CHƯƠNG MỘT

BIẾN SỐ NGẪU NHIÊN

Định nghĩa . Cho Ω là một tập hợp khác trống, $\mathcal{A}$ là một σ-đại số trên Ω, và P là một độ đo dương trên không gian đo được $(\Omega, \mathcal{A})$ sao cho $P(\Omega) = 1$. Lúc đó ta nói $(\Omega, \mathcal{A}, P)$ là một *không gian xác suất* và P là một *độ đo xác suất*.

Định nghĩa. Cho $(\Omega, \mathcal{A}, P)$ là một không gian xác suất và X là một hàm số thực trên Ω sao cho $X^{-1}((-\infty,c)) \in \mathcal{A}$ với mọi số thực c . Lúc đó ta nói X là một *hàm số thực đo được* trên Ω, hoặc X là một *biến số ngẫu nhiên* .

X là một *hàm đơn (simple function)* trên $(\Omega, \mathcal{A}, P)$ nếu và chỉ nếu có một họ hữu hạn $\{A_1, \ldots, A_m\}$ trong $\mathfrak{M}$ và một họ hữu hạn $\{\alpha_1, \ldots, \alpha_m\}$ trong $\mathbb{R}$ sao cho

$$X(x) = \sum_{k=1}^{m} \alpha_k \chi_{A_k}(x) \qquad \forall x \in \Omega.$$

trong đó $$\chi_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \in \Omega \setminus A. \end{cases}$$

Lúc đó X là một biến số ngẫu nhiên và ta đặt

$$E(X) = \sum_{k=1}^{m} \alpha_k P(A_k)$$

và gọi $E(X)$ là *kỳ vọng* (*expectation*) của biến ngẫu nhiên X

Định nghĩa. Cho $(\Omega, \mathcal{A}, P)$ là một không gian đo được, $E \in \mathcal{A}$, và X là một biến ngẫu nhiên không âm trên từ $(\Omega, \mathcal{A})$. Đặt $\mathfrak{F}(X)$ là họ các hàm đơn s trên Ω sao cho $0 \le s \le X$ và đặt

$$\int_\Omega X dP = \sup_{s \in F(X)} E(s) \quad \text{và} \quad E(X) = \int_\Omega X dP.$$

Ta gọi $E(X)$ là *kỳ vọng* của biến ngẫu nhiên X. Kỳ vọng của X có thể bằng ∞.

Cho X và Y là hai biến số ngẫu nhiên trên $(\Omega,\mathcal{A},P)$, và a và b là hai số thực. Lúc đó các hàm số sau đều là các biến ngẫu nhiên trên $(\Omega,\mathcal{A},P)$: $aX + bY$, $\sup\{X,Y\}$, $\inf\{X,Y\}$, $X^+ = \sup\{0 , X\}$, $X^- = \sup\{0 ,- X\}$ và $|X|$.

Định nghĩa. Cho X là một biến số ngẫu nhiên trên $(\Omega,\mathcal{A},P)$. Ta nói X thuộc lớp $\mathcal{L}(\Omega)$ nếu $E(|X|) < \infty$. Lúc đó ta đặt *kỳ vọng* của X như sau :

$$E(X) = E(X^+) - E(X^-) .$$

Cho X và Y thuộc lớp $\mathcal{L}(\Omega)$, và a và b là hai số thực. Lúc đó $aX + bY$ thuộc lớp $\mathcal{L}(\Omega)$.

Cho X và Y thuộc lớp $\mathscr{L}(\Omega)$, ta nói $X \sim Y$ nếu và chỉ nếu

$$P(\{\omega \in \Omega : X(\omega) \neq Y(\omega)\}) = 0 .$$

Quan hệ $\sim$ là quan hệ tương đương trong $\mathscr{L}(\Omega)$, ta ký hiệu $\mathrm{L}^1(\Omega)$ là tập hợp thương $\mathscr{L}(\Omega)/\sim$.

Cho X và Y là hai biến số ngẫu nhiên trong cùng một lớp tương đương. Ta thấy $E(X) = E(Y)$. Vì thế ta có thể xem một lớp tương đương U như là một biến số ngẫu nhiên X nếu X là một phần tử trong U, và đặt $E(U) = E(X)$.

