Discrete probability distributions

Tan Do

Vietnamese-German University

Lecture 12

In this lecture

- The binomial distribution
- The geometric distribution

Bernoulli random variables

Let X be a random variable that takes just 2 (distinct) values.

We can use X to model the outcome of a coin toss, whether a valve is open or shut, whether an item is defective or not, or any other process that has only 2 outcomes. We usually use 0 and 1 to denote the outcomes.

Let p be the probability that the outcome is 1 $(0 \leq p \leq 1)$.
Then simple calculations give

$$E(X) = p$$

and

$$\mathrm{Var}(X) = p(1 - p).$$

Such a variable X is called a Bernoulli random variable with parameter p.
An experiment that has only 2 outcomes is also referred to as a Bernoulli trial.

Definition of the binomial distribution

Let $X_1, X_2, \ldots, X_n$ be independent Bernoulli random variables whose probability of getting a 1 is p. Then

$$X = X_1 + X_2 + \ldots + X_n$$

is called a binomial random variable with parameters n and p, which is denoted by

$$X \sim B(n, p).$$

Note that X takes values $0, 1, 2, \ldots, n$ and counts the number of Bernoulli random variables that take value 1.

We usually think of the value 1 as a success, and so in other words, X counts the *number of successes* in a process consisting of n consecutive trials.

The probability mass function of a $B(n,p)$ random variable is

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

for $x=0,1,2,\ldots,n$. In particular, let $X\sim B(4,p)$. Then

$$P(X=3)=P(0111)+P(1011)+P(1101)+P(1110)$$

$$=(1-p)ppp+p(1-p)pp+pp(1-p)p+ppp(1-p)=4p^3(1-p).$$

Outcome of 4 Bernoulli trials	Probability	X = number of successes
0000	$(1-p)(1-p)(1-p)(1-p)$	0
0001	$(1-p)(1-p)(1-p)p$	1
0010	$(1-p)(1-p)p(1-p)$	1
0011	$(1-p)(1-p)pp$	2
0100	$(1-p)p(1-p)(1-p)$	1
0101	$(1-p)p(1-p)p$	2
0110	$(1-p)pp(1-p)$	2
0111	$(1-p)ppp$	3
1000	$p(1-p)(1-p)(1-p)$	1
1001	$p(1-p)(1-p)p$	2
1010	$p(1-p)p(1-p)$	2
1011	$p(1-p)pp$	3
1100	$pp(1-p)(1-p)$	2
1101	$pp(1-p)p$	3
1110	$ppp(1-p)$	3
1111	$pppp$	4

The expectation of a $B(n, p)$ random variable is

$$E(X) = E(X_1) + E(X_2) + \ldots + E(X_n) = np.$$

The variance of a $B(n, p)$ random variable is

$$\mathrm{Var}(X) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n) = np(1-p).$$

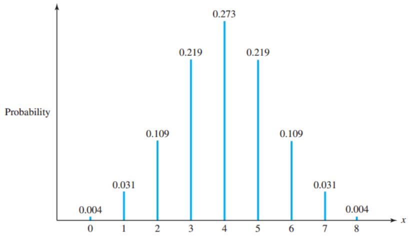
Example Consider $X \sim B(8, 0.5)$. We have

$$P(X = 3) = \binom{8}{3} \times 0.5^3 \times (1 - 0.5)^5 = 0.219$$

and

$$F(1) = P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= \binom{8}{0} \times 0.5^0 \times (1 - 0.5)^8 + \binom{8}{1} \times 0.5^1 \times (1 - 0.5)^7 = 0.004 + 0.031 = 0.35.$$



x	0	1	2	3	4	5	6	7	8
$P(X \leq x)$	0.004	0.035	0.144	0.363	0.636	0.855	0.965	0.996	1.000

Notice that X is a symmetric distribution.

More generally, $B(n, 0.5)$ is a symmetric probability distribution for any value of n. The distribution is symmetric about the expected value $n/2$.

We can look at the $B(8, 0.5)$ distribution as the distribution of the number of heads obtained in 8 tosses of a fair coin.

