Modelos Probabilísticos

Luan D. Fiorentin

Universidade Federal do Paraná
Departamento de Estatística
Laboratório de Estatística e Geoinformação

18/09/2019

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

DEST/UFPR

Sumário

Introdução

- Existem diversos **modelos probabilísticos** que procuram descrever as V. A., denominados de **distribuições de probabilidade** de V. A. (discretas ou contínuas).
- A **distribuição de probabilidade** é a descrição das probabilidades associadas com os possíveis valores de X.
- Os valores que X assume determinam o **suporte** (S) da V. A.:
  - **Variáveis discretas**: suporte em um conjunto de valores enumeráveis (finitos ou infinitos).
  - **Variáveis contínuas**: suporte em um conjunto não enumerável de valores.
- Dada uma variável de interesse, denomina-se de **distribuição de probabilidade** de alguma V. A., a regra geral que define a
  - **Função de probabilidade** (fp): para variáveis aleatórias discretas.
  - **Função densidade de probabilidade** (fdp): para variáveis aleatórias contínuas.

Introdução

- Devido a importância prática, algumas distribuições são mais estudadas.
- Para variáveis aleatórias discretas:
  - Modelo Uniforme Discreto.
  - Modelo Bernoulli.
  - Modelo Binomial.
  - Modelo Geométrico.
  - Modelo Poisson.
  - Modelo Hipergeométrico.
- Para variáveis aleatórias contínuas:
  - Modelo Uniforme Contínuo.
  - Modelo Exponencial.
  - Modelo Normal.

Sumário

Modelo Uniforme Discreto

- **Definição**: Seja X uma V. A. assumindo valores $1, 2, \ldots, k$. A variável X segue o modelo Uniforme Discreto se atribui a mesma probabilidade $1/k$ a cada um desses k valores. Então, sua função de probabilidade é dada por

$$P[X = j] = \frac{1}{k}, \quad j = 1, 2, \ldots, k.$$

Modelo Uniforme Discreto

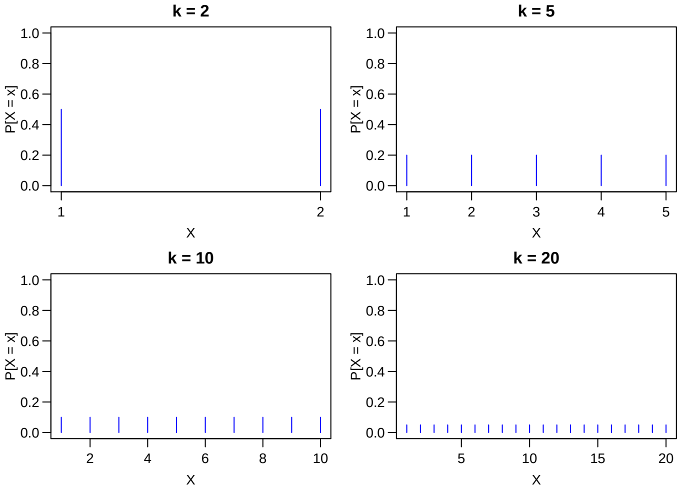
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a+1)^2-1}{12}$$

Notação: $X \sim \mathrm{U}_D[1, k]$

Modelo Uniforme Discreto



Exemplo - Modelo Uniforme Discreto

Uma rifa tem 100 bilhetes numerados de 1 a 100. Tenho 5 bilhetes consecutivos numerados de 21 a 25 e meu colega tem outros 5 bilhetes, com os números 1, 11, 29, 68 e 93. Quem tem maior possibilidade de ser sorteado?

Exemplo - Modelo Uniforme Discreto

Uma rifa tem 100 bilhetes numerados de 1 a 100. Tenho 5 bilhetes consecutivos numerados de 21 a 25 e meu colega tem outros 5 bilhetes, com os números 1, 11, 29, 68 e 93. Quem tem maior possibilidade de ser sorteado?

É comum a impressão de que "espalhar" os números é a melhor maneira de ganhar o sorteio. Entretanto, se assumir que a rifa é realizada de forma honesta, todos os números tem a mesma probabilidade de ocorrência, com valor de $1/100$ para cada um. Logo, a variável aleatória *número sorteado* segue um modelo uniforme. Portanto, ambas as pessoas tem a mesma probabilidade de $5/100$ de ganhar a rifa.

Sumário

Modelo Bernoulli

- **Definição**: Uma variável aleatória X segue o modelo Bernoulli se assume apenas os valores 0 ("fracasso") ou 1 ("sucesso"). Sua função de probabilidade é dada por

$$P[X = x] = p^x (1 - p)^{1-x}, \qquad x = 0, 1$$

onde o parâmetro $0 \leq p \leq 1$ é a probabilidade de sucesso.

