STATISTICAL INFERENCE BASED ON
THE POSSIBILITY AND BELIEF MEASURES

YUAN YAN CHEN

ABSTRACT. In statistical inference, we infer the population parameter based on
the realization of sample statistics. This can be considered in the framework of
inductive inference. We showed, in Chen (1993), that if we measure a parameter
by the possibility (or belief) measure, we can have an inductive inference similar
to the Bayesian inference in belief update. In this article we apply this inference
to statistical estimation and hypotheses evaluation (testing) for some parametric
models, and compare them to the classical statistical inferences for both one-
sample and two-sample problems.

1. INTRODUCTION

In statistical inference the treatment of a parameter is always at the center of
debate. The population parameter is a constant. However, on the basis of the
sample statistics the precise value of the parameter cannot be determined with
absolute certainty. If we evaluate the plausible values of the parameter, then we
have the concept of likelihood distribution of the parameter, as stated by Fisher
(1956), "The likelihood supplies a natural order of preference among the pos-
sibilities under consideration." Bayesian inference represents the distribution
of a parameter by a probability function. However, many authors, including
Barnard, Bartlett, and Fisher, concluded that parameters with (fiducial) dis-
tributions cannot be regarded as random variables in Kolmogoroff's sense (see
Barnard (1987)). This prompted us to introduce a concept of stationary variable
for an unknown parameter in Chen (1993) to differentiate it from the concept
of random variable, and we propose to measure a stationary variable by the
possibility measure.

Possibility theory as proposed by Zadeh (1978) is derived from the theory
of fuzzy sets. The uncertainty models similar to the possibility theory were
also considered by Shackle (1961), Cohen (1970), and Shafer (1976). The basic
operation of possibility measure is the sup operation, which is considered as a
pseudo-addition (e.g., see Ichihashi et al. (1988)). Since sup and addition have
the same algebraic property, possibility measure, although it is a weaker scale
of measurement, is parallel to probability measure. As a consequence, in statis-
tical inference, every belief representation under Bayesian inference can have a

counterpart under possibility inference, and vice versa (if we allow the improper probability function.) Both Bayesian inference and the possibility inference discussed in this article are based on the principle of inverse inference—the more plausibly a hypothesis can produce the evidence, the more likely it is to be true. However, from the logical point of view the inference from the sample to the population and the inference from the population to the sample are two different types of reasoning processes; one is an inductive reasoning and the other is a deductive reasoning. Thus, there is a need for two different kinds of uncertainty measures for the statistical inference. The calculus of probability measure has been shown to be not suitable for the logic of inductive support by many authors, e.g. Popper and Miller (1987). Thus, we opt for the possibility measure and its conjugate, belief measure, which is considered as an inductive probability by Cohen (1970), as a basis for statistical estimation and hypothesis evaluation.

The reasoning for the truth or falsity of a hypothesis is a judgment and not a decision; a decision is for anticipating an action. Without the measurement of belief we are obliged to employ decision theory for the hypothesis judgment. In classical hypothesis testing when we fail to reject a null hypothesis there is a logical ambiguity: whether we lack the evidence to make a confirmed judgment, or the null hypothesis is supported, which is very often the interpretation in statistical practice. As a result, a promising experiment (e.g. clinical trials) would be regarded as "insignificant" simply because of the shortage of sample sizes. Thus, we need a theory of support for the statistical inference.

2. Possibility measure, belief measure, and statistical inference

Possibility measure is a fuzzy measure based on evaluation (e.g. see Wang (1984)); thus the basic operation is sup (and inf). The possibility and necessity measures are also a subclass of plausibility and belief (upper and lower probabilities) measures (e.g. see Chen (1993) and Shafer (1976)). In this article we adopt the notations of plausibility (Pl) and belief (Bel) for the possibility and necessity measures.

