Jackknife variance estimation corrections

Xuelong Wang

2019-11-15

Contents

1 Jackknife variance correction

If we assume the S is a smooth functions of emperical CDF, especially a quadratic functions, then it can be shown the leading terms of $E(\tilde{Var}(S(X_1, \ldots, S_{n-1}))) \geq Var(S(X_1, \ldots, S_{n-1}))$ is a quadratic term in expecation. Therefore we could try to estimate the quadratic term and correct the bias for the jackknife variance estimation.

Define $Q_{ii'} \equiv nS - (n-1)(S_i + S_{i'}) + (n-2)S_{(ii')}$, then the correction will be

$$\hat{Var}^{corr}(S(X_1, \ldots, X_n)) = \hat{Var}(S(X_1, \ldots, X_n)) - \frac{1}{n(n-1)}\sum_{i<i'}(Q_{ii'} - \bar{Q})^2$$

where $\bar{Q} = \sum_{i<i'}(Q_{ii'})/(n(n-1)/2)$

2 Simulation study compare two GCTA and GCTA__rr

GCTA__rr is the `mixed.solve` function from `rrBLUP` r package.
Based on the following simulation results,

1. when $n < p$ case, those two methods' results are very closed to each other.
2. when $n > p$ case, in terms of effect estimation and jackknife variance estimation those two methods's reuslts are similar to each other. But for the variance corrections are quite different. That is the statistics Q of our method has a very large variance which leads to negative correction result.

2.0.1 setup

- Independent
- Normal
- $p = 100$
- $n = \{50, 75, 100, 150, 200\}$
- with interaction terms
- main effect: $Var(X^T\beta) = \{0, 8, 100\}$

2.0.2 Simulation result

2.0.3 $Var(X^T\beta) = \{0\}$

	n	MSE	est_var	est_mean	NA_main	GCTA_main_jack	GCTA_v_jack_1	GCTA_v_jack_2	GCTA_v_corr
1:	50	3.40	1.83	1.32	0	0.59	9.55	9.62	-9.28
2:	75	1.19	0.98	0.56	0	0.46	2.73	2.68	-5.64
3:	100	1.08	0.84	0.57	0	0.44	1.35	1.35	-0.77
4:	150	0.28	0.19	0.32	0	-1.09	0.86	0.67	-64.08
5:	200	0.21	0.12	0.32	0	-1.60	0.78	0.69	-46.14

	n	MSE	est_var	est_mean	NA_main	GCTA_rr_main_jack	GCTA_rr_v_jack_1	GCTA_rr_v_jack_2	GCTA_rr_v_corr
1:	50	3.40	1.83	1.32	0	0.60	9.55	9.47	-3.560
2:	75	1.19	0.98	0.56	0	0.46	2.73	2.68	-5.643
3:	100	1.08	0.84	0.57	0	0.44	1.35	1.35	-0.770
4:	150	0.28	0.19	0.33	0	-0.17	0.62	0.61	-1.204
5:	200	0.21	0.12	0.33	0	0.28	0.61	0.61	-0.041

2.0.4 $Var(X^T\beta) = \{100\}$

	n	MSE	est_var	est_mean	NA_main	GCTA_main_jack	GCTA_v_jack_1	GCTA_v_jack_2	GCTA_v_corr
1:	50	9247	1784	87	0	66	8795	8793	-3687
2:	75	10077	1863	92	0	103	5170	5109	-3122
3:	100	11839	2142	100	0	84	2072	2081	194
4:	150	10953	443	103	0	31	1280	1148	-80475
5:	200	9778	245	98	0	30	725	673	-32124

	n	MSE	est_var	est_mean	NA_main	GCTA_rr_main_jack	GCTA_rr_v_jack_1	GCTA_rr_v_jack_2	GCTA_rr_v_corr
1:	50	9247	1784	87	0	66	8795	8787	-3492
2:	75	10077	1863	92	0	103	5170	5109	-3124
3:	100	11839	2142	100	0	84	2072		
4:	150	11194	414	104	0	103	969		
5:	200	9854	238	98	0	98	616		

3:	2081	194
4:	970	158
5:	616	220

2.0.5 $Var(X^T\beta) = \{8\}$

	n	MSE	est_var	est_mean	NA_main	GCTA_main_jack	GCTA_v_jack_1	GCTA_v_jack_2	GCTA_v_corr
1:	50	90	25.8	8.0	0	8.5	74.1	73.8	-190.67
2:	75	70	13.1	7.5	0	7.5	32.1	31.9	-25.67
3:	100	68	6.3	7.8	0	7.5	13.7	13.8	-0.97
4:	150	70	4.0	8.1	0	8.4	9.2	8.1	-502.59
5:	200	65	2.5	7.9	0	7.6	4.6	2.2	-214.51

	n	MSE	est_var	est_mean	NA_main
1:	50	24.0	24.0	8.0	0
2:	75	13.8	13.8	7.9	0
3:	100	8.6	8.6	8.1	0
4:	150	3.7	3.7	8.0	0
5:	200	2.7	2.7	8.0	0

	n	MSE	est_var	est_mean	NA_main	GCTA_rr_main_jack	GCTA_rr_v_jack_1	GCTA_rr_v_jack_2	GCTA_rr_v_corr
1:	50	90	25.8	8.0	0	8.5	74.1	73.6	-177.35
2:	75	70	13.1	7.5	0	7.5	32.1	31.8	-16.78
3:	100	68	6.3	7.8	0	7.5	13.7	13.8	-0.97
4:	150	70	4.1	8.1	0	8.1	6.9	6.9	1.49
5:	200	65	2.5	7.9	0	7.9	3.9	2.0	1.38

	n	MSE	est_var	est_mean	NA_main
1:	50	23.8	23.9	8.0	0
2:	75	13.7	13.7	7.9	0
3:	100	8.6	8.6	8.1	0
4:	150	3.8	3.8	8.0	0
5:	200	2.7	2.7	8.1	0

2.0.6 correlation test $

	n	MSE	est_var	est_mean	NA_main	cor_main_jack	cor_v_jack_1
1:	50	0.0131	0.0130	0.49	0	0.49	0.0127
2:	75	0.0083	0.0083	0.50	0	0.50	0.0079
3:	100	0.0057	0.0057	0.50	0	0.50	0.0059
4:	150	0.0038	0.0038	0.50	0	0.50	0.0039
5:	200	0.0030	0.0030	0.50	0	0.50	0.0029

```
   cor_v_jack_2 cor_v_corr
1:       0.0128     0.0120
2:       0.0079     0.0076
3:       0.0059     0.0057
4:       0.0039     0.0038
5:       0.0029     0.0029
```