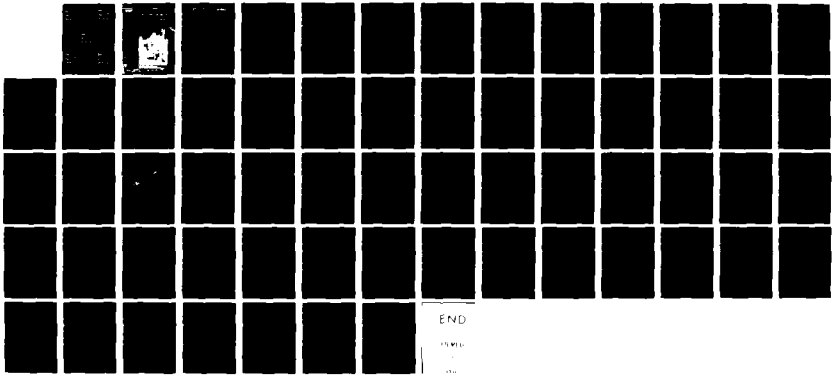
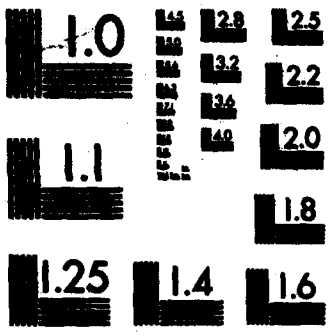
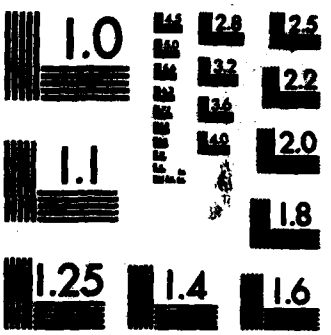
AD-A120 785 SIGNAL DETECTION FOR UNIFORM RENEWAL PROCESSES(U) SAN DIEGO STATE UNIV CA DEPT OF MATHEMATICAL SCIENCES Y CHOI ET AL. JUN 82 1-82 N00014-80-C-0208 1/1

UNCLASSIFIED F/G 17/2 NL

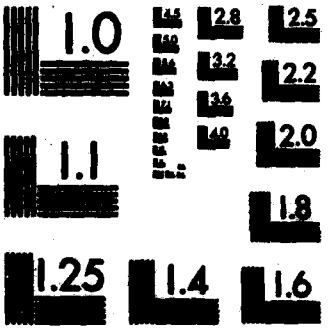




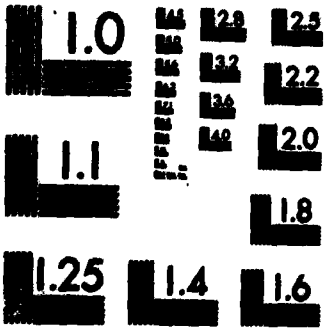
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



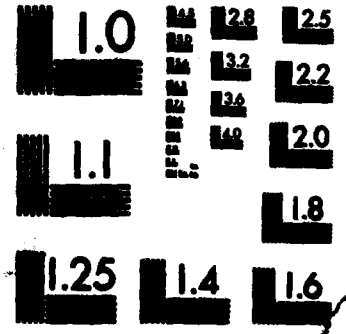
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A120785

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 1-82
2. GOVT ACCESSION NO.: AD-A120785
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): Signal Detection for Uniform Renewal Processes

5. TYPE OF REPORT & PERIOD COVERED: Technical

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Y. Choi, C.B. Bell, R. Ahmad, C. J. Park

8. CONTRACT OR GRANT NUMBER(s): N00014-80-C-0208

9. PERFORMING ORGANIZATION NAME AND ADDRESS: San Diego State University, San Diego, CA 92182

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Arlington, VA 22217

12. REPORT DATE: June 1982

13. NUMBER OF PAGES: 59

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Gaussian, Gaussian-mixture, renewal process, counting process, interarrivals, nonparametric distribution-free, parametric distribution-free, minimal sufficient statistic, maximal statistical noise, Kolmogorov-Smirnov statistic, basic data transformation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Detection problems for uniform renewal processes are represented in the formats of GOF(goodness-of-fit), 2-sample; and GOF-with-nuisance-parameter problems. The techniques are all based on the minimal sufficient statistic or its "orthogonal" complement. The statistics employed are, for the most part, that of Kolmogorov, and its modification by Lilliefors and Srinivasan. Each technique is illustrated by a numerical example.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SIGNAL DETECTION FOR UNIFORM RENEWAL PROCESSES

Y. Choi+, C. B. Bell*, R. Ahmad**, C. J. Park*

San Diego State University*

University of Strathclyde**

State University of New York, Buffalo+

DTIC
ELECTE
OCT 27 1982
S
A

1. Introduction and Summary

A major part of the statistical methodology in signal detection is directed towards models in which the data received are Gaussian or Gaussian-mixture. Under these models the classical statistical techniques and their extensions can, in essense, be adapted to signal detection problems.

Another important set of families of models evolves from replacing the Gaussian and Gaussian-mixture assumption with other parametric assumptions. These families include uniform renewal processes, homogeneous Poiss processes, and non-homogeneous Poisson processes. The signal detection methodology here is principally parametric in nature.

In this paper our aim is to investigate signal detection techniques for the uniform renewal processes. The other two cases namely homogeneous and non-homogeneous Poisson processes are also important from signal detection viewpoint, and will be dealt with elsewhere.

The general signal detection problem can be treated as a problem of testing statistical hypotheses, see Bell (1964). In practice, detection is accomplished by means of a device which receives known pure noise (PN), possible signal (PS) data, and/or artificial noise, and from these inputs "decides" (YES or NO) whether or not a signal is present. In

---

This work was principally supported by the Office of Naval Research through Grant No. N00014-80-C-0208.

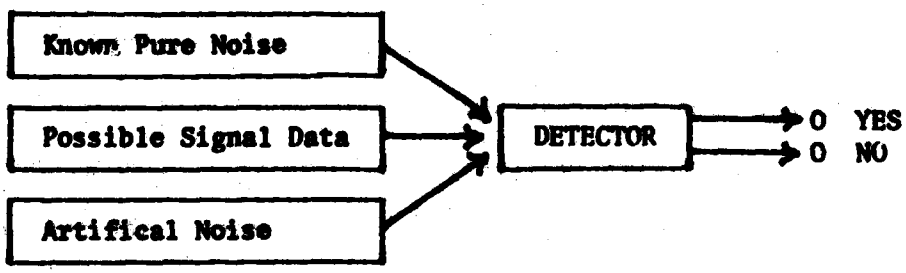
This document has been approved for public release and sale; its distribution is unlimited.

making this decision, the detector is liable to commit two types of erraneous outputs:

(i) false alarm (FA), if NO signal is present and it decides YES; and

(ii) false dismissal (FD), if YES a signal is present and it decides NO.

The detection procedure can be seen schematically from the following diagram:



The probability that the detector will produce a false alarm (PFA) will be designated by α; β will denote the probability that the detector will produce a false dismissal (PFD). Obviously, the perfect detector will be a detector which will produce $\alpha = 0$ and $\beta = 0$; but this has not yet been constructed. However, there is a considerable interest in the so-called ideal or optimal detector. This can be classified into two forms as given below.

(a) Ideal detector with given a priori probability p: This detector minimizes the probability of an error, either FA or FD, that is,

$$\varepsilon(\alpha,\beta, p) = p\beta + (1 - p)\alpha$$

where p denotes the *a priori* probability that a signal is present;

(b) Ideal detector with pre-chosen PFA, α: For this case the objective is to choose α such as .001, .01, .05 etc., and select a decision rule (detection procedure) among all the available procedures which minimizes β; in other words it maximizes $\pi = 1 - \beta$, the power of the detection procedure.

The first type of detection procedures (statistical tests) are commonly known as the Standard Bayesian Detection Procedures (SBDP); and the second type as the Classical Detection Procedures (CDP). If p is known beforehand, clearly it is advantageous to use the first type procedure, otherwise the second type.

The simplest form of the PS data received by the detector consist of discretized observations (note not necessarily discrete values) $X_1, X_2, \ldots, X_n$ of a uniform renewal process, which is a continuous time process, at times $t_1, t_2, \ldots, t_n$. To obtain optimal detectors, the following assumptions are made on the succeeding statistical analysis.

(i) The random variables $X_1, X_2, \ldots, X_n$ are statistically independent;

(ii) $X_1, X_2, \ldots, X_n$ have a common strictly increasing continuous cumulative probability function (cpf) F such that $F(x) = P(X_i \leq x)$. (For the uniform renewal process it is clear that the cpf F is monotone non-decreasing and $F(0) = 0$ and $F(+\infty) = 1$.)

