AD-A068 937 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
RANGES OF PRIOR MEASURES.(U)
JAN 79 L DEROBERTIS, J A HARTIGAN DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1915 NL

| OF |

AD
A068937

END
DATE
FILMED
7-79
DDC

LEVEL II (2)

MRC Technical Summary Report # 1915

RANGES OF PRIOR MEASURES

Lorraine DeRobertis and J. A. Hartigan

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

January 1979

Received November 21, 1978

Sponsored by

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

RANGES OF PRIOR MEASURES

Lorraine DeRobertis and J. A. Hartigan

Technical Summary Report #1915
January 1979

ABSTRACT

Partial prior knowledge is quantified by a range $R(L,U)$ of σ-finite prior measures Q satisfying $L(A) \le Q(A) \le U(A)$ for all measurable sets A, and is interpreted as a probability statement in a betting framework. The concept of conditional probability distributions is generalized to that of conditional measures, and Bayes theorem is extended to accommodate unbounded priors. According to Bayes theorem, the range $R(L,U)$ of prior measures is transformed upon observing X into a similar range $R(L_X,U_X)$ of posterior measures. Upper and lower expectations and variances induced by such ranges of measures are obtained. Under weak regularity conditions, these upper and lower posterior expectations are strongly consistent estimators. The range of posterior expectations of an arbitrary function b on the parameter space is asymptotically $b_N \pm \alpha\sigma_N + o(\sigma_N)$ where b_N and σ_N^2 are the posterior mean and variance of b induced by the upper prior measure U, and where α is a constant reflecting the uncertainty about the prior in terms of the derivative of L with respect to U.

AMS(MOS) Subject Classification: Primary 62A15; Secondary 60F15.

Key Words: Upper and lower measures, regular conditional measures, upper and lower expectations, upper and lower variances, strong consistency, approximate ranges of posterior expectations.

Work Unit No. 4 - Probability, Statistics and Combinatorics

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024. This material is based upon work supported by the National Science Foundation under Grant Nos. MCS78-09525 and MCS75-08374.

SIGNIFICANCE AND EXPLANATION

A typical example of the type of problem that can be handled by the methods discussed in this paper is the estimation of the mean and standard deviation of a normal distribution. The Bayesian method proceeds by first assuming a probability distribution for the unknown mean and standard deviation (the so-called prior). Then measurements are made in the system, giving sample information. Estimates of the unknown mean and standard deviation are made by combining the assumed prior and the sample information, producing an updated, or posterior, distribution for the parameters to be estimated (the mean and standard deviation in this case). Typically, the posterior distribution responds to increasing empirical evidence by concentrating about the true parameter values.

In practice, it is extremely demanding to require the experimenter to state a precise prior distribution for the parameters of interest; however, vague prior knowledge usually exists. Furthermore, Bayesian inference is sometimes criticized by the objective scientific community because of its dependence on subjective prior information. Thus, it is important that the sensitivity of Bayesian techniques to the prior be assessed, and that inferentially useful mathematical structures for quantifying vague prior knowledge be developed.

The purpose of this paper is to explore the inferential usefulness of specifying a range of priors. This is done in a way that is particularly easy to specify in practice and which leads to new insights into the robustness of Bayesian techniques.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

RANGES OF PRIOR MEASURES

Lorraine DeRobertis and J. A. Hartigan

1. Introduction

The personalistic Bayesian approach to statistical inference, as axiomatized by Ramsey (1926), de Finetti (1937), and Savage (1954), demonstrates that a coherent system of prior beliefs about a class of propositions must be consistent with a unique probability distribution on that class of propositions. Furthermore, coherent revision of prior opinion given empirical evidence must conform to Bayes theorem. In practice, prior knowledge is typically vague and any elicited prior distribution is only an approximation to the true one. Thus, less stringent modes of quantifying subjective information would facilitate the practicality of the Bayesian approach and would insure protection against the inferential errors that could result from the discrepancy between the elicited and true prior.

Keynes (1921) proposed that probability should be a partial ordering on pairs of propositions, induced by statements of the form "A given B is more probable than C given D." Koopman (1940) axiomatized Keynes' approach. If the observation space can be partitioned into events which are judged equally probable, Koopman developed a technique for assigning upper and lower numerical values to conditional probabilities. Good (1961) offered an axiom system for such upper and lower probabilities. Smith (1961) defined upper and lower probabilities to correspond to a range of bets which might be accepted for or against a proposition; any probability between the lower and upper probabilities is acceptable. For a continuous parameter, Smith suggested accepting any of a convex family of prior probability densities. Heath and Sudderth (1972) consider bets to be random variables on some sample space, and show that if a convex set of bounded bets contains no entirely negative bet, then there exists a finitely additive probability measure giving every bet in the set non-negative expectation; the convex set of such measures is the analogue of Smith's range of probabilities for a single event. Heath and Sudderth (1978) develop the relationship between families of acceptable bets and admissible decisions; if a decision is admissible, it must be worthwhile betting on it against other decisions.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024. This material is based upon work supported by the National Science Foundation under Grant Nos. MCS78-09525 and MCS75-08374.

Throughout the following, Θ will denote an arbitrary parameter space, and $\mathcal{B}$ a σ-field of subsets of Θ. A measure Q on $(\Theta,\mathcal{B})$ is a non-negative, countably-additive function on $\mathcal{B}$; Q is a probability measure if $Q(\Theta) = 1$. The analogue of lower and upper probabilities for a convex set C of probability measures is the collection of extreme points of the set; under suitable conditions, all probability measures in C will be expressible as mixtures of extreme probability measures. Furthermore, continuous linear functionals on C (i.e. expectations of bounded measurable functions of θ) are optimized at the extreme points of C. We consider here a convex set of measures (not necessarily probability measures) which is especially simple to specify without reference to extreme points, and on which many functionals of interest in statistical inference may be conveniently optimized.

For a given pair of σ-finite measures L, U on $(\Theta,\mathcal{B})$ satisfying $L(A) \leq U(A)$ for all $A \in \mathcal{B}$ (denoted $L \leq U$), let $R(L,U)$ be the convex set of measures Q satisfying $L \leq Q \leq U$. Since we require only that L and U be σ-finite, such useful prior measures as Lebesgue measure may be accommodated. The lower measure L and the upper measure U are direct generalizations of the lower and upper probabilities of Koopman and Smith. They are not excessively burdensome to specify, as would be a completely general convex family of measures. Bayes theorem "works" on the range $R(L,U)$ in that prior measures in $R(L,U)$ map into posterior measures ranging between those induced by L and U. Upper and lower posterior expectations, quantiles, and variances of arbitrary bets $b:\Theta \to R^1$ are obtained. And, finally, some attainable but not trivial asymptotic results are derived. Under quite weak conditions, the upper and lower posterior expectations of $b(\theta)$ are strongly consistent estimates of $b(\theta)$; furthermore, as the amount of sample information increases, the range of posterior expectations of $b(\theta)$ is approximately $b_N \pm \alpha\sigma_N$ where b_N, σ_N are the posterior expectation and standard deviation of $b(\theta)$ induced by U, and where the multiple α is determined by the precision of the prior range $R(L,U)$ of measures.

The "principle of stable estimation" put forward in Edwards, Lindman, and Savage (1963) and generalized by Dickey (1976) may be expressed in terms of ranges of prior measures.

For a real parameter θ , it is assumed that the prior density p satisfies $c \leq p(\theta) \leq (1+\beta)c$ for $\theta \in D$ and $p(\theta) \leq \gamma c$ for all θ . This is just the range of measures with lower measure $L(A) = c\mu(D \cap A)$ and upper measure $U(A) = (1+\beta)c\mu(D \cap A) + \gamma c\mu(D^c \cap A)$ where μ is Lebesque measure. Provided D has high posterior probability induced by μ , the posterior range of measures is then close to the posterior measure for a uniform prior.

