

## Ch4

Ex) #1. Consider the following data. The random variable X represent the number of days patients stayed in the hospital. Calculate Relative Frequency Probabilities.

| Number of days stayed(X) | Frequency |
|--------------------------|-----------|
| 3                        | 15        |
| 4                        | 32        |
| 5                        | 56        |
| 6                        | 19        |
| 7                        | 5         |

- #2. Simulate 50 births, where each birth results in a boy or girl.
- Sort the results, count the number of girls.
  - Based on the result in (a), calculate the probability of getting a girl when a baby is born.
  - The probability obtained in (b) is likely to be different from 0.5. Does this suggest that the computer's random number generator is defective? Why or why not?
- #3. a) Simulate rolling a single die by generating 5 integers (5 trials) between 1 and 6. Count the number of 3s that occurred and divide that number by 5 to get the empirical probability. Based on 5 trials, what is the probability of getting 3s?
- Repeat (a) for 25 trials. What is the probability of getting 3s?
  - Repeat (a) for 50 trials. What is the probability of getting 3s?
  - Repeat (a) for 100 trials. What is the probability of getting 3s?
  - Repeat (a) for 500 trials. What is the probability of getting 3s?
  - In your own words, generalize these results in a restatement of the Law of Large Numbers.

### How to Do it:

#### Calculate Relative Frequency Probabilities from the table

- In C1 enter the values of **X** and name the column **X**.
- In C2 enter the frequencies and name the column **f**.
- Select **Calc>Calculator**.
- Type **Px** in the box for Store result in variable:.
- Click in the Expression box, then double-click C2 f and click the division operator.
- Find **Sum** in the function list and click Selecton.
- Double-click C2 f . (You should see '**f**/**Sum**('f') in the Expression box.)
- Click OK.

#### Calculate Relative Frequency Probabilities from the original data in a worksheet.

- Click **Stat>Tables>Tally Individual Variables**
- Select the column needed (i.e. C1) in the Variables:
- For Display click Counts and Percents.

#### Random Data

- Click on **Calc>Random Data>Integer**
- You will generate\_\_\_\_\_ rows of data (enter the amount of rows you want)
- Store in a column (C1) or columns (C1-C10).
- Minimum Value: enter minimum, for instance, the minimum for of a die is 1
- Maximum Value: for instance, the maximum number for a die is 6

## Ch5

Ex) #1. It is known that 5% of the population is afraid of being alone at night. If a random sample of 20 Americans is selected, what is the probability that exactly 5 of them are afraid?

#2. If 63% of all women are employed outside the home,

- (a) create a binomial table for a sample of 20 women;
- (b) use the binomial table to find the probability that at least 10 are employed outside the home.

#3. For the following probability distribution table,

- (a) find the mean and standard deviation;
- (b) draw a graph for the distribution.

|      |      |      |      |      |      |
|------|------|------|------|------|------|
| x    | 0    | 1    | 2    | 3    | 4    |
| P(x) | 0.12 | 0.20 | 0.31 | 0.25 | 0.12 |

### Calculating Binomial Probability

1. Select **Calc>Probability Distribution>Binomial**.
2. Click Probability.
3. Enter Number of trials and probability of success.
4. Click the Input constant and type in the value.
5. Click **OK**

### Constructing a Binomial Distribution

1. Select **Calc>Make Patterned Data>Simple Set of Numbers**.
2. Store patterned data in C1.
3. Enter the value of 0 for the first value and enter the value of **n** for the last value.
4. Click OK
5. Select **Calc>Probability Distributions>Binomial**.
6. Choose Probability for the option and enter the value for the number of trials.
7. Enter the value for the Probability of success.
8. Check the button for Input columns then enter the column (i.e. C1).
9. Enter the desired column for Optional Storage.
10. Click OK.

### Finding Exact Values of $\mu$ and $\sigma$ for a probability Distribution

1. Enter the values of **X** in C1 and enter the corresponding probabilities in C2.
2. Select **Calc>Calculator**
3. For  $\mu$ , enter the expression **sum(C1 \* C2)** and store result in C3.
4. Select **Calc>Calculator**
5. For  $\sigma$ , enter the expression **sqrt(sum(C1\*\*2\*C2)-C3\*\*2)** and store result in C4.

### Graph a Binomial Distribution

1. Select **Graph>Scatterplot**, then Simple.
2. Double-click on C2 Ps or the Y variable and C1 X for the X variable.
3. Click [Data view], then Project lines, then [OK]. Deselect any other type of display that may be selected in this list.
4. Type an appropriate title, such as Binomial Distribution n=20, p=0.05
5. Press Tab to the Subtitle 1, (you may type in Your Name)
6. Optional: Click [Scales] then [Gridlines] then check the box for Y major ticks.
7. Click [OK] twice.

