



Stockholms
universitet

Stokastisk modellering av smitt- spridning inom ishockey

Axel Åström

Kandidatuppsats 2017:19
Matematisk statistik
Juni 2017

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stokastisk modellering av smittspridning inom ishockey

Axel Åström*

Juni 2017

Sammanfattning

I denna rapport undersöks smittspridningen inom ishockey med hjälp av en stokastisk modell definierad i diskret tid. I modellen kan individer ha kontakter på en lokal nivå respektive en global nivå. Där den lokala nivån syftar på kontakter inom lag medan kontakter på den globala nivån sker när individer inom ett lag har kontakter med personer från ett annat lag i samband med att lagen spelar match mot varandra. Huvudfrågeställningen i arbetet är således att upptäcka om epidemins slutstorlek skiljer sig mellan två olika typer av spelscheman genom att utföra diverse simuleringar. Där de båda spelscheman som används utgår från att lagen spelar match var tredje dag. Skillnaden är att i det slumpmässiga schemat lottas spelordningen mellan lagen för varje matchdag medan det verkliga schemat utgår från ett spelschema som användes i den svenska ishockeyligan SHL under säsongen 2016-2017. Resultatet blev då att slutstorleken verkade vara oberoende av vilket av de två spelscheman som användes. Däremot visade det sig att under vissa förutsättningar kunde epidemins slutstorlek minska samtidigt som smittsannolikheten inom lag ökade.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: [axel_eik95\(at\)msn.com](mailto:axel_eik95(at)msn.com). Handledare: Mia Dejifen och Pieter Trappman.

Innehåll

1. Inledning	3
2. Modellbeskrivning	4
2.1 Kontaktstruktur och population	4
2.2 Modellformulering och beteckningar	6
2.3 Övergångsfördelningar	10
2.4 Reproduktionstalet R_0	12
2.5 Sammanfattning av modellen	13
3. Simulering	14
4. Resultat och slutsats	15
5. Diskussion	24
6. Referenser	26
7. Appendix	27

1. Inledning

Under ishockeysäsongen 2014-2015 var det minst 15 stycken spelare, fördelat över olika lag i den nordamerikanska ligan NHL, som smittades av påssjukan. Det troligaste är att dessa individer har smittats av någon inom laget och att själva smittspridningen mellan lagen har skett i samband med att de olika lagen har spelat matcher mot varandra. För när man undersöker inom vilka lag som spelare har rapporterats smittade med ligans spelschema så finns det vissa mönster. Dessutom är det inom ishockeyn, i likhet med många andra sporter, förekommande att man delar vattenflaskor och handdukar samt brukar utövarna utsättas för många nära kontakter i båda omklädningsrummet och vid kamp-situationer ute på isen, kring sargerna och målburarna.[10] Men utöver det är påssjukan, som är ett virus, en smitta som är svår att skydda sig emot eftersom den är mycket smittsam och sprids genom små droppar i utandningsluften. Inkubationstiden brukar vara omkring två till tre veckor, därefter brukar man vara smittsam i ungefär sex dagar innan tydliga symtom som exempelvis svullnad i ansiktet uppstår. Efter det är man smittsam i ytterligare nio dagar. Dessutom får endast hälften av personerna som drabbas av påssjukan symtom men de kan ändå smitta andra under en lika lång tidsperiod som personerna som drabbas och får symtom. Dock kan man endast få påssjukan vid ett tillfälle innan man blir immun och utöver det ingår även vaccin mot det i vaccinationsprogrammet för barn i Sverige sedan 1982 vilket uppskattningsvis ger ett 88 procentigt skydd.[6,10]

Det kan därför vara av intresse att undersöka spridningen av en smitta inom SHL, som är den högsta svenska ishockeyligan, genom att ställa upp en stokastisk smittmodell. Anledningen till att vi väljer att kolla på SHL istället för NHL är för att NHL är en liga som är svårare att förklara. Eftersom, i skrivande stund, är det hockeyliga bestående av 30 lag från USA och Kanada. Sedan när det gäller spelschemat är lagen indelade i östra respektive västra konferensen och dessa konferenser är i sin tur uppdelade i divisioner vilket leder till att lagen exempelvis inte möter alla motsändare lika många gånger. Vidare har klubbarna samarbetsavtal med lag från andra ligor vilket gör att spelare kan hoppa fram och tillbaka mellan olika lag. Samtidigt som den svenska ligan består av 14 lag där alla möter alla fyra gånger under totalt 52 omgångar, däremot är det precis som i NHL vanligt att ett fåtal spelar pendlar mellan olika lag.[7]

Alltså för att det ska bli möjligt att göra en modell som inte är allt för komplex är vi i behov av att göra olika antagande gällande populationen och hur smittan överförs. Vi får då en modell där individer kan smittas av personer inom det egna laget eller av någon från ett annat lag i samband med att deras lag spelar match mot varandra. Där vi bestämmer att lagen spelar match var tredje dag, vilket innebär att smittprocessen pågår i maximalt 156 dagar. Det vi då ska

använda modellen till är att undersöka om smittspridningen, av en viss smitta, skiljer sig mellan två olika typer av spelscheman genom att göra simuleringar. De ena schemat som vi kallar för det verkliga schemat utgår från att lagen möts enligt samma spelordning som under SHL säsongen 2016-2017.[8] Medan det andra schemat som heter det slumpmässiga schemat utgår från att matchkombinationerna slumpas vid varje matchdag där sannolikheten för varje matchkombination är lika stor.

Dessutom är det av intresse att undersöka hur slutstorleken i den specifika populationen påverkas när man ändrar på parametrar som exempelvis smittsannolikhet inom lag. Resultatet av simuleringarna blir då att slutstorleken är oberoende av vilket spelschema som används. Däremot får vi under vissa förutsättningar att slutstorleken minskar när smittsannolikheten inom lag ökar. Vilket är ett resultat som man, till en början, kanske inte förväntar sig.

2. Modellbeskrivning

2.1 Kontaktstruktur och population

I denna modell antar vi alltså att individerna i den specifika populationen har en kontaktstruktur på två nivåer. Den ena nivån som vi betecknar som den lokala nivån beskriver kontakten mellan individer som tillhör samma lag. Medan den andra nivån som vi kallar för den globala nivån istället förklarar hur smittan sprids mellan lagen. Denna typ av kontaktstruktur liknar en modell som heter hushållsmodellen.[2] Där den största skillnaden är att i den här modellen antar vi att alla individer i populationen inte har samma sannolikhet att ha kontakt med personer från andra lag. För i vår modell spelar alla lagen match samtidigt var tredje dag. Därför antar vi att individer endast kan ha kontakter på den globala nivån under matchdagar. Vilket är rimligt med tanke på att alla lagen har sin hemvist i olika städer runt om i Sverige.[8] Således gäller det att personer inom ett lag A endast kan ha kontakt med personer från lag B om lag A och lag B spelar match mot varandra. Det betyder att vi under matchdagar har totalt sju olika matcher där en eventuell smitta kan spridas mellan de 14 lagen. Sedan om vi istället fokuserar på den lokala nivån är det rimligt att anta att individer inom ett lag har lika mycket kontakt med varandra varje dag oavsett om de spelar match eller inte. Det är inte särskilt konstigt för som spelare i SHL är ishockey ett heltidsjobb och därför träffas personer inom ett lag i princip varje dag. Det ska även påpekas att vi i denna modell inte enbart inkluderar spelare för utom spelare finns det i ett lag också tränare och materialförvaltare som kan bli smittade respektive föra smittan vidare. Därav är det lämpligt att anta att varje lag består av 30 individer och därför består populationen av totalt 420 individer.[8]

Däremot är vi i behov av att göra några antaganden för att inte göra modellen för komplicerad. Ett antagande är att alla personer har lika mycket kontakt med alla personer i sitt lag oavsett om det gäller en spelare, tränare eller materialförvaltare. På samma sätt om två lag A och B spelar match mot varandra är sannolikheten att exempelvis en person i lag A har en kontakt med någon i lag B lika stor för alla individer i lag B. Sedan eftersom vi i denna modell endast kommer att undersöka hur smittan sprids mellan spelare, ledare och materialförvaltare i de 14 olika lagen behöver vi anta att inga individer i populationen har några gemensamma kontakter utanför populationen. Vilket egentligen inte är ett särskilt troligt antagande med tanke på att exempelvis de olika spelarnas familjemedlemmar troligtvis träffas i samband med matcher. Men om vi utgår från en situation där flera personer i den specifika populationen har blivit smittade är det inte orimligt att dessa utomstående förbjuds att umgås för att minska smittspridningen. Däremot är det inget problem utifrån modellens perspektiv att personer i de olika lagen umgås med sina familjer eller utomstående vänner förutsatt att samma typ smitta som är av intresse inte får någon spridning utanför populationen. Dessutom måste vi anta att ingen person i populationen kan skadas eller smittas av någon annan sjukdom som håller dem borta från laget eftersom vi endast kommer ha en modell som tar hänsyn till den specifika smittan.

