STAT 13 Homework 2 Solutions

www.stat.ucla.edu/~dinov/courses_students.dir/08/Fall/STAT13.1.dir/assignments.html

1. The smoking status of the husband is not independent of that of the wife because P(both the husband and wife smoke) ≠ P(husband smokes) * P(wife smokes) as you can see that 0.08 ≠ 0.3 * 0.2. This comes from the fact that P(A|B) = $\frac{P(A \cap B)}{P(B)}$ by definition, and in order for the condition of independence P(A|B)=P(A) to be true, P(A∩B) has to equal P(A)*P(B).

2.

(a) P(all 20 will be cured) = $_{20}C_{20}*(0.90)^{20}*(1-0.90)^{0}$ = 0.122.

(b) P(all but 1 will be cured) = (19 out of 20 will be cured) = $_{20}C_{19}*(0.90)^{19}*(1-0.90)^{20-19}$= 0.270.

(c) & (d) P(exactly 18 will be cured) = P(exactly 20% will be cured) = $_{20}C_{18}*(0.90)^{18}*(1-0.90)^{20-18}$ = 0.285.

3.

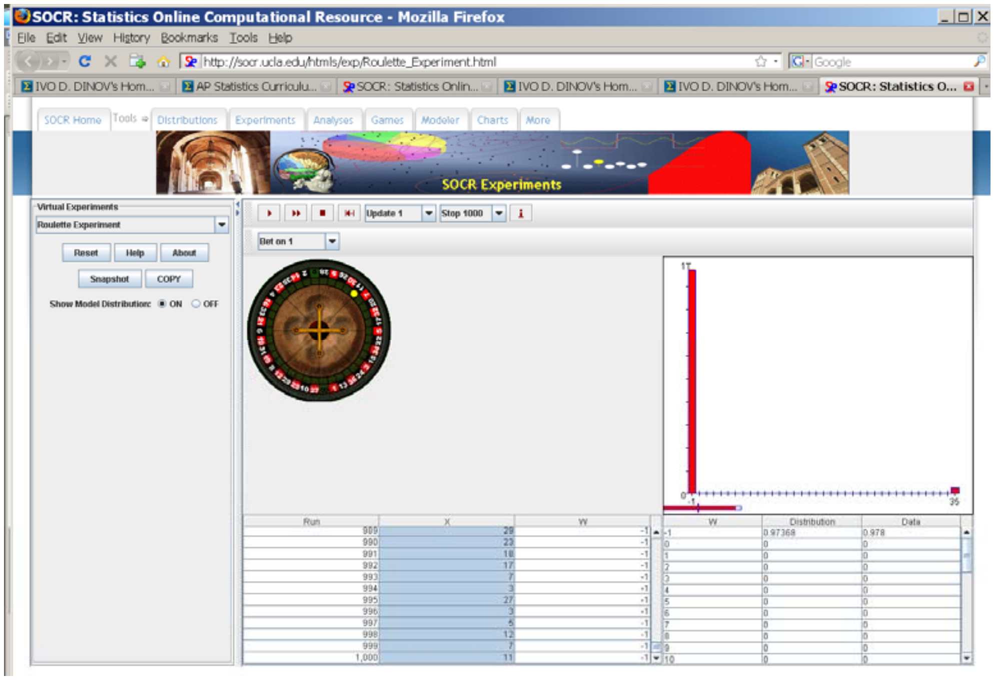
(a) P(none have high blood lead) = $_{16}C_0*(\frac{1}{8})^0*(\frac{7}{8})^{16-0}$ = 0.118

(b) P(one has high blood lead) = $_{16}C_1*(\frac{1}{8})^1*(\frac{7}{8})^{16-1}$ = 0.270

(c) P(two have high blood lead) = $_{16}C_2*(\frac{1}{8})^2*(\frac{7}{8})^{16-2}$ = 0.289

