H.G.B. ALEXANDER RESEARCH FOUNDATION

GRADUATE SCHOOL OF BUSINESS

UNIVERSITY OF CHICAGO

Judge Conference **May, 1999**
U. of Illinois

BAYESIAN ANALYSIS OF GOLF

by

Arnold Zellner
U. of Chicago

Abstract

In this paper Bayesian analysis is used to analyze some problems that arise in playing golf in what is thought to be a scientific manner. Some issues that arise are: (1) Is a scientific analysis of golf possible? (2) What concept of probability, models and inference procedures are most useful? And (3), Can Bayesian decision theoretic methods be used to help improve George Judge's and other golfers' scores? Several canonical golf problems are formulated and analyzed using Bayesian methods. Finally, frameworks for analyzing a consumer demand for golfing services and products and professional golfers' income optimization problems are provided. In the concluding section, implications for the future will be considered.

Key words: Bayesian analysis; sports statistics; golf science; decision analysis.

Bayesian Analysis of Golf

by

Arnold Zellner*
University of Chicago

1. Introduction

Golf is a world-wide sport and business. There are amateur and professional golfers and golf courses in almost all countries of the world. Manufacturers produce many products used by golfers. Educators, often golf course pros, instruct individuals on how to improve their games, i.e., lower their golf scores and play more consistently. There are many books by famous golfers and others on how to play better golf, etc. The ASA's Section on Sports Statistics occasionally has an article on golf in its annual proceedings volumes; see, e.g. Mosteller and Youtz (1992). It is also the case that many are measuring and studying the performance characteristics of golfers' clubs, balls, gloves, shoes, other equipment, health status, psychological condition and other topics. Further, there is a need to learn from past data and experience, using the methods of science, so as to be able to explain past results and predict and improve future performance. If this is done using the methods of science, according to the Unity of Science principle put forward and discussed by Pearson (1938) and Jeffreys (1998), the study of golf qualifies to be designated a science.

In Section 2 a central club choice problem and some other canonical problems will be analyzed taking account of current and past information. It will be shown how such problems can

---

*Research financed by the National Science Foundation and by income from the H.G.B. Alexander Endowment Fumd, Graduate School of Business, University of Chicago.

be solved using Bayesian methods given an appropriate loss function using just prior information and then using prior information augmented by past sample information. Some statistical models useful in analyzing these problems will be introduced and discussed. In Section 3, the problem of modeling past golf data and predicting future outcomes, e.g. next year's average score, will be analyzed using time series regression and dichotomous random variable models. Various possible inference approaches including OLS, ML, traditional Bayes, Bayesian method of moments, generalized maximum entropy, etc. will be discussed. In Section 4, economic models of golfers' behavior will be formulated and analyzed. One deals with professional golfers who play golf for a living while the second model deals with consumers who play golf for entertainment even though some golf outcomes are painful. Last, in a concluding section, a summary of results and some prospects for the future are presented.

2. Bayesian Analyses of Some Canonical Golfing Problems

The first canonical problem that is considered is the choice between a 3 iron and a 7 iron for a tee shot on a par 3 hole of length 200 yards with a green surrounded by water. A decision to use a 3 iron from the tee in an effort to hit the green may be successful. On the other hand, there is the possibility of ending up in the water with an associated penalty. An alternative is to "play it safe" by laying up with the use of a 7 iron from the tee. After the lay up shot, it may be possible to hit the green on the second shot and possibly get a moderately good score. Generally, the decision to use a 3 or a 7 iron for the tee shot is made "heuristically" taking into account past performance, current conditions, personal whims, objectives, etc. To approach the problem more formally, consider Table 1 where the possible outcomes associated with the use of a 3 or of a 7 iron from the tee are delineated. The probabilities of specific outcomes, e.g. a score of 1 on the hole are indicated by p's and q's.

Note that in Table 1, we designate outcomes in terms of possible scores, namely 1,2, 3,..., 6 strokes. Here we are assuming that any outcome above a triple bogey, that is a 6, will be scored a 6. If desired, one could extend the range of outcomes beyond 6, but we shall not do so here.

