

Thống kê Bayes

Bài tập buổi 03 - DS303.M21

Trần Phan Quốc Bảo
Sinh viên: Phạm Đức Thế - 19522253

Thứ 4, ngày 16 tháng 03 năm 2022

Bài tập 1: Phân phối nhị thức

- (a) Biết tỉ lệ suy dinh dưỡng ở trẻ em dưới 5 tuổi là 20%. Nếu chúng ta khám 15 trẻ dưới 5 tuổi. Tính xác suất để có 3 em bị suy dinh dưỡng.

Lời giải.

- Tỉ lệ suy dinh dưỡng ở trẻ em dưới 5 tuổi là 20% \Rightarrow Xác suất thành công: $p = 0,2$
- Xác suất thất bại: $q = 1 - p = 1 - 0,2 = 0,8$
- Xác suất để có 3 em bị suy dinh dưỡng trên tổng số 15 trẻ dưới 5 tuổi được khám là:

$$\begin{aligned}P(X = x) &= C_n^x \cdot p^x \cdot (1 - p)^{n-x} = C_n^x \cdot p^x \cdot q^{n-x} \\ \Rightarrow P(X = 3) &= C_{15}^3 \cdot (0,2)^3 \cdot (0,8)^{15-3} \simeq 0,25\end{aligned}$$

- (b) Tỉ lệ suy dinh dưỡng trẻ em 20% (như ví dụ trên). Tính xác suất để có ít hơn 4 em bị suy dinh dưỡng.

Lời giải.

- Tỉ lệ suy dinh dưỡng ở trẻ em dưới 5 tuổi là 20% \Rightarrow Xác suất thành công: $p = 0,2$
- Xác suất thất bại: $q = 1 - p = 1 - 0,2 = 0,8$
- Xác suất để không có em nào bị suy dinh dưỡng trên tổng số 15 trẻ dưới 5 tuổi được khám là:

$$\begin{aligned}P(X = x) &= C_n^x \cdot p^x \cdot (1 - p)^{n-x} = C_n^x \cdot p^x \cdot q^{n-x} \\ \Rightarrow P(X = 0) &= C_{15}^0 \cdot (0,2)^0 \cdot (0,8)^{15-0} \simeq 0,035\end{aligned}$$

- Xác suất để có 1 em bị suy dinh dưỡng trên tổng số 15 trẻ dưới 5 tuổi được khám là:

$$\begin{aligned}P(X = x) &= C_n^x \cdot p^x \cdot (1 - p)^{n-x} = C_n^x \cdot p^x \cdot q^{n-x} \\ \Rightarrow P(X = 1) &= C_{15}^1 \cdot (0,2)^1 \cdot (0,8)^{15-1} \simeq 0,132\end{aligned}$$

- Xác suất để có 2 em bị suy dinh dưỡng trên tổng số 15 trẻ dưới 5 tuổi được khám là:

$$\begin{aligned}P(X = x) &= C_n^x \cdot p^x \cdot (1 - p)^{n-x} = C_n^x \cdot p^x \cdot q^{n-x} \\ \Rightarrow P(X = 2) &= C_{15}^2 \cdot (0,2)^2 \cdot (0,8)^{15-2} \simeq 0,231\end{aligned}$$

- Xác suất để có 3 em bị suy dinh dưỡng trên tổng số 15 trẻ dưới 5 tuổi được khám là:

$$\begin{aligned}P(X = x) &= C_n^x \cdot p^x \cdot (1 - p)^{n-x} = C_n^x \cdot p^x \cdot q^{n-x} \\ \Rightarrow P(X = 3) &= C_{15}^3 \cdot (0,2)^3 \cdot (0,8)^{15-3} \simeq 0,25\end{aligned}$$

- Xác suất để có ít hơn 4 em bị suy dinh dưỡng trên tổng số 15 trẻ dưới 5 tuổi được khám là:

$$\begin{aligned}P(X < 4) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0,035 + 0,132 + 0,231 + 0,25 = 0,648\end{aligned}$$

Bài tập 2: Phân phối đa thức

- (a) Ta bốc 5 lá bài bất kì từ một bộ bài Tây. Tìm xác suất bốc được 1 lá cơ, 1 lá rô, 1 lá chuồn và 2 lá bích.

