3 Discrete Random Variables and Probability Distributions

Introduction

เชิงคุณภาพ เชิงปริมาณ

- whether experiment yields qualitative or quantitative outcomes, methods of statistical analysis require that we focus on certain numerical aspects of data such as
  - Sample proportion x/n
  - Mean $\overline{X}$
  - Standard deviation S
- The concept of random variable allows us to pass from

  Experimental Outcomes → Numerical Outcomes

- There are two fundamentally different types of Random Variable
  - Discrete Random Variables (ตัวแปรสุ่มแบบไม่ต่อเนื่อง)
  - Continuous Random Variables (ตัวแปรสุ่มแบบต่อเนื่อง)

3.1 Random Variables

Random Variables

- In any experiment, there are numerous characteristics that can be observed or measured, but in most cases experimenter will focus on some specific aspect or aspects of a sample
- For example : in study of commuting patterns in a metropolitan area
  - Each individual in sample might be asked about
    - commuting distance
    - number of people commuting in same vehicle
  - but not about
    - IQ
    - Income
    - Family size, and other such characteristics
- Alternatively, researcher may test a sample of components and record only the number that have failed within 1000 hours, rather than record the individual failure times

Random Variables

- In general, each outcome of experiment can be associated with a number by specifying a rule of association e.g.,
  - Number among sample of ten components that fail to last 1000 hours or
  - Total weight of baggage for a sample of 25 airline passengers
- Such a rule of association is called a **Random** **Variable**

because observed value depends on which of possible experimental outcomes results

because different numerical values are possible

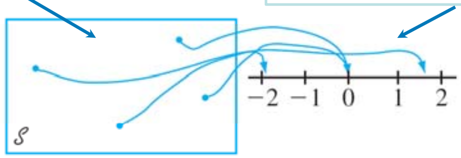


Figure 3.1 A random variable

Random Variables

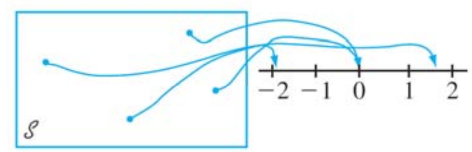
Definition

For given sample space $\mathcal{S}$ of some experiment,

Random Variable (rv) is any rule that associates a number with each outcome in $\mathcal{S}$.

In mathematical language,

Random Variable is a function whose domain is sample space and whose range is set of real numbers.



Random Variables

- Random variables are customarily denoted by uppercase letters, such as X and Y, near the end of our alphabet.

- In contrast to our previous use of a lowercase letter, such as x, to denote a variable, we will now use lowercase letters to represent some particular value of corresponding random variable.

- The notation $X(s) = x$ *means that* x *is value associated with outcome* s *by random variable* X.

Random variable X → $X(s) = x$ ← *value associated with outcome s*

outcome s

Example 1

When a student calls a university help desk for technical support, he/she will either immediately be able to speak to someone (S, for success) or will be placed on hold (F, for failure).

With $\mathcal{S} = \{S, F\}$, define an rv X by

$$X(S) = 1 \qquad X(F) = 0$$

The rv X indicates whether (1) or not (0) student can immediately speak to someone.

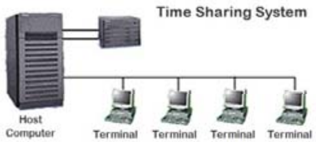
Example 3.1

- When student attempts to log on to a computer time-sharing system, either

  all ports are busy (F) in which case student will fail to obtain access,

  or else

  there is at least one port free (S), in which case student will be successful in accessing the system



With $\mathcal{S}$ = {S, F}, define an rv X by $X(S)=1$ $X(F)=0$

rv X indicates whether (1) or not (0) student can log on

rv X in ex. 3.1 was specified by explicitly listing each element of $\mathcal{S}$ and associated number

Such listing is tedious if $\mathcal{S}$ contains more than a few outcomes, but it can frequently be avoided

Example 3.2

Consider the experiment in which a telephone number in a certain area code is dialed using a random number dialer (such devices are used extensively by polling organizations), and define an rv Y by

$$Y = \begin{cases} 1 \text{ if the selected number is unlisted} \\ 0 \text{ if the selected number is listed in the directory} \end{cases}$$

For example

If 5282966 appears in the telephone directory, then Y(5282966) = 0, whereas Y(7727350) = 1 tells us that number 7727350 is unlisted

A word description of this sort is more economical than a complete listing, so we will use such a description whenever possible

In ex.3.1 and 3.2, the only possible values of this random variable were 0 and 1

Such a random variable arises frequently enough to be given a special name, after the individual who first studied it.

Random Variables

Definition

Any random variable whose only possible values are

0 and 1

is called

Bernoulli Random Variable

Random variable X $\quad X(s)=x \quad$ *value associated with outcome s*

outcome s

We will sometimes want to define and study several different random variables from the same sample space

Example 3.3

Example 2.3 described an experiment in which the number of pumps in use at each of two six-pump gas stations was determined. Define rv's *X, Y,* and *U* by

X = Total number of pumps in use at two stations

Y = Difference between the number of pumps in use at station 1 and the number in use at station 2

U = Maximum of numbers of pumps in use at two stations

outcome

If this experiment is performed and s = (2,3) results, then

$X(s)=x \Rightarrow X((2,3)) = 2+3 = 5 \Rightarrow$ The observed value of X was x = 5

$Y(s)=y \Rightarrow Y((2,3)) = 2-3 = -1 \Rightarrow$ The observed value of Y was y = -1

$U(s)=u \Rightarrow U((2,3)) = \max(2,3) = 3 \Rightarrow$ The observed value of U was u = 3

Example 3.4

In Example 2.4, we considered the experiment in which batteries were examined until a good one (S) was obtained

The sample space was $\mathcal{S} = \{S, FS, FFS, FFFS, \ldots\}$

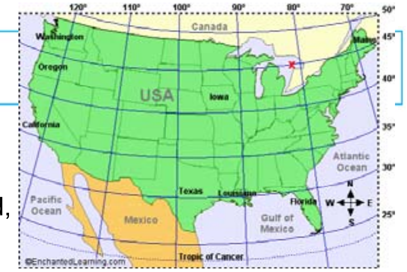
Define an random variable (rv) X by

X = Number of batteries examined before the experiment terminates

Then

$$X(S) = 1$$
$$X(FS) = 2$$
$$X(FFS) = 3$$
$$\vdots$$
$$X(FFFFFFS) = 7$$
$$\vdots$$

Example 3.5



Suppose that in some random fashion, a location (latitude and longitude) in the continental United States in selected,

Define an random variable (rv) Y by

Y = The height above sea level at the selected location

For example, if selected location were (39^0 $50^/$ N, 98^0 $35^/$ W), then we might have

$$Y((39^0 50' N, 98^0 35' W)) = 1{,}748.26\ ft$$

The largest possible value of Y is 14,494 (Mt. Whitney)

The smallest possible value of Y is -282 (Death Valley)

The set of all possible values of Y is the set of all numbers in the interval between -282 and 14,494

$$\{y : y \text{ is a number}, -282 \le y \le 14{,}494\}$$

and there are an infinite number of numbers in this interval

Two Types of Random Variables

Two Types of Random Variables

- **Discrete** Random Variable (ตัวแปรสุ่มแบบไม่ต่อเนื่อง)

- **Continuous** Random Variable (ตัวแปรสุ่มแบบต่อเนื่อง)

Definition

Discrete random variable is an random variable whose possible values either constitute a finite set or else can be listed in an infinite sequence in which there is a first element, a second element, and so on (“countably” infinite).

Two Types of Random Variables

Definition

Random variable is **continuous** if *both* of the following apply:

1. Its set of possible values consists either of
all numbers in a single interval on the number line
(possibly infinite in extent, e.g., from $-\infty$ to ∞) or
all numbers in a disjoint union of such intervals
(e.g., [0, 10] $\cup$ [20, 30]).

2. No possible value of the variable has positive probability, that is, P(X = c) = 0 for any possible value c.

Example 3.6

All random variables in Examples 3.1 –3.4 are discrete.

As another example, suppose we select married couples at random and do a blood test on each person until we find a husband and wife who both have the same Rh factor.

With
X = the number of blood tests to be performed, possible values of *X* are *D* = {2, 4, 6, 8, …}.

Since the possible values have been listed in sequence, *X* is a discrete rv.

3.2 Probability Distributions for Discrete Random Variables

การแจกแจงความน่าจะเป็นสำหรับตัวแปรสุ่มแบบไม่ต่อเนื่อง

Probability Distributions for Discrete Random Variables

Probabilities assigned to various outcomes in $\mathcal{S}$ in turn determine probabilities associated with values of any particular random variable *X*.

Probability Distribution of X says how the total probability of 1 is distributed among (allocated to) the various possible *X* values.

Probability Distributions for Discrete Random Variables

Suppose, for example, that a business has just purchased four laser printers, and

let X be number among these that require service during the warranty period.

Possible X values are then 0, 1, 2, 3, and 4.

Probability distribution will tell us

- how probability of 1 is subdivided among these five possible values
- how much probability is associated with the X value 0,
- how much is apportioned to the X value 1, and so on.

We will use the following notation for probabilities in the distribution:

$p(0)$ = Probability of the X value 0 = $P(X = 0)$

$p(1)$ = Probability of the X value 1 = $P(X = 1)$ and so on.

In general, $p(x)$ will denote probability assigned to the value x.

$$P(X = x) = p(x)$$

Example 3.7

The Cal Poly Department of Statistics has a lab with six computers reserved for statistics majors.

Let X denote number of these computers that are in use at a particular time of day.

Suppose that probability distribution of X is as given in following table;

Random variable

possible X values

x	0	1	2	3	4	5	6
$p(x)$	.05	.10	.15	.25	.20	.15	.10

$P(X = x)$

probability of each such value

Example 3.7

cont'd

We can now use elementary probability properties to calculate other probabilities of interest.

For example, probability that มากที่สุด **at most** 2 computers are in use is

x	0	1	2	3	4	5	6
$P(X = x) \rightarrow p(x)$	.05	.10	.15	.25	.20	.15	.10

$$P(X \le 2) = P(X = 0 \text{ or } 1 \text{ or } 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= p(0) + p(1) + p(2)$$
$$= 0.05 + 0.10 + 0.15$$
$$= 0.30$$

Example 3.7

cont'd

Since the event อย่างน้อย *at least 3 computers are in use* is complementary to *at most 2 computers are in use,*

which can, of course, also be obtained by adding together probabilities for the values, 3, 4, 5, and 6.

x	0	1	2	3	4	5	6
$p(x)$	.05	.10	.15	.25	.20	.15	.10

$$P(X \ge 3) = P(X = 3 \text{ or } 4 \text{ or } 5 \text{ or } 6)$$
$$= P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$$
$$= p(3) + p(4) + p(5) + p(6)$$
$$= 0.25 + 0.20 + 0.15 + 0.10$$
$$= 0.70$$

or

$$P(X \ge 3) = 1 - P(X \le 2)$$
$$= 1 - 0.30$$
$$= 0.70$$

Example 3.7

x	0	1	2	3	4	5	6
$p(x)$	.05	.10	.15	.25	.20	.15	.10

cont'd

Probability that between 2 and 5 computers *inclusive* (รวมทั้ง) are in use is

$$P(2 \le X \le 5) = P(X = 2 \text{ or } 3 \text{ or } 4 \text{ or } 5)$$
$$= P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$
$$= p(2) + p(3) + p(4) + p(5)$$
$$= 0.15 + 0.25 + 0.20 + 0.15$$
$$= 0.75$$

whereas the probability that the number of computers in use is *strictly between* 2 and 5 is

(อย่างเคร่งครัด, อย่างแน่นอน)

$$P(2 < X < 5) = P(X = 3 \text{ or } 4)$$
$$= P(X = 3) + P(X = 4)$$
$$= p(3) + p(4)$$
$$= 0.25 + 0.20$$
$$= 0.45$$

Probability Distributions for Discrete Random Variables

Definition การแจกแจงความน่าจะเป็น หรือ ฟังก์ชันความน่าจะเป็นแบบไม่ต่อเนื่อง

Probability Distribution or **Probability Mass Function** (pmf) of discrete random variable is defined for every number x by

outcome Random variable *X* *value associated with outcome s*

$$p(x) = P(X = x) = P(\text{all } s \in \mathcal{S}: X(s) = x)$$

In words, for every possible value x of random variable, **pmf** specifies **probability of observing that value** when the experiment is performed.

The conditions $p(x) \ge 0$ and $\sum_{\text{all possible } x} p(x) = 1$ are required of any pmf.

The pmf of X in the previous example was simply given in the problem description.

We now consider several examples in which various probability properties are exploited to obtain the desired distribution.

Example 3.8

- Six lots of components are ready to be shipped by a certain supplier.
- Number of defective components in each lot is as follows:

Lot	1	2	3	4	5	6
Number of defectives	0	2	0	1	2	0

- One of these lots is to be randomly selected for shipment to particular customer
- Let X be number of defectives in selected lot
- Three possible X values are 0, 1, and 2

$$X(1,3,6) = 0) \quad X(4) = 1) \quad X(2,5) = 2)$$

$$p(x) = P(X = x) = P(\text{all } s \in \{1,2,3,4,5,6\}: X(s) = x)$$

Cont.