Như vậy ta có thể xem $\mathrm{L}^1(\Omega)$ như là họ các biến số ngẫu nhiên sao cho $E(|X|) < \infty$. Ta đặt

$$E(X) = E(X^+) - E(X^-) \qquad \forall X \in \mathrm{L}^1(\Omega)$$

Cho một không gian xác suất $(\Omega,\mathcal{A},P)$ và q là một số thực > 1. Ta đặt

$$L^q(\Omega,\mathcal{A},P) = \{\, X \in L^1(\Omega) : E(|X|^q) < \infty \,\}.$$

$L^q(\Omega,\mathcal{A},P)$ là một không gian Banach với chuẩn sau

$$\| X \|_{L^q} = [E(| X |^q)]^{1/q} = [\int_\Omega | X |^q dP]^{1/q} \qquad \forall X \in L^q(\Omega,\mathcal{A},P)$$

$L^2(\Omega,\mathcal{A},P)$ là một không gian Hilbert với tích vô hướng sau

$$< X, Y > = E(XY) \qquad \forall X, Y \in L^2(\Omega,\mathcal{A},P) .$$

Nếu trong bài toán không có gì có thể gây lầm lẫn, chúng ta dùng $L^q(\Omega)$ để ký hiệu $L^q(\Omega,\mathcal{A},P)$.

Định nghĩa. Cho một không gian xác suất $(\Omega, \mathfrak{A}, P)$ và $\mathfrak{U}$ là một σ-đại số trong Ω , $\mathfrak{U} \subset \mathfrak{A}$. Lúc đó $L^2(\Omega, \mathfrak{U}, P)$ là một không gian vectơ con đóng trong $L^2(\Omega, \mathfrak{A}, P)$ và có duy nhất một ánh xạ tuyến tính T từ $L^2(\Omega, \mathfrak{A}, P)$ vào $L^2(\Omega, \mathfrak{U}, P)$ sao cho $\|T\| = 1$, $T(X) = X$ với mọi X trong $L^2(\Omega, \mathfrak{U}, P)$ và

$$< Y , X - T(X) > \; = 0 \quad \forall\, X \in L^2(\Omega, \mathfrak{A}, P),\ Y \in L^2(\Omega, \mathfrak{U}, P)$$

Ta ký hiệu $T(X)$ là $T(X|\mathfrak{U})$ và gọi là kỳ vọng có điều kiện $\mathfrak{U}$ của X.

Ta có $$\int_\Omega E(X \mid P) Y dP = \int_\Omega XY dP \qquad \forall Y \in L^2(\Omega, \mathcal{U}, P),$$

$$\int_D E(X \mid P) dP = \int_D X dP \qquad \forall D \in \mathcal{U}.$$

Định nghĩa. Cho Y là một biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Đặt $\mathcal{F}_Y$ là σ- đại số nhỏ nhất chứa họ $\{Y^{-1}(U) : U$ là một tập mở trong $\mathbb{R}\}$. Ta đặt

$$E(X|Y) = E(X\,|\mathcal{F}_Y) \qquad \forall\ X \in L^2(\Omega\ ,\mathcal{A},P).$$

Định nghĩa. Cho X là một biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Ta đặt $F_X : \mathbb{R} \to [0 , \infty]$

$$F_X(t) = P(X^{-1}((-\infty,t))) = P(\{\omega \in \Omega : X(\omega) < t \}) \quad \forall t \in \mathbb{R}.$$

Ta nói F_X là *hàm phân phối* (*distribution function*) của biến ngẫu nhiên X.

Bài toán 1.1. Tập $\{t \in \mathbb{R} : P(X^{-1}(\{t\})) > 0\}$ là một tập quá lắm đếm được.

H.D. Giả sử có k số thực khác nhau $t_1, . . ., t_k$ sao cho $P(X^{-1}(\{t_j\})) > m^{-1}$. Để ý

$$\sum_{i=1}^{k} P(X^{-1}(\{t_i\})) = P(X^{-1}(\{t_1,\cdots,t_k\})) \le 1.$$

Suy ra $k \le m$.

Bài toán 1.2. Cho một dãy đơn điệu tăng biến số ngẫu nhiên không âm $\{X_m\}$ hội tụ từng điểm về một biến số ngẫu nhiên X trong $(\Omega, \mathcal{A}, P)$. Chứng minh $\{F_{X_m}\}$ là một dãy hàm số đơn điệu giãm và hội tụ hầu hết mọi nơi về F_X trên $[0,\infty)$.

H.D. Đặt $A = \{s \in \mathbb{R} : P(X^{-1}(\{s\})) > 0\}$, ta có $\mu(A) = 0$.
Cho $t \in \mathbb{R} \setminus A$, đặt $B_m = X_m^{-1}((-\infty,t))$ và $B = X^{-1}((-\infty,t))$.
Chứng minh $\lim_{m\to\infty} \chi_{B_m}(s) = \chi_B(s) \qquad \forall s \in \mathbb{R} \setminus X^{-1}(\{t\})$.

