STAT230

PROBABILITY

Chapter 5 (a)

UNIVERSITY OF WATERLOO

Please Do Problems 5.1-5.21

Discrete Random Variables and Probability Models
Chapter 5 Objectives

- Know the definition of “Random Variable”.

- Model distributions

  - Discrete Uniform Distribution.
  - The Hypergeometric Distribution.
  - The Binomial Distribution.
  - The Negative Binomial Distribution.
  - Poisson Distribution.

UNIVERSITY OF WATERLOO

Random Variable

- When an experiment is performed, we are interested mainly in some function of the outcome as opposed to the actual outcome itself.

Example

Tossing dice, we are often interested in the sum of the two dice equal to 7 and are not really concerned about the separate values of each die,

{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), or (6, 1)}.

UNIVERSITY OF WATERLOO

Example

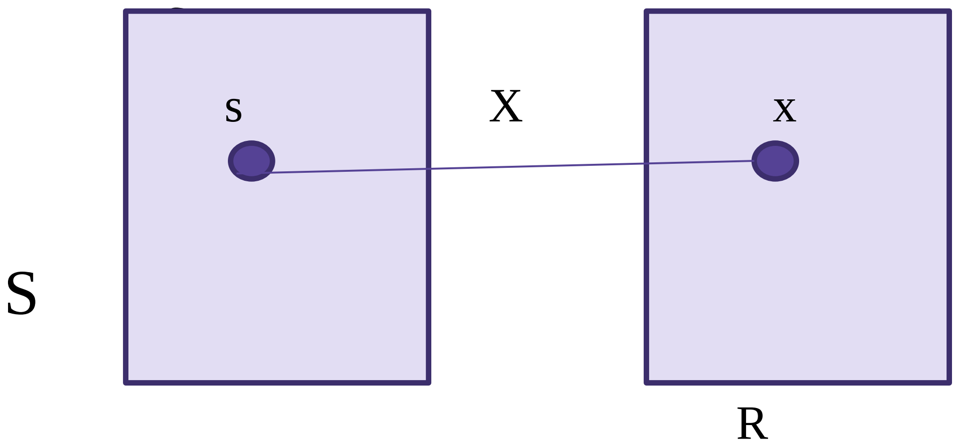
Flipping a coin, we may be interested in the total number of heads that occur and not care at all about the actual head–tail sequence that results.

http://www.studyzone.org/mtestprep/math8/e/BD05036_.gif

UNIVERSITY OF WATERLOO

- Random Variable: is a function that assigns a real number to each point in a sample space *S*.

- In mathematical language, a random variable is a function whose domain is the sample space and whose range is the set of possible values of the variable.



- The abbreviation of random variable is r.v.
- Random variables are denoted by uppercase near the end of our alphabet such as X, Y, Z,...
- Lowercase letters represent some particular value of the corresponding r.v. such as, x, y, z,...
- X(s)= x means that x is the value of the outcome s of the r.v. X.

Example

- Suppose that our experiment consists of tossing 3 fair coins.
- The sample space is

S={ HHH, THH, HTH, HHT, HTT, THT,TTH, TTT}

- Let Y denote the number of heads that appear.
- The range of the r.v Y is

$$\underset{\text{range}}{\underset{\uparrow}{A}} = \{0, 1, 2, 3\}$$

Y is a random variable taking on one of the values 0, 1, 2, and 3 with respective probabilities

P(Y = 0) = P({ TTT}) = 1/8

P(Y = 1) = P({TTH, THT, HTT }) = 3/8

P(Y = 2) = P({THH, HTH, HHT}) = 3/8

P(Y = 3) = P({HHH}) = 1/8

Two Types of Random Variables

- Discrete random variable is a r.v. whose possible values either integer or countable set in which there is a first element, a second element, and so on.

- Continuous random variable is a r.v. whose set of possible values consists of an entire interval on the number line.

Example

When a student attempts to log on a computer time-sharing system, either all ports are busy (F), or at least one port free(S),

The sample space is

$$\mathbf{S} = \{S, F\}$$

The random variable X will be

$$X(S) = 1, \quad X(F) = 0$$

Note:

- The r.v. X was specified by explicitly listing each element of S
- Any random variable whose only possible values are 0 and 1 is called a Bernoulli random variable.

Example (1)

Three automobiles are selected at random, and each is categorized as having a diesel (S) or non-diesel (F) engine.

If X is the <u>number</u> of cars among the three <u>with diesel engines.</u>

What is the sample space and its associated X values?

Sample space = {SSS, SSF, SFS, FSS, SFF, FSF, FFS, FFF}

$$X = \begin{cases} 0 & \text{for FFF} \\ 1 & \text{for SFF, FSF, FFS} \\ 2 & \text{for SSF, SFS, FSS} \\ 3 & \text{for SSS} \end{cases}$$

or it can be written as

X(FFF)=0, X(SFF)=1, X(FSF)=1,......, X(SSS)=3

Probability Function (p.*f*.) or (p.m.*f*)

- Probability Function of a random variable, X is a function

$$f\,(\mathrm{x}) = P(X = \mathrm{x}) \text{ defined for all } \mathrm{x} \in A$$

UNIVERSITY OF WATERLOO

Probability Distributions for Discrete Random Variables

- A **probability distributions** of a random variable X is a description of the probabilities associated with the possible values of X.

- The set of pairs $\{(x, f(x)) : x \in A\}$ is called the probability distribution of X.

- A **probability distributions** says how the total probability of 1 is distributed among the various value of X.

Properties of a Probability Function

$$(1) \quad f(x) \geq 0 \text{ for all } x \in A$$

$$(2) \quad \sum_{x \in A} f(x) = 1$$

- The p. *f*. can be presented nicely in tabular.
- The p. *f*. can also be displayed in line graph.
- The p. *f*. can also be displayed using histogram which is called probability histogram.
  - -The height of each rectangle is proportional to *f* (x).
  - -The base is the same for all rectangles.
- A p. *f*. may be specified as a formula.

Note:

It is not always possible to find a simple formula.

UNIVERSITY OF WATERLOO

Example(2)

For example (1)

a) Calculate the probabilities associated with each value of X?

b) Present the results using table.

