- **Binomial Probability Distribution**

Binomial Experiment:

Definition:

(1) Independently repeat n times.

(2) For each time, possible results: "Success" or "Failure".

(3) For each time: $P(\text{"Success"}) = P$, $P(\text{"Failure"}) = 1-P$

(4) We interested in: # "success" in n trials $\longrightarrow X$

(i.e, x)

$$P(X=x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, \quad x = 0, 1, \cdots, n$$

$X \sim B(n, p)$

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

- **Bernoulli Probability Distribution**

$n = 1$ (Try once) in Binomial Experiment.

Relationship between Bernoulli R.V. & Binomial R.V.

If $X_1, X_2, \cdots X_n \sim$ Bernoulli (P)

let $Y = \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$

Then $Y \sim$ Binomial (n, p). In other words, do Bernoulli Experiment for n times.

Multinomial Distribution: (If # possible outcomes > 2)

$$P(X_1 = k_1, X_2 = k_2, \cdots, X_m = k_m) = \frac{n!}{k_1! \cdot k_2! \cdots k_m!} \cdot P_1^{k_1} \cdot P_2^{k_2} \cdots P_m^{k_m}$$

- ## Hypergeometric Probability Distribution

"Success" : a

"Failure" : $N-a$

"bag" of N elements

Experiment : Randomly draw n elements without replacement.

repeat n times

Interested in : # "success" (in n trials)

$$P(X=x) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}$$

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

$$V(X) = \frac{a \cdot n \cdot (N-a)(N-n)}{N^2(N-1)}$$

- ## Negative Binomial Probability Distribution

Negative Binomial Experiment :

Definition :

Same with Binomial Exp.

(1) Independently repeat n times.

(2) For each time, possible results : "Success" or "Failure".

(3) For each time : P("Success") = P , P("Failure") = $1-P$

(4) We interested in : # trials for the K-th "success"
(i.e, Y)

$$P(Y=y) = \underbrace{\binom{y-1}{K-1} \cdot P^{K-1} \cdot (1-P)^{y-K}}_{\text{Trail 1 to } (y-1)} \cdot \underset{\text{K-th trial}}{P} \quad , \quad y = K, K+1, K+2, \dots$$

$$= \binom{y-1}{K-1} \cdot P^{K} \cdot (1-P)^{y-K}$$

$Y \sim NB(K, P)$

$$E(Y) = \frac{K}{P} \qquad V(Y) = \frac{K(1-P)}{P^2}$$

○ Geometric Probability Distribution

Special case in NB Prob. Dist.

Same experiment, but we interested in # trials to get first "success".
(i.e. Y)

$Y \sim G(P)$ $\quad P(Y=y) = P^1 \cdot (1-P)^{y-1}$, $y = 1, 2, \cdots$

$E(Y) = \frac{1}{P}$ $\quad V(Y) = \frac{1-P}{P^2}$

Summary for above:

Same Experiment, Different interested r.v. :

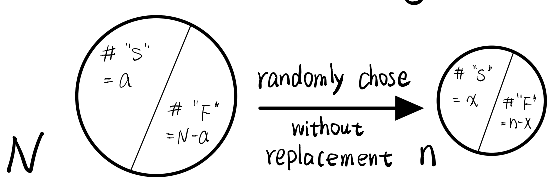
Bernoulli : 1 Trail

Binomial : n Trails # "success" = X

Geometric : Y trails to get 1-st "success"

Negative Binomial : Y trails to get K-th "success"

Hypergeometric :



"success" in n = X

Poission Probability Distribution

Poission Process : Events occur randomly & independently.

occurance per unit time / region = λ

For time period / total region t : # occurance = X

$$P(X=x) = \frac{(\lambda t)^x \cdot e^{-\lambda t}}{x!} \overset{\mu = \lambda t}{=\!=} \frac{\mu^x \cdot e^{-\mu}}{x!}, \quad \text{for } x = 0, 1, 2, \cdots$$

$X \sim \mathbb{P}(\mu)$

$E(X) = \mu$ $\quad V(X) = \mu.$

Tips :

① Hypergeometric $\xrightarrow{\approx}$ Binomial when $n \leq \frac{N}{10}$

② Binomial $\xrightarrow{\approx}$ Poission when n is large & P is small.

