Problem 1

1) $\hat{E}_p = \frac{1}{m}\sum_{x_i \in S} \|x_i\|_2^2 \cdot 1/p_i$

$$E[\hat{E}_p] = \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n} \|x_{ij}\|_2^2 \cdot \frac{1}{p_i}\cdot p_j \quad \rightarrow \quad p_i = p_j$$

$$= \frac{1}{\cancel{m}}\cdot \cancel{m}\sum_{j=1}^{n}\|x_j\|_2^2 = Cost$$

2) $\mu(S) = \frac{1}{m}\sum_{x_i \in S} x_i$ $\qquad p_i = 1/n$

$$E[\|\mu(S)\|_2^2] = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{n}\|x_{ij}\|_2^2\, p_i = \frac{1}{m^{\cancel{2}}}\cdot\cancel{m}\overbrace{\sum_{j=1}^{n}\|x_j\|_2^2}^{Cost}\cdot\frac{1}{n} = \frac{1}{m}\cdot\frac{Cost}{n}$$

3) $\hat{E}_p = \frac{n}{m}\sum_{x_i \in S}\|x_i\|_2^2 = \underline{\frac{n}{m}\sum_{S_i \in S}\|S_i = X_i\|_2^2}$, $\underline{\frac{n}{m} >> 1}$

5) $\hat{E}_p = \frac{1}{m}\sum_{x_i \in S}\|x_i\|_2^2 \cdot 1/p_i$ $\qquad p_i = \frac{\|x_i\|_2^2}{Cost}$

$$\hat{E}_p = \frac{1}{m}\cdot\sum_{i=1}^{m}\|x_i\|_2^2\cdot\frac{Cost}{\|x_i\|_2^2} = \frac{1}{\cancel{m}}\cdot\cancel{m}\cdot\cancel{\|x_i\|_2^2}\cdot\frac{Cost}{\cancel{\|x_i\|_2^2}} = Cost$$

4) $$\mu(S) = \sum_{x_i \in S}\frac{1/p_i}{\sum_{x_i \in S} 1/p_i}x_i = \sum_{x_i \in S}\frac{\frac{Cost}{\|x_i\|_2^2}}{\sum_{x_i \in S}\frac{Cost}{\|x_i\|_2^2}}x_i = \sum_{x_i \in S}\frac{\frac{1}{\|x_i\|_2^2}}{\sum_{x_i \in S}\frac{1}{\|x_i\|_2^2}}\cdot x_i$$

Using the D2 Sampling, being $c(P) = 0$ if there is one point far off from the origin and many points close to it, the sample with high probability is therefore composed only by the point far from the origin.