SOME APPLICATIONS OF THE JIŘINA SEQUENTIAL PROCEDURE TO OBSERVATIONS WITH TREND

By Sam C. Saunders

Boeing Scientific Research Laboratories

Summary. Assume that each random variable of a sequence had a density which is a Pólya frequency function of order two. To this sequence we apply the Jiřina sequential procedure to determine a tolerance interval. In this paper we find some sufficient conditions on the type of trend permissible for this sequence which enable us to show that when the Jiřina procedure is used the sampling will stop sooner and the tolerance interval cover more of the population (in a stochastic sense) than would occur in the case without trend.

Similar considerations are shown to hold when the sequences of observations have densities which have non-decreasing hazard rates.

1. The Jiřina procedure. Let us define the Jiřina procedure when it is applied specifically to the real line and introduce notation to be used subsequently. Let $X_1, X_2, \cdots$ be a sequence of continuous independent real random variables (r.v.'s) not necessarily identically distributed. The triple (η, k, D) defines a Jiřina procedure where η, k are positive integers and D is a sequence of functions determined in the following manner:

For given $n \geqq \eta$ let

$$X_{1,n} < X_{2,n} < \cdots < X_{n,n} \tag{1.1}$$

be the ordered r.v.'s determined almost surely by $X_1, \cdots, X_n$, and for notational convenience write $X^n = (X_1, \cdots, X_n)$. Then for each integer $n \geqq \eta$ we have D defined by

$$D(X^n) = \bigcup_{(i)} \{x : X_{i-1,n} < x \leqq X_{i,n}\} \tag{1.2}$$

where the union is over a preassigned set of exactly $(n - \eta)$ i's with the proviso that

$$D(X^n) \subset D(X^{n+1}) \qquad n \geqq \eta. \tag{1.3}$$

We continue sampling until the stopping event

$$B(X^{n+k}) = [D(X^n) = D(X^{n+k})] \tag{1.4}$$

occurs. In view of this definition we have also

$$B(X^{n+k}) = \bigcap_{j=1}^{k} [X_{n+j} \varepsilon D(X^n)]. \tag{1.5}$$

Let $N(X)$ be the random sample size associated with D which we define as

Received September 4, 1962.

the number of observations n drawn when $B(X^n)$ occurs for the first time. That is, letting superscript c denote complementation,

$$[N(X) = n] = \bigcap_{i=\eta}^{n-1} B^c(X^i) \cap B(X^n). \tag{1.6}$$

Thus $D(X^{N(X)})$ is the Jiřina sequential region determined by η, k, D. If P is a fixed probability measure, the *coverage* of the region (with respect to P) is the r.v. on the unit interval defined by

$$Q(X) = P[D(X^{N(X)})]. \tag{1.7}$$

The case where the X_i's are identically distributed has been studied previously, and it is known that

$$\Pr[Q(X) \geqq \beta] = (1-\beta)^\eta \exp\left[-\eta \sum_1^k \beta^j/j\right] \tag{1.8}$$

for which approximations are known and tabulations have been made. Further it is known in this case that we have the very good approximation with asymptotic equality as $k \to \infty$

$$EN(X) \doteq \eta + kS_\eta \qquad n \geqq 1 \tag{1.9}$$

where S_η is a constant which depends only upon η.

In this regard one is referred to [3] and [4] for a discussion of the properties of (1.8) and to a translation of Jiřina's original paper in [2] for its development.

2. Application to life testing. Suppose we are observing the life lengths of successively produced components with the initial manufacturing refinements and design improvements being continually incorporated in their construction. We assume the components are being improved but the degree of improvement is not known exactly or cannot be quantified in terms of the life length in service. We also assume that the degree of improvement of the component in the production run will eventually reach a plateau of development beyond which it will not progress.

