SOME BEST CHOICE PROBLEMS
WITH PARTIAL INFORMATION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Joseph David Petruccelli

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 1978

To Sherry and my parents.

ACKNOWLEDGEMENTS

I would like first of all to express thanks to my major professor Burgess Davis for his guidance in this undertaking.

I thank Professor S. M. Samuels who also contributed to my knowledge of this topic.

Last, but never least, I thank my wife Sherry for sharing the bad times as well as the good.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Petruccelli, Joseph David. Ph.D., Purdue University, August, 1978. Some Best Choice Problems with Partial Information. Major Professor: Burgess Davis.

This thesis deals with variations of the Secretary Problem.

In Chapter I the following problem is considered: let $\mathcal{F}$ be a family of continuous distribution functions and let n i.i.d. $F \in \mathcal{F}$ observations be taken sequentially with the object of selecting the largest observation. At time j the j^{th} observation must be chosen (and the process terminated) or rejected (and the process continued). No knowledge of the future is allowed, no recall of rejected observations is possible and one observation must be selected.

Let the versions of this problem obtained when the family $\mathcal{F}$ is all distribution functions, a single distribution function and any family distinct from these two be denoted respectively as the No Information, Full Information and Partial Information Best Choice Problem (hereafter N.I., F.I. and P.I. Problem).

In Chapter II the P.I. Problem is investigated when $\mathcal{F}$ is a location, scale or location-scale parameter family of distributions. For each such family a sufficient condition is obtained for there to exist stopping rules which are

asymptotically as good as the best rule in the F.I. Problem. It is of particular interest that this result applies when $\mathcal{F}$ is taken to be $\mathcal{N}$, the family of all normal distributions.

In Chapter III attention is focused on the P.I. Problem for distributional family $\mathcal{N}$. It is shown that the result of Chapter II holds for a more sophisticated stopping rule. The rules considered in Chapters II and III are invariant rules. In addition to their optimal asymptotic properties such rules are of interest since in each P.I. Problem considered the class of invariant rules contains a minimax rule. This follows from a version of the Hunt-Stein Theorem and means that a best invariant rule is minimax. The second part of Chapter III investigates a best invariant rule for the P.I. Problem defined by the family $\mathcal{N}$.

Chapter IV treats the P.I. Problem defined by the location parameter family of $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ distributions. A best invariant (hence minimax) rule is found and for small n the probability of choosing the largest observation using this rule is obtained. It is shown that this probability is asymptotically strictly between the corresponding asymptotic values for the best rules in the N.I. and F.I. Problems.

CHAPTER I

INTRODUCTION AND HISTORY

This thesis deals with variations of the Secretary Problem.

This chapter gives an historical development outlining the past variations and generalizations of the Secretary Problem which most directly relate to our own work. The chapter concludes with an outline of the main results of this thesis.

1. The Secretary Problem

The Secretary Problem has been known under a variety of names, two of them being the Dowry Problem and the Beauty Contest Problem. When known as the Secretary Problem it is stated as follows: n applicants for a secretarial position are interviewed one at a time. The applicants appear in a random order and as each is interviewed he is assigned a rank (rank 1 meaning he is the best) among those interviewed so far. At this point a decision must be made to hire him and terminate the procedure or to reject him and interview the next applicant.

The decision must be based only on the information available to the interviewer at the end of an interview.

Namely, he knows n, the total number of applicants, k, the number interviewed up to and including the applicant being considered, and the relative rank of the present applicant among those k. Once rejected an applicant may not again be considered for the position. One applicant must be chosen, meaning the n^{th} will automatically be chosen if the $n-1^{st}$ is rejected.

The object is to choose the best applicant. For this reason the Secretary Problem and variations of it for which the payoff to the observer is 1 if he chooses the best among the objects sampled and 0 otherwise, are called Best Choice Problems.

The Secretary Problem is solved by finding a stopping rule which maximizes the probability of choosing the best among the n applicants. This optimal rule is well known to be: reject the first k(n)-1 individuals and select the first applicant thereafter with relative rank 1; if no applicant among the first n-1 is chosen, choose applicant n. The value of k(n) is given by

$$(1.1) \qquad k(n) = \min\{j : \frac{1}{j} + \frac{1}{j+1} + \ldots + \frac{1}{n-1} \leq 1\}.$$

It can be shown that as $n \to \infty$

$$\frac{k(n)}{n} \to e^{-1}$$

and the probability of selecting the best applicant also approaches e^{-1}.

The origins of the Secretary Problem are hazy. Frederick Mosteller [7] heard of it in 1955 from Andrew Gleason who himself heard it elsewhere. A version first appeared in print under the name Googol in Martin Gardner's column in the February, 1960 issue of Scientific American [6], [12]. Dynkin and Yushkevich [4] offer a particularly elegant treatment of the problem.

2. The Full Information Best Choice Problem

Consider the following optimal stopping problem: let

F be a continuous cumulative distribution function (c.d.f.)

$X_1, X_2, \ldots \sim$ i.i.d. F be the observations

$L_i = \max\{X_1, \ldots, X_i\} \quad 1 \leq i \leq n$.

The X's are observed sequentially. The object (as always for Best Choice Problems) is to choose for each n the largest among $X_1, \ldots, X_n$. Thus we must find a stopping rule $\sigma_n \leq n$ for the X's such that

$$P(X_{\sigma_n} = L_n) = \sup_{\sigma \in S_n} P(X_\sigma = L_n)$$

where S_n is the class of all stopping rules σ for the X's such that $\sigma \leq n$.

If F is completely known the problem stated above is called the Full Information Best Choice Problem, hereafter abbreviated to "F.I. Problem". If F is completely unknown the problem is equivalent to the Secretary Problem. In view of the above terminology we will refer to this case as the No Information Best Choice Problem, which will be abbreviated to "N.I. Problem".

Gilbert and Mosteller [7] solved the F.I. Problem. We will now outline a solution by the standard method of backward induction. Assume F is known. Since the problem remains unaltered if we observe $F(X_1),\ldots,F(X_n)$ we may assume F is the c.d.f. of a U[0,1] distribution.

Let

$$\mathcal{S}_{j,n} = \{\tau\in\mathcal{S}_n : \tau>j\} \qquad \text{(note that } \mathcal{S}_{0,n} = \mathcal{S}_n)$$

$$\underline{X}_j = (X_1,\ldots,X_j) \qquad j = 1,2,\ldots$$

and define

$$U_{j,n}(\underline{X}_j) = P(X_j = L_n | \underline{X}_j) ,$$

$$V_{j,n}(\underline{X}_j) = \sup_{\sigma\in\mathcal{S}_{n,j}} E(U_\sigma(\underline{X}_\sigma) | \underline{X}_j) .$$

$U_{j,n}$ is the probability, given our knowledge at time j, that by choosing X_j we choose the largest of the X's. $V_{j,n}$ is the best we can expect to do if we reject X_j and continue sampling. The method of backward induction gives $V_{j,n}$ recursively as

$$V_{n,n} = U_{n,n}$$

$$\text{(1.2)} \quad V_{j,n}(\underline{X}_j) = E(\max\{U_{j+1,n}(\underline{X}_{j+1}), V_{j+1,n}(\underline{X}_{j+1})\} | \underline{X}_j) .$$

Then $$V_{0,n} = E(\max\{U_{1,n}(X_1), V_{1,n}(X_1)\})$$

$$= \sup_{\sigma\in\mathcal{S}_n} P(X_\sigma = L_n) ,$$

and the optimal stopping rule σ_n is given by

$$(1.3) \qquad \sigma_n = \min\{n, \min_{j\geq 1}\{j: U_{j,n}(\underline{X}_j) > V_{j,n}(\underline{X}_j)\}\} .$$

It can be shown that there is a sequence $1 > d_{1,n} > d_{2,n} > \ldots > d_{n,n} = 0$ such that

$$(1.4) \qquad \sigma_n = \min\{n, \min_{j\geq 1}\{j: X_j = L_j, X_j > d_{j,n}\}\} .$$

It can also be shown that

$$V_{0,n} \downarrow \alpha_0 \doteq .58 \quad \text{as} \quad n \to \infty .$$

Note that the solution to the N.I. Problem is given by (1.4) where $d_{1,n} = \ldots = d_{k(n)-1,n} = 1$, $d_{k(n),n} = \ldots = d_{n,n} = 0$, and $k(n)$ is given by (1.1). For notational purposes we will speak of the optimal rule for the F.I. (N.I.) Problem of length n as the F.I. (N.I.) Rule of length n or, where the meaning is clear, as the F.I. (N.I.) Rule.

3. Other Variations and Generalizations

The Secretary Problem has spawned numerous variations and generalizations. A few will be mentioned here.

One generalization of the Secretary Problem is known as the Rank Problem. In this problem the payoff to the observer is a non-increasing function q of the absolute rank of the chosen applicant among all applicants. Notice that the Secretary Problem corresponds to the special case of $q(x) = I_{\{1\}}(x)$ (where I_A is the indicator function of the set A).

Chow et al [2] solved this problem for linear q, the case in which it is desired to minimize the expected rank of the chosen applicant. For general q the problem is more difficult but some work has been done (Mucci [13], [14]) giving asymptotic results.

Motivated by a problem of Cayley [1], Moser [11] studied the following variation of the F.I. Problem: n i.i.d. U[0,1] random variables $X_1,\ldots,X_n$ are observed sequentially. If X_j is chosen, a payoff of X_j is obtained. To solve this problem Moser found a stopping rule $\tau \in S$, the class of all stopping rules for the X's, for which

$$EX_\tau = \sup_{\sigma \in S} EX_\sigma .$$

Guttman [8] extended Moser's results to the case in which $X_1,\ldots,X_n \sim$ i.i.d. F with F a known continuous c.d.f.

4. The Partial Information Best Choice Problem

Sakaguchi [15] and later DeGroot [3] considered the following problem: let $X_1, X_2, \ldots$ be i.i.d. random variables from a $N(\mu,1)$ distribution with unknown μ. The X's are observed sequentially and at a constant cost c per observation. If observation j is chosen the payoff is $X_j - cj$. It is desired to find a stopping rule that in some sense maximizes the expected payoff.

Both Sakaguchi and DeGroot assumed a normal prior on μ and found a sequential Bayes rule as their solution.

If in this problem it is desired to choose the largest among the first n X's then we have a Best Choice Problem which is an analogue of the F.I. and N.I. Problems. We now make the natural generalization of this example.

By the Partial Information Best Choice Problem (the P.I. Problem) we mean the Best Choice Problem (as formulated in the first paragraph of Section 2) in which it is known only that $F \in \mathcal{F}$, a family of c.d.f.'s. To distinguish the P.I. Problem from the F.I. and N.I. Problems we will require that $\mathcal{F}$ have more than one member and that $\mathcal{F}$ is not the family of all c.d.f.'s. Typically $\mathcal{F}$ will be a family of c.d.f.'s $\{F_\theta\}_{\theta \in \Theta}$ indexed by a parameter or vector of parameters. The families $\mathcal{F}$ we will consider all have this form.

One approach to the P.I. Problem is to assume prior knowledge (or a degree of belief) about the parameter θ. Stewart [17] investigated the P.I. Problem defined by the family $\mathcal{G}' = \{G_{\alpha,\beta}\}_{-\infty<\alpha<\beta<\infty}$ where $G_{\alpha,\beta}$ is the c.d.f. of the $U[\alpha,\beta]$ distribution, via this Bayesian approach.

In this thesis we investigate the P.I. Problem for other distributional families. In the last section of Chapter II we consider the P.I. Problem for the family $\{F_{\mu,\sigma}\}_{\substack{\mu \in \mathbb{R} \\ \sigma > 0}}$ where $F_{\mu,\sigma}$ is the c.d.f. of the $N(\mu, \sigma^2)$ distribution.

If we know the true value of the parameter (μ,σ) then the P.I. Problem becomes a F.I. Problem. If σ_n is the F.I. Rule of length n then (see Section 2)

$$\lim_{n\to\infty} P(X_{\sigma_n} = L_n) = \alpha_o \doteq .58 \; .$$

We show that, making no assumptions about the parameter (μ,σ), there are stopping rules τ_n with $\tau_n \leq n$ for which

$$\lim_{n\to\infty} P_{\mu,\sigma}(X_{\tau_n} = L_n) = \alpha_o$$

uniformly in (μ,σ). Thus, asymptotically at least, a Bayesian approach is unnecessary for this problem.

It is not surprising that the rules τ_n are invariant under location and scale changes and in Chapters II - IV we pursue the study of invariant rules. Such study is further justified by the fact that a version of the Hunt-Stein Theorem (Kiefer [9]) insures that a best invariant rule is also minimax for the P.I. Problems we consider, namely those defined by location, scale and location-scale parameter families of distributions.

In Chapter II we give sufficient conditions for there to exist a sequence of invariant rules $(\tau_1,\tau_2,\ldots)$ with $\tau_n \leq n$ such that

(1.5) $$\lim_{n\to\infty} P(X_{\tau_n} = L_n) = \alpha_o$$

in the P.I. Problem defined by arbitrary location, scale and location-scale parameter families of distributions. The conditions involve the shape of the upper tails of distributions in the relevant family. The proof of the sufficient conditions is constructive in that sequences of rules satisfying (1.5) are explicitly given. Applications of these sufficient conditions to families with different upper tail shapes (including the appropriate families of normal distributions) is made.

In Chapter III we restrict our attention to the P.I. Problem for the families $\{F_{\mu,1}\}_{\mu\in\mathbb{R}}$ and $\{F_{\mu,\sigma}\}_{\substack{\mu\in\mathbb{R}\\ \sigma>0}}$ where $F_{\mu,\sigma}$ is the c.d.f. of the $N(\mu,\sigma^2)$ distribution. In each case we consider a more sophisticated sequence of invariant rules than those of Chapter II and prove that these rules have the same asymptotic property (1.5). In the second part of Chapter III we discuss best invariant rules for these P.I. Problems.

Chapter IV deals with the P.I. Problem for the location parameter family $\mathcal{G} = \{G_\theta\}_{\theta\in\mathbb{R}}$ where G_θ is the $U[\theta-\frac{1}{2}, \theta+\frac{1}{2}]$ c.d.f. The distributions we consider in Chapters II and III all have decreasing upper tails. In this respect $\mathcal{G}$ is an extreme case among those we study in that its members have flat upper tails.

Indeed, this is reflected in the asymptotic result of Section 4 that if τ_n is a best invariant rule for the problem of length n then

$$\frac{1}{e} < \underline{\lim}_{n\to\infty} P(X_{\tau_n} = L_n) \le \overline{\lim}_{n\to\infty} P(X_{\tau_n} = L_n) < \alpha_o .$$

Thus asymptotically this P.I. Problem lies strictly between the N.I. and F.I. Problems.

Of interest in this vein is work done by Samuels [16] concurrent with and independent of ours. He has shown that for the P.I. Problem with distributional family $G' = \{G_{\alpha,\beta}\}_{-\infty<\alpha<\beta<\infty}$ ($G_{\alpha,\beta}$ being the $U[\alpha,\beta]$ c.d.f.) a best invariant rule, which is also minimax, is the N.I. Rule. Thus the family G' is so large in the invariant case as to be indistinguishable from the family of all distributions.

In addition this result shows that if we restrict ourselves to invariant rules then the family G is asymptotically an intermediate case of the P.I. Problem falling between that of the family G' and those of the families (among them the family of all normal distributions) which obtain the value α_o in the limit.

In the first three sections of Chapter IV a best invariant rule τ_n for the problem of length n is derived, and formulas are obtained giving $P(X_\tau=L_n)$ for stopping rules τ in a certain class (which contains τ_n) and giving decision numbers for a best invariant rule. Finally, $P(X_{\tau_n}=L_n)$ is found for small n.

CHAPTER II

ASYMPTOTIC FULL INFORMATION FOR SOME PARTIAL INFORMATION PROBLEMS

In this chapter we will consider the P.I. Problem defined for arbitrary location and/or scale parameter families. In each case the problem is invariant under the appropriate group of location, scale or location-scale changes. These considerations lead us to look at invariant stopping rules for these problems, an approach whose value is enhanced by the fact that the class of invariant rules contains a minimax rule in each case we consider (Kiefer [9]).

Recall from Chapter I that if σ_n is the F.I. Rule of length n then α_o is defined as

$$\alpha_o = \lim_{n\to\infty} P(X_{\sigma_n} = L_n) \doteq .58 .$$

This motivates the following definition:

Definition 2.1. For each n=1,2,..., let τ_n be an invariant stopping rule for the P.I. Problem of length n defined by any location, scale or location-scale parameter family $\{F_\theta\}_{\theta\in\Theta}$. The sequence of rules $\underline{\tau} = (\tau_1, \tau_2, \ldots)$ will be said to be asymptotically full information (asymptotically F.I.) if

$$\lim_{n\to\infty} P_{\theta_o}(X_{\tau_n}=L_n) = \alpha_o$$

for any fixed $\theta_o \in \Theta$. (Note that since τ_n is invariant for each n $P_\theta(X_{\tau_n}=L_n)$ is a constant independent of θ).

We will make the convention that whenever we speak of $\underline{\tau} = (\tau_1, \tau_2, \ldots)$ as being a sequence of invariant rules for a given P.I. Problem we will mean that τ_n is an invariant rule for the problem of length n (i.e., $\tau_n \leq n$).

The most significant result of this chapter is found in Section 6. There we give a sufficient condition (involving distributional upper tail shapes) for there to exist an asymptotically F.I. sequence of invariant rules in the P.I. Problem determined by an arbitrary location-scale parameter family of distributions. Of particular interest is the result that such a sequence exists for the P.I. Problem defined by the family of all normal distributions. To keep this existence proof (which, incidentally, is a constructive proof) both simple and general, the rules considered are somewhat naive. The result is obtained for a more interesting sequence of rules in Chapter III.

Section 1 of this chapter sets up the problem and proves preliminary lemmas which are also used in Chapter III. Section 2 gives a sufficient condition for the existence of asymptotically F.I. sequences of invariant rules in the P.I. Problem defined by arbitrary location parameter families of distributions. Section 4 does the same when the defining family is a scale parameter family. Both sufficient conditions involve the upper tail shapes in the defining distributional family. Each of Sections 3 and 5 applies the conditions of the preceding section to several common upper tail shapes.

The following notation holds throughout this chapter:

$\{F_\theta\}_{\theta\in\Theta}$ is a location, scale or location-scale parameter family of distributions.

$X_1, X_2, \ldots \sim$ i.i.d. F_θ, some $\theta\in\Theta$ are the observations.

$$\bar{X}_i = \frac{1}{i}\sum_{j=1}^{i} X_j \qquad i = 1,2,\ldots$$

$$S_i = \left(\frac{1}{i-1}\sum_{j=1}^{i} (X_j-\bar{X}_i)^2\right)^{\frac{1}{2}} \qquad i = 2,3,\ldots$$

$$L_i = \max\{X_1,\ldots,X_i\} \qquad i = 1,2,\ldots$$

P_θ, E_θ, Var_θ denote respectively probability, expectation and variance taken with respect to dF_θ.

1. Preliminaries

We begin with two lemmas which will be useful in Chapters II and III. Let

(2.1) $$1 \geq d_{1,n} \geq d_{2,n} \geq \cdots \geq d_{n,n} = 0$$

be the decision numbers defining the F.I. Rule when a U[0,1] distribution is assumed. Let

$$\infty \geq \delta_{1,n} \geq \delta_{2,n} \geq \cdots \geq \delta_{n,n} = -\infty$$

be the decision numbers defining the F.I. Rule when a N(0,1) distribution is assumed. Note that $F(\delta_{i,n}) = d_{i,n}$, $i=1,\ldots,n$, where F is the $N(0,1)$ c.d.f.

To obtain an exact value for $d_{i,n}$, and therefore for $\delta_{i,n}$, requires finding the roots of a polynomial of degree $n-i$ (see Gilbert and Mosteller [7]). This is impractical for large n. The following lemmas give bounds on the d's and δ's.

Lemma 2.1. Let $C_o = \dfrac{\log 2}{1+\log 2}$ and let $d_{j,n}$, $j=1,\ldots,n$, be the decision numbers given by (2.1). Then

$$1 - \frac{1}{n-j+1} \leq d_{j,n} \leq 1 - \frac{C_o}{n-j}\ , \quad j=1,\ldots,n-1\ .$$

Proof. By (1.3) and (1.4) the decision number $d_{j,n}$ is the number d at which, given $X_j=L_j=d$, we are indifferent as to whether to accept or reject X_j. Explicitly $d_{j,n}$ satisfies

$$(2.2) \qquad P(X_j=L_n \mid X_j=L_j=d_{j,n}) = \sup_{\sigma\in S_{j,n}} P(X_\sigma=L_n \mid X_j=L_j=d_{j,n})$$

where $S_{j,n}$ is the class of stopping rules σ for the X's such that $j<\sigma\leq n$. Note that

$$(2.3) \qquad P(X_j = L_n \mid X_j = L_j = d) = d^{n-j}\ .$$

Consider the stopping rule $\tau\in S_{j,n}$ given by

$$\tau = \min\{n, \min_{k>j}\{k: X_k = L_k\}\}\ .$$

For any $\ell=j+1,\ldots,n$ let

$$A_\ell = \{X_\ell > d_{j,n},\ X_m \leq d_{j,n},\ \ j < m \leq n,\ \ m \neq \ell\}.$$

It follows that

$$P(X_\tau = L_n | X_j = L_j = d_{j,n}) \geq P(\bigcup_{\ell=j+1}^{n} A_\ell | X_j = L_j = d_{j,n})$$

$$= (n-j)(1-d_{j,n}) d_{j,n}^{n-j-1} .$$

Thus by (2.2) and (2.3)

$$d_{j,n}^{n-j} \geq (n-j)(1-d_{j,n}) d_{j,n}^{n-j-1} ,$$

which implies

$$d_{j,n} \geq 1 - \frac{1}{n-j+1} .$$

We now show that

$$d_{j,n} \leq 1 - \frac{C_o}{n-j} , \text{ where } C_o = \frac{\log 2}{1+\log 2} .$$

Let $C < C_o$ and suppose $d > 1 - \frac{C}{n-j}$.

