

Expected Conditional Characteristic Function based Measures for Testing Independence 复现

刘开元

2024 年 5 月

0 注意事项

此文档记录了一部分 Expected Conditional Characteristic Function based Measures for Testing Independence 文章中的模拟复现结果。给出的代码中, DGP.R 是数据的生成函数, function.cpp 是文章中的两种方法的 rcpp 版本。ExampleX.R 是例子 X 的复现代码, 没有使用并行技术。Example7_parallel.R 是例 7 的并行版本。AA.cpp 是我自己用作测试的代码。results 文件夹中是模拟结果 (RData) 以及绘制的图像 (EPS)。8014586 文件夹是文章给出的补充材料, 作者写好了 R 版本的代码, 可以用来参考写成 Rcpp 版本, 并对比运行速度。ECCFIC2024 是建立程序包时创建的文件夹。其余几个文章是相应的参考文献, matlab 代码是 HSIC 方法作者给出的源代码。

这里复现了文章中的方法 ECCFIC(kernel 以及 slicing), DCOV 和 DISCO 方法, 后两种方法在 energy 程序包中有现成的函数可用。rcpp 版本基于 RcppEigen 编写, 由于 kernel 方法需要进行矩阵的 Cholesky 分解, 因此使用 RcppEigen 库来做速度会提升很多 (快 10 倍左右)。其余需要注意的点就是关于某些例子, 需要对 Y 进行离散化, 用来使用切片方法和 DISCO 方法, 文章中只提到切片数量是多少, 但具体如何进行切片并没有说。模拟中使用了分位数切片方法进行离散化。

1 Example 1

表 1: Example 1 的结果

Model	n	Slice	Kernel	DCOV	DISCO
(a)	25	0.0930	0.0910	0.1070	0.0970
	50	0.0710	0.1040	0.0890	0.1010
	100	0.0870	0.1010	0.1020	0.0960
(b)	25	0.0750	0.0960	0.1000	0.0910
	50	0.0950	0.1050	0.1030	0.0850
	100	0.0960	0.0860	0.0970	0.0960
(c)	25	0.0930	0.0650	0.0750	0.1090
	50	0.0930	0.1000	0.1140	0.0870
	100	0.0880	0.0880	0.0860	0.0870
(d)	25	0.1030	0.0970	0.0890	0.1010
	50	0.0960	0.1100	0.1080	0.0980
	100	0.1210	0.0860	0.1040	0.1060

