



Stockholms  
universitet

# Simultan spridning av HIV och TBC

Lisa Axelsson

Kandidatuppsats 2017:20  
Matematisk statistik  
Juni 2017

[www.math.su.se](http://www.math.su.se)

Matematisk statistik  
Matematiska institutionen  
Stockholms universitet  
106 91 Stockholm

# Simultan spridning av HIV och TBC

Lisa Axelsson\*

Juni 2017

## Sammanfattning

HIV och tuberkulos (TBC) är två av de främsta dödsorsakerna i världen och det är känt att smittorna har en samverkan. Att bära på både HIV-viruset och tuberkulosbakterien är betydligt värre än att bära på smittorna var för sig. Det här arbetet handlar om att med hjälp av stokastisk epidemimodellering och simuleringar undersöka hur denna samverkan påverkar spridningen av de två smittorna när de sprids simultant i en population i Afrika. Vi låter TBC-smittan finnas i en andel av befolkningen från början och introducerar sedan en HIV-smitta. Vi kommer att behandla sexuella spridningen av HIV-viruset mellan heterosexuella individer, därför låter vi viruset spridas på ett kontaktnätverk mellan individerna. Vi är intresserade av hur nätverkets egenskaper påverkar spridningen av HIV-viruset.

Några analytiska beräkningar görs, men fokus ligger på att undersöka smittspridningens egenskaper genom simuleringar. TBC-spridningen analyserar vi genom dess så kallade endemiska nivåer, andelen av populationen som bär på smittan. Resultatet av HIV-spridningen utvärderas med sannolikheten att den introducerade smittan leder till en epidemi, storleken på denna epidemi och dödstalet.

Det visar sig att när HIV-introduceras i en befolkning där TBC finns så ökar inte antalet individer som bär på tuberkulosbakterien, däremot blir fler sjuka. Ett HIV-virus som introduceras i en befolkning med TBC har både större sannolikhet att leda till en epidemi och leder i genomsnitt till ett större antal HIV-smittade än om populationen hade varit TBC-fri. På vilket sätt dödstalet påverkas är oklart. På vilket sätt kontaktnätverket påverkar spridning av HIV beror på hur lätt smittan sprids i en population i allmänhet; en smitta som är lätt att behandla sprids bättre när antalet kontakter individerna har stor variation medan en svårbehandlad smitta sprids bättre på ett nätverk som är mer homogent.

---

\*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.  
E-post: [liax3385@student.su.se](mailto:liax3385@student.su.se). Handledare: Mia Deijfen och Pieter Trapman.

## **Förord**

Detta arbete utgör en kandidatuppsats i matematisk statistik om 15 högskolepoäng. Jag vill tacka mina handledare Mia Deijfen och Pieter Trapman för handledning och inspiration under arbetets gång.

# Innehållsförteckning

<b>1</b>	<b>Introduktion</b>	<b>4</b>
<b>2</b>	<b>Modellering</b>	<b>4</b>
2.1	Nätverksmodell för HIV-spridning . . . . .	4
2.1.1	Generering av nätverksgraf . . . . .	5
2.2	Modell för HIV-spridning . . . . .	5
2.3	Modell för TBC-spridning . . . . .	6
2.4	Samspel . . . . .	7
<b>3</b>	<b>Intressanta storheter</b>	<b>7</b>
3.1	TBC . . . . .	7
3.1.1	Endemiska nivåer . . . . .	7
3.2	HIV . . . . .	8
3.2.1	Utbrottsstorlek . . . . .	8
3.2.2	$R_0$ . . . . .	8
3.2.3	Utbrottssannolikhet . . . . .	10
3.2.4	Antal döda . . . . .	10
<b>4</b>	<b>Simuleringar &amp; Resultat</b>	<b>10</b>
4.0.1	Tester . . . . .	10
4.1	Val av parametervärden . . . . .	11
4.1.1	Kantfördelningen . . . . .	12
4.2	Simulering av smittoförlopp . . . . .	12
4.3	TBC . . . . .	13
4.4	HIV . . . . .	13
4.4.1	Kantfördelning . . . . .	13
4.5	Simultan spridning . . . . .	16
4.5.1	Påverkan på TBC . . . . .	16
4.5.2	Påverkan på HIV . . . . .	17
4.5.3	$R_0$ . . . . .	17
<b>5</b>	<b>Slutsatser och diskussion</b>	<b>19</b>
<b>6</b>	<b>Referenser</b>	<b>21</b>

# 1 Introduktion

Humant immunbristvirus, HIV, är ett virus som sedan sin upptäckt har orsakat ca. 35 miljoner dödsfall och trots att det årliga antalet dödsfall minskar är det fortfarande en av de ledande dödsorsakerna i världen. Om viruset förblir obehandlat leder det till ett förstört immunförsvar och med stor sannolikhet död till följd av andra infektioner. 36.7 miljoner beräknades leva med infektionen år 2015, majoriteten i östra och subsahariska Afrika. Den vanligaste dödsorsaken hos HIV-smittade är tuberkulos, TBC. TBC är en bakterieinfektion som även den är en av de ledande dödsorsakerna i världen. 2015 orsakade bakterien 1.8 miljoner dödsfall och en tredjedel av världens befolkning beräknas bära på bakterien. Majoriteten av smittspridningen sker i Afrika och sydöstra Asien. De allra flesta som bär på tuberkulosbakterien blir dock aldrig sjuka. Ungefär en tiondel av de smittade utvecklar en aktiv och smittsam infektion under sin livstid, en siffra som är betydligt högre bland HIV-smittade. Koinfektion med HIV gör den också mer allvarlig och dödlig. (WHO, 2016; UNAIDS, 2016)

Målet med den här uppsatsen är att modellera spridningen av HIV och TBC med fokus på Afrika och analysera den med hjälp av framför allt simuleringar. Främst är vi intresserade av hur spridningen av ett HIV-virus som introduceras i en befolkning påverkas av att TBC är närvarande. HIV kan spridas på flera sätt, men vi kommer endast att behandla den sexuella spridningen mellan heterosexuella män och kvinnor. Sexuella kontakter mellan individer representeras med en nätverksgraf på vilken viruset sprids. Vi är utöver huvudproblemet även intresserade av hur fördelningen av antalet sexuella partners påverkar spridningen av HIV.

Eftersom att många förenklingar görs i modelleringen är målet inte att så realistiskt som möjligt spegla smittspridningen, men slutsatserna bör ändå kunna ge aningar om hur smittorna samspelar i verkligheten.

I kapitel 2.1 definieras kontaktnätverket för spridning av HIV och i kapitel 2 definieras modeller för spridning av smittorna. I kapitel 3 går vi igenom vilka storheter vi kommer att använda för att analysera smittspridningarna. Kapitel 4 börjar med en förklaring av simuleringarna och val av parametervärden, sedan presenteras resultaten av dessa simuleringar. I kapitel 5 går vi igenom resultaten, diskuterar vad som kan ha påverkat dessa och vad framtida undersökningar inom området skulle kunna inkludera som inte ingår i det här arbetet.

## 2 Modellering

HIV-viruset kan spridas sexuellt, via blod eller från mor till barn vid födsel och amning. I många delar av världen sprids viruset främst mellan homosexuella män och via sprutdelning. Även i Afrika är dessa grupper överrepresenterade men den sexuella spridningen mellan heterosexuella individer beräknas ändå stå för majoriteten av spridningen (Anderson & May, 1991, kap. 11.3.9), därför är det den vi väljer att undersöka. Det gör också problemet tillräckligt enkelt för att kunna dra vissa analytiska slutsatser. Vi får dessutom möjlighet att använda olika smittorisker för könen, något som är relevant eftersom att infektionsrisken är dubbelt så stor från man till kvinna jämfört med det motsatta. Att vi studerar den heterosexuella spridningen innebär att alla inte kan smitta alla, därför är det passande att sprida viruset på ett bipartit kontaktnätverk.

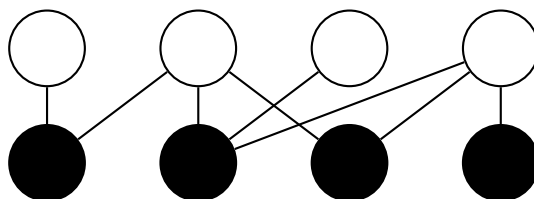
Tuberkulos sprids genom luften, till exempel genom att en smittsam person hostar. En rimlig förenkling är då att alla kan smitta alla. Vi behöver alltså inte skapa något nätverk för tuberkulosspridning. Det tillåter oss att enkelt göra analytiska beräkningar.

