Förstånd och missförstånd kring de stora talens lag

Olle Häggström*

Ett av ganska få begrepp från sannolikhetsteorin som trängt in i det allmänna medvetandet, är **de stora talens lag**. Men hur många vet egentligen vad den innebär? Jag vill med dessa rader försöka räta ut en del frågetecken, och skingra en del missuppfattningar, som florerar kring denna berömda matematiska sats.

Ibland stöter man på hänvisningar till de stora talens lag i sammanhang som inte alls har med denna att göra[1]. Vanligare är dock att den som hänvisar till lagen ändå har någon sorts uppfattning av vad den handlar om, och ofta formuleras den ungefär så här:

$$\text{Tur och otur jämnar i längden ut sig.} \tag{1}$$

Det ligger mycket i denna formulering, även om den som vi skall se inte är helt invändningsfri.

För att föra en mer precis diskussion om de stora talens lag behöver vi införa en del matematiska begrepp. En **stokastisk variablel** kan vi tänka på som en storhet som beror på slumpen. För att modellera ett tärningskast kan vi t.ex. införa en stokastisk variabel X, som kan anta värdet 1, 2, 3, 4, 5 eller 6, och som representerar tärningens utfall. Om tärningen är rättvis har vart och ett av utfallen sannolikheten $\frac{1}{6}$, och vi skriver

$$\mathbf{P}[X = k] = \frac{1}{6} \quad \text{för } k = 1, \ldots, 6.$$

P är (under inflytande från franskan och engelskan) den gängse matematiska beteckningen för sannolikheter.

Om vi istället vill modellera slantsingling kan vi låta den stokastiska variableln X ha de möjliga värdena 0 (för klave) och 1 (för krona). Om slanten är symmetrisk får vi[2]

$$\mathbf{P}[X = 0] = \mathbf{P}[X = 1] = 0{,}5\,.$$

En viktig storhet i anslutning till en stokastisk variabel X är dess **väntevärde** $\mathbf{E}[X]$ (**E** som i *expectation*) som anger vad vi i genomsnitt väntar oss av X. Om alla utfall är lika sannolika, fås väntevärdet som medelvärdet av de möjliga utfallen. För slantsinglingen får vi alltså väntevärdet $\mathbf{E}[X] = (0 + 1)/2 = 0{,}5$, och för tärningen fås

*Matematisk statistik, Chalmers tekniska högskola och Göteborgs universitet, 412 96 Göteborg, `olleh@math.chalmers.se`, `http://www.math.chalmers.se/~olleh/`

[1] Exemeplvis lät en berömd (och av mig mycket uppskattad!) TV-sportreporter från Laholm ett par gånger under OS i Sydney undslippa sig uttrycket "detta strider mot alla stora talens lag" när något han bedömt som synnerligen osannolikt – det kan ha rört sig om att två simmare gjorde på hundradelen samma tid – likväl hade inträffat.

[2] Denna modell har visat sig användbar, men kan också (liksom vår modell för tärningskast) kritiseras. Om man t.ex. gör en tillräckligt noggrann mätning av myntets utgångsläge och -hastighet då det kastas, kan man inte längre hävda att utfallet är slumpmässigt, eftersom dess rörelse bestäms av Newtons lagar, som är deterministiska. För en intressant diskussion om denna aspekt på slantsingling, se De Groot, M. (1986) A Conversation with Persi Diaconis, *Statistical Science* **1**, 319–334.

$\mathbf{E}[X] = (1+2+3+4+5+6)/6 = 3{,}5$. (Om olika utfall har olika sannolikhet blir väntevärdet istället ett *viktat* medelvärde av de olika utfallen.)

Antag nu att vi singlar vår slant upprepade gånger, och symboliserar utfallen med stokastiska variabler $X_1, X_2, \ldots$ som var och en alltså antar värdet 0 eller 1 med sannolikhet 0,5 vardera. Vi antar vidare att de olika singlingarna är **oberoende**, vilket innebär att, då vi singlar slanten den n:te gången, vi inte har någon som helst ledning av $X_1, \ldots, X_{n-1}$ om vi vill prediktera X_n. Detta antas gälla för alla n. Vårt antagande är fysikasliskt mycket rimligt[3], eftersom ju slanten knappast kan ha något "minne" av tidigare singlingar.

