TN2004-1: Epidyne – At first blush

George Rebane, PhD
4 April 2020

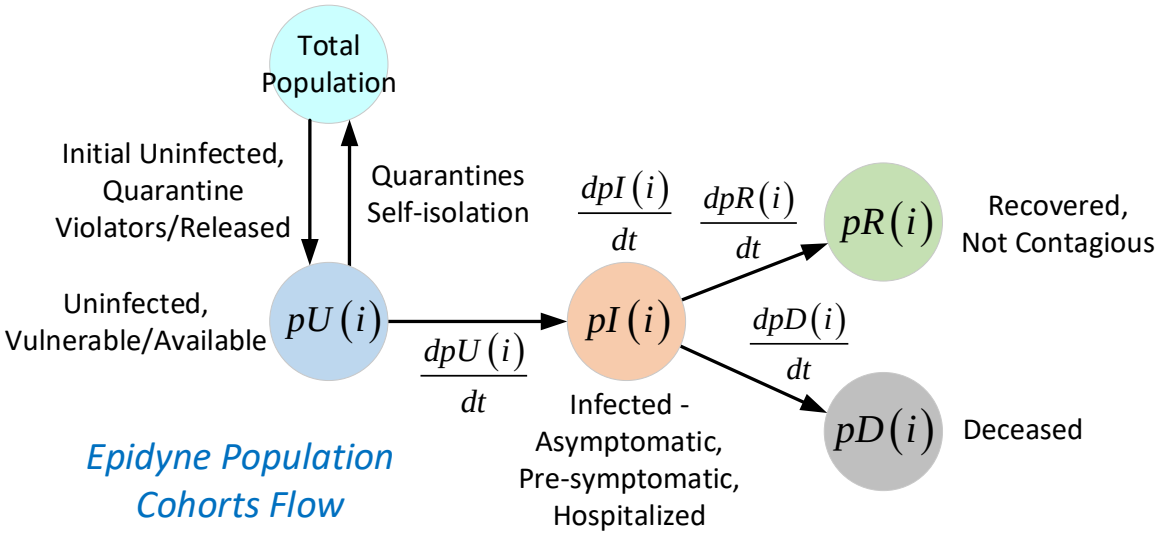
This technical note documents the derivation and mathematical model of a simple epidemic spread model. It incorporates sufficient complexity to make it a useful tool to study and understand the dynamics of the spread of an infectious disease within a target population. With an appropriate set of realworld inputs, Epidyne can also be used to generate and analyze the predicted spread of such diseases. It is also an extremely facile model in the sense that it has been implemented as a MS Excel™ spreadsheet, and in the Matlab™ programming language.

Motivation. In light of the daily reported dilemmas of widely divergent epidemic spread models that are being consulted by our national Covid-19 (C19) response planners, and also the background reports of their various vintages and provenances that make them all more or less difficult to access and use, let alone upgrade or even maintain, I decided to develop my own version so that I could play out various recorded data and policy alternatives, and inform readers of my weblog *Rebane's Ruminations* of results I considered insightful, informative, or at least interesting.

Objective. My objective was to develop a model that was sufficiently complex to simulate and predict realworld experiences and outcomes, but no more than that. The model had to have an adequately rich input space (think of them as control knobs) that could be set to represent both policy alternatives (such as time/size dependent quarantines, and regional populations), and also incorporate field measured data (such as infection rates dependent on virulence and vectors, contagion durations, and onset time lags). Most certainly the model had to be able to handle the critical infection rate dynamics driven by the process known as 'herd immunity'. Epidyne has been programmed in both an Excel™ spreadsheet and the system development language Matlab™. A portion of the spreadsheet model is shown below.