Vi kommer att använda oss av stokastiska modeller för smittspridning. Vi vill bland annat undersöka sannolikheten att en introducerad HIV-smitta leder till ett utbrott och fördelningarna för antalet smittade och antalet döda i epidemier, något som endast är möjligt för stokastiska modeller. Stokastiska modeller är också passande att analysera med hjälp av simuleringar. Vi kommer att låta spridningen ske i diskret tid. I det här avsnittet definieras nätverksmodellen för spridning av HIV samt de två modellerna för spridning av HIV och TBC och hur de påverkar varandra.

### 2.1 Nätverksmodell för HIV-spridning

En nätverksgraf består av ett antal noder samt kanter som sammanbinder noderna. I vårt fall representerar noder individer i populationen och kanter sexuella kontakter mellan dessa. Eftersom att spridningen som studeras är mellan heterosexuella individer kommer vi att använda oss av en

bipartit graf. En bipartit graf är en graf där noderna delas in i två typer (här är det män och kvinnor) och där kanter endast kan sammanbinda två noder som har olika typ. Under simuleringarna kommer smittan alltså endast att kunna spridas via grafens kanter. För en given individ kallar vi alla individer som har en gemensam kant med denna för individens kontakter. Varje individ tilldelas ett antal kontakter utefter en bestämd fördelning, nodernas gradfördelning. Vem som paras ihop med vem bestäms uniformt slumpmässigt. Vi kommer framför allt att välja fördelningar som ger många noder få kanter medan ett fåtal noder får betydligt fler kanter, så kallade tungsvansade fördelningar.



Figur 1: Exempel på en bipartit graf

### 2.1.1 Generering av nätverksgraf

Grafen för spridning av HIV genereras på följande sätt, liknande Newman (2002):

1. Bestäm populationsstorlek och representera varje individ med en nod.
2. Dela populationen i två delar: män och kvinnor. Vi väljer att dela populationen i två lika stora delar.
3. Tilldela varje individ (nod) ett antal halvkanter utifrån en bestämd fördelning. En halvkant är en kant som endast är sammankopplad med en nod.
4. Välj slumpmässigt en kvinnlig halvkant och en manlig halvkant och ersätt dem med en hel kant mellan de två individerna. Fortsätt tills ena sidan har slut på halvkanter.
5. Släng bort kanter som blir över. Populationen är tillräckligt stor för att fördelningen av antal kanter per individ ska bli approximativt oförändrad trots att ett fåtal kanter försvinner.

Av Newman (2002) beskrivs en liknande generering av ett nätverk, men i en homogen population där exakta analytiska resultat är av intresse, därför är hanteringen av överblivna kanter där striktare. Vi kommer inte att vara lika intresserade av exakta analytiska resultat för HIV-spridning på nätverket och nöjer oss därför med att gradfördelningen är approximativ.

## 2.2 Modell för HIV-spridning

För att kunna välja en lämplig modell för spridningen tittar vi på sjukdomsförloppets egenskaper. När en individ blir smittad med viruset är det inte alltid sjukdomen märks, men blir trots det tidigt smittsam. När viruset har upptäckts finns behandling att få. Behandlingen gör inte att de smittade kan bli helt friska, men dagens behandling är så pass effektiv att de smittade kan leva ett normalt liv med en smittsamhet som är så liten att den kan ses som försumbar. Om behandling inte fås kommer HIV så småningom leda till att kroppen inte längre kan försvara sig från infektioner. Det är det här stadiet som definierar AIDS, som med stor sannolikhet leder till döden.

En smittspridning av den här typen kan förenklas till en så kallad SIR-modell för smittspridning. SIR står för "susceptible", "infectious" och "removed", de tre stadierna en individ kan befinna sig i. S, I och R kommer i fortsättningen att representera andelarna av populationen som befinner sig i de olika grupperna. Vi kommer även att tillåta död, därför inför vi även en grupp D. Ur HIV-spridningens perspektiv är det likvärdigt med att bli behandlad men eftersom att vi kommer att sprida två smittor simultant är det relevant att skilja på grupperna, en individ med behandlad HIV kan fortfarande ingå i spridningen av tuberkulos.

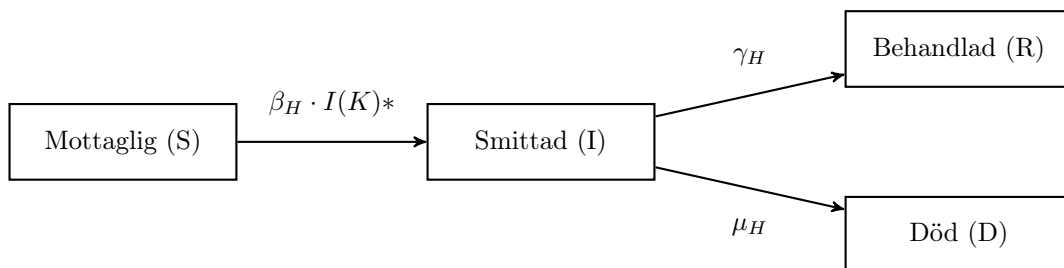
S: Mottaglig (Susceptible)  
 I: Smittsam (Infectious)  
 R: Behandlad (Removed)  
 D: Död

I vår smittspridning kommer alla individer initialt att vara mottagliga utom en slumpmässigt vald person som är smittsam. I varje tidssteg kan individerna flyttas mellan grupperna med bestämda sannolikheter per tidssteg:

$\beta_{HK}$ : Sannolikheten att en given mottaglig kvinna blir smittad av en given smittsam kontakt  
 $\beta_{HM}$ : Sannolikheten att en given mottaglig man blir smittad av en given smittsam kontakt  
 $\gamma_H$ : Sannolikheten att en given smittsam individ får behandling  
 $\mu_H$ : Sannolikheten att en given smittsam individ dör i AIDS

Figur 2 visar de olika stadierna och övergångssannolikheterna.

I en spridningsmodell där ingen kan återgå till att bli mottaglig och där inga nya individer tillkommer i populationen kommer smittan så småningom att försvinna. Nivåerna börjar minska när andelen mottagliga är få jämfört med antalet smittsamma och när ingen smittsam individ finns kvar i populationen är smittan utdöd. Vi kommer alltså att undersöka ett sjukdomsförlopp under en begränsad tid. Att låta populationen vara sluten bortsett från dödsfallen är en grov förenkling av verkligheten som diskuteras i kapitel 5, det är dock en förenkling som krävs för att kontaktnätverket ska bli enkelt hanterbart.



\*  $I(K)$  är antal smittsamma kontakter.

Figur 2: Spridningsmodell för HIV

### 2.3 Modell för TBC-spridning

När en individ blir smittad med TBC blir denna varken sjuk eller smittad på en gång utan har bakterien vilande i kroppen. De flesta utvecklar aldrig en aktiv sjukdom. Om bakterien däremot blir aktiv, ofta till följd av ett försvagat immunförsvar, blir individen både sjuk och smittsam. Behandling finns. Den är effektiv och gör att individen blir helt frisk igen, däremot ger den inte immunitet mot att bli infekterad igen. Utan behandling är tuberkulos dödlig.

En modell som kan beskriva ett sjukdomsförlopp av den här typen är en SEIS-modell. I modellen finns tre stadier, plus i vårt fall död. Även här kommer S, E, I och D att representera andelar av populationen.

S: Mottaglig (Susceptible)  
 E: Latent (Exposed)  
 I: Smittsam (Infectious)  
 D: Död

Parametrarna som bestämmer sannolikheterna för en individ att övergå från en grupp till en annan i varje tidssteg är:

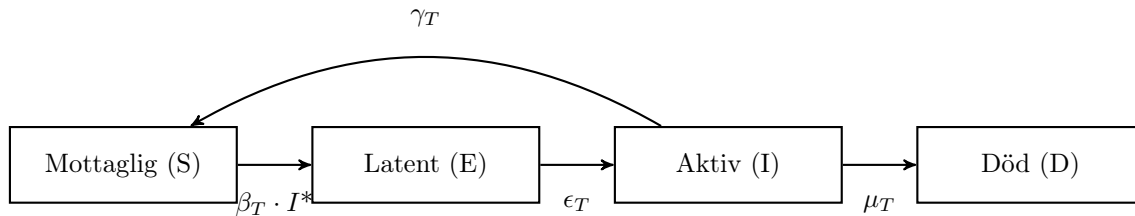
$\beta_T/n_t$ : Sannolikheten att bli smittad av en given smittsam individ  
 $\epsilon_T$ : Sannolikheten att övergå från latent till aktiv smitta  
 $\gamma_T$ : Sannolikheten att få behandling



$\mu_T$ : Sannolikheten att dö av TBC

Här skalas sannolikheten att bli smittad med  $n_t$ , antalet levande individer i tidssteget. Den totala sannolikheten att bli smittad beror alltså på andelen smittsamma individer i populationen. Figur 3 visar de olika stadierna och övergångssannolikheterna.

