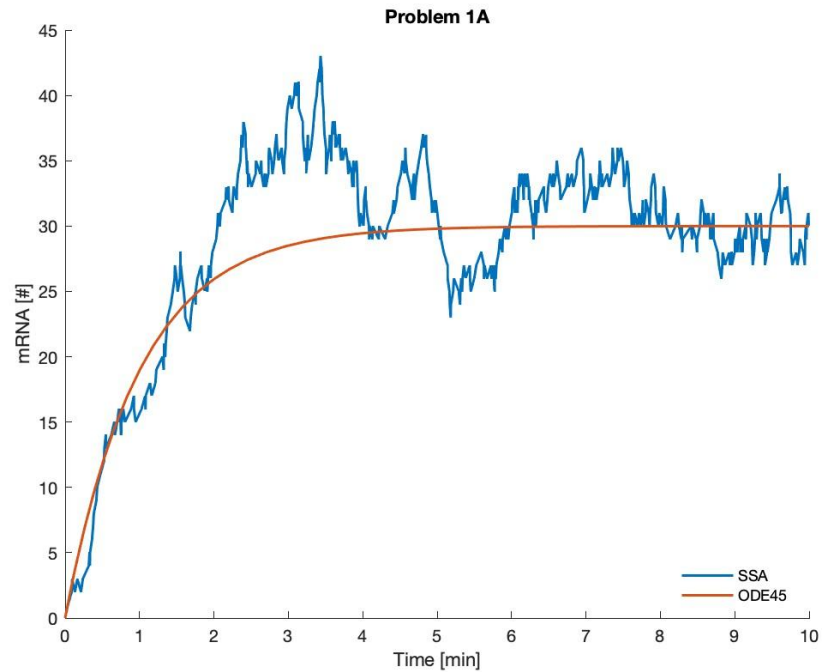Robert Heeter
BIOE 446
29 September 2023

Lab 5: Stochastic Gene Expression and Cellular Variability

ANSWERS
    1. Simulation of stochastic gene expression dynamics (DEMO)
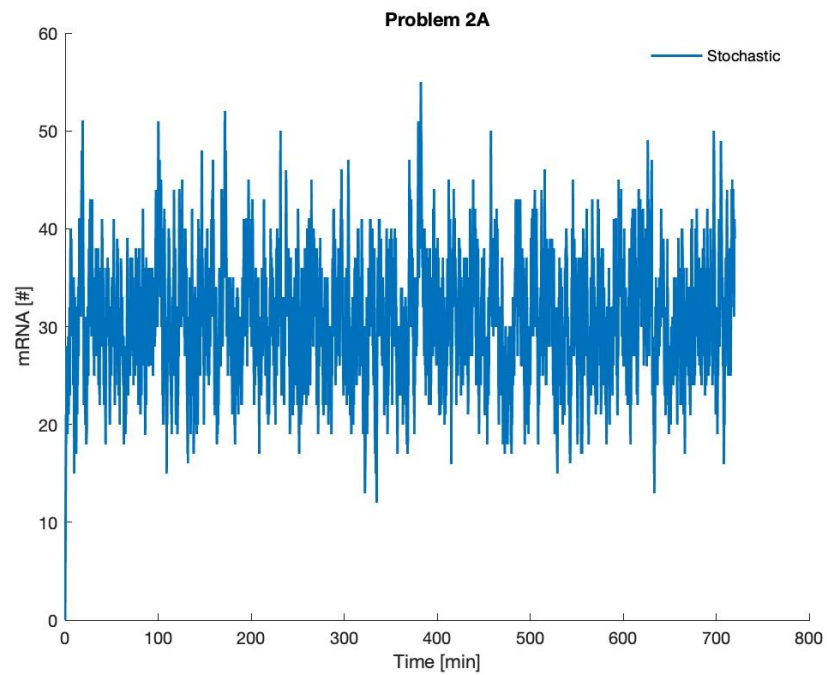   a. Figure

