

# 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 train from Rdata fileloading test from Rdata file
```

Summary of the Kaggle Rain *training* data set, which has 8476966 rows and 24 columns.

There are 5707878 rows with at least one NA, which is 67.333973% of the data set.

There are 828265 unique Ids in the data set, representing an hourly rain gauge measurement.

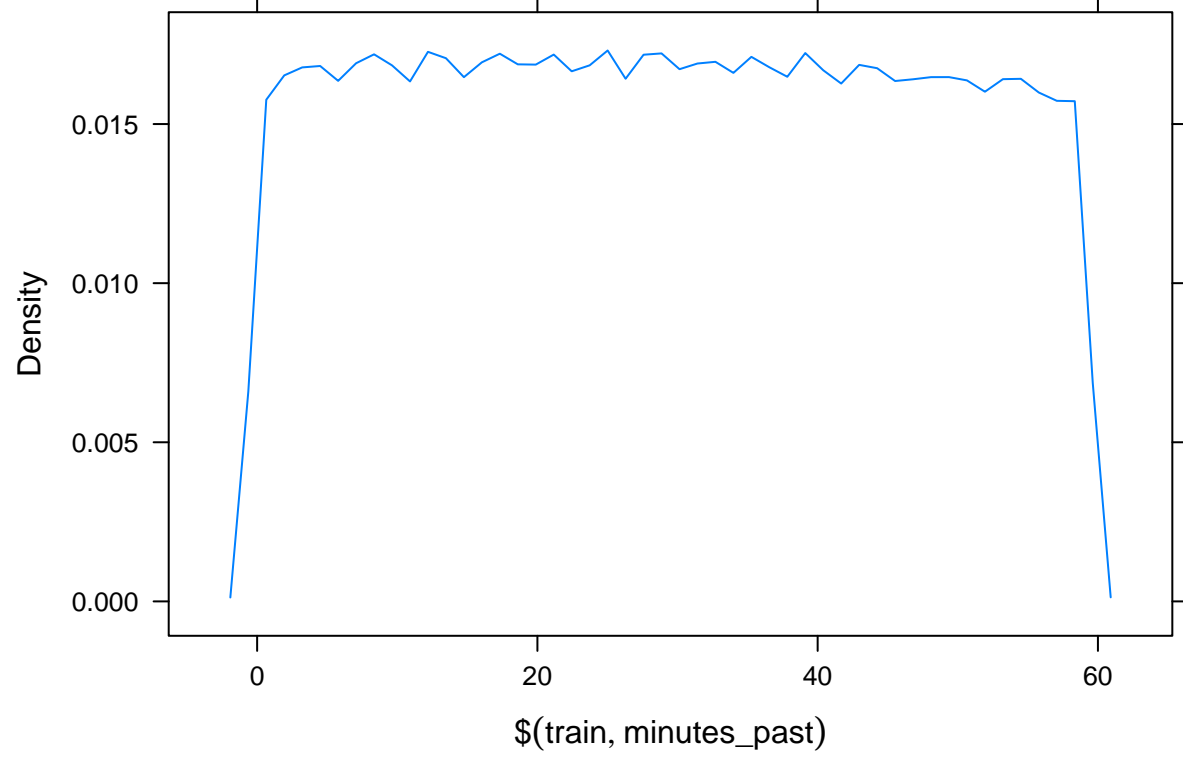
## Variable Summary

##	nmiss	min	mean	max	P01
## Id	0	2.00	591071.44	1180945.00	12689.000
## minutes_past	0	0.00	29.34	59.00	0.000
## radardist_km	0	0.00	9.63	21.00	1.000
## Ref	2127591	-31.00	22.93	71.00	-2.500
