

Mục lục

Lời mở đầu	1
1 Kiến thức cơ bản về xác suất	2
1.1 Các tiên đề cơ bản	2
1.2 Xác suất có điều kiện và sự độc lập	3
1.3 Biến ngẫu nhiên	4
1.4 Giá trị kỳ vọng và phương sai	6
1.5 Kỳ vọng và phương sai của tổng các biến ngẫu nhiên	9
1.6 Hàm moment tổng quát và biến đổi Laplace	18
1.7 Kỳ vọng có điều kiện	21
1.8 Biến ngẫu nhiên phân phối mũ	27
1.9 Các định lý về giới hạn	35
1.9.1 Luật số lớn yếu	35
1.9.2 Luật số lớn mạnh	35
1.9.3 Định lý giới hạn trung tâm	36
1.9.4 Định lý giới hạn trung tâm cho khối đa diện Gauss	37
1.10 Bất đẳng thức Markov	38
1.11 Bất đẳng thức Chebyshev	40
1.11.1 Phát biểu theo lý thuyết độ đo	41
1.11.2 Phát biểu theo xác suất	41
2 Bất đẳng thức Chernoff và ứng dụng	43
2.1 Bất đẳng thức Chernoff	43
2.1.1 Cận Chernoff	43
2.1.2 Bất đẳng thức Chernoff	44
2.2 Cận Chernoff cho tổng các biến ngẫu nhiên Bernoulli độc lập.	45

3	Xấp xỉ Poisson	47
3.1	Biến ngẫu nhiên Bernoulli	47
3.1.1	Biến ngẫu nhiên Bernoulli và biến ngẫu nhiên nhị thức	47
3.1.2	Các tính chất của biến ngẫu nhiên nhị thức	49
3.2	Biến ngẫu nhiên Poisson	50
3.2.1	Biến ngẫu nhiên Poisson và mẫu Poisson	50
3.2.2	Tổng Poisson của các biến ngẫu nhiên	54
3.3	Đánh giá độ chính xác khi xấp xỉ tổng các biến ngẫu nhiên Bernoulli độc lập bởi một biến ngẫu nhiên Poisson	56
	Kết luận	60
	Tài liệu tham khảo	61

Lời mở đầu

Toán học là cơ sở nền tảng cho sự ra đời và phát triển của Tin học, và ngày nay cùng với sự phát triển của xã hội, vai trò của Toán học ngày càng trở nên quan trọng trong các lĩnh vực ứng dụng thực tế, đặc biệt là trong Tin học. Do đó, tác giả đã được giao đề tài “**Ứng dụng xác suất trong Tin học**” để làm luận văn tốt nghiệp thạc sĩ. Luận văn gồm 3 chương:

Chương 1 Kiến thức cơ bản về xác suất. Chương này giới thiệu các kiến thức cơ sở về xác suất. Đồng thời, tác giả có giới thiệu một số kết quả mới về định lý giới hạn trung tâm cho khối đa diện Gauss.

Chương 2 Bất đẳng thức Chernoff và ứng dụng. Chương này giới thiệu một cách tổng quát về bất đẳng thức Chernoff và ứng dụng của nó, đồng thời cũng giới thiệu cách tính cận Chernoff cho tổng các biến ngẫu nhiên Bernoulli độc lập.

Chương 3 Xấp xỉ Poisson. Chương này giới thiệu kỹ hơn về biến ngẫu nhiên nhị thức và biến ngẫu nhiên Poisson cùng các tính chất đặc biệt của các biến ngẫu nhiên này.

Do kiến thức còn có những hạn chế và thiếu sót, vì vậy tác giả kính mong nhận được các ý kiến đóng góp của các thầy cô, cùng bạn bè đồng nghiệp để bản luận văn này được hoàn thiện hơn.

Qua đây, tác giả xin được gửi lời biết ơn sâu sắc nhất đến người hướng dẫn khoa học **GS TS. Nguyễn Duy Tiến** đã tạo mọi điều kiện và chỉ bảo tận tình trong suốt quá trình tác giả nghiên cứu để hoàn thành được bản luận văn. Đồng thời, tác giả cũng xin gửi lời biết ơn chân thành đến các thầy cô trong khoa Toán - Cơ - Tin trường Đại học KHTN - Đại học Quốc Gia Hà Nội đã truyền đạt những kiến thức quý báu và tạo mọi điều kiện để tác giả hoàn thành quá trình nghiên cứu luận văn. Cuối cùng, tác giả xin được gửi lời cảm ơn chân thành đến tập thể lớp Cao học khóa 2006 – 2008, bạn bè và người thân đã luôn động viên, giúp đỡ tác giả hoàn thành bản luận văn này.

Hà Nội, năm 2009

Học viên

Nguyễn Thị Hà Tĩnh

Chương 1

Kiến thức cơ bản về xác suất

1.1 Các tiên đề cơ bản

Mặc dù ra đời từ thế kỷ *XVII*, nhưng mãi đến năm 1929 nhà Toán học người Nga, A.N.Kolmogorov mới xây dựng được cơ sở Toán học chắc chắn cho lý thuyết xác suất bằng cách đề xuất hệ tiên đề cho lý thuyết xác suất dựa trên lý thuyết tập hợp và độ đo.

Xét một phép thử nào đó. Gọi $\mathcal{S}$ là tập tất cả các kết quả có thể có của phép thử, tập này được gọi là **không gian mẫu** (sample space). Một biến cố A là một tập con của không mẫu $\mathcal{S}$. Ta gán cho biến cố A một số $P(A)$, được gọi là **xác suất** (probability) của biến cố A , và số này phải thỏa mãn các tiên đề sau.

Tiên đề 1.1.1.

$$0 \leq P(A) \leq 1$$

Tiên đề 1.1.2.

$$P(S) = 1$$

Tiên đề 1.1.3. Nếu $A_1, A_2, \dots$ là dãy các biến cố xung khắc từng đôi một bất kỳ, thì

$$P\left(\bigcup_{i=1}^{i=\infty} A_i\right) = \sum_{i=1}^{i=\infty} P(A_i)$$

Từ hệ tiên đề trên, ta có thể suy ra một số tính chất cơ bản của xác suất.

- (a) Nếu $A \subset B$ thì $P(A) \leq P(B)$
- (b) $P(A^C) = 1 - P(A)$, với $A^C = S \setminus A$ là biến cố đối của A
- (c) $P(A \cup B) = P(A) + P(B) - P(AB)$

1.2 Xác suất có điều kiện và sự độc lập

Xét một phép thử là tung một đồng xu hai lần, mỗi lần tung có thể xuất hiện mặt ngửa hoặc mặt sấp. Không gian mẫu của phép thử này là

$$\mathcal{S} = \{(n, n), (n, s), (s, n), (s, s)\}$$

Giả sử rằng các kết quả trên có khả năng xuất hiện như nhau, và vì thế nên có xác suất là $1/4$. Giờ ta giả sử rằng lần tung thứ nhất xuất hiện mặt sấp: với điều kiện này, xác suất để cả hai lần tung đều xuất hiện mặt sấp là bao nhiêu? Để tính được xác suất này, ta xét như sau: với điều kiện lần tung thứ nhất xuất hiện mặt sấp thì lần tung thứ hai sẽ có nhiều nhất hai khả năng, cụ thể là $\{(s, n), (s, s)\}$. Vì hai trường hợp này có khả năng xảy ra như nhau nên chúng có xác suất bằng nhau, và do tổng các xác suất có điều kiện này phải có tổng bằng 1 nên với điều kiện lần tung thứ nhất xuất hiện mặt sấp thì xác suất để cả hai lần tung đều xuất hiện mặt sấp là $1/2$.

Nếu ký hiệu A là biến cố cả hai lần tung đều xuất hiện mặt sấp và B là biến cố lần tung đầu tiên xuất hiện mặt sấp, thì xác suất ta vừa tính ở trên được gọi là xác suất có điều kiện của A với điều kiện B đã xảy ra, kí hiệu là $P(A | B)$. Một công thức tổng quát cho $P(A | B)$ có thể đạt được theo phương pháp tương tự như đã sử dụng trong ví dụ về tung đồng xu ở trên. Cụ thể, nếu biến cố B đã xảy ra thì để biến cố A cũng xảy ra, ta cần phải có biến cố A và B cùng xảy ra; nghĩa là biến cố AB xảy ra. Vì các khả năng đều xét trong trường hợp B đã xảy ra, nên B trở thành một không gian mẫu mới và vì thế xác suất để biến cố AB xảy ra trong không gian mẫu mới này sẽ bằng xác suất của AB chia cho xác suất của B . Nghĩa là, với $P(B) > 0$

$$P(A | B) = \frac{P(AB)}{P(B)}$$

Tương tự như ví dụ trên, xác suất có điều kiện $P(A | B)$ của A với điều kiện của B đã xảy ra, nói chung không bằng $P(A)$, xác suất không có điều kiện của A . Trong trường hợp đặc biệt, khi $P(A | B) = P(A)$ thì ta nói rằng A độc lập với B . Từ định nghĩa của $P(A | B)$, ta thấy rằng điều kiện để A độc lập với B là

$$P(AB) = P(A)P(B)$$

Khái niệm độc lập có thể mở rộng cho hơn hai biến cố như sau: các biến cố $A_1, \dots, A_n$ được gọi là độc lập với nhau nếu

$$P(A_{i_1}, \dots, A_{i_r}) = P(A_{i_1})P(A_{i_2})\dots P(A_{i_r})$$

với $i_1, \dots, i_r$ là một tập con bất kỳ của $\{1, \dots, n\}$.

1.3 Biến ngẫu nhiên

Xét một phép thử có không gian mẫu $\mathcal{S}$ mà trên đó các xác suất được định nghĩa theo các biến cố của $\mathcal{S}$. Một **biến ngẫu nhiên** X là một hàm thực của các phần tử của không gian mẫu. Cho một tập số thực $\mathcal{C}$ bất kỳ, xác suất để X nhận giá trị trong tập hợp $\mathcal{C}$ là bằng với xác suất của biến cố $X^{-1}(\mathcal{C})$. Nghĩa là,

$$P(X \in \mathcal{C}) = P(X^{-1}(\mathcal{C}))$$

với $X^{-1}(\mathcal{C})$ là biến cố bao gồm tất cả các kết quả $s \in \mathcal{S}$ sao cho $X(s) \in \mathcal{C}$.

Hàm phân phối

Hàm phân phối F của biến ngẫu nhiên X được định nghĩa bởi

$$F(x) = P(X \leq x) = P\{X \in (-\infty, x]\}$$

với mọi số thực x . Ta kí hiệu $\bar{F}$ để biểu diễn $1 - F$, tức là

$$\bar{F}(x) = 1 - F(x) = P\{X > x\}$$

Một biến ngẫu nhiên được gọi là **rời rạc** nếu tập giá trị có thể có của nó là một tập hữu hạn hoặc vô hạn đếm được. Cho X là một biến ngẫu nhiên rời rạc, ta định nghĩa **hàm khối lượng xác suất** $p(x)$ của X là

$$p(x) = P\{X = x\}$$

Nếu $x_i, i \geq 0$, biểu diễn các giá trị có thể có của X thì

$$F(x) = \sum_{i: x_i \leq x} p(x_i)$$

Một biến ngẫu nhiên được gọi là **liên tục** nếu tồn tại một hàm $f(x)$, được gọi là **hàm mật độ xác suất** của X , sao cho với một tập số $\mathcal{C}$ bất kỳ

$$P\{X \in \mathcal{C}\} = \int_{\mathcal{C}} f(x)dx$$

Mối quan hệ giữa hàm phân phối F và hàm mật độ xác suất f của biến ngẫu nhiên liên tục X được biểu diễn như sau

$$F(y) = P\{X \in (-\infty, y]\} = \int_{-\infty}^y f(x)dx$$

Lấy đạo hàm hai vế, ta có

$$F'(y) = f(y)$$

Ta sẽ có một biểu diễn trực giác hơn cho hàm mật độ xác suất bằng chú ý rằng

$$P\{a - \varepsilon/2 < X < a + \varepsilon/2\} = \int_{a-\varepsilon/2}^{a+\varepsilon/2} f(x)dx \approx \varepsilon f(a)$$

khi ε là nhỏ. Nói cách khác, xác suất để X nhận giá trị trong một khoảng nào đó có độ dài ε quanh điểm a là xấp xỉ $\varepsilon f(a)$, như vậy $f(a)$ là một độ đo để xem biến ngẫu nhiên X gần a cỡ nào.

Trong nhiều trường hợp ta không chỉ quan tâm hàm phân phối của từng biến ngẫu nhiên, mà còn là mối quan hệ giữa hai hoặc nhiều hơn các biến ngẫu nhiên. Để định rõ mối quan hệ giữa hai biến ngẫu nhiên X và Y , ta định nghĩa hàm phân phối đồng thời của chúng như sau:

$$F(x, y) = P\{X \leq x, Y \leq y\}$$

Nghĩa là, $F(x, y)$ là xác suất để cả $X \leq x$ và $Y \leq y$. Hàm phân phối của từng biến ngẫu nhiên X và Y có thể đạt được từ hàm phân phối đồng thời bằng cách

$$F_X(x) = \lim_{y \rightarrow \infty} F(x, y), \quad F_Y(y) = \lim_{x \rightarrow \infty} F(x, y)$$

với F_W là kí hiệu hàm phân phối của biến ngẫu nhiên W .

Nếu cả X và Y là biến ngẫu nhiên rời rạc, thì ta định nghĩa hàm khối lượng xác suất đồng thời của chúng là

$$p(x, y) = P\{X = x, Y = y\}$$

Ta nói rằng X và Y là đồng thời liên tục với hàm mật độ xác suất đồng thời $f(x, y)$, nếu với các tập số thực bất kỳ $\mathcal{C}$ và $\mathcal{D}$

$$P\{X \in \mathcal{C}, Y \in \mathcal{D}\} = \int_{\mathcal{C}} \int_{\mathcal{D}} f(x, y) dy dx$$

Nếu X và Y là đồng thời liên tục thì

$$P\{X \in \mathcal{C}\} = P\{X \in \mathcal{C}, Y \in (-\infty, \infty)\} = \int_{\mathcal{C}} \int_{-\infty}^{\infty} f(x, y) dy dx$$

Điều này cho thấy X là liên tục với hàm mật độ

$$f_X(x) = \int_{-\infty}^{\infty} f(y) dy$$

Hai biến ngẫu nhiên được gọi là **độc lập** nếu với các tập số thực bất kỳ $\mathcal{C}$ và $\mathcal{D}$

$$P\{X \in \mathcal{C}, Y \in \mathcal{D}\} = P\{X \in \mathcal{C}\}P\{Y \in \mathcal{D}\}$$

Đẳng thức trên được suy ra từ

$$F(x, y) = F(x)F(y)$$

với mọi x và y . Hơn nữa, với mọi x và y , các biến ngẫu nhiên X và Y sẽ là độc lập trong trường hợp rời rạc nếu

$$P\{X = x, Y = y\} = P\{X = x\}P\{Y = y\}$$

và trong trường hợp đồng thời liên tục nếu

$$f(x, y) = f_X(x)f_Y(y)$$

Ta cũng có thể định nghĩa hàm phân phối đồng thời của một số bất kỳ các biến ngẫu nhiên $X_1, \dots, X_n$ bởi

$$F(x_1, \dots, x_n) = P\{X_1 \leq x_1, \dots, X_n \leq x_n\}$$

với mọi $x_1, \dots, x_n$.

1.4 Giá trị kỳ vọng và phương sai

Trong lý thuyết xác suất (và đặc biệt là đánh bạc), **giá trị kỳ vọng** (expectation) của một biến ngẫu nhiên là tổng xác suất của mỗi kết quả có thể của thử nghiệm nhân với giá trị của kết quả đó. Như vậy, nó biểu diễn giá trị trung bình mà người ta "mong đợi" thắng cược nếu đặt cược liên tục nhiều lần với khả năng thắng cược là như nhau. Chú ý rằng, bản thân giá trị đó có thể không được mong đợi theo nghĩa thông thường, nó có thể ít có khả năng xảy ra hoặc không thể xảy ra. Một trò chơi hoặc một tình huống trong đó giá trị kỳ vọng cho người chơi bằng 0 (không có thắng hoặc thua chung cuộc) được gọi là một "trò chơi công bằng" (fair game).

Định nghĩa toán học

Thông thường, nếu X là một biến ngẫu nhiên xác định trên một không gian xác suất $(\Omega, \mathbb{P})$, thì giá trị kỳ vọng của X , ký hiệu là $\mathbb{E}[X]$, được định nghĩa như sau

$$\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P}$$

trong đó sử dụng tích phân Lebesgue. Lưu ý rằng không phải mọi biến ngẫu nhiên đều có một giá trị kỳ vọng, do có thể không tồn tại tích phân (ví dụ phân bố Cauchy). Hai

biến ngẫu nhiên có cùng phân bố xác suất sẽ có giá trị kỳ vọng bằng nhau.

Nếu X là một biến ngẫu nhiên rời rạc nhận các giá trị $x_i, i \geq 1$, thì kỳ vọng của X được tính như sau

$$\mathbb{E}[X] = \sum_i x_i \mathbb{P}\{X = x_i\}$$

Ví dụ 1.1 a. Biến ngẫu nhiên chỉ tiêu của biến cố A , ký hiệu I_A hoặc $I\{A\}$ được xác định bằng công thức sau

$$I_A = \begin{cases} 1 & \text{nếu } A \text{ xảy ra} \\ 0 & \text{nếu } A^C \text{ xảy ra} \end{cases}$$

Từ định nghĩa của I_A , ta suy ra

$$\mathbb{E}[I_A] = 1 \cdot \mathbb{P}(A) + 0 \cdot \mathbb{P}(A^C) = \mathbb{P}(A)$$

Một biến ngẫu nhiên mà chỉ nhận hai giá trị 0 hoặc 1 thường được gọi là biến ngẫu nhiên Bernoulli. Như vậy, giá trị kỳ vọng của một biến ngẫu nhiên Bernoulli chính là xác suất để biến ngẫu nhiên nhận giá trị 1.

□

Ví dụ 1.1 b. Xét một dãy các phép thử độc lập, mỗi phép thử sẽ thành công với xác suất $p, 0 < p < 1$, hoặc thất bại với xác suất $1 - p$. Nếu X biểu diễn cho số thứ tự của phép thử mà là thành công đầu tiên, thì X được gọi là biến ngẫu nhiên có phân phối hình học với tham số p . Hãy tính giá trị kỳ vọng của X .

Lời giải. X sẽ bằng n nếu tất cả $n - 1$ phép thử đầu tiên đều thất bại và phép thử thứ n là thành công. Vì các phép thử là độc lập nên

$$\mathbb{P}\{X = n\} = p(1 - p)^{n-1}, n \geq 1$$

Do đó

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} np(1 - p)^{n-1} = 1/p$$

với đẳng thức trên có được là do, với $0 < a < 1$,

$$\sum_{n=1}^{\infty} na^{n-1} = \sum_{n=1}^{\infty} \frac{d}{da} a^n = \frac{d}{da} \sum_{n=1}^{\infty} a^n = \frac{1}{(1 - a)^2}$$

□

Nếu X là một biến ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$, thì giá trị kỳ vọng của X được tính như sau

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Giờ giả sử ta quan tâm đến giá trị kỳ vọng của biến ngẫu nhiên $g(X)$ với $g(\cdot)$ là một hàm tùy ý nào đó. Vì $g(X)$ sẽ nhận giá trị $g(x)$ khi X nhận giá trị x , nên có thể thấy ngay rằng $\mathbb{E}[g(X)]$ là trung bình trọng lượng của các giá trị có thể có của $g(x)$, với trọng lượng của $g(x)$ chính là xác suất (hoặc mật độ xác suất trong trường hợp liên tục) để X nhận giá trị x . Kết quả trên thể hiện trong mệnh đề sau.