2 Example 2

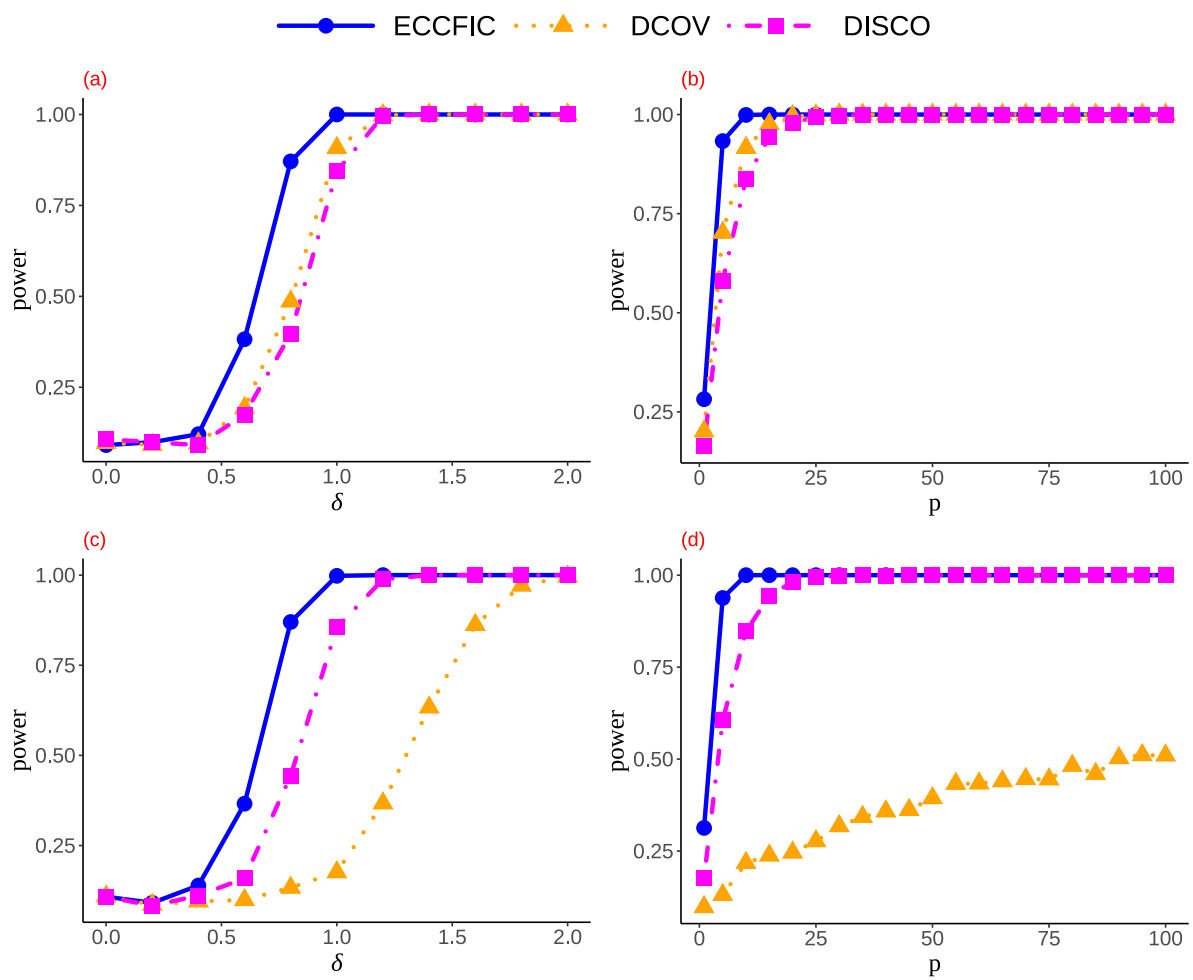


图 1: Example 2 的结果

3 Example 3

4 Example 4

表 2: Example 1 的结果

# of slices	2	5	10	23	46	115
ECCFIC	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050
DISCO	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050

