Chapter VII

Some Special Discrete Distributions

D. 7. 1. (*Hypergeometric Distribution*)
A discrete variable X has a hypergeometric distribution if its probability function is of the form

$$P(X = x) = p_i = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}},$$

$$x = 0, 1, \ldots, n; \quad n \le M \le N .$$

R. 7. 1.
The probability function of the hypergeometric distribution is formally equivalent to the sampling scheme *without replacement*.

R. 7. 2.
The requirements for a hypergeometric experiment are as follows:

1. There must be a fixed number of trials.
2. Trials must be independent.
3. All outcomes of trials must be of two categories.
4. Probabilities must remain constant for each trial.

T. 7. 1.
Let X have a hypergeometric distribution. Then

$$E(X) = n \cdot \frac{M}{N}$$

$$= n \cdot p .$$

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

$$= \frac{N-n}{N-1} n \cdot p \cdot q .$$

Proof:

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

$$D^2(X) = E\left(X^2\right) - \left(E(X)\right)^2$$

$$= \sum_{x=0}^{n} \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} \cdot x^2 - \left(n \cdot \frac{M}{N}\right)^2$$

$$= n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N-n}{N-1}$$

$$= \frac{N-n}{N-1} \cdot n \cdot p \cdot q \,.$$

<u>Ex. 7. 1.</u>

A box contains 25 items, 10 of which are defective. A sample of two items will be taken, *without replacement:*

1. Find the probability and the distribution function.
2. What is the probability that

   a) none of the items is defective?

   b) at most one of the items is defective?

3. Calculate the expected value, the variance, and the standard deviation of the random variable.

Solution:

Let X denote the number of defective items in the sample. Then, we have

$$N = 25, \quad M = 10, \quad n = 2\,.$$

1.

$$P(X = 0) = \frac{\binom{10}{0} \cdot \binom{25-10}{2-0}}{\binom{25}{2}} = \frac{7}{20} = 0.35 \cdot$$

$$P(X = 1) = \frac{\binom{10}{1} \cdot \binom{25-10}{2-1}}{\binom{25}{2}} = \frac{10}{20} = 0.50$$

$$P\left(X=2\right)=\frac{\binom{10}{2}\cdot\binom{25-10}{2-2}}{\binom{25}{2}}=\frac{3}{20}=0.15$$

Probability function:

x_i	0	1	2
$P(X=x_i)$	0.35	0.50	0.15

Distribution function:

$$F(x)=\begin{cases}0.00 & \text{when} \quad -\infty<x\le 0\\ 0.35 & \text{when} \quad 0<x\le 1\\ 0.85 & \text{when} \quad 1<x\le 2\\ 1.00 & \text{when} \quad 2<x<+\infty\end{cases}$$

2.

a)

$$P(X=0)=0.35\,.$$

b)

$$P\left(X<2\right)=F(2)=0.85.$$

3.

$$p=\frac{M}{N}=\frac{10}{25}=\frac{2}{5}, \qquad q=1-\frac{2}{5}=\frac{3}{5},$$
$$E\left(X\right)=n\cdot p=2\cdot\frac{2}{5}=0.8\,,$$
$$D^2\left(X\right)=\frac{N-n}{N-1}\cdot n\cdot p\cdot q=\frac{25-2}{25-1}\cdot 2\cdot\frac{2}{5}\cdot\frac{3}{5}=0.46\,,$$
$$D(X)=0.678232998\,.$$

D. 7. 2. (***Binomial Distribution***)

A discrete variable X has a binomial distribution if its probability function is of the form

$$P\left(X=x\right)=p_i=\binom{n}{x}\cdot p^x\cdot q^{n-x}, \quad x=0,\ 1,\ldots,n\,.$$

R. 7. 3.

The probability function of the binomial distribution is formally equivalent to the sampling scheme with replacement.

R. 7. 4.

The requirements for a binomial experiment are as follows:

1. There must be a fixed number of trials.
2. Trials must be independent.
3. All outcomes of trials must be of two categories.
4. Probabilities must remain constant for each trial.

T. 7. 2

Let X have a binomial distribution. Then

$$E(X) = n \cdot p\,,$$

$$D^2(X) = n \cdot p \cdot q\,.$$

Proof:

$$E(X) = \sum_{x=0}^{n} x \cdot \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} = n \cdot p\,,$$

$$D^2(X) = \sum_{x=0}^{n} x^2 \cdot \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} - n^2 \cdot p^2 = n \cdot p \cdot (1-p) = n \cdot p \cdot q\,.$$

Ex. 7. 2.

