Simulation Results: Size

Contents

We review results from our simulations for size.

0.1 Prepare Working Environment

We begin by loading the relevant R packages:

```
library(kableExtra)
options(knitr.table.format = 'html')
```

We then define directories for relevant scripts and data:

```
dir_main <- dirname(dirname(rstudioapi::getActiveDocumentContext()$path))
dir_src <- file.path(dir_main, 'source_scripts')
source(file.path(dir_src, 'define_directories.R'))
```

All scenarios in this simulation setting share the following parameter values:

```
size_or_power <- "size"
alpha <- 0.05
num_permutations <- 1000
num_replicates <- 5000
values_for_error_distribution <- c('normal', 'cauchy')
values_for_error_corr_strength <- c(0, 0.5)
```

0.2 Continuous Features

We generate a table of empirical size results for each test method from the simulated data for size scenarios where $\boldsymbol{X}$ was modeled as a continuous random vector:

```
x_type <- 'cts'
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T, caption = caption) %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  add_header_above(outer_header_row) %>%
  pack_rows(index = row_group_index)
```

The sample sizes considered in these scenarios are small ($n = 50$ and $n = 100$) in order to investigate whether AMKAT can reasonably control the type I error rate under small sample sizes when using the PhiMr filter (especially when using multiple test statistics to estimate a P-value).

Table 1: Empirical size (percentage) across 5000 simulated data replicates using continuous features. All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components				Correlated error components			
	$p = 500$		$p = 2000$		$p = 500$		$p = 2000$	
	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
Multivariate normal errors								
AMKAT-PhiMr-Q1	5.06	5.04	4.86	4.88	4.94	4.24	5.34	5.06
AMKAT-PhiMr-Q2	4.08	4.20	3.82	4.48	4.62	3.84	4.60	4.78
AMKAT-PhiMr-Q4	3.70	3.96	3.28	4.20	4.42	3.58	4.22	4.82
AMKAT-PhiMr-Q8	3.52	3.62	3.02	3.92	4.00	3.56	3.82	4.70
AMKAT-PhiMr-Q16	3.44	3.44	2.88	3.80	4.18	3.62	3.68	4.52
AMKAT-PhiMr-Q32	3.38	3.34	2.82	3.52	4.20	3.52	3.54	4.40
AMKAT-PhiMr-Q64	3.40	3.34	2.84	3.56	4.14	3.48	3.48	4.42
AMKAT-PhiMr-Q128	3.48	3.44	2.82	3.68	3.98	3.38	3.44	4.48
OMGA	0.50	0.28	0.24	0.04	0.40	0.14	0.24	0.12
DKAT	5.00	5.46	4.60	5.22	4.66	4.68	4.92	5.36
DKAT-LPK	4.70	5.26	4.56	5.20	5.02	4.32	5.18	5.44
Multivariate Cauchy errors								
AMKAT-PhiMr-Q1	4.96	4.96	4.70	5.30	4.50	5.02	5.24	5.18
AMKAT-PhiMr-Q2	4.12	4.32	4.12	4.00	4.08	4.20	4.36	4.62
AMKAT-PhiMr-Q4	3.46	3.94	3.24	3.50	3.70	3.78	4.12	4.18
AMKAT-PhiMr-Q8	3.32	3.64	3.18	3.22	3.52	3.84	4.04	4.16
AMKAT-PhiMr-Q16	3.32	3.50	3.24	3.20	3.34	3.76	3.88	4.00
AMKAT-PhiMr-Q32	3.20	3.24	3.14	3.14	3.30	3.66	3.58	4.00
AMKAT-PhiMr-Q64	3.14	3.28	3.06	3.10	3.36	3.58	3.70	3.96
AMKAT-PhiMr-Q128	3.20	3.34	2.92	3.14	3.28	3.54	3.68	3.92
OMGA	0.10	0.02	0.00	0.00	0.00	0.00	0.00	0.00
DKAT	5.38	4.86	5.24	4.68	4.88	5.26	4.60	5.56
DKAT-LPK	5.28	4.82	5.00	5.10	4.50	5.42	5.26	5.36

Multivariate normal errors

When the random error vector ϵ was simulated as multivariate normal, we see that AMKAT with the PhiMr filter had a well-controlled type I error rate when using a single test statistic ($Q = 1$) for P-value estimation. By contrast, OMGA, whose P-value is based on asymptotic results, was highly conservative under the samples sizes considered. DKAT, which approximates the permutation null distribution of its test statistic by a Pearson type III distribution whose parameters are estimated via moment matching, had well-controlled type I error rate.

When using the average of multiple test statistics ($Q \geq 2$) under the PhiMr filter to estimate a P-value, AMKAT's type I error rate gradually became (slightly) more conservative as Q increased; this trend stabilized at or before $Q = 64$, beyond which the type I error rate held constant at a value typically of around 3.5%.

