STAT230

PROBABILITY

(Chapter 7)

UNIVERSITY OF WATERLOO

Expected Value and Variance
Chapter 7: Chapter objectives

- Understand the types of descriptive statistics.
- Understand the expectation of a random variable.
- Understand the mean and the variance of some discrete distributions.

Please Do Chapter 7 Problems 7.1-7.14

Why do we need descriptive statistics?

Aggregate data into meaningful information.

445 446 397 226
388 3445 188 1002
477 62 432 54 12
98 345 22 45 88 39
774 92 472 565 999
1 34 882 545 4022
827 572 597 364

$$\bar{x} = \ldots$$

UNIVERSITY OF WATERLOO

Summarizing Data on Random Variables

Types of descriptive statistics:

- Representing a data using visual techniques
  - Tables
  - Graphs

- Numerical summary measures for data sets
  - Central Tendency
  - Variation

Example (1)

- Suppose we were to observe cars crossing a toll bridge, and record the number, X, of people in each car.

- Suppose in a small study, data on 25 cars were collected.

- Present the data using the **frequency distribution, which gives the number of times (the "frequency") each** value of X occurred.

Port Mann Bridge (2012)

Ten lanes of British Columbia

Frequency distribution

X	Frequency Count	Frequency
1	~~||||~~ |	6
2	~~||||~~ |||	8
3	~~||||~~	5
4	|||	3
5	||	2
6	|	1
Total		25

❑ Pictorial Methods: Histograms

- The purpose of a histogram is to put numerical information into graphic form so it is easier to understand.

- Histograms are good summaries of data because they show the variability in the observed outcomes very clearly.

UNIVERSITY OF WATERLOO

Constructing a Histogram

1- Determine the frequency and relative frequency of each x.

Frequency of a value = number of times the value occurs.

$$\text{Relative frequency of a value} = \frac{\text{number of times the value occurs}}{\text{number of observations in the data set}}$$

UNIVERSITY OF WATERLOO

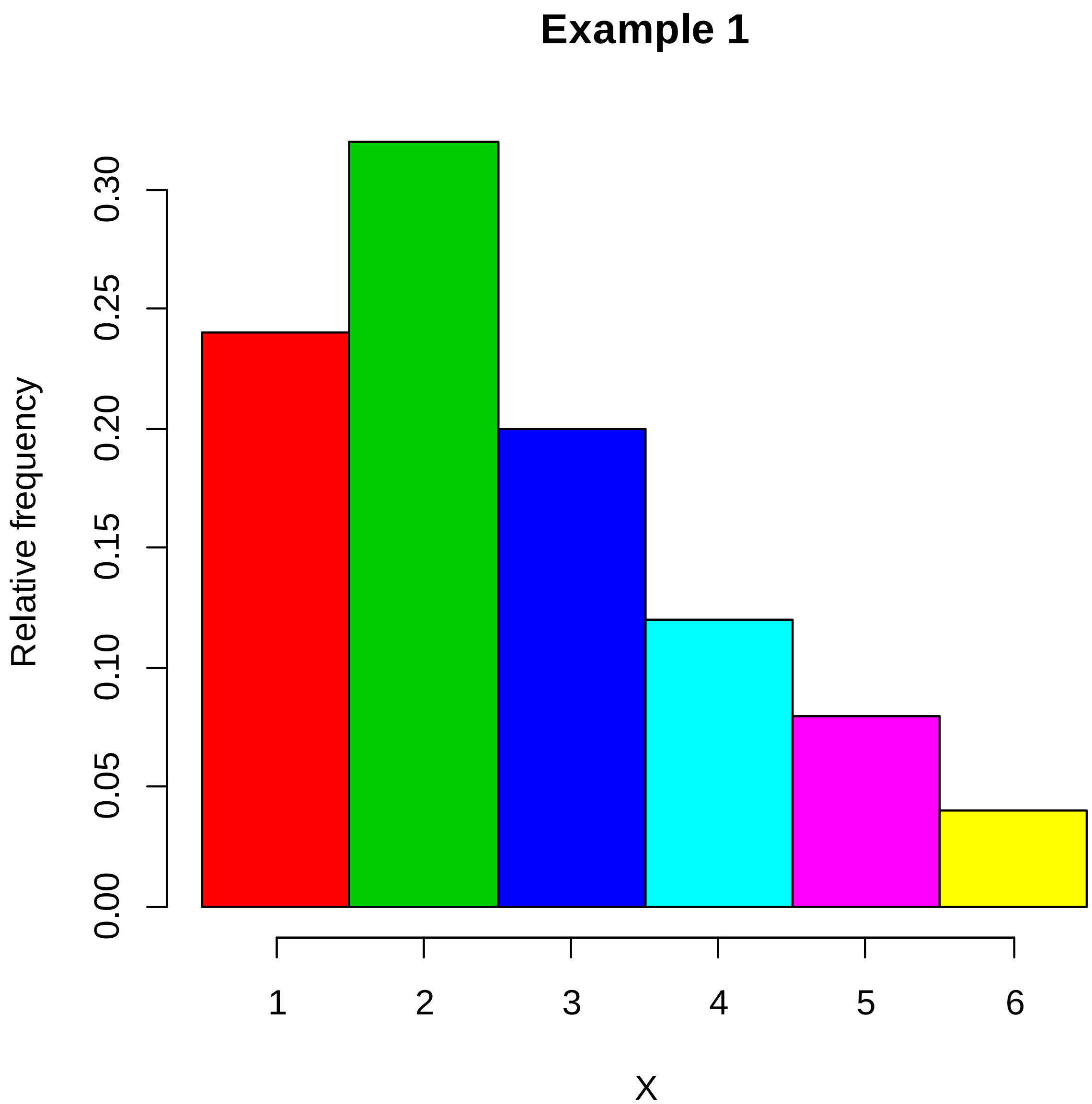
2- Mark possible x values on horizontal scale.

3- Above each value of x, draw a rectangle whose height is the relative frequency or the frequency of that value.

```
# R code for example 1 (toll bridge)

X<-  c( 1,1,1,1,1,1,2,2,2,2,2,2,2,2,3,3,3,3,3,4,4,4,5,5,6)

require(MASS)

truehist(X, f =F, col=c(2:7), ylab="Relative frequency", h=1,
  x0=0.5, main="Example 1")
```

UNIVERSITY OF WATERLOO

Relative frequency (probability) histogram



Example (2)

