**Universidade de São Paulo**

**Faculdade de Filosofia, Letras e Ciências Humanas**

**Departamento de Ciência Política**

**FLP0406 – Métodos e Técnicas de Pesquisa em Ciência Política**

1º semestre / 2019

Prof. Glauco Peres da Silva

**Laboratório 7 - Teorema do Limite Central**

No laboratório de hoje, tentaremos entender os conceitos relacionados ao Teorema do Limite Central.

Para isso, inicialmente, considere o gráfico que consta da aula anterior no qual há informações sobre diversas distribuições de probabilidade (no moodle, está no link *Famílias de Distribuição (Casella e Berger)*). Baseado nele, quais distribuições você poderia utilizar para modelar as seguintes variáveis:

1. Um lançamento de um dado?

Uniforme discreta, Binomial, Bernoulli.

1. Duração de uma guerra?

Depende se considerarmos como discreta ou contínua. Também do número de casos analisados, se for grande o suficiente podemos aproximar para contínua.

Normal (porém nesse caso estamos assumindo uma parcela de dados fictícia), Quiquadrado

1. Número de chamadas necessárias até completar 20 questionários?

Binomial.

Número de deputados reeleitos no estado *s*?

Pode variar de 0 até o número de deputados (assumindo que todos tentaram se reeleger).

Poisson (porém usa-se muito em casos concentrados em 0).

1. Variação no peso do indivíduo *i* no ano *t*? (Não está assumindo qual é o número de indivíduos)

Normal.

1. Duração da vida (humana)?

Pode-se pensar em normal, mas é provável que tenha picos, então pode não corresponder inteiramente.

1. Proporção da renda de uma pessoa *i* gasta em queijo?

Beta.

Cada uma destas variáveis foi modelada por uma distribuição distinta. De que maneira, então, o Teorema do Limite Central nos ajuda? Para essa tarefa, acesse inicialmente o site:

<http://onlinestatbook.com/stat_sim/sampling_dist/>

Clique no botão “Begin” a esquerda.

Na tela que surgirá, escolha a distribuição “Uniforme” na caixa suspensa superior à direita. Aparecerá uma distribuição em cinza com as estatísticas (Média, mediana, desvio padrão, assimetria e curtose) à esquerda. Nos três gráficos abaixo, serão apresentados a seleção amostral (1º) e a média amostral (2º e 3º, quando for o caso). Para ilustrar, clique em “Animated”, botão que está à direita do 1º gráfico.

Assim, selecione a média (Mean) no segundo e terceiro gráficos, mas escolha N=5 no 2º e N=25 no 3º gráfico.

1. Simule uma amostra de 5 dados. Registre seus resultados estimados para a média. Clique novamente no botão “5” até que sua amostra seja de tamanho 20. Registre os resultados estimados para a média. Compare os quatro resultados obtidos (média n=5, 5 amostras; n=5, 20 amostras; n=25, 5 amostras; n=25 20 amostras);

Amostra de 5 dados:

2º

Reps=5

mean=15.05

median=14.00

sd=4.36

skew=0.28

kurtosis=-1.77

3º

Reps=5

mean=16.33

median=16.00

sd=0.79

skew=1.00

kurtosis=2.62

Amostra de 20 dados:

2º

Reps=20

mean=13.48

median=12.50

sd=4.13

skew=0.53

kurtosis=-0.67

3º

Reps=20

mean=15.94

median=16.00

sd=1.62

skew=-0.16

kurtosis=0.71

1. Comente os resultados obtidos acima. Qual deles está mais próximo da média populacional e por que?

O resultado mais próximo é a distribuição com 20 amostra com tamanho 25. Isso se deve pela lei dos números grandes. Essa lei diz que o aumento do n tende a diminuir a distância entre os parâmetros da média e os verdadeiros parâmetros da população, como o média.

Apague os resultados obtidos até aqui, clicando em “Clear lower 3”.

1. Simule novamente as estimativas da média, mas agora usando 100 amostras. Para isso clique no botão “5” vinte vezes. Antes, marque “fit normal” nos dois últimos gráficos.
2. Compare os resultados gráficos obtidos acima em termos precisão. Qual dos dois procedimentos é mais preciso? Por que?
3. Encontre uma medida de precisão nos dois gráficos e compare com a resposta dada no item acima.

Apague os resultados obtidos até aqui e simule novamente as estimativas de média, agora usando 10.000 amostras.

1. Compare os novos resultados com os obtidos em d) e e). Comente.
2. Em uma pesquisa baseado em dados quantitativos, em geral, temos apenas uma única amostra. Diga, então, como o Teorema contribui para a realização do trabalho.

Ela garante que temos uma distribuição normal, e se tivéssemos uma amostra, podemos dizer que veio de uma distribuição normal. O IC diz que a variável foi normalizada.