Modelo Bernoulli

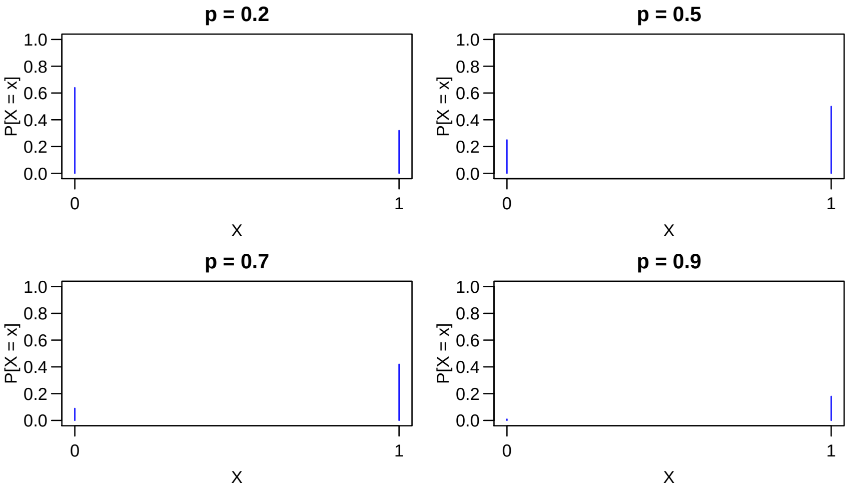
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = p$$

$$\sigma^2 = p(1 - p)$$

Notação: $X \sim \text{Ber}(p)$

Modelo Bernoulli



Exemplo - Modelo Bernoulli

No lançamento de uma moeda, considere cara como o evento de sucesso (1), e coroa como fracasso (0).

- Qual a probabilidade de sair cara, sendo que $p = 1/2$?

Exemplo - Modelo Bernoulli

No lançamento de uma moeda, considere cara como o evento de sucesso (1), e coroa como fracasso (0).

- Qual a probabilidade de sair cara, sendo que $p = 1/2$?

X	$P[X = x]$	$p = 1/2$
0	$1 - p$	$1/2$
1	p	$1/2$

Sumário

Modelo Binomial

- **Definição**: Seja um experimento realizado dentro das seguintes condições:

1. São realizados n "ensaios" de Bernoulli independentes.
2. Cada ensaio só pode ter dois resultados possíveis: "sucesso" ou "fracasso".
3. A probabilidade p de sucesso em cada ensaio é constante.

- Vamos associar a V. A. X o número de sucessos em n ensaios de Bernoulli. Portanto X poderá assumir os valores $0, 1, \ldots, n$.

Modelo Binomial

- Vamos determinar a distribuição de probabilidade de X por meio da probabilidade de um número genérico x de sucessos.
- Suponha que ocorram sucessos (1) apenas nas x primeiras provas, e fracassos (0) nas $n-x$ provas restantes

$$\underbrace{1, 1, 1, \ldots, 1}_{x}, \underbrace{0, 0, 0, \ldots, 0}_{n-x}.$$

- Como as provas são independentes, a probabilidade de ocorrência de x sucessos em n tentativas é uma extensão do modelo de Bernoulli para n ensaios, ou seja,

$$\underbrace{p \cdot p \cdot p \cdots p}_{x} \cdot \underbrace{(1-p) \cdot (1-p) \cdot (1-p) \cdots (1-p)}_{n-x} = p^x (1-p)^{n-x}.$$

Modelo Binomial

- Porém, o evento: "x sucessos em n provas" pode ocorrer de diferentes maneiras (ordens) distintas, todas com a mesma probabilidade.
- Como o número de ordens é o número de combinações de n elementos tomados x a x, então a probabilidade de ocorrerem x sucessos em n provas de Bernoulli será então a distribuição binomial, dada por

$$P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n,$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o **coeficiente binomial**, que dá o número total de combinações possíveis de n elementos, com x sucessos.

Modelo Binomial

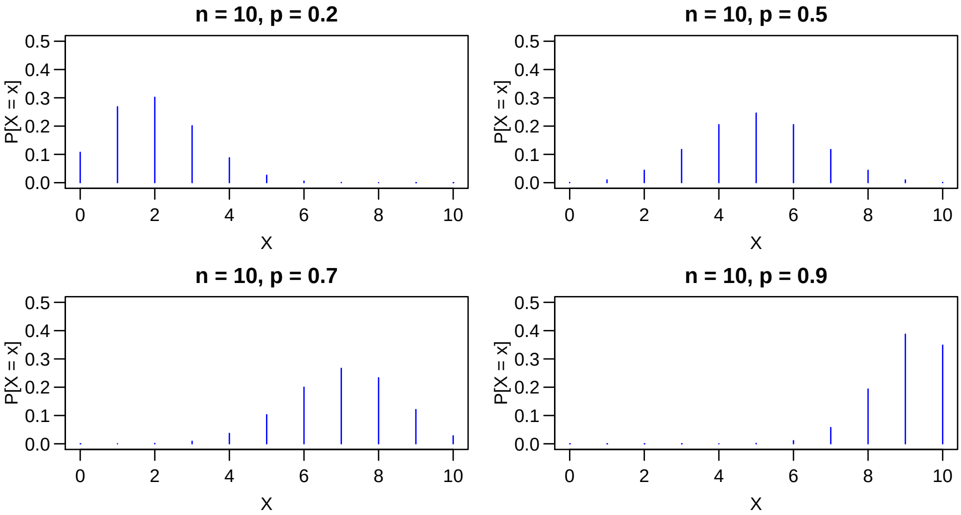
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = np$$

$$\sigma^2 = np(1 - p)$$

Notação: $X \sim B(n, p)$

Modelo Binomial



Exemplo - Modelo Binomial

Um processo de produção opera a uma taxa de 5% de peças produzidas não conformes. A cada hora uma amostra de 5 unidades do produto é retirada, e o número de peças não conformes (defeituosas) registrado.