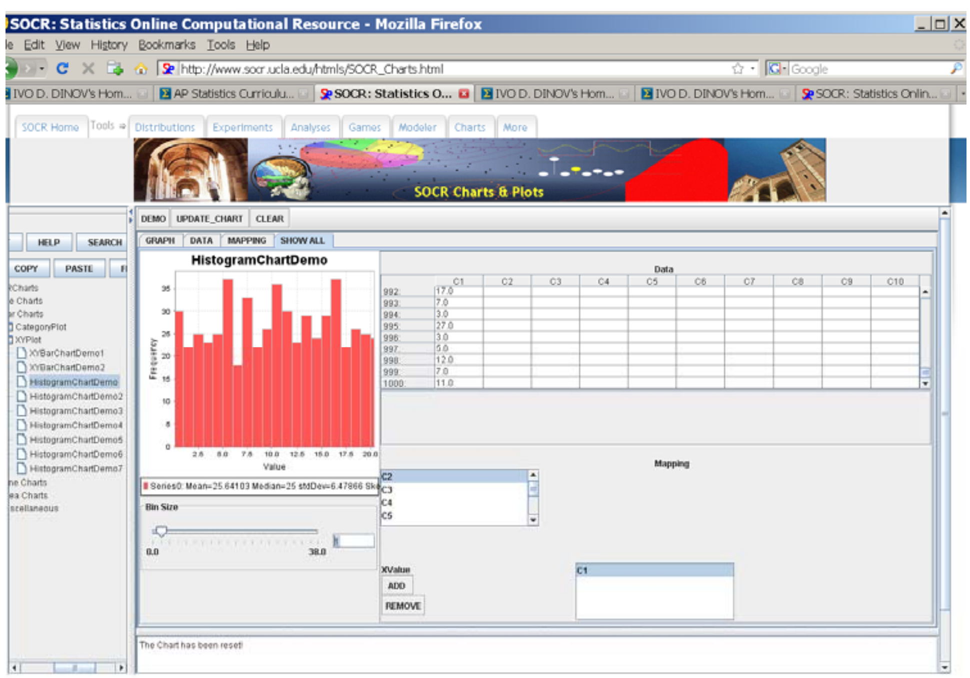
(d) P(three or more have high blood lead) = 1 – P(none or one or two have high blood lead) = 1 – 0.118 – 0.270 – 0.289 = 0.323

4.
After running a simulation, count the number of observations that have a value less than 20 and divide that by the total number of observations. Answers may vary.



The figure above illustrates the results of 1,000 experiments (you need to do this with 10 and 100 trials first!) The data of interest is in the second column (containing the outcomes of all 1,000 trials). The question is how many of these 1,000 outcomes are < 20 (i.e., 0, 00, 1, 2, …, 19).

You can copy and paste this column into, say, SOCR Histogram Chart and obtain the frequencies of all 38 possible outcomes (see chart below). Finally adding all the frequencies for these outcomes (0, 00, 1, 2, …, 19), and dividing the sum by 1,000 will give you the estimate of the desired probability of P(Outcome <20). In this case, P(Outcome <20) ≈ 510/1,000 = 0.51!



5.

(a) Since there are 2 ways of satisfying the condition, $2*(0.3)^{10}*(0.7) = 0.000008$

(b) There are 3 scenarios you have to consider (#A = #T = 0 for first term, 1 for second term, or 2 for last term): $(1-0.3-0.3)^5 + \frac{5!}{1!*1!*3!}(0.3)(0.3)(1-0.3-0.3)^3$

$+\frac{5!}{2!*2!*1!}(0.3)^2(0.3)^2(1-0.3-0.3) = 0.223$

(c) We first need to set up the right spinner parameters (number of segments, 4, and their probabilities, 0.3, 0.3, 0.2, 0.2, respectively). Then run the experiment 100 times (of course you can do 1,000 or more times, but this specific problem just asks for 100 trials. More trials will produce more reliable estimates, indeed). Then you need to count the number of outcomes that satisfy the conditions of parts (a) and (b) of this problem. Remember these could be numbers between 0, 1, 2, …, 100. Then the proportions of these numbers over 100 will give you the empirical estimates of the probabilities of interest. Does that make sense? So, if one of the events in *not* observed in the 100 trials then the empirical probability = 0, if it's observed only once, then the probability is 1/100 = 0.01. And so forth.

