MMA 863 – Assignment 1

This assignment includes material from Sessions 1 to 3 inclusive. You may also benefit from session 6, which can be done at any time.

The results should be provided in a single ‘main’ document (probably a PDF) that is print-ready with a single supporting Excel file if required. You should assume that any supporting documentation may not be reviewed, so your entire answer should be in the main document. The main document may contain scanned versions of hand-drawn diagrams / formulas if you like.

1. Kingston is becoming increasingly known for the potholes in its streets. It seems to me that driving along you encounter on average about one every 50 m of driving, though they are randomly distributed. Most of the potholes are small, but about 10% of them are large enough to be a bit dangerous. Those dangerous ones will flatten a tire about 0.1% of the time (i.e. if hit ‘just right’, which itself is, of course, random.)
   1. I live 1 km from the Kingston campus. If I drive to school and back home, what is the probability that I will encounter more than 3 dangerously large potholes? Be sure to document your relevant assumptions.

Poisson: (1) (potholes) per (50 m)

10% of potholes are dangerous **dp = dangerous potholes**

Poisson: (.1) (dp) per (50 m)

Poisson: (4) (dp) per (2k)

=1-POISSON.DIST(3,4,TRUE) = .567

Therefore, there is approximately 56.7% chance that you will encounter more than 3 dangerously large potholes

Chance of flat tire if hit

* 1. Suppose I drive to school and back home 200 times in a year. Develop a model to estimate the probability that I will flatten a tire due to a pothole. Be sure to list any necessary and reasonable assumptions, if any.

Success = not getting a flat

Number of success = Number of trials (to find probability of zero flats)

1 Trial = 1 dp

Average dp per trip =4; 200 trips = 800 dps

Number of trials = 800

% success = (1- chance of getting a flat for each time encountering a dp) = 1 -.1% = .999

So, we are testing the chance of getting more than zero flats when encountering 800 dp.

Prob more than zero flats = 1 – P(zero flats)

=1-BINOM.DIST(800,800,0.999,FALSE) = 0.550850851 = 55.1%

1. Where I live it rains about 20% of days, but it seldom rains two days in a row – if it rains on one day, there is only a 5% chance it will rain on the next. It is Friday morning and it looks like it will rain today, in fact, I’d guess there is an 80% chance that it will rain. The rain should not really bother me as I am going to fly to Zurich for two days, so I will not actually see it rain.
   1. What is the probability it will rain tomorrow?

Tree diagram

Probability that it rains tomorrow =

Probability that it rains tomorrow given that it rains today

+ Probability that it rains tomorrow given that it doesn’t rain today

= (.2)(.05) + (.8)(.2) = .01 +.16 = **.17**

* 1. What is the probability it will rain both today and tomorrow?

= Probability that it will rain today X Probability that it will rain tomorrow given that it rained today

= .8 X .05 = **.04**

Note: this makes sense that it is lower than a. If it rains today, it decreases the chance of raining tomorrow.

* 1. Suppose I come back and my neighbor tells me it rained on Saturday, what is the probability it rained on Friday?

**Note: we know the answer will be much lower than 50%, since the chance of it raining two days in a row is low**

**Probability Rain Saturday | Rain Friday**

Tree diagram

For 100 Fridays:

80 Fridays it did not rain. For the following 80 Saturdays, it rained 16 times (20%)

20 Fridays it did rain. For the following 20 Saturdays, it rained 1 day (5%)

It rained 17 times. 16/17 times it did not rain on the preceding Friday.

The probability that it rained on Friday 1/17 = 0.0588 = **5.9%**

1. Winter Wonderland: The driveway at my cottage is literally 1 km long. It is quite a peaceful drive in the summer but in the winter it can be a treacherous frozen nightmare. Last week, after the big ice storm, I drove up to the lake and realized I couldn’t drive down the driveway because it was too icy for the car. It was even dangerous for walking and I anticipated a 5% chance I’d slip and fall during the first 100 m alone. Nevertheless, it was important to go in because I needed to know how much propane we had left for heating the place. (You may be interested in knowing that the daily consumption of propane is distributed with a mean of 6.5 L with a standard deviation of 2 L.) Clearly I was going to have to walk all the way in and then all the way back out later that afternoon – I was not happy!
   1. SOLVE THIS TWO DISTINCT WAYS AND COMPARE ANSWERS: So I started walking in, only to realize, half way to the cottage, that I’d forgotten my keys and had to go back to the car to get them. Given that I ‘got lucky’ and had not fallen by the time I realized I had forgotten my keys, what is the probability I would finish the day without slipping?