1. Qual a probabilidade de duas peças defeituosas serem verificadas numa realização da inspeção?
2. A empresa será multada se dois ou mais itens defeituosos forem verificados na próxima inspeção. Qual a probabilidade de multa?

Exemplo - Modelo Binomial

As probabilidades são completamente caracterizadas pela informações dos parâmetros. Para calcular a probabilidade de 2 sucessos em um $B(n = 5, p = 0,05)$

1. Qual a probabilidade de duas peças defeituosas serem verificadas numa realização da inspeção?

$$P(X = 2) = \binom{5}{2} 0,05^2 (1 - 0,05)^{5-2} = 0,021.$$

2. A empresa será multada se dois ou mais itens defeituosos forem verificados na próxima inspeção. Qual a probabilidade de multa?

$$1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)]$$

$$1 - P(X < 2) = 1 - [0,774 + 0,204] = 0,023.$$

Sumário

Modelo Geométrico

- **Definição**: Considere o número (k) de ensaios Bernoulli **que precedem o primeiro sucesso**. Nesse caso, dizemos que a variável aleatória X tem distribuição Geométrica de parâmetro p, e sua função de probabilidade tem a forma

$$P(X = k) = p(1-p)^k, \quad k = 0, 1, 2, \ldots .$$

onde $0 \leq p \leq 1$ é a probabilidade de sucesso.

Modelo Geométrico

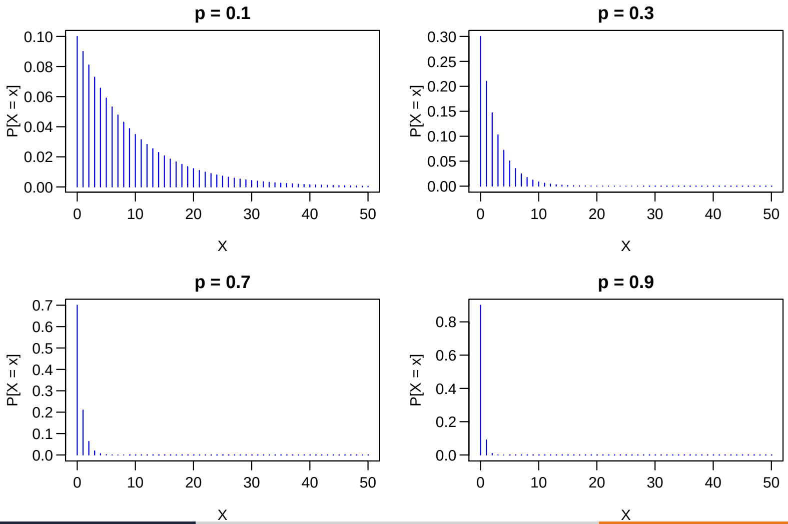
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \frac{1-p}{p}$$

$$\sigma^2 = \frac{1-p}{p^2}$$

Notação: $X \sim \mathrm{G}(p)$.

Modelo Geométrico



Exemplo - Modelo Geométrico

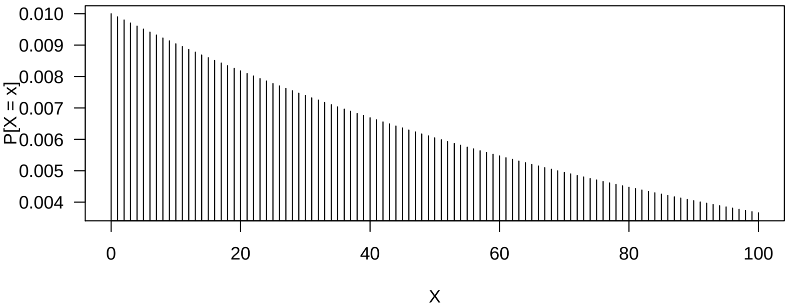
Uma linha de produção está sendo analisada para efeito de controle da qualidade das peças produzidas. Tendo em vista o alto padrão requerido, a produção é interrompida para regulagem toda vez que uma peça defeituosa é observada. Se $0,01$ é a probabilidade da peça ser defeituosa, estude o comportamento da variável Q, *quantidade de peças boas produzidas antes da primeira defeituosa*.

