Monte-Carlo simulations - Conditional Poisson Distribution

March 26, 2019

# Contents

1 Description

To investigate small-sample performance of the bias correction, we carried out Monte-Carlo simulations. To this purpose, we generate datasets according to description of the model. A dataset $\boldsymbol{X}^k$ constitutes of N samples where every sample is a 0-1 vector which carries a randomly generated MOI from conditional Poisson distribution. The sample sizes N are chosen in an increasing fashion in order to see the improvement in performance in an asymptotic manner. For each combination of values for set of model parameters, $S = 10,000$ datasets are generated. Afterwards, results for each dataset is derived, i.e., MLE, TBCMLE (theoretical bias-corrected MLE) and HBCMLE (heuristic bias-corrected MLE). Finally, considering the empirical distribution of these statistics, the mean and variance of each statistic from $10,000$ datasets are calculated.
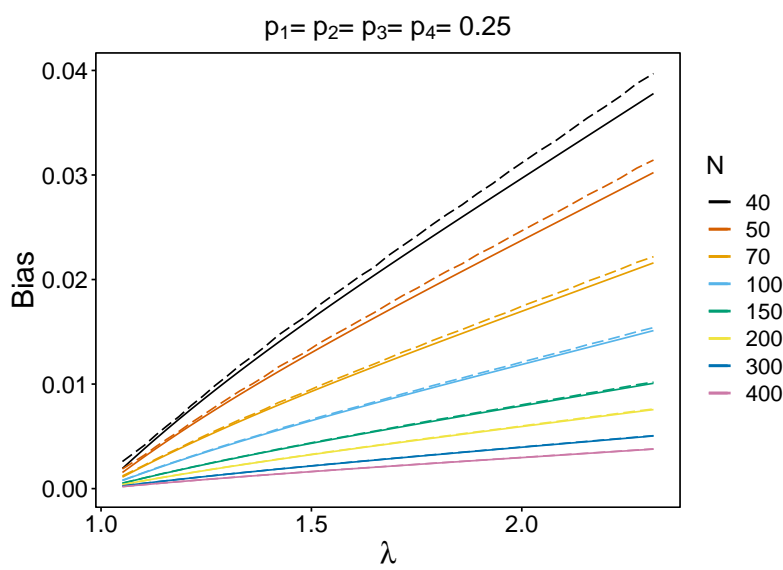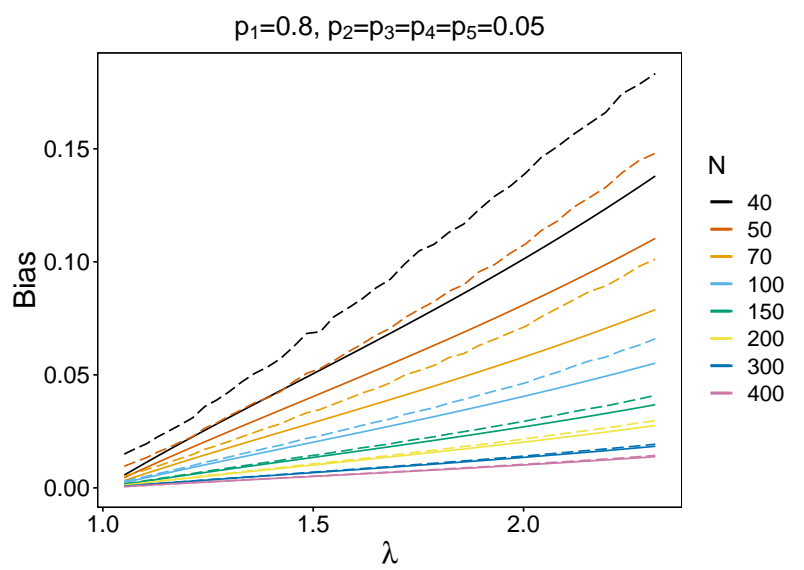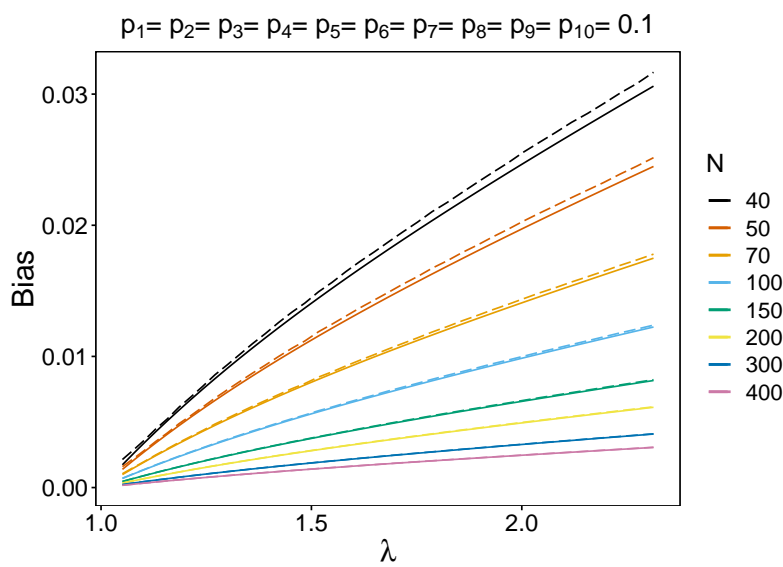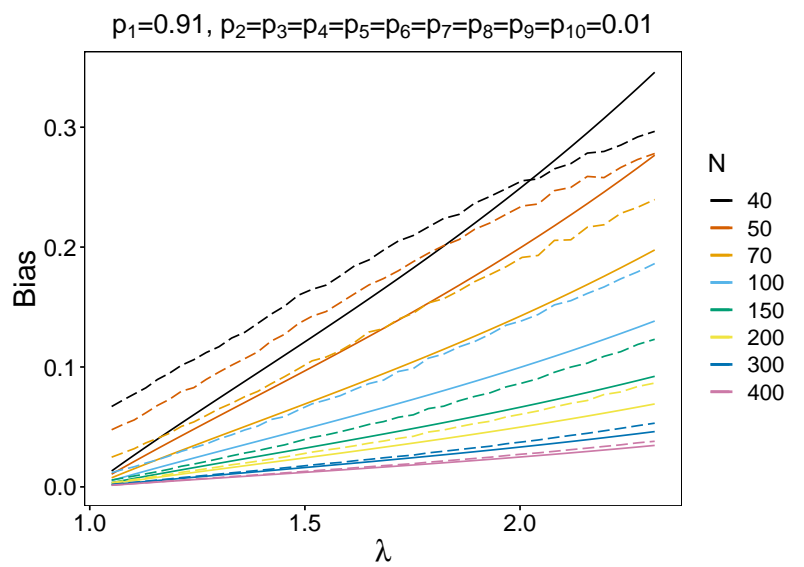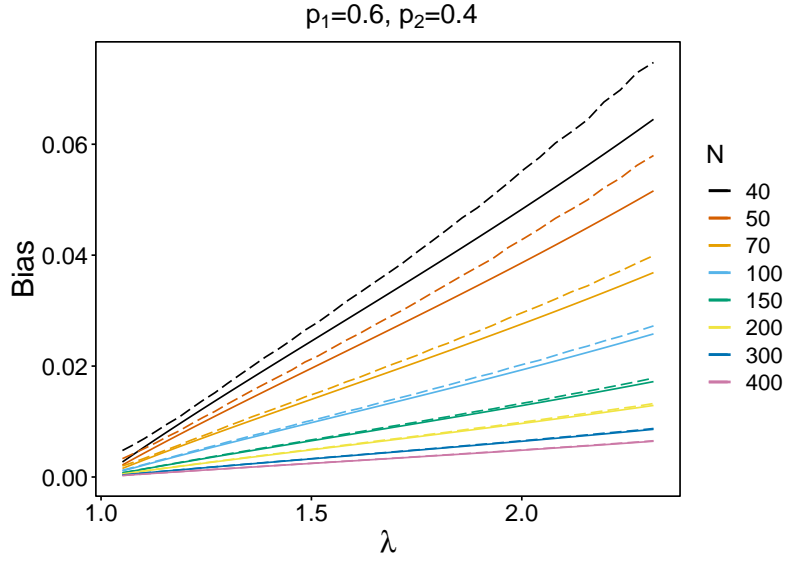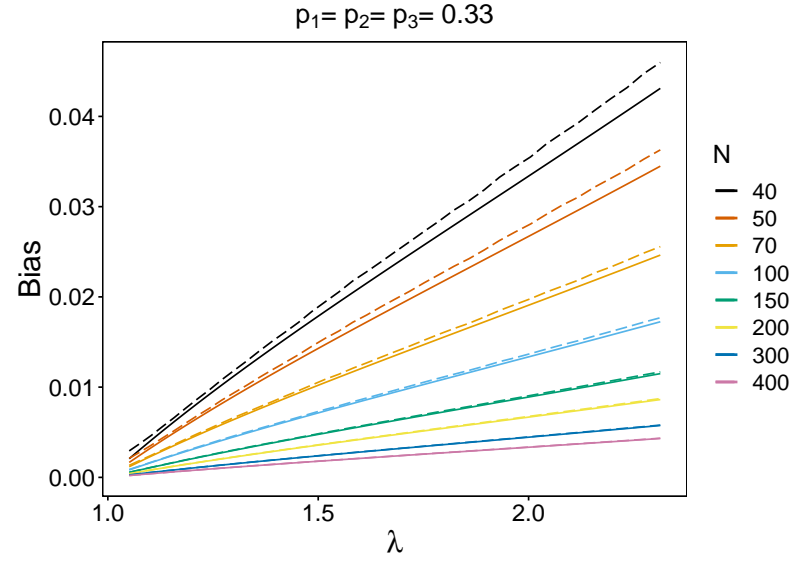
2 Bias function
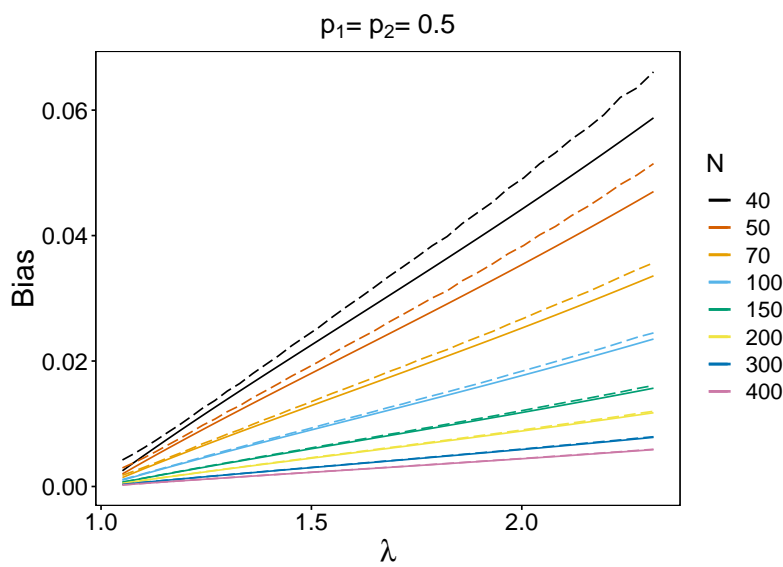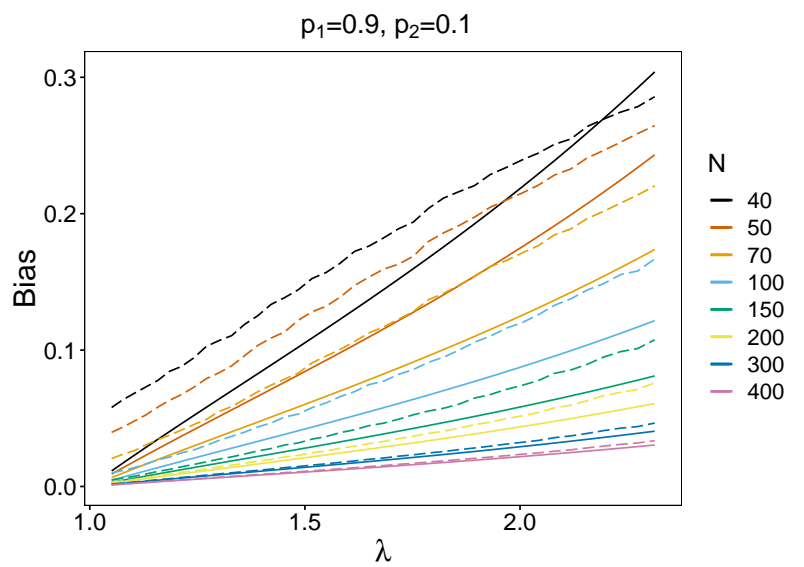
Figure 1: **Bias function**. The figure shows the behavior of average approximation of theoretical bias from $S = 10,000$ simulated datasets (solid lines) versus the theoretical bias at the true parameter (long-dashed lines). Different sample sizes are specified with different colors.
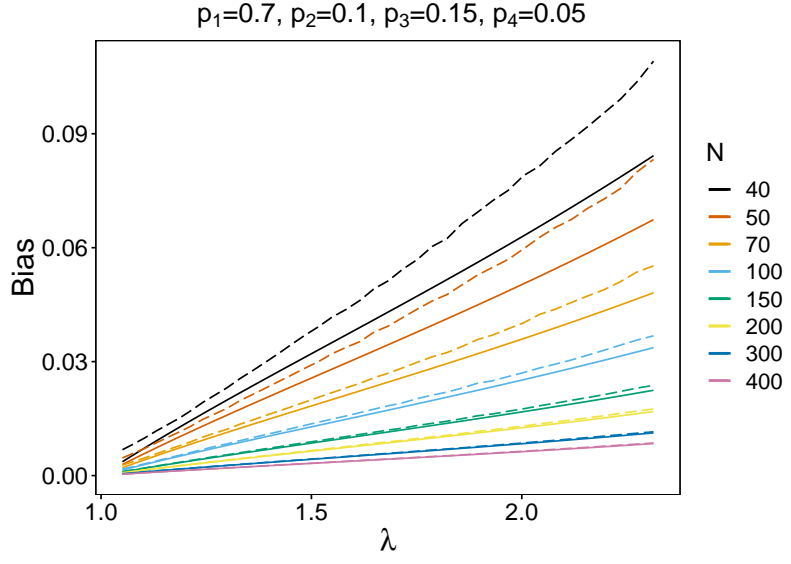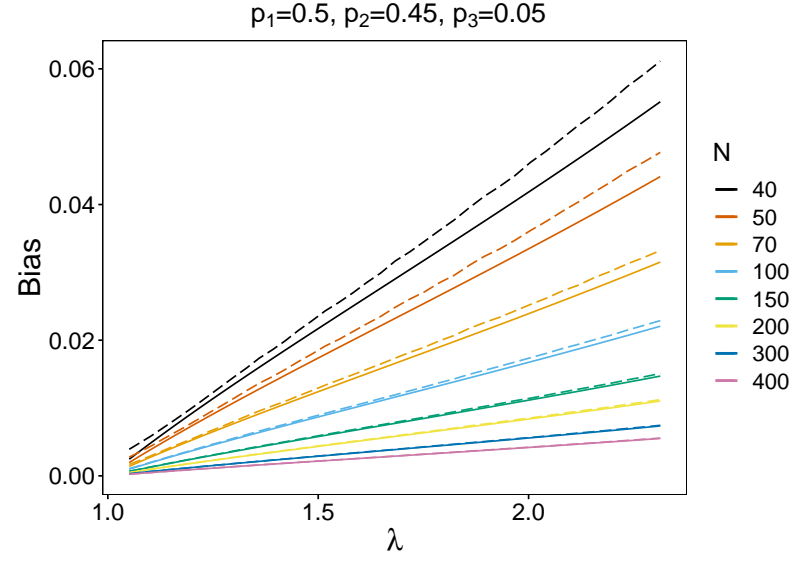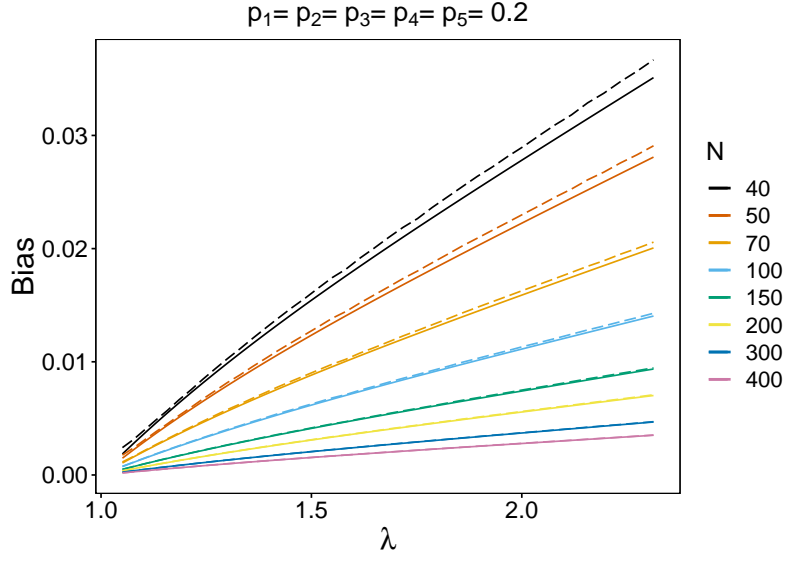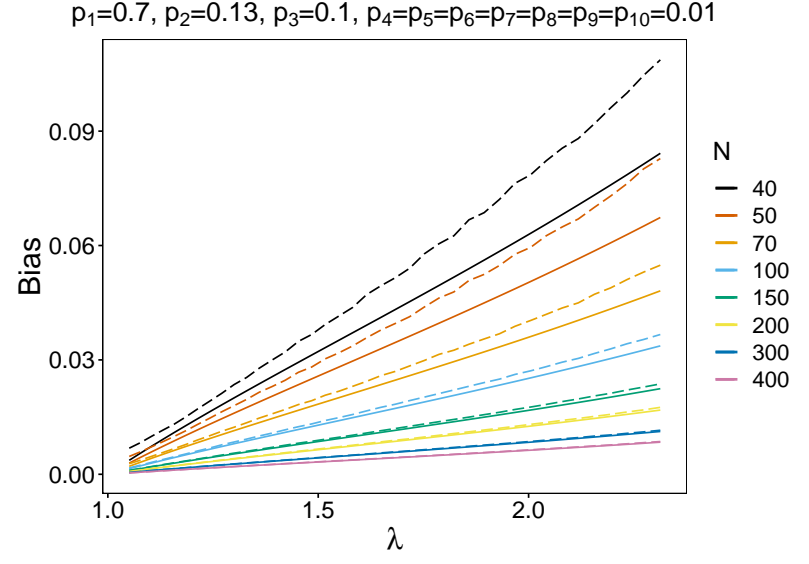
A $p_1=0.6, p_2=0.4$

