0.1 Introduction: What is Concentration

1. Introduction : What is "concentration" here?

Suppose that X r.v. has expected value $E(X)=\mu$, and variance $E[(X-\mu)^2]=\sigma^2$. Then chebychev's inequality :

$$Pr(|X-\mu|\geq t) \leq \sigma^2/t^2 \quad \text{for any } t>0.$$

Thus for $t \gg \sigma$, $Pr(|X-\mu|\geq t)$ is small. However, we want prob. of large deviations to be very small - " X is strongly concentrated around μ. One concentration is shown :

> [Theorem 1.1] Let $X_1, \cdots, X_n$ be independent binary random variables, with $Pr(X_k=1)=p$ and $Pr(X_k=0)=1-p$ for all k. Let $S_n=\sum X_k$, then for any $t\geq 0$,
>
> $$Pr(|S_n-np|\geq nt) \leq 2e^{-2nt^2}$$

Since $\sigma^2=D(S_n)=np(1-p)=\frac{n}{4}$ when $p=\frac{1}{2}$, due to Chebychev's inequality : $Pr(|S_n-np|\geq nt) \leq 1/(4nt^2)$, not small enough.

As an example - Quicksort : one of the sorting algorithm. The variance of the time taken is not to large, and so large deviations from the average are very not likely.

There are two main approaches for proving concentration results - the bounded difference (martingale method) and the method of Talagrand. We stick to bounded discrete "time"