SF1911: Statistik för bioteknik
Föreläsning 4.

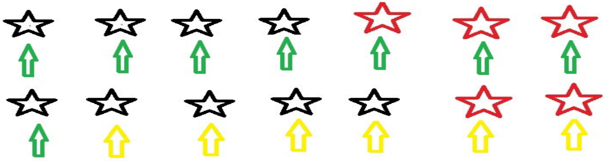
TK

7.11.2017

Lärandemål

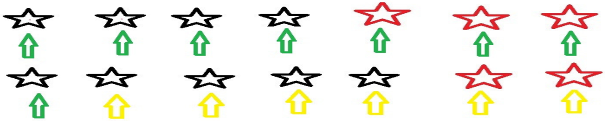
- Betingad sannolikhet (definition, betydelse)
- Oberoende händelser
- Lagen om total sannolikhet
- Bayes' sats (definition, användning, betydelse)
  - a priori sannolikhet
  - posteriori sannolikhet
- Bayes' sats i kliniska tillämpningar

In the figure we have 14 (abstract) peas: nine have white flowers, five have red flowers, six have yellow pods, eight have green pods.



Two peas are selected at **random** and **without replacement**. **Q**: What is the probability for the event that the first selection has a green pod and the second selection has a yellow pod ?

Conditionally Probable Events: a Mendelian introduction



- First selection: $P(\text{green pod}) = \frac{8}{14}$
- Second selection: $P(\text{yellow pod}) = \frac{6}{13}$
- $P($ 1st pea with green pod and 2nd pea yellow pod$) = \frac{8}{14} \cdot \frac{6}{13} \approx$ 0.264.

KTH Matematik

Conditionally Probable Events

Two peas are selected at **random** and **without replacement**.

$$P(\text{ 1st pea green pod and 2nd pea with yellow pod}) = \frac{8}{14} \cdot \frac{6}{13} \approx 0.264.$$

The key point here is that **we must adjust the probability of the second event to reflect the outcome of the first event**. One pea was removed in the first selection, there are only 13 left to choose for the second selection.

(Notation for Conditional Probability)

$P(B \mid A)$ represents the probability of the event B occurring after it is assumed that the event A has already occurred.

We read $B \mid A$ as " B given A " or as " event B after that event A has already occurred " .

Multiplication Rule

$$P\left(A \cap B\right) = P(A) \cdot P(B \mid A)$$

But also

$$P\left(A \cap B\right) = P(B) \cdot P(A \mid B)$$

Now we have in fact already used the multiplication rule

$$P(\text{ 1st pea green pod and 2nd pea yellow pod}) =$$

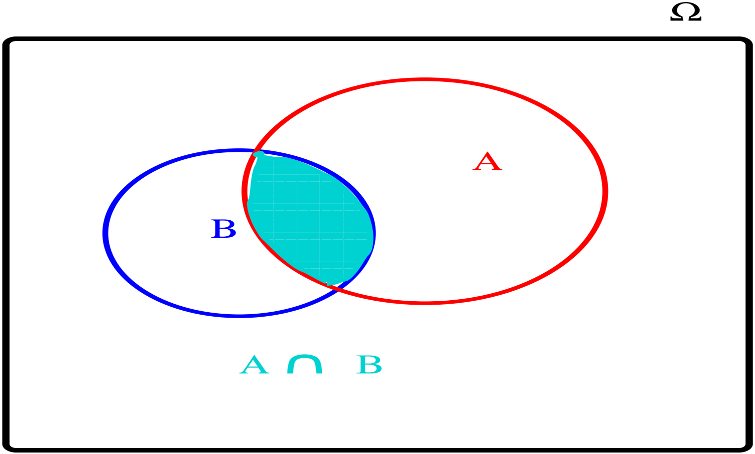
$$P(\text{ 1st pea green pod}) \cdot Pr(\text{2nd pea yellow pod} \mid \text{1st pea green pod }) =$$

$$\frac{8}{14} \cdot \frac{6}{13}$$

Generellt: Betingad sannolikhet

Inledning (betingad sannolikhet)

Låt A och B vara två händelser, dvs $A, B \subset \Omega$. Vad är sannolikheten, betecknad med $P(B \mid A)$, för B då vi vet att A har inträffat?



$P(B \mid A)$

Vi är ledda oss till följande definition.

