Poder Estatístico

Paulo Passos

Motivação

- Você planeja realizar um estudo, mas gostaria de saber se a amostra do seu desenho experimental será capaz de fornecer evidências suficientes em favor da sua hipótese.
- Você pretende buscar financiamento para o seu projeto. Os avaliadores da comissão responsável por decidir sobre o se é ou não uma boa ideia financiar o seu projeto precisam de algo que os ajude a decidir em seu favor.

Definição de Poder Estatístico

É a probabilidade de rejeitar corretamente a hipótese nula quando esta for, de fato, falsa.

Aqui, definiremos o poder estatístico como a probabilidade associada ao evento $A = \{$ *"rejeição da falsa hipótese nula"* $\}$:

$$\mathbb{P}(A) = 1 - \beta$$

Onde β é a probabilidade do evento complementar, ou seja, $A^C = \{$ *"aceitação da falsa hipótese nula"* $\}$

Exemplo Ilustrativo

Suponha que o governo aplique um teste para averiguar o desempenho dos alunos em matemática. Com o modelo de ensino vigente, os alunos tem pontuação média de 75. A variância dessa pontuação igual a 10.

- Um pesquisador gostaria de propor um novo método, utilizado em outro país, para substituir o vigente.
- O pesquisador sugere que o método seja aplicado à uma amostra de 25 alunos como forma de averiguar se o método pode resultar em um desempenho melhor. Uma amostra de alunos do mesmo tamanho, sob o método vigente, será utilizada para comparação.
- **É desconhecido por parte do experimentador, mas o novo método resulta em uma pontuação média maior, cerca de 80, mas apresenta variância idêntica ao método vigente**.

Exemplo Ilustrativo

Para o cálculo do poder estatístico, o pesquisador **PRECISARÁ** fazer uma série de **SUPOSIÇÕES**

- Supor que o novo método tem variância igual ao anterior (verdade)
- Supor que o desempenho de ambos os teste seguem uma distribuição normal (em muitos casos, verdade)

SUPOSIÇÕES precisam de amparo! Por exemplo, de um experimento parecido conduzido em amostra da população do país de onde o método é originário.

Entendendo as Suposições

O que significa que o desempenho dos alunos sumetidos ao teste tem distribuição normal?

Significa que se toda a população de alunos fosse submetida a um dos métodos e o desempenho desses alunos fosse registrado:

- Ao sortear um desses alunos ao acaso, a probabilidade desse aluno apresentar um desempenho x específico é de $\mathbb{P}(X = x)$ onde $X \sim N(\mu, \sigma)$.
- Aqui, X representa a **variável aleatória** correspondente ao experimento de sortear ao acaso uma nota de desempenho dentro da população.
- A probabilidade associada a cada desempenho tem formato de uma função normal com parâmetros (μ, σ).

Entendendo as Suposições

Ou seja, temos:

$$\mathbb{P}(X = x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = N(\mu, \sigma)$$

E porque isso é importante? Porque pelo **teorema da soma de var. aleatórias independentes e normalmente distribuidas** temos que:

- Se sorteamos, ao acaso e com substituição, amostras dentro dessa população, a probabilidade de obtermos uma média amostral específica $\bar{x}$ é:

$$\mathbb{P}(\bar{X} = \bar{x}) = N(\mu, \frac{\sigma}{\sqrt{n}})$$

onde n é o tamanho da amostra.

Entendendo as Suposições

O teste que pretendemos aplicar envolve:

- Comparar distribuições **amostrais** de populações com distribuição normal.
- O teste irá indicar se existe uma probabilidade menor ou maior que o **nível de significância** das amostras serem provenientes de populações idênticas.