EpidyneTest03		Scenario00	ap =	20,000,000	rImax =	2	popI0 =	60	kI =	2.944438979				
gjr - 3apr20			nI =	4	fI =	0.95	popR0 =	5,000,000						
ref gjr 22mar20			rD =	0.01	pfoffset =	0	popD0 =	-						
© 2020 George J. Rebane, All Rights Reserved														
				Min popU	*Max popI*	*Max popR*	*Max popD*							
				7,575,942	5,710,293	12,349,758	74,240							
	Cum Qtine	*Quarantine*	*T (week)*	*popU*	*popI*	*popR*	*popD*	*popTot*	*popfrac*	*d*	*rI*	*pUdot*	*pIdot*	*pIRdot*	*pIDdot*
	-		1	14,999,940	60	5,000,000	-	20,000,000	2.0000	5.8887	1.994	120	120	-	-
	0		2	14,999,820	180	5,000,000	-	20,000,000	1.9999	5.8885	1.994	239	239	-	-
	0		3	14,999,581	419	5,000,000	-	20,000,000	1.9997	5.8879	1.994	477	477	-	-
	0		4	14,999,103	897	5,000,000	-	20,000,000	1.9993	5.8868	1.994	952	952	-	-
	0		5	14,998,151	1,849	5,000,000	-	20,000,000	1.9985	5.8845	1.994	1,899	1,779	119	1
	0		6	14,996,252	3,628	5,000,119	1	20,000,000	1.9970	5.8801	1.994	3,548	3,308	237	2
	0		7	14,992,705	6,936	5,000,356	4	20,000,000	1.9942	5.8717	1.994	6,598	6,121	473	5
	0		8	14,986,106	13,057	5,000,828	8	20,000,000	1.9889	5.8562	1.994	12,207	11,255	943	10
	0		9	14,973,900	24,312	5,001,771	18	20,000,000	1.9792	5.8277	1.994	22,443	20,545	1,880	19

Epidyne models the time history of the sizes and change rates (time derivatives) of four inter-related population cohorts that comprise a regional population into which an infectious disease is introduced that has the potential of becoming an epidemic. These cohorts are comprised of *pU* - those uninfected yet vulnerable; pI - those currently infected; *pR* - those who have recovered and are no longer infectious; and *pD* - the deceased. *pU* may also be mediated by timed quarantines which can reduce the size of *pU* or increase it when the quarantines end or are (partially) ignored before the epidemic runs its course.



As shown above, the population flows from vulnerable uninfected cohort to the infected cohort, which then feeds the non-decreasing recovered and the deceased cohorts. The asymptomatic and pre-symptomatic infected continue to reduce the uninfected population at a variable 'reproduction' rate *rI* (cf. herd immunity). The course of the disease in an infected individual is completed within *nI* weeks after which the individual has either recovered or died. The mortality rate of the infected is given by *rD*.

Herd immunity is a form of indirect protection from an infectious disease that occurs when a large percentage of the originally uninfected and vulnerable have been reduced by the increasing complementary percentage of the immune (recovered or vaccinated), the currently infectious, and the deceased. This type of virtual or probabilistic immunity provides a significant measure of protection for individuals not yet immune by reducing the encounter chances between the infectious and those uninfected, available, and therefore vulnerable. This has the effect of reducing the aggregate infection rate ('reproduction') within the modeled target population.

The infected individual is a threat and/or a resource drain for a limited time. This is an important point to keep in mind when listening to C19 reports and modeling the spread of the disease. Almost all media reports cite the growing number of the infected as if they were still in circulation infecting the remaining vulnerable cohort. Nothing could be further from the truth. An infected C19 individual will infect others for at most two weeks, most likely closer to one week, after which time he either recovers or enters the healthcare system (i.e. is constructively quarantined) and then either recovers or dies. The way to estimate how many infectious people may still be running around among us is to take the difference of today's infected and those of, say, a week ago, adjusting for growth. It is this short interval of contagion that dictates the end of epidemics which seldom infect the entire vulnerable population.

Before we look at some C19 epidemic scenarios, it's important to understand how spread models, and specifically Epidyne, calculate a dynamic infection rate, *rI*, in the presence of time-varying herd immunity. The discussion that follows will also refer to the Epidyne math model which is a set of iterative equations that compute the population cohort sizes and their rates of change (time derivatives) as presented below.

Cohort Populations Calcs

$$pU(i)=\left[pU(i-1)-\frac{dpU(i-1)}{dt}-pQ(i)\right]\left\{\left[pU(i-1)-pQ(i)\right]>\frac{dpU(i-1)}{dt}\right\}\left[\frac{dpU(i-1)}{dt}>0\right]+\ldots$$

$$\left[pU(i-1)-pQ(i)\right]\left[\frac{dpU(i-1)}{dt}\le 0\right];\ \ pU(1)=a_p;$$

$$pI(i)=\left[pI(i-1)+\frac{dpI(i>n_I)}{dt}\right]\left[pI(i-1)>-\frac{dpI(i>n_I)}{dt}\right]\left[\frac{dpI(i>n_I)}{dt}\neq 0\right]\left\{\left[pI(i-1)+\frac{dpI(i>n_I)}{dt}\right]\le a_p\right\}+\ldots$$

$$\left[a_p-pR(i)-pD(i)\right]\left\{\left[pI(i-1)+\frac{dpI(i>n_I)}{dt}\right]>a_p\right\};\ \ pI(1)=pI_0;$$

