Teaching insurance concepts and developing problem solving skills through statistical simulation

Ed Pappanastos
Troy University

Courtney Baggett
Butler University

Edwin H. Duett
Troy University

ABSTRACT

Most colleges and schools of business list some version of problem solving, analytical reasoning, or critical thinking skills as assessable program outcomes. This is reasonable given that employers of business students often cite these as essential skills. This article examines how Excel can be used to build simulation models to illustrate an important insurance concept – the law of large numbers. Teaching in this fashion results in a better understanding of the concept and the development of better problem solving/modeling skills.

Key Words: Insurance Education, Law of Large Numbers, Simulation, Problem Solving
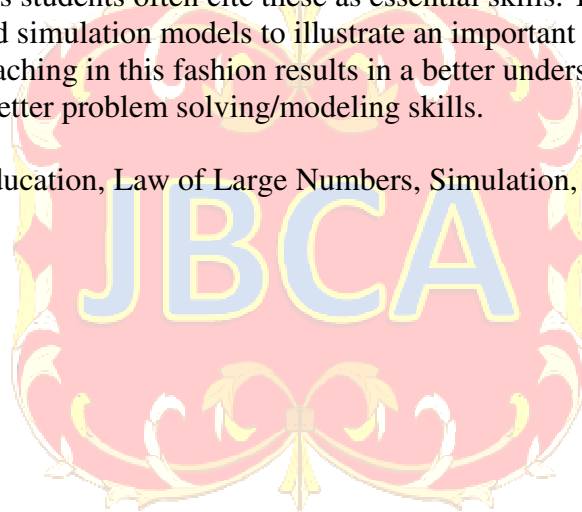
INTRODUCTION

The primary function of colleges and universities is to prepare their students for successful careers. In order to accomplish this, colleges and universities must focus on skill sets that are demanded by employers. While the particular skill sets may vary among employers, there is a consistent desire of employers of business graduates, insurance graduates in particular, to possess problem solving and critical thinking skills. A clear example of this is GEICO's recent search for graduates to join their Auto Damage Management Development Program (GEICO, 2015). The qualifications for entry into the program explicitly state that applicants must "… be well-developed in analytical, problem-solving and decision-making skills …". In addition, one on-line employment site (Target Jobs, 2016) lists problem solving skills as one of the six most important skills for someone working in financial services and insurance and while another site (Inside Careers, 2016) lists it as being one of the top five skills.

While colleges and universities strive to instill these skills in graduates, studies continue to show that employers are dissatisfied with the problem-solving and critical thinking skills of college graduates. It was reported (Belkin, 2015) that the American Association of Colleges and Universities found that 90% of business owners surveyed evaluated recent college graduates as being "poorly prepared for the workforce in such areas as critical thinking, communication and problem solving."

As colleges and universities continue to bridge the gap between desired skills of employers and the actual skills of graduates, it is important that new approaches to teaching be used. Instead of the traditional lecture and test approach, students need to be given activities that develop problem solving and analytical skills. The remainder of this article demonstrates how an important insurance concept can be taught while allowing students to hone their problem solving and analytical skill sets.

INSURANCE CONCEPT

Insurance is a risk transfer mechanism whose financial viability is contingent on the ability of the insurance company to accurately identify a large, homogenous class of independent risk exposures and accurately determine the insurance premium for that class (Rejda, 2017).

Identifying a large, homogenous class of independent risk exposures serves three purposes. First, the result of having a homogenous class is that the underlying loss distributions for each exposure in the class are identical. Second, the fact that the risk exposures are independent means that losses are uncorrelated. From an insurance perspective, this is important as this ensures that probability of a loss for one exposure does not change given a loss by another exposure. Third, the result of having a large class is that the law of large numbers applies, which in turn allows the insurance company to minimize its objective risk.

The law of large numbers is an often misinterpreted principle. It is often equated to the concept of a particular event "being due to happen." For example, suppose someone is flipping a coin that is assumed to be a fair coin (50% change of heads and 50% chance of tails). Upon observing several tails in a row, it is often claimed that the law of large numbers suggests that a "head is due" and that the next flip "has to result in a head." This incorrect reasoning lies in the failure to recognize that coin flips are independent and that, if the coin is in fact fair, the probability of getting a head remains at 50%.

