NRI, IDI, cNRI, dAUC, Bootstrap Confidence Intervals, Type I Error, Power

Comparison of Evaluation Index for improving model performance using Bootstrap Confidence Intervals

Purpose

- NRI, IDI, cNRI, dAUC 지표 타당성 비교
- 각 지표에 대한 구간 추정 결과,
  1) 실제로 모형 성능 향상이 없을 때 있다고 판단할 확률**(1종 오류)이 통제되는지**
  2) 모형 성능 향상이 명백해 짐에 따라 **얼마나 빠르게 성능 향상이 있다고 판단하는지**
  3) 위 두가지를 판단하기 위해 추정한 **신뢰 구간의 포함 확률이 적절한지**
  다양한 상황의 모의실험을 통해 비교

Methodology

- 특정 질병과 관련된 예후 인자를 식별하기 위한 지표로 NRI, IDI, cNRI, dAUC 존재 (Pencina et al. 2011)
- **정규성 가정 하에서** 가설 검정 시, 귀무가설 하에서 정규성 가정이 위배돼 **제1종 오류가 너무 크거나 검정력이 매우 작은 상황 발생**
  → **Bootstrap 신뢰구간** 사용하여 해결 (Shao et al. 2015, Olga et al. 2017)
- **포함비율(coverage probability)**와 **검정력 함수(power function)**

$$\text{포함비율} = P\left(\{CI_{lower} \le \text{모수값} \le CI_{upper}\}\right)$$

$$\text{검정력 함수} = 1 - P\left(\{CI_{lower} \le 0 \le CI_{upper}\}\right)$$

Definition

1) 지표 정의

M_1: 기존 모형

M_2: 새로운 바이오 마커 추가한 모형

$D = \{0,1\}$: 질병의 여부를 나타내는 이진 변수

$$\textbf{NRI} = [P(up|D=1) - P(down|D=1)] + [P(down|D=0) - P(up|D=0)]$$

$$\textbf{c.NRI} = E\{sign(P(M_2) - P(M_1))|D=1\} - E\{sign(P(M_2) - P(M_1))|D=0\}$$

$$\textbf{IDI} = E\{P(M_2) - P(M_1)|D=1\} - E\{P(M_2) - P(M_1)|D=0\}$$

$$\boldsymbol{\Delta}\textbf{AUC} = AUC(M_2) - AUC(M_1)$$

2) 신뢰구간

Asymptotic - 점근적 정규성 근사를 적용한 신뢰구간

$$\hat{\theta} - z_{\frac{\alpha}{2}}\widehat{se}(\hat{\theta}) < \theta_0 < \hat{\theta} + z_{\frac{\alpha}{2}}\widehat{se}(\hat{\theta})$$

Boot I - Asymptotic 신뢰구간에서 $\widehat{se}(\hat{\theta})$을 붓스트랩 표준오차로 대체한 신뢰구간

$$\hat{\theta} - z_{\frac{\alpha}{2}}\widehat{se}^*(\hat{\theta}) < \theta_0 < \hat{\theta} + z_{\frac{\alpha}{2}}\widehat{se}^*(\hat{\theta})$$

Boot II - $\tau = |\hat{\theta} - \theta_0|$의 붓스트랩 버전인 $\tilde{\tau} = |\tilde{\theta} - \hat{\theta}|$을 사용한 신뢰구간

$$\hat{\theta} - Q_{\tilde{\tau}}(1-\alpha) < \theta_0 < \hat{\theta} + Q_{\tilde{\tau}}(1-\alpha)$$

1) Simulation

Scenario I

- True Model : $\log\frac{p}{1-p} = 1 - 2X_1 + 3kX_2$

  $X_i \sim N(1,2^2)$, $i = 1,2$

  $k = \{0,0.01,0.05,0.1,0.2,0.3,0.4,0.6,0.8,1\}$
- $M_1: \log\frac{p}{1-p} = \alpha_0 + \alpha_1 X_1$

  $M_2: \log\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$
- 단순한 두 모형 비교

Scenario II

- True Model : $\log\frac{p}{1-p} = -3 + 0.02X_1 + 0.9X_2 + 0.01X_3 + 0.01X_4 - 0.01X_5 + 0.2X_6 + 1.5kX_7$

  $X_1 \sim N(40,12^2)$, $X_i \sim Ber(0.5), i = 2,6$, $X_3 \sim Unif(120,320)$, $X_4 \sim N(110,1^2)$

  $X_5 \sim Unif(30,70)$, $X_7 \sim N(1,2^2)$

  $k = \{0,0.01,0.05,0.1,0.2,0.3,0.4,0.6,0.8,1\}$
- $M_1: \log\frac{p}{1-p} = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5 + \alpha_6 X_6$

  $M_2: \log\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$
- 각 설명 변수는 나이, 성별, 총 콜레스테롤, SBP, HDL, 흡연여부와 같은 범위를 갖음.

