MAR 30% missing - Probit

```
# sample MCAR dataset from PUMS
source("../../utils/sampleMAR30.R")
n = 3000
missing_col = c(1,3,7,9,10,11)
set.seed(2)

output_list <- sampleMAR30(n)
df <- output_list[['df']]
df_observed <- output_list[['df_observed']]

apply(is.na(df_observed), MARGIN = 2, mean)
```

```
##       VEH        MV        NP      RMSP       ENG     MARHT      SCHL    RACNUM
## 0.2926667 0.0000000 0.2816667 0.0000000 0.0000000 0.0000000 0.3213333 0.0000000
##      AGEP       WKL     PINCP
## 0.3330000 0.3100000 0.2890000
```

Ordinal bayesian nonparametric model

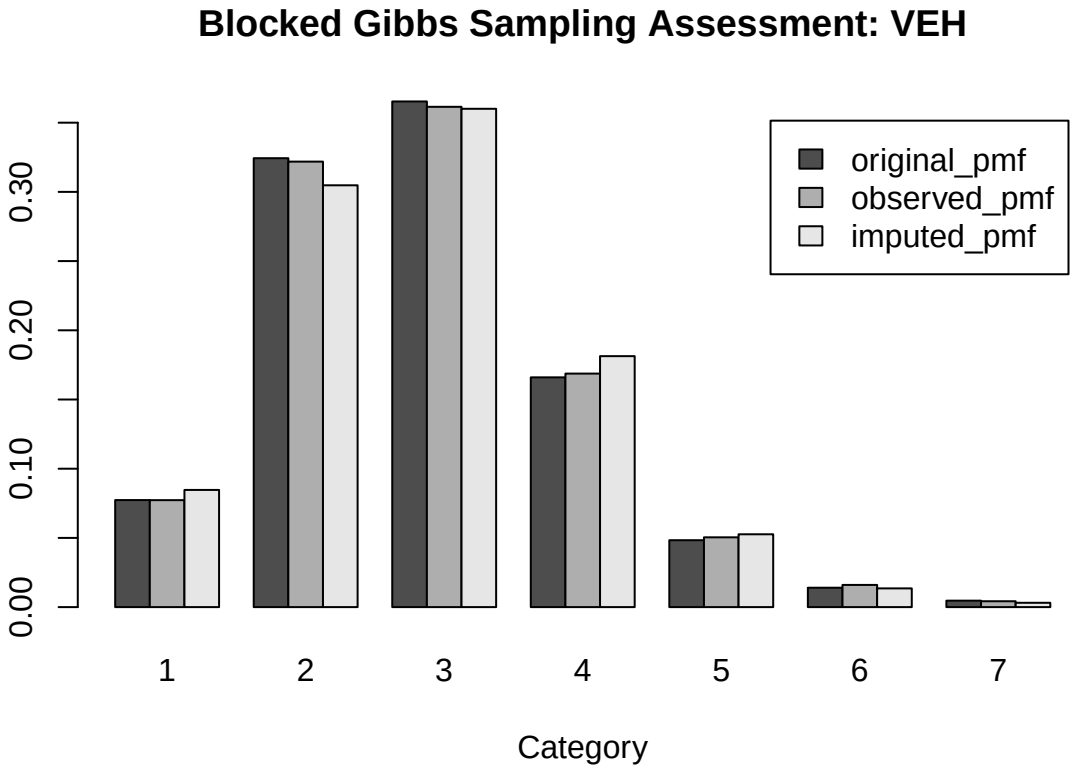
```
source("../../probitBayes.R")
N = 40
Mon = 300
B = 300
thin.int = 1
# function(y, N = 40, Mon = 2000, B = 300, thin.int = 5, seed = 0)
output_list <- probitBayesImputation(df_observed, N, Mon, B, thin.int)
```