Then

$$P(X_j = L_n | X_j = L_j = d) > (1 - \frac{C}{n-j})^{n-j} > \frac{1}{2} .$$

For any $\sigma \in S_{j,n}$

$$P(X_\sigma = L_n | X_j = L_j = d) \leq P(X_{j+\ell} > X_j, \text{ some } 1 \leq \ell \leq n-j | X_j = L_j = d)$$

$$< 1 - (1 - \frac{C}{n-j})^{n-j} < \frac{1}{2} .$$

Thus

$$\sup_{\sigma \in S_{j,n}} P(X_\sigma = L_n | X_j = L_j = d) < P(X_j = L_n | X_j = L_j = d) .$$

By (2.2) then, $d_{j,n} < 1 - \frac{C}{n-j}$ for each such C and the desired inequality is proved. //

<u>Lemma 2.2.</u> Let $y < \int_1^\infty e^{-\frac{u^2}{2}} du$. For each such y let

$$(2.4)\qquad a(y) = \frac{\log\left(\dfrac{1 - 2\log y}{(-2\log y)^{\frac{1}{2}}-1}\right)}{(-2\log y)^{\frac{1}{2}}-1} .$$

Then if x(y) is defined by

$$y = \int_{x(y)}^{\infty} e^{-\frac{u^2}{2}}\, du ,$$

$$(2.5)\qquad (-2\log y)^{\frac{1}{2}} - a(y) < x(y) \le (-2\log y)^{\frac{1}{2}} .$$

Proof. The right inequality in (2.5) follows from the relation

$$y = \int_{x(y)}^{\infty} e^{-\frac{u^2}{2}}\, du \le \int_{x(y)}^{\infty} ue^{-\frac{u^2}{2}}\, du = e^{-\frac{x^2(y)}{2}} .$$

To obtain the left inequality we first consider the bound

$$\frac{x(y)}{1+x^2(y)}\, e^{-\frac{x^2(y)}{2}} < \int_{x(y)}^{\infty} e^{-\frac{u^2}{2}}\, du = y$$

which is well known (see McKean [10], p4). It follows that

$$x^2(y) + 2\log\frac{1+x^2(y)}{x(y)} > -2\log y .$$

Since x increases faster than $\log\left(\frac{1+x^2}{x}\right)$ and $1 > \log\left(\frac{5}{2}\right)$, we have

$$x + \frac{1}{2} > \log\left(\frac{1+x^2}{x}\right) \qquad \text{if } x > \frac{1}{2} .$$

It then follows that

$$(x(y)+1)^2 > x^2(y) + 2\log\frac{1+x^2(y)}{x(y)} > -2\log y$$

which implies

$$(2.6)\qquad x(y) > (-2\log y)^{\frac{1}{2}} - 1 .$$

Now

$$(2.7)\qquad 2a(y)x(y) + a^2(y) > 2a(y)x(y)$$

$$> 2\log\left(\frac{1-2\log y}{(-2\log y)^{\frac{1}{2}}-1}\right)$$

$$> 2\log\frac{1+x^2(y)}{x(y)}\ ,$$

the last inequality following from (2.6) and the right inequality of (2.5).

From (2.7) we obtain

$$(x(y)+a(y))^2 > x^2(y) + 2\log\frac{1+x^2(y)}{x(y)} > -2\log y$$

which yields

$$x(y) > (-2\log y)^{\frac{1}{2}} - a(y)\ . \qquad //$$

Lemmas 2.1 and 2.2 give bounds on the δ's. Let

$$y_{j,n} = \int_{\delta_{j,n}}^{\infty} e^{-\frac{u^2}{2}} du = \sqrt{2\pi}\,(1-d_{j,n})\ .$$

By lemma 2.1

$$\frac{\sqrt{2\pi}\,C_o}{n-j} \le y_{j,n} \le \frac{\sqrt{2\pi}}{n-j+1} \quad \text{where } C_o = \frac{\log 2}{1+\log 2}\ .$$

We note that a(y) given by (2.4) is an increasing function of y, $0 < y < e^{-\frac{1}{2}}$.

Hence Lemma 2.2 implies

$$(2.8)\qquad \left(-2\log\frac{\sqrt{2\pi}}{n-j+1}\right)^{\frac{1}{2}} - a\left(\frac{\sqrt{2\pi}}{n-j+1}\right) \le \delta_{j,n} \le \left(-2\log\frac{\sqrt{2\pi}\,C_o}{n-j}\right)^{\frac{1}{2}}$$

$$\text{for } n-j > \frac{\sqrt{2\pi}}{\int_1^\infty e^{-\frac{u^2}{2}}\,du} - 1 \qquad (\text{i.e. } n-j \geq 6)\ .$$

Let $\mathbb{Z}^+$ denote the positive integers.

We end this section of preliminaries by defining

$$\mathcal{D} = \{\psi: \mathbb{Z}^+ \to \mathbb{Z}^+ \text{ s.t. } \psi \text{ is non-decreasing, } \psi(n) = o(n),\ \psi(n) \leq n \text{ for all } n\}$$

$$\mathcal{D}' = \{\psi \in \mathcal{D} \text{ s.t. } \psi(n) \to \infty \text{ as } n \to \infty\}\ .$$

2. Location Parameter Families

In this section it is assumed that the family of distributions defining the P.I. Problem is the location parameter family $\{F_\theta\}_{\theta \in \mathbb{R}}$, where $F_\theta(x) = F(x-\theta)$, for some continuous c.d.f. F. It is also assumed that $\int_{\mathbb{R}} x^2 dF(x) < \infty$.

Our approach will be to use a small proportion of the observations to estimate the true value of the location parameter θ. If $\hat{\theta}$ is the estimate so obtained we will then apply the F.I. Rule for distribution $F_{\hat{\theta}}$ to the remaining observations.

Let $h, g \in \mathcal{D}'$ with $h(n) + g(n) < n$ for each n.

We may assume (see Remarks at the end of this section) that $E_0 X_1 = 0$. Then $\bar{X}_k$ is an unbiased estimate of θ, and $\mathrm{Var}_\theta\, \bar{X}_k = \frac{\mathrm{Var}_0\, X_1}{k}$. Thus

$$(2.9) \qquad P_\theta(|\bar{X}_k - \theta| > \gamma) < \frac{\mathrm{Var}_0\, X_1}{k\,\gamma^2}\ .$$

We will assume further that $\theta = 0$.

Let $\zeta_{i,n}$ $i = 1,\ldots,n$ be the decision numbers for the F.I. Rule when the F_0 distribution is assumed. Thus $\zeta_{i,n} = F_0^{-1}(d_{i,n})$, the d's being defined by (2.1).

Define two sequences of stopping rules, $\underline{R}_1$ and $\underline{R}_2$ as follows: for each n=1,2,... let

(2.10) $$R_{1,n} = R_{1,n}(h,g) = \min\{n-g(n), \min_{k>h(n)}\{k: X_k = L_k, X_k > \zeta_{k,n}\}\}$$

(2.11) $$R_{2,n} = R_{2,n}(h,g) = \min\{n-g(n), \min_{k>h(n)}\{k: X_k = L_k, X_k - \overline{X}_{h(n)} > \zeta_{k,n}\}\}.$$

Let $\underline{R}_i = (R_{i,1}, R_{i,2}, \ldots)$, $i = 1,2$.

It will be noticed that $R_{1,n}$ is the F.I. Rule of length n, assuming distribution F_0, altered so that it does not select any of the first h(n) or last g(n) observations. Since $h(n), g(n) = o(n)$ it is clear that $\underline{R}_1$ is asymptotically F.I.

$R_{2,n}$ is a stopping rule analogous to $R_{1,n}$ in that it ignores the same $h(n)+g(n)$ observations for selection purposes. However $R_{2,n}$ is invariant under location changes (i.e. for any $C \in \mathbb{R}$, $R_{2,n}$ will take the same value whether $X_1,\ldots,X_n$ or $X_1+C,\ldots,X_n+C$ are observed).

In what follows we obtain a sufficient condition for $\underline{R}_2$ to be asymptotically F.I. The condition involves the upper tail shape of $F_0 = F$. Our method is to compare $\underline{R}_2$ with the sequence $\underline{R}_1$.

$R_{1,n}$ and $R_{2,n}$ can disagree at observation X_j, $j> h(n)$ only if

$$\zeta_{j,n} < X_j \leq \zeta_{j,n} + \overline{X}_{h(n)}$$

or

$$\zeta_{j,n} + \overline{X}_{h(n)} < X_j \leq \zeta_{j,n} .$$

Thus if $\gamma_n > 0$ for each $n=1,2,\ldots,$

(2.12)
$$P_0(\{\{R_{1,n}=j,R_{2,n}\neq j\} \cup \{R_{1,n}\neq j,R_{2,n}=j\}\} \cap \{|\overline{X}_{h(n)}|<\gamma_n\})$$
$$\leq P_0(\zeta_{j,n}-\gamma_n<X_j<\zeta_{j,n}+\gamma_n)$$
$$= F_0(\zeta_{j,n}+\gamma_n)-F_0(\zeta_{j,n}-\gamma_n).$$

By (2.12) then

(2.13) $$P_0(\{X_{R_{1,n}}\neq X_{R_{2,n}}\} \cap \{|\overline{X}_{h(n)}|<\gamma_n\})$$
$$\leq \sum_{j=h(n)+1}^{n-g(n)} (F_0(\zeta_{j,n}+\gamma_n)-F_0(\zeta_{j,n}-\gamma_n)).$$

If γ_n can be chosen so that the right hand side of (2.13) goes to 0 as $n\to\infty$ and

(2.14) $$\lim_{n\to\infty} P_0(\{X_{R_{1,n}}\neq X_{R_{2,n}}\} \cap \{|\overline{X}_{h(n)}|>\gamma_n\}) = 0 ,$$

then it will follow that $\underline{R}_2$ is asymptotically F.I.

But by (2.9), in order for (2.14) to hold it suffices that $\lim_{n\to\infty} h(n)\gamma_n^2 = \infty$. We state these conditions as:

<u>Theorem 2.1</u>. In the P.I. Problem defined by a location parameter family $\{F_\theta\}_{\theta\in\mathbb{R}}$ with $E_0X_1=0$ and $E_0X_1^2<\infty$ the sequence of invariant rules $\underline{R}_2(h,g)$ defined by (2.11) is asymptotically

F.I. provided there exists a sequence $\{\gamma_n\}_{n=1}^{\infty}$ of positive numbers for which the two conditions

$$\text{i)} \quad \lim_{n\to\infty} h(n)\gamma_n^{\ 2} = \infty$$

$$\text{ii)} \quad \lim_{n\to\infty} \sum_{j=h(n)+1}^{n-g(n)} (F_0(\zeta_{j,n}+\gamma_n)-F_0(\zeta_{j,n}-\gamma_n)) = 0$$

are satisfied.

The role that h(n) plays in the foregoing formulation is necessary and clear: we need h(n) observations to supply an estimate of θ. Letting g be an arbitrary function in $\mathcal{D}'$ will sometimes facilitate computations in applications of this result (e.g. the normal case treated in the next section). However knowing that $\underline{R}_2(h,g)$ is asymptotically F.I. is enough to insure that $\underline{R}_2(h,0)$ is asymptotically F.I.

Corollary 2.1. Under the assumptions of Theorem 2.1, the sequence of invariant rules $\underline{R}_2(h,0)$ is asymptotically F.I.

Proof. The rule $R_{2,n}(h,g)$ will choose L_n when $R_{2,n}(h,0)$ does not choose it only if $X_{n-g(n)} = L_n$. Thus $P_0(X_{R_{2,n}(h,0)} \neq L_n, X_{R_{2,n}(h,g)} = L_n) \leq \frac{1}{n}$, which implies

$$P_0(X_{R_{2,n}(h,0)} = L_n) \geq P_0(X_{R_{2,n}(h,g)} = L_n) - \frac{1}{n} ,$$

and the result follows. //

Remarks

Since the Chebychev bound (2.9) applies evenly to all distributional families having finite second moment, the theorem distinguishes among the various families through condition ii) which is really a condition on the shape of the upper tail of the distribution. The more rapidly the tail approaches 0, the more likely we are to obtain an asymptotically F.I. rule.

While Theorem 2.1, giving only a sufficient condition, cannot tell us which tail shapes do not approach 0 quickly enough to guarantee an asymptotically F.I. rule, it can verify that the decrease is fast enough for some common tail shapes as the next section shows.

In two of the applications in the next section we apply condition ii) to families $\{F_\theta\}_{\theta\in\mathbb{R}}$ in which $E_0 X_1 \neq 0$. It can be shown that if the conditions of the theorem are satisfied and if $E_0 X_1 \neq 0$ then the sequence of rules $(R'_{2,2}, R'_{2,3}, \ldots)$ where

$$R'_{2,n} = \min\{n, \min_{k>h(n)} \{k: X_k = L_k,\ X_k - (\overline{X}_{h(n)} - E_0 X_1) > \zeta_{k,n}\}\}$$

is asymptotically F.I.

3. Location Parameter Families - Applications

Normal Tail Shapes

Suppose $F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du$.

Then (2.4) and (2.8) give us

$$\left(2 \log\frac{n-j+1}{\sqrt{2\pi}}\right)^{\frac{1}{2}} - a_n \leq \zeta_{j,n} = \delta_{j,n} \leq \left(2 \log\frac{n-j}{\sqrt{2\pi}\ C_o}\right)^{\frac{1}{2}}$$

where $\quad a_n = O\left(\frac{\log\log g(n)}{(\log g(n))^{\frac{1}{2}}}\right)$.

Choose γ_n satisfying $\gamma_n < \zeta_{j,n}$, $h(n) < j \leq n-g(n)$.
Then, noting that $x^{\frac{1}{2}} < \frac{x}{2}$ if $x \geq 4$,

$$F_0(\zeta_{j,n}+\gamma_n)-F_0(\zeta_{j,n}-\gamma_n) \leq \frac{2\gamma_n}{\sqrt{2\pi}}\, e^{\frac{-((2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-(a_n+\gamma_n))^2}{2}}$$

$$\leq \frac{2\gamma_n}{n-j+1}\, e^{(a_n+\gamma_n)(2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}}$$

$$\leq \frac{2\gamma_n}{n-j+1}\, e^{(a_n+\gamma_n)\log\frac{n-j+1}{\sqrt{2\pi}}}$$

$$= \frac{2\gamma_n}{(2\pi)^{\frac{a_n+\gamma_n}{2}}(n-j+1)^{1-(a_n+\gamma_n)}}.$$

Thus if $\lim\limits_{n\to\infty}\gamma_n = 0$,

$$(2.15)\quad \overline{\lim_{n\to\infty}}\sum_{j=h(n)+1}^{n-g(n)}(F_0(\zeta_{j,n}+\gamma_n) - F_0(\zeta_{j,n}-\gamma_n))$$

$$\leq \overline{\lim_{n\to\infty}}\frac{2\gamma_n}{(2\pi)^{\frac{a_n+\gamma_n}{2}}}\sum_{j=h(n)+1}^{n-g(n)}\frac{1}{(n-j+1)^{1-(a_n+\gamma_n)}}$$

$$\leq \overline{\lim_{n\to\infty}}\, 2\gamma_n\int_{h(n)+1}^{n}\frac{dx}{(n-x+1)^{1-\varepsilon}} \leq \overline{\lim_{n\to\infty}}\frac{2\gamma_n n^{\varepsilon}}{\varepsilon}$$

for any $1 > \varepsilon > 0$.

So if we choose γ_n and $h \in \mathcal{D}'$ such that $\gamma_n n^{\varepsilon} \to 0$ and $h(n)\ \gamma_n^2 \to \infty$ as $n \to \infty$, the conditions of Corollary 2.1 will be met. For example we may take $\gamma_n = n^{-2\varepsilon}$ and $h(n) = n^{5\varepsilon}$ for an arbitrarily small $\varepsilon > 0$.

Exponentially Shaped Tails

Suppose $F_0(x) = 1-e^{-x}$, $x > 0$. Then

$$\log\ (n-j+1) \le \zeta_{j,n} \le \log\ \left(\frac{n-j}{C_o}\right) \qquad j=1,\dots,n-1\ .$$

Assume also $0 < \gamma_n < \zeta_{n-1,n}$ for each n, so that

$$F_0(\zeta_{j,n}+\gamma_n) - F_0(\zeta_{j,n}-\gamma_n) = e^{-\zeta_{j,n}}(e^{\gamma_n}-e^{-\gamma_n})$$

$$\le \frac{e^{\gamma_n}-e^{-\gamma_n}}{n-j+1} \le \frac{2}{n-j+1}\ \frac{\gamma_n}{1-\gamma_n}\ ,$$

$$j=1,\dots,n-1\ .$$

Assuming $g(n) \ge 1$ for all n we have

$$\sum_{j=h(n)+1}^{n-g(n)} (F_0(\zeta_{j,n}+\gamma_n)-F_0(\zeta_{j,n}-\gamma_n)) \le \frac{\gamma_n}{1-\gamma_n} \sum_{j=h(n)+1}^{n-g(n)} \frac{2}{n-j+1} \to 0$$

if $\gamma_n \log n \to 0$ as $n \to \infty$. In this case we could take $\gamma_n = [\log^{-2} n]$ and $h(n) = [\log^5 n]$ (where $[\cdot]$ is the greatest integer function).

Tails Decreasing Like Inverse Powers

Suppose $F_0(x) = 1-x^{1-t}$, $t>2$, $x \ge 1$. Then $(n-j+1)^{\frac{1}{t-1}} \le \zeta_{j,n} \le \left(\frac{n-j}{C_o}\right)^{\frac{1}{t-1}}$. Assume $0 < \gamma_n < \zeta_{n-1,n}$ for all n. We have

$$F_0(\zeta_{j,n}+\gamma_n) - F_0(\zeta_{j,n}-\gamma_n) = 2\, F_0'(\eta)\, \gamma_n$$

$$\text{some } \eta \in (\zeta_{j,n}-\gamma_n,\ \zeta_{j,n}+\gamma_n)$$

$$= \frac{2(t-1)\gamma_n}{\eta^t} \leq \frac{2(t-1)\gamma_n}{(\zeta_{j,n}-\gamma_n)^t}$$

$$\leq \frac{2(t-1)\gamma_n}{((n-j+1)^{\frac{1}{t-1}}-\gamma_n)^t} \leq \frac{2(t-1)\gamma_n}{(n-j+1)^{\frac{2t}{2t-1}}}$$

if $\gamma_n < (n-j+1)^{\frac{1}{t-1}} (1-(n-j+1)^{-\frac{1}{(t-1)(2t-1)}})$.

It is clear that we may easily choose γ_n to satisfy this last condition.

Thus if $g \geq 1$ for all n,

$$\sum_{j=h(n)+1}^{n-g(n)} (F_0(\zeta_{j,n}+\gamma_n)-F_0(\zeta_{j,n}-\gamma_n)) \leq 2(t-1)\gamma_n \sum_{j=h(n)+1}^{n-g(n)} \frac{1}{(n-j+1)^{\frac{2t}{2t-1}}} \to 0$$

if $\gamma_n \to 0$ as $n \to \infty$.

It is then an easy matter to find an h for which

$\lim_{n\to\infty} h(n)\, \gamma_n^2 = \infty$.

Remark

If we consider the location parameter family $\{G_\theta\}_{\theta\in\mathbb{R}}$ where G_θ is the c.d.f. of the $U[\theta-\frac{1}{2}, \theta+\frac{1}{2}]$ distribution, then conditions i.) and ii.) of Theorem 2.1 cannot both hold.

4. Scale Parameter Families

In this section we assume the family of distributions defining the P.I. Problem is the scale parameter family $\{F_\theta\}_{\theta>0}$ where $F_\theta(x) = F(\frac{x}{\theta})$ for some continuous c.d.f. F. We will require that $\int_{\mathbb{R}} x^4 dF(x) < \infty$. We will also assume (see Remarks at the end of this section) that $E_1X_1 = 0$, $Var_1X_1 = 1$.

Under this assumption $E_\theta\, S_k^2 = Var_\theta X_1 = \theta^2$ so that S_k^2 is an unbiased estimate of θ^2. It can be shown that

$$Var_1 S_k^2 = \frac{1}{k}(E_1X_1^4-1)(1+o(1)) \text{ as } k \to \infty.$$

Thus for $\xi \in (0,1)$

$$(2.16)\quad P_\theta(S_k^2>(1+\xi)^2\theta^2 \text{ or } S_k^2<(1-\xi)^2\theta^2) \le P_1(|S_k^2-1|>2\xi-\xi^2) \le \frac{E_1X_1^4(1+o(1))}{k(2\xi-\xi^2)^2}.$$

We follow the approach of Section 2. Assume $\theta = 1$. Let $\zeta_{i,n}$ $i=1,\ldots,n$ be the decision numbers for the F.I. Rule when the F_1 distribution is assumed. Then for these ζ values and $h,g\in \mathcal{D}'$, $h(n)+g(n)<n$, let $R_{1,n}$ be given by (2.10) for each n. Define $R_{3,n}(h,g)$ as

$$(2.17)\quad R_{3,n}=R_{3,n}(h,g)=\min\{n-g(n), \min_{k>h(n)} \{k: X_k=L_k, \frac{X_k}{S_{h(n)}}>\zeta_{k,n}\}\}.$$

We define the sequences $\underline{R}_1$, $\underline{R}_3$ as was done in Section 2. $R_{3,n}$ is a rule invariant under scale changes.