## Ref_5x5_10th	3192978	-32.00	19.95	62.50	-4.500
## Ref_5x5_50th	2120484	-32.00	22.61	69.00	-2.500
## Ref_5x5_90th	925685	-28.50	25.90	72.50	0.500

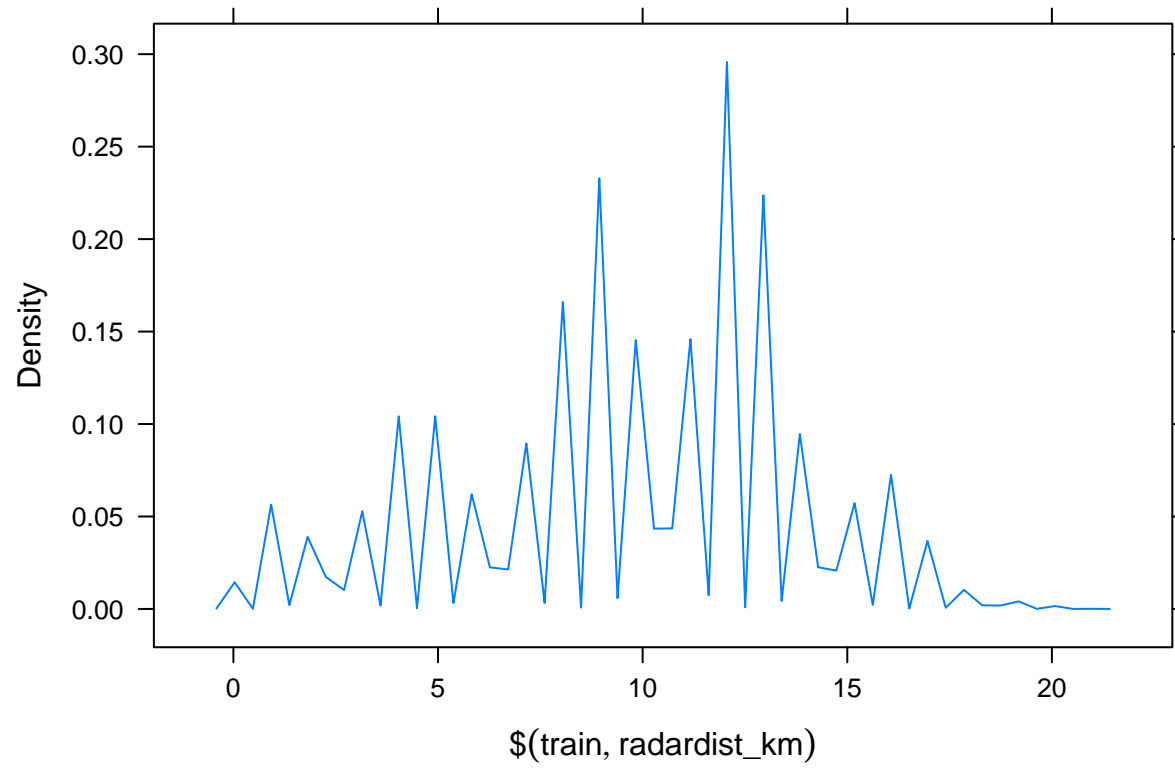
## RefComposite	1760623	-32.00	24.71	92.50	-0.500
## RefComposite_5x5_10th	2721293	-31.00	22.16	66.00	-2.500
## RefComposite_5x5_50th	1765303	-27.50	24.42	71.00	-0.500
## RefComposite_5x5_90th	647763	-25.00	27.37	93.50	1.500
## RhoHV	3542050	0.21	0.97	1.05	0.512
## RhoHV_5x5_10th	4343812	0.21	0.92	1.05	0.302
## RhoHV_5x5_50th	3540398	0.21	0.97	1.05	0.598
## RhoHV_5x5_90th	2571382	0.21	1.02	1.05	0.828
## Zdr	3542050	-7.88	0.54	7.94	