Statistical Inference for Ranks of Health Care Facilities

in the Presence of Ties and Near Ties

Minge Xie

Department of Statistics
Rutgers, The State University of New Jersey

Supported in part by NSF, NSA and DHS

RUTGERS

Institutional Ranking

- We often need to evaluate performance of institutions, which "inevitably" lead to "institutional ranking" (Goldstein and Spiegelhalter, '96)
  - health care facilities, schools, financial institutions, etc.
- There is an emphasis for "interval estimation" to display uncertainty associated with rankings (GS, '96, among others)
- Two approaches are recommended for deriving confidence intervals of ranks (GS, '96)
  - Frequentist bootstrap sampling method
  - Bayesian computational method

  -- Many publications (thousands) have used these approaches!

RUTGERS

VHA Study

- An observational study on diabetes patients in Veteran Health Administration (VHA) facilities
  - Objective is to evaluate the quality of care for diabetes patients in VA hospitals across the US
- Study population
  - Veterans who used the VHA facilities in Fiscal Year 1999 or 2000 and in at least 1 year received a diabetes medication, plus other inclusive/exclusive criteria
  - → This identified 250,317 patients in 79 VHA facilities.
- Our goal in this talk is to construct confidence intervals for the ranks of the facilities
  - Based on "poor A1c control" rates (Pogach et al. 2005).

RUTGERS

Population and Sample ranks

Interested to rank *k* institutions through a specific characteristic described by parameters $\theta_1, \ldots, \theta_k$.

- Population rank of the *i*th institution is

$$R_i = 1 + \sum_{j \neq i} G(\theta_j - \theta_i) = 1 + \sum_{j \neq i} \mathbf{1}_{(\theta_j < \theta_i)} + \frac{1}{2} \sum_{j \neq i} \mathbf{1}_{(\theta_j = \theta_i)},$$

where $G(t) = \mathbf{1}_{(t<0)} + (1/2)\mathbf{1}_{(t=0)}$, and $\mathbf{1}_{(A)}$ is the 0–1 indicator function.

- Sample rank is an estimator of the population rank

$$\hat{R}_{in} = 1 + \sum_{j \neq i} G(\hat{\theta}_{jn} - \hat{\theta}_{in}) = 1 + \sum_{j \neq i} \mathbf{1}_{(\hat{\theta}_{jn} \leq \hat{\theta}_{in})}$$

with $\hat{\theta}_{in} = \theta_i + o_P(1)$

RUTGERS

Bootstrap Confidence Intervals

Let $\hat{R}_i^*$ be the sample rank from a bootstrap sample

- Conventional bootstrap approaches (Efron&Tibshirani,'94).
  - I. Centered bootstrap percentile intervals derived from $\hat{R}_i - R_i \sim_{dist} \hat{R}_i^* - \hat{R}_i$
  - II. Regular bootstrap percentile intervals directly derived from the sample bootstrap distribution of $\hat{R}_i^*$
    - The intervals from Method II are also asymptotically equivalent to a set of Bayesian credible intervals.

  Note: These two methods are the same, if the bootstrap distribution is symmetric; but they need not be the same otherwise.

RUTGERS

Problems

- Potential problems:
  - Conventional bootstrap methods may not work!
    - Population rank is a discrete parameter (a "difficult parameter" by Snijders (1996))
  - Bayesian credible intervals do not guarantee frequentist coverage
- Lack of theoretical results on rankings in frequentist inference, except in a simple no tie case

RUTGERS

Asymptotic result in no tie case

When $\theta_1, \ldots, \theta_k$ are all fixed and distinct.