2. Ranges of Measures

Given a pair Q_1, Q_2 of measures on $(\Theta,\mathcal{B})$, we say $Q_1 \leq Q_2$ if $Q_1(A) \leq Q_2(A)$ for all $A \in \mathcal{B}$. Let L and U be σ-finite measures on $(\Theta,\mathcal{B})$ satisfying $L \leq U$. The range of measures $R(L,U)$ consists of all measures Q such that $L \leq Q \leq U$. Equivalently, if L and U have densities ℓ and u with respect to some σ-finite measure ν on $(\Theta,\mathcal{B})$, then $R(L,U)$ consists of measures with densities q with respect to ν satisfying $\ell(\theta) \leq q(\theta) \leq u(\theta)$ $[\nu]$. If $Q \in R(L,U)$, then the odds ratio $Q(A)/Q(B)$ for $A,B \in \mathcal{B}$ is bounded by $U(A)/L(B)$ whenever that bound is well-defined. The measure L will be called the lower measure and the measure U the upper measure.

A real-valued, $\mathcal{B}$-measurable function b on Θ is called a bet, and $b(\theta)$ the payoff received if θ occurs. A probability statement is defined as the acceptance of some convex class of bets containing no strictly negative bet. For a measure Q on $(\Theta,\mathcal{B})$, let $Q(b)$ denote the integral of b with respect to Q. The measure Q, interpreted as a probability statement, is taken to mean the acceptance of all bets b for which $Q(|b|) < \infty$ and $Q(b) \geq 0$. Of course, Q_1 and Q_2 accept the same bets if and only if, for some constant $\alpha > 0$, $Q_1(A) = \alpha Q_2(A)$ for all $A \in \mathcal{B}$; thus proportional measures are equivalent probability statements.

A bet b is $R(L,U)$-nonnegative if $Q(|b|) < \infty$ and $Q(b) \geq 0$ for every Q in $R(L,U)$. The range of measures $R(L,U)$, interpreted as a probability statement, is taken to mean the acceptance of all $R(L,U)$-nonnegative bets. The class of all measures proportional to some member of $R(L,U)$ is, thus, equivalent to the probability statement $R(L,U)$. For any bet b, define $b^+(\theta) = b(\theta)$ if $b(\theta) \geq 0$, $b^+(\theta) = 0$ if $b(\theta) < 0$, and define $b^-(\theta) = b(\theta) - b^+(\theta)$. Since $Q(b) = Q(b^+) + Q(b^-) \geq L(b^+) + U(b^-)$ for all Q in $R(L,U)$, we see that a bet b is $R(L,U)$-nonnegative if and only if $U(|b|) < \infty$ and $L(b^+) + U(b^-) \geq 0$. A bet b is $R(L,U)$-positive if $L(b^+) + U(b^-) > 0$.

3. Ranges of Posterior Measures

Bayes Theorem states that if $\{P_\theta | \theta \in \Theta\}$ is a family of probability measures defined on a σ-field F of subsets of a sample space X, where P_θ has density $f(X|\theta)$ with respect to a σ-finite measure ν on (X, F), and if Q_0 is a probability measure on (Θ, B), then the unique probability measure on $(X \times \Theta, F \times B)$ having θ-marginal measure Q_0 and regular conditional distribution $\{P_\theta | \theta \in \Theta\}$ given B is the probability measure Q with density $f(X|\theta)$ with respect to $\nu \times Q_0$. Furthermore, the regular conditional distribution of Q given F has density $f(X|\theta)/\int f(X|\theta)dQ_0(\theta)$ with respect to Q_0. It is straightforward to extend Bayes Theorem to the class of improper σ-finite priors Q_0 for which the induced measure Q has σ-finite X-marginal measure, as, for example, in the case where $X|_\theta \sim N(\theta,1)$ and Q_0 is the uniform prior. However, if the X-marginal of Q is not σ-finite then it is possible that $\nu\{X \in X | \int f(X|\theta)dQ_0(\theta) = \infty\}$ is positive, so that a conditional probability measure for θ given X cannot be defined.

For example, if $P_\theta(X=1) = \theta$, $P_\theta(X=0) = 1-\theta$, and $dQ_0/d\theta = \theta^{-1}(1-\theta)^{-1}$, $0 \leq \theta \leq 1$, then $\int f(X|\theta)dQ_0(\theta) = \int \theta^{X-1}(1-\theta)^{-X}d\theta = \infty$ for all X. Thus, the posterior probability density "prescribed" by Bayes Theorem equals zero if $0 < \theta < 1$ and is indeterminate at $\theta = 0$ and $\theta = 1$. Nevertheless, the measure $dQ_X(\theta) \equiv \theta^{X-1}(1-\theta)^{-X}d\theta$ is a probability statement about θ according to which, given X, all bets $b(\theta)$ such that $\int b(\theta)dQ_X(\theta) \geq 0$ are acceptable.

To accommodate all σ-finite priors we will generalize the concept of a regular conditional probability measure to that of a regular conditional measure which will be interpreted as a conditional probability statement.

Definition 3.1. Let Q be a measure on (Ω, S) and let D be a sub-σ-field of S. A family $\{Q_\omega | \omega \in \Omega\}$ of measures on (Ω, S) is a regular conditional measure given D if for every S-measurable, Q-integrable function b

i) $Q_\omega(b)$ is D-measurable and finite $[Q]$;

ii) $Q_\omega(gb) = g(\omega)Q_\omega(b)$ $[Q]$ for every bounded D-measurable function g;

iii) $Q_\omega(b) = 0$ $[Q]$ implies $Q(b) = 0$.

The existence of such a family requires regularity conditions similar to those guaranteeing the existence of a regular conditional probability measure (see Renyi (1970)).

If $Q(\Omega) = 1$ and $Q_\omega(\Omega) = 1$ [Q] then $Q(gb) = Q(gQ_{(\cdot)}(b))$ for all bounded $\mathcal{D}$-measurable g and Q-integrable b ; thus $\{Q_\omega | \omega \in \Omega\}$ is a regular conditional probability given $\mathcal{D}$. In particular, $Q_{(\cdot)}(b)$ is [Q]-unique. For unbounded measures Q we prove the following:

Lemma 3.2. Assume Q is σ-finite on $\mathcal{S}$, and let $\{Q_\omega|\omega \in \Omega\}$, $\{Q'_\omega|\omega \in \Omega\}$ be regular conditional measures for Q given $\mathcal{D} \subset \mathcal{S}$. Then there exists a $\mathcal{D}$-measurable function k such that $k(\omega) > 0$ for every $\omega \in \Omega$ and $Q_\omega(b) = k(\omega)Q'_\omega(b)$ [Q] for every Q-integrable b.

