~~05~~ 06.1

BINOMIAL RANDOM VARIABLES

KHAN ACADEMY – Type of Distribution
Statistics And probability – Random variables

Binomial Variables

- It is made up of finite number and independent trials.
- Each trial can be classified as either success or failure.
- Fixed number of trials
- Probability of Success in each trial is constant.

} Conditions for Binomial Variable

Binomial Variable example →

X = Number of heads after 10 flip of my coin.

$P(H) = 0.5$ $P(T) = 0.5$, Trials = 10. Probability will not change in any trial
They are independent

Non Binomial Variable example →

Y = Number of kings after taking 2 cards from standard deck without replace

$P(\text{King on } 1^{st} \text{ trial}) = \frac{4}{52}$, Second picks

if we get king — $\frac{3}{51}$
if we do not get king — $\frac{4}{\cancel{52}\,51}$

Probability is changing
So Non binomial Variable.
They are not independent

Suppose, we say pick a card with replacement.

$P(\text{King on } 1^{st} \text{ trial}) = \frac{4}{52}$; $P(\text{King on } 2^{nd} \text{ trial}) = \frac{4}{52}$. Probability is same, independent, finite trial. So Binomial Variable.

A manager oversees 11 female employees and 9 male employees. They need to pick 3 of these employee to go on a business trip, so the manager places all 20 names in a hat and choosen at random. Let X = the number of female employees choosen.

Is X a binomial variable? Why or why not?

$P(\text{Female on first pick}) = \frac{11}{20}$, Second picks

if we got female — $\frac{10}{19}$
if we not got female — $\frac{11}{19}$

Probability is changing.
So not binomial random variable
They are not Independent.

Few other examples –

i) In a game involving a standard deck of 52 playing cards, an individual draws 7 cards without replacement. Let Y = Number of aces drawn. NOT BINOMIAL VARIABLE (not independent trial)

ii) 60% of a certain species of tomato live after transplanting from pot to garden. Eli transplants 16 of those tomato plants. Assume that the plants live independently of each other. Let T = Number of tomato plants that live. BINOMIAL VARIABLE

iii) In a game of luck, a turn consist of a player continuing to roll a pair of six sided dice until they roll a double (two of the same face value). Let X = the number of rolls in one turn. NOT BINOMIAL VARIBLE (No fixed number of trials)

03.8
06.2

Eg - if Probability (score) = 70% or 0.7 and Probability (Miss) = 30% or 0.3

P(Exactly 2 scores in 6 attempts)?

Score → S   Miss → M

First scenario → S S M M M M $= (0.7)^2 (0.3)^4$
0.7 0.7 0.3 0.3 0.3 0.3

Second scenario → M S S M M M M $= (0.7)^2 (0.3)^4$
~~0.7~~ 0.7 0.7 0.3 0.3 0.3 0.3

All possible scenario → Total Number of events = 6, Possible events = 2
(Total position)   (Selected position)

$= {}^6C_2 = \frac{6!}{4!2!}$

Total scenario $= {}^6C_2 \times [(0.7)^2 \times (0.3)^4] = 0.05 = 5\%$

P(Exactly 2 scores in 6 attempts) = 5%

~~Gener~~ Generalizing K scores in n attempts →

Suppose $P(\text{Score}) = P$ and $P(\text{Not Score}) = 1-P$

$P(\text{Exactly } K \text{ scores in } n \text{ attempts}) = \binom{n}{K} p^K (1-p)^{n-K}$

Binomial distribution formula.

or

$$\boxed{P(\text{Exactly } K \text{ score in } n \text{ attempt}) = ({}^nC_K)\, p^K (1-p)^{n-K}}$$

Binomial Probability Distribution

X = Number of Score when take 6 throws.

Prob (Score) = 0.7
Prob (Miss) = 0.3

zero score

$P(x=0) = ({}^6C_0) \times (0.7)^0 (0.3)^6 \cong 0.001 = 0.1\%$

$P(x=1) = ({}^6C_1) (0.7)^1 (0.3)^5 \cong 0.01 = 1.0\%$

$P(x=2) = ({}^6C_2) (0.7)^2 (0.3)^4 \cong 0.06 = 6.0\%$

$P(x=3) = ({}^6C_3) (0.7)^3 (0.3)^3 \cong 0.185 = 18.5\%$

$P(x=4) = ({}^6C_4) (0.7)^4 (0.3)^2 = 0.324 = 32.4\%$

$P(x=5) = ({}^6C_5) (0.7)^5 (0.3)^1 = 0.303 = 30.3\%$

$P(x=6) = ({}^6C_6) (0.7)^6 (0.3)^0 = 0.118 = 11.8\%$

Here is some trend

No trend.