Cho $\alpha_1 < \alpha_2 < \ldots < \alpha_m$ và $\{A_1, \ldots, A_m\}$ là m tập đo được rời nhau và có phần hội là Ω. Cho X là một biến số ngẫu nhiên có dạng

$$X(x) = \sum_{k=1}^{m} \alpha_k \chi_{A_k}(x) \qquad \forall x \in \Omega,$$

Lúc đó

$$F_X = \chi_{[\alpha_m, \infty)} + \sum_{i=1}^{m-1} \Big[\sum_{j=1}^{i} P(A_j)\Big] \chi_{(\alpha_i, \alpha_{i+1}]}$$

$$F_X = \chi_{[\alpha_m, \infty)} + \sum_{i=1}^{m-1} \Big[\sum_{j=1}^{i} P(A_j)\Big] \chi_{(\alpha_i, \alpha_{i+1}]}$$

Bài toán 1.3. Cho $q \in (1,\infty)$, $0 < \alpha_1 < \alpha_2 < \ldots < \alpha_m$ và $\{A_1, \ldots, A_m\}$ là m tập đo được rời nhau và có phần hội là Ω. Cho X là một biến số ngẫu nhiên có dạng

$$X(x) = \sum_{k=1}^{m} \alpha_k \chi_{A_k}(x) \qquad \forall x \in \Omega.$$

Chứng minh $$E(X^q) = q\int_0^\infty s^{q-1}(1 - F_X(s))ds$$

$$X^p(x) = \sum_{k=1}^{m} \alpha_k^p \chi_{A_k}(x) \qquad \forall x \in \Omega.$$

$$E(X^q) = E\left(\sum_{i=1}^{m} \alpha_i^q \chi_{A_i}\right) = \sum_{i=1}^{m} \alpha_i^q P(A_i) \qquad (1)$$

$$F_X = \chi_{[\alpha_m,\infty)} + \sum_{i=1}^{m-1}[\sum_{j=1}^{i} P(A_j)]\chi_{(\alpha_i,\alpha_{i+1}]}$$

$$q\int_0^\infty s^{q-1}(1-P_X(s))ds =$$

$$= q\int_0^\infty [1-(s^{q-1}(\chi_{[\alpha_m,\infty)} + \sum_{i=1}^{m-1}[\sum_{j=1}^{i} P(A_j)]\chi_{(\alpha_i,\alpha_{i+1}]})(s)]ds$$

$$= q\int_0^{\alpha_m} \sum_{i=1}^{m-1}[1-\sum_{j=1}^{i} P(A_j)]s^{q-1}\chi_{(\alpha_i,\alpha_{i+1}]})(s)ds$$

$$= \sum_{i=1}^{m-1}[1-\sum_{j=1}^{i} P(A_j)]q\int_{\alpha_i}^{\alpha_{i+1}} s^{q-1}ds$$

$$= \sum_{i=1}^{m-1}[1-\sum_{j=1}^{i} P(A_j)](\alpha_{i+1}^q - \alpha_i^q)$$

$$= \sum_{i=1}^{m-1}[1-\sum_{j=1}^{i}P(A_j)](\alpha_{i+1}^{q}-\alpha_{i}^{q})$$

$$= \alpha_m^q - \sum_{i=1}^{m-1}\sum_{j=1}^{i}P(A_j)(\alpha_{i+1}^{q}-\alpha_{i}^{q})$$

$$= \alpha_m^q - \sum_{i=1}^{m-1}\sum_{j=1}^{i}P(A_j)\alpha_{i+1}^{q} + \sum_{i=1}^{m-1}\sum_{j=1}^{i}P(A_j)\alpha_{i}^{q}$$

$$= \alpha_m^q - \sum_{j=1}^{m-1}\sum_{i=j}^{m-1}P(A_j)\alpha_{i+1}^{q} + \sum_{j=1}^{m-1}\sum_{i=j}^{m-1}P(A_j)\alpha_{i}^{q}$$

$$= \alpha_m^q - \sum_{j=1}^{m-1}P(A_j)[\sum_{i=j+1}^{m}\alpha_i^q - \sum_{i=j}^{m-1}\alpha_i^q] = \alpha_m^q - \sum_{j=1}^{m-1}P(A_j)[\alpha_m^q - \alpha_j^q]$$

$$= \alpha_m^q - \sum_{j=1}^{m-1}P(A_j)\alpha_m^q + \sum_{j=1}^{m-1}P(A_j)\alpha_j^q = \sum_{j=1}^{m}P(A_j)\alpha_j^q$$

Bài toán 1.4. Cho $q \in (1,\infty)$ và X là một biến số ngẫu nhiên không âm trên $(\Omega, \mathcal{A}, P)$. Chứng minh $X \in L^q(\Omega)$ nếu và chỉ nếu

$$q\int_0^\infty s^{q-1}(1-F_X(s))ds \ < \ \infty,$$

Và lúc đó $$E(X^q) \ = \ q\int_0^\infty s^{q-1}(1-F_X(s))ds \ .$$

H.D. Chọn một dãy biến ngẫu nhiên đơn không âm $\{X_m\}$ đơn điệu tăng hội tụ về X. Chứng minh $\{X_m{}^q\}$ đơn điệu tăng hội tụ về X^q, và $\{1-P_{X_m}\}$ đơn điệu tăng và hội tụ hầu hết mọi nơi vê $(1 - P_X)$. Giải bài toán cho các biến ngẫu nhiên đơn trước.