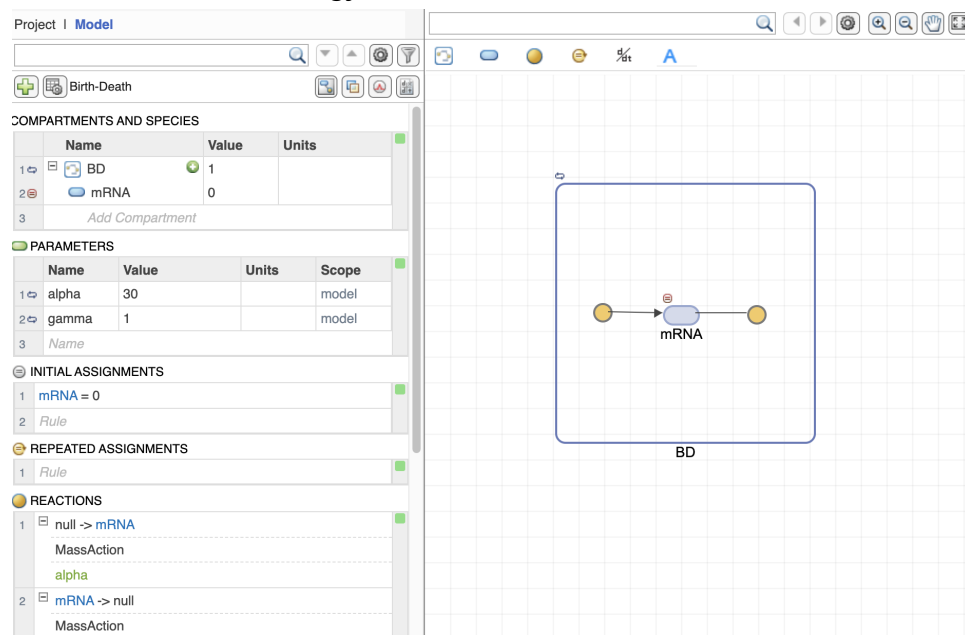   b. Mean M = 30.4960
           Variance M = 31.0042

    2. Modeling stochastic gene expression with SimBiology
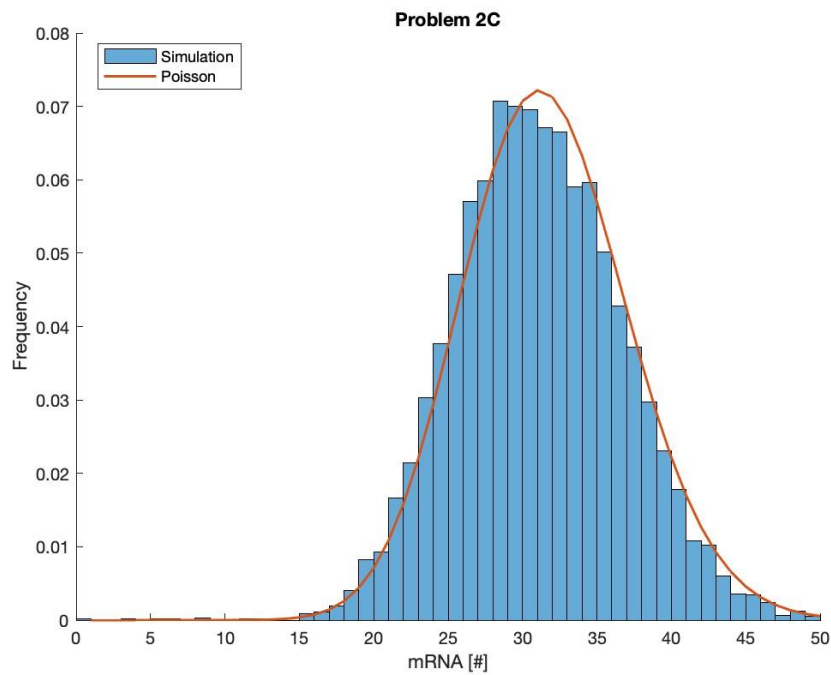   a. Figure
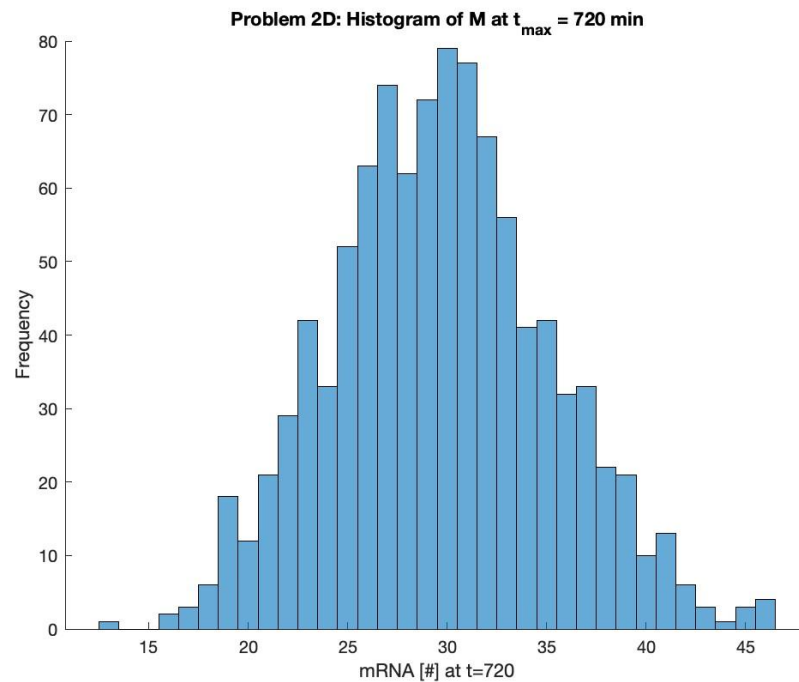
Problem 2A

