4.6. Testes de diferenças com amostras grandes

Até agora, vimos técnicas para testar hipóteses sobre parâmetros de uma população (a média μ ou a proporção π). Estes são os testes básicos, os primeiros que foram desenvolvidos por estatísticos, e os primeiros a serem estudados pelos estudantes. São usados, por exemplo, no controle de qualidade: testes de médias são feitos para verificar se as dimensões de um produto estão de acordo com os valores especificados pelo projeto, ou se a resistência de um material atende ao requerido pelas normas técnicas; testes de proporções são feitos para verificar se a proporção de peças defeituosas produzidas por uma fábrica está dentro do especificado por um contrato.

Nas maioria das aplicações nas ciências e na tecnologia, porém, os testes que mais interessam são os que fazem *comparações* entre duas ou mais populações. Grande parte da pesquisa feita na Medicina, por exemplo, visa comparar as características dos pacientes submetidos a um tratamento A com as dos pacientes submetidos a um tratamento B. Nas Engenharias, testes são feitos para comparar processos de fabricação e verificar qual deles consegue melhores resultados (melhor qualidade do produto final, menor custo, maior rapidez, etc.).

Do ponto de vista estatístico, estes testes comparam as características de duas populações imaginárias, cada uma representada por uma amostra. Por exemplo, se um novo tratamento é proposto para reduzir o nível de colesterol de pacientes, um teste deverá ser feito para verificar se a redução média conseguida por este tratamento difere da redução média conseguida por um outro tratamento já existente. Iremos supor duas populações: a de pacientes hipoteticamente tratados pelo novo método, e a de pacientes hipoteticamente tratados pelo método já existente. Um teste estatístico será feito para comparar *dois* parâmetros, as médias que caracterizam estas duas populações, a partir dos resultados obtidos em duas amostras, uma de cada população.

Nos testes que vimos anteriormente, a hipótese nula afirmava que o parâmetro de interesse tinha um valor determinado (por exemplo, no problema sobre peso de peixes,

$$H_o : \mu = 30$$

Este valor se tornava, por hipótese, a média da distribuição amostral a partir da qual era feito o teste. Se estamos comparando duas populações, porém, não é possível criar uma hipótese nula que atribua valores aos *dois* parâmetros simultaneamente; o que podemos fazer é criar uma variável que indique quão similares (ou não) são os parâmetros das duas populações, e fazer um teste sobre esta nova variável.

Há duas formas de fazer isto. A primeira é criar a variável *diferença* entre os dois parâmetros, que representaremos por δ (letra *delta* minúscula do alfabeto grego, equivalente ao d do alfabeto latino). Por exemplo, num teste para comparar as médias de duas populações (como no problema do tratamento para redução do colesterol mencionado acima), a hipótese nula será de que esta diferença é nula:

$$H_0 : \quad \delta = \mu_1 - \mu_2 = 0$$

onde μ_1 e μ_2 são as médias das duas populações. (Na maior parte das aplicações, a hipótese nula é a de que esta diferença é igual a zero; daí o nome de "hipótese nula"). O teste visa verificar se a diferença D observada entre as duas médias das amostras

$$D = \overline{X}_1 - \overline{X}_2$$

é *significativa*; isto é, que ela é tão grande que dificilmente poderia teria ocorrido por acaso, se a hipótese nula fosse verdadeira. Testes de diferenças também são feitos para comparar duas proporções. A hipótese nula é a de que as duas proporções são iguais:

$$H_0 : \delta = \pi_1 - \pi_2 = 0$$

e o teste é feito a partir da diferença entre as proporções encontradas nas amostras,

$$D = P_1 - P_2 .$$

Um outra forma de comparar dois parâmetros é criar a variável *razão* entre eles; se estes parâmetros forem iguais, a razão deverá ser igual à unidade. Representaremos esta razão por ρ (letra grega *ro*, equivalente ao *r* latino). Este tipo de hipótese é usado, por exemplo, na comparação de variâncias de duas populações. No controle de qualidade de produção, se quisermos comparar as variâncias de dois processos de fabricação para uma mesma peça (a variância é um dos indicadores da qualidade, pois as dimensões de uma peça não podem variar muito em relação ao que foi requerido no projeto), a hipótese nula será

$$H_0 : \quad \rho = \frac{\sigma_1^2}{\sigma_2^2} = 1$$

Por que existem estas duas formas? Em parte, por causa da conveniência matemática. É fácil mostrar que a distribuição amostral da diferença entre duas médias é uma distribuição normal, se estas médias têm distribuição normal, pois a diferença entre duas variáveis normais também é normal (seção 3.4.4.3). Por outro lado, a diferença entre duas variâncias não tem distribuição amostral conhecida, mas a razão entre duas variâncias tem. Existe também um motivo prática: às vezes os resultados obtidos de uma forma são mais úteis do que os obtidos da outra forma. Por exemplo, num teste para comparar os riscos de dois tratamentos médicos (o *risco* é a probabilidade de ocorrência de um evento determinado), poderíamos testar a *diferença* entre eles (como os riscos são probabilidades, a diferença entre eles tem distribuição normal). Esta diferença, porém, não tem grande interesse para os médicos, já que as probabilidades em geral são muito pequenas; mais útil é a *razão* entre estes riscos, isto é, quantas vezes um deles é maior do que o outro. (Por exemplo, se um tratamento tem um risco de 0,005 e o outro tem um risco de 0,010, a diferença entre eles é de apenas 0,005, o que parece não significar muito; mas a razão entre eles é 2, o que indica que um tratamento tem o *dobro* do risco do outro.)

Nas próximas duas seções veremos o teste da diferença de médias (seção 4.6.1) e o da diferença de proporções (seção 4.6.2) de duas populações.