  → **현실에 대응하는 모형**

Scenario III

- True Model : $\log\frac{p}{1-p} = 1 + 0.5X_1 - 2kX_2 + 1.5kX_3 + 3kX_4 - 2.7kX_5 - kX_6$

  $X_i \sim N(1,2^2)$, $i = 1,2,3,4,5,6$

  $k = \{0,0.01,0.05,0.1,0.2,0.3,0.4,0.6,0.8,1\}$
- $M_1: \log\frac{p}{1-p} = \alpha_0 + \alpha_1 X_1$

  $M_2: \log\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6$
- $k = 0$일 때, M_2가 매우 **과적합되는** 양상을 보일 것이며,

  $k \neq 0$에서는, M_1이 M_2에 비해 훨씬 안좋은 모형이 될 것임

Scenario IV

- True Model : $\log\frac{p}{1-p} = 1 + 2X_1 - 0.9kX^{2k+1}$

  $X_i \sim Unif(0,2)$, $k = \{0,0.01,0.05,0.1,0.2,0.3,0.4,0.6,0.8,1\}$
- $M_1: \log\frac{p}{1-p} = \alpha_0 + \alpha_1 X_1$

  $M_2: \log\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2^{2k+1}$
- 두 모형이 **동일한 설명변수를 사용하지만,**

