

SPSS Program Notes

Biostatistics: A Guide to Design, Analysis, and Discovery Second Edition

by Ronald N. Fortofer, Eun Sul Lee, Mike Hernandez

Chapter 4: Probability and Life Tables

Program Note 4.1 – Simulation study to find probability of common birthdays

Below we display some simple commands to create ten columns of simulated data. The first line in SPSS syntax shown below is the **SET SEED** command. The **SET SEED** command when set equal to a specific value sets the seed for the pseudo-random number generator so that the same sequence of pseudo-random numbers can be reproduced each time the SPSS syntax is used. The command **UNIFORM(x)** is a random number generator that produces values from a uniform distribution between 0 and **x**. In our case below, we use **UNIFORM(1)** to acquire values from a uniform distribution between 0 and 1. We proceed by multiplying the random number by 364 and rounding the resulting product to the nearest whole number using the SPSS command **RND(x)**. Finally by adding one, we obtain values starting with 1 and ending with 365 representing each day of the year.

The variable **student** is used to identify the 30 students in each of the 10 classes. Each class is denoted by the variables **class1** to **class10** which store birthdays for each student. To do this begin with the **INPUT PROGRAM** command as shown in the SPSS syntax below. Notice that we used two **LOOP** statements. The first or outer **LOOP** identifies the each of the 30 students. Before getting to the second or inner **LOOP** which is used to generate random number from 1 to 365, we use the **VECTOR** statement to refer to more than one variable by creating a vector which allows us to store an array of variables. For example, we use **VECTOR CLASS(10)** where the vector name **CLASS(10)** refers to the variables **class1**, **class2**, **class3**, ..., **class10**. Notice that after the **COMPUTE** command, the statement **CLASS(I)** is used to assign a random number between 1 and 365 to the first student in each of the 10 classes. It then repeats this process 29 times until all 30 students in each of the 10 classes have been assigned a random number between 1 and 365. Note that each **LOOP** is closed with an **END LOOP** statement, each **INPUT PROGRAM** is closed with an **END INPUT PROGRAM** statement, and **END CASE** outputs the values obtained from the inner **LOOP** before addressing the next value in the outer **LOOP**.

SPSS syntax:

```
SET SEED = 123.  
INPUT PROGRAM.  
LOOP STUDENT = 1 TO 30.  
  VECTOR CLASS(10).  
  LOOP I = 1 TO 10.  
    COMPUTE CLASS(I) = RND(364*UNIFORM(1)) + 1.  
  END LOOP.  
END CASE.  
END LOOP.  
END FILE.  
END INPUT PROGRAM.  
  
LIST CASES.
```

In the SPSS output below, there are nine out of the classes that have duplicates.

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help



1 : student		1										
	student	class1	class2	class3	class4	class5	class6	class7	class8	class9	class10	i
1	1.00	274.00	118.00	66.00	331.00	131.00	81.00	287.00	146.00	46.00	69.00	11.00
2	2.00	284.00	160.00	353.00	97.00	261.00	203.00	194.00	315.00	53.00	314.00	11.00
3	3.00	239.00	281.00	257.00	121.00	191.00	341.00	18.00	225.00	240.00	137.00	11.00
4	4.00	215.00	115.00	197.00	104.00	193.00	204.00	37.00	41.00	213.00	247.00	11.00
5	5.00	177.00	54.00	274.00	351.00	277.00	351.00	35.00	248.00	326.00	104.00	11.00
6	6.00	242.00	292.00	290.00	117.00	285.00	199.00	258.00	304.00	36.00	27.00	11.00
7	7.00	93.00	86.00	290.00	250.00	95.00	313.00	173.00	251.00	257.00	200.00	11.00
8	8.00	28.00	229.00	273.00	301.00	35.00	185.00	129.00	74.00	40.00	82.00	11.00
9	9.00	365.00	4.00	88.00	9.00	308.00	248.00	6.00	126.00	310.00	44.00	11.00
10	10.00	275.00	351.00	348.00	82.00	106.00	127.00	303.00	245.00	114.00	224.00	11.00
11	11.00	319.00	272.00	314.00	317.00	234.00	306.00	341.00	77.00	306.00	127.00	11.00
12	12.00	4.00	189.00	270.00	152.00	195.00	230.00	96.00	31.00	94.00	48.00	11.00
13	13.00	189.00	351.00	229.00	353.00	231.00	150.00	291.00	208.00	61.00	86.00	11.00
14	14.00	148.00	346.00	206.00	180.00	245.00	69.00	143.00	168.00	99.00	237.00	11.00
15	15.00	284.00	232.00	117.00	291.00	119.00	224.00	122.00	158.00	187.00	33.00	11.00
16	16.00	168.00	11.00	262.00	310.00	221.00	273.00	133.00	164.00	240.00	177.00	11.00
17	17.00	147.00	321.00	74.00	304.00	65.00	161.00	151.00	28.00	351.00	167.00	11.00
18	18.00	188.00	220.00	334.00	287.00	36.00	168.00	204.00	39.00	237.00	358.00	11.00
19	19.00	179.00	192.00	127.00	322.00	211.00	159.00	363.00	106.00	81.00	293.00	11.00
20	20.00	4.00	108.00	152.00	215.00	266.00	277.00	287.00	214.00	219.00	252.00	11.00
21	21.00	328.00	15.00	263.00	62.00	364.00	31.00	216.00	320.00	163.00	183.00	11.00
22	22.00	352.00	78.00	279.00	156.00	299.00	52.00	175.00	108.00	212.00	77.00	11.00
23	23.00	340.00	237.00	72.00	287.00	324.00	244.00	210.00	316.00	91.00	176.00	11.00
24	24.00	261.00	174.00	65.00	324.00	3.00	282.00	79.00	204.00	213.00	342.00	11.00
25	25.00	153.00	71.00	165.00	37.00	112.00	321.00	307.00	199.00	285.00	281.00	11.00
26	26.00	272.00	144.00	158.00	288.00	105.00	296.00	320.00	265.00	92.00	288.00	11.00
27	27.00	120.00	149.00	356.00	251.00	147.00	260.00	282.00	98.00	297.00	140.00	11.00
28	28.00	164.00	302.00	161.00	252.00	261.00	154.00	9.00	316.00	281.00	228.00	11.00
29	29.00	49.00	189.00	352.00	253.00	160.00	363.00	318.00	53.00	279.00	252.00	11.00
30	30.00	199.00	165.00	142.00	8.00	108.00	208.00	312.00	163.00	222.00	364.00	11.00