



Stockholms
universitet

Slutstorlekar hos Reed Frost-epidemier på slumpgrafer skapade enligt den vik- tade konfigurationsmodellen

Anders Carlsson

Kandidatuppsats 2012:1
Matematisk statistik
Februari 2012

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Slutstorlekar hos Reed Frost-epidemier på slumpgrafer skapade enligt den viktade konfigurationsmodellen

Anders Carlsson*

Februari 2012

Sammanfattning

I denna uppsats studerar vi epidemier på slumpgrafer. Epidemier på grafer kan t.ex. användas för att modellera en verklig sjukdomsepidemi eller ett rykte som sprids i en population av människor men kan också modellera data som skickas på ett datanätverk. Matematiskt är epidemin en stokastisk process som utspelar sig på en graf som i sin tur konstrueras med hjälp av en stokastisk modell. Epidemimodellen vi använder är *Reed Frost-modellen* och modellen för att konstruera grafen är den *viktade konfigurationsmodellen*. Vid studier av en epidemi, är man ofta intresserad av under vilka förutsättningar ett stort utbrott kan ske och, givet att ett stort utbrott skett, hur många individer som blivit smittade. Denna uppsats undersöker det senare d.v.s. storleken på ett utbrott av en Reed Frost-epidemi på en slumpgraf som konstruerats efter den viktade konfigurationsmodellen. Istället för analytiska beräkningar använder vi Matlab för att simulera några exempel på grafer och epidemier från vilka vi skattar slutstorlekarna av epidemiutbrotten. Vi beräknar också reproduktionstalet R_0 , för varje exempel, vilket är det förväntade antalet nya smittfall som genereras av en given smittad individ i början av tidsförloppet. Från simuleringarna ser vi att det inte verkar finnas några andra generella samband mellan R_0 och slutstorlekarna av epidemierna förutom att $R_0 > 1$ implicerar att ett stort utbrott är möjligt. Vi ser fall där R_0 är avtagande medan slutstorlekarna är växande och vice versa. Vi ger också exempel i vilka förekomsten av höggradiga noder, vilket typiskt gör att epidemin har lättare att ta fart, kan tas ut genom små sannolikheter för smitta mellan höggradiga noder.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: anders.c@telia.com. Handledare: Maria Deijfen.

Abstract

In this thesis we study epidemics on random graphs. Epidemics on graphs can be used for example to model a real life epidemic disease or a rumour spreading in a human population but could also model data being sent on a computer network. Mathematically, an epidemic is a stochastic process which takes place on a graph which in turn is constructed using a stochastic model. The epidemic model we use is the *Reed Frost model* and the model used to construct the graph is the *weighted configuration model*. When studying an epidemic, one is often interested in under what circumstances a large outbreak can occur and, given that a large outbreak has occurred, how many individuals that has ultimately been infected. This thesis investigates the latter, that is, the outbreak sizes of Reed Frost-epidemics on a random graph constructed according to the weighted configuration model. Rather than an analytic approach we use Matlab to simulate some examples of graphs and epidemics from which we can estimate the final sizes of the epidemic outbreaks. We also calculate the basic reproduction number R_0 , for each example which is basically the expected number of new cases generated by a given infective in the beginning of the time course. From the simulations we see that there seems to be no other general connection between R_0 and the final sizes of the epidemics other than that $R_0 > 1$ implies that a large outbreak is possible. We see cases where R_0 is decreasing while the outbreak sizes are increasing and vice versa. We also give examples in which the presence of high degree nodes, which typically makes the epidemic take off more easily, can be cancelled out by a small probability of infection between high degree nodes.

Tack

Jag vill tacka min handledare Maria Deijfen för att ha gett mig en inblick i detta intressanta forskningsområde och för all hjälp med att genomföra denna uppsats.

Innehåll

| | | |
|----------|------------------------------------------------------|-----------|
| 1 | Inledning | 3 |
| 2 | Beskrivning av modell | 3 |
| 2.1 | Reed Frost-epidemier | 3 |
| 2.2 | Konfigurationsmodellen | 5 |
| 2.3 | Epidemistorheter | 6 |
| 2.4 | Uttryck för R_0 i konfigurationsmodellen | 7 |
| 2.5 | Från konfigurationsmodellen till τ | 8 |
| 3 | Metod | 8 |
| 4 | Simuleringar av underliggande grafer | 9 |
| 4.0.1 | Exempel 1a | 9 |
| 4.0.2 | Exempel 1b | 10 |
| 4.1 | Exempel 2 | 11 |
| 4.1.1 | Exempel 2a | 12 |
| 4.1.2 | Exempel 2b | 12 |
| 5 | Simulering av epidemier | 12 |
| 5.1 | Exempel 3 | 17 |
| 5.2 | Exempel 4 | 17 |
| 5.3 | Exempel 5 | 18 |
| 6 | Diskussion | 21 |
| A | Härledning | 23 |
| B | Figurer för Exempel 1, 2 och 3 | 23 |

1 Inledning

I den här uppsatsen kommer vi studera epidemier på nätverk. Nätverk kan vara sociala strukturer, t.ex. vänskapsrelationer, men det kan också vara andra typer av nät, t.ex. datornätverk. Epidemin i fråga kan vara en sjukdom hos människor, som en förkylning, men det skulle också kunna vara t.ex. ett rykte som sprids. I datorvärlden kan epidemin vara ett datorvirus eller någon typ av spridning av information. I modeller för en epidemi finns det ofta brytpunkter där antalet smittade exploderar i antal. Utifrån detta ställer man sig frågorna:

Under vilka förutsättningar exploderar epidemin i storlek?

Om epidemin exploderar, hur stort blir utbrottet?

Den senare frågan är det vi kommer avhandla i denna uppsats, i fallet då det underliggande nätverket är en stokastisk viktad graf skapad enligt konfigurationsmodellen och epidemin är en Reed Frost-epidemi. Konfigurationsmodellen konstruerar vårt nätverk i form av en viktad graf med hjälp av en gradfördelning, en viktfördelning och en smittsannolikhetsfunktion. Nodernas grader bestäms efter en given gradfördelning och vikterna på kanterna som sitter fast i noderna bestäms med hjälp av en viktfördelning som tillåts bero på nodernas grad. Vi bestämmer med hjälp av smittsannolikhetsfunktionen en sannolikhet för smittöverföring för var och en av de viktade kanterna i grafen. Vi startar sedan en Reed Frost-epidemi på grafen. Reed Frost-epidemin börjar med att en nod smittas, den för vidare smittan med de sannolikheter som bestämts utifrån vikterna. De grannar som blir smittade kan i sin tur sprida smittan vidare etc. När det sedan står klart hur epidemin spridit sig bestämmer vi hur stor andel av populationen som smittats sammanlagt. I [1] beskrivs en metod där man genom genererande funktioner tar fram ett ekvationssystem vars minsta lösning ger slutstorleken på epidemin. Dock är denna teknik så pass krånglig att vi istället valt att göra simuleringar av epidemierna. Vi kommer bredvid samtliga simuleringsexempel även att räkna fram storheten R_0 , reproduktionstalet, som säger hur många förväntade nya smittfall en smittad nod ger upphov till. Den första frågan ovan kan kort besvaras med *epidemin riskerar att explodera då $R_0 > 1$* . Då R_0 är stor har smittan lättare att ta fart vilket gör att det är naturligt att tänka sig att R_0 kan ge en indikation på hur stora epidemierna blir till slut. Vi kommer se flera simuleringsexempel där R_0 ökar samtidigt som epidemiernas storlekar ökar, men vi kommer även att se det omvända. Vi kommer också illustrera att effekten av höga gradtal kan kompenseras av en lägre smittsannolikhet på kanterna mellan de höggradiga noderna och vice versa.

2 Beskrivning av modell

I det här avsnittet kommer vi beskriva hur modellen är uppbyggd genom att definiera Reed Frost-epidemier, Reed Frost-epidemier på grafer och konfigurationsmodellen.

2.1 Reed Frost-epidemier

Vi börjar med att formulera den ursprungliga definitionen av en Reed Frost-epidemi.

Definition 1 En Reed Frost-epidemi $\{RF(t), t = 1, 2, \dots\}$ är en stokastisk process i diskret tid där $RF(t)$ är mängden av individer som är smittade vid tiden t . Om individen v_i är smittad vid tiden t löper varje mottaglig individ v_j sannolikheten $\frac{\gamma}{n}$ att bli smittad av v_i vid tiden $t+1$. En individ v som är smittad vid tiden t slutar från $t+1$ att vara mottaglig för ny smitta och påverkar därefter inte epidemiprocessen. Processen pågår tills dess att inga nya individer smittas.

Från definitionen följer att populationen av individer är homogen d.v.s alla är lika mottagliga för smittan och lika benägna att sprida smittan vidare. Det följer också att alla kan smitta alla. Vi inskränker nu den senare förutsättningen genom att införa en social struktur som bestämmer vilka individer som har kontakt med vilka och därmed också vilka som kan smitta vilka. Detta gör vi genom att införa en viktad graf G och låta Reed Frost-epidemin utspela sig på denna.

Definition 2 En viktad graf G definieras som en trippel av mängder $G = G(V, E, W)$ där $V = \{v_1, v_2, \dots, v_n\}$ betecknar nodmängden, $E = \{e_{\{ij\}}; i, j \in \{1, 2, \dots, n\}\}$ betecknar kantmängden och $W = \{w_e; e \in E\}$ betecknar viktmängden.

I grafen representeras individerna av noderna och de möjliga smittvägarna av kanterna. Vikterna kommer användas för att genom en funktion $\pi(w)$ tilldela en sannolikhet att smittan går från en nod till en annan via en kant med vikten w . Vi ger nu en definition av en Reed Frost-epidemi på en graf:

Definition 3 En Reed Frost-epidemi $\{RF(t), t = 1, 2, \dots\}$ på en viktad graf $G=(V, E, W)$ är en stokastisk process i diskret tid där $RF(t)$ är mängden av individer som är smittade vid tiden t . Om individen v_i är smittad vid tiden t löper varje mottaglig individ v_j som är granne till v_i sannolikheten $\pi(w)$ att bli smittad av v_i vid tiden $t + 1$ där w är vikten på kanten mellan v_i och v_j . En nod v som är smittad vid tiden t slutar från $t+1$ att vara mottaglig för ny smitta och påverkar därefter inte epidemiprocessen. Processen pågår tills dess att inga nya individer smittas

Grafen som beskriver den sociala strukturen kallar vi för en *underliggande graf* och betecknar den G . Utgångsläget för Reed Frost-epidemin som vi beskrev i Definition 1, där alla kan smitta alla, motsvaras av grafen bestående av n noder, en kant mellan varje möjligt par av noder $\{v_i, v_j\}, i \neq j$, och där alla kanter har samma vikt. Grafen blir följaktligen en *komplett graf* och $\pi \equiv \gamma/n$. Antalet kanter i den kompletta underliggande grafen är $\binom{n}{2}$ st.

Vi startar med att smitta en slumpvis vald nod i populationen och vi är ute efter en beskrivning av hur många som blivit smittade då smittan har dött ut och inte sprids längre. Vi konstruerar nu en ny graf där vi utgår från G och vi gör ett slumpförsök med sannolikhet $\pi(w_e)$ att lyckas, för varje kant $e \in E$ för att se om smitta skulle kunna överföras. Eftersom smittöverföringarna är oberoende kan vi utföra slumpförsöken i vilken ordning som helst. Lyckas slumpförsöket så behåller vi kanten, annars tar vi bort den. Grafen vi får kvar är vad vi kallar en *uttunnad graf* och betecknas G_E . Grafen G_E består inte nödvändigtvis av en enda stor komponent utan kan bestå av flera stycken mindre komponenter. Om vi betraktar en nod v i en komponent C så består komponenten C av de noder som skulle smittas om v någon gång smittas. Vi påminner om att man inte

behöver tänka på *smittan* eller *epidemin* som en sjukdom utan att det lika gärna skulle kunna motsvara t.ex. ett datornätverks informationsspridning. Uttunnningen (d.v.s ingen överföring av smittan) svarar då mot att motsvarande kommunikationslänkar slås ut. Med den tolkningen består v 's komponent i den uttunnade grafen, av de noder som v fortfarande kan ha kontakt med i händelse av att vissa länkar slagits ut.

I denna uppsats kommer vi att studera storlekarna på komponenterna i den uttunnade grafen i fallet då den underliggande grafen genereras stokastiskt enligt en generaliserad version av *konfigurationsmodellen*. Med denna modell kommer vi kunna kontrollera nodernas gradfördelning och dessutom kunna ha gradberoende vikter på kanterna.

