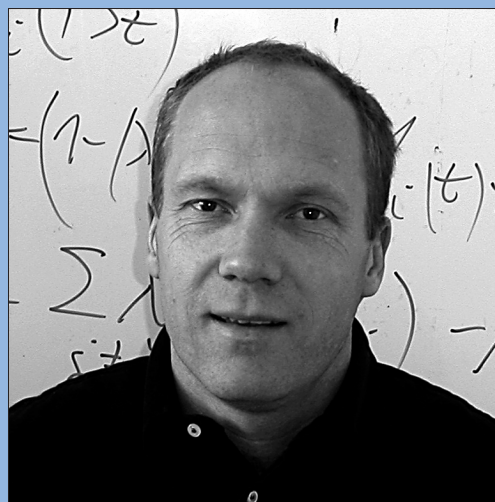


## Tom Britton

” *Avvägningen mellan att göra en matematisk modell enkel, för att kunna räkna på den, eller avancerad, för att bättre efterlikna verkligheten, är en svår balansgång.*

Jag började läsa matematik vid högskolan av en anledning som jag tror är vanlig oberoende av ämne, nämligen att jag hade en mycket bra matematiklärare, Ingmar Carlsson vid Södra Latin, som dessutom påpekade att jag hade fallenhet för ämnet och borde läsa vidare. Tack Ingmar! Väl vid högskolan tyckte jag att ämnet var roligt och det gick bra så jag fortsatte mot en magisterexamen. De sista två åren av studierna skulle man profilera sig mot någon del av ämnet.

Att jag valde matematisk statistik, dvs sannolikhetsteori och statistisk teori, var nog inte bara en slump – jag tyckte att denna slumpens matematik var mycket spännande. Det som fascinerade mig mest var för det första vilket kraftigt verktyg slumpen är för att beskriva verkliga fenomen. Inte så att det ofta förekommer ”äkta slump” (jag vet själv inte om det ens finns någon sådan) utan snarare att slump på ett mycket effektivt sätt ”tar hand” om de delar vi inte lyckas fånga upp i vår matematiska modell. Om vi tex planterar ett antal till



synes identiska frön av någon växtsort och odlar upp dem under så identiska former vi kan, så kommer växternas tillväxt inte att bli lika, utan den varierar enligt den välkända normalfördelningen. Att det kommer att bli så (jag har i och för sig aldrig utfört experimentet) beror på att hur noggranna vi än är kommer det att finnas små skillnader i fröna och deras omgivning, och många (små) skillnader brukar tillsammans skapa normalfördelade avvikelser.

Det andra i ämnet som fascinerade mig var de regler och system som den, i första betraktelsen, kaotiska slumpen följer. Att många små variationer tillsammans ger normalfördelningen som i exemplet ovan är en sådan. Denna egenskap kallas centrala gränsvärdessatsen. En annan dylik lag är att medelvärdet av många likartade slumpmässigheter inte längre är särskilt slumpmässigt. Denna ”de stora talens lag” ger sig till känna t ex genom att det i princip räcker att i en opinionsundersökning tillfråga några tusen slumpmässigt utvalda av Sveriges invånare för att med god precision veta vad hela landets invånare tycker.

Nu, när jag forskar och undervisar, inser jag också två nya egenskaper hos sannolikhetssteori och statistisk teori som jag gillar. Det första är att det är väldigt lätt att levandegöra teorin med exempel, något som underlättar och motiverar i undervisning. Det andra är att man som forskare har lätt att finna samarbeten med forskare inom andra ämnesområden. Merparten vetenskapsområden använder sig nämligen av sannolikhetssteori och /eller statistiska modeller, och att samarbeta med forskare i andra ämnesområden är något jag uppskattar alltmer.

Jag är född 1965 och uppvuxen i Stockholm. Jag disputerade vid Stockholms universitet 1996 och var därefter postdoktorstipendiat i Australien ett år och lektor/forskare vid Uppsala universitet i 5 år. Sedan 2003 är jag professor i matematisk statistik vid Stockholms universitet där jag forskar, handleder och undervisar. Min forskning rör sannolikhetssteori och statistiska metoder med tillämpningar inom biologi och medicin. Framför allt har jag arbetat med modeller och statistiska metoder för hur smittsamma sjukdomar sprids i en befolkning och hur man drar slutsatser om växtarters evolutionära träd utifrån DNA-sekvenser från växterna.