  모형에 차이가 있는 경우

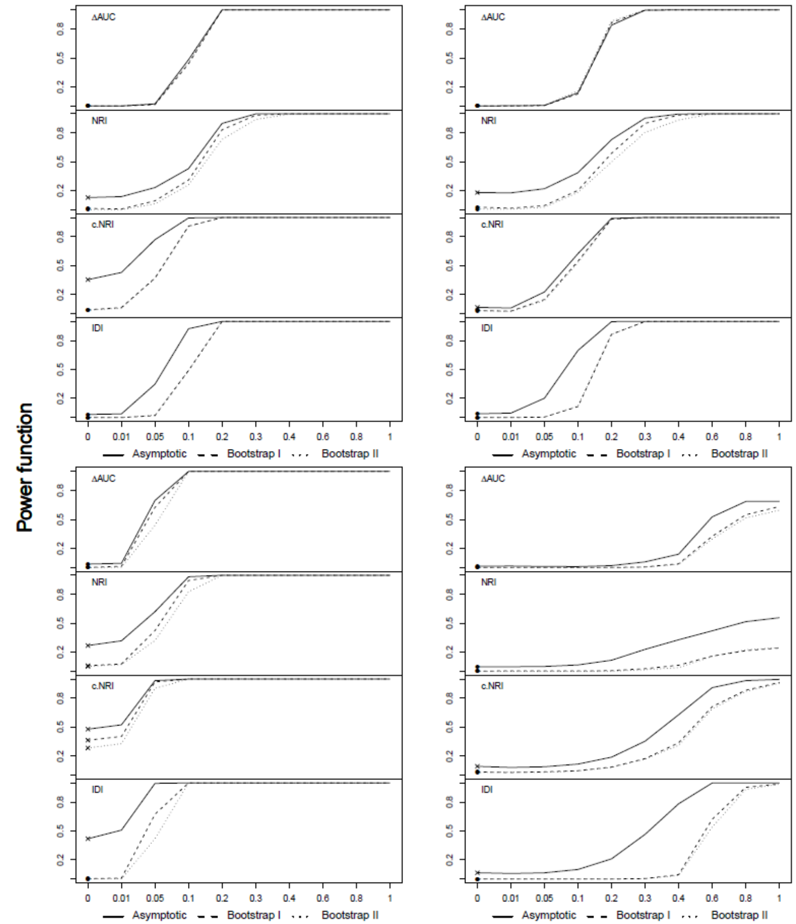
2) Results

k	n	dAUC TRUE	dAUC Asymptotic CP	dAUC Asymptotic P	dAUC I CP	dAUC I P	dAUC II CP	dAUC II P	NRI TRUE	NRI Asymptotic CP	NRI Asymptotic P	NRI I CP	NRI I P	NRI II CP	NRI II P	cNRI TRUE	cNRI Asymptotic CP	cNRI Asymptotic P	cNRI I CP	cNRI I P	cNRI II CP	cNRI II P	IDI TRUE	IDI Asymptotic CP	IDI Asymptotic P	IDI I CP	IDI I P	IDI II CP	IDI II P
0	500	0.0000	1	0	1	0	1	0	0.0000	0.872	0.128	0.984	0.016	0.996	0.004	0.0000	0.646	0.354	0.96	0.04	0.963	0.037	0.0000	0.968	0.032	1	0	1	0
0.01	500	0.0001	0.993	0.001	1	0	1	0	0.0007	0.789	0.138	0.995	0.009	0.999	0.002	0.0644	0.683	0.429	0.972	0.064	0.975	0.058	0.0002	0.913	0.038	1	0	1	0
0.05	500	0.0015	0.885	0.022	0.951	0.015	0.98	0.018	0.0050	0.865	0.232	0.96	0.094	0.969	0.063	0.3119	0.669	0.77	0.936	0.371	0.966	0.367	0.0056	0.75	0.347	0.922	0.02	0.975	0.024
0.1	500	0.0060	0.929	0.482	0.941	0.44	0.948	0.448	0.0177	0.889	0.429	0.938	0.31	0.947	0.26	0.5811	0.699	0.996	0.954	0.911	0.958	0.908	0.0217	0.751	0.924	0.931	0.49	0.943	0.489
0.2	500	0.0227	0.945	1	0.945	1	0.944	1	0.0619	0.9	0.899	0.938	0.833	0.945	0.733	0.9513	0.786	1	0.943	1	0.946	1	0.0781	0.765	1	0.946	1	0.946	1
0.3	500	0.0473	0.946	1	0.948	1	0.947	1	0.1203	0.922	0.998	0.948	0.984	0.96	0.94	1.1685	0.872	1	0.959	1	0.962	1	0.1536	0.777	1	0.952	1	0.948	1
0.4	500	0.0762	0.941	1	0.942	1	0.938	1	0.1836	0.905	1	0.936	1	0.945	0.995	1.3064	0.863	1	0.943	1	0.946	1	0.2346	0.768	1	0.951	1	0.951	1
0.6	500	0.1356	0.957	1	0.958	1	0.957	1	0.3045	0.905	1	0.956	1	0.965	1	1.4792	0.92	1	0.959	1	0.962	1	0.3862	0.747	1	0.946	1	0.948	1
0.8	500	0.1884	0.952	1	0.95	1	0.949	1	0.4071	0.842	1	0.946	1	0.955	1	1.5861	0.923	1	0.956	1	0.959	1	0.5076	0.742	1	0.952	1	0.953	1
1	500	0.2313	0.958	1	0.958	1	0.959	1	0.4887	0.815	1	0.953	1	0.964	1	1.6593	0.923	1	0.944	1	0.946	1	0.5986	0.767	1	0.951	1	0.948	1

