One-Day Short Course on Bayesian Modeling, Inference and Prediction

Preface, Tentative Syllabus/Outline, and 1: Background and Basics

David Draper

Department of Applied Mathematics and Statistics
University of California, Santa Cruz

draper@ams.ucsc.edu
http://www.ams.ucsc.edu/~draper

Sponsored by the Chicago Chapter of the American Statistical Association

Fri 18 March 2005, 8am–5.30pm

Holiday Inn Chicago Mart Plaza
350 North Orleans St., Chicago, IL

© 2005 David Draper (all rights reserved)
Preface

This material provides a thorough introduction to (or review of) basic methodological ideas and applications in Bayesian modeling, inference and prediction, based on a series of case studies and assuming no previous exposure to Bayesian ideas or methods.

The case studies are drawn from medicine (diagnostic screening for HIV; hospital-specific prediction of patient-level mortality rates; hospital length of stay for premature births; a randomized controlled trial of in-home geriatric assessment) and the physical sciences (measurement of physical constants), but the methods illustrated will apply to a broad range of subject areas in the natural and social sciences.

The course is intended mainly for applied statisticians and will focus on methods and applications rather than theory; an understanding of probability and statistics at the level typically required for a Master’s degree in statistics will provide sufficient mathematical background for participants.

Extensive details required for carrying out the analyses are provided below, including hardcopy of a number of sessions with a frequently-used statistical computing package (R), a leading symbolic computing package (Maple), and one of the two most widely available packages for fitting Bayesian models (BUGS/WinBUGS).
Tentative Syllabus/Outline

8.00–9.00am: Check-in and coffee.

9.00–9.30am: (1) Background and basics.

Quantification of uncertainty: classical, frequentist, and Bayesian definitions of probability. Subjectivity and objectivity. Case study: Diagnostic screening for HIV.

Sequential learning; Bayes’ Theorem. Inference (science) and decision-making (policy and business).

Bayesian decision theory; coherence. Maximization of expected utility.

9.30–11.00am: (2) Exchangeability and conjugate modeling.


Review of frequentist modeling and maximum-likelihood inference.

Exchangeability as a Bayesian concept parallel to frequentist independence.

Prior, posterior, and predictive distributions.

Inference and prediction. Coherence and calibration.

Conjugate analysis. Comparison with frequentist modeling.
Tentative Syllabus/Outline (continued)

11.00–11:15am: Coffee break.

11.15–noon: (3) Integer-valued outcomes; Poisson modeling.

Case Study: Hospital length of stay for birth of premature babies.

noon–12.30pm: (4) Continuous outcomes; Gaussian modeling.

Multivariate unknowns; marginal posterior distributions.
Case Study: Measurement of physical constants (NB10).

The exponential family; conjugate priors.

12.30–1.30pm: Lunch break.

1.30–3.30pm: (5) Simulation-based computation.

IID sampling; rejection sampling.


User-friendly implementation of Gibbs and Metropolis-Hastings sampling via BUGS and WinBUGS.
Case Study: the NB10 data revisited.

MCMC implementation strategies.
3.30–3.45pm: Coffee break.


Poisson fixed-effects modeling. Case study: a randomized controlled trial of in-home geriatric assessment.

Additive and multiplicative treatment effects.

Expansion of a simple model that does not satisfy all diagnostic checks, by embedding it in a richer class of models of which it’s a special case.

Random-effects Poisson regression. Hierarchical modeling with latent variables as an approach to mixture modeling.


4.45pm–5.30pm: (7) Bayesian model specification.

Predictive diagnostics. Model selection as a decision problem.

Bayesian cross-validation as an approach to diagnostics: comparing outcomes from omitted cases with their predictive distributions given the rest of the data. 3CV: 3-way cross-validation.

The log score as a model-selection method, and its relationship to the deviance information criterion (DIC). Case study: continuation of IHGA example.
1: Background and Basics

1.1 Quantification of uncertainty: Classical, frequentist and Bayesian definitions of probability

Case study: Diagnostic screening for HIV

Widespread screening for HIV has been proposed by some people in some countries (e.g., the U.S.).

Two blood tests that screen for HIV are widely available: ELISA, which is relatively inexpensive (roughly $20) and fairly accurate; and Western Blot (WB), which is considerably more accurate but costs quite a bit more (about $100).

