J. R. Statist. Soc. A,
(1974), **137**, *Part* 2, p. 227

Reliability of Subjective Evaluations in a High Incentive Situation

By ARTHUR E. HOERL and HERBERT K. FALLIN

University of Delaware *U.S. Army Material Systems Analysis Agency*

SUMMARY

This paper presents the results of an analysis of horse race data collected from Aqueduct and Belmont Park in 1970. These data are used to demonstrate the reliability of subjective evaluations when incentive is offered to the subjects.

Keywords: SUBJECTIVE PROBABILITY; PERSONAL PROBABILITY; INCENTIVE

1. INTRODUCTION

1.1. *Background*

ON May 19th, 1965, Professor Cedric A. B. Smith (1965) read a paper entitled "Personal probability and statistical analysis" before the Royal Statistical Society. Professor Smith considered a hypothetical example of the effect of some treatments, i.e. fertilizer or sprays, on some fruit trees in order to illustrate some methods using personal or subjective probability. Professor E. S. Pearson raised a question during a discussion on Professor Smith's paper as to what progress has been made at this date regarding the arguments for and against the Bayesian (subjective vs objective) approach. Addressing the question himself, Professor Pearson concluded in part: "In the first place, our discussions are still nearly always based on what are, after all, somewhat artificial examples. These serve to illustrate the mechanics of applying a specific piece of theory, but they tell us little of the extent of which this theory is usable and helpful to the practical statistician. We still, I feel, lack any adequate reporting on real case histories which would show how the working statistician, as distinct from the writer on theory, does or can in fact use these and other methods in marshalling the information that leads him to make real recommendations in practice."

A very recent study by Pearn (1973) discusses the application of subjective interpretation of risks in genetic counselling. As stated by Pearn, "One essential part of the genetic counselling interview is the communication to patients of the recurrence risks to relatives, usually to future children." But even this study, although indicative of areas of application, does not contain actual validation of the subjective procedure.

1.2. *Purpose and Scope*

The purpose of this paper is to demonstrate the reliability of subjective estimates in a high incentive situation using data that may be regarded as real case histories in which actual outcomes are known. The data used comprised the results of all the thoroughbred horse races run at Aqueduct and Belmont Park in 1970 (a total of 1,825 races).

2. SUBJECTIVE PROBABILITY

2.1. *Definition*

The subjective probability of an event can be defined as the degree of belief an individual has that that event will occur. It has sometimes been interpreted in the

TABLE 1

Comparison of subjective probabilities of winning and actual frequencies of wins as a function of odds rank of horse

No. of entries	*No. of races*		*Ranking by odds* 1	2	3	4	5	6	7	8	9	10	11	12	*Computed* χ^2	$\psi^2_{0.95}$
5	69	Subj. prob.	0·42	0·25	0·17	0·11	0·06								3·2	9·5
		Obs. freq.	0·41	0·30	0·20	0·07	0·03									
6	181	Subj. prob.	0·36	0·23	0·17	0·12	0·08	0·04							10·5	11·1
		Obs. freq.	0·43	0·21	0·20	0·11	0·03	0·02								
7	312	Subj. prob.	0·33	0·22	0·16	0·12	0·09	0·06	0·03						5·7	12·6
		Obs. freq.	0·34	0·21	0·16	0·12	0·08	0·08	0·02							
8	352	Subj. prob.	0·31	0·20	0·15	0·12	0·09	0·06	0·04	0·03					7·1	14·1
		Obs. freq.	0·33	0·25	0·13	0·09	0·07	0·06	0·04	0·02						
9	283	Sub. prob.	0·30	0·20	0·15	0·11	0·09	0·06	0·05	0·03	0·02				13·1	15·5
		Obs. freq.	0·35	0·15	0·17	0·13	0·08	0·06	0·02	0·01	0·02					
10	241	Subj. prob.	0·29	0·19	0·14	0·11	0·08	0·06	0·05	0·03	0·02	0·02			5·0	16·9
		Obs. freq.	0·31	0·17	0·16	0·10	0·07	0·07	0·06	0·04	0·02	0·01				
11	154	Subj. prob.	0·27	0·18	0·14	0·11	0·08	0·07	0·05	0·04	0·03	0·02	0·01		11·8	18·3
		Obs. freq.	0·27	0·18	0·19	0·08	0·05	0·05	0·05	0·05	0·04	0·04	0·01			
12	233	Subj. prob.	0·26	0·17	0·13	0·10	0·08	0·07	0·05	0·04	0·03	0·02	0·02	0·01	3·2	19·7
		Obs. freq.	0·28	0·14	0·17	0·12	0·10	0·06	0·02	0·05	0·03	0·03	0·01	0·00		

