F4 Grundläggande sannolikhetslära

Christian Tallberg

Avdelningen för Nationalekonomi/Statistik

Karlstads universitet

Vad menar vi med sannolikhet?

- *Sannolikhet* är ett kvantitativt mått på hur säkert det är att en händelse skall inträffa.
- Sannolikheten för en händelse är ett tal mellan 0 och 1.

Man kan (grovt sett) definiera sannolikheter på två sätt.

1. Sannolikhet definierat som relativa frekvenser (objektiva sannolikheter)

   Utifrån den frekventistiska definitionen talar man om sannolikheter att en viss *händelse* skall inträffa i samband med ett *slumpförsök*. Med ett slumpförsök menas ett experiment eller företeelse, som kan upprepas under identiska förhållanden, och där resultatet vid varje enskild upprepning inte kan förutsägas med säkerhet.

   - Tolkning: Sannolikheten för en händelse anger andelen gånger som händelsen inträffar vid en mycket lång serie upprepningar (ett oändligt antal) av slumpförsöket. Dvs, sannolikheten är den *relativa frekvensen* vid ett oändligt antal upprepade slumpförsök.

   Exempel på slumpförsök:

   - Tärningskast $(1, 2, 3, 4, 5, 6?)$
   - Lottdragning (Vinst eller ej?)
   - Slumpmässigt urval från en population (Vilka blir utvalda?)
   - Befruktning av äggcell (Pojke eller flicka?)
   - Radioaktivt sönderfall (Antal partiklar under ett visst tidsintervall?)
   - Industriell tillverkning av en enhet (Riktig eller felaktig?)

2. Subjektiv sannolikhetsdefinition

   Sannolikheter är *personliga* och mäter grad av övertygelse "degree of belief". Olika personer kan ha olika stark tilltro till ett och samma påstående. Vi *slipper begränsa* oss till händelser där sannolikheten kan fastställas med upprepade identiska experiment.

   Exempel:

   - Sannolikheten för regn i morgon.
   - Sannolikheten att Djurgården vinner allsvenskan i år.
   - Sannolikheten att det blir fred på jorden.
   - Sannolikheten att Gud existerar.

På den här kursen är det den *frekventistiska sannolikhetsuppfattningen* som gäller.

Utfallsrum för ett slumpförsök

- Resultatet av ett slumpförsök kallas för ett *utfall* (outcome).
- Mängden av möjliga utfall kallas *utfallsrummet* (sample space), och betecknas med $\mathbf{\Omega}$.

Utfallen måste vara definierade så att *ett och endast ett* utfall inträffar varje gång försöket utförs.

Exempel på utfallsrum:

- *Försök*: Kast med en tärning.

  Vi kan observera: Antal prickar

  Utfallsrum: $\mathbf{\Omega} = \{1, 2, 3, 4, 5, 6\}$

- *Försök*: Befruktning av en äggcell.

  Vi kan observera: Kön

  Utfallsrum: $\mathbf{\Omega} = \{pojke, flicka\}$

- *Försök*: Ett kast med ett mynt.

  Vi kan observera: Vilken sida som kommer upp.

  Utfallsrum: $\mathbf{\Omega} = \{krona, klave\}$

- *Försök*: Två kast med ett mynt.

  Vi kan observera: Resultat i första och andra kastet.

  Utfallsrum: $\mathbf{\Omega} = \{(kr, kr), (kr, kl)\,(kl, kr)\,(kl, kl)\}$

- *Försök*: Två kast med ett mynt.

  Vi kan observera: Antal krona.

  Utfallsrum: $\mathbf{\Omega} = \{0, 1, 2\}$

- *Försök*: Slumpmässigt välja ut två personer från en grupp av fyra personer: *A, B, C, D*.

  Vi kan observera: Vilka två som blir utvalda.

  Utfallsrum: $\mathbf{\Omega} = \left\{ \begin{array}{c} (A, B), (A, C), (A, D), \\ (B, C), (B, D), (C, D) \end{array} \right\}$

Händelse

- En delmängd av utfallen kallas en *händelse* (event). Dvs händelsen inträffar om och endast om något av utfallen i delmängden inträffar.

