

3-8 Hypergeometric Distribution

Definition

A set of N objects contains

K objects classified as successes

$N - K$ objects classified as failures

A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$.

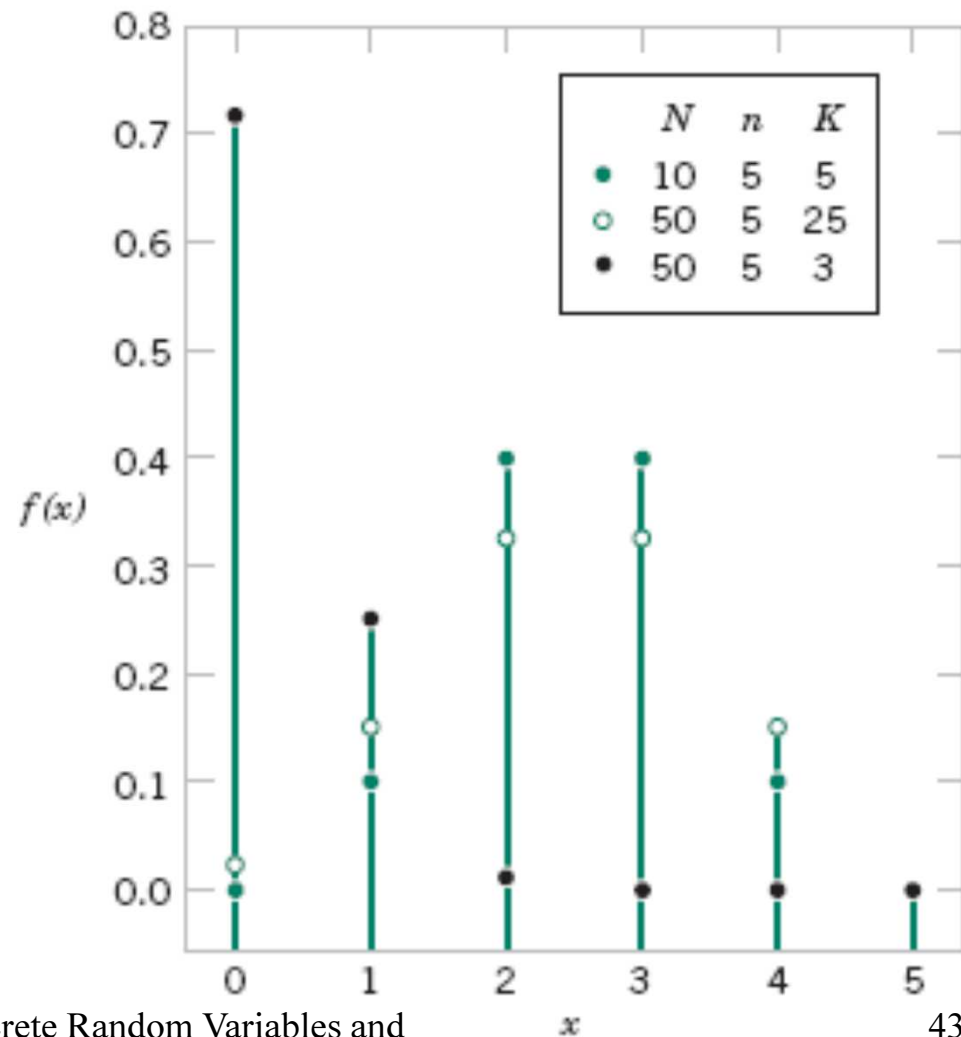
Let the random variable X denote the number of successes in the sample. Then X is a **hypergeometric random variable** and

$$f(x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} \quad x = \max\{0, n + K - N\} \text{ to } \min\{K, n\} \quad (3-13)$$

3-8 Hypergeometric Distribution

Figure 3-12.

Hypergeometric distributions for selected values of parameters N , K , and n .



3-8 Hypergeometric Distribution

Example 3-27

A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in the next state. If four parts are selected randomly and without replacement, what is the probability they are all from the local supplier?

Let X equal the number of parts in the sample from the local supplier. Then, X has a hypergeometric distribution and the requested probability is $P(X = 4)$. Consequently,

$$P(X = 4) = \frac{\binom{100}{4} \binom{200}{0}}{\binom{300}{4}} = 0.0119$$

3-8 Hypergeometric Distribution

Example 3-27

What is the probability that two or more parts in the sample are from the local supplier?

$$\begin{aligned} P(X \geq 2) &= \frac{\binom{100}{2} \binom{200}{2}}{\binom{300}{4}} + \frac{\binom{100}{3} \binom{200}{1}}{\binom{300}{4}} + \frac{\binom{100}{4} \binom{200}{0}}{\binom{300}{4}} \\ &= 0.298 + 0.098 + 0.0119 = 0.408 \end{aligned}$$

What is the probability that at least one part in the sample is from the local supplier?

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{100}{0} \binom{200}{4}}{\binom{300}{4}} = 0.804$$

3-8 Hypergeometric Distribution

Mean and Variance

If X is a hypergeometric random variable with parameters N, K , and n , then

$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1 - p) \left(\frac{N - n}{N - 1} \right) \quad (3-14)$$

where $p = K/N$.

Here p is interpreted as the proportion of successes in the set of N objects.

