LECTURE 16

Agenda:

① Hypergeometric distribution
② Summary

HYPERGEOMETRIC DISTRIBUTION

Suppose we have N dichotomous objects, i.e, N objects each of which is Type I or Type II. Suppose we know that k of the objects are Type I.

<u>Experiment:</u> Draw n objects (WOR) from N objects.

X = # of objects drawn of Type I.

Note that the set of possible values that X can take is $\{0, 1, 2, ---, k-1, k\}$. Hence,

$$\mathcal{X} = \text{Range}(X) = \{0, 1, ---, k-1, k\}.$$

$$P(X=x) = \frac{\overbrace{\binom{k}{x}\binom{N-k}{n-x}}^{\text{\# of ways of choosing } n \text{ objects out of } N \text{ objects (WOR) such that } x \text{ are of type I}}}{\underbrace{\binom{N}{n}}_{\text{\# of ways of choosing } n \text{ objects out of } N \text{ objects.}}}$$

$x = 0, 1, ---, k.$

Since $\sum_{x=0}^{k} P(X=x) = 1$, it follows that

$$\sum_{x=0}^{k} \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = 1. \quad \text{——} \quad (*)$$

Note that the identity holds for any fixed integers N, k, n.

$$E(X) = \sum_{x=0}^{k} x \, \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=1}^{k} x \, \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=1}^{k} \frac{k\binom{k-1}{x-1}\binom{N-k}{n-x}}{\frac{N}{n}\binom{N-1}{n-1}} \qquad \left(\begin{array}{l} \because x\binom{k}{x} = k\binom{k-1}{x-1} \\ \binom{N}{n} = \frac{N}{n}\binom{N-1}{n-1} \end{array} \right)$$

$$= \frac{nk}{N} \sum_{x=1}^{k} \frac{\binom{k-1}{x-1}\binom{(N-1)-(k-1)}{(n-1)-(x-1)}}{\binom{N-1}{n-1}}$$

$$= \frac{nk}{N} \sum_{x'=0}^{k-1} \frac{\binom{k-1}{x'\,\text{[redacted]}}\binom{(N-1)-(k-1)}{(n-1)-\text{[redacted]}x'\,\text{[redacted]}}}{\binom{N-1}{n-1}},$$

where $x' = x-1$

But using (*) with $N-1$, $k-1$, $n-1$ in place of N, k, n, we get

$$\sum_{x'=0}^{k-1} \frac{\binom{k-1}{x'}\binom{(N-1)-(k-1)}{(n-1)-x'}}{\binom{N-1}{n-1}} = 1.$$

Hence, $E(X) = \frac{nk}{N}$.

Similarly,

$$V(X) = n\left(\frac{k}{N}\right)\left(1-\frac{k}{N}\right)\left(\frac{N-n}{N-1}\right).$$

<u>Example:</u> Two positions are open in a company. Ten men and five women have applied for a job, and all are equally qualified. The manager randomly picks two people to fill the position. What is the probability that one man and one woman are chosen?

This is a hypergeometric experiment with Type I = man, Type II = woman, $N = 15$, $k = 10$, $n = 2$.

X = # of men chosen for the two positions

$$P(\text{One man, one woman}) = P(X=1) = \frac{\binom{10}{1}\binom{5}{1}}{\binom{15}{2}}$$

$$= \frac{10 \times 5}{\frac{15 \times 14}{2!}}$$

$$= \frac{10}{21}.$$

BINOMIAL APPROXIMATION TO HYPERGEOMETRIC

It can be proved that as $N \to \infty$ and $\frac{k}{N} \to p$

$$P(X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \longrightarrow \binom{n}{x} p^{x} (1-p)^{n-x}.$$

Hence if the number of objects N is very large, the hypergeometric random variable can be approximated by a binomial random variable with parameters n and $p = \frac{k}{N}$.

But unless the question says that N is large and you can approximate the experiment by a binomial experiment, always assume that the experiment is a hypergeometric experiment.

SUMMARY

In the past 3 weeks, we have studied 6 random variables. Here is a summary of their basic properties.

BERNOULLI RANDOM VARIABLE

Experiment: Any experiment with two outcomes, success and failure.

$$X = \begin{cases} 0 & \text{if Outcome is failure} \\ 1 & \text{if Outcome is success} \end{cases}$$

Parameters: p = Probability of success

$\mathcal{X}$ = Range $(X) = \{0, 1\}$.

P.M.F. : $P(X = x) = p^{x}(1-p)^{1-x}$, $x = 0, 1$.

$E(X) = p$

$V(X) = p(1-p)$

BINOMIAL RANDOM VARIABLE

Experiment: Repeat n Bernoulli trials (independently)

X = # of successes

Parameters : n = Repetitions, p = Probability of success in a single Bernoulli experiment.

$\mathcal{X}$ = Range $(X) = \{0, 1, 2, ---, n-1, n\}$

P.M.F. : $P(X = x) = \binom{n}{x} p^{x}(1-p)^{n-x}$,

$x = 0, 1, ---, n$

$E(X) = np$

$V(X) = np(1-p)$

GEOMETRIC RANDOM VARIABLE

Experiment: Repeat Bernoulli trials (independently) until first success.

X = # of failures before first successes

Parameters: p = Probability of success in a single Bernoulli experiment

$\mathcal{X} = \text{Range}(X) = \{0, 1, 2, \text{----------}\}$

P.M.F. : $P(X = x) = (1-p)^x p$, $x = 0, 1, 2, \text{-------}$

$$E(X) = \frac{1-p}{p}$$

$$V(X) = \frac{1-p}{p^2}$$

NEGATIVE BINOMIAL RANDOM VARIABLE

Experiment: Repeat Bernoulli trials (independently) until r^{th} success.

X = # of failures before the r^{th} success

Parameters: r = Number of successes after which we stop the experiment

p = Probability of success in a singe Bernoulli trial.

$\mathcal{X} = \text{Range}(X) = \{0, 1, 2, \text{----------}\}$

P.M.F. : $P(X = x) = \binom{x+r-1}{x} p^r (1-p)^x$,

$x = 0, 1, 2, \text{-------}\ldots$

$$E(X) = \frac{r(1-p)}{p}$$

$$V(X) = \frac{r(1-p)}{p^2}$$

POISSON RANDOM VARIABLE

<u>Experiment:</u> This random variable arises from experiments which can be approximated by Binomial experiments with LARGE n, SMALL p and $np \to \lambda > 0$. It is generally used to model the number of times a certain event occurs in a given time frame or a given area.

Parameters: $\lambda > 0$.

Range(X) $= \{0, 1, 2, \text{---------}\}$

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

$E(X) = \lambda$

$V(X) = \lambda$

HYPERGEOMETRIC RANDOM VARIABLE

<u>Experiment:</u> Draw n objects from N objects of two types, type I and type II. The objects are drawn without replacement.

X = # of objects of type I

Parameters:
N = Total number of objects
n = Number of objects drawn
k = Total number of objects of Type I.

$\mathfrak{X} = \text{Range}(X) = \{0, 1, ---, k\}$

P.M.F: $P(X = x) = \dfrac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, ---, k$

$E(X) = \dfrac{nk}{N}$

$V(X) = n\left(\dfrac{k}{N}\right)\left(1 - \dfrac{k}{N}\right)\left(\dfrac{N-n}{N-1}\right)$