1sp xấu từ lô II.

Lời giải

Gọi A_j (j = 0,1, 2, 3) là biến cố có j sản phẩm tốt và (3-j) sản phẩm xấu có trong 3 sản phẩm được chọn ra từ lô II. Khi đó A_0, A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi. Theo công thức Bernoulli ta có:

$$P(A_0) = C_3^0p^0q^3 = (0,4)^3 = 0,064;$$
$$P(A_1) = C_3^1p^1q^2 = 3(0,6)^1(0,4)^2 = 0,288;$$
$$P(A_2) = C_3^2p^2q^1 = 3(0,6)^2(0,4)^1 = 0,432;$$
$$P(A_3) = C_3^3p^3q^0 = (0,6)^3 = 0,216.$$

a) Gọi A là biến cố lấy được 1sp tốt, 1sp xấu từ lô I.
Theo công thức xác suất đầy đủ, ta có:

$$P(A) = P(A_0)P(A/A_0) + P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + P(A_3)P(A/A_3).$$

Từ giả thiết ta suy ra trong lô I có 15.60% = 9 sp tốt và 6 sp xấu. Do đó theo công thức tính xác suất lựa chọn, ta có:

$$P(A/A_0) = \frac{C_9^1C_9^1}{C_{18}^2} = \frac{81}{153};$$
$$P(A/A_1) = \frac{C_{10}^1C_8^1}{C_{18}^2} = \frac{80}{153};$$
$$P(A/A_2) = \frac{C_{11}^1C_7^1}{C_{18}^2} = \frac{77}{153};$$
$$P(A/A_3) = \frac{C_{12}^1C_6^1}{C_{18}^2} = \frac{72}{153}.$$

Suy ra xác suất cần tìm là: $P(A) = 0,5035$

b) Gọi B là biến cố lấy được 1sp tốt, 1sp xấu từ lô I, trong đó sp tốt có trong lô I từ trước. Theo công thức xác suất đầy đủ, ta có:

$$P(B) = P(A_0)P(B/A_0) + P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + P(A_3)P(B/A_3).$$

Ta có:

$$P(B/A_0) = \frac{C_9^1C_9^1}{C_{18}^2} = \frac{81}{153};$$
$$P(B/A_1) = \frac{C_9^1C_8^1}{C_{18}^2} = \frac{72}{153};$$
$$P(B/A_2) = \frac{C_9^1C_7^1}{C_{18}^2} = \frac{63}{153};$$
$$P(B/A_3) = \frac{C_9^1C_6^1}{C_{18}^2} = \frac{54}{153}.$$

Suy ra xác suất cần tìm là: $P(B) = 0,4235$.

c) Giả sử đã lấy được 1sp tốt, 1sp xấu từ lô I. Khi đó biến cố A đã xảy ra. Do đó xác suất đã lấy được 2sp tốt, 1sp xấu từ lô II trong trường hợp này chính là XS có điều kiện $P(A_2/A)$. Theo công thức Bayes, ta có:

$$P(A_2 / A) = \frac{P(A_2)P(A / A_2)}{P(A)} = \frac{0,432.\frac{77}{153}}{0,5035} = 0,4318.$$

-------------- ◦ -------------

BÀI GIẢI

XÁC SUẤT THỐNG KÊ

(GV: Trần Ngọc Hội – 2009)

CHƯƠNG 2

ĐẠI LƯỢNG NGẪU NHIÊN VÀ PHÂN PHỐI XÁC SUẤT

Bài 2.1: Nước giải khát được chở từ Sài Gòn đi Vũng Tàu. Mỗi xe chở 1000 chai bia Sài Gòn, 2000 chai coca và 800 chai nước trái cây. Xác suất để 1 chai mỗi loại bị bể trên đường đi tương ứng là 0,2%; 0,11% và 0,3%. Nếu không quá 1 chai bị bể thì lái xe được thưởng.

a) Tính xác suất có ít nhất 1 chai bia Sài Gòn bị bể.

b) Tính xác suất để lái xe được thưởng.

c) Lái xe phải chở ít mất mấy chuyến để xác suất có ít nhất một chuyến được thưởng không nhỏ hơn 0,9?

Lời giải

Tóm tắt:

Loại	Bia Sài Gòn	Coca	Nước trái cây
Số lượng/chuyến	1000	2000	800
Xác suất 1 chai bể	0,2%	0,11%	0,3%

- Gọi X_1 là ĐLNN chỉ số chai bia SG bị bể trong một chuyến. Khi đó, X_1 có phân phối nhị thức $X_1 \sim B(n_1,p_1)$ với $n_1 = 1000$ và $p_1 = 0,2\% = 0,002$. Vì n_1 khá lớn và p_1 khá bé nên ta có thể xem X_1 có phân phân phối Poisson:

$$X_1 \sim P(a_1) \text{ với } a_1 = n_1p_1 = 1000.0,002 = 2, \text{ nghĩa là}$$
$$X_1 \sim P(2).$$

- Tương tự, gọi X_2, X_3 lần lượt là các ĐLNN chỉ số chai bia coca, chai nước trái cây bị bể trong một chuyến. Khi đó, X_2, X_3 có phân phối Poisson:

$$X_2 \sim P(2000.0,0011) = P(2,2);$$
$$X_3 \sim P(800.0,003) = P(2,4).$$

a) Xác suất có ít nhất 1 chai bia Sài Gòn bị bể là

$$P(X_1 \geq 1) = 1 - P(X_1 = 0) = 1 - \frac{e^{-2}2^0}{0!} = 1 - e^{-2} = 0,8647.$$

b) Tính xác suất để lái xe được thưởng.

Theo giả thiết, lái xe được thưởng khi có không quá 1 chai bị bể, nghĩa là

$$X_1 + X_2 + X_3 \leq 1.$$

Vì $X_1 \sim P(2)$; $X_2 \sim P(2,2)$; $X_3 \sim P(2,4)$ nên $X_1 + X_2 + X_3 \sim P(2+2,2 + 2,4) = P(6,6)$

Suy ra xác suất lái xe được thưởng là:

$$P(X_1 + X_2 + X_3 \leq 1) = P[(X_1 + X_2 + X_3 =0) + P(X_1 + X_2 + X_3 = 1)] = \frac{e^{-6,6}(6,6)^0}{0!} + \frac{e^{-6,6}(6,6)^1}{1!} = 0,0103.$$

c) Lái xe phải chở ít mất mấy chuyến để xác suất có ít nhất một chuyến được thưởng không nhỏ hơn 0,9?

Gọi n là số chuyến xe cần thực hiện và A là biến cố có ít nhất 1 chuyến được thưởng. Yêu cầu bài toán là xác định n nhỏ nhất sao cho $P(A) \geq 0,9$.

Biến cố đối lập của A là: $\bar{A}$ không có chuyến nào được thưởng.

Theo câu b), xác suất để lái xe được thưởng trong một chuyến là p = 0,0103. Do đó theo công thức Bernoulli ta có:

$$P(A) = 1 - P(\bar{A}) = 1 - q^n = 1 - (1 - 0,0103)^n$$
$$= 1 - (0,9897)^n.$$

Suy ra

$$P(A) \geq 0,9 \Leftrightarrow 1 - (0,9897)^n \geq 0,9$$
$$\Leftrightarrow (0,9897)^n \leq 0,1$$
$$\Leftrightarrow n\ln(0,9897) \leq \ln 0,1$$
$$\Leftrightarrow n \geq \frac{\ln 0,1}{\ln(0,9897)} \approx 222,3987$$
$$\Leftrightarrow n \geq 223.$$

Vậy lái xe phải chở ít nhất là 223 chuyến.

Bài 2.2: Một máy tính gồm 1000 linh kiện A, 800 linh kiện B và 2000 linh kiện C. Xácsuất hỏng của ba linh kiện đó lần lượt là 0,02%; 0,0125% và 0,005%. Máy tính ngưng hoạt động khi số linh kiện hỏng nhiều hơn 1. Các linh kiện hỏng độc lập với nhau.

a) Tính xácsuất để có ít nhất 1 linh kiện B bị hỏng.

b) Tính xác suất để máy tính ngưng hoạt động.

c) Giả sử trong máy đã có 1 linh kiện hỏng. Tính xác suất để máy tính ngưng hoạt động.

Lời giải

Tóm tắt:

Loại linh kiện	A	B	C
Số lượng/1máy	1000	800	2000
Xác suất 1linh kiện hỏng	0,02%	0,0125%	0,005%

- Gọi X_1 là ĐLNN chỉ số linh kiện A bị hỏng trong một máy tính. Khi đó, X_1 có phân phối nhị thức $X_1 \sim B(n_1,p_1)$ với $n_1 = 1000$ và $p_1 = 0,02\% = 0,0002$. Vì n_1 khá lớn và p_1 khá bé nên ta có thể xem X_1 có phân phân phối Poisson:

$$X_1 \sim P(a_1) \text{ với } a_1 = n_1p_1 = 1000.0,0002 = 0,2, \text{ nghĩa là}$$

$$X_1 \sim P(0,2).$$

- Tương tự, gọi X_2, X_3 lần lượt là các ĐLNN chỉ số linh kiện B, C bị hỏng trong một máy tính. Khi đó, X_2 , X_3 có phân phối Poisson như sau:

$$X_2 \sim P(800.0,0125\%) = P(0,1);$$

$$X_3 \sim P(2000.0,005\%) = P(0,1).$$

a) Xác suất có ít nhất 1 linh linh kiện B bị hỏng là:

$$P(X_2 \geq 1) = 1 - P(X_2 = 0) = 1 - \frac{e^{-0,1}(0,1)^0}{0!} = 1 - e^{-0,1} = 0,0952.$$

b) Tính xác suất để máy tính ngưng hoạt động.

Theo giả thiết, máy tính ngưng hoạt động khi số linh kiện hỏng nhiều hơn 1, nghĩa là khi

$$X_1 + X_2 + X_3 > 1.$$

Vì $X_1 \sim P(0,2)$;$X_2 \sim P(0,1)$; $X_3 \sim P(0,1)$ nên $X_1 + X_2 + X_3 \sim P(0,2+0,1 + 0,1) = P(0,4)$

Suy ra xác suất để máy tính ngưng hoạt động là:

$P(X_1 + X_2 + X_3 > 1) = 1 - P(X_1 + X_2 + X_3 \leq 1)$

$= 1- [P(X_1 + X_2 + X_3 = 0) + P(X_1 + X_2 + X_3 = 1)] =$

$1 - \frac{e^{-0,4}(0,4)^0}{0!} - \frac{e^{-0,4}(0,4)^1}{1!}$

$= 1-1,4.e^{-0,4} = 0,0615 = 6,15\%.$

c) Giả sử trong máy đã có 1 linh kiện hỏng. Khi đó máy tính ngưng hoạt động khi có thêm ít nhất 1 linh kiện hỏng nữa, nghĩa là khi

$$X_1 + X_2 + X_3 \geq 1.$$

Suy ra xác suất để máy tính ngưng hoạt động trong trường hợp này là:

$P(X_1 + X_2 + X_3 \geq 1) = 1 - P(X_1 + X_2 + X_3 < 1) = 1- P(X_1 + X_2 + X_3 = 0)$

$$= 1 - \frac{e^{-0,4}(0,4)^0}{0!} = 1-e^{-0,4} = 0,3297 = 32,97\%.$$

Bài 2.3: Trọng lượng của một loại sản phẩm được quan sát là một đại lượng ngẫu nhiên có phân phối chuẩn với trung bình 50kg và phương sai $100kg^2$. Những sản phẩm có trọng lượng từ 45kg đến 70kg được xếp vào loại A. Chọn ngẫu nhiên 100 sản phẩm (trong rất nhiều sản phẩm). Tính xác suất để

a) có đúng 70 sản phẩm loại A.

b) có không quá 60 sản phẩm loại A.

c) có ít nhất 65 sản phẩm loại A.

Lời giải

Trước hết ta tìm xác suất để một sản phẩm thuộc loại A.

Gọi X_0 là trọng lượng của loại sản phẩm đã cho. Từ giả thiết ta suy ra X_0 có phân phối chuẩn $X_0 \sim N(\mu_0, \sigma_0^2)$ với $\mu_0 = 50$, $\sigma_0^2 = 100$ ($\sigma_0 = 10$). Vì một sản phẩm được xếp vào loại A khi có trọng lượng từ 45kg đến 70kg nên xác suất để một sản phẩm thuộc loại A là $P(45 \le X_0 \le 70)$.

Ta có

$$P(45 \le X_0 \le 70) = \varphi(\frac{70-\mu_0}{\sigma_0}) - \varphi(\frac{45-\mu_0}{\sigma_0}) = \varphi(\frac{70-50}{10}) - \varphi(\frac{45-50}{10})$$
$$= \varphi(2) - \varphi(-0,5) = \varphi(2) + \varphi(0,5) = 0,4772 + 0,1915 = 0,6687.$$

(Tra bảng giá trị hàm Laplace ta được φ (2) = 0,4772; φ (0,5) = 0,1915).

Vậy xác suất để một sản phẩm thuộc loại A là p =0,6687.

Bây giờ, kiểm tra 100 sản phẩm. Gọi X là số sản phẩm loại A có trong 100 sản phẩm được kiểm tra, thì X có phân phối nhị thức $X \sim B(n,p)$ với n = 100, p = 0,6687. Vì n = 100 khá lớn và p = 0,6687 không quá gần 0 cũng không quá gần 1 nên ta có thể xem X có phân phối chuẩn như sau:

$$X \sim N(\mu, \sigma^2)$$

với $\mu = np = 100.0,6687 = 66,87$;

$\sigma = \sqrt{npq} = \sqrt{100.0,6687.(1-0,6687)} = 4,7068.$

a) Xác suất để có 70 sản phẩm loại A là:

$$P(X = 70) = \frac{1}{\sigma} f(\frac{70-\mu}{\sigma}) = \frac{1}{4,7068} f(\frac{70-66,87}{4,7068})$$
$$= \frac{1}{4,7068} f(0,66) = \frac{0,3209}{4,7068} = 0,0681 = 6,81\%.$$

(Tra bảng giá trị hàm Gauss ta được f(0,66) = 0,3209).

b) Xác suất để có không quá 60 sản phẩm loại A là:

$$P(0 \le X \le 60) = \varphi(\frac{60-\mu}{\sigma}) - \varphi(\frac{0-\mu}{\sigma}) = \varphi(\frac{60-66,87}{4,7068}) - \varphi(\frac{0-66,87}{4,7068})$$
$$= \varphi(-1,46) - \varphi(-14,21) = -\varphi(1,46) + \varphi(14,21) = -\varphi(1,46) + \varphi(5)$$
$$= -0,4279 + 0,5 = 0,0721 = 7,21\%.$$

(Tra bảng giá trị hàm Laplace ta được φ (14,21) = φ (5) = 0,5; φ(1,46) = 0,4279).

c) Xác suất để có ít nhất 65 sản phẩm loại A là:

$$P(65 \le X \le 100) = \varphi(\frac{100-\mu}{\sigma}) - \varphi(\frac{65-\mu}{\sigma}) = \varphi(\frac{100-66,87}{4,7068}) - \varphi(\frac{65-66,87}{4,7068})$$
$$= \varphi(7,0388) - \varphi(-0,40) = \varphi(5) + \varphi(0,4) = 0,5 + 0,1554 = 0,6554 = 65,54\%.$$

(Tra bảng giá trị hàm Laplace ta được φ (7,7068)≈ φ (5) = 0,5; φ(0,4) = 0,1554).

