STA421
Foundations of Bayesian Methodology
FS22

Małgorzata Roos
Department of Biostatistics at Epidemiology, Biostatistics and Prevention Institute,
University of Zurich, Hirschengraben 84, CH-8001 Zurich, Switzerland
malgorzata.roos@uzh.ch

February 22, 2022

Contents

Chapter 1

Lecture 1: Classical vs Bayesian paradigms and conditional probability

1.1 Overview of the lecture

Bayesian methods combine prior knowledge with observed data and are powerful tools for data analysis in many domains of science. However, underlying concepts, derivations, and computations can be challenging. This lecture reviews fundamental concepts of Bayesian methodology and provides an accessible introduction to theoretical and practical tools with medical applications. A successful participant will be able to apply Bayesian methods in other areas of research.

Probability calculus	Distributions	Change of variables formula
Priors	MC sampling	Asymptotics

Bayes	**Classical**
Posterior $\propto$ Likelihood $\times$ Prior	Likelihood

Conjugate Bayes	MCMC sampling	Bayesian logistic regression
Predictive distributions	JAGS	Bayesian meta-analysis
Prior elicitation	CODA	Bayesian model selection

Table 1.1: Foundations of Bayesian Methodology: content of the lecture.

1.2 Overview of the individual project

Table 2 of Baeten et al. [2013] provides results of a Bayesian analysis of ASA20 responders at week 6 for Secukinumab and Placebo. This case-control study considers Ankylosing spondylitis in an experimental treatment with Secukinumab (monoclonal antibody) and uses historical controls. The primary binary endpoint ASAS20 indicates patients with a 20% response according to the Assessment of Spondylo Arthritis international Society criteria for improvement at week 6.

A classical clinical trial would for example use a 1:1 sampling with n=24 patients in the treatment group and n=24 patients in the placebo group. This Bayesian analysis uses a smaller number of patients. It applies a 4:1 study design with n=24 patients in the treatment group and only n=6 patients in the placebo group, but uses 8 similar historical placebo-controlled clinical trials to derive an informative prior for the placebo group instead.

Potential benefits of Bayesian analysis

- Reduces the number of placebo patients in the new trial
- Decreases costs
- Shortens trials duration ($\rightarrow$ faster decision)
- Facilitates recruitment ($\rightarrow$ faster decision)
- Can be more ethical in some situations

Secukinumab		**Placebo**
Sample size computation		
		Bayesian meta-analysis Prior elicitation
Beta(0.5, 1)		Beta(11, 32)
Data		Data
Posterior (S)		Posterior (P)
	Posterior probability of superiority	

Table 1.2: A sketch of analysis steps leading to the results provided in Table 2 of Baeten et al. [2013]. For your individual project you are asked to conduct this analysis in several small steps and provide a report of your findings.

- An intermediate study can be conducted at any timepoint

Potential dangers of Bayesian analysis

- Posteriors hinge on the prior elicited for the placebo group
- The prior elicited for the placebo group depends on the prior for the between-study precision in a Bayesian meta-analysis

1.3 History

The history of both the Bayesian and the classical approaches to statistics is intertwined. This section reviews the most relevant historical facts.

Bayes

INDUCTIVE LOGIC
(θ) before $\longleftarrow$ after ($\mathbf{y}$)
before: possible, probable causes
after: effects, results

- James Bernoulli (1713)
- Reverend Thomas Bayes (1763)
- Laplace (1812)

Bayes Theorem
timeline $A \to B$:

$$P[A \mid B] = \frac{P[B \mid A]P[A]}{P[B]}$$

or

$$P[\theta \mid \mathbf{y}] = \frac{P[\mathbf{y} \mid \theta]P[\theta]}{P[\mathbf{y}]}$$

or

$$P[H \mid D] = \frac{P[D \mid H]P[H]}{P[D]}$$

- quantification of evidence
- Bayes factor

1940 Physics
1950 MCMC Metropolis Hastings
1980 Gibbs Sampling
1990 WinBUGS
... OpenBUGS, JAGS, Stan, INLA,
Variational Bayes, bayesmeta

Classical

DEDUCTIVE LOGIC
(θ) before $\longrightarrow$ after ($\mathbf{y}$)
before: causes
after: results
general rules, promises (θ) lead to certain results and conclusions ($\mathbf{y}$)

- Pearson, Galton - 1890, 1900
- Gosset, Fisher - 1910, 1920
- Pearson, Neyman - 1930

Likelihood
timeline $A \to B$:
Only interested in $P[B \mid A]$ or $P[\mathbf{y} \mid \theta]$

- 95% confidence intervals
- tests
- p-values
- statistical programs

Nowadays, parallel usage of Bayesian and classical paradigms is quite common. See, for example, your individual project motivated by Baeten et al. [2013].