$$P(\hat{R}_{in} = R_i) \to 1 \quad \text{as } n \to \infty$$

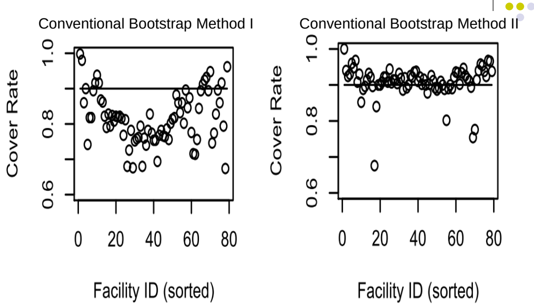
➔Implies: *When there is no tie to the facility i, the confidence intervals for $\boldsymbol{R}_i$ at any level always have asymptotically 100% coverage probability as long as the interval contains the estimator $\hat{R}_i$, regardless of the methods used.*

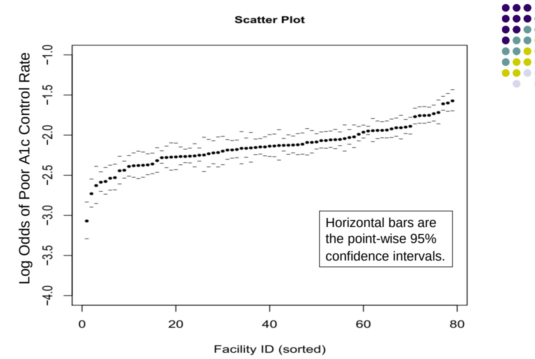
- This result of 100% coverage does not agree with most practical settings!

Table 1. "Poor A1c Control" data of 79 VHA facilities

Facility ID	Poor A1c control rate $\hat{p}_i$	No. of cases n_i	Facility ID	Poor A1c control rate $\hat{p}_i$	No. of cases n_i	Facility ID	Poor A1c control rate $\hat{p}_i$	No. of cases n_i	Facility ID	Poor A1c control rate $\hat{p}_i$	No. of cases n_i
1	0.0444	1,803	21	0.0939	3,811	41	0.1057	2,128	61	0.1247	13,598
2	0.0612	2,270	22	0.0940	6,641	42	0.1063	1,176	62	0.1251	1,758
3	0.0673	1,188	23	0.0943	5,577	43	0.1065	6,828	63	0.1255	3,020
4	0.0700	4,143	24	0.0946	2,252	44	0.1066	2,851	64	0.1256	3,624
5	0.0706	2,195	25	0.0953	6,902	45	0.1068	3,810	65	0.1260	3,254
6	0.0733	2,443	26	0.0956	1,015	46	0.1072	2,417	66	0.1276	2,719
7	0.0738	2,290	27	0.0969	1,558	47	0.1078	2,365	67	0.1291	2,779
8	0.0799	1,827	28	0.0978	2,208	48	0.1099	1,519	68	0.1292	1,648
9	0.0803	4,842	29	0.0982	1,293	49	0.1100	2,081	69	0.1301	4,573
10	0.0839	3,170	30	0.0998	1,423	50	0.1106	4,468	70	0.1313	3,230
11	0.0846	2,234	31	0.1010	2,940	51	0.1122	4,037	71	0.1456	3,380
12	0.0848	1,780	32	0.1011	3,649	52	0.1123	3,685	72	0.1472	3,044
13	0.0851	2,069	33	0.1017	3,620	53	0.1132	3,808	73	0.1474	2,952
14	0.0854	3,032	34	0.1029	1,078	54	0.1134	2,751	74	0.1477	3,576
15	0.0863	3,072	35	0.1030	2,990	55	0.1137	5,436	75	0.1502	3,182
16	0.0897	2,108	36	0.1035	1,266	56	0.1147	1,482	76	0.1521	1,407
17	0.0927	7,853	37	0.1040	4,502	57	0.1161	4,668	77	0.1666	3,812
18	0.0929	2,626	38	0.1045	4,513	58	0.1169	1,950	78	0.1680	2,321
19	0.0933	1,640	39	0.1046	3,059	59	0.1203	4,672	79	0.1718	1,624
20	0.0935	1,733	40	0.1055	2,284	60	0.1231	5,785			

Cover rate of 90% confidence intervals in 79 facilities





➔ The facilities cannot be distinguished from their neighboring facilities!

