

## **Unit 3**

### **Randomness in Data**

# Topic 11

## Probability

### In-Class Activities

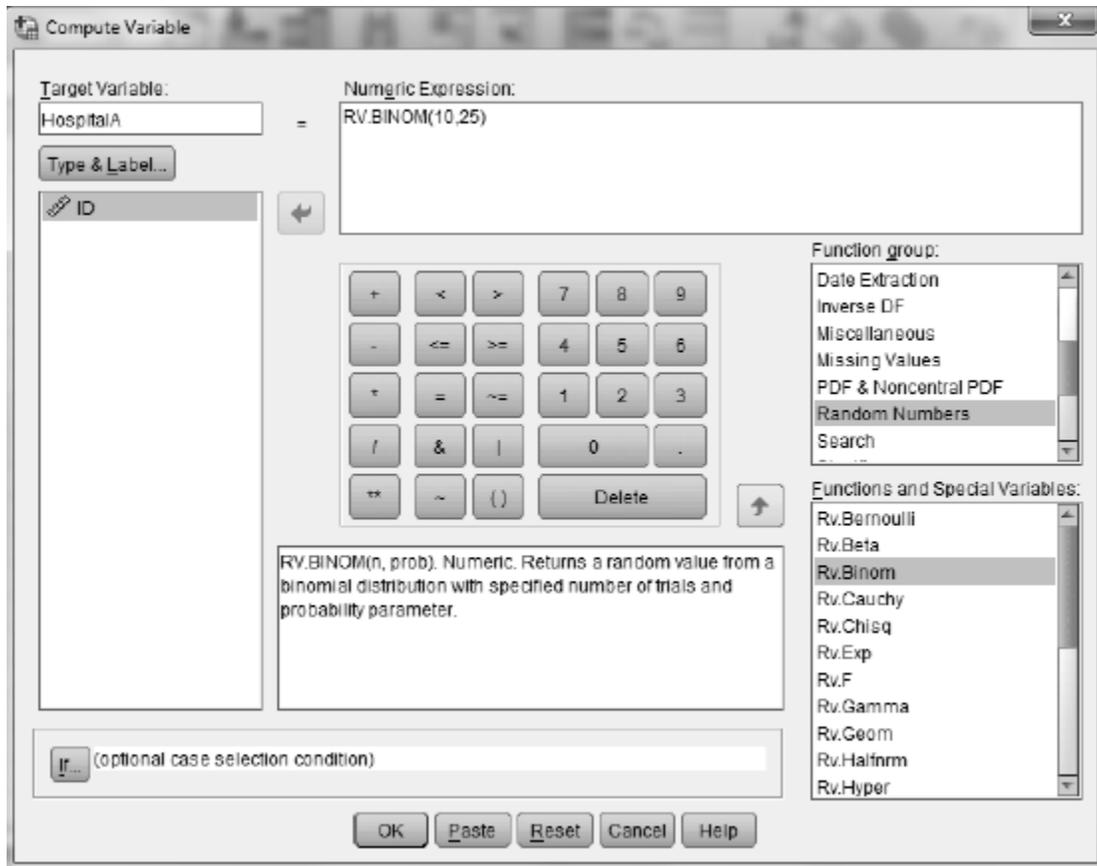
There are no in-class activities that require the use of SPSS Statistics.

### Exercises

#### Exercise 11-18

#### 11-17, 11-18

- a. Generate the 1000 ice cream prices as follows:
  1. Open a new Data view. Select **File > Open > Syntax** and open the SPSS syntax file **CreateID.SPS**. Highlight all the lines in the file and run it by clicking on the green right pointing arrow. This creates a new Data view with 1000 cases in it. The variable **IDs** is a dummy variable that just tells SPSS that there are 1000 cases.
  2. Select **Transform > Compute Variable** to open the dialog box shown below. Make the **Target Variable** *RV*. Select **Random Numbers** in the **Function group** box. Select **Rv.Uniform** in the **Functions and Special Variables** box and move it to the **Numeric Expression** box. Make the **Numeric Expression** box contain  $Rv.Uniform(0,6)$ . Click **OK** to make a variable containing randomly generated values equally spread out over the range 0 to 6. Make sure the measure of this variable is scale.
  3. Select **Transform > Compute Variable** and enter **Toss1** in the **Target Variable** box. Enter  $1+(RV \geq 1)+(RV \geq 2)+(RV \geq 3)+(RV \geq 4)+(RV \geq 5)$ . Click **OK** to create the results of the first toss. Make sure the measure of this variable is scale.
  4. Repeat part 2. When asked if you want to change the existing variable, click Yes.
  5. Repeat part 3 entering *Toss2* in the **Target Variable** box. Make sure the measure of this variable is scale.
  6. Select **Transform > Compute Variable**. Enter *Price* in the **Target Variable** box and enter  $(Toss1 > Toss2)*(10*Toss1+Toss2) + (Toss1 \leq Toss2)*(10*Toss2 + Toss1)$ . This variable contains the ice cream price. Make sure the measure of this variable is scale.