$$pR(i>n_I)=\left[pR(i-1)+\frac{dpR(i-1)}{dt}\right]\left\{\left[\frac{dpR(i-1)}{dt}\right]\ge 0\right\}+pR(i-1)\left[\frac{dpR(i-1)}{dt}<0\right];\ \ pR(1:n_I)=pR_0;$$

$$pD(i\le n_I)=\left[pD(i-1)+\frac{dpD(i-1)}{dt}\right]\left\{\left[\frac{dpD(i-1)}{dt}\right]\ge 0\right\}+pD(i-1)\left[\frac{dpD(i-1)}{dt}<0\right];\ \ pD(1:n_I)=pD_0;$$

$$pTot(i)=pU(i)+pI(i)+pR(i)+pD(i);$$

Cohort Rate of Change Calcs

$$f_p(i)=\frac{\left\{pU(i)-\left[pI(i)+pR(i)+pD(i)\right]\right\}}{pI(i)+pR(i)+pD(i)}+f_{po};\ \ \delta(i)=k_I f_p(i);\ \ rI(i)=\frac{rI_{MAX}}{1-e^{-\delta(i)}};$$

$$\frac{dpU(i)}{dt}=\min\left\{rI(i)\left[pI(i)-pI(i-1)\right]\left[pI(i)>pI(i-1)\right],pU(i)\right\};\ \ \frac{dpU(1)}{dt}=rI_{MAX}\,pI(1);$$

$$\frac{dpR(i)}{dt}=\left[1-rID\right]\frac{dpU(i-n_I)}{dt}\left[\frac{dpU(i-n_I)}{dt}>0\right];\ \ \frac{dpR(1:n_I)}{dt}=0;$$

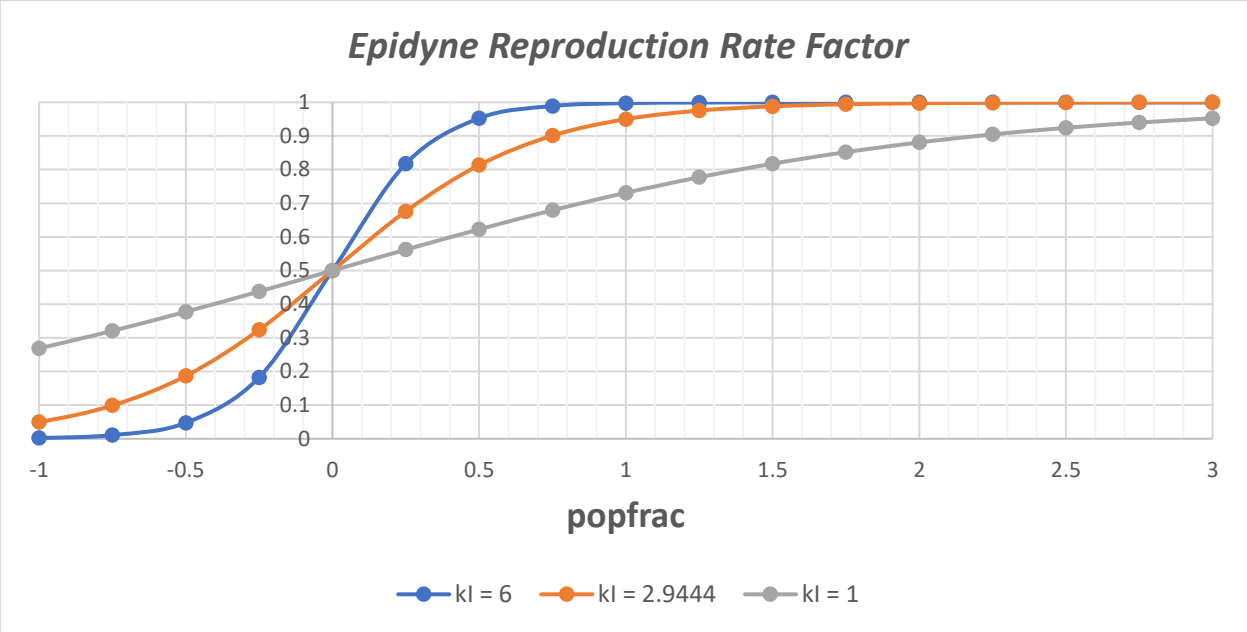
$$\frac{dpD(i)}{dt}=rID\frac{dpU(i-n_I)}{dt}\left[\frac{dpU(i-n_I)}{dt}>0\right];\ \ \frac{dpD(1:n_I)}{dt}=0;$$

$$\frac{dpI(i)}{dt}=\left[\frac{dpU(i)}{dt}-\frac{dpR(i)}{dt}-\frac{dpD(i)}{dt}\right]\left[pI(i)>0\right];$$

