Bus 701: Advanced Statistics

Harald Schmidbauer

İSTANBUL BİLGİ ÜNİVERSİTESİ

About These Slides

- The present slides are not self-contained; they need to be explained and discussed. They contain only a small part of the course Bus 701.

- Even though being a “work in progress” and subject to revision, the slides constitute copyrighted material.
  If you want to reproduce or copy anything from the slides, please ask:

Harald Schmidbauer	**harald** at **hs-stat** dot **com**
Angi Rösch	**angi.r** at **t-online** dot **de**

- The slides were produced using LaTeX and R (the R project; www.R-project.org) on a Linux system.

- R files used for this course are available upon request.

Chapter 6:

Probability and Its Philosophical Background

6.1 Ideas of Probability

Example 1.

Consider a die, and the statement:

> "The probability that this die falls '6' is 25%."

- This statement is either true or false.
- This statement is about how the world is.
- We imagine someone could explain this, invoking the geometry of the die.
- We could do experiments to see if the statement is true.

6.1 Ideas of Probability

Example 1.

Here, probability is related to:

- frequency
- tendency
- propensity
- disposition
- symmetry

Probability in this sense is about how the world is.

We speak of a frequency-type probability.

6.1 Ideas of Probability

Example 2. Consider the statement:

> "Taking all the evidence into account, the probability of a recession of global economy in 2009 is 25%."

If this statement is true. . .

- . . . it is not because of how the world is.
- . . . it is because of how well the evidence supports the hypothesis:

  > "There will be a recession of global economy in 2009."

- No experiment is possible to test the statement.

6.1 Ideas of Probability

Example 2.

Here, probability is related to:

- belief
- evidence
- confidence
- credibility

Probability in this sense is about what we know.

We speak of a belief-type probability.

6.1 Ideas of Probability

Further considerations.

- We are often switching back and forth between frequency-type and belief-type probabilities.
- We can take the frequency-type probability as a belief-type probability in the single case, when we are ignorant of anything else.
  ("frequency principle")
- Many applied situations can be seen from both perspectives.

6.1 Ideas of Probability

Example: A weather forecast.

A meteorologist says:

> "The probability of rain tomorrow is 30%."

- A frequency dogmatist's point of view: . . .
- A belief dogmatist's point of view: . . .

6.2 Theories about Probability

Belief-type theories: Personal probability.

This approach assumes that every person has their own probability.

- Frank P. Ramsey (1903–1930)
- Bruno de Finetti (1906–1985)
- Leonard J. Savage (1917–1971)

6.2 Theories about Probability

Belief-type theories: Logical probability.

This approach assumes that probability is relative to available evidence; i.e.: Any person would find it reasonable to believe this.

- John Meynard Keynes (1883–1946)
- Rudolf Carnap (1891–1970)

6.2 Theories about Probability

Frequency-type theories: Probability as limiting frequency.

This approach emphasizes what can be seen.

- John Venn (1834–1923)
- Richard von Mises (1883–1953)

6.2 Theories about Probability

Frequency-type theories: Probability as propensity.

This approach emphasizes the underlying structures — physical or geometric properties.

- Karl Popper (1902–1994)

6.2 Theories about Probability

Assessing probabilities.

- Whatever concept of probability is used in an application, it must be clear how probabilities are assigned to events.

- How is empirical information processed?
  Concepts also differ in this respect.

- We shall outline some ideas for two concepts.

6.3 The Bayesian Approach

Bayesianism.

Hacking:

> "Belief-type dogmatists who make exclusive use of belief-type probabilities and place great emphasis on Bayes's rule are often called Bayesians."

6.3 The Bayesian Approach

Bayesianism.

We shall now see, using personal probabilities:

1. How to use numbers to represent personal beliefs.
2. Why these numbers should satisfy
   the basic rules of probability.
3. How to learn from experience, using Bayes's rule.

6.3 The Bayesian Approach

Calibrating a personal probability, using a bet.

- Betting on/against an event or proposition E at odds $r/(1-r)$, with total stake €S:

	E	$\neg E$
bet on E	win €$(1-r)S$	lose €rS
bet against E	lose €$(1-r)S$	win €rS
abstain from betting	status quo	status quo

- An option is called fair if the agent is indifferent between it and abstaining.

- A fair betting rate is the agent's personal probability of E.

6.3 The Bayesian Approach

Example.

- Let $E =$ "This coin will fall H when tossed the next time."