2.2 Konfigurationsmodellen

Vi ska i detta avsnitt beskriva den modell som vi kommer generera grafer utefter i resten av uppsatsen, nämligen den *viktade konfigurationsmodellen*. Vi börjar dock med att beskriva den enklare *grundläggande konfigurationsmodellen*.

Fixera antalet noder n , låt D_1, \dots, D_n beteckna nodernas grader och låt dessa anta positiva heltal oberoende enligt sannolikhetsfunktionen $p(d)$. Betrakta en nods grad $D = d$ som antalet halvkanter som är kopplade till noden. Koppla ihop halvkanterna slumpmässigt två och två så att de bildar hela kanter. Om antalet halvkanter är udda blir vi tvungna att lägga till eller ta bort en halvkant eftersom varje helkant uppkommer från sammankoppling av två halvkanter. Vi tar också bort eventuella loopar och dubbelkanter, alltså två kanter mellan ett och samma par av noder. Då n är stort inträffar loopar och dubbelkanter sällan och i [2] visas att gradfördelningen förblir densamma även efter borttagning av loopar och dubbelkanter då $n \rightarrow \infty$. Resultatet blir en s.k. enkel graf där varje komponent oftast är av storlek 2 eller större (eftersom D alltid utfaller positivt, är de enda situationerna där en komponent endast består av en enda nod resultat av borttagning av loopar). Vi bestämmer sedan smittsannolikheten π för att två sammankopplade noder ska smitta varandra. Smittsannolikheten är alltså densamma utmed alla kanter i grafen.

Vi utökar nu denna modell så att den genererar en graf med vikter på varje kant. Vi gör som ovan med skillnaden att när vi genererat graderna på noderna så ger vi halvkanterna varsin vikt enligt en heltalsvärd fördelning som för en nod med grad $D = d$ har sannolikhetsfunktion $q(w|d)$. Tilldelning av vikter tillåts alltså bero på gradtalet hos den nod som halvkanterna tillhör. För varje unikt värde av w kopplar vi slumpvis ihop halvkanterna med vikten w två och två så att de bildar hela kanter. Om, för något värde på w , antalet halvkanter med vikt w är ett udda tal, tar vi bort en halvkant med vikt w till en slumpvis vald nod. Vi tar också bort loopar och dubbelkanter. Sannolikheten att den ena noden hos en kant med vikt w överför sin eventuella smitta till den andra bestäms enligt en funktion $\pi(w)$. Notera att tilldelning av vikter har två effekter jämfört med den grundläggande konfigurationsmodellen, dels påverkar den vilka noder (nodgrader) som kan sammankopplas med vilka, dels gör den att smittsannolikheterna till/från en nod beror av vikterna. Vi kommer ge exempel längre fram som illustrerar detta.

2.3 Epidemistorheter

Låt \mathbb{M}_θ vara en modell för att generera en graf G , där θ är en parameter i en eller flera dimensioner och $|G| = n$. Då en Reed Frost-epidemi startats på G genom att en slumpvis utvald nod v smittats, låt Z_n vara antalet noder som smittats då smittspridningen stannat. Som vi såg tidigare är Z_n lika med antalet noder som ligger i samma komponent som v i den uttunnade grafen G_E .

Vi definierar begreppet *stort utbrott* som att en positiv andel av populationen smittats asymptotiskt d.v.s $\frac{Z_n}{n} > 0$ då $n \rightarrow \infty$. Vi definierar den kritiska storheten $R_0 = R_0(\theta)$, *reproduktionstalet*, som en icke-negativ storhet som beror av modellparametrarna där $R_0 > 1 \iff P(\text{stort utbrott}) > 0$.

I Reed Frost-epidemier definieras R_0 mer specifikt med hjälp av förgreningsapproximationer. Vi börjar med att ge en definition av en förgreningsprocess:

Definition 4 *En förgreningsprocess $\{FP(t), t = 1, 2, \dots\}$ är en stokastisk process där en individ v dör efter 1 tidsenhet och då lämnar efter sig X antal nya individer där $X \in F$ med $E[X] = \mu$, $V[X] = \sigma^2$ för någon icke negativ diskret fördelning F . Dessa individer ger i sin tur upphov till nya individer enligt samma fördelning, oberoende av varandra. $FP(t)$ består av alla individer som uppkommit från individerna i $FP(t - 1)$.*

Om vi startar en förgreningsprocess med en st ursprungsindivid och konstruerar en graf där varje individ är en nod och varje kant $\{v_i, v_j\}$ betyder att v_i är avkomma till v_j , blir resultatet ett träd. Ett *träd* inom grafteorin är en graf utan cykler. Från teorin för förgreningsprocesser kommer följande resultat:

$$\begin{aligned} \lim_{t \rightarrow \infty} P(|FP(t)| < \infty) &= r < 1 \text{ då } \mu > 1 \\ \lim_{t \rightarrow \infty} P(|FP(t)| > 0) &= 0 \text{ då } \mu < 1 \end{aligned}$$

Betrakta nu en Reed Frost-epidemi på en graf genererad enligt konfigurationsmodellen (viktad eller ej). Vi väljer en nod v i en komponent C i den uttunnade grafen G_E och utforskar hur smittan sprider ut sig. Följer vi smittvägarna kan vi se resultatet som ett träd från en förgreningsprocess så länge som vi inte hittar några cykler i komponenten. Eftersom konfigurationsmodellen slumpvis parar ihop halvkanter med samma vikt i stora populationer är korta cykler osannolika och vi får då trädliknande strukturer i G_E . Ju längre bort från smittkällan vi kommer desto större blir sannolikheten att vi hittar en cykel och då uppfylls inte kravet om att antalet nya smittfall/individer är oberoende och likafördelade. Förgreningsapproximationen kommer alltså stämma sämre och sämre ju längre tiden går och fler blivit smittade. Om vi ska kunna få ett stort utbrott, $\{\frac{Z_n}{n}\} > 0$, så måste $P(Z_n < \infty) \rightarrow 0$ då $n \rightarrow \infty$. Eftersom en förgreningsprocess antingen dör ut eller exploderar måste vi kräva att $\mu > 1$ för att ett stort utbrott ska vara möjligt, där μ anger förväntat antal nya smittfall som uppkommer från en smittad individ i början av smittskeendet. Detta gör att vi väljer:

$$R_0 := \mu.$$

Vi söker nu en storhet som säger något om andelen smittade i populationen. Om $P(\text{stort utbrott}) = q$, för något $q > 0$, gäller enligt [5] i allmänhet att $\{\frac{Z_n}{n}\}$ konvergerar

mot en tvåpunktsfördelning T där

$$P(T = 0) = 1 - q$$

$$P(T = \tau) = q$$

Utifrån detta resultat definieras τ som andelen smittade i en population givet ett stort utbrott. I termer av vår uttunnade graf G_E är τ den relativa storleken av den största komponenten i G_E asymptotiskt då $n \rightarrow \infty$. För ett rigoröst resonemang kring teorin för epidemistorheterna och strukturen hos grafen G_E , se [3], [4].

2.4 Uttryck för R_0 i konfigurationsmodellen

Vi ska nu skissa på hur man kommer fram till ett uttryck för R_0 i den grundläggande konfigurationsmodellen utan uttunning (d.v.s $\pi \equiv 1$). Vi har gradfördelning D med sannolikhetsfunktion $p(d)$ som vi här tillåter utfalla på hela icke-negativa heltalslinjen och $E[D] = \mu, V[D] = \sigma^2$. Vi börjar med att betrakta startnoden för smittan, kalla den v_1 . Antal avkommor/smittfall X_1 från v_1 kommer att ha samma fördelning som D eftersom v_1 är den första noden och alla dess grannar motsvaras av avkommor. Från v_1 sticker det ut D_1 st halvkanter som leder till nya noder. Om man väljer slumpmässigt bland en mängd med halvkanter är sannolikheten att hitta en halvkant som leder till en nod med grad k proportionell mot $k \cdot p(k) \cdot \pi = k \cdot p(k)$. Vi normerar så att vi får en sannolikhetsfunktion och benämner denna fördelning \tilde{D} med sannolikhetsfunktion $\tilde{p}(k)$:

$$\tilde{p}(k) = \frac{k \cdot p(k)}{\sum_{j=0}^{\infty} j \cdot p(j)} = \frac{k \cdot p(k)}{\mu}$$

Fördelningen för \tilde{D} kallas för *size biased distribution*. Så länge cykler inte existerar så kommer en smittad nod endast ha fått sin smitta från en källa. I termer från teorin för förgreningsprocesser säger vi att av en godtycklig nods v :s d st grannar utgörs en st av v :s mamma och $d - 1$ st av v :s barn. Antalet avkommor kommer alltså att ha samma fördelning som $\tilde{D} - 1$ så länge som vi inte hittar några cykler. Alltså får vi följande uttryck för $R_0 = "E(\text{antal smittfall som uppkommer från en nod})"$:

$$R_0 = E[\tilde{D} - 1] = \sum_{k=0}^{\infty} k \cdot \frac{k \cdot p(k)}{\mu} - 1 = \frac{\sigma^2 + \mu^2}{\mu} - 1 = \mu + \frac{\sigma^2 - \mu}{\mu} \quad (1)$$

När vi hamnar i en situation då vi hittar en cykel stämmer det inte längre att vi har ett träd och då slutar förgreningsapproximationen stämma. Som vi tidigare nämnt gäller förgreningsapproximationen längre ju större n är eftersom chansen att knyta tillbaka till en nod som redan finns i trädet är mindre ju större population vi har att välja bland. Av denna anledning kommer smittskeendet att approximeras av en förgreningsprocess i början av smittskeendet då en förhållandevis liten andel av populationen är smittad.

Man kan på liknande sätt konstruera ett uttryck för R_0 för den viktade konfigurationsmodellen, genom att arbeta med förgreningsprocesser med flera typer (i denna uppsats

är typ lika med gradtal) av individer. Enligt [1] är R_0 det största egenvärdet till matrisen $M = \{M_{dk}\}_{d,k \geq 2}$ med:

$$M_{dk} = (d-1) \sum_w \frac{\pi(w)q(w|d)q(w|k)kp(k)}{\sum_j q(w|j)jp(j)} \quad (2)$$

där d och k anger gradtal och w kantvikter. Här motsvarar elementet M_{dk} förväntat antal avkommor av grad k till en nod av grad d .

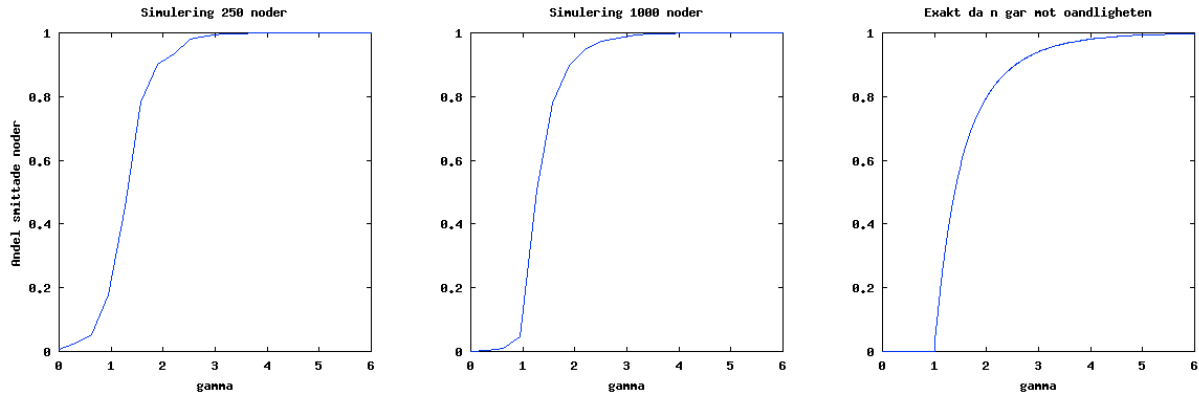
2.5 Från konfigurationsmodellen till τ

Vi kommer att simulera Reed Frost-epidemier på grafer av storleken n . Vi bygger först upp en *underliggande graf* enligt konfigurationsmodellen och tunnar sedan ut den med hjälp av smittsannolikheterna på kanterna till en *uttunnad graf* G_E . Givet en godtyckligt vald nod v är komponenten som v ligger i i G_E de noder som skulle smittas om v smittas. Storleken på denna komponent kallas för Z_n , precis som i avsnittet för epidemistorheter. Genom att approximera början av Reed Frost-epidemin med en förgreningsprocess med intensiteten μ fås $R_0 = \mu$ där $R_0 > 1$ innebär att det finns en positiv sannolikhet för ett stort utbrott. Ett stort utbrott motsvaras i sin tur av en stor komponent i grafen G_E , d.v.s. en komponent som asymptotiskt utgör en positiv andel av hela grafen. I [3] visas att en sådan stor komponent är asymptotiskt unik (d.v.s. det finns högst en). Slutstorleken τ motsvaras av den relativa storleken hos den stora komponenten.

