

**Universidade de São Paulo**  
**Faculdade de Filosofia, Letras e Ciências Humanas**  
**Departamento de Ciência Política**  
**FLS 5028**  
**Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política**  
**FLP0406**  
**Métodos e Técnicas de Pesquisa em Ciência Política**  
**1º semestre / 2018**  
**Prof. Glauco Peres da Silva**  
**LISTA DE EXERCÍCIOS 05**  
**Data de entrega: 16/04/2018 (noturno) e 18/04/2018 (vespertino)**

**Exercício 1 (2 pontos)**

As alternativas abaixo apresentam conceitos importantes sobre famílias de distribuição de probabilidade. Para cada uma delas marque “verdadeiro” (V) ou “falso” (F). Para as assinaladas como “falsas” aponte e corrija as inconsistências encontradas.

- a) (F) A Distribuição Binomial é uma distribuição para variáveis contínuas e necessita pelo menos três resultados possíveis para ser executada.

A Distribuição Binomial é uma distribuição para variáveis discretas e necessita somente, dois resultados possíveis para ser executada.

- b) (F) A Distribuição Binomial Negativa conta o número de sucessos em um número fixo de provas de Bernoulli.

Essa definição corresponde à Distribuição Binomial comum, a Binomial Negativa conta o número de provas de Bernoulli necessário para obter um número fixo de falhas.

- c) (F) Uma suposição importante feita pelos pesquisadores que assumem que seus dados uma Distribuição de Poisson é a de que, para pequenos intervalos de tempo, a probabilidade de uma chegada (evento) é proporcional ao tempo de espera. Essa distribuição possui apenas o parâmetro  $\lambda$ , que representa a média, enquanto a variância é igual a  $\lambda/2$ .

Média e variância são iguais a  $\lambda$ .

- d) (V) A Distribuição Beta atribui probabilidade 1 para um intervalo finito, dessa forma, ela é geralmente utilizada para modelar proporções, uma vez que elas são representadas por valores restritos entre 0 e 1.

## Exercício 2 (4 pontos)

A seguir, serão apresentados três problemas de pesquisa. Leia atentamente cada um dos problemas, de modo a responder corretamente os itens propostos.

- I. Determinado(a) pesquisador(a) deseja saber se um prefeito será ou não removido do cargo em decorrência de um processo de cassação.
- II. Determinado(a) pesquisador(a) pretende analisar o número de projetos apresentados pelos parlamentares na Câmara dos Deputados em determinado ano. Assuma que as ocorrências da variável são independentes.
- III. Tendo em mãos uma lista de prefeitos cassados nos últimos 20 anos no estado de São Paulo, um(a) pesquisador(a) deseja conhecer o tempo médio, em meses, que um prefeito cassado se mantém no cargo. Assuma que o tempo no cargo seja contabilizado em dias, desse modo, 45 dias equivalem a 1,5 mês (todos os meses possuem 30 dias).

a) Para cada uma das situações, indique: I. qual é a variável dependente; II. quais são os possíveis valores que a variável pode assumir; III. se a variável é contínua ou discreta.

**(1 ponto)**

### Problema de pesquisa I

A variável dependente refere-se à destituição ou não do prefeito do cargo podendo, portanto, assumir apenas dois valores, são eles, sim e não, comumente categorizados como, respectivamente, 1 e 0. É uma variável discreta, visto que seu espaço amostral compreende duas categorias e valores decimais não fazem sentido.

### Problema de pesquisa II

A variável dependente refere-se ao número de projetos apresentados na Câmara dos Deputados por ano, podendo assumir valores iguais ou maiores que zero. É uma variável discreta, visto que valores decimais não fazem sentido (não é possível apresentar 0,5 projeto).

### Problema de pesquisa III

A variável dependente refere-se ao número de meses que um prefeito se mantém no cargo.

Visto que a variável dependente foi mensurada em meses, ela pode assumir valores que vão do 0,0333 (caso o prefeito fique apenas um dia no cargo) até 48 meses. É, portanto, uma variável contínua, visto que compreende valores decimais.

b) Para cada uma das situações indique:

I. Uma distribuição probabilística apropriada, demonstrando porque ela se adequa a cada um dos problemas de pesquisa. Dica: verifique se a variável é discreta ou contínua, conforme item anterior;

II. Outra variável que possua a mesma distribuição e, mais uma vez, explique porque a variável escolhida pode ser modelada pela distribuição probabilística selecionada

**(2 pontos)**

Problema de pesquisa I

A distribuição probabilística apropriada é a Binomial. Dado que a variável pode assumir apenas dois valores (é discreta), os quais são mutuamente exclusivos (não é possível que o prefeito seja destituído e mantido no cargo ao mesmo tempo) e com ocorrências independentes, a distribuição Binomial deve ser a escolhida. Diversas outras variáveis podem ser modeladas a partir da distribuição Binomial: votar ou não votar em determinada eleição (assumindo que os cidadãos não são influenciados pela decisão de votar ou não de outros, ou seja, assumindo independência); a jogada de uma moeda (cara ou coroa); o sexo de bebês (menina ou menino), entre tantas outras.