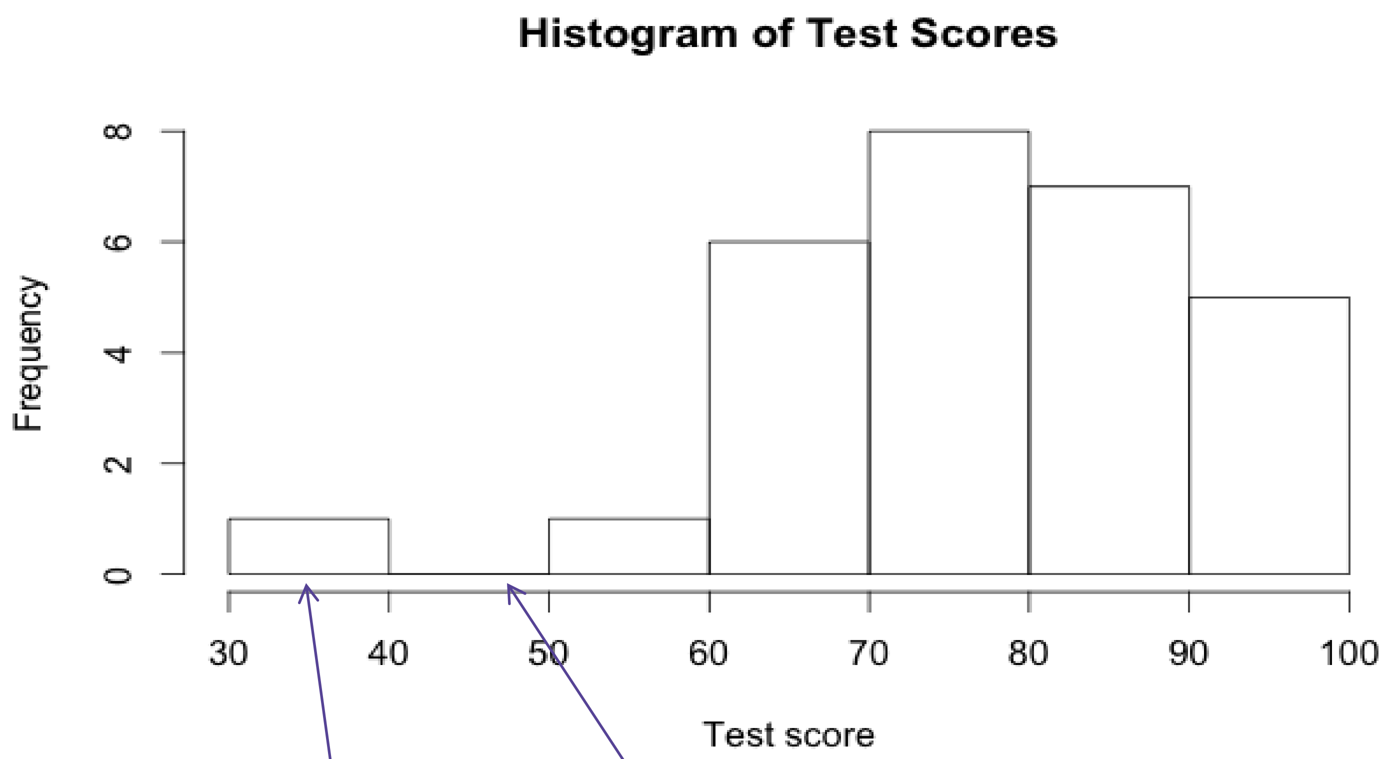
Create a histogram for the following list of Statistic exam scores:

75, 95, 60, 93, 85, 84, 76, 92, 62, 83, 80, 90, 64, 75, 79, 32, 78, 64, 98, 73, 88, 61, 82, 68, 79, 78, 80, 55

The frequency distribution is below

X	Frequency
30-40	1
40-50	0
50-60	1
60-70	6
70-80	8
80-90	7
90-100	5

UNIVERSITY OF WATERLOO



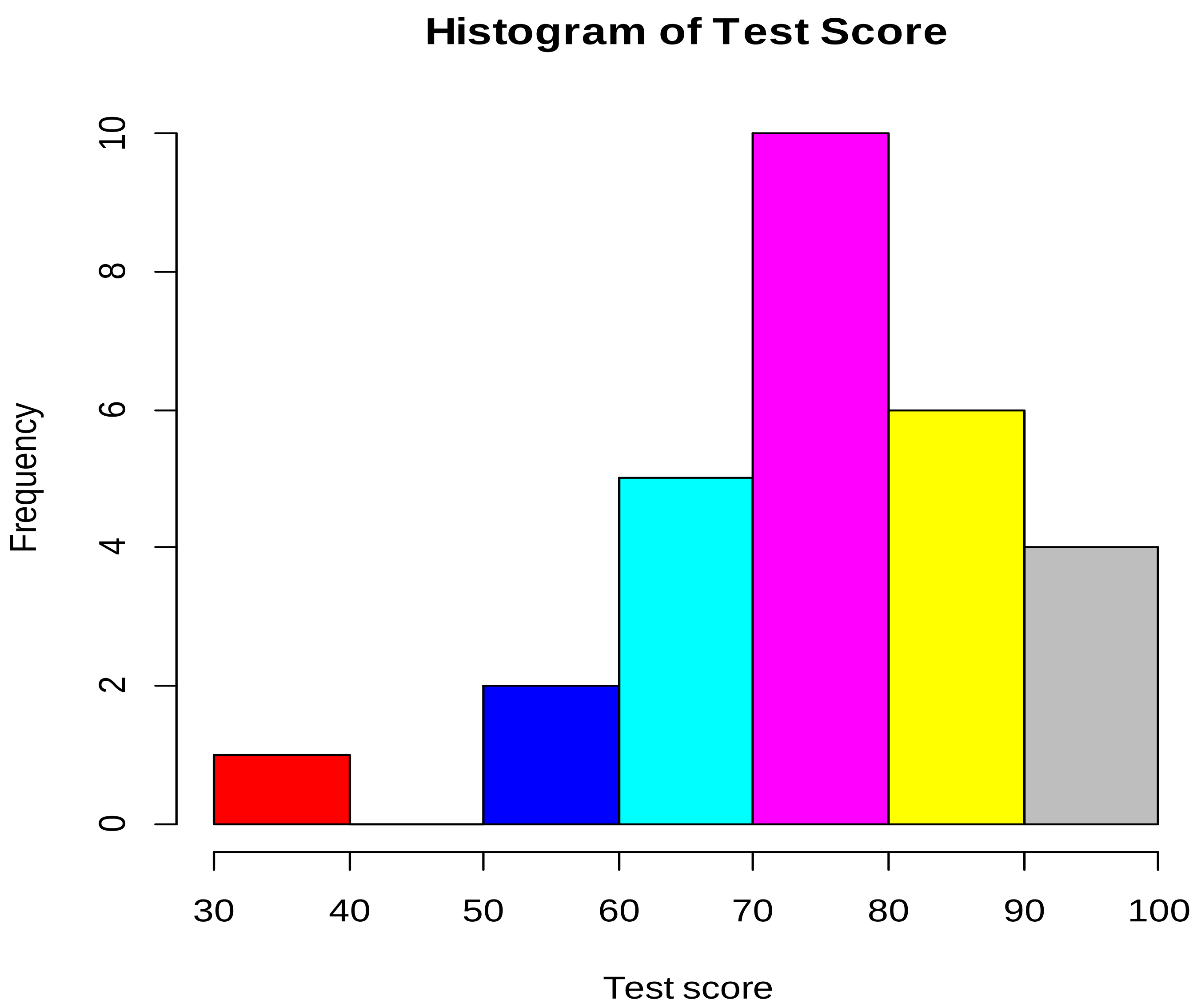
Histogram of Test Scores
Frequency
0
2
4
6
8
30
40
50
60
70
80
90
100
Test score