Second, the p's and the q's are probabilities associated with possible outcomes. As in all problems, it is important to define probability in order to understand what it is we are discussing. Here, we, along with many others, shall use the degree of reasonable belief definition of probability. That is, with given background information, probability is a numerical measure of the degree of confidence in a proposition , e.g. that use of a 3 iron will result in a score of three. See Jeffreys (1998, Ch 7) for a perceptive analysis and discussion of alternative definitions of probability including the axiomatic, long run frequency and hypothetical infinite population definitions. He concludes that the "degree of reasonable belief" definition, a "subjective" concept is most satisfactory for work in science. And in connection with the current problem, it seems difficult to define a long run frequency since golfers tend to age in the long run and aging has an effect on the value of a probability. As regards the axiomatic definition, we have six possible outcomes in the present problem and to assign each a probability of 1/6 is absurd. (Would that the probability of a 1 were 1/6.)

Using the degree of reasonable belief definition, we have the problem of assigning values to the p's and q's, the "assessment problem." If we have little background information, we might use a "diffuse" or "noninformative" prior. One possibility is to assign the values to the p's and q's that maximize entropy à la Jaynes (1968). If we do we get all the p's = 1/6 and all the q's = 1/6, i.e. uniform prior probability mass functions. With use of a 7 iron, we know that we can not hit the green short of a miracle, and thus the probability of scoring a one is practically 0 As Jaynes would be quick to point out, additional information is available that should be used in deriving a maxent prior. For example, we could seek the p_i's that maximize H = $-\sum_{i=1}^{6} p_i \ln p_i$

subject to $\sum_{i=1}^{6} p_i = 1$, $a_i, < p_i < b_i$, i =1, 2, . . ., 6, $\sum_{i=1}^{6} p_i i = m$, a given value, $\sum^{6} (i - m)^2 p_i = v$, a given value, and possibly other side conditions. Using numerical procedures, the optimal p_i's can be computed. Note, in contrast to many other procedures for producing prior probabilities, Jaynes' procedure does not require knowing the form of the likelihood function. While the maxent procedure can be employed, here we shall proceed to assign prior values to the p_i's and q_i's heuristically taking account of available information regarding current weather conditions, mental state, physical state, recent past performance, etc. Given that we have done so, it is possible to make probability statements about possible outcomes, e.g., given that a 3 iron is selected, the probability of scoring above 3 is equal to the sum of the probabilities of scoring a 4, a 5 and a 6. Further, the predicted mean score, denoted by ES, given that we use a 3 iron is:

$$\mathrm{ES}|3 = \sum_{i=1}^{6} i\, p_i \qquad \textbf{(2.1)}$$

Similarly ,

$$\mathrm{ES}\ |7 = \sum_{i=1}^{6} i\, q_i \qquad \textbf{(2.2)}$$

If our problem were just a prediction problem, (2.1) and (2.2) provide optimal point predictions relative to a squared error prediction loss function. Or if we employ a 0-1 loss function, the optimal point prediction would be the outcomes associated with the largest values of the p's and of the q's. However, our problem is a club selection decision problem, not a pure prediction problem. To solve it, we consider the problem of choosing an appropriate utility function and face problems similar to those faced by Tinbergen (1954) who was concerned about the appropriate social welfare function to use in making economic policy. See also Zellner (1973) for consideration of the effects of errors in formulating utility or loss functions. Here we might assume that a player's utility increases the lower the score. If utility is linear in score, S, i.e., U(S) = a -bS, with a,b >0, expected utility is higher for choice of the 3 iron if mean score in (2.1) is lower than that in (2.2). The rule that emerges in this case is: CHOOSE THE CLUB FOR WHICH EXPECTED SCORE IS LOWER OR LOWEST if more than two choices are being made.

However, there will be many cases in which a player's utility structure will be different from the linear structure described above. To tailor the analysis to the individual player, assume that U(i) denotes the utility of scoring i, i= 1,2,3,4,5 and 6, with U(i)> U(i + j) for j>0. Note that some players associate enormous values to U(1), the utility of a hole in one and this can rule out linearity of the utility function in score and the use of a 7 iron to lay up on a par 3 hole. Formally, these considerations are taken into account by comparing the following expected utilities,

$$\text{EU given 3} = \sum_{i=1}^{6} p_i U(i) \qquad \textbf{(2.3)}$$

and

$$\text{EU given 7} = \sum_{i=1}^{6} q_i U(i) \qquad \textbf{(2.4)}$$

If (2.3) is larger than (2.4), the optimal choice is a 3 iron while if (2.4) is larger, the optimal choice is a 7 iron, using the maximization of expected utility framework. By assigning values to the p's, q's and U(i)'s and computing expected utilities as given in (2.3) and (2.4), we have arrived at an operational, optimal rule for club selection.

