ECSE 304-305B Solution to Assignment 2 Winter 2007

Question 2.1 Let a coin have a probability $\frac{2}{3}$ of coming down Heads in a toss and $\frac{1}{3}$ Tails. Generalize the method in the example below to this biased coin case so as to find the probability of getting three or more heads in five tosses. (This is an example of the use of a *multinomial distribution with parameter* r *=2*.)

Express your answer in terms of fractions and then evaluate it to three decimal places.

Example: Find the probability of getting four or more heads in six tosses of an unbiased coin.

Solution: $P(4 \text{ or more heads in 6 tosses}) = P(4) + P(5) + P(6)$ where,

$$P(k) = P(k \text{ heads in 6 tosses}) = \binom{6}{k} / 2^6$$

so $P(4 \text{ or more heads in 6 tosses}) = (15 + 6 + 1)/2^6 = 11/32$.

Question 2.1 (Solution)

Let $p = \frac{2}{3}$ be the probability of heads, then, we find that

$$P(k \text{ heads in 5 tosses}) = P(k) = \binom{5}{k} p^k (1-p)^{5-k}$$

so

$$P(3 \text{ or more heads in 5 tosses}) = P(3) + P(4) + P(5) = 0.790$$

Question 2.2 Assume that the probability of any particular child being a girl in a given group of families is $\frac{6}{10}$ and the probability the child is a boy is $\frac{4}{10}$. By a double application of the method used to solve Question 2.1, find the probability that among five families, each with six children, at least three of the families have five or more girls? Express your answer in terms of fractions and then evaluate it to three decimal places

Question 2.2 (Solution)

For a single family, let $p = \frac{6}{10}$ be the probability of a girl, then, we find that

$$P(k \text{ girls in 6 children}) = P(k) = \begin{pmatrix} 6 \\ k \end{pmatrix} p^k (1-p)^{6-k}$$

so

$$P(5 \text{ or more girls in 6 children}) = P(5) + P(6) = \frac{436}{1869}$$

Next, for a given group of families, let $p = \frac{436}{1869}$ be the probability of 5 or more girls per family, then, we find that

$$P(k) = \begin{pmatrix} 5 \\ k \end{pmatrix} p^k (1-p)^{5-k}$$

so

$$P(3 \text{ or more families have five or more girls}) = P(3) + P(4) + P(5) = \frac{664}{7661} = 0.0867$$

Question 2.3

A graph G has N nodes $\{n_i; 1 \leq i \leq N\}$ and a set U of undirected edges. G is called a *clique* if each of its N nodes is connected by exactly one edge to every other node. For a clique graph G with $N > 2$ give expressions for the number of:

(a) Distinct unordered edges (n, n') in G, where n, n' are distinct nodes in G.

(b) Distinct unordered edge pairs $((n, n'), (m, m'))$ in G , where (n, n'),(m, m') are distinct edges.

(c) Distinct unordered edges pairs in G which connect at a node, i.e. are of the form $((n', n), (n, n''))$ where n, n', n'' are distinct nodes.

(d) Verify your answers to (b) and (c) for a triangle and for a clique graph on four nodes.

Question 2.3 (Solution)

(a) If there are N nodes in G, then there are

$$C_2^N = \frac{N(N-1)}{2}$$

distinct undirected edges.

(b) It follows that there are

$$\begin{aligned} C_2^{(C_2^N)} &= \frac{C_2^N(C_2^N - 1)}{2} \\ &= \frac{1}{2}\frac{N(N-1)}{2}\left(\frac{N(N-1)}{2} - 1\right) \\ &= \frac{N(N-1)(N^2 - N - 2)}{8} \end{aligned}$$

distinct unordered pairs of distinct edges.

(c) Each node is connected to all the other nodes in the clique, i.e., each node is connected to $N-1$ nodes and thus is the vertex to $N-1$ edges. Thus the number of pairs of edges that connect through any particular node is C_2^{N-1}. The total number of nodes is N. Thus the total number of distinct unordered edge pairs which conntect through a node is given by:

$$\begin{aligned} N \times C_2^{N-1} &= \frac{N((N-1)!)}{(N-3)!2!} \\ &= \frac{N(N-1)(N-2)}{2} \end{aligned}$$

(d) For a triangle (Figure 1), we have $N = 3$, hence there are

$$\frac{C_2^C(C_2^N - 1)}{2} = \frac{1}{2} \cdot 3 \cdot (3 - 1) = 3$$

unordered edge pairs. There are also

$$\frac{N(N-1)(N-2)}{2} = \frac{3 \cdot 2 \cdot 1}{2} = 3$$

unordered ege pairs that connect through a node.