1 data

0 data

```
# R code for the test score example

X <-
  c(75,95,60,93,85,84,76,92,62,83,80,90,64,75,79,32,78,64,98,73,
  88,61,82,68,79,78,80,55)

hist(X, col=c(2:8),ylab="Frequency",main="Histogram of Test
  Score")
```

UNIVERSITY OF WATERLOO

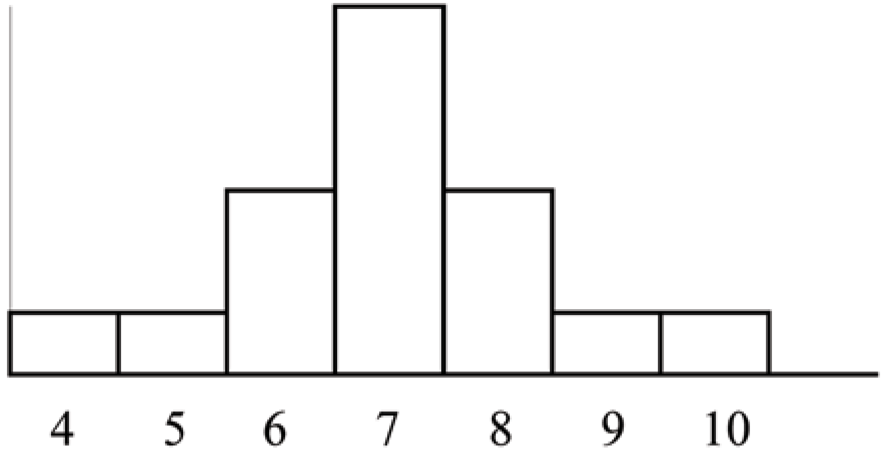

Histogram of Test Score
Frequency
0
2
4
6
8
10
30
40
50
60
70
80
90
100
Test score

Describing the Shape of a Histogram

- The shape of a histogram has several characteristics:

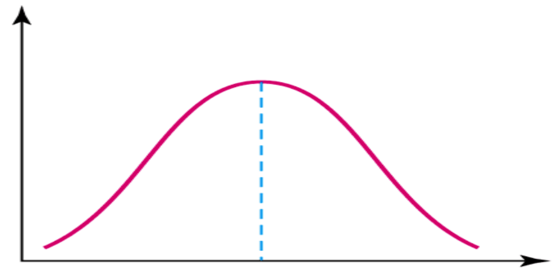
- Symmetry if draw line through center, picture on one side would be mirror image of picture on other side.

Example: bell-shaped data set.



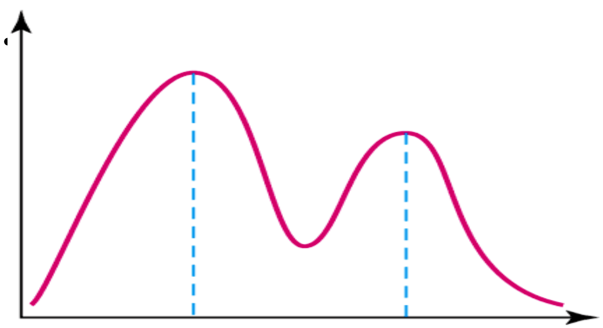
- Unimodal: single prominent peak.

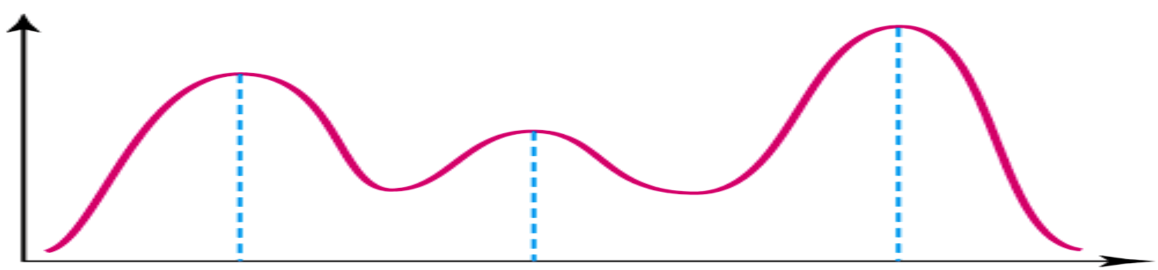
  (A local maximum in a chart.)



- Bimodal: two prominent peaks.

- Multimodal: more than two prominent peaks.





UNIVERSITY OF WATERLOO

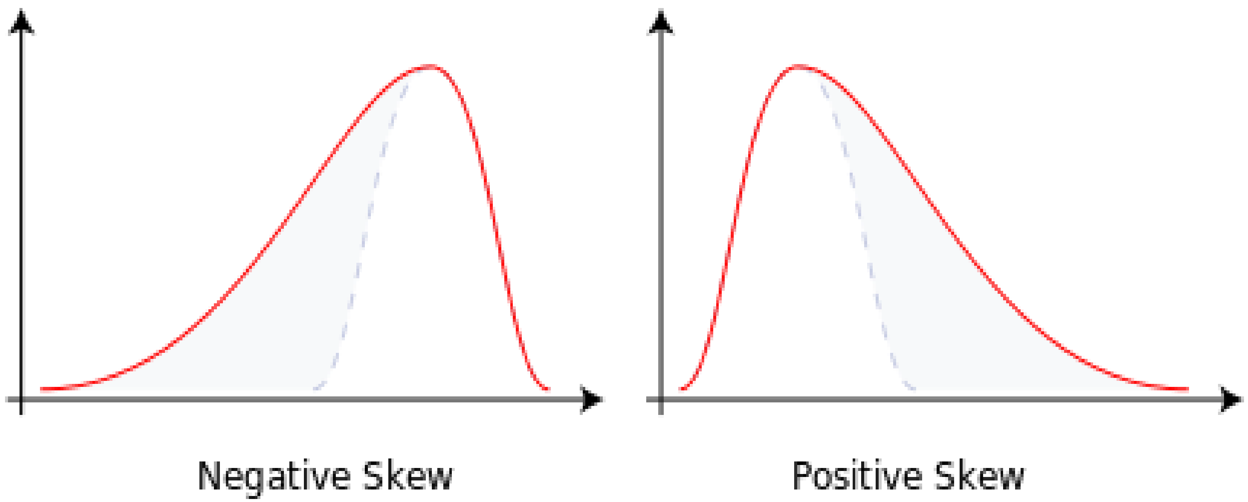
- Skewness: Whether or not the data is pulled to one side.

