The Annals of Statistics
1981, Vol. 9, No. 2, 235–244

BAYESIAN INFERENCE USING INTERVALS OF MEASURES[1]

By Lorraine DeRobertis and J. A. Hartigan

University of Wisconsin-Madison and National Bureau of Standards and Yale University

Partial prior knowledge is quantified by an interval $I(L, U)$ of σ-finite prior measures Q satisfying $L(A) \le Q(A) \le U(A)$ for all measurable sets A, and is interpreted as acceptance of a family of bets. 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 interval $I(L, U)$ of prior measures is transformed upon observing X into a similar interval $I(L_X, U_X)$ of posterior measures. Upper and lower expectations and variances induced by such intervals of measures are obtained. Under weak regularity conditions, as the amount of data increases, 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 determined by the density of L with respect to U reflecting the uncertainty about the prior.

1. Introduction. In practice, prior knowledge is typically vague and any elicited prior distribution is only an approximation to the true one. Less stringent modes of quantifying prior information would make the Bayes approach more practical and would protect 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 nonnegative 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.

Throughout the following, Θ will denote an arbitrary parameter space, and $\mathscr{B}$ a σ-field

Received April 1979; revised January 1980.

[1] 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.

AMS 1970 *subject classifications.* Primary 62A15; secondary 60F15.

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

of subsets of Θ. A *measure* Q on $(\Theta, \mathscr{B})$ is a nonnegative, countably-additive function on $\mathscr{B}$; 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, and continuous linear functionals on C (i.e., expectations of bounded measurable functions of θ) are optimized at the extreme points of C.

For a given pair of σ-finite measures L, U on $(\Theta, \mathscr{B})$ satisfying $L(A) \le U(A)$ for all $A \in \mathscr{B}$ (denoted $L \le U$), let $I(L, U)$ be the convex set of measures Q satisfying $L \le Q \le U$. 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 interval $I(L, U)$ in that prior measures in $I(L, U)$ map into posterior measures ranging between those induced by L and U. Upper and lower posterior expectations, and variances of arbitrary bets $b: \Theta \to R^1$ are obtained. Under quite weak conditions, as the amount of data increases, the upper and lower posterior expectations of $b(\theta)$ are strongly consistent estimates of $b(\theta)$; 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 interval $I(L, U)$ 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 intervals of prior measures. For a real parameter θ, it is assumed that the prior density p satisfies $c \le p(\theta) \le (1 + \beta)c$ for $\theta \in D$ and $p(\theta) \le \gamma c$ for all θ. This is just the interval 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 Lebesgue measure. Provided D has high posterior probability induced by μ, the posterior interval of measures is then close to the posterior measure for a uniform prior.

2. Intervals of measures.

Given a pair Q_1, Q_2 of measures on $(\Theta, \mathscr{B})$, we say $Q_1 \le Q_2$ if $Q_1(A) \le Q_2(A)$ for all $A \in \mathscr{B}$. Let L and U be σ-finite measures on $(\Theta, \mathscr{B})$ satisfying $L \le U$. The *interval of measures* $I(L, U)$ consists of all measures Q such that $L \le Q \le U$. Equivalently, if L and U have densities l and u with respect to some σ-finite measure ν on $(\Theta, \mathscr{B})$, then $I(L, U)$ consists of measures with densities q with respect to ν satisfying $l(\theta) \le q(\theta) \le u(\theta)[\nu]$, where $[\nu]$ signifies "ν almost surely." If $Q \in I(L, U)$, then the odds ratio $Q(A)/Q(B)$ for $A, B \in \mathscr{B}$ 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, $\mathscr{B}$-measurable function b on Θ is called a *bet*, and $b(\theta)$ the *payoff* received if θ occurs. For a measure Q on $(\Theta, \mathscr{B})$, 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 $Q(b) \ge 0$. Of coure, 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 $A \in \mathscr{B}$; thus proportional measures are equivalent probability statements.

