

多位点单倍体型分数检验

采用计算分数 (score) 的方法, 检测多位点单倍体型与表型的相关性, 筛选出最可能导致疾病发生的单倍体型。表型特征可以是两分类型 (有与无)、有序多分类型、正态分布的定量型和符号泊松分布的类型。检验包括对几个位点单倍体 型的全局检验与特异单倍型的分数试验, 并允许调整其它协变量。效应包括相加性、显性或隐性模型。此外该模块还可有一个选项来让计算机随机组合数据, 计算模拟的 p 值, 这对于小样本量或偏离假定分布的数据尤为重要。

关于分数检验详见: Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425-34。

例 1, 某研究选取血管紧张素原 (angiotensinogen, AGT) 基因启动子-217, -152, -20, -6, 内含子 1 的+31, 共 5 个位点, 对 100 例正常人进行了 SNP 位点基因型检测, 点击[这里](#)下载数据, 其中 0, 1, 2 分别代表 5 个位点的基因型: G217-A 为 AA, AG 和 GG; G152-A 为 AA, AG 和 GG, A-20C 为 AA, AC 和 CC, A-6G 为 AA, AG 和 GG, C+31T 为 CC, CT 和 TT, 并记录了 2 个表型变量 phenol 与 disease。试用分数检验分析这 5 个位点 SNP 单体型与两个表型的关系。(下载数据:

http://www.empowerstats.com/empowerStats/exdata/hardyweinberg_data.xls)

输入界面:

单倍体型分数检验 ?

标题:

选择分析对象:

表型变量

变量名	分布类型	联系函数
PHENO1	Gaussian	
DISEASE	Binomial	

基因型/等位基因变量

变量
G217.A
G152.A
A.20C
A.6G
C.31T

选择调整变量:

变量

选择分层变量:

 Simulation p-value

基因型/等位基因变量类型

输出结果:

表型: PHEN01

整体分数检验

模型	Global score	自由度	P 值
相加模型	2.49946	7.00000	0.92714
显性模型	2.32257	7.00000	0.93985
隐形模型	0.00000	1.00000	0.99912

特定单倍体型分数

G217. A	G152. A	A. 20C	A. 6G	C. 31T	单倍体型频率	单倍体型. 分数	P 值
相加模型							
1	1	1	1	1	0.54471	0.34839	0.72755
1	1	1	1	2	0.09724	-0.51059	0.60964
1	1	1	2	1	0.12638	-0.64637	0.51804
1	1	2	1	1	0.02738	1.12703	0.25973
1	2	1	1	2	0.04431	0.70691	0.47963
2	1	2	2	1	0.03128	-0.33148	0.74028
2	1	2	2	2	0.03740	-0.02832	0.97741
显性模型							
1	1	1	1	1	0.54471	0.66783	0.50424
1	1	1	1	2	0.09724	-0.21130	0.83265
1	1	1	2	1	0.12638	-0.73790	0.46057
1	1	2	1	1	0.02738	1.12703	0.25973
1	2	1	1	2	0.04431	0.41349	0.67925
2	1	2	2	1	0.03128	-0.33148	0.74028
2	1	2	2	2	0.03740	-0.02832	0.97741
隐形模型							
1	1	1	1	1	0.54471	-0.00110	0.99912

表型: DISEASE

整体分数检验

模型	Global score	自由度	P 值
相加模型	2.49946	7.00000	0.92714
显性模型	2.32257	7.00000	0.93985
隐形模型	0.00000	1.00000	0.99912

特定单倍体型分数

G217. A	G152. A	A. 20C	A. 6G	C. 31T	单倍体型频率	单倍体型. 分数	P 值
---------	---------	--------	-------	--------	--------	----------	-----

相加模型							
1	1	1	1	1	0.54471	0.34839	0.72755
1	1	1	1	2	0.09724	-0.51059	0.60964
1	1	1	2	1	0.12638	-0.64637	0.51804
1	1	2	1	1	0.02738	1.12703	0.25973
1	2	1	1	2	0.04431	0.70691	0.47963
2	1	2	2	1	0.03128	-0.33148	0.74028
2	1	2	2	2	0.03740	-0.02832	0.97741
显性模型							
1	1	1	1	1	0.54471	0.66783	0.50424
1	1	1	1	2	0.09724	-0.21130	0.83265
1	1	1	2	1	0.12638	-0.73790	0.46057
1	1	2	1	1	0.02738	1.12703	0.25973
1	2	1	1	2	0.04431	0.41349	0.67925
2	1	2	2	1	0.03128	-0.33148	0.74028
2	1	2	2	2	0.03740	-0.02832	0.97741
隐形模型							
1	1	1	1	1	0.54471	-0.00110	0.99912

结果解释：

上述两个表现型 PHEN01 与 DISEASE 与这 5 个 SNP 的单倍体型分数检验均无统计学显著性。即这 5 个 SNP 的单倍体型与这两个表现型无关。