A PROBABILISTIC ANALYSIS OF THE "UNFAIR" EURO COIN

Jonathan E. Hamaker

Department of Electrical and Computer Engineering
Mississippi State University

hamaker@isip.msstate.edu

This analysis was originally presented in

D. J. C. MacKay, "Belgian euro coins: 140 heads in 250 tosses - suspicious?," available from *http://www.inference.phy.cam.ac.uk/mackay/abstracts/euro.html*, University of Cambridge, Department of Physics, January 2002.

Modifications have been made to provide detail and interpretation for a novice level student with some experience with probability theory.

A STATISTICAL "FACT"?

From *The Guardian,* Friday, January 4, 2002:

> "When spun on edge 250 times, a Belgian one-euro coin came up heads 140 times and tails 110. 'It looks very suspicious to me,' said Barry Blight, a statistics lecturer at the London School of Economics. 'If the coin were unbiased the chance of getting a result as extreme as that would be less than 7%.'"

Questions to ask:

- What is a "fair" coin?
- Where did Blight get 7% from?
- What hypothesis did he make from his analysis?
- Can we prove/disprove his hypothesis?

WHAT IS A FAIR COIN?

- A fair coin is one where P(Heads) = P(Tails) = 1/2.
- Probability of seeing M Heads when tossing a coin N times, when p is the probability of seeing a head (1/2 for a fair coin) is given by the binomial distribution:

$$P(M|N, p) = \binom{N}{M} p^M (1-p)^{N-M}$$

- $\binom{N}{M}$ is the number of ways that you could split N data samples up into two sets, one of length M and one of length N-M.
- p^M is the probability that a grouping of M elements will have all been heads
- $(1-p)^{N-M}$ is the probability that a grouping of N-M elements will have all been 'not heads' or tails.

WHERE DID BLIGHT GET 7% FROM?

- What is the probability of seeing 140 heads in 250 tosses with a fair coin?

$$P\left(M = 140 \middle| N = 250, p = \frac{1}{2}\right) = \binom{250}{140}\left(\frac{1}{2}\right)^{140}\left(\frac{1}{2}\right)^{250 - 140} = 0.0084$$

- What is the probability of seeing this large of a discrepancy (or worse) for an unbiased coin?

$$P = P\left(M \geq 140 \middle| N = 250, p = \frac{1}{2}\right) + P\left(M \leq 110 \middle| N = 250, p = \frac{1}{2}\right)$$

$$= 2P\left(M \geq 140 \middle| N = 250, p = \frac{1}{2}\right)$$

$$= \sum_{k = 140}^{250} P\left(M = k \middle| 250, \frac{1}{2}\right) = 0.064$$

WHAT HYPOTHESIS CAN ONE DRAW?

- "Looks very suspicious": i.e. it seems that the coin may be biased.
- How can we test this hypothesis that the coin is biased?

Define H0: the model (hypothesis) that the coin is fair
Define H1: the model (hypothesis) that the coin is biased

Infer using the probability ratio: $\frac{P(H1|D)}{P(H0|D)} > 1?$

Rewrite as $\frac{P(H1|D)}{P(H0|D)} = \frac{P(D|H1)P(H1)}{P(D|H0)P(H0)}$

If we have no prior preference for H1 or H0 (i.e. P(H1) = P(H0)) then we can use the "evidence", $P(D|H_*)$ to rank the alternative hypotheses. If our suspicion is true then we would expect the evidence for H1 to overwhelm the evidence for H0.