Definition

Låt A och B vara två händelser. Antag att $P(A) > 0$. Sannolikheten för B betingat av A betecknas med $P(B \mid A)$ och definieras som

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

KTH Matematik

Multiplication Rule

If $P(A) > 0$

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

and if $P(B) > 0$

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

Probability: independent events

$$P\,(\;A \cap B\;) = P(A) \cdot P(B \mid A)$$

But if A and B are **independent** events

$$P\,(A \cap B) = P(A) \cdot P(B).$$

Hence, if A and B are independent events

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

i.e., we need not adjust the probability of B to reflect the outcome of A

Oberoende händelser

Intuitivt är två händelser A och B oberoende om inträffandet av A inte ger någon information om huruvida B inträffar eller ej. I formler betyder detta

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

Allmänt gäller ju

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \quad \text{om } P(A) > 0.$$

Oberoende händelser

Multiplikation med $P(A)$ leder oss till följande definition:

Definition

Två händelser A och B är oberoende om

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

Definitionen ovan kräver inget villkor om positiva sannolikheter.

Komplement till oberoende händelser är oberoende

Vi har

$$P(A^c \cap B^c) = P((A \cup B)^c) = 1 - P(A \cup B)$$

$$= 1 - P(A) - P(B) + P(A)P(B) = (1 - P(A)) - P(B)(1 - P(A))$$

$$= (1 - P(A))(1 - P(B)) = P(A^c)P(B^c).$$

KTH Matematik

Betingad sannolikhet: exempel

Exempel

(Kast med röd och vit tärning)
$A =$ *summan av ögonen är högst 4.*
$B_k =$ *vita tärningen visar* k *ögon.*
$P(B_k \mid A) = 0$ *om* $k \geq 4$.
Möjliga utfall, m, är 36: $(v, r), v, r = 1, \ldots 6$, *dvs* $(1, 1), (1, 2), \ldots (6, 6)$.
Gynnsamma utfall för A, är 6*: (1,1), (1,2), (1,3), (2,1), (2,2), (3,1).*
Gynnsamma utfall för $A \cap B_k$, *är* $4 - k$*:* $(v, r), v = k, r = 1, \ldots 4 - k$,
dvs $(k, 1), (k, 2), \ldots (k, 4 - k)$ *om* $k < 4$.

Betingad sannolikhet: exempel forts.

Exempel

Klassiska sannolikhetsdefinitionen ger

$$P(A) = \frac{6}{36} \quad och \quad P(A \cap B_k) = \frac{4-k}{36}.$$

Detta ger, för $k < 4$,

$$P(B_k \mid A) = \frac{4-k}{6} = \begin{cases} \frac{3}{6} = \frac{1}{2} & k = 1 \\ \frac{2}{6} = \frac{1}{3} & k = 2 \\ \frac{1}{6} & k = 3. \end{cases}$$

KTH Matematik

Betingad sannolikhet

Ofta är det lättare att ange värden till betingade sannolikheter än till obetingade, och vi utnyttjar definitionen "baklänges".

Exempel

En ohederlig person har två tärningar, en äkta och en falsk som alltid ger 6 ögon. Han väljer slumpmässigt den ena. Vad är sannolikheten för 5 resp. 6 ögon. Låt oss betrakta fallet med sex ögon. Intuitivt bör gälla att sannolikheten är

$$\frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot 1 = \frac{1}{12} + \frac{6}{12} = \frac{7}{12}.$$

KTH Matematik

Lagen om total sannolikhet

Mera systematiskt gäller följande sats

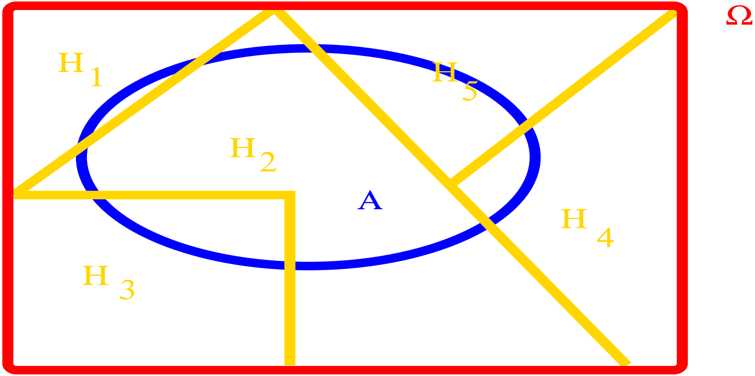
Sats

(Lagen om total sannolikhet)
Om $H_1, \ldots, H_n$ är disjunkta händelser, har positiv sannolikhet och uppfyller hela Ω, så gäller för varje händelse $A \subset \Omega$ att

