

Kaggle Rain Training Set - Variable Summary

Scott Mark

October 7, 2015

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:data.table':
##
##   between, last
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

## loading from CSV
Read 0.0% of 13765201 rows
Read 6.7% of 13765201 rows
Read 13.4% of 13765201 rows
Read 20.1% of 13765201 rows
Read 26.7% of 13765201 rows
Read 33.5% of 13765201 rows
Read 40.4% of 13765201 rows
Read 47.4% of 13765201 rows
Read 54.1% of 13765201 rows
Read 61.0% of 13765201 rows
Read 67.9% of 13765201 rows
Read 74.4% of 13765201 rows
Read 80.8% of 13765201 rows
Read 87.3% of 13765201 rows
Read 93.7% of 13765201 rows
Read 13765201 rows and 24 (of 24) columns from 1.163 GB file in 00:00:19
```

Summary of the Kaggle Rain *training* data set, which has 13765201 rows and 24 columns.

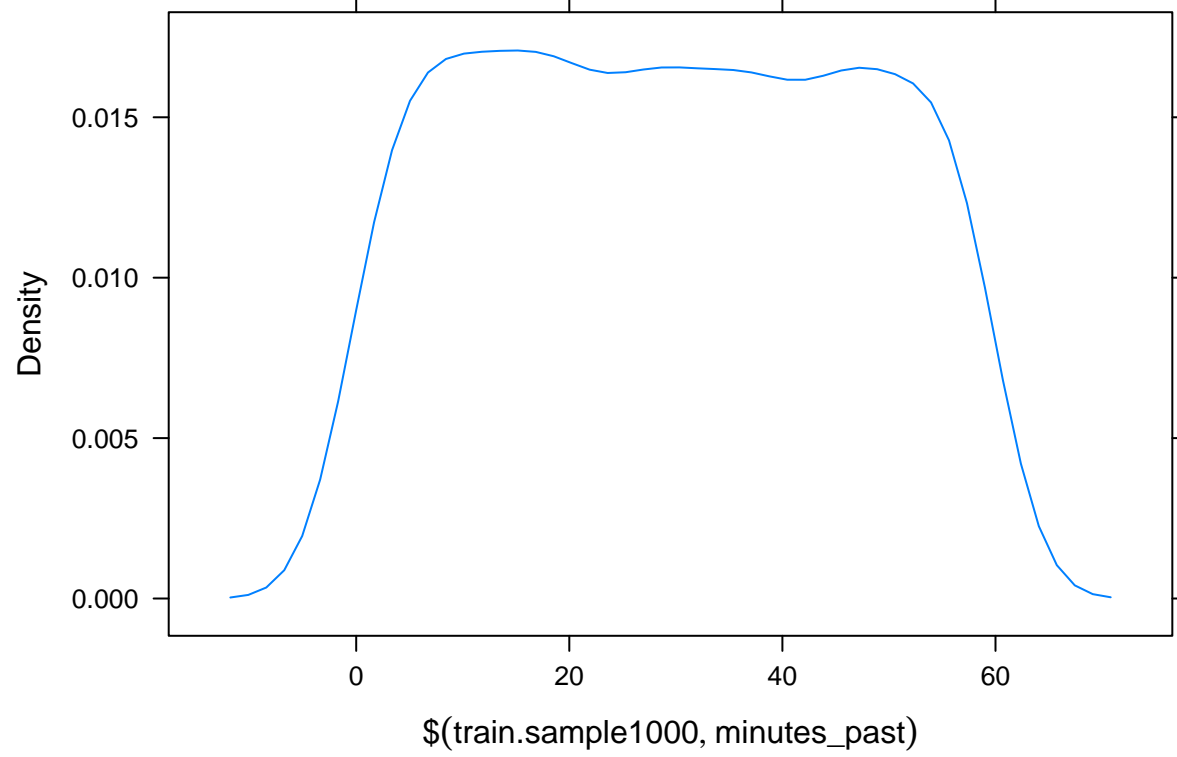
There are 10996113 rows with at least one NA, which is 79.8834176% of the data set.

Variable Summary

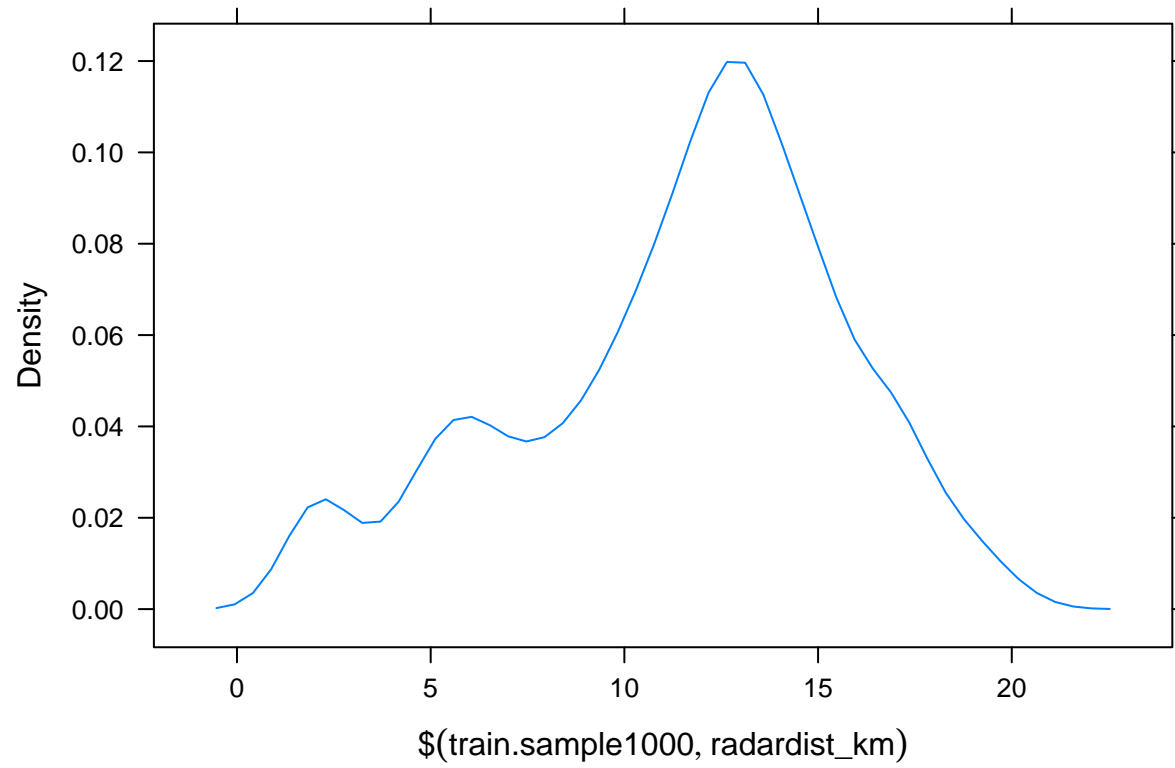
##	nmiss	min	mean	max	P01	P05	P25
## Id	0	1.00	48.50	97.00	2.000	5.000	23.000
## minutes_past	0	0.00	29.54	59.00	0.000	3.000	14.000
## radardist_km	0	2.00	11.49	20.00	2.000	3.000	9.000
## Ref	521	