Alltså kan individer smittas på två nivåer och det brukar benämnas som en heterogent mixad kontaktstruktur.[4] Anledningen till varför vi gör det, istället för att använda en homogent mixad kontaktstruktur, är för att vi antar att smittsamheten mellan två individer är högre om de spelar i samma lag än om de spelar i olika lag. Eftersom det i verkligheten troligtvis är vanligare att de smittsamma kontakterna sker i exempelvis omklädningsrummet istället för ute på isen. Däremot är det också viktigt att vi tar hänsyn till att denna kontaktstruktur inte fångar upp verklighetens komplexitet fullt ut. Eftersom det skulle leda till en allt för komplicerad modell som eventuellt skulle bli mycket svårtolkad. För när man har allt för många parametrar som kan variera i en stokastisk modell kan det vara svårt att se vilka parametrar som egentligen påverkar slutresultatet. Vi väljer därför att nöja oss med denna kontaktstruktur då vi anser att den troligtvis fångar upp det mest väsentliga av smittspridningen inom och mellan lag.

Vidare vill vi förtydliga att det är en stokastisk modell som används istället för en deterministisk modell för att försöka förklara smittspridningen. Däremot har en deterministisk modell, vilket är en modell som utesluter slumpmässighet, några fördelar. Framför allt är de enklare att analysera och kan därför vara mer komplexa. Eftersom när det gäller de stokastiska modellerna behöver de vara ganska enkla och därav inte helt realistiska om de ska vara matematiskt hanterbara. Däremot finns det skäl till att stokastiska modeller är att föredra när deras analys är möjlig. Exempelvis är det mer naturligt att anta att spridningen av en

smitta mellan två individer är stokastisk till skillnad från att huruvida en spridning kommer att ske eller inte redan är bestämt.[2] Det gäller också vid simulering av stokastiska modeller att smittprocessen kan dö ut innan en epidemi uppstår trots att det så kallade reproduktionstalet, som vi förklarar noggrannare i avsnitt 2.4, är över 1. Eftersom de slumpmässiga effekterna kan vara betydande om man studerar en population bestående av relativt få individer och initialt en smittsam individ. Medan det för deterministiska modeller gäller att en epidemi uppstår om reproduktionstalet är större än 1 respektive dör ut innan en epidemi uppstår om reproduktionstalet är mindre än 1.[3]

2.2 Modellformulering och beteckningar

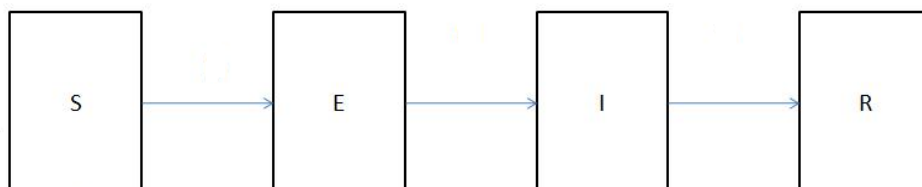
Modellen ska alltså användas för att undersöka hur en smitta sprids inom och mellan lag och om vi utgår från hur ett vanligt virus kan se ut är det lämpligt att anta att modellen har fyra stycken tillståndskategorier som en individ kan befinna sig i. Dessa betecknas S(Susceptible), E(Exposed), I(Infective) och R(Removed). Där de olika beteckningarna står för:

- S: Mottagliga individer i populationen som inte blivit infekterade.
- E: Individer som befinner sig i inkubationstadiet och kan därför inte smitta eller uppvisa sjukdomssymptom.
- I: Infekterade individer som kan smitta andra i sin omgivning
- R: Individer som återhämtat sig eller plockats bort då de har uppvisat sjukdomssymptom. Dessa kan således inte smittas eller smitta andra i populationen.

Det gäller också att modellen definieras i diskret tid där vi antar att varje tidsteg representerar en dag. Det betyder att händelserna, det vill säga förflyttningarna mellan tillstånden, inträffar eller inträffar inte vid ett specifikt tidpunkt varje dag.[10] Det skulle däremot vara mer realistiskt att använda sig av en kontinuerlig tid eftersom en smittöverföring kan ske när som helst när två personer har kontakt med varandra. Men med tanke på att individerna inom vår population inte har kontakt med varandra dygnet runt innebär det att smittsannolikheterna skulle behöva variera över tiden vilket i sin tur skulle resultera i att modellen blir mer komplicerad. Så genom att använda diskreta tidsteg får vi en modell som blir enklare att simulera. Dessutom anser vi att det ger en tillräcklig beskrivning av hur det ser ut i verkligheten. Eftersom de flesta av individerna i den specifika populationen endast har kontakt med varandra under några timmar av dygnet på ett liknande sätt i princip varje dag. Sedan när det gäller data över exempelvis sjukdomars inkubationstid respektive smittsamhetstid brukar dessa skattningar

ges i diskret tid.

Det innebär att vi har en SEIR-modell där en individ kan förflyttas mellan de olika tillstånden på ett sätt som illustreras i Figur 1 nedan.



Figur 1: En SEIR modell

I Figur 1 ser vi att från tillstånd S är det, vid varje tidsteg, endast möjligt att stanna i S eller flyttas till E . På samma sätt kan en individ som befinner sig i E respektive I , vid varje tidsteg, stanna i de tidigare nämnda tillstånden eller förflyttas till I respektive R . Vi kan också se att tillstånd R är ett absorberande tillstånd vilket definieras som att sannolikheten för att en individ som befinner sig i tillstånd R stannar i R är lika med ett för varje givet tidsteg.[5] Där vi har antagit att de som befinner sig i tillstånd R är personer som har tillfrisknat eller personer som är smittsamma men har fått tydliga symptom och därför flyttats till karantän.

Således har vi fyra stycken tillståndskategorier som en individ kan befinna sig i och det kan därför vara lämpligt att införa användbara beteckningar innan vi går vidare med att formulera modellen. Dessa beteckningar presenteras tillsammans med en kortare beskrivning i Tabell 1 nedan. Där vi för tydlighetens vill nämna att $t = 1, 2, \dots, 156$ eftersom modellen är definierad i diskret tid.

Beteckning	Beskrivning
π_L	Sannolikheten att en mottaglig individ i S smittas av en infekterad individ från samma lag per tidsteg
π_G	Sannolikheten att en mottaglig individ i S smittas av en infekterad individ från från motståndarlaget under matchdag
$P_{SE}^{(t)}$	Sannolikheten att en individ flyttas från S till E vid tid t givet en matchfri dag
$P_{SEM}^{(t)}$	Sannolikheten att en individ flyttas från S till E vid tid t givet en matchdag
P_{EI}	Sannolikheten att en individ flyttas från E till I vid ett givet tidsteg
P_{IR}	Sannolikheten att en individ flyttas från I till R vid ett givet tidsteg
S_t	Antal individer i S vid tid t
E_t	Antal individer i E vid tid t
I_t	Antal individer i I vid tid t
R_t	Antal individer i R vid tid t
S_{ti}	Antal individer i S vid tid t för lag i
E_{ti}	Antal individer i E vid tid t för lag i
I_{ti}	Antal individer i I vid tid t för lag i
R_{ti}	Antal individer i R vid tid t för lag i
$Q_{SE}^{(t)}$	Totala antalet individer som flyttas från S till E vid tid t
$Q_{EI}^{(t)}$	Totala antalet individer som flyttas från E till I vid tid t
$Q_{IR}^{(t)}$	Totala antalet individer som flyttas från I till R vid tid t
$Q_{SE_i}^{(t)}$	Antalet individer för lag i som flyttas från S till E vid tid t
$Q_{EI_i}^{(t)}$	Antalet individer för lag i som flyttas från E till I vid tid t
$Q_{IR_i}^{(t)}$	Antalet individer för lag i som flyttas från I till R vid tid t

Tabell 1: Användbara beteckningar

Men utöver dessa beteckningar i Tabell 1 ska vi också förtydliga att $i = 1, 2, \dots, n$ där n i det här fallet är lika med 14 eftersom vi antagit att individerna i populationen är indelade i 14 stycken olika lag. Det gäller därför för varje tidsteg att

$$S_t = \sum_{i=1}^n S_{ti} \quad E_t = \sum_{i=1}^n E_{ti} \quad I_t = \sum_{i=1}^n I_{ti} \quad R_t = \sum_{i=1}^n R_{ti}$$

$$Q_{SE}^{(t)} = \sum_{i=1}^n Q_{SE_i}^{(t)} \quad Q_{EI}^{(t)} = \sum_{i=1}^n Q_{EI_i}^{(t)} \quad Q_{IR}^{(t)} = \sum_{i=1}^n Q_{IR_i}^{(t)}$$

Dessutom är det även vanligt att man i epidemimodeller antar att populationstorleken är konstant under utbrottet.[3] Vilket också är ett rimligt antagande

i vår modell eftersom själva tidsrymden är ganska kort då säsongen i vår modell består av 156 dagar. Däremot är det i verkligheten inte särskilt vanligt att lagen består av exakt samma personer under en hel säsong eftersom spelare försvinner respektive tillkommer på grund av exempelvis skador och spelarövergångar. Men i det här fallet anser vi att dessa förändringar kan försummas då det vanligtvis endast handlar om ett fåtal individer i varje lag.[8]

Detta innebär alltså att om vi bestämmer att den totala populationsstorleken är lika med N har vi för varje tidsteg t att $N = S_t + E_t + I_t + R_t$. Dessutom om antalet individer i varje lag i är lika med M gäller det även att $M = S_{ti} + E_{ti} + I_{ti} + R_{ti}$ för alla i och t . Där N och M i detta fall är lika med 420 respektive 30.