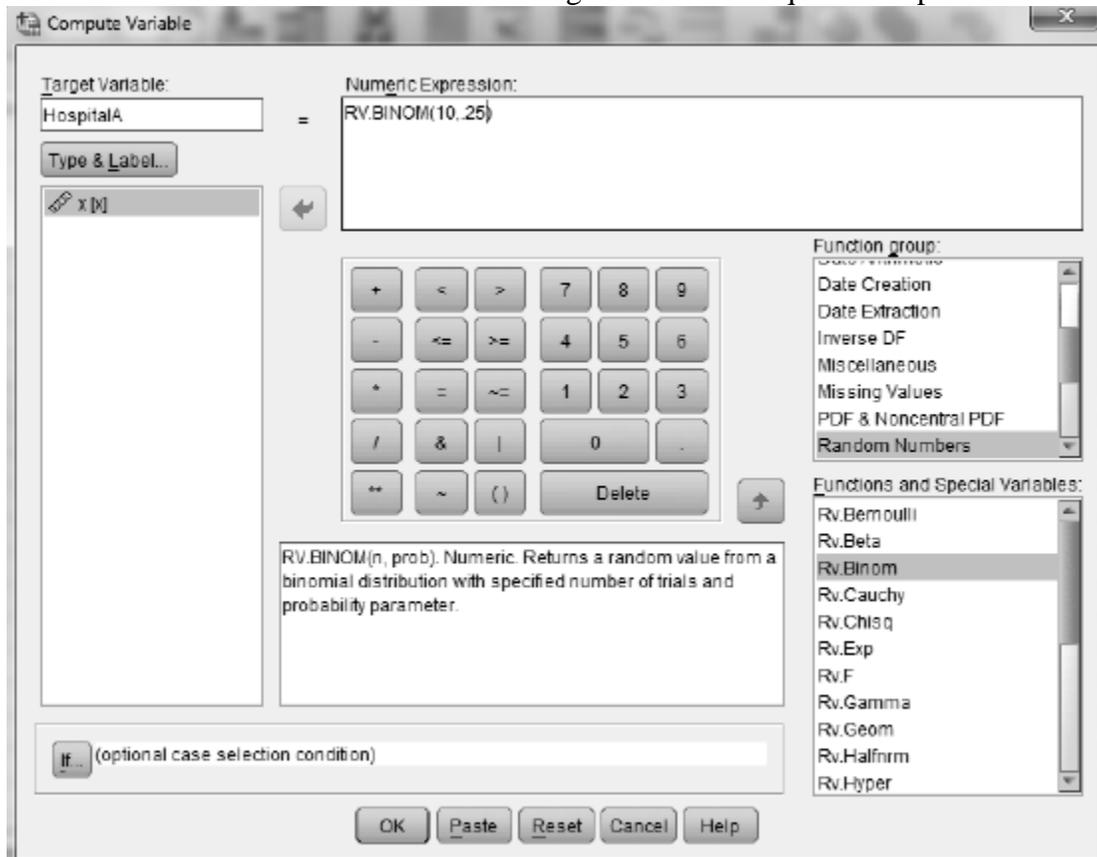
7. Use the Chart Builder to draw a histogram of *Price*. Make the vertical axis be percents.
- b. The mean *Price* appears with the histogram. Is the average price close to the expected value? Should it be? Explain.

### Exercise 11-19: Hospital Births

Generate the 365 days of births for the two hospitals requested in part a as follows.

1. Open a new Data view. Select **File > Open > Syntax** and open the SPSS syntax file **CreateID.SPS**. In the line *LOOP ID = 1 TO 1000*, change the 1000 to 365. Highlight all the lines in the file and run it by clicking the right pointing green arrow. This creates a new Data view with 365 cases. The variable *ID* is a dummy variable used to set the number of cases.
2. Select **Transform > Compute Variable** to open the dialog box shown below.
3. Enter *HospitalA* in the **Target Variable** box.
4. Select **Random Numbers** in the **Function group** box.
5. Move **Rv.Binom** to the **Numeric Expression** box.
6. Make the numeric expression be *Rv.Binom(10,.25)* as shown above.
7. Click **OK** to create a variable containing the births at Hospital A. Make sure the measure of this variable is scale

8. Repeat parts 2-7 to create a variable containing the births at Hospital B. Enter *HospitalB* in the **Target Variable** box and enter *Rv.Binom(50,.25)* in the **Numeric Expression** box. Make sure the measure of this variable is scale.
9. Use the Chart Builder to create the histograms that are requested in part a.