[The technical reader will note the extensive use of logicals in the equations. These evaluate to either 1 (True) or 0 (False) depending on whether or not their underlying condition is met. Surprisingly, the use of logicals in expressing the mathematics of complex dynamic systems is not yet widely practiced, with programmers succumbing to the use of the more cumbersome IF and WHILE statements. Embedding logicals directly into the mathematics and its resulting software implementations allows the expression of both in much more clear, concise, and complete format. It is remarkable that so few people know that Excel spreadsheet equations also accept logicals.]

Dynamic Reproduction Rate. Epidyne computes the time-varying reproduction rate, or the average number of vulnerable people a newly infected person will infect in one week, as a fraction of *rImax*, the input maximum infection rate. The fraction – spreadsheet '*popfrac*', first

equation under Rate of Change Calcs – is a sigmoid function that varies from one to zero. The sigmoid is shown below. Its argument *d* (spreadsheet) or δ in the equations is calculated as a scaled value of a population fraction that is the normalized difference between the remaining vulnerable and sum of the non-vulnerable (i.e. the infected, recovered, and deceased) population cohorts. Initially *d* (see spreadsheet above) is very large since there are many vulnerable and few infected as shown in the model. As the epidemic spreads, more people transit from the vulnerable to the non-vulnerable causing *popfrac* to decrease through zero to its minimum value of -1.



This fraction is scaled by the input constant *kI* with the value of 2.9444. *kI* determines how rapidly the sigmoid decreases from one to zero as shown above. The *kI* value for the current model was selected so as to make the infection rate equal to 95% of *rImax*, the input max infection rate, when the remaining vulnerable population was twice as large as those removed from the initial vulnerable population. As seen from the model, at that point $f_p(t) = 1$ and the *popfrac* sigmoid achieves this value as shown in the red trace above. The *popfrac* sigmoid can be shifted laterally as needed, perhaps, to more accurately fit observed field data, through input of the *popfrac* offset constant *pfoffset* > 0. *pfoffset* can be used to delay the onset of herd immunity to the point where it can be completely eliminated from affecting the spread of the epidemic.

Fitting Epidyne to data. The Matlab version of Epidyne is coded so that it can be readily converted into a function subroutine that accepts its inputs in vector form and outputs all the population and rate arrays that it computes. This allows the model to be customized into a predictive tool through the use of any of several optimizing techniques ranging from numerical gradient search to various forms of evolutionary programming.

Sample Scenarios. In the remaining figures we illustrate a series of six epidemic scenarios computed with various sets of inputs. The scenarios assume that the starting populations of 100M and 50M uninfected and vulnerable citizens that reflect the state of our nation which already self-quarantined most of its 330M citizens. All scenarios start with an initial cohort of 100 infected people at the time that self-quarantining goes into effect. The major input parameter variations to note are the impact of various *rImax* values on the extent of the resulting epidemic in both numbers and timing. (The illustrated graphics from Excel have little overshoot transitions that are artifacts of the coding of spline algorithms used to connect the discrete weekly computed values.)