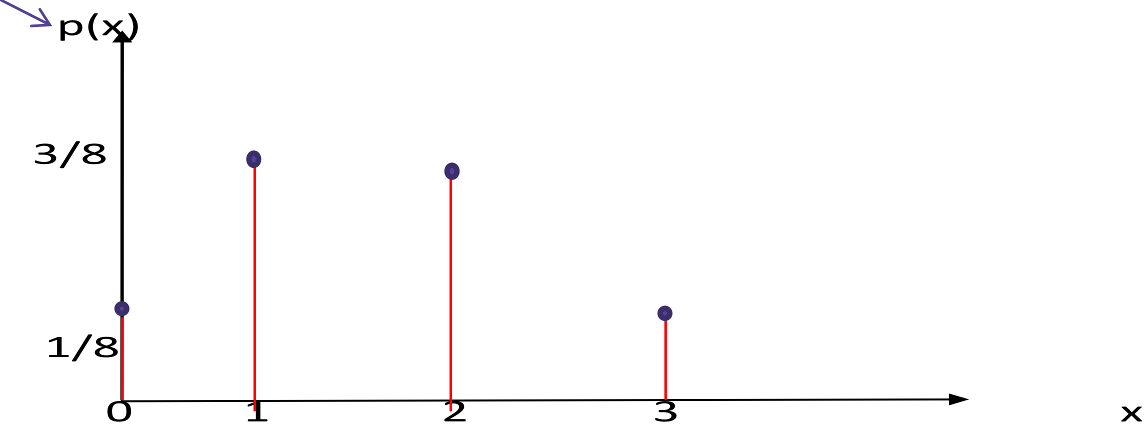
c) Display the result in line graph.

Sample space ={ SSS, SSF, SFS, FSS, SFF, FSF, FFS, FFF}

$P(X=0) = f(0) = 1/8$
$P(X=1) = f(1) = 3/8$
$P(X=2) = f(2) = 3/8$
$P(X=3) = f(3) = 1/8$

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

$f(x)$



Example

A random variable X that can take on values 1, 2, 3, 4 and 5 according to the following probability function:

$$f(\mathrm{x}) = (kx)/(2+x)$$

a) Determine the value of k.

b) And hence calculate the $f(2) = P(X=2)$.

$$\sum_{x\in A} f(x) = 1$$

$$\sum_{x=1}^{5} f(x) = \sum_{x=1}^{5} (kx)/(2+x) = 1$$

$$1 = k(1)/(2+1) + k(2)/(2+2)+\ldots+k(5)/(2+5)$$

$$1 = 1182\,k/420$$

$$k = 420/1182 = 0.355$$

$$f(x) = \begin{cases} (0.355x)/(2+x) & \text{for } x=1,2,..,5 \\ 0 & \text{o/w} \end{cases}$$

UNIVERSITY OF WATERLOO

(b) $f(2) = P(X=2)$

$f(\mathrm{x}) = (0.355x)/(2+x)$

$f(2) = (\ 0.355(2)\ )\ /(2+2)$

$f(2) = \quad 0.1775$

Example

The probability function of a random variable X is given by

$$f(x) = \frac{c\ \lambda^x}{x!} \qquad for\ x = 0, 1, 2, 3, \ldots$$

Where λ is some positive value.

a) Determine the value of c.

b) Find P(X= 0).

c) Find P(X > 2).

a) $$f(x) = \frac{c\ \lambda^x}{x!} \qquad for\ x = 0, 1, 2, 3, \ldots$$

$$\sum_{all\ x} f(x) = \sum_{x=0}^{\infty} \frac{c\ \lambda^x}{x!} = 1$$

$$\sum_{all\ x} f(x) = c \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = 1$$

$$c\ e^{\lambda} = 1$$

$$c = \frac{1}{e^{\lambda}}$$

$$f(x) = \frac{\frac{1}{e^{\lambda}}\ \lambda^x}{x!} \quad x = 0, 1, 2, \ldots$$

Exponential Series

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

UNIVERSITY OF WATERLOO

b) P(X = 0)

$$f(x) = \frac{\frac{1}{e^{\lambda}}\ \lambda^{x}}{x!} \quad x = 0, 1, 2, \dots$$

$$f(0) = P(X = 0) = \frac{e^{-\lambda}\ \lambda^{0}}{0!} = e^{-\lambda}$$

c) $P(X > 2) = 1 - P(x \le 2)$

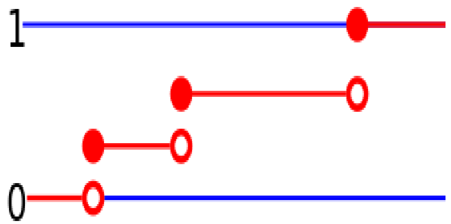
$$= 1 - \{ P(x = 0) + P(x = 1) + P(x = 2) \}$$

$$f(x) = \frac{\frac{1}{e^{\lambda}} \lambda^x}{x!} \quad x = 0, 1, 2, \ldots$$

$$P(X > 2) = 1 - \{ e^{-\lambda}\lambda^0 / 0! + e^{-\lambda}\lambda^1 / 1! + e^{-\lambda}\lambda^2 / 2! \}$$
$$= 1 - \{e^{-\lambda} + e^{-\lambda}\lambda + e^{-\lambda}\lambda^2 / 2\}$$
$$= 1 - \{e^{-\lambda} (1 + \lambda + \lambda^2 / 2)\}$$

UNIVERSITY OF WATERLOO

The Cumulative Distribution function (***c.d.f***)



F(x) of discrete r.v. for variable X with p. *f.*
*f (*x) is defined for every number x ∈ R by

$$F(x) = P(X \le x) = \sum_{u:\, u \le x} f(u)$$

Properties of the Cumulative Distribution Function

Let

F(x) denotes the probability that the random variable X takes on a value that is less than or equal to x.

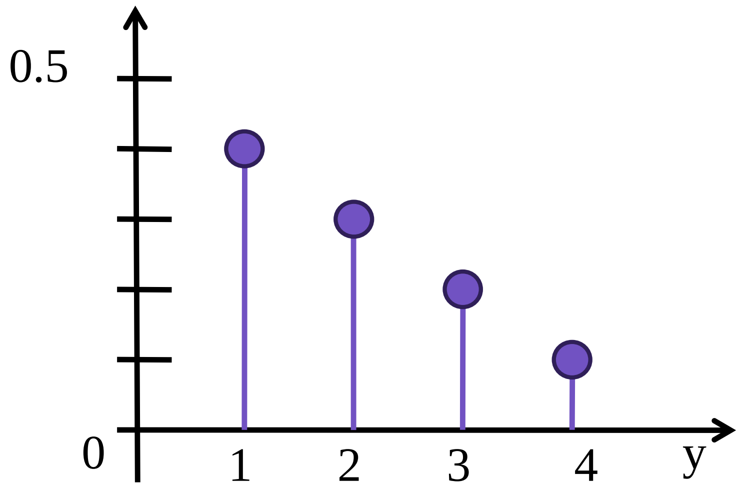
- F is a nondecreasing function; that is, if $a < b$, then $F(a) \leq F(b)$.
- $0 \leq F(x) \leq 1$, for all $x \in R$
- $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$.

