The Annals of Probability
1975, Vol. 3, No. 2, 375–379

ON CONDITIONAL MEDIANS[1]

BY R. J. TOMKINS

University of Regina

This article examines the properties of conditional medians, especially vis-à-vis those of conditional expectations.

1. Introduction. While the notion of a conditional median has been defined (see Loève (1963) page 385) and employed (for example, see Csörgö (1970) page 316), the literature does not appear to contain an *exposition* of the properties of conditional medians. This paper aims to present such an exposition.

Let $(\Omega, \mathscr{F}, P)$ be a probability space and let $\mathscr{G}_0$ be the trivial sigma-field. Let R represent the set of real numbers and let $I(A)$ denote the indicator function of an event $A \in \mathscr{F}$.

DEFINITION. Let X be a random variable (rv) and let $\mathscr{G}$ be a sigma-field with $\mathscr{G} \subseteq \mathscr{F}$. An rv M is called a *conditional median* (CM) of X with respect to $\mathscr{G}$ if and only if

(CM1) M is $\mathscr{G}$-measurable, and
(CM2) $P(X \geqq M \mid \mathscr{G}) \geqq \frac{1}{2} \leqq P(X \leqq M \mid \mathscr{G})$ almost surely (a.s.).

A conditional median M shall be denoted by $\mu(X \mid \mathscr{G})$.

It seems natural to say that a CM M is *unique* if, given any other CM M^*, we have $M = M^*$ a.s. Note that $\mu(X \mid \mathscr{G}_0)$ is just a usual (constant) median of X which is, of course, not necessarily unique; it is easy to concoct examples to show that this is also the case when $\mathscr{G}$ is not trivial. Hence, inequalities and equalities involving CM's will be carefully explained.

Loève ((1963) page 385) has defined M to be a CM of X if and only if (CM2) holds. Thus our definition is tougher.

Note that if ϕ is any strictly monotone function on R and if M is any CM of X, then $\phi(M)$ is a CM of $\phi(X)$.

THEOREM 1. *Let $\mathscr{G}$ be a sigma-field. Then every* rv *X has a* CM *with respect to $\mathscr{G}$.*

PROOF. Assume that $P(X \leqq x \mid \mathscr{G})(\omega)$, where $x \in R$ and $\omega \in \Omega$, is a regular conditional distribution function for X given $\mathscr{G}$ (cf. Ash (1972) page 263). For each $\omega \in \Omega$, define $M(\omega) \equiv \inf\{x \in R : P(X \leqq x \mid \mathscr{G})(\omega) > \frac{1}{2}\}$. Clearly M is a rv and (CM1) holds. Moreover, for any (fixed) positive integer n, and any

Received November 9, 1973; revised June 14, 1974.

[1] This work was supported by a grant from the National Research Council of Canada.

AMS 1970 *subject classifications.* Primary 60A05, 60E05; Secondary 60F99.

Key words and phrases. Conditional median, Jensen's inequality.

integer k,

$$P(X < M - n^{-1} \mid \mathscr{G}) \leqq P(X \leqq (k-1)/n \mid \mathscr{G}) \leqq \tfrac{1}{2} \leqq P(X \leqq k/n \mid \mathscr{G}) \leqq P(X \leqq M + n^{-1} \mid \mathscr{G})$$

a.s. on the event $[k - 1 < nM \leqq k]$. Hence $P(X \leqq M + n^{-1} \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. and $P(X \geqq M - n^{-1} \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. for all $n \geqq 1$. (CM2) now follows readily from the monotone convergence theorem for conditional expectations (Loève, page 348). □

REMARK. Our definition of CM can be shown to be equivalent to that of Brunk and Johansen (1970) so that Theorem 1 is a consequence of their Theorem 1.5. Theorem 1 may also be obtained by proving it first for simple rv and then employing Lemma 1 (below) twice.

2. Some properties of conditional medians. Since conditional medians are measures of central tendency in their own right, it behooves us to examine their properties in light of the well-known traits of conditional expectations. We will commence such an examination with a convergence lemma.