(iii) If the PS data are PN (pure noise), then $F = F_0$; a specified uniform renewal process (URP) cpf; and if the PS data are noise-plus-signal (N + S), then

DTIC COPY INSPECTED 2
Dist Special
A

$F = F_1 \neq F_0$, where F_1 is still URP cpf but different from F_0.

The implications of the various assumptions are discussed in detail in Bell (1964) in connection with the signal detection problems. With the above preliminary considerations, it is now possible to introduce the three major statistical models of detectors to be investigated in subsequent sections.

MODEL I:: PN_1: - The pure noise cpf $F_0 \equiv U(0, \theta_0)$, $0 < \theta_0 < \infty$, is known and there are available PS data $X_1, X_2, \ldots, X_n$ whose cpf $F_1 \equiv U(0, \theta_1)$, $0 < \theta_1 < \infty$, is unknown. In this case we want to detect the null hypothesis; $H_0(PN_1)$: $\theta = \theta_0$ against alternative; $H_a(N + S)$: $\theta \neq \theta_0$. One-sided versions of the alternative hypothesis will be (i) $H_a^+(N + S)$: $\theta > \theta_0$, or (ii) $H_a^-(N + S)$: $\theta < \theta_0$.

MODEL II: PN_2: The problem here consists of having an ideal detection procedure for $H_0[F = U(0,\theta);\ \theta$ unknown] against $H_a[F \neq U(0,\theta)]$.

MODEL III: PN_3: In this case the basic data (BD) set is $Z = (X_1, X_2, \ldots, X_m;\ Y_1, Y_2, \ldots, Y_n)$ or $(W_1, W_2, \ldots, W_m;\ V_1, V_2, \ldots, V_n)$, where $W_j = X_1 + \ldots + X_j$ and $V_j = Y_1 + \ldots + Y_j$. Let $X \sim U(0,\theta_1)$ and $Y \sim U(0,\theta_2)$, where θ_1 and θ_2 are unknown.

We want a detection procedure for $H_0(PN_3)$: $\theta_1 = \theta_2$ against $H_a(N + S)$: $\theta_1 \neq \theta_2$. Similarly, one can define $H_a^+(N + S)$: $\theta_1 > \theta_2$ and $H_a^-(N + S)$: $\theta_1 < \theta_2$.

Since the underlying process is a uniform renewal stochastic process, instead of the probability distributions F_0, F_1, F_2 etc. We shall reformulate our null and alternative hypotheses in terms of probability laws $\mathcal{L}_0$, $\mathcal{L}_1$, $\mathcal{L}_2$ etc. This is done in section three when the suitable detection procedures are discussed.

The organization of the paper is as follows.

In section two basic statistical concepts are briefly outlined. Next, in Section three optimal detection procedures for (i) PN_1 versus $N + S$ and (ii) PN_2 versus $N + S$ are developed. The detection statistics here are of the Kolmogorov-Smirnov-type, Lilliefors- and Srinivasan-type, and some which are based on maximal statistical noise. In Section four PN_3 versus $N + S$ is dealt in detail.

2. Basic Statistical Concepts

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with probability distribution function $F(x)$, and set $W_m = X_1 + X_2 + \ldots + X_m$ for $m = 1, 2, \ldots, n$ with $W_0 \equiv 0$. In stochastic processes terminology we say $\{W_n, n \geq 1\}$ is a renewal process and $\{X_n, n \geq 1\}$ is the dual inter-arrival time process. In the sequal we assume $F(0) = 0$; of course, one could allow an atom

at the origin, but nothing would be gained and we would have to exclude the case of a distribution F concentrated at the origin.

<u>Definition 2.1</u>. If $X_1, X_2, \ldots, X_n$ are independent identically distributed $U(0,\theta)$ random variables, then the process $W_1, W_2, \ldots, W_j, \ldots$ with $W_j = \sum_{i=1}^{j} X_i$ $(W_0 \equiv 0)$ is called the uniform renewal process (URP).

The random variable W_j denotes the waiting time to the jth event. Let $\Omega(\text{URP})$ denote the family of all uniform renewal processes' probability distributions, that is

$$\Omega(\text{URP}) = \{F(\cdot): \ F = U(0,\theta), \quad \theta > 0\}.$$

Consider the situation like models PN_1, PN_2 and PN_3 in which the data received is $\underset{\sim}{X} = (X_1, \ldots, X_m)$ or $\underset{\sim}{Z} = (X_1, \ldots, X_m; \ Y_1, \ldots, Y_n)$. For many practical situations the decision as to whether a signal is present or not is optimally based on the data solely through the minimal sufficient statistics M-S-S, $S(\underset{\sim}{X})$ or $S(\underset{\sim}{Z})$. However, in many other situations one needs a quantity complementary to the M-S-S; this is called the maximal statistical noise (M-S-N) and is denoted by $N(\underset{\sim}{X})$ or $N(\underset{\sim}{Z})$ depending upon the model in the problem. In what follows we write $\underset{\sim}{Z}$ for the generic data point.

(A) <u>Basic Data Transformation and Maximal Statistical Noise</u>

<u>Definition 2.2</u>. Let $\delta(\underset{\sim}{Z}) = [N(\underset{\sim}{Z}), S(\underset{\sim}{Z})]$ be one-to-one almost everywhere

with the M-S-S, $S(\underset{\sim}{Z})$, and the entity $N(\underset{\sim}{Z})$ be statistically independent of $S(\underset{\sim}{Z})$. For the family of interest $\Omega(\text{URP})$, (i) $\delta(\underset{\sim}{Z})$ is called the basic data transformation (BDT); and (ii), $N(\underset{\sim}{Z})$ is called the maximal statistical noise (M-S-N).

<u>Example 2.1</u>. Let $X_1, X_2, \ldots, X_m$ be the inter-arrival times of URP in the family $\Omega(\text{URP})$. Then the maximum likelihood estimator for θ is $\hat{\theta} = X_{(m)}$, the largest order statistic of the generic data point. Clearly, M-S-S is $X_{(m)}$ and a uniformly minimum variance unbiased estimate for θ is $\theta^* = (\frac{m+1}{m})X_{(m)}$. Now, the vector of random variables

$$\underset{\sim}{V}^* = (V_1^*, V_2^*, \ldots, V_{m-1}^*) = \{\frac{X_{(1)}}{X_{(m)}}, \frac{X_{(2)}}{X_{(m)}}, \ldots, \frac{X_{(m-1)}}{X_{(m)}}\}$$

is not a M-S-N. However, the augmented vector variable defined by

$$(\underset{\sim}{V}^*, \underset{\sim}{R}) = \{V_1^*, \ldots, V_{m-1}^*; \; R(X_1), \ldots, R(X_m)\}$$

is indeed M-S-N, here $R(X_j)$ denotes the rank of the random variable X_j for $j = 1, 2, \ldots, m$.

(B) <u>Types of Distribution-free-ness</u>

There are three distinct types of distribution-free statistics which arise for many types of detection problems.

<u>Definition 2.3</u>. (i) A statistic $T(\underset{\sim}{Z})$ is called nonparametric distribution-free (NPDF) w.r.t. a family Ω^* of stochastic processes if there exists a single distribution function $Q(\cdot)$ such that for all probability

[illegible] $\mathcal{X}$ in Ω^* $P(T(\underset{\sim}{Z}) \leq t | \mathcal{X}) = Q(t)$.

(ii) A family of statistics, $T_1(\underset{\sim}{Z}; \mathcal{X})$, which is indexed by the elements of Ω^* is called parametric distribution-free (PDF) w.r.t. Ω^* if there exists a distribution function $Q_{T_1}(\cdot) \equiv Q_1(\cdot)$, such that for all $\mathcal{X}$ in Ω^* and all t_1,

$$P(T_1(\underset{\sim}{Z}; \mathcal{X}) \leq t_1 | \mathcal{X}) = Q_1(t_1).$$

(Note that each NPDF statistic is PDF).

(iii) A PDF statistic $T(\cdot,\cdot)$ is called strongly distribution-free (SDF) w.r.t. Ω^* (and in the PN-situation: $\mathcal{X} = \mathcal{X}_0$) if the FDR (false dismissal rate) depends on $\mathcal{X}_1$ only through $\mathcal{X}_1 \circ \mathcal{X}_0^{-1}$, where the symbol $\circ$ denotes the composition of the involved functions.

Now, after these preliminaries, we proceed to discuss some optimal detection procedures for the family of renewal stochastic processes with uniform inter-arrivals.

