MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2008)

Problem Set 4: Solutions [1]
Due: October 1, 2008

1. We are given the following information:

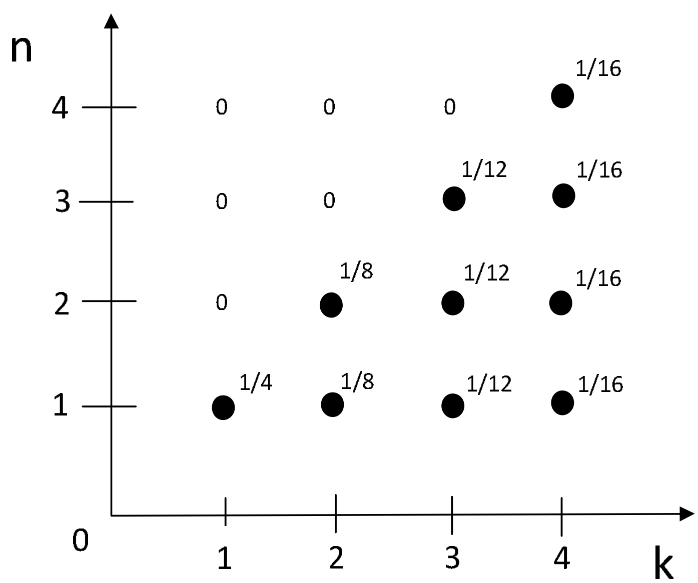
$$p_K(k) = \begin{cases} 1/4, & \text{if } k = 1,2,3,4; \\ 0, & \text{otherwise} \end{cases}$$

$$p_{N|K}(n \mid k) = \begin{cases} 1/k, & \text{if } n = 1,\ldots,k; \\ 0, & \text{otherwise} \end{cases}$$

   (a) We use the fact that $p_{N,K}(n,k) = p_{N|K}(n \mid k)p_K(k)$ to arrive at the following joint PMF:

$$p_{N,K}(n,k) = \begin{cases} 1/(4k), & \text{if } k = 1,2,3,4 \text{ and } n = 1,\ldots,k; \\ 0, & \text{otherwise} \end{cases}$$

   The joint PMF $p_{N,K}(n,k)$ is plotted below.



   (b) The marginal PMF $p_N(n)$ is given by the following formula:

$$p_N(n) = \sum_k p_{N,K}(n,k) = \sum_{k=n}^{4} \frac{1}{4k}$$

   On simplification this yields

$$p_N(n) = \begin{cases} 1/4 + 1/8 + 1/12 + 1/16 = 25/48, & n = 1; \\ 1/8 + 1/12 + 1/16 = 13/48, & n = 2; \\ 1/12 + 1/16 = 7/48, & n = 3; \\ 1/16 = 3/48, & n = 4; \\ 0, & \text{otherwise.} \end{cases}$$

[1] Published September 23, 2008

(c) The conditional PMF is

$$p_{K|N}(k \mid 2) = \frac{p_{N,K}(2,k)}{p_N(2)} = \begin{cases} 6/13, & k=2; \\ 4/13, & k=3; \\ 3/13, & k=4; \\ 0, & \text{otherwise.} \end{cases}$$

(d) Let A be the event $2 \le N \le 3$. We first find the conditional PMF of K given A.

$$\begin{aligned} p_{K|A}(k) &= \frac{\mathbf{P}(K=k,A)}{\mathbf{P}(A)} \\ \mathbf{P}(A) &= p_N(2)+p_N(3) = \frac{5}{12} \\ \mathbf{P}(K=k,A) &= \begin{cases} \frac{1}{8}, & k=2; \\ \frac{1}{12}+\frac{1}{12}, & k=3; \\ \frac{1}{16}+\frac{1}{16}, & k=4; \\ 0, & \text{otherwise} \end{cases} \\ p_{K|A}(k) &= \begin{cases} \frac{3}{10}, & k=2; \\ \frac{2}{5}, & k=3; \\ \frac{3}{10}, & k=4; \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Because the conditional PMF of K given A is symmetric around $k = 3$, we know $\boxed{\mathbf{E}[K \mid A] = 3}$. We now find the conditional variance of K given A.