A bet b is $I(L, U)$-*nonnegative* if $Q(|b|) < \infty$ and $Q(b) \ge 0$ for every Q in $I(L, U)$. The interval of measures $I(L, U)$, interpreted as a probability statement, is taken to mean the acceptance of all $I(L, U)$-nonnegative bets. The class of all measures proportional to some member of $I(L, U)$ is, thus, equivalent to the probability statement $I(L, U)$. For any bet b, define $b^+(\theta) = b(\theta)$ if $b(\theta) \ge 0$, $b^+(\theta) = 0$ if $b(\theta) < 0$, and define $b^-(\theta) = b(\theta) - b^+(\theta)$. Let $B(\theta) = 1$ if $b(\theta) \ge 0$, else $B(\theta) = 0$. Since $Q(b) = Q(b^+) + Q(b^-) \ge L(b^+) + U(b^-)$ for all Q in $I(L, U)$, and since $Q_0(b) = L(b^+) + U(b^-)$, where $Q_0(b') = L(Bb') + U[(1 - B)b']$ is such that $Q_0 \in I(L, U)$, we see that a bet b is $I(L, U)$-nonnegative if and only if $U(|b|) < \infty$ and $L(b^+) + U(b^-) \ge 0$. A bet b is defined to be $I(L, U)$-positive if $L(b^+) + U(b^-) > 0$.

3. Intervals of posterior measures. Bayes theorem states that if $\{P_\theta | \theta \in \Theta\}$ is a family of probability measures defined on a σ-field $\mathscr{F}$ of subsets of a sample space $\mathscr{X}$, where for each θ, P_θ has density $f(X|\theta)$ with respect to a σ-finite measure ν on $(\mathscr{X}, \mathscr{F})$, and if Q_0 is a probability measure on $(\Theta, \mathscr{B})$, and $f(X|\theta)$ is $\mathscr{F} \times \mathscr{B}$ measurable as a function of X and θ, then the unique probability measure on $(\mathscr{X} \times \Theta, \mathscr{F} \times \mathscr{B})$ having θ-marginal measure Q_0 and regular conditional distribution $\{P_\theta | \theta \in \Theta\}$ given $\mathscr{B}$ 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 $\mathscr{F}$ has density $f(X|\theta)/\int f(X|\theta)\, dQ_0(\theta)$ with respect to Q_0, given X. 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 $\nu\{X \in \mathscr{X} | \int f(X|\theta)\, dQ_0(\theta) = \infty\}$ 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 $dQ_0/d\theta = \theta^{-1}(1 - \theta)^{-1}$, $0 \le \theta \le 1$, 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 $dQ_X(\theta) \equiv \theta^{X-1}(1 - \theta)^{-X} d\theta$ may be interpreted to give a probability statement about θ according to which, given X, all bets $b(\theta)$ such that $\int b(\theta)\, dQ_X(\theta) \ge 0$ are acceptable.

To accommodate all σ-finite priors the concept of a regular conditional probability measure is generalized 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 $(\Omega, \mathscr{S})$ and let $\mathscr{D}$ be a sub-σ-field of $\mathscr{S}$. A family $\{Q_\omega | \omega \in \Omega\}$ of measures on $(\Omega, \mathscr{S})$ is a regular conditional measure given $\mathscr{D}$ if for every $\mathscr{S}$-measurable, Q-integrable function b

(i) $Q_\omega(b)$ is $\mathscr{D}$-measurable and finite $[Q]$;
(ii) $Q_\omega(gb) = g(\omega)Q_\omega(b)[Q]$ for every bounded $\mathscr{D}$-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 Loeve (1955), page 363).

If $\{Q_\omega | \omega \in \Omega\}$ is a regular conditional measure given $\mathscr{D}$, and if $Q(\Omega) = 1$ and $Q_\omega(\Omega) = 1[Q]$, then $\{Q_\omega | \omega \in \Omega\}$ is a regular conditional probability given $\mathscr{D}$. For, taking $b \equiv 1$, condition (ii) of Definition 3.1 together with the monotone convergence theorem implies that $Q_\omega(h) = h(\omega)[Q]$ for any $\mathscr{D}$-measurable function h. Letting $h(\omega) = g(\omega)Q_\omega(b)$ for arbitrary bounded $\mathscr{D}$-measurable g and Q-integrable b, we obtain $Q_\omega(gQ_{(\cdot)}(b)) = g(\omega)Q_\omega(b) = Q_\omega(gb)[Q]$, where the last equality follows from (ii). Condition (iii) then implies that $Q(gb) = Q(gQ_{(\cdot)}(b))$; thus, $\{Q_\omega | \omega \in \Omega\}$ is a regular conditional probability given $\mathscr{D}$. In particular, $Q_{(\cdot)}(b)$ is (Q)-unique. For unbounded measures Q we prove the following:

LEMMA 3.2. *Assume Q is a σ-finite measure on $\mathscr{S}$, and let $\{Q_\omega | \omega \in \Omega\}$, $\{Q'_\omega | \omega \in \Omega\}$ be regular conditional measures for Q given $\mathscr{D} \subset \mathscr{S}$. Then there exists a $\mathscr{D}$-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 $\mathscr{S}$ is equivalent to the existence of an $\mathscr{S}$-measurable, Q-integrable function b_0 such that $b_0(\omega) > 0$ for every $\omega \in \Omega$. Since Q_ω and 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 nonnegative, $\mathscr{S}$-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 $C(\omega) \le M$. 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 $Q_\omega(Cb_0 - b) = 0[Q]$. Hence, $Q(Cb_0 - b) = 0$; 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 $A \in \mathscr{D}$, 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 $\mathscr{S}$-measurable Q-integrable b. For arbitrary $\mathscr{S}$-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 *conditional probability statement*, 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.

Bayes' theorem for unbounded priors. Let $(\mathscr{X}, \mathscr{F})$ be a sample space, and let $\{P_\theta | \theta \in \Theta\}$ *be a family of probability measures on* $(\mathscr{X}, \mathscr{F})$ *with densities* $f(X|\theta)$ *with respect to a* σ*-finite measure* ν *on* $(\mathscr{X}, \mathscr{F})$. *Let* Q_0 *be a* σ*-finite prior measure on* $(\Theta, \mathscr{B})$. *Assume that* $f(X|\theta)$ *is measurable with respect to* $\mathscr{F} \times \mathscr{B}$.

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

Viewing $\mathscr{B}$ as a sub-σ-field of $\mathscr{F} \times \mathscr{B}$ in the obvious way, we next show that $\{P_\theta | \theta \in \Theta\}$ is a regular conditional measure of Q given $\mathscr{B}$ in the sense that the family of measures $\{Q_{X,\theta} | X \in \mathscr{X}, \theta \in \Theta\}$ defined by $Q_{X,\theta}(A \times B) \equiv P_\theta(A) 1_B(\theta)$ for $A \in \mathscr{F}$, $B \in \mathscr{B}$ satisfies conditions (i), (ii) and (iii) of Definition 3.1. Let b be any $\mathscr{F} \times \mathscr{B}$-measurable, Q-integrable function. Then $Q_{X,\theta}(b) = \int_{\mathscr{X}} b(t, \theta) f(t|\theta) d\nu(t) \equiv P_\theta(b)$ is $\mathscr{B}$-measurable by Fubini's theorem. If g is any bounded $\mathscr{B}$-measurable on $\mathscr{X} \times \Theta$ then, without loss of generality, g is a function of θ and so $P_\theta(bg) = \int_{\mathscr{X}} b(t, \theta) g(\theta) f(t|\theta) d\nu(t) = g(\theta) P_\theta(b)$. 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 $(\mathscr{X} \times \Theta, \mathscr{F} \times \mathscr{B})$ which agrees with Q_0 on $\mathscr{B}$ and for which $\{P_\theta | \theta \in \Theta\}$ is a regular conditional measure given $\mathscr{B}$. Then Q' is σ-finite, as is Q. If $C \in \mathscr{F} \times \mathscr{B}$ 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 $\{Q_X | X \in \mathscr{X}\}$ is a regular conditional measure of Q given $\mathscr{F}$, in the sense that the family of measures $\{Q'_{X,\theta} | X \in \mathscr{X}, \theta \in \Theta\}$ defined by $Q'_{X,\theta}(A \times B) = 1_A(X) \int_B f(X|t) dQ_0(t)$ for $A \in \mathscr{F}$, $B \in \mathscr{B}$ satisfies conditions (i), (ii) and (iii). Fubini's theorem implies that $Q_X(b)$ is $\mathscr{F}$-measurable for every Q-integrable b. And obviously $Q_X(bg) = g(X) Q_X(b)$ for any bounded $\mathscr{F}$-measurable g. Lastly, suppose $Q_X(b) = 0[Q]$. By Fubini's theorem, $Q(b) = \int \int b(X, t) f(X|t) dQ_0(t) d\nu(X) = \int Q_X(b) d\nu(X) = 0$. □

