Bayes law. Sensitivity, specificity.

Jon Michael Gran
Department of Biostatistics, UiO

MF9130 Introductory course in statistics
Tuesday 24.05.2011

Overview

Bayes law. Sensitivity and specificity

(Aalen chapter 3.9-3.10, Kirkwood and Sterne chapter 14.4 and 36.2)

- Conditional probability
- Law of total probability
- **Bayes law**
- Uncertainty in diagnostic testing: **Sensitivity, specificity**, positive and negative predictive value
- Bayesian statistics

Conditional probability

Repetition of conditional probability...

- The **conditional probability** of B given that A has occurred is written
$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$
- Example:

  Probability of a child dying in crib death: ca $1/2000$
  If a sibling has previously died in crib death: about $6/2000$
  The latter is a conditional probability

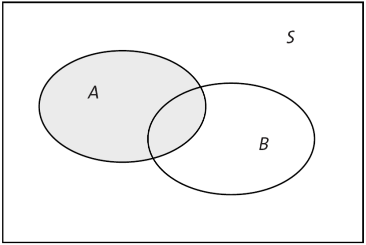
Law of total probability

Example: Gender of twins

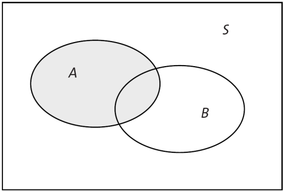
- We want to find **the probability of two twins having the same gender**
- Monozygotic twins have the same gender, while dizygotic twins are like any other siblings
- We have take into consideration if the twins are monozygotic or not $\rightarrow$ use **the law of total probability**

Some notation...

- S is the set of all events
- The **union** of A and B:
  $A \cup B =$ all events that are in A **or** B
- The **intersection** of A and B:
  $A \cap B =$ all events that are both in A **and** B
- The **complement** of A: $\bar{A} =$ all events **not in** A



Figur: Venn diagram, where $A = (A \cap B) \cup (A \cap \bar{B})$



Deriving the law of total probability

- In the previous slide we saw that **any event A can be divided in two** with regard to another event B: $A = (A \cap B) \cup (A \cap \bar{B})$
- Because **the two events are disjunct** we can write: $P(A) = P(A \cap B) + P(A \cap \bar{B})$
- Using the multiplicative rule, we get **the law of total probability**:

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

Example: Gender of twins (cont.)

- A = Both twins have the same gender
  B = The twins are monozygotic
- **Want to find $P(A)$**
- The probability of twins being monozygotic, $P(B)$, is $1/3$
- The law of total probability gives us:

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|\bar{B})P(\bar{B}) \\ &= 1 \cdot 1/3 + 1/2 \cdot 2/3 \\ &= 2/3 \end{aligned}$$

- The probability that two twins have the same gender is 0.67

Bayes law

Example: Gender of twins (cont.)

- What if we now want to find **the probability that two twins of the same gender are monozygotic**?
- In others words, what is $P(B|A)$?
- **Bayes law** - Thomas Bayes (1702-1761)

Theory: Bayes law

- Remember the definition of **conditional probability**,

  $P(B|A) = \frac{P(A \cap B)}{P(A)}$,
- **the multiplicative rule**,

  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$,
- and **the law of total probability,**
  $P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$

Theory: Bayes law

- Remember the definition of **conditional probability**, $P(B|A) = \frac{P(A \cap B)}{P(A)}$,
- **the multiplicative rule**,

  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$,
- and **the law of total probability,** $P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$
- Combining the three, putting the last two into the first one, gives us **Bayes law**:

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

Example: Gender of twins (cont.)

- Remember:
  A = Both twins have the same gender
  B = The twins are monozygotic
- What is **the probability that two twins of the same gender are monozygotic**? In others words, what is $P(B|A)$?
- Bayes law gives us:

$$
\begin{aligned}
P(B|A) &= \frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|\bar{B})P(\bar{B})} \\
&= \frac{1 \cdot 1/3}{1 \cdot 1/3 + 1/2 \cdot 2/3} \\
&= 1/2
\end{aligned}
$$

Uncertainty in diagnostic testing: Sensitivity, specificity, positive and negative predictive value

An application of Bayes law (among many)

- Bayes law plays a central part when doing calculations on **uncertainty in diagnostic testing**

Examples of uncertainty in diagnostic testing:

- Mammography: How certain is the breast cancer diagnosis?
- HIV-testing: How certain is a positive test?
- Lie detector: How certain is it that the person is lying?