k	n	dAUC TRUE	dAUC Asymptotic CP	dAUC Asymptotic P	dAUC I CP	dAUC I P	dAUC II CP	dAUC II P	NRI TRUE	NRI Asymptotic CP	NRI Asymptotic P	NRI I CP	NRI I P	NRI II CP	NRI II P	cNRI TRUE	cNRI Asymptotic CP	cNRI Asymptotic P	cNRI I CP	cNRI I P	cNRI II CP	cNRI II P	IDI TRUE	IDI Asymptotic CP	IDI Asymptotic P	IDI I CP	IDI I P	IDI II CP	IDI II P
0	500	0.0000	0.999	0.001	1	0	1	0	0.0000	0.82	0.18	0.972	0.028	0.993	0.007	0.0000	0.934	0.066	0.967	0.033	0.969	0.031	0.0000	0.961	0.039	0.999	0.001	0.999	0.001
0.01	500	0.0001	0.994	0.003	1	0	0.998	0.002	0.0011	0.818	0.177	0.984	0.018	0.994	0.007	0.0237	0.969	0.059	0.981	0.027	0.982	0.025	0.0002	0.937	0.046	1	0	1	0
0.05	500	0.0035	0.859	0.006	0.998	0.006	0.997	0.01	0.0066	0.866	0.22	0.972	0.045	0.986	0.028	0.1202	0.971	0.226	0.986	0.148	0.99	0.137	0.0038	0.697	0.201	0.997	0.005	0.998	0.005
0.1	500	0.0137	0.899	0.138	0.912	0.126	0.936	0.158	0.0220	0.847	0.387	0.949	0.203	0.969	0.182	0.2376	0.921	0.625	0.935	0.546	0.957	0.531	0.0149	0.709	0.698	0.925	0.114	0.943	0.121
0.2	500	0.0486	0.922	0.841	0.918	0.867	0.918	0.884	0.0753	0.9	0.731	0.945	0.588	0.963	0.498	0.4600	0.946	0.992	0.967	0.985	0.969	0.983	0.0544	0.735	0.998	0.94	0.866	0.943	0.866
0.3	500	0.0936	0.926	0.997	0.914	0.998	0.912	0.998	0.1458	0.898	0.954	0.948	0.9	0.968	0.806	0.6561	0.932	1	0.946	1	0.957	1	0.1097	0.731	1	0.94	0.999	0.94	0.999
0.4	500	0.1398	0.919	1	0.916	1	0.914	1	0.2206	0.909	0.997	0.958	0.986	0.982	0.935	0.8248	0.94	1	0.959	1	0.963	1	0.1728	0.754	1	0.942	1	0.94	1
0.6	500	0.2200	0.93	1	0.922	1	0.926	1	0.3573	0.885	1	0.964	1	0.983	0.997	1.0824	0.937	1	0.955	1	0.957	1	0.2993	0.789	1	0.947	1	0.946	1
0.8	500	0.2796	0.92	1	0.902	1	0.923	1	0.4656	0.855	1	0.952	1	0.982	1	1.2599	0.924	1	0.944	1	0.952	1	0.4081	0.777	1	0.941	1	0.939	1
1	500	0.3221	0.917	1	0.906	1	0.926	1	0.5478	0.837	1	0.959	1	0.985	1	1.3841	0.928	1	0.954	1	0.961	1	0.4949	0.821	1	0.955	1	0.952	1

k	n	dAUC TRUE	dAUC Asymptotic CP	dAUC Asymptotic P	dAUC I CP	dAUC I P	dAUC II CP	dAUC II P	NRI TRUE	NRI Asymptotic CP	NRI Asymptotic P	NRI I CP	NRI I P	NRI II CP	NRI II P	cNRI TRUE	cNRI Asymptotic CP	cNRI Asymptotic P	cNRI I CP	cNRI I P	cNRI II CP	cNRI II P	IDI TRUE	IDI Asymptotic CP	IDI Asymptotic P	IDI I CP	IDI I P	IDI II CP	IDI II P
0	500	0.0000	0.964	0.036	0.999	0.001	1	0	0.0000	0.731	0.269	0.94	0.06	0.948	0.052	0.0000	0.521	0.479	0.636	0.364	0.716	0.284	0.0000	0.581	0.419	0.995	0.005	0.999	0.001
0.01	500	0.0012	0.972	0.044	0.993	0.013	0.999	0.003	0.0031	0.736	0.318	0.931	0.077	0.943	0.071	0.0796	0.762	0.524	0.846	0.403	0.869	0.328	0.0015	0.634	0.509	0.995	0.009	0.998	0.002
0.05	500	0.0268	0.952	0.698	0.964	0.626	0.983	0.439	0.0436	0.818	0.621	0.9	0.425	0.921	0.321	0.3797	0.914	0.984	0.953	0.971	0.96	0.904	0.0349	0.746	0.995	0.967	0.674	0.981	0.419
0.1	500	0.0879	0.963	0.999	0.967	0.999	0.969	0.999	0.1425	0.848	0.984	0.924	0.945	0.939	0.828	0.6967	0.935	1	0.962	1	0.969	1	0.1240	0.766	1	0.951	1	0.966	1
0.2	500	0.2114	0.944	1	0.947	1	0.943	1	0.3506	0.815	1	0.902	1	0.925	1	1.1091	0.934	1	0.957	1	0.963	1	0.3351	0.744	1	0.94	1	0.946	1
0.3	500	0.2938	0.962	1	0.967	1	0.97	1	0.4996	0.782	1	0.906	1	0.915	1	1.3373	0.923	1	0.949	1	0.957	1	0.4946	0.766	1	0.932	1	0.938	1
0.4	500	0.3449	0.949	1	0.959	1	0.963	1	0.5996	0.674	1	0.853	1	0.878	1	1.4764	0.92	1	0.945	1	0.95	1	0.6007	0.758	1	0.934	1	0.938	1
0.6	500	0.4001	0.951	1	0.945	1	0.945	1	0.7192	0.569	1	0.76	1	0.794	1	1.6351	0.917	1	0.94	1	0.948	1	0.7244	0.749	1	0.929	1	0.933	1
0.8	500	0.4279	0.945	1	0.922	1	0.939	1	0.7864	0.496	1	0.693	1	0.769	1	1.7214	0.916	1	0.947	1	0.952	1	0.7915	0.788	1	0.935	1	0.941	1
1	500	0.4442	0.931	1	0.884	1	0.967	1	0.8289	0.43	1	0.696	1	0.749	1	1.7752	0.924	1	0.954	1	0.954	1	0.8327	0.797	1	0.942	1	0.946	1