B $p_1= p_2= p_3= 0.33$

C $p_1=0.7, p_2=0.1, p_3=0.15, p_4=0.05$

D $p_1=0.5, p_2=0.45, p_3=0.05$

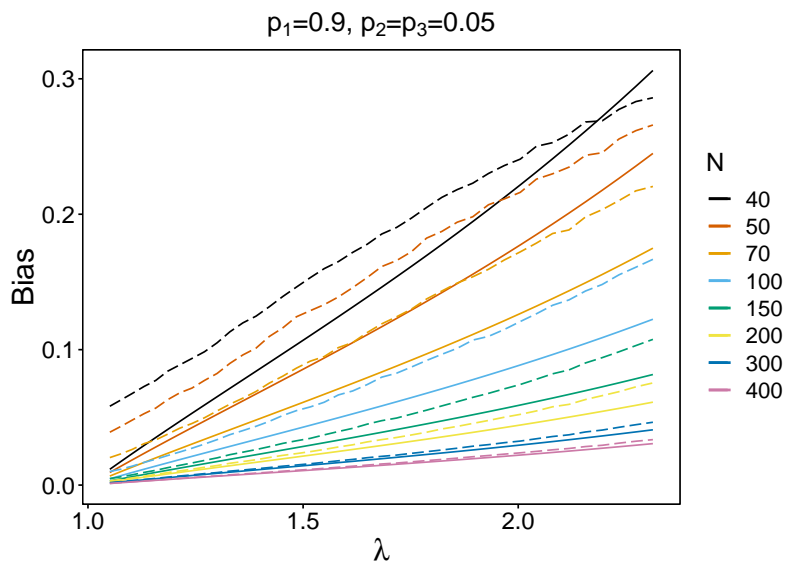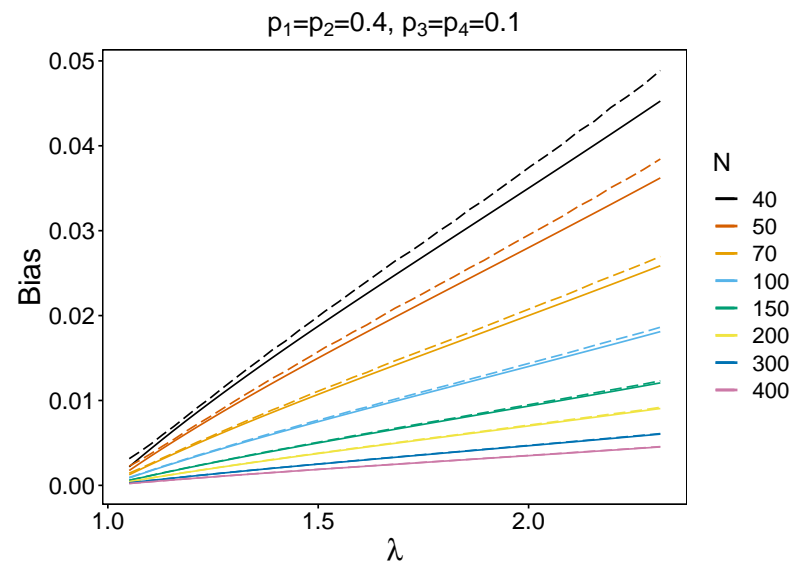E $p_1= p_2= p_3= p_4= p_5= 0.2$