$$\begin{aligned} var(K \mid A) &= \mathbf{E}[(K - \mathbf{E}[K \mid A])^2 \mid A] \\ &= \frac{3}{10}\cdot(2-3)^2 + \frac{2}{5}\cdot 0 + \frac{3}{10}\cdot(4-3)^2 \\ &= \boxed{\tfrac{3}{5}} \end{aligned}$$

2. (a) Use the total probability theorem by conditioning on the number of questions that Professor Right has to answer. Let A be the event that she gives all wrong answers in a given lecture, let B_1 be the event that she gets one question in a given lecture, and let B_2 be the event that she gets two questions in a given lecture. Then

$$\mathbf{P}(A) = \mathbf{P}(A|B_1)\mathbf{P}(B_1) + \mathbf{P}(A|B_2)\mathbf{P}(B_2).$$

From the problem statement, she is equally likely to get one or two questions in a given lecture, so $\mathbf{P}(B_1) = \mathbf{P}(B_2) = \frac{1}{2}$. Also, from the problem statement, $\mathbf{P}(A|B_1) = \frac{1}{4}$, and, because of independence, $\mathbf{P}(A|B_2) = (\frac{1}{4})^2 = \frac{1}{16}$. Thus we have

$$\mathbf{P}(A) = \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{16}\cdot\frac{1}{2} = \frac{5}{32}.$$

(b) Let events A and B_2 be defined as in the previous part. Using Bayes's Rule:

$$\mathbf{P}(B_2|A) = \frac{\mathbf{P}(A|B_2)\mathbf{P}(B_2)}{\mathbf{P}(A)}.$$

From the previous part, we said $\mathbf{P}(B_2) = \frac{1}{2}$, $\mathbf{P}(A|B_2) = \frac{1}{16}$, and $\mathbf{P}(A) = \frac{5}{32}$. Thus

$$\mathbf{P}(B_2|A) = \frac{\frac{1}{16} \cdot \frac{1}{2}}{\frac{5}{32}} = \frac{1}{5}.$$

As one would expect, given that Professor Right answers all the questions in a given lecture incorrectly, it's more likely that she got only one question rather than two.

(c) We start by finding the PMFs for X and Y. The PMF $p_X(x)$ is given from the problem statement:

$$p_X(x) = \begin{cases} \frac{1}{2}, & if x \in \{1, 2\} \\ 0, & otherwise. \end{cases}$$

The PMF for Y can be found by conditioning on X for each value that Y can take on. Because Professor Right can be asked at most two questions in any lecture, the range of Y is from 0 to 2. Looking at each possible value of Y, we find

$$p_Y(0) = \mathbf{P}(Y = 0|X = 1)\mathbf{P}(X = 1) + \mathbf{P}(Y = 0|X = 2)\mathbf{P}(X = 2) = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{16} \cdot \frac{1}{2} = \frac{5}{32},$$

$$p_Y(1) = \mathbf{P}(Y = 1|X = 1)\mathbf{P}(X = 1) + \mathbf{P}(Y = 1|X = 2)\mathbf{P}(X = 2) = \frac{3}{4} \cdot \frac{1}{2} + 2 \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{9}{16},$$

$$p_Y(2) = \mathbf{P}(Y = 2|X = 1)\mathbf{P}(X = 1) + \mathbf{P}(Y = 2|X = 2)\mathbf{P}(X = 2) = 0 \cdot \frac{1}{2} + (\frac{3}{4})^2 \cdot \frac{1}{2} = \frac{9}{32}.$$

Note that when calculating $\mathbf{P}(Y = 1|X = 2)$, we got $2 \cdot \frac{3}{4} \cdot \frac{1}{4}$ because there are two ways for Professor Right to answer one question right when she's asked two questions: either she answers the first question correctly or she answers the second question correctly. Thus, overall

$$p_Y(y) = \begin{cases} 5/32, & \text{if } y = 0; \\ 9/16, & \text{if } y = 1; \\ 9/32, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

Now the mean and variance can be calculated explicitly from the PMFs:

