Chương 4: BiẾN NGẪU NHIÊN LIÊN TỤC

1

BNN liên tục: BNN X thuộc loại liên tục khi tập hợp các giá trị thực mà X nhận được từ phép thử là tập hợp vô hạn không đếm được (tức là, tập chứa vô số giá trị liên tiếp nhau trên trục số thực $\mathbb{R}$) và $P\{X = c\} = 0, \forall c \in \mathbb{R}$.

Ví dụ: Y là chiều cao của một học sinh nữ được chọn ngẫu nhiên từ các học sinh nữ cao từ 1,5m đến 1,6m ở khu vực K.

$\Rightarrow$ Y là BNN liên tục có tập giá trị là đoạn [1,5 ; 1,6].

Tương tự, các BNN biểu diễn thời gian, kích thước, nồng độ,… (nói chung là kết quả phép đo) là các BNN liên tục.

2

4.1 Hàm mật độ xác suất của BNN liên tục

Định nghĩa : Cho *X* là BNN liên tục, khi đó hàm mật độ xác suất của *X*, thường được kí hiệu là $f(u)$, là hàm số một biến thỏa mãn các tính chất sau:

i) $f(u) \geq 0 \quad , \forall u \in \mathbb{R}$

ii) $\int_{-\infty}^{+\infty} f(u)du = 1$

iii) $P(a \leq X \leq b) = \int_{a}^{b} f(u)du$

Chú ý: Khi X là BNN liên tục thì P{X = c} = 0 với bất kỳ số thực c nào.

3

Ý nghĩa của giá trị hàm mật độ xác suất *f*(*u*) trong phân phối xác suất của BNN liên tục *X*:

$$f(u) = \lim_{\triangle u \to 0^+} \frac{P\{u - \triangle u < X < u\}}{\triangle u} = \lim_{\triangle u \to 0^+} \frac{P\{u < X < u + \triangle u\}}{\triangle u}$$

tại mỗi điểm *u* thuộc tập giá trị của *X* mà các giới hạn này bằng nhau (nếu các giới hạn này khác nhau thì chúng lần lượt là giới hạn một bên của *f* (*u*) tại điểm *u*).

$f(u) = 0$ tại các điểm *u* không thuộc tập giá trị của *X*.

4.2 Hàm phân phối xác suất:

Định nghĩa: Hàm phân phối xác suất của biến ngẫu nhiên *X*, thường được kí hiệu là $F(t)$, là hàm số một biến thực sau đây:

$$F(t) = P(X \leq t) \quad , \forall t \in R$$

$F(t)$ cũng được gọi là "hàm phân phối tích lũy" hay "hàm xác suất tích lũy".

5

Tính chất của hàm phân phối xác suất:

i/ $F(t) \in [0,1] \quad , \quad F(-\infty) = \lim\limits_{t \to -\infty} F(t) = 0 \quad , \quad F(+\infty) = \lim\limits_{t \to +\infty} F(t) = 1$

ii/ $F(t)$ không giảm trên ℝ.

iii/ Với *X* là BNN liên tục thì $F(t)$ liên tục (hai phía) trên ℝ.

6

Liên hệ giữa hàm mật độ xác suất và hàm phân phối xác suất:

1/ $F(t) = \int_{-\infty}^{t} f(u)du$; $F'(t) = f(t)$ với $F(t)$ là hàm ppxs và $f(u)$ là hàm mđxs của BNN liên tục *X*.

2/ $P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b)$

$$= \int_{a}^{b} f(u)du$$

$$= F(b) - F(a)$$

7

4.3 Kỳ vọng, phương sai, phân vị của BNN liên tục:

Kỳ vọng của BNN liên tục *X* (kí hiệu là E[*X*]) được xác định như sau:

$$E\left[X\right] = \int_{-\infty}^{+\infty} u.f(u)du$$ với $f(u)$ là hàm mđxs của *X*.