Exemplo - Modelo Geométrico

$$P(Q = k) = 0.01 \times 0.99^k, \quad k = 0, 1, 2, \ldots$$

Usando o R, temos que

```
x = c(0:100)
dx = dgeom(x = x, prob = 0.01)
plot(x, dx, type = "h", xlab = "X", ylab = "P[X = x]")
```



Sumário

Modelo Poisson

- **Definição:** Seja um experimento realizado nas seguintes condições:

1. As ocorrências são independentes.
2. As ocorrências são aleatórias.
3. A variável aleatória X é o número de ocorrências de um evento **ao longo de algum intervalo** (de tempo ou espaço).

- Denominamos esse experimento de **processo de Poisson**.
- Vamos associar a variável aleatória X o número de ocorrências em um intervalo. Portanto, X poderá assumir os valores $0, 1, \ldots$ (sem limite superior).

Modelo Poisson

- Uma variável aleatória X segue o modelo de Poisson se surge a partir de um processo de Poisson, e sua **função de probabilidade** for dada por

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

  onde o parâmetro $\lambda > 0$ é a taxa média de ocorrências em um intervalo de tempo ou espaço.
- Aplicável no caso de eventos que ocorrem sob **taxa constante**.

Modelo Poisson

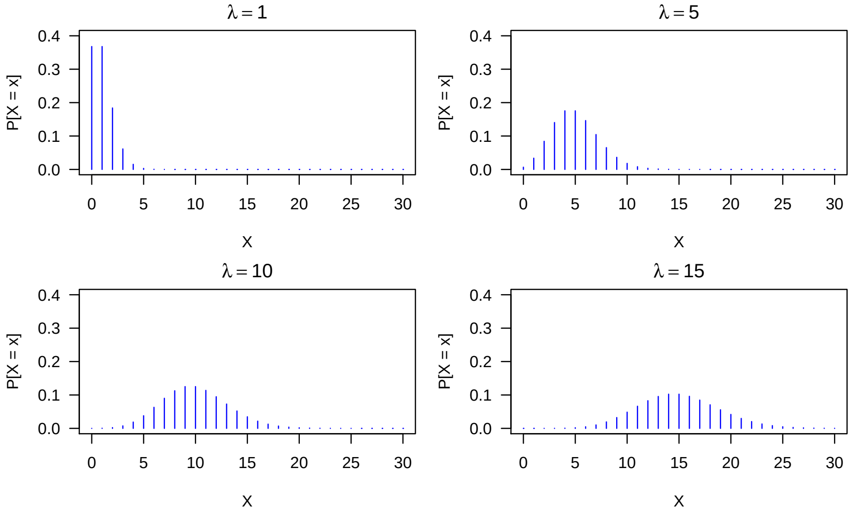
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

Notação: $X \sim P(\lambda)$.

Modelo Poisson



Exemplo - Modelo Poisson

A emissão de partículas radioativas têm sido modelada através de uma distribuição de Poisson, com o valor do parâmetro dependendo da fonte utilizada.

Suponha que o número de partículas alfa, emitidas por minuto, seja uma variável aleatória seguindo o modelo Poisson com parâmetro 5, isto é, a taxa média de ocorrência é de 5 emissões a cada minuto.

1. Qual a probabilidade de haver duas emissões em um minuto?

$$P(X = 2) = \frac{5^2 e^{-5}}{2!} = 0,084.$$

2. Calcule a probabilidade de haver mais de 2 emissões em um minuto?

$$1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)]$$

$$1 - P(X < 2) = 1 - [0,007 + 0,034] = 0,959.$$

Sumário

Modelo Hipergeométrico

- **Definição**: Considere um conjunto de N objetos dos quais D são do tipo I e $N - D$ são do tipo II. Para um sorteio de n objetos $n < N$, feito ao acaso e sem reposição, defina X como o número de objetos de tipo I selecionados.
- A variável aleatória X segue o modelo Hipergeométrico e sua função de probabilidade é dada por

$$P[X = x] = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}},$$

onde: $x = 0, 1, 2, ..., min(n, D)$ e $x = max\{0, n - (n - D)\} \leq x \leq min\{n, D\}$.

Modelo Hipergeométrico

A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \frac{nD}{N},$$

$$\sigma^2 = \frac{nD}{N}\left(\frac{N-D}{N}\right)\left(\frac{N-n}{N-1}\right).$$

Notação: $X \sim HG(D, N, n)$

Exemplo - Modelo Hipergeométrico

Considere que, num lote de 20 peças, existam o total de 4 peças defeituosas. Selecionando-se 5 dessas peças sem reposição.