For a clique with $N = 4$ nodes (Figure 2), there are

$$\frac{C_2^C(C_2^N - 1)}{2} = \frac{1}{2} \cdot 6 \cdot (6-1) = 15$$

unordered edge pairs.

There are also

$$\frac{N(N-1)(N-2)}{2} = \frac{4 \cdot 3 \cdot 2}{2} = 12$$

unordered ege pairs that connect through a node.

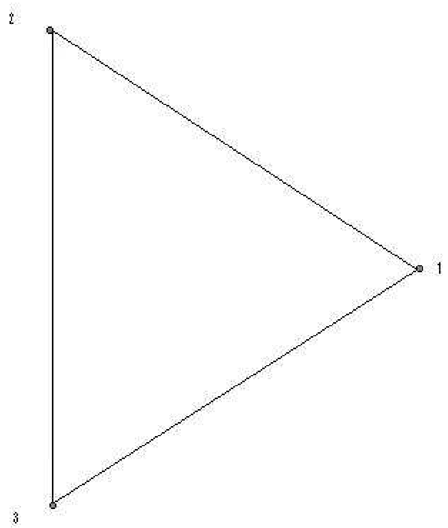

Figure 1: Triangle

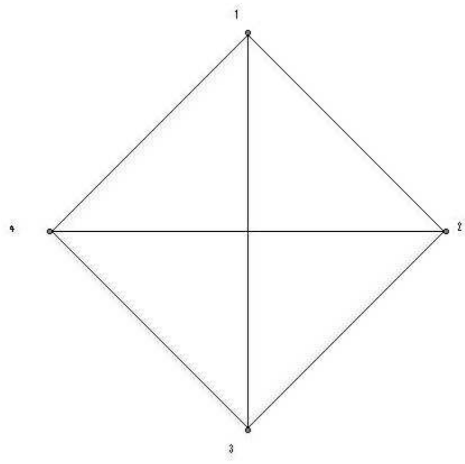

Figure 2: Clique of size 4

Question 2.4

In a specific application of the Partition Function of statistical thermodynamics, a substance in a vessel has $n = 10^{10}$ particles at a temperature $\mu = 1/100$, in appropriate units. There are four energy levels $e_i = i^2$, for $i = 1, 2, 3, 4$.

(a) Using the approximate version of Sterling's formula in the lecture notes, find an expression for the ratio of the probability of the most likely (corresponding to the most equal) distribution of particles to the probability of the most unlikely distribution of particles. Do this respecting the Conservation of Particles Constraint (1), but neglecting the Energy Conservation Constraint (2).

(b) Find an expression for the number of particles in each energy level in the most likely configuration (subject to (1) and (2)) by applying the formulas for each n_i in the notes.

(c) Give a formula for the energy per particle in the most likely configuration (subject to (1) and (2)) by using the formula for $\frac{E}{N}$ in the notes.

NB Check web vista for the statistical thermodynamics formulas (typos now corrected).

Question 2.4 (Solution)

(a) The most likely distribution is found by solving the following problem:

$$Max \;\; P(n_1, n_2, n_3, n_4) = \frac{N!}{n_1!, n_2!, n_3! n_4!}$$
Subject to: $n_1 + n_2 + n_3 + n_4 = N$

Using Lagrange Multiplier λ and Stirling's Approximation, the most likely distribution is when $n_i = \frac{N}{4} = \frac{10^{10}}{4}, i = 1, 2, 3, 4$.

Thus,

$$Max(P) = \frac{10^{10}!}{(2.5 \times 10^9!)^4}$$

The least likely distribution of particles is when all the particles are in one energy level. Thus $Min(P) = \frac{10^{10}!}{10^{10}!} = 1$. The ratio of the probability of the most likely distribution to the least likely distribution is:

$$\frac{Max(P)}{Min(P)} = \frac{10^{10}!}{(2.5 \times 10^9!)^4}$$

To simplify the above expression we take its logarithm and apply Stirling's Approximation.

$$\begin{aligned} log\frac{Max(P)}{Min(p)} &= logMax(P) - logMin(P) \\ &= log(10^{10}!) - 4log(2.5 \times 10^9!) \\ &\approx 10^{10}(log10^{10} - 1) - 10^{10}(log(2.5 \times 10^9) - 1) \\ &= 10^{10} log(4) = log(4^{10^{10}}) \end{aligned}$$

Thus, $\frac{Max(P)}{Min(P)} \approx 4^{10^{10}}$.

