Lecture 15 **Math 20 Fall 2014, Dartmouth College**

1 Independent trials

Let S_n be the sum of n Bernoulli trials $X_1, \dots, X_n$, each with $p = 1/6$.

1. Find $\mu = E(X_i)$ and $\sigma^2 = V(X_i)$.

$$\mu = p, \ \sigma^2 = pq$$

2. For any n, find $E(S_n)$ and $V(S_n)$.

$$E(S_n) = nE(X_i) = \mu, \ V(S_n) = nV(S_n) = n\sigma^2.$$

3. Find $E(\frac{S_n}{n})$ and $V(\frac{S_n}{n})$. What does $\frac{S_n}{n}$ describe?

$$E(\frac{S_n}{n}) = \frac{1}{n}E(S_n) = \mu, \ V(\frac{S_n}{n}) = \frac{1}{n^2}V(S_n) = \frac{\mu}{n}.$$

$\frac{S_n}{n}$ is the proportion of trials, out of n, that turned out to be successful.

Theorem 1 (Chebyshev inequality)
Let X be a random variable with expected value $\mu = E(X)$ and let $\epsilon > 0$ be any positive real number. Then:

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

The probability that X differs from μ by ϵ or more is bounded by $V(X)/\epsilon^2$

Using Chebyshev inequality, find the bounds on the following:

$P(|S_{360} - 60| \geq 10) \leq \frac{1}{2}$

$P(|S_{360} - 60| \geq 20) \leq \frac{1}{8}$

$P(S_{3600} - 600| \geq 100) \leq \frac{1}{20}$

$P(|S_{3600} - 600| \geq 200) \leq \frac{1}{80}$

$P(|\frac{S_{360}}{360} - \frac{1}{6}| \geq \frac{1}{36}) \leq \frac{1}{2}$

$P(|\frac{S_{360}}{360} - \frac{1}{6}| \geq \frac{2}{36}) \leq \frac{1}{8}$

$P(\frac{S_{3600}}{3600} - \frac{1}{6}| \geq \frac{1}{36}) \leq \frac{1}{20}$

$P(|\frac{S_{3600}}{3600} - \frac{1}{6}| \geq \frac{2}{36}) \leq \frac{1}{80}$

In other words, as the number of trials increases, the probability that the average differs from the mean by any specific amount gets smaller. This property is called the *Law of Large Numbers.*

2 Law of Large Numbers

Let $X_1, X_2, \ldots, X_n$ be an independent trials process where each trial has finite expected value $\mu = E(X_i)$ and finite variance $\sigma^2 = V(X_i)$. Let $S_n = X_1 + X_2 + \cdots + X_n$. Then for any $\epsilon > 0$, Chebyshev inequality on S_n/n implies:

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

Equivalently,

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

Notice that as n grows, $E(S_n/n)$ stays the same, it's always μ. But $V(S_n/n)$ is:

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

which gets smaller and smaller. So for the case of Bernoulli trials with $p = 1/6$, we get:

$$E(\frac{S_n}{n}) = \mu = \frac{1}{6}$$

$$V(\frac{S_n}{n}) = \frac{npq}{n^2}\frac{5}{36n}.$$

LAW OF LARGE NUMBERS: AS WE PERFORM MORE AND MORE TRIALS, THEN WITH PROBABILITY APPROACHING 1 THE PROPORTION OF SUCCESSFUL TRIALS APPROACHES p.

3 Uniform density

This works exactly the same way in the continuous case. For instance, let $X_1, \ldots, X_i$ each be chosen uniformly at random from [0,1]. We have: $\mu = E(X_i) = 1/2$, $\sigma^2 = V(X_i) = 1/12$. Then if $S_n = X_1 + \cdots + X_n$. Find:

1. $E(\frac{S_n}{n}) = \frac{1}{2}$

2. $V(\frac{S_n}{n}) = \frac{1}{12n}$

Then, fill in the appropriate form of Chebyshev inequality for S_n being the sum of n uniformly distributed random variables on [0,1]:

For any $\epsilon > 0$,

$$P(|\frac{S_n}{n} - \frac{1}{2}| \geq \epsilon) \leq \frac{1}{12n\epsilon^2},$$

Which clearly satisfies:

$$\lim_{n \to \infty} P(|\frac{S_n}{n} - \frac{1}{2}| \geq \epsilon) = 0.$$

4 Normal density

Suppose we choose n numbers at random, independently, using a standard normal distribution. Then for each number, $\mu = 0,\ \sigma^2 = 1$. Let S_n be the sum of those numbers.

$$E(S_n) = 0$$

$$V(S_n) = n$$

$$E(\frac{S_n}{n}) = 0$$

$$V(\frac{S_n}{n}) = \frac{1}{n}$$

And therefore by Chebyshev inequality,

$$P(|\frac{S_n}{n}| \geq \epsilon)\frac{1}{n\epsilon^2}.$$

5 Case when LLN fails

If variance is infinite, the Law of Large Numbers fails. Consider the density $f(x) = \frac{1}{x^2}$ for $x \in (1, \infty)$. Then neither $E(X)$ nor $V(X)$ exist, and we cannot find the bounds necessary for the LLN.