The Annals of Probability
1975, Vol. 3, No. 3, 569-572

SOME MARTINGALES ASSOCIATED WITH SAMPLE SPACINGS

BY MAREK KANTER

Sir George Williams University

A certain sequence of statistics based on sample spacings is shown to constitute a martingale. This fact is applied to prove an inequality relating to a process of picking points randomly but not independently.

Suppose m points $X_1, \cdots, X_m$ are randomly chosen in $(0, l)$. Let $X_{(1)}, \cdots, X_{(m)}$ denote the same points arranged in ascending order, and for $i = 1, \cdots, m+1$ write $D_i = X_{(i)} - X_{(i-1)}$ to denote the spacings between these points. (We set $X_{(0)} \equiv 0$, $X_{(m+1)} \equiv l$.)

For $\lambda \in (0, l)$ we form the random variables $V_m(\lambda, l) = \sum_{i=1}^{m+1} |D_i - \lambda|$ and $V_m^+(\lambda, l) = \sum_{D_i \geqq \lambda} (D_i - \lambda)$. The principal theorem of this paper states that both $V_m(\lambda, l)$ and $V_m^+(\lambda, l)$ form a martingale sequence when suitably normalized. Denote $V_m(\lambda, 1)$ and $V_m^+(\lambda, 1)$ by $V_m(\lambda)$ and $V_m^+(\lambda)$ respectively. We were led to study $V_m(\lambda)$, because $V_m(1/(m+1))$ has an obvious interpretation as a measure of how close the points $X_1, \cdots, X_m$ are to being equidistributed. We shall call $V_m(1/(m+1))$ the *sample total variation of the points* $X_1, \cdots, X_m$ and we shall call $V_m(\lambda)$ the *sample total variation from* λ *of the points* $X_1, \cdots, X_m$. The quantity $V_m^+(\lambda)$ is also interesting; one interpretation would be that it measures how much "usable" space there remains after m points X_i have been picked (if one specifies that space is usable if it is of length exceeding λ and contains none of the points X_i.)

LEMMA 1. *Suppose $X_1, \cdots, X_m$ are independent random variables uniformly distributed on $(0, l)$. Define $V_m(\lambda, l)$ and $V_m^+(\lambda, l)$ in the same way as above. Then for $m \geqq 0$,*

$$\text{(a)} \qquad E(V_m(\lambda, l)) = m\lambda + (\lambda - l) \qquad \text{if } \lambda > l$$
$$= m\lambda + (\lambda - l) + 2l(1 - \lambda/l)(1 - \lambda/l)^m \qquad \text{if } \lambda \leqq l$$

while

$$\text{(b)} \qquad E(V_m^+(\lambda, l)) = 0 \qquad \text{if } \lambda > l$$
$$= l(1 - \lambda/l)^{m+1} \qquad \text{if } \lambda \leqq l\,.$$

PROOF. We prove (a), from which (b) follows since it is easily seen that

$$\text{(c)} \qquad V_m(\lambda, l) + l - \lambda(m+1) = 2V_m^+(\lambda, l)\,.$$

We note that $m = 0$ corresponds to no observations and in that case (a) holds trivially. Furthermore if $\lambda > l$ the result is clear. Finally if $\lambda \leqq l$, the expectation

Received February 8, 1974; revised June 10, 1974.
AMS 1970 *subject classifications.* Primary 60G45; Secondary 62G30.
Key words and phrases. Martingales, sample spacings, total variation.

in (a) equals $(m+1)E(|\lambda - D_1|)$ by exchangeability. All that is left is to compute $E(|\lambda - D_1|)$. □

REMARK 1. In [2] it is shown that $V_m(1/(m+1))$ converges in probability to $2/e$. We note that (a) implies $E(V_m(1/(m+1))) = 2(1-(m+1)^{-1})^{m+1}$.