3. Detection Procedures for Renewal Processes with Uniform Interarrivals

(A) Standard Detection Procedures for a Uniform Distribution (PN_1: MODEL 1)

Before we develop some detection procedures for renewal processes with uniform inter-arrivals with nuisance parameter θ, it seems worthwhile to point out some other detection procedures in the case of (PN_1) $U(0,\theta)$, $\theta > 0$, distribution. Let $\underset{\sim}{Z} = (X_1, \ldots, X_m)$ be a sample from the uniform distribution on $(0,\theta)$. For detecting H_0: $\theta \leq \theta_0$

against H_a: $\theta > \theta_0$ any test, $\phi(z)$, is uniformly most powerful (UMP) at the significance level α for which $E_{\theta_0} \phi(\underset{\sim}{Z}) = \alpha$, $E_\theta \phi(\underset{\sim}{Z}) \leq \alpha$ for $\theta \leq \theta_0$; and $\phi(\underset{\sim}{Z}) = 1$ when $X_{(m)} = \max(X_1, \ldots, X_m)$ is greater than θ_0 and zero otherwise [Lehmann (1959, p. 110)].

For detecting H_0: $\theta = \theta_0$ against H_a: $\theta \neq \theta_0$ a unique UMP unbiased test exists and is given by $\phi(\underset{\sim}{Z}) = 1$ when $X_{(m)} > \theta_0$ or $X_{(m)} \leq \theta_0 \, m\sqrt{\alpha}$, and $\phi(\underset{\sim}{Z}) = 0$ otherwise. Similarly, a UMP detection procedure for the case H_0: $\theta = \theta_0$ against H_a: $\theta < \theta_0$ can be developed and combined with the one-sided case discussed above. [Note: A variation of the first situation is PN_1 versus $N + S$, that is H_0: $\mathcal{L} = \mathcal{L}_0 \in \Omega(\text{URP})$ versus H_a: $\mathcal{L} <^* \mathcal{L}_0$ (where $<^*$ indicates that $X_{\mathcal{L}}$ is stochastically larger than $X_{\mathcal{L}_0}$); and the second case for URP becomes H_0: $\mathcal{L} = \mathcal{L}_0$ against H_a: $\mathcal{L} \neq \mathcal{L}_1$ ($\mathcal{L} \in \Omega(\text{URP})$).]

<u>Remark 3.1</u>. Consider $Y_1, \ldots, Y_m$ to be a random sample from the family of two-parameter exponential densities given by

$$\{f(x; \alpha, \beta) = \alpha e^{-\alpha(x-\beta)}, \quad x \geq \beta\}.$$

Clearly $(X_1, \ldots, X_m)$ with $X_j = e^{-\alpha Y_j}$ is a random sample from the uniform distribution $U(0, e^{-\alpha\beta})$. Now, one can easily develop the UMP detection procedure for H_0: $\beta = \beta_0$ against H_a: $\beta \neq \beta_0$ when α is assumed known; and also one can determine the UMP detection procedure for H_0: $(\alpha, \beta) = (\alpha_0, \beta_0)$ against the alternative H_a: $\alpha > \alpha_0$, $\beta < \beta_0$.

<u>Remark 3.2</u>. Suppose that in a detection problem we are interested to detect

that certain signals (events) occur uniformly over a stated time interval such as (say) 35 minutes, one hour, one day etc. If the total time interval is divided into N equal parts and p_j denotes the probability of an occurrence (presence of a signal) in the jth subinterval, the detection problem becomes H_0: $p_j = N^{-1}$ for $j = 1, 2, \ldots, N$ against the alternative H_a: $p_j \neq N^{-1}$ for $j = 1, 2, \ldots, N$. Then the detection rule can be based on the statistic (a chi-squared statistic)

$$T^* = mN \sum_{j=1}^{N} (q_j - N^{-1})^2$$

where q_j is the relative frequency of a signal occurrence in the j th subinterval. The approximate power of the detection procedure is given by the probability of rejection

$$\int_c^\infty \chi^2_{N-1} (\omega; \lambda^2) d\omega,$$

where $\chi^2_{N-1}(\cdot; \lambda)$ is the noncentral chi-squared density with $N - 1$ degrees of freedom and noncentrality parameter $\lambda^2 = mN \sum_{j=1}^{N} (p_j - N^{-1})^2$.

Here, the critical point c is determined by the expression

$$\int_c^\infty \chi^2_{N-1} (\omega) d\omega = \alpha$$

under the null hypothesis.

<u>Remark 3.3</u> [Lehmann (1959, p. 308)]

Suppose that in the above set-up (Remark 3.2), where the hypothesis of uniform distribution is being detected, the alternatives of interest

are those of cyclic movement (perhaps, not an unusual situation in some signal detection problems), which may be represented at least approximately by a sine-wave function

$$p_j = N^{-1} + \rho \int_{2\pi(i-1)/N}^{2\pi i/N} \sin(\omega-\theta^*)d\omega, \quad j = 1,2, \ldots, N$$

Here, ρ is the amplitude and θ^* the phasing of the cycle disturbance. By setting $\xi = \rho\cos\theta^*$, $\eta = \rho\sin\theta^*$ one gets

$$p_j = N^{-1}(1 + a_j\xi + b_j\eta)$$

where

$$a_j = 2N \sin\frac{\pi}{N} \sin(2j - 1)\frac{\pi}{N},$$

$$b_j = -2N \sin\frac{\pi}{N} \cos(2j - 1)\frac{\pi}{N}.$$

The quantities $\hat{\xi}, \hat{\eta}$ which minimize $N \sum_{j=1}^{N} (q_j - p_j)^2$ (subject to the fact that equations for p_j define a surface) are

$$\hat{\xi} = N \sum_{j=1}^{N} a_j q_j \Big| \sum_{j=1}^{N} a_j^2 ; \qquad \hat{\eta} = N \sum_{j=1}^{N} b_j q_j \Big| \sum_{j=1}^{N} b_j^2 .$$

For $N > 2$, after some algebraic simplification the detection rule becomes

$$2m[\sum_{j=1}^{N} q_j \sin(2j - 1)\frac{\pi}{N}]^2 + 2m[\sum_{j=1}^{N} q_j \cos(2j - 1)\frac{\pi}{N}]^2 > c$$

where the number of degrees of freedom of the left-hand side expression, a

chi-squared statistic, is $d = 2$. The noncentral parameter determining the approximate power, that is for fixed PFA α minimizing PFD β for $N + S$ alternative cpf F_1, is

$$\lambda^2 = m(\xi N \sin \frac{\pi}{N})^2 + m\,(\eta N \sin \frac{\pi}{N})^2$$

$$= m\rho^2 N^2 \sin^2 \frac{\pi}{N}\,.$$

Now, we return to the uniform renewal processes with given interarrival times when the underlying parameter becomes a nuisance parameter.

(B) <u>The Kolmogorov-Smirnov-Type Detection Procedures for Renewal Processes with Uniform Interarrivals (Nuisance Parameters Case)</u>

The object here is to develop the Kolmogorov-Smirnov-type detection procedures for "H_0: $F = U(0,\theta)$ for some $\theta > 0$", that is the case PN_1: $\mathcal{X}_F \in \Omega(URP)$, where $X_1, X_2, \ldots, X_m, \ldots$ are the interarrival "times" of a renewal process with distribution law $\mathcal{X}_F \equiv \mathcal{X}$

If the value of θ were specified as θ_0 then, in addition to statistics employed in section (A), a natural Kolmogorov-Smirnov statistic is

$$D_m = \sup_x |F(x, \theta_0) - F_m(x)|,$$

where $F_m(x)$ is the sample cpf. However, in the situation under discussion the value of θ is unknown, and, therefore, θ is a nuisance parameter. In the case of nuisance parameter structure, we define two

variants of Kolmogorov-Smirnov-type statistics. These are called, respectively, Lilliefors-type statistics and Srinivasan-type statistics; see Lilliefors (1967, 1969), Srinivasan (1970) and Lieberman and Resnikoff (1955) for some special cases.

<u>Definition 3.1</u>. (i) Let $\underset{\sim}{Z} = (X_1, \ldots, X_m)$ be a random sample with distribution function F_θ, and let $F_m(x)$ be its sample cumulative distribution, where θ is unknown. The statistic

$$\hat{D}_m = \sup_x |\hat{F}_\theta(x) - F_m(x)| ,$$

where $\hat{F}_\theta(x)$ is the maximum likelihood estimator (MLE) of $\hat{F}_\theta(x)$, is called Lilliefors-type statistic.

(ii) With the same notation as above, the statistic

$$\tilde{D}_m = \sup_x |F_m(x) - \tilde{F}_\theta(x)| ,$$

where $\tilde{F}_\theta(x) = E\{S(X_1)|T\}$ with $S(X_1) = I_{\{X_1 \leq x_1\}}$, the indicator function, and T as a sufficient statistic for θ, is called Srinivasan-type statistic.

The above two types of statistics are examples of NPDF statistics w.r.t. an appropriate family of distributions Ω^*, such as {normals}, {exponentials}, and Ω(URP).

Let a probability law $\mathcal{L}_0(X; \theta_1, \theta_2, \ldots, \theta_k)$, say, with distribution function $F_0(x; \theta_1, \ldots, \theta_k)$ be such that $T_1, T_2, \ldots, T_k$ are the joint sufficient statistics for the vector parameter $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k)$.

