5.3.4. Teste de Kruskall-Wallis

5.3.4. 1. Introdução

O teste de Kruskall-Wallis (*KW*) testa a hipótese nula de que um conjunto de *k* populações tenham a mesma distribuição de probabilidades, contra a hipótese alternativa de que nem todas estas populações tenham a mesma distribuição.

A ANOVA (seção **5.3.1**) faz algo parecido: testa a hipótese nula de que várias populações tenham a mesma média. Como o pressuposto básico é o de que estas populaões têm distribuição normal e mesma variância, se elas tiverem as mesmas médias terão distribuições idênticas (distribuições normais de mesma média e mesma variância são idênticas entre si). A ANOVA é um teste *paramétrico*, baseado nos parâmetros média e variância, e exige que as distribuições sejam normais. O teste *KW* é uma alternativa *não-paramétrica* que pode ser usada quando as distribuições não são normais, pois requer apenas que a variável seja medida pelo menos numa escala ordinal.

A idéia em que se baseia este teste é a mesma que vimos nos testes de Wilcoxon / Mann-Whitney e dos postos sinalizados de Wilcoxon: os valores originais da variável são substituídos por seus *postos*, e o cálculo da significância das somas de postos encontradas nas amostras é feita por Análise Combinatória (se as amostras forem pequenas) ou por aproximação por um modelo conhecido de VAC (normal ou qui-quadrado).

5.3.4. 2. Exemplo 1

Para ilustrar o funcionamento do teste, usaremos um exemplo simples, com dados simulados. Suponha que foi feito um estudo comparando os pesos medidos nos animais de amostras de uma mesma espécie, que vivem em três regiões diferentes. Os valores encontrados estão mostrados na Tab. **1**.

Tabela 1

	pesos (g)			postos		
	amostra 1	amostra 2	amostra 3	amostra 1	amostra 2	amostra 3
	316	336	303	7	11	5
	319	356	287	9	17	3
	362	315	295	19	6	4
	334	351	266	10	15	1
	358	353	318	18	16	8
	343	345	347	12	13	14
			283			2
soma dos postos de cada tratamento →				75	78	37
média dos postos de cada tratamento →				12,5	13,0	5,29
desvio da média de cada tratamento em relação a média geral →				2,5	3,0	– 4,71

Os passos para a realização do teste *KW* são:

1. Atribuir a cada valor observado nas amostras o posto correspondente (três colunas à direita da tabela);

2. Somar os postos de cada amostra, e calcular a média destes postos (penúltima linha da tabela).

$$\overline{R}_1 = \frac{\Sigma_1}{n_1} = \frac{75}{6} = 12{,}5 \qquad \overline{R}_2 = \frac{\Sigma_2}{n_2} = \frac{78}{6} = 13{,}0 \qquad \overline{R}_3 = \frac{\Sigma_3}{n_3} = \frac{37}{7} = 5{,}29$$

3. Calcular a *média geral* (média dos postos de todas as observações)

Se temos um total de n observações, divididas entre as várias amostras, a soma dos postos destas observações pode ser calculada por:

$$soma = 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

No exemplo, como $n = 19$, a média geral pode ser calculada por:

$$\overline{R} = \frac{soma}{n} = \frac{n(n+1)}{2} \times \frac{1}{n} = \frac{n+1}{2} = \frac{190}{19} = 10$$

4. Calcular o desvio da média dos postos de cada amostra em relação à média geral (última linha da tabela). Os desvios são dados por:

$$d_1 = \overline{R}_1 - \overline{R} = 12{,}5 - 10{,}0 = 2{,}5$$
$$d_2 = \overline{R}_2 - \overline{R} = 13{,}0 - 10{,}0 = 3{,}0$$
$$d_3 = \overline{R}_3 - \overline{R} = 5{,}29 - 10{,}0 = -4{,}71$$

Se a hipótese nula for verdadeira (as três populações de onde vieram estas amostras têm a mesma distribuição), esperaríamos que as médias dos postos das amostras fossem semelhantes entre si, e os desvios delas em relação à média geral fossem pequenos. No exemplo, os desvios das amostras 1 e 2 parecem semelhantes, mas o da amostra 3 parece ser bem diferente. Será que este desvio pode ser considerado significativamente diferente dos outros dois?

O teste *KW* usa como estatística de teste o valor *H*, definido como:

$$H = \frac{12}{n(n+1)}\left[n_1 d_1^2 + n_2 d_2^2 + \ldots + n_k d_k^2\right]$$

Para amostras pequenas, os valores críticos de *H* podem ser encontrados em tabelas, construídas por meio de Análise Combinatória. Fazer estas tabelas porém é muito trabalhoso, uma vez que é preciso que haja tabelas diferentes para cada número de amostras, e para cada número de observações em cada amostra; a maioria dos livros, por isso, não traz estas tabelas. Para amostras grandes, pode ser demonstrado que *H* tem distribuição que tende para uma distribuição qui-quadrado com *n-1* graus de liberdade, e o valor-p do *H* encontrado pode ser calculado a partir daí.

Os programas de computador sempre usam as aproximações por qui-quadrado, ao invés de usar os valores exatos. Fazendo o cálculo no R, obtemos:

```
kruskal.test(list(amostra1,amostra2,amostra3))

        Kruskal-Wallis rank sum test
        data:  pesos
        Kruskal-Wallis chi-squared = 7.8023, df = 2, p-value = 0.02022
```

Nesta saída do R, note que o valor da estatística de teste (chamada de *Kruskal-Wallis chi-squared*) está bem próximo do valor *H* que calculamos acima. O valor-p deste resultado é menor que $\alpha=0.05$, e o resultado pode então ser considerado significativo (mas não muito). Há evidência, mas não muito forte, de que as três populações não tenham a mesma distribuição.