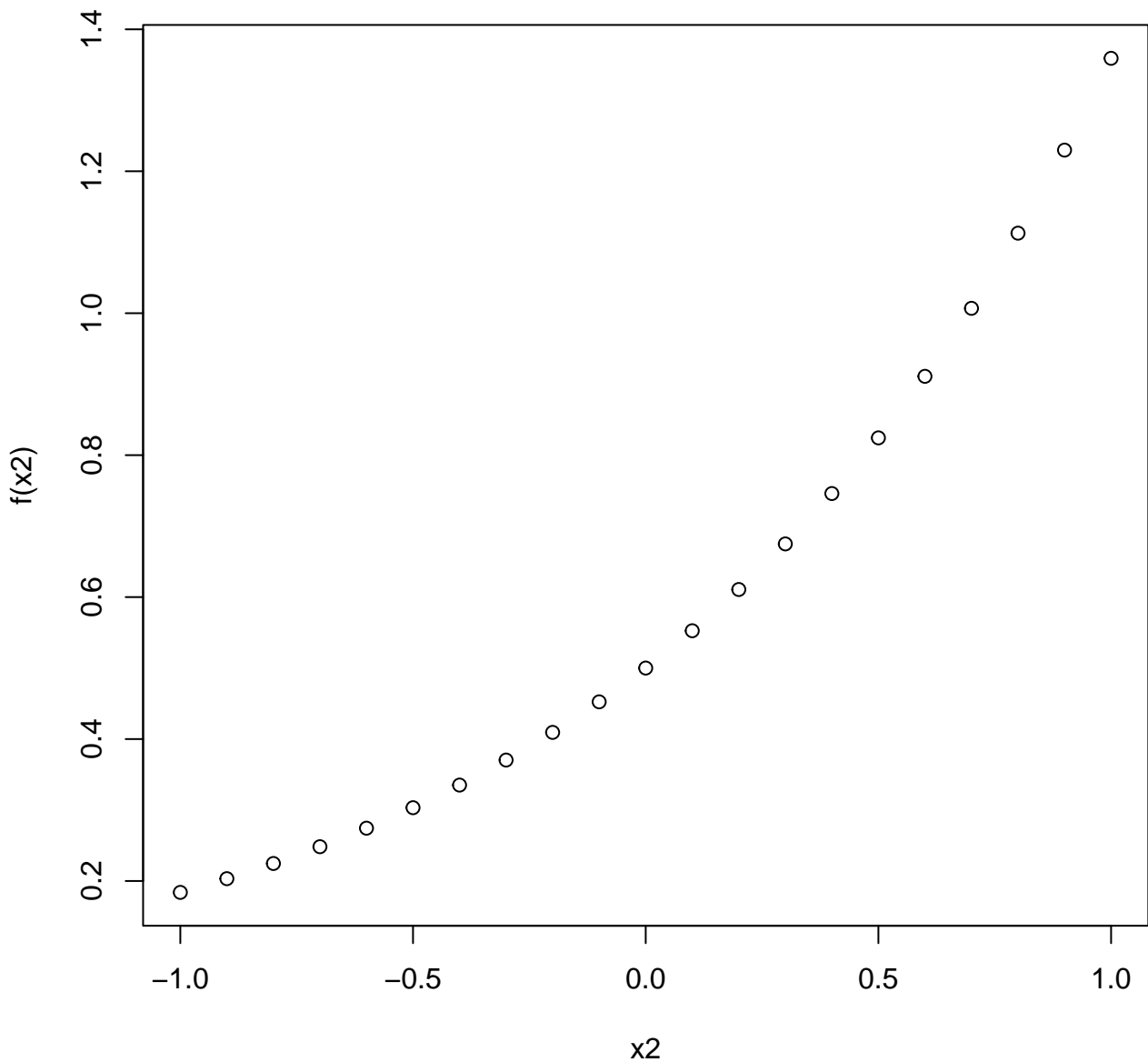
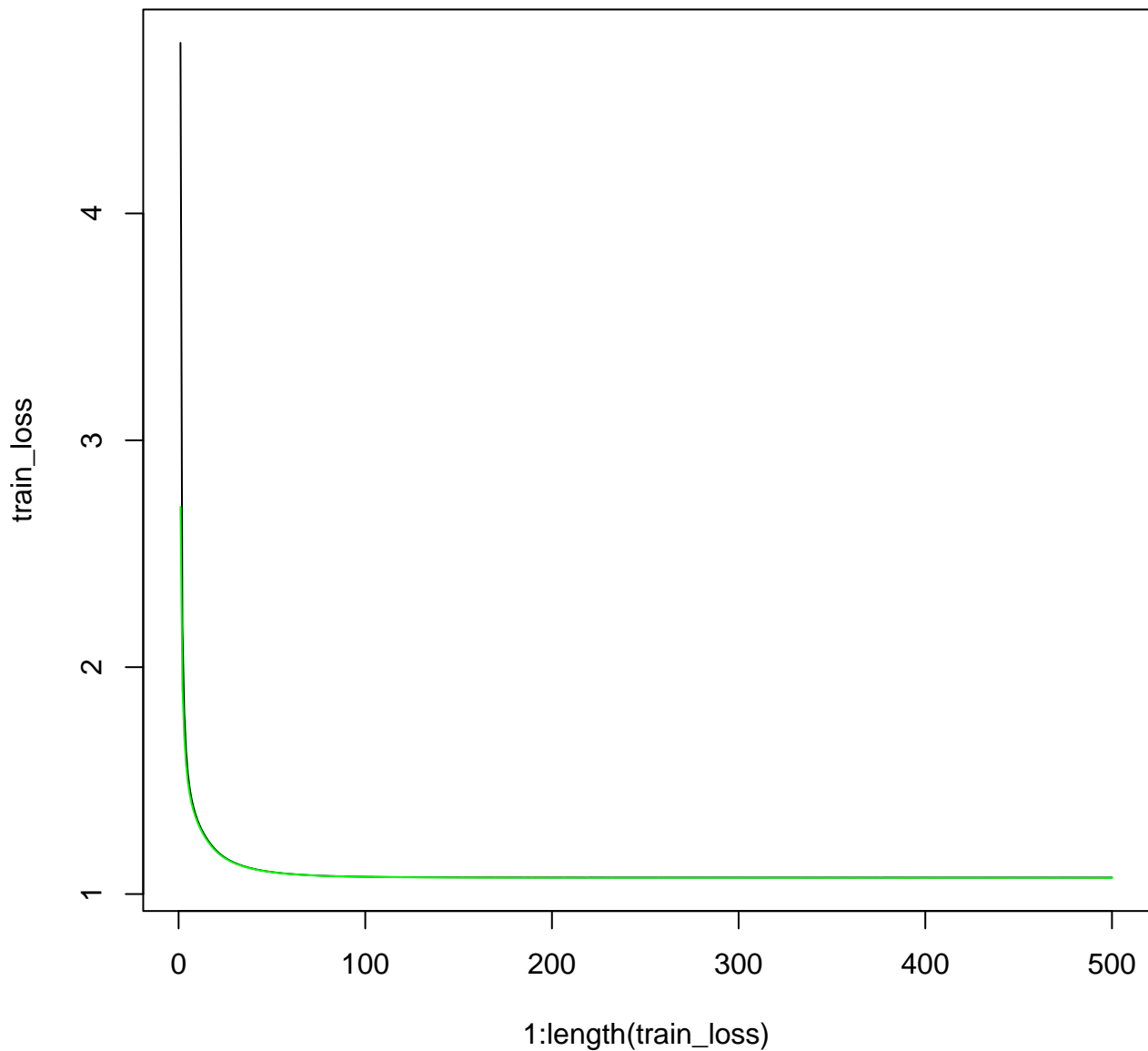