$R_{1,n}$ and $R_{3,n}$ disagree at observation X_j, $j > h(n)$ only if

$$S_{h(n)}\ \zeta_{j,n} < X_j < \zeta_{j,n}$$

or

$$\zeta_{j,n} < X_j < S_{h(n)}\ \zeta_{j,n}\ .$$

Thus, taking $\xi_n \in (0,1)$ for each n,

$$(2.18)\quad P_1(\{\{R_{1,n}=j, R_{2,n}\neq j\} \cup \{R_{1,n}\neq j, R_{2,n}=j\}\} \cap \{|S_{h(n)}-1|<\xi_n\})$$

$$\leq |F_1((1+\xi_n)\ \zeta_{j,n}) - F_1((1-\xi_n)\zeta_{j,n})|\ .$$

(The absolute value is necessary since $\zeta_{j,n}$ may be negative). By (2.16), (2.18) and the discussion leading to Corollary 2.1 we may write

Theorem 2.2. In the P.I. Problem defined by a scale parameter family $\{F_\theta\}_{\theta>0}$ with $Var_1 X_1 = 1$ and $E_1 X_1^{\ 4} < \infty$, the sequence of invariant rules $\underline{R}_3(h,0)$ defined by (2.17) is asymptotically F.I. provided there exists a sequence $\{\xi_n\}_{n=1}^{\infty}$, $0 < \xi_n < 1$ and a $g \in \mathcal{D}'$, for which the two conditions

i) $\lim_{n\to\infty} h(n)\xi_n^{\ 2} = \infty$

ii) $\lim_{n\to\infty} \sum_{j=h(n)+1}^{n-g(n)} |F_1((1+\xi_n)\zeta_{j,n}) - F_1((1-\xi_n)\zeta_{j,n})| = 0,$

are satisfied.

Remark

Echoing the last remark of Section 2 we note that if $Var_1X_1 \neq 1$ then the sequence of rules $\underline{R}_3'$ given by

$$R_{3,n}'=\min\{n,\min_{k>h(n)}\{k:X_k=L_k,\frac{(Var_1X_1)^{\frac{1}{2}}X_k}{S_{h(n)}}>\zeta_{k,n}\}\}$$

is asymptotically F.I. if the conditions of Theorem 2.2 are met.

5. Scale Parameter Families - Applications

We now consider the application of Theorem 2.2 to scale parameter families in which F_1 has those tail shapes considered in Section 3.

Normal Tail Shapes

Suppose $F_1(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}du$.

Then, taking a_n as in Section 3,

$$F_1((1+\xi_n)\zeta_{j,n})-F_1((1-\xi_n)\zeta_{j,n})$$

$$\leq \frac{2\xi_n\zeta_{j,n}}{\sqrt{2\pi}}\, e^{\frac{-((1-\xi_n)((2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-a_n))^2}{2}}$$

$$\leq \frac{2\xi_n(2\log\frac{n-j}{\sqrt{2\pi}\,C_o})^{\frac{1}{2}}}{(n-j+1)^{(1-\xi_n)^2}}\, e^{a_n(1-\xi_n)(2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}}$$

$$\leq \frac{2\xi_n\left(2\log\frac{n-j}{\sqrt{2\pi}\,C_o}\right)^{\frac{1}{2}}}{(n-j+1)^{(1-\xi_n)^2}}\, e^{a_n(1-\xi_n)\log\frac{n-j+1}{\sqrt{2\pi}}}$$

$$\leq \frac{2\xi_n\left(2\log\frac{n-h(n)}{\sqrt{2\pi}\,C_o}\right)^{\frac{1}{2}}}{(2\pi)^{\frac{a_n(1-\xi_n)}{2}}\,(n-j+1)^{(1-\xi_n)(1-\xi_n-a_n)}} .$$

Thus

$$(2.19)\quad \sum_{j-h(n)+1}^{n-g(n)} \left(F_1((1+\xi_n)\zeta_{j,n})-F_1((1-\xi_n)\zeta_{j,n})\right)$$

$$\leq \frac{2\xi_n\left(2\log\frac{n-h(n)}{\sqrt{2\pi}\,C_o}\right)^{\frac{1}{2}}}{(2\pi)^{\frac{a_n(1-\xi_n)}{2}}} \sum_{j=h(n)+1}^{n-g(n)} \frac{1}{(n-j+1)^{(1-\xi_n)(1-\xi_n-a_n)}} \to 0$$

if $\xi_n n^{\varepsilon} \to 0$ for some $\varepsilon > 0$ as $n \to \infty$.

Exponentially Shaped Tails

If $F_1(x) = 1 - e^{-x} \quad x > 0$

then (see Section 3)

$$F_1((1+\xi_n)\zeta_{j,n}) - F_1((1-\xi_n)\zeta_{j,n}) = e^{-\zeta_{j,n}}\left(e^{\xi_n\zeta_{j,n}} - e^{-\xi_n\zeta_{j,n}}\right)$$

$$\leq \frac{2}{n-j+1} \sum_{\ell=1}^{\infty} \frac{\xi_n^{2\ell-1}}{(2\ell-1)!}\left(\log\frac{n-j}{C_o}\right)^{2\ell-1} ,$$

for $h(n) < j \leq n-g(n)$, and $g(n) \geq 1$.

Thus

$$\sum_{j=h(n)+1}^{n-g(n)} (F_1((1+\xi_n)\zeta_{j,n})-F_1((1-\xi_n)\zeta_{j,n}))$$

$$\leq \sum_{\ell=1}^{\infty} \sum_{j=h(n)+1}^{n-g(n)} \frac{2\xi_n^{2\ell-1}(\log\frac{n-j}{C_o})^{2\ell-1}}{(2\ell-1)!\,(n-j)}$$

$$\leq \sum_{\ell=1}^{\infty} \frac{2\xi_n^{2\ell-1}}{(2\ell-1)!} \int_1^{\frac{n-h(n)+1}{C_o}} \frac{(\log u)^{2\ell-1}}{u}\,du$$

$$= 2\sum_{\ell=1}^{\infty} \frac{\xi_n^{2\ell-1}(\log\frac{n-h(n)+1}{C_o})^{2\ell}}{(2\ell)!} \to 0$$

if $\xi_n(\log n)^2 \to 0$.

Tails Decreasing Like Inverse Powers

Suppose $F_1(x) = 1-x^{1-t} \quad t > 5,\ x \geq 1,$

and $(1-\xi_n)\zeta_{n-1,n} > 1$. Then, recalling computations done with this distribution in Section 3,

$$F_1((1+\xi_n)\zeta_{j,n})-F_1((1-\xi_n)\zeta_{j,n}) = 2F_1'(\eta)\zeta_{j,n}\xi_n$$

$$\eta\in((1-\xi_n)\zeta_{j,n},(1+\xi_n)\zeta_{j,n})$$

$$= \frac{2(t-1)\zeta_{j,n}\xi_n}{\eta^t} \leq \frac{2(t-1)\zeta_{j,n}\xi_n}{(1-\xi_n)^t\zeta_{j,n}^t}$$

$$= \frac{2(t-1)\xi_n}{(1-\xi_n)^t\zeta_{j,n}^{t-1}} \leq \frac{2(t-1)\xi_n}{(1-\xi_n)^t(n-j+1)}$$

Thus

$$\sum_{j=h(n)+1}^{n-g(n)} (F_1((1+\xi_n)\zeta_{j,n})-F_1((1-\xi_n)\zeta_{j,n})) \leq \frac{2(t-1)\xi_n}{(1-\xi_n)^t} \log(n-h(n)) \to 0$$

if $\xi_n \log n \to 0$ as $n \to \infty$.

6. Location-Scale Parameter Families

In this section we assume the family of distributions defining the P.I. Problem is the location-scale parameter family $\{F_{\mu,\sigma}\}_{\substack{\mu\in\mathbb{R}\\ \sigma>0}}$ where $F_{\mu,\sigma}(x) = F(\frac{x-\mu}{\sigma})$ for some continuous c.d.f. F. As in Section 4 we assume $\int_{\mathbb{R}} x^4 dF(x) < \infty$. We also assume (see Remark following Theorem 2.3) that $E_{0,1}X_1 = 0$ and $Var_{0,1}X_1 = 1$. Let $\gamma_n > 0$, $1 > \xi_n > 0$ for each n. Then

$$P_{\mu,\sigma}\left(\left|\frac{\overline{X}_k-\mu}{\sigma}\right| > \gamma_n\right) \leq \frac{1}{k\gamma_n^2}, \quad k \geq 1,$$

and

$$P_{\mu,\sigma}(S_k^2>(1+\xi_n)^2\sigma^2 \text{ or } S_k^2<(1-\xi_n)^2\sigma^2) \leq P_{0,1}(|S_k^2-1|>2\xi_n-\xi_n^2)$$

$$\leq \frac{(E_{0,1}X_1^4-1)(1+o(1))}{k(2\xi_n-\xi_n^2)^2}$$

as $k \to \infty$.

We assume now that $(\mu,\sigma) = (0,1)$. Let $\{\zeta_{j,n}\}_{j=1}^{n}$ be the decision numbers for F.I. Rule of length n when a $F_{0,1}$ distribution is assumed. Then define the stopping rule $R_{1,n}(h,g)$ for $h,g\in\mathcal{D}'$, $h(n)+g(n)<n$, by (2.10). Define the stopping rule $R_{4,n}(h,g)$ as

(2.20)

$$R_{4,n}=R_{4,n}(h,g)=\min\{n-g(n), \min_{k>h(n)} \{k:X_k=L_k, \frac{X_k-\overline{X}_{h(n)}}{S_{h(n)}} > \zeta_{k,n}\}\} .$$

$R_{4,n}$ is a rule similar to $R_{1,n}$ but invariant under location-scale changes. Let $\underline{R}_1$, $\underline{R}_4$ be the sequences formed from the above rules.

R_1 and R_4 disagree at observation X_j only if

$$S_{h(n)}\zeta_{j,n} + \overline{X}_{h(n)} < X_j < \zeta_{j,n}$$

or

$$\zeta_{j,n} < X_j < S_{h(n)}\zeta_{j,n} + \overline{X}_{h(n)} .$$

This gives us

$$P_{0,1}(\{\{R_{1,n}=j,R_{4,n}\neq j\} \cup \{R_{1,n}\neq j,R_{4,n}=j\}\} \cap \{|\overline{X}_{h(n)}|<\gamma_n\} \cap \{|S_{h(n)}-1|<\xi_n\})$$

$$\leq \max\{|F_{0,1}((1-\xi_n)\zeta_{j,n}-\gamma_n)-F_{0,1}((1+\xi_n)\zeta_{j,n}+\gamma_n)|,$$

$$|F_{0,1}((1-\xi_n)\zeta_{j,n}+\gamma_n)-F_{0,1}((1+\xi_n)\zeta_{j,n}-\gamma_n)|\} .$$

Thus we have

Theorem 2.3. In the P.I. Problem defined by the location-scale parameter family $\{F_{\mu,\sigma}\}_{\mu\in\mathbb{R}, \sigma>0}$ with $E_{0,1}X_1=0$, $E_{0,1}X_1^2=1$, and $E_{0,1}X_1^4 < \infty$, the sequence of invariant rules $\underline{R}_4(h,0)$ given by (2.20) is asymptotically F.I. provided there exist sequences $\{\gamma_n\}_{n=1}^{\infty}, \{\xi_n\}_{n=1}^{\infty}$, $\gamma_n>0$, $1>\xi_n>0$, and a $g\in\mathcal{D}'$ such that both conditions

i) $\lim_{n\to\infty} h(n)\xi_n^2 = \lim_{n\to\infty} h(n)\gamma_n^2 = \infty$

ii) $\lim_{n\to\infty} \sum_{j=h(n)+1}^{n-g(n)} \max\{|F_{0,1}((1+\xi_n)\zeta_{j,n}+\gamma_n)-F_{0,1}((1-\xi_n)\zeta_{j,n}-\gamma_n)|,$

$|F_{0,1}((1-\xi_n)\zeta_{j,n}+\gamma_n)-F_{0,1}((1+\xi_n)\zeta_{j,n}-\gamma_n)|\}=0$

are satisfied.

Remarks

The easiest method for handling the case where $E_{0,1}X_1\neq 0$ and/or $Var_{0,1}X_1\neq 1$ is to find μ',σ' for which $E_{\mu',\sigma'}X_1=0$, $Var_{\mu',\sigma'}X_1=1$ and reformulate condition ii.) by substituting $F_{\mu',\sigma'}$ for $F_{0,1}$ and $\zeta'_{j,n}=F^{-1}_{\mu',\sigma'}(d_{j,n}) = \sigma' F^{-1}_{0,1}(d_{j,n}) + \mu'$ for $\zeta_{j,n}$.

This method, with a glance at Sections 3 and 5, will easily give the result for the exponential and inverse power upper tail shapes studied there. For the remainder of this section we confine ourselves to the location-scale family of all normal distributions. That is, we assume

$$F_{0,1}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du .$$

Then condition ii.) becomes

(2.21)

$$\lim_{n\to\infty} \left[\sum_{j=h(n)+1}^{n-g(n)} (F_{0,1}((1+\xi_n)\zeta_{j,n}+\gamma_n)-F_{0,1}((1+\xi_n)\zeta_{j,n})) \right.$$
$$+ \sum_{j=h(n)+1}^{n-g(n)} (F_{0,1}((1+\xi_n)\zeta_{j,n})-F_{0,1}((1-\xi_n)\zeta_{j,n}))$$

$$+\sum_{j=h(n)+1}^{n-g(n)} (F_{0,1}((1-\xi_n)\zeta_{j,n})-F_{0,1}((1-\xi_n)\zeta_{j,n}-\gamma_n))\Bigg]=0.$$

The first two sums of (2.21) are dominated by their respective analogues in Sections 3 and 5 ((2.15) and (2.19)) and therefore converge to 0 if $\gamma_n n^{\varepsilon}\to 0$ and $\xi_n n^{\varepsilon}\to 0$ for some $1>\varepsilon>0$ as $n\to\infty$. Therefore it remains only to show that the last sum has 0 as its limit.

$$F_{0,1}((1-\xi_n)\zeta_{j,n})-F_{0,1}((1-\xi_n)\zeta_{j,n}-\gamma_n) \le \frac{\gamma_n}{\sqrt{2\pi}} e^{\dfrac{-[(1-\xi_n)((2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-a_n)-\gamma_n]^2}{2}}$$

$$\le \frac{\gamma_n}{\sqrt{2\pi}} \left(\frac{\sqrt{2\pi}}{n-j+1}\right)^{(1-\xi_n)^2} \left(\frac{n-j+1}{\sqrt{2\pi}}\right)^{(a_n(1-\xi_n)^2+\gamma_n(1-\xi_n))}$$

and it may be seen from similar computations in Sections 3 and 5 that choosing γ_n, ξ_n such that $\gamma_n n^{\varepsilon} \to 0$ and $\xi_n n^{\varepsilon} \to 0$ for some $1 > \varepsilon > 0$ produces the desired convergence to 0.

Let $\alpha_n = P(X_{\sigma_n}=L_n)$ where the distribution of the X's is known and σ_n is the correct F.I. Rule of length n. Let $\beta_{j,n} = P_{\theta_o}(X_{R_{j,n}}=L_n)$ $j=2,3,4$, where θ_o is a fixed parameter value for the appropriate distributional family (location, scale or location-scale).

Theorems 2.1-2.3 give us a sufficient condition in order that $\lim_{n\to\infty}(\alpha_n-\beta_{j,n})=0$ for each $j=2,3,4$. The computations of Sections 3, 5 and 6 tell us more. Implicit in these computations is an upper bound for $\alpha_n-\beta_{j,n}$.

Let us consider this bound for the normal location, scale and location-scale parameter families.

The bound obtained clearly depends on our choice of h, g, γ_n and ξ_n, and its value can be divided into two components. One is $\frac{g(n)}{n} + \frac{h(n)}{n}$, the probability that the largest observation occurs among the $h(n) + g(n)$ observations our rule neglects. The second is obtained from condition ii.) of each theorem and represents our estimate of how well we can do with the other $n-h(n)-g(n)$ observations.

By (2.15) we have for the location parameter case a bound smaller than $\frac{2\gamma_n n^{(a_n+\gamma_n)}}{a_n+\gamma_n}$. Let $\varepsilon>0$, $\delta>2\varepsilon$, $h(n)=[n^\delta]$, $g(n)=[n^\nu]$, $\gamma_n=n^{-\varepsilon}$. Then this bound is less than

$$b_n = C_1 \frac{2n^{\rho(n,\nu,\varepsilon)+n^{-\varepsilon}}}{n^\varepsilon \rho(n,\nu,\varepsilon)+1}$$

for a constant C_1 and

$$\rho(n,\nu,\varepsilon) = \frac{\log\log n + \log \nu}{\nu^{\frac{1}{2}} (\log n)^{\frac{1}{2}}}$$

It can be seen from (2.19) that for the same values of h, g and $\xi_n = n^{-\varepsilon}$,

$$\alpha_n - \beta_{3,n} \le C_2 (\log n)^{\frac{1}{2}} b_n + \frac{1}{n^{1-\delta}} + \frac{1}{n^{1-\nu}} .$$

From (2.21) we obtain the estimate

$$\alpha_n - \beta_{4,n} \leq C_3 b_n + C_2 (\log n)^{\frac{1}{2}} b_n + \frac{1}{n^{1-\delta}} + \frac{1}{n^{1-\nu}}$$

if we take h, g, γ_n, ξ_n as above.

CHAPTER III

THE NORMAL PROBLEM

This chapter deals with the versions of the P.I. Problem determined by the two families of normal distributions $\{F_{\mu,1}\}_{\mu\in\mathbb{R}}$ and $\{F_{\mu,\sigma}\}_{\substack{\mu\in\mathbb{R}\\ \sigma>0}}$ where $F_{\mu,\sigma}$ is the c.d.f. of the $N(\mu,\sigma^2)$ distribution. From now on these versions will be known respectively as the $N(\mu,1)$ Problem and the $N(\mu,\sigma^2)$ Problem.

Recall that in Chapter II we proved that there are sequences of asymptotically F.I. invariant rules for each of these problems. In Section 1 of this chapter we consider sequences of rules having a more satisfactory form than those of Chapter II, and show that these sequences also are asymptotically F.I. Section 2 considers best invariant rules for each case. Unfortunately the expressions involved in obtaining exact decision numbers and probabilities of winning for these best invariant rules are intractable even for small n. In view of this, the chapter concludes with a discussion of some approximate procedures we have investigated for small n.

The following notation will be used in this chapter:

$$X_1, X_2, \ldots \sim \text{i.i.d. } N(\mu,\sigma^2) \text{ (or } N(\mu,1))$$

are the observations.

$\overline{X}_i$, S_i, L_i, $\{d_{i,n}\}_{i=1}^{n}$, $\{\delta_{i,n}\}_{i=1}^{n}$, $\mathcal{D}$, $\mathcal{D}'$ are as defined in Chapter II.

$$Y_2 = \frac{X_2-\overline{X}_2}{S_2}, \quad Y_i = \frac{X_i-\overline{X}_{i-1}}{S_{i-1}} \qquad i=3,\ldots,n.$$

$$\underline{Y}_i = (Y_2,\ldots,Y_i) \qquad i=2,\ldots,n$$

$$Z_i = X_i-\overline{X}_{i-1} \qquad i=2,\ldots,n$$

$$\underline{Z}_i = (Z_2,\ldots,Z_i) \qquad i=2,\ldots,n.$$

$F_{\mu,\sigma}$= c.d.f. of $N(\mu,\sigma^2)$ distribution.

1. Asymptotic Results

In this section we first consider a class C_5 of sequences of invariant stopping rules for the $N(\mu,1)$ Problem which are asymptotically F.I. As in Chapter II, our approach is to compare the sequences of rules in C_5 with the sequence of optimal rules for the F.I. Problem. We then compare each sequence in a subclass of C_5 with an analogous sequence of invariant rules for the $N(\mu,\sigma^2)$ Problem to show that the latter sequence is also asymptotically F.I.

Assume $X_1,X_2,\ldots \sim$ i.i.d. $N(0,1)$. Let $h,g\in\mathcal{D}$ be such that $h(n)+g(n)<n$ for each n. We define for each n the stopping rules

$$R_{1,n}=R_{1,n}(h,g)=\min\{n-g(n),\min_{k>h(n)}\{k:X_k=L_k,X_k>\delta_{k,n}\}\}$$

$$R_{5,n}=R_{5,n}(h,g)=\min\{n-g(n),\min_{k>h(n)}\{k:X_k=L_k,Z_k>\delta_{k,n}\}\}$$

$$R_{6,n}=R_{6,n}(h,g)=\min\{n-g(n),\min_{k>h(n)}\{k:X_k=L_k,Y_k>\delta_{k,n}\}\}.$$

$R_{1,n}$ is the same modified F.I. Rule considered in Chapter II. $R_{5,n}$ is a stopping rule analogous to $R_{1,n}$ but invariant under location changes. $R_{6,n}$ is similar to the others but is invariant under location and scale changes. Notice that $R_{5,n}$ differs from $R_{2,n}$ of Chapter II (and $R_{6,n}$ differs from $R_{3,n}$) in that the former constantly updates its estimate of the parameter while the latter obtains an estimate using a fixed number of observations and never changes that estimate. The greater complexity of the former makes it more difficult to work with, but the more desirable form of this rule rewards our extra effort.

As in Chapter II each pair (h,g) defines a sequence of rules

$$\underline{R}_i = \underline{R}_i(h,g) = (R_{i,3}(h,g),\ R_{i,4}(h,g),\ \ldots) \quad i=1,5,6.$$

Then

$$\mathcal{C}_i = \{\underline{R}_i(h,g),h,g\in\mathcal{D}\} \qquad i=1,5,6$$

are the classes of sequences of rules whose asymptotic behavior we will investigate.

As was noted in Chapter II, each sequence in $\mathcal{C}_1$ is asymptotically F.I. We will show that each sequence in $\mathcal{C}_5$ is asymptotically F.I. by comparing these sequences with their counterparts in $\mathcal{C}_1$. We begin with a computational lemma.

<u>Lemma 3.1</u>. Let $a>0$, $X_1, X_2, \ldots \sim$ i.i.d. $N(0,1)$.

