McGill University

Department of Epidemiology
and Biostatistics

Bayesian Analysis for the Health Sciences

Course EPIB-682

Lawrence Joseph

Intro to Bayesian Analysis for the Health Sciences – EPIB-682 – 2 credits

Instructor:	Lawrence Joseph
Email address:	Lawrence.Joseph@mcgill.ca (best way to reach me)
Home page:	http://www.medicine.mcgill.ca/epidemiology/Joseph/
Telephone:	934-1934 X 44713
Address:	Division of Clinical Epidemiology Royal Victoria Hospital V Building Room V2.10

Course Objectives and Topics Covered: To provide researchers with an introduction to practical Bayesian methods. Topics will include Bayesian philosophy, simple univariate models, linear and logistic regression and hierarchical models. Numerical techniques including Monte Carlo integration, sampling importance resampling (SIR), and the Gibbs sampler will be covered, including programming in R and WinBUGS.

Place and Time: September 7 to November 30, 2017. Thursdays, 12:30 PM to 2:30 PM. Room 25, Purvis Hall, 1020 Pine Avenue West, corner Peel Street.

Assessment: Five assignments of approximately 5 questions each. Each assignment is worth 20%. There will be no exams.

Textbook (reference only): A. Gelman, J. Carlin, H. Stern and D. Rubin, Bayesian Data Analysis, 2nd Edition, Chapman and Hall, 2003.

Prerequisites: At least two previous courses in statistics, including topics such as inferences for means and proportions, and linear and logistic regression. Differential and integral calculus. If you are unsure you have sufficient background, please speak to the instructor.

<u>Bayesian Analysis in the Health Sciences</u>

Course Outline – EPIB–682

Date	Topic Covered
Thurs Sept 7	Introduction/Evaluation/Motivation/Background
Thurs Sept 14	Basic Elements of Bayesian Analysis
Thurs Sept 21	Bayesian Philosophy
Thurs Sept 28	Simple Univariate Models
Thurs Oct 5	Computation and Numerical Methods I - Introduction
Thurs Oct 12	Computation and Numerical Methods II - Monte Carlo Integration
Thurs Oct 19	Computation and Numerical Methods III - SIR Algorithm
Thurs Oct 26	Computation and Numerical Methods IV - Gibbs sampler and WinBUGS
Thurs Nov 2	Computation and Numerical Methods V - More on WinBUGS
Thurs Nov 9	Bayesian Linear and Logistic Regression
Thurs Nov 16	Hierarchical Linear and Logistic Regression
Thurs Nov 23	Prior Distributions - Prior Selection and Elicitation
Thurs Nov 30	Model Selection in Regression - Bayes Factors

Bayesian Probabilities - Discrete Case of Bayes Theorem

It is easy to get confused between Bayesian analysis as an inferential paradigm, and Bayes Theorem as a basic way to manipulate discrete probabilities. Let us first consider the discrete case:

Suppose we are considering a test for cancer:

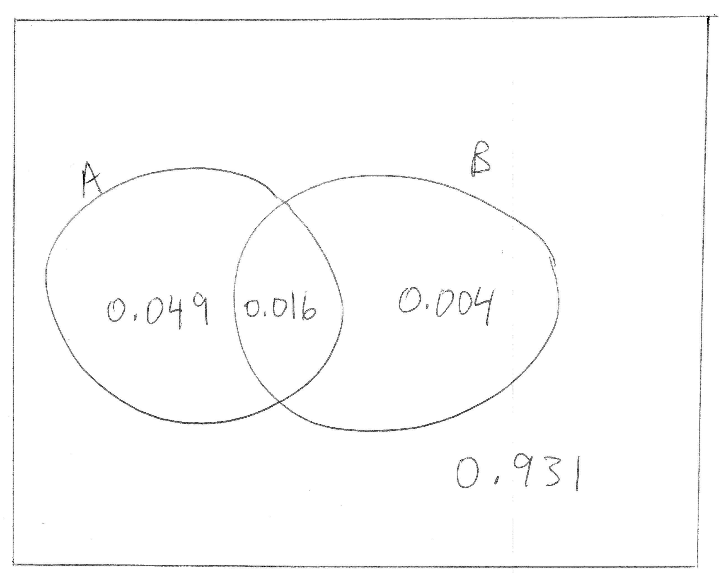
Let A = the event that a test is positive.
Let B = the event of actually having cancer.

Suppose we know that:

- $P(A|B^c) = 0.05$, and so $P(A^c|B^c) = 1 - 0.05 = 0.95$
- $P(A^c|B) = 0.20$, and so $P(A|B) = 1 - 0.20 = 0.80$
- $P(B) = 0.02$, and so $P(B^c) = 0.98$

(a) What is the probability of cancer given that the test is positive?
(b) What is the probability of cancer given that the test is negative?