Lời giải.

- Gọi p_1, p_2, p_3, p_4 lần lượt là xác suất bốc được lá cơ, lá rô, lá chuồn và lá bích.
- Ta có: $p_1 = p_2 = p_3 = p_4 = \frac{1}{4} = 0,25$
- Ta bốc 5 lá bài $\Rightarrow n = 5$
- Xác suất bốc được 1 lá cơ, 1 lá rô, 1 lá chuồn và 2 lá bích là:

$$\begin{aligned}P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) &= \frac{n!}{\prod_{i=1}^k x_i!} \prod_{i=1}^k p_i^{x_i} \\ \Rightarrow P(X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 2) &= \frac{5!}{1!1!1!2!} (0,25) \cdot (0,25) \cdot (0,25) \cdot (0,25)^2 \\ &= \frac{15}{256} \simeq 0,059\end{aligned}$$

- (b) Giả sử ta có một hộp bi gồm 10 viên bi. Trong đó có 2 bi đỏ, 3 bi xanh và 5 bi vàng. Ta sẽ bốc 4 viên bi ra khỏi hộp bi trên cùng một lúc. Hỏi xác suất ta bốc được 2 bi xanh và 2 bi vàng là bao nhiêu.

Lời giải.

- Gọi p_1, p_2, p_3 lần lượt là xác suất bốc được bi đỏ, bi xanh và bi vàng.
- Ta có: $p_1 = \frac{2}{10} = 0,2; p_2 = \frac{3}{10} = 0,3; p_3 = \frac{5}{10} = 0,5$
- Ta sẽ bốc 4 viên bi ra khỏi hộp bi trên cùng một lúc $\Rightarrow n = 4$
- Xác suất bốc được 2 bi xanh và 2 bi vàng là:

$$\begin{aligned}P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) &= \frac{n!}{\prod_{i=1}^k x_i!} \prod_{i=1}^k p_i^{x_i} \\ \Rightarrow P(X_1 = 0, X_2 = 2, X_3 = 2) &= \frac{4!}{0!2!2!} (0,2)^0 \cdot (0,3)^2 \cdot (0,5)^2 = 0,135\end{aligned}$$

Bài tập 3: Phân phối Poisson

Một trạm cho thuê xe ô tô có 3 xe. Mỗi chiếc xe cho thuê được với giá 100 nghìn/ngày. Giả sử yêu cầu thuê xe của trạm là biến ngẫu nhiên X có phân phối Poisson với tham số $\lambda = 3$.

- (a) Tính xác suất trong một ngày có 3 khách thuê.

Lời giải.

- Xác suất trong một ngày có 3 khách thuê là:

$$\begin{aligned}P(X = k) &= C_n^k p^k (1-p)^{n-k} \simeq \frac{\lambda^k}{k!} e^{-\lambda} \\ \Rightarrow P(X = 3) &= \frac{3^3}{3!} e^{-3} \simeq 0,224\end{aligned}$$

- (b) Tính tiền lãi trung bình trạm thu được trong một ngày.

Lời giải.

Gọi Y là biến ngẫu nhiên tiền lãi. Ta sẽ chia thành nhiều trường hợp để xét bài toán trên.

