Các phân phối xác suất thường gặp

Hoàng Văn Hà
University of Science, VNU - HCM
hvha@hcmus.edu.vn

Nội dung

1. Các phân phối rời rạc
   - Phân phối nhị thức
   - Phân phối Poisson
   - Phân phối hình học
2. Các phân phối liên tục
   - Phân phối đều
   - Phân phối mũ
   - Phân phối chuẩn
3. Định lý giới hạn trung tâm
   - Sự hội tụ theo phân phối
   - Định lý giới hạn trung tâm
   - Ứng dụng của định lý giới hạn trung tâm

Nội dung

1. Các phân phối rời rạc
   - Phân phối nhị thức
   - Phân phối Poisson
   - Phân phối hình học
2. Các phân phối liên tục
   - Phân phối đều
   - Phân phối mũ
   - Phân phối chuẩn
3. Định lý giới hạn trung tâm
   - Sự hội tụ theo phân phối
   - Định lý giới hạn trung tâm
   - Ứng dụng của định lý giới hạn trung tâm

Nội dung

Biến ngẫu nhiên Bernoulli

Định nghĩa 1 (Biến ngẫu nhiên Bernoulli)

Thực hiện một phép thử, ta quan tâm đến biến cố A. Nếu biến cố A xảy ra (thành công) thì X nhận giá trị là 1 ($X = 1$)*, ngược lại biến ngẫu nhiên X nhận giá trị* 0*. Phép thử này gọi là phép thử Bernoulli. Giả sử xác suất xảy ra biến cố A là p,* $0 < p < 1$

$$\mathbb{P}(A) = \mathbb{P}(X = 1) = p$$

và

$$\mathbb{P}(\bar{A}) = \mathbb{P}(X = 0) = 1 - p = q.$$

Khi đó biến ngẫu nhiên X được gọi là biến ngẫu nhiên có phân phối Bernoulli với tham số p, ký hiệu $X \sim B(1, p)$.

Phân phối Bernoulli

Ví dụ 1

Các phép thử sau đây cho kết quả là một biến ngẫu nhiên Bernoulli

- Tung ngẫu nhiên một đồng xu: $X = 1$ nếu xuất hiện mặt sấp, $X = 0$ nếu xuất hiện mặt ngửa.

Phân phối Bernoulli

Ví dụ 1

Các phép thử sau đây cho kết quả là một biến ngẫu nhiên Bernoulli

- Tung ngẫu nhiên một đồng xu: $X = 1$ nếu xuất hiện mặt sấp, $X = 0$ nếu xuất hiện mặt ngửa.
- Kiểm tra ngẫu nhiên một sản phẩm trong lô hàng: $X = 1$ nếu gặp được sản phẩm tốt, $X = 0$ nếu gặp được sản phẩm kém.

Phân phối Bernoulli

Ví dụ 1

Các phép thử sau đây cho kết quả là một biến ngẫu nhiên Bernoulli

- Tung ngẫu nhiên một đồng xu: $X = 1$ nếu xuất hiện mặt sấp, $X = 0$ nếu xuất hiện mặt ngửa.
- Kiểm tra ngẫu nhiên một sản phẩm trong lô hàng: $X = 1$ nếu gặp được sản phẩm tốt, $X = 0$ nếu gặp được sản phẩm kém.
- Trả lời ngẫu nhiên 1 câu trắc nghiệm: $X = 0$ nếu trả lời đúng, $X = 1$ nếu trả lời sai.

Phân phối Bernoulli

Ví dụ 1

Các phép thử sau đây cho kết quả là một biến ngẫu nhiên Bernoulli

- Tung ngẫu nhiên một đồng xu: $X = 1$ nếu xuất hiện mặt sấp, $X = 0$ nếu xuất hiện mặt ngửa.
- Kiểm tra ngẫu nhiên một sản phẩm trong lô hàng: $X = 1$ nếu gặp được sản phẩm tốt, $X = 0$ nếu gặp được sản phẩm kém.
- Trả lời ngẫu nhiên 1 câu trắc nghiệm: $X = 0$ nếu trả lời đúng, $X = 1$ nếu trả lời sai.
- Mua vé số: $X = 0$ nếu trúng số, $X = 1$ nếu không trúng số.

Phân phối Bernoulli

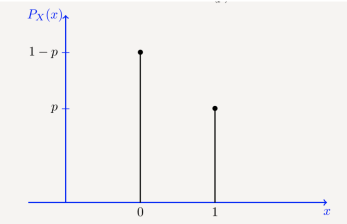
Bảng phân phối xác suất của biến ngẫu nhiên $X \sim B(1, p)$ có dạng

X	1	0
$\mathbb{P}$	p	$1-p$

Phân phối Bernoulli

Bảng phân phối xác suất của biến ngẫu nhiên $X \sim B(1, p)$ có dạng

X	1	0
$\mathbb{P}$	p	$1-p$

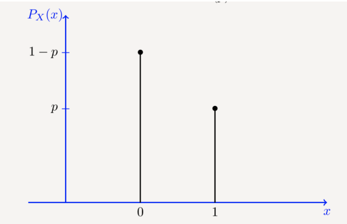


Phân phối xác suất của biến ngẫu nhiên Bernoulli

Phân phối Bernoulli

Bảng phân phối xác suất của biến ngẫu nhiên $X \sim B(1, p)$ có dạng

X	1	0
$\mathbb{P}$	p	$1-p$



Phân phối xác suất của biến ngẫu nhiên Bernoulli

Dựa vào bảng phân phối xác suất của biến ngẫu nhiên X ta dễ dàng tính được

$$\mathbb{E}(X) = p,$$
$$\mathbb{V}ar(X) = p(1-p).$$

Phân phối nhị thức

Định nghĩa 2 (Binomial distribution)

Thực hiện n phép thử Bernoulli độc lập với xác suất thành công trong mỗi phép thử là p. Gọi X là số lần thành công (biến cố A xảy ra) trong n phép thử thì

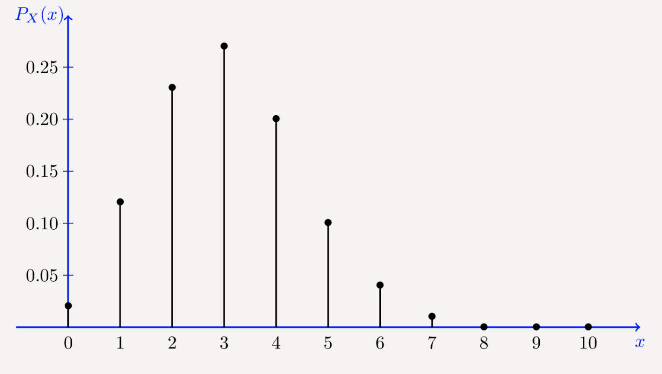
$$X = X_1 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

với $X_i \sim B(1, p)$, $i = 1, \ldots, n$. *Khi đó X là biến ngẫu nhiên rời rạc nhận các giá trị trong tập* $S = \{0, \ldots, n\}$ *và xác suất*

$$p_X(k) = \mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k}, \quad k \in S. \tag{1}$$

Biến ngẫu nhiên X được định nghĩa như trên gọi là biến ngẫu nhiên nhị thức với các tham số n và p, ký hiệu $X \sim B(n, p)$.

Phân phối nhị thức



Phân phối xác suất của $X \sim B(10, 0.3)$

Phân phối nhị thức

Ví dụ 2

Trong một nhà máy sản xuất vi mạch điện tử, biết rằng tỷ lệ vi mạch không đạt chất lượng là 5%. Kiểm tra ngẫu nhiên 15 vi mạch. Tính xác suất

(a) *Có đúng 7 vi mạch không đạt chất lượng.*

(b) *Có ít nhất 1 vi mạch không đạt chất lượng.*

Phân phối nhị thức

Ví dụ 2

Trong một nhà máy sản xuất vi mạch điện tử, biết rằng tỷ lệ vi mạch không đạt chất lượng là 5%. Kiểm tra ngẫu nhiên 15 vi mạch. Tính xác suất

(a) *Có đúng 7 vi mạch không đạt chất lượng.*

(b) *Có ít nhất 1 vi mạch không đạt chất lượng.*

Ví dụ 3

Giả sử màu mắt đen của một người được quy định bởi bởi một cặp gen A và a, trong đó A là gen trội và a là gen lặn. Kiểu hình AA gọi là trội thuần chủng, aa là lặn thuần chủng, và Aa là kiểu hình lai. Người có kiểu hình AA hay Aa sẽ có màu mắt giống nhau (màu đen). Những đứa con sẽ nhận mỗi một gen từ bố mẹ. Giả sử một cặp bố mẹ có màu mắt đen với kiểu hình lai Aa có 4 con, thì xác suất 3 trong 4 đứa trẻ có màu mắt giống bố mẹ là bao nhiêu?

Phân phối nhị thức

Định lý 1 (Các đặc trưng của BNN có phân phối nhị thức)

Nếu X là biến ngẫu nhiên có phân phối nhị thức $B(n, p)$ thì

i) $\mathbb{E}(X) = np$.

ii) $\mathbb{V}ar(X) = np(1-p)$.

