CHƯƠNG 2.
ĐẠI LƯỢNG NGẪU NHIÊN

Ở chương 1, ta đã nghiên cứu các loại biến cố và phương pháp tính xác suất xảy ra của các biến cố đó. Nó cho phép ta chuyển sang nghiên cứu khái niệm trung tâm của lý thuyết xác suất, đó là khái niệm về biến ngẫu nhiên.

ĐẠI LƯỢNG NGẪU NHIÊN

Giả sử sau phép thử, đại lượng X sẽ nhận giá trị nào đó. Nếu giá trị của đại lượng X là ngẫu nhiên, không thể xác định được trước khi thực hiện phép thử thì ta gọi X là đại lượng ngẫu nhiên (biến ngẫu nhiên).

VÍ DỤ. Kiểm tra 3 sản phẩm.

X là số phế phẩm.

X là **biến ngẫu nhiên** vì trước khi thực hiện phép thử, ta không xác định được giá trị của X mà chỉ biết X có thể nhận giá trị: 0, 1, 2, 3.

PHÂN LOẠI

Biến ngẫu nhiên được gọi là rời rạc nếu các giá trị có thể có của nó lập nên một tập hợp hữu hạn hoặc vô hạn đếm được (gồm các giá trị rời nhau).

VÍ DỤ. **X là số người vào mua hàng tại một siêu thị trong một ngày.**

X = 0, 1, 2, ...

X là biến ngẫu nhiên rời rạc vì các giá trị có thể có của nó lập nên một tập hợp đếm được.

VÍ DỤ. Một phân xưởng có 5 máy hoạt động. Y là số máy hỏng trong một ca.

Y là biến ngẫu nhiên rời rạc với các giá trị có thể có là:

0, 1, 2, 3, 4, 5.

Biến ngẫu nhiên được gọi là liên tục nếu các giá trị có thể có của nó lấp đầy một khoảng trên trục số thực R.

Đối với biến ngẫu nhiên liên tục, ta không thể liệt kê được tất cả các giá trị có thể có của nó.

VÍ DỤ. Bắn một phát súng vào bia.

X là khoảng cách từ điểm chạm của viên đạn đến tâm bia.

X là biến ngẫu nhiên **liên tục**, vì ta không thể kể ra được tất cả các giá trị có thể có của nó. Ta chỉ có thể nói rằng các giá trị có thể có của X nằm trong khoảng (a, b) nào đó.

VÍ DỤ. **X là sai số khi đo lường một đại lượng vật lý. X là biến ngẫu nhiên liên tục.**

Y là kích thước của chi tiết do một máy sản xuất ra. Y là biến ngẫu nhiên liên tục.

Z là năng suất lúa vụ mùa của một tỉnh. Z là biến ngẫu nhiên liên tục.

Có thể nói rằng, rất nhiều các đại lượng mà ta gặp trong thực tế là các biến ngẫu nhiên và chúng sẽ là biến ngẫu nhiên rời rạc hoặc liên tục.

Các biến ngẫu nhiên thường được ký hiệu bởi X, Y, Z, ...

Các giá trị mà biến ngẫu nhiên nhận thường được ký hiệu bởi x, y, z, ...

QUY LUẬT PHÂN PHỐI XÁC SUẤT CỦA BIẾN NGẪU NHIÊN

Ta có thể nghĩ rằng chỉ cần xác định các giá trị có thể có của một biến ngẫu nhiên là đủ để xác định biến ngẫu nhiên ấy.

Tuy nhiên điều này chưa đủ.

Trong thực tế có những đại lượng rất khác nhau mà các giá trị có thể có của chúng lại giống nhau.

Hơn nữa việc các biến ngẫu nhiên nhận một giá trị nào đó trong kết quả của phép thử chỉ là một biến cố ngẫu nhiên, do đó nếu chỉ mới biết được các giá trị có thể có của nó thì ta mới nắm được rất ít thông tin về biến ngẫu nhiên ấy.

Vì vậy, ta còn phải xác định các xác suất tương ứng các giá trị có thể có của biến ngẫu nhiên để hoàn toàn xác định nó.

Quy luật phân phối xác suất của biến ngẫu nhiên là mối quan hệ giữa các giá trị có thể có của nó và các xác suất tương ứng với các giá trị đó.

Người ta sử dụng ba phương pháp để mô tả quy luật phân phối xác suất của biến ngẫu nhiên:

BẢNG PHÂN PHỐI XÁC SUẤT

HÀM MẬT ĐỘ XÁC SUẤT

HÀM PHÂN BỐ XÁC SUẤT

BẢNG PHÂN PHỐI XÁC SUẤT

Giả sử X là một biến ngẫu nhiên rời rạc nhận các giá trị x_1, x_2, ... với các xác suất tương ứng là p_1, p_2, ... Bảng phân phối xác suất của X có dạng:

X	x_1	x_2	...
Xác suất	p_1	p_2	...

$$\sum_i p_i = 1$$

VÍ DỤ. Một hộp có 12 viên ngọc trai, trong đó 8 viên màu trắng và 4 viên màu hồng. Lấy ngẫu nhiên không hoàn lại từ hộp ra 2 viên. X là số viên màu trắng trong 2 viên lấy ra. Lập bảng phân phối xác suất của X.

Hộp 12 viên: 8 viên màu trắng, 4 viên màu hồng.

Lấy ra 2 viên => n = C^2_{12}

X là số viên màu trắng trong 2 viên lấy ra.

X = 0, 1, 2 => X là biến ngẫu nhiên rời rạc.