Example 3.8

$$p(x) = P(X = x) = P(\text{all } s \in \{1,2,3,4,5,6\}: X(s) = x)$$

- Of the six equally likely simple events,

Three result in X=0 One in X=1

Lot	1	2	3	4	5	6
Number of defectives	0	2	0	1	2	0

Two in X=2

- Then

$$p(0) = P(X = 0) = P(\text{lot 1 or 3 or 6 is sent}) = \frac{3}{6} = 0.500$$ Probability of 0.5 is distributed to X value 0

$$p(1) = P(X = 1) = P(\text{lot 4 is sent}) = \frac{1}{6} = 0.167$$ Probability of 0.167 is placed on X value 1

$$p(2) = P(X = 2) = P(\text{lot 2 or 5 is sent}) = \frac{2}{6} = 0.333$$ Probability of 0.333 is associated with X value 2

- Value of X along with their probabilities collectively specify the pmf
- If this experiment were repeated over and over again, in the long run
  - X = 0 occur one-half of the time,
  - X = 1 one-sixth of the time, and
  - X = 2 one-third of the time

Example 3.9

$p(x) = P(X = x) = P(all\ s \in \{desktop, laptop\} : X(s) = x)$

- Consider whether the next person buying a computer at a university book store buys a laptop or a desktop model

Let

$$X = \begin{cases} 1 & if\ the\ customer\ purchases\ a\ laptop\ computer \\ 0 & if\ the\ customer\ purchases\ a\ desktop\ computer \end{cases}$$

- If 20% of all purchasers during that week select a laptop, the pmf for X is

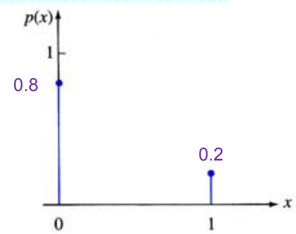
$$p(0) = P(X = 0) = P(next\ customer\ purchases\ a\ desktop\ model) = 0.8$$

$$p(1) = P(X = 1) = P(next\ customer\ purchases\ a\ laptop\ model) = 0.2$$

$$p(x) = P(X = x) = 0\ for\ x \neq 0\ or\ 1$$

- An equivalent description is

$$p(x) = \begin{cases} 0.8 & if\ x = 0 \\ 0.2 & if\ x = 1 \\ 0 & if\ x \neq 0\ or\ 1 \end{cases}$$



Line graph for pmf of Example 3.9

Example 3.10

$p(y) = P(Y = y) = P(all\ s \in \{a, b, c, d, e\} : Y(s) = y)$

- Consider a group of five potential blood donors – a, b, c, d, and e – of whom only a and b have type O+ blood.
- Five blood samples, one from each individual, will be typed in random order until and O+ individual is identified
- Let rv Y = Number of typings necessary to identify an O+ individual.
- Then the pmf of Y is

$$p(1) = P(Y = 1) = P(a\ or\ b\ typed\ first) = \frac{2}{5} = 0.4$$

$$p(2) = P(Y = 2) = P(c, d, or\ e\ first, and\ then\ a\ or\ b)$$

$$= P(c, d, or\ e\ first) \cdot P(a\ or\ b\ next \mid c, d, or\ e\ first) = \frac{3}{5} \cdot \frac{2}{4} = 0.3$$

$$p(3) = P(Y = 3) = P(c, d, or\ e\ first\ and\ second, and\ then\ a\ or\ b)$$

$$= \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{2}{3}\right) = 0.2$$

$$p(4) = P(Y = 4) = P(c, d, and\ e\ all\ done\ first) = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right) = 0.1$$

$$p(y) = 0 \quad if\ y \neq 1, 2, 3, 4$$

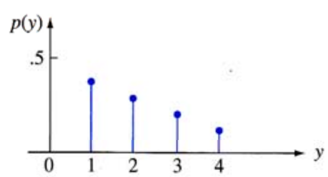
Cont.

Example 3.10

- In tabular form, the pmf is

y	1	2	3	4
p(y)	0.4	0.3	0.2	0.1

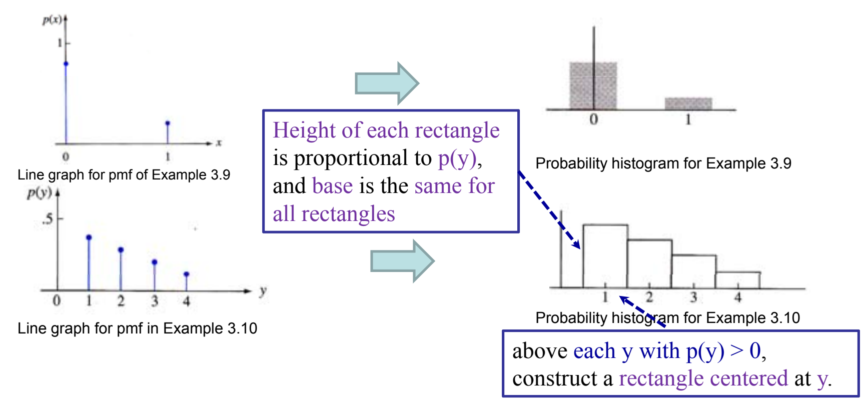
- where any y value not listed receives zero probability



Line graph for pmf in Example 3.10

Probability Histogram

- Another useful pictorial representation of a pmf, called a **Probability Histogram**, is similar to histograms



- When possible values are equally spaced, base is frequently chosen as distance between successive y values

Example 3.11

It is often helpful to think of pmf as specifying mathematical model for discrete pupulation

- Consider selecting at random student who is among 15,000 registered for current term at Mega University
- Let X = number of courses for which the selected student is registered, and suppose that X has the following pmf

x	1	2	3	4	5	6	7
p(x)	0.01	0.03	0.13	0.25	0.39	0.17	0.02

- One way to view this situation is to think of population as consisting of 15,000 individuals, each having his or her own X value; the proportion with each X value is given by p(x)
- an alternative viewpoint is to forget about students and think of population itself as consisting of X values:
- There are some 1s in population, some 2s, …, and finally some 7s
- Population then consists of number 1,2,…,7 (so is discrete), and p(x) give a model for distribution of population values
- Once we have such population model, we will use it to compute values of pupulation characteristics (e.g., mean) and make inferences about such characteristics

Parameter of a Probability Distribution

ถึงข้อ ๒ อันเก่ากวน ๗-ดิ้น

Parameter of a Probability Distribution

- In Example 3.9, the pmf of Bernoulli rv X was
  - $p(0) = 0.8$ and
  - $p(1) = 0.2$ because 20% of all purchasers selected a desktop computer.
- At another store, it may be the case that
  - $p(0) = 0.9$ and
  - $p(1) = 0.1$.
- More generally, pmf of any Bernoulli rv can be expressed in the form
  - $p(1) = \alpha$ and
  - $p(0) = 1 - \alpha$, where $0 < \alpha < 1$.
- Because pmf depends on the particular value of α we often write **$p(x; \alpha)$** rather than just **$p(x)$**:

$$p(x; \alpha) = \begin{cases} 1 - \alpha & \text{if } x = 0 \\ \alpha & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases} \qquad (3.1)$$

Parameter of a Probability Distribution

Then each choice of a in Expression (3.1) yields a different pmf.

Definition

p(1) = 0.8 p(0) = 0.2	p(1) = 0.9 p(0) = 0.1	p(1) = 0.4 p(0) = 0.6	⇨	p(1) = α p(0) = 1- α

- Suppose $p(x)$ depends on quantity that can be assigned any one of a number of possible values, with each different value determining different probability distribution.
- Such a quantity is called a **Parameter** of Distribution.
- Collection of all probability distributions for **different values of parameter** is called a **Family** of probability distributions.

Parameter of a Probability Distribution

$$p(x;\alpha)=\begin{cases}1-\alpha & \text{if } x=0\\ \alpha & \text{if } x=1\\ 0 & \text{otherwise}\end{cases} \qquad (3.1)$$

- Quantity α in Expression (3.1) is a parameter.
- Each different number α between 0 and 1 determines different member of the family of distributions
- Two such number are

$$p(x;0.6)=\begin{cases}0.4 & if\ x=0\\ 0.6 & if\ x=1\\ 0 & otherwise\end{cases} \quad \text{and} \quad p(x;0.5)=\begin{cases}0.5 & if\ x=0\\ 0.5 & if\ x=1\\ 0 & otherwise\end{cases}$$

- Each probability distribution for a Bernoulli random variable has the form of Expression (3.1), so it is called the **Bernoulli Family of Distributions**

Example 3.12

Starting at a fixed time, we observe the gender of each newborn child at a certain hospital until a boy (B) is born.

Let $p = P(B)$,

assume that successive births are independent, and define random variable X by x = number of births observed.

$$p(1)=P(X=1)=P(B)=p$$

$$p(2)=P(X=2)=P(GB)=P(G)\cdot P(B)=(1-p)\cdot p$$

$$p(3)=P(X=3)=P(GGB)=P(G)\cdot P(G)\cdot P(B)=(1-p)^2\cdot p$$

$$\vdots$$

Continuing in this way, a general formula emerges:S

Quantity p represents number between 0 and 1 and is **Parameter of Probability Distribution**

$$p(x)=\begin{cases}(1-p)^{x-1}p & x=1,2,3,\ldots\\ 0 & otherwise\end{cases} \qquad (3.2)$$

Expression (3.2) describes the family of ***Geometric Distributions***.

การแจกแจงทางเรขาคณิต

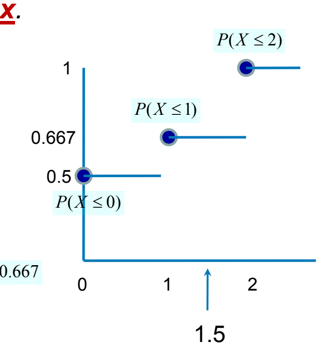
Cumulative Distribution Function

ฟังก์ชันการแจกแจงสะสม

Cumulative Distribution Function

- For some fixed value x, we often wish to compute probability that observed value of X will be **at most x**.
- For example, pmf in Example 3.8 was

$$p(x)=\begin{cases}0.500 & x=0\\ 0.167 & x=1\\ 0.333 & x=2\\ 0 & otherwise\end{cases}$$



- Probability that X is at most 1 is then

$$P(X\le 1)=P(X=0)+P(X=1)=p(0)+p(1)=0.500+0.167=0.667$$

- In this example, $P(X\le 1.5)\ iff\ X\le 1$, so

$$P(X\le 1.5)=P(X\le 1)=0.667$$

- Similarly,

$$P(X\le 0)=P(X=0)=0.5$$

$$P(X\le 0.75)=P(X=0)=0.5$$

Cumulative Distribution Function

And in fact for any x satisfying $0 \le x < 1$, $P(X \le x) = 0.5$.

The largest possible X value is 2, so

$$p(x) = \begin{cases} 0.500 & x = 0 \\ 0.167 & x = 1 \\ 0.333 & x = 2 \\ 0 & otherwise \end{cases}$$

$P(X \le 2) = 1 \quad P(X \le 3.7) = 1 \quad P(X \le 20.5) = 1$

and so on.

Notice that $P(X < 1) < P(X \le 1)$

since the latter includes the probability of the X value 1, whereas the former does not.

More generally, when X is discrete and x is a possible value of the variable, $P(X < x) < P(X \le x)$.

Cumulative Distribution Function

Definition

Cumulative Distribution Function (cdf) $F(x)$ of a discrete rv variable X with pmf $p(x)$ is defined for every number x by

$$F(x) = P(X \le x) = \sum_{y: y \le x} p(y) \qquad \textbf{(3.3)}$$

For any number x, F(x) is the probability that the observed value of X will be **at most x**

Example 3.13

A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory.

The accompanying table gives

distribution of Y = the amount of memory in a purchased drive:

ขนาดของหน่วยความจำของ Flash Drive ที่ขาย

y	1	2	4	8	16
$p(y)$	.05	.10	.35	.40	.10

$$F(y) = P(Y \le y)$$

Cumulative Distribution Function (cdf)

Example 3.13

cont'd

Let's first determine $F(y)$ for each of the five possible values of Y:

y	1	2	4	8	16
$p(y)$	.05	.10	.35	.40	.10

$$F(y) = P(Y \le y)$$

$$\begin{aligned} F(1) &= P(Y \le 1) \\ &= P(Y = 1) \\ &= p(1) \\ &= 0.05 \end{aligned}$$

$$\begin{aligned} F(2) &= P(Y \le 2) \\ &= P(Y = 1 \text{ or } 2) \\ &= p(1) + p(2) = 0.05+0.10 \\ &= 0.15 \end{aligned}$$

Example 3.13

cont'd

$$F(4) = P(Y \le 4)$$
$$= P(Y = 1 \text{ or } 2 \text{ or } 4)$$
$$= p(1) + p(2) + p(4)$$
$$= 0.05 + 0.10 + 0.35$$
$$= 0.50$$

y	1	2	4	8	16
$p(y)$	.05	.10	.35	.40	.10

$$F(y) = P(Y \le y)$$

$$F(8) = P(Y \le 8)$$
$$= p(1) + p(2) + p(4) + p(8)$$
$$= 0.05 + 0.10 + 0.35 + 0.40 + 0.10$$
$$= 0.90$$

$$F(16) = P(Y \le 16)$$
$$= p(1) + p(2) + p(4) + p(8) + p(16)$$
$$= 0.05 + 0.10 + 0.35 + 0.40 + 0.10$$
$$= 1$$

Example 3.13

cont'd

Now for any other number y,

$F(y)$ will equal the value of F at the closest possible value of Y to the left of y.