Definition. Let $\mathscr{B}$ be a Borel σ-algebra on Ω. A function $\text{Pl} : \mathscr{B} \to [0, 1]$ is *possibility measure* if

$$
\begin{aligned}
&\text{(i)} && \text{Pl}(\varnothing) = 0, \ \text{Pl}(\Omega) = 1, \\
&\text{(ii)} && A_n \uparrow A \to \text{Pl}(A_n) \uparrow \text{Pl}(A), \\
&\text{(iii)} && \text{Pl}(A \cup B) = \text{Pl}(A) \vee \text{Pl}(B), \ \forall A, B \in \mathscr{B}, \\
&&& A \cap B = \varnothing \ \text{(fuzzy additivity)},
\end{aligned}
$$

(2.1)

where $\vee$ is a sup operator. Note that the condition $A \cap B = \varnothing$ in (iii) can also be dropped. A conjugate measure is a necessity measure or a belief measure, which is defined by $\text{Bel}(A) = 1 - \text{Pl}(\overline{A})$. Under the belief measure we have

$$
\begin{aligned}
&\text{(i)}' && \text{Bel}(\varnothing) = 0, \ \text{Bel}(\Omega) = 1, \\
&\text{(ii)}' && A_n \downarrow A \to \text{Bel}(A_n) \downarrow \text{Bel}(A), \\
&\text{(iii)}' && \text{Bel}(A \cap B) = \text{Bel}(A) \wedge \text{Bel}(B), \ \forall A, B \in \mathscr{B},
\end{aligned}
$$

(2.2)

where $\wedge$ is an inf operator.

Similarly to the probability measure, the possibility measure can be characterized by a distribution function. Let $l(\theta) = \text{Pl}(\{\theta\})$; then

$$(2.3) \qquad \text{Pl}(A) = \sup_{\theta \in A} l(\theta).$$

The function $l(\theta)$ is called the likelihood function in Chen (1993). Note that the likelihood function plays a role similar to that of the probability function and the operator sup or $\vee$ acts as $\int$ or $\sum$ in probability measure.

Under the possibility model, if we believe A is true, but we are not absolutely certain, then we have $\text{Bel}(A) = s$, $0 < s < 1$, where s is the confidence level. The rest of the belief or doubt, $1 - s$, is attributed to $\text{Pl}(\overline{A})$. Since by (iii)' and (iii) we have $\text{Bel}(\overline{A}) = 0$ and $\text{Pl}(A) = 1$, these two values do not need to be specified. Thus, although the possibility model belongs to the class of upper and lower probabilities model, it is actually a single-value "probability" model in disguise. There is no need to embed a probability measure in between belief and possibility measures.

For the statistical inference under possibility and belief measures we state a few theorems from Chen (1993).

Theorem 2.1 (Chen (1993)). *If $X_1, \ldots, X_n | \theta \sim$ i.i.d. $P(x|\theta)$, and $l(\theta)$ is prior belief for θ, then the posterior belief for $\theta | x_1, \ldots, x_n$ is*

$$(2.4) \qquad l(\theta | x_1, \ldots, x_n) = l(\theta) p(x_1, \ldots, x_n | \theta) / \sup_{\theta \in \Theta} l(\theta) p(x_1, \ldots, x_n | \theta).$$

If $l(\theta) = 1$, $\forall \theta \in \Theta$, which is known as a vacuous belief, then (2.4) reduces to $l(\theta | x_1, \ldots, x_n) = kp(x_1, \ldots, x_n | \theta)$. Thus, $l(\theta | x_1, \ldots, x_n)$ can be considered as a normalized "likelihood function", $L(\theta | x_1, \ldots, x_n) = p(x_1, \ldots, x_n | \theta)$.

Theorem 2.2. *If $l(\theta) = 1$, $\forall \theta \in \Theta$, and $g : \theta \to \eta$ is a function onto Θ', then $l(\eta) = 1$, $\forall \eta \in \Theta'$.*

Proof. $l(\eta) = \sup_{g(\theta) = \eta} l(\theta) = 1$, $\forall \eta \in \Theta'$.