P(X = 0) = P(2 viên hồng) = C^2_4 / C^2_{12}

P(X = 1) = P(1 viên trắng, 1 viên hồng) = $C^1_8 \, C^1_4 / C^2_{12}$

P(X = 2) = P(2 viên trắng) = C^2_8 / C^2_{12}

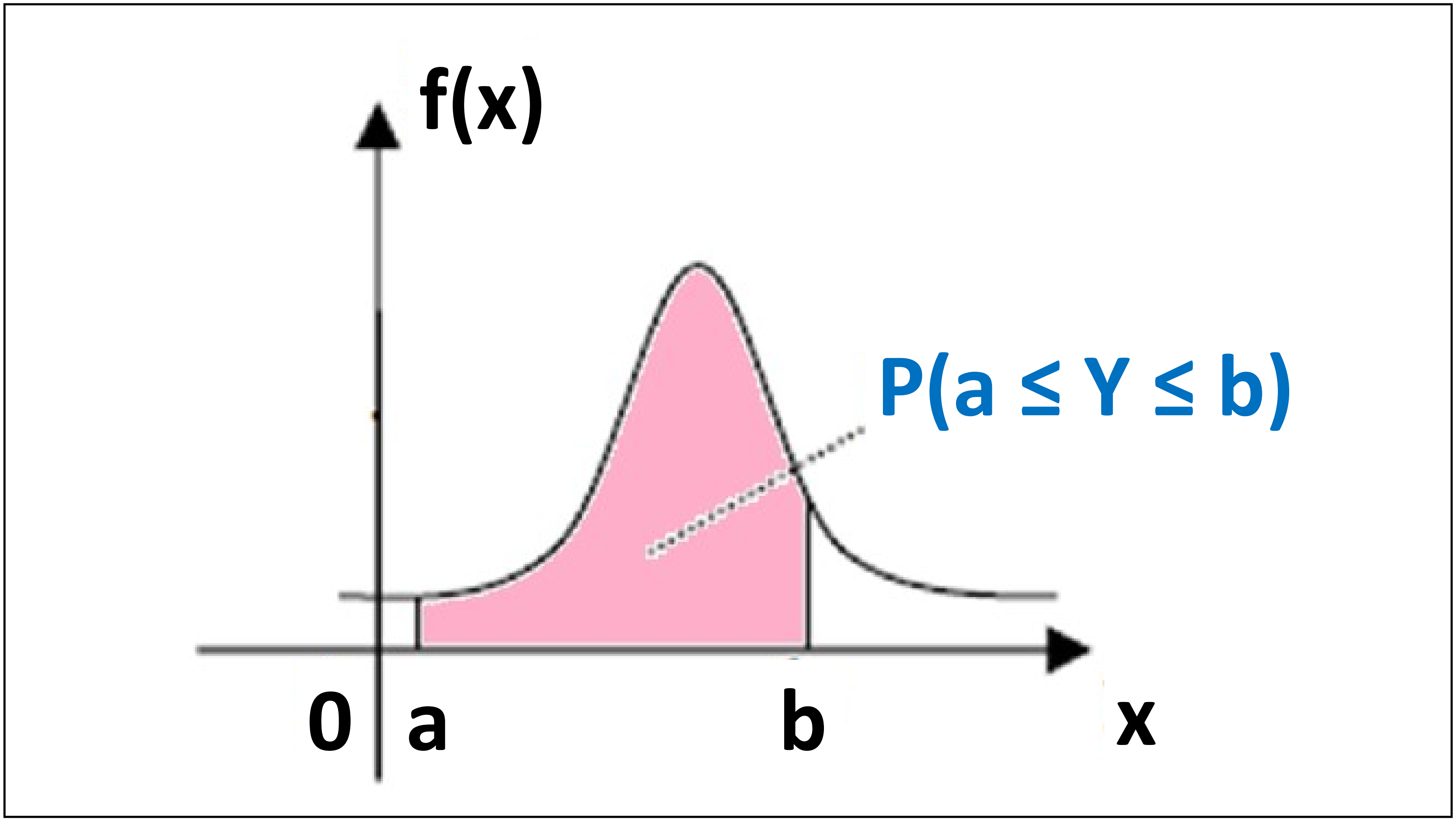
Bảng phân phối xác suất của X:

X	0	1	2
	6/66	32/66	28/66

HÀM MẬT ĐỘ XÁC SUẤT

Giả sử Y là một biến ngẫu nhiên liên tục. Hàm f(x) gọi là hàm mật độ xác suất của Y nếu:

$$P(a \leq Y \leq b) = \int_a^b f(x)\,dx$$


f(x)
P(a ≤ Y ≤ b)
0
a
b
x

TÍNH CHẤT.

$$f(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Ý NGHĨA. **Hàm mật độ xác suất của biến ngẫu nhiên liên tục Y trong mỗi khoảng (a, b) cho biết mức độ tập trung xác suất trong khoảng đó.**

HÀM PHÂN BỐ XÁC SUẤT

Hàm phân bố xác suất của biến ngẫu nhiên Y được định nghĩa là:

$$F(x) = P(Y < x)$$

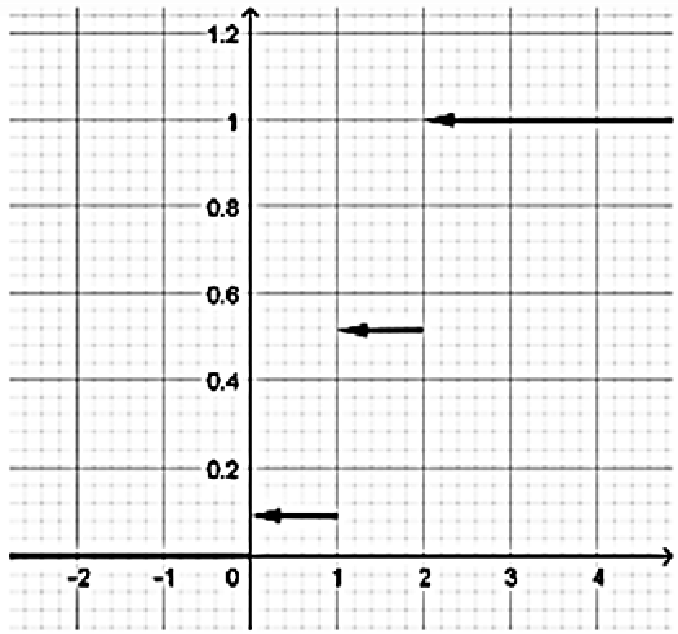
Hàm phân bố dùng để xác định qui luật phân phối xác suất cho cả biến ngẫu nhiên rời rạc và liên tục.

VÍ DỤ. Biến ngẫu nhiên Y có bảng phân phối xác suất:

Y	0	1	2
	6/66	32/66	28/66

Hàm phân bố xác suất của Y là:

$$F(x) = P(Y < x) = \begin{cases} 0 & \text{khi } x \le 0 \\ 6/66 & \text{khi } 0 < x \le 1 \\ 38/66 & \text{khi } 1 < x \le 2 \\ 1 & \text{khi } x > 2 \end{cases}$$



Đồ thị hàm phân bố xác suất của Y

CÁC THAM SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN

Như đã biết, quy luật phân phối xác suất của biến ngẫu nhiên (dưới dạng bảng phân phối xác suất, hàm mật độ xác suất hay hàm phân bố xác suất) hoàn toàn xác định biến ngẫu nhiên.

Như vậy, khi đã xác định được quy luật phân phối xác suất của một biến ngẫu nhiên thì ta đã nắm được toàn bộ thông tin về biến ngẫu nhiên đó.

Tuy nhiên, trong thực tế ta không chỉ cần đến những thông tin đó mà còn phải quan tâm đến những thông tin cô đọng phản ánh tổng hợp những đặc trưng quan trọng nhất của biến ngẫu nhiên được nghiên cứu.

