

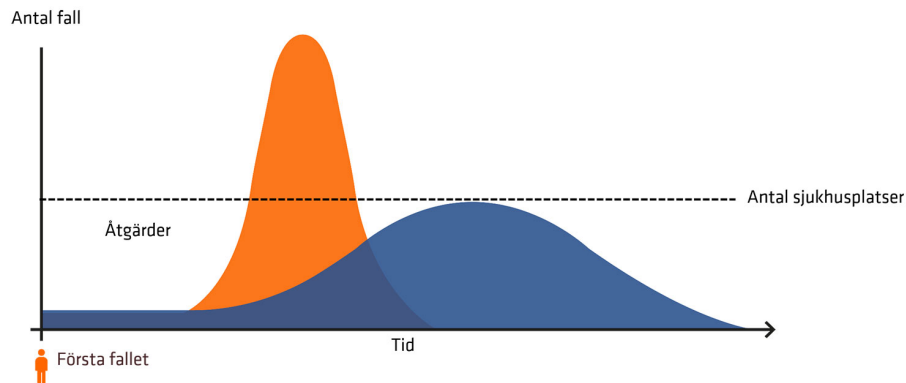
Pandemins Matematik

Vetenskapens Hus
Anna Karlhede

September 2020

1 Introduktion

Vi kan använda matematiska modeller för att beskriva smittspridning. Detta är viktigt för att kunna förutspå vad som kommer hända i framtiden och förbereda vården och samhället. Dessa modeller kan också användas till att undersöka effekten av olika åtgärder så som vaccinering. Det finns en mängd olika matematiska modeller för att beskriva smittspridning, [här](#) kan du läsa om några olika. Kärnan i dessa modeller är *differentialekvationer*, och det är genom att lösa dessa ekvationer som vi kan skapa de infektionskurvor vi har kunnat se överallt det senaste halvåret, speciellt i 'flatten the curve'-bilder, lik Figur 1.



Figur 1: Typisk 'flatten the curve'-bild från FHM.

Även om ni väljer att inte lösa dessa differentialekvationer så är det viktigt att *förstå* dom. Så - låt oss kika lite på dom!

2 Differentialekvationer

Definitionen av en differentialekvation är helt enkelt att det är en ekvation som beskriver ett samband mellan en okänd funktion och minst en derivata av den funktionen. En derivata beskriver i vilken takt funktionen ändrar sig. Mer exakt är en funktions derivata i en viss punkt lutningen på funktionen i just den punkten.

2.1 Typiskt gymnasieexempel

Ett typiskt exempel från Matte 5 i gymnasiet (som ni kanske ännu inte läst) är bakterietillväxt. Låt $y(t)$ vara mängden bakterier vid tidpunkt t . Förändringshastigheten, alltså hur snabbt bakteriekulturen ökar, kan då beskrivas genom differentialekvationen

$$y'(t) = k \cdot y(t), \quad (1)$$

där k är någon proportionalitetskonstant. Det ekvationen ovan säger, i vardagligt språk, är att hur snabbt mängden bakterierna ökar ($y'(t)$) är proportionellt mot antalet bakterier som finns just nu ($y(t)$). Alltså, ju fler bakterier vi har, desto snabbare ökar antalet bakterier! Se till så att du förstår varför detta är rimligt. Lösningen till denna ekvation är en *exponentialkurva*. Vi har alltså exponentiell tillväxt (= den snabbaste typen av tillväxt).

Detta är en *första ordningens* differentialekvation eftersom vi enbart har en förstaderivata ($y'(t)$).

Detta exempel är taget [härifrån](#). Kika där för att lära dig mer/uppfriska minnet.

3 SIR-modellen

Låt oss nu kika på differentialekvationerna som utgör den klassiska SIR-modellen. I denna modell är alltså alla människor antingen mottagliga för sjukdomen (*susceptible*), har sjukdomen (*infected*), eller har antingen tillfrisknat och blivit immuna eller dött (*removed/recovered*).

Våra tre grupper kan beskrivas med följande funktioner;

- $S(t)$ betecknar antalet människor som är mottagliga för sjukdomen vid tid t .
- $I(t)$ betecknar antalet människor som är sjuka vid tid t .
- $R(t)$ betecknar antalet människor som inte längre kan få sjukdomen vid tid t .