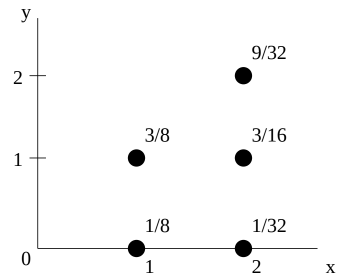
$$\mathbf{E}[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2},$$

$$\text{var}(X) = \left(1 - \frac{3}{2}\right)^2 \frac{1}{2} + \left(2 - \frac{3}{2}\right)^2 \frac{1}{2} = \frac{1}{4},$$

$$\mathbf{E}[Y] = 0 \cdot \frac{5}{32} + 1 \cdot \frac{9}{16} + 2 \cdot \frac{9}{32} = \frac{9}{8},$$

$$\text{var}(Y) = \left(0 - \frac{9}{8}\right)^2 \frac{5}{32} + \left(1 - \frac{9}{8}\right)^2 \frac{9}{16} + \left(2 - \frac{9}{8}\right)^2 \frac{9}{32} = \frac{27}{64}.$$

(d) The joint PMF $p_{X,Y}(x, y)$ is plotted below. There are only five possible (x, y) pairs. For each point, $p_{X,Y}(x, y)$ was calculated by $p_{X,Y}(x, y) = p_X(x)p_{Y|X}(y|x)$.



(e) By linearity of expectations,

$$\mathbf{E}[Z] = \mathbf{E}[X + 2Y] = \mathbf{E}[X] + 2\mathbf{E}[Y] = \frac{3}{2} + 2 \cdot \frac{9}{8} = \frac{15}{4}.$$

Calculating $\text{var}(Z)$ is a little bit more tricky because X and Y are not independent; therefore we *cannot* add the variance of X to the variance of $2Y$ to obtain the variance of Z. (X and Y are clearly not independent because if we are told, for example, that $X = 1$, then we know that Y cannot equal 2, although normally without any information about X, Y could equal 2.)

To calculate $\text{var}(Z)$, first calculate the PMF for Z from the joint PDF for X and Y. For each (x, y) pair, we assign a value of Z. Then for each value z of Z, we calculate $p_Z(z)$ by summing over the probabilities of all (x, y) pairs that map to z. Thus we get

$$p_Z(z) = \begin{cases} 1/8, & \text{if } z = 1; \\ 1/32, & \text{if } z = 2; \\ 3/8, & \text{if } z = 3; \\ 3/16, & \text{if } z = 4; \\ 9/32, & \text{if } z = 6; \\ 0, & \text{otherwise.} \end{cases}$$

In this example, each (x, y) mapped to exactly one value of Z, but this does not have to be the case in general. Now the variance can be calculated as:

$$\text{var}(Z) = \frac{1}{8}\left(1 - \frac{15}{4}\right)^2 + \frac{1}{32}\left(2 - \frac{15}{4}\right)^2 + \frac{3}{8}\left(3 - \frac{15}{4}\right)^2 + \frac{3}{16}\left(4 - \frac{15}{4}\right) + \frac{9}{32}\left(6 - \frac{15}{4}\right)^2 = \frac{43}{16}.$$

(f) Let Y be the number of questions she will answer wrong in a randomly chosen lecture. We can find $\mathbf{E}[Y]$ by conditioning on whether the lecture is in math or in science. Let M be the event that the lecture is in math, and let S be the event that the lecture is in science. Then

$$\mathbf{E}[Y] = \mathbf{E}[Y|M]\mathbf{P}(M) + \mathbf{E}[Y|S]\mathbf{P}(S).$$

Since there are an equal number of math and science lectures and we are choosing randomly among them, $\mathbf{P}(M) = \mathbf{P}(S) = \frac{1}{2}$. Now we need to calculate $\mathbf{E}[Y|M]$ and $\mathbf{E}[Y|S]$ by finding the respective conditional PMFs first. The PMFs can be determined in an manner analagous to how we calculated the PMF for the number of correct answers in part (c).