Note that the Bayes approach is also based on the likelihood. Therefore, all problems for classical inference such as uncertainty about the sampling model, randomness of the data (outliers) and model complexity propagate.

However, Bayes needs more work. For example, priors $P[\theta]$ must be elicitated from contextual information. Contextual information is usually provided by mean, standard deviation, minimum, maximum (range). A good understanding of properties of different distributions is necessary in order to define a correct $P[\theta]$ prior. See, for example, distributions zoo: Leemis and McQueston [2008]. Moreover, good communication with experts is necessary to get the correct information. Bayesian computation comprises:

- Conjugate analyses
- MCMC sampling: R, JAGS, OpenBUGS, Stan
- Bayesian numerical approximations: INLA, bayesmeta

Recommended reading: Bayarri and Berger [2004] and Martin et al. [2020]. You can also check interactive visualizations Seeing Theory `http://students.brown.edu/seeing-theory/index.html`

1.4 Probability calculus

The probability calculus is based on three axioms:

$P[A] \geq 0$

$P[A] = 1$ if A is true

$P[A \text{ or } B] = P[A \cup B] = P[A] + P[B]$ if $A \cap B = \emptyset$ and $P[A \text{ and } B] = P[A \cap B] = 0$ (events A and B are mutually exclusive).

There are several important properties of probabilities.

Conditional probability

$$P[A \mid B] = \frac{P[A \text{ and } B]}{P[B]} = \frac{P[A \cap B]}{P[B]}, \tag{1.1}$$

given that $P[B] > 0$.

Two events A and B are called independent if the occurrence of B does not change the probability of A

$$P[A \mid B] = P[A]$$

and vice versa

$$P[B \mid A] = P[B].$$

Thus,

$$P[A \text{ and } B] = P[A]P[B].$$

Note that from Equation (1.1)

$$P[A \text{ and } B] = P[A \mid B]P[B] = P[B \mid A]P[A].$$

This observation leads to the **Bayes theorem**

$$P[A \mid B] = \frac{P[B \mid A]P[A]}{P[B]}. \tag{1.2}$$

Assume that event A has a disjoint, complementary event A^c such that $P[A]+P[A^c] = 1$. Conditional probabilities behave like ordinary probabilities, so that we have

$$P[A \mid B] + P[A^c \mid B] = 1.$$

This leads to the simplest version of the law of total probability:

$$P[B] = P[B \mid A]P[A] + P[B \mid A^c]P[A^c].$$

Therefore, the **Bayes theorem** from Equation (1.2) can be rewritten as

$$P[A \mid B] = \frac{P[B \mid A]P[A]}{P[B \mid A]P[A] + P[B \mid A^c]P[A^c]}. \tag{1.3}$$

For formulas applying to more than two events see Held and Sabanés Bové [2020, Sections A.1.1–A.1.2].

Note that there is a link between probability P and odds O:

$$O = \frac{P}{1-P}$$

and

$$P = \frac{O}{1+O}.$$

We can obtain the odds form for the Bayes theorem

$$\frac{P[A \mid B]}{P[A^c \mid B]} = \frac{P[B \mid A]}{P[B \mid A^c]}\frac{P[A]}{P[A^c]}, \tag{1.4}$$

by dividing Equation (1.2) by the same equation applied to the disjoint, complementary event A^c instead of A.

The ratio

$$\frac{P[A \mid B]}{P[A^c \mid B]} = \frac{P[A \mid B]}{1 - P[A \mid B]}$$

is called posterior odds, and

$$\frac{P[A]}{P[A^c]} = \frac{P[A]}{1 - P[A]}$$

is called prior odds, and the ratio

$$\frac{P[B \mid A]}{P[B \mid A^c]}$$

is the Bayes factor (likelihood ratio).