Định lý 2

Nếu $X \sim B(n, p)$ và $Y \sim B(m, p)$ là hai biến ngẫu nhiên độc lập. Đặt $Z = X + Y$, ta có $Z \sim B(n+m, p)$.

Phân phối nhị thức

Ví dụ 4

Một học sinh làm một bài thi trắc nghiệm có 60 câu hỏi, mỗi câu có 5 đáp án và chỉ có 1 đáp án đúng. Biết rằng học sinh không học bài và đánh ngẫu nhiên toàn bộ bài thi. Tính xác suất:

(a) Học sinh làm đúng ít nhất 1 câu.

(b) Học sinh làm đúng 30 câu.

(b) Số câu trả lời đúng trung bình mà học sinh làm được là bao nhiêu? Tính độ lệch chuẩn của số câu trả lời đúng.

Phân phối nhị thức

Ví dụ 5

Trong một nhà máy, hàng đóng thành kiện, mỗi kiện 10 sản phẩm, trong đó có 3 phế phẩm. Khi kiện hàng được giao cho khách hàng, khách hàng sẽ lấy ngẫu nhiên ra 2 sản phẩm trong kiện để kiểm tra. Nếu cả hai sản phẩm đều tốt, kiện hàng sẽ được nhận, ngược lại kiện hàng sẽ bị trả lại. Gọi X là số kiện hàng được nhận trong số 50 kiện hàng giao cho khách hàng.

(a) Tính xác suất có 40 kiện hàng được nhận.

(b) Tính $\mathbb{E}(X)$ và $\mathbb{V}ar(X)$.

Phân phối nhị thức

Ví dụ 6

Trên mỗi chuyến bay, có thể có những hành khách bỏ chuyến bay mặc dù đã đặt vé trước. Một hãng hàng không bán ra 125 vé cho một chuyến bay chỉ có 120 ghế. Xác suất một hành khác vắng mặt là 0.10 và độc lập với các hành khách khác.

(a) Tính xác suất tất cả những hành khác có mặt tại sân bay đều có thể thực hiện chuyến bay (có đủ ghế).

(b) Xác suất máy bay cất cánh mà còn dư ghế ngồi là bao nhiêu?

Nội dung

Phân phối Poisson

Ví dụ 7

Giả sử ta cần truyền đi n bit qua một kênh truyền tín hiệu số. Gọi X là số bit lỗi trong n bit được truyền. Khi xác suất một bit bị lỗi là hằng số và việc truyền tín hiệu độc lập, X có phân phối nhị thức. Gọi p là xác suất một bit truyền đi bị lỗi. Đặt $\lambda = np$ thì $\mathbb{E}(X) = np = \lambda$ và

$$\mathbb{P}(X = x) = C_n^x p^x (1-p)^{n-x} = C_n^x \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Phân phối Poisson

Ví dụ 7

Giả sử ta cần truyền đi n bit qua một kênh truyền tín hiệu số. Gọi X là số bit lỗi trong n bit được truyền. Khi xác suất một bit bị lỗi là hằng số và việc truyền tín hiệu độc lập, X có phân phối nhị thức. Gọi p là xác suất một bit truyền đi bị lỗi. Đặt $\lambda = np$ thì $\mathbb{E}(X) = np = \lambda$ và

$$\mathbb{P}(X = x) = C_n^x p^x (1-p)^{n-x} = C_n^x \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Giả sử số bit truyền đi tăng lên và xác suất một bit lỗi giảm xuống sao cho np không đổi. Nghĩa là, n và p đồng thời tăng và giảm sao cho $\mathbb{E}(X) = \lambda$ là hằng số.

Phân phối Poisson

Ví dụ 7

Giả sử ta cần truyền đi n bit qua một kênh truyền tín hiệu số. Gọi X là số bit lỗi trong n bit được truyền. Khi xác suất một bit bị lỗi là hằng số và việc truyền tín hiệu độc lập, X có phân phối nhị thức. Gọi p là xác suất một bit truyền đi bị lỗi. Đặt $\lambda = np$ thì $\mathbb{E}(X) = np = \lambda$ và

$$\mathbb{P}\left(X = x\right) = C_n^x p^x (1-p)^{n-x} = C_n^x \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Giả sử số bit truyền đi tăng lên và xác suất một bit lỗi giảm xuống sao cho np không đổi. Nghĩa là, n và p đồng thời tăng và giảm sao cho $\mathbb{E}(X) = \lambda$ là hằng số. Ta có thể chỉ ra

$$\lim_{n\to\infty} \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots$$

Định nghĩa

Định nghĩa 3 (Poisson distribution)

Biến ngẫu nhiên rời rạc X nhận các giá trị từ $0, 1, 2, \ldots$ gọi là có phân phối Poisson với tham số λ, ký hiệu $X \sim P(\lambda)$ nếu

$$p_X(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2 \ldots \tag{2}$$

Định nghĩa

Định nghĩa 3 (Poisson distribution)

Biến ngẫu nhiên rời rạc X nhận các giá trị từ $0, 1, 2, \ldots$ gọi là có phân phối Poisson với tham số λ, ký hiệu $X \sim P(\lambda)$ nếu

$$p_X(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2 \ldots \tag{2}$$

Kỳ vọng và phương sai của X lần lượt bằng

$$\mathbb{E}(X) = \lambda, \ \ \mathbb{V}ar(X) = \lambda.$$

Định nghĩa

Một số biến ngẫu nhiên mô tả các sự kiện sau thường được xem là tuân theo phân phối Poisson:

- số lỗi in trong một (hoặc một số) trang sách
- số lỗ rò rỉ trên mỗi 100m đường ống nước
- số người sống lâu trên 100 tuổi trong một cộng đồng dân cư
- số khách hàng chờ ở một quầy dịch vụ của ngân hàng trong vòng 30 phút
- số tai nạn hoặc sự cố giao thông xảy ra tại một điểm giao thông trong một ngày
- ...

Định nghĩa

Một số biến ngẫu nhiên mô tả các sự kiện sau thường được xem là tuân theo phân phối Poisson:

- số lỗi in trong một (hoặc một số) trang sách
- số lỗ rò rỉ trên mỗi 100m đường ống nước
- số người sống lâu trên 100 tuổi trong một cộng đồng dân cư
- số khách hàng chờ ở một quầy dịch vụ của ngân hàng trong vòng 30 phút
- số tai nạn hoặc sự cố giao thông xảy ra tại một điểm giao thông trong một ngày
- . . .

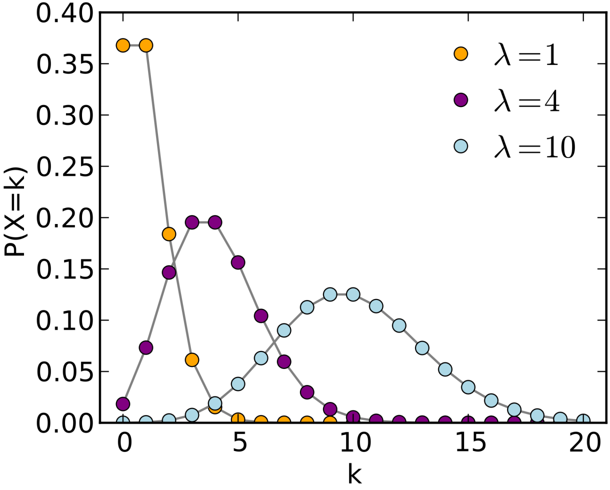
Các biến ngẫu nhiên được sử dụng để mô tả, "đếm" số lần xảy ra của một biến cố, sự kiện nào đó xảy ra trong một khoảng thời gian (hoặc không gian) và thỏa một số điều kiện (các điều kiện này thường thỏa mãn trong thực tế) thường được mô tả bằng phân phối Poisson.

Định nghĩa

Xét t là một khoảng thời gian cố định. Giả sử rằng số biến cố trung bình xảy ra trong một đơn vị thời gian bằng r. Đặt $\lambda = rt$ và $X =$ số biến cố xảy ra trong khoảng thời gian t. Khi đó ta có

$$\mathbb{P}(X = k) = \frac{(rt)^k e^{-rt}}{k!} = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, 2, \ldots$$

Phân phối xác suất



Ví dụ

Ví dụ 8

Giả sử số lỗi in trong một trang của một quyển sách có phân phối Poisson với tham số $\lambda = \frac{1}{2}$. Tính xác suất có ít nhất một lỗi in trong trang này.

Ví dụ 9

Giả sử số người đến rút tiền tại một máy ATM tuân theo phân phối Poisson. Trung bình có 10 người đến rút tiền tại máy trong một giờ. Đặt $X =$ số khách hàng đến rút tiền tại máy ATM này từ 10g00 đến 11g30. Tính $\mathbb{P}(5 < X \leq 10)$.

Ví dụ

Ví dụ 10

Số cuộc điện thoại gọi đến một tổng đài điện thoại trong một giờ có phân phối Poisson với $\lambda = 10$. Tính xác suất

(a) Có 5 cuộc điện thoại gọi đến trong một giờ.

(b) Có nhiều nhất 3 cuộc điện thoại gọi đến trong một giờ.

(c) Có 15 cuộc điện thoại gọi đến trong hai giờ.