Suppose we apply the Jiřina procedure to this sequence of observations. For instance, we might agree to stop sampling when we have obtained for the first time 15 observations which exceed the minimum life length obtained in the sample. From the tables and results in [3] we know that if in fact there was no trend, i.e., the life lengths were not being improved, then we could expect to stop in 27.7 observations and we would have a distribution of the coverage Q which has the values given in the display below. On the other hand, if there is in fact a very decided improving trend, it is not inconceivable that we could stop early, say less than twenty observations. Suppose for the sake of comparison we consider the coverage Q_1 of the minimum of a fixed sample of size $n = 25$ in the case of no trend. Now Q_1 has distribution $\Pr[Q_1 \geqq \beta] = 1 - \beta^n$

	$\Pr(Q_1 \geqq x)$	$\Pr(Q \geqq x)$
$x = .8$	.99	.99
$x = .9$	.93	.97
$x = .95$	.72	.73
$x = .99$	.22	.24

It is the purpose of this note to set out certain conditions under which in the case with trend the coverage is improved even though the sample size is decreased. The fact that under certain conditions the coverage of a sample with trend is stochastically larger than the sample without and the sample size is stochastically smaller for a sample with trend than a sample without is the rationale for the use of the Jiřina procedure.

It should be noted that these results do not state the conditions under which the Jiřina procedure is superior to a non-sequential procedure. Exact results in this direction would depend upon making explicit enough assumptions about the trend that the distribution of coverage and the expected sample size can be computed and compared for the two procedures. The results we have merely suggest the superiority of the Jiřina procedure in the case of trend.

We now state more precisely our contention.

Let $Y_1, \cdots, Y_n, \cdots$ be a sequence of independent non-negative r.v.'s, which we may regard as representing life lengths, with Y_n having continuous distribution F_n, $n = 1, 2, 3, \cdots$.

We assume that $Y_1 \leqq Y_2 \leqq \cdots$, the inequalities meant in the stochastic sense, which is equivalent to

$$F_1 \geqq F_2 \geqq \cdots. \tag{2.1}$$

Let $X_1, \cdots, X_n, \cdots$ be a sequence of independent r.v.'s on the positive real line indentically distributed with continuous distribution F, and we further assume that $\lim_{n\to\infty} F_n = F$.

It follows from the above assumptions that there exists a sequence of order preserving transformations $\tau_1, \tau_2, \cdots$ such that $\tau_i(Y_i) = X$ where we mean stochastic equality to a r.v. X with distribution F, that is, $F_i = F\tau_i$. It further follows from (2.1) that $\tau_i \geqq \tau_{i+1}$ $i = 1, 2, \cdots$.

We assume additionally that the functions τ_i are differentiable and $\tau_i' \leqq \tau_{i+1}'$, $i = 1, 2, \cdots$.

We make the intuitively appealing assumption:

Both the X's and the Y's have an increasing hazard (failure) rate.

Definition. A distribution F with density f is IHR iff $\delta = f/(1 - F)$ is non decreasing. Under the same condition we say the density f is IHR. In either case a density exists and we say the random variable has a (weakly) *increasing hazard rate.*

We do not attempt to justify this assumption but it is a common one. We now state

THEOREM 1. *If the Jiřina procedure* $(1, k, D)$ *is applied to both the* X *and* Y *sequences where* $D(X^n) = [X_{1,n}, \infty]$, *then the random sample size* N_Y *and the coverage with respect to* F, *say* $Q(Y)$, *for the* Y *sequence are, respectively, stochastically smaller and larger than the sample size* N_X *and coverage with respect to* F, *say* $Q(X)$, *associated with the* X *sequence.*

The proof will appear as a consequence of the general results to follow.

3. The main result. Let $X_1, \cdots, X_n, \cdots$ be independent random variables on the real line with X_n having distribution F_n. Let t be an order preserving transformation, i.e., continuous and strictly increasing, and for some fixed integer m we define

$$Y_j = t(X_j) \qquad j = 1, \cdots, m$$
$$Y_{m+j} = X_{m+j} \qquad j = 1, 2, \cdots. \tag{3.1}$$

Let us write $t(X^n) = (tX_1, \cdots, tX_n)$ for convenience.

Let D be a Jiřina sequential tolerance function as defined in Part 2 with given parameters η, k. We make the assumption that

$$D(tX^n) = t[D(X^n)] \tag{3.2.1}$$

and

$$t[D(X^n)] \supset D(Y^n) \supset D(X^n). \tag{3.2.2}$$

We have

LEMMA 1. *The stopping events are invariant under* t, *i.e.,*

$$B(tX^n) = B(X^n). \tag{3.3}$$

PROOF. This follows immediately from (2.6) and (3.2).