Phương sai của BNN liên tục X (kí hiệu là D(X) hay Var(X)) được tính như sau:

$$D\left(X\right) \equiv Var(X) = E\left[\left(X - E[X]\right)^2\right] = E[X^2] - E^2[X]$$

__Chú ý:__ trung bình và phương sai của BNN liên tục cũng có đầy đủ các tính chất như trung bình và phương sai của BNN rời rạc.

$E\left[X^2\right] = \int_{-\infty}^{+\infty} u^2.f\left(u\right)du$ với *f*(*u*) là hàm mđsx của *X*.

Độ lệch tiêu chuẩn $\sigma_X \equiv \sigma(X) = \sqrt{D(X)}$

8

Phân vị thứ 100p (với $0 \le p \le 1$) trong phân phối xác suất của BNN liên tục X, kí hiệu η_p, là giá trị thực thỏa mãn $F(\eta_p) = P(X \le \eta_p) = p$,

hay tương đương là $$\int_{-\infty}^{\eta_p} f(u)\,du = p$$

với $F(t)$ là hàm phân phối xác suất và $f(u)$ là hàm mật độ xác suất của X.

Nhận xét: Khi $0 < p < 1$ thì chắc chắn η_p phải thuộc tập giá trị của X.

Ví dụ 1: Cho biến ngẫu nhiên liên tục X có hàm mật độ xác suất:

$$f(x) = \begin{cases} m(1-x^2) & ,\ khi\ \ -1 \le x \le 1 \\ 0 & ,\ khi\ \ x > 1\ hay\ x < -1 \end{cases}$$

a) Tìm (số thực) m và viết biểu thức hàm phân phối xác suất của X.
b) Tính E[X], D(X), σ(X).
c) Tính $P\left\{\left(X > \frac{1}{2}\right) \middle| \left(X \le \frac{3}{4}\right)\right\}$, $P\left\{\left(X > \frac{1}{2}\right) \middle| \left(X > \frac{1}{5}\right)\right\}$

Giải Ví dụ 1a:

$$\begin{cases} f(x) \ge 0, \forall x \in \mathbb{R} \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases} \Leftrightarrow \begin{cases} m\left(1-x^2\right) \ge 0, \forall x \in [-1;1] \\ \int_{-\infty}^{-1} 0dx + \int_{-1}^{1} m\left(1-x^2\right)dx + \int_{1}^{+\infty} 0dx = 1 \end{cases} \Leftrightarrow m = \frac{3}{4}$$

Vậy hàm mật độ xác suất của X là:

$$f(x) = \begin{cases} \frac{3}{4}(1-x^2) & , \ khi \ -1 \le x \le 1 \\ 0 & , \ khi \ x > 1 \ hay \ x < -1 \end{cases}$$

11

Giải Ví dụ 1a (tiếp):

Hàm phân phối xác suất của X là: $F(t) = \int_{-\infty}^{t} f(x)dx$

Nếu t < -1 thì $(-\infty;t] \subset (-\infty;-1] \Rightarrow F(t) = \int_{-\infty}^{t} f(x)dx = \int_{-\infty}^{t} 0dx = 0$

Nếu -1 ≤ t ≤ 1 thì $\begin{cases} (-\infty;t] = (-\infty;-1] \cup [-1;t] \\ [-1;t] \subset [-1;1] \end{cases} \Rightarrow F(t) = \int_{-\infty}^{t} f(x)dx = \int_{-\infty}^{-1} 0dx + \int_{-1}^{t} \frac{3}{4}\left(1-x^2\right)dx = \frac{2+3t-t^3}{4}$

Nếu 1 < t thì

$$\begin{cases} (-\infty;t] = (-\infty;-1] \cup [-1;1] \cup [1;t] \\ [1;t] \subset [1;+\infty] \end{cases} \Rightarrow F(t) = \int_{-\infty}^{t} f(x)dx = \int_{-\infty}^{-1} 0dx + \int_{-1}^{1} \frac{3}{4}\left(1-x^2\right)dx + \int_{1}^{t} 0dx = 1$$

Vậy: $$F(t) = \begin{cases} 0 & \text{khi } t < -1 \\ \frac{2+3t-t^3}{4} & \text{khi } -1 \le t \le 1 \\ 1 & \text{khi } t > 1 \end{cases}$$