Remark: This Bayes factor is a measure of evidence [Held and Ott, 2018] of the null hypothesis H_0 against an alternative hypothesis H_A, when we replace events A and A^c in Equation (1.4) by H_0 and H_A.

Remark: One can also derive a conditional version of the Bayes theorem

$$P[A \mid B, I] = \frac{P[A \mid I]P[B \mid A, I]}{P[B \mid I]},$$

where I is an additional piece of information.

Recommended reading: Held and Sabanés Bové [2020]: Sections 6.1 and 6.2, A1, A1.1, A1.2, A2.1, A2.2, A2.3. See also Rouder and Morey [2019] for a deeper insight into the meaning of the Bayes theorem.

1.5 Example: Breast cancer and diagnostic tests

This section demonstrates on one medical example that $P[D^+ \mid T^+] \neq P[T^+ \mid D^+]$.

Assumptions

Prevalence of breast cancer: $P[D^+] = 0.045$

Sensitivity: $P[T^+ \mid D^+] = 0.866$

Specificity: $P[T^- \mid D^-] = 0.968$

$$\text{Full bivariate distribution } P[T \cap D] = \begin{array}{c} \\ T^+_{\text{yes}} \\ T^-_{\text{no}} \end{array} \begin{array}{c} \begin{array}{cc} D^+_{\text{yes}} & D^-_{\text{no}} \end{array} \\ \begin{pmatrix} 0.03897 & 0.03056 \\ 0.00603 & 0.92444 \end{pmatrix} \end{array}$$

$$\text{marginal distribution for Test } P[T] = \begin{array}{c} T^+ \\ T^- \end{array} \begin{pmatrix} 0.06953 \\ 0.93047 \end{pmatrix}$$

$$\text{marginal distribution for Disease } P[D] = \begin{array}{c} D^+ \\ D^- \end{array} \begin{pmatrix} 0.045 \\ 0.955 \end{pmatrix}$$

$$P[D^- \mid T^-] = \frac{P[D^- \cap T^-]}{P[T^-]} = \frac{0.92444}{0.93047} = 0.994$$

$$P[D^+ \mid T^+] = \frac{P[D^+ \cap T^+]}{P[T^+]} = \frac{0.03897}{0.06953} = 0.56 \neq P[T^+ \mid D^+]$$

Remark: Another way of computing

$$P[D^+ \mid T^+] = \frac{P[D^+]P[T^+ \mid D^+]}{P[T^+]} = \frac{P[D^+]P[T^+ \mid D^+]}{P[D^+]P[T^+ \mid D^+] + P[D^-]P[T^+ \mid D^-]}$$

1.6 Overview of the classical statistic

The classical statistic is based on the likelihood (Figure 1.1).

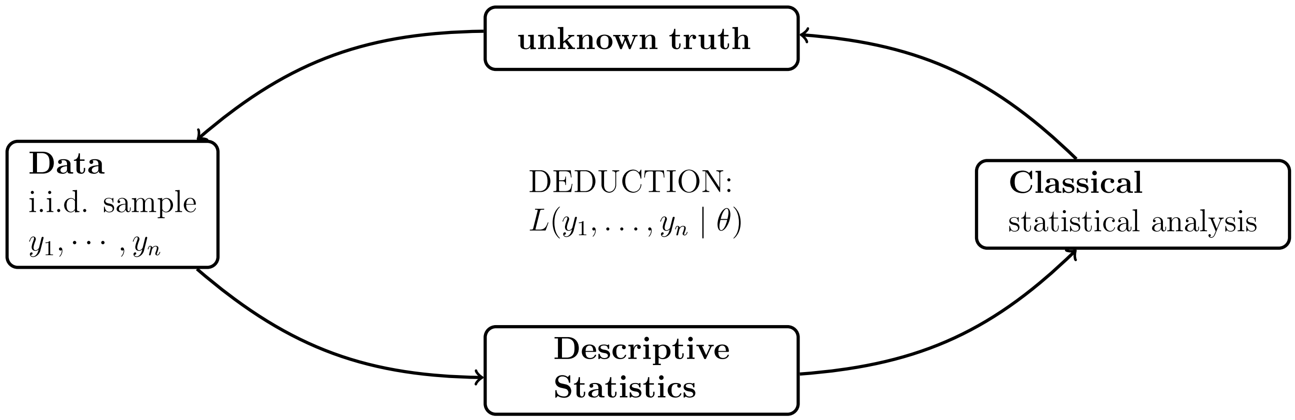

Figure 1.1: Overview of the classical statistic

1.6.1 Example: Primary outcome follows normal distribution

Data $Y \overset{i.i.d.}{\sim} \mathrm{N}(\mu, \sigma^2)$ and $\theta = \mu$.