3 Metod

I denna uppsats använder vi ett matlabprogram för att simulera epidemier på grafer efter den viktade konfigurationsmodellen. Slutresultatet av en simulering är den uttunnade grafen G_E . När vi har G_E räknas antalet noder i den största komponenten vilket som vi såg tidigare konvergerar mot storheten τ då $n \rightarrow \infty$. I denna uppsats har vi simulerat med 1000 st noder (om inget annat anges specifikt i exemplet). Typiskt har vi gjort 10 iterationer för varje uppsättning av parametrar för att sedan variera en av parametrarna längs x-axeln och studera medelvärdet av slutstorlekarna i de 10 iterationerna. I Figur 1 illustreras ett exempel på en simulering där gradfördelningen är $Po(\gamma)$ med samma vikt på alla kanter och utan uttunning (vi har alltså simulerat den ursprungliga konfigurationsmodellen med $Po(\gamma)$ -grader). Vi varierar γ på x-axeln. Längst till höger i Figur 1 visas τ uträknat teoretiskt då $n \rightarrow \infty$, som i detta enkla fall visar sig vara den största lösningen till ekvationen $1 - \tau = e^{-\gamma\tau}$. Vi kan också konstatera att $R_0 = \{\text{enligt formel 1}\} = \frac{\gamma + \gamma^2}{\gamma} - 1 = \gamma$. Notera att $Po(\gamma)$ också är den asymptotiska gradfördelningen i en grafmodell med n noder där varje par av noder sammanlänkas oberoende med sannolikhet γ/n . Denna modell är nära relaterad till det inledande skedet av en homogen Reed Frost-epidemi.

I den exakta uträkningen i Figur 1 ser vi att vi får stora komponenter vid $\gamma > 1$. I de två olika simuleringarna blir brytpunkten något otydligare. Epidemierna för $\gamma < 1$ är inte asymptotiskt stora utbrott, men utbrotten får ändå än viss storlek i förhållande till den ändliga populationen vi har i simuleringarna. Även om vi inte i simuleringarna kan



Figur 1: Slutstorlek för stor komponent från graf med gradfördelningen $Po(\gamma)$

se *exakt* när utbrotten övergår till en asymptotiskt positiv andel av populationen kan vi se en brytpunkt kring $\gamma = 1$ i stil med den som vi ser för den exakta uträkningen.

4 Simuleringar av underliggande grafer

I de första simuleringarna undersöks storleken hos de stora komponenterna för olika gradfördelningar och viktfordelningar. Vi kommer i dessa exempel att låta funktionen $\pi(w) \equiv 1$ för alla aktuella w d.v.s vi gör ingen uttunning. Dessa exempel illustrerar hur vi genom olika viktfordelningar kan styra vilka noder som kan kopplas till vilka och hur detta påverkar komponentstorlekarna i grafen. Modellerna som exempel 1a, 3 och 4 bygger på är hämtade från [1].

4.0.1 Exempel 1a

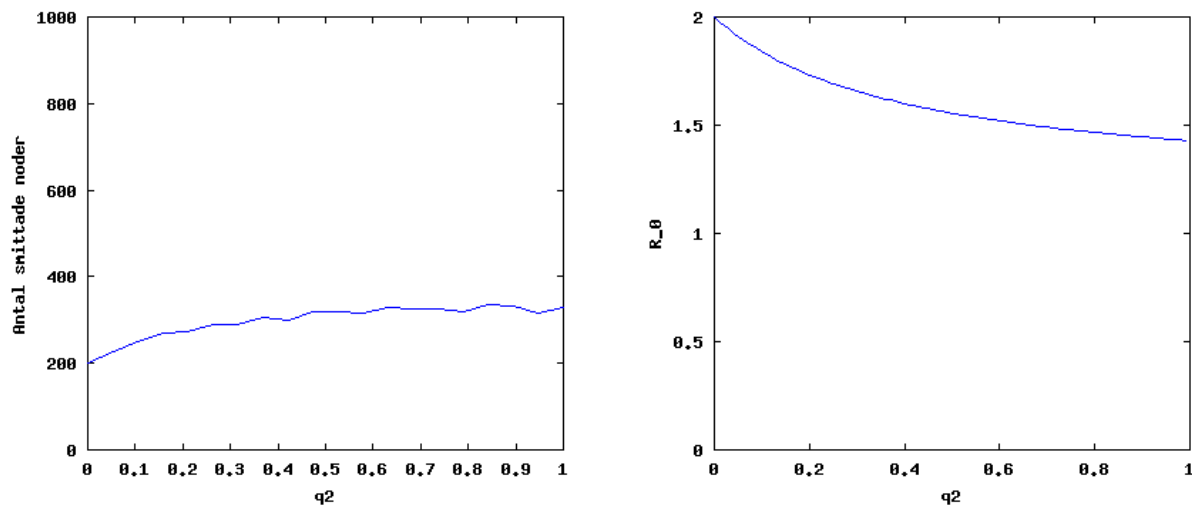
I första simuleringen vars resultat syns i Figur 2 tilldelar vi noderna graden 1 eller graden 3 enligt fördelningen:

$$\begin{aligned} p(1) &= p_1 = 0.8 \\ p(3) &= 1 - p_1 = 0.2 \end{aligned}$$

Vi ger kanterna tre möjliga vikter enligt fördelningen:

$$\begin{aligned} q(w = 1|d = 1) &= 1 - q_1 = 0.7 \\ q(w = 2|d = 1) &= q_1 = 0.3 \\ q(w = 2|d = 3) &= q_2 \\ q(w = 3|d = 3) &= 1 - q_2 \\ \pi(w) &\equiv 1 \end{aligned}$$

Notera att både grad-ett-noderna (kallas här efter för 1-noder) och 3-noderna kan få vikten 2 på sina halvkanter vilket gör att det finns möjlighet för 1-noder att sitta ihop med 3-noder då $q_2 > 0$. I Figur 2, visas en simulering då q_2 varieras. Till höger i figuren ser



Figur 2: Exempel 1a. Gradfördelningen $P(1) = 0.8, P(2) = 0.2$. Viktfördelningen: $q(w = 1|d = 1) = 0.7, q(w = 2|d = 1) = 0.3, q(w = 2|d = 3) = q_2, q(w = 3|d = 3) = 1 - q_2$

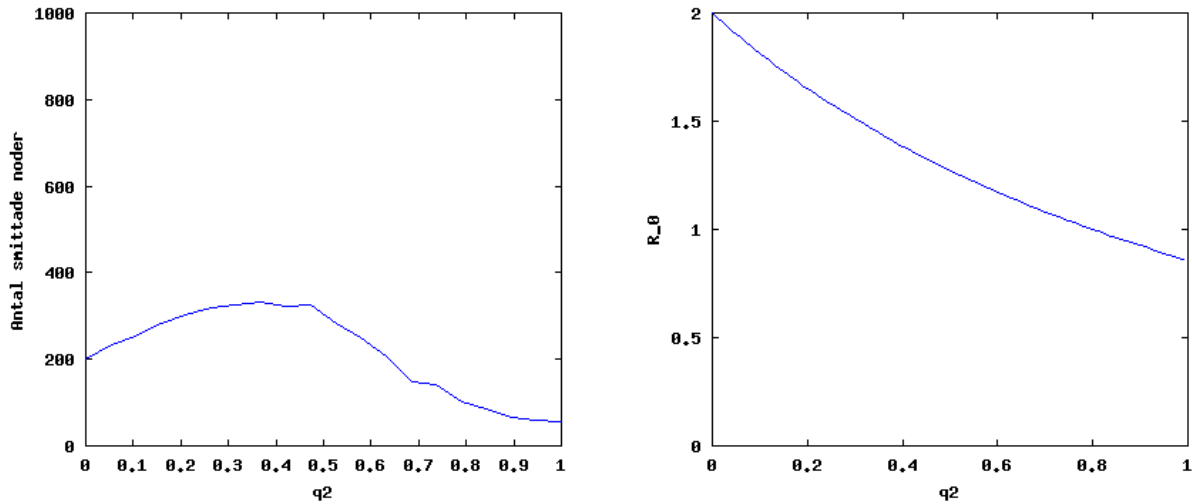
vi värden på R_0 uträknat enligt (2). Vi kan observera att trots att *reproduktionstalet* R_0 minskar så ökar storlekarna på de största komponenterna i smittgraferna. Detta kan förstås genom följande resonemang. Då $q_2 = 0$ kommer ingen 1-nod sitta ihop med någon 3-nod. Alltså kommer 1-noderna sitta ihop två och två och 3-noderna kommer sitta ihop i stora klumpar, sannolikt kommer alla 3-noder ligga i samma komponent. Då vi ökar q_2 kommer 1-noderna att successivt blandas in i jättekomponenten med 3-noder så att den blir större.

4.0.2 Exempel 1b

Vi ändrar värdet på q_1 till $q_1 = 1$ så att uppställningen blir:

$$\begin{aligned}
 p(1) &= p_1 = 0.8 \\
 p(3) &= 1 - p_1 = 0.2 \\
 q(w = 1|d = 1) &= 1 - q_1 = 0 \\
 q(w = 2|d = 1) &= q_1 = 1 \\
 q(w = 2|d = 3) &= q_2 \\
 q(w = 3|d = 3) &= 1 - q_2 \\
 \pi(w) &\equiv 1
 \end{aligned}$$

I Figur 3 visas simulering med 20 iterationer på varje q_2 -värde. Här ser vi att jättekomponenten blir större fram tills $q_2 = 0.5$ för att därefter minska i storlek. I detta exempel har vi ökat på q_1 för att på detta sätt ytterligare blanda 1-noder med 3-noder. Vad vi då ser i Figur 3 efter att $q_2 = 0.5$ är att det blandas in så många 1-noder att de bryter sönder jättekomponenten, alltså 1-noder fångar upp 3-noder ur jättekomponenten för att bilda mindre komponenter. Observera att $q_1 = 1, q_2 = 1$ innebär att alla halvkanter får vikten 2 vilket är ekvivalent med att man genererat de underliggande graferna med den grundläggande konfigurationsmodellen.



Figur 3: Exempel 1b. Gradfördelningen $P(1) = 0.8, P(2) = 0.2$. Viktfördelningen: $q(w = 1|d = 1) = 0, q(w = 2|d = 1) = 1, q(w = 2|d = 3) = q_2, q(w = 3|d = 3) = 1 - q_2$

Figur 14 i appendix visar en skattning av variationskoefficienterna $\frac{\sigma}{\mu}$ för de olika värdena på q_2 med parametrarna ovan. Vi kan se att utfallet av epidemin blir mer oviss ju närmare 1 som R_0 ligger. Andelarna av de olika noderna i den stora komponenten visas i Figur 15.

4.1 Exempel 2

I detta exempel visar vi en annan variant där vi med parametern α bestämmer om vikterna fördelas beroende på graden hos de noder halvkanterna sitter fast i eller helt oberoende. Medan vi i förra exemplet ändrade andelen av de olika vikterna kommer vi i detta exempel använda oss av en viktfördelning som hela tiden håller fördelningen för den totala andelen av de olika vikterna konstant, men ändrar beroendet steglöst. Vi börjar med att välja gradfördelningen:

$$\begin{aligned} p(1) &= p_1 \\ p(2) &= p_2 \\ &\dots \\ p(m) &= p_m \end{aligned}$$

Vi väljer α så att då $\alpha = 0$ kommer varje halvkant som sitter fast i en k -nod få samma tal k som vikt. Då $\alpha = 0$ kommer därför sannolikheten för att en slumpmässigt vald halvkant har vikt w vara

$$\tilde{p}(w) = \frac{w \cdot p(w)}{\sum_{j=0}^{\infty} j \cdot p(j)}$$

vilket är samma fördelning som vi nämnde i avsnittet om R_0 , fördelningen kallas för *size biased distribution*. Vi vill att fördelningen för de totala andelarna av de olika vikterna ska vara konstant och när vi ändrar α ska endast graden av beroende ändras. Detta får

oss att välja viktfordelningen på följande vis:

$$q(w|d) = \begin{cases} (1 - \alpha) + \alpha \cdot \tilde{p}(w) & ,w = d \\ \alpha \cdot \tilde{p}(w) & ,w \neq d \end{cases}$$

I appendix visas att andelen av de olika vikterna har samma fördelning oberoende av α . Även i dessa exempel är $\pi \equiv 1$.

