Distribuições Discretas

Prof. José Roberto Silva dos Santos

Depto. de Estatística e Matemática Aplicada - UFC

Fortaleza, 20 de maio de 2022

Sumário

- Se $X \sim B(n,p)$ então
  - $\mathbb{E}(X) = np$ (Demonstração)
  - $\mathrm{Var}(X) = np(1-p)$ (Demonstração)
  - $\mathbb{E}(X^k) = np\mathbb{E}[(Y+1)^{k-1}]$ sendo que $Y \sim B(n-1,p)$.

Distribuição Binomial

Proposição:

Se $X \sim B(n,p)$ então

$$\mathbb{P}(X = x+1) = \frac{p}{1-p}\frac{n-x}{x+1}\mathbb{P}(X = x)$$

- Suponha que uma válvula eletrônica, instalada em determinado circuito, tenha probabilidade $0,2$ de funcionar mais do que 500 horas. Se testarmos 20 válvulas:

(a) Qual será a probabilidade de que, exatamente 2 válvulas, funcionem mais do que 500 horas?

(b) Qual o número esperado de válvulas que funcionaram mais do que 500 horas?

(c) Qual o desvio-padrão?

- **Solução**: Seja X a v.a que representa o número de válvulas que funcionem mais do que 500 horas. Admitindo que, $X \sim B(20; 0,2)$ temos que

$$
\begin{aligned}
\mathbb{P}(X = 2) &= \binom{20}{2}(0,2)^2(0,8)^{20-2} \\
&= 190 \times (0,2)^2 \times (0,8)^{18} = 0,137
\end{aligned}
$$

- O número esperado de válvulas funcionando mais do que 500 horas é dado por

$$\mathbb{E}(X) = 20 \times 0,2 = 4, \text{ com desvio-padrão}$$

$$\mathrm{DP}(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{20 \times 0,2 \times 0,8} \approx 1,79.$$

Exemplo

- O controle de qualidade de uma determinada industria verificou que a probabilidade de que uma peça, produzida por determinada máquina, seja defeituosa é $0,2$. Se 10 peças produzidas por essa máquina forem escolhidas ao acaso, responda:
  (a) Qual a probabilidade de ser encontrada no máximo uma peça defeituosa?

  (b) Qual a esperança e o desvio-padrão do número de peças defeituosas?

  (c) Se nenhuma peça defeituosa for encontrada, a máquina está operando em perfeitas condições. Se somente uma peça defeituosa for encontrada, a máquina precisará de "reparos leves"com um custo de R$ 10,00. Já se forem encontradas mais do que uma, a máquina precisará de "reparos maiores"ao custo de R$ 50,00. Qual o custo de manutenção esperado para essa máquina?

Exemplo

Roda da Fortuna

- O jogo a seguir, conhecido como *roda da fortuna*, é bastante popular em muitos parques e casinos. Um jogador aposta em um número de 1 a 6. Três dados são então lançados, e se o número apostado sair i vezes, $i = 1, 2, 3$, então o jogador ganha i unidades; se o número apostado não sair em nenhum dos dados, então o jogador perde 1 unidade. Este jogo é justo para o jogador?

Fonte: Ross (2010)

Sumário

Distribuição Hipergeométrica

- Este é um modelo para amostragem *sem reposição* de uma população com um número finito de elementos que podem ser classificados em duas categorias mutuamente excludentes.
- Detalhes:
  - N objetos
  - r possuem uma característica A
  - $N - r$ possuem uma característica B
  - um grupo de n elementos é escolhido ao acaso, dentre os N possíveis, sem reposição
- Objetivo: calcular a probabilidade de que este grupo de n elementos contenha x elementos com a característica A.

Distribuição Hipergeométrica

- Seja X a v.a representando o número de elementos com a característica A, dentre os n elementos selecionados. Dizemos que X tem distribuição hipergeométrica e sua distribuição de probabilidade é dada por:

$$p(x) = \mathbb{P}(X = x) = \frac{\begin{pmatrix} r \\ x \end{pmatrix} \begin{pmatrix} N - r \\ n - x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}} \quad \forall$$

  $\max\{0, n - (N - r)\} \leq x \leq \min\{r, n\}$.

- **Notação**: $X \sim \text{Hip}(N, n, r)$

- Se $X \sim \text{Hip}(N, n, r)$ então

$$\mathbb{E}(X^k) = n\frac{r}{N}\mathbb{E}((Y+1)^{k-1}) \text{ em que } Y \sim \text{Hip}(N-1, n-1, r-1).$$

Dessa forma,

$$\mathbb{E}(X) = n\frac{r}{N} \text{ e}$$

$$\text{Var}(X) = n\frac{r}{N}(1-\frac{r}{N})(\frac{N-n}{N-1}).$$

- O termo $\frac{N-n}{N-1}$ é o chamado *fator de correção em amostras finitas*.

- Aplicação: Controle de Qualidade
  Suponha um lote com $N = 100$ elementos a ser analisado. São escolhidas $n = 5$ peças sem reposição. Sabendo que neste lote de 100 elementos, $r = 10$ são defeituosos, a probabilidade de não se obter nenhuma peça defeituosa na amostra retirada é:

$$\mathbb{P}(X = 0) = \frac{\begin{pmatrix} 10 \\ 0 \end{pmatrix}\begin{pmatrix} 100 - 10 \\ 5 - 0 \end{pmatrix}}{\begin{pmatrix} 100 \\ 10 \end{pmatrix}} = \frac{\begin{pmatrix} 90 \\ 5 \end{pmatrix}}{\begin{pmatrix} 100 \\ 10 \end{pmatrix}} \approx 0,584$$

Aproximação da Hipergeométrica pela Binomial

- Uma interessante propriedade da distribuição Hipergeométrica pode ser observada quando N é grande em comparação a n. Nesta situação, as retiradas serão "quase" independentes.
- Quando isso ocorre a distribuição Hipergeométrica se aproxima de uma distribuição binomial com parâmetros n e r/N. Ou seja,

$$\mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \approx \binom{n}{x}(\frac{r}{N})^x(1 - \frac{r}{N})^{n-x}$$

- Em geral, essa é uma boa aproximação quando $\frac{n}{N} \leq 0,1$.

Tabela: Valores de $p(0)$ para diferentes valores de N, n e r sob os modelos Hipergeométrico e Binomial

(N, n, r)	n/N	$\text{Hip}(N, n, r)$	$B(n, r/N)$
$(50, 15, 10)$	0,300	0,018	0,035
$(70, 15, 10)$	0,214	0,074	0,099
$(90, 15, 10)$	0,167	0,145	0,171
$(110, 15, 10)$	0,136	0,215	0,239
$(130, 15, 10)$	0,115	0,280	0,301
$(150, 15, 10)$	0,100	0,337	0,355
$(170, 15, 10)$	0,088	0,387	0,403
$(190, 15, 10)$	0,079	0,430	0,444
$(200, 15, 10)$	0,075	0,450	0,463