(d) Có 5 cuộc điện thoại gọi đến trong 30 phút.

Xấp xỉ pp nhị thức bằng pp Poisson

Định lý 3

Cho $X \sim B(n,p)$, *nếu* $n \to \infty$ *và* $p \to 0$ *sao cho* $np \to \lambda$ *thì*

$$\mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Trong thực tế, phân phối Poisson sẽ xấp xỉ tốt cho phân phối nhị thức khi $n \geq 100$ và $np \leq 10$.

Xấp xỉ pp nhị thức bằng pp Poisson

Định lý 3

Cho $X \sim B(n,p)$, *nếu* $n \to \infty$ *và* $p \to 0$ *sao cho* $np \to \lambda$ *thì*

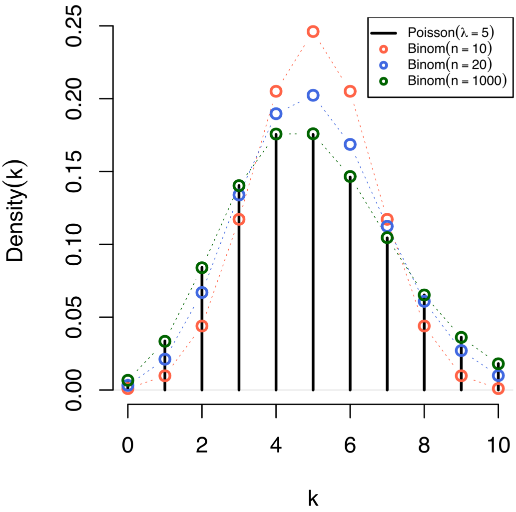
$$\mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Trong thực tế, phân phối Poisson sẽ xấp xỉ tốt cho phân phối nhị thức khi $n \geq 100$ và $np \leq 10$.

Ví dụ 11

Trong một đợt tiêm chủng cho trẻ em ở một khu vực, biết xác suất một trẻ bị phản ứng với thuốc sau khi tiêm là 0,001. Thực hiện tiêm cho 2000 trẻ, tính xác suất có nhiều nhất 1 trẻ bị phản ứng với thuốc sau khi tiêm.

Xấp xỉ pp nhị thức bằng pp Poisson



Nội dung

1. Các phân phối rời rạc
   - Phân phối nhị thức
   - Phân phối Poisson
   - Phân phối hình học
2. Các phân phối liên tục
   - Phân phối đều
   - Phân phối mũ
   - Phân phối chuẩn
3. Định lý giới hạn trung tâm
   - Sự hội tụ theo phân phối
   - Định lý giới hạn trung tâm
   - Ứng dụng của định lý giới hạn trung tâm

Phân phối hình học (Geometric distribution)

- Xét một thí nghiệm chỉ có hai khả năng: *thành công* với xác suất $p(0 < p < 1)$ và *thất bại* với xác suất $1 - p$. Đặt $X =$ số lần thí nghiệm cho đến khi gặp được lần *thành công* đầu tiên. Các lần thí nghiệm được thực hiện độc lập với nhau.
- Ví dụ:
  - tung một đồng xu cho đến khi gặp được mặt hình
  - trong một lô hàng có rất nhiều sản phẩm, số sản phẩm cần chọn cho đến khi gặp được sản phẩm hỏng đầu tiên
  - số lần thử cho đến khi ném được bóng vào rổ của một người chơi bóng rổ (giả sử xác suất ném trúng rổ của người này ở mỗi lần ném không đổi)
  - ...
- Một biến ngẫu nhiên X được định nghĩa như trên gọi là biến ngẫu nhiên có phân phối hình học (geometric distribution).

Phân phối hình học

Định nghĩa 4

Một biến ngẫu nhiên X được gọi là có phân phối hình học với tham số p, ký hiệu $X \sim G(p)$ (hoặc $X \sim \mathrm{Geometric(p)}$), nếu phân phối xác suất của nó cho bởi

$$p_X(k) = p(1-p)^{k-1}, \quad k = 1, 2, 3, \ldots \tag{3}$$

với $0 < p < 1$. Kỳ vọng và phương sai của X được cho bởi

$$\mathbb{E}(X) = \frac{1}{p}, \quad \mathbb{V}ar(X) = \frac{1-p}{p^2}.$$

Phân phối hình học

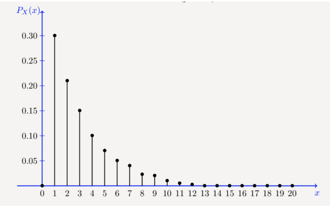
Định nghĩa 4

Một biến ngẫu nhiên X được gọi là có phân phối hình học với tham số p, ký hiệu $X \sim G(p)$ (hoặc $X \sim \mathrm{Geometric(p)}$), nếu phân phối xác suất của nó cho bởi

$$p_X(k) = p(1-p)^{k-1}, \quad k = 1, 2, 3, \ldots \tag{3}$$

với $0 < p < 1$. Kỳ vọng và phương sai của X được cho bởi

$$\mathbb{E}(X) = \frac{1}{p}, \quad \mathbb{V}ar(X) = \frac{1-p}{p^2}.$$



Phân phối xác suất của $X \sim G(0.3)$

Ví dụ

Ví dụ 12

Biết rằng xác suất một vận động viên bắn súng bắn trúng tâm ở mỗi lần bắn bằng 0.6.

a) Tính xác suất để vận động viên bắn trúng tâm ở lần bắn thứ 3.

b) Số lần bắn trung bình bằng bao nhiêu để bắn trúng tâm.

Ví dụ 13 (St. Petersburg Paradox)

Một người tung một đồng xu cho đến khi xuất hiện mặt hình ở lần đầu tiên. Nếu mặt hình xuất hiện ở lần tung thứ n, người này sẽ thắng 2^n\$. Gọi X là số tiền người này thắng được.

a) *Chứng tỏ rằng* $\mathbb{E}(X) = +\infty$.

b) *Bạn có bằng lòng trả 1 triệu \$ để tham gia trò chơi này?*

c) *Bọn có bằng lòng trả 1 triệu \$ cho mỗi lần chơi nếu bạn được quyền chơi lâu cho đến khi bạn muốn dừng trò chơi?*

Một số phân phối rời rạc khác

- Phân phối siêu bội (Hypergeometric distribution)
- Phân phối nhị thức âm (Negative Binomial distribution)
- ...

Nội dung

Nội dung

Phân phối đều

Định nghĩa 5 (Uniform distribution)

Biến ngẫu nhiên liên tục X được gọi là có phân phối đều trên đoạn $[a; b]$, ký hiệu $X \sim U([a; b])$, nếu hàm mật độ xác suất của X có dạng

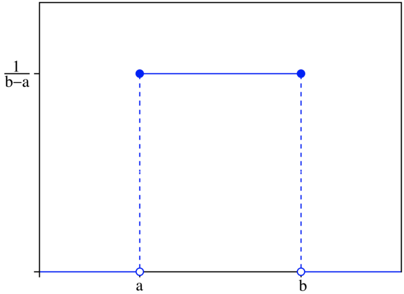
$$f(x) = \begin{cases} \dfrac{1}{b-a} & \textit{khi } x \in [a, b] \\ 0 & \textit{nơi khác} \end{cases}$$

Phân phối đều

Định nghĩa 5 (Uniform distribution)

Biến ngẫu nhiên liên tục X được gọi là có phân phối đều trên đoạn $[a; b]$, ký hiệu $X \sim U([a; b])$, nếu hàm mật độ xác suất của X có dạng

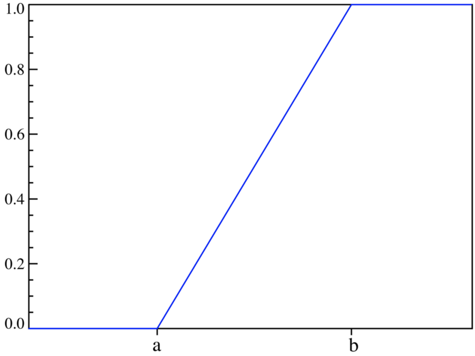
$$f(x) = \begin{cases} \dfrac{1}{b-a} & \textit{khi } x \in [a, b] \\ 0 & \textit{nơi khác} \end{cases}$$



Phân phối đều - Hàm phân phối

Hàm phân phối xác suất của $X \sim U([a;b])$ cho bởi

$$F(x) = \begin{cases} 0 & \text{khi } x < a \\ \dfrac{x-a}{b-a} & \text{khi } x \in [a,b] \\ 1 & \text{khi } x > b \end{cases}$$



Kỳ vọng và phương sai của pp đều

Định lý 4 (Các đặc trưng của biến ngẫu nhiên có phân phối đều)

Cho X là biến ngẫu nhiên có phân phối đều trên $[a, b]$ $(X \sim U([a; b]))$ thì

i) *Kỳ vọng* $\mathbb{E}(X) = \dfrac{a+b}{2}$.

ii) *Phương sai* $\mathbb{V}ar(X) = \dfrac{(b-a)^2}{12}$.