En smittspridning som följer en SEIS-modell är i vanliga fall endemisk, det innebär att sjukdomen alltid finns i en viss andel av befolkningen och har mycket liten sannolikhet att försvinna om den väl får fäste. Andelen kallas smittans endemiska nivå. Då kan det tyckas orimligt att tillåta död i en annars sluten population eftersom att populationen då bara kan minska, men aldrig öka. Anledningen till att vi ändå vill tillåta död i modellen är att vi har ett begränsat tidsperspektiv (tiden då HIV-utbrottet pågår), då kommer det inte att vara något stort problem så länge dödligheten inte väljs som alltför hög i relation till sannolikheten att få behandling. Att vi väljer att skala smittosannolikheten mot befolkningsstorleken gör också att de endemiska nivåerna bibehålls trots att befolkningen minskar. Detta känns också intuitivt rimligt; sannolikheten att bli smittad av en given smittsam individ bör vara större i en liten by på 20 personer än i en stad med en miljon invånare.



\* I=andelen smittsamma individer i befolkningen

Figur 3: Spridningsmodell för TBC

## 2.4 Samspel

Det här arbetet handlar till stor del om att undersöka samspelet mellan HIV och tuberkulos. I modellen finns samspelet på två ställen:

1. Risken att latent TBC blir aktiv ( $\epsilon_T$ ) ökar för individer med HIV.
2. Risken att dö i HIV ( $\mu_H$ ) ökar för individer med aktiv TBC

Detta är de två tydliga samband som är kända mellan HIV och TBC. Det kan diskuteras om det är risken att dö av HIV som ökar med TBC eller tvärtom, men eftersom att vi främst är intresserade av att studera TBCs påverkan på HIV är detta den mest passande modellen. Även i verkligheten klassificeras personer som dör med koinfektion som att ha dött av HIV. (WHO, 2016)

Det kan även finnas andra faktorer som indirekt påverkar, men ingenting sådant är tydligt nog för att inkludera i modellen.

## 3 Intressanta storheter

När vi studerar resultatet av smittspridning finns det ett antal intressanta storheter att undersöka, dessa presenteras nedan.

### 3.1 TBC

#### 3.1.1 Endemiska nivåer

De endemiska nivåerna i en stokastisk endemi i en stor population kan approximeras med en deterministisk modell i kontinuerlig tid. I den deterministiska modellen antar vi en stor och sluten

population. Eftersom att död ingår i vår modell är vår population inte sluten, men eftersom att vi skalar smittorisken mot antalet levande individer är övergångssannolikheterna oberoende av populationsstorleken och dödstal kan ignoreras i beräkningen av endemiska nivåer så länge populationen fortfarande är stor.

I stället för att individer går mellan infektionsgrupper med sannolikheter räknar vi i den deterministiska, kontinuerliga modellen med att andelar av populationen går mellan grupper med intensiteter:

$\beta$ : Intensiteten med vilken en given mottaglig individ blir smittad av en given smittsam individ.

$\epsilon$ : Intensiteten med vilken smittan går från latent till smittsam.

$\gamma$ : Behandlingsintensitet

Smittspridningen brukar definieras med ett ekvationssystem som bygger på intensiteterna med vilka individer flyttar mellan grupperna. Lösningen då  $T \rightarrow \infty$  ger de endemiska nivåerna för S, E och I:

$$\begin{aligned}\frac{dS}{dT} &= -\beta SI + \gamma I \\ \frac{dE}{dT} &= \beta SI - \epsilon E \\ \frac{dI}{dT} &= \epsilon E - \gamma I\end{aligned}$$

Andelen mottagliga individer (S) minskar då mottagliga blir smittade och ökar när smittsamma får behandling. Andelen latent individer (E) ökar då mottagliga blir smittsamma och minskar då latent smitta utvecklas till aktiv smitta. Andelen smittsamma individer (I) ökar då latent smitta utvecklas till aktiv och minskar då smittsamma får behandling. Ekvationerna bygger på den allmänna deterministiska epidemimodellen som beskrivs av Andersson & Britton (2008, kapitel 1.4) och utökas med en latent fas samt låter behandlade individer direkt återgå till att bli smittsamma (Anderson & May, 1991, kapitel 3.4.1 & 3.4.6).

För att få fram de endemiska nivåerna för intressanta parametervärden löser vi ekvationssystemet numeriskt.

## 3.2 HIV

### 3.2.1 Utbrottsstorlek

Utbrottsstorleken definieras som antalet individer som blir smittade innan viruset dör ut. För ett utbrott som följer en SIR-modell i en stor population gäller det enligt "*threshold limit theorem*" (Andersson & Britton, 2008, kapitel 4) att antalet aningen är litet eller normalfördelat kring en deterministisk andel av befolkningen. När endast ett fåtal blir smittade sägs det vara ett litet utbrott, annars kallas det ett stort utbrott eller en epidemi.

För att undersöka utbrottsstorleken gör vi flera simuleringar av HIV-spridning på nätverket och noterar hur många som smittades i varje simulering. Oftast är vi endast intresserade av de fall som ledde till stora utbrott men det kan även vara intressant att undersöka den empiriska sannolikhetsfördelningen som fås av alla simuleringar.

### 3.2.2 $R_0$

Reproduktionstalet  $R_0$  är en central storhet inom epidemiologi. Den brukar för homogena populationer på homogena grafer (grafer där nodernas gradtal har liten variation) definieras som *det förväntade antalet personer en smittsam individ smittar i en fullt mottaglig population*. För smittspridning på grafer där antalet kontakter som i vårt fall har stor variation är inte denna definition tillräcklig. En definition av  $R_0$  som tar hänsyn till heterogeniteten i antal kontakter beskrivs av Olinky & Stone (2004):

$$R_0 = \rho_0(1 + C_V^2) \quad (1)$$

Där  $\rho_0$  reproduktionstalet för smittan om den hade spridits på en homogen graf och  $C_V^2$  är variationskoefficienten (variansen delat med väntevärdet i kvadrat) för gradfördelningen.

Förutom att vår gradfördelning är tungsvansad måste vi ta itu med att vår graf är bipartit med olika smittosannolikhet för män och kvinnor. Vi hanterar problemet liksom Olinky & Stone (2004) och projicerar nätverket på den manliga populationen. Vi är alltså till att börja med intresserade av  $\rho_0 = \text{det förväntade antalet män som smittas av kvinnor som smittats av en smittsam man i en fullt mottaglig population}$ . Det kommer att visa sig att en projektion på den kvinnliga befolkningen ger samma resultat.

För att beräkna  $\rho_0$  behöver vi veta det genomsnittliga antalet män varje man gemensamma kontakter med, vi kallar talet  $E[K_2]$ . Vi behöver också veta sannolikheten att en given smittsam man smittar en given mottaglig man genom en kvinna. Sannolikheten kommer vara produkten av sannolikheterna för de två händelserna i kedjan. Vi betecknar talet  $\sigma$ . Sedan kan vi alltså beräkna

$$\rho_0 = E[K_2]\sigma \quad (2)$$

Antalet män en man har gemensamma kontakter med,  $E[K_2]$  är produkten av det förväntade antalet kontakter en man har ( $E[K]$ ) och det förväntade antalet kontakter en av mannens kontakter har ( $E[\tilde{K}]$ ).  $E[K]$  definieras som en modellparameter och är alltså känd.  $E[K] \neq E[\tilde{K}]$  på grund av så kallad *size bias*. Size bias innebär här att en kvinna som har kontakt med den initialt smittade mannen i genomsnitt har fler kontakter än en slumpmässigt vald kvinna, det beror på att en individ med många kontakter har större sannolikhet att vara en kontakt till den initialt smittade. Om den vanliga gradfördelningen har fördelningsfunktion  $P(K=k)$  har gradfördelningen med size bias fördelningsfunktion

$$P(\tilde{k} = k) = \frac{k \cdot P(K = k)}{E(K)}$$

Vi har alltså att

$$E(K_2) = E(K) \cdot \sum_{i=1}^{\infty} \frac{i^2 \cdot P(K = k)}{E(K)} \quad (3)$$

Väntevärdena beräknar vi numeriskt.

När vi ska beräkna  $\sigma$  ska vi beräkna sannolikheten att smitta överförs från en smittsam till en mottaglig individ innan perioden då överföring är möjlig tar slut. Det som kan leda till att infektionen inte hinner ske är (sannolikheter per tidssteg):

- Den smittsamma får behandling (med sannolikhet  $\gamma_H$ )
- Den smittsamma dör av HIV (med sannolikhet  $\mu_H(1 + I_T)$ )
- Den smittsamma dör av TBC (med sannolikhet  $\mu_T \cdot I_T$ )
- Den mottagliga grannen dör av TBC (med sannolikhet  $\mu_T \cdot I_T$ )

Där  $I_T$  är sannolikheten att en individ har aktiv tuberkulos, vilken är lika med tuberkulos endemiska nivå. Dessa händelser är inte nödvändigtvis exklusiva, utan flera kan ske i samma tidssteg enligt Venn-diagrammet i Figur 4.