Bài 2.4: Sản phẩm trong một nhà máy được đóng thành từng kiện, mỗi kiện gồm 14 sản phẩm trong đó có 8 sản phẩm loại A và 6 sản phẩm loại B. Khách hàng chọn cách kiểm tra như sau: từ mỗi kiện lấy ra 4 sản phẩm; nếu thấy số sản phẩm thuộc loại A nhiều hơn số sản phẩm thuộc loại B thì mới nhận kiện đó; ngược lại thì loại kiện đó. Kiểm tra 100 kiện (trong rất nhiều kiện). Tính xác suất để

a) có 42 kiện được nhận.

b) có từ 40 đến 45 kiện được nhận.

c) có ít nhất 42 kiện được nhận.

Lời giải

Trước hết ta tìm xác suất để một kiện được nhận.
Theo giả thiết, mỗi kiện chứa 14 sản phẩm gồm 8A và 6B. Từ mỗi kiện lấy ra 4 sản phẩm; nếu thấy số sản phẩm A nhiều hơn số sản phẩm B, nghĩa là được 3A,1B hoặc 4A, thì mới nhận kiện đó. Do đó xác suất để một kiện được nhận là:

$$P_4(3 \le k \le 4) = P_4(3) + P_4(4) = \frac{C_8^3 C_6^1}{C_{14}^4} + \frac{C_8^4 C_6^0}{C_{14}^4} = 0,4056$$

Vậy xác suất để một kiện được nhận là p = 0,4056.

Bây giờ, kiểm tra 100 kiện. Gọi X là số kiện được nhận trong 100 kiện được kiểm tra, thì X có phân phối nhị thức $X \sim B(n,p)$ với n = 100, p = 0,4056. Vì n = 100 khá lớn và p = 0,4056 không quá gần 0 cũng không quá gần 1 nên ta có thể xem X có phân phối chuẩn như sau:

$$X \sim N(\mu, \sigma^2)$$

với $\mu = np = 100.0,4056 = 40,56$;

$\sigma = \sqrt{npq} = \sqrt{100.0,4056.(1-0,4056)} = 4,9101.$

a) Xác suất để có 42 kiện được nhận là:

$$P(X=42)=\frac{1}{\sigma}f(\frac{42-\mu}{\sigma})=\frac{1}{4,9101}f(\frac{42-40,56}{4,9101})=\frac{1}{4,9101}f(0,29)$$
$$=\frac{0,3825}{4,9101}=0,0779=7,79\%.$$

(Tra bảng giá trị hàm Gauss ta được f(0,29) = 0,3825).

b) Xác suất để có từ 40 đến 45 kiện được nhận là

$$P(40\le X\le 45)=\varphi(\frac{45-\mu}{\sigma})-\varphi(\frac{40-\mu}{\sigma})=\varphi(\frac{45-40,56}{4,9101})-\varphi(\frac{40-40,56}{4,9101})$$
$$=\varphi(0,90)-\varphi(-0,11)=\varphi(0,90)+\varphi(0,11)=0,3159+0,0438=0,3597=35,97\%.$$

(Tra bảng giá trị hàm Laplace ta được φ (0,9) = 0,3519; φ (0,11) = 0,0438).

c) Xác suất để có ít nhất 42 kiện được nhận là

$$P(42\le X\le 100)=\varphi(\frac{100-\mu}{\sigma})-\varphi(\frac{42-\mu}{\sigma})=\varphi(\frac{100-40,56}{4,9101})-\varphi(\frac{42-40,56}{4,9101})$$
$$=\varphi(12)-\varphi(0,29)=0,50-0,1141=0,3859=38,59\%.$$

(Tra bảng giá trị hàm Laplace ta được φ(12) = φ(5) = 0,5; φ(0,29) = 0,1141).

Bài 2.5: Sản phẩm trong một nhà máy được đóng thành từng kiện, mỗi kiện gồm 10 sản phẩm Số sản phẩm loại A trong các hộp là X có phân phối như sau:

X	6	8
P	0,9	0,1

Khách hàng chọn cách kiểm tra như sau: từ mỗi kiện lấy ra 2 sản phẩm; nếu thấy cả 2 sản phẩm đều loại A thì mới nhận kiện đó; ngược lại thì loại kiện đó. Kiểm tra 144 kiện (trong rất nhiều kiện).

a) Tính xác suất để có 53 kiện được nhận.
b) Tính xác suất để có từ 52 đến 56 kiện được nhận.
c) Phải kiểm tra ít nhất bao nhiêu kiện để xác suất có ít nhất 1 kiện được nhận không nhỏ hơn 95%?

Lời giải

Trước hết ta tìm xác suất p để một kiện được nhận.
Gọi C là biến cố kiện hàng được nhận. Ta cần tìm p = P(C).
Từ giả thiết ta suy ra có hai loại kiện hàng:
Loại I: gồm 6A, 4B chiếm 0,9 = 90%.
Loại II: gồm 8A, 2B chiếm 0,1 = 10%.
Gọi A_1, A_2 lần lượt là các biến cố kiện hàng thuộc loại I, II. Khi đó A_1, A_2 là một hệ đầy đủ, xung khắc từng đôi và ta có

$$P(A_1) = 0,9;\ \ P(A_2) = 0,1.$$

Theo công thức xác suất đầy đủ ta có:

$$P(C) = P(A_1)\,P(C/A_1) + P(A_2)\,P(C/A_2).$$

Theo giả thiết, từ mỗi kiện lấy ra 2 sản phẩm; nếu cả 2 sản phẩm thuộc loại A thì mới nhận kiện đó. Do đó:

$$P(C/A_1)=P_2(2)=\frac{C_6^2C_4^0}{C_{10}^2}=\frac{1}{3};$$

$$P(C/A_2)=P_2(2)=\frac{C_8^2C_2^0}{C_{10}^2}=\frac{28}{45}.$$

Suy ra P(C) = 0,9. (1/3) + 0,1.(28/45) = 0,3622.
Vậy xác suất để một kiện được nhận là p = 0,3622.

Bây giờ, kiểm tra 144 kiện. Gọi X là số kiện được nhận trong 144 kiện được kiểm tra, thì X có phân phối nhị thức $X \sim B(n,p)$ với n = 144, p = 0,3622. Vì n = 144 khá lớn và p = 0,3622 không quá gần 0 cũng không quá gần 1 nên ta có thể xem X có phân phối chuẩn như sau:

$$X \sim N(\mu, \sigma^2)$$

với $\mu = np = 144.0,3622 = 52,1568;$

$$\sigma=\sqrt{npq}=\sqrt{144.0,3622.(1-0,3622)}=5,7676.$$

a) Xác suất để có 53 kiện được nhận là P(X=53) = 6,84% (Tương tự Bài 21).
b) Xác suất để có từ 52 đến 56 kiện được nhận là $P(52 \le X \le 56)$ = 26,05% (Tương tự Bài 21).
c) Phải kiểm tra ít nhất bao nhiêu kiện để xác suất có ít nhất 1 kiện được nhận không nhỏ hơn 95%?

Gọi n là số kiện cần kiểm tra và D là biến cố có ít nhất 1 kiện được nhận. Yêu cầu bài toán là xác định n nhỏ nhất sao cho $P(D) \ge 0,95$.

Biến cố đối lập của D là $\overline{D}$: không có kiện nào được nhận.
Theo chứng minh trên, xác suất để một kiện được nhận là $p = 0,3622$. Do đó
Theo công thức Bernoulli ta có:

$$P(D) = 1 - P(\overline{D}) = 1 - q^n = 1 - (1 - 0,3622)^n = 1 - (0,6378)^n.$$

Suy ra

$$\begin{aligned} P(D) \geq 0,95 &\Leftrightarrow 1 - (0,6378)^n \geq 0,95 \\ &\Leftrightarrow (0,6378)^n \leq 0,05 \\ &\Leftrightarrow n\ln(0,6378) \leq \ln 0,05 \\ &\Leftrightarrow n \geq \frac{\ln 0,05}{\ln(0,6378)} \approx 6,6612 \\ &\Leftrightarrow n \geq 7. \end{aligned}$$

Vậy phải kiểm tra ít nhất 7 kiện.

Bài 2.6: Một máy sản xuất sản phẩm với tỉ lệ sản phẩm đạt tiêu chuẩn là 80% và một máy khác cũng sản xuất loại sản phẩm này với tỉ lệ sản phẩm đạt tiêu chuẩn là 60%. Chọn ngẫu nhiên một máy và cho sản xuất 100 sản phẩm. Tính xác suất để

a) có 70 sản phẩm đạt tiêu chuẩn.
b) có từ 70 đến 90 sản phẩm đạt tiêu chuẩn.
c) có không ít hơn 70 sản phẩm đạt tiêu chuẩn.

Lời giải

Gọi X là ĐLNN chỉ số sản phẩm đạt tiêu chuẩn trong 100 sản phẩm.
A_1, A_2 lần lượt là các biến cố chọn được máy 1, máy 2.
Khi đó A_1, A_2 là một hệ đầy đủ, xung khắc từng đôi và ta có:

$$P(A_1) = P(A_2) = 0,5.$$

Theo công thức xác xuất đầy đủ, với mỗi $0 \leq k \leq 100$, ta có:

$$\begin{aligned} P(X = k) &= P(A_1)P(X{=}k/A_1) + P(A_2)P(X{=}k/A_2) \\ &= \frac{1}{2}P(X{=}k/A_1) + \frac{1}{2}P(X{=}k/A_2) \end{aligned} \qquad (1)$$

Như vậy, gọi X_1, X_2 lần lượt là các ĐLNN chỉ số sản phẩm đạt tiêu chuẩn trong trường hợp chọn được máy 1, máy 2. Khi đó:

- (1) cho ta $P(X = k) = \frac{1}{2}P(X_1{=}k) + \frac{1}{2}P(X_2{=}k)$

- X_1 có phân phối nhị thức $X_1 \sim B(n_1,p_1)$ với $n_1 = 100$, $p_1 = 80\% = 0,8$. Vì $n_1 = 100$ khá lớn và $p_1 = 0,8$ không quá gần 0 cũng không quá gần 1 nên ta có thể xem X_1 có phân phối chuẩn như sau:
$$X_1 \sim N(\mu_1, \sigma_1^2)$$
với $\mu_1 = n_1p_1 = 100.0,8 = 80$;
$\sigma_1 = \sqrt{n_1p_1q_1} = \sqrt{100.0,8.0,2} = 4.$
- X_2 có phân phối nhị thức $X_2 \sim B(n_2,p_2)$ với $n_2 = 100$, $p_2 = 60\% = 0,60$. Vì $n_2 = 100$ khá lớn và $p_2 = 0,60$ không quá gần 0 cũng không quá gần 1 nên ta có thể xem X_2 có phân phối chuẩn như sau:
$$X_2 \sim N(\mu_2, \sigma_2^2)$$
với $\mu_2 = n_2p_2 = 100.0,60 = 60$;
$\sigma_2 = \sqrt{n_2p_2q_2} = \sqrt{100.0,60.0,40} = 4,8990.$

a) Xác suất để có 70 sản phẩm đạt tiêu chuẩn là:

$$P(X = 80) = \frac{1}{2}P(X_1{=}70) + \frac{1}{2}P(X_2{=}70) = \frac{1}{2}\frac{1}{\sigma_1}f(\frac{70-\mu_1}{\sigma_1}) + \frac{1}{2}\frac{1}{\sigma_2}f(\frac{70-\mu_2}{\sigma_2})$$

$$= \frac{1}{2}.\frac{1}{4}f(\frac{70-80}{4}) + \frac{1}{2}.\frac{1}{4,8990}f(\frac{70-60}{4,8990}) = \frac{1}{2}.\frac{1}{4}f(-2,5) + \frac{1}{2}.\frac{1}{4,8990}f(2,04)$$

$$= \frac{1}{2}.\frac{1}{4}0,0175 + \frac{1}{2}.\frac{1}{4,8990}0,0498 = 0,000727$$

b) Xác suất để có từ 70 đến 90 sản phẩm đạt tiêu chuẩn là:

$$P(70 \leq X \leq 90) = \frac{1}{2}P(70 \leq X_1 \leq 90) + \frac{1}{2}P(70 \leq X_2 \leq 90)$$

$$= \frac{1}{2}[\varphi(\frac{90-\mu_1}{\sigma_1}) - \varphi(\frac{70-\mu_1}{\sigma_1})] + \frac{1}{2}[\varphi(\frac{90-\mu_2}{\sigma_2}) - \varphi(\frac{70-\mu_2}{\sigma_2})]$$

$$= \frac{1}{2}[\varphi(\frac{90-80}{4}) - \varphi(\frac{70-80}{4})] + \frac{1}{2}[\varphi(\frac{90-60}{4,899}) - \varphi(\frac{70-60}{4,899})]$$

$$= \frac{1}{2}[\varphi(2,5) - \varphi(-2,5) + \varphi(6,12) - \varphi(2,04)]$$

$$= \frac{1}{2}(0,49379 + 0,49379 + 0,5 - 0,47932)$$

$$= 0,50413$$

c) Xác suất có không ít hơn 70 sản phẩm đạt tiêu chuẩn là

$$P(70 \leq X \leq 100) = 0,5072$$

(Tương tự câu b)

Bài 2.7: Một máy sản xuất sản phẩm với tỉ lệ phế phẩm là 1% và một máy khác cũng sản xuất loại sản phẩm này với tỉ lệ phế phẩm là 2%.

Chọn ngẫu nhiên một máy và cho sản xuất 1000 sản phẩm. Tính xác suất để

a) có 14 phế phẩm.

b) có từ 14 đến 20 phế phẩm.

Lời giải

Gọi X là ĐLNN chỉ số phế phẩm trong 1000 sản phẩm.

A_1, A_2 lần lượt là các biến cố chọn được máy 1, máy 2.

Khi đó A_1, A_2 là một hệ đầy đủ, xung khắc từng đôi và ta có:

$$P(A_1) = P(A_2) = 0{,}5.$$

Theo công thức xác xuất đầy đủ, với mỗi $0 \le k \le 100$, ta có:

$$P(X = k) = P(A_1)P(X=k/A_1) + P(A_2)P(X= k/A_2) = \frac{1}{2}P(X=k/A_1)+\frac{1}{2}P(X=k/A_2) \quad (1)$$

Như vậy, gọi X_1, X_2 lần lượt là các ĐLNN chỉ số phế phẩm trong trường hợp chọn được máy 1, máy 2. Khi đó:

- (1) cho ta $P(X = k) = \frac{1}{2}P(X_1=k)+\frac{1}{2}P(X_2=k)$
- X_1 có phân phối nhị thức $X_1 \sim B(n_1,p_1)$ với $n_1 = 1000$ và $p_1 = 1\% = 0{,}001$. Vì n_1 khá lớn và p_1 khá bé nên ta có thể xem X_1 có phân phân phối Poisson:

  $X_1 \sim P(a_1)$ với $a_1 = n_1p_1 = 1000.0{,}01 = 10$, nghĩa là $X_2 \sim P(10)$.