Instead, what the law of large numbers does say is, that as the number of flips increases, the observed results more closely match what is expected. So, as the number of flips increases, the percentage of heads will more closely approximate 50%. From an insurance perspective, the law of large numbers means that as the number of exposures becomes large, the insurance company can more accurately forecast total losses for the class because the variance in the loss distribution decreases. As a result, the insurance company can more accurately determine the proper insurance premium for the class.

The remainder of this article focuses on building Excel models to simulate the law of large numbers. A variety of situations are considered and the required model changes are presented for each variation.

SIMULATION MODELS

Coin Example

The simplest example for students to model that will reinforce the concept of the law of large number is to model the results of successively flipping a coin (Doane, 2016). In order to do this in Excel, only two functions (RAND and IF) and three different Excel formulas are necessary.

The function RAND() is a predefined function that generates a random, uniform variable between 0 and 1. To simulate the flipping of a fair coin, half of the random numbers need to be assigned to HEADS and the other half need to be assigned to tails. This can be accomplished by using the IF function. The IF function has three arguments: IF(argument 1, argument 2, argument 3). The first argument is a condition that is evaluated as TRUE or FALSE. If argument 1 is evaluated as TRUE, then argument 2 is the value of the cell with the IF function; otherwise, argument 3 is the value of the cell. Arguments 2 and 3 can be a calculation, a number, or a label. To simulate the random flipping of a coin, the RAND function can be used within the IF function as follows:

=IF(RAND()<0.5,"HEADS","TAILS")

Upon entering this formula, Excel (1) randomly generates a number between 0 and 1 and (2) evaluates to see if this randomly generated number is less than 0.5. If the value is less than 0.5, then HEADS is placed in the cell with the IF function; otherwise, TAILS is placed in the cell. This statement just needs to be replicated for the number coin flips that need to be simulated. Figure 1 (Appendix) illustrates a portion of a sequence of 250 flips (note the formula in the formula bar).

The next step is to create a running total of the number of HEADS (assuming we are wanting to simulate the percentage of heads over a number of flips). This can be accomplished by using two IF functions. First, an initial value must be determined. This is accomplished by using the following function: =IF(C4="HEADS",1,0). This function places a 1 in cell D4 if the first flip is HEADS; otherwise, a 0 is placed in D4. This is illustrated in Figure 2 (Appendix).

Next, another IF function must be used to (1) evaluate the next flip and (2) update the running total for the number of HEADS. This is accomplished by using the following function: =IF(C5="HEADS",D4+1,D4)

So, if the next flip is a HEADS, the previous total is increased by one. If the next flip is a TAILS, then new total is simply the previous total. Figure 3 (Appendix) illustrates the use of this formula. Note that the total does not change for the second and third flips because both are TAILS. This formula needs to be copied down for all 250 flips.

The next step is to calculate the percentage of heads after each individual flip. This is accomplished by simply dividing the current Total Heads by the Flip number as shown in Figure 4 (Appendix).

The final step is to chart the Percent HEADS over time. As shown in Figure 5 (Appendix), this plot and clearly demonstrates the law of large numbers. Given a fair coin, it is expected that 50% of the flips will be HEADS. When there are only a few flips, there is significant variability in what is observed and what is expected. However, as the number of flips increases, the Percent HEADS much more closely approximates what is expected.

An infinite number of simulations can be generated by repeated hitting F9 (the recalculation key). A chart illustrating an additional simulation is provided in Figure 6 (Appendix). Note, once again, there is a tremendous amount of variation when the number flips is small, but as the number of flips increases, the Percent HEADS approaches 50%. The simulations include only 250 flips. This would not be considered a large class by any means. However, it is interesting to not how quickly the percentage starts to level off around 50%.

Life Insurance Example

Only a slight modification needs to be made to use an insurance application to illustrate the law of large numbers. For example, consider a one-year term life insurance policy for a 70-year old male. According to the Social Security Administration (Social Security Administration, 2013), the probability that an average 70-year old male will die before turning 71 is 0.023528. So, the previous simulation model only requires one change – replace 0.50 with 0.023528 – and, of course the verbiage needs to change. The resulting simulation model is shown in Figure 7 (Appendix).