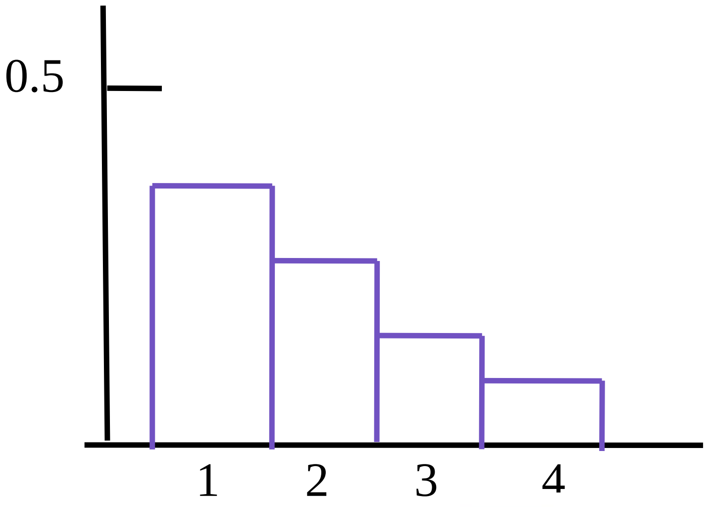
UNIVERSITY OF WATERLOO

Example (3)

Line Graph and Probability Histogram

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





What is the F(y) for the p. *f*. of Y

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

$F(1) = P(Y \leq 1) = P(Y=1) = f(1) = 0.4$

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

$F(3) = P(Y \leq 3) = P(Y=1 \text{ or } 2 \text{ or } 3) = f(1) + f(2) + f(3)$

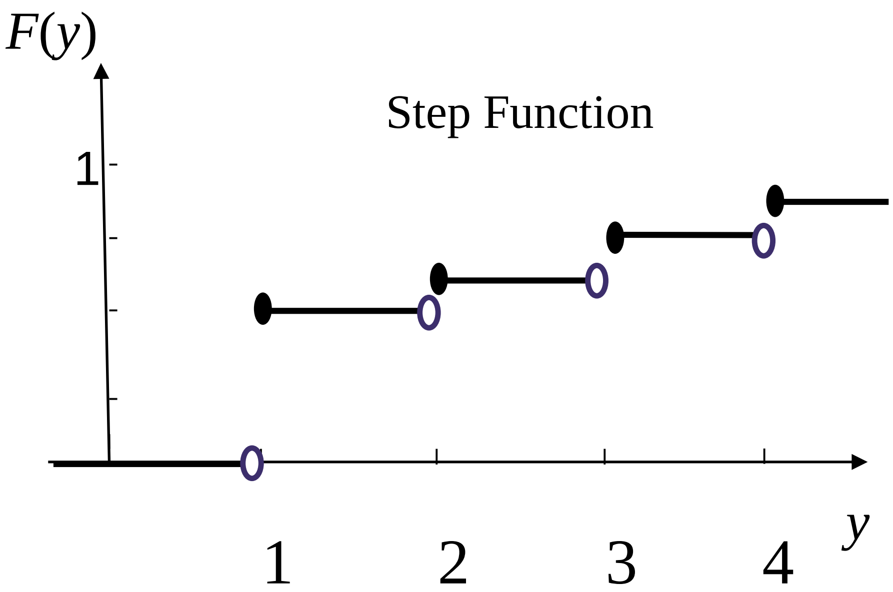
$= 0.4+0.3+0.2 = 0.9$

$F(4) = P(Y \leq 4) = P(Y=1 \text{ or } 2 \text{ or } 3 \text{ or } 4)$

$= f(1) + f(2) + f(3) + f(4)$

$= 0.4+0.3+0.2+0.1 = 1$

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



y	1	2	3	4
f (y)	0.4	0.3	0.2	0.1
F(y)	0.4	0.7	0.9	1

Note: 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.

Example $F(2.7) = P(Y \leq 2.7)$

$= P(Y \leq 2) = F(2) = 0.7$

Example $F(1.9) = P(Y \leq 1.9)$

$= P(Y \leq 1) = F(1) = 0.4$

- Proposition

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

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

(a-) the largest possible X value that is strictly less than a.

Examples

- $P(3 \leq x \leq 6) = P(x = 3, 4, 5, 6) = F(6) - F(2)$

- $P(3 < x \leq 6) = F(6) - F(3)$

- $P(X=3) = F(3) - F(2)$

Note

- $F(x)$ can be obtained from $f(x)$.

- The opposite is also true

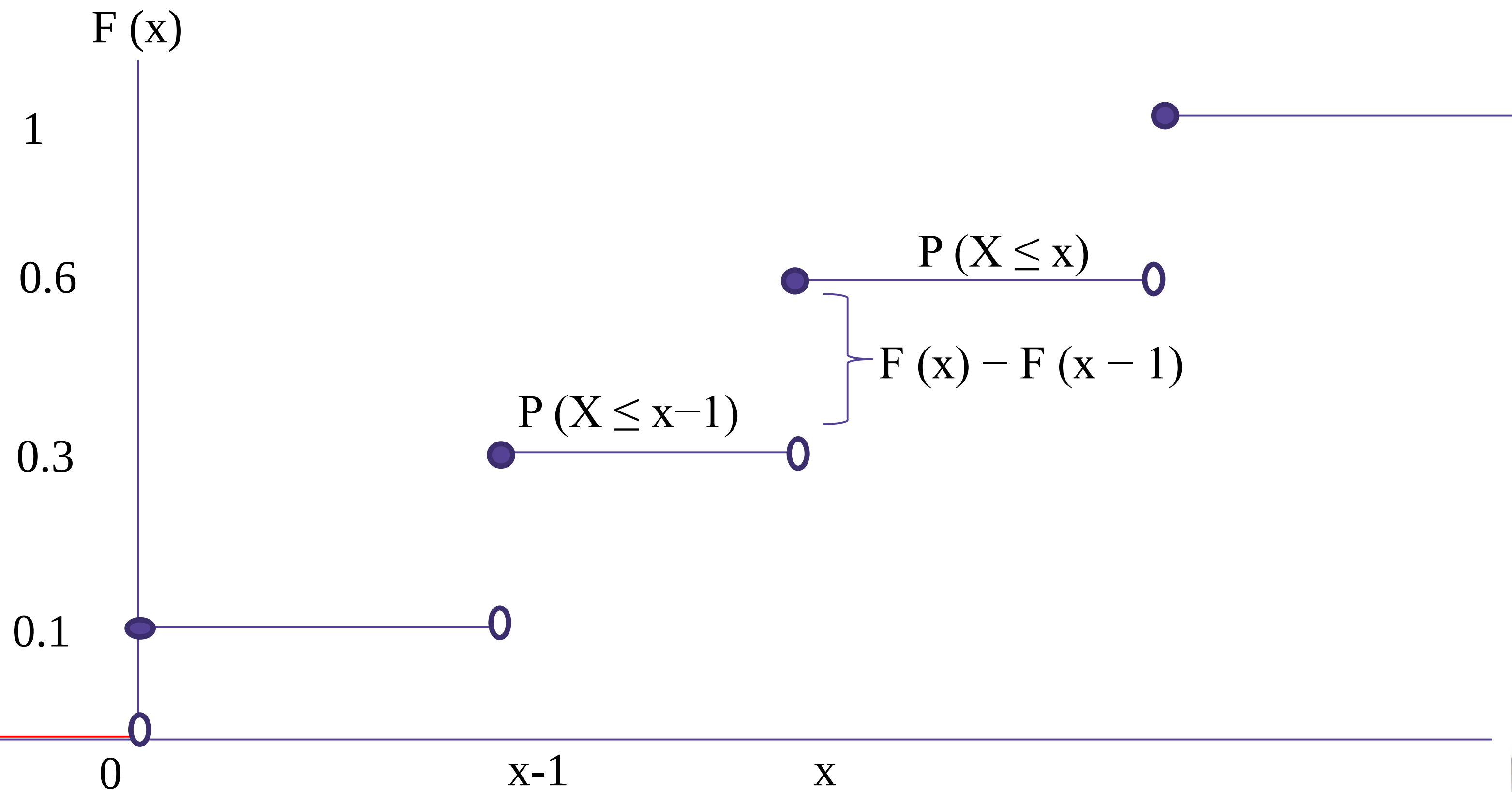
- If X takes on integer values then for values x such that $x \in A$ and $x - 1 \in A$,