4.1.1 Exempel 2a

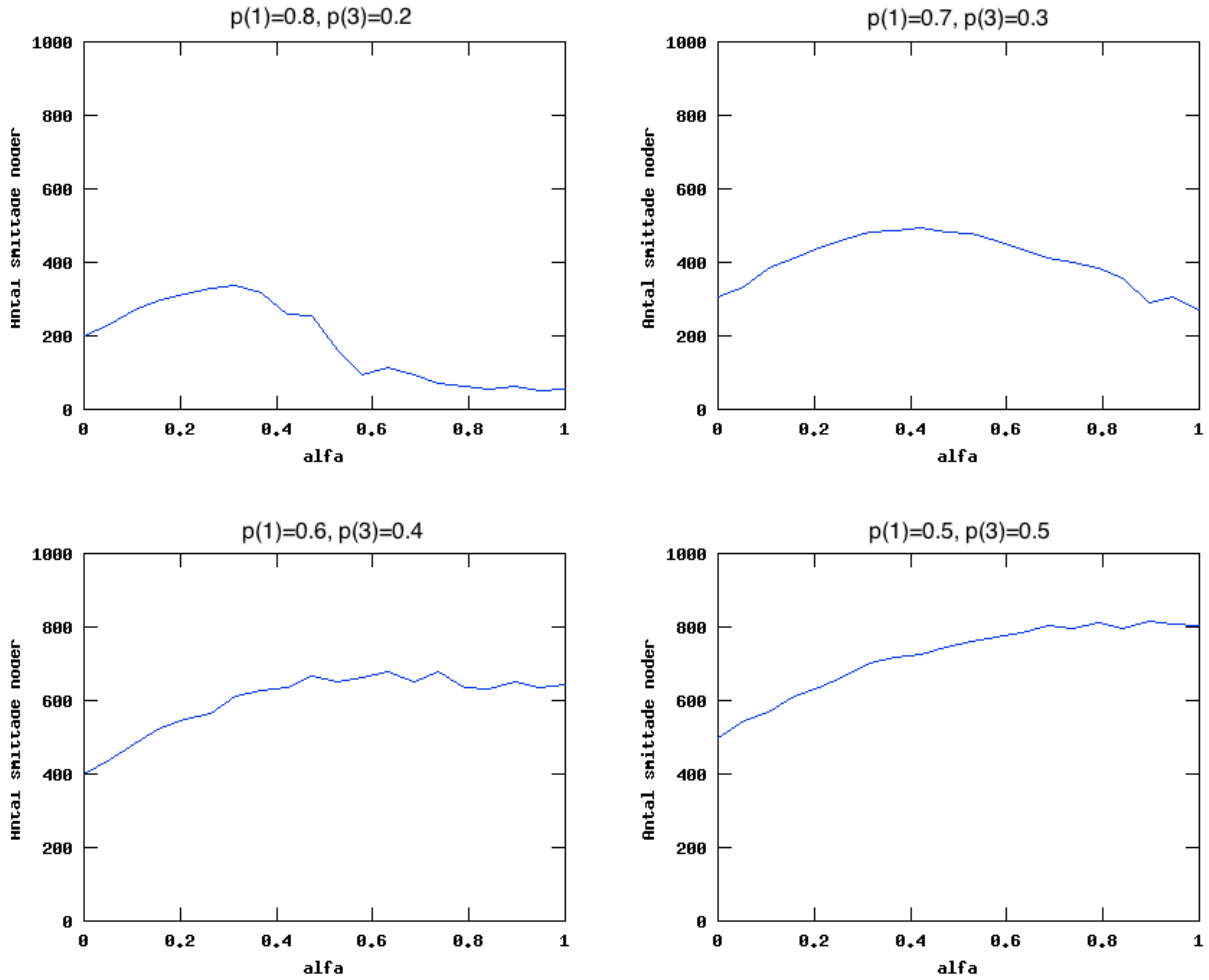
I Figur 4 är simuleringarna gjorda med två möjliga gradtal (1 och 3) och därmed två möjliga vikter. I Figur 5 visas R_0 -värdena. Här är alltså fördelningen för andelarna av 3-noder och 1-noderna konstant samt fördelningen för andelarna av kanterna med vikt 3 och kanterna med vikt 1 är konstant genom alla simuleringar. Då $\alpha = 0$ sitter alla 1-viktkanterna mellan 1-noder och alla 3-viktkanterna mellan 3-noder. Detta gör att 3-noderna bildar en egen stor komponent och 1-noderna sätter ihop sig i par. Då α ökar hamnar istället en del kanter mellan 1-noder och 3-noder vilket kan påverka den maximala komponentstorleken på två sätt. Antingen ansluter ett förhållandevis litet antal 1-noder till 3-nodskomponenten vilket gör denna större om inga 1-noder ansluter. Men som vi ser i de övre plottarna så bryts jättekomponenten sönder då ett förhållandevis stort antal 1-noder ansluter. Effekten är densamma som i Exempel 1 men med skillnaden att andelarna av de olika kantvikterna hålls konstanta och bara ändrar kopplingsmönster med hjälp av α . Andelarna i den största komponenten kan utläsas av Figur 16.

4.1.2 Exempel 2b

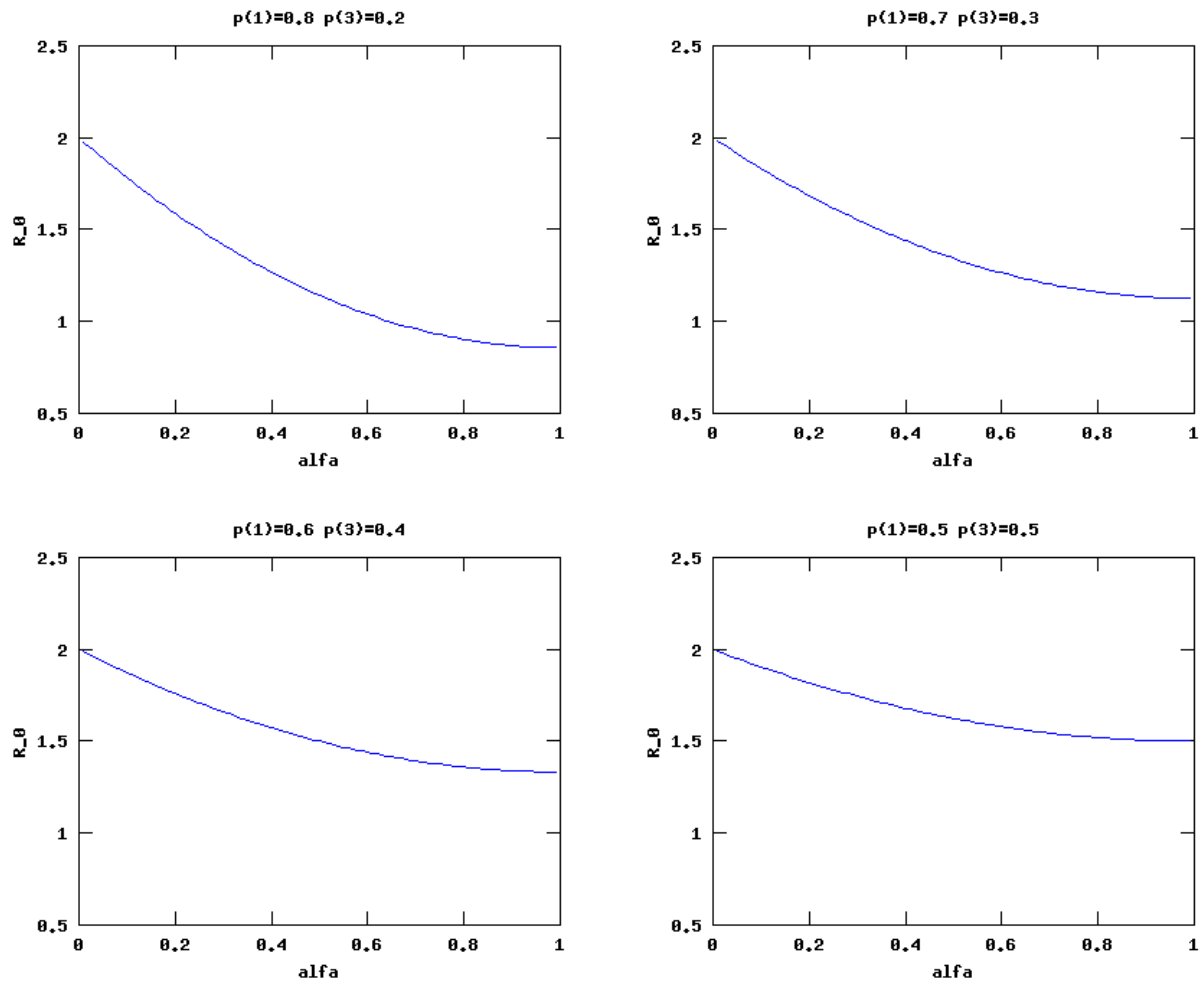
I Figur 6 är simuleringarna gjorda med fyra möjliga gradtal (1, 2, 3 och 4) och därmed fyra möjliga vikter. I Figur 7 visas R_0 för samma val av fördelningar. I de fyra olika plottarna flyttar vi succesivt sannolikhetsmassan i fördelningen mer mot de högre graderna. Effekten liknar den som vi ser i 4.1.1. Andelarna av de olika nodgraderna i den stora komponenten återfinns i Figur 17. Då $\alpha = 0$ är det endast möjligt för noder av samma grad att kopplas till varandra, vilket innebär att grafen egentligen har sönderfallit i fyra stycken separata komponenter. I detta skede ser vi från andelarna i Figur 17 att den stora komponenten utgörs av komponenten med endast 2-noder, förutom i fallet då vi har gradfördelningen $p(1) = 0.4, p(2) = 0.3, p(3) = 0.2, p(4) = 0.1$ då vi ser att 3-nodskomponenten blir större än 2-nodskomponenten i vissa iterationer. Situationen ändras genast då $\alpha > 0$ då det kan förekomma kanter som binder samman de olika komponenterna, vilket syns som en snabb ökning i slutstorlekarna i Figur 6.

5 Simulering av epidemier

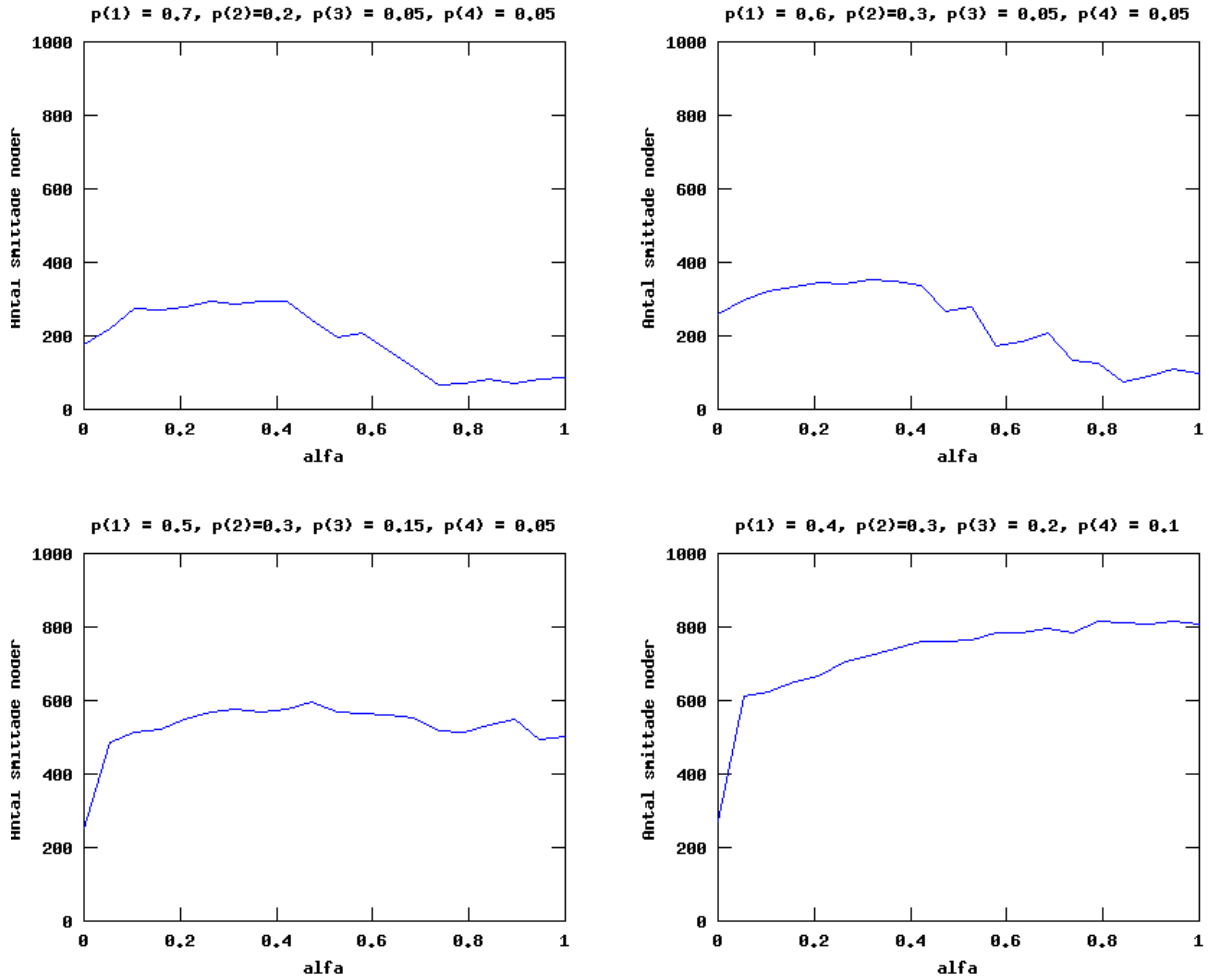
I detta avsnitt simulerar vi underliggande grafer och dessutom en Reed Frost-epidemi på densamma genom att tunna ut den underliggande grafen som vi beskrivit tidigare.