- If you believe the coin is unbiased, this bet is fair:

	E	$\neg E$
bet on E	win €1	lose €1
bet against E	lose €1	win €1

- If you believe the coin is unbiased, this bet is unfair:

	E	$\neg E$
bet on E	win €1	lose €2
bet against E	lose €1	win €2

(When would you find this bet fair?)

6.3 The Bayesian Approach

Example.

- What is wrong with this bet:

	E	$\neg E$
bet on E	win €1	lose €2
bet against E	lose €2	win €1
abstain from betting	status quo	status quo

- Bets can be combined to guarantee a sure gain (whatever the coin is like).

- Coherence is violated.

6.3 The Bayesian Approach

The Bayes theorem and its purpose.

- The Bayes theorem states that

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

- $P(A)$: prior probability of A
  (before we know if B occurs or not)

- $P(A|B)$: posterior probability of A
  (after we know B has occurred)

- The prior $P(A)$ is augmented to the posterior $P(A|B)$, using the empirical evidence that B has occurred.

6.3 The Bayesian Approach

Example 1: Sampling from boxes.

- Two boxes with 5 balls each:
  - Box I: 2 white, 3 black.
  - Box II: 4 white, 1 black.
- A ball is drawn from one of the boxes. This ball is black. It is not known from which box the ball was drawn.
- Probability that the drawn ball comes from Box I?

6.3 The Bayesian Approach

Example 1: Sampling from boxes.

- Define events:

$$\begin{aligned} A_1 &= \text{Box I was chosen,} \\ A_2 &= \text{Box II was chosen,} \\ B &= \text{the drawn ball is black.} \end{aligned}$$

- Problem: Find $P(A_1|B)$!

- Solution: Apply

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2)}$$

6.3 The Bayesian Approach

Example 1: Sampling from boxes.

- To apply

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2)},$$

  we need the prior probabilities: $P(A_1)$, $P(A_2)$.

- We set: $P(A_1) = P(A_2) = \frac{1}{2}$ (reflecting ignorance!)

