

Índices de significância para tabelas de contingência

Natália Lombardi de Oliveira¹, Adriano Polpo de Campos²

1. Aluna de Estatística – DEs/UFSCar ; *nat.nlo@gmail.com

2. Professor Doutor do Depto.de Estatística, UFSCar, São Carlos/SP

Palavras Chave: *índices de significância, tabelas de contingência, p-valor*

Introdução

Testes de hipótese é a ferramenta estatística mais utilizada em diversos campos da ciência. O p-valor é um índice amplamente utilizado; porém, é um índice realmente razoável? E seriam as aproximações assintóticas, utilizadas nos testes de homogeneidade por exemplo, adequadas? Com isto em mente, estudou-se os índices de significância para o teste de homogeneidade em tabelas de contingência 2x2. Utilizando a estatística de razão de verossimilhanças, desenvolveu-se um índice de significância (P-valor) para testar esta hipótese. Realizou-se um estudo de comparação para entender o comportamento deste índice proposto e compará-lo a diversos índices conhecidos, tanto frequentistas como Bayesianos.

Resultados e Discussão

Sejam X e Y duas variáveis aleatórias que podem ser representadas na Tabela1. Queremos testar se as variáveis X e Y são homogêneas com respeito às categorias C1 e C2. Considerando n e m fixos e conhecidos, assume-se que $X \sim \text{Binomial}(p,n)$ e $Y \sim \text{Binomial}(q,m)$. A hipótese de homogeneidade pode ser escrita como $H: p = q$.

Tabela 1. Tabela de contingência 2x2.

	C1	C2	Total
X	x	n-x	n
Y	y	m-y	m

Denotaremos a verossimilhança por $L(p,q|x,y,n,m)$ e a estatística do teste de razão de verossimilhanças por $\text{Lamb}(x,y)$. Os índices estudados e posteriormente comparados são o P-valor proposto, o p-valor referente ao teste qui-quadrado, o p-valor referente ao teste da razão de verossimilhanças e um índice Bayesiano, o e-valor assintótico e exato (Full Bayesian Significance Test - FBST). Por questões de espaço, neste resumo será construído apenas o P-valor. Este valor não depende de aproximações assintóticas, ou seja, propõe-se uma maneira exata de obter a distribuição da estatística de teste.

O P-valor baseia-se na idéia de utilizar a distribuição preditiva de X,Y antes de se observar qualquer dado. Este índice é uma forma alternativa de calcular um p-valor exato para o teste de razão de verossimilhanças. Para calcular este índice, precisa-se obter a distribuição amostral da estatística de teste sob H. O objetivo é encontrar uma distribuição para a tabela sob H que não dependa do vetor de parâmetros (p,q). Para isso, considerou-se (p,q) parâmetro nuisance em $L(p,q|x,y,n,m)$ e, para eliminá-lo, integrou-se nestes parâmetros. Denotaremos esta integral por $h(x,y)$. Para chegar a uma distribuição de probabilidades, precisa-se obter uma constante normalizadora. Esta constante dá-se pela soma

de $h(x,y)$ para todas as possíveis tabelas com marginais n, m. Com isso, o cálculo do P-valor segue diretamente:

$$P\text{-valor} = P(\text{Lamb}(X,Y) < \text{Lamb}(x,y) | H).$$

Para melhor entender os índices de significância para testar homogeneidade, realizou-se um estudo de comparação. Vale ressaltar que não é um estudo de simulação, dado que construiu-se todas as possíveis tabelas e possíveis índices para cada tamanho amostral considerado. Este estudo foi realizado de duas maneiras. Primeiramente, construiu-se todos os índices para cada tabela e desenhou-se os índices indexados pela estatística $\text{Lamb}(x,y)$ considerando 6 diferentes tamanhos amostrais. Neste estudo, notou-se que os índices exatos e assintóticos são muito similares, mesmo ao considerar tamanhos amostrais pequenos. Por último, o segundo estudo considerou a relação existente entre o e-valor assintótico e o p-valor assintótico do teste de razão de verossimilhanças e verificou-se que esta relação também existe (aproximadamente) para o e-valor exato e o P-valor proposto, também considerando os mesmos 6 tamanhos amostrais.

Conclusões

Desenvolveu-se o P-valor, um índice para testar hipóteses de homogeneidade em tabelas de contingência 2x2 baseado no Teste de Razão de Verossimilhanças. O estudo de comparação indicou que o índice proposto comporta-se de maneira muito similar ao p-valor assintótico e ao e-valor, mesmo para tamanhos amostrais pequenos. A desvantagem deste índice é que precisa-se avaliar a probabilidade de todas as tabelas do espaço amostral para calculá-lo. Graças à boa aproximação deste índice aos assintóticos em condições inesperadas, pode-se considerar que esta aproximação produz resultados confiáveis mesmo para pequenas amostras. Como trabalho futuro, propõe-se estender o P-valor para tabelas de dimensões maiores e diferentes testes, como o de independência. Além disso, sugere-se também comparar os resultados com outras alternativas de cálculos exatos de p-valor.

Agradecimentos

Agradeço à FAPESP pelo apoio financeiro para desenvolver este trabalho (projeto 2012/16669-4).

1. G. Casella and R. Berger. *Statistical Inference*. Duxbury Press, 2nd edition, 2001.
2. M. A. Diniz, C. A. B. Pereira, A. Polpo, J. Stern, and S. Wechsler. Relationship between Bayesian and frequentist significance indices. *International Journal for Uncertainty Quantification*, 2(2):161–172, 2012. doi: 10.1615/Int.J.UncertaintyQuantification.2012003647.
3. C. A. B. Pereira and S. S. Wechsler. On the concept of p-value. *Brazilian Journal of Probability and Statistics*, 7:159–177, 1993.
4. C. A. B. Pereira, J.M. Stern, and S. Wechsler. Can a significance test be genuinely Bayesian? *Bayesian Analysis*, 3(1):19–100, 2008.
5. C.A.B. Pereira and J.M. Stern. Evidence and credibility: a full Bayesian test of precise hypothesis. *Entropy*, 1:104–115, 1999.