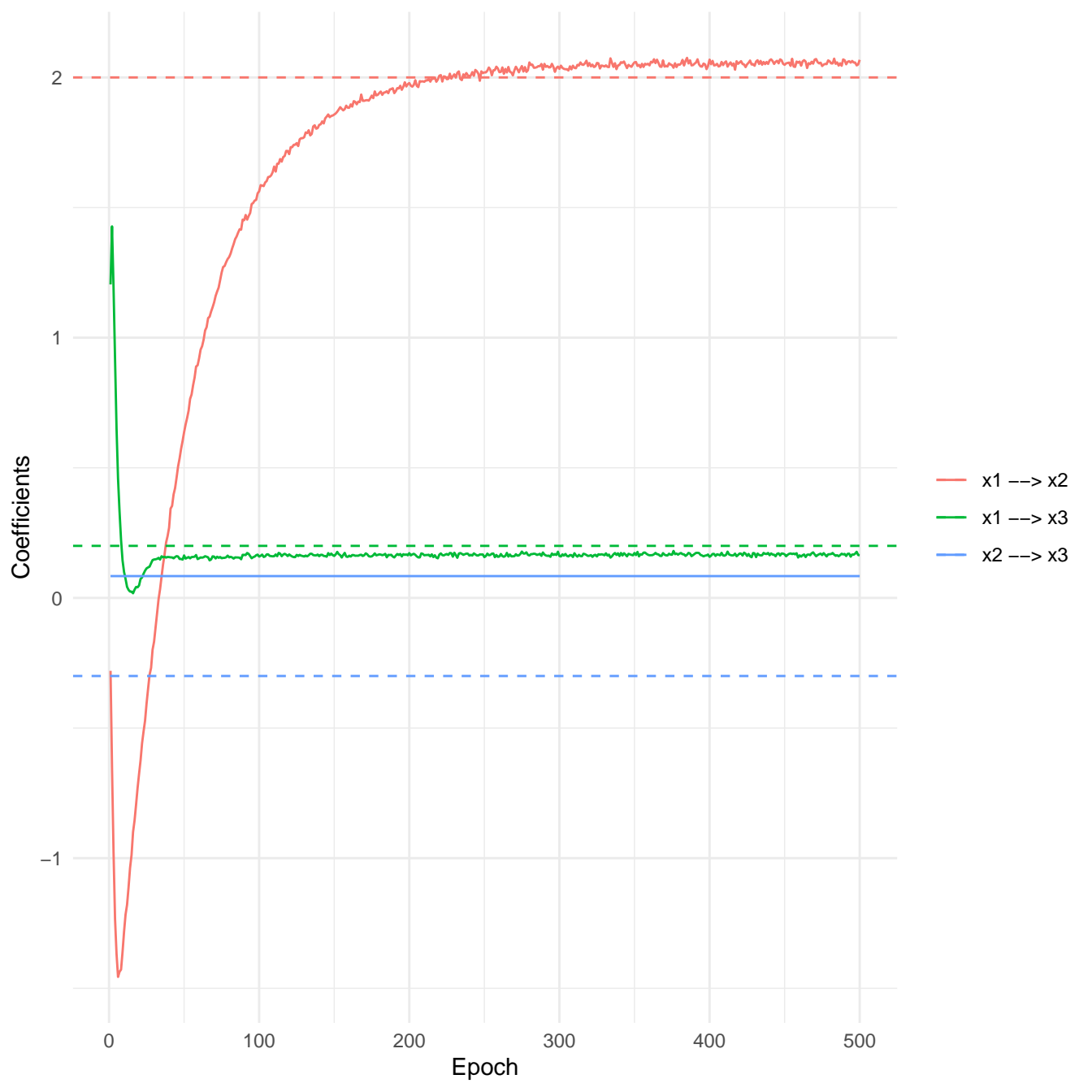


