Math 373 Lecture 7

Exam 1, one week from today, Lectures 1-8

ROUNDING. Round percentages to the nearest digit, all other answers to the nearest two significant decimal places. Exception, for percentages near 0, leave the first significant 10th's or 100th's digit. If none, write .00 .

33.33	33.33%	.2449	.245%	.025%	.0025%
33.33	33%	.24	.3%	.03%	.00%

Don't use rounded numbers in calculations, round only final answers.

When to use binomial formulas with random samples

In a binomial experiment of n trials, the probability p of success must not change from trial to the next. If the experiment draws n samples from a population of N items, the process of removing the sample items changes the population size and the probability of success. When $n < (.05)N$, the sample size is comparatively small. We ignore these changes and the sampling process is approximately binomial.

■ A crate has 100 apples (population size $N = 100$), five of which are rotten. We select a sample of n apples and count the number x of rotten apples in the sample. Find $P(x=k)$, μ, and σ.

Suppose we select 3 apples (sample size $n = 3$), one at a time. Then $n = 3 < 5 = (.05)100 = (.05)N$. $\therefore$ may use the binomial formula with $p = 5/100 = .05$ and $q = .95$.

$$P(x=k) = \frac{n!}{k!(n-k)!}p^k q^{n-k}, \quad \mu = np, \quad \sigma = \sqrt{npq}\,.$$

Now suppose 10 apples (sample size $n = 10$) apples from 100 apples (population size $N = 100$) with 5 rotten. Then $n = 10 > 5 = (.05)100 = (.05)N$. The fact that withdrawing a sample reduces the population can not be ignored. Thus we must use the following *hypergeometric* formulas for $P(x=k)$, μ, and σ.

Hypergeometric Distributions

Suppose we have a population of N balls with M red balls (the "successes") and $N-M$ blue balls ("failures"). Suppose a sample of n balls is withdrawn. Let x be the number of red balls (x = # of successes).

There are $\binom{N}{n} = C_n^N = \frac{N!}{n!(N-n)!}$ ways of choosing n sample balls from a population of N balls.

For $k = 0, 1, \ldots, n$, how many ways are there of choosing n balls with exactly k red balls? To choose such a set, you must choose k red balls and $n-k$ blue balls. The total number of ways this can be done is $\binom{M}{k}\binom{N-M}{n-k}$.

Thus $P(x=k)$, the probability that k sample balls are red,

= (number of ways of drawing k red balls and $n-k$ blue balls) / (total number of ways of drawing n balls)

$= \binom{M}{k}\binom{N-M}{n-k} / \binom{N}{n}$.

When choosing samples, why should we ever use a binomial approximation instead of the more exact hypergeometric distribution? In the hypergeometric case, $P(x = k)$ involves $N!$ which is usually be too big to calculate. The binomial formula case uses only the smaller $n!$.

HYPERGEOMETRIC THEOREM.

Suppose a population has N items. Of the N items, M are labeled "success" and $N-M$ are "failures".

A sample of n items is drawn.

Let x = the number of successes in the n item sample. Then

- $\mu = n\left(\frac{M}{N}\right)$,
- $\sigma = \sqrt{n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)}$,
- $P(x=k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ for $k = 0, 1, \ldots, n$.

The probability of success changes only when more than one sample item is withdrawn. Thus if only one element is sampled this should be just the binomial distribution. In this case the sample size is $n = 1$, the probability of drawing a success is $p = M/N$, the probability of failure is $q = (N-M)/N$, and μ and σ simplify to the binomial case:

$$\mu = n\left(\frac{M}{N}\right) = np,$$

$$\sigma = \sqrt{n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)} = \sqrt{npq}\,.$$

What if the entire population is drawn? Then $n = N$ and there is only one way the sample can be drawn and this one sample has M successes. Since there is only this one sample, the mean is $\mu = M$, the std. dev. is $\sigma = 0$. This is just what the formulas give:

$$\mu = n\left(\frac{M}{N}\right) = N\left(\frac{M}{N}\right) = M,$$

$$\sigma = \sqrt{N\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-N}{N-1}\right)} = 0.$$

■ In a shipment of 100 apples, 5 are rotten. A sample of 4 apples is taken and x is the number of rotten apples in the sample.

Find the average number of rotten apples in the sample?

Find the std. dev.

Find the probability that all the sample is rotten?

Find the probability that none are?

Find $P(x \le 1)$.