University of Wisconsin-Madison

Bayesian Analysis of Health Status and Quality of Life Data

(Preliminary) Example from the Beaver Dam Studies

Dennis Fryback, Janel Hanmer
Department of Population Health Sciences
University of Wisconsin-Madison

Karen Cruickshanks, Ronald Klein, Barbara Klein
Dept of Opthalmology and Visual Sciences
University of Wisconsin-Madison

June 29, 2004 ISOQOL Symposium
"Stating the Art: Advancing Outcomes Research Methodology and Clinical Applications

University of Wisconsin-Madison

OR…

"Quest for QA^LE"

University of Wisconsin-Madison

- Much of research is about 2 things:
  - Asking how likely is some hypothesis, H, in light of data
    - usually H expressed in terms of value of a parameter or statistic
      - H : treatment effect A = treatment effect B
      - H : the true correlation, ρ, is >0.6
      - H : (treatment A cost – treatment B cost) > $0
  - Estimating the value of a parameter
    - what is best estimate of p, the 30-day survival with tPA?
    - what is the average yearly cost of treatment A?
    - **How many QALYs will be accrued in the next 10 years by a 56.7 y.o male who self-rates his health as "Good" on the EVGFP scale?**
- The debate:
  - Bayesians: "Frequentists can't do the first one"
  - Frequentists: "Bayesians too subjective about the second one"

University of Wisconsin-Madison

Goal of this talk

- Can't solve the debate, so…
- Illustrate what I'm most interested in:

  Estimation of parameters given data

- May work in some reflections on the debate.

University of Wisconsin-Madison

Some history…

Rev. Thomas Bayes writes a paper noting a trivial consequence of the definitions of joint and conditional probabilities

$$\Pr(A \mid B)\cdot\Pr(B)=\Pr(A,B)=\Pr(B \mid A)\cdot\Pr(A)$$

$$\Pr(A \mid B)=\frac{\Pr(A,B)}{\Pr(B)}$$

$$\Pr(A \mid B)=\frac{\Pr(B \mid A)\cdot\Pr(A)}{\Pr(B)}$$

Bayes' Theorem

1763 1800 1900 2000

University of Wisconsin-Madison

Some history…

Hypothesis-data form:

$$\Pr(H \mid Data)=\frac{\Pr(Data \mid H)\cdot\Pr(H)}{\Pr(Data)}$$

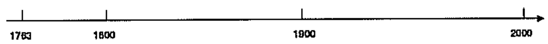
1763 1800 1900 2000

University of Wisconsin-Madison

Some history...

Estimation of quantity □ from data

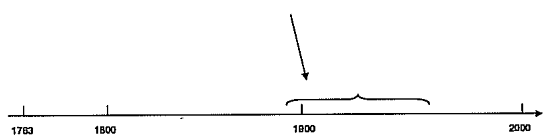
$$\Pr(\theta \mid Data) = \frac{\Pr(Data \mid \theta) \cdot \Pr(\theta)}{\Pr(Data)}$$



University of Wisconsin-Madison

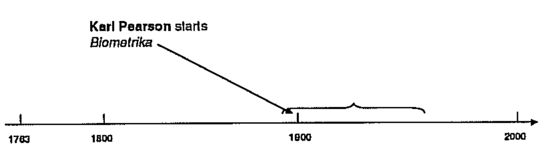
Some history...

Much debate about analysis using Bayes' theorem



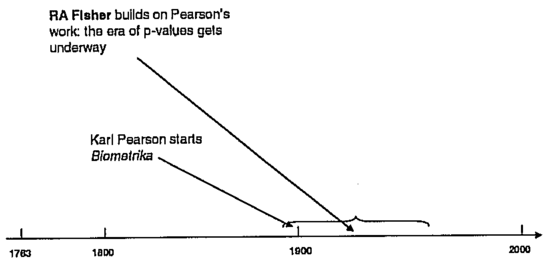
University of Wisconsin-Madison

Some history...



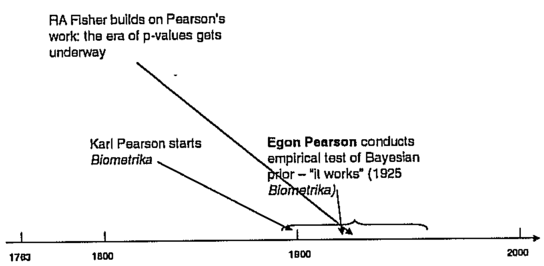
University of Wisconsin-Madison

Some history...