A new patient comes to You, a physician, with symptoms that suggest he may be HIV positive (Good, 1950: You = a generic person making uncertainty assessments).

Questions

• Is it appropriate to use the language of probability to quantify Your uncertainty about the proposition $A = \{\text{this patient is HIV positive}\}$?

• If so, what kinds of probability are appropriate, and how would You assess $P(A)$ in each case?

• What strategy (e.g., ELISA, WB, both?) should You employ to decrease Your uncertainty about $A$? If You decide to run a screening test, how should Your uncertainty be updated in light of the test results?
The Meaning of Probability

Statistics might be defined as the study of uncertainty: how to measure it, and what to do about it, and probability as the part of mathematics (and philosophy) devoted to the quantification of uncertainty.

The systematic study of probability is fairly recent in the history of ideas, dating back to about 1650 (e.g., Hacking, 1975).

In the last 350 years three main ways to define probability have arisen (e.g., Oakes, 1990):

- **Classical**: Enumerate elemental outcomes (EOs) in a way that makes them equipossible on, e.g., symmetry grounds, and compute \( P_C(A) = \text{ratio of } n_A = (\text{number of EOs favorable to } A) \text{ to } n = (\text{total number of EOs}) \).

- **Frequentist**: Restrict attention to attributes \( A \) of events: phenomena that are inherently repeatable under “identical” conditions; define \( P_F(A) = \text{limiting value of relative frequency with which } A \text{ occurs as the number of repetitions } \to \infty \).

- **Personal, or “Subjective”, or Bayesian**: Imagine betting with someone about the truth of proposition \( A \), and ask Yourself what odds \( O_{\text{You}} \) (in favor of \( A \)) You would need to give or receive in order that You judge the bet fair; then (for You) \( P_{B: \text{You}}(A) = \frac{O_{\text{You}}}{(1+O_{\text{You}})} \).

Other approaches not covered here include logical (e.g., Jeffreys, 1961) and fiducial (Fisher, 1935) probability.
Strengths and Weaknesses

Each of these probability definitions has general advantages and disadvantages:

- **Classical**
  - **Plus:** Simple, when applicable (e.g., idealized coin-tossing, drawing colored balls from urns, etc.).
  - **Minus:** The only way to define “equipossible” without a circular appeal to probability is through the principle of insufficient reason—You judge EOs equipossible if You have no grounds (empirical, logical, or symmetrical) for favoring one over another—but this leads to paradoxes (e.g., assertion of equal uncertainty is not invariant to the choice of scale on which it’s asserted).

- **Frequentist**
  - **Plus:** Mathematics relatively tractable.
  - **Minus:** Only applies to inherently repeatable events, e.g., (as of October 2004) $P_F$(George W. Bush will be re-elected in 2004) was (strictly speaking) undefined.

- **Bayesian**
  - **Plus:** All forms of uncertainty are in principle quantifiable with this approach.
  - **Minus:** There’s no guarantee that the answer You get by querying Yourself about betting odds will retrospectively be seen by You or others as “good” (but how should the quality of an uncertainty assessment itself be assessed?).
Application to HIV Screening

\[ P(A) = P(\text{this patient is HIV-positive}) =? \]

Data are available from medical journals on prevalence of HIV-positivity in various subsets of \( \mathcal{P} = \{\text{all humans}\} \) (e.g., it’s higher in gay people, and lower in older people).

All three probabilistic approaches require You to use Your judgment to identify the recognizable subpopulation \( \mathcal{P}_{\text{this patient}} \) (Fisher, 1956; Draper et al., 1993): this is

\[ \text{the smallest subset to which this patient belongs for which the HIV prevalence differs from that in the rest of } \mathcal{P} \text{ by an amount You judge as large enough to matter in a practical sense.} \]

Within \( \mathcal{P}_{\text{this patient}} \) You regard HIV prevalence as close enough to constant that the differences aren’t worth bothering over, but the differences between HIV prevalence in \( \mathcal{P}_{\text{this patient}} \) and its complement matter to You.

Here \( \mathcal{P}_{\text{this patient}} \) might consist of everybody who matches this patient on (e.g.) gender, age (category, e.g., 25–29), and sexual orientation.

\[ \text{NB} \] This is a modeling choice based on judgment; different reasonable people might make different choices; thus probability modeling in the real world is inherently subjective (see page 11 below for more details).