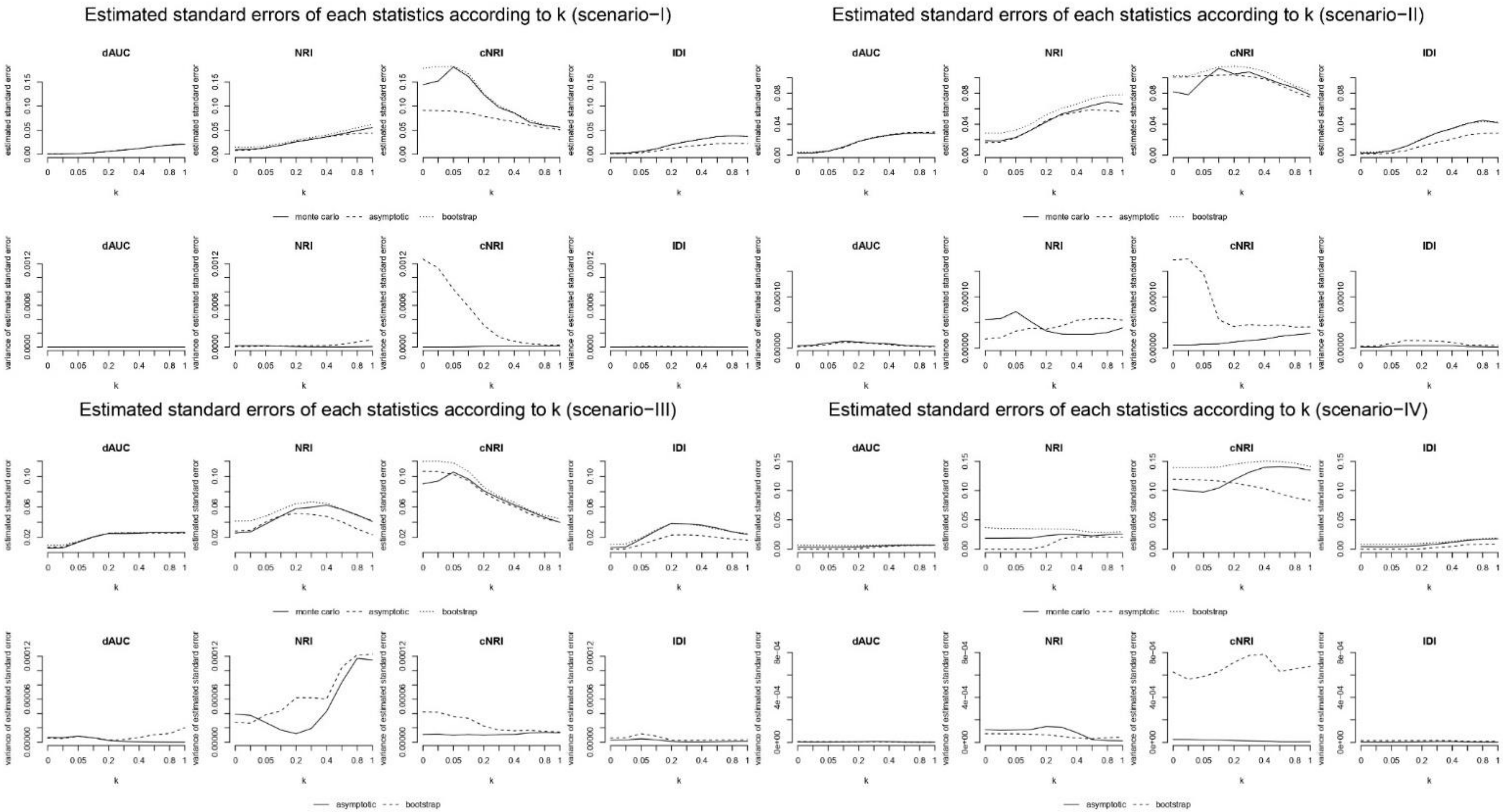
k	n	dAUC TRUE	dAUC Asymptotic CP	dAUC Asymptotic P	dAUC I CP	dAUC I P	dAUC II CP	dAUC II P	NRI TRUE	NRI Asymptotic CP	NRI Asymptotic P	NRI I CP	NRI I P	NRI II CP	NRI II P	cNRI TRUE	cNRI Asymptotic CP	cNRI Asymptotic P	cNRI I CP	cNRI I P	cNRI II CP	cNRI II P	IDI TRUE	IDI Asymptotic CP	IDI Asymptotic P	IDI I CP	IDI I P	IDI II CP	IDI II P
0	500	0.0000	0.986	0.014	1	0	1	0	0.0000	0.951	0.049	0.998	0.002	0.998	0.002	0.0000	0.908	0.092	0.966	0.034	0.971	0.029	0.0000	0.934	0.066	1	0	1	0
0.01	500	0.0000	0.379	0.016	0.998	0.002	0.999	0.002	0.0001	0.27	0.048	0.994	0.006	0.999	0.001	0.0119	0.937	0.081	0.972	0.028	0.974	0.031	0.0000	0.857	0.058	1	0	1	0
0.05	500	0.0001	0.354	0.011	0.999	0.001	0.999	0.001	0.0008	0.259	0.051	0.995	0.005	0.999	0.001	0.0298	0.936	0.088	0.976	0.035	0.979	0.03	0.0002	0.248	0.065	1	0	1	0
0.1	500	0.0003	0.371	0.01	0.999	0.001	1	0.001	0.0016	0.312	0.067	0.996	0.004	0.999	0.001	0.0624	0.945	0.116	0.977	0.047	0.984	0.045	0.0006	0.276	0.1	1	0	1	0
0.2	500	0.0012	0.505	0.02	1	0	1	0	0.0039	0.44	0.116	0.993	0.008	0.999	0.002	0.1215	0.935	0.188	0.989	0.085	0.996	0.081	0.0027	0.368	0.21	1	0	1	0