The probability of obtaining exactly 4 heads in 8 tosses is

$$P(X = 4) = 0.273.$$

The probability of obtaining at least 6 heads in 8 tosses is

$$P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.885 = 0.145.$$

The expected number of heads obtained in 8 tosses is

$$E(X) = np = 8 \times 0.5 = 4.$$

The variance is

$$\mathrm{Var}(X) = np(1 - p) = 8 \times 0.5 \times (1 - 0.5) = 2.$$

Example Suppose that a fair die is rolled 8 times and that \$1 is won each time a 5 or a 6 is scored.

The distribution of the winnings is then given by the $B(8, 1/3)$ distribution since there is a probability of $p = 1/3$ of winning \$1 on each of the 8 rolls of the die.

The probability that exactly \$4 is won is

$$P(X = 4) = 0.171.$$

The probability that no more than \$2 is won is

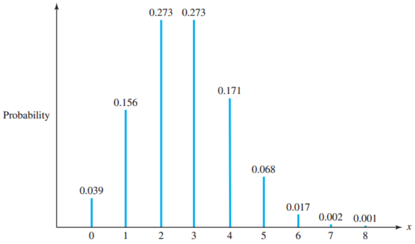
$$P(X \leq 2) = 0.468.$$

The expected winnings are

$$E(X) = np = 8 \times \frac{1}{3} = \$2.67.$$

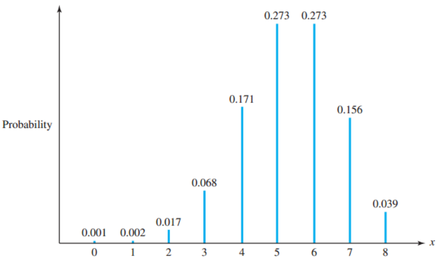
The variance is

$$\mathrm{Var}(X) = np(1-p) = 8 \times \frac{1}{3} \times \frac{2}{3} = 1.78.$$



x	0	1	2	3	4	5	6	7	8
$P(X \leq x)$	0.039	0.195	0.468	0.741	0.912	0.980	0.997	0.999	1.000

The following figure illustrates the probability mass function and cumulative distribution function of a $B(8, 2/3)$ random variable.



x	0	1	2	3	4	5	6	7	8
$P(X \leq x)$	0.001	0.003	0.020	0.088	0.259	0.532	0.805	0.961	1.000

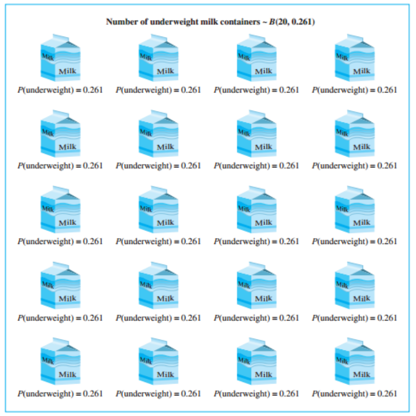
Note that the $B(8, 2/3)$ distribution is the reflection of the $B(8, 1/3)$ distribution about $x = 4$.

In general, let $X \sim B(n, p)$ and $Y = n - X$. Then $Y \sim B(n, 1 - p)$.

That is, if X counts the number of successes in n Bernoulli trials with the probability of a success being p, then $Y = n - X$ counts the number of failures in n Bernoulli trials with the probability of a failure being $1 - p$.

Example (Milk container contents) Recall that there is a probability of 0.261 that a milk container is underweight. Suppose that the milk containers are shipped to retail outlets in boxes of 20 containers. What is the distribution of the number of underweight containers in a box?

Answer Each milk container represents a Bernoulli trial with a constant probability $p = 0.261$ of being underweight.



Assume that the milk contents of 2 different milk containers are independent of each other.

Let X be the number of underweight containers in a box. Then X has binomial distribution with $n = 20$ and $p = 0.261$.

The expected number of underweight cartons in a box is

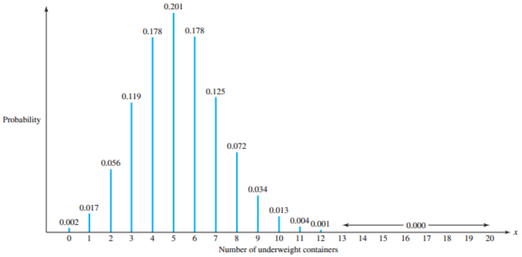
$$E(X) = np = 20 \times 0.261 = 5.22$$

and the variance is

$$\mathrm{Var}(X) = np(1 - p) = 20 \times 0.261 \times 0.739 = 3.86.$$

So the standard deviation is $\sigma = \sqrt{3.86} = 1.96$.