Phân phối đều

Ví dụ 14

Những chuyến xe buýt số 8 đến một trạm trước ký túc xá khu A mỗi 15 phút kể từ 7 giờ sáng, nghĩa là xe buýt sẽ ghé qua trạm lúc 7:00, 7:15, 7:30, ... Nếu thời gian đến trạm của một sinh viên có phân phối đều trong khoảng từ 7:00 đến 7:30, hãy tính xác suất sinh viên này phải chờ

(a) ít hơn 5 phút để bắt một chuyến bus.

(b) nhiều hơn 10 phút để bắt một chuyến bus.

Nội dung

Phân phối mũ

Định nghĩa 6 (Exponential distribution)

Biến ngẫu nhiên T gọi là có phân phối mũ, ký hiệu $T \sim Exp(\lambda)$, $\lambda > 0$, nếu nó có hàm mật độ xác suất

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{nếu } t \geq 0 \\ 0 & \text{nếu } t < 0 \end{cases} \tag{4}$$

trong đó

- *λ: số biến cố trung bình xảy ra trong một đơn vị thời gian*
- *t: số đơn vị thời gian cho đến biến cố kế tiếp*

Phân phối mũ

Định nghĩa 6 (Exponential distribution)

Biến ngẫu nhiên T gọi là có phân phối mũ, ký hiệu $T \sim Exp(\lambda)$, $\lambda > 0$, nếu nó có hàm mật độ xác suất

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{nếu } t \geq 0 \\ 0 & \text{nếu } t < 0 \end{cases} \tag{4}$$

trong đó

- *λ: số biến cố trung bình xảy ra trong một đơn vị thời gian*
- *t: số đơn vị thời gian cho đến biến cố kế tiếp*

- Phân phối mũ thường được dùng để mô tả phân phối của khoảng thời gian cho đến khi một biến cố cụ thể nào đó xảy ra. Ví du: khoảng thời gian chờ cho đến khi một xe buýt đến trạm, khoảng thời gian cho đến khi một thiết bị điện tử bị hỏng, khoảng thời gian một khách hàng chờ đến lượt phục vụ tại một quầy dịch vụ ở ngân hàng, v.v.

Các đặc trưng của phân phối mũ

Hàm phân phối của T:

$$F(t) = \mathbb{P}(T \leq t) = 1 - e^{-\lambda t}, \quad t \geq 0. \tag{5}$$

Các đặc trưng của phân phối mũ

Hàm phân phối của T:

$$F(t) = \mathbb{P}(T \leq t) = 1 - e^{-\lambda t}, \quad t \geq 0. \tag{5}$$

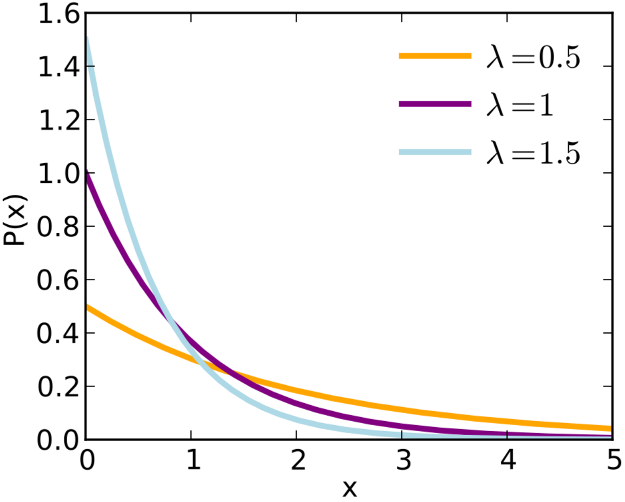
Định lý 5

Nếu $T \sim Exp(\lambda)$ thì kỳ vọng và phương sai của T lần lượt bằng

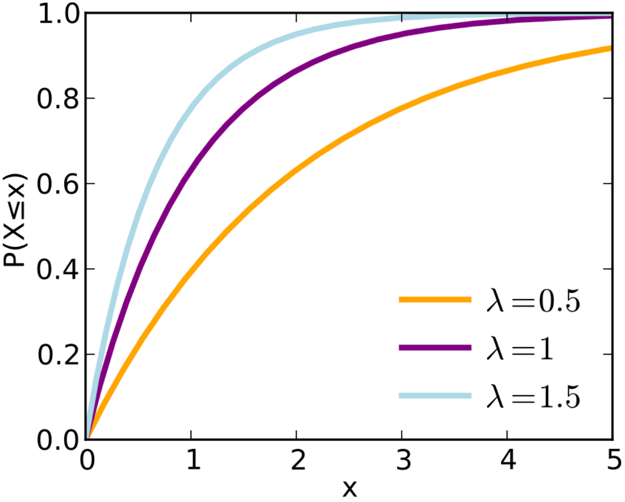
$$\mathbb{E}(T) = \frac{1}{\lambda},$$

$$\mathbb{V}ar(T) = \frac{1}{\lambda^2}.$$

Phân phối mũ - Hàm mật độ



Phân phối mũ - Hàm phân phối



Phân phối mũ

Ví dụ 15

Trong một mạng máy tính ở một công ty, biết rằng số người dùng đăng nhập vào mạng trong một giờ có phân phối Poisson với trung bình bằng 25.

(a) Tính xác suất không có người dùng nào đăng nhập trong khoảng thời gian 6 phút.

(b) Tính xác suất lần đăng nhập kế tiếp cách lần đăng nhập đầu từ 2 đến 3 phút.

Phân phối mũ

Ví dụ 16

Số khách hàng đến làm thủ tục tại một quầy dịch vụ ở ngân hàng với tỷ lệ là 15 người một giờ. Hỏi xác suất thời gian giữa 2 khách hàng liên tiếp đến quầy dịch vụ ít hơn 3 phút là bao nhiêu?

Ví dụ 17

Trong một nhà máy sản xuất linh kiện điện tử, biết tuổi thọ của một mạch điện là biến ngẫu nhiên có phân phối mũ với tuổi thọ trung bình 6.25 năm. Nếu thời gian bảo hành của sản phẩm là 5 năm. Hỏi tỷ lệ sản phẩm bảo hành của nhà máy là bao nhiêu?

Nội dung

Phân phối chuẩn

Định nghĩa 7 (Normal distribution)

Biến ngẫu nhiên liên tục X nhận giá trị trong khoảng $(-\infty, +\infty)$ được gọi là có phân phối chuẩn tham số μ, σ nếu hàm mật độ xác suất có dạng

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad -\infty < x < +\infty \tag{6}$$

trong đó μ, σ là hằng số và $\sigma > 0$, $-\infty < \mu < +\infty$, ký hiệu $X \sim \mathcal{N}\left(\mu; \sigma^2\right)$.

Phân phối chuẩn

Định nghĩa 7 (Normal distribution)

Biến ngẫu nhiên liên tục X nhận giá trị trong khoảng $(-\infty, +\infty)$ được gọi là có phân phối chuẩn tham số μ, σ nếu hàm mật độ xác suất có dạng

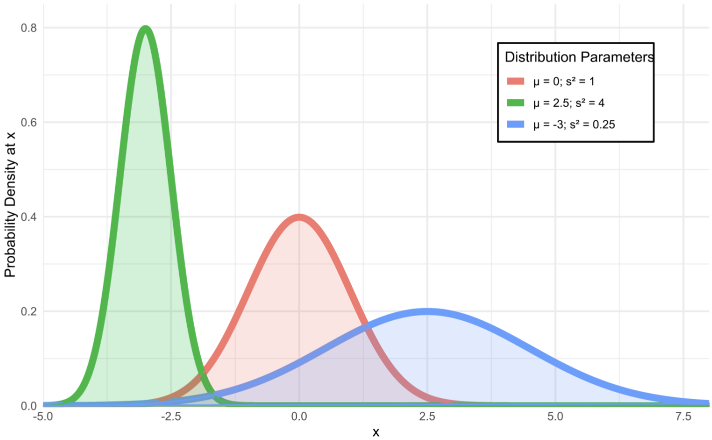
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad -\infty < x < +\infty \tag{6}$$

trong đó μ, σ là hằng số và $\sigma > 0$, $-\infty < \mu < +\infty$, ký hiệu $X \sim \mathcal{N}\left(\mu; \sigma^2\right)$.