The measures Q_X, $X \in \mathscr{X}$ 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 $\mathscr{B}$-measurable bet b is acceptable whenever b is Q_0-integrable and $Q_0(b) \geq 0$. It should be noted that such popular bets as $b \equiv 1$ are excluded by the requirement of Q_0-integrability. Bets b such that b and $-b$ are acceptable are fair bets. Now consider Bayes theorem, interpreted on bets which are measurable functions on $\mathscr{F} \times \mathscr{B}$. For the conditional measures P_θ, any $\mathscr{F} \times \mathscr{B}$-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 $\mathscr{B}$-measurable bet $P_\theta(b)$. Bayes theorem states that there exists a unique measure Q on $(\mathscr{X} \times \Theta, \mathscr{F} \times \mathscr{B})$ such that b is Q-acceptable (i.e., $Q(b) \geq 0$) if and only if 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, $X \in \mathscr{X}$, 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 $\{P_\theta \mid \theta \in \Theta\}$. See also Freedman and Purves (1969).

If $Q_0 \in I(L, U)$, then $Q_X \in I(L_X, U_X)$. For if $L(A) \leq Q_0(A) \leq U(A)$ for all $A \in \mathscr{B}$ then it follows, since $f(X \mid \theta) \geq 0$, that $\int_A f(X \mid \theta) dL \leq \int_A f(X \mid \theta) dQ_0 \leq \int_A f(X \mid \theta) dU$ all $A \in \mathscr{B}$, and so that $L_X(A) \leq Q_X(A) \leq U_X(A)$ for all $A \in \mathscr{B}$. Thus, Bayes' theorem "works" for intervals of prior measures in that an interval of prior measures is transformed given X into an interval 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 and Bayes risks. For the $\mathscr{B}$-measurable, U-integrable bets b and c, with c $I(L, U)$-positive, consider the range of integral ratios $Q(b)/Q(c)$ for $Q \in I(L, U)$.

THEOREM 4.1. *$\inf\{Q(b)/Q(c) \mid Q \in I(L, U)\}$ is the unique solution λ of $U(b - \lambda c)^- + L(b - \lambda c)^+ = 0$. $\sup\{Q(b)/Q(c) \mid Q \in I(L, U)\}$ is the unique solution λ of $U(b - \lambda c)^+ + L(b - \lambda c)^- = 0$.*

PROOF. Let $\lambda_0 = \inf\{Q(b)/Q(c) \mid Q \in I(L, U)\}$, $c_1 = \inf\{Q(c) \mid Q \in I(L, U)\}$, and $c_2 = \sup\{Q(c) \mid Q \in I(L, U)\}$. Then $0 < c_1 < c_2 < \infty$ and $|\lambda_o| < \infty$. Since $U(b - \lambda c)^- + L(b - \lambda c)^+ = \inf\{Q(b - \lambda c) \mid Q \in I(L, U)\}$, it follows that $\lambda_o \geq \lambda$ if and only if $U(b - \lambda c)^- + L(b - \lambda c)^+ \geq 0$. Moreover, for any $\epsilon > 0$, $\lambda + \epsilon/c_1 \leq \lambda_o$ implies $\epsilon \leq U(b - \lambda c)^- + L(b - \lambda c)^+$ which in turn implies $\lambda + \epsilon/c_2 \leq \lambda_o$; thus; $\lambda_o > \lambda$ if and only if $U(b - \lambda c)^- + L(b - \lambda c)^+ > 0$. Hence, λ_o is the unique solution of $U(b - \lambda c)^- + L(b - \lambda c)^+ = 0$. □

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

The posterior mean $Q(\theta f(X \mid \theta))/Q(f(X \mid \theta))$, $Q \in I(L, U)$, 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 $1 \leq k \leq 10$. 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.