Således kommer smittprocessen, som startar i $t = 0$, inledas med att en individ i populationen väljs ut slumpmässigt och flyttas till tillstånd I medan de övriga $N - 1$ individerna börjar i tillstånd S . Vi har därav vid $t = 0$ att

$$S_0 = N - 1 \quad E_0 = 0 \quad I_0 = 1 \quad R_0 = 0$$

Alltså för exempelvis $t = 1$ respektive $t = 2$ kan den infekterade individ endast smitta personer inom det egna laget. Eftersom vi har bestämt att lagen spelar match var tredje dag och det leder då till att smittan som tidigast kan överföras till individer i ett annat lag vid tidsteg $t = 3$. Därefter fortsätter processen på samma sätt med två stycken matchfria dagar följt av en matchdag där smittan kan överföras till andra lag. Det gäller då precis som vi såg i Figur 1 att en mottaglig individ som smittas först flyttas till inkubationstillståndet E , därifrån till smittsamhetstillståndet I innan individen tillslut hamnar i det absorberande tillståndet R . Men utöver detta är det också viktigt att nämna att på grund av vårt antagande gällande att ingen i populationen kan smittas av någon utanför populationen betyder det att smittprocessen har upphört när

$$S_t = N - k \quad E_t = 0 \quad I_t = 0 \quad R_t = k$$

för $1 \leq k \leq N$. Där k i det här fallet representerar den såkallade slutstorleken, det vill säga hur många som smittas totalt.[3] Men eftersom vi har en process som endast pågår i 156 dagar kan de innebära att individer fortfarande befinner sig i tillstånden E eller I när smittprocessen upphört. Speciellt när de förväntade antal dagar som en individ befinner sig i tillstånd E respektive tillstånd I ökar. Däremot är det inte ett problem eftersom när säsongen är slut är det rimligt att anta att individerna undviker att träffa varandra tills smittspridningen har död ut. I dessa fall gäller det istället att slutstorleken är lika med summan av antalet individer som befinner sig i tillstånden E , I och R vid tidsteg t lika med 156.

2.3 Övergångsfördelningar

Vi har alltså antagit diskreta tidsteg och en konstant populationsstorlek. Vidare ska vi nu titta på fördelningarna av antalet personer som flyttas mellan de fyra olika tillståndskategorierna. Först och främst har vi tidigare antagit att sannolikheten till kontakt mellan individer skiljer sig beroende på om det är en matchdag eller en matchfri dag. Det betyder alltså att överföringsfördelningarna av antalet individer som flyttas från S till E kommer vara olika för matchdagar respektive matchfria dagar.

- Matchfri dag: Under en matchfri dag har vi tidigare kommit fram till att en infekterad individ som befinner sig i tillstånd I endast kan smitta individer inom samma lag som befinner sig i tillstånd S . Således är sannolikheten att en mottaglig individ undviker att bli smittad av en infekterad inom samma lag under ett tidsteg lika med $1 - \pi_L$. Därav är sannolikheten att en mottaglig individ undviker att bli smittad av alla infekterade individer inom individens lag i vid tidsteg t lika med $1 - P_{SE}^{(t)} = (1 - \pi_L)^{I_{ti}}$. Vilket är ekvivalent med $P_{SE}^{(t)} = 1 - (1 - \pi_L)^{I_{ti}}$ som är sannolikheten att en mottaglig individ i tillstånd S blir smittad under en matchfri dag vid tid t . Vi antar därför att antalet individer som flyttas från S till E vid tid t inom lag i är binomialfördelat enligt:

$$Q_{SE_i}^{(t)} \sim Bin(S_{ti}, P_{SE}^{(t)})$$

- Matchdag: Under matchdag har vi antagit att individerna har lika mycket kontakt inom laget som om det vore en matchfri dag. Dessutom har de kontakt med individerna i motståndarlaget med en likafördelad sannolikhet. Det gäller därför att sannolikheten att en mottaglig individ i lag i underviker att bli smittad av de infekterade individerna i det egna laget i respektive de infekterade individerna i motståndarlaget j är lika med $1 - P_{SE_M}^{(t)} = (1 - \pi_L)^{I_{ti}}(1 - \pi_G)^{I_{tj}}$. Vilket innebär att sannolikheten att en mottaglig individ i lag i ska smittas under en matchdag då lag i möter lag j är lika med $P_{SE_M}^{(t)} = 1 - (1 - \pi_L)^{I_{ti}}(1 - \pi_G)^{I_{tj}}$. Det i sin tur resulterar till att antalet individer som flyttas från S till E vid tid t , givet matchdag, inom lag i är binomialfördelat enligt:

$$Q_{SE_i}^{(t)} \sim Bin(S_{ti}, P_{SE_M}^{(t)})$$

Denna modell för att förklara spridningen av en smittsam sjukdom mellan individer liknar en epidemimodell som kallas för Reed-Frost modellen. Den utgår från en SIR och brukar anses vara den enklaste epidemimodellen, vilket däremot

inte betyder att den är lätt att förklara matematiskt. Några antagande som görs i Reed-Frost är att alla individer är lika mottagliga för smittan och lika benägna att sprida smittan vidare. Sedan är det en stokastisk process i diskret tid där antalet personer som smittas vid varje tidsteg är binomialfördelat med en sannolikhet som beror på hur många smittsamma kontakter individen har. Det gäller även alla infekterade kan smitta alla vid varje tidsteg, vilket inte gäller i vår modell eftersom vi har kommit fram till att de personer som en infekterad individ har kontakt med bland annat beror på om det är en matchfri dag eller en matchdag. Dessutom har man i Reed-Frost modellen antagit att tiden som en infekterad individ befinner sig i tillstånd I är deterministisk som i sin tur är ett antagande som vi inte kommer att ha i vår modell.[2]

I vår modell antar vi att tiden som en individ befinner sig i tillstånd E respektive tillstånd I är stokastisk och anledningen till att vi gör det är främst för att vi anser att dessa tider brukar skilja sig mellan olika individer. Därav har vi att sannolikheten att en individ som befinner sig i E flyttas till I vid varje tidsteg är lika med P_{EI} . Således är sannolikheten att individen ska stanna i tillstånd E lika med $(1 - P_{EI})$. Det gäller också att värdet för P_{EI} är konstant, det vill säga oberoende av t , eftersom sannolikheten P_{EI} inte beror på antal kontakter som en individ har vid tid t . Därför antar vi att antalet dagar D som en smittad individ befinner sig i tillstånd E är geometriskt fördelad så $D_E \sim Geo(P_{EI})$ med väntevärdet $E[D_E] = \frac{1}{P_{EI}}$. Det leder då till att antalet individer som flyttas från E till I vid tidsteg t för lag i är binomialfördelat enligt

$$Q_{EI_i}^{(t)} \sim Bin(E_{ti}, P_{EI})$$

Vidare vill vi nu bestämma fördelningen för antal personer som flyttas från I till R vid tid t . Liksom tidigare börjar vi med att konstatera att sannolikheten att en individ som befinner sig i I flyttas till R respektive stannar i I är lika med P_{IR} respektive $(1 - P_{IR})$ oberoende av t . Vilket innebär att även P_{IR} är konstant och därför har vi att fördelningen för antal dagar som en smittad individ befinner sig i I är geometriskt fördelad enligt $D_I \sim Geo(P_{IR})$ med väntevärdet $E[D_I] = \frac{1}{P_{IR}}$. Således har vi att antalet individer som flyttas från I till R vid tidsteg t för lag i är binomialfördelat så att

$$Q_{IR}^{(t)} \sim Bin(I_{ti}, P_{IR})$$

2.4 Reproduktionstalet R_0

Inom epidemiologi är reproduktionstalet R_0 en intressant storhet som är relaterad till huruvida en stor epidemi är möjlig eller inte. Det brukar tolkas som "genomsnittligt antal smittsamma kontakter man har under sin smittoperiod".[3] Dessutom gäller det att när R_0 är mindre än ett dör smittspridningen ut innan en epidemi uppstår. Medan det finns en risk för epidemi om R_0 är större än ett. Men på grund av slumpmässiga effekter kan slutstorleken i dessa fall ändå bli liten.