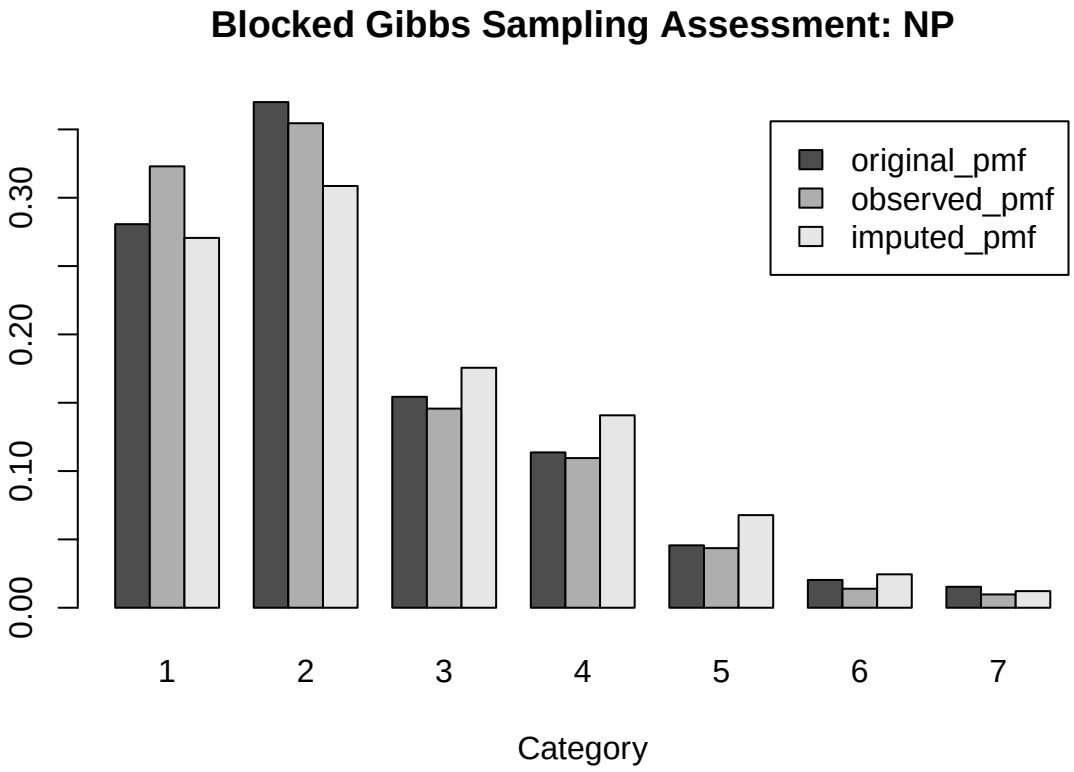
```
sampled_y <- output_list[['sampled_y']]
sampled_z <- output_list[['sampled_z']]
```