Density $f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_i - \mu)^2\right\}$.

Likelihood

$$
\begin{aligned}
L(y_1, \dots, y_n \mid \mu) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_i - \mu)^2\right\} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right\},
\end{aligned} \tag{1.5}
$$

and the log-likelihood

$$
\log L(y_1, \dots, y_n \mid \mu) = n \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2.
$$

In order to derive an estimator of μ, compute

$$
\frac{d \log L(y_1, \dots, y_n \mid \mu)}{d\mu} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(y_i - \mu)(-1)\Big|_{\mu = \hat{\mu}} = 0,
$$

$$
\sum_{i=1}^{n} y_i - n\hat{\mu} = 0.
$$

Thus,

$$
\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i.
$$

One can also derive that $\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{\mu})^2}$.

1.6.2 Example: Primary outcome follows Bernoulli distribution

Let $Y \overset{i.i.d.}{\sim} \text{Be}(p)$ with $\theta = p$ and $P[Y = 0] = 1 - p$ and $P[Y = 1] = p$.

Density $f(y_i) = p^{y_i}(1-p)^{1-y_i}$.

Likelihood

$$\begin{aligned} L(y_1, \dots, y_n \mid p) &= \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i} \\ &= p^{\sum_{i=1}^{n} y_i}(1-p)^{n-\sum_{i=1}^{n} y_i}, \end{aligned} \tag{1.6}$$

Log-likelihood

$$\log L(y_1, \dots, y_n \mid p) = \sum_{i=1}^{n} y_i \log p + \left(n - \sum_{i=1}^{n} y_i\right) \log(1-p).$$

In order to derive an estimator of p, compute

$$\frac{d \log L(y_1, \dots, y_n \mid p)}{dp} = \sum_{i=1}^{n} y_i \frac{1}{p} + \left(n - \sum_{i=1}^{n} y_i\right) \frac{1}{1-p}(-1)\Bigg|_{p=\hat{p}} = 0,$$

$$\sum_{i=1}^{n} y_i \frac{1}{\hat{p}} - \left(n - \sum_{i=1}^{n} y_i\right) \frac{1}{1-\hat{p}} = 0,$$

$$\sum_{i=1}^{n} y_i (1-\hat{p}) = \left(n - \sum_{i=1}^{n} y_i\right) \hat{p},$$

$$\sum_{i=1}^{n} y_i = n\hat{p},$$

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} y_i.$$

1.7 Overview of the Bayesian methodology

There is a consent that probability calculus leading to the Bayes formula in Equation (1.2) is objective. Bayesian methodology extends the classical approach based on the likelihood and considers

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}.$$

More specifically,

$$P[\theta \mid y_1, \dots, y_n] \propto \mathrm{L}(y_1, \dots, y_n \mid \theta) \times P[\theta].$$

Figure 1.2 provides an overview of the Bayesian methodology and its relation to the classical statistics.