Proof. Q σ-finite on $\mathcal{S}$ is equivalent to the existence of an $\mathcal{S}$-measurable, Q-integrable function b_0 such that $b_0(\omega) > 0$ for every $\omega \in \Omega$. Since Q_ω, Q'_ω are countably additive, $Q_\omega(b_0) > 0$ and $Q'_\omega(b_0) > 0$ for all $\omega \in \Omega$. Define $k(\omega) = Q_\omega(b_0)/Q'_\omega(b_0)$. Let b be any non-negative, $\mathcal{S}$-measurable, Q-integrable function, and let $C(\omega) = Q_\omega(b)/Q_\omega(b_0)$, $C'(\omega) = Q'_\omega(b)/Q'_\omega(b_0)$. Define $C_M(\omega) = 0$ if $C(\omega) > M$, $C_M(\omega) = C(\omega)$ if $C(\omega) \leq M$. Then $Q_\omega(C_M b_0 - b) = C_M(\omega)Q_\omega(b_0) - Q_\omega(b)$ [Q]. Letting $M \uparrow \infty$, we find by the monotone convergence theorem that $Q_\omega(C\, b_0 - b) = 0$ [Q]. Hence, $Q(C\, b_0 - b) = 0$; similarly, $Q(C'b_0 - b) = 0$. Thus, $Q(b_0(C-C')) = 0$. Let $A = \{\omega | C(\omega) > C'(\omega)\}$. Applying the above argument to the function $1_A b$, and noting that $Q_\omega(1_A b) = 1_A(\omega)Q_\omega(b)$ and $Q'_\omega(1_A b) = 1_A(\omega)Q'_\omega(b)$ since $A \in \mathcal{D}$, we find that $Q(b_0 1_A(C-C')) = 0$. But $b_0(C-C')$ is strictly positive on A ; hence, $Q(A) = 0$. Similarly, $Q\{\omega|C(\omega) < C'(\omega)\} = 0$, and so $C(\omega) = C'(\omega)$ [Q]. We have shown, then, that $Q_\omega(b) = k(\omega)Q'_\omega(b)$ [Q] for any nonnegative $\mathcal{S}$-measurable Q-integrable b. For arbitrary $\mathcal{S}$-measurable Q-integrable $b = b^+ + b^-$, the desired equality holds since it holds for each of b^+ and b^- . ∎

A regular conditional measure $\{Q_\omega | \omega \in \Omega\}$, interpreted as a <u>conditional probability statement</u>, is taken to mean the acceptance of all Q-integrable bets b for which $Q_\omega(b) \geq 0$ [Q]. By Lemma 3.2, conditional probability statements are unique.

<u>Bayes theorem for unbounded priors</u>:

Let $(\mathcal{X}, \mathcal{F})$ be a sample space, and let $\{P_\theta | \theta \in \Theta\}$ be a family of probability measures on $(\mathcal{X}, \mathcal{F})$ with densities $f(x|\theta)$ with respect to a σ-finite measure ν on $(\mathcal{X}, \mathcal{F})$. Let Q_0 be a σ-finite prior measure on $(\Theta, \mathcal{B})$.

Then there exists a unique measure Q on $(\mathcal{X} \times \Theta, F \times B)$ such that Q agrees with Q_0 on B and such that $\{P_\theta | \theta \in \Theta\}$ is a regular conditional measure of Q given B. Q is σ-finite on $F \times B$. A regular conditional measure of Q given F is $\{Q_X | X \in \mathcal{X}\}$ where $Q_X(b) \equiv \int_\Theta b(X,\theta) f(X|\theta) dQ_0(\theta)$.

Proof. Let Q have density $f(X|\theta)$ with respect to $\nu \times Q_0$. Tonelli's Theorem implies that for every $B \in B$, $Q(\mathcal{X} \times B) = \int_{\theta \in B} \int_{\mathcal{X}} f(X|\theta) d\nu(X) dQ_0(\theta) = Q_0(B)$. Since Q_0 is σ-finite on B, Q is σ-finite on $F \times B$.

Viewing B as a sub-σ-field of $F \times B$ in the obvious way, we next show that $\{P_\theta | \theta \in \Theta\}$ is a regular conditional measure of Q given B in the sense that the family of measures $\{Q_{X,\theta} | X \in \mathcal{X}, \theta \in \Theta\}$ defined by $Q_{X,\theta}(A \times B) \equiv P_\theta(A) 1_B(\theta)$ for $A \in F, B \in B$ satisfies conditions i, ii, iii of Definition 3.1. Let b be any $F \times B$-measurable, Q-integrable function. Then $Q_{X,\theta}(b) = \int_{\mathcal{X}} b(t,\theta) f(t|\theta) d\nu(t) \equiv P_\theta(b)$ is B-measurable by Fubini's Theorem. If g is any bounded B-measurable on $\mathcal{X} \times \Theta$ then, without loss of generality, g is a function only of θ and so $P_\theta(bg) = \int_{\mathcal{X}} b(t,\theta) g(\theta) f(t|\theta) d\nu(t) = g(\theta) P_\theta(b)$. Finally, if $P_\theta(b) = 0$ $[Q]$ then $P_\theta(b) = 0$ $[Q_0]$ and, by Fubini's Theorem, $Q(b) = \int_\Theta P_\theta(b) dQ_0(\theta) = 0$.

Now, suppose Q' is a measure on $(\mathcal{X} \times \Theta, F \times B)$ which agrees with Q_0 on B and for which $\{P_\theta | \theta \in \Theta\}$ is a regular conditional measure given B. Then Q' is σ-finite, as is Q. If $C \in F \times B$ and $Q(C)$, $Q'(C)$ are both finite then $Q(C) = Q(P_\theta(1_C)) = Q'(P_\theta(1_C)) = Q'(C)$. Since Q and Q' are both σ-finite, it follows that $Q = Q'$.

Finally, we show that $\{Q_X | X \in \mathcal{X}\}$ is a regular conditional measure of Q given F, in the sense that the family of measures $\{Q'_{X,\theta} | X \in \mathcal{X}, \theta \in \Theta\}$ defined by $Q'_{X,\theta}(A \times B) = 1_A(X) \int_B f(X|t) dQ_0(t)$ for $A \in F, B \in B$ satisfies conditions i,iii,iii. Fubini's Theorem implies that $Q_X(b)$ is F-measurable for every Q-integrable b. And obviously $Q_X(bg) = g(X) Q_X(b)$ for any bounded F-measurable g. Lastly, suppose $Q_X(b) = 0$ $[Q]$. Let $A = \{Q_X(b) = 0\}$. By Fubini's Theorem, $Q(b) = Q(1_A b) = \int 1_A(X) Q_X(b) d\nu(X) = 0$. ∎

The measures Q_X, $X \in \mathcal{X}$ are called posterior measures, given the observation X, induced by the prior measure Q_0. We interpret unbounded measures to make probability statements as follows. For the prior Q_0, the $\mathcal{B}$-measurable bet b is *acceptable* whenever b is Q_0-integrable and $Q_0(b) \geq 0$. Bets b such that b and $-b$ are acceptable are fair bets. For the *conditional measures* P_θ, any $\mathcal{F} \times \mathcal{B}$-measurable bets b such that $P_\theta(b) \geq 0$ $[Q_0]$ are acceptable. Any bet b is the sum of the P_θ-fair bet $b(\theta)-P_\theta(b)$ and the $\mathcal{B}$-measurable bet $P_\theta(b)$. Bayes Theorem states that there exists a unique measure Q on $(\mathcal{X} \times \Theta, \mathcal{F} \times \mathcal{B})$ such that b is Q-acceptable (i.e. $Q(b) \geq 0$) if and only if b is the sum of a P_θ-fair bet and a Q_0-acceptable bet. A subset of these Q-acceptable bets consists of all bets acceptable according to the posterior measures Q_X, $X \in \mathcal{X}$, namely, all bets b such that $Q_X(b) \geq 0$ $[Q]$. By Lemma 3.2, the set of posterior acceptable bets is uniquely determined by Q_0 and $\{P_\theta | \theta \in \Theta\}$. See also Freedman and Purves (1969).

If $Q_0 \in R(L,U)$, then $Q_X \in R(L_X,U_X)$. For if $L(A) \leq Q_0(A) \leq U(A)$ for all $A \in \mathcal{B}$ then it follows, since $f(X|\theta) \geq 0$, that $L_X(A) \leq Q_X(A) \leq U_X(A)$ for all $A \in \mathcal{B}$. Thus, Bayes Theorem "works" for ranges of prior measures in that a range of prior measures is transformed given X into a range of posterior measures. Indeed, $dL/dU = dL_X/dU_X$, so that the density dL/dU, which indicates how far apart the lower and upper measures are, is unaffected by the *observation* X.