Right-skewed (positively skewed) :

The right or upper tail is stretched out

Left-skewed (negatively skewed) :

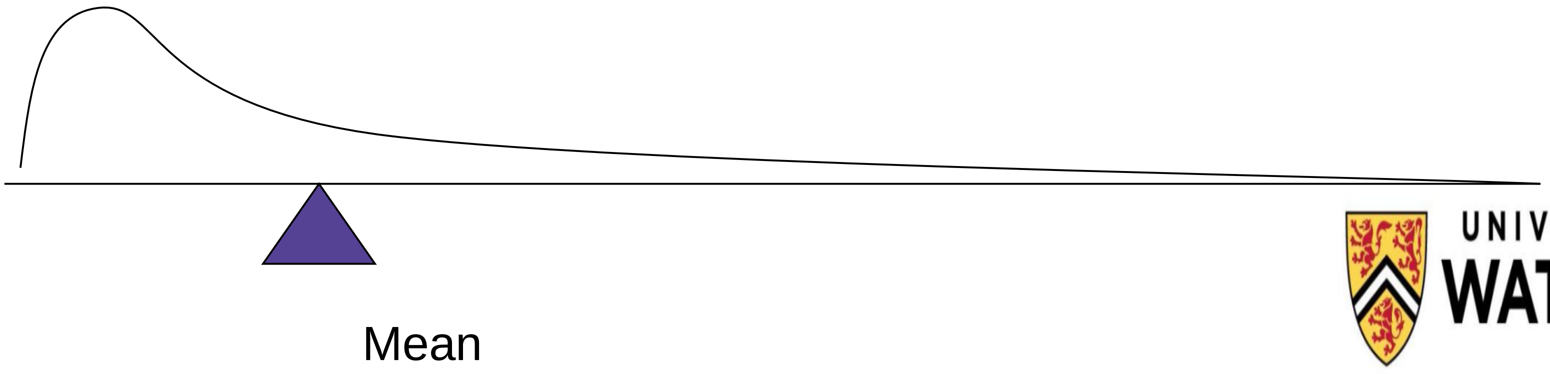
The stretching is to the left



Measures of Location

1- Arithmetic Mean: The arithmetic average value of observations.

- The sample mean: measures the location , balance, center of a sample.
- The sum of a list of numbers divided by the size of the list.



If the n observations in a random sample are denoted by $x_1, x_2, \ldots, x_n$, the sample mean is

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

For the N observations in a population denoted by $x_1, x_2, \ldots, x_N$, the population mean is analogous to a probability distribution as

$$\mu = \sum_{i=1}^{N} x_i \cdot f(x) = \frac{\sum_{i=1}^{N} x_i}{N}$$

where "N" is the populations size

UNIVERSITY OF WATERLOO

Example (Toll Bridge)

X	Frequency Count	Frequency
1	$\cancel{||||}$ |	6
2	$\cancel{||||}$ |||	8
3	$\cancel{||||}$	5
4	|||	3
5	||	2
6	|	1
Total		25

The arithmetic mean is

$$\frac{(6*1) + (8*2) + \ldots + (1*6)}{25} = \frac{65}{25} = 2.6$$

There was an average of 2.6 persons per car.

Notes

- The sample mean ($\overline{x}$) is not in general an integer, even though X is.
- We would not expect to get precisely the same sample mean if we get another sample from the same population.

2- Median: The middle value

a- Can be found by arranging all the observations from lowest value to highest value and picking the middle one.

b- If there are two middle numbers, then take their mean

For the ordered values

$\tilde{x} =$

The single middle value if n is odd

= (n+1) / 2

The average of the two middle values if n is even

= average of {n / 2} and { (n/2) +1}

The population median is $\tilde{\mu}$

UNIVERSITY OF WATERLOO

Example (3)

6, 8, 1, 5, 4, 8, 11, 3

ordered values

1, 3, 4, 5, 6, 8, 8,11

Mean = 5.75,

Median = 5.5

http://www.mynamesnotmommy.com/wp-content/uploads/2013/05/question-mark.png

The Sample Mode

The **mode** is the value that appears most often in a set of data.

Example

What is the mode of the sample

(a) 1, 3, 6, 6, 6, 6, 7, 7, 12, 12, 17

Mode= 6.

(b) 1, 1, 2, 4, 4

- Mode is not unique (1 and 4) the dataset may be said to be Bimodal.
- A set with more than two modes may be described as Multimodal.

(c) 1, 1, 2, 4, 4, 3, 9, 9, 10, 15

(d) 1, 1, 2, 2, 4, 4

Mode is not unique (1, 2 and 4) the dataset may be said to be Uniform.

(e) 1, 1, 2, 2, 4, 4, 3, 3, 9, 9, 10, 10, 15, 15

(1, 2, 3, 4, 9, 10, 15 are all modes)

(f) 1, 2, 4, 3, 9, 10, 15

No mode.

