COMP 580 – Assignment 4
Section 2: Adaptive Sampling and Random Sampling

Saurav Kumar Gupta (SG275)

December 2, 2025

Setup

We have a set of non–negative numbers

$$S = \{w_1, w_2, \ldots, w_n\},$$

and we want to compute the sum

$$T = \sum_{i=1}^{n} w_i.$$

Assume that element w_i is sampled with probability p_i, where

$$p_i > 0, \qquad \sum_{i=1}^{n} p_i = 1.$$

Sampling with these probabilities induces a distribution D on S. We form a sampled set SS by sampling according to these probabilities. The assignment defines the estimator

$$\hat{T} \;=\; \sum_{w_j \in SS} \frac{w_j}{p_j}.$$

We are asked to:

- Prove that $\hat{T}$ is an unbiased estimator of T.
- Calculate $\mathrm{Var}(\hat{T})$ in terms of the w_i's and p_i's.
- Comment on how to design the p_i's to obtain smaller variance than random sampling with $p_i = 1/n$.

Let I_i be the indicator random variable of whether w_i is included in SS:

$$I_i = \begin{cases} 1, & \text{if } w_i \in SS, \\ 0, & \text{otherwise.} \end{cases}$$

Then $\mathbb{P}(I_i = 1) = p_i$ and

$$\hat{T} = \sum_{i=1}^{n} I_i \, \frac{w_i}{p_i}.$$

Using linearity of expectation,

$$\begin{aligned}\mathbb{E}[\hat{T}] &= \mathbb{E}\left[\sum_{i=1}^{n} I_i \frac{w_i}{p_i}\right] = \sum_{i=1}^{n} \mathbb{E}\left[I_i \frac{w_i}{p_i}\right] \\ &= \sum_{i=1}^{n} \frac{w_i}{p_i} \mathbb{E}[I_i] = \sum_{i=1}^{n} \frac{w_i}{p_i} p_i \\ &= \sum_{i=1}^{n} w_i = T.\end{aligned}$$

Thus

$$\boxed{\mathbb{E}[\hat{T}] = T,}$$

so $\hat{T}$ is an unbiased estimator of T.

Write

$$\hat{T} = \sum_{i=1}^{n} Y_i, \qquad \text{where} \qquad Y_i := I_i \frac{w_i}{p_i}.$$

We assume that the sampling of different elements is independent, so that the I_i (and hence the Y_i) are independent.

2.1 Variance of a single term

The random variable Y_i takes value $\frac{w_i}{p_i}$ with probability p_i and 0 with probability $1-p_i$. We already have $\mathbb{E}[Y_i] = w_i$. Also,

$$\mathbb{E}[Y_i^2] = \left(\frac{w_i}{p_i}\right)^2 \mathbb{P}(I_i = 1) + 0^2\,\mathbb{P}(I_i = 0) = \left(\frac{w_i}{p_i}\right)^2 p_i = \frac{w_i^2}{p_i}.$$

Therefore

$$\begin{aligned}\mathrm{Var}(Y_i) &= \mathbb{E}[Y_i^2] - \big(\mathbb{E}[Y_i]\big)^2 \\ &= \frac{w_i^2}{p_i} - w_i^2 \\ &= w_i^2\left(\frac{1}{p_i} - 1\right) = w_i^2\,\frac{1-p_i}{p_i}.\end{aligned}$$

2.2 Variance of the estimator

Since the Y_i are independent,

$$\begin{aligned}\mathrm{Var}(\hat{T}) &= \mathrm{Var}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \mathrm{Var}(Y_i) \\ &= \sum_{i=1}^{n} w_i^2\,\frac{1-p_i}{p_i}.\end{aligned}$$

Equivalently,

$$\mathrm{Var}(\hat{T}) = \sum_{i=1}^{n}\left(\frac{w_i^2}{p_i} - w_i^2\right) = \left(\sum_{i=1}^{n} \frac{w_i^2}{p_i}\right) - \sum_{i=1}^{n} w_i^2.$$

3. Choice of p_ip-i and comparison to uniform sampling

3.1 Uniform random sampling

Under uniform random sampling, $p_i = 1/n$ for all i. Plugging into the variance expression,

$$
\begin{aligned}
\mathrm{Var}_{\text{uniform}}(\hat{T}) &= \sum_{i=1}^{n} w_i^2 \, \frac{1-\frac{1}{n}}{\frac{1}{n}} \\
&= \sum_{i=1}^{n} w_i^2 \, (n-1) \\
&= (n-1) \sum_{i=1}^{n} w_i^2 .
\end{aligned}
$$

3.2 Designing p_i for smaller variance

From the general variance formula,

$$
\mathrm{Var}(\hat{T}) = \sum_{i=1}^{n} w_i^2 \, \frac{1-p_i}{p_i} = \left(\sum_{i=1}^{n} \frac{w_i^2}{p_i} \right) - \sum_{i=1}^{n} w_i^2 ,
$$

we see that for fixed w_i the term that depends on the sampling design is

$$
F(p_1, \ldots, p_n) = \sum_{i=1}^{n} \frac{w_i^2}{p_i} ,
$$

subject to

$$
p_i > 0, \qquad \sum_{i=1}^{n} p_i = 1.
$$

To minimize the variance of $\hat{T}$, we minimize F under these constraints. Using a Lagrange multiplier λ,

$$
\mathcal{L}(p_1, \ldots, p_n, \lambda) = \sum_{i=1}^{n} \frac{w_i^2}{p_i} + \lambda \left(\sum_{i=1}^{n} p_i - 1 \right) .
$$

For each i,

$$
\frac{\partial \mathcal{L}}{\partial p_i} = -\frac{w_i^2}{p_i^2} + \lambda = 0 \quad \Longrightarrow \quad p_i^2 = \frac{w_i^2}{\lambda} \quad \Longrightarrow \quad p_i \propto w_i .
$$

Thus, to obtain smaller variance than uniform sampling, we should choose sampling probabilities p_i so that larger w_i have larger p_i. In particular, the choice

$$
p_i = \frac{w_i}{\sum_{j=1}^{n} w_j}
$$

makes F (and hence $\mathrm{Var}(\hat{T})$) as small as possible under the given constraints, and therefore gives better (smaller) variance than uniform random sampling when the w_i are not all equal.