Class Notes 3-4-2019

• The Law of Large Numbers (LLN):
  Although the LLN can/will be proved without using the CLT we give here a first look at it.
  The LLN states that for any \( p \in [0, 1] \) that for any \( \epsilon > 0 \) we have
  \[
P \left( \left| \frac{S_n}{n} - p \right| < \epsilon \right) \to 1 \quad \text{as } n \to \infty
  \]
  For \( p = 0 \) and \( p = 1 \) this is trivially true since \( P(S_n/n = 0) = 1 \) for \( p = 0 \) and \( P(S_n/n = 1) = 1 \)
  for \( p = 1 \). For \( 0 < p < 1 \) we have from the CLT
  \[
P \left( \left| \frac{S_n}{n} - p \right| < \epsilon \right) = P \left( \left| \frac{\bar{S}/n - p}{\sqrt{p(1-p)/n}} \right| < \frac{\epsilon}{\sqrt{p(1-p)/n}} \right) \approx P \left( \left| Z \right| < \frac{\sqrt{n}\epsilon}{\sqrt{p(1-p)}} \right)
  \]
  \[
  = 2\Phi \left( \frac{\sqrt{n}\epsilon}{\sqrt{p(1-p)}} \right) - 1 \to 1 \quad \text{as } n \to \infty
  \]
  HW 7, Problem 2 lets you work out a specific example case.

• Sampling with and without Replacement:
  Suppose we have a large population of size \( N \) from which we sample at random \( n \) items
  without replacement, either one by one or as a single unordered grab of \( n \). Then the outcome
  \( \omega \) is either an ordered \( n \)-tuple, denoted by \( \omega \), each with probability
  \[
P(\omega) = \frac{1}{N \cdot (N-1) \cdot (N-2) \cdots (N-n+1)} = \frac{1}{(N)_n}
  \]
  or an unordered set of \( n \) items representing all \( n! \) permutations of the previous \( \omega \). We will
  denote this set by \( \{\omega\} \) and it has probability
  \[
P^*(\{\omega\}) = \frac{n!}{N \cdot (N-1) \cdot (N-2) \cdots (N-n+1)} = \frac{1}{(N)_n}
  \]
  Suppose now that our population consists of two types of items, say \( M \) red balls and \( N - M \)
  blue balls. We are interested in the random variable \( X \) that counts the number of red balls in \( \omega \)
  or in \( \{\omega\} \). We can find \( P(X = k) = P(\omega : X(\omega) = k) \) or \( P^*(X = k) = P^*(\{\omega\} : X(\{\omega\}) = k) \).
  Any \( n \)-tuple \( \omega \) with \( k \) red balls in specified position and with blue balls in the other positions
  can be obtained in
  \[
  M \cdot (M-1) \cdots (M-k+1) \cdot (N-M) \cdot (N-M-1) \cdots (N-M-(n-k-1)) = (M)_k(N-M)_{n-k}
  \]
  ways. As far as \( P(X = k) = P(\omega : X(\omega) = k) \) is concerned there are \( \binom{n}{k} \) \( n \)-tuples \( \omega \) which
  have \( k \) red balls and \( n - k \) blues balls in one of the chosen \( \binom{n}{k} \) positions of the \( n \)-tuple. Thus
  \[
P(X = k) = \frac{\binom{n}{k}(M)_k(N-M)_{n-k}}{(N)_n} = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}} = P^*(X = k)
  \]
  where the \( = P^* \) derives from the fact that in order to get a set \( \{\omega\} \) with exactly \( k \) red balls I
  have to grab \( k \) from \( M \) red balls in one of \( \binom{M}{k} \) ways and combine that with a grab of \( n - k \)
blue balls from $N - M$ in $\binom{N-M}{n-k}$ possible ways. This equality $P(X = k) = P^*(X = k)$ should have been made clear in the beginning of this class and it is predicated on the fact that the event $X = k$ does not involve any order. This common distribution of $X$ is referred to as the hypergeometric distribution and we write $X \sim \text{Hypergeom}(N, M, n)$. See HW 4, Problem 4.

