Chapter 6

# Applying the Concepts: Excel Instructions

## Relate Variability to Uncertainty and Predictability

Currently not available in Excel.

## Estimating Likelihood of Values in Data

1. Open penguins.csv file in Excel.
2. Select the data.
3. Find the column header for body\_mass\_g in Column F.
4. Select the data in Column F (rows 1 through 334 in Column F only).
5. Insert a Histogram.
   1. With Column F highlighted, navigate to the Insert tab.
   2. Click on Insert Statistic Chart→Histogram.

Compute Probability Estimates:

1. Compute the proportion of penguins with body\_mass\_g between 3500 and 5500
   1. In cell J2, type the following formula:

=COUNTIFS(F2:F334, ">=3500", F2:F334, "<=5500")/COUNT(F2:F334)

1. Compute the proportion of penguins with body\_mass\_g of at least 6000
   1. In cell J3, type the following formula:

=COUNTIF(F2:F334, ">=6000")/COUNT(F2:F334)

## Predictability From Histograms

Create Histograms:

1. Open skeeball.csv file in Excel.
2. Insert → Chart Group → Statistic Chart → Histogram
3. Right-click the chart area.
   1. Click “Select Data”
   2. Legend Series (Series) → Edit
      1. Series values → Select the scores corresponding to Darc (B2:B1001)
      2. Click OK
   3. Click OK.
   4. Change the Chart Title to “Darc” by clicking on “Chart Title” and writing over it.
4. Repeat steps 2-3 for Qpawnz (B1002:B2001) and Ember (B2002:3001).

Calculate Average Scores:

1. Highlight both columns A and B (including column headers).
2. In the “Insert” tab.
   1. Click “PivotTable”
   2. Click “OK”
3. In the PivotTable sheet
   1. Drag person to the “Rows” box
   2. Drag scores to the “Values” box
   3. Click the dropdown next to “Sum of scores”
      1. Click “Value Field Settings…”
      2. Select “Average”
      3. Click “OK”

Compute Probability Estimates:

1. For Darc: In cell E2, type:

=COUNTIFS(A2:A3001, "=Darc", B2:B3001, ">400")/1000

1. For Qpawnz: In cell E3, type:

=COUNTIFS(A2:A3001, "=Qpawnz", B2:B3001, ">400")/1000

1. For Ember: In cell E4, type:

=COUNTIFS(A2:A3001, "=Ember", B2:B3001, ">400")/1000

## Predictability with Two Variables

1. Open the treats\_hour\_age.csv dataset in Excel.
2. Select the data.
   1. Click on the column headers for both day\_hour (column A) and treats (Column B) to select all the data in those columns.
3. Insert a scatterplot.
   1. With Columns A and B highlighted, go to the Insert tab on the Excel ribbon.
   2. Click on the Scatter (X,Y) icon in the Charts group.
   3. Choose the first scatter chart option, a simple scatter chart with no lines.
4. Repeat steps 2-3 for the age (column C) and treats (Column B) variables.

## Estimation from Different Sampling Methods

Calculate Statistics for Simple Random Sample:

1. Open the student\_gpa\_srs.csv file in Excel.
2. Highlight both columns A and B (including column headers)
3. In the “Insert” tab.
   1. Click “PivotTable”
   2. Click “OK”
4. In the PivotTable sheet
   1. Drag class to the “Rows” box
   2. Drag gpa to the “Values” box
      1. Click the dropdown next to “Sum of scores”
      2. Click “Value Field Settings…”
      3. Select “Average”
      4. Click “OK”
   3. Drag gpa to the “Values” box (for a second time)
      1. Click the dropdown next to “Sum of scores”
      2. Click “Value Field Settings…”
      3. Select “Count”
      4. Click “OK”

Calculate Statistics for Stratified Random Sample:

1. Open the student\_gpa\_strat.csv file in Excel.
2. Highlight both columns A and B (including column headers)
3. In the “Insert” tab.
   1. Click “PivotTable”
   2. Click “OK”
4. In the PivotTable sheet
   1. Drag class to the “Rows” box
   2. Drag gpa to the “Values” box
      1. Click the dropdown next to “Sum of scores”
      2. Click “Value Field Settings…”
      3. Select “Average”
      4. Click “OK”
   3. Drag gpa to the “Values” box (for a second time)
      1. Click the dropdown next to “Sum of scores”
      2. Click “Value Field Settings…”
      3. Select “Count”
      4. Click “OK”

Calculate Statistics for Balanced Stratified Random Sample:

1. Open the student\_gpa\_bal\_strat.csv file in Excel.
2. Highlight both columns A and B (including column headers)
3. In the “Insert” tab.
   1. Click “PivotTable”
   2. Click “OK”
4. In the PivotTable sheet
   1. Drag class to the “Rows” box
   2. Drag gpa to the “Values” box
      1. Click the dropdown next to “Sum of scores”
      2. Click “Value Field Settings…”
      3. Select “Average”
      4. Click “OK”
   3. Drag gpa to the “Values” box (for a second time)
      1. Click the dropdown next to “Sum of scores”
      2. Click “Value Field Settings…”
      3. Select “Count”
      4. Click “OK”

## Simulate Flipping a Loaded Coin

Currently not available in Excel.

## Simulate Samples

Currently not available in Excel.