

Grundläggande begrepp i sannolikhetslära

Anders Tengstrand, Örebro universitet

I dagens samhälle är både sannolikhetsläran och statistiken mer närvarande än någonsin. Vi konfronteras dagligen med prognoser som bygger på statistiska undersökningar och sannolikhetsberäkningar. Det kan gälla marknadsundersökningar, opinionsmätningar, väderleksutsikter och instrument för att förutse utvecklingen på börsen. Statistiska metoder används vid mätningar inom alla vetenskaper från humaniora till naturvetenskap och för att bedöma deras tillförlitlighet är sannolikhetsläran ett effektivt redskap. Statistik och sannolikhetslära har både blivit en viktig medborgarkunskap och ett oundgängligt redskap för specialister inom vitt skilda områden. I denna text skall vi gå igenom några begrepp i sannolikhetsläran, som blivit statistikens teoretiska grund.

Att räkna med sannolikheter

”Min bedömning är att det är mer sannolikt att det regnar än att det snöar i morgon.”, ”Den som vinner fajten om ekonomin vinner med stor sannolikhet valet.”, ”Sannolikheten att vinna på Lotto är mycket liten.” Orden ”sannolikhet”, ”sannolikt” och ”osannolikt” använder vi i privata och offentliga samtal när vi vill förmedla att vi tror på eller inte tror på att något skall inträffa. Vi använder också orden för att jämföra troligheten för olika händelseförlopp. Ibland sätter vi t.o.m. siffror på graden av övertygelse som i påståendet ”Jag är till 80 procent säker på att jag klarade provet”. Oftast bygger dessa uttalanden mer eller mindre på vad man kallar magkänsla. Ord som sannolikhet, sannolikt, säkerhet, chans och risk används intuitivt precis som många av de abstrakta begrepp som berikar vårt språk. Begreppen preciseras inte – det är inte heller nödvändigt och kanske inte ens önskvärt för att kommunikationen skall fungera. De som deltar i samtalet vet att bedömningarna är godtyckliga och oprecisa.

Medan begreppet ”sannolikhet” som det används i vår vardag är vagt och abstrakt så måste det preciseras och konkretiseras om det skall användas vid matematisk modellering. Ett område som är särskilt lämpat för att göra mer exakta sannolikhetsbedömningar är hasardspel. Det är naturligt att försöka uppskatta chanserna att vinna samtidigt som spelreglerna ger avgränsningar och gör objektiva mätningar möjliga. Det var också frågor kring hasardspel som var starten för en matematisk sannolikhets teori.

I denna modul kommer elevernas uppfattningar om slump, chans och sannolikhet att diskuteras. Förhoppningsvis kommer diskussionerna och aktiviteterna att fördjupa insikterna om hur eleverna tänker, om hur deras förståelse kan fördjupas och hur deras uppfattning om centrala begrepp kan förtydligas. För att på ett meningsfullt sätt kunna delta i en sådan diskussion krävs en gemensam terminologi och beskrivningar av de viktigaste sannolikhetsbegrepp som används i litteraturen. Denna text ägnas åt detta. Det är viktigt att tillägga att

begreppen inte skall presenteras i denna form för eleverna utan att de utgör en viktig grundkunskap för lärare.

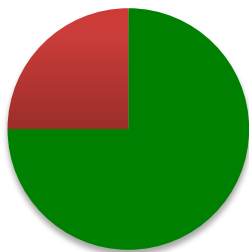
Terminologi

Innan vi går vidare behövs en fungerande terminologi. Ett ord som är centralt i sannolikhetsläran är *slump*. Med slump menar vi något som är oförutsägbart. När vi singlar slant, kastar en träning eller drar ett kort ur en kortlek gör vi ett *slumpförsök*. Ett slumpförsök resulterar i *utfall* och utfallen är oförutsägbara. Alla möjliga utfall i ett slumpförsök bildar tillsammans ett *utfallsrum*. Vid kast med mynt består utfallsrummet av krona och klave, vid kast med tärning består det av ”att få en etta”, ”att få en tvåa”, ”att få en trea”, ”att få en fyra”, ”att få en femma” och ”att få en sexa” och om vi drar kort ur en vanlig kortlek så består utfallsrummet av de 52 olika korten. Vi skall också tala om händelser. En *händelse* består av ett eller ett flertal utfall. Vid kast med tärning består händelsen ”tärningen visar ett jämnt antal prickar” av de tre utfallen ”att få en tvåa”, ”att få en fyra” och ”att få en sexa”. Händelsen ”antalet prickar är färre än 3” består av de båda utfallen ”att få en etta” och ”att få en tvåa”. Om vi drar kort ur en kortlek består händelsen ”kortet är hjärter” av 13 olika utfall. Det kan vara angeläget att påpeka att varje utfall också utgör en händelse med ett enda utfall, exempelvis ”att får hjärter kung”.

