***Blair’s Calculation of Infection Probability***

**by SPC**

In one of the more ominous scenes of *John Carpenter's The Thing*, the scientist Blair is sitting at his computer where he is running a simulation of how the Thing’s cells might assimilate and duplicate the cells of another creature. After finishing the simulation, the computer then performs a calculation of the chances that some people at Outpost #31 are really Things in disguise. It dutifully reports to Blair the results of its calculation:

**PROBABILITY THAT ONE OR MORE TEAM MEMBERS MAY BE INFECTED BY INTRUDER ORGANISM:**

**- 75% -**

Being a big fan of *The Thing* who’s also a grad student studying mathematics, I couldn't help but wonder at the accuracy of Blair's number. Now I realize that when Carpenter and his production crew were putting together the movie they probably didn't whip out their calculators or dust off any math books that happened to be lying around. The 75% figure was just a nice rounded-off number that someone cooked up to communicate to the audience the urgency of the situation at Outpost #31. But still, I had to wonder. Could we possibly do a real probability calculation based on what is shown in the film? And if so, how far off would the number be from the 75% shown on Blair's computer?

Once my curiosity got started, there was no stopping it. So I went ahead and worked through the calculations. I dusted off an old textbook that I used in teaching an undergraduate course in probability and statistics and started perusing through the theorems and formulas to see if anything applied to the hypothetical question of Thing infection. Sure enough, it didn't take long before I found the perfect equation, and the number that it gave was very surprising.

What follows is a high school-level description of how to do Blair's calculation of the infection probability. I have tried to keep the complicated jargon and terminology to a bare minimum, choosing rather to use ordinary words to describe the concepts involved. Still, some parts of this don’t exactly make for the lightest reading. If you have trouble understanding a certain point, then don’t let it bother you. Just move on and maybe it will become clear as you read more. So put your thinking caps on and let's see how Blair could've arrived at his 75% chance that one or more members of his team were alien duplicates of the original people.

***Finding a Formula***

The first thing we must do is come up with some kind of formula or equation that adequately models the situation that Blair was facing. To do this, we reduce Blair's problem to a fairly simple question: if infection occurred at an average rate since the Thing’s escape from the ice block, then what are the chances that one or more infections have happened to the men of Outpost #31? It turns out that mathematicians have worked out a formula for answering precisely this sort of problem. It is called the Poisson ("pwah-sahn") distribution and, written out in words, it says:

**P(N) = (e to power of -R) \* (R to power of N) / (N factorial)**

Let's explain what these letters and symbols mean. The asterisk \* means multiply. The backslash / means divide. P(N) is read "P of N" and denotes the probability of having N infections. The number "e" is a famous mathematical constant whose value is about 2.718. The letter R is the average infection rate (which we will measure as # of infections per day), and N is the number of possible infections. The phrase "N factorial" is interesting. For example, "3 factorial" means 1 x 2 x 3 = 6 and "5 factorial" means 1 x 2 x 3 x 4 x 5 = 120. So "factorial" tells us to multiply all the numbers up to the given number. Written out in a more symbolic form, the Poisson formula looks like:

**P(N) = exp(-R) \*RN / N!**

This equation provides an accurate estimate for the probability of having N number of infections. Since there are only 12 men at Outpost #31, the number of Thing infections can range from 0 up to 12. This means the variable N can take on values 0, 1, 2, 3, 4, ..., 12. If N = 0, then no infections have occurred, and everyone is safe. If N = 12, then everyone at Outpost #31 would be a Thing.

So when Blair asks for the probability of one or more Thing infections, he's dealing with all the cases where N = 1, 2, 3, 4, ..., 12. That is, he's considering all the possibilities except N = 0. Now the rules of probability say that, in order to get the total value, Blair needs to add up the numbers from each of these cases. That is, Blair should do the following:

**P(1 or more infections) = P(1) + P(2) + P(3) + P(4) + ... + P(12)**

Now, if done by hand, this calculation can be time-consuming since it requires Blair to use the Poisson formula twelve times and then add up all those answers to get the final number. Perhaps this is why the film shows Blair using his computer to do the calculation, for a machine could run through these numbers much faster than any human.

But fortunately, there is a simple short cut that greatly reduces the amount of work involved. It comes from the fact that all probabilities must add up to 1.00 or 100%. In the case of Blair's calculation, this gives:

**P(0) + P(1) + P(2) + P(3) + ... + P(12) = 1**

So if we subtract P(0) from both sides, we get:

**P(1) + P(2) + P(3) + P(4) + ... + P(12) = 1 - P(0)**

Now look closely at this equation, especially at the left-hand side of it. We’ve already seen that what's on the left is just the probability of one or more infections! So this expression tells us exactly what Blair is looking for. It says the probability of one or more infections is the same as the number 1 subtracted by the probability of no infections P(0). So we get the following simple equation:

**P(1 or more infections) = 1 - P(no infections) = 1 - P(0)**

So, Blair’s problem has been greatly reduced to this fairly simple expression. But what's even more is that, when we let N = 0, the Poisson equation for P(N) also greatly simplifies. Plugging 0 into the formula gives:

**P(0) = (e to power of -R) \* (R to power of 0) / (0 factorial)**

Now, math students the world over know that a number raised to the power of 0 is just 1 and that "0 factorial" is also 1. So the formula for P(0) becomes:

**P(0) = (e to power of -R) \* 1 / 1 = (e to power of -R) = exp(-R)**

We can now plug this expression into 1 - P(0) to get the following equation for Blair's probability question:

**P(1 or more infections) = 1 – P(0) = 1 - exp(-R)**

Remember that e is equal to about 2.718 and R is just the rate of Thing infections. This formula can be easily done by virtually any handheld calculator (at least scientific calculators).