4. Bounds on Integral Ratios

For the $\mathcal{B}$-measurable, U-integrable bets b and c, with c $R(L,U)$-positive, consider the range of integral ratios $Q(b)/Q(c)$ for $Q \in R(L,U)$. Of particular interest is the bet $c(\theta) = f(X|\theta)$ for fixed X, corresponding to posterior expectations.

Theorem 4.1. Let λ_1 and λ_2 be the infimum and supremum, respectively, of $Q(b)/Q(c)$ for $Q \in R(L,U)$. For $\lambda \in R^1$, define $J_1(\lambda) = U(b-\lambda c)^- + L(b-\lambda c)^+$ and $J_2(\lambda) = U(b-\lambda c)^+ + L(b-\lambda c)^-$. Then $J_i(\lambda) < 0$ if and only if $\lambda_i < \lambda$; and $J_i(\lambda) > 0$ if and only if $\lambda_i > \lambda$. Thus λ_i is the unique solution of $J_i(\lambda) = 0$.

Proof. Let $c_1 = \inf\{Q(c)|Q \in R(L,U)\}$, $c_2 = \sup\{Q(c)|Q \in R(L,U)\}$. Since c is U-integrable and $R(L,U)$-positive, c_1 and c_2 are finite and positive. Furthermore, $|Q(b)/Q(c)| \leq U(|b|)/c_1 < \infty$ for all $Q \in R(L,U)$, so that λ_1 and λ_2 are finite.

Since $J_1(\lambda) = \inf\{Q(b-\lambda c)|Q \in R(L,U)\}$, $\lambda \leq \lambda_1$ if and only if $0 \leq J_1(\lambda)$. Moreover, for any $\varepsilon > 0$, $\lambda + \varepsilon/c_1 \leq \lambda_1$ implies $\varepsilon \leq J_1(\lambda)$ which in turn implies $\lambda + \varepsilon/c_2 \leq \lambda_1$. Thus, $\lambda < \lambda_1$ if and only if $0 < J_1(\lambda)$, $\lambda > \lambda_1$ if and only if $0 > J_1(\lambda)$, and so λ_1 is the unique solution of $J_1(\lambda) = 0$. The result for λ_2 follows similarly. ∎

In particular, if $U(\Theta) < \infty$ then for any $A \in \mathcal{B}$, $\sup\{Q(A)/Q(\Theta)|Q \in R(L,U)\} = U(A)/(U(A) + L(A^c))$ and $\inf\{Q(A)/Q(\Theta)|Q \in R(L,U)\} = L(A)/(L(A) + U(A^c))$. For any $\mathcal{B}$-measurable bet b, we define the lower and upper distribution functions of b by

$$F_1(t) = \inf\{Q(A_t)/Q(\Theta)|Q \in R(L,U)\} \quad \text{and}$$

$$F_2(t) = \sup\{Q(A_t)/Q(\Theta)|Q \in R(L,U)\}$$

where $A_t = \{b(\theta) \leq t\}$. If $U(A_t)$ is continuous and strictly increasing in t, then, for $0 < \alpha < 1$, the lower and upper α-percent points of b are the unique solutions of $F_2(t) = \alpha$ and $F_1(t) = \alpha$, respectively. In the special case where $U = kL$ for some $k > 1$, these are simply the $\alpha/(\alpha+k-\alpha k)$ and $\alpha k/(1+\alpha k-\alpha)$ percent points of b with respect to L.

Example 4.2. Let $\Theta = R^1$ and suppose that L is Lebesgue measure and $U = kL$ for some constant $k > 1$; we are supposing, then, that the prior measure of any set does not exceed k times the prior measure of any set of the same Lebesgue measure. If $X \sim N(\theta,\sigma_0^2)$ given θ, with σ_0 known, then the posterior measure range is $R(L_X,U_X)$

where L_X has density $f(X|\theta) = (\sqrt{2\pi}\sigma_0)^{-1}\exp[-\frac{1}{2}(X-\theta)^2/\sigma_0^2]$ with respect to Lebesgue measure and U_X has density $k\, f(X|\theta)$.

The upper and lower posterior distribution functions of θ are

$$F_2(t) = \frac{k\,\Phi\,[(t-X)/\sigma_0]}{1 + (k-1)\Phi[(t-X)/\sigma_0]}$$

and

$$F_1(t) = \frac{\Phi[(t-X)/\sigma_0]}{k - (k-1)\Phi[(t-X)/\sigma_0]}, \quad t \in R^1,$$

where Φ is the standard normal cdf. For $0 < \alpha < 1$, the upper and lower posterior α-percent points of θ are, respectively, the $\alpha k/(1+\alpha k-\alpha)$ and $\alpha/(\alpha+k-\alpha k)$ percent points of $N(X,\sigma_0^2)$.

The posterior mean $Q(\theta f(X|\theta))/Q(f(X|\theta))$, $Q \in R(L,U)$, has minimum value satisfying $U_X(\theta-\lambda)^- + L_X(\theta-\lambda)^+ = 0$, and maximum value satisfying $U_X(\theta-\lambda)^+ + L_X(\theta-\lambda)^- = 0$. It is easily seen that this range of posterior means is $[X-\sigma_0\gamma(k),\ X+\sigma_0\gamma(k)]$ where $\gamma(k)$ satisfies

$$k\gamma = (k-1)[\phi(\gamma) + \gamma\Phi(\gamma)]$$

and ϕ,Φ are the standard normal density and cdf. Table 1 displays values of $\gamma(k)$ for $1 \le k \le 10$. It is seen that a substantial amount of variation in the prior has only a minor effect on the posterior mean of θ, as compared to the variability of the estimate X due to the data.

Table 1

k	1	1.25	1.50	1.75	2	2.5	3	4	5	6	7	8	9	10
γ(k)	0	.089	.162	.223	.276	.364	.436	.549	.636	.707	.766	.817	.862	.901

5. Bounds on Variances

For a $\mathcal{B}$-measurable bet b on Θ, we will characterize the range of the variance of b, $V_Q(b) \equiv Q(|b|^2)/Q(\Theta) - |Q(b)/Q(\Theta)|^2$, for $Q \in R(L,U)$ when $U(|b|^2) < \infty$ and $U(\Theta) < \infty$. We note first that without loss of generality we may assume $(\Theta, \mathcal{B}) = (R^1, \mathcal{B}_0)$, where $\mathcal{B}_0$ is the Borel σ-field on R^1, and $b(\theta) = \theta$. For, let $\mathcal{B}(b)$ denote the sub-σ-field of $\mathcal{B}$ generated by $\{b^{-1}(A) | A \in \mathcal{B}_0\}$; and let $R_b(L,U)$ be the family of measures Q on $(\Theta, \mathcal{B}(b))$ satisfying $L(B) \le Q(B) \le U(B)$ for all $B \in \mathcal{B}(b)$. Clearly, $R(L,U) \subset R_b(L,U)$. Now, the mapping $Q \to Q^*$, where $Q^*(A) = Q(b^{-1}(A))$ for $A \in \mathcal{B}_0$, defines a one-to-one correspondence between $R(L^*,U^*)$ and $R_b(L,U)$ with respect to which V_Q is invariant. Furthermore, if $Q \in R_b(L,U)$ then there exists some $\mathcal{B}(b)$-measurable function α such that $0 \le \alpha(\theta) \le 1$ [U] and $Q(B) = L(\alpha 1_B) + U((1-\alpha)1_B)$ for every $B \in \mathcal{B}(b)$. Defining $Q'(C) = L(\alpha 1_C) + U((1-\alpha)1_C)$ for $C \in \mathcal{B}$, we see that $Q' \in R(L,U)$ and $V_{Q'}(b) = V_Q(b)$. Hence, the range of $V_Q(b)$, $Q \in R(L,U)$ equals the range of $\int |t-a|^2 dQ^*(t)/\int dQ^*(t)$, $Q^* \in R(L^*,U^*)$ where $a = \int t\, dQ^*(t)/\int dQ^*(t)$. Throughout this section, then, we will assume without loss of generality that $(\Theta, \mathcal{B}) = (R^1, \mathcal{B}_0)$ and $b(\theta) = \theta$.

