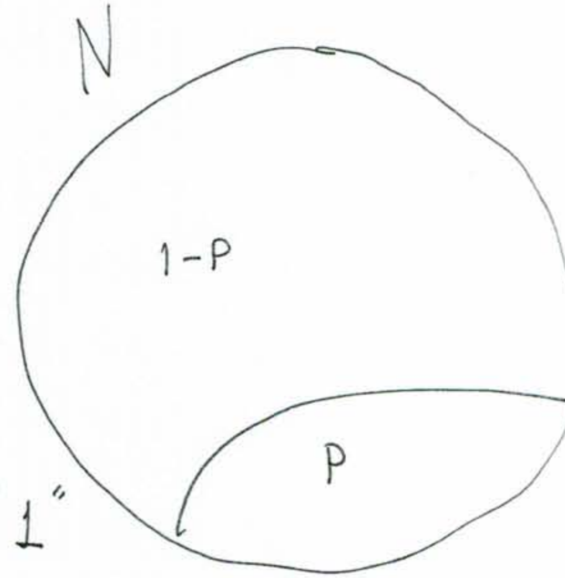
1 9-8-10

Chernoff Bounds

Motivation



N is a very large population

(eg U.S. citizens
web pages over WWW
traffic over ISP etc)

p fraction of the population votes/is type "1"

1-p " "0"

We want a "very good estimate" $\tilde{p}$ of p.

In particular, for given α and ε (very small)

we want $\boxed{\Pr\left[\,|p-\tilde{p}| > \alpha\,\right] < \varepsilon}$

Algorithm

> Take n independent samples $X_1, \dots, X_n$
>
> Set $X = \sum_{i=1}^{n} X_i$
>
> Set $\tilde{p} = \dfrac{X}{n}$

QUESTION How large does n have to be so that our estimate is "very good"?

(Recall, in election polls, $n \simeq 2{,}000$ samples predict outcome very accurately!)

9-8-10

Let X_i, $1 \le i \le n$ be independent Poisson trials

such that $\Pr[X_i = 1] = p_i$

$\Pr[X_i = 0] = 1 - p_i$

Let $X = \sum_{i=1}^{n} X_i$

Let $\mu = E[X]$

<u>Theorem 1</u> For any $\delta > 0$ $\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$

<u>Theorem 2</u> For any $1 > \delta > 0$ $\Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$

<u>Theorem 3</u> For any $0 < \delta < 2e - 1$ $\Pr[X > (1+\delta)\mu] < e^{-\frac{\delta^2}{4}\mu}$

<u>Theorem 4</u> For any $1 > \delta > 0$ $\Pr[X < (1-\delta)\mu] < e^{-\frac{\delta^2}{2}\mu}$

<u>Theorem 5</u> For any $0 < \delta < 1$ $\Pr[|X - \mu| > \delta\mu] < 2e^{-\frac{\delta^2\mu}{4}}$