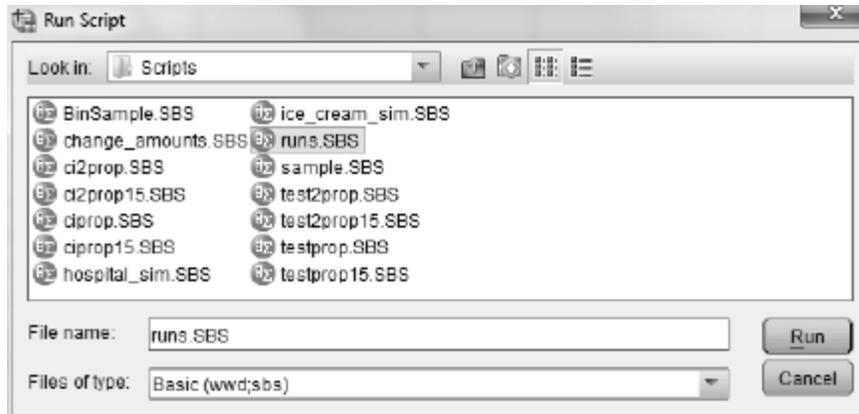


Complete parts b-e as requested in your main textbook.

### Exercise 11-21: Runs and “Hot” Streaks

Complete part a as directed in your main textbook. For part b proceed as follows:

1. Open a new Data view and create an empty variable *Long\_Run*.
2. Select **Utilities > Run Script** to open the dialog box shown below.
3. Locate and highlight the SPSS script file **Runs.SBS** and click **Run**.



The output from the script contains enough information to complete parts c-g.

## Topic 12

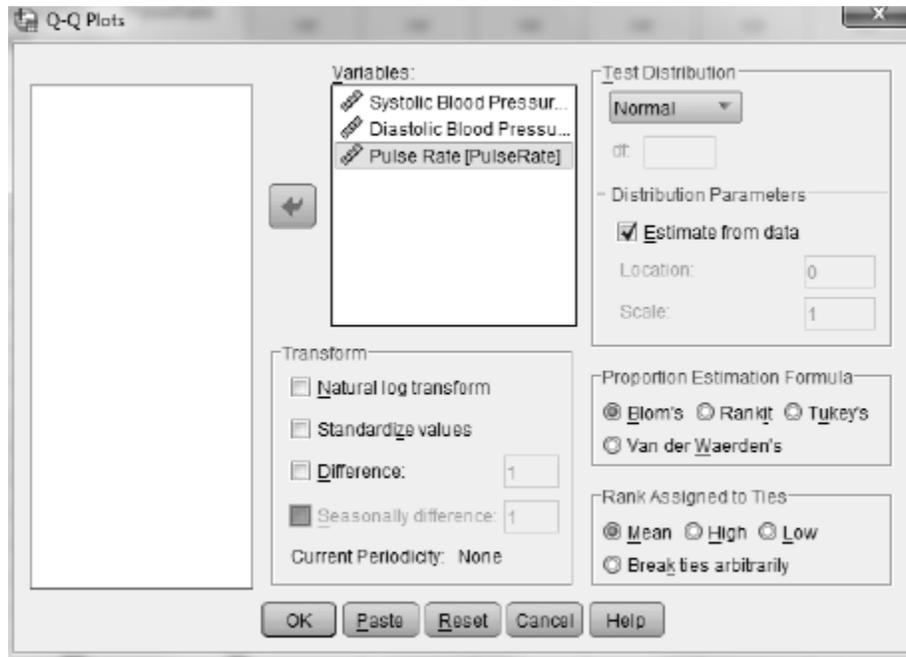
### Normal Distributions

#### In-Class Activities

##### Activity 12-3: Blood Pressure and Pulse Rate Measurements

Complete parts a and b as directed in your main textbook. Follow the instructions below to have SPSS create the requested normal probability plots.

1. Open the SPSS data file `BloodPressures.SAV`.
2. Select **Analyze > Descriptive Statistics > Q-Q Plots** to open the dialog box shown below.



3. Move the three variables to the **Variables** box.
4. Click **OK** and the normal probability plots will be drawn in an Output window.