The following notation will prove useful. For $-\infty < a < \infty$, define $\sigma_1^2(a) = \inf\{Q(|\theta-a|^2)/Q(\Theta) | Q \in R(L,U)\}$ and $\sigma_2^2(a) = \sup\{Q(|\theta-a|^2)/Q(\Theta) | Q \in R(L,U)\}$. And for $\sigma > 0$ let $J_1(a,\sigma) = U((|\theta-a|^2 - \sigma^2)^-) + L((|\theta-a|^2-\sigma^2)^+)$ and $J_2(a,\sigma) = L((|\theta-a|^2-\sigma^2)^-) + U((|\theta-a|^2-\sigma^2)^+)$. By Theorem 4.1, $\sigma_i(a)$ is the unique nonnegative solution of $J_i(a,\sigma) = 0$.

Next, for $0 \le \lambda \le 1$ define $Q^1_{\lambda,a} \in R(L,U)$ by

$$Q^1_{\lambda,a}(A) = \begin{cases} U(A) & \text{if} \quad A \subset \{|\theta-a| < \sigma_1(a)\} \\ L(A) & \text{if} \quad A \subset \{|\theta-a| > \sigma_1(a)\} \\ \lambda L(A)+(1-\lambda)U(A) & \text{if} \quad A = \{a-\sigma_1(a)\} \\ \lambda U(A)+(1-\lambda)L(A) & \text{if} \quad A = \{a+\sigma_1(a)\}. \end{cases}$$

Define $Q^2_{\lambda,a}$ similarly, with the roles of L and U reversed and with $\sigma_1(a)$ replaced by $\sigma_2(a)$. Note that

$$\sigma_i^2(a) = Q^i_{\lambda,a}(|\theta-a|^2)/Q^i_{\lambda,a}(\Theta)$$

for $i = 1,2$ and for all $0 \le \lambda \le 1$. Furthermore, if a satisfies $Q^i_{\lambda,a}(\theta-a) = 0$ for some $0 \le \lambda \le 1$ then $\sigma_i^2(a)$ is the variance of θ with respect to $Q^i_{\lambda,a}$.

Theorem 5.1. Let $A_1 = \{a | Q^1_{\lambda,a}(\theta-a) = 0$ for all $0 \le \lambda \le 1\}$ and $A_2 = \{a | Q^2_{\lambda,a}(\theta-a) = 0$ for some $0 \le \lambda \le 1\}$. Then A_1 and A_2 are nonempty and

i) $\inf\{V_Q(\theta) | Q \in R(L,U)\} = \inf\{\sigma_1^2(a) | a \in A_1\}$

ii) $\sup\{V_Q(\theta) | Q \in R(L,U)\} = \inf\{\sigma_2^2(a) | a \in A_2\}$.

Remark. In general, A_1 and A_2 may contain more than one value. However, (ii) implies that $\sigma_2^2(a)$ is constant on A_2.

Proof of Theorem 5.1.

It is readily seen that σ_1 and σ_2 are continuous and that $\lim_{a\to\pm\infty} \sigma_i(a) = \infty$; thus $\inf\{\sigma_1(a) | -\infty < a < \infty\}$ and $\inf\{\sigma_2(a) | -\infty < a < \infty\}$ are, in fact, attained. We now show that a minimizes $\sigma_i(a)$ only if $a \in A_i$, from which it follows that A_1 and A_2 are nonempty.

Consider, first, σ_2. If a minimizes σ_2 then $J_2(a \pm \delta, \sigma_2(a)) \ge 0$ for all $\delta > 0$. Let

$$G_2(a,\sigma) = \int_{|\theta-a|<\sigma} (\theta-a)dL(\theta) + \int_{|\theta-a|>\sigma} (\theta-a)dU(\theta) .$$

It is straightforward to show that for all $-\infty < a < \infty$ and $0 < \delta < \sigma$,

$$J_2(a+\delta,\sigma) = J_2(a,\sigma) - 2\delta\{G_2(a,\sigma) + \sigma[L(a+\sigma) - U(a-\sigma)] + o_\delta(1)\}$$

and

$$J_2(a-\delta,\sigma) = J_2(a,\sigma) + 2\delta\{G_2(a,\sigma) + \sigma[U(a+\sigma) - L(a-\sigma)] + o_\delta(1)\}$$

where $o_\delta(1) \to 0$ as $\delta \downarrow 0$. Since $J_2(a,\sigma_2(a)) = 0$ we see that a minimizes $\sigma_2(a)$ only if

$$\sigma_2(a)[L(a-\sigma_2(a)) - U(a+\sigma_2(a))] \le G_2(a,\sigma_2(a)) \le \sigma_2(a)[U(a-\sigma_2(a)) - L(a+\sigma_2(a))]$$

or, equivalently, only if there exists some $0 \le \lambda \le 1$ such that $Q^2_{\lambda,a}(\theta-a) = 0$ (i.e. only if $a \in A_2$). Thus, $A_2 \ne \phi$; and $\inf\{\sigma_2^2(a) | -\infty < a < \infty\} = \inf\{\sigma_2^2(a) | a \in A_2\}$.

Next, consider $\sigma_1(a)$; define

$$G_1(a,\sigma) = \int_{|\theta-a|<\sigma} (\theta-a)dU(\theta) + \int_{|\theta-a|>\sigma} (\theta-a)dL(\theta) .$$

An analysis similar to that for $\sigma_2(a)$ demonstrates that a minimizes $\sigma_1(a)$ only if

$$\sigma_1(a)[U(a-\sigma_1(a)) - L(a+\sigma_1(a))] \le G_1(a,\sigma_1(a)) \le \sigma_1(a)[L(a-\sigma_1(a)) - U(a+\sigma_1(a))].$$

Since $L \le U$, this is equivalent to $Q^1_{\lambda,a}(\theta-a) = 0$ for all $0 \le \lambda \le 1$. Thus, $A_1 \ne \phi$; and $\inf\{\sigma_1^2(a) \mid -\infty < a < \infty\} = \inf\{\sigma_1^2(a) \mid a \in A_1\}$.