Then $P\left(\sup_{1\le k\le n}\left|\frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}}\right|>a\right) \le 4\left(1-F_{0,1}\left(\frac{a}{\sqrt{2}}\right)\right)\frac{\log 2n}{\log 2}$.

<u>Proof</u>

(3.1)
$$P\left(\sup_{1\le k\le 2^j}\left|\frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}}\right|>a\right) \le \sum_{i=1}^{j} P\left(\sup_{2^{i-1}\le k\le 2^i}\left|\frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}}\right|>a\right)$$

$$\le \sum_{i=1}^{j} P\left(\sup_{1\le k\le 2^i}\left|\sum_{\ell=1}^{k} X_\ell\right|>2^{\frac{i-1}{2}}a\right)$$

$$\le 2\sum_{i=1}^{j} P\left(\left|\sum_{\ell=1}^{2^i} X_\ell\right|>2^{\frac{i-1}{2}}a\right) \quad \text{(reflection principle)}$$

$$= 2\sum_{i=1}^{j} P\left(|X_1|>\frac{a}{\sqrt{2}}\right)$$

$$\le 4j\left(1-F_{0,1}\left(\frac{a}{\sqrt{2}}\right)\right) .$$

Suppose $2^{j-1}<n\le 2^j$. Then $(j-1)\log 2 < \log n \le j \log 2$

which implies $\frac{\log n}{\log 2} \le j < \frac{\log 2n}{\log 2}$.

This along with (3.1) gives the result. //

For $0<\alpha<1$ Lemmas 3.1 and 2.2 imply

$$P\left(\sup_{1\le k\le n}\left|\frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}}\right| > 2\left(\log\frac{4\log 2n}{\sqrt{2\pi}\,\alpha\,\log 2}\right)^{\frac{1}{2}}\right) < \alpha.$$

Let $C_\alpha = \log \dfrac{4}{\sqrt{2\pi}\alpha\log 2}$.

Fix $b \in (0,\frac{1}{4})$ and choose $h_b(n)$, $g_b(n) \in \mathcal{D}'$ such that

$$(3.2)\qquad h_b(n) \geq \min\{1 \leq j \leq n-1 : \frac{2(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(j-1)^{\frac{1}{2}}} < b\}$$

and

$$(3.3)\qquad (\log\log 2n)^{\frac{1}{1-4b}} = o(g_b(n)) \ .$$

(Naturally we assume n is large enough for each pair (α,b) so that these expressions can be satisfied.)

By (2.8) for n-j large enough we have

$$(3.4)\qquad (2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}} - b < \delta_{j,n} \ .$$

Specifically, for large n and $j<n-g_b(n)$ (3.4) holds.

To simplify notation we will denote

$$R_{i,n}(h_b(n),g_b(n)) \quad \text{as} \quad R_i$$

and

$$\underline{R}_i(h_b(n),g_b(n)) \quad \text{as} \quad \underline{R}_i \qquad i=1,5 \ .$$

Rules R_1 and R_5 choose differently at X_j only if

$$\delta_{j,n} + \overline{X}_{j-1} < X_j < \delta_{j,n}$$

or

$$\delta_{j,n} < X_j < \delta_{j,n} + \overline{X}_{j-1} \ .$$

If $\sup\limits_{1\leq k\leq n} \left| \dfrac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}} \right| \leq 2(\log\log 2n + C_\alpha)^{\frac{1}{2}}$

then $$|\bar{X}_{j-1}| \leq \frac{2(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(j-1)^{\frac{1}{2}}} \qquad j=2,\ldots,n+1 .$$

Hence, recalling (3.2) and (3.4),

$$(3.5) \qquad P(\{X_{R_1} \neq X_{R_5}\} \cap \{\sup_{1 \leq k \leq n} \left| \frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}} \right| < 2(\log\log 2n + C_\alpha)^{\frac{1}{2}}\})$$

$$\leq \sum_{j=h_b(n)+1}^{n-g_b(n)} P(\delta_{j,n} - \frac{2(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(j-1)^{\frac{1}{2}}} < X_j < \delta_{j,n})$$

$$\leq \sum_{j=h_b(n)+1}^{n-g_b(n)} \frac{1}{\sqrt{2\pi}} \frac{2(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(j-1)^{\frac{1}{2}}} e^{-\frac{((2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-2b)^2}{2}}$$

$$= K(b)(\log\log 2n + C_\alpha)^{\frac{1}{2}} \sum_{j=h_b(n)+1}^{n-g_b(n)} \frac{e^{2b(2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}}}{(n-j+1)(j-1)^{\frac{1}{2}}}$$

$$\leq K(b)(\log\log 2n + C_\alpha)^{\frac{1}{2}} \sum_{j=h_b(n)+1}^{n-g_b(n)} \frac{1}{(n-j+1)^{1-2b}(j-1)^{\frac{1}{2}}}$$

$$\leq \frac{K(b)(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(g_b(n)+1)^{\frac{1}{2}-2b}} \int_1^{n-g_b(n)} \frac{dx}{(n-x)^{\frac{1}{2}}x^{\frac{1}{2}}} .$$

Now $$\int_1^{n-g_b(n)} \frac{dx}{(n-x)^{\frac{1}{2}}x^{\frac{1}{2}}} = 2\tan^{-1}\sqrt{\frac{x}{n-x}}\,\Bigg|_1^{n-g_b(n)} \to \pi \text{ as } n\to\infty.$$

So using assumption (3.3),

$$\lim_{n\to\infty} P(\{X_{R_1} \neq X_{R_5}\} \cap \{\sup_{1\leq k\leq n} \left|\frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}}\right| < 2(\log\log 2n + C_\alpha)^{\frac{1}{2}}\}) = 0.$$

We have shown that for all n

$$P(\{X_{R_1} \neq X_{R_5}\} \cap \{\sup_{1\leq k\leq n} \left|\frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}}\right| > 2(\log\log 2n + C_\alpha)^{\frac{1}{2}}\}) \leq \alpha.$$

Thus $\overline{\lim}_{n\to\infty} P(X_{R_{1,n}} \neq X_{R_{5,n}}) \leq \alpha$ for all $\alpha > 0$, and $\underline{R}_5$ is asymptotically F.I.

Recalling the argument used in the proof of Corollary 2.1 it is clear that if R_5' is defined as $R_{5,n}'(h) = R_{5,n}(h,0)$ then $\underline{R}_5'$ is also asymptotically F.I. For this problem we can say more. The following lemma shows that the sequence $\underline{R}_5(2,0)$ is asymptotically F.I.

<u>Lemma 3.2</u>. Let $h\in\mathcal{D}'$ with $h(n) = o((n \log n)^{\frac{1}{2}})$. Then

i.) $\lim_{n\to\infty} P(Z_j < \delta_{j,n},\ j=2,\ldots,h(n)) = 1$

ii.) $\lim_{n\to\infty} P(X_j < \delta_{j,n},\ j=1,\ldots,h(n)) = 1$.

<u>Proof</u>. Since the Z's are independent (this is easily seen by applying Basu's Theorem),

$$P(Z_j<\delta_{j,n},\ j=2,\ldots,h(n)) = \prod_{j=2}^{h(n)} P(X_j-\overline{X}_{j-1}<\delta_{j,n})$$

$$= \prod_{j=2}^{h(n)} P(X_1<(\frac{j-1}{j})^{\frac{1}{2}}\delta_{j,n}) \geq \prod_{j=2}^{h(n)} \left(1- \frac{j^{\frac{1}{2}}e^{-\frac{j-1}{2j}\delta_{j,n}^2}}{(2\pi)^{\frac{1}{2}}(j-1)^{\frac{1}{2}}\delta_{j,n}}\right),$$

since for $x > 0$

$$\int_x^\infty e^{-\frac{u^2}{2}} du < \frac{1}{x} \int_x^\infty u e^{-\frac{u^2}{2}} = \frac{e^{-\frac{x^2}{2}}}{x} .$$

Let $\Lambda_{n,j} = \log\left(1 - \frac{j^{\frac{1}{2}} e^{-\frac{j-1}{2j}\delta_{j,n}^2}}{(2\pi)^{\frac{1}{2}}(j-1)^{\frac{1}{2}}\delta_{j,n}}\right)$.

By the above,

$$\log P(Z_j < \delta_{j,n},\ j=2,\ldots,h(n)) \geq \sum_{j=2}^{h(n)} \Lambda_{n,j} .$$

To prove i.) it suffices to show that

$$(3.6) \qquad \lim_{n\to\infty} \sum_{j=2}^{h(n)} \Lambda_{n,j} = 0.$$

$$\sum_{j=3}^{h(n)} \Lambda_{n,j} = - \sum_{j=3}^{h(n)} \sum_{k=1}^{\infty} \frac{j^{\frac{k}{2}}}{k(2\pi(j-1))^{\frac{k}{2}}\delta_{j,n}^k} e^{-\frac{k(j-1)}{2j}\delta_{j,n}^2} .$$

By (2.8), then,

$$\sum_{j=3}^{h(n)} \Lambda_{n,j}$$

$$\geq - \sum_{k=1}^{\infty} \sum_{j=3}^{h(n)} \frac{j^{\frac{k}{2}}}{(j-1)^{\frac{k}{2}}((2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-o(1))^k} \left(\frac{\sqrt{2\pi}}{n-j+1}\right)^{\frac{k(j-1)}{j}(1-o(1))}$$

$$\geq - \sum_{k=1}^{\infty} \frac{1}{(\log\frac{n-h(n)+1}{\sqrt{2\pi}})^{\frac{k}{2}}} \sum_{j=3}^{h(n)} \left(\frac{3\sqrt{2\pi}}{2(n-j+1)}\right)^{\frac{k}{2}} \quad \text{for large } n$$

$$\geq - \sum_{k=1}^{\infty} h(n) \left(\frac{3\sqrt{2\pi}}{2(\log\frac{n-h(n)+1}{\sqrt{2\pi}})(n-h(n)+1)} \right)^{\frac{k}{2}}$$

$$= \frac{-h(n)\, O\left(\frac{1}{(n \log n)^{\frac{1}{2}}}\right)}{1 - O\left(\frac{1}{(n \log n)^{\frac{1}{2}}}\right)} \to 0 \qquad \text{as } n \to \infty.$$

Clearly $\lim_{n\to\infty} \Lambda_{n,2} = 0$, so that (3.6) is true and i.) is proven. The same method works for showing ii.). //

By the definition of $h_b(n)$ it is clear that for any $b\in(0,\frac{1}{4})$ and $\alpha>0$ we may choose $h_b(n) = o((n \log n)^{\frac{1}{2}})$ so that Lemma 3.2 applies. Then i.) of Lemma 3.2 implies that the sequence of rules $\underline{R}_5'' = \underline{R}_5\ (2,0)$ is asymptotically F.I. We state this result as:

<u>Theorem 3.1.</u> In the $N(\mu,1)$ Problem the sequence of rules $\underline{R}_5'' = (R_{5,2}'', R_{5,3}'', \ldots)$ given by

$$R_{5,n}'' = \min\{n, \min_{k\geq 2}\{k: X_k=L_k, Z_k > \delta_{k,n}\}\}$$

is asymptotically F.I.

Note that Theorem 3.1 implies that all sequences in C_5 are asymptotically F.I.

The next theorem is a version of Theorem 3.1 for the $N(\mu,\sigma^2)$ Problem.

<u>Theorem 3.2.</u> In the $N(\mu,\sigma^2)$ Problem the sequence of invariant rules $\underline{R}_6(h(n),0)$ where $h(n)=O(n^{\varepsilon})$, some $\varepsilon>0$, is asymptotically F.I.

Theorem 3.2 states that sequences of rules in the class $\mathcal{C}_6' = \{\underline{R}_6(h,g): h(n)=O(n^{\varepsilon}), \text{ some } \varepsilon>0\} \subset \mathcal{C}_6$ are asymptotically F.I. In preparation for its proof we state and prove

Lemma 3.3. Let $0<\gamma<\frac{1}{2}$. For a suitable constant $C(\gamma)$, $h(n)\in\mathcal{D}'$ and $\xi_n = C(\gamma)\left(\frac{\log h(n)}{h(n)}\right)^{\gamma}$,

$$\lim_{n\to\infty} P(1-\xi_n<S_j<1+\xi_n,\ h(n)<j\le n) = 1.$$

Proof. Assume $a_j<j-2$, $U_j\sim\chi_j^2$. Then

$$(3.7)\qquad P(U_j<a_j) = \frac{1}{\Gamma(\frac{j}{2})2^{\frac{j}{2}}}\int_0^{a_j} u^{\frac{j-2}{2}} e^{-\frac{u}{2}}du$$

$$= \frac{(\frac{j-2}{2})^{\frac{j}{2}}}{\Gamma(\frac{j}{2})}\int_0^{\frac{a_j}{j-2}} [te^{-t}]^{\frac{j-2}{2}} dt$$

$$\le \frac{\left(\frac{a_j}{2}\right)^{\frac{j}{2}}}{\Gamma(\frac{j}{2})e^{\frac{a_j}{2}}} .$$

Recall Stirling's formula: $1 \le \frac{k!}{(\frac{k}{e})^k(2\pi k)^{\frac{1}{2}}} \le 1 + \frac{1}{12k-1}$.

If j is even then $j=2k$ and $\Gamma(\frac{j}{2}) = (k-1)!$ Thus

$$(3.8)\qquad \Gamma(\tfrac{j}{2}) \ge \frac{(k-1)^{k-1}}{e^{k-1}}(2\pi(k-1))^{\frac{1}{2}} = (\tfrac{j-2}{2})^{\frac{j-2}{2}}\frac{(\pi(j-2))^{\frac{1}{2}}}{e^{\frac{j-2}{2}}} .$$

If j is odd then for $j \ge 4$ $\Gamma(\frac{j}{2}) > \Gamma(\frac{j-1}{2})$ so that

$$(3.9)\qquad \Gamma(\tfrac{j}{2}) > (\tfrac{j-3}{2})^{\frac{j-3}{2}}\frac{(\pi(j-3))^{\frac{1}{2}}}{e^{\frac{j-3}{2}}} .$$

For $j>4$ the right hand side of (3.9) is smaller than the right hand side of (3.8) which implies that (3.9) holds for all $j>4$.

From (3.7) we may then write

$$(3.10)\quad P(U_j<a_j) \le \frac{e^{\frac{j-3}{2}}\left(\frac{a_j}{2}\right)^{\frac{j}{2}}}{\pi^{\frac{1}{2}}(j-3)^{\frac{1}{2}}\left(\frac{j-3}{2}\right)^{\frac{j-3}{2}} e^{\frac{a_j}{2}}} = \frac{1}{(2e)^{\frac{3}{2}}\pi^{\frac{1}{2}}} \frac{e^{\frac{j}{2}} a_j^{\frac{j}{2}} e^{-\frac{a_j}{2}}}{(j-3)^{\frac{j-2}{2}}} .$$

Fix $0<\alpha<1$. Since $(j-1)S_j^2 \sim \chi^2_{j-1}$, we propose to find for each large n a $\xi_n>0$ such that

$$(3.11)\quad P(U_j<(1-\xi_n)j) \le \frac{3\alpha}{\pi^2 j^2}, \qquad j>h(n) .$$

It would then follow that for each n

$$P(S_j^2>1-\xi_n,\ h(n)<j\le n) > 1 - \frac{\alpha}{2}$$

and, in fact, this probability would approach 1 as $n \to \infty$.

By (3.10) it suffices to find ξ_n satisfying

$$(3.12)\quad \frac{e^{\frac{j}{2}}(1-\xi_n)^{\frac{j}{2}} j^{\frac{j}{2}} e^{-\frac{(1-\xi_n)j}{2}}}{(2e)^{\frac{3}{2}}\pi^{\frac{1}{2}}(j-3)^{\frac{j-2}{2}}} \le \frac{3\alpha}{\pi^2 j^2}, \qquad j>h(n) .$$

(3.12) can be rewritten as

$$\frac{(j-3)}{(2e)^{\frac{3}{2}}\pi^{\frac{1}{2}}} \left[(1+\frac{3}{j-3})(1-\xi_n)e^{\xi_n}\right]^{\frac{j}{2}} \le \frac{3\alpha}{\pi^2 j^2}$$

or

$$\xi_n^{\,2}\left[1+\frac{2\xi_n}{3}+\frac{2\xi_n^{\,2}}{4}+\ldots\right] \geq \frac{4}{j}\log\left(\left(\frac{\pi}{2e}\right)^{\frac{3}{2}}\frac{j^2(j-3)}{3\alpha}\right)$$

$$+ 2\log\left(1+\frac{3}{j-3}\right), \quad j>h(n).$$

In particular (3.11) will be satisfied if we choose

(3.13)

$$\xi_n \geq \left[\frac{4}{h(n)}\log\left(\left(\frac{\pi}{2e}\right)^{\frac{3}{2}}\frac{h^2(n)(h(n)-3)}{3\alpha}\right) + 2\log\left(1+\frac{3}{h(n)-3}\right)\right]^{\gamma},\ 0<\gamma\leq\frac{1}{2}.$$

Since $(1-\xi_n)^2 < 1-\xi_n$ we have for ξ_n satisfying (3.13),

$$\lim_{n\to\infty} P(S_j > 1-\xi_n,\ h(n)<j\leq n) = 1.$$

We now find $0<\xi_n<1$ for which

$$\lim_{n\to\infty} P(S_j < 1+\xi_n,\quad h(n)<j\leq n) = 1.$$

Wallace [18] derived bounds on the upper tail of the χ^2 distribution in terms of the standard normal distribution. His Theorem 4.1 gives

$$P(U_j>t) < \frac{\left(\frac{j}{2}\right)^{\frac{j-1}{2}} e^{-\frac{j}{2}}}{\Gamma\left(\frac{j}{2}\right)} \int_{\left(t-j-j\log\left(\frac{t}{j}\right)\right)^{\frac{1}{2}}}^{\infty} e^{-\frac{u^2}{2}}\,du\ .$$

Recall that $x>0$ implies $\int_x^{\infty} e^{-\frac{u^2}{2}}\,du < \frac{e^{-\frac{x^2}{2}}}{x}$.

Then by (3.9) and the above

$$P(U_j>(1+\xi_n)j) < \frac{\left(\frac{j}{2}\right)^{\frac{j-1}{2}} e^{-\frac{j}{2}}}{\Gamma\left(\frac{j}{2}\right)} \int_{\left(\xi_n j-j\log(1+\xi_n)\right)^{\frac{1}{2}}}^{\infty} e^{-\frac{u^2}{2}}\,du$$

$$\leq \frac{(\frac{j}{2})^{\frac{j-1}{2}} e^{-\frac{j}{2}} e^{\frac{j-3}{2}}}{(\frac{j-3}{2})^{\frac{j-3}{2}} \pi^{\frac{1}{2}} (j-3)^{\frac{1}{2}}} \frac{(e^{-\xi_n}(1+\xi_n))^{\frac{j}{2}}}{j^{\frac{1}{2}}(\xi_n-\log(1+\xi_n))^{\frac{1}{2}}}$$

$$= \frac{e^{-\frac{3}{2}}(1+\frac{3}{j-3})^{\frac{j-2}{2}} (e^{-\xi_n}(1+\xi_n))^{\frac{j}{2}}}{(2\pi)^{\frac{1}{2}}(\xi_n{}^2 - \frac{2\xi_n{}^3}{3} + \frac{2\xi_n{}^4}{4} - \cdots)^{\frac{1}{2}}}$$

$$\leq \frac{(1+\frac{3}{j-3})^{\frac{1}{2}} (e^{-\xi_n}(1+\xi_n))^{\frac{j}{2}}}{(2\pi)^{\frac{1}{2}}\xi_n(1-\frac{2\xi_n}{3} + \frac{2\xi_n{}^2}{4} - \cdots)^{\frac{1}{2}}}$$

$$\leq \frac{3\alpha}{\pi^2 j^2}$$

if

$$(3.14) \qquad \log \frac{3(2\pi)^{\frac{1}{2}}\alpha\xi_n(1-\frac{2\xi_n}{3} + \frac{2\xi_n{}^2}{4} - \cdots)^{\frac{1}{2}}}{\pi^2 j^2} \geq$$

$$-\frac{j}{2}(\frac{\xi_n{}^2}{2} - \frac{\xi_n{}^3}{3} + \ldots) + \frac{1}{2}\log(1+\frac{3}{j-3}) .$$

For large n (3.14) will hold for $j>h(n)$ if we choose

$$(3.15) \qquad \xi_n = 2\left(\frac{\log h(n)}{h(n)}\right)^{\gamma}(1+o(1)) \qquad 0<\gamma<\frac{1}{2} .$$

To see this we first note that the left hand side of (3.14) is increasing in j with respect to its right hand side if $\frac{1}{j}(2-\frac{3}{2(j-3)}) < \frac{1}{2}(\frac{\xi_n{}^2}{2} - \frac{\xi_n{}^3}{3} + \cdots)$. So by choosing ξ_n as in (3.15) this relation is certainly satisfied for $j\geq h(n)$ if n is large. Thus it suffices to show (3.14) for $j=h(n)$

and ξ_n as in (3.15). That is, we must show $\log c +$ $\log \xi_n + \log(1-o(1)) - 2\log h(n) \geq \frac{-h(n)}{4}\xi_n^2(1-o(1)) + o(1)$, or

$$\log c' + \gamma \log\log h(n) - (2+\gamma)\log h(n)$$
$$\geq -(h(n))^{1-2\gamma}(\log h(n))^{2\gamma}(1-o(1)) + o(1).$$

This last becomes clear upon dividing both sides by $(h(n))^{1-2\gamma}$ and letting n become large.