5 Example 5

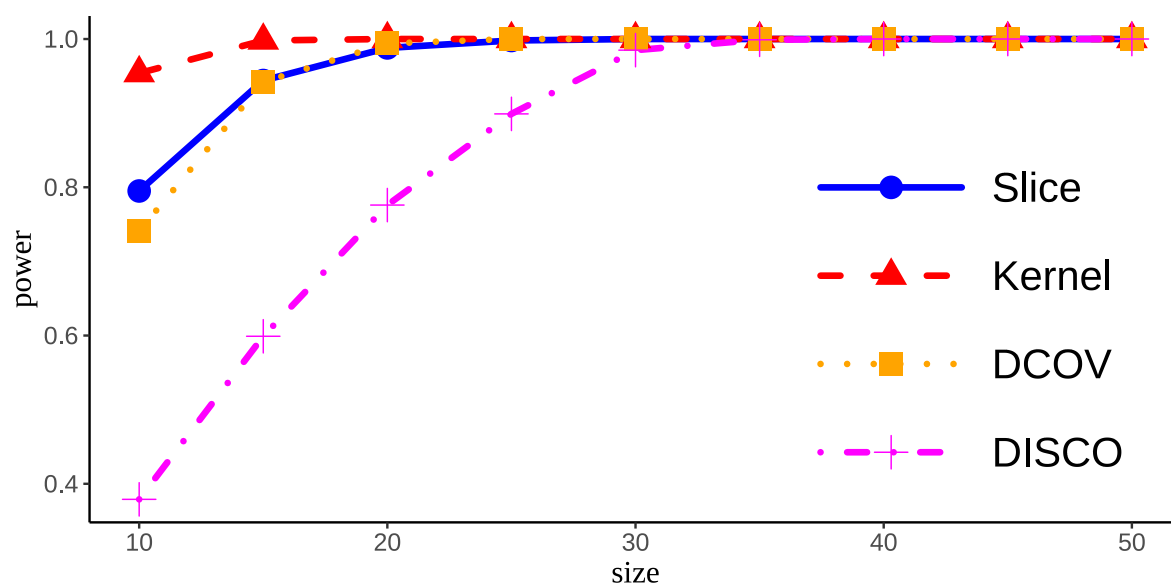


图 2: Example 5 的结果

6 Example 6

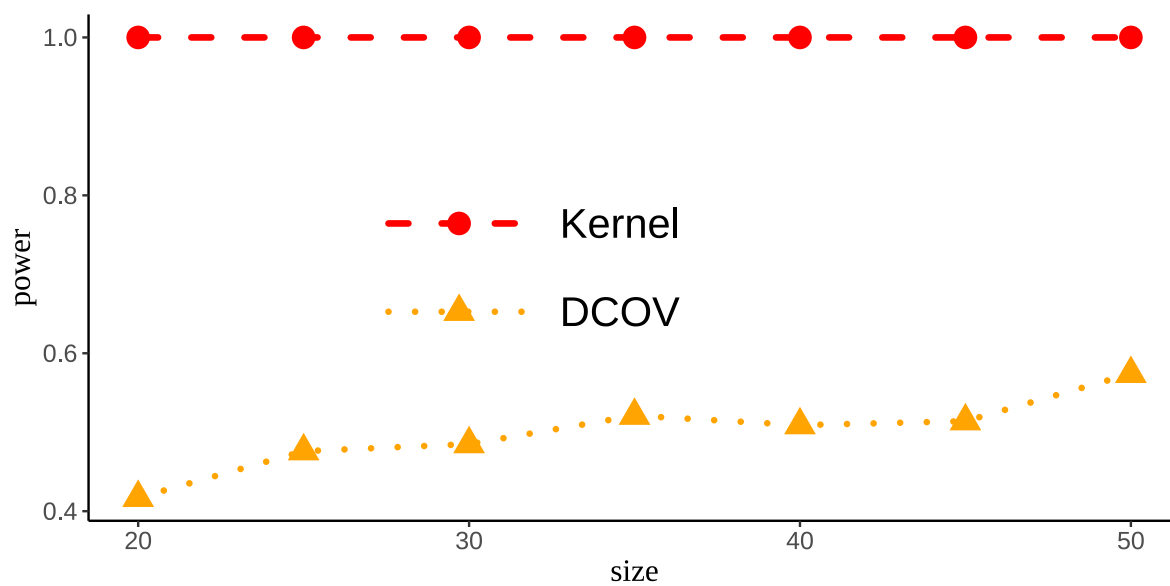


图 3: Example 6 的结果

7 Example 7

表 3: Example 7 的结果

Model	n	Slice	Kernel	DCOV	DISCO
A(1)	10	0.2270	0.2230	0.1940	0.1130
	20	0.3220	0.4660	0.4360	0.1650
	50	0.4130	0.6040	0.6530	0.2580
A(2)	10	0.1890	0.1910	0.1760	0.0920
	20	0.3480	0.4470	0.3850	0.2330
	50	0.2340	0.2800	0.2660	0.1150
A(3)	10	0.1380	0.3470	0.2520	0.1190
	20	0.4360	0.7530	0.6690	0.2990
	50	0.6380	0.9530	0.9690	0.6180
B(1)	10	0.1450	0.2710	0.2440	0.1360
	20	0.4280	0.7110	0.6620	0.4360
	50	0.2790	0.5060	0.4400	0.2610
B(2)	10	0.2290	0.6240	0.4080	0.1680
	20	0.9030	0.9870	0.9750	0.6750
	50	0.9980	1.0000	1.0000	0.9970
B(3)	10	0.1630	0.4450	0.3930	0.1470
	20	0.8810	0.9840	0.9740	0.9270
	50	0.6480	0.8590	0.8130	0.6410