LEMMA 2. *Always*

$$B(X^{n+k}) \subset B(Y^{n+k}) \qquad n = 1, 2, \cdots. \tag{3.3.1}$$

PROOF. We have three cases (i) $n + k \leqq m$, (ii) $n \leqq m < n + k$, (iii) $m < n$.

Case (i): By (3.3) we have $B(X^{n+k}) = B(tX^{n+k}) = B(Y^{n+k})$.

Case (ii): By definition $B(X^{n+k}) = [D(tX^n) = D(tX^{n+k})]$, but by (3.2.1) $D(tX^{n+k}) \supset D(Y^{n+k})$. Thus since $D(tX^n) = D(Y^n)$, we have $B(X^{n+k})$ which implies $B(Y^{n+k})$.

Case (iii): Now $B(X^{n+k}) = \bigcap_{j=1}^{k} [X_{n+j} \varepsilon D(X^n)]$ by (2.7), but $D(Y^n) \supset D(X^n)$ by (3.2.1). Thus $B(X^{n+k}) \subset \bigcap_{j=1}^{k} [X_{n+j} \varepsilon D(Y^n)] = \bigcap_{j=1}^{k} [Y_{n+j} \varepsilon D(Y^n)]$.

We now have quite generally

THEOREM 2. *Let* $N(Y)$ *and* $N(X)$ *be the random sample sizes for the Jiřina procedure applied to the sequences* $Y_1, Y_2, \cdots$, *and* $X_1, X_2, \cdots$, *respectively, defined above. Then* $N(Y)$ *is stochastically smaller than* $N(X)$.

PROOF. By (3.3.1) we have $B(X^n) \subset B(Y^n)$. Hence taking complements we thus have

$$\bigcap_{i=1}^{n} B^c(X^i) \supset \bigcap_{i=1}^{n} B^c(Y^i), \tag{3.3.2}$$

where $B(X^i) = \phi$ for $i \leqq \eta + k$. Thus we have $[N(X) \geqq n] \supset [N(Y) \geqq n]$, which proves the result.

We now must discuss more specifically conditions under which we can have $Q(X)$ stochastically smaller than $Q(Y)$.

DEFINITION. A non-negative function f is a *Pólya frequency of order* 2 (PF_2) iff $f(x) = e^{-\psi(x)}$ where ψ is convex.

We now state without proof some well known facts.

Fact 1: f is IHR iff $1 - F$ is PF_2 iff F is PF_2.

Fact 2: f is PF_2 implies f is IHR.

Fact 3: F is IHR iff F_D is IHR where $F_D(x) = 1 - F(-x)$.

These statements are for example proved by Barlow, Marshall and Proschan in [1] and we follow their notation.

We now establish

LEMMA 3. *If H, G are non-decreasing non-negative functions and F is an* IHR *distribution and t is a differentiable function such that $t(x) \geqq x$ and $t'(x) \geqq 1$ for every x, then for each ξ*

$$\int_{-\infty}^{t^{-1}(\xi)} H(y)G(ty)\,dF(y) \leqq \int_{-\infty}^{\xi} H(y)G(y)\,dF(y). \tag{3.4}$$

PROOF. Since G can be approximated by an increasing sequence of linear combinations of increasing non-negative step functions, it is sufficient to prove (3.4) with G replaced by $c(\cdot, a)$ where $c(x, y) = 1$ if $x \geqq y$ and zero elsewhere. Thus we need prove that

$$\int_{t^{-1}(a)}^{t^{-1}(\xi)} H(y)\,dF(y) \leqq \int_{a}^{\xi} H(y)\,dF(y).$$

But similarly it is sufficient to replace H by $c(\cdot, b)$. We then have two cases (3.4.1) $b \geqq a$ and (3.4.2) $t^{-1}(a) \leqq b < a$. So we need prove

$$\int_{b}^{t^{-1}(\xi)} dF(y) \leqq \int_{b}^{\xi} dF(y) \tag{3.4.1}$$

which is obvious since $t^{-1}(\xi) \leqq \xi$. We now prove

$$\int_{b}^{t^{-1}(\xi)} dF(y) \leqq \int_{a}^{\xi} dF(y) \tag{3.4.2}$$

which will be done since $F(b) \leqq F(a)$, if we show

$$\int_{t^{-1}(a)}^{t^{-1}(\xi)} dF(x) = Ft^{-1}(\xi) - Ft^{-1}(a) \leqq \int_{a}^{\xi} dF(x) = F(\xi) - F(a).$$