F $p_1=0.7, p_2=0.13, p_3=0.1, p_4=p_5=p_6=p_7=p_8=p_9=p_{10}=0.01$
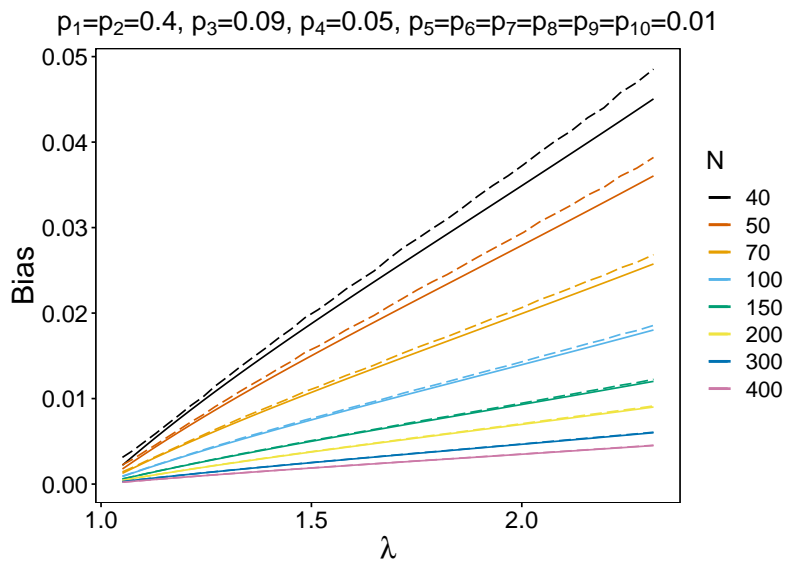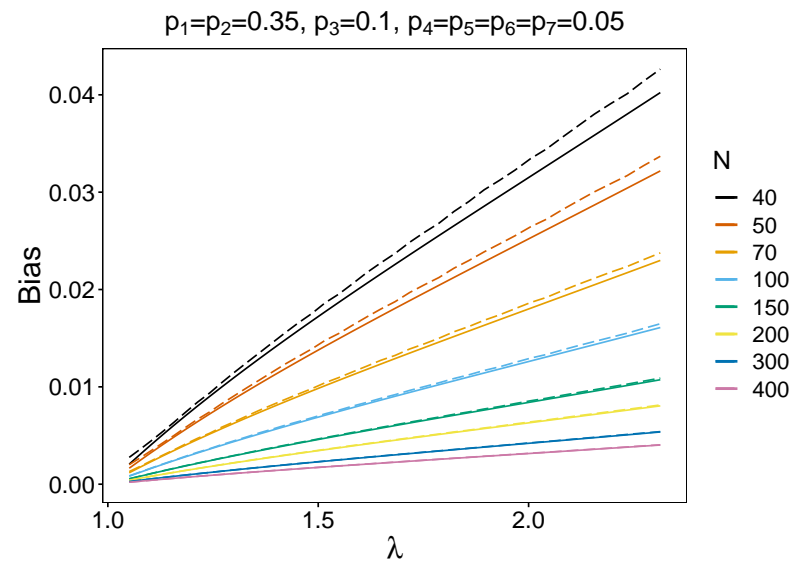
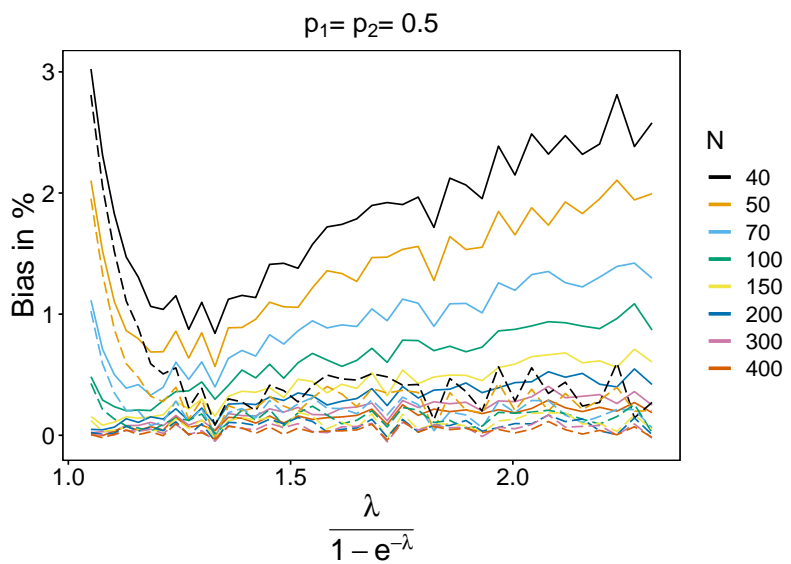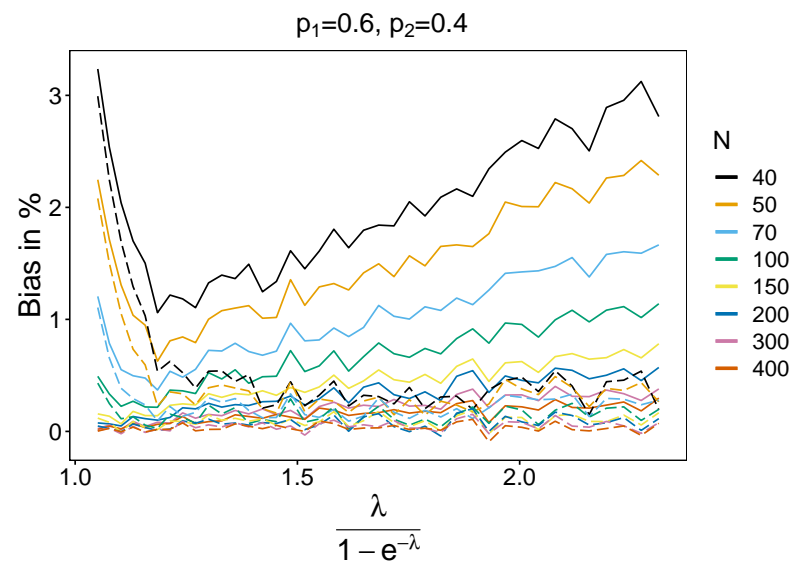Figure 2: **Bias function**. Same as Figure 1

