Thống kê ứng dụng

Đỗ Lân

dolan@tlu.edu.vn
Đại học Thủy Lợi

Ngày 2 tháng 7 năm 2018

Nội dung môn học

1. Tổng quan về Thống kê
2. Thu thập dữ liệu
3. Tóm tắt và trình bày dữ liệu bằng bảng và đồ thị
4. Tóm tắt dữ liệu bằng các đại lượng thống kê mô tả
5. **Xác suất căn bản và biến ngẫu nhiên**
6. Phân phối của tham số mẫu và ước lượng tham số tổng thể
7. Kiểm định giả thuyết về tham số một tổng thể
8. Kiểm định giả thuyết về tham số hai tổng thể
9. Phân tích phương sai
10. Kiểm định phi tham số
11. Kiểm định chi - bình phương

Phần V

Xác suất và biến ngẫu nhiên

Mục lục

1. Xác suất là gì?
   - Những khái niệm cơ bản của xác suất
     - Không gian mẫu, biến cố sơ cấp và biến cố
     - Quan hệ giữa các biến cố
   - Định nghĩa xác suất
     - Khái niệm chung về xác suất
     - Định nghĩa cổ điển về xác suất
     - Định nghĩa thống kê về xác suất
     - Định nghĩa xác suất theo nghĩa chủ quan
   - Một số qui tắc tính xác suất
     - Qui tắc cộng xác suất
     - Qui tắc xác suất có điều kiện
     - Công thức xác suất đầy đủ
2. Biến ngẫu nhiên
   - Phân loại biến ngẫu nhiên
   - Phân phối xác suất của biến ngẫu nhiên rời rạc
   - Phân phối xác suất của biến ngẫu nhiên liên tục
   - Trung bình, phương sai của biến ngẫu nhiên
   - Chữa một số bài tập về xác suất và biến ngẫu nhiên
   - Một số phân phối quan trọng
     - Phân phối nhị thức
     - Phân phối chuẩn

Example

1. Một lớp học cần có ít nhất bao nhiêu sinh viên để có 2 sinh viên có cùng sinh nhật?

Example

1. Một lớp học cần có ít nhất bao nhiêu sinh viên để có 2 sinh viên có cùng sinh nhật? $\longrightarrow$ Cần 367 sinh viên để có chắc chắn 2 sinh viên cùng sinh nhật.

Example

1. Một lớp học cần có ít nhất bao nhiêu sinh viên để có 2 sinh viên có cùng sinh nhật? $\longrightarrow$ Cần 367 sinh viên để có chắc chắn 2 sinh viên cùng sinh nhật.
2. Thực tế, có tới 99% các lớp học có từ 60 sinh viên là có 2 sinh viên cùng sinh nhật.

Example

1. Một lớp học cần có ít nhất bao nhiêu sinh viên để có 2 sinh viên có cùng sinh nhật? $\longrightarrow$ Cần 367 sinh viên để có chắc chắn 2 sinh viên cùng sinh nhật.
2. Thực tế, có tới 99% các lớp học có từ 60 sinh viên là có 2 sinh viên cùng sinh nhật. Tại sao lại như vậy?
   $\longrightarrow$ Kiến thức xác suất sẽ cho ta câu trả lời.

Example

1. Chơi đề khả năng thua hay thắng nhiều hơn? Nếu chơi khoảng 100 lần, mỗi lần mua một số 10K thì lỗ hay lãi khoảng bao nhiêu?

Example

1. Chơi đề khả năng thua hay thắng nhiều hơn? Nếu chơi khoảng 100 lần, mỗi lần mua một số 10K thì lỗ hay lãi khoảng bao nhiêu?
2. Có chiến lược chơi nào có lãi không?

Example

1. Chơi đề khả năng thua hay thắng nhiều hơn? Nếu chơi khoảng 100 lần, mỗi lần mua một số 10K thì lỗ hay lãi khoảng bao nhiêu?
2. Có chiến lược chơi nào có lãi không?

$\longrightarrow$ Nội dung phần biến ngẫu nhiên sẽ giải đáp và cho bạn một chiến lược chơi có lãi.

Khái niệm

Xác suất là một bộ phận của toán học nghiên cứu các hiện tượng ngẫu nhiên để tìm ra quy luật ẩn sau những ngẫu nhiên đó.

Khái niệm

Xác suất là một bộ phận của toán học nghiên cứu các hiện tượng ngẫu nhiên để tìm ra quy luật ẩn sau những ngẫu nhiên đó.
$\longrightarrow$ *Giúp lí giải các hiện tượng hay gặp trong tự nhiên và xã hội.*

Khái niệm

Xác suất là một bộ phận của toán học nghiên cứu các hiện tượng ngẫu nhiên để tìm ra quy luật ẩn sau những ngẫu nhiên đó.
$\longrightarrow$ Giúp lí giải các hiện tượng hay gặp trong tự nhiên và xã hội.
$\longrightarrow$ Giúp vạch ra các thông tin có ích cho việc ra quyết định, vạch ra được các chiến lược có xu hướng có lợi, . . .

1. Xác suất là gì?
   - Những khái niệm cơ bản của xác suất
     - Không gian mẫu, biến cố sơ cấp và biến cố
     - Quan hệ giữa các biến cố
   - Định nghĩa xác suất
     - Khái niệm chung về xác suất
     - Định nghĩa cổ điển về xác suất
     - Định nghĩa thống kê về xác suất
     - Định nghĩa xác suất theo nghĩa chủ quan
   - Một số qui tắc tính xác suất
     - Qui tắc cộng xác suất
     - Qui tắc xác suất có điều kiện
     - Công thức xác suất đầy đủ
2. Biến ngẫu nhiên
   - Phân loại biến ngẫu nhiên
   - Phân phối xác suất của biến ngẫu nhiên rời rạc
   - Phân phối xác suất của biến ngẫu nhiên liên tục
   - Trung bình, phương sai của biến ngẫu nhiên
   - Chữa một số bài tập về xác suất và biến ngẫu nhiên
   - Một số phân phối quan trọng
     - Phân phối nhị thức
     - Phân phối chuẩn

Phép thử

Definition

Phép thử là một quá trình diễn biến có xuất hiện các kết quả. Phép thử được gọi là ngẫu nhiên nếu ta không thể dự báo trước kết quả nào sẽ xảy ra.

Phép thử

Definition

Phép thử là một quá trình diễn biến có xuất hiện các kết quả. Phép thử được gọi là ngẫu nhiên nếu ta không thể dự báo trước kết quả nào sẽ xảy ra.

Example (Hãy lấy 4 ví dụ về phép thử)

Biến cố

Definition

Một kết quả của phép thử được gọi là một **biến cố** và thường được kí hiệu là $A, B, C, \ldots$ Biến cố đôi khi còn được gọi là "sự kiện".

phép thử chọn ngẫu nhiên một bạn trong danh sách sau:

Tuổi	Tên	Giới tính
21	An	Nam
22	Bình	Nữ
21	Hoa	Nữ
23	Lâm	Nam
23	Hổ	Nam

Xét biến cố A : "người được chọn là nam", B : "người được chọn là Hoa".
$\longrightarrow$ A có thể chia thành nhiều biến cố như là: chọn được An, chọn được Lâm, người được chọn là Hổ.
$\longrightarrow$ biến cố B không chia nhỏ hơn được thành biến cố nào nữa. Biến cố như biến cố B ta gọi là biến cố sơ cấp.

Definition

Một biến cố mà không thể chia thành những biến cố nhỏ hơn được gọi là **biến cố sơ cấp**.
Tập hợp tất cả các biến cố sơ cấp được gọi là **không gian mẫu**, kí hiệu là Ω.

Example

Với phép thử gieo một con xúc xắc, kết quả của phép thử có thể là:

1. $A =$ "mặt trên cùng của con xúc xắc là mặt 2 chấm", - đây là biến cố sơ cấp,
2. $B =$ "mặt trên cùng của con xúc xắc có số chấm chẵn". - đây không phải là biến cố sơ cấp

Example

Từ mỗi phép thử nói trên lập bốn biến cố, trong đó, mô tả một biến cố sơ cấp.

Definition

Biến cố không thể là biến cố không bao giờ xảy ra, kí hiệu $\emptyset$. **Biến cố chắc chắn** là biến cố luôn xảy ra.

Definition

Biến cố không thể là biến cố không bao giờ xảy ra, kí hiệu $\emptyset$. **Biến cố chắc chắn** là biến cố luôn xảy ra.

Example

Xét phép thử tung hai con xúc xắc (là loại xúc xắc 6 mặt, số chấm trên các mặt từ 1 đến 6) và quan sát số chấm của mặt trên cùng. Khi đó biến cố: "Tổng số chấm mặt trên cùng của hai xúc xắc là 13" là biến cố không thể; còn biến cố "Tổng số chấm mặt trên cùng của hai xúc xắc lớn hơn 1" là biến cố chắc chắn.

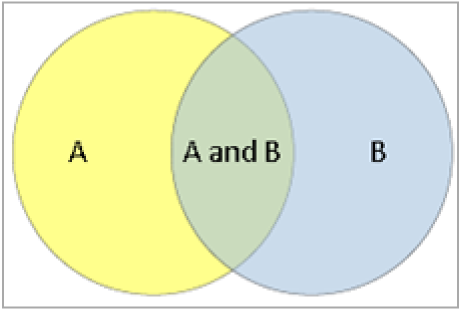
Example

Xét phép thử: chọn ngẫu nhiên một sinh viên lớp này biến cố "sinh viên được chọn trên 16 tuổi" là biến cố chắc chắn. Còn biến cố "sinh viên được chọn trên 140 tuổi" là biến cố không thể.

Example

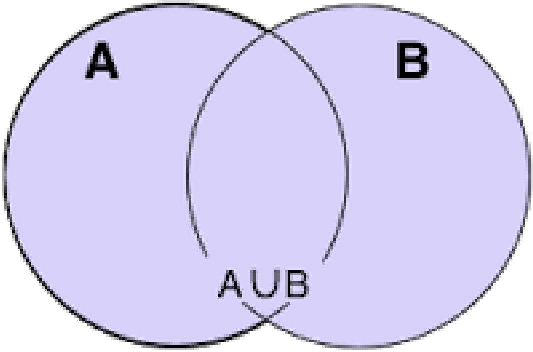
Trong mỗi phép thử đã lập ở trên, hãy nêu biến chắc chắn và biến cố không thể.

