The Annals of Mathematical Statistics
1968, Vol. 39, No. 6, 2149–2153

WEAK CONVERGENCE OF A SEQUENCE OF QUICKEST DETECTION PROBLEMS[1]

By Donald L. Iglehart[2] and Howard M. Taylor

Cornell University

1. Introduction and summary. Consider a production process which exists in one of two states, in-control and out-of-control. Production begins in-control and while there a chance event occurs after each item is produced so that the probability of remaining in-control is $(1 - \pi)$ and the probability of a transition out-of-control is π. Once out-of-control the process remains there until trouble is removed.

Associated with each item produced is a measurable characteristic or quality which is a Gaussian random variable whose mean depends on the underlying process state. Let the mean in-control be μ_0, the mean out-of-control be μ_1 and the common variance be σ^2. It costs K units to repair the process and the operating cost rate in-control is c_0 and out-of-control, c_1. The true process state is unknown and a rule is desired which specifies when to repair, based on the quality history, so as to minimize some function of the cost components.

This model of a production process was first studied by Girshick and Rubin (1952). They show that when the criterion is long run time average total cost to repair plus operation, the optimal rule is of the form: "Stop and repair after production of the kth item if and only if $Z(k) \geqq \zeta$" for some critical value ζ, where $Z(k)$ is a monotonic function of the posterior probability, given the observed history, that the process will be out of control for the $k + 1$st item. Unfortunately, no easy method for computing the critical value ζ was given.

Shiryaev (1963) studied a similar model but in continuous time and with a slightly different criterion. The continuous time process was a diffusion process and the operating characteristics of the control procedure could be found by solving a second order linear differential equation, thus enabling the appropriate critical value ζ to be found.

Taylor (1967) applied the continuous time computations of Shiryaev to the discrete time model of Girshick and Rubin and plotted the optimal choice for ζ as a function of the costs and other system parameters. No proof of the validity of the continuous time approximations to the discrete time process was given, however. This paper fills this gap by exhibiting a sequence of Girshick and Rubin processes that converges to the continuous process of Shiryaev, and by showing that the corresponding cost functionals also converge. This procedure of obtaining a limiting continuous process from a sequence of discrete processes

---

Received 18 January 1968; revised 11 July 1968.

[1] This work was supported by the Office of Naval Research, Nonr-401(55), and the National Science Foundation, GK-856.

[2] Now at Stanford University.

is very naturally carried out as an application of the theory of weak convergence of probability measures; cf. Billingsley (1968).

More precisely, let $(X_i : i = 1, 2, \cdots)$ be a sequence of independent Gaussian random variables with zero mean and unit variance. Let $S_0 = 0$, $S_k = X_1 + \cdots + X_k$ for $k = 1, 2, \cdots$, and define for $n = 1, 2, \cdots$ and $k = 0, 1, \cdots, n$, $X_k(k) = S_k/n^{\frac{1}{2}}$. Let T_n be a geometric random variable with $P[T_n = j] = (\pi/n)(1 - \pi/n)^{j-1}$ for $j = 1, 2, \cdots$, and $n = [\pi] + 1, [\pi] + 2, \cdots$. Define

$$\begin{aligned} \theta_n(k) &= 0 && \text{for } 0 \leqq k < T_n \\ &= n^{-1}(k - T_n + 1) && \text{if } T_n \leqq k, \end{aligned}$$

where $k = 0, 1, \cdots, n$. Finally let $Y_n(k) = \theta_n(k) + X_n(k)$. One interprets $Y_n(k)$ as the cumulative sum up to item k of observations in a truncated (at n) Girshick and Rubin process where $\mu_0 = 0$, $\mu_1 = 1/n$, $\sigma^2 = 1/n$ and the probability of leaving control between any two items is $\pi_n = \pi/n$.