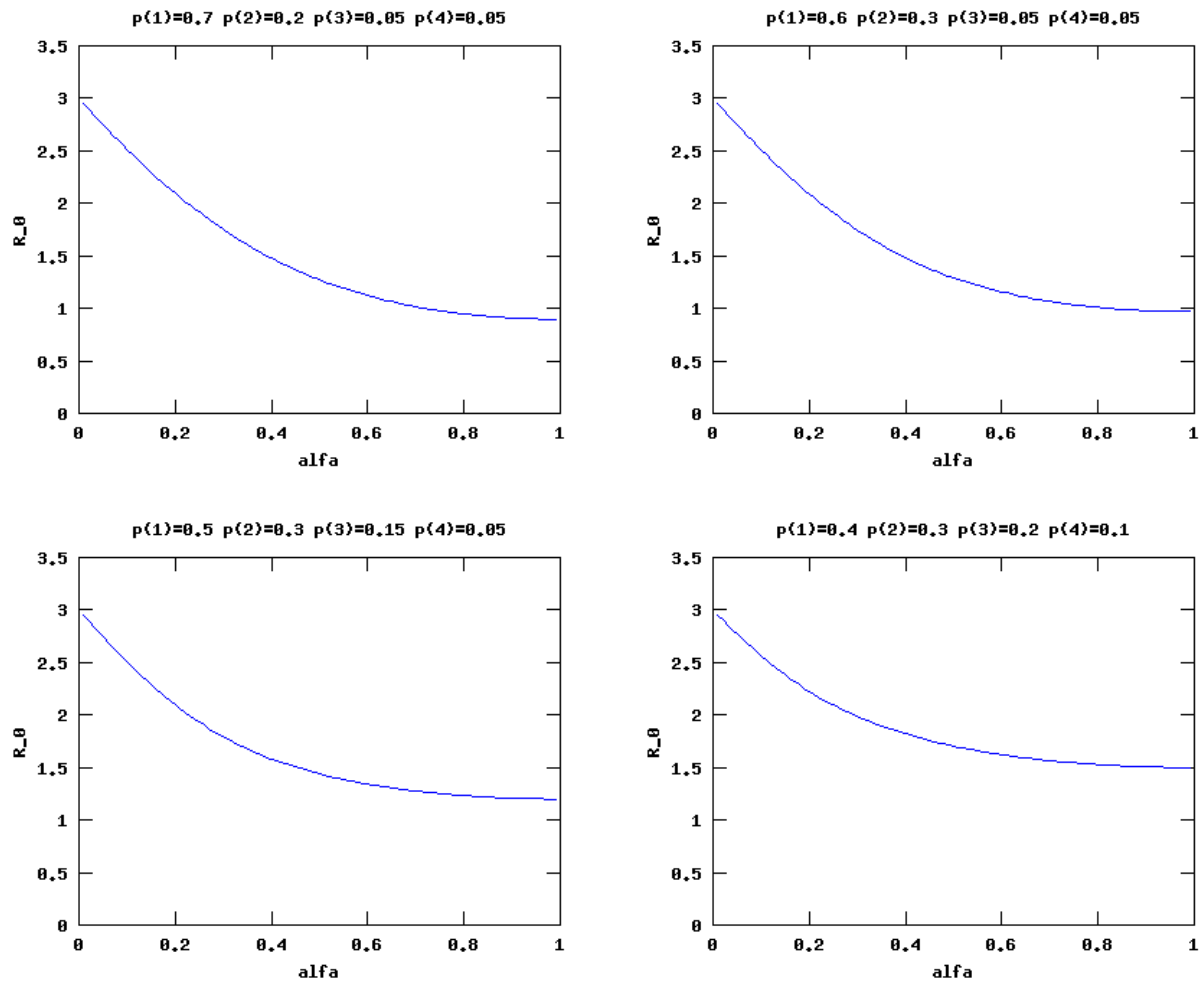
Figur 4: Exempel 2a slutstorlekarna. Gradfördelningen står över varje plott. Viktfördelningen är: $q(w|d) = (1 - \alpha) + \alpha \cdot \tilde{p}(w)$ för $w = d$ och $q(w|d) = \alpha \cdot \tilde{p}(w)$ för $w \neq d$.



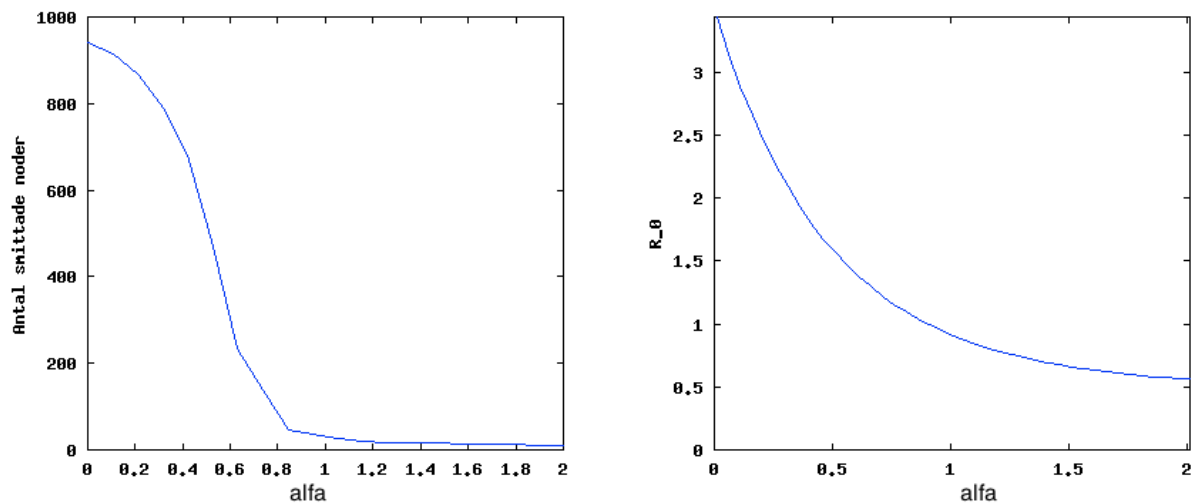
Figur 5: Exempel 2a R_0 . Gradfördelningen står över varje plott. Viktfördelningen är: $q(w|d) = (1 - \alpha) + \alpha \cdot \tilde{p}(w)$ för $w = d$ och $q(w|d) = \alpha \cdot \tilde{p}(w)$ för $w \neq d$.



Figur 6: Exempel 2b slutstorlekarna. Gradfördelningen står över varje plott. Viktfördelningen är: $q(w|d) = (1 - \alpha) + \alpha \cdot \tilde{p}(w)$ för $w = d$ och $q(w|d) = \alpha \cdot \tilde{p}(w)$ för $w \neq d$.



Figur 7: Exempel 2b R_0 . Gradfördelningen står över varje plott. Viktfördelningen är: $q(w|d) = (1 - \alpha) + \alpha \cdot \tilde{p}(w)$ för $w = d$ och $q(w|d) = \alpha \cdot \tilde{p}(w)$ för $w \neq d$.



Figur 8: Exempel 3a. Gradfördelningen: $D \in Po(4)$ betingat på att $D \in \{1, 2, \dots, 200\}$. Viktfördelningen: $q(w = 1|d) = 1 - d^{-\alpha}$, $q(w = 2|d) = d^{-\alpha}$, Smittsannolikhetsfunktionen: $\pi(1) = 0.1, \pi(2) = 0.7$.

5.1 Exempel 3

Vi väljer här gradfördelningen till $D \in Po(4)$ betingat på att $D \in \{1, 2, \dots, 200\}$. Vi har två möjliga vikter med fördelningen:

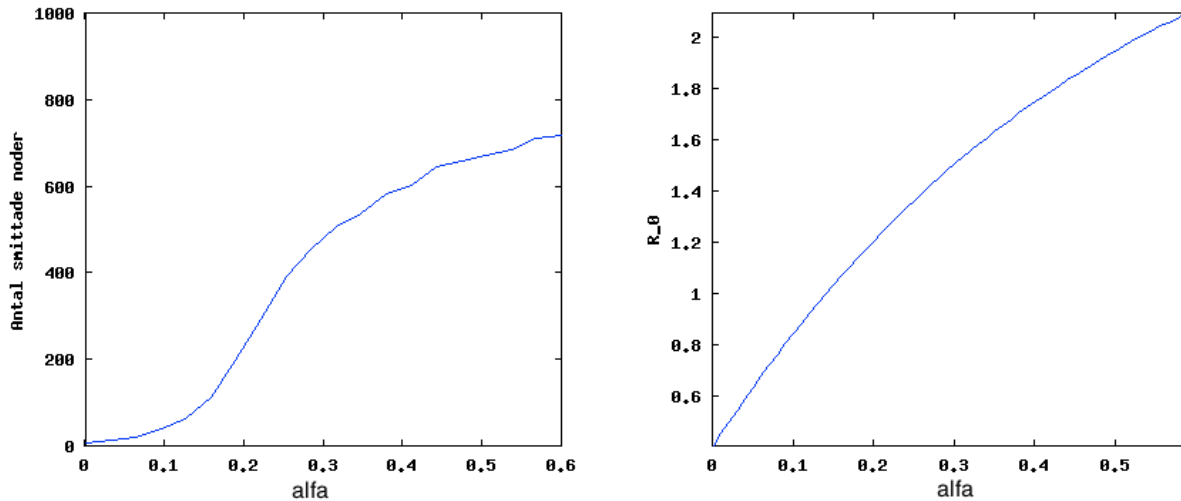
$$\begin{aligned} q(w = 1|d) &= 1 - d^{-\alpha} \\ q(w = 2|d) &= d^{-\alpha} \end{aligned}$$

för $\alpha > 0$. Fördelningen med vilken vi fördelar vikter till halvkanter tenderar att tilldela halvkanter hos låggradiga noder vikt 2 och halvkanter hos höggradiga noder vikt 1. I Figur 18 (se appendix) visas funktionen $q(w = 1|d) = 1 - d^{-\alpha}$ för några olika α . Observera att $q(w = 2|d = 1) = 1$ för alla $\alpha > 0$ och färre halvkanter hos höggradiga noder tenderar att få vikt 2 ju större α är. Vi väljer $\pi(1) = 0.1, \pi(2) = 0.7$ vilket tillsammans med viktfordelningen innebär att kanterna hos de låggradiga noderna tenderar att ha högre smittsannolikheter än kanterna hos de höggradiga. Ju större α är desto färre höggradiga noders kanter tenderar att överföra smitta vilket leder till att epidemierna blir mindre ju större α är, se Figur 8.

I Figur 9 är situationen omvänd till $\pi(1) = 0.7, \pi(2) = 0.1$ vilket alltså innebär att kanter hos låggradiga noder tenderar att sprida smitta mindre än kanter hos höggradiga. Detta får till effekt att då fler höggradiga får vikt 2 sprids smittan lättare och epidemierna blir större då α ökar.