LEMMA 1. *Let $\mathscr{G}$ be a sigma-field. Let $\{X_n\}$, $\{M_n\}$, $n \geqq 1$, be sequences of* rv *such that*

(a) *M_n is a* CM *of X_n with respect to $\mathscr{G}$, and*

(b) rv *X, M exist such that $X_n \to X$ in probability* (i.p.) *and $M_n \to M$* i.p.

Then M is a CM *of X with respect to $\mathscr{G}$.*

PROOF. Note first that M is $\mathscr{G}$-measurble. To start with, assume $X_n \to X$ a.s. and $M_n \to M$ a.s. Clearly $I(X \geqq M) \geqq \limsup_{n\to\infty} I(X_n \geqq M_n)$ a.s. Hence, by Fatou's lemma for conditional expectations (Loève, page 348), $\frac{1}{2} \leqq \limsup_{n\to\infty} P(X_n \geqq M_n \mid \mathscr{G}) \leqq P(X \geqq M \mid \mathscr{G})$ a.s. Similarly $P(X \leqq M \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. Finally, if (b) holds, a subsequence $\{n_k\} \uparrow \infty$ exists satisfying $X_{n_k} \to X$ a.s. and $M_{n_k} \to M$ a.s. so that the preceding argument applies. □

REMARK 1. Lemma 1 remains true if (b) is replaced by the weaker condition (b'): $P(\cdot \mid \mathscr{G})$ is a regular conditional probability, $M_n \to M$ i.p., and an event $C \in \mathscr{F}$ exists such that $P(C) = 1$ and, for each $\omega \in C$, the distributions $P(X_n \leqq x \mid \mathscr{G})(\omega)$ converge completely to $P(X \leqq x \mid \mathscr{G})(\omega)$. This follows, by Exercise 2 on page 214 of Loève.

REMARK 2. It is worthy of note that, even when $X_n \to X$ a.s., it is not necessary that a sequence of respective CM's $\{M_n\}$ exists and converges to a CM of X. For example, let Y be uniformly distributed on $(-\frac{1}{2}, \frac{1}{2})$, let $\mathscr{G} = \mathscr{G}_0$, and define $X_n \equiv I(Y \leqq (-1)^n/n)$, $n \geqq 1$, and $X \equiv I(Y < 0)$. Then $X_n \to X$ a.s., and, for each $n \geqq 1$, X_n has a unique median $M_n = 0$ or 1 accordingly as n is odd or even. But, clearly, $\{M_n\}$ does not converge.

THEOREM 2. *Let X, Y be* rv *and let $\mathscr{G}$ be a sigma-field. Let $\mu(X \mid \mathscr{G})$ be any* CM *of X. If Y is $\mathscr{G}$-measurable, then the following statements hold.*

(i) *If* $\mu(Y \mid \mathscr{G})$ *is any* CM *of* Y *then* $\mu(Y \mid \mathscr{G}) = Y$ a.s.

(ii) $\mu(X \mid \mathscr{G}) + Y$ *is a* CM *of* $X + Y$; *i.e.* $\mu(X + Y \mid \mathscr{G}) = \mu(X \mid \mathscr{G}) + Y$.

(iii) $Y\mu(X \mid \mathscr{G})$ *is a* CM *of* XY; *i.e.* $\mu(XY \mid \mathscr{G}) = Y\mu(X \mid \mathscr{G})$.

PROOF. To prove (i), note that $P(Y \geqq \mu(Y \mid \mathscr{G}) \mid \mathscr{G}) = I(Y \geqq \mu(Y \mid \mathscr{G})) \geqq \frac{1}{2}$ a.s. Similarly, $I(Y \leqq \mu(Y \mid \mathscr{G})) \geqq \frac{1}{2}$ a.s. These inequalities are true only if $\mu(Y \mid \mathscr{G}) \leqq Y \leqq \mu(Y \mid \mathscr{G})$ a.s., establishing (i).