- X_2 có phân phối nhị thức $X_2 \sim B(n_2,p_2)$ với $n_2 = 1000$ và $p_2 = 2\% = 0{,}002$. Vì n_2 khá lớn và p_2 khá bé nên ta có thể xem X_2 có phân phân phối Poisson:

  $X_1 \sim P(a_2)$ với $a_2 = n_2p_2 = 1000.0{,}02 = 20$, nghĩa là $X_2 \sim P(20)$.

a) Xác suất để có 14 phế phẩm là:

$$P(X = 14) = \frac{1}{2}P(X_1=14)+\frac{1}{2}P(X_2=14) = \frac{1}{2}\frac{e^{-10}10^{14}}{14!}+\frac{1}{2}\frac{e^{-20}20^{14}}{14!} = 0{,}0454$$

b) Xác suất để có từ 14 đến 20 phế phẩm là:

$$P(14 \le X \le 20) = \frac{1}{2}P(14 \le X_1 \le 20)+\frac{1}{2}P(14 \le X_2 \le 20)$$

$$= \frac{1}{2}\sum_{k=14}^{20}\frac{e^{-10}10^k}{k!}+\frac{1}{2}\sum_{k=14}^{20}\frac{e^{-20}20^k}{k!} = 31{,}35\%$$

Bài 2.8: Một xí nghiệp có hai máy I và II. Trong ngày hội thi, mỗi công nhân dự thi được phân một máy và với máy đó sẽ sản xuất 100 sản phẩm. Nếu số sản phẩm loại A không ít hơn 70 thì công nhân đó sẽ được thưởng. Giả sử đối với công nhân X, xác suất sản xuất được 1 sản phẩm loại A với các máy I và II lần lượt là 0,6 và 0,7.

a) Tính xác suất để công nhân X được thưởng.

b) Giả sử công nhân X dự thi 50 lần. Số lần được thưởng tin chắc nhất là bao nhiêu?

Lời giải

Gọi Y là ĐLNN chỉ số sản phẩm loại A có trong 100 sản phẩm được sản xuất.

A_1, A_2 lần lượt là các biến cố chọn được máy I, máy II.

Khi đó A_1, A_2 là một hệ đầy đủ, xung khắc từng đôi và ta có:

$$P(A_1) = P(A_2) = 0{,}5.$$

Theo công thức xác xuất đầy đủ, với mỗi $0 \le k \le 100$, ta có:

$$P(Y = k) = P(A_1)P(Y=k/A_1) + P(A_2)P(Y= k/A_2) = \frac{1}{2}P(Y=k/A_1)+\frac{1}{2}P(Y=k/A_2) \quad (1)$$

Như vậy, gọi X_1, X_2 lần lượt là các ĐLNN chỉ số sản phẩm loại A có trong 100 sản phẩm được sản xuất trong trường hợp chọn được máy I, máy II. Khi đó:

- (1) cho ta $P(Y = k) = \frac{1}{2}P(X_1=k)+\frac{1}{2}P(X_2=k)$
- X_1 có phân phối nhị thức $X_1 \sim B(n_1,p_1)$ với $n_1 = 100$, $p_1 = 0{,}6$. Vì $n_1 = 100$ khá lớn và $p_1 = 0{,}6$ không quá gần 0 cũng không quá gần 1 nên ta có thể xem X_1 có phân phối chuẩn như sau:

  $$X_1 \sim N(\mu_1, \sigma_1^2)$$

  với $\mu_1 = n_1p_1 = 100.0{,}6 = 60$;

  $\sigma_1 = \sqrt{n_1p_1q_1} = \sqrt{100.0,6.0,4} = 4,8990.$

- X_2 có phân phối nhị thức $X_2 \sim B(n_2,p_2)$ với $n_2 = 100$, $p_2 = 0{,}7$. Vì $n_2 = 100$ khá lớn và $p_2 = 0{,}7$ không quá gần 0 cũng không quá gần 1 nên ta có thể xem X_2 có phân phối chuẩn như sau:

  $$X_2 \sim N(\mu_2, \sigma_2^2)$$

  với $\mu_1 = n_2p_2 = 100.0{,}7 = 70$;

  $\sigma_2 = \sqrt{n_2p_2q_2} = \sqrt{100.0,7.0,3} = 4,5826.$

a) Xác suất để công nhân X được thưởng là:

$$P(70 \le Y \le 100) = \frac{1}{2}P(70 \le X_1 \le 100) + \frac{1}{2}P(70 \le X_2 \le 100)$$
$$= \frac{1}{2}[\varphi(\frac{100-\mu_1}{\sigma_1}) - \varphi(\frac{70-\mu_1}{\sigma_1})] + \frac{1}{2}[\varphi(\frac{100-\mu_2}{\sigma_2}) - \varphi(\frac{70-\mu_2}{\sigma_2})]$$
$$= \frac{1}{2}[\varphi(\frac{100-60}{4,899}) - \varphi(\frac{70-60}{4,899})] + \frac{1}{2}[\varphi(\frac{100-70}{4,5826}) - \varphi(\frac{70-70}{4,5826})]$$
$$= \frac{1}{2}[\varphi(8,16) - \varphi(2,04) + \varphi(6,55) - \varphi(0)] = \frac{1}{2}(0,5 - 0,47932 + 0,5) = 0,2603$$

b) Giả sử công nhân X dự thi 50 lần. Số lần được thưởng tin chắc nhất là bao nhiêu?

Gọi Z là ĐLNN chỉ số lần công nhân X được thưởng. Khi đó Z có phân phối nhị thức Z ~ B(n,p) với n = 50, p = 0,2603. Số lần được thưởng tin chắc nhất chính là Mod(Z). Ta có:

$$\begin{aligned} Mod(Z) = k &\Leftrightarrow np - q \le k \le np - q + 1 \\ &\Leftrightarrow 50.0,2603 - 0,7397 \le k \le 50.0,2603 - 0,7397 + 1 \\ &\Leftrightarrow 12,2753 \le k \le 13,2753 \ \Leftrightarrow k = 13 \end{aligned}$$

Vậy số lần được thưởng tin chắc nhất của công nhân X là 13 lần.

Bài 2.9: Trong ngày hội thi, mỗi chiến sĩ sẽ chọn ngẫu nhiên một trong hai loại súng và với khẩu súng chọn được sẽ bắn 100viên đạn. Nếu có từ 65 viên trở lên trúng bia thì được thưởng. Giả sử đối với chiến sĩ A, xác suất bắn 1 viên trúng bia bằng khẩu súng loại I là 60% và bằng khẩu súng loại II là 50%.
a) Tính xác suất để chiến sĩ A được thưởng.
b) Giả sử chiến sĩ A dự thi 10 lần. Hỏi số lần được thưởng tin chắc nhất là bao nhiêu?
c) Chiến sĩ A phải tham gia hội thi ít nhất bao nhiêu lần để xác suất có ít nhất một lần được thưởng không nhỏ hơn 98%?

Lời giải

Gọi X là ĐLNN chỉ số viên trúng trong 100 viên được bắn ra.
Gọi A_1, A_2 lần lượt là các biến cố chọn được khẩu súng loại I, II.
Khi đó A_1, A_2 là một hệ đầy đủ, xung khắc từng đôi và ta có:
$$P(A_1) = P(A_2) = 0,5.$$
Theo công thức xác xuất đầy đủ, với mỗi $0 \le k \le 100$, ta có:

$$P(X = k) = P(A_1)P(X=k/A_1) + P(A_2)P(X=k/A_2)$$
$$= \frac{1}{2}P(X=k/A_1) + \frac{1}{2}P(X=k/A_2) \qquad (1)$$

Như vậy, gọi X_1, X_2 lần lượt là các ĐLNN chỉ số viên trúng trong 100 viên được bắn ra trong trường hợp chọn được khẩu loại I, II. Khi đó:

- (1) cho ta $P(X = k) = \frac{1}{2}P(X_1=k) + \frac{1}{2}P(X_2=k)$
- X_1 có phân phối nhị thức $X_1 \sim B(n_1,p_1)$ với $n_1 = 100$, $p_1 = 0,6$. Vì $n_1 = 100$ khá lớn và $p_1 = 0,6$ không quá gần 0 cũng không quá gần 1 nên ta có thể xem X_1 có phân phối chuẩn như sau:
  $$X_1 \sim N(\mu_1, \sigma_1^2)$$
  với $\mu_1 = n_1p_1 = 100.0,6 = 60$;
  $\sigma_1 = \sqrt{n_1p_1q_1} = \sqrt{100.0,6.0,4} = 4,8990.$
- X_2 có phân phối nhị thức $X_2 \sim B(n_2,p_2)$ với $n_2 = 100$, $p_2 = 0,5$. Vì $n_2 = 100$ khá lớn và $p_2 = 0,5$ không quá gần 0 cũng không quá gần 1 nên ta có thể xem X_2 có phân phối chuẩn như sau:
  $$X_2 \sim N(\mu_2, \sigma_2^2)$$
  với $\mu_1 = n_2p_2 = 100.0,5 = 50$;
  $\sigma_2 = \sqrt{n_2p_2q_2} = \sqrt{100.0,5.0,5} = 5.$

a) Xác suất để chiến sĩ A được thưởng là:
$$P(65 \le X \le 100) = \frac{1}{2}P(65 \le X_1 \le 100) + \frac{1}{2}P(65 \le X_2 \le 100)$$
$$= \frac{1}{2}[\varphi(\frac{100-\mu_1}{\sigma_1}) - \varphi(\frac{65-\mu_1}{\sigma_1})] + \frac{1}{2}[\varphi(\frac{100-\mu_2}{\sigma_2}) - \varphi(\frac{65-\mu_2}{\sigma_2})]$$
$$= \frac{1}{2}[\varphi(\frac{100-60}{4,899}) - \varphi(\frac{65-60}{4,899})] + \frac{1}{2}[\varphi(\frac{100-50}{5}) - \varphi(\frac{65-50}{5})]$$
$$= \frac{1}{2}[\varphi(8,16) - \varphi(1,02) + \varphi(10) - \varphi(3)] = \frac{1}{2}(0,5 - 0,34614 + 0,5 - 0,49865) = 0,0776.$$

b) Giả sử chiến sĩ A dự thi 10 lần. Số lần được thưởng tin chắc nhất là bao nhiêu?
Gọi Y là ĐLNN chỉ số lần chiến sĩ A được thưởng. Khi đó Y có phân phối nhị thức Y ~ B(n,p) với n = 10, p = 0,0776. Số lần được thưởng tin chắc nhất chính là mod(Y). Ta có:

$$\begin{aligned} mod(Y) = k &\Leftrightarrow np - q \le k \le np - q + 1 \\ &\Leftrightarrow 10.0,0776 - 0,9224 \le k \le 10.0,0776 - 0,9224 + 1 \\ &\Leftrightarrow -0,1464 \le k \le 0,8536 \ \Leftrightarrow k = 0 \end{aligned}$$

Vậy số lần được thưởng tin chắc nhất của chiến sĩ A là 0 lần, nói cách khác, thường là chiến sĩ A không được thưởng lần nào trong 10 lần tham gia.

c) Chiến sĩ A phải tham gia hội thi ít nhất bao nhiêu lần để xác suất có ít nhất một lần được thưởng không nhỏ hơn 98%?

Gọi n là số lần tham gia hội thi và D là biến cố có ít nhất 1 lần được thưởng. Yêu cầu bài toán là xác định n nhỏ nhất sao cho $P(D) \geq 0,98$.
Biến cố đối lập của D là $\bar{D}$: không có lần nào được thưởng.
Theo chứng minh trên, xác suất để một lần được thưởng là p = 0,0776. Do đó
Theo công thức Bernoulli ta có:

$$P(D) = 1 - P(\bar{D}) = 1 - q^n = 1 - (1 - 0,0776)^n = 1 - (0,9224)^n.$$

Suy ra

$$\begin{aligned} P(D) \geq 0,98 &\Leftrightarrow 1 - (0,9224)^n \geq 0,98 \\ &\Leftrightarrow (0,9224)^n \leq 0,02 \\ &\Leftrightarrow n \ln 0,9224 \leq \ln 0,02 \\ &\Leftrightarrow n \geq \frac{\ln 0,02}{\ln 0,9224} \approx 48,43 \\ &\Leftrightarrow n \geq 49. \end{aligned}$$

Vậy chiến sĩ A phải tham gia hội thi ít nhất là 49 lần.

Bài 2.10: Một người thợ săn bắn 4 viên đạn. Biết xác suất trúng đích của mỗi viên đạn bắn ra là 0,8. Gọi X là đại lượng ngẫu nhiên chỉ số viên đạn trúng đích.
a) Tìm luật phân phối của X.
b) Tìm kỳ vọng và phương sai của X.

Lời giải

a) Ta thấy X có phân phối nhị thức X~ B(n,p) với n = 4, p = 0,8. X là ĐLNN rời rạc nhận 5 giá trị: 0, 1, 2, 3 , 4. Luật phân phối của X có dạng:

X	0	1	2	3	4
P	p_0	p_1	p_2	p_3	p_4

Theo công thức Bernoulli ta có:

$$P(X=0) = C_4^0(0,8)^0(0,2)^4 = 0,0016;$$
$$P(X=1) = C_4^1(0,8)^1(0,2)^3 = 0,0256;$$
$$P(X=2) = C_4^2(0,8)^2(0,2)^2 = 0,1536;$$
$$P(X=3) = C_4^3(0,8)^3(0,2)^1 = 0,4096;$$
$$P(X=4) = C_4^4(0,8)^4(0,2)^0 = 0,4096.$$

Vậy luật phân phối của X là:

X	0	1	2	3	4
P	0,0016	0,0256	0,1536	0,4096	0,4096

b) Tìm kỳ vọng và phương sai của X.
- Kỳ vọng: M(X) = np = 3,2.
- Phương sai: D(X) = npq = 0,64.

Bài 2.11: Có hai lô hàng I và II, mỗi lô chứa rất nhiều sản phẩm. Tỉ lệ sản phẩm loại A có trong hai lô I và II lần lượt là 70% và 80%. Lấy ngẫu nhiên từ mỗi lô 2 sản phẩm.
a) Tính xác suất để số sản phẩm loại A lấy từ lô I lớn hơn số sản phẩm loại A lấy từ lô II.
b) Gọi X là số sản phẩm loại A có trong 4 sản phẩm được lấy ra. Tìm kỳ vọng và phương sai của X.