By letting $o(1) = \frac{1}{2}(\frac{\log h(n)}{h(n)})^{\gamma}$ in (3.15) we have $(1+\xi_n) = (1+(\frac{\log h(n)}{h(n)})^{\gamma})^2$. Thus, from (3.13) and the above, by choosing a suitable constant $C(\gamma)$ and setting $\xi_n = C(\gamma)(\frac{\log h(n)}{h(n)})^{\gamma}$ we have

$$\lim_{n\to\infty} P(1-\xi_n < S_j < 1+\xi_n,\quad h(n)<j\leq n) = 1 . \qquad //$$

<u>Proof of Theorem 3.2.</u>

The spirit of this proof is close to that of the previous theorem. Let us denote

$$R_{i,n}(h,g) \quad \text{by} \quad R_i$$

and

$$\underline{R}_i(h,g) \quad \text{by} \quad \underline{R}_i \qquad i=5,6 ,$$

where we assume $h,g\in\mathcal{D}'$. For a given pair (h,g) R_5 and R_6 can differ in their choice or rejection of X_j only when

$$S_{j-1}\,\delta_{j,n} < X_j-\overline{X}_{j-1} < \delta_{j,n}$$

or

$$\delta_{j,n} < X_j-\overline{X}_{j-1} < S_{j-1}\delta_{j,n} .$$

Let ξ_n satisfy the conditions of Lemma 3.3. Then

$$\{\{R_5=j, R_6\neq j\} \cup \{R_5\neq j, R_6=j\}\} \cap \{|S_{j-1}-1|<\xi_n\}\}$$

$$\subset \{(1-\xi_n)\delta_{j,n} < X_j-\overline{X}_{j-1} < (1+\xi_n)\delta_{j,n}\},$$

and therefore

$$P(\{X_{R_5}\neq X_{R_6}\} \cap \{|S_{j-1}-1|<\xi_n\},\ h(n)<j\leq n-g(n)\})$$

$$\leq P(\bigcup_{j=h(n)+1}^{n-g(n)} \{(1-\xi_n)\delta_{j,n}<X_j-\overline{X}_{j-1}<(1+\xi_n)\delta_{j,n}\})$$

$$= 1-\prod_{j=h(n)+1}^{n-g(n)} P(X_j-\overline{X}_{j-1} \notin ((1-\xi_n)\delta_{j,n},(1+\xi_n)\delta_{j,n}))$$

$$= 1-\prod_{j=h(n)+1}^{n-g(n)} P(X_1\notin ((\tfrac{j-1}{j})^{\frac{1}{2}}(1-\xi_n)\delta_{j,n},(\tfrac{j-1}{j})^{\frac{1}{2}}(1+\xi_n)\delta_{j,n})) .$$

Let $\lambda_j = P(X_1 \in ((\frac{j-1}{j})^{\frac{1}{2}}(1-\xi_n)\delta_{j,n},(\frac{j-1}{j})^{\frac{1}{2}}(1+\xi_n)\delta_{j,n}))$

and assume $\xi_n n^{\varepsilon} \to 0$ for some $\varepsilon>0$ as $n\to\infty$. Then for

$h(n)<j\leq n-g(n)$

$$\lambda_j \leq (\frac{j-1}{j})^{\frac{1}{2}}\frac{2\xi_n}{(2\pi)^{\frac{1}{2}}}(2\log\frac{n-j}{\sqrt{2\pi}\,C_o})^{\frac{1}{2}} e^{-\frac{j-1}{2j}(1-\xi n)^2((\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-o(1))^2} .$$

So

$$(3.16)\qquad 0 \geq \sum_{j=h(n)+1}^{n-g(n)} \log(1-\lambda_j) \geq$$

$$-\sum_{k=1}^{\infty}\sum_{j=h(n)+1}^{n-g(n)} \frac{2^k\xi_n^{\,k}(j-1)^{\frac{k}{2}}(2\log\frac{n-j}{\sqrt{2\pi}\,C_o})^{\frac{k}{2}}}{k\, j^{\frac{k}{2}}\,(2\pi)^{\frac{k}{2}}} e^{-\frac{k(j-1)}{2j}(1-\xi_n)^2((2\log\frac{n-j+1}{\sqrt{2\pi}})^{\frac{1}{2}}-o(1))^2}$$

$$\geq -\sum_{k=1}^{\infty} \frac{2^{\frac{k}{2}}\xi_n^k}{\pi^{\frac{k}{2}}k} \sum_{j=h(n)+1}^{n-g(n)} \left\{\left(\frac{\sqrt{2\pi}}{n-j+1}\right)^{\frac{j-1}{j}(1-o(1))} \left(2\log\frac{n-j}{\sqrt{2\pi}\,C_o}\right)^{\frac{1}{2}}\right\}^k$$

$$\geq -\sum_{k=1}^{\infty} \frac{2^{\frac{k}{2}}\xi_n^k}{\pi^{\frac{k}{2}}k} \sum_{j=h(n)+1}^{n-g(n)} \left(\frac{\sqrt{2\pi}}{n-j+1}\right)^{k(1-\eta)} \left(2\log\frac{n-j}{\sqrt{2\pi}\,C_o}\right)^{\frac{k}{2}}$$

for $0<\eta<\varepsilon$ and large n.

Since

$$\sum_{j=h(n)+1}^{n-g(n)} \frac{\left(\log\frac{n-j}{\sqrt{2\pi}\,C_o}\right)^{\frac{1}{2}}}{(n-j+1)^{1-\eta}} \quad \text{dominates}$$

$$\sum_{j=h(n)+1}^{n-g(n)} \frac{\left(\log\frac{n-j}{\sqrt{2\pi}\,C_o}\right)^{\frac{k}{2}}}{(n-j+1)^{k(1-\eta)}} \quad \text{for all } k > 1$$

and since the asymptotic behavior of the former quantity is between that of n^η and $n^\eta \log n$ we see that

$$\lim_{n\to\infty} \sum_{j=h(n)+1}^{n-g(n)} \log(1-\lambda_j) = 0, \text{ and therefore}$$

$$\lim_{n\to\infty} P(\{X_{R_{5,n}} \neq X_{R_{6,n}}\} \cap \{|S_{j-1}-1|<\xi_n,\ h(n)<j\leq n\}) = 0.$$

By Lemma 3.3 then

(3.17) $$\lim_{n\to\infty} P(X_{R_{5,n}} \neq X_{R_{6,n}}) = 0$$

and thus $\underline{R}_6$ is asymptotically F.I.

Choosing $h(n) = [n^{\frac{\varepsilon}{\gamma}+\delta}]$ for $1 - \frac{\varepsilon}{\gamma} > \delta > 0$ ($[\cdot]$ is the greatest integer function) implies that $n^{\varepsilon}\xi_n \to 0$ as $n\to\infty$ and therefore that (3.17) holds.

Finally, using the same argument as in the proof of Corollary 2.1, it is easily seen that we may take $g(n) \equiv 0$ and still have $\underline{R}_6(h,0)$ asymptotically F.I. //

Remarks

As was indicated earlier exact values for the decision numbers $\{\delta_{i,n}\}_{i=1}^{n}$ are difficult to obtain even for moderate values of n. Gilbert and Mosteller [7] found approximations to the δ's which are close to the true values, are easily computed, and give for the F.I. Problem a sequence $\underline{\tau} = (\tau_1, \tau_2, \ldots)$ of stopping rules such that $\lim_{n\to\infty} P(X_{\tau_n} = L_n) = \alpha_o$. It can be shown that Theorems 3.1 and 3.2 are valid when the δ's are replaced by these approximations in the rules R_5 and R_6. Thus there are easily computed asymptotically F.I. Rules for both the $N(\mu,1)$ and $N(\mu,\sigma^2)$ Problems.

As was the case with the theorems of Chapter II, Theorems 3.1 and 3.2 implicitly give rates of convergence of $P(X_{R_{5,n}} = L_n)$ and $P(X_{R_{6,n}} = L_n)$ to the value α_o.

In the first case we know

$$\delta_{j,n} \geq \left(2 \log \frac{n-j+1}{\sqrt{2\pi}}\right)^{\frac{1}{2}} - o_g(1) \quad \text{where, if } j \leq n-g(n).$$

Lemma 2.2 implies

$$o_g(1) = C_g(n) \frac{\log\log g(n)}{(\log g(n))^{\frac{1}{2}}} \text{ with}$$

$$\lim_{n\to\infty} C_g(n) = \frac{1}{2\sqrt{2}} .$$

If in the proof of Theorem 3.1 we let

$b_n = C_g(n) \dfrac{\log\log g(n)}{(\log g(n))^{\frac{1}{2}}}$ take the place of b then (3.3)

becomes

$$(\log\log 2n)^{t(n)} = o(g(n))$$

$$\text{where} \quad t(n) = \frac{(\log g(n))^{\frac{1}{2}}}{(\log g(n))^{\frac{1}{2}} - 4C_g(n) \log\log g(n)} .$$

This will be true if for some $\varepsilon > 0$ $(\log\log n)^{1+\varepsilon} = o(g(n))$.

We also require that $h \in \mathcal{D}'$ be such that

$$h(n) \geq \min\left\{1 \leq j \leq n-1 : \frac{2(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(j-1)^{\frac{1}{2}}} < C_g(n) \frac{\log\log g(n)}{(\log g(n))^{\frac{1}{2}}}\right\}$$

or equivalently

$$h(n) \geq \frac{4(\log\log 2n + C_\alpha) \log g(n)}{C_g^2(n) (\log\log g(n))^2} + 1 .$$

This makes it clear that the minimum size of h depends on g. If, for example, $g(n) = [\log n]$ we may take

$$h(n) \leq \frac{\beta_1 (\log\log n)^2}{(\log\log\log n)^2}$$

for β_1 a suitable constant. If $g(n)=[n^r]$ for some $0<r<1$ then we may take

$$h(n) \leq \frac{\beta_2 \log n}{\log \log n}$$

for β_2 a suitable constant.

If we use the above values $b=b_n$ in the computations following (3.5) we have

$$P(\{X_{R_1} \neq X_{R_5}\} \cap \{\sup_{1 \leq k \leq n} \left| \frac{\sum_{\ell=1}^{k} X_\ell}{\sqrt{k}} \right| < 2(\log\log 2n + C_\alpha)^{\frac{1}{2}}\})$$

$$\leq C_n \frac{(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(g(n)+1)^{\frac{1}{2} - 2b_n}}$$

where $C_n \leq 2\pi$ for all n.

Therefore

$$\alpha_o - P(X_{R_{5,n}} = L_n) \leq \frac{g(n)+h(n)}{n} + C_n \frac{(\log\log 2n + C_\alpha)^{\frac{1}{2}}}{(g(n)+1)^{\frac{1}{2} - 2b_n}} .$$

If we turn our attention to $\alpha_o - P(X_{R_{6,n}} = L_n)$ we see from (3.16) that

$$0 \geq \sum_{j=h(n)+1}^{n-g(n)} \log (1-\lambda_j)$$

$$\geq \frac{-2^{\frac{3}{2}} \xi_n \left(\log \frac{n-h(n)}{\sqrt{2\pi}\, C_o}\right)^{\frac{1}{2}} n^{\eta}(1-o(1))}{(2\pi)^{\frac{\eta}{2}} \eta}$$

$$- \sum_{k=2}^{\infty} \frac{2^{\frac{k}{2}} \xi_n{}^k \left(2 \log\left(\frac{n-h(n)}{\sqrt{2\pi}\, C_o}\right)\right)^{\frac{1}{2}} (2\pi)^{\frac{k(1-\eta)}{2}}}{\pi^{\frac{k}{2}}\, k} \int_1^{n-g(n)+1} \frac{dx}{(n-x+1)^{1-\eta}}$$

$$\geq \frac{-2^{\frac{3}{2}} \xi_n \left(\log \frac{n-h(n)}{\sqrt{2\pi}\, C_o}\right)^{\frac{1}{2}}}{(2\pi)^{\frac{\eta}{2}}\, \eta}\, n^{\eta} (1+o(1))$$

$$\geq -\, K(\eta) \left(\log \frac{n-h(n)}{\sqrt{2\pi}\, C_o}\right)^{\frac{1}{2}} n^{\eta} \xi_n$$

for large n.

Thus $P(\{X_{R_{6,n}} \neq X_{R_{5,n}}\} \cap \{|S_{j-1}-1| < \xi_n,\ h(n) \leq j \leq n\})$

$$\leq 1 - e^{-K(\eta) \left(\log \frac{n-h(n)}{\sqrt{2\pi}\, C_o}\right)^{\frac{1}{2}} n^{\eta} \xi_n}.$$

2. Best Invariant Rules

So far our search for asymptotically F.I. sequences of rules has led us to invariant sequences in the $N(\mu,1)$ and $N(\mu,\sigma^2)$ Problems. We showed in Chapter II that there exist asymptotically F.I. invariant sequences of rules in these problems, and in the first section of this chapter we obtained the same result for a more sophisticated class of rules.

Based upon these results it is natural to inquire about the best that one can do with invariant rules in these problems. This section pursues this question. The existence of best invariant rules is guaranteed by the theory of sequential decision problems and such rules can be constructed by the method of backward induction (see Ferguson [5] whose treatment we follow here). We begin by setting the problem in a formal decision theoretic framework, and then move on to deriving the form of a best invariant rule. Difficulties arise in trying to obtain an exact rule since the expressions giving us this best invariant rule prove to be intractable for even small values of n. Our efforts to cope with these difficulties are the topic of the next section.

Fix $n \geq 2$. For each $2 \leq i \leq n$ define the group of location-scale changes on $\mathbb{R}^i$ as

$$G_i = \{g_{ab}:\mathbb{R}^i \to \mathbb{R}^i \text{ where } g_{ab}(x_1,\ldots,x_i) = (ax_1+b,\ldots,ax_i+b),\ a>0,\ b\in\mathbb{R}\}.$$

G_i will be considered the group of transformations on the space of observations $(X_1,\ldots,X_i)$. The group G_i induces a group $\overline{G}$ acting on the parameter space $\Theta = \{(\mu,\sigma)\}_{\substack{\mu\in\mathbb{R}\\ \sigma>0}}$ as $\overline{g}_{ab}(\mu,\sigma) = (a\mu+b,a\sigma)$. $\overline{G}$ is induced by requiring that the distribution of $g_{ab}(X_1,\ldots,X_i)$ under F_θ be the same as that of $(X_1,\ldots,X_i)$ under $F_{\overline{g}_{ab}(\theta)}$. A third group required in this formulation is a group $\tilde{G}$ of transformations on the action space A. Because our actions in this problem consist of choosing one and only one of $X_1,\ldots,X_n$ as our candidate for the largest observation we will take $A = \{1,\ldots,n\}$ where action i corresponds to the action "choose X_i". We will take the group $\tilde{G}$ to be the identity transformation on A.

A sequential decision rule consists of a stopping rule and a terminal decision rule. The stopping rule tells us when to stop sampling and the terminal decision rule tells us what action to take once we have stopped sampling. The conditions of the problem allow us to simplify this structure. Because no recall is allowed and because we must accept or reject observation X_i as it is observed, we need only specify that stopping at X_i requires that we take action i. Thus the terminal decision rule is given by $\delta_i(X_1,\ldots,X_i) = i$ and a solution of the problem will be a determination of when to optimally stop.

If we stop at X_i the gain to the observer is $I(X_i=L_n)$; that is, 1 if X_i is the largest observation and 0 otherwise. Clearly this gain is invariant under the action of groups $\overline{G}$ and $\tilde{G}$ on the parameter and action spaces. This fact along with the existence of the group $\overline{G}$ assures us that the $N(\mu,\sigma^2)$ Problem of length n is, in Ferguson's terminology, an invariant sequential decision problem truncated at n.

To find a best invariant rule our next step is to find a maximal invariant under G_i for i=2,...,n. It is easy to show that $\underline{Y}_i$ is such a maximal invariant. In addition an application of Basu's Theorem shows that the components $Y_2,...,Y_i$ are mutually independent.

In what follows it is helpful to refer to the backward induction method for the F.I. Problem. An outline of this is given in Section 2 of Chapter I. We assume a N(0,1) distribution without loss of generality. $\underline{Y}_i$ is our invariant state of knowledge at step i. Given $\underline{Y}_i$, our expected gain by stopping (and therefore choosing X_i) is

$$U_{i,n}(\underline{Y}_i) = P(X_i=L_n|\underline{Y}_i) \ .$$

If, on the other hand, we elect to observe X_{i+1} the best we can expect to do is given by

$$E(V_{i+1,n}(\underline{Y}_{i+1})|\underline{Y}_i)$$

where V is defined recursively as

$$V_{n,n} = U_{n,n}$$

$$(3.18) \quad V_{i+1,n}(\underline{Y}_{i+1}) = \max\{U_{i+1,n}(\underline{Y}_{i+1}), E(V_{i+2,n}(\underline{Y}_{i+2})|\underline{Y}_{i+1})\}$$

Then a best invariant rule for the $N(\mu,\sigma^2)$ Problem of length n is given by

$$(3.19) \quad \tau_n = \min\{n, \min_{j\geq 2}\{j: U_{j,n}(\underline{Y}_j) > E(V_{j+1,n}(\underline{Y}_{j+1})|\underline{Y}_j)\}\}.$$

The independence of the Y's and of Y_j and $X_{j+1},...,X_n$ allows us to write

$$U_{n-1,n}(\underline{Y}_{n-1}) = I(X_{n-1}=L_{n-1})P(X_n<X_{n-1}|\underline{Y}_{n-1})$$

$$= I(X_{n-1}=L_{n-1})P(X_n<X_{n-1}|Y_{n-1})$$

$$= I(X_{n-1}=L_{n-1})T_{n-1,n}(Y_{n-1}) .$$

$$E(V_{n,n}(\underline{Y}_n)|\underline{Y}_{n-1}) =$$

$$I(X_{n-1}=L_{n-1})E(P(X_n\geq X_{n-1}|Y_n,Y_{n-1})|Y_{n-1})$$

$$+ I(X_{n-1}\neq L_{n-1})E(P(X_n=L_n|\underline{Y}_n)|\underline{Y}_{n-1})$$

$$= I(X_{n-1}=L_{n-1})H_{n-1,n}(Y_{n-1})+I(X_{n-1}\neq L_{n-1})G_{n-1,n}(\underline{Y}_{n-1}) .$$

Since $G_{n-1} > 0$ (3.19) implies

$$\tau_n I(\tau_n>n-2)=(\min\{n,\min_{j>n-2}\{j:X_j=L_j,T_{j,n}(Y_j)>H_{j,n}(Y_j)\}\})I(\tau_n>n-2) .$$

Continuing to the next step in the induction,

$$U_{n-2,n}(\underline{Y}_{n-2}) = I(X_{n-2}=L_{n-2})P(X_{n-1},X_n<X_{n-2}|\underline{Y}_{n-2})$$

$$= I(X_{n-2}=L_{n-2})P(X_{n-1},X_n<X_{n-2}|Y_{n-2})$$

$$= I(X_{n-2}=L_{n-2})T_{n-2,n}(Y_{n-2}) .$$

$$E(V_{n-1,n}(\underline{Y}_{n-1})|\underline{Y}_{n-2}) =$$

$$E(I(X_{n-1}=L_{n-1})\max\{P(X_n<X_{n-1}|Y_{n-1}),P(X_n\geq X_{n-1}|Y_{n-1})\}$$

$$+ I(X_{n-1}\neq L_{n-1})P(X_n=L_n|\underline{Y}_{n-1})|\underline{Y}_{n-2})$$

$$=E(I(X_{n-2}=L_{n-2})I(X_{n-1}\geq X_{n-2})\max\{T_{n-1,n}(Y_{n-1}),H_{n-1,n}(Y_{n-1})\}|\underline{Y}_{n-2})$$

$$+ E(I(X_{n-2}=L_{n-2})I(X_{n-1}<X_{n-2})P(X_n\geq X_{n-2}|\underline{Y}_{n-1})|\underline{Y}_{n-2})$$

$$+ E(I(X_{n-2}\neq L_{n-2})[I(X_{n-1}=L_{n-1})\max\{T_{n-1,n}(Y_{n-1}),H_{n-1,n}(Y_{n-1})\}$$

$$+ I(X_{n-1}\neq L_{n-1})P(X_n=L_n|\underline{Y}_{n-1})]|\underline{Y}_{n-2})$$

$$= I(X_{n-2}=L_{n-2})E(I(X_{n-1}\geq X_{n-2})\max\{T_{n-1,n}(Y_{n-1}),H_{n-1,n}(Y_{n-1})\}|Y_{n-2})$$

$$+ I(X_{n-2}=L_{n-2})E(I(X_{n-1}<X_{n-2})P(X_n\geq X_{n-2}|Y_{n-1},Y_{n-2})|Y_{n-2})$$

$$+ I(X_{n-2}\neq L_{n-2})G_{n-2,n}(\underline{Y}_{n-2})$$

$$= I(X_{n-2}=L_{n-2})H_{n-2,n}(Y_{n-2})+I(X_{n-2}\neq L_{n-2})G_{n-2,n}(\underline{Y}_{n-2}) .$$

Letting $M_{n-1,n}(Y_{n-1}) = \max\{T_{n-1,n}(Y_{n-1}),H_{n-1,n}(Y_{n-1})\}$,

$$H_{n-2,n}(Y_{n-2})=E(I(X_{n-1}\geq X_{n-2})M_{n-1,n}(Y_{n-1})|Y_{n-2})$$

$$+E(I(X_{n-1}<X_{n-2})P(X_n\geq X_{n-2}|Y_{n-1},Y_{n-2})|Y_{n-2}).$$

Since $G_{n-2,n} > 0$ we have

$$\tau_n I(\tau_n>n-3)=(\min\{n,\min_{j>n-3}\{j:X_j=L_j,T_{j,n}(Y_j)>H_{j,n}(Y_j)\}\})I(\tau_n>n-3) .$$

In general we find for $k=1,\dots,n-2$

$$U_{n-k,n}(\underline{Y}_{n-k})=I(X_{n-k}=L_{n-k})P(X_j<X_{n-k},j=n-k+1,\dots,n|\underline{Y}_{n-k})$$

$$=I(X_{n-k}=L_{n-k})P(X_j<X_{n-k},j=n-k+1,\dots,n|Y_{n-k})$$

$$=I(X_{n-k}=L_{n-k})T_{n-k,n}(Y_{n-k}) .$$

The V's are computed recursively according to (3.18) giving

$$E(V_{n-k+1,n}(\underline{Y}_{n-k+1})|\underline{Y}_{n-k}) =$$

$$I(X_{n-k}=L_{n-k})H_{n-k,n}(Y_{n-k}) + I(X_{n-k}\neq L_{n-k})G_{n-k,n}(\underline{Y}_{n-k}).$$

Letting

$$M_{j,n}(Y_j)=\max\{H_{j,n}(Y_j),T_{j,n}(Y_j)\},$$

$$\Phi^{(k)}_{j,n}(Y_{n-k})$$
$$= E(\prod_{\ell=1}^{k-j-1} I(X_{n-k+\ell}<X_{n-k})I(X_{n-j}\geq X_{n-k})M_{n-j,n}(Y_{n-j})|Y_{n-k})$$

$$j=1,\ldots,k-1 ,$$

and

$$\Phi^{(k)}_{k,n}(Y_{n-k})=E(\prod_{\ell=1}^{k-1} I(X_{n-k+\ell}<X_{n-k})I(X_n\geq X_{n-k})|Y_{n-k}) ,$$

we have

$$H_{n-k,n}(Y_{n-k}) = \sum_{j=1}^{k} \Phi^{(k)}_{j,n} (Y_{n-k}) .$$

Since $G_{n-k} > 0$ we obtain

$$\tau_n I(\tau_n>n-k-1)$$

$$= (\min\{n, \min_{j>n-k-1} \{j:X_j=L_j,T_{j,n}(Y_j)>H_{j,n}(Y_j)\}\})I(\tau_n>n-k-1)$$

By taking k=n-1 we obtain the best invariant rule given by (3.19).