As regards the above solution to the club selection problem, note that we have relied on the maxim, act so as to maximize expected utility when making decisions under uncertainty, a precept supported by analyses of Ramsey, Morgenstern and von Neumann, Savage, Friedman, et al. If other approaches, e.g. those of Meginniss (1977) or Machina (1987), that involve making the utility function depend on the probabilistic structure, that is the p's and q's above, are considered more appropriate, solutions will be somewhat different. Note, that we have assigned the same utility to a birdie, that is i=2, whether it is made by use of 2 iron or by use of 7 iron. Some would prefer a birdie made using a 2 iron to one made luckily after a 7 iron tee shot. Clearly, we could elaborate the utility structure to take account of such considerations and solve for the optimal acts within these other frameworks. Last, it may be the case that the outcome space, here a set of 6 mutually exclusive and exhaustive outcomes is not specified in great enough detail. As an alternative outcome space, consider that shown in Fig. 1. Here we have allowed for much greater detail in describing alternative possible outcomes. Given that we can assess the probabilities shown in Fig. 1 and have an assignment of utilities to scores, we can evaluate the probabilities of alternative outcomes, compute expected utilities and choose between the 3 and 7 iron tee shots in an optimal fashion. Whether use of more complex utility specifications and the more complex outcome space shown in Fig. 1 will produce better results in practice is an empirical matter that can not unfortunately be settled deductively. See Mockus (1989) and Mockus, Mockus and Mockus (1991) for algorithms for computing Bayesian solutions to discrete optimization problems, such as that described in Figure 1.

Another interesting canonical problem is whether use of a 3 iron will lead to a lower or equal score than that provided by use of a 7 iron, i.e. Pr $[S(3) \leq S(7)] = P$ and $1-P$ is the probability on the outcome $S(3)>S(7)$. Here we can define a dichotomous random variable and use past data and prior information to make inferences about possible values of P. For example, if a binomial model for the outcomes is considered appropriate, given past data, various inference procedures can be employed to make inferences about possible values of P. In the Bayesian approach, noninformative or informative priors can be employed. Also, when there is some question about assuming a binomial process for the outcomes, the BMOM approach has been applied in Zellner (1997) to produce a postdata density for P that differs somewhat from that yielded by a traditional Bayesian approach using, say, a Bayes-Laplace uniform prior.

For example, if in 5 cases it is observed that $S(3) \leq S(7)$, and it is assumed that the trials are independent binomial trials, the likelihood function is just P^5 and the maximum likelihood estimate is equal to 1. On other hand, with a uniform prior, the posterior density for P is $6P^5$, with posterior mean 6/7, the Laplace Rule of Succession result, and the modal value of the posterior density is equal to 1. Thus if quadratic loss is used, the optimal point estimate is 6/7 = .857, whereas if a 0-1 loss structure is used, the optimal point estimate is 1, the modal value. Application of the BMOM approach that does not require specification of a prior or likelihood function yields the result that the postdata probability that P=1 is .871 and a postdata mean for P is .944, slightly higher than the Bayes-Laplace posterior mean, 6/7= .857; see Zellner (1997a,b) for discussion of this and other BMOM analyses and references to the literature. With the complete posterior densities available in the traditional Bayesian and BMOM approaches, they can be used along with a 2x2 loss structure to choose between the two alternatives in such a way as to minimize expected loss.