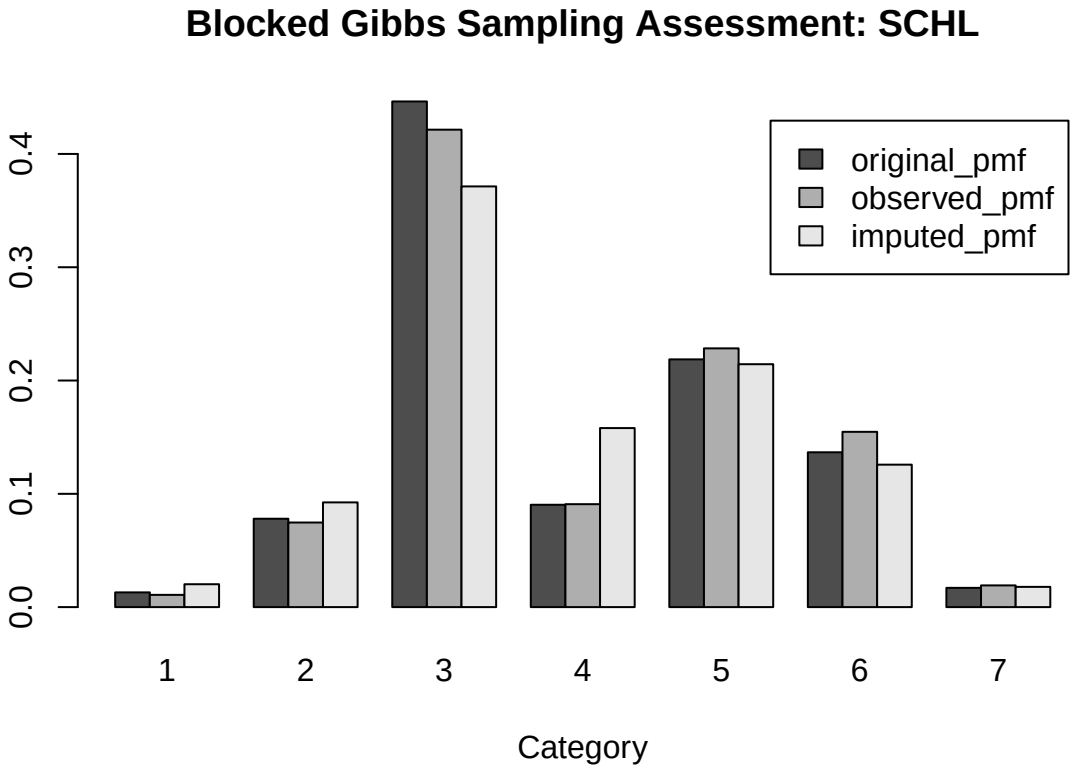
Diagnostics

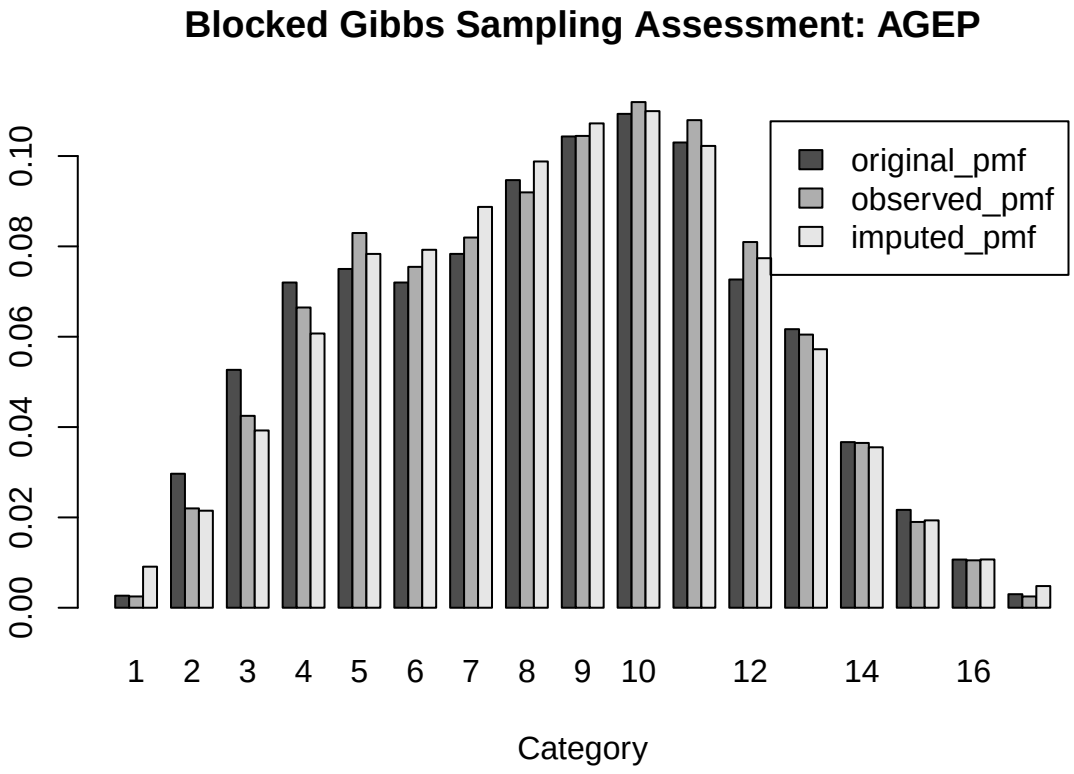
Assess bivariate joint distribution

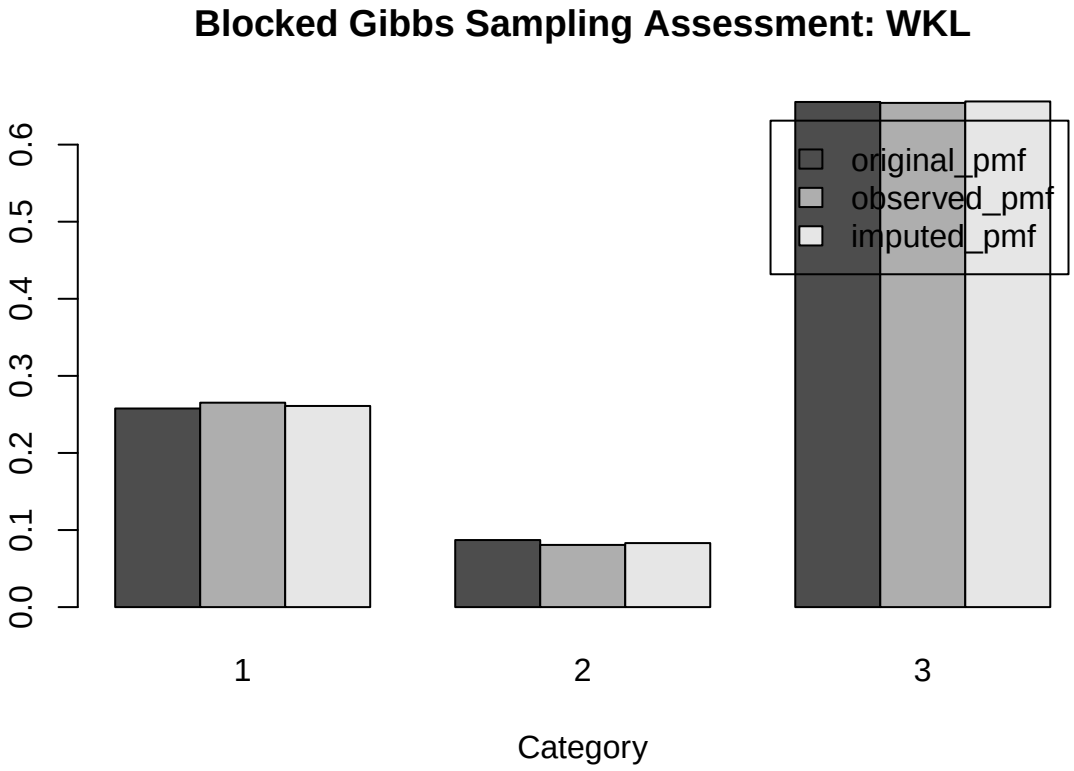