- Qual seria a probabilidade de 2 defeituosas terem sido escolhidas?

$$P[X = 2] = \frac{\binom{4}{2}\binom{16}{3}}{\binom{20}{5}} = 0,217.$$

Sumário

Modelo Uniforme Contínuo

- **Definição**: uma variável aleatória X tem distribuição Uniforme contínua no intervalo $[a, b]$, $a < b$, se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{se } a \leq x \leq b \\ 0 & \text{caso contrário} \end{cases}.$$

Modelo Uniforme Contínuo

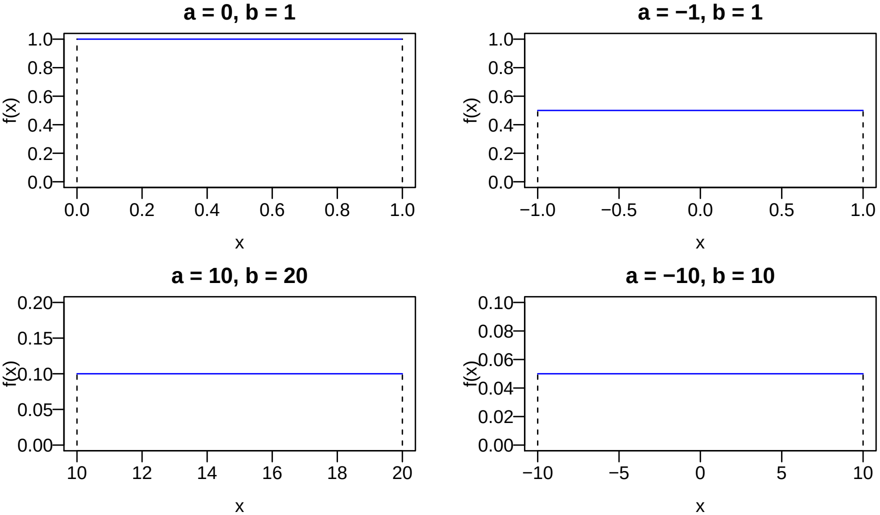
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a)^2}{12}.$$

Notação: $X \sim U[a, b]$

Modelo Uniforme contínuo



Exemplo - Modelo Uniforme contínuo

Com o objetivo de verificar a resistência à pressão de água, os técnicos de qualidade de uma empresa inspecionam os tubos de PVC produzidos. Os tudos inspecionados têm 6 metros de comprimento e são submetidos a grandes pressões até o aparecimento do primeiro vazamento, cuja distância a uma das extremidades (fixada à priori) é anotada para fins de análises.

Escolhe-se um tubo ao acaso para ser inspecionado. Queremos calcular a probabilidade de que o vazamento esteja a no máximo 1 metro das extremidades (V. A. $X \sim U(a = 0, b = 6)$).

Exemplo - Modelo Uniforme contínuo

Com o objetivo de verificar a resistência à pressão de água, os técnicos de qualidade de uma empresa inspecionam os tubos de PVC produzidos. Os tudos inspecionados têm 6 metros de comprimento e são submetidos a grandes pressões até o aparecimento do primeiro vazamento, cuja distância a uma das extremidades (fixada à priori) é anotada para fins de análises.

Escolhe-se um tubo ao acaso para ser inspecionado. Queremos calcular a probabilidade de que o vazamento esteja a no máximo 1 metro das extremidades (V. A. $X \sim U(a = 0, b = 6)$).

$$P(X \in \{[0,1]\} \cup \{[5,6]\}) = P(0 \leq X \leq 1) + P(5 \leq X \leq 6),$$

$$P(X \in \{[0,1]\} \cup \{[5,6]\}) = \int_0^1 \frac{1}{6} dx + \int_5^6 \frac{1}{6} dx = \frac{1}{6} - 0 + \frac{6}{6} - \frac{5}{6} = \frac{1}{3}.$$

Importante: os intervalos $[0,1]$ e $[5,6]$ são disjuntos e, portanto, a probabilidade da sua união é a soma das probabilidades de ocorrência de cada intervalo.

Sumário

Modelo Exponencial

- **Definição**: uma variável aleatória contínua X, assumindo valores não negativos, segue o modelo exponencial com parâmetro $\alpha > 0$ se sua densidade é dada por

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{se } x \geq 0 \\ 0 & \text{caso contrário.} \end{cases}$$

Modelo Exponencial

A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \frac{1}{\alpha}$$

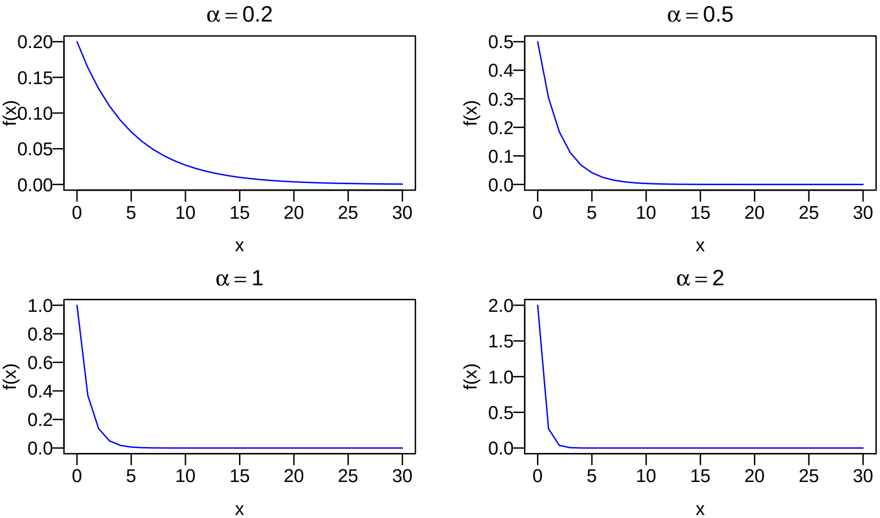
$$\sigma^2 = \frac{1}{\alpha^2}.$$

Notação: $X \sim \text{Exp}(\alpha)$

Propriedade da distribuição Exponencial

$P(a < X < b) = \int_a^b \alpha e^{-\alpha x}\, dx = e^{-\alpha a} - e^{-\alpha b}.$