-2.938
## Zdr_5x5_10th	4343812	-7.88	-0.72	7.94	-4.250
## Zdr_5x5_50th	3540398	-7.88	0.34	7.94	-2.188
## Zdr_5x5_90th	2571382	-7.88	2.07	7.94	-0.875
## Kdp	4294331	-96.04	0.04	179.75	-10.900
## Kdp_5x5_10th	5048184	-80.79	-3.48	3.52	-13.260
## Kdp_5x5_50th	4289685	-78.77	-0.47	12.80	-9.470
## Kdp_5x5_90th	3424190	-100.20	4.08	144.60	-6.310
## Expected	0	0.01	24.32	33017.73	0.010
##	P05	P25	P50	P75	
## Id	64610.000	294470.000	591485.000	889741.750	
## minutes_past	3.000	14.000	29.000	44.000	
## radardist_km	2.000	7.000	10.000	12.000	
## Ref	6.500	16.000	22.500	29.500	
## Ref_5x5_10th	5.000	14.000	20.000	26.000	
## Ref_5x5_50th	6.500	16.000	22.500	29.000	
## Ref_5x5_90th	8.500	18.000	25.500	33.500	
## RefComposite	8.000	17.500	24.000	31.500	
## RefComposite_5x5_10th	6.500	16.000	22.000	28.500	
## RefComposite_5x5_50th	8.000	17.500	24.000	31.500	
## RefComposite_5x5_90th	9.500	19.500	27.000	35.000	
## RhoHV	0.835	0.972	0.992	1.002	
## RhoHV_5x5_10th	0.662	0.915	0.958	0.982	
## RhoHV_5x5_50th	0.858	0.975	0.992	0.998	
## RhoHV_5x5_90th	0.978	0.998	1.012	1.052	
## Zdr	-1.375	-0.188	0.375	1.062	
## Zdr_5x5_10th	-2.312	-1.125	-0.625	-0.188	
## Zdr_5x5_50th	-0.750	-0.062	0.250	0.688	
## Zdr_5x5_90th	0.375	1.062	1.688	2.625	
## Kdp	-5.630	-1.410	