Sannolikhet och statistik II

Johan Jonasson

Mars 2019

Uppräkneliga och ändliga utfallsrum

$S = \{u_1, u_2, \ldots\}$.

Exempelvis om vi singlar slant till första klaven är det naturligt att ha $S = \{1, 2, 3, \ldots\}$.

Antag att $p_1, p_2, \ldots$ är sådana att $p_i \geq 0$ för all n och $\sum_{i=1}^{\infty} p_i = 1$. Låt för all $A \in \mathcal{P}(A)$,

$$\mathbb{P}(A) = \sum_{i:u_i \in A} p_i.$$

Då är $\mathbb{P}$ ett sannolikhetsmått och $\mathbb{P}(\{u_i\}) = p_i$.

Exempel: Slantsingling till första klaven, $S = \{1, 2, 3, \dots\}$. Med en rättvis slant har man $p_n = 1/2^n$, dvs

$$p_1 = \frac{1}{2},\ p_2 = \frac{1}{4},\ p_3 = \frac{1}{8},\ \dots.$$

Vi får exempelvis

$$\mathbb{P}(\{\text{högst tre kast krävs}\}) = \mathbb{P}(\{1, 2, 3\}) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}.$$

Det klassiska sannolikhetsmåttet

Här är $S = \{u_1, u_2, \dots, u_n\}$ och $p_i = \frac{1}{n}$.
Då blir

$$\mathbb{P}(A) = \frac{\text{antal utfall i } A}{\text{totala antalet utfall}} = \frac{|A|}{n} = \frac{|A|}{|S|}.$$

Mycket vanlig situation.

Exempel. Om man slår en tärning tre gånger, vad är sannolikheten att få exakt en sexa?

$S = \{(x,y,z) : x,y,z \in \{1,2,3,4,5,6\}\}$. Det gäller att $|S| = 6^3 = 216$. Vi söker $\mathbb{P}(A)$ för

$$A = \{(x,y,z) \in S : \text{exakt en av } x,\ y,\ z \text{ är } 6\}.$$

Vi har $|A| = 1 \cdot 5 \cdot 5 + 5 \cdot 1 \cdot 5 + 5 \cdot 5 \cdot 1 = 75$. Alltså

$$\mathbb{P}(A) = \frac{75}{216} = \frac{25}{72}.$$

Exempel. I en hatt ligger tre kort: rött-rött, rött-svart, svart-svart. Ett kort dras på måfå och läggs med på måfå vald sida upp. Vad är sannolikheten att den andra sidan har samma färg?

Kalla korten a, b och c. Vi kan ta

$$S = \{\text{a sida 1, a sida 2, b sida 1, b sida 2, c sida 1, c sida 2}\}.$$

Vi söker $\mathbb{P}(A)$ där

$$A = \{\text{a sida 1, a: sida 2, c sida 1, c sida 2}\}.$$

Alltså blir svaret

$$\mathbb{P}(A) = \frac{4}{6} = \frac{2}{3}.$$

Vi behöver kunna räkna antal element i en mängd.

Kombinatorik

Multiplikationsprincipen: Om man utför r stycken experiment i tur och ordning och experimenten kan utfalla på n_1, n_2, ... respektive n_r sätt, så är det totala antalet utfall för alla experiment tillsammans

$$n_1 n_2 \dots n_r.$$

Exempel. Födelsedagsproblemet. En skolklass har r elever. Vad är sannolikheten att alla har olika födelsedagar?

$$S = \{(x_1, x_2, \dots, x_r) : \forall i : x_i \in \{1, 2, \dots, 365\}\}.$$

$$A = \{(x_1, x_2, \dots, x_r) \in S : \forall i \neq j : x_i \neq x_j\}.$$

$$\mathbb{P}(A) = \frac{365 \cdot 364 \cdot 363 \cdot (365 - r + 1)}{365^r}.$$

(Detta är mindre än $1/2$ då $r \geq 23$.)

Detta var ett exempel på att: Antalet sätt att välja r element ur en mängd med n element utan återläggning med hänsyn till ordning är

$$n(n-1)(n-2)\ldots(n-r+1) = \frac{n!}{(n-r)!} =: (n)_r.$$

Detta följer av mutiplikationsprincipen.

Antalet sätt att välja r element ur en mängd av n element utan återläggning utan hänsyn till ordning betecknas

$$\binom{n}{r}.$$

Enligt multiplikationsprincipen gäller

$$(n)_r = \binom{n}{r} \cdot r!$$

dvs

$$\binom{n}{r} = \frac{(n)_r}{r!} = \frac{n!}{r!(n-r)!}.$$

Exempel. Att välja tre kort ur en kortlek kan göras, med hänsyn till ordning, på

$$52 \cdot 51 \cdot 50 = 132600$$

sätt. Utan hänsyn till ordning kan de väljas på

$$\binom{52}{3} = \frac{52 \cdot 51 \cdot 50}{1 \cdot 2 \cdot 3} = 52 \cdot 17 \cdot 25 = 22100$$

sätt.

