What You See is What You Get

Methodological Component

Collin Nolte

March 8, 2023

Outline

- Mean structure assumptions
- Alternative methods
  - Modified bootstrap
  - Permutation test
- FWER control
- Power

Mean Assumptions

Observed data:

$$y_{it} = f(t|\theta_i) + \epsilon_{it}$$

Homogeneous Means

For all subjects i, j in group $g = 1, \ldots, G$,

$$\theta_i = \theta_j$$

Heterogeneous Means

Subject i in group $g = 1, \ldots, G$ follows

$$\theta_i \sim N(\mu_g, V_g)$$

with no presumption that $\theta_i = \theta_j$

Heterogeneous Bootstrap

Differs from original `bdots` bootstrap in that it samples subjects with replacement

This gives distribution for the bth bootstrap estimate in group g

$$\theta_{bg}^{(het)} \sim N\left(\mu_g, \frac{1}{n_g} V_g + \frac{1}{n_g^2} \sum s_i^2\right)$$

This compares with homogeneous bootstrap which samples without replacement,

$$\theta_{bg}^{(hom)} \sim N\left(\mu_g, \frac{1}{n_g^2} \sum s_i^2\right).$$

Heterogeneous Bootstrap

The change in algorithm (and resulting distribution) are only things that change

Still construct test statistic

$$T_t = \frac{(\overline{p}_{1t} - \overline{p}_{2t})}{\sqrt{s_{1t}^2 + s_{2t}^2}}$$

at each time point, with FWER being controlled with modified Bonferroni adjustment

Permutation Test

Begin by constructing test statistic on observed data,

$$T_t^{(p)} = \frac{|\overline{p}_{1t} - \overline{p}_{2t}|}{\sqrt{s_{1t}^2 + s_{2t}^2}}$$

Then proceeded similarly to a standard permutation test by permuting labels for group membership at each permutation. At each permutation, retain the maximum test statistic, giving a null distribution of P values denoted $\widetilde{T}$

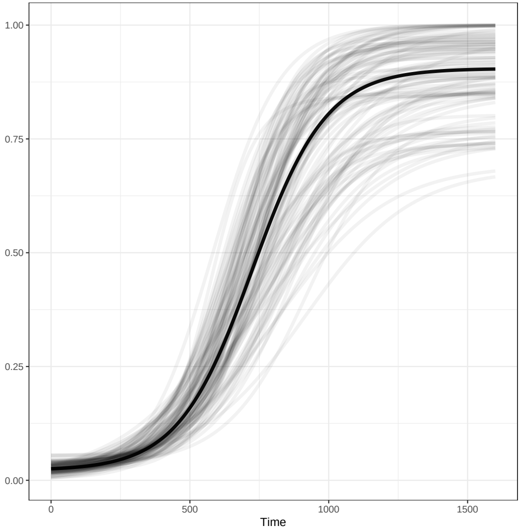
Letting $\widetilde{T}_\alpha$ be the $1 - \alpha$ quantile of $\widetilde{T}$, significant regions will be those in which

$$T_t^{(p)} \geq \widetilde{T}_\alpha$$

FWER Simulation

- Logistic function from 0-1600
- With and without AR(1) error
- For heterogeneous means, sampled from empirical distribution from VWP with NH subjects
- Paired data only for heterogeneous means, subjects differed only in error term
- 25 subjects in each group (50 total), 100 simulations

FWER Simulation



FWER Results (Unpaired)

$\theta_i = \theta_j$	AR Error	AR(1) Specified	Hom Bootstrap	Het Bootstrap	Permutation
No	Yes	Yes	0.06	0.01	0.08
No	Yes	No	0.87	0.08	0.00
No	No	Yes	0.08	0.00	0.06
No	No	No	0.15	0.02	0.01
Yes	Yes	Yes	0.92	0.03	0.05
Yes	Yes	No	0.96	0.02	0.08
Yes	No	Yes	0.99	0.05	0.03
Yes	No	No	1.00	0.05	0.06

Table 1: FWER for empirical parameters (unpaired)

FWER Results (Paired)

$\theta_i = \theta_j$	AR Error	AR(1) Specified	Hom Bootstrap	Het Bootstrap	Permutation
Yes	Yes	Yes	0.49	0.02	0.01
Yes	Yes	No	0.94	0.03	0.02
Yes	No	Yes	0.72	0.02	0.00
Yes	No	No	0.74	0.04	0.00

Table 2: FWER for empirical parameters (paired)

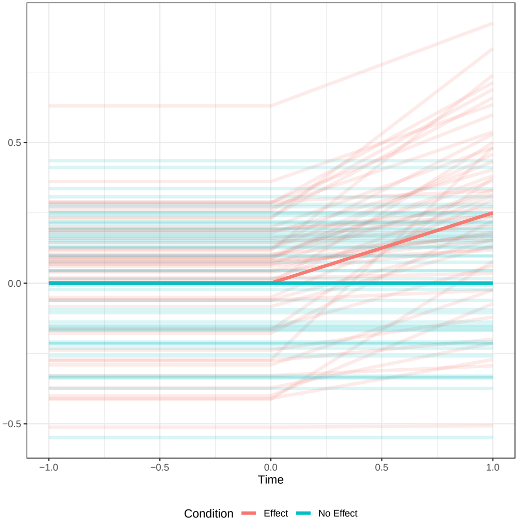
Power Simulation

- Fit using piecewise function

$$y = \begin{cases} b & x < 0 \\ mx + b & x \geq 0 \end{cases}$$

- cases with AR specification ugh
- I have other sims that just aren't presented
- 25 subjects in each group, 1000 simulations
- Metrics include FWER, family wise type II error, and a distribution of where differences first detected (explain)