Định nghĩa. Cho X là một biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Ta nói X *rời rạc* (discrete) nếu $X(\Omega)$ là một tập con quá lắm đếm được $\{x_i\}_{i\in I}$ trong $\mathbb{R}$.

Lúc đó đặt $p : \mathbb{R} \to [0, \infty]$:

$$p(s) = P(X^{-1}(\{s\})) = P(\{\omega \in \Omega : X(\omega) = s\}) \quad \forall s \in \mathbb{R}.$$

Ta gọi p là *hàm mật độ* (*the probability mass function*) của biến ngẫu nhiên X, và ta có

$$F_X(t) = \sum_{x_i < t} p(x_i) \qquad \forall t \in \mathbb{R}.$$

$$F_X(t) = P(X^{-1}((-\infty, t])) = P(\{\omega \in \Omega : X(\omega) \le t\}) \quad \forall t \in \mathbb{R}.$$

Định nghĩa. Cho X là một biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Ta nói X *liên tục* (continuous) nếu có một hàm số thực không âm liên tục từng mảnh trên $\mathbb{R}$ sao cho

$$F_X(t) = \int_{-\infty}^{t} p(s)ds \qquad \forall\ t \in \mathbb{R}.$$

Ta gọi p là *hàm mật độ* (*probability mass function*) của biến ngẫu nhiên X.

$$F_X(t) = P(X^{-1}((-\infty, t])) = P(\{\omega \in \Omega : X(\omega) \le t\}) \quad \forall t \in \mathbb{R}.$$

Vai trò hàm phân phối F_X và hàm mật độ xác suất p có vai trò rất quan trọng đối với biến số ngẫu nhiên X. Căn cứ vào p ta gọi tên các biến số ngẫu nhiên như sau :

• Nếu $X(\Omega) = \{0,1,2, \ldots n\}$ và

$$p(t) = \begin{cases} \dfrac{t!}{n!(n-t)!} p^t (1-p)^{n-t} & \forall t \in \{0,1,2,\ldots,n\}, \\ 0 & \forall t \in \mathbb{R} \setminus \{0,1,2,\ldots,n\}, \end{cases}$$

X được gọi là biến số ngẫu nhiên có phân phối nhị thức $B(n,p)$, và được ký hiệu $X \sim B(n,p)$.

• Nếu $X(\Omega) = \{0,1,2, \ldots, n, \ldots\}$ và có một số thực $\lambda > 0$

$$p(t) = \begin{cases} \dfrac{\lambda^t e^{-\lambda}}{t!} & \forall t \in \{0,1,2,\ldots,n,\ldots\}, \\ 0 & \forall t \in \mathbb{R} \setminus \{0,1,2,\ldots,n,\ldots\}, \end{cases}$$

X được gọi là biến số ngẫu nhiên có phân phối Poisson $P(\lambda)$, và được ký hiệu $X \sim P(\lambda)$.

- Nếu $X(\Omega) = \mathbb{R}$ và có hai số thực a và b sao cho $b > a$ và

$$p(t) = \begin{cases} \dfrac{1}{b-a} & \forall t \in [a,b], \\ 0 & \forall t \in \mathbb{R} \setminus [a,b], \end{cases}$$

X được gọi là biến số ngẫu nhiên có phân phối đều (uniformly distributed) $U(a;b)$, và được ký hiệu $X \sim U(a;b)$

- Nếu $X(\Omega) = \mathbb{R}$ và có hai số thực μ và σ sao cho $\sigma > 0$

và

$$p(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \qquad \forall t \in \mathbb{R},$$

X được gọi là biến số ngẫu nhiên có phân phối chuẩn (normal distributed) $N(\mu;\, \sigma^2)$, và được ký hiệu $X \sim N(\mu;\sigma^2)$

