

1 Poisson approximation

a)

We assume that each ball is unique (each one holds a unique identifier), so the total number of different orders to put $3 \cdot n$ balls in n bins are the permutation of them, so $(3 \cdot n)!$. In this case, we care only for those different orders that a bin receives an exact triplet of balls in it. The total number of those is $(3!)^n$, since we care about all the exact 3-way permutations that could happen in each one of those n bins. So, the total number of different orders in our case is $\frac{(3 \cdot n)!}{(3!)^n}$. In addition, the probability that one of those $\frac{(3 \cdot n)!}{(3!)^n}$ different orders is chosen given that we pick a ball uniformly and independently is $(\frac{1}{n})^{3n}$. So, the requested probability is:

$$p_1 = \frac{(3 \cdot n)!}{(3!)^n} \cdot \left(\frac{1}{n}\right)^{3n}$$

b)

We define Y_i as the total number of balls i^{th} bin received. In "Poisson world" $Y_i \sim \text{Poisson}(\mu)$, where $\mu = \frac{3 \cdot n}{n} = 3$, so $Y_i \sim \text{Poisson}(3)$. Given that by definition $Pr[Y_i = 3] = \frac{e^{-3} \cdot 3^3}{3!} = \frac{e^{-3} \cdot 3^2}{2}$ and all the Y_i are mutually independent we imply that:

$$p_2 = (Pr[Y_i = 3])^n = \left(\frac{e^{-3} \cdot 3^2}{2}\right)^n = \frac{e^{-3 \cdot n} \cdot 3^{2 \cdot n}}{2^n}$$

c)

Based on results from subquestions (a) and (b) we can easily imply that:

$$\frac{p_2}{p_1} = \frac{\frac{e^{-3 \cdot n} \cdot 3^{2 \cdot n}}{2^n}}{\frac{(3 \cdot n)!}{(3!)^n} \cdot \left(\frac{1}{n}\right)^{3n}} = \frac{3^{-3n} \cdot (3n)^{3n}}{(3n)!} \quad (1)$$

Using the definition of Poisson probability we have from equation (1) that $Pr[Y_i = 3n] = \frac{e^{-3n} \cdot 3^{3n}}{(3n)!} = \frac{p_2}{p_1}$

d)

The theorem 5.6 refers that: "The distribution of $(Y_1^{(m)}, \dots, Y_n^{(m)})$ conditioned on $\sum_i Y_i^{(m)} = k$ is the same as $(X_1^{(m)}, \dots, X_n^{(m)})$, regardless of the value of m ."

In our case, $k = 3n$. We define as X_i the binomial random variable of having exactly 3 balls in i^{th} bin. So we can easily imply that based on subquestion (a) $p_1 = Pr[(X_1^{(m)} = 3) \cap \dots \cap (X_n^{(m)} = 3)]$, and based on subquestion (b) $p_2 = Pr[(Y_1^{(m)} = 3) \cap \dots \cap (Y_n^{(m)} = 3)]$. Using the theorem 5.6 we know that:

$$p_1 = Pr[(X_1^{(m)} = 3) \cap \dots \cap (X_n^{(m)} = 3)] = Pr[(Y_1^{(m)} = 3) \cap \dots \cap (Y_n^{(m)} = 3) | \sum_{i=1}^n Y_i^{(m)} = 3n] \quad (2)$$

Based on one discussion problem we solved, the sum of n Poisson distribution random variables is a Poisson distribution random variable with $\mu = \sum_{i=1}^n \mu_i$. In our case $Y = \sum_{i=1}^n Y_i^{(m)}$ with $\mu_Y = \sum_{i=1}^n \mu_{Y_i}$. Using this result and the definition of conditional probabilities we can imply that equation (2) will be:

$$\begin{aligned}
 p_1 &= Pr[(Y_1^{(m)} = 3) \cap \dots \cap (Y_n^{(m)} = 3) | \sum_{i=1}^n Y_i^{(m)} = 3n] \implies \\
 p_1 &= Pr[(Y_1^{(m)} = 3) \cap \dots \cap (Y_n^{(m)} = 3) | Y = 3n] \implies \\
 p_1 &= \frac{Pr[(Y_1^{(m)} = 3) \cap \dots \cap (Y_n^{(m)} = 3) \cap (Y = 3n)]}{Pr[Y = 3n]} \implies \\
 p_1 &= \frac{Pr[(Y_1^{(m)} = 3) \cap \dots \cap (Y_n^{(m)} = 3)]}{Pr[Y = 3n]} \implies \\
 p_1 &= \frac{p_2}{Pr[Y = 3n]} \implies \frac{p_2}{p_1} = Pr[Y = 3n]
 \end{aligned}$$