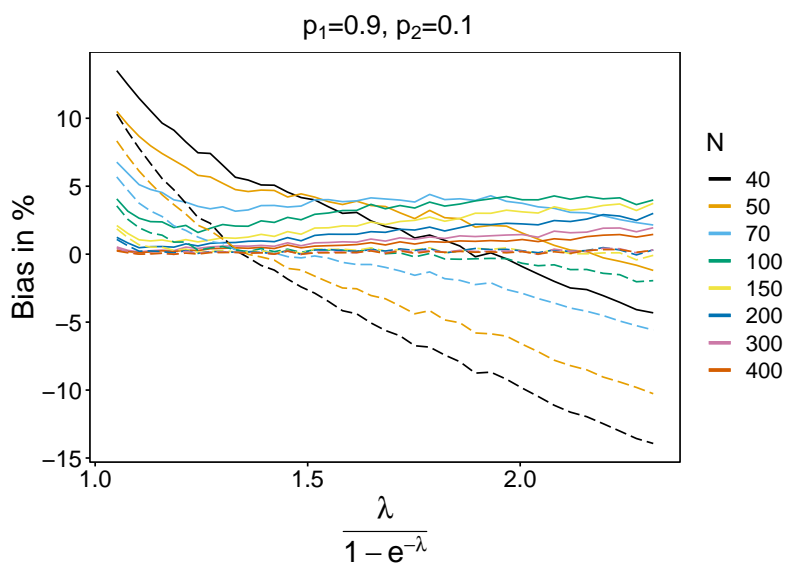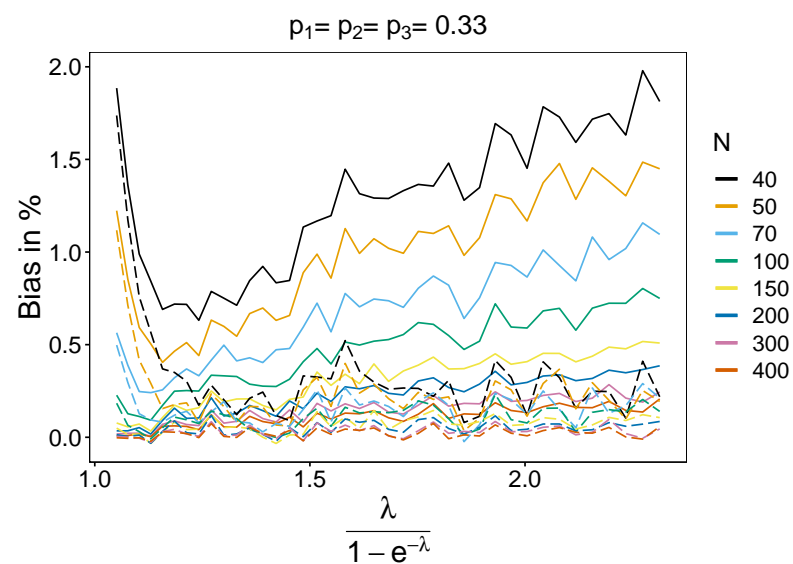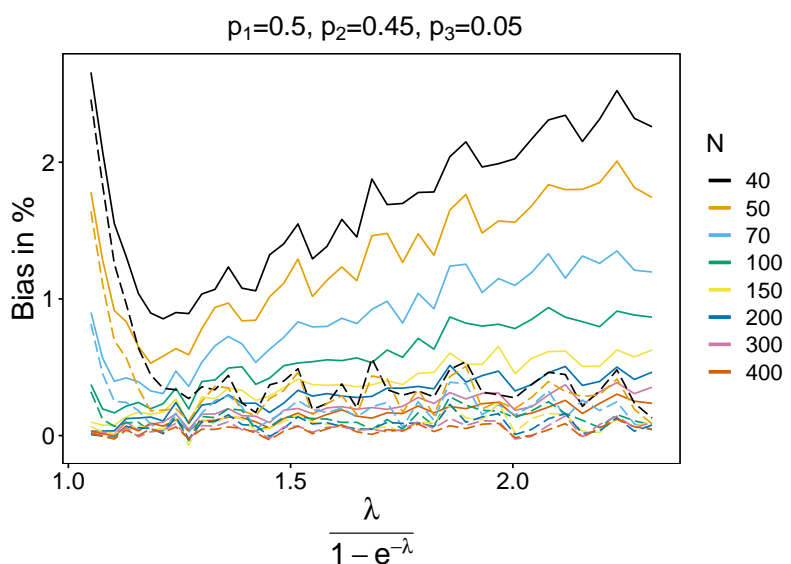Figure 3: **Bias function**. Same as Figure 1
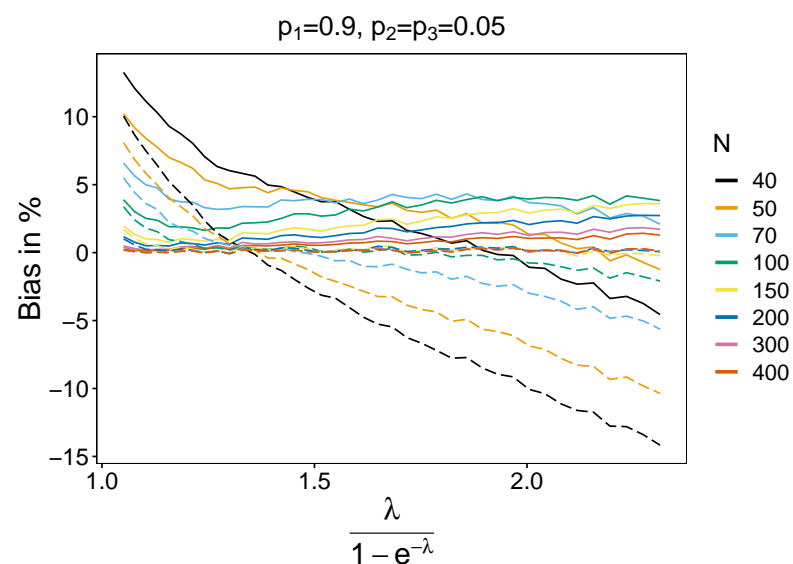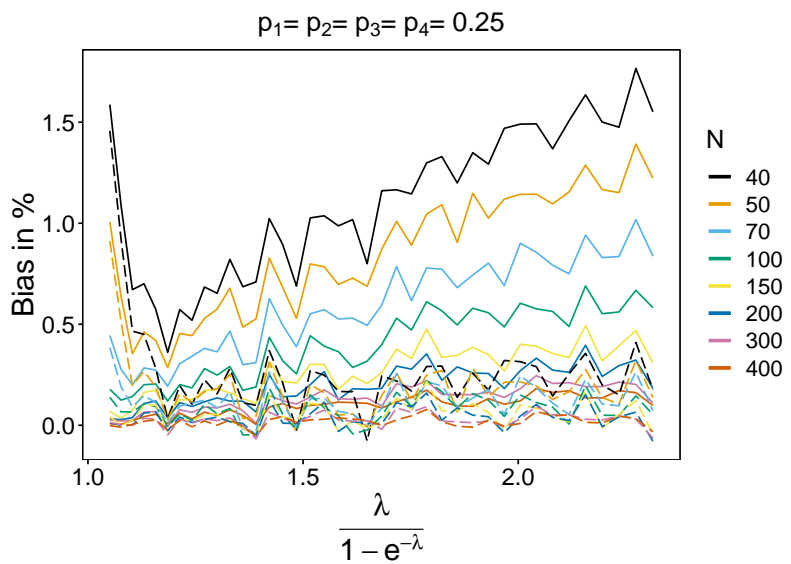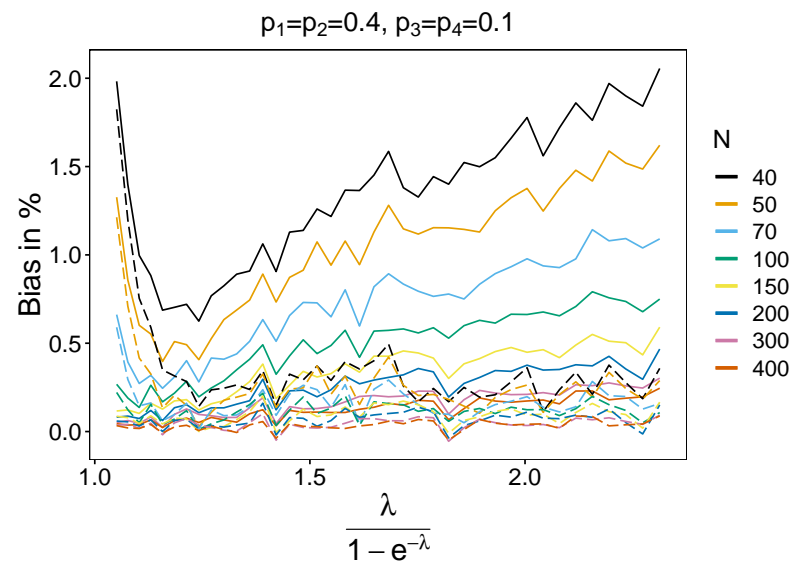
3 Bias of $\hat{\lambda}$

