MAE221 - Probabilidade I

Prof. Fábio Machado

Lista 3 - simulações

1. O problema do **colecionador de cupons** pergunta: se amostras unitárias (com reposição) são selecionadas repetidamente e independentemente de uma coleção de n itens distintos ("cupons" ou "figurinhas"), qual é a esperança de X, número de selecões necessárias para que cada item seja selecionado pelo menos uma vez? Vamos decompor a contagem em unidades mais simples. Seja $G_1 = 1$ e, para $k = 2, \dots, n$, seja G_k o número de seleções necessárias para passar de $k-1$ para k itens distintos. Assim, com $G - 1 = 1$ temos que $X = G_1 + \cdots + G_n$. Observe que na segunda seleção, a probabilidade de termos um item diferente do primeiro, é $(n-1)/n$. Lembrando da independência entre as várias seleções e do fato que os itens selecionados serão recolocados na coleção, percebe-se que $G_2 \sim \mathcal{G}((n-1)/n)$. Em geral, podemos concluir $G_i \sim \mathcal{G}((n-(i-1))/n)$ e por consequência, que $\mathbb{E}[G_i] = n/(n-(i-1))$. Assim

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} G_i] = \sum_{i=1}^{n} \mathbb{E}[G_i] = n \sum_{i=1}^{n} (1/(n-(i-1)))$$

Se lembrarmos que (veja série harmônica)

$$\sum_{i=1}^{n} \frac{1}{i} \approx \int_1^n \frac{1}{x} dx = \ln n,$$

concluiremos que $\mathbb{E}[X] \sim n \ln n$.

2. Outra questão interessante, ainda no problema do **colecionador de cupons**, é sobre Y, o número de itens distintos observados em t seleções de uma lista de n itens distintos. Podemos tomar X como $I_1 + \cdots + I_n$, onde para $k = 1, \dots n$

$$I_k = \begin{cases} 1, & \text{se o item } k \text{ está entre os } t \text{ itens selecionados} \\ 0, & \text{se o item } k \text{ não está entre os } t \text{ itens selecionados} \end{cases}$$

Assim, temos que

$$\mathbb{E}[I_k] = P[\text{item } k \text{ esteja presente entre os } t \text{ itens}] = 1 - (1 - \frac{1}{n})^t$$

e como consequencia

$$\mathbb{E}[Y] = \mathbb{E}[I_1 + \cdots + I_n] = n[1 - (1 - \frac{1}{n})^t]$$

.

Esta solução nos remete a outro problema prático interessante. Suponha uma reserva florestal onde habitam onças nativas. A administração da reserva deseja estimar, ao final de cada ano, o número de onças que vivem na reserva. Para isto eles colocam algumas camêras que enviam fotografias para uma central, cada vez que um animal de grande porte passa em seu raio de ação. Durante uma semana foram enviadas várias fotografias sendo 100 destas, de onças. A análise destas $t = 100$ fotos, indicou que se tratavam de $d = 50$ onças distintas. Usando a fórmula recém obtida

$$d \approx \mathbb{E}[Y] = n[1 - (1 - \frac{1}{n})^t]$$

com $d = 50$ e $t = 100$. Precisamos resolver o número a esquação acima para obter uma estimativa do número de onças na reserva florestal.

3. Suponha um album com 10 figurinhas. Faça simulações em número suficiente para estimar a distribuição de X.

a) Apresente o gráfico para a distribuição de probabilidades da variável X estimada pelas suas simulções;

b) Apresente os valores encontrados para *mínimo, mediana, média, máximo* da variável aleatória X.

c) Baseado no que foi feito para encontrar $\mathbb{E}[X]$, calcule o $DP[X]$ e compare este valor com a estimativa baseada nas simulações.

4. Ainda supondo um álbum de 10 figurinhas, Faça simulações em número suficiente para estimar a distribuição de Y para $t = 25$.

   a) Apresente o gráfico para a distribuição de probabilidades da variável Y estimada pelas suas simulções;

   b) Apresente o gráfico para $\mathbb{E}[Y]$ em função de t.

   c) Apresente os valores encontrados para *mínimo, mediana, média, máximo* da variável aleatória Y.

4. O **problema do aniversário** é um clássico da probabilidade. O matemático Richard von Mises (irmão do economista Ludwig von Mises), em 1939, perguntou: "Quantas pessoas devem estar em uma sala para que a probabilidade de haver pelo menos duas pessoas fazendo aniversário no mesmo dia (sem levar em consideração o ano de nascimento) seja superior a 0,5?"

   a) Assuma todos os anos tendo 365 dias e considere uma distribuição uniforme sobre todos estes 365 dias para ser o dia do nascimento de uma pessoa escolhida ao acaso em uma multidão. Faça simulações em número suficiente para estimar esta probabilidade para $n = 20, 30, ..., 100$ pessoas.

   b) Encontre o menor número de pessoas que resolve a dúvida de Richard Von Mises.