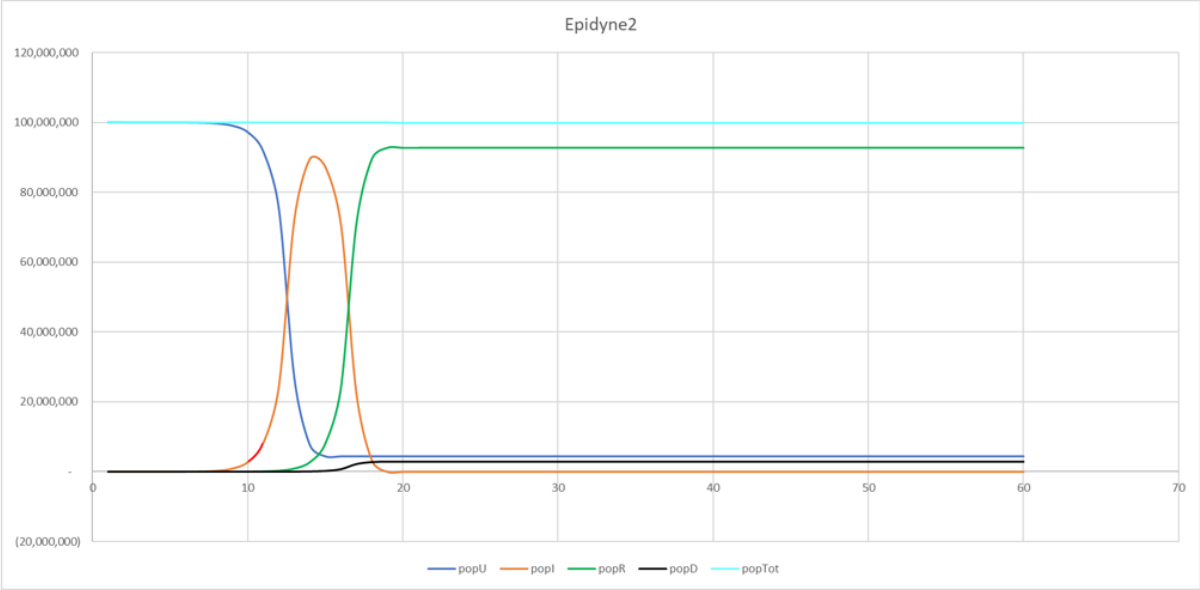
California's low death rate conjecture. Before presenting the scenario series, it may be useful to consider why California so far has had such low infection (*pI*) and deceased (*pD*) numbers, given that it is the most populous state with large urban areas, large homeless population, and a third of the country's welfare recipients. Several people (e.g. academic Victor Davis Hanson) have suggested that the reason may be that C19 landed in the state early since California is the major gateway into the US from Asia. If it turns out true that many asymptomatic and pre-symptomatic travelers arriving last fall from Asia had already infected a large number of Californians before recovering (and some dying of non-C19 attributed causes), then when the recognizable pandemic hit earlier this year it encountered a California with an already large recovered and immune cohort (*pR*) in its population.

Add to that the state's early imposition of the 'stay at home' dictum, which at least half of California's 40M obeyed, then it is easy to see why this year's 'main C19 invasion' could have encountered a population that already had a large degree of herd immunity. The above graphic, showing the top portion of the spreadsheet model, illustrates how Epidyne may be used to model such a high level of initial herd immunity within a population that mostly adheres to a self-quarantine policy. In that scenario we assume that this March 20M Californians chose not to stay at home and become part of the cohort that avails itself to the newly arrived infected (here starting at *pI0* = 60). However, 25% or 5M of these folks are already immune and captured by setting *pR0* = 5M. The effect of this herd immunity is immediately seen in the greatly reduced initial value of *popfrac*.

When this particular 'California scenario' is run, Epidyne projects that of those 20M people an additional 7.5M will recover, with another 7.5M never becoming infected before the epidemic runs its course. That means that only 7.5M Californians were infected, with the overwhelming share of them recovering and joining the immune ranks. The total C19 deaths will number under 75,000. The epidemic will end in about six months (25 weeks) from onset, and reach its highest intensity (5.7M infected) in week 19. The highest weekly death rate will reach a little over 19,000 in week 21 before it plummets as the epidemic burns itself out. We note that with significant herd immunity at the outset, the infection rate (*rI*) already starts below *rImax* = 2, and then rapidly reduces to an unsustainable rate of less than 0.5 during weeks 16 through 20.