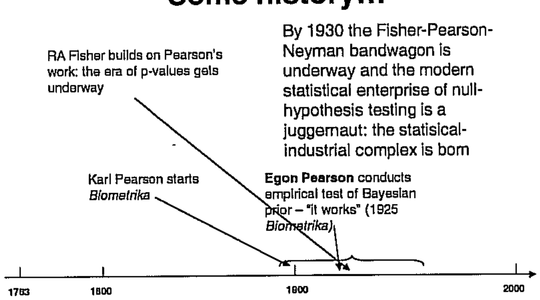
University of Wisconsin-Madison

Some history...



University of Wisconsin-Madison

Some history...

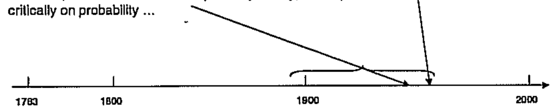


University of Wisconsin-Madison

Some history...

LJ Savage: *The Foundations of Statistics*, 1954. Simultaneous axiomatization of probability and utility from subjective decision-making viewpoint

Von Neumann & Morgenstern: *Theory of Games and Economic Behavior*, 1947. Axioms to underpin utility theory, but depends critically on probability ...



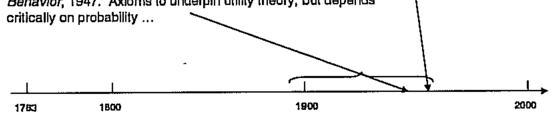
University of Wisconsin-Madison

Some history...

Origins of preference-based health outcome measures – health-related quality of life metrics for medical decision making and for cost-utility analysis.

LJ Savage: *The Foundations of Statistics*, 1954. Simultaneous axiomatization of probability and utility from subjective decision-making viewpoint

Von Neumann & Morgenstern: *Theory of Games and Economic Behavior*, 1947. Axioms to underpin utility theory, but depends critically on probability ...

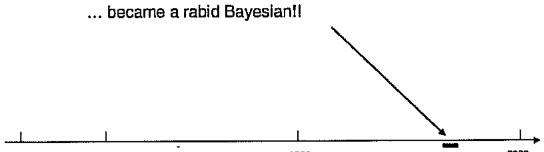


University of Wisconsin-Madison

A more personal history...

1969-1974 – I learned Bayesian statistics, psychology and math of decision theory.

... became a rabid Bayesian!!



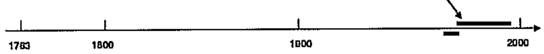
University of Wisconsin-Madison

A more personal history...

By 1975 three things were apparent:

- Not possible to publish without p-values
- Not possible to calculate posterior probabilities for any interesting statistical models anyway ... the integrals could not be solved.
- I was not good enough statistician to contribute to the statistical theory

So I used frequentist statistics and got on with applied work.



University of Wisconsin-Madison

A more personal history...

1988: I needed to know the QALE for an average community dwelling person of arbitrary age, and starting in a given health state.

Moreover, I needed to know a full probability distribution for this quantity so I could do a cost-utility analysis properly.

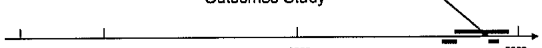
Such a quantity did not exist and there were not even data to estimate it from.

University of Wisconsin-Madison

A more personal history...

1989: I talked a colleague into letting me use his longitudinal cohort in Beaver Dam, Wisconsin, to get community quality of life estimates.

1990-92: Beaver Dam Health Outcomes Study



University of Wisconsin-Madison

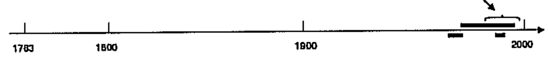
Meanwhile on the Bayesian front...

1990's: very active time of Bayesian progress

MCMC applied to Bayesian modeling (early 1990s):
At last! Can compute real posteriors from complex data models!!! .. that is if you are a programming wizard.

Late 1990s: Spiegelhalter and colleagues at MRC introduce BUGS software. Now, maybe even *I* can do MCMC.

1997: I'm reborn a Bayesian!



University of Wisconsin-Madison

Today

- Join together my 16-year pursuit for an estimate of QALE and my renewed interest in Bayesian statistics.
- You are first to see posterior probability distribution for 10-year QALE of older adults of specified age, sex, self-rated health based on Beaver Dam data.
- These estimates did not exist before June 27, 2004 (approximately 1:30 am)
- Hope to whet your appetite to learn more about Bayesian methods.