The same analysis applies to the $N(\mu,1)$ Problem. A best invariant rule will be obtained if the Y's in the above argument are replaced by the corresponding Z's.

It seems clear that there are decision numbers $\{b_{j,n}\}_{n=2}^{n}$ for each $n\geq 3$ such that

$$\tau_n = \min\{n,\min_{j\geq 2}\{j:X_j=L_j,Y_j>b_{j,n}\}\}.$$

Though we offer no formal proof of this assertion, we will outline the reason we believe it to be true.

If $X_j=L_j$ then X_j is Y_j sample standard deviations above the sample mean. A large value of Y_j indicates that to the best of our (invariant) knowledge at time j, X_j is large relative to its true distribution. In particular if one is willing to choose an X_j which is x sample standard deviations above the sample mean then one should be willing to choose an X_j which is $x+\delta$ sample standard deviations above the sample mean for any $\delta>0$.

Finally, we note that for n=1,2 a best invariant rule is based only on relative ranks and therefore is the N.I. rule.

3. Small Sample Sizes

The intractability of the expressions defining the best invariant rules of the last section makes it extremely difficult (if not impossible) to obtain exact values of $P(X_{\tau_n}=L_n)$ for even small n. Our approach to this difficulty has been to guess a seemingly reasonable invariant rule σ_n and then to obtain an estimate of $P(X_{\sigma_n}=L_n)$ through a Monte Carlo procedure.

For reasons given in the last section we expect a good invariant rule to be of the form

$$\sigma_n = \min\{n, \min_{j\geq 2}\{j: X_j=L_j, Y_j>b_{j,n}\}\},$$

and consequently we confine ourselves to rules of this form. Guessing a reasonable rule is then equivalent to guessing a reasonable set of decision numbers $\{b_{j,n}\}_{j=2}^{n-1}$.

We consider two rules σ_1 and σ_2 obtained in this way, where for i=1,2

$$\sigma_{i,n} = \min\{n, \min_{j\geq 3}\{j: X_j = L_j, Y_j > c_{j,n}^{(i)}\}\} .$$

The decision numbers $c_{j,n}^{(i)}$ are defined as follows:

We first note that $\sqrt{\frac{j-1}{j}}\, Y_j$ has a Student's t distribution with j-2 degrees of freedom. Let T_j be the c.d.f. of Y_j. Let $\{d_{j,n}^{(1)}\}_{j=1}^{n-1}$ be the decision numbers for the F.I. Rule when a U[0,1] distribution is assumed.

Consider the P.I. Problem defined by the location parameter family of distributions $\{G_\theta\}_{\theta\in\mathbb{R}}$ where G_θ is the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ c.d.f. In Chapter IV a best invariant rule of length n is shown to be

$$\tau_n = \min\{n, \min_{j\geq 2}\{j: X_j = L_j, R_j > d_{j,n}^{(2)}\}\}$$

where R_j is the range of $X_1,\ldots,X_j$. The numbers $d_{j,n}^{(2)}$ are computable.

We define $c_{j,n}^{(i)} = T_j^{-1}(d_{j,n}^{(i)})$, i=1,2.

For n=25, 50 and i=1,2 the quantity $P(X_{\sigma_{i,n}} = L_n)$ was estimated by a Monte Carlo procedure using the CDC 6500 computer at Purdue. For each of n=25, 50, 50,000 pseudo random sequences of n N(0,1) observations were generated using the RVP routine. The proportion of successes (i.e. the proportion of sequences in which rule $\sigma_{i,n}$ selected the largest observation) was computed in each case. These numbers are displayed in Table 3.1.

To facilitate comparison Table 3.1 also lists the true probability of choosing the largest for both the N.I. and F.I. Rules of lengths 25 and 50. These numbers are taken from Gilbert and Mosteller [7]. As the true value for the F.I. Rule of length 25 is not listed in [7] the value listed in Table 3.1 has been obtained by interpolation.

It can be seen that both rules σ_1 and σ_2 offer definite improvements over the N.I. Rule for each n. It must be mentioned that as we do not know the values for the best invariant rule we cannot say exactly how significant the improvement is. It is certainly not impressive when compared with the F.I. values.

For both sample sizes σ_2 improves upon σ_1. Perhaps this is because the decision numbers $d_{j,n}^{(2)}$ take into account some of the same variability (due to an unknown location parameter) as is found in the $N(\mu,\sigma^2)$ Problem.

Table 3.1. Estimated Probability of Choosing L_n for Rules $\sigma_{1,n}$ and $\sigma_{2,n}$; True Probability of Choosing L_n for the N.I. and F.I. Rules

		σ_1	σ_2	N.I.	F.I.
Sample	25	.3961	.4294	.3809	.5914 (approx.)
Size n	50	.4187	.4442	.3743	.5857

CHAPTER IV

THE UNIFORM PROBLEM

In this chapter we investigate the P.I. Problem determined by the location parameter family of distributions $\{G_\theta\}_{\theta\in\mathbb{R}}$, where G_θ is the c.d.f. of the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ distribution. As in the previous chapters our approach is to consider stopping rules that are invariant under location changes. For the problem of length n we obtain the form of a best invariant rule τ_n (which is also minimax - see Chapter I), and we derive exact expressions for obtaining the optimal decision numbers and for finding $P_\theta(X_{\tau_n}=L_n)$ (which, since τ_n is invariant, is constant in θ) for $n=1,\ldots,50$.

In the last section we show that the sequence $\underline{\tau}=(\tau_1,\tau_2,\ldots)$ asymptotically improves upon the sequence of optimal rules for the N.I. Problem but is not asymptotically F.I. (i.e. $\frac{1}{e} < \underline{\lim}_{n\to\infty} P_\theta(X_{\tau_n}=L_n) \leq \overline{\lim}_{n\to\infty} P_\theta(X_{\tau_n}=L_n)<\alpha_o$).

To set notation for this chapter let:

$X_1,X_2,\ldots \sim$ i.i.d. $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ be the observations

$\underline{Y}_i = (X_2-X_1,X_3-X_1,\ldots,X_i-X_1) \qquad i=2,3,\ldots$

$L_i = \max\{X_1,\ldots,X_i\}$

$D_i = \min\{X_1,\ldots,X_i\}$.

$M_i = \frac{L_i+D_i}{2}$ (the midrange of $X_1,\ldots,X_i$)

$R_i = L_i-D_i$ (the range of $X_1,\ldots,X_i$)

G_θ = c.d.f. of the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ distribution

$\mathcal{F}$ = {invariant stopping rules for the X's}

$\mathcal{F}_{j,n} = \{\sigma\in\mathcal{F}: j < \sigma \leq n\}$.

$E_\theta(P_\theta)$ denotes expectation (probability) taken with respect to dG_θ.

1. The Form of a Best Invariant Rule

The P.I. Problem defined by the location parameter family $\{G_\theta\}_{\theta\in\mathbb{R}}$ (hereafter known as the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ Problem) is clearly invariant under location changes and a decision theoretic formulation which exactly parallels that of the $N(\mu,1)$ Problem may be given (see Chapter III Section 2).

It is easily seen that $\underline{Y}_i$ is a maximal invariant under $\mathcal{G}_i$, the group of location changes on $\mathbb{R}^i$. We also note that the range R_i is $\underline{Y}_i$ measurable.

We will assume $\theta=0$ and $n\geq 3$. The joint distribution of R_i and M_i is then given by the density

(4.1) $$h(M_i,R_i) = i(i-1)\ R_i^{i-2} I_{\left[-\frac{1-R_i}{2},\ \frac{1-R_i}{2}\right]}(M_i) I_{[0,1]}(R_i),$$

from which it may be seen that the distribution of R_i is given by the density

(4.2) $$f(R_i) = i(i-1)R_i^{i-2}(1-R_i)I_{[0,1]}(R_i)\ .$$

The distribution of M_i given $\underline{Y}_i$ is the same as that of M_i given R_i which, we infer from (4.1) and (4.2), is uniform on $\left[-\frac{1-R_i}{2}, \frac{1-R_i}{2}\right]$.

Following the same steps as in Chapter III Section 2 we define for the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ Problem of length n,

$$U_{i,n}(\underline{Y}_i) = P(X_i=L_n|\underline{Y}_i)$$

$$= I(X_i=L_i)P(X_\ell<X_i, \quad i<\ell\leq n|\underline{Y}_i)$$

$$= I(X_i=L_i)P(X_\ell<\frac{R_i}{2} + M_i, \ i<\ell\leq n|R_i)$$

$$= I(X_i=L_i) \ \frac{{1-R_i}^{n-i+1}}{(n-i+1)(1-R_i)}$$

$$= I(X_i=L_i) \quad T_{i,n}(R_i) \ .$$

Note that $T_{i,n}$ is an increasing function.

Using backward induction we define $V_{i,n}$ recursively as

$$V_{n,n} = U_{n,n}$$

$$V_{i,n}(\underline{Y}_i) = E(\max\{U_{i+1,n}(\underline{Y}_{i+1}),V_{i+1,n}(\underline{Y}_{i+1})\}|\underline{Y}_i) \ .$$

A best invariant rule is then

$$\tau_n = \min\{n,\min_{j\geq 2}\{j:U_{j,n}(\underline{Y}_j) > V_{j,n}(\underline{Y}_j)\}\}. \tag{4.3}$$

As defined $V_{j,n}(\underline{Y}_j) = \sup_{\sigma\in F_{j,n}} E_0(U_\sigma(\underline{Y}_\sigma)|\underline{Y}_j)$.

We now show that $V_{j,n}(\underline{Y}_j)$ is a function only of R_j.

<u>Lemma 4.1</u>. $V_{j,n}(\underline{Y}_j) = V_{j,n}(R_j)$ for each $j=2,\ldots,n$.

<u>Proof</u>. The relation of R_{j+1} and R_j is given by

$$R_{j+1} = R_j I(M_j-\frac{R_j}{2}\leq X_{j+1}\leq M_j+\frac{R_j}{2})+(M_j+\frac{R_j}{2}-X_{j+1})I(X_{j+1}<M_j-\frac{R_j}{2})$$

$$+ \ (X_{j+1}-M_j+\frac{R_j}{2})\,I(X_j>M_j+\frac{R_j}{2}) \ .$$

If j=n-1 the result is clear since

$$V_{n-1,n}(\underline{Y}_{n-1}) = P(X_n>L_{n-1}|\underline{Y}_{n-1})$$

$$= P(X_n>M_{n-1}+\frac{R_{n-1}}{2}|R_{n-1})$$

$$= \frac{1-R_{n-1}}{2} = V_{n-1,n}(R_{n-1}) \ .$$

Suppose $V_{j,n}(\underline{Y}_j) = V_{j,n}(R_j)$ for $j=n-1,n-2,\ldots,n-m+1$, $2\le m\le n-2$.

Then

(4.4) $V_{n-m,n}(\underline{Y}_{n-m})$

$$= E(\max\{U_{n-m+1,n}(\underline{Y}_{n-m+1}),V_{n-m+1,n}(R_{n-m+1})\}|\underline{Y}_{n-m})$$

$$= E(V_{n-m+1,n}(R_{n-m})\,I(M_{n-m}-\frac{R_{n-m}}{2}\le X_{n-m+1}\le M_{n-m}+\frac{R_{n-m}}{2})$$

$$+ \ V_{n-m+1,n}(M_{n-m}+\frac{R_{n-m}}{2}-X_{n-m+1})\,I(X_{n-m+1}<M_{n-m}-\frac{R_{n-m}}{2})$$

$$+ \ \max\{T_{n-m+1,n}(X_{n-m+1}-M_{n-m}+\frac{R_{n-m}}{2}),V_{n-m+1,n}(X_{n-m+1}-M_{n-m}+\frac{R_{n-m}}{2})\}$$

$$I(X_{n-m+1}>M_{n-m}+\frac{R_{n-m}}{2})\,|R_{n-m})$$

$$= V_{n-m,n}(R_{n-m}) \ . \qquad //$$

We can give a more explicit expression for $V_{n-m,n}(R_{n-m})$. From (4.4) and (4.1) we have

(4.5)

$$V_{n-m,n}(R_{n-m}) = R_{n-m}V_{n-m+1,n}(R_{n-m})$$

$$+ \frac{1}{1-R_{n-m}} \int_{-\frac{1-R_{n-m}}{2}}^{\frac{1-R_{n-m}}{2}} \Bigg[\int_{-\frac{1}{2}}^{M_{n-m}-\frac{R_{n-m}}{2}} V_{n-m+1,n}(M_{n-m}+\frac{R_{n-m}}{2}-x)\,dx$$

$$+ \int_{M_{n-m}+\frac{R_{n-m}}{2}}^{\frac{1}{2}} \max\{V_{n-m+1,n}(x-M_{n-m}+\frac{R_{n-m}}{2}), T_{n-m+1,n}(x-M_{n-m}+\frac{R_{n-m}}{2})\}dx \Bigg] dM_{n-m}$$

$$= R_{n-m}V_{n-m+1,n}(R_{n-m}) + \frac{1}{1-R_{n-m}} \int_{R_{n-m}}^{1} (1-u)\,[V_{n-m+1,n}(u)$$

$$+ \max\{V_{n-m+1,n}(u), T_{n-m+1,n}(u)\}]du.$$

We now derive the form of a best invariant rule. Let

$$h_{j,n}(x) = (1-x)\; V_{j,n}(x)$$

$$g_{j,n}(x) = (1-x)\; T_{j,n}(x) = \frac{1-x^{n-j+1}}{n-j+1}$$

for $j=2,\ldots,n$, $x\in[0,1]$.

Lemma 4.2. i.) $\lambda_{n-1,n} = 0$ is the unique solution of $h_{n-1,n}(x) = g_{n-1,n}(x)$. If $2\leq j<n-1$ there is a unique solution $\lambda_{j,n}\in(0,1)$ of the equation $h_{j,n}(x) = g_{j,n}(x)$.

For each $2\leq j<n$:

ii.) $1 > \lambda_{j,n} > \lambda_{j+1,n} > \ldots > \lambda_{n-1,n} = 0.$

iii.) $h_{j,n}, g_{j,n}$ are decreasing functions on [0,1].

iv.) $\frac{h'_{j,n}(x)}{g'_{j,n}(x)} > 1$ for $0 \leq x \leq \lambda_{j,n}$.

v.) $V_{j,n}(x)$ is decreasing for $x \geq \lambda_{j+1,n}$.

vi.) $h_{j,n}(\lambda_{j+1,n}) > g_{j,n}(\lambda_{j+1,n})$.

Proof. The proof is by induction. To simplify notation we consider n fixed and omit the second subscript on λ, h, g, T and V.

We begin by showing i.) - vi.) hold for j=n-1 and n-2. It is easily verified that

$$V_{n-1}(x) = \frac{1-x}{2} \qquad V_{n-2}(x) = \frac{1-x^2}{2}$$

$$T_{n-1}(x) = \frac{1+x}{2} \qquad T_{n-2}(x) = \frac{1+x+x^2}{3}$$

Thus i.) and ii.) hold for j=n-1,n-2 and $\lambda_{n-1}=0$, $\lambda_{n-2} \doteq .29$. iii.) and v.) are immediate for j=n-1,n-2 as is vi.) for j=n-2. By the above

$$\frac{h'_{n-1}(x)}{g'_{n-1}(x)} = \frac{1}{2x} > 1 \qquad \text{for } x = \lambda_{n-1} = 0.$$

$$\frac{h'_{n-2}(x)}{g'_{n-2}(x)} = \frac{1}{2x^2} + \frac{1}{x} - \frac{3}{2} > 1 \qquad \text{for } 0 \leq x \leq \lambda_{n-2},$$

so that iv.) holds for j=n-1,n-2.

Next we show that if i.) is satisfied for index j+1 then vi.) is satisfied for index j. Let

$$f_{j+1}(u) = \max\{h_{j+1}(u), g_{j+1}(u)\}, \quad 1 \leq j < n.$$

(4.5) implies

$$(4.6) \qquad h_j(x) = x\, h_{j+1}(x) + \int_x^1 (h_{j+1}(u)+f_{j+1}(u))\,du.$$

Suppose i.) is satisfied for index j. We claim:

$$(4.7) \qquad h_j(x) > g_j(x) \qquad \text{if } x < \lambda_j ,$$

$$h_j(x) < g_j(x) \qquad \text{if } 1 > x > \lambda_j .$$

To see this first note that (4.7) is equivalent to

$$(4.8) \qquad V_j(x) > T_j(x) \qquad \text{if } x < \lambda_j ,$$

$$V_j(x) < T_j(x) \qquad \text{if } 1 > x > \lambda_j .$$

It is clear that $V_{n-1}(1) = 0$. Suppose that $V_j(1) = 0$ for $j=n-1,n-2,\ldots,n-k+1$. Then

$$V_{n-k}(1) = V_{n-k+1}(1) - \left[\frac{d}{dx}\int_x^1 (h_{n-k+1}(u)+f_{n-k+1}(u))\,du\right]_{x=1} = 0.$$

Thus $V_j(1) = 0$ for all $2 \le j < n$. (Note: the probabilistic meaning of V_j makes this intuitively clear).

From the expressions for T,

$$T_j(0) = \frac{1}{n-j+1}, \qquad T_j(1) = 1, \qquad 2 \le j < n.$$

$$V_j(0) = \int_0^1 (h_{j+1}(u)+f_{j+1}(u))\,du$$

$$\ge \int_0^1 g_{j+1}(u)\,du = \frac{1}{n-j+1} = T_j(0), \quad 2 \le j < n.$$

Thus $V_{j+1}(0) \ge \frac{1}{n-j}$, and since V_{j+1} is continuous and non-negative on [0,1], so is h_{j+1}. Thus

$$V_j(0) \ge \int_0^1 h_{j+1}(u)\,du + \frac{1}{n-j+1} > T_j(0) .$$

T_j is also continuous on [0,1] for $2\leq j<n$. Therefore if λ_j is the unique root of $V_j(x) = T_j(x)$ in (0,1) (and therefore, by the above, the unique root in [0,1]), (4.8) is satisfied, implying (4.7) is also satisfied.

Now if i.) (and therefore (4.7)) holds for index j+1 then

$$h_j(\lambda_{j+1}) = \lambda_{j+1}h_{j+1}(\lambda_{j+1}) + \int_{\lambda_{j+1}}^{1} (h_{j+1}(u)+g_{j+1}(u))du$$

$$= \frac{\lambda_{j+1}(1-\lambda_{j+1})^{n-j}}{n-j} + \int_{\lambda_{j+1}}^{1} (h_{j+1}(u)+g_{j+1}(u))du$$

$$> \frac{\lambda_{j+1}(1-\lambda_{j+1})^{n-j}}{n-j} + \int_{\lambda_{j+1}}^{1} \frac{1-u^{n-j}}{n-j}\,du$$

$$= \frac{\lambda_{j+1}(1-\lambda_{j+1})^{n-j}}{n-j} + \frac{1-\lambda_{j+1}}{n-j} - \frac{1-\lambda_{j+1}^{n-j+1}}{(n-j)(n-j+1)}$$

$$= \frac{1-\lambda_{j+1}^{n-j+1}}{n-j+1} = g_j(\lambda_{j+1}).$$

Thus vi.) holds for index j.

Let $3\leq m<n-1$ and assume i.)-v.) hold for $j=n-1,n-2,\ldots,n-m+1$. We will show that they hold for $j=n-m$. By the above this will show vi.) is valid for $2\leq j<n$ and the lemma will be proved.

Consider iii.) It is clear that g_j is a decreasing function on [0,1] for $2\leq j<n$. By (4.6) and the induction hypothesis

$$h'_{n-m}(x) = x\, h'_{n-m+1}(x) - f_{n-m+1}(x) < 0,$$

and therefore iii.) holds for j=n-m.