Suppose that $r(d, \theta)$ is the loss incurred if decision $d \in \mathscr{D}$ is taken when θ is true. For a measure Q satisfying $0 < Q(1) < \infty$, the *Bayes risk* corresponding to r is $\beta(r, Q) = \inf\{Q(r(d, \theta))/Q(1) \mid d \in \mathscr{D}\}$. Assume that $0 < L(1) \leq U(1) < \infty$ and that $r(d, \theta)$ is $\mathscr{B}$-

TABLE 1
Half interval, γ(k), of posterior means as a function of maximum prior ratio k.

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

measurable and U-integrable for every $d \in \mathscr{D}$. For $-\infty < \lambda < \infty$, define

$$\beta_1(\lambda) = \inf_{d \in \mathscr{D}}[U(r(d, \theta) - \lambda)^- + L(r(d, \theta) - \lambda)^+]$$

and

$$\beta_2(\lambda) = \inf_{d \in \mathscr{D}}[L(r(d, \theta) - \lambda)^- + U(r(d, \theta) - \lambda)^+].$$

THEOREM 4.3. *The lower Bayes risk* $\inf\{\beta(r, Q) \mid Q \in I(L, U)\}$ *is the unique solution* λ_1 *of* $\beta_1(\lambda) = 0$. *The upper Bayes risk* $\sup\{\beta(r, Q) \mid Q \in I(L, U)\}$ *is no greater than the unique solution* λ_2 *of* $\beta_2(\lambda) = 0$.

PROOF. It is straightforward to show that, for $\lambda \in R^1$ and $\delta > 0$, $-\delta U(1) \leq \beta_i(\lambda + \delta) - \beta_i(\lambda) \leq -\delta L(1)$ for $i = 1, 2$. Thus, β_1 and β_2 are continuous and strictly decreasing and, hence, have unique zeros; denote these by λ_1 and λ_2 respectively. Then $\inf\{\beta(r, Q) \mid Q \in I(L, U)\} = \sup\{\lambda \in R^1 \mid Q(r(d, \theta) - \lambda) \geq 0 \text{ for all } d \in \mathscr{D}, Q \in I(L, U)\} = \sup\{\lambda \in R^1 \mid L(r(d, \theta) - \lambda)^+ + U(r(d, \theta) - \lambda)^- \geq 0 \text{ for all } d \in \mathscr{D}\} = \sup\{\lambda \in R^1 \mid \beta_1(\lambda) \geq 0\} = \lambda_1$. And $\sup\{\beta(r, Q) \mid Q \in I(L, U)\} = \inf\{\lambda \in R^1 \mid \beta(r, Q) \leq \lambda \text{ for all } Q \in I(L, U)\} \leq \inf\{\lambda \in R^1 \mid \sup_{Q \in I(L,U)} Q(r(d, \theta) - \lambda) < 0 \text{ for some } d \in \mathscr{D}\} = \inf\{\lambda \in R^1 \mid L(r(d, \theta) - \lambda)^- + U(r(d, \theta) - \lambda)^+ < 0 \text{ for some } d \in \mathscr{D}\} = \inf\{\lambda \in R^1 \mid \beta_2(\lambda) < 0\} = \lambda_2$.