Ties and Near Ties

Assume, w.l.g., we are interested in facility 1, and suppose $\hat{\theta}_{in} = \theta_i + n^{-1/2}(Z_{in} + o_P(1)),\ Z_{in} \sim N(0, \sigma_i^2),\ i \leq k$

- Tie set:

$$\Theta_T = \{j : \theta_j = \theta_1,\ j = 2, 3, \ldots, k\}$$

- Near tie case:
  - Near tie definition:

$$|\theta_j - \theta_1| = O(n^{-\frac{1}{2}})$$

  - Near tie set:

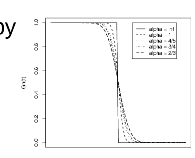
$$\Theta_N = \left\{ j : |\theta_j - \theta_1| \ = O(n^{-\frac{1}{2}}),\ j = 2, 3, \ldots, k \right\}$$

Problem in tie or near tie case

- The sample rank $\hat{R}_1$ is not even a consistent estimator of the population rank R_1:

$$\hat{R}_1 - R_1 = \sum_{j \in \Theta_T} \left\{ \mathbf{1}_{\left(Z_{jn} - Z_{1n} \leq o_P(1)\right)} - \frac{1}{2} \right\} + o_p(1) \not\to 0.$$

- Solution: Define smooth ranks by replacing the discrete indicator function $G(t) = \mathbf{1}_{(t<0)} + (1/2)\mathbf{1}_{(t=0)}$ with a smooth function $G_n(t)$!



(Xie, Singh and Zhang, 2009)

Smooth Rank Definition

- Define smooth population and sample ranks:

$$R_1^{(smooth)} = 1 + \sum_{j=2}^{k} G_n(\theta_j - \theta_1)$$

$$\hat{R}_1^{(smooth)} = 1 + \sum_{j=2}^{k} G_n(\hat{\theta}_{jn} - \hat{\theta}_{1n})$$

where $G_n(t) = 1 - \Phi(t/\tau_n)$ with $\tau_n \propto n^{-\alpha/2}$, for some $0 < \alpha < 1$.

Remark: these smooth ranks are closely related to the Bayesian ranks defined by Laird and Louis ('89), as well as the p-values for the tests $\mathrm{H}_0\colon\ \theta_j \le \theta_1$ vs $\mathrm{H}_1\colon\ \theta_j > \theta_1$.

RUTGERS UNIVERSITY

Summary of Theoretical Results (I)

No tie (distinct parameters) or ties case, but not near ties case

- Population rank:
  - $|R_1^{(smooth)} - R_1| \to 0$ at an exponential rate.
- Sample rank:
  - $\hat{R}_1^{(smooth)} \to R_1^{(smooth)}$ and $\to R_1$ at the rate $n^{-1/2}/\tau_n$
  - $\hat{R}_1^{(smooth)}$ is asymptotically normally distributed
  - Bootstrap interval based on $\hat{R}_1^{(smooth)}$ provides a consistent coverage probability for R_1 and $R_1^{(smooth)}$.

RUTGERS UNIVERSITY

Summary of Theoretical Results (II)

Near ties case (in presence of near ties)

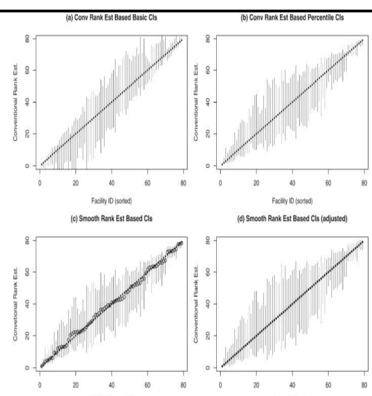
- Population rank:
  - $|R_1^{(smooth)} - R_1| \not\to 0$ but the difference is bounded.
- Sample rank:
  - $\hat{R}_1^{(smooth)} \to R_1^{(smooth)}$ but $\not\to R_1$
  - $\hat{R}_1^{(smooth)}$ is asymptotically normally distributed
  - Bootstrap interval based on $\hat{R}_1^{(smooth)}$ provides a consistent coverage probability for $R_1^{(smooth)}$. This interval can be modified to get a conservative confidence interval for R_1 .