Lời giải

Gọi X_1, X_2 lần lượt là các ĐLNN chỉ số sp loại A có trong 2 sp được chọn ra từ lô I, II. Khi đó

- X_1 có phân phối nhị thức $X_1 \sim B(n_1, p_1)$; $n_1 = 2$; $p_1 = 70\% = 0,7$

với các xác suất định bởi:

$$P(X_1 = k) = C_2^k(0,7)^k(0,3)^{2-k}$$

Cụ thể

X_1	0	1	2
P	0,09	0,42	0,49

- X_2 có phân phối nhị thức $X_2 \sim B(n_2, p_2)$; $n_2 = 2$; $p_2 = 80\% = 0,8$

với các xác suất định bởi:

$$P(X_2 = k) = C_2^k(0,8)^k(0,2)^{2-k}$$

Cụ thể

X_2	0	1	2
P	0,04	0,32	0,64

a) Xác suất để số sản phẩm loại A lấy từ lô I lớn hơn số sản phẩm loại A lấy từ lô II là:

$P(X_1 \geq X_2) = P[(X_1 =2)(X_2 =0)+ (X_1 =2)(X_2 =1)+ (X_1 =1)(X_2 =0)]$

$= P(X_1 =2)P(X_2 =0)+ P(X_1 =2)P(X_2 =1)+ P(X_1 =1)P(X_2 =0) =$ 0,1932.

b) Gọi X là số sp loại A có trong 4 sp chọn ra . Khi đó

$$X = X_1 + X_2$$

Vì X_1 , X_2 độc lập nên ta có:

- Kỳ vọng của X là $M(X) = M(X_1) + M(X_2) = n_1p_1 + n_2p_2 = 3$
- Phương sai của X là $D(X) = D(X_1) + D(X_2) = n_1p_1q_1 + n_2p_2q_2 =$ 0,74.

Bài 2.12: Cho hai hộp I và II, mỗi hộp có 10 bi; trong đó hộp I gồm 6 bi đỏ, 4 bi trắng và hộp II gồm 7 bi đỏ, 3 bi trắng. Rút ngẫu nhiên từ mỗi hộp hai bi.

a) Tính xác suất để được hai bi đỏ và hai bi trắng.

b) Gọi X là đại lượng ngẫu nhiên chỉ số bi đỏ có trong 4 bi được rút ra. Tìm luật phân phối của X.

Lời giải

Gọi X_1, X_2 lần lượt là các ĐLNN chỉ số bi đỏ có trong 2 bi được chọn ra từ hộp I, hộp II. Khi đó

- X_1 có phân phối siêu bội $X_1 \sim H(N_1, N_{1A}, n_1)$; $N_1 = 10$; $N_{1A}= 6$; $n_1 =$ 2 với các xác suất định bởi:

$$P(X_1 = k) = \frac{C_6^k C_4^{2-k}}{C_{10}^2}.$$

Cụ thể

X_1	0	1	2
P	6/45	24/45	15/45

- X_2 có phân phối siêu bội $X_2 \sim H(N_2, N_{2A}, n_2)$; $N_2 = 10$; $N_{2A} = 7$; n_2 = 2

với các xác suất định bởi:

$$P(X_2 = k) = \frac{C_7^k C_3^{2-k}}{C_{10}^2}.$$

Cụ thể

X_2	0	1	2
P	3/45	21/45	21/45

Gọi X là đại lượng ngẫu nhiên chỉ số bi đỏ có trong 4 bi được rút ra. Khi đó

$$X = X_1 + X_2$$

Bảng giá trị của X dựa vào X_1, X_2 như sau:

X X_2 / X_1	0	1	2
0	0	1	2
1	1	2	3
2	2	3	4

a) Xác suất để được 2 bi đỏ và 2 bi trắng là:

$P(X = 2) = P[(X_1=0) (X_2=2)+ (X_1=1) (X_2=1)+ (X_1=2) (X_2=0)]$

$= P(X_1=0) P(X_2=2)+ P(X_1=1)P(X_2=1)+ P(X_1=2)P(X_2=0)]$

$= (6/45)(21/45) + (24/45)(21/45) + (15/45)(3/45) = 1/3.$

b) Luật phân phối của X có dạng:

X	0	1	2	3	4
P	p_0	p_1	p_2	p_3	p_4

trong đó:

$p_0 = P(X = 0)= P(X_1 =0) P(X_2 = 0) = 2/225;$

$p_1 = P(X = 1)= P(X_1 =0) P(X_2 = 1) + P(X_1 =1) P(X_2 = 0)= 22/225;$

$p_2 = P(X = 2) = 1/3;$

$p_3 = P(X = 3)= P(X_1 =1) P(X_2 = 2) + P(X_1 =2) P(X_2 = 1)= 91/225;$

$p_4 = P(X = 4)= P(X_1 =2) P(X_2 = 2) = 7/45.$

Vậy luật phân phối của X là :

X	0	1	2	3	4
P	2/225	22/225	1/3	91/225	7/45

Bài 2.13: Một máy sản xuất sản phẩm với tỉ lệ phế phẩm 10%. Một lô hàng gồm 10 sản phẩm với tỉ lệ phế phẩm 30%. Cho máy sản xuất 3 sản phẩm và từ lô hàng lấy ra 3 sản phẩm. Gọi X là số sản phẩm tốt có trong 6 sản phẩm này.

a) Tìm luật phân phối của X.

b) Không dùng luật phân phối của X, hãy tính M(X), D(X).

Lời giải

Gọi X_1, X_2 lần lượt là các ĐLNN chỉ số sp tốt có trong 3 sản phẩm do máy sản xuất; do lấy từ lô hàng. Khi đó X_1, X_2 độc lập và ta có:

- X_1 có phân phối nhị thức $X_1 \sim B(n_1, p_1)$; $n_1 = 3$; $p_1 = 0,9$. Cụ thể ta có:

$$P(X_1 = 0) = C_3^0 p^0 q^2 = (0,1)^3 = 0,001;$$
$$P(X_1 = 1) = C_3^1 p^1 q^2 = 3(0,9)(0,1)^2 = 0,027;$$
$$P(X_1 = 2) = C_3^2 p^2 q^1 = 3(0,9)^2(0,1) = 0,243;$$
$$P(X_1 = 3) = C_3^3 p^3 q^0 = (0,9)^3 = 0,729.$$

- X_2 có phân phối siêu bội $X_2 \sim H(N_2, N_{2A}, n_2)$; $N_2 = 10$; $N_{2A} = 7$; $n_2 = 3$ (vì lô hàng gồm 10 sản phẩm với tỉ lệ phế phẩm là 30%, nghĩa là lô hàng gồm 7 sản phẩm tốt và 3 sản phẩm xấu). Cụ thể ta có:

$$P(X_2 = 0) = \frac{C_7^0 C_3^3}{C_{10}^3} = \frac{1}{120};$$
$$P(X_2 = 1) = \frac{C_7^1 C_3^2}{C_{10}^3} = \frac{21}{120};$$
$$P(X_2 = 2) = \frac{C_7^2 C_3^1}{C_{10}^3} = \frac{63}{120};$$
$$P(X_2 = 3) = \frac{C_7^3 C_3^0}{C_{10}^3} = \frac{35}{120}.$$

a) Ta có $X = X_1 + X_2$. Luật phân phối của X có dạng:

X	0	1	2	3	4	5	6
P	p_0	p_1	p_2	p_3	p_4	p_5	p_6

trong đó:

$p_0 = P(X = 0) = P(X_1 = 0)P(X_2 = 0) = 1/120000;$

$p_1 = P(X = 1) = P(X_1 = 0)P(X_2 = 1) + P(X_1 = 1)P(X_2 = 0) = 1/2500;$

$p_2 = P(X = 2) = P(X_1 = 0)P(X_2 = 2) + P(X_1 = 1)P(X_2 = 1) + P(X_1 = 2)P(X_2 = 0)$
$= 291/40000$

$p_3 = P(X = 3) = P(X_1 = 0)P(X_2 = 3) + P(X_1 = 1)P(X_2 = 2) + P(X_1 = 2)P(X_2 = 1)$
$+ P(X_1 = 3)P(X_2 = 0) = 473/7500$

$p_4 = P(X = 4) = P(X_1 = 1)P(X_2 = 3) + P(X_1 = 2)P(X_2 = 2) + P(X_1 = 3)P(X_2 = 1)$
$= 10521/40000$

$p_5 = P(X = 5) = P(X_1 = 2) P(X_2 = 3) + P(X_1 = 3)P(X_2 = 2) = 567/1250$

$p_6 = P(X = 6) = P(X_1 = 3)P(X_2 = 3) = 1701/8000.$

Vậy luật phân phối của X là:

X	0	1	2	3	4	5	6
P	1/120000	1/2500	291/40000	473/7500	10521/40000	576/1250	1701/8000

b) Vì $X = X_1 + X_2$ và X_1, X_2 độc lập nên ta có:

- Kỳ vọng của X là
$$M(X) = M(X_1) + M(X_2) = n_1p_1 + n_2 p_2 = 4,8 \text{ (với } p_2 = N_{2A}/N_2)$$
- Phương sai của X là
$$D(X) = D(X_1) + D(X_2) = n_1p_1q_1 + n_2 p_2q_2(N_2-n_2)/(N_2-1) = 0,76.$$

Bài 2.14: Cho hai hộp I và II, mỗi hộp có 10 bi; trong đó hộp I gồm 8 bi đỏ, 2 bi trắng và hộp II gồm 6 bi đỏ, 4 bi trắng. Rút ngẫu nhiên từ hộp I hai bi bỏ sang hộp II, sau đó rút ngẫu nhiên từ hộp II ba bi.

a) Tính xác suất để được cả 3 bi trắng.

b) Gọi X là đại lượng ngẫu nhiên chỉ số bi trắng có trong ba bi được rút ra từ hộp II. Tìm luật phân phối của X. Xác định kỳ vọng và phương sai của X.

Lời giải

Gọi X là ĐLNN chỉ số bi trắng có trong 3 bi rút ra từ hộp II.

A_i (i = 0, 1, 2) là biến cố có i bi trắng và (2-i) bi đỏ có trong 2 bi lấy ra từ hộp I. Khi đó A_0, A_1, A_2 là hệ biến cố đầy đủ, xung khắc từng đôi và ta có:

$$P(A_0) = \frac{C_2^0 C_8^2}{C_{10}^2} = \frac{28}{45};$$
$$P(A_1) = \frac{C_2^1 C_8^1}{C_{10}^2} = \frac{16}{45};$$
$$P(A_2) = \frac{C_2^2 C_8^0}{C_{10}^2} = \frac{1}{45}.$$

Với mỗi k = 0, 1, 2, 3 theo công thức xác suất đầy đủ, ta có

$$P(X = k) = P(A_0)P(X = k/A_0) + P(A_1)P(X = k/A_1) + P(A_2)P(X = k/A_2)$$

a) Xác suất để được cả ba bi trắng là:

$$P(X = 3) = P(A_0)P(X = 3/A_0) + P(A_1)P(X = 3/A_1) + P(A_2)P(X = 3/A_2)$$

Mà

$$P(X = 3/A_0) = \frac{C_4^3 C_8^0}{C_{12}^3} = \frac{4}{220};$$
$$P(X = 3/A_1) = \frac{C_5^3 C_7^0}{C_{12}^3} = \frac{10}{220};$$
$$P(X = 3/A_2) = \frac{C_6^3 C_6^0}{C_{12}^3} = \frac{20}{220}.$$

nên P(X= 3) = 73/2475.

b) Luật phân phối của X có dạng:

X	0	1	2	3
P	p_0	p_1	p_2	p_3

trong đó, tương tự như trên ta có:

$$p_0 = P(X = 0) = \frac{28}{45}.\frac{C_4^0 C_8^3}{C_{12}^3} + \frac{16}{45}.\frac{C_5^0 C_7^3}{C_{12}^3} + \frac{1}{45}.\frac{C_6^0 C_6^3}{C_{12}^3} = 179/825;$$
$$p_1 = P(X = 1) = \frac{28}{45}.\frac{C_4^1 C_8^2}{C_{12}^3} + \frac{16}{45}.\frac{C_5^1 C_7^2}{C_{12}^3} + \frac{1}{45}.\frac{C_6^1 C_6^2}{C_{12}^3} = 223/450;$$
$$p_2 = P(X = 2) = \frac{28}{45}.\frac{C_4^2 C_8^1}{C_{12}^3} + \frac{16}{45}.\frac{C_5^2 C_7^1}{C_{12}^3} + \frac{1}{45}.\frac{C_6^2 C_6^1}{C_{12}^3} = 1277/4950;$$

p_3 = P(X= 3) = 73/2475.

Suy ra luật phân phối của X là:

X	0	1	2	3
P	179/825	223/450	1277/4950	73/2475

Từ đó suy ra kỳ vọng của X là M(X) = 1,1 và phương sai của X là D(X) = 0,5829.

Bài 2.15: Có ba lô sản phẩm, mỗi lô có 20 sản phẩm. Lô thứ i có i+4 sản phẩm loại A (i = 1, 2, 3).

a) Chọn ngẫu nhiên một lô rồi từ lô đó lấy ra 3 sản phẩm. Tính xác suất để trong 3 sản phẩm được lấy ra có đúng 1 sản phẩm loại A.

b) Từ mỗi lô lấy ra 1 sản phẩm. Gọi X là tổng số sản phẩm loại A có trong 3 sản phẩm được lấy ra. Tìm luật phân phối của X và tính Mod(X), M(X), D(X).

Lời giải

a) Gọi C là biến cố trong 3 sản phẩm được lấy ra có đúng 1 sản phẩm loại A.

Gọi A_1, A_2, A_3 lần lượt là các biến cố chọn được lô I, II, III. Khi đó A_1, A_2, A_3 là một hệ đầy đủ, xung khắc từng đôi và $P(A_1) = P(A_2) = P(A_3) = 1/3$.

Theo công thức xác suất đầy đủ, ta có:

$$P(C) = P(A_1)P(C/A_1) + P(A_2)P(C/A_2) + P(A_3)P(C/A_3)$$

Theo Công thức xác suất lựa chọn:

$$P(C/A_1) = \frac{C_5^1 C_{15}^2}{C_{20}^3} = \frac{525}{1140};$$
$$P(C/A_2) = \frac{C_6^1 C_{14}^2}{C_{20}^3} = \frac{546}{1140};$$
$$P(C/A_3) = \frac{C_7^1 C_{13}^2}{C_{20}^3} = \frac{546}{1140}.$$

Suy ra P(C)= 0,4728.

b) Luật phân phối của X có dạng:

X	0	1	2	3
P	p_0	p_1	p_2	p_3

Gọi B_j (j = 1, 2, 3) là biến cố lấy được sp loại A từ lô thứ j. Khi đó B_1, B_2, B_3 độc lập và

$$P(B_1) = \frac{5}{20}; P(\bar{B}_1) = \frac{15}{20};$$
$$P(B_2) = \frac{6}{20}; P(\bar{B}_2) = \frac{14}{20};$$
$$P(B_3) = \frac{7}{20}; P(\bar{B}_3) = \frac{13}{20}.$$

Ta có

$-"X = 0" = \bar{B}_1\bar{B}_2\bar{B}_3 \Rightarrow P(X = 0) = P(\bar{B}_1)P(\bar{B}_2)p(\bar{B}_3) = 273/800$

$-"X = 1" = B_1\bar{B}_2\bar{B}_3 + \bar{B}_1B_2\bar{B}_3 + \bar{B}_1\bar{B}_2B_3 \Rightarrow$

$P(X = 1) = P(B_1)P(\bar{B}_2)P(\bar{B}_3) + P(\bar{B}_1)P(B_2)P(\bar{B}_3) + P(\bar{B}_1)P(\bar{B}_2)P(B_3) = 71/160$

$-"X = 2" = B_1B_2\bar{B}_3 + B_1\bar{B}_2B_3 + \bar{B}_1B_2B_3 \Rightarrow$

$P(X = 2) = P(B_1)P(B_2)P(\bar{B}_3) + P(B_1)P(\bar{B}_2)P(B_3) + P(\bar{B}_1)P(B_2)P(B_3) = 151/800$

$-"X = 3" = B_1B_2B_3 \Rightarrow P(X = 3) = P(B_1)P(B_2)P(B_3) = 21/800$

Vậy luật phân phối của X là

X	0	1	2	3
P	273/800	71/160	151/800	21/800

Từ luật phânphối của X ta suy ra mode, kỳ vọng và phương sai của X :

- Mode: Mod(X) = 1.
- Kỳ vọng: M(X) = 0,9.
- Phương sai: D(X) = 0,625.