$$p_{Y|S}(y) = \begin{cases} \frac{1}{2} \cdot \frac{3}{4} + \frac{1}{2}(\frac{3}{4})^2 = 21/32, & \text{if } y = 0; \\ \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot 2 \cdot \frac{1}{4} \cdot \frac{3}{4} = 5/16, & \text{if } y = 1; \\ \frac{1}{2} \cdot 0 + \frac{1}{2}(\frac{1}{4})^2 = 1/32, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

$$p_{Y|M}(y) = \begin{cases} \frac{1}{2}\cdot\frac{9}{10} + \frac{1}{2}(\frac{9}{10})^2 = 171/200, & \text{if } y = 0; \\ \frac{1}{2}\cdot\frac{1}{10} + \frac{1}{2}\cdot 2\cdot\frac{1}{10}\cdot\frac{9}{10} = 7/50, & \text{if } y = 1; \\ \frac{1}{2}\cdot 0 + \frac{1}{2}(\frac{1}{10})^2 = 1/200, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

Therefore

$$\mathbf{E}[Y|S] = 0\cdot\frac{21}{32} + 1\cdot\frac{5}{16} + 2\cdot\frac{1}{32} = \frac{3}{8},$$

$$\mathbf{E}[Y|M] = 0\cdot\frac{171}{200} + 1\cdot\frac{7}{50} + 2\cdot\frac{1}{200} = \frac{3}{20}.$$

This implies that

$$\mathbf{E}[Y] = \frac{3}{20}\cdot\frac{1}{2} + \frac{3}{8}\cdot\frac{1}{2} = \frac{21}{80}.$$

3. All possible outcomes are:
(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (2, 4),
(3, 1), (3, 2), (3, 3), (3, 4), (4, 1), (4, 2), (4, 3), and (4, 4)

Given that the sum of the down-face values is greater than the product of the down-face values, our universe is restricted to:
(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (3, 1), and (4, 1)
Then we have that:

$$p_X(x) = \begin{cases} \frac{1}{7} & \text{if } x = 1 \\ \frac{2}{7} & \text{if } x = 2 \\ \frac{2}{7} & \text{if } x = 3 \\ \frac{2}{7} & \text{if } x = 4 \\ 0 & \text{otherwise} \end{cases}$$

Let $Z = X^2$. Then we have:

$$p_Z(z) = \begin{cases} \frac{1}{7} & \text{if } z = 1 \\ \frac{2}{7} & \text{if } z = 4 \\ \frac{2}{7} & \text{if } z = 9 \\ \frac{2}{7} & \text{if } z = 16 \\ 0 & \text{otherwise} \end{cases}$$

$E[Z] = \frac{1}{7}(1) + \frac{2}{7}(4) + \frac{2}{7}(9) + \frac{2}{7}(16) = \frac{59}{7}$.
$Var(Z) = \frac{1}{7}(\frac{52}{7})^2 + \frac{2}{7}(\frac{31}{7})^2 + \frac{2}{7}(\frac{-4}{7})^2 + \frac{2}{7}(\frac{-53}{7})^2 = \frac{1468}{49} \approx 29.96$.

4. Define the following two random variables:

$$\begin{aligned} Y &= \text{number of undergraduate students who get an A} \\ Z &= \text{number of graduate students who get an A} \end{aligned}$$

Each of these is a binomial random variable. Their PMFs are:

$$\begin{aligned} p_Y(y) &= \binom{250}{y}\left(\frac{1}{3}\right)^y\left(\frac{2}{3}\right)^{250-y} \\ p_Z(z) &= \binom{50}{z}\left(\frac{1}{2}\right)^{50} \end{aligned}$$

(a) We wish to find the PMF of $X = Y + Z$.

$$\begin{aligned} p_X(x) &= \sum_{(y,z)|(y+z)=x} p_{Y,Z}(y,z) \\ &= \sum_{(y,z)|(y+z)=x} p_Y(y)p_Z(z) \qquad (Y \text{ and } Z \text{ are independent}) \\ &= \begin{cases} \sum_{i=0}^{\min(x,50)} \binom{50}{i}\left(\frac{1}{2}\right)^{50}\binom{250}{x-i}\left(\frac{1}{3}\right)^{x-i}\left(\frac{2}{3}\right)^{250-x+i}, & 0 \le x \le 250; \\ \sum_{i=x-250}^{50} \binom{50}{i}\left(\frac{1}{2}\right)^{50}\binom{250}{x-i}\left(\frac{1}{3}\right)^{x-i}\left(\frac{2}{3}\right)^{250-x+i}, & 251 \le x \le 300; \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Notice how closely the PMF of X defined as a function of the PMF of Y and Z resembles the convolution operation.

