Grundläggande matematisk statistik

Axiomsystem, Betingad sannolikhet, Total sannolikhet, Bayes sats

Uwe Menzel, 2018
uwe.menzel@matstat.org
www.matstat.org

1

Slumpförsök

Ett slumpförsök är ett försök där utgången inte kan förutses eller kontrolleras:
- kasta en tärning
- dra en kula ur en urna (med flera olikfärgade kulor)
- myntkast (klave eller krona)
- blodtryck av en slumpmässigt vald person
- livslängd av en glödlampa
- radioaktivt sönderfall: antalet signaler inom en minut
- mätning av diameter, temperatur, koncentration, …

Utfall

Resultatet av ett slumpförsök kallas ett utfall:
- ögontalet efter tärningskast: $\omega = \{1\}$
- roulette: $\omega = \{24\}$
- myntkast: $\omega = \{krona\}$
- radioaktivt sönderfall: : $\omega = \{109\}$
- kast med två tärningar: $\omega = \{(1, 5)\}$
- mätt livslängd av en glödlampa: $\omega = 1245.6\ timmar$
- opinionsundersökning: en tillfälligt utvald väljare kommer att rösta på parti A

2

Utfallsrum

Mängden av alla möjliga utfall av ett slumpförsök kallas utfallsrum:
- ögontalet efter tärningskast: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- roulette: $\Omega = \{1, 2, 3, . . , 36\}$
- myntkast: $\Omega = \{0, 1\}$ $\quad$ ($0 =$ klave; $1=$ krona)
- kast med två tärningar: : $\Omega = \{(1,1),\ (1,2),\ (1,3)\ \ldots\ (6,6)\}$
- livslängd av en glödlampa: $\Omega = \{0, +\infty\}$ $\quad$ (?)

Händelse

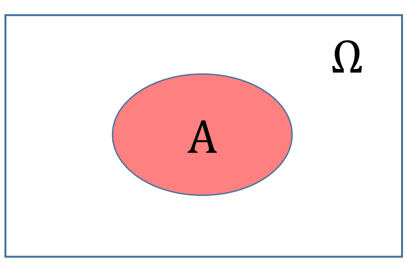
En händelse är en samling av utfall (en delmängd av utfallsrummet):
- tärningskast: $A = \{2, 4, 6\}$ (jämt ögontal)
- (händelse A inträffar om **ett** utfall som tillhör A inträffar)
- tärningskast: $B = \{1, 2, 3\}$ (ögontalet mindre än 4)
- kast med två tärningar: : $A = \{(1,1),\ (1,2,),\ (2,1)\}$ (summan av ögonen högst 3)
- radioaktivt sönderfall: : $C = \{0,\ 1,\ 2,\ \ldots, 10\}$ $\quad$ (högst 10 sönderfall)

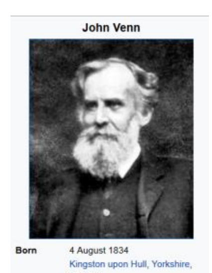
www.matstat.org

3

Venn-diagram

Utfallsrummet och händelser kan illustreras med hjälp av ett Venn-diagram:





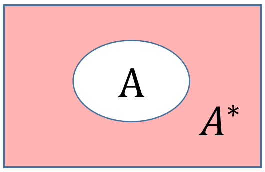
- rektangel: hela utfallsrummet (Ω)
- ellips: en händelse (A; del av utfallsrummet)

www.matstat.org

4

Sammansättning av händelser

Komplementhändelse:

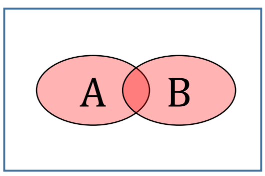


exempel (tärning) :
$A = \{2, 4, 6\}$ (jämnt tal)

komplementhändelse:
$A^* = \{1, 3, 5\}$ (ojämnt tal)

händelse och komplementhändelse ”täcker” utfallsrummet, antingen A eller A* måste inträffa

Union av händelser:



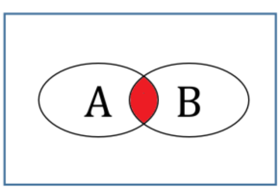
$A = \{2, 4, 6\}$
$\mathrm{B} = \{1, 2, 3\}$

union:
$A \cup B = \{1, 2, 3, 4, 6\}$
”A eller B eller båda inträffar”
(inte helt detsamma som i vardagsspråket)

5

Sammansättning av händelser

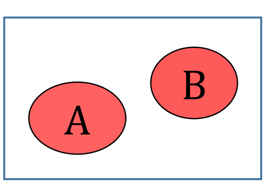
Snitthändelse:



$A = \{2, 4, 6\}$
$\mathrm{B} = \{1, 2, 3\}$

snitt:
$A \cap B = \{2\}$
”både A och B inträffar”

Disjunkta (oförenliga) händelser:



$A = \{1, 6\}$
$\mathrm{B} = \{4, 5\}$

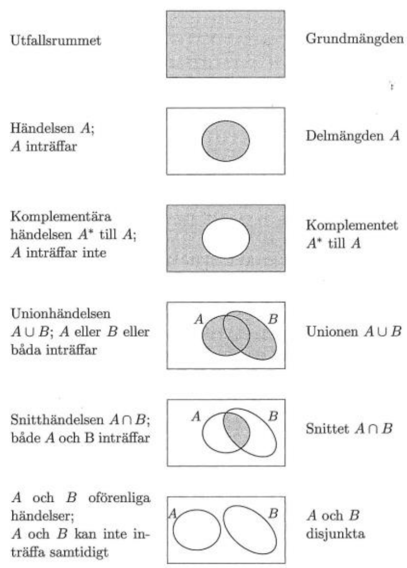
snitt:
$A \cap B = \emptyset$
inga gemensamma utfall

www.matstat.org

6

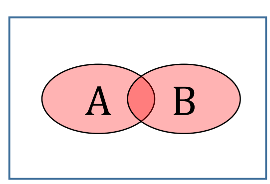
Jämförelse med mängdlära*

Bild tagen från:
G. Blom et al.