Solution 1:

RY BINOM.DIST(0,25,0.05,TRUE) =0.277389573 or

MH BINOM.DIST(25,25,0.95,FALSE) = 0.27739 = 27.7%

Using the binomial distribution, the probability of finishing the day without slipping is approximately 27.7%

Logic:

One trial = 100m of walking

25 trials to complete the days walking. (500 back to car, 1 km to cottage, 1 km back out)

5% chance of falling = 95% chance of not falling

Number of successes = 25

Number of trials = 25

Probability of success = .95

Probability mass function (FALSE) used to solve for exactly 25 successes.

Solution 2:

RY POISSON.DIST.DIST(0,25,0.05,TRUE) = 0.286504797 or

MH POISSON.DIST(0,1.25,FALSE) = 0.286505 = 28.7%

Using the poisson distribution, the probability of finishing the day without slipping is approximately 28.7%

Logic:

(.05) (falls) per (100m)

=(1.25)(falls) per (2.5km)

x = 0

mean = 1.25

Probability mass function (FALSE) used to solve for exactly 0 occurrences.

* 1. What assumptions do I need to make in the questions above? How reasonable are they / which ones do you suspect are violated?

Binomial assumptions

Specific number of trials. This condition is met.

Identical trials. This assumption is violated. It is very likely that some portions of the driveway are relatively easy to walk (increasing the chance of success), while other portions are more challenging (decreasing the chance of success).

Independent trials. This assumption is violated. Dr Rogers will develop skill at not falling on the treacherous driveway with experience, making the chance of success increase on later trials. On the other hand, Dr Rogers may become overconfident and/or fatigued as the journey progresses, decreasing the chance of success on later trials.

Two outcomes. This is a reasonable assumption, as only one of the two outcomes will occur. However, the assumption does not account for the chance of multiple falls in one trial. As a result, this model would be less accurate in predicting the probability of more than one fall during the day.

Maximum number of occurrences is known. This assumption is violated, as Dr Rogers could fall an unknown amount of times in each trial.

Poisson assumptions

The poisson distribution deals with situations where things happen randomly at a given rate over time or space. The rules for the Poisson are that events occur independently over time at a constant random rate and only one event can occur at a time

Independent events. Violated as per binomial assumptions above.

Constant random rate of time. Violated, as per the identical trial reasoning above.

Only one event can occur at a time.

One trial = 100m of walking

25 trials to complete the days walking. (500 back to car, 1 km in to cottage, 1 km back out)

5% chance of falling = 95% chance of not falling

Binomial

Number of successes = 25

Number of trials = 25

Probability of success = .95

Want the probability of exactly 25 successes. Therefore, use the probability mass function, FALSE in the formula

=BINOM.DIST(25,25,0.95,FALSE) = 0.27739 = 27.7%

Using the binomial distribution, the probability of finishing the day without slipping is approximately 27.7%

1. As the owner of a growing online company, I have considerable intellectual property tied up in various bits of data. My technology team has proposed a system involving 3 identical hard drives. If all three drives fail at the same time, my company goes offline. That would be bad.

This may sound like a lot of technology, but we feel we need this kind of redundancy to ensure that our system is available 24 hrs. per day for 365 days a year. Drive failures are rare with new equipment, but we bought some refurbished equipment so each drive has a 5% chance of failing per day. If it fails, it can be repaired when the system goes into maintenance mode just before midnight and be ready for the next day.

* 1. Using language your manager is likely to understand, explain why this could be modelled as a binomial, develop and use that model to determine the probability of having more than 50 drive failures in a year?

Independent events: prob of each drive failing/succeeding is not affected by the other drives

Identical experiments: the percent of failure is the same every day of the year

Two outcomes, success and failure; Success = drive does not fail, Failure = drive fails

Prob of more than 50 failures = 1- cumulative prob of 50 (use TRUE)

# trials = 3 X 365 =1,095

% fail = .05

=1-BINOM.DIST(50,1095,0.05,TRUE) = 0.717559812 = 71.8%

Note: intuitive answer is 5% chance means on average each drive fails once in 20 days

Average number of fails in a year is 54.75. Therefore, we expect the chance of having at least 50 fails to be comfortably over 50%.

* 1. What is the probability that I will go offline at least once in a year?

Prob that all go offline in 1 day = .05 X .05 X .05 = 0.000125

Number of trials = 365

At least one offline event per year = 1- Prob(cumulative 1) =1 - BINOM.DIST(0, 365, 0.000125,TRUE) = 4.46% chance of going offline at least once

* 1. Would it be better to have a system with 2 brand new identical drives, each of which only has a 2% chance of failure?

Trials = 365

Prob that both go offline in 1 day = .02 X .02 = .0004

At least one offline event per year = 1- Prob(cumulative 1)

=1-BINOM.DIST(0,365,0.0004,TRUE) =.13586 = 13.6%