There is no such trend in output.

Only trend is → ${}^6C_0 = \frac{6!}{0!6!} = 1$.

${}^6C_1 = \frac{6!}{5!\,1!} = \frac{6}{1} = 6$

${}^6C_2 = \frac{6!}{4!\,2!} = \frac{6\times5}{2} = 15$

${}^6C_3 = \frac{6!}{3!\,3!} = \frac{6\times5\times4}{3\times2} = 20$.

${}^6C_4 = \frac{6!}{2!\,4!} = \frac{6\times5}{2} = 15$

${}^6C_5 = \frac{6!}{1!\,5!} = 6$

${}^6C_6 = \frac{6!}{0!\,6!} = 1$

There is no symmetry because of uneven probability (0.7, 0.3).

If we have same probability (0.5, 0.5) we will get symmetry. [Next example]

06.3 X = Number of Heads, flipping coin 5 times

Possible outcomes of 5 flips $= 2 \times 2 \times 2 \times 2 \times 2 = 2^5 = 32$.

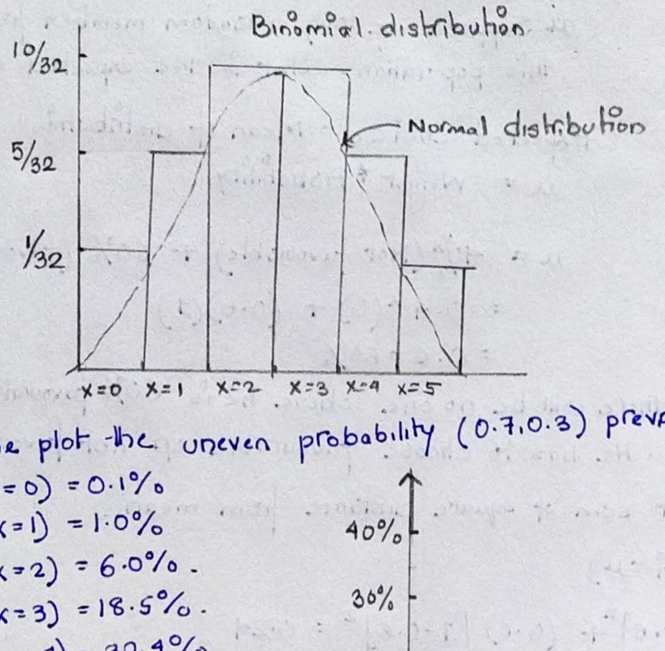
$P(X=0)$ = Zero head = all tail = TTTTT = $\frac{\text{Possible coin outcome}}{\text{Total possibility}} = \frac{1}{32}$.

Otherway is, n = no of flip = 5, K = no of head = 0.
to find possible outcome $^nC_K = {}^5C_0 = \frac{5!}{0!\,5!} = 1$. so $P(X=0) = \frac{1}{32}$.

$P(X=1) = \frac{^5C_1}{32} = \frac{5}{32}$ $\quad P(X=2) = \frac{^5C_2}{32} = \frac{10}{32}$ $\quad P(X=3) = \frac{^5C_3}{32} = \frac{10}{32}$

$P(X=4) = \frac{^5C_4}{32} = \frac{5}{32}$ $\quad P(X=5) = \frac{^5C_5}{32} = \frac{1}{32}$.

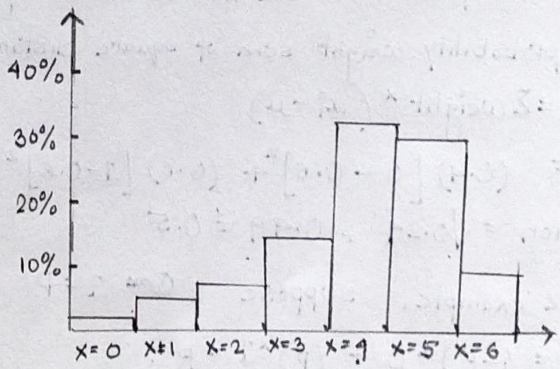
so when we observe we see the symmetry $\frac{1}{32}, \frac{5}{32}, \frac{10}{32}, \frac{10}{32}, \frac{5}{32}, \frac{1}{32}$.
It is because we have equal probability of happening.



Normal distribution is continuous case.
Binomial distribution is categorical case of Normal distribution.

Even we plot the uneven probability (0.7, 0.3) prevpus to previous example.

$P(X=0) = 0.1\%$
$P(X=1) = 1.0\%$
$P(X=2) = 6.0\%$.
$P(X=3) = 18.5\%$.
$P(X=4) = 32.4\%$
$P(X=5) = 30.3\%$.
$P(X=6) = 11.8\%$



Eg - I have a 0.35 probability of making a ~~throw~~ free throw. What is the probability of making 4 out of 7 free throws?