Mệnh đề 1.4.1.

$$\mathbb{E}[g(X)] = \begin{cases} \sum g(x)\mathbb{P}\{X = x\} & \text{nếu } X \text{ rời rạc} \\ \int_{-\infty}^{\infty} g(x)f(x)dx & \text{nếu } X \text{ liên tục với hàm mật độ } f \end{cases}$$

Nếu hàm $g(x) = ax + b$, mệnh đề 1.1.1 chỉ ra rằng với các hằng số a, b

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

Phương sai

Phương sai của một biến ngẫu nhiên X , ký hiệu là $Var(X)$, được định nghĩa như sau

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

Ví dụ 1.1 c. Tìm phương sai của biến ngẫu nhiên Bernoulli X có $\mathbb{E}[X] = p$.

Lời giải. Vì

$$(X - \mathbb{E}[X])^2 = \begin{cases} (1 - p)^2 & \text{với xác suất } p \\ (0 - p)^2 & \text{với xác suất } 1 - p \end{cases}$$

nên

$$Var(X) = (1 - p)^2p + p^2(1 - p) = p(1 - p)$$

□

Từ trên ta có phương sai của X là

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

và với các hằng số a và b bất kỳ,

$$Var(aX + b) = a^2Var(X)$$

Trong lý thuyết xác suất, phương sai của một biến ngẫu nhiên là một độ đo sự phân tán của biến ngẫu nhiên quanh giá trị trung bình của nó, nó hàm ý các giá trị của biến

ngẫu nhiên đó thường ở cách giá trị kỳ vọng bao xa.

Phương sai của biến ngẫu nhiên giá trị thực là moment trung tâm bậc hai. Phương sai của một biến ngẫu nhiên là bình phương của độ lệch tiêu chuẩn. Định nghĩa này áp dụng cho cả biến ngẫu nhiên rời rạc và liên tục.

1.5 Kỳ vọng và phương sai của tổng các biến ngẫu nhiên

Một mở rộng của Mệnh đề (1.4.1) cho trường hợp hai chiều là: nếu g là một hàm của các biến ngẫu nhiên X và Y , thì

$$\mathbb{E}[g(X, Y)] = \begin{cases} \sum_x \sum_y g(x, y) \mathbb{P}\{X = x, Y = y\} & \text{nếu } X, Y \text{ rời rạc} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy & \text{nếu } X, Y \text{ đồng thời liên tục với mật độ } f \end{cases}$$

Nếu cho $g(x, y) = x + y$ thì từ trên dễ dàng thấy

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Tổng quát, với một số hữu hạn bất kỳ các biến ngẫu nhiên, bằng phương pháp quy nạp, ta có

$$\mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}[X_i] \quad (1.1)$$

Ví dụ 1.1 d. Thuật toán sủi bọt (The Bubble Sort)

Giả sử ta có một tập gồm n giá trị phân biệt $x_1, \dots, x_n$ mà ta muốn đặt chúng theo thứ tự tăng dần, hay còn gọi là sắp xếp chúng. Sủi bọt là một thuật toán ta có thể dùng. Ta sẽ bắt đầu với một dãy số cho trước nào đó, tuần tự đi qua các phần tử của dãy, và sẽ đổi chỗ cặp phần tử nào mà không theo thứ tự tăng dần như ta mong muốn. Cụ thể, đầu tiên giá trị thứ nhất và thứ hai sẽ được so sánh với nhau, ta sẽ đổi chỗ chúng nếu giá trị thứ hai là nhỏ hơn; sau đó giá trị mới ở vị trí thứ hai sẽ được so sánh với giá trị ở vị trí thứ ba, và hai giá trị này sẽ đổi chỗ nếu giá trị sau nhỏ hơn giá trị trước,... Tiếp tục như vậy cho đến khi giá trị cuối cùng trong dãy được so sánh và cũng có thể đổi chỗ nếu cần. Quá trình này được lặp lại và được tiếp tục cho đến khi tất cả các giá trị của dãy đều được sắp xếp. Chẳng hạn, nếu dãy ban đầu được cho như sau

5 3 8 7 0 9 6 4 1

thì các lần sắp xếp liên tiếp trong lượt so sánh đầu tiên sẽ như sau (với gạch ngang trên đầu là chỉ giá trị đang được so sánh trực tiếp với giá trị đứng ngay sau):

$$\begin{array}{cccccccc}
 3 & \bar{5} & 8 & 7 & 0 & 9 & 6 & 4 & 1 \\
 3 & 5 & \bar{8} & 7 & 0 & 9 & 6 & 4 & 1 \\
 3 & 5 & 7 & \bar{8} & 0 & 9 & 6 & 4 & 1 \\
 3 & 5 & 7 & 0 & \bar{8} & 9 & 6 & 4 & 1 \\
 3 & 5 & 7 & 0 & 8 & \bar{9} & 6 & 4 & 1 \\
 3 & 5 & 7 & 0 & 8 & 6 & \bar{9} & 4 & 1 \\
 3 & 5 & 7 & 0 & 8 & 6 & 4 & \bar{9} & 1 \\
 3 & 5 & 7 & 0 & 8 & 6 & 4 & 1 & \bar{9}
 \end{array}$$

Để thấy rằng, sau lượt so sánh đầu tiên, giá trị lớn nhất sẽ chính là giá trị nằm ở vị trí cuối cùng. Như vậy, ở lượt so sánh thứ hai ta không cần phải xét đến giá trị cuối cùng này, và sau lượt so sánh thứ hai ta sẽ có được hai giá trị lớn nhất của dãy nằm ở hai vị trí cuối cùng. Kết quả này cũng tương tự cho các lần so sánh tiếp theo và cứ như thế, các giá trị được sắp xếp theo giá trị tăng dần. Thuật toán này được gọi là **Thuật toán sủi bọt** vì cách di chuyển các giá trị nhỏ hơn lên phía trên gọi cho ta hình ảnh giống các bọt bong bóng nổi lên trong nước.

Thuật toán này kết thúc ngay cả khi không có một sự đổi chỗ nào xảy ra ở ngay lượt so sánh đầu tiên, hoặc cũng có khi phải qua tới tổng số $n - 1$ lượt so sánh. Vì ở lượt so sánh thứ i ta cần thực hiện $n - i$ lần so sánh, nên trong trường hợp xấu nhất ta sẽ phải cần tới $n - 1 + n - 2 + \dots + 1 = n(n - 1)/2$ lần so sánh. Nếu ta ký hiệu 1 thay thế cho giá trị nhỏ nhất, 2 thay cho giá trị thứ hai, và cứ tiếp tục như thế, thì trường hợp xấu nhất sẽ xảy ra nếu dãy ban đầu được cho là

$$n, n - 1, n - 2, \dots, 3, 2, 1$$

Tuy nhiên, không có lý do gì đặc biệt để cho rằng dãy được cho ban đầu sẽ có các phần tử xếp theo chiều giá trị giảm dần, nên điều này sẽ khiến ta nghĩ đến số lần so sánh cần thiết trung bình khi mà dãy ban đầu được cho một cách ngẫu nhiên. Vì thế, sẽ là hợp lý nếu ta coi dãy được cho có một thứ tự bất kỳ trong $n!$ thứ tự ngẫu nhiên có thể có. Nếu ký hiệu X là số lần so sánh cần đến khi thuật toán sủi bọt kết thúc, thì ta quan tâm $\mathbb{E}[X]$ là giá trị kỳ vọng của X . Trong khi rất khó để tính rõ ràng $\mathbb{E}[X]$, thì ta lại có thể đạt được các cận cho nó. Đầu tiên, vì với mỗi thứ tự ngẫu nhiên ban đầu,

$$X \leq n(n - 1)/2$$

nên

$$\mathbb{E}[X] \leq \frac{n(n-1)}{2} \quad (1.2)$$

Để đạt được một cận khác nữa, ta cần đến khái niệm **số nghịch thế của một hoán vị**. Cho một hoán vị bất kỳ $i_1, \dots, i_n$ của $1, 2, \dots, n$, ta nói rằng cặp có thứ tự (i, j) là một **ngịch thế** (inversion) của hoán vị này nếu $i < j$ và j đứng trước i trong hoán vị. Chẳng hạn, hoán vị

$$2, 4, 1, 5, 6, 3$$

có năm nghịch thế, cụ thể: $(1, 2), (1, 4), (3, 4), (3, 5), (3, 6)$. Cuối cùng, vì các giá trị trong các cặp nghịch thế là cần được đổi chỗ (nghĩa là được so sánh) nên số lần so sánh cần thiết nếu dùng thuật toán sủi bọt sẽ ít nhất phải bằng số nghịch thế của dãy cho ban đầu. Nghĩa là, nếu gọi I là số nghịch thế thì

$$X \geq I$$

Từ đó suy ra

$$\mathbb{E}[X] \geq \mathbb{E}[I] \quad (1.3)$$

Nhưng nếu, $i < j$ thì ta đặt

$$I(i, j) = \begin{cases} 1 & \text{nếu } (i, j) \text{ là một nghịch thế của dãy cho ban đầu} \\ 0 & \text{nếu ngược lại} \end{cases}$$

thì ta có

$$I = \sum_j \sum_{i < j} I(i, j)$$

Do vậy, theo tính chất của kỳ vọng ta có

$$\mathbb{E}[I] = \sum_j \sum_{i < j} \mathbb{E}[I(i, j)] \quad (1.4)$$

Bây giờ, nếu $i < j$, ta có

$$\mathbb{E}[I(i, j)] = \mathbb{P}\{j \text{ đứng trước } i \text{ trong dãy cho ban đầu}\}$$

Nhưng nếu thứ tự của dãy cho ban đầu là một thứ tự bất kỳ trong $n!$ thứ tự ngẫu nhiên có thể có, thì có thể thấy rằng khả năng i đứng trước j cũng ngang bằng khả năng j đứng trước i , nên

$$\mathbb{E}[I(i, j)] = \frac{1}{2}$$

Do vậy, vì có C_n^2 cặp i, j mà $i < j$ nên từ ?? suy ra

$$\mathbb{E}[I] = \frac{C_n^2}{2} = \frac{n(n-1)}{4}$$

Từ phương trình (1.2), (1.3) và phương trình trên ta được

$$\frac{n(n-1)}{4} \leq \mathbb{E}[X] \leq \frac{n(n-1)}{2}$$

Như vậy, nếu với n đủ lớn, số lần so sánh trung bình cần đến khi sắp xếp n giá trị bởi thuật toán sủi bọt sẽ dao động từ $n^2/4$ đến $n^2/2$.

□

Để đưa ra được công thức phương sai cho một tổng, ta cần đến khái niệm **hiệp phương sai**.

Hiệp phương sai (covariance) của hai biến ngẫu nhiên X và Y được định nghĩa như sau

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

Sau đây là các tính chất cơ bản của hiệp phương sai.

Mệnh đề 1.5.1. Cho các biến ngẫu nhiên X, Y, Z bất kỳ và hằng số c :

1. $Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
2. $Cov(X, X) = Var(X)$
3. $Cov(X, Y) = Cov(Y, X)$
4. $Cov(cX, Y) = cCov(X, Y)$
5. $Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$

Tính chất 5 tổng quát lên là

$$Cov\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right) = \sum_{i=1}^n \sum_{j=1}^n Cov(X_i, Y_j)$$

Công thức này có thể được dùng để suy ra phương sai cho tổng các biến ngẫu nhiên

$$\begin{aligned} Var\left(\sum_{i=1}^n X_i\right) &= Cov\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n Cov(X_i, X_j) \\ &= \sum_{i=1}^n Cov(X_i, X_i) + \sum_{i=1}^n \sum_{j \neq i}^n Cov(X_i, X_j) \\ &= \sum_{i=1}^n Var(X_i) + 2 \sum_{i=1}^n \sum_{j < i}^n Cov(X_i, X_j) \end{aligned} \quad (1.5)$$

Vì $Cov(X, Y) = 0$ khi X và Y độc lập, nên từ phương trình trên suy ra nếu các biến ngẫu nhiên độc lập thì

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i)$$

Ví dụ 1.1 e. Số thành công trong một dãy n phép thử độc lập, mỗi phép thử có kết quả thành công với xác suất p hoặc thất bại với xác suất $1 - p$, được gọi là một biến ngẫu nhiên nhị thức với các tham số n và p . Hãy tính giá trị kỳ vọng và phương sai của biến ngẫu nhiên này.

Lời giải. Nếu X là một biến ngẫu nhiên nhị thức với các tham số n và p , thì

$$\mathbb{P}\{X = i\} = C_n^i p^i (1 - p)^{n-i}, i = 0, \dots, n \quad (1.6)$$

Trong khi ta có thể tính giá trị kỳ vọng và phương sai của X trực tiếp từ (1.6), thì sẽ dễ dàng hơn nếu ta dùng biểu diễn

$$X = \sum_{i=1}^n X_i$$

với

$$X_i = \begin{cases} 1 & \text{nếu phép thử } i \text{ là thành công} \\ 0 & \text{nếu phép thử } i \text{ là thất bại} \end{cases}$$

Vì mỗi X_i là một biến ngẫu nhiên Bernoulli với

$$\mathbb{E}[X_i] = p, \quad Var(X_i) = p(1 - p)$$

nên

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i]$$

$$Var(X) = \sum_{i=1}^n Var(X_i) = np(1 - p)$$

với giả thiết các X_i độc lập để thỏa mãn phương sai của một tổng bằng tổng các phương sai.

□

Ví dụ 1.1 f. Bài toán sưu tập vé. Giả sử có m loại vé khác nhau, và tại mỗi thời điểm người ta chỉ mua một vé là một loại bất kỳ trong các loại này. Nếu gọi X là số vé mà một người cần mua theo thứ tự để có mỗi loại ít nhất một vé, tìm giá trị kỳ vọng và phương sai của X .

Lời giải. Để xác định $\mathbb{E}[X]$, ta gọi X_i là số vé cần thiết thêm vào, sau khi đã có i loại vé

phân biệt, cho đến khi có thêm một loại vé mới, $i = 0, \dots, m-1$. Khi đó,

$$X = \sum_{i=1}^{m-1} X_i$$

Bây giờ, khi đã có i loại vé phân biệt, mỗi vé mới sẽ là một loại mới với xác suất $(m-i)/m$. Như vậy, X_i là một biến ngẫu nhiên hình học với tham số $(m-i)/m$. Do đó,

$$\mathbb{E}[X] = \sum_{i=1}^{m-1} \mathbb{E}[X_i] = \sum_{i=1}^{m-1} \frac{m}{m-i} = m \sum_{i=1}^{m-1} \frac{1}{i}$$

Hơn nữa, vì dễ thấy rằng các biến ngẫu nhiên $X_i, i = 0, \dots, m-1$ độc lập, nên ta có

$$\text{Var}(X) = \sum_{i=1}^{m-1} \text{Var}(X_i) = m \sum_{i=1}^{m-1} \frac{i}{(m-i)^2}$$

với điều kiện trên có được là do phương sai của một biến ngẫu nhiên hình học có tham số p là $(1-p)/p^2$.

□

Chú ý. Ta có thể chỉ ra rằng

$$\lim_{m \rightarrow \infty} \left(\sum_{i=1}^m 1/i - \log(m) \right) = \gamma$$

với $\gamma \approx 0.57721$ được gọi là hằng số Euler, từ đó ta được

$$\mathbb{E}[X] \sim m \log(m)$$

vì

$$\mathbb{E}[X] = \sum_{i=1}^{m-1} \mathbb{E}[X_i] = \sum_{i=1}^{m-1} \frac{m}{m-i} = m \sum_{i=1}^{m-1} \frac{1}{i}$$

Ở đây, ta nói rằng $a_m \sim b_m$ khi $\lim_{m \rightarrow \infty} \frac{a_m}{b_m} = 1$. Hơn nữa, ta cũng có thể chỉ ra rằng

$$\lim_{m \rightarrow \infty} \sum_{i=1}^m \frac{1}{i^2} = \frac{\pi^2}{6}$$

Ta có

$$\sum_{i=1}^{m-1} \frac{i}{(m-i)^2} = \sum_{i=1}^{m-1} \frac{m}{(m-i)^2} - \sum_{i=1}^{m-1} \frac{m-i}{(m-i)^2}$$

nên

$$\text{Var}(X) \sim m^2 \pi^2 / 6$$

Ví dụ 1.1 g. Một bài toán sưu tập vé khác. Với giả thiết như trong ví dụ 1.1e, hãy tìm kỳ vọng và phương sai của số loại phân biệt trong một tập gồm n vé.

Lời giải. Ký hiệu X là số loại phân biệt, và

$$X = \sum_{i=1}^m X_i$$

với

$$X_i = \begin{cases} 1 & \text{nếu có vé loại } i \text{ trong bộ sưu tập} \\ 0 & \text{nếu ngược lại} \end{cases}$$

Vì X_i là một biến ngẫu nhiên Bernoulli, nên ta có

$$\begin{aligned} \mathbb{E}[X_i] &= 1 - \left(\frac{m-1}{m}\right)^n \\ \text{Var}(X_i) &= \left(\frac{m-1}{m}\right)^n \left(1 - \left(\frac{m-1}{m}\right)^n\right) \end{aligned} \quad (1.7)$$

Suy ra

$$\mathbb{E}[X] = \sum_{i=1}^m \mathbb{E}[X_i] = m \left[1 - \left(\frac{m-1}{m}\right)^n\right]$$

Hơn nữa, vì $X_i X_j$ cũng là Bernoulli, với $i \neq j$, nên

$$\mathbb{E}[X_i X_j] = \mathbb{P}\{X_i X_j = 1\} = \mathbb{P}(A_i A_j)$$

với A_k là biến cố sưu tập có chứa ít nhất một vé loại k . Vì

$$\begin{aligned} \mathbb{P}(A_i A_j) &= 1 - \mathbb{P}((A_i A_j)^C) \\ &= 1 - \mathbb{P}(A_i^C \cup A_j^C) \\ &= 1 - \mathbb{P}(A_i^C) - \mathbb{P}(A_j^C) + \mathbb{P}(A_i^C A_j^C) \\ &= 1 - 2 \left(\frac{m-1}{m}\right)^n + \left(\frac{m-2}{m}\right)^n \end{aligned}$$

Do đó, với $i \neq j$,

$$\begin{aligned} \text{Cov}(X_i, X_j) &= 1 - 2 \left(\frac{m-1}{m}\right)^n + \left(\frac{m-2}{m}\right)^n - \left[1 - \left(\frac{m-1}{m}\right)^n\right]^2 \\ &= \left(\frac{m-2}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n} \end{aligned} \quad (1.8)$$

Như vậy, từ các phương trình (1.5), (1.7) và (1.8), ta có

$$\begin{aligned} \text{Var}(X) &= m \left(\frac{m-1}{m}\right)^n \left(1 - \left(\frac{m-1}{m}\right)^n\right) + m(m-1) \left[\left(\frac{m-2}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n}\right] \\ &= m \left(\frac{m-1}{m}\right)^n + m(m-1) \left(\frac{m-2}{m}\right)^n - m^2 \left(\frac{m-1}{m}\right)^{2n} \quad \square \end{aligned}$$

Ví dụ 1.1 h Bước nhảy (Runs). Giả sử $X_1, \dots, X_n$ là một dãy các biến ngẫu nhiên nhị thức độc lập, và mỗi biến X_i sẽ nhận với xác suất p . Một dãy con liên tiếp cực đại gồm toàn giá trị 1 được gọi là một bước nhảy. Chẳng hạn, dãy

$$1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0$$

là có bước nhảy 3. Gọi R là số bước nhảy, hãy tìm $\mathbb{E}[R], \text{Var}(R)$.