Modelo Exponencial



Exemplo - Modelo Exponencial

Uma indústria fabrica lâmpadas especiais que ficam em operação continuamente.

A empresa oferece a seus clientes a garantia de reposição, caso a lâmpada dure menos de 50 horas. A vida útil dessas lâmpadas é modelada através da distribuição Exponencial com parâmetro $1/8000$.

1. Qual a proporção de troca por defeito de fabricação.
2. Qual a duração média das lâmpadas?

Exemplo 6 - Modelo Exponencial

1. Qual a proporção de troca por defeito de fabricação.

$$\begin{aligned} P(X < 50) &= \int_0^{50} \alpha e^{-\alpha x}\, dx \\ &= \int_0^{50} \frac{1}{8000} e^{-\frac{1}{8000}x}\, dx \\ &= e^{-\frac{1}{8000}0} - e^{-\frac{1}{8000}50} && \text{(usando a propriedade)} \\ &= 1 - 0.994 \\ &= 0.006 \end{aligned}$$

2. Qual a duração média das lâmpadas?

$$E(X) = \mu = \frac{1}{\alpha} = \frac{1}{1/8000} = 8000.$$

Sumário

Modelo Normal

- **Definição**: Dizemos que uma variável aleatória X segue o modelo normal se sua fdp é dada por

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

onde $\mu \in \mathbb{R}$ é a média da população, $\sigma \in \mathbb{R}^+$ é o desvio-padrão populacional.

Modelo Normal

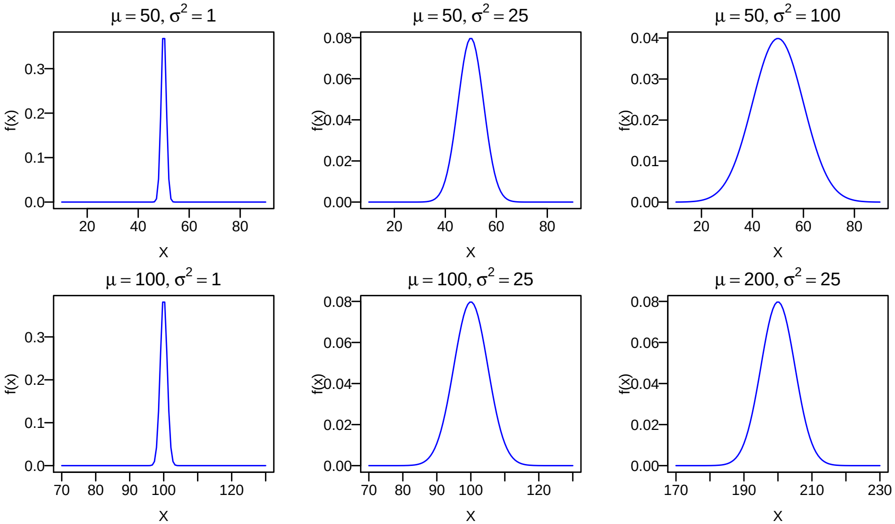
A média (esperança) e a variância são dadas, respectivamente, pelas expressões

$$\mu = \frac{\sum_{i=i}^{n} X_i}{n}$$

$$\sigma^2 = \frac{\sum_{i=1}^{n} X_i^2}{n} - \bar{X}^2.$$

Notação: $X \sim \mathrm{N}(\mu, \sigma^2)$

Modelo Normal



Modelo Normal

Características da curva normal:

- É **simétrica** em relação à μ.
- O ponto máximo (moda) de $f(x)$ é o ponto $x = \mu$.
- Os pontos de inflexão da função são $\mu - \sigma$ e $\mu + \sigma$.
- A área total sob a curva é 1 ou 100%.
- A curva é **assintótica** em relação ao eixo x.