2.16: Một người có 5 chìa khóa bề ngoài rất giống nhau, trong đó chỉ có 2 chìa mở được cửa. Người đó tìm cách mở cửa bằng cách thử từng chìa một cho đến khi mở được cửa thì thôi (tất nhiên, chìa nào không mở được thì loại ra). Gọi X là số chìa khóa người đó sử dụng. Tìm luật phân phối của X. Hỏi người đó thường phải thử bao nhiêu chìa mới mở được cửa? Trung bình người đó phải thử bao nhiêu chìa mới mở được cửa?

Lời giải

Ta thấy X là ĐLNN rời rạc nhận 4 giá trị: 1, 2, 3, 4. Luật phân phối của X có dạng:

X	1	2	3	4
P	p_1	p_2	p_3	p_4

Gọi A_j (j = 1,2, 3, 4) là biến cố chìa khóa chọn lần thứ j mở được cửa. Khi đó:

P(X=1) = P(A_1) = 2/5.

$P(X = 2) = P(\bar{A}_1A_2) = P(\bar{A}_1)P(A_2/\bar{A}_1) = (3/5)(2/4) = 3/10;$

$P(X = 3) = P(\bar{A}_1\bar{A}_2A_3) = P(\bar{A}_1)P(\bar{A}_2/\bar{A}_1)P(A_3/\bar{A}_1\bar{A}_2) = (3/5)(2/4)(2/3) = 1/5$

$$P(X = 4) = P(\bar{A}_1\bar{A}_2\bar{A}_3A_4) = P(\bar{A}_1)P(\bar{A}_2/\bar{A}_1)P(\bar{A}_3/\bar{A}_1\bar{A}_2)P(A_4/\bar{A}_1\bar{A}_2\bar{A}_3)$$
$$= (3/5)(2/4)(1/3)(2/2) = 1/10$$

Vậy luật phân phối của X là:

X	1	2	3	4
P	2/5	3/10	1/5	1/10

Từ luật phân phối trên ta suy ra:

- Mode của X là Mod(X) = 1.
- Kỳ vọng của X là $M(X) = \sum x_i p_i = 2$.

Vậy người đó thường phải thử 1 chìa thì mở được cửa. Trung bình người đó phải thử 2 chìa mới mở được cửa.

Bài 2.17: Một người thợ săn có 5 viên đạn. Người đó đi săn với nguyên tắc: nếu bắn trúng mục tiêu thì về ngay, không đi săn nữa. Biết xác suất trúng đích của mỗi viên đạn bắn ra là 0,8. Gọi X là đại lượng ngẫu nhiên chỉ số viên đạn người ấy sử dụng trong cuộc săn.

a) Tìm luật phân phối của X.

b) Tìm kỳ vọng và phương sai của X.

Lời giải

a) Ta thấy X là ĐLNN rời rạc nhận 5 giá trị: 1, 2,..., 5. Luật phân phối của X có dạng:

X	1	2	3	4	5
P	p_1	p_2	p_3	p_4	p_5

Gọi A_j (j = 1,2,..., 5) là biến cố viên đạn thứ j trúng đích. Khi đó:

$$P(A_j) = 0,8; P(\bar{A}_j) = 0,2$$

Ta có:

$P(X=1) = P(A_1) = 0,8.$

$P(X = 2) = P(\bar{A}_1 A_2) = P(\bar{A}_1)P(A_2) = 0,2.0,8 = 0,16;$

$P(X = 3) = P(\bar{A}_1\bar{A}_2 A_3) = P(\bar{A}_1)P(\bar{A}_2)P(A_3) = 0,2.0,2.0,8 = 0,032;$

$P(X = 4) = P(\bar{A}_1\bar{A}_2\bar{A}_3 A_4) = P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)P(A_4) = 0,2.0,2.0,2.0,8 = 0,0064;$

$P(X = 5) = P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)P(\bar{A}_4) = 0,2.0,2.0,2.0,2 = 0,0016.$

Vậy luật phân phối của X là:

X	1	2	3	4	5
P	0,8	0,16	0,032	0,0064	0,0016

b) Từ luật phân phối của X ta suy ra:

- Kỳ vọng của X là M(X) = 1,2496.
- Phương sai của X là D(X) = 0,3089.

Bài 2.18: Một người thợ săn có 4 viên đạn. Người đó đi săn với nguyên tắc: nếu bắn 2 viên trúng mục tiêu thì về ngay, không đi săn nữa. Biết xác suất trúng đích của mỗi viên đạn bắn ra là 0,8. Gọi X là đại lượng ngẫu nhiên chỉ số viên đạn người ấy sử dụng trong cuộc săn.

a) Tìm luật phân phối của X.

b) Tìm kỳ vọng và phương sai của X.

Lời giải

a) Ta thấy X là ĐLNN rời rạc nhận 3 giá trị: 2, 3, 4. Luật phân phối của X có dạng:

X	2	3	4
P	p_2	p_3	p_4

Gọi A_j (j = 1,2, 3, 4) là biến cố viên đạn thứ j trúng đích. Khi đó:

$$P(A_j) = 0,8; P(\bar{A}_j) = 0,2$$

Ta có:

$P(X = 2) = P(A_1 A_2) = P(A_1)P(A_2) = 0,8.0,8 = 0,64;$

$P(X = 3) = P(\bar{A}_1 A_2 A_3 + A_1\bar{A}_2 A_3) = P(\bar{A}_1 A_2 A_3) + P(A_1\bar{A}_2 A_3)$

$= P(\bar{A}_1)P(A_2)P(A_3) + P(A_1)P(\bar{A}_2)P(A_3) = 0,2.0,8.0,8 + 0,8.0,2.0,8 = 0,256$

$P(X = 4) = P(\bar{A}_1\bar{A}_2\bar{A}_3 + A_1\bar{A}_2\bar{A}_3 + \bar{A}_1 A_2\bar{A}_3 + \bar{A}_1\bar{A}_2 A_3)$

$= P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3) + P(A_1)P(\bar{A}_2)P(\bar{A}_3) + P(\bar{A}_1)P(A_2)P(\bar{A}_3) + P(\bar{A}_1)P(\bar{A}_2)P(A_3)$

$= 0,2.0,2.0,2 + 0,8.0,2.0,2 + 0,2.0,8.0,2 + 0,2.0,2.0,8 = 0,104$

Vậy luật phân phối của X là:

X	2	3	4
P	0,64	0,256	0,104

b) Từ luật phân phối của X ta suy ra:

- Kỳ vọng của X là M(X) = 2,464.
- Phương sai của X là D(X) = 0,456704.

BÀI GIẢI
XÁC SUẤT THỐNG KÊ
(GV: Trần Ngọc Hội – 2009)

CHƯƠNG 4

KIỂM ĐỊNH GIẢ THIẾT

Bài 4.1. Trọng lượng của một sản phẩm theo qui định là 6kg. Sau một thời gian sản xuất, người ta tiến hành kiểm tra 121 sản phẩm và tính được trung bình mẫu là 5,975kg và phương sai mẫu hiệu chỉnh 5,7596kg^2. Sản xuất được xem là bình thường nếu các sản phẩm có trọng lượng trung bình bằng trọng lượng qui định. Với mức ý nghĩa 5%, hãy kết luận về tình hình sản xuất.

Lời giải

Gọi X là trọng lượng của một sản phẩm. Giả thiết cho ta:

- Cỡ mẫu n = 121.
- Kỳ vọng mẫu của X là $\overline{X} = 5,975$ (kg).
- Phương sai mẫu hiệu chỉnh của X là $S^2 = 5,7596(kg^2)$.
- Độ lệch mẫu hiệu chỉnh của X là $S = 2,3999(kg)$.

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$\mathbf{H_0}: \mu = 6 \text{ với giả thiết đối } \mathbf{H_1}: \mu \neq 6.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X}-\mu_0)\sqrt{n}}{S} = \frac{(5,975-6)\sqrt{121}}{2,3999} = -0.1146.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1-\alpha)/2 = 0,95/2 = 0,475$$

ta được $z_\alpha = 1,96$.

Bước 3: Kiểm định. Vì $|z| = 0,1146 < 1,96 = z_\alpha$ nên ta chấp nhận giả thiết H_0: $\mu = 6$.

Kết luận: Với mức ý nghĩa 5%, tình hình sản xuất được xem là bình thường.

Bài 4.2. Trọng lượng của một sản phẩm có phân phối chuẩn với trọng lượng trung bình là 500g. Sau một thời gian sản xuất, người ta nghi ngờ trọng lượng trung bình của loại sản phẩm này có xu hướng giảm nên tiến hành kiểm tra 25 sản phẩm và thu được kết quả sau:

Trọng lượng (g)	480	485	490	495	500	510
Số sản phẩm	2	3	8	5	3	4

Với mức ý nghĩa 3%, hãy kết luận điều nghi ngờ trên có đúng hay không.

Lời giải

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 3\% = 0,03$:

$$\mathbf{H_0}: \mu = 500 \text{ với giả thiết đối } \mathbf{H_1}: \mu < 500.$$

Ta có:

Xi	480	485	490	495	500	510
ni	2	3	8	5	3	4

$$n = 25; \ \sum X_i n_i = 12350; \ \sum X_i^2 n_i = 6102800.$$

- Kỳ vọng mẫu của X là

$$\overline{X} = \frac{1}{n}\sum X_i n_i = 494(g).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \overline{X}^2 = (8,7178)^2(g^2).$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1}\hat{S}^2 = (8,8976)^2(g^2).$$

Vì $n < 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X}-\mu_0)\sqrt{n}}{S} = \frac{(494-500)\sqrt{25}}{8,8976} = -3,3717.$$

Bước 2: Đặt $k = n - 1 = 24$. Tra bảng phân phối Student ứng với k = 24 và $2\alpha = 0,06$ ta được $t_{2\alpha} = 1,974$.

Bước 3: Kiểm định. Vì $-z = 3,3717 > 1,974 = t_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu = 500$, nghĩa là chấp nhận H_1: $\mu < 500$.

Kết luận: Với mức ý nghĩa 3%, điều nghi ngờ trọng lượng trung bình của loại sản phẩm này có xu hướng giảm là đúng.

Bài 4.3. Năng suất lúa trung bình của những vụ trước là 5,5tấn/ha. Vụ lúa năm nay người ta áp dụng một phương pháp kỹ thuật mới cho toàn bộ diện tích trồng lúa trong vùng. Điều tra năng suất 100ha lúa, ta có bảng số liệu sau:

Năngsuất (tạ/ha)	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
Diện tích (ha)	7	12	18	27	20	8	5	3

Với mức ý nghĩa 1%, hãy kết luận xem phương pháp kỹ thuật mới có làm tăng năng suất lúa trung bình của vùng này hay không?

Lời giải

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 1\% = 0{,}01$:

$\mathbf{H_0}$: $\mu = 55$ với giả thiết đối $\mathbf{H_1}$: $\mu > 55$.

(5,5tấn = 55tạ).

Ta có:

X_i	42,5	47,5	52,5	57,5	62,5	67,5	72,5	77,5
n_i	7	12	18	27	20	8	5	3

$$n = 100;\ \sum X_i n_i = 5750;\ \sum X_i^2 n_i = 337475.$$

- Kỳ vọng mẫu của X là

$$\overline{X} = \frac{1}{n}\sum X_i n_i = 57,5(\text{tạ}).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \overline{X}^2 = (8,2765)^2(\text{tạ}^2).$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1}\hat{S}^2 = (8,3182)^2(\text{tạ}^2).$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S} = \frac{(57,5 - 55)\sqrt{100}}{8,3182} = 3,0055.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2{,}33$.

Bước 3: Kiểm định. Vì $z = 3{,}0055 > 2{,}33 = z_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu = 55$, nghĩa là chấp nhận H_1: $\mu > 55$.

Kết luận: Với mức ý nghĩa 1%, phương pháp kỹ thuật mới làm tăng năng suất lúa trung bình của vùng này.

Bài 4.4. Một công ty dự định mở một siêu thị tại một khu dân cư. Để đánh giá khả năng mua hàng của dân cư trong khu vực, người ta tiến

hành điều tra về thu nhập của 100 hộ trong khu vực và có bảng số liệu sau:

Thu nhập bình quân (ngàn/người/tháng)	150	200	250	300	350
Số hộ	8	15	38	22	17

Theo bộ phận tiếp thị thì siêu thị chỉ hoạt động có hiệu quả tại khu vực này khi thu nhập bình quân hàng tháng của các hộ tối thiểu là vào khoảng 250ngàn/người/tháng. Vậy theo kết quả điều tra trên, công ty có nên quyết định mở siêu thị tại khu vực này hay không với mức ý nghĩa 5%?

Lời giải

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0{,}05$:

$\mathbf{H_0}$: $\mu = 250$ với giả thiết đối $\mathbf{H_1}$: $\mu > 250$.

Ta có:

Xi	150	200	250	300	350
ni	8	15	38	22	17

$$n = 100;\ \sum X_i n_i = 26250;\ \sum X_i^2 n_i = 7217500.$$

- Kỳ vọng mẫu của X là

$$\overline{X} = \frac{1}{n}\sum X_i n_i = 262,5.$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \overline{X}^2 = (57,1730)^2.$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1}\hat{S}^2 = (57,4610)^2.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S} = \frac{(262,5 - 250)\sqrt{100}}{57,4610} = 2,1754.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1{,}65$.

Bước 3: Kiểm định. Vì $z = 2{,}1754 > 1{,}65 = z_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu = 250$, chấp nhận giả thiết H_1: $\mu > 250$, nghĩa là thu nhập bình quân của các hộ cao hơn 250ngàn/người/tháng.

Kết luận: Với mức ý nghĩa 5%, công ty nên quyết định mở siêu thị tại khu vực này.

