

Teste de Mann-Whitney

Material obtido em: http://www.apis2.com.br/?page_id=262

Quando se dispõe de uma amostra pequena e a variável numérica não apresenta sabidamente uma variação normal (ou não dá para ser verificada satisfatoriamente), ou ainda, quando não há homogeneidade das variâncias (embora exista uma correção no teste t que considera as variâncias desiguais), o teste t não é apropriado. Para exemplificar uma situação onde o teste t acusaria falsamente uma associação estatisticamente significativa, imagine que em um dos dois grupos se observe um *outlier* (valor muito discrepante). Em função desse único valor, em sendo muito maior do que os outros, o grupo a que ele pertence apresentará uma média elevada, o que aumentará a estatística do teste t , com um conseqüente p -valor associado pequeno.

Nessa situação, pode-se utilizar o teste não paramétrico de Mann-Whitney. O teste de Mann-Whitney teste foi desenvolvido primeiramente por F. Wilcoxon em 1945, para comparar tendências centrais de duas amostras independentes de tamanhos iguais. Em 1947, H.B. Mann e D.R. Whitney generalizaram a técnica para amostras de tamanhos diferentes. O **teste de Mann-Whitney** (*Wilcoxon rank-sum test*) é indicado para comparação de dois grupos não pareados para se verificar se pertencem ou não à mesma população e cujos requisitos para aplicação do teste t de Student não foram cumpridos. Na verdade, verifica-se se há evidências para acreditar que valores de um grupo A são superiores aos valores do grupo B. O teste U pode ser considerado a versão não paramétrica do teste t , para amostras independentes. Ao contrário do teste t , que testa a igualdade das médias, o teste de Mann-Whitney (U) testa a igualdade das medianas. Os valores de U calculados pelo teste avaliam o grau de entrelaçamento dos dados dos dois grupos após a ordenação. A maior separação dos dados em conjunto indica que as amostras são distintas, rejeitando-se a hipótese de igualdade das medianas.

Testes NÃO PARAMÉTRICOS

Variável Qualitativa
com escala ordinal
ou
Variável Quantitativa
sem distribuição normal



A lógica do teste é a mesma do teste t : calcula-se uma certa estatística de teste e obtém-se o p -valor a partir da distribuição amostral dessa estatística sob H_0 . A diferença é que ao invés de construir essa estatística com dados originais, eles são previamente convertidos em postos (ordenações). A vantagem é que, com isso, as suposições de normalidade e homogeneidade das variâncias não são necessárias, permitindo mais generalidade aos resultados. Perceba também que aquele *outlier* perde sua influência nessa abordagem, sendo apenas o maior valor da amostra. Outra vantagem desse teste é que ele pode ser aplicado em uma variável que seja originalmente ordinal, enquanto que o teste t exige uma escala pelo menos intervalar. Uma desvantagem é que ao substituir os dados originais por postos, “joga-se fora” alguma informação. Se as condições para o teste t são satisfeitas, ainda assim poder-se-ia usar o teste de Mann-Whitney, mas ele não seria tão poderoso quanto o teste t .

A estatística U , que é a base para a decisão sobre a aceitação ou não da hipótese de nulidade é calculada da seguinte maneira:

- É formado um conjunto W , com todos os dados das duas amostras (A e B);
- O conjunto W é ordenado de forma crescente;
- Anota-se a ordem de cada elemento deste conjunto;
- Separam-se novamente as amostras A e B ;
- O valor de U é a soma das ordens da amostra A .

Quanto mais baixo for o valor de U , maior será a evidência de que as populações são diferentes. Isso se explica porque U é a soma de ordens, portanto seu valor será baixo se na categoria A estiverem os primeiros da ordem (obviamente em B estarão os dados de ordem superior). É claro então que neste caso se evidencia uma diferença entre as populações.

Se as diferenças entre as situações forem aleatórias, como é postulado pela hipótese nula, então os resultados devem ser aproximadamente os mesmos e, conseqüentemente, as ordens devem ser também

aproximadamente as mesmas para as duas situações. Se houver uma preponderância de ordens altas ou baixas numa situação ou na outra, então é porque a diferença no total dos resultados ordenados para cada situação é devida aos efeitos previstos da variável independente e não ao acaso. Se a soma total das ordens for muito baixa para uma das situações, então terá de haver uma preponderância de ordens elevadas na outra situação. Quanto menor for U mais significativas serão as diferenças entre as ordens das duas situações.

Para amostras Pequenas (nenhum $n > 20$)

A estatística do teste Mann-Whitney (U) é calculada a partir dos tamanhos amostrais de cada grupo e a soma dos postos de um deles. Existe uma distribuição teórica e exata para a estatística U especialmente desenvolvida para amostras pequenas (nenhum $n > 20$), cuja tabela pode ser vista em Siegel.

Para amostras grandes

Por outro lado, quando os tamanhos amostrais forem de pelo menos dez em cada grupo ($n > 20$), a aproximação pela distribuição normal para a estatística T1 já é razoável.