12

Giải Ví dụ 1b:

$$E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\int_{-1}^{1} x.\frac{3}{4}(1-x^2)dx=0$$

$$E(X^2)=\int_{-\infty}^{+\infty} x^2 f(x)dx=\int_{-1}^{1} x^2.\frac{3}{4}(1-x^2)dx=2.\frac{3}{4}\left(\frac{x^3}{3}-\frac{x^5}{5}\right)\Bigg|_0^1=\frac{1}{5}$$

$$D(X)=E(X^2)-E^2(X)=\frac{1}{5}.$$

$$\sigma(X)=\sqrt{D(X)}=\frac{1}{\sqrt{5}}$$

13

Giải Ví dụ 1c:

$$P\left\{\left(X>\tfrac{1}{2}\right)\Big|\left(X\le\tfrac{3}{4}\right)\right\}=\frac{P\left\{\left(X>\tfrac{1}{2}\right).\left(X\le\tfrac{3}{4}\right)\right\}}{P\left\{X\le\tfrac{3}{4}\right\}}$$

$$=\frac{P\left\{\tfrac{1}{2}<X\le\tfrac{3}{4}\right\}}{P\left\{X\le\tfrac{3}{4}\right\}}=\frac{\int_{1/2}^{3/4} f(x)dx}{\int_{-\infty}^{3/4} f(x)dx}=\frac{\int_{1/2}^{3/4}\tfrac{3}{4}\left(1-x^2\right)dx}{\int_{-\infty}^{-1}0dx+\int_{-1}^{3/4}\tfrac{3}{4}\left(1-x^2\right)dx}=\ldots$$

$$P\left\{\left(X>\tfrac{1}{2}\right)\Big|\left(X>\tfrac{1}{5}\right)\right\}=\frac{P\left\{\left(X>\tfrac{1}{2}\right).\left(X>\tfrac{1}{5}\right)\right\}}{P\left\{X>\tfrac{1}{5}\right\}}=\frac{P\left\{X>\tfrac{1}{2}\right\}}{P\left\{X>\tfrac{1}{5}\right\}}=\frac{\int_{1/2}^{+\infty} f(x)dx}{\int_{1/5}^{+\infty} f(x)dx}=\frac{\int_{1/2}^{1}\tfrac{3}{4}\left(1-x^2\right)dx}{\int_{1/5}^{1}\tfrac{3}{4}\left(1-x^2\right)dx}=\ldots$$

Ví dụ 2: Cho X là BNN liên tục có hàm phân phối xác suất:

$$F(t)=\begin{cases} a-e^{-2t} & \text{khi } t\ge 0 \\ 0 & \text{khi } t<0 \end{cases}$$

Hãy tìm các hằng số *a, b* và hàm mật độ xác suất của X.

Giải Ví dụ 2:

F(*t*) cần phải liên tục trên $\mathbb{R}$. Mà khi $t < 0$ hay $t > 0$ thì *F*(*t*) đều trở thành hàm số sơ cấp liên tục, do đó chỉ cần thêm *F*(*t*) liên tục tại điểm $t = 0$.

$$\Rightarrow \lim_{t\to 0^-} F(t)=\lim_{t\to 0^+} F(t)=F(0) \quad \Rightarrow \quad \lim_{t\to 0^-} 0=\lim_{t\to 0^+}(a-e^{-2t})=a-e^0 \quad \Rightarrow \quad a=1$$

Vậy $F(t)=\begin{cases} 1-e^{-2t} & \text{khi } t\ge 0 \\ 0 & \text{khi } t<0 \end{cases}$ $\Rightarrow$ Hàm mđxs: $f(t)=F'(t)=\begin{cases} 2e^{-2t} & \text{khi } t>0 \\ undefined & \text{khi } t=0 \\ 0 & \text{khi } t<0 \end{cases}$

4.4 PHÂN PHỐI ĐỀU

Định nghĩa: Biến ngẫu nhiên liên tục *X* có phân phối đều trên [*a*;*b*] khi hàm mật độ xác suất của *X* có biểu thức như sau:

$$f_X(u)=\begin{cases} \dfrac{1}{b-a} & \text{khi } u\in[a;b] \\ 0 & \text{khi } u\notin[a;b] \end{cases}$$

Nhận xét: Khi *X* có phân phối đều trên [*a*;*b*] thì *X* nhận mỗi giá trị thực thuộc [*a*;*b*] với cùng một mức xác suất.