For a fixed real x_1 define the random variable $S(X_1) = I_{\{X_1 \le x_1\}}$, where I_A is the indicator function of the set A. Clearly $S(X_1)$ is an unbiased estimator of $F_0(x; \theta_1, \theta_2, \ldots, \theta_k)$ under H_0: $F = F_0$. Consequently, from the general theory of sufficient statistics it follows that the statistic

$$\tilde{F}_{\underset{\sim}{\theta}}(x) = E\{S(X_1)|T_1, \ldots, T_k\}$$

is an unbiased estimator of $F_0(x; \underset{\sim}{\theta})$ with a smaller variance than $S(X_1)$. If in addition $(T_1, \ldots, T_k)$ are complete, then $\tilde{F}_{\underset{\sim}{\theta}}(x)$ is the unique MVUE of $F_0(x; \underset{\sim}{\theta})$. If the distribution of statistics like $\hat{D}_m$ or $\tilde{D}_m$ does not depend on $\underset{\sim}{\theta}$, then a statistic of this type would be an appropriate statistic for detecting a composite hypothesis F.

Now, we specialize these statistics to the case of renewal processes with uniform interarrivals, where the nuisance parameter is $\theta > 0$.

<u>MODEL: PN_2.</u>

In the treatment of signal detection problems the objective here is to detect

$$PN_2: \quad \mathcal{L}(\underset{\sim}{Z}) \in \Omega(URP) \quad \text{against} \quad (N + S): \mathcal{L}(\underset{\sim}{Z}) \notin \Omega(URP),$$

where $\underset{\sim}{Z} = (X_1, \ldots, X_m)$ are the interarrival times.

Recall that the general rule in constructing detection procedures for problems like PN_1, PN_2 and PN_3 is:

(i) to use the M-S-S when the PN formulation entails a completely specified stochastic process law (as in the case with MODEL 1: PN_1), and

(ii) employ the M-S-N when the PN formulation entails membership in a family of stochastic processes laws (this will be the case for PN_2 and PN_3).

(a) <u>Lilliefors-Type Detection Statistics</u>

This statistic was earlier defined to be of the form:

$$\hat{D}_m = \sup_x |\hat{F}_\theta(x) - F_m(x)|,$$

where $\hat{F}_\theta(x)$ is the MLE of $F_\theta(x)$, and $F_m(x)$ is the cumulative sample distribution, that is, $F_m(x) = (\text{number of } X_j\text{'s} \leq x)|m = m^{-1} \sum_{j=1}^{m} \varepsilon\,(x - X_j)$, where $\varepsilon(u) = 1$ if $u \geq 0$ and zero otherwise.

The MLE of θ is $X_{(m)} = \max(X_1, \ldots, X_m)$, which also happens to be M-S-S. Note that $T \equiv X_{(m)}$ is an example of a PDF statistic and $T(Z; \mathcal{L}_{\theta_0}) = X_{(m)}|\,\theta$ is an example of a NPDF statistic. Clearly, we have

$$\hat{F}_\theta(x) = \begin{cases} 0 & \text{if } x < 0, \\ \dfrac{x}{X_{(m)}}, & 0 \leq x \leq X_{(m)} \\ 1 & \text{if } x > X_{(m)} \end{cases}$$

Therefore, the Lilliefors type detection statistic can be written as

$$\hat{D}_m = \sup_x |F_m(x) - \hat{F}_\theta(x)|$$

$$= \sup_{0 \leq x \leq X_{(m)}} \left| \frac{1}{m} \sum_{j=1}^{m} \varepsilon\,[x - X_{(j)}] - \frac{x}{X_{(m)}} \right|$$

$$\overset{\mathcal{L}}{=} \sup_{0 \le x \le X_{(m)}} \left| \frac{1}{m} \sum_{j=1}^{m-1} \varepsilon \left[\frac{x}{X_{(m)}} - \frac{X_{(j)}}{X_{(m)}} \right] - \frac{x}{X_{(m)}} \right|$$

$$\overset{\mathcal{L}}{=} \sup_{0 \le u \le 1} \left| \frac{1}{m} \sum_{j=1}^{m-1} \varepsilon \left[u - U(j) \right] - u \right|$$

when $u = x/X_{(m)}$ and $U_{(j)} = X_{(j)}/X_{(m)}$ and $\overset{\mathcal{L}}{=}$ means "it has the same probability $\mathcal{L}$ or distribution".

Before we give the actual decision rule, we sum up the basic distribution structure as below.

<u>Theorem 3.1</u>. Under $H_0(PN_2)$, the following are valid:

(i) The statistic $T(Z, \mathcal{L}_{\theta_0}) = \frac{X_{(m)}}{\theta}$ has the distribution given by $P\{X_{(m)} | \theta \le u\} = u^n, \quad 0 < u < 1.$

(ii) If $R(X_j)$ denotes the rank of X_j then the vector variable $\underset{\sim}{R} = [R(X_1), \ldots, R(X_m)]$ is distributed uniformly over the permutations of $\{1, 2, \ldots, m\}$. This we write as $\underset{\sim}{R} \sim D - U\{S_m(1, 2, \ldots, m)\}$.

(iii) The vector

$$\underset{\sim}{V}^* = \left[\frac{X_{(1)}}{X_{(m)}}, \ldots, \frac{X_{(m-1)}}{X_{(m)}} \right] = (V_1^*, \ldots, V_{m-1}^*)$$

has the same law as $[U_{(1)}, \ldots, U_{(m-1)}]$, where $U_1, \ldots, U_{m-1}$ are i.i.d. random variables from $U(0,1)$. That is, $\underset{\sim}{V}^* \overset{\mathcal{L}}{=} [U_{(1)}, \ldots, U_{(m-1)}]$, where $U_{(j)}$ is the jth order statistic.

<u>Decision Rule</u>. For detecting PN_2: $\mathcal{L} \in \Omega(URP)$ versus $N + S$: $\mathcal{L} \notin \Omega(URP)$;

decide $N \to S$, if and only if, $\hat{D}_m > d^*$, where $P\{\hat{D}_m \leq d^*\} = 1 - \alpha$, and $\alpha = P\{\text{False Alarm}\}$.

Lilliefors (1967, 1969) have given some critical values in other cases but not for a uniform renewal process. Choi (1981) has computed some tables for this case. However, we give in the Appendix an improved version of Choi's table which can be used in practice as well as some illustrative examples.

(b) Srinivasan-Type Detection Statistics

Consider the same detection problem as in subsection (a) above. This type of statistic (other than URP) was originally suggested by Lieberman and Resnikoff (1955) and later on Srinivasan (1970) investigated some cases in detail. Here, the statistics is

$$\tilde{D}_m = \sup_x |F_m(x) - \tilde{F}_\theta(x)|$$

where $\tilde{F}_\theta(x)$ is the MVUE (minimum variance unbiased estimate) of $F_\theta(x)$. If $\mathcal{X} \in \Omega(\text{URP})$, then by the theory of sufficiency, one gets

$$\begin{aligned}
\tilde{F}_\theta(x) &= E\{I_{[X_1 \leq x]} | X_{(m)}\} \\
&= P\{X_1 \leq x | X_{(m)}\} \\
&= m^{-1} \sum_{j=1}^{m} P(X_j \leq x | X_{(m)}) \\
&= m^{-1} \sum_{j=1}^{m} P(X_{(j)} \leq x | X_{(m)}) \\
&= m^{-1} \{\sum_{j=1}^{m-1} P(X_{(j)} \leq x | X_{(m)}) + P(X_{(m)} \leq x | X_{(m)})\}
\end{aligned}$$

$$= m^{-1} \sum_{j=1}^{m-1} P(X_{(j)} \leq x | X_{(m)})$$

Finally,

$$\tilde{F}_\theta(x) = \begin{cases} \frac{m-1}{m} \frac{x}{X_{(m)}}, & 0 < x < X_{(m)} \\ 1, & x \geq X_{(m)} \\ 0, & \text{otherwise.} \end{cases}$$

Consequently, the detection statistic of this type becomes

$$\tilde{D}_m = \sup_x |F_m(x) - \tilde{F}_\theta(x)|$$

$$= \sup_{0 \leq x \leq X_{(m)}} \left| m^{-1} \sum_{j=1}^{m-1} \varepsilon \left[\frac{x}{X_{(m)}} - \frac{X_{(j)}}{X_{(m)}}\right] - \left(\frac{m-1}{m}\right) \frac{x}{X_{(m)}} \right|$$

$$= \sup_{0 \leq u < 1} \left(\frac{m-1}{m}\right) \left|(m-1)^{-1} \sum_{j=1}^{m-1} \varepsilon \left[(u - u_{(j)}) - u\right]\right|.$$

Thus, we note that the detection statistic $\tilde{D}_m \overset{\mathcal{L}}{=} \left(\frac{m-1}{m}\right) D_{m-1}$, the Kolmogorov-Smirnov statistic with size $m - 1$, for all $m \geq 2$ provided $H_0(PN_2)$: $\mathcal{L} \in \Omega(URP)$ holds.