University of Wisconsin-Madison

Strategy

- Longitudinal data 1991-2002 in Beaver Dam cohort
  - 4 observations of SF-36 (transformed to SF-6D pseudo utility)
  - Continuous monitoring of mortality
- From this cohort, simultaneously estimate:
  - Parametric survival function conditioned on sex, starting age, sex, and initial self-rated health
  - Autoregressive estimates of SF-6D utilities (segmented linear functions) across time, conditioned on sex, starting age, and initial self-rated health.
- Compute $\Pr[QALE_{10}(\text{sex, initial age, initial health})]$

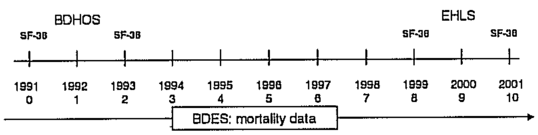
$$= \iint (S(t \mid sex, age, health) QALY(t \mid sex, age, health)) \cdot p(\vec{\theta} \mid Data) dt d\vec{\theta}$$

where $\vec{\theta}$ [illegible]

University of Wisconsin-Madison

10-year Longitudinal Data

- 76% of non-institutionalized adults over the age of 45 in Beaver Dam, WI, agreed to participate in the Beaver-Dam Eye Studies beginning in 1987 (n=4926). Mortality recorded by BDES through 2002.
- A random subset (n=1430) participated in the Beaver Dam Health Outcomes Study. BDHOS comprised two waves in 1991 and 1993 using SF-36 among other instruments.
- All BDES subjects invited to participate in the Epidemiology of Hearing Loss Study in 1992-3. EHLS completed two further waves in 1999 and 2001 also using SF-36.



University of Wisconsin-Madison

Demographics

	Year 0	Year 2	Year 8	Year 10
# active	1430	1142	938	844
# dead	0	80	261	332
# alive, but missing	0	208	231	254
% male	41.3	39.9	39.0	37.7
Min age	46	48	54	56
Max age	90	92	97	100
Median age	64	66	69	71

University of Wisconsin-Madison

A note about the "missing"

- Beaver Dam study was study of non-institutionalized persons.
- "Missing" are those who were:
  - Unable to be located-
  - Unable to be scheduled for interview
  - Institutionalized (nursing home, jail, long-term hospital)
  - Unable to respond (per caregiver) due to dementia
  - Refused
    - "Too sick"
    - "Too busy"
    - "Have done enough"
    - "Don't like the research"
    - ...
- For present purposes, these were considered missing at random.
- Note that some "missing" in observations at early year did in fact participate at later year observation

University of Wisconsin-Madison

Exploring Categorical Outcomes using EVGFP question

	Year 0	Year 2	Year 8	Year 10
Excellent	207	139	106	67
Very Good	538	403	327	266
Good	527	458	379	353
Fair	137	129	88	102
Poor	21	13	17	28
Missing	0	208	252	282
Dead	0	80	261	332

University of Wisconsin-Madison

Describing Change Over Time

We can use transition matrices to describe status change between observations. Below, we see how those who said they were "Very Good" at year 0 answered the same question 2 years later.

Year 2

Year 0	E	V	G	F	P	Dead	Missing
Excellent							
Very Good	.09	.46	.26	.03	.00	.03	.13
Good							
Fair							
Poor							

Some individuals report better health

Year 0 to Year 2

	E	V	G	F	P	Dead	Missing
Excellent	.41	.34	.09	.01	.00	.04	.11
Very Good	.09	.46	.26	.03	.00	.03	.13
Good	.02	.15	.50	.10	.00	.05	.17
Fair	.00	.02	.24	.37	.05	.16	.16
Poor	.00	.00	.05	.33	.24	.19	.19

Year 2 to Year 8

	E	V	G	F	P	Dead	Missing
Excellent	.42	.35	.04	.03	.00	.06	.09
Very Good	.09	.41	.28	.01	.00	.08	.12
Good	.00	.19	.45	.07	.01	.13	.15
Fair	.00	.05	.18	.26	.05	.31	.16
Poor	.00	.00	.00	.23	.23	.31	.23
Dead	.00	.00	.00	.00	.00	1.0	.00
Missing	.04	.11	.16	.05	.01	.15	.47

University of Wisconsin-Madison

Estimating Transition Matrices

- Transition matrices can be estimated with multinomial logistic regression.
- Multinomial logistic regression can accommodate covariates.
- Compare transition matrices from year 2 to year 8 for those younger than 60 and those older than 70.