(b) Assuming the conservation of particles constraint and the conservation of energy constraint $n_i = \frac{Ne^{-\mu e_i}}{\sum_{j=1}^{4} e^{-\mu e_j}}$. Thus,

$$n_1 = \frac{10^{10} e^{-\frac{1}{100}}}{e^{-\frac{1}{100}} + e^{-\frac{4}{100}} + e^{-\frac{9}{100}} + e^{\frac{-16}{100}}} \approx 2.664 \times 10^9$$

$$n_2 = \frac{10^{10} e^{-\frac{4}{100}}}{e^{-\frac{1}{100}} + e^{-\frac{4}{100}} + e^{-\frac{9}{100}} + e^{\frac{-16}{100}}} \approx 2.585 \times 10^9$$

$$n_3 = \frac{10^{10} e^{-\frac{9}{100}}}{e^{-\frac{1}{100}} + e^{-\frac{4}{100}} + e^{-\frac{9}{100}} + e^{\frac{-16}{100}}} \approx 2.459 \times 10^9$$

$$n_4 = \frac{10^{10} e^{-\frac{16}{100}}}{e^{-\frac{1}{100}} + e^{-\frac{4}{100}} + e^{-\frac{9}{100}} + e^{\frac{-16}{100}}} \approx 2.293 \times 10^9$$

(c) The energy per particle in the most likely distribution using both constraints is:

$$\frac{E}{N} = \frac{\sum_{j=1}^{4} j^2 e^{\frac{-j^2}{100}}}{\sum_{j=1}^{4} e^{\frac{-j^2}{100}}} \approx 7.181$$

Question 2.5

The elevator of a four-floor building leaves the first floor with six passengers and stops at all of the remaining three floors. If it is equally likely that a passenger gets of at any of these three floors, what is the probability that, at each stop, of the elevator at least one passenger departs?

Hint: This is Problem 28, Chapter 2, Section 2.2 of GS. Use the counting of permutations method which is used to solve Problem 27 (answers in the back of the book). No explicit probabilities are used until you obtain the solution.

Question 2.5 (Solution)

(a) Solution by Multinomial Counting

Let event E={At each stop, at least one passenger departs at each stop}.

The number of distinct ways that 6 distinct passengers can get off the elevator so that the condition defining event E is satisfied is given by: the sum $|E|$ of the terms appearing in the multinomial expression $(s+t+r)^6$ as coefficients of $s^i t^j r^k$ with $i \geq 1, , j \geq 1, k \geq 1$.

(Note that the groups that get off at any floor are unordered).

Hence,

$$|E| = 6 \times \left(\frac{6!}{3!2!1!}\right) + 3 \times \left(\frac{6!}{4!1!1!}\right) + \left(\frac{6!}{2!2!2!}\right)$$

Since the total number $|S|$ of distinct ways that unordered groups can get off at each floor is given by $|S| = 3^6$, the EPP gives $P(E) = \frac{|E|}{|S|} = \frac{20}{27}$

(b) Solution by Combinatorial Probability

Let p be the probability that a passenger gets off at a particular floor. Then $p = \frac{1}{3}$.

We define the events E,A, B, C as follows:

E={At each stop, at least one passenger departs at each stop}

A={At least one passenger gets off at the second floor}

B={At least one passenger gets off at the third floor}

C={At least one passenger gets off at the fourth floor}

Then the probability of at least one passenger departing at each floor is:

$P(E) = P(A \cap B \cap C) = 1 - P((A \cap B \cap C)^c) = 1 - P(A^c \cup B^c \cup C^c)$

Now $P(A^c \cup B^c \cup C^c)$ can be expanded as:

$$P(A^c \cup B^c \cup C^c) = P(A^c) + P(B^c) + P(C^c) - P(A^c \cap B^c) - P(A^c \cap C^c) - P(B^c \cap C^c) + P(A^c \cap B^c \cap C^c)$$

$P(A^c \cap B^c \cap C^c)$ is the probability that none of the passengers departs at any of the floors which is zero.

$P(A^c)$ is the probability that none of the passengers gets off on the second floor, which is $(1-p)^6$. Similary, $P(B^c) = (1-p)^6$ and $P(C^c) = (1-p)^6$.