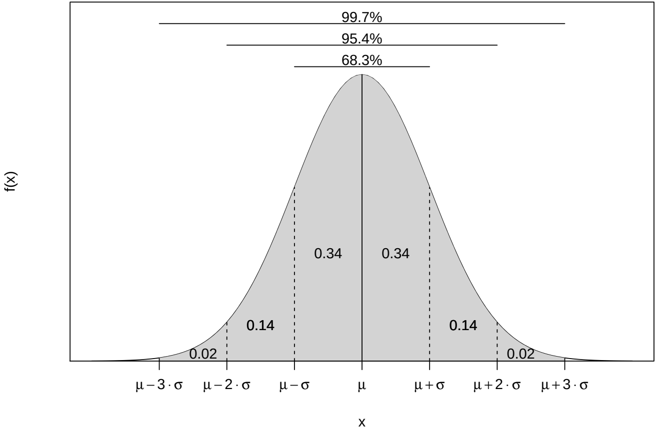
Modelo Normal

- Para qualquer variável aleatória normal X valem as seguintes relações:

$$
\begin{aligned}
P[X > \mu] &= P[X < \mu] \\
P[\mu - \sigma < X < \mu + \sigma] &\approxeq 0,683 \\
P[\mu - 2\sigma < X < \mu + 2\sigma] &\approxeq 0,954 \\
P[\mu - 3\sigma < X < \mu + 3\sigma] &\approxeq 0,997
\end{aligned}
$$

- Portanto, 6σ é frequentemente referida como a **largura** de uma distribuição normal.
- Métodos mais avançados de integração podem ser utilizados para mostrar que a área sob a função densidade de probabilidade normal de $-\infty < x < \infty$ é igual a 1.

Regra empírica para uma distribuição normal



Modelo Normal

- Para obter uma probabilidade do modelo normal, devemos calcular a área entre os pontos a e b, ou seja,

$$P[a < X < b] = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

- No entanto, essa função não possui forma fechada, e o cálculo de probabilidades pode ser feito apenas por aproximações numéricas.
- Para contornar esse problema, os valores de probabilidade são obtidos para uma distribuição normal padrão (Z) com $\mu = 0$ e $\sigma^2 = 1$,

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

- A vantagem é que podemos fazer uma única tabela com as integrais aproximadas de Z, ao invés de uma tabela para cada par (μ, σ^2).

Modelo Normal

- Se $Z \sim \mathrm{N}(0, 1)$, então sua fdp é

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z)^2\right],$$

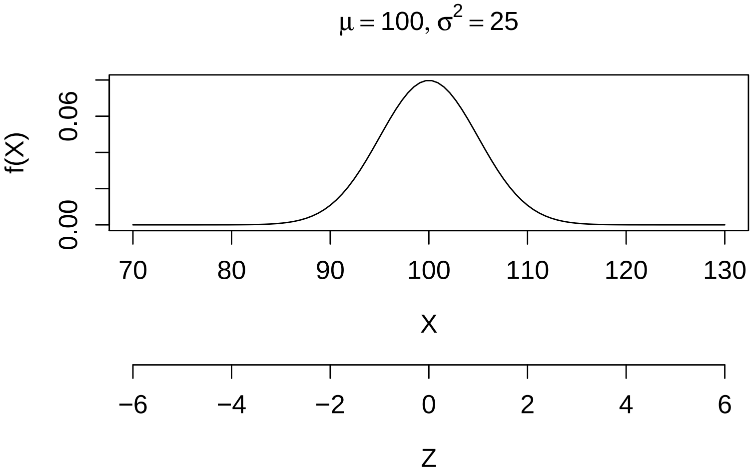
e para se obter a probabilidade de Z estar entre a e b,

$$P[a < Z < b] = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z)^2\right] dz.$$

- As integrais (áreas) para valores de Z entre 0,00 e 3,99 estão na tabela. Portanto, para qualquer valor de X entre a e b, podemos calcular a probabilidade correspondente por meio da transformação

$$P[a < X < b] = P\left[\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right].$$

Modelo Normal



Modelo Normal

Exemplo de uso da tabela

- Calcule as probabilidades (áreas):
  - $P(0 < Z < 2)$
  - $P(Z > 2)$
  - $P(Z < -2)$
  - $P(2,0 < Z < 2,5)$
  - $P(-2,61 < Z < 2,43)$
  - $P(Z > -1,63)$
  - Qual é o valor de c tal que $P(0 < Z < c) = 0,4$?
  - Qual é o valor de d tal que $P(Z > d) = 0,8$?

Modelo Normal

Exemplo de uso da tabela

- Calcule as probabilidades (áreas):
  - $P(0 < Z < 2) = 0,4772$
  - $P(Z > 2) = 0,0228$
  - $P(Z < -2) = 0,0228$
  - $P(2,0 < Z < 2,5) = 0,0165$
  - $P(-2,61 < Z < 2,43) = 0,9879$
  - $P(Z > -1,63) = 0,9484$
  - Qual é o valor de c tal que $P(0 < Z < c) = 0,4$? $c = 1,281$
  - Qual é o valor de d tal que $P(Z > d) = 0,8$? $c = -0,841$

Exemplo - Modelo Normal

Doentes sofrendo de certa moléstia são submetidos a um tratamento intensivo. O tempo de cura foi modelado por uma densidade Normal de média 15 e desvio padrão 2 (em dias).