COMPUTING THE EVIDENCE

- Marginalize the evidence over the adjustable parameter, p

$$P(D|H_*) = \int_0^1 P(D|p, H_*)P(p|H_*)dp.$$

- For H1, $P(D|H1) = \int_0^1 \binom{250}{140} p^{140}(1-p)^{110} P(p|H1)dp$
- How should we set the prior probability on the coin bias, p?
- A first analysis would be to use a uniform prior on p - i.e. we have no knowledge as to how much the coin is biased if at all so we assume all biases equally likely. The result will, thus, be constant for all M!

$$P(p|H1) = 1; \quad \int_0^1 P(p|H1) = \int_0^1 1 = 1$$

$$P(D|H1) = \int_0^1 \binom{250}{140} p^{140}(1-p)^{110} dp = 0.00398$$

COMPUTING THE EVIDENCE

- Marginalize the evidence over the adjustable parameter, p

$$P(D|H_*) = \int_0^1 P(D|p, H_*)P(p|H_*)dp.$$

- For H0, $P(D|H0) = \int_0^1 \binom{250}{140} p^{140}(1-p)^{110} P(p|H0)dp$
- Note that the prior probability, $P(p|H0)$, is a unit impulse at $p = \frac{1}{2}$.
- Using the sifting theorem:

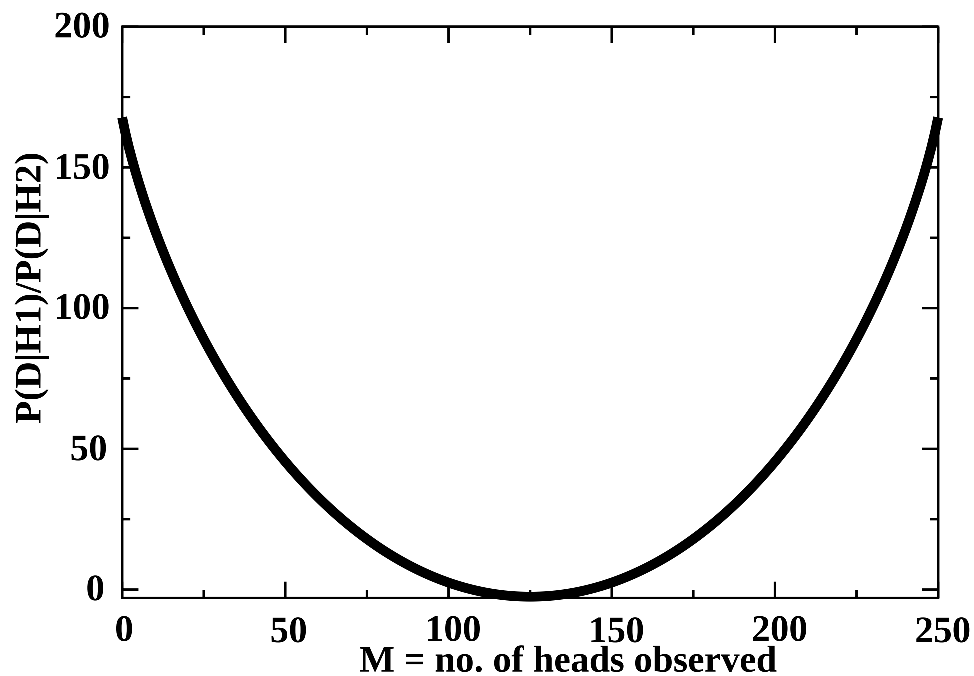
$$P(D|H0) = \binom{250}{140}\left(\frac{1}{2}\right)^{140}\left(1-\frac{1}{2}\right)^{110} = 0.00836$$

HYPOTHESIS TEST

- So which one is a more probable explanation according to the evidence?

$$\frac{P(D|H1)}{P(D|H0)} = \frac{0.00398}{0.00836} = 0.476$$

- Uh-Oh. Wasn't this supposed to be a biased coin? In fact there is weak evidence leaning toward an unbiased coin (2 to 1).