Giao của hai biến cố



Giao của hai biến cố A và B được kí hiệu là $A \cap B$ hoặc AB.

Hợp của hai biến cố



Hợp của hai biến cố A, B được kí hiệu là $A \cup B$ hoặc $A + B$.

Biến cố đồng khả năng

Definition

Hai biến cố A và B được gọi là đồng khả năng nếu như chúng cùng khả năng xuất hiện.

Hai biến cố xung khắc

Definition

Hai biến cố A và B được gọi là xung khắc khi và chỉ khi chúng không thể cùng xảy ra trong một phép thử:

$$AB = \emptyset$$

Hai biến cố xung khắc

Definition

Hai biến cố A và B được gọi là xung khắc khi và chỉ khi chúng không thể cùng xảy ra trong một phép thử:

$$AB = \emptyset$$

Example

Hãy mô tả một cặp biến cố xung khắc từ một trong 4 phép thử đã mô tả.

Definition

Hệ các biến cố $A_1, A_2, \ldots, A_n$ được gọi là xung khắc từng đôi nếu như hai biến cố bất kì trong các biến cố trên xung khắc với nhau:

$$A_iA_j = \emptyset, \ \forall i \neq j$$

Xác suất

Dùng để đo khả năng xảy ra của một biến cố, thỏa mãn:

- Xác suất của một biến cố A, kí hiệu $P(A)$, là một con số thuộc đoạn $[0; 1]$.
- Xác suất một biến cố càng gần 1 thì nó càng có khả năng xảy ra, càng gần 0 thì càng ít khả năng xảy ra.
- Biến cố chắc chắn có xác suất là 1, biến cố không thể hay còn gọi là biến cố rỗng có xác suất là 0.
- Hai biến cố A, B xung khắc thì: $P(A \cup B) = P(A) + P(B)$.

Definition

Một biến cố ω được gọi là kết quả thuận lợi cho biến cố A nếu kết quả của phép thử là ω thì biến cố A xảy ra.

Definition

Một biến cố ω được gọi là kết quả thuận lợi cho biến cố A nếu kết quả của phép thử là ω thì biến cố A xảy ra.

Example

Tung một con xúc xắc, ta gọi A là biến cố "Mặt trên cùng có số chấm chẵn", B là biến cố "Mặt trên cùng của xúc xắc là mặt có 2 chấm". Khi đó nếu B xảy ra thì A sẽ xảy ra. Vậy B là kết quả thuân lợi cho biến cố A.

Definition

Trong một phép thử có n kết quả đồng khả năng và xung khắc, trong đó m kết quả có lợi cho biến cố A, khi đó xác suất (theo nghĩa cổ điển) của biến cố A là tỉ số:

$$P(A) = \frac{m}{n} = \frac{|A|}{|\Omega|} = \frac{\text{Số kết quả có lợi cho A}}{\text{Số kết quả đồng khả năng có thể xảy ra}}.$$

Definition

Trong một phép thử có n kết quả đồng khả năng và xung khắc, trong đó m kết quả có lợi cho biến cố A, khi đó xác suất (theo nghĩa cổ điển) của biến cố A là tỉ số:

$$P(A) = \frac{m}{n} = \frac{|A|}{|\Omega|} = \frac{\text{Số kết quả có lợi cho A}}{\text{Số kết quả đồng khả năng có thể xảy ra}}.$$

Example

Tung một đồng xu cân đối 2 hai lần, các kết quả có thể xảy ra là $\Omega = \{SS, SN, NS, NN\}$, các kết quả đã liệt kê này xuất hiện với cùng khả năng vì đồng xu là cân đối. Khi đó biến cố A: "xuất hiện mặt S trong hai lần tung" có 3 kết quả có lợi là SS, SN, NS. Do đó $P(A) = 3/4 = 0.75$

Example

Một lớp có 30 nam, 40 nữ. Gọi ngẫu nhiên một sinh viên lên bảng. Xác suất để sinh viên được chọn là nam là bao nhiêu?

Definition

Tiến hành n phép thử được tiến hành trong những điều kiện giống nhau. Gọi k là số lần biến cố A xuất hiện. Khi số phép thử n đủ lớn, tỉ số $\frac{k}{n}$ sẽ dao động quanh một con số. Ta gọi con số này là xác suất của biến cố A theo nghĩa thống kê.

Definition

Tiến hành n phép thử được tiến hành trong những điều kiện giống nhau. Gọi k là số lần biến cố A xuất hiện. Khi số phép thử n đủ lớn, tỉ số $\frac{k}{n}$ sẽ dao động quanh một con số. Ta gọi con số này là xác suất của biến cố A theo nghĩa thống kê.

Example

Chạy thử nghiệm 200 lần một ứng dụng game, một nhân viên ở bộ phận test thấy có 6 lần game này làm treo máy. Vậy xác xuất làm treo máy của phiên bản đang thử nghiệm là $\approx \dfrac{6}{200}$.

Definition

Theo nghĩa chủ quan, xác suất của một biến cố được đưa ra dựa vào sự nhạy cảm hoặc khả năng phân tích, phán đoán của người xác định xác suất.

Theorem

Cho A và B là hai biến cố của cùng một phép thử. Khi đó ta có

$$P(A+B) = P(A) + P(B) - P(AB).$$

Example

Một khảo sát 2000 sinh viên mới nhập học cho thấy 1500 sinh viên đồng ý nộp tiền qua thẻ BIDV, 1200 bạn đồng ý nộp trực tiếp, trong số đó có 900 sinh viên đồng ý cả việc nộp qua thẻ lẫn nộp trực tiếp (tức là thích thế nào cũng được). Số sinh viên còn lại không thích nộp qua BIDV cũng như không thích nộp trực tiếp (mà muốn qua thẻ ngân hàng khác). Chọn ngẫu nhiên một trong 2000 sinh viên trên. Xác suất để sinh viên này đồng ý nộp qua thẻ BIDV hoặc nộp trực tiếp.

Lưu ý

1. Như đã nói ở định nghĩa: khi A, B là hai biến cố xung khắc ta có:

$$P(A+B) = P(A) + P(B).$$

   Từ tính chất này chứng minh được định lí phía trên.

2. Khi xét $\bar{A}$ và A là hệ xung khắc thì $P(A+\bar{A}) = 1$ nên:

$$P(\bar{A}) = 1 - P(A).$$

3. Một cách mở rộng, nếu các biến cố $A_1, \ldots, A_n$ xung khắc từng đôi thì ta có:

$$P(A_1 + \ldots + A_n) = P(A_1) + \ldots + P(A_n).$$

Example

Theo số liệu thống kê ở một vùng, tỷ lệ dân truy cập internet nhiều chiếm 40%, tỷ lệ dân xem ti vi nhiều chiếm 50%, tỷ lệ dân vừa truy cập internet nhiều vừa xem tivi nhiều chiếm 10%. Chọn ngẫu nhiên một người dân ở vùng trên.

1. Tính xác suất để người này truy cập internet nhiều hoặc xem tivi nhiều.
2. Tính xác suất để người này không truy cập internet nhiều.

Example

Theo số liệu thống kê ở một vùng, tỷ lệ dân truy cập internet nhiều chiếm 40%, tỷ lệ dân xem ti vi nhiều chiếm 50%, tỷ lệ dân vừa truy cập internet nhiều vừa xem tivi nhiều chiếm 10%. Chọn ngẫu nhiên một người dân ở vùng trên.

1. Tính xác suất để người này truy cập internet nhiều hoặc xem tivi nhiều.
2. Tính xác suất để người này không truy cập internet nhiều.

Để tiện cho việc trình bày, ta gọi A là biến cố người được chọn truy cập internet nhiều, B là biến cố người được chọn xem tivi nhiều. Khi đó:

Example

Theo số liệu thống kê ở một vùng, tỷ lệ dân truy cập internet nhiều chiếm 40%, tỷ lệ dân xem ti vi nhiều chiếm 50%, tỷ lệ dân vừa truy cập internet nhiều vừa xem tivi nhiều chiếm 10%. Chọn ngẫu nhiên một người dân ở vùng trên.

1. Tính xác suất để người này truy cập internet nhiều hoặc xem tivi nhiều.
2. Tính xác suất để người này không truy cập internet nhiều.

Để tiện cho việc trình bày, ta gọi A là biến cố người được chọn truy cập internet nhiều, B là biến cố người được chọn xem tivi nhiều. Khi đó:

1. Xác suất để người này truy cập internet nhiều hoặc xem tivi nhiều là $P(A+B) = P(A) + P(B) - P(AB) = 0.4 + 0.5 - 0.1 = 0.8$
2. Xác suất để người này không truy cập internet nhiều là $P(\bar{A}) = 1 - P(A) = 1 - 0.4 = 0.6$.

Example

Tung một đồng xu hai lần, gọi A là biến cố cả hai lần đều được mặt sấp, B là biến cố: lần tung đầu tiên được mặt sấp. Khi đó các kết quả đồng khả năng là $\Omega = \{SS, NN, SN, NS\}, A = \{SS\}$, do đó$P(A) = 1/4$. Nhưng nếu biết biến cố B xảy ra, khi đó chỉ có hai khả năng là SN và SS. Xác suất để A xảy ra khi này là $1/2$ và ta nói: xác suất để xảy ra A khi B đã xảy ra là $1/2$.

Definition

Xác suất của biến cố A được tính với điều kiện biến cố B đã xảy ra gọi là xác suất có điều kiện của A khi B đã xảy ra và kí hiệu là $P(A|B)$.

Theorem

Xác suất có điều kiện của B khi A đã xảy ra ($P(A) > 0$), được tính bởi công thức:

$$P(B|A) = \frac{P(AB)}{P(A)} \tag{1}$$

Do vậy với A và B là hai biến cố tùy ý, từ công thức (1) ta có

$$P(AB) = P(A)P(B|A) = P(B)P(A|B) \tag{2}$$

Example

Trong một lễ hóa trang tại một hội trường có hai phòng tách biệt. Phòng thứ nhất có 20 nam và 25 nữ, phòng thứ 2 có 30 nữ và 20 nam. Một người vừa từ phòng thứ nhất sang phòng thứ 2. Chọn ngẫu nhiên một người trong phòng 2. Tính xác suất để:

1. Người được chọn là nam, biết người vừa sang là nữ.
2. Người được chọn là nam.

Definition

Hai biến cố A và B được gọi là độc lập với nhau nếu nếu việc xảy ra hay không xảy ra của biến cố này không làm thay đổi xác suất xảy ra của biến cố kia và ngược lại.

