Capítulo 2

Sobre a natureza das relações

Prelúdio: Quem precisa do valor p?

O racional apresentado no capítulo anterior é diretamente relacionado ao método hipotético-dedutivo e seus princípios filosóficos.
Apesar dos inúmeros avanços citados, a interpretação do valor p não é muito intuitiva.
Envolve mensurar quão improváveis são as observações em um cenário hipotético na vigência da hipótese nula.
Sua tradução (errada) mais popular é de que representa *"a chance de o resultado deste estudo estar errado"*.

O arcabouço descrito no capítulo anterior é suficiente para produzir um trabalho científico críptico para leigos.

Ao seguir receitas pré-definidas (formulação de H_0 e H_1, cálculo de estatísticas e valores p), um texto parece estar em conformação com os padrões acadêmicos, mesmo que a hipótese elementar em torno do objeto de pesquisa seja simplória. Assim, inadvertidamente, priorizamos a forma e relegamos a segundo plano o miolo de propostas científicas.
Trabalhos de pouca originalidade recebem grande atenção pelo rigor por atributos quantitativos (e.g. tamanho amostral grande, valor p baixo), enquanto criativos e revolucionários experimentos menores levam anos ou décadas até atingirem a comunidade.
Outro efeito colateral é a busca por valores p que rejeitem H_0, desprezando precedentes teóricos e premissas probabilísticas (múltiplos testes).

A difícil interpretabilidade do valor p e as armadilhas frequentes envolvidas no processo de inferência levaram a comunidade científica a questionar a hegemonia desse parâmetro. Há uma presente tendência a abandonar o valor p e o limite $p < 0.05$ como critérios canônicos.

No próximo capítulo, vamos conhecer argumentos contundentes ao método hipotético dedutivo.
Por enquanto, basta sabermos que é sempre vantajoso obter outras informações, complementares ou alternativas.

Neste capítulos, vamos aprender a estimar (1) a magnitude da diferença entre duas amostras e (2) quão relacionados são valores pareados (e.g. peso e altura).

Tamanho de efeito

O tamanho de efeito nos ajuda a expressar magnitude.
Retomando o exemplo anterior, de que adianta uma diferença significativa entre o tamanho dos bicos dos pássaros, se ela for de $0.00001mm$?
Ainda, existem casos em que estudos pequenos sugerem efeitos importantes, porém o tamanho amostral não fornce poder estatístico suficiente para rejeição da hipótese nula.

Além de saber quão improvável é a diferença observada, é natural imaginarmos o quão grande ela é.

Uma medida bastante popular é o *D de Cohen (Cohen's D)*.

É um parâmetro que expressa a magnitude da diferença sem usar unidades de medida.
Uma torcedora de futebol conta (feliz) a um amig que seu time favorito venceu com placar de 4 x 1 (gols).
Porém, esse amigo acompanha basquetebol e está acostumado a placares como 102 x 93 (cestas).
Como é possível comparar gols com cestas? Qual vitória é representa pontuações mais discrepantes: 4 x 1 ou 102 x 93?

O problema aqui é que as pontuações se comportam de maneiras diferentes entre os esportes. O D de Cohen consiste em expressar essa diferença em desvios-padrão. Bastante simples:
$D_{cohen} = \frac{mu_1 - mu_2}{\sigma_{pooled}}$

Como vimos no teste t de medidas independentes, o desvio-padrão balanceado entre amostras pode ser dado por:
$\sigma_{pooled} = \sqrt{\frac{(n_1-1)\sigma_1^2+(n_2-1)\sigma_2^2}{(n_1-1)+(n_2-1)}}$

Usando a biblioteca *effects*, podemos calcular diretamente:

```
library(effects)
# O dataset galapagos_birds foi criado no capitulo 1
cohen.d(galapagos_birds$X1,galapagos_birds$X2)

Cohen's d

d estimate: -5.460017 (large)
95 percent confidence interval:
    lower     upper 1
-5.954047 -4.965987
```

Cohen propôs algumas faixas para classificar a magnitude desses efeitos:

	Pequeno	Médio	Grande
Cohen's D	0-0.2	0.2-0.5	0.5 - 0.8

Assim, podemos atualizar nossos resultados anteriores, reportando também o tamanho de efeito da diferença e seu intervalo de confiança.

Correlações