**Homework 2**

**Total Points 20**

1. The data below shows the number of prisoners who escaped from 10 prisons over a 5 year period. Hand Calculate the mean, median, mode, range, variance and standard deviation of the distribution

(SHOW YOUR WORK / CALCULATIONS)

**Mode = 5,7**

**Median=5**

**Mean= 4.8**

**Range=7**

**Variance=39.47**

**Standard deviation = 6.28**

X X

2 1 1

4 2 2

5 3 3

3 4 4

7 5 5

7 5 6

5 6 7

6 7 8

1 7 9

8 8 10

1. **MODE:**

The mode is defined **AS THE MOST FREQUENTLY OCCURRING SCORE** in a group of scores.

**Mode = 5,7**

1. **MEDIAN:**

( N + 1) / 2

Step 1: Arrange the numbers in ascending or descending order.

Step 2: Find the median position

**(10+1)/2 = (11)/2 = 5.5**

Step 3: So the median position is between 5th and 6th number

The number between the 5th and 6th number is 5 and 5

Therefore,

**Median=5**

1. **MEAN:**

**WE ARE ONLY DEALING WITH SAMPLE MEAN IN CHAPTER 4 AND 5.**

**A STATISTIC is a numerical descriptive measure of a sample.**

**Mean = ∑X / n**

**Where ∑X is the sum of all values in the data set and**

**n is the sample size.**

∑X=2+4+5+3+7+7+5+6+1+8= 48

N=10

Mean=48/10

**Mean= 4.8**

1. **RANGE**

Range simply is defined as the difference between the highest and the lowest value.

Range=8-1=7

**Range=7**

1. **Variance**

Step 1: Calculate mean.

(**Mean = ∑X / n Where ∑X is the sum of all values in the data set and**

**n is the sample size.)**

Mean=4.8

Step 2: Form a column to calculate (X - ). (X - ) is referred to as **mean deviation**

X (X - )

1 (1-4.8=--3.8)

2 (2-4.8=--2.8)

3 (3-4.8=-1.8)

4 (4-4.8-=-0.8)

5 (5-4.8=0.2)

5 (5-4.8=0.2)

6 (6-4.8=1.2)

7 (7-4.8=2.2)

7 (7-4.8=2.2)

8 (8-4.8=3.2)

Step 4: Find the total or ∑(X - )2

X (X - ) (X - )2

1 (1-4.8=--3.8) -3.82=11.24

2 (2-4.8=--2.8) -2.82=7.84

3 (3-4.8=-1.8) -1.82=3.24

4 (4-4.8-=-0.8) -0.82=0.64

5 (5-4.8=0.2) 0.22= 0.04

5 (5-4.8=0.2) 0.22= 0.04

6 (6-4.8=1.2) 1.22= 1.44

7 (7-4.8=2.2) 2.22=4.84

7 (7-4.8=2.2) 2.22=4.84

8 (8-4.8=3.2) 3.22=10.24

∑ =44.4

Step 5: calculate s2 = ∑(X - )2 / (n-1). We are calculating the sample variance so divide the total squared deviation by (n-1).

**NOTE: The symbol for population** variance is σ2 and for **the sample variance** it is **s2.**

s2 = ∑(X - )2 / (n-1)

s2 = 44.4/(10-1)

s2 =44.4/9

s2 =39.466=39.47

The calculated variance for this distribution is 39.47

**Variance=39.47**

1. **Standard Deviation**

Calculation of standard deviation is very simple.

Simply take the square root of the variance.

Standard deviation = s = √39.47 = 6.282=6.28

**Standard deviation = 6.28**

1. Hand calculate mean, median, mode, range, variance and standard deviation for the following distribution: X represents how many brothers and sisters do you have and f represents frequency (total N =338)

(SHOW YOUR WORK / CALCULATIONS)

X f fx cf

0 39 0 39

**1 137 137 176**

2 79 158 255

3 39 117 294

4 17 68 311

5 13 65 324

6 6 36 330

7 4 38 334

8 1 8 335

9 1 9 336

12 1 12 337

15 1 15 338

∑663 ∑338

(N =338)

1. **Mean =1.96**
2. **Median Position =168.5**
3. **Median =1**
4. **Mode =1**
5. **Range =15**
6. **Variance=3.16**
7. **Standard deviation=1.78**
8. **Mean:**

Mean = ∑X / n

Where ∑X is the sum of all values in the data set and

n is the sample size.

∑X=663

N=338

**Mean = ∑X / n**

Mean=663/338

Mean=1.961

**Mean=1.96**

1. **Median Position & C. Median**

Cumulative frequency is denoted by cf.

***The data is arranged in descending order.***

Create a cumulative frequency column.