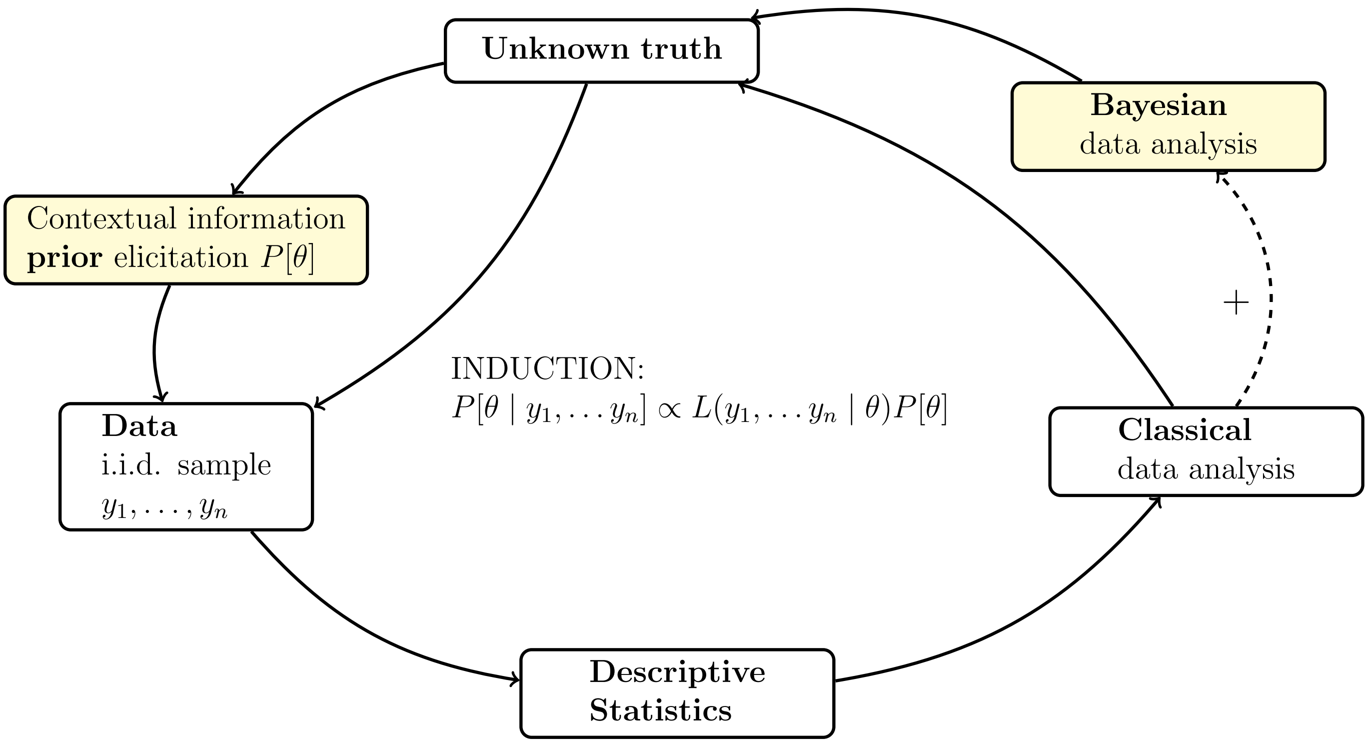

Figure 1.2: Overview of the Bayesian methodology. Fields with yellow background correspond to Bayes-specific steps.

1.7.1 Bayes factors and p-values

The Bayesian methodology enables an independent view of classical hypothesis testing. In applications, the estimation of the credibility of a conclusion expressed by the probability of H_0 given the data is usually of primary interest. The Bayes factor directly quantifies whether the data have increased or decreased the odds of H_0. Thus, Bayes factors facilitate direct conclusions about the probability of H_0 given the data, provided that both null H_0 and alternative H_1 hypotheses have been specified.

On page 70, Held and Sabanés Bové [2020] define the p-value: the probability, under the assumption of the null hypothesis H_0, of obtaining a result equal to or more extreme than what was actually observed. A p-value is computed under the assumption that the null hypothesis H_0 is true. It does not allow for conclusions about the probability of H_0 given the data. A particular p-value can be obtained either for a large study with a small effect or for a small study with a large effect. Thus, the p-value does not say anything about the actual effect or evidence that such an effect exists.

Consider a significance test with a point null hypothesis $H_0 : \theta = \theta_0$. The alternative hypothesis can be either simple $H_1 : \theta = \theta_1 \neq \theta_0$ or composite $H_1 : \theta \neq \theta_0$. For a composite H_1 a prior distribution $f(\theta \mid H_1)$ must be specified.