The proof of (ii) is trivial, so proceed to the proof of (iii). On the ($\mathscr{G}$-measurable) event $[Y = 0]$, $P(XY \geqq Y\mu(X \mid \mathscr{G}) \mid \mathscr{G}) = P(0 \geqq 0 \mid \mathscr{G}) = 1 > \frac{1}{2}$ a.s. On the event $[Y > 0]$, $P(XY \geqq Y\mu(X \mid \mathscr{G}) \mid \mathscr{G}) = P(X \geqq \mu(X \mid \mathscr{G}) \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. Finally, on $[Y < 0]$, $P(XY \geqq Y\mu(X \mid \mathscr{G}) \mid \mathscr{G}) = P(X \leqq \mu(X \mid \mathscr{G}) \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. That $P(XY \leqq Y\mu(X \mid \mathscr{G}) \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. follows similarly. □

THEOREM 3. *Let* X, Y *be* rv, *let* $a \in R$, *let* $\mathscr{G}$ *be a sigma-field, and let* $\mu(X \mid \mathscr{G})$ *be any* CM *of* X. *Then the following assertions are true.*

(i) *If* M *is any* CM *of the constant* rv a, *then* $M = a$ a.s.

(ii) $a\mu(X \mid \mathscr{G})$ *is a* CM *of* aX; *i.e.* $\mu(aX \mid \mathscr{G}) = a\mu(X \mid \mathscr{G})$.

(iii) *If* $X \geqq Y$ a.s. *then a* CM $\mu(Y \mid \mathscr{G})$ *exists such that* $\mu(X \mid \mathscr{G}) \geqq \mu(Y \mid \mathscr{G})$.

(iv) *If* $X \geqq a$ a.s. *then* $\mu(X \mid \mathscr{G}) \geqq a$ a.s.

(v) *If* X *is independent of* $\mathscr{G}$, *then every median* $\mu(X \mid \mathscr{G}_0)$ *of* X *is a* CM *of* X *with respect to* $\mathscr{G}$.

(vi) *If, for each* $n \geqq 1$, $\mu_n(X \mid \mathscr{G})$ *is a* CM *of* X *and if* M *is a* $\mathscr{G}$*-measurable* rv *satisfying* $\inf_{n \geqq 1} \mu_n(X \mid \mathscr{G}) \leqq M \leqq \sup_{n \geqq 1} \mu_n(X \mid \mathscr{G})$ a.s., *then* M *is also a* CM *of* X.

(vii) *There exist* CM'*s* M^* *and* M^{**} *of* X *such that* $M^* \leqq M \leqq M^{**}$ a.s. *for all* CM'*s* M *of* X.

PROOF. (i) and (ii) are immediate from theorem 2 by taking $Y = a$, while (iv) follows from (i) and (iii).

To prove (iii), let M be any CM of Y and define $\mu(Y \mid \mathscr{G}) = \min(M, \mu(X \mid \mathscr{G}))$. Clearly $\mu(Y \mid \mathscr{G})$ is $\mathscr{G}$-measurable, $\mu(Y \mid \mathscr{G}) \leqq \mu(X \mid \mathscr{G})$, and $P(Y \geqq \mu(Y \mid \mathscr{G}) \mid \mathscr{G}) \geqq P(Y \geqq M \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. Moreover $P(Y \leqq \mu(Y \mid \mathscr{G}) \mid \mathscr{G}) = P(Y \leqq M \mid \mathscr{G}) \geqq \frac{1}{2}$ on the event $[M < \mu(X \mid \mathscr{G})]$ while, on the complement of that event, $P(Y \leqq \mu(Y \mid \mathscr{G}) \mid \mathscr{G}) = P(Y \leqq \mu(X \mid \mathscr{G}) \mid \mathscr{G}) \geqq P(X \leqq \mu(X \mid \mathscr{G}) \mid \mathscr{G}) \geqq \frac{1}{2}$ a.s. Thus $\mu(Y \mid \mathscr{G})$ is indeed a CM of Y.