Hence it is sufficient to prove that $Ft^{-1}(x) - F(x)$ is a non-increasing function, which disposes of the third case $b < t^{-1}(a)$ also. Let $\Delta(x) = -\ln [1 - F(x)]$, we see $Ft^{-1}(x) - F(x) = [1 - F(x)][1 - \exp\{-\Delta t^{-1}(x) + \Delta(x)\}]$. Thus we need only establish that $\Delta(x) - \Delta t(x)$ is non-increasing, but $\delta = \Delta'$ and $\delta(x) \leqq \delta[t(x)]t'(x)$, which proves the result since δ is non-decreasing.

COROLLARY 2. *If H, G are non-increasing non-negative functions and F is* IHR *and t is a differentiable function such that $t(x) \leqq x$ and $t'(x) \geqq 1$ for each x, then for each ξ*

$$(3.5) \qquad \int_{t^{-1}(\xi)}^{\infty} H(y)G(ty)\,dF(y) \leqq \int_{\xi}^{\infty} H(y)G(y)\,dF(y).$$

PROOF. In (3.4) replace y by $-y$, replace $t(x)$ by $-t(-x)$, and replace ξ by $-\xi$, and the result follows by a change of designation utilizing Fact 3.

We now prove

LEMMA 4. *If $X_1, \cdots, X_n$ are* r.v.'s *with densities $f_1, \cdots, f_n$ which are* PF_2 *then both g and h are* PF_2 *where*

$$(3.6) \qquad g(x) = \Pr [x < X_n < \cdots < X_1]$$

$$(3.6.1) \qquad h(x) = \Pr [X_1 < \cdots < X_n < x].$$

PROOF. By definition

$$g(x) = \int_x^{\infty} \cdots \int_{x_4}^{\infty} \left\{ \int_{x_3}^{\infty} [1 - F_1(x_1)]\, f_2(x_2)\, dx_2 \right\} f_3(x_3)\, dx_3 \cdots f_n(x_n)\, dx_n\,.$$

Keeping in mind Facts 1 and 2, f_1, being PF_2, is IHR and hence $1 - F_1$ is PF_2. Thus since f_2 is PF_2 the product $(1 - F_1)f_2$ is PF_2 and thus IHR. Therefore the integral in braces is PF_2. This argument repeats.

To prove (3.6.1) simply replace X_i by $-X_i$ and x by $-x$ and keep Fact 3 in mind.

We are now in a position to prove

THEOREM 3. *If $X_1, X_2, \cdots$ is a sequence of* r.v.'s *with densities which are* PF_2 *and $Y_1, Y_2, \cdots$ is a sequence determined by some transformation t as in* (3.1), *then sufficient conditions that $Q(X)$ be stochastically smaller than $Q(Y)$ are*

(a) *t is a differentiable function such that both $t(x) \geqq x$, $t'(x) \geqq 1$ for all x and for some positive integer η $D(X^n) = (-\infty, X_{n-\eta+1,n})$ for all $n \geqq \eta$ or*

(b) *t is a differentiable function such that both $t(x) \leqq x$, $t'(x) \geqq 1$ for all x and for some positive integer η $D(X^n) = (X_{\eta,n}, \infty)$ for all $n \geqq \eta$.*

PROOF. One checks that for both Cases (a) and (b) we have (3.2.1) and (3.2.2) satisfied. It is necessary to show that $\Pr [Q(Y) \leqq \beta] \leqq \Pr [Q(X) \leqq \beta]$ for all $\beta \in (0, 1)$ but by setting $S_n(Y) = [Q(Y^{n+k}) \leqq \beta, N(Y) = n + k]$ for $n \geqq \eta$ it is sufficient to show that $\Pr S_n(Y) \leqq \Pr S_n(X)$ for $n \geqq \eta$.