Theorem

Để kiểm tra xem hai biến cố A, B có độc lập nhau không ta kiểm tra một trong các điều kiện sau:

- $P(A|B) = P(A)$.
- $P(A|\bar{B}) = P(A)$.
- $P(AB) = P(A)P(B)$.

Example

Điều tra mức độ thường xuyên xem các chương trình thể thao của 500 cặp vợ chồng, ta thu được bảng số liệu sau:

Mức độ chồng	Mức độ vợ: Thường xuyên	Mức độ vợ: Không thường xuyên
Thường xuyên	100	150
Không thường xuyên	50	200

Example

Điều tra mức độ thường xuyên xem các chương trình thể thao của 500 cặp vợ chồng, ta thu được bảng số liệu sau:

Mức độ chồng	Mức độ vợ: Thường xuyên	Mức độ vợ: Không thường xuyên
Thường xuyên	100	150
Không thường xuyên	50	200

1. Tính xác suất để người chồng trong cặp được chọn thường xuyên xem chương trình thể thao.
2. Tính xác suất để người chồng trong cặp được chọn thường xuyên xem chương trình thể thao biết rằng người vợ thường xem chương trình thể thao.
3. Hai biến cố "Vợ xem chương trình thể thao" và "Chồng xem chương trình thể thao" có độc lập với nhau không?

Example

Hai xe bus hoạt động độc lập nhau trên một tuyến đường. Xác suất để xe thứ nhất đến trạm chờ gần nhà An trong khoảng 6h30 - 7h là 60%, xác suất tương ứng cho xe thứ 2 là 80%. Tính xác suất để:

1. Cả hai xe cùng đến trong khoảng thời gian trên.
2. Ít nhất một trong hai xe đến trong khoảng thời gian trên.
3. Cả hai xe đều không đến trong khoảng thời gian trên.

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$
3. *Đánh thứ tự 70 người trong danh sách*
4. *Có bao nhiêu khả năng cho ngày sinh nhật của 70 người trên?*

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$
3. *Đánh thứ tự 70 người trong danh sách*
4. *Có bao nhiêu khả năng cho ngày sinh nhật của 70 người trên?*
   $\longrightarrow 366^{70}$

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$
3. *Đánh thứ tự 70 người trong danh sách*
4. *Có bao nhiêu khả năng cho ngày sinh nhật của 70 người trên?*
   $\longrightarrow 366^{70}$
5. *Có bao nhiêu khả năng xảy ra mà ở đó 70 người có sinh nhật hoàn toàn khác nhau (không cặp nào giống nhau)*

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$
3. *Đánh thứ tự 70 người trong danh sách*
4. *Có bao nhiêu khả năng cho ngày sinh nhật của 70 người trên?*
   $\longrightarrow 366^{70}$
5. *Có bao nhiêu khả năng xảy ra mà ở đó 70 người có sinh nhật hoàn toàn khác nhau (không cặp nào giống nhau)*
   $\longrightarrow 366.365\ldots(366-69)$

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$
3. *Đánh thứ tự 70 người trong danh sách*
4. *Có bao nhiêu khả năng cho ngày sinh nhật của 70 người trên?*
   $\longrightarrow 366^{70}$
5. *Có bao nhiêu khả năng xảy ra mà ở đó 70 người có sinh nhật hoàn toàn khác nhau (không cặp nào giống nhau)*
   $\longrightarrow 366.365\ldots(366-69)$
6. $P(\bar{A}) = \dfrac{366.365\ldots(366-69)}{366^{70}}$

Example

Giả sử một lớp học môn xác suất thống kê của chúng ta có 70 sinh viên. Hãy tính xác suất để có ít nhất hai người cùng sinh nhật.

Solution

1. *A: "trong số 70 sinh viên trên có ít nhất hai người cùng sinh nhật"*
2. $P(A) = 1 - P(\bar{A})$
3. *Đánh thứ tự 70 người trong danh sách*
4. *Có bao nhiêu khả năng cho ngày sinh nhật của 70 người trên?*
   $\longrightarrow 366^{70}$
5. *Có bao nhiêu khả năng xảy ra mà ở đó 70 người có sinh nhật hoàn toàn khác nhau (không cặp nào giống nhau)*
   $\longrightarrow 366.365\ldots(366-69)$
6. $P(\bar{A}) = \dfrac{366.365\ldots(366-69)}{366^{70}} \longrightarrow P(A) = 0.999$

Công thức xác suất đầy đủ

Definition

Một hệ gồm n biến cố $\{A_1, A_2, \ldots, A_n\}$ được gọi là hệ biến cố đầy đủ nếu chúng đôi một xung khắc và hợp của chúng là biến cố chắc chắn.

Nói riêng hệ 2 biến cố $\{A, \bar{A}\}$ là đầy đủ, vì A và $\bar{A}$ là xung khắc và $A \cup \bar{A} = \Omega$.

Công thức xác suất đầy đủ

Theorem

Cho hệ biến cố đầy đủ $\{A_1, A_2, \ldots, A_n\}$. *Khi đó một biến cố* X *bất kì có xác suất được tính theo công thức:*

$$P(X) = P(X|A_1).P(A_1) + \ldots + P(X|A_n).P(A_n) = \sum_{k=1}^{n} P(X|A_k).P(A_k)$$

Example

Giả sử trong số nữ có 80% dùng smartphone, tỉ lệ này trong nhóm nam là 75%. Biết rằng tỉ lệ nam chiếm 52% dân số đất nước. Tính tỉ lệ người dùng smartphone trên tổng số dân.

Theorem (Công thức Bayes)

Cho hệ biến cố đầy đủ $\{A_1, A_2, \ldots, A_n\}$*. Khi đó nếu biết* X *đã xảy ra thì xác suất để* A_i *xảy ra là:*

$$P(A_i|X) = \frac{P(XA_i)}{P(X)} = \frac{P(X|A_i).P(A_i)}{\sum\limits_{k=1}^{n} P(X|A_k).P(A_k)} \tag{3}$$

Công thức Bayes

Example

Trong lúc mất điện tại một trung tâm thương mại, một tên trộm đã giật nữ trang của một cô gái, hệ thống an ninh đã phong tỏa và toàn bộ 500 người đang có mặt đều bị giữ lại thẩm vấn. Vì không kịp thoát ra ngoài nên kẻ trộm đã bỏ đồ nữ trang vào thùng rác. Không có camera nào ghi lại được hình ảnh lúc đó.

Một máy kiểm tra phát hiện nói dối được sử dụng. Thí nghiệm đã chỉ cứ 100 người nói dối thì nó chỉ phát hiện được 95 người.

Đồng thời thử 100 người nói thật, nó vẫn trả kết quả có 2 trong số đó nói dối.

Một người đang được thẩm vấn với sự tham gia của máy phát hiện nói dối trên. Anh ta phủ nhận việc lấy trộm đồ.

Công thức Bayes

Example

Trong lúc mất điện tại một trung tâm thương mại, một tên trộm đã giật nữ trang của một cô gái, hệ thống an ninh đã phong tỏa và toàn bộ 500 người đang có mặt đều bị giữ lại thẩm vấn. Vì không kịp thoát ra ngoài nên kẻ trộm đã bỏ đồ nữ trang vào thùng rác. Không có camera nào ghi lại được hình ảnh lúc đó.
Một máy kiểm tra phát hiện nói dối được sử dụng. Thí nghiệm đã chỉ cứ 100 người nói dối thì nó chỉ phát hiện được 95 người.
Đồng thời thử 100 người nói thật, nó vẫn trả kết quả có 2 trong số đó nói dối.
Một người đang được thẩm vấn với sự tham gia của máy phát hiện nói dối trên. Anh ta phủ nhận việc lấy trộm đồ.

1. Xác suất để anh ta bị máy kết luận nói dối là bao nhiêu?
2. Kết quả cuối cùng máy kết luận anh ta nói dối. Vậy, xác suất để anh ta là kẻ trộm là bao nhiêu?

Câu hỏi

Một khảo sát 400 bạn trẻ về sự yêu thích đến rạp chiếu phim vào cuối tuần được kết quả như sau:

	Nam	*Nữ*
Thích	*110*	*90*
Không thích	*120*	*80*

Chọn ngẫu nhiên một bạn trong 400 bạn đó. Tính xác suất để:

1. Bạn này thích đi xem phim vào cuối tuần.
2. Bạn này thích đi xem phim vào cuối tuần, biết bạn này là nữ.

Hai biến cố "người được chọn là nam" và "người được chọn thích tới rạp xem phim vào cuối tuần" có độc lập nhau không?

Câu hỏi

Một siêu thị nhập rau sạch từ 3 gia đình A, B, C cùng trong một hợp tác xã sản suất theo tiêu chuẩn VietGap. Tỉ trọng nhập từ mỗi gia đình là 20 %, 30 % và 50 %. Tỉ lệ rau bị hỏng của mỗi gia đình sau 3 ngày bảo quản tương ứng là $5\%, 3\%, 2\%$.

Câu hỏi

Một siêu thị nhập rau sạch từ 3 gia đình A, B, C cùng trong một hợp tác xã sản suất theo tiêu chuẩn VietGap. Tỉ trọng nhập từ mỗi gia đình là 20 %, 30 % và 50 %. Tỉ lệ rau bị hỏng của mỗi gia đình sau 3 ngày bảo quản tương ứng là $5\%, 3\%, 2\%$.