To prove (v), let $m \in R$ be any $\mathscr{G}_0$-median of X. Then, by independence, $P(X \geqq m \mid \mathscr{G}) = P(X \geqq m) \geqq \frac{1}{2} \leqq P(X \leqq m \mid \mathscr{G})$.

In view of the definition of CM and of part (ii), it will be enough to show that (vi) holds in the case $M = \sup_{n \geqq 1} \mu_n(X \mid \mathscr{G})$.

Note that, given any two CM's M_1, M_2 of X, $\max(M_1, M_2)$ is also a CM of X; this can be proved readily by an argument similar to the proof of (iii). Hence, letting $M_n \equiv \max_{k \leqq n} \mu_k(X \mid \mathscr{G})$, $n \geqq 1$, it follows by induction that each M_n is a CM of X. Moreover $M_n \uparrow M$, where M is easily seen to be a rv. Hence M is a CM of X by Lemma 1.

For (vii), note that it will suffice to establish the existence of M^{**}, since that of M^* follows by part (ii). To this end, let $\mathscr{M}$ be the collection of all CM's of X with respect to $\mathscr{G}$, and let M^{**} be the essential supremum of $\mathscr{M}$. (vii) is an immediate consequence of part (vi) and Theorem 1.5 of Chow, Robbins, and Siegmund (1971). □

REMARK. If X is independent of $\mathscr{G}$, it does not necessarily follow that every CM of X is constant (cf. (v) above). For example, let X and Y be independent, identically distributed (i.i.d.) rv with $P[X = 1] = P[X = 0] = \frac{1}{2}$, and let $\mathscr{G}$ be the sigma-field generated by Y; then Y is a CM of X.

THEOREM 4 (*Jensen's inequality for conditional medians*). *Let ϕ be any real-valued convex function which is defined on an open interval I, bounded or unbounded, in R. Then, given any sigma-field $\mathscr{G}$, any* rv *X with $P[X \in I] = 1$, and any* CM *$\mu(X \mid \mathscr{G})$, a* CM *of the* rv *$\phi(X)$ exists satisfying $\mu(\phi(X) \mid \mathscr{G}) \geqq \phi(\mu(X \mid \mathscr{G}))$* a.s.

In particular, CM's *$\mu(|X| \mid \mathscr{G})$ and $\mu(X^2 \mid \mathscr{G})$ exist such that $\mu(|X| \mid \mathscr{G}) \geqq |\mu(X \mid \mathscr{G})|$* a.s. *and $\mu(X^2 \mid \mathscr{G}) \geqq \mu^2(X \mid \mathscr{G})$* a.s.

PROOF. Note that real sequences $\{a_n\}$ and $\{b_n\}$ exist such that $\phi(x) = \sup_{n \geqq 1} (a_n x + b_n)$ for all $x \in I$ (see Ash (1972) page 286). Since $P[X \in I] = 1$, $\phi(X) \geqq a_n X + b_n$ a.s. for all $n \geqq 1$. By Theorem 2(ii) and (iii), $a_n \mu(X \mid \mathscr{G}) + b_n$ is a CM of $a_n X + b_n$. Moreover, Theorem 3(iii) shows that a CM $\mu_n(\phi(X) \mid \mathscr{G})$ exists which is a.s. at least as great as $a_n \mu(X \mid \mathscr{G}) + b_n$.

Let $\mu(\phi(X) \mid \mathscr{G}) \equiv \sup_{n \geqq 1} \mu_n(\phi(X) \mid \mathscr{G})$; this is a CM of $\phi(X)$ by Theorem 3(vi). Note that $\mu(X \mid \mathscr{G}) \in I$ a.s. by Theorem 3(iv). Hence $\mu(\phi(X) \mid \mathscr{G}) \geqq \sup_{n \geqq 1} (a_n \mu(X \mid \mathscr{G}) + b_n) = \phi(\mu(X \mid \mathscr{G}))$ a.s.