Nếu $X \sim \mathcal{N}(\mu, \sigma^2)$, ta có

$$\mathbb{E}(X) = \mu,$$
$$\mathbb{V}ar(X) = \sigma^2.$$

Phân phối chuẩn



Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:
  - Đồ thị có dạng chuông (bell shape)

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:
  - Đồ thị có dạng chuông (bell shape)
  - Phân phối đối xứng

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:
  - Đồ thị có dạng chuông (bell shape)
  - Phân phối đối xứng
  - Trung bình = trung vị (median) = yếu vị (mode)

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:
  - Đồ thị có dạng chuông (bell shape)
  - Phân phối đối xứng
  - Trung bình = trung vị (median) = yếu vị (mode)
  - Vị trí của phân phối được xác định bởi kỳ vọng μ

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:
  - Đồ thị có dạng chuông (bell shape)
  - Phân phối đối xứng
  - Trung bình = trung vị (median) = yếu vị (mode)
  - Vị trí của phân phối được xác định bởi kỳ vọng μ
  - Độ phân tán được xác định bởi độ lệch tiêu chuẩn σ

Phân phối chuẩn - Tính chất

- Phân phối chuẩn là một trong những phân phối quan trọng nhất, được dùng để mô tả phân phối của nhiều biến ngẫu nhiên trong thực tế, như chiều cao/cân nặng của một người, tổng doanh thu của một công ty, điểm thi của sinh viên, sai số của một phép đo, v.v. Bên cạnh đó, định lý giới hạn trung tâm (central limit theorem) đã chứng tỏ rằng, phân phối chuẩn là phân phối xấp xỉ của nhiều phân phối khác như nhị thức, tổng các biến ngẫu nhiên độc lập, v.v.
- Một số tính chất của phân phối chuẩn:
  - Đồ thị có dạng chuông (bell shape)
  - Phân phối đối xứng
  - Trung bình = trung vị (median) = yếu vị (mode)
  - Vị trí của phân phối được xác định bởi kỳ vọng μ
  - Độ phân tán được xác định bởi độ lệch tiêu chuẩn σ
  - Xác định trên $\mathbb{R}$

Ví dụ 18

Đường kính của một chi tiết máy do một máy tiện sản xuất có phân phối chuẩn với kỳ vọng 20mm, phương sai $(0.2mm)^2$. Tính xác suất lấy ngẫu nhiên một chi tiết

a) có đường kính trong khoảng 19.9mm đến 20.3mm.

b) có đường kính sai khác với kỳ vọng không quá 0.3mm.

Phân phối chuẩn tắc

Định nghĩa 8 (Standard normal distribution)

Biến ngẫu nhiên Z được gọi là có phân phối chuẩn tắc nếu nó có phân phối chuẩn với tham số $\mu = 0$ và $\sigma^2 = 1$, ký hiệu $Z \sim \mathcal{N}(0, 1)$.

Phân phối chuẩn tắc

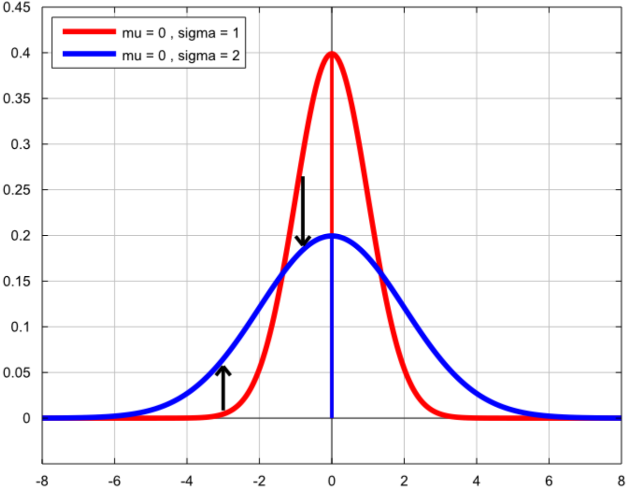
Định nghĩa 8 (Standard normal distribution)

Biến ngẫu nhiên Z được gọi là có phân phối chuẩn tắc nếu nó có phân phối chuẩn với tham số $\mu = 0$ và $\sigma^2 = 1$, ký hiệu $Z \sim \mathcal{N}(0, 1)$.

Theo quy ước, hàm phân phối của biến ngẫu nhiên chuẩn hóa được ký hiệu là $\Phi(z)$, tức

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx$$

Phân phối chuẩn tắc



Phân phối chuẩn tắc

Theo định lý về tính tuyến tính của phân phối chuẩn, nếu $X \sim \mathcal{N}\left(\mu;\sigma^2\right)$ thì $\dfrac{X-\mu}{\sigma}$ có phân phối chuẩn hóa hay

$$\frac{X-\mu}{\sigma} \sim N\left(0,1\right)$$

Phân phối chuẩn tắc

Theo định lý về tính tuyến tính của phân phối chuẩn, nếu $X \sim \mathcal{N}\left(\mu; \sigma^2\right)$ thì $\dfrac{X-\mu}{\sigma}$ có phân phối chuẩn hóa hay

$$\frac{X-\mu}{\sigma} \sim N(0,1)$$

Dựa vào tính chất này ta có thể tính xác suất của biến ngẫu nhiên $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$.

$$\mathbb{P}(X \le b) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

Phân phối chuẩn tắc

Theo định lý về tính tuyến tính của phân phối chuẩn, nếu $X \sim \mathcal{N}\left(\mu;\sigma^2\right)$ thì $\dfrac{X-\mu}{\sigma}$ có phân phối chuẩn hóa hay

$$\frac{X-\mu}{\sigma} \sim N(0,1)$$

Dựa vào tính chất này ta có thể tính xác suất của biến ngẫu nhiên $X \sim \mathcal{N}\left(\mu,\sigma^2\right)$.

$$\mathbb{P}(X \le b) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

Tương tự, với $a \le b$ thì

$$\mathbb{P}(a < X \le b) = \mathbb{P}(X \le b) - \mathbb{P}(X \le a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Ví dụ 19

Cho $X \sim \mathcal{N}(10, 4)$, tính các xác suất sau

(a) $\mathbb{P}(X < 13)$

(b) $\mathbb{P}(X > 9)$

(c) $\mathbb{P}(6 < X < 14)$

(d) $\mathbb{P}(2 < X < 4)$

(e) $\mathbb{P}(-2 < X < 8)$

Phân phối chuẩn

Định lý 6 (Tính "tuyến tính" của phân phối chuẩn)

Nếu biến ngẫu nhiên X có phân phối chuẩn với kỳ vọng μ, phương sai σ^2. Đặt $Y = aX + b$, (a, b là hằng số và $a \neq 0$), thì Y có phân phối chuẩn với kỳ vọng $a\mu + b$ và phương sai $a^2\sigma^2$.

Phân phối chuẩn

Định lý 6 (Tính "tuyến tính" của phân phối chuẩn)

Nếu biến ngẫu nhiên X có phân phối chuẩn với kỳ vọng μ, phương sai σ^2. Đặt $Y = aX + b$, (a, b là hằng số và $a \neq 0$), thì Y có phân phối chuẩn với kỳ vọng $a\mu + b$ và phương sai $a^2\sigma^2$.

Định lý 7

Nếu các biến ngẫu nhiên $X_1, \ldots, X_n$ là độc lập và nếu X_i có phân phối chuẩn với kỳ vọng μ_i và phương sai σ_i^2, $(i = 1, 2, \ldots, n)$, thì tổng $X_1 + \cdots + X_n$ có phân phối chuẩn với kỳ vọng là $\mu_1 + \cdots + \mu_n$ và phương sai là $\sigma_1^2 + \cdots + \sigma_n^2$.

Phân phối chuẩn

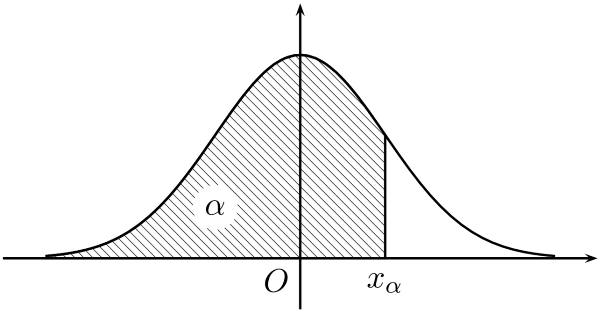
Hệ quả 8

Nếu các biến ngẫu nhiên $X_1, \dots, X_n$ là độc lập và X_i có phân phối chuẩn với kỳ vọng μ_i và phương sai σ_i^2, $(i = 1, \dots, n)$. $a_i, \dots, a_n$ và b là các hằng số sao cho có ít nhất một $a_i \neq 0$, thì biến ngẫu nhiên $a_1X_1 + \cdots + a_nX_n + b$ có phân phối chuẩn với kỳ vọng $a_1\mu_1 + \cdots + a_n\mu_n$ và phương sai $a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$.

Phân vị chuẩn

Định nghĩa 9 (Phân vị chuẩn (normal quartile))

Cho biến ngẫu nhiên $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, *phân vị mức* α, *ký hiệu* x_α, *là giá trị của biến ngẫu nhiên* X *thỏa mãn điều kiện* $\mathbb{P}\left(X \leq x_\alpha\right) = \alpha$



Ví dụ 20

Cho $X \sim \mathcal{N}(10, 4)$, tìm x sao cho

(a) $\mathbb{P}(X > x) = 0.5$

(b) $\mathbb{P}(X > x) = 0.95$

(c) $\mathbb{P}(x < X < 10) = 0.2$

(d) $\mathbb{P}(-x < X - 10 < x) = 0.95$

(e) $\mathbb{P}(-x < X - 10 < x) = 0.99$

Một số phân phối liên tục khác

- Phân phối Gamma
- Phân phối Cauchy
- Phân phối Beta

Một số phân phối liên tục thường gặp trong thống kê:

- Phân phối Chi bình phương (Chi-square distribution)
- Phân phối Student t
- Phân phối Fisher

Nội dung

Giới thiệu

- Trong phần này, chúng ta sẽ khảo sát về Định lý giới hạn trung tâm (Central Limit Theorem - CLT). Định lý giới hạn trung tâm là một trong những kết quả quan trọng nhất của lý thuyết xác suất, nó phát biểu rằng, dưới một số điều kiện nhất định, tổng của một số lớn các biến ngẫu nhiên có phân phối xấp xỉ theo phân phối chuẩn.