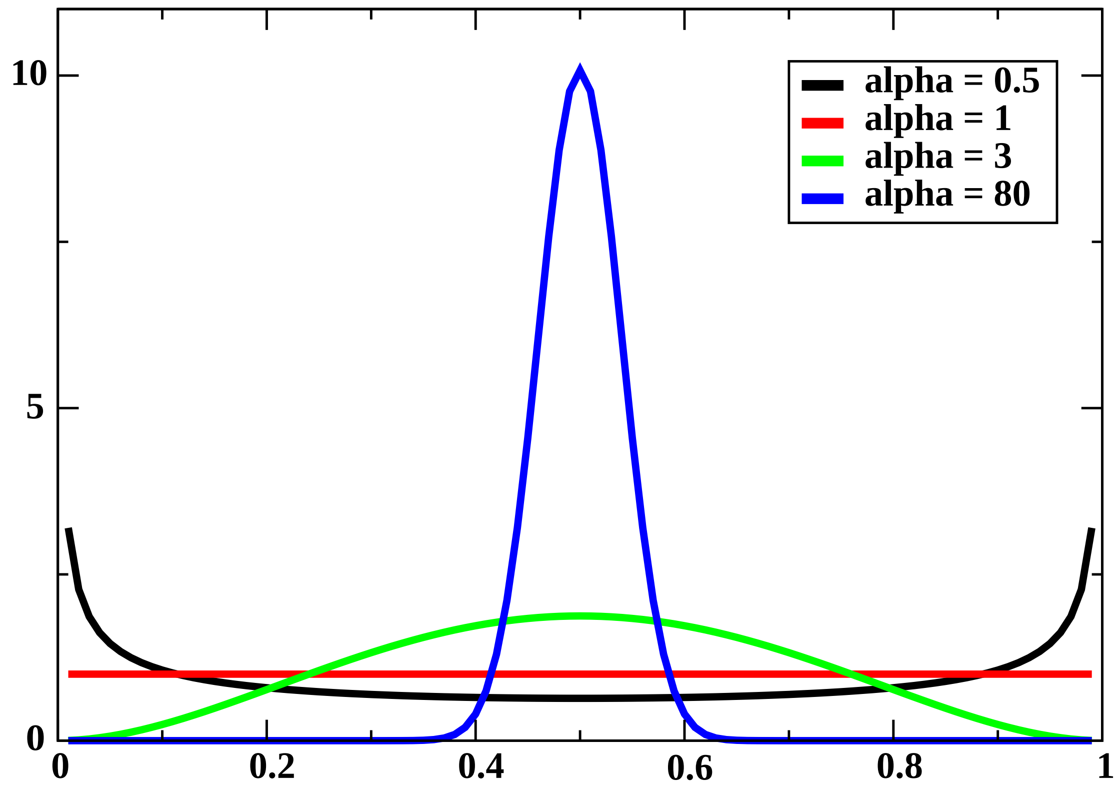
- What happened? An objection to bayesian methods is the choosing of "arbitrary" prior distributions. For H1, we chose a uniform distribution. What if we had a prior belief that the bias was not uniform across [0,1]? How could we modify our hypothesis to take this into account?

BETA DISTRIBUTION

- If we make the prior for H1 be a beta distribution:

$$P(p|H1,\alpha) = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2}p^{\alpha-1}(1-p)^{\alpha-1}, \quad \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1}e^{-x}dx$$

- α gives us an adjustable parameter that we can use to set the region of our prior belief.



