

IIT Madras
ONLINE DEGREE

Statistics for Data Science - 1
Professor. Usha Mohan
Department of Management Studies
Indian Institute of Technology, Madras
Lecture No.10.2

Binomial Distribution- Independent and Identically Distributed Bernoulli Trials

(Refer Slide Time: 00:01)

Statistics for Data Science -1
 ↳ The Binomial Experiment
 ↳ Independent and identically distributed Bernoulli trials



$n=3$ independent trials

Prob of success = p

- ▶ Let $n = 3$ independent Bernoulli trials.
- ▶ Let p is probability of success.

Outcome $P \cdot (SSS)$
 ↓
 Prob of 1st trial success and Prob 2nd trial success and Prob 3rd trial success
 ↓
 $p \times p \times p = p^3$
 $P(SSS) = p \times p \times p = p^3$
 $P(Sff) = p \times (1-p) \times (1-p) = p(1-p)^2$
 $P(ffs) = (1-p) \times (1-p) \times p = (1-p)^2 p$



Statistics for Data Science -1
 ↳ The Binomial Experiment
 ↳ Independent and identically distributed Bernoulli trials



$n=3$ independent trials

- ▶ Let $n = 3$ independent Bernoulli trials.
- ▶ Let p is probability of success.
- ▶ The probabilities of outcomes of the independent trials are

S.No	Outcome	Probabilities
1	3 (s,s,s)	$p \times p \times p$ Independence
2	2 (s,s,f)	$p \times p \times (1-p)$
3	2 (s,f,s)	$p \times (1-p) \times p$
4	1 (s,f,f)	$p \times (1-p) \times (1-p)$
5	2 (f,s,s)	$(1-p) \times p \times p$
6	1 (f,s,f)	$(1-p) \times p \times (1-p)$
7	1 (f,f,s)	$(1-p) \times (1-p) \times p$
8	0 (f,f,f)	$(1-p) \times (1-p) \times (1-p)$



You can see that I can enumerate or I can find out the probability of each of these outcomes and to compute this probability, I use the fact that the independence of the trials, so I list out the probability of each of these outcomes. So, you can see that each of these outcomes I have listed out the probabilities.

But, what is a binomial random variable? The binomial random variable was again a model for counts what did I count I counted the total number of successes in the outcomes. Now, if I am going to map it I see in this outcome I have 3 successes. In this outcome I have 2, in this I have 2, in this I have 1, in this I have 2, in this I have 1, in this I have 1, in this I have no success. So, if what this counts is the number of successes in these outcomes.


(Refer Slide Time: 01:08)

Statistics for Data Science -1
 ↳ The Binomial Experiment
 ↳ Independent and identically distributed Bernoulli trials

$n=3$ independent trials, X = number of successes

- ▶ Let $n = 3$ independent Bernoulli trials.
- ▶ Let p is probability of success.
- ▶ Let X = number of successes in 3 independent trials.
- ▶ The probability distribution of X

X	0	1	2	3
$P(X = i)$	$(1 - p)^3$	$3 \times p \times (1 - p)^2$	$3 \times p^2 \times (1 - p)$	p^3



So, I can map this X , or the random variable to be the number of successes so I have a 3, 2, 2, 1, 2, 1, 1, 0 the probability is remain the same. So, I now I can see that the chance of X , where X is the number of successes taking the value 3, 3 appears only 1 with just p^3 . Now, the chance again they are independent they are identically distributed the chance of X taking the value 2 is equal to this plus this plus this. So, how many times does that occur? 1, 2, 3 so 1, 2, 3 so it would be $3 \times p^2 \times (1 - p)$.

Similarly, I see that the chance of X taking the value 1 equal to $p \times (1 - p)^2$ and that appears 3 times and the chance of X taking the value 0 or no success is $(1 - p)^3$. So, I have now what is what I refer to as the probability distribution of the number of successes in $n = 3$ independent Bernoulli trials with a p which is a probability of success, $(X = i) = (1 - p)^3, 3 \times p \times (1 - p)^2, 3 \times p^2 \times (1 - p), p^3$.

(Refer Slide Time: 3:06)

Statistics for Data Science -1
 The Binomial Experiment
 Independent and identically distributed Bernoulli trials


$n=3$ independent trials, X = number of successes

- Let $n = 3$ independent Bernoulli trials.
- Let p is probability of success.
- X = number of successes in 3 independent trials.
- The probabilities of outcomes of the independent trials are

Handwritten notes: $p(x=3) = p^3$, $p(x=2) = 3p^2(1-p)$, $p(x=1) = 3p(1-p)^2$, $p(x=0) = (1-p)^3$

S.No	Outcome	Number of successes	Probabilities
1	(s,s,s)	3	$p \times p \times p \rightarrow p^3$
2	(s,s,f)	2	$p \times p \times (1-p) \checkmark$
3	(s,f,s)	2	$p \times (1-p) \times p \checkmark$
4	(s,f,f)	1	$p \times (1-p) \times (1-p)$
5	(f,s,s)	2	$(1-p) \times p \times p \checkmark$
6	(f,s,f)	1	$(1-p) \times p \times (1-p)$
7	(f,f,s)	1	$(1-p) \times (1-p) \times p$
8	(f,f,f)	0	$(1-p) \times (1-p) \times (1-p)$

Handwritten notes on the left: 3 (with arrow to row 1), $\binom{3}{2} = \frac{3!}{2!1!} = 3$, $\binom{3}{1} = 3$



Now, what I want you to notice is, how did I get this 3, now when I look at the number of successes is equal to 2. Now, what I mean by number of successes equal to 2 is the way I can have 3 trials, I need 2 of these trials to be successful. So, I am in other words I am choosing 2 positions out of the 3 positions, where I can have a success and I know I can do this in 3C_2 ways which is $\frac{3!}{2!1!}$ which is 3 and that reflects in this 3. Similarly, choosing 1 success is also can be done in 3C_1 way, which is again the same as 3 and that reflects here.