Tính chất: $E[X]=\dfrac{a+b}{2}$; $D(X)=\dfrac{(b-a)^2}{12}$

16

Ví dụ: Bạn A có khả năng như nhau để đến trường tại các thời điểm trong khoảng thời gian từ 6:40 AM đến 7:00 AM hàng ngày. Tính xác suất trong một ngày bất kỳ thì thấy bạn A đến trường trước 6:55 AM.

Giải:
Đặt *X* = khoảng thời gian (phút) tính từ lúc 6:40 AM đến thời điểm bạn A đến trường.
(***Nhận xét:*** *Vì bạn A có khả năng như nhau để đến trường tại các thời điểm trong khoảng thời gian từ 6:40 AM đến 7:00 AM hàng ngày nên X có cùng mức xác suất để nhận mỗi giá trị thực thuộc [0 ; 20]*)

$\Rightarrow$ *X* có phân phối đều trên [0 ; 20]

$\Rightarrow$ Hàm mật độ xác suất của *X* là: $f_X(u) = \begin{cases} \dfrac{1}{20} & \text{khi } u \in [0;20] \\ 0 & \text{khi } u \notin [0;20] \end{cases}$

$$\Rightarrow \quad P\{X < 15\} = \int_{-\infty}^{15} f_X(u)\,du = \int_0^{15} \frac{1}{20}\,du = 0,75$$

17

4.6 PHÂN PHỐI MŨ

Định nghĩa: Biến ngẫu nhiên liên tục *X* có phân phối mũ tham số λ (λ > 0) khi hàm mật độ xác suất của *X* có biểu thức như sau:

$$f_X(u) = \begin{cases} \lambda.\mathrm{e}^{-\lambda u} & \text{khi } u \ge 0 \\ 0 & \text{khi } u < 0 \end{cases}$$

Tính chất: 1) $E[X] = \dfrac{1}{\lambda}$; $D(X) = \dfrac{1}{\lambda^2}$

2) $P\{(X > k + t) \,|\, (X > t)\} = P\{X > k\}$ (tính không nhớ)

18

Lưu ý: Khi X có phân phối mũ tham số λ thì:

- Hàm phân phối xác suất của X là $F_X(t) = P\{X \le t\} = \begin{cases} 1 - e^{-\lambda t} & \text{khi } t \ge 0 \\ 0 & \text{khi } t < 0 \end{cases}$

- Hàm phân phối bù của X là $F_X^c(t) = P\{X > t\} = 1 - F_X(t) = \begin{cases} e^{-\lambda t} & \text{khi } t \ge 0 \\ 1 & \text{khi } t < 0 \end{cases}$

Ví dụ: Tuổi thọ (năm) của một loại sản phẩm là biến ngẫu nhiên X có phân phối mũ với giá trị trung bình là 3 năm.

a) Nếu các sản phẩm có tuổi thọ không quá 1 năm được nhà sản xuất bảo hành thì xác suất cần bảo hành với mỗi sản phẩm loại này là bao nhiêu?

b) Nhà sản xuất cần đặt ra thời hạn bảo hành là bao nhiêu để khả năng mỗi sản phẩm loại này phải bảo hành không vượt quá 10%.