Interestingly, when using multiple test statistics ($Q \geq 2$), AMKAT's type I error rate was typically closer to nominal when error components were correlated rather than uncorrelated. For $n = 100$ and $p = 2000$, all AMKAT variations maintained type I error rates above 4.4% when the error components were correlated; a similar situation was also observed for $n = 50$ and $p = 500$, where no rates were below 3.98% when error components were correlated. Consistent with this pattern, the one setting in which AMKAT had type I error rates below its typical levels was $n = 50$ and $p = 2000$ with uncorrelated error components, where the type I error rate stabilized at around 2.8% for $Q \geq 32$.

Multivariate Cauchy errors

When the random error vector ϵ was simulated as multivariate Cauchy, AMKAT had type I error control similar to the case when ϵ was simulated as multivariate normal. The multivariate Cauchy distribution (i.e., multivariate Student's t with 1 degree of freedom) is a pathological distribution whose mean vector and covariance matrix are undefined, and was used to assess the robustness of AMKAT's type I error control against deviations from our model's assumption of the random error vector having a mean vector and covariance matrix with finite entries; this assumption guarantees certain asymptotic properties of some estimators used in our test, such as the ratio consistency of the estimator $\hat{\sigma}^2_{\tau_j}$ for the asymptotic variance of the statistic τ_j, corresponding to the jth response component, that appears in the test statistic used by AMKAT.

We observe from our results that a multivariate Cauchy error distribution had little to no impact on AMKAT's type I error control. While DKAT's type I error control was also robust under multivariate Cauchy errors, OMGA's type I error rate was highly conservative relative to scenarios with multivariate normal errors.

0.3 Discrete Features

We generate a table of empirical size results for each test method from the simulated data for size scenarios where $\boldsymbol{X}$ was modeled as a discrete random vector representing additive-encoded SNP-set data:

```
x_type <- 'snp'
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T, caption = caption) %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

Multivariate normal errors

From the above table, we observe that AMKAT maintained reasonably well-controlled type I error rate in all settings and across all values for the number Q of test statistics used under PhiMr for P-value estimation. DKAT had well-controlled type I error rate, while OMGA was conservative at the sample sizes considered, though not as much so as was observed for the case with a set of continuous features.

Table 2: Empirical size (percentage) across 5000 simulated data replicates using a simulated SNP-set ($p = 567$). All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components			Correlated error components		
	$n = 30$	$n = 50$	$n = 70$	$n = 30$	$n = 50$	$n = 70$
Multivariate normal errors						
AMKAT-PhiMr-Q1	5.12	5.20	5.28	4.90	4.88	4.88
AMKAT-PhiMr-Q2	4.58	4.70	4.74	4.66	4.76	4.74
AMKAT-PhiMr-Q4	4.16	4.20	4.48	4.44	4.64	4.46
AMKAT-PhiMr-Q8	4.20	4.02	4.36	4.52	4.56	4.42
AMKAT-PhiMr-Q16	4.00	3.94	4.20	4.44	4.42	4.34
AMKAT-PhiMr-Q32	3.96	3.88	4.28	4.34	4.42	4.36
AMKAT-PhiMr-Q64	4.00	3.98	4.24	4.38	4.42	4.40
AMKAT-PhiMr-Q128	3.92	4.00	4.14	4.32	4.46	4.34
OMGA	2.66	1.70	1.24	2.26	1.74	1.66
DKAT	4.82	4.96	4.82	4.62	4.86	5.74
DKAT-LPK	4.90	4.68	4.92	4.88	5.04	5.18
Multivariate Cauchy errors						
AMKAT-PhiMr-Q1	5.12	4.52	5.22	5.14	5.44	4.58
AMKAT-PhiMr-Q2	4.58	4.36	4.76	4.80	4.92	4.34
AMKAT-PhiMr-Q4	4.32	4.06	4.22	4.34	4.78	4.18
AMKAT-PhiMr-Q8	4.06	3.74	4.16	4.14	4.68	4.00
AMKAT-PhiMr-Q16	4.08	3.80	4.18	4.12	4.64	3.92
AMKAT-PhiMr-Q32	3.96	3.78	4.30	4.20	4.54	3.98
AMKAT-PhiMr-Q64	4.02	3.80	4.32	4.28	4.62	3.94
AMKAT-PhiMr-Q128	3.90	3.78	4.26	4.14	4.52	3.86
OMGA	0.14	0.06	0.00	0.18	0.08	0.00
DKAT	5.36	5.06	4.88	5.76	5.58	4.68
DKAT-LPK	5.48	4.74	5.04	5.40	5.82	4.58

Multivariate Cauchy errors

AMKAT and DKAT had robust type I error control under a multivariate Cauchy error distribution, while OMGA was extremely conservative.