Những thông tin cô đọng phản ánh từng phần về biến ngẫu nhiên được gọi là các tham số đặc trưng.

Các tham số đặc trưng của biến ngẫu nhiên được chia thành ba loại sau:

Các tham số đặc trưng cho xu hướng trung tâm của biến ngẫu nhiên (kỳ vọng, trung vị, giá trị tin chắc nhất, ...)

Các tham số đặc trưng cho độ phân tán của biến ngẫu nhiên (phương sai, độ lệch chuẩn, hệ số biến thiên, ...)

Các tham số đặc trưng cho dạng phân phối xác suất.

KỲ VỌNG
PHƯƠNG SAI
ĐỘ LỆCH CHUẨN
GIÁ TRỊ TIN CHẮC NHẤT

KỲ VỌNG

Sau khi biết phân phối xác suất của biến ngẫu nhiên X, người ta muốn biết giá trị trung bình mà biến ngẫu nhiên X nhận.

Nếu X là biến ngẫu nhiên rời rạc thì kỳ vọng được định nghĩa là:

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

Nếu Y là biến ngẫu nhiên liên tục với hàm mật độ f(x) thì kỳ vọng là:

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

VÍ DỤ. Nghiên cứu thu nhập của 500 công nhân ngành may.

Thu nhập (triệu đồng/tháng)	7	8	9	10	11	12	14
Số công nhân	50	70	150	120	55	30	25

Thu nhập trung bình của một công nhân:

(7 . 50 + 8 . 70 + 9 . 150 + 10 . 120 + 11 . 55 + 12 . 30 + 14 . 25) / 500 = 9,55

Gọi X là thu nhập của công nhân ngành may. X có phân phối xác suất:

X (triệu đồng/tháng)	7	8	9	10	11	12	14
	0,1	0,14	0,3	0,24	0,11	0,06	0,05

E(X) = 7 . 0,1 + 8 . 0,14 + 9 . 0,3 + 10 . 0,24 + 11 . 0,11 + 12 . 0,06 + 14 . 0,05

= 9,55

Khái niệm kỳ vọng lúc đầu xuất hiện trong các trò chơi may rủi để tính giá trị mà người chơi mong đợi sẽ nhận được.

Hiện nay khái niệm này được áp dụng rộng rãi trong nhiều lĩnh vực kinh doanh và quản lý như một tiêu chuẩn để ra quyết định trong tình huống cần lựa chọn giữa nhiều chiến lược khác nhau. Tiêu chuẩn này thường được biểu diễn dưới dạng lợi nhuận kỳ vọng hay doanh số kỳ vọng để làm căn cứ lựa chọn chiến lược kinh doanh.

TÍNH CHẤT.

$E(C) = C$

$E(CX) = C . E(X)$

$E(X \pm Y) = E(X) \pm E(Y)$

$E(XY) = E(X) E(Y)$ nếu X, Y độc lập.

PHƯƠNG SAI

Trong thực tế, nhiều khi chỉ xác định kỳ vọng toán của biến ngẫu nhiên thì chưa đủ để xác định biến ngẫu nhiên đó. Ta còn phải xác định mức độ phân tán của các giá trị của biến ngẫu nhiên xung quanh giá trị trung bình của nó nữa.

Chẳng hạn, khi nghiên cứu biến ngẫu nhiên là năng suất lúa của một địa phương nào đó thì năng suất lúa trung bình (kỳ vọng) mới chỉ phản ánh được một khía cạnh của đại lượng đó mà thôi. Mức độ biến động về năng suất của các thửa ruộng khác nhau xung quanh giá trị trung bình cũng là một khía cạnh quan trọng cần nghiên cứu.

Phương sai của biến ngẫu nhiên X được định nghĩa là:

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

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

Nếu X là biến ngẫu nhiên rời rạc thì phương sai D(X) bằng:

$$(x_1^2 p_1 + x_2^2 p_2 + \ldots) - [E(X)]^2$$

Nếu Y là biến ngẫu nhiên liên tục với hàm mật độ f(x) thì phương sai bằng:

$$D(Y) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx - [E(Y)]^2$$

VÍ DỤ. X là thu nhập của công nhân ngành may.

X (triệu đồng/tháng)	7	8	9	10	11	12	14
	0,1	0,14	0,3	0,24	0,11	0,06	0,05

$E(X) = 9{,}55$

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

$= (7^2 . 0{,}1 + 8^2 . 0{,}14 + 9^2 . 0{,}3 + 10^2 . 0{,}24 + 11^2 . 0{,}11 + 12^2 . 0{,}06$

$+ 14^2 . 0{,}05) - [9{,}55^2]$

$= 2{,}7075$

Cùng với kỳ vọng, phương sai có những ứng dụng to lớn trong nhiều lĩnh vực thực tiễn. Nếu như trong kỹ thuật, phương sai đặc trưng cho mức độ phân tán của các chi tiết gia công hay sai số của thiết bị thì trong quản lý và kinh doanh, nó đặc trưng cho mức độ rủi ro của các quyết định.

TÍNH CHẤT.

$D(C) = 0$

$D(CX) = C^2 . D(X)$

$D(X \pm Y) = D(X) + D(Y)$ nếu X, Y độc lập.

ĐỘ LỆCH CHUẨN

Độ lệch chuẩn của biến ngẫu nhiên X được định nghĩa là:

$$\sigma(X) = \sqrt{D(X)}$$

VÍ DỤ. X là thu nhập của công nhân ngành may.

X (triệu đồng/tháng)	7	8	9	10	11	12	14
	0,1	0,14	0,3	0,24	0,11	0,06	0,05

E(X) = 9,55

D(X) = 2,7075

σ(X) ≈ 1,6454

Độ lệch chuẩn σ(X) cũng biểu thị mức độ phân tán các giá trị của biến ngẫu nhiên X xung quanh giá trị trung bình E(X) nhưng có điểm thuận tiện hơn phương sai là độ lệch chuẩn σ(X) có cùng đơn vị đo với X và E(X).

VÍ DỤ.

Biến ngẫu nhiên X (kg).

Kỳ vọng E(X) (kg).

Phương sai D(X) (kg^2).

Độ lệch chuẩn σ(X) (kg).

