Problem Set 3 Solutions

1. *Count the number of different letter arrangements you can make by changing the order of the letters in the word* **aardvark** *(count the original word, too).*

   Answer: $\binom{8}{3}\binom{5}{2}3! = \frac{8!}{3!2!1!1!1!} = 3360$.

   Solution: choose 3 places out of the total 8 for the 3 "a"-s (the $\binom{8}{3}$ multiplier), then choose 2 places out of the remaining 5 for the 2 "r"-s (the $\binom{5}{2}$ multiplier), then place "d","v", and "k" in the remaining 3 spaces (the 3! multiplier).

   Equivalently, we can take the number of ways to order 8 *different* letters (8!), and then divide it by the numbers of ways to re-order groups of 3 (letter "a"), 2 (letter "r"), 1 (letter "d"), 1 (letter "v"), 1 (letter "k").

2. *A candy factory has an endless supply of red, orange, yellow, green, blue, black, white, and violet jelly beans. The factory packages the jelly beans into jars in such a way that each jar has 200 beans, equal number of red and orange beans, equal number of yellow and green beans, one more black bean than the number blue beans, and three more violet beans than the number of white beans. One possible color distribution, for example, is a jar of 50 yellow, 50 green, one black, 48 white, and 51 violet jelly beans. As a marketing gimmick, the factory guarantees that no two jars have the same color distribution. What is the maximum number of jars the factory can produce?*

   Answer: $\binom{101}{3} = 166650$.

   Solution: Let $N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8$ denote, respectively, the numbers of red, orange, yellow, green, blue, black, white, and violet jelly beans in a jar. There is a one-to-one correspondence

   $$x = (x_1, x_2, x_3, x_4) \mapsto N = (x_1, x_1, x_2, x_2, x_3, x_3 + 1, x_4, x_4 + 3)$$

   between the non-negative integer solutions $x = (x_1, x_2, x_3, x_4)$ of the equation

   $$x_1 + x_2 + x_3 + x_4 = 98,$$

   and the sequences $N = (N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8)$ of non-negative integers N_i satisfying the conditions

   $$N_2 = N_1,\ N_4 = N_3,\ N_6 = N_5 + 1,\ N_8 = N_7 + 3,\ \sum_{i=1}^{8} N_i = 200$$

   (i.e. possible color arrangements). According to the answer in problem 2 of Recitation 4, the number of possible solutions x is $\binom{101}{3}$.

3. *Vivek offered Dwayne and Star two dollars, on the condition that they decide who gets the money by tossing a fair coin (actually, he even provided the coin and volunteered to do the tossing) until either the total number of tails reaches 20 (in which case Star gets the money) or the total number of heads reaches 20 (in which case Dwayne gets the money). At it turned out, Vivek was so eager to resume his lecture that he terminated the game in the middle. At the time the game was stopped, the count has reached 7 tails to 4 heads. Now, Vivek wants to divide the two dollars between Dwayne and Star in proportion to the probability of their winning, assuming*

the game would continue starting with the 7 to 4 count. How much money should Dwayne get according to this criterion?

Answer: $200 \sum_{i=0}^{12} 2^{-16-i} \binom{15+i}{i} \approx 57$ cents. In general, with initial score $s : d$ and playing to n wins, the probability of Dwayne winning is $\sum_{i=0}^{n-s-1} 2^{d-n-i} \binom{n-d+i-1}{i}$.

Solution: let A_i (with $i \in \{0, 1, \ldots, n-s-1\}$) be the event of Dwayne winning after exactly i "tails". The events A_i are disjoint, hence

$$\mathbf{P}(\text{Dwayne wins}) = \sum_{i=0}^{n-s-1} \mathbf{P}(A_i).$$

A particular event A_i is the outcome of $n - d + i$ coin tosses, of which exactly i were "tails", and the last one os a "head". Since, for a particular assignment of outcomes to each toss, the probability is 2^{d-n-i}, and there are $\binom{n-d+i-1}{i}$ possible ways of assigning i "tail" outcomes to the first $n + d + i - 1$ tosses, $\mathbf{P}(A_i) = 2^{d-n-i} \binom{n-d+i-1}{i}$.