(b) Define random variables $X_1, X_2, \ldots, X_{300}$ by

$$X_i = \begin{cases} 1, & \text{if student } i \text{ gets A;} \\ 0, & \text{if student } i \text{ does not get A.} \end{cases}$$

Then $X = X_1 + X_2 + \cdots + X_{300}$, and since $\mathbf{E}[X] = \sum_{i=1}^{300} \mathbf{E}[X_i]$, the problem reduces to computing $\mathbf{E}[X_i]$.

The computation of $\mathbf{E}[X_i]$ is simplified by the total expectation theorem. Consider the numbering of the students to be random, irrespective of the student being an undergraduate or graduate. To have short expressions in the following, let A_i be the event that student i is an undergraduate. Then

$$\mathbf{E}[X_i] = \mathbf{E}[X_i \mid A_i]\mathbf{P}(A_i) + \mathbf{E}[X_i \mid A_i^c]\mathbf{P}(A_i^c) = \frac{1}{3}\cdot\frac{250}{300} + \frac{1}{2}\cdot\frac{50}{300} = \frac{13}{36}$$

and $\mathbf{E}[X] = 300\mathbf{E}[X_i] = 325/3$.

(c) $\mathbf{E}[W] = E[X] + 2 = 325/3 + 2 = 331/3$

5. The problem statement says that a topping i is added to a pizza independently of all other toppings and the toppings on all other pizzas. Therefore, the number of pizzas with a certain topping i is independent of the number of pizzas with another topping j, and we may say that:

$$P_{N_1,N_2,N_3,N_4}(n_1,n_2,n_3,n_4) = P_{N_1}(n_1)P_{N_2}(n_2)P_{N_3}(n_3)P_{N_4}(n_4)$$

Now we have to find the distribution $P_{N_i}(n_i)$ for each topping i. The distribution will be binomial where we equate each of the n pizzas with an independent trial and p_i with the probability of success for each trial. Because $p_i = 2^{-i}$, we get:

$$P_{N_i}(n_i) = \binom{n}{n_i} 2^{-in_i}(1-2^{-i})^{n-n_i}, 0 \le n_i \le n$$

and

$$P_{N_1,N_2,N_3,N_4}(n_1,n_2,n_3,n_4) = \prod_{i=1}^{i=4} \binom{n}{n_i} 2^{-in_i}(1-2^{-i})^{n-n_i}, 0 \le n_1,n_2,n_3,n_4 \le n$$

G1†. Let X_1, X_2, ..., X_n be independent, identically distributed (IID) random variables.

We note that

$$E[X_1 + \cdots + X_n | X_1 + \cdots + X_n = x_0] = x_0.$$

It follows from the linearity of expectations that

$$\begin{aligned} x_0 &= E[X_1 + \cdots + X_n | X_1 + \cdots + X_n = x_0] \\ &= E[X_1 | X_1 + \cdots + X_n = x_0] + \cdots + E[X_n | X_1 + \cdots + X_n = x_0] \end{aligned}$$

Because the X_i's are identically distributed, we have the following relationship.

$$E[X_i | X_1 + \cdots + X_n = x_0] = E[X_j | X_1 + \cdots + X_n = x_0], \text{ for any } 1 \le i \le n, 1 \le j \le n.$$

Therefore,

$$\begin{aligned} nE[X_1 | X_1 + \cdots + X_n = x_0] &= x_0 \\ E[X_1 | X_1 + \cdots + X_n = x_0] &= \frac{x_0}{n}. \end{aligned}$$

G2†. (a) For each $i = 1, 2, \ldots, n$, let X_i be the number of beans in jar i.

The crucial observation is that X_i has a binomial PMF. Consider each bean. It chooses a jar independently and uniformly. Thus, the probability the bean lands in jar i is $\frac{1}{n}$. There are m jelly beans, so X_i is distributed like a binomial random variable with parameters m and $\frac{1}{n}$. Thus, we obtain

$$\begin{aligned} p_{X_i}(k) &= \binom{m}{k}\left(\frac{1}{n}\right)^k \left(1-\frac{1}{n}\right)^{m-k}, k = 0, 1, \ldots, m \\ \mathbf{E}[X_i] &= \frac{m}{n} \end{aligned}$$

(b) For each $k = 0, 1, \ldots, m$, let Y_k be the number of jars that have exactly k beans.