GIÁ TRỊ TIN CHẮC NHẤT

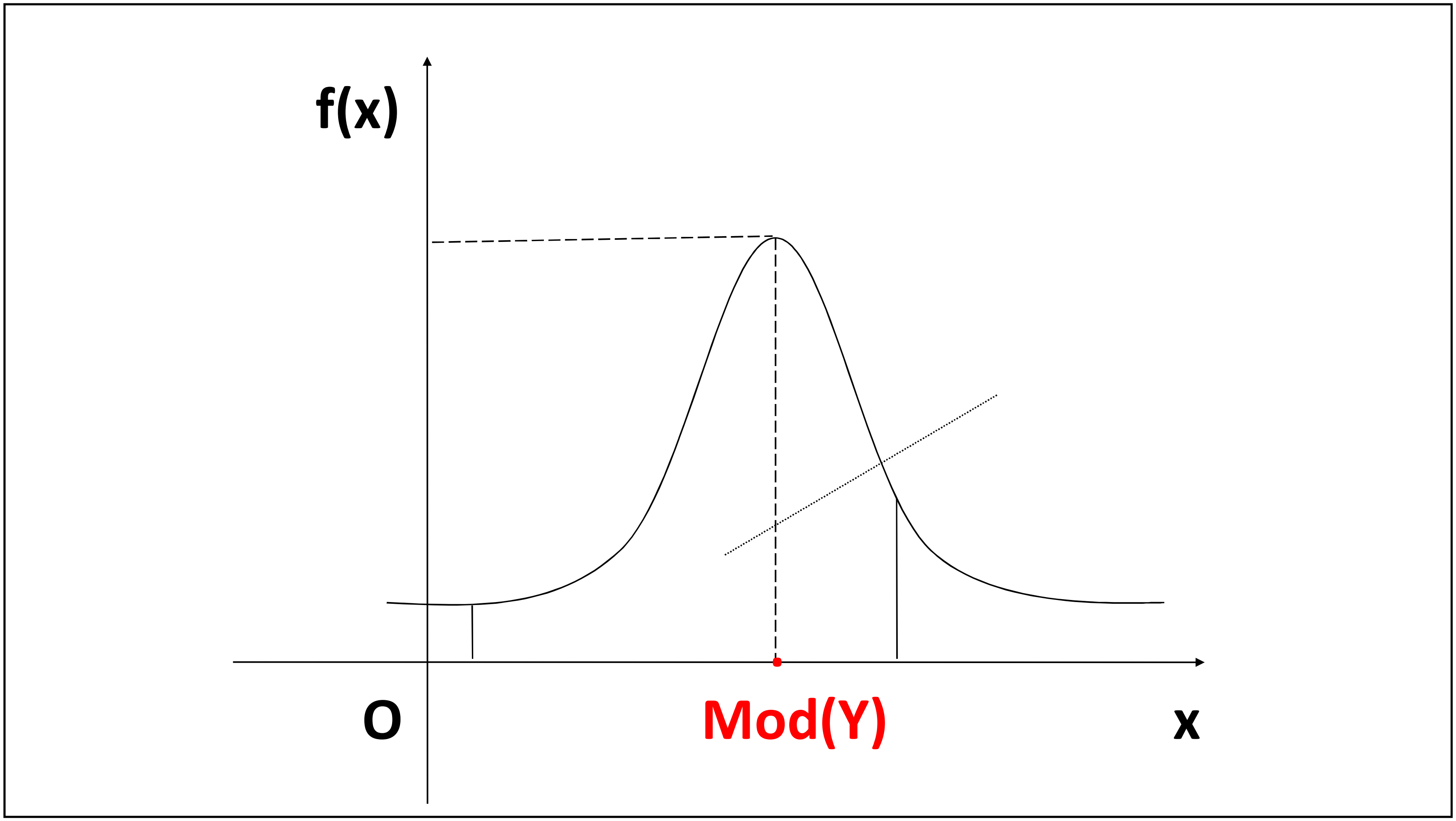
Nếu X là biến ngẫu nhiên rời rạc thì giá trị tin chắc nhất là giá trị X nhận với xác suất lớn nhất trong bảng phân phối xác suất. Ký hiệu: Mod (X)

VÍ DỤ. X là thu nhập của công nhân ngành may.

X (triệu đồng/tháng)	7	8	9	10	11	12	14
	0,1	0,14	0,3	0,24	0,11	0,06	0,05

Mod (X) = 9

Nếu Y là biến ngẫu nhiên liên tục có hàm mật độ f(x) thì giá trị tin chắc nhất là giá trị Y nhận mà tại đó hàm mật độ f(x) đạt giá trị lớn nhất.


f(x)
O
Mod(Y)
x

NHẬN XÉT.

Mod (X) chính là giá trị có khả năng nhiều nhất trong các giá trị mà X có thể nhận.

MỘT SỐ QUY LUẬT PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

PHÂN PHỐI NHỊ THỨC
PHÂN PHỐI ĐA THỨC
PHÂN PHỐI POISSON
PHÂN PHỐI SIÊU BỘI
PHÂN PHỐI HÌNH HỌC
PHÂN PHỐI CHUẨN
PHÂN PHỐI ĐỀU
PHÂN PHỐI MŨ
PHÂN PHỐI χ^2
PHÂN PHỐI STUDENT
PHÂN PHỐI FISHER – SNEDECOR

PHÂN PHỐI NHỊ THỨC

Tiến hành n phép thử độc lập.

Trong mỗi phép thử, xác suất xảy ra biến cố A là P(A) = p.

(Dãy phép thử Bernoulli).