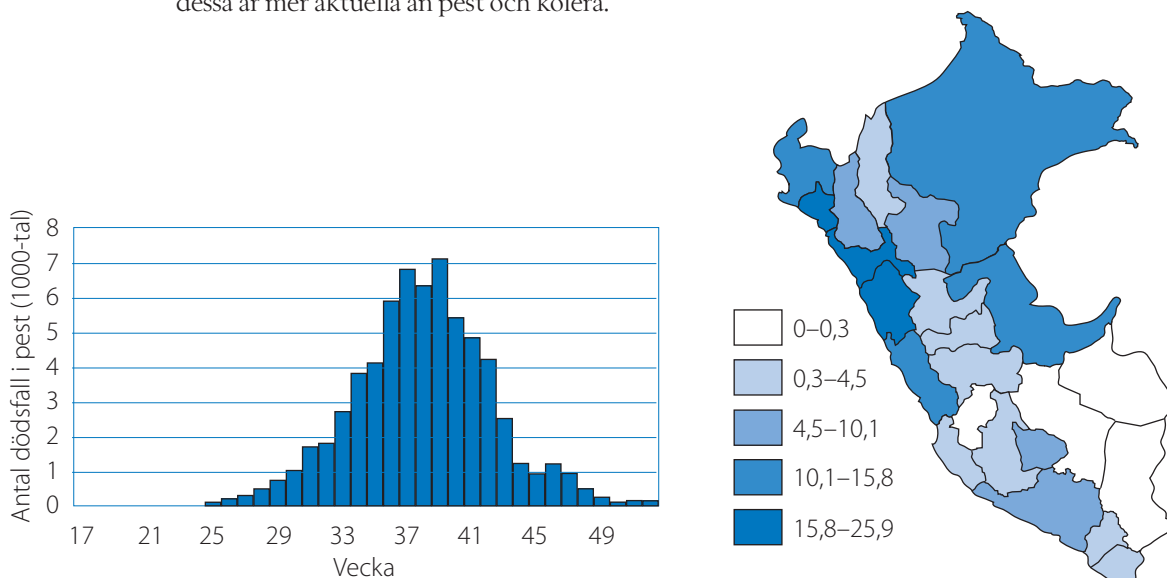
# Pest, kolera och matematik

– vad kan matematik och statistik lära oss om smittsamma sjukdomars utbredning?

Detta kapitel tar upp frågeställningar om smittsamma sjukdomars utbredning i en befolkning och hur utbredningen minskar vid vaccinering. Med hjälp av relativt enkla matematiska modeller kommer vi fram till slutsatser om hur stor andel som smittas i en befolkning, givet att vi vet sjukdomens *reproduktionstal* – ett mått på dess smittsamhet. Ur ett statistiskt perspektiv kan man därför skatta reproduktionstalet genom att observera hur många som smittades. Den viktigaste användningen av denna skattning är vid vaccination: om vi vet reproduktionstalet kan vi bestämma hur många som måste vaccineras för att undvika framtida utbrott. Genom att observera hur många som smittades under ett utbrott i en viss befolkning kan vi således skatta hur många som minst måste vaccineras i en liknande befolkning för att denna skall klara sig undan ett utbrott. För att kunna tillgodogöra sig allt i detta kapitel krävs matematik motsvarande gymnasiekompetens. Inga speciella kunskaper om smittsamma sjukdomar förutsätts. Ett visst mått av nyfikenhet skadar inte, något som gäller nästan all matematik.

Pest och kolera är två bland många infektionssjukdomar som har orsakat mycket mänskligt lidande genom tiderna. Pest är numera utrotad medan kolera fortfarande förekommer relativt frekvent i tredje världen. Till vänster i figur 1 ser du data från ett utbrott av pest i London 1665 medan kartan till höger visar antal dödsfall i ett kolerautbrott i Peru 1991.

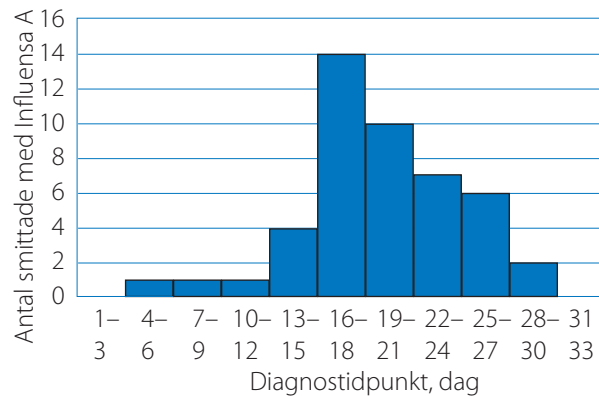
Eftersom dessa och andra sjukdomar drabbar många är det önskvärt att lära sig om utbredningen av smittsamma sjukdomar och hur man kan minska densamma. Jag kommer i kapitlet att beskriva en smittsam sjukdoms utbredning i en befolkning med hjälp av enkla matematiska modeller, och utifrån dessa försöker jag besvara frågor om hur man drar slutsatser för konkreta sjukdomsutbrott. Jag exemplifierar med influensa och mässling eftersom dessa är mer aktuella än pest och kolera.