As a classicist You would then (a) use this definition to establish equipossibility within \( \mathcal{P}_{\text{this patient}} \), (b) count \( n_A = \) (number of HIV-positive people in \( \mathcal{P}_{\text{this patient}} \)) and \( n = \) (total number of people in \( \mathcal{P}_{\text{this patient}} \)), and (c) compute \( P_C(A) = \frac{n_A}{n} \).
As a **frequentist** You would (a) equate $P(A)$ to $P(\text{a person chosen at random with replacement (i.e., independent identically distributed (IID) sampling) from } \mathcal{P}_{\text{this patient is HIV-positive}})$, (b) imagine **repeating** this random sampling indefinitely, and (c) conclude that the limiting value of the relative frequency of HIV-positivity in these repetitions would also be $P_F(A) = \frac{n_A}{n}$.

**NB** Strictly speaking we’re not allowed in the frequentist approach to talk about $P(\text{this patient is HIV-positive})$—either he is or he isn’t.

In the frequentist paradigm we can only talk about the **process** of sampling people like him from $\mathcal{P}_{\text{this patient}}$.

As a **Bayesian**, with the information given here You would regard this patient as **exchangeable** (de Finetti, e.g., 1964, 1974/5) with all other patients in $\mathcal{P}_{\text{this patient}}$—meaning informally that You judge Yourself equally uncertain about HIV-positivity for all the patients in this set—and this judgment, together with the axioms of **coherence**, would also yield $P_B:You(A) = \frac{n_A}{n}$ (although I’ve not yet said why this is so).

**Exchangeability** and **coherence** will be defined and explored in more detail in what follows.

Note that with the same information base the three approaches in this case have led to the same answer, although the **meaning** of that answer depends on the approach, e.g., frequentist probability describes the **process** of observing a repeatable event whereas Bayesian probability is an attempt to **quantify Your uncertainty about** something, repeatable or not.
Subjectivity and Objectivity

The classical and frequentist approaches have sometimes been called objective, whereas the Bayesian approach is clearly subjective, and—since objectivity sounds like a good goal in science—this has sometimes been used as a claim that the classical and frequentist approaches are superior.

I’d argue, however, that in interesting applied problems of realistic complexity, the judgment of equivalence or similarity (equipossibility, IID, exchangeability) that’s central to all three theories makes them all subjective in practice.

Imagine, for example, that you were given data on HIV prevalence in a large group of people, along with many variables (possible predictors) that might or might not be relevant to identifying the recognizable subpopulations.

You and other reasonable people working independently might well differ in your judgments on which of these predictors are relevant (and how they should be used in making the prediction), and the result could easily be noticeable variation in the estimates of $P$(HIV positive) obtained by you and the other analysts, even if you and the other people all attempt to use “objective” methods to arrive at these judgments (there are many such methods, and they don’t always lead to the same conclusions).

Thus the assessment of complicated probabilities is inherently subjective—there are “judgment calls” built into probabilistic and statistical analysis.

With this in mind attention in all three approaches should evidently shift away from trying to achieve “objectivity” toward two things: (1) the explicit statement of the assumptions and judgments made on the way to your probability assessments, so that other people may consider their plausibility, and (2) sensitivity analyses exploring the mapping from assumptions to conclusions.

(To a Bayesian saying that $P_B(A)$ is objective just means that lots of people more or less agree on its value.)
1.2 Sequential Learning; Bayes’ Theorem

Let’s say that, with this patient’s values of relevant demographic variables, the prevalence of HIV estimated from the medical literature, $P(A) = P(\text{he’s HIV-positive})$, in his recognizable subpopulation is about $\frac{1}{100} = 0.01$.

To improve this estimate by gathering data specific to this patient, You decide to take some blood and get a result from ELISA.

Suppose the test comes back positive—what is Your updated $P(A)$?

Bayesian probability has that name because of the simple updating rule attributed to Thomas Bayes (1763), who was one of the first people to define conditional probability and make calculations with it:

Bayes’ Theorem

for propositions

$$P(A|D) = \frac{P(A) P(D|A)}{P(D)}$$

(actually—Stigler, 1986; Bernardo and Smith, 1994—Bayes only stated and worked with a special case of this; the general form was first used by Laplace, 1774).