The following figure represents the p.m.f of X:



x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	$\cdots$	20
$P(X \leq x)$	0.002	0.019	0.075	0.194	0.372	0.573	0.751	0.876	0.948	0.982	0.995	0.999	1.000	1.000	$\cdots$	1.000

The probability that a box contains exactly 7 underweight containers is

$$P(X=7) = \binom{20}{7} \times 0.261^7 \times 0.739^{13} = 0.125.$$

The probability that a box contains no more than 3 underweight containers is

$$P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 0.1935.$$

Games of chance

Coin tossing Suppose that a fair coin is tossed n times.

Let X be the number of heads obtained. Then $X \sim B(n, 0.5)$.

We have

$$E(X) = np = \frac{n}{2}$$

and

$$\mathrm{Var}(X) = np(1-p) = \frac{n}{4}.$$

The *proportion* of heads obtained is $Y = \frac{X}{n}$ which has

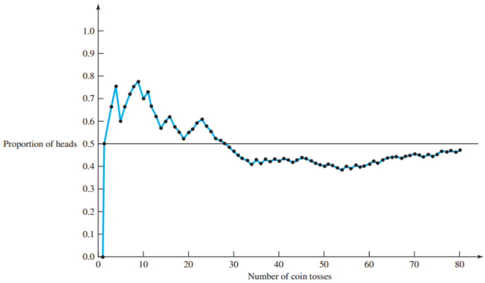
$$E(Y) = \frac{E(X)}{n} = \frac{1}{2} \quad \text{and} \quad \mathrm{Var}(Y) = \frac{\mathrm{Var}(X)}{n^2} = \frac{1}{4n}.$$

Implication In practice, as the coin is tossed more and more times, the *number* of heads may vary widely about the expected value $n/2$.

But there will be a tendency for the *proportion* of heads to become closer and closer to the value $1/2$.

n	Result of nth coin toss	Number of heads in first n coin tosses	Proportion of heads in first n coin tosses
1	tail	0	0.000
2	head	1	0.500
3	head	2	0.667
4	head	3	0.750
5	tail	3	0.600
6	head	4	0.667
7	head	5	0.714
8	head	6	0.750
9	head	7	0.778
10	tail	7	0.700
11	head	8	0.727
12	tail	8	0.667
13	tail	8	0.615
14	tail	8	0.571
15	head	9	0.600
16	head	10	0.625
17	tail	10	0.588
18	tail	10	0.556
19	tail	10	0.526
20	head	11	0.550
21	head	12	0.571
22	head	13	0.591
23	head	14	0.609
24	tail	14	0.583
25	tail	14	0.560
26	tail	14	0.538
27	tail	14	0.519
28	tail	14	0.500
29	tail	14	0.483
30	tail	14	0.467
31	tail	14	0.452
32	tail	14	0.438
33	tail	14	0.424
34	tail	14	0.412
35	head	15	0.429
36	tail	15	0.417
37	head	16	0.432
38	tail	16	0.421
39	head	17	0.436
40	tail	17	0.425

n	Result of nth coin toss	Number of heads in first n coin tosses	Proportion of heads in first n coin tosses
41	head	18	0.439
42	tail	18	0.429
43	tail	18	0.419
44	head	19	0.432
45	head	20	0.444
46	tail	20	0.435
47	tail	20	0.426
48	tail	20	0.417
49	tail	20	0.408
50	tail	20	0.400
51	head	21	0.412
52	tail	21	0.404
53	tail	21	0.396
54	tail	21	0.389
55	head	22	0.400
56	tail	22	0.393
57	head	23	0.404
58	tail	23	0.397
59	head	24	0.407
60	head	25	0.417
61	head	26	0.426
62	tail	26	0.419
63	head	27	0.429
64	head	28	0.438
65	head	29	0.446
66	tail	29	0.439
67	tail	29	0.433
68	head	30	0.441
69	head	31	0.449
70	head	32	0.457
71	tail	32	0.451
72	tail	32	0.444
73	head	33	0.452
74	tail	33	0.446
75	head	34	0.453
76	head	35	0.461
77	tail	35	0.455
78	head	36	0.462
79	tail	36	0.456
80	head	37	0.463



In general, we have the following.

