5.3.2. Teste dos sinais

O *teste dos sinais* é o mais simples dos testes não-paramétricos, e pode ser usado quando as amostras são pareadas e não muito pequenas. O teste é baseado apenas nos sinais (positivos ou negativos) das diferenças observadas em cada par de observações; é o único teste não-paramétrico que veremos que não se baseia em *postos*.

5.3.2.1. Exemplo

Suponha que pesquisadores desejam verificar qual de duas marcas de refrigerante é preferida pelo público de uma região, C-Cola ou P-Cola. Para isto, fazem uma pesquisa de campo com uma amostra de 20 voluntários, que devem experimentar os dois refrigerantes, oferecidos em dois copos que não têm nenhum sinal que identifique o conteúdo, e depois dar uma nota entre 1 a 5 para cada. Os resultados dão mostrados na Tab. 1. A variável de teste (X) será o número de pares em que a nota dada ao refrigerante P-Cola foi maior do que a dada ao C-Cola (ou seja, o número de vezes em que o sinal da diferença entre as notas foi positivo).

Tabela 1. Notas dadas pelos participantes do estudo a cada marca

participantes >	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	20	soma
C-Cola	5	2	4	3	4	4	2	1	3	2	2	1	3	5	4	5	5	3	5	1	dos
P-Cola	3	3	5	5	3	5	3	5	4	5	4	2	5	3	2	3	4	5	4	5	sinais
sinal da diferença	–	+	+	+	–	+	+	+	+	+	+	+	+	–	–	–	–	+	–	+	13

O que o teste do sinal faz é responder à pergunta: se na população não houver diferença na preferência por uma ou outra das marcas, e as escolhas dos participantes forem inteiramente aleatórias, qual é a probabilidade de que seja obtido um valor de X como este que encontramos, ou um valor mais extremo? (ou seja, qual seria o valor-p deste X?).

O teste parte da hipótese nula de que a probabilidade de os participantes escolherem uma ou outra marca sejam as mesmas; a hipótese alternativa será a de que a probabilidade sejam diferentes, o que leva a um teste bilateral. O teste é feito usando uma distribuição binomial com $n = 20$, e $p = 0.5$, e $x = 13$ (ou $x = 7$, o que dá no mesmo; tanto faz contar as diferenças positivas ou as negativas, já que o teste é bilateral).

Note que este é exatamente o mesmo problema que já foi analisado na seção 4.3.1: se uma tachinha equilibrada é lançada 20 vezes, qual é a probabilidade de que ela caia 13 ou mais vezes com a ponta para cima?

No R, o teste pode ser feito usando a função `binom.test`, como mostrado a seguir:

```
n=20; x= 13; binom.test(x,n)
Exact binomial test
data:  x and n
number of successes = 13, number of trials = 20
      p-value = 0.2632
alternative hypothesis:
      true probability of success is not equal to 0.5
95 percent confidence interval:
      0.4078115 0.8460908
```

O valor-p obtido é $p = 0,2632$. O resultado do teste é não-significativo; não há evidência nestes dados de que haja na população alguma preferência por uma marca em relação à outra.

O teste pode ser feito também simplesmente usando a distribuição binomial, calculada pelo R na função `pbinom`, já conhecida:

```
x= 13; size=20; prob=0.5
pvalue=2*pbinom(x-1,size,prob); pvalue
```

Note que no argumento da função `pbinom` temos que introduzir a valor *x-1*, em lugar de *x*. No problema acima, por exemplo, queremos calcular a soma de todas as probabilidades dos valores $X \leq 12$, para obtermos a probabilidade de $X \geq 13$.

5.3.2.2. Comentários

(i) Poder do teste

Este teste desperdiça grande parte da informação contida nos dados, pois considera apenas o sinal da diferença entre as notas dadas, não a magnitude desta diferença (uma diferença = 1 vale tanto quanto uma diferença = 4; o que importa é que ambas sejam positivas). Isto faz que este teste tenha um *poder* muito baixo; um teste de maior poder, que usa tanto os sinais quanto a magnitude das diferenças, é o teste de Wilcoxon, que será visto a seguir.

Por outro lado, o fato de este teste não requerer os valores numéricos das diferenças permite que ele seja usado mesmo quando as variáveis são medidas numa escala *ordinal*, e não numa escala quantitativa. Por exemplo, poderíamos ter perguntado a cada participante do estudo acima qual das duas marcas de refrigerantes preferia (em vez de lhes pedir que dessem notas a cada refrigerante, o que é mais trabalhoso). As respostas seriam as mostradas na Tab. 2, e o resultado do teste seria idêntico ao anterior.

Tabela 2. Marca preferida por cada participante

participantes >	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	20	totais
C-Cola	•				•									•	•	•	•		•		7
P-Cola		•	•	•		•	•	•	•	•	•	•	•					•		•	13

(ii) Empates (diferenças nulas)

Se as diferenças entre as observações em um par for nula, e portanto não tiver sinal (por exemplo, se no estudo acima um participante der a mesma nota a ambas as marcas), este par deve ser excluído da análise. Isto evidentemente vai reduzir um pouco o poder do teste, pois a amostra ficará menor. Se o número de empates em um estudo for muito grande, provavelmente será melhor não usar este teste para analisar os resultados.

(iii) Amostras grandes

Para amostras grandes, pode ser usada a aproximação normal para a binomial (veja seção **3.3.5**). Outra opção é e a de analisar os resultados usando um teste de *Chi-quadrado* (este teste porém não será visto neste curso).