Bài 4.5. Để nghiên cứu nhu cầu của một loại hàng, người ta tiến hành khảo sát nhu cầu của mặt hàng này ở 400 hộ. Kết quả như sau:

Nhu cầu (kg/tháng)	0	0-1	1-2	2-3	3-4	4-5	5-6	6-7
Số hộ	10	35	86	132	78	31	18	10

Giả sử khu vực đó có 4000 hộ. Nếu cho rằng nhu cầu trung bình về mặt hàng này của toàn khu vực là 14tấn/tháng thì có chấp nhận được không với mức ý nghĩa 2%?

Lời giải

Khi cho rằng nhu cầu trung bình về mặt hàng này của toàn khu vực là 14tấn/tháng, nghĩa là nhu cầu trung bình về mặt hàng này của một hộ trong một tháng là

$$\frac{14\text{tấn}}{4000} = \frac{14000\text{kg}}{4000} = 3{,}5\text{kg}.$$

Do đó đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 2\% = 0{,}02$:

$\mathbf{H_0}$: $\mu = 3{,}5$ với giả thiết đối $\mathbf{H_1}$: $\mu \neq 3{,}5$.

Ta có:

Xi	0	0,5	1,5	2,5	3,5	4,5	5,5	6,5
ni	10	35	86	132	78	31	18	10

$$n = 400;\ \sum X_i n_i = 1053;\ \sum X_i^2 n_i = 3577{,}5.$$

- Kỳ vọng mẫu của X là
$$\overline{X} = \frac{1}{n}\sum X_i n_i = 2{,}6325.$$
- Phương sai mẫu của X là:
$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \overline{X}^2 = (1{,}4190)^2.$$
- Phương sai mẫu hiệu chỉnh của X là:
$$S^2 = \frac{n}{n-1}\hat{S}^2 = (1{,}4208)^2.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có
$$z = \frac{(\overline{X}-\mu_0)\sqrt{n}}{S} = \frac{(2{,}6325-3{,}5)\sqrt{400}}{1{,}4208} = -12{,}2114.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả
$$\varphi(z_\alpha) = (1-\alpha)/2 = 0{,}98/2 = 0{,}49$$
ta được $z_\alpha = 2{,}33$.

Bước 3: Kiểm định. Vì $|z| = 12{,}2114 > 2{,}33 = z_\alpha$ nên ta bác bỏ giả thiết H_0: $\mu = 3{,}5$, chấp nhận giả thiết H_1: $\mu \neq 3{,}5$.

Kết luận: Với mức ý nghĩa 5%, không thể cho rằng nhu cầu trung bình về mặt hàng này của toàn khu vực là 14tấn/tháng.

Bài 4.6. Trọng lượng của một loại gà công nghiệp ở một trại chăn nuôi có phân phối chuẩn. Trọng lượng trung bình khi xuất chuồng năm trước là 2,8kg/con. Năm nay, người ta sử dụng một loại thức ăn mới. Cân thử 25 con khi xuất chuồng người ta tính được trung bình mẫu là 3,2kg và phương sai mẫu hiệu chỉnh 0,25kg^2.

a) Với mức ý nghĩa 5%, hãy kết luận xem loại thức ăn mới có thực sự làm tăng trọng lượng trung bình của đàn gà hay không?

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

Lời giải

Gọi X là trọng lượng của một con gà sau khi sử dụng loại thức ăn mới. Giả thiết cho ta:

- X có phân phối chuẩn.
- Cỡ mẫu n = 25.
- Kỳ vọng mẫu của X là $\overline{X} = 3{,}2$(kg).
- Phương sai mẫu hiệu chỉnh của X là $S^2 = 0{,}25(\text{kg}^2)$.
- Độ lệch mẫu hiệu chỉnh của X là $S = 0{,}5$(kg).

a) Với mức ý nghĩa 5%, hãy kết luận xem loại thức ăn mới có thực sự làm tăng trọng lượng trung bình của đàn gà hay không?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0{,}05$:

$\mathbf{H_0}$: $\mu = 2{,}8$ với giả thiết đối $\mathbf{H_1}$: $\mu > 2{,}8$.

Vì $n < 30$; X có phân phối chuẩn; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có
$$z = \frac{(\overline{X}-\mu_0)\sqrt{n}}{S} = \frac{(3{,}2-2{,}8)\sqrt{25}}{0{,}5} = 4.$$

Bước 2: Đặt $k = n - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0{,}1$ ta được $t_{2\alpha} = t_{2\alpha}^k = 1{,}711$.

Bước 3: Kiểm định. Vì $z = 4 > 1{,}711 = t_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu = 2{,}8$, ghĩa là chấp nhận giả thiết H_1: $\mu > 2{,}8$.

Kết luận: Với mức ý nghĩa 5%, loại thức ăn mới thực sự làm tăng trọng lượng trung bình của đàn gà.

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

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$\mathbf{H_0}: \mu = 3,3 \text{ với giả thiết đối } \mathbf{H_1}: \mu \neq 3,3.$$

Vì $n < 30$; X có phân phối chuẩn; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S} = \frac{(3,2 - 3,3)\sqrt{25}}{0,5} = -1.$$

Bước 2: Đặt $k = n - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $\alpha = 0,05$ ta được $t_\alpha = t_\alpha^k = 2,064$.

Bước 3: Kiểm định. Vì $|z| = 1 < 2,064 = t_\alpha$ nên ta chấp nhận giả thiết H_0: $\mu = 3,3$.

Kết luận: Với mức ý nghĩa 5%, báo cáo trọng lượng trung bình khi xuất chuồng là 3,3kg/con là chấp nhận được.

Bài 4.7. Chiều cao trung bình của 100 nam sinh lớp 12 ở một trường trung học nội thành là 1,68m với độ lệch mẫu hiệu chỉnh 6cm. Trong khi kiểm tra 120 nam sinh lớp 12 ở một huyện ngoại thành thì chiều cao trung bình là 1,64m với độ lệch mẫu hiệu chỉnh 5cm. Với mức ý nghĩa 1%, có thể kết luận rằng nam sinh nội thành thực sự cao hơn nam sinh ngoại thành hay không?

Lời giải

Gọi X, Y (cm) lần lượt là chiều cao của nam sinh nội thành và nam sinh ngoại thành. Bài toán trên chính là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$\mathbf{H_0}: \mu_X = \mu_Y \text{ với giả thiết đối } \mathbf{H_1}: \mu_X > \mu_Y.$$

1) Đối với X, giả thiết cho ta:

- Cỡ mẫu $n_X = 100$.
- Kỳ vọng mẫu của X là $\overline{X} = 168(\text{cm})$.
- Độ lệch mẫu hiệu chỉnh của X là $S_X = 6(\text{cm})$.

2) Đối với Y, giả thiết cho ta:

- Cỡ mẫu $n_Y = 120$
- Kỳ vọng mẫu của Y là $\overline{Y} = 164(\text{cm})$.
- Độ lệch mẫu hiệu chỉnh của Y là $S_Y = 5(\text{cm})$.

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = \frac{168 - 164}{\sqrt{\frac{6^2}{100} + \frac{5^2}{120}}} = 5,3059.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 5,3059 > 2,33 = z_{2\alpha}$ nên ta bác bỏ giả thiết $\mathbf{H_0}$: $\mu_X = \mu_Y$, nghĩa là chấp nhận $\mathbf{H_1}$: $\mu_X > \mu_Y$.

Kết luận: Với mức ý nghĩa 1%, có thể kết luận rằng nam sinh nội thành thực sự cao hơn nam sinh ngoại thành.

Bài 4.8. Một hợp tác xã trồng thử hai giống lúa, mỗi giống trên 30 thửa ruộng và được chăm sóc như nhau. Cuối vụ thu hoạch ta được số liệu như sau:

	Năng suất trung bình (tạ/ha)	Độ lệch mẫu hiệu chỉnh
Giống lúa 1	45	2,5
Giống lúa 2	46,5	4,0

a) Với mức ý nghĩa 2%, có thể xem năng suất của hai giống lúa trên là như nhau hay không?

b) Với mức ý nghĩa 2%, có thể xem năng suất của giống lúa 2 cao hơn của giống lúa 1 hay không?

Lời giải

Gọi X, Y (tạ/ha) lần lượt là năng suất của giống lúa 1 và 2. Khi đó:

1) Đối với X, giả thiết cho ta:

- Cỡ mẫu $n_X = 30$.
- Kỳ vọng mẫu của X là $\overline{X} = 45$.
- Độ lệch mẫu hiệu chỉnh của X là $S_X = 2,5$.

2) Đối với Y, giả thiết cho ta:

- Cỡ mẫu $n_Y = 30$.
- Kỳ vọng mẫu của Y là $\overline{Y} = 46,5$.
- Độ lệch mẫu hiệu chỉnh của Y là $S_Y = 4$.

a) Với mức ý nghĩa 2%, có thể xem năng suất của hai giống lúa trên là như nhau hay không?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa 2% = 0,02:

$$\mathbf{H_0}: \mu_X = \mu_Y \text{ với giả thiết đối } \mathbf{H_1}: \mu_X \neq \mu_Y.$$

Vì $n_X = n_Y = 30$ nên ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = \frac{45 - 46,5}{\sqrt{\frac{2,5^2}{30} + \frac{4^2}{30}}} = -1,7418.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,98/2 = 0,49$$

ta được $z_\alpha = 2,33$.

Bước 3: Kiểm định. Vì $|z| = 1,7418 < 2,33 = z_\alpha$ nên ta chấp nhận giả thiết $\mathbf{H_0}$: $\mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 2%, có thể xem năng suất của hai giống lúa trên là như nhau.

b) Với mức ý nghĩa 2%, có thể xem năng suất của giống lúa 2 cao hơn của giống lúa 1 hay không?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa α = 2% = 0,02:

$$\mathbf{H_0}: \mu_X = \mu_Y \text{ với giả thiết đối } \mathbf{H_1}: \mu_X < \mu_Y.$$

Vì $n_X = n_Y = 30$ nên ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = -1,7418.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$$

ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $-z = 1,7418 < 2,06 = z_{2\alpha}$ nên ta chấp nhận giả thiết $\mathbf{H_0}$: $\mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 2%, chưa thể xem năng suất của giống lúa 2 cao hơn của giống lúa 1.

Bài 4.9. Một máy sản xuất tự động, lúc đầu tỉ lệ sản phẩm loại A là 45%. Sau khi áp dụng một phương pháp sản xuất mới, người ta lấy ra 400 sản phẩm để kiểm tra thì thấy có 215 sản phẩm loại A. Với mức ý nghĩa 5%, hãy kết luận xem phương pháp mới có thực sự làm tăng tỉ lệ sản phẩm loại A hay không?

Lời giải

Từ giả thiết ta suy ra:

- Cỡ mẫu n = 400.
- Số sản phẩm loại A có trong mẫu là m = 215.
- Tỉ lệ mẫu sản phẩm loại A là $F_n = m/n = 215/400 = 0,5375$.

Ta đưa bài toán về bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại A với mức ý nghĩa α = 5% = 0,05:

$$H_0: p = 45\% = 0,45 \text{ với giả thiết đối } H_1: p > 0,45.$$

Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0(1-p_0)}} = \frac{(0,5375 - 0,45)\sqrt{400}}{\sqrt{0,45(1-0,45)}} = 3,5176.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 3,5176 > 1,65 = z_{2\alpha}$ nên ta bác bỏ giả thiết H_0: p = 0,45, nghĩa là chấp nhận H_1: p > 0,45.

Kết luận: Với mức ý nghĩa 5%, phương pháp mới thực sự làm tăng tỉ lệ sản phẩm loại A.

Bài 4.10. Thống kê 10650 trẻ sơ sinh ở một địa phương người ta thấy có 5410 bé trai.

a) Với mức ý nghĩa 3%, hỏi có sự khác biệt về tỉ lệ sinh bé trai và bé gái hay không?

b) Với mức ý nghĩa 1%, hỏi tỉ lệ sinh bé trai có thực sự cao hơn tỉ lệ sinh bé gái hay không?

Lời giải

Từ các giả thiết của bài toán ta suy ra:

1) Khi khảo sát tỉ lệ bé trai p_1:
   - Cỡ mẫu $n_1 = 10650$.
   - Số bé trai là $m_1 = 5410$.
   - Tỉ lệ bé trai $F_{n1} = 5410/10650$.

2) Khi khảo sát tỉ lệ bé gái p_2:
   - Cỡ mẫu $n_2 = 10650$.
   - Số bé gái là $m_2 = 10650 - 5410 = 5240$.
   - Tỉ lệ bé gái $F_{n2} = 5240/10650$.

3) $p_0 = 0,5$.

a) Với mức ý nghĩa 3%, hỏi có sự khác biệt về tỉ lệ sinh bé trai và bé gái hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa α = 3% = 0,03:

$$\mathbf{H_0}: p_1 = p_2 \ (= p_0) \text{ với giả thiết đối } \mathbf{H_1}: p_1 \neq p_2$$

Ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\frac{5410}{10650} - \frac{5240}{10650}}{\sqrt{0,5(1-0,5)\left(\frac{1}{10650} + \frac{1}{10650}\right)}} = 2,3296.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,97/2 = 0,485$$

ta được $z_\alpha = 2,17$.

Bước 3: Kiểm định. Vì $|z| = 2,3296 > 2,17 = z_\alpha$ nên ta bác bỏ giả thiết H_0: $p_1 = p_2$, nghĩa là chấp nhận H_1: $p_1 \neq p_2$.

Kết luận: Với mức ý nghĩa 3%, có sự khác biệt về tỉ lệ sinh bé trai và bé gái.

b) Với mức ý nghĩa 1%, hỏi tỉ lệ sinh bé trai có thực sự cao hơn tỉ lệ sinh bé gái hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$\mathbf{H_0}$: $p_1 = p_2$ với giả thiết đối $\mathbf{H_1}$: $p_1 > p_2$

Ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 2,3296.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 2,3296 < 2,33 = z_{2\alpha}$ nên ta chấp nhận giả thiết H_0: $p_1 = p_2$.

Kết luận: Với mức ý nghĩa 1%, chưa thể nói tỉ lệ sinh bé trai thực sự cao hơn tỉ lệ sinh bé gái.

Bài 4.11. Bệnh A có thể chữa bằng hai loại thuốc H và K. Công ty sản xuất thuốc H tuyên bố tỉ lệ bệnh nhân khỏi bệnh do dùng thuốc của họ là 85%. Người ta dùng thử thuốc H chữa cho 250 bệnh nhân thì thấy có 210 người khỏi bệnh, dùng thử thuốc K cho 200 bệnh nhân thì thấy có 175 người khỏi bệnh.

a) Với mức ý nghĩa 1% có thể kết luận thuốc K có khả năng chữa bệnh A tốt hơn thuốc H hay không?

b) Xét xem hiệu quả chữa bệnh của thuốc H có đúng như công ty quảng cáo với mức ý nghĩa 5% hay không.