$$f(x) = F(x) - F(x - 1)$$

- This says that $f(x)$ is the size of the jump in $F(x)$ at the point x.

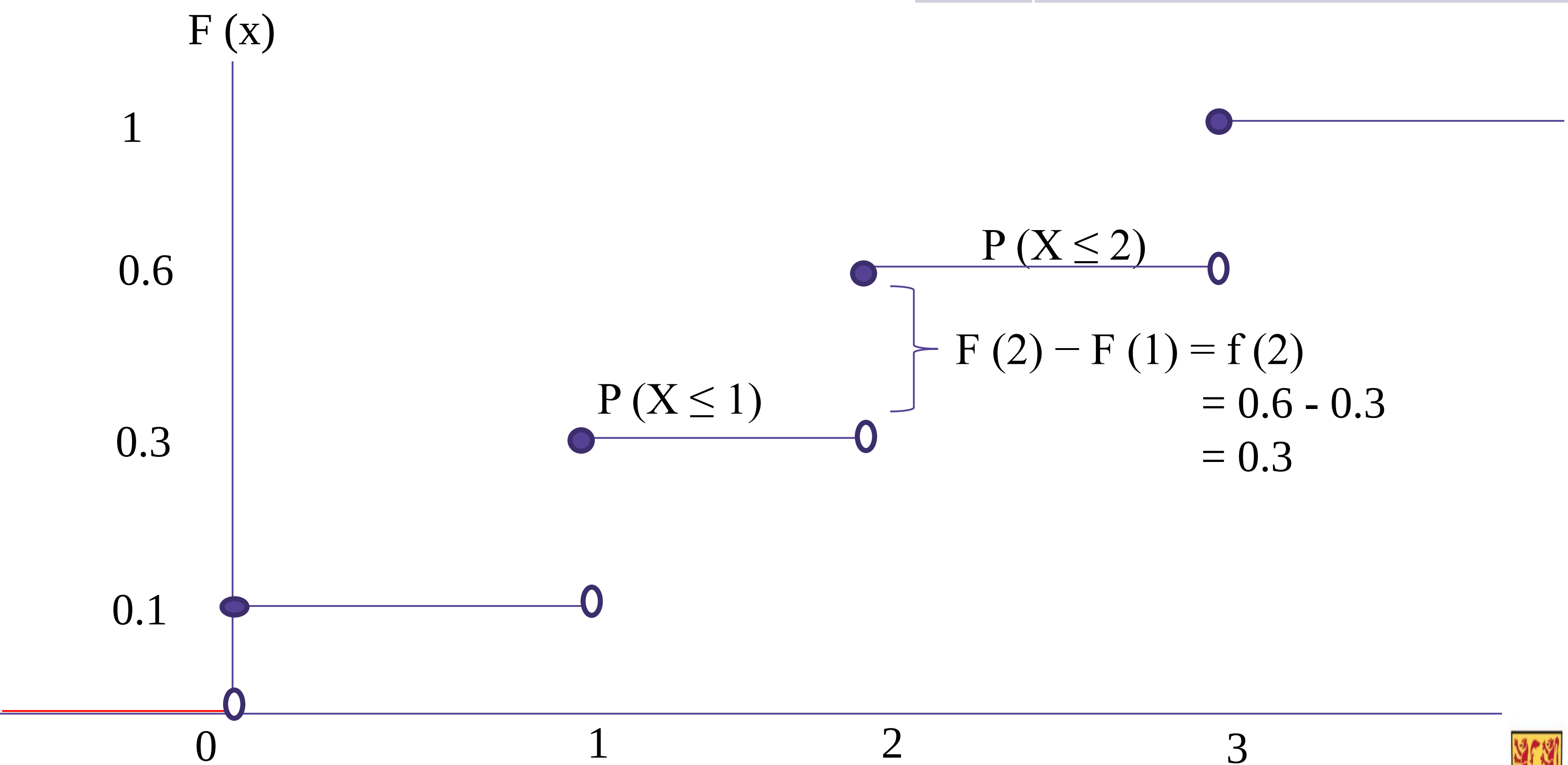
To prove this, just note that

$$F(x) - F(x-1) = P(X \le x) - P(X \le x-1) = P(X = x)$$



Example

y	0	1	2	3
$f(y)$	0.1	0.2	0.3	0.4
F(y)	0.1	0.3	0.6	1



Example

Suppose that N balls labelled 1, 2, . . . , N are placed in a box, and n balls ($n \leq N$) are randomly selected without replacement.

Define the r.v.

$$X = \text{largest number selected}$$

First find the c.d.f, $F(x)$ and then find the p. *f*. , $f(x)$ of X.

First find $F(x) = P(X \le x)$. Noting that $X \le x$ if and only if all n balls selected are from the set {1, 2, . . . , x}, we get

$$F(x) = \frac{\binom{x}{n}}{\binom{N}{n}} \quad \text{for } x = n, n+1, \ldots N$$

We can now find

$$f(x) = F(x) - F(x-1)$$

$$= \frac{\binom{x}{n} - \binom{x-1}{n}}{\binom{N}{n}} = \frac{\binom{x-1}{n-1}}{\binom{N}{n}}$$

UNIVERSITY OF WATERLOO

$$f(x) = \begin{cases} \dfrac{\binom{x-1}{n-1}}{\binom{N}{n}} & \text{for } x = n, n+1, \ldots N \\ 0 & \text{o/w} \end{cases}$$

Examples of Discrete Probability Distributions:

- Discrete Uniform Distribution.
- The Hypergeometric Distribution.
- The Binomial Distributions.
- The Negative Binomial Distribution.
- Poisson Distributions.