2 Poisson diagnostics

We assume that A,B and C are independent Poisson random variables with means $p_1\mu, p_2\mu$ & $p_3\mu$ respectively. Given that this is true, in order to prove that X,Y and Z are independant Poisson random variables we have to show that the joint distribution of A,B and C is the same with the joint distribution of X,Y and Z. Moreover, through discussions we have shown that the sum of Poisson random variables is a Poisson random variable with expected value equal to the sum of the respected random variables,so we define $D = X + Y + Z$. Also, we define the joint distribution of X,Y,Z as:

$$Pr[(X = x) \cap (Y = y) \cap (Z = z)] = Pr[(X = x) \cap (Y = y) \cap (D = x + y + z)] \quad (3)$$

The above expression implies that if $\{D = x + y + z\}$ event happens then $\{Z = z\}$ happens too. So equation (3) takes the following form based on product rule:

$$Pr(X, Y, Z) = Pr(\{X = x\} | \{Y = y\} \cap \{D = x + y + z\}) \cdot Pr(\{Y = y\} | \{D = x + y + z\}) \cdot Pr(\{D = x + y + z\}) \quad (4)$$

Due to definition of conditional probabilities and the idea that the Poisson approximation to Binomial distribution we have the following expressions for the equation (4):

$$\begin{aligned}
 Pr(X, Y, Z) &= Pr(\{X = x\} | \{Y = y\} \cap \{D = x + y + z\}) \cdot Pr(\{Y = y\} | \{D = x + y + z\}) \cdot Pr(\{D = x + y + z\}) \implies \\
 &= \binom{x+z}{x} \cdot \left(\frac{p_1}{1-p_2}\right)^x \cdot \left(\frac{1-p_1-p_2}{1-p_2}\right)^z \cdot \binom{x+y+z}{y} \cdot p_2^y \cdot (1-p_2)^{x+z} \cdot \frac{e^{-\mu} \cdot \mu^{x+y+z}}{(x+y+z)!} \implies \\
 &= \frac{(x+z)!}{x! \cdot z!} \cdot \left(\frac{p_1}{1-p_2}\right)^x \cdot \left(\frac{1-p_1-p_2}{1-p_2}\right)^z \cdot \frac{(x+y+z)!}{(x+z)! \cdot y!} \cdot p_2^y \cdot (1-p_2)^{x+z} \cdot \frac{e^{-\mu} \cdot \mu^{x+y+z}}{(x+y+z)!} \implies \\
 &= \frac{1}{x! \cdot z!} \cdot \left(\frac{p_1}{1-p_2}\right)^x \cdot \left(\frac{1-p_1-p_2}{1-p_2}\right)^z \cdot \frac{1}{y!} \cdot p_2^y \cdot (1-p_2)^{x+z} \cdot e^{-\mu} \cdot \mu^{x+y+z} \xrightarrow{1=p_3+p_2+p_1} \\
 &= \frac{1}{x! \cdot z!} \cdot p_1^x \cdot (1-p_1-p_2)^z \cdot \frac{1}{y!} \cdot p_2^y \cdot e^{-\mu \cdot (p_1+p_2+p_3)} \cdot \mu^{x+y+z} \xrightarrow{p_3=1-p_1+p_2} \\
 &= \frac{1}{x! \cdot z!} \cdot p_1^x \cdot p_3^z \cdot \frac{1}{y!} \cdot p_2^y \cdot e^{-\mu \cdot (p_1+p_2+p_3)} \cdot \mu^{x+y+z} = \frac{e^{-\mu \cdot p_1} \cdot (\mu p_1)^x}{x!} \cdot \frac{e^{-\mu \cdot p_2} \cdot (\mu p_2)^y}{y!} \cdot \frac{e^{-\mu \cdot p_3} \cdot (\mu p_3)^z}{z!} \implies \\
 &= Pr(\{A = a\}) \cdot Pr(\{B = b\}) \cdot Pr(\{C = c\}) = Pr(A, B, C)
 \end{aligned}$$

3 Poisson Fish

a)

We define the Poisson random variable $N \sim \text{Poisson}(s)$, which describes the number of samples it gets as input the algorithm A. We define the event $E_{fail}^{A'}$ as the event of A' algorithm's failure. Algorithm A' uses algorithm A as a black box. As the exercise describes the algorithm A' fails when the total number of samples feed in algorithm A exceed $2s$. So this probability is bounded as follows using the Chebyshev's inequality :

$$Pr(E_{fail}^{A'}) = Pr(N > 2s) = Pr(N - s > s) \leq Pr(|N - s| > s) \leq \frac{Var(N)}{s^2} = \frac{s}{s^2} = \frac{1}{s}$$

We define Y and Y' as random variables of the event of estimate number of species in the pond as a result from algorithm A and algorithm A'. If we define S to be every set of possible outputs S including "fail", then through law of total probability we can imply that:

$$Pr(Y \in S) = Pr(Y \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'}) + Pr(Y \in S | \tilde{E}_{fail}^{A'}) \cdot Pr(\tilde{E}_{fail}^{A'}) \quad (5)$$

Also applying the law of total probability to $Pr(Y' \in S)$ we have:

$$Pr(Y' \in S) = Pr(Y' \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'}) + Pr(Y' \in S | \tilde{E}_{fail}^{A'}) \cdot Pr(\tilde{E}_{fail}^{A'}) \quad (6)$$