Giới thiệu

- Trong phần này, chúng ta sẽ khảo sát về Định lý giới hạn trung tâm (Central Limit Theorem - CLT). Định lý giới hạn trung tâm là một trong những kết quả quan trọng nhất của lý thuyết xác suất, nó phát biểu rằng, dưới một số điều kiện nhất định, tổng của một số lớn các biến ngẫu nhiên có phân phối xấp xỉ theo phân phối chuẩn.
- Trước tiên, ta sẽ tìm hiểu về sự hội tụ theo phân phối (convergence in distribution) của một dãy các biến ngẫu nhiên, sau đó sẽ phát biểu định lý giới hạn trung tâm và xét các ứng dụng của nó.

Nội dung

Sự hội tụ theo phân phối

Định nghĩa 10

Một dãy các biến ngẫu nhiên $X_1, X_2, X_3, \ldots$ được gọi là hội tụ theo phân phối (in distribution) về một biến ngẫu nhiên X, kí hiệu là

$$X_n \xrightarrow{d} X,$$

nếu

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x),$$

với mọi x mà tại đó $F_X(x)$ liên tục.

Sự hội tụ theo phân phối

Ví dụ 21

Xét $X_2, X_3, X_4, \ldots$ *là một dãy các biến ngẫu nhiên có hàm phân phối*

$$F_{X_n}(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Chứng tỏ rằng X_n *hội tụ theo phân phối về một biến ngẫu nhiên có phân phối* $\mathrm{Exp}(1)$.

Sự hội tụ theo phân phối

Ví dụ 21

Xét $X_2, X_3, X_4, \ldots$ *là một dãy các biến ngẫu nhiên có hàm phân phối*

$$F_{X_n}(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Chứng tỏ rằng X_n *hôi tụ theo phân phối về một biến ngẫu nhiên có phân phối* Exp(1).

Giải ví dụ 1:

- Xét $X \sim \text{Exp}(1)$. Ta có với $x \leq 0$,

$$F_{X_n}(x) = F_X(x) = 0, \quad \text{với } n = 2, 3, 4, \ldots$$

Sự hội tụ theo phân phối

Ví dụ 21

Xét $X_2, X_3, X_4, \ldots$ *là một dãy các biến ngẫu nhiên có hàm phân phối*

$$F_{X_n}(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Chứng tỏ rằng X_n *hôi tụ theo phân phối về một biến ngẫu nhiên có phân phối* $\mathrm{Exp}(1)$.

Giải ví dụ 1:

- Xét $X \sim \mathrm{Exp}(1)$. Ta có với $x \leq 0$,
$$F_{X_n}(x) = F_X(x) = 0, \quad \text{với } n = 2, 3, 4, \ldots$$
- Với $x > 0$, ta có
$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty} \left(1 - \left(1 - \frac{1}{n}\right)^{nx}\right) = 1 - \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^{nx}$$
$$= 1 - e^{-x} = F_X(x), \quad \text{với mọi } x.$$

Sự hội tụ theo phân phối

Ví dụ 21

Xét $X_2, X_3, X_4, \dots$ là một dãy các biến ngẫu nhiên có hàm phân phối

$$F_{X_n}(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Chứng tỏ rằng X_n hôi tụ theo phân phối về một biến ngẫu nhiên có phân phối $\mathrm{Exp}(1)$.

Giải ví dụ 1:

- Xét $X \sim \mathrm{Exp}(1)$. Ta có với $x \le 0$,
$$F_{X_n}(x) = F_X(x) = 0, \quad \text{với } n = 2, 3, 4, \dots$$
- Với $x > 0$, ta có
$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty}\left(1 - \left(1 - \frac{1}{n}\right)^{nx}\right) = 1 - \lim_{n\to\infty}\left(1 - \frac{1}{n}\right)^{nx}$$
$$= 1 - e^{-x} = F_X(x), \quad \text{với mọi } x.$$
- Vậy, $X_n \xrightarrow{d} X$.

Nội dung

Định lý giới hạn trung tâm

Định lý 9 (Central Limit Theorem)

Xét $X_1, \ldots, X_n$ là một dãy các biến ngẫu nhiên độc lập và cùng phân phối với kỳ vọng $\mathbb{E}[X_i] = \mu < \infty$ và phương sai $\mathbb{V}ar(X_i) = \sigma^2 < \infty$. Đặt

$$S_n = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i.$$

và

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

thì ta có $Z_n \xrightarrow{d} \mathcal{N}(0,1)$, tức là

$$\lim_{n\to\infty} \mathbb{P}(Z_n \leq x) = \Phi(x), \quad \forall x \in \mathbb{R},$$

với $\Phi(x)$ là hàm phân phối xác suất của phân phối chuẩn tắc $\mathcal{N}(0,1)$.

Định lý giới hạn trung tâm

- Một trong những đặc điểm thú vị của định lý giới hạn trung tâm là ta không cần bất cứ điều kiện gì về phân phối của dãy $X_1, X_2, \ldots, X_n$. Các biến ngẫu nhiên này có thể tuân theo phân phối rời rạc, liên tục hoặc phân phối hỗn hợp (mixed distribution). Nhưng tổng $S_n = \sum_{i=1}^{n} X_i$ đều tuân theo phân phối chuẩn khi n đủ lớn.

Định lý giới hạn trung tâm với phân phối rời rạc

Ví dụ 22 (Minh họa CLT cho phân phối rời rạc)

- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{B}(p)$. *Đặt* $S_n = \sum_{i=1}^{n} X_i$ *thì* $\mathbb{E}[S_n] = np$ *và* $\mathbb{V}ar(S_n) = np(1-p)$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - np}{\sqrt{np(1-p)}},$$

khi n tăng dần.

Định lý giới hạn trung tâm với phân phối rời rạc

Ví dụ 22 (Minh họa CLT cho phân phối rời rạc)

- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{B}(p)$. *Đặt* $S_n = \sum_{i=1}^n X_i$ *thì* $\mathbb{E}[S_n] = np$ *và* $\mathbb{V}ar(S_n) = np(1-p)$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - np}{\sqrt{np(1-p)}},$$

khi n tăng dần.

- $n = 1$:

$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

Hàm khối xác suất (pmf) của Z_1

Định lý giới hạn trung tâm với phân phối rời rạc

Ví dụ 22 (Minh họa CLT cho phân phối rời rạc)

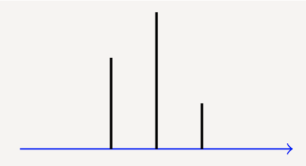
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{B}(p)$. *Đặt* $S_n = \sum_{i=1}^{n} X_i$ *thì* $\mathbb{E}[S_n] = np$ *và* $\mathbb{V}ar(S_n) = np(1-p)$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - np}{\sqrt{np(1-p)}},$$

khi n tăng dần.

- $n = 2$:

$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$



Hàm khối xác suất (pmf) của Z_2

Định lý giới hạn trung tâm với phân phối rời rạc

Ví dụ 22 (Minh họa CLT cho phân phối rời rạc)

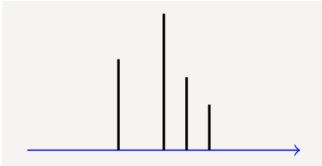
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{B}(p)$. *Đặt* $S_n = \sum_{i=1}^n X_i$ *thì* $\mathbb{E}[S_n] = np$ *và* $\mathbb{V}ar(S_n) = np(1-p)$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - np}{\sqrt{np(1-p)}},$$

khi n tăng dần.

- $n = 3$:

$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$



Hàm khối xác suất (pmf) của Z_3

Định lý giới hạn trung tâm với phân phối rời rạc

Ví dụ 22 (Minh họa CLT cho phân phối rời rạc)

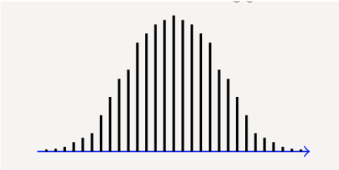
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{B}(p)$. *Đặt* $S_n = \sum_{i=1}^n X_i$ *thì* $\mathbb{E}[S_n] = np$ *và* $\mathbb{V}ar(S_n) = np(1-p)$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - np}{\sqrt{np(1-p)}},$$

khi n tăng dần.

- $n = 30$:

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$



Hàm khối xác suất (pmf) của Z_{30}

Định lý giới hạn trung tâm với phân phối liên tục

Ví dụ 23 (Minh họa CLT cho phân phối liên tục)

- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{U}([0,1])$. *Đặt* $S_n = \sum_{i=1}^{n} X_i$ *thì* $\mathbb{E}[S_n] = n/2$ *và* $\mathbb{V}ar(S_n) = n/12$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}},$$

khi n tăng dần.