X f Cumulative Frequency

15 1 39+137+79+39+17+13+6+4+1+1+1+1=338

12 1 39+137+79+39+17+13+6+4+1+1+1=337

9 1 39+137+79+39+17+13+6+4+1+1=336

8 1 39+137+79+39+17+13+6+4+1=335

7 4 39+137+79+39+17+13+6+4=334

6 6 39+137+79+39+17+13+6=330

5 13 39+137+79+39+17+13=324

4 17 39+137+79+39+17=311

3 39 39+137+79+39=294

2 79 39+137+79=255

1 137 39+137=176

0 39 39

*Once you get the cumulative frequency you have to find the mid point or median position.*

The median position can be find out by using the formula (N + 1) / 2

(N =338)

Therefore, (N + 1) / 2= (338+1)/2=337/2=168.5

Check where the 168.5 number lies in cumulative frequency.

**The Median Position is 168.5**

The 168.5 score is 1.

**SO median is 1.**

1. **Mode**

The mode is defined **AS THE MOST FREQUENTLY OCCURRING SCORE** in a group of scores.

The most frequently occurring score is 1 with a frequency of 137.

Mode=1

1. **Range**

Range simply is defined as the difference between the highest and the lowest value.

X f

0 39

**1 137**

2 79

3 39

4 17

5 13

6 6

7 4

8 1

9 1

12 1

15 1

Range=15-0=15

**Range=15**

1. **Variance**

Step 1: Calculate mean.

Mean = ∑X / n

Where ∑X is the sum of all values in the data set and

n is the sample size.

∑X=663

N=338

**Mean = ∑X / n**

Mean=663/338

Mean=1.961

**Mean=1.96**

Step 2: Form a column to calculate (X - ). (X - ) is referred to as **mean deviation**

X f (X - ). **(X-**)2 **f(X-**)2 **f(X-**)2

0 39 0-1.96= (-1.96) 3.84 39(3.84) 149.76

**1 137 1**-1.96= (-.96) 0.92 137(0.92) 126.04

2 79 2-1.96= 0.04 0 79(0) 0

3 39 3-1.96=1.04 1.08 39(1.08) 42.12

4 17 4-1.96=2.04 4.16 17(4.16) 70.72

5 13 5--1.96=3.04 9.24 13(9.24) 120.12

6 6 6-1.96=4.04 16.32 6(16.32) 97.92

7 4 7-1.96=5.04 25.40 4(25.4) 101.60

8 1 8-1.96=6.04 36.48 1(36.48) 36.48

9 1 9-1.96=7.04 49.56 1(49.56) 49.56

12 1 12-1.96=10.04 100.80 1(100.80) 100.80

15 1 15-1.96=13.04 170.04 1(170.04) 170.04

**∑=1065.16**

(N =338)

Variance = ∑f(X-)2 / (n-1)

Variance = 1065.16/ (338-1)

Variance = 1065.16/337

Variance = 3.160=3.16

1. **Standard deviation**

Standard deviation = square root of 3.16

Standard deviation = 1.777=1.78

1. Hand calculate mean, median, mode, range, variance and standard deviation for the following distribution.

(SHOW YOUR WORK / CALCULATIONS)

|  |
| --- |
| Parking Tickets (X) |
| 5 |
| 3 |
| 0 |
| 1 |
| 0 |
| 3 |
| 3 |
| 1 |

**Mean=2**

**Mode=3**

**Median=2**

**Range=5**

**Variance=4.71**

**Standard deviation=**2.17

**Mean:**

**WE ARE ONLY DEALING WITH SAMPLE MEAN IN CHAPTER 4 AND 5.**

**A STATISTIC is a numerical descriptive measure of a sample.**

**Mean = ∑X / n**

**Where ∑X is the sum of all values in the data set and**

**n is the sample size.**

**∑X=5+3+0+1+0+3+3+1=16**

**N=8**

**∑X / n=16/8=2**

**Mean=2**

**Mode:**

The mode is defined **AS THE MOST FREQUENTLY OCCURRING SCORE** in a group of scores.

Mode=3

**Median:**

Step 1: Arrange the numbers in ascending or descending order.

|  |
| --- |
| Parking Tickets (X) |
| 5 |
| 3 |
| 0 |
| 1 |
| 0 |
| 3 |
| 3 |
| 1 |
|  |

Parking Tickets:

5

3

3

3

1

1

0

0

Step 2: Find the median position

( N + 1) / 2

N=8

(8+1)/2

9/2

4.5

Median Position = 4.5

Step 3: So the median position is between 4th and 5th number

The number between the 4th and 5th number is 2

Therefore,

**Median=2**

**Range:**

Range simply is defined as the difference between the highest and the lowest value.