-2.50	23.03	56.50	-0.500	5.500	17.000
## Ref_5x5_10th	591	-7.50	19.81	56.50	-3.960	2.500	14.000
## Ref_5x5_50th	516	-3.50	22.60	58.50	0.415	6.000	17.000
## Ref_5x5_90th	440	-1.50	26.61	63.00	4.795	8.975	20.000
## RefComposite	486	-1.00	25.13	59.00	2.065	8.500	19.000
## RefComposite_5x5_10th	542	-4.50	22.31	56.50	-1.000	6.350	16.500
## RefComposite_5x5_50th	481	-1.50	24.84	59.00	3.000	8.000	18.500
## RefComposite_5x5_90th	420	3.50	28.45	63.00	5.895	11.000	21.500
## RhoHV	660	0.48	0.99	1.05	0.756	0.885	0.982
## RhoHV_5x5_10th	698	0.47	0.94	1.03	0.565	0.815	0.928
## RhoHV_5x5_50th	660	0.71	0.99	1.05	0.853	0.945	0.988
## RhoHV_5x5_90th	612	0.21	1.02	1.05	0.967	0.985	0.998
## Zdr	660	-2.50	0.58	7.94	-2.014	-1.125	0.000
## Zdr_5x5_10th	698	-4.62	-0.58	1.50	-2.999	-2.125	-1.109
## Zdr_5x5_50th	660	-2.12	0.41	7.94	-0.951	-0.438	0.000
## Zdr_5x5_90th	612	-2.00	1.86	7.94	0.288	0.562	1.062
## Kdp	711	-23.87	0.06	14.62	-13.384	-5.254	-1.410
## Kdp_5x5_10th	748	-17.91	-3.41	1.06	-12.639	-8.780	-4.230
## Kdp_5x5_50th	714	-22.46	-0.18	3.51	-4.936	-2.460	-0.710
## Kdp_5x5_90th	659	-13.23	3.76	25.05	-2.758	0.000	1.760
## Expected	0	0.01	152.37	2976.63	0.010	0.254	0.350
##	P50	P75	P95	P99	variance	stdev	