Lời giải. Đặt I_i bằng 1 nếu có một bước nhảy bắt đầu tại X_i , và bằng 0 nếu ngược lại. Với $i = 1, \dots, n$, khi đó

$$R = \sum_{i=1}^n I_i \quad (1.9)$$

Vì

$$\mathbb{E}[I_1] = \mathbb{P}\{X_1 = 1\} = p$$

$$\mathbb{E}[I_i] = \mathbb{P}\{X_{i-1} = 0, X_i = 1\} = p(1 - p) \text{ với } i \geq 1$$

nên

$$\mathbb{E}[R] = \sum_{i=1}^n \mathbb{E}[I_i] = p + (n - 1)p(1 - p)$$

Giờ ta cần tính $\text{Var}(R)$. Giả sử $n \geq 2$, từ phương trình (1.9) ta có

$$\text{Var}(R) = \sum_{i=1}^n \text{Var}(I_i) + 2 \sum_{j=2}^n \sum_{i < j} \text{Cov}(I_i, I_j)$$

Vì I_i là một biến ngẫu nhiên Bernoulli nên

$$\text{Var}(I_i) = \mathbb{E}[I_i](1 - \mathbb{E}[I_i])$$

Ngoài ra, có thể thấy rằng I_i và I_j sẽ độc lập khi $i < j - 1$, từ đó

$$\text{Cov}(I_i, I_j) = 0, \quad i < j - 1$$

Bên cạnh đó, vì $I_{j-1}I_j = 0$, nên

$$\text{Cov}(I_{j-1}, I_j) = -\mathbb{E}[I_{j-1}\mathbb{E}[I_j]] = -\mathbb{E}[I_{j-1}p(1 - p)]$$

Như vậy, khi $n \geq 2$, ta được

$$\begin{aligned} \text{Var}(R) &= \text{Var}(I_1) + \sum_{j=2}^n \text{Var}(I_j) + 2\text{Cov}(I_1, I_2) + 2 \sum_{j=3}^n \text{Cov}(I_{j-1}, I_j) \\ &= p(1 - p) + (n - 1)p(1 - p)[1 - p(1 - p)] - 2p^2(1 - p) - 2(n - 2)p^2(1 - p)^2 \end{aligned}$$

Nếu ta đặt

$$c = p(1-p)[1-p(1-p)] - 2p^2(1-p)^2$$

thì ta có thể viết lại như sau

$$\text{Var}(R) = (n-2)c + p(1-p)^2(2-p)$$

Một cách khác để tính $\mathbb{E}[R]$ và $\text{Var}(R)$ là dùng phương pháp đệ quy. Nếu kí hiệu R_j là số bước nhảy trong dãy $X_1, \dots, X_j$ thì rõ ràng

$$R_j = R_{j-1} + I_j \quad (1.10)$$

Như vậy, với $j \geq 2$,

$$\mathbb{E}[R_j] = \mathbb{E}[R_{j-1}] + p(1-p)$$

Bắt đầu từ $\mathbb{E}[R_1] = p$, ta suy ra được

$$\mathbb{E}[R_n] = p + (n-1)(1-p)$$

Tương tự, với $j \geq 2$, từ (1.10) suy ra

$$\text{Var}(R_j) = \text{Var}(R_{j-1}) + \text{Var}(I_j) + 2\text{Cov}(R_{j-1}, I_j)$$

Mà

$$\begin{aligned} \text{Cov}(R_{j-1}, I_j) &= \text{Cov}(R_{j-2} + I_{j-1}, I_j) \\ &= \text{Cov}(R_{j-2}, I_j) + \text{Cov}(I_{j-1}, I_j) \\ &= \text{Cov}(I_{j-1}, I_j) \\ &= -\mathbb{E}[I_{j-1}]\mathbb{E}[I_j] \end{aligned}$$

Do vậy, với $j \geq 2$,

$$\text{Var}(R_j) = \text{Var}(R_{j-1}) + p(1-p)[1-p(1-p)] - 2p(1-p)\mathbb{E}[I_{j-1}]$$

Vì $\text{Var}(R_1) = p(1-p)$, đệ quy ta được

$$\begin{aligned} \text{Var}(R_2) &= p(1-p) + p(1-p)[1-p(1-p)] - 2p^2(1-p) \\ &= p(1-p)^2(2-p) \end{aligned}$$

Với $j \geq 3$, đệ quy ta được

$$\text{Var}(R_j) = \text{Var}(R_{j-1}) + c$$

Vậy nên

$$\text{Var}(R_n) = \text{Var}(R_2) + (n - 2)c$$

Bây giờ ta nói đến một ứng dụng của bài toán trên, đó là cho phần chỉ dẫn của một quyển sách. Phần chỉ dẫn là phần gồm một danh sách các từ nào đó có trong quyển sách và ứng với mỗi từ là phần liệt kê tất cả các trang sách có xuất hiện từ đó. Giả sử mục chỉ dẫn sẽ liệt kê các trang theo phân đoạn; nghĩa là, nếu một từ nào đó xuất hiện ở các trang 2, 3, 4, 7, 8, 9, 10 và 12, thì trong bảng chỉ dẫn nó sẽ được viết thành ba phân đoạn 2 – 4, 7 – 10, 12. Giờ giả sử ta quan tâm đến số phân đoạn trong phần chỉ dẫn của một từ nào đó, với giả thiết từ này sẽ xuất hiện độc lập trên mỗi trang, trong tổng số n trang của quyển sách, với xác suất p . Nếu kí hiệu X_i bằng 1 nếu từ này xuất hiện ở trang i , thì có thể thấy ngay rằng số phân đoạn của các trang cho từ này chính là bằng số bước nhảy của dãy $X_1, \dots, X_n$, do đó trung bình và phương sai của nó sẽ được tính được nhờ các kết quả trên.

1.6 Hàm moment tổng quát và biến đổi Laplace

Hàm moment tổng quát của biến ngẫu nhiên X được định nghĩa như sau

$$\phi(t) = \mathbb{E}[e^{tX}]$$

Nó được gọi là hàm moment tổng quát của X vì moment các bậc khác nhau của X có thể suy từ việc đạo hàm liên tiếp $\phi(t)$, và xét các đạo hàm này tại $t = 0$. Chẳng hạn

$$\phi'(t) = \mathbb{E}[Xe^{tX}]$$

$$\phi''(t) = \mathbb{E}[X^2 e^{tX}]$$

và tổng quát

$$\phi^{(n)}(t) = \mathbb{E}[X^n e^{tX}]$$

Xét tại $t = 0$ thì

$$\phi^{(n)}(0) = \mathbb{E}[X^n]$$

Ví dụ 1.1 i. Phân phối hình học. Biến ngẫu nhiên X được gọi là có phân phối hình học với tham số p nếu

$$\mathbb{P}\{X = n\} = pq^{n-1}, n = 1, 2, \dots$$

với $q = 1 - p$. Như vậy, hàm moment tổng quát của X là

$$\begin{aligned}\phi(t) &= \mathbb{E}[e^{tX}] \\ &= \sum_{n=1}^{\infty} e^{tn} p q^{n-1} \\ &= p e^t \sum_{n=1}^{\infty} (q e^t)^{n-1} \\ &= \frac{p e^t}{1 - q e^t}\end{aligned}$$

Đạo hàm ta được

$$\begin{aligned}\phi'(t) &= \frac{(1 - q e^t) p e^t + p e^t q e^t}{(1 - q e^t)^2} = \frac{p e^t}{(1 - q e^t)^2} \\ \phi''(t) &= \frac{(1 - q e^t)^2 p e^t + 2 p e^t (1 - q e^t) q e^t}{(1 - q e^t)^4}\end{aligned}$$

Xét tại $t = 0$ ta được

$$\begin{aligned}\mathbb{E}[X] &= \frac{p}{(1 - q)^2} = \frac{1}{p} \\ \mathbb{E}[X^2] &= \frac{(1 - q)^2 p + 2 p (1 - q) q}{(1 - q)^4} = \frac{p^2(p + 2q)}{(1 - q)^4} = \frac{1 + q}{p^2}\end{aligned}$$

Do vậy,

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X] = \frac{1 - p}{p^2}$$

□

Nếu X và Y độc lập thì

$$\mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} e^{tY}] = \mathbb{E}[e^{tX}] \mathbb{E}[e^{tY}]$$

Nói cách khác, hàm moment tổng quát của một tổng các biến ngẫu nhiên độc lập bằng tích các hàm moment tổng quát của từng biến ngẫu nhiên thành phần.

Một tính chất quan trọng khác của hàm moment tổng quát đó là có một tương ứng một-một giữa hàm moment tổng quát với hàm phân phối của một biến ngẫu nhiên. Như vậy, một hàm moment tổng quát xác định duy nhất một hàm phân phối.

Ví dụ 1.1 j. Một biến ngẫu nhiên X được gọi là có **phân phối chuẩn** với trung bình μ và phương sai σ^2 nếu hàm mật độ xác suất của nó là

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, -\infty < x < +\infty$$

Khi $\mu = 0, \sigma = 1$, ta nói X có **phân phối chuẩn tiêu chuẩn**.

Hàm moment tổng quát của biến ngẫu nhiên có phân phối chuẩn tiêu chuẩn Z là

$$\begin{aligned}\mathbb{E}[e^t Z] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x^2-2tx)/2} dx \\ &= e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x-t)^2/2} dx \\ &= e^{t^2/2}\end{aligned}$$

Bây giờ, ta thấy nếu Z là một biến ngẫu nhiên có phân phối chuẩn tiêu chuẩn, thì $X = \sigma Z + \mu$ sẽ là biến ngẫu nhiên có phân phối chuẩn với trung bình μ và phương sai σ^2 ; như vậy,

$$\mathbb{E}[e^{tx}] = \mathbb{E}[e^{t(\sigma Z + \mu)}] = e^{t\mu} \mathbb{E}[e^{t\sigma Z}] = e^{\mu t + \sigma^2 t^2/2}$$

Giả sử X và Y là các biến ngẫu nhiên có phân phối chuẩn độc lập với trung bình μ_x và μ_y , và phương sai σ_x^2 và σ_y^2 ; khi đó

$$\mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}] \mathbb{E}[e^{tY}] = \exp\{(\mu_x + \mu_y)t + (\sigma_x^2 + \sigma_y^2)t^2/2\}$$

Do tính duy nhất của hàm moment tổng quát, nên từ trên ta thấy rằng tổng của các biến ngẫu nhiên có phân phối chuẩn độc lập sẽ là một biến ngẫu nhiên có phân phối chuẩn. □

Cho biến ngẫu nhiên không âm X , ta định nghĩa **biến đổi Laplace** $g(t), t > 0$, của X như sau

$$g(t) = \phi(-t) = \mathbb{E}[e^{-tX}]$$

Ta thấy rằng, biến đổi Laplace xét tại t chính là moment tổng quát xét tại $-t$. Biến đổi Laplace có điểm thuận lợi hơn hàm moment tổng quát là khi biến ngẫu nhiên là không âm, tức là $X \geq 0$ và $t \geq 0$, thì

$$0 \leq e^{-tX} \leq 1$$

Nghĩa là, biến đổi Laplace luôn có giá trị nằm giữa 0 và 1, từ đó cho thấy các biến ngẫu nhiên không âm mà có cùng biến đổi Laplace thì cũng phải có cùng phân phối.

1.7 Kỳ vọng có điều kiện

Nếu X và Y là hai biến ngẫu nhiên rời rạc, thì **hàm khối lượng xác suất có điều kiện** của X với điều kiện $Y = y$ được định nghĩa bởi

$$\mathbb{P}\{X = x \mid Y = y\} = \frac{\mathbb{P}\{X = x, Y = y\}}{\mathbb{P}\{Y = y\}}$$

và kỳ vọng có điều kiện của X với điều kiện $Y = y$ được định nghĩa bởi

$$\mathbb{E}[X \mid Y = y] = \sum_x \mathbb{P}\{X = x \mid Y = y\}$$

Tương tự, nếu X và Y có hàm mật độ đồng thời $f(x, y)$, thì **hàm mật độ xác suất có điều kiện** của X với điều kiện $Y = y$ được định nghĩa bởi

$$f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$

và kỳ vọng có điều kiện của X với điều kiện $Y = y$ được định nghĩa bởi

$$\mathbb{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y) dx$$

Chú ý rằng, trong tất cả các công thức trên vẫn đúng trong trường hợp không có điều kiện ngoại trừ rằng bây giờ các xác suất đang được tính toán với điều kiện là biến cố $Y = y$.

Giả sử $\mathbb{E}[X \mid Y]$ là một hàm theo biến Y (kí hiệu là $\mathbb{E}[X \mid Y = y]$), một kết quả vô cùng quan trọng là

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] \quad (1.11)$$

Nếu Y là một biến ngẫu nhiên rời rạc, thì (1.11) chính là

$$\mathbb{E}[X] = \sum_x \mathbb{E}[X \mid Y = y] \mathbb{P}\{Y = y\}$$

Khi Y liên tục thì (1.11) chính là

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \mathbb{E}[X \mid Y = y] f_Y(y) dy$$

Trong trường hợp X và Y có hàm mật độ đồng thời $f(x, y)$,

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbb{E}[X \mid Y = y] f_Y(y) dy &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy \\ &= \mathbb{E}[X] \end{aligned}$$

Trường hợp rời rạc chứng minh tương tự. Như vậy, ta thấy $\mathbb{E}[X]$ chính là trọng lượng trung bình của các kỳ vọng của các kỳ vọng có điều kiện của X với điều kiện $Y = y$, trong đó mỗi số hạng $\mathbb{E}[X | Y = y]$ có trọng lượng bằng xác suất (hoặc mật độ xác suất trong trường hợp liên tục) để $Y = y$. Kết quả này thường được dùng để tính giá trị kỳ vọng của X một cách dễ dàng hơn bằng việc đặt điều kiện lên giá trị của biến ngẫu nhiên thứ hai Y . Các ví dụ sau sẽ minh họa cho điều này.

Ví dụ 1.1 k. Tổng một số ngẫu nhiên các biến ngẫu nhiên. Giả sử $X_1, X_2, \dots$ là một dãy các biến ngẫu nhiên độc lập có cùng phân phối, có hàm moment tổng quát $\phi_X(t) = \mathbb{E}[\exp\{tX_i\}]$, và giả sử N là một biến ngẫu nhiên nhận giá trị nguyên không âm, độc lập với các X_i . Tìm trung bình và phương sai của

$$S = \sum_{i=1}^N X_i$$

Lời giải. Để giải bài toán trên, đầu tiên ta xác định hàm moment tổng quát của S . Vì có thể tính được hàm moment tổng quát của tổng một số cố định các biến ngẫu nhiên độc lập, nên ta sẽ bắt đầu bằng việc đặt điều kiện lên giá trị của N ,

$$\begin{aligned} \mathbb{E}[\exp\{tS\} | N = n] &= \mathbb{E}\left[\exp\left\{t \sum_{i=1}^N X_i\right\} | N = n\right] \\ &= \mathbb{E}\left[\exp\left\{t \sum_{i=1}^n X_i\right\} | N = n\right] \\ &= \mathbb{E}\left[\exp\left\{t \sum_{i=1}^n X_i\right\}\right] \quad \text{do sự độc lập giữa } N \text{ và } X_i \\ &= (\phi_X(t))^n \end{aligned}$$

Do đó,

$$\mathbb{E}[\exp\{tS\} | N] = (\phi_X(t))^N$$

Suy ra

$$\phi_S(t) = \mathbb{E}[\exp\{tS\}] = \mathbb{E}[(\phi_X(t))^N]$$

Để xác định giá trị kỳ vọng và phương sai của S , đạo hàm $\phi_S(t)$ ta được

$$\begin{aligned} \phi'_S(t) &= \mathbb{E}[N(\phi_X(t))^{N-1} \phi'_X(t)] \\ \phi''_S(t) &= \mathbb{E}[N(\phi_X(t))^{N-1} \phi''_X(t) + N(N-1)(\phi_X(t))^{N-2} (\phi'_X(t))^2] \end{aligned}$$

Xét tại $t = 0$ cho ta

$$\mathbb{E}[S] = \mathbb{E}[N\mathbb{E}[X]] = \mathbb{E}[N]\mathbb{E}[X]$$

và

$$\begin{aligned}\mathbb{E}[S^2] &= \mathbb{E}[N\mathbb{E}[X^2]] + N(N-1)\mathbb{E}^2[X] \\ &= \mathbb{E}[N]Var(X) + \mathbb{E}[N^2]\mathbb{E}^2[X]\end{aligned}$$

Vậy

$$\begin{aligned}Var(S) &= \mathbb{E}[S^2] - \mathbb{E}^2[S] \\ &= \mathbb{E}[N]Var(X) + \mathbb{E}^2[X]Var(N)\end{aligned}$$

□

Ví dụ 1.1 1. Xét một danh sách gồm m phần tử, với m là một số lớn, và giả sử ta quan tâm đến số phần tử phân biệt trong danh sách. Nghĩa là, nếu ký hiệu $n(i)$ là số lần mà phần tử ở vị trí thứ i xuất hiện trong danh sách, thì ta cần tính

$$d = \sum_{i=1}^m \frac{1}{n(i)}$$

Gọi X là biến ngẫu nhiên mà nó nhận một trong các giá trị $1, 2, \dots, m$, và ta sẽ tìm d bằng cách xác định số lần phần tử nằm ở vị trí X xuất hiện trong danh sách. Vì

$$\begin{aligned}\mathbb{E}\left[\frac{1}{n(X_i)}\right] &= \sum_{i=1}^m \frac{1}{n(i)} \mathbb{P}\{X = i\} \\ &= \frac{1}{m} \sum_{i=1}^m \frac{1}{n(i)}\end{aligned}$$

nên

$$\mathbb{E}\left[\frac{m}{n(X)}\right] = d$$

Như vậy, bằng việc đưa vào các biến ngẫu nhiên $X_1, \dots, X_k$ độc lập, mỗi biến nhận một trong các giá trị $1, 2, \dots, m$, ta có thể dùng

$$\frac{m}{k} \sum_{i=1}^k \frac{1}{n(X_i)}$$

để tìm d .