Teoretisk sannolikhet

Ett mått på sannolikheten för en händelse, där varje utfall anses ha samma chans, är kvoten mellan antalet gynnsamma utfall och det totala antalet möjliga utfall. Sannolikheten för en händelse är då alltid ett tal mellan 0 och 1 och sannolikheten för den händelse som består av alla utfall är lika med 1. De kraven kommer alla de sannolikhetsbegrepp vi introducerar att uppfylla. Sannolikheten att få krona vid slantsingling är $1/2$. Sannolikheten för varje utfall vid tärningskast är $1/6$ och sannolikheten för att få ett jämnt antal prickar är $3/6 = 1/2$. Sannolikheten för varje utfall när man på måfå drar ett kort ur en kortlek är $1/52$ och sannolikheten för att få hjärter är $13/52 = 1/4$.

I vissa hasardspel kan man inte ge de olika utfallen samma sannolikhet. Vi har väl alla någon gång spelat på ett chokladhjul som är indelat i olika zoner. Några zoner ger bättre utdelning än andra och några kanske inte ger någon utdelning alls. Anta att vi har ett hjul med bara två zoner där den ena är röd och den andra grön. Den röda zonen upptar en kvartscirkel och den återstående delen är grön. Om pilen stannar på den röda zonen blir vinsten en relativt dyr chokladask. Om den stannar på den gröna blir det ingen utdelning alls och insatsen är förlorad. Spelet på vårt chokladhjul är ett slumpförsök med två utfall, ”grönt” och ”rött”. Men i detta fall är det orimligt att ge de båda utfallen samma sannolikhet. Sannolikheten för ”grönt” bör bedömas vara tre gånger så stor som sannolikheten för ”rött”. Eftersom vi vill att summan



Den röda zonen upptar en kvartscirkel och den återstående delen är grön. Om pilen stannar på den röda zonen blir vinsten en relativt dyr chokladask. Om den stannar på den gröna blir det ingen utdelning alls och insatsen är förlorad. Spelet på vårt chokladhjul är ett slumpförsök med två utfall, ”grönt” och ”rött”. Men i detta

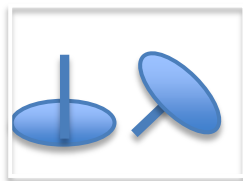
fall är det orimligt att ge de båda utfallen samma sannolikhet. Sannolikheten för ”grönt” bör bedömas vara tre gånger så stor som sannolikheten för ”rött”. Eftersom vi vill att summan

av sannolikheterna skall vara 1 så skall vi sätta sannolikheten för ”grönt” lika med $3/4$ och sannolikheten för ”rött” lika med $1/4$.

I de exempel som vi har diskuterat har vi kunnat tilldela varje utfall en sannolikhet med utgångspunkt från förutsättningarna. När vi kastar ett fullständigt symmetriskt mynt eller tärning är det rimligt att varje utfall har samma sannolikhet och samma sak gäller om vi på måfå väljer ett kort ur en kortlek. Om vi spelar på ett chokladhjul bestämmer indelningen av hjulet sannolikheterna. Om vi som i exemplen i förväg kan bestämma sannolikheterna för de olika utfallen utifrån slumpförsökets förutsättningar talar vi om ett *teoretiskt sannolikhetsbegrepp*. Ett specialfall är det *klassiska sannolikhetsbegreppet* där varje utfall har samma sannolikhet. Vi skall i nästa avsnitt se att det teoretiska sannolikhetsbegreppet inte alltid är tillämbart.

Experimentell sannolikhet och de stora talens lag

Vi har hittills diskuterat slumpförsök där det varit möjligt att teoretiskt bestämma sannolikheten för varje utfall med utgångspunkt från förutsättningarna. Men tyvärr är det svårt för att inte säga omöjligt i många fall. Låt oss ta ett exempel. Istället för att kasta ett mynt eller en tärning så kastar vi ett häftstift. Det finns då två möjliga utfall: Häftstiftet hamnar med spetsen uppåt eller nedåt. Vi bortser från möjligheten att häftstiftet kommer att balansera



på spetsen. Sannolikheterna för de båda utfallen är antagligen inte lika stora och de beror på häftstiftets form. Här är det svårt att teoretiskt beräkna sannolikheterna för de båda utfallen. Det skulle kräva en noggrann analys av stiftets form, material m.m. för att därefter möjligen kunna dra slutsatser om sannolikheterna. Men detta är i praktiken genomförbart.