Appendix : Prove $E(X) = np$ when $X \sim B(n,p)$

$$P(X=x) = \frac{n!}{x!\,(n-x)!} \cdot p^x \cdot (1-p)^{n-x} \qquad E(X) = \sum x \cdot P(x)$$

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

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} \cdot p^x \cdot (1-p)^{n-x} = np \boxed{\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} \cdot (1-p)^{n-x}}$$

$$= 1 = \sum_{x-1=0}^{n} P(x-1)$$

$$(X-1) \sim B(n-1, p)$$

If it's not straight for you, let $x-1=t$, $n-1=m$

$$\boxed{\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} \cdot (1-p)^{n-x}} = \sum_{t=0}^{m} \frac{m!}{t!\,(m-t)!} \cdot p^t \cdot (1-p)^{m-t} = 1$$

Definition of f(x)

$x=1$ to n

$t=0$ to $n-1=m$ since $t=x-1$

OR. use $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k}$

$$\sum_{t=0}^{m} \frac{m!}{t!\,(m-t)!} \cdot p^t \cdot (1-p)^{m-t} = \sum_{t=0}^{m} \binom{m}{t} \cdot p^t \cdot (1-p)^{m-t} = [p + (1-p)]^m = 1$$

In all solutions below, an undefined RV X always corresponds to the one that was obviously intended.

1. According to a survey by the administrative Management Society, 40% of U.S. companies give employees 4 weeks of vacation after they have been with the company for 15 years. Find the probability that among 6 companies surveyed at random, the number that give employees 4 weeks of vacation after 15 years of employment is

   (a) anywhere from 2 to 5

   (b) fewer than 3

"S": give employees 4 weeks ...

n (# trails) = 6 $\qquad P(\text{"S"}) = 0.4, \quad P(\text{"F"}) = 0.6$

(a) X = # "S" in n $\qquad X \sim B(n=6, p=0.4)$

$$P(2 \leq X \leq 5) = 1 - P(X=0) - P(X=1) - P(X=6)$$
$$= 1 - \binom{6}{0} \cdot p^0 \cdot (1-p)^6 - \binom{6}{1} \cdot p^1 \cdot (1-p)^{6-1} - \binom{6}{6} \cdot p^6 \cdot (1-p)^{6-6}$$
$$= 1 - 1 \times 0.4^0 \times 0.6^6 - 6 \times 0.4 \times 0.6^5 - 1 \times 0.4^6 \times 0.6^0$$
$$\approx 0.7626$$

Book Table A1

OR $P(2 \leq X \leq 5) = P(X \leq 5) - P(X \leq 1) \approx 0.9959 - 0.233 = 0.7626.$

(b) $P(X < 3) = P(X \leq 2) \approx 0.5443$

2. One prominent physician claims that 70% of those with lung cancer are chain smokers. If his assertion is correct,

   (a) find the probability that 10 of such patients recently admitted to a hospital, fewer than half are chain smokers.

   (b) find the probability that 20 such patients recently admitted to a hospital, fewer than half are chain smokers.

"S": Patients with lung cancer are chain smokers.

X = # "S" in n patients.

(a) $n = 10.$ $\qquad X \sim B(n=10, p=0.7)$

$$P(X < 5) = P(X \leq 4) = 0.0473$$

(b) $n = 20$ $\qquad X \sim B(n=20, p=0.7)$

$$P(X < 10) = P(X \leq 9) = 0.0171$$

3. According to a study published by a group of University of Massachusetts sociologists, approximately 60% of the Valium users in the state of Massachusetts first took Valium for psychological problems. Find the probability that among the next 8 users from this state who are interviewed,

   (a) exactly 3 began taking Valium for psychological problems

   (b) at least 5 began taking Valium for psychological problems

"S": Take Valium for psychological problems.

$P(\text{"S"}) = 0.6 \qquad n=8 \qquad X = \#\text{"S" in 8 users.} \qquad X \sim B(n=8, p=0.6)$

(a) $P(X=3) = \binom{n}{x} p^x \cdot (1-p)^{n-x} = \binom{8}{3} \cdot 0.6^3 \cdot 0.4^5 \approx 0.124$

(b) $P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.4059 = 0.5941$

4. The probability that a patient recovers from a delicate heart operation is 0.9. What is the probability that exactly 5 of the next 7 patients having this operation survive?