Lời giải

Từ các giả thiết của bài toán ta suy ra:

1) Đối với loại thuốc H:
- Cỡ mẫu $n_1 = 250$.
- Số bệnh nhân khỏi bệnh: 210.
- Tỉ lệ mẫu bệnh nhân khỏi bệnh $F_{n1} = 210/250 = 0,84$.

2) Đối với loại thuốc K:
- Cỡ mẫu $n_2 = 200$.
- Số bệnh nhân khỏi bệnh: 175.
- Tỉ lệ mẫu bệnh nhân khỏi bệnh $F_{n2} = 175/200 = 0,875$.

3) $p_0 = \frac{n_1F_{n1} + n_2F_{n2}}{n_1 + n_2} = \frac{250.0,84 + 200.0,875}{250 + 200} = \frac{385}{450}$.

a) Với mức ý nghĩa 1% có thể kết luận thuốc K có khả năng chữa bệnh A tốt hơn thuốc H hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$\mathbf{H_0}$: $p_1 = p_2$ với giả thiết đối $\mathbf{H_1}$: $p_1 < p_2$

Ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0,84 - 0,875}{\sqrt{\frac{385}{450}\left(1 - \frac{385}{450}\right)\left(\frac{1}{250} + \frac{1}{200}\right)}} = -1,0495.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $-z = 1,0495 < 2,33 = z_{2\alpha}$ nên ta chấp nhận giả thiết H_0: $p_1 = p_2$.

Kết luận: Với mức ý nghĩa 1%, không thể kết luận thuốc K có khả năng chữa bệnh A tốt hơn thuốc H.

b) Xét xem hiệu quả chữa bệnh của thuốc H có đúng như công ty quảng cáo với mức ý nghĩa 5% hay không.

Đây là bài toán kiểm định giả thiết về tỉ lệ p_1 các bệnh nhân khỏi bệnh A khi được điều trị bằng thuốc H với mức ý nghĩa $\alpha = 5\% = 0,05$:

H_0: $p_1 = 85\% = 0,85$ với giả thiết đối H_1: $p_1 < 0,85$.

Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_{n1} - p_{01})\sqrt{n_1}}{\sqrt{p_{01}q_{01}}} = \frac{(0,84 - 0,85)\sqrt{250}}{\sqrt{0,85(1-0,85)}} = -0,4428.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $-z = 0,4428 < 1,65 = z_{2\alpha}$ nên ta chấp nhận giả thiết H_0: $p_1 = 0,85$.

Kết luận: Với mức ý nghĩa 5%, hiệu quả chữa bệnh của thuốc H đúng như công ty quảng cáo.

Bài 4.12. Để khảo sát chỉ tiêu X của một loại sản phẩm, người ta quan sát một mẫu và có kết qủa sau:

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

Những sản phẩm có chỉ tiêu X từ 19cm trở xuống được gọi là những sản phẩm loại B.

a) Giả sử trung bình tiêu chuẩn của chỉ tiêu X là 29cm. Hãy cho nhận xét về tình hình sản xuất với mức ý nghĩa 2%.

b) Bằng phương pháp sản xuất mới, sau một thời gian người ta thấy giá trị trung bình của chỉ tiêu X của những sản phẩm loại B là 16cm. Hãy cho kết luận về phương pháp sản xuất mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

c) Một tài liệu thống kê cũ cho rằng tỉ lệ những sản phẩm loại B là 12%. Hãy nhận định về tài liệu này với mức ý nghĩa 5%.

Lời giải

Lập bảng:

X_i	13	17	21	25	29	33	37
n_i	8	9	20	16	16	13	18

Ta có:

$$n = 100;\ \sum X_i n_i = 2636;\ \sum X_i^2 n_i = 75028.$$

- Kỳ vọng mẫu của X là

$$\overline{X} = \frac{1}{n}\sum X_i n_i = 26,36(\text{cm}).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \overline{X}^2 = (7,4452)^2(\text{cm}^2).$$

- Phương sai mẫu đã hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1}\hat{S}^2 = (7,4827)^2(\text{cm}^2).$$

a) Giả sử trung bình tiêu chuẩn của chỉ tiêu X là 29cm. Hãy cho nhận xét về tình hình sản xuất với mức ý nghĩa 2%.

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 2\% = 0,02$:

$$\mathbf{H_0}: \mu = 29 \text{ với giả thiết đối } \mathbf{H_1}: \mu \neq 29.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S} = \frac{(26,36 - 29)\sqrt{100}}{7,4827} = -3,5281.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,98/2 = 0,49$$

ta được $z_\alpha = 2,33$.

Bước 3: Kiểm định. Vì $|z| = 3,5281 > 2,33 = z_\alpha$ nên ta bác bỏ giả thiết H_0: $\mu=29$, nghĩa là chấp nhận H_1: $\mu \neq 29$.

Kết luận: Với mức ý nghĩa 1%, tình hình sản xuất không bình thường vì giá trị trung bình của chỉ tiêu X không đúng tiêu chuẩn.

b) Bằng phương pháp sản xuất mới, sau một thời gian người ta thấy giá trị trung bình của chỉ tiêu X của những sản phẩm loại B là 16cm. Hãy cho kết luận về phương pháp sản xuất mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_B = M(X_B)$ của chỉ tiêu $X = X_B$ của những sản phẩm loại B với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$\mathbf{H_0}: \mu_B = 16 \text{ với giả thiết đối } \mathbf{H_1}: \mu_B \neq 16.$$

Ta lập bảng số liệu của X_B:

X_{Bi}	13	17
n_{Bi}	8	9

Từ bảng trên ta tính được:

$$n_B = 17;\ \sum X_{Bi} n_{Bi} = 257;\ \sum X_{Bi}^2 n_{Bi} = 3,953.$$

- Kỳ vọng mẫu của X_B là

$$\overline{X}_B = \frac{1}{n}\sum X_{Bi} n_{Bi} = 15,1176\ (cm).$$

- Phương sai mẫu của X_B là:

$$\hat{S}_B^2 = \frac{1}{n}\sum X_{Bi}^2 n_{Bi} - \overline{X}_B^2 = (1,9965)^2(\text{cm}^2).$$

- Phương sai mẫu đã hiệu chỉnh của X_B là:

$$S_B^2 = \frac{n_B}{n_B - 1}\hat{S}_B^2 = (2,0580)^2\ (\text{cm}^2).$$

Vì $n_B < 30$, X_B có phân phối chuẩn, $\sigma^2_B = D(X_B)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X}_B - \mu_0)\sqrt{n_B}}{S_B} = \frac{(15{,}1176 - 16)\sqrt{17}}{2{,}0580} = -1{,}7678.$$

Bước 2: Đặt $k = n_B - 1 = 16$. Tra bảng phân phối Student ứng với $k = 16$ và $\alpha = 0{,}01$ ta được $t_\alpha = 2{,}921$.

Bước 3: Kiểm định. Vì $|z| = 1{,}7678 < 2{,}921 = t_\alpha$ nên ta chấp nhận giả thiết H_0: $\mu_B = 16$.

Kết luận: Với mức ý nghĩa 2%, phương pháp mới không có tác dụng làm thay đổi giá trị trung bình của chỉ tiêu X_B của các sản phẩm loại B.

c) Một tài liệu thống kê cũ cho rằng tỉ lệ sản phẩm loại B là 12%. Hãy nhận định về tài liệu này với mức ý nghĩa 5%.

Đây là bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại B với mức ý nghĩa $\alpha = 5\% = 0{,}05$:

H_0: $p = 12\% = 0{,}12$ với giả thiết đối H_1: $p \neq 0{,}12$.

Ta có qui tắc kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0 q_0}} = \frac{(0{,}17 - 0{,}12)\sqrt{100}}{\sqrt{0{,}12(1 - 0{,}12)}} = 1{,}5386.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0{,}95/2 = 0{,}475$$

ta được $z_\alpha = 1{,}96$.

Bước 3: Kiểm định. Vì $|z| = 1{,}5386 < 1{,}96 = z_\alpha$ nên ta chấp nhận giả thiết H_0: $p = 0{,}12$.

Kết luận: Với mức ý nghĩa 5%, tài liệu cũ về tỉ lệ sản phẩm loại B còn phù hợp.

Bài 4.13. Để khảo sát chiều cao X của một giống cây trồng, người ta quan sát một mẫu và có kết qủa sau:

X(cm)	95-105	105-115	115-125	125-135	135-145	145-155	155-165
Số cây	10	10	15	30	10	10	15

a) Một tài liệu thống kê cũ cho rằng chiều cao trung bình của giống cây trồng trên là 127cm. Hãy cho kết luận về tài liệu đó với mức ý nghĩa 1%.

b) Những cây trồng có chiều cao từ 135cm trở lên được gọi là những cây "cao". Trước đây, tỉ lệ những cây "cao" của loại cây trồng trên là 40%. Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy cho kết luận về kỹ thuật mới với mức ý nghĩa 5%.

c) Những cây trồng có chiều cao từ 105cm đến 125cm được gọi là những cây loại A. Bằng phương pháp mới, sau một thời gian người ta thấy chiều cao trung bình của những cây loại A là 119,5cm. Hãy cho kết luận về phương pháp mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

Lời giải

X_i	100	110	120	130	140	150	160
n_i	10	10	15	30	10	10	15

Ta có:

$$n = 100; \ \sum X_i n_i = 13100; \ \sum X_i^2 n_i = 1749000.$$

- Kỳ vọng mẫu của X là

$$\overline{X} = \frac{1}{n}\sum X_i n_i = 131(\text{cm}).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \overline{X}^2 = (18{,}1384)^2(\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1}\hat{S}^2 = (18{,}2297)^2(\text{cm}^2).$$

a) Một tài liệu thống kê cũ cho rằng chiều cao trung bình của giống cây trồng trên là 127cm. Hãy cho kết luận về tài liệu đó với mức ý nghĩa 1%.

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 1\% = 0{,}01$:

H_0: $\mu = 127$ với giả thiết đối H_1: $\mu \neq 127$

Vì $n \geq 30$; σ^2 chưa biết, nên ta có qui tắc kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X} - \mu_0)\sqrt{n}}{S} = \frac{(131 - 127)\sqrt{100}}{18{,}2297} = 2{,}1942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0{,}99/2 = 0{,}495$$

ta được $z_\alpha = 2{,}58$.

Bước 3: Kiểm định. Vì $|z| = 2{,}1942 < 2{,}58 = z_\alpha$ nên ta chấp nhận H_0: $\mu = 127$.

Kết luận: Với mức ý nghĩa 1%, tài liệu cũ về chiều cao trung bình của giống cây trồng trên còn phù hợp với thực tế.

b) Những cây trồng có chiều cao từ 135cm trở lên được gọi là những cây "cao". Trước đây, tỉ lệ những cây "cao" của loại cây trồng trên là 40%.

Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy cho kết luận về kỹ thuật mới với mức ý nghĩa 5%.

Đây là bài toán kiểm định giả thiết về tỉ lệ p các cây cao với mức ý nghĩa $\alpha = 5\% = 0,05$:

H_0: p = 40% = 0,4 với giả thiết đối H_1: $p \neq 0,4$

Ta có qui tắc kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,35-0,4)\sqrt{100}}{\sqrt{0,4(1-0,4)}} = -1,0206.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1-\alpha)/2 = 0,95/2 = 0,475$$

ta được $z_\alpha = 1,96$.

Bước 3: Kiểm định. Vì $|z| = 1,0206 < 1,96 = z_\alpha$ nên ta chấp nhận giả thiết H_0: p = 0,4.

Kết luận: Với mức ý nghĩa 5%, phương pháp mới không có tác dụng làm thay đổi tỉ lệ các cây cao.

c) Những cây trồng có chiều cao từ 105cm đến 125cm được gọi là những cây loại A. Bằng phương pháp mới, sau một thời gian người ta thấy chiều cao trung bình của những cây loại A là 119,5cm. Hãy cho kết luận về phương pháp mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_A = M(X_A)$ của chiều cao $X = X_A$ của các cây loại A với mức ý nghĩa $\alpha = 1\% = 0,01$:

H_0: $\mu_A = 119,5$ với giả thiết đối H_1: $\mu_A \neq 119,5$.

Ta lập bảng số liệu của X_A:

X_{Ai}	110	120
N_{Ai}	10	15

Từ bảng trên ta tính được:

$$n_A = 25;\ \sum X_{Ai}n_{Ai} = 2900;\ \sum X_{Ai}^2 n_{Ai} = 337000.$$

- Kỳ vọng mẫu của X_A là

$$\overline{X}_A = \frac{1}{n}\sum X_{Ai}n_{Ai} = 116(\text{cm}).$$

- Phương sai mẫu của X_A là:

$$\hat{S}_A^2 = \frac{1}{n}\sum X_{Ai}^2 n_{Ai} - \overline{X}_A^2 = (4,8990)^2(\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X_A là:

$$S_A^2 = \frac{n_A}{n_A - 1}\hat{S}_A^2 = 5^2(\text{cm}^2).$$

Vì $n_A = 25 < 30$, X_A có phân phối chuẩn, $\sigma^2_A = D(X_A)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\overline{X}_A - \mu_0)\sqrt{n_A}}{S_A} = \frac{(116-119,5)\sqrt{25}}{5} = -3,5.$$

Bước 2: Đặt $k = n_A - 1 = 24$. Tra bảng phân phối Student ứng với k = 24 và $\alpha = 0,01$ ta được $t_\alpha = t_\alpha^k = 2,797$.

Bước 3: Kiểm định. Vì $|z| = 3,5 > 2,797 = t_\alpha$ nên ta bác bỏ giả thiết H_0: $\mu_A = 119,5$, nghĩa là chấp nhận H_1: $\mu_A \neq 119,5$. Cụ thể, ta nhận định $\mu_A < 119,5$ (vì $\overline{X}_A = 116 < 119,5$).

Kết luận: Với mức ý nghĩa 1%, phương pháp mới có tác dụng làm thay đổi chiều cao trung bình của các cây loại A, theo hướng làm tăng chiều cao trung bình của các cây loại này.

Bài 4.14. Cho các số liệu như Bài 4.13.

a) Giả sử trung bình tiêu chuẩn của chiều cao X là 125cm. Có thể khẳng định rằng việc canh tác làm tăng chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 1% hay không?

b) Giả sử trung bình tiêu chuẩn của chiều cao X là 134cm. Có thể khẳng định rằng việc canh tác làm giảm chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 2% hay không?

c) Sau khi áp dụng phương pháp canh tác mới, người ta thấy chiều cao trung bình của các cây loại A là 114cm. Hãy kết luận xem phương pháp mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 3% (Giả sử X có phân phối chuẩn).

d) Trước đây, chiều cao trung bình của các cây loại A là 120cm. Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy kết luận xem kỹ thuật mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 2% (Giả sử X có phân phối chuẩn).

e) Sau khi áp dụng một phương pháp sản xuất, người ta thấy tỉ lệ cây loại A là 35%. Hãy kết luận xem phương pháp mới có làm tăng tỉ lệ cây loại A lên hay không với mức ý nghĩa 2% .

f) Một tài liệu thống kê cũ cho rằng tỉ lệ cây loại A là 20%. Hãy xét xem việc canh tác có làm tăng tỉ lệ cây loại A hay không với mức ý nghĩa 5%?