Figure 4: **Bias in %**-The figure shows the bias in % of different versions of bias-corrected MLE of mean MOI parameter along with MLE of $\psi = \frac{\lambda}{1-e^{-\lambda}}$ for different lineage frequency distributions. Different colors correspond to different sample sizes N. The solid, long-dashed and dot-dashed lines correspond to MLE, TBCMLE (theoretical bias-corrected MLE) and HBCMLE (heuristic bias-corrected MLE), respectively.
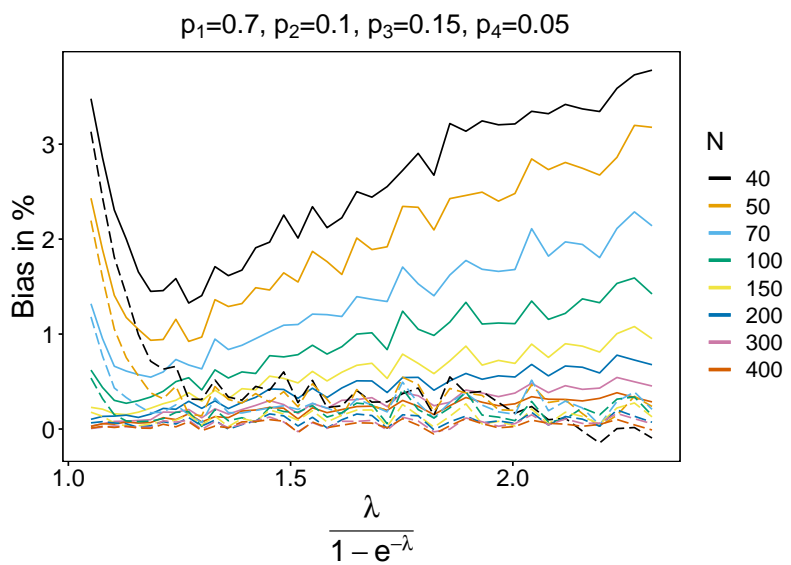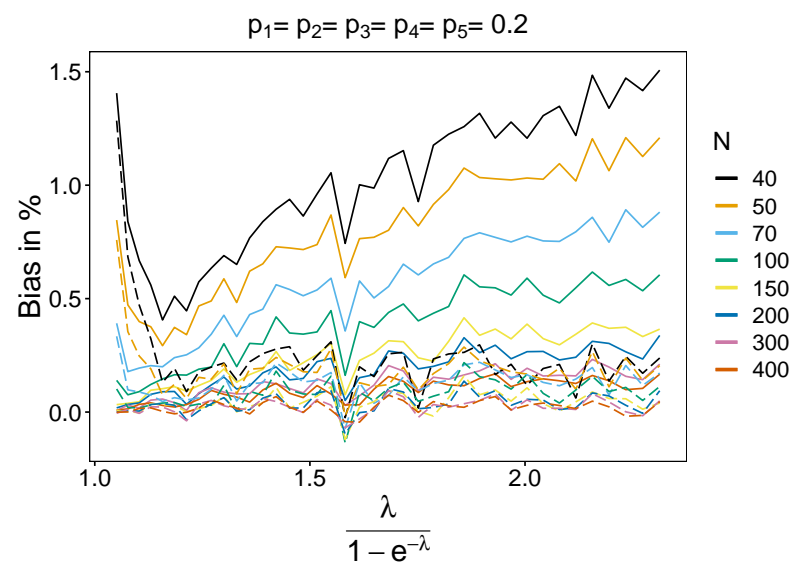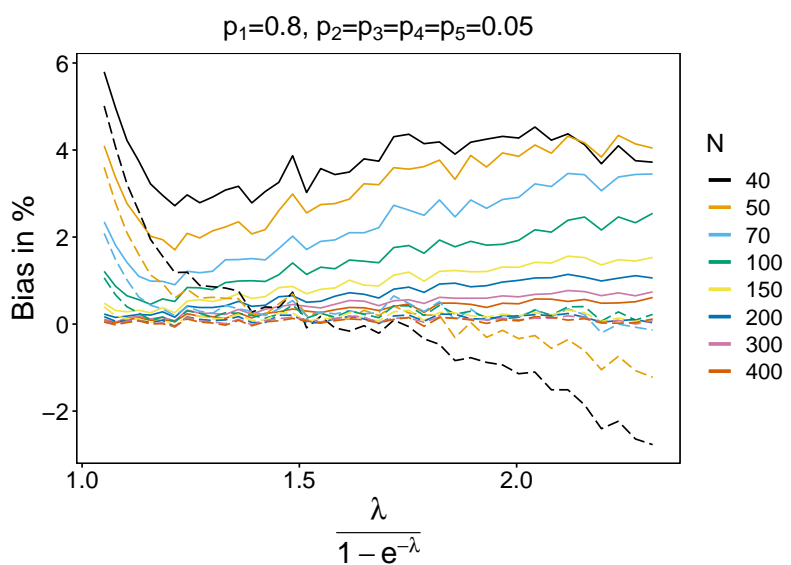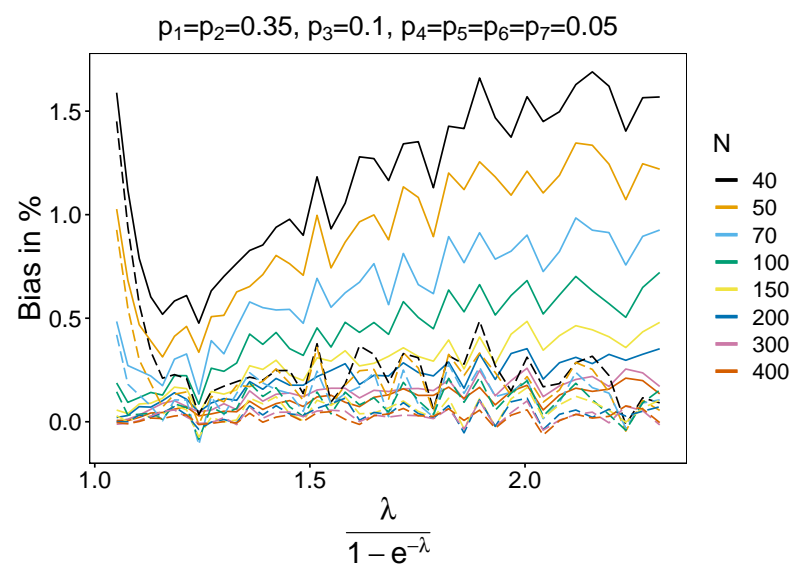
A

p₁= p₂= p₃= p₄= 0.25

B

p₁=p₂=0.4, p₃=p₄=0.1

C

p₁=0.7, p₂=0.1, p₃=0.15, p₄=0.05

D

p₁= p₂= p₃= p₄= p₅= 0.2

E

p₁=0.8, p₂=p₃=p₄=p₅=0.05

F

p₁=p₂=0.35, p₃=0.1, p₄=p₅=p₆=p₇=0.05

Figure 5: Same as Figure 4

A

$p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.1$

B

$p_1 = p_2 = 0.4, p_3 = 0.09, p_4 = 0.05, p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.01$

C

$p_1 = 0.7, p_2 = 0.13, p_3 = 0.1, p_4 = p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.01$