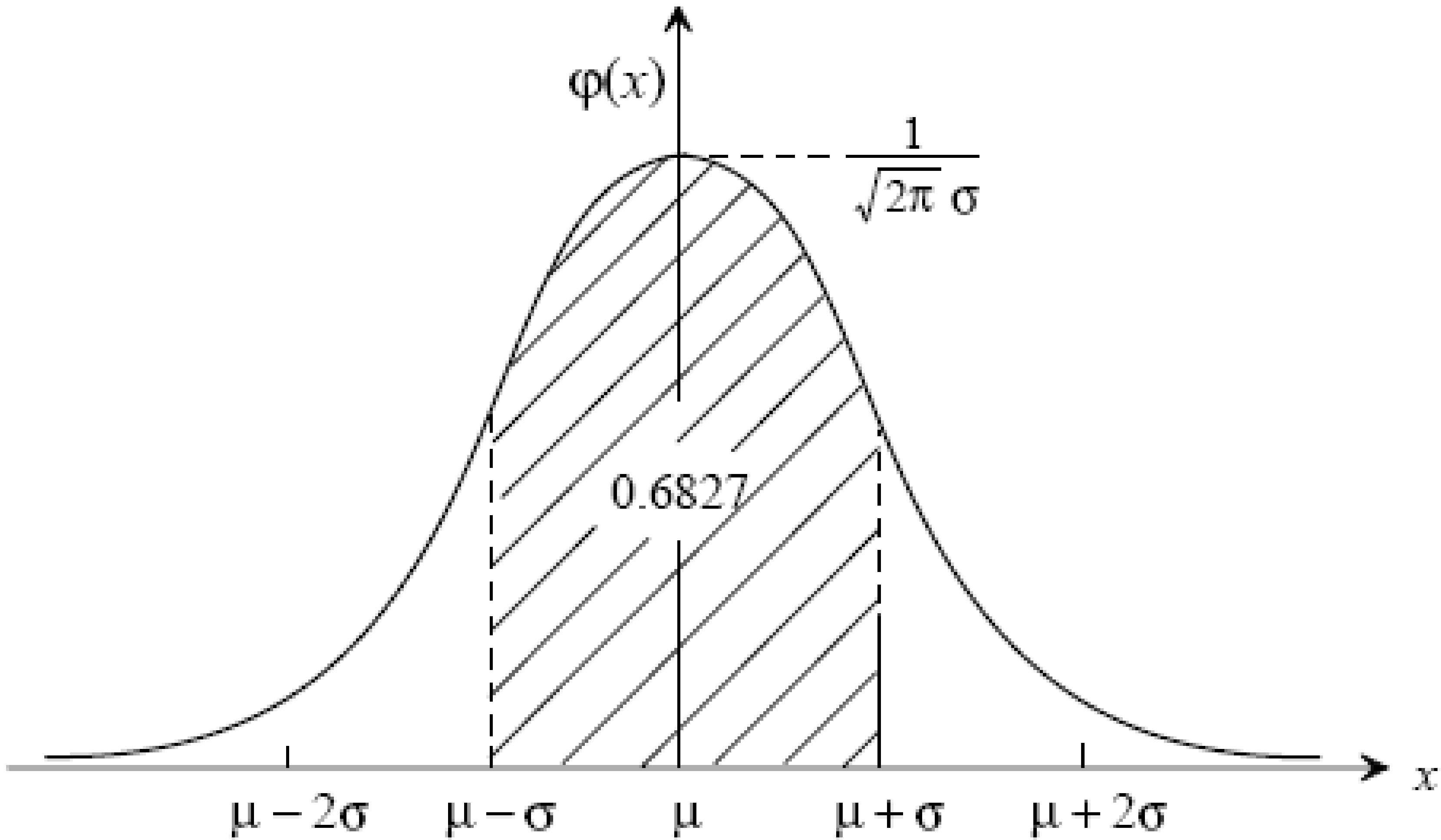

Figure 1.7 Density of the normal distribution (Gaussian bell curve)

Thí dụ. On average, only 0.01% of trout eggs will develop into adult fishes. What is the probability p_a that at least three adult fishes arise from 40,000 eggs?

Theo lối làm toán bình thường chúng ta làm các bước sau :

- Số trứng phát triển thành cá trưởng thành là :

$$40.000 \times 0{,}0001 = 4.$$

- Vậy xác suất để có ít nhất có 3 trứng trong mỗi lô 40.000 trứng phát triển thành cá trưởng thành là 100%.

Nhưng thực tế có các lô 40.000 trứng không phát triển được một con cá trưởng thành nào cả. Tại sao vậy?

Thí dụ . On average, only 0.01% of trout eggs will develop into adult fishes. What is the probability p_a that at least three adult fishes arise from 40,000 eggs?

Thực ra con số 0.01% chỉ là xác suất mà người ta có được được qua rất nhiều lần lấy mẫu thử nghiệm, chứ không thể nào có con số chính xác như vậy được.