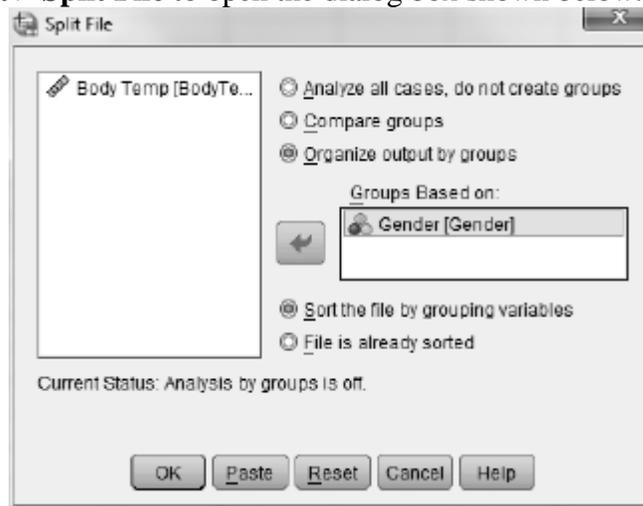
## Exercises

### Exercise 12-19: Body Temperatures

**12-1, 12-19, 14-3, 14-18, 15-9, 19-7, 19-8, 20-11, 22-10, 23-3**

Use SPSS to produce the histograms and normal probability plots requested in parts a and b. The data are stored in the SPSS file `BodyTemps.SAV`. To create separate graphs for men and women you will need to split the file. Follow the instructions given below.

1. Select **Data > Split File** to open the dialog box shown below.



2. Click **Organize output by groups** and move the variable *Gender* to the **Groups Based on** box.
3. Click **OK** and all subsequent analysis will be done separately for men and for women.
4. If you want later to analyze the whole file return to the **Split File** dialog box and click **Analyze all cases, do not create groups** and click **OK**.

### Exercise 12-20: Natural Selection

**10-1, 10-6, 10-7, 12-20, 22-21, 23-3**

Use SPSS to create the normal probability plots requested in part a. See Exercise 12-19 for instructions on splitting the file. The data are stored in the SPSS file `Bumpus.SAV`.

### Exercise 12-21: Honda Prices

**7-10, 10-19, 12-21, 28-14, 28-15, 29-10, 29-11**

Use SPSS to create the requested normal probability plots. The data are stored in the SPSS file `HondaPrices.SAV`.

## **Topic 13**

### **Sampling Distributions: Proportions**

There are no activities in this topic that require the use of SPSS statistics. You can use the `Reeses Pieces` applet for those parts that tell you to use technology. Just disregard the candy context of the applet.

## Topic 14

### Sampling Distributions: Means

#### In-Class Activities

##### Activity 14-1: Coin Ages

12-16, 14-1, 14-2, 19-15

- h. Enter the data into SPSS and use SPSS to compute the mean and standard deviation.
- j. Use SPSS to produce the requested dotplot.

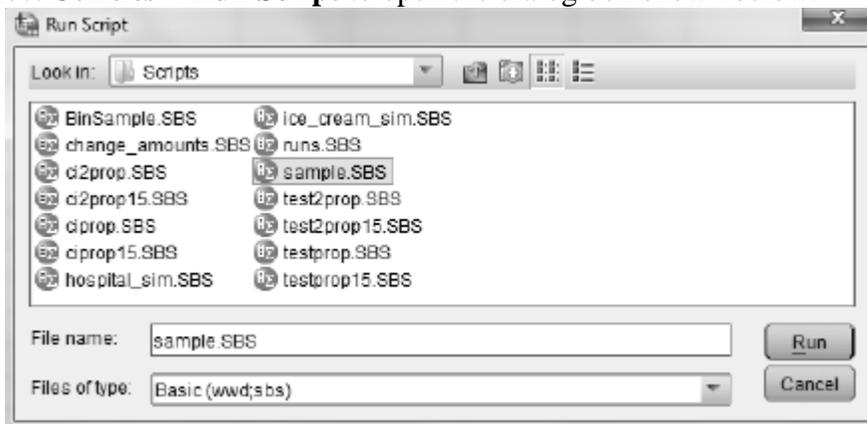
#### Exercises

##### Exercise 14-5: Heart Rates

Use SPSS to draw the normal probability plot requested in part b. See Activity 12-3 for instructions. The data are stored in the SPSS file `HeartRate.SAV`.

Use the SPSS script `Sample.SBS` to draw the samples requested in parts c and e.

1. Select **Utilities > Run Script** to open the dialog box shown below.



2. Highlight *sample.SBS* and click **Run**.
3. Enter the desired number of samples (1000) and the desired sample size (3 in part c and 10 in part e)
4. Click **OK** to run the script. Be patient. It will take some time to complete.

