基于Renyi检验的两ROC曲线比较

汇报人：余林 3157042040
指导老师：欧春泉教授

目 录

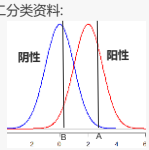
01 研究背景
02 方法介绍
03 模拟研究
04 实例分析
05 结论与讨论

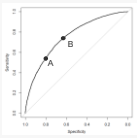
1 研究背景

▲二分类资料:

T 诊断结果	D（患病情况） D+	D-	合计
T+	a	b	a+b
T-	c	d	c+d
合计	a+c	b+d	n

▲非二分类资料:





1 研究背景

▲AUC

假设检验：H&M法、Delong法等

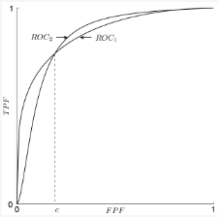
▲现存问题

Q1：当两ROC曲线相交时，使 用AUC作为评价指标得到的结论是否可靠？

Q2：使用pAUC作为评价指标存在缺陷。

1、区域选择问题

2、丢失部分信息



2 方法介绍

▲Renyi检验

（1）计算 t_j时刻前累计加权差之和 $Z(t_j)$

$$Z(t_j)=\sum_{t_k\le t_j} w(t_k)\left[d_{ik}-Y_{ik}\left(\frac{d_k}{Y_k}\right)\right], j=1,2,\ldots r$$

（2）取 $|Z(t_j)|$的上确界来构造统计量。检验统计量$Q=\sup\{|Z(t_j)|, t_j\le\tau\}/\sigma(\tau)$

▲统计量构造

（1）计算每个诊断界值点 T_j 下实际新增阳性数P_{ik} 和理论新增阳性数P_k

$$Z'(T_j)=\sum_{T_k\le T_j} w(T_k)\left[P_{ik}-P_{ik}\left(\frac{P_k}{N_k}\right)\right], j=1,2,\ldots r$$

（2）取累计加权值中数值最大值和方差值构建统计量， $Q'=\sup\{|Z'(T_j)|, T_j\le\tau\}/\sigma'(\tau)$

2 方法介绍

▲统计量分布

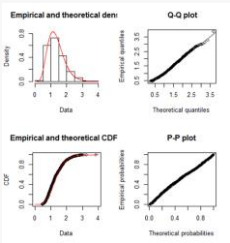
重抽样得到统计量的频数分布图

↓

观察直方图，总结出可能服从的分布

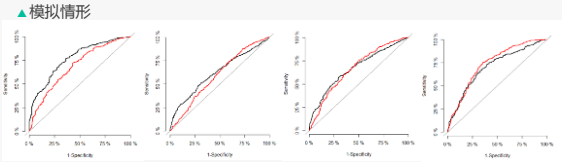
↓

利用拟合函数对其进行拟合



3 模拟研究

▲模拟情形



▲模拟数据的产生

--诊断结果服从广义伽马分布

--重抽样得到统计量Q的分布

--显著性水平α=0.05

▲模拟评价指标

1、 一类错误

2、 检验效能

3 模拟研究

▲参数设置

--无相交时

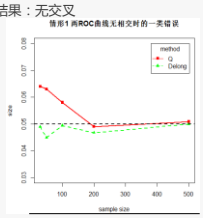
μ	$\mu_{case1}=\mu_{case2}=2$ $\mu_{case1}=2$ $\mu_{case2}=1.25$ $\mu_{control1}=\mu_{control2}=1$
σ	1
样本量	$n_{case1}=n_{case2}=n_{control1}=n_{control2}=30,50,100,200,500$ $n_{case1}=n_{case2}=100$，$n_{control1}=n_{control2}=50$ $n_{case1}=n_{case2}=50$，$n_{control1}=n_{control2}=100$
模拟次数	10000 次