Sannolikheten att någon av dessa sker i ett visst tidssteg är alltså

$$\Gamma = \mu_H(1 + I_T) + \gamma_H + 2I_T\mu_T - \mu_H(1 + I_T)\mu_T I_T - 2\gamma_H\mu_T I_T - (\mu_T I_T)^2 + \gamma_H(\mu_T I_T)^2$$

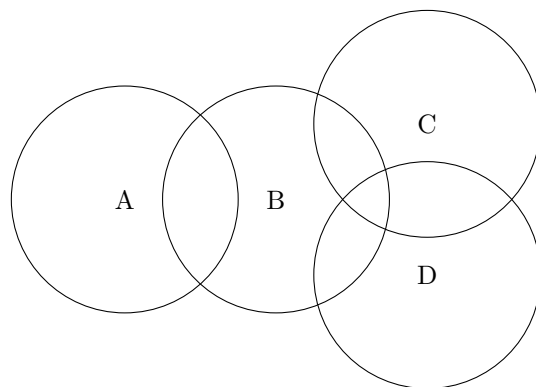
Sannolikheten att den smittsamma individen smittar den mottagliga i ett tidssteg är som vanligt  $\beta_{H_M}$  eller  $\beta_{H_F}$  beroende på kön.

Som vår simulering är uppbyggd kan smitta ske i samma tidssteg som någon av de fyra händelserna ovan sker.  $\sigma$  är alltså sannolikheten att smitta sker innan eller i samma tidssteg som någon av de fyra händelserna, i både fallet man till kvinna och i fallet kvinna till man. Detta betyder att:

$$\sigma = \beta_{H_K}/(\beta_{H_K} + \Gamma - \beta_{H_K}\Gamma) \cdot \beta_{H_M}/(\beta_{H_M} + \Gamma - \beta_{H_M}\Gamma) \quad (4)$$

Vi kan nu beräkna  $\rho_0$  enligt ekvation (2),(3) och (4).  $C_V^2$  kan skrivas som  $C_V^2 = E(K_2^2)/E(K)^2 - 1$  och beräknas på liknande sätt som  $E(K_2)$  ovan.

Vi har nu alla delar för att beräkna vårt reproduktionstal enligt ekvation (1).



A: Individen dör av HIV  
 B: Kontakten dör av TBC  
 C: Individen får HIV-behandling  
 D: Individen dör av TBC

Figur 4: Venn-diagram över situationer som kan leda till smitta inte längre kan överföras från en given smittsam individ till en given mottaglig kontakt

Framför allt på grund av att vår gradfördelning inte är exakt på grund av hur nätverket genererades i kapitel 2.1.1 kommer inte heller reproduktionstalet att vara exakt, men det kommer att vara tillräckligt nära för att vara en intressant storhet.

Enligt *threshold limit theorem* (se kapitel 3.2.1) är en epidemi endast möjlig om  $R_0 > 1$ . Om  $R_0 < 1$  kan alltså ett stort utbrott omöjligt ske. Enligt (Olinky & Stone) gäller det att  $R_0$  alltid är större än 1 om grafen är tillräckligt heterogen. Vi kommer dock endast att undersöka grafer vars gradfördelning har ändlig varians, det kommer alltså inte vara omöjligt att få ett reproduktionstal under 1, men för vissa nätverksgrafer bör det vara ovanligt.

### 3.2.3 Utbrottssannolikhet

Som vi vet är sannolikheten för en epidemi 0 då  $R_0 < 1$ . När reproduktionstalet är större finns en positiv sannolikhet för ett större utbrott. Vi beräknar den sannolikheten genom att göra flera simuleringar av HIV-spridning på nätverket och beräkna hur stor andel av simuleringarna som leder till stora utbrott.

### 3.2.4 Antal döda

Något som till stor del bestämmer hur allvarlig en epidemi är är dödstalet. Vi tittar på de simuleringar i kapitel 3.2.3 som leder till stora utbrott och noterar hur många av de smittade som dör på grund av HIV. Individer som dör enbart på grund av tuberkulos räknas inte med.

## 4 Simuleringar & Resultat

Till att börja med vill vi undersöka hur smittspridningarna beter sig för smittorna var för sig. För tuberkulos vill vi veta om andelen smittade över tid stämmer överens med de analytiskt beräknade endemiska nivåerna. För HIV kommer vi först att undersöka hur valet av kantfördelning påverkar spridningen, sedan vill vi få fram värden på storheterna i kapitel 3.2. När vi gör simuleringar med simultan spridning vill vi främst jämföra resultatet av HIV-spridning med och utan TBC närvarande i populationen. Vi vill även undersöka hur nivåerna av TBC påverkas av ett HIV-utbrott.

Vi använder Matlab för att göra beräkningar och simuleringar.

### 4.0.1 Tester

Antalet simuleringar som leder till stora utbrott är binomialfördelat, så för att ta reda på om skillnader i utbrottssannolikhet är signifikanta använder vi z-testet (Alm & Britton, kapitel 7.6.4.

Jämförelse av proportioner). Testet bygger på att vi kan approximera antalet utbrott (och därmed skillnaden) med normalfördelningar. Approximationen är tillåten när

$$(n_1 + n_2)\hat{p}(1 - \hat{p}) > 5$$

där  $n_1$  och  $n_2$  är antalet simuleringar,  $\hat{p} = (x + y)/(n_1 + n_2)$  och  $x_1$  och  $x_2$  är antalet simuleringar som leder till stora utbrott.

För att ta reda på om skillnader i utbrottsstorlek och antalet döda är signifikanta använder vi Welch's t-test (Alm & Britton, kapitel 7.6.2: Inferens om  $\mu_1 - \mu_2$  då varianserna är okända), det kan vi göra eftersom att *threshold limit theorem* (Kapitel 3.2.1) säger att storleken på stora utbrott är normalfördelade. För att använda testet krävs inte lika varians.

När vi testar använder vi signifikansnivå 5%. Vi kommer att göra testerna för flera olika parametervärden, därför låter vi signifikansnivån för varje test i en grupp av tester vara  $0.05/m$  där  $m$  är antalet parametervärden som testas. Detta för att signifikansnivån för testet i sin helhet ska vara 5%.

## 4.1 Val av parametervärden

Parametrar för smittspridningen är svåra att välja eftersom att de beror på många faktorer och inte går att läsa direkt ur statistik. De varierar dessutom i verkligheten geografiskt och mellan individer. Vi kommer därför att välja parametervärdena något godtyckligt och fokusera på beteenden hos smittorna snarare än absoluta värden. Då kan vi också välja parametrar som gör att andelen smittade blir relativt stor vilket gör att resultaten blir tydligare och lättare att studera. Framför allt kan vi se till att de koinficerade inte blir för få.

Till att börja med används värdena i tabellen nedan på populationsstorleken och parametrarna som beskrivs i kapitel 2.2 och 2.3. Sannolikheterna är per tidssteg. Ett tidssteg är tänkt att vara i ungefär en månad.

	Parameter	Värde	
Populationsstorlek	n	10000	
<b>TBC</b>			
Smittosannolikhet	$\beta_T/n_t$	$0.6/n_t$	20 gånger så hög för koinficerade
Aktiveringssannolikhet	$\epsilon_T$	0.04	
Behandlingssannolikhet	$\gamma_T$	0.12	
Dödssannolikhet	$\mu_T$	0.001	
<b>HIV</b>			
Smittosannolikhet (Män)	$\beta_{HM}$	0.02	Dubbelt så hög för koinficerade
Smittosannolikhet (Kvinnor)	$\beta_{HK}$	0.04	
Behandlingssannolikhet	$\gamma_H$	0.06	
Dödssannolikhet	$\mu_H$	0.03	

Populationsstorleken väljs till 10000, det gör den tillräckligt stor för att kunna approximera resultat för stora populationer men samtidigt liten nog för att kunna simuleras på en rimlig tid. Smittosannolikheten för TBC,  $\beta_T/n_t$  beskriver sannolikheten att som mottaglig individ bli smittad av en given smittsam individ. Sannolikheten skalar mot populationsstorleken i tidssteget,  $n_t$ . Den totala sannolikheten att bli smittad är ett minus sannolikheten att undkomma smitta från alla smittsamma:  $1 - (1 - \beta_T/n_t)^{I_t}$ . Här beskriver  $I_t$  antalet smittsamma individer vid tidpunkten.

Parametervärdena för TBC-spridning är inspirerade av siffror från WHO (2016) som beskriver TBC-läget i Afrika år 2015. I en population på 989 miljoner upptäcktes 2.75 miljoner nya fall (av aktiv TBC) varav 31% med HIV. 0.45 miljoner dog i TBC utan HIV i kroppen och 0.3 miljoner dog i TBC med viruset. I Afrika är latent TBC mycket vanligare än i resten av världen och beräknas finnas i runt 80% av befolkningen.