Trong hoàn cảnh lơ mơ thế này, ta phải giải toán với triết lý lơ mơ.

Ta dùng thống kê để giải bài này. Nếu chúng ta làm k thử nghiệm nuôi từng lô gồm 40000 trứng, không gian mẫu $\Omega = \{1, \ldots, k\}$, đặt $\mathcal{A}$ là $\mathcal{P}(\Omega)$, và $X(m)$ là số cá có được trong lần thử nghiệm thứ m. Lúc đó $X(\Omega) = \{1, \ldots, 40000\}$, X có phân bố nhị thức $B(40000, 0{,}0001)$ và có hàm mật độ là

$$p(j)=P(\{X=j\})=\frac{40000}{j!(40000-j)!}(0,0001)^{j}(0,9999)^{40000-j}$$

Dùng công thức này để tính rất khó, ta xấp nó bằng hàm mật độ của phân phối Poisson $P(\lambda)$ với $\lambda = 0,0001\times 40000$

$$p(j)\approx\frac{4^{j}}{j!}e^{-4} \qquad \forall\ j\in\{1,\ldots,40000\}.$$

Vậy xác suất để có ít nhất ba con cá nở từ 40000 trứng là

$$p = 1 - p(0) - p\,(1) - p\,(2) \approx 1 - 0,0183 - 0,0733 - 0,1465$$
$$= 0,7619.$$

Như vậy với kiến thức của thông kê, chúng ta sẽ lập các dự án khả thi và có kết quả hơn.

Qua thí dụ này, ta thấy đôi khi chúng ta không quá bận tâm đến $(\Omega,\mathcal{A},P)$, mà chỉ để ý đến $X(\Omega)$, dạng hàm phân phối của X và hàm mật độ tương ứng xác định trên $X(\Omega)$.

Vì thế trong nhiều bài toán chúng ta chẳng cần đề cập đến $(\Omega,\mathcal{A},P)$ mà vẫn nói về các biến ngẫu nhiên. Trong các trường hợp này chúng ta sẽ làm việc với các yếu tố sau : $X(\Omega)$, hàm phân phối P_X và hàm mật độ p tương ứng với X.

Cách chọn hàm phân phối cho từng bài toán dựa vào kinh nghiệm thực tiễn về mức độ phù hợp của chúng trong các áp dụng thống kê.

Sau đây chúng ta sẽ thấy có thể tính toán một số tích phân liên hệ với biến ngẫu nhiên X chỉ dựa trên P_X và p .

Bài toán 1.5. Cho X là một biến ngẫu nhiên rời rạc trong không gian xác suất $(\Omega, \mathcal{A}, P)$, $X \in L^1(\Omega)$, $X(\Omega)$ là một tập con quá lắm đếm được $\{x_i\}_{i\in I}$ trong $\mathbb{R}$, và p là hàm mật độ của X. Chứng minh

$$E(X) = \sum_{i\in I} x_i p(x_i) \equiv \sum_{i\in I} x_i P(X^{-1}(\{x_i\})).$$

H.D. Nếu $I = \mathbb{N}$, xét dãy $\{\sum_{i=1}^{m} x_i \chi_{X^{-1}(\{x_i\})}\}$

Bài toán 1.6. Cho X là một biến ngẫu nhiên liên tục không âm với hàm mật đô p trong không gian xác suất $(\Omega, \mathcal{A}, P)$, và $X \in L^q(\Omega)$. Chứng minh

$$E(X^q) = q\int_0^\infty t^{q-1}[1 - P_X(t)]dt = \int_0^\infty s^q p(s)ds \quad .$$

Bài toán 1.7. Cho X là một biến ngẫu nhiên liên tục không âm với hàm mật đô liên tục p trong không gian xác suất $(\Omega,\mathcal{A}, P)$ sao cho $\lim\limits_{N\to\infty} N^q[1-F_X(N)]=0$.

Chứng minh $X \in L^q(\Omega)$ và

$$E(X^q)=q\int_0^\infty t^{q-1}[1-F_X(t)]dt=\int_0^\infty s^q p(s)ds \qquad .$$

H.D. Đặt $u(t) = t^q$ và $v(t) = 1 - F_X(t)$. Lúc đó $u'(t) = qt^{q-1}$ và $v'(t) = -p(t)$. Áp dụng tích phân từng phần

$$q\int_0^N t^{q-1}[1-F_X(t)]dt = N^q[1-F_X(N)] + \int_0^N s^q p(s)ds \qquad \forall N\in\mathbb{N} \quad .$$

