

Corona virus modell

Lars Åström

Mars 2020

1 Introduktion

I denna rapport presenteras en modell som har använts och anpassats för att prediktera antalet fall av Corona viruset Covid-19. Syftet med denna prediktion är att ha en skattning för hur många sjukvårdsplatser som behövs som mest under sjukdomsförloppet.

Det skulle gå att prediktera antalet personer med hjälp av statistisk regression eller liknande där man sätter upp en statistisk modell. Detta skulle kunna vara exempelvis exponentiell eller linjär tillväxt. Nackdelen med dessa två modeller är att antalet sjuka patienter just nu går mot oändligheten, vilket inte är rimligt långsiktigt eftersom personer bli immuna.

Modellen som används i denna rapport har tagit sin inspiration från Lotka-Volterras ekvationer, vilka beskriver hur många individer det finns av två arter (som kallas byten och rovdjur) över tid. Denna utgår från att antalet rovdjur ökar då antalet byten är stort och vice versa. Se https://en.wikipedia.org/wiki/Lotka%E2%80%93Volterra_equations.

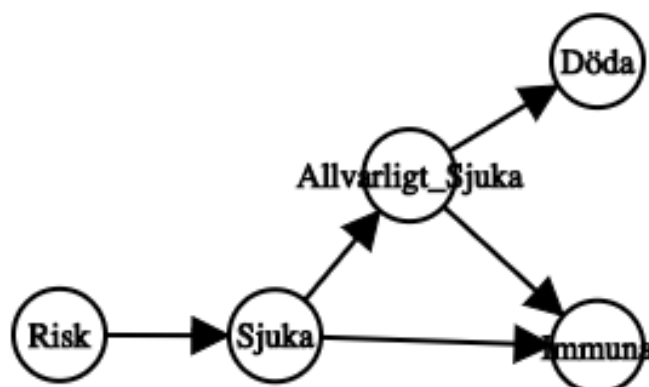
2 Modellen

2.1 Modellens uppbyggnad

I modellen säger vi att det finns fem olika kategorier av individer vid varje given tidpunkt. Dessa kategorier är beskrivna i följande tabell:

Kategori	Beskrivning
Risk	De som riskerar bli sjuka, alltså ej haft sjukdomen än
Sjuka	De som just nu har sjukdomen
Allvarligt sjuka	De som just nu har sjukdomen och uppsöker vård, detta är en delmängd av alla sjuka
Immuna	De som har haft sjukdomen och nu är immuna
Döda	De som avlidit med sjukdomen

Enligt vår modell färdas individer genom följande:



Figur 2.1: Figur som visar hur en individ färdas mellan de olika kategorierna.

Som vi ser färdas individerna endast enligt pilarnas riktning och går aldrig tillbaka. I vår modell antar vi också att "Allvarligt sjuka" är en delmängd av alla sjuka. Vi antar i modelleringen att alla individer finns i exakt en av grupperna "Risk", "Sjuka", "Immuna" och "Döda".

2.2 Beskrivning av flödet mellan kategorierna

Vi börjar med att införa beteckningar för hur många individer det finns i varje kategori. För att göra detta låter vi:

$$\begin{aligned} r(t) &= \text{Antal i Risk vid tiden } t \\ s(t) &= \text{Antal Sjuka vid tiden } t \\ a(t) &= \text{Antal Allvarligt Sjuka vid tiden } t \\ i(t) &= \text{Antal Immuna vid tiden } t \\ d(t) &= \text{Antal Döda vid tiden } t \end{aligned}$$

Vi mäter tiden t i dagar. Notera att $r(t) + s(t) + i(t) + d(t)$ antas vara konstant för alla t , eftersom vi antar att alla individer befinner sig i exakt en av dessa fyra grupper.

Vi går igenom varje pil för sig i figur 2.1 för att förklara antaganden.

Antalet personer som går från Risk till Sjuka antar vi är proportionellt mot antalet möten mellan sjuka och riskindivider. Det betyder att vi antar att antalet som går från Risk till

Sjuka kan beskrivas som $\beta r(t)s(t)$, där β är en positiv konstant. Detta kommer från att antalet möten mellan individer från Risk och Sjuka antas vara proportionellt mot $r(t)s(t)$.