Assess trivariate joint distribution

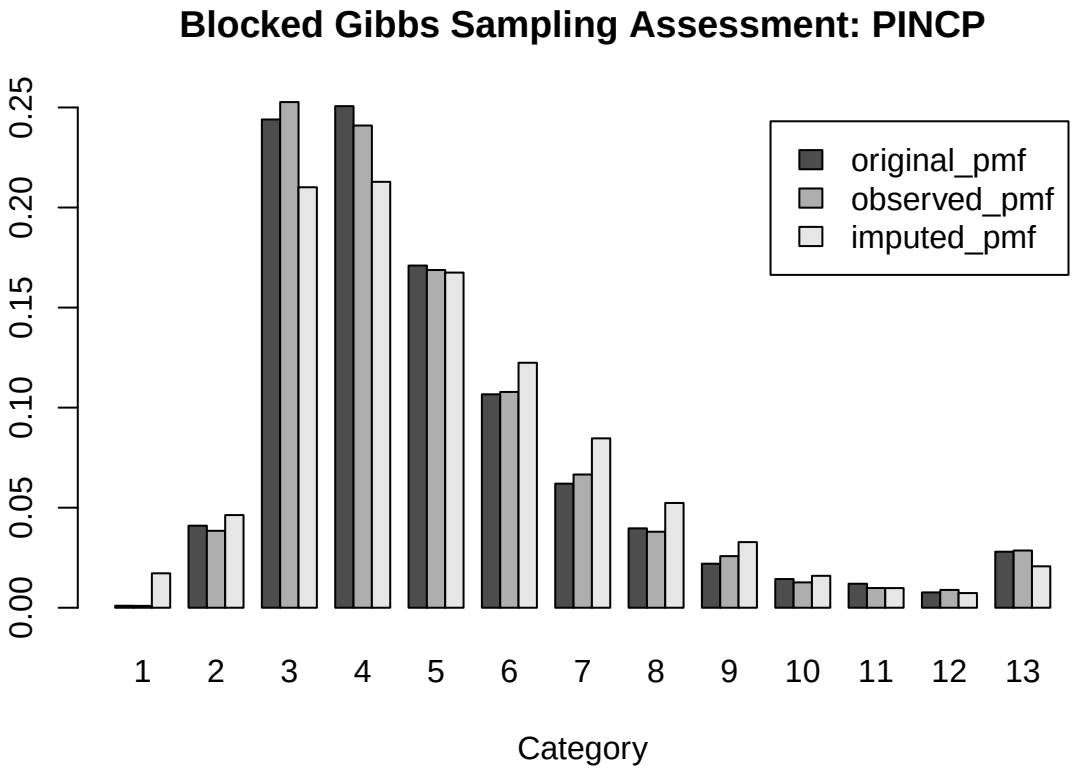

Blocked Gibbs Sampling Assessment: VEH
original_pmf
observed_pmf
imputed_pmf
0.00
0.10
0.20
0.30
1
2
3
4
5
6
7
Category