Problema de pesquisa II

A distribuição probabilística apropriada é a de Poisson. Visto que é uma variável discreta que se realiza no tempo e as ocorrências são independentes (é assumido que um parlamentar não apresenta um projeto só porque outro também apresentou, por exemplo), as ocorrências da variável dependente podem ser modeladas a partir da distribuição de Poisson. Outras variáveis que obedecem à distribuição de Poisson: número de chamadas recebidas por minuto em um callcenter; número de dias com chuva em um mês; número de gols em uma partida de futebol, número de pacientes que dão entrada em hospitais por hora, entre outras.

Problema de pesquisa III

Uma distribuição probabilística adequada é a normal. A média de tempo que um prefeito se mantém no cargo é uma variável contínua e, como o interesse do pesquisador é pela

média, esta pode ser modelada a partir da distribuição normal. Quando o interesse é pela média, diversas distribuições que não possuem distribuição simétrica podem ser consideradas normais. Podem ser modeladas a partir da distribuição normal: renda salarial média; proporção dos brasileiros que aprovam o governo; valor médio gasto por candidato em campanhas eleitorais, entre outras.

c) Em relação ao problema de pesquisa esboçado no item I, assuma que a probabilidade de afastamento do prefeito do cargo em decorrência de uma cassação seja conhecida e equivalha a 0,68. Qual é a probabilidade da não ocorrência do afastamento do prefeito durante seu mandato? Indique a distribuição de probabilidade apropriada, sua fórmula matemática correspondente e substitua os valores fornecidos na questão na fórmula para obter o resultado. **(1 ponto)**

A distribuição apropriada é a Bernoulli.

Sucesso (ocorrência do fenômeno, codificado como 1) = 0,68; Fracasso (não ocorrência do fenômeno, codificado como 0) = 0,32.

$$\Pr(Y_i = 1) = \pi = 0,68$$

$$\Pr(Y_i = 0) = 1 - \pi = 1 - 0,68 = 0,32$$

ou

$$Y_i \sim f_{\text{Bern}}(y_i | \pi) = \pi^{y_i} (1 - \pi)^{1 - y_i} \text{ para } y_i = 0, 1, 0 \text{ em caso contrário}$$

$$Y_i \sim f_{\text{Bern}}(y_i | \pi) = \pi^0 (1 - 0,68)^{1 - 0}$$

$$Y_i \sim f_{\text{Bern}}(y_i | \pi) = 1 (0,32)^1$$

$$Y_i \sim f_{\text{Bern}}(y_i | \pi) = 0,32$$

### Questão 3 (4 pontos)

1. Para cada uma das variáveis a seguir, indique a função de distribuição adequada.

Justifique sua resposta:

a. Resultado de um sorteio da Mega Sena;

Hipergeométrica. O sorteio da mega sena seleciona 6 números, em que a ordem não importa, mas é sem reposição.

b. Altura média do brasileiro;

Normal. No intervalo possível de alturas, há uma infinidade de tamanhos diferentes, o que permite identificar a variável como contínua.

- c. Porcentagem de votos obtida por cada partido em cada urna nas eleições para deputado federal no Brasil;

Beta. A variável em referência é uma porcentagem - portanto é contínua e está limitada entre 0 e 1.

- d. Duração de uma guerra;

Aqui depende de como se considera o tempo. Se a duração for medida em anos - em números inteiros - uma Poisson atende. Se for medida em dias ou em frações ainda menores, a variável se torna contínua. Nesse caso, uma Gama se adequa.

- e. Número de chamadas necessárias até completar 20 questionários.

Binomial negativa. É como a binomial, mas neste caso contam-se o número de ocorrências até determinado evento.

2. Para cada uma das funções de distribuição a seguir, indique quais são os parâmetros, com as respectivas definições (média, variância, etc.):

- a. Binomial negativa;

Os parâmetros aqui são o número de sucessos, de falhas e a probabilidade de sucesso.

- b. Uniforme discreta;

Os parâmetros a e b indicam o intervalo no qual a variável ocorre.

- c. Gama;

Os parâmetros representam a escala e o shape (formato)

- d. Lognormal;

Média e variância são os parâmetros

- e. Chi-quadrado.

O único parâmetro são os graus de liberdade

#### Questão 4 - Pós-graduação

Acesse o seguinte site:

<http://www.math.uah.edu/stat/dist/index.html>.

Procure em “Apps” pelo “Special distribution simulator”

Cuidado: Há um link chamado Special distribution calculator!

Vamos fazer um exercício para entender o que significa a convergência de uma série em outra.

No box superior, escolha a distribuição Poisson. Defina  $\lambda=3$ .

- a) Dê uma interpretação para o valor de  $\lambda$ .

O parâmetro de uma distribuição de Poisson significa sua média e sua variância. Neste caso, ambas são 3.

No box a direita, escolha “Stop: 100”. Este comando pedirá ao software que simule 100 observações de uma variável desta distribuição.

- b) Copie e cole o gráfico resultante em sua resposta. Comente os seus achados.

A distribuição encontrada não é idêntica à distribuição esperada para uma Poisson com lambda igual a 3.

Repita o procedimento acima, mas agora ajuste no box a direita para “Stop: 1000”

- c) Copie o novo gráfico aqui. Comente os resultados atuais em comparação com o resultado anterior.

Com o aumento do n, a distribuição da função vai se aproximando da distribuição esperada.

- d) Para que a distribuição desta variável convirja para uma distribuição normal, quais modificações você precisa fazer? Por que? Faça as modificações necessárias e simule a nova distribuição (copie o novo gráfico se quiser).

O parâmetro precisa tender ao infinito para que a convergência ocorra