Một khó khăn của phương pháp trên là việc xác định giá trị $n(X)$, nó phụ thuộc nhiều vào kích thước của danh sách. Có một cách hiệu quả hơn như sau: gọi X là biến ngẫu nhiên mà nó nhận một trong các giá trị $1, \dots, m$, và ta xét phần tử nằm ở vị trí X trong danh sách. Lướt qua danh sách một lượt, bắt đầu từ phần tử đầu tiên cho đến khi bắt gặp phần tử giống với phần tử đang xét, nếu lần đầu tiên ta bắt gặp phần tử này là ở vị

trí X thì đặt $I = 1$; ngược lại đặt $I = 0$. (Chẳng hạn, giả sử $X = 1200$ và phần tử ở vị trí 1200 là A , nếu lần xuất hiện đầu tiên của A trong danh sách là tại vị trí 44 thì $I = 0$). Giả sử cho biết $n(X)$, thì vì phần tử ta đang xét là phần tử nằm ở vị trí X , sẽ xuất hiện tại mỗi vị trí trong số $n(X)$ vị trí mà nó xuất hiện, nên

$$\mathbb{E}[I \mid n(X)] = \frac{1}{n(X)}$$

Lấy kỳ vọng hai vế ta được

$$\mathbb{E}[I] = \mathbb{E}\left[\frac{1}{n(X)}\right] = \frac{d}{m}$$

Như vậy, có thể tìm được giá trị d bằng cách lặp lại m lần quá trình trên để được m lần giá trị trung bình của I .

□

Ví dụ 1.1 m. Bài toán đóng gói (A Bin Packing Problem). Giả sử ta có các vật nào đó mà trọng lượng của chúng độc lập với nhau và có phân phối đều trên khoảng $(0, 1)$, cần được xếp vào một dãy các thùng mà mỗi thùng chỉ có thể chứa tối đa một đơn vị trọng lượng. Các vật được xếp một cách tuần tự, đầu tiên từ thùng 1, cho đến khi có một vật nào đó mà có trọng lượng vượt quá trọng lượng cho phép có thể thêm vào thùng, nghĩa là nếu thêm vật này vào thì thùng sẽ có trọng lượng vượt quá 1. Tại thời điểm này, thùng 1 sẽ được đóng gói lại. Các vật tiếp khác được xếp tiếp theo vào thùng 2 và quá trình cứ tiếp tục như vậy cho đến hết. Chẳng hạn, nếu trọng lượng của sáu vật nào đó là 0.40, 0.36, 0.44, 0.22, 0.10, 0.62 thì các vật 1 và 2 sẽ được xếp vào thùng 1, các vật 3, 4, 5 sẽ được xếp vào thùng 2, vật 6 sẽ là vật đầu tiên được xếp vào thùng 3. Gọi B_i là số vật được xếp vào thùng i , ta quan tâm

$$\mathbb{E}\left[\sum_{i=1}^n B_i\right] = \sum_{i=1}^n \mathbb{E}[B_i]$$

là số vật trung bình được xếp vào n thùng đầu tiên.

Ta bắt đầu bằng việc xác định số vật trung bình được xếp vào thùng 1. Bây giờ nếu ta ký hiệu $U_1, U_2, \dots$ là một dãy các biến ngẫu nhiên độc lập và có cùng phân phối đều trên khoảng $(0, 1)$, biểu diễn lần lượt cho các trọng lượng và đặt

$$N = \min \left\{ n : \sum_{i=1}^n U_i > 1 \right\}$$

thì rõ ràng $N - 1$ là số vật có trong thùng 1. Để xác định $\mathbb{E}[N]$, ta xét bài toán tổng quát hơn là tìm

$$m(x) = \mathbb{E}[N(x)] \quad \text{với } 0 \leq x \leq 1$$

với

$$N(x) = \min \left\{ n : \sum_{i=1}^n U_i > x \right\}$$

Nghĩa là, $N(x)$ chính là số biến ngẫu nhiên, có phân phối đều trên khoảng $(0, 1)$, tối thiểu mà ta cần thêm vào cho đến khi tổng của chúng vượt quá x , $0 \leq x \leq 1$. Đặt điều kiện lên U_1 cho ta

$$\mathbb{E}[N(x)] = \int_0^1 \mathbb{E}[N(x) \mid U_1 = y] dy \quad (1.12)$$

Tuy nhiên,

$$\mathbb{E}[N(x) \mid U_1 = y] = \begin{cases} 1 & \text{nếu } y > x \\ 1 + m(x - y) & \text{nếu } y \leq x \end{cases} \quad (1.13)$$

Điều trên có thể thấy trực tiếp khi $y > x$; khi $y \leq x$, vì nếu biến ngẫu nhiên đầu tiên nhận giá trị y , thì tại thời điểm này ta có được một biến ngẫu nhiên và vẫn phải thêm vào các biến ngẫu nhiên khác cho đến khi tổng của các biến ngẫu nhiên này vượt quá $x - y$. Thay (1.13) vào phương trình (1.12) ta được

$$\begin{aligned} m(x) &= 1 + \int_0^x m(x - y) dy \\ &= 1 + \int_0^x m(u) du \end{aligned}$$

Đạo hàm ta được

$$m'(x) = m(x)$$

hay

$$\frac{m'(x)}{m(x)} = 1$$

Lấy tích phân ta được

$$\log(m(x)) = x + c$$

hay

$$m(x) = ke^x$$

Vì $m(0) = 1$, suy ra $k = 1$; nghĩa là,

$$m(x) = e^x$$

Như vậy,

$$\mathbb{E}[B_1] = \mathbb{E}[N(1) - 1] = m(1) - 1 = e - 1$$

và vì thế số vật trung bình xếp trong thùng 1 là $e - 1$. (Một điều thú vị là từ trên ta có thể thấy số biến ngẫu nhiên, có phân phối đều trên khoảng $(0, 1)$, trung bình mà có tổng

vượt 1 chính bằng e !).

Bây giờ ta xét vật đầu tiên được xếp vào thùng i , với $i > 1$. Vì đây là vật có trọng lượng lớn hơn dung lượng còn lại của thùng trước, nên suy ra trọng lượng của nó là một biến ngẫu nhiên có phân phối đều trên $(0, 1)$ với điều kiện lớn hơn trọng lượng (ngẫu nhiên) nào đó. Tuy nhiên, vì trọng lượng của vật này lớn hơn giá trị của một biến ngẫu nhiên nào đó có phân phối đều trên $(0, 1)$, nên số vật trung bình xếp chung thùng với nó sẽ ít hơn so với thùng 1 (là thùng mà vật đầu tiên được xếp vào có trọng lượng ngẫu nhiên phân phối đều trên $(0, 1]$). Nghĩa là,

$$\mathbb{E}[B_i] \leq \mathbb{E}[B_1] = e - 1$$

Suy ra

$$\mathbb{E} \left[\sum_{i=1}^n B_i \leq n(e - 1) \right]$$

Để đạt được một cận tốt hơn, giả sử ngay khi ta gặp một vật mà nó nặng quá để xếp được vào thùng đang xếp, thì không những ta sẽ đặt vật này vào thùng mới mà ta còn đóng gói ngay thùng mới này lại. (Chẳng hạn, nếu năm vật đầu tiên có trọng lượng là 0.40, 0.36, 0.44, 0.22, 0.90, thì các vật một và hai sẽ được xếp vào thùng 1, vật ba sẽ được xếp vào thùng 2, vật bốn được xếp vào thùng 3, còn vật năm sẽ là vật đầu tiên được xếp vào thùng 4). Giả sử gọi B_i^* là số vật được xếp vào thùng i theo cách này, thì

$$\mathbb{E}[B_i^*] = \begin{cases} 1 & \text{nếu } n \text{ chẵn} \\ e - 1 & \text{nếu } n \text{ lẻ} \end{cases}$$

Rõ ràng là

$$\sum_{i=1}^n B_i \geq \sum_{i=1}^n B_i^*$$

nên ta đạt được một cận tốt hơn

$$\mathbb{E} \left[\sum_{i=1}^n B_i \right] \geq \begin{cases} \frac{ne}{2} & \text{nếu } n \text{ chẵn} \\ \frac{(n-1)e}{2} + e - 1 & \text{nếu } n \text{ lẻ} \end{cases}$$

Hơn nữa, vì $e - 1 \approx 1.718$, nên với n chẵn

$$1.359n \leq \mathbb{E} \left[\sum_{i=1}^n B_i \right] \leq 1.718n$$

□

1.8 Biến ngẫu nhiên phân phối mũ

Biến ngẫu nhiên liên tục không âm X được gọi là có phân phối mũ với tham số λ nếu hàm mật độ của nó có dạng

$$f(x) = \lambda e^{-\lambda x}, x \geq 0$$

Hàm phân phối của nó có dạng

$$F(x) = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}$$

Hàm moment tổng quát của nó là

$$\begin{aligned} \phi(t) &= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{\lambda - t}, \quad t < \lambda \end{aligned}$$

Đạo hàm ta được

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^2}, \quad \phi''(t) = \frac{2\lambda}{(\lambda - t)^3}$$

Suy ra

$$\mathbb{E}[X] = 1/\lambda, \quad \text{Var}(X) = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$$

Một tính chất quan trọng của biến ngẫu nhiên phân phối mũ đó là nó không có trí nhớ, với ta gọi một biến ngẫu nhiên không âm X là không có trí nhớ nếu, với mọi $s > t, t > 0$,

$$\mathbb{P}\{X > s + t \mid X > t\} = \mathbb{P}\{X > s\} \quad (1.14)$$

Nếu ta xem X như là thời gian sống của một vật nào đó, thì phương trình trên nói rằng xác suất để vật này sống sót ít nhất $s + t$ giờ, với điều kiện nó đã sống sót t giờ, là bằng xác suất để nó sống sót ít nhất s giờ. Nói cách khác, nếu vật này sống sót tới thời điểm t , thì phân phối của tổng thời gian sống sót còn lại của nó là giống với phân phối thời gian sống lúc đầu của nó; nghĩa là, vật này không nhớ được rằng nó đã sống sót được đến thời điểm t . Phương trình (1.14) tương đương

$$\mathbb{P}\{X > s + t\} = \mathbb{P}\{X > s\} \mathbb{P}\{X > t\}$$

Vì phương trình trên được thỏa mãn khi X có phân phối mũ (do $e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t}$) nên suy ra các biến ngẫu nhiên phân phối mũ là các biến ngẫu nhiên không có trí nhớ.

Ví dụ. Xét một bưu điện trong đó có hai bàn điện thoại. Giả sử cô A bước vào bưu điện để gọi điện thì cô thấy rằng ông B đang gọi điện ở một bàn và bàn còn lại thì ông C đang

gọi. Giả sử thêm rằng cô A sẽ sử dụng điện thoại ngay sau khi ông B hoặc C gọi xong. Nếu coi tổng thời gian khách hàng gọi điện trên một bàn điện thoại là có phân phối mũ với trung bình $1/\lambda$, thì xác suất để cô A là người cuối cùng, trong ba người trên, rời bưu điện là bằng bao nhiêu?

Lời giải. Câu trả lời đạt được bằng cách sau: Xét thời điểm A tìm thấy được bàn điện thoại rồi đầu tiên để gọi. Tại thời điểm này, hoặc B hoặc C phải có một người vừa rời đi, còn người kia vẫn đang gọi. Tuy nhiên, do tính không có trí nhớ của biến ngẫu nhiên phân phối mũ, nên tổng thời gian mà người còn lại đang gọi điện là có phân phối mũ với trung bình $1/\lambda$. Nghĩa là, giống như ông ấy chỉ vừa mới bắt đầu gọi điện tại thời điểm này. Như vậy, do tính đối xứng, xác suất để ông ấy vừa gọi điện xong trước cô A phải bằng $1/2$. $\square$

Tính không có trí nhớ còn được minh họa bởi **hàm tốc độ hỏng** của phân phối mũ. Xét một biến ngẫu nhiên dương liên tục X có hàm phân phối F và hàm mật độ f . Hàm tốc độ hỏng $r(t)$ được định nghĩa bởi

$$r(t) = \frac{f(t)}{F(t)} = \frac{f(t)}{1 - F(t)}$$

Để giải thích hàm tốc độ hỏng $r(t)$, ta giả sử một vật nào đó, có thời gian sống X , đã sống sót được t giờ, và bây giờ ta muốn biết xác suất để nó không thể sống sót thêm thời gian dt nữa là bao nhiêu. Nghĩa là, ta quan tâm xác suất $\mathbb{P}\{X \in (t, t + dt) \mid X > t\}$. Ta có

$$\begin{aligned} \mathbb{P}\{X \in (t, t + dt) \mid X > t\} &= \frac{\mathbb{P}\{X \in (t, t + dt), X > t\}}{\mathbb{P}\{X > t\}} \\ &= \frac{\mathbb{P}\{X \in (t, t + dt)\}}{\mathbb{P}\{X > t\}} \\ &\approx \frac{f(t)dt}{F(t)} \\ &= r(t)dt \end{aligned}$$

Bây giờ giả sử thời gian sống của vật nào đó là có phân phối mũ. Khi đó, do tính không có trí nhớ của biến ngẫu nhiên phân phối mũ, nên thời gian sống còn lại của một vật đã có t năm tuổi là có phân phối mũ giống với phân phối của thời gian sống của một vật mới. Như vậy, hàm tốc độ hỏng là một hằng số. Thật vậy,

$$r(t) = \frac{f(t)}{F(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

Vậy hàm tốc độ hỏng của phân phối mũ chính bằng tham số λ của nó, thường còn được gọi là **tốc độ của phân phối mũ**.

Vì $\frac{d}{ds}\bar{F}(s) = -f(s)ds$, và $\log(\bar{F}(0)) = 0$, nên tích phân hai vế phương trình

$$r(s) = \frac{f(s)}{\bar{F}(s)}$$

ta được

$$\int_0^t r(s)ds = -\log(\bar{F}(t))$$

hay

$$\bar{F}(t) = \exp \left\{ - \int_0^t r(s)ds \right\}$$

Do vậy, hàm tốc độ hỏng $r(t)$ xác định một cách duy nhất hàm phân phối F . Vì một biến ngẫu nhiên không có trí nhớ sẽ có tốc độ hỏng bằng hằng số, do định nghĩa của nó, nên từ phương trình trên ta thấy duy nhất các biến ngẫu nhiên phân phối mũ là không có trí nhớ.

Nếu $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập và có phân phối mũ với các tốc độ $\mu_1, \dots, \mu_n$, thì biến ngẫu nhiên nhỏ nhất trong chúng có phân phối mũ với tốc độ bằng tổng các tốc độ trên. Thật vậy,

$$\begin{aligned} \mathbb{P}\{\min X_i > x\} &= \mathbb{P}\{\min X_i > x, \text{ với mọi } i = 1, \dots, n\} \\ &= \prod_{i=1}^n \mathbb{P}\{X_i > x\} \quad (\text{do sự độc lập}) \\ &= \prod_{i=1}^n e^{-\mu_i x} \\ &= \exp \left\{ - \sum_{i=1}^n \mu_i x \right\} \end{aligned} \tag{1.15}$$

Một điều hữu ích về giá trị nhỏ nhất trong tập các biến ngẫu nhiên phân phối mũ đó là nó độc lập với việc sắp thứ tự các biến ngẫu nhiên trong tập này. Để thấy rõ điều này, ta giả sử $i_1, \dots, i_n$ là một hoán vị của $1, \dots, n$. Khi đó,

$$\begin{aligned} \mathbb{P}\{\min X_i > x \mid X_{i_1} < \dots < X_{i_n}\} &= \frac{\mathbb{P}\{\min X_i > x, X_{i_1} < \dots < X_{i_n}\}}{\mathbb{P}\{X_{i_1} < \dots < X_{i_n}\}} \\ &= \frac{\mathbb{P}\{\min X_i > x\} \mathbb{P}\{X_{i_1} < \dots < X_{i_n} \mid \min X_i > x\}}{\mathbb{P}\{X_{i_1} < \dots < X_{i_n}\}} \\ &= \mathbb{P}\{\min X_i > x\} \end{aligned}$$

với phương trình cuối cùng là dùng tính chất không có trí nhớ của biến ngẫu nhiên phân phối mũ để suy ra được rằng

$$\mathbb{P}\{X_{i_1} < \dots < X_{i_n} \mid \min X_i > x\} = \mathbb{P}\{X_{i_1} < \dots < X_{i_n}\}$$

Để tính xác suất X_j là giá trị nhỏ nhất, đặt điều kiện lên X_j ta được

$$\begin{aligned}\mathbb{P}\{X_j = \min_i X_i\} &= \int_0^\infty \mathbb{P}\{X_j = \min_i X_i \mid X_j = x\} \mu_j e^{-\mu_j x} dx \\ &= \int_0^\infty \mathbb{P}\{X_i > x, \text{ với mọi } i \neq j \mid X_j = x\} \mu_j e^{-\mu_j x} dx \\ &= \int_0^\infty \prod_{i \neq j} e^{-\mu_i x} \mu_j e^{-\mu_j x} dx \\ &= \frac{\mu_j}{\sum_{i=1}^n \mu_i}\end{aligned}$$

Nghĩa là, xác suất để X_j là nhỏ nhất trong n biến ngẫu nhiên phân phối mũ độc lập chính là bằng tốc độ của X_j chia cho tổng tất cả các tốc độ của n biến ngẫu nhiên phân phối mũ này.

Ví dụ. Phân tích thuật toán tham lam cho bài toán phân công. Có một nhóm gồm n người được phân công n công việc, với mỗi người được phân công một công việc. Nếu người i được phân công việc j thì ta mất một chi phí là $C_{i,j}$. Như vậy, ta có một tập gồm n^2 giá trị $C_{i,j}, i, j = 1, \dots, n$. Bài toán phân công cổ điển là làm sao để tổng chi phí là thấp nhất. Tuy nhiên, khá là khó để xác định được chính xác phân công tối ưu nhất. Ta sẽ xem xét hai thuật toán sau để tìm ra được phân công tối ưu. Thuật toán đầu tiên bắt đầu bằng việc phân công người 1 với công việc mà làm cho chi phí là thấp nhất, nghĩa là người 1 sẽ được phân công công việc j_1 , với $C_{1,j_1} = \min_j C_{1,j}$. Sau đó, loại công việc này ra, và người 2 lại được phân công công việc mà làm cho chi phí là thấp nhất. Cứ như thế cho đến khi tất cả n người đều đã được phân công. Theo cách này, ta luôn chọn được công việc tốt nhất cho từng người, vì thế ta sẽ gọi nó là **thuật toán tham lam A**.

Thuật toán thứ hai ta gọi là **thuật toán tham lam B**. Xét tất cả n^2 giá trị chi phí có thể có và chọn cặp i_1, j_1 sao cho C_{i_1,j_1} là giá trị nhỏ nhất. Khi đó, ta phân công cho người i_1 công việc j_1 . Sau đó, loại bỏ tất cả các giá trị chi phí mà liên quan đến cả người i_1 hoặc công việc j_1 , (nghĩa là còn lại $(n-1)^2$ giá trị chi phí) và cứ tiếp tục như thế cho đến hết. Nghĩa là, tại mỗi bước ta chọn ra một người và một công việc sao cho chi phí là thấp nhất trong số tất cả người và công việc còn lại chưa được phân công. Với giả thiết $C_{i,j}$ tạo nên một tập gồm n^2 biến ngẫu nhiên phân phối mũ độc lập, mỗi biến ngẫu nhiên có trung bình 1, thuật toán nào trong hai thuật toán trên cho tổng chi phí trung bình thấp hơn?