$$P(A) = \sum_{i=1}^{n} P(H_i) P(A \mid H_i).$$

KTH Matematik

Lagen om total sannolikhet



Bevis. Vi utnyttjar en av De Morgans regler

$$P(A) = P(A \cap \Omega) = P(A \cap (H_1 \cup \ldots \cup H_n)) = P((A \cap H_1) \cup \ldots \cup (A \cap H_n))$$

och sedan regel (c) från första föreläsningen, ty $(A \cap H_i) \cap (A \cap H_j) = \emptyset$ om $i \neq j$,

$$= \sum_{i=1}^{n} P(A \cap H_i) = \sum_{i=1}^{n} P(H_i) P(A \mid H_i).$$

$\square$

Bayes' sats

Vi ska nu ge en viktig sats om "vändning" av händelserna i betingade sannolikheter.

Sats

(Bayes' sats) *Under samma villkor som i lagen om total sannolikhet gäller*

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{\sum_{j=1}^{n} P(H_j)P(A \mid H_j)}.$$

Bevis.

$$P(H_i \mid A) = \frac{P(H_i \cap A)}{P(A)} = \frac{P(H_i \cap A)}{P(H_i)} \cdot \frac{P(H_i)}{P(A)} = P(A \mid H_i) \cdot \frac{P(H_i)}{P(A)}.$$

Lagen om total sannolikhet tillämpad på $P(A)$ ger resultatet.

Bayes' sats: tolkning

Sannolikheten $P(H_i)$ kallas **a priori sannolikhet** för H_i. Sannolikheten $P(H_i \mid A)$ kallas **posteriori sannolikhet** för H_i. Bayes' sats visar alltså hur a priori sannolikhet omvandlas till posteriori sannolikhet.

Thomas Bayes ca. 1701 – 1761

Bayes' sats: exempel

Låt oss gå tillbaka till exemplet om falskspelaren. Sätt
$A = 6$ ögon.
$H_1 =$ äkta tärningen.
$H_2 =$ falska tärningen.
Då gäller

$$P(A) = P(H_1)P(A \mid H_1) + P(H_2)P(A \mid H_2) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot 1 = \frac{7}{12},$$

som i exemplet. Bayes' sats ger vidare

$$P(H_1 \mid A) = \frac{P(H_1 \cap A)}{P(A)} = P(A \mid H_1) \cdot \frac{P(H_1)}{P(A)} = \frac{1}{6}\frac{1}{2}\frac{12}{7} = \frac{1}{7}$$

och

$$P(H_2 \mid A) = \frac{P(H_2 \cap A)}{P(A)} = P(A \mid H_2) \cdot \frac{P(H_2)}{P(A)} = 1 \cdot \frac{1}{2}\frac{12}{7} = \frac{6}{7}$$

vilket kanske inte är lika lätt att inse rent intuitivt.

Bayes' sats & inlärning

Vi har N slantar i en urna. $N - 1$ av dessa är hederliga i den meningen att de har två olika sidor, krona och klave. En av slantarna är falsk i den meningen att den har krona på båda sidor.
Någon plockar på måfå en av slantarna utan att vi får inspektera densamma, och singlar den k ggr. Utfallet blir k kronor. Vad är sannolikheten att den valda slanten är den falska?

Beteckningar: H_1 = hederlig slant, H_2 = falsk slant. A_k = händelsen med k st. kronor i k singlingar.
Sannolikheter enligt uppgiften: $P(H_1) = \frac{N-1}{N}$ (antalet gynnsamma fall/totalantalet fall). $P(H_2) = \frac{1}{N}$. $P(A_k \mid H_1) = \frac{1}{2^k}$ (antalet gynnsamma fall/totalantalet fall, senare alt. p.g.a oberoende) och $P(A_k \mid H_2) = 1$. Sökt: $P(H_2 \mid A_k)$.