Power Simulation

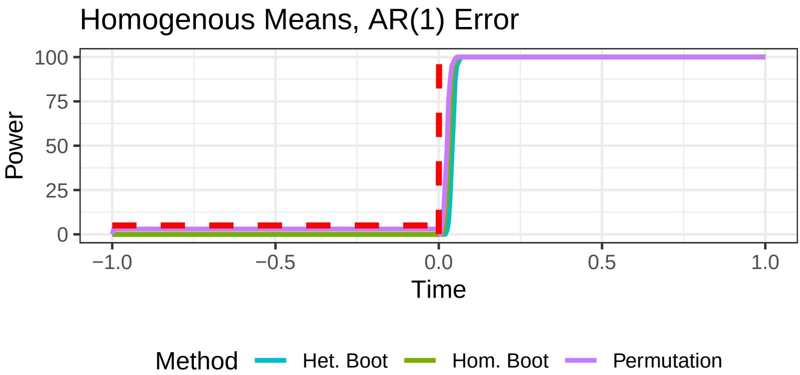


Power Results

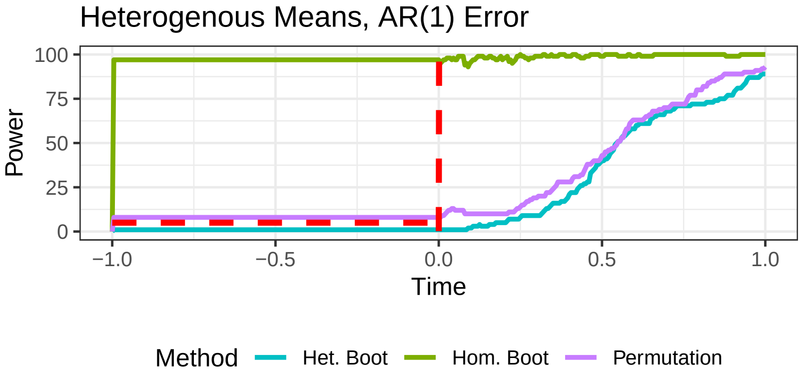
Method	Het.	AR(1)	α	β	1 - α - β	1st Qu.	Median	3rd Qu.
Hom. Boot	No	Yes	0.00	0.00	1.00	0.025	0.030	0.035
Het. Boot	No	Yes	0.00	0.00	1.00	0.035	0.040	0.045
Perm	No	Yes	0.03	0.00	0.97	0.015	0.025	0.025
Hom. Boot	Yes	No	0.96	0.00	0.04	0.005	0.008	0.010
Het. Boot	Yes	No	0.00	0.10	0.90	0.403	0.513	0.690
Perm	Yes	No	0.03	0.05	0.92	0.378	0.515	0.681
Hom. Boot	Yes	Yes	0.97	0.00	0.03	0.008	0.010	0.010
Het. Boot	Yes	Yes	0.01	0.10	0.89	0.420	0.525	0.690
Perm	Yes	Yes	0.08	0.03	0.89	0.360	0.540	0.705

Table 3: Power for methods

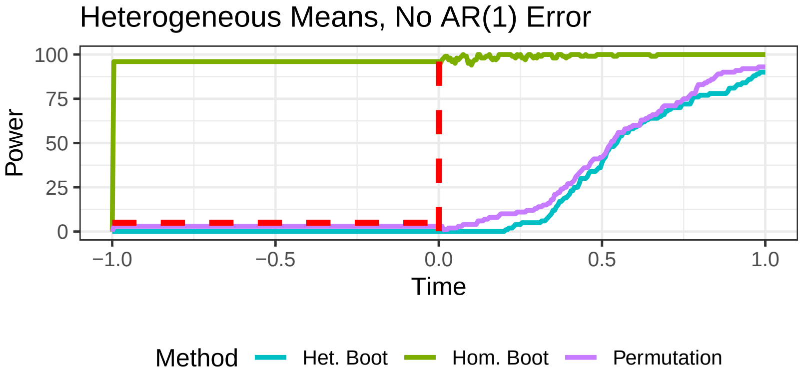
Power (Homogeneous Means)



Power (Heterogeneous Means, AR(1))



Power (Heterogeneous Means, No AR(1))



Conclusions

- New methods adequately control FWER under myriad of situations
- Also have comparable power under homogeneous means assumptions
- In general, permutation performs closest to nominal alpha while having slightly improved power (compared to heterogeneous bootstrap)
- Sampling without replacement will be removed from `bdots` package as there is no situation in which is drastically outperforms any of the either two