1 Student Life

In an attempt to avoid having to do laundry often, Marcus comes up with a system. Every night, he designates one of his shirts as his dirtiest shirt. In the morning, he randomly picks one of his shirts to wear. If he picked the dirtiest one, he puts it in a dirty pile at the end of the day (a shirt in the dirty pile is not used again until it is cleaned). When Marcus puts his last shirt into the dirty pile, he finally does his laundry, and again designates one of his shirts as his dirtiest shirt (laundry isn't perfect) before going to bed. This process then repeats.

(a) If Marcus has n shirts, what is the expected number of days that transpire between laundry events? Your answer should be a function of n involving no summations.

(b) Say he gets even lazier, and instead of organizing his shirts in his dresser every night, he throws his shirts randomly onto one of n different locations in his room (one shirt per location), designates one of his shirts as his dirtiest shirt, and one location as the dirtiest location. In the morning, if he happens to pick the dirtiest shirt, *and* the dirtiest shirt was in the dirtiest location, then he puts the shirt into the dirty pile at the end of the day and does not throw any future shirts into that location and also does not consider it as a candidate for future dirtiest locations (it is too dirty). What is the expected number of days that transpire between laundry events now? Again, your answer should be a function of n involving no summations.

a) $X = X_1 + X_2 + \dots + X_n$ $\qquad X_i \sim \text{Geometric}\left(\frac{1}{n-i+1}\right)$

$$E[X] = E[X_1] + \dots + E[X_n]$$
$$= n + (n-1) + \dots + 2 + 1 = \frac{n(n+1)}{2}$$

b) $Y = Y_1 + Y_2 + \dots + Y_n$ $\qquad Y_i = \text{Geometric}\left(\frac{1}{(n-i+1)^2}\right)$

$$E[Y] = E[Y_1] + \dots + E[Y_n]$$
$$= n^2 + (n-1)^2 + \dots + 2^2 + 1^2 = \frac{n(n+1)(2n+1)}{6}$$

2 Planetary Party

(a) Suppose we are at party on a planet where every year is 2849 days. If 30 people attend this party, what is the exact probability that two people will share the same birthday? You may leave your answer as an unevaluated expression.

$$1 - \frac{2849}{2849} \cdot \frac{2848}{2849} \cdot \frac{2847}{2849} \cdot \dots \cdot \frac{2820}{2849} = 1 - \frac{2849!}{2819!\, 2849^{30}}$$

(b) From lecture, we know that given n bins and m balls, $\mathbb{P}[\text{no collision}] \approx \exp(-m^2/(2n))$. Using this, give an approximation for the probability in part (a).

$n = 2849$
$m = 30$ $\qquad P[\text{collision}] \approx 1 - e^{-\frac{30^2}{2 \cdot 2849}} \approx 0.854$

(c) What is the minimum number of people that need to attend this party to ensure that the probability that any two people share a birthday is at least 0.5? You can use the approximation you used in the previous part.

$$P[\text{no collisions}] \approx e^{-\frac{m^2}{2\cdot 2849}} \leq 0.5$$

$$-\frac{m^2}{2\cdot 2849} \leq \ln(0.5)$$

$$m \geq \sqrt{2\cdot 2849 \ln(0.5)}$$

$$m \geq 62.845 \qquad m \geq 63$$

(d) Now suppose that 70 people attend this party. What the is probability that none of these 70 individuals have the same birthday? You can use the approximation you used in the previous parts.

$$P[\text{no collision}] \approx e^{-\frac{70^2}{2\cdot 2849}} \approx 0.423$$

3 Throwing Balls into a Depth-Limited Bin

Say you want to throw n balls into n bins with depth $k-1$ (they can fit $k-1$ balls, after that the bins overflow). Suppose that n is a large number and $k = 0.1n$. You throw the balls randomly into the bins, but you would like it if they don't overflow. You feel that you might expect not too many balls to land in each bin, but you're not sure, so you decide to investigate the probability of a bin overflowing.

(a) Count the number of ways we can select k balls to put in the first bin, and then throw the remaining balls randomly. You should assume that the balls are distinguishable.

$$\binom{n}{k} n^{n-k}$$

(b) Argue that your answer in (a) is an upper bound for the number of ways that the first bin can overflow.

The expression overcounts the number of ways that more than k balls go into bin 1.

(c) Calculate an upper bound on the probability that the first bin will overflow.

$$\frac{\binom{n}{k} n^{n-k}}{n^n} = \frac{\binom{n}{k}}{n^k}$$

(d) Upper bound the probability that some bin will overflow. [*Hint*: Use the union bound.]

A_i = event that bin i overflows

$$P[A_1 \cup \ldots \cup A_n] \le \sum_{i=1}^{n} P[A_i] \le \frac{n\binom{n}{k}}{n^k}$$

(e) How does the above probability scale as n gets really large?

$$n\frac{\binom{n}{k}}{n^k} = n\frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{k!\,n^k} \le n\cdot\frac{n^k}{k!\,n^k} = \frac{n}{k!} = \frac{n}{(0.1n)!} \to 0$$

as $n \to \infty$