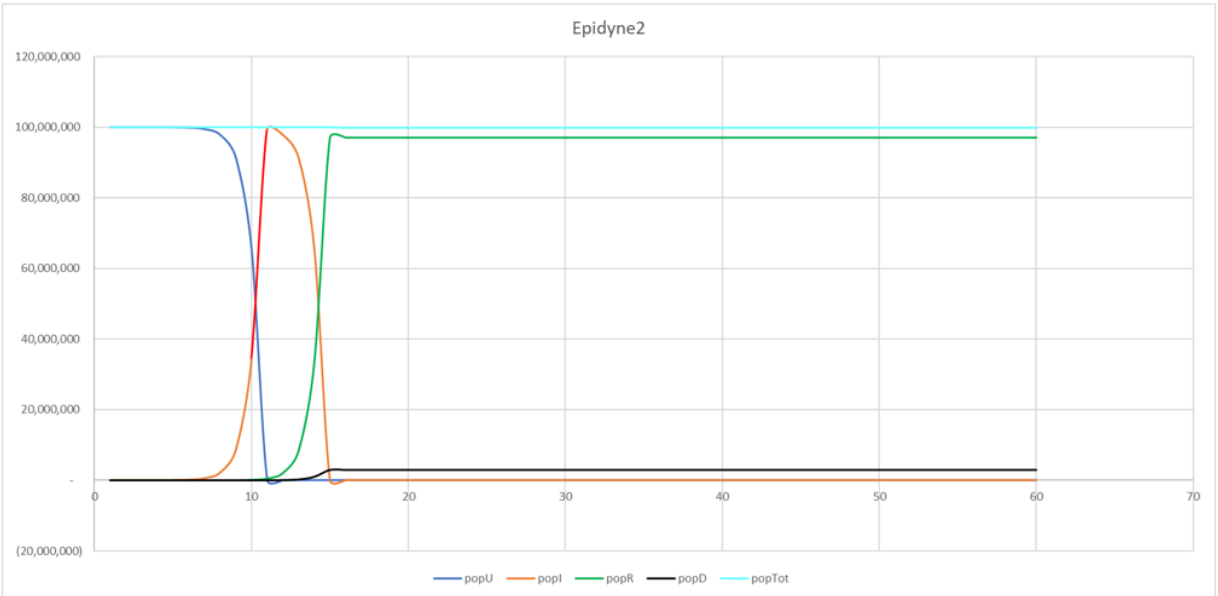
Scenario01	ap =	100,000,000	rImax =	3	popI0 =	100
	nI =	4	fI =	0.95		
	rD =	0.03	kI =	2.944438979		

Uninfected	*Max Infected*	*Recovered*	*Deceased*	*MaxI wks*	*Duration wks*
4,315,196	89,596,599	92,814,260	2,870,544	14	20



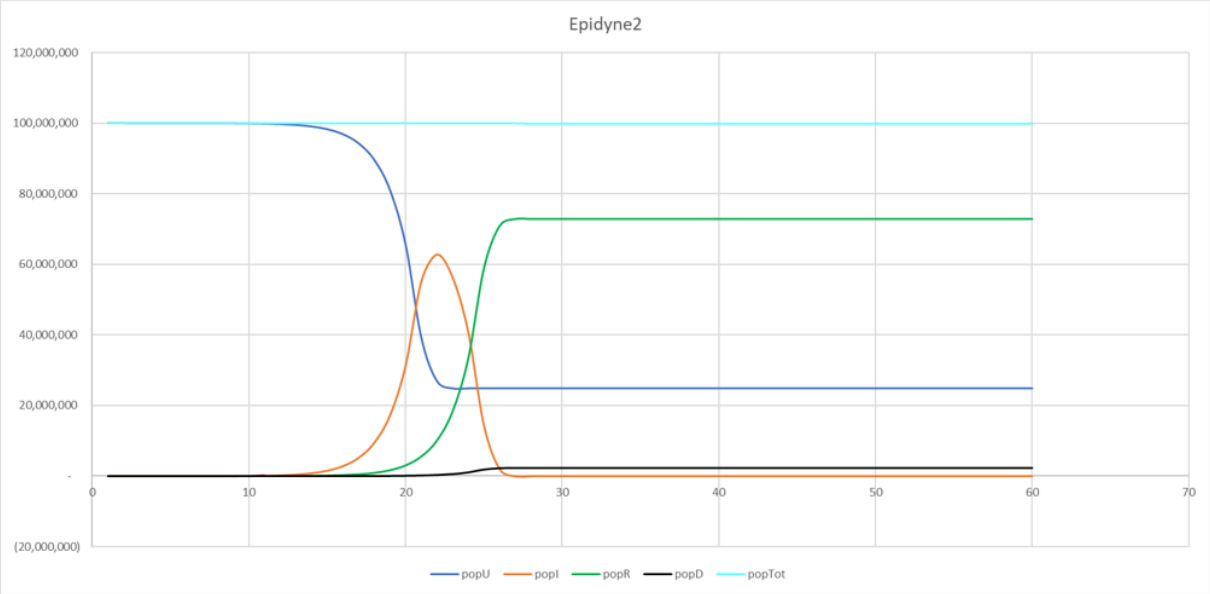
Scenario02	ap =	100,000,000	rImax =	4	popI0 =	100
	nI =	4	fI =	0.95		
	rD =	0.03	kI =	2.944438979		

Uninfected	*Max Infected*	*Recovered*	*Deceased*	*MaxI wks*	*Duration wks*
-	99,455,700	97,000,000	3,000,000	11	16