Lời giải. Đầu tiên ta xét thuật toán A. Gọi C_i là chi phí ứng với người $i, i = 1, \dots, n$. Rõ ràng, C_1 là giá trị nhỏ nhất trong n biến ngẫu nhiên phân phối mũ độc lập, mỗi biến ngẫu nhiên có tốc độ 1; bởi phương trình (1.15) nên C_1 sẽ là biến ngẫu nhiên phân phối mũ có tốc độ n . Tương tự, C_2 là giá trị nhỏ nhất trong $n-1$ biến ngẫu nhiên phân phối

mũ độc lập với tốc độ 1 nên nó sẽ là một biến ngẫu nhiên phân phối mũ có tốc độ $n - 1$. Tổng quát, C_i sẽ có phân phối mũ với tốc độ $n - i + 1, i = 1, \dots, n$. Như vậy, tổng chi phí trung bình tính theo thuật toán A là

$$\mathbb{E}_A[\text{tổng chi phí}] = \mathbb{E}[C_1 + \dots + C_n] = \sum_{i=1}^n \frac{1}{i}$$

Giờ ta xét thuật toán B . Gọi C_i là chi phí cho cặp người-việc thứ i được phân công bởi thuật toán này. Vì C_1 là nhỏ nhất trong tất cả n^2 giá trị $C_{i,j}$, nên C_1 có phân phối mũ với tốc độ n^2 . Bây giờ, do tính chất không có trí nhớ của biến ngẫu nhiên phân phối mũ, nên các $C_{i,j}$ mà lớn hơn C_1 sẽ là các biến ngẫu nhiên phân phối mũ độc lập, có tốc độ 1. Do vậy, C_2 bằng C_1 cộng với giá trị nhỏ nhất trong $(n - 1)^2$ biến ngẫu nhiên phân phối mũ độc lập có tốc độ 1. Tương tự, C_3 bằng C_2 cộng với giá trị nhỏ nhất trong $(n - 2)^2$ biến ngẫu nhiên phân phối mũ tốc độ 1, và cứ như thế. Vậy nên, ta có

$$\begin{aligned}\mathbb{E}[C_1] &= 1/n^2 \\ \mathbb{E}[C_2] &= \mathbb{E}[C_1] + 1/(n - 1)^2 \\ \mathbb{E}[C_3] &= \mathbb{E}[C_2] + 1/(n - 2)^2 \\ &\dots \\ \mathbb{E}[C_j] &= \mathbb{E}[C_{j-1}] + 1/(n - j + 1)^2 \\ &\dots \\ \mathbb{E}[C_n] &= \mathbb{E}[C_{n-1}] + 1\end{aligned}$$

Suy ra

$$\begin{aligned}\mathbb{E}[C_1] &= 1/n^2 \\ \mathbb{E}[C_2] &= 1/n^2 + 1/(n - 1)^2 \\ &\dots \\ \mathbb{E}[C_n] &= 1/n^2 + 1/(n - 1)^2 + \dots + 1\end{aligned}$$

Cộng tất cả các $\mathbb{E}[C_i]$ lại với nhau ta được

$$\mathbb{E}_B[\text{tổng chi phí}] = n/n^2 + (n - 1)/(n - 1)^2 + \dots + 1 = \sum_{i=1}^n \frac{1}{i}$$

Như vậy, tổng chi phí trung bình khi dùng hai thuật toán trên là như nhau.

Ví dụ. Ta có n vật, với các trọng lượng khác nhau, cần được phân phối vào hai thùng sao cho trọng lượng chênh lệch giữa hai thùng là nhỏ nhất. Một thuật toán khá đơn giản

là ta sẽ đặt vật 1 vào thùng 1 và vật 2 vào thùng 2, sau đó đặt lần lượt từng vật trong các vật còn lại vào thùng hiện đang có trọng lượng thấp hơn. Chẳng hạn, nếu $n = 6$, và các vật lần lượt có các trọng lượng là 5, 8, 12, 6, 2, 10, thì vật 1 sẽ được đặt vào thùng 1, vật 2 được đặt vào thùng 2, vật 3 được đặt vào thùng 1, vật 4, 5, 6 được đặt vào thùng 2 và trọng lượng chênh lệch giữa hai thùng lúc này là bằng 9.

(a) Tìm trọng lượng trung bình giữa hai thùng khi n vật có trọng lượng là các biến ngẫu nhiên độc lập cùng phân phối mũ với tốc độ λ .

(b) Tìm trọng lượng trung bình của thùng nặng hơn.

Lời giải. Không cần bất kỳ tính toán nào ta cũng có thể kết luận rằng trọng lượng chênh lệch trung bình giữa hai thùng là $1/\lambda$. Đó là vì mỗi lần có một vật mới được đặt vào một thùng nào đó, do khối lượng chênh lệch giữa hai thùng là một biến ngẫu nhiên không có trí nhớ, nên nó có phân phối mũ với tốc độ λ . (Chẳng hạn, xét bài toán tương tự: Có hai đèn pin, mỗi đèn cần một pin để sáng, và giả sử ta có trong tay n pin mà thời gian chạy của chúng là các biến ngẫu nhiên có phân phối mũ cùng tốc độ λ . Giả sử, mỗi lần có một pin nào đó chạy hết thì nó sẽ được thay thế bởi một pin mới. Khi đó, có thể thấy ngay rằng do tính chất không có trí nhớ của biến ngẫu nhiên phân phối mũ nên tại thời điểm mà chỉ có một đèn đang sáng thì thời gian sáng còn lại của nó sẽ là một biến ngẫu nhiên có phân phối mũ với tốc độ λ).

Để xác định trọng lượng trung bình của thùng nặng hơn, ta gọi X là trọng lượng của thùng nặng hơn, và Y là trọng lượng của thùng nhẹ hơn. Khi đó, vì $X + Y$ là tổng trọng lượng của tất cả các vật, nên

$$\mathbb{E}[X] + \mathbb{E}[Y] = n/\lambda$$

Hơn nữa, từ câu (a) ta có

$$\mathbb{E}[X] - \mathbb{E}[Y] = 1/\lambda$$

Do vậy,

$$\mathbb{E}[X] = \frac{n+1}{2\lambda}, \quad \mathbb{E}[Y] = \frac{n-1}{2\lambda}$$

□

Ví dụ. Một biến thể của thuật toán cho bài toán trong ví dụ trên, thường cho kết quả cải tiến hơn, là đầu tiên ta sẽ sắp xếp các vật theo thứ tự trọng lượng giảm dần, sau đó áp dụng thuật toán. Với cách này, tìm trọng lượng chênh lệch giữa các thùng khi $n = 3$, biết rằng trọng lượng các vật có phân phối mũ với cùng tốc độ λ .

Lời giải. Ký hiệu $X_{(i)}$ là trọng lượng của vật nhỏ thứ i trong 3 trọng lượng, $i = 1, 2, 3$. Do tính không có trí nhớ của biến ngẫu nhiên phân phối mũ, cùng với kết quả giá trị nhỏ nhất trong tập các biến ngẫu nhiên phân phối mũ thì cũng có phân phối mũ với tốc độ

bằng tổng tốc độ của các biến ngẫu nhiên trong tập, nên ta có

$X_{(1)}$	là biến ngẫu nhiên phân phối mũ có tốc độ	3λ
$X_{(2)} - X_{(1)}$	là biến ngẫu nhiên phân phối mũ có tốc độ	2λ
$X_{(3)} - X_{(2)}$	là biến ngẫu nhiên phân phối mũ có tốc độ	λ

Hơn nữa, cũng từ tính chất không có trí nhớ của biến ngẫu nhiên phân phối mũ mà ba biến ngẫu nhiên trên là độc lập với nhau. Do đó, nếu gọi Y và Z là các biến ngẫu nhiên độc lập và có phân phối mũ với các tốc độ tương ứng là λ và 3λ , thì ta thấy

$$\begin{aligned}
 \mathbb{E}[|X_{(3)} - (X_{(1)} + X_{(2)})|] &= \mathbb{E}[|X_{(3)} - X_{(1)} - X_{(2)}|] \\
 &= \mathbb{E}[|Y - Z|] \\
 &= \mathbb{E}[|Y - Z| \mid Y < Z] \mathbb{P}\{Y < Z\} + \mathbb{E}[|Y - Z| \mid Y > Z] \mathbb{P}\{Y > Z\} \\
 &= \mathbb{E}[Z - Y \mid Y < Z] \frac{1}{4} + \mathbb{E}[Y - Z \mid Y > Z] \frac{3}{4} \\
 &= \frac{1}{3\lambda} \frac{1}{4} + \frac{1}{\lambda} \frac{3}{4} \quad (\text{do tính không có trí nhớ}) \\
 &= \frac{5}{6\lambda}
 \end{aligned}$$

□

Ví dụ tiếp theo sẽ giới thiệu một cách tiếp cận mang tính xác suất cho việc thử lại các đồng nhất thức đa thức.

Ví dụ. Xét $n + 1$ vật có thời gian sống là $X, Y_i, i = 1, \dots, n$ độc lập với nhau, với X có phân phối mũ tốc độ λ , và $Y_i, i = 1, \dots, n$ có phân phối mũ cùng tốc độ μ . Tìm xác suất để vật có thời gian sống X sống sót lâu nhất.

Lời giải. Xác suất cần tìm là

$$p = \mathbb{P}\{X = \max(X, Y_1, \dots, Y_n)\}$$

Ta định nghĩa các biến cố $A_1, \dots, A_n$ như sau

$$A_i = \{ \text{giá trị nhỏ thứ } i \text{ trong dãy } X, Y_1, \dots, Y_n \text{ là một } Y_k \text{ nào đó} \}$$

Khi đó

$$p = \mathbb{P}\{A_1 A_2 \dots A_n\} = \mathbb{P}\{A_1\} \mathbb{P}\{A_2 \mid A_1\} \dots \mathbb{P}\{A_n \mid A_1 \dots A_{n-1}\}$$

Bây giờ, giả sử đồng nhất mỗi vật với biến ngẫu nhiên chỉ thời gian sống của nó. Dùng tính chất không trí nhớ của biến ngẫu nhiên phân phối mũ, ta suy ra được với điều kiện j vật đầu tiên bị chết là các vật Y , thì tại thời điểm vật thứ j bị chết, thời gian sống của

vật X sẽ có phân phối mũ với tốc độ λ và thời gian sống của $n - j$ vật Y còn lại sẽ có phân phối mũ với tốc độ μ . Do vậy,

$$\mathbb{P}\{A_n \mid A_1 \dots A_{n-1}\} = \frac{(n-j)\mu}{(n-j)\mu + \lambda}$$

Suy ra

$$p = \prod_{j=1}^n \frac{(n-j)\mu}{(n-j)\mu + \lambda} = \prod_{i=1}^n \frac{i\mu}{i\mu + \lambda}$$

Tuy nhiên, có một cách tiếp cận khác để giải bài toán này, đó là đặt điều kiện lên X .

$$\begin{aligned} p &= \int_0^\infty \mathbb{P}\{X = \max(X, Y_1, \dots, Y_n) \mid X = x\} \lambda e^{-\lambda x} dx \\ &= \int_0^\infty \mathbb{P}\{Y_i < x, i = 1, \dots, n\} \lambda e^{-\lambda x} dx \quad (\text{do sự độc lập}) \\ &= \int_0^\infty (1 - e^{-\mu x})^n \lambda e^{-\lambda x} dx \\ &= \int_0^\infty \sum_{i=0}^n C_n^i (-1)^i e^{-i\mu x} \lambda e^{-\lambda x} dx \\ &= \sum_{i=0}^n (-1)^i C_n^i \lambda \int_0^\infty e^{-i\mu x} e^{-\lambda x} dx \\ &= \sum_{i=0}^n \frac{(-1)^i C_n^i \lambda}{\lambda + i\mu} \end{aligned}$$

Từ hai phương trình biểu diễn p , ta có đồng nhất thức

$$\prod_{i=1}^n \frac{i\mu}{i\mu + \lambda} = \sum_{i=0}^n \frac{(-1)^i C_n^i \lambda}{\lambda + i\mu}$$

Bây giờ giả sử, với một cách độc lập, ta muốn thử lại đồng nhất thức trên. Rõ ràng đồng nhất thức trên tương đương với

$$\prod_{i=1}^n (\lambda + i\mu) \left[\prod_{i=1}^n \frac{i\mu}{i\mu + \lambda} - \sum_{i=0}^n \frac{(-1)^i C_n^i \lambda}{\lambda + i\mu} \right] = 0$$

Do vế trái là một đa thức, nên đồng nhất thức trên biểu diễn một đa thức riêng theo λ và μ đồng nhất bằng 0. Nhưng vì một đa thức khác không theo hai (hoặc một số hữu hạn) biến không thể đồng nhất bằng 0 trên một tập có độ đo (Lebesgue) dương, nên nếu đa thức là không đồng nhất bằng 0, thì xác suất để nó bằng 0 tại các giá trị μ và λ ngẫu nhiên nào đó là bằng 0. Do vậy, tất cả điều chúng ta cần làm để thiết lập đẳng thức này là tạo ra hai biến ngẫu nhiên độc lập, có phân phối đều trên $(0, 1)$, λ sẽ nhận giá trị của biến ngẫu nhiên thứ nhất, μ sẽ nhận giá trị của biến ngẫu nhiên thứ hai, và kiểm tra xem đồng nhất thức trên có thỏa mãn với các giá trị của μ và λ không. Nếu có, ta có thể kết luận, với xác suất 1, đồng nhất thức trên đúng với mọi μ và λ .

□

1.9 Các định lý về giới hạn

Luật số lớn được đưa ra vào thế kỷ *XVII*. Luật số lớn chỉ ra rằng khi ta chọn ngẫu nhiên các giá trị (mẫu thử) trong một dãy các giá trị (quần thể), nếu kích thước dãy mẫu thử càng lớn thì các đặc trưng thống kê (trung bình, phương sai...) của mẫu thử càng "gần" với các đặc trưng thống kê của quần thể. Các nhà Toán học phân biệt hai phát biểu khác nhau của Luật số lớn, là Luật số lớn yếu và Luật số lớn mạnh.

1.9.1 Luật số lớn yếu

Luật số lớn yếu còn được gọi là **Định lý Khintchine**: nếu $X_1, \dots, X_n$ là n biến ngẫu nhiên độc lập, cùng phân phối xác suất với phương sai hữu hạn và kỳ vọng $\mathbb{E}[X]$ thì, với mọi số thực ε dương, xác suất để khoảng cách giữa trung bình tích lũy $\frac{X_1 + \dots + X_n}{n}$ và kỳ vọng $\mathbb{E}[X]$ lớn hơn ε là tiến về 0 khi n tiến đến vô cùng, tức

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mathbb{E}[X] \right| \geq \varepsilon \right\} = 0$$

Định lý được chứng minh bằng cách sử dụng Bất đẳng thức *Bienaymé – Tchebyshev* của Chebyshev, mà sẽ được trình bày rõ hơn trong phần sau.

$$\mathbb{P}\{|Y - \mathbb{E}[Y]| \geq \varepsilon\} \leq \frac{\text{Var}(Y)}{\varepsilon^2}$$

Biến ngẫu nhiên $Y_n = \frac{X_1 + \dots + X_n}{n}$ có kỳ vọng

$$\mathbb{E}[Y_n] = \frac{n\mathbb{E}[X]}{n} = \mathbb{E}[X]$$

và phương sai

$$\text{Var}(Y_n) = \frac{n\text{Var}(X_n)}{n^2} = \frac{\text{Var}(X_n)}{n}$$

Từ bất đẳng thức Chebyshev, ta có

$$\mathbb{P} \left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mathbb{E}[X] \right| \geq \varepsilon \right\} \leq \frac{\text{Var}(Y_n)}{n\varepsilon^2}$$

Vế phải tiến về 0 khi n tiến ra vô cùng. Định lý được chứng minh.

Theo định nghĩa hội tụ của biến ngẫu nhiên, thì Y_n hội tụ theo xác suất về $\mathbb{E}[X]$.

1.9.2 Luật số lớn mạnh

Nếu $X_1, X_2, \dots, X_n$ là n biến ngẫu nhiên độc lập, cùng phân phối xác suất với kỳ vọng $\mathbb{E}[X]$ hữu hạn, thì

$$\mathbb{P} \left\{ \lim_{n \rightarrow +\infty} \frac{X_1 + \dots + X_n}{n} = \mathbb{E}[X] \right\} = 1$$

Nghĩa là trung bình tích lũy $\frac{X_1 + \dots + X_n}{n}$ hội tụ hầu chắc chắn về $\mathbb{E}[X]$.

Như vậy, Luật số lớn (yếu hay mạnh) khẳng định rằng khi n đủ lớn, trung bình số học có thể xấp xỉ kỳ vọng.

1.9.3 Định lý giới hạn trung tâm

Trong xác suất, Định lý giới hạn trung tâm là một định lý nổi tiếng và có vai trò quan trọng, nó là kết quả về sự hội tụ yếu của một dãy các biến ngẫu nhiên. Với định lý này, ta có kết quả là tổng của các biến ngẫu nhiên độc lập và phân phối đồng nhất theo cùng một phân phối xác suất sẽ hội tụ về một biến ngẫu nhiên nào đó.

Trong trường hợp đơn giản nhất, các biến ngẫu nhiên là độc lập, có cùng kỳ vọng và phương sai. Một cách tổng quát, tổng của các biến ngẫu nhiên sẽ tăng vô định khi số biến ngẫu nhiên tăng. Do đó để có một kết quả hữu hạn, ta hạn chế sự tăng của tổng bằng cách lấy tổng trừ đi giá trị trung bình và rút gọn bằng cách chia cho căn bậc hai của phương sai. Với một số điều kiện nữa thì phân phối xác suất của biến ngẫu nhiên giản lược sẽ hội tụ về một phân phối chuẩn.

Sự hội tụ được đảm bảo trong trường hợp đơn giản này. Tuy nhiên cũng tồn tại sự hội tụ trong trường hợp các biến ngẫu nhiên không cùng phân phối, nhưng vẫn phải đảm bảo điều kiện không có biến ngẫu nhiên nào có phân phối trội hơn hoặc gây ảnh hưởng đến phân phối của các biến ngẫu nhiên khác. Điều này được đảm bảo bởi điều kiện Lindeberg và điều kiện Lyapunov. Một số phiên bản khác của định lý cũng cho phép sự phụ thuộc yếu giữa các biến ngẫu nhiên.

Ở đây chỉ đề cập đến Định lý giới hạn trung tâm cho trường hợp các phân phối có phương sai hữu hạn.

Nếu $X_1, X_2, \dots$ là các biến ngẫu nhiên độc lập và có cùng phân phối với kỳ vọng μ và độ lệch tiêu chuẩn σ , thì $S_n = \sum_{i=1}^n X_i$ sẽ có phân phối xấp xỉ phân phối chuẩn, khi n đủ lớn.

Như đã biết, S_n có kỳ vọng $n\mu$ và độ lệch tiêu chuẩn $\sigma/\sqrt{n}$. Để làm rõ hơn sự hội tụ này, ta đặt

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

để có được kỳ vọng và độ lệch tiêu chuẩn của Z_n lần lượt là 0 và 1.