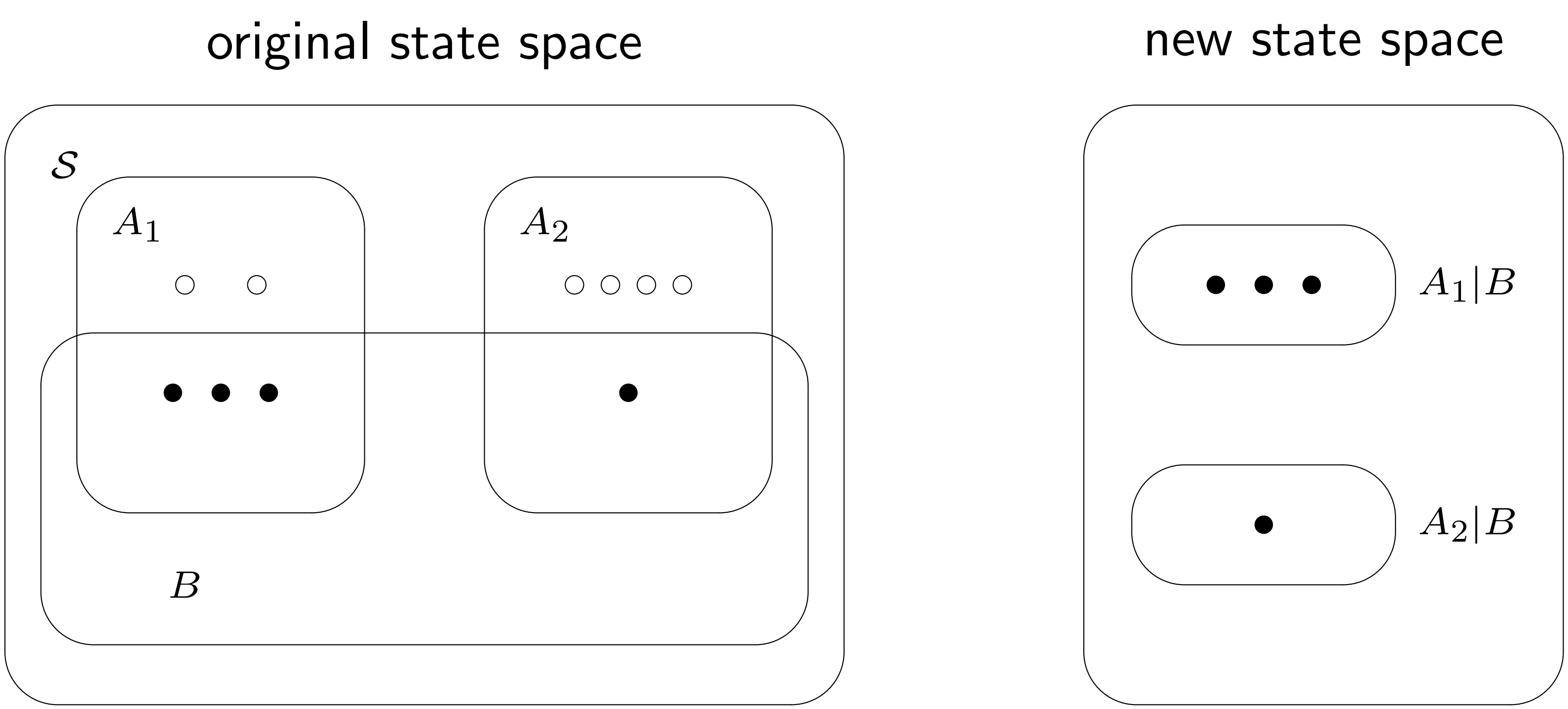
- Then,

$$P(A_1|B) = \frac{\frac{3}{5} \cdot 0.5}{\frac{3}{5} \cdot 0.5 + \frac{1}{5} \cdot 0.5} = 0.75.$$

6.3 The Bayesian Approach

Example 1: Sampling from boxes.

Illustration of Bayes's theorem:



6.3 The Bayesian Approach

Example 2: HIV tests.

- No HIV test is 100% accurate:
  It may produce a false positive or a false negative result.

- A person is tested. Define events:

  A: person is HIV infected;
  $\bar{A}$: person is not HIV infected;

  B: test result is positive;
  $\bar{B}$: test result is negative.

6.3 The Bayesian Approach

Example 2: HIV tests.

- Probabilities and their names in epidemiology:

$$\begin{aligned} \text{sensitivity} &= P(B|A), \\ \text{specificity} &= P(\bar{B}|\bar{A}), \\ \text{prevalence} &= P(A), \end{aligned} \qquad \begin{aligned} \text{PPV} &= P(A|B), \\ \text{NPV} &= P(\bar{A}|\bar{B}). \end{aligned}$$

- Suppose a person has tested positive. What is $P(A|B)$?

- Bayes:

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

6.3 The Bayesian Approach

Example 2: HIV tests.

- For a numerical example, let's assume:
  $P(B|A) = 0.98, \;\; P(\bar{B}|\bar{A}) = 0.99.$

- Scenario 1: Low HIV prevalence: $P(A) = 0.001$. Then:

$$P(A|B) = \frac{0.98 \cdot 0.001}{0.98 \cdot 0.001 + 0.01 \cdot 0.999} = 0.089.$$

  91% false positives! Most probably no infection in spite of positive test.

- Scenario 2: High HIV prevalence: $P(A) = 0.1$. Then:

$$P(A|B) = \frac{0.98 \cdot 0.1}{0.98 \cdot 0.1 + 0.01 \cdot 0.9} = 0.916.$$

  8.4% false positives. Less uncertainty, PPV is much higher.

6.3 The Bayesian Approach

Example 2: HIV tests.

- What we can learn from the test result depends on what we already know or assume to know.
- The Bayes theorem processes the information contained in the experiment (the test).
- A positive test result need not indicate infection.
- A negative test result need not indicate the absence of infection.

6.3 The Bayesian Approach

Learning about a parameter.

- Let $p \in [0, 1]$ be the probability of event A; it is unknown.
- p is conceived as a random variable.
- Knowledge about p is expressed in the distribution of p.
- As empirical evidence becomes available, the distribution of p is modified.

6.3 The Bayesian Approach

Learning about a parameter.

- Suppose the event A occurred x times in n trials.