This theorem indicates that a vacuous likelihood function is still vacuous under transformation. This property does not hold for a uniform probability function. Thus, Bayesian inference has difficulty incorporating the concept of vacuous belief; furthermore, a uniform prior does not always exist in Bayesian inference.

If we define the weight of evidence in the same way as Shafer (1976),

$$(2.5) \qquad W(H|E) = -\log(1 - \text{Bel}(H|E)),$$

then this weight of evidence is also equivalent to Bayesian weight of evidence defined by Good (1950). If $\text{Pl}(H) = 1$, $\text{Pl}(\overline{H}) = 1$ and $P(E|H) \geq P(E|\overline{H})$, by (2.4) we have

$$(2.6) \qquad \log(P(E|\overline{H}) / P(E|H)) = -\log(1 - \text{Bel}(H|E)),$$

where $\log(P(E|\overline{H}) / P(E|H))$ is Good's definition of weight of evidence. And if $P(E|H) < P(E|\overline{H})$, then by Good's definition, the support of H given E is negative, which is equivalent to assigning a positive support to $\overline{H}$.

Theorem 2.3 (Chen (1993)). *Let $X_1|\theta$ and $X_2|\theta$ be independent random variables with probability function $p_1(x_1|\theta)$ and $p_2(x_2|\theta)$ respectively, and $l(\theta) = 1$, $\forall \theta \in \Theta$. If $\sup_{\theta \in \Theta} l_1(\theta|x_1) l_2(\theta|x_2) > 0$, then*

$$(2.7) \qquad l(\theta|x_1, x_2) = l_1(\theta|x_2) l_2(\theta|x_2) / \sup_{\theta \in \Theta} l_1(\theta|x_1) l_2(\theta|x_2).$$

Equation (2.7) is known as the likelihood rule. Note that (2.7) holds even if x_1 and x_2 are vectors. Thus, if we have likelihood functions from two independent samples, they can also be combined by (2.7). This provides a simple method for a meta-analysis.

We now look at the likelihood rule from an evidential point of view.

Corollary. *Let H and $\overline{H}$ be two rival hypotheses, and let E_1, E_2 be two independent evidences.*

  (i) *If E_1 and E_2 both support H, then $W(H|E_1, E_2) = W(H|E_1) + W(H|E_2)$.*

  (ii) *If E_1 supports H, E_2 supports $\overline{H}$ and $W(H|E_1) > W(\overline{H}|E_2)$, then $W(H|E_1, E_2) = W(H|E_1) - W(\overline{H}|E_2)$.*

Proof. By (2.7) $\text{Bel}(H|E_1) = s_1$ and $\text{Bel}(H|E_2) = s_2 \Rightarrow \text{Bel}(H|E_1, E_2) = 1 - (1 - s_1)(1 - s_2)$, and $\text{Bel}(H|E_1) = s_1$, $\text{Bel}(\overline{H}|E_2) = s_2$, and $s_1 > s_2 \Rightarrow \text{Bel}(H|E_1, E_2) = 1 - (1 - s_1)/(1 - s_2)$. So we have the results.

From this corollary we see that a belief update under the likelihood rule reduces to simple addition or subtraction of weights of evidence.

Remark. Shafer (1976) also suggests measuring a parameter by the possibility and belief measures, which he called a consonant belief function. But he thought that dissonant evidence could produce partial supports to rival hypotheses. Thus, he proposed using Dempster's rule to combine the statistical evidence. As a result the combined belief of each sample is not the same as the belief of the combined sample; this has been much criticized. Smets (1982) pointed out that if we use the likelihood rule instead of Dempster's rule this difficulty can be resolved.

