The Annals of Statistics
1974, Vol. 2, No. 4, 829–832

A RATE OF CONVERGENCE OF A DISTRIBUTION CONNECTED WITH INTEGRAL REGRESSION FUNCTION ESTIMATION

BY GARY MAKOWSKI

Marquette University

Brunk studied integral regression functions and has obtained strong laws and limiting distributions for estimators of these functions. In this note we will study additional conditions that ensure a rate of convergence of the distribution function of the maximum absolute difference of an integral regression function and its estimator, suitably normalized, to the distribution function of a normalized maximum absolute value of partial sums of random variables. These results are corollaries of convergence results obtained by Sawyer and Rosenkrantz.

1. Introduction. We will define and discuss integral regression functions. What immediately follows is very similar to Section 2 of [1].

Suppose that associated with each point t of the unit interval there is a univariate distribution $D(t)$ with mean $\mu(t)$; $\mu(\cdot)$ is called the regression function. Let $\{t_n\}$ be a sequence of numbers in [0, 1], not necessarily distinct, to be called observation points. For each n, let $Y_n(t_n)$ denote a random variable having the distribution associated with t_n, so that $EY_n(t_n) = \mu(t_n)$; and let the random variables $\{Y_n(t_n)\}$ be independent. Write $h_j(\cdot)$ the indicator function of $[t_j, 1]$ and set

$$S_n(t) = \sum_{j=1}^{n} Y_j(t_j)h_j(t) , \qquad t \in [0, 1] .$$

Define s_n^2 to be the variance of $S_n(1) = \sum_{j=1}^{n} Y_j(t_j)$. Let $F_n(\cdot)$ denote the "empirical distribution function" of $\{t_1, \cdots, t_n\}$. For a given probability distribution function F with support in [0, 1] set

$$M(t) = \int_{[0,t]} \mu(v)\, dF(v)$$

for each t in [0, 1]. M is called the integral regression function. Also let

$$M_n(t) = ES_n(t)/n = \int_{[0,t]} \mu(v)\, dF_n(v) .$$

We will take $S_n(t)/n$ as our estimator of $M(t)$. Brunk [1] has obtained sufficient conditions for the a.s. convergence of $D_n = \sup_t |S_n(t)/n - M(t)|$ to zero. (Here and elsewhere we will write $\sup_t$ in place of $\sup_{0 \leq t \leq 1}$.) He also obtained the limiting distribution of D_n.

2. Maximum absolute value of partial sums. In what follows let $\{X_n\}$ be a sequence of independent random variables that are centered at expectations. Let

Received December 1972; revised September 1973.
AMS 1970 *subject classifications.* Primary 60F05, 62G05, 62G10; Secondary 60G50.
Key words and phrases. Integral regression function, maximum absolute value of partial sum, independent observations regression model.

$S_n = \sum_{j=1}^n X_j$ and write σ_n^2 for the variance of S_n. Define $G_n(x)$ and $G(x)$ by

$$G_n(x) = P[\max_{1 \leqq k \leqq n} |S_k| \leqq x\sigma_n]$$
$$G(x) = (4/\pi) \sum_0^\infty [(-1)^n/(2n+1)] \exp[-(2n+1)^2\pi^2/(8x^2)], \qquad x > 0$$
$$= 0, \qquad x \leqq 0.$$

This note is a corollary of the following result obtained for $2 < p \leqq 4$ by Rosenkrantz [2] and for $5 < p < \infty$ by Sawyer [3].

THEOREM 1. *If $\sigma_n^2/n \geqq J_1$ and $\sum_{k=1}^n E(|X_k|^p)/n \leqq J_2$ where J_1 and J_2 are positive constants and p satisfies either $2 < p \leqq 4$ or $5 < p < \infty$, then*

$$\sup_x |G_n(x) - G(x)| \leqq A(\log n)^{\frac{1}{2}} n^{-\frac{1}{2}(p-2)/(p+1)} = \Psi(n)$$

for some positive constant A whose value depends only on J_1, J_2 and p.

3. Integral regression functions. In applying the preceding to integral regression functions set $D_n' = \sup_t |S_n(t)/n - M_n(t)|$, $\Delta_n = \sup_t |M_n(t) - M(t)|$, $d_n = \sup_t |F_n(t) - F(t)|$, and observe that

$$D_n' - \Delta_n \leqq D_n \leqq D_n' + \Delta_n. \tag{1}$$

Thus

$$P[D_n' \leqq (x - \Delta_n/s_n)s_n] \leqq P[D_n \leqq xs_n] \leqq P[D_n' \leqq (x + \Delta_n/s_n)s_n]$$

and

$$\begin{aligned}
\sup_x &|P[D_n \leqq xs_n] - G(x)| \\
&\leqq \max\{\sup_x |P[D_n' \leqq (x + \Delta_n/s_n)s_n] - G(x)|, \\
&\qquad \sup_x |P[D_n' \leqq (x - \Delta_n/s_n)s_n] - G(x)|\} \\
&\leqq \max\{\sup_x |P[D_n' \leqq (x + \Delta_n/s_n)s_n] - G(x + \Delta_n/s_n)| \\
&\qquad + \sup_x |G(x + \Delta_n/s_n) - G(x)|, \\
&\qquad \sup_x |P[D_n' \leqq (x - \Delta_n/s_n)s_n] - G(x - \Delta_n/s_n)| \\
&\qquad + \sup_x |G(x - \Delta_n/s_n) - G(x)|\} \\
&= \sup_x |P[D_n' \leqq xs_n] - G(x)| + \sup_x |G(x + \Delta_n/s_n) - G(x)|.
\end{aligned}$$

Since $G(\cdot)$ satisfies a Lipschitz condition we can write $\sup_x |G(x + \Delta_n/s_n) - G(x)| \leqq A_1 \Delta_n/s_n$ for some positive constant A_1. Hence

$$\sup_x |P[D_n \leqq xs_n] - G(x)| \leqq \sup_x |P[D_n' \leqq xs_n] - G(x)| + A_1\Delta_n/s_n. \tag{2}$$

Theorem 1 and the above discussion enable us to state the following result.