D

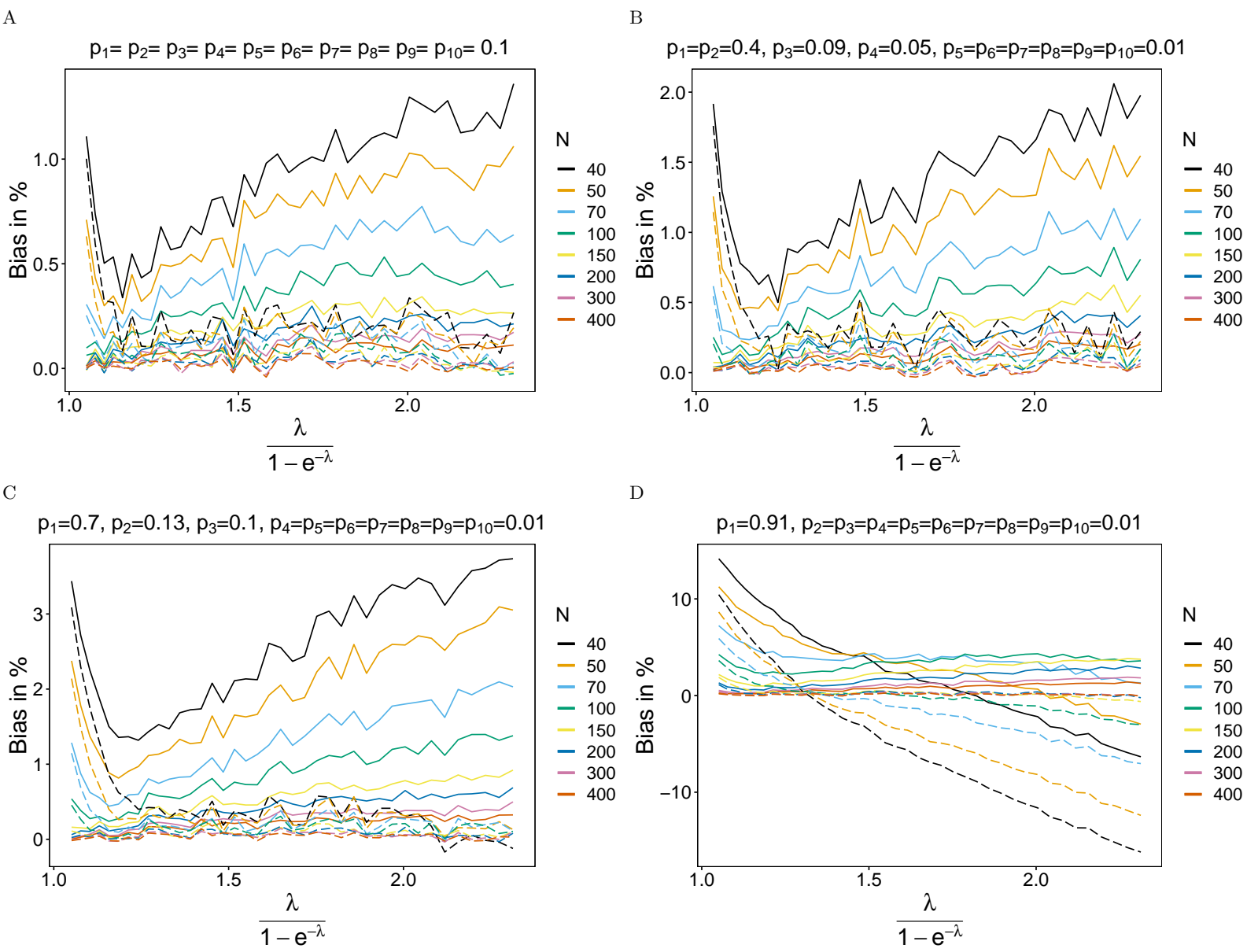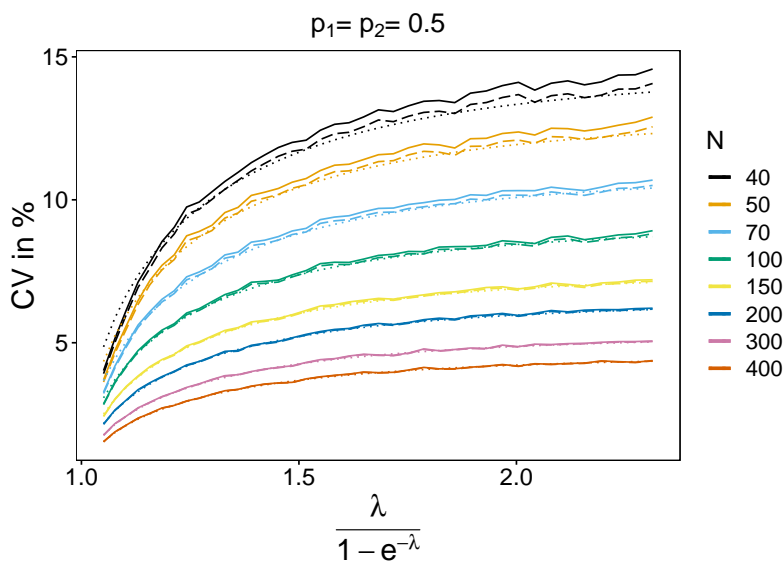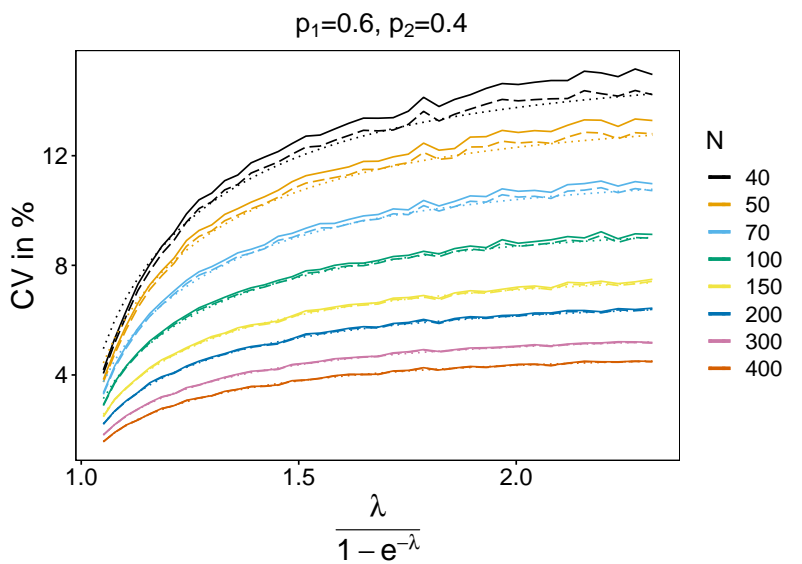$p_1 = 0.91, p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.01$
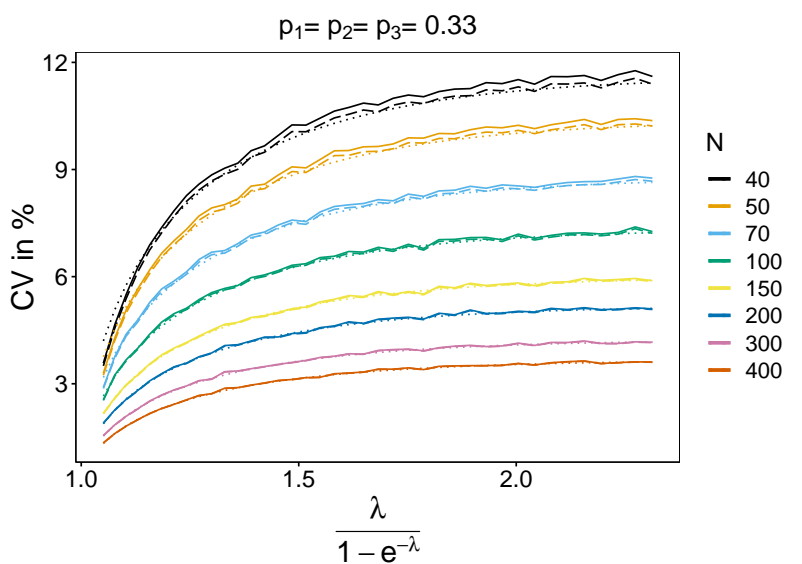
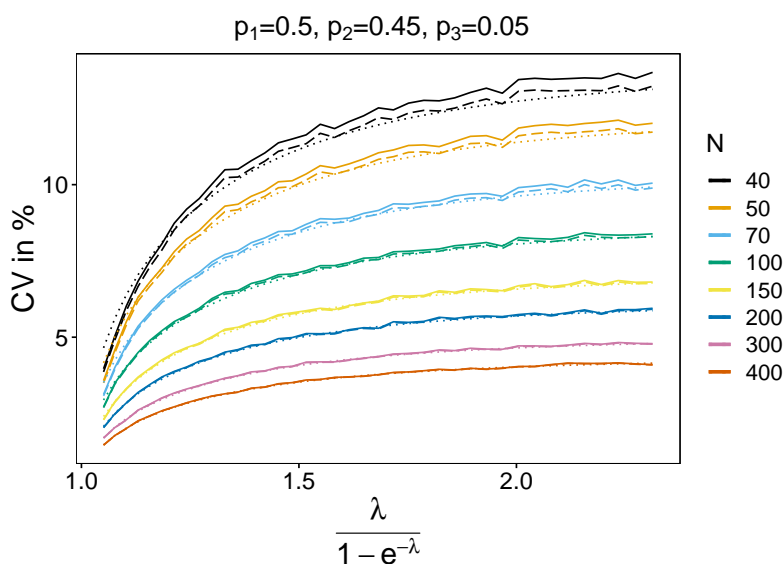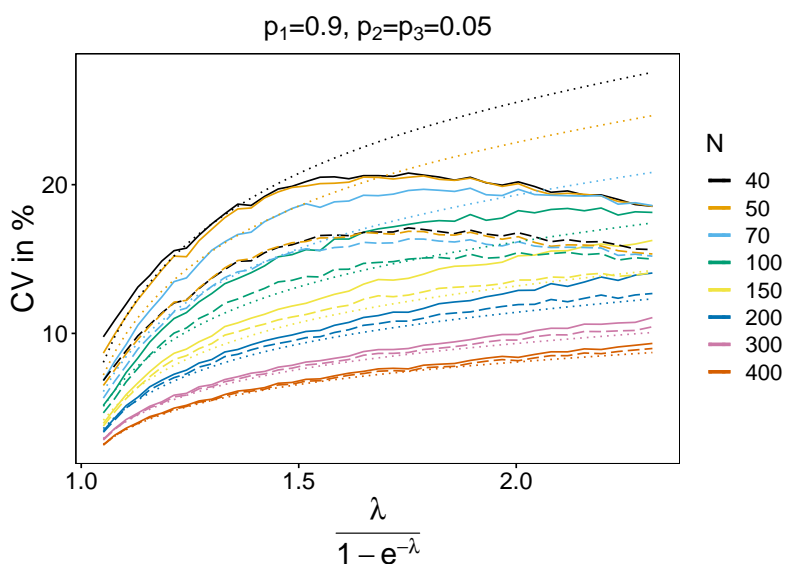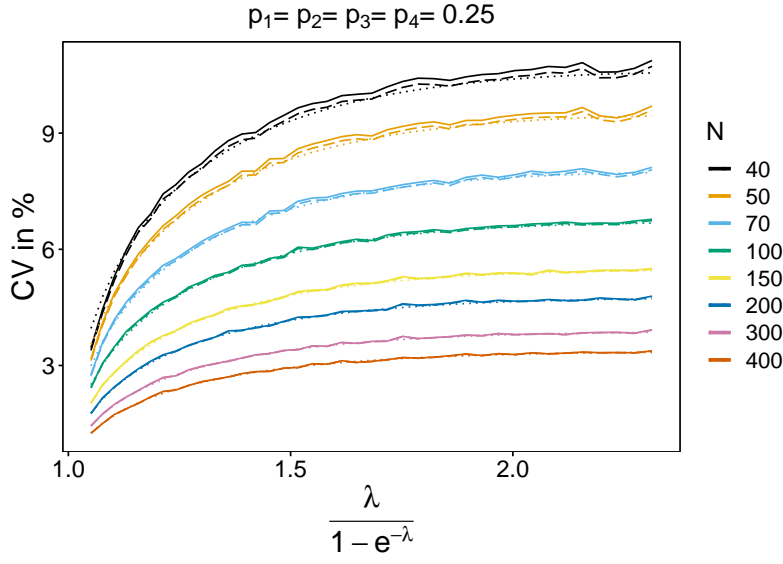Figure 6: Same as Figure 4