# Topic 15

## Central Limit Theorem and Statistical Inference

### In-Class Activities

There are no in-class activities that require the use of SPSS Statistics.

#### Exercises 15-9: Body Temperatures

**12-1, 12-19, 14-3, 14-18, 15-9, 19-7, 19-8, 20-11, 22-10, 23-3**

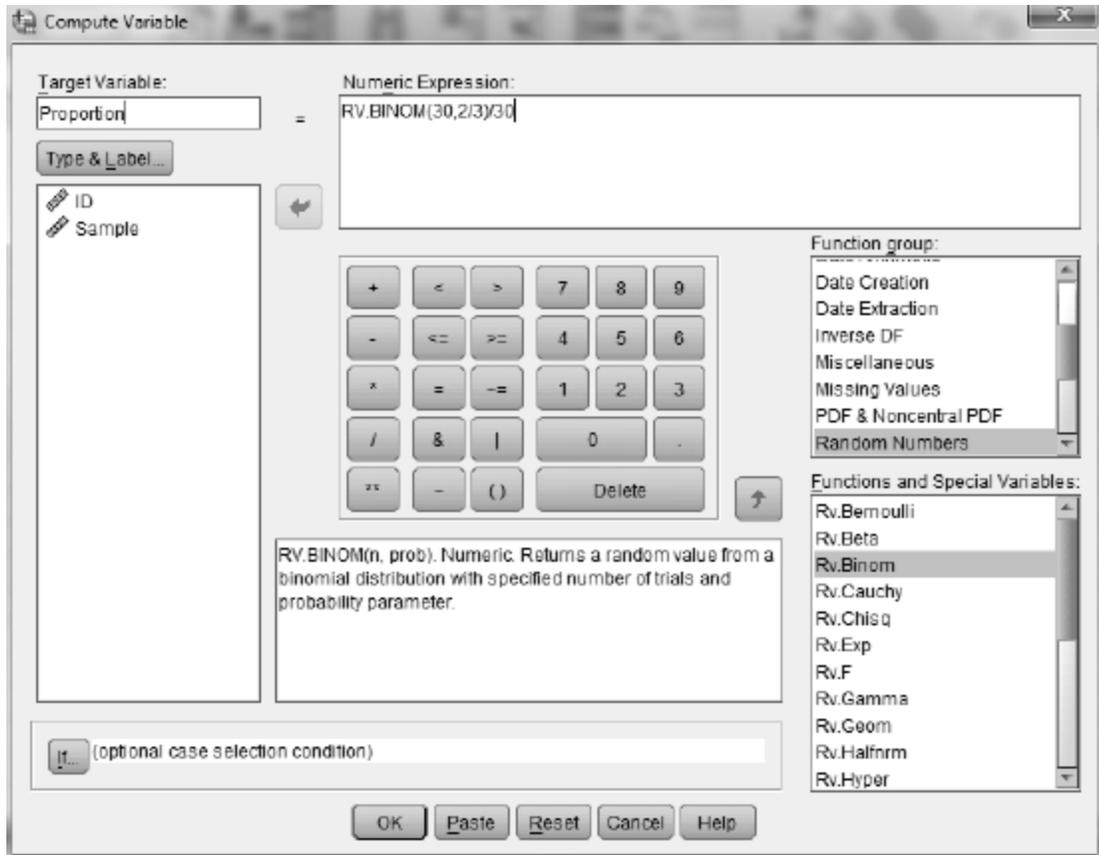
Use SPSS to create the graphical displays requested in part a and to compute the mean and standard deviation requested in part b.

#### Exercise 15-13: Distinguishing Colas

**13-13, 15-13, 17-24, 18-9**

Do the following to simulate the samples requested in part c.

1. Open a new Data view.
2. Run the syntax file `CreateID.SPS`. See Activity 11-18 for instructions.
3. Select **Transform > Compute Variable** to open the dialog box shown below.
4. Enter *Proportion* in the **Target Variable** box.
5. Highlight *Random Numbers* in the **Function group** box.
6. Select *Rv.Binom* in the **Functions and Special Variables** box and move it to the **Numeric Expression** box.
7. Replace the first question mark with *30*, and replace the second question mark with *2/3*. Divide by 30. The **Numeric Expression** box should contain *Rv.Binom(30,2/3)/30*.
8. Click **OK**. This produces a variable *Sample* containing the proportion of correct guesses out of 30 for a subject that can guess the correct cola two-thirds of the time.



Use SPSS to create the histogram requested in part d.