7

De Morgans lagar*



$(A \cap B)^* = A^* \cup B^*$

$(A \cup B)^* = A^* \cap B^*$

Kan ses genom att finna respektive ytor i Venn-diagrammet

De Morgans lagar gäller också för fler än två händelser.

www.matstat.org

8

Den klassiska sannolikhetsdefinitionen

$\Omega = \{u_1, u_2, u_3, \ldots, u_m\}$ utfallsrummet

$P(u_i) = {}^{1}/_{m}$ antar att alla möjliga utfall u_i är lika sannolika!

m = antalet möjliga utfall
g = antalet utfall som ingår i händelse A (= antalet ”gynsamma” utfall)

Sannolikheten (slh.) för händelse A betecknas med $P(A)$:

$$\boxed{P(A) = \frac{g}{m}}$$ klassiska sannolikhetsdefinitionen

Exempel 1: tärning, söker slh. att ögontalet blir ett jämt tal:

$$A = \{2, 4, 6\} \Rightarrow P(A) = \frac{g}{m} = \frac{3}{6} = \underline{\underline{\frac{1}{2}}}$$

www.matstat.org

9

Den klassiska sannolikhetsdefinitionen

Begrepp slumpvariabel:

- En slumpvariabel (s.v.) är en storhet som är obestämd **före** slumpförsöket och erhåller ett värde **efter** försöket.

Exempel 2: två tärningar kastas samtidigt

slumpvariabel X: ögontalet för tärning 1
slumpvariabel Y: ögontalet för tärning 2

Låt händelse A vara: $X + Y \leq 4$ (summan av ögontalet är mindre eller lika med 4)

Söker:
$P(A) = P(X + Y \leq 4)$ (slh. att summan av ögontalet är mindre eller lika med 4)

Vi behöver g (antalet ”gynsamma” utfall) och m (antalet ”möjliga” utfall)

www.matstat.org

10

Den klassiska sannolikhetsdefinitionen

Fortsättning ex 2: två tärningar kastas samtidigt. Söker $P(A) = P(X + Y \leq 4)$:

Vi behöver g (antalet ”gynsamma” utfall) och m (antalet ”möjliga” utfall).
I tabellen visas summan $X + Y$ för varje värde av X och Y:

Y\X	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

tärning 1

tärning 2

g = 6 (grön)
m = 36

Svar:

$$P(A) = P(X + Y \leq 4) = \frac{g}{m} = \frac{6}{36} = \frac{1}{6}$$

11

Den klassiska sannolikhetsdefinitionen

Exempel 3: En urna innehåller 100 kulor; därav 90 vita och 10 blåa.

Slumpförsök: dra en kula och kom ihåg färgen.
Slumpvariabel X: färgen på den dragna kulan.
Låt **händelse B** vara: den dragna kulan är blå.
Låt **händelse V** vara: den dragna kulan är vit.
$\boldsymbol{P(B)}$: sannolikheten att händelse B inträffar, dvs. att den dragna kulan är blå.
$\boldsymbol{P(V)}$: sannolikheten att händelse V inträffar, dvs. att den dragna kulan är vit.

Svar:

$$P(B) = P(X = blå) = \frac{g}{m} = \frac{10}{100} = \frac{1}{10} = 10\%$$

$$P(V) = P(X = vit) = \frac{g}{m} = \frac{90}{100} = \frac{9}{10} = 90\%$$

Anmärkning: Summan av båda slh:na är ett, för att:

- båda händelser utesluter varandra
- en av de två händelserna **måste** inträffa

$$P(B) + P(V) = 1$$

12

Kombinatorik*

... behövs för att räkna ut antalet gynsamma utfall (g) och antalet möjliga utfall (m).

1. Permutationer
2. Multiplikationsprincipen
3. Dragning med återläggning och med hänsyn till ordning
4. Dragning utan återläggning men med hänsyn till ordning
5. Dragning utan återläggning utan hänsyn till ordning

www.matstat.org

13

1.) Permutationer

På hur många sätt kan man anordna (sortera) k olika element ?

Exempel: numrerade kulor 1 2 3

$k = 2$

$$\left.\begin{array}{l} 1,2 \\ 2,1 \end{array}\right\} 1 \cdot 2 = 2 \text{ möjligheter}$$

$k = 3$

$$\left.\begin{array}{l} 1,2,3 \\ 1,3,2 \\ 2,1,3 \\ 2,3,1 \\ 3,1,2 \\ 3,2,1 \end{array}\right\} 1 \cdot 2 \cdot 3 = 6 \text{ möjligheter}$$

Allmänt: k element kan anordnas på k! (”k fakultet”) olika sätt.

$$M = k! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot k$$

Anmärkning: Den exakta terminologin för detta är ”antalet permutationer av k element bland k”. I föreläsningen använder jag bara uttrycket ”permutationer”.

www.matstat.org

14

2.) Multiplikationsprincipen

Multiplikationsprincipen:
Om åtgärd 1 kan utföras på n_1 sätt och åtgärd 2 på n_2 sätt, så finns det $n_1 \cdot n_2$ sätt att kombinera båda åtgärderna. Generalisering på 3 åtgärder blir $n_1 \cdot n_2 \cdot n_3$ osv.

$$\left.\begin{array}{ll} \text{åtgärd 1} & n_1 \text{ möjligheter} \\ \text{åtgärd 2} & n_2 \text{ möjligheter} \\ \text{åtgärd 3} & n_3 \text{ möjligheter} \\ \ldots & \\ \ldots & \\ \text{åtgärd k} & n_k \text{ möjligheter} \end{array}\right\}$$

antalet kombinerade åtgärder:

$M = n_1 \cdot n_2 \cdot n_3 \cdot \ldots \cdot n_k$

Exempel 1: 2 tärningar

$$\left.\begin{array}{l} n_1 = 6 \\ n_2 = 6 \end{array}\right\} \quad M = 6 \cdot 6 = \underline{\underline{36}}$$

www.matstat.org

15

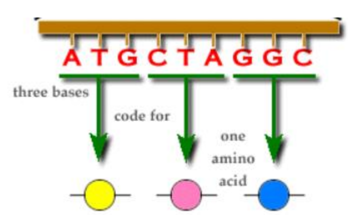
2.) Multiplikationsprincipen

Exempel 2: svenskt bilnummer

26 bokstäver
10 siffror

$$M = 26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 26^3 \cdot 10^3 = \underline{\underline{17.576.000}}$$