At this point we have to point out that $Pr(Y' \in S | \tilde{E}_{fail}^{A'}) = Pr(Y \in S | \tilde{E}_{fail}^{A'})$ since if we draw less than $2s$ then algorithm A' does not fail and given that algorithm A' uses the algorithm A, that means that both probabilities describes the same event. Therefore, using the absolute value of the subtraction of the above equations (5) and (6), the factors including the above probabilities will be eliminated and leaving us as a result:

$$\begin{aligned} |Pr(Y \in S) - Pr(Y' \in S)| &= |Pr(Y \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'}) - Pr(Y' \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'})| \implies \\ &= Pr(E_{fail}^{A'}) \cdot |Pr(Y \in S | E_{fail}^{A'}) - Pr(Y' \in S | E_{fail}^{A'})| \end{aligned}$$

Based on the fact that the absolute difference between any probabilities could be at most 1 and thanks to $Pr(E_{fail}^{A'}) = \frac{1}{s}$ the above equation takes its final form:

$$|Pr(Y \in S) - Pr(Y' \in S)| = Pr(E_{fail}^{A'}) \cdot |Pr(Y \in S | E_{fail}^{A'}) - Pr(Y' \in S | E_{fail}^{A'})| \leq \frac{1}{s} < \frac{2}{s}$$

b)

In contrast to subquestion (a), the requested algorithm A' is a Poisson algorithm that uses a standardized algorithm A to satisfy the claim. So the algorithm A' is the following. We draw N samples from Poisson distribution and if $N < s$ then we fail since algorithm A takes exactly s samples in order to work. Otherwise, if $N \geq s$ then we give s samples to algorithm A to process. Based on this algorithm A', both algorithm A' and algorithm A satisfy the claim.

c)

We define the Poisson random variable $N \sim \text{Poisson}(2 \cdot s)$, which describes the number of samples it gets as input the algorithm A'. We define the event $E_{fail}^{A'}$ as the event of A' algorithm's failure. Algorithm A' uses algorithm A as a black box. As the subquestion (b) describes the algorithm A' fails when the total number of samples feed in algorithm A is less than s . So this probability is bounded as follows using the Chernoff's inequality for Poisson distributions:

$$Pr(E_{fail}^{A'}) = Pr(N < s) \leq \frac{e^{-\mu} \cdot (e \cdot \mu)^s}{s^s} = \frac{e^{-2 \cdot s} \cdot (e \cdot 2 \cdot s)^s}{s^s} = e^{-2 \cdot s} \cdot e^s \cdot 2^s = e^{-s} \cdot 2^s = \left(\frac{2}{e}\right)^s$$

We define Y and Y' as random variables of the event of estimate number of species in the pond as a result from algorithm A and algorithm A'. If we define S to be every set of possible outputs S including "fail", then through law of total probability we can imply that:

$$Pr(Y \in S) = Pr(Y \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'}) + Pr(Y \in S | \tilde{E}_{fail}^{A'}) \cdot Pr(\tilde{E}_{fail}^{A'}) \quad (7)$$

Also applying the law of total probability to $Pr(Y' \in S)$ we have:

$$Pr(Y' \in S) = Pr(Y' \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'}) + Pr(Y' \in S | \tilde{E}_{fail}^{A'}) \cdot Pr(\tilde{E}_{fail}^{A'}) \quad (8)$$

At this point we have to point out that $Pr(Y' \in S | \tilde{E}_{fail}^{A'}) = Pr(Y \in S | \tilde{E}_{fail}^{A'})$ since if we draw more than s then algorithm A' does not fail and given that algorithm A' uses the algorithm A, that means that both probabilities describes the same event. Therefore, using the absolute value of the subtraction of the above equations (7) and (8), the factors including the above probabilities will be eliminated and leaving us as a result:

$$\begin{aligned} |Pr(Y \in S) - Pr(Y' \in S)| &= |Pr(Y \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'}) - Pr(Y' \in S | E_{fail}^{A'}) \cdot Pr(E_{fail}^{A'})| \implies \\ &= Pr(E_{fail}^{A'}) \cdot |Pr(Y \in S | E_{fail}^{A'}) - Pr(Y' \in S | E_{fail}^{A'})| \end{aligned}$$

Based on the fact that the absolute difference between any probabilities could be at most 1 and thanks to $Pr(E_{fail}^{A'}) = \left(\frac{2}{e}\right)^s$ the above equation takes its final form:

$$|Pr(Y \in S) - Pr(Y' \in S)| = Pr(E_{fail}^{A'}) \cdot |Pr(Y \in S | E_{fail}^{A'}) - Pr(Y' \in S | E_{fail}^{A'})| \leq \left(\frac{2}{e}\right)^s \leq \frac{2}{s}$$

Given that $\frac{2}{e} < 1$ then or sufficient large number of s , the expression $\left(\frac{2}{e}\right)^s$ is going to be less than $\frac{2}{s}$.