## Id	48.000	73.000	93.000	96.000	808.80	28.44	
## minutes_past	29.000	45.000	57.000	59.000	303.37	17.42	
## radardist_km	12.000	14.000	17.000	19.000	17.37	4.17	
## Ref	23.000	30.000	37.000	42.500	91.13	9.55	
## Ref_5x5_10th	20.000	26.000	33.300	40.460	86.50	9.30	
## Ref_5x5_50th	23.000	29.500	35.925	43.500	89.09	9.44	
## Ref_5x5_90th	27.500	34.000	42.500	48.205	98.26	9.91	
## RefComposite	24.500	32.000	40.500	46.435	97.14	9.86	

## RefComposite_5x5_10th	22.000	28.500	37.500	43.930	89.57	9.46
## RefComposite_5x5_50th	24.500	31.500	40.000	47.000	93.03	9.65
## RefComposite_5x5_90th	28.500	36.000	45.000	50.605	104.20	10.21
## RhoHV	0.995	1.008	1.052	1.052	0.00	0.06
## RhoHV_5x5_10th	0.972	0.988	0.998	1.008	0.01	0.08
## RhoHV_5x5_50th	0.995	1.002	1.042	1.052	0.00	0.03
## RhoHV_5x5_90th	1.012	1.052	1.052	1.052	0.00	0.05
## Zdr	0.500	1.062	2.194	4.412	1.31	1.14
## Zdr_5x5_10th	-0.500	0.062	0.625	0.999	0.79	0.89
## Zdr_5x5_50th	0.312	0.750	1.375	2.496	0.54	0.73
## Zdr_5x5_90th	1.562	2.250	4.188	7.720	1.70	1.30
## Kdp	0.350	1.760	5.122	9.193	14.49	3.81
## Kdp_5x5_10th	-2.470	-1.635	-0.710	-0.350	7.49	2.74
## Kdp_5x5_50th	0.000	0.700	1.760	2.811	3.41	1.85
## Kdp_5x5_90th	2.820	5.280	9.870	13.390	