Now,

$$\inf_{Q\in R(L,U)} V_Q(\theta) = \inf_{Q\in R(L,U)} \inf_{-\infty < a < \infty} Q(|\theta-a|^2)/Q(\Theta) =$$

$$\inf_{-\infty < a < \infty} \inf_{Q\in R(L,U)} Q(|\theta-a|^2)/Q(\Theta) = \inf_{-\infty < a < \infty} \sigma_1^2(a) = \inf_{a\in A_1} \sigma_1^2(a)$$

which proves (i). And

$$\sup_{Q\in R(L,U)} V_Q(\theta) = \sup_{Q\in R(L,U)} \inf_{-\infty < a < \infty} Q(|\theta-a|^2)/Q(\Theta) \le$$

$$\inf_{-\infty < a < \infty} \sup_{Q\in R(L,U)} Q(|\theta-a|^2)/Q(\Theta) =$$

$$\inf_{-\infty < a < \infty} \sigma_2^2(a) = \inf_{a\in A_2} \sigma_2^2(a) \le \sup_{Q\in R(L,U)} V_Q(\theta)$$

where the last inequality follows from the fact that $a \in A_2$ implies $\sigma_2^2(a) = V_{Q^2_{\lambda,a}}(\theta)$ for some $0 \le \lambda \le 1$; thus statement (ii) is proved. ∎

Example 5.2. Consider the normal location problem of Example 4.2, and let V_X^1, V_X^2 denote the minimum and maximum posterior variance of θ as the prior ranges over $R(L,U)$. Applying Theorem 5.1 to $R(L_X,U_X)$, and noting that $A_1 = A_2 = \{X\}$ in this case, we find that V_X^1 and V_X^2 are, respectively, the unique solutions of $U_X((|\theta-X|^2-V)^-) + L_X((|\theta-X|^2-V)^+) = 0$ and $L_X((|\theta-X|^2-V)^-) + U_X((|\theta-X|^2-V)^+) = 0$. Defining $H(\Delta,c) = \Delta\phi(\Delta) + (\Delta^2-1)[\Phi(\Delta)-c]$, it is readily seen that $V_X^1 = \Delta_1^2\sigma_0^2$ where Δ_1 is the unique solution of $H(\Delta,c_1) = 0$ with $c_1 = (k-2)/(2k-2)$. And $V_X^2 = \Delta_2^2\sigma_0^2$ where Δ_2 is the unique solution of $H(\Delta,c_2) = 0$ with $c_2 = (2k-1)/(2k-2)$. Table 2 gives values of Δ_1 and Δ_2 for $1 \le k \le 10$. We see, for example, that if the prior varies about the uniform prior by up to a factor of two, the posterior standard deviation is affected by no more than 17% of the standard error σ_0 of the data.

Table 2

k	Δ_1	Δ_2	k	Δ_1	Δ_2
1	1	1	4	.697	1.360
1.25	.947	1.055	5	.654	1.421
1.50	.904	1.100	6	.621	1.472
1.75	.870	1.140	7	.574	1.515
2	.840	1.174	8	.592	1.552
2.5	.792	1.233	9	.552	1.585
3	.754	1.282	10	.535	1.615

Example 5.3. The mixture parameter λ of Theorem 5.1 is, of course, relevant only if U is not continuous. In the discrete case, the functions σ_1^2 and σ_2^2 are easily minimized since they are piecewise quadratic. For example, if $\Theta = \{\theta_1, \theta_2\}$ with $\theta_1 < \theta_2$, and if we denote $L_i = L(\theta_i)$, $U_i = U(\theta_i)$, $i = 1,2$, then $\sigma_2^2(a) = (|\theta_1 - a|^2 L_1 + |\theta_2 - a|^2 U_2)/(L_1 + U_2)$ if $a \leq (\theta_1+\theta_2)/2$, and $\sigma_2^2(a) = (|\theta_1 - a|^2 U_1 + |\theta_2 - a|^2 L_2)/(U_1 + L_2)$ if $a \geq (\theta_1+\theta_2)/2$. If $L_1 \leq U_2$ and $L_2 \leq U_1$, we find $\min \sigma_2^2(a) = (\theta_2 - \theta_1)^2/4$, attained at $a = (\theta_1+\theta_2)/2$. These are the variance and mean of the uniform measure $Q^2_{\lambda,a} \in R(L,U)$, where $a = (\theta_1+\theta_2)/2$ and $\lambda = (U_2 - L_1)/(U_1 + U_2 - L_1 - L_2)$.

6. Asymptotic Behavior of Ranges of Posterior Expectations.

Let the observation X_i have sample space (X_i, F_i) $i = 1,2,\ldots,$ and define $X_\infty = \prod_1^\infty X_i$, $F_\infty = \prod_1^\infty F_i$. Suppose $\{P_\theta | \theta \in \Theta\}$ is a family of probability distribution on (X_∞, F_∞) such that $P_\theta(A)$ is B-measurable for all $A \in F_\infty$. Let Q be any measure on (Θ, B) and let $Y, Y_1, Y_2, \ldots$ be any sequence of random variables on $(X_\infty \times \Theta, F_\infty \times B)$. The notation $Y = 0$ [Q] will mean that, except for some set of θ-values of Q-measure zero, $P_\theta\{\underset{\sim}{X} \in X_\infty | Y(\underset{\sim}{X}, \theta) = 0\} = 1$. Similarly, the notation $Y_N \to Y$ [Q] will mean that, except for some set of θ-values of Q-measure zero, $Y_N(\underset{\sim}{X}, \theta)$ converges to $Y(\underset{\sim}{X}, \theta)$ with P_θ-probability one.

Assume further that, for all θ, the $(X_1, \ldots, X_N)$-marginal distribution of P_θ has density $f(X_1, \ldots, X_N | \theta)$ with respect to some σ-finite measure ν_N on $(X_1 \times \ldots \times X_N, F_1 \times \ldots \times F_N)$. For a B-measurable bet b and a measure Q on (Θ, B), we will denote $Q_N(b) = \int b(\theta) f(X_1, \ldots, X_N | \theta) dQ(\theta)$; and for $A \in B$, define $Q_N(A) = Q_N(1_A)$.

If Q is a probability measure and $Q(|b|) < \infty$, the Martingale Convergence Theorem (see, for example, Doob (1949)) implies that $Q_N(b)/Q_N(1) \to b(\theta)$ [Q] for any F_∞-measurable bet b. In particular $\lim_{M\to\infty} \overline{\lim_N} Q_N(|b| 1_{|b|>M})/Q_N(1) = 0$ [Q]; and $Q_N(A)/Q_N(1) \to 1_A(\theta)$ [Q] if $1_A(\theta)$ is F_∞-measurable. We now show that these conditions are essentially sufficient to extend the martingale result to σ-finite measures Q.

Lemma 6.1. Assume Q is a σ-finite measure on (Θ, B); let $\Theta = \bigcup_1^\infty A_i$ with $A_i \in B$ disjoint and $0 < Q(A_i) < \infty$. Suppose

i) $b(\theta)$ is F_∞-measurable;

ii) $Q_N(1) < \infty$ $[\nu_N]$ for sufficiently large N;

iii) $Q_N(A_i)/Q_N(1) \to 1_{A_i}(\theta)$ [Q] for $i = 1,2,\ldots$;

iv) $\lim_{M\to\infty} \overline{\lim_N} Q_N(|b| 1_{|b|>M})/Q_N(1) = 0$ [Q].

Then $Q_N(b)/Q_N(1) \to b(\theta)$ [Q].

Proof. Note, first, that if $A \in B$, $0 < Q(A) < \infty$, and $Q(|b 1_A|) < \infty$ then, by the Martingale Convergence Theorem applied to the probability measure $Q_A(D) \equiv Q(A \cap D)/Q(A)$ for $D \in B$, it follows that $Q_N(b 1_A)/Q_N(A)$ converges to $b(\theta)$ with P_θ-probability one for all $\theta \in A_0 \subset A$ where $Q(A - A_0) = 0$.

Now, suppose $|b(\theta)| \leq M$ for all θ. Then $Q(|b1_{A_i}|) < \infty$ for every i. For any fixed $\theta \in A_i$ such that $Q_N(b1_{A_i})/Q_N(A_i) \to b(\theta)$ and $Q_N(A_i)/Q_N(1) \to 1$ with P_θ-probability one, it follows that

$$|Q_N(b1_{A_i^c})/Q_N(1)| \leq M\ |1 - Q_N(A_i)/Q_N(1)| \to 0$$

and so also that $Q_N(b)/Q_N(1) \to b(\theta)$ with P_θ-probability one. By (iii), then, we obtain $Q_N(b)/Q_N(1) \to b(\theta)$ $[Q]$.