- Händelser betecknas med stora bokstäver, A,B, C, ...

Tärningsexempel 1: Ett kast med en symmetrisk tärning

Utfallsrummet är

$$\mathbf{\Omega} = \{1, 2, 3, 4, 5, 6\}$$

Exempel på händelser och delmängder:

- *Händelse*: $A =$ Att få ett udda antal prickar

  Delmängd: $A = \{1, 3, 5\}$

- *Händelse*: $B =$ Att få högst tre prickar

  Delmängd: $B = \{1, 2, 3\}$

- *Händelse*: $C =$ Att få ett jämnt antal prickar

  Delmängd: $C = \{2, 4, 6\}$

- *Händelse*: $D =$ Att inte få en sexa

  Delmängd: $D = \{1, 2, 3, 4, 5\}$

Mer om händelser

Med symboler och begrepp från mängdläran kan vi bilda nya händelser och uttrycka egenskaper hos händelser.

- *Unionhändelse*: $A \cup B$. Minst en av händelserna A eller B.

- *Snitthändelse*: $A \cap B$. Både A och B.

- *Komplementhändelse*: $A^* =$ icke-A.

- *Tomma mängden*: $\emptyset = \{\}$ (innehåller inga element, dvs $\emptyset$ är en omöjlig händelse som aldrig kan inträffa).

- Låt $A_1, A_2, ..., A_n$ vara delmängder i $\mathbf{\Omega}$. Om händelserna

  $$A_i \cap A_j = \emptyset$$

  kan de inte inträffa samtidigt. De kallas *disjunkta* eller *ömsesidigt uteslutande*.

Forts. tärningsexempel 1:

Då blir:

$$A \cup B = \{1, 2, 3, 5\}$$

$$A \cap B = \{1, 3\}$$

$$\begin{aligned} A^* &= \{2, 4, 6\} \\ B^* &= \{4, 5, 6\} \end{aligned}$$

$$A \cup A^* = \mathbf{S}$$

$$A \cap A^* = \emptyset$$

De Morgans lagar:

$$\begin{aligned} (A \cup B)^* &= A^* \cap B^* = \{4, 6\} \\ (A \cap B)^* &= A^* \cup B^* = \{2, 4, 5, 6\} \end{aligned}$$

I ett slumpförsök med utfallsrummet S motsvaras varje utfall av en sannolikhet. Sannolikheten att en händelse A skall inträffa får vi som summan av sannolikheterna för alla utfall som tillhör mängden A (dvs alla utfall som innebär att händelsen A inträffar). Sannolikheten att händelsen A_i skall inträffa betecknas med $P(A_i)$ och uppfyller följande axiom (Kolmogorov):

1. $0 \leqslant P(A) \leqslant 1$

2. $P(\mathbf{\Omega}) = 1$

3. Om $A_1, A_2, ...$är parvis oförenliga händelser gäller att

$$P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$$

4. $P(A^*) = 1 - P(A)$

5. $P(\emptyset) = 0$

Likformig sannolikhetsfördelning

- Om alla utfall i utfallsrummet har samma sannolikhet, har vi en *likformig* sannolikhetsfördelning.

- Om $\mathbf{\Omega}$ har m stycken utfall, så innebär en likformig sannolikhetsfördelning att sannolikheten för varje utfall (w_i) är $1/m$.

Exempel: Tärningskast med utfallsrummet: $\mathbf{\Omega} = \{1, 2, 3, 4, 5, 6\}$,

$$P(1) = ... = P(6) = \frac{1}{6}.$$

- Vid slumpförsök med likformig sannolikhetsfördelning kan sannolikheten för en händelse A beräknas så här

$$P(A) = \frac{g}{m},$$

där g är antalet utfall som är gynnsamma för A.

- Detta kallas ibland för den klassiska sannolikhetsdefinitionen, och gäller alltså bara för likformig sannolikhetsfördelning.