UNIVERSITY OF WATERLOO

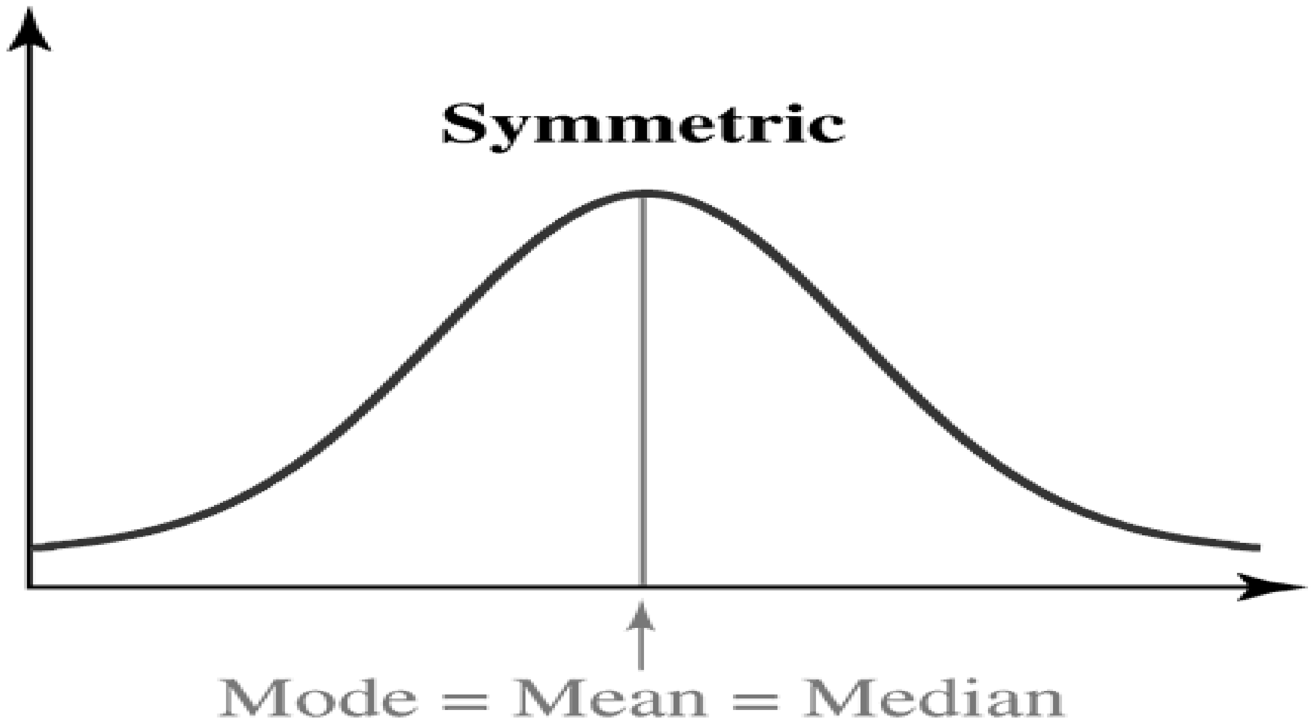
Outliers

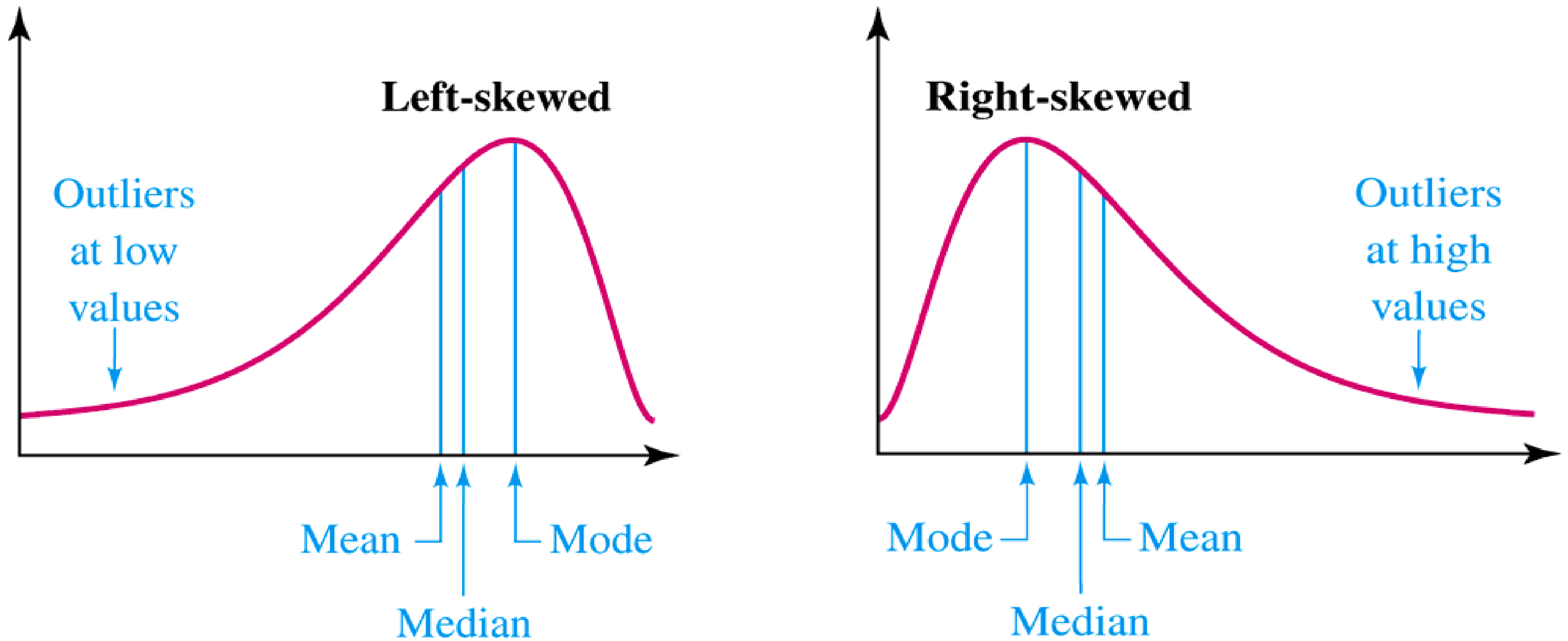
- Sometimes we get values that are very different from the rest of the data.

- May indicate something worth investigating
  - Possibly an error in our data?
  - Perhaps just a natural outlier.

  - Outliers affect the mean the most. Median and mode are robust statistics (not affected much by outliers).