For example,

y	1	2	4	8	16
$p(y)$	.05	.10	.35	.40	.10

$$F(y) = P(Y \le y)$$

$$F(2.7) = P(Y \le 2.7)$$
$$= P(Y \le 2)$$
$$= F(2)$$
$$= 0.15$$

$$F(7.999) = P(Y \le 7.999)$$
$$= P(Y \le 4)$$
$$= F(4)$$
$$= 0.50$$

Example 3.13

cont'd

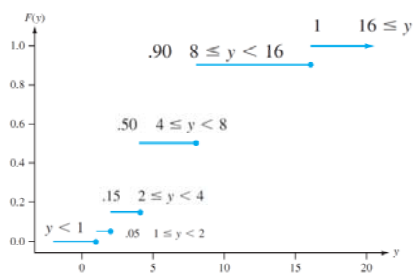
If y is less than 1, $F(y) = 0$ [e.g. $F(.58) = 0$],

and if y is at least 16, $F(y) = 1$ [e.g. $F(25) = 1$].

The cdf is thus

y	1	2	4	8	16
$p(y)$	.05	.10	.35	.40	.10

$$F(y) = \begin{cases} 0 & y < 1 \\ .05 & 1 \le y < 2 \\ .15 & 2 \le y < 4 \\ .50 & 4 \le y < 8 \\ .90 & 8 \le y < 16 \\ 1 & 16 \le y \end{cases}$$



A graph of this cdf

Example 3.13

The pmf of Y (the number of blood typings) in Example 3.10 was

y	1	2	3	4
p(y)	0.4	0.3	0.2	0.1

We first determine F(y) for each value in set {1, 2, 3, 4} of possible values :

$$F(1) = P(Y \le 1) = P(Y = 1) = p(1) = 0.4$$

$$F(2) = P(Y \le 2) = P(Y = 1 \text{ or } 2) = p(1) + p(2) = 0.4 + 0.3 = 0.7$$

$$F(3) = P(Y \le 3) = P(Y = 1 \text{ or } 2 \text{ or } 3) = p(1) + p(2) + p(3) = 0.4 + 0.3 + 0.7 = 0.9$$

$$F(4) = P(Y \le 4) = P(Y = 1 \text{ or } 2 \text{ or } 3 \text{ or } 4) = p(1) + p(2) + p(3) + p(4) = 0.4 + 0.3 + 0.7 + 0.1 = 1.0$$

Cont.

Example 3.13

y	1	2	3	4
p(y)	0.4	0.3	0.2	0.1

Now for any other number y,

F(y) will equal the value of F at the closest possible value of Y to the left of y

For example,

$$F(2.7) = P(Y \le 2.7) = P(Y \le 2) = p(1) + p(2) = 0.4 + 0.3 = 0.7$$

$$F(3.999) = P(Y \le 3.999) = P(Y \le 3) = p(1) + p(2) + p(3) = 0.4 + 0.3 + 0.2 = 0.9$$

The cdf is thus

$$F(y) = \begin{cases} 0 & if\ y < 1 \\ 0.4 & if\ 1 \le y < 2 \\ 0.7 & if\ 2 \le y < 3 \\ 0.9 & if\ 3 \le y < 4 \\ 1 & if\ 4 \le y \end{cases}$$

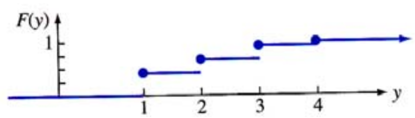


Figure 3.5 A graph of the cdf of Example 3.13

For X a discrete rv, the graph of $F(x)$ will have a jump at every possible value of X and will be flat between possible values. Such a graph is called a **step function**.

Example 3.14

Example 3.14 In Example 3.12, any positive integer was a possible X value, and the pmf was

$$p(x) = \begin{cases} (1-p)^{x-1}p & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

For any positive integer x,

$$F(x) = \sum_{y \le x} p(y) = \sum_{y=1}^{x} (1-p)^{y-1} p = p \sum_{y=0}^{x-1} (1-p)^y \tag{3.4}$$

Geometric Series

$$a + ar + ar^2 + \ldots + ar^{n-1} = \sum_{k=0}^{n-1} ar^k = a \cdot \frac{1-r^n}{1-r}$$

To evaluate this sum, we use the fact that the partial sum of a geometric series is

$$\sum_{y=0}^{k} a^y = \frac{1-a^{k+1}}{1-a}$$

Using this in Equation (3.4), with $a = 1 - p$ and $k = x - 1$, gives

$$\boxed{F(x) = p \cdot \frac{1-(1-p)^x}{1-(1-p)} = 1 - (1-p)^x} \quad x \text{ a positive integer}$$

Cont.

Example 3.14

Since F is constant in between positive integers,

$$F(x) = \begin{cases} 0 & x < 1 \\ 1 - (1-p)^{[x]} & x \ge 1 \end{cases} \tag{3.5}$$

where $[x]$ is the largest integer $\le x$ (e.g., $[2.7] = 2$).

if p = 0.51 as in the birth example, then probability of having to examine at most five births to see the first boy is F(5)

$F(5) = 1\text{-}(1\text{-}0.51)^5 = 1\text{-}0.49^5 = 1\text{-}0.0282 = 0.9718$

$F(10) = 1\text{-}(1\text{-}0.51)^{10} = 1\text{-}0.49^{10} = 1\text{-}0.000798 = 0.999202$ $\quad F(10) \approx 1.0000.$

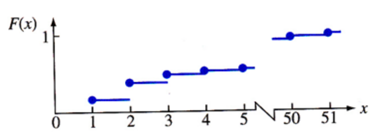


Figure 3.6 A graph of $F(x)$ for Example 3.14

Cont.

Example 3.14

In examples thus far, the cdf has been derived from the pmf. This process can be reversed to obtain the pmf from the cdf whenever the latter function is available. For example, consider again the rv of Example 3.7 (the number of pumps in use at a gas station); possible X values are 0, 1, . . . , 6. Then

$$\begin{aligned} p(3) &= P(X = 3) \\ &= [p(0) + p(1) + p(2) + p(3)] - [p(0) + p(1) + p(2)] \\ &= P(X \le 3) - P(X \le 2) \\ &= F(3) - F(2) \end{aligned}$$

More generally, the probability that X falls in a specified interval is easily obtained from the cdf. For example,

$$\begin{aligned} P(2 \le X \le 4) &= p(2) + p(3) + p(4) \\ &= [p(0) + \cdots + p(4)] - [p(0) + p(1)] \\ &= P(X \le 4) - P(X \le 1) \\ &= F(4) - F(1) \end{aligned}$$

Notice that $P(2 \le X \le 4) \ne F(4) - F(2)$. This is because the X value 2 is included in $2 \le X \le 4$, so we do not want to subtract out its probability. However, $P(2 < X \le 4) = F(4) - F(2)$ because $X = 2$ is not included in the interval $2 < X \le 4$.

The Cumulative Distribution Function

For X a discrete rv, the graph of $F(x)$ will have a jump at every possible value of X and will be flat between possible values. Such a graph is called a **step function.**

Proposition

For any two numbers a and b with $a \leq b$,

$$P(a \leq X \leq b) = F(b) - F(a-)$$

where "$a-$" represents the largest possible X value that is strictly less than a.

Cumulative Distribution Function

In particular, if the only possible values are integers and if a and b are integers, then

$$P(a \leq X \leq b) = P(X = a \text{ or } a + 1 \text{ or} \ldots \text{ or } b)$$

$$= F(b) - F(a - 1)$$

Taking $a = b$ yields $P(X = a) = F(a) - F(a - 1)$ in this case.

Cumulative Distribution Function

The reason for subtracting $F(a-)$ rather than $F(a)$ is that we want to include $P(X = a)$

$F(b) - F(a)$; gives $P(a < X \leq b)$.

This proposition will be used extensively when computing binomial and Poisson probabilities in Sections 3.4 and 3.6.

Example 15

- Let X = the number of days of sick leave taken by a randomly selected employee of a large company during a particular year.
- If the maximum number of allowable sick days per year is 14, possible values of X are 0, 1, . . . , 14.

With $F(0) = 0.58$,
$F(1) = 0.72$,
$F(2) = 0.76$,
$F(3) = 0.81$,
$F(4) = 0.88$,
$F(5) = 0.94$,

$$P(2 \leq X \leq 5) = P(X = 2, 3, 4, \text{ or } 5)$$
$$= F(5) - F(1)$$
$$= 0.94 - 0.72$$
$$= 0.22$$

and

$$P(X = 3) = F(3) - F(2)$$
$$= 0.81 - 0.76$$
$$= 0.05$$

3.3 Expected Values

The Expected Value of *X*

Expected Values

- Consider university having 15,000 students
- Let X = number of courses for which randomly selected student is registered
- pmf of X follows

x	1	2	3	4	5	6	7
$p(x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

- since p(1)=0.01, we know that (0.01) (15,000) = 150 of student are registered for one course

Expected Values Cont.

$$\textit{average no. of courses per student} = \frac{\textit{total no. of courses taken by all students}}{\textit{total no. of students}}$$

(*average value of X in Population*)

x	1	2	3	4	5	6	7
$p(x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

Each of 150 students is taking one course

150 contribute 1(150) courses to total

Each of 450 students is taking two courses

450 contribute 2(450) courses to total

$$\textit{Population averge value of } X = \frac{1(150) + 2(450) + 3(1950) + \cdots + 7(300)}{15{,}000} = 4.57$$

$$1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) + 4 \cdot p(4) + 5 \cdot p(5) + 6 \cdot p(6) + 7 \cdot p(7)$$

Expected Values

Cont.

Population averge value of X $= 1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) + 4 \cdot p(4) + 5 \cdot p(5) + 6 \cdot p(6) + 7 \cdot p(7)$

To compute population average value of X,
only possible values of X along with their probabilities are needed

Population size is irrelevant as long as pmf is given

Average or mean value of X is then weighted average of possible values 1, 2, …, 7
Where weights are probabilities of those values

Expected Value of X

Definition

Let X be discrete random variable with set of possible values D and pmf $p(x)$

Expected Value or **Mean value** of X, denoted by

- $E(X)$ or
- μ_X

is

X f O
o outcome x pop

$$\boxed{E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)}$$

When it is clear to which X the expected value refers, μ rather than μ_X is often used.

Example 3.16

Consider a university having 15,000 students and
let X = of courses for which randomly selected student is registered.
The pmf of X follows.

x	1	2	3	4	5	6	7
$p(x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

$$\boxed{E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)}$$

$$
\begin{aligned}
E(X) = \mu_X = \mu &= [1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) + 4 \cdot p(4) + 5 \cdot p(5) + 6 \cdot p(6) + 7 \cdot p(7)] \\
&= [(1)(0.01) + (2)(0.03) + (3)(0.13) + (4)(0.25) + (5)(0.39) + (6)(0.17) + (7)(0.02)] \\
&= 0.01 + 0.06 + 0.39 + 1.00 + 1.95 + 1.02 + 0.14 \\
&= 4.57
\end{aligned}
$$

If we think of population as consisting of the X values 1, 2, . . . , 7, then $\mu = 4.57$ is **Population Mean**.

In the sequel, we will often refer to μ as the ***Population Mean*** rather than the **Mean of *X* in the population.**

Example 3.16

cont'd

Notice that

μ here is not 4, the ordinary average of 1, . . . , 7,
because distribution puts more weight on 4, 5, and 6 than on other X values.

x	1	2	3	4	5	6	7
$p(x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

Example 3.17

APGAR เป็นการตรวจที่ใช้ไปทั่วโลกในห้องคลอดเพื่อ ที่จะประเมินสุขภาพทั่วไป และความสมบูรณ์ของทารกแรกเกิด

○ Just after birth, each newborn child is rated on a scale called Apgar Scale. Possible ranges are 0, 1, …, 10, with child's rating determined by color (สีผิว), muscle tone (ความตึงตัวของกล้ามเนื้อ), respiratory effort (ความสามารถในการหายใจ), heartbeat (ภาวะการเต้นของหัวใจ), and reflex irritability (การตอบสนองต่อการกระตุ้น)

(the best possible score is 10).

○ Let X be Apgar score of a randomly selected child born at a certain hospital during next year, and suppose that pmf of X is

x	0	1	2	3	4	5	6	7	8	9	10
$p(x)$	.002	.001	.002	.005	.02	.04	.18	.37	.25	.12	.01

Then mean of value X is

$$E(X) = \mu = 0(.002) + 1(.001) + 2(.002) + \cdots + 8(.25) + 9(.12) + 10(.01) = 7.15$$

○ Again μ is not a possible value of variable X.

○ Also, because variable refers to future child, there is no concrete existing population to which μ refers.

○ Instead, we think of pmf as model for conceptual population consisting of value 0, 1, 2, … ,10.

○ Mean value of this conceptual population is $\mu = 7.15$.

Example 3.18

Let X=1, if a randomly selected component needs warranty service and = 0 otherwise.

Then X is Bernoulli random variable with pmf

$$p(x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \\ 0 & x \neq 0, 1 \end{cases}$$

from which

$$E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0(1-p) + 1(p) = p.$$

○ That is, expected value of X is just probability that X takes on value 1.

○ If we conceptualize a population consisting of *0*s in proportion *1-p* and *1*s in proportion *p*, then

○ Population average is $\mu = p$

$$p(1) = P(X=1) = P(B) = p$$

Ex. 3.12

$$p(2) = P(X=2) = P(GB) = P(G) \cdot P(B) = (1-p) \cdot p$$

$$p(3) = P(X=3) = P(GGB) = P(G) \cdot P(G) \cdot P(B) = (1-p)^2 \cdot p$$

$\vdots$

Example 3.19

The general form for pmf of X =number of children born up to and including the first boy is

$$p(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

From the definition

$$E(X) = \sum_D x \cdot p(x) = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = p \sum_{x=1}^{\infty} \left[-\frac{d}{dp}(1-p)^x \right] \qquad (3.9)$$

If we interchange the order of taking derivative and summation, the sum is that geometric series.