Parametervärdena för HIV-spridning är inspirerade av siffror från UNAIDS (2016) som beskriver HIV-läget i södra och östra Afrika 2015. I en population på 479 miljoner levde då 19 miljoner då med HIV, 54% med behandling. 0.96 miljoner nya fall upptäcktes och 0.47 miljoner dog i AIDS-relaterade sjukdomar, en tredjedel av dessa var TBC.

Hur våra parametervärden beräknades kommer inte att presenteras eftersom det är något godtyckligt. Dock är samtliga parametrar för HIV mycket högre än i verkligheten eftersom att viruset annars skulle finnas i del av vår befolkning som är för liten för att pålitligt kunna analysera.

Aktiveringssannolikheten för TBC valdes även den som mycket högre än i verkligheten av samma anledning. Anledningen till att överdriva spridningen är framför allt för att vi vill att en tillräckligt stor del av befolkningen ska vara koinficerade. Med verkliga parametrar skulle den delen vara obetydligt liten. Eftersom att fler än i verkligheten nu blir smittade av TBC sänker vi även risken att dö så att antalet inte blir för stort.

När vi sedan gör simuleringar kommer chansen att få behandling för HIV,  $\gamma_H$ , vara den vi ändrar mest på när vi vill undersöka hur HIV-smittan påverkas av att vara olika allvarlig. Detta för att det känns rimligt att chansen att få behandling är den parameter som skiljer sig mest runt om i världen och är lättast att påverka.

#### 4.1.1 Kantfördelningen

För att skapa den bipartita nätverksgraf för HIV-spridning behöver vi bestämma egenskaperna för detta. Vi har i kapitel 2.1 valt att ge varje nod ett antal kanter efter en bestämd gradfördelning och sedan koppla ihop noderna slumpmässigt utefter dessa. Vi låter för enkelhetens skull gradfördelningen vara densamma för män och kvinnor. I en mer komplicerad modell skulle det dock kunna vara av intresse att låta fördelningen variera men låta väntevärdena vara lika.

Gradfördelningen för en nod beskriver alltså hur många sexuella partners en individ har. Det är rimligt att anta att de flesta under sin livstid har ett relativt litet antal partners medan några få har betydligt fler. Vi vill alltså ha en fördelning där sannolikheten för ett litet antal partners är stor, men där det samtidigt finns en sannolikhet för ett mycket större antal. En fördelning med en sådan egenskap kallas tungsvansad. För att kunna styra tyngden på fördelningens svans kommer vi att jämföra olika Poissonfördelningar med stokastiska, Paretofördelade väntevärden där vi sedan lägger till 1 (individer utan partners inte är intressanta i simuleringen). Till Paretofördelningen kan vi välja ett väntevärde  $\mu$  och en parameter  $\alpha$  som styr tyngden på fördelningens svans (lägre värde ger tyngre svans). Dessa egenskaper överförs till Poissonfördelningen.

Paretofördelningen har sannolikhetsfunktion

$$f(x) = \alpha s^\alpha / x^{\alpha+1}$$

$$s \leq \infty, 0 < s, 0 < \alpha$$

Där  $s = \mu(\alpha - 1)/\alpha$ .

För ett ändligt väntevärde krävs att  $\alpha > 1$  och för ändlig varians krävs att  $\alpha > 2$ . Poissonfördelningen har sannolikhetsfunktion

$$P(K = k | \Lambda = \lambda) = \lambda^k e^{-\lambda} / k!$$

$$0 \leq k, 0 < \lambda$$

där i vårt fall  $\lambda$  har en Paretofördelning.

Vi låter väntevärdet vara 5 och svanstyngdparametern  $\alpha$  vara 2.2. Värdet på svanstyngden är tagen från en studie som refereras av Newman (2002) som kommer fram till att antalet sexuella partners hos svenskar har en fördelning med en svanstyngd i närheten av det värdet. Även om vår undersökning fokuserar på Afrika är 2.2 ett bättre val än ett helt godtyckligt värde.

	Parameter	Värde
Förväntat antal kontakter	$E[K]$	5
Svanstyngd	$\alpha$	2,2

## 4.2 Simulering av smittoförlopp

Ett smittoförlopp simuleras här på följande sätt:

1. Generera nätverksgraf för spridning av HIV enligt kapitel 2.1.1.
2. Smitta en person med HIV

3. Beräkna de endemiska nivåerna för TBC och smitta slumpmässigt en andel av befolkningen enligt dessa
4. I varje tidssteg, gå igenom alla individer och:
  - (a) Om individen har smittsam HIV
    - i. Smitta varje mottaglig kontakt med sannolikhet  $\beta_{H_M}$  eller  $\beta_{H_K}$  beroende på kön
    - ii. Få behandling med sannolikhet  $\gamma_H$   
Annars dö med sannolikhet  $\mu_H$  ( $20\mu_H$  om individen har smittsam TBC)
  - (b) Om individen är mottaglig för TBC:  
Bli smittad med sannolikhet  $\beta_T I_T n_t$  ( $I_T n_t$  är antalet individer med smittsam TBC)
  - (c) Om individen har latent TBC:  
Övergå till smittsam TBC med sannolikhet  $\epsilon_T$  ( $2\epsilon_T$  om individen har smittsam HIV)
  - (d) Om individen har smittsam TBC:  
Få behandling med sannolikhet  $\gamma_T$   
Annars dö med sannolikhet  $\mu_T$

Simuleringen fortsätter tills spridningen av HIV har slutat eller fram tills en bestämd tidpunkt. Nivåerna av smittorna noteras i varje tidssteg om det är intressant för vad som undersöks.

### 4.3 TBC

När vi undersöker spridningen av TBC börjar vi med att låta Matlab lösa differentialekvationerna i kapitel 3.1.1 för att få reda på vad vi kan förvänta oss. För de valda parametervärdena får vi följande endemiska nivåer:

S=0.2  
E=0.6  
I=0.2

Andelen av befolkningen som bär på bakterien stämmer alltså bra överens med de 80% som estimeras bära på den i Afrika. Dock har vi hela 20% med aktiv smitta, betydligt högre än i verkligheten, som vi önskade. När en så pass stor andel av befolkningen är smittsam kommer koinfektionerna med HIV inte att bli för få.

När vi gör simuleringarna av endast TBC låter vi startnivåerna skilja sig från de förväntade endemiska nivåerna för att tydligare se konvergensen, senare kommer vi dock att starta på de endemiska nivåerna.

I Figur 6 plottas resultatet av en simulering med startnivåer [0.7,0.2,0.1] för mottagliga, latent och smittsamma som andelar av den levande befolkningen. Som väntat konvergerar nivåerna mot [0.2,0.6,0.2]. Det totala antalet döda i tidssteg 300 uppkom till 532.

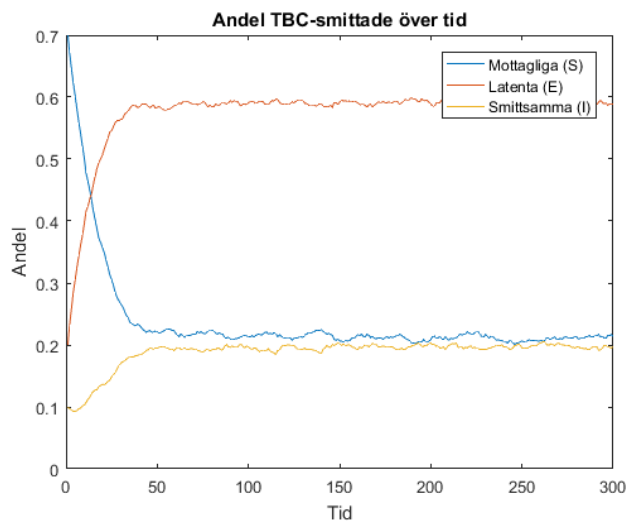
### 4.4 HIV

När vi simulerar ett HIV-utbrott med de valda parametrarna ser det ut som förväntat. I Figur 6 ser vi att antalet smittsamma först ökar för att sedan minska och försvinna. I slutet av epidemin är alla antingen friska, behandlade eller döda. 3532 personer blev totalt smittade och 1160 dog.