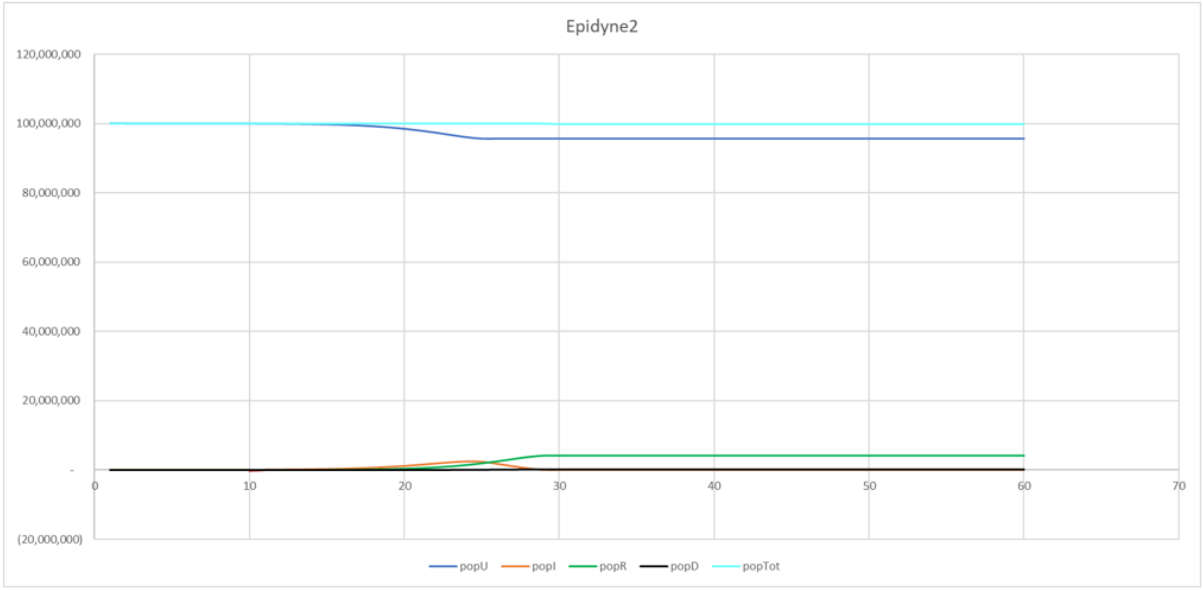
Scenario03	ap =	100,000,000	rImax =	2	popI0 =	100
	nI =	4	fI =	0.95		
	rD =	0.03	kI =	2.944438979		

Uninfected	*Max Infected*	*Recovered*	*Deceased*	*MaxI wks*	*Duration wks*
24,800,944	62,716,992	72,943,085	2,255,972	22	28



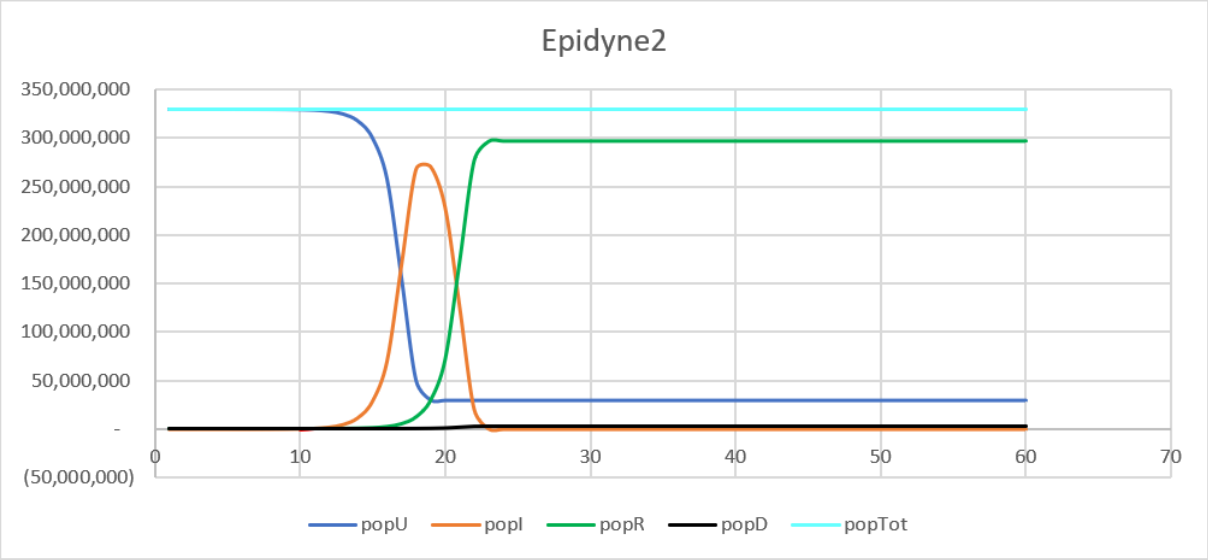
Scenario04	ap =	100,000,000	rImax =	1.8	popI0 =	100
	nI =	4	fI =	0.95		
	rD =	0.03	kI =	2.944438979		