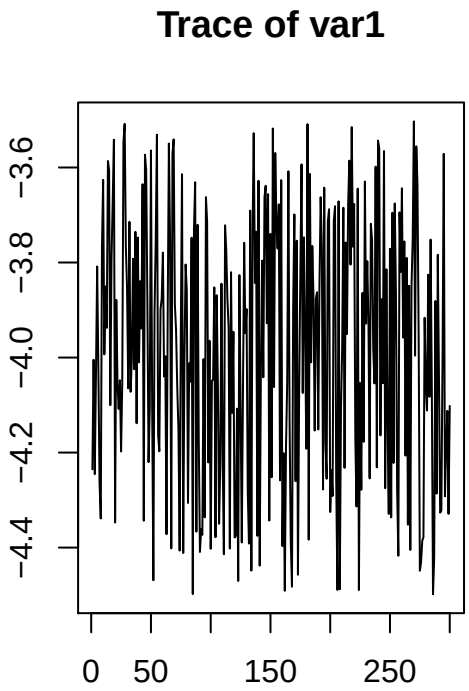

Blocked Gibbs Sampling Assessment: NP
original_pmf
observed_pmf
imputed_pmf
0.00
0.10
0.20
0.30
1
2
3
4
5
6
7
Category

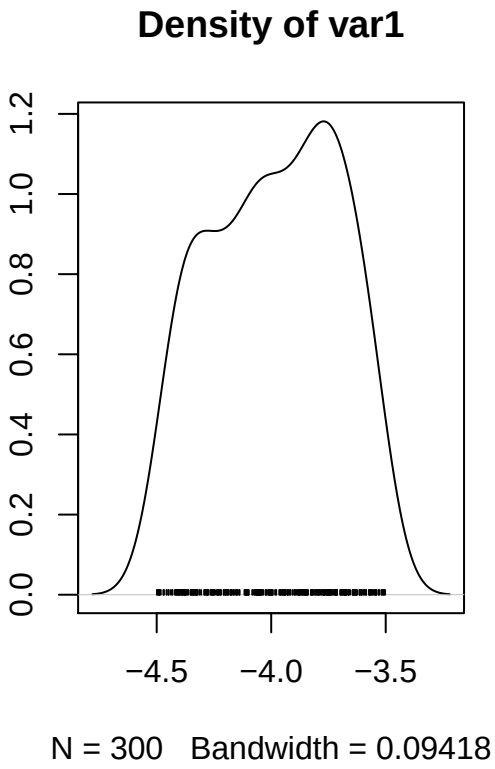

Blocked Gibbs Sampling Assessment: SCHL
original_pmf
observed_pmf
imputed_pmf
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
6
7
Category

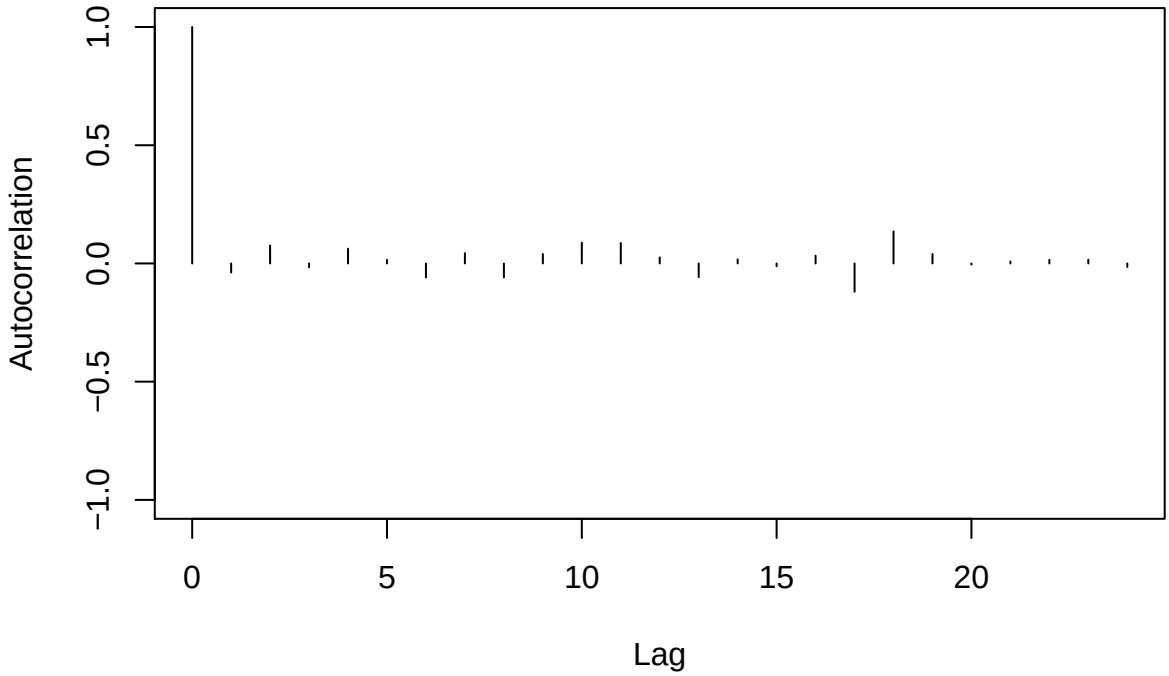

Blocked Gibbs Sampling Assessment: AGEP
original_pmf
observed_pmf
imputed_pmf
0.00
0.02
0.04
0.06
0.08
0.10
1
2
3
4
5
6
7
8
9
10
12
14
16
Category