Bài toán 1.8. Cho X là một biến số ngẫu nhiên trên $(\Omega,\mathcal{A},P)$ có phân phối chuẩn $N(0,\sigma)$. Lúc đó $E(|X|^q) < \infty$ với mọi q trong $[1, \infty)$ và

$$E(|X|^q) == (\frac{2^{q/2}}{\sqrt{\pi}}\int_0^\infty t^q \exp(-t^2)dt)\sigma^q$$

H.D. Để ý $F_X(t) = \dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_{-\infty}^t \exp(-\frac{s^2}{2\sigma^2})ds \qquad \forall t \in \mathbb{R}.$

Suy ra hàm phân phối của biến ngẫu nhiên $|X|$ có dạng

$$p(t) = \frac{2}{\sigma\sqrt{2\pi}}\exp(-\frac{t^2}{2\sigma^2})\chi_{[0,\infty)}(t) \qquad \forall t \in \mathbb{R},$$

Để ý

$$\frac{2N^q}{\sigma\sqrt{2\pi}}\int_N^\infty \exp(-\frac{t^2}{2\sigma^2})dt \le \frac{N^q}{\sigma\sqrt{2\pi}}\int_N^\infty \exp(-\frac{t}{2\sigma^2})dt = \frac{4\sigma N^q}{\sqrt{2\pi}}\exp(-\frac{N}{2\sigma^2})$$

Suy ra

$$E(|X|^q) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty s^q \exp(-\frac{s^2}{2\sigma^2})ds$$

$$= \frac{1}{\sqrt{\pi}}\int_0^\infty (\sqrt{2}\sigma t)^q \exp(-t^2)dt = (\frac{2^{q/2}}{\sqrt{\pi}}\int_0^\infty t^q \exp(-t^2)dt)\sigma^q$$

Định nghĩa. Cho $(\Omega, \mathcal{A}, P)$ là một không gian xác suất. Cho $\mathcal{U}$ và $\mathcal{V}$ là hai σ-đại số trên Ω sao cho $\mathcal{U}$ và $\mathcal{V}$ đều bị chứa trong $\mathcal{A}$. Ta nói $\mathcal{U}$ và $\mathcal{V}$ độc lập trong $(\Omega, \mathcal{A}, P)$ nếu

$$P(A \cap B) = P(A)P(B) \qquad \forall\ A \in \mathcal{U},\ B \in \mathcal{V}.$$

Thí dụ . Cho $\Omega = [0,1]\times[01]$, $\mathcal{A}$ và P là σ-đại số và độ đo Lebesgue hạn chế trên Ω. Đặt

$\mathcal{U} = \{\ A\times[0,1] : A$ Lebesgue đo được chứa trong $[0,1]\}$,

$\mathcal{V} = \{[0,1]\times B : B$ Lebesgue đo được chứa trong $[0,1]\}$.

Lúc đó $\mathcal{U}$ và $\mathcal{V}$ độc lập trong $(\Omega,\mathcal{A}, P)$.

Định nghĩa. Cho X là một biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Đặt $\mathcal{F}_X$ là σ- đại số nhỏ nhất chứa họ $\{X^{-1}(U) : U$ là một tập mở trong $\mathbb{R}\}$.

Ta thấy :

(*i*) $\mathcal{F}_X \subset \mathcal{A}$,

(ii) Đặt P_X là thu hẹp của P trên $\mathcal{F}_X$. Lúc đó $(\Omega, \mathcal{F}_X, P_X)$ là một không gian đo được với độ đo P_X,

(iii) Có một dãy hàm đơn trong $(\Omega, \mathcal{F}_X)$ hội tụ P-hầu hết mọi nơi về X.

Bài toán 1.9. Cho $\mathcal{U}$ và $\mathcal{V}$ độc lập trong một không gian xác suất $(\Omega, \mathcal{A}, P)$. Cho f là một hàm số thực đo được trong $(\Omega,\mathcal{U})$ và g là một hàm số thực đo được trong $(\Omega,\mathcal{V})$. Giả sử f và g khả tích trên $(\Omega, \mathcal{A}, P)$. Chứng minh

$$E(fg) = E(f)\, E(g)\ .$$

H.D. Trước hết xét f và g là các hàm đơn. Sau đó xấp xỉ f và g bằng các hàm đơn. Dùng định lý hội tụ đơn điệu Lebesgue cho trường hợp f và g là các hàm không âm. Từ đó suy ra trường hợp tổng quát.

