homework: Jaccard similarity in Min-hash

Denote the Jaccard similarity of columns $J(C_1, C_2) = \frac{|C_1 \cap C_2|}{|C_1 \cup C_2|}$ and the similarity of their signatures in Min-hashing $\hat{J}(C_1, C_2) = \frac{1}{k} \sum_{i=1}^{k} \mathbb{1}[h_{\pi_i}(C_1) = h_{\pi_i}(C_2)]$, k is the number of hash functions.

Please prove:

> $$Pr(\left|J - \hat{J}\right| \leq \epsilon) > 1 - \delta$$

Where ϵ, δ satisfies $k = O(\frac{1}{\epsilon^2 \delta})$.

Hint: You can use Chebyshev's inequality or Chernoff bound.