DGP influence of x2 on x3

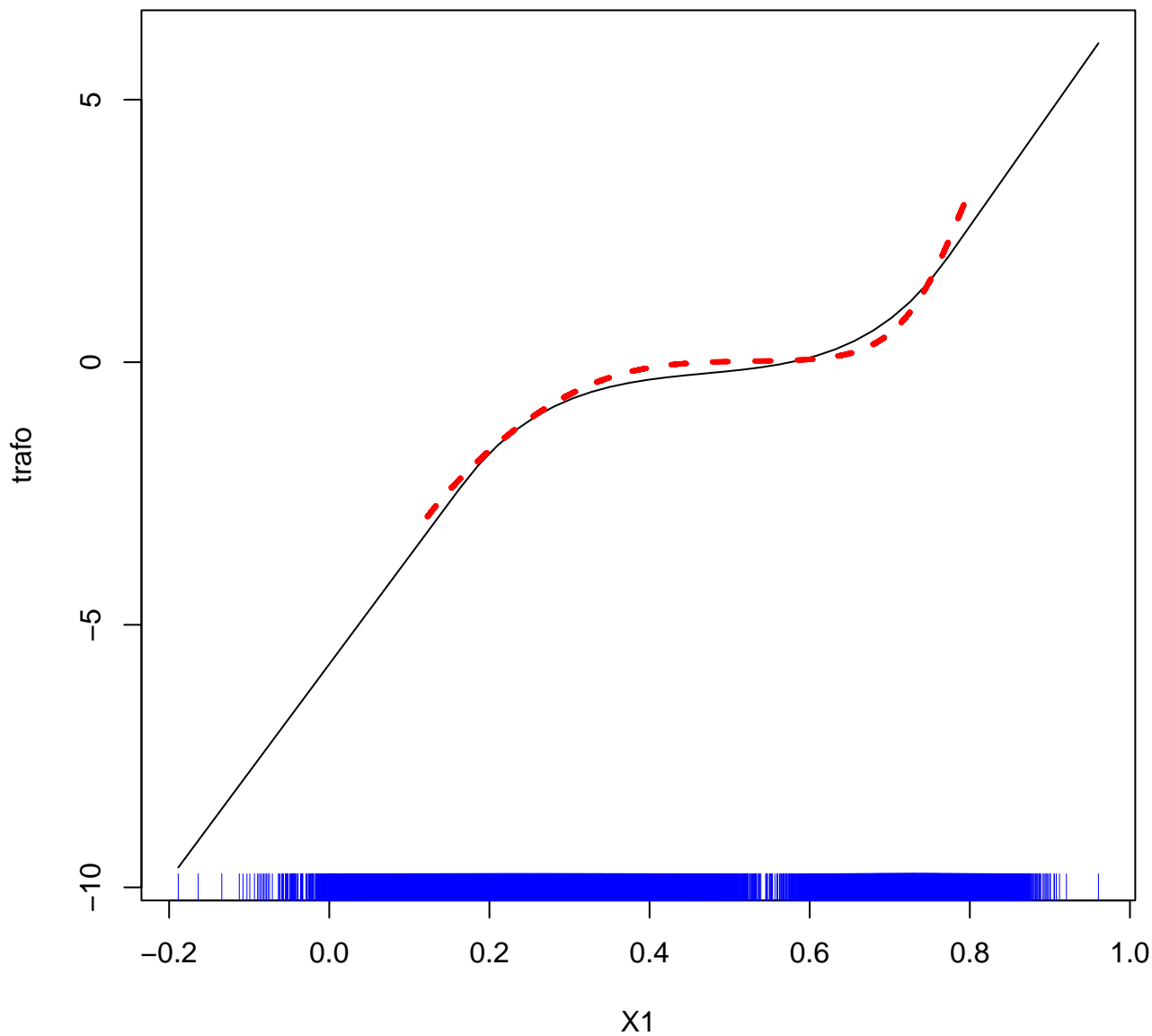


Normal Training (green is valid)

