Homework 7, CS 177

Applications of Probability in Computer Science: Winter 2007

Due in-class Thursday March 8th

Suggested Reading

- Chapter 2.6, 2.7 on the exponential and Gaussian (Normal) pdfs.
- Chapter 5 on simulation methods.

Your solutions should be submitted electronically to the Homework 7 dropbox for EEE. The MATLAB functions should be submitted using the function names below and the remainder of your solutions should be submitted in a single Word document or PDF file. In cases where numerical answers are required provide your answers to 4 decimal places.

Problem 1

Implement a function in MATLAB called `sim_exp.m` that simulates data from an exponential pdf. Please use comments liberally in your code. The header for your function should be exactly as follows:

```
function data = sim_exp(n, lambda);
%  data = sim_exp(n, lambda)
%
%  simulates n data points from an exponential density function
%
% INPUTS:
%   n: number of data points simulated
%   lambda: parameter of the exponential distribution
%
% OUTPUT:
%   data: n x 1 vector of simulated data points
%
%                              [Student name here], CS 177, Winter 2007


% rest of the code follows below.....
```

You can use the function `rand.m` to generate n uniformly distributed random numbers, and Chapter 5 of the text will explain how to transform these numbers into exponentially distributed random numbers.

- Set lambda = 10 and generate a histogram of $n = 1000$ data points, using 20 bins in the histogram (use `hist.m` to generate the histogram). Superimpose on the same plot the true pdf (you can use `plot.m` and the command `hold on` to superimpose the pdf on the histogram). Briefly comment on the results.
- For the same data you generated for part 1, calculate the empirical mean and empirical variance of the simulated data[1]. Compare these empirical numbers to the values we would theoretically expect and briefly comment.
- Repeat the 2 steps above for lambda= 0.1.

[1] "empirical" here means "as calculated from the data"—to get the empirical variance first calculate the empirical mean and then plug it into your calculation for the empirical variance

Problem 2

Implement a function in MATLAB called `sim_gaussian.m` that simulates data from a Gaussian (Normal) pdf. Please use comments liberally in your code. The header for your function should be exactly as follows:

```
function data = sim_gaussian(n, mean, variance);
%  data = sim_gaussian(n, mean, variance)
%
%  simulates n data points from a Gaussian density function
%
% INPUTS:
%   n: number of data points simulated
%   mean: mean parameter of the Gaussian density
%   variance: variance parameter of the Gaussian density
%
% OUTPUT:
%   data: n x 1 vector of simulated data points
%
%                              [Student name here], CS 177, Winter 2007


% rest of the code follows below.....
```

You can use the function `rand.m` to generate n uniformly distributed random numbers. Chapter 5 of the text explains how to transform these numbers into Gaussian distributed random numbers with mean 0 and variance 1. You will need to further figure out how to transform these $N(0,1)$ data points into data with values of mean and variance as specified in the input arguments (recall how we defined the variable z in class as the "standardized" Gaussian/Normal random variable).

- Set mean=10 and variance=4 and generate a histogram of $n = 1000$ data points, using 20 bins in the histogram (use `hist.m` to generate the histogram). Superimpose on the same plot the true pdf. Briefly comment on the results.
- For the same data you generated for part 1, calculate the empirical mean and empirical variance of the simulated data. Compare these empirical numbers to the values we would theoretically expect and briefly comment.
- Repeat the 2 steps above for $n = 50$ and briefly comment.

Problem 3

We have a switch on a network where packets arrive in an unpredictable manner. Assume that the time t between packet arrivals is exponential with parameter λ.

1. If $\lambda = 100$ packets per microsecond, what is the probability that a randomly selected time-interval between packets will be less than 2 microseconds?

2. If $\lambda = 100$ packets per microsecond, what is the probability that a randomly selected time-interval between packets will be greater than 1 millisecond?

3. Let p be the probability that the gap in time between any pair of packets is greater than 10 microseconds. What value of λ do we need so that $p = 0.01$?

Problem 4

A disk manufacturer collects data on how long it takes before its disks fail. It finds that after 1 year 80% of its disks have not failed. After 2 years 30% of the original disks have not failed. Do you think the time to failure follows an exponential distribution? Clearly explain how you arrived at your answer.