We can draw a diagram as below:



From the diagram, we see that

$$P(B|A) = \frac{0.016}{0.016 + 0.049} = .2462$$

and

$$P(B|A^c) = \frac{0.004}{0.004 + 0.931} = .0043$$

Alternatively, we can use Bayes Theorem, which states:

$$P(B|A) = \frac{P(B) \times P(A|B)}{P(B) \times P(A|B) + P(B^c) \times P(A|B^c)}$$

Plugging in the numbers, we can check that the solutions are the same. For example,

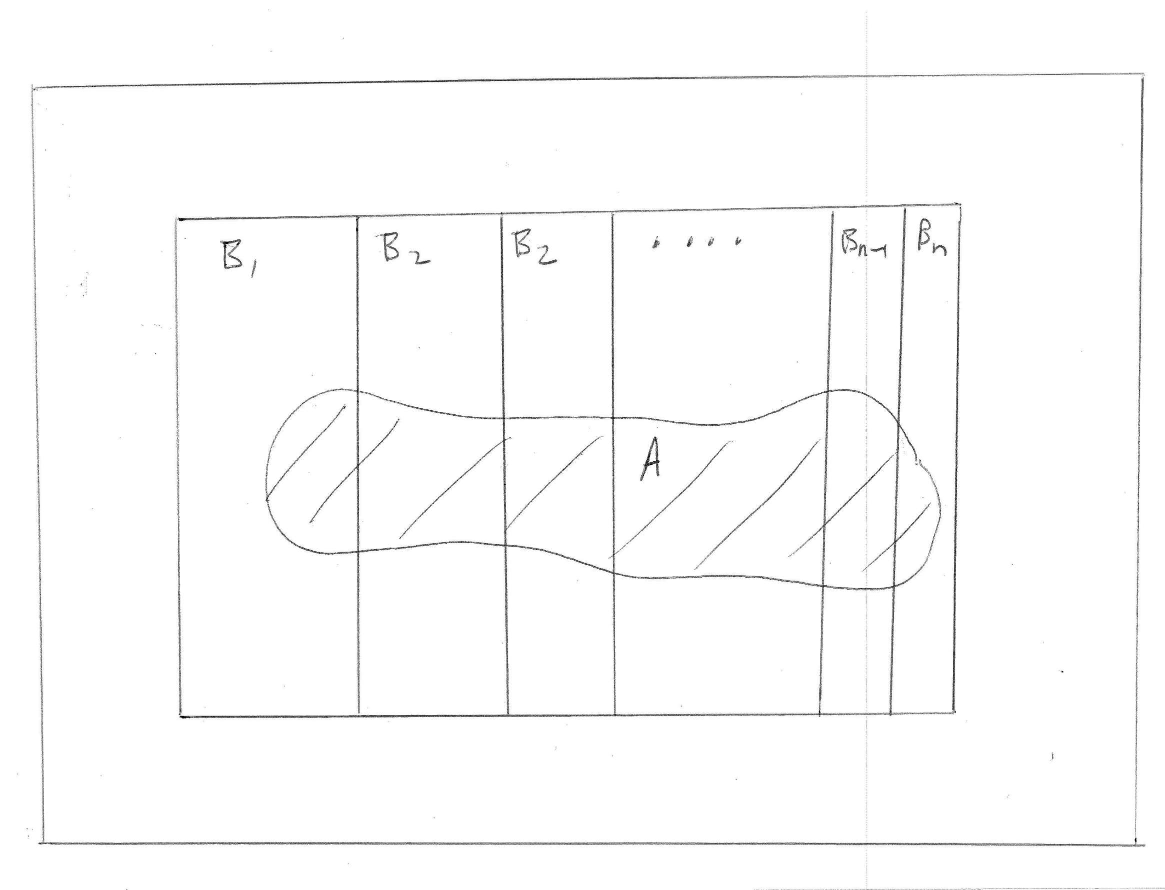
$$P(B|A) = \frac{P(B) \times P(A|B)}{P(B) \times P(A|B) + P(B^c) \times P(A|B^c)} = \frac{0.02 \times 0.80}{0.02 \times 0.80 + 0.98 \times 0.05} = .2462.$$

Switching the roles of A and A^c in the above formula yields

$$P(B|A^c) = \frac{P(B) \times P(A^c|B)}{P(B) \times P(A^c|B) + P(B^c) \times P(A^c|B^c)} = 0.0043$$

Note that before the test is performed, the probability that a person has cancer is 0.02, but that these probabilities are "updated" in a natural way, once the test results become available.

Bayes Theorem may be generalized to the case where the event B has more than two possible outcomes, say $B_1,\ B_2, \ldots, B_n$.