Let $U_n(k)$ be the posterior probability that the production process will be out-of-control for the production of the $(k + 1)$st item, assuming no repair is made. It is convenient to take as our control variable, $Z_n(k)$, the monotone function of $U_n(k)$ given by

$$Z_n(k) = U_n(k)/(1 - U_n(k)).$$

Then using the recursion for $U_n(k)$ (cf. Taylor (1967)) it is easy to show by induction that $Z_n(0) = 0$ and for $k = 1, 2, \cdots, n$

$$Z_n(k) = \pi_n(1 - \pi_n)^{-1} \textstyle\sum_{j=0}^{k-1} (1 - \pi_n)^{-j} \exp\{Y_n(k) - Y_n(k - j) - j/2n\}.$$

Since our analysis will treat probability measures on $C[0, 1]$, the space of continuous, real-valued functions on $[0, 1]$ with the metric (ρ) of uniform convergence, it is convenient to introduce continuous versions of the processes $\{X_n(k)\}$, $\{\theta_n(k)\}$, $\{Y_n(k)\}$, and $\{Z_n(k)\}$. We define $x_n(t) = X_n(k) + (nt - k)(X_n(k + 1) - X_n(k))$ where $kn^{-1} \leqq t \leqq (k+1)n^{-1}$, $k = 0, 1, \cdots, n - 1$, and $0 \leqq t \leqq 1$. Observe that $x_n(t) = X_n(k)$ for $t = k/n$ and is defined by linear interpolation for other values of t. In a similar manner define $\theta_n(t)$, $y_n(t)$, and $z_n(t)$. Clearly, $y_n(t) = x_n(t) + \theta_n(t)$. Let ζ be a fixed positive number and $z \varepsilon C[0, 1]$, then define $\tau(z) = \inf\{t: 0 \leqq t \leqq 1, z(t) = \zeta\}$ if the set $\{t: 0 \leqq t \leqq 1, z(t) = \zeta\}$ is not empty and 1 otherwise. A cost functional whose expected value may be used to measure long run time average cost of operation plus repair (Taylor, (1967)) is given by

$$C_\zeta^{(n)}(z_n) = \textstyle\sum_{k=0}^{[n\tau(z_n)]} \{c(n^{-1})z_n(k/n)(1 + z_n(k/n))^{-1} - \gamma/n\}$$

where γ is a constant related to K, the repair cost.

For the continuous version of Shiryaev (1963) let $\{x(t): 0 \leqq t \leqq 1\}$ be Brownian motion and T, an exponential random variable, independent of $\{x(t): 0 \leqq t \leqq 1\}$, with density $\pi e^{-\pi t}$, $t > 0$. Define the continuous time process $(\theta(t); 0 \leqq t \leqq 1)$ by

$$\theta(t) = 0 \quad \text{if} \quad t \leqq T$$
$$= t - T \quad \text{if} \quad t > T.$$

Let $y(t) = \theta(t) + x(t)$. The control variable here is $z(t) = \pi \int_0^t \exp\{\pi s + y(t) - y(t-s) - s/2\}\, ds$ (Shiryaev (1963)), and the cost functional is $C_\xi(z) = \int_0^{\tau(z)} [cz(s)(1 + z(s))^{-1} - \gamma]\, ds$. This is the process whose operating characteristics concerning $C_\xi(z)$ are computed in Taylor (1967) and assumed to hold approximately for the discrete time process.

Our principal objective in this paper is to establish in a rigorous fashion the manner in which the discrete processes converge to the continuous process. We show that the measures induced on $\mathcal{C}$, the Borel sets of $C[0, 1]$, by the processes $\{y_n(t): 0 \leqq t \leqq 1\}$ and $\{z_n(t): 0 \leqq t \leqq 1\}$ converge weakly as $n \to \infty$ to the measures induced by $\{y(t): 0 \leqq t \leqq 1\}$ and $\{z(t): 0 \leqq t \leqq 1\}$ respectively. As a consequence, we also show that the distributions of $\tau(z_n)$ and $C_\xi^{(n)}(z_n)$ converge weakly to the distributions of $\tau(z)$ and $C_\xi(z)$, respectively.

2. Weak Convergence of the Process $\{y_n(\cdot)\}$ and $\{z_n(\cdot)\}$. In this section we shall discuss the weak convergence of the sequences $\{y_n(\cdot)\}$ and $\{z_n(\cdot)\}$ to $y(\cdot)$ and $z(\cdot)$ respectively. For background material the reader is referred to the book of Billingsley (1968), whose notation and terminology we shall follow. We begin with some definitions.