5.2 Exempel 4

I detta exempel väljer vi gradfördelningen till $Po(8)$ betingat på att den utfaller positivt och viktfordelningen $Po(8/d)$ där d är gradtalet för noden halvkanter sitter fast i. Lägga märke till att för en godtyckligt vald nod är väntevärdet av summan av dess vikter konstant. Smittsannolikheten sätts till $\pi(w) = 1 - (1 - s)^{w+1}$. Denna π -funktion uppkommer



Figur 9: Exempel 3b. Gradfördelningen: $D \in Po(4)$ betingat på att $D \in \{1, 2, \dots, 200\}$. Viktfördelningen: $q(w = 1|d) = 1 - d^{-\alpha}$, $q(w = 2|d) = d^{-\alpha}$, Smittsannolikhetsfunktionen: $\pi(1) = 0.7$, $\pi(2) = 0.1$.

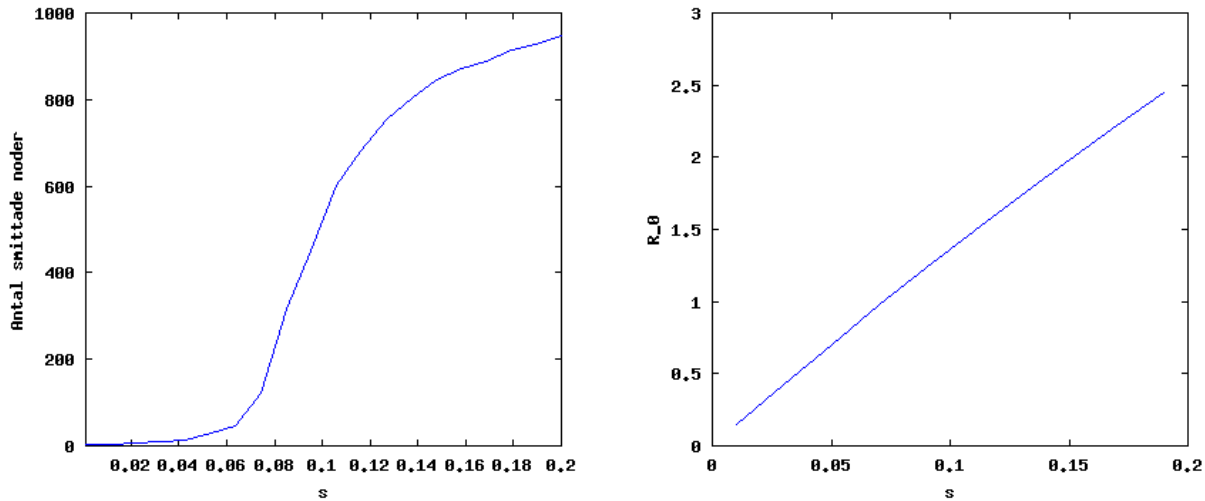
naturligt då vi ser talet $k := w + 1$ som antalet smittförsök mellan två noder som binds ihop av kanten med vikten w och talet s som sannolikheten för smitta för varje sådant försök.

I Figur 10 visas slutstorlekarna och R_0 . Vi ser att efter att s ökat så mycket att ett stort utbrott är möjligt sker en ökning av komponentstorlekarna med lika tydlig brytpunkt som den i Figur 1.

5.3 Exempel 5

Vi utvidgar idén i Exempel 2 där vi med hjälp av en parameter α bestämde huruvida noderna endast kunde sitta ihop med noder av samma grad eller om vi även tillät andra kombinationer. Nu väljer vi fördelningen $Po(10.5)$ betingat på att $D \in \{1, 2, \dots, 30\}$. Vi delar in nodernas grad i tre grupper och tillåter endast kanterna att gå mellan noder i samma grupp då $\alpha = 0$. Grupperna är gjorda så att noderna med små gradtal hamnar i grupp 1, noderna med typiska gradtal hamnar i grupp 2 och de med stora gradtal hamnar i grupp 3. Vi kommer med hjälp av en parameter β att justera vilka kanter som kommer ha högst smittsannolikhet. Vi låter noderna med grad 1 till 8 vara i grupp 1, noderna med grad 9, 10, 11 vara i grupp 2 och noderna med grad 12 till 30 vara i grupp 3. Indelningen kan beskrivas som att vi definierar en funktion $h(d)$ som tar ett gradtal som argument och vars funktionsvärde är numret på gruppen som graden tillhör:

$$h(d) = \begin{cases} 1, & d \in \{1, 2, \dots, 8\} \\ 2, & d \in \{9, 10, 11\} \\ 3, & d \in \{12, 13, \dots, 30\} \end{cases}$$



Figur 10: Exempel 4. Gradfördelningen: $Po(8)$ betingat på att den utfaller positivt. Viktfördelningen $Po(8/d)$. Smittsannolikhetsfunktionen: $\pi(w) = 1 - (1 - s)^{w+1}$.

Uppdelningen ovan är gjord så att andelen noder i de olika grupperna ska vara ungefär lika stora:

$$\begin{aligned} p(1) + \dots + p(8) &\approx 0.28 \\ p(9) + p(10) + p(11) &\approx 0.36 \\ p(12) + \dots + p(30) &\approx 0.36 \end{aligned}$$

Precis som i Exempel 2 kommer $\alpha = 0$ innebära att vi tilldelar vikterna helt deterministiskt efter graden av noden halvkanterna sitter fast i. Vi tilldelar vikten $h(d)$ till alla halvkanter från noder av grad d och de totala andelarna av de olika vikterna kommer då vara:

$$\tilde{p}^h(w) = \frac{\sum_{\{d; h(d)=w\}} d \cdot p(d)}{\sum j \cdot p(j)}.$$

Utifrån $\tilde{p}^h(w)$, en variant på *the size biased distribution*, bestämmer vi viktfordelningen som:

$$q(w|d) = \begin{cases} (1 - \alpha) + \alpha \cdot \tilde{p}^h(w), & h(d) = w \\ \alpha \cdot \tilde{p}^h(w), & h(d) \neq w \end{cases}$$

Man kan se detta upplägg som att vi delar in en population i tre stycken grupper där individerna endast har kontakt inom gruppen då $\alpha = 0$ men vi kan blanda upp dessa genom att öka α . Vi räknar också ut sannolikheterna för de tre vikterna:

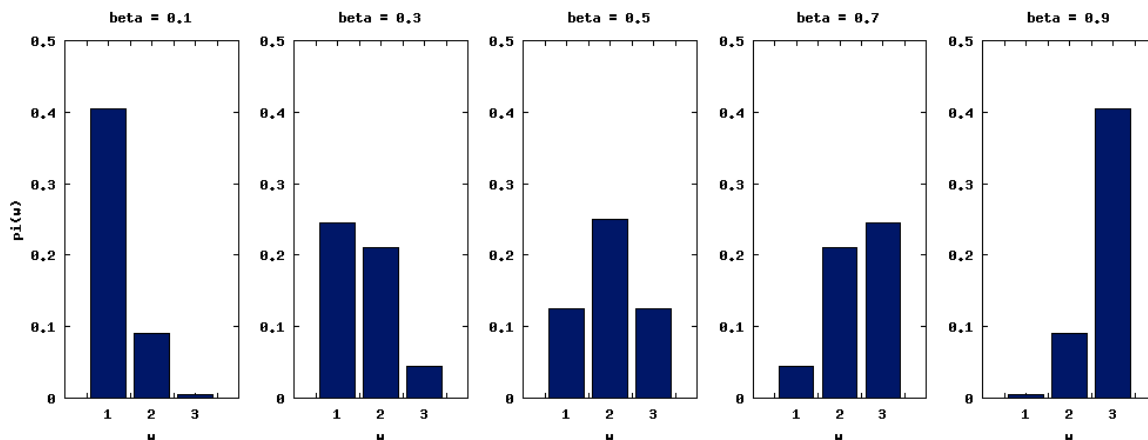
$$\begin{aligned} \tilde{p}^h(1) &\approx 0.18 \\ \tilde{p}^h(2) &\approx 0.34 \\ \tilde{p}^h(3) &\approx 0.48 \end{aligned}$$

Dessa sannolikheter kommer (analogt med Exempel 2) vara konstanta för alla α .

Som π -funktion väljer vi:

$$\pi(w) = c \cdot \binom{2}{w-1} \beta^{w-1} \cdot (1-\beta)^{2-(w-1)} \text{ för } w = 1, 2, 3$$

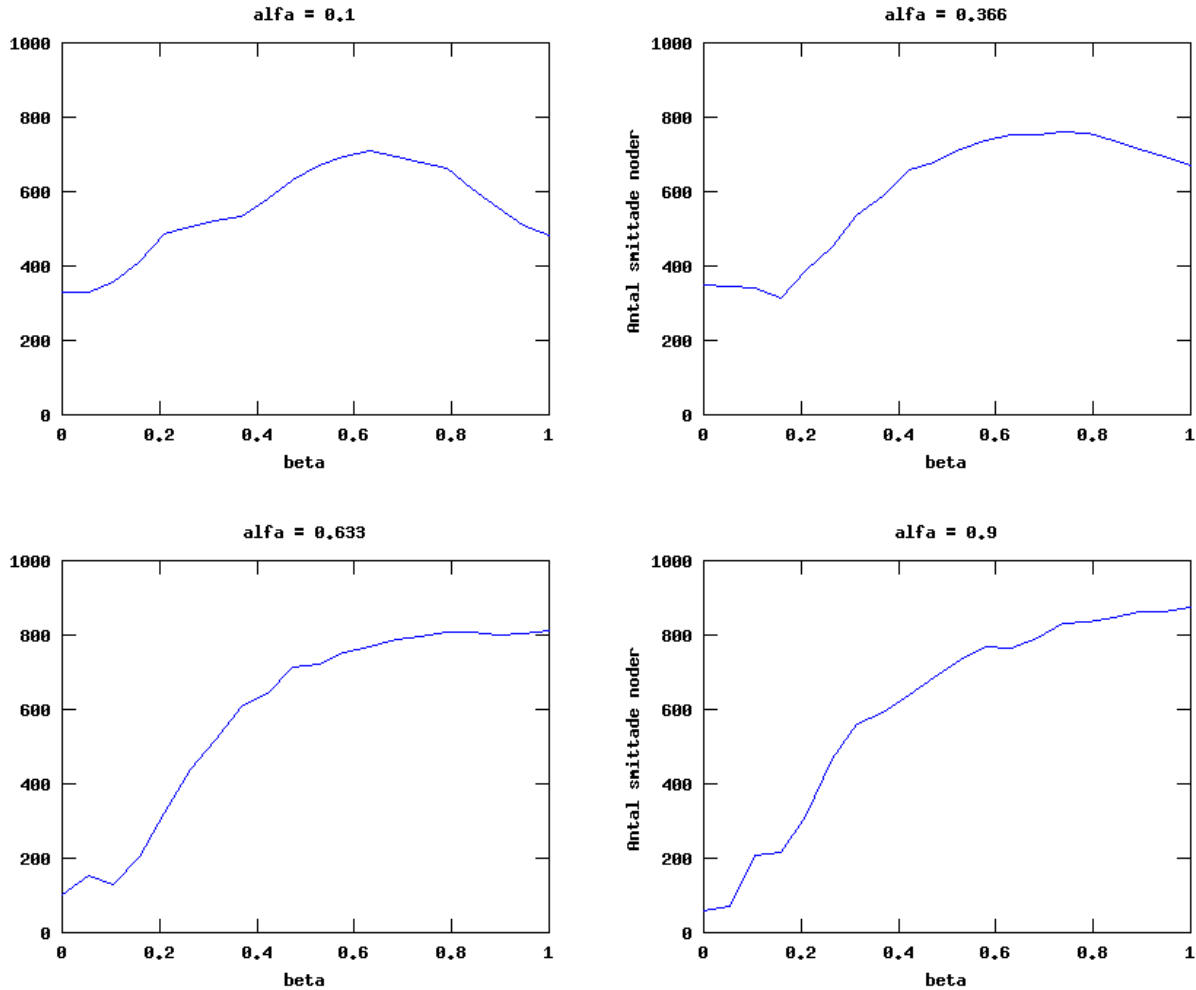
Observera att $\sum_{w=1}^3 \pi(w) = c$ för alla $\beta \in [0, 1]$. Funktionen $\pi(w)$ är vald så att vi med hjälp av parametern β kan förskjuta tunnningssannolikheten mellan de tre olika vikterna. I Figur 11 visas hur förskjutningen av uttunningsannolikheten sker för $c = 0.5$. I Figur 12



Figur 11: Exempel 5. Smittsannolikhetsfunktionen: $\pi(w) = c \cdot \binom{2}{w-1} \beta^{w-1} \cdot (1-\beta)^{2-(w-1)}$ för $w = 1, 2, 3$. β används för att ändra vilken vikt som ska ha störst smittsannolikhet.

har vi låtit β variera på x-axlarna, α växla mellan plottarna och c vara konstant $c = 0.5$. I Figur 13 visas R_0 .