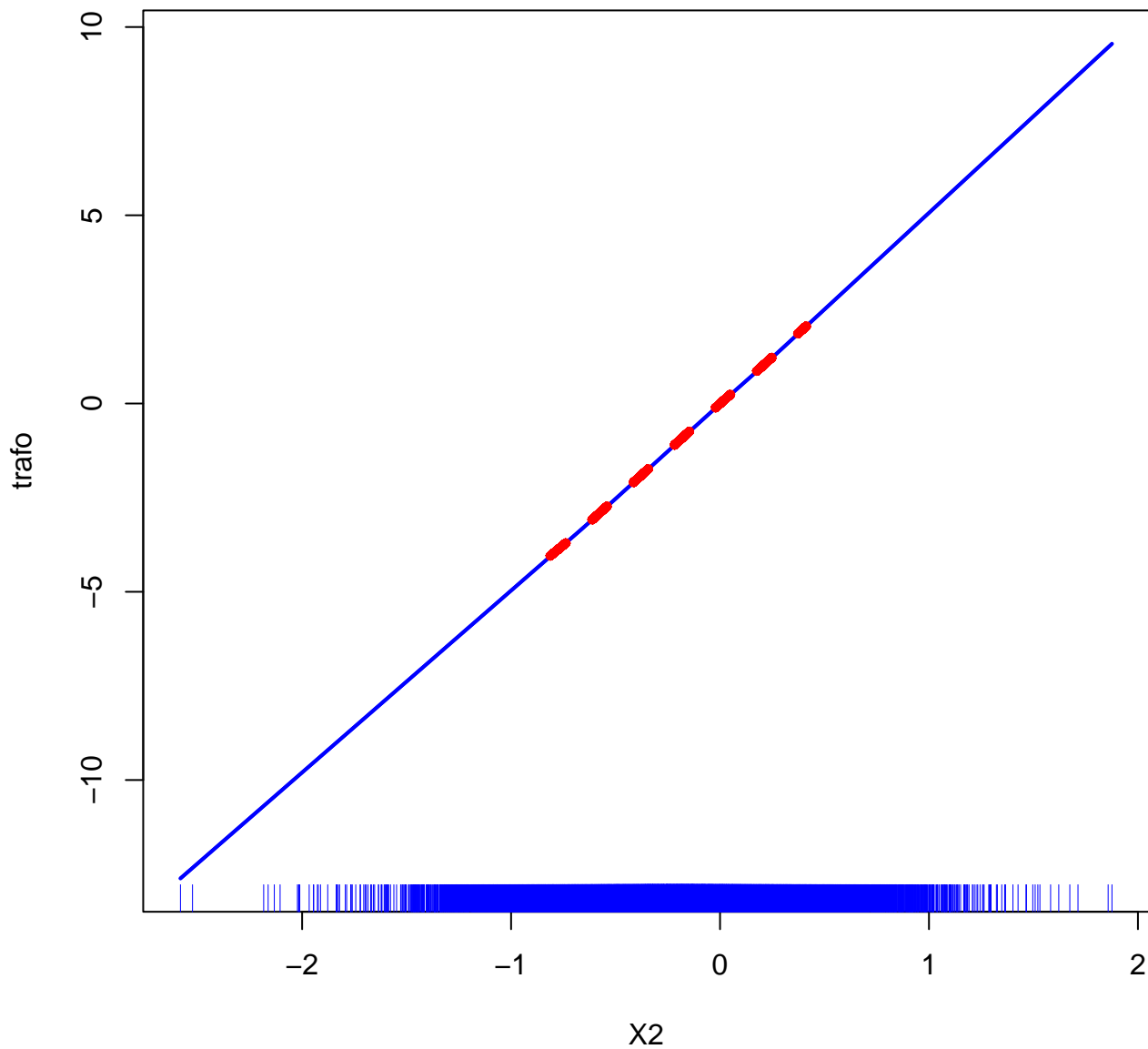




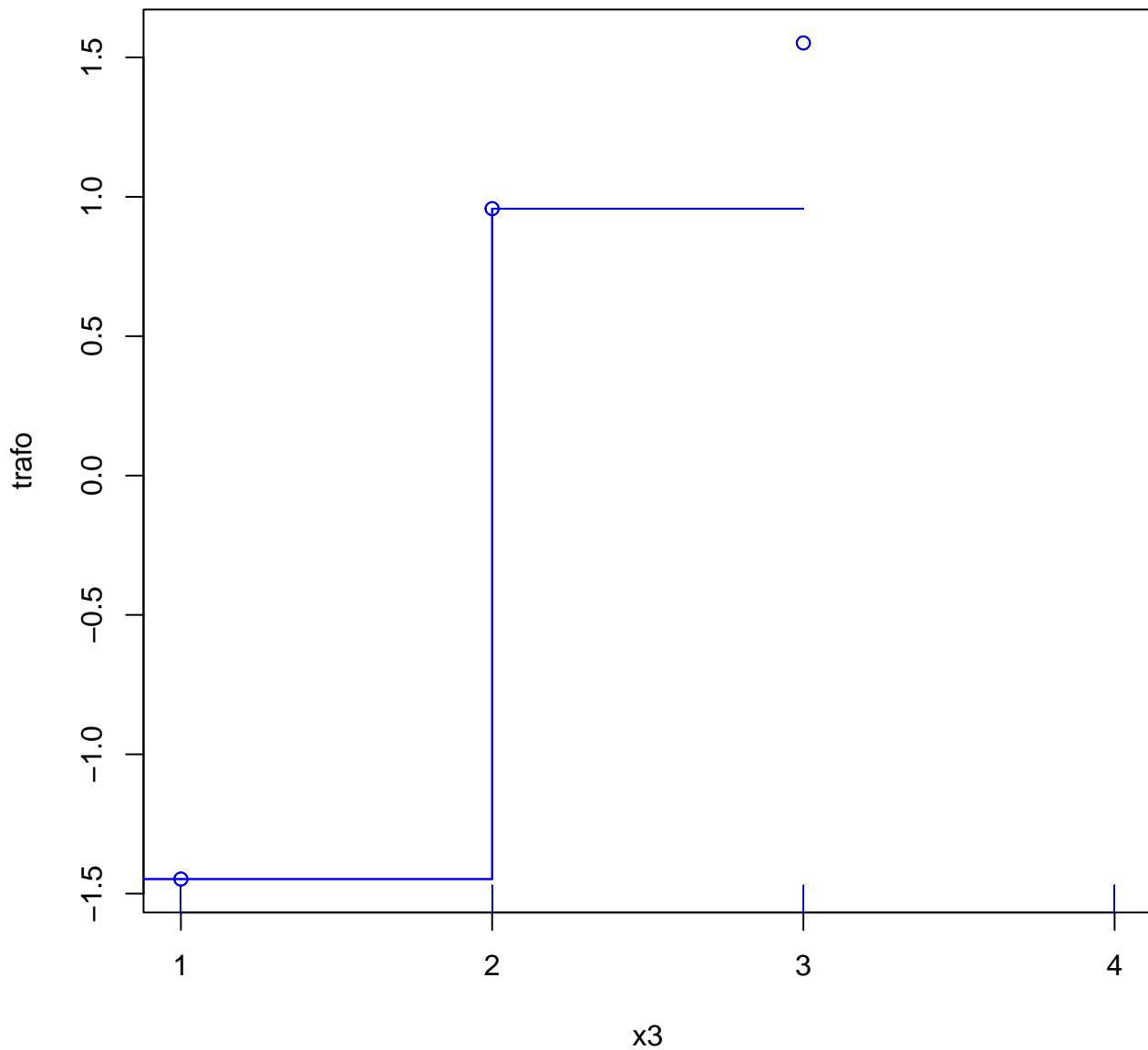
Black: COLR, Red: Our Model



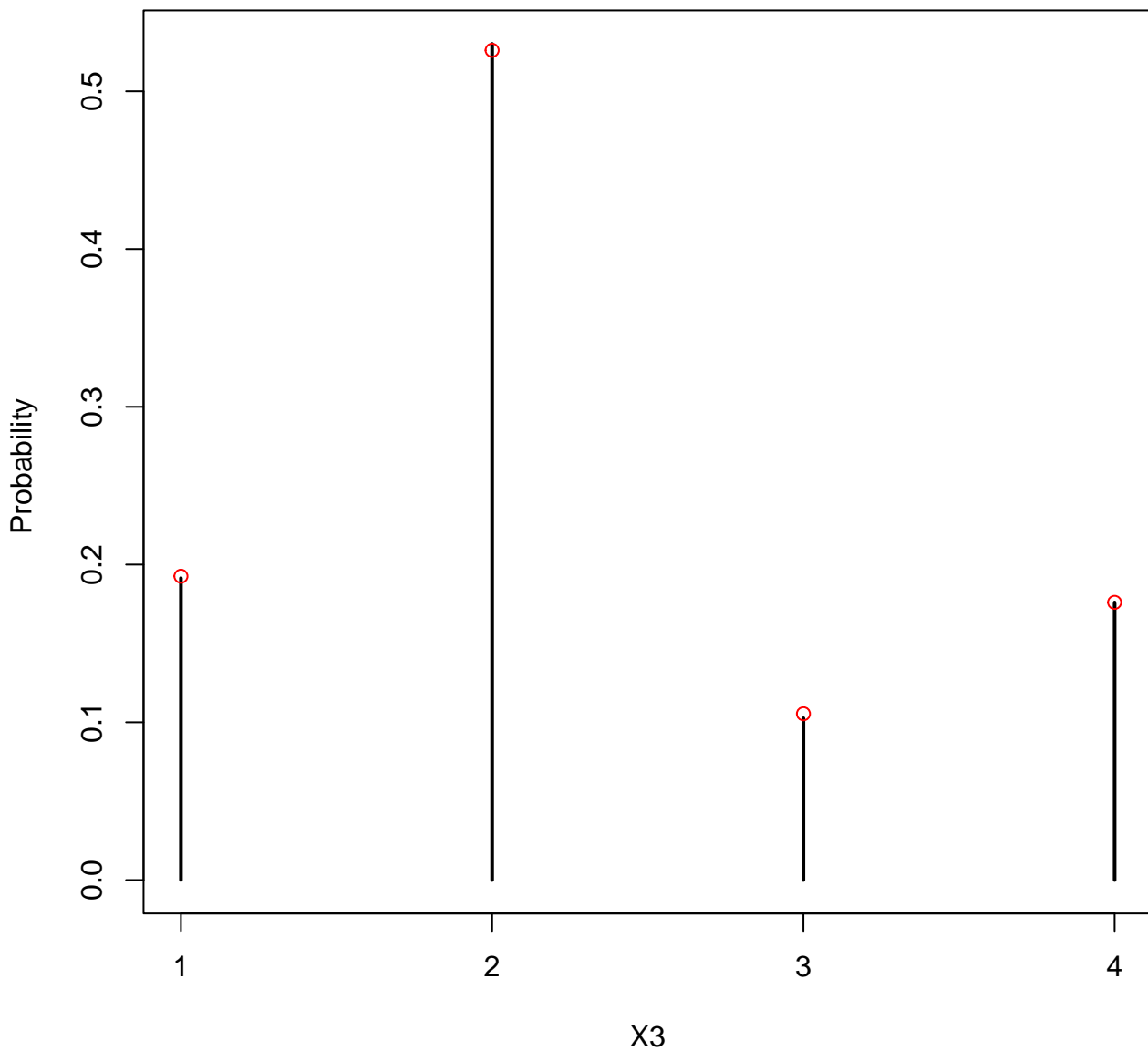
$h_I(X_2)$ Black: COLR, Red: Our Model



