

CH 4: PROBABILITY AND COUNTING RULES

- Empirical Probabilities or Relative Frequency Probabilities can be calculated either from summary data or from raw data.
 - Raw Data: We have already seen how to perform the calculation for raw data (section 2.1 above); Use the Tally function:
 1. Click on **Stat** → **Tables** → **Tally Individual Variables**
 2. Select the desired column and place it in the **Variables** window
 3. Check the **Counts** and **Percents** check boxes.
 - Summary Data: Assume that the values are in column C1 (named “X”) and that the frequencies are in column C2 (named “f”).

	C1	C2	C3	C4
	X	f		
1	5	26		
2	6	12		
3	7	34		
4	8	24		
5	9	18		
6	11	50		
7	11	48		
8	12	13		
9	13	30		
10	14	45		
11				
12				
13				

Figure 3.3

- Click on **Calc** → **Calculator**
- Type P(X) in the box for storing the result and
- Enter the following expression in the Expression box:
 - Find the **Round** function in the **function** list and click **Select**.
 - Double click C2 f on the left window containing the available columns
 - Click the Division operator “/”
 - Find the **Sum** function and click **Select**.
 - Double click C2 f
 - Enter 3 for the second argument of the Round function
 - The Expression box should contain: **Round(‘f’ / Sum(‘f’), 3)**
 - Click OK. The result should look like the one in Figure 3.4

	C1	C2	C3	C4	C5
	X	f	P(X)		
1	5	26	0.087		
2	6	12	0.040		
3	7	34	0.113		
4	8	24	0.080		
5	9	18	0.060		
6	11	50	0.167		
7	11	48	0.160		
8	12	13	0.043		
9	13	30	0.100		
10	14	45	0.150		
11					
12					
13					

Picture 3.4

- Law of Large Numbers - Simulation

We will simulate the probability of obtaining a 5 when we roll a “fair” die. We know that the theoretical probability is $1/6$ or $0.\overline{16}$.

We simulate the roll of die by generating a random integer between 1 and 6

- Click on **Calc** → **Random Data** → **Integer**
- The box **Generate _____ rows of data** will contain the number of trials for the experiment. In our case we will try the experiment 10, 50, 250, and 1000 times
- **Store in column(s)**: outcome of each iteration of the experiment. In our case the experiment consists of rolling once a single die. We will use columns C1, C2, C3, and C4 successively for each round of the simulation
- **Minimum Value**: In our case enter 1
- **Maximum Value**: In our case enter 6

We can use the following technique to calculate the empirical probabilities for each simulation:

- **Stat** → **Tables** → **Tally Individual Variables**
- In the Variables window enter the four columns containing the outcomes of the simulation: C1 C2 C3 C4.
- Check the **Counts** and **Percents** check boxes, then click OK
- The output will be placed in the Session window. For our simulation the results were 0.3, 0.12, 0.147, 0.173. Your results will be different than these as they will be different every time you run the simulation (this is the nature of random number generation).