1. *Tính tỉ lệ rau hỏng sau 3 ngày.*
2. *Một khách hàng đã mua rau tại khu vực rau đã được bảo quản trong 3 ngày và phàn nàn qua điện thoại rằng có một mớ rau của họ bị hỏng bên phần lõi. Hỏi xác suất để mớ rau này của gia đình A là bao nhiêu?*

Example

Xổ số 6/45 Xổ số mega 6/45 là hình thức xổ được du nhập từ Mỹ, được mở thưởng đầu tiên vào tối 20/7/2016 tại Việt Nam. Hình thức chơi đơn giản nhất là bạn chọn 6 con số từ 01 đến 45. Cơ cấu giải như sau:
Giải ba: 30.000 đồng (trùng khớp 3 con trong 6 con)
Giải nhì: 300.000 đồng (trùng khớp 4 con trong 6 con)
Giải nhất: 10.000.000 đồng (trùng khớp 5 con trong 6 con)
Giải đặc biệt: tối thiểu 12 tỷ đồng và tích lũy (trùng khớp 6 con)
Sự trùng khớp ở đây không cần theo thứ tự

Xổ số 6/45 ngày 2 - 9 - 2016

KẾT QUẢ TRÚNG THƯỞNG MEGA 6/45

Kỳ quay thưởng #00020 | Ngày quay thưởng 02/09/2016

09 13 24 35 36 41

Giá trị Jackpot

29.958.428.500 đồng

Giải thưởng	Trùng khớp	Số lượng giải	Giá trị giải (đồng)
Jackpot	○○○○○○	0	29.958.428.500
Giải nhất	○○○○○	17	10.000.000
Giải nhì	○○○○	649	300.000
Giải ba	○○○	11459	30.000

Kết quả ngày 2 - 9 - 2016

Giải thưởng	Trùng khớp	Số lượng giải	Giá trị giải (vnđ)
Jackpot	Trùng 6 số	0	29.958.428.500
Giải nhất	Trùng 5 số	17	10.000.000
Giải nhì	Trùng 4 số	649	300.000
Giải ba	Trùng 3 số	11459	30.000

Câu hỏi

Bạn tham gia dự thưởng một bộ số. Tính xác xuất để bạn được giải:

1. *đặc biệt*
2. *nhất.*
3. *nhì.*
4. *ba.*

1. Xác suất là gì?
   - Những khái niệm cơ bản của xác suất
     - Không gian mẫu, biến cố sơ cấp và biến cố
     - Quan hệ giữa các biến cố
   - Định nghĩa xác suất
     - Khái niệm chung về xác suất
     - Định nghĩa cổ điển về xác suất
     - Định nghĩa thống kê về xác suất
     - Định nghĩa xác suất theo nghĩa chủ quan
   - Một số qui tắc tính xác suất
     - Qui tắc cộng xác suất
     - Qui tắc xác suất có điều kiện
     - Công thức xác suất đầy đủ
2. Biến ngẫu nhiên
   - Phân loại biến ngẫu nhiên
   - Phân phối xác suất của biến ngẫu nhiên rời rạc
   - Phân phối xác suất của biến ngẫu nhiên liên tục
   - Trung bình, phương sai của biến ngẫu nhiên
   - Chữa một số bài tập về xác suất và biến ngẫu nhiên
   - Một số phân phối quan trọng
     - Phân phối nhị thức
     - Phân phối chuẩn

Biến ngẫu nhiên

Khái niệm

Việc gán chô mỗi kết quả của một phép thử một con số theo một quy tắc nào đó cho ta một **biến ngẫu nhiên**.

Biến ngẫu nhiên

Khái niệm

Việc gán chô mỗi kết quả của một phép thử một con số theo một quy tắc nào đó cho ta một **biến ngẫu nhiên**.

Example

Chọn ngẫu nhiên một lớp ở TLU, gọi X là số sinh viên đang học ở lớp đó. Khi đó X là một biến ngẫu nhiên.

Example

Xét phép thử: tung hai đồng xu.
Ký hiệu S là trường hợp đồng xu xuất hiện mặt số và N là trường hợp mặt còn lại, tại gọi là mặt ngửa, ta có các kết quả có thể là: SS,SN,NS,NN.
Gọi X là đại lượng chỉ số mặt ngửa xuất hiện trong phép thử.

Example

Xét phép thử: tung hai đồng xu.

Ký hiệu S là trường hợp đồng xu xuất hiện mặt số và N là trường hợp mặt còn lại, tại gọi là mặt ngửa, ta có các kết quả có thể là: SS,SN,NS,NN.

Gọi X là đại lượng chỉ số mặt ngửa xuất hiện trong phép thử.

Ta có $X(SS) = 0 \qquad X(SN) = 1 \qquad X(NS) = 1 \qquad X(NN) = 2.$

Biến ngẫu nhiên

Example

Xét phép thử: tung hai đồng xu.
Ký hiệu S là trường hợp đồng xu xuất hiện mặt số và N là trường hợp mặt còn lại, tại gọi là mặt ngửa, ta có các kết quả có thể là: SS,SN,NS,NN.
Gọi X là đại lượng chỉ số mặt ngửa xuất hiện trong phép thử.
Ta có $X(SS) = 0 \qquad X(SN) = 1 \qquad X(NS) = 1 \qquad X(NN) = 2.$
Như vậy X có thể nhận ba giá trị: 0, 1, 2 tùy thuộc vào kết quả của phép thử. X ở đây là một biến ngẫu nhiên.

Biến ngẫu nhiên

Example

Biến X trong mỗi tình huống sau là biến ngẫu nhiên:

1. X là số ô tô một cửa showroom của Toyota Cầu Giấy bán được trong một tháng được chọn ngẫu nhiên.

Biến ngẫu nhiên

Example

Biến X trong mỗi tình huống sau là biến ngẫu nhiên:

1. X là số ô tô một cửa showroom của Toyota Cầu Giấy bán được trong một tháng được chọn ngẫu nhiên.
2. X là điểm thi môn XSTK của một sinh viên được chọn ngẫu nhiên trong lớp.

Biến ngẫu nhiên

Example

Biến X trong mỗi tình huống sau là biến ngẫu nhiên:

1. X là số ô tô một cửa showroom của Toyota Cầu Giấy bán được trong một tháng được chọn ngẫu nhiên.
2. X là điểm thi môn XSTK của một sinh viên được chọn ngẫu nhiên trong lớp.
3. X là số ôtô qua đường Nguyễn Xiển trong một ngày được chọn ngẫu nhiên.

Biến ngẫu nhiên

Example

Biến X trong mỗi tình huống sau là biến ngẫu nhiên:

1. X là số ô tô một cửa showroom của Toyota Cầu Giấy bán được trong một tháng được chọn ngẫu nhiên.
2. X là điểm thi môn XSTK của một sinh viên được chọn ngẫu nhiên trong lớp.
3. X là số ôtô qua đường Nguyễn Xiển trong một ngày được chọn ngẫu nhiên.
4. X là chiều cao của một người trưởng thành nào đó đươc chọn ngẫu nhiên.

Biến ngẫu nhiên

Example

Biến X trong mỗi tình huống sau là biến ngẫu nhiên:

1. X là số ô tô một cửa showroom của Toyota Cầu Giấy bán được trong một tháng được chọn ngẫu nhiên.
2. X là điểm thi môn XSTK của một sinh viên được chọn ngẫu nhiên trong lớp.
3. X là số ôtô qua đường Nguyễn Xiển trong một ngày được chọn ngẫu nhiên.
4. X là chiều cao của một người trưởng thành nào đó đươc chọn ngẫu nhiên.
5. X là số thành viên của một gia đình được chọn ngẫu nhiên ở Hà Nội.

Biến ngẫu nhiên

Example

Biến X trong mỗi tình huống sau là biến ngẫu nhiên:

1. X là số ô tô một cửa showroom của Toyota Cầu Giấy bán được trong một tháng được chọn ngẫu nhiên.
2. X là điểm thi môn XSTK của một sinh viên được chọn ngẫu nhiên trong lớp.
3. X là số ôtô qua đường Nguyễn Xiển trong một ngày được chọn ngẫu nhiên.
4. X là chiều cao của một người trưởng thành nào đó đươc chọn ngẫu nhiên.
5. X là số thành viên của một gia đình được chọn ngẫu nhiên ở Hà Nội.
6. Chọn một thang máy ngẫu nhiên. Gọi X là số lần thang máy này lên đến tầng 8 trong tuần này.

Phân loại biến ngẫu nhiên

1. Biến ngẫu nhiên rời rạc là biến ngẫu nhiên mà chỉ nhận giá trị: $x_1, x_2, \ldots$
2. Biến ngẫu nhiên liên tục là những biến ngẫu nhiên mà giá trị của nó có thể phủ kín một khoảng (A, B) trên tập số thực.

Example

1. Chọn một phòng học ngẫu nhiên tại ĐHTL. Gọi X là số sinh viên đang học ở lớp đó. Khi đó X là biến ngẫu nhiên rời rạc.
2. Sút một quả bóng trên sân. Gọi X là khoảng cách tính đến chính xác mà quả bóng đi được. Khi đó X có thể mang các giá trị rất "lẻ": 45m, 45.1m, 45.0001m, ...

Phân loại biến ngẫu nhiên

1. Biến ngẫu nhiên rời rạc là biến ngẫu nhiên mà chỉ nhận giá trị: $x_1, x_2, \ldots$
2. Biến ngẫu nhiên liên tục là những biến ngẫu nhiên mà giá trị của nó có thể phủ kín một khoảng (A, B) trên tập số thực.

Example

1. Chọn một phòng học ngẫu nhiên tại ĐHTL. Gọi X là số sinh viên đang học ở lớp đó. Khi đó X là biến ngẫu nhiên rời rạc.
2. Sút một quả bóng trên sân. Gọi X là khoảng cách tính đến chính xác mà quả bóng đi được. Khi đó X có thể mang các giá trị rất "lẻ": 45m, 45.1m, 45.0001m, $\ldots .Y$ có thể lấy giá trị quét đầy một khoảng số chẳng hạn (0,100). Y là biến ngẫu nhiên liên tục.

Phân phối xác suất của biến ngẫu nhiên rời rạc

Definition

Cho một biến ngẫu nhiên rời rạc X. Khi ta mô tả tất cả các giá trị X có thể nhận cùng với xác suất tương ứng để X nhận các giá trị đó, ta nói ta có quy luật phân phối xác suất của X.

Phân phối xác suất của biến ngẫu nhiên rời rạc

Definition

Cho một biến ngẫu nhiên rời rạc X. Khi ta mô tả tất cả các giá trị X có thể nhận cùng với xác suất tương ứng để X nhận các giá trị đó, ta nói ta có quy luật phân phối xác suất của X.

Có thể mô tả bằng công thức hoặc bảng như:

Giá trị của X	x_1	x_2	x_3	$\ldots$	x_k
Xác suất	p_1	p_2	p_3	$\ldots$	p_k

Example

Xét phép thử tung hai lần một đồng xu, gọi X là số lần xuất hiện mặt N.