11.69	3.42
## Expected	1.270	4.128	1308.101	2500.123	221921.82	471.09

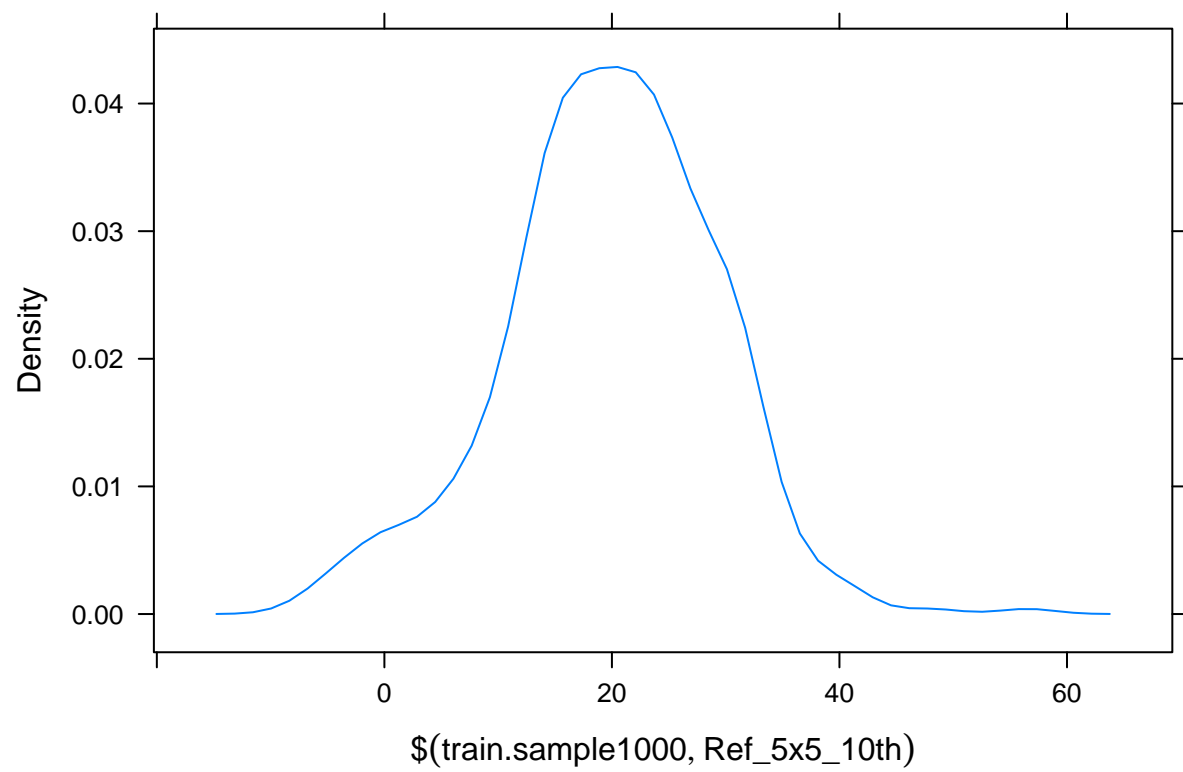
minutes_past



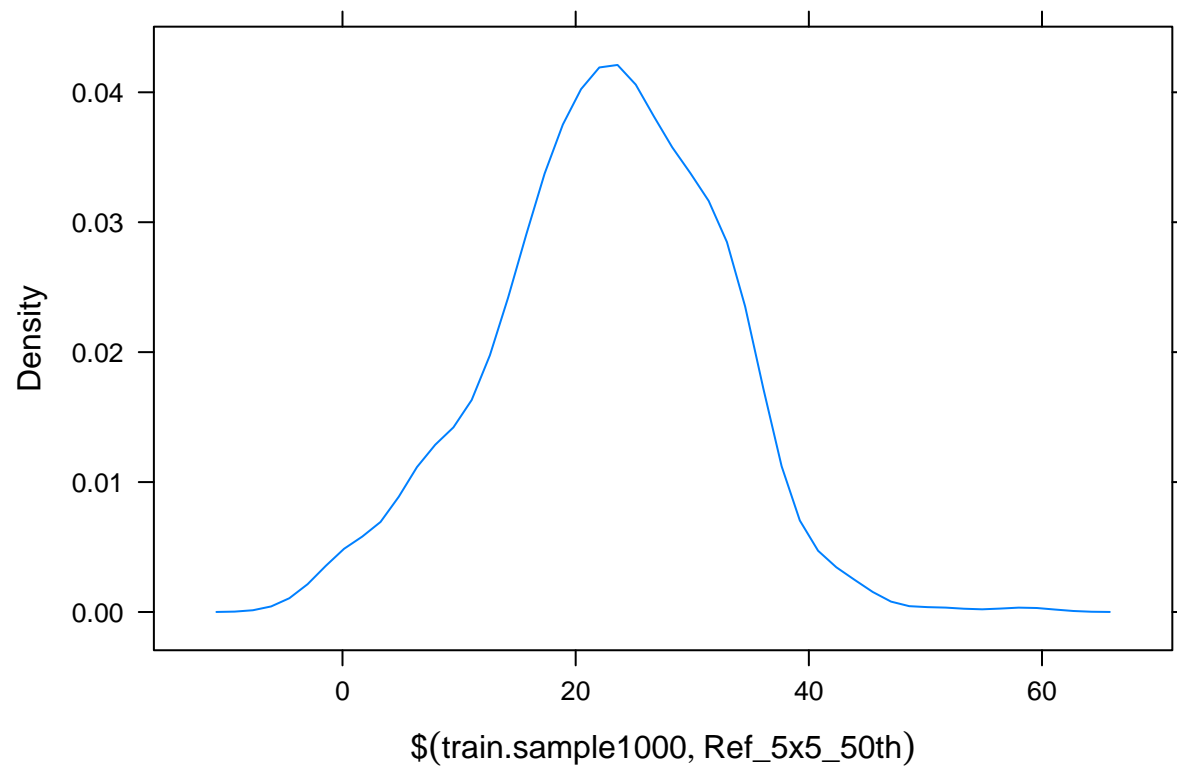
radardist_km



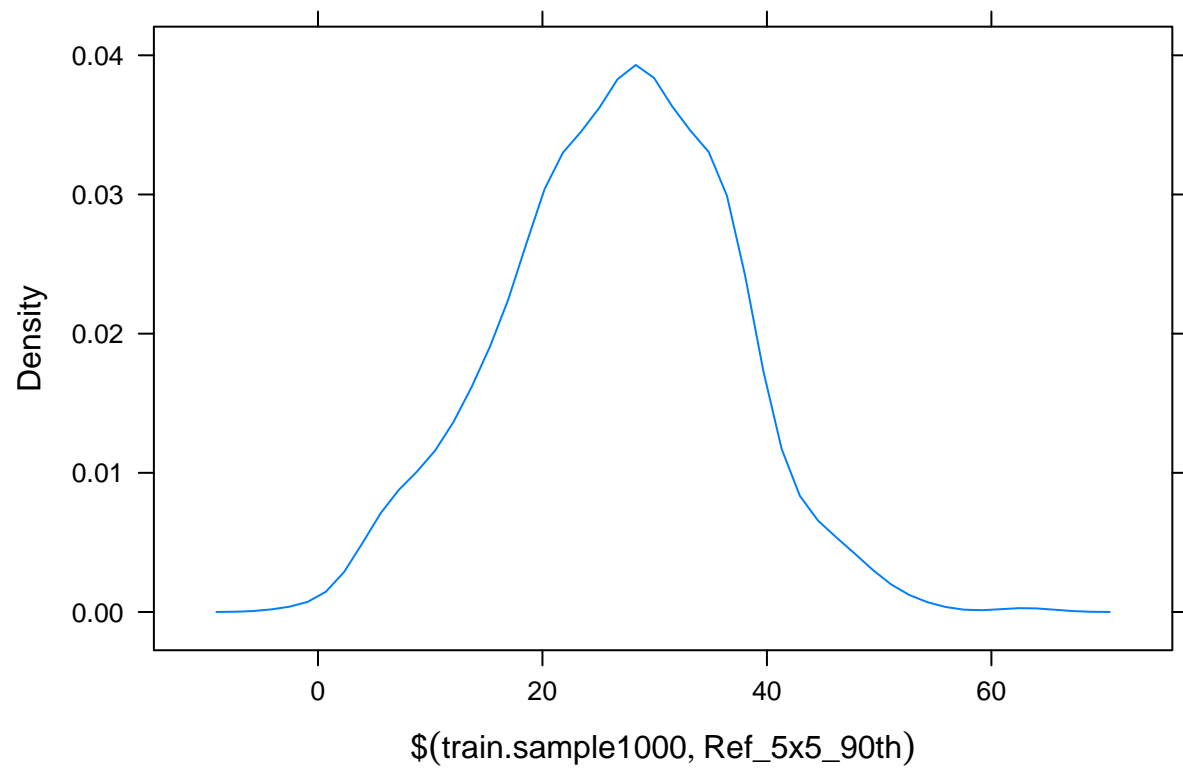
Ref_5x5_10th



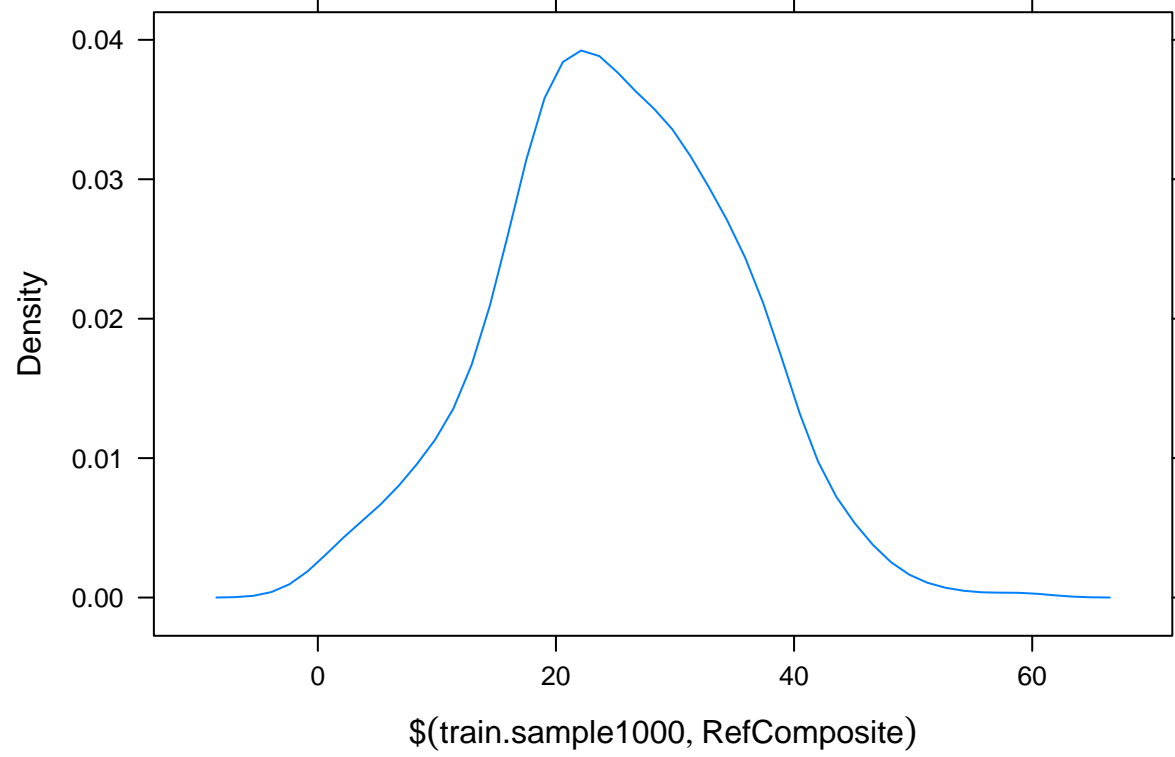
Ref_5x5_50th



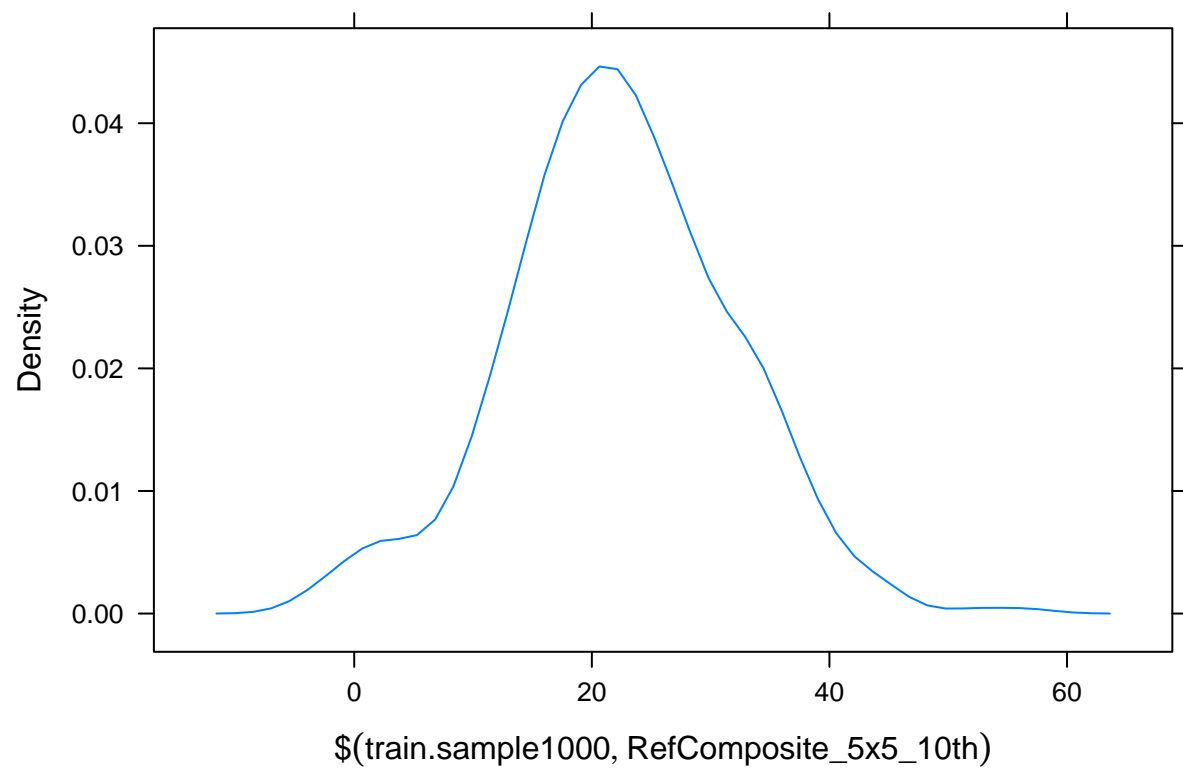
Ref_5x5_90th



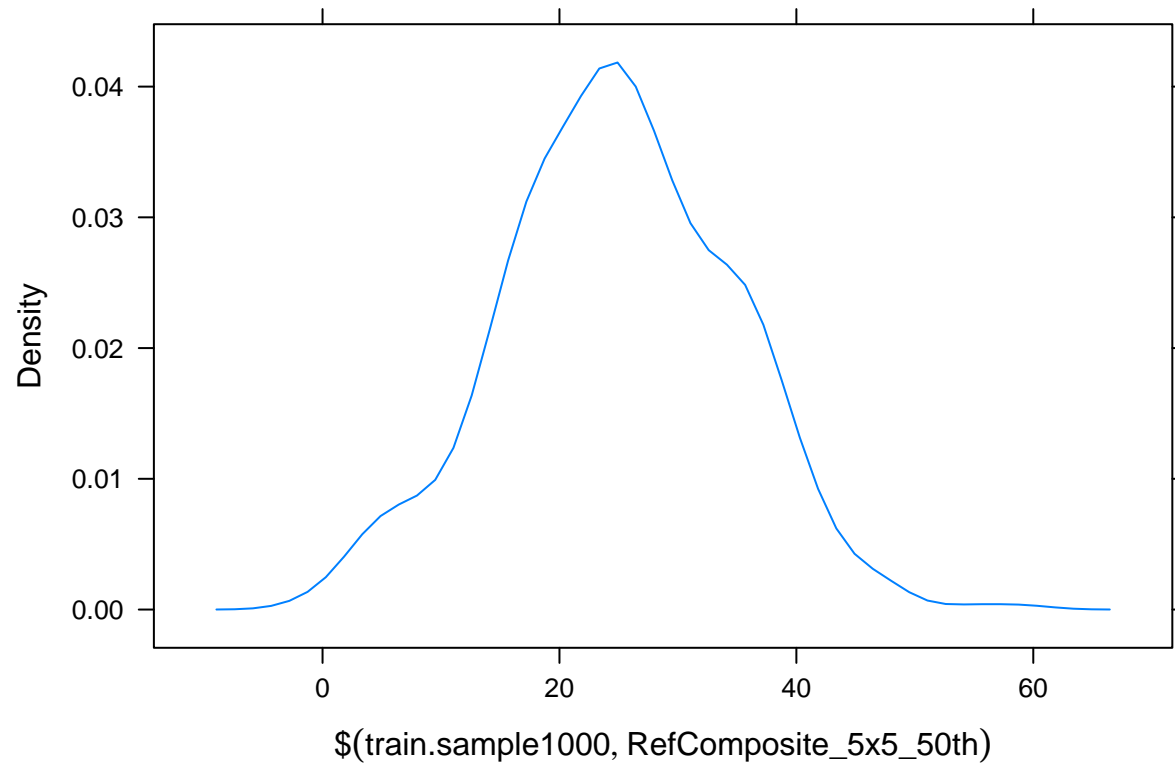
RefComposite



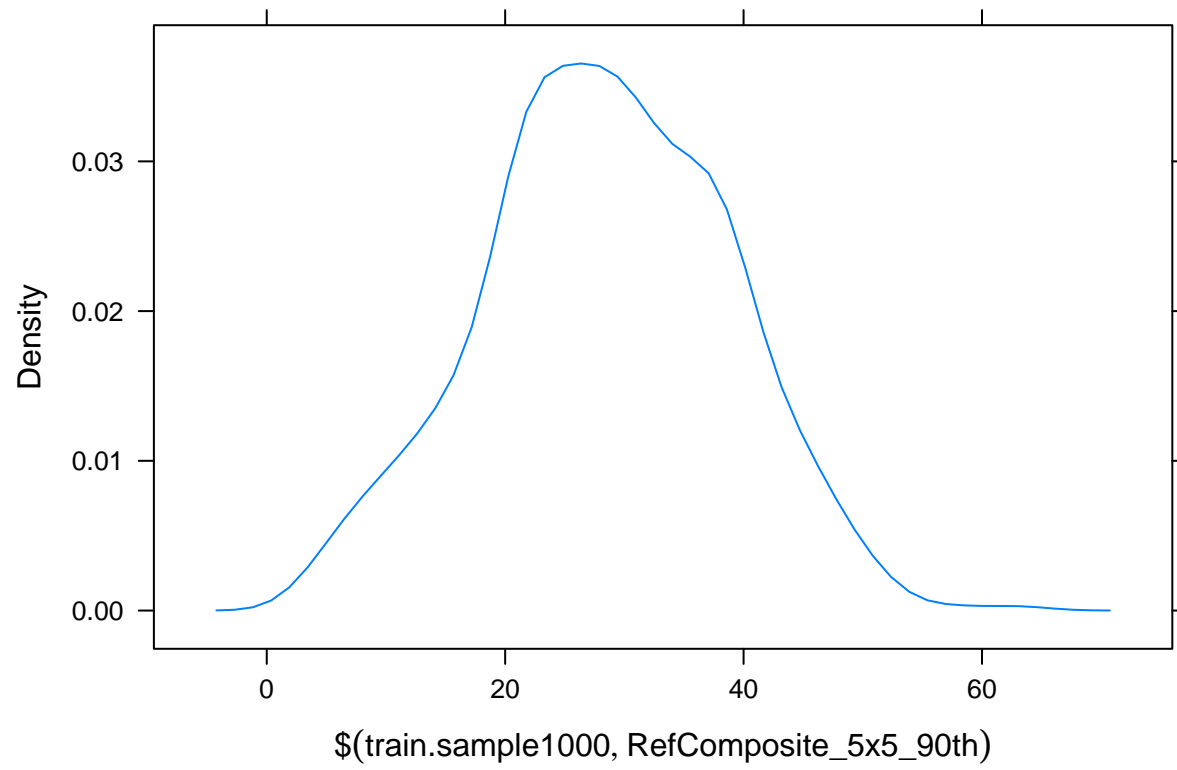
RefComposite_5x5_10th



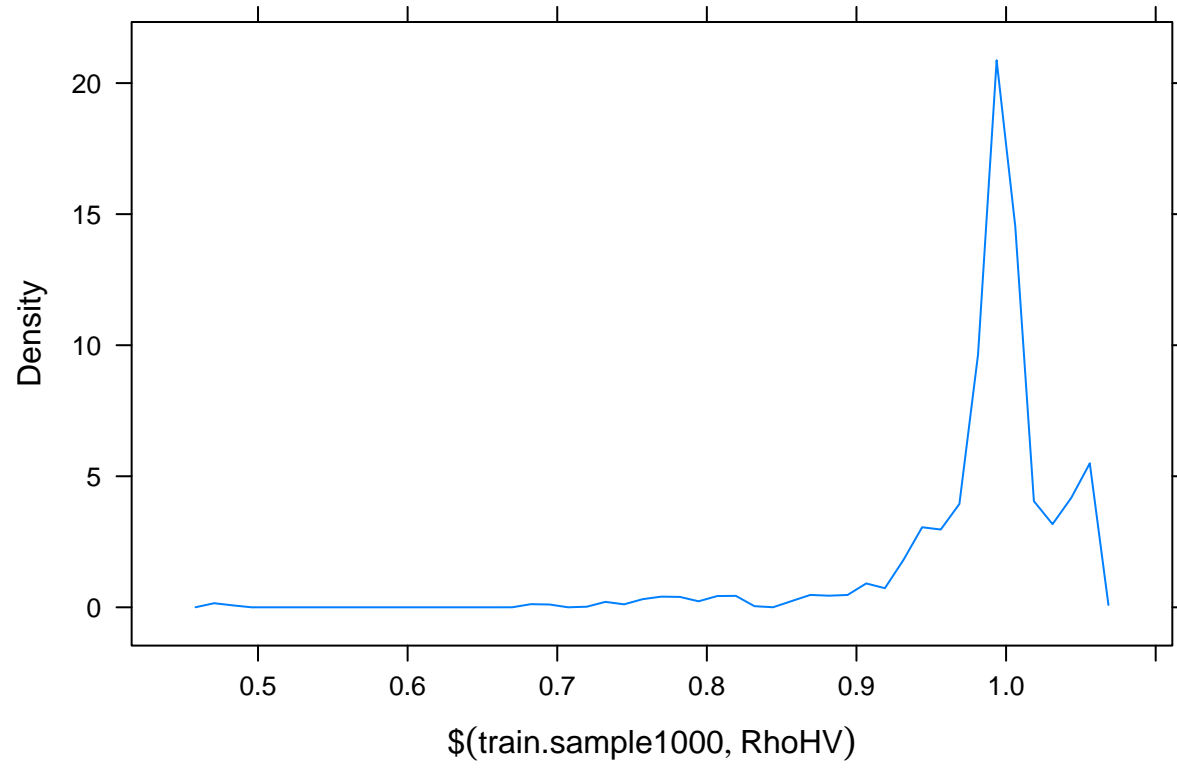
RefComposite_5x5_50th



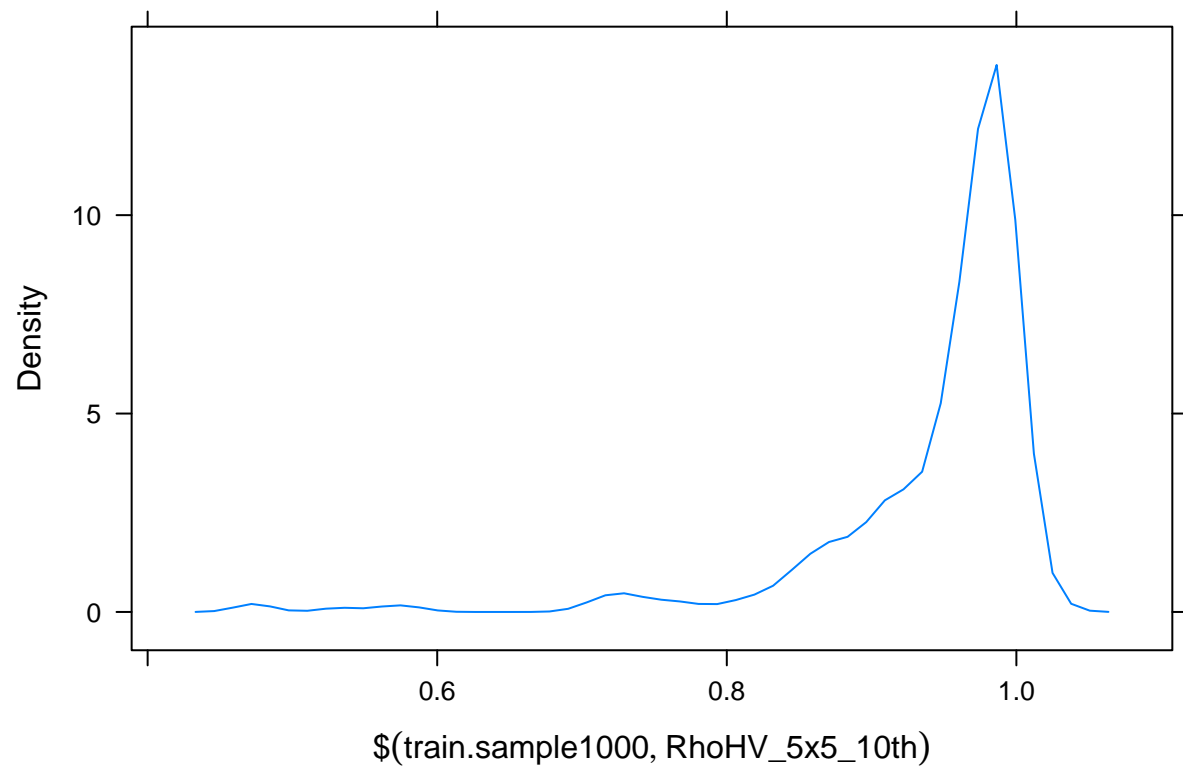
RefComposite_5x5_90th