Nếu phân phối của Z_n hội tụ về phân phối chuẩn $N(0, 1)$ khi n tiến đến vô cùng (tức là hội tụ theo phân phối), thì cũng có nghĩa là: nếu Φ là hàm phân phối tích lũy của $N(0, 1)$, thì với mọi số thực z ,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{Z_n < z\} = \Phi(z)$$

hay một cách tương đương

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} < z \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-x^2/2} dx$$

Định lý trên cho thấy dù X_i có thể là biến ngẫu nhiên rời rạc hay liên tục bất kỳ, nhưng tổng chuẩn hóa Z_n của chúng, khi n đủ lớn, sẽ tiệm cận với phân phối tiêu chuẩn. Điều này giải thích vì sao phân phối chuẩn lại phổ biến và quan trọng trong thực tế. Định lý giới hạn trung tâm được xem như là **hòn đá tảng** của thống kê. Trong thực tế, với n xấp xỉ 30, ta có thể ứng dụng Định lý giới hạn trung tâm. Dưới đây, ta đưa ra các kết quả chính khi áp dụng Định lý giới hạn trung tâm cho khối đa diện ngẫu nhiên Gauss.

1.9.4 Định lý giới hạn trung tâm cho khối đa diện Gauss

Lấy ngẫu nhiên n điểm độc lập trong $\mathbb{R}^d$ theo phân phối chuẩn tắc. Khi đó, bao lồi K_n của n điểm trên được gọi là khối đa diện ngẫu nhiên Gauss. Ta sẽ chứng minh rằng thể tích và số mặt của K_n thỏa mãn định lý giới hạn trung tâm. Dưới đây là các kết quả chính.

Ký hiệu $\Psi_d = \Psi$ là phân phối chuẩn tắc trong $\mathbb{R}^d$, và hàm mật độ của nó là

$$\psi_d = \psi = \frac{1}{(2\pi)^{d/2}} \exp \left\{ \frac{-x^2}{2} \right\}$$

trong đó, $x^2 = |x|^2$ là bình phương của chuẩn Euclide của x trong $\mathbb{R}_d$. Ta sẽ chỉ sử dụng ký hiệu này với $d \geq 2$, với $d = 1$ thì phân phối chuẩn tắc có hàm mật độ là

$$\phi = \frac{1}{(2\pi)^{1/2}} \exp \left\{ \frac{-x^2}{2} \right\}$$

với hàm phân phối Φ .

Với $d \geq 2$ và lấy ngẫu nhiên một tập gồm n điểm độc lập $X_n = \{x_1, x_2, \dots, x_n\}$ từ $\mathbb{R}_d$ theo phân phối chuẩn tắc Ψ , khi đó bao lồi $K_n = \text{Conv}(x_1, x_2, \dots, x_n)$ của n điểm trên được gọi là **khối đa diện ngẫu nhiên Gauss**. Đây là một trong những mô hình chính của lý thuyết khối đa diện ngẫu nhiên, được đưa ra bởi *Rényi* và *Sulanke* vào những năm 1960. Mục đích chính của lý thuyết này là để nghiên cứu các phân phối của các đặc trưng chính (ví dụ như thể tích) của khối đa diện ngẫu nhiên.

Một "hòn đá tảng" trong lý thuyết xác suất là định lý giới hạn trung tâm. Một dãy các biến ngẫu nhiên X_n thỏa mãn định lý giới hạn trung tâm nếu với mọi t ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{X_n - \mathbb{E}X_n}{\sqrt{\text{Var}X_n}} \leq t \right) - \Phi(t) = 0$$

Một phỏng đoán tự nhiên và quan trọng trong lý thuyết các khối đa diện ngẫu nhiên là các đặc trưng chính của K_n thỏa mãn định lý giới hạn trung tâm, khi $n \rightarrow \infty$. Sự phỏng đoán này đã có trong khoảng vài chục năm, và một vài kết quả đặc biệt đã được chứng minh. Ở đây, ta sẽ chỉ đưa ra kết quả liên quan đến thể tích và số mặt của khối đa diện ngẫu nhiên (có lẽ đây là hai tham số được quan tâm nhất).

Cho một đa diện lồi K , kí hiệu $Vol(K)$ và $f_s(K)$ là thể tích và số mặt của khối đa diện K ứng với không gian có số chiều s .

Định lý 1.9.1. *Cho d là một số nguyên cố định lớn hơn 2. Khi đó, tồn tại một hàm $\varepsilon(n) \rightarrow 0$ khi $n \rightarrow \infty$ thỏa mãn:*

$$\left| \mathbb{P} \left(\frac{Vol(K_n) - \mathbb{E}Vol(K_n)}{\sqrt{VarVol(K_n)}} \leq t \right) - \Phi(t) \right| \leq \varepsilon(n)$$

với mọi giá trị t .

Định lý 1.9.2. *Cho d là một số nguyên cố định lớn hơn 2 và s là một số nguyên không âm nhỏ hơn hoặc bằng $d - 1$. Khi đó, tồn tại một hàm $\varepsilon(n) \rightarrow 0$ khi $n \rightarrow \infty$ thỏa mãn:*

$$\left| \mathbb{P} \left(\frac{f_s(K_n) - \mathbb{E}f_s(K_n)}{\sqrt{Varf_s(K_n)}} \leq t \right) - \Phi(t) \right| \leq \varepsilon(n)$$

với mọi giá trị t .

Chú ý rằng, trong hai định lý trên ta có thể lấy $\varepsilon(n) = (\log n)^{-(d-1)/4+o(1)}$

1.10 Bất đẳng thức Markov

Trong lý thuyết xác suất, Bất đẳng thức Markov cho ta xác suất cận trên để một hàm không âm của một biến ngẫu nhiên lớn hơn hoặc bằng một hằng số dương nào đó. Bất đẳng thức này được đặt tên theo nhà toán học người Nga Andrey Markov.

Bất đẳng thức Markov liên quan đến giá trị kỳ vọng trong xác suất, và nó chỉ ra được các cận, tuy có độ chính xác không cao lắm, nhưng vẫn là hữu ích đối với hàm phân phối tích lũy của biến ngẫu nhiên.

Với ngôn ngữ lý thuyết độ đo, Bất đẳng thức Markov được phát biểu rằng: nếu (X, Ω, μ) là một không gian đo được, f là một hàm đo được nhận giá trị thực, và $t > 0$, thì

$$\mu(\{x \in X \mid |f(x)| \geq t\}) \leq \frac{1}{t} \int_X |f| d\mu$$

Đối với trường hợp đặc biệt là khi không gian này có độ đo 1 (tức là không gian xác suất), thì ta có thể phát biểu lại như sau: nếu X là một biến ngẫu nhiên và $a > 0$, thì

$$\mathbb{P}\{|X| \geq a\} \leq \frac{\mathbb{E}[|X|]}{a}$$

Chứng minh.

Đầu tiên ta sẽ chứng minh cho trường hợp không gian xác suất vì trường hợp này đơn giản hơn.

Trường hợp đặc biệt (lý thuyết xác suất). Cho biến cố E bất kỳ, gọi I_E là hàm chỉ tiêu của E . Như vậy, $I_{(|X| \geq a)} = 1$ nếu biến cố $|X| \geq a$ xảy ra, và $I_{(|X| \geq a)} = 0$ nếu biến cố $|X| < a$ xảy ra. Khi đó, với $a > 0$,

$$aI_{(|X| \geq a)} \leq |X|$$

Lấy kỳ vọng hai vế, ta được

$$\mathbb{E}[aI_{(|X| \geq a)}] \leq \mathbb{E}[|X|]$$

Giờ lại để ý rằng, vế trái của bất phương trình trên lại tương đương với

$$a\mathbb{E}[I_{(|X| \geq a)}] = a\mathbb{P}\{|X| \geq a\}$$

Như vậy ta được

$$a\mathbb{P}\{|X| \geq a\} \leq \mathbb{E}[|X|]$$

và vì $a > 0$, nên ta có thể chia hai vế cho a .

Trường hợp tổng quát (lý thuyết độ đo). Cho A là một tập đo được bất kỳ, đặt 1_A là hàm chỉ tiêu của tập A , nghĩa là $1_A(x) = 1$ nếu $x \in A$, và bằng 0 nếu ngược lại. Nếu tập A_t là tập được định nghĩa bởi $A_t = \{x \in X \mid |f(x)| \geq t\}$, thì

$$0 \leq t1_{A_t} \leq |f|1_{A_t} \leq |f|$$

Do đó

$$\int_X t1_{A_t} d\mu \leq \int_{A_t} |f| d\mu \leq \int_X |f| d\mu$$

Giờ lại chú ý rằng vế trái của bất đẳng thức trên tương đương với

$$t \int_X 1_{A_t} d\mu = t\mu(A_t)$$

Do vậy

$$t\mu(\{x \in X \mid |f(x)| \geq t\}) \leq \int_X |f| d\mu$$

và vì $t > 0$ nên ta có thể chia hai vế bất đẳng thức trên cho t , vậy

$$\mu(\{x \in X \mid |f(x)| \geq t\}) \leq \frac{1}{t} \int_X |f| d\mu$$

□

Ví dụ. Bất đẳng thức Boole. Đặt $I\{A_i\}, i = 1, \dots, n$, là hàm chỉ tiêu của các biến cố $A_i, i = 1, \dots, n$. Ký hiệu

$$N = \sum_{i=1}^n I\{A_i\}$$

là số biến cố xảy ra trong n biến cố trên. Áp dụng bất đẳng thức Markov, ta có

$$\mathbb{P}\{N \geq 1\} \leq \mathbb{E}[N] \quad (1.16)$$

Tuy nhiên, vì $N \geq 1$ có nghĩa là có ít nhất một biến cố xảy ra, và

$$\mathbb{E}[N] = \sum_{i=1}^n \mathbb{E}[I\{A_i\}] = \sum_{i=1}^n \mathbb{P}(A_i)$$

nên ta thấy bất đẳng thức (1.16) là tương đương với bất đẳng thức sau, được gọi là **Bất đẳng thức Boole**,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i)$$

□

1.11 Bất đẳng thức Chebyshev

Bất đẳng thức Chebyshev, được đặt theo tên nhà toán học Pafnuty Chebyshev, phát biểu rằng đối với một số liệu mẫu hoặc một hàm phân phối xác suất bất kỳ, hầu hết các giá trị là gần với giá trị trung bình, hơn nữa còn chỉ ra mô tả định lượng cho khái niệm "hầu hết" và "gần với".

Đặc biệt ta có: tính từ tâm giá trị trung bình, có không quá $1/4$ các giá trị lớn hơn 2 lần độ lệch tiêu chuẩn, không quá $1/9$ các giá trị lớn hơn 3 lần độ lệch tiêu chuẩn, không quá $1/25$ các giá trị lớn hơn 5 lần độ lệch tiêu chuẩn,... Tổng quát, có không quá $1/k^2$ các giá trị lớn hơn k lần độ lệch tiêu chuẩn tính từ tâm giá trị trung bình.

Bất đẳng thức Chebyshev có thể được phát biểu một cách tổng quát theo lý thuyết độ đo, phát biểu theo ngôn ngữ xác suất là một trường hợp đặc biệt, trường hợp không gian có độ đo 1.

1.11.1 Phát biểu theo lý thuyết độ đo

Giả sử (X, Ω, μ) là một không gian đo được, và f là một hàm đo được nhận giá trị thực trên X . Khi đó, với mọi số thực $t > 0$ bất kỳ,

$$\mu(\{x \in X : |f(x)| \geq t\}) \leq \frac{1}{t^2} \int_X f^2 d\mu$$

Tổng quát hơn, nếu g là một hàm đo được không âm nhận giá trị thực, không giảm trên miền giá trị của f , thì

$$\mu(\{x \in X : |f(x)| \geq t\}) \leq \frac{1}{(gt)} \int_X g \circ f d\mu$$

với $g(t)$ được định nghĩa như sau

$$g(t) = \begin{cases} t^2 & \text{nếu } t \geq 0 \\ 0 & \text{nếu } t < 0 \end{cases}$$

1.11.2 Phát biểu theo xác suất

Giả sử X là một biến ngẫu nhiên có kỳ vọng μ và phương sai hữu hạn σ^2 . Khi đó, với mọi số thực $k > 0$ bất kỳ,

$$\mathbb{P}\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

Ta chỉ quan tâm trường hợp $k > 1$, vì với trường hợp $k \leq 1$ là hiển nhiên đúng. Phát biểu trên tương đương với

$$\mathbb{P}\{|X - \mu| \geq \alpha\} \leq \frac{\sigma^2}{\alpha^2}$$

Ví dụ như khi $k = \sqrt{2}$, ta sẽ thấy rằng có ít nhất một nửa các giá trị nằm trong khoảng $(\mu - \sqrt{2}\sigma, \mu + \sqrt{2}\sigma)$.

Định lý này là hữu ích, cho dù các cận đưa ra có thể có độ chính xác không cao lắm, nhưng vì nó có thể áp dụng cho biến ngẫu nhiên có phân phối bất kỳ, và vì ta có thể tính các cận này trong khi không cần biết nhiều thông tin về phân phối của biến ngẫu nhiên ngoại trừ giá trị kỳ vọng và phương sai.

Ví dụ. Giả sử X là biến ngẫu nhiên có phân phối mũ với tham số λ . Ta có $\mathbb{E}[X] = 1/\lambda$, $\mathbb{D}(X) = 1/\lambda^2$. Khi đó, cận Chebyshev là

$$\mathbb{P}\left\{\left|X - \frac{1}{\lambda}\right| \geq a\right\} \leq \frac{1}{\lambda^2 a^2}$$

Chứng minh.

Trường hợp lý thuyết độ đo tổng quát. Đặt $A_t = \{x \in X \mid f(x) \geq t\}$, và gọi 1_{A_t} là hàm chỉ tiêu của tập A_t . Khi đó, dễ thấy rằng

$$0 \leq g(t)1_{A_t} \leq g \circ f 1_{A_t} \leq g \circ f$$

Do đó

$$g(t)\mu(A_t) = \int_X g(t)1_{A_t}d\mu \leq \int_{A_t} g(t) \circ f d\mu \leq \int_X g(t) \circ f d\mu$$

Chia hai vế bất đẳng thức trên cho $g(t)$, ta được bất đẳng thức cần chứng minh.

Trường hợp xác suất. Bất đẳng thức Markov phát biểu rằng: với mọi biến ngẫu nhiên nhận giá trị thực Y và với mọi số dương a , ta có

$$\mathbb{P}\{|Y| \geq a\} \leq \frac{\mathbb{E}[|Y|]}{a}$$

Một cách để chứng minh bất đẳng thức Chebyshev là áp dụng bất đẳng thức Markov với biến ngẫu nhiên $Y = (X - \mu)^2$ và $a = (\sigma k)^2$.

Ta cũng có thể chứng minh trực tiếp như sau: Cho A là một biến cố bất kỳ, gọi I_A là hàm chỉ tiêu của A . Khi đó,

$$\begin{aligned} \mathbb{P}\{|X - \mu| \geq k\sigma\} &= \mathbb{E}[I_{|X - \mu| \geq k\sigma}] \\ &= \mathbb{E}[I_{[(X - \mu)/(k\sigma)]^2 \geq 1}] \\ &\leq \mathbb{E}\left[\left(\frac{X - \mu}{k\sigma}\right)^2\right] = \frac{1}{k^2} \frac{\mathbb{E}[(X - \mu)^2]}{\sigma^2} = \frac{1}{k^2} \end{aligned}$$

□

Chương 2

Bất đẳng thức Chernoff và ứng dụng

2.1 Bất đẳng thức Chernoff

Herman Chernoff đã tìm ra một cách để suy ra được cận cho dãy các biến ngẫu nhiên độc lập. Tất cả các cận được suy ra từ cách này của ông ngày nay được biết đến như là cận Chernoff.

Phương pháp của Chernoff tập trung quanh việc xác định cận cho một biến ngẫu nhiên X bằng cách nghiên cứu biến ngẫu nhiên e^{tX} , thay vì X . Có nhiều dạng cận Chernoff, ta thường gặp dạng cơ bản nhất là dạng cộng, hoặc dạng tiện lợi hơn là dạng nhân. **Cận Chernoff** là một hệ quả từ Bất đẳng thức Markov.

2.1.1 Cận Chernoff

Giả sử X có hàm moment tổng quát $\phi(t) = \mathbb{E}[e^{tX}]$. Khi đó, với mọi $c > 0$ bất kỳ,

$$\mathbb{P}\{X \geq c\} \leq e^{-tc}\phi(t), \text{ nếu } t > 0$$

$$\mathbb{P}\{X \leq c\} \leq e^{-tc}\phi(t), \text{ nếu } t < 0$$

Chứng minh.

Với $t > 0$,

$$\mathbb{P}\{X \geq c\} = \mathbb{P}\{e^{tX} \geq e^{tc}\} \leq \mathbb{E}[e^{tX}]e^{-tc} \quad (\text{do bất đẳng thức Markov})$$

Tương tự với $t < 0$.

□

Vì cận Chernoff xác định với mọi t , nên ta có thể đạt được cận tốt nhất bằng cách chọn t sao cho cực tiểu $e^{-tc}\phi(t)$. Nghĩa là, với $t > 0$,

$$\mathbb{P}\{X \geq c\} = \mathbb{P}\{e^{tX} \geq e^{tc}\} \leq \mathbb{E}[e^{tX}]e^{-tc}$$

nên

$$\mathbb{P}\{X \geq c\} \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{tc}}$$

Tương tự, với $t < 0$,

$$\mathbb{P}\{X \leq c\} = \mathbb{P}\{e^{tX} \geq e^{tc}\} \leq \mathbb{E}[e^{tX}]e^{-tc}$$

nên

$$\mathbb{P}\{X \leq c\} \leq \min_{t<0} \frac{\mathbb{E}[e^{tX}]}{e^{tc}}$$

Ví dụ. Nếu X là biến ngẫu nhiên có phân phối Poisson với trung bình λ , thì $\phi(t) = \exp\{\lambda(e^t - 1)\}$. Do vậy, cận Chernoff sẽ là

$$\mathbb{P}\{X \geq n\} \leq \exp\{\lambda(e^t - 1) - nt\}$$

Giá trị của t mà làm cực tiểu vế phải là giá trị mà làm cực tiểu $\lambda(e^t - 1) - nt$, chính là giá trị t để $e^t = n/\lambda$. Để t không âm thì cần $n/\lambda \geq 1$. Vậy

$$\mathbb{P}\{X \geq n\} \leq \exp\left\{\lambda\left(\frac{n}{\lambda}\right)\right\}(\lambda/n)^n = \frac{e^{-\lambda}(\lambda e)^n}{n^n}, n \geq \lambda$$

□

2.1.2 Bất đẳng thức Chernoff

Một trường hợp đặc biệt của cận Chernoff là Bất đẳng thức Chernoff. Giả sử $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập có $\mathbb{E}[X_i] = 0$ và $|X_i| \leq 1$ với mọi i . Đặt

$$X = \sum_{i=1}^n X_i$$

và giả sử σ^2 là phương sai của X . Khi đó,

$$\mathbb{P}\{|X| \geq k\sigma\} \leq 2e^{-k^2/4}, \text{ với mọi } 0 \leq k \leq 2\sigma$$

Ví dụ. Giả sử $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập. Đặt

$$X = \sum_{i=1}^n X_i$$

Khi đó, nếu $\mathbb{P}\{X_i = 1\} = \mathbb{P}\{X_i = -1\} = 1/2$ và $a > 0$, thì

$$\mathbb{P}\{X \geq a\} \leq e^{-a^2/2n}$$

Từ đó suy ra

$$\mathbb{P}\{|X| \geq a\} \leq 2e^{-a^2/2n}$$

□

2.2 Cận Chernoff cho tổng các biến ngẫu nhiên Bernoulli độc lập.

Bổ đề 2.2.1. Cho $0 \leq p \leq 1$, khi đó

$$pe^{t(1-p)} + (1-p)e^{-tp} \leq e^{t^2/8} \quad (2.1)$$

Bổ đề 2.2.1 được dùng để tìm cận Chernoff cho tổng các biến ngẫu nhiên Bernoulli độc lập.

