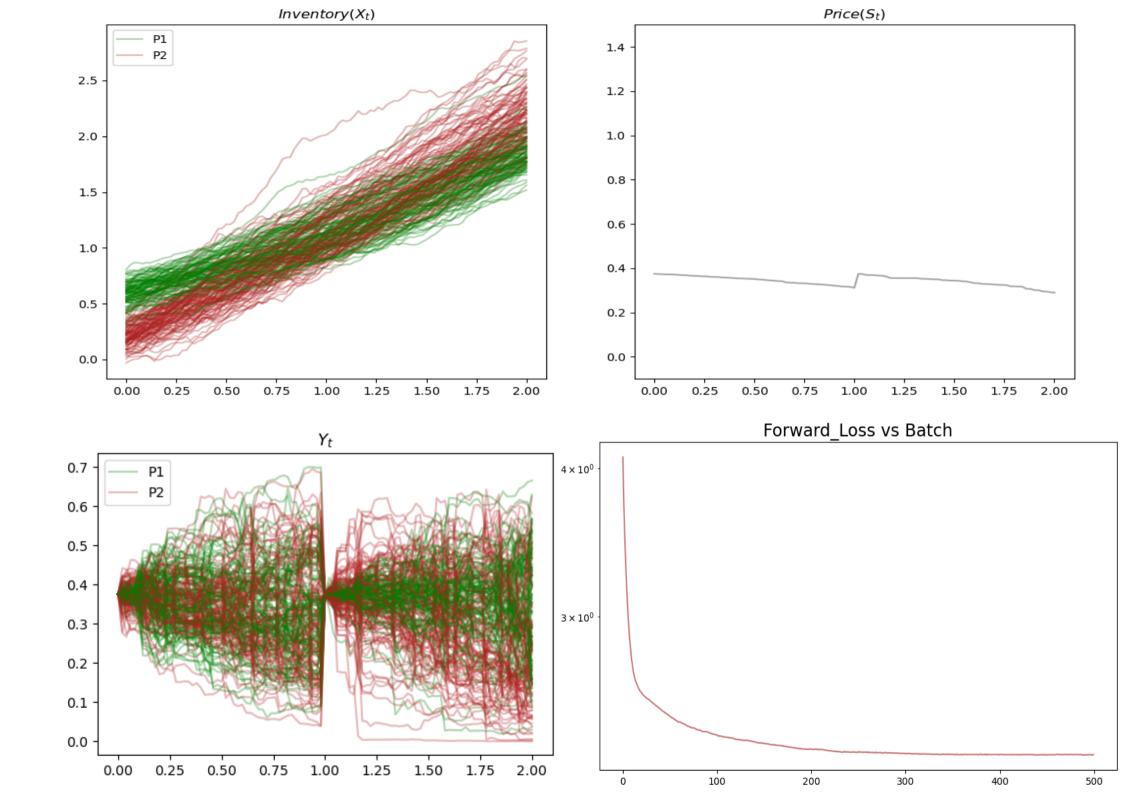
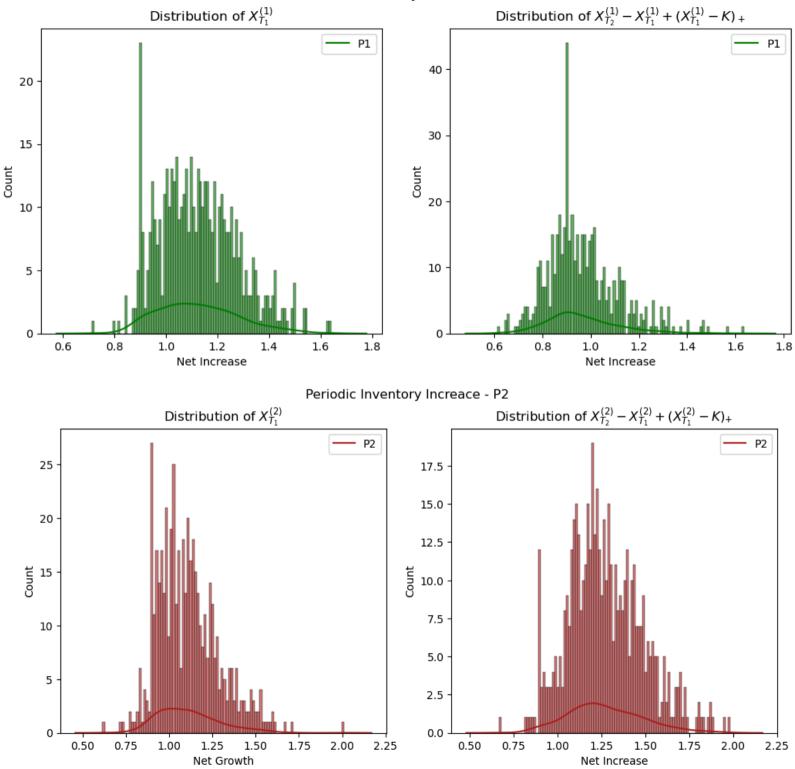