0.3	500	0.0031	0.679	0.059	0.999	0.006	0.999	0.005	0.0064	0.629	0.229	0.984	0.028	0.994	0.017	0.1773	0.9	0.354	0.974	0.173	0.987	0.168	0.0069	0.477	0.465	0.994	0.005	0.996	0.005
0.4	500	0.0058	0.819	0.139	0.969	0.039	0.987	0.034	0.0086	0.734	0.329	0.976	0.064	0.98	0.039	0.2348	0.861	0.629	0.961	0.337	0.982	0.311	0.0141	0.587	0.783	0.939	0.046	0.984	0.037
0.6	500	0.0113	0.919	0.527	0.96	0.325	0.971	0.297	0.0124	0.783	0.422	0.944	0.159	0.957	0.156	0.3477	0.784	0.911	0.952	0.713	0.959	0.688	0.0365	0.674	0.999	0.946	0.622	0.961	0.54
0.8	500	0.0133	0.929	0.687	0.947	0.549	0.959	0.516	0.0197	0.811	0.517	0.949	0.215	0.967	0.225	0.4596	0.74	0.984	0.952	0.884	0.963	0.872	0.0530	0.687	1	0.941	0.954	0.946	0.932
1	500	0.0136	0.935	0.689	0.946	0.634	0.958	0.596	0.0239	0.81	0.558	0.959	0.243	0.986	0.246	0.5586	0.725	0.997	0.951	0.965	0.956	0.956	0.0604	0.666	1	0.944	0.989	0.946	0.987



각 Scenario에서의 모의실험 결과

Scenario III을 제외한 나머지 모의실험에서는 모든 지표가 제1종 오류와 검정력에 큰 문제가 없었지만, Scenario III에서 $\Delta AUC, IDI$를 제외한 $NRI, cNRI$는 Bootstrap 신뢰구간을 사용하더라도 제1종 오류를 통제하지 못하는 것을 볼 수 있다.

2) Results



각 지표의 monte carlo simulation을 통한 표준오차, 공식 기반의 표준오차, Bootstrap 표준오차

일반적으로 $NRI, cNRI$는 세 표준오차 간의 차이가 크기 때문에 앞선 모의실험에서 문제가 생긴 것으로 생각할 수 있다.