# Yale University
# S&DS 551, Spring 2023
# Homework 1

**Chang Feng (Felix) Zhou cz397**

## Problem 1.

### (a) Mean

Let $X_1, \dots, X_{n-k}$ be random variables denoting the eval of each of the $n-k$ remaining balls and $X = \sum_{i=1}^{n-k} X_i$ be their sum. By the linearity of expectation,

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^{n-k} \mathbb{E}[X_i] \\
&= \boxed{(n-k)\frac{n+1}{2}}.
\end{aligned}$$

### (b) Variance

We compute the second moment in similar fashion. Indeed,

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{i,j=1}^{n-k} \mathbb{E}[X_i X_j] \\
&= \sum_{\ell=1}^{n-k} \mathbb{E}[X_\ell^2] + 2 \sum_{1 \leq i < j \leq n-k} \mathbb{E}[X_i X_j].
\end{aligned}$$

For any $\ell \in [n-k]$,

$$\begin{aligned}
\mathbb{E}[X_\ell^2] &= \frac{1}{n} \sum_{i=1}^{n} i^2 \\
&= \frac{(n+1)(2n+1)}{6} \\
\sum_{\ell=1}^{n-k} \mathbb{E}[X_\ell^2] &= \frac{(n-k)(n+1)(2n+1)}{6}.
\end{aligned}$$

Now, there are $\binom{n}{2}$ possible pairs $1 \le i < j \le n$ and $\binom{n-k}{2}$ pairs of indices from $[n-k]$. It follows that

$$\begin{aligned}
2 \sum_{1 \le i < j \le n-k} \mathbb{E}[X_i X_j] &= \frac{\binom{n-k}{2}}{\binom{n}{2}} \sum_{1 \le i, j \le n : i \neq j} ij \\
&= \frac{(n-k)(n-k-1)}{n(n-1)} \left( \sum_{1 \le i, j \le n} ij - \sum_{\ell=1}^{n} \ell^2 \right) \\
&= \frac{(n-k)(n-k-1)}{n(n-1)} \left( \left[ \sum_{i=1} i \right]^2 - \sum_{j=1}^{n} j^2 \right) \\
&= \frac{(n-k)(n-k-1)}{n(n-1)} \left( \frac{n^2(n+1)^2}{4} - \frac{n(n+1)(2n+1)}{6} \right) \\
&= \frac{(n-k)(n-k-1)}{n(n-1)} \left( \frac{n(n+1)[3n(n+1) - 2(2n+1)]}{12} \right) \\
&= \frac{(n-k)(n-k-1)}{n(n-1)} \left( \frac{n(n+1)(3n^2 - n - 2)}{12} \right).
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{(n-k)(n+1)(2n+1)}{6} + \frac{(n-k)(n-k-1)}{n(n-1)} \left( \frac{n(n+1)(3n^2 - n - 2)}{12} \right) - (n-k)^2 \frac{(n+1)^2}{4} \\
&= \boxed{\frac{k(n+1)(n-k)}{12}}.
\end{aligned}$$

**Problem 2.**

**Problem 3.**

**Problem 4.**

**Problem 5.**