After sum is computed, derivative is taken, and final results is

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

If *p is near 1*, we expect to see a boy very soon, whereas if *p is near 0*, we expect many births before the first boy.

$$a_n = \frac{1}{1-r} \quad \cancel{\frac{1}{1-\frac{1}{2}}} \quad \frac{1}{\frac{1}{2}} = 2$$

Cont.

Example 3.19

$\frac{1}{2}$ ✓ $\quad$ $\frac{1}{2}$ × $\frac{1}{2}$ × $\frac{1}{2}$ ✓

$\frac{1}{2}$ × $\frac{1}{2}$ ✓

For $p = 0.5$, $E(X) = 1/p = 1/0.5 = 2$

There is another frequently used interpretation of μ. Consider the pmf

$$p(x) = \begin{cases} (.5) \cdot (.5)^{x-1} & \text{if } x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

○ This is pmf of X = the number of tosses of a fair coin necessary to obtain the first H (a special case of Example 3.19).

○ Suppose we observe value *x* from this pmf (toss a coin until H appears), then observe independently another value (keep tossing), then another, and so on.

○ If after observing a very large number of *x* values, we average them, resulting sample average will be very near to $\mu = 2$.

○ This is , μ can be interpreted as the long-run average observed value of X when the experiment is performed repeatedly

Example 3.20

Let X, the number of interviews student has prior to getting a job, have pmf

$$p(x) = \begin{cases} k/x^2 & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

where k is chosen so that $\sum_{x=1}^{\infty}\left(\frac{k}{x^2}\right) = 1$

The expected value of X is $\boxed{E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)}$

$$\mu = E(X) = \sum_{x=1}^{\infty} x \cdot \frac{k}{x^2} = k \sum_{x=1}^{\infty} \frac{1}{x} \qquad (3.10)$$

Harmonic Series

The Variance of X

The Variance of X

○ Expected value of X describes where probability distribution is centered

○ Using physical analogy of placing point mass p(x) at value x on one-dimensional axis, if axis were then supported by a fulcrum placed at μ, there would be no tendency for axis to tilt.

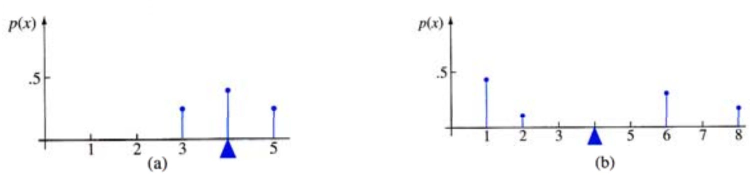


Figure 3.7 Two different probability distributions with $\mu = 4$

○ Although both distributions have the same center μ,

- ○ distribution of Figure 3.7(b) has greater spread or variability or dispersion than does that of Figure 3.7(a).

○ We will use **Variance of X** to assess the amount of variability in (distribution of) X

The Variance of X

Definition

Let X have pmf $p(x)$ and expected value μ.

Then **Variance** of X, denoted by $V(X)$ or σ^2_X, or just σ^2, is

$$V(X) = \sum_D (x-\mu)^2 \cdot p(x) = E\left[(X-\mu)^2\right]$$

Standard Deviation (SD) of X is

$$\sigma_X = \sqrt{\sigma_X^2}$$

○ If most of probability distribution is close to μ, then σ^2 will be relatively small.

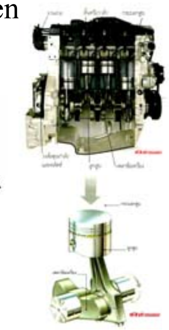
○ However, if there are x values far from μ that have large $p(x)$, then σ^2 will be quite large.

Example 3.24

x	4	6	8
$p(x)$	.5	.3	.2

If X is the number of cylinders on the next car to be tuned at a service facility

[p(4)=0.5, p(6) = 0.3, p(8) = 0.2, from which μ = 5.4], then



$$V(X) = \sum_D (x-\mu)^2 \cdot p(x) = E\left[(X-\mu)^2\right]$$

$$V(X) = \sigma^2 = \sum_{x=4}^{8} (x - 5.4)^2 \cdot p(x)$$

$$= (4 - 5.4)^2(.5) + (6 - 5.4)^2(.3) + (8 - 5.4)^2(.2) = 2.44$$

The standard deviation of X is $\sigma = \sqrt{2.44} = 1.562$.

When pmf $p(x)$ specifies a mathematical model for the distribution of population values, both σ^2 and σ measure the spread of values in population; σ^2 is population variance, and σ is population standard deviation.

A Shortcut Formula for σ^2

A Shortcut Formula for σ^2 $V(X) = \sum_D (x-\mu)^2 \cdot p(x) = E\left[(X-\mu)^2\right]$

The number of arithmetic operations necessary to compute σ^2 can be reduced by using an alternative formula.

Proposition $V(X) = \sum_D (X-\mu)^2 \cdot p(x) = E\left[(X-\mu)^2\right]$

$$V(X) = \sigma^2 = \left[\sum_D x^2 \cdot p(x)\right] - \mu^2 = E(X^2) - [E(X)]^2$$

$E(\square) = \square \cdot p(\square)$

In using this formula, $E(X^2)$ is computed first without any subtraction; then $E(X)$ is computed, squared, and subtracted (once) from $E(X^2)$.

A Shortcut Formula for σ^2

Proof of the Shortcut Formula

Expand $(x-\mu)^2$ in the definition of σ^2 to obtain $x^2 - 2\mu x + \mu^2$, and then carry $\sum$ through to each of the three terms:

$$\sigma^2 = \sum_D x^2 \cdot p(x) - 2\mu \cdot \sum_D x \cdot p(x) + \mu^2 \sum_D p(x)$$

$$= E(X^2) - 2\mu \cdot \mu^2 = E(X^2) - \mu^2$$

Example 3.25

The pmf of the number of cylinders X on next car to be tuned at a certain facility was given in Example 3.24 as

$p(4) = 0.5$, $p(6) = 0.3$, and $p(8) = 0.2$, from which $\mu = 5.4$, and

$$V(X) = \sigma^2 = \left[\sum_D x^2 \cdot p(x)\right] - \mu^2 = E(X^2) - [E(X)]^2$$

x	4	6	8
$p(x)$	.5	.3	.2

$E(X) = \sum_D x\, p(x)$

$= 4(0.5) + 6(0.3) + 8(0.2)$

$= 2 + 1.8 + 1.6$

$= 5.4$

Thus

$$E(X^2) = (4^2)(0.5) + (6^2)(0.3) + (8^2)(0.2) = 31.6$$

$$V(X) = \sigma^2 = E(X^2) - [E(X)]^2$$

$$= 31.6 - (5.4)^2 = 2.44$$

$31.6 - (5.4)^2 = \ldots$

As in Example 3.24.

3.4 Binomial Probability Distribution

The Binomial Probability Distribution

There are many experiments that conform either exactly or approximately to the following list of requirements:

1. Experiment consists of sequence of n smaller experiments called *trials*, where n is fixed in advance of experiment.

   รู้ n แน่นอน

2. Each trial can result in one of the same two possible outcomes (dichotomous trials), which we generically denote by success (S) and failure (F).

   มีผลลัพธ์ 2 แบบเท่านั้น

3. Trials are independent, so that outcome on any particular trial does not influence outcome on any other trial.

   เหตุการณ์ไม่เกี่ยวข้องกัน

The Binomial Probability Distribution

4. Probability of success $P(S)$ is <u>constant</u> from trial to trial; we denote this probability by p.

รู้ค่า p

Definition

An experiment for which Conditions 1–4 are satisfied is called a **Binomial Experiment.**

Example 3.27

- The same coin is tossed successively and independently n times.
- We arbitrarily use
  - S to denote outcome H (heads) and
  - F to denote the outcome T (tails).
- Then this experiment satisfies Conditions 1–4.

- Tossing a thumbtack n times, with
  - S = point up and
  - F = point down, also results in a binomial experiment.

Many experiments involve a sequence of independent trials for which there are more than two possible outcomes on any one trial.

Binomial experiment can then be created by dividing the possible outcomes into two groups.

Example 3.28

- Color of pea seeds is determined by a single genetic locus.
- If two alleles at this locus are *AA* or *Aa* (genotype), then pea will be yellow (phenotype), and
- if allele is *aa*, pea will be green.
- Suppose we pair of 20 *Aa* seeds and cross the two seeds in each of ten pairs to obtain ten new genotypes
- Call each new genotype a *success* S if it is *aa* and a failure otherwise
- Then with this identification of S and F, the experiment is binomial with $n=10$ and $p=P$ *(aa genotype)*
- If each member of pair is equally likely to contribute *a or A*, then $p=P(a).P(a)$

$$p = P(a).P(a) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

Example 3.29

- Suppose a certain city has 50 licensed restaurants, of which 15 currently have at least one serious health code violation and the other 35 have no serious violation.

- There are five inspectors, each of whom will inspect one restaurant during the coming week.

- Name of each restaurant is written on a different slip of paper, and after slips are thoroughly mixed, each inspector in turn draws one of the slips *without replacement*.

- Label the ith trial as a success if the ith restaurant selected (i=1,2,…,5) has no serious violations. Then

Example 3.29

Cont.

$$P(S\ on\ 1^{st}\ trial) = \frac{35}{50} = 0.7$$
$$P(S\ on\ 2^{nd}\ trial \mid S\ on\ 1^{st}\ trial) = \frac{34}{49} = 0.6938$$
$$P(S\ on\ 3^{rd}\ trial \mid S\ on\ 2^{nd}\ trial) = \frac{33}{48} = 0.6875$$
$$P(S\ on\ 4^{th}\ trial \mid S\ on\ 3^{rd}\ trial) = \frac{32}{47} = 0.6800$$
$$P(S\ on\ 5^{th}\ trial \mid S\ on\ 4^{th}\ trial) = \frac{31}{46} = 0.6739$$

$$P(S \text{ on first trial}) = \frac{35}{50} = .70$$

$$\begin{aligned} P(S \text{ on second trial}) &= P(SS) + P(FS) \\ &= P(\text{second } S \mid \text{first } S)\, P(\text{first } S) \\ &\quad + P(\text{second } S \mid \text{first } F)\, P(\text{first } F) \\ &= \frac{34}{49}\cdot\frac{35}{50} + \frac{35}{49}\cdot\frac{15}{50} = \frac{35}{50}\left(\frac{34}{49}+\frac{15}{49}\right) = \frac{35}{50} = .70 \end{aligned}$$

Similarly, it can be shown that $P(S \text{ on } i\text{th trial}) = .70$ for $i = 3, 4, 5$. However,

$$P(S \text{ on fifth trial} \mid SSSS) = \frac{31}{46} = .67$$

whereas

$$P(S \text{ on fifth trial} \mid FFFF) = \frac{35}{46} = .76$$

The experiment is not binomial because the trials are not independent. In general, if sampling is without replacement, the experiment will not yield independent trials. If each slip had been replaced after being drawn, then trials would have been independent, but this might have resulted in the same restaurant being inspected by more than one inspector. ■

Example 3.30

A certain state has 500,000 licensed drivers, of whom 400,000 are insured. A sample of 10 drivers is chosen without replacement. The ith trial is labeled S if the ith driver chosen is insured. Although this situation would seem identical to that of Example 3.29, the important difference is that the size of the population being sampled is very large relative to the sample size. In this case

$$P(S \text{ on } 2 \mid S \text{ on } 1) = \frac{399{,}999}{499{,}999} = .80000$$

and

$$P(S \text{ on } 10 \mid S \text{ on first } 9) = \frac{399{,}991}{499{,}991} = .799996 \approx .80000$$

These calculations suggest that although the trials are not exactly independent, the conditional probabilities differ so slightly from one another that for practical purposes the trials can be regarded as independent with constant $P(S) = .8$. Thus, to a very good approximation, the experiment is binomial with $n = 10$ and $p = .8$. ■

We will use the following rule of thumb in deciding whether a "without-replacement" experiment can be treated as a binomial experiment.

The Binomial Probability Distribution

Rule

Consider sampling without replacement from a dichotomous population of size N.

If sample size (number of trials) n is at most 5% of population size, experiment can be analyzed as though it were exactly a **Binomial Experiment**.

Ex.

$N = 50$, $n = 5$ ➔ $n/N = 5/50 = 0.1 > 0.05$: binomial experiment is not a good approximation

$N = 500{,}000$, $n = 10$ ➔ $n/N = 10/500{,}000 = 0.00002 < 0.05$: binomial experiment is a good approximation

Binomial Random Variable and Distribution

The Binomial Random Variable and Distribution

In most binomial experiments, it is the total number of S's, rather than knowledge of exactly which trials yielded S's, that is of interest.

Definition

Binomial Random Variable *X* associated with a binomial experiment consisting of n trials is defined as

X = the **number** of $\boldsymbol{S}$'s among the $\boldsymbol{n}$ **trials**

ตัวแปรสุ่มไบโนเมียล X คือ จำนวนครั้งของความสำเร็จในการทดลอง n ครั้ง

The Binomial Random Variable and Distribution

Suppose, for example, that $n = 3$.

Then there are eight possible outcomes for the experiment:

SSS SSF SFS SFF FSS FSF FFS FFF

From definition of X,

$X(SSS)=3$, $X(SSF)=2$, $X(SFS)=2$, $X(SFF)=1$, $X(FSS)=2$, $X(FSF)=1$, $X(FFS)=1$, and $X(FFF)=0$

Possible values for X in an ***n*-trial** experiment are

$x = 0, 1, 2, 3, \ldots, n$

We will often write **X~Bin (*n*, *p*)** to indicate that X is binomial rv based on **n trials** with **success probability *p***

Binomial Random Variable and Distribution

Notation

Because **pmf** of binomial rv X depends on two parameters n and p, we denote the pmf by $b(x; n, p)$.