Note that $P[H_1] = 1 - P[H_0]$ and $P[y] = f(y \mid H_0)P[H_0] + f(y \mid H_1)P[H_1]$. The Bayes formula for H_0

$$P[H_0 \mid y] = \frac{f(y \mid H_0)P[H_0]}{P[y]} \tag{1.7}$$

divided by the Bayes formula for H_1

$$P[H_1 \mid y] = \frac{f(y \mid H_1)P[H_1]}{P[y]} \tag{1.8}$$

render

$$\frac{P[H_0 \mid y]}{P[H_1 \mid y]} = BF_{01}(y)\frac{P[H_0]}{P[H_1]}, \tag{1.9}$$

where

$$BF_{01}(y) = \frac{f(y \mid H_0)}{f(y \mid H_1)}. \tag{1.10}$$

Note that the Bayes factor $BF_{01}(y)$ transforms the prior odds $P[H_0]/P[H_1]$ into posterior odds $P[H_0 \mid y]/P[H_1 \mid y]$ in the light of the data y. $BF_{01}(y)$ is a direct quantitative measure of how data y have increased or decreased the odds of H_0 and is referred to as the strength of evidence for or against H_0. The evidence against the null hypothesis H_0 is provided by small Bayes factors $BF_{01}(y) < 1$. The evidence in favor of the null hypothesis H_0 is provided by large Bayes factors $BF_{01}(y) > 1$. Table 2 of Held and Ott [2018] provides a categorization of Bayes factors $BF_{01}(y) \leq 1$ into levels of evidence against H_0: weak (1 to 1/3), moderate (1/3 to 1/10), substantial (1/10 to 1/30), strong (1/30 to 1/100), very strong (1/100 to 1/300), and decisive (< 1/300).

$BF_{01}(y)$ is the ratio of the likelihood $f(y \mid H_0) = f(y \mid \theta = \theta_0)$ of the observed data y under the null hypothesis H_0 and the marginal likelihood

$$f(y \mid H_1) = \int f(y \mid \theta)f(\theta \mid H_1)d\theta \tag{1.11}$$

under the alternative hypothesis H_1. Equation (1.11) is useful for composite alternative hypotheses H_1. It is the average likelihood $f(y \mid \theta)$ with respect to the prior distribution $f(\theta \mid H_1)$ for θ under the alternative H_1, which is called marginal likelihood (prior predictive distribution at the observed data). For a simple alternative, Equation (1.11) reduces to the likelihood $f(y \mid H_1) = f(y \mid \theta = \theta_1)$ and the $BF_{01}(y)$ reduces to a likelihood ratio.

Once we know $BF_{01}(y)$, we can solve the formula in Equation (1.9) for the posterior probability of H_0. Note that

$$\frac{P[H_0 \mid y]}{1 - P[H_0 \mid y]} = BF_{01}(y)\frac{P[H_0]}{P[H_1]}.$$

Thus,

$$P[H_0 \mid y] = \frac{BF_{01}(y)\frac{P[H_0]}{P[H_1]}}{1 + BF_{01}(y)\frac{P[H_0]}{P[H_1]}}. \tag{1.12}$$

Note that Bayes factors facilitate multiple hypothesis comparisons because they can be updated sequentially:

$$BF_{01}(y)BF_{12}(y) = \frac{f(y \mid H_0)}{f(y \mid H_1)}\frac{f(y \mid H_1)}{f(y \mid H_2)} = \frac{f(y \mid H_0)}{f(y \mid H_2)} = BF_{02}(y).$$

The minimum Bayes factor is the smallest Bayes factor within a certain class of alternative hypotheses. Minimum Bayes factors are very interesting because they quantify the maximal evidence of a p-value against a point H_0 within a certain class of alternative hypotheses.

The Bayesian approach provides a way of transforming p-values to direct measures of evidence against the null hypothesis expressed by Bayes factors. This transformation is called calibration. Held and Ott [2018] consider different transformations of p-values to minimum Bayes factors and show that minimum Bayes factors provide less evidence against the null hypothesis than the corresponding p-value might suggest. They also demonstrate that many techniques have been proposed to calibrate p-values and there is no consensus which calibration is the optimal one.
Recommended reading: Held and Sabanés Bové [2020] Sections 3.3 and 7.2.1, Goodman [1999b], Goodman [1999a], Held and Ott [2018], and `pCalibrate` package.

Example: Discuss `pCalibrate` to show the calibration of p-values by Bayes factors on the border between the classical and the full Bayes analysis.