context of a gambling situation as the ratio of the amount of money a gambler is willing to risk on a particular outcome to the total amount of money he has available. On the other hand, Bertrand Russell (1948) has warned .... "Subjective certainty, therefore, is not a guarantee of truth, or even of a high degree of credibility."

2.2. *Applied to Horse Races*

In any event, the amount of money bet on a horse is directly proportional to the confidence the betting population has in that horse, and is inversely proportional to the odds (as computed by the totalizator) of that horse winning the race. Therefore, the subjective probability that a horse wins a race can be computed as the normalized reciprocal of that quantity comprising the odds to a dollar of that horse winning plus one.

3. COMPARISON OF SUBJECTIVE PROBABILITIES VS ACTUAL FREQUENCIES

3.1. *Results*

Table 1 presents a comparison of subjective probabilities of winning and actual frequencies of wins as a function of odds rank for races with 5–12 entries. Also shown are computed χ^2 values with the appropriate values of the χ^2 statistic for the 0·95 confidence level.

3.2. *Assessment of Results*

The results indicate good agreement between the expected and actual values. In no case is the null hypothesis that the subjective probabilities are the correct theoretical frequencies rejected. A slight tendency to overpredict at the right tail is noted and is interpreted to mean that more money is bet on the long shots than is consistent with their frequency to wins. It is also indicative of futile attempts by the racing fans to recoup previous losses.

However, the reliability of subjective evaluations, when the subject has incentive to estimate well, has been clearly demonstrated.

3.3. *Further Results*

A summary of the mean order of finish is presented in Table 2 as a function of odds rank of horses for races with 5–12 entries. These data are seen to be amazingly

TABLE 2

Mean order of finish as a function of odds rank of horse

No. of entries	*No. of races*	*Ranking by odds*											
		1	2	3	4	5	6	7	8	9	10	11	12
5	69	2·1	2·4	2·9	3·4	4·1							
6	181	2·2	2·9	3·2	3·6	4·2	4·9						
7	312	2·8	3·2	3·7	4·0	4·3	4·6	5·4					
8	352	2·8	3·2	3·9	4·2	4·7	5·1	5·7	6·4				
9	283	3·1	3·6	4·1	4·6	5·1	5·3	6·0	6·4	7·1			
10	241	3·1	4·0	4·3	5·1	5·3	5·6	6·2	6·5	7·0	7·9		
11	154	3·8	4·0	4·7	5·2	5·7	5·8	6·3	6·9	7·2	7·8	8·5	
12	233	3·9	4·6	5·1	5·4	6·0	6·2	6·7	7·2	7·6	7·7	8·7	9·1

consistent in that for each level of number of entries, the mean order of finish is a monotonically non-decreasing function of the composite group ranking of the horses. This example demonstrates that individuals with incentive can on the average successfully discriminate small differences in items (such as the relatively small inate differential rates of speed of thoroughbred race horses).

REFERENCES

PEARN, J. H. (1973). Patient's subjective interpretation of risks offered in genetic counselling. *J. Med. Genetics*, **10**, 129–134. (The authors are indebted to the referee for bringing to their attention the useful bibliography contained in this reference.)

RUSSELL, B. (1948). *Human Knowledge, Its Scope and Limits* (Part V). New York: Simon & Schuster.

SMITH, C. A. B. (1965). Personal probability and statistical analysis. *J. R. Statist. Soc.* A, **128**, 469–499.