Year 2 to Year 8 <60 yo

	E	V	G	F	P	Dead
Excellent	.47	.47	.06	.00	.00	.00
Very Good	.13	.57	.26	.02	.00	.01
Good	.03	.27	.53	.09	.02	.06
Fair /Poor	.01	.09	.45	.19	.05	.21

Year 2 to Year 8 >70 yo

	E	V	G	F	P	Dead
Excellent	.05	.25	.39	.09	.01	.21
Very Good	.06	.29	.38	.08	.01	.17
Good	.03	.18	.37	.11	.02	.29
Fair /Poor	.01	.06	.21	.10	.02	.60

University of Wisconsin-Madison

- Multinomial matrices can be estimated with either frequentist or Bayesian techniques.
  - Lots to show, but no time.
- Used here to point out that the long term course of an individual can vary according to age and to self-rated health at year 0. Sex matters too.

University of Wisconsin-Madison

Survival Model: Weibull hazard

- Weibull hazard density: W(shape, scale)
  - Shape parameter must be >0; =1 is exponential hazard
  - We let shape vary, with gamma density function and diffuse prior.
- We fit Weibull scale using Bayesian regression:

$$\ln(scale_i) = a + a_{sex} sex_i + a_{age} age_i + a_{f/p} I_{f/p,i} + a_{good} I_{good,i} + \varepsilon_i$$

Red = covariates for individual *i*

Blue = constants to be fit

Individual error term

Zero mean; constant variance across Ss

University of Wisconsin-Madison

Prior distributions for Weibull parameters

- Shape ~ gamma(1, .001)
- a, a_{sex}, a_{age}, $a_{f/p}$, a_g ~ normal(0,1000)
- e ~ normal(mu, sigma)
  - mu = 0 (fixed)
  - 1/sigma² ~ gamma(.001, .001)

University of Wisconsin-Madison

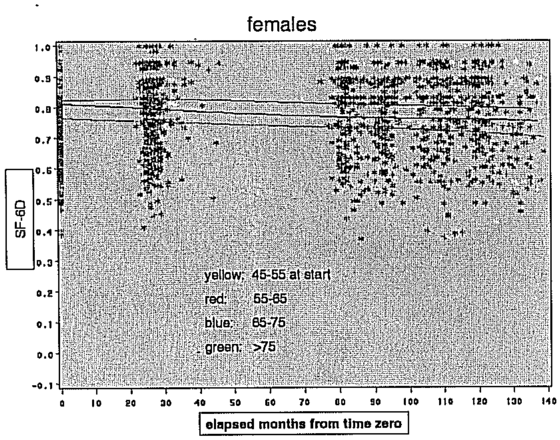
Posterior densities for parameters

			Percentiles of distribution		
parameter	mean	sd	2.5%	median	97.5%
a	-11.89	0.57	-12.98	-11.88	-10.84
a_{age}	0.007	.00049	0.0061	0.0070	0.0080
a_{sex}	-0.63	0.11	-0.83	-0.63	-0.41
$a_{f/p}$	0.83	0.15	0.53	0.83	1.11
a_g	0.26	0.12	0.023	0.26	0.50
shape	1.14	0.055	1.036	1.14	1.25

University of Wisconsin-Madison

Quality of Life data

- SF-36 → SF-6D at times 0, 2, 8, 10.
  - defined SF-6D = 1 as = ".99"
  - dead = 0
- Slowly decreasing with age and over time longitudinally; males > females
- characterized by small trends, and a lot of variance



University of Wisconsin-Madison

- Fit AR(1) to data for those alive at each time point. Define logit($SF6D_t$) = LQ_t
  - $LQ_{0,i}$ ~ normal($\mu_{0,i}$, s_0)

    $$\mu_{0,i} = b_0 + b_{sex} sex_i + b_{age} age_{0,i} + b_{fp} I_{fp} + b_g I_g$$

  - $LQ_{1,i}$ ~ normal($\mu_{1,i}$, s_1)

    $$\mu_{1,i} = b_1 + slope_{1,i} \cdot LQ_{0,i}$$

    $$slope_{1,i} = b_{1,sex} sex_i + b_{1,age} age_{0,i} + b_{1,fp,i} I_{fp,i} + b_{1,g,i} I_{g,i}$$

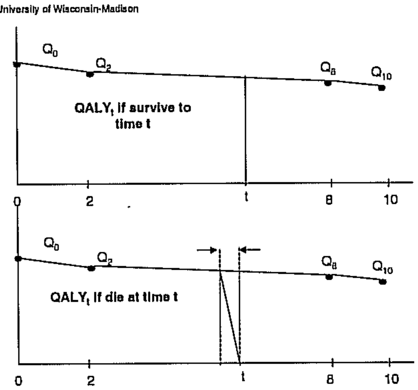
  - $LQ_{2,i}$ ~ normal($\mu_{2,i}$, s_2)
  - $LQ_{3,i}$ ~ normal($\mu_{3,i}$, s_3) } etc.

