Escola Superior de Agricultura “Luiz de Queiroz”
Universidade de São Paulo

Variáveis Aleatórias Contínuas

Professora Renata Alcarde Sermarini

Piracicaba
Abril 2016

Distribuição normal

- Um dos modelos mais importantes de uma distribuição contínua de probabilidade;
- Representa, com boa aproximação, muitos fenômenos da natureza;
- Alguns exemplos de variáveis aleatórias contínuas que seguem distribuição normal (geralmente):
  - Peso: de matéria seca, de raiz, de animais, de pessoas, de frutos, de sacas de café,...
  - Altura: de árvores, plantas, animais;
  - DAP;
  - Produtividade: de cana-de-açúcar, de soja,...
  - Erros de medida em geral.

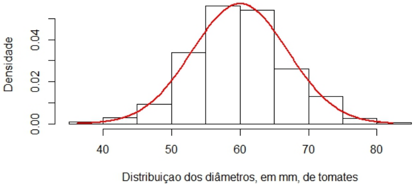
Distribuição normal

Um problema:

Foi criado no estado de São Paulo um programa para melhoria dos padrões comerciais e embalagens de hortifrutigranjeiros.

- Como parte desse programa pretende-se estabelecer um sistema de classificação para tomates oblongos quanto ao diâmetro.
- Foi feito, então, um levantamento amostral envolvendo diferentes variedades, propriedades, cidades e épocas, observando-se um calibre médio de tomates de 60 mm, variância de 49 mm^2 e distribuição conforme a figura a seguir.

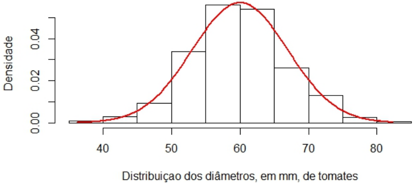
Distribuição normal



- Sistema I:

Classificação	Diâmetro	Porcentagem esperada
Pequeno	até 50 mm	____%
Médio	De 50 a 60 mm	____%
Grande	acima de 60 mm	____%

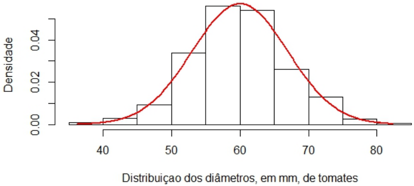
Distribuição normal



- Sistema II:

Classificação	Diâmetro	Porcentagem esperada
Pequeno	até ___ mm	20%
Médio	De ___ a ___ mm	60%
Grande	acima de ___ mm	20%

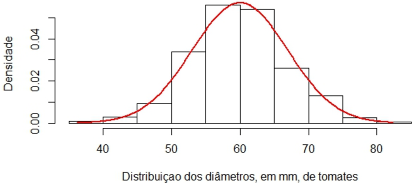
Distribuição normal



Observações:

- As observações estão mais concentradas em torno do valor central e a concentração vai diminuindo a medida que os valores vão aumentando ou diminuindo;
- Distribuição em forma de sino;
- Distribuição simétrica em torno do seu ponto central;

Distribuição normal



Observações:

- As distribuições amostrais de estatísticas como médias e proporções podem ser aproximadas pela distribuição normal ⇒ Inferência estatística
- Distribuições binomial e Poisson ⇒ aproximação através da distribuição normal
- Denominação: distribuição gaussinana ⇒ Karl F. Gauss (1777-1855).

Distribuição normal

Definição

Dizemos que uma variável aleatória X tem distribuição normal, com parâmetros μ e σ, em que $-\infty < \mu < \infty$ e $\sigma > 0$, se sua função densidade de probabilidade for dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

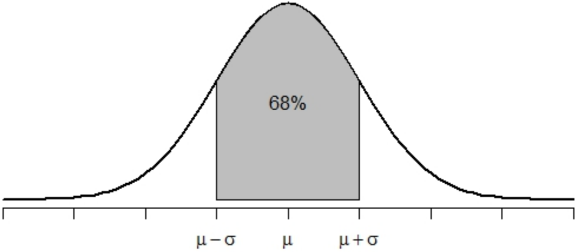
Notação: $X \sim N(\mu, \sigma^2)$.