Då $\alpha = 0.1$ är de från början vattentäta skotten mellan grupperna något upplösta och vissa kanter kommer gå mellan de tre grupperna. Då β är nära ändpunkterna 0 eller 1 kommer grupp 1 respektive grupp 3 att tunnas ut så pass mycket att grupp 1 respektive grupp 3 mer eller mindre upplöses vilket gör att de inte kommer utgöra en betydande del i den stora komponenten. Däremot då β är runt 0.5 kommer alla grupper tunnas ut ungefär lika mycket och eftersom kanter mellan grupperna är möjliga kan de sitta ihop och tillsammans bilda en stor komponent. Ju större α desto mindre tenderar vi att få en formation av tre grupper. Istället har vi tre vikter som kan sitta mellan vilket par av noder som helst. Då α är stort och β litet kommer kanterna med vikt 1 vara de som troligast överför smitta men de är också minst till antalet. Det stora antalet kanter med vikt 2 och 3 med betydligt mindre smittsannolikhet, verkar som kilar snarare än kopplingar vilket bryter upp komponenterna till småbitar. Här ser vi att höggradiga noder, som annars är en bidragande faktor till att en epidemi kan ta fart, tas ut av den låga smittsannolikheten. Då β ökar förskjuts smittsannolikheten så att de kanter som är i majoritet är de som mest sannolikt överför smitta och komponenterna blir stora.

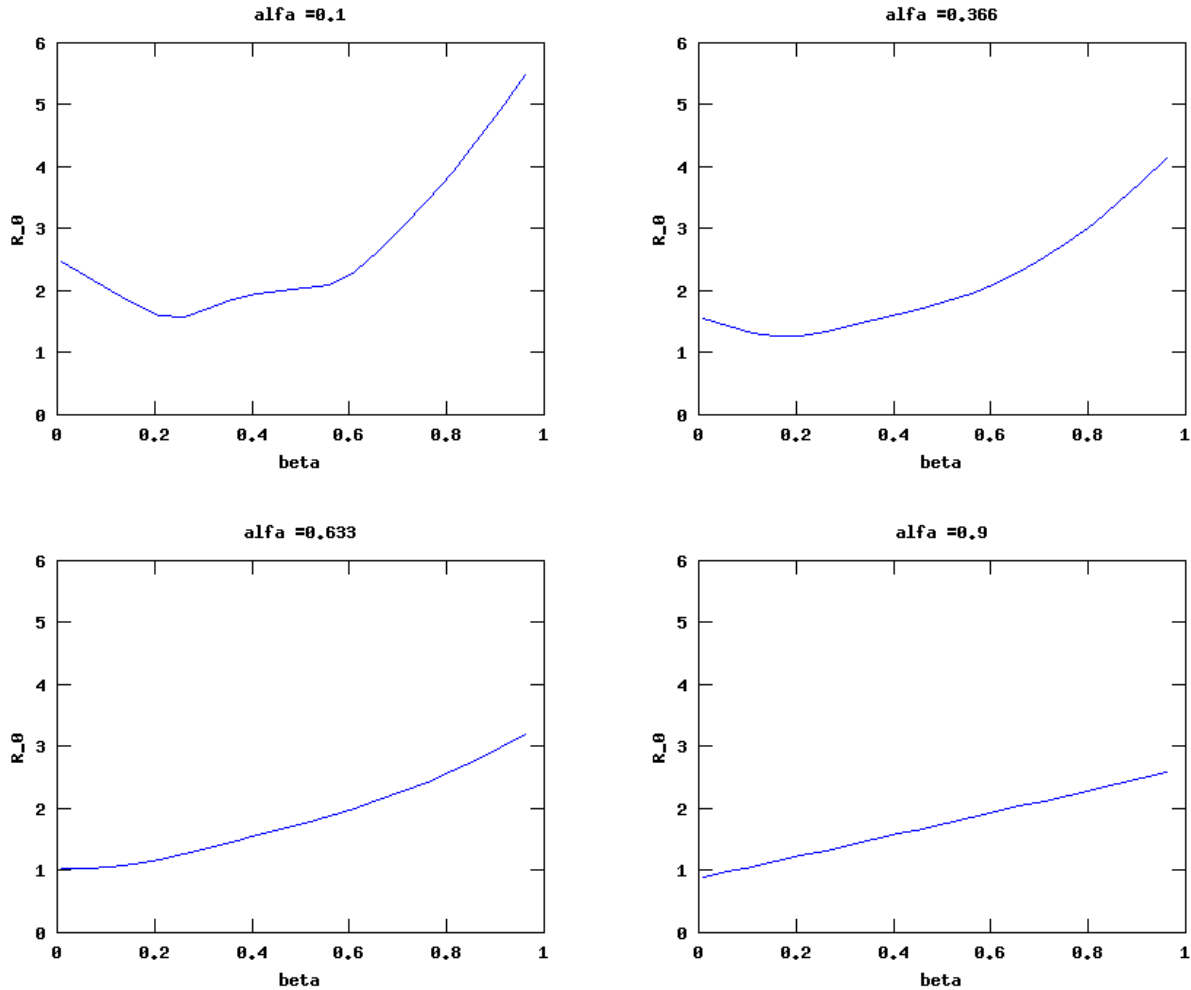


Figur 12: Exempel 5 slutstorlekarna. Gradfördelning: $Po(10.5)$ betingat på att $D \in \{1, 2, \dots, 30\}$. Viktfördelning: $q(w|d) = (1 - \alpha) + \alpha \cdot \tilde{p}^h(w)$ för $h(d) = w$, $q(w|d) = \alpha \cdot \tilde{p}^h(w)$ för $h(d) \neq w$. Smittsannolikhetsfunktion: $\pi(w) = c \cdot \binom{2}{w-1} \beta^{w-1} \cdot (1 - \beta)^{2-(w-1)}$ för $w = 1, 2, 3$. $c = 0.5$.

6 Diskussion

Vi har i denna uppsats främst sett att det inte generellt föreligger något närmare samband mellan slutstorlekarna och storheten R_0 än att $R_0 > 1 \implies P(\text{stort utbrott}) > 0$, vilket är själva definitionen av R_0 . Vi har sett exempel där R_0 ökar då τ , slutstorleken givet ett stort utbrott, minskar och tvärtom. I en graf med många höggradiga noder har en smitta lättare att ta fart än i en graf med lägre grad på noderna. Vi har sett att man kan kompensera en graf med höga grader genom att ge dessa noder kanter med låg smittsannolikhet.

Utvidgningen från den grundläggande konfigurationsmodellen till den viktade konfigurationsmodellen innebär att man kan införa vikter på kanterna så att noder med lika grad tenderar att sitta ihop i större utsträckning. Det vore intressant att studera en ännu något mer utvidgad tillämpning av den viktade konfigurationsmodellen där varje nod inte enbart karaktäriseras av dess grad. Man skulle kunna skilja på noder av samma grad



Figur 13: Exempel 5 R_0 . Gradfördelning: $Po(10.5)$ betingat på att $D \in \{1, 2, \dots, 30\}$. Viktfördelning: $q(w|d) = (1 - \alpha) + \alpha \cdot \tilde{p}^h(w)$ för $h(d) = w$, $q(w|d) = \alpha \cdot \tilde{p}^h(w)$ för $h(d) \neq w$. Smittsannolikhetsfunktion: $\pi(w) = c \cdot \binom{2}{w-1} \beta^{w-1} \cdot (1 - \beta)^{2-(w-1)}$ för $w = 1, 2, 3$. $c = 0.5$.

genom att dela in dem i olika typer. Dessa olika typer kan ha tendenser att få olika vikter på sina halvkanter. Detta kan i sin tur leda till att nya varianter av grupperingar kan uppkomma. Matematiskt innebär det att vi skapar typerna t_1, \dots, t_a . Vi lägger till typfördelningen T till gradfördelningen D . Efter att vi tilldelat en grad till varje nod tilldelar vi en av de a st typerna till varje nod där fördelningen vi använder för att tilldela typerna till noderna kan bero på graderna. Vi byter ut sannolikhetsfördelningen för vikterna $q(w|d) = P(W = w|D = d)$ mot $q(w|t) = P(W = w|T = t)$. I termer av denna modell är den tidigare modellen, som vi studerat denna uppsats, det specialfall där T är en bijektiv funktion av D . I det mer generella fallet kan vi låta T och D vara mer eller mindre beroende av varandra. I och med att vi inför typer är det inte längre säkert att två noder av samma grad får samma vikter tilldelade till sina halvkanter och därmed tillåts nya kopplingsmönster uppstå.

En annan modifikation man kan tänka sig införa är att då vi tilldelar halvkanterna sina vikter så ger vi dessutom varje halvkant antingen beskaffenheten *högerhalvkant* eller

vänsterhalvkant, också beroende på nodens grad. När vi sedan parar ihop halvkanterna till helkanter väljs inte två halvkantar av samma vikt helt slumpmässigt, utan vi väljer istället en högerhalvkant och en vänsterhalvkant med den aktuella vikten och parar ihop. Vi bör då kräva att för varje w gäller att fördelningen för antalet högerhalvkantar med vikt w har samma väntevärde som fördelningen för antalet vänsterhalvkantar med med vik w . Med detta tillägg kan vi konstruera riktade grafer där t.ex. riktningen är från vänsterhalvkant till högerhalvkant. Man kan även komma ifrån att lika noder tenderar att kopplas till varandra. T.ex. kan man skapa 1-noder som aldrig kopplas till andra 1-noder utan endast kopplas till höggradiga noder, vilket liknar strukturen hos datornätverk där klienter aldrig är kopplade direkt till varandra utan där de kommunicerar via hubbar.

Referenser

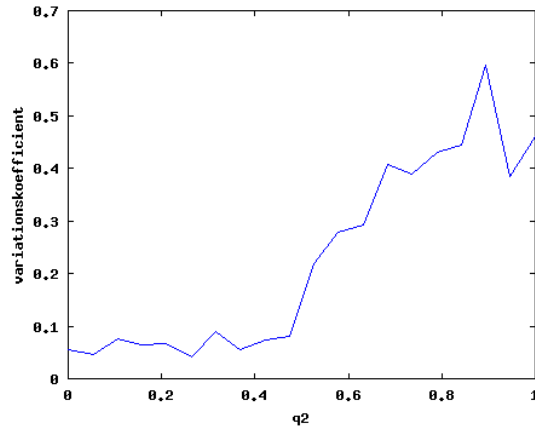
- [1] Britton, T., Deijfen, M., Liljeros, F. (2011): A weighted configuration model and inhomogeneous epidemics, *Journal of Statistical Physics* **145**, 1268-1384.
- [2] Britton, T., Deijfen, M., Martin-Löf, A. (2006): Generating simple random graphs with prescribed degree distribution, *Journal of Statistical Physics* **124**, 1377-1397.
- [3] Molloy, M., Reed, B. (2000): A critical point for random graphs with a given degree sequence, *Rand. Struct. Alg.* **6**, 161-179.
- [4] Molloy, M., Reed, B. (2000): The size of the giant component of a random graphs with a given degree sequence, *Combinatorics Probability Computing* **7**, 295-305.
- [5] Scalia-Tomba, G.(1985): Asymptotic final size distribution for sum chain binomial processes, *Adv. Appl. Prob.* **17** 477-495.

