function \[N1,N2]\ = \text{runsimulation}(\lambda,t,\mu)
% A function for running one simulation in problem 1(a) 
% and get a sample of \((N1,N2)\).
%
% Inputs:
% \(\lambda\) intensity of the infection process
% \(t\) time at which we "stop" the process 
% \(\mu\) mean for the incubation times
% Outputs:
% \(N1\) number of symptomless infected people at time \(t\)
% \(N2\) number of infected people that show symptoms at time \(t\)

% Number of infected people at time \(t\)
\(N\ = \text{poissrnd}(\lambda*1)\);

% Generate the arrival times for the infections (unordered)
\(T\ = \ t*\text{rand}(1,N)\);
% Generate the incubation times
Tinc = exprnd(mu,1,N);

% times at which symptoms show
Tsymshown = T+Tinc;

% Number of infected people that show symptoms
% at time t
N2 = sum(Tsymshown <= t);

% Number of symptomless infected people at time t
N1 = N-N2;
Show Empirical vs. Exact Distribution

lambda = 50;
t = 20;

% incubation times
mu = 8;

% number of simulations
B = 10000;
N1 = zeros(1,B);
N2 = zeros(1,B);

% run the simulations
for k=1:B,
    [N1(k),N2(k)] = runsimulation(lambda,t,mu);
end
[F1,X1] = ecdf(N1);

% Estimate the lambda for N1
lambda1 = lambda*mu*(1-exp(-t/mu));

N1cdf = poisscddf(X1,lambdal);

figure(1);
plot(X1,F1,'x',X1,N1cdf,'o');
xlabel('N1(t): Number of symptomless individuals at t=20');
ylabel('Cumulative frequency');
legend('Empirical','Exact');
print -depsc Nlempiricalcdf.eps
X = min(N1):max(N1);
F = histc(N1,min(N1):max(N1));
% normalize
F = F/sum(F);

N1pdf = poisspdf(X,lambda1);

figure(2);
plot(X,F,'x',X,N1pdf,'o');
xlabel('N1(t): Number of symptomless individuals at t=20');
legend('Empirical','Exact');
print -depsc N1empiricalpdf.eps
Accuracy of the Prediction

\[ p = \frac{\lambda_1}{(\lambda_0 t)}; \]
\[ q = 1-p; \]

% The prediction:
\[ \hat{N}_1 = \frac{p}{q} N_2; \]

% The prediction errors:
\[ \text{predErrors} = N_1 - \hat{N}_1; \]

% Calculating the mean and standard deviation
% of the prediction errors:
\[ \text{empMeanPredErrors} = \text{mean}(\text{predErrors}) \]
\[ \text{empStdPredErrors} = \text{std}(\text{predErrors}) \]

\[ \text{formulaStdPredErrors} = \sqrt{\frac{\lambda_0 p t}{q}} \]
figure(4);
hist(predErrors,15);
xlabel('Prediction errors');ylabel('Count');
print -deps predErrorsEmpiricalpdf.eps

% calculate the percentiles
prctile(predErrors,2.5)
prctile(predErrors,97.5)

Formula for std error: 24.0872
Estimated 95% confidence interval for prediction error $[-48.16, 48.16]$

Actual percentiles: $-46.9275$ and $47.5870$

Excellent agreement