Pode-se demonstrar que:

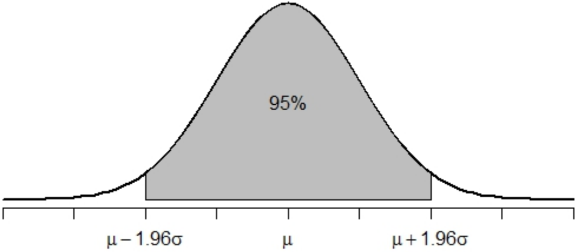
- $f_X(x) > 0$
- $\int_{-\infty}^{+\infty} f_X(x)dx = 1$
- $\mathrm{E}(X) = \mu$
- $\mathrm{Var}(X) = \sigma^2$
- $\mathrm{DP}(X) = \sigma$

- $f_X(x)$ é simétrica ao redor de μ, ou seja, $f(\mu - x) = f(\mu + x)$ para todo x

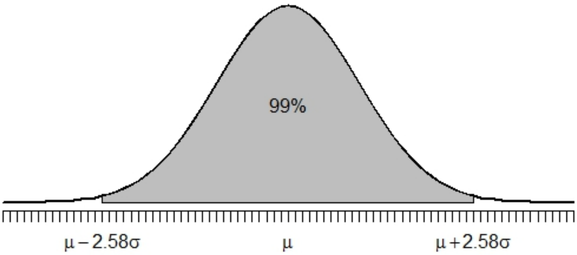
Distribuição normal



Distribuição normal

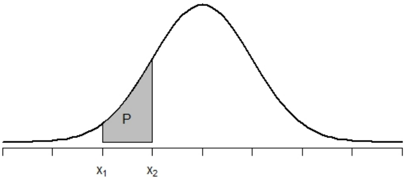


Distribuição normal



Distribuição normal

A probabilidade de uma variável aleatória com distribuição normal tomar um valor entre dois pontos quaisquer, x_1 e x_2, tal que $x_1 < x_2$, é igual a área sob a curva normal compreendida entre os dois pontos.



$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Distribuição normal

Sabendo-se que uma variável $X =$ diâmetro, em mm, de um tomate tem distribuição $N(60, 49)$, calcular:

(a) $P(X < 50)$

(b) $P(50 < X < 60)$

(c) $P(X > 60)$

Cálculo da integral $\Rightarrow$ métodos numéricos

$$\Downarrow$$

Distribuição normal padrão

Distribuição normal

Distribuição normal padrão

Se X uma variável aleatória com distribuição $N(\mu, \sigma^2)$, então a variável aleatória Z, definida por:

$$Z = \frac{X - \mu}{\sigma},$$

tem uma distribuição $N(0, 1)$, cuja função densidade de probabilidade é dada por:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad z \in \mathbb{R}.$$

Observações:

- A nova distribuição tem média correspondente a origem e desvio padrão como medida de afastamento da média;
- $\mathrm{E}(Z) = \mu_Z = 0$ e $\mathrm{Var}(Z) = \sigma_Z^2 = 1$;
- Os valores correspondentes a $P(x_1 < X < x_2) = P(z_1 < Z < z_2)$ estão descritos em uma única tabela.

Distribuição normal

Exercício: Sabendo-se que $Z \sim N(0,1)$, usando a tabela da distribuição normal padrão, calcular:

(a) $P(0 < Z < 2,14)$

(b) $P(-3,01 < Z < 0)$

(c) $P(-3,01 < Z < 2,14)$

(d) $P(Z > 0)$

(e) $P(Z > 1,00)$

(f) $P(Z < -1,00)$

Distribuição normal

Agora podemos calcular as probabilidades associadas aos intervalos correspondentes a variável X = diâmetro, em *mm*, de um tomate tem distribuição $N(60, 49)$.

