MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Civil and Environmental Engineering

1.017 Computing and Data Analysis for Environmental Applications

---

Problem Set 3: Conditional Probability, Bayes Theorem, and Random Variables
(Solutions provided at end of each problem)
Due: Thursday, October 2, 2003

Please turn in a hard copy of your MATLAB source file as well as all printed outputs (tables, plots, etc.) required to solve the problem.

Problem 1

There are three modes of transporting material from Boston to Houston: by land (either by highway or railway), sea, or air. Half of the materials are transported by land, 20% by sea, and the rest by air. Also, 60% of all land transportation is by highway and the rest is by rail shipments.

Cargo traveling by each mode of transportation has a different probability of being damaged during shipment. The percentages of damaged cargo are as follows: 12% by highway, 3% by rail, 5% by sea, and 2% by air.

Answer the following two questions both analytically and through a virtual experiment using MATLAB.

  a.) What percentage of all cargoes may be expected to be undamaged?
  b.) If a damaged cargo is received, what is the probability that it was shipped by land? By sea? By air?

Problem 1 Solution:

```
% Pset 3 -- Problem 1:
clear all
% Probability cargo transported by each method:
p_land=.5;
p_sea=.2;
p_air=.3;
% out of land transportation:
p_land_highway=.6;
p_land_rail=.4;
% total probabilities for land transpo
p_highway=p_land*p_land_highway;
p_rail=p_land*p_land_rail;
```

```
% probability of damaged cargo
damage_highway=.12;
damage_rail=.03;
damage_sea=.05;
damage_air=.02;

% what percentage of cargo is undamaged?
% if damaged, what is probablilty shipped by each method
(land, sea, air)
nrep=100000;
damaged=zeros(nrep,2);
for i=1:nrep
    shipment=rand(1,1);
    if shipment<p_highway
        cargo=rand(1,1);
        damaged(i,1)=cargo<damage_highway;
        damaged(i,2)=1;        % 1 means shipped by land
    elseif shipment>=p_highway&shipment<(p_highway+p_rail)
        cargo=rand(1,1);
        damaged(i,1)=cargo<damage_rail;
        damaged(i,2)=1;        % 1 means shipped by land
    elseif
shipment>=(p_highway+p_rail)&shipment<(p_highway+p_rail+p_s
ea)
        cargo=rand(1,1);
        damaged(i,1)=cargo<damage_sea;
        damaged(i,2)=2;        % 2 means shipped by sea
    elseif shipment>=(p_highway+p_rail+p_sea)
        cargo=rand(1,1);
        damaged(i,1)=cargo<damage_air;
        damaged(i,2)=3;        % 3 means shipped by air
    end
end

p_undamaged=1-sum(damaged(:,1))/nrep
% numbers that are both damaged and
% by each transportation mode:
d_land=damaged(:,1)==1&damaged(:,2)==1;
d_sea=damaged(:,1)==1&damaged(:,2)==2;
d_air=damaged(:,1)==1&damaged(:,2)==3;
p_damaged_land=sum(d_land)/sum(damaged(:,1))
p_damaged_sea=sum(d_sea)/sum(damaged(:,1))
p_damaged_air=sum(d_air)/sum(damaged(:,1))
```

Problem 2

The National Ambient Air Quality Standard (NAAQS) for Ozone is 0.08 parts per million. In Los Angeles during the month of August, the probability that this standard is exceeded on any given day is approximately 0.3.

Plot the Probability Mass Function (PMF) and the Cumulative Distribution Function (CDF) of the number of days during the month of August that exceed the NAAQS for Ozone in LA **by hand**.

Perform a Monte Carlo simulation using MATLAB to simulate 1000 months of August. Use the results to plot the PMF and CDF from your simulation. What is the probability that the NAAQS will be exceeded on at least 5 out of 31 days? What is the probability that the NAAQS will be exceeded on fewer than 15 days? On between 2 and 10 days?

Problem 2 Solution:

```
% Problem Set 3 Problem 2
clear all
close all
% Ozone Problem
nrep=100000;
randoms=rand(nrep,31);
exceed=randoms<0.3;
tot_d_over=sum(exceed,2);
for i=1:32
    count(i)=sum(tot_d_over==(i-1));
end
% Plot the PMF
figure
stem(0:31,count/nrep)
title('PMF for the Number of Days in August Exceeding Ozone
Standard')
xlabel('Number of Days in August Exceeding Standard')
ylabel('Frequency')
% Plot the CDF
figure
cdfplot(tot_d_over)
title('CDF for the Number of Days in August Exceeding Ozone
Standard')
xlabel('Number of Days in August Exceeding Standard')
ylabel('F(x)')
```