Last, in connection with dichotomous random outcomes, e.g. shoot bogey or below versus shoot above bogey, with associated probabilities, P and 1 –P respectively, it is very possible that P does not remain constant from trial to trial because of a variety of factors that are not kept constant that affect play, namely, weather, condition of course, health, age, etc. Thus, as in models for dichotomous random variable models, e.g. logit or probit models, we assume

P = P(x(t)'b), t = 1,2,...,T, where x(t) is a vector of measurable input variables and b is a vector of parameters with unknown values. See Judge, et. al (1985), Zellner and Rossi (1984), and Albert and Chib (1993) for Bayesian and non-Bayesian analyses of logit, probit and other models for dichotomous random variables. Zellner and Rossi (1984) implement the Bayesian approach to such models using importance function Monte Carlo integration techniques to compute exact finite sample results and compare them with approximate large sample Bayesian and non-Bayesian results with the finding that the large sample approximate results are not very good in small to moderate-sized samples. Albert and Chib (1993) employ data augmentation and Markov Chain Monte Carlo (MCMC) techniques to analyze this class of models as well as more general models. Thus, finite sample Bayesian procedures are available to compute exact posterior distributions for the coefficients of the input variables and predictive probability mass functions.

In the traditional Bayesian and ML approaches, it is necessary to choose a "link function," e.g. normal for the probit model, logistic for the logit model, Of course, posterior odds can be computed to compare and/or combine models using different link functions, e.g. normal, logistic, Student t, etc. Further, diagnostic checking using realized error terms, as in Albert and Chib

(1993), Hong (1989), Chaloner and Brant (1988), Zellner and Moulton (1985) and Zellner (1975) can reveal outlier, non-independence, functional form and other problems. It is also possible to investigate the possible effects of past performance on current performance by including measures of past performance in the input variables. If such effects are present, they may indicate that there is some truth to the "hot hand" or better in the present context, "hot club" hypothesis. Last, modelling annual average scores using regression or time series transfer function models is possible. Let $y' = (y_1, y_2, \ldots, y_T)$ be a vector of annual average scores for years 1, 2. . . , T, and X be a T x*k* matrix of rank *k* on *k* input variables, e.g. age, health status, etc., and $y = X\beta + u$ be the relation linking input variables to y where β is a *k*x1 vector of parameters and u is an Tx1 vector of realized error terms. For given y and X, we can write $y = XE\beta + Eu$, where E = subjective expectation operator given the data and models. If we assume $X'Eu = 0$, that is that the columns of X are orthogonal to the vector Eu, we have $\hat{b} = (X'X)^{-1}xy = E\beta + (X'X)^{-1}X'Eu = E\beta$. Thus under the above orthogonality assumption, the least squares estimate $\hat{b}$ is equal to Eβ, the "post data mean of β" that is an optimal estimate relative to a quadratic loss function. With a second assumption regarding the form of the covariance matrix for the realized error term vector u, given and discussed in Zellner (1994, 1997), Green and Strawderman (1996) and Tobias and Zellner (1997), the post data covariance matrix for β, denoted by $\text{Var}(\beta) = E\left(b - \hat{b}\right)\left(b - \hat{b}\right)' = (X'X)^{-1}s^2$, where $s^2 = \left(y - X\hat{b}\right)'\left(y - X\hat{b}\right)/(n-k)$.

Then the maximum entropy density for β given the data and the two assumptions above is $N[\hat{b}, (X'X)^{-1}s^2]$, a multivariate normal density with mean $\hat{b}$ and covariance matrix

$(X'X)^{-1}s^2$. See, e.g. articles in Fomby and Hill (1997) and in the *American Journal of Agricultural Economics,* August *1999* for comparisons of the Bayesian method of moments (BMOM) approach to the traditional Bayesian approach and to the generalized maximum entropy (GME) approach of Golan, Judge and Miller (1996).

Of course many variants of the above regression model can be considered. For example, the parameter vector β might not be constant over a life cycle. Then for each of the sub-periods of the life cycle, we may assume a relation $y_a = X_a b_a + u_a$, a = 1, 2, . . . , m sub-periods. Similarly we may want to utilize observations relating to different seasons, e.g. spring, summer, autumn and winter and have a specific regression for such seasonal scores rather than just use seasonal dummy variables in the overall regression mentioned above. As is well known, Bayesian posterior odds can be employed to choose among alternative models and/or combine them. See Judge et al (1985) for discussion of various models selection procedures and Min and Zellner (1993) for Bayesian model selection and combining techniques and their application.