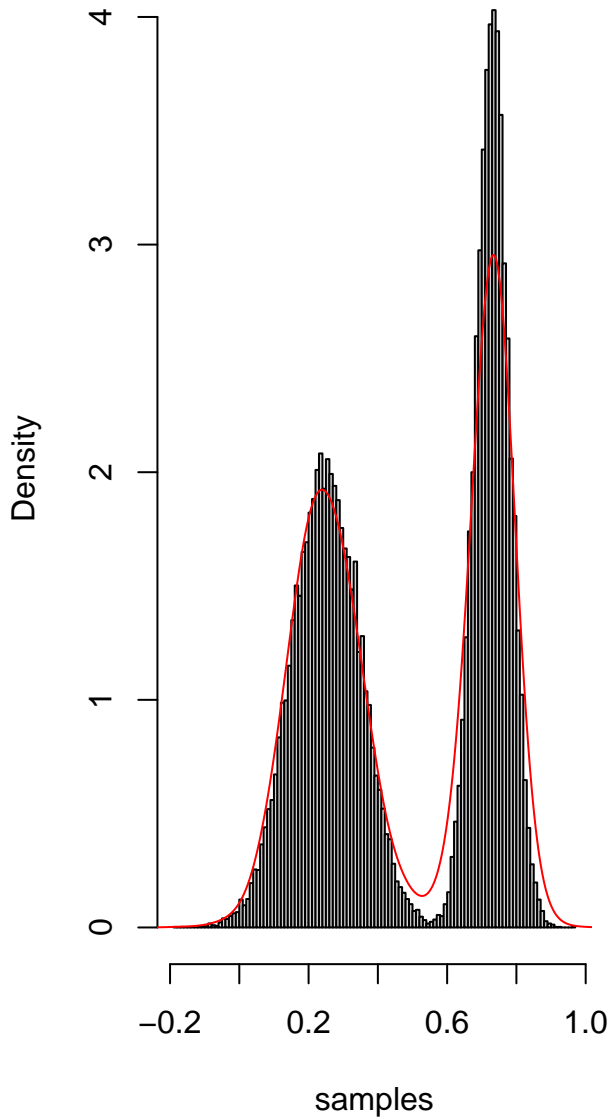
$h_l(X_3)$ Polr (blue) our Model (red)



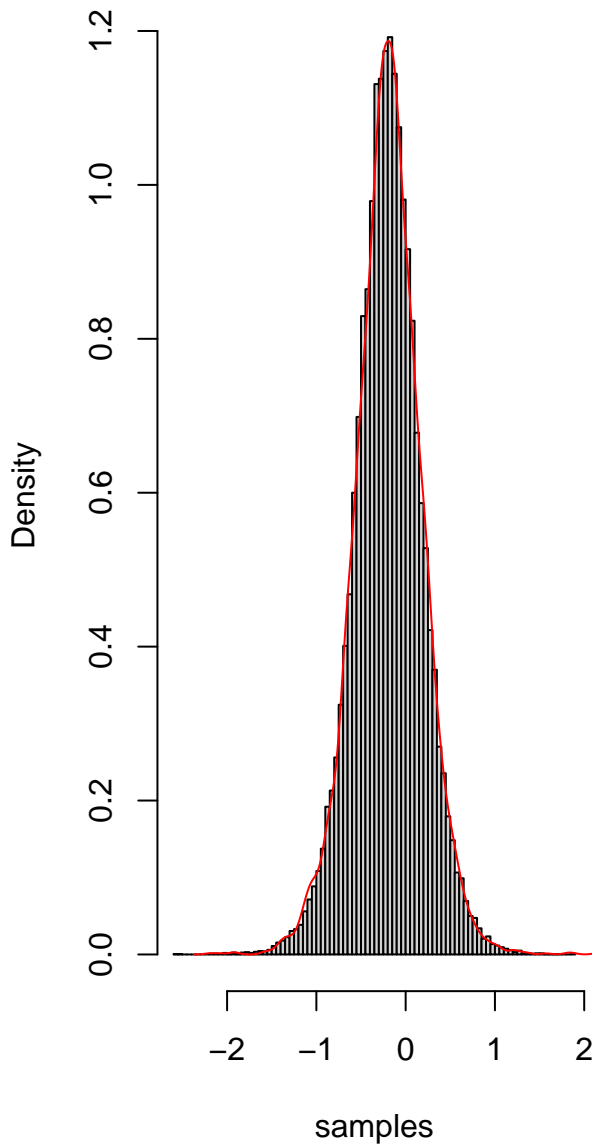
Black = Observations, Red samples from TRAM-DAG