Phân phối xác suất của biến ngẫu nhiên rời rạc

Example

Xét phép thử tung hai lần một đồng xu, gọi X là số lần xuất hiện mặt N.Ta có:

Biến cố	SS	SN	NS	NN
Giá trị của X	0	1	1	2

Example

Xét phép thử tung hai lần một đồng xu, gọi X là số lần xuất hiện mặt N.Ta có:

Biến cố	SS	SN	NS	NN
Giá trị của X	0	1	1	2

Phân phối xác suất của X:

Giá tri của X	0	1	2

Phân phối xác suất của biến ngẫu nhiên rời rạc

Example

Xét phép thử tung hai lần một đồng xu, gọi X là số lần xuất hiện mặt N.Ta có:

Biến cố	SS	SN	NS	NN
Giá trị của X	0	1	1	2

Phân phối xác suất của X:

Giá tri của X	0	1	2
Xác suất	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Giả sử biến ngẫu nhiên X có bảng phân phối xác suất sau:

Giá trị của X	x_1	x_2	x_3	$\dots$	x_k
Xác suất	p_1	p_2	p_3	$\dots$	p_k

Khi đó ta có:

1. các số p_i nằm giữa 0 và 1.
2. tổng $p_1 + p_2 + \ldots + p_k = 1$.

Example

Trong ví dụ sút bóng phía trên, P(X=45)=? Rõ ràng là nó chỉ là một giá trị trong vô số các giá trị có thể đạt được.

Example

Trong ví dụ sút bóng phía trên, P(X=45)=? Rõ ràng là nó chỉ là một giá trị trong vô số các giá trị có thể đạt được.

1. Cứ giả dụ là X lấy các giá trị với khả năng như nhau cho đơn giản. Khi đó P(X=45)=0.

Phân phối xác suất của biến ngẫu nhiên liên tục

Example

Trong ví dụ sút bóng phía trên, P(X=45)=? Rõ ràng là nó chỉ là một giá trị trong vô số các giá trị có thể đạt được.

1. Cứ giả dụ là X lấy các giá trị với khả năng như nhau cho đơn giản. Khi đó P(X=45)=0.
   Nếu như vậy thì tổng xác suất trên toàn bộ sẽ là 0. Mâu thuẫn!
2. Sự thật là khả năng quả bóng lăn khoảng 3m sẽ thấp hơn so với khoảng 30 m.

Example

Trong ví dụ sút bóng phía trên, P(X=45)=? Rõ ràng là nó chỉ là một giá trị trong vô số các giá trị có thể đạt được.

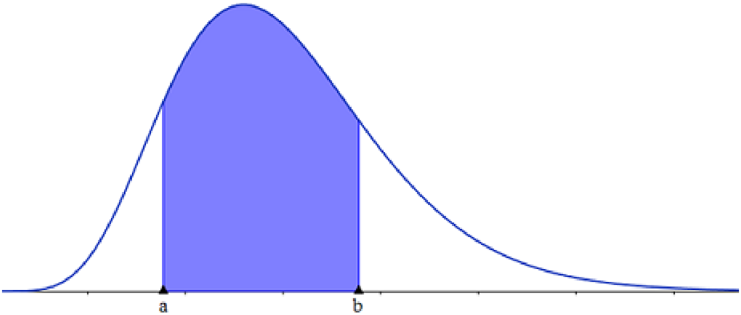
1. Cứ giả dụ là X lấy các giá trị với khả năng như nhau cho đơn giản. Khi đó P(X=45)=0.
   Nếu như vậy thì tổng xác suất trên toàn bộ sẽ là 0. Mâu thuẫn!
2. Sự thật là khả năng quả bóng lăn khoảng 3m sẽ thấp hơn so với khoảng 30 m.
   Làm sao để mô tả phân phối xác suất của X?

Phân phối xác suất của biến ngẫu nhiên liên tục

Definition

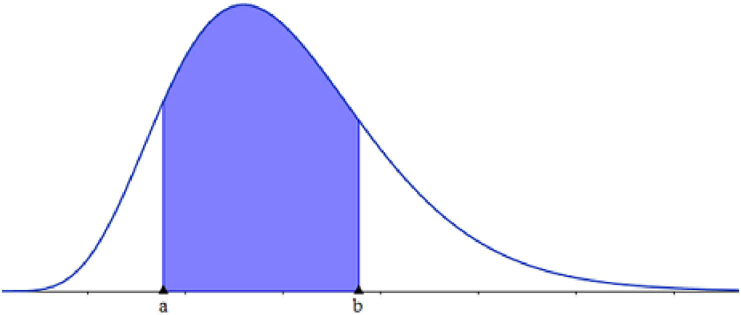
Một cách tổng quát, phân phối xác suất của biến ngẫu nhiên liên tục được mô tả qua các xác suất dạng $P(a < X < b)$. Ở đây ta xét trường hợp mà các xác suất này có thể được tính từ một hàm (qua một công thức tích phân), gọi là hàm mật độ của X.

Đường cong giới hạn là đồ thị của hàm mật độ $f(x)$, còn diện tích của phần được tô chính là $P(a < X < b)$.



Phân phối xác suất của biến ngẫu nhiên liên tục

Đường cong giới hạn là đồ thị của hàm mật độ $f(x)$, còn diện tích của phần được tô chính là $P(a < X < b)$.

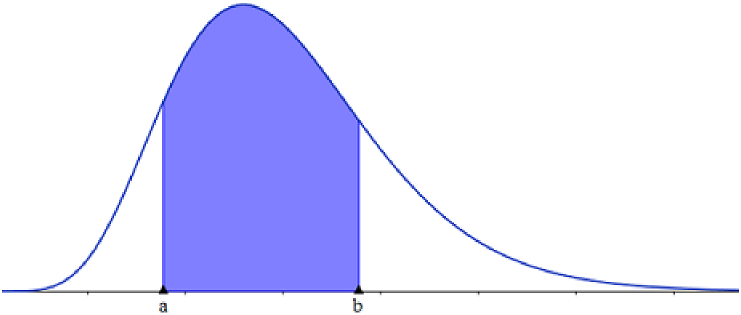


$f(x)$ phải thỏa mãn hai điều kiện:

1. $f(x) \geq 0 \ \forall x$.
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Phân phối xác suất của biến ngẫu nhiên liên tục

Đường cong giới hạn là đồ thị của hàm mật độ $f(x)$, còn diện tích của phần được tô chính là $P(a < X < b)$.



$f(x)$ phải thỏa mãn hai điều kiện:

1. $f(x) \geq 0 \; \forall x$.
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Như vậy, $P(a < X < b) = \int_a^b f(x)dx$.

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

$$5 \quad 6 \quad 8 \quad 3 \quad 5 \quad 8 \quad 8 \quad 6 \quad 4 \quad 5$$

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất					

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

5 6 8 3 5 8 8 6 4 5

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10				

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

5	6	8	3	5	8	8	6	4	5

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10	1/10			

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

$$5 \quad 6 \quad 8 \quad 3 \quad 5 \quad 8 \quad 8 \quad 6 \quad 4 \quad 5$$

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10	1/10	3/10		

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

5 6 8 3 5 8 8 6 4 5

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10	1/10	3/10	2/10	

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

$$5 \quad 6 \quad 8 \quad 3 \quad 5 \quad 8 \quad 8 \quad 6 \quad 4 \quad 5$$

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10	1/10	3/10	2/10	3/10

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

$$5 \quad 6 \quad 8 \quad 3 \quad 5 \quad 8 \quad 8 \quad 6 \quad 4 \quad 5$$

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10	1/10	3/10	2/10	3/10

Ta thấy rằng lương trung bình của 10 người trên là

$$(5+6+8+3+5+8+8+6+4+5)/10 = 5.8$$

Example

Giả sử lương tháng của 10 người được cho ở dãy sau (đơn vị triệu):

$$5 \quad 6 \quad 8 \quad 3 \quad 5 \quad 8 \quad 8 \quad 6 \quad 4 \quad 5$$

Chọn ngẫu nhiên một người trong số 10 người, gọi X là lương tháng vừa qua của người đó. Ta có:

X	3	4	5	6	8
Xác suất	1/10	1/10	3/10	2/10	3/10

Ta thấy rằng lương trung bình của 10 người trên là

$$(5+6+8+3+5+8+8+6+4+5)/10 = 5.8$$

Ta cũng có thể tính giá trị này từ bảng phân phối xác suất của X:

$$\frac{1}{10} \times 3 + \frac{1}{10} \times 4 + \frac{3}{10} \times 5 + \frac{2}{10} \times 6 + \frac{3}{10} \times 8 = 5.8$$

Definition

Một cách tổng quát, nếu **biến ngẫu nhiên rời rạc** X có phân phối xác suất

Giá trị của X	x_1	x_2	x_3	$\ldots$	x_k
Xác suất	p_1	p_2	p_3	$\ldots$	p_k

thì kì vọng của X là

$$E(X) = x_1 \times p_1 + x_2 \times p_2 + \ldots + x_k \times p_k$$

Trung bình, phương sai của biến ngẫu nhiên

Definition

Một cách tổng quát, nếu **biến ngẫu nhiên rời rạc** X có phân phối xác suất

Giá trị của X	x_1	x_2	x_3	$\ldots$	x_k
Xác suất	p_1	p_2	p_3	$\ldots$	p_k

thì kì vọng của X là

$$E(X) = x_1 \times p_1 + x_2 \times p_2 + \ldots + x_k \times p_k$$

Đối với **biến ngẫu nhiên liên tục** X có hàm mật độ $f(x)$, kì vọng được tính theo công thức:

$$E(X) = \int_{-\infty}^{+\infty} x.f(x)dx$$

Trung bình, phương sai của biến ngẫu nhiên

Để đo mức độ phân tán các giá trị của X, ta dùng khái niệm phương sai của biến ngẫu nhiên:

$$V(X) = E(X^2) - [E(X)]^2$$

trong đó, nếu X là biến rời rạc:

$$E(X^2) = x_1^2 \times p_1 + x_2^2 \times p_2 + \ldots + x_k^2 \times p_k$$

Để đo mức độ phân tán các giá trị của X, ta dùng khái niệm phương sai của biến ngẫu nhiên:

$$V(X) = E(X^2) - [E(X)]^2$$

trong đó, nếu X là biến rời rạc:

$$E(X^2) = x_1^2 \times p_1 + x_2^2 \times p_2 + \ldots + x_k^2 \times p_k$$

còn trong trường hợp X liên tục:

$$E(X^2) = \int_{-\infty}^{+\infty} x^2.f(x)dx$$

Example

Một người tham gia trò cờ bạc 2 đen một đỏ như sau: có hai quân bài đen và một quân đỏ. Người chủ trò sẽ tráo 3 quân và sau đó cho người chơi lật 1 quân. Người chơi sẽ thắng nếu lật được quân đỏ và số tiền được nhận là 100 nghìn.