P (free throw) = 0.35 $\quad$ P (not free throw) = 0.65.
n = 7 free throws total, K = 4 free throw $\Rightarrow {}^7C_4$.

Free	Free	Free	Free	Not Free	Not Free	Not Free	← Combination 1
0.35	0.35	0.35	0.35	0.65	0.65	0.65	$= (0.35)^4 (0.65)^3$.

For all combination $= (^7C_4)(0.35)^4(0.65)^3$ Ans
$= 0.14$.

Calculator step
binompdf(7, 0.35, 4)
↑ n ↑ p ↑ success

What is the probability of making less than 5 free throws?

$P(X \leq 5) = P(X=4)$ ~~and~~ or $P(X=3)$ or $P(X=2)$ or $P(X=1)$ or $P(X=0)$

$$= \left[{}^7C_4 (0.35)^4 (0.65)^3\right] + \left[{}^7C_3 (0.35)^3 (0.65)^4\right] + \left[{}^7C_2 (0.35)^2 (0.65)^5\right]$$

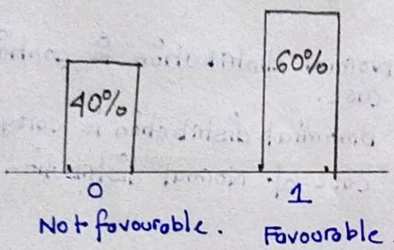
$$\left[{}^7C_1 (0.35)^1 (0.65)^6\right] + \left[{}^7C_0 (0.35)^0 (0.65)^7\right]$$

Mean and Variance of Bernoulli Distribution example

Suppose we ~~can~~ took survey of all the population who supported Modi for Prime Minister. You can be favourable or unfavourable.

From survey we got ~~unfav~~ unfavourable = 40% . favourable = 60% .
(Consider this as 0) (Consider this as 1)

~~A~~ Expected value of this distribution = Mean = μ.



~~A~~ Suppose we took a random member from this population. What is the expected choice?.

Expected choice = Mean of distribution

μ = Weight * Probability

suppose Not favourable = 0
Favourable = 1

μ = 40% * (Not favourable) + 60% (favourable).

$= 0.4 \times (0) + (0.6)(1)$

$= 0.6 = 60\%$

So this is impossible because there will be no one where he is 60% favourable and remaing 40% not favourable. He has to choose favourable or not favourable.

For variance, probability weight sum of square distance from mean.

$$\sigma^2 = \Sigma \text{weight} \times (x_i - \mu)$$

$$\sigma^2 = (0.4)\left[0 - 0.6\right]^2 + (0.6)\left[1 - 0.6\right]^2 = 0.24$$

Standard deviation $= \sqrt{0.24} = 0.49 \cong 0.5$

Generalizing above example, suppose ~~o~~ $1 = P$ then $0 = 1 - P$. because sum of $P = 1$.

Mean $= \mu = (1-P) \times 0 + (P) \times 1 = P$.

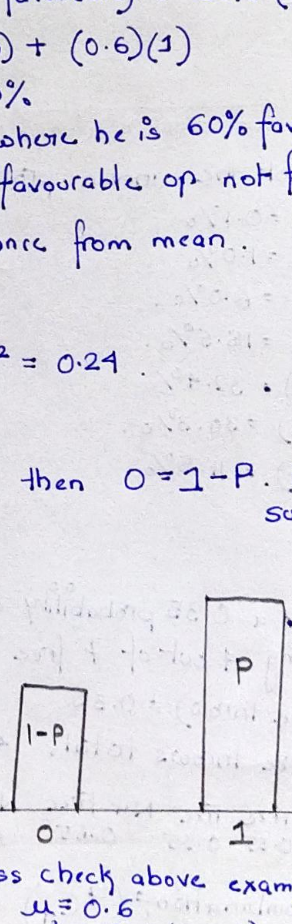
$$\text{Variance} = \sigma^2 = (1-P)(0-P)^2 + P(1-P)^2$$

$$= (1-P)(-P)^2 + P(1-P)^2$$

$$= (1-P)(P^2) + P(1 - 2P + P^2)$$

$$= P^2 - \cancel{P^3} + P - 2P^2 + \cancel{P^3}$$

$$= P - 2P$$



Variance $= \sigma^2 = P(1-P)$

Standard Deviation $= \sigma = \sqrt{P(1-P)}$

we can cross check above example
$P = 0.6$; $\mu = 0.6$
$\sigma^2 = 0.6(1-0.6) = 0.6 \times 0.4 = 0.24$
$\sigma = \sqrt{0.24} = 0.5$

06.5 Expected value of a binomial variable.

X = Number of success after n trials where P(success) for each trial is p.