Exempel 3: DNA-triplett kvävebaser: A, C, G, T



$$\underbrace{\boxed{?} \quad \boxed{?} \quad \boxed{?}}_{}$$

$$\underbrace{n_1 = 4 \quad n_2 = 4 \quad n_3 = 4}$$

$$M = 4^3 = \underline{\underline{64}}$$

En DNA-triplett kan teoretiskt koda 64 olika aminosyror. (Det finns dock bara 20.)

www.matstat.org

16

2.) Multiplikationsprincipen

Exempel 4: Hur många 4-siffriga decimaltal finns det?

$$\underbrace{\overset{\boxed{?}}{n_1 = 10} \quad \overset{\boxed{?}}{n_2 = 10} \quad \overset{\boxed{?}}{n_3 = 10} \quad \overset{\boxed{?}}{n_4 = 10}}_{\mathrm{M} = 10^4 = 10.000}$$

Detta är ett speciellt fall av multiplikationsprincipen, där det finns **samma** antal valmöjligheter (n) i varje kategori. Antalet kombinerade åtgärder blir alltså

$$\boxed{\mathbf{M} = \boldsymbol{n^k}}$$

17

Olika antaganden

I alla föregående exempel antog vi att:

- elementen får upprepas, t. ex. siffran 1215 kan förekomma (ettan förekommer 2 gånger)
- ordning spelar roll, dvs. objekt betraktas som olika om elementen förekommer i annorlunda ordningsföljd, t. ex. ansåg vi händelse $\{1, 2, 3\}$ och händelse $\{1, 3, 2\}$ som två enskilda händelser

Vi ska nu med hjälp av en urnmodell undersöka hur stor antalet möjliga åtgärder blir om:

1. element får upprepas och ordningsföljd räknas, som ovan (”dragning med återläggning, med hänsyn till ordning”)
2. element får **icke** upprepas men ordningsföljd räknas fortfarande (”dragning utan återläggning, men med hänsyn till ordning”)
3. element får **icke** upprepas och ordningsföljd räknas **inte** (” dragning utan återläggning och utan hänsyn till ordning”)

18

3.) Dragning med återläggning, med hänsyn till ordning

Man kan föreställa sig en urna med n numrerade kulor.

Slumpförsok:
- dra en kula, lägg tillbaka
- dra en andra kula, lägg tillbaka,
- osv.
- dra k kulor totalt

Fråga: Hur många olika sifferkombinationer kan man få om dessutom ordningsföljden räknas?

Svar: ”Med återläggning” betyder att vi har n möjligheter (siffror) i varje dragning. Om vi drar k kulor totalt blir alltså antalet olika sifferkombinationer:

$\boxed{\mathbf{M} = \boldsymbol{n^k}}$ (enligt multiplikationsprincipen)

Låt n = 9 och k = 3 → möjliga resultat är t. ex.:
112; 113; 123; 321; 478; 444; 978; 789; ...

OBS: Siffrorna kan upprepas, dessutom räknas ordningen, dvs. både 978 och 789 räknas som två enskilda händelser.

www.matstat.org

19

4.) Dragning utan återläggning, med hänsyn till ordning

Slumpförsok:
- dra en kula, lägg inte tillbaka
- dra en andra kula, lägg inte tillbaka,
- osv.
- dra k kulor totalt (lägg inte tillbaka någon kula)

Fråga: Hur många olika sifferkombinationer kan man få om ordningsföljden räknas?

Svar: ”Utan återläggning” betyder att siffrorna inte kan upprepas, dvs. i första dragningen har man n möjligheter; i andra bara $(n-1)$ möjligheter, i 3:e dragningen bara $(n-2)$ möjligheter osv. Antalet åtgärder blir:

$$\boxed{\mathbf{M} = \boldsymbol{n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}}}$$

Låt n = 9 och k = 3 → möjliga resultat är t. ex. (strykna tal kan ej förekomma!):
~~112~~; ~~113~~; 123; 321; 478; ~~444~~; 978; 789; ...

OBS: Ordningen räknas fortfarande, dvs. 978 och 789 räknas som enskilda händelser

www.matstat.org

20

5.) Dragning utan återläggning, utan hänsyn till ordning

Slumpförsok:
- dra en kula, lägg inte tillbaka
- dra en andra kula, lägg inte tillbaka,
- osv.
- dra k kulor totalt (lägg inte tillbaka någon kula)

Fråga: Hur många olika sifferkombinationer kan man få om ordningsföljden inte räknas (t. ex. om kulorna ändå blir sorterade efteråt)?

Svar: Siffrorna inte kan upprepas – som ovan. Dessutom måste alla möjligheter som bara skiljer sig på ordningsföljden tas bort – detta är k! (”k fakultet”).
Vi måste därför dela det ovanstående antalet med k!

$$\boxed{\mathbf{M} = \frac{\boldsymbol{n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)}}{\boldsymbol{k!}} = \frac{\boldsymbol{n!}}{\boldsymbol{k! \cdot (n-k)!}} = \boldsymbol{\binom{n}{k}}}$$ Binomialkoefficient

Låt n = 9 och k = 3 → möjliga resultat är t. ex. (strykna tal kan ej förekomma!):
~~112~~; ~~113~~; 123; ~~321~~; 478; ~~444~~; 978; ~~789~~; …

www.matstat.org

21

Dragning utan återläggning, utan hänsyn till ordning

Exempel: dragning utan återläggning (utan repetition), utan hänsyn till ordning: n = 35; k = 7 (35 siffror, 7 kommer att dras)

$$\mathrm{M} = \binom{n}{k} = \binom{35}{7} = 6.724.520$$

Det finns fler än 6.7 miljoner möjligheter att dra 7 siffror utav 35, om återläggning inte sker och ordningsföljd inte räknas.
(= antalet möjligheter på **Lotto**)

www.matstat.org

22

Exempel

Blom et al., s. 23

Ett femsiffrigt slumptal väljs (00000; 00001; 00003; … ; 99999).

a) Hur stor är slh. att talet bara innehåller 0 och 1?
b) Hur stor är slh. att alla siffror är olika?