0.000	1.750	
## Kdp_5x5_10th	-8.430	-4.580	-2.820	-1.760	

## Kdp_5x5_50th	-3.530	-0.710	0.000	0.350
## Kdp_5x5_90th	0.000	2.070	3.520	5.640
## Expected	0.254	0.350	1.270	3.556
##	P95	P99	variance	stdev
## Id	1120198.000	1168568.000	1.160216e+11	340619.39
## minutes_past	56.000	59.000	2.963600e+02	17.22
## radardist_km	16.000	18.000	1.597000e+01	4.00
## Ref	40.500	49.000	1.072300e+02	10.36
## Ref_5x5_10th	35.000	42.000	8.479000e+01	9.21
## Ref_5x5_50th	39.000	47.000	1.010600e+02	10.05
## Ref_5x5_90th	44.500	53.000	1.234200e+02	11.11
## RefComposite	43.000	51.000	1.142700e+02	10.69
## RefComposite_5x5_10th	38.000	45.500	9.414000e+01	9.70
## RefComposite_5x5_50th	42.000	50.000	1.086700e+02	10.42
## RefComposite_5x5_90th	47.000	55.000	1.320800e+02	11.49
## RhoHV	1.052	1.052	1.000000e-02	0.09
## RhoHV_5x5_10th	0.995	1.002	2.000000e-02	0.13
## RhoHV_5x5_50th	1.025	1.052	1.000000e-02	0.08
## RhoHV_5x5_90th	1.052	1.052	0.000000e+00	0.05
## Zdr	3.000	6.562	2.280000e+00	1.51
## Zdr_5x5_10th	0.562	1.312	1.010000e+00	1.01
## Zdr_5x5_50th	1.562	3.750	8.800000e-01	0.94
## Zdr_5x5_90th	5.625	7.938	2.790000e+00	1.67
## Kdp	5.630	10.860	1.497000e+01	3.87
## Kdp_5x5_10th	-0.710	0.000	7.800000e+00	2.79
## Kdp_5x5_50th	1.760	2.820	5.120000e+00	2.26
## Kdp_5x5_90th	