## Ch6

Ex) #1. Central Limit Theorem

Simulate 200 random samples of size 25 from a normally distributed population with a mean of 56 and a standard deviation of 12.

- (a) Determine the sample mean of each of the 200 samples.
- (b) Construct a histogram of the 200 sample means. Does it appear to be normally distributed?
- (c) Compute descriptive statistics for the 200 sample means.
- (d) What is the mean of the 200 sample means? Is it close to the population mean? Should it be?
- (e) What is the standard deviation of the 200 sample means? Is it close to the population standard deviation? Should it be?

#2. Find the Area to the left of  $Z=1.39$

#3. Find the z value for a cumulative probability of 0.025.

### How to Do it:

#### **Generating N Random Samples of size K from a Normally Distributed Population**

##### **1. Calc>Random Data>Normal**

2. Generate \_\_\_\_\_ rows of data (enter the number of samples)
3. Store in a column C1-CK (Here **K** is the sample size)
4. Enter Mean and Standard Deviation.
5. Click OK.

#### **Checking Data for Normality:**

- Click on **Stat > Basic Statistics > Display Descriptive Statistics**
- Double-click on the column intended as the “Variables”
- Click **Graphs** and then click the **Histogram of data, with normal curve**, check the box
- Click **OK** twice

#### **Determining Cumulative Probabilities for Normal Distribution:**

- Click on **Calc > Probability Distributions > Normal**
- Chose Cumulative probability
- Type in **Mean and the Standard Deviation**
- Check **Input Constant**, enter in the number
- Click **OK**

#### **Inverse Cumulative Probabilities for Normal Distribution:**

- **Calc > Probability Distributions > Normal**
- Check in the box **Inverse Cumulative Probability**
- Type in **Mean and the Standard Deviation**
- Check **Input Constant**, enter in the number
- Click **OK**

**Answer Key Ch4-Ch6 Handout**

**Ch4**

#1. Relative Frequency Probabilities.

| Number of days stayed(X) | Frequency(f) | Px       |
|--------------------------|--------------|----------|
| 3                        | 15           | 0.118110 |
| 4                        | 32           | 0.251969 |
| 5                        | 56           | 0.440945 |
| 6                        | 19           | 0.149606 |
| 7                        | 5            | 0.039370 |

#2. Simulate 50 births, where each birth results in a boy or girl.

**Tally for Discrete Variables: C1**

| C1 | Count | Percent |
|----|-------|---------|
| 0  | 29    | 58.00   |
| 1  | 21    | 42.00   |
| N= | 50    |         |

- a) 21Girls(Answers vary.) 1s for baby girls and 0s for baby boys
- b)  $21/50=42$
- c) No, 0.5 is only the theoretical probability. As the number of trials increases, the empirical probability of an event will approach the theoretical probability.

#3. a)

| C2 | Count | Percent |
|----|-------|---------|
| 2  | 1     | 20.00   |
| 3  | 1     | 20.00   |
| 4  | 1     | 20.00   |
| 6  | 2     | 40.00   |
| N= | 5     |         |

**Tally for Discrete Variables: C3**

b)

| C3 | Count | Percent |
|----|-------|---------|
| 1  | 6     | 24.00   |
| 2  | 4     | 16.00   |
| 3  | 5     | 20.00   |
| 4  | 1     | 4.00    |
| 5  | 5     | 20.00   |
| 6  | 4     | 16.00   |
| N= | 25    |         |

c)

**Tally for Discrete Variables: C4**

| C4 | Count | Percent |
|----|-------|---------|
| 1  | 10    | 20.00   |
| 2  | 11    | 22.00   |
| 3  | 6     | 12.00   |
| 4  | 8     | 16.00   |
| 5  | 6     | 12.00   |
| 6  | 9     | 18.00   |
| N= | 50    |         |

d)

**Tally for Discrete Variables: C5**

| C5 | Count | Percent |
|----|-------|---------|
| 1  | 22    | 22.00   |

Math 227 Elementary Statistics-Minitab Handout Ch4, Ch5, and Ch6

|    |     |       |
|----|-----|-------|
| 2  | 15  | 15.00 |
| 3  | 11  | 11.00 |
| 4  | 13  | 13.00 |
| 5  | 23  | 23.00 |
| 6  | 16  | 16.00 |
| N= | 100 |       |

e)

**Tally for Discrete Variables: C6**

| C6 | Count | Percent |
|----|-------|---------|
| 1  | 81    | 16.20   |
| 2  | 74    | 14.80   |
| 3  | 74    | 14.80   |
| 4  | 84    | 16.80   |
| 5  | 105   | 21.00   |
| 6  | 82    | 16.40   |
| N= | 500   |         |

**(Extra for N=1000)**

**Tally for Discrete Variables: C7**

| C7 | Count | Percent |
|----|-------|---------|
| 1  | 160   | 16.00   |
| 2  | 154   | 15.40   |
| 3  | 167   | 16.70   |
| 4  | 168   | 16.80   |
| 5  | 177   | 17.70   |
| 6  | 174   | 17.40   |
| N= | 1000  |         |

f) **Law of large numbers:** when a probability experiment is repeated a large number of times, the relative frequency probability of an outcome will approach its theoretical probability.