1. Giả sử mỗi lần thua người chơi phải trả 60 nghìn. Hỏi, nếu chơi rất nhiều lần, trung bình lợi nhuận anh ta trong mỗi lần chơi đó sẽ là bao nhiêu?
2. Số tiền phải trả mỗi lần thua là bao nhiêu để tổng thể những người chơi có lãi?

Solution

Gọi X là lợi nhuận (nghìn đồng) của một lần chơi được chọn ngẫu nhiên.

Solution

Gọi X là lợi nhuận (nghìn đồng) của một lần chơi được chọn ngẫu nhiên.

1. $P(X = -60) = 2/3$ *và* $P(X = 100) = 1/3$*:*

Solution

Gọi X là lợi nhuận (nghìn đồng) của một lần chơi được chọn ngẫu nhiên.

1. $P(X = -60) = 2/3$ *và* $P(X = 100) = 1/3$*:*

$$E(X) = -60.\frac{2}{3} + 100.\frac{1}{3} = -\frac{20}{3}$$

Lợi nhuận trung bình mỗi lần là $-\dfrac{20}{3}$ *nghìn: Lỗ*

Solution

Gọi X là lợi nhuận (nghìn đồng) của một lần chơi được chọn ngẫu nhiên.

1. $P(X = -60) = 2/3$ *và* $P(X = 100) = 1/3$*:*

$$E(X) = -60.\frac{2}{3} + 100.\frac{1}{3} = -\frac{20}{3}$$

Lợi nhuận trung bình mỗi lần là $-\dfrac{20}{3}$ *nghìn: Lỗ*

2. *Gọi a là số tiền phải trả cho một lần chơi, ta có:*

X	-a	100
p	2/3	1/3

Solution

Gọi X là lợi nhuận (nghìn đồng) của một lần chơi được chọn ngẫu nhiên.

1. $P(X = -60) = 2/3$ *và* $P(X = 100) = 1/3$*:*

$$E(X) = -60.\frac{2}{3} + 100.\frac{1}{3} = -\frac{20}{3}$$

Lợi nhuận trung bình mỗi lần là $-\dfrac{20}{3}$ *nghìn: Lỗ*

2. *Gọi a là số tiền phải trả cho một lần chơi, ta có:*

X	-a	100
p	2/3	1/3

$$E(X) = \frac{-2a}{2} + \frac{100}{3} = \frac{100-2a}{2}$$

Do đó, tổng thể người chơi lãi khi và chỉ khi $E(X) > 0$,

Solution

Gọi X là lợi nhuận (nghìn đồng) của một lần chơi được chọn ngẫu nhiên.

1. $P(X = -60) = 2/3$ *và* $P(X = 100) = 1/3$*:*

$$E(X) = -60.\frac{2}{3} + 100.\frac{1}{3} = -\frac{20}{3}$$

Lợi nhuận trung bình mỗi lần là $-\frac{20}{3}$ *nghìn: Lỗ*

2. *Gọi a là số tiền phải trả cho một lần chơi, ta có:*

X	-a	100
p	2/3	1/3

$$E(X) = \frac{-2a}{2} + \frac{100}{3} = \frac{100 - 2a}{2}$$

Do đó, tổng thể người chơi lãi khi và chỉ khi $E(X) > 0$*, tức là khi* $a < 50$*. Vậy, mỗi lần chơi chỉ phải trả ít hơn 50 nghìn.*

Example

Giả sử bạn mua một số đề. Nếu mua 1 nghìn mà "trúng" thì được 70 nghìn.

1. Gọi X là lợi nhuận khi bỏ ra một nghìn để mua một số đề. Hãy tính lợi nhuận trung bình thu được.
2. Nếu bỏ ra 50 nghìn mua một số thì lợi nhuận trung bình là boa nhiêu? Phương sai là bao nhiêu?

1. Ta xét một cách chơi khác như sau: Cũng với số tiền 50 nghìn như trên, nhưng rải đều mua 50 số. Gọi Y là lợi nhuận thu được. Lập bảng phân phối cho Y. Tính $E(Y), V(Y)$. So sánh với $E(X), V(X)$
2. Khảo sát số đề 11 tháng qua, lọc ra những số đề ít xuất hiện. Thực hiện chiến thuật rải đều cho 1 tháng tiếp theo với cách chơi lũy tiến, tức là: số tiền mua hôm sau phải đủ lớn để nếu trúng thu lại được toàn bộ lỗ và có một chút lãi. Chiến lược này liệu có hiệu quả?

Example

Trong một trò chơi truyền hình, một người chơi được chọn một trong 100 cánh cửa, phần quà sẽ là vật được ghi phía sau cánh cửa. Biết rằng đằng sau 99 cánh cửa là "1 con dê", chỉ có một cánh cửa là "Một xe Audi Q8". Bạn là một người chơi may mắn được tham gia mở cửa nhận quà và bạn mong mở được cửa có xe. Bạn chọn một cửa.
Sau đó, vị MC mở lần lượt 98 cửa có ghi "1 con dê", còn giữ lại 2 cửa (trong đó có cửa mà bạn đã chọn). MC hỏi bạn có đổi cánh cửa để lấy cánh cửa còn lại hay không?
Hỏi nếu bạn đổi cửa lấy cửa kia khả năng nhận được có xe có tăng không? vì sao?

Example

Một anh chàng thích cả hai cô bạn gái Huệ và Hồng nhưng không biết quyết định sẽ tỏ tình với cô nào. Anh ta đến nhà hai cô gái một cách ngẫu nhiên bằng cách ra trạm xe bus gần nhà, nếu gặp xe đến nhà cô nào trước thì lên xe đó. Mỗi tuyến xe đến nhà hai cô gái đều có đều đặn 15 phút một lần. Sau một thời gian dài, anh ta nhận ra mình đến nhà bạn Hồng nhiều hơn gấp đôi số lần đến nhà bạn Huệ. Hãy cho một giải thích về điều này.

Câu hỏi

Một người đang phân tích hai kết hoạch lập startup và sau đó nhượng quyền với thông tin như sau:

1. *Starup 1: Nếu thành công, với khoản đầu tư 15 nghìn đô sẽ mang lại 31.25 nghìn đô khi nhượng lại, nếu thất bại, việc thanh lí vớt vát lại 10 nghìn đô. Theo thống kê đã có 5000 startup như thế và số thành công lên tới 2000.*
2. *Starup 2: Nếu thành công, với khoản đầu tư 15 nghìn đô sẽ mang lại 100 nghìn đô khi nhượng lại, nếu thất bại, việc thanh lí vớt vát lại 5 nghìn đô. Theo thống kê đã có 2000 startup như thế và số thành công là 200.*

Dựa trên thông tin về lợi nhuận kì vọng và phương sai, nên lựa chọn phương án nào?

Phân phối nhị thức

Definition

Phép thử nhị thức là phép thử có các đặc điểm sau:

Phân phối nhị thức

Definition

Phép thử nhị thức là phép thử có các đặc điểm sau:

- Phép thử bao gồm n thử nghiệm giống hệt nhau.

Phân phối nhị thức

Definition

Phép thử nhị thức là phép thử có các đặc điểm sau:

- Phép thử bao gồm n thử nghiệm giống hệt nhau.
- Mỗi thử nghiệm này chỉ có kết quả là "thành công" hoặc "thất bại".

Phân phối nhị thức

Definition

Phép thử nhị thức là phép thử có các đặc điểm sau:

- Phép thử bao gồm n thử nghiệm giống hệt nhau.
- Mỗi thử nghiệm này chỉ có kết quả là "thành công" hoặc "thất bại".
- Xác suất "thành công" trong mỗi phép thử đều là p (và xác suất "thất bại" là $q = 1 - p$).

Phân phối nhị thức

Definition

Phép thử nhị thức là phép thử có các đặc điểm sau:

- Phép thử bao gồm n thử nghiệm giống hệt nhau.
- Mỗi thử nghiệm này chỉ có kết quả là "thành công" hoặc "thất bại".
- Xác suất "thành công" trong mỗi phép thử đều là p (và xác suất "thất bại" là $q = 1 - p$).
- Các thử nghiệm là độc lập lẫn nhau (kết quả của thử nghiệm này không ảnh hưởng tới kết quả của thử nghiệm khác).

Phân phối nhị thức

Example

1. Tung môt đồng xu 10 lần, trong mỗi lần tung, xác suất xuất hiện mặt S trong mỗi lần đều là 0.5.

Example

1. Tung một đồng xu 10 lần, trong mỗi lần tung, xác suất xuất hiện mặt S trong mỗi lần đều là 0.5.
2. Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25.

Phân phối nhị thức

Example

1. Tung môt đồng xu 10 lần, trong mỗi lần tung, xác suất xuất hiện mặt S trong mỗi lần đều là 0.5.
2. Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25.
3. Chọn 5 sản phẩm từ một lô hàng (có rất nhiều sản phẩm) mà tỉ lệ đạt tiêu chuẩn là 84%. Khi đó, xác suất sản phẩm là đạt tiêu chuẩn trong mỗi lần chọn đều có thể coi là 0.84.

Phân phối nhị thức

Example

1. Tung một đồng xu 10 lần, trong mỗi lần tung, xác suất xuất hiện mặt S trong mỗi lần đều là 0.5.
2. Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25.
3. Chọn 5 sản phẩm từ một lô hàng (có rất nhiều sản phẩm) mà tỉ lệ đạt tiêu chuẩn là 84%. Khi đó, xác suất sản phẩm là đạt tiêu chuẩn trong mỗi lần chọn đều có thể coi là 0.84.

Làm việc nhóm

Câu hỏi

Một thí sinh không học bài, trả lời ngẫu nhiên, độc lập từng câu hỏi trong một phiếu trắc nghiệm gồm n câu hỏi với 4 phương án lựa chọn ở mỗi câu và chỉ có một phương án đúng. Gọi X là số câu đúng. Lập bảng phân phối xác suất của X và tính EX, VX trong các trường hợp sau:

1. $n = 1$.
2. $n = 2$.
3. $n = 4$.
4. $n = 10$.