Bayes' sats & inlärning

H_1 = hederlig slant, H_2 = falsk slant. A_k = händelsen med k st. kronor i k singlingar.
$P(H_1) = \frac{N-1}{N}$, $P(H_2) = \frac{1}{N}$. $P(A_k \mid H_1) = \frac{1}{2^k}$ och $P(A_k \mid H_2) = 1$. Sökt: $P(H_2 \mid A_k)$.

Lagen om total sannolikhet (LTS) ger

$$P(A_k) = P(A_k \mid H_1)\, P(H_1) + P(A_k \mid H_2)\, P(H_2)$$

$$= \frac{1}{2^k}\frac{N-1}{N} + \frac{1}{N} = \frac{2^k + N - 1}{2^k N}$$

Bayes' sats ger

$$P(H_2 \mid A_k) = \frac{P(A_k \mid H_2)\, P(H_2)}{P(A_k)}$$

$$= \frac{2^k}{2^k + N - 1}.$$

KTH Matematik

Bayes' sats & inlärning

H_1 = hederlig slant, H_2 = falsk slant. A_k = händelsen med k st. kronor i k singlingar.
$P(H_1) = \frac{N-1}{N}$, $P(H_2) = \frac{1}{N}$, $P\left(A_k \mid H_1\right) = \frac{1}{2^k}$ och $P\left(A_k \mid H_2\right) = 1$.

$$P\left(H_2 \mid A_k\right) = \frac{2^k}{2^k + N - 1}.$$

Observera att

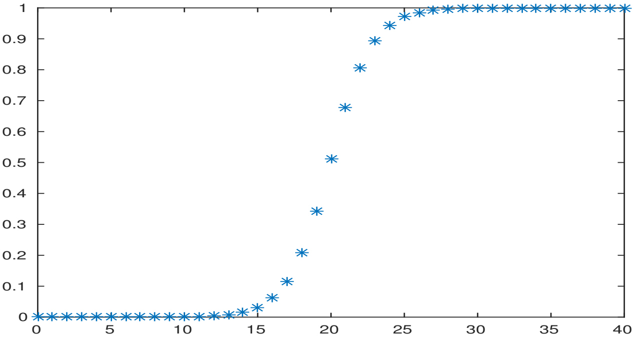
$$P\left(H_2 \mid A_0\right) = \frac{1}{N} = P(H_2).$$

KTH Matematik

Bayes' sats & en inlärningskurva

I figuren plottas med $*$ den sökta sannolikheten som funktion av k för $k = 0, \ldots, 40$ och $N = 1000000$

$$P(H_2 \mid A_k) = \frac{2^k}{2^k + N - 1}.$$



Bayes' sats i diagnostiken

General terminology: $T+$ = positive test result in a diagnostic test, $T-$ = negative test result in a diagnostic test. $D+$ = has a disease, $D-$ = does not have a disease.

- **Sensitivity** $= P(T+ \mid D+)$
- **Specificity** $= P(T- \mid D-)$
- **True positive (predicted value)** $= P(D+ \mid T+)$
- **True negative (predicted value)** $= P(D- \mid T-)$

- **Overall accuracy** $= P(D+) \cdot \text{sensitivity} + P(D-) \cdot \text{specificity}$
- **False Positive** $P(\text{False positive}) = P(T+ \mid D-) = 1 - \text{specificity}$
- **False Negative** $P(\text{False negative}) = P(T- \mid D+) = 1 - \text{sensitivity}$

Bayes' rule tells amongst other things that

$$P(D+|T+) = \frac{P(T+ \mid D+)P(D+)}{P(T+)}$$

and

$$P(D-|T-) = \frac{P(T- \mid D-)P(D-)}{P(T-)}.$$

KTH VETENSKAP OCH KONST

KTH Matematik

Bayes' sats: screening för cancer

Vi räknar på det exempel som ges av John Allen Paulos: The Math behind Screening Tests. What a positive result really means. *Scientific American* , January 2012. http://www.scientificamerican.com/article.cfm?id=weighing-the-positives

Låt oss anta att 0.4 % av befolkningen drabbas av en form av cancer. Ett diagnostiskt test ger positivt resultat i 99.5 % av fallen, om en individ har cancer och ger positivt resultat i 1 % av fallen, om en individ inte har cancer. Om vi plockar på måfå en vuxen individ ur en stor population och testet ger ett positivt svar, vad är sannolikheten för att individen har den aktuella formen av cancer ? Beteckna $D+ =$ individen har cancer. $D- =$ individen har inte cancer. $T+$ positivt testresultat. Vi har

$P(D+) = 0.004$, $P(D-) = 0.996$, $P(T+|D+) = 0.995$, $P(T+|D-) = 0.01$

Sökt är $P(D+|T+)$.