***Finding the Infection Rate***

It’s okay if you haven’t been able to follow all the steps of our reasoning above. Just understand what the last formula is essentially saying. It’s telling us that Blair’s question ultimately depends on just one number: the daily rate of Thing infections. If we figure out how many infections on average occur per day, then we can just plug that number into the last formula above and we get the answer to Blair’s probability question. For example, suppose R is on average 1 infection per day. Then, according to the formula, this would give a probability:

**P(1 or more) = 1 - exp(-R)= 1 - exp(-1)= 1 - 0.3679 = 0.6321 = 63%**

So, let’s find a good value for the Thing’s infection rate R. Once we know a number for R, we can plug it in and get the answer to Blair’s question.

Since R is a rate of infection, all we have to do is figure out the average number of Thing infections per day. This is precisely what Blair would've had to do. So, let’s put ourselves in Blair's shoes and work only with the information he knew as he did his computer calculation.

Blair first needed to get an idea of the total number of Thing infections that occurred so far. He would have to make a count of how many people and dogs had been infected. Blair would’ve included the dog-Thing and the Norwegian-Thing plus the 2 dogs from Outpost #31's own kennel that were being assimilated when the men so rudely interrupted the process. This gives a total of 4 infections.

But there is also the matter of the 6 missing Norwegians (only 4 of the 10 Norwegians were accounted for) and an unknown number of other Norwegian dogs. How many of these were also infected? Blair would have to make an educated guess. It seems reasonable that, out of this missing group, at least 2 more Norwegians and their dogs were assimilated. This would bring the total number of infections to 4 + 2 = 6. (Note that greater numbers than this would only make the final probability go higher. So it‘s safe to be a little on the conservative side here.)

Having found an estimate for the number of infections, Blair next has to estimate how many days over which these infections took place. For better or worse, this involves a bit more guesswork. Blair knew exactly how long the Thing had been at Outpost #31. The dog-Thing had arrived at around 8:30 AM the previous day, and the time shown on Blair’s computer was 1736 hours or about 5:30 PM (see also [Timeline](file:///F%3A%5CUsers%5CChefles.TONY_WORK_PC%5CDesktop%5CO31%20backup%5Cpublic_html%5Coutpost31.com%5CFanThings%5CFanEssays%5Cmovie%5Ctimeline.html)). This means that by the time Blair did his simulation 33 hours had passed since the Thing’s arrival at Outpost #31. This is the same as 1.375 days.

But there’s also the question of how many days was the Thing at the Norwegian camp after it escaped from the ice block. A very low number would be just 1 or 2 days. Very high values would be around 5 or 6 days. (Although Blair obviously didn’t know it, Outpost #31 itself would only last late into the fifth day.) Like any good scientist, Blair would avoid the extremes in favor of a normal, intermediate value. This would be somewhere around 3 or 4 days. Thus, the total estimated time that the Thing had been infecting men and dogs prior to Blair’s calculation is somewhere between 3 + 1.375 = 4.375 days and 4 + 1.375 = 5.375 days.

We are now ready to find R, the average rate of Thing infections per day. To get this number, we just divide the estimated number of infections by the estimated number of days. This means that R is somewhere between:

**R = 6 infections / 4.375 days = 1.371 infections per day**

**R = 6 infections / 5.375 days = 1.116 infections per day**

***Finding the Probability***

We are finally in a position to determine Blair’s probability. As mentioned before, to get an answer all we have to do is plug the value for R into the equation we found. Since the estimated R occurs over a small range of possible values, from 1.116 up to 1.371, it is best to make a table of possible R values with their corresponding probabilities:

|  |  |  |
| --- | --- | --- |
| **Infection Rate R** | **PoissonFormula** | **CorrespondingProbability** |
| **1.116** | **1 - exp(-1.116) = 1 - 0.3276 = 0.6724** | **67%** |
| **1.15** | **1 - exp(-1.15)= 1 - 0.3166 = 0.6844** | **68%** |
| **1.20** | **1 - exp(-1.20)= 1 - 0.3012 = 0.6988** | **70%** |
| **1.25** | **1 - exp(-1.25) = 1 - 0.2865 = 0.7135** | **71%** |
| **1.30** | **1 - exp(-1.30) = 1 - 0.2725 = 0.7275** | **73%** |
| **1.35** | **1 - exp(-1.35) = 1 - 0.2592 = 0.7408** | **74%** |
| **1.371** | **1 - exp(-1.371) = 1 - 0.2539 = 0.7461** | **75%** |

So, the possible values range from a low of 67% to a high of 75%. Since Blair was apparently looking for just one number, he would’ve instructed his computer to find the highest, or worst case, probability just so he could remain on the safe side of his situation. The computer dutifully reported that, given the data Blair had entered, its search through all the possible values had determined that the probability of one or more team members being infected was 75%.

Incredibly, the number chosen by The Thing’s producers and director is exactly what the real mathematics of the situation would‘ve produced. Quite accidentally, the number seen in the film is precisely the correct value. Now what are the chances of that happening?