COMPUTING THE EVIDENCE

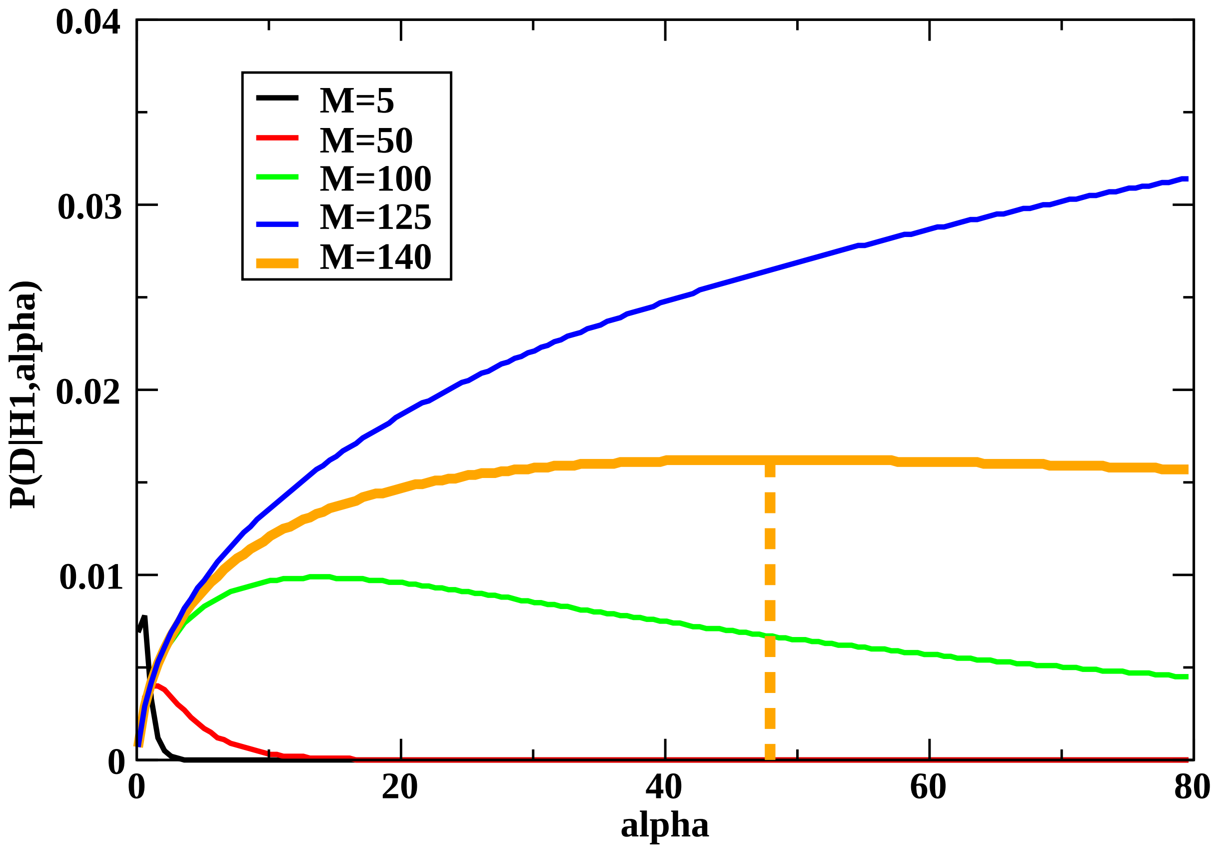
- For H1, $P(D|H1) = \int_0^1 \binom{250}{140} p^{140}(1-p)^{110} P(p|H1, \alpha)\, dp$
- Using the Beta distribution gives

$$P(D|H1) = \int_0^1 \binom{250}{140} p^{140}(1-p)^{110} \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} p^{\alpha-1}(1-p)^{\alpha-1}\, dp$$

- Rearranging yields:

$$P(D|H1) = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} \binom{250}{140} \int_0^1 p^{140+\alpha-1}(1-p)^{110+\alpha-1}\, dp$$

HYPOTHESIS TEST



- Even at the value of α most amenable to H1, $\alpha = 47.9$, the evidence ratio is only a factor of 2 - again no strong conclusion can be drawn as to which hypothesis is better.

$$\frac{P(D|H1,\alpha)}{P(D|H0)} = \frac{0.01622}{0.00836} = 1.94$$

A CHEATING PRIOR

- What if we set the prior so that it exactly matches the data? In other words set $P(p|H1)$ as an impulse function centered at $p = \dfrac{140}{250}$.

- Again, using the sifting theorem, we get

$$P(D|H1) = \binom{250}{140}\left(\frac{140}{250}\right)^{140}\left(1 - \frac{140}{250}\right)^{110} = 0.05078$$

- This gives the highest evidence ratio possible for this data:

$$\frac{P(D|H1)}{P(D|H0)} = \frac{0.05078}{0.00836} = 6.07$$