Referring to a result in [3] we have that

$$S_n(Y) = \left[Q(Y^n) \leqq \beta, \bigcap_{i=\eta}^{n-1} B^c(Y^i), Y_n \notin D(Y^{n-1}), \bigcap_{j=1}^{k} Y_{n+j} \in D(Y^n) \right].$$

In words this formula merely says that stopping at exactly $(n + k)$ observations and having a coverage of less than β can be accomplished only by not having stopped on or before $(n-1)$ observations, the nth observation falls outside the tentative region and the succeeding k observations fall inside the region constructed after n observations, which has a coverage of less than β.

Again we consider three cases (i) $n+k \leqq m$ (ii) $n \leqq m < n+k$ and (iii) $m < n$.

Case (i): $n+k \leqq m$. By (3.1) we have immediately that

$$S_n(Y) = \left[P(tDX^n) \leqq \beta, \bigcap_{i=\eta}^{n-1} B^c(X^i), X_n \not\varepsilon D(X^{n-1}), \bigcap_{j=1}^{k} X_{n+j} \varepsilon D(X^n)\right],$$

and by (3.2.2) we have $S_n(Y) \subset S_n(X)$.

Case (ii): $n \leqq m < n+k$. We have

$$S_n(Y) = \left[P(tDX^n) \leqq \beta, \bigcap_{i=\eta}^{n-1} B^c(X^i), X_n \not\varepsilon D(X^{n-1}), \bigcap_{j=n+1}^{m} X_j \varepsilon D(X^n), \bigcap_{j=m+1}^{n+k} X_j \varepsilon tD(X^n)\right]. \tag{3.7}$$

Now we must be more specific. In what follows we shall consider the Case (a), $D(X^n) = (-\infty, X_{n-\eta+1,n})$, but every step may be duplicated for Case (b). Thus (3.7) becomes

$$S_n(Y) = \left\{F(tX_{n-\eta+1,n}) \leqq \beta, \bigcap_{i=n}^{n-1} B^c(X^i), [X_n > X_{n-\eta,n-1}], \bigcap_{j=m+1}^{n+k} [X_j < tX_{n-\eta+1,n}], \bigcap_{j=n+1}^{m} [X_j < X_{n-\eta+1,n}]\right\}.$$

Since for $S_n(Y)$ to have occurred we must have had occur one of $[X_{i_1} < X_{i_2} < \cdots < X_{i_n}] = K_{(i)}$ where (i) is a certain one of the permutations of the indices $(1, \cdots, n)$. Thus let $F^{-1}(\beta) = \xi$

$$\Pr S_n(Y) = \sum_{(i)} \int_{-\infty}^{t^{-1}(\xi)} \Pr[X_{i_1} < \cdots < X_{i_{n-\eta}} < y] \cdot \Pr[y < X_{i_{n-\eta+2}} < \cdots < X_{i_n}] \cdot \prod_{j=m+1}^{n+k} F_j(ty) \cdot \prod_{j=n+1}^{n} F_j(y)\, dF_{i_{n-\eta+1}}(y) \tag{3.7.1}$$

but (3.7.1) of the general form $\int_{-\infty}^{t^{-1}(\xi)} H(y)G(ty)\cdot f(y)\,dy$, where H and G are monotone increasing and f is PF_2 so that Lemma 3 applies, and this proves the result.

We remark that if $\eta = 1$ the second probability statement under the integral in (3.7.1) is gone, in which case Lemma 3 applies with only the assumption of IHR distributions instead of PF_2 densities.