EXAMPLE 4.4. Consider the normal location problem of Example 4.2, and let $\mathscr{D} = R^1$ and $r(d, \theta) = |d - \theta|^2$. For $Q \in I(L, U)$, $\beta(r, Q_X)$ is the posterior variance of θ corresponding to Q; its range is bounded by the solutions of

$$\inf_{d \in \mathscr{D}} U_X[k(|d - \theta|^2 - \lambda_1)^- + (|d - \theta|^2 - \lambda_1)^+] = 0$$

and

$$\inf_{d \in \mathscr{D}} U_X[(|d - \theta|^2 - \lambda_2)^- + k(|d - \theta)^2 - \lambda_2)^+] = 0.$$

Because U_X is symmetric about X, the optimal decision for both equations is $d = X$. The solutions λ_1 and λ_2 are thus each posterior variances of elements of $I(L, U)$ so that the bounds of Theorem 4.3 are, in this case, sharp. Defining $H(\Delta, c) = \Delta\, \phi(\Delta) + (\Delta^2 - 1)(\Phi(\Delta) - c)$, it develops that $\lambda_1 = \sigma_0^2\Delta_1^2$ and $\lambda_2 = \sigma_0^2\Delta_2^2$ where Δ_1 is the solution of $H(\Delta, c) = 0$ with $c = (k - 2)/(2k - 2)$, and Δ_2 is the solution of $H(\Delta, c) = 0$ with $c = (2k - 1)/(2k - 2)$. Table 2 gives Δ_1 and Δ_2 for $1 \leq k \leq 10$.

5. Asympototic behavior of ranges of posterior expectations. Let the observation X_i have sample space $(\mathscr{X}_i, \mathscr{F}_i)$ $i = 1, 2, \cdots$, and define $\mathscr{X}_\infty = \prod_1^\infty \mathscr{X}_i$, $\mathscr{F}_\infty = \prod_1^\infty \mathscr{F}_i$. Suppose $\{P_\theta \mid \theta \in \Theta\}$ is a family of probability distributions of $(\mathscr{X}_\infty, \mathscr{F}_\infty)$ such that $P_\theta(A)$ is $\mathscr{B}$-measurable for all $A \in \mathscr{F}_\infty$. Let Q be any measure on $(\Theta, \mathscr{B})$ and let $Y, Y_1, Y_2, \cdots$ be any

TABLE 2
Interval (Δ_1, Δ_2) *of posterior standard deviation as a function of maximum prior ratio* k.

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

sequence of random variables on $(\mathscr{X}_\infty \times \Theta, \mathscr{F}_\infty \times \mathscr{B})$. The notation $Y = 0[Q]$ will mean that, except for some set of θ-values of Q-measure zero, $P_\theta\{\mathbf{X} \in \mathscr{X}_\infty \mid Y(\mathbf{X}, \theta) = 0\} = 1$. Similarly, the notation $Y_N \rightarrow Y[Q]$ will mean that, except for some set of θ-values of Q-measure zero, $Y_N(\mathbf{X}, \theta)$ converges to $Y(\mathbf{X}, \theta)$ with P_θ-probability one. Similarly Y is $\mathscr{F}_\infty$-measurable $[Q]$ if h is $\mathscr{F}_\infty$-measurable and $P_\theta[Y = h] = 1[Q]$.

Assume further that, for all θ, the $(X_1, \cdots, X_N)$-marginal distribution of P_θ has density $f(X_1, \cdots, X_N \mid \theta)$ with respect to some σ-finite measure ν_N on $(\mathscr{X}_1 \times \cdots \times \mathscr{X}_N, \mathscr{F}_1 \times \cdots \times \mathscr{F}_N)$. For a $\mathscr{B}$-measurable bet b and a measure Q on $(\Theta, \mathscr{B})$, we will denote $Q_N(b) = \int b(\theta) f(X_1, \cdots, X_N \mid \theta) dQ(\theta)$; and for $A \in \mathscr{B}$, define $Q_N(A) = Q_N(1_A)$.

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

LEMMA 5.1 *Assume Q is a σ-finite measure on $(\Theta, \mathscr{B})$; let $\Theta = \cup_1^\infty A_i$ with $A_i \in \mathscr{B}$ disjoint and* $0 < Q(A_i) < \infty$. *Suppose*

(i) $b(\theta)$ *is $\mathscr{F}_\infty$-measurable* $[Q]$;
(ii) $Q_N(1) < \infty$ $[\nu_N]$ *for sufficiently large N*;
(iii) $Q_N(A_i)/Q_N(1) \rightarrow 1_{A_i}(\theta)[Q]$ *for* $i = 1, 2, \cdots$;
(iv) $\lim_{M\rightarrow\infty}\lim\sup_N Q_N(\mid b \mid 1_{|b|>M})/Q_N(1) \rightarrow 0[Q]$.