In this case, the Bayes Theorem is

$$P(B_k|A) = \frac{P(B_k) \times P(A|B_k)}{\sum_{i=1}^{n} P(B_i) \times P(A|B_i)}, \quad k = 1, 2, \ldots, n.$$

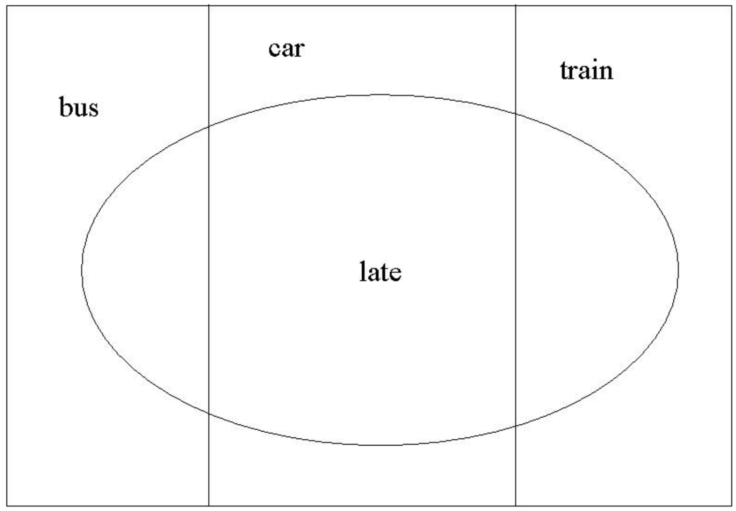
Here is an example for this case:

Suppose that Bob can decide to go to work by one of three modes of transportation, car, bus, or commuter train. Because of high traffic, if he decides to go by car, there is a 50% chance he will be late. If he goes by bus, which has special reserved lanes but is sometimes overcrowded, the probability of being late is only 20%. The commuter train is almost never late, with a probability of only 1%, but is more expensive than the bus.

(a) Suppose that Bob is late one day, and his boss wishes to estimate the probability that he drove to work that day by car. Since he does not know which mode of transportation Bob usually uses, he gives a prior probability of $\frac{1}{3}$ to each of the three possibilities. What is the boss' estimate of the probability that Bob drove to work?

(b) Suppose that a coworker of Bob's knows that he almost always takes the commuter train to work, never takes the bus, but sometimes, 10% of the time, takes the car. What is the coworkers probability that Bob drove to work that day, given that he was late?

Solution: The Venn diagram would be:



(a) We have the following information given in the problem:

$$\begin{aligned} Pr\{\text{ bus }\} = Pr\{\text{ car }\} = Pr\{\text{ train }\} &= \frac{1}{3} \\ Pr\{\text{ late } \mid \text{ car }\} &= 0.5 \\ Pr\{\text{ late } \mid \text{ train }\} &= 0.01 \\ Pr\{\text{ late } \mid \text{ bus }\} &= 0.2 \end{aligned}$$

We want to calculate $Pr\{\text{ car } \mid \text{ late }\}$.

By Bayes Theorem, this is

$Pr\{\text{ car } \mid \text{ late }\}$

$$\begin{aligned} &= \frac{Pr\{\text{ late } \mid \text{ car }\}Pr\{\text{ car }\}}{Pr\{\text{ late } \mid \text{ car }\}Pr\{\text{ car }\} + Pr\{\text{ late } \mid \text{ bus }\}Pr\{\text{ bus }\} + Pr\{\text{ late } \mid \text{ train }\}Pr\{\text{ train }\}} \\ &= \frac{0.5 \times 1/3}{0.5 \times 1/3 + 0.2 \times 1/3 + 0.01 \times 1/3} \\ &= 0.7042 \end{aligned}$$

(b) Repeat the identical calculations as the above, but instead of the prior probabilities being $\frac{1}{3}$, we use $Pr\{\text{ bus}\} = 0$, $Pr\{\text{car}\} = 0.1$, and $Pr\{\text{ train }\} = 0.9$. Plugging in to the same equation with these three changes, we get $Pr\{\text{ car } \mid \text{ late }\} = 0.8475$

This is a simple theorem in probability, having nothing to do with drawing inferences from a data set, that *everybody* uses. Bayes Theorem creates no controversy whatsoever (not that Bayesian inference is so controversial nowadays).

Bayesian Inference - Continuous Case of Bayes Theorem

The above discrete version is different from the continuous version of Bayes Theorem, in that it is typically used for drawing inferences, as an alternative to the freqeuntist approach that leads to p-values and confidence intervals. The continuous version of Bayes Theorem looks like this:

$$\text{posterior distribution} = \frac{\text{likelihood of the data} \times \text{prior distribution}}{\text{a normalizing constant}},$$

or

$$f(\theta|x) = \frac{f(x|\theta) \times f(\theta)}{\int f(x|\theta) \times f(\theta) d\theta,}$$

or, forgetting about the normalizing constant,

$$f(\theta|x) \propto f(x|\theta) \times f(\theta).$$

Thus we "update" the prior distribution to a posterior distribution after seeing the data via Bayes Theorem.

