

Additional problem for chapter 1:

3. Random Publishers is deciding whether to publish a new marketing-research textbook, which they hope might be able to take a substantial market share from Churchill's famous marketing-research book.

Let Q denote the unknown fraction of all marketing-research teachers who would prefer Random's new book over Churchill's book.

If the business editor at Random knew Q , then she would say that every teacher has a conditionally independent Q probability of preferring Random's new book over Churchill.

Unfortunately, the editor does not know Q . In fact, the Random editor's uncertainty about Q can be represented by uniform distribution over the interval from 0 to 1. That is, up to any given number of digits, the editor thinks that Q is equally likely to be any number between 0 and 1.

To get more information about Q , the editor has paid 8 randomly sampled marketing-research teachers to carefully read the new book and compare it to Churchill. Three months from now, each of these teachers will send the editor a report indicating whether he or she would prefer to this new book over Churchill's.

- (a) Make a spreadsheet model to simulate the editor's uncertainty about the unknown fraction Q and the results that she will get from her sample of eight readers. Make a table of data from at least 500 simulations of your model to answer the following questions.
- (b) Estimate the conditional probability that $Q > .5$ (that is, a majority in the overall population will prefer the new book) given that 6 out the eight sampled prefer the new book.
- (c) Estimate the conditionally median value of Q given that 6 out the eight sampled prefer the new book.
- (d) Make a chart showing the (inverse) cumulative distribution given that 6 out of the eight sampled prefer the new book.
- (e) Make a table and a chart indicating, for each number K from 0 to 8,
 - (i) the value that would have conditional cumulative probability 0.10 for Q given that K out of the eight sampled teachers prefer the new book (that is, the number q such that $P(Q < q | K \text{ prefer}) = 0.10$)
 - (ii) the conditional median value for Q given that K out of the eight sampled teachers prefer the new book (that is, the number q such that $P(Q < q | K \text{ prefer}) = 0.50$)
 - (iii) the value that would have conditional cumulative probability 0.90 for Q given that K out of the eight sampled teachers prefer the new book (that is, the number q such that $P(Q < q | K \text{ prefer}) = 0.90$)

Additional problem for Chapter 2.

Recall these random variables from problem 3:

(a) $=IF(RAND()>0.3,2,0)+IF(RAND()>0.6,3,0)$

(b) $=IF(AND(0.3<RAND(),RAND()<0.4),1,0)$

(c) $=IF(RAND()<0.6,IF(RAND()<0.5,1,2),3)$

4. Make a simulation table containing data from 100 simulations of each of the random variables that were described in parts (a), (b), and (c) of problem 3. For each random variable, answer the questions (i)-(iv) based on the results in this simulation table.

(i) Estimate of the expected value of the random variable from these 100 simulations.

(ii) Estimate the standard deviation of the random variable from these 100 simulations.

(iii) Give a 95% confidence interval for the expected value of the random variable, based on this simulation data.

(iv) If the radius of the 95% confidence interval in part (iii) is more than 10% of the estimated expected value in part (i), then estimate how many more simulations you should need to get a 95% confidence interval that has a radius less than 10% of the estimated expected value. (That is, estimate how large the sample size should be so that the confidence interval has the form $m \pm r$ where $r < 0.1 * m$.)

(v) Generate the additional number of simulations that you estimated in part (iv) (if any more were necessary to get a the radius of the 95% confidence interval to be less than 10% of the estimated expected value), and put a copy of this data directly below the data from your original 100 simulations. Then compute a new 95% confidence interval for the expected value based on all your simulation data for this random variable.

(vi) Now go back to the probability distributions and compute the true expected value of this random variable. Is it in your 95% confidence intervals?