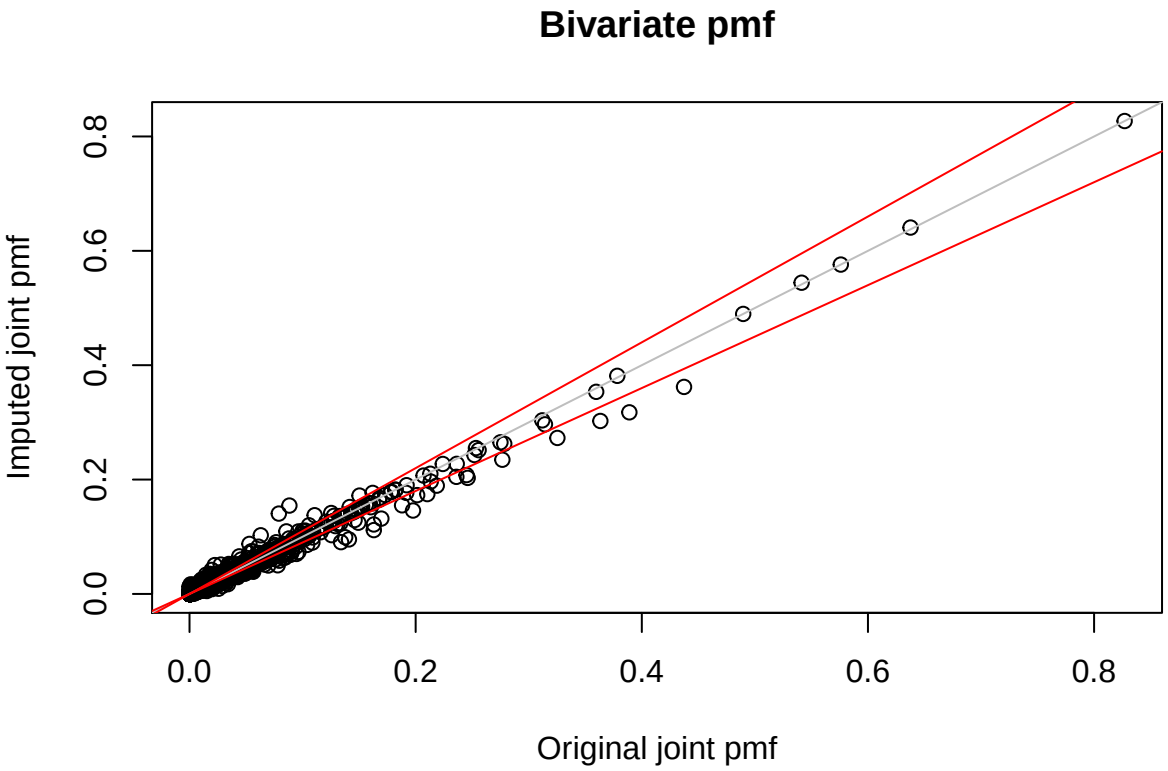

Blocked Gibbs Sampling Assessment: WKL
0.0
0.1
0.2
0.3
0.4
0.5
0.6
original_pmf
observed_pmf
imputed_pmf
1
2
3
Category