- Trường hợp không xe nào được thuê:

$$P(Y = 0) = P(X = 0) = \frac{3^0}{0!} e^{-3} \simeq 0,05$$

- Trường hợp n xe được thuê với $1 \leq n \leq 3$:

$$P(Y = 100n) = P(X = n) = \frac{3^n}{n!} e^{-3}$$

Tiền lãi kì vọng (trung bình) của nhà xe là:

$$\begin{aligned}
 E(Y) &= \sum_{i=0}^3 P(Y = 100i) \cdot 100i \\
 &= P(Y = 100 \cdot 0) \cdot 100 \cdot 0 + P(Y = 100 \cdot 1) \cdot 100 \cdot 1 + P(Y = 100 \cdot 2) \cdot 100 \cdot 2 + P(Y = 100 \cdot 3) \cdot 100 \cdot 3 \\
 &= \frac{3^0}{0!} e^{-3} \cdot 100 \cdot 0 + \frac{3^1}{1!} e^{-3} \cdot 100 \cdot 1 + \frac{3^2}{2!} e^{-3} \cdot 100 \cdot 2 + \frac{3^3}{3!} e^{-3} \cdot 100 \cdot 3 \\
 &\simeq 0 + 14,94 + 44,81 + 67,21 = 126,96 \text{ (nghìn/ngày)}
 \end{aligned}$$

- (c) Cũng là với bài toán trên, nếu ta thêm tiền thuế mỗi ngày cho quản lí thị trường là 50 k/ngày thì lợi nhuận trung bình của nhà xe sẽ là bao nhiêu?

Lời giải.

Tiền lãi kì vọng (trung bình) của nhà xe là:

$$\begin{aligned}
 E(Y) &= \sum_{i=0}^3 P(Y = 100i) \cdot 100i - 50P(Y = 100i) \\
 &= P(Y = 100 \cdot 0) \cdot 100 \cdot 0 - 50P(Y = 100 \cdot 0) + P(Y = 100 \cdot 1) \cdot 100 \cdot 1 - 50P(Y = 100 \cdot 1) \\
 &+ P(Y = 100 \cdot 2) \cdot 100 \cdot 2 - 50P(Y = 100 \cdot 2) + P(Y = 100 \cdot 3) \cdot 100 \cdot 3 - 50P(Y = 100 \cdot 3) \\
 &= \frac{3^0}{0!} e^{-3} \cdot 100 \cdot 0 - 50 \cdot \frac{3^0}{0!} e^{-3} + \frac{3^1}{1!} e^{-3} \cdot 100 \cdot 1 - 50 \cdot \frac{3^1}{1!} e^{-3} \\
 &+ \frac{3^2}{2!} e^{-3} \cdot 100 \cdot 2 - 50 \cdot \frac{3^2}{2!} e^{-3} + \frac{3^3}{3!} e^{-3} \cdot 100 \cdot 3 - 50 \cdot \frac{3^3}{3!} e^{-3} \\
 &\simeq 0 - 2,49 + 14,94 - 7,47 + 44,81 - 11,20 + 67,21 - 11,20 = 94,6 \text{ (nghìn/ngày)}
 \end{aligned}$$

Bài tập 4: Phân phối chuẩn

Năng suất của một loại cây ăn quả là một biến ngẫu nhiên phân phối chuẩn với năng suất trung bình là 15 kg/cây và độ lệch chuẩn là 2 kg. Cây đạt tiêu chuẩn hàng hóa là cây có năng suất tối thiểu 12 kg. Tính tỉ lệ cây đạt tiêu chuẩn hàng hóa.

Lời giải.

- Gọi X là năng suất của một loại cây ăn quả.
- Ta có $X \sim N(\mu, \sigma^2)$ với $\mu = 15$ và $\sigma = 2$
- Tỉ lệ cây đạt tiêu chuẩn hàng hóa là:

$$\begin{aligned}
 P(X \geq 12) &= P\left(\frac{x - \mu}{\sigma} \geq \frac{12 - \mu}{\sigma}\right) = \frac{1}{2} - \Phi_0\left(\frac{12 - 15}{2}\right) = \frac{1}{2} - \Phi_0(-1,5) \\
 &= \frac{1}{2} - 0,0668 = 0,4332
 \end{aligned}$$