Let u be a measurable mapping (random element) from a probability space $(\Omega, \mathcal{B}, \mathcal{P})$ into $C[0, 1]$. The *distribution* of u is the probability measure $P = \mathcal{P}u^{-1}$ on $(C[0, 1], \mathcal{C})$. We say that a sequence $\{u_n\}$ of random elements of $C[0, 1]$ converges weakly to the random element u if $E[h(u_n)] \to E[h(u)]$ as $n \to \infty$ for all bounded continuous functions h on $C[0, 1]$ and we write $u_n \Rightarrow u$.

Now for the processes $\{x_n\}$ and x defined in Section 1 it is known by Donsker's theorem that $x_n \Rightarrow x$; cf. Donsker (1951) and Billingsley (1968, Section 10). Also it is easy to see that $P[T_n/n > t] = (1 - \pi_n)^{[nt]+1} \to e^{-\pi t} = P[T > t]$, or $T_n/n \to_L T$. Therefore using Skorohod (1956) one can construct a probability space and random variables $\{T_n^*, T^*\}$ with the same distribution functions as $\{T_n, T\}$, but for which $T_n^*/n \to T^*$ a.s. Hence $\rho(\theta_n^*, \theta^*) \to 0$ a.s., where the random elements θ_n^* and θ^* are defined like θ_n and θ but in terms of T_n^* and T^*. Thus we have $\theta_n^* \Rightarrow \theta^*$ and therefore $\theta_n \Rightarrow \theta$.[3] Since $y_n = x_n + \theta_n$ and the random elements x_n and θ_n are independent, an application of the continuous mapping theorem (cf., Billingsley (1968), Section 5) completes the proof of

THEOREM 1. $\quad y_n \Rightarrow y.$

To prove that $z_n \Rightarrow z$ we use the same method used to show $\theta_n \Rightarrow \theta$. Since $y_n \Rightarrow y$, there exists a probability space and random elements $\{y_n^*, y^*\}$ of $C[0, 1]$ having the same distributions as $\{y_n, y\}$ for which $\rho(y_n^*, y^*) \to 0$ a.s. Define z_n^* and z^* like z_n and z but in terms of y_n^* and y^*. In order to show that

[3] We are grateful to the referee for indicating this simpler proof of the fact that $\theta_n \Rightarrow \theta$.

$\rho(z_n^*, z^*) \to 0$ a.s., we introduce the estimates

$$z^*(t) \leqq \pi \sum_{j=0}^{[nt]} n^{-1} \exp\{\pi s_j' + y^*(t) - y^*(t - s_j') - s_j'/2\}$$

and

$$z_n^*(t) \geqq \pi \sum_{j=0}^{[nt]} n^{-1} \exp\{-(j+1)\ln(1-\pi_n) + y_n^*([nt]n^{-1}) - y_n^*([nt]n^{-1} - s_j) - s_j/2\} - O(n^{-1}),$$

where $s_j = j/n$, $s_j \leqq s_j' \leqq s_{j+1}$, and

$$\exp\{\pi s_j' + y(t) - y(t - s_j') - s_j'/2\} = \max_{s_j \leqq s \leqq s_{j+1}} [\exp\{\pi s + y(t) - y(t-s) - s/2\}].$$

Using the estimate $e^a - e^b \leqq (a - b)e^{a \vee b}$ we have

$$z^*(t) - z_n^* \leqq \pi \sum_{j=0}^{[nt]} n^{-1}[\pi(s_j' - s_j) + y^*(t) - y_n^*([nt]n^{-1}) - y^*(t - s_j') + y_n^*([nt]n^{-1} - s_j) - \tfrac{1}{2}(s_j' - s_j)] \times \exp\{O(1)\}.$$

Now using the fact that $\rho(y_n^*, y^*) \to 0$ a.s., we obtain the fact that $\max_{0 \leqq t \leqq 1} (z^*(t) - z_n^*(t)) \to 0$ a.s. A similar argument shows that the $\max_{0 \leqq t \leqq 1} (z_n^*(t) - z^*(t)) \to 0$ a.s. and thus $\rho(z_n^*, z^*) \to 0$ a.s. Consequently, $z_n^* \Rightarrow z^*$ and therefore we have