Låt oss definiera, för varje n,

$$S_n = X_1 + X_2 + \cdots + X_n \tag{2}$$

och

$$M_n = \frac{S_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}. \tag{3}$$

S_n är alltså *antalet* krona efter n kast med myntet, medan M_n är *andelen* krona. (S och M står för summa respektive medelvärde.) De stora talens lag säger nu att om vi fortsätter singla slanten i evighet, så får vi gränsvärdet

$$\lim_{n\to\infty} M_n = 0{,}5$$

med sannolikhet 1. Vi kan alltså vara säkra på att andelen krona i det långa loppet kommer att svänga in allt närmare 0,5. Detta innebär bland annat att

$$\mathbf{P}[0{,}49 < M_n < 0{,}51]$$

alltmer närmar sig 1 då $n \to \infty$. Detsamma gäller för

$$\mathbf{P}[0{,}5 - \varepsilon < M_n < 0{,}5 + \varepsilon], \tag{4}$$

hur litet $\varepsilon > 0$ vi än väljer.

Mer generellt säger de stora talens lag att om vi har en följd $X_1, X_2, \ldots$ av stokastiska variabler med samma fördelning, och väntervärde m, så får vi

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = m$$

med sannolikhet 1. Den första versionen av de stora talens lag erhölls av Jacob Bernoulli 1713, då han bevisade att (4) gäller för slantsinglingsexemplet. För oberoende stokastiska variabler med en och samma fördelning fick de stora talens lag sin slutliga formulering av Andrej Kolmogorov 1933 – alltså mer än 200 år senare! Än idag kan man i matematiska tidskrifter varje månad stöta på bevis av de stora talens lag i nya situationer med olika typer av beroenden mellan de ingående stokastiska variablerna.

I vår modell för slantsinglingar vet vi alltså att andelen krona M_n kommer att ligga nära 0,5 för tillräckligt stora n. Betyder detta att *antalet* krona S_n kommer att vara nära $\frac{n}{2}$? I absoluta termer är svaret nej. Exempelvis kan man bevisa att $|S_n - \frac{n}{2}| > 100$ med en sannolikhet som går mot 1 då $n \to \infty$ (och i detta påstående kan talet 100 bytas

[3]Även detta antagande har visat sig fungera väl i praktiken, men är naturligtvis åter möjligt att kritisera; jämför föregående fotnot.

ut mot vilket ändligt tal som helst). Det är först när vi dividerar med n som skillnaden blir liten.

Ett annat exempel: vi tänker oss att vi köper en lottsedel i veckan, i ett lotteri som har samma vinstplan varje vecka. Varje lott kostar 20 kr, och den genomsnittliga vinstutbetalningen per lott är 10 kr. Kan vi i denna situation lite på att tesen (1) om tur och otur gäller? Låt X_n vara vår nettovinst (dvs vinstutbetalning minus lottpris) vecka n. X_n kan då ses som en stokastisk variabel med väntevärde

$$\mathbf{E}[X_n] = 10 - 20 = -10 .$$

Definiera S_n och M_n som i (2) och (3), så att S_n alltså är vår totala vinst de n första veckorna, och M_n är vinsten per vecka. Nu kan vi, med stöd av de stora talens lag, dra slutsatesn att M_n alltmer kommer att närma sig väntevärdet -10 då n ökar. I denna mening är alltså påståendet (1) alldeles riktigt[4]. Däremot är det *inte* sant att den totala vinsten S_n kommer att hålla sig i närheten av dess förväntade värde $-10n$. En sådan tolkning av (1) är alltså en överdriven och omotiverad tilltro till det långa loppets utjämning av tur och otur. I själva verket kommer diskrepansen $S_n - (-10n)$ mellan total vinst och förväntad total vinst att fluktuera inom allt vidare ramar då n ökar.