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

PROOF. Note, first, that if $A \in \mathscr{B}$, $0 < Q(A) < \infty$, and $Q(\mid b1_A \mid) < \infty$ then, by the martingale convergence theorem applied to the probability measure $Q_A(D) \equiv Q(A \cap D)/Q(A)$ for $D \in \mathscr{B}$, it follows that $Q_N(b1_A)/Q_N(A)$ converges to $b(\theta)$ with P_θ-probability one for all θ in A, except for a subset of A of Q-measure zero.

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

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

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

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

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

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

REMARK. Condition (i) may be weakened slightly to requiring that $b(\theta)$ is estimable: $t_N(X_1, \cdots, X_N) \rightarrow b(\theta)[Q]$ for some sequence of estimators t_N.

If $X_1, X_2, \cdots$ are independent and identically distributed given θ, with sample space $(\mathscr{X}, \mathscr{F})$ where $(\mathscr{X}, \mathscr{F})$ and $(\Theta, \mathscr{B})$ 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 $\mathscr{F}_\infty$-measurable function h on $\mathscr{X}_\infty$ such that $P_\theta(h(\mathbf{X}) = \theta) = 1$ for every $\theta \in \Theta$. Hence, any $\mathscr{B}$-measurable bet $b(\theta)$ satisfies condition (i).

THEOREM 5.2. *Let L and U be mutually absolutely continuous lower and upper measures on $(\Theta, \mathscr{B})$. Suppose the density $l(\theta)$ of L with respect to U is $\mathscr{F}_\infty$-measurable $[Q]$. If the measure U and the bet $b(\theta)$ satisfy the assumptions of Lemma* 5.1, *then*

$$\sup_{Q\in I(L,U)} \mid Q_N(b)/Q_N(1) - b(\theta) \mid \rightarrow 0[U].$$

PROOF. Let A be the set of $\theta \in \Theta$ for which $l(\theta) > 0$ and for which, with P_θ-probability one, $U_N(l(b-\lambda)^+ + (b-\lambda)^-)/U_N(1) \to l(\theta)(b(\theta)-\lambda)^+ + (b(\theta)-\lambda)^-$ and $U_N((b-\lambda)^+ + l(b-\lambda)^-)/U_N(1) \to (b(\theta)-\lambda)^+ + l(\theta)(b(\theta)-\lambda)^-$ for all rational λ. By Lemma 5.1, $U(A^c) = 0$. Fix $\theta \in A$ and suppose $\alpha_1 < b(\theta) < \alpha_2$, with α_1 and α_2 rational. Then $(b(\theta) - \alpha_2)^+ + l(\theta)(b(\theta)-\alpha_2)^- < 0$. Thus, with P_θ-probability one, $U_N((b-\alpha_2)^+ + l(b-\alpha_2)^-) < 0$ eventually which, by Theorem 4.1, implies $\sup\{Q_N(b)/Q_N(1) \mid Q \in I(L, U)\} < \alpha_2$ eventually. Similarly, $\inf\{Q_N(b)/Q_N(1) \mid Q \in I(L, U)\} > \alpha_1$ eventually with P_θ-probability one. Since α_1 and α_2 are arbitrary, the theorem is proved. □

DEFINITION 5.3. Let S be the set of functions $g: R^1 \to R^1$ that are bounded and continuous; 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 5.4. *Assume that the conditions of Theorem* 5.2 *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*

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

and

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

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

$$\begin{aligned}
&\lim_{M\to\infty}\limsup_N U_N(|Z_N(\lambda)|\,1_{|Z_N(\lambda)|>M})/U_N(1)\\
&\quad\le \lim_{M\to\infty}\lim_{N\to\infty} U_N(Z_N^2(\lambda)1_{|Z_N(\lambda)|>M})/U_N(1)\\
&\quad= \lim_{M\to\infty}\lim_{N\to\infty}(1+\lambda^2 - U_N(Z_N^2(\lambda)1_{|Z_N(\lambda)|\le M})/U_N(1)\}\\
&\quad= \lim_{M\to\infty}\left\{1+\lambda^2 - \int_{|t-\lambda|\le M}(t-\lambda)^2\phi(t)\,dt\right\}\\
&\quad= 0[U].
\end{aligned}$$