RUTGERS UNIVERSITY

Numerical Example: VHA Study

- Four types of rank confidence intervals are constructed
  1) Conventional bootstrap confidence intervals I
  2) Conventional bootstrap confidence intervals II (asymptotic Bayesian credible intervals)
  3) Confidence intervals derived from smooth ranks
  4) Conservative confidence intervals derived from smooth ranks
- The tuning parameter in the smooth rank method is tuned using the trade-off between the coverage probability rates and the interval lengths (or the differences between the smooth and conventional ranks)

RUTGERS UNIVERSITY

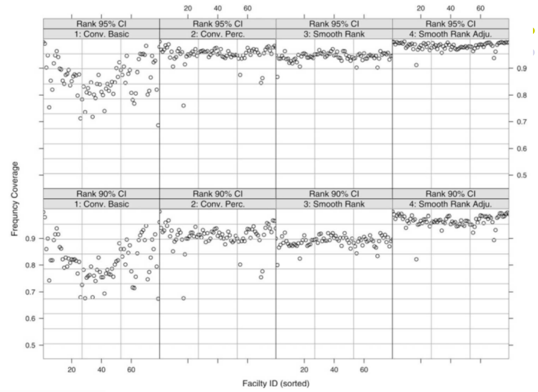


RUTGERS UNIVERSITY

Simulation Studies

A. Near tie case:
   Assume the parameter values listed in Table 1 are their "true" values, and simulation 1000 "observed" data sets using the sample sizes listed in Table 1. Based on these 1000 "observed" data sets, we construct 1000 confidence intervals and see how many of them cover their corresponding "true" ranks.
B. No tie case:
   Same as in A, except that we increase the sample size in each facility 20,000 times.
C. True tie (but no near tie) case:
   Same as in B, except that we group the 79 facilities into one group of 9 facilities and seven groups of 10 facilities. In each of group, we use their median value their "true" parameter value.

RUTGERS UNIVERSITY



Near tie case

Table 2. Summary statistics of the median lengths and coverages of the rank confidence intervals for 79 facilities in the simulations

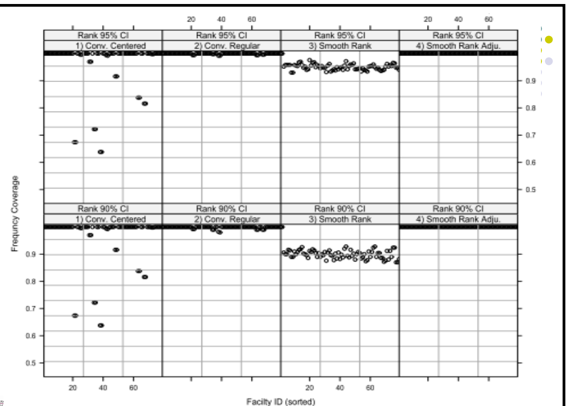
Panel 1: Near Ties Case (sample sizes = n_i)

		1) Conv. Centered		2) Conv. Perc.		3) Smooth Rank		4) Smooth Rank Adj.	
		90%	95%	90%	95%	90%	95%	90%	95%
Length	Min.	0.00	0.00	0.00	0.00	0.29	0.51	2.00	2.00
	First Qu.	12.50	14.50	12.50	14.50	11.99	14.33	16.50	18.50
	Median	18.00	22.00	18.00	22.00	17.87	21.47	23.00	26.00
	Mean	18.57	21.96	18.57	21.96	17.83	21.10	23.30	26.61
	Third Qu.	25.00	29.50	25.00	29.50	23.74	28.03	31.00	35.50
	Max.	38.00	43.00	38.00	43.00	35.43	41.04	44.00	50.00
Coverage	MSE_4	0.9601	0.9904	0.1690	0.0840	0.0447	0.0244	0.3902	0.1031
Coverage	MSE_5	0.9340	0.9841	0.1140	0.0644	0.0410	0.0220	0.0061	0.0016

