Weighted Multiplicative Spanners. We saw a greedy algorithm for finding a multiplicative spanner of an unweighted graph in lecture. Recall a $k$-multiplicative spanner $H = (V, E')$ of a given unweighted graph $G = (V, E)$ on $n$ nodes, is a subgraph (so $E' \subseteq E$) for which for all pairs $u, v$ of vertices in $V$, $d_G(u, v) \leq d_H(u, v) \leq k \cdot d_G(u, v)$. In this problem we will find a multiplicative spanner $H = (V, E')$ in a weighted graph $G = (V, E)$ on $n$ nodes, where each edge $e \in E$ has a positive edge weight $w_e$. Consider the following algorithm:

1. Initialize $E'$ to $\emptyset$
2. Let $E = \{e_1 = \{u_1, v_1\}, e_2 = \{u_2, v_2\}, \ldots, e_m = \{u_m, v_m\}\}$ be such that
   \[ w_{e_1} \leq w_{e_2} \leq w_{e_3} \leq \cdots \leq w_{e_m}. \]
3. For $i = 1, 2, \ldots, m$,
   (a) If the distance between $u_i$ and $v_i$ in $H = (V, E')$ is more than $k \cdot w_{e_i}$, then add the edge $e_i$ to $E'$, otherwise discard the edge.
4. Output $H = (V, E')$.

1. Argue that $H$ is a $k$-multiplicative spanner.

2. Argue that for any choices of the weights $w_e$, the girth (minimum cycle length) of $H$ is at least $k + 2$. 
3. What is an upper bound on the number of edges in $H$?

**The Variance of CountSketch.** Recall in lecture we introduced the CountSketch, which is a random linear map $S$ from $\mathbb{R}^n$ to $\mathbb{R}^k$, for $k = \Theta(1/\epsilon^2)$, defined as follows. Let $h : \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, k\}$ be a 2-wise independent hash function, and $\sigma : \{1, 2, \ldots, n\} \rightarrow \{-1, 1\}$ be a 4-wise independent hash function. Then for $i = 1, 2, \ldots, k$, we have $(Sx)_i = \sum_{j \text{ s.t. } h(j) = i} \sigma(j)x_j$, where $x$ is the $n$-dimensional input vector.

In lecture, we showed $E[\|Sx\|^2] = \|x\|^2$, and claimed that $\text{Var}[\|Sx\|^2] = O(\|x\|^4/k)$. We saw that these statements, by Chebyshev’s inequality, imply $\Pr[|\|Sx\|^2 - \|x\|^2| > \epsilon\|x\|^2] \leq \frac{1}{10}$. Prove that $\text{Var}[\|Sx\|^2] \leq 2\|x\|^4/k$. 

(1)
Case 1: if \( j_1 \neq j_2 \), then
\[ E[\delta(h(i_1) = j_1) \delta(h(i_2) = j_1) \delta(h(i_3) = j_2) \delta(h(i_4) = j_2)] = 0 \]
since the same index \( i \) cannot hash to more than one bucket. If \( j_1 = j_2 \), then
\[ E[\delta(h(i_1) = j_1) \delta(h(i_2) = j_1) \delta(h(i_3) = j_2) \delta(h(i_4) = j_2)] = 1/k \]
so (1) simplifies to
\[ \sum_{k=1}^{n} \sum_{i=1}^{n} (1/k) x_4 i = \sum_{i=1}^{n} x_4 i. \]

Case 2: \( E[\delta(h(i_1) = j_1) \delta(h(i_2) = j_1) \delta(h(i_3) = j_2) \delta(h(i_4) = j_2)] = 1/k^2 \), and so (1) simplifies to
\[ \sum_{k=1}^{n} \sum_{i_1,j_2} (1/k^2) x_2 i_1 x_2 i_2 \leq \|x\|_4 - \sum_{i=1}^{n} x_4 i. \]

Case 3: if \( j_1 \neq j_2 \), then
\[ E[\delta(h(i_1) = j_1) \delta(h(i_2) = j_1) \delta(h(i_3) = j_2) \delta(h(i_4) = j_2)] = 0 \]
since the same index \( i \) cannot hash to more than one bucket. If \( j_1 = j_2 \), then (1) simplifies to
\[ \sum_{k=1}^{n} \sum_{i_1 \neq i_2} (1/k^2) x_2 i_1 x_2 i_2 \leq 1/k \|x\|_4. \]

Case 4: is analogous to case 3. For completeness: if \( j_1 \neq j_2 \), then
\[ E[\delta(h(i_1) = j_1) \delta(h(i_2) = j_1) \delta(h(i_3) = j_2) \delta(h(i_4) = j_2)] = 0 \]
since the same index \( i \) cannot hash to more than one bucket. If \( j_1 = j_2 \), then (1) simplifies to
\[ \sum_{k=1}^{n} \sum_{i_1 \neq i_2} (1/k^2) x_2 i_1 x_2 i_2 \leq 1/k \|x\|_4. \]

Summing over the four cases, (1) is upper bounded as
\[ \|x\|_4 + 2/k \|x\|_4 \]. Hence,
\[ \text{Var}[\|Sx\|_2] = E[(\|Sx\|_2)^2] - (E[\|Sx\|_2])^2 \leq 2/k \|x\|_4. \]