Giải: X là biến ngẫu nhiên có phân phối mũ với tham số $\lambda = \frac{1}{E[X]} = \frac{1}{3}$

a) $P\{X \le 1\} = 1 - e^{-\lambda.1} = 1 - e^{\frac{-1}{3}} \approx 0,2835 = 28,35\%$

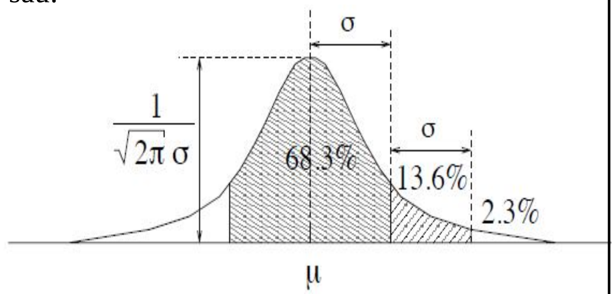
b) Gọi u (năm) là thời hạn bảo hành sản phẩm, ta có khả năng một sản phẩm phải bảo hành không vượt quá 10% khi $P(X \le u) \le 10\%$

$\Leftrightarrow F_X(u) \le 10\%$ $\Leftrightarrow$ $1 - e^{-\frac{u}{3}} \le 0,1$ $\Leftrightarrow$ $u \le -3ln0,9$ năm ≈ 115 ngày

4.5 PHÂN PHỐI CHUẨN

Định nghĩa: Biến ngẫu nhiên liên tục X có phân phối chuẩn tham số μ ($\mu \in \mathbf{R}$) và σ^2 ($\sigma^2 \in \mathbf{R}^+$) khi hàm mật độ xác suất của X có biểu thức như sau:

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}, \quad \forall t \in (-\infty; \infty)$$

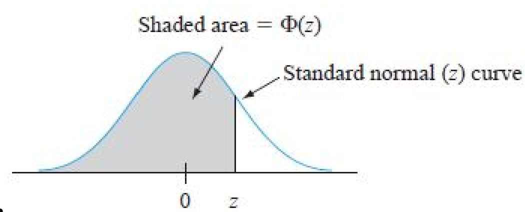


Kí hiệu: $X \sim N(\mu; \sigma^2)$

Tính chất: 1) $E[X] = \mu$; $D(X) = \sigma^2$; $\sigma(X) = \sigma = \sqrt{\sigma^2}$

2) Nếu có các BNN độc lập $X_i \sim N(\mu_i; \sigma_i^2),\ i = \overline{1;n}$, thì $X = \sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n}\mu_i ; \sum_{i=1}^{n}\sigma_i^2\right)$

Phân phối chuẩn tắc $N(0\ ;\ 1)$:



Biến ngẫu nhiên liên tục Z ~ N(0 ; 1) thì có:

+ Hàm mật độ xác suất là $f_Z(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$,

+ Hàm phân phối xác suất là $\Phi(z) = P\{Z \le z\} = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} dt$, $\forall z \in (-\infty; \infty)$

+ Hàm phân phối bù là $Q(z) = P\{Z > z\} = 1 - \Phi(z) = \Phi(-z)$, $\forall z \in (-\infty; \infty)$

+ Các giá trị đặc trưng: $E[Z] = 0$; $D(Z) = 1$

Phân vị mức (ý nghĩa) α, hay cũng là phân vị thứ $100(1-\alpha)$, của phân phối chuẩn tắc, kí hiệu z_α, là giá trị thực thỏa mãn $\Phi(z_\alpha) = P(Z \le z_\alpha) = 1 - \alpha$.

Cách bấm máy tính bỏ túi Casio-fx570ES Plus để tra giá trị của các hàm Φ(*z*) và *Q*(*z*):

MODE	3	1	AC				
SHIFT	1	5	1	nhập z	)	=	(được Φ(z))
SHIFT	1	5	3	nhập z	)	=	(được *Q*(z))

Chú ý: với z ≤ - 4 thì có thể xem như Φ(z) = 0;
với z ≥ 4 thì có thể xem như Φ(z) = 1.