Binomial Random Variable and Distribution

- Consider first the case $n = 4$ for which each outcome, its probability, and corresponding x value are listed in Table 3.1.

Outcome	x	Probability	Outcome	x	Probability
SSSS	4	p^4	FSSS	3	$p^3(1-p)$
SSSF	3	$p^3(1-p)$	FSSF	2	$p^2(1-p)^2$
SSFS	3	$p^3(1-p)$	FSFS	2	$p^2(1-p)^2$
SSFF	2	$p^2(1-p)^2$	FSFF	1	$p(1-p)^3$
SFSS	3	$p^3(1-p)$	FFSS	2	$p^2(1-p)^2$
SFSF	2	$p^2(1-p)^2$	FFSF	1	$p(1-p)^3$
SFFS	2	$p^2(1-p)^2$	FFFS	1	$p(1-p)^3$
SFFF	1	$p(1-p)^3$	FFFF	0	$(1-p)^4$

Table 3.1 Outcomes and Probabilities for a Binomial Experiment with four Trials

- For example,

$$\begin{aligned} P(SSFS) &= P(S)\cdot P(S)\cdot P(F)\cdot P(S) \qquad (\textit{indenpendent trials}) \\ &= p\cdot p\cdot(1-p)\cdot p \qquad (\textit{cons}\tan t\ P(S)) \\ &= p^3\cdot(1-p) \end{aligned}$$

Binomial Random Variable and Distribution

- In this special case, we wish $b(x; 4, p)$ for $x = 0,1,2,3,$ and 4
- For $b(3; 4, p)$, and $x = 3$
- Sum probabilities associated with each such outcome

$$b(3;4,p) = P(FSSS) + P(SFSS) + P(SSFS) + P(SSSF) = 4p^3\cdot(1-p)$$

- There are four outcomes with $x=3$ and each has probability $p^3(1-p)$
  - *Order of S*'s and *F*'s is not important, but only the number of *S*'s, so

$$b(3;4,p) = \left\{\begin{matrix} \textit{number of outcomes} \\ \textit{with } X=3 \end{matrix}\right\}\cdot\left\{\begin{matrix} \textit{probability of any particular outcome} \\ \textit{with } X=3 \end{matrix}\right\}$$

Binomial Random Variable and Distribution

- Similarly, $b(2; 4, p) = 6p^2(1-p)^2$, which is also product of the number of outcomes with $X=2$ and probability of any such outcome

- In general,

$$b(x;4,p) = \left\{ \begin{matrix} \text{number of sequences of} \\ \text{length } n \text{ consisting of } x\ S\text{'s} \end{matrix} \right\} \cdot \left\{ \begin{matrix} \text{probability of any} \\ \text{particular such sequence} \end{matrix} \right\}$$

Binomial Random Variable and Distribution

- Since ordering of S's and F's is not important, the second factor in the previous equation is

$$p^x(1-p)^{n-x}$$

(e.g., the first x trials resulting in S and the last n-x resulting in F.)

- First factor is the number of ways of choosing x of n trials to be S's

Number of combinations of size x that can be constructed from n distinct objects (trial here).

The Binomial Random Variable and Distribution

THEOREM

$$b(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0,1,2,\ldots,n \\ 0 & \text{otherwise} \end{cases}$$

Example 3.31

- Each of six randomly selected cola drinkers is given glass containing cola S and one containing cola F.
- Glasses are identical in appearance except for a code on the bottom to identify cola.
- Suppose there is actually no tendency among cola drinkers to prefer one cola to the other.

Then $p = P$(a selected individual prefers S) = 0.5, so with
X = the number among the six who prefer S,
$X \sim \text{Bin}(6,0.5)$.

Thus

$$P(X=3) = b(3;\ 6,\ 0.5) = \binom{6}{3}(0.5)^3(0.5)^3 = 20(0.5)^6 = 0.313$$

Example 3.31

cont'd

Probability that at least three prefer S is

$$P(X \geq 3) = \sum_{x=3}^{6} b(x; 6, 0.5)$$

$$= \sum_{x=3}^{6} \binom{6}{x} (0.5)^x (0.5)^{6-x}$$

$$= 0.656$$

and Probability that at most one prefers S is

$$P(X \leq 1) = \sum_{x=0}^{1} b(x; 6, 0.5)$$

$$= 0.109$$

Using Binomial Tables

Using Binomial Tables

Even for a relatively small value of n, the computation of binomial probabilities can be tedious.

Appendix Table A.1 tabulates the cdf $F(x) = P(X \leq x)$ for n = 5, 10, 15, 20, 25 in combination with selected values of p.

Notation

For $X \sim \text{Bin}(n, p)$, *the cdf will be denoted by*

$$P(X \leq x) = B(x; n, p) = \sum_{y=0}^{x} b(y; n, p) \qquad x = 0,1,2,\ldots,n$$

Various other probabilities can then be calculated using the proposition on cdf's. (from Section 3.2 : Probability Distributions for Discrete Random Variables)

Using Binomial Tables

Table A.1 Cumulative Binomial Probabilities

$$P(X \leq x) = B(x; n, p) = \sum_{y=0}^{x} b(y; n, p) \qquad x = 0,1,2,\ldots,n$$

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

a. $n = 5$

	p: 0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
x = 0	.951	.774	.590	.328	.237	.168	.078	.031	.010	.002	.001	.000	.000	.000	.000
1	.999	.977	.919	.737	.633	.528	.337	.188	.087	.031	.016	.007	.000	.000	.000
2	1.000	.999	.991	.942	.896	.837	.683	.500	.317	.163	.104	.058	.009	.001	.000
3	1.000	1.000	1.000	.993	.984	.969	.913	.812	.663	.472	.367	.263	.081	.023	.001
4	1.000	1.000	1.000	1.000	.999	.998	.990	.969	.922	.832	.763	.672	.410	.226	.049

b. $n = 10$

	p: 0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
x = 0	.904	.599	.349	.107	.056	.028	.006	.001	.000	.000	.000	.000	.000	.000	.000
1	.996	.914	.736	.376	.244	.149	.046	.011	.002	.000	.000	.000	.000	.000	.000
2	1.000	.988	.930	.678	.526	.383	.167	.055	.012	.002	.000	.000	.000	.000	.000
3	1.000	.999	.987	.879	.776	.650	.382	.172	.055	.011	.004	.001	.000	.000	.000
4	1.000	1.000	.998	.967	.922	.850	.633	.377	.166	.047	.020	.006	.000	.000	.000
5	1.000	1.000	1.000	.994	.980	.953	.834	.623	.367	.150	.078	.033	.002	.000	.000
6	1.000	1.000	1.000	.999	.996	.989	.945	.828	.618	.350	.224	.121	.013	.001	.000
7	1.000	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.474	.322	.070	.012	.000
8	1.000	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.756	.624	.264	.086	.004
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.944	.893	.651	.401	.096

Using Binomial Tables

$$P(X \le x) = B(x;n,p) = \sum_{y=0}^{x} b(y;n,p) \qquad x = 0,1,2,\ldots,n$$

c. $n = 15$

x \ p	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
2	1.000	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
3	1.000	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
4	1.000	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
5	1.000	1.000	.998	.939	.852	.722	.402	.151	.034	.004	.001	.000	.000	.000	.000
6	1.000	1.000	1.000	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
7	1.000	1.000	1.000	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
8	1.000	1.000	1.000	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
9	1.000	1.000	1.000	1.000	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000
10	1.000	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.314	.164	.013	.001	.000
11	1.000	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.539	.352	.056	.005	.000
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.764	.602	.184	.036	.000
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.920	.833	.451	.171	.010
14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.987	.965	.794	.537	.140

(continued)

Using Binomial Tables

Table A.1 Cumulative Binomial Probabilities *(cont.)*

$$B(x;n,p) = \sum_{y=0}^{x} b(y;n,p)$$

d. $n = 20$

$$P(X \le x) = B(x;n,p) = \sum_{y=0}^{x} b(y;n,p) \qquad x = 0,1,2,\ldots,n$$

x \ p	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.818	.358	.122	.012	.003	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.983	.736	.392	.069	.024	.008	.001	.000	.000	.000	.000	.000	.000	.000	.000
2	.999	.925	.677	.206	.091	.035	.004	.000	.000	.000	.000	.000	.000	.000	.000
3	1.000	.984	.867	.411	.225	.107	.016	.001	.000	.000	.000	.000	.000	.000	.000
4	1.000	.997	.957	.630	.415	.238	.051	.006	.000	.000	.000	.000	.000	.000	.000
5	1.000	1.000	.989	.804	.617	.416	.126	.021	.002	.000	.000	.000	.000	.000	.000
6	1.000	1.000	.998	.913	.786	.608	.250	.058	.006	.000	.000	.000	.000	.000	.000
7	1.000	1.000	1.000	.968	.898	.772	.416	.132	.021	.001	.000	.000	.000	.000	.000
8	1.000	1.000	1.000	.990	.959	.887	.596	.252	.057	.005	.001	.000	.000	.000	.000
9	1.000	1.000	1.000	.997	.986	.952	.755	.412	.128	.017	.004	.001	.000	.000	.000
10	1.000	1.000	1.000	.999	.996	.983	.872	.588	.245	.048	.014	.003	.000	.000	.000
11	1.000	1.000	1.000	1.000	.999	.995	.943	.748	.404	.113	.041	.010	.000	.000	.000
12	1.000	1.000	1.000	1.000	1.000	.999	.979	.868	.584	.228	.102	.032	.000	.000	.000
13	1.000	1.000	1.000	1.000	1.000	1.000	.994	.942	.750	.392	.214	.087	.002	.000	.000
14	1.000	1.000	1.000	1.000	1.000	1.000	.998	.979	.874	.584	.383	.196	.011	.000	.000
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.994	.949	.762	.585	.370	.043	.003	.000
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.984	.893	.775	.589	.133	.016	.000
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.965	.909	.794	.323	.075	.001
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.992	.976	.931	.608	.264	.017
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.997	.988	.878	.642	.182

(continued)

Using Binomial Tables

Table A.1 Cumulative Binomial Probabilities *(cont.)*

$$B(x;n,p) = \sum_{y=0}^{x} b(y;n,p)$$

e. $n = 25$

$$P(X \le x) = B(x;n,p) = \sum_{y=0}^{x} b(y;n,p) \qquad x = 0,1,2,\ldots,n$$

x \ p	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.778	.277	.072	.004	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.974	.642	.271	.027	.007	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.998	.873	.537	.098	.032	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	1.000	.966	.764	.234	.096	.033	.002	.000	.000	.000	.000	.000	.000	.000	.000
4	1.000	.993	.902	.421	.214	.090	.009	.000	.000	.000	.000	.000	.000	.000	.000
5	1.000	.999	.967	.617	.378	.193	.029	.002	.000	.000	.000	.000	.000	.000	.000
6	1.000	1.000	.991	.780	.561	.341	.074	.007	.000	.000	.000	.000	.000	.000	.000
7	1.000	1.000	.998	.891	.727	.512	.154	.022	.001	.000	.000	.000	.000	.000	.000
8	1.000	1.000	1.000	.953	.851	.677	.274	.054	.004	.000	.000	.000	.000	.000	.000
9	1.000	1.000	1.000	.983	.929	.811	.425	.115	.013	.000	.000	.000	.000	.000	.000
10	1.000	1.000	1.000	.994	.970	.902	.586	.212	.034	.002	.000	.000	.000	.000	.000
11	1.000	1.000	1.000	.998	.980	.956	.732	.345	.078	.006	.001	.000	.000	.000	.000
12	1.000	1.000	1.000	1.000	.997	.983	.846	.500	.154	.017	.003	.000	.000	.000	.000
13	1.000	1.000	1.000	1.000	.999	.994	.922	.655	.268	.044	.020	.002	.000	.000	.000
14	1.000	1.000	1.000	1.000	1.000	.998	.966	.788	.414	.098	.030	.006	.000	.000	.000
15	1.000	1.000	1.000	1.000	1.000	1.000	.987	.885	.575	.189	.071	.017	.000	.000	.000
16	1.000	1.000	1.000	1.000	1.000	1.000	.996	.946	.726	.323	.149	.047	.000	.000	.000
17	1.000	1.000	1.000	1.000	1.000	1.000	.999	.978	.846	.488	.273	.109	.002	.000	.000
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.993	.926	.659	.439	.220	.009	.000	.000
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.971	.807	.622	.383	.033	.001	.000
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.910	.786	.579	.098	.007	.000
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.967	.904	.766	.236	.034	.000
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.968	.902	.463	.127	.002
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.993	.973	.729	.358	.026
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.996	.928	.723	.222

Example 3.32

- Suppose that *20%* of all copies of a particular textbook fail a certain binding strength test.
  การทดสอบความแข็งแรงของการเข้าปกหนังสือ
- Let X denote the number among 15 randomly selected copies that fail the test.
- Then X has a binomial distribution with $n = 15$ and $p = 0.2$

1. Probability that <u>at most 8 fail</u> the test is $P(X \le 8) = \sum_{y=0}^{8} b(y;15,0.2) = B(8;15,0.2)$

0.999

c. $n = 15$

x \ p	0.01	0.05	0.10	$\boxed{0.20}$	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
2	1.000	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
3	1.000	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
4	1.000	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
5	1.000	1.000	.998	.939	.852	.722	.402	.151	.034	.004	.001	.000	.000	.000	.000
6	1.000	1.000	1.000	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
7	1.000	1.000	1.000	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
$\boxed{8}$	1.000	1.000	1.000	$\boxed{.999}$	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
9	1.000	1.000	1.000	1.000	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000

Example 3.32

2. Probability that exactly 8 fail is

$$P(X=8)=P(X\le 8)-P(X\le 7)=B(8;15{,}0.2)-B(7;15,\,0.2)$$

c. $n = 15$

		p														
		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
	2	1.000	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
	3	1.000	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
	4	1.000	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
	5	1.000	1.000	.998	.939	.852	.722	.402	.151	.034	.004	.001	.000	.000	.000	.000
	6	1.000	1.000	1.000	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
x	7	1.000	1.000	1.000	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
	8	1.000	1.000	1.000	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
	9	1.000	1.000	1.000	1.000	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000

$$\left.\begin{array}{l} B(8;15{,}0.2)=0.999 \\ B(7;15{,}0.2)=0.996 \end{array}\right\} \Rightarrow P(X=8)=0.999-0.996=0.003$$

Example 3.32

3. Finally, Probability that **between** 4 and 7, **inclusive**, fail is

$$\begin{aligned} P(4\le X\le 7) &= P(X=4,5,6,\,or\;7) \\ &= P(X\le 7)-P(X\le 3) \\ &= B(7;15,\,0.2)-B(3;15,\,0.2) \\ &= 0.996-0.648 \\ &= 0.348 \end{aligned}$$

c. $n = 15$

		p														
		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
	2	1.000	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
	3	1.000	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
	4	1.000	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
	5	1.000	1.000	.998	.939	.852	.722	.402	.151	.034	.004	.001	.000	.000	.000	.000
	6	1.000	1.000	1.000	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
x	7	1.000	1.000	1.000	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
	8	1.000	1.000	1.000	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
	9	1.000	1.000	1.000	1.000	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000

Notice that this latter probability is the difference between entries in the $x = 7$ and $x = 3$ rows, *not* the $x = 7$ and $x = 4$ rows. ■

Mean and Variance of *X*

The Mean and Variance of *X*

For ***n* = 1**, Binomial distribution becomes Bernoulli distribution.