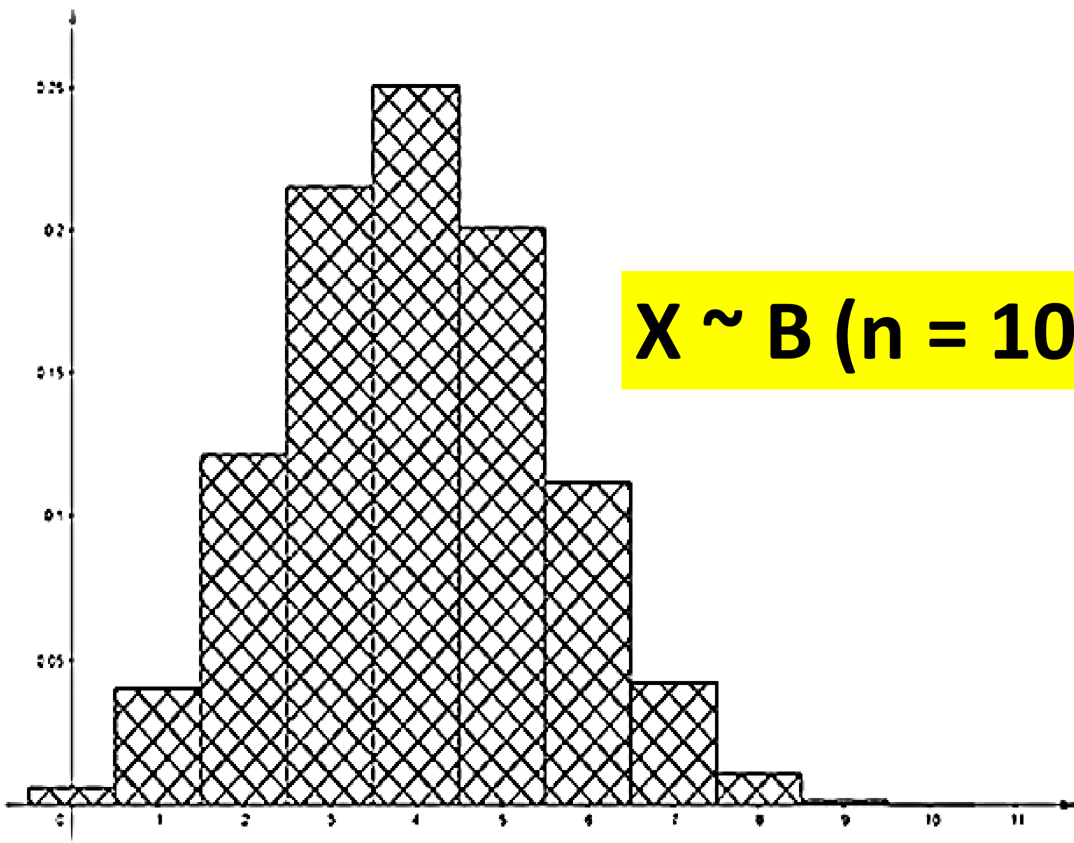
X là số lần biến cố A xảy ra trong dãy phép thử.

X = 0, 1, …, n => X là biến ngẫu nhiên rời rạc.

CÔNG THỨC BERNOULLI: $P(X = k) = C^k_n \, p^k (1 - p)^{n-k}$

X được gọi là có phân phối nhị thức với tham số n và p.

Ký hiệu: $X \sim B(n, p)$



$X \sim B\ (n = 10,\ p = 0{,}4)$

TÍNH CHẤT.

Nếu $X \sim B(n, p)$ thì:

$$E(X) = np$$

$$D(X) = np(1 - p)$$

$$np + p - 1 \leq Mod(X) \leq np + p$$

VÍ DỤ. **Một máy sản xuất 200 sản phẩm trong một ngày. Xác suất để máy sản xuất ra phế phẩm là 0,053.**

a) Tính xác suất có 5 phế phẩm trong một ngày.

b) Tìm số phế phẩm trung bình và số phế phẩm tin chắc nhất của máy đó trong một ngày?

c) Biết chi phí sản xuất một sản phẩm là 1 USD, giá bán một chính phẩm là 2 USD và phế phẩm bị loại bỏ. Trung bình lợi nhuận từ một máy trong một ngày là bao nhiêu?

a) Kiểm tra 200 sản phẩm là 200 phép thử Bernoulli với xác suất gặp phế phẩm trong mỗi phép thử là $p = 0,053$.

X là số phế phẩm trong một ngày.

$X = 0, 1, \ldots, 200$.

$X \sim B(n, p)$ với $n = 200$ và $p = 0,053$.

Xác suất có 5 phế phẩm trong một ngày:

$$P(X = 5) = C^k_n \, p^k \, (1 - p)^{n-k}$$

$$= C^5_{200} \, 0,053^5 \, (1 - 0,053)^{195} \approx 0,0259$$

b) Số phế phẩm trung bình trong một ngày:

$E(X) = np = 200 \,.\, 0{,}053 = 10{,}6$

Số phế phẩm tin chắc nhất trong một ngày:

$np + p - 1 \leq Mod\,(X) \leq np + p$

$9{,}653 \leq Mod\,(X) \leq 10{,}653$

$Mod\,(X) = 10.$

c) Trung bình lợi nhuận từ một máy trong một ngày:

E (lợi nhuận từ một máy trong một ngày)

$= E\,[(200 - X)\,.\,2 - 200]$

$= E\,(200 - 2X)$

$= 200 - 2\,.\,E(X)$

$= 200 - 2\,.\,10{,}6$

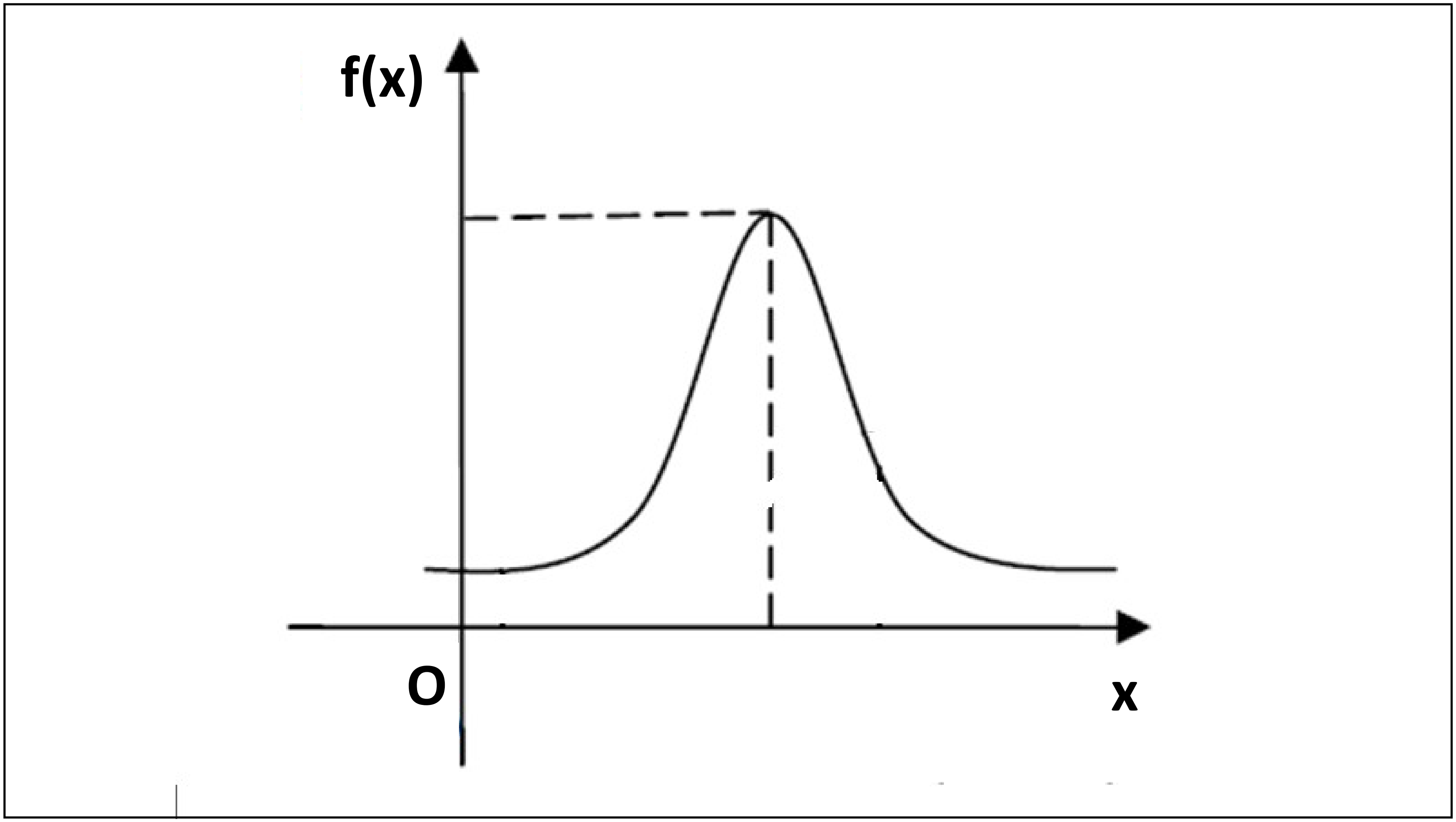
$= 178{,}8$ (USD)