THEOREM 1. *Let $X_1, \cdots, X_{m+1}$ be random variables with values in $(0, 1)$. No other assumptions are made on $(X_1, \cdots, X_m)$ but X_{m+1} is assumed to be uniformly distributed in $(0, 1)$ and independent of $(X_1, \cdots, X_m)$. For $\lambda \in (0, 1)$ form the random variable $V_m(\lambda)$ from $(X_1, \cdots, X_m)$ and the random variable $V_{m+1}(\lambda)$ from $(X_1, \cdots, X_{m+1})$.*

We claim that

$$E(V_{m+1}(\lambda) \mid X_1, \cdots, X_m) = (1-\lambda)V_m(\lambda) + \lambda^2(m+1) .$$

PROOF. Let $X'_{(1)} < \cdots < X'_{(m+1)}$ be the ascending rearrangement of the random variables $(X_1, \cdots, X_{m+1})$ and define $D_i' = X'_{(i)} - X'_{(i-1)}$ for $i = 1, \cdots, m+2$. (As usual $X'_{(0)} = 0$ and $X'_{(m+2)} = 1$.) we have

$$E(V_{m+1}(\lambda) \mid X_1, \cdots, X_m) = \textstyle\sum_{i=1}^{m+1} E(V^*_{m+1}(\lambda, D_i') \mid X_1, \cdots, X_m)$$

where $V^*_{m+1}(\lambda, D_i')$ stands for the contribution from the interval $(X_{(i-1)}, X_{(i)})$ to the total sum $\sum_{i=1}^{m+2} |D_i' - \lambda|$. By Lemma 1, we have

$$(*) \qquad \begin{aligned} V_{m+1}(\lambda, D_i) &= B_i + \lambda(\lambda - D_i) & &\text{if } \lambda > D_i \\ &= B_i + (\lambda - D_i) + 2(1-\lambda/D_i)(1-\lambda/D_i)^{B_i} & &\text{if } \lambda \leqq D_i \end{aligned}$$

where $B_i = 1$ if $X_{m+1} \in (X_{(i-1)}, X_{(i)})$; otherwise $B_i = 0$. We note that

$$E(B_i \mid X_1, \cdots, X_m) = D_i .$$

Taking the conditional expectation of both sides of (*) we get

$$(**) \qquad \begin{aligned} E(V_{m+1}(\lambda, D_i) \mid X_1, \cdots, X_m) &= D_i\lambda + (\lambda - D_i) & &\text{if } \lambda > D_i \\ &= D_i\lambda + (\lambda - D_i) \\ &\quad + 2(D_i - \lambda)(1-\lambda) & &\text{if } \lambda \leqq D_i . \end{aligned}$$

We conclude that

$$\begin{aligned} E(V_m(\lambda) \mid X_1, \cdots, X_m) &= \textstyle\sum_{D_i < \lambda} [D_i\lambda + (\lambda - D_i)] \\ &\quad + \textstyle\sum_{D_i \geqq \lambda} [D_i\lambda + (\lambda - D_i) + 2(1-\lambda)(D_i - \lambda)] \\ &= 2(1-\lambda) \textstyle\sum_{D_i \geqq \lambda} (D_i - \lambda) + (m+2)\lambda - 1 \\ &= 2(1-\lambda)V_m^+(\lambda) + (m+2)\lambda - 1 . \end{aligned}$$

By (c), we conclude the last expression is equal to

$$(1-\lambda)(V_m(\lambda) + 1 - \lambda(m+1)) + (m+2)\lambda - 1$$

which simplifies to

$$(1-\lambda)V_m(\lambda) + \lambda^2(m+1) .$$ □

THEOREM 2. *Let now $X_1, X_2, \cdots, X_m, \cdots$ be independent and uniform in $(0, 1)$.*

Then for fixed $\lambda \in (0, 1)$ *the sequences*

$$\frac{V_m(\lambda) + 2(1-\lambda)^{m+1} - (m+1)\lambda + 1}{(1-\lambda)^{m+1}}$$

and $V_m^+(\lambda)/(1-\lambda)^{m+1}$ *are martingales.*

PROOF. By direct computation using Theorem 1 and (c). □

REMARK 2. Let $K_m = \max_{1 \leq j \leq m+1} \{D_j\}$. It is clear that $K_m \downarrow 0$ a.s. This implies that for any $\lambda > 0$ we have $V_m^+(\lambda) = 0$ a.s. for m sufficiently large. We conclude that the martingale $V_m^+(\lambda)/(1-\lambda)^{m+1}$ is not L^p bounded for any $p > 1$ (otherwise Doob's martingale theorem would guarantee a.s. convergence to 1).