THEOREM 2. $$z_n \Rightarrow z.$$

3. Convergence in distribution of the cost functionals, $C_n(\zeta)$. Let ζ be a fixed positive number and let $\tau(y)$ be the first passage time to ζ for the path $y \varepsilon C[0, 1]$ as defined in Section 1. Since τ is continuous almost everywhere with respect to the measure induced by the process z, an application of the continuous mapping theorem yields $\tau(z_n) \to_L \tau(z)$. For $z \varepsilon C[0, 1]$ define the cost functional $C_\zeta(z)$ as in Section 1. Another application of the continuous mapping theorem yields $C_\zeta(z_n) \to_L C_\zeta(z)$. Next we show that $C_\zeta^{(n)}(z_n) - C_\zeta(z_n) \to_P 0$, where

$$C_\zeta^{(n)}(z_n) = \sum_{k=0}^{[n\tau(z_n)]} n^{-1}\{c[z_n(kn^{-1})/(1 + z_n(kn^{-1}))] - \gamma\}.$$

Using Riemann approximating sums one can easily show that $C_\zeta^{(n)}(y) \to C_\zeta(y)$ uniformly for y in a compact set of $C[0, 1]$. Let P_n be the measure on $\mathcal{C}$ induced by z_n. Since $z_n \Rightarrow z$ and C[0, 1] is a complete separable space, Prohorov's (1956) theorem states that for every $\epsilon > 0$ there exists a compact subset, K_ϵ, of $C[0, 1]$ for which $P_n(K_\epsilon) \geqq 1 - \epsilon$ for all n. Thus the

$$\begin{aligned} P_n\{|C_\zeta^{(n)}(z_n) - C_\zeta(z_n)| > \epsilon\} &= P_n\{|C_\zeta^{(n)}(z_n) - C_\zeta(z_n)| > \epsilon \mid z_n \varepsilon K_\epsilon\} \\ &\quad \times P_n(K_\epsilon) + P_n\{|C_\zeta^{(n)}(z_n) - C_\zeta(z_n)| > \epsilon \mid z_n \varepsilon K_\epsilon\} \times P_n(K_\epsilon^c) \\ &\leqq P_n\{|C_\zeta^{(n)}(z_n) - C_\zeta(z_n)| > \epsilon \mid z_n \varepsilon K_\epsilon\} + \epsilon. \end{aligned}$$

Since $C_\xi^{(n)}(\cdot) \to C_\xi$ uniformly on compact sets, we have $C_\xi^{(n)}(z_n) - C_\xi(z_n) \to_P 0$. Finally, an application of Theorem 4.1 of Billingsley (1968) yields

THEOREM 3. $C_\xi^{(n)}(z_n) \to_L C_\xi(z)$.

4. Remarks. We have exhibited a sequence of discrete time control processes which converges to a continuous time process. While we have worked with the time interval [0, 1], clearly the results are true for any interval [0, T]. In our particular case the practical conclusions are that when $(\mu_1 - \mu_0)/\sigma$ and π are small the continuous results of Taylor (1967) may be used to approximately evaluate a control limit for the discrete time Girshick and Rubin (1952) model. More generally, as Chernoff (1967) has noted, approximating discrete time problems by continuous time problems often makes it possible to apply the methods of partial differential equations to problems of sequential decisions. The path we have followed should be valid in justifying such approximations in other control problems.

REFERENCES

[1] BILLINGSLEY, P. (1968). *Weak Convergence of Probability Measures.* To be published by John Wiley and Sons.

[2] CHERNOFF, H. (1967). Optimal stochastic control. Paper delivered at the American Mathematical Society 1967 Summer Seminar on Mathematics of the Decision Sciences.

[3] DONSKER, M. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6**.

[4] GIRSHICK, M. A. and RUBIN, H. (1952). A Bayes approach to a quality control model. *Ann. Math. Statist.* **23** 114–125.

[5] PROHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Prob. Appl.* **1** 157–214.

[6] SHIRYAEV, A. N. (1963). On optimum methods in quickest detection problems. *Theor. Prob. Appl.* **8** 22–46.

[7] SKOROHOD, A. (1956). Limit theorems for stochastic processes. *Theor. Prob. Appl.* **1** 262–290. (English translation.)

[8] TAYLOR, H. M. (1967). Statistical control of a Gaussian process. *Technometrics* **9** 29–42.