PHÂN PHỐI CHUẨN

Biến ngẫu nhiên liên tục Y nhận giá trị trong R được gọi là biến ngẫu nhiên có phân phối chuẩn nếu hàm mật độ của nó có dạng:

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

Ký hiệu: Y ~ N (μ, σ²)


f(x)
O
x

HÀM MẬT ĐỘ PHÂN PHỐI CHUẨN f(x):

Có dạng hình chuông.

Nhận trục hoành làm tiệm cận ngang.

Đạt cực đại tại một điểm.

Có một trục đối xứng (đường thẳng đứng đi qua điểm cực đại).

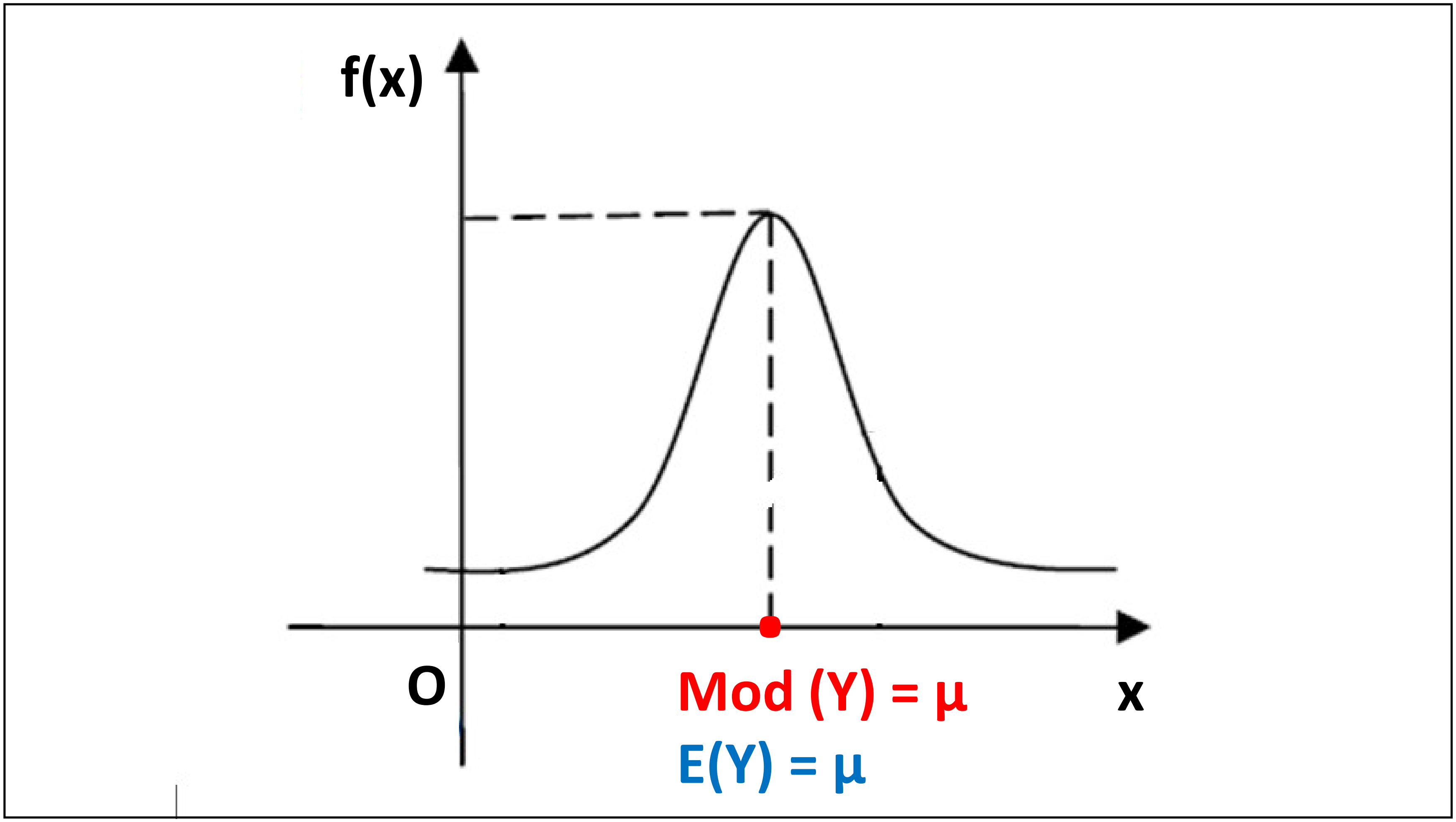
TÍNH CHẤT.

Nếu $Y \sim N(\mu, \sigma^2)$ thì:

$$E(Y) = \mu$$

$$D(Y) = \sigma^2$$

$$Mod\,(Y) = \mu$$



f(x)
O
Mod (Y) = μ
E(Y) = μ
x

ỨNG DỤNG CỦA PHÂN PHỐI CHUẨN

Phân phối chuẩn do nhà toán học Gauss tìm ra năm 1809 nên gọi là phân phối Gauss.

Quy luật phân phối chuẩn là quy luật phân phối xác suất được áp dụng rất rộng rãi trong thực tế. Trong nhiều lĩnh vực của khoa học và đời sống, ta đều gặp các biến ngẫu nhiên phân phối chuẩn.