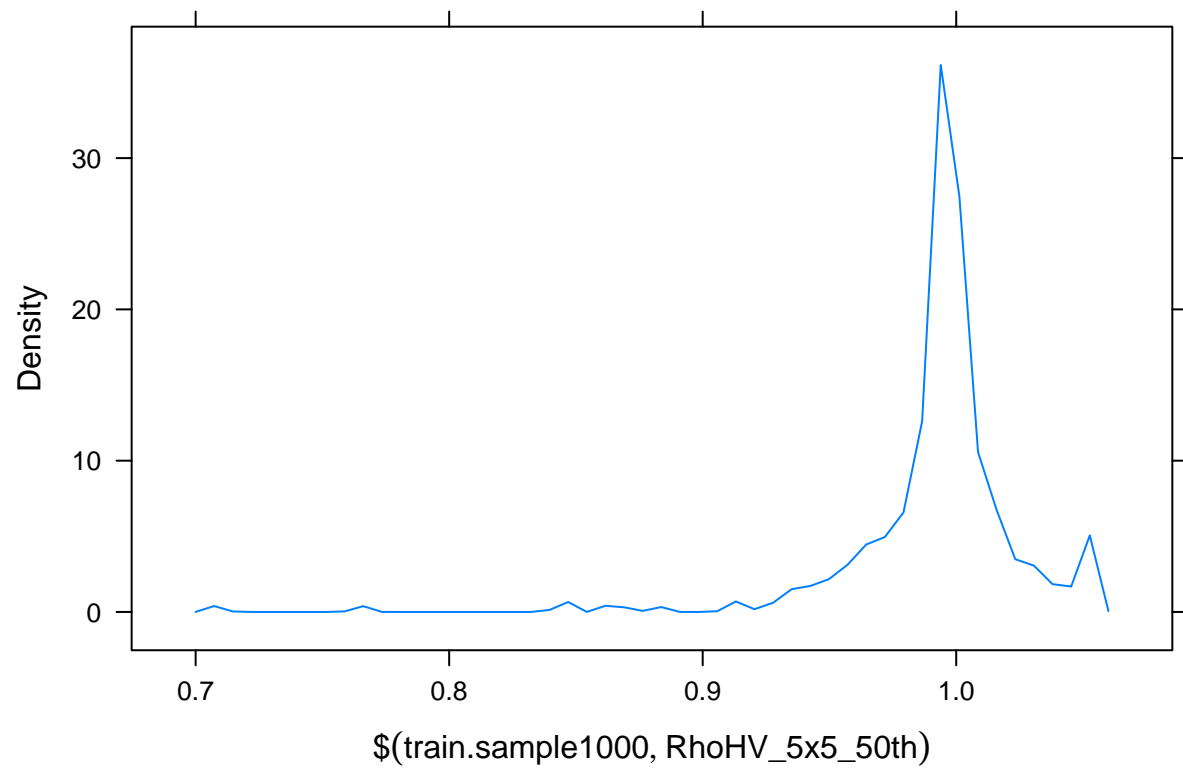
RhoHV



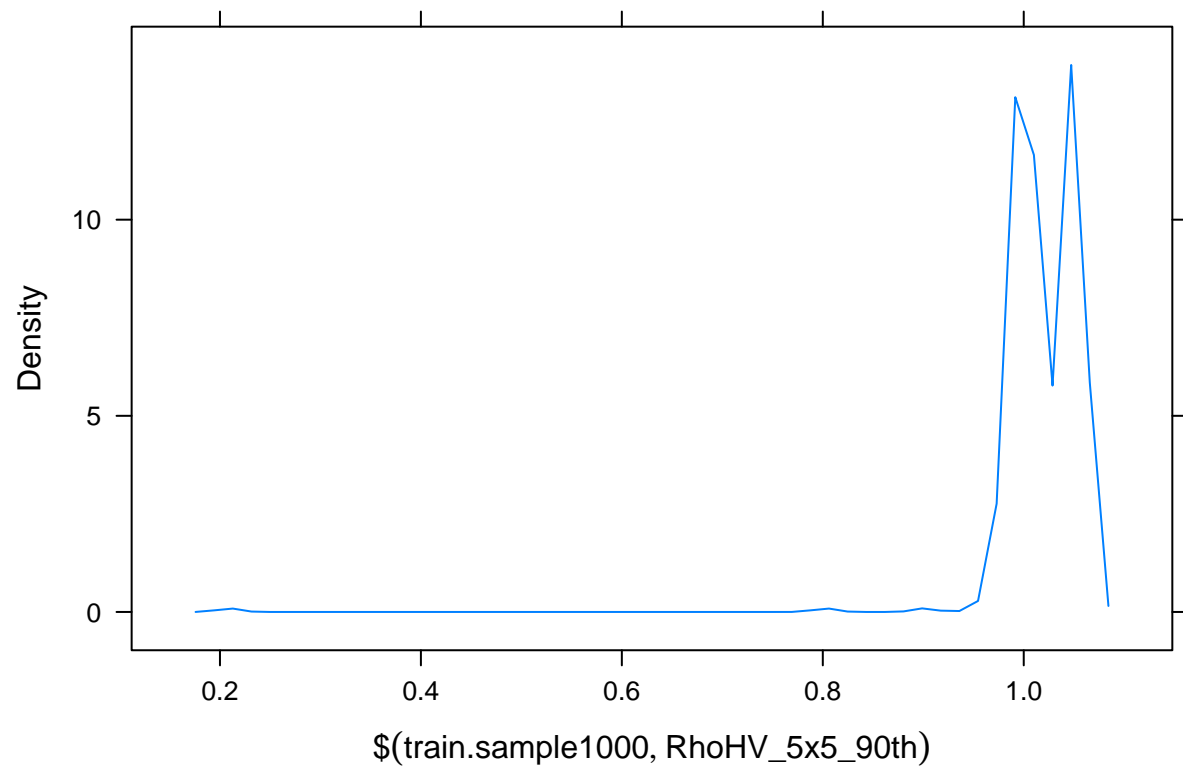
RhoHV_5x5_10th



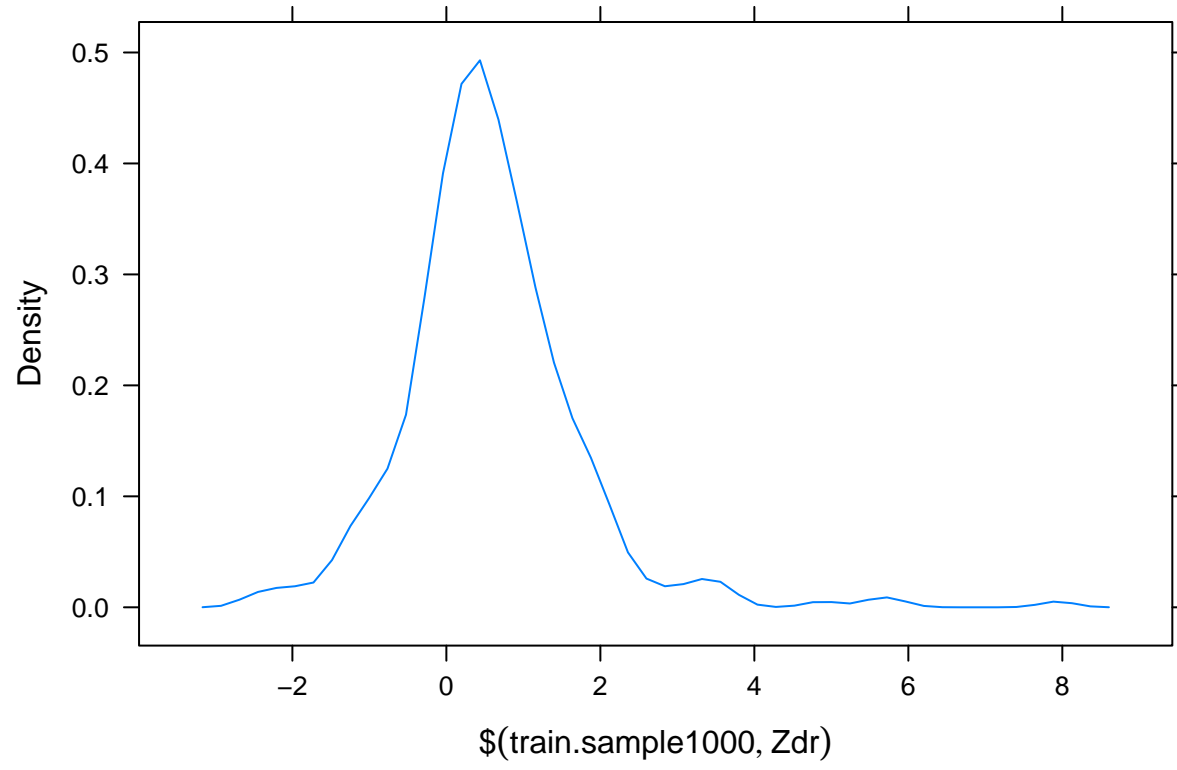
RhoHV_5x5_50th



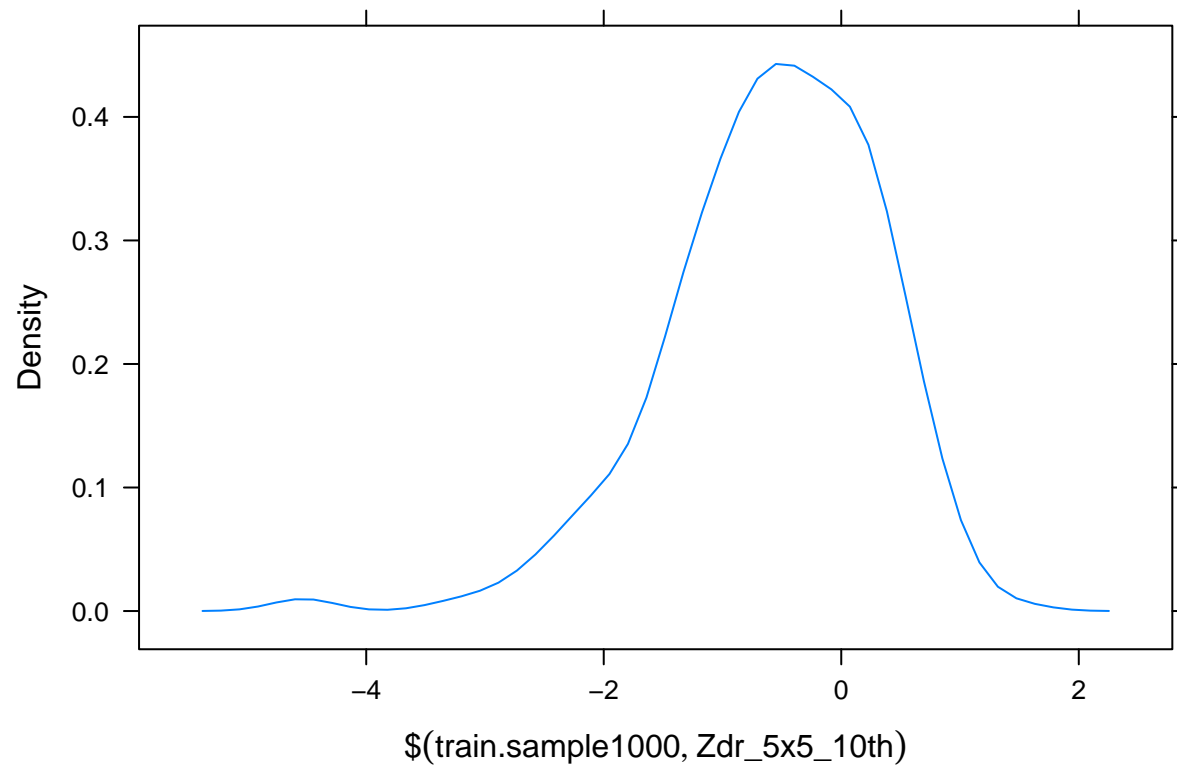
RhoHV_5x5_90th



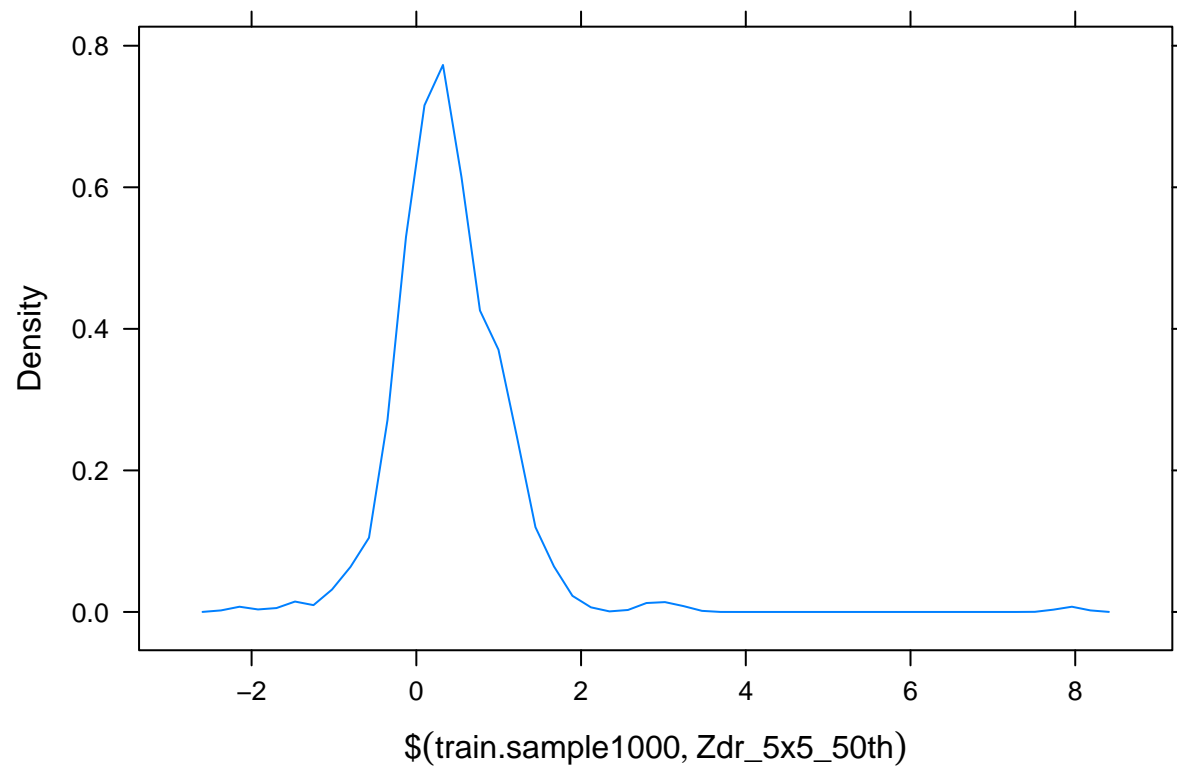
Zdr



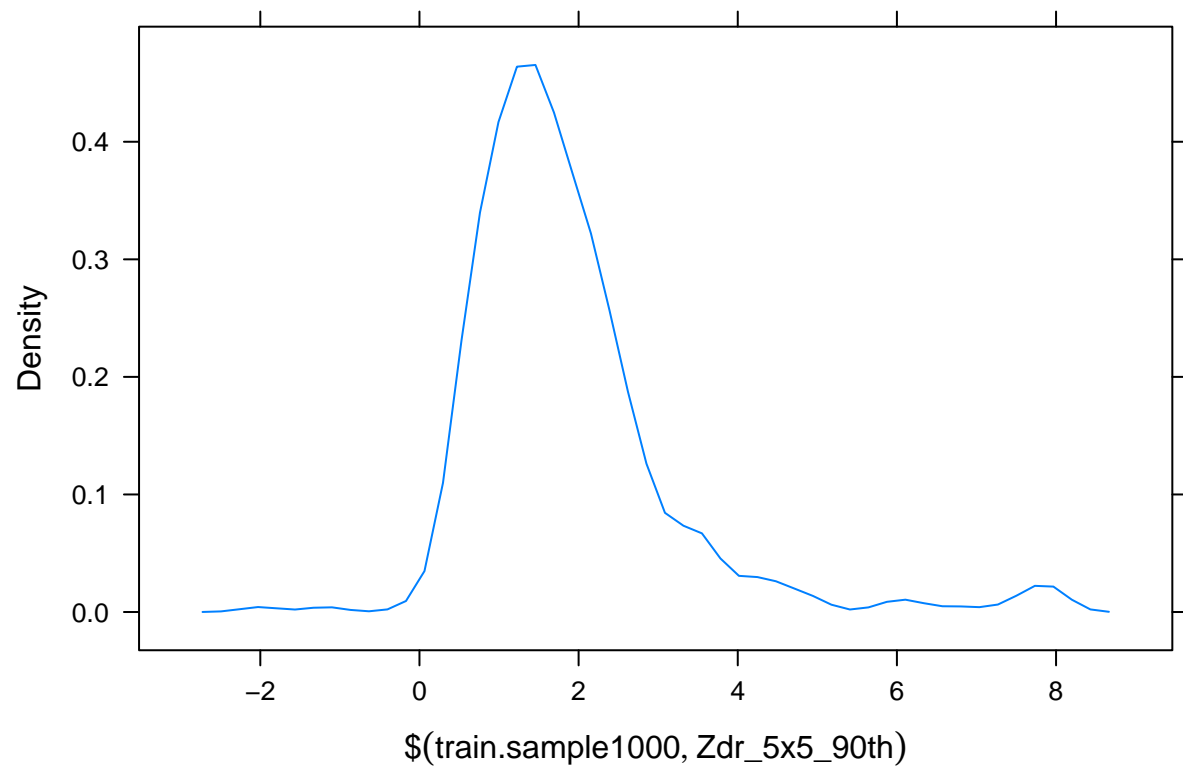
Zdr_5x5_10th



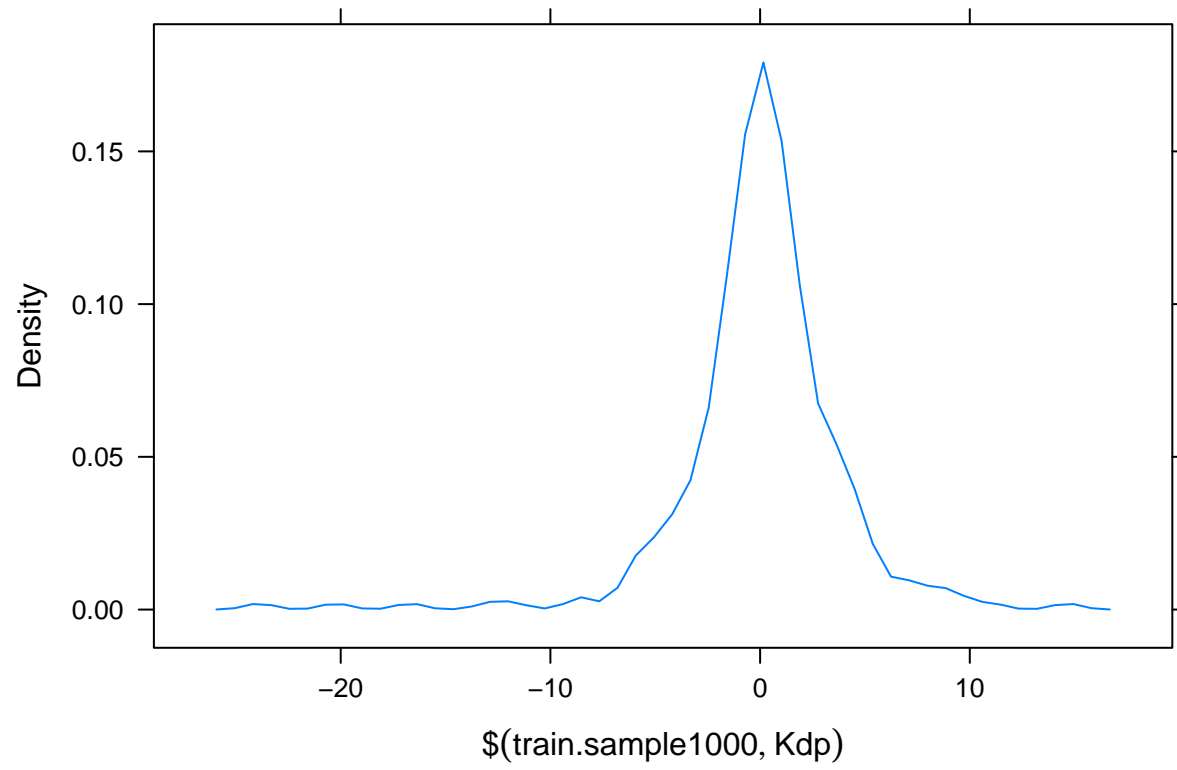
Zdr_5x5_50th



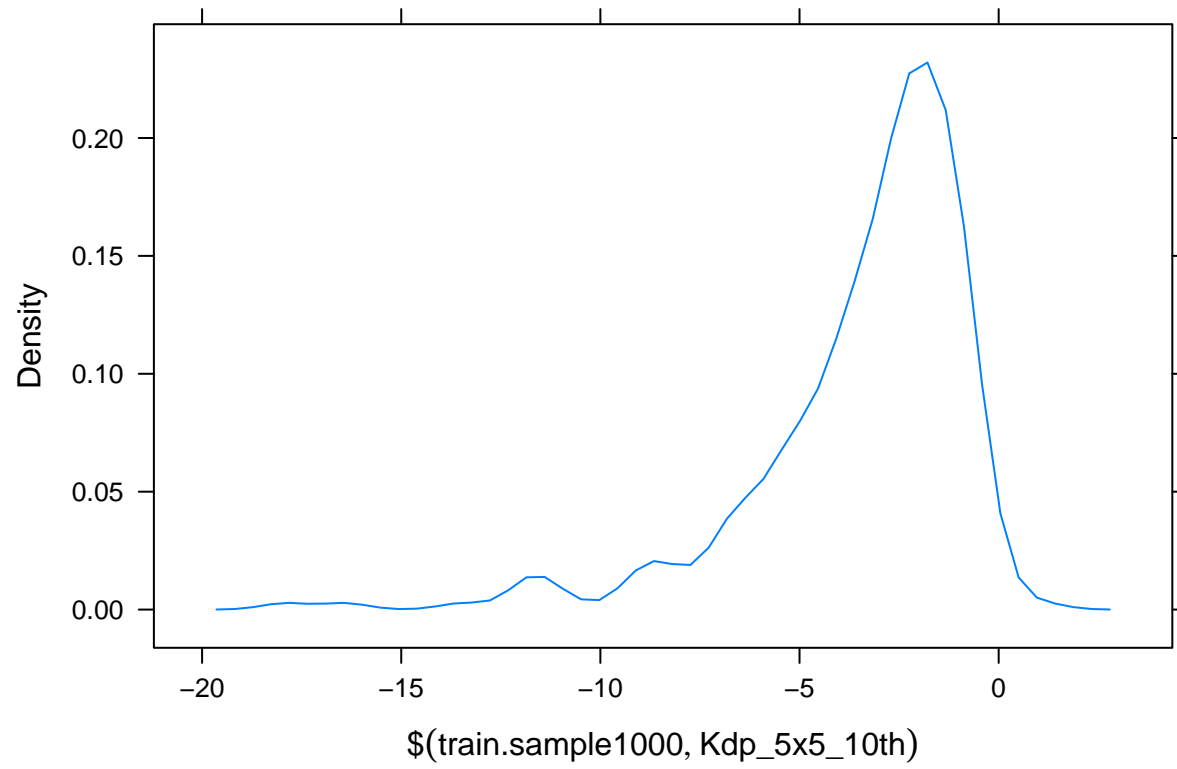
Zdr_5x5_90th



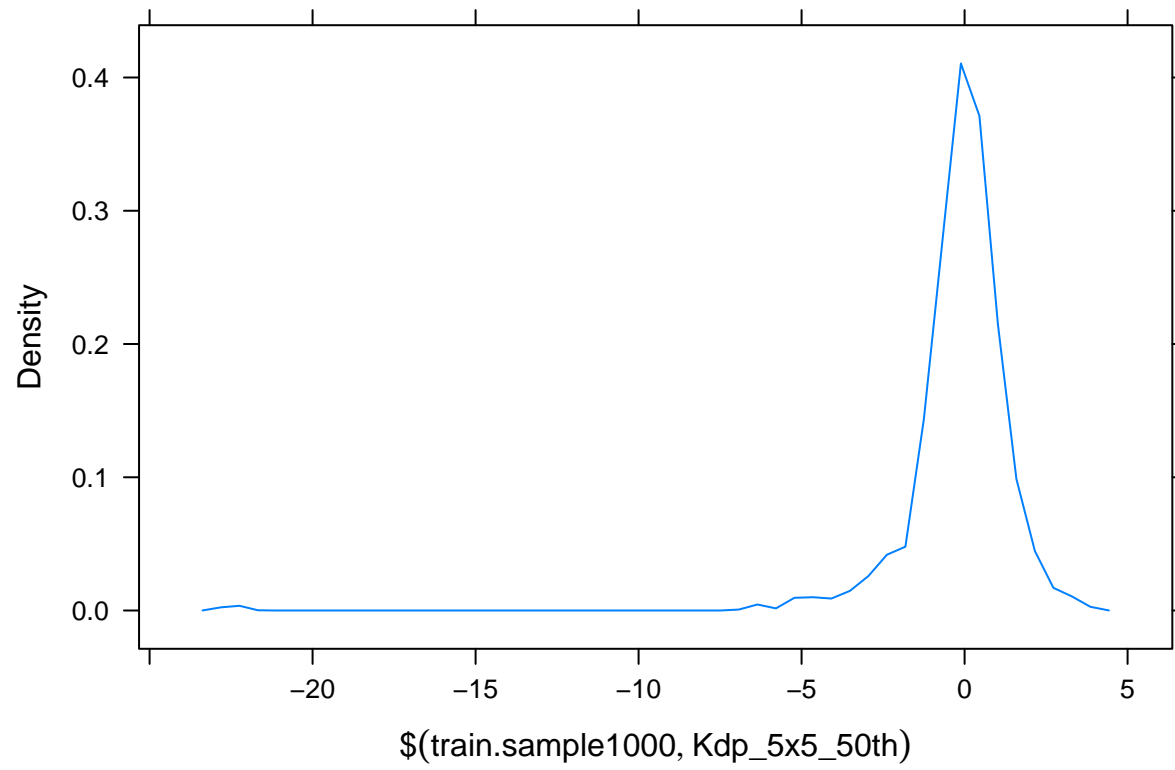
Kdp



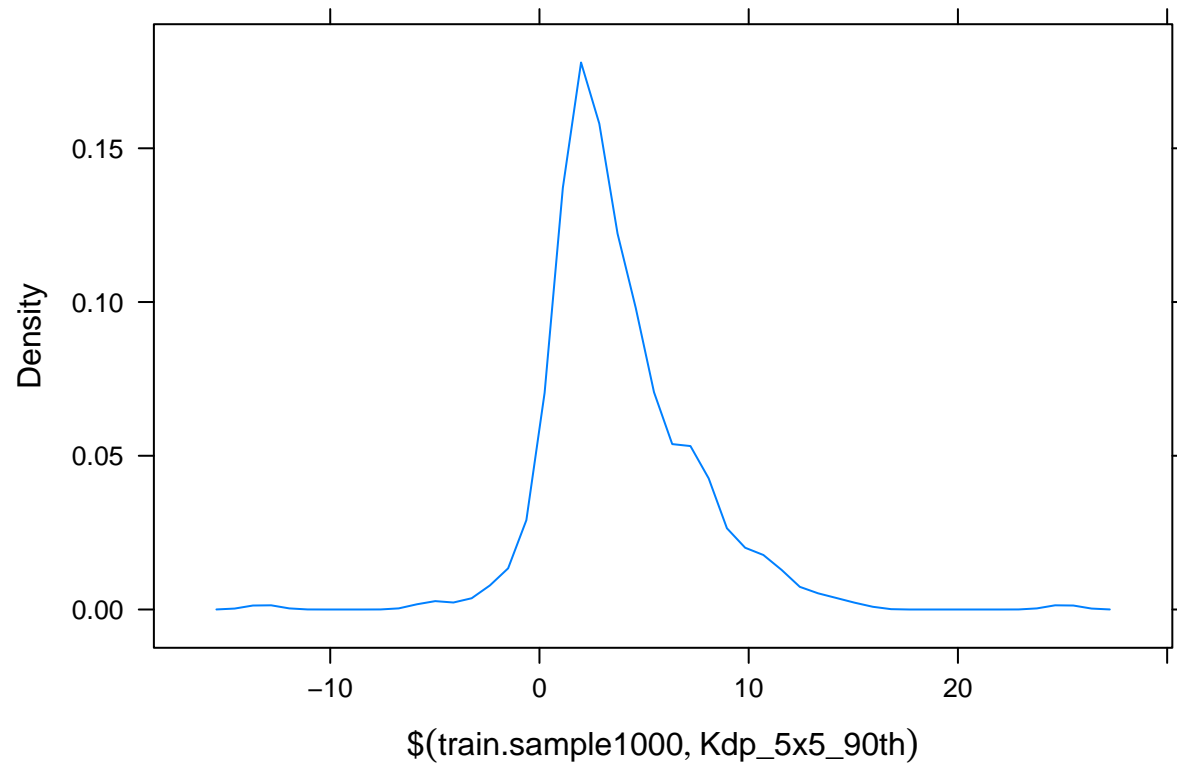
Kdp_5x5_10th



Kdp_5x5_50th



Kdp_5x5_90th



Expected