<u>Decision Rule</u>. For detecting PN_2: $\mathcal{L} \in \Omega(URP)$ against $N + S$: $\mathcal{L} \notin \Omega(URP)$; decide $N + S$, if and only if, $\tilde{D}_m > \tilde{d}$, where $P\{\tilde{D}_m \leq \tilde{d}\} = 1 - \alpha$, and α is PFA.

Srinivasan (1970) computed critical values for the normal and exponential cases, and Choi (1981) for the uniform Case. For critical values in connection with detection statistic $\tilde{D}_m$, one can use the standard Kolmogorov-Smirnov table through the relation $\tilde{D}_m = \left(\frac{m-1}{m}\right) D_{m-1}$.

(c) <u>A Detection Statistic Based on Maximal Statistical Noise.</u>

We have already defined a M-S-N statistic, $N(\underset{\sim}{Z})$, in definition 2.2. For the URP case, that is for $PN_2: \mathcal{L} \in \Omega(URP)$ the maximal statistical noise (M-S-N) statistic is

$$(\underset{\sim}{V}^*, \underset{\sim}{R}) = (V_1^*, V_2^*, \ldots, V_{m-1}^*; \; R_1, \ldots, R_m)$$

$$= [\frac{X_{(1)}}{X_{(m)}}, \ldots, \frac{X_{(m-1)}}{X_{(m)}}; \; R(X_1), \ldots, R(X_m)],$$

where $R(X_j)$ denotes the rank of X_j. Note that in PN_2 situation we have a NPDF structure for the underlying statistic which leads us to the natural detection statistic candidate $(\underset{\sim}{V}^*, \underset{\sim}{R})$. Furthermore, we shall see that if a detection procedure does not involve the parameter in the model under consideration, then it is solely based on M-S-N statistic. This motivates one to define a third detection statistic D_m^* as follows:

$$D_m^* = \sup_{0 \le u \le 1} |(m-1)^{-1} \sum_{j=1}^{m-1} \varepsilon\,(u - u_{(j)}) - u|,$$

where $u = x/X_{(m)}$ and $U_{(j)} = X_{(j)}/X_{(m)}$. Obviously, D_m^* is the statistic which is based on M-S-N, $(\underset{\sim}{V}^*, \underset{\sim}{R}^*)$. The statistic D_m^* has the Kolmogorov-Smirnov distribution for a sample of size $m - 1$, that is, $D_m^* = D_{m-1}$ under $PN_2: \mathcal{L} \in \Omega(URP)$. For critical values one can use the standard K-S table with the sample size $m - 1$.

<u>Decision Rule</u>. For detecting $PN_2: \mathcal{L} \in \Omega(URP)$ versus $N + S: \mathcal{L} \notin \Omega(URP)$; decide $N + S$, if and only if, $D_m^* \equiv D_{m-1} > d$ where $P\{D_{m-1} \le d\} = 1 - \alpha$

and $\alpha = P\{\text{False Alarm}\}$.

In essence, the detection procedure structure for the Model II: PN_2 can be summarized by the following result.

<u>Theorem 3.2</u>. Under $H_0: \mathcal{L} \in \Omega(\text{URP})$:

(i) $D_m^* = D_{m-1}$, for $m \geq 2$;

(ii) $\tilde{D}_m = (\frac{m-1}{m}) D_{m-1}$, for all $m \geq 2$,

(iii) $\hat{D}_m = \sup_{0 \leq u < 1} |(\frac{m-1}{m}) F_{m-1}(u) - u|$, where $F_{m-1}(\cdot)$ is the sample cpf based on $u_1, \ldots, u_{m-1}$ which are i.i.d. $U(0,1)$.

<u>Proof</u>. (i) This is an immediate consequence of the fact that $(V_1^*, \ldots, V_{m-1}^*)$ has the same distribution as the $(m - 1)$ uniform order statistics $(U_{(1)}, \ldots, U_{(m-1)})$, see theorem 3.1 part (iii).

(ii) That $\tilde{D}_m$ has the same distribution as $(\frac{m-1}{m}) D_{m-1}$ follows from the fact that:

$$\hat{D}_m = \sup_{0 \leq z \leq X_{(m)}} \left| \frac{1}{m} \sum_{j=1}^{m} \varepsilon (z - X_{(j)}) - (\frac{m-1}{m}) \frac{z}{X_{(m)}} \right|$$

$$= \sup_{0 \leq u < 1} (\frac{m-1}{m}) \left| \frac{1}{m-1} \sum_{j=1}^{m-1} \varepsilon (u - U_{(j)}) - u \right|$$

where $u = z/X_{(m)}$ and $U_{(j)} = X_{(j)}/X_{(m)}$.

(iii) Clearly,

$$\hat{D}_m = \sup_{0 \le z \le X_{(m)}} \left| \frac{1}{m} \sum_{j=1}^{m} \varepsilon\,(z - X_{(j)}) - \frac{z}{X_{(m)}} \right|$$

$$= \sup_{0 \le z < X_{(m)}} \left| \frac{1}{m} \sum_{j=1}^{m} \varepsilon\,[\frac{z}{X_{(m)}} - \frac{X_{(j)}}{X_{(m)}}] - \frac{z}{X_{(m)}} \right|$$

$$= \sup_{0 \le u < 1} \left| \frac{1}{m} \sum_{j=1}^{m-1} \varepsilon\,(u - U_{(j)}) - u \right|$$

$$= \sup_{0 \le u < 1} \left| (\frac{m-1}{m})\, F_{m-1}(u) - u \right|,$$

where $U_1, U_2, \ldots, U_{m-1}$ are i.i.d. $U(0,1)$. This establishes the theorem.

Since $\frac{m-1}{m} \to 1$ as $m \to \infty$, it follows that:

<u>Theorem 3.3</u>. Asymptotically (as $m \to \infty$) the statistics D^*, $\tilde{D}$ and $\hat{D}$ have the same distribution as the usual Kolmogorov-Smirnov statistic D_∞.

4. Detection Procedures for MODEL III: PN_3

The received basic data (BD) set is $\underset{\sim}{Z} = (X_1, \ldots, X_m; Y_{m+1}, \ldots, Y_{m+n}) = (Z_1, \ldots, Z_N)$. The detection problem for this situation is

$$H_0(PN_3): \mathcal{X}_1 = \mathcal{X}_2 (\mathcal{X}_i \in \Omega(URP),\ i = 1, 2),$$

against

$$H_a(N + S): \mathcal{X}_1 \neq \mathcal{X}_2\ (\mathcal{X}_i \in \Omega(URP),\ i = 1, 2).$$

For this case the natural detection statistic will be a NPDF statistic and should depend on the M-S-N, $N(\underset{\sim}{Z}) = (\underset{\sim}{V}^*, \underset{\sim}{R})$, where

$$(\underset{\sim}{V}^*, \underset{\sim}{R}) = (V_1^*, V_2^*, \ldots, V_{N-1}^*; \; R_1, R_2, \ldots, R_N)$$

$$= \{\frac{Z_{(1)}}{Z_{(N)}}, \frac{Z_{(2)}}{Z_{(N)}}, \ldots, \frac{Z_{(N-1)}}{Z_{(N)}}; \; R(Z_1), \ldots, R(Z_N)\}$$

Under H_0, $\underset{\sim}{Z}_{(\cdot)} = (Z_{(1)}, \ldots, Z_{(N)})$ and $\underset{\sim}{R}$ are independent since one can write their joint p.d.f. as equal to the product of the marginal p.d.fs., that is,

$$f_{\underset{\sim}{Z}(\cdot)}(\underset{\sim}{z}_{(\cdot)}, \underset{\sim}{r}) = f_{\underset{\sim}{Z}_{(\cdot)}}(\underset{\sim}{z}_{(\cdot)}) \cdot f_{\underset{\sim}{R}}(\underset{\sim}{r})$$

Note that, under H_0,

$$f_{Z_{(\cdot)}}(\underset{\sim}{z}_{(\cdot)}) = N! \prod_{j=1}^{N} f_{Z_1}(z_{(j)}) I_A$$

where $A = \{(z_{(1)}, \ldots, z_{(N)}): \; 0 \leq z_{(1)} < z_{(2)} < \ldots < z_{(N)} \leq \theta = \theta_1 = \theta_2\}$; and $\underset{\sim}{R}$ is distributed uniformly over the permutation set $\{S_N(1, 2, \ldots, N)\}$. Furthermore, $\underset{\sim}{Z}_{(N)}$ and $\underset{\sim}{V}^*$ are independent; and so are $\underset{\sim}{R}(\underset{\sim}{Z})$ and $\underset{\sim}{V}^*$.

In this case we shall develop the likelihood ratio detection procedure, and show that the resulting detection statistics is based on the maximal statistical noise (M-S-N), $N(\underset{\sim}{Z}) = (\underset{\sim}{V}^*, \underset{\sim}{R})$.

