MATH F113 (Probability and Statistics)

Chandra Shekhar
Associate Professor

Department of Mathematics
BITS Pilani, Pilani Campus, Rajasthan 333 031
Email: chandrashekhar@pilani.bits-pilani.ac.in
Mobile: 9414492349

Chapter 1

What is Statistics?
Interpretation of probability
Determination of probability
Sampling without and with replacement
Sampling without and with order
Permutation and combination

Chapter 2

Axiomatic definition of probability
Conditional probability
Independence
Total probability
Bayes Theorem

Chapter 3

Discrete distribution
Geometric distribution
Binomial distribution
Bernoulli distribution
Hyper-geometric distribution
Uniform distribution
Poisson distribution

Binomial Approximation
(for Hyper-geometric Distribution)

- **When the sample size n is small compared to population size N, we can use binomial distribution even when sampling is without replacement.**
- **This is done if $n/N \leq 0.05$. The parameters are n and $p = r/N$.**

Binomial Approximation
(for Hyper-geometric Distribution)

- If n is small relative to N, then the composition of the sampled group does not change much from trial to trial even we are keeping the sampled items

Example 3.7.3: During a course of an hour, 1000 bottles of beer are filled by a particular machine. Each hour a sample of 20 bottles is randomly selected and number of ounces of beer per bottle is checked. Let X denote the number of bottles selected that are under filled. Suppose during a particular hour, 100 under filled bottles are produced.

Example 3.7.3: Find the probability that at least 3 under filled bottles will be among those sampled.

Using Hyper-geometric Distribution

Required Probability

$$
\begin{aligned}
& P(X \geq 3) \\
=& 1 - P(X \leq 2) \\
=& 1 - P(X = 0) - P(X = 1) - P(X = 2) \\
=& 1 - \frac{\binom{100}{0}\binom{900}{20}}{\binom{1000}{20}} - \frac{\binom{100}{1}\binom{900}{19}}{\binom{1000}{20}} - \frac{\binom{100}{2}\binom{900}{18}}{\binom{1000}{20}} \\
=& 0.3224
\end{aligned}
$$

Using Binomial Approximation

Check $n/N = 20/1000 = 0.02 \leq 0.05$ **(true)**

hence,

$n = 20$, $p = r/N = 100/1000 = 0.1$

Required Probability

$$\begin{aligned} P(X \geq 3) =& 1 - P(X \leq 2) \\ =& 1 - F(2) = 1 - 0.6769 \\ =& 0.3231 \end{aligned}$$

Exercise 59: A distributor of computer software wants to obtain some customer feedback concerning its newest package. 3000 customers have purchased the package. Assume that 600 of these customers are dissatisfied with the product. 20 customers are randomly sampled and questioned about the package.

Let X denote the number of dissatisfied customers sampled.
(a) Find the density for X? (b) Find $E(X)$ and $Var(X)$. (c) Set up the calculations needed to find $P(X \leq 3)$. (d) Use the binomial tables to approximate $P(X \leq 3)$.

(a) Find the density for X?

$X \sim$ **Hyper-geometric Distribution with**

$N = 3000, r = 600, n = 20)$

$$P(X = x) = f(x) = \frac{\binom{600}{x}\binom{2400}{20-x}}{\binom{3000}{20}}$$

$max(0, 20 - (3000 - 600)) \leq x \leq min(20, 600)$

$x = 0, 1, 2, ..., 19, 20$

(b) Find $E(X)$ and $Var(X)$.

$$\begin{aligned}
E(x) =& n\,\frac{r}{N} = 20\,\frac{600}{3000} = 4\\
Var(x) =& n\,\frac{r}{N}\,\frac{N-r}{N}\,\frac{N-n}{N-1}\\
=& 20\,\frac{600}{3000}\,\frac{2400}{3000}\,\frac{2980}{2999}\\
=& 3.1797
\end{aligned}$$

(c) Set up the calculations needed to find $P(X \leq 3)$.

$$P(X \leq 3) = \sum_{x=0}^{3} \frac{\binom{600}{x}\binom{2400}{20-x}}{\binom{3000}{20}}$$

(d) Use the binomial tables to approximate $P(X \leq 3)$**.**

Check $n/N = 20/3000 = 0.0067 \leq 0.05$ **(true)**

hence, $n = 20$, $p = r/N = 600/3000 = 0.2$

Required Probability

$$P(X \leq 3) = F(3) = 0.4114$$

Note: Sometimes population size is large but not known. Proportion of favorable population is given. Then we can use binomial distribution for both sampling with or without replacement where p is the proportion of favorable population.

Poisson Approximation
(for Binomial Distribution)

- **If a binomial random variable X has parameter p very small ($p \leq 0.01$) and n very large ($n \geq 100$) so that $np = k$ is moderate, then X can be approximated by a Poisson random variable Y with parameter k**

Exercise 81: A computer terminal can pick up an erroneous signal from the keyboard that does not show up on the screen. This creates a silent error that is difficult to detect. Assume that for a particular keyboard the probability that this will occur per entry is 1/1000. In 12,000 entries find the probability that no silent error occur.

Find the probability of at least one silent error.

X: **Number of silence error out of 12,000 entries**

$X \sim BD(n = 12,000, p = 1/1000)$

Using Poisson approximation

$X \sim PD(k = np = 12)$

P[no silent error occur]

$P[X = 0] = e^{-12}$

P[at least one silent error]

$P[X \geq 1] = 1 - P[X = 0] = 1 - e^{-12}$