We now apply the preceding results to the following problem. Let X_1 be picked randomly from (0, 1) and consider the subintervals $(0, X_1)$ and $(X_1, 1)$. Pick X_2 randomly from the larger of these two subintervals, thereby obtaining three subintervals $(0, X_{(1)})$, $(X_{(1)}, X_{(2)})$, and $(X_{(2)}, 1)$. This method of picking the X_n's was mentioned by Kakutani at a lecture in Berkeley in 1973. He conjectured that if the X_n are picked in the above fashion (we shall call it the "purposeful" method of picking from now on) then the points $X_1, \cdots, X_m$ tend to become equidistributed as $m \to \infty$. We do not prove his conjecture, but support it by showing that the sample variation of the points $X_1, \cdots, X_m$ when picked purposefully is dominated by the sample total variation of the points $X_1, \cdots, X_m$ when picked randomly.

LEMMA 2. *Let* $X_1, \cdots, X_m$ *be random variables with values in* (0, 1), *no other assumptions being made on their distribution. Fix i in* $\{1, \cdots, m\}$ *and let* $(X_1, \cdots, X_m, X_{m+1})$ *denote that random* $(m+1)$ *vector such that the conditional distribution of* X_{m+1} *given* $(X_1, \cdots, X_m)$ *is uniform in* $(X_{(i-1)}, X_{(i)})$. *Then for any* $\lambda > 0$, $E(V_{m+1}(\lambda))$ *is minimized as a function of i whenever* $X_{(i)} - X_{(i-1)}$ *is maximal.*

PROOF. For $y \in (0, 1)$ let $h_i(y)$ denote the value of $V_{m+1}(\lambda)$ if $X_{m+1} = X_{(i-1)} + y(X_{(i)} - X_{(i-1)})$. Let $D_i = X_{(i)} - X_{(i-1)}$; then $h_i(y) = V_m(\lambda) + |yD_i - \lambda| + |(1-y)D_i - \lambda| - |D_i - \lambda|$. It is easily obtained that $f(x) \equiv |yx - \lambda| + |(1-y)x - \lambda| - |x - \lambda|$ is decreasing in x for fixed y, hence we conclude that $h_i(y)$ is minimized if $X_{(i)} - X_{(i-1)}$ is maximal. Since $E(V_{m+1}(\lambda)) = \int_0^1 h_i(y)\,dy$ we are done. □

For clarity we shall use $\hat{X}_1, \cdots, \hat{X}_m, \cdots$ to stand for an infinite sequence of random variables generated "purposefully".

THEOREM 3. *Let* $\hat{V}_m(\lambda)$ *denote the sample total variation from* λ *after m picks from the sequence* $(\hat{X}_1, \cdots, \hat{X}_m, \cdots)$. *Then the sequence*

$$\frac{\hat{V}_m(\lambda) + 2(1-\lambda)^{m+1} - (m+1)\lambda + 1}{(1-\lambda)^{m+1}}$$

is a supermartingale, for $\lambda \in (0, 1)$.

PROOF. For any m let $(\hat{X}_1, \cdots, \hat{X}_m, X_{m+1})$ denote that random $m+1$ vector such that $\hat{X}_1, \cdots, \hat{X}_m$ are picked sequentially from (0, 1) in the purposeful fashion, and then X_{m+1} is picked uniformly from (0, 1). Let $V'_{m+1}(\lambda)$ denote the sample total variation from λ of the sequence $(\hat{X}_1, \cdots, \hat{X}_m, X_{m+1})$. By Lemma 3 we have that

$$E(\hat{V}_{m+1}(\lambda) \mid \hat{X}_1, \cdots, \hat{X}_m) \leqq E(V'_{m+1}(\lambda) \mid \hat{X}_1, \cdots, \hat{X}_m)$$

and the result then follows from Theorems 1 and 2. □

COROLLARY.

$$E\left(\hat{V}_m\left(\frac{1}{m+1}\right)\right) \leqq 2\left(1 - \frac{1}{m+1}\right)^{m+1}.$$

PROOF. $\hat{V}_0(\lambda) = E(\hat{V}_0)) = 1 - \lambda$ since no points have yet been randomly picked. It follows that

$$\frac{E(\hat{V}_m(\lambda)) + 2(1-\lambda)^{m+1} - (m+1)\lambda + 1}{(1-\lambda)^{m+1}} \leqq 4$$

for all $\lambda \in (0, 1)$, by the submartingale property proved in Theorem 3. Now let $\lambda = (m+1)^{-1}$. □

REFERENCES

[1] FELLER, W. (1971). *Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

[2] SHERMAN, B. (1950). A random variable related to the spacings of sample values. *Ann. Math. Statist.* **21** 339-361.

DEPARTMENT OF MATHEMATICS
SIR GEORGE WILLIAMS UNIVERSITY
MONTREAL 107, CANADA