Däremot är det komplicerat att bestämma R_0 analytiskt för modellen som vi har definierat eftersom vi har en kontaktstruktur på två nivåer. Där strukturen för hur smittan sprids mellan lagen framför allt försvårar beräkningarna. Men vi ska ändå illustrera en metod för hur R_0^L inom lag kan beräknas analytiskt under förutsättningarna att vi endast har ett lag bestående av $n = 30$ individer där endast en given individ kan smitta de mottagliga individerna. Vilket betyder att de individer som smittas av den givna individen, i princip, flyttas direkt till tillstånd R utan att smitta några andra mottagliga individer. Det i sin tur innebär att vi får ett rimligt värde för R_0^L om latensperioden är relativt lång samtidigt som sannolikheten π_L är låg. Vi får därav att

$$R_0^L = (n - 1)P(A)$$

där $P(A)$ är sannolikheten att en smittsam individ smittar en given annan individ i populationen. Vi får alltså enligt lagen om total sannolikhet[1] att

$$P(A) = \sum_{t=1}^{\infty} (1 - P_{IR})^{t-1} P_{IR} P(A|B_t)$$

där B_t innebär att individen återhämtar sig dag t . Vilket betyder att

$$P(A|B_t) = 1 - (1 - \pi_L)^t$$

Således kan vi få fram att reproduktionstalet R_0^L för denna situation är lika med

$$\begin{aligned} R_0^L &= (n - 1) \sum_{t=1}^{\infty} (1 - P_{IR})^{t-1} P_{IR} (1 - (1 - \pi_L)^t) \\ &= \frac{n - 1}{1 - P_{IR}} \sum_{t=1}^{\infty} (1 - P_{IR})^t P_{IR} (1 - (1 - \pi_L)^t) \\ &= \frac{n - 1}{1 - P_{IR}} \left[1 - \frac{P_{IR}}{1 - (1 - P_{IR})(1 - \pi_L)} \right] \\ &= \frac{n - 1}{1 - P_{IR}} \left[\frac{\pi_L (1 - P_{IR})}{P_{IR} + \pi_L (1 - P_{IR})} \right] \end{aligned}$$

En viktig anledning till varför man vill skaffa sig en god uppfattning om reproduktionstalet R_0 är för att det kan användas för att bestämma hur stor andel som måste vaccineras för att ett framtida utbrott ska undvikas. Genom att man vaccinerar tillräckligt många så att R_0 blir mindre än ett.[3] Men, som vi tidigare har sagt, kan det vara svårt att bestämma R_0 analytiskt. En metod som då istället kan användas för att komma fram till hur man ska göra för att undvika framtida epidemier är att undersöka kritiska gränser. Det kan exempelvis göras genom att man ger alla parametrar utom en, som vi kallar för ρ , ett värde som överensstämmer med verkligheten. Därefter gör man en simulering där man låter ρ variera medan de andra parametrarna är konstanta. Från simuleringsresultatet undersöker man då för vilka värden på ρ som det finns en risk för att en epidemi uppstår. Det innebär alltså att om man kan se till att ρ är lägre än dessa värden så kan man undvika att ett stort utbrott sker. Däremot när det gäller vår modell har vi problem med att bestämma skattningar för parametrarna smittsannolikhet inom lag π_L och smittsannolikhet mellan lag π_G . Vilket leder till att vi inte ska dra allt för stora slutsatser kring specifika värden. Utan istället fokusera mer på att undersöka hur en ökning respektive en minskning påverkar slutstorleken.

2.5 Sammanfattning av modellen

Vi ska nu sammanfatta vår modellbeskrivning genom att ta upp de viktigast antagandena. Innan vi går vidare med diverse simuleringar och resultat.

- Population: Populationen som vi undersöker består av 420 individer som vi antar är homogena, det vill säga lika mottagliga för att bli smittade och lika benägna att sprida smittan vidare. Det gäller även att dessa individer är uppdelade i 14 stycken lag bestående av 30 individer vardera. Dessutom har vi antagit att populationsstorleken är konstant och att ingen person kan skadas eller smittas av någon annan sjukdom som håller individen borta från laget.
- Kontaktstruktur: Vi har en kontaktstruktur på två nivåer, en lokal nivå och en global nivå. Dagligen har en given individ kontakt med en lagkamrat med samma sannolikhet för alla individer inom samma lag oberoende. Medan en given individ från lag A endast kan ha kontakt med individer från lag B om lag A och lag B spelar match mot varandra. Där vi antagit att alla lagen spelar match var tredje dag. Dessutom kan ingen individ i populationen smittas av den intressanta smittan av någon utanför populationen.

- Övergångsfördelningar: Modellen definieras i diskret tid, där ett tidsteg representerar en dag. Sedan har vi i modellen antagit fyra stycken övergångsfördelningar som beskriver antalet individer som flyttas mellan de givna tillstånden för varje lag i vid tid t . Vi har då kommit fram till att dessa är binomialfördelade med parametrar som beror på individernas fördelning över tillstånden samt några givna sannolikheter. Det här resultaten är särskilt användbara nu när vi ska genomföra olika simuleringar.

3. Simulering

Simuleringarna görs i programmet R och koden är skriven i språket R, dessutom har inga programpaket används.[9] Vi har även delat in koden i två delar eftersom vi är intresserade av att undersöka två stycken olika spelscheman, det vill säga ett verkligt och ett slumpmässigt. Skillnaden mellan de två koderna är alltså att koden för de slumpmässiga schemat slumpar fram motståndare för respektive lag, under matchdag, med samma sannolikhet för varje matchkombination. Medan vi för det verkliga schemat har angett varje lags motståndare för respektive matchdag manuellt, i koden, med utgångspunkt från spelschemat för SHL under säsongen 2016/2017.[8]

I simuleringsprocessen har vi valt att representera varje lag som en matris bestående av 30 rader och fyra kolumner. Där varje rad representerar en individ och kolumnerna förklarar vilket tillstånd som individen befinner sig i genom att alla element är lika med ett eller noll beroende på om individen befinner sig i tillståndet eller inte (se Figur A1 i appendix). Det gäller sedan att modellen har två matchfria dagar följt av en matchdag tills alla lagen har spelat 52 matcher vardera. Vi har därför skapat en funktion som används för att undersöka om individerna byter tillstånd, under en matchfri dag, genom att för varje individ slumpa ett likformigt fördelat slumpvariabel $X \sim U(0, 1)$. Sedan om utfallet är mindre än $P_{SE}^{(t)}$, P_{EI} eller P_{IR} beroende på vilket tillstånd som individen befinner sig i byter individen tillstånd. Det gäller då att P_{EI} och P_{IR} är konstanta under hela processen medan $P_{SE}^{(t)}$ beror på hur många individer som är smittsamma i den givna individens lag.

På samma sätt finns det också en funktion som flyttar individer mellan tillstånd under matchdagar. Skillnaden är bara att $P_{SEM}^{(t)}$ i det här fallet, till skillnad från $P_{SE}^{(t)}$, också beror på antalet individer som är smittsamma i den givna individens motståndarlag. Alltså har vi två funktioner som kan flytta individer mellan tillstånd beroende på om det är matchdag eller en matchfri dag. Vilket innebär att vi ursprungligen utgår från en kod, för respektive spelschema, som används för att undersöka hur många individer som har smittas efter 156 dagar. I denna kod kan

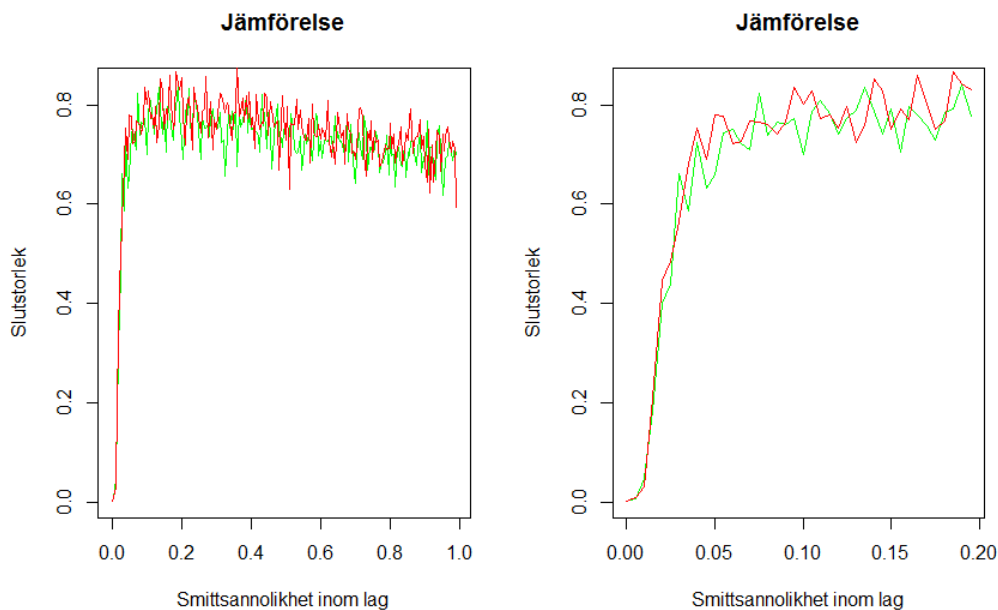
sedan de fyra variablerna π_L , π_G , P_{EI} och P_{IR} variera beroende på vad som är av intresse att undersöka. Där det största problemet är att bestämma ett värde för de två olika smittsannolikheterna som överensstämmer med verkligheten.