Case (iii): $m < n$. We have by definition

$$S_n(Y) = \left\{P[D(Y^n)] \leqq \beta, \bigcap_{i=\eta}^{n-1} B^c(Y^i), X_n \not\varepsilon D(Y^{n-1}), \bigcap_{j=1}^{k} X_{n+j} \varepsilon D(Y^n)\right\}$$

but utilizing (3.3.2), and by (3.2.2) that $D(Y^{n-1}) \supset D(X^{n-1})$, there follows $[X_n \not\in D(Y^{n-1})] \subset [X_n \not\in D(X^{n-1})]$. Since we also have $tD(X^n) \supset D(Y^n)$, we obtain

$$S_n(Y) \subset \left\{P[tD(X^n)] \leqq \beta, \bigcap_{1=\eta}^{n-1} B^c(X^i), X_n \not\in D(X^n), \bigcap_{j=1}^{k} X_{n+j} \in tD(X^n)\right\},$$

but by Case (ii) we see that the probability of the right hand side does not exceed Pr $[S_n(X)]$. This concludes the proof.

THEOREM 4. *If* $X_1, X_2, \cdots$ *is a sequence of independent identically distributed* r.v.'s *with* PF$_2$ *densities, and* $\tau_1, \tau_2, \cdots$ *is a sequence of order preserving differentiable transformations such that* $\tau_1' \leqq \tau_2' \leqq \cdots$ *and either* (3.8.1) *or* (3.8.2) *holds where*

$$(3.8.1) \qquad \tau_i \leqq \tau_{i+1}, \qquad \tau_i(x) \to x,$$

$$(3.8.2) \qquad \tau_i \geqq \tau_{i+1}, \qquad \tau_i(x) \to x,$$

then

$$(3.8.3) \qquad Y_1 = \tau_1^{-1}(X_1), \qquad Y_2 = \tau_2^{-1}(X_2), \cdots$$

is a sequence of r.v.'s *for which, in Case* (3.8.1), *the Jiřina upper tolerance interval* (*Case* (3.8.2) *the lower tolerance interval*) *can be established with both stochastically smaller sample size and stochastically larger coverage than for the X-sequence.*

THEOREM 5. *In Theorem* 4 *the assumption of* PF$_2$ *densities can be relaxed to* IHR *densities, if the tolerance intervals are restricted to the minimum and maximum of the sample in Case* (3.8.1) *and* (3.8.2) *respectively.*

PROOF. Let $t_1, t_2, \cdots$ be a sequence of order preserving transformation such that $\lim_{m\to\infty} t_m t_{m-1} \cdots t_j = \tau_j^{-1}$ (juxtaposition indicates composition). Thus there follows

$$(3.8.4) \qquad \tau_j^{-1} = \tau_{j+1}^{-1}(t_j),$$

or equivalently

$$(3.8.5) \qquad \tau_{j+1}(\tau_j^{-1}) = t_j .$$

Thus (3.8) is equivalent with $t_j' \geqq 1$. To see this use (3.8.5);

$$t_j' \geqq 1 \text{ iff } \tau_{j+1}'[t_j^{-1}(x)]/\tau_j'[\tau_j^{-1}(x)] \geqq 1.$$

Since the τ_j's are order preserving, they must have non-negative derivatives, and the result is proved. But further (3.8.1) is equivalent with $\tau_{j+1}^{-1} \leqq \tau_j^{-1}$ which is equivalent with $t_j(x) \geqq x$ by (3.8.5). Similarly (3.8.2) iff $t_j(x) \leqq x$. Thus we may apply Theorem 3 after applying each one of the t_j's, and since at each step the distribution of coverage and sample size are monotone increasing and decreasing, respectively, this proves the result.

REFERENCES

[1] Barlow, R. E., Marshall, A. W. and Proschan, F. (1963). Properties of probability distributions with monotone hazard rate. *Ann. Math. Statist.* **34** 375–389.

[2] Jiřina, Miroslav (1952 and 1953). Sequential estimation of distribution-free tolerance limits (in Russian). *Czechoslovak Math. J.* **77** 211–232 and correction **78** 283. (Also (in English) in *Selected Translations in Mathematical Statistics and Probability* **1** (1961) 145–156. Amer. Math. Soc.)

[3] Saunders, Sam C. (1963). On the sample size and coverage for the Jiřina sequential procedure. *Ann. Math. Statist.* **34** 847–856.

[4] Saunders, Sam C. (1960). Sequential tolerance regions. *Ann. Math. Statist.* **31** 198–216.