Vidare antar vi att andelen som är Allvarligt Sjuka, alltså uppsöker sjukvård, är en lika stor andel av alla sjuka. Låt denna andel kallas γ , alltså har vi $a(t) = \gamma s(t)$.

Vi antar att individer är sjuka under L dagar vilket gör att L dagar efter att en individ blir sjuk blir denna antingen immun eller död. Antag att andelen som dör är α . Från detta ges hur förändringen av antal immuna och döda beror på hur många individer som gick från risk till sjuka var L dagar tidigare.

Dessa beskrivningar kan sammanfattas med följande matematiska formler. Notera att för en funktion $f(t)$ betecknar $f'(t)$ derivatan, alltså förändringen, av funktionen.

$$i'(t) = (1 - \alpha)\beta r(t - L)s(t - L) \quad (1)$$

$$d'(t) = \alpha\beta r(t - L)s(t - L) \quad (2)$$

$$s'(t) = \beta s(t)r(t) - \beta r(t - L)s(t - L) \quad (3)$$

$$r'(t) = -\beta s(t)r(t) \quad (4)$$

$$a(t) = \gamma s(t) \quad (5)$$

För att underlätta beräkningar diskretiserar vi modellen och låter f_x beteckna värdet av f vid dag x , alltså $f(x)$. Skillnaden är att nu kan vi bara betrakta f_x för heltal x , inte vilket värde som helst. Derivatan av en funktion $f'(x)$ ersätter vi med $f_x - f_{x-1}$. Vi ändrar också lite grann i ekvationerna från 1 för att hela tiden kunna skapa en skattning från en dag till nästa. Denna skillnad är att vi antar att antalet som går från Risk till Sjuka dag t är proportionellt mot antalet möten mellan personer i Risk och Sjuka dag $t - 1$, vilket inte förefaller orimligt. Det ger oss:

$$i_t - i_{t-1} = (1 - \alpha)\beta r_{t-L}s_{t-L} \quad (6)$$

$$d_t - d_{t-1} = \alpha\beta r_{t-L}s_{t-L} \quad (7)$$

$$s_t - s_{t-1} = \beta s_{t-1}r_{t-1} - \beta r_{t-L}s_{t-L} \quad (8)$$

$$r_t - r_{t-1} = -\beta s_{t-1}r_{t-1} \quad (9)$$

$$a_t = \gamma s_t \quad (10)$$

Dessa antaganden och ekvationer ger ett sätt att givet hur många som finns i respektive kategori en dag förutspå dagen efter. Denna procedur kan sedan upprepas godtyckligt långt vilket ger att en fullständig tidsserie kan skapas. För att få en bra prediktion krävs dock att (förutom att antagandena inte är för långt ifrån verkligheten) att parametrarna α, β, γ, L väljs på ett korrekt sätt. Dessa kan väljas på åtminstone två sätt. Det första är att betrakta vad expertis säger om exempelvis dödlighet (alltså α) och antal dagar som individer är sjuka (L). Utöver dessa värden behöver vi ange antalet personer som har sjukdomen när vi startar tidsserien, låt denna variabel kallas K .

Det som vi kan observera är a_t och d_t . Vi vill alltså välja parametrarna så att dessa värden stämmer överens med observationerna som finns. När vi väljer parametrar har vi i denna

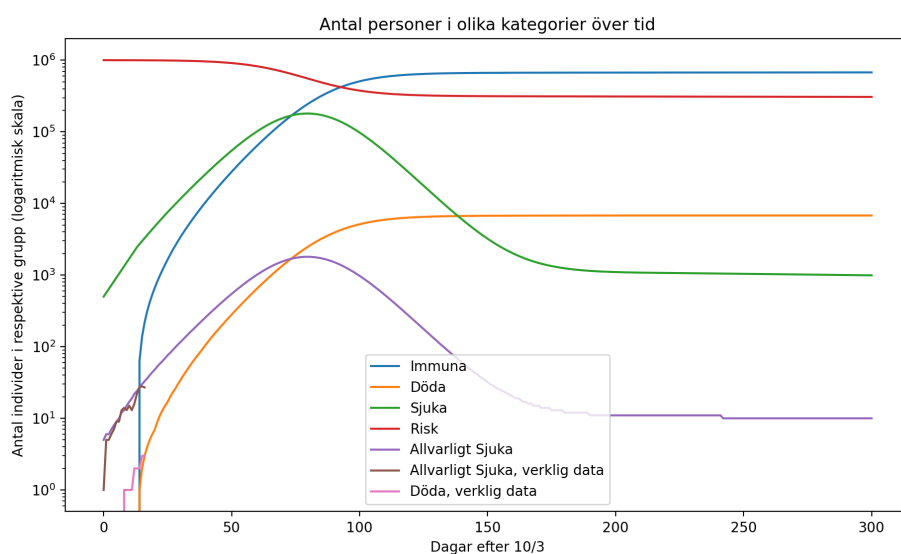
modell försökt att så långt som möjligt gå på vad expertisen säger samt använda sunt förnuft när vi väljer parametrarna. Dock måste det poängteras att det är väldigt osäkra prediktioner som ges.