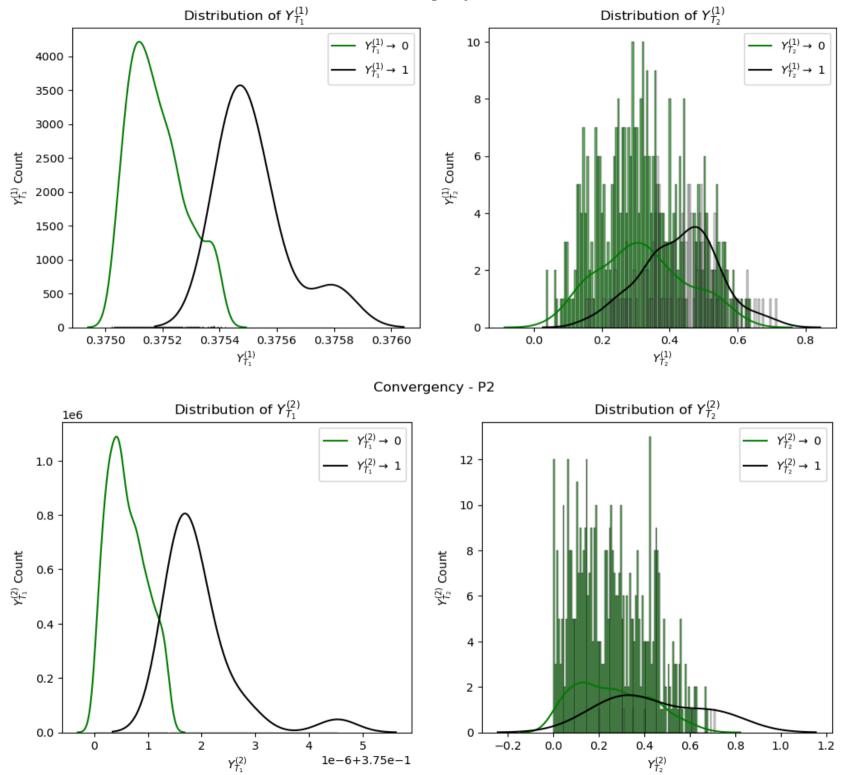
target_type='indicator',trick='logit',loss_type='BCEWithLogitsLoss',delta=0.03,w=0.75,lr=0.001)|

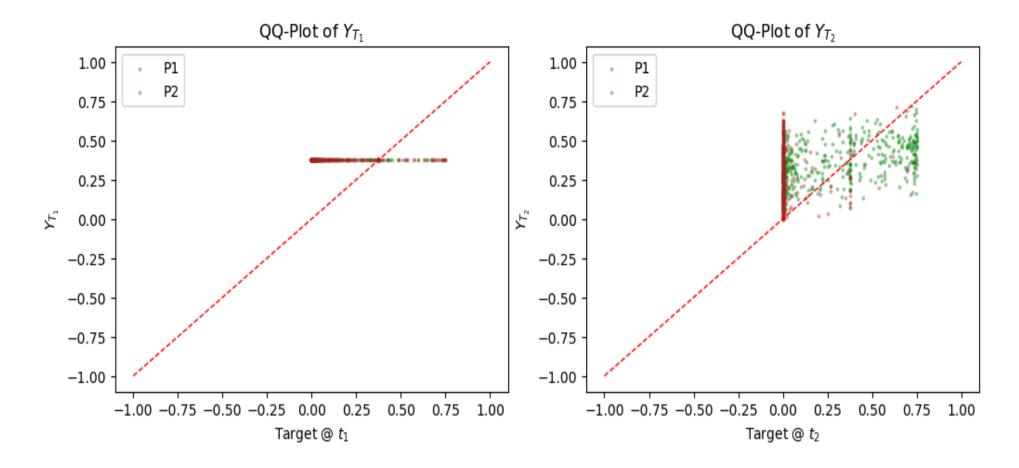
500 steps Started @ August 30 - 21:10:41 Saved @ August 31 - 01:33:46