1. Calcule a proporção de pacientes que demorarão mais de 17 dias para se recuperar.
2. Calcule a probabilidade um paciente selecionado ao acaso demorar menos de 20 dias para se recuperar.
3. Qual o tempo máximo necessário para a recuperação de 25% dos pacientes?
4. Se 100 pacientes forem escolhidos ao acaso, qual seria o número esperado de doentes curados em menos de 11 dias?

Exemplo - Modelo Normal

1. $P(X > 17) = P\left(\frac{X-15}{\sqrt{4}} > \frac{17-15}{\sqrt{4}}\right) = P(Z > 1) = 0,1584.$
2. $P(X < 20) = P\left(\frac{X-15}{\sqrt{4}} > \frac{20-15}{\sqrt{4}}\right) = P(Z > 2,5) = 0,9938.$
3. $P(X > 17) = P\left(\frac{X-15}{\sqrt{4}} > \frac{t-15}{\sqrt{4}}\right) = 0,25 \ \frac{t-15}{\sqrt{4}} = -1,67 \rightarrow t = 13,66.$
4. $$P(X < 11) = P\left(\frac{X-15}{\sqrt{4}} > \frac{11-15}{\sqrt{4}}\right) = P(Z < -2) = 0,0228.$$

   Para 100 pacientes o número esperado com o tempo de cura inferior a 11 dias será $100 \cdot 0,0228 \approx 2$ pacientes.

Normal como aproximação da binomial

- A distribuição Normal é uma das mais importantes na Estatística:
  - Muitos fenômenos aleatórios se comportam próximos à essa distribuição
  - Pode ser usada como aproximação para outras distribuições

Se $X \sim \text{Bin}(n, p)$ então $E(X) = np$ e $Var(X) = np(1-p)$.

Podemos aproximar a binomial pela normal, usando

$$Y \sim \text{N}(\mu = np, \sigma^2 = np(1-p)),$$

em geral, quando $np \geq 5$ e $np(1-p) \geq 5$.

Exemplo 6.11

Estudo do Sindicato dos Bancários indica que cerca de 30% dos funcionários de banco têm problemas de estresse, provenientes das condições de trabalho. Numa amostra de 200 bancários, qual seria a probabilidade de pelo menos 50 com essa doença?

Exemplo 6.11

Estudo do Sindicato dos Bancários indica que cerca de 30% dos funcionários de banco têm problemas de estresse, provenientes das condições de trabalho. Numa amostra de 200 bancários, qual seria a probabilidade de pelo menos 50 com essa doença?

Temos então $X \sim \text{Bin}(200, 0.3)$, e a probabilidade seria

$$P(X \geq 50) = \sum_{k=50}^{200} \binom{200}{k} 0.3^k 0.7^{200-k}$$

que é difícil de calcular sem computador.

Exemplo 6.11

Mas $E(X) = np = 60$ e $Var(X) = np(1-p) = 42$. Assim, temos $Y \sim \mathrm{N}(60, 42)$, de modo que

$$P(X \geq 50) \approx P(Y \geq 50 - 0.5) = P\left(\frac{Y-60}{\sqrt{42}} \geq \frac{49.5-60}{\sqrt{42}}\right)$$
$$= P(Z \geq -1.62) = 0.9474$$

- **Obs.:** o valor -0.5 é o fator de correção de continuidade.

Usando o R:

```
## Cálculo exato pela binomial
pbinom(49, size = 200, prob = 0.3, lower.tail = FALSE)
```

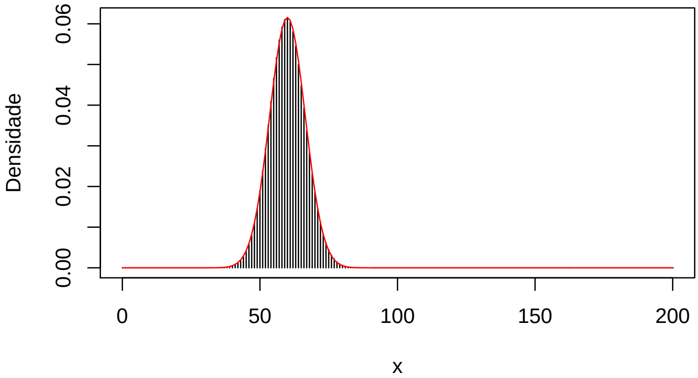
```
# [1] 0.9494082
```

```
## Aproximação pela Normal
pnorm(50-0.5, mean = 60, sd = sqrt(42), lower.tail = FALSE)
```

```
# [1] 0.9474037
```

Exemplo 6.11

Aproximação de $X \sim \text{Bin}(200, 0.3)$ com $Y \sim \text{N}(60, 42)$.



Sumário

Exercícios recomendados

- Seção 3.2: Ex. 1, 2, 3, 4, 5 e 6.
- Seção 3.3: Ex. 1, 2, 3, 4, 5 e 6.
- Seção 3.4: Ex. 3, 14, 16, 17, 18, 21, 25, 26 e 27.
- Seção 6.2: Ex. 1, 2, 3, 4, 6, 7, 8 e 9.
- Seção 6.3: Ex. 25, 27, 28, 29 e 33.