Fortsättning tärningsexempel 1:

$$
\begin{aligned}
P(A) &= \frac{g}{m} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} \\
P(B) &= \frac{g}{m} = \frac{3}{6} = \frac{1}{2} \\
P(A \cap B) &= \frac{g}{m} = \frac{2}{6} = \frac{1}{3} \\
P(A \cup B) &= \frac{g}{m} = \frac{4}{6} = \frac{2}{3} \\
P(A^*) &= 1 - P(A) = 1 - \frac{1}{2} = \frac{1}{2} \\
P(B^*) &= 1 - P(B) = 1 - \frac{1}{2} = \frac{1}{2}
\end{aligned}
$$

Elementär kombinatorik

Har man ibland nytta av vid sannolikhetsberäkningar.

Exempel:

Vi har en matsedel med tre förrätter, två huvudrätter och fyra efterrätter. På hur många sätt kan en trerätters måltid komponeras?

$$n_1 \cdot n_2 \cdot n_3 = 3 \cdot 2 \cdot 4 = 24$$

- *Kombinatoriskt problem 1*: Man skall i ordning utföra k operationer. Antalet möjligheter vid de olika operationerna är $n_1, n_2, ..., n_k$.

- *Multiplikationsprincipen*: Antal möjliga sätt att utföra de olika operationerna är $n_1 \cdot n_2 \cdot ... \cdot n_k$.

Exempel: På en personsökare finns fem lampor, numrerade 1 till 5, som oberoende av varandra kan vara släckta, tända med rött sken eller tända med grönt sken. Då en person söks, tänds minst en lampa, med en viss kombination för varje person. T ex

Lampa	1	2	3	4	5
	släckt	grön	grön	släckt	röd

Hur många personer kan sökas med denna anläggning?

$$3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 - 1 = 3^5 - 1 = 242$$

Ett arrangemang av n olika objekt i en *bestämd ordning* kallas för en *permutation* av objekten.

- *Kombinatoriskt problem 2*: Hur många permutationer (ordnade kombinationer) kan man välja av n olika element?

- Antalet olika sätt att permutera n olika element på är

  $n(n-1)(n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1 = n!$ (n-fakultet).

- $n!$ definieras enbar för heltal. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

- För $n = 0$ definieras $0! = 1$.

Exempel:

På hur många sätt kan man rangordna personerna Andersson, Bergman och Cesar?

Extrauppgift (lös på egen hand): På hur många sätt kan man permutera n element, om vissa av dem är lika? T ex, hur många ord med nio bokstäver kan man bilda av ordet *statistik*? Svar: 15120.

- *Kombinatoriskt problem 3*: På hur många sätt kan vi välja ut r objekt från n objekt och *ordna dem* ($r \leq n$)?

- Antalet *ordnade delmängder* när r element väljs från n element är

$$n\,(n-1)\cdot\ldots\cdot(n-r+1) = \frac{n!}{(n-r)!}$$

Exempel:

Bland fyra personer skall två väljas ut och rangordnas. På hur många sätt kan detta ske?

$n = 4$; $r = 2$

- *Kombinatoriskt problem 4*: På hur många sätt kan vi välja ut r objekt från n objekt ($r \leq n$) om vi struntar i ordningen?

- Antalet (*icke-ordnade delmängder*) när r element väljs från n element är

$$\frac{n!}{r!\,(n-r)!} = \binom{n}{r}$$ (uttalas "n över r") eller binomialkoefficienten.

Exempel:

Bland fyra personer skall två väljas ut (Obs! ingen rangordning). På hur många sätt kan detta ske?

$n = 4$; $r = 2$

Extrauppgift (lös på egen hand): a) I en förening bestående av tio män och femton kvinnor skall en styrelse av fem väljas ut (Obs! ingen rangordning). På hur många sätt kan detta ske? Svar: 53130

b) Hur många sätt kan man välja styrelse så att två är män och tre är kvinnor? Svar: 45455

c) Vad är sannolikheten att styrelsen kommer bestå av tre kvinnor och fyra män? Svar: 0.385

Sammanfattning

	Med återl.	Utan återl.
Med häns t ordn	n^r	$\frac{n!}{(n-r)!}$
Utan häns t ordn	$\binom{n+r-1}{r}$	$\binom{n}{r}$

Ännu mer om händelser

Tärningsexempel 2: *Två* kast med en symmetrisk sexsidig tärning.