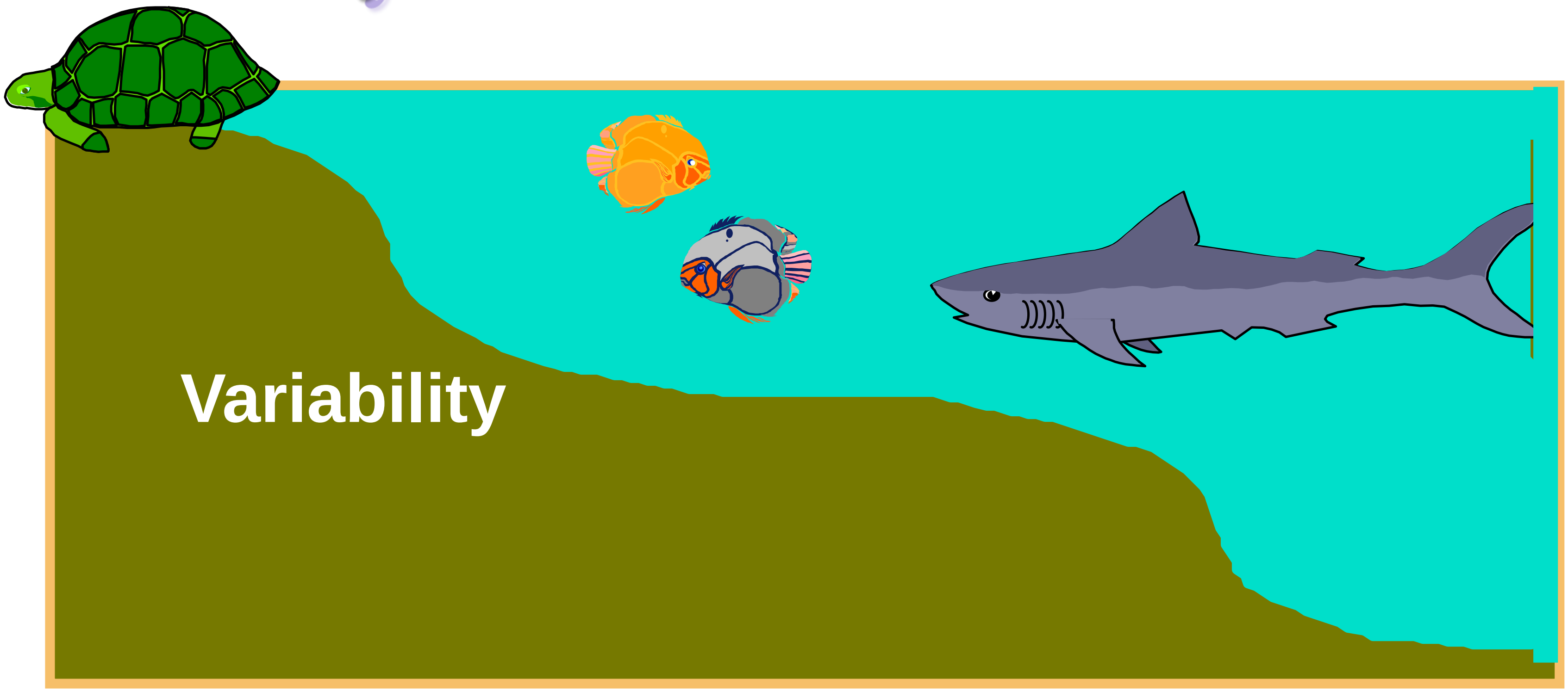
UNIVERSITY OF WATERLOO

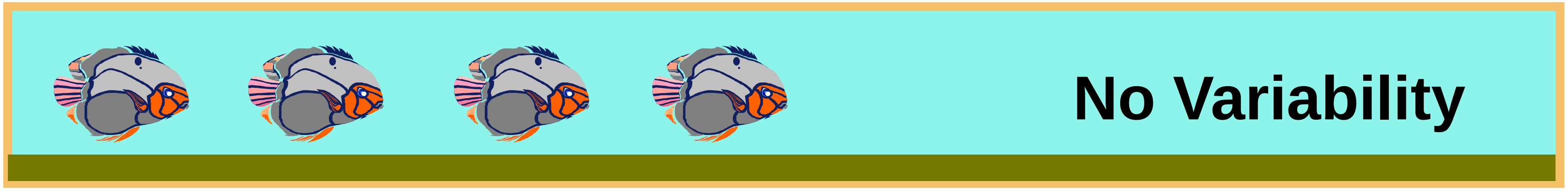

Symmetric
Mode = Mean = Median



UNIVERSITY OF WATERLOO

Variability





Sample Standard Deviation

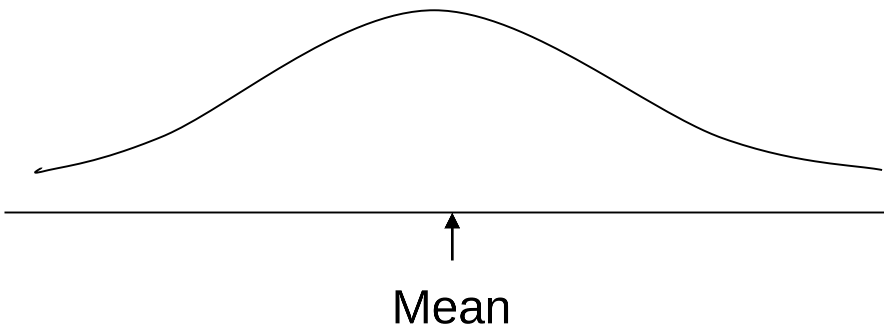
- The standard deviation is used to describe the variation (measure of dispersion) around the mean.

- To get the standard deviation of a **SAMPLE** of data:

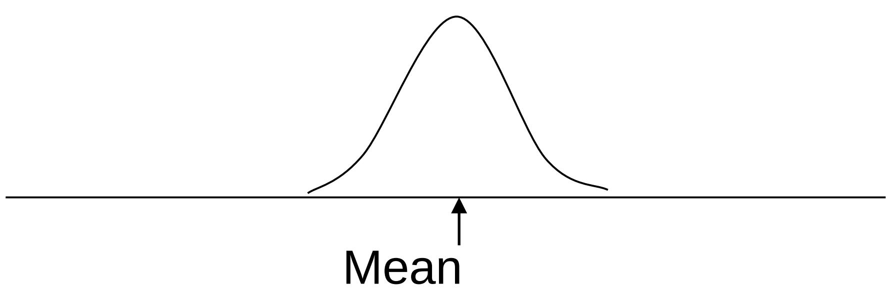
- Calculate the **variance** S^2

- Take the square root to get the **standard deviation** S**.**

The larger the S, the further the individual cases are from the mean.



The smaller the S, the closer the individual scores are to the mean.



Note

- Although variance is a useful measure of spread, its units are units squared.

- The standard deviation (square root of the variance) is more intuitive, because it has the same units as the raw data and the mean.

- Like the mean, the s.d. will be inflated by an outlier.

Mean and Variance of a Discrete Random Variables

- Used to summarize the probability distribution of a random variable X.

- The expectation (also called the mean or the expected value) of a discrete random variable X with probability function $f(x)$.

- The mean is a measure of the center of the probability distribution.

UNIVERSITY OF WATERLOO

- The *expectation* of X is also often denoted by the Greek letter μ.

- The mean of the discrete random variable X is a weighted average of the possible values of X, with weights equal to the probabilities.

If X is a discrete r.v. then

- The expected value or the mean of X is

$$\mu_x = E(X) = \sum_x x \ f(x)$$

- The Variance of X , denoted as σ^2 or V(X), is

$$\sigma^2 = V(X) = E\,(X - \mu)^2 = \sum_x (x - \mu)^2 \ f(x) = \sum x^2 \ f(x) - \mu^2$$

$$= E(X^2) - [E(X)]^2$$

The standard deviation (SD) of X is $\sqrt{\sigma^2} = \sigma$

UNIVERSITY OF WATERLOO

Prove that $Var(X) = E(X^2) - \mu^2$

$$Var(X) = E\{(X - \mu)^2\}$$

$$= E(X^2) - 2\mu E(X) + \mu^2 \quad (\mu \text{ is constant})$$

$$= E(X^2) - 2\mu^2 + \mu^2 \quad (\text{since } E(X) = \mu)$$

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

Note

$$\mathrm{Var}(x) = E[\, X(X-1)\,] + \mu - \mu^2$$

Proof :

let

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

$$\begin{aligned}
\mathrm{Var}(X) &= E\,(X^2) - \mu^2 \\
&= E\,[\, X(X-1) + X \,] - \mu^2 \\
&= E\,[\, X(X-1)] + E(X) - \mu^2 \\
&= E\,[\, X(X-1)] + \mu - \mu^2
\end{aligned}$$

UNIVERSITY OF WATERLOO

Example (4)

Suppose that UW have 15,000 students and let X denotes the number of courses for which a randomly selected student is registered.

x	1	2	3	4	5	6	7	Total
$f(x)$	0.01	0.03	0.13	0.25	0.39	0.17	0.02	
Number registered	150	450	1950	3750	5850	2550	300	15000

Calculate the average number of courses per student.