Let us next consider sampling $n$ items with replacement from the same population and let $X$ again denote the number of red balls. Then we are dealing with $n$ independent trials and with success (red ball chosen in a given trial) probability $p = M/N$ and we have $X \sim \text{Bin}(n, p)$. We write

$$\tilde{P}(X = k) = \binom{n}{k} \left( \frac{M}{N} \right)^k \left( 1 - \frac{M}{N} \right)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Let $A$ denote the set of all $n$-tuples with distinct elements, assuming all $N$ items are labeled $1, 2, \ldots, N$. Then

$$\tilde{P}(A) = \frac{N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot (N-n+1)}{N^n} \to 1 \quad \text{as } N \to \infty, \ n \text{ staying fixed}$$

and since $\tilde{P}(A^c) \to 0$ as $N \to \infty$ we then have

$$\tilde{P}(X = k) = \tilde{P}(X = k|A)\tilde{P}(A) + \tilde{P}(X = k|A^c)\tilde{P}(A^c) \approx \tilde{P}(X = k|A)$$

For $n$-tuples $\omega \in A$ we have

$$\tilde{P}(\omega|A) = \frac{\tilde{P}(\omega)}{\tilde{P}(A)} = \frac{1}{\frac{N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot (N-n+1)}{N^n}} = \frac{1}{(N)_n} = P(\omega)$$

and thus

$$\tilde{P}(X = k) \approx \tilde{P}(X = k|A) = \tilde{P}(\{\omega : X(\omega) = k\}|A) = P(\omega : X(\omega) = k) = P(X = k) = P^*(X = k)$$

Thus the “two sampling without replacement” schemes give approximately the same distribution to $X$ as the sampling with replacement, provided $N$ is large. In order for the binomial distribution $\text{Bin}(n, p = M/N)$ to stay stable we would also need to assume that $p = M/N$ stay stable and stay away from zero or one. In that case the CLT applied to the binomial distribution $\tilde{P}$ equally applies to the nonreplacement schemes, where the sequential selections (if we sample one by one) are certainly no longer independent.

**The Poisson Distribution and Poisson Approximation to the Binomial:**

If $S_n \sim \text{Bin}(n, p)$ with $p \approx 0$ or $p \approx 1$ presents a problem for the CLT, at least for moderate $n$. In that case the Poisson distribution provides a very useful alternative for $p \approx 0$ and for $p \approx 1$ we switch to $\tilde{S}_n = n - S_n \sim \text{Bin}(n, 1 - p)$, counting the number of failures instead.

A random variable $X$ has a Poisson distribution with parameter $\lambda > 0$, and we write $X \sim \text{Poisson}(\lambda)$ or $X \sim \text{Pois}(\lambda)$, when

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$

Its mean and variance are given by

$$E(X) = \lambda \quad \text{and} \quad \text{var}(X) = \lambda$$
\[ E(X) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = \lambda \]

\[ E(X(X-1)) = \sum_{k=0}^{\infty} k(k-1) \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=2}^{\infty} k(k-1) \frac{e^{-\lambda} \lambda^k}{k!} = \lambda^2 \sum_{k=2}^{\infty} \frac{e^{-\lambda} \lambda^{k-2}}{(k-2)!} = \lambda^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = \lambda^2 \]

Thus \( \text{var}(X) = E(X(X-1)) + E(X) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \)

- **Law of Rare Events Or Law of Small Numbers**
  Let \( \lambda > 0 \) and consider positive integer \( n \) with \( \lambda/n < 1 \). Let \( S_n \sim \text{Bin}(n, \lambda/n) \). Then
  \[
  \lim_{n \to \infty} P(S_n = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{for} \quad k = 0, 1, 2, \ldots
  \]

Proof:
\[
P(S_n = k) = \binom{n}{k} \left( \frac{\lambda}{n} \right)^k \left( 1 - \frac{\lambda}{n} \right)^{n-k} = \frac{n(n-1) \cdots (n-k+1) \lambda^k}{n^k} \frac{1 - \frac{\lambda}{n}}{k!} \left( 1 - \frac{\lambda}{n} \right)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{as} \quad n \to \infty \quad \text{for any fixed} \quad k
\]