Ofta pratar vi om *andelen* av befolkningen, istället för de exakta siffrorna. Andelen får vi ju helt enkelt genom att dela mängden med det totala antalet människor N (i till exempel Sverige). Då får vi istället de snarlika funktionerna nedan.

- $s(t) = S(t)/N$ betecknar *andelen* av befolkningen som är mottaglig för sjukdomen vid tid t .
- $i(t) = I(t)/N$ betecknar andelen av befolkningen som är sjuka vid tid t .
- $r(t) = R(t)/N$ betecknar andelen av befolkningen som inte längre kan få sjukdomen vid tid t .

Övning: Ungefär hur kommer var och en av $s(t)$, $i(t)$ och $r(t)$ att förändras över tid? Försök rita kurvorna!

Övning: Varför gäller det att $s(t)+i(t)+r(t) = 1$ och att $S(t)+I(t)+R(t) = N$ vid varje tidpunkt t ?

Vi kan ta reda på vad var och en av dessa funktioner är vid varje given tidpunkt genom att lösa differentialekvationerna som utgör SIR-modellen, som vi snart ska gå in på.

3.1 Antaganden

I SIR-modellen gör en ett antal antaganden. Dessa är:

- Inga nya tillkommer till den mottagliga gruppen (vi ignorerar nya födslar och invandring). Genom att folk blir sjuka försvinner det folk från denna grupp, och det är enda sättet som antalet mottagliga förändras. Vi ignorerar alltså utvandring och dödsfall som orsakas av andra orsaker.
- Antalet som insjuknar varje dag beror på antalet som är mottagliga för sjukdomen, antalet som redan är sjuka, samt mängden kontakt mellan mottagliga och sjuka. Låt denna mängd beskrivas med b som anger hur många människor en sjuk individ har smittsam kontakt med varje dag (alltså sådan kontakt att sjukdomen sprids vidare givet att kontaktpersonen är mottaglig för den).

(Ibland skrivs b som produkten av transmissionssannolikheten och kontaktfrekvensen (*eng*: transmission probability, contact rate))
- En viss andel k tillfrisknar varje dag.

3.2 Mottagliga för sjukdomen, $S(t)$ & $s(t)$

$S(t)$ betecknar antalet människor som är mottagliga för sjukdomen vid varje given tidpunkt.

Övning: Vad betyder $\frac{dS}{dt}$? Notera att detta är detsamma som $S'(t)$.

$S'(t) = \frac{dS}{dt}$ anger hur snabbt $S(t)$ ändras, alltså hur många fler/färre som är mottagliga för sjukdomen vid varje tidpunkt (exempelvis varje dag). Detta är ju detsamma som hur många som insjuknar vid varje tidpunkt (tex dag). Den första differentialekvationen i SIR-modellen beskriver just detta, enligt

$$\frac{dS}{dt} = -b\frac{S(t)}{N}I(t) = -bs(t)I(t), \quad (2)$$

Övning: Pausa här och fundera över varför ekvationen ser ut som den gör. Försök översätta den till vardagligt språk! Varför är det ett minustecken? Var kommer faktorn $I(t)$ ifrån?

Ledning: Ökar eller minskar antalet mottagliga? Vad påverkar hur många som blir sjuka? Kika på vårt andra antagande! Kommer alla b människor en sjuk individ har smittsam kontakt med vara mottagliga för sjukdomen?

Svar: Minustecknet kommer ifrån att antalet människor som är mottagliga för sjukdomen *minskar* vid varje tidpunkt. Hur många som insjuknar vid varje tidpunkt beror ju på hur många som redan är sjuka, $I(t)$, och hur stor andel av de b smittsamma kontakter som sker med mottagliga individer, $S(t)/N$.

Motsvarande differentialekvation för $s(t)$ ges av

$$\frac{ds}{dt} = -b\frac{s(t)}{N}i(t) = -bs(t)i(t). \quad (3)$$

3.3 Tillfrisknade och avlidna

Differentialekvationen för andelen tillfrisknade och avlidna, betecknad med $r(t)$ (notera: *andelen*) ges av

$$\frac{dr}{dt} = ki(t) \quad (4)$$

Övning: Varför ser den ut såhär? Verkar det rimligt? Kom ihåg våra antaganden!