3. COMPARISON OF LIKELIHOOD INFERENCES AND CLASSICAL INFERENCES

In the next two sections we will compare the possibility inferences with the classical statistical inferences based on decision and sampling theory. We refer to the inference based on (2.4) with vacuous prior as the likelihood inference since it is close to Fisher's concept of *mathematical likelihood*, and it also satisfies the likelihood principles advocated by Birnbaum (1962) and Edwards (1972), etc. With the likelihood function of possibility measure to portray the likelihood, we can have joint likelihood, marginal likelihood and prior belief; thus, many early criticisms of the likelihood inference can now be resolved.

The likelihood inference measures the whole hypothesis space. Several hypotheses can be simultaneously entertained; there is no need to set up the null and alternative hypotheses. However, many problems of classical hypothesis testing still can be applied. If $\text{Bel}(H_1) \geq 1 - \alpha$, and α is small, then it is logically equivalent to rejecting a null hypothesis in classical hypothesis testing. But, unlike the classical hypothesis testing, we can also accept H_0 as true, if $\text{Bel}(H_0) \geq 1 - \alpha$, although this seldom holds for a point null hypothesis.

The precise location of the true parameter cannot be ascertained from the likelihood function. However, we have confidence that it lies in a certain region, which will be called a likelihood interval.

Definition. Let θ be a stationary variable on Θ. A subset $A \subset \Theta$ is a likelihood interval for θ with a confidence level $1 - \alpha$ if

$$(3.1) \qquad A = \{\theta | l(\theta) \geq \alpha\}.$$

This kind of interval is equivalent to an α-cut set in fuzzy set theory. From (3.1) obviously we have $\mathrm{Bel}(\theta \in A) \geq 1 - \alpha$. A likelihood interval is a rational belief on where the parameter is located after the samples are realized, while a classical confidence interval is a predicted coverage of the parameter before the samples are selected. A discussion of these conceptual differences can be found in Hacking (1975). Note that the concept of likelihood interval was also realized by Fisher (1956).

We first look at the binomial case. Let x be the observed outcome from a binomial distribution $B(n, p)$. A likelihood interval for p with confidence level $1 - \alpha$ is (p_L, p_U), where p_L and p_U are the solutions of

$$(3.2) \qquad p^x(1 - p)^{n-x}\{(x/n)^x((n - x)/n)^{n-x}\}^{-1} = \alpha.$$

Table 1 shows a few interval estimates for sample sizes $n = 10, 50,$ and 100.

TABLE 1. Likelihood limits for binomial p with confidence level $1 - \alpha$

α	x	$n = 10$		x	$n = 50$		x	$n = 100$	
		p_L	p_U		p_L	p_U		p_L	p_U
.01	1	.0004	.554	5	.018	.273	10	.033	.214
.05		.002	.455		.027	.234		.042	.189
.1		.004	.403		.033	.214		.048	.176
.01	3	.031	.756	15	.135	.560	30	.176	.448
.05		.055	.678		.161	.528		.197	.418
.1		.072	.634		.176	.448		.209	.403
.01	5	.113	.887	25	.295	.705	50	.352	.648
.05		.164	.836		.332	.688		.379	.621
.1		.197	.803		.352	.648		.394	.606

Next, we look at the normal cases. Let $x_1, \ldots, x_n$ be the observed outcome from a normal distribution $N(\mu, \sigma^2)$. First we assume σ is known and μ is to be estimated. Then the likelihood limits for μ at confidence level $1 - \alpha$ are the solutions of

$$(3.3) \qquad \exp\{-n(\overline{x} - \mu)^2/2\sigma^2\} = \alpha.$$

Thus they are

$$(3.4) \qquad \overline{x} \pm c_\alpha \sigma/\sqrt{n},$$

where $c_\alpha = (-2 \ln \alpha)^{1/2}$.