Proportion of success in Bernoulli trials

If the random variable X counts the number of successes in n independent Bernoulli trials with a constant success probability p, so that $X \sim B(n, p)$, then the *proportion* of successes $Y = X/n$ has an expected value and variance of

$$E(Y) = p \quad \text{and} \quad \mathrm{Var}(Y) = \frac{p(1-p)}{n}.$$

The variance of the proportion Y decreases as the number of trials n increases. So there is a tendency for Y to become closer and closer to the success probability p as the number of trials n increases.

Definition of the geometric distribution

Consider a sequence of independent Bernoulli trials with a constant success probability p.

We are now interested in counting the number of trials performed until the *first* success occurs.

This defines the so-called geometric random variable X with parameter p.

The p.m.f of X is given by

$$P(X = x) = (1-p)^{x-1}\, p$$

for $x = 1, 2, 3, 4, \ldots$.

The expected value is

$$E(X) = \frac{1}{p}$$

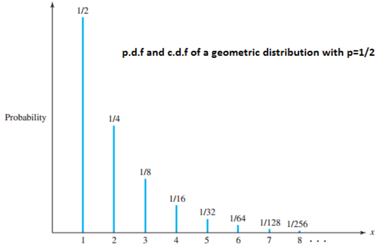
and the variance is

$$\mathrm{Var}(X) = \frac{1-p}{p^2}.$$

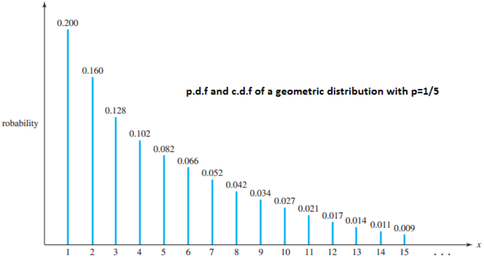
The c.d.f is

$$P(X \le x) = \sum_{i=1}^{x} P(X = i) = \sum_{i=1}^{x} (1-p)^{i-1}\, p = p\Big(1 + (1-p) + (1-p)^2 + \ldots + (1-p)^{x-1}\Big)$$

$$= p \times \frac{1-(1-p)^x}{1-(1-p)} = 1-(1-p)^x.$$



x	1	2	3	4	5	6	7	8	···
$P(X \le x)$	1/2	3/4	7/8	15/16	31/32	63/64	127/128	255/256	···



x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	···
$P(X \le x)$	0.200	0.360	0.488	0.590	0.672	0.738	0.790	0.832	0.866	0.893	0.914	0.931	0.945	0.956	0.965	···

Example (Telephone ticket sales) Telephone ticket sales for a popular event are handled by a bank of telephone salespersons who start accepting calls at a specified time. In order to get through to an operator, a caller has to be lucky enough to place a call at just the right time when a salesperson has become free from a previous client. Suppose that the chance of this is 0.1. What is the distribution of the number of calls that a person needs to make until a salesperson is reached?

Answer In this problem, a placing of a call represents a Bernoulli trial with a success probability of $p = 0.1$. Here we define that reaching a salesperson is a success.

Let X be the number of phone calls that a person makes until a salesperson is reached. Then X is a sequence of the above Bernoulli trials performed until the the *first* success. Therefore X is a random variable with the geometric distribution of parameter $p = 0.1$.

The probability that a caller gets through on the fifth attempt is

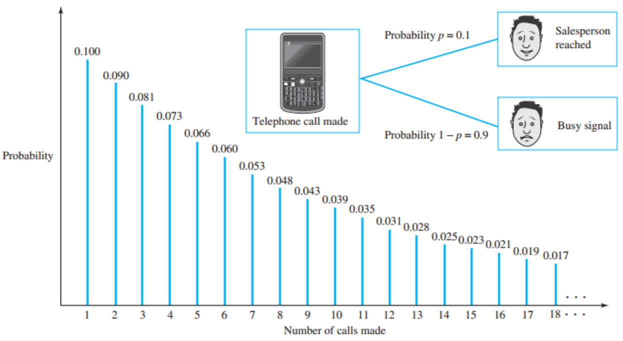
$$P(X = 5) = 0.9^4 \times 0.1 = 0.066.$$

The expected number of calls needed to get through to a salesperson is

$$E(X) = \frac{1}{p} = \frac{1}{0.1} = 10.$$

The probability that 15 or more calls are needed is

$$P(X \geq 15) = 1 - P(X \leq 14) = 1 - (1 - 0.9^{14}) = 0.9^{14} = 0.229.$$



Definition of the negative binomial distribution

Next we consider situations in which we count the number of trials performed until the rth success occurs.