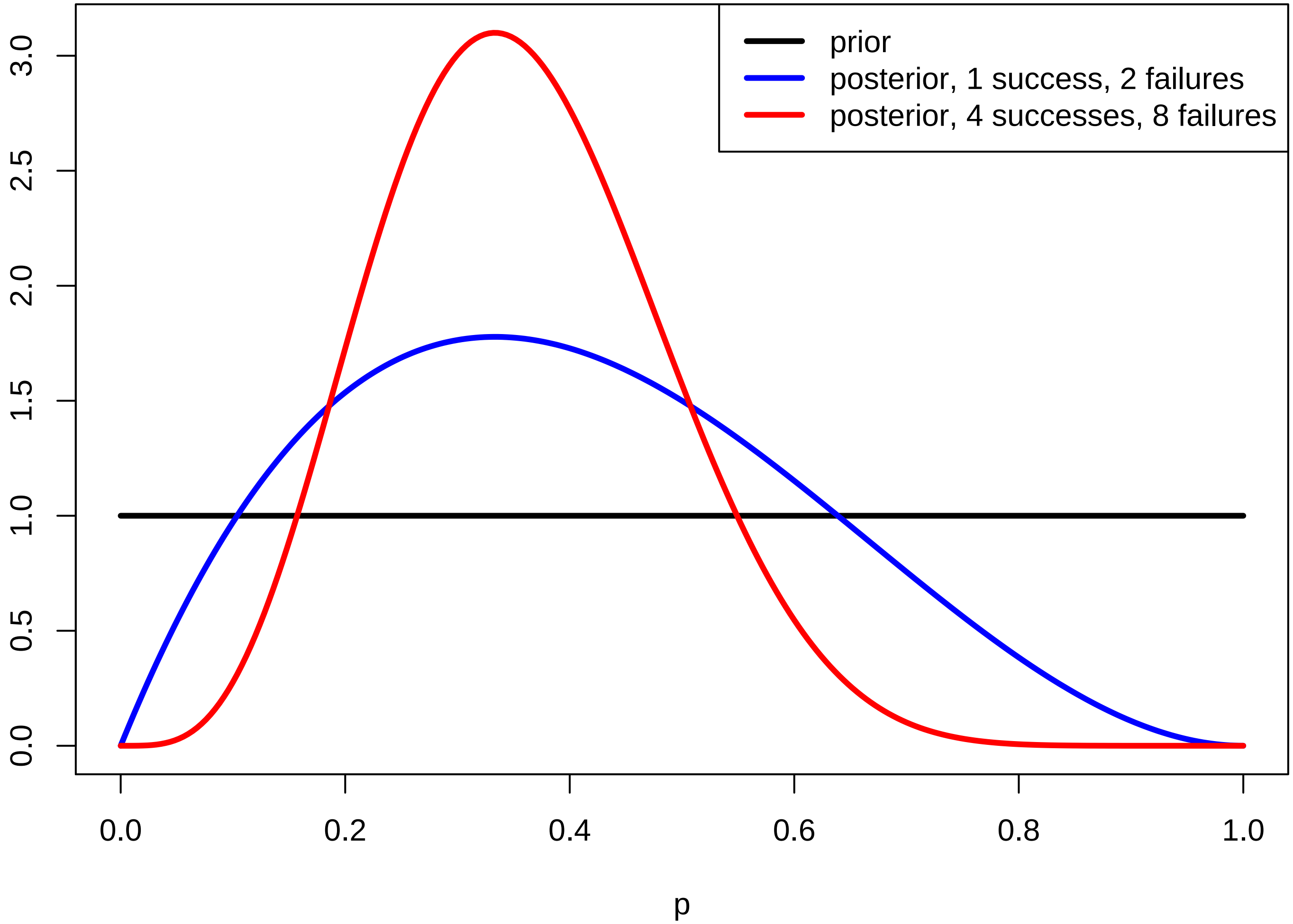
- Then:

$$\underbrace{f(p|x)}_{\text{posterior}} \propto \underbrace{p^x(1-p)^{n-x}}_{\text{likelihood}} \times \underbrace{f(p)}_{\text{prior}}$$

- If $f(p) > 0$ for all $p \in [0, 1]$, the sequence of posteriors after n trials converges, with a limit independent of the prior.

6.3 The Bayesian Approach

Example of prior and posteriors.



6.4 Probability as Frequency

Repeated experiments.

- Idea: Experiments can be repeated.
- The notion of iid random variables is very important.
  (As model itself or as an ingredient for a more complex model.)
- High emphasis on limit theorems.
  (What happens as an experiment is repeated very often?)
- Many quality criteria for methods of statistical inference are based on frequency.

6.4 Probability as Frequency

Repeated experiments.

- The two most important limit theorems in this context are:
  - the law of large numbers
  - the central limit theorem (CLT)

  (See Chapter 9.)

- This course is based on what *today* seems to be "mainstream statistics".