Law of Large Numbers

May 9, 2006

- An intuitive way to view the probability of a certain outcome is the frequency with which that outcome occurs in the long run.
- We defined probability mathematically as a value of a distribution function for the random variable representing the experiment.
- The Law of Large Numbers shows that this model is consistent with the frequency interpretation of probability.

Chebyshev Inequality

Theorem. *Let X be a discrete random variable with expected value $\mu = E(X)$, and let $\epsilon > 0$ be any positive real number. Then*

$$P(|X - \mu| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2} \ .$$

Example

- Let X by any random variable with $E(X) = \mu$ and $V(X) = \sigma^2$.

- Then, if $\epsilon = k\sigma$, Chebyshev's Inequality states that

$$P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2} .$$

- Thus, for any random variable, the probability of a deviation from the mean of more than k standard deviations is $\leq 1/k^2$.

- Chebyshev's Inequality is the best possible inequality in the sense that, for any $\epsilon > 0$, it is possible to give an example of a random variable for which Chebyshev's Inequality is in fact an equality.

- Chebyshev's Inequality is the best possible inequality in the sense that, for any $\epsilon > 0$, it is possible to give an example of a random variable for which Chebyshev's Inequality is in fact an equality.

- Given $\epsilon > 0$, choose X with distribution

$$p_X = \begin{pmatrix} -\varepsilon & -\varepsilon \\ 1/2 & 1/2 \end{pmatrix} .$$

Then $E(X) = 0$, $V(X) = \epsilon^2$, and

$$P(|X - \mu| \geq \epsilon) = \frac{V(X)}{\epsilon^2} = 1 .$$

Law of Large Numbers

Theorem. *Let* X_1, X_2, *. . .* , X_n *be an independent trials process, with finite expected value* $\mu = E(X_j)$ *and finite variance* $\sigma^2 = V(X_j)$. *Let* $S_n = X_1 + X_2 + \cdots + X_n$. *Then for any* $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \to 0$$

as $n \to \infty$. *Equivalently,*

$$P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) \to 1$$

as $n \to \infty$.