Chẳng hạn trong công nghiệp, người ta đã xác định được rằng kích thước của các chi tiết do các nhà máy sản xuất ra sẽ phân phối chuẩn nếu quá trình sản xuất diễn ra bình thường.

Trong nông nghiệp, năng suất của cùng một loại cây trồng tại các thửa ruộng khác nhau cũng phân phối chuẩn.

Năng suất lao động của các công nhân có cùng tay nghề và làm cùng một công việc như nhau cũng phân phối chuẩn.

Nhu cầu về các loại hàng hóa khác nhau cũng phân phối chuẩn…

Người ta ghi nhận rằng các năng lực về trí tuệ và thể lực con người cũng phân phối theo quy luật chuẩn.

Thậm chí cả một số chỉ tiêu về sinh lý của những người cùng giới (chẳng hạn chiều cao, vòng ngực, chiều dài cánh tay…) cũng phân phối theo quy luật chuẩn.

Sự nhận biết này cho phép lập kế hoạch sản xuất quần áo may sẵn, sản xuất hàng loạt sao cho đáp ứng một cách hợp lý nhất kích cỡ của người mua, tránh tình trạng thừa thiếu do không vừa kích cỡ.

$Y \sim N(\mu, \sigma^2)$

Ta cần tính xác suất:

$$P(a \le Y \le b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Hai tham số μ và σ có ý nghĩa rất quan trọng trong phân phối chuẩn. Khi μ và σ thay đổi, hàm mật độ xác suất f(x) cũng thay đổi.

Khi μ thay đổi thì dạng của đường cong f(x) không thay đổi song nó sẽ chuyển dịch sang phải hoặc sang trái theo trục Ox. Khi μ tăng lên, đồ thị sẽ dịch sang phải. Còn khi μ giảm, đồ thị sẽ dịch sang trái.

Khi σ thay đổi thì dạng của đồ thị sẽ thay đổi theo. Nếu σ tăng lên thì đồ thị sẽ thấp xuống và phình ra, còn khi σ giảm đi thì đồ thị sẽ cao lên và nhọn thêm.

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

Hàm f(x) đơn giản nhất (đặc biệt nhất) khi $\mu = 0$ và $\sigma = 1$.

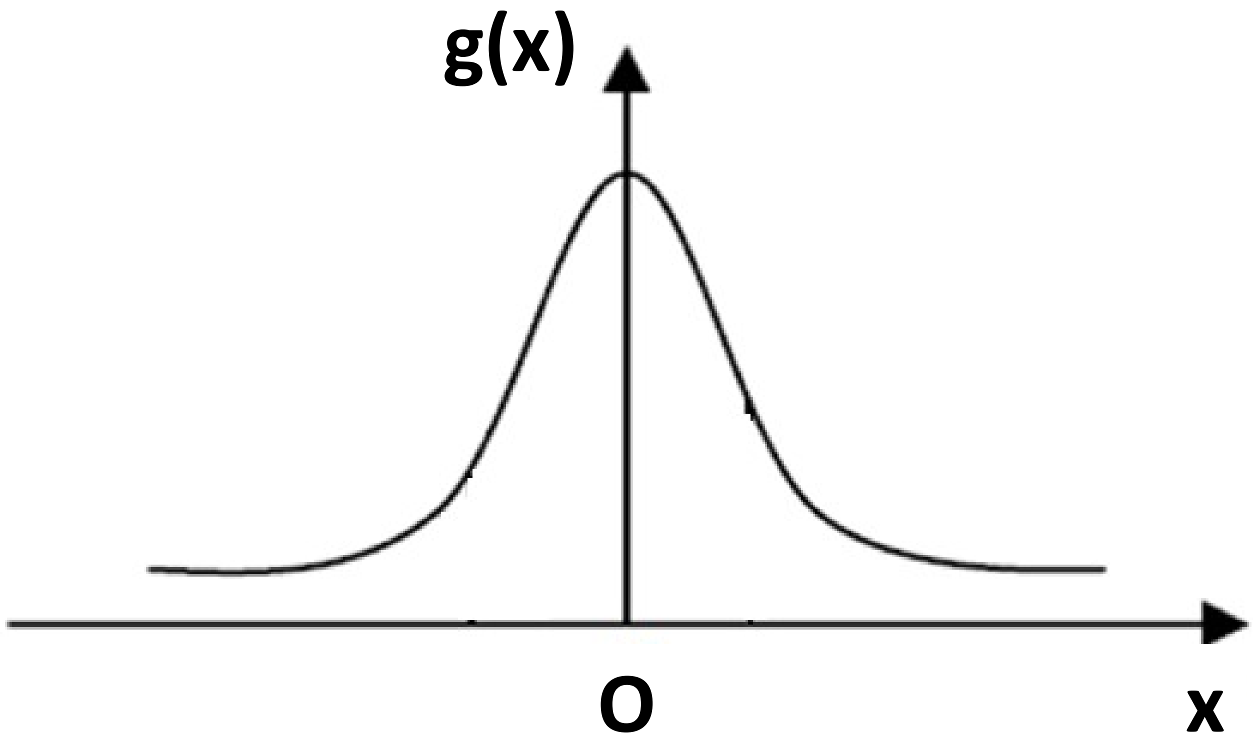
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

PHÂN PHỐI CHUẨN TẮC

Biến ngẫu nhiên liên tục Z nhận giá trị trong R được gọi là biến ngẫu nhiên có phân phối chuẩn tắc nếu hàm mật độ của nó có dạng:

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

Ký hiệu: Z ~ N (0, 1)



Hàm mật độ phân phối chuẩn tắc

TÍNH CHẤT.

Nếu $Z \sim N(0, 1)$ thì:

$$E(Z) = 0$$

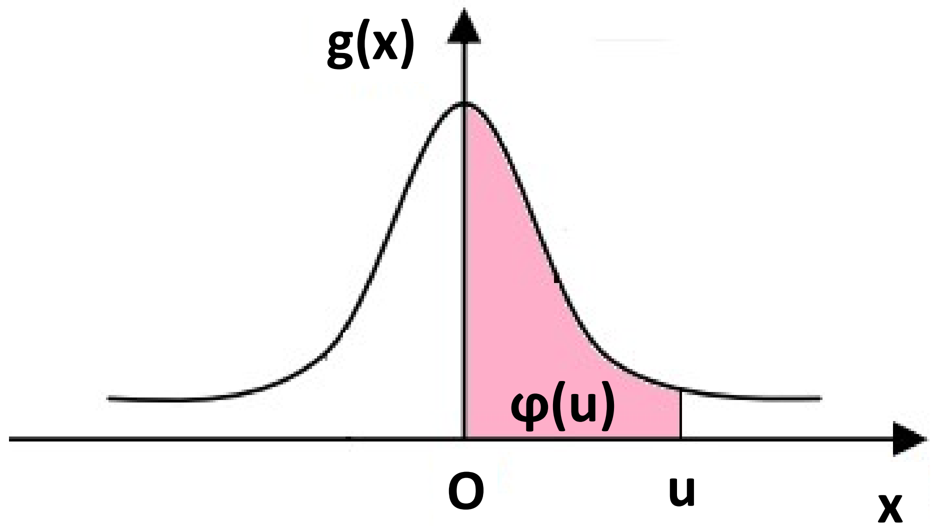
$$D(Z) = 1$$

$$Mod(Z) = 0$$

BẢNG 1. GIÁ TRỊ TÍCH PHÂN HÀM MẬT ĐỘ CHUẨN TẮC: $\varphi(u) = \int_0^u \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
4.0	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

$$\varphi(u) = \int_0^u \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx$$