TABLE 2. Coefficients of c_α and the corresponding $z_{\alpha/2}$

coefficient	$\alpha = .1$	$\alpha = .05$	$\alpha = .02$	$\alpha = .01$	$\alpha = .005$
c_α	2.146	2.448	2.797	3.035	3.255
$z_{\alpha/2}$	1.645	1.960	2.326	2.576	2.807

Table 2 shows the comparison of the coefficients c_α and the corresponding coefficients $z_{\alpha/2}$ of the classical method, where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution.

If σ is unknown, then both μ and σ are stationary variables; we need a joint likelihood function of (μ, σ). Since the marginal likelihood $l(\mu|x_1, \ldots, x_n) = \sup_{0 < \sigma < \infty} l(\mu, \sigma|x_1, \ldots, x_n)$, the likelihood limits for μ at confidence level $1 - \alpha$ are the solutions of

$$(3.5) \qquad \{1 + (\overline{x} - \mu)^2)/\hat{\sigma}^2\}^{-n/2} = \alpha,$$

where $\hat{\sigma}^2 = \sum_{i=1}^n (x_i - \overline{x})^2/n$ is the MLE of σ^2. So they are

$$(3.6) \qquad \overline{x} \pm (\alpha^{-2/n} - 1)^{1/2} \hat{\sigma},$$

or equivalently

$$(3.7) \qquad \overline{x} \pm c_{n,\alpha} s/\sqrt{n},$$

where $s = \{\sum_{i=1}^n (x_i - \overline{x})^2/(n-1)\}^{1/2}$ and $c_{n,\alpha} = \{(n-1)(\alpha^{-2/n} - 1)\}^{1/2}$.

Table 3 shows the coefficients of $c_{n,\alpha}$ in comparison with the $t_{n-1,\alpha/2}$ of Student's t-values. We notice that as $n \to \infty$, Table 2 and Table 3 give the same result, as expected.

TABLE 3. Student's t-values and the corresponding likelihood coefficients

	$n = 5$		$n = 10$		$n = 20$		$n = 30$		$n \to \infty$	
α	$t_{n-1,\alpha/2}$	$c_{n,\alpha}$	$t_{n-1,\alpha/2}$	$c_{n,\alpha}$	$t_{n-1,\alpha/2}$	$c_{n,\alpha}$	$t_{n-1,\alpha/2}$	$c_{n,\alpha}$	$t_{n-1,\alpha/2}$	$c_{n,\alpha}$
.1	2.132	2.459	1.833	2.294	1.729	2.218	1.699	2.194	1.645	2.146
.05	2.776	3.043	2.262	2.718	2.093	2.576	2.045	2.532	1.960	2.448
.02	3.747	3.889	2.821	3.268	2.539	3.016	2.462	2.940	2.326	2.797
.01	4.604	4.609	3.250	3.689	2.861	3.334	2.756	3.228	2.576	3.035
.005	5.595	5.413	3.690	4.119	3.174	3.643	3.038	3.505	2.807	3.255

4. TWO-SAMPLE PROBLEMS

In two-sample statistical inference we are interested in evaluating the relation between parameters of two populations. Mathematically, a relation can be considered as a subset of a product space.

Definition. If (θ_1, θ_2) is the joint likelihood function of θ_1 and θ_2, and $R \subset \Theta_1 \times \Theta_2$ is a relation, then

$$(4.1) \qquad \mathrm{Pl}((\theta_1, \theta_2) \subset R) = \sup_{(\theta_1, \theta_2) \subset R} l(\theta_1, \theta_2)$$

is a likelihood relation between θ_1 and θ_2 with respect to R.