Man skulle alltså kunna säga att koden är uppbyggd efter den teoretiska modellformuleringen. Däremot när man programmerar och arbetar med relativt stora koder är det också viktigt att tänka på tidskomplexiteten genom att göra effektiva algoritmer. En förbättring som vi skulle kunnat gjort för att spara tid är att utnyttja alla individer är lika, det vill säga lika benägna att smittas och att föra smittan vidare. Eftersom det betyder att vi istället skulle kunnat respresentera lagen som vektorer bestående av fyra element, ett för respektive tillstånd. Där antalet som befinner sig i vardera tillstånd beskrivs av vektorns element. Anledning till att detta skulle gjort att simuleringen gått fortare är för att antalet beräkningar skulle minskat markant. Däremot anser vi ändå att beräkningarna med de nuvarande programmen går relativt fort, vilket är anledningen till att denna förändring aldrig blev av.

4. Resultat och slutsats

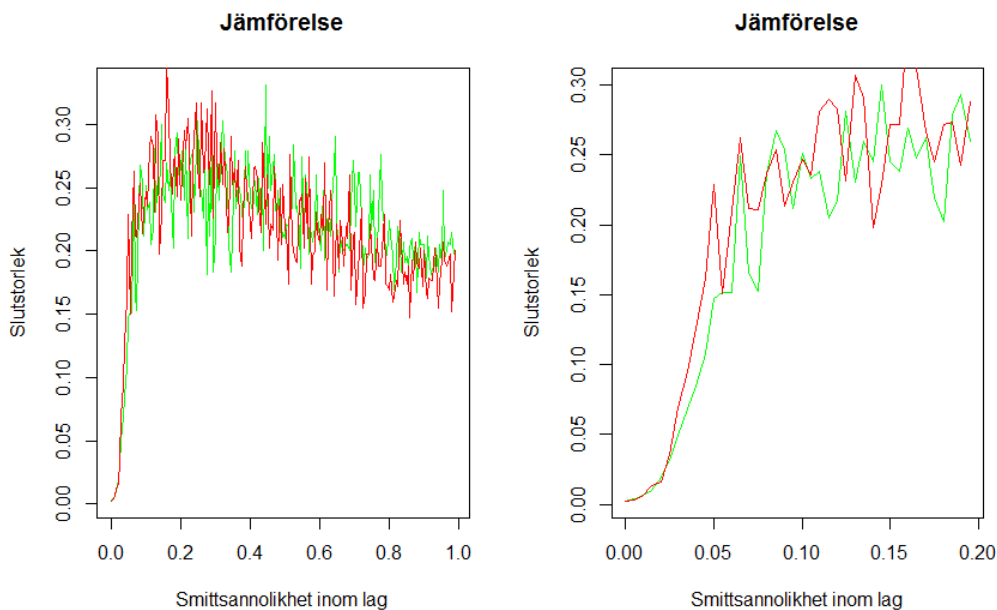
Graferna i detta avsnitt illustrerar bland annat hur medelvärdet av slutstorlekarna varierar när parametrarna smittsannolikhet inom lag π_L respektive smittsannolikhet mellan lag π_G förändras. Samt visar några av dem hur medelvärdet av antal smittade beror på tiden. Där det för samtliga grafer gäller att de gröna representerar det slumpmässiga schemat medan de röda representerar det verkliga schemat.

Vi börjar med att undersöka hur slutstorleken förändras, i vår modell, när smittsannolikheten inom lag ökar. Det första som då måste bestämmas är vilka värden som de tre övriga parametrarna ska fixeras till. För den första simuleringen sätter vi då $P_{EI} = 1$, $P_{IR} = 0.2$ och $\pi_G = 0.002$.



Figur 2: Illustrerar medelvärdet av slutstorleken för 30 simuleringar för varje parameteruppsättning. När $P_{EI} = 1$, $P_{IR} = 0.2$, $\pi_G = 0.002$ och π_L ökar från noll till ett med steglängd 0.005.

I Figur 2 kan vi då se att slutstorleken inte ser ut att bero på vilket spelschema som används. Dessutom verkar det som att slutstorleken växer fort tills den når ett gränsvärde. Det är därför av intresse att göra ytterligare en simulering där smittsannolikheten inom lag varierar. Skillnad är att vi sätter P_{IR} lika med 0.5, vilket betyder att vi ändrar de förväntade antal dagar som en person är smittsam från fem dagar till två dagar.

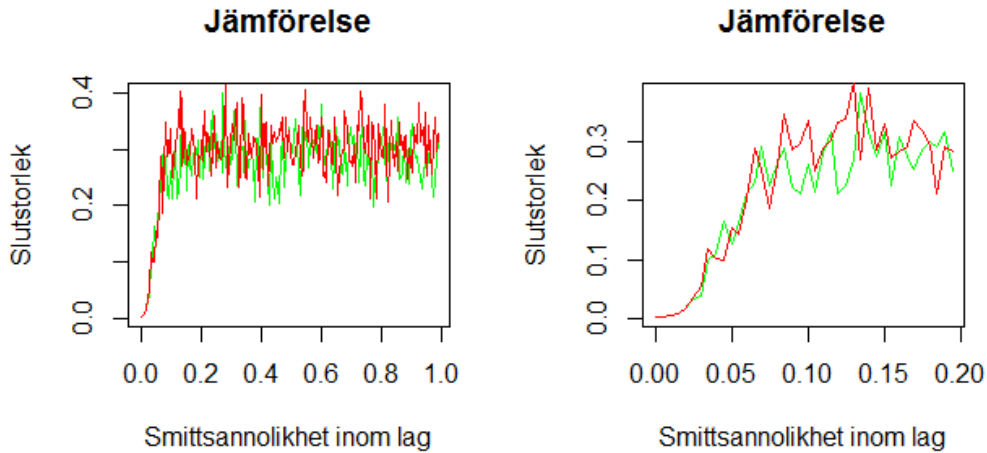


Figur 3: Illustrerar medelvärden av slutstorleken för 30 simuleringar för varje parameteruppsättning. När $P_{EI} = 1$, $P_{IR} = 0.5$, $\pi_G = 0.002$ och π_L ökar från noll till ett med steglängd 0.005.

I Figur 3 kan vi då, precis som tidigare i Figur 2, se att slutstorleken inte beror på vilket spelschema som används. Däremot verkar det som att slutstorleken faktiskt minskar när smittsannolikheten inom lag ökar. Vilket också bekräftas när man gör en liknande simulering med smittsannolikheten mellan lag lika med 0.01 (Se Figur A2 i appendix). En möjlig anledning till varför detta inträffar är för att i dessa simuleringar som illustreras i Figur 3 och Figur A2 är de förväntade antal dagar som en individ är smittsam lika med två. Samtidigt som inkubationstiden är en dag, vilket då resulterar i att ett lag faktiskt endast kan bestå av individer som befinner sig i tillstånd R redan tre till fyra dagar efter att den första individen i laget har smittats förutsatt att smittsannolikheten inom lag är hög. Därav har det infekterade laget, i princip, endast en match på sig att sprida smittan vidare till något annat lag. Så när smittsannolikheten mellan lag är relativt låg leder det till att slutstorleken minskar när smittsannolikheten inom lag ökar.