Because the probability of a death during the one-year term is so low, it is necessary to increase the number of exposures/policies to get a true representation of the law of large numbers. In this example 1,000 one-year term policies have been simulated. The percent of policies where a death occurred is plotted show in Figure 8 (Appendix).

Even though the probability of each individual policy holder dying during the year remains at 0.023528, the actual percent dying varies tremendously when there are a low number of risk exposures. As with flipping the coin, the simulation illustrates that as the number of risk exposures increases, the observed results very closely approximate the expected results.

Automobile Insurance Example

Unlike life insurance where there are only two outcomes, no loss and complete loss, automobile insurance involves risk exposures that have a continuum of possible losses. From a liability standpoint, there is no upper limit on the amount of a loss. From a property loss standpoint, the range of outcomes is from no loss to the value the automobile. Consider the physical loss distribution shown in Table 1 (Appendix) for a class of automobile exposures (modeled as a discrete distribution for simplicity) where the insured value of the car is $15,000.

The expected loss of this distribution (which is a long-run value – i.e., a product of the law of large numbers) is found by summing the products of the value of the random variable and its respective probability. In this case, the expected value is:

$$E(X) = \sum x_i P(x_i).$$

The expected loss for this distribution is $625, as indicated in Figure 9 (Appendix).

Simulating this random variable requires a little more work than simulating the binary events of flipping a coin (heads/tails) or a life insurance policy (live/die). The Excel function =RAND() is still needed to generate random values between 0 and 1. However, the resulting random values need to be converted to losses of $0, $500, $1,000, $5,000, and $15,000 in a manner that is consistent with the probability distribution. That is, 70% of the random values must result in a $0 loss, 15% must result in a $500 loss, ... This can be accomplished by using the LOOKUP function. In generic terms, the LOOKUP function has two arguments and takes the following form: =LOOKUP(what,where). In this case, a random value is "what" is going to be looked up and a table needs to be created to identify "where" it is to be looked up.

Given this probability distribution the lookup table would take the form indicated in Figure 10 (Appendix). Column B represents values to which the random value is to be compared. This column is created by entering a 0 in the first cell and then by adding each probability from the loss distribution. For example, the probability of $0 loss is 0.70. The second value in Column B is found by (0 + 0.70). The next probability in the loss distribution is 0.15. So, the third value in Column B is found by (0.70 + 0.15). This process is continued for all but the last probability in the loss distribution.

Using the LOOKUP function, a random variable is compared to the entries in the first column, one at a time, in order to find a 'match.' A 'match' is found by identifying the largest value in the first column that is greater than or equal to the random variable. For example, suppose that the result of =RAND() is 0.83. By definition, all random values will be at least 0. As a result, 0.83 is compared to the second entry in the first column, 0.70. Since 0.83 is greater than 0.70, 0.83 is compared to the next entry in the first column. Because the 0.83 is NOT greater than or equal to 0.85, 0.83 is considered a 'match' with 0.70. Now the LOOKUP function can be used to identify the loss as a $500 loss. Under this "mapping" arrangement, every value that is at least 0.70 and less than 0.85 will be identified as a $500 loss. The range 0.70 to 0.85 represents 15% of the values generated by =RAND(). Therefore, 15% of the losses will be identified as $500 losses.

The complete "mapping" arrangement is illustrated in Figure 11 (Appendix). Note that the areas are not proportionally accurate, but the range of random values are mapped to the losses in a manner that is consistent with the stated probability distribution.

A sample of the Excel simulation model to simulate the automobile example is provided in Figure 12 (Appendix). In Column C are the individual outcomes for each auto policy. Note that the reference to the lookup table does not include the headings and that $ are used to make use of absolute cell references. Using absolute cell reference allows for the formula to be copied without the cell references changing.

Instead of calculating the percentage of losses (deaths in the life insurance example), it is more practical to calculate the average loss. As indicated in Figure 13 (Appendix) plots the average loss as the number of policies increases. Once again, the law of large numbers is demonstrated. As the number of policies increases, the average loss more closely approximates

the expected loss. Recall from Figure 8, the theoretical average loss is $625 – the simulated average loss has quickly approached this value.