$A =$ Högst två prickar i första kastet

$B =$ Minst sju prickar sammanlagt

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$
\begin{aligned}
P(A) &= \frac{g}{m} = \frac{12}{36} = \frac{1}{3} = 0.33 \\
P(B) &= \frac{g}{m} = \frac{21}{36} = \frac{7}{12} = 0.58 \\
P(A \cap B) &= \frac{3}{36} = \frac{1}{12} = 0.08 \\
P(A \cup B) &= \frac{30}{36} = \frac{5}{6} = 0.83 \\
P(A \cap B^*) &= P(A) - P(A \cap B) = \frac{1}{3} - \frac{1}{12} \\
&= \frac{4}{12} - \frac{1}{12} = \frac{1}{4} \\
P(A^* \cap B) &= P(B) - P(A \cap B) = \frac{7}{12} - \frac{1}{12} \\
&= \frac{6}{12} = \frac{1}{2} \\
P(A^* \cap B^*) &= P(A \cup B)^* = 1 - P(A \cup B) \\
&= 1 - \frac{5}{6} = \frac{1}{6}
\end{aligned}
$$

Additionssatsen

Om $P(A), P(B)$ och $P(A \cap B)$ är kända, så kan sannolikheten att minst en av händelserna A och B inträffar beräknas med hjälp av *additionssatsen*:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

I tärningsexempel 2 kan $P(A \cup B)$ då beräknas som

$$P(A \cup B) = \frac{12}{36} + \frac{21}{36} - \frac{3}{36} = \frac{5}{6} = 0.83$$

För *disjunkta* (ömsesidigt uteslutande) händelser gäller att

$$P(A \cup B) = P(A) + P(B).$$

Betingad sannolikhet

Ibland vill vi veta hur stor sannolikheten är för en händelse B ifall vi vet att en händelse A redan har inträffat. Detta kallas för den *betingade* sannolikheten för B, givet att A har inträffat, och ges av

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \qquad P(A) > 0.$$

På motsvarande ges en *betingade* sannolikheten för A, givet att B har inträffat av

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) > 0.$$

Exempel:

	Rökare	**Icke-rökare**	**Summa**
Män	320	360	680
Kvinnor	72	248	320
Summa	392	608	1000

En person väljs slumpmässigt från populationen (dvs med lika sannolikhet för samtliga att bli valda). Vad är sannolikheten att personen

a) är rökare?

b) är en man?

c) är rökare och är en man?

d) är rökare om vi vet att det är en man?

e) är rökare om vi vet att det är en kvinna?

f) är en man om vi vet att det är en rökare?

Låt A = händelsen "rökare" och B = händelsen "Man"

$$\begin{aligned}
P(A) &= \frac{392}{1000} = 0.39 \\
P(B) &= \frac{680}{1000} = 0.68 \\
P(A \cap B) &= \frac{320}{1000} = 0.32 \\
P(A|B) &= \frac{320}{680} = 0.47 \\
P\left(A \middle| \bar{B}\right) &= \frac{72}{320} = 0.22 \\
P(B|A) &= \frac{320}{392} = 0.82
\end{aligned}$$

Fortsättning tärningsexempel 2: Beräkna $P(A|B) =$

$$\begin{aligned}
& P(\text{Högst två pr i kast 1} \,|\text{Minst sju pr sammanl}) \\
=\ & \frac{P(\text{Högst två pr i kast 1 och Minst sju pr sammanl})}{P(\text{Minst sju prickar sammanlagt})} \\
=\ & \frac{3}{21} = \frac{1}{7} = 0.14
\end{aligned}$$

Multiplikationssatsen

Om $P(A), P(B|A)$ är kända, så kan vi beräkna $P(A \cap B)$. Från definitionen av betingad sannolikhet följer nämligen att

$$P(A \cap B) = P(A) P(B|A).$$

Om $P(B), P(A|B)$ är kända, beräknas $P(A \cap B)$ istället från

$$P(A \cap B) = P(B) P(A|B).$$