Trabalho 3 - Introdução a Modelagem Computacional

Lucas Müller e Pedro Henrique Santos

Julho 2018

Sumário

1 Modelo de simulação da urna de Polya

Modelou-se uma urna U com N cores diferentes, representadas por letras. A cada vez, retira-se uma bola ao acaso e sua letra é observada, recolocando-a na urna juntamente com mais k bolas da mesma letra. Tal procedimento é denominado como um evento. Alguns modelos computacionais foram criados para analisar algumas questões.

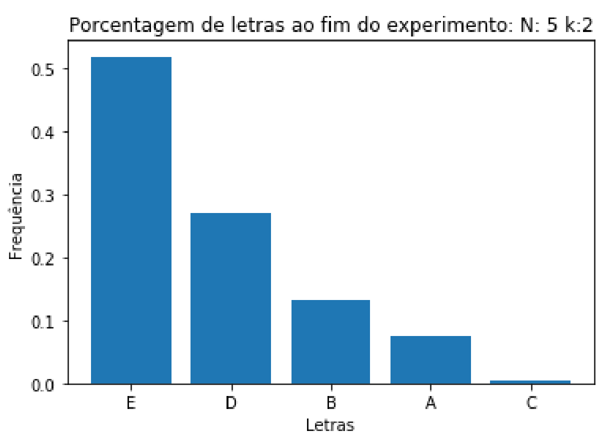
Todos os testes foram realizados com a semente = 1, para que se consiga obter um bom parâmetro de comparação entre os testes.

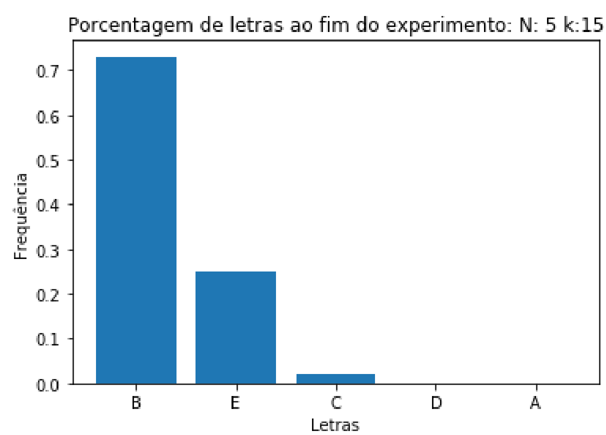
1.1 Para diferentes valores de N e k e interferências dos mesmos ao longo do experimento

Como primeiro experimento, analisou-se os parâmetros N e k, que representam,respectivamente, a quantidade de letras iniciais e o número de cópias que serão recolocadas na urna ao fim de um evento.

Para essa análise, plotou-se gráficos com a distribuição das N bolas após 10.000 eventos. Foram utilizados 4 valores de N e k, que foram: N = [5,10,20,25] e k = [2,5,10,15].

Nota-se que para valores superiores de "k" a diferença na distribuição das letras se torna maior, o que faz bastante sentido, uma vez que quanto maior o "k", a bola que é sorteada incialmente tem muito mais chance de ser sorteada novamente, já que é inserida "k" cópias dela na urna. Os gráficos abaixam mostram claramente essa disparidade.





Já para o parâmetro N, percebe-se que quanto maior, maior a chance da distribuição ficar um pouco mais uniforme. Isso se dá pois após a primeira bola ser sorteada, ainda existe uma maior probabilidade de outras bolas, que não a escolhida, de serem escolhidas no próximo evento, já que a quantidade de bolas com letras diferentes é maior.

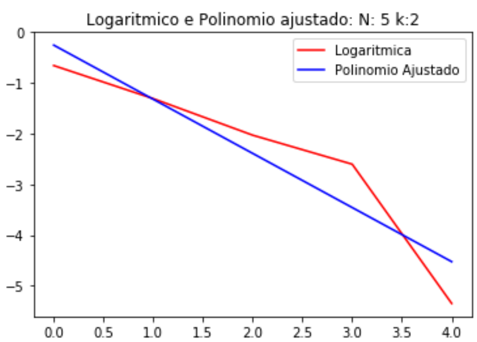
Com intuito de deixar o relatório mais legível, nem todos os gráficos serão colocados aqui, porém todos podem ser vistos na pasta em anexo.