"S": Patients recover from a delicate heart operation.

$P(\text{"S"}) = 0.9 \qquad n=7 \qquad X = \#\text{"S" in 7 patients.}$

$X \sim B(n=7, p=0.9)$

$P(X=5) = \binom{7}{5} \cdot 0.9^5 \cdot 0.1^2 \approx 0.124$

5. According to a genetics theory, a certain cross of guinea pigs will result in red, black and white offspring in ratio 8 : 4 : 4. Find the probability that among 8 offsprings, 5 will be red, 2 black and 1 white.

$P(\text{"r"}) : P(\text{"b"}) : P(\text{"w"}) = 8:4:4. \Rightarrow P(\text{"red"}) = \frac{8}{16} = \frac{1}{2} \quad \cdots$

$X_1 = \#\text{"red"} \qquad X_2 = \#\text{"black"} \qquad X_3 = \#\text{"white"} \qquad n=8$

$P(X_1=5, X_2=2, X_3=1) = \frac{8!}{5! \cdot 2! \cdot 1!} \times \left(\frac{8}{16}\right)^5 \times \left(\frac{4}{16}\right)^2 \times \left(\frac{4}{16}\right)^1 \approx 0.0820$

6. The probabilities are 0.4, 0.2, 0.3 and 0.1 respectively, that a delegate to a certain convention arrived by air, bus, automobile or train. What is the probability that among 9 delegates randomly selected at this convention, 3 arrived by air, 3 arrived by bus, 1 arrived by automobile and 2 arrived by train?

$P(\text{"air"}) = 0.4,\ P(\text{"bus"}) = 0.2,\ P(\text{"automobile"}) = 0.3,\ P(\text{"train"}) = 0.1$

$X_1 = \#$ delegates arrived by air, $X_2 = \#$ delegates arrived by bus

$X_3 = \#$ delegates arrived by automobile, $X_4 = \cdots$ by train

$$P(X_1=3,\ X_2=3,\ X_3=1,\ X_4=2)$$
$$= \frac{9!}{3!\cdot 3!\cdot 1!\cdot 2!} \times 0.4^3 \times 0.2^3 \times 0.3^1 \times 0.1^2 = 0.00774$$

7. From a lot of 10 missiles, 4 are selected at random and fired. If the lot contains 3 defective missiled that will not fire, what is the probability that

(a) all 4 will fire?

(b) at most 2 will fire?

$N=10,\quad n=4,\quad a=7,\quad N-a=3$

"S": will fire "F"

$$P(X=x) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}},\ x=1,2,3,4$$

$X = \#$ "S" in $a=7$ missiles. "Hypergeometric"

$$= \frac{\binom{7}{x}\binom{3}{4-x}}{\binom{10}{4}}$$

(a) $$P(X=4) = \frac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \frac{1}{6}$$

(b) $$P(X\le 2) = P(X=1)+P(X=2)$$ OR $$= 1-P(X=4)-P(X=3)$$

$$= \frac{\binom{7}{1}\binom{3}{3}}{\binom{10}{4}} + \frac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}} \qquad = 1 - \frac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} - \frac{\binom{7}{3}\binom{3}{1}}{\binom{10}{4}}$$

$$= \frac{1}{3} \qquad = \frac{1}{3}$$

8. What is the probability that a waitress will refuse to serve alcoholic beverages to only 2 minors if she randomly checks the IDs of 5 among 9 students, 4 of whom are minors?

"Hypergeometric" "S": minors

$N=9\quad a=4\quad n=5$

$$P(X=x) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}$$

$X = \#$ minors refused to serve.

$$P(X=2) = \frac{\binom{4}{2}\binom{5}{3}}{\binom{9}{5}} = \frac{10}{21}$$

9. It is estimated that 4000 of the 10000 voting residents of a town are against a new sales tax. If 15 eligible voters are selected at random and asked their opinion, what is the probability that at most 7 favor the new tax?