CHÚ Ý.

$$\varphi(-x) = -\varphi(x)$$

$$\varphi(+\infty) = 0,5$$

$$\varphi(-\infty) = -0,5$$

Từ X phân phối N (μ, σ^2) suy ra

$$Z = \frac{X - \mu}{\sigma}$$ **phân phối N (0, 1).**

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

NHẬN XÉT.

$$P(X > a) = P(a < X < +\infty) = \varphi(+\infty) - \varphi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(X < b) = P(-\infty < X < b) = \varphi\left(\frac{b-\mu}{\sigma}\right) - \varphi(-\infty)$$

CHÚ Ý. **Biến ngẫu nhiên X liên tục thì:** $P(X = a) = P(X = b) = 0$.

Do đó

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

VÍ DỤ. **Trọng lượng của một loại sản phẩm là biến ngẫu nhiên X (kg) có phân phối chuẩn N (10, 0,25).**

a) Tìm tỉ lệ sản phẩm có trọng lượng từ 9,5 kg đến 11 kg?

b) Tìm tỉ lệ sản phẩm có trọng lượng nặng hơn 10,5 kg?

c) Tìm tỉ lệ sản phẩm có trọng lượng nhẹ hơn 9 kg?

a) X ~ N (μ = 10, σ^2 = 0,25) => σ = 0,5

Tỉ lệ sản phẩm trọng lượng 9,5 kg đến 11 kg:

$$P(9,5 \le X \le 11) = \varphi\left(\frac{11-10}{0,5}\right) - \varphi\left(\frac{9,5-10}{0,5}\right)$$

$$= \varphi(2) - \varphi(-1)$$

$$= 0,4772 - (-0,3413)$$

$$= 0,8185$$

b) Tỉ lệ sản phẩm nặng hơn 10,5 kg:

$$P(X > 10,5) = \varphi(+\infty) - \varphi\left(\frac{10,5-10}{0,5}\right)$$

$$= \varphi(+\infty) - \varphi(1)$$

$$= 0,5 - 0,3413$$

$$= 0,1587$$

c) Tỉ lệ sản phẩm nhẹ hơn 9 kg:

$$P(X < 9) = \varphi\left(\frac{9-10}{0,5}\right) - \varphi(-\infty)$$

$$= \varphi(-2) - \varphi(-\infty)$$

$$= (-0,4772) - (-0,5)$$

$$= 0,0228$$

QUAN HỆ GIỮA PHÂN PHỐI NHỊ THỨC VÀ PHÂN PHỐI CHUẨN

Biến ngẫu nhiên X có phân phối nhị thức B(n, p).

Nếu n lớn, p không quá gần 0 và 1 thì khi đó:

X có phân phối xấp xỉ chuẩn N (μ, σ^2)

với $\mu \approx np$, $\sigma^2 \approx npq$.

$$P(a \le X \le b) \approx \varphi\left(\frac{b-\mu}{\sigma}\right) - \varphi\left(\frac{a-\mu}{\sigma}\right)$$

Khi sử dụng quy luật nhị thức, nếu n khá lớn thì việc tính toán theo công thức Bernoulli sẽ gặp khó khăn. Lúc đó nếu p nhỏ đến mức np ≈ npq thì có thể dùng quy luật Poisson thay thế cho quy luật nhị thức. Song nếu p lại không nhỏ (p > 0,1) thì không thể dùng quy luật Poisson để thay thế được. Lúc đó có thể dùng quy luật chuẩn để thay thế cho quy luật nhị thức.

VÍ DỤ. **Xác suất sinh được em bé trai là 0,48. Tính xác suất sao cho trong 300 trường hợp sắp sinh:**

a) Số bé trai vào khoảng từ 150 đến 170?

b) Số bé trai từ 170 trở lên?

a) Kiểm tra 300 em bé sắp sinh là 300 phép thử Bernoulli.

X là số bé trai trong 300 bé sắp sinh. X = 0, 1, 2, …, 300.

$X \sim B(n, p)$ với n = 300 và p = 0,48.

Xác suất số bé trai vào khoảng từ 150 đến 170 là:

$$P(150 \le X \le 170) = P(X = 150) + P(X = 151) + \dots + P(X = 170)$$

$$= (C^{150}_{300} \,.\, 0{,}48^{150} \,.\, 0{,}52^{150}) + \dots + (C^{170}_{300} \,.\, 0{,}48^{170} \,.\, 0{,}52^{130})$$

Ta thấy n = 300 khá lớn, p = 0,48 không quá gần 0 và 1.

Do đó, có thể coi X có phân phối xấp xỉ chuẩn $N(\mu, \sigma^2)$

với $\mu \approx np = 300 \,.\, 0{,}48 = 144$

$\sigma^2 \approx npq = 300 \,.\, 0{,}48 \,.\, 0{,}52 \Rightarrow \sigma \approx 8{,}6533$

$$P(150 \leq X \leq 170) \approx \varphi\left(\frac{170-144}{8,6533}\right) - \varphi\left(\frac{150-144}{8,6533}\right)$$

$$\approx \varphi(3,00) - \varphi(0,69)$$

$$= 0,4987 - 0,2549$$

$$= 0,2438$$

b) Xác suất số bé trai từ 170 trở lên là:

$P(X \geq 170) = P(X = 170) + P(X = 171) + \ldots + P(X = 300)$

$= (C^{170}_{300} \cdot 0{,}48^{170} \cdot 0{,}52^{130}) + \ldots + (0{,}48^{300})$

Xấp xỉ sang phân phối chuẩn:

$$P(X \geq 170) \approx \varphi(+\infty) - \varphi\left(\frac{170 - 144}{8{,}6533}\right)$$

$$\approx \varphi(+\infty) - \varphi(3{,}00)$$

$$= 0{,}5 - 0{,}4987$$

$$= 0{,}0013$$