For unbounded b, define $b_M(\theta) = b(\theta)$ if $|b(\theta)| \leq M$ and $b_M(\theta) = 0$ otherwise. Then $P_\theta(Q_N(b_M)/Q_N(1) \to b_M(\theta)$ for all $M = 1,2,\ldots) = 1$ $[Q]$. Fix θ; then for all $M > |b(\theta)|$,

$$|Q_N(b)/Q_N(1) - b(\theta)| \leq |Q_N(b_M)/Q_N(1) - b_M(\theta)| + Q_N(|b|1_{|b|>M})/Q_N(1) .$$

Thus, (iv) implies that $Q_N(b)/Q_N(1) \to b(\theta)$ $[Q]$. ∎

Remark. Clearly, assumption (i) may be replaced by the requirement that $b(\theta)$ be equivalent to some F_∞-measurable function $h(\underset{\sim}{X})$ in the sense that $P_\theta(h(\underset{\sim}{X}) = b(\theta)) = 1$ $[Q]$. If $b(\theta)$ is estimable in that $t_N(X_1,\ldots,X_N) \to b(\theta)$ $[Q]$ for some sequence of estimators t_N, then $b(\theta)$ is equivalent to the F_∞-measurable function $\lim_{N\to\infty} t_N(X_1,\ldots,X_N)$.

If $X_1, X_2, \ldots$ are independent and identically distributed given θ, with sample space (X,F) where (X,F) and (Θ,B) are isomorphic to Borel subsets of complete separable metric spaces, and if $\theta_1 \neq \theta_2$ implies $P_{\theta_1} \neq P_{\theta_2}$, then Doob (1949) showed that there exists some F_∞-measurable function h on X_∞ such that $P_\theta(h(\underset{\sim}{X}) = \theta) = 1$ for every $\theta \in \Theta$. Hence, any B-measurable bet $b(\theta)$ is, in this case, F_∞-measurable.

We now prove that upper and lower posterior expectations determined by the range $R(L,U)$ of prior measures are strongly consistent estimators:

Theorem 6.2. Let L and U be mutually absolutely continuous lower and upper measures on (Θ,B). Suppose the density $\ell(\theta)$ of L with respect to U is F_∞-measurable. If the measure U and the bet $b(\theta)$ satisfy the assumptions of Lemma 6.1, then

$$\sup_{Q \in R(L,U)} |Q_N(b)/Q_N(1) - b(\theta)| \to 0 \quad [U].$$

Proof. Let A be the set of $\theta \in \Theta$ for which $\ell(\theta) > 0$ and for which, with P_θ-probability one, $U_N(\ell(b-\lambda)^+ + (b-\lambda)^-)/U_N(1) \to \ell(\theta)(b(\theta)-\lambda)^+ + (b(\theta)-\lambda)^-$ and $U_N((b-\lambda)^+ + \ell(b-\lambda)^-)/U_N(1) \to (b(\theta)-\lambda)^+ + \ell(\theta)(b(\theta)-\lambda)^-$ for all rational λ. By Lemma

6.1, $U(A^c) = 0$. Fix $\theta \in A$ and suppose $\alpha_1 < b(\theta) < \alpha_2$, with α_1 and α_2 rational. Then $(b(\theta) - \alpha_2)^+ + \ell(\theta)(b(\theta)-\alpha_2)^- < 0$. Thus, with P_θ-probability one, $U_N((b-\alpha_2)^+ + \ell(b-\alpha_2)^-) < 0$ eventually which, by Theorem 4.1, implies $\sup\{Q_N(b)/Q_N(1) | Q \in R(L,U)\} < \alpha_2$ eventually. Similarly, $\inf\{Q_N(b)/Q_N(1) | Q \in R(L,U)\} > \alpha_1$ eventually with P_θ-probability one. Since α_1 and α_2 are arbitrary, the theorem is proved. ∎

Convenient approximations to the lower and upper posterior expectations of $b(\theta)$ are available when $b(\theta)$ is asymptotically normal in the following sense:

Definition 6.3. Let S be the set of functions $g:R^1 \to R^1$ that are bounded and continuous almost everywhere with respect to Lebesgue measure; and let $\phi(t)$ be the standard normal density. The bet $b(\theta)$ is asymptotically normal under U if for all $g \in S$,

$$U_N(g((b-b_N)/\sigma_N))/U_N(1) \to \int_{-\infty}^{\infty} g(t)\phi(t)dt \ [U],$$

where $b_N = U_N(b)/U_N(1)$ and $\sigma_N^2 = U_N(b^2)/U_N(1)-b_N^2$.

Recall from Example 4.2 that for $k > 1$, the constant $\gamma(k)$ is the unique solution of $\int_{-\infty}^{\infty}(k(t-\gamma)^+ + (t-\gamma)^-)\phi(t)dt = 0$.

Theorem 6.4. *Assume that the conditions of Theorem 6.2 hold and that $b(\theta)$ is asymptotically normal under U, with mean b_N and variance σ_N^2. Let $k_N = U_N(1)/L_N(1)$. Then*

$$\sigma_N^{-1}\left[\sup_{Q \in R(L,U)} Q_N(b)/Q_N(1) - (b_N + \sigma_N\gamma(k_N))\right] \to 0 \ [U]$$

and

$$\sigma_N^{-1}\left[\inf_{Q \in R(L,U)} Q_N(b)/Q_N(1) - (b_N - \sigma_N\gamma(k_N))\right] \to 0 \ [U].$$

Proof. We will prove the desired result for the maximum posterior expectation of b; that for the minimum posterior expectation is similar.

For real λ, define $Z_N(\lambda) = (b(\theta)-b_N)/\sigma_N - \lambda$. Then $U_N(Z_N^2(\lambda))/U_N(1) = 1 + \lambda^2$ for all N. From the asymptotic normality of $b(\theta)$ under U, we find

$$\lim_{M\to\infty} \overline{\lim_N}\, U_N(|Z_N(\lambda)| 1_{|Z_N(\lambda)|>M})/U_N(1) \le \lim_{M\to\infty} \lim_{N\to\infty} U_N(Z_N^2(\lambda) 1_{|Z_N(\lambda)|>M})/U_N(1) =$$

$$\lim_{M\to\infty} \lim_{N\to\infty} \{1+\lambda^2 - U_N(Z_N^2(\lambda) 1_{|Z_N(\lambda)|\le M})/U_N(1)\} = \lim_{M\to\infty} \{1+\lambda^2 - \int_{|t-\lambda|\le M} (t-\lambda)^2\phi(t)dt\} = 0 \ [U].$$

Hence,

$$\overline{\lim_N} \left|U_N(Z_N^+(\lambda))/U_N(1) - \int(t-\lambda)^+\phi(t)dt\right| \leq \lim_{M\to\infty} \overline{\lim_N}\ U_N(|Z_N(\lambda)|1_{|Z_N(\lambda)|>M})/U_N(1) +$$

$$\lim_{M\to\infty} \overline{\lim_N} \left|U_N(Z_N^+(\lambda)1_{|Z_N(\lambda)|\leq M})/U_N(1) - \int_{|t-\lambda|\leq M} (t-\lambda)^+\phi(t)dt\right|$$

$$+ \lim_{M\to\infty} \left|\int_{|t-\lambda|\leq M} (t-\lambda)^+\phi(t)dt - \int(t-\lambda)^+\phi(t)dt\right| = 0 \quad [U] .$$