For example, for part (b), you can use this ***Excel*** function to compute the number of trials where the Number of A's (1's) is equal to the number of T's (2's). In the applet, the 3rd column (N) contains the results of the 100 trials. Copy and paste these data in an Excel Column (say column "D") and use this function in cell E5 and then drag this sell down to E100. In the

E5 =ABS(COUNTIF(D1:D5, "1") -COUNTIF(D1:D5, "2"))

…

E100 =ABS(COUNTIF(D96:D100, "1") -COUNTIF(D96:D100, "2"))

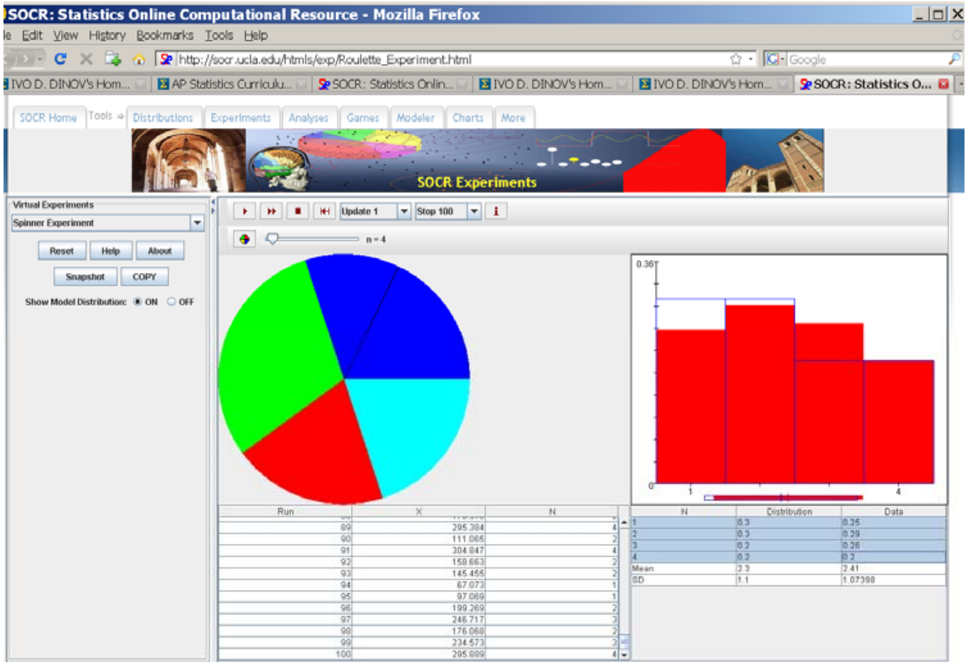
Finally in cell E101, we can count the number of cells in E5 – E100 that are equal to zero (i.e., the number of 1's and 2's, in a list of 5 consecutive trials, are equal):

E101=COUNTIF(E5:E100,"0")

In this example the answer is 33, out of the 100 experiments had the same number of 1's and 2's (i.e., same number of A's and T's). Thus the empirical estimate of this probability is **0.33**. Contrast this with the theoretical probability which we computed above (0.223).

Note on empirically estimating the probability in part (a) using only 100 trials: You most likely will **not** observe the event {A random drawing of 10 A's in a row in a sample of 11 randomly chosen nucleotides} in 100 trials, as its theoretical probability is small (0.000008), so your estimate will be 1/100 = 0.0.

Answers may vary depending on your simulation.


SOCR: Statistics Online Computational Resource - Mozilla Firefox
Edit
View
History
Bookmarks
Tools
Help
http://socr.ucla.edu/htmls/exp/Roulette_Experiment.html
SOCR Home
Tools
Distributions
Experiments
Analyses
Games
Modeler
Charts
More
SOCR Experiments
Virtual Experiments
Spinner Experiment
Reset
Help
About
Snapshot
COPY
Show Model Distribution:
ON
OFF
Update 1
Stop 100
Run
X
N
Distribution
Data
Mean
SD