3.4 Infekterade av sjukdomen, $I(t)$ & $i(t)$

Som nämnt i en övning ovan gäller att $s(t) + i(t) + r(t) = 1$ (se till så att du förstår varför!). Från detta har vi

$$\frac{ds(t)}{dt} + \frac{di(t)}{dt} + \frac{dr(t)}{dt} = 0. \quad (5)$$

Övning: Förklara varför detta gäller.

Från allt vi gjort tidigare, kan vi då få differentialekvationen för $i(t)$, som ges av

$$\frac{di}{dt} = bs(t)i(t) - ki(t) \quad (6)$$

Övning: Hur kom vi fram till denna ekvation?

Ledning: Känner du igen termerna till höger om likhetstecknet?

3.5 Sammanställning

Vi har alltså tre differential ekvationer enligt nedan;

$$\frac{ds}{dt} = -bs(t)i(t) \quad (7)$$

$$\frac{di}{dt} = bs(t)i(t) - ki(t) \quad (8)$$

$$\frac{dr}{dt} = ki(t). \quad (9)$$

Vi kan lösa dessa med hjälp av en dator för att ta reda på hur många/hur stor andel mottagliga/sjuka/tillfrisknade+avlidna vi har vid varje tidpunkt. Men vi behöver en sista sak - nämligen de initiala villkoren. Dessa är hur många som är mottagliga/sjuka/tillfrisknade+avlidna vid tidpunkt 0. Om vi tar Sverige som exempel, med en befolkning på ca 10,23 miljoner, så skulle dessa kunna vara $S(0) = 10.230.000$, $I(0) = 5$, och $R(0) = 0$. Lösningarna beror naturligtvis också på k och b , som varierar över tid och beror på vilken sjukdom det är.

Slutligen kan vi nämna att reproduktionstalet R_0 , som Tom Britton pratade om i videon, ges av

$$R_0 = \frac{b}{k}. \quad (10)$$

4 Pandemins Matematik

Vi har ovan kikat närmre på en av de matematiska modellerna som beskriver smittspridning. Detta är den 'enklaste' (med det sagt är den fortfarande väldigt användbar) och det finns en mängd olika, som gör olika komplicerade antaganden. Ni behöver inte arbeta med den här modellen, eller lösa differentialekvationer för den delen - men det är viktigt att förstå bakgrunden till, och principen för, modellerna.

Möjliga 'fokus' för ett gymnasiearbete kan vara, men är ej begränsat till:

- Beräkna en modells infektionskurvor (alltså lös differentialekvationerna med hjälp av en dator) och jämför med faktiskt data från FMH.
- Undersök olika riskgrupper på något vis. Möjliga förslag skulle kunna vara:
 - Litteraturstudie om/varför vissa riskgrupper är mer mottagliga för olika sjukdomar.
 - Jämför data från olika riskgrupper. Är vissa riskgrupper mer mottagliga för exempelvis covid-19? Notera att detta förutsätter att denna data finns att tillgå.
- Jämför olika matematiska modeller. Detta kan göras såväl teoretiskt som 'praktiskt'. Vag gör de för antaganden? Hur ser differentialekvationerna ut? Hur påverkar det resultatet? Hur ser lösningarna ut?
 - Det har nyligen rapporterats fall där människor som tidigare haft covid-19 har insjuknat igen. Då är SIS-modellen mer lämplig (denna antar att efter att en har tillfrisknat från sjukdomen så är en igen mottaglig för den).
- Hur föll tidigt gjorda prediktioner ut?
- Litteraturstudie av något lämpligt papper på området. Eventuellt ett arbete av Tom Britton, kanske [detta](#).

Kom ihåg att det är *ert* gymnasiearbete och att ni därför får välja vilket fokus/ämne som helst! Detta är bara några förslag. Gymnasiearbetet är stort, så det är inte konstigt om det känns övermäktigt i stunder. Notera också att ni inte behöver jobba med just covid-19. Om ni vill lösa differentialekvationerna själva (med hjälp av en dator) så kommer det krävas att ni programmerar lite. Har ni aldrig programmerat tidigare kan det kännas väldigt svårt. Men - gymnasiearbetet är som sagt stort, så låt det ta tid! Om ni vill lära er programmera så finns det massor av resurser online (och kanske även på era skolor?). Vi på VH kan också guida er i rätt riktning. Om ni pluggar vidare inom natur/matematik/teknik är det också väldigt värdefullt att ha sett lite programmering tidigare.

Lycka till!!!