THEOREM 2. *If $\{Y_j(t_j)\}_{j=1}^n$ satisfies the hypotheses of Theorem 1 and μ is continuous on $[0, 1]$, then there is a positive constant A_2 such that*

$$\begin{aligned}
\sup_x |P[D_n \leqq xs_n] - G(x)| &\leqq \Psi(n) + A_2 d_n/n^{\frac{1}{2}} \\
&\leqq \Psi(n) + A_2/n^{\frac{1}{2}}.
\end{aligned} \tag{3}$$

PROOF. Since D_n' is just a maximum absolute value of partial sums of the random variables $\{Y_1(t_1) - \mu(t_1), Y_2(t_2) - \mu(t_2), \cdots, Y_n(t_n) - \mu(t_n)\}$ when they are permuted in some way (see [1], page 179) we can use Theorem 1 to conclude that

$$\sup_x |P[D_n' \leqq xs_n] - G(x)| \leqq \Psi(n) .$$

Since $\mu(\cdot)$ is continuous on [0, 1], $\Delta_n \leqq A_3 d_n \leqq A_3$ for some positive constant A_3. The conclusion follows from (2) and the hypothesis $\sigma_n^2/n \geqq J_1$ of Theorem 1.

It is interesting to note that $\Psi(n)^{-1} = o(n^{\frac{1}{2}})$. Thus for the rate $\Psi(n)$ and the variance s_n^2 that is bounded away from zero, by (2), the boundedness of Δ_n is all that is needed to ensure a $2\Psi(n)$ convergence rate for $\sup_x |P[D \leqq xs_n] - G(x)|$; this rate obtains whether or not F_n converges to F.

The situation when the observation points $\{T_n\}$ are random variables and the distribution of Y_n given $[T_n = t_n]$ is denoted by $D(t_n)$ (so that $E(Y_n \mid T_n = t) = \mu(t)$) is called an independent observations regression model and is considered by Brunk [1]. We now define $F_n(\cdot)$ to be the empirical distribution function of $T_1, \cdots, T_n$. Following the argument of Corollary 2.5 [1] one can consider the case where Theorem 2 holds on a set of points $\{t_n\}$ of probability 1 and derive a convergence rate for $\sup_x |P[D_n \leqq xs_n] - G(x)|$:

THEOREM 3. *In an independent observations regression model assume that for positive constants* p, J_1, *and* J_2 *with probability* 1 *for every* n *the distributions* $D(T_1)$, $D(T_2), \cdots, D(T_n)$ *satisfy the hypothesis of Theorem* 2. *Then* (3) *holds and*

$$\sup_x |P[D_n \leqq xs_n] - G(x)| \leqq 2\Psi(n)$$

for n *sufficiently large.*

INDICATION OF PROOF. (See [1] for definitions of terms.) Set $C = \{\omega_1 = (t_1, t_2, \cdots) \in \Omega_1$: for positive constants p, J_1 and J_2, given in the hypothesis of Theorem 3, for every n, the distributions $D(t_1), D(t_2), \cdots, D(t_n)$ satisfy the hypotheses of Theorem 2$\}$ and $B(x, \omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in [D_n \leqq xs_n]\}$. By assumption $P_1(C) = 1$, and

$$\begin{aligned}\sup_x |P[D_n \leqq xs_n] - G(x)| &= \sup_x |\textstyle\int_C (P_2^1(\omega_1, B(x, \omega_1)) - G(x))P_1(d\omega_1)| \\ &\leqq \textstyle\int_C \sup_x |P_2^1(\omega_1, B(x, \omega_1)) - G(x)| P_1(d\omega_1) \\ &\leqq \textstyle\int_C [\Psi(n) + A_2/n^{\frac{1}{2}}] P_1(d\omega_1) \\ &= \Psi(n) + A_2/n^{\frac{1}{2}} \leqq 2\Psi(n)\end{aligned}$$

for n sufficiently large.

REMARK. The continuity of μ can be replaced by the following condition "μ is bounded, and continuous on all but a finite number of points at which μ is left continuous," provided that the integrals $\int_{[0,t]} \mu(v)\, dF(v)$ and $\int_{[0,t]} \mu(v)\, dF_n(v)$ are interpreted as Darboux–Stieltjes integrals.

REFERENCES

[1] BRUNK, H. D. (1969). Estimation of isotonic regression. *Nonparametric Techniques in Statistical Inference*, 177–195. Cambridge Univ. Press.

[2] ROSENKRANTZ, W. A. (1967). On rates of convergence for the invariance principle. *Trans. Amer. Math. Soc.* **129** 542–552.
[3] SAWYER, S. (1967). A uniform rate of convergence for the maximum absolute value of partial sums in probability. *Comm. Pure Appl. Math.* **20** 647–658.

DEPARTMENT OF MATHEMATICS AND STATISTICS
MARQUETTE UNIVERSITY
MILWAUKEE, WISCONSIN 53233