Uninfected	*Max Infected*	*Recovered*	*Deceased*	*MaxI wks*	*Duration wks*
95,637,146	2,423,808	4,231,969	130,886	24	30



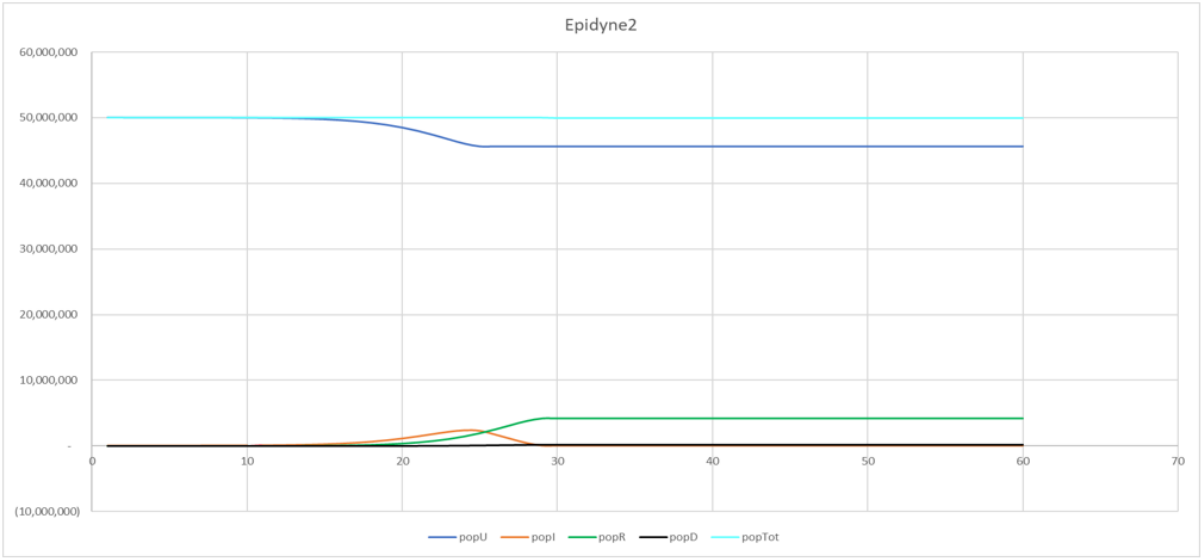
Scenario05	ap =	50,000,000	rImax =	2	popI0 =	100
	nI =	4	fI =	0.95		
	rD =	0.03	kI =	2.944438979		

Uninfected	*Max Infected*	*Recovered*	*Deceased*	*MaxI wks*	*Duration wks*
12,031,589	31,643,821	36,829,358	1,139,052	21	27



Scenario06	ap =	50,000,000	rImax =	1.8	popI0 =	100
	nI =	4	fI =	0.95		
	rD =	0.03	kI =	2.944439		

Uninfected	*Max Infected*	*Recovered*	*Deceased*	*MaxI wks*	*Duration wks*
45,637,146	2,423,808	4,231,969	130,886	24	30



[Disclaimer – I am not an epidemiologist. I am a retired systems scientist and professional engineer with degrees in Physics, Estimation, Optimization, and Control Theory, Complex Dynamic Systems, and Bayesian Inference and Machine Learning. I spent my career in technical enterprises designing and modeling the performance of complex dynamic systems both for the Department of Defense and commercial enterprises. Some of these systems included animal subsystems. I have also taught at the university level as an adjunct professor. Epidyne is presented here only as an educational tool to study the potential behavior of epidemic spread that conforms to the stated assumptions underlying the model. Epidyne is admittedly what is often called a 'toy model', and therefore enjoys an ease of access and facile use not available with the larger and more cumbersome models from academe. However, in evaluating its performance re realworld scenarios, Epidyne will speak for itself. That is one of the fortunate aspects of developed science and technology; critics of this work need only show how/where Epidyne fails in its conclusions and predictions, and there it may easily be compared with any of the gaggle of contending epidemic spread models which today are used to drive national and state C19 response policies. I may be contacted through the comment stream of this post on *Rebane's Rumination*s by readers who leave their return email address when they log on to comment. Their email address will not display.]