Concepts of diagnostic testing

- **Sensitivity**: the probability of revealing that a person is "ill"
  - Mammography: 98%
  - HIV-testing: 70-90%
  - Lie detector: 76%
- **Specificity**: the probability of revealing that a person is "well"
  - Mammography: 99.8%
  - HIV-testing: 90-95%
  - Lie detector: 63%
- **Positive predictive value**: The probability that the person has the disease given a positive test
- **Negative predictive value**: The probability that the person does not have the disease given a negative test

Example: Validity of mammography

- From the Norwegian Medical Journal, 1990: 372 women with a lump in the breast has been referred to surgical clinic

		Mammography	
		Benign	Malign
Final diagnosis	Benign	331	16
	Malign	3	22

Sensitivity: $22/(3+22) = 88\%$
Specificity: $331/(331+16) = 95\%$
Positive predictive value: $22/(16+22) = 58\%$
Negative predictive value: $331/(331+3) = 99.1\%$

The concepts of diagnostic testing in the form of conditional probabilities

- Sensitivity: $P(pos.|ill)$
- Specificity: $P(neg.|well)$
- Positive predictive value: $P(ill|pos.)$
- Negative predictive value: $P(well|neg.)$

Bayes law is used to compute sensitivity, specificity and prevalence to positive or negative predictive value

Example: HIV testing

- Testing for antibodies of the HIV virus
  - Positive result: test shows antibodies
  - Negative result: test does not show antibodies
- False positives:
  - Test error
  - Antibodies from related virus
  - Probability of error: 0.2%
- False negatives:
  - Test error
  - Antibodies not yet produced in sufficient quantity
  - Probability of error: 2%

Example: Computation for one single ELISA test

- Sensitivity: 98%, Specificity: 99.8%
- Prevalence: assumed to be 0.1%
- 100 000 persons:
  - No. of HIV infected: $100000 \cdot 0.001 = 100$
  - No. of true positives: $100 \cdot 0.98 = 98$
  - No. of false negatives: 2
  - No. of false positives: $99900 \cdot (1 - 0.998) = 200$
  - Positive predictive value: $\frac{98}{98+200} = 33\%$
    Just a third of the positives are infected

Special case of Bayes law: Computation of positive predictive value

- **Positive predictive value** (PPV) is calculated from Bayes law by the following formula

$$PPV = \frac{sens \cdot prev}{sens \cdot prev + (1 - spes) \cdot (1 - prev)}$$

- You have to know the value of the sensitivity (sens), specificity (spes) and prevalence (prev)

The corresponding formula for negative predictive value

- $NPV = \frac{spes \cdot (1 - prev)}{(1 - sens) \cdot prev + spes \cdot (1 - prev)}$

Example: Calculation for the HIV test

For the HIV test the sensitivity is 98% and the specificity 99.8%. With 0.1% prevalence we get:

- $PPV = \frac{0.98 \cdot 0.001}{0.98 \cdot 0.001 + (1-0.998) \cdot (1-0.001)} = 0.33$
- $NPV = \frac{0.998 \cdot (1-0.001)}{(1-0.98) \cdot 0.001 + 0.998 \cdot (1-0.001)} = 0.99998$

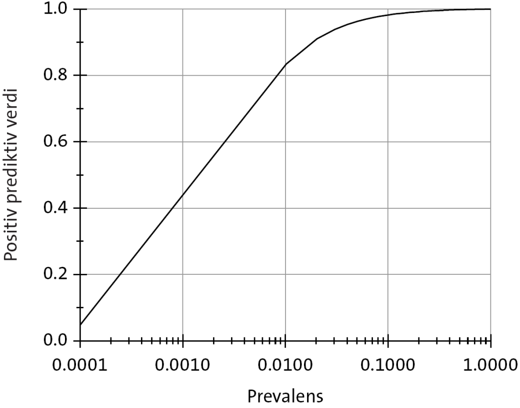
The positive predictive value, $P(ill|pos.)$, depending on the prevalence for the HIV test