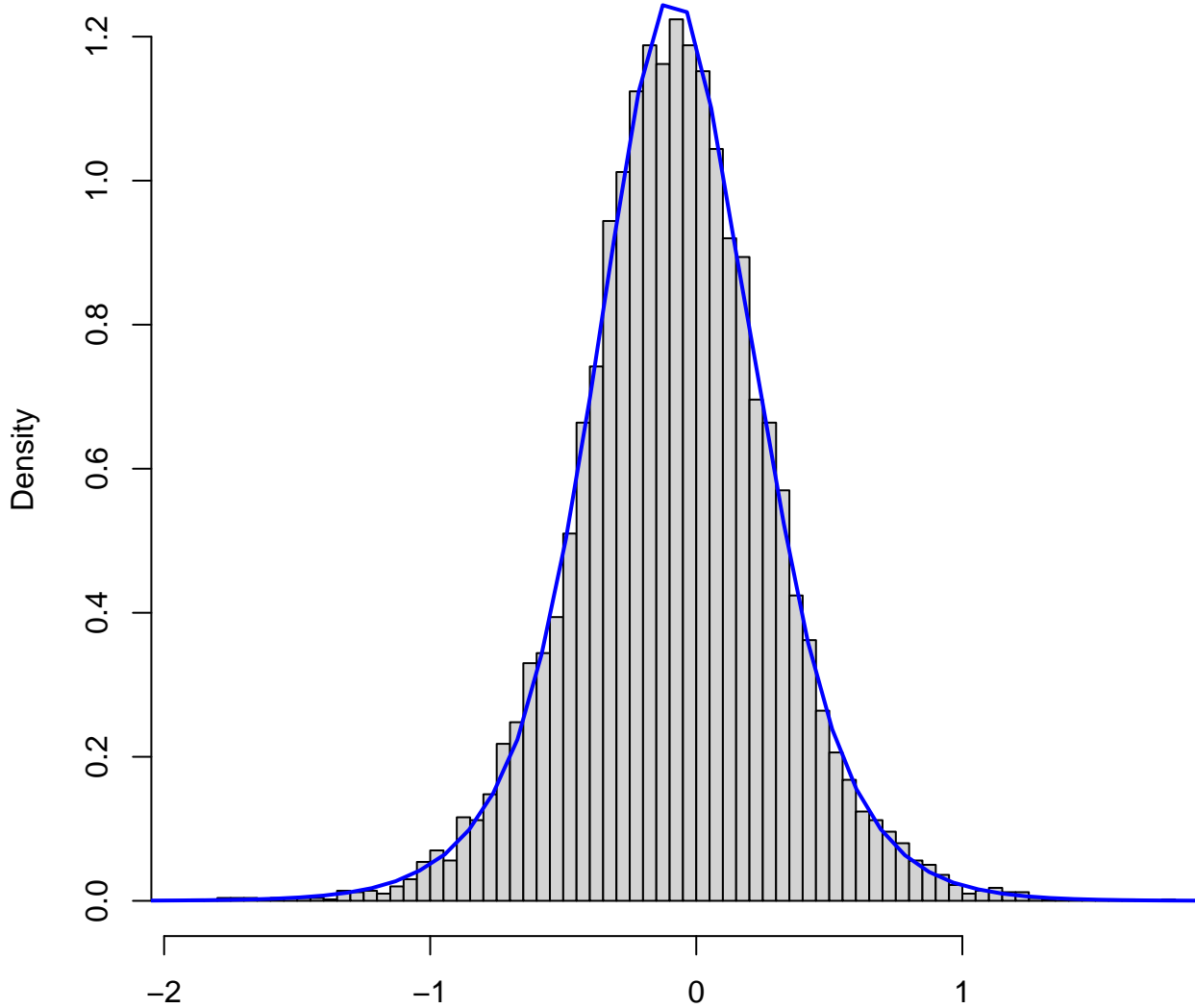
X1 red: ours, black: data



X2 red: ours, black: data

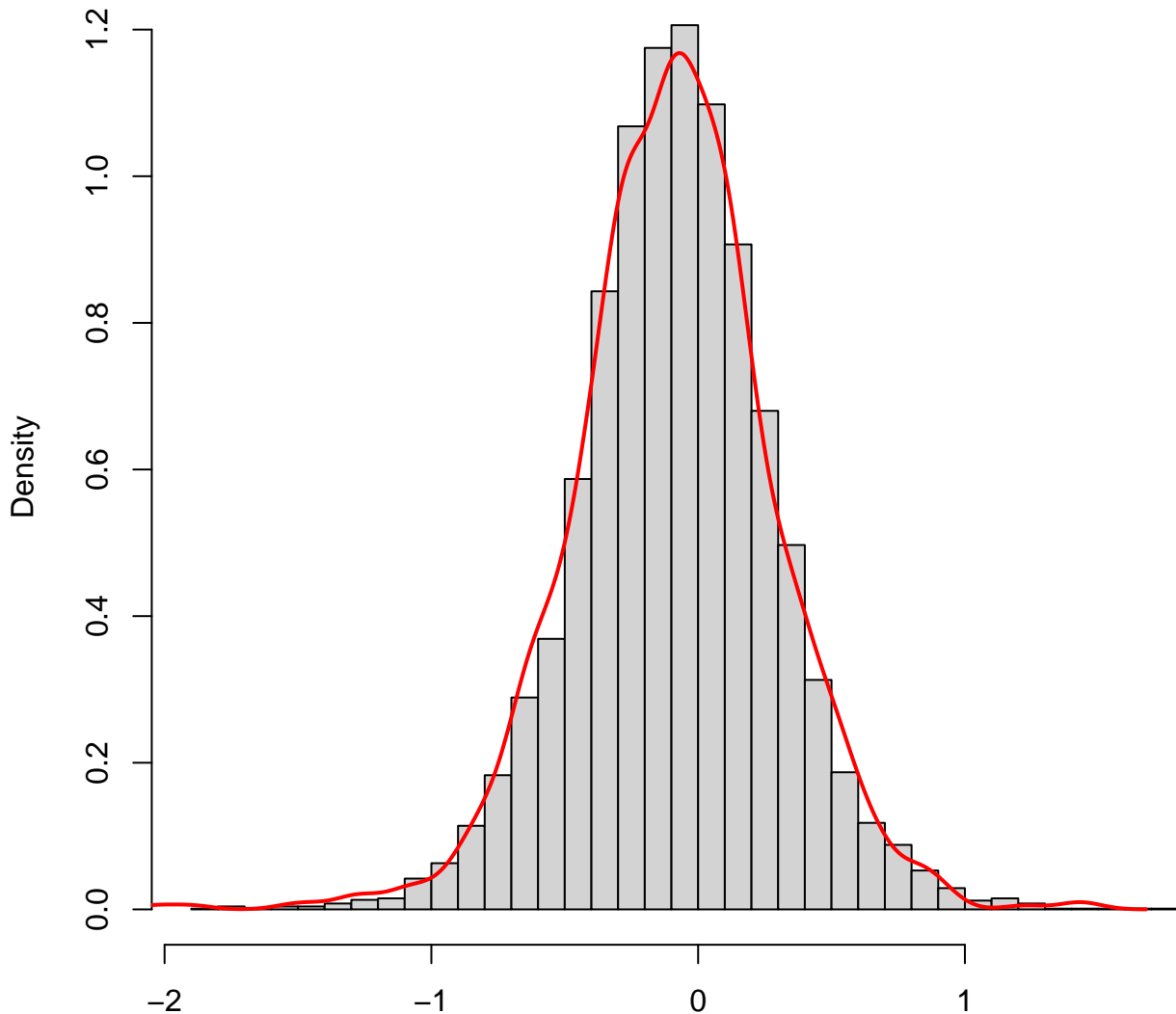


