

Fallet Lucia

Olle Häggström

I mars 2003 dömdes den då 41-åriga Lucia de Berk av en domstol i Nederländerna till livstids fängelse för fyra mord och ytterligare ett antal mordförsök. Efter överklagan och ny rättegång dömdes hon på nytt till samma straff i juni 2004, men listan över mord hade då utökats till sju. Alltmedan de Berk år efter år satt inspärrad förekom emellertid omfattande ifrågasättanden av bevisningen, vilket till slut ledde till att fallet öppnades på nytt i oktober 2008, och i april 2010 förklarades hon icke skyldig till mordanklagelserna.

Vad gör då detta kriminalfall värt att rapportera om i *Qvintensen*? Här vill jag peka i första hand på de statistiska argument som de ursprungliga domarna byggde på, och i andra hand på det långtgående engagemang som flera ledande matematiska statistiker i Nederländerna uppvisat till försvar för de Berk och för rättsstaten – ett engagemang som till sist bar frukt i den friande domen.

Historien började i september 2001 med att en sjukhusdirektör Smits på Juliana Kinderziekenhuis i den Haag fattade misstankar om det stora antalet ”incidenter” (i den lösa betydelsen oväntade dödsfall) under sjuksköterskan Lucia de Berks skift. Inom ett par veckor var både polis och sensationspress inkopplade på fallet. Den statistiska analys av den förhöjda incidensfrekvensen som kom att ligga till grund för de fällande domarna mot de Berk bestod i en tillämpning av Fishers exakta test för (2×2)-kontingenstabeller, men gjordes inte av någon statistiker utan av rättspsykologen Hans Elffers. Denne tycks ha haft statistiska färdigheter tillräckliga för att trycka på rätt knappar för att få ut riktiga p-värden för sådana test, men inte så mycket mer. Låt mig nämna några av de statistiska tillkortakommandena (härutöver föreligger väldiga problem med den ytterst informella datainsamling som ligger till grund och som inte längre ens kan rekonstrueras).¹

För det första utför Elffers, utan närmare diskussion eller tillfredställande korrigerings, sitt hypotestest på just de data som gav upphov till den ursprungliga misstanken mot de Berk och därmed till valet av noll- och mothypotes. Detta är ett typiskt fall av att blåsa upp evidensen genom att använda samma data två gånger.

För det andra gör Elffers tre test, baserade på tre olika vårdavdelningar där de Berk gått sina skift, och väger sedan samman evidensen från de tre testen på ett härresande felaktigt vis. Han multiplicerar nämligen ihop de tre p-värdena han får, och behandlar produkten som vore den själv ett p-värde. Att det inte går an inses t.ex. av att produkten av n oberoende p-värden under nollhypotesen får väntevärde inte $\frac{1}{2}$ som sig bör utan $1/2^n$. Härav följer att, med Elffers metod, vilken oskyldig sköterska som helst kan fällas på godtycklig signifikansnivå om blott hon arbetat på tillräckligt många olika avdelningar. (Idén att multiplicera ihop p-värdena är dock inte fel i sig, men man måste då jämföra den erhållna produkten med fördelningen hos produkten av n oberoende variabler likformigt fördelade på intervallet $[0,1]$. Detta kallas Fishers metod.)²

För det tredje finns starka invändningar att göra mot nollhypotesen Elffers testar, nämligen att incidenterna på en given avdelning infaller enligt en homogen Poissonprocess (dvs i viss mening maximalt slumpmässigt). Speciellt antas Poissonintensiteten vara oberoende av vilken sköterska som går sitt pass. Detta kan inte försvaras ens under antagandet att inga mord på patienter förekommer. Exempelvis är välkänt att dödsfallsfrekvensen beror på diverse säsons-, vecko- och dygnsvariationer

1 För närmare redogörelser för vad som gått snett i rättsprocessen i allmänhet och i Elffers amatörstatistiska analys i synnerhet, se t.ex. Meester, R., van Lambalgen, M., Collins, M. och Gill, R. (2007) On the (ab)use of statistics in the legal case against the nurse Lucia de B, *Law, Probability & Risk* 5, 233-250, samt Groeneboom, P. (2007) Lucia de Berk and the amateur statisticians, <http://pietg.wordpress.com/2007/05/28/lucia-de-berk-and-the-amateur-statisticians/>

2 Se t.ex. Wikipedia-artikeln om Fishers metod: http://en.wikipedia.org/wiki/Fisher's_method

– samtidigt som inte alla sköterskor lika ofta går t.ex. nattpass. Vidare betar sig olika sköterskor olika på sätt som kan påverka dödsfallsfrekvensen, t.ex. när det gäller benägenhet att tillkalla läkare. Därtill finns en bland sköterskor välkänd (av vetenskapen ignorerad, men knappast orimlig) tendens hos patienter att hellre dö då den sköterska de känner särskild närhet till eller förtroende för är närvarande. Och så vidare. Statistikerna Richard Gill, Piet Groeneboom och Peter de Jong har nyligen studerat en modell för sådan heterogenitet.³ Under deras modell – som liksom alla modeller inte är perfekt verklighetstrogen, men den är knappast heller mindre realistisk än de antaganden som krävs för Elffers kalkyl – erhålles ett p-värde om 1/26 (att jämföra med Elffers värde, 1 på 342 miljoner). Detta pekar alltså mot att om 1000 sjuksköterskor skulle plockas på måfå och utsättas för samma slags granskning som de Berk fick utstå, så skulle kanske uppemot 40 av dem dömas för mord.

Listan över fel och egendomligheter i processen mot de Berk kan göras längre, och om den statistiska analysen kan lugnt sägas att den är så bristfällig att den förtjänar beteckningen pseudovetenskap. Det stämmer till eftertanke att ett sådant rättsövergrepp kunnat ske till och med i en rättsstat som Nederländerna. Att Lucia de Berks liv är slaget i spillror kan inte göras ogjort, men man kan i alla fall hoppas att rättssystemen både i Nederländerna och andra länder kan dra en och annan lärdom inför framtiden. En sådan skulle kunna vara att det är att föredra att anlita professionella statistiker, hellre än amatörstatistiker, för de statistiska överväganden som behövs för att ta ställning i viktiga rättsfall.

3 Gill, R., Groeneboom, P. och de Jong, P. (2010) Elementary statistics on trial (the case of Lucia de Berk), preprint, <http://www.math.leidenuniv.nl/~gill/hetero6.pdf>