The special cases follow by taking $I = R$, and $\phi(x) = |x|$ or $\phi(x) = x^2$. □

THEOREM 5. *Let X, Y be* rv *and let $\mathscr{G}$ be a sigma-field. Then*

(i) *it is not always possible to find* CM's *satisfying $\mu(X + Y \mid \mathscr{G}) = \mu(X \mid \mathscr{G}) + \mu(Y \mid \mathscr{G})$;*

(ii) *it is not always possible to find* CM's *satisfying $\mu(\mu(X \mid \mathscr{G}) \mid \mathscr{G}_1) = \mu(X \mid \mathscr{G}_1)$ where $\mathscr{G}_1 \subseteqq \mathscr{G}$ is a sigma-field;*

(iii) *it is not always possible to find* CM's *satisfying $\mu^2(|XY| \mid \mathscr{G}) \leqq \mu(X^2 \mid \mathscr{G})\mu(Y^2 \mid \mathscr{G})$; i.e. the Cauchy–Schwarz inequality does not always hold for* CM's.

PROOF. The counterexamples given below only involve rv with unique CM's.

That (i) holds is well known. In fact, it holds even if $XY = 0$: let X_1, X_2 be i.i.d. with $P[X_1 = 0] = \frac{1}{3} = 1 - P[X_1 = 1]$, put $X = X_1 X_2$, $Y = X_1 - X_1 X_2$ so that $\mu(X \mid \mathscr{G}_0) = \mu(Y \mid \mathscr{G}_0) = 0$ while $\mu(X + Y \mid \mathscr{G}_0) = \mu(X_1 \mid \mathscr{G}_0) = 1$.

For (ii), define X, Y such that $P[Y = k] = \frac{1}{3}$, $k = 0, 1, 2$, $P(X = 1 \mid Y = k) = \frac{3}{8} = 1 - P(X = 0 \mid Y = k)$, $k = 0, 1$ and $P(X = 1 \mid Y = 2) = \frac{7}{8} = 1 - P(X = 0 \mid Y = 2)$. Let $\mathscr{G}$ be the sigma-field generated by Y. Since $P(X = 1) = \frac{13}{24}$, $\mu(X \mid \mathscr{G}_0) = 1$. But $\mu(X \mid \mathscr{G}) = \mu(X \mid Y) = I(Y = 2)$ so that $\mu(\mu(X \mid \mathscr{G}) \mid \mathscr{G}_0) = 0$.

To prove (iii), let X, Y be i.i.d. rv with $P[X = 1] = .6 = 1 - P[X = 2]$. Then $\mu(X \mid \mathscr{G}_0)\mu(Y \mid \mathscr{G}_0) = 1 \cdot 1 = 1$. But $P[XY = 1] = .36$ and $P[XY = 2) = .48$ so $\mu(XY \mid \mathscr{G}_0) = 2$. Hence, by a remark in Section 1, $\mu^2(|XY| \mid \mathscr{G}_0) = \mu^2(XY \mid \mathscr{G}_0) > \mu^2(X \mid \mathscr{G}_0)\mu^2(Y \mid \mathscr{G}_0) = \mu(X^2 \mid \mathscr{G}_0)\mu(Y^2 \mid \mathscr{G}_0)$. □

3. Acknowledgment. The author thanks the referee for pointing out the relevance of the Brunk and Johansen paper, and for other useful comments and suggestions.

REFERENCES

[1] Ash, Robert B. (1972). *Real Analysis and Probability*. Academic Press, New York.

[2] Brunk, H. D. and Johansen, S. (1970). A generalized Radon-Nikodym derivative. *Pacific J. Math.* **34** 585–617.

[3] Chow, Y. S., Robbins, Herbert and Siegmund, David (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, New York.

[4] Csörgő, M. (1970). On a law of the iterated logarithm for strongly multiplicative systems. *Acta Math. Acad. Sci. Hungar.* **21** 315–321.

[5] Loève, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, Princeton.

Department of Mathematics
University of Regina
Regina, Saskatchewan, Canada