Screenshot of SimBiology GUI model

b. Mean M = 30.6160
   Variance M = 31.2913

c. Figure

Problem 2C

The distribution of M is nearly identical to the Poisson distribution.
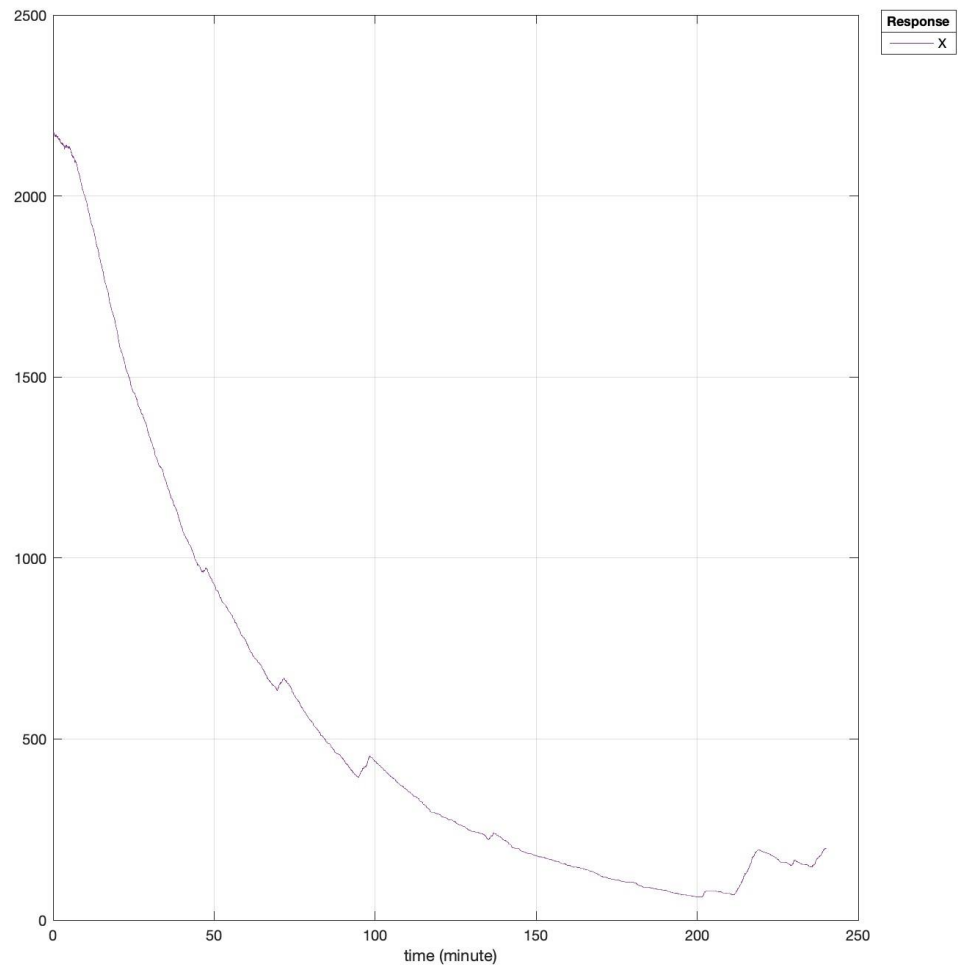
d. Figure