To return to the first problem considered in Table 1 in which we had 6 possible outcomes for each choice of club, obviously such a problem can be modelled using a multinomial model. Then past sample and current prior information can be combined using Bayes' theorem to yield a posterior probability mass function for the underlying probabilities, the p's and q's shown in Table 1. With such posterior probability mass functions available, it is possible to use them to solve the club choice problems considered above. In addition, it is possible to go beyond multinomial or polytomous (or "polychotomous") constant parameter models to multinomial logit or probit or polytomous random variable models with covariates. See Judge, et al. (1985), McCulloch and Rossi (1991) and Albert and Chib (1993) for Bayesian and non-Bayesian procedures for analyzing this class of models. For non-Bayesian analyses, it is usually difficult to get exact finite sample results and thus users generally rely on asymptotic approximations in contrast to Bayesian analyses that yield finite sample results. See Zellner and Rossi (1984) and Albert and Chib (1993) for some comparisons of finite sample and asymptotic results.

3. Selected Economic Models of Golf

In this section we consider some canonical economic models of golf. First we take up an analysis of some problems facing a professional golfer and then turn to consider a consumer model for non-professional golfers, such as George Judge, who play golf for pleasure.

Table 3 provides information relevant for a professional golfer who is to play in a particular tournament. Some possible outcomes are that he or she finishes first, second, third, etc. or out of the money. Also shown in Table 3 are the probabilities that the professional golfer associates with each possible outcome along with the winnings for each outcome. Then, net expected winnings, NEW, associated with participating in this tournament are given by NEW = $\Sigma p_i W_i$ – Expected Costs, where expected costs include equipment, time, travel, lodging and other expenses associated with participation in this particular tournament. Knowing the value of NEW will help a golf pro to decide whether to participate in this particular tournament. Further, if the decision facing the golf pro is whether to participate in tournament A or tournament B, he or she can compute NEW for each tournament and participate in the one with the larger value of NEW. Finally, calculation of NEW for each of the tournaments in which a pro participates during a year and summing such values over the tournaments provides an estimate of annual net earnings that may be useful for income tax purposes.

As a component of an optimizing model for a professional golfer, assume that score, denoted by S is given by S = f(x, q, u), where x is a vector of inputs, e.g., practice time, equipment, instruction time, psychological consultation, etc., q is a vector of parameters and u is a stochastic error term. Then let W(S) denote winnings, here assumed just a function of score, S. Then, net expected winnings (NEW) is given by NEW = $EW(S) - a'x,$ where a' is a vector of input prices and $a'x$ is the total cost of the inputs. Not that the expectation of W(S) can involve use of a prior or posterior density for q and parameters of the density for u, the random error in the score function above. Then by maximizing NEW, the optimal input vector, x* is obtained, which when substituted in the function for NEW gives the optimal NEW. Also, summing the x*'s over golf pros for given factor prices, provides demand functions for golf inputs such as balls, clubs, instruction, etc. Empirical and theoretical work is needed to determine an appropriate functional form for the score function, the probability density function for u, and a predictive density for S. Given these inputs, the above, one period optimization problem can be solved. Further, multi-period problems can be formulated and solved using a multi-period predictive density function for S and anticipated future factor prices.

As regards a model of a "consumer" golfer who plays golf for entertainment, the following version of the Becker consumer model, as elaborated by Verma (1980) to take account of informational inputs and by Marsh and Zellner (1996) to incorporate random shocks in production functions to produce an economic random utility model (ERUM) may be useful in modeling and predicting the behavior of consumer golfers. Let such a consumer's random utility function be given by U[c(1),c(2). . . , c(m)] where c(i) is a random consumption "characteristic", assumed generated by c(i) = f[x(i),u(i)], the stochastic production function for the ith consumption characteristic where x(i) is a vector of input variables, time, information, market goods, e.g., golf balls, etc., and u(i) is a stochastic error term for i = 1, 2, . . ., m. Then it is possible to maximize EU with respect to the elements of the x(i)'s subject to the usual budget and time constraints (even though some golfers seem to violate the time constraint in a mystical way) to obtain the optimal x's, the market demands of an individual consumer which, when aggregated over consumers, provides the market demands for the elements of the x vectors and the "optimal" total time that golfers spend on courses playing golf given market prices for inputs, the price of time, income, etc.