KTH Matematik

$$P(D+) = 0.004, P(D-) = 0.996,$$

$$P(T+|D-) = 0.995, P(T+|D-) = 0.01$$

Bayes' sats ger $P(D+|T+)$ som

$$P(D+|T+) = \frac{P(T+|D+)P(D+)}{P(T+|D+)P(D+) + P(T+|D-)P(D-)} =$$

$$= \frac{0.995 \cdot 0.004}{0.995 \cdot 0.004 + 0.01 \cdot 0.996} = 0.2855.$$

En oväntat låg posteriori sannolikhet ! Varför är 72% falska positiva ändå ett rimligt svar ? Vilken slutsats drar vi om screening ?

Positive likelihood ratio (LR_+) is defined as

$$\mathrm{LR}_+ = \frac{P(T+ \mid D+)}{P(T+ \mid D-)}$$

The posterior odds of $D+$ given $T+$ is

$$\frac{P(D+ | T+)}{P(D- | T+)} = \mathrm{LR}_+ \times \frac{P(D+)}{P(D-)}.$$

Here $\frac{P(D+)}{P(D-)}$ is the prior odds of $D+$.

KTH Matematik

$$\mathrm{LR}_+ = \text{Sensitivity / (1 - Specificity)}$$

(Negative) likelihood ratio (LR_-) is defined as

$$\mathrm{LR}_- = \frac{P(T- \mid D+)}{P(T- \mid D-)} = \text{ (1 - Sensitivity) / Specificity}$$

$$\frac{P(D+ \mid T-)}{P(D- \mid T-)} = \mathrm{LR}_- \times \frac{P(D+)}{P(D-)}.$$

KTH Matematik

En ny version av Bayes' sats

$$P(\text{ True positive predicted value }) =$$

$$\frac{P(\text{ D+ }) \cdot \text{ sensitivity}}{P(\text{ D+ }) \cdot \text{sensitivity } + P(\text{ D- })(1 - \text{specificity})}.$$

$$P(\text{ True negative predicted value }) =$$

$$\frac{P(\text{ D- }) \cdot \text{ specificity}}{P(\text{ D- }) \cdot \text{specificity } + P(\text{ D+ })(1 - \text{sensitivity })}.$$

En ny version av Bayes' sats: härledning

$$P(\text{ True positive predicted value }) = P(D+ \mid T+) = \frac{P(T+ \mid D+)P(D+)}{P(T+)}$$

$$= \text{ Sensitivity} \cdot \frac{P(D+)}{P(T+)}.$$

The law of total probability (lagen om total sannolikhet) yields

$$P(T+) = P(T+ \mid D+)P(D+) + P(T+ \mid D-)P(D-)$$

and

$$= \text{Sensitivity} \cdot P(D+) + (1 - \text{Specificity}) \cdot P(D-).$$

KTH Matematik

En ny version av Bayes' sats: härledning

This gives

$$P(\text{ True positive predicted value }) =$$

$$\frac{P(\text{ D+ }) \cdot \text{ sensitivity}}{P(\text{ D+ }) \cdot \text{sensitivity } + P(\text{ D- })(1 - \text{specificity})}.$$

KTH Matematik

En ny version av Bayes' sats: härledning

The second case

$$P(\text{ True negative predicted value }) = \frac{P(T- \mid D-)P(D-)}{P(T-)} = \frac{\text{specificity} \cdot P(D-)}{P(T-)}.$$

Law of total probability again,

$$P(T-) = P(T- \mid D+)P(D+) + P(T- \mid D-)P(D-) = (1 - \text{sensitivity } \cdot P(D+) + \text{specificity } \cdot P(\text{ D- })$$

and we have the desired formula.

Vem skall få huvudrollen i en kommande Hollywoodfilm om Tom Bayes?