Problem 2D: Histogram of M at t_{max} = 720 min

3. Expanding the transcriptional bursting model
a. Screenshot of SimBiology GUI model

	Name	Value	Units
1	MP	1	
2	mRNA	2	
3	P0	0	
4	P1	1	
5	X	2174	
6	Add Compartment		

PARAMETERS

	Name	Value	Units	Scope
1	kf	5		model
2	gamma1	0.2300		model
3	ktsn	20		model
4	alpha	0.1000		model
5	kb	5		model
6	gamma2	0.0200		model
7	Name			

INITIAL ASSIGNMENTS

1	P1 = 1
2	P0 = 0
3	mRNA = 2
4	X = 2174
5	Rule

REPEATED ASSIGNMENTS
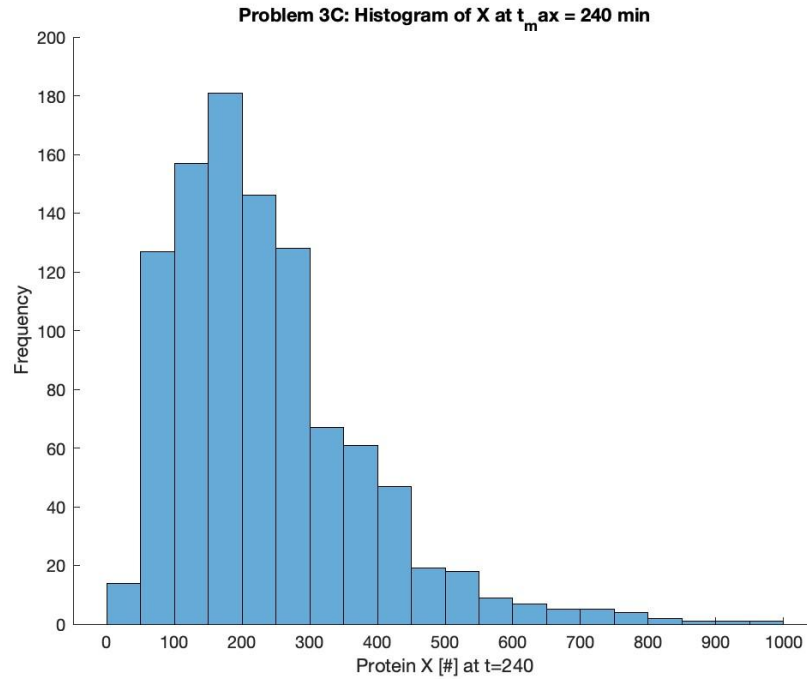
1	Rule

Figure

Mean X = 1137

The average value of X (1137) is about half of its initial value of 2174. Looking at the figure, over time the amount of protein X decreases.
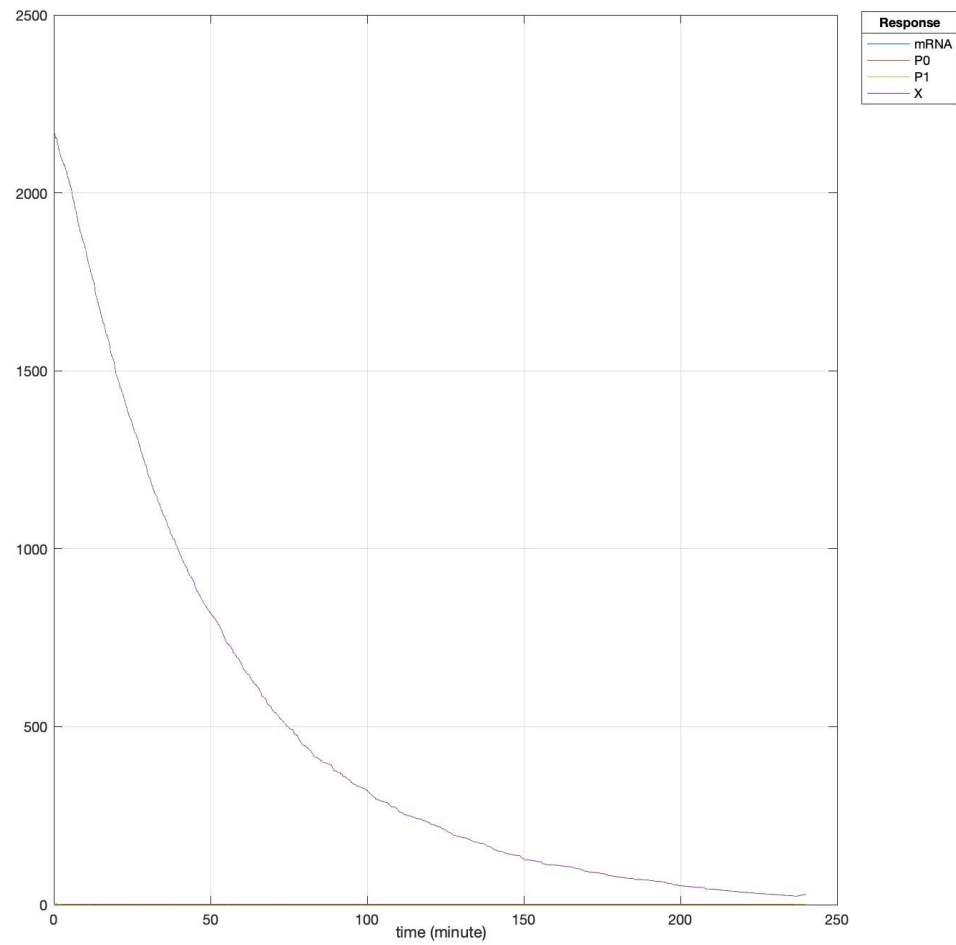
b. No output.