$P(A^c \cap B^c)$ is the probability that none of the passengers gets off on the second or third floors, i.e. all the passengers get off on the fourth floor, thus $P(A^c \cap B^c) = p^6$. Similarly, $P(A^c \cap C^c) = p^6$ and $P(B^c \cap C^c) = p^6$.

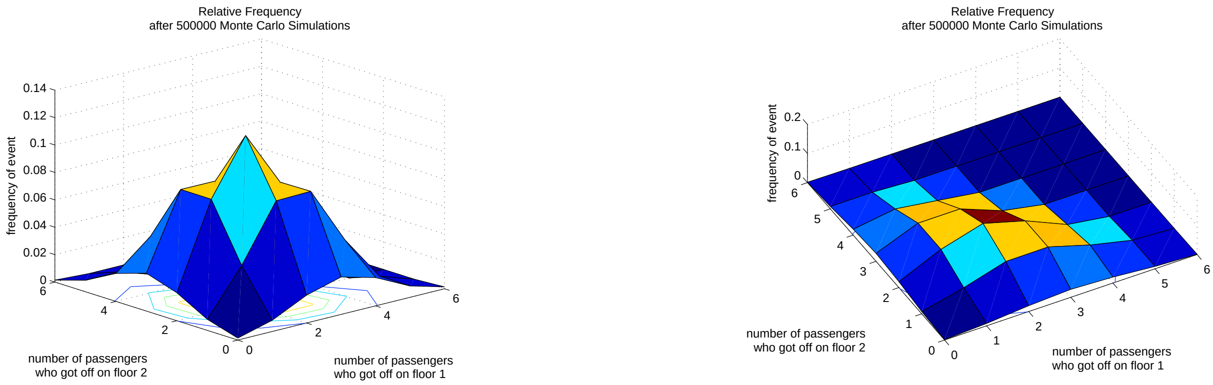

Figure 3: Two views of the relative frequency graphs.

Finally,

$$
\begin{aligned}
P(E) &= 1 - 3(1-p)^6 + 3(p)^6 \\
&= 1 - 3(\frac{2}{3})^6 + 3(\frac{1}{3})^6 \\
&= 1 - \frac{7}{27} = \frac{20}{27} \\
&\approx 0.741
\end{aligned}
$$

To confirm the results, we can perform Monte Carlo Simulations (see Matlab code below). The results are as shown in the table.

Number of Simulations	Relative Frequency of Event	Difference with $\frac{20}{27}$
5	0	0.7407
50	0.74	0.0007
500	0.742	0.0013
5000	0.7454	0.0047
50000	0.74156	0.0008
500000	0.74037	0.0004
5000000	0.74086	0.0001

The Monte Carlo Simulations also enable us to view the relative frequencies.

```
function ass2_07_sol_q5(nmcs)
if nargin == 1
    %do nothing
```

```
else
    nmcs = 500000; %num of monte carlo simulation
end
np = 6; %num of passengers

mcs = ceil(3*rand(np,nmcs)); %monte carlo simulation

%all(mcs-f): will return a vector where entries are 0 if a passenger gets
%off on floor f or 1 if no passenger gets off on floor f.
%or(or(all(mcs-1),all(mcs-2)),all(mcs-3)): will return a vector where
%entries are 0 if at least 1 passenger gets off on every floors and 0 ow.
oos = not(or(or(all(mcs-1),all(mcs-2)),all(mcs-3)));  %outcome of simulation

%display the approximate answer to Q2.5
disp(['Q2.5 ans ' num2str(nmcs) ': ' num2str(sum(oos)/length(oos))]);

npwgo1 = sum(mcs==1); %num passengers who got off on floor 1
npwgo2 = sum(mcs==2);
rf = zeros(np+1); %relative frequency
for i = 1:nmcs;
    rf(npwgo1(i)+1,npwgo2(i)+1) = rf(npwgo1(i)+1,npwgo2(i)+1)+1;
end
rf = rf./nmcs;
X = repmat(0:np,np+1,1);
Y = repmat([0:np]',1,np+1);
surfc(X,Y,rf);
title({'Relative Frequency',['after ' num2str(nmcs) ' Monte Carlo Simulations']});
xlabel({'number of passengers','who got off on floor 1'});
ylabel({'number of passengers','who got off on floor 2'});
zlabel('frequency of event');
```