- The probability that a positive result is true:

Prevalence	PPV
1/10000	5%
1/1000	33%
1/100	83%
1/10	98%

- Norway: Prevalence 1/1000, PPV $= 0.33$
- Injecting drug users: Prevalence 1/50, PPV $= 0.91$
- Cities in central Africa: Prevalence 1/4, PPV $= 0.99$

HIV test: How does the prevalence influence the positive predictive value



Figur 3.4 Positiv prediktiv verdi av en HIV-test som funksjon av smitte-prevalensen. (Merk at skalaen for prevalensen er logaritmisk.)

The importance of the prevalence

- The **risk of false positives depends strongly on the prevalence** – it is greatest for rare diseases and smaller for more common diseases
- False positives are **a big problem in mass screening** for disease. It could be that the majority of the positives are false positives

Kappa

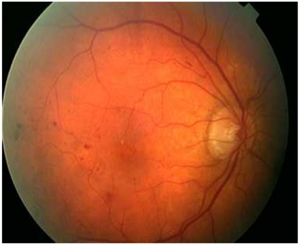
- Measure of agreement between evaluations

Kappa

- Measure of agreement between evaluations

Example: Reliability of clinical investigation

		Second clinician Little/no	Second clinician Moder./severe
First clinician	Little/no:	46	10
	Moder./severe:	12	32



- Taken from Sackett et al: Clinical Epidemiology (Little, Brown and Company, 1985). Photographs of the retina in 100 patients evaluated by two clinicians with respect to occurrence of retinopathy
- Observed agreement: $(46 + 32) / 100 = 78\%$

Finding Kappa using the expected table

- Table with marginals. Example: $(56 \times 58)/100 = 32.5$

Observed:

46	10	56
12	32	44
58	42	100

Expected:

32.5	23.5	56
25.5	18.5	44
58	42	100

- Expected agreement: $(32.5 + 18.5)/100 = 51\%$
- Measure corrected for random agreement:

$$\text{Kappa} = \frac{78 - 51}{100 - 51} = 0.55$$

Interpretation of kappa

Kappa	Strength
<0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Very good

Bayesians statistics

Bayesians statistics: Subjective probabilities

- Although a consept used in everyday live, **probability** is difficult to define exactly
- **Frequentist** definition: The probability of an event is the proportion of times that the event would occur in a large number of similar trials
  - Estimation is completly data driven
- **Bayesian** definition: The size of the probability represent ones degree of belief in the occurence of an event
  - Driven by your data AND your prior belief

Where does Bayes come in?

- Bayes law is used to calculate such probabilities based on our prior belief:
$$P(\theta|\text{data}) = \frac{P(\text{data}|\theta)P(\theta)}{P(\text{Data})}$$
- θ refers to the parameters in your model (mean, variance, etc.)
- The **prior** distribution $P(\theta)$ is where you put in your prior beliefs
- What you want to estimate is the **posterior** distribution $P(\theta|\text{data})$, the probability distribution of the model parameters given your data
- The more data you have, the more will it dominate over your prior belief

Bayesian statistics and applications

- When you have **prior knowledge** about your problem, you get to actually use this information
  - Was (at least) controversial
- But also, when you know little (or nothing) there is many **methodological advantages** with Bayesian statistics
  - Non informative priors: Pretend to know something (very little): A prior around 0 with great uncertainty. Usually the results are the same as when using frequentistic methods
  - Easier to use than frequentistic methods in many advanced problems
- Bayesian statistics is not really relevant for simpler problems like in this course, but you will (probably) **at some point come across papers using Bayesian approaches**

Summary

Key words

- Conditional probability
- Law of total probability
- Bayes law
- Sensitivity, specificity
- Positive predictive value, negative predictive value
- Kappa

Notation

- $P(A)$ and $P(\bar{A})$
- $P(A|B)$
- $P(A \cup B)$ and $P(A \cap B)$