We will see many examples of its use later in the course.

Aspirin		Tylenol	
Cured	Not Cured	Cured	Not Cured
5	5	5	5
6	4	5	5
6	4	4	6
7	3	4	6
8	2	4	6
8	2	3	7
9	1	3	7
$\vdots$	$\vdots$	$\vdots$	$\vdots$
10	0	0	10


MOZ
ART
BEET
VEN


TEA
MILK


HEADS!
HEADS!
HEADS
TAILS!

Effects of Therapeutic Touch On Tension Headache Pain

ELIZABETH KELLER • VIRGINIA M. BZDEK

Therapeutic touch (TT) is a modern derivative of the laying on of hands that involves touching with the intent to help or heal. This study investigated the effects of TT on tension headache pain in comparison with a placebo simulation of TT. Sixty volunteer subjects with tension headaches were randomly divided into treatment and placebo groups. The McGill-Melzack Pain Questionnaire was used to measure headache pain levels before each intervention, immediately afterward, and 4 hours later. A Wilcoxon signed rank test for differences indicated that 90% of the subjects exposed to TT experienced a sustained reduction in headache pain, $p < .0001$. *An average 70% pain reduction was sustained over the 4 hours following TT, which was twice the average pain reduction following the placebo touch. Using a Wilcoxon rank sum test, this was statistically significant,* $p < .01$. *Study results indicated that TT may have potential beyond a placebo effect in the treatment of tension headache pain.*

Therapeutic touch (TT), a modern version of the laying on of hands, was introduced into nursing by Krieger (1975). It does not entail belief in the method or in any other precept on the part of its recipients to be effective (Krieger, 1979). TT may or may not involve contact with the physical body, but contact is said always to be made with the energy field of the healing process (Boguslawski, 1979; Krieger, 1975, 1981).

Background of the Study

Therapeutic touch is based on the philosophy of holism (Krieger, 1981; Weber, 1981) and general systems theory (Battista, 1977). Holism is represented in nursing science by Roger's (1970) theory of unitary man. According to this theory, all persons are highly complex fields of various forms of life energy. These fields of energy are coextensive with the universe and in constant interaction and exchange with surrounding energy fields. The functional basis of TT lies in the direction of life energy through the hands of the therapist to the recipient who may then internalize this energy, use it to restore balance, and thereby self-heal (Boguslawski, 1979; Krieger, 1979, 1981). The predominant theory in recent TT literature concerning the source of the transferred energy is that the therapist serves as a conduit, a channel, so that environmental energy may be transferred to the recipient (Boguslawski, 1979; Weber, 1981). To be recognized as a realistic and tenable phenomenon TT must be considered within a holistic context.

Nurse researchers have investigated TT since Krieger (1976) demonstrated increased hemoglobin

of three groups of 30 hosp diovascular patients: the TT, casual touch, or nc group receiving TT shov cant reduction in state a on the Spielberger Se Questionnaire (Spielberg & Lushene, 1970) p compared with their scores and with the post the other two groups, p

Quinn (1982) replic anxiety study with 60 c patients, but replaced (pulse-taking) with pl which was a simulation ments without energy t of the subjects received tact during either Quinn reasoned that if for TT is an exchange between human fields beyond the skin, direct tact would not be nec indicated that the n group demonstrated lower posttest anxiety s noncontact placebo g degree of anxiety Quinn's noncontact T almost identical to He the physical contact T

Randolph (1984) physiologic response college students to a st film while receiving ei


20, 19, 18,
17, 16, 15,...
HEAL!
HEAL!

Neigh!

Definitely, it's a cocktail!
Nah, it's seafood.
No way. It's a source of nitrites.


BAYES 'R' US
Specializing In:
Customized complex modelling
Posteriors made to measure
Ask about our optimal design package
We require:
Prior(s)
Data
Model(s)
Design
Our guarantee: Direct probability statements for your parameters, accounting for current data and your view of past information.
FREQ CORP INC.
P-VALUES
CONFIDENCE INTERVALS
AIC
STEP METHODS
CLEARANCE SALE ON ALL OUR P-VALUES
SUPPLIED BY CLIENT:
DATA
MODEL(S)
DESIGN
GUARANTEED:
LONG RUN ERROR RATES AVERAGED OVER ALL OUR CUSTOMERS (BUT NO GUARANTEE FOR YOUR DATA).

FRESHER BECAUSE MORE PEOPLE EAT THEM

MORE PEOPLE EAT THEM BECAUSE THEY ARE FRESHER