Here is a MATLAB function solving the problem in two different ways, one of which is the approach presented above.

```
function ps03_3(d,s,n)
% function ps03_3(d,s,n)
%
% solves problem 3 of PS3 using 2 different approaches
% d - Dwayne's score at interruption (default 4)
% s - Star's score at interruption (default d+3)
% n - number of points required to win (default s+d+9)

% default arguments
if nargin<1, d=4; end
if nargin<2, s=d+3; end
if nargin<3, n=s+d+9; end

% analytical expression
a=0;
for i=0:n-s-1,
    a=a+2^(d-n-i)*nchoosek(n-d+i-1,i);
end

% a table of probabilities
% p(i,j) = probability of getting to (s+i-1):(d+j-1) score
p=zeros(n-s,n-d);
p(1,:)=2.^(0:-1:d-n+1);
p(:,1)=2.^(0:-1:s-n+1);
for i=2:n-s,
    for j=2:n-d,
        p(i,j)=0.5*(p(i-1,j)+p(i,j-1));
    end
```

```
end
b=0.5*sum(p(:,n-d));  % probability of Dwayne winning
fprintf('  probability: %f (compare to %f)\n',b,a)
fprintf('  Dwayne gets %d cents\n',round(200*b))
```

4. *A fair m-sided die (with sides numbered 1 through m) is tossed until getting "1" k times (not necessarily in a row). Define random variable N as the total number of tosses required (unless "1" never happens for the kth time, in which case we set $N = -1$). Find the PMF of N.*

   Answer: $p_N(n) = \binom{n-1}{k-1}\frac{(m-1)^{n-k}}{m^n}$ for $n \in \{k, k+1, k+2, \dots\}$ (and $\mathbf{P}(N = -1) = 0$).

   Solution: reaching k "1"-s at step n means getting "1" at step n, and getting $k-1$ "1"-s in the first $n-1$ steps. The probability of this is

$$\frac{1}{m} \cdot \binom{n-1}{k-1} \cdot \left(\frac{1}{m}\right)^{k-1} \left(\frac{m-1}{m}\right)^{n-k} = p_N(n).$$

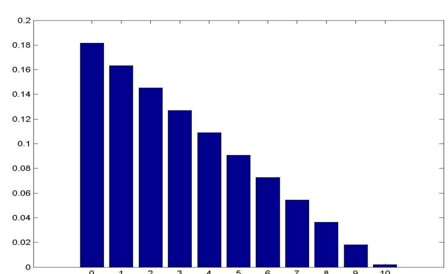
   To show that $\mathbf{P}(N = -1) = 0$, note that the probability of not getting exactly r "1"-s in the first n steps equals $\binom{n}{r}(m-1)^{n-r}m^{-n}$, and has an upper bound

$$\binom{n}{r}(m-1)^{n-r}m^{-n} \le \frac{1}{r!(m-1)^r} n^r \left(\frac{m-1}{m}\right)^n$$

   which converges to zero as $n \to \infty$ (because every increasing exponent outgrows every polynomial), the probability of not stopping in n trials converges to zero as $n \to \infty$.

5. *Mary, Tom, and Jerry park their cars in a parking lot that consists of $N > 2$ parking spaces, arranged in a circular fashion and numbered sequentially clock-wise from 1 to N, so that, for example, parking space number 1 is next to parking spaces number 2 and N. Assume that each possible triplet of (different) parking locations is equally likely. Let X be the random variable (taking integer values between 0 and $\frac{N-3}{3}$) defined as the minimum of the three non-negative integers representing the numbers of parking spaces between the cars of Mary, Tom, and Jerry. For example, if $N = 100$ and Mary, Tom, and Jerry have parked their cars in slots number 2, 20, and 97 respectively, then $X = 4$ is the number of parking slots (98, 99, 100, and 1 if listed clock-wise) between the cars of Mary and Jerry. Find the probability mass function of X, and sketch it graphically.*

   Answer: $p_X(x) = \frac{6(N-3x-3)}{(N-1)(N-2)}$ for $x < \frac{N-3}{3}$. When N is divisible by 3, and $x = \frac{N-3}{3}$, $p_X(x) = \frac{2}{(N-1)(N-2)}$. The sample plot is for $N = 33$.