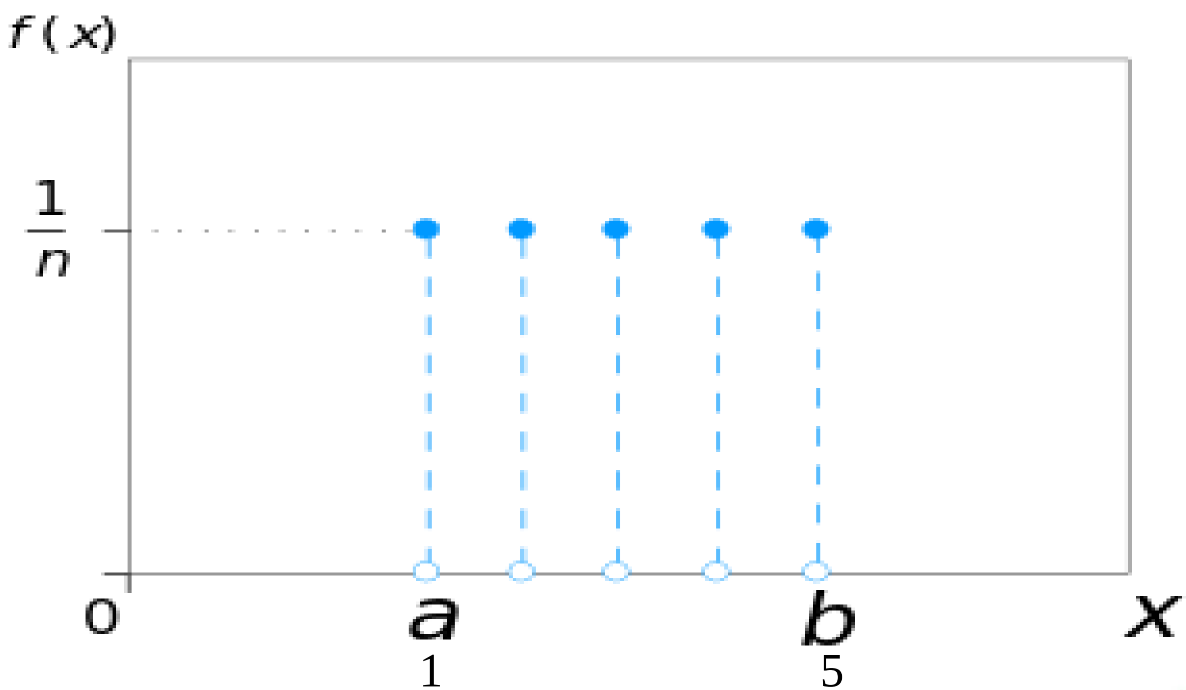
UNIVERSITY OF WATERLOO

Discrete Uniform Distribution

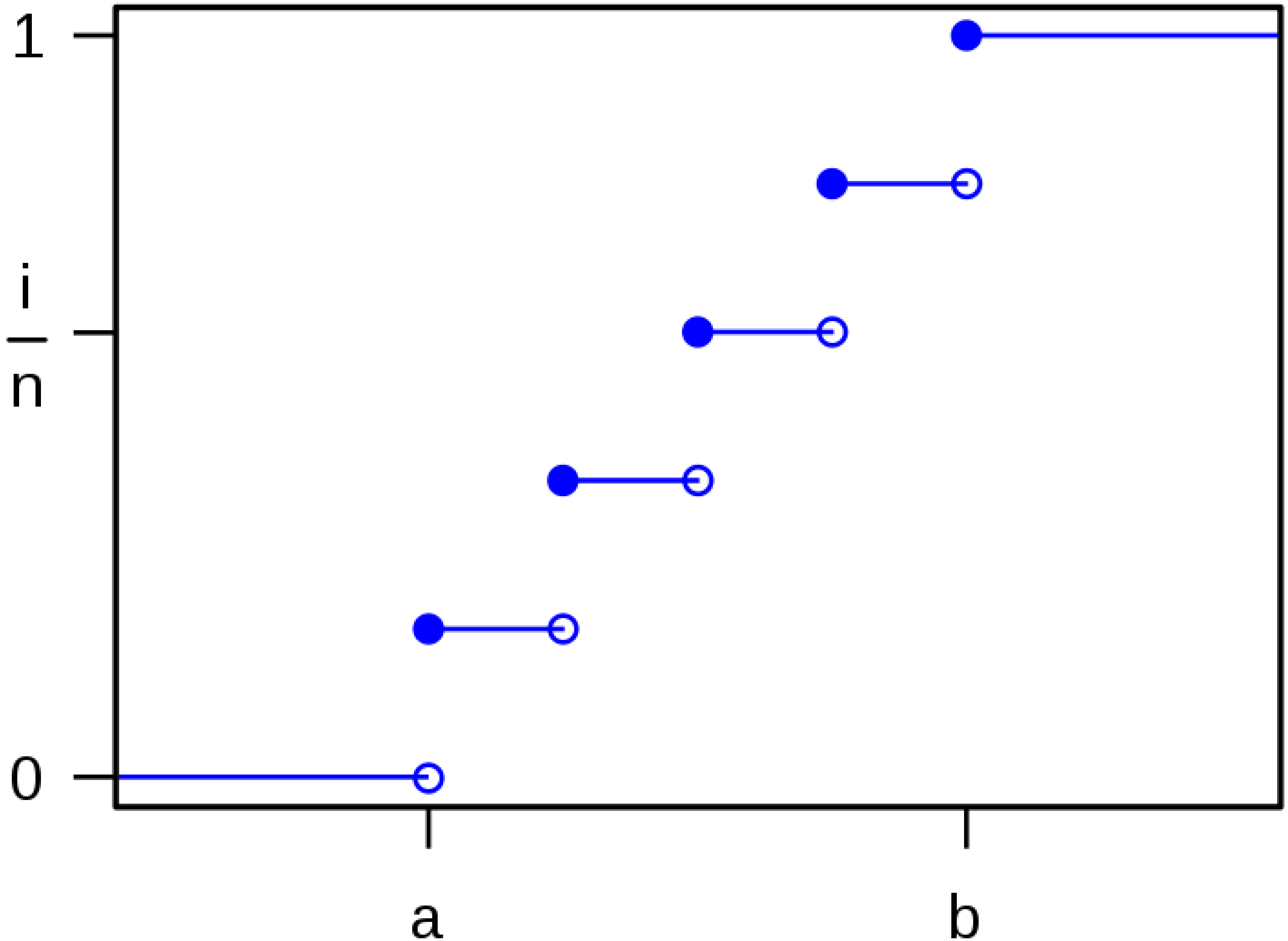
- A random variable X has a discrete uniform distribution if each of the n values in its range, say, a, a+1, a+2, …, b has equal probability . Then,

$$f(x) = P(X = x) = \begin{cases} \dfrac{1}{b - a + 1} = \dfrac{1}{n} & \text{for } x = a, a+1, \ldots, b \quad n = b - a + 1 \\ 0 & \text{otherwise} \end{cases}$$