Chính tắc hóa phân phối Chuẩn và cách tính xác suất theo hàm Φ(*z*):

Khi $X \sim N\left(\mu\,;\sigma^2\right)$ thì $Z = \dfrac{X-\mu}{\sigma} \sim N\left(0\,;1\right)$. Do đó, ∀[a ; b] ⊂ (-∞ ; ∞) ta có:

$$\begin{aligned} P\{a \le X \le b\} &= P\left\{\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right\} \\ &= P\left\{\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right\} \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

23

Xấp xỉ phân phối Nhị thức *B*(*n* ; *p*) bởi phân phối Chuẩn *N*(μ ; σ^2):

Định lý DeMoivre-Laplace: Với $S_{n,p}$ là biến ngẫu nhiên có phân phối Nhị thức với tham số *n* và *p* thì mỗi khi cố định *p* tại một giá trị thuộc (0 ; 1) và chọn bất kỳ $c \in \mathbb{R}$ ta đều có:

$$\lim_{n\to\infty} P\left\{\frac{S_{n,p}-np}{\sqrt{np(1-p)}} \le c\right\} = \Phi(c).$$

Suy ra từ Định lý DeMoivre-Laplace rằng khi biến ngẫu nhiên *X* ~ *B*(*n* ; *p*) mà *n* đủ lớn và *p* không quá gần 0 đồng thời cũng không quá gần 1 thì có thể xem như $X \sim N\left(\mu\,;\sigma^2\right)$ với $\mu = n.p$ và $\sigma^2 = n.p.(1-p)$.

Xấp xỉ trên rất tốt khi $n.p \ge 10$ đồng thời $n.p.(1-p) \ge 10$.

24

Ví dụ 1: Điểm thi môn Toán của trường đại học A là biến ngẫu nhiên có phân phối chuẩn với trung bình là 6,78 và độ lệch chuẩn là 0,46. Lấy ngẫu nhiên 200 bài thi môn Toán của sinh viên trường A. Tính xác suất trong số 200 bài lấy ra số bài có điểm từ 6 đến 8 điểm không dưới 25 bài.

Giải: Đặt X là điểm một bài thi môn Toán của sinh viên trường A.

Xác suất để một bài thi môn Toán của sinh viên trường A có điểm trong khoảng từ 6 đến 8 là:

$$P(6 < X < 8) = P\left(-\frac{39}{23} < \frac{X - 6,78}{0,46} < \frac{61}{23}\right) = \Phi\left(\frac{61}{23}\right) - \Phi\left(-\frac{39}{23}\right) = 0,951$$

Đặt Y = số lượng bài thi đạt từ 6 đến 8 điểm trong 200 bài thi Toán của sv trường A.
Ta có $Y \sim B(200 ; 0,951) \Rightarrow$ có thể xem như $Y \sim N(190,2 ; 9,3198)$

$$\Rightarrow P(200 \geq Y \geq 25) = P\left(3.0528 \geq \frac{Y - 190,2}{\sqrt{9,3198}} \geq -54,1136\right)$$

$$= \Phi(3.0528) - \Phi(-54,1136) \approx 0.9989$$

25

Ví dụ 2: Tuổi thọ (năm) của một loại máy là biến ngẫu nhiên X có phân phối mũ với giá trị trung bình là 3 năm. Biết các máy có tuổi thọ không quá 1 năm được nhà sản xuất bảo hành, tính xác suất trong 100 máy loại này có từ 20 đến 40 máy cần bảo hành.

Giải: Trong ví dụ cuối mục "4.6 PHÂN PHỐI MŨ" ta đã tính: $P\{X \leq 1\} \approx 0,2835$
Xác suất này cũng chính là xác suất cần bảo hành của mỗi máy loại này.

Đặt Y = số lượng máy cần bảo hành trong 100 máy loại này.
Ta có Y ~ B(100 ; 0,2835) $\Rightarrow$ có thể xem như Y ~ N(28,35 ; 20,3128)

$$\Rightarrow P(20 \leq Y \leq 40) = P\left(\frac{-8,35}{\sqrt{20,3128}} \leq \frac{Y - 28,35}{\sqrt{20,3128}} \leq \frac{11,65}{\sqrt{20,3128}}\right)$$

$$= \Phi\left(\frac{11,65}{\sqrt{20,3128}}\right) - \Phi\left(\frac{-8,35}{\sqrt{20,3128}}\right) = \ldots$$

26