Team Problems

Problems 1&2

Mathematics for Computer Science

MIT 6.042J/18.062J

Deviation of Repeated Trials

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*

*taken from Grinstead \& Snell, http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html *Introduction to Probability*, American Mathematical Society, p. 310.

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.

Repeated Trials

Random variable R with mean μ

n independent observations of R

$$R_1, \cdots, R_n$$

Repeated Trials

take average:

$$A_n ::= \frac{R_1 + \cdots + R_n}{n}$$

close to `true ratio' with prob. ?

$$\Pr\{|A_n - \mu| \leq \underbrace{x}_{\text{as close as } x>0}\} \text{ ?}$$

Repeated Trials

$$\Pr\{|A_n - \mu| \leq x\}$$

Even `stupidest man' knows this prob. gets bigger as n gets bigger

—but *how big?*

Does it "exceed… any given degree of certainty"?

That is, does it approach 1?

Weak Law of Large Numbers

$$\lim_{n\to\infty} \Pr\{|A_n - \mu| \leq x\} = 1$$

YES

$$\boxed{\lim_{n\to\infty} \Pr\{|A_n - \mu| > x\} = 0}$$

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.

Jacob D. Bernoulli (1659 – 1705)

Weak Law of Large Numbers

$$\boxed{\lim_{n\to\infty} \Pr\{|A_n - \mu| > x\} = 0}$$

Will be an easy Corollary of Chebyshev and properties of variance.

Repeated Trials

$$\boxed{\mathrm{E}[A_n]} = \mathrm{E}\left[\frac{R_1 + \cdots + R_n}{n}\right] = \frac{\mathrm{E}[R_1] + \cdots + \mathrm{E}[R_n]}{n} = \frac{n\mu}{n} \boxed{= \mu}$$

Repeated Trials

by Chebyshev:

$$\Pr\{|A_n - \mu| > x\} \le \frac{\mathrm{Var}[A_n]}{x^2}$$

so need only show $\mathrm{Var}[A_n] \to 0$

Repeated Trials

what is $\mathrm{Var}[A_n]$?

let σ^2 ::= $\mathrm{Var}[R]$

Repeated Trials

$$\boxed{\mathrm{Var}[A_n]}$$
$$= \mathrm{Var}[(R_1 + \cdots + R_n)/n]$$
$$= (\mathrm{Var}[R_1] + \cdots + \mathrm{Var}[R_n])/n^2$$
$$= \frac{n\sigma^2}{n^2} = \underbrace{\boxed{\frac{\sigma^2}{n}}}$$

$\to 0$ as $n \to \infty$

QED

Analysis of the Proof

proof only used

- $R_1,\ldots,R_n$ have same finite mean, μ
- and finite variance, σ^2
- and variances add:

$$\mathrm{Var}[R_1 + \cdots + R_n]$$
$$= \mathrm{Var}[R_1] + \cdots + \mathrm{Var}[R_n]$$

which follows from *pairwise* independence

Pairwise Independent Sampling

Let $R_1,\ldots,R_n$ be pairwise independent random vars with the same finite mean, μ, and variance, σ^2. Let $A_n ::= (R_1 + \cdots + R_n)/n$. Then

$$\boxed{\Pr\{|A_n - \mu| > x\} \leqslant \frac{1}{n}\left(\frac{\sigma}{x}\right)^2}$$

Pairwise Independent Sampling

The punchline:

we now know how big a sample is needed to estimate the mean of any* random variable to within any* desired tolerance and to any* degree of confidence.

* Var[rand. var] < ∞, tolerance > 0, confidence < 100%

Team Problems

Problems 3&4

 May 16, 2007 lec 14W.19