Section 3.8 : Hypergeometric Probability Distribution

Thursday, January 31, 2019 2:40 PM

note: we are omitting 3.7: Geometric / Negative Binomial

---

remember the rule to approximate with a binomial if you are sampling without replacement

$\frac{n}{N} \geq \frac{1}{20}$ Cannot approximate with binomial

n = number of trials
N = population size

why? because you need p to be more-or-less constant to approximate with a binomial

so, what if you are sampling without replacement and

$\frac{n}{N} \geq \frac{1}{20}$?

use hypergeometric probability distribution instead

---

hypergeometric distribution:

population with a total number of N contains M successes and $N-M$ failures

choose without replacement

the probability of exactly k successes in a random sample of size n is

$$P(X=k) = \frac{{}_M C_k \; {}_{N-M} C_{n-k}}{{}_N C_n}$$

with mean

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

and variance:

$$\sigma^2 = n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

↑ this is a correction factor for the finite population

---

example: A case of wine (12 bottles in total) has 5 bottles which contain spoiled wine. If

three bottles are randomly sampled, what is the probability distribution of x, the number of spoiled bottles sampled?

- sampling without replacement $N = 12$, $n = 3$

hypergeometric (sample size $\geq$ 5% of population)

$N = 12$
$n = 3$
$M = 5$ (success is spoiled wine)
$N - M = 7$
$k = x$

so $p(x) = \dfrac{{}_5C_x \; {}_7C_{3-x}}{{}_{12}C_3}$

x	p(x)	
0	$7/44$	≈ 0.159091
1	$21/44$	≈ 0.477273
2	$7/22$	≈ 0.318182
3	$1/22$	≈ 0.045455

note: sum is 1

what is the mean value of x?

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

$= 3\left(\frac{5}{12}\right) = 1.25$

what is the standard deviation of x?

$$\sigma^2 = n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

$$= 3\left(\frac{5}{12}\right)\left(\frac{7}{12}\right)\left(\frac{9}{11}\right)$$

$\sigma = 0.772$ ≈ 0.77 or 0.8

a case of wine will be rejected if, when three bottles are randomly sampled, one or more bottles is found to be spoiled. What is the probability that a case with 5 spoiled bottles will be accepted?

$P(x=0) = 15.9\%$ (oops)