(a) <u>Likelihood Ratio Detection Procedure</u>

Let $\Omega_0 = \{(\theta_1, \theta_2): \; \theta_1 = \theta_2 = \theta, \; \theta_i > 0, \; i = 1, 2\}$ and $\Omega_1 = \{(\theta_1, \theta_2): \; \theta_1 \neq \theta_2, \; \theta_i > 0, \; i = 1, 2\}$. Define the set Ω_1^*, instead of Ω_1 when we replace "$\theta_1 \neq \theta_2$" by "$\theta_1 > \theta_2$", and similarly Ω_1' is defined by replacing "$\theta_1 \neq \theta_2$" by "$\theta_2 > \theta_1$". Now, in PN_3

situation we can detect for three distinct cases:

(i) $H(\Omega_0)$ against $H(\Omega_1)$, a two-tailed alternative;

(ii) $H(\Omega_0)$ against $H(\Omega_1^*)$, a one-tail alternative;

(iii) $H(\Omega_0)$ versus $H(\Omega_1')$, a one-tail alternative.

We shall develop the case (ii) . the case (iii) is similar, and part (i), of course, combines cases (ii) and (iii) in the alternative. The likelihood ratio detection procedure principle states that the decision rule is base on "YES N + S", if and only if,

$$\frac{\sup\limits_{\theta \in \Omega_1} L(z, \theta)}{\sup\limits_{\theta \in \Omega_0} L(z, \theta)} > a \quad \text{or} \quad < b,$$

where the existence of constants a and b is guaranteed by the likelihood ratio detection theory. For the case (ii) there exits a constant, say k, such that by the L.R.T.

$$k < \frac{\hat{L}_{\Omega_1^*}}{\hat{L}_{\Omega_0}} = \frac{[X_{(m)}]^{-m}\,[Y_{(n)}]^{-n}}{[Z_{(N)}]^{-N}} = [\frac{Z_{(N)}}{X_{(m)}}]^m\,[\frac{Z_{(N)}}{Y_{(n)}}]^n$$

$$\equiv Q^*.$$

Clearly,

$$Q^* = \begin{cases} [\frac{Y_{(n)}}{X_{(m)}}]^m, & Y_{(n)} = Z_{(N)} \\ [\frac{X_{(m)}}{Y_{(n)}}]^n, & X_{(m)} = Z_{(N)}. \end{cases}$$

Therefore the critical region, C^* can be written as

$$C^* = \{k_1 < [\frac{Y_{(n)}}{X_{(m)}}]^m ; \quad X_{(m)} < Y_{(n)}\} \cup$$

$$\cup \{k_2 < [\frac{X_{(m)}}{Y_{(n)}}]^n ; \quad X_{(m)} > Y_{(n)}\}.$$

Or,

$$C^* = \{ \frac{\max[X_{(m)}, Y_{(n)}]}{\min[X_{(m)}, Y_{(n)}]} > k_1^{-m} , \quad Y_{(n)} = Z_{(N)}\} \cup$$

$$\cup \{ \frac{\max[X_{(m)}, Y_{(n)}]}{\min[X_{(m)}, Y_{(n)}]} > k_2^{-n}, \quad X_{(m)} = Z_{(N)}\}.$$

(The exact expression for k_1 and k_2 will be given later on). Finally, we can write the above expressions as

$$C^* = \{ \frac{\max[X_{(m)}, Y_{(n)}]}{\min[X_{(m)}, Y_{(n)}]} > k^*\}.$$

Let $T = \frac{\max[X_{(m)}, Y_{(n)}]}{\min[X_{(m)}, Y_{(n)}]}$. Before, we give the probability distribution of T under Ω_0 and Ω_1^* we first want to establish that statistic T depends on $N(\underset{\sim}{Z}) = (\underset{\sim}{V}^*, \underset{\sim}{R})$. For this write

$$N(\underset{\sim}{Z}) = (\underset{\sim}{V}^*, \underset{\sim}{R}) = (\underset{\sim}{V}^*, \underset{\sim}{R}_X, \underset{\sim}{R}_Y),$$

$$m^* = \max \{R(Z_1), \ldots, R(Z_m)\} = \max \{\underset{\sim}{R}_X\},$$

$$n^* = \max \{R(Z_{m+1}), \ldots, R(Z_N)\} = \max \{\underset{\sim}{R}_Y\}.$$

Now, notice that

$$\min\left(\frac{Z_{(m^*)}}{Z_{(N)}}, \frac{Z_{(n^*)}}{Z_{(N)}}\right) = \begin{cases} \dfrac{Z_{(m^*)}}{Z_{(N)}}, & Y_{(n)} = Z_{(N)} \Longleftrightarrow n^* = N \\[2ex] \dfrac{Z_{(n^*)}}{Z_{(N)}}, & X_{(m)} = Z_{(N)} \Longleftrightarrow m^* = N. \end{cases}$$

Thus, we conclude that $T = \max(V^*_{m^*}, V^*_{n^*})$, which shows that T indeed depends on $N(Z)$, the maximal statistical noise.

(A) <u>The Distribution of the Statistic T under Ω_0.</u>

For convenience in this section we shall set $X = X_{(m)}$ and $Y_{(n)} = Y$; and similarly $Z_{(N)} = Z$. First note that $t > 1$, now

$$H(t) = P\{T \le t\} = P\{T \le t;\ X = Z\} +$$

$$+ P\{T \le t;\ Y = Z\}$$

$$= P\{\tfrac{X}{Y} \le t;\ X > Y\} + P\{\tfrac{Y}{X} \le t;\ X < Y\}$$

$$= P\{\tfrac{X}{t} \le Y < X\} + P\{X < Y \le Xt\}$$

$$= P\{\tfrac{X}{t} \le Y \le Xt\} = P\{\tfrac{1}{t} \le \tfrac{Y}{X} \le t\}.$$

The integration region can be sketched as in the Diagram (1) below

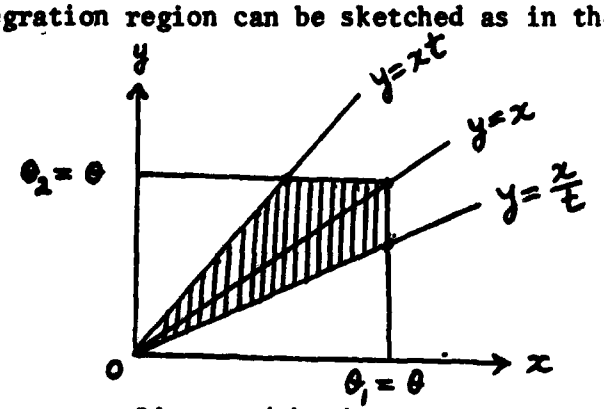


Diagram (1) $(\theta_1 = \theta_2 = \theta)$

Now,

$$H_{m=n}(t;\ \Omega_0) = \int_0^\theta \left[\int_{x/t}^{x} g(y)dy\right]f(x)dx + \int_0^\theta \left[\int_{y/t}^{y} f(x)dx\right]g(y)dy$$

$$= \int_0^\theta \left[G(x) - G(\tfrac{x}{t})\right]f(x)dx + \int_0^\theta \left[F(y) - F(\tfrac{y}{t})\right]g(y)dy$$

Or, one gets:

$$H_{m=n}\ (t;\ \Omega_0) = 1 - t^{-m}$$

Hence, $H_{m=n}(t;\ \Omega_0) = \begin{cases} 0, & t \leq 1 \\ 1 - t^{-m}, & 1 < t < \infty \end{cases}$

Similarly, $H_{m \neq n}(t;\Omega_0) = \begin{cases} 0, & t \leq 1 \\ 1 - \frac{1}{m+n}\left[\frac{m}{t^n} + \frac{n}{t^m}\right], & 1 \leq t < \infty \end{cases}$

(B) <u>The Distribution of Statistic T under Ω_1^* $(\theta_2 < \theta_1)$.</u>

Let the notation be the same as in the case (A) above. Define the sets A_1 and A_2 as follows:

$$A_1 = \{T \leq t;\ X > Y\} = \{\tfrac{X}{Y} \leq t;\ X > Y\} = \{Y < X \leq tY\}$$

$$A_2 = \{T \leq t;\ X < Y\} = \{\tfrac{Y}{X} \leq t;\ X < Y\} = \{X < Y \leq tX\}.$$

The integration region to compute $H(t;\ \Omega_1^*)$ is somewhat complex, and therefore we explicitly illustrate it in the two diagrams below.