Let $X_1, \ldots, X_{n_1}|\theta_1$ be i.i.d. $\sim p_1(x|\theta_1)$ and $Y_1, \ldots, Y_{n_2}|\theta_2$ be i.i.d. $\sim$ $p_2(x|\theta_2)$; the X's and Y's are independent. If the prior beliefs $l(\theta_1)$ and $l(\theta_2)$ are vacuous, then the joint likelihood function of posterior belief is

(4.2)
$$l(\theta_1, \theta_2|x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2})$$
$$= \frac{p_1(x_1, \ldots, x_{n_1}|\theta_1)p_2(y_1, \ldots, y_{n_2}|\theta_2)}{\sup_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2} p_1(x_1, \ldots, x_{n_1}|\theta_1)p_2(y_1, \ldots, y_{n_2}|\theta_2)},$$

or equivalently
(4.3)
$$l(\theta_1, \theta_2|x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) = l_1(\theta_1|x_1, \ldots, x_{n_1})l_2(\theta_2|y_1, \ldots, y_{n_2}).$$

We first look at the binomial case. Let x and y be the observed outcomes from two independent binomial distributions $B(n_1, p_1)$ and $B(n_2, p_2)$ respectively. Then we have

(4.4) $\mathrm{Pl}(p_1 = p_2|x, y) = \hat{p}^{x+y}(1 - \hat{p})^{n_1+n_2-x-y}/\hat{p}_1^x(1 - \hat{p}_1)^{n_1-x}\hat{p}_2^y(1 - \hat{p}_2)^{n_2-y},$

where $\hat{p} = (x + y)/(n_1 + n_2)$, $\hat{p}_1 = x/n_1$ and $\hat{p}_2 = y/n_2$.

Remark. Fisher (1956) also derived a formula the same as (4.4) from a likelihood argument.

Next we consider the normal cases. Let $x_1, \ldots, x_{n_1}$ and $y_1, \ldots, y_{n_2}$ be the observed outcomes from two independent normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. First we assume σ_1 and σ_2 are known; then we have
(4.5)
$$\mathrm{Pl}(\mu_1 = \mu_2|x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) = \exp\{-(\bar{x} - \bar{y})^2/2(\sigma_1^2/n_1 + \sigma_2^2/n_2)\},$$

where $\bar{x} = \sum x_i/n_1$, $\bar{y} = \sum y_j/n_2$ are the two sample means.

In comparison with the classical two-sample test we have

$$\mathrm{Bel}(\mu_1 \neq \mu_2|x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) \geq 1 - \alpha,$$

if

(4.6)
$$|\bar{x} - \bar{y}| \geq (-2\ln\alpha)^{1/2}(\sigma_1^2/n_1 + \sigma_2^2/n_2)^{1/2}.$$

Again, in statistical practice, the difference between the likelihood inference and the classical inference is the constants, $c_\alpha = (-2\ln\alpha)^{1/2}$ versus $z_{\alpha/2}$.

If σ_1 and σ_2 are unknown, based on the likelihood inference we have

(4.7)
$$\mathrm{Pl}(\mu_1 = \mu_2|x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2})$$
$$= \sup_\mu\{(1 + (\bar{x} - \mu)^2)/\hat{\sigma}_1^2\}^{-n_1/2}\{(1 + (\bar{y} - \mu)^2)/\hat{\sigma}_2^2\}^{-n_2/2},$$

where $\hat{\sigma}_1^2 = \sum(x_i - \bar{x})^2/n_1$ and $\hat{\sigma}_2 = \sum(y_j - \bar{y})^2/n_2$ are the MLE of σ_1^2 and σ_2^2. Note that there is no explicit formula for $\hat{\mu}$; however, the supremum of (4.7) can be obtained by a numerical approximation.

Next let us discuss a simple relationship between the one-sample and two-sample likelihood inference.

Theorem 4.1. *Let θ_1 and θ_2 be two stationary variables on Θ. If* $\mathrm{Bel}(\theta_1 \in A) \geq 1 - \alpha$, $\mathrm{Bel}(\theta_2 \in B) \geq 1 - \alpha$, *and $A \cap B = \varnothing$, then we have* $\mathrm{Bel}(\theta_1 \neq \theta_2) \geq 1 - \alpha$.
Proof. Since $A \cap B = \varnothing$, we have $\mathrm{Bel}(\theta_1 \neq \theta_2) \geq \mathrm{Bel}(A \times \Theta \cap B \times \Theta) = \mathrm{Bel}(A \times \Theta) \wedge \mathrm{Bel}(B \times \Theta) \geq 1 - \alpha$.

