Card Drawing

BIOF2014

```r
set.seed(1234);
```

Experimental procedure

1. Ensure that you have a deck of 52 cards (with jokers removed).
2. Shuffle the deck well.
3. Draw 10 cards.
4. Among the drawn cards, how many are face cards (Jack, Queen, or King)? Record this as realizations from the random variable X.
5. Place all cards back into the deck, and repeat steps 1-5 for 10 rounds.
6. Collect the realizations from all groups together. If there are 12 groups, there now should be 120 samples of X.
7. Plot a histogram with a bin size of 1 (i.e. empirical probability mass function) of X.

We've just create a distribution known as the hypergeometric distribution!

Hypergeometric distribution

A hypergeometric distribution is useful in finite population sampling where the probability of sampling each element of the population is uniform.

The hypergeometric distribution describes the probability of observing k successes from an *unordered* draw of sample size n from a finite population with containing K successes from a population size of N.

The hypergeometric pmf is given by

$$P_X(k \mid N, K, n) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \quad k \in \{0, 1, \dots, K\},$$

where $k \leq K$ and $n - k \leq N - K$.

To derive the hypergeometric distribution, we need to determine two numbers:

M, the total number of possible combinations of samples. m, the number of possible combinations that satisfy the specified number of success k.

Then, the probability of observing k successes is $\frac{m}{M}$.

The total number of possible combinations of a sample of n from a population of N is $M = \binom{N}{n}$.

As for m, the k observed successes come from K successes in the population, and there are $\binom{K}{k}$ combinations for the observed successes. Conversely, the $n-k$ observed failures come from $N-K$ failures in the population, and there are $\binom{N-K}{n-k}$ combinations for these observed failures.

Therefore, $m = \binom{K}{k}\binom{N-K}{n-k}$.

Finally,

$$P_X(k) = \frac{m}{M} = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}. \quad \blacksquare$$

Notably, the hypergeometric distribution forms the basis for the well-known Fisher's exact test.

Binomial distribution

We can show that as $N \to \infty$ while keeping $\frac{K}{N}$ fixed, Hypergeometric $(k \mid N, K, n)$ converges to Binomial $\left(k \mid n, \frac{K}{N}\right)$.

Questions

Assume that you have a standard deck of 52 playing cards.

1. What is the probability of drawing a face card from a deck?
2. What is the difference in the probability of drawing an Ace of spades vs. a Two of clubs from a deck?
3. Plot out the theoretical hypergeometric pmf we would expect to observe in the card drawing experiment we performed above. Compare the empirical and the theoretical pmf.