c. Figure

Problem 3C: Histogram of X at t_{max} = 240 min

The distribution of final protein X at time 240 min appears to be skewed to the right, indicating that by the end of the 240 min simulation, most protein X is degraded (low amount) but in a few cases a high amount may remain.

d. Mean X = 232.2020
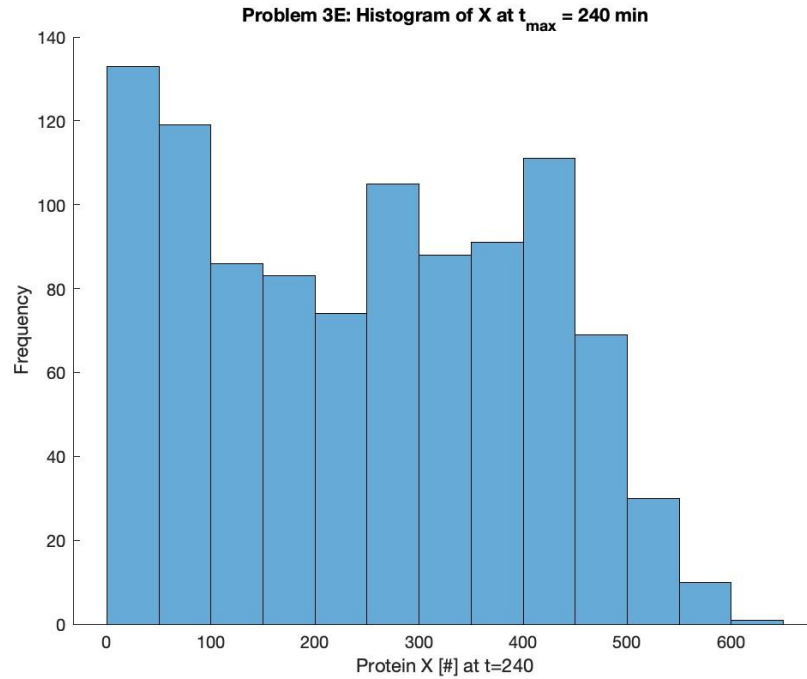   Standard deviation X = 134.7695
   Coefficient of variation X = 0.5804

e. Figure

Mean X = 941.5

Again, the average value of X (941.5) is about half of its initial value of 2174.