(a) $P(X < 50)$

(b) $P(50 < X < 60)$

(c) $P(X > 60)$

Assim, as porcentagens esperadas são dadas por:

Classificação	Diâmetro	Porcentagem esperada
Pequeno	até 50 *mm*	___%
Médio	De 50 a 60 *mm*	___%
Grande	acima de 60 *mm*	___%

Distribuição normal

Exercício: Calcular os valores de X correspondentes às porcentagens esperadas, em que $X =$ diâmetro, em *mm*, de um tomate e tem distribuição $N(60, 49)$.

Classificação	Diâmetro	Porcentagem esperada
Pequeno	até ___ *mm*	20%
Médio	De ___ a ___ *mm*	60%
Grande	acima de ___ *mm*	20%

Distribuição normal

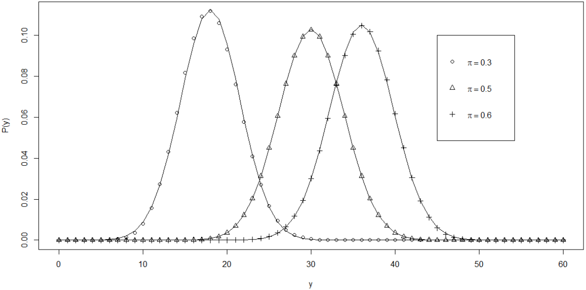
Exercício: O comprimento X, em cm, de *Litopenaeus schmitti* (camarão marinho), em condições normais na Lagoa do Ibiraquera, tem distribuição aproximadamente normal, com média de 6,0 cm e variância de 0,2 cm^2.

(a) Qual o intervalo simétrico em torno da média, que conterá 75% dos comprimentos dos camarões?

(b) Qual o comprimento c, que é superado por 7% dos camarões?

Distribuição normal

Seja Y uma variável aleatória representando o número de sucessos em um total de n ensaios independentes e π a probabilidade de ocorrer sucesso em um ensaio. Então $Y \sim B(n; \pi)$. Observe os seguintes gráficos:



A aproximação normal com média $\mu = n\pi$ e variância $\sigma^2 = n\pi(1-\pi)$ aproxima-se bem das distribuições binomiais apresentadas!

Distribuição normal

Quando a aproximação é boa?

Quando a probabilidade π de ocorrer sucesso não está muito próxima de 0 ou de 1 e o número n de ensaios é grande, de tal modo que $n\pi \geq 20$.

O cálculo da probabilidade pela normal é feito utilizando-se uma distribuição $N(n\pi, n\pi(1-\pi))$.

Distribuição normal

Correção de continuidade

Consiste em somar e/ou subtrair 0,5 aos limites do intervalo para o qual desejamos calcular as probabilidades.

- Em muitas situações práticas o cálculo das probabilidades pode ser realizado sem levarmos em conta a correção de continuidade;
- Ignorar a correção para os casos em que $0,30 < \pi < 0,75$ e n maior do que 400.

Distribuição normal

Exercício: Sabe-se que a probabilidade de um indivíduo inoculado contra o surto de gripe vir a ter uma reação séria indesejável é de 0,05. Usando a aproximação normal à distribuição binomial, calcule a probabilidade de que mais de 16 indivíduos dentre 200 indivíduos inoculados tenham tais reações.

Distribuição normal

Exercício: Os ovos da produção de uma granja são classificados em grandes ou pequenos, conforme seu diâmetro. Verificou-se que 45% dos ovos são considerados grandes. Supondo que os ovos são colocados em caixas com 60 unidades, aleatoriamente, pergunta-se:

(a) Em que porcentagem esperada de caixas teremos pelo menos 50% de ovos grandes? (50% é igual a 30 ovos).

(b) Em que porcentagem de caixas teremos exatamente 50% de ovos grandes?