Proof

- Since X_1, X_2, ..., X_n are independent and have the same distributions,

$$V(S_n) = n\sigma^2 \ ,$$

$$V(\frac{S_n}{n}) = \frac{\sigma^2}{n} \ .$$

$$E(\frac{S_n}{n}) = \mu \ .$$

Proof

- Since X_1, X_2, ..., X_n are independent and have the same distributions,

$$V(S_n) = n\sigma^2 \ ,$$

$$V(\frac{S_n}{n}) = \frac{\sigma^2}{n} \ .$$

$$E(\frac{S_n}{n}) = \mu \ .$$

- By Chebyshev's Inequality, for any $\epsilon > 0$,

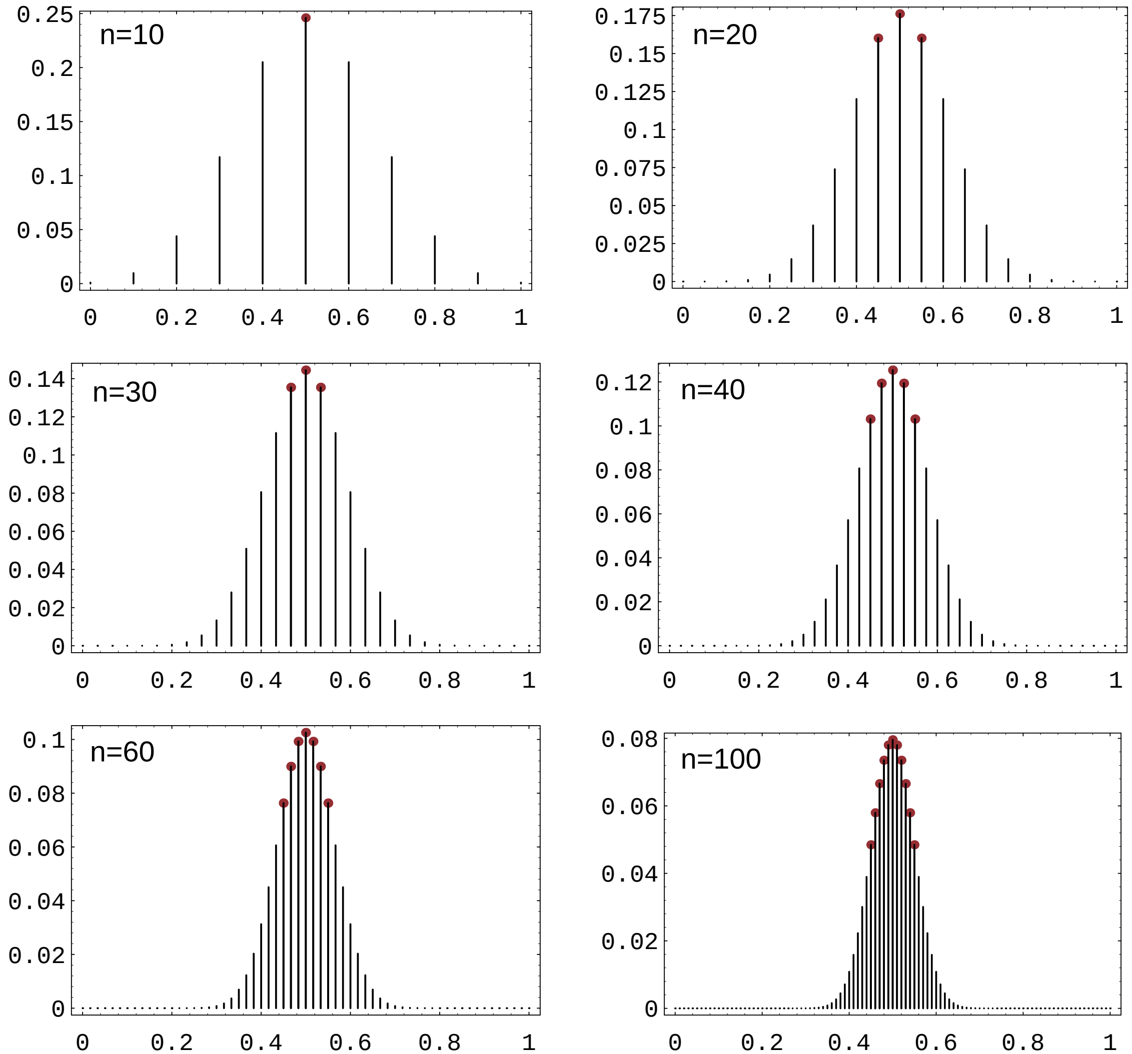
$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} \ .$$

Law of Averages

- Consider the important special case of Bernoulli trials with probability p for success.
- Let $X_j = 1$ if the jth outcome is a success and 0 if it is a failure.
- Then $S_n = X_1 + X_2 + \cdots + X_n$ is the number of successes in n trials and $\mu = E(X_1) = p$.
- The Law of Large Numbers states that for any $\epsilon > 0$

$$P\left(\left|\frac{S_n}{n} - p\right| < \epsilon\right) \to 1$$

as $n \to \infty$.


n=10
0 0.05 0.1 0.15 0.2 0.25
0 0.2 0.4 0.6 0.8 1
n=20
0 0.025 0.05 0.075 0.1 0.125 0.15 0.175
0 0.2 0.4 0.6 0.8 1
n=30
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14
0 0.2 0.4 0.6 0.8 1
n=40
0 0.02 0.04 0.06 0.08 0.1 0.12
0 0.2 0.4 0.6 0.8 1
n=60
0 0.02 0.04 0.06 0.08 0.1
0 0.2 0.4 0.6 0.8 1
n=100
0 0.02 0.04 0.06 0.08
0 0.2 0.4 0.6 0.8 1

Problem

- Show that the estimate

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \frac{1}{4n\epsilon^2}.$$

Problem

- We have two coins: one is a fair coin and the other is a coin that produces heads with probability 3/4.
- One of the two coins is picked at random, and this coin is tossed n times.
- Let S_n be the number of heads that turns up in these n tosses.
- Does the Law of Large Numbers allow us to predict the proportion of heads that will turn up in the long run?
- After we have observed a large number of tosses, can we tell which coin was chosen?

- How many tosses suffice to make us 95 percent sure?

The Continuous Case

- **(Chebyshev Inequality)** Let X be a continuous random variable with density function $f(x)$. Suppose X has a finite expected value $\mu = E(X)$ and finite variance $\sigma^2 = V(X)$. Then for any positive number $\epsilon > 0$ we have

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}\ .$$

Law of Large Numbers

Theorem. *Let X_1, X_2, . . . , X_n be an independent trials process with a continuous density function f, finite expected value μ, and finite variance σ^2. Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of the X_i. Then for any real number $\epsilon > 0$ we have*

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0 \ ,$$

or equivalently,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) = 1 \ .$$