Figure

Problem 3E: Histogram of X at t_{max} = 240 min

The mean amount of protein X with these new parameters (slower promoter switching, slower translation, faster transcription) is lower, indicating a higher overall degradation rate of X, and the histogram appears to be less skewed. There is also less noise in this system, potentially due to decreased mRNA conversion and translational 'bursting'.
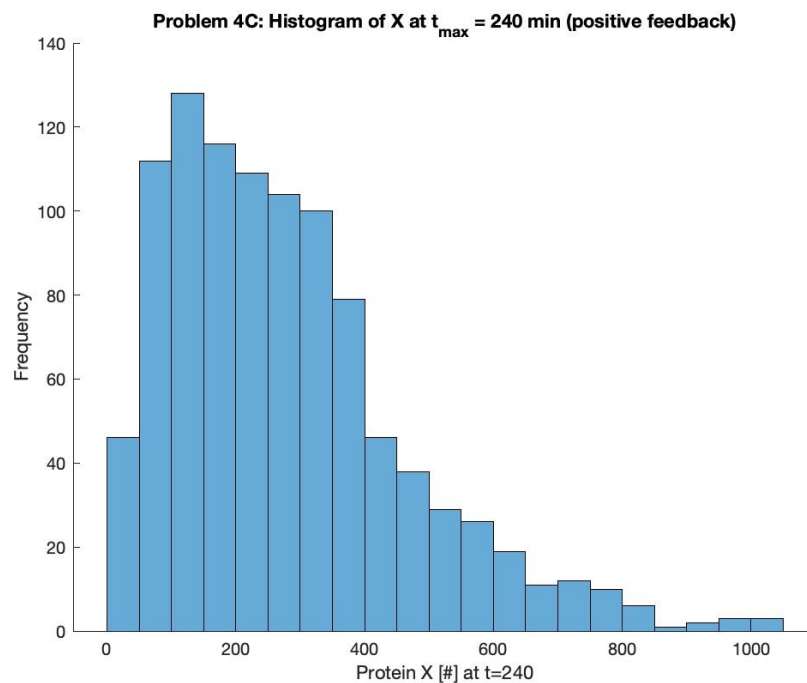
4. Feedback and regulatory noise
   a. Screenshot of SimBiology GUI model
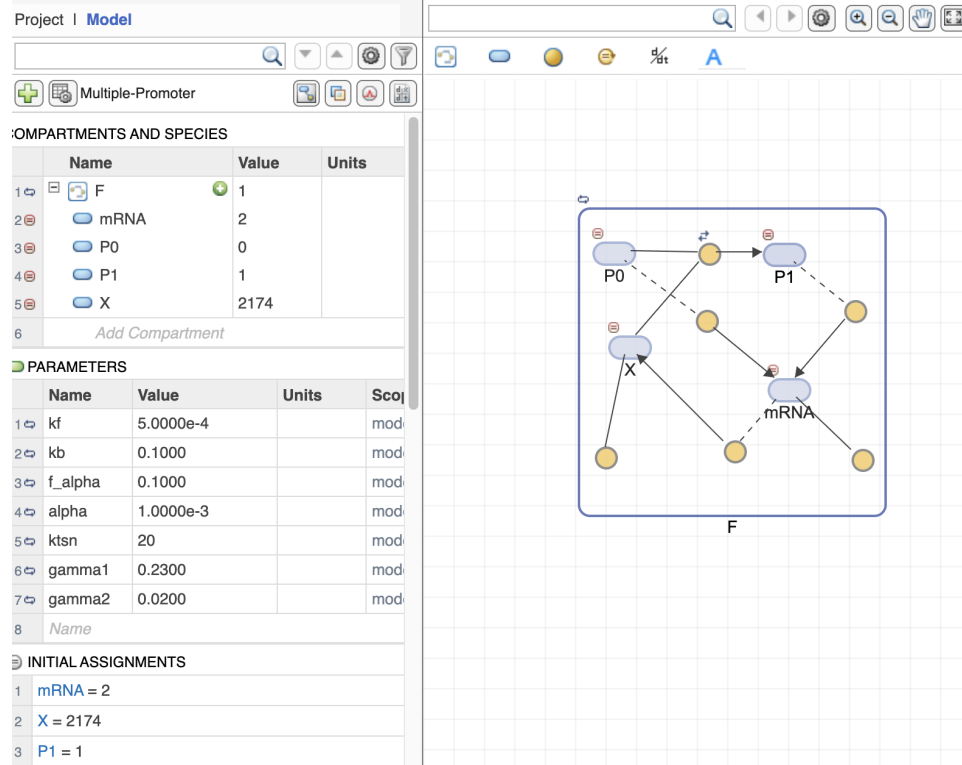
   b. No output.

   c. Figure

```
Mean X = 277.9890
Standard deviation X = 184.6857
Coefficient of variation X = 0.6644
```

Adding positive feedback seems to have skewed the distribution of X more towards the right, likely because the positive feedback favors the production of protein X. There is a higher final average amount of protein X at time 240 min as well (277 vs. 232 in problem 3 without feedback). There is also a higher standard deviation and coefficient of variation with the positive feedback.

d. No output.