Convergency - P1





Save The Models

Load The Models

Saved @ August 31 - 01:33:46

```
In [ ]: import numpy as np
        import torch as torch
        import torch.nn as nn
        import torch.nn.functional as F
        import torch.optim as optim
        import seaborn as sns
        import pandas as pd
        import matplotlib.pyplot as plt
        from datetime import datetime
        import random
        from scipy.stats import norm
        import os
        import pathlib
        from Model import *
        from utils import *
        torch.autograd.set detect anomaly(True)
        start time=datetime.now().strftime('%B %d - %H:%M:%S')
In [ ]: #Global parameters
        GlobalParams1=Params(param_type='k1',target_type='indicator',trick='logit',loss_type='BCEWithLogitsLoss',delta=0.03,w=0.75,lr=0.001)
        dB1 = SampleBMIncr(GlobalParams=GlobalParams1)
        init x1 = Sample Init(GlobalParams=GlobalParams1)
        init c1= torch.zeros like(init x1)
        GlobalParams2=Params(param type='k2',target type='indicator',trick='logit',loss type='BCEWithLogitsLoss',delta=0.03,w=0.75,lr=0.001)
        dB2 = SampleBMIncr(GlobalParams=GlobalParams2) ## TODO: same dB?????
        init_x2 = Sample_Init(GlobalParams=GlobalParams2)
        init c2= torch.zeros like(init x2)
        NT1=GlobalParams1.NT1
        NT2=GlobalParams1.NT2
        dt=GlobalParams1.dt
        device=GlobalParams1.device
        learning rate = GlobalParams1.lr
        #Forward Loss
        forward_losses = []
        #How many batches
        MaxBatch= 500
        #How many optimization steps per batch
        OptimSteps= 25
        #Train on a single batch?
        single batch = True
```

```
#Set up main models for y0, yt1, and z (z will be list of models)
        v0 model main1 = Network(scaler type='sigmoid')
        yt1 model main1 = Network(scaler type='sigmoid')
        zy models main1 = [Network() for i in range(NT2)]
        main models1=Main Models(GlobalParams=GlobalParams1)
        main models1.create(y0 model=y0 model main1,
                            yt1 model=yt1 model main1.
                            zy models=zy models main1,
                            forward loss=forward losses.
                            dB=dB1,
                            init x=init x1.
                            init c=init c1)
        v0 model main2 = Network(scaler type='sigmoid')
        vt1 model main2 = Network(scaler type='sigmoid')
        zy models main2 = [Network() for i in range(NT2)]
        main models2=Main Models(GlobalParams=GlobalParams2)
        main_models2.create(y0_model=y0_model_main2,
                            yt1 model=yt1 model main2,
                            zv models=zv models main2,
                             forward loss=forward losses,
                            dB=dB2,
                            init x=init x2,
                            init c=init c2)
        pop1 dict={'dB':dB1,
                    'init x':init x1 ,
                    'init c':init c1 .
                    'GlobalParams':GlobalParams1,
                    'main models':main models1}
        pop2 dict={'dB':dB2,
                    'init x':init x2,
                    'init c':init c2 ,
                    'GlobalParams':GlobalParams2,
                    'main models':main models2}
In [ ]: #Define optimization parameters
```

```
params=[]
params = list(main_models1.y0_model.parameters())+\
    list(main_models1.yt1_model.parameters())+\
    list(main_models2.y0_model.parameters())+\
    list(main_models2.yt1_model.parameters())

for i in range(NT2):
    params += list(main_models1.zy_models[i].parameters())
    params += list(main_models2.zy_models[i].parameters())

#Set up optimizer and scheduler
optimizer = optim.Adamax(params, lr=learning_rate)
```

```
scheduler = torch.optim.lr scheduler.StepLR(optimizer, step size=100, gamma=0.95)
for k in range(0.MaxBatch):
    print("Batch Number: ", k+1)
    sloss=0
    #optimize main network wrt the foward loss
    for l in range(0,0ptimSteps):
        optimizer.zero grad()
        loss = get foward loss(pop1 dict=pop1 dict, pop2 dict=pop2 dict)
        loss.backward()
        # torch.nn.utils.clip grad norm (parameters=params,max norm=0.7)
        optimizer.step()
        scheduler.step()
        nloss = loss.detach().numpy()
        sloss += nloss
        # print('OptimStep: '+ str(l+1))
        # print('forward loss: ' + str(nloss))
   avgloss = sloss/OptimSteps
   print("Average Error Est: ", avgloss)
    forward losses.append(avgloss)
   #Generate a new batch if using multiple batches
    if(not single batch):
        dB1 = SampleBMIncr(GlobalParams=GlobalParams1)
        init x1 = Sample Init(GlobalParams=GlobalParams1)
        init c1= torch.zeros like(init x1)
        pop1 dict={'dB':dB1,
                'init x':init x1 ,
                'init_c':init_c1 ,
                'GlobalParams':GlobalParams1,
                'main models':main models1}
        dB2 = SampleBMIncr(GlobalParams=GlobalParams2) ## TODO: same dB??????
        init x2 = Sample Init(GlobalParams=GlobalParams2)
        init c2= torch.zeros like(init x2)
        pop2_dict={'dB':dB2,
                'init x':init x2 ,
                'init c':init c2 ,
                'GlobalParams':GlobalParams2,
                'main_models':main_models2}
end time=datetime.now().strftime('%B %d - %H:%M:%S')
```