Bổ đề 2.2.2. Cho $X_i, i = 1, \dots, n$ là các biến ngẫu nhiên Bernoulli độc lập, đặt $X = \sum_{i=1}^n X_i$. Khi đó, với mọi $a > 0$

$$\mathbb{P}\{X - \mathbb{E}[X] \geq a\} \leq e^{-2a^2/n} \quad (2.2)$$

$$\mathbb{P}\{X - \mathbb{E}[X] \leq -a\} \leq e^{-2a^2/n} \quad (2.3)$$

Chứng minh.

Với $a > 0$ và $t > 0$ bất kỳ

$$\begin{aligned} \mathbb{P}\{X - \mathbb{E}[X] \geq a\} &= \mathbb{P}\{e^{t(X - \mathbb{E}[X])} \geq e^{ta}\} \\ &\leq e^{-ta} \mathbb{E}[e^{t(X - \mathbb{E}[X])}] \\ &= e^{-ta} \mathbb{E}\left[\exp\left\{\sum_{i=1}^n t(X_i - \mathbb{E}[X_i])\right\}\right] \\ &= e^{-ta} \mathbb{E}\left[\prod_{i=1}^n e^{t(X_i - \mathbb{E}[X_i])}\right] \\ &= e^{-ta} \prod_{i=1}^n \mathbb{E}[e^{t(X_i - \mathbb{E}[X_i])}] \quad (\text{do tính chất độc lập}) \end{aligned} \quad (2.4)$$

Vì X_i là các biến ngẫu nhiên Bernoulli độc lập với tham số $(1, p)$ nên theo (2.1), ta có

$$\begin{aligned} \mathbb{E}[e^{t(X_i - \mathbb{E}[X_i])}] &= pe^{t(1-p)} + (1-p)e^{-tp} \\ &\leq e^{t^2/8} \end{aligned}$$

Sử dụng kết quả này vào (2.4), ta được

$$\mathbb{P}\{X - \mathbb{E}[X] \geq a\} \leq e^{-ta} e^{nt^2/8}$$

Lấy $t = 4a/n$ ta được (2.2).

Bằng lập luận tương tự ta cũng chứng minh được (2.3). $\square$

Ví dụ. Giả sử có n quan sát Bernoulli độc lập, xác suất thành công của mỗi quan sát là p . Tìm giới hạn xác suất sao cho số thành công ít nhất là $1 + \alpha$ lần giá trị kỳ vọng của chúng, với $\alpha > 0$.

Lời giải.

Ký hiệu X_n là số phép thử thành công. Khi đó, áp dụng bất đẳng thức (2.2), ta có

$$\mathbb{P}\{X_n - \mathbb{E}[X_n] \geq \alpha \mathbb{E}[X_n]\} \leq \exp\{-2(\alpha \mathbb{E}[X_n])^2/n\}$$

Vì $\mathbb{E}[X_n] = np$, nên

$$\mathbb{P}\{X_n - \mathbb{E}[X_n] \geq \alpha \mathbb{E}[X_n]\} \leq \exp\{-2n\alpha^2 p^2\}$$

Điều này chỉ ra rằng, xác suất trên sẽ giảm tới 0 (theo hàm mũ) khi $n \rightarrow \infty$. □

Nếu X là biến ngẫu nhiên nhị thức với các tham số (n, p) , thì các cận Chernoff (2.2) và (2.3) sẽ là

$$\mathbb{P}\{|X - np| \geq a\} \leq 2e^{-2a^2/n}$$

Khi p nhỏ, các cận Chernoff có thể được cải tiến tốt hơn như sau

$$\mathbb{P}\{|X - np| \geq a\} \leq 2e^{-a^2/(3np)}$$

Chương 3

Xấp xỉ Poisson

3.1 Biến ngẫu nhiên Bernoulli

3.1.1 Biến ngẫu nhiên Bernoulli và biến ngẫu nhiên nhị thức

Giả sử có một phép thử, hay một thí nghiệm, mà kết quả của nó được xem như là hoặc một thành công, hoặc một thất bại. Nếu ta đặt $X = 1$ khi kết quả phép thử là một thành công và $X = 0$ khi kết quả phép thử là một thất bại, thì hàm khối lượng xác suất của X là

$$\begin{aligned}p(0) &= \mathbb{P}\{X = 0\} = 1 - p \\p(1) &= \mathbb{P}\{X = 1\} = p\end{aligned}\tag{3.1}$$

với $p, 0 \leq p \leq 1$, là xác suất để phép thử thành công.

Một biến ngẫu nhiên X được gọi là **biến ngẫu nhiên Bernoulli** (được đặt tên theo nhà toán học người Thụy Sĩ James Bernoulli) nếu hàm khối lượng xác suất của nó được cho bởi phương trình (3.1) với $p \in (0, 1)$ nào đó.

Bây giờ giả sử có n phép thử độc lập, mỗi phép thử có kết quả thành công với xác suất p và thất bại với xác suất $1 - p$, được thực hiện. Nếu gọi X biểu diễn cho số thành công xảy ra trong n phép thử trên, thì X được gọi là một **biến ngẫu nhiên nhị thức** với các tham số (n, p) . Như vậy, biến ngẫu nhiên Bernoulli chính là biến ngẫu nhiên nhị thức với các tham số $(1, p)$.

Hàm khối lượng xác suất của một biến ngẫu nhiên nhị thức với các tham số (n, p) được cho bởi

$$p(i) = C_n^i p^i (1 - p)^{n-i}\tag{3.2}$$

Phương trình (3.2) có được đầu tiên là do xác suất của một dãy riêng bất kỳ n kết quả gồm i thành công và $n - i$ thất bại là bằng $p^i (1 - p)^{n-i}$, do giả thiết các phép thử độc

lập. Và vì có C_n^i dãy khác nhau của n kết quả mà trong đó có i thành công và $n - i$ thất bại, nên ta có phương trình (3.2). Chẳng hạn, nếu $n = 4, i = 2$, thì có $C_4^2 = 6$ cách để trong bốn phép thử này có hai thành công, cụ thể,

$$(tc, tc, tb, tb), (tc, tb, tc, tb), (tc, tb, tb, tc), (tb, tc, tc, tb), (tb, tc, tb, tc), (tb, tb, tc, tc)$$

(với giả sử kết quả (tc, tc, tb, tb) có nghĩa là hai phép thử đầu tiên thành công và hai phép thử sau bị thất bại). Vì mỗi một trong các kết quả này là có xác suất xảy ra bằng $p^2(1 - p)^2$, nên xác suất cần tìm cho biến cố có hai thành công trong bốn phép thử sẽ là $C_4^2 p^2 (1 - p)^2$.

Chú ý rằng, vì tổng tất cả các xác suất là bằng 1, nên

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^n C_n^i p^i (1 - p)^{n-i} = [p + (1 - p)]^n = 1$$

Ví dụ. Có năm đồng xu cân bằng được tung lên, nếu giả thiết các kết quả là độc lập, hãy tìm hàm khối lượng xác suất của số mặt ngửa xuất hiện.

Lời giải. Nếu gọi X là số mặt ngửa xuất hiện, thì X là một biến ngẫu nhiên nhị thức với các tham số $(n = 5, p = 1/2)$. Do đó, từ phương trình (3.2), ta có

$$\mathbb{P}\{X = 0\} = C_5^0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

$$\mathbb{P}\{X = 1\} = C_5^1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = \frac{5}{32}$$

$$\mathbb{P}\{X = 2\} = C_5^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = \frac{10}{32}$$

$$\mathbb{P}\{X = 3\} = C_5^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32}$$

$$\mathbb{P}\{X = 4\} = C_5^4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^1 = \frac{5}{32}$$

$$\mathbb{P}\{X = 5\} = C_5^5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^0 = \frac{1}{32}$$

□

3.1.2 Các tính chất của biến ngẫu nhiên nhị thức

Bây giờ ta sẽ nghiên cứu các tính chất của một biến ngẫu nhiên nhị thức với các tham số n và p . Đầu tiên, ta tính giá trị kỳ vọng và phương sai của nó.

$$\begin{aligned}\mathbb{E}[X^k] &= \sum_{i=0}^n i^k C_n^i p^i (1-p)^{n-i} \\ &= \sum_{i=1}^n i^k C_n^i p^i (1-p)^{n-i}\end{aligned}$$

Dùng đồng nhất thức

$$iC_n^i = nC_{n-1}^{i-1}$$

suy ra

$$\begin{aligned}\mathbb{E}[X^k] &= np \sum_{i=1}^n i^{k-1} C_{n-1}^{i-1} p^{i-1} (1-p)^{n-i} \\ &= np \sum_{j=0}^{n-1} (j+1)^{k-1} C_{n-1}^j p^j (1-p)^{n-1-j} \quad (\text{với } j = i-1) \\ &= np \mathbb{E}[(Y+1)^{k-1}]\end{aligned}$$

với Y là một biến ngẫu nhiên nhị thức với các tham số $n-1, p$. Thế $k=1$ vào phương trình trên, ta được

$$\mathbb{E}[X] = np$$

Thế $k=2$ vào phương trình trên, ta được

$$\begin{aligned}\mathbb{E}[X^2] &= np \mathbb{E}[Y+1] \\ &= np[(n-1)p + 1]\end{aligned}$$

Vì $\mathbb{E}[X] = np$ nên

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= np[(n-1)p + 1] - (np)^2 \\ &= np(1-p)\end{aligned}$$

Vậy, nếu X là một biến ngẫu nhiên nhị thức với các tham số n và p , thì

$$\begin{aligned}\mathbb{E}[X] &= np \\ \text{Var}(X) &= np(1-p)\end{aligned}$$

Mệnh đề sau sẽ chỉ rõ hàm khối lượng xác suất của một biến ngẫu nhiên nhị thức đầu tiên sẽ tăng lên rồi sau đó giảm xuống như thế nào.

Mệnh đề 3.1.1. Nếu X là một biến ngẫu nhiên nhị thức với các tham số (n, p) , với $0 < p < 1$, thì khi k chạy từ 0 đến n , $\mathbb{P}\{X = k\}$ đầu tiên sẽ tăng dần một cách đơn điệu và sau đó giảm dần một cách đơn điệu, đạt giá trị lớn nhất khi k là số nguyên lớn nhất mà nhỏ hơn hoặc bằng $(n + 1)p$.

Chứng minh.

Ta sẽ chứng minh mệnh đề bằng việc xét $\mathbb{P}\{X = k\}/\mathbb{P}\{X = k - 1\}$ và việc xác định xem với giá trị nào của k thì phân thức trên lớn hơn hoặc nhỏ hơn 1. Ta có

$$\begin{aligned} \frac{\mathbb{P}\{X = k\}}{\mathbb{P}\{X = k - 1\}} &= \frac{\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}}{\frac{n!}{(n-k+1)!(k-1)!}p^{k-1}(1-p)^{n-k+1}} \\ &= \frac{(n-k+1)p}{k(1-p)} \end{aligned}$$

Do vậy, $\mathbb{P}\{X = k\} \geq \mathbb{P}\{X = k - 1\}$ khi và chỉ khi

$$(n - k + 1)p \geq k(1 - p)$$

hay tương đương, nếu và chỉ nếu

$$k \leq (n + 1)p$$

và mệnh đề được chứng minh. □

3.2 Biến ngẫu nhiên Poisson

3.2.1 Biến ngẫu nhiên Poisson và mẫu Poisson

Một biến ngẫu nhiên X được gọi là một **biến ngẫu nhiên Poisson** với tham số $\lambda > 0$ nếu

$$\mathbb{P}\{X = i\} = e^{-\lambda} \lambda^i / i!, \quad i = 0, 1, \dots$$

Hàm moment tổng quát của nó được suy ra như sau:

$$\begin{aligned} \phi(t) &= \mathbb{E}[e^{tX}] \\ &= \sum_{i=0}^{\infty} e^{it} e^{-\lambda} \lambda^i / i! \\ &= e^{-\lambda} \sum_{i=0}^{\infty} (e^t \lambda)^i / i! \\ &= \exp\{\lambda(e^t - 1)\} \end{aligned}$$

Đạo hàm ta được

$$\begin{aligned}\phi'(t) &= \lambda e^t \exp\{\lambda(e^t - 1)\} \\ \phi''(t) &= (\lambda e^t)^2 \exp\{\lambda(e^t - 1)\} + \lambda e^t \exp\{\lambda(e^t - 1)\}\end{aligned}$$

Suy ra

$$\begin{aligned}\mathbb{E}[X] &= \phi'(0) = \lambda \\ \text{Var}(X) &= \phi''(0) - \mathbb{E}^2[X] = \lambda^2 + \lambda - \lambda^2 = \lambda\end{aligned}$$

Biến ngẫu nhiên Poisson thường được biết đến như là một xấp xỉ cho phân phối của biến ngẫu nhiên nhị thức (n, p) khi n lớn, và p nhỏ. Nếu X là một biến ngẫu nhiên nhị thức (n, p) , thì với $\lambda = np$, ta có

$$\begin{aligned}\mathbb{P}\{X = i\} &= \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\ &= \frac{n!}{i!(n-i)!} (\lambda/n)^i (1-\lambda/n)^{n-i} \\ &= \frac{n(n-1)\dots(n-i+1)}{n^i} \frac{\lambda^i (1-\lambda/n)^n}{i! (1-\lambda/n)^i}\end{aligned}$$

Với n lớn và $p = \lambda/n$ nhỏ,

$$\mathbb{P}\{X = i\} \approx e^{-\lambda} \lambda^i / i!$$

Từ trên suy ra, số thành công trong n phép thử độc lập, trong đó mỗi phép thử là thành công với xác suất p , xấp xỉ một biến ngẫu nhiên Poisson với tham số $\lambda = np$.

Hơn nữa, có thể phát biểu điều này cho trường hợp rộng hơn, đó là không cần các phép thử phải có cùng xác suất thành công, mà chỉ cần các xác suất này là nhỏ. Để có được phát biểu trên cho trường hợp này, ta sẽ giả sử các phép thử là độc lập, phép thử thứ i thành công với xác suất p_i , với các $p_i, i = 1, \dots, n$ là nhỏ. Nếu đặt $X_i = 1$ khi phép thử thứ i thành công, và bằng 0 khi ngược lại, thì số thành công, ký hiệu là X , có thể được biểu diễn bởi

$$X = \sum_{i=1}^n X_i$$

Vì mỗi X_i là một biến ngẫu nhiên Bernoulli (hay nhị thức), nên hàm moment tổng quát của nó là

$$\mathbb{E}[e^{tX_i}] = p_i e^t + 1 - p_i = 1 + p_i(e^t - 1)$$

Bây giờ, dùng kết quả là, với $|x|$ nhỏ,

$$e^x \approx 1 + x$$

nên suy ra được rằng vì $p_i(e^t - 1)$ là nhỏ khi p_i nhỏ, nên

$$\mathbb{E}[e^{tX_i}] \approx \exp\{p_i(e^t - 1)\}$$

Vì hàm moment tổng quát của một tổng các biến ngẫu nhiên độc lập thì bằng tích của các hàm moment tổng quát của chúng, nên từ trên suy ra

$$\mathbb{E}[e^{tX}] \approx \prod_{i=1}^n \exp\{p_i(e^t - 1)\} = \exp\left\{\sum_i p_i(e^t - 1)\right\}$$

với giả thiết X xấp xỉ một biến ngẫu nhiên Poisson với trung bình $\sum_i p_i$.

Tiếp theo, ta xét đến việc các phép thử không những không có cùng xác suất thành công để có số thành công là có phân phối xấp xỉ phân phối Poisson, mà còn không cần đến điều kiện độc lập, tức là giả sử sự độc lập giữa các phép thử là **yếu**. Chẳng hạn, nhắc lại **Bài toán phù hợp**, với n người, mỗi người chọn một mũ một cách ngẫu nhiên các mũ như là việc thực hiện n phép thử, với ta gọi phép thử thứ i là thành công nếu người i chọn được chính cái mũ của mình, thì với A_i là biến cố phép thử thứ i thành công, ta có

$$\mathbb{P}(A_i) = \frac{1}{n} \quad \text{và} \quad \mathbb{P}(A_i | A_j) = \frac{1}{n-1}, j \neq i$$

Do đó, khi xét các phép thử là không độc lập, thì nghĩa là giữa chúng xuất hiện sự phụ thuộc yếu, khi n lớn. Chính vì điều này, cộng với xác suất để phép thử thành công là nhỏ, khi n lớn, suy ra được rằng số người chọn đúng mũ của chính mình sẽ có phân phối xấp xỉ phân phối Poisson với trung bình 1.

Xét một dãy gồm n phép thử, với phép thử thứ i có xác suất thành công là p_i , phát biểu "*nếu tất cả các p_i đều nhỏ và các phép thử là độc lập với nhau từng đôi một hoặc ít nhất là phụ thuộc 'yếu', thì số thành công xuất hiện trong dãy phép thử này sẽ có phân phối xấp xỉ phân phối Poisson với trung bình $\sum_{i=1}^n p_i$* " được gọi là **mẫu Poisson**.

Ví dụ. Bài toán ngày sinh nhật. Đây là một bài toán cổ điển, trong xác suất gọi là Bài toán sinh nhật, trong đó giả thiết rằng mỗi một người trong số n người sẽ có ngày sinh một cách độc lập với nhau và là một trong 365 ngày của năm. Bài toán là hãy xác định giá trị nhỏ nhất của n để có ít nhất hai người có cùng ngày sinh. Vì xác suất để n ngày sinh của n người trên đều khác nhau là bằng $365 \cdot 364 \cdots (365 - n + 1)/365^n$, nên dễ thấy rằng khi $n = 23$ thì xác suất để tất cả n người này đều có ngày sinh khác nhau là nhỏ hơn $1/2$.

Ta có thể xấp xỉ xác suất trên bằng cách dùng mẫu Poisson như sau: ta coi mỗi cặp trong số C_n^2 cặp hai người $i \neq j$ là một phép thử, và ta nói rằng phép thử i, j là một thành công nếu người i và j có cùng ngày sinh. Nếu đặt A_{ij} là biến cố phép thử i, j thành công, thì

C_n^2 biến cố A_{ij} là không độc lập với nhau hoàn toàn, mà giữa chúng có sự phụ thuộc yếu. (Thực vậy, dễ thấy rằng chúng độc lập với nhau từng đôi một, tức là hai biến cố bất kỳ A_{ij} và A_{rs} sẽ độc lập với nhau)

Do $\mathbb{P}(A_{ij}) = 1/365$, nên sẽ hợp lý để giả sử rằng số thành công là có phân phối xấp xỉ Poisson với trung bình $C_n^2/365 = n(n-1)/730$. Như vậy,

$$\begin{aligned}\mathbb{P}\{\text{không có hai người nào có cùng ngày sinh}\} &= \mathbb{P}\{\text{không thành công}\} \\ &\approx \exp\{-n(n-1)/730\}\end{aligned}$$

Số nguyên nhỏ nhất để xấp xỉ trên nhỏ hơn $1/2$ là giá trị n sao cho

$$n(n-1) \geq 730 \log(2) = 505.997$$

Giải ra ta được kết quả chính xác $n = 23$.