För att kunna mäta sannolikheten även i detta mer komplicerade fall så återgår vi till det enklare fallet med tärning. Vi kastar en tärning ett antal gånger och noterar hur många gånger vi får en sexa. Det antalet kallar vi *absoluta frekvensen* för en sexa. Om vi dividerar absoluta frekvensen med totala antalet kast får vi ett tal som vi kallar för *relativa frekvensen* för en sexa. Den relativa frekvensen för en sexa är alltså andelen sexor av det totala antalet kast. Om vi gör 100 kast och 24 av dessa visar en sexa så är alltså den relativa frekvensen $24/100 = 0,24$. Den relativa frekvensen varierar naturligtvis när antalet kast ökar. Vi gör en tabell och kommer kanske fram till följande:

Antal kast	1	5	10	100	200	500	1000
Antal sexor	0	1	3	24	40	84	170
Relativ frekvens	0	0,2	0,3	0,24	0,2	0,168	0,170

Man kan teoretiskt bevisa att de relativa frekvenserna närmar sig den teoretiska sannolikheten $1/6$ om antalet kast växer obegränsat. De relativa frekvenserna stabiliseras om försöket upprepas ett stort antal gånger. Detta resultat kallas *De stora talens lag*. Det ger oss en idé om hur vi skall bestämma sannolikheten för de olika utfallen vid kast med häftstift. Vi kan

kasta häftstiftet ett stort antal gånger och beräkna de relativa frekvenserna för t.ex. utfallet ”spetsen upp”. De kommer då att så småningom stabiliseras kring ett tal och det talet kallar vi för sannolikheten för ”spetsen upp”.

Vi kan konstruera ett nytt *experimentellt sannolikhetsbegrepp* där sannolikheten för ett utfall vid ett slumpförsök bestäms genom att upprepa försöket ett stort antal gånger. Varje försök skall utföras under identiska förhållanden och de skall vara oberoende av varandra. Den relativa frekvensen av ett utfall antas stabilisera sig kring ett värde efter ett stort antal försök och det värdet kallar vi utfallets sannolikhet.

Det experimentella sannolikhetsbegreppet ger oss möjlighet att mäta sannolikheter i helt andra sammanhang än de som har med hasardspel att göra. Om vi t.ex. vill bestämma sannolikheterna för att ett nyfött barn är en flicka eller en pojke kan vi studera ett stort antal födselar och bestämma de relativa frekvenserna för flicka respektive pojke. Det finns ett rikt statistiskt material som vi kan använda oss av. Om vi baserar vår uppskattning på barn födda i Sverige under 2010 så är sannolikheterna för att ett nyfött barn är en flicka respektive pojke ungefär lika med 0,487 respektive 0,513.

Ett subjektivt sannolikhetsbegrepp

Det teoretiska sannolikhetsbegreppet kan användas när vi kan bestämma de olika utfallens sannolikhet utifrån de givna förutsättningarna. Begreppet har stora begränsningar och ett viktigt tillämpningsområde är hasardspel. Det experimentella sannolikhetsbegreppet förutsätter att slumpförsöket kan upprepas. Men när vi i dagligt tal använder ordet ”sannolikhet” eller ”chans” har det ofta en betydligt vagare betydelse. Att en kärnkraftsolycka inträffar med en viss procents sannolikhet är inte praktiskt möjligt att verifiera genom teoretiska beräkningar utifrån de givna förutsättningarna. Det kan inte heller visas genom att utföra upprepade försök. När en läkare ger en diagnos så är det alltid med en viss osäkerhet och den anges med ett tal mellan 0 och 1. Bakom diagnosen ligger ofta en kombination av medicinsk kunskap och erfarenhet, men man kan knappast tala om att sannolikhetsmättet har sin grund i upprepade försök och stabiliteten hos de relativa frekvenserna. Karakteristiskt för denna typ av sannolikhetsbedömningar är att de förändras när informationen ökar.

Det är möjligt att konstruera en logiskt uppbyggd matematisk teori med ett sannolikhetsbegrepp som kan tillämpas i de situationer vi beskrivit ovan och där sannolikheten av ett påstående anger i vilken grad vi är övertygad om dess riktighet. Ett sådant sannolikhetsbegrepp kallas *subjektivt* eller *induktivt*. En sådan teori är komplicerad och det finns inga möjligheter att gå närmare in på den. Vi kommer inte att använda det subjektiva sannolikhetsbegreppet i aktiviteterna i denna modul.

Referenser

- Batanero, C., Henry, M., & Parzys, B. (2005). The nature of chance and probability. I G. A. Jones (red.), *Exploring probability in school: challenges for teaching and learning* (pp. 15-37). Dordrecht: Kluwer Academic Publishers.