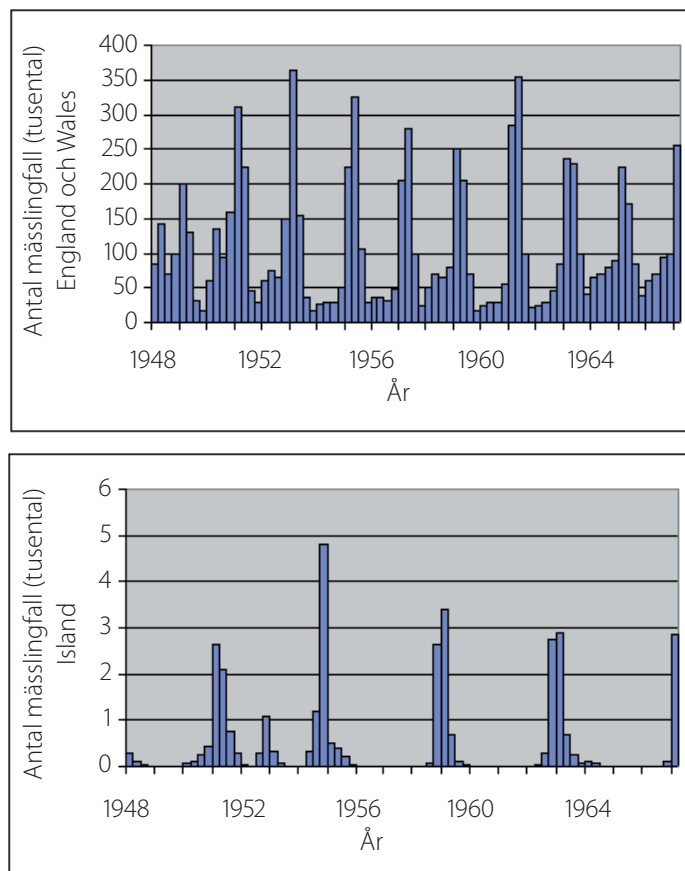
*Figur 1. Det vänstra diagrammet visar antalet dödsfall (i tusental) orsakade av pest vecka för vecka i London under ett utbrott 1665–1666. Kartan till höger beskriver antal kolerafall per 1000 invånare i respektive region i Peru under ett kolerautbrott 1991. Källa: Anderson & May respektive WHO.*

Figur 2 illustrerar ett typiskt influensautbrott, detta bland anställda och boende på ett ålderdomshem i Rochester, Minnesota, under november/december 1996. Som synes är kurvan inte helt olik den för pest i London 1665 i figur 1 – utbrott av detta slag där många smittas under kort tid varefter utbrottet klingar av brukar kallas epidemi.

Under utbrottet i Minnesota smittades 43 individer av totalt 129 boende och personal, dvs 33%. Man kan ställa ett flertal frågor relaterade till diagrammet. Varför smittades endast 33% och inte alla? Varför växer kurvan så hastigt i början för att sedan dö ut lika snabbt trots att endast en bråkdel smittades? Hur många hade man behövt vaccinera för att undvika utbrottet, eller snarare, hur många bör vaccineras nästa gång en influensaepidemi är på gång för att ett utbrott ska undvikas?



*Figur 2. Ett influensautbrott bland boende och personal på ett ålderdomshem i Minnesota 1996. Totala antalet smittade var 43 av 129 möjliga. Källa: Kuhle & Breese, 1998:3, Annals of long-term care.*



*Figur 3. Diagrammen visar antalet mässlingsfall under perioden 1948–1967, innan vaccinering startades, i England och Wales (övre diagrammet) respektive på Island (nedre diagrammet). Observera att skalorna är olika på de två y-axlarna. Källa: Anderson & May*

Vi tar nu ett nytt exempel, som illustreras i figur 3. Denna gång visar de två diagrammen i figuren antalet mässlingsfall under en 20-årsperiod, 1948–1967, i England och Wales, respektive på Island. Även dessa diagram ger upphov till ett flertal naturliga frågor: Varför verkar utbrotten ske periodiskt, dessutom med olika periodlängd (2 respektive 4 år för England och Island)? Varför försvinner mässlingen helt mellan epidemierna på Island medan sjukdomen hela tiden finns kvar i England – ett sk *endemiskt* tillstånd? Hur många måste här vaccineras om man vill få bort mässlingen ur landet?