Giả sử bây giờ ta lại muốn biết n cần lớn bao nhiêu để xác suất không có ba người nào trong n người có cùng ngày sinh là nhỏ hơn $1/2$. Trong khi việc tính toán để có được xác suất chính xác giờ đây là một bài toán tổ hợp phức tạp, thì đơn giản hơn là ta đi tìm một xấp xỉ tốt cho nó. Đầu tiên ta hình dung ứng với mỗi bộ ba trong số C_n^3 bộ ba i, j, k là một phép thử, với phép thử i, j, k được xem là một thành công nếu cả ba người i, j, k có cùng ngày sinh. Giả thiết số thành công là xấp xỉ một biến ngẫu nhiên Poisson với trung bình bằng

$$C_n^3 \mathbb{P}\{\text{cả ba người } i, j, k \text{ có cùng ngày sinh}\} = \frac{C_n^3}{365^2}$$

Dễ kiểm tra lại xác suất gần đúng của biến cố không có thành công nào sẽ nhỏ hơn $1/2$ khi $n \geq 84$. Tuy nhiên, ta có thể cải tiến xấp xỉ Poisson này để cho kết quả tốt hơn dựa trên sự phụ thuộc yếu giữa các phép thử. Chẳng hạn, xác suất có điều kiện để phép thử $1, 2, 4$ là một thành công, với điều kiện phép thử $1, 2, 3$ thành công, bằng $1/365$.

Một cách tiếp cận hay hơn là ta giả sử ứng với mỗi ngày trong số 365 ngày của năm là một phép thử, với phép thử i được gọi là thành công nếu có ba người trở lên trong số n người có ngày sinh là ngày i . Vì số người có ngày sinh nhật vào ngày i là một biến ngẫu nhiên nhị thức với các tham số $n, p = 1/365$, nên p_i , xác suất để phép thử i thành công, là bằng

$$p_i = 1 - \sum_{j=0}^2 C_n^j (1/365)^j (364/365)^{n-j}$$

Vì ta có thể thấy một cách trực giác rằng 365 phép thử này là phụ thuộc yếu, nên xác suất để không có thành công nào có thể được xấp xỉ bởi $\mathbb{P}_n = \exp\{-365p_i\}$. Một tính toán sau đó cho thấy

$$\mathbb{P}_{84} = 0.54, \mathbb{P}_{87} = 0.502, \mathbb{P}_{88} = 0.495$$

Vì thế việc chỉ ra rằng với 88 người, xác suất để có ít nhất ba người có cùng ngày sinh là xấp xỉ 0.5. (Thực tế, tính toán chính xác chỉ ra được 88 quả thực là số người ít nhất mà ta cần.)

□

Ví dụ. Giả sử có m quả bóng được đặt vào n cái bình, với mỗi quả bóng được đặt vào bình i một cách độc lập với xác suất p_i , $\sum_{i=1}^n p_i = 1$. Nếu ta gọi X_i là biến ngẫu nhiên chỉ tiêu cho biến cố bình i là trống không, nghĩa là X_i bằng 1 nếu bình i trống không và bằng 0 nếu ngược lại, thì số bình trống không, gọi là X , có thể được biểu diễn bởi

$$X = \sum_{i=1}^n X_i$$

Nếu m đủ lớn để mỗi xác suất

$$\mathbb{P}\{X_i = 1\} = (1 - p_i)^m$$

là nhỏ, thì ta có thể xem được rằng X có phân phối xấp xỉ phân phối Poisson với trung bình $\lambda = \sum_{i=1}^n (1 - p_i)^m$. Nghĩa là,

$$\mathbb{P}\{X = i\} \approx e^{-\lambda} \lambda^i / i!, \quad i \geq 0$$

Bảng sau so sánh các xấp xỉ trên so với các xác suất chính xác khi $n = 4, m = 20$, và $p_i = i/10$.

i	$\mathbb{P}\{X = i\}$	Giá trị xấp xỉ
0	0.8669	0.8746
1	0.1323	0.1171
2	0.0008	0.0079

3.2.2 Tổng Poisson của các biến ngẫu nhiên

Cho $X_1, X_2, \dots$ là dãy các biến ngẫu nhiên độc lập, cùng phân phối và có hàm phân phối F , và giả sử rằng dãy biến ngẫu nhiên này độc lập với biến ngẫu nhiên N , là biến ngẫu nhiên Poisson có trung bình λ . Khi đó, biến ngẫu nhiên

$$S = \sum_{i=1}^N X_i$$

được gọi là một biến ngẫu nhiên **tổng Poisson** với tham số Poisson λ và hàm phân phối thành phần F .

Hàm moment tổng quát của biến ngẫu nhiên tổng Poisson S được xác định bằng việc lấy điều kiện trên N .

$$\begin{aligned}\mathbb{E}[e^{tS}] &= \sum_{n=0}^{\infty} \mathbb{E}[e^{tS} \mid N = n] \mathbb{P}\{N = n\} \\ &= \sum_{n=0}^{\infty} \mathbb{E}[e^{t(X_1 + \dots + X_n)} \mid N = n] e^{-\lambda} \lambda^n / n! \\ &= \sum_{n=0}^{\infty} \mathbb{E}[e^{t(X_1 + \dots + X_n)}] e^{-\lambda} \lambda^n / n! \quad (3.3)\end{aligned}$$

$$\begin{aligned}&= \sum_{n=0}^{\infty} (\phi_X(t))^n e^{-\lambda} \lambda^n / n! \\ &= \exp\{\lambda(\phi_X(t) - 1)\} \quad (3.4)\end{aligned}$$

trong đó

$$\phi_X(t) = \mathbb{E}[e^{tX_i}]$$

và phương trình (3.3) được suy ra từ tính chất độc lập của N với dãy biến ngẫu nhiên $X_i, i \geq 1$.

Áp dụng phương trình (3.4), ta được

$$\begin{aligned}\mathbb{E}[S] &= \lambda \mathbb{E}[X] \\ \text{Var}(S) &= \lambda \mathbb{E}[X^2]\end{aligned}$$

trong đó X có phân phối F .

Xét trường hợp đặc biệt khi F là một phân phối rời rạc, khi đó ta có thể biểu diễn lại biến ngẫu nhiên S như là một biểu diễn tuyến tính của các biến ngẫu nhiên Poisson. Trước tiên, ta xét Mệnh đề sau:

Mệnh đề 3.2.1. *Giả sử N , số các phép thử sẽ xảy ra, là một biến ngẫu nhiên Poisson với trung bình λ , và giả sử mỗi phép thử thứ $j, j = 1, \dots, k$, xảy ra độc lập với xác suất $p_j, \sum_{j=1}^k p_j = 1$. Nếu kí hiệu $N_j, j = 1, \dots, k$, là số lần phép thử thứ j xảy ra, thì các biến ngẫu nhiên $N_j, j = 1, \dots, k$, là các biến ngẫu nhiên Poisson với các trung bình tương ứng là $\lambda p_j, j = 1, \dots, k$.*

Chứng minh.

Cho các số nguyên không âm bất kỳ $n_1, n_2, \dots, n_k$, và đặt $n = \sum_{j=1}^k n_j$. Khi đó, vì $N = \sum_{j=1}^k N_j$, nên ta có

$$\mathbb{P}\{N_j = n_j, j = 1, \dots, k\} = \mathbb{P}\{N_j = n_j, j = 1, \dots, k \mid N = n\} \mathbb{P}\{N = n\}$$

Giả sử rằng, có n phép thử xảy ra, vì mỗi phép thử thứ j là độc lập với xác suất $p_j, j = 1, \dots, k$, nên suy ra $N_1, \dots, N_k$ có một phân phối nhiều chiều với các tham số n và $p_1, \dots, p_k$. Vì vậy,

$$\mathbb{P}\{N_j = n_j, j = 1, \dots, k \mid N = n\} = \frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k}$$

nên suy ra rằng

$$\begin{aligned} \mathbb{P}\{N_j = n_j, j = 1, \dots, k\} &= \frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k} e^{-\lambda} \lambda^n / n! \\ &= \prod_{j=1}^k e^{-\lambda p_j} (\lambda p_j)^{n_j} / n_j! \end{aligned}$$

Suy ra điều phải chứng minh. □

Tiếp theo, ta xét $S = \sum_{i=1}^N X_i$ là một biến ngẫu nhiên tổng Poisson và X_i là các biến ngẫu nhiên rời rạc thỏa mãn

$$\mathbb{P}\{X_i = j\} = p_j, \sum_{j=1}^k p_j = 1$$

Nếu kí hiệu N_j là số lần X_i bằng j , khi đó ta có thể biểu diễn lại biến ngẫu nhiên tổng Poisson S như sau

$$S = \sum_{j=1}^k j N_j$$

trong đó, theo Mệnh đề (3.2.1) thì các biến ngẫu nhiên $N_j, j = 1, \dots, k$, là các biến ngẫu nhiên Poisson độc lập với trung bình tương ứng là $\lambda p_j, j = 1, \dots, k$. Ta có thể tính toán trực tiếp kỳ vọng và phương sai của biến ngẫu nhiên tổng Poisson S như sau:

$$\mathbb{E}[S] = \sum_{j=1}^k j \mathbb{E}[N_j] = \sum_{j=1}^k j \lambda p_j = \lambda \mathbb{E}[X]$$

và

$$Var(S) = \sum_{j=1}^k j^2 Var(N_j) = \sum_{j=1}^k j^2 \lambda p_j = \lambda \mathbb{E}[X^2]$$

3.3 Đánh giá độ chính xác khi xấp xỉ tổng các biến ngẫu nhiên Bernoulli độc lập bởi một biến ngẫu nhiên Poisson

Trong phần này ta sẽ xem xét việc xấp xỉ một tổng các biến ngẫu nhiên Bernoulli độc lập bởi một biến ngẫu nhiên Poisson là chính xác cỡ nào. Giả sử ta muốn xấp xỉ một

tổng các biến ngẫu nhiên Bernoulli độc lập với các trung bình tương ứng $p_1, p_2, \dots, p_n$. Đầu tiên, bắt đầu với một dãy $Y_1, Y_2, \dots, Y_n$ các biến ngẫu nhiên Poisson độc lập, với Y_i có trung bình p_i , ta sẽ xây dựng một dãy các biến ngẫu nhiên Bernoulli độc lập $X_1, X_2, \dots, X_n$ với các tham số $p_1, p_2, \dots, p_n$ sao cho

$$\mathbb{P}\{X_i \neq Y_i\} \leq p_i^2, \quad \text{với mỗi } i$$

Đặt $X = \sum_{i=1}^n X_i$ và $Y = \sum_{i=1}^n Y_i$, ta sẽ dùng điều kiện trên để kết luận được rằng

$$\mathbb{P}\{X \neq Y\} \leq \sum_{i=1}^n p_i^2$$

Cuối cùng, ta sẽ chỉ ra rằng từ bất đẳng thức trên suy ra được, với một tập các số thực A bất kỳ,

$$|\mathbb{P}\{X \in A\} - \mathbb{P}\{Y \in A\}| \leq \sum_{i=1}^n p_i^2$$

Và vì X là tổng của các biến ngẫu nhiên Bernoulli độc lập và Y là một biến ngẫu nhiên Poisson, nên từ bất đẳng thức trên sẽ cho ta cận mà ta mong muốn.

Để làm rõ quá trình trên, gọi $Y_i, i = 1, \dots, n$ là các biến ngẫu nhiên Poisson độc lập với các trung bình tương ứng p_i . Ký hiệu, $U_1, U_2, \dots, U_n$ là các biến ngẫu nhiên độc lập, đồng thời độc lập với Y_i , và thỏa mãn

$$U_i = \begin{cases} 0 & \text{với xác suất } (1 - p_i)e^{p_i} \\ 1 & \text{với xác suất } 1 - (1 - p_i)e^{p_i} \end{cases}$$

Định nghĩa trên cho ta ứng dụng của bất đẳng thức

$$e^{-p} \geq 1 - p$$

với giả thiết $(1 - p_i)e^{p_i} \leq 1$.

Bây giờ ta định nghĩa các biến ngẫu nhiên $X_i, i = 1, \dots, n$ bởi

$$X_i = \begin{cases} 0 & \text{nếu } Y_i = U_i = 0 \\ 1 & \text{nếu ngược lại} \end{cases}$$

Để ý rằng

$$\mathbb{P}\{X_i = 0\} = \mathbb{P}\{Y_i = 0\}\mathbb{P}\{U_i = 0\} = e^{-p_i}(1 - p_i)e^{p_i} = 1 - p_i$$

$$\mathbb{P}\{X_i = 1\} = 1 - \mathbb{P}\{X_i = 0\} = p_i$$

Rõ ràng, nếu $X_i = 0$ thì Y_i cũng phải bằng 0 (do định nghĩa của X_i). Do đó, ta thấy

$$\begin{aligned}
 \mathbb{P}\{X_i \neq Y_i\} &= \mathbb{P}\{X_i = 1, Y_i \neq 1\} \\
 &= \mathbb{P}\{Y_i = 0, X_i = 1\} + \mathbb{P}\{Y_i > 1\} \\
 &= \mathbb{P}\{Y_i = 0, U_i = 1\} + \mathbb{P}\{Y_i > 1\} \\
 &= e^{-p_i}[1 - (1 - p_i)e^{p_i}] + 1 - e^{-p_i} - p_i e^{-p_i} \\
 &= p_i - p_i e^{-p_i} \\
 &\leq p_i^2 \quad (\text{vì } 1 - e^{-p} \leq p)
 \end{aligned}$$

Đặt $X = \sum_{i=1}^n X_i$ và $Y = \sum_{i=1}^n Y_i$ và chú ý rằng X là tổng các biến ngẫu nhiên Bernoulli độc lập và Y là biến ngẫu nhiên Poisson với giá trị kỳ vọng $\mathbb{E}[Y] = \mathbb{E}[X] = \sum_{i=1}^n p_i$. Cũng dễ ý rằng, từ $X \neq Y$ suy ra $X_i \neq Y_i$ với một vài giá trị i nào đó, vì thế

$$\begin{aligned}
 \mathbb{P}\{X \neq Y\} &\leq \mathbb{P}\{X_i \neq Y_i, \text{ với một vài giá trị } i \text{ nào đó}\} \\
 &\leq \sum_{i=1}^n \mathbb{P}\{X_i \neq Y_i\} \quad (\text{do bất đẳng thức Boole}) \\
 &\leq \sum_{i=1}^n p_i^2
 \end{aligned}$$

Với biến cố B bất kỳ, ký hiệu I_B là biến chỉ tiêu của biến cố B và được định nghĩa bởi

$$I_B = \begin{cases} 1 & \text{nếu } B \text{ xảy ra} \\ 0 & \text{nếu ngược lại} \end{cases}$$

Chú ý rằng, với một tập số thực A bất kỳ,

$$I_{\{X \in A\}} - I_{\{Y \in A\}} \leq I_{\{X \neq Y\}}$$

Ta có điều này là do một biến chỉ tiêu chỉ có thể nhận giá trị 0 hoặc 1, nên vế trái bằng 1 chỉ khi $I_{\{X \in A\}} = 1$ và $I_{\{Y \in A\}} = 0$, suy ra $X \in A$ và $Y \notin A$, nghĩa là $X \neq Y$, vì vậy vế phải cũng bằng 1. Lấy kỳ vọng hai vế bất đẳng thức trên, ta được

$$\mathbb{P}\{X \in A\} - \mathbb{P}\{Y \in A\} \leq \mathbb{P}\{X \neq Y\}$$

Thay đổi vai trò của X và Y với cách làm tương tự, ta được

$$\mathbb{P}\{Y \in A\} - \mathbb{P}\{X \in A\} \leq \mathbb{P}\{X \neq Y\}$$

Vì vậy, ta có thể kết luận |

$$|\mathbb{P}\{Y \in A\} - \mathbb{P}\{X \in A\}| \leq \mathbb{P}\{X \neq Y\}$$

Như vậy, ta có thể chứng minh được, với $\lambda = \sum_{i=1}^n p_i$,

$$\left| \mathbb{P} \left\{ \sum_{i=1}^n X_i \in A \right\} - \sum_{i \in A} \frac{e^{-\lambda} \lambda^i}{i!} \right| \leq \sum_{i=1}^n p_i^2$$

Chú ý. Khi tất cả p_i đều bằng p , thì X trở thành biến ngẫu nhiên nhị thức. Do đó, từ chứng minh trên ta có thể suy ra được rằng với một tập các số nguyên không âm A bất kỳ,

$$\left| \sum_{i \in A} C_n^i n^i (1-p)^{n-i} - \sum_{i \in A} \frac{e^{-np} (np)^i}{i!} \right| \leq np^2$$

□

Kết luận

Luận văn tiến hành nghiên cứu vấn đề rất cơ bản của lý thuyết xác suất, bên cạnh đó luận văn cũng đưa ra được một số tính chất đặc biệt của biến ngẫu nhiên có phân phối mũ, mối quan hệ đặc biệt giữa biến ngẫu nhiên nhị thức và biến ngẫu nhiên Poisson, cũng như một số mô hình xác suất thường được áp dụng trong Tin học, đây là vấn đề mang tính thực tiễn cao. Ngoài ra, luận văn cũng giới thiệu một số kết quả mới liên quan đến định lý giới hạn trung tâm của khối đa diện Gauss và các kết quả liên quan đến độ chính xác khi xấp xỉ một tổng biến ngẫu nhiên Bernoulli độc lập bằng một biến ngẫu nhiên Poisson.

Tuy nhiên, luận văn mới chỉ dừng lại ở việc nghiên cứu các kiến thức cơ bản và đưa ra một số kết quả về xác suất thường được áp dụng trong thực tế mà chưa đi sâu vào nghiên cứu lĩnh vực này. Vì vậy, tác giả kính mong nhận được sự đóng góp ý kiến của các thầy cô, bạn bè đồng nghiệp và những người quan tâm đến các lĩnh vực này để tác giả hoàn thiện hơn về mặt kiến thức, cũng như hoàn chỉnh bản luận văn để có thể ứng dụng trong thực tiễn.

Tài liệu tham khảo

- [1] Nguyễn Duy Tiến, Vũ Viết Yên, *Lý thuyết xác suất*, Nhà xuất bản Giáo dục, 2000.
- [2] Nguyễn Duy Tiến, *Giải tích ngẫu nhiên, phần 3*, Nhà xuất bản đại học Quốc gia Hà Nội, 2001.
- [3] Văn Vũ, Imre Bárány, *Central limit theorems for Gaussian polytopes*, Institute of Mathematical Statistics, 2007.
- [4] AD Barbour, Louis H.Y Chen, *Stein's Method and Applications (Vol 4, 5)*, World Scientific, 2005.
- [5] Douglas C. Montgomery, George C. Runger, *Applied Statistics and Probability for Engineers*, John Wiley, 2003.
- [6] M. KELBERT, Y. SUHOV, *Probability and Statistics by Example, Vol 1*, Cambridge University Press, 2005.
- [7] Sheldon M. Ross, *Probability models for computer science*, San Francisco New York Boston, 2002.
- [8] T.T Soong, *Fundamentals of Probability and Statistics for engineers*, John Wiley, 2004.