Hints, Assignment 2: CS 215, Fall 2024

Prof. Sunita Sarawagi

Released: 11 August, 2024
Hard Deadline: 11:59 PM, 13 September, 2024

§ 1C
The coefficient of z^n in $G(z)^k$ is $\sum_{i_1+i_2+\cdots+i_k=n} \prod_{j=1}^{k} P\left[X_j = i_j\right]$. Now, what is the probability that $X = n$?

§ 2D
How to simulate is simple; maintaining position and randomly incrementing or decrementing it h times works. How to do it *fast*?
`np.random.choice([-1, 1], shape=(?, ?)` can get you a lot of random numbers $\in \{1, -1\}$ quick. `np.sum` should be able to get you the final position x. There is a one line solution to simulate N balls to depth h composing the two `numpy` solutions.

§ 3C
The second moment of the data is the mean of the squares of the data values. Once you compute the values of the first and second moments (in terms of n and p) of $\text{Bin}(n, p)$, GPT can explain how `fsolve` can give you a best-fit n & p.

§ 3E
Should you take the log first or the mean/product? Take utmost care that your log likelihood function does not go to $-\infty$ if the probability of a particular data point is 0. Maybe add a `1e-8` appropriately?

§ 4D
(1) Expectation is linear. (2) You know from Q1A and the MGF definition what the MGF of a Bernoulli variable is. Then, $1 + x \le \exp(x)$ and Q4C are your friends. (3) JEE problem.