CE062 - Tópicos em Biometria

Silva, J.P; Taconeli, C.A.

13 de agosto, 2019

Cálculo de Tamanho de Amostra

Introdução

Estudos estatísticos são sempre melhores quando são cuidadosamente planejados.

- O problema deve ser cuidadosamente definido e operacionalizado.
- As unidades experimentais ou observacionais devem ser selecionadas da população apropriada.
- A aleatorização deve ser feita corretamente.
- Os procedimentos devem ser seguidos corretamente.
- Instrumentos confiáveis devem ser usados para obter as medidas.

Introdução

Finalmente o estudo deve ser de tamanho adequado.

- Deve ser “grande suficiente” para que um efeito de dada magnitude que seja de significância científica seja também de significância estatística.
- Importante, também, que não seja “tão grande” para que um efeito de pouca significância científica seja estatisticamente detectável.

Introdução

O tamanho amostral é importante por razões econômicas:

- Um estudo com tamanho amostral menor que o necessário pode ser um desperdício de recursos por não ter a capacidade de produzir resultados úteis.
- Um estudo com tamanho amostral exagerado usa mais recursos que o necessário.

Num experimento envolvendo seres humanos, o tamanho amostral é questão de grande importância por razões éticas:

- Com tamanho amostral exagerado, um número desnecessário de indivíduos é exposto a um tratamento potencialmente perigoso, ou não são beneficiadas de um potencialmente benéfico.

Introdução

Há vários enfoques para o tamanho amostral:

- Pode-se especificar o comprimento desejado do intervalo de confiança e determinar o tamanho amostral que atende tal meta;
- Um dos enfoques mais populares envolve estudar o poder de um teste de hipóteses.

Esse é o enfoque que será utilizado aqui.

Introdução

De forma geral, o enfoque baseado no poder envolve os seguintes elementos:

1. Especifique um teste de hipóteses sobre um parâmetro θ (junto com um modelo probabilístico para os dados).
2. Especifique um nível de significância α do teste.
3. Especifique um *tamanho de efeito* $\tilde{\theta}$ que seja de interesse científico.
4. Obtenha valores históricos ou estimativas dos outros parâmetros necessários para cálculo da função poder do teste.
5. Especifique um valor alvo $\tilde{\pi}$ do poder do teste quando $\theta = \tilde{\theta}$.

Introdução

O poder do teste é uma função $\pi(\theta, n, \alpha, \ldots)$, em que:

- n é o tamanho amostral, e
- e a parte "$\ldots$" refere-se aos parâmetros adicionais do passo 4.

O tamanho amostral necessário é o menor inteiro n tal que $\pi(\tilde{\theta}, n, \alpha, \ldots) \geq \tilde{\pi}$.

Exemplo

Suponha que planejamos conduzir um experimento de duas amostras para comparar um tratamento com um controle.

A variável resposta é pressão sanguínea sistólica, medida com um esfigmomanômetro padrão.

Espera-se que o tratamento reduza a pressão sanguínea.

Temos um teste unilateral: $H_0 : \mu_T = \mu_C$ versus $H_1 : \mu_T < \mu_C$, em que μ_T é a pressão média para o grupo tratamento e μ_C é a pressão média para o grupo controle.

O parâmetro $\theta = \mu_T - \mu_C$ é o efeito a ser testado; e escrevemos $H_0 : \theta = 0$ e $H_1 : \theta < 0$.

Exemplo

As metas do experimento especificam que queremos detectar uma situação em que a média do tratamento é 15 mmHg menor que a do grupo controle; i.e., o tamanho do efeito é $\tilde{\theta} = -15$.

Tal efeito deve ser detectado com 80% de poder ($\tilde{\pi} = 0,80$) e nível de significância $\alpha = 0,05$.

Experiência passada com experimentos similares sugere que os dados sejam normalmente distribuídos com $\sigma = 20$ mmHg.

Usaremos um teste *t* para duas amostras (com variância combinada) e *n* igual para cada grupo.