Similarly, $U_N(Z_N^-(\lambda))/U_N(1) \to \int(t-\lambda)^-\phi(t)dt$ [U]. Let $c_N = 1/k_N = U_N(\ell)/U_N(1)$. Since $0 \leq \ell(\theta) \leq 1$ [U] and $\ell(\theta)$ is F_∞-measurable, Lemma 6.1 implies that $c_N \to \ell(\theta)$ [U]. Thus, for every real λ we have

$$U_N(Z_N^+(\lambda) + c_N Z_N^-(\lambda))/U_N(1) \to \int(t-\lambda)^+\phi(t)dt + \ell(\theta)\int(t-\lambda)^-\phi(t)dt \quad [U].$$

Furthermore, by the Cauchy-Schwartz inequality,

$$\left|\frac{U_N((c_N-\ell)Z_N^-(\lambda))}{U_N(1)}\right|^2 \leq \frac{U_N(|c_N-\ell|^2)}{U_N(1)} \cdot \frac{U_N(Z_N^2(\lambda))}{U_N(1)} = (1+\lambda^2)[U_N(\ell^2)/U_N(1) - c_N^2] \to 0 \quad [U]$$

since, by Lemma 6.1, $U_N(\ell^2)/U_N(1) \to \ell^2(\theta)$ [U]. Hence, for every real λ,

$$U_N(Z_N^+(\lambda) + \ell Z_N^-(\lambda))/U_N(1) \to \int(t-\lambda)^+\phi(t)dt + \ell(\theta)\int(t-\lambda)^-\phi(t)dt \quad [U].$$

Now, let A denote the set of $\theta \in \Theta$ for which $\ell(\theta) > 0$ and for which, with P_θ-probability one, $U_N(Z_N^+(\lambda) + \ell Z_N^-(\lambda))/U_N(1)$ converges to $\int(t-\lambda)^+\phi(t)dt + \ell(\theta)\int(t-\lambda)^-\phi(t)dt$ for all rational λ. Then $U(A^c) = 0$. Fix $\theta \in A$, and suppose $\alpha_1 < \gamma(1/\ell(\theta)) < \alpha_2$ with α_1 and α_2 rational. Then $\int(t-\alpha_2)^+\phi(t)dt + \ell(\theta)\int(t-\alpha_2)^-\phi(t)dt < 0$. Hence, with P_θ-probability one, $U_N(Z_N^+(\alpha_2) + \ell Z_N^-(\alpha_2)) < 0$ eventually, which, by Theorem 4.1, implies that $\sup\{Q_N((b-b_N)/\sigma_N)/Q_N(1) \mid Q \in R(L,U)\} < \alpha_2$ eventually. Similarly,

$$\sup\{Q_N((b-b_N)/\sigma_N)/Q_N(1) \mid Q \in R(L,U)\} > \alpha_1$$

eventually with P_θ-probability one. Since α_1, α_2 are arbitrary, we have shown that

$$\sigma_N^{-1}\{\sup\{Q_N(b)/Q_N(1) \mid Q \in R(L,U)\} - b_N\} \to \gamma(1/\ell(\theta)) \quad [U] .$$

Furthermore, $k_N \to 1/\ell(\theta)$ [U] and $\gamma(\cdot)$ is continuous; thus, $\gamma(k_N) \to \gamma(1/\ell(\theta))$ [U], and the theorem is proved. ∎

Theorem 6.4 permits close approximations of ranges of posterior expectations of arbitrary bets. When θ is the true parameter value, the upper and lower posterior expectations of b are $b_N \pm \sigma_N \gamma(1/\ell(\theta)) + o(\sigma_N)$. Thus, if the prior density with respect to U is uncertain up to a factor of, say, 2 in the sense that $\ell(\theta) \geq \frac{1}{2}\,[U]$, then the range of posterior expectations of b is $b_N \pm .276\,\sigma_N + o(\sigma_N)$ for any bet b satisfying the weak conditions of the theorem.

REFERENCES

[1] DE FINETTI, B. (1937). Foresight: Its logical laws, its subjective sources. 93-158, in Studies in Subjective Probability (1964), ed. Kyburg, H. E. and Smokler, H. E., Wiley, New York.

[2] DICKEY, J. M. (1976). Approximate posterior distributions. J. Am. Stat. Assoc. 71, 680-689.

[3] DOOB, J. L. (1949). Applications of the theory of martingales. Colloques Internationaux de Centre National de la Recherche Scientifique Paris, 22-28.

[4] EDWARDS, W., LINDMAN, H. and SAVAGE, L.J. (1963). Bayesian statistical inference for psychological Research. Psychological Reveiw 70, 193-242.

[5] FREEDMAN, D. A. and PURVES, R. A. (1969). Bayes methods for bookies. Ann. Math. Statist. 40, 1177-1186.

[6] GOOD, I. J. (1961). Weight of evidence, causality, and false-alarm probabilities. Fourth London Symp. on Information Theory, Butterworths, 125-136.

[7] HEATH, D. C. and SUDDERTH, W. D. (1972). On a theorem of de Finetti, oddsmaking and game theory. Ann. Math. Statist. 43, 2072-2077.

[8] HEATH, D. C. and SUDDERTH, W. D. (1978). On finitely additive priors, coherence, and extended admissibility. Ann. Statist. 6, 333-345.

[9] KEYNES, J. M. (1921). A Treatise on Probability. Macmillan, London.

[10] KOOPMAN, B. O. (1940). The axioms and algebra of intuitive probability. Annals of Mathematics 41, 269-278.

[11] RAMSEY, F. P. (1926). Truth and Probability. 61-92, in Studies in Subjective Probability (1964), ed. Kyburg, H. E. and Smokler, H. E., Wiley, New York.

[12] RENYI, A. (1970). Probability Theory. American Elsevier, New York.

[13] SAVAGE, L. J. (1954). The Foundations of Statistics. Wiley, New York.

[14] SMITH, C. A. B. (1961). Consistency in statistical inference and decision. J. Roy. Stat. Soc. Ser. B 23, 1-25.

LDR:JAH/db

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 1915 / 2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) RANGES OF PRIOR MEASURES.	5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Lorraine DeRobertis and J. A. Hartigan	8. CONTRACT OR GRANT NUMBER(s) ~~MCS75-08374~~ (15) DAAG29-75-C-0024, NSF-MCS78-09525
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS #4-Probability, Statistics and Combinatorics
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below	12. REPORT DATE (11) January 1979
	13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) MRC-TSR-1915	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. (12) 25 p.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(9) **Technical summary rept.,**

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Upper and Lower measures, regular conditional measures, upper and lower expectations, upper and lower variances, strong consistency, approximate ranges of posterior expectations.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Partial prior knowledge is quantified by a range $R(L,U)$ of σ-finite prior measures Q satisfying $L(A) \leq Q(A) \leq U(A)$ for all measurable sets A, and is interpreted as a probability statement in a betting framework. The concept of conditional probability distributions is generalized to that of conditional measures, and Bayes theorem is extended to accommodate unbounded priors. According to Bayes theorem, the range $R(L,U)$ of prior measures is transformed upon observing X into a similar range $R(L_X,U_X)$ of posterior measures. Upper and lower expectations and variances induced by such ranges of measures are obtained. Under weak

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

221 200

Abstract (continued)

regularity conditions, these upper and lower posterior expectations are strongly consistent estimators. The range of posterior expectations of an arbitrary function b on the parameter space is asymptotically $b_N \pm \alpha\sigma_N + o(\sigma_N)$ where b_N and σ_N^2 are the posterior mean and variance of b induced by the upper prior measure U, and where α is a constant reflecting the uncertainty about the prior in terms of the derivative of L with respect to U.