Probability function n = 5 where n = b − a + 1



http://en.wikipedia.org/wiki/Uniform_distribution_%28discrete%29



Discrete uniform cumulative distribution function

n = 5

http://en.wikipedia.org/wiki/Uniform_distribution_%28discrete%29

Probability Function

- There are $b - a + 1$ values X can take so the probability at each of these values must be

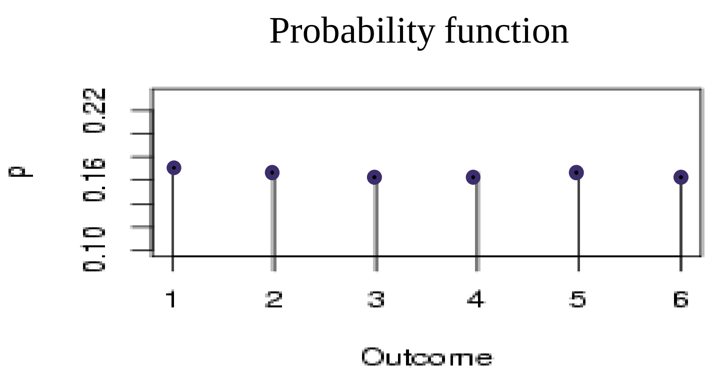
$$\frac{1}{b-a+1}$$ in order that

$$\sum_{x=a}^{b} f(x) = 1.$$

Example: Rolling a Dice

Suppose a fair die is thrown once and let X be the number on the face.

x	*f*(x)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6



$n = 6$ where $n = b - a + 1$

$$f(x) = P(X = x) = \begin{cases} \dfrac{1}{6} & x = 1, \cdots, 6 \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = P(X \le x) = \begin{cases} 0 & x < 1 \\ [x]/6 & 1 \le x < 6 \\ 1 & x \ge 6 \end{cases}$$

Example

Let X be the largest number when a die is rolled 3 times. First find the c.d.f, F(x) and then find the p. *f*. , *f* (x) of X.

S={ (1,1,1), (1,1,2),..., (6,6,6)}

There are 6^3 possible outcomes with each has a probability of 1/ 216

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ [x]^3 /216 & \text{for } 1 \leq x < 6 \\ 1 & \text{for } x \geq 6 \end{cases}$$

$$f(x) = F(x) - F(x-1)$$

$$= \frac{x^3 - (x-1)^3}{216}$$

$$= \frac{x^3 - \{ x^3 + 3x^2(-1) + 3x(-1)^2 + (-1)^3\}}{216}$$

$$f(x) = \begin{cases} \dfrac{3x^2 - 3x + 1}{216} & \text{for } x=1, \ldots, 6 \\ 0 & \text{o/w} \end{cases}$$

The Hypergeometric Distribution

- "n" objects in a sample taken from a finite population of size N.
- Sample taken without replacement.
- Objects can be classified into two distinct types: success(S) or failure(F).
- There are **r success in the population.**
- Let X be the number of successes obtained.

UNIVERSITY OF WATERLOO

Notation

N	The number of items in the population.
r	The number of "S" in the population
N-r	The number of "F" in the population
n	The number of items in the sample
x	The number of "S" in the sample
n-x	The number of "F" in the sample

UNIVERSITY OF WATERLOO

X has a Hypergeometric distribution, is given by

$$F(x) = P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad x = 0, 1, \ldots, \min(n, r)$$

Where $\max(0,\ n-N+r) \leq x \leq \min(n, r)$

Example

Different computers are checked from 10 in the department. 4 of the 10 computers have illegal software loaded. What is the probability that 2 of the 3 selected computers have illegal software loaded?

$N = 10 \qquad n = 3$

$r = 4 \qquad x = 2$

$$P(X=2) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = \frac{(6)(6)}{120} = 0.3$$

The probability that 2 of the 3 selected computers have illegal software loaded is 0.30, or 30%.

Example

Suppose that a batch of 100 items contains 6 that are defective and 94 that are not defective. If X is the number of defective items in a randomly drawn sample of 10 items from the batch, find

(a) $P(X = 0)$ (The probability that non of the 10 is defective)

(b) $P(X > 2)$

(a)

$$P(X = 0) = \frac{\binom{6}{0}\binom{94}{10}}{\binom{100}{10}}$$

(b)

(The probability that at most 2 of the 10 is defective)

$$P(X > 2) = 1 - P(X \leq 2)$$

$$= 1 - \left\{ \frac{\binom{6}{0}\binom{94}{10} + \binom{6}{1}\binom{94}{9} + \binom{6}{2}\binom{94}{8}}{\binom{100}{10}} \right\}$$

UNIVERSITY OF WATERLOO

Example

In Lotto 6/49 a player selects a set of six numbers (with no repeats) from the set {1, 2,…, 49}.

In the lottery draw six numbers are selected at random. Find the probability function for X, the number from your set which are drawn.

$N = 49$ $\quad$ $n = 6$

$r = 6$ $\quad$ $x = 0, 1, 2, 3, 4, 5, 6$

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{6}{x}\binom{43}{6-x}}{\binom{49}{6}} \qquad x = 0, 1, 2, 3, 4, 5, 6$$

Exercise

If X has a Hypergeometric distribution, show that $\sum_{\text{all } x} f(x) = 1$.

UNIVERSITY OF WATERLOO

Binomial Distribution

- There are many experiments for which the results of each trial can be reduced to two outcomes: success and failure.

- Tossing a coin: tail or head.

- Testing a product: good (non defective) or defective item.

UNIVERSITY OF WATERLOO

Jacob Bernoulli

- Bernoulli was born in Basel, Switzerland, 6 January 1655.
- Died 16 August 1705. (aged 50)

Physical Setup:

1. The experiment is repeated for a fixed number of trials, where n is fixed in advance of the experiment.

2. There are only two possible outcomes of interest for each trial. The outcomes can be classified as a success S or as a failure F.

3. The trials are independent of the other trials.

4. The probability of successful P(S) is constant from trial to trial denoted by p.

Notation for Binomial Distribution

n	The number of times a trial is repeated
$p = P(S)$	The probability of success in a single trial.
$q = P(F)$	The probability of failure in a single trial ($q = 1 - p$)
X	The number of successes in n trials: $x = 0, 1, 2, 3, \ldots n$.

Note: if n=1, the Binomial distribution becomes the Bernoulli distribution.