A Härledning

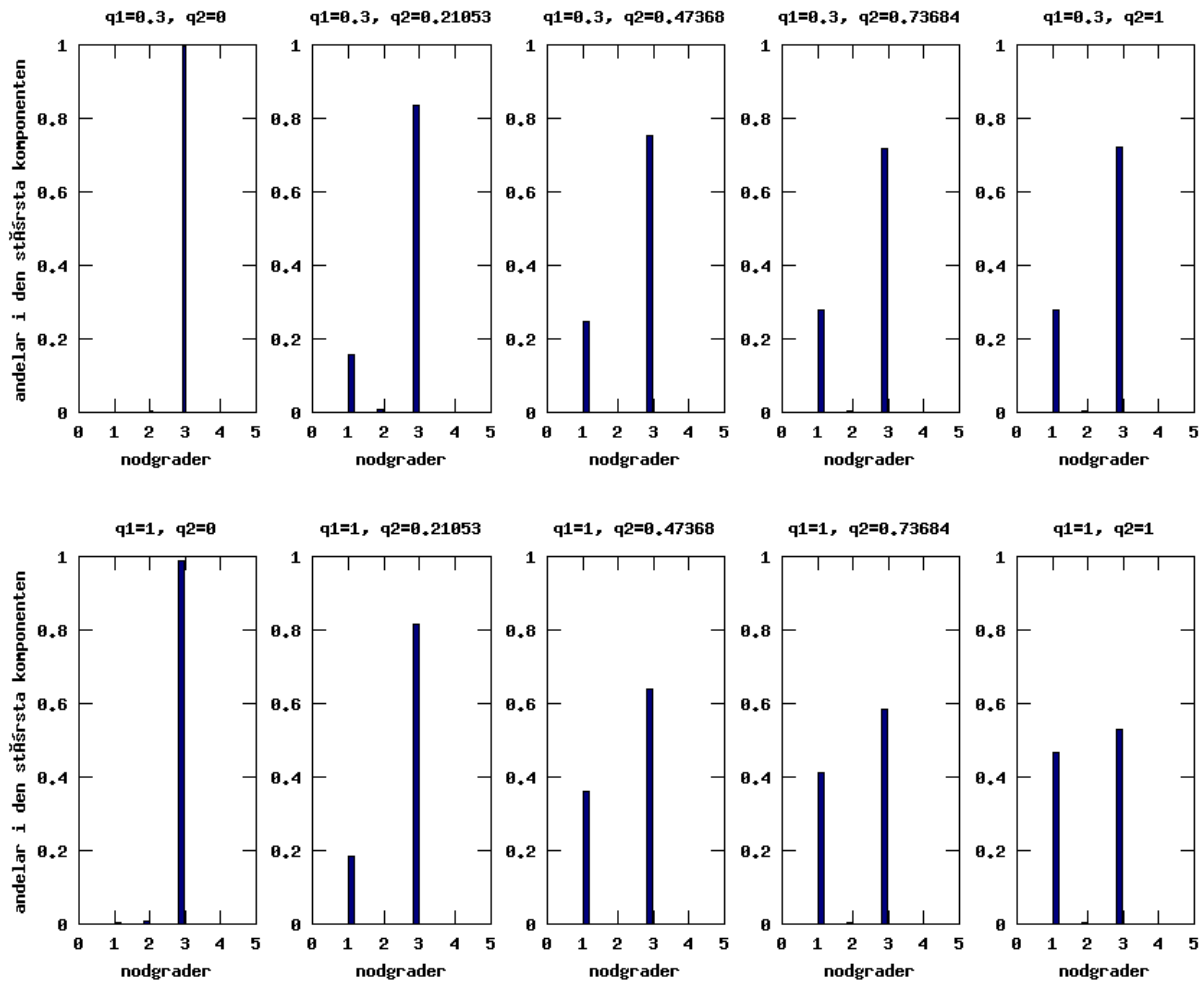
Vi definierar \tilde{D} som gradtalet på en nod som väljs genom att man väljer en halvkant på måfå, d.v.s den har fördelningen *the sized biased distribution*.

$$\begin{aligned}
 P(W = w) &= P(W = w | \tilde{D} = d) \cdot P(\tilde{D} = d) + P(W = w | \tilde{D} \neq d) \cdot P(\tilde{D} \neq d) = \\
 &= [(1 - \alpha) + \alpha P(\tilde{D} = w)] \cdot P(\tilde{D} = w) + \alpha P(\tilde{D} = w) \cdot P(\tilde{D} \neq w) = \\
 &= (1 - \alpha)P(\tilde{D} = w) + \alpha P(\tilde{D} = w) \cdot [P(\tilde{D} = w) + P(\tilde{D} \neq w)] = \\
 &= (1 - \alpha)P(\tilde{D} = w) + \alpha P(\tilde{D} = w) \cdot 1 = P(\tilde{D} = w) = \tilde{p}(w)
 \end{aligned}$$

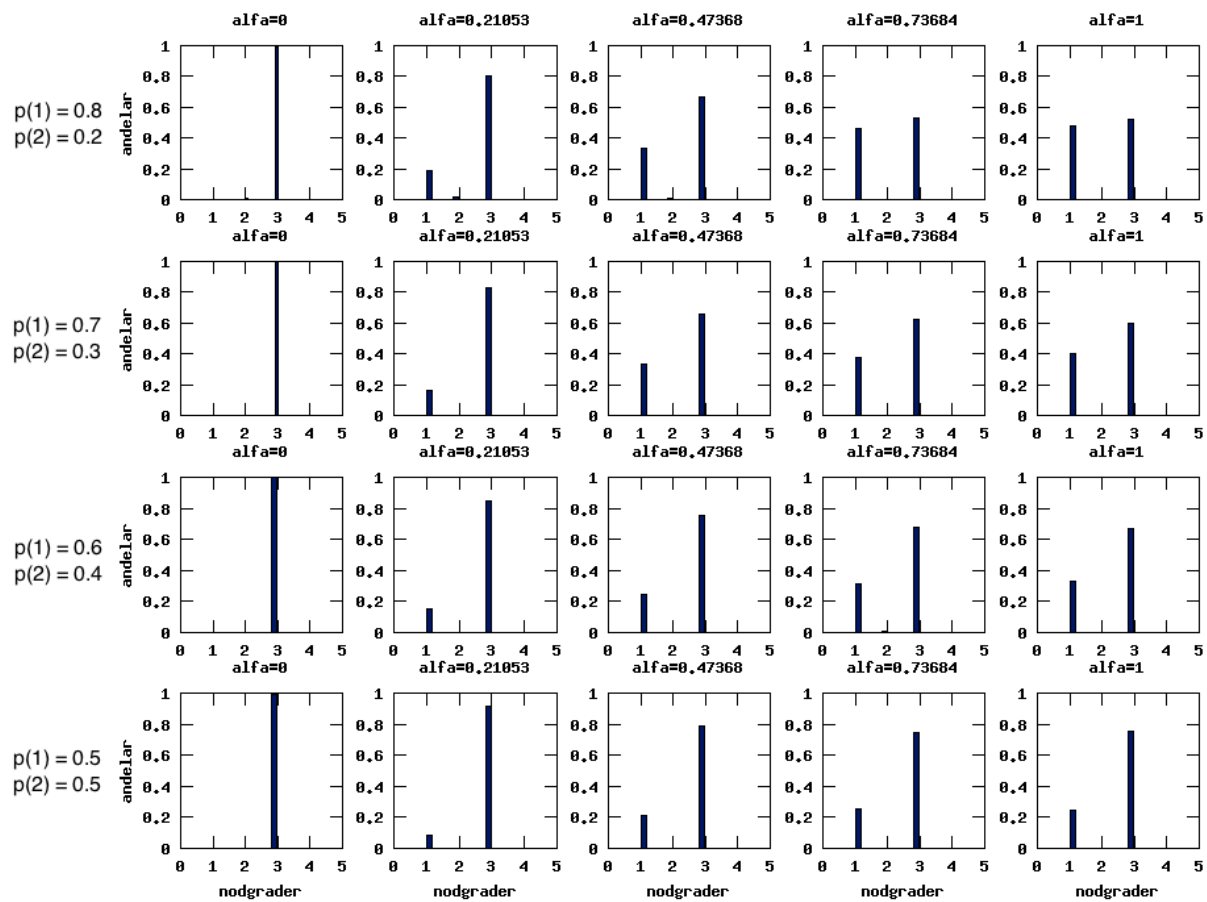
B Figurer för Exempel 1, 2 och 3



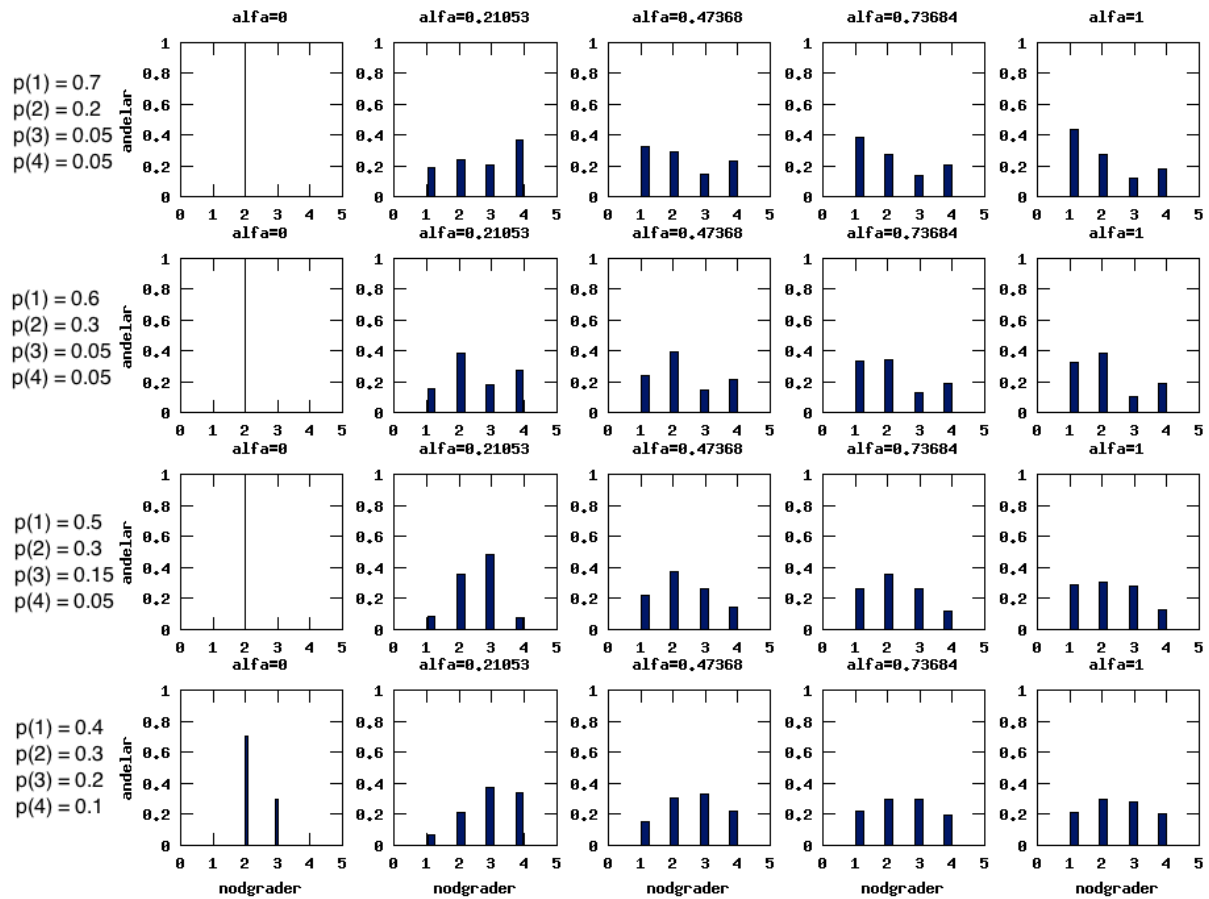
Figur 14: Exempel 1b variationskoefficient. Gradfördelningen $P(1) = 0.8, P(2) = 0.2$. Viktfördelningen: $q(w = 1|d = 1) = 0, q(w = 2|d = 1) = 1, q(w = 2|d = 3) = q_2, q(w = 3|d = 3) = 1 - q_2$



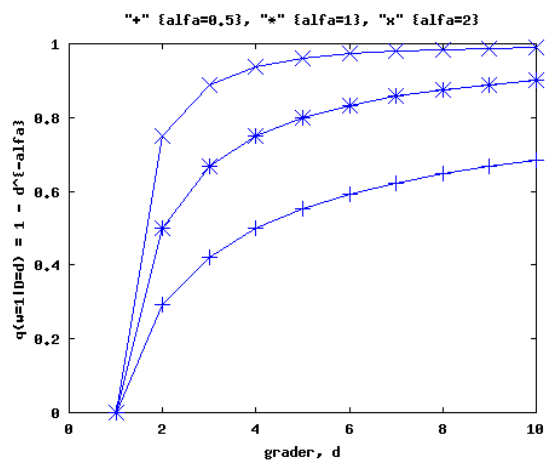
Figur 15: Exempel 1 andelar av de olika noderna i den stora komponenten. 1a överst, 1b underst



Figur 16: Exempel 2a, andelar av de olika noderna i den stora komponenten.



Figur 17: Exempel 2b, andelar av de olika noderna i den stora komponenten.



Figur 18: Exempel 3 viktfordelning: $q(w = 1|d) = 1 - d^{-\alpha}$