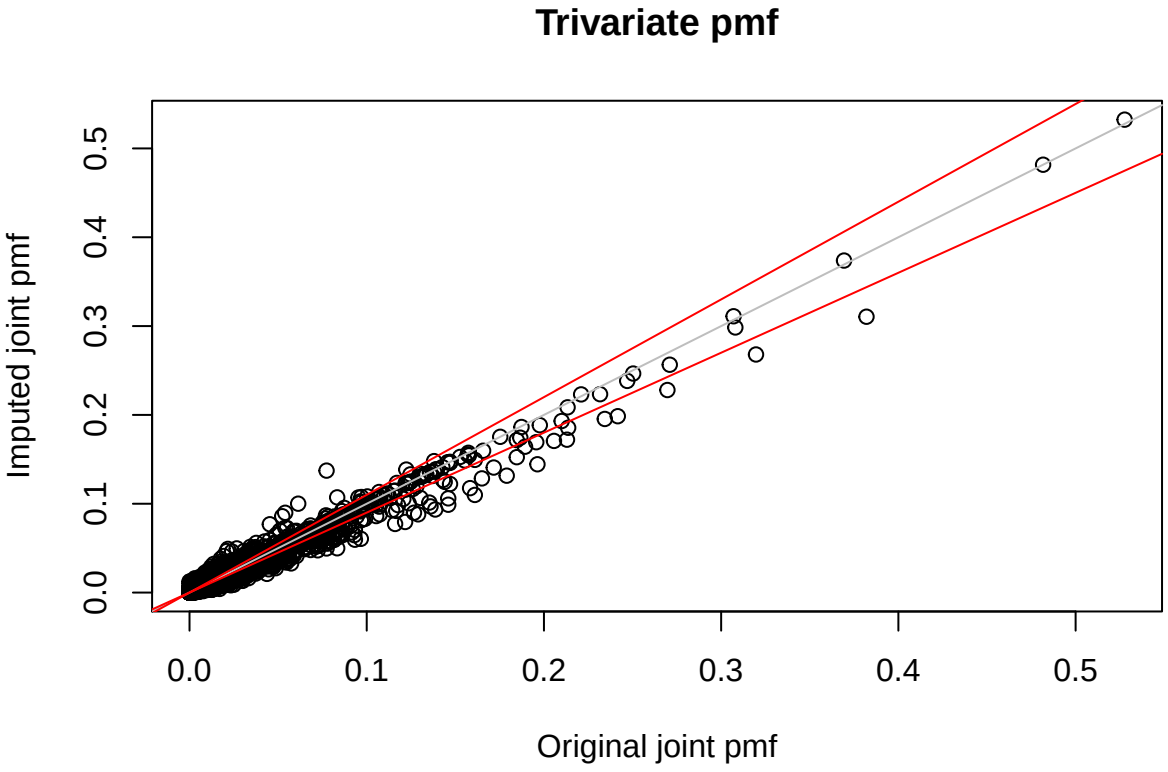

Blocked Gibbs Sampling Assessment: PINCP
0.00
0.05
0.10
0.15
0.20
0.25
original_pmf
observed_pmf
imputed_pmf
1
2
3
4
5
6
7
8
9
10
11
12
13
Category