Here is a MATLAB function which checks that the PMF values sum up to 1, and produces the plot.

```
function ps03_5(N)
% function ps03_5(N)
%
% sanity check for ps03.5

if nargin<1, N=33; end

n=floor(N/3);
x=0:(n-1);
p=(6/((N-1)*(N-2)))*(N-3-3*x);
if N==3*n,
    p(n)=2/((N-1)*(N-2));
end

fprintf(' check: 1=%f\n',sum(p))
close(gcf);bar(x,p);
```

Solution: the sample space Ω in this problem can be viewed as the set of assignments of three letters "M", "T", and "J" to three different locations among N points uniformly spaced on a circle. The total number of such assignments is $N(N-1)(N-2)$. Depending on how many times the minimal number of spaces between the three cars is achieved, the sample space can be partitioned into two subsets (note that X in these definitions is the random variable defined in the problem formulation, *not* a parameter):

A_{12}: there is a pair of cars with more than X parking spaces between them;

A_3: all three pairs of cars have X parking spaces between them.

Note that event A_{12} is contained in the event $3X+3<N$, while A_3 is the same set as the event $3X+3=N$ (in particular, it is empty unless N is divisible by 3).

To select a car parking arrangement in A_{12} with $X=x$, one chooses the order in which the cars are parked, starting with the first (clockwise) of the cars which have a neighbor at distance X (6 possibilities), the position of the first (clockwise) car (N possibilities), and, finally, the position of the remaining car ($N-3x-3$ possibilities) which proves that the set $A_{12}\cap(X=x)$ has $6N(N-3x-3)$ elements.

To select an element in A_3, one must assume that N is divisible by 3, and choose the clockwise order of the cars (2 possibilities), as well as the position of Mary's car (N possibilities), which proves that the set A_3 has $2N$ elements.

6. *The PMF of random variable X is given by $p_X(k)=\frac{2k}{n^2+n}$ for $k\in\{1,2,\ldots,n\}$. Find the PMF of $Y=f(X)$, where $f(x)=\min\{x,n-x\}$.*

   Answer: Y takes integer values from the interval $[0,n/2]$, and

$$p_Y(y)=\begin{cases}\frac{2}{n+1}, & y<n/2,\\ \frac{1}{n+1}, & y=n/2\ (n\text{ is even}).\end{cases}$$

Solution: each possible integer value $y < n/2$ of Y corresponds to two possible values ($x = y$ and $x = n - y$) of X (assuming for convenience that X takes value $x = 0$ with probability $0 = \frac{2x}{n^2+n}$). The only other possible value of Y is $y = n/2$ (corresponds to a single value $x = n/2$, and only when n is even).

7. [**6.431 question**] *X and Y are random variables with probability mass functions such that $p_X(0) = p_Y(0) = 1/3$, and $p_X(1) = p_Y(1) = 2/3$. Random variable Z is defined by $Z = X - Y$.*

   (a) *What are the possible values of $p_Z(0)$?*

   Answer: $p_Z(0) \in [1/3, 1]$.

   Solution: for $i, j \in \{0, 1\}$ let A_{ij} be the event "$X = i$, $Y = j$". Let $p_{ij} = \mathbf{P}(A_{ij})$. The problem formulation states that

$$p_{00} + p_{01} = \frac{1}{3}, \quad p_{11} + p_{01} = \frac{2}{3}, \quad p_{00} + p_{01} + p_{10} + p_{11} = 1.$$

   This means that every probability is uniquely defined by $r = p_{01}$:

$$p_{00} = \frac{1}{3} - r, \; p_{11} = \frac{2}{3} - r, \; p_{10} = r.$$

   Hence $r \in [0, 1/3]$, and

$$p_Z(0) = p_{00} + p_{11} = 1 - 2r \in [1/3, 1].$$

   To verify explicitly that all values $r = p_{01} \in [0, 1/3]$ are possible, construct the sample space $\Omega = \{0,1\} \times \{0,1\} = \{(i,j) : \; i,j \in \{0,1\}\}$, and define the probability function on the all subsets of Ω by setting the single point set probabilities

$$\mathbf{P}(\{(0,0)\}) = \frac{1}{3} - r, \; \mathbf{P}(\{(1,1)\}) = \frac{2}{3} - r, \; \mathbf{P}(\{(0,1)\}) = \mathbf{P}(\{(1,0)\}) = r.$$

   This definition satisfies all axioms of probability for all $r \in [0, 1/3]$, which shows that all values in the interval $[1/3, 1]$ are possible for $p_Z(0)$.

   (b) *Assuming $q = p_Z(0)$ is given, what is the PMF of Z?*

   Answer: Z takes values from the set $\{-1, 0, 1\}$, with probabilities $0.5(1-q)$, q, and $0.5(1-q)$ respectively.

   Solution: by assumption, the only possible values for X and Y are 0 or 1. The relation between the probabilities of the events "$X = i$, $Y = j$" are computed in the solution of part (a).