<u>Svar</u>: Alla möjliga ”dragna” tal kan anses som lika sannolika. Vi kan alltså använda den klassiska sannolikhetsdefinitionen:

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

Antalet möjliga utfall: $\mathrm{m} = n^k = 10^5$ för att vi väljer 5 gånger och varje gång bland 10 siffror ($n = 10$ och $k = 5$). Siffrorna kan uprepas och ordningsföljden räknas.

a) Vi har två gynsamma möjligheter (0 och 1) för varje siffra (5 siffror). Antalet gynsamma utfall är alltså $g = n^k = 2^5$. Sannolikheten att talet bara innehåller 0 och 1 är därmed $P(A) = \frac{g}{m} = \frac{2^5}{10^5} = \underline{\underline{0.00032}}$

b) Vi ”drar” 5 siffror. Först väljer vi mellan 10 siffror, sedan bara mellan 9 (den 1:a får ju inte längre förekomma), sedan mellan 8, osv…
Antalet gynsamma utfall är alltså
$g = n \cdot (n-1) \cdot \ \dots \cdot (n-k+1) = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 30240.$
Slh. att alla siffror är olika är därmed $P(A) = \frac{g}{m} = \frac{30240}{10^5} = \underline{\underline{0.3024}}$

23

Urnmodell, sammanfattning

- n numrerade kulor
- k kulor dras

Antaganden	Antalet möjligheter	exempel
med återläggning, med hänsyn till ordning	$\mathbf{M} = \boldsymbol{n^k}$	kombinationslås
utan återläggning, med hänsyn till ordning	$\mathbf{M} = \frac{\boldsymbol{n}!}{(\boldsymbol{n}-\boldsymbol{k})!}$	se exempel ovan
utan återläggning, utan hänsyn till ordning	$\boldsymbol{M} = \binom{\boldsymbol{n}}{\boldsymbol{k}}$	lotto

www.matstat.org

24

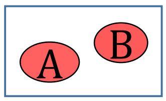
Kolmogorows axiomer

A: händelse
$P(A)$: sannolikhet att händelse A inträffar

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$ om A, B disjunkta

Axiom 3 gäller också för fler än 2 händelser

axiom 3:



$A \cap B = \emptyset$

<u>Exempel</u>: axiom 3, tärning $\quad A = \{1, 2\}$
$B = \{5, 6\}$ (A och B är disjunkta)

$A \cup B = \{1, 2, 5, 6\} \Rightarrow P(A \cup B) = \frac{g}{m} = \frac{4}{6}$

$P(A) = \frac{g}{m} = \frac{2}{6} \quad P(B) = \frac{g}{m} = \frac{2}{6} \Rightarrow P(A \cup B) = P(A) + P(B)$ ✓

25

Komplementsats

Axiom 3: $P(A \cup B) = P(A) + P(B)$ om $A \cap B = \emptyset$ (disjunkt)

Låt $B = A^*$ (gäller ju allmänt)

$P(A \cup A^*) = P(A) + P(A^*)$ från axiom 3

$P(\Omega) = P(A) + P(A^*)$ ty $\Omega = A \cup A^*$

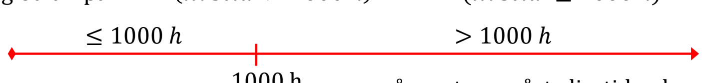
$1 = P(A) + P(A^*)$ ty $\Omega = A \cup A^*$ axiom 2

$\boxed{P(A^*) = 1 - P(A)}$ komplementsats (används mycket ofta!)

<u>Ex1</u>, tärning: A: får en etta
A*: får inte en etta

$P(etta) = 1 - P(ingen\ etta)$
$\frac{1}{6} = 1 - \frac{5}{6}$

<u>Ex2</u>, glödlampa: $P(livstid > 1000\ h) = 1 - P(livstid \leq 1000\ h)$

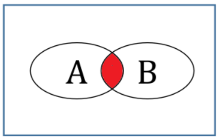


någonstans måste livstiden hamna

26

Additionssats

$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$ gäller allmänt



$P(A) + P(B)$: snittet räknas två gånger, måste subtraheras en gång

<u>Exempel</u>, tärning: $A = \{1, 2, 3\}$ $\quad B = \{3, 4, 5\}$

$A \cap B = \{3\} \quad A \cup B = \{1, 2, 3, 4, 5\}$

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

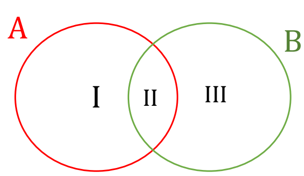
$\frac{5}{6} = \frac{3}{6} + \frac{3}{6} - \frac{1}{6}$ stämmer! ✓

www.matstat.org

27

Additionssats

<u>Härledning</u> med hjälp av axiomerna:



I, II, III disjunkta

$P(A \cup B) = P(I) + P(II) + P(III)$ axiom 3

$= P(I) + P(II) + P(III) + P(II) - P(II)$ ("nolladdition")

$= \underbrace{P(I) + P(II)}_{P(A)} + \underbrace{P(III) + P(II)}_{P(B)} - P(A \cap B)$ q. e. d.