10.730	16.680	1.720000e+01	4.15
## Expected	20.574	370.840	6.065078e+04	246.27

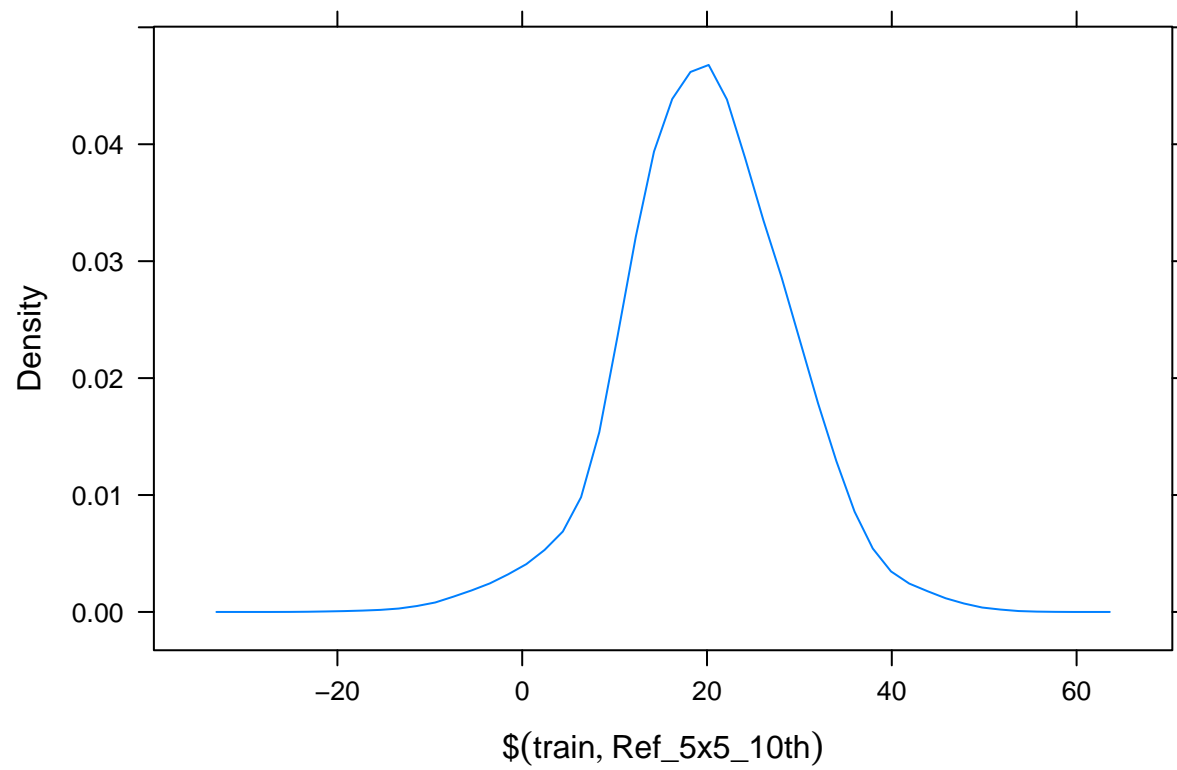
minutes\_past



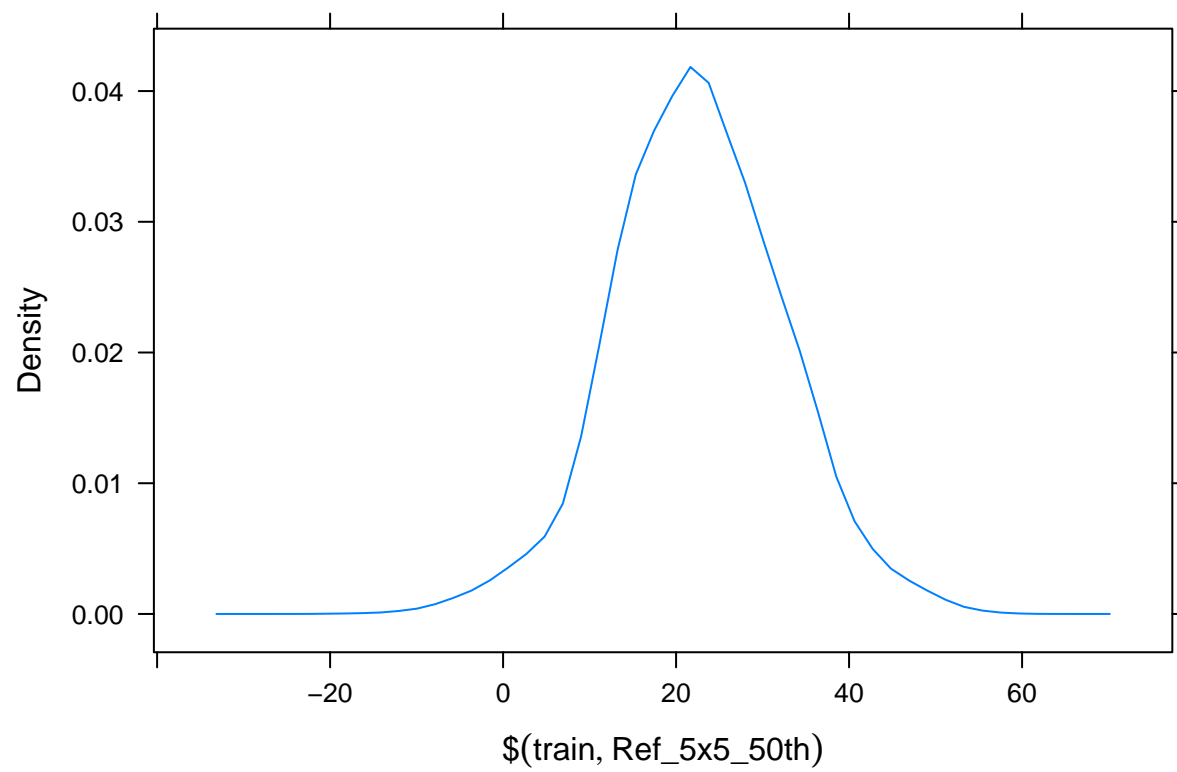
radardist\_km



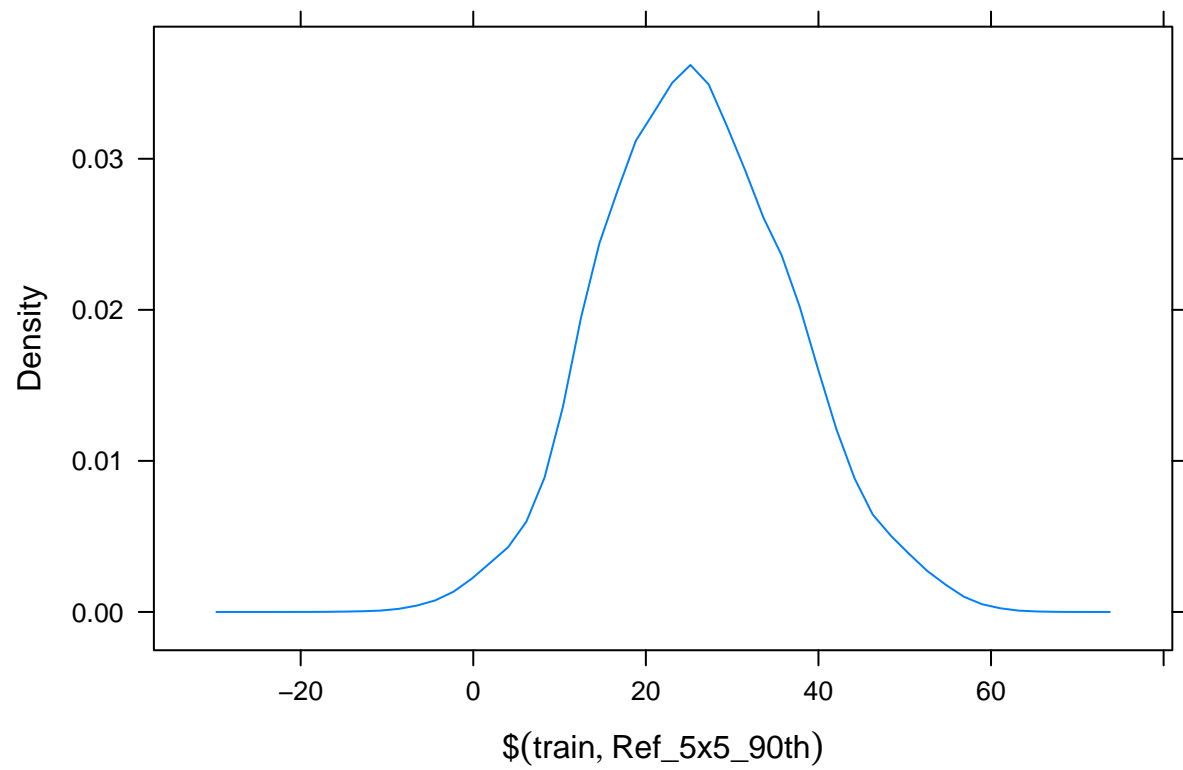
Ref\_5x5\_10th



Ref\_5x5\_50th

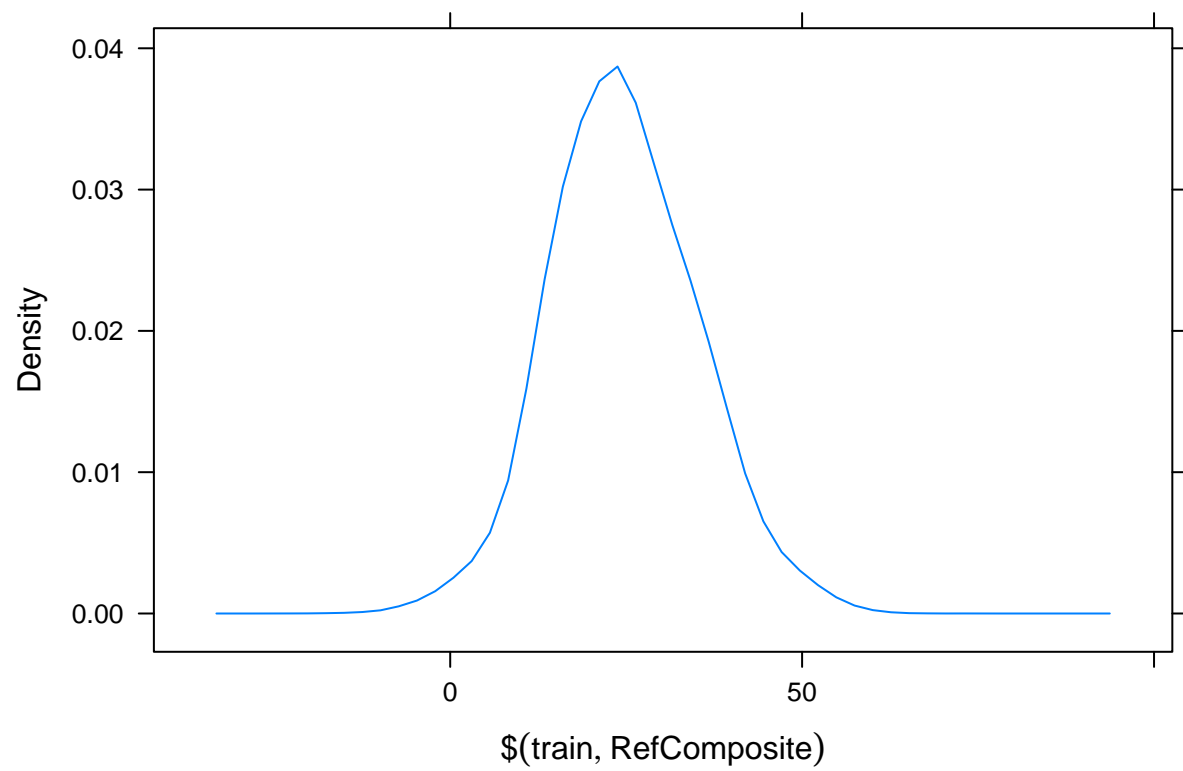


Ref\_5x5\_90th

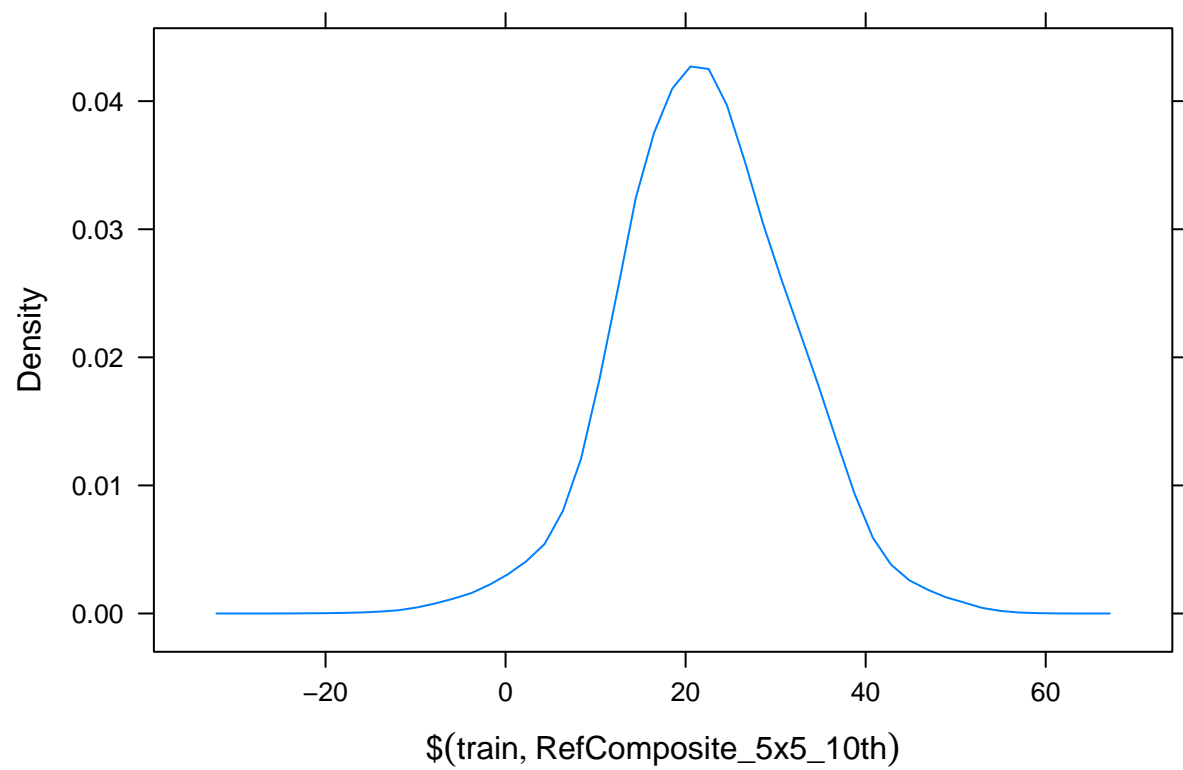




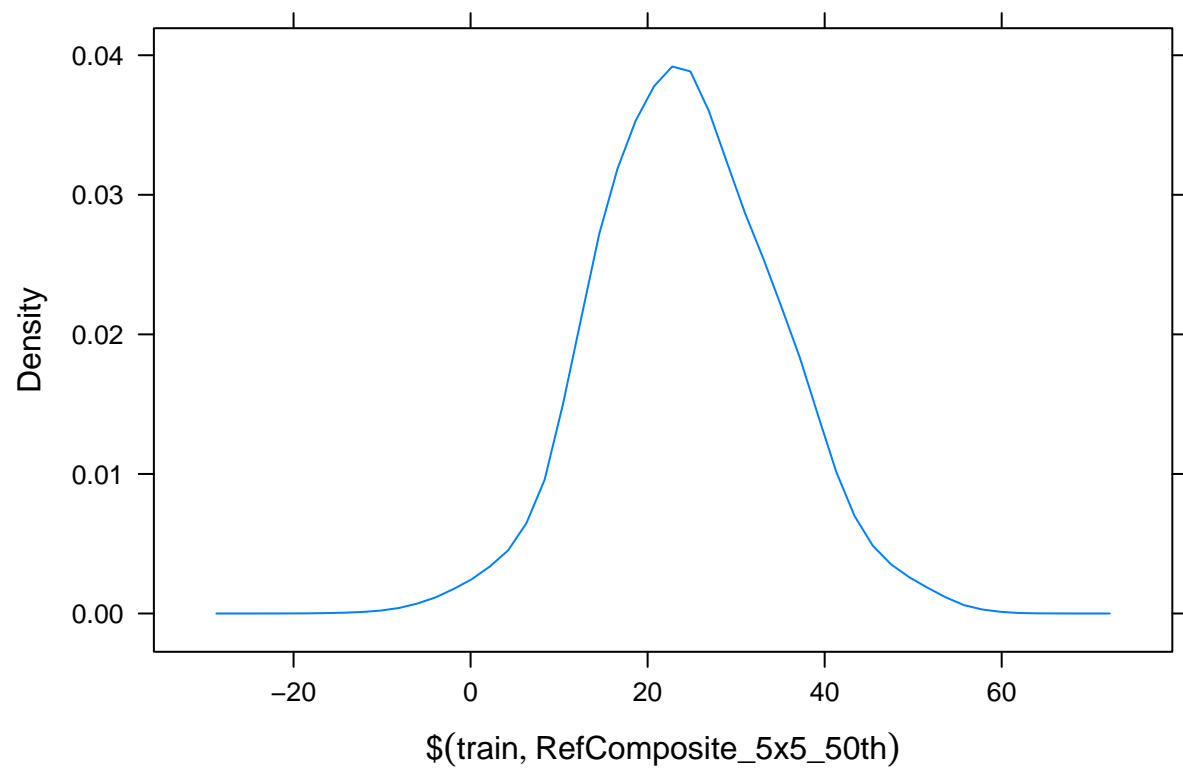
RefComposite



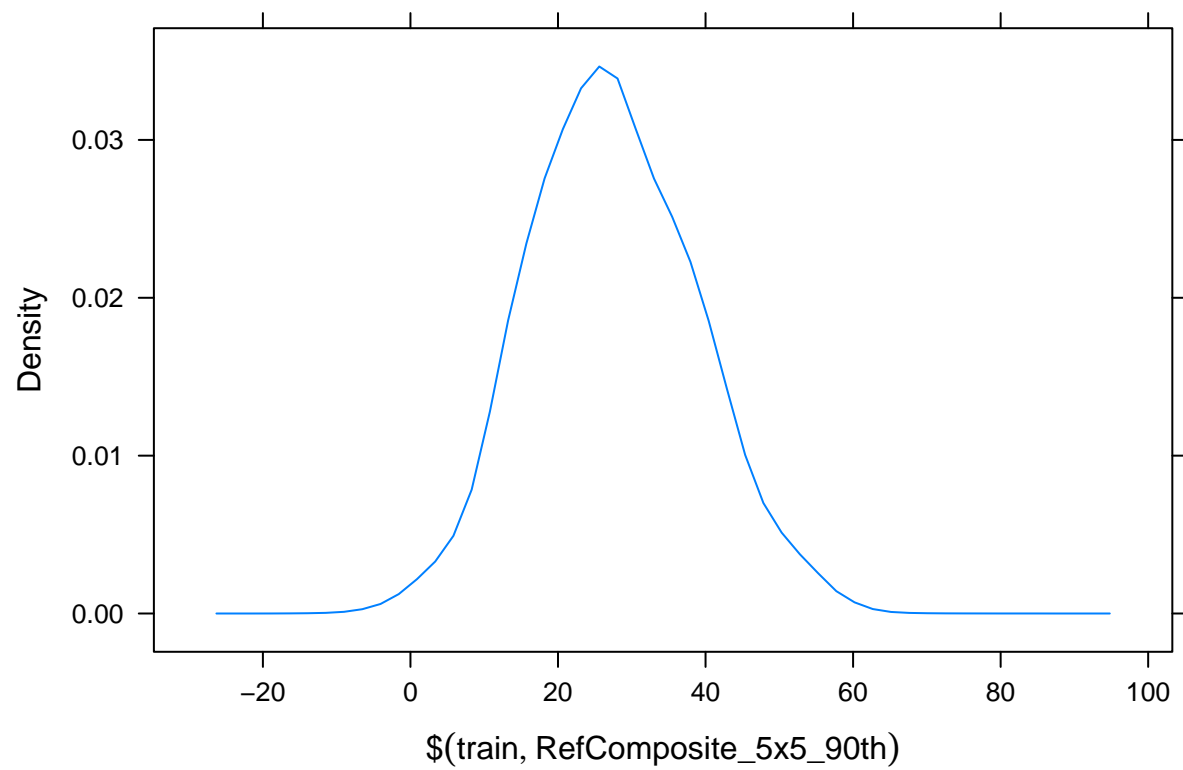
RefComposite\_5x5\_10th



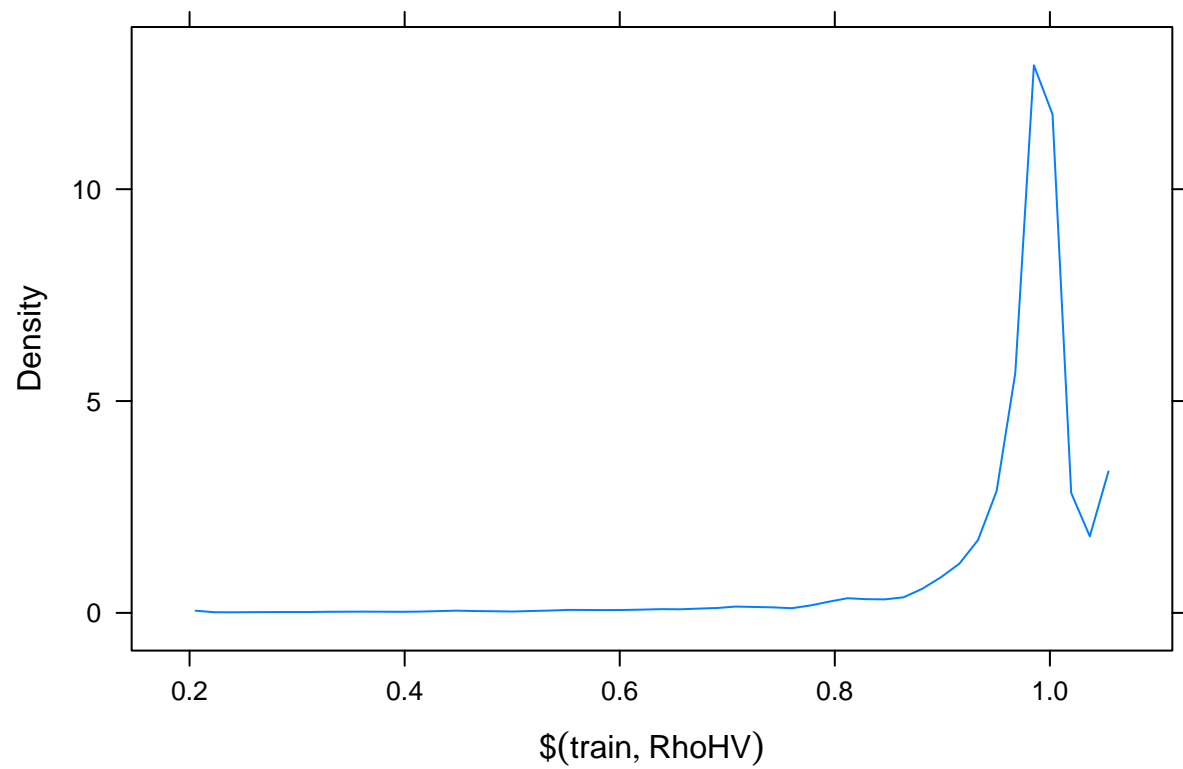
RefComposite\_5x5\_50th



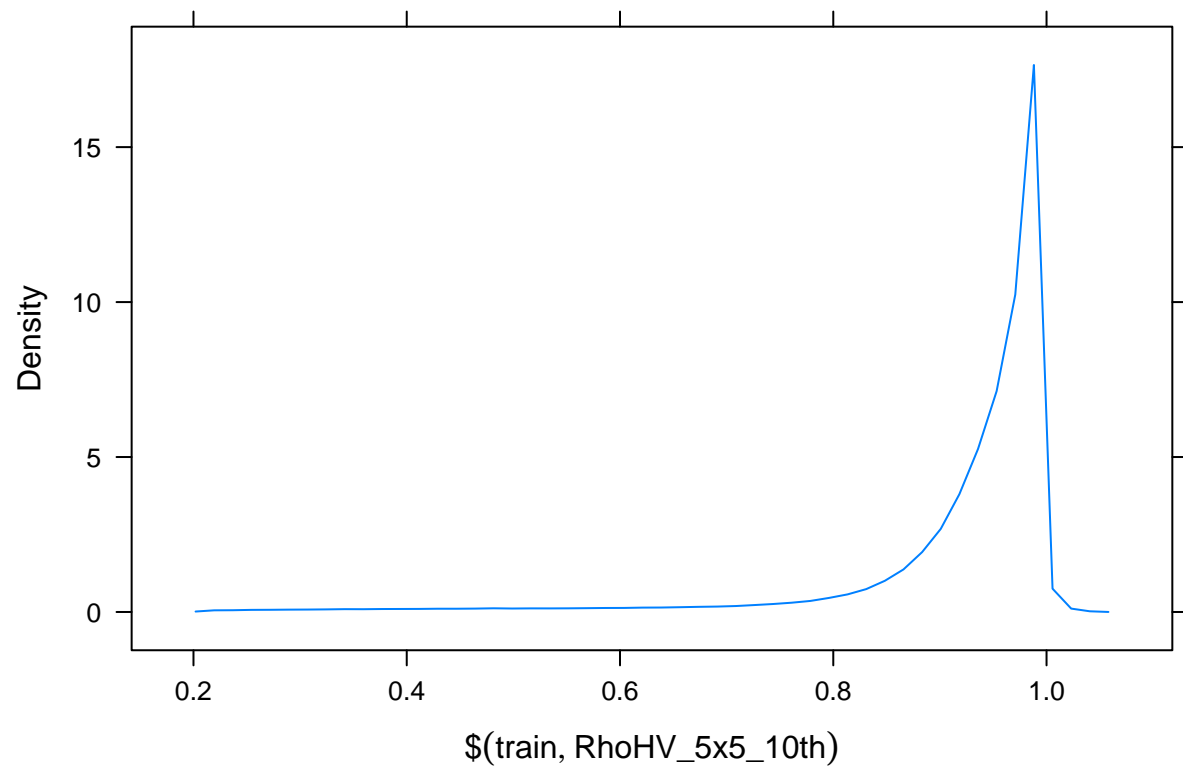