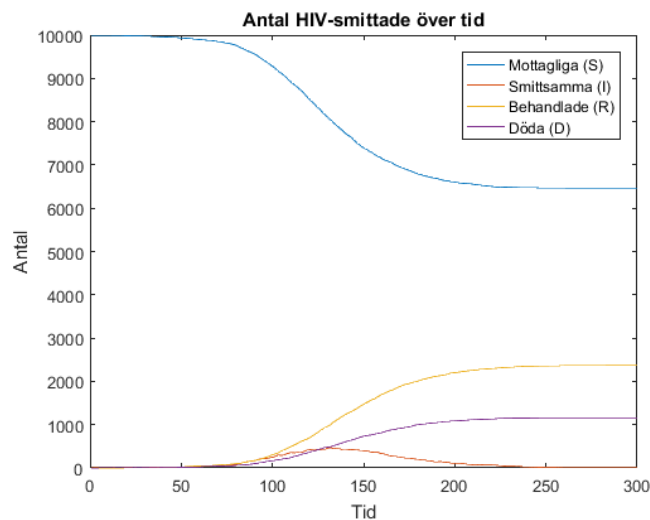
I Figur 7 ser vi slutstorlekar för 1000 simulerade HIV-utbrott. Som väntat är resultatet uppdelat i två: små och stora utbrott. De stora utbrotts storlek ska enligt *threshold limit theorem* (kapitel 3.2.1) vara normalfördelade och det ser ut att stämma. För valda parametervärden ledde endast en fjärdedel av simuleringarna till stora utbrott. Huruvida det är realistiskt är svårt att säga eftersom att virusets epidemistatus är en huvudanledning till varför vi studerar det.

#### 4.4.1 Kantfördelning

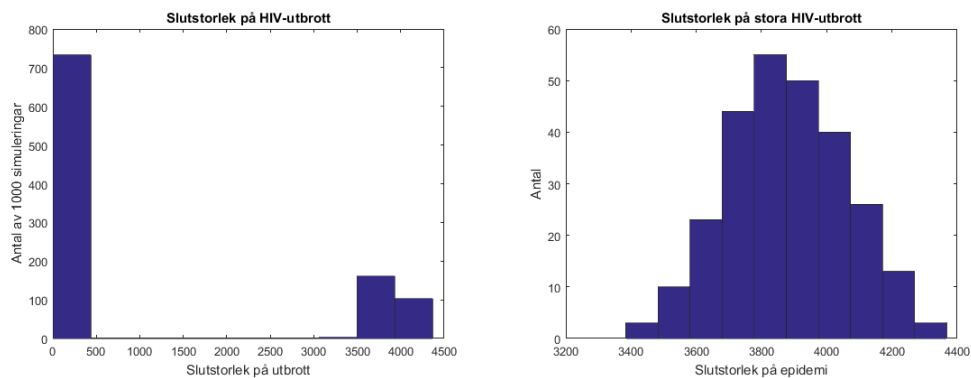
Vi kommer nu att undersöka hur olika gradfördelningar (fördelningar av antal sexuella partners) påverkar spridningen av viruset. Vi kommer att jämföra de typer av Poisson-Pareto-fördelningar som beskrivs i kapitel 4.4.1 för olika svanstyngder. Vi kommer även att titta på en vanlig Poissonfördelning som är ekvivalent med en Poisson-Pareto-fördelning med  $\alpha = \infty$ . I Figur 4.4.1 kan



Figur 5: Andelen mottagliga (S), latenta (E) och smittsamma (I) av levande över 300 tidssteg



Figur 6: Antal mottagliga (S), smittsamma (I), behandlade (R) och döda (D) under en HIV-epidemi

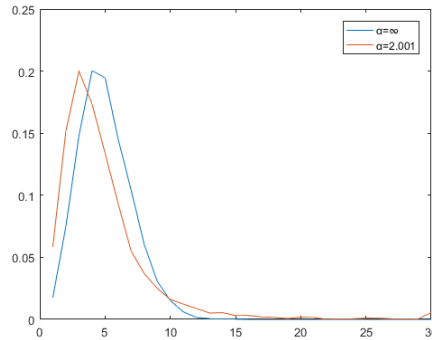


Figur 7: Histogram över slutstorlekar på 1000 HIV-utbrott. Histogrammet till höger visar den högra stapeln i histogrammet till vänster.

vi se sannolikhetsfördelningarna för två olika svanstyngder. I figuren ser skillnaden liten ut, men



när svansen är tung är de största värdena betydligt större, något som inte syns i grafen. Figuren genererades från 10000 punkter i varje fördelning varav den tungsvansade hade 30 observationer utanför det presenterade området medan den andra inte hade några. Vi kommer att vara intresserade av huruvida det spelar någon roll att några få individer har ovanligt många partners.



Figur 8: Empiriskt genererade sannolikhetsfunktioner för Poisson-Pareto-fördelningen med svanstyngder 2.001 och  $\infty$

I Tabell 1 presenteras resultatet av simuleringar med fem olika svanstyngder och två olika behandlingssannolikheter. För varje kombination av parametervärden görs 1000 simuleringar. Förväntat antal partners är som vanligt satt till 5. TBC är inte närvarande.

$\alpha$	$\gamma_H$	$R_0$	Andel utbrott	Medelvärde utbrottsstorlek	Medelvärde antal döda
$\infty$	0.02	6.53	54%	7315	4386
100		6.53	53%	7310	4384
3		17.0	55%	7174	4304
2.2		107	53%	6894	4139
2.001		223	54%	6829	4098
$\infty$		0.09	1.86	3%	156
100	1.86		3%	194	48
3	4.84		8%	1238	309
2.2	30.4		15%	2092	523
2.001	63.5		17%	2590	268

Table 1: Resultat av simuleringar med olika svanstyngd på kantfördelningen

Svansens tyngd tycks ha mycket liten påverkan på utbrottssannolikheten när få får behandling, men stor påverkan då många får behandling. En tyngre svans tycks då leda till att betydligt fler epidemier sker då behandlingssannolikheten är hög.

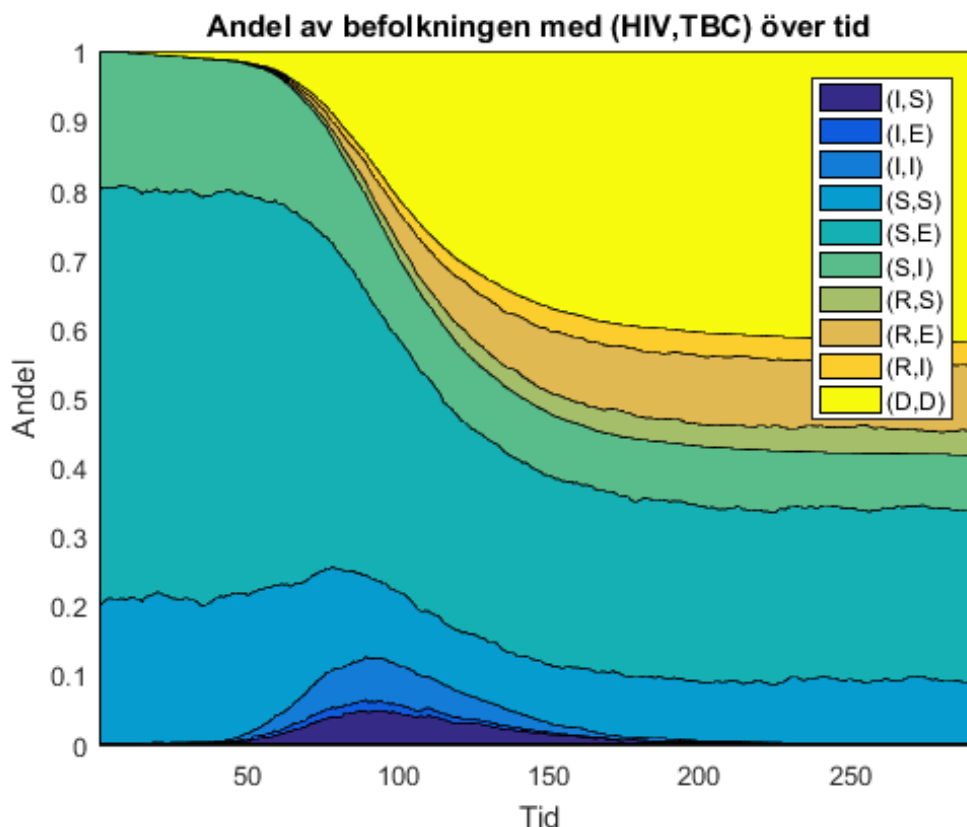
Efter att ha bekräftat att normalapproximation är tillåten gör vi z-tester för att avgöra om andelen utbrott skiljer sig signifikant för olika svanstyngder. För  $\gamma_H = 0.09$  är skillnaden signifikant för samtliga kombinationer och rimliga signifikansnivåer utom när vi jämför  $\alpha = \infty$  med  $\alpha = 100$  och  $\alpha = 2.001$  med  $\alpha = 2.2$ , vilket inte är helt oväntat då dessa fördelningar ligger nära varandra. Att de två fördelningar med högt värde på  $\alpha$  har mycket lika  $R_0$ -värden bekräftar detta. För  $\gamma_H = 0.02$  kan testet inte visa på att svanstyngd gör någon signifikant skillnad för utbrottssannolikheten. För behandlingssannolikheten 0.001 (som inte visas i tabellen) går det dock att visa på att nätverk med tungsvansade gradfördelningar har en signifikant lägre risk för epidemi.