3-8 Hypergeometric Distribution

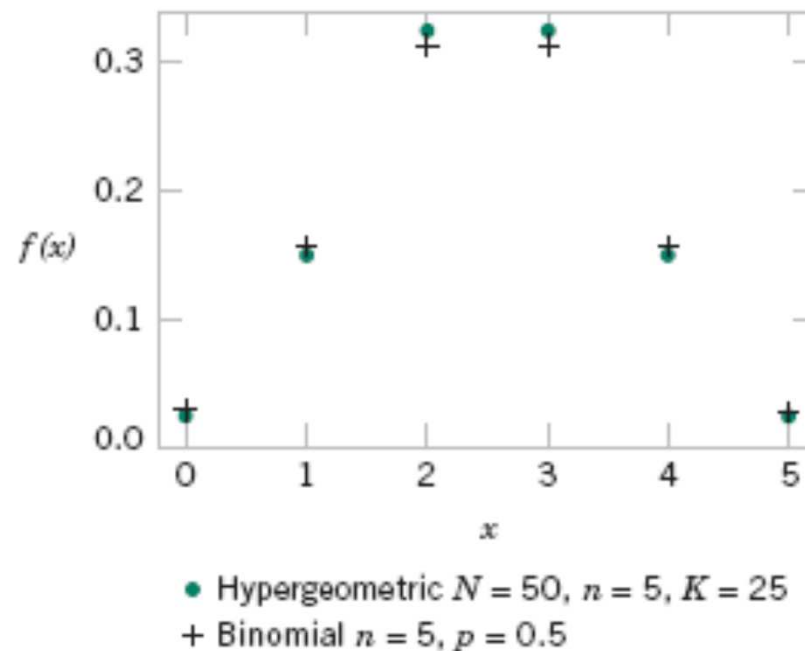
Finite Population Correction Factor

The term in the variance of a hypergeometric random variable

$$\frac{N - n}{N - 1} \quad (3-15)$$

is called the finite population correction factor.

3-8 Hypergeometric Distribution



	0	1	2	3	4	5
Hypergeometric probability	0.025	0.149	0.326	0.326	0.149	0.025
Binomial probability	0.031	0.156	0.321	0.312	0.156	0.031

Figure 3-13. Comparison of hypergeometric and binomial distributions.

3-9 Poisson Distribution

Example 3-30

Consider the transmission of n bits over a digital communication channel. Let the random variable X equal the number of bits in error. When the probability that a bit is in error is constant and the transmissions are independent, X has a binomial distribution. Let p denote the probability that a bit is in error. Let $\lambda = pn$. Then, $E(x) = pn = \lambda$ and

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Now, suppose that the number of bits transmitted increases and the probability of an error decreases exactly enough that pn remains equal to a constant. That is, n increases and p decreases accordingly, such that $E(X) = \lambda$ remains constant. Then, with some work, it can be shown that

$$\binom{n}{x} \left(\frac{\lambda}{n}\right)^x \rightarrow 1 \quad \left(1 - \frac{\lambda}{n}\right)^{-x} \rightarrow 1 \quad \left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda}$$

so that

$$\lim_{n \rightarrow \infty} P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

Also, because the number of bits transmitted tends to infinity, the number of errors can equal any non-negative integer. Therefore, the range of X is the integers from zero to infinity.

3-9 Poisson Distribution

Definition

Given an interval of real numbers, assume events occur at random throughout the interval. If the interval can be partitioned into subintervals of small enough length such that

- (1) the probability of more than one event in a subinterval is zero,
- (2) the probability of one event in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and
- (3) the event in each subinterval is independent of other subintervals, the random experiment is called a **Poisson process**.

The random variable X that equals the number of events in the interval is a **Poisson random variable** with parameter $0 < \lambda$, and the probability mass function of X is

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \dots \quad (3-16)$$

3-9 Poisson Distribution

Consistent Units

It is important to **use consistent units** in the calculation of probabilities, means, and variances involving Poisson random variables. The following example illustrates unit conversions. For example, if the

average number of flaws per millimeter of wire is 3.4, then the average number of flaws in 10 millimeters of wire is 34, and the average number of flaws in 100 millimeters of wire is 340.

3-9 Poisson Distribution

Example 3-33

Contamination is a problem in the manufacture of optical storage disks. The number of particles of contamination that occur on an optical disk has a Poisson distribution, and the average number of particles per centimeter squared of media surface is 0.1. The area of a disk under study is 100 squared centimeters. Find the probability that 12 particles occur in the area of a disk under study.

Let X denote the number of particles in the area of a disk under study. Because the mean number of particles is 0.1 particles per cm^2

$$E(X) = 100 \text{ cm}^2 \times 0.1 \text{ particles/cm}^2 = 10 \text{ particles}$$

Therefore,

$$P(X = 12) = \frac{e^{-10} 10^{12}}{12!} = 0.095$$

3-9 Poisson Distribution

Example 3-33

The probability that zero particles occur in the area of the disk under study is

$$P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$$

Determine the probability that 12 or fewer particles occur in the area of the disk under study. The probability is

$$P(X \leq 12) = P(X = 0) + P(X = 1) + \cdots + P(X = 12) = \sum_{i=0}^{12} \frac{e^{-10} 10^i}{i!}$$

3-9 Poisson Distribution

Mean and Variance

If X is a Poisson random variable with parameter λ , then

$$\mu = E(X) = \lambda \quad \text{and} \quad \sigma^2 = V(X) = \lambda \quad (3-17)$$

IMPORTANT TERMS AND CONCEPTS

Bernoulli trial	Finite population	Mean-function of a	Probability mass
Binomial distribution	correction factor	discrete random	function
Cumulative probability	Geometric distribution	variable	Standard deviation-
distribution function-	Hypergeometric distri-	Negative binomial	discrete random
discrete random	bution	distribution	variable
variable	Lack of memory	Poisson distribution	Variance-discrete
Discrete uniform distri-	property-discrete	Poisson process	random variable
bution	random variable	Probability distribution-	
Expected value of a	Mean-discrete random	discrete random	
function of a random	variable	variable	
variable			