3 Resultat

Genom att använda följande parametrar:

Parameter	Värde
α	0.01
β	$1.3 \cdot 10^{-7}$
γ	0.01
K	500
L	14

ges följande prediktion:



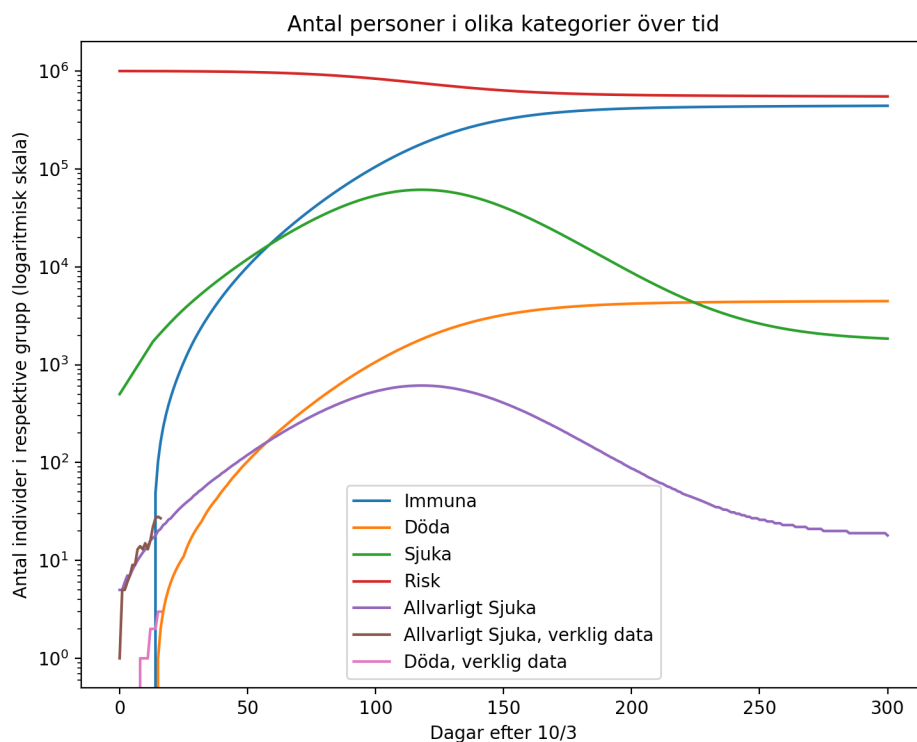
Figur 3.1: Prediktion av antal i de olika grupperna över tid. Denna prediktion är mycket osäker.

Denna prediktion är alltså väldigt osäker och beror på dels modellens validitet i dess antaganden och dels valet av parametrar. Dessutom har många olika parametervärden testats och vi har testat tills datan är anpassad. Det är alltså inte alls konstigt att prediktionerna är väldigt nära den verkliga datan. Dock kan fortfarande modellen säga något om framtiden om vi tror att den kan vara bra med dessa parametrar.

För att belysa att mycket saker kan ändras när parametervärden ändras så visar vi här hur det ser ut med följande parametrar:

Parameter	Värde
α	0.01
β	$1.0 \cdot 10^{-7}$
γ	0.01
K	500
L	14

Vi har alltså endast bytt värde på β från $1.3 \cdot 10^{-7}$ till $1.0 \cdot 10^{-7}$. Det har gjort att toppen på hur många individer som kommer vara allvarligt sjuka har dels flyttats från att inträffa ca 20 maj och vara på ca 2000 individer till att inträffa ca 1 juli och vara på ca 600 personer.