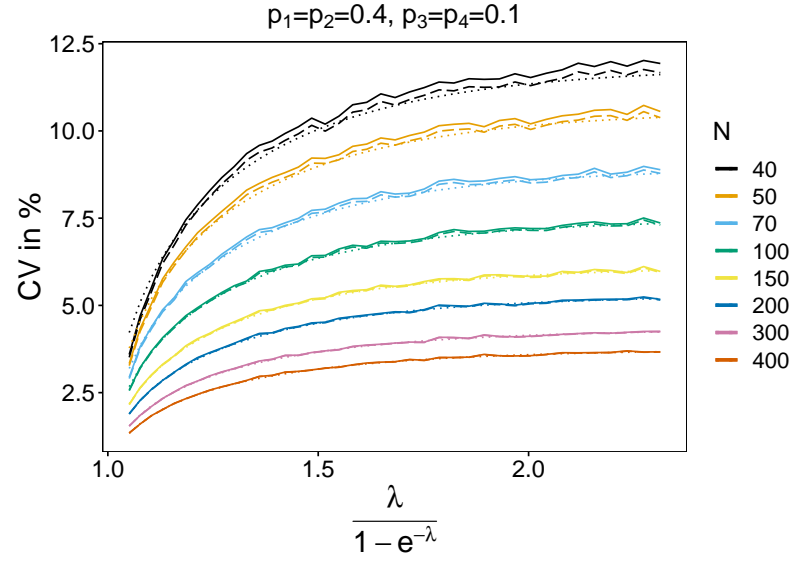# 4 CV of $\hat{\lambda}$

A

B

C
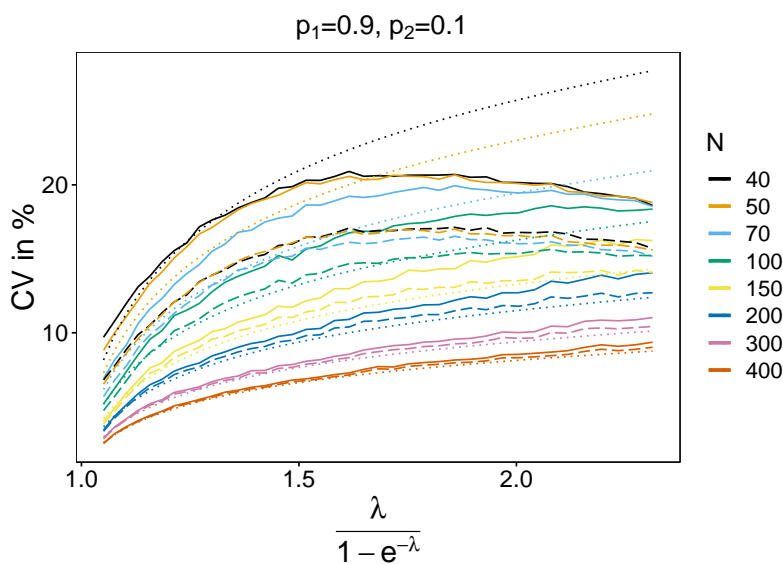
D

E

F

Figure 7: Same as Figure 4 but for coefficient of variation.

A $p_1 = p_2 = p_3 = p_4 = 0.25$

B $p_1 = p_2 = 0.4, \; p_3 = p_4 = 0.1$

C $p_1 = 0.7, \; p_2 = 0.1, \; p_3 = 0.15, \; p_4 = 0.05$

D $p_1 = p_2 = p_3 = p_4 = p_5 = 0.2$

E $p_1 = 0.8, \; p_2 = p_3 = p_4 = p_5 = 0.05$

F $p_1 = p_2 = 0.35, \; p_3 = 0.1, \; p_4 = p_5 = p_6 = p_7 = 0.05$

CV in %

$\dfrac{\lambda}{1 - e^{-\lambda}}$

N: 40, 50, 70, 100, 150, 200, 300, 400

Figure 8

9

A

$p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.1$

B

$p_1 = p_2 = 0.4,\ p_3 = 0.09,\ p_4 = 0.05,\ p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.01$

C

$p_1 = 0.7,\ p_2 = 0.13,\ p_3 = 0.1,\ p_4 = p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.01$

D

$p_1 = 0.91,\ p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = p_8 = p_9 = p_{10} = 0.01$

Figure 9