Det gäller därför för parametern smittsannolikhet mellan lag, när man jämför Figur 3 med Figur A2, att den har en stor inverkan på hur stor smittspridningen blir. Men för att få en bättre förståelse för modellen vill vi också undersöka hur den förväntade tiden i inkubationstillståndet E påverkar slutstorleken. Vi börjar därför med att sätta $P_{EI} = 0.2$, $P_{IR} = 0.2$ och $\pi_G = 0.002$ vilket innebär att

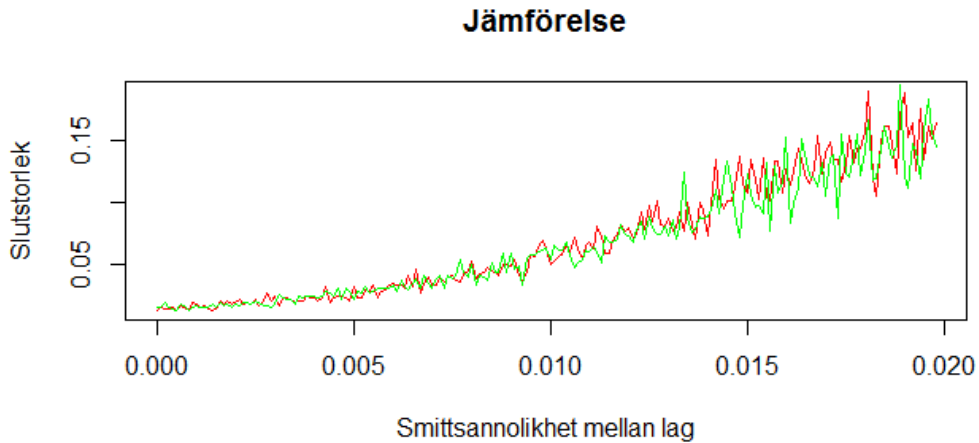
de förväntade antal dagar som en individ befinner sig i tillstånd E är lika med fem. Vi får då ett resultat som överensstämmer ganska bra med resultatet som presenteras i Figur 2 (se Figur A3 i appendix). Vilket tyder på att inkubationstiden inte har särskilt stor inverkan på slutstorleken när smittsamhetstiden är relativt lång. Därav är det också av intresse att göra ytterligare en simulering där vi minskar den förväntade smittsamhetstiden till två dagar och med samma värde på de övriga parametererna som simuleringen innan.



Figur 4: Illustrerar medelvärdet av slutstorleken för 30 simuleringar för varje parameteruppsättning. När $P_{EI} = 0.2$, $P_{IR} = 0.5$, $\pi_G = 0.002$ och π_L ökar från noll till ett med steglängd 0.005.

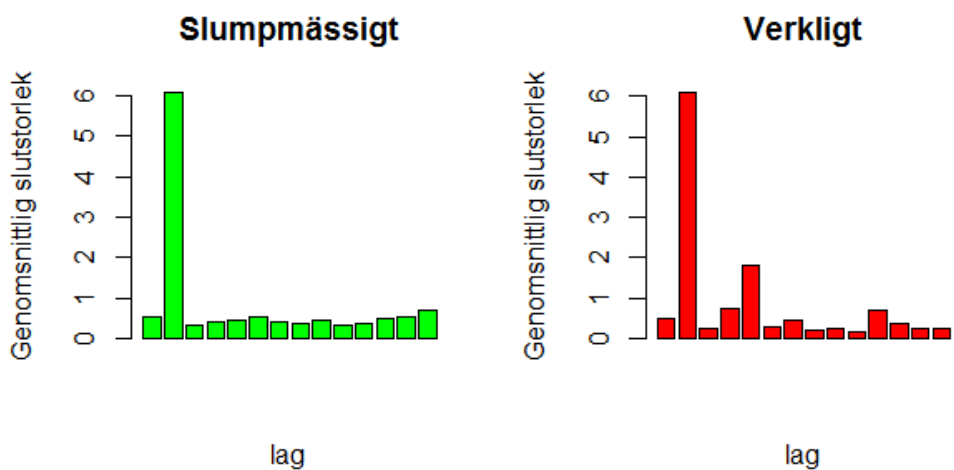
Det vi då kan se i Figur 4 är först och främst att slutstorlekarna för både det verkliga och det slumpmässiga schemat verkar konvergera mot ett gränsvärde. Vilket innebär att slutstorleken inte minskar när smittsannolikheten inom lag ökar. Det tyder på att vårt tidigare resonemang gällande resultaten som vi såg i Figur 3 och Figur A2 verkar stämma. Men utöver detta kan det också vara av intresse att förtydliga vad vi menar med kritiska gränser, som vi nämnde i avsnitt 2.4. Däremot är det i simuleringresultaten ovan svårt att se några tydliga gränser eftersom slutstorleken växer fort i början av alla plottar. Men när man studerar den högra bilden i Figur 4 ser det ut som att R_0 är mindre än ett när smittsannolikheten inom lag är mindre än ungefär 0.02. Dock måste vi också ha i åtanke att teorin kring exempelvis R_0 bygger på asymptotiska resultat. Medan vi i detta fall endast har tittat på medelvärdet av 30 simuleringar för varje parameteruppsättning.

Vidare vill vi nu istället undersöka ett simuleringsresultat när smittsannolikheten mellan lag varierar. Vi låter då $P_{EI} = 1$, $P_{IR} = 0.5$ och $\pi_L = 0.02$. Därefter kommer smittsannolikheten mellan lag variera mellan 0 och 0.02 eftersom vi antagit π_L är större än π_G och resultatet illustreras i Figur 5 nedan.



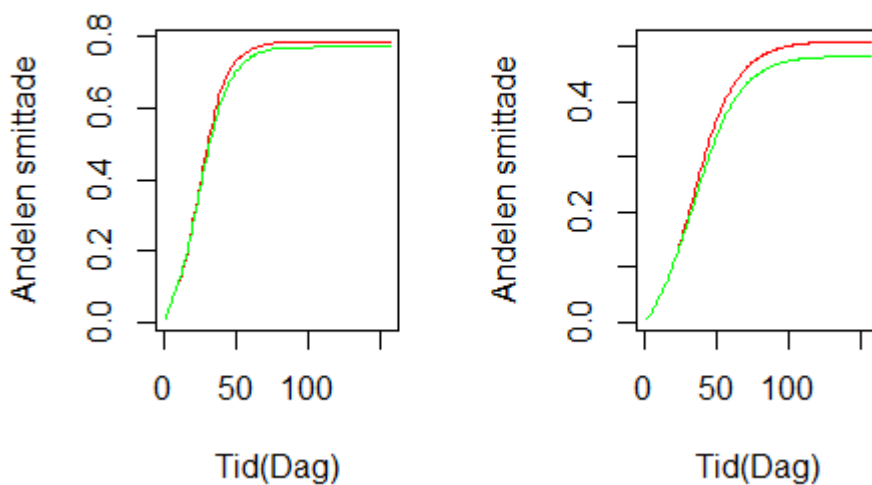
Figur 5: Illustrerar medelvärdet av slutstorleken för 150 simuleringar för varje parameteruppsättning. När $P_{EI} = 1$, $P_{IR} = 0.5$, $\pi_L = 0.02$ och π_G ökar från noll till 0.02 med steglängd 0.0001. Dessutom har vi har $R_0^L = 1.137255$.

En anledning till varför dessa värden valdes är för att det är av intresse att undersöka en kritisk gräns för smittsannolikheten mellan lag. Alltså om vi utgår från en smitta som betar sig som ovan verkar det som, om man utgår från Figur 5, att smittspridningen näst intill isoleras till ett lag om smittsannolikheten mellan lag är mindre än 0.005. Vilket också bekräftas i stapeldiagrammet nedan. Som visar fördelningen över antalet individer som har smittats i respektive lag när $P_{EI} = 1$, $P_{IR} = 0.5$, $\pi_L = 0.02$ och $\pi_G = 0.005$.



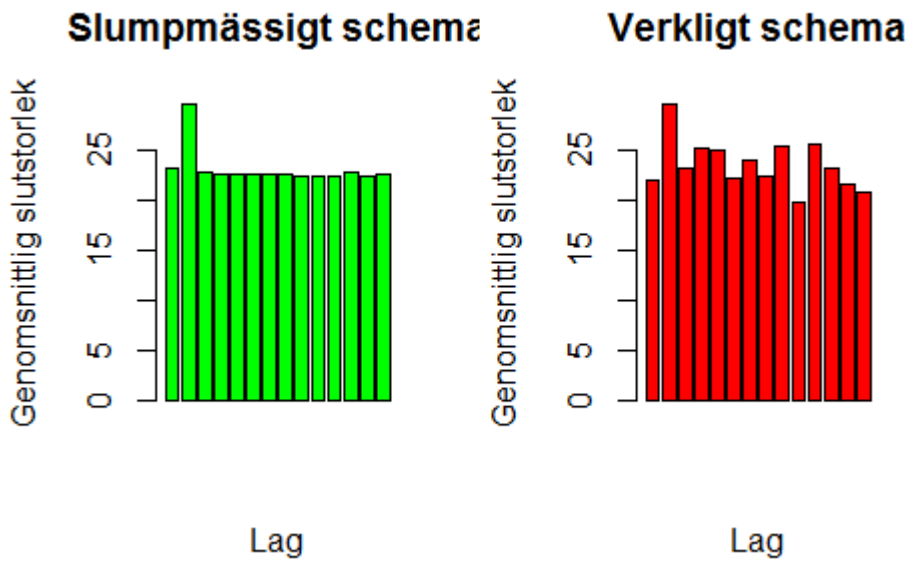
Figur 6: Två stapeldiagram som visar fördelningen över hur många som smittas i varje lag för respektive spelschema. När simuleringen har upprepats 1000 gånger för varje tidsteg, samt var $\pi_L = 0.02$, $\pi_G = 0.005$, $P_{IR} = 0.5$ och $P_{EI} = 1$. Dessutom har vi att $R_0^L = 1.137255$.