In the usual application of this $A$ is an unknown quantity (such as the truth value of some proposition) and $D$ stands for some data relevant to Your uncertainty about $A$:

$$P(\text{unknown|data}) = \frac{P(\text{unknown}) P(\text{data|unknown})}{\text{normalizing constant}}$$

posterior $= c \cdot \text{prior} \cdot \text{likelihood}$
Bayes’ Theorem (continued)

The terms **prior** and **posterior** emphasize the sequential nature of the learning process: \( P(\text{unknown}) \) was Your uncertainty assessment before the data arrived; this is updated multiplicatively on the probability scale by the **likelihood** \( P(\text{data}|\text{unknown}) \), and renormalized so that total probability remains 1.

Writing the Theorem both for \( A \) and (not \( A \)) and combining gives a (perhaps even more) **useful** version:

**Bayes’ Theorem in odds form**

\[
\frac{P(A|\text{data})}{P(\text{not } A|\text{data})} = \frac{P(A)}{P(\text{not } A)} \cdot \frac{P(\text{data}|A)}{P(\text{data}|\text{not } A)}
\]

**posterior odds** = **prior odds** \( \cdot \) **Bayes factor**

Other names for the Bayes factor are the **data odds** and the **likelihood ratio**, since this factor measures the relative plausibility of the data given \( A \) and (not \( A \)).

Applying this to the HIV example requires additional information about **ELISA** obtained by screening the blood of people with known HIV status:

- **sensitivity** = \( P(\text{ELISA positive}|\text{HIV positive}) \) and
- **specificity** = \( P(\text{ELISA negative}|\text{HIV negative}) \)

In practice **ELISA**’s operating characteristics are (or at least seem) **rather good**—sensitivity about 0.95, specificity about 0.98—so you might well expect that

\[ P(\text{this patient HIV positive}|\text{ELISA positive}) \]

would be **close to 1**.
Inference and Decision-Making

Here the updating produces a surprising result (if you’ve never seen this sort of thing before): the Bayes factor comes out

\[ B = \frac{\text{sensitivity}}{1 - \text{specificity}} = \frac{0.95}{0.02} = 47.5, \]

which sounds like strong evidence that this patient is HIV positive, but the prior odds are quite a bit stronger the other way \( \frac{P(A)}{1-P(A)} = 99 \text{ to } 1 \) against HIV, leading to posterior odds of \( \frac{99}{47.5} \approx 2.08 \) against HIV, i.e.,

\[ P(\text{HIV positive|data}) = \frac{1}{1+\text{odds}} = \frac{95}{293} \approx 0.32 \text{ (!).} \]

The reason for this is that ELISA was designed to have a vastly better false negative rate—\( P(\text{HIV positive|ELISA negative}) = \frac{5}{9707} \approx 0.00052 \approx 1 \text{ in } 1941 \)—than false positive rate—\( P(\text{HIV negative|ELISA positive}) = \frac{198}{293} \approx 0.68 \approx 2 \text{ in } 3. \)

This in turn is because ELISA’s developers judged that it’s far worse to tell somebody who’s HIV positive that they’re not than the other way around (reasonable for using ELISA for, e.g., blood bank screening).

This false positive rate would make widespread screening for HIV based only on ELISA a truly bad idea.

Formalizing the consequences of the two types of error in diagnostic screening would require quantifying misclassification costs, which shifts the focus from (scientific) inference (the acquisition of knowledge for its own sake: Is this patient really HIV-positive?) to decision-making (putting that knowledge to work to answer a public policy or business question, e.g.: What use of ELISA and Western Blot would yield the optimal screening strategy?).
1.3 Bayesian Decision Theory

Axiomatic approaches to rational decision-making date back to Ramsay (1926), with von Neumann and Morgenstern (1944) and Savage (1954) also making major contributions.

The ingredients of a general decision problem (e.g., Bernardo and Smith, 1994) include

- A set \( \{a_i, i \in I\} \) of available actions, one of which You will choose;

- For each action \( a_i \), a set \( \{E_j, j \in J\} \) of uncertain outcomes describing what will happen if You choose action \( a_i \);

- A set \( \{c_j, j \in J\} \) of consequences corresponding to the outcomes \( \{E_j, j \in J\} \); and

- A preference relation \( \leq \), expressing Your preferences between pairs of available actions (\( a_1 \leq a_2 \) means "\( a_1 \) is not preferred by You to \( a_2 \)").