e. Figure

Problem 4E: Histogram of X at t_{max} = 240 min (negative feedback)
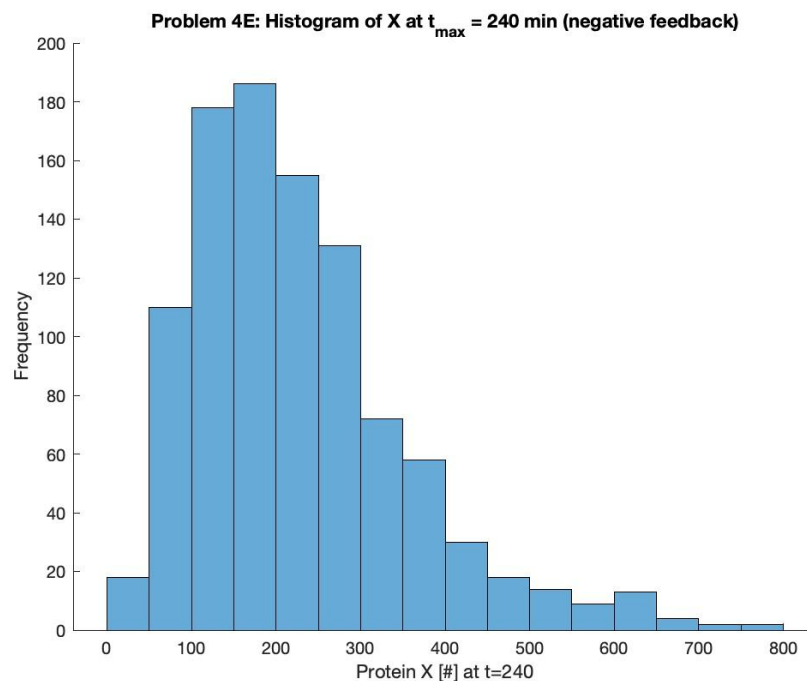
```
Mean X = 226.0810
Standard deviation X = 127.0107
Coefficient of variation X = 0.5618
```

The addition of repression (negative feedback) reduces the right skew of the distribution of X, likely because the negative feedback favors the greater degradation/reduced production of X. There is a lower average amount of protein X at time 240 min as well (226) compared to both the positive feedback case (277) and no feedback case (232). The standard deviation and coefficient of variation more closely match the no feedback case in problem 3 and are both lower than the positive feedback case in part 4A.