Examples

Decide whether the experiment is a Binomial experiment. If it is, specify the values of n, p and q and list the possible values of the random variable, X. If it is not, explain why.

1. Example

A certain surgical procedure has an 65% chance of success. A doctor performs the procedure on four patients. The random variable represents the number of successful surgeries.

The experiment is a Binomial experiment because it satisfies the four conditions of a Binomial experiment.

In the experiment

- Each surgery represents one trial and there are four surgeries
- Each surgery is independent of the others.
- Also, there are only two possible outcomes for each surgery—either the surgery is a success or it is a failure.
- The probability of success for each surgery is 0.65.

$n = 4$, $p = 0.65$, $q = 1 – 0.65 = 0.35$,

$x = 0, 1, 2, 3, 4$

UNIVERSITY OF WATERLOO

2. Example

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 violations.

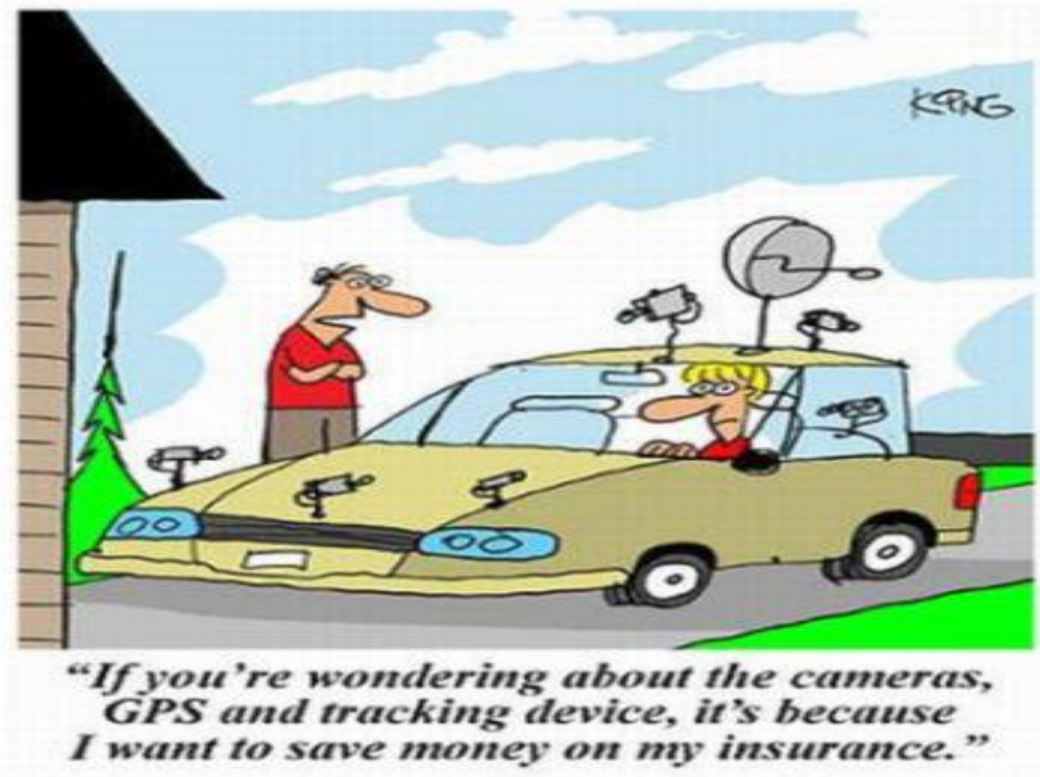
There are 5 inspectors, each of whom will inspect one restaurant during the coming week. The name of each restaurant is written on a different slip of paper and after the slips are thoroughly mixed, each inspector in turn draws one of the slips without replacement. Label the ith trial as a success if the ith restaurant selected (i = 1, 2, .., 5) has no serious violations.

UNIVERSITY OF WATERLOO

The experiment is not Binomial because the trials are not independent.

3. Example

Suppose a certain city 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.



UNIVERSITY OF WATERLOO

- The size of the population being sampled is very large relative to the sample size.

Binomial Probability Function

There are several ways to find the probability of x successes in n trials of a binomial experiment.

The probability of exactly x successes in n trials is:

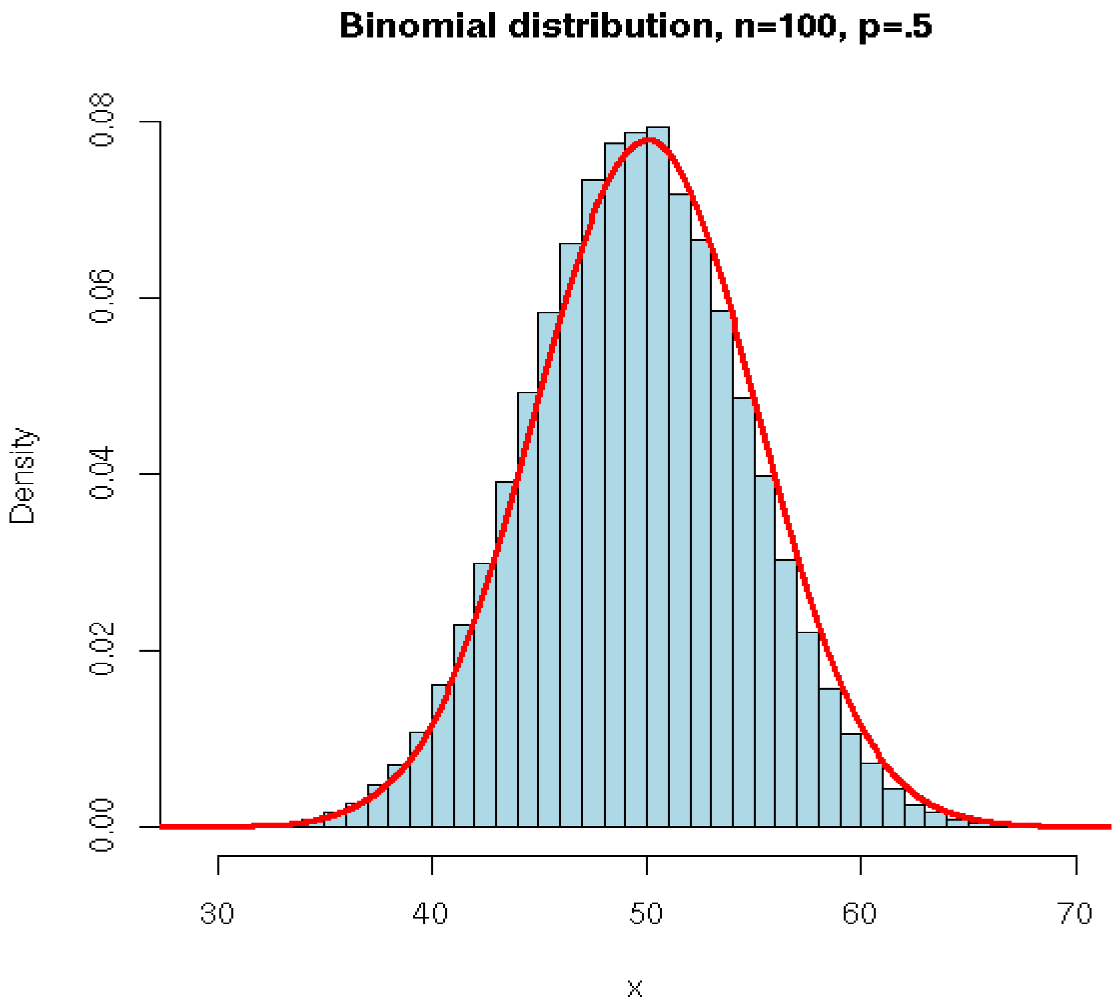
$$f(x) = P(X=x) = b(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0,1,2,\ldots,n \\ 0, & otherwise \end{cases}$$