Define \( a_1 \sim a_2 \) ("\( a_1 \) and \( a_2 \) are equivalent" to You) iff \( a_1 \leq a_2 \) and \( a_2 \leq a_1 \).

This preference relation induces a qualitative ordering of the uncertain outcomes (\( E \leq F \) means "\( E \) is not more likely than \( F \)"), because if You compare two dichotomized possible actions, involving the same consequences and differing only in their uncertain outcomes, the fact that You prefer one action to another means that You must judge it more likely that if You take that action the preferred consequence will result.
Coherence

Within this framework You have to make further assumptions—the coherence axioms—to ensure that Your actions are internally consistent.

Informally (see Bernardo and Smith, 1994, for the formalism) these are:

- An axiom insisting that You be willing to express preferences between simple dichotomized possible actions ({a, not a});

- A transitivity axiom in which (for all actions a, a₁, a₂, a₃) a ≤ a, and if a₁ ≤ a₂ and a₂ ≤ a₃ then a₁ ≤ a₃; and

- An axiom based on the sure-thing principle: if, in two situations, no matter how the first comes out the corresponding outcome in the second is preferable, then You should prefer the second situation overall.

This puts ≤ on a sound footing for qualitative uncertainty assessment, but does not yet imply how to quantify—it's like being able to say that one thing weighs less than another but not to say by how much.

To go further requires a fourth assumption, analogous to the existence of a set of reference standards (e.g., an official kg weight, half-kg, etc.) and the ability to make arbitrarily precise comparisons with these standards:

- An axiom guaranteeing that for each outcome E there exists a standard outcome S (e.g., “idealized coin lands heads”) such that E ∼ S.

This framework implies the existence and uniqueness of a (personal) probability \(P_{B:YoU}\) (abbreviated \(P\)), mapping from outcomes \(E\) to [0,1] and corresponding to the judgments in Your definition of ≤, and a utility function \(U_{YoU}\) (abbreviated \(U\); large values preferred, say), mapping from consequences \(c\) to \(R\) and quantifying Your preferences.
“Dutch Book”

This has all been rather abstract.

Three concrete results arising from this framework may make its implications clearer:

- Bayes' original definition of personal probability is helpful in thinking about how to quantify uncertainty. Pretending that consequences are monetary (e.g., US$), to say that \( P_{B: \text{You}}(E) = p \) for some uncertain outcome \( E \) whose truth value will be known in the future is to say that You're indifferent between (a) receiving \( p \cdot m \) for sure (for some hypothetical (and reasonably small) amount of money \( m \)) and (b) betting with someone in such a way that you will get \( m \) if \( E \) turns out to be true and nothing if not (you can use this to estimate \( P_{B: \text{You}}(E) \)).

- Any coherent set of probability judgments must satisfy the standard axioms and theorems of a finitely additive probability measure:

- \( 0 \leq P(E) \leq 1 \) and \( P(E^c) = 1 - P(E) \);

- \( P(E_1 \text{ or } \ldots \text{ or } E_J) = \sum_{j \in J} P(E_j) \) for any finite collection \( \{E_j, j \in J\} \) of disjoint outcomes;

- \( P(E \text{ and } F) = P(E) \cdot P(F) \) for any two independent outcomes (informally, \( E \) and \( F \) are independent if Your uncertainty judgments involving one of them are unaffected by information about the other); and

- Conditional probability has a natural definition in this setup, corresponding to the updating of Your uncertainty about \( E \) in light of \( F \), and with this definition \( P(E|F) = \frac{P(E \text{ and } F)}{P(F)} \).

Otherwise (de Finetti, 1964) someone betting with You on the basis of Your probability judgments can make “Dutch book” against you, i.e., get You to agree to a series of bets that are guaranteed to lose You money.

Thus coherent Bayesian probability obeys the same laws as with the classical and frequentist approaches (apart from a technical issue about finite versus countable additivity).
Maximization of Expected Utility

- Nothing so far has said clearly what choice to make in a decision problem if you wish to avoid incoherence.

If the outcomes were certain you’d evidently choose the action that maximizes your utility function, but since they’re not the best action must involve weighing both your probabilities for the uncertain outcomes and the utilities you place on their consequences.