Avslutningsvis ska vi också se hur antalet smittade för respektive spelschema beror på tiden. Vi väljer då att titta närmare på två olika parameteruppsättningar. För den ena är $P_{EI} = 1$, $P_{IR} = 0.2$, $\pi_L = 0.1$ och $\pi_G = 0.002$ och för den andra är $P_{EI} = 0.5$, $P_{IR} = 0.5$, $\pi_L = 0.05$ och $\pi_G = 0.005$.

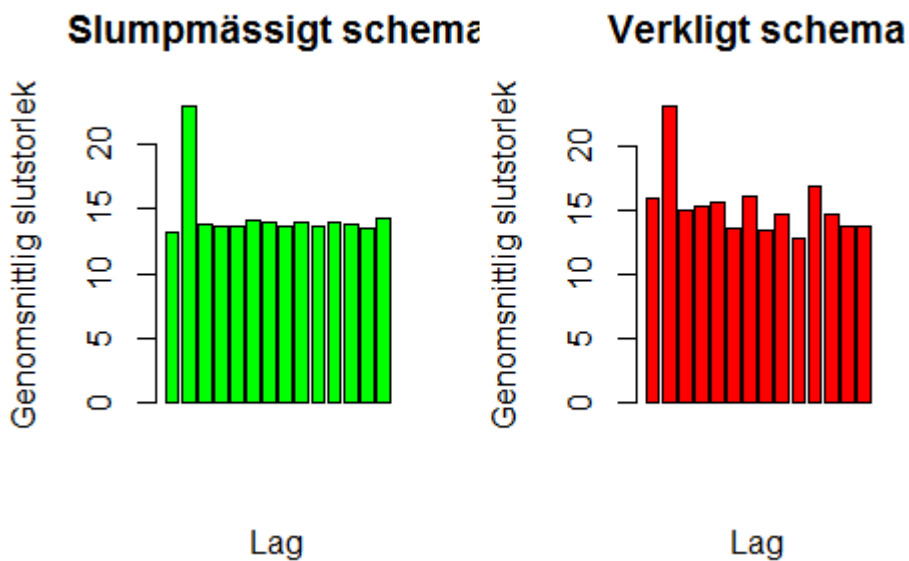


Figur 7: Visar två stycken plottar över tid, där de gröna kurvorna representerar det slumpmässiga schema och de röda kurvorna representerar det verkliga schemat. Sedan för den vänstra bilden är $\pi_L = 0.1$, $\pi_G = 0.002$, $P_{IR} = 0.2$ och $P_{EI} = 1$ samtidigt som värdet på dessa parametrar i den högra bilden är $\pi_L = 0.05$, $\pi_G = 0.005$, $P_{IR} = 0.5$ och $P_{EI} = 0.5$. Dessutom har vi att $R_0^L = 3.295455$ för den vänstra bilden medan $R_0^L = 2.761905$ för den högra bilden.

Från Figur 7 ser vi då att det inte heller verkar finnas någon signifikant skillnad mellan hur snabbt individerna i populationen smittas för respektive spelschema. Däremot när vi istället tittar på stapeldiagrammen i Figur 8 och Figur 9 nedan, som illustrerar medelvärdet över antal individer som blivit smittade för respektive lag.



Figur 8: Två stapeldiagram som visar fördelningen över hur många som smittas i varje lag för respektive spelschema. När simuleringen har upprepats 1000 gånger för varje tidsteg, samt var $\pi_L = 0.1$, $\pi_G = 0.002$, $P_{IR} = 0.2$ och $P_{EI} = 1$. Dessutom har vi har $R_0^L = 3.295455$.



Figur 9: Två stapeldiagram som visar fördelningen över hur många som smittas i varje lag för respektive spelschema. När simuleringen har upprepats 1000 gånger för varje tidsteg, samt var $\pi_L = 0.05$, $\pi_G = 0.005$, $P_{IR} = 0.5$ och $P_{EI} = 0.5$. Dessutom har vi har $R_0^L = 2.761905$.

Då kan vi se att det verkar finns en skillnad mellan det slumpmässiga och det verkliga spelschemat. Eftersom i stapeldiagrammen för det slumpmässiga schemat är antalet smittade i respektive lag ungefär lika utom i laget där smittan började. Medan det för stapeldiagrammen för det verkliga schemat är en större spridning mellan lagen. Anledningen till de resultaten beror troligtvis på att i det slumpmässiga schemat är spelordningen slumpmässig för varje simulering vilket gör att staplarna konvergerar mot samma resultat enligt stora talens lag.[1] Samtidigt som spelschemat är lika för varje simulering när man utgår från det verkliga schemat vilket exempelvis kan leda till samma lag möter laget där smittspridningen inleds i ett skede när många individer i det laget är smittsamma. Däremot hade troligtvis det här resultatet blivit annorlunda om vi hade slumpat laget där smittspridningen skulle inledas innan varje simulering. Istället för att vi slumpade ett lag och därefter valde att smittan skulle börja där för varje simulering som gjordes.

5. Diskussion

I föregående avsnitt kom vi fram till några resultat utifrån diverse simuleringar. Vidare är det nu av intresse att, utifrån dessa resultat, analysera eventuella konsekvenser samt diskutera vilka slutsatser vi egentligen kan dra med tanke på modellens utformning.

När det gäller själva huvudfrågeställningen om smittspridningen är beroende av vilket spelschema som används. Visade det sig att slutstorleken respektive hur snabbt smittan spred sig, i princip, var lika för de två olika spelscheman som undersöktes. Därav anser vi att det inte finns någon särskild anledning att föra diskussionen kring det vidare. Däremot vill vi nämna att det kanske skulle blivit en skillnad om det verkliga schemat hade varit mer realistiskt. Det vill säga olika tidsintervall mellan matcherna samt att alla lagen inte alltid spelar match samma dag.

Men utöver detta ville vi också undersöka hur slutstorleken påverkas när man ändrar värdet på de olika parameterarna. Vi kunde då se att slutstorleken verkar bero på själva parameteruppsättningen och inte enbart värdet på varje parameter var för sig, möjligtvis kan det vara så att smittsannolikheten mellan lag bara påverkar slutstorleken och inte utseendet på graferna ovan. Anledningen till det är troligtvis vår modells struktur där lagen spelar match var tredje dag. För exempelvis när det gäller plottarna där smittsannolikheten inom lag varierade visade det sig att när den förväntade inkubationstiden och smittsamhetstiden var kort så minskades slutstorleken när smittsannolikheten inom lag ökade. Däremot kunde vi genom att öka den förväntade inkubationstiden få ett resultat där slutstorleken istället växte till en specifik nivå för att sedan vara oförändrad samtidigt som smittsannolikheten inom lag fortsatte att öka.