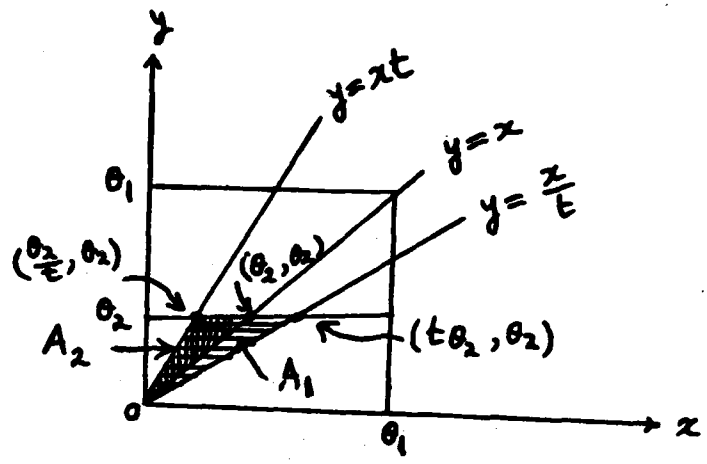


Diagram (3) $(\theta_1 > \theta_2;\ 1 < t < \frac{\theta_1}{\theta_2})$

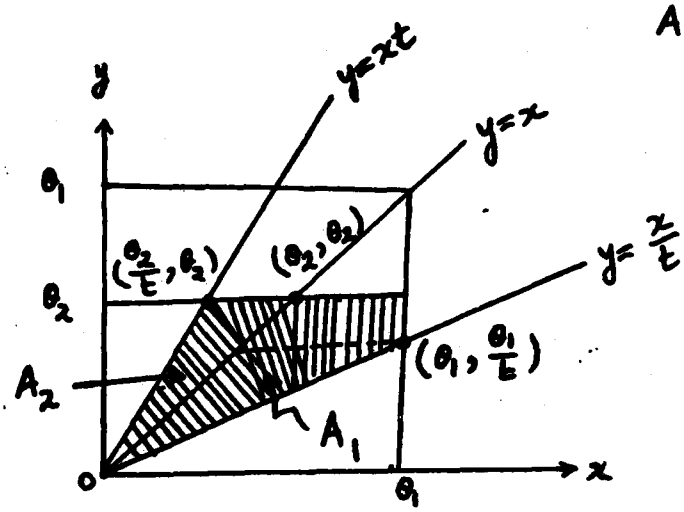


Diagram (2) $(t > 1;\ \theta_1 > \theta_2;\ \frac{\theta_1}{\theta_2} < t)$ Diagram (3) $(\theta_1 > \theta_2;\ 1 < t < \frac{\theta_1}{\theta_2})$

What we want to compute are the following probabilities:

$$H(t;\Omega_1^*) = \begin{cases} 0, & t \le 1 \\ P(A_1) + P(A_2), & 1 < t \le \frac{\theta_1}{\theta_2} \\ P(A_1) + P(A_2), & \frac{\theta_1}{\theta_2} \le t < \infty \end{cases}$$

We shall consider $P(A_2)$, $P(A_1)$ with restriction $1 < t \le \frac{\theta_1}{\theta_2}$, and $P(A_1)$ under the constraint $\frac{\theta_1}{\theta_2} \le t < \infty$, respectively.

$$P(A_2) = \int_0^{\theta_2} [\int_{y/t}^{y} f(x)]\ g(y)dy$$

$$= \int_0^{\theta_2} [F(y) - F(\frac{y}{t})]\ \frac{n\ y^{n-1}}{\theta_2^n}\ dy$$

$$= \frac{n}{\theta_2^n} \int_0^{\theta_2} [(\frac{y}{\theta_1})^m - (\frac{y}{t\theta_1})^m]\ y^{n-1}\ dy$$

$$= \frac{n}{N} (\frac{\theta_2}{\theta_1})^m\ [1 - t^{-m}] \qquad (1)$$

Under the Diagram (3), we obtain

$$P(A_1) = \int_0^{\theta_2} [\int_y^{ty} f(x)dx]\ g(y)dy$$

$$= \int_0^{\theta_2} [(\frac{ty}{\theta_1})^m - (\frac{y}{\theta_1})^m]\ \frac{ny^{n-1}}{\theta_2^n}\ dy$$

$$= \frac{n}{N} (\frac{\theta_2}{\theta_1})^m\ [t^m - 1] \qquad (2)$$

Next, we compute $P(A_1)$ under the Diagram (2)

$$P(A_1) = \int_0^{\theta_1/t} [\int_y^{ty} f(x)dx]\ g(y)dy + \int_{\theta_{1/t}}^{\theta_2} [\int_y^{\theta_1} f(x)dx]g(y)dy$$

$$= \int_0^{\theta_1/t} [(\frac{ty}{\theta_1})^m - (\frac{y}{\theta_1})^m]\ \frac{ny^{n-1}}{\theta_2^n}\ dy +$$

$$\int_{\theta_{1/t}}^{\theta_2} [(\frac{\theta_1}{\theta_1})^m - (\frac{y}{\theta_1})^m]\ \frac{ny^{n-1}}{\theta_2^n}\ dy$$

After some simplification, one gets

$$P(A_1) = 1 - \frac{n}{N}\left(\frac{\theta_2}{\theta_1}\right)^m - \left(\frac{\theta_1}{\theta_2}\right)^m \frac{m}{Nt^n} \qquad (3)$$

Finally, combining expressions (1), (2) and (3), we obtain:

$$H(t;\ \Omega_1^*) = \begin{cases} 0, & t \leq 1 \\ \frac{n}{N}\left(\frac{\theta_2}{\theta_1}\right)^m [t^m - t^{-m}], & 1 \leq t \leq \frac{\theta_1}{\theta_2} \\ 1 - \frac{n}{N}\left(\frac{\theta_2}{\theta_1}\right)^m t^{-m} - \frac{m}{N}\left(\frac{\theta_1}{\theta_2}\right)^n t^{-n}, & \frac{\theta_1}{\theta_2} \leq t < \infty \end{cases}$$

<u>Detection Decision Rule</u>. For detecting PN_3: $\mathcal{X}_1 = \mathcal{X}_2$ ($\mathcal{X}_i \in \Omega(URP)$, $i = 1,2$) against $N + S$: $\mathcal{X}_1 \neq \mathcal{X}_2$, $\mathcal{X}_i \in \Omega(URP)$; decide $N + S$, if and only if,

$$\sup \int_{C^*} dH(t;\ \Omega_0) \leq \alpha$$

where α is P {False Alarm}. This test is optimal in the sense that it minimizes β, the PFD among all detection procedures for fixed α. In other words, the above detection rule maximizes the power

$$1 - P\ \{\text{False Dismissal}\} = 1 - \sup \int_{C^*} dH(t;\ \Omega_1^*),$$

where C^* is the complimentary region to the acceptance region C, in which one decides in favor of PN_3.

Similarly one can devise decision rules in cases Ω_0 against Ω_1' and Ω_0 versus Ω_1.

It turns out that the detection statistic T is equivalent to the statistic $T_1 = X_{(m)}/Y_{(n)}$, which has distribution under the null hypothesis as

$$H(z) = \begin{cases} \frac{n}{N} z^m, & 0 \leq z \leq 1 \\ 1 - \frac{m}{N} z^{-n}, & 1 \leq z < \infty. \end{cases}$$

This statistic was also treated by Murty (1955) when testing two uniform distributions in a parametric setting. Rider (1951) for the same parametric problem gave a statistic which is the ratio of two ranges. The detection statistic proposed by Rider has less power than either of T or T_1 statistics. Khatri (1960) investigates the problem of testing the equality of parameters in k uniform populations. Khatri extends previous results by Roy's (1953) union-intersection principle; again this is in a parametric setting. Some critical values are given by Murty (for $n = 10 = m$) and Khatri takes larger sample sizes. Barr (1966) also gives some results on testing the equality of uniform and related distributions. Another equivalent statistic to statistics T and T_1 is

$$T_2 = [\frac{Z_{(N)}}{X_{(m)}}]^m \, [\frac{Z_{(N)}}{Y_{(n)}}]^n.$$

For $m = n$, the three critical regions given below can easily be seen to be the same. These regions are:

$$C_1^* = \{T_2 > k_1\}$$

$$C_2^* = \{T_1 < a_2 \text{ or } > b_2\}$$

$$C_3^* = \{T > k_3\}$$

Clearly $(\alpha_0)^{-1/m} = k_3 = b_2 = \frac{1}{a_2} = k_1$. Thus $C_1^* \equiv C_2^* \equiv C_3^*$. This is also the case for $m \neq n$. However, if the critical region is two-sided then the level of significance will be $2\alpha_0$ instead of α_0.

The power of detection statistics $T \equiv T_1 \equiv T_2$ is discussed in the following example which is a slight modification of Murty's result.