1.2 Lei de Zipf e coeficiente angular para cada valor de N e k testados

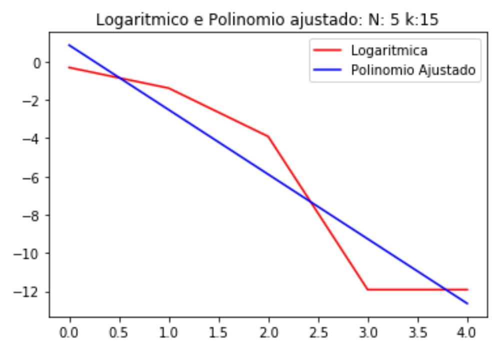
A partir dos valores obtidos nos experimentos, foi "criado" um polinômio de grau 1, que segue a equação de uma reta

$$y = ax + b \tag{1}$$

Onde deseja-se encontrar o coeficiente angular, que seria o termo que determina a inclinação da reta. Em outras palavras o "a". Percebe-se que o modelo segue a lei de Zipf, uma vez que tal polinômio apresenta comportamento linear, com um dado coeficiente angular.



(a) Coeficiente Angular: -1.068



(b) Coeficiente Angular: -3.374

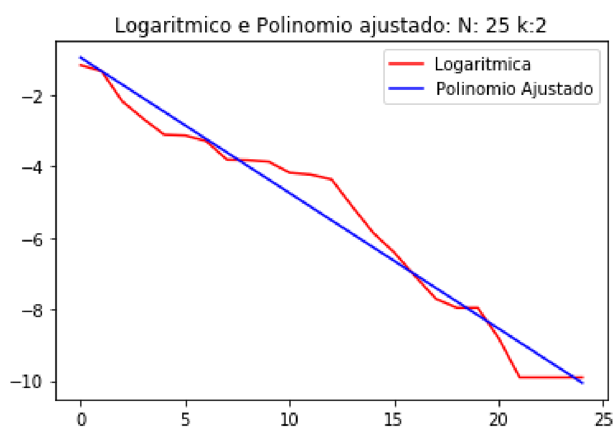


Figure 1: Coeficente Angular: -0.379

N \ k	2	5	10	15
5	-1.068	-1.597	-3.164	-3.374
10	-0.970	-0.644	-1.335	-1.499
20	-0.342	-0.485	-0.612	-0.517
25	-0.379	-0.414	-0.483	-0.531

Table 1: Coeficiente angular da reta y=ax+b

1.3 Relação N/k e coeficiente angular

Feitos os experimentos, percebe-se que possui uma relação entre o fator N/k e o coeficiente angular. Quanto menor o fator N/k, maior a disparidade entre a quantidade de cada letra na urna ou seja, poucos com muito e muitos com pouco, portanto a inclinação da reta deve ser maior, indicando um coeficiente angular maior, em módulo. O que faz bastante sentido, uma vez que quanto maior o k, menor o fator N/k e maior o desequilíbrio entre as quantidades de letras, como já visto anteriormente. Uma análise equivalente pode ser feita ao N.

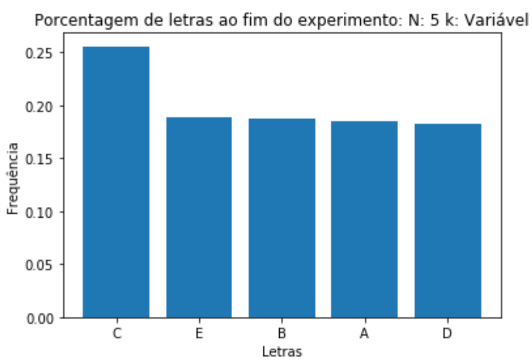
1.4 Modelo de urna no qual os valores de k são inversamente proporcionais ao número de bolas da letra escolhida

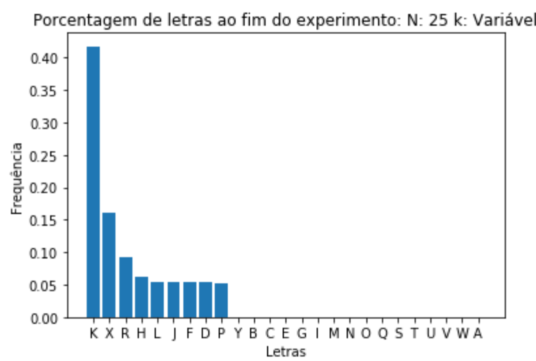
Foi implementado um modelo de urna no qual os valores de k são inversamente proporcionais ao número de bolas da letra escolhida. Foram feitos testes com os mesmos valores de N e k e foi possível identificar que no geral, a distribuição das letras tende a ficar mais uniforme, ou seja, a discrepância entre as quantidades de cada letra é bem menor, isso fica ainda mais evidente em gráficos com valores de N menores.

