Some Crazy Distribution

Notebook:	INIAD Statistics		
Created:	10/19/2018 12:00 PM	**Updated:**	11/14/2018 10:45 PM
Author:	danganhvu1998		

- ## Hypergeometric distribution

  Hypergeometric distribution

  - There are two attributes, A and B. You have N materials that consist of M materials of attribute A, and (N-M) materials of attribute B. Now, suppose that you take out n materials from this population, and regard the number *x* of materials of attribute A (Ofcourse, then the number of attribute B is (*n*-*x*).).
    - $f(x) = {}_M C_x \cdot {}_{N-M} C_{n-x} / {}_N C_n$
  - Expected value: $E(X) = n(M/N)$
  - Variance: $V(X) = n\{M(N-M)/N^2\}\{(N-n)/(N-1)\}$

  - The ratio of attribute A is $p = M/N$. Then, as $N \to \infty$,
    - Expected value: $E(X) = np$
    - Variance: $V(X) = np(1-p)$

  - Hypergeometric distribution【Answer】

    - Suppose that there are 1000 fish in a lake, 200 of which has red marks. Now, if you catch 5 fish from this lake, find the probability that the number of marked fish is
    - (i)0 (ii) 1.

    $f(x) = {}_M C_x \cdot {}_{N-M} C_{n-x} / {}_N C_n$

    $N = 1000,\ M = 200,\ n = 5$

    $f(0) = {}_{800} C_5 / {}_{1000} C_5 = 0.32686$

    $f(1) = {}_{200} C_1 \cdot {}_{800} C_4 / {}_{1000} C_5 = 0.41063$

    Used for the resource survey.

  -

- ## Binomial distribution

Binomial distribution

- Suppose that a certain trial has two results (say, S and F, for instance). Each result occurs with probability p and (1-p).

- If you repeat such trials independently n times under the same condition, it is called as the *Bernoulli trial*.

➢ The probability of Shappensxtimes, and also Fhappens $(n-x)$ times :

➢ $f(x) = {}_nC_x\, p^x\, (1-p)^{n-x}$

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

Example of binomial distribution 【Answer】

- Suppose that there are a large amount of fish in a lake.

The fish with red marks account for the ratio of 0.2 among the whole fish. Now, if you catch 5 fish from this lake, find the probability that the number of marked fish is

- (i)0 (ii) 1.

$$f(x) = {}_nC_x\, p^x\, (1-p)^{n-x}$$
$$f(0) = (0.8)^5 = 0.32768$$
$$f(1) = 5\,(0.2)^1\,(0.8)^4 = 0.40960$$
$$f(2) = 10\,(0.2)^2\,(0.8)^3 = 0.20480$$
$$f(3) = 10\,(0.2)^3\,(0.8)^2 = 0.05120$$
$$f(4) = 5\,(0.2)^4\,(0.8)^1 = 0.00640$$
$$f(5) = (0.2)^5 = 0.00032$$

Close to the values of hypergeometric distribution.

- **Poisson distribution**

Poisson distribution

Consider the case $n \to \infty$ and $p \to 0$ so that $np \to \lambda$.

Then, for each x, the following statement holds

(Poisson's Law of Small Numbers).

$$ {}_nC_x \, p^x (1-p)^{n-x} \to e^{-\lambda} \cdot \lambda^x / x! $$

$$ f(x) = e^{-\lambda} \cdot \lambda^x / x! \quad (\lambda > 0,\ x = 0, 1, 2, \ldots) $$

: Poisson distribution denoted as $Po(\lambda)$.

$$ f(3) = e^{-2} \cdot 2^3 / 3! = 0.180447 $$

- $E(X) = \lambda \quad (\fallingdotseq np)$
- $V(X) = \lambda \quad (\fallingdotseq np(1-p))$

Poisson distribution depends only on λ.

Exercise

In a certain store, 4 customers arrive every one hour on average. Under the assumption that the arrival of customers is subject to the Poisson distribution, find the probability that 3 customers will arrive in one hour.

- **Answer: $f(3) = e^{-\lambda}.\lambda^x/x! = e^{-4}.4^3/3! = 0.195$**

- **Geometric distribution**

Geometric distribution

- Focus on the number of failure until the first success in the Bernoulli trial.
- Let the number of the trials be *x*.
  - S: once, F: $(x-1)$ times. The probability of such a situation is…
  - $f(x) = p(1-p)^{x-1}$
  - $f(x) = p \cdot q^{x-1} \quad (q = 1-p)$: Geometric distribution
- $E(X) = 1/p$
- $V(X) = q/p^2$

Formular At least 1 Success after X trial

$$F(X \le x) = 1 - (1-p)^x$$

- In Japan, large earth quakes occur with probability of 0.04 within a year. Then, by using the geometric distribution, find the expected value of their intervals. In addition, find the probability that it occurs within 10 years.
  - **Answer:**
    - Expected Value: 25 years
    - Probability Occurs Within 10 years: $1 - (1-0.04)^{10} = 0.335$

- **Exponential distribution**

Exponential distribution

- Let a certain event happen λ times on average within an unit time interval.
- Then, the probability that the interval of events, say, x, is subject to the exponential distribution.
  - ➢ $f(x) = \lambda e^{-\lambda x} \quad (x \geqq 0), \quad 0 \quad (x < 0)$
- The cumulative distribution is:
  - ➢ $F(x) = P(X \leqq x) = 1 - e^{-\lambda x} \quad (x \geqq 0)$
    $= 0 \quad (x < 0)$
- $E(X) = 1/\lambda$
- $V(X) = 1/\lambda^2$

Relationship between exponential and Poisson distributions

- Poisson distribution concerns the *numbers* of events within an unit time interval.
- Exponential distribution concerns the *interval* between the events.

- You observe the same events from different point of view.
  - ➢ Poisson distribution: numbers, Exponential: interval
  - ➢ The expected value of exponential distribution is $1/\lambda$.
  - ➢ Variance is $1/\lambda^2$.

【Exercise 6: Exponential distribution】【Answer】

- Let the interval of customer arrival at the gate subject to the exponential distribution, whose expected value is 30 [seconds] in a station. Then, find the probability that the interval is 1 minutes or more.

Let the unit time interval be 1 minutes. Then,

$\lambda = 2$[customers／minute], $x = 1$[minute].

$P(X>1) = 1-(1-e^{-2\times1}) = e^{-2} = 0.1353$

➢ $F(x) = P(X \leqq x) = 1 - e^{-\lambda x}$

- Let the interval of customer arrival at the gate subject to the exponential distribution, whose expected value is 30 [seconds] in a station. Then, find the probability that the interval is 1 minute or more, and 2 minutes or less.

Let the unit time interval be 1 minutes. Then,

$\lambda = 2$[customers／minute], $x = 1$[minute].

$P(1<X<2) = P(X<2) - P(X<1) = (1-e^{-2\times2}) - (1-e^{-2\times1})$

$= e^{-2} - e^{-4} = \underline{0.117}$