CONCLUSION

This article demonstrates how statistical simulation is a valuable tool for teaching the law of large numbers, an extremely important insurance concept. In addition to teaching this important concept, statistical simulation helps develop problem solving/analytical skills. Other insurance concepts such as adverse selection, correlated risk exposures, and property coinsurance penalties can be taught using simulation.

REFRENCES

Belkin, D. (2015, Jan 17). U.S. news: Skills gap found in college students. Wall Street Journal Retrieved from http://search.proquest.com/docview/1646105560?accountid=38769

Doane, David P., and Lori Seward. *Applied Statistics in Business and Economics*. New York: McGraw-Hill, 2016. Print.

GEICO invites college graduates to join management development program. (2015, Oct 17). Business Wire Retrieved from http://search.proquest.com/docview/1722627252?accountid=38769

Inside Careers. *Skills & Training: Five Key Insurance Skills*. Retrieved from http://www.insidecareers.co.uk/career-advice/insurance-skills/

Rejda, George E., and Michael J. McNamara. *Principles of Risk Management and Insurance*. Boston: Pearson, 2017. Print.

Social Security. *Actuarial Life Table, 2013*. Retrieved from https://www.ssa.gov/OACT/STATS/table4c6.html

Target Jobs. *The top skills retail banking, insurance and actuarial employers* want. Retrieved from https://targetjobs.co.uk/career-sectors/financial-services-and-insurance/advice/284627-the-top-skills-retail-banking-insurance-and-actuarial-employers-want

APPENDIX

Figure 1. Simulating One Flip.

C4	⋮	✕	✓	*fx*	=IF(RAND()<0.5,"HEADS","TAILS")

◢	A	B	C	D	E	F	G
1							
2							
3		Flip	Outcome				
4		1	HEADS				
5		2	TAILS				
6		3	TAILS				
7		4	HEADS				
8		5	HEADS				

Figure 2. Initial Count

D4	⋮	✕	✓	*fx*	=IF(C4="HEADS",1,0)

◢	A	B	C	D	E	F	G
1							
2							
3		Flip	Outcome	Total HEADS			
4		1	HEADS	1			
5		2	TAILS				
6		3	TAILS				
7		4	HEADS				
8		5	HEADS				

Figure 3. Calculating the Running Total

D5	⋮	✕	✓	*fx*	=IF(C5="HEADS",D4+1,D4)

◢	A	B	C	D	E	F	G
1							
2							
3		Flip	Outcome	Total HEADS			
4		1	HEADS	1			
5		2	TAILS	1			
6		3	TAILS	1			
7		4	HEADS	2			
8		5	HEADS	3			

Figure 4. Calculating the percent HEADS.

		Total	Percent
Flip	Outcome	HEADS	HEADS
1	HEADS	1	100%
2	TAILS	1	50%
3	TAILS	1	33%
4	HEADS	2	50%
5	HEADS	3	60%

E4 =D4/B4

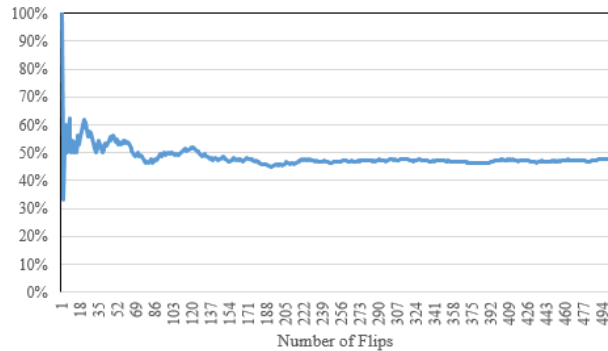Figure 5. Plot of Percent HEADS over time.

Number of Flips

Figure 6. Additional Plot of Percent HEADS over time.