O cálculo do valor do k foi feito da seguinte maneira:

$$k = \frac{1}{R} \tag{2}$$

onde R é a razão entre a quantidade de letras da bola escolhida dentro da urna e a quantidade total de bolas na urna. Isso faz com que k seja inversamente proporcional a quantidade de bolas da letra escolhida na urna, ou seja, quanto mais bolas de uma determinada letra na urna, menor o k, e mais lentamente ocorrerá aumento dessa letra na urna, e vice versa.





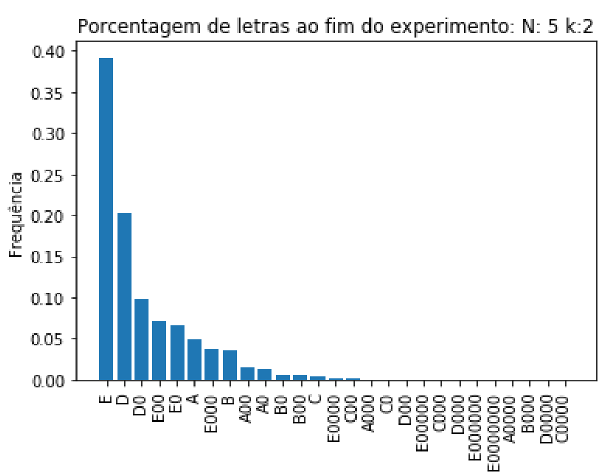
1.5 Modelo adaptado com duas urnas

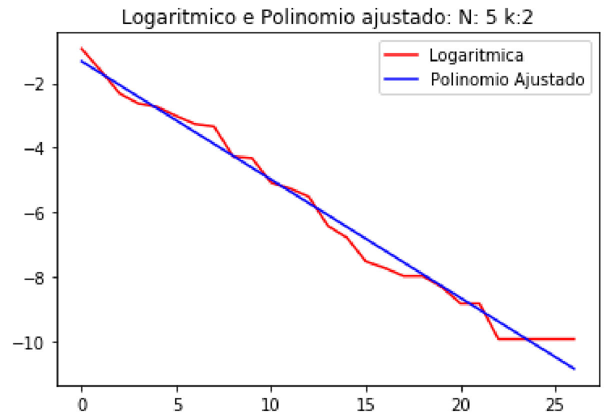
Foi feito um modelo modificado no qual utiliza-se duas urnas, U_1 e U_2 com as seguintes regras. Retira-se ao acaso uma bola de cada urna e observa-se as letras. Caso sejam iguais, cada urna recebe k bolas, senão a bola retirada da U_1 é devolvida para U_2 e vice versa. Nesse caso utilizou-se o valor de N = [5,25] e k = [2].

Os gráficos de barras mostram nitidamente que a discrepância nos valores é muito alta, ou seja, ambas as urnas ficam com muitas bolas de apenas uma letra. Isso faz sentido, uma vez que apenas quando as bolas retiradas são iguais elas são multiplicadas e reinseridas nas urnas, o que leva a urna a ficar majoritariamente com bolas de uma letra. É interessante notar que em algumas ocasiões, uma das urnas ficam sem bolas de determinada letra, o que se deve ao fato de serem eventualmente trocadas quando bolas de letras diferentes são sorteadas.

1.6 Modelo de inovação

Foi feito um modelo de invovação, que consiste no seguinte método: Quando uma bola com uma letra for retirada pela primeira vez, k cópias dessa bola é inserida na urna, como antes, porém são criadas Q novas bolas derivadas da bola sorteada, onde o parâmetro Q é a quantidade de inovações criadas a cada possível evento. Caso a bola já tenha sido retirada, o método é o mesmo dos experimentos anterior.
Como utilizou-se de letras para representar as cores das bolas, foi instituído a criação das inovações como a letra + um número. Por exemplo, caso a bola que tenha sido sorteada pela primeira vez foi a B, as inovações derivadas dessa bola serão nomeadas de B00, B01 e assim sucessivamente. Foram feitos testes para alguns valores de N,k e Q. Porém com valores altos de Q ou k, a quantidade de inovações fica muito alta, e o gráfico fica ilegível. Os testes podem ser facilmente reproduzidos trocando os parâmetros no código. A fim de análise, o teste feito foi com N=5, k=2 e Q=1.





É possível notar que tal modelo permanece seguindo a lei de Zipf, e possui um coeficiente angular de -0.3654. Percebe-se também que algumas inovações "deram certo", no exemplo do gráfico acima, a D0 por exemplo, foi a terceira letra que mais aparece na urna, representando em média 10% de toda a urna, seguida pela inovação E00 e E0, respectivamente. Importante salientar que quando maior o número de inovações, maior a "concorrência", e mais difícil de uma delas se sobrepor sobre outras.