Expected Value (X) = n p. (Mean of distribution)

Suppose a trial is a throw, ~~so~~ P(Success) = 30% = 0.3.

Suppose n = 10, then. Expected Value (X) = (10) * (0.3) = 3.

which make sense, out of 10 trials we will get 3 success as per probability.

~~Expect~~ Variance (X) = np(1-p), variance = 10 (0.3) (0.7) = 2.1.

Standard deviation (X) = $\sqrt{np(1-p)}$

Eg - A company produces processing chips for cell phones. At one of its large factories 2% of the chip produced are defective in some way. A quality check involves randomly selecting and testing 500 chips. What are the mean and standard deviation of the number of defective processing chips in samples?

X = Number of defective in 500 chips.

It is binomial because n = 500, one chip is not depends on other chip. Independent event.

P(defective chip) = 0.02.

Mean (X) = Expected Value (X) = E (X) = n.p = 500 * (0.02) = 10

Standard Deviation (X) = $\sigma(X) = \sqrt{500 (0.02) (0.98)} = \sqrt{9.8} = 3.13$

Geometric Random Variable

X = Number of 6's after 12 rolls of fair die. Y = Number of rolls until get 6 on fair die.

X is a binomial random variable.

	X	Y
- trial outcome success / failure.	✓	✓
- trial results independent.	✓	✓
- Fixed number of trials.	✓	X → We do not know after how much roll we will get 6.
- Same probability of each event.	✓	✓

So, if a problem meet most of the condition of binomial random variable excluding the condition on fixed number of trials then that is Geometric Random variable.

In Binomial - We check how many success in finite number of trials.

In Geometric - We ~~also~~ check how many trials until we get success.

Jeremiah makes 25% of the three point shots he attempts. For a warm up, Jeremiah like to shoot three point shoots until he successfully make one. Let M be the number of shots it takes Jeremiah to successfully make his first three point shot. Assume that the results of each shot are independent. Find the probability that Jeremiah's first successful shot occurs on his 3^{rd} attempt. You may round your answer to nearest hundredth.

- Seems like a Geometric Random Variable,

$P(\text{Shot} = 3^{rd} \text{ attempt}) = (\text{Miss first shot})$ and $(\text{Miss second shot})$ and (3 pointer shot)

$= 3/4 \times 3/4 \times 1/4$

$P(\text{Miss shot}) = 1 - 25\% = 3/4$

$= 9/64 \cong 0.14$

$P(3 \text{ point shot}) = 25\% = 1/2$

14% chance first successful at 3^{rd} attempt

Emelia register vehicles for Department of Transportation. Sports utility vehicle (SUV) makes 12% of the vehicle she register. Let V be the number of vehicles Emelia registers in a day until she first register an SUV. Assume the type of vehicle is independent.

Find the probability that Emila register more than 4 vehicle she registers an SUV.

→ V = number of vehicles Emila registers in a day until she first registers an SUV

Suppose first customer books SUV, V = 0.

Suppose first one buy any other but second one book SUV, V = 1.

This is geometric random variable.

$P(V > 4) = P(\text{Not SUV}_{\text{1st booking}})$ and $(\text{Not SUV}_{\text{2nd booking}})$ and $(\text{Not SUV}_{\text{3rd booking}})$ and $(\text{Not SUV}_{\text{4th booking}})$

$= P(\text{first 4 cars not SUV}) = (0.88)^4 = 0.60 = 60\%.$

Lilyana runs a cake decorating business, for which 10% of her orders come over the telephone. Let C be the number of cake orders Lilyana receives in a month until she gets an order over the telephone. Assume the method of placing each cake order is independent.

Find the probability that it takes fewer than 5 orders for Lilyana to get her first telephone order of the month?

$P(C < 5) = P(\text{first order is telephone}) + P\left(\begin{array}{c}\text{first} \neq \text{telephon} \\ \text{second} = \text{telephone}\end{array}\right) + P\left(\begin{array}{c}\text{first, second} \neq t \\ \text{third} = \text{telephon}\end{array}\right) + P\left(\begin{array}{c}1,2,3 \neq t \\ 4 = t\end{array}\right)$

$= 0.1 + (0.9)(0.1) + (0.9)^2(0.1) + (0.9)^3(0.1) = 0.34$

Another way, $1 - P\left(\begin{array}{c}\text{No telephone} \\ \text{order in first 4}\end{array}\right) = 1 - (0.9)^4 = 0.34$