We write: **X ~ Binomial (n, p)**

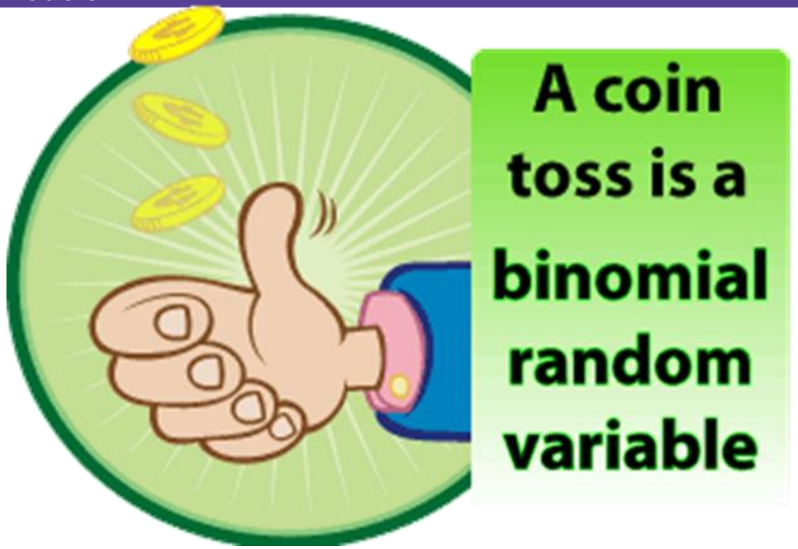
{reads: "X is distributed Binomially with parameters n and p}.

UNIVERSITY OF WATERLOO

f(x) increases to a maximum value near np and then decreases thereafter.



Example



http://faculty.darden.virginia.edu/Pfeiferp/Statisticsinbusiness/cointoss.png

(a) Toss the same fair coin 20 times, what's the probability of getting exactly 10 heads?

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

$$= \binom{20}{10} (.5)^{10} (.5)^{10} = .176$$

UNIVERSITY OF WATERLOO

(b)Toss a fair coin 20 times, what's the probability of getting at most 2 or fewer heads?

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

$F(2)$

$$\binom{20}{0}(.5)^0(.5)^{20} = \frac{20!}{20!0!}(.5)^{20} \quad +$$

$$\binom{20}{1}(.5)^1(.5)^{19} = \frac{20!}{19!1!}(.5)^{20} \quad +$$

$$\binom{20}{2}(.5)^2(.5)^{18} = \frac{20!}{18!2!}(.5)^{20}$$

$$= 1.8x10^{-4}$$

UNIVERSITY OF WATERLOO

Example

A survey indicates that 41% of Canadian women consider walking as their favorite sport. You randomly select four women and ask them if walking is their favorite sport. Find the probability that

(1) Exactly two of them respond yes.

1) Exactly two of them respond yes.

$$P(2) = {}_4C_2 (0.41)^2 (0.59)^{4-2} =$$

$$\frac{4!}{(4-2)!2!}(0.41)^2(0.59)^{4-2} =$$

$$\frac{24}{4}(.1681)(.3481) =$$

$$6(.1681)(.3481) = .35109366$$

(2) At least two of them respond yes.

2) At least two of them respond yes. X

$$P(\text{X} \geq 2) = P(2) + P(3) + P(4)$$

$$P(2) = {}_4C_2(0.41)^2(0.59)^{4-2} = .351093$$
$$P(3) = {}_4C_3(0.41)^3(0.59)^{4-3} = 0.162653$$
$$P(4) = {}_4C_4(0.41)^4(0.59)^{4-4} = 0.028258$$

$$= .351093 + .162653 + 028258$$
$$\approx 0.542$$

UNIVERSITY OF WATERLOO

(3) Fewer than two of them respond yes.

(3) Fewer than two of them respond yes.

$$P(x < 2) = P(0) + P(1)$$

$$P(0) = {}_4C_0(0.41)^0(0.59)^{4-0} = 0.121174$$
$$P(1) = {}_4C_1(0.41)^1(0.59)^{4-1} = 0.336822$$

$$= .121174 + .336822$$
$$\approx 0.458$$

UNIVERSITY OF WATERLOO

Example

If X has a Binomial distribution, show that $\sum_{\text{all } x} f(x) = 1$.

Binomial vs. Hypergeometric Distribution

The main difference is that the

Binomial requires INDEPENDENT trials, and the probability of success(p) is the same in each trial.

- Hypergeometric, the draws are made from a finite number of objects (N) WITHOUT replacement, the trials are NOT INDEPENDENT.

Example

Suppose we have 15 boxes the same shape and size and colour with no labels, but 6 are laptops and 9 are tablets. We randomly pick 8 boxes and open them. Find the probability 3 are laptops.

The correct solution uses Hypergeometric.

Let

X = number of laptops picked

$$f(3) = P(X = 3) = \frac{\binom{6}{3}\binom{9}{5}}{\binom{15}{8}} = 0.396$$

UNIVERSITY OF WATERLOO

If we incorrectly used Binomial, we'd get

$$f(3) = \binom{8}{3}\left(\frac{6}{15}\right)^3\left(\frac{9}{15}\right)^5$$

$$= 0.279$$

- As expected, this is a poor approximation since we're picking over half of a fairly small collection of boxes.

If we had 1500 boxes - 600 laptop and 900 tablet, we're not likely to get the same box again even if we did replace each of the 8 boxes after opening it.

$$f(3) = \binom{8}{3}\left(\frac{600}{1500}\right)^3\left(\frac{900}{1500}\right)^5 = 0.279$$

is a very good approximation.

$$f(3) = P(X = 3) = \frac{\binom{600}{3}\binom{900}{5}}{\binom{1500}{8}} = 0.2794$$

UNIVERSITY OF WATERLOO