$P(A \cup B) \leq P(A) + P(B)$ Booles olikhet

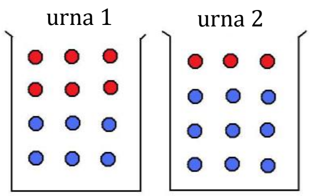
www.matstat.org

28

Betingad sannolikhet

beräkna slh. för händelse B om man vet att händelse A redan har inträffat
Symbol: $\boldsymbol{P(B|A)}$ "sannolikhet för B givet A"

<u>Exempel</u> : dra kulor ur två urnor



urna 1: 6 röda av 12
urna 2: 3 röda av 12

steg 1: väljer slumpmässigt en urna $\Omega_U = \{1, 2\}$
steg 2: väljer slumpmässigt en kula ur denna urna $\Omega_F = \{röd, blå\}$

$P(F = röd \mid U = 1) = \dfrac{6}{12}$ slh. att dra en röd kula givet att urna 1 valts

$P(F = röd \mid U = 2) = \dfrac{3}{12}$ slh. att dra en röd kula givet att urna 2 valts

29

Betingad sannolikhet

<u>Exempel</u> : (Blom et al. s. 22) kortspel: 52 kort, 4 ess

- händelse **A**: dra ett ess i första dragningen
- händelse **B**: dra ett ess i andra dragningen

$P(A) = \dfrac{g}{m} = \dfrac{4}{52}$

$P(B|A) = \dfrac{g}{m} = \dfrac{3}{51}$ 2:a dragningen: 51 kort kvar, därav 3 ess (A har inträffat, dvs. ett ess är borta)

$P(B|A^*) = \dfrac{g}{m} = \dfrac{4}{51}$ 2:a dragningen: 51 kort kvar, därav 4 ess (A har inte inträffat, dvs. alla ess är kvar)

$P(B) = ?$ total sannolikhet, se nedan

30

Betingad sannolikhet

Exempel : två tärningar
- slumpvariabel **X** : ögontalet för tärning 1
- slumpvariable **Y** : ögontalet för tärning 2

$P(X+Y=10 \mid X=5)=\frac{1}{6}$ (Y **måste** vara 5 för att summan blir 10)

$P(X+Y=10)=\frac{3}{6}$ tre gynnsamma utfall: (5, 5) ; (4, 6) ; (6, 4)

$P(X+Y=10 \mid X=1)=0$ (om X bara är ett kan summan inte bli 10)

www.matstat.org

31

Betingad sannolikhet

$\boxed{P(A|B)=\frac{P(A\cap B)}{P(B)}}$ Definition "**betingad sannolikhet**"

Exempel : tärning
- händelse B : jämnt ögontal
- händelse A : ögontalet blir två

$\left.\begin{array}{l} B=\{2,4,6\} \\ \mathrm{A}=\{2\} \end{array}\right\} A\cap B=\{2\}$

$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{1/6}{1/2}=\frac{1}{6}\cdot\frac{2}{1}=\frac{1}{3}$

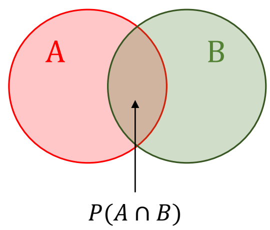
Om man vet att ögontalet blivit jämnt så är slh. 1/3 att ögontalet blev en tvåa

www.matstat.org

32

Betingad sannolikhet

Intuitiv förklaring:



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

Om man vet att B har inträffat, reduceras utfallsrummet till den arean som symboliserar händelse B (grön). Sannolikheten att händelse A inträffar under dessa omständigheter är kvoten $P(A \cap B)/P(B)$**.** Detta är alltså lika med den betingade sannolikheten $P(B|A)$**.**

www.matstat.org

33

Betingad sannolikhet

Exempel : två tärningar

- slumpvariabel **X** : ögontalet för tärning 1
- slumpvariable **Y** : ögontalet för tärning 2

OBS!: Ibland skrivs ”komma” istället för ”∩”

$$P(X+Y=10 \mid X=5) = \frac{P(X+Y=10\,,\; X=5)}{P(X=5)} = \frac{P(X=5, Y=5)}{P(X=5)} = \frac{{}^{1}/_{36}}{{}^{1}/_{6}} = \underline{\underline{\frac{1}{6}}}$$

... samma som vi fick ovan

www.matstat.org

34

Betingad sannolikhet

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

Definition ”betingad sannolikhet”

Efter omformning:

$P(A \cap B) = P(B) \cdot P(A|B)$ Om A och B ska inträffa, måste först B inträffa, sedan måste A inträffa givet att B redan har inträffat.

$P(A \cap B) = P(A) \cdot P(B|A)$ ... eller tvärtom.

<u>Exempel</u> : (Blom et al. s. 22) kortspel: 52 kort, 4 ess

- händelse **A**: dra ett ess i första dragningen
- händelse **B**: dra ett ess i andra dragningen

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{4}{52} \cdot \frac{3}{51} = \underline{\underline{0.0045}}$$

www.matstat.org

35

Betingad sannolikhet

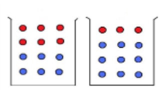
Betingade sannolikheter är sannolikheter, alla lagar gäller också för dem:

$P(A^*) + P(A) = 1$ komplementsats

$P(A^*|\,S) + P(A|\,S) = 1$ ”conditioning through”

När händelse S har inträffat måste antingen A eller A* inträfffa.

<u>Exempel</u> : två urnor



steg 1: väljer slumpmässigt en urna $\Omega_U = \{1, 2\}$
steg 2: väljer slumpmässigt en kula ur denna urna $\Omega_F = \{röd, blå\}$

Om det är givet att urna 1 valts, så måste färgen antingen blir röd eller blå:

$$P(\underbrace{blå}_{A^*} \mid \underbrace{U = 1}_{S}) + P(\underbrace{röd}_{A} \mid \underbrace{U = 1}_{S}) = 1$$

36

Oberoende händelser

händelser A, B

$\boxed{P(B \mid A) = P(B)}$ när detta gäller kallas A och B oberoende

Exempel : $P(L \mid R) = P(L)$

L: vinner på lotto R: regnar i Stockholm

Exempel : två tärningar
- slumpvariabel X : ögontalet tärning 1
- slumpvariabel Y : ögontalet tärning 2

$P(Y = 2 \mid X = 1) = P(Y = 2) = \frac{1}{6}$ $\Rightarrow$ X, Y är oberoende

37

Oberoende händelser

Om händelser A och B är oberoende:

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

$\boxed{P(A \cap B) = P(A) \cdot P(B)}$ om A, B oberoende (används mycket ofta!)

gäller även för flera händelser:

$\boxed{P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)}$ om A, B, C är parvis oberoende

Exempel : två tärningar

- slumpvariabel **X** : ögontalet för tärning 1
- slumpvariable **Y** : ögontalet för tärning 2

$P(X = 1, Y = 5) = P(X = 1) \cdot P(Y = 5) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

pga. oberoendet

38

Oberoende händelser

Oberoendet leder till flera följdsatser:

$P(B \mid A) = P(B)$ ⇨ $P(A \mid B) = P(A)$ bevisas med hjälp av definitionen för betingad slh.)