Hence

$$\begin{aligned}
&\limsup_N \left| U_N(Z_N^+(\lambda))/U_N(1) - \int (t-\lambda)^+\,\phi(t)\,dt \right|\\
&\le \lim_{M\to\infty}\limsup_N U_N(|Z_N(\lambda)|1_{|Z_N(\lambda)|>M})/U_N(1)\\
&\quad+ \lim_{M\to\infty}\limsup_N \left| U_N(Z_N^+(\lambda)1_{|Z_N(\lambda)|\le M})/U_N(1) - \int_{|t-\lambda|\le M}(t-\lambda)^+\phi(t)\,dt \right|\\
&\quad+ \lim_{M\to\infty}\left| \int_{|t-\lambda|\le M}(t-\lambda)^+\phi(t)\,dt - \int (t-\lambda)^+\phi(t)\,dt \right|\\
&= 0\ [U].
\end{aligned}$$

Similarly, $U_N(Z_N^-(\lambda))/U_N(1) \to \int(t-\lambda)^-\phi(t)dt[U]$. Let $c_N = 1/k_N = U_N(l)/U_N(1)$. Since $0 \le l(\theta) \le 1$ $[U]$ and $l(\theta)$ is $\mathscr{F}_\infty$-measurable $[Q]$, Lemma 5.1 implies that $c_N \to l(\theta)[U]$. Thus, for every real λ we have

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

Furthermore, by the Cauchy-Schwartz inequality,

$$\left|\frac{U_N((c_N - l)Z_N^-(\lambda))}{U_N(1)}\right|^2 \le \frac{U_N(|c_N - l|^2)}{U_N(1)} \cdot \frac{U_N(Z_N^2(\lambda))}{U_N(1)}$$

$$= (1+\lambda^2)[U_N(l^2)/U_N(1) - c_N^2] \to 0[U]$$

since, by Lemma 5.1, $U_N(l^2)/U_N(1) \to l^2(\theta)[U]$. Hence, for every real λ,

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

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

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

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

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

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

Theorem 5.4 permits close approximations of intervals of posterior expectations of arbitrary bets. When θ is the true parameter value, the upper and lower posterior expectations of b are $b_N \pm \sigma_N\gamma(1/l(\theta)) + o(\sigma_N)$. Thus, if the prior density with respect to U is uncertain up to a factor of, say, 2 in the sense that $l(\theta) = ½[U]$, then the interval 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 thorem.

Acknowledgment. We are grateful for the work of a referee who suggested numerous improvements in terminology and proofs.

REFERENCES

[1] DICKEY, J. M. (1976). Approximate posterior distributions. *J. Amer. Statist. Assoc.* **71** 680–689.

[2] DOOB, J. L. (1949). Applications of the theory of martingales. *Colloques Internationaux de Centre National de la Recherche Scientific Paris* 22–28.

[3] EDWARDS, W., LINDMAN, H. AND SAVAGE, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Rev.* **70** 193–242.

[4] FREEDMAN, D. A. AND PURVES, R. A. (1969). Bayes methods for bookies. *Ann. Math. Statist.* **40** 1177–1186.

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

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

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

[8] KEYNES, J. M. (1921). *A Treatise on Probability.* Macmillan, London.
[9] KOOPMAN, B. O. (1940). The axioms and algebra of intuitive probability. *Ann. Math.* **41** 269–278.
[10] LOEVE, M. (1955). *Probability Theory.* Van Nostrand, Princeton.
[11] SMITH, C. A. B. (1961). Consistency in statistical inference and decision. *J. Roy. Stat. Soc. Ser. B* **23** 1–25.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN-MADISON
1210 W. DAYTON ST.
MADISON, WISCONSIN 53706

DEPARTMENT OF STATISTICS
YALE UNIVERSITY
BOX 2179, YALE STATION
NEW HAVEN, CONNECTICUT 06520