Example 3.18 Let $X = 1$ if a randomly selected component needs warranty service and $= 0$ otherwise. Then X is a Bernoulli rv with pmf

$$p(x)=\begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & x\ne 0,1 \end{cases}$$

- Mean value of a Bernoulli variable is $\mathbf{E(X)}=\boldsymbol{\mu}=\boldsymbol{p}$, so the expected number of S's on any single trial is $\boldsymbol{p}$.
- Since a binomial experiment consists of ***n* trials**, intuition suggests that for $X\sim \text{Bin}(n,p)$, $\boldsymbol{E(X)=np}$,

product of number of trials and probability of success on single trial.

Mean and Variance of X

Proposition

$$\begin{aligned} &\textit{if } X \sim Bin(n,p), \textit{then} \\ &\quad E(X) = np, \\ &\quad V(X) = np(1-p), \textit{ and} \\ &\quad \sigma_X = \sqrt{npq} \qquad (\textit{where } q = 1-p) \end{aligned}$$

Proof of E(X)

Proof #1

$$\begin{aligned} E(X) &= \sum_{x=0}^{n} x \cdot p(x) \\ &= \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=1}^{n} x \cdot \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=1}^{n} n \cdot \binom{n-1}{x-1} p^x (1-p)^{n-x} \\ &= np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x} \\ &= np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{(n-1)-(x-1)} \\ &= np \sum_{j=0}^{m} \binom{m}{j} p^{j} (1-p)^{m-j} \\ &= np \end{aligned}$$

$x = 0; \quad 0 \cdot \binom{n}{0} p^0 (1-p)^{n-0} = 0$

$$\begin{aligned} x \cdot \binom{n}{x} &= x \cdot \frac{n!}{x!(n-x)!} \\ &= \frac{n!}{(x-1)!(n-x)!} \\ &= n \cdot \frac{(n-1)!}{(x-1)!(n-x)!} \\ &= n \cdot \binom{n-1}{x-1} \end{aligned}$$

$n - x = n - x - (1-1) = n - 1 - x + 1 = (n-1) - (x-1)$

let $m = n-1$ *and* $j = x-1$ $\Rightarrow$ $x = j+1$

$x = 1$
$j + 1 = 1$
$j = 0$

Example 3.34

- If 75% of all purchases at a certain store are made with a credit card
- X is the number among ten randomly selected purchases made with a credit card, then

$n=10, \ p=0.75, \ q=0.25$)

$$X \sim Bin(10, 0.75)$$

Thus,

$$\begin{aligned} E(X) &= np \\ &= (10)(0.75) \\ &= 7.5 \end{aligned}$$

$$\begin{aligned} V(X) &= npq \\ &= np(1-p) \\ &= (10)(0.75)(0.25) \\ &= 1.875 \end{aligned}$$

$$\begin{aligned} \sigma &= \sqrt{V(X)} \\ &= \sqrt{1.875} \end{aligned}$$

Again, even though X can take on only integer values, $E(X)$ need not be integer.

If we perform a large number of independent binomial experiments, each with $n = 10$ trials and $p = 0.75$, then the average number of S's per experiment will be close to 7.5.

3.5 Hypergeometric Distribution and Negative Binomial Distribution

Hypergeometric and Negative Binomial Distributions

- Hypergeometric and Negative Binomial distributions are both related to Binomial distribution.
- **Binomial distribution** is approximate probability model for sampling without replacement from finite dichotomous (S–F) population provided sample size n is small relative to population size N;
- **Hypergeometric distribution** is exact probability model for number of S's in sample.
- **Binomial random variable X** is number of S's when **number n of trials is fixed**, whereas
- **Negative Binomial distribution** arises from **fixing the number of S's desired** and letting number of trials (n) be random.

ชอบ slide จารย์มาก เนื้อแน่นเกิ๊ย

Hypergeometric Distribution

Hypergeometric Distribution

Assumptions leading to hypergeometric distribution are as follows:

1. Population or set to be sampled consists of N individuals, objects, or elements (a *finite* population).

2. Each individual can be characterized as success (S) or failure (F), and there are **M successes** in population.

3. Sample of n individuals is selected without replacement in such a way that each subset of size n is equally likely to be chosen.

- Random variable of interest is X = number of S's in sample.
- Probability distribution of X depends on the parameters n, M, and N,
- so we wish to obtain $P(X = x) = h(x; n, M, N)$.

Example 3.35

- During a particular period a university's Information Technology office received 20 service orders for problems with printers, of which N
  - M 8 were laser printers and
  - n 12 were inkjet models.
- Sample of 5 of these service orders is to be selected for inclusion in customer satisfaction survey.
- Suppose that 5 are selected in completely random fashion, so that any particular subset of size 5 has the same chance of being selected as does any other subset.
- What is probability that exactly x (x = 0, 1, 2, 3, 4, or 5) of selected service orders were for inkjet printers?

Example 3.35

cont'd

- Here, population size is $N = 20$, sample size is $n = 5$, and number of S's (inkjet = S) and F's in population are $M = 12$ and $N - M = 8$, respectively.

ทำแบบไม่มีสูตร: $X=0 \rightarrow \frac{\binom{12}{0}\binom{8}{5}}{\binom{20}{5}}$

- Consider value $x = 2$.
- Because all outcomes (each consisting of 5 particular orders) are equally likely,

$X=1$ $\frac{\binom{12}{1}\binom{8}{4}}{\binom{20}{5}}$

sample size (n) population size (N)

$$P(X=2) = h(2; 5, 12, 20) = \frac{\text{number of outcomes having } X = 2}{\text{number of possible outcomes}}$$

x number of S's (M)

$x=2$ $\frac{\binom{12}{2}\binom{8}{3}}{\binom{20}{5}}$

$x=3$ $\frac{\binom{12}{3}\binom{8}{2}}{\binom{20}{5}}$

$x=4$ $\frac{\binom{12}{4}\binom{8}{1}}{\binom{20}{5}}$

$x=5$ $\frac{\binom{12}{5}\binom{8}{0}}{\binom{20}{5}}$

Example 3.35

cont'd

- Number of possible outcomes in experiment is number of ways of selecting 5 from 20 objects without regard to order

$$\text{No. of possible outcomes} = \binom{20}{5}$$

- To count number of outcomes having $X = 2$,

$$\text{There are } \binom{12}{2} \text{ ways of selecting 2 of inkjet orders}$$

- for each such way

$$\text{There are } \binom{8}{3} \text{ ways of selecting 3 laser orders to fill out sample}$$

- From product rule $\text{No. of outcomes with } X = 2 = \binom{12}{2}\binom{8}{3}$

$$P(X=2) = h(2; 5, 12, 20) = \frac{\text{number of outcomes having } X = 2}{\text{number of possible outcomes}} = \frac{\binom{12}{2}\binom{8}{3}}{\binom{20}{5}} = \frac{77}{323} = 0.238$$

Hypergeometric Distribution

- In general, if sample size n is smaller than number of successes in population (M), then the largest possible X value is n.

sample size population size

$$P(X = x) = h(x; n, M, N)$$

number of S's

$\boxed{n}$ $\boxed{M}$

- However, if $M < n$ (e.g., a sample size of 25 and only 15 successes in the population), then X can be at most M.

$$P(X = x) = h(x; n, M, N) = h(x; 25, 15, N)$$

Hypergeometric Distribution

sample size population size

$P(X = x) = h(x; n, M, N)$

number of S's

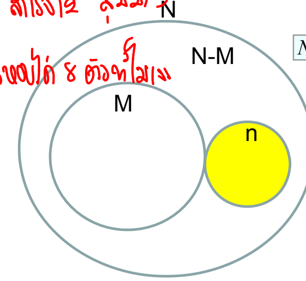
- Similarly, whenever number of population failures $(N - M)$ exceeds sample size (n), the smallest possible X value is 0 (since all sampled individuals might then be failures). $\Rightarrow$ $\boxed{N - M > n}$

ไปอยู่ในเซตที่ไม่สนใจหมดก็ได้ $x=0$

Ex. N=20, M=12, n=5 $\Rightarrow$ $\boxed{20 - 12 > 5}$ $\Rightarrow$ $\boxed{8 > 5}$

ประชากร=20 สำเร็จ12 สุ่มมา 5

มีโอกาสหยิบได้ 8 ตัวที่ไม่สน



$\boxed{N - M > n}$

smallest possible X value is 0

Hypergeometric Distribution

sample size · population size · number of S's

$P(X = x) = h(x; n, M, N)$

○ However, if $N - M < n$, the smallest possible X value is $n - (N - M)$.

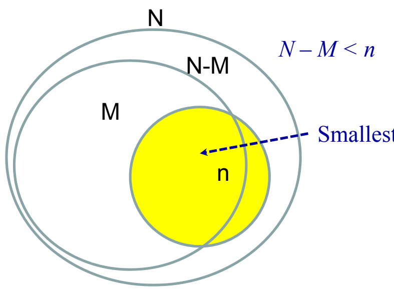
Ex. N=20, M=18, n=5 ⇨ $\boxed{20 - 18 < 5}$ ⇨ $\boxed{2 < 5}$

18 สำเร็จ 2 ไม่สำเร็จ

หยิบมา 5 ยังไงก็ต้องเจอสำเร็จ

เพราะมีของขอบเขตของชิ้น

smallest possible X value is $\boxed{n - (N - M) = 5 - (20 - 18) = 3}$



Smallest possible X value is $n - (N - M)$.

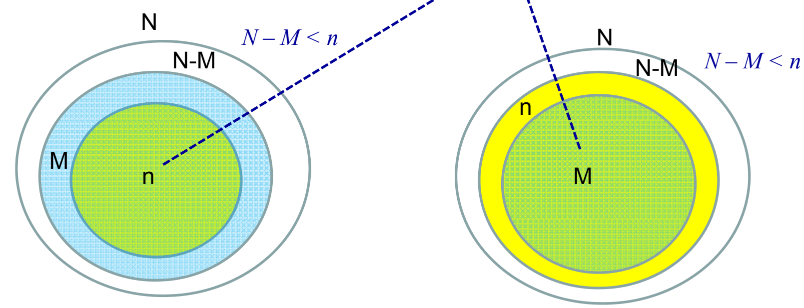
Hypergeometric Distribution

sample size · population size · number of S's

$P(X = x) = h(x; n, M, N)$

○ Thus, the possible values of X satisfy the restriction

$$\max(0,\ n - (N - M)) \le x \le \min(n, M).$$



○ An argument parallel to that of previous example gives pmf of X.

Hypergeometric Distribution

X = เหตุการณ์ที่สำเร็จที่สุ่มมา M จากประชากร

Proposition

If X is number of S's in a completely random sample of size n drawn from population consisting of M S's and $(N - M)$ F's, then probability distribution of X, called **Hypergeometric Distribution**, is given by

ที่ไม่สำเร็จ

Number of Outcomes having X=x

สำเร็จ

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad \textbf{(3.15)}$$

possible outcome

สูตรทั่วไป

Number of possible outcomes

for x, an integer, satisfying $\max(0,\ n - N + M) \le x \le \min(n, M)$.

Hypergeometric Distribution

In Example 3.35, $n = 5$, $M = 12$, and $N = 20$, so $h(x; 5, 12, 20)$ for $x = 0, 1, 2, 3, 4, 5$ can be obtained by substituting these numbers into Equation (3.15).