A box contains 25 items, 10 of which are defective. A sample of two items will be taken, *with replacement:*

1. Find the probability and the distribution function.
2. What is the probability that

   a. none of the items is defective?

   b. at most one of the items is defective?

3. Calculate the expected value, the variance, and the standard deviation of the random variable.

Solution

Let X denote the number of defective items in the sample. Then, we have

$$n = 2, \quad p = \frac{M}{N} = \frac{10}{25} = \frac{2}{5}, \quad q = \frac{3}{5}.$$

1.

$$P(X=0)=\binom{2}{0}\cdot\left(\frac{2}{5}\right)^{0}\cdot\left(\frac{3}{5}\right)^{2-0}=\frac{9}{25}=0.36\,.$$

$$P(X=1)=\binom{2}{1}\cdot\left(\frac{2}{5}\right)^{1}\cdot\left(\frac{3}{5}\right)^{2-1}=\frac{12}{25}=0.48$$

$$P(X=2)=\binom{2}{2}\cdot\left(\frac{2}{5}\right)^{2}\cdot\left(\frac{3}{5}\right)^{2-2}=\frac{4}{25}=0.16$$

Probability function:

x_i	0	1	2
$P(X=x_i)$	0.36	0.48	0.16

Distribution function:

$$F(x)=\begin{cases}0.00 & \text{when} \quad -\infty < x \le 0\\ 0.36 & \text{when} \quad 0 < x \le 1\\ 0.84 & \text{when} \quad 1 < x \le 2\\ 1.00 & \text{when} \quad 2 < x < +\infty\end{cases}$$

2.

a)

$$P(X=0)=0.36\,.$$

b)

$$P(X<2)=F(2)=0.84.$$

3.

$$E(X)=n\cdot p=2\cdot\frac{2}{5}=0.8\,,$$
$$D^2(X)=n\cdot p\cdot q=2\cdot\frac{2}{5}\cdot\frac{3}{5}=0.48\,,$$
$$D(X)=0.692820323\,.$$

T. 7. 3.

$$\lim_{N\to+\infty}\frac{\binom{M}{x}\cdot\binom{N-M}{n-x}}{\binom{N}{n}}=\binom{n}{x}\cdot p^{x}\cdot q^{n-x}.$$

R. 7. 5.
For a "sufficiently" large N , the hypergeometric distribution can be approximated by the binomial distribution. It will be recommended to use the following rule of thumb:

> "If $10\cdot n\leq N$, then the hypergeometric distribution can be approximated by the binomial distribution."

Ex. 7. 3.
A manufacturer of car tires reports that among a shipment of 5000 sent to a local distributor, 1000 are slightly blemished. If one purchases 10 of these tires at random from the distributor, what is the probability that exactly 3 will be blemished?

Solution:
We have

$$N=5000,\qquad M=1000,\qquad n=10.$$

$$P(X=3)=p_3=\frac{\binom{1000}{3}\cdot\binom{5000-1000}{10-3}}{\binom{5000}{10}}=0.201477715\cdot$$

Because of

$$10n=10\cdot 10=100\leq 5000=N$$

the probability can also be calculated as follows:

$$n=10,\quad p=\frac{M}{N}=\frac{1000}{5000}=\frac{1}{5},\quad q=\frac{4}{5},$$

$$P(X=3)=p_3=\binom{10}{3}\cdot\left(\frac{1}{5}\right)^{3}\cdot\left(\frac{4}{5}\right)^{10-3}=0.201326592.$$

R. 7. 6.

$$\lim_{N\to+\infty}\frac{N-n}{N-1}\cdot n\cdot p\cdot q=\lim_{N\to+\infty}\frac{1-\frac{n}{N}}{1-\frac{1}{N}}\cdot n\cdot p\cdot q=n\cdot p\cdot q.$$

D. 7. 3. ***(Poisson Distribution)***

A discrete variable X has a Poisson distribution if its probability function is of the form

$$P(X = x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda},$$
$$x = 0, 1, \ldots, n .$$