Lời giải

Ta có:

- Cỡ mẫu là n = 100.
- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n}\sum X_i n_i = 131(\text{cm}).$$

- Phương sai mẫu của X là

$$\hat{S}^2 = \frac{1}{n}\sum X_i^2 n_i - \bar{X}^2 = (18,1384)^2(\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X là

$$S^2 = \frac{n}{n-1}\hat{S}^2 = (18,2297)^2(\text{cm}^2).$$

a) Giả sử trung bình tiêu chuẩn của chiều cao X là 125cm. Có thể khẳng định rằng việc canh tác làm tăng chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 1% hay không?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 1\% = 0,01$:

H₀: $\mu = 125$ với giả thiết đối **H₁**: $\mu > 125$.

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X}-\mu_0)\sqrt{n}}{S} = \frac{(131-125)\sqrt{100}}{18,2297} = 3,2913.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả $\varphi(z_{2\alpha}) = (1-2\alpha)/2 = 0,98/2 = 0,49$ ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 3,2913 > 2,33 = z_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu=125$, nghĩa là chấp nhận H_1: $\mu > 125$.

Kết luận: Với mức ý nghĩa 1%, có thể kết luận rằng việc canh tác làm tăng chiều cao trung bình của giống cây trồng trên.

b) Giả sử trung bình tiêu chuẩn của chiều cao X là 134cm. Có thể khẳng định rằng việc canh tác làm giảm chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 2% hay không?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 2\% = 0,02$:

H₀: $\mu = 134$ với giả thiết đối **H₁**: $\mu < 134$.

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X}-\mu_0)\sqrt{n}}{S} = \frac{(131-134)\sqrt{100}}{18,2297} = -1,6457.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả $\varphi(z_{2\alpha}) = (1-2\alpha)/2 = 0,96/2 = 0,48$ ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $-z = 1,6457 < 2,06 = z_{2\alpha}$ nên ta chấp nhận giả thiết H_0: $\mu = 134$.

Kết luận: Với mức ý nghĩa 2%, không thể kết luận rằng việc canh tác làm giảm chiều cao trung bình của giống cây trồng trên.

c) Sau khi áp dụng phương pháp canh tác mới, người ta thấy chiều cao trung bình của các cây loại A là 114cm. Hãy kết luận xem phương pháp mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 3% (Giả sử X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_A = M(X_A)$ của chỉ tiêu $X = X_A$ của các cây loại A với mức ý nghĩa $\alpha = 3\% = 0,03$:

H₀: $\mu_A = 114$ với giả thiết đối **H₁**: $\mu_A > 114$.

Ta lập bảng số liệu của X_A:

X_{Ai}	110	120
N_{Ai}	10	15

Từ bảng trên ta tính được:

$$n_A = 25; \sum X_{Ai} n_{Ai} = 2900; \sum X_{Ai}^2 n_{Ai} = 337000.$$

- Kỳ vọng mẫu của X_A là

$$\bar{X}_A = \frac{1}{n}\sum X_{Ai} n_{Ai} = 116(\text{cm}).$$

- Phương sai mẫu của X_A là:

$$\hat{S}_A^2 = \frac{1}{n}\sum X_{Ai}^2 n_{Ai} - \bar{X}_A^2 = (4,8990)^2(\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X_A là:

$$S_A^2 = \frac{n_A}{n_A - 1}\hat{S}_A^2 = 5^2(\text{cm}^2).$$

Vì $n_A < 30$, X_A có phân phối chuẩn, $\sigma^2_A = D(X_A)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X}_A - \mu_0)\sqrt{n_A}}{S_A} = \frac{(116-114)\sqrt{25}}{5} = 2.$$

Bước 2: Đặt $k = n_A - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0,06$ ta được $t_{2\alpha} = 1,974$.

Bước 3: Kiểm định. Vì $z = 2 > 1,974 = t_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu_A = 114$, nghĩa là chấp nhận H_1: $\mu_A > 114$.

Kết luận: Với mức ý nghĩa 3%, phương pháp mới làm giảm chiều cao trung bình của các cây loại A.

d) Trước đây, chiều cao trung bình của các cây loại A là 120cm. Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy kết luận xem kỹ thuật mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 2% (Giả sử X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_A = M(X_A)$ của chỉ tiêu $X = X_A$ của các cây loại A với mức ý nghĩa $\alpha = 2\% = 0,02$:

$\mathbf{H_0}$: $\mu_A = 120$ với giả thiết đối $\mathbf{H_1}$: $\mu_A < 120$.

Vì $n_A < 30$, X_A có phân phối chuẩn, $\sigma^2_A = D(X_A)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có $z = \dfrac{(\bar{X}_A - \mu_0)\sqrt{n_A}}{S_A} = \dfrac{(116-120)\sqrt{25}}{5} = -4.$

Bước 2: Đặt $k = n_A - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0,04$ ta được $t_{2\alpha} = 2,1715$.

Bước 3: Kiểm định. Vì $-z = 4 > 2,1715 = t_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $\mu_A = 120$, nghĩa là chấp nhận H_1: $\mu_A < 120$.

Kết luận: Với mức ý nghĩa 2%, kỹ thuật mới làm giảm chiều cao trung bình của các cây loại A.

e) Sau khi áp dụng một phương pháp sản xuất, người ta thấy tỉ lệ cây loại A là 35%. Hãy kết luận xem phương pháp mới có làm tăng tỉ lệ cây loại A lên hay không với mức ý nghĩa 2% .

Đây là bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại A với mức ý nghĩa $\alpha = 2\% = 0,02$:

H_0: $p = 35\% = 0,35$ với giả thiết đối H_1: $p < 0,35$.

Ta có tỉ lệ mẫu các cây loại A là $F_n = 25/100 = 0,25$. Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,25-0,35)\sqrt{100}}{\sqrt{0,35(1-0,35)}} = -2,0966.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$$

ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $-z = 2,0966 > 2,06 = z_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $p = 0,35$, nghĩa là chấp nhận H_1: $p < 0,35$.

Kết luận: Với mức ý nghĩa 2%, phương pháp mới làm tăng tỉ lệ cây loại A.

f) Một tài liệu thống kê cũ cho rằng tỉ lệ cây loại A là 20%. Hãy xét xem việc canh tác có làm tăng tỉ lệ cây loại A hay không với mức ý nghĩa 5%?

Đây là bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại A với mức ý nghĩa $\alpha = 5\% = 0,05$:

H_0: $p = 20\% = 0,20$ với giả thiết đối H_1: $p > 0,20$.

Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,25-0,20)\sqrt{100}}{\sqrt{0,20(1-0,20)}} = 1,25.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 1,25 < 1,65 = z_{2\alpha}$ nên ta chấp nhận giả thiết H_0: $p = 0,20$.

Kết luận: Với mức ý nghĩa 5%, việc canh tác không làm tăng tỉ lệ các cây loại A.

Bài 4.15. Để khảo sát đường kính của một chi tiết máy người ta kiểm tra một số sản phẩm của hai nhà máy. Trong kết quả sau đây, X là đường kính của chi tiết máy do nhà máy 1 sản xuất còn Y là đường kính của chi tiết máy do nhà máy 2 sản xuất. Những sản phẩm có chi tiết máy nhỏ hơn 19cm được xếp vào loại C.

X(cm)	11-15	15-19	19-23	23-27	27-31	31-35	35-39
Số sản phẩm	9	19	20	26	16	13	18

Y(cm)	13-16	16-19	19-22	22-25	25-28	28-31	31-34
Số sản phẩm	7	9	25	26	18	15	11

a) Có thể kết luận rằng đường kính trung bình của một chi tiết máy do hai nhà máy sản xuất bằng nhau hay không với mức ý nghĩa 1%?

b) Có thể cho rằng đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất lớn hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất hay không với mức ý nghĩa 5%?

c) Xét xem đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất có nhỏ hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất hay không với mức ý nghĩa 2%?

d) Với mức ý nghĩa 4%, tỉ lệ sản phẩm loại C do hai nhà máy sản xuất có như nhau không?

e) Với mức ý nghĩa 3%, có thể cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất lớn hơn tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất hay không?

f) Hãy nhận xét về ý kiến cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất nhỏ hơn tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất với mức ý nghĩa 5%?

Lời giải

1) Đối với X ta có bảng số liệu:

Xi	13	17	21	25	29	33	37
ni	9	19	20	26	16	13	18

Ta có:

$$n_X = 121;\ \sum X_i n_{Xi} = 3069;\ \sum X_i^2 n_{Xi} = 84337.$$

- Kỳ vọng mẫu của X là
$$\overline{X} = \frac{1}{n_X}\sum X_i n_{Xi} = 25,3636(\text{cm}).$$
- Phương sai mẫu của X là
$$\hat{S}_X^2 = \frac{1}{n_X}\sum X_i^2 n_{Xi} - \overline{X}^2 = (7,3271)^2(\text{cm}^2).$$
- Phương sai mẫu hiệu chỉnh của X là
$$S_X^2 = \frac{n_X}{n_X - 1}\hat{S}_X^2 = (7,3575)^2(\text{cm}^2).$$
- Tỉ lệ sản phẩm loại C là
$$F_{Xn} = \frac{m_X}{n_X} = \frac{9+19}{121} = 0,2314.$$

2) Đối với Y ta có bảng số liệu:

Yi	14,5	17,5	20,5	23,5	26,5	29,5	32,5
ni	7	9	25	26	18	15	11

Ta có:

$$n_Y = 111;\ \sum Y_i n_{Yi} = 2659,5;\ \sum Y_i^2 n_{Yi} = 66405,75.$$

- Kỳ vọng mẫu của Y là
$$\overline{Y} = \frac{1}{n_Y}\sum Y_i n_{Yi} = 23,9595(\text{cm}).$$
- Phương sai mẫu của Y là
$$\hat{S}_Y^2 = \frac{1}{n_Y}\sum Y_i^2 n_{Yi} - \overline{Y}^2 = (4,9188)^2(\text{cm}^2).$$
- Phương sai mẫu hiệu chỉnh của Y là
$$S_Y^2 = \frac{n_Y}{n_Y - 1}\hat{S}_Y^2 = (4,9411)^2(\text{cm}^2).$$
- Tỉ lệ sản phẩm loại C là

$$F_{Yn} = \frac{m_Y}{n_Y} = \frac{7+9}{111} = 0,1441.$$

a) Có thể kết luận rằng đường kính trung bình của một chi tiết máy do hai nhà máy sản xuất bằng nhau hay không với mức ý nghĩa 1%?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa α = 1% = 0,01:

$$\mathbf{H_0}: \mu_X = \mu_Y \text{ với giả thiết đối } \mathbf{H_1}: \mu_X \neq \mu_Y.$$

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = \frac{25,3636 - 23,9595}{\sqrt{\frac{(7,3575)^2}{121} + \frac{(4,9411)^2}{111}}} = 1,7188.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,99/2 = 0,495$$

ta được z_α = 2,58.

Bước 3: Kiểm định. Vì $|z|$ = 1,7188 < 2,58 = z_α nên ta chấp nhận giả thiết $\mathbf{H_0}$: $\mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 1%, có thể xem đường kính trung bình của một chi tiết máy do hai nhà máy sản xuất là bằng nhau.

b) Có thể cho rằng đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất lớn hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất hay không với mức ý nghĩa 5%?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa α = 5% = 0,05:

$$\mathbf{H_0}: \mu_X = \mu_Y \text{ với giả thiết đối } \mathbf{H_1}: \mu_X > \mu_Y.$$

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = 1,7188.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha}$ = 1,65.

Bước 3: Kiểm định. Vì z = 1,7188 > 1,65 = $z_{2\alpha}$ nên ta bác bỏ giả thiết $\mathbf{H_0}$: $\mu_X = \mu_Y$, nghĩa là chấp nhận H_1: $\mu_X > \mu_Y$.

Kết luận: Với mức ý nghĩa 5%, có thể xem đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất lớn hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất.

c) Xét xem đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất có nhỏ hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất hay không với mức ý nghĩa 2%?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 2\% = 0,02$:

$$\mathbf{H_0}: \mu_X = \mu_Y \text{ với giả thiết đối } \mathbf{H_1}: \mu_X > \mu_Y.$$

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = 1,7188.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$$

ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $z = 1,7188 < 2,06 = z_{2\alpha}$ nên ta chấp nhận giả thiết $\mathbf{H_0}: \mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 2%, chưa thể xem đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất nhỏ hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất.

d) Với mức ý nghĩa 4%, tỉ lệ sản phẩm loại C do hai nhà máy sản xuất có như nhau không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 4\% = 0,04$:

$$\mathbf{H_0}: p_1 = p_2 \text{ với giả thiết đối } \mathbf{H_1}: p_1 \neq p_2$$

Ta kiểm định như sau:

Bước 1: Ta có:

$$p_0 = \frac{n_1F_{n1} + n_2F_{n2}}{n_1 + n_2} = \frac{28 + 16}{121 + 111} = 0,1897.$$

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1 - p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0,2314 - 0,1441}{\sqrt{0,1897(1 - 0,1897)\left(\frac{1}{121} + \frac{1}{111}\right)}} = 1,6942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,96/2 = 0,48$$

ta được $z_\alpha = 2,06$.

Bước 3: Kiểm định. Vì $|z| = 1,6942 < 2,06 = z_\alpha$ nên ta chấp nhận giả thiết H_0: $p_1 = p_2$.

Kết luận: Với mức ý nghĩa 4%, có thể xem tỉ lệ sản phẩm loại C do hai nhà máy sản xuất là như nhau.

e) Với mức ý nghĩa 3%, có thể cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất lớn hơn tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 3\% = 0,03$:

$$\mathbf{H_0}: p_1 = p_2 \text{ với giả thiết đối } \mathbf{H_1}: p_1 > p_2$$

Ta kiểm định như sau:

Bước 1: Tương tự câu d), ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1 - p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 1,6942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,94/2 = 0,47$$

ta được $z_{2\alpha} = 1,88$.

Bước 3: Kiểm định. Vì $z = 1,6942 < 1,88 = z_{2\alpha}$ nên ta chấp nhận giả thiết H_0: $p_1 = p_2$.

Kết luận: Với mức ý nghĩa 3%, chưa thể cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất lớn hơn tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất.

f) Hãy nhận xét về ý kiến cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất nhỏ hơn tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất với mức ý nghĩa 5%?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$\mathbf{H_0}: p_1 = p_2 \text{ với giả thiết đối } \mathbf{H_1}: p_1 > p_2$$

Ta kiểm định như sau:

Bước 1: Tương tự câu d), ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1 - p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 1,6942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 1,6942 > 1,65 = z_{2\alpha}$ nên ta bác bỏ giả thiết H_0: $p_1 = p_2$, nghĩa là chấp nhận H_1: $p_1 > p_2$.

Kết luận: Với mức ý nghĩa 5%, có thể chấp nhận ý kiến cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất nhỏ hơn tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất.