⇨ $P(A^* \mid B) = P(A^*)$

⇨ $P(B^* \mid A) = P(B^*)$

⇨ $P(A^* \mid B^*) = P(A^*)$ komplementära händelser är också oberoende

<u>OBS</u>!: Blanda inte ihop oförenliga (disjunkta) och oberoenede händelser:

	Snitt	Betingade slh.
Oförenliga	$P(A \cap B) = \emptyset$	$P(A\|B) = 0$
Oberoende	$P(A \cap B) = P(A) \cdot P(B)$	$P(A\|B) = P(A)$

går inte ihop går inte ihop

39

Oberoende händelser

<u>Exempel :</u> tärning, Blom s. 32

$A = \{1, 3, 5\}$ ojämnt tal $P(A) = {}^1/_2$

$B = \{4, 5, 6\}$ minst 4 $P(B) = {}^1/_2$

$C = \{3, 6\}$ delbar med 3 $P(C) = {}^1/_3$

a) Är A och B oberoende ?

$A \cap B = \{5\}$ $P(A \cap B) = {}^1/_6 \neq P(A) \cdot P(B)$ $\rightarrow$ beroende !

b) Är A och C oberoende ?

$A \cap C = \{3\}$ $P(A \cap C) = {}^1/_6 = P(A) \cdot P(C)$ $\rightarrow$ oberoende !

40

Oberoende händelser

<u>Exempel</u>: tre elektroniska komponenter K_A, K_B, K_C

- händelse A: K_A slutar fungera $P(A) = 0.2$
- händelse B: K_B slutar fungera $P(B) = 0.05$
- händelse A: K_C slutar fungera $P(C) = 0.1$

inom det första året

Vi antar att händelserna A, B och C är parvis oberoende.

a) Slh. att alla tre slutar fungera inom det första året ?

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) = 0.2 \cdot 0.05 \cdot 0.1 = \underline{\underline{0.001}}$$

b) Slh. att K_A och K_B slutar fungera, men inte K_C ?

$$P(A \cap B \cap C^*) = P(A) \cdot P(B) \cdot P(C^*) = P(A) \cdot P(B) \cdot [1 - P(C)] = 0.2 \cdot 0.05 \cdot 0.9$$

www.matstat.org

41

Oberoende händelser

c) Slh. att exakt två komponenter slutar fungera ?

$$P(exakt\ 2) = P[(A \cap B \cap C^*) \cup (A \cap B^* \cap C) \cup (A^* \cap B \cap C)]$$

”eller” ”eller”

händelserna i parenteserna är disjunkta

$$= P(A \cap B \cap C^*) + P(A \cap B^* \cap C) + P(A^* \cap B \cap C)$$ axiom 3

$$= P(A) \cdot P(B) \cdot P(C^*) + P(A) \cdot P(B^*) \cdot P(C) + P(A^*) \cdot P(B) \cdot P(C)$$ pga. oberoendet

$$= 0.2 \cdot 0.05 \cdot 0.9 + 0.2 \cdot 0.95 \cdot 0.1 + 0.8 \cdot 0.05 \cdot 0.1 = \cdots$$

d) Slh. att minst två komponenter slutar fungera ?

$$P(minst\ 2) = P(alla\ tre) + P(exakt\ 2)$$ ”+” p.g.a. ”eller”

del a) del c)

e) Slh. att alla fungerar hela första året ?

$$P(ingen\ slutar) = P(A^* \cap B^* \cap C^*) = P(A^*) \cdot P(B^*) \cdot P(C^*) = 0.8 \cdot 0.95 \cdot 0.9$$

www.matstat.org

42

Sannolikhet att ”minst en” händelse inträffar

$\boxed{P(\text{"}minst\ en\text{"}) = 1 - P(\text{"}ingen\text{"})}$ följer från komplementsatsen

$P(S) = 1 - P(S^*)$ komplementsats

Låt $S = A \cup B \cup C$ A eller B eller C → ”minst en”

$S^* = (A \cup B \cup C)^* = A^* \cap B^* \cap C^*$ (de Morgan) ej A och ej B och ej C → ”ingen”

$P(A \cup B \cup C) = 1 - P(A^* \cap B^* \cap C^*)$ $P(\text{"}minst\ en\text{"}) = 1 - P(\text{"}ingen\text{"})$

Se exempel ”tre elektroniska komponenter” ovan: slh. att ”minst en” komponent slutar fungera …

$$P(\text{"}ingen\text{"}) = P(A^*) \cdot P(B^*) \cdot P(C^*) = 0.8 \cdot 0.95 \cdot 0.9 = \underline{\underline{0.684}}$$

$$P(\text{"}minst\ en\text{"}) = 1 - 0.684 = \underline{\underline{0.316}}$$

www.matstat.org

43

Sannolikhet att ”minst en” händelse inträffar

Mera allmänt: n oberoende händelser $A_1, A_2, \ldots, A_n$

$$P(A_1) = p_1$$
$$P(A_2) = p_2$$
$$\ldots$$
$$P(A_n) = p_n$$

Händelse **M**: minst en av dem inträffar
Händelse **M***: ingen av dem inträffar

$$P(M) = 1 - P(M^*)$$

$$P(M) = 1 - \underbrace{(1-p_1) \cdot (1-p_2) \cdot \ldots \cdot (1-p_n)}_{\text{slh. att ingen inträffar}}$$