Både för hög och låg behandlingsgrad ser det ut att finnas mönster i utbrottsstorlek och antal döda. För liten behandlingsgrad tycks en tyngre svans leda till mindre utbrott och färre döda medan svanstyngd ser ut att ha motsatt effekt då behandlingsgraden är hög.

Welch's t-test visar att alla skillnader är signifikanta (p-värde under 0.0001) utom de mellan svanstyngderna 100 och  $\infty$ . Utbrottsstorlek och antal döda ökar alltså med ökad svanstyngd när många får behandling och minskar med ökad svanstyngd när färre får behandling, ett mönster som liknar det hos sannolikheten för stora utbrott. En förklaring kan vara att antalet kontakter en smittad i genomsnitt behöver ha för att sprida viruset vidare skiljer sig för olika behandlingsgrad. För hög behandlingsgrad, där det kan tänkas krävas att smittan når en individ med ett stort antal kontakter är detta mer sannolikt på en graf med tungsvansad gradfördelning. För låg behandlingsgrad krävs det dock endast att smittan når individer som inte tillhör den grupp individer med lägst gradtal, detta är mer sannolikt på en graf med en lättare svans eftersom att en tungsvansad graf förutom att ha fler individer med höga gradtal även har flera individer med låga gradtal för att kompensera i väntevärdet. Samma resonemang kan användas för att förklara mönstret hos utbrottssannolikheterna.

## 4.5 Simultan spridning

När vi studerar simultan smittspridning börjar vi med att studera ett enda smittförlopp som ledde till en HIV-epidemi över tid. Vi låter behandlingssannolikheten för HIV  $\gamma_H$  vara 0.02, något lägre än annars för att tydligare kunna se eventuella effekter. I Figur 9 ser vi hur storleken på grupperna förändras över tid. Vi ser att HIV-epidemin betar sig som tidigare; den ökar för att sedan minska och försvinna. Vi ser också att andelen av populationen som bär på latent TBC är betydligt lägre bland de med smittsam HIV jämfört med den friska befolkningen, vilket troligtvis beror på att de som har både latent TBC och smittsam HIV snabbt övergår till att ha aktiv TBC.

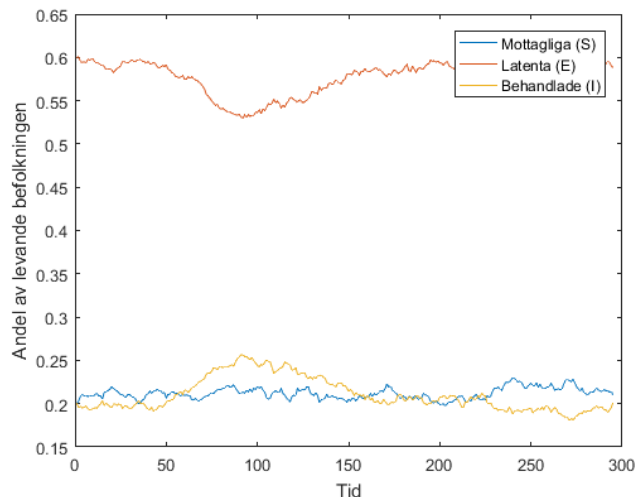


Figur 9: Antal smittade över tid under ett HIV-utbrott med närvarande TBC

### 4.5.1 Påverkan på TBC

I Figur 10 ser vi tydligare vad som händer med TBC-nivåerna under ett HIV-utbrott. Andelen friska individer förblir nära oförändrad medan andelen med latent TBC minskar och aktiv TBC

ökar. Att fler blir aktivt sjuka beror på att det är just det som HIV påverkar hos TBC-bakterien: risken att övergå från latent till aktiv smitta. Det är alltså inte fler som blir smittade av TBC under en HIV-epidemi utan bara fler som blir sjuka och visar symptom. Efter utbrottet återgår smittan till de tidigare endemiska nivåerna.



Figur 10: TBC-nivåer under HIV-utbrott

#### 4.5.2 Påverkan på HIV

För att undersöka hur en närvarande tuberkulosendemi påverkar spridningen av HIV gör vi simuleringar med och utan TBC och jämför resultaten. Vi gör också simuleringar där TBC närvarar, men i sig själv inte är dödlig. Detta för att kunna bekräfta att eventuella effekter inte bara beror på TBCs dödlighet. Vi låter också behandlingssannolikheten variera för att kunna undersöka om tuberkulos påverkan beror på hur allvarligt HIV-utbrottet är. För varje val av parametrar gör vi 1000 simuleringar av smittoförlopp och noterar de intressanta storheterna. Vi använder behandlingssannolikheter mellan 0.01 och 0.14 med 0.01-intervall.

#### 4.5.3 $R_0$

I Tabell 2 nedan redovisas värden på  $R_0$  för några av de undersökta parametervärdena. Värdena på  $R_0$  är något lägre när TBC är närvarande.

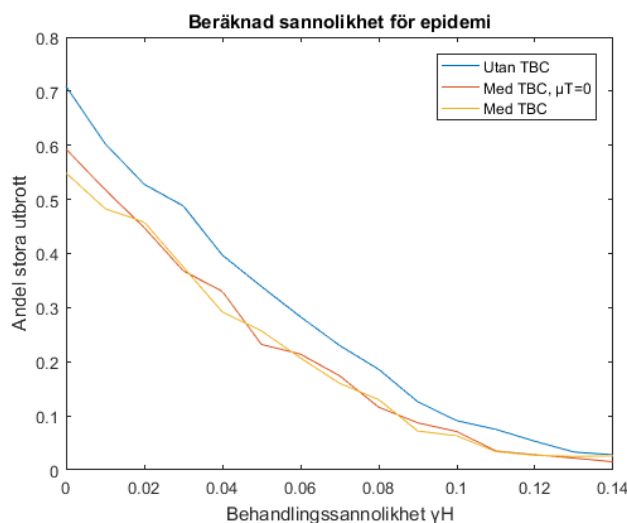
$\gamma_H$	0	0.04	0.09	0.14
Med TBC	156	60	28	16
Utan TBC	191	68	30	17

Table 2: Värden på  $R_0$  för olika behandlingssannolikhet, med och utan TBC i befolkningen.

**Utbrottssannolikhet** Figur 11 visar de beräknade sannolikheterna att ett HIV-virus som introduceras i en befolkning leder till ett stort utbrott. Vi ser att sannolikheten tycks vara större utan TBC i befolkningen. Med hjälp av z-test bekräftar vi att skillnaden i utbrottssannolikhet för simuleringar med och utan TBC är signifikant för alla nivåer av behandlingssannolikhet upp till och med 0.08. De högre behandlingssannolikheternas värden är något osäkra i allmänhet eftersom att det är svårare att definiera gränsen för ett stort utbrott när få blir smittade. Någon skillnad mellan TBC med och utan dödlighet går inte att se i dessa simuleringar.

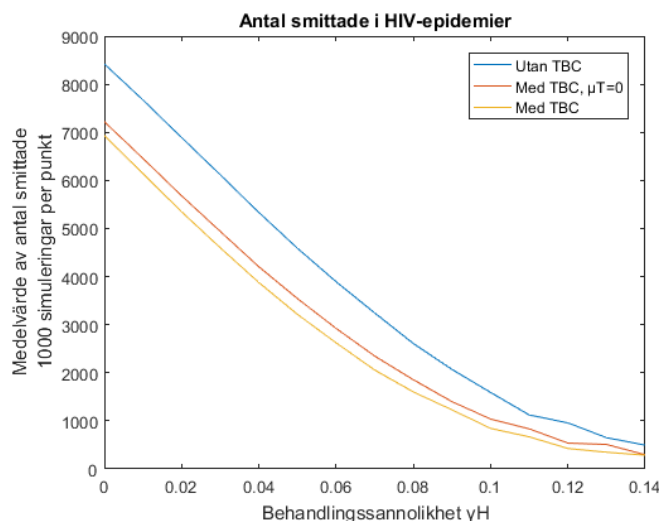
Skillnaden i sannolikhet beror troligtvis på att risken att dö i HIV ökar för en individ med TBC, alltså är sannolikheten större att individen dör innan viruset hinner spridas vidare. Att vi inte ser någon effekt av att dödligheten i TBC sätts till 0 beror troligtvis på att TBCs dödlighet

har valts som liten redan från början, annars är det rimligt att tro att även den dödligheten skulle påverka sannolikheten för stora utbrott. Det kan också bero på att vår stickprovsstorlek är relativt liten.



