MAT 105 - Summary of Lecture 18
Thursday, March 10, 2016

> *The conception of chance enters in the very first steps of scientific activity in virtue of the fact that no observation is absolutely correct. I think chance is a more fundamental conception that causality; for whether in a concrete case, a cause-effect relation holds or not can only be judged by applying the laws of chance to the observation.* (Max Born)

Limit Laws

Recall the following notation: if $X_1, X_2, \ldots$ are independent random variables with the same distribution (think same experiment, repeated over and over, with X_1 being the first outcome, X_2 the second etc.) with mean $\mu = E[X_1] = E[X_2] = \cdots$ and variance $\sigma^2 = Var(X_1) = Var(X_2) = \cdots$, we say they are **i.i.d.**, short for **independent identically distributed** random variables. Let

$$S_n = X_1 + X_2 + \cdots + X_n.$$

Law of Large Numbers (LLN)

The Law of Large Numbers intuitively says that the average of multiple occurrences of a random variable approaches the mean (expected value) of the random variable. More precisely,

Theorem (Law of Large Numbers): If $X_1, X_2, \ldots$ are independent trials of an experiment, each trial having the same distribution with expectation μ (finite) and variance σ^2, then $\frac{S_n}{n}$ converges to μ that is,

$$\frac{S_n}{n} \approx \mu \text{ for } n \text{ very large.}$$

Examples:

1. Suppose a fair coin is tossed many times. Suppose it comes up Heads in all of the first 100 flips.
   (a) What percentage of Heads will there be in the next 900 flips? Does LLN say that there will be about 400, so that we average in 1000 flips is about 500? or does LLN say there will be about 450, which is half of the remaining 900 tosses? The answer is 450, since we can only average over the randomness of the next 900 flips. That means we will arrive at about $100 + 450 = 550$ Heads in 1000 tosses. We get an average of
   $$\frac{100 + 450}{1000} = 55\% \text{ Heads.}$$

   We encountered this question earlier in the semester and saying that "I am due for a win" because I am averaging too many losses is gambler's fallacy.

   (b) What happens if we play more? Still 100 rounds have passed, but we play 9,900 more rounds. Then LLN suggests we should expect about 9900/2 more Heads which brings us to $5,050$ Heads in 10,000 rounds and an average of
   $$\frac{100 + 4,950}{10,000} = 50.5\% \text{ Heads.}$$

   (c) Playing even longer, say 99,900 rounds, we expect $49,950$ Heads which leads to an average of
   $$\frac{100 + 49,950}{100,000} = 50.05\% \text{ Heads.}$$

We can see that the more we play, the closer we get to the mean for this random variable. This justifies the expression "The Law of Large Numbers swallows but does not compensate." That is, playing more will slowly erase the negative effect of the first 100 unlucky outcomes, but it will never completely erase the bad outcomes, in the sense that the average is expected to stay above 50%.

2. Roll a die repeatedly. Count the number of times a 6 comes up. Let $X_1 = 1$ if the 1st roll lands on 6 and zero otherwise. Then $X_1 = Bernoulli(1/6)$. Similarly, we let $X_2, X_3, \cdots$ be $Bernollui(1/6)$ associated with rolls $2, 3, \cdots$. Then
$$E[X_1] = E[X_2] = \cdots = \frac{1}{6} \text{ and } Var(X_1) = Var(X_2) = \cdots = \frac{5}{36}.$$
   (a) If we roll the die 10 times, $n = 10$ and we expect $\frac{S_{10}}{10} = \frac{X_1 + \cdots + X_{10}}{10} \approx \frac{1}{6}$

   (b) If we roll the die 100 times, $n = 100$ and we expect $\frac{S_{100}}{100} = \frac{X_1 + \cdots + X_{100}}{100} \approx \frac{1}{6}$, where the approximation should be better than in (a).