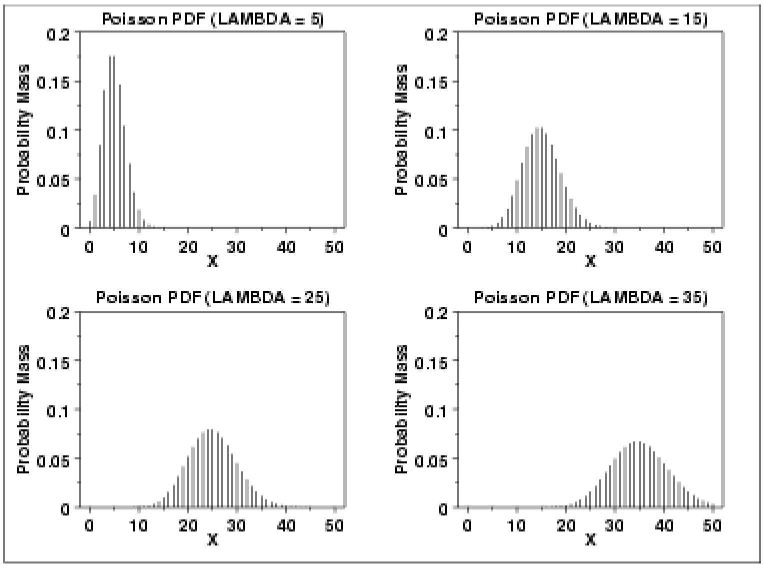
T. 7. 4.

Let X have a Poisson distribution. Then

$$E(X) = D^2(X) = n \cdot p = \lambda .$$

R. 7. 7.

The following is the plot of the Poisson probability function for four values of λ



Ex. 7. 4.

Assume that the number of defects per square meter of a certain type of cloth manufactured by a mill is measured as no defects, one defect, two defects, and so on. In the average, the number of defects is 0.5.

Compute the probabilities that a square meter will have

a) no defects
b) one defect
c) two defects.

Solution:

We have $\lambda = 0.5$.

a)

$$P(X = 0) = \frac{0.5^0}{0!} \cdot e^{-0.5} = 0.606530659 .$$

b)

$$P(X=1)=\frac{0.5}{1!}\cdot e^{-0.5}=0.303265329.$$

c)

$$P(X=2)=\frac{0.5^2}{2!}\cdot e^{-0.5}=0.075816332.$$

T. 7. 5.

$$\lim_{\substack{n\to+\infty\\ n\cdot p=\lambda}}=\binom{n}{x}\cdot p^x\cdot q^{n-x}=\frac{\lambda^x}{x!}\cdot e^{-\lambda}.$$

Proof:

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

$$=\frac{\lambda^x}{x!}\cdot\frac{\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)\cdot\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}.$$

The statement of the theorem follows directly from:

$$\lim_{n\to+\infty}\left(1-\frac{\lambda}{n}\right)^n=e^{-\lambda},\qquad\qquad \lim_{n\to+\infty}\left(1-\frac{\lambda}{n}\right)^x=1.$$

R. 7. 8.
The binomial distribution can be approximated by the Poisson distribution for n "sufficiently" large ($n\to\infty$) while $n\cdot p=\lambda$ remaining constant. That is why the Poisson distribution is also known as the "distribution of rare events".
It will be recommended to use the following rule of thumb:

> "If
>
> $n\cdot p\le 10$ and $n\ge 1500p$,
>
> then the binomial distribution can be approximated by the Poisson distribution."

Ex 7. 5.
80 people work in a factory. The probability that one of them becomes sick in winter is estimated to be 0.05.
What is the probability that at least 11 people will become sick as a result of a flue epidemic?

Solution:
Denote the number of sick people by X . Using the binomial distribution with $n = 80$ and $p = 0.05$, we shall have to calculate

$$P(X \geq 11) = 1 - P(X < 11) = 1 - F(11) = 1 - \sum_{x=0}^{10} \binom{80}{x} \cdot 0.05^x \cdot (1 - 0.05)^{80-x}.$$

A rather „tedious“ job. However, this will not be necessary, since because of

$$n \cdot p = 80 \cdot 0.05 = 4 < 10 \quad \text{and} \quad n = 80 \geq 1500 \cdot 0.05 = 75 = 1500 \cdot p$$

the probability can be found by using the Poisson distribution:

$$\lambda = n \cdot p = 80 \cdot 0.05 = 4,$$
$$P(X \geq 11) = 1 - P(X < 11) = 1 - F(11) = 1 - \sum_{x=0}^{10} \frac{4^x}{x!} \cdot e^{-4} = 1 - 0.997158 = 0.002842$$

(Last revised: 30.06.09)