

Estatística não paramétrica

Fernando Lucambio

Departamento de Estatística
Universidade Federal do Paraná

28 de fevereiro de 2024

Estrutura:

- ▶ <http://leg.ufpr.br/~lucambio/CE050/20241S/CE050.html>
- ▶ quarta-feira 19h, sexta-feira 20:45h
- ▶ 4 trabalhos assíncronos

Bibliografia:

- ▶ <http://leg.ufpr.br/~lucambio/Nonparam/Nparam.html>
- ▶ outras referências mencionadas na página da disciplina

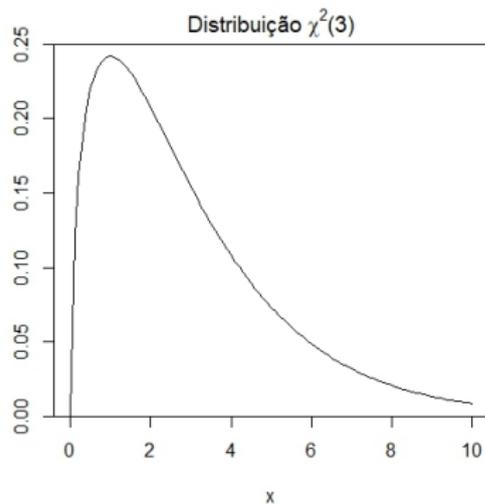
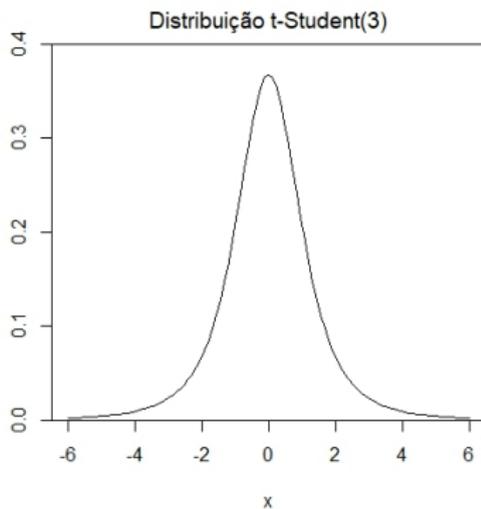
Em todos os problemas de inferência estatística considerados, assumimos que a distribuição da variável aleatória amostrada seja conhecida a menos, talvez, para alguns parâmetros. Na prática, entretanto, a forma funcional da distribuição é raramente ou nunca conhecida. Por conseguinte, é desejável conceber alguns procedimentos que estejam livres desta hipótese relativa à distribuição.

Para entender a ideia de estatística não-paramétrica, primeiro requeremos uma compreensão de conceitos da estatística básica paramétrica. Conceitos elementares introduzem o conceito de teste de significância estatística com base na distribuição amostral de uma estatística particular. Em resumo, se tivermos um conhecimento básico da distribuição subjacente de uma variável, poderemos fazer previsões sobre como, em amostras repetidas de tamanho igual, essa estatística específica se comportará, isto é, como será distribuída.

Estudamos aqui alguns procedimentos que são comumente referidos como métodos sem distribuição ou não paramétricos. O termo livre de distribuição refere-se ao fato de que nenhuma suposição é feita sobre a distribuição subjacente, exceto que a função de distribuição sendo amostrada seja absolutamente contínua ou puramente discreta. O termo não paramétrico refere-se ao fato de não haver parâmetros envolvidos no sentido tradicional do termo parâmetro utilizado até o momento.

Grosseiramente falando, um procedimento não-paramétrico é um procedimento estatístico que possui certas propriedades desejáveis que mantêm suposições relativamente leves em relação às populações subjacentes das quais os dados são obtidos.

Nos dois exemplos seguintes mostramos distribuições conhecidas que são livres de parâmetros.



O desenvolvimento repetido e contínuo de procedimentos estatísticos não paramétricos nas últimas décadas deve-se às seguintes vantagens de técnicas não paramétricas:

- ▶ Métodos não-paramétricos exigem poucas suposições sobre as populações subjacentes das quais os dados são obtidos. Em particular, os procedimentos não paramétricos abandonam a suposição tradicional de que as populações subjacentes sejam normais.
- ▶ Os procedimentos não paramétricos permitem que o usuário obtenha p-valores exatos para testes, probabilidades de cobertura exatas para intervalos de confiança, taxas exatas de erros experimentais para procedimentos de comparação múltipla e probabilidades exatas de cobertura para faixas de confiança sem confiar nas suposições de que as populações subjacentes sejam normais.

- ▶ As técnicas não paramétricas são frequentemente, embora nem sempre, mais fáceis de aplicar do que as suas contrapartes teóricas normais.
- ▶ Os procedimentos não paramétricos são geralmente muito fáceis de entender.
- ▶ Embora, à primeira vista, a maioria dos procedimentos não-paramétricos pareça sacrificar muito as informações básicas nas amostras, investigações de eficiência teórica mostraram que esse não é o caso. Normalmente, os procedimentos não-paramétricos são apenas ligeiramente menos eficientes do que os seus concorrentes de teoria normal quando as populações subjacentes são normais e podem ser moderadamente ou muito mais eficientes que os concorrentes quando as populações subjacentes não são normais.

Mas têm desvantagens:

- ▶ Métodos não paramétricos são relativamente insensíveis a observações distantes.
- ▶ Os procedimentos não paramétricos são aplicáveis em muitas situações em que os procedimentos teóricos normais não podem ser utilizados. Muitos procedimentos não-paramétricos exigem apenas as classificações das observações em vez da magnitude real das observações, enquanto os procedimentos paramétricos exigem as magnitudes.
- ▶ Nem todos os procedimentos desenvolvidos na estatística paramétrica podem ser aplicados à estatística não-paramétrica.

Estrutura:

I. Estimação não paramétrica

I.1. Estimação de densidades

II. Problemas de amostra única

II.1 Testes de bondade de ajuste

II.2 O problema de posição

II.3 Testes de aleatoriedade

II.4 Métodos de reamostragem

III. Procedimentos em k amostras

III.1 Procedimentos de amostra única e com amostras pareadas

III.2 O problema geral de duas amostras

III.3 Medidas de Associação em Classificações Múltiplas

III.4 Coeficientes de correlação

IV. Regressão Não Paramétrica