Mathematics for Computer Science

MIT 6.042J/18.062J

The Law of Large Numbers

What the mean means

The mean value of a fair die roll is 3.5, but we will never roll 3.5. So why do we care what the mean is?

We believe that after many rolls, the average roll will be near 3.5.

What probability means

$$\Pr[\text{roll } 6] = \frac{1}{6}$$

We believe that after many rolls, the fraction of 6's will be near 1/6.

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.

Dice Rolls

n rolls of fair die

$$\Pr[\text{roll } 6] = \frac{1}{6}$$

$$\text{Avg } \#6\text{'s} ::= \frac{\#6\text{'s rolled}}{n}$$

Dice Rolls

n rolls of fair die

Bernoulli: we believe intuitively that

$$\frac{\#6\text{'s rolled}}{n} \to \frac{1}{6} \text{ as } n \to \infty$$

Dice Rolls

n rolls of fair die

of course, an unlucky average might be way off, but that's unlikely. how unlikely?

$$\frac{\#6\text{'s rolled}}{n} \to \frac{1}{6} \text{ as } n \to \infty$$

Pr[Average = 1/6 ± %]

n	±10%
6	0.4
60	0.26
600	0.72
1200	0.88
3000	0.98
6000	0.999

Pr bigger with # rolls

Pr[Average = 1/6 ± %]

n	±10%	±5%
6	0.4	0.4
60	0.26	0.14
600	0.72	0.41
1200	0.88	0.56
3000	0.98	0.78
6000	0.999	0.91

Pr bigger with # rolls

Pr smaller for better %

Pr[Average = 1/6 ± %]

n	±10%	±5%
6	0.4	0.4
60	0.26	0.14
600	0.72	0.41
1200	0.88	0.56
3000	0.98	0.78
6000	0.999	0.91

Pr bigger with # rolls

Pr smaller for better %

Pr[Average = 1/6 ± %]

n	±10%	±5%
3000	0.98	0.78

If you rolled 3000 times and did not get 450—550 6's

Pr[Average = 1/6 ± %]

n	±10%	±5%
3000	0.98	0.78

If you rolled 3000 times and did not get 450—550 6's you can be 98% confident your die is loaded

Pr[Average $= 1/6 \pm$ %]

If you rolled 3000 times and did not get 475—525 6's, you can be 78% confident your die is loaded

n	±10%	±5%
3000	0.98	0.78

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.

What Bernoulli means

Random var R with mean μ.

What Bernoulli means

Random var R with mean μ. Make n "trial observations" of R and take the average

Repeated Trials

Mutually independent, identically distributed (i.i.d) random variables

$$R_1, \cdots, R_n$$

with mean μ

Repeated Trials

take average:

$$A_n ::= \frac{R_1 + R_2 + \cdots + R_n}{n}$$

Bernoulli question: is average probably close to μ if n is big?

Repeated Trials

take average:

$$A_n ::= \frac{R_1 + R_2 + \cdots + R_n}{n}$$

probably close to μ

$$\Pr[\,|A_n - \mu| \le \underbrace{\delta}\,] = ?$$

as close as $\delta > 0$

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.

Bernoulli question:

$$\lim_{n\to\infty} \Pr[\,|A_n - \mu| \le \delta\,] = ?$$

Weak Law of Large Numbers

$$\boxed{\lim_{n\to\infty} \Pr[\,|A_n - \mu| \le \delta\,] = 1}$$

Weak Law of Large Numbers

$$\boxed{\lim_{n\to\infty} \Pr[\,|A_n - \mu| > \delta\,] = 0}$$

will follow easily by Chebyshev & variance properties