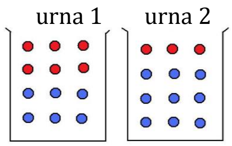
Om alla p_i är lika, dvs. $p_i = p$ för alla i : $\boxed{P(M) = 1 - (1-p)^n}$

www.matstat.org

44

Lagen om total sannolikhet

<u>Exempel</u> : två urnor, se ovan



urna 1: 6 röda av 12
urna 2: 3 röda av 12
steg 1: väljer en urna $\Omega_U = \{1, 2\}$
steg 2: väljer en kula ur denna urna $\Omega_F = \{röd, blå\}$

<u>Problem</u>: $P(röd) = ??$

$$\left.\begin{aligned} P(röd \mid U_1) &= {}^{6}/_{12} \\ P(röd \mid U_2) &= {}^{3}/_{12} \end{aligned}\right\}$$

de betingade sannolikheterna
vikta med slh:na att urna 1 eller 2 har valts

$$P(röd) = P(röd \mid U_1) \cdot P(U_1) + P(röd \mid U_2) \cdot P(U_2)$$ viktat medelvärde

vikter, adderas till 1, t. ex. $P(U_1) = 0.5 \;\; P(U_2) = 0.5$

$$P(röd) = \frac{6}{12} \cdot \frac{1}{2} + \frac{3}{12} \cdot \frac{1}{2} = \underline{\underline{\frac{3}{8}}}$$

45

Lagen om total sannolikhet

Man kan ”härleda” lagen på följande sätt:

$$P(röd) = P[\; (U_1 \cap röd) \;\cup\; (U_2 \cap röd) \;]$$

”och” ”eller” ”och”

$$= P(U_1 \cap röd) + P\,(U_2 \cap röd)$$ pga. axiom 3 (”eller” → ”+”)

använd definition bet. slh.: $P(A \cap B) = P(A) \cdot P(B \mid A)$

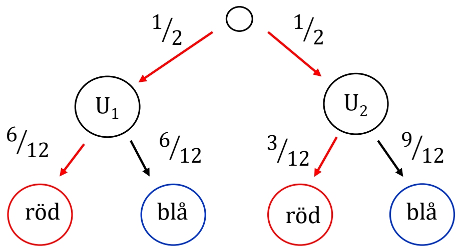
$$= P(röd \mid U_1) \cdot P(U_1) + P(röd \mid U_2) \cdot P(U_2)$$

som ovan

46

Lagen om total sannolikhet

Träddiagram för urn-exemplet :



$$\frac{1}{2}+\frac{1}{2}=1$$

$$\frac{6}{12}+\frac{6}{12}=1$$

$$\frac{3}{12}+\frac{9}{12}=1$$

summan av alla de utgående slh:na från en knop måste vara 1

$$P(röd)=\frac{1}{2}\cdot\frac{6}{12}+\frac{1}{2}\cdot\frac{3}{12}=\underline{\underline{\frac{3}{8}}}$$

två "vägar" som leder till en röd kula

www.matstat.org

47

Lagen om total sannolikhet

... gäller också för flera möjliga händelser:

Exempel : tre urnor 

$$P(röd)=P(röd\mid U_1)\cdot P(U_1)+P(röd\mid U_2)\cdot P(U_2)+P(röd\mid U_3)\cdot P(U_3)$$

"röd" kan realiseras på tre sätt

$P(U_1)+P(U_2)+P(U_3)=1$ vikterna måste adderas till 1

Allmänt, n urnor (eller vad det nu handlar om ...):

$$\boxed{P(R)=\sum_{i=1}^{n}P(R\mid U_i)\cdot P(U_i)}$$

Lagen om total sannolikhet:

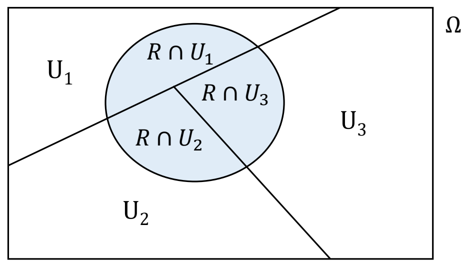
- U_i måste vara oförenliga $U_i\cap U_j=\emptyset$
- U_i måste uttömma utfallsrummet: $\sum P(U_i)=1$

www.matstat.org

48

Lagen om total sannolikhet

- U_i måste vara oförenliga $U_i \cap U_j = \emptyset$
- måste uttömma utfallsrummet: $\sum P(U_i) = 1$



$$P(R) = P(R \cap U_1) + P(R \cap U_2) + P(R \cap U_3)$$

$$= P(R \mid U_1) \cdot P(U_1) + P(R \mid U_2) \cdot P(U_2) + P(R \mid U_3) \cdot P(U_3)$$

genom att använda definitionen för bet. slh.: $P(A \cap B) = P(A) \cdot P(B \mid A)$

Lagen om total sannolikhet

Exempel: flermaskinstillverkning (Blom et al. s. 30)

En fabrik tillverkar någon produkt, på flera olika maskiner.

- maskin 1 ($\mathbf{M_1}$): 25% av produkterna tillverkas, 5% defekt
- maskin 2 ($\mathbf{M_2}$): 35% av produkterna tillverkas , 4% defekt
- maskin 3 ($\mathbf{M_3}$): 40% av produkterna tillverkas, 2% defekt

Alla produkter blandas i en container.
Slumpförsök: väljer slumpmässigt en produkt ur containern.
Sökes: slh. att denna produkt är defekt

$$\left.\begin{array}{l} P(fel \mid M_1) = 0.05 \\ P(fel \mid M_2) = 0.04 \\ P(fel \mid M_3) = 0.02 \end{array}\right\} \text{de betingade sannolikheterna, måste viktas med respektive andelar}$$

$$P(fel) = P(fel \mid M_1) \cdot P(M_1) + P(fel \mid M_2) \cdot P(M_2) + P(fel \mid M_3) \cdot P(M_3)$$

$$\left.\begin{array}{l} P(M_1) = 0.25 \\ P(M_2) = 0.35 \\ P(M_3) = 0.40 \end{array}\right\} \text{"vikter"} \quad \sum_i P(M_i) = 1$$

$$P(fel) = 0.05 \cdot 0.25 + 0.04 \cdot 0.35 + 0.02 \cdot 0.4 = \underline{\underline{0.0345}}$$