**Ch5**

**#1. Calculating Binomial Probability**

Binomial with  $n = 20$  and  $p = 0.05$

$x$   $P(X = x)$   
 5 0.0022446 (0.002 using the binomial distribution table)

**2. Constructing a binomial Distribution**

a. Binomial with  $n = 20$  and  $p = 0.63$

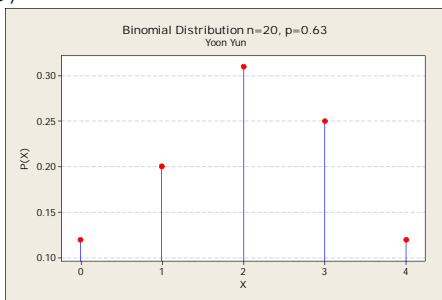
| X  | P(X)     |
|----|----------|
| 0  | 0.000000 |
| 1  | 0.000000 |
| 2  | 0.000001 |
| 3  | 0.000013 |
| 4  | 0.000094 |
| 5  | 0.000513 |
| 6  | 0.002184 |
| 7  | 0.007437 |
| 8  | 0.020578 |
| 9  | 0.046719 |
| 10 | 0.087503 |
| 11 | 0.135446 |
| 12 | 0.172969 |
| 13 | 0.181240 |
| 14 | 0.154299 |
| 15 | 0.105090 |
| 16 | 0.055918 |
| 17 | 0.022403 |
| 18 | 0.006358 |
| 19 | 0.001139 |
| 20 | 0.000097 |

b.  $P(\text{at least 10 are employed outside the home}) = P(X=10) + P(X=11) + P(X=12) + P(X=13) + P(X=14) + P(X=15) + P(X=16) + P(X=17) + P(X=18) + P(X=19) + P(X=20) = 0.922462$

**3. a.**

| X | P(X) | Mean | SD      |
|---|------|------|---------|
| 0 | 0.12 | 2.05 | 1.18638 |
| 1 | 0.20 |      |         |
| 2 | 0.31 |      |         |
| 3 | 0.25 |      |         |
| 4 | 0.12 |      |         |

**b)**

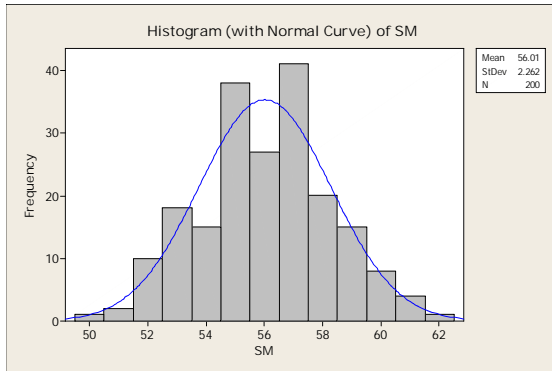


**Ch6**

1. Central Limit Theorem

a. C26

b.



Yes, it appears to be normally distributed.

c.

**Descriptive Statistics: SM**

| Variable | N   | N* | Mean   | SE Mean | StDev | Minimum | Q1     | Median | Q3     |
|----------|-----|----|--------|---------|-------|---------|--------|--------|--------|
| SM       | 200 | 0  | 56.012 | 0.160   | 2.262 | 50.125  | 54.564 | 56.147 | 57.455 |

| Variable | Maximum |
|----------|---------|
| SM       | 62.042  |

d.  $\mu_{\bar{x}} = 56.021$

Yes, it is close to  $\mu = 56$ . yes, according to the Central Limit Theorem,  $\mu_{\bar{x}} = \mu$

e.  $\sigma_{\bar{x}} = 2.262$  No, it is not close to  $\sigma = 12$ . However, it is close to  $\frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{25}} = 2.4$ . According to the Central Limit

Theorem,  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

2. Cumulative Distribution Function

Normal with mean = 0 and standard deviation = 1

| x    | P( X <= x ) |
|------|-------------|
| 1.39 | 0.917736    |

3. Inverse Cumulative Distribution Function

Normal with mean = 0 and standard deviation = 1

| P( X <= x ) | x        |
|-------------|----------|
| 0.025       | -1.95996 |