CODE
```matlab
close all
clc

%% Challenge Problem 1
disp('CHALLENGE PROBLEM 1')
clear

x = rand;
Tnext = (1/5)*log(1/x);
Rind = rand;

if Rind < 2/5
    n = 4;
else
    n = 2;
end
disp(n)


%% Problem 1: Simulation of stochastic gene expression dynamics (DEMO)
disp('PROBLEM 1')
clear

%% Problem 1, Part A
% initialize
M0 = 0;
alpha = 30;
gamma = 1;
tmax = 720;

i = 1;
M(1) = 0;
t(1) = 0;

t_current = 0;
m_current = 0;

% SSA
while t_current < tmax
    % calculate production and degradation
    production = alpha;
    degradation = gamma*m_current;
    total_rate = production + degradation;

    % determine wait time until the next event
    tau = exprnd(1/total_rate);
    t_current = t_current + tau;

    % determine which event occurred (i.e. birth or death)
    if (rand < production/total_rate)
```

```matlab
            m_current = m_current + 1;
        else
            m_current = m_current - 1;
        end

        i = i+1;
        M(i) = m_current;
        t(i) = t_current;
end

% ode solution
[t_ode, M_ode] = ode45(@(t,m) alpha-gamma*m, [0 10], 0);

% plot results
figure(1)
hold on
plot(t(t<=10), M(t<=10), LineWidth=1.5, DisplayName='SSA')
plot(t_ode, M_ode, LineWidth=1.5, DisplayName='ODE45')
xlabel('Time [min]')
ylabel('mRNA [#]')
title('Problem 1A')
legend(location='best')
legend box off

%% Problem 1, Part B
M_reg = interp1(t, M, 0:0.1:720, 'previous');
mean_M_reg = mean(M_reg)
var_M_reg = var(M_reg)


%% Problem 2: Modeling stochastic gene expression with SimBiology
disp('PROBLEM 2')
clear

%% Problem 2, Part A
simbio = load('problem2a.mat');

figure(2)
hold on
plot(simbio.problem2a.Time, simbio.problem2a.Data, LineWidth=1.5,
DisplayName='Stochastic')
xlabel('Time [min]')
ylabel('mRNA [#]')
title('Problem 2A')
legend(location='best')
legend box off

%% Problem 2, Part B
t = simbio.problem2a.Time;
M = simbio.problem2a.Data;
M_reg = interp1(t, M, 0:0.1:720, 'previous');
```

```matlab
mean_M_reg = mean(M_reg)
var_M_reg = var(M_reg)

%% Problem 2, Part C
poisson_dist = poisspdf(0:720, mean(M_reg));

% plot histogram for M
figure(3)
hold on
histogram(M_reg, 'normalization', 'pdf', DisplayName='Simulation')
plot(poisson_dist, LineWidth=1.5, DisplayName='Poisson')
xlim([0 50])
xlabel('mRNA [#]')
ylabel('Frequency')
title('Problem 2C')
legend(location='best')

%% Problem 2, Part D
project = sbioloadproject('Birth-Death.sbproj');
model = project.m1;
config = getconfigset(model);
set(config, 'SolverType', 'ssa');

M_ens = sbioensemblerun(model, 1000);

y = zeros(1000, 1);
for i = 1:1000
    y(i) = M_ens(i).Data(end);
end

% plot results
figure(4)
hold on
histogram(y)
xlabel('mRNA [#] at t=720')
ylabel('Frequency')
title('Problem 2D: Histogram of M at t_{max} = 720 min')


%% Problem 3: Expanding the transcriptional bursting model
disp('PROBLEM 3')
clear

%% Problem 3, Part A
simbio = load('problem3a.mat');
X = simbio.problem3a.Data(:,4);
mean_X = mean(X)

%% Problem 3, Part B
project = sbioloadproject('Multiple-Promoter.sbproj');
model = project.m1;
```

```matlab
config = getconfigset(model);
set(config, 'SolverType', 'ssa');

X_ens = sbioensemblerun(model, 1000);

%% Problem 3, Part C
X = zeros(1000, 1);
for i = 1:1000
    data = X_ens(i).Data(:,4);
    X(i) = data(end);
end

% plot results
figure(5)
hold on
histogram(X)
xlabel('Protein X [#] at t=240')
ylabel('Frequency')
title('Problem 3C: Histogram of X at t_{max} = 240 min')

%% Problem 3, Part D
mean_X = mean(X)
stdev_X = std(X)
co_var_X = stdev_X/mean_X

%% Problem 3, Part E
simbio = load('problem3e.mat');
X = simbio.problem3e.Data(:,4);
mean_X = mean(X)

project = sbioloadproject('Multiple-Promoter.sbproj');
model = project.m1;
config = getconfigset(model);
set(config, 'SolverType', 'ssa');

X_ens = sbioensemblerun(model, 1000);

X = zeros(1000, 1);
for i = 1:1000
    data = X_ens(i).Data(:,4);
    X(i) = data(end);
end

% plot results
figure(6)
hold on
histogram(X)
xlabel('Protein X [#] at t=240')
ylabel('Frequency')
title('Problem 3E: Histogram of X at t_{max} = 240 min')
```

```matlab
%% Problem 4: Feedback and regulatory noise
disp('PROBLEM 4')
clear

%% Problem 4, Part A
% No MATLAB code required.

%% Problem 4, Part B
project = sbioloadproject('Feedback.sbproj');
model = project.m1;
config = getconfigset(model);
set(config, 'SolverType', 'ssa');

X_ens = sbioensemblerun(model, 1000);

%% Problem 4, Part C
X = zeros(1000, 1);
for i = 1:1000
    data = X_ens(i).Data(:,4);
    X(i) = data(end);
end

% plot results
figure(7)
hold on
histogram(X)
xlabel('Protein X [#] at t=240')
ylabel('Frequency')
title('Problem 4C: Histogram of X at t_{max} = 240 min (positive feedback)')

mean_X = mean(X)
stdev_X = std(X)
co_var_X = stdev_X/mean_X

%% Problem 4, Part D
project = sbioloadproject('Feedback.sbproj');
model = project.m1;
config = getconfigset(model);
set(config, 'SolverType', 'ssa');

X_ens = sbioensemblerun(model, 1000);

%% Problem 4, Part E
X = zeros(1000, 1);
for i = 1:1000
    data = X_ens(i).Data(:,4);
    X(i) = data(end);
end

% plot results
```

```matlab
figure(8)
hold on
histogram(X)
xlabel('Protein X [#] at t=240')
ylabel('Frequency')
title('Problem 4E: Histogram of X at t_{max} = 240 min (negative feedback)')

mean_X = mean(X)
stdev_X = std(X)
co_var_X = stdev_X/mean_X


%% Functions

% None
```