(Refer Slide Time: 3:54)


Statistics for Data Science -1
 The Binomial Experiment
 Independent and identically distributed Bernoulli trials

n independent trials, X = number of successes

Handwritten notes: $1 \ 2 \ 3 \dots n$, $S \ P$, $F \ 1-p$

Handwritten notes: $(s, s, s, \dots, s) \rightarrow p \times p \times \dots \times p = p^n$, $(s, s, \dots, s, f) \rightarrow p \times p \times \dots \times p \times (1-p) = p^{n-1}(1-p)$

Handwritten notes: $(f, f, \dots, f) \rightarrow (1-p) \times (1-p) \times \dots \times (1-p) = (1-p)^n$



So, in other words if I am going to extend the notion of 3 independent trials to n independent trials, again I have trial 1, trial 2, trial 3 up to n independent trial each 1 of them results in a success or failure the probability of getting a success is p and failure is $(1 - p)$. So, the possible outcome is I could have all the n successes, the probability with which it happens is $p \times p \times p \dots$ which is p^n .

The other extreme is all failure, which can happen with probability $(1 - p) \times (1 - p) \times (1 - p) \dots$ which is $(1 - p)^n$. I could have my first $(n - 1)$ successes and the last failure this would happen with $p \times p \times p \dots \times (1 - p)$ which is $p^{n-1} \times (1 - p)$. In other words, I have to choose $(n - 1)$ positions from n positions and that can be done in n ways or ${}^nC_{n-1}$ way.

(Refer Slide Time: 05:16)

Statistics for Data Science -1
 The Binomial Experiment
 Independent and identically distributed Bernoulli trials

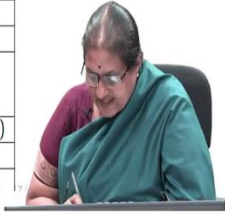
$P(X=2) = {}^nC_2 p^2 (1-p)^{n-2}$

n independent trials, X = number of successes

- ▶ Let there be n independent Bernoulli trials.
- ▶ Let p is probability of success.
- ▶ X = number of successes in n independent trials.
- ▶ The probabilities of outcomes of the independent trials are

S.No	Outcome	Number of successes	Probabilities
1	(s,s,...,s)	n	$p \times p \times \dots \times p$
2	(s,s,...,f)	$n-1$	$p \times p \times \dots \times (1-p)$
3	(s,...,f,s)	$n-1$	$p \times \dots \times p \times (1-p) \times p$
\vdots	\vdots	\vdots	\vdots
2^{n-2}	(f,...,s,f)	1	$(1-p) \times (1-p) \dots \times p \times (1-p)$
2^{n-1}	(f,f,...,s)	1	$(1-p) \times (1-p) \dots \times p$
2^n	(f,f,...,f)	0	$(1-p) \times (1-p) \dots \times (1-p)$

$X: 0, 1, 2, \dots, n$
 $P(X=0) = (1-p)^n, P(X=1) = {}^nC_1 p (1-p)^{n-1}, \dots, P(X=n) = p^n$



So, I can continue it in this form and I can show that the way I have obtain for each of these outcomes and I have n independent trials. So, I have my variable X , which now takes the value 0, 1, 2 up to n . Now it takes the value 0 with probability $(1 - p)^n$, it takes the value n or all of them or successes with probability p^n . It takes the value 1 it tells me that if I take the value 1 all I have $(n - 1)$ failures and 1 success this success can occur in any 1 of these points I can choose it in 1 way, I have $p \times (1 - p)^{n-1}$.

Similarly, $P(X = i)$, I have i successive p^i , $(n - i)$ failures $(n - i)$ and I can choose this i successes as nC_i way.

(Refer Slide Time: 6:27)



n independent trials, X = number of successes

- ▶ Consider any outcome that results in a total of i successes.
 - ▶ This outcome will have a total of i successes and $(n - i)$ failures.
 - ▶ Probability of i success and $(n - i)$ failures = $p^i \times (1 - p)^{(n-i)}$
- ▶ There number of different outcomes that result in i successes and $(n - i)$ failures = $\binom{n}{i}$
- ▶ The probability of i successes in n trials is given by

$$P(X = i) = \binom{n}{i} \times p^i \times (1 - p)^{(n-i)} \quad i = 0, 1, 2, \dots, n$$



So, this gives me the general formula for the probability of i successes and $(n - i)$ failure to be $p^i \times (1 - p)^{n-i}$ and this appears in nC_i outcomes. Hence, the probability of i successes in n trials is ${}^nC_i \times p^i \times (1 - p)^{n-i}$. This is the probability mass function of a binomial random variable, i can take any value 0, 1, 2 up to n .

(Refer Slide Time: 07:15)



Section summary

- ▶ Independent and identically distributed distribution
- ▶ n independent trials
- ▶ Probability of i successes and $(n - i)$ failures in n independent trials.

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$$



So, this is how we can see that independent and identically distributed Bernoulli random variables, how the binomial distribution arises naturally out of these n independent trials, where probability of i successes and $(n - i)$ failures is given by the expression $P(X = i) = {}^nC_i \times p^i \times (1 - p)^{n-i}$. this is the key derivation of the binomial probability mass function.