Antalet sätt att välja r element ur en mängd med n element med återläggning och med hänsyn till ordning är förstås n^r.

Antal sätt att välja r element med återläggning men utan hänsyn till ordning är lika med antalet ickenegativa heltalslösningar till ekvationen

$$x_1 + x_2 + \ldots + x_n = r.$$

Detta är antalet sätt att placera ut $n-1$ väggar bland r bollar, vilket kan ske på

$$\binom{n-1+r}{r} = \binom{n-1+r}{n-1}$$

sätt.

Det sista resultatet är inte så tillämpbart utan mest ett exempel.

Betingad sannolikhet och oberoende

Exempel. Två tärningar slås. Vad är sannolikheten för två sexor? Tärningarna påverkar inte varandra så om A_i är händelsen att tärning i visar sexa ($i = 1, 2$), borde vi ha

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2).$$

Exempel. Vad är sannolikheten att det regnar både den 1 och den 2 augusti i år? Låt R_i vara händelsen att det regnar den i:te augusti. Eventuell kunskap om att det regnar den 1/8 påverkar i hög grad hur sannolikhet vi tror att det är att det regnar den 2/8, dvs

$$\mathbb{P}(R_1 \cap R_2) \neq \mathbb{P}(R_1)\mathbb{P}(R_2)$$

utan snarare

$$\mathbb{P}(R_1 \cap R_2) = \mathbb{P}(R_1) \cdot \text{sannolikheten för } R_2 \text{ givet att } R_1 \text{ inträffar.}$$

Definition: För två händelser A och B ges sannolikheten för B givet A av

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

Definition: Två händelser A och B sägs vara oberoende om

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Notera att om $\mathbb{P}(A) > 0$ så är A och B oberoende om och endast om $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Exempel: Slå en blå tärning och en gul tärning. Låt

$$A = \{\text{blå tärning sexa}\} = \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$$

$$B = \{\text{gul tärning sexa}\} = \{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6)\}$$

$$C = \{\text{summan är 7}\} = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$$D = \{\text{summan är 8}\} = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$$

Rimligt modell att anta att alla 36 utfall är lika sannolika. Vi har då

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/6, \ \mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = 1/36$$

så A och B är oberoende och A och C är oberoende och B och C är oberoende. (Men $\mathbb{P}(A \cap B \cap C) = 0$ så det är inte rimligt att säga att trippeln A, B, C är oberoende.)

Vidare har vi

$$\mathbb{P}(D) = 5/36,\ \mathbb{P}(A \cap D) = 1/36$$

vilket ger

$$\mathbb{P}(D|A) = \frac{\mathbb{P}(A \cap D)}{\mathbb{P}(A)} = \frac{1/36}{1/6} = \frac{1}{6} \neq \mathbb{P}(D)$$

så A och D är inte oberoende.

Proposition: Om A och B är oberoende, så är A och B^c oberoende.

Bevis:

$$\begin{aligned}\mathbb{P}(A \cap B^c) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(A)(1 - \mathbb{P}(B)) \\ &= \mathbb{P}(A)\mathbb{P}(B^c).\end{aligned}$$

Proposition: Fixera A och låt för varje $B \in \mathcal{P}(S)$,

$$Q(B) = \mathbb{P}(B|A).$$

Då är Q ett sannolikhetsmått.

Bevis:

(i) Eftersom $0 \leq \mathbb{P}(A \cap B) \leq \mathbb{P}(A)$ gäller att $0 \leq \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \leq 1$, dvs $0 \leq Q(B) \leq 1$.

(ii) $Q(S) = \mathbb{P}(S|A) = \frac{\mathbb{P}(A \cap S)}{\mathbb{P}(A)} = 1$.

(iii) Låt $B_1, B_2, \ldots$ vara disjunkta. Då gäller att

$$\begin{aligned} Q(\cup_n B_n) &= \mathbb{P}(\cup_n B_n|A) \\ &= \frac{\mathbb{P}(\cup_n (A \cap B_n))}{\mathbb{P}(A)} \\ &= \frac{\sum_n \mathbb{P}(A \cap B_n)}{\mathbb{P}(A)} \\ &= \sum Q(B_n). \end{aligned}$$

Alltså är en betingad sannolikhet verkligen en sannolikhet, så att t.ex.

$$\mathbb{P}(B_1 \setminus B_2|A) = \mathbb{P}(B_1|A) - \mathbb{P}(B_1 \cap B_2|A)$$

och

$$\mathbb{P}(B_1 \cup B_2|A) = \mathbb{P}(B_1|A) + \mathbb{P}(B_2|A) - \mathbb{P}(B_1 \cap B_2|A).$$

o.s.v.

En tankebild är att genom att betinga på att A inträffar så skär man bort A^c från utfallsrummet utan att förändra de inbördes sannolikheterna av delmängder till A.