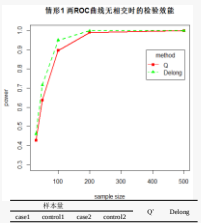
--有相交时

λ_1	$\lambda_{case1}=1.3575$ $\lambda_{control1}=-1.3575$， $\lambda_{case2}=0.1653$ $\lambda_{control2}=-0.1653$
λ_2	后部交叉：$\lambda_{case1}=0.012$ $\lambda_{control1}=0.02$ $\lambda_{case2}=0.014$ $\lambda_{control2}=0.014$ 中部交叉：$\lambda_{case1}=0.02$ $\lambda_{control1}=0.03$， $\lambda_{case2}=0.014$ $\lambda_{control2}=0.014$ 前部交叉：$\lambda_{case1}=0.02$ $\lambda_{control1}=0.012$ $\lambda_{case2}=0.014$ $\lambda_{control2}=0.014$
λ_3	$\lambda_{case1}=0.009695$ $\lambda_{control1}=0.009695$， $\lambda_{case2}=0.009695$ $\lambda_{control2}=0.028$
λ_4	$\lambda_{case1}=0.028$ $\lambda_{control1}=0.028$， $\lambda_{case2}=0.028$ $\lambda_{control2}=0.009695$
样本量	$n_{case1}=n_{case2}=n_{control1}=n_{control2}=30,50,100,200,500$
模拟次数	外循环 5000 次 内循环 1000 次

3 模拟研究

▲模拟结果：无交叉



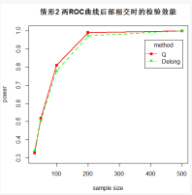


α					
case1	control1	case2	control2	Q'	Delong
100	50	100	50	0.067	0.0531
50	100	50	100	0.059	0.0508

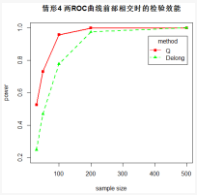
样本量					
case1	control1	case2	control2	Q'	Delong
100	50	100	50	0.259	0.28065
50	100	50	100	0.236	0.2229

3 模拟研究

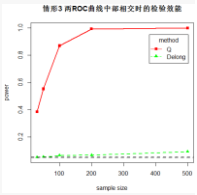
▲模拟结果：有相交



后部相交



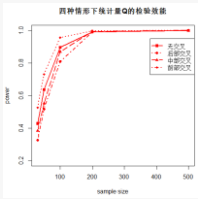
前部相交

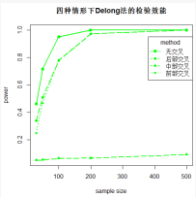


中部相交

3 模拟研究

▲总结



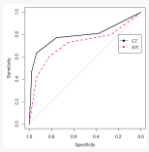


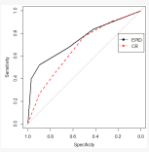
1、 无交叉时，Delong法能较好地控制一类错误，其检验效能也高于统计量Q'

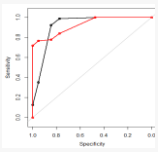
2、 有交叉时：统计量Q' 的检验效能均高于Delong法
Delong法：在前后部相交时检验效能差异不大，中部交叉时不适用；
统计量Q' ：在前部和中部相交时表现更好

4 实例分析

▲实例数据







▲结果

	实例1		实例2		实例3	
	统计量	P值/95%分位数	统计量	P值/95%分位数	统计量	P值/95%分位数
Delong	-2.953	0.003	-4.3978	<0.001	-0.741	0.459
Q	16.672	2.355	2.288777	1.888	7.948	1.544

5 结论与讨论

▲结论

1、在实际应用中，先绘制ROC曲线图，观察两ROC曲线是否存在相交的情况，以及其相交情形

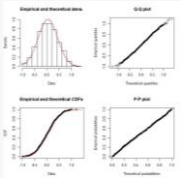
2、若无相交，则可使用AUC来进行统计推断；若存在相交，建议使用两ROC曲线比较的方法

▲讨论

统计量的分布:

$Z'(T_j) = \sum_{T_k \le T_j} w(T_k)\left[P_{ik} - P_{ik}\left(\frac{P_k}{N_k}\right)\right]$; w取总阳性数的倒数

$Q' = \sup\{|Z'(T_j)|, T_j \le \tau\}/\sigma'(\tau)$ $\xrightarrow{2.5\log}$



THANKS!