Define indicator random variables $I_1, I_2, \ldots, I_n$ as follows: I_i is 1 if jar i has exactly k beans, and 0 otherwise. With this definition, $Y_k = \sum_{i=1}^n I_i$. By linearity of expectation, we see that $\mathbf{E}[Y_k] = \sum_{i=1}^n \mathbf{E}[I_i]$.

To calculate $\mathbf{E}[I_i]$, note that $\mathbf{E}[I_i] = \mathbf{P}(X_i = k) = p_{X_i}(k)$. From (a), we know that $p_{X_i}(k) = \binom{m}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{m-k}$. This means $\mathbf{E}[Y_k] = n\cdot\binom{m}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{m-k}$ and $\mathbf{E}[Y_0] = n \cdot \left(1-\frac{1}{n}\right)^m$.

(c) To determine the desired probability, we will compute $1 - \mathbf{P}$(some jar is empty). Define events $A_i, i = 1, 2, \ldots, n$ such that A_i is the event $\{X_i = 0\}$, i.e. that jar i is empty. Then, $\mathbf{P}$(some jar is empty) $= \mathbf{P}(A_1 \cup A_2 \ldots \cup A_n)$. We will use the *inclusion-exclusion* principle to calculate the probability of the union of the A_i (see chapter 1, problem 12 for a detailed discussion of the inclusion-exclusion principle).

The inclusion-exclusion formula states

$$\mathbf{P}(A_1 \cup A_2 \ldots \cup A_n) = \sum_{i=1}^{n} \mathbf{P}(A_i) - \sum_{1 \le i < j \le n} \mathbf{P}(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} \mathbf{P}(A_i \cap A_j \cap A_k) +$$

$$\ldots + (-1)^{n-1} \mathbf{P}(A_1 \cap A_2 \ldots \cap A_n).$$

Let us calculate $\mathbf{P}(A_1 \cap A_2 \ldots \cap A_k)$ for any $k \le n$. This intersection represents the event that the first k jars are all empty. The probability the first k jars are empty is simply the probability that every bean misses these k jars, which is $\left(1 - \frac{k}{n}\right)^m$. By symmetry, this formula works for any fixed set of k jars. Plugging this into the inclusion-exclusion formula, we get

$$\mathbf{P}(A_1 \cup A_2 \ldots \cup A_n) = \sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^m - \sum_{1 \le i < j \le n} \left(1 - \frac{2}{n}\right)^m + \sum_{1 \le i < j < k \le n} \left(1 - \frac{3}{n}\right)^m +$$

$$\ldots + (-1)^{n-1} \left(1 - \frac{n}{n}\right)^m.$$

To simplify this expression, consider the first sum. There are n terms in the sum, so the first sum is just $n\left(1 - \frac{1}{n}\right)^m$. In the next sum, there are $\binom{n}{2}$ terms. In general, the k^{th} sum has $\binom{n}{k}$ terms. Thus, we get

$$\mathbf{P}(A_1 \cup A_2 \ldots \cup A_n) = n\left(1 - \frac{1}{n}\right)^m - \binom{n}{2}\left(1 - \frac{2}{n}\right)^m + \binom{n}{3}\left(1 - \frac{3}{n}\right)^m + \ldots$$

$$+ \binom{n}{n}(-1)^{n-1}\left(1 - \frac{n}{n}\right)^m.$$

We subtract this probability from 1 to get the final answer,

$\mathbf{P}$(every jar receives at least one jelly bean)

$$\begin{aligned}
&= 1 - \mathbf{P}(\text{some jar is empty}) \\
&= 1 - \mathbf{P}(A_1 \cup A_2 \ldots \cup A_n) \\
&= 1 - \left(n\left(1 - \frac{1}{n}\right)^m - \binom{n}{2}\left(1 - \frac{2}{n}\right)^m + \ldots + \binom{n}{n}(-1)^{n-1}\left(1 - \frac{n}{n}\right)^m\right) \\
&= \sum_{k=0}^{n-1} (-1)^k \binom{n}{k}\left(1 - \frac{k}{n}\right)^m
\end{aligned}$$

(We can leave out the $k = n$ term since that term is always 0.)