Do($X_1=0.2$) X_2



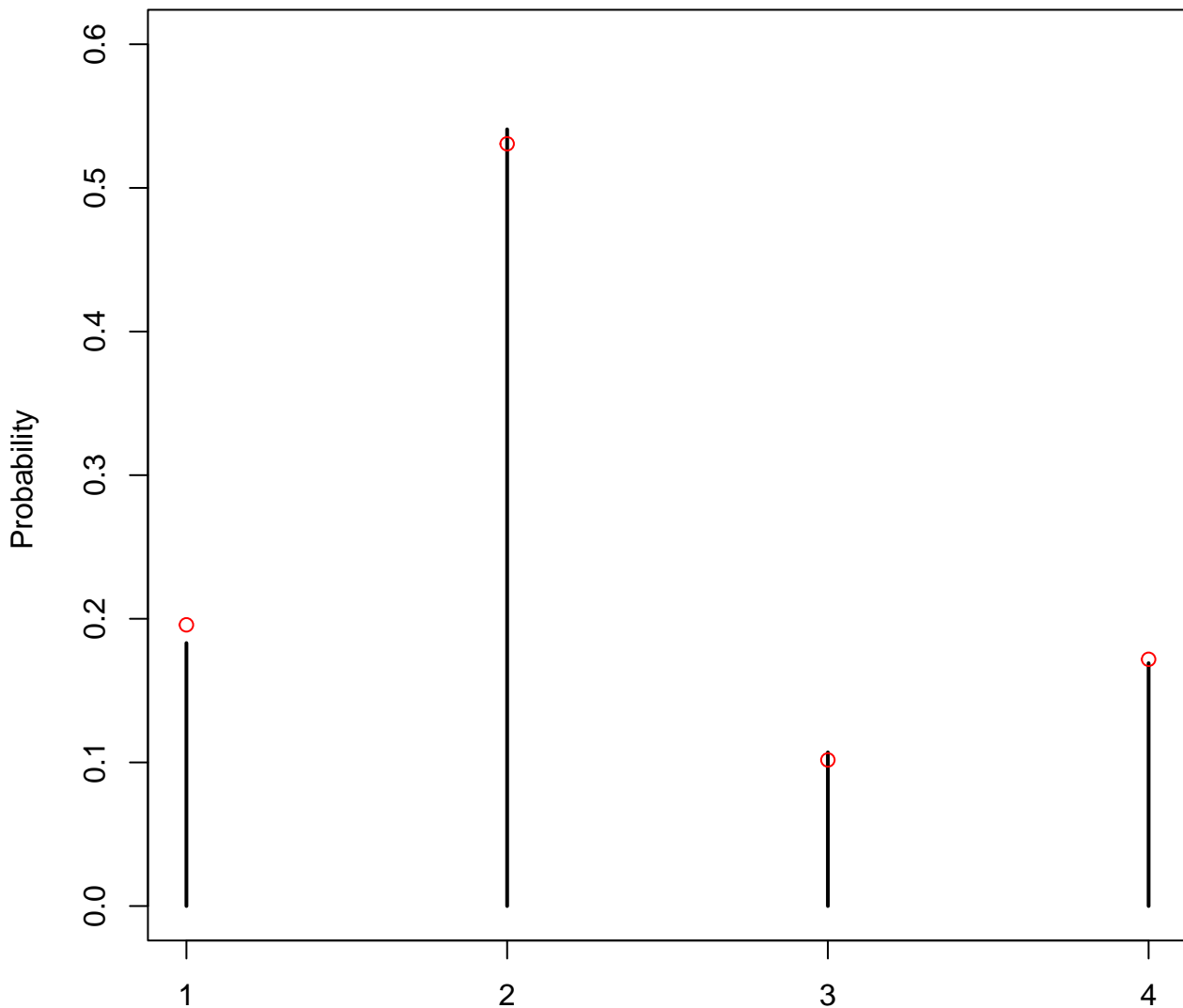
samples
Histogram from DGP with do. Blue: Colr