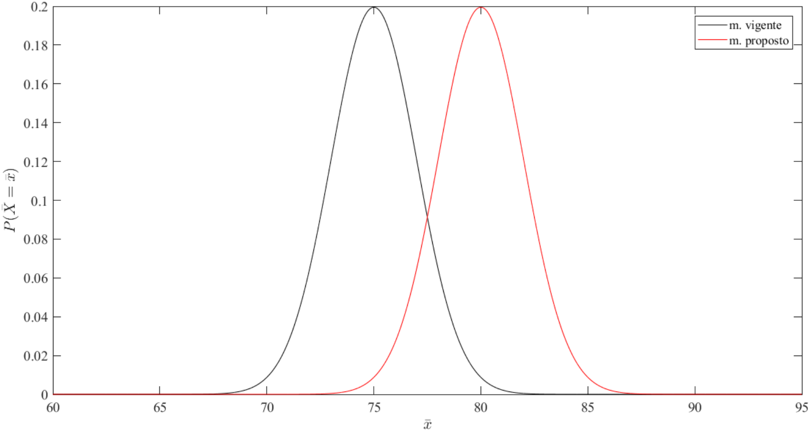
Comparando as distribuições

Podemos comparar essas distribuições visualmente com o código abaixo:

```matlab
% grafico das distribuicoes

x = 60:0.01:95;
y1 = pdf('normal',x,75,10/sqrt(25));
y2 = pdf('normal',x,80,10/sqrt(25));

figure
plot(x,y1,'k')
hold
plot(x,y2,'r')
xlabel('$\bar{x}$','interpreter','latex','FontSize',14)
ylabel('$P (\bar{X} = \bar{x})$', 'interpreter', 'latex','FontSize',14)
a = get(gca,'XTickLabel');
set(gca,'XTickLabel',a,'FontName','Times','fontsize',14)
legend('m. vigente', 'm. proposto')

% valor correspondente a alpha

p = normcdf(x,75,10/sqrt(25));
alpha_value = x(find(p > 0.95, 1));
plot(alpha_value*ones(100,1),linspace(0,max(y1)),'b--')
legend('m. vigente', 'm. proposto', '\alpha')
```

Comparando as distribuições

- Funções densidade de probabilidade.



Comparando as distribuições

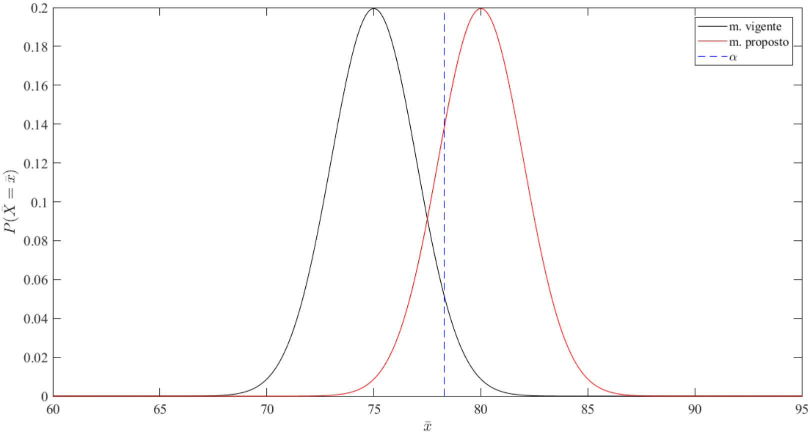
Um nivel de significância $\alpha = 0.05$ significa que:

- Se o novo método resultar num desempenho médio que seja superior (teste unicaudal) ao ponto em que a área à esquerda da distribuição soma 0.95, consideraremos que existe uma probabilidade suficientemente pequena desse valor ser oriundo de uma distribuição idêntica a do método vigente.
- Ou seja, vamos considerar que a média é, com grande probabilidade, oriunda de uma distribuição com média maior, a distribuição de um método mais eficiente.

Comparando as distribuições

Método vigente (preto) e método proposto (vermelho).

- método vigente: para $\alpha = 0.05$, temos $\bar{x} = 78.29$



Comparando as distribuições

Se considerarmos a distribuição do método método proposto, veremos que a área da distribuição do método proposto superior a $\bar{x} = 78.29$ é de 0.837.

- $\mathbb{P}(A) = 0.837$ é, portanto, o poder do teste.

Ué? Então o cálculo do poder é tirado de uma distribuição que eu como experimentador não conheço!?

- Elementar, meus caros! É isso mesmo.

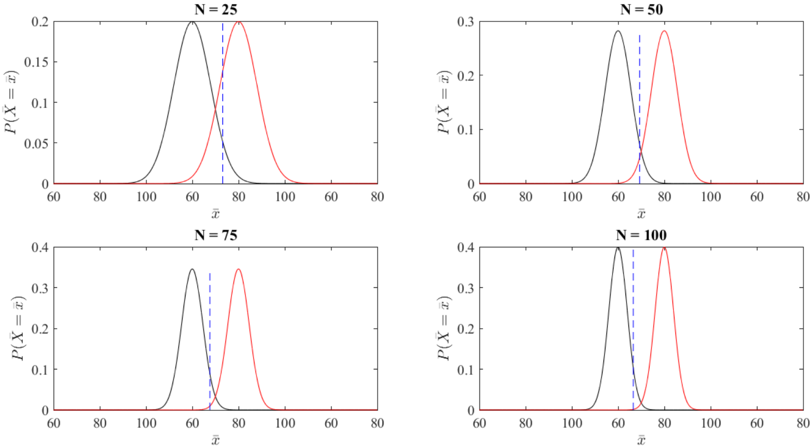
Comparando as distribuições

Como vocês podem imaginar, várias coisas podem influenciar o poder estatístico:

- A diferença entre as média das distribuições. (pode ser sugerida a partir de outros experimentos)
- A variância. (pode ser sugerida a partir de outros experimentos)
- O tamanho da amostra. (está sobre o controle do experimentador)

Comparando as distribuições

- ex.: tamanho da amostra.



Quase concluindo

Então, quando você tem experimentos semelhantes você pode sugerir que:

- Variância será igual no próximo experimento.
- Que espera-se que a diferença entre as médias seja de aproximadamente L.

Obs.: Caso você suponha que a diferença é pequena, se o cálculo do poder der um resultado elevado mesmo com uma amostra pequena. Ponto positivo!

Mas existe alguma forma de estimar esses parâmetros a serem utilizados? **Sim**.

Cálculo de intervalo de confiança (IC) para um parâmetro

O intervalo de confiânça (IC) para um parâmetro θ (ex.: $\theta = \mu$) é um intervalo $\theta_1 \leq \theta \leq \theta_2$ que contém θ com probabilidade $\mathbb{P} = \gamma$. É estimado com base em uma amostra.

O parâmetro γ é escolhido pelo experimentador, em geral 0.95 ou 0.99, depende do **risco** de estar errado.

Exemplo, para um $\gamma = 0.95$, uma em cada 20 amostras não terá o parâmetro dentro do intervalo estimado.

Cálculo de intervalo de confiança (IC) para um parâmetro

O cálculo do IC faz uso do seguinte teorema.

Teorema

Sejam $X_1, \ldots, X_N$ variáveis aleatórias normais, independentes com média μ e variância σ^2, então é assegurado:

- A variável aleatória $\bar{X} = \frac{X_1 + \ldots + X_N}{N}$ tem média μ e variância $\frac{\sigma^2}{n}$.
- A variável aleatório Z tem média 0 e variância 1. onde Z é:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}}$$

Cálculo de intervalo de confiança (IC) para um parâmetro

A variável Z é conhecida como **variável normal padrão**, tem *função de distribuição de probabilidade* conhecida:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

Assim como **função cumulativa de probabilidade** conhecida $\Phi(x)$. Essa última é que é realmente utilizada nos cálculos. Os cálculos são feitos sobre Z ao invés de $\bar{X}$ para facilitar.

Cálculo de intervalo de confiança (IC) para um parâmetro

Resumindo bastante, resolvemos para nossa icógnita c a seguinte equação:

$$\mathbb{P}(-c \leq Z \leq c) = \Phi(c) - \Phi(-c) = \gamma$$

O que equivale a resolver a seguinte equação:

$$\mathbb{P}(\bar{X} - \frac{c\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + \frac{c\sigma}{\sqrt{N}}) = \Phi(\bar{X} + \frac{c\sigma}{\sqrt{N}}) - \Phi(\bar{X} - \frac{c\sigma}{\sqrt{N}}) = \gamma$$

Ex.: Para $c = 1.96$, $\mathbb{P}(-c \leq Z \leq c) = 0.95$.

Cálculo de intervalo de confiança (IC) para um parâmetro

Por fim, perceba que o intervalo:

$$\bar{X} - \frac{c\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + \frac{c\sigma}{\sqrt{N}}$$

Tem tamanho:

$$L = 2\frac{c\sigma}{\sqrt{N}}$$

você pode determinar c de acordo com o γ de escolha. Então, sabendo o valor do desvio padrão, você pode achar o N necessário para que seu intervalo tenha o comprimento L que você achar melhor (dentro do possível).

Você pode utilizar o ponto médio desse intervalo, por exemplo, como média no cálculo do poder estatístico.

Concluindo

- Método anterior é uma estratégia bem elegante, mas precisa de uma amostra para você se basear.
- Como dito, você também pode apenas supor que o efeito é pequeno e ver se mesmo assim, sobre essa suposição, sua amostra é adequada.