Phân phối nhị thức

Theorem

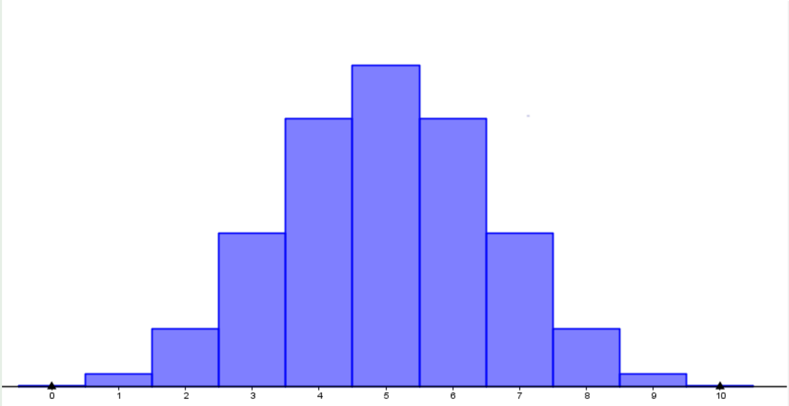
Gọi $X =$ *số lần "thành công" trong n thử nghiệm của phép thử nhị thức. thì* X *được gọi là biến ngẫu nhiên có phân phối nhị thức (biến ngẫu nhiên nhị thức), ký hiệu* $X \sim B(n, p)$*. Khi đó, xác suất để có k lần thành công trong n thử nghiệm là*

$$P(X = k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

Phân phối nhị thức

Example

Tung môt đồng xu 10 lần, trong mỗi lần tung, xác suất xuất hiện mặt S là 0.5. Gọi X là số mặt sấp, $X \sim B(10, 0.5)$.
Phân phối của X được minh họa như hình sau đây:



k	P(X = k)
0	0.001
1	0.0098
2	0.0439
3	0.1172
4	0.2051
5	0.2461
6	0.2051
7	0.1172
8	0.0439
9	0.0098
10	0.001

Theorem

Nếu $X \sim B(n, p)$ *thì trung bình của* X *là* $EX = np$, *phương sai của* X *là* $VX = np(1 - p)$.

Example

Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25. Khi đó gọi X là số câu trả lời đúng thì $X \sim B(100, 0.25)$.

Khi đó:

1. Xác suất đển trả lời đúng 45 câu hỏi là

Example

Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25. Khi đó gọi X là số câu trả lời đúng thì $X \sim B(100, 0.25)$.
Khi đó:

1. Xác suất đển trả lời đúng 45 câu hỏi là
$$P(X = 45) = C_{100}^{45} \times 0.25^{45} \times 0.75^{55} = 6.67 \times 10^{-6}$$
2. Xác suất để trả lời đúng dưới 50 câu là

Phân phối nhị thức

Example

Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25. Khi đó gọi X là số câu trả lời đúng thì $X \sim B(100, 0.25)$.

Khi đó:

1. Xác suất đến trả lời đúng 45 câu hỏi là
$$P(X = 45) = C_{100}^{45} \times 0.25^{45} \times 0.75^{55} = 6.67 \times 10^{-6}$$
2. Xác suất để trả lời đúng dưới 50 câu là
$P(X < 50) = C_{100}^{0} \times 0.25^{45} \times 0.75^{100} + C_{100}^{1} \times 0.25^{1} \times 0.75^{95} + \ldots + C_{100}^{49} \times 0.25^{45} \times 0.75^{51} = 0.9999999$
3. Số câu trả lời đúng trung bình là

Phân phối nhị thức

Example

Một thí sinh không học bài, chọn đáp án ngẫu nhiên tất cả các câu trong một đề gồm 100 câu hỏi trắc nghiệm (mỗi câu có 4 phương án lựa chọn). Mỗi lần chọn, khả năng đúng đều là 0.25. Khi đó gọi X là số câu trả lời đúng thì $X \sim B(100, 0.25)$.

Khi đó:

1. Xác suất đển trả lời đúng 45 câu hỏi là
$$P(X = 45) = C_{100}^{45} \times 0.25^{45} \times 0.75^{55} = 6.67 \times 10^{-6}$$
2. Xác suất để trả lời đúng dưới 50 câu là
$P(X < 50) = C_{100}^{0} \times 0.25^{45} \times 0.75^{100} + C_{100}^{1} \times 0.25^{1} \times 0.75^{95} + \ldots + C_{100}^{49} \times 0.25^{45} \times 0.75^{51} = 0.9999999$
3. Số câu trả lời đúng trung bình là $EX = 100 \times 0.25 = 25$ câu.

Giả sử $X \sim B(n, p)$

- $P(X = k) =$ **dbinom(k,n,p)**
- $P(X \leq k) =$ **pbinom(k,n,p)**
- $P(X > k) =$ **pbinom(k,n,p,F)**

Lưu ý rằng X mang các giá trị nguyên nên

$$P(X < k) = P(X \leq k - 1)$$

$$P(X \geq k) = P(X > k - 1)$$

Câu hỏi

Giả sử tỉ lệ qua môn XSTK của sinh viên TLU là 80%. Chọn ngẫu nhiên 10 thí sinh chuẩn bị thi môn này. Tính xác suất để:

1. *Tất cả các thí sinh đều thi qua môn.*
2. *Có hơn một nửa trong số trên qua môn.*
3. *Trung bình có bao nhiêu bạn qua?*

Câu hỏi

Một thí sinh không học bài, trả lời ngẫu nhiên, độc lập từng câu hỏi trong một bài thi quá trình trắc nghiệm gồm 25 câu hỏi với 4 phương án lựa chọn ở mỗi câu và chỉ có một phương án đúng. Với mỗi câu, nếu đúng, được tính 0.5 điểm, nếu sai, bị trừ 0.2 điểm. Gọi X là số câu đúng, Y là số điểm đạt được.

1. *Tính EY.*
2. *Tính xác suất để đạt điểm từ 4 trở lên.*
3. *Việc trả lời ngẫu nhiên trong mỗi câu hỏi là lợi hay thiệt? Vì sao? Nếu trong một câu hỏi bạn loại được một phương án thì có nên chọn ngẫu nhiên một trong các đáp án còn lại không? Vì sao?*

Sinh viên phải làm hai bài kiểm tra quá trình với cấu trúc và cách tính điểm như trên. Nếu một sinh sinh viên không học bài và trả lời ngẫu nhiên tất cả các câu hỏi thì xác suất để sinh viên đó được cả hai lần trên 4 điểm là bao nhiêu?

Phân phối chuẩn

Definition

Một biến ngẫu nhiên liên tục X gọi được gọi là có phân phối chuẩn nếu hàm mật độ $f(x)$ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

trong đó μ, σ là các hằng số.

Phân phối chuẩn

Definition

Một biến ngẫu nhiên liên tục X gọi được gọi là có phân phối chuẩn nếu hàm mật độ $f(x)$ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

trong đó μ, σ là các hằng số.

Có thể chứng minh được rằng khi đó $E(X) = \mu$, $V(X) = \sigma^2$.

Phân phối chuẩn

Definition

Một biến ngẫu nhiên liên tục X gọi được gọi là có phân phối chuẩn nếu hàm mật độ $f(x)$ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

trong đó μ, σ là các hằng số.

Có thể chứng minh được rằng khi đó $E(X) = \mu$, $V(X) = \sigma^2$.
Do vậy, biết $f(x) \Rightarrow$ kì vọng và phương sai của nó.

Phân phối chuẩn

Definition

Một biến ngẫu nhiên liên tục X gọi được gọi là có phân phối chuẩn nếu hàm mật độ $f(x)$ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

trong đó μ, σ là các hằng số.

Có thể chứng minh được rằng khi đó $E(X) = \mu$, $V(X) = \sigma^2$.
Do vậy, biết $f(x) \Rightarrow$ kì vọng và phương sai của nó.
Ngược lại, biết hai $\mu, \sigma \Rightarrow f(x)$.

Phân phối chuẩn

Definition

Một biến ngẫu nhiên liên tục X gọi được gọi là có phân phối chuẩn nếu hàm mật độ $f(x)$ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

trong đó μ, σ là các hằng số.

Có thể chứng minh được rằng khi đó $E(X) = \mu, V(X) = \sigma^2$.
Do vậy, biết $f(x) \Rightarrow$ kì vọng và phương sai của nó.
Ngược lại, biết hai $\mu, \sigma \Rightarrow f(x)$.
Do vậy, biến ngẫu nhiên X có phân phối chuẩn đặc trưng bởi 2 tham số: trung bình μ, phương sai σ^2.

Phân phối chuẩn

Definition

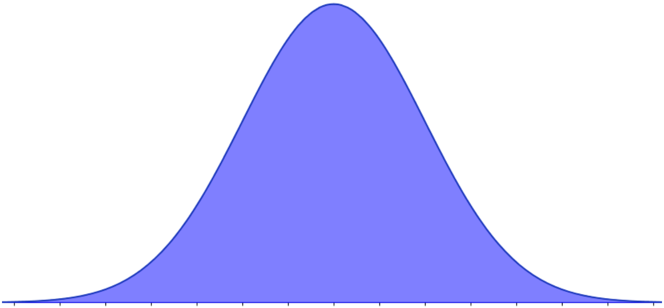
Một biến ngẫu nhiên liên tục X gọi được gọi là có phân phối chuẩn nếu hàm mật độ $f(x)$ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

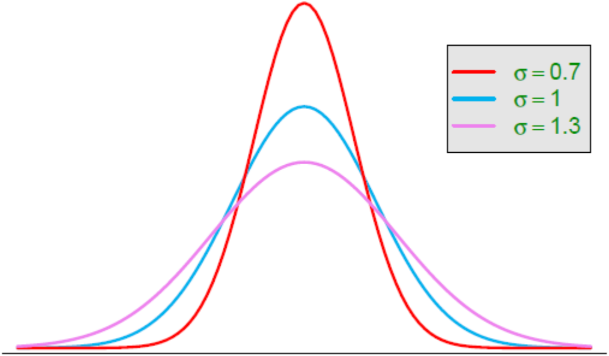
trong đó μ, σ là các hằng số.

Có thể chứng minh được rằng khi đó $E(X) = \mu$, $V(X) = \sigma^2$.
Do vậy, biết $f(x) \Rightarrow$ kì vọng và phương sai của nó.
Ngược lại, biết hai $\mu, \sigma \Rightarrow f(x)$.
Do vậy, biến ngẫu nhiên X có phân phối chuẩn đặc trưng bởi 2 tham số: trung bình μ, phương sai σ^2.
Để kí nói rằng X có phân phối chuẩn ta chỉ cần viết là $X \sim N(\mu, \sigma^2)$.