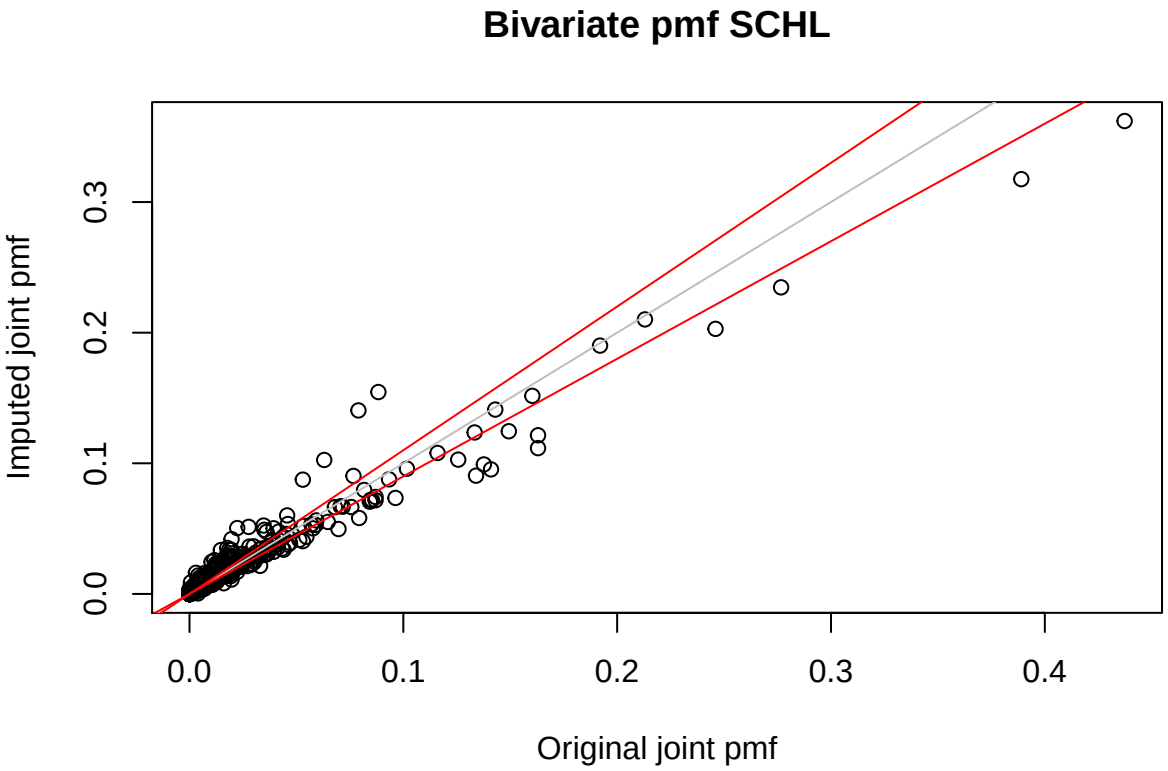

Trace of var1
−4.4
−4.2
−4.0
−3.8
−3.6
0
50
150
250
Iterations

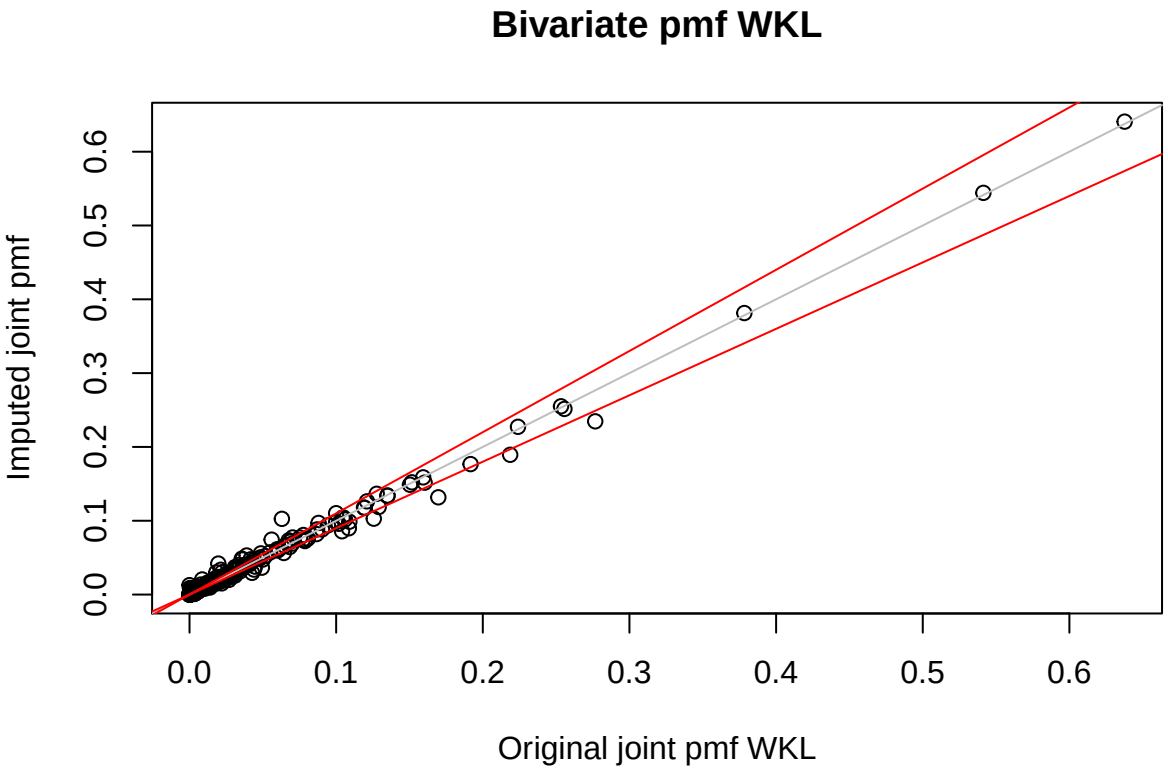

Density of var1
0.0
0.2
0.4
0.6
0.8
1.0
1.2
−4.5
−4.0
−3.5
N = 300 Bandwidth = 0.09418


Autocorrelation
−1.0
−0.5
0.0
0.5
1.0
0
5
10
15
20
Lag


Bivariate pmf
Imputed joint pmf
0.0
0.2
0.4
0.6
0.8
Original joint pmf


Trivariate pmf
Imputed joint pmf
0.0
0.1
0.2
0.3
0.4
0.5
Original joint pmf


Bivariate pmf SCHL
Imputed joint pmf
0.0
0.1
0.2
0.3
0.0
0.1
0.2
0.3
0.4
Original joint pmf


Bivariate pmf WKL
Imputed joint pmf
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Original joint pmf WKL