4. Summary and Conclusions

It has been pointed out that the study of golf, as with any other topic, is a science if scientific methods for learning from data, making predictions and solving decision problems are employed, a general point made by Pearson (1938) and Jeffreys (1939) many years ago. Further, to facilitate further work in this area, that will undoubtedly benefit George Judge, a canonical decision problem involving club choice has been structured and solved using Bayesian decision theory and a degree of reasonable belief definition of probability. Further, some models for analyzing individuals' golf scores have been considered that probably will be useful in analyzing past data on individuals' golf scores. Then too, the analyses of several professional golfers' problems, as yet not too relevant for George Judge, and of consumer golfers' problems, have been presented that may be of help to economists and business personnel in their analysis of the professional recreational golf industries. With improved models and methods for the analysis of past data, prediction and decision-making, golfers and the golf industry will probably perform better in the future.

Table 1

Outcomes, Probabilities and Utilities Associated With Use of a 3 or 7 Iron Tee Shot on a Par 3 Hole

Club	Outcomes 1	2	3	4	5	6*	Expected Scores and Utilities
3 Iron							
Probabilities	p_1	p_2	p_3	p_4	p_5	p_6	$ES_3 = \sum_{i=1}^{6} p_i i$
Utilities	$U_3(1)$	$U_3(2)$	$U_3(3)$	$U_3(4)$	$U_3(5)$	$U_3(6)$	$EU_3 = \sum_{i=1}^{6} p_i U_3(i)$
7 Iron							
Probabilities	q_1	q_2	q_3	q_4	q_5	q_6	$ES_7 = \sum_{i=1}^{6} q_i i$
Utilities	$U_7(1)$	$U_7(2)$	$U_7(3)$	$U_7(4)$	$U_7(5)$	$U_7(6)$	$ES_7 = \sum_{i=1}^{6} q_i U_7(i)$

*Any score higher than a triple bogey, i.e., a 6, is recorded as a 6.

Table 2

Probabilities of Scores on Each of N Plays of a Given Hole

Trial	Scores 1 2 3 4 5 . . . . . J
	Probabilities

1	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	. . . . .p_{1J}
2	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}	. . . . .p_{2J}
3	p_{31}	p_{32}	p_{33}	p_{34}	p_{35}	. . . . .p_{3J}
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
N	p_{N1}	p_{N2}	p_{N3}	p_{N4}	p_{N8}	p_{NJ}

Table 3

Tournament Outcomes, Associated Probabilities and Winnings for a Professional Golfer in a Tournament

Golfer's Place	**1st**	**2nd**	**3rd**	**4th**	**. . .**	**m'th**	**Out of money**

Probabilities	p_1	p_2	p_3	p_4	. . .	p_m	p_o
Winnings	W_1	W_2	W_3	W_4	. . .	W_m	0

Expected Winnings $= \sum_{i=1}^{m} p_i W_i = p'W$

Net Expected Winnings $= p'W - \text{Costs}$

Table 4

Tournament Outcomes, Associated Probabilities and Winnings for a Professional Golfer in Two Tournaments

A. Tournament 1

Golfer's Place	**1st**	**2nd**	**3rd**	**4th**	**. . .**	**m'th**	**Out of money**
Probabilities	p_{11}	p_{12}	p_{13}	p_{14}	. . .	p_{1m}	p_{1o}
Winnings	W_{11}	W_{12}	W_{13}	W_{14}	. . .	W_{1m}	0