   (c) If we roll the die 1000 times, $n = 1000$ and we expect $\frac{S_{1000}}{1000} = \frac{X_1 + \cdots + X_{1000}}{1000} \approx \frac{1}{6}$, with a better approximation than in (a) or (b).

Central Limit Theorem (CLT)

Another result dealing with large sums of random variables is the Central Limit Theorem.

Theorem (Central Limit Theorem): If $X_1, X_2, \ldots$ are independent trials of an experiment, each trial having the same distribution with expectation μ (finite) and variance σ^2 with $0 < \sigma^2 < \infty$, and we let $S_n = X_1 + X_2 + \cdots + X_n$, then
$$P\left(a \leq \frac{S_n - n\mu}{\sqrt{n}\,\sigma} \leq b\right) \approx \phi(b) - \phi(a) \text{ for } n \text{ very large.}$$
This says that the sum of a large number of random variables (independent and with the same distribution) can be approximated by the normal distribution. In fact, this says that for n very large, $S_n \approx N(n\mu, n\sigma^2)$. One needs only to **normalize** S_n by subtracting its mean and dividing by its standard deviation in order to use the z-table and approximate using a standard normal.

Examples:

1. Flip a fair coin 400 times. Let S_{400} =the number of heads. Find the probability that we rolled more than 213 heads.

   Let $X_1 = 1$ if the 1st coin lands on Heads, and 0 otherwise.

   Let $X_2 = 1$ if the 2nd coin lands on Heads, and 0 otherwise.

   Let $X_3 = 1$ if the 3rd coin lands on Heads, and 0 otherwise, etc

   Let $X_{400} = 1$ if the 400th coin lands on Heads, and 0 otherwise.

   The mean for each of these random variables $X_1, X_2, \cdots$ is $\mu = 1/2$ and standard deviation $\sigma = \sqrt{Var(X_1)} = \sqrt{(1/2)(1 - 1/2)} = 1/2$. Then, when $n = 400$, $S_{400} = X_1 + X_2 + ... + X_{400}$ and by the CLT,
$$P\left(S_{400} > 213\right) = P\left(\frac{S_{400} - 400(1/2)}{\sqrt{400}(1/2)} > \frac{213 - 400(1/2)}{\sqrt{400}(1/2)}\right) = P(z > 1.3) = 1 - \phi(1.3) = 1 - .9032 = .0968$$
   Note that this is exactly the approximation of the binomial with a normal random variables, as illustrated in the last lecture.

2. Roll a fair 6-sided die 400 times. Let S_{400} =the sum of values on the die. Find the probability that we rolled a sum of at least 1500.

   Let $X_1, X_2, \ldots$ be the value on die $1, 2, \ldots$. Then they have the same mean and variance, so we only find them for X_1. Since $X_1 \in \{1, 2, 3, 4, 5, 6\}$,

$$E[X_1] = 1 \cdot P(X_1 = 1) + 2 \cdot P(X_1 = 2) + \cdots + 6 \cdot P(X_1 = 6) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$\begin{aligned} Var(X_1) &= (1 - 3.5)^2 \cdot P(X_1 = 1) + (2 - 3.5)^2 \cdot P(X_1 = 2) + \cdots + (6 - 3.5)^2 \cdot P(X_1 = 6) \\ &= (1 - 3.5)^2 \cdot \frac{1}{6} + (2 - 3.5)^2 \cdot \frac{1}{6} + \cdots + (6 - 3.5)^2 \cdot \frac{1}{6} = \frac{35}{12} \end{aligned}$$

   Therefore, by CLT

$$P\left(S_{400} \geq 1500\right) = P\left(\frac{S_{400} - 400(3.5)}{\sqrt{400}\sqrt{35/12}} \geq \frac{1500 - 400(3.5)}{\sqrt{400}\sqrt{35/12}}\right) = 1 - \phi(2.92) = 1 - .9982 = .0018$$

One question to keep in mind is how large does n have to be for a good approximation by a normal? And furthermore, what do we mean by good approximation? We'll partly answer these questions in next lecture.