"Binomial." "S" : favor the new tax

$P(\text{"S"}) = \frac{6000}{10,000} = 0.6$ $n=15$

X = # favors in $n=15$ voters $X \sim B(n=15, p=0.6)$

$P(X \leq 7) = 0.2131$

Also, you can view it as a hypergeometric case :

$N = 10000$, $a = 6000$, $n = 15$

But you will find it's difficult to calculate and the final result is close to using Binomial Distribution.

When $n \leq \frac{N}{10}$, we can use Binomial Distribution to approximate Hypergeometric Distribution.

10. The probability that a person living in a certain city owns a dog is estimated to be 0.3. Find the probability that

a the sixth person randomly interviewed in that city is the first one to own a dog.

b the tenth person randomly interviewed in that city is the fifth one to own a dog.

(a) "Geometric" "S": own a dog.

$P(\text{"S"}) = 0.3$ Y = # interviews to find the first "S"

$$\boxed{P(Y=y) = p^1 \cdot (1-p)^{y-1}, \quad y = 1, 2, \cdots}$$

$P(Y=6) = 0.3^1 \cdot 0.7^{6-1} = 0.0504$

(b) "Negative Binomial" Y = # interviews to find k-th "S"

$$\boxed{P(Y=y) = \binom{y-1}{k-1} \cdot p^k \cdot (1-p)^{y-k}} \quad Y \sim NB(k=5, p=0.3)$$

$$P(Y=10) = \binom{10-1}{5-1} \cdot 0.3^5 \cdot 0.7^{10-5} = \binom{9}{4} \cdot 0.3^5 \cdot 0.7^5 = 0.0515$$

11. A scientist inoculates mice, one at a time, with a disease germ until he finds 2 that have contracted the disease. If the probability of contracting the disease is 1/6, what is the probability that 8 mice are required?

"Negative Binomial" $Y = \#$ mice contracted the disease.

$Y \sim NB(K=2, P=\frac{1}{6})$

$P(Y=8) = \binom{8-1}{2-1} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^{8-2} = \binom{7}{1}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^6 = 0.0651$

12. The average number of field mice per acre in a 5-acre field is estimated to be 12. Assuming a Poisson-distribution, find the probability that fewer than 7 field mice are found

a on a given acre;

b on 2 of the next 3 acres inspected.

(a) "Poission" $\mu = 12$ $X = \#$ field mice found on a 5-acre field.

$X \sim P(\mu=12)$

$P(X<7) = P(X \leq 6) = 0.0458$

(b) "Binomial" "S": <7 field mice found on a 5-acre field.

$n=3$, $P(\text{"S"}) = 0.0458$

$X = \#$ of next acres inspected have <7 field mice

$X \sim B(n=3, P=0.0458)$

$P(X=2) = \binom{3}{2} \cdot (0.0458)^2 \cdot (0.9542)^1 = 0.0060$

13. Assume that the probability that a randomly infected person will die from COVID-19, is 0.002. Find the probability that fewer than 5 of the next 2000 infected people will die from it.

"Binomial" "S": Infected person die from COVID-19

$n = 2000$, $X = \#$ person die from COVID-19

$X \sim B(n=2000, p=0.002)$

$P(X<5) = P(X \leq 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$

When n is large an p is small, we can use "Poission" to replace "Binomial".

$\mu = np = 2000 \times 0.002 = 4$

$X \sim B(n=2000, p=0.002) \approx \mathbb{P}(\mu=4)$

$P(X<5) = P(X \leq 4) = 0.6288$

14. Suppose that, on average, 1 person in 1000 makes a numerical error in preparing his or her income tax return. If 10000 forms are selected at random and examined, find the probability that 6, 7 or 8 of the forms contain an error.

"S": there is an error in tax return.

$P(\text{"S"}) = \frac{1}{1000} = 0.001$ $\quad n = 10^4$ $\quad X = $ # forms contain an error.

$X \sim B(10000, 0.001) \approx \mathbb{P}(10)$ $\quad \mu = n \cdot p = 10^4 \times 0.001 = 10$

$X \sim \mathbb{P}(10)$

$P(6 \leq X \leq 8) = P(X \leq 8) - P(X \leq 5) = 0.3328 - 0.0671 = 0.2657$