RefComposite\_5x5\_90th



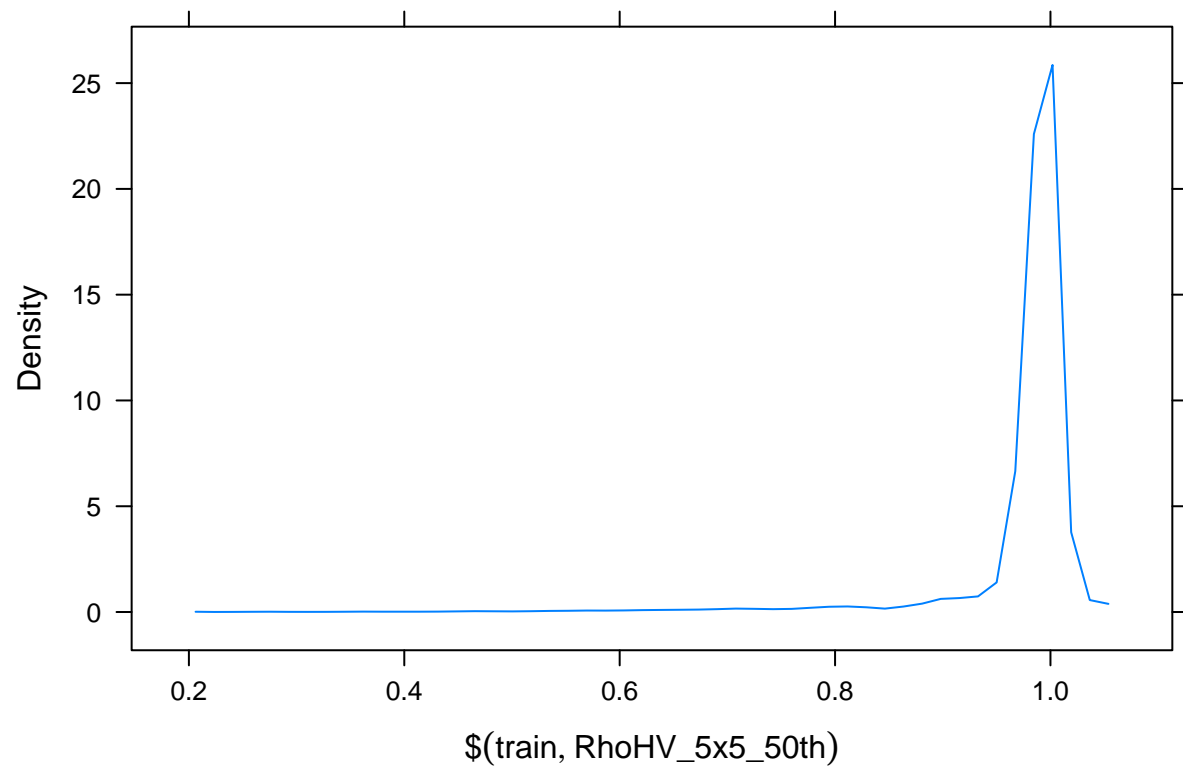
RhoHV



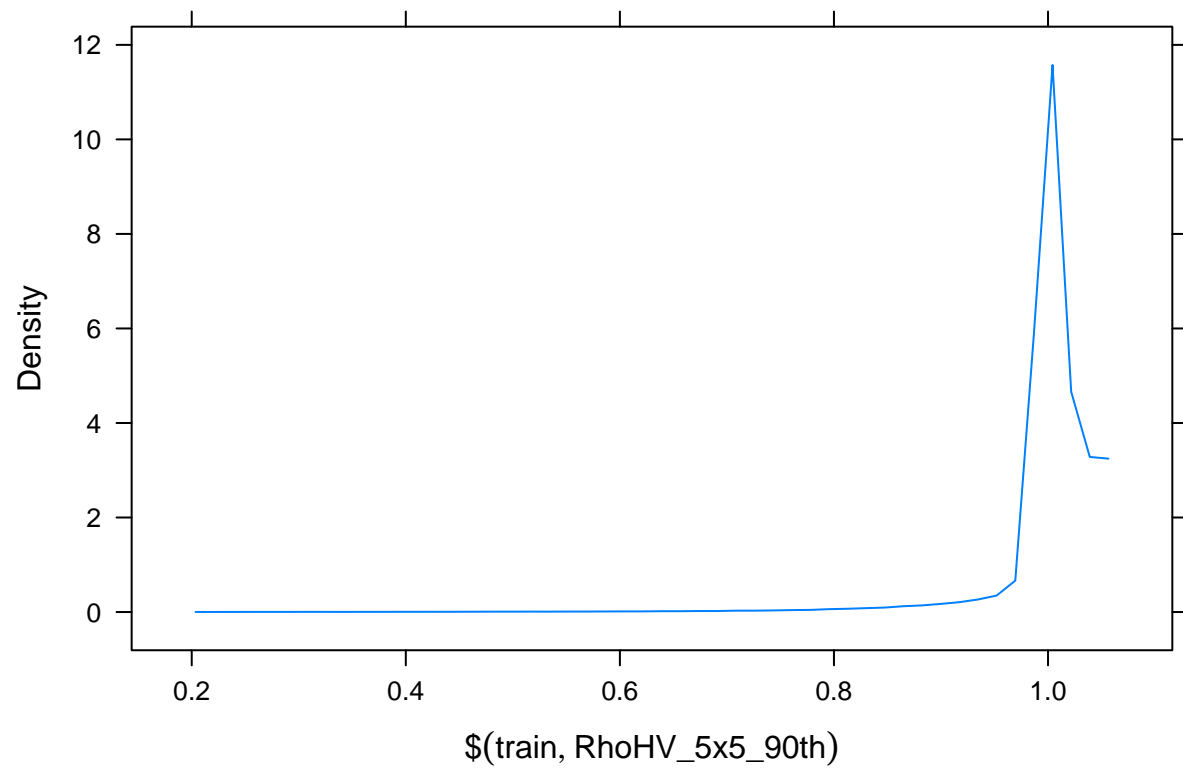
RhoHV\_5x5\_10th



RhoHV\_5x5\_50th

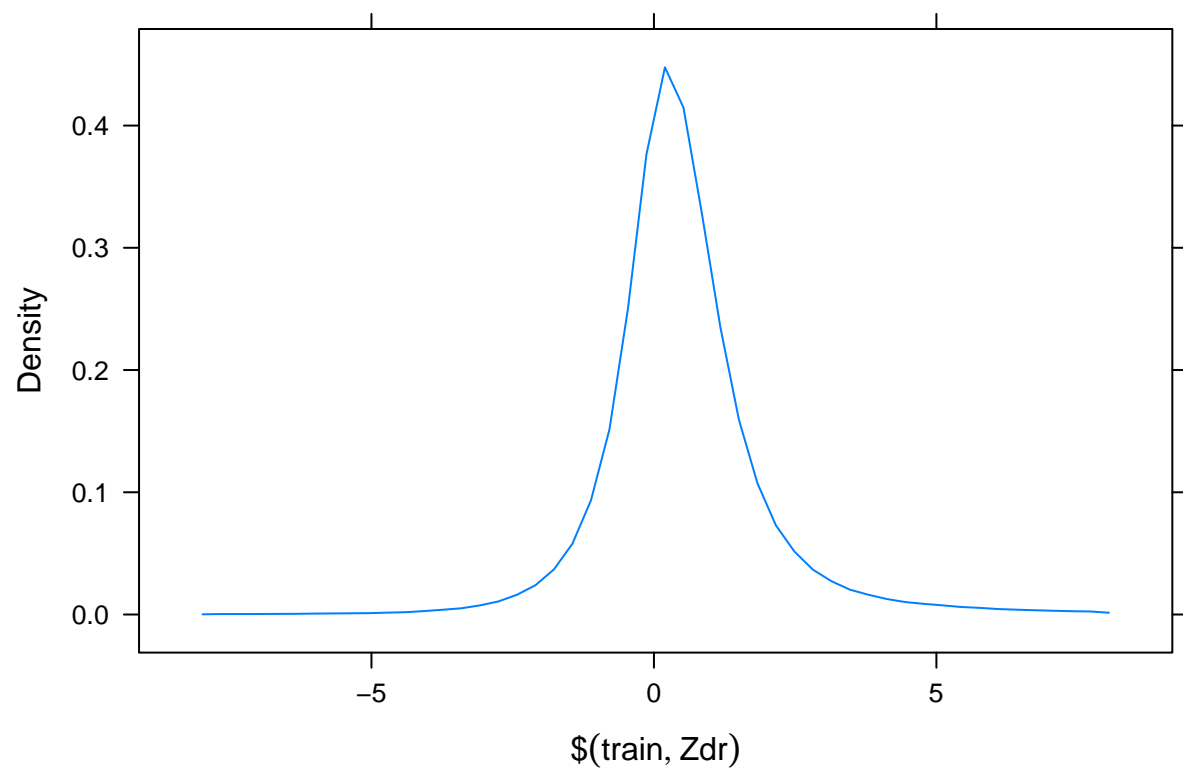


RhoHV\_5x5\_90th

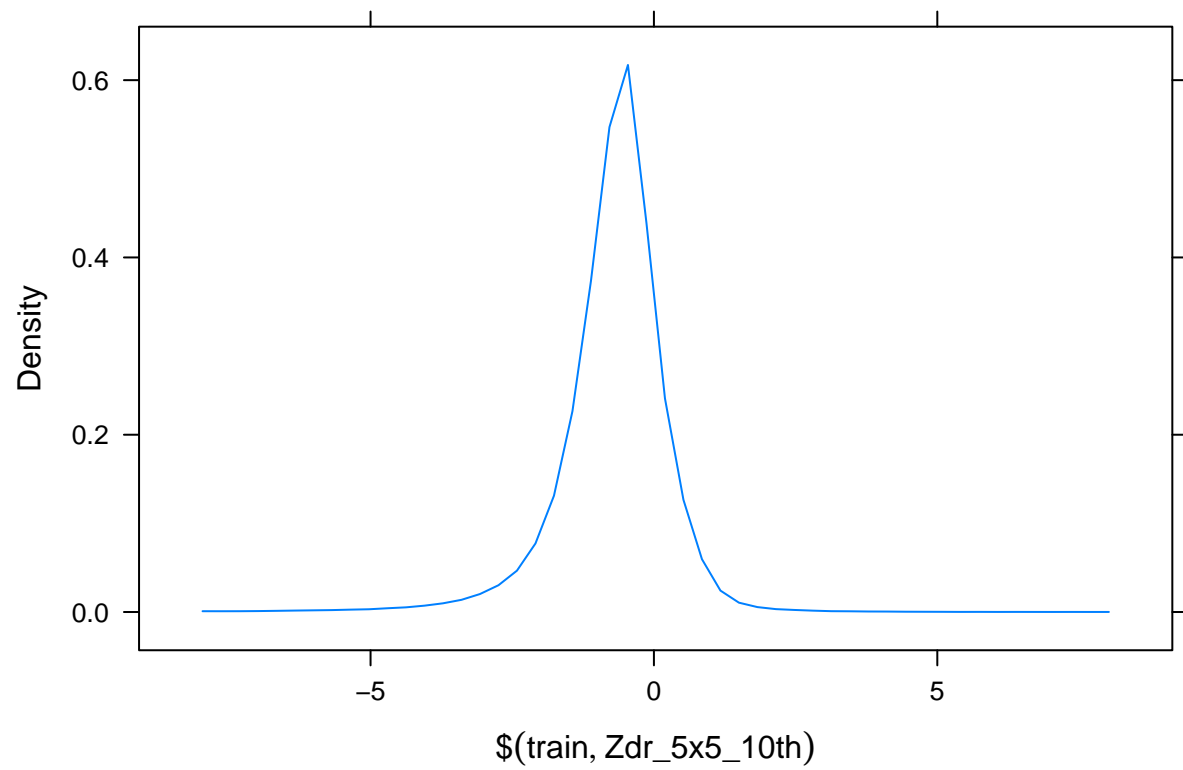




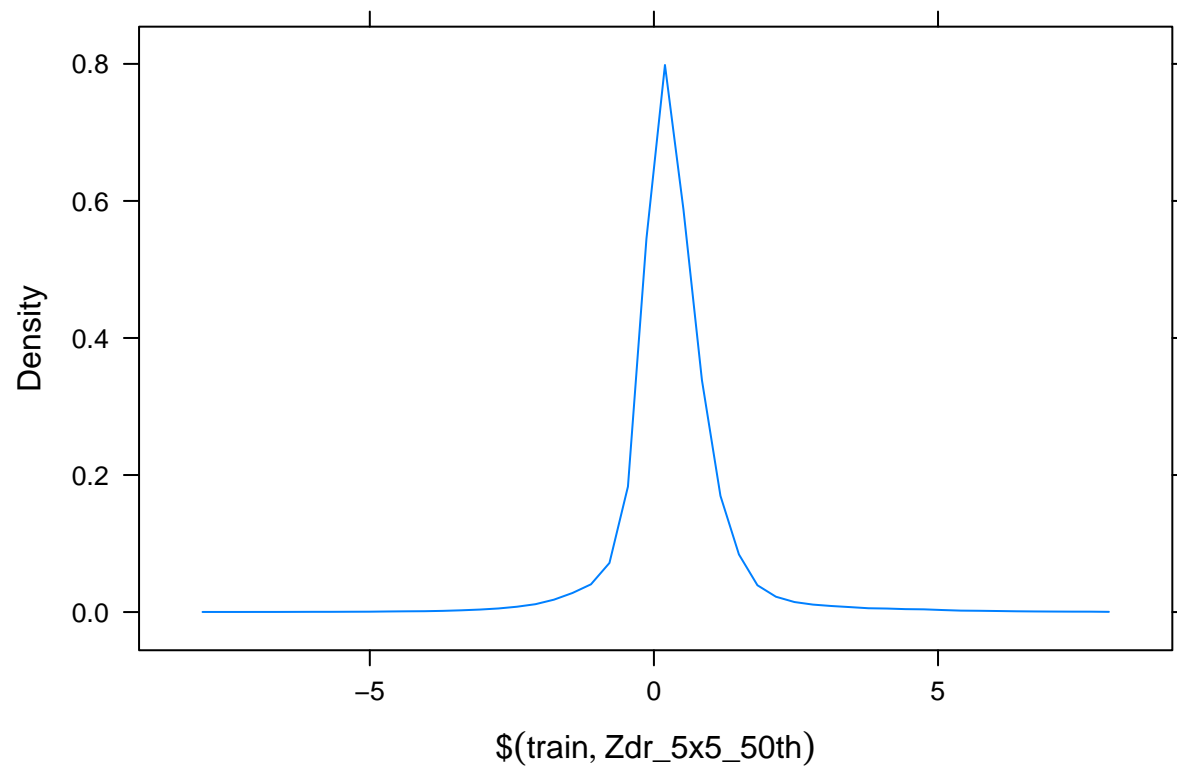
Zdr



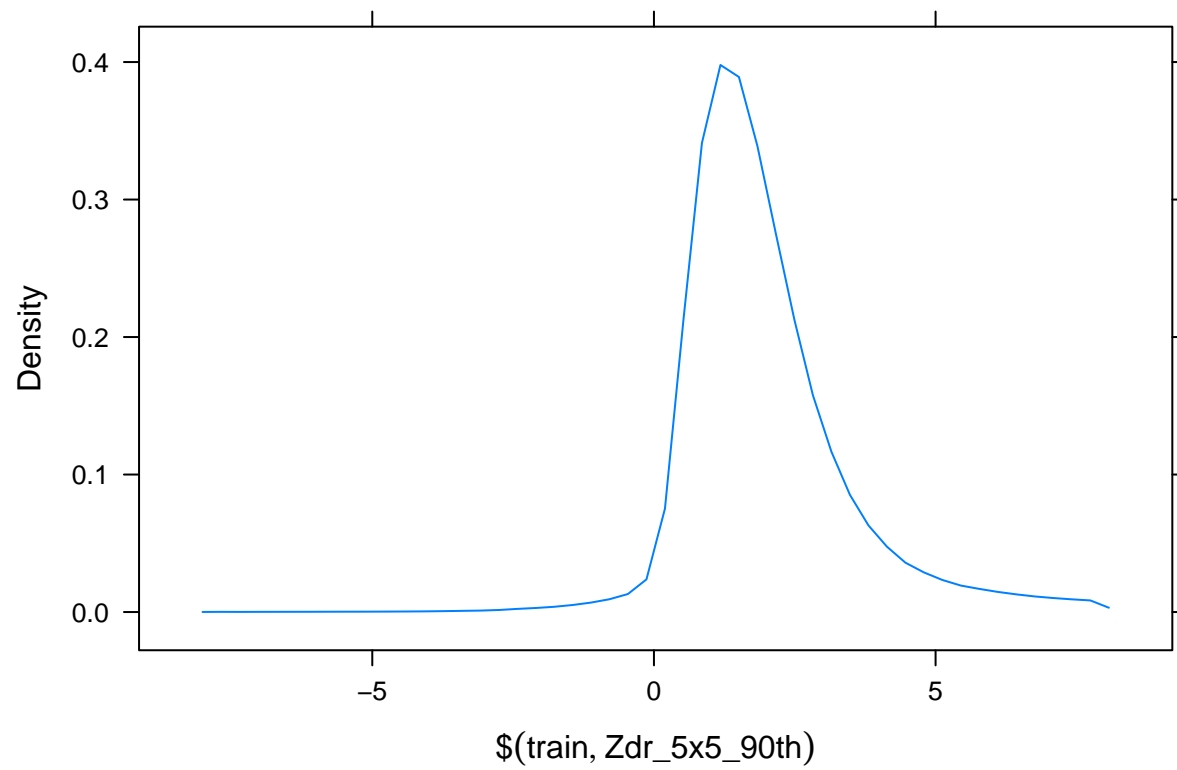
Zdr\_5x5\_10th



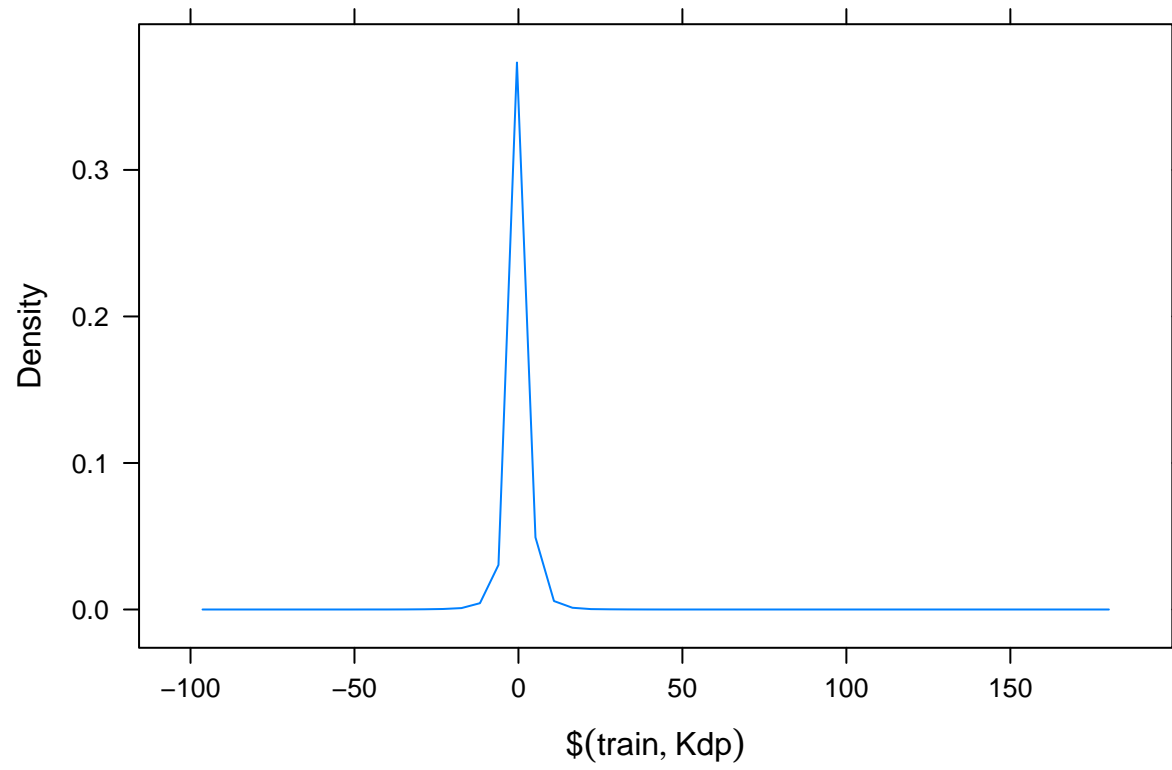
Zdr\_5x5\_50th



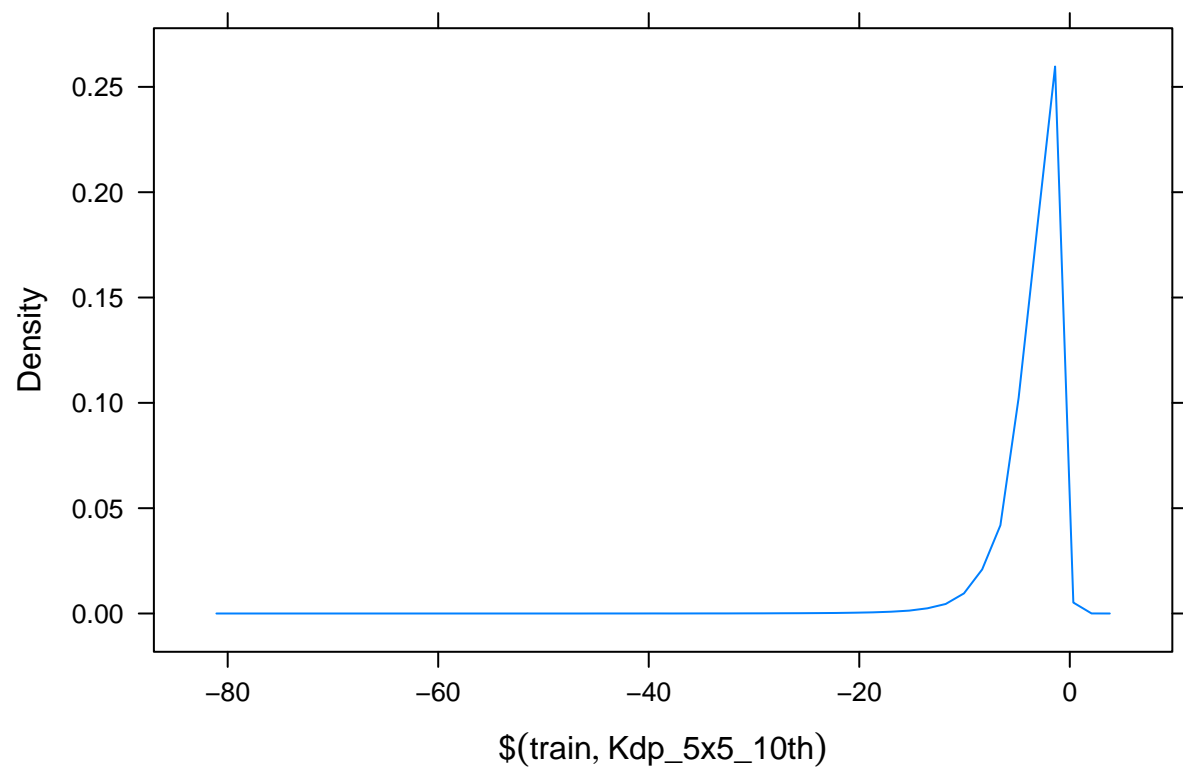
Zdr\_5x5\_90th



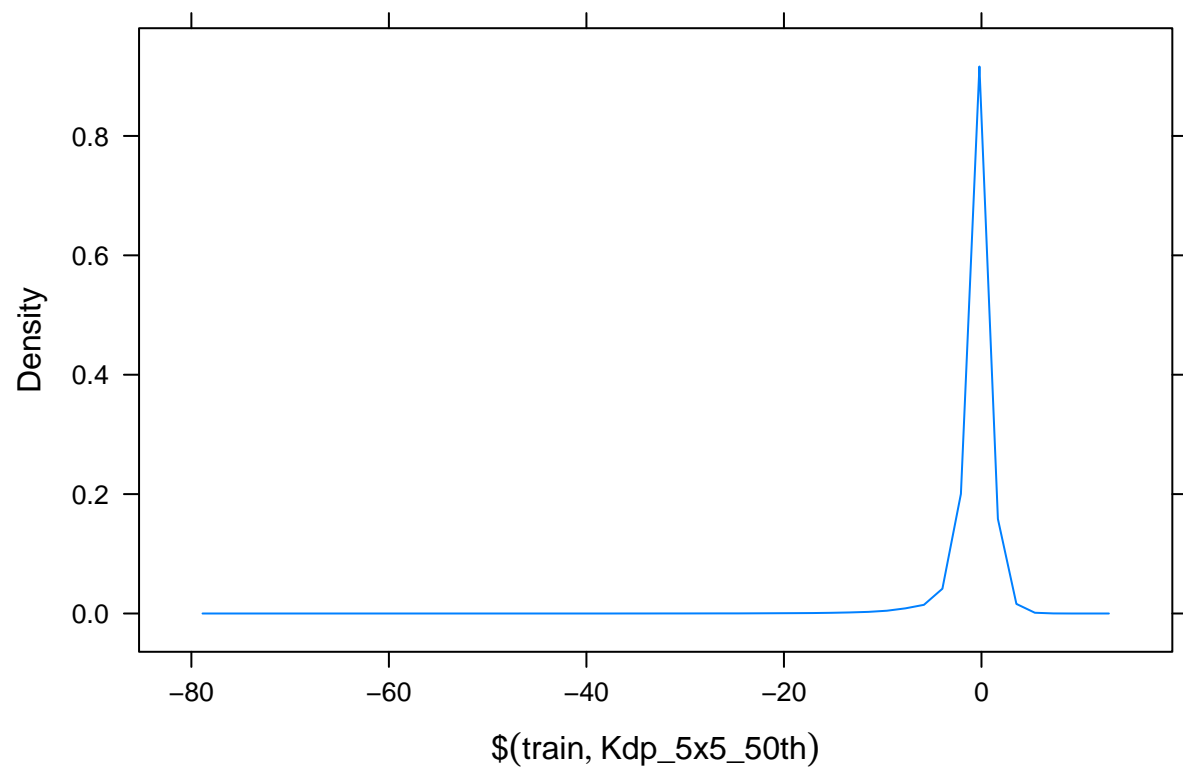
Kdp



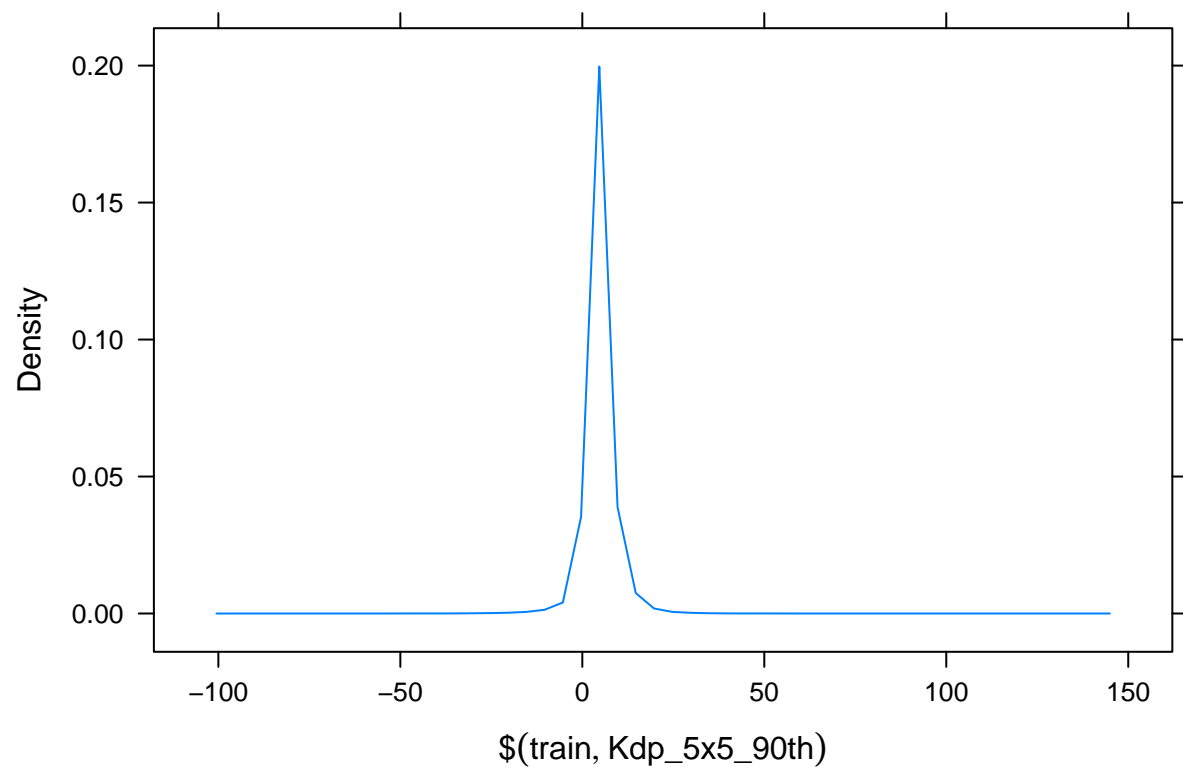
Kdp\_5x5\_10th



Kdp\_5x5\_50th



Kdp\_5x5\_90th





Expected

