Métodos Estatísticos Básicos

Aula 10 - Distribuições de probabilidade

Regis A. Ely

Departamento de Economia
Universidade Federal de Pelotas

14 de junho de 2021

Conteúdo

Distribuição binomial

Variável aleatória: número de sucessos obtidos em n realizações de experimentos de Bernoulli com probabilidade de sucesso p

Distribuição binomial: $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ para $x = 0, 1, 2, ..., n$

Valor esperado: $E(x) = np$

Variância: $V(x) = np(1-p)$

Exemplo no R

Suponha que haja 10 questões de múltipla escolha em uma prova de estatística. Cada pergunta tem 5 respostas possíveis e apenas uma delas está correta. Encontre a probabilidade de um aluno responder aleatoriamente a todas as perguntas e obter:

1. Exatamente cinco respostas corretas

```
dbinom(5, size = 10, prob = 0.2)
```

```
[1] 0.02642412
```

2. Cinco ou mais respostas corretas

```
1 - pbinom(4, size = 10, prob = 0.2)
```

```
[1] 0.0327935
```

Distribuição de Poisson

Variável aleatória: número de ocorrências de certo evento obtidas em um determinado período de tempo, sendo o parâmetro λ o número esperado de ocorrências

Distribuição de Poisson: $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ para $x = 0, 1, ..., n$

Valor esperado: $E(x) = \lambda$

Variância: $V(x) = \lambda$

Exemplo no R

Se a cada 5 minutos chegar em média 50 pacientes em uma UTI de um hospital, qual a probabilidade de:

1. Exatamente 10 pacientes chegarem a UTI no próximo minuto?

```
dpois(10, lambda = 10)
```

```
[1] 0.12511
```

2. Cinco pacientes ou mais chegarem a UTI no próximo minuto?

```
1 - ppois(4, lambda = 10)
```

```
[1] 0.9707473
```

Distribuição geométrica

Variável aleatória: número de repetições necessárias para obter o primeiro sucesso em experimentos de Bernoulli com probabilidade p

Distribuição geométrica: $f(x) = q^{x-1}p$ para $x = 1, 2...$

Valor esperado: $E(x) = \frac{1}{p}$

Variância: $V(x) = \frac{q}{p^2}$

Exemplo no R

Se 20% dos motores elétricos produzidos em uma empresa apresentam falhas. Se selecionarmos cinco motores ao acaso qual a probabilidade de encontrarmos quatro motores perfeitos antes de encontrarmos um defeituoso?

```r
dgeom(4, prob = 0.2)
```

```
[1] 0.08192
```

Distribuição de Pascal

Variável aleatória: número de repetições necessárias para obter o r-ésimo sucesso em experimentos de Bernoulli com probabilidade p (generalização da distribuição geométrica)

Distribuição de Pascal (binomial negativa): $f(x) = \binom{x-1}{r-1} p^r q^{x-r}$ para $x = r, r+1, r+2...$

Valor esperado: $E(x) = \frac{r}{p}$

Variância: $V(x) = \frac{rq}{p^2}$

Exemplo no R

Uma empresa de petróleo tem uma chance de 20% de encontrar petróleo ao perfurar um poço. Qual a probabilidade de que a empresa perfure 7 poços e encontre petróleo 3 vezes?

```
dnbinom(7-3, size = 3, prob = 0.2)
```

```
[1] 0.049152
```

Distribuição hipergeométrica

Variável aleatória: número de peças defeituosas em uma amostra de n peças retiradas de um total de N peças que continham r defeituosas

Distribuição hipergeométrica: $f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$ para $x = 0, 1, 2...$

Valor esperado: $E(x) = np$

Variância: $V(x) = npq\frac{N-n}{N-1}$

Exemplo no R

Qual é a probabilidade de selecionar 14 bolas vermelhas de uma amostra de 20 retiradas de uma urna contendo 70 bolas vermelhos e 30 verdes?

```r
dhyper(14, 70, 30, 20)
```

```
[1] 0.2140911
```

Distribuição uniforme

Variável aleatória: um ponto em um intervalo de reta dos números reais

Distribuição uniforme: $f(x) = \frac{1}{b-a}$ para $a \leq x \leq b$ e 0 caso contrário

Valor esperado: $E(x) = \frac{a+b}{2}$

Variância: $V(x) = \frac{(b-a)^2}{12}$

Exemplo no R

Um ponto é escolhido ao acaso no segmento de reta $[1,3]$. Qual a probabilidade de que o ponto esteja entre $1,5$ e 2?

```r
punif(2, min=1, max=3)-punif(1.5, min=1, max=3)
```

```
[1] 0.25
```

Distribuição exponencial

Variável aleatória: tempo ou distância necessária para ocorrências de um processo de Poisson

Distribuição exponencial: $f(x) = \alpha e^{-\alpha x}$ para $x > 0$ e 0 para $x \leq 0$

Valor esperado: $E(x) = \frac{1}{\alpha}$

Variância: $V(x) = \frac{1}{\alpha^2}$

Exemplo no R

Suponha que o tempo médio de atendimento de um caixa de supermercado seja de três minutos. Encontre a probabilidade de o tempo de atendimento de um cliente ser concluído pelo caixa em menos de dois minutos.

```r
pexp(2, rate = 1/3)
```

```
[1] 0.4865829
```

Distribuição normal

Variável aleatória: valores pertencentes ao conjunto dos números reais obtidos de experimentos aleatórios

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

Valor esperado: $E(x) = \mu$

Variância: $V(x) = \sigma^2$

Distribuição normal

- Os parâmetros da distribuição normal devem satisfazer as seguintes condições: $-\infty < \mu < \infty$ e $\sigma > 0$
- Se X tem distribuição normal com média μ e variância σ^2, denotamos $X \sim N(\mu, \sigma^2)$
- Se $X \sim N(0, 1)$ dizemos que X tem uma distribuição normal reduzida, de modo que:

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

Distribuição normal

- Se $X \sim N(\mu, \sigma)$ e $Y = aX + b$, então $Y \sim N(a\mu + b, a^2\sigma^2)$
- Se $X \sim N(\mu, \sigma)$ e $Y = (X - \mu)/\sigma$, então $Y \sim N(0, 1)$
- Se $X \sim N(\mu, \sigma)$ então:

$$P(a \leq X \leq b) = P(\tfrac{a-\mu}{\sigma} \leq Y \leq \tfrac{b-\mu}{\sigma}) = F(\tfrac{b-\mu}{\sigma}) - F(\tfrac{a-\mu}{\sigma})$$

onde F é a distribuição acumulada da normal reduzida, $N(0, 1)$

- Temos também vale que $F(-x) = 1 - F(x)$

Exemplo no R

Suponha que a altura em centímetros de uma amostra de estudantes de estatística seja distribuída normalmente com $E(X) = 170$ e $V(X) = 100$. Qual a probabilidade de que um aluno escolhido ao acaso tenha altura menor do que 160cm?

$$P(X < 160) = P(\frac{X-170}{10} < \frac{160-170}{10}) = F(-1) = 1 - F(1) = 0,159$$

```
pnorm(160, mean = 170, sd = 10)
```

```
## [1] 0.1586553
```

Teorema do Limite Central

Seja $X_1, X_2, \ldots, X_n, \ldots$ uma sequência de variáveis aleatórias independentes, com $E(X_i) = \mu_i$ e $V(X_i) = \sigma_i^2$ para $i = 1, 2, \ldots$ Fazendo $X = X_1 + X_2 + \ldots + X_n$, então, sob determinadas condições gerais:

$$Z_n = \frac{X - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \sim N(0, 1)$$