Định lý giới hạn trung tâm với phân phối liên tục

Ví dụ 23 (Minh họa CLT cho phân phối liên tục)

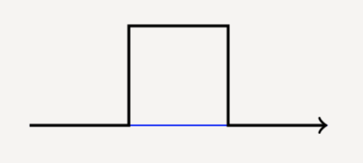
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{U}([0,1])$. *Đặt* $S_n = \sum_{i=1}^{n} X_i$ *thì* $\mathbb{E}[S_n] = n/2$ *và* $\mathbb{V}ar(S_n) = n/12$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}},$$

khi n tăng dần.

- $n = 1$:

$$Z_1 = \frac{X_1 - \frac{1}{2}}{\sqrt{\frac{1}{12}}}$$



Hàm mật độ xác suất (pdf) của Z_1

Định lý giới hạn trung tâm với phân phối liên tục

Ví dụ 23 (Minh họa CLT cho phân phối liên tục)

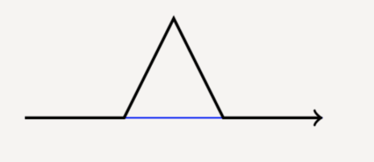
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{U}([0,1])$. *Đặt* $S_n = \sum_{i=1}^{n} X_i$ *thì* $\mathbb{E}[S_n] = n/2$ *và* $\mathbb{V}ar(S_n) = n/12$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}},$$

khi n tăng dần.

- $n = 2$:

$$Z_2 = \frac{X_1 + X_2 - 1}{\sqrt{\frac{2}{12}}}$$



Hàm mật độ xác suất (pdf) của Z_2

Định lý giới hạn trung tâm với phân phối liên tục

Ví dụ 23 (Minh họa CLT cho phân phối liên tục)

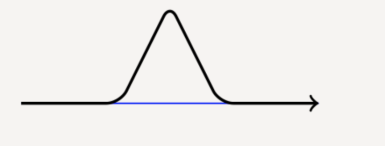
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{U}([0,1])$. *Đặt* $S_n = \sum_{i=1}^{n} X_i$ *thì* $\mathbb{E}[S_n] = n/2$ *và* $\mathbb{V}ar(S_n) = n/12$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}},$$

khi n tăng dần.

- $n = 3$:

$$Z_3 = \frac{X_1 + X_2 + X_3 - \frac{3}{2}}{\sqrt{\frac{3}{12}}}$$



Hàm mật độ xác suất (pdf) của Z_3

Định lý giới hạn trung tâm với phân phối liên tục

Ví dụ 23 (Minh họa CLT cho phân phối liên tục)

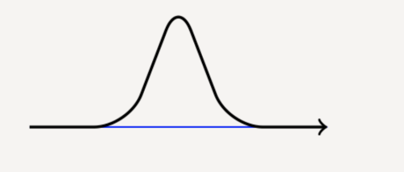
- *Xét* $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathcal{U}([0,1])$. *Đặt* $S_n = \sum_{i=1}^n X_i$ *thì* $\mathbb{E}[S_n] = n/2$ *và* $\mathbb{V}ar(S_n) = n/12$.
- *Ta khảo sát phân phối của*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{V}ar(S_n)}} = \frac{X_1 + \cdots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}},$$

khi n tăng dần.

- $n = 30$:

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - \frac{30}{2}}{\sqrt{\frac{30}{12}}}$$



Hàm mật độ xác suất (pdf) của Z_{30}

Nội dung

1. Các phân phối rời rạc
   - Phân phối nhị thức
   - Phân phối Poisson
   - Phân phối hình học
2. Các phân phối liên tục
   - Phân phối đều
   - Phân phối mũ
   - Phân phối chuẩn
3. Định lý giới hạn trung tâm
   - Sự hội tụ theo phân phối
   - Định lý giới hạn trung tâm
   - Ứng dụng của định lý giới hạn trung tâm

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:
  - Sai số các phép đo trong phòng thí nghiệm thường được giả sử tuân theo phân phối chuẩn.

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:
  - Sai số các phép đo trong phòng thí nghiệm thường được giả sử tuân theo phân phối chuẩn.
  - Trong xử lý tín hiệu, $Y = s(X) + \epsilon$, ta thường giả sử thành phần nhiễu ϵ tuân theo phân phối chuẩn (Gaussian noise).

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:
  - Sai số các phép đo trong phòng thí nghiệm thường được giả sử tuân theo phân phối chuẩn.
  - Trong xử lý tín hiệu, $Y = s(X) + \epsilon$, ta thường giả sử thành phần nhiễu ϵ tuân theo phân phối chuẩn (Gaussian noise).
  - Trong tài chính, phần trăm thay đổi về giá của một loại tài sản nào đó thường được mô hình bởi phân phối chuẩn.

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:
  - Sai số các phép đo trong phòng thí nghiệm thường được giả sử tuân theo phân phối chuẩn.
  - Trong xử lý tín hiệu, $Y = s(X) + \epsilon$, ta thường giả sử thành phần nhiễu ϵ tuân theo phân phối chuẩn (Gaussian noise).
  - Trong tài chính, phần trăm thay đổi về giá của một loại tài sản nào đó thường được mô hình bởi phân phối chuẩn.
- Định lý giới hạn trung tâm cũng rất hữu ích về mặt tính toán. Giả sử trong vấn đề mà ta quan tâm có liên quan đến việc tính tổng của một số lượng rất lớn các biến độc lập có cùng phân phối. Đôi khi việc tính toán trực tiếp sẽ rất khó khăn, trong trường hợp này ta có thể sử dụng định lý giới hạn trung tâm để đơn giản việc tính toán.

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:
  - Sai số các phép đo trong phòng thí nghiệm thường được giả sử tuân theo phân phối chuẩn.
  - Trong xử lý tín hiệu, $Y = s(X) + \epsilon$, ta thường giả sử thành phần nhiễu ϵ tuân theo phân phối chuẩn (Gaussian noise).
  - Trong tài chính, phần trăm thay đổi về giá của một loại tài sản nào đó thường được mô hình bởi phân phối chuẩn.
- Định lý giới hạn trung tâm cũng rất hữu ích về mặt tính toán. Giả sử trong vấn đề mà ta quan tâm có liên quan đến việc tính tổng của một số lượng rất lớn các biến độc lập có cùng phân phối. Đôi khi việc tính toán trực tiếp sẽ rất khó khăn, trong trường hợp này ta có thể sử dụng định lý giới hạn trung tâm để đơn giản việc tính toán.

Ứng dụng của định lý giới hạn trung tâm

- Trong nhiều ứng dụng thực tế, một biến ngẫu nhiên mà ta quan tâm có thể là tổng của nhiều biến ngẫu nhiên độc lập. Trong các trường hợp này, ta có thể sử dụng định lý giới hạn trung tâm để sử dụng phân phối chuẩn. Một số ví dụ:
  - Sai số các phép đo trong phòng thí nghiệm thường được giả sử tuân theo phân phối chuẩn.
  - Trong xử lý tín hiệu, $Y = s(X) + \epsilon$, ta thường giả sử thành phần nhiễu ϵ tuân theo phân phối chuẩn (Gaussian noise).
  - Trong tài chính, phần trăm thay đổi về giá của một loại tài sản nào đó thường được mô hình bởi phân phối chuẩn.
- Định lý giới hạn trung tâm cũng rất hữu ích về mặt tính toán. Giả sử trong vấn đề mà ta quan tâm có liên quan đến việc tính tổng của một số lượng rất lớn các biến độc lập có cùng phân phối. Đôi khi việc tính toán trực tiếp sẽ rất khó khăn, trong trường hợp này ta có thể sử dụng định lý giới hạn trung tâm để đơn giản việc tính toán.

Trong thực hành, thông thường định lý giới hạn trung tâm sẽ thỏa mãn với $n \geq 30$.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

- Xét $X \sim \mathcal{B}(n, p)$ với $\mathbb{E}[X] = np$ và $\mathbb{V}ar(X) = np(1-p)$.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

- Xét $X \sim \mathcal{B}(n, p)$ với $\mathbb{E}[X] = np$ và $\mathbb{V}ar(X) = np(1-p)$.
- Khi n lớn, chẳng hạn $n = 300$ và giả sử ta cần tính $\mathbb{P}(50 \le X \le 250)$. Vì X là phân phối rời rạc, nên việc tính trực tiếp

$$\mathbb{P}(50 \le X \le 250) = \mathbb{P}(X = 50) + \mathbb{P}(X = 51) + \cdots + \mathbb{P}(X = 250)$$

sẽ rất khó khăn. Trong trường hợp này ta áp dụng định lý giới hạn trung tâm để đơn giản việc tính toán.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

- Xét $X \sim \mathcal{B}(n, p)$ với $\mathbb{E}[X] = np$ và $\mathbb{V}ar(X) = np(1-p)$.
- Khi n lớn, chẳng hạn $n = 300$ và giả sử ta cần tính $\mathbb{P}(50 \leq X \leq 250)$. Vì X là phân phối rời rạc, nên việc tính trực tiếp

$$\mathbb{P}(50 \leq X \leq 250) = \mathbb{P}(X = 50) + \mathbb{P}(X = 51) + \cdots + \mathbb{P}(X = 250)$$

  sẽ rất khó khăn. Trong trường hợp này ta áp dụng định lý giới hạn trung tâm để đơn giản việc tính toán.
- Theo ĐL GHTT, nếu ta đặt

$$Z = \frac{X - np}{\sqrt{np(1-p)}}, \text{ thì } Z \overset{\text{approx}}{\sim} \mathcal{N}(0,1) \text{ khi } n \to \infty.$$