The previous theorem indicates that if the likelihood intervals with confidence level $1 - \alpha$ for θ_1 and θ_2 do not overlap, then we can infer that θ_1 and θ_2 are not equal with a confidence level at least $1 - \alpha$. Note that the independence assumption of θ_1 and θ_2 is not required here.

5. CONCLUDING REMARKS

It is always an important philosophical issue if there is a mathematical model of *inductive probability*, different from the model of *mathematical probability*. In this article we show that the possibility and belief measures can be employed as *inductive probability* in statistical inference. Although we only address the parametric models, a similar approach can be applied to a nonparametric model by considering a population distribution F as a "stationary process"; we can obtain an inference similar to Ferguson (1973).

Fisher promoted the tests of significance, but he held the viewpoint that statistical inference is a part of the inductive inference. In Fisher (1935) he stated, "The inferences of classical theory of probability are all deductive in character....*mathematical likelihood* appear[s] to take its place as a measure of rational belief when we reason from the sample to the population." He anticipated an uncertainty model of *mathematical likelihood* different from the model of *mathematical probability* (see also Fisher (1936), (1956)). Although he never developed a rigorous theory for the *mathematical likelihood*, he did derive some results similar to the likelihood inference discussed in this article.

REFERENCES

G. A. Barnard (1987), *R. A. Fisher—a true Bayesian?*, Internat. Statist. Rev. **55**, 183–189.

A. Birnbaum (1962), *On the foundations of statistical inference*, J. Amer. Statist. Assoc. **57**, 269–326.

Y. Y. Chen (1993), *Bernoulli trials: from a fuzzy measure point of view*, J. Math. Anal. Appl. **175**, 392–404.

L. J. Cohen (1970), *The implications of induction*, Methuen, London.

A. W. F. Edwards (1972), *Likelihood*, Cambridge Univ. Press, Cambridge.

T. S. Ferguson (1973), *A Bayesian analysis of some nonparametric problems*, Ann. Statist. **1**, 209–230.

R. A. Fisher (1935), *The logic of inductive inference*, J. Roy. Statist. Soc. **98**, 39–54.

——— (1936), *Uncertain inference*, Proc. Amer. Acad. Arts and Sciences **71**, 245–258.

——— (1956), *Statistical methods and scientific inference*, Oliver & Boyd, Edinburgh.

I. J. Good (1950), *Probability and the weighing of evidence*, Griffin, London.

I. Hacking (1975), *The emergence of probability*, Cambridge Univ. Press, Cambridge.

H. Ichihashi, H. Tanaka, and K. Asai (1988), *Fuzzy integrals based on pseudo-additions and multiplications*, J. Math. Anal. Appl. **130**, 354–364.

K. Popper and D. W. Miller (1987), *Why probabilistic support is not inductive*, Philos. Trans. Roy. Soc. London Ser. A **321**, 569–591.

G. L. S. Shackle (1961), *Decision, order and time in human affairs*, Cambridge Univ. Press, Cambridge.

G. Shafer (1976), *A mathematical theory of evidence*, Princeton Univ. Press, Princeton, NJ.

P. Smets (1982), *Discussion of 'Belief functions and parametric models'*, J. Roy. Statist. Soc. Ser. B **44** 343.

Z. Wang (1984), *Fuzzy measure and measure of fuzziness*, J. Math. Anal. Appl. **104**, 589–601.

L. A. Zadeh (1978), *Fuzzy sets as a basis for a theory of possibility*, Fuzzy Sets and Systems **1**, 3–28.

U.S. ARMY CONCEPT ANALYSIS AGENCY, BETHESDA, MARYLAND 20814-2797
E-mail address: `chen@caa.army.mil`