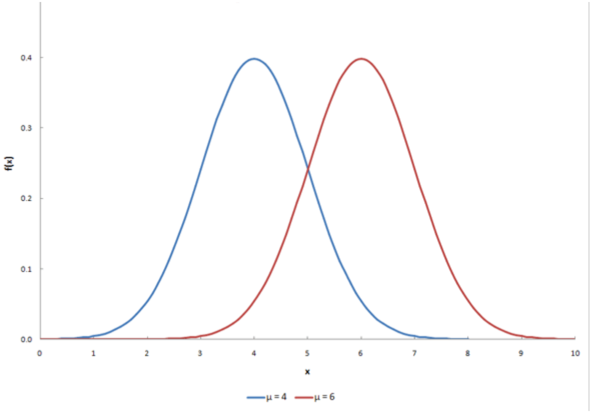
Hình dáng của phân phối chuẩn



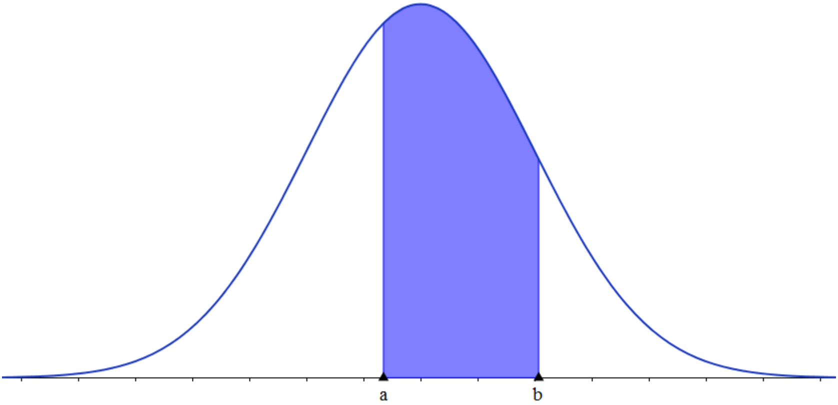
Hình dáng của phân phối chuẩn



Hình dáng của phân phối chuẩn



$P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = \ldots$



Phân phối chuẩn tắc

Definition

Khi trung bình $\mu = 0$, phương sai $\sigma^2 = 1$ thì ta nói X có phân phối chuẩn tắc.

Phân phối chuẩn tắc

Definition

Khi trung bình $\mu = 0$, phương sai $\sigma^2 = 1$ thì ta nói X có phân phối chuẩn tắc.

Theorem

Nếu X có phân phối $N(\mu, \sigma^2)$ thì biến ngẫu nhiên Z xác định bởi $Z = \dfrac{X - \mu}{\sigma}$ *có phân phối chuẩn tắc.*

Example

1. Chỉ số IQ của một người được chọn ngẫu nhiên là biến ngẫu nhiên tuân theo phân phối chuẩn (thường là $N(100, 15)$).

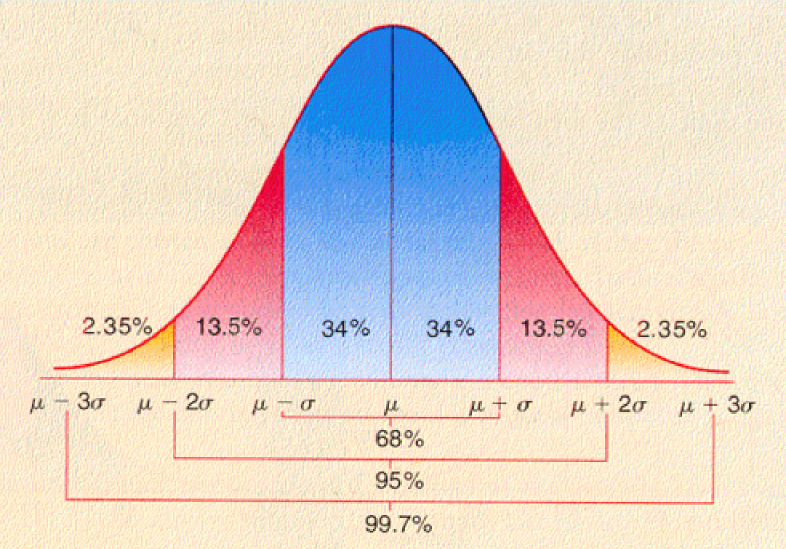
Example

1. Chỉ số IQ của một người được chọn ngẫu nhiên là biến ngẫu nhiên tuân theo phân phối chuẩn (thường là $N(100, 15)$).
2. Điểm thi Toefl của tổng thể những người thi tuân theo phân phối chuẩn.

Example

1. Chỉ số IQ của một người được chọn ngẫu nhiên là biến ngẫu nhiên tuân theo phân phối chuẩn (thường là $N(100, 15)$).
2. Điểm thi Toefl của tổng thể những người thi tuân theo phân phối chuẩn.
3. Trọng lượng của đàn cá (cùng một giống) trong một hồ tuân theo phân phối chuẩn.
4. . . .

Example



Phân phối chuẩn trong R

Trong R, ta có thể tính được xác suất của phân phối chuẩn $X \sim N(\mu, \sigma^2)$ trong một khoảng nào đó bằng các câu lệnh sau:

1. $P(X \leq a) = P(X < a) = pnorm(a, \mu, \sigma)$.
2. $P(X \geq a) = P(X > a) = pnorm(a, \mu, \sigma, FALSE)$.
3. $P(a \leq X \leq b) = P(b < X < a) = P(a \leq X < b) = P(a < X \leq b) =$

$$= pnorm(b, \mu, \sigma) - pnorm(a, \mu, \sigma)$$

Phân phối chuẩn trong R

Trong R, ta có thể tính được xác suất của phân phối chuẩn $X \sim N(\mu, \sigma^2)$ trong một khoảng nào đó bằng các câu lệnh sau:

1. $P(X \leq a) = P(X < a) = pnorm(a, \mu, \sigma)$.
2. $P(X \geq a) = P(X > a) = pnorm(a, \mu, \sigma, FALSE)$.
3. $P(a \leq X \leq b) = P(b < X < a) = P(a \leq X < b) = P(a < X \leq b) =$

$$= pnorm(b, \mu, \sigma) - pnorm(a, \mu, \sigma)$$

Bài toán ngược lại:

1. Ta muốn tìm x_0 để $P(X \leq x_0) = a \Rightarrow x_0 = qnorm(a, \mu, \sigma)$.
2. Nếu cần tìm x_1 để $P(X \geq x_1) = a \Rightarrow x_1 = qnorm(a, \mu, \sigma, FALSE)$.

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là $2kg$ độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là $2kg$ độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Solution

1. $P(1.5 \leq X \leq 2.5) = P(X \leq 2.5) - P(X < 1.5) =$

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là $2kg$ độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Solution

1. $P(1.5 \leq X \leq 2.5) = P(X \leq 2.5) - P(X < 1.5) = pnorm(2.5, 2, 0.5) - pnorm(1.5, 2, 0.5) = 0.6826895$

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là $2kg$ độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Solution

1. $P(1.5 \leq X \leq 2.5) = P(X \leq 2.5) - P(X < 1.5) = pnorm(2.5, 2, 0.5) - pnorm(1.5, 2, 0.5) = 0.6826895$
2. $P(X < 1) = pnorm(1, 2, 0.5)$

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là $2kg$ độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Solution

1. $P(1.5 \leq X \leq 2.5) = P(X \leq 2.5) - P(X < 1.5) = pnorm(2.5, 2, 0.5) - pnorm(1.5, 2, 0.5) = 0.6826895$
2. $P(X < 1) = pnorm(1, 2, 0.5) = 0.02275013$
3. $P(X \geq x_0) = 0.2$

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là $2kg$ độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Solution

1. $P(1.5 \leq X \leq 2.5) = P(X \leq 2.5) - P(X < 1.5) = pnorm(2.5, 2, 0.5) - pnorm(1.5, 2, 0.5) = 0.6826895$
2. $P(X < 1) = pnorm(1, 2, 0.5) = 0.02275013$
3. $P(X \geq x_0) = 0.2 \Rightarrow x_0 = qnorm(0.2, 2, 0.5, F)$

Example

Giả sử tại một hồ nuôi trọng lượng của tổng thể các con cá có phân phối chuẩn với trung bình là 2kg độ lệch chuẩn 0.5.

1. Xác suất để bắt được một con cá có trọng lượng từ 1.5 đến 2.5kg là bao nhiêu?
2. Tỉ lệ cá nhỏ hơn 1 kg là bao nhiêu?
3. Nếu ta muốn bắt 20% số lượng cá, mà phải là những con cá to nhất, thì ta phải bắt những con cá có trọng lượng từ bao nhiêu trở lên?

Solution

1. $P(1.5 \leq X \leq 2.5) = P(X \leq 2.5) - P(X < 1.5) =$ $pnorm(2.5, 2, 0.5) - pnorm(1.5, 2, 0.5) = 0.6826895$
2. $P(X < 1) = pnorm(1, 2, 0.5) = 0.02275013$
3. $P(X \geq x_0) = 0.2 \Rightarrow x_0 = qnorm(0.2, 2, 0.5, F) = 2.420811$

Câu hỏi (Bạn có bao nhiêu người bạn thân?)

Theo một khảo sát trên 1100 người lớn cho thấy phân phối số người bạn thân của mẫu này xấp xỉ phân phối chuẩn với trung bình là 9, độ lệch chuẩn là 2.5.

1. *Chọn một người trưởng thành ngẫu nhiên.*
   1. *Tính xấp xỉ xác suất để người đó có số bạn thân từ 10 trờ lên.*
   2. *Tính xấp xỉ xác suất để người đó có từ 7 đến 11 bạn thân.*
2. *Chọn một trong số 10 % những người có số bạn thân nhiều nhất. Anh ta phải có ít nhất bao nhiêu bạn thân trở lên?*
3. *Chọn 6 người trưởng thành ngẫu nhiên. Gọi X là số người trong số 6 người đó có ít nhất 10 bạn thân. X tuân theo phân phối gì? Tính xác suất để có 4 trong số 6 người trên có ít nhất 10 người bạn thân.*