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$P(X = 0) = h(0; 5, 12, 20) = \frac{\binom{12}{0}\binom{20-12}{5-0}}{\binom{20}{5}} \qquad P(X = 3) = h(3; 5, 12, 20) = \frac{\binom{12}{3}\binom{20-12}{5-3}}{\binom{20}{5}}$$

$$P(X = 1) = h(1; 5, 12, 20) = \frac{\binom{12}{1}\binom{20-12}{5-1}}{\binom{20}{5}} \qquad P(X = 4) = h(4; 5, 12, 20) = \frac{\binom{12}{4}\binom{20-12}{5-4}}{\binom{20}{5}}$$

$$P(X = 2) = h(2; 5, 12, 20) = \frac{\binom{12}{2}\binom{20-12}{5-2}}{\binom{20}{5}} \qquad P(X = 5) = h(5; 5, 12, 20) = \frac{\binom{12}{5}\binom{20-12}{5-5}}{\binom{20}{5}}$$

Example 3.36

M

○ Five individuals from animal population thought to be near extinction in a certain region have been caught, tagged, and released to mix into population

n

○ After they have had an opportunity to mix, random sample of 10 of these animals is selected.

จำนวนสัตว์ที่ถูกติดแถบป้ายในการสุ่มตัวอย่างครั้งที่ 2

○ Let X = the number of tagged animals in the second sample

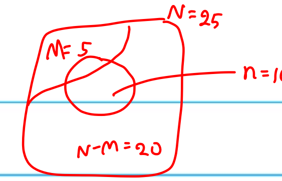
N

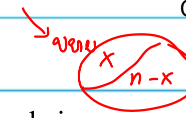
○ If there are actually 25 animals of this type in region,

○ what is probability that

a) X=2?

b) X<= 2?





Cont.

Example 3.36

○ Parameter values are $n = 10$, $M = 5$ (5 tagged animals in population), and $N = 25$, so

$$P(X=x)=h(x;n,M,N)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \Rightarrow P(X=x)=h(x;10,5,25)=\frac{\binom{5}{x}\binom{25-5}{10-x}}{\binom{25}{10}} \quad x = 0, 1, 2, 3, 4, 5$$

(a) *What is probability that $X = 2$?*

$$P(X=2)=h(2;10,5,25)=\frac{\binom{5}{2}\binom{25-5}{10-2}}{\binom{25}{10}}=\frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}}=0.385$$

(b) *What is probability that $X \le 2$?*

$$P(X\le 2)=P(X=0,1,or\ 2)=\sum_{x=0}^{2}h(x;10,5,25)=\frac{\binom{5}{0}\binom{20}{10-0}}{\binom{25}{10}}+\frac{\binom{5}{1}\binom{20}{10-1}}{\binom{25}{10}}+\frac{\binom{5}{2}\binom{20}{10-2}}{\binom{25}{10}}=\frac{\binom{5}{0}\binom{20}{10}}{\binom{25}{10}}+\frac{\binom{5}{1}\binom{20}{9}}{\binom{25}{10}}+\frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}}$$

$$=0.057+0.257+0.385=0.699$$

$E(X) = np$ ← pop of success

$V(X) = np(1-p)$ } binomial

ถ้า hype

$E(X) = n\left(\frac{M}{N}\right)$ ← ความสำเร็จ

$V(X) = n\left(\frac{M}{N}\right)\left(1-\frac{M}{N}\right)\cdots$

* p คือ pop ของความสำเร็จ

Hypergeometric Distribution

As in binomial case, there are simple expressions for $E(X)$ and $V(X)$ for hypergeometric random variable's.

Proposition

Mean and Variance of Hypergeometric rv X having pmf $h(x; n, M, N)$ are

$$E(X)=n\cdot\frac{M}{N} \qquad V(X)=\left(\frac{N-n}{N-1}\right)\cdot n\cdot\frac{M}{N}\cdot\left(1-\frac{M}{N}\right)$$

Ratio M/N is the proportion of S's in population.

If we replace M/N by p in $E(X)$ and $V(X)$, we get

$$E(X)=np \qquad V(X)=\left(\frac{N-n}{N-1}\right)\cdot np\cdot(1-p) \qquad \textbf{(3.16)}$$

Hypergeometric Distribution

Mean and Variance of Binomial Distribution	Mean and Variance of Hypergeometric Distribution
$E(X)=np$ $\quad V(X)=np(1-p)$	$E(X)=np$ $\quad V(X)=\left(\frac{N-n}{N-1}\right)\cdot np\cdot(1-p)$

○ Means of binomial and hypergeometric rv's are equal

○ Variances of the two rv's differ by the factor $(N - n)/(N - 1)$, often called **Finite Population Correction Factor.**

○ This factor is less than 1, $\frac{(N-n)}{(N-1)}<1$

so hypergeometric rv has smaller variance than does binomial rv.

○ The correction factor can be written $\frac{\left(1-\frac{n}{N}\right)}{1-\frac{1}{N}}$

, which is approximately 1 when n is small relative to N.

Example 3.37 (from example 3.36)

○ Five individuals from an animal population thought to be near extinction in a certain region have been caught, tagged, and released to mix into the population.

สัตว์ N ตัว มี 5 ตัวใกล้extinct

○ After they have had an opportunity to mix, a random sample of 10 of these animals is selected.

$n=10$

$0,1,2,3,4,5$

○ Let X = the number of tagged animals in second sample.

○If there are actually 25 animals of this type in the region,

$N=25$

what is the ***E*(*X*)** and ***V*(*X*)**?

$E(X) = np$

$= 10\left(\frac{5}{25}\right)$

$= 2$

$V(X) = \left(\frac{N-n}{N-1}\right) np(1-p)$

$=1$

Example 3.37

cont'd

In the animal-tagging example,

$n = 10$, $M = 5$, and $N = 25$, so $\frac{M}{N} = p = \frac{5}{25} = 0.2$

$$\boxed{E(X) = n \cdot \frac{M}{N}} \Rightarrow \boxed{E(X) = np} \Rightarrow \boxed{E(X) = 10(0.2) = 2}$$

$$\boxed{V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)} \Rightarrow \boxed{V(X) = \left(\frac{N-n}{N-1}\right) \cdot np \cdot (1-p)}$$

$$\begin{aligned} V(X) &= \left(\frac{25-10}{25-1}\right) \cdot 10(0.2) \cdot (1-0.2) \\ &= \left(\frac{15}{24}\right) \cdot 10(0.2) \cdot (0.8) \\ &= (0.625)(1.6) \\ &= 1 \end{aligned}$$

Example 3.37

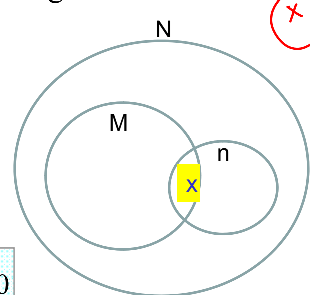
cont'd

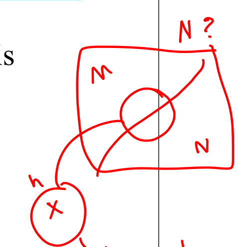
○ Suppose population size N is <u>not actually known</u>, so value x is observed and we wish to estimate N.

○ It is reasonable to equate the observed sample proportion of S's, x/n, with population proportion, M/N, giving the estimate

$$\boxed{\frac{x}{n} = \frac{M}{N}} \Rightarrow \boxed{\hat{N} = \frac{M \cdot n}{x}}$$

เหมือนเป็นอัตราส่วน





สามารถประมาณค่าของ N ได้

○ If $M = 100$, $n = 40$, and $x = 16$, then

$$\boxed{\hat{N} = \frac{M \cdot n}{x} = \frac{(100)(40)}{16} = 250}$$

Negative Binomial Distribution

สนใจจำนวนครั้งที่สำเร็จมากกว่า

คล้ายกับเรขาคณิต

ต้องการแค่ตัวเดียว เหมือนการหยิบ

battery

Negative Binomial Distribution

- Negative binomial random variable and distribution are based on experiment satisfying the following conditions:

1. Experiment consists of a sequence of independent trials.

2. Each trial can result in either a success (S) or a failure (F).

3. Probability of success is constant from trial to trial, so for $i = 1, 2, 3, \ldots$.

4. Experiment continues (trials are performed) until a total of r successes have been observed, where r is a specified positive integer.

Negative Binomial Distribution

- Random variable of interest is

จำนวนการทดลองที่ไม่สำเร็จก่อนหน้าความสำเร็จครั้งที่ r

$\boxed{X = \text{the number of failures that precede the } r^{th} \text{ success}}$

- X is called a **Negative Binomial Random Variable** because, in contrast to binomial random variable, number of successes is fixed and the number of trials is random.

Binomial Random Variable กำหนดจำนวนครั้งของการทดลอง (Trial) ที่แน่นอน
จำนวนครั้งของความสำเร็จขึ้นอยู่กับความสนใจ

Negative Binomial Distribution

จำนวนครั้งของความสำเร็จ

- Possible values of X are 0, 1, 2, ...
- Let $nb(x; r, p)$ denote pmf of X.
- Consider $nb(7; 3, p) = P(X = 7)$, probability that exactly 7 F's occur before the 3[rd] S.
- In order for this to happen, the 10[th] trial must be an S and there must be exactly 2 S's among the first 9 trials. Thus

$$nb(7;3,p) = \left\{\binom{9}{2}\cdot p^2(1-p)^7\right\}\cdot p = \binom{9}{2}\cdot p^3(1-p)^7 \Rightarrow \binom{\overset{x+r-1}{7+3-1}}{\underset{r-1}{3-1}}\cdot p^{\overset{r}{3}}(1-p)^{\overset{x}{7}}$$

- Generalizing this line of reasoning gives the following formula for negative binomial pmf.

$P(E) = \frac{n(E)}{n(S)}$

$n(E) = P(E)\,n(S)$

$\binom{x+(r-1)}{r-1} p^{r-1}(1-p)^x \cdot p$

Negative Binomial Distribution

Proposition

The pmf of negative binomial random variable X with parameters r = number of S's and $p = P(S)$ is

$$nb(x;r,p) = \binom{x+r-1}{r-1} p^r (1-p)^x \qquad x = 0, 1, 2, \ldots$$

จำนวนครั้งของความไม่สำเร็จ

จำนวนครั้งของความสำเร็จ

ความน่าจะเป็นของความสำเร็จ

Example 3.38

กุมารแพทย์ รับสมัคร/คัดเลือก/สรรหา

- A pediatrician wishes to recruit 5 couples, each of whom is expecting their first child, to participate in a new natural หลักการคลอดบุตรตามธรรมชาติ childbirth regimen.
- Let $p = P$(a randomly selected couple agrees to participate).
- If $p = 0.2$, what is **probability** that **15 couples** must be asked **before 5 are found** who agree to participate?

ตนที่ 14 สุดท้าย → สลับแบบนี้ 4 $\binom{14}{4}$

- That is, with S = {agrees to participate},
- what is **probability** that **10 F's occur before the fifth S**?
- Substituting $r = 5$, $p = 0.2$, and $x = 10$ into $nb(x; r, p)$ gives

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

$$nb(10; 5, 0.2) = \binom{14}{4}(0.2)^5(1-0.2)^{10} = 0.034$$

Example 3.38

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

cont'd

- Probability that **at most** 10 F's are observed (at most 15 couples are asked) is

$x = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$

$r = 5$

$p = 0.2$

binary ∈ Binomial

$$P(X \le 10) = \sum_{x=0}^{10} nb(x; 5, 0.2)$$
$$= (0.2)^5 \sum_{x=0}^{10} \binom{x+4}{4} (0.8)^x$$
$$= 0.164$$

ถ้าแทนตัวเลือก $\binom{n-1}{1}(1-p)$?

Negative Binomial Distribution

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

- In some sources, negative binomial random variable is taken to be the number of trials $X + r$ rather than the number of failures.
- In the special case $r = 1$, the pmf is

$$nb(x; 1, p) = p(1-p)^x \quad x = 0, 1, 2, \ldots \qquad \textbf{(3.18)}$$

- In Example 3.12 (slide 39), we derived pmf for the number of trials necessary to obtain the first S, and pmf there is similar to Expression (3.18).
- Both X = number of F's and Y = number of trials (= 1 + X) are referred to in the literature as **Geometric Random Variables**, and pmf in Expression (3.18) is called **Geometric Distribution.**

A B → B เกิดขึ้น → ไม่มีผลต่อ A

A B → ทำให้ P(A) change

Negative Binomial Distribution

Proposition

If X is a negative binomial random variable with pmf $nb(x; r, p)$, then

$$E(X) = \frac{r(1-p)}{p} \qquad V(X) = \frac{r(1-p)}{p^2}$$

จบบทของเรา ← binomial → รู้จำนวนการทดลองแน่ๆ, สุ่มแบบใส่คืน, อิสระต่อกัน, $\frac{n}{N} \le 5\%$ การทดลองเป็นอิสระต่อกัน, p(success) ต้องเท่ากัน] → ประมาณค่าความน่าจะเป็น

hyper geom → N, n, x head, n−x

exact pop model → กำหนดที่ชัดเจน

neg -bi → จำนวนครั้งไม่แน่นอน ทำจนสำเร็จครบ r, เปลี่ยน ตัวแปร

อันเก่าล้วนเกิดขึ้นจากการทดลอง → ดูกราฟ

3.6 Poisson Probability Distribution

Poisson Probability Distribution

○ Binomial, Hypergeometric, and Negative Binomial Distributions were all derived by starting with experiment consisting of trials or draws and applying laws of probability to various outcomes of experiment.

○ There is no simple experiment on which Poisson distribution is based, though we will shortly describe how it can be obtained by certain limiting operations.

○ In contrast to Binomial and Hypergeometric Distributions, Poisson Distribution spreads probability over *all* non-negative integers, an infinite number of possibilities.