Range=5-0=5

Range=5

**Variance:**

Step 1: Calculate mean.

(Mean = ∑X / n Where ∑X is the sum of all values in the data set and

n is the sample size.)

∑X=5+3+0+1+0+3+3+1=16

N=8

∑X / n=16/8=2

**Mean=2**

Step 2: Form a column to calculate (X - ). (X - ) is referred to as **mean deviation**

Parking Tickets: (X - )

5 5-2=3

3 3-2=1

3 3-2=1

3 3-2=1

1 1-2=(-1)

1 1-2=(-1)

0 0-2=(-2)

0 0-2=(-2)

Step 4: Find the total or ∑(X - )2

Parking Tickets: (X - ) ∑(X - )2

5 5-2=3 32=9

3 3-2=1 12=1

3 3-2=1 12=1

3 3-2=1 12=1

1 1-2=(-1) -12=1

1 1-2=(-1) -12=1

0 0-2=(-2) -22=4

0 0-2=(-2) -22=4

∑=33

Step 5: calculate s2 = ∑(X - )2 / (n-1).

s2 = ∑(X - )2 / (n-1).

s2 =33/(n-1)

s2 =33/(8-1)

s2 =33/7

s2 =4.714=4.71

Variance=4.71

**Standard deviation**

Calculation of standard deviation is very simple.

Simply take the square root of the variance.

Standard deviation = √4.71=2.170=2.17

Standard deviation =2.17

1. Let us consider the following number of delinquent acts committed by a sample of 20 junior high school students. Calculate mode, median, mean range, variance and standard deviation.

(SHOW YOUR WORK / CALCULATIONS)

|  |  |  |
| --- | --- | --- |
| Delinquent Acts (X) | Frequency (f) | F (X) |
| 5 | 3 | 5(3)=15 |
| 4 | 6 | 4(6)=24 |
| 2 | 9 | 2(9)=18 |
| 0 | 2 | 0(2)=0 |

**Mean=2.85**

**Median=2**

**Mode=2**

**Variance = 2.6**

**Standard deviation=0.51**

**Range=5**

**Mean:**

Mean = ∑X / n

Where ∑X is the sum of all values in the data set and

n is the sample size.

**X Y f(x) c(f)**

**5 3 15 17+3=20**

**4 6 24 11+6=17**

**2 9 18 2+9=11**

**0 2 0 0+2=2**

**57 20**

∑X=57

N=20

Mean = ∑X / n

Mean = 57/20=2.85

Mean=2.85

**Median**

*Once you get the cumulative frequency you have to find the mid point or median position.*

The median position can be find out by using the formula (N + 1) / 2

Therefore, (N + 1) / 2=

N=20

(20+1)/2=

21/2=

10.5

Check where the 10.5 number lies in cumulative frequency.

**X Y f(x) c(f)**

**5 3 15 17+3=20**

**4 6 24 11+6=17**

**2 9 18 2+9=11**

**0 2 0 0+2=2**

**57 20**

**The Median Position is 10.5**

The 10.5 score is 2.

**SO median is 2.**

1. **Mode**

The mode is defined **AS THE MOST FREQUENTLY OCCURRING SCORE** in a group of scores.

The most frequently occurring Delinquent Acts (X) score is 2 with a frequency of 9.

**Mode=2**

**Variance**

Step 1: Calculate mean.

Mean = ∑X / n

Where ∑X is the sum of all values in the data set and

n is the sample size.

**X Y f(x) c(f)**

**5 3 15 17+3=20**

**4 6 24 11+6=17**

**2 9 18 2+9=11**

**0 2 0 0+2=2**

**57 20**

∑X=57

N=20

Mean = ∑X / n

Mean = 57/20=2.85

**Mean=2.85**

Step 2: Form a column to calculate (X - ). (X - ) is referred to as **mean deviation**

X Y f(x) c(f) (X - )

5 3 15 17+3=20 5-2.85=2.15

4 6 24 11+6=17 4-2.85=1.15

2 9 18 2+9=11 2-2.85=(-0.85)

0 2 0 0+2=2 0-2.85=(-2.85)

57 20

Step 4: Find the total or ∑(X - )2

X Y f(x) c(f) (X - ) ∑(X - )2

5 3 15 17+3=20 5-2.85=2.15 2.142=4.579=4.58

4 6 24 11+6=17 4-2.85=1.15 1.152=1.322=1.32

2 9 18 2+9=11 2-2.85=(-0.85) (-0.85) 2=0.722=0.72

0 2 0 0+2=2 0-2.85=(-2.85) (-2.85) 2=8.122=8.12

57 20 ∑=14.74

N=57

Variance = ∑f(X-)2 / (n-1)

Variance=14.74/(57-1)

Variance=14.74/56

Variance=0.263

**Variance=0.26**

1. **Standard deviation**

Standard deviation = square root of 0.26

Standard deviation =0.509

**Standard deviation =0.51**

1. **Range**

Range simply is defined as the difference between the highest and the lowest value.

|  |  |  |
| --- | --- | --- |
| Delinquent Acts (X) | Frequency (f) | F (X) |
| 5 | 3 | 5(3)=15 |
| 4 | 6 | 4(6)=24 |
| 2 | 9 | 2(9)=18 |
| 0 | 2 | 0(2)=0 |

The highest value of x is 5 and the lowest value of x is 0.