Expected Winnings $= \sum_{i=1}^{m} p_{1i} W_{1i} \, p_1 W_1$

Net Expected Winnings $= p_1 W_1$ - Costs

B. Tournament 2

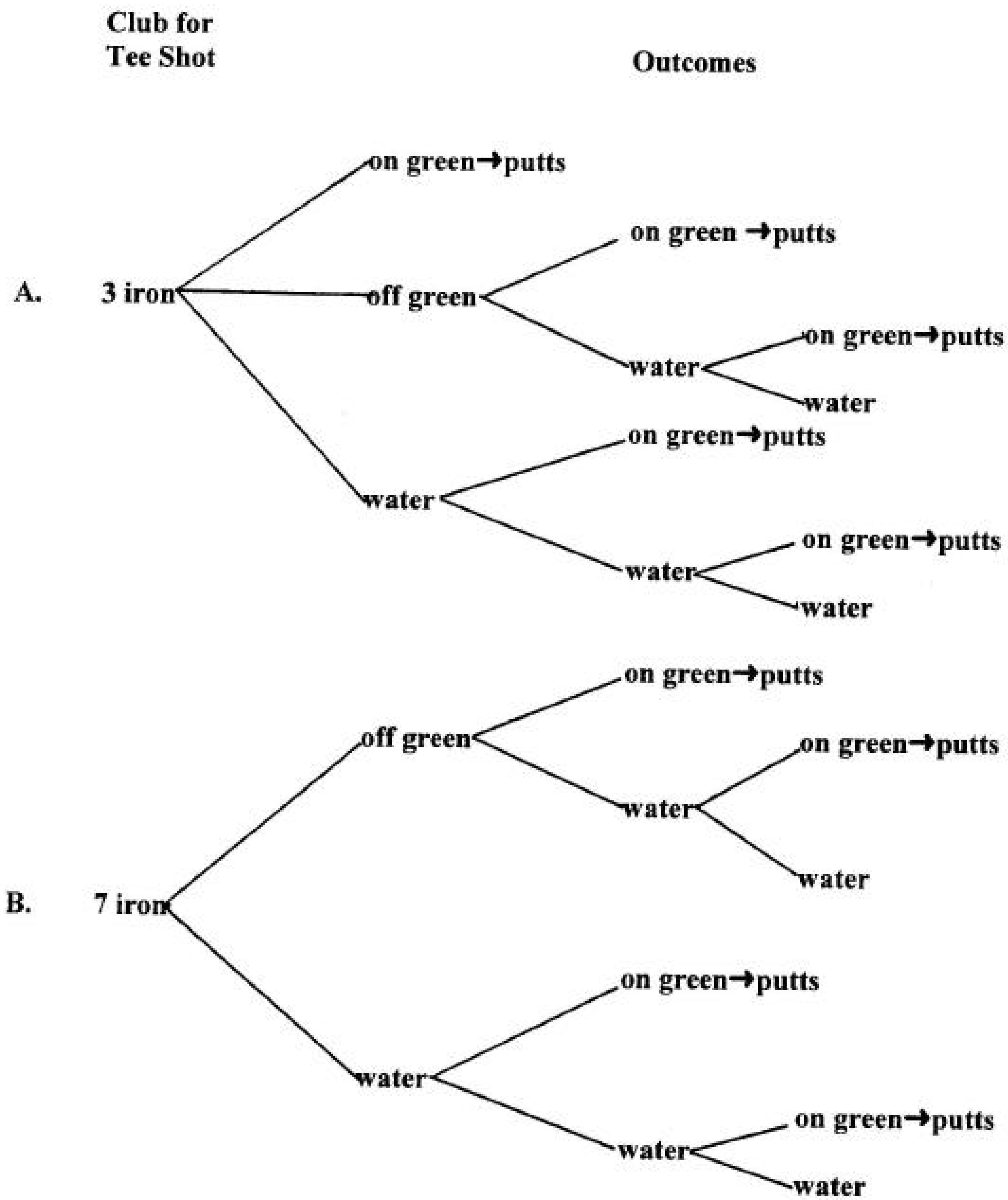
Probabilities	p_{21}	p_{22}	p_{23}	p_{24}	. . .	p_{2m}	p_{2o}
Winnings	W_{21}	W_{22}	W_{23}	W_{24}	. . .	W_{2m}	0

Expected Winnings $= \sum_{i=1}^{m} p_{2i} W_{2i} = p_2 W_2$

Net Expected Winnings $= p_2 W_2 - \text{Costs}_2$

Figure 1

Outcome Tree for Choice Between a 3 Iron and a 7 Iron Tee Shot on a Par 3 Hole[a]



[a] It is assumed that this green is completely surrounded by water. Further, "putts" denotes the possible outcomes 0, 1, 2, 3, . . . constrained by the condition that maximal score is 6.

References

Albert, J.H. and S. Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of American Statistical Association*, 88, 669-679.

Chaloner, K. and R. Brant (1988), "A Bayesian Approach to Outlier Detection and Residual Analysis," *Biometrika*, 75, 651-659.