Lagen om total sannolikhet

Exempel: diagnostiskt test för en viss sjukdom (Blom et al. s. 31)

- händelse **S** : en slumpmässigt vald person är sjuk
- händelse **S*** : en slumpmässigt vald person är inte sjuk
- $\boldsymbol{P}(+)$: slh. att en slumpmässigt vald person blir positivt testad
- $\boldsymbol{P}(-)$: slh. att en slumpmässigt vald person blir negativt testad
- $\boldsymbol{P}(\boldsymbol{S}) = \boldsymbol{p}$: andelen sjuka personer i befolkningen

Sensitiviteten och specificiteten för testet är bekant:

- $P(+ \mid S) = 0.9999$; sensitivitet = slh. att testet blir positivt om personen är sjuk
- $P(- \mid S^*) = 0.995$; specificitet = slh. att testet blir negativt om personen inte är sjuk

Båda dessa värden borde vara nära ett om testet är bra !

Sannonlikhet för ett positivt test:

$$P(+) = P(+\mid S) \cdot P(S) + P(+ \mid S^*) \cdot P(S^*) \quad \text{total slh.}$$

$$= P(+\mid S) \cdot p + [1 - P(- \mid S^*)] \cdot (1 - p) \quad \text{enligt komplementsatsen}$$

a) Hur stor är slh. att en slumpmässigt vald person testas positivt om $\boldsymbol{p} = \mathbf{0.2}$?

$P(+) = 0.9999 \cdot 0.2 + 0.005 \cdot 0.8 = \mathbf{0.204}$ (lite högre än den verkliga andelen sjuka)

b) Hur stor är slh. att en slumpmässigt vald person testas positivt om $\boldsymbol{p} = \mathbf{0.001}$?

$P(+) = 0.9999 \cdot 0.001 + 0.005 \cdot 0.999 = \mathbf{0.006}$

- 6 gånger högre än den verkliga andelen sjuka !
- många ”falskt positiva” om sjukdomen är sällsynt!

www.matstat.org

51

Bayes sats

En hel gren inom statistik: ”Bayesian statistics”

Thomas Bayes

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

$$P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

$$\boxed{P(B \mid A) = \frac{P(A \mid B) \cdot P(B)}{P(A)}} \quad \text{Bayes sats}$$

”vänder om” betingade sannolikheter : vet $P(A \mid B)$ → söker $P(B \mid A)$

www.matstat.org

52

Bayes sats

Exempel : flermaskinstillverkning (se ovan)

Problem: En kund påträffar en felaktig komponent. Vad är slh. att den kom från maskin 1 ?

känd känd

$$P(M_1 \mid fel) = \frac{P(fel \mid M_1) \cdot P(M_1)}{P(fel)} \qquad \text{Bayes sats}$$

total slh.

$$P(M_1 \mid fel) = \frac{0.05 \cdot 0.25}{0.0345} = 0.36 = \underline{\underline{36\%}}$$

www.matstat.org

53

Bayes sats

Exempel : diagnostiskt test (se ovan)

En person testas positivt (denna händelse är alltså givet). Hur stor är sannolikheten att personen verkligen är sjuk om man antar att 0.1% av befolkningen bär på sjukdomen (dvs. $\boldsymbol{p = 0.001}$)

Vi söker alltså $\boldsymbol{P(S \mid +)}$ - slh. att sjukdomen föreligger givet att testet var positivt !

känd känd

$$P(S \mid +) = \frac{P(+ \mid S) \cdot P(S)}{P(+)} \qquad \text{Bayes sats}$$

total slh.

$$P(S \mid +) = \frac{0.9999 \cdot 0.001}{0.006} = 0.167 \approx \underline{\underline{17\%}}$$

OBS! Bara 17% av de positivt testade är verkligen sjuka ! (sällsynt sjukdom)

www.matstat.org

54

Bayes sats

Exempel : alkohol i trafiken (OBS!: siffrorna i detta exempel är påhittade, men kanske inte helt orealistiska)

”20% av alla trafikolyckor sker i samband med alkohol”
Är denna information meningsfull ? – Knappast ! (80% av olyckorna sker utan alkohol ...)

$P(alk \mid ol.) = 20\%$ bara olyckor ($ol.$) räknas in – fel!

Bättre att jämföra: $P(ol. \mid alk) \rightleftarrows P(ol. \mid alk^*)$

gissning!

$$P(ol. \mid alk) = \frac{P(alk \mid ol.) \cdot P(ol.)}{P(alk)} = \frac{0.2 \cdot 10^{-4}}{10^{-3}} = 0.2 \cdot 0.1 = \underline{\underline{2\%}}$$

$$P(ol. \mid alk^*) = \frac{P(alk^* \mid ol.) \cdot P(ol.)}{P(alk^*)} = \frac{0.8 \cdot 10^{-4}}{1 - 10^{-3}} \approx 0.8 \cdot 10^{-4} = \underline{\underline{0.008\%}}$$

alk* betyder alltså ”ingen alkohol” (komplement)

55

Bayes sats

Exempel: lungcancer och rökning

”90% av alla lungcancerdrabbade är rökare”

Bättre att jämföra: $P(cancer \mid rökare) \rightleftarrows P(cancer \mid rökare^*)$

$rökare^* =$ icke-rökare

beräknas som exemplet ”alkohol i trafiken”

56

Sammanfattning

Klassiska sannolikhetsdefinition: $P(A) = \frac{g}{m}$

Kolmogorows axiomer

Komplementsats $P(A^*) = 1 - P(A)$

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

Betingade sannolikhet $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

Oberoende händelser $P(A \cap B) = P(A) \cdot P(B)$ eller $P(B \mid A) = P(B)$

Lagen om total slh. $P(R) = \sum_i P(R \mid M_i) \cdot P(M_i)$ $\sum_i M_i = 1$ M_i disjunkta

Bayes sats $P(B \mid A) = \frac{P(A \mid B) \cdot P(B)}{P(A)}$ ”vänder om” betingade slh.

57