Number of Flips
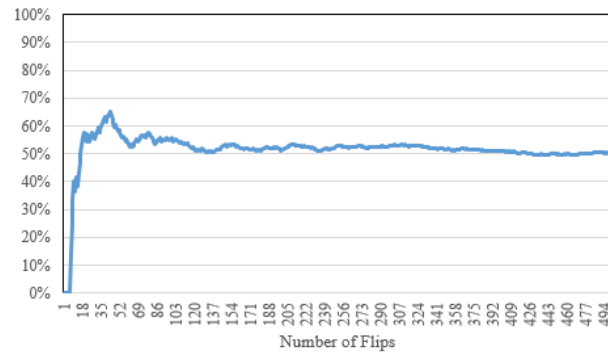
Figure 7. One-year Term Life Insurance Simulation Model

D4 =IF(C4="Died",1,0)

		Total	Percent
Policy	Outcome	Deaths	of Deaths
1	Lived	0	0.00%
2	Lived	0	0.00%
3	Lived	0	0.00%
4	Lived	0	0.00%
5	Lived	0	0.00%

Figure 8. Percent of Deaths as Number of Policies Increases
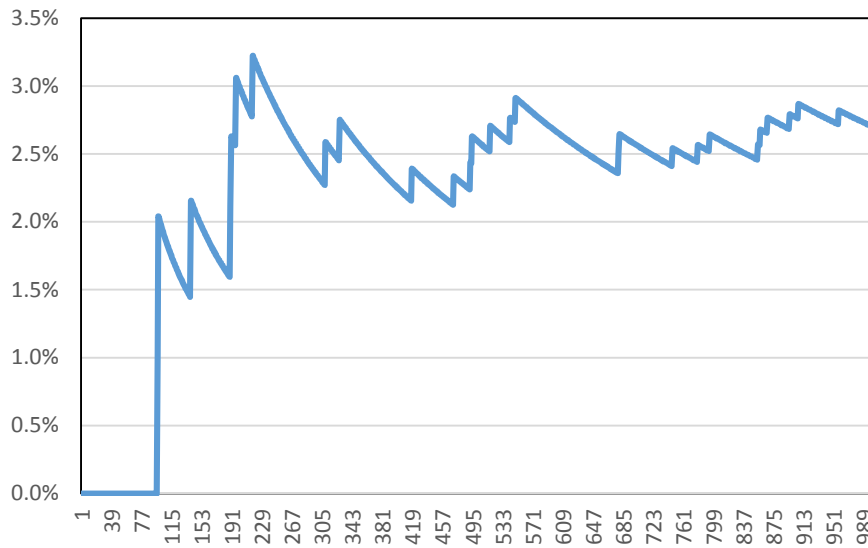
Table 1. Automobile Physical Loss Distribution

Loss	Probability
$0	0.70
$500	0.15
$1,000	0.10
$5,000	0.03
$15,000	0.02

Figure 9. Expected Value of Loss Distribution.

Figure 10. Loss Lookup Table

	A	B	C	D
1				
2		RAND()	Loss	
3		0	$0	
4		0.70	$500	
5		0.85	$1,000	
6		0.95	$5,000	
7		0.98	$15,000	
8				

Figure 11. Random Value Mapping to Loss Distribution.
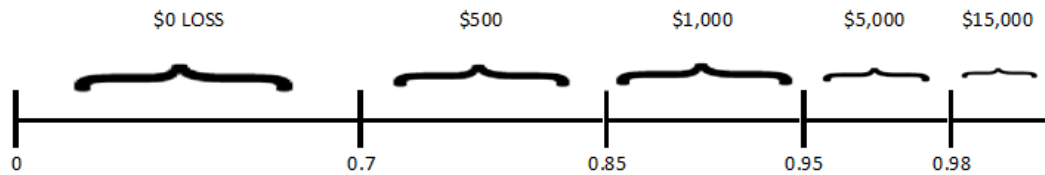
Figure 12. Sample of Excel Automobile Policy Simulation.

C3   fx   =LOOKUP(RAND(),$G$4:$H$8)

	A	B	C	D	E	F	G	H
1								
2		Auto Policy	Loss	Cumulative Loss	Average Loss			
3		1	$0	$0	$0		RAND()	Loss
4		2	$0	$0	$0		0	$0
5		3	$1,000	$1,000	$333		0.7	$500
6		4	$0	$1,000	$250		0.85	$1,000
7		5	$0	$1,000	$200		0.95	$5,000
8		6	$0	$1,000	$167		0.98	$15,000
9		7	$1,000	$2,000	$286			

Figure 13. Average Automobile Loss as the Number of Policies Increases.