Hur ser insvängningen mot väntevärdet ut i praktiken? Det varierar naturligtvis. I Figur 1 visas en datorsimulering av slantsinglingsexemplet, och hur andelen krona M_n kan bete sig som en funktion av antalet slantsinglingar n, för $n = 1, \ldots, 100$. För $n = 100$ ser vi att $M_n = 0{,}55$ (dvs 55 krona på totalt 100 kast), vilket är hyfsat nära 0,5, men likväl en bit ifrån. Detta är ett relativt normalt utfall, vilket kan förstås av Figur 2, som är ett diagram över den exakta fördelningen för M_{100}. Tydligen är $n = 100$ inte tillräckligt för att komma *riktigt* nära 0,5.

När vi nu har kastat myntet 100 gånger (eller låtit datorn göra jobbet!) och fått 55 krona, så kan vi stanna upp och spekulera lite över hur det skall gå i ett fortsatt slantsinglande. Låt mig besvara några naturliga och närliggande frågor kring detta.

Fråga: Har jag förstått saken rätt, att de stora talens lag garanterar att kurvan kommer att svänga in allt närmare den streckade linjen $M_n = 0{,}5$ om vi håller på tillräckligt länge?

Svar: Ja, det är alldeles riktigt.

Fråga: Vad händer om vi nöjer oss med totalt 200 kast? Är det då troligt att vi har kommit närmare 0,5?

Svar: Ja, det stämmer.

Fråga: Det är med andra ord troligt att vi får fler klave än krona bland de kommande 100 kasten? I alla fall mer troligt än att vi skulle få fler krona än klave?

Svar: Nej!

Fråga: Va? Du sa ju nyss att vi troligen kommer att närma oss 0,5! Då måste ju M_n sjunka!

[4] Vi kan också dra slutsatsen att det i längden *med säkerhet* är olönsamt att köpa lotter. Spelets inbyggda orättvisa kan inga maximer som (1) rå på.

Svar: Just det, M_n kommer troligen att sjunka.

Fråga: Men du sa ju nyss att det *inte* är troligt att vi får fler klave än krona!?

Svar: Det är riktigt.

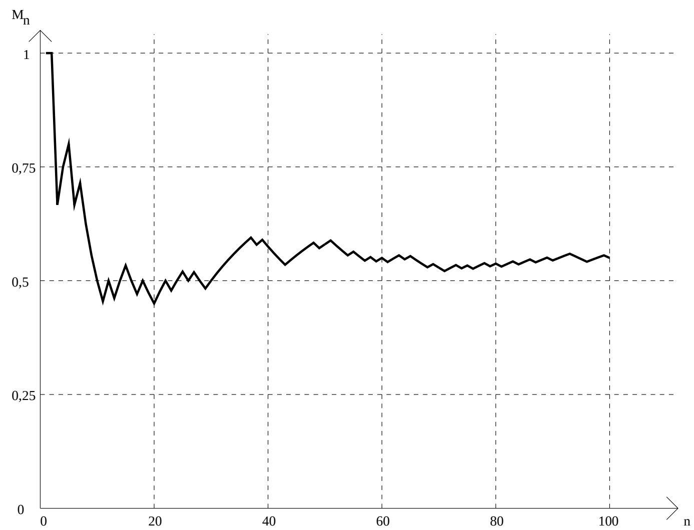
Fråga: Men hur går detta ihop? Om M_n skall sjunka så måste vi väl ha fler klave än krona?

Svar: Nej, det är just det som inte behövs! Om vi t.ex. får lika många krona som klave (50 av varje) bland de 100 nästkommande kasten, så blir totala antalet krona efter 200 kast 105, och M_{200} blir $\frac{105}{200} = 0{,}525$, vilket är lägre än M_{100}. Det kan rentav bli så att vi får *fler* krona än klave på de nästa 100 kasten, och att M_n *ändå* sjunker! Låt oss säga att de 100 nästa kasten ger 53 krona och 47 klave: då blir $M_{200} = \frac{55+53}{200} = 0{,}54$, vilket åter är en sänkning jämfört med M_{100}.