Poisson Probability Distribution

Definition

Discrete random variable ***X*** is said to have **Poisson Distribution** with parameter λ ($\lambda > 0$) if **pmf of *X*** is

e represents base of natural logarithm system; its numerical value is approximately 2.71828.

$$p(x;\lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \qquad x = 0, 1, 2, 3, \ldots$$

อัตรา / หน่วย

○ Value of λ is frequently rate per unit time or per unit area

○ Because λ must be positive; $p(x; \lambda) > 0$ for all possible x values

Poisson Probability Distribution

○ The fact that $\Sigma\, p(x; \lambda) = 1$ is consequence of the Maclaurin series expansion of e^{λ}, which appears in most calculus texts :

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \qquad (3.19)$$

○ If the two extreme terms in (3.19) are multiplied by $e^{-\lambda}$ and then this quantity is moved inside the summation on the far right, the result is

$$1 = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} \quad \Rightarrow \quad \sum_{x=0}^{\infty} p(x;\lambda) = 1$$

○ which shows that $p(x; \lambda)$ fulfills the second condition necessary for specifying pmf

ต่อไป

Example 3.39

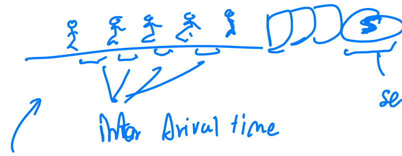
$$p(x;\lambda) = \frac{e^{-\lambda}\cdot\lambda^x}{x!} \qquad x = 0, 1, 2, 3, \ldots$$

- Let X denote the number of creatures of a particular type captured in a trap during a given time period.
- Suppose that X has Poisson distribution with $\lambda = 4.5$, so on average traps will contain 4.5 creatures.
- Probability that a trap contains **exactly five** creatures is

$$P(X=5) = p(5; 4.5) = \frac{e^{-4.5}(4.5)^5}{5!} = 0.1708$$

- Probability that a trap has **at most five** creatures is

$$P(X \le 5) = \sum_{x=0}^{\infty} p(x; 4.5) = \sum_{x=0}^{5} \frac{e^{-4.5}(4.5)^x}{x!}$$
$$= e^{-4.5}\left[1 + 4.5 + \frac{(4.5)^2}{2!} + \frac{(4.5)^3}{3!} + \frac{(4.5)^4}{4!} + \frac{(4.5)^5}{5!}\right]$$
$$= 0.7029$$



การแจกแจงแบบ exponential

ค่าอะไรสักอย่าง

Poisson Distribution as a Limit

ใช้

binomial → รู้ขนาดตัวอย่าง , p(success)

np = λ ในครั้งของ

Poisson Distribution as a Limit

- The rationale for using Poisson distribution in many situations is provided by the following proposition.

Proposition

binomial สามารถใช้ poisson ประมาณ ค่าได้

- Suppose that in Binomial pmf $b(x; n, p)$,
- we let $n \to \infty$ and $p \to 0$ in such a way that np approaches a value $\lambda > 0$.
- Then $\boldsymbol{b(x; n, p) \to p(x; \lambda)}$.

- According to this proposition, *in any binomial experiment in which* ***n* is large** *and* ***p* is small**, $\boldsymbol{b(x; n, p) \approx p(x; \lambda)}$,
- *where* $\boldsymbol{\lambda = np}$. จำเลขนี้
- As a rule of thumb, this approximation can safely be applied if **$n > 50$ and $np < 5$**.

ค่าเฉลี่ย

Example 3.40

- If publisher of nontechnical books takes great pains to ensure that its books are free of typographical errors,
- so that probability of any given page containing **at least one** such **error** is 0.005 and
- errors are independent from page to page,
- what is probability that one of its 400-page novels will contain
  1) **exactly** one page with errors?
  2) **at most** three pages with errors?

- With S denoting page containing at least one error and F an error-free page,
- the number X of pages containing at least one error is a binomial random variable with $n = 400$ and $p = 0.005$, so

$$\lambda = np = (400)(0.005) = 2$$

ก็ใช้ binomial $\binom{400}{1}(0.005)^1(1-0.005)^{400-1}$; [illegible]

Example 3.40

$p(x;\lambda) = \frac{e^{-\lambda}\cdot\lambda^x}{x!} \quad x = 0, 1, 2, 3, \ldots$

cont'd

What is probability that one of its 400-page novels will contain **exactly** one page with errors?

$$P(X = 1) = b(1; 400, 0.005) \approx p(1; 2) = \frac{e^{-2}(2)^1}{1!} = 0.270671$$

Binomial value is $b(1; 400, 0.005) = 0.270669$, so approximation is very good.

Example 3.40

$p(x;\lambda) = \frac{e^{-\lambda}\cdot\lambda^x}{x!} \quad x = 0, 1, 2, 3, \ldots$

cont'd

What is probability that one of its 400-page novels will contain **at most** three pages with errors?

$$P(X \le 3) = P(X = 0, 1, 2, or, 3)$$
$$\approx \sum_{x=0}^{3} p(x, 2) = \sum_{x=0}^{3} \frac{e^{-2}\cdot 2^x}{x!}$$
$$= \frac{e^{-2}\cdot 2(0)}{0!} + \frac{e^{-2}\cdot 2(1)}{1!} + \frac{e^{-2}\cdot 2(2)}{2!} + \frac{e^{-2}\cdot 2(3)}{3!}$$

= 0.135335 + 0.270671 + 0.270671 +0 .180447
= 0.8571

and this again is quite close to the Binomial value

$P(X \le 3) = 0.8576$

binomial → รู้แต่ตัวอย่าง

Poisson Distribution as a Limit

Table 3.2 shows Poisson distribution for **$\lambda = 3$** along with three binomial distributions with ***np* = 3**

x	$n = 30, p = .1$	$n = 100, p = .03$	$n = 300, p = .01$	Poisson, $\lambda = 3$
0	0.042391	0.047553	0.049041	0.049787
1	0.141304	0.147070	0.148609	0.149361
2	0.227656	0.225153	0.224414	0.224042
3	0.236088	0.227474	0.225170	0.224042
4	0.177066	0.170606	0.168877	0.168031
5	0.102305	0.101308	0.100985	0.100819
6	0.047363	0.049610	0.050153	0.050409
7	0.018043	0.020604	0.021277	0.021604
8	0.005764	0.007408	0.007871	0.008102
9	0.001565	0.002342	0.002580	0.002701
10	0.000365	0.000659	0.000758	0.000810

Table 3.2 Comparing the Poisson and Three Binomial Distributions

Poisson Distribution as a Limit

- Figure 3.8 (from S-Plus) plots Poisson along with first two Binomial Distributions.
- Approximation is of limited use for $n = 30$, the accuracy is better for $n = 100$ and much better for $n = 300$.

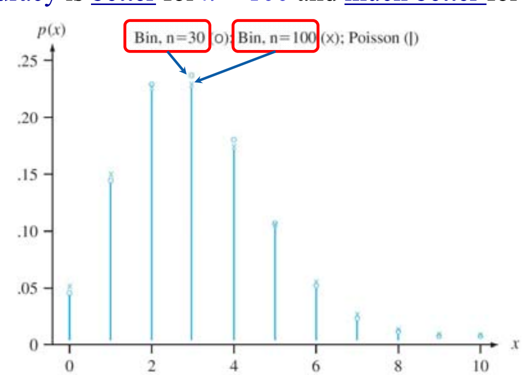


Figure 3.8 Comparing a Poisson and two binomial distributions

Poisson Distribution as a Limit

Appendix Table A.2 exhibits cdf $F(x; \lambda)$ for $\lambda = 0.1, 0.2, \ldots, 1, 2, \ldots, 10, 15$, and 20.

Table A.2 Cumulative Poisson Probabilities

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^y}{y!}$$

	λ									
x	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368
1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736
2	1.000	.999	.996	.992	.986	.977	.966	.953	.937	.920
3		1.000	1.000	.999	.998	.997	.994	.991	.987	.981
4				1.000	1.000	1.000	.999	.999	.998	.996
5							1.000	1.000	1.000	.999
6										1.000

Poisson Distribution as a Limit

Table A.2 Cumulative Poisson Probabilities *(cont.)*

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^y}{y!}$$

	λ										
x	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	15.0	20.0
0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
6	.995	.966	.889	.762	.606	.450	.313	.207	.130	.008	.000
7	.999	.988	.949	.867	.744	.599	.453	.324	.220	.018	.001
8	1.000	.996	.979	.932	.847	.729	.593	.456	.333	.037	.002
9		.999	.992	.968	.916	.830	.717	.587	.458	.070	.005
10		1.000	.997	.986	.957	.901	.816	.706	.583	.118	.011
11			.999	.995	.980	.947	.888	.803	.697	.185	.021
12			1.000	.998	.991	.973	.936	.876	.792	.268	.039
13				.999	.996	.987	.966	.926	.864	.363	.066
14				1.000	.999	.994	.983	.959	.917	.466	.105
15					.999	.998	.992	.978	.951	.568	.157
16					1.000	.999	.996	.989	.973	.664	.221
17						1.000	.998	.995	.986	.749	.297
18							.999	.998	.993	.819	.381
19							1.000	.999	.997	.875	.470
20								1.000	.998	.917	.559
21									.999	.947	.644
22									1.000	.967	.721
23										.981	.787
24										.989	.843
25										.994	.888
26										.997	.922
27										.998	.948
28										.999	.966
29										1.000	.978
30											.987
31											.992
32											.995
33											.997
34											.999
35											.999
36											1.000

Poisson Distribution as a Limit

For example, if $\lambda = 2$, then $P(X \leq 3) = F(3; 2) = 0.857$ as in example 3.40, whereas $P(X = 3) = F(3; 2) - F(2; 2) = 0.857 - 0.677 = 0.180$.

Table A.2 Cumulative Poisson Probabilities *(cont.)*

$$F(x; \lambda) = \sum_{y=0}^{x} \frac{e^{-\lambda}\lambda^y}{y!}$$

	λ										
x	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	15.0	20.0
0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
6	.995	.966	.889	.762	.606	.450	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Alternatively, many statistical computer packages will generate $p(x; \lambda)$ and $F(x; \lambda)$ upon request.

Mean and Variance of X

Mean and Variance of X

- Since $b(x; n, p) \to p(x; \lambda)$ as $n \to \infty$, $p \to 0$, $np \to \lambda$,
- Mean and Variance of binomial variable should approach those of Poisson variable.
- These limits are $np \to \lambda$ and $np(1 - p) \to \lambda$.

Proposition

If X has a Poisson distribution with parameter λ, then

$$\boxed{E(X) = V(X) = \lambda}$$

These results can also be derived directly from the definitions of mean and variance.

Example 3.41

Example 3.39 continued…

Both **expected** number of creatures trapped and **variance** of the number trapped equal **4.5**, and

$$\boxed{E(X) = V(X) = \lambda}$$

Standard deviation (SD)

$$\begin{aligned} \sigma_X &= \sqrt{\lambda} \\ &= \sqrt{4.5} \\ &= 2.12 \end{aligned}$$

Poisson Process

Poisson Process

- A very important application of Poisson distribution arises in connection with **occurrence of events** of some type **over time**.

เหตุการณ์ที่เกิดขึ้นในช่วงเวลา

- Events of interest might be
  - visits to a particular website,
  - pulses of some sort recorded by counter,
  - email messages sent to a particular address,
  - accidents in an industrial facility, or
  - cosmic ray showers observed by astronomers at a particular observatory.

Poisson Process

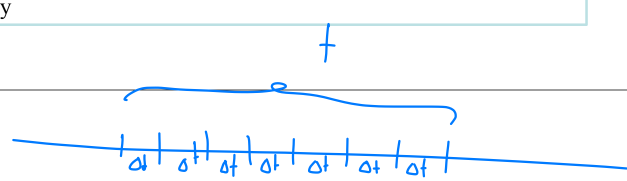
We make following assumptions about the way in which events of interest occur:

1. There exists parameter $\alpha > 0$ such that for any short time interval of length Δt,
   probability that exactly one event occurs is received is $\alpha \cdot \Delta t + o(\Delta t)$*

2. Probability of more than one event occurring during Δt is $\boldsymbol{o(\Delta t)}$
[which, along with Assumption 1, implies that
probability of no events during Δt is $1 - \alpha \cdot \Delta t - o(\Delta t)$]

3. The number of events occurring during the time interval $\boldsymbol{\Delta t}$ is independent of the number that occur prior to this time interval.

- * Quantity is $o(\Delta t)$ (read "little o of delta t" is as Δt *approaches* 0, so does $o(\Delta t)/\Delta t$
- That is, $o(\Delta t)$ is even more negligible (approaches 0 faster) than Δt itself.
- The $(\Delta t)^2$ has this property



Poisson Process

- Informally, Assumption 1 says that for a short interval of time, probability of receiving single event (occurring) is approximately proportional to the length of time interval,

  where α is the constant of proportionality.

- Now let $\boldsymbol{P_k(t)}$ denote probability that $\boldsymbol{k}$ events will be observed during any particular time interval of length $\boldsymbol{t}$.

Poisson Process

Proposition

$$P_k(t) = e^{-\alpha t} \cdot \frac{(\alpha t)^k}{k!}$$

- so that the number of events during time interval of length t is Poisson random variable with parameter $\lambda = \alpha t$.
- Expected number of events during any such time interval is then $\boldsymbol{\alpha t}$, so
- expected number during a unit interval of time is α.

- **Occurrence of events over time** as described is called a ***Poisson Process***; **parameter $\boldsymbol{\alpha}$** specifies ***rate*** for process.

Example 42

- Suppose pulses arrive at counter at average rate of **six per minute**, so that $\alpha = 6$.

- To find probability that in 0.5-min interval **at least** one pulse is received,
- note that the number of pulses in an interval has a Poisson distribution with parameter $\alpha t = 6(0.5) = 3$
(0.5 min is used because α is expressed as a rate per minute).

- Then with $\boldsymbol{X}$ = the number of pulses received in **30-sec** (0.5-min) interval,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{e^{-3}(3)^0}{0!} = 0.950$$