Det mest intressanta resultatet som dök upp var alltså att under vissa förutsättningar, som vi nämde ovan, visade det sig att slutstorleken minskades samtidigt som smittsannolikheten inom lag ökade. Vilket på sätt och vis faktiskt är ett resultat som talar emot vaccinering. För genom vaccinering kan man bland annat minska smittsannolikheten. Därav om man om man utgår från en situation där en smitta sprids under liknande förutsättningar och där smittsannolikheten är väldigt hög så skulle vaccinering kunna resultera till att slutstorleken bara blir större. Däremot är det viktigt att inte enbart stirra sig blind på resultatet innan man drar sådana slutsatser med tanke på att modellen inte förklarar hela verkligheten. Eftersom om man utgår från en situation där smittsannolikheten inom lag är väldigt hög skulle det kunna leda till att vissa lag ibland består av få individer som kan spela match. Med tanke på att i verkligheten flyttas man inte från smittsam till återhämtad utan man stannar först i ett karantäntillstånd innan man är fullt återställd. Därav skulle eventuellt många matcher inte ens

genomföras på grund av för få spelare. Därför skulle det kanske vara intressant om man lagt till ett sådant tillstånd i modellen och om ett lag då har för många spelare som inte kan spela väljer man att ställa in den matchen. Dock skulle flera inställda matcher vara ett stort problem för ligan eftersom det finns många ekonomiska intressen som styr och därför tror vi att ligan eventuellt skulle bestämma att alla individer i populationen sätts i karantän under en vecka för att se till att smittspridningen upphör. Men utöver det här finns det även andra anledningar till varför man ska låta individerna vaccineras istället för att se till att alla individer smittas fort. Eftersom om man tar påssjukan som ett exempel kan den också leda till följsjukdommar som hjärnhinneinflammation eller bukspottkörtelinflammation.[6]

Det var egentligen inget resultat som är oförklarligt och utom resultatet som vi nyligen diskuterade var inget av resultaten särskilt häpnadsväckande. Därav skulle det vara tänkvärt att göra nya simuleringar där man gör några förändringar i modellen för att göra den lite mer realistisk. En förändring är man inför det där karantäntillståndet samt använder sig av det för att ta hänsyn till hur många individer som inte har tillfrisknat. Eftersom det är rimligt att anta att om individerna i populationen har färre kontakter borde det resultera till en högre smittsannolikhet på grund av att man då tillbringar mer tid med varje person. Däremot om den förändringen ska leda till något användbart resultat är det också viktigt att ha tillgång till realistiska skattningar av de två smittsannolikheterna för den smittan som ska undersökas, vilket inte är särskilt enkelt att beräkna. Dessutom kan resultaten förändras lite om ett subkliniskt tillstånd, det vill säga ett tillstånd för individer som är smittade utan att uppvisa symptom, läggs till. För som vi nämnde tidigare i inledningen gällde det att cirka 50 procent av alla individer som smittas av påssjukan inte uppvisar några symptom men kan ändå sprida smittan vidare.[6] Vilket då skulle innebära att vissa individer som smittas får en längre smittsamhetstid än de som sätts i karantän.

Avslutningsvis vill vi också poängtera att denna typ av modell, där individer har kontakter på en lokal nivå dagligen och en global nivå under vissa specifika dagar, också kan användas för att undersöka en eventuell smittspridning inom andra populationer. En lämplig population skulle exempelvis vara brottare eftersom de tränar i en klubb och under vissa dagar möter de brottare från andra klubbar. Dessutom har man även kunnat påvisa att olika hudsjukdomar kan spridas mellan personer som brottas.[11]

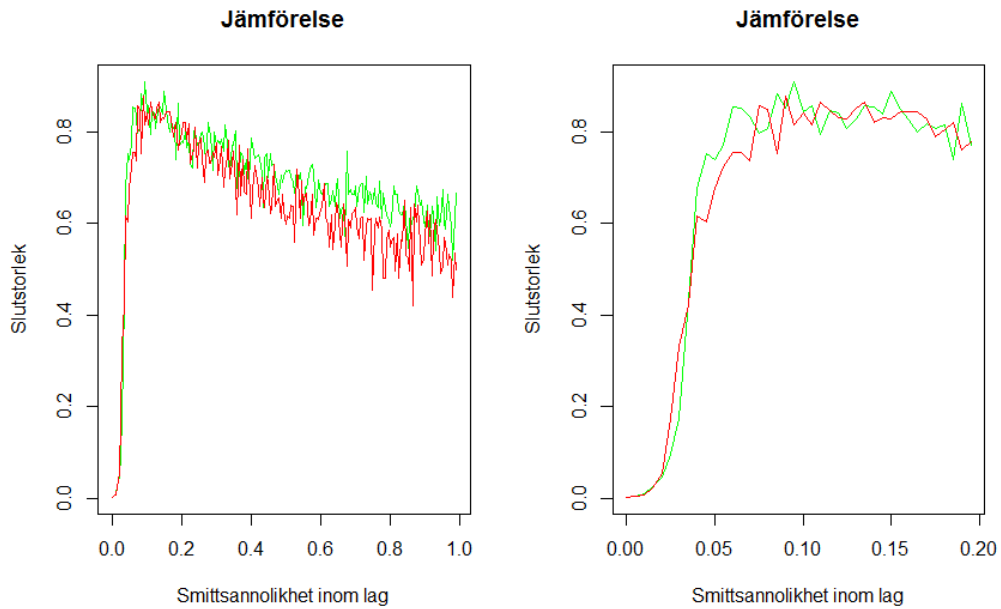
6. Referenser

- [1] S.E.Alm, T.Britton: *Stokastik*, 2014, Liber.
- [2] H.Andersson, T.Britton: *STOCHASTIC EPIDEMIC MODELS AND THEIR STATISTICAL ANALYSIS*, February 2000.
- [3] T.Britton: *Pest kolera och matematik eller Vad kan matematik och statistik lära oss om smittsamma sjukdomars utbredning?*, 2008, NCM.
- [4] T.Britton: *Stochastic epidemic models: a survey*, 2010, Math. Biosci, 225, 24-35.
- [5] S.M.Ross: *Introduction to probability models*, 2014, Academic Press
- [6] 1177 vårdguiden. Senast ändrad 2016-06-15. Påssjuka
<https://www.1177.se/Skane/Fakta-och-rad/Sjukdomar/Passjuka/>
(Hämtad 2017-04-02)
- [7] NHL. 2016. Alla lagen i National Hockey League
<https://www.nhl.com/> (Hämtad 2017-04-02)
- [8] Svenska Ishockeyförbundet. 2017. Official statistics.
<http://stats.swehockey.se> (Hämtad 2017-05-10)
- [9] R Core Team. 2017. The R Project for Statistical Computing
<http://www.R-project.org/> (Hämtad 2017-05-09)
- [10] Macleans. 2014-12-10. Why the NHL is the ultimate breeding ground for mumps
<http://www.macleans.ca/society/health/the-nhl-is-the-ultimate-breeding-ground-for-mumps/>
(Hämtad 2017-03-05)
- [11] Privatmedicin. Hudproblem hos idrottare
<http://www.privatmedicin.se/medicinboken/idrottsmedicin-och-idrottsskador/hudproblem-hos-idrottare/> (Hämtad 2017-05-09)

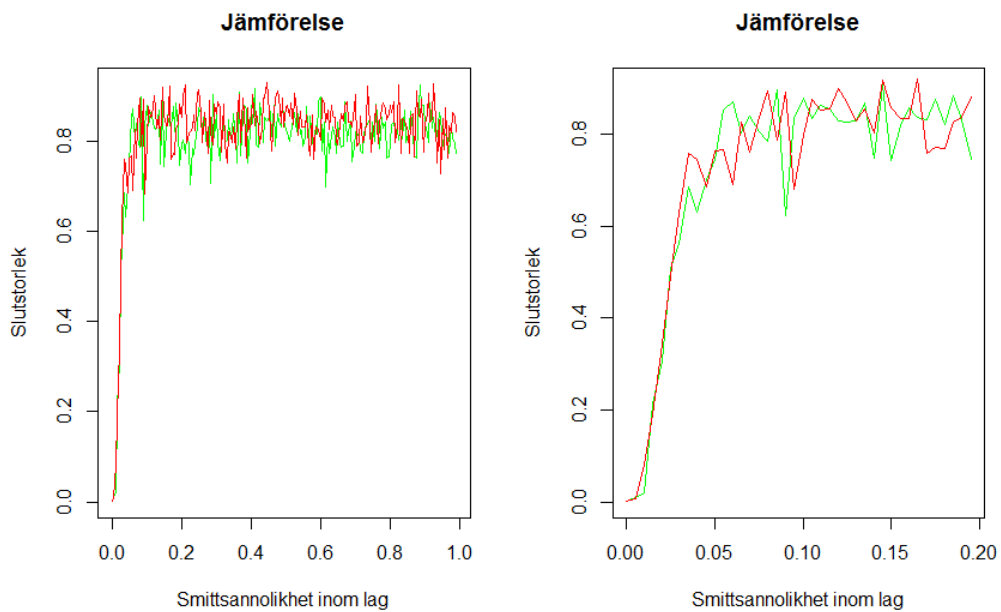
7. Appendix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Figur A1: En matris som förklarar vilket tillstånd som individerna inom ett lag befinner sig i vid tid t , förutsatt att laget bara har sex individer. I det här fallet ser vi då exempelvis att fyra individer befinner sig i tillstånd S vid tid t .



Figur A2: Illustrerar medelvärdet av slutstorleken för 30 simuleringar för varje parameteruppsättning. När $P_{EI} = 1$, $P_{IR} = 0.5$, $\pi_G = 0.01$ och π_L ökar från noll till ett med steglängd 0.005.



Figur A3: Illustrerar medelvärdet av slutstorleken för 30 simuleringar för varje parameteruppsättning. När $P_{EI} = 0.2$, $P_{IR} = 0.2$, $\pi_G = 0.01$ och π_L ökar från noll till ett med steglängd 0.005.