Therefore, 5-0=5

**The Range is 5.**

**Excel**

**Points 15**

**MEASURES OF CENTRAL TENDENCY** *are helpful in summarizing important feature namely* ***how typical the middle score is****, in a given distribution*.

We also need some information about how the score differ from measures that reflect how different or variable scores are from each other or from some central score are called **MEASURES OF DISPERSION**, ***because they tell us how dispersed or scattered the scores in a distribution are.***

﻿“To calculate descriptive statistics, open the NCSD data set file. As outlined in the Data Dictionary NCSD file, the W3 variable refers to “Worry: Getting mugged.” If we want to produce descriptive statistics for W3, we need to highlight the data in the column E. Click on E2, hold the shift key, and press the arrow key down until all cells through E1006 are highlighted. Once the data are selected, click on “Data” and then on “Data analysis.” Choose “Descriptive Statistics” and a pop-up screen will appear (Figure 4.11). Instructions on how to add in Data Analysis are provided in chapter 2, p. 43. Enter E2:E1006 in “Input Range” and check on “Summary statistics.” Excel will generate a summary column in a separate sheet as shown below. In this table, Excel provides central tendency information including mean, mode, median, and sum. To find the frequency distribution information, refer back to instructions provided in chapter 3” (Vito. 78).

1. With State Data Set I use Excel to get measures of central tendency and dispersion.

Calculate mean, median, mode, range, variance and standard deviation using Excel

1. Number of Juveniles Arrested (under 18) 1997 (Variable name juvarr)
2. Number of Adults Arrested (under 18) 1997 (Variable name adultarr)
3. Number of Juveniles arrested violent crime (under 18) 1997 (Variable name juvarrv)
4. Number of Adults arrested for violent crime, 1997 (variable name aduarrv)

Paste the Excel output

Data: [www.waveland.com](http://www.waveland.com) – Student resources – End of chapter Excel data sets – 3b State Data for Chapters 3, 4 and 5

1. Number of Juveniles Arrested (under 18) 1997 (Variable name juvarr)



1. Number of Adults Arrested (under 18) 1997 (Variable name adultarr)



1. Number of Juveniles arrested violent crime (under 18) 1997 (Variable name juvarrv)



1. Number of Adults arrested for violent crime, 1997 (variable name aduarrv)

|  |  |
| --- | --- |
| *ADURRV* | |
|  |  |
| Mean | 9019.36957 |
| Standard Error | 2833.39747 |
| Median | 5071 |
| Mode | #N/A |
| Standard Deviation | 19217.0366 |
| Sample Variance | 369294496 |
| Kurtosis | 36.6783329 |
| Skewness | 5.79562582 |
| Range | 129814 |
| Minimum | 152 |
| Maximum | 129966 |
| Sum | 414891 |
| Count | 46 |

1. Using the following table on Number of robberies reported to the Miami-Dade Police Department, 2001 – 2005 calculate measures of central tendency and measures of dispersion using Excel. Enter the data in Excel and calculate Mean, median, mode, variance, standard deviation and range for # of visitor Robberies Reported (X)

**Year # of visitor Robberies Reported (X)**

**2001 25**

**2002 10**

**2003 16**

**2004 30**

**2005 12**

**Copy and paste the Excel output**

|  |  |
| --- | --- |
|  |  |
| Mean | 18.6 |
| Standard Error | 3.84187454 |
| Median | 16 |
| Mode | #N/A |
| Standard Deviation | 8.59069264 |
| Sample Variance | 73.8 |
| Kurtosis | -2.0372574 |
| Skewness | 0.52745052 |
| Range | 20 |
| Minimum | 10 |
| Maximum | 30 |
| Sum | 93 |
| Count | 5 |

1. Using the following table on Number of reported Homicides, 1995 to 2000, calculate measures of central tendency and measures of dispersion using Excel. Calculate Mean, median, mode, variance, standard deviation and range of # of Reported Homicides (X)

**Year # of Reported Homicides (x)**

**1995 99**

**1996 86**

**1997 58**

**1998 58**

**1999 48**

**2000 50**

Copy and paste Excel output