X2 | Do(X1=0.2)



samples
Histogram from DGP with do. red:TRAM_DAG

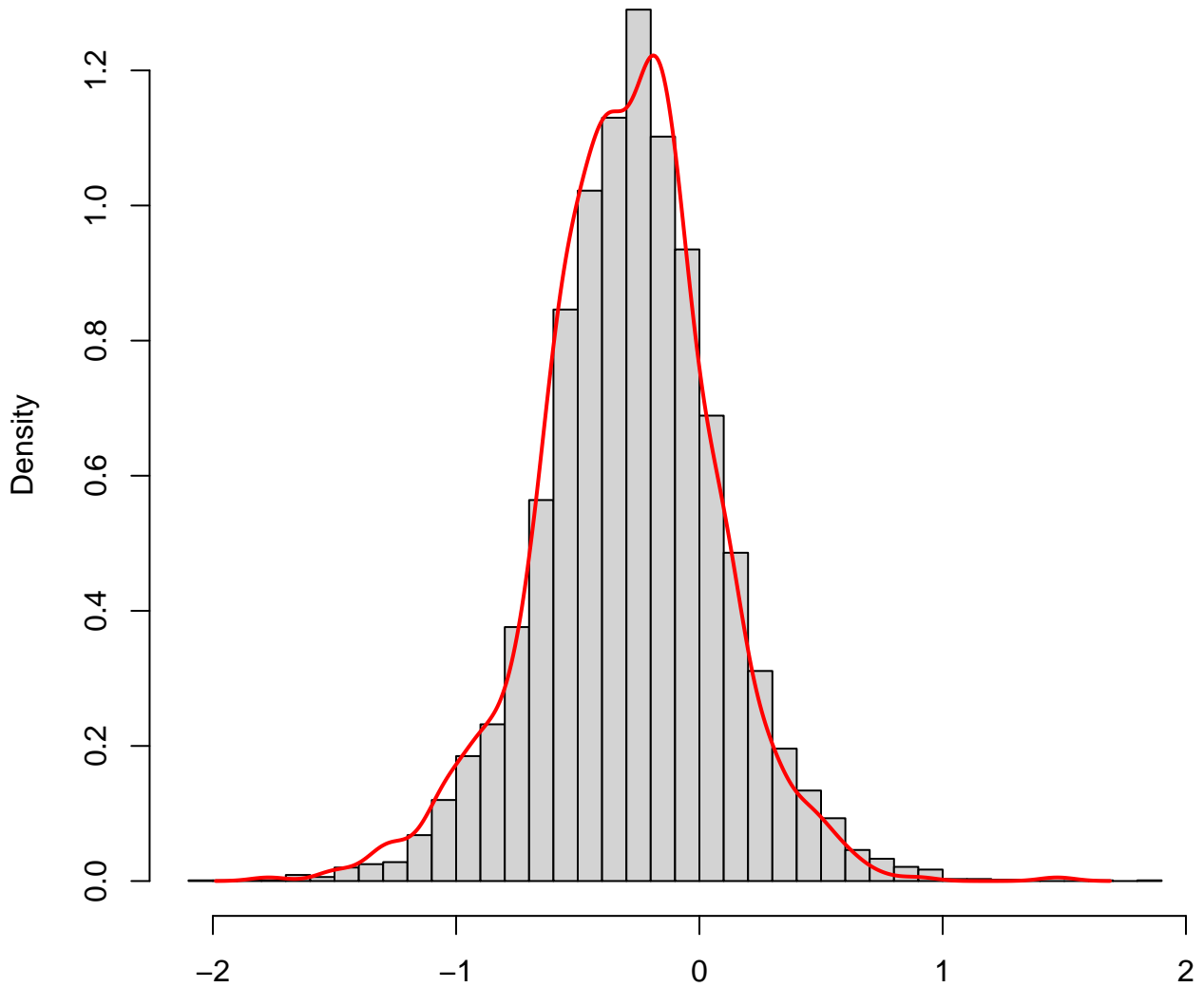
X3 | do(X1=0.2)



X_3

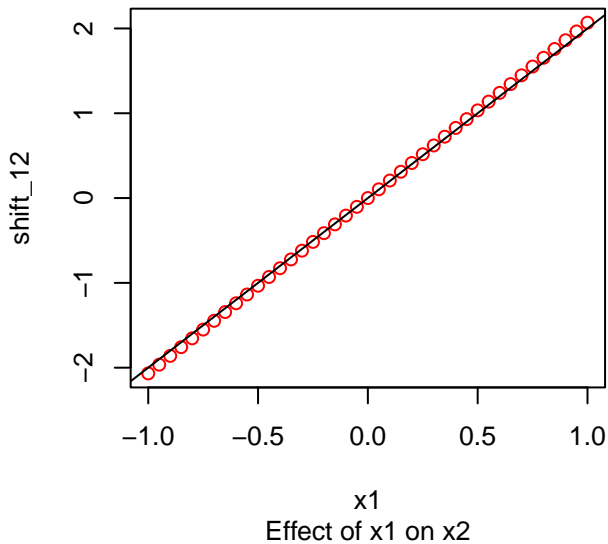
Black DGP with do. red:TRAM_DAG

X2 | Do(X1=0.7)

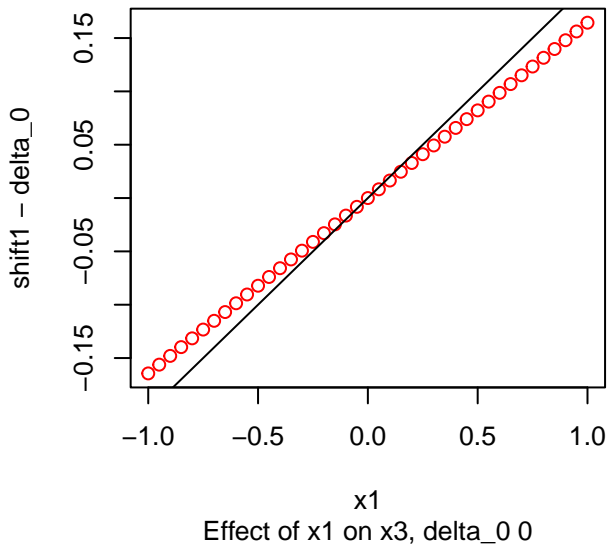


samples
Histogram from DGP with do. red:TRAM_DAG

LS-Term (black DGP, red Ours)



LS-Term (black DGP, red Ours)



CS-Term (black DGP $f_2(x)$, red Ours)