University of Wisconsin-Madison

Parameters

- In all, 27 parameters in the model
- Note that these were fit as independent parameters, but in future we could incorporate inter-correlations.
- Since both survival hazard and autoregressive process for QoL are referenced to same individual-level covariates, survival and QoL will be correlated in predictions if they are in the data.

University of Wisconsin-Madison



University of Wisconsin-Madison

Posterior predictions

- Model parameters are fitted from data so are uncertain. We have joint posterior distribution over all 27 parameters.
  - sampled 7,500 vectors of parameters from joint posterior distribution.
  - these represent 7,500 realizations of the model of the data.
- For each realization:
  - Numerically integrate

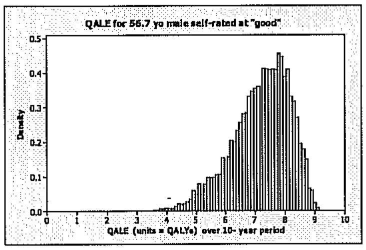
$$E_{10}(QALY) = \int_0^{10} (1 - S_i(t)) QALY_i(t)\,dt$$

  - Where $S_i(t)$ is survival function and $QALY_i(t)$ is QALYs accrued for person with covariates i to time t.
- Finally, average across realizations to integrate out uncertainty about model parameters. .... Whew!

University of Wisconsin-Madison

The first distribution over $QALE_{10}$

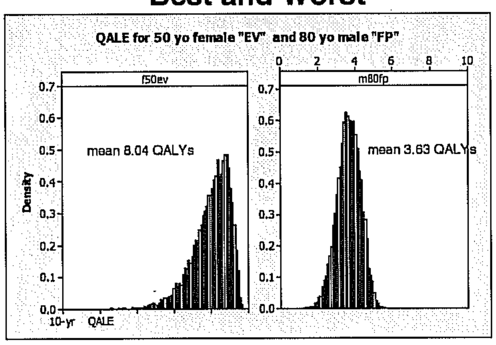
56.7 y.o. male, self-rated health "good".



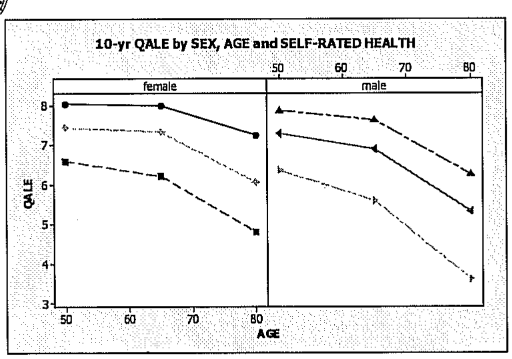
Mean: 7.19 QALYs
central 95% interval:
(4.94 , 8.68)

University of Wisconsin-Madison

Best and Worst



University of Wisconsin-Madison



Note: We have the full predictive distribution around each point

University of Wisconsin-Madison

Limitations

- I'm pretty sure these results are ok – but model only recently fit and have not been proofed by anyone else
- Arbitrary interval for "slope to death" in $QALY_t$ – model this more realistically with data on QALY in 6 mo. before death.
- Missing data: here presumed MAR (really: ignored). We should to model missingness directly
  - Particularly long-term institutionalization
- Survival has increased over the past 10 years, so these probably underestimate survival a bit.
- One small community in Wisconsin

University of Wisconsin-Madison

Future

- Model with non-independent regression coefficients
- Inform survival estimates with prior based on cohorts from Berkeley Life Tables and with geographically and temporally relevant life tables.
  (Rosenberg, Fryback, Lawrence: "Computing population-based estimates of health-adjusted life expectancy." *Med Decis Making* 1999)
- I think this last step – use of informative priors – is not possible with standard frequentist approach.

University of Wisconsin-Madison

Summary

- Bayesian analysis has come a long way
  - Can compute posteriors for very complex models
  - Still not easy, but do-able
- Bayesian analysis results in more information about quantities of interest – full predictive distributions for further use in computations such as cost-effectiveness and cost-utility analysis
  - Can use prior relevant information to inform posterior estimates
  - Can average out uncertainty about estimated parameters, to incorporate uncertainty about the analytic result into the distributions.
- "Life is good, but your mileage may vary"

  -WF Lawrence, Jr.