Fomby, T. and R. Hill, eds. (1997), *Advances in Econometrics: Applying Maximum Entropy to Econometric Problems*, Vol. 12, Greenwich: Jai Press

Golan, A., G. Judge and D. Miller (1996), *Maximum Entropy Econometrics*, Chichester: John Wiley & Sons, Ltd.

Green, E. and W. Strawderman (1996), "A Bayesian Growth and Yield Model for Slash Pine Plantations," *Journal of Applied Statistics*, 23, 285-299.

Hong, C. (1989), "Forecasting Real Output Growth Rates and Cyclical Properties of Models: A Bayesian Approach," Ph.D. thesis, Department of Economics, University of Chicago.

Jaynes, E.T. (1968), "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, 227-241

Jeffreys, H. (1998), *Theory of Probability*,reprinted 3rd ed. In Oxford Classic Texts, London: Oxford University Press.

Judge, G.G., et al (1985), *The Theory and Practice of Econmetrics* (2nd ed.), New York: Wiley.

Machina, M. (1987), "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives*, 1, 121-155.

McCulloch, R. and P.E. Rossi (1991), "A Bayesian Analysis of the Multinomial Probit Model," Working Paper, Graduate School of Business, University of Chicago.

Meginniss, J. (1977), "Alternatives to the Expected Utility Rule," Ph.D. thesis, Graduate School Of Business, University of Chicago.

Min, C. and A. Zellner (1993), "Bayesian and Non-Bayesian Methods for Combining Models And Forecasts with Applications to Forecasting International Growth Rates," *Journal of Econometrics*, 56, 89-118.

Mockus, J. (1989), *Bayesian Approach to Global Optimization*, Dordrecht-Longon-Boston: Kluwer Academic Publishers.

Mockus, A., J. Mockus and L. Mockus (1994), "Bayesian Approach to Discrete Optimization," *Technical Report, Institute of Mathematics and Informatics*, Vilnius, 21 pp.

Mosteller, F. and C. Yotz (1992), "Professional Golf Scores Are Poisson on the Final Tournament Days," *American Statistical Association Proceedings of the Sports Statistics Section*, 39-51.

Pearson, K. (1938), *The Grammar of Science*, London: Everyman Edition.

Tinbergen, J. (1954), *Centralization and Decentralization in Economic Policy*, Amsterdam: North-Holland Publishing Co.

Tobias, J. and A. Zellner (1997), "Further Results on Bayesian Method of Moments Analysis of the Multiple Regression Model," H.G.B. Alexander Research Foundation, Graduate School of Business, University of Chicago, to appear in the *International Economic Review.*

Verma, V. (1980), "A Price Theoretic Approach to the Specification and Estimation of the Sales Advertising Function," *Journal of Business*, 53, July 1980.

Zellner, A. (1973), "The Quality of Quantitative Economic Policy-making When Targets and Costs of Change are Mis-specified," in *Selected Readings in Econometrics and Economic Theory: Essays in Honour of Jan Tinbergen*, Part 2, Macmillan, 147-164, reprinted in Zellner (1997b).

______ (1997a), "The Bayesian Method of Moments (BMOM): Theory and Applications," in T. Fomby and R. Hill (eds.), *Advances in Econometrics*, Vol. 12, 85-105.

______ (1997b), *Bayesian Analysis in Econometrics and Statistics: The Zellner View and Papers* Cheltenham, Economists of the Twentieth Century Series, Cheltenham, UK: Edward Elgar Publishing Ltd.

______ and L. Marsh (1996), "Bayesian Optimization Procedures for Firms Facing Discrete Product Demand, Economic Random Utility Models and Industry Equilibrium," presented to the American Statistical Association meeting, 1995, and in *Proceedings of the Section on Bayesian Statistical Science*

______ and B.R. Moulton (1985), "Bayesian Regression Diagnostics with Applications to International Consumption and Income Data," *Journal of Econometrics*, 29, 187-211 and Reprinted in Zellner (1997b).

______ and P.E. Rossi (1984) "Bayesian Analysis of Dichotomous Quantal Response Models," *Journal of Econometrics*, 25, 365-393, reprinted in Zellner (1997b).