Định nghĩa. Cho X và Y là hai biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Ta nói X và Y độc lập với nhau nếu và chỉ nếu $\mathcal{F}_X$ và $\mathcal{F}_Y$ độc lập trong $(\Omega, \mathcal{A}, P)$, nghĩa là

$$P(A \cap B) = P(A)P(B) \qquad \forall\ A \in \mathcal{F}_X,\ B \in \mathcal{F}_Y$$

Bài toán 1.10. Cho $\Omega = [0,1]\times[01]$, $\mathcal{A}$ và P là σ-đại số và độ đo Lebesgue hạn chế trên Ω. Cho f và g là hai hàm số thực Lebesgue đo được trên Ω. Với mọi (x,y) trong Ω, ta đặt $X(x,y) = f(x)$, và $Y(x,y) = g(y)$. Chứng minh X và Y độc lập với nhau.

H.D. Để ý với mọi A trong $\mathcal{F}_X$, có một tập D Lebesgue đo được trong $[0,1]$ sao cho $A = D \times [01]$.

Bài toán 1.11. Cho X và Y là hai biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Nếu Y là một ánh xạ hằng. Chứng minh X và Y độc lập với nhau.

H.D. Để ý $\mathcal{F}_Y = \{\Omega, \phi\}$ luôn luôn độc lập với $\mathcal{F}_X$.

Bài toán 1.12. Cho X và Y là hai biến ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Giả sử X và Y khả tích trên Ω , và độc lập với nhau. Chứng minh

$$E(XY) = E(X)\, E(Y)$$

H.D. Để ý X và Y đo được trên $(\Omega, \mathcal{F}_X)$ và $(\Omega, \mathcal{F}_X)$.

Bài toán 1.13. Cho $\mathcal{U}_1, \ldots, \mathcal{U}_m$ là m σ-đại số độc lập trong một không gian xác suất $(\Omega, \mathcal{A}, P)$. Chứng minh

$$P(A_1 \cap \cdots \cap A_m) = P(A_1) \cdots P(A_m) \qquad \forall A_1 \in \mathcal{U}_1, \cdots, A_m \in \mathcal{U}_m .$$

H.D. Qui nạp toán học

VECTƠ NGẪU NHIÊN

Định nghĩa . Cho X_1 , ..., X_m là m biến ngẫu nhiên trong không gian xác suất (Ω, $\mathcal{A}$, P). Ta gọi $X = (X_1 , ..., X_m)$ là một vectơ ngẫu nhiên *m*-chiều trong (Ω, A, P).

Định nghĩa . Cho $X = (X_1 , ..., X_m)$ là một vectơ ngẫu nhiên trong không gian xác suất (Ω, $\mathcal{A}$, P). Ta đặt

$$F_X(t_1,\ldots,t_m) = P(A_{1,t_1} \cap \cdots \cap A_{m,t_m}) \qquad \forall \ t_1,\ldots,t_m \in \mathbb{R},$$

ở đây $A_{k,t_k} = \{x \in \Omega : X_k(x) < t_k\}$ với mọi $k \in \{1,\ldots,m\}$.

Ta gọi F_X là hàm phân phối của vectơ ngẫu nhiên X. Tương ứng ta cũng định nghĩa được hàm mật độ p_X có tính chất sau

• Nếu $X_1, \ldots, X_m$ là các biến số ngẫu nhiên rời rạc tương ứng với các họ $\{x_{1,j}\}, \ldots, \{x_{m,j}\}$, thì

$$p_X(x_{1,j_1},\ldots,x_{m,j_m}) = P(\{x : X_1(x) = x_{1,j_1},\ldots,X_m(x) = x_{m,j_m}\})$$

• Nếu $X_1, \ldots, X_m$ là các biến số ngẫu nhiên liên tục, thì

$$F_X(x_1,\ldots,x_m) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_m} p_X(s_1,\ldots,s_m)ds_m\ldots ds_1.$$

Nếu $X_1, \ldots, X_m$ độc lập thì các công thức trên trở thành

$$p_X(x_{1,j_1},\ldots,x_{m,j_m}) = p_{X_1}(x_{1,j_1})\ldots p_{X_m}(x_{m,j_m}),$$

$$F_X(x_1,\ldots,x_m) = \int_{-\infty}^{x_1} p_{X_1}(s_1)dx_1\cdots\int_{-\infty}^{x_m} p_{X_m}(s_m)dx_m.$$

Định nghĩa . Cho $X = (X_1, \ldots, X_m)$ là một vectơ ngẫu nhiên trong không gian xác suất $(\Omega, \mathcal{A}, P)$. Ta đặt

$$\mu = (\mu_1, \ldots, \mu_m) = (\int_\Omega X_1 dP, \ldots, \int_\Omega X_m dP),$$

$$a_{ij} = \text{covariance}(X_i X_j) \equiv \int_\Omega (X_i - \mu_i)(X_j - \mu_j) dP$$

$$E((X-\mu)(X-\mu)^T) = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}$$