$$\mu_x = \frac{1(150)+2(450)+3(1950)+4(3750)+5(5850)+6(2550)+7(300)}{15000}$$

$$= 4.57$$

Or

$$\mu_x = 1\,f(1)+2\,f(2)+\ldots\ldots.+7\,f(7)$$

$$= 1(0.01)+2(0.03)+\ldots..+7(0.02)$$

$$= 4.57$$

- Note: μ is not a possible value of the variable X

The Expected Value of a Function

If X is a discrete r.v. with set of possible values D and p.f $f(x)$ then the expected value of any function g(x), denoted by E[g(x)] or $\mu_{g(x)}$, is computed as

$$E[\, g(X)\,] = \sum_{\text{all } x} g(x)\; f(x)$$

Example(5)

Note: Use example (4)

Let g(x)= 10X+2, What is the expected value of g(x)?

$$E[\, g(X)\,] = \sum_{\text{all } x} g(x)\ f(x)$$

g(x) = 10X +2

g(x) = 12, 22, 32, 42, 52, 62, 72

E(g(x))= 12(0.01)+ 22(0.03)+...+72(0.02)= 47.7

UNIVERSITY OF WATERLOO

Rule of Expected value

$$E\,(a\,X + b) = a\,E(X) + b$$

Example (6)

Calculate the expected value in example (5) using this rule

$g(x) = 10X+2$

$E(\,g(x)) = E(10X+2) = 10E(X) + 2$

$$= 10(4.57) + 2 = 47.7$$

UNIVERSITY OF WATERLOO

Properties of Expectation

1- For constants a and b,

$$E\,[a\,g(X) + b] = a\,E\,[g(X)] + b$$

2. For constants a and b and functions g_1 and g_2, it is also easy to

$$E\,[a\,g_1(X) + b\,g_2(X)] = a\,E\,[g_1(X)] + b\,E\,[g_2(X)]$$

Example (7)

Calculate the mean of X

a)

x	1	2	3
f(x)	0.2	0.6	0.2

b)

x	1	2	5
f(x)	0.3	0.6	0.1

UNIVERSITY OF WATERLOO

mean(a) = $1(0.2) + 2(0.6) + 3(0.2) = 2$

mean(b) = $1(0.3) + 2(0.6) + 5(0.1) = 2$

Note:

Although both distributions have the same center of μ, the distribution in (b) has greater spread or variability or dispersion than does that of (a).

Rule of Variance

$$V(a X + b) = a^2 \sigma^2_x$$

$$\sigma_{ax+b} = |a| \sigma_x$$

Note :

The addition of the constant b does not affect the variance, because b change the location (mean value) but, not the spread of values.

Proof:

Let $Y = aX + b$

$$\sigma^2_Y = E[\,(Y - \mu_Y)^{2}]$$

$$= E\ [\ \ (a\,X + b) - (a\,\mu_X + b)\ \ ]^2\ \}$$

$$= E\ [\ (a\,X - a\,\mu_X)^2\ ]$$

$$= a^2\ E\ [\ (X - \mu_X)^2\ ]$$

$$= \ \ a^2\ \sigma^2_x$$

Example (8)

a. Calculate the Var(X) where $\mu_x = 5.4$

b. Calculate the Var(2X+3)

c. V(4-2X)

x	4	6	8
f(x)	0.5	0.3	0.2

$$V(X) = E(x - \mu)^2$$

$$= \sum (x - \mu)^2 f(x) = \sum x^2 f(x) - \mu^2$$

$$= E(X^2) - [E(X)]^2$$

a.

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

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

OR

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

$$\sigma^2 = E(X^2) - E(X)^2 = 31.6 - (5.4)^2 = 2.44$$

b. V(2X+3)= ?

V(a X + b) = a^2 V(x)

V(2X+3)= 4(2.44)

(c) $V(4-2X) = V(4) + (-2)^2\, V(X)$

$$= 0 + 4(2.44)$$

Some Applications of Expectation

- Because expectation is an average value, it is frequently used in problems where costs or profits are connected with the outcomes of a random variable X.

- It is also used a lot as a summary statistic for probability distributions; for example,

    - The expected life (expectation of lifetime) for a person.
    - The expected return on an investment.

Example: Expected Winnings in a Lottery

- A small lottery sells 1000 tickets numbered 000, 001, . . . , 999; the tickets cost \$10 each.
- When all the tickets have been sold the draw takes place: this consists of a simple ticket from 000 to 999 being chosen at random.
- For ticket holders the prize structure is as follows:

UNIVERSITY OF WATERLOO

- Your ticket is drawn - win $5000.
- Your ticket has the same first two number as the winning ticket, but the third is different – win $100.
- Your ticket has the same first number as the winning ticket, but the second number is different - win $10.
- All other cases - win nothing.

- Let the random variable X represent the winnings from a given ticket.

- Find E(X).

UNIVERSITY OF WATERLOO

Solution:

- The possible values for X are 0, 10, 100, 5000 (dollars).

$$f(0) = 0.9, \quad f(10) = 0.09, \quad f(100) = 0.009, \quad f(5000) = 0.001$$

- The expected winnings are thus the expectation of X, or

$$E(X) = \sum_{\text{all } x} x\, f(x) = \$6.80$$

- Thus, the gross expected winnings per ticket are \$6.80.

- Since a ticket costs \$10 your expected net winnings are negative, -\$3.20 (i.e. an expected loss of \$3.20).

UNIVERSITY OF WATERLOO

Note

- The random variable associated with a given problem may be defined in different ways but the expected winnings will remain the same.

- For example, instead of defining X as the amount won we could have defined X = 0, 1, 2, 3 as follows:

$X = 3$ all 3 digits of number match winning ticket

$X = 2$ 1st 2 digits (only) match

$X = 1$ 1st digit (but not the 2nd) match

$X = 0$ 1st digit does not match

- Now, we would define the function $g(x)$ as the winnings when the outcome $X = x$ occurs.

$$g(0) = 0,\ g(1) = 10,\ g(2) = 100,\ g(3) = 5000$$

- The expected winnings are then

$$E(g(X)) = \sum_{x=0}^{3} g(x)f(x) = \$6.80,$$

the same as before.

UNIVERSITY OF WATERLOO

Example: Diagnostic Medical Tests

Suppose we have two cheap tests and one expensive test, with the following characteristics.

All three tests are positive if a person has the condition (there are no “false negatives”), but the cheap tests give “false positives”.

Let a person be chosen at random, and let

D = {person has the condition}. The three tests are

Test 1: $P(\text{positive test} \mid \overline{D}) = 0.05$; test costs \$5.00

Test 2: $P(\text{positive test} \mid \overline{D}) = 0.03$; test costs \$8.00

Test 3: $P(\text{positive test} \mid D) = 0$; test costs \$40.00

We want to check a large number of people for the condition, and have to choose among three testing strategies:

(i) Use Test 1, followed by Test 3 if Test 1 is positive.