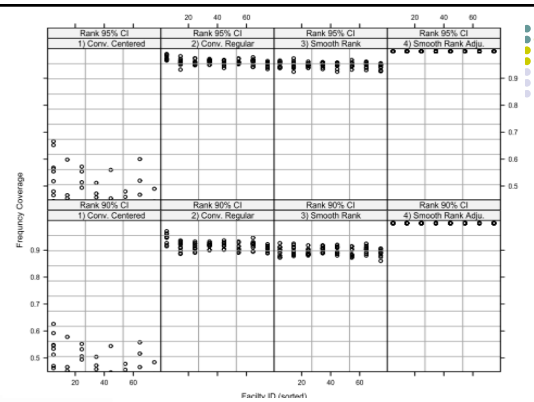
$$\mathrm{MSE}_4(\beta) = (1/79)\sum_{i=1}^{79}(c_i - \gamma)^2$$

$$\mathrm{MSE}_5(\beta) = (1/79)\sum_{i=1}^{79}(c_i - \gamma)^2 \mathbf{1}_{(c_i<\gamma)}$$

γ: significance level
c_i: coverage probability



No tie case



Tie case (no near ties)

Panel 3: True Ties Case (sample sizes = $20{,}000n_i$)

		1) Conv. Centered		2) Conv. Perc.		3) Smooth Rank		4) Smooth Rank Adj.	
		90%	95%	90%	95%	90%	95%	90%	95%
Length	Min.	5.0000	6.0000	5.0000	6.0000	0.7501	0.8931	6.0000	6.0000
	First Qu.	6.0000	7.0000	6.0000	7.0000	0.9084	1.0810	6.0000	6.0000
	Median	7.0000	8.0000	7.0000	8.0000	0.9930	1.1810	6.0000	6.0000
	Mean	6.8730	7.6200	6.8730	7.6200	0.9961	1.1850	6.0000	6.3860
	Third Qu.	7.0000	8.0000	7.0000	8.0000	1.0610	1.2620	6.0000	7.0000
	Max.	8.0000	9.0000	8.0000	9.0000	1.2600	1.4980	6.0000	7.0000
Coverage	MSE_4	19.4443	23.2261	0.0537	0.0195	0.0182	0.0077	0.7900	0.1975
	MSE_5	19.4443	23.2261	0.0021	0.0008	0.0103	0.0034	0.0000	0.0000

Conclusions and Discussions

- Unlike the conventional ranks which are intrinsically discrete and difficult to make inference, the proposed smooth ranks have nice properties and do no have the same technical problem.
- The simulation of increasing sample sizes 20,000 times to create no tie case also re-enforces our belief that the tie and near tie assumption may be appropriate for the data in Table 1 and also many other data in practice.
- The development has been extended to risk adjustment models
- The technique of replacing a discrete indicate function by a smooth function may have other applications.

References

- Xie, M., Singh, K. and Zhang, C.-H. (2009). Confidence intervals for population ranks in the presence of ties or near ties. *Journal of the American Statistical Association.* Vol. 104, 775-787.
- Goldstein, H., and Spiegelhalter, D.J. (1996), League tables and their limitations: statistical Issues in comparisons of institutional performance (with discussion) *JRSS*-A, 159, 385–443.
- Hall, P. and Miller, H. (2009). Using the bootstrap to quantify the authority of an empirical ranking. *Annals of Statistics*. In press.

RUTGERS
UNIVERSITY