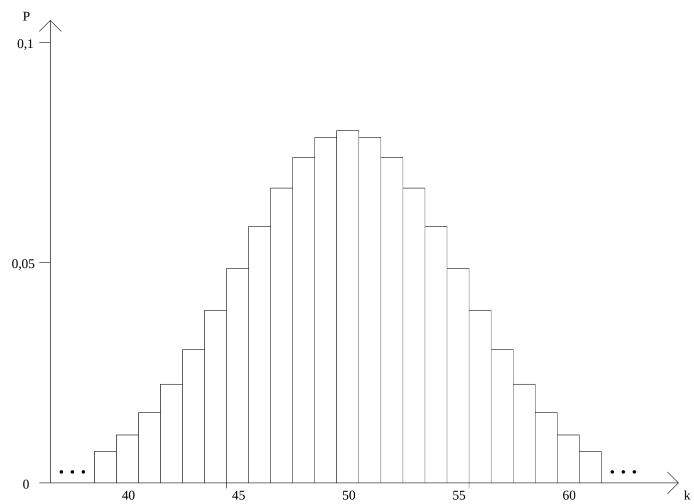
Denna påhittade dialog illustrerar det som jag tror är det vanligaste missförståndet med de stora talens lag, nämligen att om vi har fått alltför många krona i det förflutna (fler än vad som förutspås av väntevärdet), så kommer detta att korrigeras genom en ökad tendens hos myntet att visa klave i framtida kast. Men detta är en grov övertolkning[5] av de stora talens lag, och den strider i själva verket mot en av förutsättningarna: slantsinglingarna antogs ju vara oberoende av varandra!

Tendensen hos M_n att svänga ner mot 0,5 då $M_{100} = 0{,}55$ beror alltså *inte* på någon ökad "längtan" efter klave hos myntet. Istället kan vi förstå saken på följande vis. För att bevara status quo till tidpunkten $n = 200$, krävs lika många krona i kasten $101, 102, \ldots, 200$ som i de 100 första, dvs 55 stycken. Då fås nämligen $M_{200} = \frac{55+55}{200} = 0{,}55$. Fler än 55 krona resulterar i en ökning av andelen krona, medan färre än 55 resulterar i en minskning av denna andel. Samtidigt vet vi att fördelningen för antalet krona i kasten $101, 102, \ldots, 200$ är densamma som fördelningen för antal krona i de 100 första kasten. Denna fördelning är (av uppenbara skäl) symmetrisk kring värdet 50 – se Figur 2 igen – och sannolikheten för att ha färre än 55 krona är därför betydligt större än den att ha fler.

[5] Med tanke på det (som jag tidigare påpekade) fysikaliskt orimliga i att myntet skulle ha något minne av de gamla utfallen, som påverkar de kommande, tycker jag att denna övertolkning närmast är att betrakta som vidskepelse. Men vi har väl alla åtminstone någon gång stött på någon som vid roulettebordet hävdar att "så mycket svart som det blivit på sistone, är det nog dags för rött nu", eller som med lottokupongen i hand påpekar att "det var länge sedan nummer 19 drogs, så denna vecka tar jag med den"? Denna typ av sannolikhetsteoretiska missförstånd är nog i själva verket en bidragande orsak till att svenska folket gör slut på en så förskräckande stor del av sina hushållsekonomier på olika hasardspel.



Figur 1. En datorsimulering av hur andelen krona M_n kan bete sig för en serie av 100 slantsinglingar. Efter ganska häftiga fluktuationer i början, kan man i det här fallet redan kring $n = 15$ få för sig att insvängningen mot 0,5 kommit ganska långt på vägen, men så, vid $n = 30$ till och med 37, får vi inte mindre än åtta krona i rad (sådant händer emellanåt!), och M_n glider iväg på ett sätt som tar lång tid att ”reparera”. Den intresserade läsaren rekommenderas varmt att göra om experimentet fár att få fler exempel på hur insvängningen mot $0,5$ kan te sig.



Figur 2. Sannolikheten att få exakt k krona på 100 kast, som en funktion av k. Notera att $S_{100} = 50$ är det mest sannolika utfallet, men att $S_{100} = 55$ på intet vis är någon extrem händelse.