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

- Xét $X \sim \mathcal{B}(n, p)$ với $\mathbb{E}[X] = np$ và $\mathbb{V}ar(X) = np(1-p)$.
- Khi n lớn, chẳng hạn $n = 300$ và giả sử ta cần tính $\mathbb{P}(50 \leq X \leq 250)$. Vì X là phân phối rời rạc, nên việc tính trực tiếp

$$\mathbb{P}(50 \leq X \leq 250) = \mathbb{P}(X = 50) + \mathbb{P}(X = 51) + \cdots + \mathbb{P}(X = 250)$$

  sẽ rất khó khăn. Trong trường hợp này ta áp dụng định lý giới hạn trung tâm để đơn giản việc tính toán.
- Theo ĐL GHTT, nếu ta đặt

$$Z = \frac{X - np}{\sqrt{np(1-p)}}, \text{ thì } Z \overset{\text{approx}}{\sim} \mathcal{N}(0,1) \text{ khi } n \to \infty.$$

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

- Xét $X \sim \mathcal{B}(n,p)$ với $\mathbb{E}[X] = np$ và $\mathbb{V}ar(X) = np(1-p)$.
- Khi n lớn, chẳng hạn $n = 300$ và giả sử ta cần tính $\mathbb{P}(50 \leq X \leq 250)$. Vì X là phân phối rời rạc, nên việc tính trực tiếp

$$\mathbb{P}(50 \leq X \leq 250) = \mathbb{P}(X = 50) + \mathbb{P}(X = 51) + \cdots + \mathbb{P}(X = 250)$$

  sẽ rất khó khăn. Trong trường hợp này ta áp dụng định lý giới hạn trung tâm để đơn giản việc tính toán.
- Theo ĐL GHTT, nếu ta đặt

$$Z = \frac{X - np}{\sqrt{np(1-p)}}, \text{ thì } Z \overset{\text{approx}}{\sim} \mathcal{N}(0,1) \text{ khi } n \to \infty.$$

  Vậy,

$$\begin{aligned}\mathbb{P}(a \leq X \leq b) &= \mathbb{P}\left(\frac{a-np}{\sqrt{np(1-p)}} \leq \frac{X-np}{\sqrt{np(1-p)}} \leq \frac{b-np}{\sqrt{np(1-p)}}\right)\\ &= \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right),\end{aligned}$$

  với $\Phi(\cdot)$ là hàm phân phối xác suất tính lũy của $\mathcal{N}(0,1)$.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Ví dụ 24

Trong một cuộc bầu cử thị trưởng ở một thành phố, người ta ước tính được rằng có 40% người dân thành phố ủng hộ ứng cử viên A. Trong một đợt khảo sát, người ta chọn ngẫu nhiên 200 người để hỏi ý kiến về việc bỏ phiếu, hỏi xác suất gặp được từ 90 đến 120 người ủng hộ ứng cử viên A là bao nhiêu?

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Ví dụ 24

Trong một cuộc bầu cử thị trưởng ở một thành phố, người ta ước tính được rằng có 40% người dân thành phố ủng hộ ứng cử viên A. Trong một đợt khảo sát, người ta chọn ngẫu nhiên 200 người để hỏi ý kiến về việc bỏ phiếu, hỏi xác suất gặp được từ 90 đến 120 người ủng hộ ứng cử viên A là bao nhiêu?

Giải ví dụ 24:

- Đặt $X =$ số người đồng ý bỏ phiếu cho ứng viên A tron 200 người được hỏi ý kiến. Thì $X \sim \mathcal{B}(200, 0.4)$.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Ví dụ 24

Trong một cuộc bầu cử thị trưởng ở một thành phố, người ta ước tính được rằng có 40% người dân thành phố ủng hộ ứng cử viên A. Trong một đợt khảo sát, người ta chọn ngẫu nhiên 200 người để hỏi ý kiến về việc bỏ phiếu, hỏi xác suất gặp được từ 90 đến 120 người ủng hộ ứng cử viên A là bao nhiêu?

Giải ví dụ 24:

- Đặt $X =$ số người đồng ý bỏ phiếu cho ứng viên A tron 200 người được hỏi ý kiến. Thì $X \sim \mathcal{B}(200, 0.4)$.
- Ta cần tính $\mathbb{P}(90 \leq X \leq 120)$. Nếu tính toán trực tiếp sử dụng công thức của phân phối nhị thức, ta cần tính

$$\mathbb{P}(90 \leq X \leq 120) = \mathbb{P}(X = 90) + \mathbb{P}(X = 91) + \cdots + \mathbb{P}(X = 120) = \ldots,$$

$\Rightarrow$ khó khăn tính toán. Ta áp dụng ĐL GHTT.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Giải ví dụ 24:

- Đặt $\mu = \mathbb{E}[X] = np = 200 \times 0.4 = 80$ và $\sigma = \sqrt{np(1-p)} = \sqrt{200 \times 0.4 \times 0.6} = 6.93$.

Áp dụng: xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Giải ví dụ 24:

- Đặt $\mu = \mathbb{E}[X] = np = 200 \times 0.4 = 80$ và $\sigma = \sqrt{np(1-p)} = \sqrt{200 \times 0.4 \times 0.6} = 6.93$.
- Áp dụng ĐL GHTT, ta có

$$\begin{aligned}
\mathbb{P}(90 \leq X \leq 120) &= \mathbb{P}\left(\frac{90-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{120-\mu}{\sigma}\right) \\
&= \mathbb{P}\left(\frac{90-80}{6.93} \leq \frac{X-80}{6.93} \leq \frac{120-80}{6.93}\right) \\
&= \Phi(5.77) - \Phi(1.44) \approx 1 - 0.9251 = 0.0749.
\end{aligned}$$

Xấp xỉ với hiệu chỉnh liên tục (continuity correction)

Để tăng độ chính xác khi thực hiện xấp xỉ, ta có thể thực hiện việc hiệu chỉnh liên tục (continuity correction) như sau: trừ đi 0.5 cho chặn dưới và cộng thêm 0.5 cho chặn trên, tức là

- Tính $\mathbb{P}(a - 0.5 < X < b + 0.5)$ thay vì $\mathbb{P}(a < X < b)$.
- Tương tự, tính $\mathbb{P}(X < c + 0.5)$ thay vì $\mathbb{P}(X < c)$.

Xấp xỉ với hiệu chỉnh liên tục (continuity correction)

Để tăng độ chính xác khi thực hiện xấp xỉ, ta có thể thực hiện việc hiệu chỉnh liên tục (continuity correction) như sau: trừ đi 0.5 cho chặn dưới và cộng thêm 0.5 cho chặn trên, tức là

- Tính $\mathbb{P}(a - 0.5 < X < b + 0.5)$ thay vì $\mathbb{P}(a < X < b)$.
- Tương tự, tính $\mathbb{P}(X < c + 0.5)$ thay vì $\mathbb{P}(X < c)$.

Ví dụ 25

Giải lại ví dụ 24, sử dụng hiệu chỉnh liên tục.

Xấp xỉ với hiệu chỉnh liên tục (continuity correction)

Để tăng độ chính xác khi thực hiện xấp xỉ, ta có thể thực hiện việc hiệu chỉnh liên tục (continuity correction) như sau: trừ đi 0.5 cho chặn dưới và cộng thêm 0.5 cho chặn trên, tức là

- Tính $\mathbb{P}(a - 0.5 < X < b + 0.5)$ thay vì $\mathbb{P}(a < X < b)$.
- Tương tự, tính $\mathbb{P}(X < c + 0.5)$ thay vì $\mathbb{P}(X < c)$.

Ví dụ 25

Giải lại ví dụ 24, sử dụng hiệu chỉnh liên tục.

Trong các tính toán sau này (cũng như trong bài thi), ta luôn sử dụng hiệu chỉnh liên tục khi thực hiện xấp xỉ chuẩn.

Ví dụ 26

Trong một kênh truyền tín hiệu số, biết rằng xác suất một bit bị lỗi khi truyền là 1×10^{-5}. Nếu 16 triệu bit được truyền đi, hỏi xác suất có hơn 150 lỗi là bao nhiêu?

Ví dụ 27

Sản phẩm do một nhà máy sản xuất được đóng thành từng kiện. Mỗi kiện 10 sản phẩm, trong đó có 6 sản phẩm tốt và 4 sản phẩm xấu. Khách hàng chọn cách kiểm tra như sau: từ mỗi kiện chọn ngẫu nhiên 3 sản phẩm. Nếu thấy có ít nhất 2 sản phẩm tốt thì nhận kiện đó, ngược lại thì loại kiện đó. Kiểm tra 140 kiện hàng trong số rất nhiều kiện. Tính xác suất để có

(a) *93 kiện được nhận.*

(b) *Từ 90 đến 110 kiện được nhận.*