This defines the so-called negative binomial random variable X with 2 parameters r and p.

The p.m.f of X is given by

$$P(X = x) = \binom{x-1}{r-1} (1-p)^{x-r} \, p^r$$

for $x = r, r+1, r+2, r+3, \ldots$.

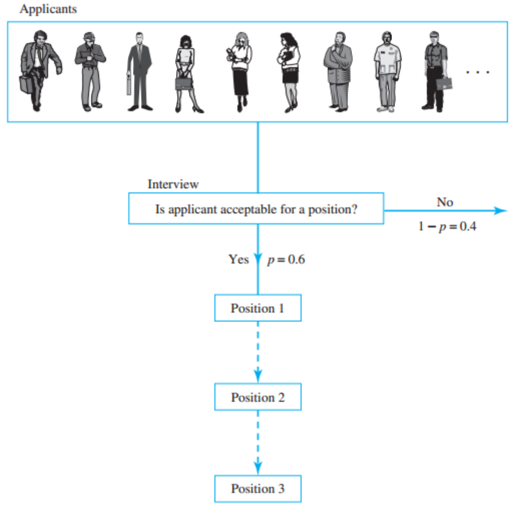
The expected value is

$$E(X) = \frac{r}{p}$$

and the variance is

$$\mathrm{Var}(X) = \frac{r(1-p)}{p^2}.$$

Example (Personnel recruitment) Suppose that a company wishes to hire 3 new workers and that each applicant interviewed has a probability of 0.6 of being found acceptable. What is the distribution of the total number of applicants that the company needs to interview?

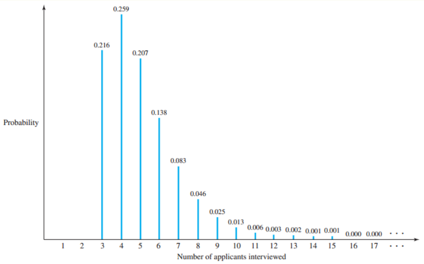


Answer In this problem, an interview represents a Bernoulli trial with a success being that the applicant is accepted for the position.
The success probability is therefore $p = 0.6$.

Let X be the number of interviews undertaken until 3 applicants are chosen. Then X is a sequence of the above Bernoulli trials performed until the third success. Therefore X is a random variable with negative geometric distribution of 2 parameters $r = 3$ and $p = 0.6$.

The probability that exactly 6 applicants interviewed is

$$P(X = 6) = \binom{5}{2} \times 0.4^3 \times 0.6^3 = 0.138.$$



The probability that at most 6 applicants interviewed is

$$P(X \leq 6)$$

$$= P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$$

$$= 0.216 + 0.256 + 0.207 + 0.138 = 0.82.$$

The expected number of interview is $E(X) = \frac{r}{p} = \frac{3}{0.6} = 5$.

Games of chance

Die rolling If a fair die is repeatedly thrown, the number of throws made until a 6 is obtained has a geometric distribution with parameter $p = 1/6$.

The expected number of throws made until a 6 is obtained is therefore

$$E(X) = \frac{1}{p} = 6.$$

The probability that a 6 is *not* obtained in the first six throws is

$$P(X \geq 7) = 1 - P(X \leq 6) = 1 - \left(1 - \left(\frac{5}{6}\right)^6\right) = \left(\frac{5}{6}\right)^6 = 0.335.$$

The distribution of the number of throws made until a 6 is obtained for the *second* time has a negative binomial distribution with parameters $p = 1/6$ and $r = 2$.

The expected number of throws required is

$$E(X) = \frac{r}{p} = 12.$$

The probability that two 6s are obtained in the first four throws is

$$P(X \leq 4) = P(X = 2) + P(X = 3) + P(X = 4)$$

$$= 0.028 + 0.046 + 0.058 = 0.132.$$