Fórmula para Cálculo: Cenário Simples

Seja $n_i,\ i = 1, 2$, o tamanho amostral no grupo i e σ o desvio padrão (assumido igual nos dois grupos) e considere $n_1/n_2 = \kappa$ para algum κ, a razão de alocação.

Para o caso de σ conhecido, o tamanho amostral pode ser encontrado por (teste bilateral):

$$
\begin{aligned}
n_1 &= \kappa n_2 \\
n_2 &= \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2 (1 + 1/\kappa)}{\tilde{\theta}^2},
\end{aligned}
$$

em que $z_{\alpha/2}$ e z_β são quantis da distribuição normal padrão.

Fórmula para Cálculo: Cenário Simples

Para o caso em questão, a variância deverá ser estimada dos dados e será aplicado um teste *t*.

Mesmo nesta situação simples o tamanho amostral não pode mais ser obtido por fórmula fechada.

De forma geral, o tamanho amostral só poderá ser obtida após algumas simplificações (aproximações) e por meio de cálculos iterativos.

Exemplo: Tamanho Amostral

```r
power.t.test(delta = 15, sd = 20, sig.level = 0.05,
             power = 0.8, alternative="one.sided")
```

```
     Two-sample t test power calculation

              n = 22.69032
          delta = 15
             sd = 20
      sig.level = 0.05
          power = 0.8
    alternative = one.sided

NOTE: n is number in *each* group
```

Exemplo: Poder Real

```r
power.t.test(delta = 15, sd = 20, sig.level = 0.05,
             n = 23, alternative="one.sided")
```

```
     Two-sample t test power calculation

              n = 23
          delta = 15
             sd = 20
      sig.level = 0.05
          power = 0.8048559
    alternative = one.sided

NOTE: n is number in *each* group
```

Exemplo

Na definição do poder do teste, queremos ter uma chance razoável de detectar o tamanho do efeito estabelecido. Um valor de $0,80$ é muito comum – alguns autores sugerem poder maior, como $0,85$ ou $0,90$.

À medida que o poder aumenta, contudo, o tamanho amostral aumenta numa taxa crescente.

No exemplo, um poder de $\tilde{\pi} = 0,95$ necessita de um tamanho amostral de $n = 40$ – cerca de 75% mais do que é necessário para um poder de $0,80$.

Exemplo: Tamanho Amostral

```
power.t.test(delta = 15, sd = 20, sig.level = 0.05,
             power = 0.95, alternative="one.sided")
```

```
     Two-sample t test power calculation

              n = 39.17515
          delta = 15
             sd = 20
      sig.level = 0.05
          power = 0.95
    alternative = one.sided

NOTE: n is number in *each* group
```

Tamanho Amostral e Poder

Algumas dificuldades...

- Quem nos disse que a meta era detectar uma diferença média de 15 mmHg?
- Quem nos disse que $\sigma = 20$, se estamos apenas planejando o experimento e os dados nem foram coletados ainda?

Tais *inputs* do tamanho amostral são geralmente complicados...

- Obter um tamanho de efeito de importância científica requer o conhecimento do pesquisador responsável pelo estudo.
- Novamente é bastante importante a discussão com o especialista.

Escolha do Tamanho do Efeito

Tamanho do Efeito

Importante no problema do tamanho amostral é definir um tamanho de efeito de interesse científico.

Essa é uma tarefa do pesquisador envolvido no estudo.

O problema é que o pesquisador nem sempre sabe o que está sendo perguntado, ou não reconhece como uma questão que seja de sua responsabilidade responder.

Questões concretas

P: "Quais resultados você espera ver?"

Podemos ter um limite *superior* para $\tilde{\theta}$.

O pesquisador provavelmente não faria o estudo se não espera que os resultados sejam cientificamente significativos. Podemos estabelecer um limite *inferior* para o tamanho amostral.

P: "Um efeito de metade dessa magnitude seria de interesse científico?"

Mas metade de $\tilde{\theta}$ aproximadamente vai quadruplicar o tamanho amostral...

