Mathematics for Computer Science

MIT 6.042J/18.062J

Deviation of Repeated Trials

 December 12, 2005 lec15M.1

Jacob D. Bernoulli (1659 – 1705)

 December 12, 2005 lec15M.2

Jacob D. Bernoulli (1659 – 1705)

Even the stupidest man---by some instinct of nature *per se* and by no previous instruction (this is truly amazing) -- knows for sure that the more observations ...that are taken, the less the danger will be of straying from the mark.

---*Ars Conjectandi* (The Art of Guessing), 1713*

Ars Conjectandi by Jacob Bernoulli,
quoted in Introduction to Probability by Charles Grinstead and J. Laurie Snell,
published by the American Mathematical Society, Providence RI, in 1997. The book is freely available here:
http://www.dartmouth.edu/%7Echance/teaching_aids/books_articles/probability_book/amsbook.mac.pdf

 December 12, 2005 lec15M.3

Jacob D. Bernoulli (1659 – 1705)

It certainly remains to be inquired whether after the number of observations has been increased, the probability…of obtaining the true ratio…finally exceeds any given degree of certainty; or whether the problem has, so to speak, its own asymptote---that is, whether some degree of certainty is given which one can never exceed.

Ars Conjectandi by Jacob Bernoulli, quoted in Introduction to Probability by Charles Grinstead and J. Laurie Snell, published by the American Mathematical Society, Providence RI, in 1997. The book is freely available here: http://www.dartmouth.edu/%7Echance/teaching_aids/books_articles/probability_book/amsbook.mac.pdf

 December 12, 2005 lec15M.4

Deviation from the Mean

Pr{observed value *far from* expected value}

is *SMALL*

How small?

 December 12, 2005 lec15M.5

Deviation from the Mean

Observed value means random variable, R.

far from may mean:

- distance or
- amount above (or below)

 December 12, 2005 lec15M.6

Markov Bound

$$\Pr\{R \text{ above } \underbrace{x}_{\text{far}}\} \le \frac{\mu}{\underbrace{x}_{\text{small}}}$$

Chebychev Bound

$$\Pr\{\underbrace{|R-\mu|}_{\text{distance}} > \underbrace{x}_{\text{far}}\} \le \frac{\sigma^2}{\underbrace{x^2}_{\text{small}}}$$

Binomial Bound

$$\Pr\{|B_{n,p}-\mu| > x\} \le e^{\frac{-x^2\mu}{2}}$$

Weak Law of Large Numbers

A_n ::= Avg. of n independent trials

μ::= E[single trial]

$$\lim_{n\to\infty}\left[\Pr\{\underbrace{|A_n-\mu|}_{\text{distance}} > \underbrace{\varepsilon}_{\text{far}}\}\right] \overset{?}{=} \underbrace{0}_{\text{small}}$$

Jacob D. Bernoulli (1659 – 1705)

Therefore, this is the problem which I now set forth and make known after I have pondered over it for twenty years. Both its novelty and its very great usefulness, coupled with its just as great difficulty, can exceed in weight and value all the remaining chapters of this thesis.

The Principle Behind:

- Estimation (polling)
- Algorithm analysis
- Design against failure
- Communication thru noise
- Gambling

Not Usable as Stated

Need to know the

rate of convergence to 0

for any application.

Repeated Trials

$X_1, \cdots, X_n$ independent

with mean, μ, and variance σ^2

$$A_n ::= (X_1 + \cdots + X_n)/n$$

$$E[A_n] = n\mu/n = \mu$$

Repeated Trials

$$\mathrm{Var}[X_1 + \cdots + X_n] = n\sigma^2$$

(by independence)

$$\mathrm{Var}[A_n] = n\sigma^2/n^2 = \frac{\sigma^2}{n}$$

decreases with # trials

Repeated Trials

So by Chebychev

$$\Pr\{|A_n - \mu| > \varepsilon\} \le (\sigma/\varepsilon)^2 \cdot \underbrace{\frac{1}{n}}_{\to 0}$$

as $n \to \infty$

Weak Law of Large Numbers

Therefore

$$\lim_{n\to\infty} \left[\Pr\{|A_n - \mu| > \varepsilon\}\right] = 0$$

QED