To prove v.) for j=n-m it suffices to show $V_{n-m}(x)$ is decreasing for $x \geq \lambda_{n-m+1}$. Recall that

$$V_{n-m}(x) = \sup_{\sigma \in \mathcal{F}_{n-m,n}} P_0(X_\sigma = L_n | R_{n-m} = x)$$
$$= P_0(X_{\sigma'} = L_n | R_{n-m} = x)$$

where the induction hypothesis implies that

$$\sigma' = \min\{n, \min_{k>n-m} \{k: X_k = L_k, R_k > \lambda_k\}\}.$$

If $x \geq \lambda_{n-m+1} > \lambda_{n-m+2} > \ldots > \lambda_{n-1} = 0$ then

$$P_0(X_{\sigma''} = L_n | R_{n-m} = x) = P_0(X_{\sigma'} = L_n | R_{n-m} = x)$$

where

$$\sigma'' = \min\{n, \min_{k>n-m} \{k: X_k = L_k\}\}.$$

Now for $y > x - \frac{1}{2}$

$$P_0(X_{\sigma''} = L_n | L_{n-m} = y,\ R_{n-m} = x)$$

$$= \sum_{k=1}^{m} \frac{1}{k} P \text{ (exactly } k \text{ of } \{X_i\}_{i=n-m+1}^{n} \text{ are larger than } y).$$

The latter is clearly decreasing in y. Given $R_{n-m} = x$, $L_{n-m} \sim U[x - \frac{1}{2}, \frac{1}{2}]$, so that

$$V_{n-m}(x) = \frac{1}{1-x} \int_{x-\frac{1}{2}}^{\frac{1}{2}} P_0(X_{\sigma''} = L_n | L_{n-m} = y,\ R_{n-m} = x)\, dy$$

which is decreasing in x. Thus v.) holds for j=n-m.

We now show that iv.) holds for $x \leq \lambda_{n-m+1}$ and j=n-m, i.e.

$$\frac{h'_{n-m}(x)}{g'_{n-m}(x)} > 1 \qquad \text{for } 0 \leq x \leq \lambda_{n-m+1}.$$

Note that $g_j'(x) = -x^{n-j}$. Thus if $x \leq \lambda_{n-m+1}$ the induction hypothesis for i.), iii.) and iv.) gives

$$\frac{h'_{n-m}(x)}{g'_{n-m}(x)} = \frac{h'_{n-m+1}(x)}{g'_{n-m+1}(x)} - \frac{h_{n-m+1}(x)}{g'_{n-m}(x)} > 1.$$

This means that h_{n-m} is decreasing more rapidly than g_{n-m} on $[0,\lambda_{n-m+1}]$. This, along with v.) and the fact that T_{n-m} is increasing on [0,1], gives us i.) for j=n-m. As noted earlier vi.) follows for j=n-m-1.

Since $h_{n-m}(\lambda_{n-m+1}) > g_{n-m}(\lambda_{n-m+1})$, we see that the equation $h_{n-m}(x) = g_{n-m}(x)$ has no roots in $[0,\lambda_{n-m+1}]$. Thus there is exactly one root in $(\lambda_{n-m+1},1)$. This proves ii.) for j=n-m.

For $x \in (\lambda_{n-m+1},\lambda_{n-m}]$ V_{n-m} is decreasing, T_{n-m} is increasing and $T_{n-m}(x) \leq V_{n-m}(x)$ by what has been shown so far. Therefore

$$0 > h'_{n-m}(x) = -V_{n-m}(x) + (1-x)\ V'_{n-m}(x)$$

$$0 > g'_{n-m}(x) = -T_{n-m}(x) + (1-x)\ T'_{n-m}(x)$$

which implies

$$\frac{h'_{n-m}(x)}{g'_{n-m}(x)} > 1 \qquad \text{for } x \leq \lambda_{n-m}.$$

Thus iv.) holds for j=n-m and the induction is complete. //

Lemma 4.2 leads us immediately to

<u>Theorem 4.1.</u> A best invariant rule for the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ Problem of length $n\geq 3$ is $\tau_n=\min\{n,\min_{k\geq 2}\{k:X_k=L_k,R_k>\lambda_{k,n}\}\}$, where $1>\lambda_{2,n}>\lambda_{3,n}>\ldots>\lambda_{n-1,n} = 0$.

Proof. Immediate from (4.3) and Lemma 4.2. //

Remark

For n=1,2 the best invariant rule is clearly the N.I. Rule. Since $\lambda_{n-1,n} = 0$ this is also the case for n=3. The rule of Theorem 4.1 differs from the N.I. Rule when $n \geq 4$.

2. $P(X_\tau = L_n)$ for Stopping Rules τ in a Certain Class

It is our purpose in this section to derive a formula giving $P_0(X_\tau = L_n)$ for a certain class of invariant stopping rules τ in the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ Problem of length n. The class considered is

$$\mathcal{T}_n = \{\tau : \tau = \min\{n, \min_{k\geq 2}\{k : X_k = L_k, R_k > \beta_{k,n}\}\},$$
$$1 \geq \beta_{2,n} \geq \beta_{3,n} \geq \cdots \geq \beta_{n-1,n} = 0\}.$$

By Theorem 4.1 we know that $\mathcal{T}_n$ contains a best invariant rule. For notational simplicity in what follows we assume n is fixed and omit the second subscript on the decision numbers.

Let $1 \geq u_2 \geq u_3 \geq \cdots \geq u_{n-1} = 0$ be any set of decision numbers and let $\tau \in \mathcal{T}_n$ be the stopping rule defined by these numbers. Assume $\theta=0$ and define the sets

$$A_2 = \{R_2 > u_2\}$$

$$A_j = \{R_{j-1} \leq u_{j-1}, R_j > u_j\}, \quad 3 \leq j \leq n-1.$$

Clearly $A_i \cap A_j = \phi$, $i \neq j$ and $P(\bigcup_{j=2}^{n-1} A_j) = 1$. Define $A_n = \phi$. On A_j, $\tau \geq j$. Let $3 \leq j \leq r \leq n-1$.

(4.9)

$P(\{\tau>r\} \cap A_j)$

$$=P(X_\nu \le L_{j-1}, \nu=j,\ldots,r, R_{j-1} \le u_{j-1}, R_j > u_j)$$

$$=P(X_\nu \le M_{j-1}+\frac{R_{j-1}}{2}, \nu=j,\ldots,r, R_{j-1} \le u_j, M_{j-1}+\frac{R_{j-1}}{2}-X_j>u_j, X_j<M_{j-1}-\frac{R_{j-1}}{2})$$

$$+P(X_\nu \le M_{j-1}+\frac{R_{j-1}}{2}, \nu=j,\ldots,r, u_j<R_{j-1} \le u_{j-1}).$$

Since $M_{j-1} - \frac{R_{j-1}}{2} < M_{j-1} + \frac{R_{j-1}}{2} - u_j$ if and only if $R_{j-1} > u_j$, the right hand side of (4.9) can be written as

$$P(X_\nu \le M_{j-1}+\frac{R_{j-1}}{2}, \nu=j+1,\ldots,r, R_{j-1} \le u_j, X_j<M_{j-1}+\frac{R_{j-1}}{2}-u_j)$$

$$+P(X_\nu \le M_{j-1}+\frac{R_{j-1}}{2}, \nu=j,\ldots,r, u_j<R_{j-1} \le u_{j-1})$$

= (using (4.1))

$$(j-1)(j-2)\int_0^{u_j} R_{j-1}^{j-3} \int_{u_j-\frac{1+R_{j-1}}{2}}^{\frac{1-R_{j-1}}{2}} (M_{j-1}+\frac{R_{j-1}}{2}+\frac{1}{2})^{r-j} (M_{j-1}+\frac{R_{j-1}}{2}+\frac{1}{2}-u_j)\, dM_{j-1}\, dR_{j-1}$$

$$+(j-1)(j-2)\int_{u_j}^{u_{j-1}} R_{j-1}^{j-3} \int_{-\frac{1-R_{j-1}}{2}}^{\frac{1-R_{j-1}}{2}} (M_{j-1}+\frac{R_{j-1}}{2}+\frac{1}{2})^{r-j+1} dM_{j-1}\, dR_{j-1}$$

$$=(j-1)\left[\frac{u_j^{j-2}-u_j^r}{r-j+2} - \frac{u_j^{j-1}-u_j^r}{r-j+1}\right]+(j-1)\frac{u_{j-1}^{j-2}-u_j^{j-2}}{r-j+2} - (j-1)(j-2)\frac{u_{j-1}^r-u_j^r}{r(r-j+2)}.$$

Let $p_{j,r}(u_{j-1},u_j)$ denote this last quantity.

Let Q be the complement of the set $\{\tau=n, X_n \ne L_n\}$. For r,j as above $\{\tau>r\} \cap Q \cap A_j = \{\tau=k,\ X_k=L_k,\ \text{some } r<k\le n\} \cap A_j$. Given $\{\tau>r\} \cap Q \cap A_j$, the probability that

$X_{r+1} = \max\{X_{r+1},\ldots,X_n\}$ is $\frac{1}{n-r}$. By (4.9) and the computations following it we see that

$$P(\{\tau>r\}\cap Q\cap A_j) = p_{j,r}(u_{j-1},u_j) - p_{j,n}(u_{j-1},u_j).$$

Combining these we obtain

$$P(\{X_\tau=X_{r+1}=L_n\}\cap A_j) = \frac{1}{n-r}(p_{j,r}(u_{j-1},u_j)-p_{j,n}(u_{j-1},u_j)).$$

Since $\{X_\tau=X_{r+1}=L_n\}\subset \bigcup_{j=2}^{r+1} A_j$ it remains only to consider $\{X_\tau=X_{r+1}=L_n\}\cap A_j$, $j=2,r+1$.

We first look at

$$(4.10)\quad P(\{X_\tau=X_{r+1}=L_n\}\cap A_{r+1}) =$$

$$P(R_r\le u_{r+1}<R_{r+1}, X_j\le X_{r+1}=L_{r+1},\ j=r+2,\ldots,n)$$

$$+\ P(u_{r+1}<R_r\le u_r, X_j\le X_{r+1}=L_{r+1},\ j=r+2,\ldots,n)\ .$$

Recalling that, given R_r, L_r is distributed uniformly on $[R_r-\frac{1}{2},\frac{1}{2}]$, and letting $a\ge 0$ we have

$$P(X_{r+1}\le b, X_{r+1}-L_r=a\,|\,R_r)$$

$$= E(I_{[R_r-\frac{1}{2},b-a]}(L_r)\,|\,R_r)$$

$$= \frac{b-a-R_r+\frac{1}{2}}{1-R_r} I_{[R_r-\frac{1}{2}+a,\frac{1}{2}]}(b)\ I_{[0,1-R_r]}(a)\ .$$

Thus the density (note that the total probability does not equal 1)

$$P(X_{r+1}=b, X_{r+1}-L_r=a\,|\,R_r)=\frac{1}{1-R_r} I_{[R_r-\frac{1}{2}+a,\frac{1}{2}]}(b)\ I_{[0,1-R_r]}(a)\ .$$

Hence

$$(4.11)\quad P(X_j \le X_{r+1}, j=r+2,\ldots,n, X_{r+1}-L_r=a \mid R_r)$$

$$= \left[\int_{R_r-\frac{1}{2}+a}^{\frac{1}{2}} \frac{(b+\frac{1}{2})^{n-r-1}}{1-R_r}\, db\right] I_{[0,1-R_r]}(a)$$

$$= \frac{1-(R_r+a)^{n-r}}{(n-r)(1-R_r)}\; I_{[0,1-R_r]}(a) \;.$$

If $R_r \le u_{r+1}$ we can write

$$P(X_j \le X_{r+1}, j=r+2,\ldots,n, X_{r+1}-L_r > u_{r+1}-R_r \mid R_r)$$

$$= \int_{u_{r+1}-R_r}^{1-R_r} \frac{1-(R_r+a)^{n-r}}{(n-r)(1-R_r)}\, da$$

$$= \frac{(1-u_{r+1}) - \dfrac{1-u_{r+1}^{n-r+1}}{n-r+1}}{(n-r)(1-R_r)} \;.$$

Thus the first summand on the right hand side of (4.10),

$$P(R_r \le u_{r+1} < R_{r+1}, X_j \le X_{r+1} = L_{r+1},\quad j=r+2,\ldots,n)$$

$$= r(r-1)\left(\int_0^{u_{r+1}} R_r^{\,r-2}\, dR_r\right)\left(\frac{(1-u_{r+1}) - \dfrac{1-u_{r+1}^{n-r+1}}{n-r+1}}{n-r}\right)$$

$$= \frac{r\, u_{r+1}^{r-1}}{n-r}\left(\frac{n-r}{n-r+1} - u_{r+1} + \frac{u_{r+1}^{n-r+1}}{n-r+1}\right) \;.$$

Let $q_{r+1,1}(u_r, u_{r+1})$ denote this quantity.

To obtain an expression for the second summand on the right hand side of (4.10) we use (4.11) and write

$$P(X_j \le X_{r+1} = L_{r+1},\ j=r+2,\ldots,n \mid R_r)$$

$$= \frac{1}{(n-r)(1-R_r)} \int_0^{1-R_r} [1-(R_r+a)^{n-r}]\,da$$

$$= \frac{1}{(n-r)(1-R_r)} \left[(1-R_r) - \frac{1-R_r^{n-r+1}}{n-r+1}\right] .$$

Integrating with respect to the density of R_r on (u_{r+1}, u_r) we obtain the desired result as

$$\frac{r(r-1)}{n-r} \int_{u_{r+1}}^{u_r} \left[(1-R_r)R_r^{r-2} - \frac{(1-R_r^{n-r+1})R_r^{r-2}}{n-r+1}\right] dR_r$$

$$\frac{r(u_r^{r-1}-u_{r+1}^{r+1})}{n-r+1} - \frac{(r-1)(u_r^r - u_{r+1}^r)}{n-r} + \frac{r(r-1)(u_r^n-u_{r+1}^n)}{n(n-r)(n-r+1)} .$$

Let $q_{r+1,2}(u_r, u_{r+1})$ denote this quantity. Let

$$q_{r+1}(u_r, u_{r+1}) = P(\{X_\tau = X_{r+1} = L_n\} \cap A_{r+1})$$

$$= q_{r+1,1}(u_r, u_{r+1}) + q_{r+1,2}(u_r, u_{r+1}) .$$

Note that $q_n \equiv 0$.

To obtain $P(\{X_\tau = X_{r+1} = L_n\} \cap A_2)$ we let

$$p_{2,r}(u_2) = P(\{\tau > r\} \cap Q \cap A_2)$$

$$= P(\{\tau > r\} \cap A_2) - P(\{\tau = n, X_n \ne L_n\} \cap A_2) .$$

Now

$$P(\{\tau>r\}\cap A_2) = \frac{1}{2}P(X_\nu \le M_2+\frac{R_2}{2}, \nu=3,\ldots,r, R_2>u_2)$$

$$= \int_{u_2}^{1} \int_{-\frac{1-R_2}{2}}^{\frac{1-R_2}{2}} (M_2+\frac{R_2}{2}+\frac{1}{2})^{r-2}\, dM_2 dR_2$$

$$= \frac{1}{r-1}(1-u_2-\frac{1-u_2^r}{r}) = \frac{1}{r} - \frac{u_2}{r-1} - \frac{u_2^r}{r(r-1)} .$$

$$P(\{\tau=n, X_n \ne L_n\}\cap A_2) = \frac{1}{2}P(X_\nu \le M_2+\frac{R_2}{2}, \nu=3,\ldots,n, R_2>u_2)$$

$$= \frac{1}{n} - \frac{u_2}{n-1} + \frac{u_2^n}{n(n-1)} .$$

So $\quad p_{2,r}(u_2) = (\frac{1}{r}-\frac{1}{n}) - u_2(\frac{1}{r-1}-\frac{1}{n-1}) + \frac{u_2^r}{r(r-1)} - \frac{u_2^n}{n(n-1)}$.

Let $\quad p(k) = P(X_\tau = X_k = L_n), \quad 2 \le k \le n.$

$$p(2) = \frac{1}{2}P(X_\nu \le M_2+\frac{R_2}{2}, \; \nu=3,\ldots,n, R_2>u_2) = \frac{1}{n} - \frac{u_2}{n-1} - \frac{u_2^n}{n(n-1)} .$$

For $k \ge 3$, $\quad p(k) = \sum_{j=2}^{k} P(\{X_\tau = X_k = L_n\}\cap A_j)$.

Thus we have

<u>Theorem 4.2</u>. In the $U[\theta-\frac{1}{2}, \theta+\frac{1}{2}]$ Problem let $\tau \in \mathcal{T}_n$ be an invariant stopping rule determined by the decision numbers $1 \ge u_2 \ge u_3 \ge \ldots \ge u_{n-1} = 0$. Then

$$P(X_\tau = L_n) = \sum_{k=2}^{n} p(k) \quad \text{where}$$

(regarding vacuous sums as 0)

$$p(2) = \frac{1}{n} - \frac{u_2}{n-1} + \frac{u_2^n}{n(n-1)}$$

$$p(r+1) = \frac{p_{2,r}(u_2)}{n-r} + q_{r+1}(u_r,u_{r+1})$$

$$+ \frac{1}{n-r} \sum_{j=3}^{r} (p_{j,r}(u_{j-1},u_j)-p_{j,n}(u_{j-1},u_j)),$$

$$n>r\geq 2.$$

3. Computing Decision Numbers for a Best Invariant Rule

Recall the best invariant rule for the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ Problem of length n given by Theorem 4.1:

$$\tau_n = \min\{n,\min_{k\geq 2}\{k:X_k=L_k,R_k>\lambda_{k,n}\}\}.$$

In this section we will obtain each of the n-2 decision numbers $\{\lambda_{i,n}\}_{i=2}^{n-1}$ as a root of a polynomial. We are able to compute the λ's with the aid of a high speed computer and have done so for n=1,...,50. The results are displayed in Table 4.1. Using these numbers in the formula of Theorem 4.2 we then obtain $P(X_{\tau_n}=L_n)$. These values are displayed for selected n in Table 4.2.

Let $\binom{k}{\ell,m,n}$ denote the trinomial coefficient $\frac{k!}{\ell!m!n!}$.

Theorem 4.3. Let $\{\lambda_{i,n}\}_{i=2}^{n-1}$ be the decision numbers for the best invariant rule of Theorem 4.1. $\lambda_{n-k,n}$ is a root of the equation

$$\sum_{r=0}^{k+1} a_r\lambda^r = 0, \tag{4.12}$$

where

$$a_{k+1} = \frac{1}{k+1} \left(\sum_{i=1}^{k} \frac{(-1)^{i+1}}{i} \binom{k}{i} + 1\right) \tag{4.13}$$

(4.14) $$a_0 = \sum_{i=1}^{k} \frac{1}{i}\binom{k}{i} \sum_{j=0}^{i} (-1)^j \binom{i}{j} \frac{1}{k+j-i+1} - \frac{1}{k+1}$$

(4.15) $$a_r = \sum_{i=k-r+1}^{k} \frac{(-1)^{r-k+i}}{(r)(i)} \binom{k}{k-i,k-r+1,r-k+i-1}, \quad 0<r<k+1.$$

Proof. Assume $\theta = 0$. If we observe $X_{n-k} = L_{n-k}$ we will be indifferent to choosing X_{n-k} or rejecting it if

(4.16) $$P(X_{n-k}=L_n \mid R_{n-k}=\lambda) = \sup_{\tau \in F_{n-k,n}} P(X_\tau=L_n \mid R_{n-k}=\lambda).$$

The value of λ satisfying (4.16) will be $\lambda_{n-k,n}$.

We assume n is fixed and drop the second subscript on the λ's. If $X_{n-k}=L_{n-k}$ and $R_{n-k}=\lambda_{n-k}$ then by rejecting X_{n-k} our rule can only choose the largest observation if it is the first observation following X_{n-k} which exceeds X_{n-k}. Let E_i be the event that exactly i of $X_{n-k+1},\ldots,X_n$ exceed $L_{n-k}, i=1,\ldots,k$.

$$P(E_i \mid R_{n-k} = \lambda_{n-k})$$

$$= \binom{k}{i} P(X_\ell > L_{n-k}, \ell=n-k+1,\ldots,n-k+i, X_\ell \le L_{n-k}, \ell=n-k+i+1,\ldots,n \mid R_{n-k} = \lambda_{n-k})$$

$$= \binom{k}{i} \frac{1}{1-\lambda_{n-k}} \int_{-\frac{1-\lambda_{n-k}}{2}}^{\frac{1-\lambda_{n-k}}{2}} \left(M_{n-k}+\frac{\lambda_{n-k}}{2}+\frac{1}{2}\right)^{k-i} \left(\frac{1}{2} - \frac{\lambda_{n-k}}{2}-M_{n-k}\right)^{i} dM_{n-k}$$

$$= \binom{k}{i} \frac{1}{1-\lambda_{n-k}} \int_{\lambda_{n-k}}^{1} u^{k-i}(1-u)^i du.$$

Given $\{E_i, R_{n-k}=\lambda_{n-k}\}$, the probability is $\frac{1}{i}$ that the first of these i observations larger than L_{n-k} is also the largest of all the observations. Thus

P (exactly i of $\{X_\ell\}_{\ell=n-k+1}^{n}$ are larger than L_{n-k} and the first of these i is equal to $L_n | R_{n-k} = \lambda_{n-k}$)

$$= \frac{1}{i} \binom{k}{i} \frac{1}{1-\lambda_{n-k}} \int_{\lambda_{n-k}}^{1} u^{k-i} (1-u)^{i} du.$$

If, on the other hand, we elect to choose $X_{n-k}=L_{n-k}$, the probability it is equal to L_n is

$$P(L_{n-k}=L_n | R_{n-k}=\lambda_{n-k})$$

$$= P(X_\ell \leq L_{n-k}, \ell=n-k+1,\ldots,n | R_{n-k}=\lambda_{n-k})$$

$$= \frac{1}{1-\lambda_{n-k}} \int_{-\frac{1-\lambda_{n-k}}{2}}^{\frac{1-\lambda_{n-k}}{2}} (M_{n-k} + \frac{\lambda_{n-k}}{2} + \frac{1}{2})^{k} dM_{n-k}$$

$$= \frac{1-\lambda_{n-k}^{k+1}}{(k+1)(1-\lambda_{n-k})}.$$

So (4.16) becomes

$$\sum_{i=1}^{k} \frac{1}{i} \binom{k}{i} \int_{\lambda_{n-k}}^{1} u^{k-i} (1-u)^{i} du = \int_{\lambda_{n-k}}^{1} u^{k} du$$

or

$$\sum_{i=1}^{k} \frac{1}{i} \binom{k}{i} \sum_{j=0}^{i} (-1)^{j} \binom{i}{j} \frac{1-\lambda_{n-k}^{k+j-i+1}}{k+j-i+1} = \frac{1-\lambda_{n-k}^{k+1}}{k+1} .$$

Simplifying, we obtain the form of this equation given by (4.12) - (4.15). //

An interesting sidelight of this theorem is that the decision numbers $\lambda_{n-k,n}$ depend only upon the number of observations remaining and not on n, in contrast to the numbers for the $N(\mu,\sigma^2)$ Problem. This is so since R_{n-k} contains the same amount of information about θ no matter how many observations it took to obtain this information.