(ii) Use Test 2, followed by Test 3 if Test 2 is positive.

(iii) Use Test 3.

Determine the expected cost per person under each of strategies (i), (ii) and (iii).

We will then choose the strategy with the lowest expected cost.

It is known that about 0.001 of the population have the condition (i.e. $P(D) = 0.001$, $P(\bar{D}) = .999$).

UNIVERSITY OF WATERLOO

Solution:

- Define the random variable X as follows (for a random person who is tested):

$$X = 1 \text{ if the initial test is negative}$$
$$X = 2 \text{ if the initial test is positive}$$

Also let g(x) be the total cost of testing the person.

The expected cost per person is then

$$E[g(X)] = \sum_{x=1}^{2} g(x)\, f(x)$$

UNIVERSITY OF WATERLOO

The probability function $f(x)$ for X and function g(x) differ for strategies (i), (ii) and (iii).

Consider for example strategy (i). Then

$$f(2) = P(X = 2) = P(\text{initial test positive})$$

$$= P(D) + P(\text{positive} \mid \overline{D})P(\overline{D})$$

$$= 0.001 + (0.05)(.999) = 0.0510$$

The rest of the probabilities, associated values of g(X) and E[g(X)] are obtained below.

(i)

$f(1) = P\ (X = 1) = 1 - f\ (2) = 1 - 0.0510 = 0.949$

$g(1) = 5,\ \ g(2) = 45$

$E[g(X)] = 5(.949) + 45(.0510) = \$7.04$

(ii)

$f(2) = 0.001 + (0.03)(0.999) = 0.03097$

$f(1) = 1 - f(2) = 0.96903$

$g(1) = 8, \;\; g(2) = 48$

$E[g(X)] = 8(.96903) + 48(.03097) = \$9.2388$

(iii)

$f(2) = 0.001, \;\; f(1) = 0.999$

$g(2) = g(1) = 40$

$E[g(X)] = \$40.00$

Thus, its cheapest to use strategy (i).

UNIVERSITY OF WATERLOO

Examples of Discrete Probability Distributions:

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

UNIVERSITY OF WATERLOO

The Mean and Variance of X

If X follows a binomial distribution with parameters n and p: **X ~ Binomial (n, p)**

Then: $\mu_x = E(X) = np$

$$\sigma_x^2 = Var(X) = np(1-p)$$

$$\sigma_x = SD(X) = \sqrt{np(1-p)}$$

Example

Let X= number of heads in 100 tosses of a fair coin X ~ Binomial (100, 0.5)

$E(x) = 100 * 0.5 = 50$,

$Var(X) = 100 * 0.5 * 0.5 = 25$, $SD(X) = 5$

UNIVERSITY OF WATERLOO

Mean of Binomial distribution

Let X ~ *Binomial*(n, p). Find E(X).

Solution:

$$\mu = E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^{x} (1-p)^{n-x}$$

- When x = 0 the value of the expression is 0.
- We can therefore begin our sum at x = 1.
- Provided $x \neq 0$, we can expand x! as x(x − 1)!

UNIVERSITY OF WATERLOO

$$\mu = \sum_{x=1}^{n} \frac{n(n-1)\,!}{(x-1)!\,[(n-1)-(x-1)]!}\; p\;\; p^{\,x-1}\; (1-p)^{(n-1)-(x-1)}$$

$$= \; np(1-p)^{\,n-1} \sum_{x=1}^{n} \binom{n-1}{x-1} \left(\frac{p}{1-p}\right)^{x-1}$$

Let y = x − 1 in the sum, to get

$$= np(1-p)^{n-1} \sum_{y=0}^{n-1} \binom{n-1}{y} \left(\frac{p}{1-p}\right)^{y}$$

$$= np(1-p)^{n-1} \left(1 + \frac{p}{1-p}\right)^{n-1}$$

by the Binomial Theorem

$$\sum_{x=0}^{n} \binom{n}{x} a^x = (1+a)^n$$

$$= np(1-p)^{n-1} \frac{(1-p+p)^{n-1}}{(1-p)^{n-1}}$$

$$= np$$

Exercise page 130 Course Notes

Show that the variance of Binomial random variable is equal to $np(1-p)$.

The Mean and variance of the Hypergeometric rv. X

E(X) = n (r / N)

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

where p=r/N

Note:

- The means of the Binomial and Hypergeometric rv's are equal.
- The variances of the two r.v.'s differ by the factor (N-n)/(N-1) (called finite population correction factor)

UNIVERSITY OF WATERLOO

The Mean and the Variance of Poisson Distribution

Mean $\boxed{\mu = \lambda t}$

Variance $\boxed{\sigma^2 = \lambda t}$

Standard Deviation $\boxed{\sigma = \sqrt{\lambda t}}$

The mean and the variance are equal $E(X) = V(X) = \lambda t$

Mean of the Poisson Distribution

Let X have a Poisson distribution where λ is the average rate of occurrence and the time interval is of length t. Find $\mu = E(X)$.

Solution:

The probability function of X is

$$f(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

$$\mu = E(X) = \sum_{x=0}^{\infty} \frac{x\,(\lambda t)^x e^{-\lambda t}}{x!}$$

UNIVERSITY OF WATERLOO

- As in the Binomial example, we can eliminate the term when $x = 0$ and expand $x!$ as $x(x - 1)$ for $x = 1, 2, \cdots, \infty$.

$$\mu = E(X) = \sum_{x=1}^{\infty} \frac{x\,(\lambda t)^x e^{-\lambda t}}{x!}$$

$$= \sum_{x=1}^{\infty} \frac{x\,(\lambda t)^x e^{-\lambda t}}{x(x-1)!}$$

$$= \sum_{x=1}^{\infty} (\lambda t)\, e^{-\lambda t} \frac{(\lambda t)^{x-1}}{(x-1)!}$$

UNIVERSITY OF WATERLOO

$$\mu = (\lambda t)\, e^{-\lambda t} \sum_{x=1}^{\infty} \frac{(\lambda t)^{x-1}}{(x-1)!}$$

Let $y = x - 1$ in the sum,

$$= (\lambda t)\, e^{-\lambda t}\, e^{\lambda t}$$

$$= \lambda t$$

$$= \mu$$

Note:

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

UNIVERSITY OF WATERLOO

Exercise page 133 Course Notes

Show that the variance of Poisson random variable is equal to $\lambda t = \mu$.

Uniform Distribution (Discrete)

- Mean= $\dfrac{a+b}{2}$

- Variance = $\dfrac{(b-a+1)^2-1}{12}$

<u>NB(k, p)</u> where $0 < p < 1,\ q = 1 - p$

$$\text{Mean} = \frac{k\,(1-p)}{p}$$

$$\text{Variance} = \frac{k\,(1-p)}{p^2}$$

<u>Geometric(p)</u> where $0 < p < 1,\ q = 1 - p$

$$\text{Mean} = \frac{q}{p}$$

$$\text{Variance} = \frac{q}{p^2}$$