INSTITUTO FEDERAL DE SANTA CATARINA
CAMPUS SÃO JOSÉ
COORDENADORIA DE ÁREA DE TELECOMUNICAÇÕES
ENGENHARIA DE TELECOMUNICAÇÕES

PRE29006	BOLAS E URNAS	2017.2

A principal referência para este trabalho é a seguinte obra:

C. Dennis O'Shaughnessy. *Hypergeometric and Negative Hypergeometric Distributions*. Disponível em `http://math.usask.ca/~oshaughn/103Hypergeometric.pdf`.

Introdução

Considere uma urna com t bolas, sendo g verdes e r vermelhas, com $g + r = t$. Quando uma bola é retirada **com reposição**, significa que esta bola volta para a urna antes da próxima ser retirada. Quando uma bola é retirada **sem reposição**, significa que esta bola não volta para a urna antes da próxima ser retirada.

Sejam os seguintes experimentos probabilísticos, em que n e k são parâmetros dos experimentos.

- **Experimento 1a.** Retira-se, uma por uma, um total de n bolas da urna, **com reposição**.
- **Experimento 1b.** Retira-se, uma por uma, um total de n bolas da urna, **sem reposição**.
- **Experimento 2a.** Retira-se, uma por uma, bolas da urna, **com reposição**. O experimento se encerra imediatamente após a k-ésima bola verde é retirada.
- **Experimento 2b.** Retira-se, uma por uma, bolas da urna, **sem reposição**. O experimento se encerra imediatamente após a k-ésima bola verde é retirada

Defina as seguintes variáveis aleatórias:

- **Experimentos 1a e 1b.** K = número de bolas verdes retiradas no experimento.
- **Experimentos 2a e 2b.** N = número total de bolas retiradas no experimento.

Tais variáveis aleatórias estão relacionadas com as seguintes distribuições discretas:

- **Experimentos 1a.** Distribuição binomial.
- **Experimentos 1b.** Distribuição hipergeométrica.
- **Experimentos 2a.** Distribuição binomial negativa (ou de Pascal).
- **Experimentos 2b.** Distribuição hipergeométrica negativa.

Funções massa de probabilidade

As PMFs das variáveis aleatórias associadas aos Experimentos 1a e 1b são dadas, respectivamente, por

$$p_K(k) = \binom{n}{k} p^k q^l \qquad \text{e} \qquad p_K(k) = \frac{\binom{g}{k}}{\binom{t}{n}} \binom{r}{l},$$

para $k \in \mathbb{N}$, e as PMFs das variáveis aleatórias associadas aos Experimentos 2a e 2b são dadas, respectivamente, por

$$p_N(n) = \binom{n-1}{k-1} p^k q^l \qquad \text{e} \qquad p_N(n) = \frac{\binom{g}{k-1}}{\binom{t}{n-1}} \frac{g-k+1}{t-n+1} \binom{r}{l},$$

para $n \in \mathbb{N}$.
Nas expressões acima,

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!\,(n-k)!}, & \text{se } 0 \leq k \leq n, \\ 0, & \text{caso contrário,} \end{cases}$$

denota o coeficiente binomial,

$$l \triangleq n - k,$$

e

$$p = \frac{g}{t} \qquad \text{e} \qquad q = \frac{r}{t}$$

denotam as proporções originais de bolas verdes e vermelhas, respectivamente.

Médias

As médias das variáveis aleatórias associadas aos Experimentos 1a e 1b são, respectivamente,

$$\mathrm{E}[K] = np \qquad \text{e} \qquad \mathrm{E}[K] = np,$$

e das variáveis aleatórias associadas aos Experimentos 2a e 2b são, respectivamente,

$$\mathrm{E}[N] = \frac{k}{p} \qquad \text{e} \qquad \mathrm{E}[N] = k\frac{t+1}{g+1}.$$

Variâncias

As variâncias das variáveis aleatórias associadas aos Experimentos 1a e 1b são, respectivamente,

$$\mathrm{var}[K] = npq \qquad \text{e} \qquad \mathrm{var}[K] = npq\frac{t-n}{t-1},$$

e das variáveis aleatórias associadas aos Experimentos 2a e 2b são, respectivamente,

$$\mathrm{var}[N] = k\frac{q}{p^2} \qquad \text{e} \qquad \mathrm{var}[N] = k\frac{(t-g)(t+1)(g+1-k)}{(g+1)^2(g+2)}.$$

Exercícios

1. Escreva uma função que implementa uma única realização da variável aleatória K do Experimento 1a, para (t, g, n) genéricos. Em seguida, utilize as funções escritas com $(t, g, n) = (8, 4, 6)$ e com $(t, g, n) = (100, 20, 95)$.
   - Plote figuras contendo a função massa de probabilidade teórica, bem como aquela obtida via simulação de Monte Carlo.
   - Calcule os valores teóricos da média e da variância, bem como aqueles obtidos via simulação.
2. Repita o exercício anterior, agora considerando o Experimento 1b.
3. Escreva uma função que implementa uma única realização da variável aleatória N do Experimento 2a, para (t, g, k) genéricos. Em seguida, utilize as funções escritas com $(t, g, k) = (8, 6, 4)$ e com $(t, g, k) = (100, 20, 20)$.
   - Plote figuras contendo a função massa de probabilidade teórica, bem como aquela obtida via simulação de Monte Carlo.
   - Calcule os valores teóricos da média e da variância, bem como aqueles obtidos via simulação.
4. Repita o exercício anterior, agora considerando o Experimento 2b.