

We imagine dividing the set  $S$  into 20 *quantiles*  $Q_1, \dots, Q_{20}$ , where  $Q_i$  consists of all elements that have at least  $.05(i-1)n$  elements less than them, and at least  $.05(20-i)n$  elements greater than them. Choosing the sample  $S'$  is like throwing a set of numbers at random into bins labeled with  $Q_1, \dots, Q_{20}$ .

Suppose we choose  $|S'| = 40,000$  and sample with replacement. Consider the event  $\mathcal{E}$  that  $|S' \cap Q_i|$  is between 1800 and 2200 for each  $i$ . If  $\mathcal{E}$  occurs, then the first nine quantiles contain at most 19,800 elements of  $S'$ , and the last nine quantiles do as well. Hence the median of  $S'$  will belong to  $Q_{10} \cup Q_{11}$ , and thus will be a (.05)-approximate median of  $S$ .

The probability that a given  $Q_i$  contains more than 2200 elements can be computed using the Chernoff bound (4.1), with  $\mu = 2000$  and  $\delta = .1$ ; it is less than

$$\left[ \frac{e^{.05}}{(1.05)^{(1.05)}} \right]^{10000} < .0001.$$

The probability that a given  $Q_i$  contains fewer than 1800 elements can be computed using the Chernoff bound (4.2), with  $\mu = 2000$  and  $\delta = .1$ ; it is less than

$$e^{-(.5)(.1)(.1)2000} < .0001.$$

Applying the Union Bound over the 20 choices of  $i$ , the probability that  $\mathcal{E}$  does not occur is at most  $(40)(.0001) = .004 < .01$ .

---

<sup>1</sup>ex835.763.619