1.7.2 Priors

The use of prior distributions for Bayesian analysis can be controversial. Therefore, a good understanding of different distributions is very important.

- Discussion of different distributions Leemis and McQueston [2008].
- Monte Carlo (MC) simulations vs true parameters (expectation and variance).
- The Change-of-Variables Formula Held and Sabanés Bové [2020, Section A.2.3].

  Assume a one-to-one and differentiable transformation $g(\dot{})$. Assume that the random variable Y with probability density function $f_Y(y)$ is a transformation of a continuous random variable X with probability density function $f_X(x)$, where g is a one-to-one and differentiable transformation and $Y = g(X)$. Then

$$f_Y(y) = f_X\big(g^{-1}(y)\big)\left|\frac{dg^{-1}(y)}{dy}\right|.$$

1.8 Worksheet 1

Probability calculus	**Distributions**	Change of variables formula
Priors	**MC sampling**	Asymptotics

Bayes	**Classical**
Posterior $\propto$ Likelihood $\times$ Prior	**Likelihood**

Conjugate Bayes	MCMC sampling	Bayesian logistic regression
Predictive distributions	JAGS	Bayesian meta-analysis
Prior elicitation	CODA	Bayesian model selection

Table 1.3: Foundations of Bayesian Methodology: content of the lecture relevant for Worksheet 1.

Acknowledgement:

We thank Sona Hunanyan for typing an earlier version of this script for the Bayesian Data Analysis lecture in the spring term 2018.

Secukinumab		**Placebo**
Classical Sample size computation		
		Bayesian meta-analysis Prior elicitation
Beta(0.5, 1)		Beta(11, 32)
Data		**Data**
	Classical analysis	
Posterior (S)		Posterior (P)
	Posterior probability of superiority	

Table 1.4: Individual project: A sketch of analysis steps leading to the results provided in Table 2 of Baeten et al. [2013]. For your individual project you are asked to conduct this analysis in several small steps and provide a report of your findings.

Bibliography

D. Baeten, X. Baraliakos, J. Braun, J. Sieper, P. Emery, D. van der Heijde, I. McInnes, J. van Laar, R. Landewé, P. Wordsworth, J. Wollenhaupt, H. Kellner, J. Paramarta, J. Wei, A. Brachat, S. Bek, D. Laurent, Y. Li, Y.A. Wang, A.P. Bertolino, S. Gsteiger, A.M. Wright, and W. Hueber. Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. *The Lancet*, 382:1705–1713, 2013.

M.J. Bayarri and J.O. Berger. An interplay of Bayesian and frequentist analysis. *Statistical Science*, 19(1):58–80, 2004.

S.N. Goodman. Toward evidence-based medical statistics. 2: The Bayes Factor. *Annals of Internal Medicine*, 130(12):1005–1013, 1999a. URL `https://doi.org/10.7326/0003-4819-130-12-199906150-00019`.

S.N. Goodman. Toward evidence-based medical statistics. 1: The p value fallacy. *Annals of Internal Medicine*, 130(12):995–1004, 1999b. URL `https://doi.org/10.7326/0003-4819-130-12-199906150-00008`.

L Held and M. Ott. On p-values and Bayes Factors. *Annual Review of Statistics and Its Application*, 5:393–419, 2018. URL `https://doi.org/10.1146/annurev-statistics-031017-100307`.

L. Held and D. Sabanés Bové. *Likelihood and Bayesian Inference: With Applications in Biology and Medicine.* Springer, 2020. URL `https://www.springer.com/gp/book/9783662607916`.

L.M. Leemis and J.T. McQueston. Univariate Distribution Relationships. *The American Statistician*, 62(1):45–53, 2008. doi: 10.1198/000313008X270448. URL `https://doi.org/10.1198/000313008X270448`.

G.M. Martin, D.T. Frazier, and C.P. Robert. Computing Bayes: Bayesian computation from 1763 to the 21st century. *arXiv:2004.06425*, 2020. URL `https://arxiv.org/abs/2004.06425`.

J.N. Rouder and R.D. Morey. Teaching Bayes' theorem: Strength of evidence as predictive accuracy. *The American Statistician*, 73(2):186–190, 2019. URL `https://doi.org/10.1080/00031305.2017.1341334`.