It’s a direct implication of the framework here that the form this weighing should take is simple and clear:

**Maximization of Expected Utility (MEU)**

Given your probability and utility judgments, your decision-making is coherent iff for each action $a_i$, with associated uncertain outcomes $\{E_j, j \in J\}$ and consequences $\{c_j, j \in J\}$, you compute the expected utility $EU_i = \sum_{j \in J} U(c_j) P(E_j)$ and choose the action that maximizes $\{EU_i, i \in I\}$.

**Example: HIV screening.** As a simplified version of this problem consider choosing between two actions:

- $a_1$: Obtain ELISA results at a cost of $c_1 = $20; if positive conclude this patient is HIV+, if negative conclude HIV–.

- $a_2$: Same as $a_1$ except if ELISA comes out positive, obtain Western Blot (WB) results at an additional cost of $c_2 = $100; if WB is positive conclude HIV+, if negative conclude HIV–.
HIV Screening

With action $a_1$ the probabilities, uncertain outcomes, and utilities are as follows:

<table>
<thead>
<tr>
<th>Probability</th>
<th>True HIV Status</th>
<th>ELISA Status</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>.0095</td>
<td>+</td>
<td>+</td>
<td>$-c_1$</td>
</tr>
<tr>
<td>.0005</td>
<td>+</td>
<td>-</td>
<td>$-c_1 - L_I$</td>
</tr>
<tr>
<td>.0198</td>
<td>-</td>
<td>+</td>
<td>$-c_1 - L_{II}$</td>
</tr>
<tr>
<td>.9702</td>
<td>-</td>
<td>-</td>
<td>$-c_1$</td>
</tr>
</tbody>
</table>

Here $L_I$ and $L_{II}$ are the false negative (false positive) monetary losses suffered by this patient if he really is HIV+ (HIV–) but ELISA says he is HIV– (HIV+).

The expected utility with action $a_1$ is thus

$EU_1 = .0095(-c_1) + .0005(-c_1 - L_I) + \ldots + .9702(-c_1)$

$= -(c_1 + .0005L_I + .0198L_{II})$ .

The corresponding table for action $a_2$ is:

<table>
<thead>
<tr>
<th>Prob.</th>
<th>True HIV Status</th>
<th>ELISA Status</th>
<th>WB Status</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>.00945</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>$-c_1 - c_2$</td>
</tr>
<tr>
<td>.00005</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>$-c_1 - c_2 - L_I$</td>
</tr>
<tr>
<td>.00004</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>$-c_1 - L_I$</td>
</tr>
<tr>
<td>.00046</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>$-c_1 - L_I$</td>
</tr>
<tr>
<td>.0001</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>$-c_1 - c_2 - L_{II}$</td>
</tr>
<tr>
<td>.0197</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>$-c_1 - c_2$</td>
</tr>
<tr>
<td>.00095</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>$-c_1$</td>
</tr>
<tr>
<td>.96925</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$-c_1$</td>
</tr>
</tbody>
</table>
HIV Screening (continued)

These probabilities arise from WB’s design (the goal was to have about the same false negative rate as ELISA and a much lower false positive rate (about 0.1), leading to a slightly worse sensitivity (0.949) but much improved specificity (0.999)).

The expected utility with action $a_2$ comes out

$$EU_2 = 0.00945(-c_1 - c_2) + \ldots + 0.9604(-c_1)$$

$$= -(c_1 + 0.0293c_2 + 0.00055L_I + 0.0001L_{II}) .$$

By MEU You should prefer $a_2$ to $a_1$ iff $EU_2 > EU_1$, i.e., iff

$$0.0197L_{II} - 0.00005L_I - 0.0293c_2 > 0 .$$

Thus $a_2$ becomes more desirable as the loss suffered with a false positive (negative) increases (decreases), and less desirable as WB’s cost increases, all of which makes good sense.

It’s interesting to note that with a modest value for $L_{II}$ (e.g., $1,000), the monetary advantage from taking action $a_2$ is quite small even with a realistically huge value for $L_I$ (e.g., $100,000, which leads to an edge for $a_2$ of only about $12).

This is due to the extremely low false negative rate for both tests—$L_I$ would have to be over $335,000 for $a_1$ to dominate!
1.4 References