Vi ska nu se hur matematik, sannolikhets teori och statistik kan användas för att besvara frågeställningar vid smittsamma sjukdomars utbredning, likt dem som nämndes ovan. Vi kommer att förenkla verkligheten med ett flertal antaganden och studera enkla matematiska modeller som beskriver hur en smitta sprider sig i en befolkning. Förhoppningen är att resultaten även gäller i den mer komplicerade verkligheten. Som så ofta är fallet med matematiska modeller visar det sig att samma modeller även kan användas för att beskriva andra förlopp, i detta fall t ex hur kunskap och rykten sprids i en befolkning men även hur datavirus sprids bland datorer. Här koncentrerar vi oss på smittsamma sjukdomar och kallar modellerna epidemimodeller.

## Viktiga egenskaper hos smittsamma sjukdomar

Jag vill sätta upp matematiska modeller som beskriver hur en smitta sprids i en befolkning för att kunna få svar på de frågor som formulerades i förra avsnittet och som motiverades av figurerna. Jag fokuserar på sjukdomar som sprids från människa till människa och utesluter därmed infektionssjukdomar som sprids via mat, t ex salmonella, eller via djur, t ex malaria. En smitta kan endast spridas från en person till en annan om den första personen är *smittsam* och den andre är *mottaglig*. Utöver dessa två tillstånd kan man även vara *immun*, dvs man är inte smittsam men kan ej heller smittas. De två möjliga övergångarna är att gå från mottaglig till smittsam (när man smittas), samt att upphöra att vara smittsam och bli immun (när man tillfrisknar). Denna klass av modeller brukar kallas SIR efter engelskans motsvarande tillståndsnamn och övergångar: Susceptible – Infectious – Recovered.

Utöver dessa tre tillstånd, och övergångarna däremellan, är den viktigaste matematiska egenskapen hos en smittsam sjukdom hur stor chans det är att en smittsam individ under sin smittoperiod smittar andra individer. Andra egenskaper, som hur själva smittan sprids, symptom på sjukdomen mm, har bara indirekt betydelse för sjukdomens utbredning. Inkubationstid, latensperiod och infektionsperiodens längd påverkar tidsdynamiken av sjukdomens utbredning men däremot inte utsagor om epidemins storlek eller den endemiska nivån (andelen smittsamma vid endemiskt tillstånd). Om vi tex jämför två sjukdomar där den ena är hälften så smittsam per tidsenhet, men man är smittsam dubbelt så länge, kommer i genomsnitt lika många att smittas av de två sjukdomarna även om förloppet är snabbare för den första sjukdomen.

Somliga sjukdomar har så kallade subkliniska fall, dvs individer som smittas men inte blir smittsamma. Likaså finns det sjukdomar där inte alla blir immuna. Att bortse från dessa fall är dock ingen större inskränkning eftersom de brukar vara sällsynta. Den stora förenklingen jag gör nedan är att anta att alla individer betar sig likadant och har samma förutsättningar att få sjukdomen och sprida den vidare. Konsekvensen blir att vi antar att man smittar varje annan individ med samma sannolikhet, något som förstås inte är sant i verkliga livet eftersom det tex är större chans att smitta familjemedlemmar och nära vänner jämfört med personer man inte känner. Förenklingen gör det dock väsentligt enklare att räkna i den matematiska modellen. Det finns mer avancerade matematiska modeller som tillåter varierande sannolikheter att smitta olika individer beroende på om de ingår i samma familj eller liknande. Även modeller där somliga individer är mer eller mindre mottagliga eller smittsamma förekommer. I en större befolkning bör man dock försöka hålla nere antalet olika sannolikheter i sin modell; att låta alla par av individer ha olika smittosannolikhet skulle göra det praktiskt ogenomförbart att använda modellen.

Avvägningen mellan att göra en matematisk modell enkel, för att kunna räkna på den, eller avancerad, för att bättre efterlikna verkligheten, är en svår balansgång. En matematisk modellerare försöker alltid ta med de "viktigaste" egenskaperna i det hon studerar, men utelämnar mindre viktiga detaljer. Vilka egenskaper som är "viktiga" är inte alltid lätt att avgöra, varför en modell bör granskas kritiskt innan dess slutsatser används för att dra slutsatser om verkligheten.