Figur 3.2: Prediktion av antal i de olika grupperna över tid. Denna prediktion är mycket osäker.

4 Utveckling av parametrarnas innebörd

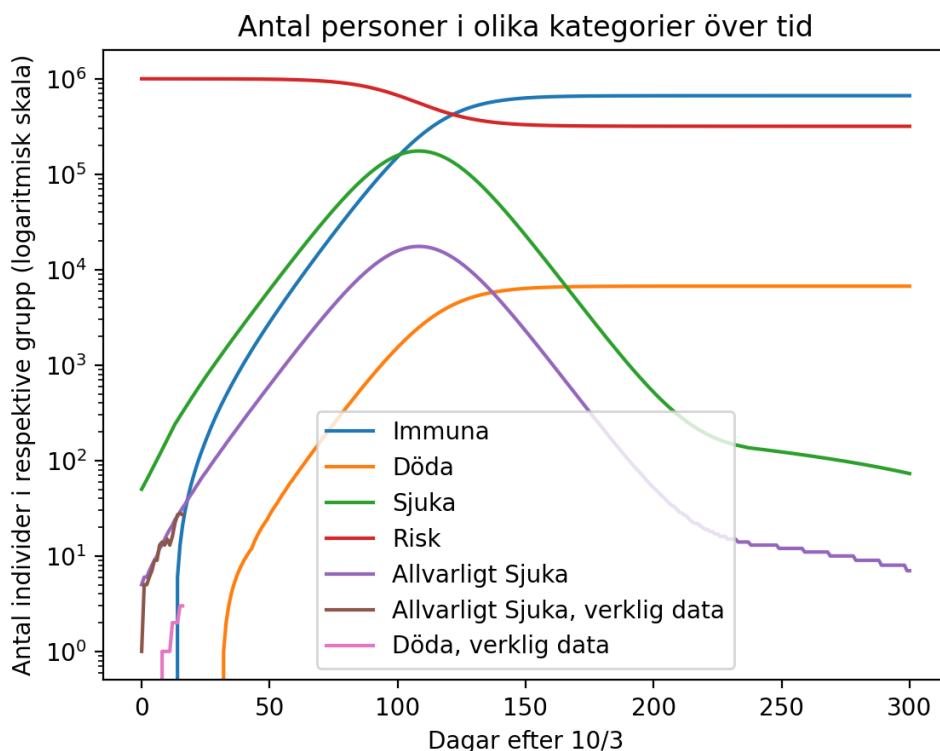
Vi ska här beskriva varje parameter lite till för att förklara dess innebörd och hur den påverkar förloppet av de olika kategoriernas storlek.

α är andelen som dör när de inte längre har sjukdomen. Denna parameter ändrar bara förhållandet mellan antalet immuna och antalet döda.

β är en parameter som i princip säger hur många individer varje sjuk individ smittar. Ett högt β innebär att en sjuk individ smittar många från riskgruppen. Detta ger ett

snabbare förlopp av hela händelseförloppet samt att det maximala antalet allvarligt sjuka vid toppen blir högre. Notera att β kan bero på individers vanor och rörelsemönster, alltså påverkas denna parameter över tid av exempelvis politiska beslut.

γ är andelen av alla individer som är sjuka som uppsöker vård. Denna parameter är väldigt svår att skatta och veta vad den borde vara. Om denna ökas (samtidigt som K minskas för att få lika många som uppsökt vård som vi observerat) så infaller toppen senare och blir högre, alltså att fler blir allvarligt sjuka.



Figur 4.1: Prediktion av antal i de olika grupperna över tid. Här används ett större γ än i de tidigare prediktionerna. Notera att antalet allvarligt sjuka vid toppen är väldigt stort.

L är antalet dagar en individ är sjuk. Större L innebär att varje sjuk individ hinner smitta fler andra.

K är antalet personer som var sjuka den 10 mars. Denna siffra är väldigt svårskattad. Vi kan ha en hyfsad skattning av $K \cdot \gamma$ eftersom vi observerat detta, men förloppet blir väldigt olika beroende på vilken av dessa parametrar som är stor.