Figur 11: Andel av 1000 simuleringar som leder till HIV-epidemi

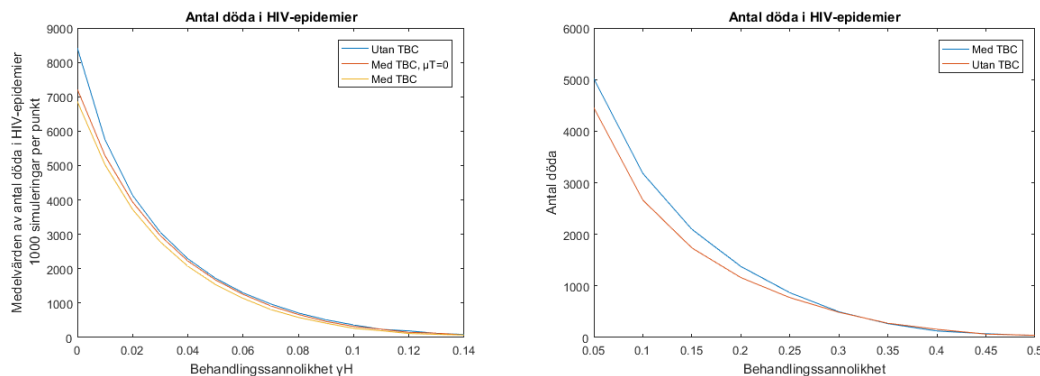
**Utbrottsstorlek** I Figur 12 ser vi att även storleken på epidemierna minskar när TBC är närvarande i befolkningen. Welch's T-test visar att skillnaden är signifikant för alla undersökta behandlingssannolikheter under 0.13 mellan försök med och utan TBC. Anledningen är antagligen densamma som ovan; individer hinner dö innan viruset sprids vidare. Skillnaderna i utbrottsstorlek mellan de två typerna av modeller med TBC är signifikant för alla undersökta behandlingssannolikheter under 0.11 vilket visar att även den ökade dödlighet för koinficerade spelar roll för spridningen av HIV. Att högre behandlingssannolikheter inte ger signifikans lär bero både på att stickprovsstorlekarna (antal stora utbrott) är mindre samt att definitionen av stora utbrott är osäker när ingen tydlig gräns går mellan stora och små utbrott.



Figur 12: Genomsnittliga slutstorlekar på HIV-epidemier

**Antal döda** I den vänstra grafen i Figur 13 ser vi att fler dör när TBC är frånvarande. I Den högra grafen ser vi samma simuleringar gjorda för andra parametervärden, men där gäller det motsatta. Vilket fall som är värst beror alltså på hur stor risken att dö i TBC är samt hur stor

effekt TBC har på risken att dö i HIV. Welch's T-test visar att skillnaden är signifikant olika för alla tre modeller när behandlingssannolikheten är under 0.1. Eftersom att vår modell har en låg risk att dö i TBC men leder till att en större del blir sjuka är det svårt att säga vilket fall som gäller i verkligheten.



Figur 13: Antal döda i HIV-epidemier för olika parametervärden och behandlingssannolikheter

## 5 Slutsatser och diskussion

Vi har undersökt smittspridning av TBC, HIV samt av de båda simultant och kan konstatera att det två smittorna påverkar varandras spridning.

Under ett HIV-utbrott ökar andelen individer med smittsam TBC i en befolkning samtidigt som andelen med latent TBC minskar. Andelen friska individer tycks vara oförändrad. När HIV-utbrottet är över återgår andelarna till de tidigare endemiska nivåerna. Att bakterien inte sprider sig till större delar av befolkningen kanske inte är vad man intuitivt förväntar sig, men så tycks vara fallet. TBC blir alltså inte vanligare, men allvarligare, under ett HIV utbrott.

Det visar sig att HIV utbrott blir både mindre och mindre sannolika när TBC närvarar. Detta beror på att TBC leder till större dödlighet som i sin tur förhindrar att viruset sprids vidare. TBC förhindrar alltså att HIV får maximal spridning. På vilket sätt TBC bidrar till det totala antalet döda i verkligheten går inte att avgöra med vår modell på grund av de orealistiska parametrarna, effekten tycks bero på val av parametrar. I vår modell är det dock fler som dör med HIV när TBC är frånvarande.

Fördelningen av antalet sexuella partners visar sig spela en stor roll i hur stor en HIV-epidemi blir. På vilket sätt påverkan sker beror på hur lätt smittan sprids i allmänhet; en smita som är lätt att behandla sprids bättre när antalet kontakter individerna har har stor variation medan en svårbehandlad smitta sprids bättre på ett nätverk som är mer homogent.

I framtida studier finns det framför allt två intressanta riktningar för fördjupning. Den första är en mer analytisk riktning som fokuserar på att studera vår förenklade modell mer ingående och hitta teoretiska resultat gällande simultan smittspridning. Den andra är att utöka modellen för att bättre spegla verkligheten och göra simuleringar utan att nödvändigtvis försöka komma fram till teoretiska resultat. Det finns många sätt att göra modellen mer verklighetstrogen, vi kommer nu att gå igenom några exempel på dessa.

I det här arbetet gjordes inga ingående undersökningar av vilka parametervärden som bäst speglar verkligheten men det skulle vara intressant att försöka komma fram till mer sanningsenliga värden och göra om försök utefter dessa.

Många förenklade antaganden gjordes i vår modell. Flera av dessa skulle vara intressanta att lätta på i studier av mer komplexa modeller.

Den största förenklingen vi gjorde var att anta en sluten befolkning bortsett från dödsfallen. Eftersom att både HIV och TBC har långa smittförlopp kan vi gissa att det är den främsta anledningen till att modellen avviker från verkligheten. Ett tecken på det är att HIV i många länder har endemisk status, något som inte är möjligt i en SIR-modell utan demografi. Att vi valde att göra förenklingen beror främst på att det är svårt att integrera det bipartita kontaktnätverket i en sådan modell. Att sprida en epidemi i en population där den samspelar med en endemisk smitta är dessutom ett intressant problem i sig att studera.

I kontaktnätverket för spridning av HIV antogs en gradfördelning som var lika för båda könen. Detta verkar inte helt stämma överens med verkligheten. I (Anderson & May) diskuteras hur fördelningen varierar mellan könen i Afrika, som vi har fokuserat på. Kvinnorna kan antas ha en större spridning än män, de flesta kvinnor har färre sexuella partners än de flesta män medan ett fåtal (ofta prostituerade) har betydligt fler än resten av befolkningen. Det är känt att dessa individer har en stor roll i spridningen av HIV. Dessutom antog vi att kanterna parades ihop helt slumpmässigt, men i verkligheten är det finns det ofta beroenden som ger upphov till ytterligare strukturer i grafen.

Vi har inte heller tagit hänsyn till att människor ofta är monogama i perioder och inte har samma sannolikhet att ha sexuell kontakt med alla möjliga partners vid alla tidpunkter.

Även i spridningen av TBC skulle man kunna använda ett nätverk, till exempel en hushållsmodell där individer kan ha kontakt med alla i befolkningen men där individer oftare har kontakt med individer i samma hushåll.

En annan förenkling som har gjorts är för smittoförloppen. Vi har antagit samma övergångssannolikheter i alla tidssteg. I verkligheten ser det annorlunda ut. Framför allt för HIV, där smittsamheten är betydligt högre i början av den smittsamma perioden och i slutet, när individen har nått AIDS.

Något som man i verkligheten försöker göra för att förhindra spridningen är också att testa individer med den ena sjukdomen för att ta reda på om de även bär på den andra. Om detta görs är det alltså möjligt att sannolikheten att den okända sjukdomen upptäcks och blir behandlad ökar vid koinfektion, en effekt som vår modell inte tar hänsyn till.

Trots dessa förenklingar är de valda modellerna intressanta att studera, både för sin egen skull och för de aningar de ger om hur HIV och tuberkulos sprids i verkligheten.

## 6 Referenser

Alm, S. E. & Britton, T. (2008): Stokastik: Sannolikhetsteori och statistikteori med tillämpningar. Liber.

Anderson, R. M. & May, R. M. (1991): Infectious Diseases of Humans, Dynamics and Control, Oxford Science Publications.

Andersson, H. & Britton, T. (2000): Stochastic Epidemic Models and Their Statistical Analysis. Springer.

Olinky, R & Stone, L (2004): Unexpected epidemic thresholds in heterogeneous networks: The role of disease transmission. Phys. Rev. E Vol 70, 030902.

Newman, M. E. J. (2002): The spread of epidemic disease on networks. Phys. Rev. E Vol 66:1, 016128.

UNAIDS (2016): Fact Sheet November 2016. Hämtad från <http://www.unaids.org/en/resources/fact-sheet>

WHO (2016): Global Tuberculosis Report 2016. Hämtad från [http://www.who.int/tb/publications/global\\_report/en/](http://www.who.int/tb/publications/global_report/en/)