Vários cálculos de n para várias propostas pode ajudar

Podemos tentar uma seleção de tamanhos de efeito e seu poder correspondente:

- "Com 25 observações, teremos 50% de chance de detectar uma diferença de $9,4$ mmHg, e 90% de chance de detectar uma diferença de 16.8 mmHg."
- "Se você puder bancar mais 6 indivíduos por tratamento, poderá detectar uma diferença de 15 mmHg com 90% de poder."

Vários cálculos de n para várias propostas pode ajudar

```
power.t.test(delta = 9.4, sd = 20, sig.level = 0.05,
             n = 25, alternative="one.sided")
```

```

     Two-sample t test power calculation 

              n = 25
          delta = 9.4
             sd = 20
      sig.level = 0.05
          power = 0.4973543
    alternative = one.sided

NOTE: n is number in *each* group
```

Vários cálculos de n para várias propostas pode ajudar

```
power.t.test(delta = 16.8, sd = 20, sig.level = 0.05,
             n = 25, alternative="one.sided")
```

```
     Two-sample t test power calculation

              n = 25
          delta = 16.8
             sd = 20
      sig.level = 0.05
          power = 0.9002053
    alternative = one.sided

NOTE: n is number in *each* group
```

Vários cálculos de n para várias propostas pode ajudar

```
power.t.test(delta = 15, sd = 20, sig.level = 0.05,
             n = 31, alternative="one.sided")
```

```

     Two-sample t test power calculation 

              n = 31
          delta = 15
             sd = 20
      sig.level = 0.05
          power = 0.8987293
    alternative = one.sided

NOTE: n is number in *each* group
```

Vários cálculos de *n* para várias propostas pode ajudar

E o que *não* pode ser detectado:

P: “Qual a variação de irrelevância clínica?”

P: “Se você fosse o paciente, os benefícios de se reduzir a pressão em 15 mmHg compensariam o custo, inconveniência, e potenciais efeitos colaterais desse tratamento?”

Escolha da Variância Correta

Escolha da Variância

Funções poder envolvem parâmetros não relacionados com as hipóteses. Geralmente envolvem uma ou mais variâncias.

No exemplo queremos saber a variância residual das medidas no experimento de duas amostras.

Opções são buscar a variância da experiência do pesquisador, ou usar dados históricos, ou conduzir um estudo piloto.

"Qual a variação natural da pressão? Quais são os maiores e menores valores de pressão que você já viu?"

Escolha da Variância

Dados históricos ou de uma amostra piloto não precisam seguir o mesmo desenho do estudo planejado; mas deve-se tomar cuidado para que a variância correta seja estimada.

Exemplo: o fabricante do esfigmomanômetro pode ter publicado resultados de teste mostrando que o desvio padrão das leituras é de $2,5$ mmHg.

Esse número não é apropriado para uso na determinação do tamanho amostral: reflete variações nas leituras feitas no mesmo indivíduo sob condições idênticas.

A variação residual no experimento da pressão inclui a variação entre indivíduos!

Escolha da Variância

A determinação e consideração das fontes de variação em estudos passados é bastante importante.

Incluem: atributos do paciente (sexo, idade, fatores de risco, demográficos, etc.), instrumentos, como, quando e quem administra os medicamentos e coleta os dados, e outros fatores.

Suponha que temos dados passados de um experimento em que homens e mulheres foram aleatorizados separadamente a grupos que receberam diferentes regimes de exercícios; e que a resposta é a pressão medida usando instrumentos idênticos àqueles que planejamos usar.

Escolha da Variância

Isso nos fornece dados úteis para planejar o novo estudo – mas temos que ter cuidado.

Exemplo: a variância residual do estudo anterior não inclui variações devido ao sexo.

Se o novo estudo inclui indivíduos de ambos os sexos, então a variação devido ao sexo deve ser incluída na variância do erro a ser usada no planejamento do tamanho amostral.

A mesma pessoa toma todas as medidas? Se isso é feito por várias pessoas – o treinamento delas é comparável?

Todos esses fatores podem afetar a variância do erro!