We used the results of Theorem 4.3 and the CDC 6500 computer at Purdue to obtain the decision numbers $\{\lambda_{j,n}\}_{j=2}^{n-1}$ for n=3,...,50. To find the roots of the defining polynomials (4.12) the double precision library routine RPOLY was used. The values obtained are listed in Table 4.1 according to the number of observations remaining (since by the above for $n_1,n_2>k$, $\lambda_{n_1-k,n_1} = \lambda_{n_2-k,n_2}$).

Using the formulas of Theorem 4.2 and the λ values from Table 4.1 we were then able to obtain $P(X_{\tau_n}=L_n)$ for n=1,...,50. These values as well as the corresponding values for the N.I. and F.I. Rules are given in Table 4.2 for representative n.

4. Asymptotics

In this section we will show that for the stopping rule τ_n of Theorem 4.1

$$\frac{1}{e} < \varliminf_{n\to\infty} P_\theta(X_{\tau_n}=L_n) \leq \varlimsup_{n\to\infty} P_\theta(X_{\tau_n}=L_n) < \alpha_o$$

for any $\theta\in\mathbb{R}$. We begin with a lemma.

Table 4.1. Decision Numbers for the Best Invariant Rule of Theorem 4.1

n-k	λ_k	n-k	λ_k	n-k	λ_k
1	0	17	.88147	33	.93743
2	.28990	18	.88774	34	.93922
3	.46275	19	.89338	35	.94091
4	.56971	20	.89848	36	.94252
5	.64163	21	.90312	37	.94403
6	.69312	22	.90735	38	.94547
7	.73175	23	.91123	39	.94684
8	.76177	24	.91480	40	.94814
9	.78577	25	.91809	41	.94938
10	.80539	26	.92114	42	.95056
11	.82173	27	.92396	43	.95169
12	.83553	28	.92660	44	.95276
13	.84736	29	.92905	45	.95379
14	.85760	30	.93135	46	.95478
15	.86656	31	.93350	47	.95572
16	.87445	32	.93552	48	.95663

Table 4.2. Probability of Choosing L_n for the N.I. Rule, the F.I. Rule, and the Best Invariant Rule of Theorem 4.1.

n	N.I.	Best Invariant	F.I.
1	1.00000	1.00000	1.00000
2	.50000	.50000	.75000
3	.50000	.50000	.68429
4	.45833	.48305	.65540
5	.43333	.47311	.69392
10	.39869	.45393	.60870
15	.38940	.44762	.59898
20	.38420	.44449	.59420
30	.37865	.44137	.58947
40	.37574	.43981	.58713
50	.37427	.43888	.58573

Lemma 4.3. Let $\{\lambda_{k,n}\}_{k=2}^{n-1}$ be the decision numbers defining τ_n in Theorem 4.1. Then

$$1 - \frac{2\sqrt{3}}{n-k-1} \le \lambda_{k,n} \le 1 - \frac{\log 2}{n-k+1} \qquad k=2,\ldots,n-2 .$$

Proof. For simplicity of notation fix n and let $\lambda_k = \lambda_{k,n}$. Assume $\theta=0$ and recall that

$$P(X_k=L_n | X_k=L_k, R_k=\lambda) = \frac{1-\lambda^{n-k+1}}{(n-k+1)(1-\lambda)} .$$

As in the proof of Lemma 2.1, if

$$P(X_k=L_n | X_k=L_k, R_k=\lambda) > \frac{1}{2} \quad \text{then for any } \sigma\in\mathcal{F}_{k,n}$$

$$P(X_\sigma=L_n | X_k=L_k, R_k=\lambda) \le$$

$$P(X_j>L_k, \text{ some } j>k | X_k=L_k, R_k=\lambda) < \frac{1}{2} .$$

Suppose $\lambda = 1 - \frac{\log 2}{n-k+1}$. Then $(1-\frac{\log 2}{n-k+1})^{n-k+1} < \frac{1}{2}$ implies

$$\frac{1-\lambda^{n-k+1}}{(n-k+1)(1-\lambda)} = \frac{1-(1-\frac{\log 2}{n-k+1})^{n-k+1}}{\log 2} > \frac{1}{2} .$$

Since $\frac{1-\lambda^{n-k+1}}{(n-k+1)(1-\lambda)}$ is increasing in λ,

$$\lambda_k \le 1 - \frac{\log 2}{n-k+1} .$$

To show that $\lambda_k \ge 1 - \frac{2\sqrt{3}}{n-k-1}$ we recall that λ_k is a solution of

$$(4.17) \qquad \sum_{i=1}^{n-k} \frac{1}{i} \binom{n-k}{i} \int_\lambda^1 u^{n-k-i}(1-u)^i du = \int_\lambda^1 u^{n-k} du .$$

Integrating $(n-k)\int_\lambda^1 u^{n-k-1}(1-u)du$ by parts gives

$$(n-k) \int_{\lambda}^{1} u^{n-k-1}(1-u)du = -\lambda^{n-k}(1-\lambda)+\int_{\lambda}^{1} u^{n-k}du.$$

Therefore to satisfy (4.17) λ_k must satisfy

$$\lambda^{n-k}(1-\lambda) = \sum_{i=2}^{n-k} \frac{1}{i} \binom{n-k}{i} \int_{\lambda}^{1} u^{n-k-i}(1-u)^{i}du$$

$$> \sum_{i=2}^{n-k} \frac{1}{i(i+1)} \binom{n-k}{i} \lambda^{n-k-i}(1-\lambda)^{i+1} .$$

So we must have

$$(4.18)\quad \lambda_k^{n-k} > \sum_{i=2}^{n-k} \frac{1}{i(i+1)} \binom{n-k}{i} \lambda_k^{n-k-i}(1-\lambda_k)^{i} .$$

Suppose $\lambda_k = 1 - \frac{c}{n-k-1}$, $c > 2\sqrt{3}$.

Then $\sum_{i=2}^{n-k} \frac{1}{i(i+1)} \binom{n-k}{i} \lambda_k^{n-k-i} (1-\lambda_k)^{i}$

$$\geq \frac{(n-k)(n-k-1)}{12} \lambda_k^{n-k-2} (1-\lambda_k)^{2}$$

$$= \frac{(n-k)(n-k-1)}{12} \left(1-\frac{c}{n-k-1}\right)^{n-k-2} \left(\frac{c}{n-k-1}\right)^{2} > \left(1-\frac{c}{n-k-1}\right)^{n-k} = \lambda_k^{n-k}$$

since $\frac{(n-k)(n-k-1)}{12} > \left(\frac{n-k-1-c}{c}\right)^{2}$.

This contradicts (4.18) and so

$$\lambda_k \geq 1 - \frac{2\sqrt{3}}{n-k-1} . \qquad //$$

We now assume $\theta=\frac{1}{2}$ so that the X's are U[0,1] random variables. Let ω_n, σ_n be the optimal N.I. and F.I. Rules of length n respectively, for the observations $X_1, X_2, \ldots$ The next theorem shows that if τ_n is the best invariant rule of Theorem 4.1,

$$(4.19) \quad \frac{1}{e} = \lim_{n\to\infty} P(X_{\omega_n}=L_n) < \underline{\lim}_{n\to\infty} P(X_{\tau_n}=L_n) \le \overline{\lim}_{n\to\infty} P(X_{\tau_n}=L_n)$$

$$< \lim_{n\to\infty} P(X_{\sigma_n}=L_n) = \alpha_o \doteq .58 .$$

Our approach to the leftmost inequality of (4.19) is to take advantage of the fact that ω_n ignores more than $\frac{1}{3}$ of the observations. Among these observations significant improvements can be gained by using a simple invariant rule.

To obtain the rightmost inequality in (4.19) we note that

$$\tau_n = \min\{n, \min_{k\ge 2}\{k: X_k=L_k, R_k>\lambda_{k,n}\}\}$$

$$= \min\{n, \min_{k\ge 2}\{k: X_k=L_k>\lambda_{k,n}+D_k\}\}.$$

Looked at in this way τ_n appears to have the same form as σ_n but with random decision numbers $\lambda_{k,n} + D_k$. Our approach is based on the premise that the random quantities D_k introduce too much variation to allow $\underline{\tau} = (\tau_1, \tau_2, \ldots)$ to be asymptotically F.I.

Theorem 4.4. For τ_n the best invariant rule of Theorem 4.1 and any $\theta \in \mathbb{R}$

$$\frac{1}{e} < \underline{\lim}_{n\to\infty} P_\theta(X_{\tau_n}=L_n) \le \overline{\lim}_{n\to\infty} P_\theta(X_{\tau_n}=L_n) < \alpha_o .$$

Proof. Assume $\theta=\frac{1}{2}$. Let σ_n be the F.I. Rule of length n when a U[0,1] distribution is assumed. Let 0<a<1 and for $[\cdot]$ the largest integer function let $a_n=[an]$. Let $\gamma_n = \frac{2\sqrt{3}}{n-a_n-1}$. Let $\{\lambda_{k,n}\}_{k=2}^{n-1}$ be the decision numbers for τ_n.

Lemma 4.3 implies that $\lambda_{k,n} > 1 - \gamma_n$, $k=2,\ldots,a_n$.

Let $x_n = 1 - \frac{\varepsilon}{n}$ where $0 < \varepsilon < 1$ is a number to be determined.

Note $x_n > \lambda_{k,n}$ for all $k=3,\ldots,n-1$ and all $n \geq 3$.

Consider the stopping rule

$$\nu_n = \begin{cases} \min\limits_{2\leq k\leq a_n} \{k: X_k = L_k > x_n, D_k > \gamma_n\}, & \text{if such } k \text{ exists} \\ n, & \text{if no such } k \text{ exists.} \end{cases}$$

Let $\tau_n' = \min\{\nu_n, \tau_n\}$. Clearly

$$P(X_{\sigma_n} = L_n) \geq P(X_{\tau_n'} = L_n) \quad , \quad n \geq 2.$$

We will show that there is a $v > 0$ such that

(4.20) $\quad P(X_{\tau_n'} = L_n) \geq P(X_{\tau_n} = L_n) + v \qquad$ as $n \to \infty$.

Let

$$B_{k,n} = \{\tau_n' = \nu_n = k\} \qquad k=2,\ldots,a_n.$$

Then $\bigcup\limits_{k=2}^{a_n} B_{k,n} = \{\tau_n' = \tau_n\}^c$.

On $B_{k,n}$ $\quad X_{\tau_n'} = X_k = L_k > x_n$ and therefore

(4.21) $\quad P(X_{\tau_n'} = L_n | B_{k,n}) \geq x_n^n \qquad k=2,\ldots,a_n$.

Also, since $\tau_n > k$ on $B_{k,n}$,

(4.22) $\quad P(X_{\tau_n} = L_n | B_{k,n}) \leq P(X_j > x_n, \text{ some } k<j\leq n) \leq 1 - x_n^n$.

Fix $0<\delta<1$ and let $0<\varepsilon<1$ be such that

$$e^{-\frac{\varepsilon}{1-\varepsilon}} > \frac{1+\delta}{2} \qquad \left(\text{equivalently, } \varepsilon < \frac{\log\frac{2}{1+\delta}}{1+\log\frac{2}{1+\delta}}\right)$$

Then $x_n = 1 - \frac{\varepsilon}{n}$ means

$$x_n^n = \left(1 - \frac{\varepsilon}{n}\right)^n \geq e^{-\frac{\varepsilon}{1-\varepsilon}} > \frac{1+\delta}{2}$$

and $1 - x_n^n < \frac{1-\delta}{2}$.

So

$$(1 - x_n^n) + \delta < x_n^n$$

and therefore by (4.21) and (4.22)

$$P(X_{\tau_n'} = L_n | B_{k,n}) > P(X_{\tau_n} = L_n | B_{k,n}) + \delta \qquad k=2,\ldots,a_n .$$

Since $\{\tau_n' = \tau_n\} = (\bigcup_{k=2}^{a_n} B_{k,n})^c$ we have

$$P(X_{\tau_n'} = L_n | (\bigcup_{k=2}^{a_n} B_{k,n})^c) = P(X_{\tau_n} = L_n | (\bigcup_{k=2}^{a_n} B_{k,n})^c)$$

and therefore

$$P(X_{\tau_n'} = L_n) > P(X_{\tau_n} = L_n) + \delta\, P(\bigcup_{k=2}^{a_n} B_{k,n}) .$$

We complete the proof of (4.20) by showing that
$\underline{\lim}_{n\to\infty} P(\bigcup_{k=2}^{a_n} B_{k,n}) > 0.$

Consider the set

$$E_n = \{D_{a_n} > \gamma_n,\ X_k > x_n,\ \text{some } 2 \leq k \leq a_n,\ X_1 < x_n\}.$$

It is easily verified that $E_n \subset \bigcup_{k=2}^{a_n} B_{k,n}$.

$$P(E_n) = (x_n - \gamma_n)[(1-\gamma_n)^{a_n - 1} - (x_n - \gamma_n)^{a_n - 1}]$$

$$\to e^{-\frac{2\sqrt{3}\,a}{1-a}}(1 - e^{-a\varepsilon}) > 0 \qquad \text{as } n\to\infty$$

and therefore (4.20) follows.

Let ω_n be the N.I. Rule of length n. That is, $\omega_n = \min\{n, \min_{k\geq k(n)} \{k: X_k = L_k\}\}$ where $\lim_{n\to\infty} \frac{k(n)}{n} = \frac{1}{e}$. We want to show

$$\underline{\lim}_{n\to\infty} P(X_{\tau_n} = L_n) > \lim_{n\to\infty} P(X_{\omega_n} = L_n) = \frac{1}{e}.$$

For $n>\varepsilon>0$, let

$$\xi_{n,\varepsilon} = \min\{n, \min_{k\geq 2}\{k: R_k > 1-\frac{\varepsilon}{n}, X_k = L_k\}\}$$

$$\rho_{n,\varepsilon} = \min\{\xi_{n,\varepsilon}, \omega_n\}.$$

For each $n>\varepsilon>0$ $\rho_{n,\varepsilon}$ and $\xi_{n,\varepsilon}$ are invariant rules. We will show there is an $\varepsilon>0$ such that

$$\underline{\lim}_{n\to\infty} P(X_{\rho_{n,\varepsilon}} = L_n) > \frac{1}{e}.$$

Let $A_{n,\varepsilon} = \{\rho_{n,\varepsilon} = \xi_{n,\varepsilon} < k(n)\}$. In order for X_{ω_n} to equal L_n on the set $A_{n,\varepsilon}$ at least one of $\{X_k\}_{k=k(n)+1}^{n}$ must exceed $1-\frac{\varepsilon}{n}$. Thus

$$P(X_{\omega_n} = L_n | A_{n,\varepsilon}) \leq 1 - (1-\frac{\varepsilon}{n})^{n-k(n)} \to 1-e^{-\varepsilon(1-\frac{1}{e})} \quad \text{as } n\to\infty.$$

On the other hand $X_{\xi_{n,\varepsilon}}$ will equal L_n if $X_j < X_{\xi_{n,\varepsilon}}$ for $\xi_{n,\varepsilon} < j \leq n$. The probability of this is greater than $(1-\frac{\varepsilon}{n})^n$ on the set $A_{n,\varepsilon}$. Thus $P(X_{\rho_{n,\varepsilon}} = L_n | A_{n,\varepsilon}) \geq (1-\frac{\varepsilon}{n})^n \to e^{-\varepsilon}$ as $n\to\infty$.

Since $\{\rho_{n,\varepsilon} = \omega_n\} = A^c_{n,\varepsilon}$ we have

$$P(X_{\rho_{n,\varepsilon}} = L_n | A^c_{n,\varepsilon}) = P(X_{\omega_n} = L_n | A^c_{n,\varepsilon}).$$

Clearly

$$\underline{\lim}_{n\to\infty} P(A_{n,\varepsilon}) = \delta(\varepsilon) > 0 \quad \text{for each } \varepsilon>0.$$

If we choose $\varepsilon>0$ such that

$$1 - e^{-\varepsilon(1-\frac{1}{e})} < e^{-\varepsilon}$$

then

$$\lim_{n\to\infty} P(X_{\rho_{n,\varepsilon}} = L_n) - \lim_{n\to\infty} P(X_{\omega_n} = L_n) \geq$$

$$(e^{-\varepsilon} - (1-e^{-\varepsilon(1-\frac{1}{e})}))\,\delta(\varepsilon) > 0 \;. \qquad //$$

Remarks

Theorem 4.4, the results of Chapter II and the work of Samuels previously cited (Chapter I Section 4) show that asymptotically the $U[\theta-\frac{1}{2},\theta+\frac{1}{2}]$ Problem is an intermediate case lying between the extremes represented by the $N(\mu,\sigma^2)$ Problem and the $U[\alpha,\beta]$ Problem.

It would be gratifying to know the true value of $\lim_{n\to\infty} P(X_{\tau_n}=L_n)$ or even that such a limit exists. From Table 4.2 it seems that $P(X_{\tau_n}=L_n)$ is decreasing in n (this holds not only for the values of n displayed in Table 4.2 but also for all $n=1,\ldots,50$). This is in keeping with the results in the F.I. and N.I. Problems and suggests that a good estimate of $\lim_{n\to\infty} P(X_{\tau_n}=L_n)$ may be obtainable by extrapolating these known results.

For example, in the N.I. and F.I. Problems if one considers $P(X_{\omega_n}=L_n)$ and $P(X_{\sigma_n}=L_n)$ as linear in $\frac{1}{n}$ one obtains very good estimates of their limits by extrapolation. In the case of full information the extrapolated value is .580121 while the true value is .580164. In the no information case the extrapolated value is .36759 while the true value is .36787. Both results were obtained by linear extrapolation in $\frac{1}{n}$ from the values n = 40, 50.

If the same procedure is applied to the values in Table 4.2 we obtain an estimate of $\lim_{n\to\infty} P(X_{\tau_n} = L_n) = .43516$.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Cayley, A., Mathematical questions and their solutions, Educational Times, 22, (1875) pp. 18-19.

[2] Chow, Y.S., Moriguti, S., Robbins, H., and Samuels, S.M., Optimal selection based on relative rank, Israel Journal of Mathematics, Vol. 2, No. 2, (1964) pp. 81-90.

[3] De Groot, M., Some problems of optimal stopping, Journal of the Royal Statistical Society - Series B, Vol. 30, No. 1, (1968) pp. 108-122.

[4] Dynkin, E.B. and Yushkevich, A.A., Markov Processes Theorems and Problems, Plenum Press, New York (1969) pp. 87-145.

[5] Ferguson, T.S., Mathematical Statistics: A Decision Theoretic Approach, Academic Press, New York, (1967) pp. 340-348.

[6] Fox, J.H. and Marnie, L.G. in Martin Gardner's column: Mathematical Games, Scientific American, 202, (1960) pp. 150, 153.

[7] Gilbert, J.P. and Mosteller, F., Recognizing the maximum of a sequence, Journal of the American Statistical Association, Vol. 61, (1966) pp. 35-73.

[8] Guttman, I., On a problem of L. Moser, Canadian Mathematical Bulletin, Vol. 3, No. 1, (1960) pp. 35-39.

[9] Kiefer, J., Invariance, minimax sequential estimation, and continuous time processes, Annals of Mathematical Statistics, Vol. 28, No. 3, (1957) pp. 573-601.

[10] McKean, H.P., Stochastic Integrals, Academic Press, New York, (1969) p. 4.

[11] Moser, L., On a problem of Cayley, Scripta Mathematica, Vol. 22, (1956) pp. 289-292.

[12] Moser, L. and Pounder, J.R. in Martin Gardner's column: Mathematical Games, Scientific American, 202, (1960) pp. 178, 181.

[13] Mucci, A.G., Differential equations and optimal choice problems, Annals of Statistics, Vol. 1, No. 1, (1973) pp. 104-113.

[14] Mucci, A.G., On a class of secretary problems, Annals of Probability, Vol. 1, No. 3, (1973) pp. 417-427.

[15] Sakaguchi, M., Dynamic programming of some sequential sampling design, Journal of Mathematical Analysis and Applications, Vol. 2, No. 3, (1961) pp. 446-466.

[16] Samuels, S.M., Minimax stopping rules when the distribution is uniform with unknown endpoints, Technical Report No. 523, Purdue University, Department of Statistics (1978).

[17] Stewart, T.J., Optimal selection from a random sequence with learning of the underlying distribution, Journal of the American Statistical Association, (1978), (to appear).

[18] Wallace, D.L., Bounds on normal approximations to Student's and the chi-square distributions, Annals of Mathematical Statistics, Vol. 30, No. 4, (1959) pp. 1121-1130.

VITA

VITA

Joseph David Petruccelli was born on October 4, 1949 in Hartford, Connecticut, the first child of Guy and Theresa Petruccelli. After graduation from high school in 1967 he attended Boston College earning an A.B. in Mathematics in June, 1971. After a year off for good behavior he enrolled in the graduate mathematics program at Purdue in 1972, receiving an M.S. degree in August, 1974. He has spent the last four years as a student in the Department of Statistics at Purdue.