<u>Example 4.1</u>. Let $(X_1, X_2, \ldots, X_m)$ be a random sample of size m from the family

$$\Omega_1(\text{URP}) = \{\mathcal{X}_1: \ f(x; \mathcal{X}_1) = (\tau\theta^*)^{-1} I_{(0,\tau\theta^*)}(x); \ \tau > 1\},$$

and $(Y_1, Y_2, \ldots, Y_n)$ be another random sample of size n from the class

$$\Omega_2(\text{URP}) = \{\mathcal{X}_2: \ f(\ ; \mathcal{X}_2) = (\theta^*)^{-1} I_{(0,\theta^*)}(y)\}.$$

To detect PN_3: $\mathcal{X}_1 = \mathcal{X}_2$ versus $N + S$: $\mathcal{X}_1 \in \Omega_1(\text{URP})$ and $\mathcal{X}_2 \in \Omega_2(\text{URP})$, the most powerful detection rule is obtained by the likelihood ratio principle via statistics T, T_1 or T_2. In this case the rejection region is

$$C^* = \{T_1 \geq t_\alpha\}[\text{or } \{T_1 < t_\alpha\} = C],$$

where $t_\alpha = \sqrt[n]{\frac{m}{(m+n)\alpha}}$, where α is critical level. The power, π, of the test is given by

$$\pi = P_{H_a}\{T_1 \geq t_\alpha\} = \begin{cases} \frac{m}{N}(\tau/t_\alpha)^n, & 1 < \tau < t_\alpha \\ 1 - \frac{n}{N}(t_\alpha/\tau)^m, & \tau > t_\alpha. \end{cases}$$

If $0 < \tau < 1$, then the MP detection rule has the critical region: $\tilde{C} = \{T_1 \leq t_\alpha\}$, where $t_\alpha = \sqrt[m]{\frac{(m+n)\alpha}{m}}$, and hence the power of the test is

$$\tilde{\pi} = P\{T_1 \leq t_\alpha | N + S\}$$
$$= \begin{cases} \frac{n}{N}\left(\frac{t_\alpha}{\tau}\right)^m, & t_\alpha < \tau < 1 \\ 1 - \frac{m}{N}\left(\frac{\tau}{t_\alpha}\right)^n, & t_\alpha > \tau \end{cases}$$

<u>A General Expression for the Power.</u>

If we are detecting $H(\Omega_0) = \{(\mathcal{L}_1, \mathcal{L}_2): \mathcal{L}_1 = \mathcal{L}_2 \in \Omega(\text{URP})\}$ against $H(\Omega_1^*) = \{(\mathcal{L}_1, \mathcal{L}_2): \mathcal{L}_1 \neq \mathcal{L}_2;\ \mathcal{L}_i \in \Omega(\text{URP}),\ i = 1, 2,$ and $\theta_1 > \theta_2\}$, then by using the statistic T one can get the power function $\pi(\theta_1, \theta_2)$ as below.

Since the critical region is of the form $C^* = \{T \geq t_\alpha\}$ where $t_\alpha = \sqrt[n]{\frac{m}{(m+n)\alpha}}$, the power function of this UMPU test becomes:

$$P\{T \geq t_\alpha\} = \int_{t_\alpha}^{\theta_1/\theta_2} dH(t;\Omega_1^*) = \frac{n}{N}\left(\frac{\theta_2}{\theta_1}\right)^m \left[\left(\frac{\theta_1}{\theta_2}\right)^m - t_\alpha^m\right] -$$
$$- \left[\left(\frac{\theta_1}{\theta_2}\right)^{-m} - t_\alpha^{-m}\right]\}, \text{ if } 1 \leq t_\alpha \leq \frac{\theta_1}{\theta_2};$$

and

$$P\{T \geq t_\alpha\} = \int_{t_\alpha}^{\infty} dH(t;\ \Omega_1^*)$$

$$= \frac{1}{N}\{n(\frac{\theta_2}{\theta_1})^m\ t_\alpha^{-m} - m(\frac{\theta_1}{\theta_2})^n\ t_\alpha^{-n}\}\ ,\quad \frac{\theta_1}{\theta_2} \leq t_\alpha < \infty\ .$$

<u>An Alternative Approach which is Equivalent to the Detector Statistic T.</u>

As before let $\Omega_0 = \{(\theta_1, \theta_2):\ \theta_1 = \theta_2\}$ and $\Omega_1' = \{(\theta_1, \theta_2):$ $0 < \theta_1 < \theta_2 < \infty\}$. Let

$$X_{(m)} = \max\ \{X_1, X_2, \ldots, X_m\}$$

$$Y_{(n)} = \max\ \{Y_1, Y_2, \ldots, Y_m\}$$

$$Z^*_{(N)} = \max\ [X_{(m)}, Y_{(n)}],\quad m + n = N$$

To detect $H(\Omega_0) = \{(\mathcal{L}_1, \mathcal{L}_2):\ \mathcal{L}_1 = \mathcal{L}_2 \in \Omega(\text{URP})\}$ versus $H(\Omega_1') = \{(\mathcal{L}_1, \mathcal{L}_2):\ \mathcal{L}_1 \neq \mathcal{L}_2;\ \mathcal{L}_i \in \Omega(\text{URP})$ for $i = 1, 2$ and $\theta_1 < \theta_2\}$, consider

$$\frac{f_{\Omega_1'}\ (x_{(\cdot)}|Z^*_{(N)})}{f_{\Omega_0}\ (x_{(\cdot)}|Z^*_{(N)})} \equiv \phi(T_3),\quad \text{say,}$$

where ϕ is an arbitrary function which gives rise to a test statistic T by the (conditional) L.R. (likelihood ratio) principle.

After some computation, one can show that

$$\phi(T_3) = 1, \quad 0 < z^*_{(\cdot)} < \theta_1$$

$$= [\frac{m\, x^{m-1}_{(\cdot)}}{\theta_1^m} I_{(0,\theta_1)}(x_{(\cdot)})] \cdot$$

$$\cdot [\frac{mn}{N} (\frac{x_{(\cdot)}}{z^*_{(\cdot)}})^{m-1} \frac{1}{z^*_{(\cdot)}} I_{(0,\, z^*_{(\cdot)}}(x_{(\cdot)}) + \frac{m}{N} I_{(0,\, z^*_{(\cdot)})}(x_{(\cdot)})],$$

$$\theta_1 < z^*_{(\cdot)} < \theta_2$$

$$= \frac{m\, x^{m-1}_{(\cdot)}\ \theta_1^{-m}\ I_{(0,\theta_1)}(x_{(\cdot)})}{\frac{mn}{N} \frac{(x_{(\cdot)})^{m-1}}{(z^*_{(\cdot)})^m} I_{(0,z^*_{(\cdot)})}(x_{(\cdot)})\ [1 + \frac{m}{N} I_{(0,z^*_{(\cdot)})}(x_{(\cdot)}) + \frac{mn}{N} \frac{x^{m-1}_{(\cdot)}}{z^{*m}_{(\cdot)}} I_{(0,\theta)}(x_{(\cdot)})]}$$

$$\theta_2 < z^*_{(\cdot)} < \infty.$$

Clearly $\dfrac{f_{\Omega_1'}(x_{(\cdot)}|Z^*_{(\cdot)})}{f_{\Omega_0}(x_{(\cdot)}|Z^*_{(\cdot)})} = \phi(T_3)$ is monotone decreasing in $x_{(\cdot)}$ hence by the statistical testing theory [Lehmann (1959)], the uniformly minimum FDR unbiased procedure of size α for detecting $H(\Omega_0)$ against $H(\Omega_1')$ is given by

$$\phi(X_{(m)}) = \begin{cases} 1, & \text{if } X_{(m)} \le c(Z^*_{(\cdot)}) \\ 0, & \text{otherwise,} \end{cases}$$

where $\phi(x_{(\cdot)})$ satisfies

$$\int_0^{z^*_{(\cdot)}} \phi(x_{(\cdot)})\, f_{\Omega_0}(x_{(\cdot)}, z_{(\cdot)})\, dx_{(\cdot)} = \alpha .$$

The critical region is obtained by the expression

$$\int_0^{C(z^*_{(\cdot)})} \left(\frac{mn}{N}\right) \frac{x^{m-1}_{(\cdot)}}{z^{*m}_{(\cdot)}}\, dx_{(\cdot)} = \alpha$$

or

$$C(Z^*_{(\cdot)}) = Z^*_{(\cdot)}\, \alpha^{1/m} \left(\frac{N}{n}\right)^{1/m}$$

If we develop a detection procedure this way for detecting $H(\Omega_0)$ against $H(\Omega_1^*)$, then the critical region turns out to be the same as for the statistic T. Thus the test based on T_3 given $Z^*_{(\cdot)}$ is equivalent to the detection procedures T, T_1, T_2.

<u>Decision Rule</u>. For detecting PN_3 versus $N + S$: $H(\Omega_1')$; decide $N + S$, if and only if,

$$X_{(m)} \leq \alpha^{1/m} \left(\frac{N}{n}\right)^{1/m} Z^*_{(\cdot)} \quad ,$$

$\alpha = P$ {False Alarm}.

(Remark: Detection statistics T, T_1, T_2, T_3 are all SDF wrt $H(\Omega_0)$.)

REFERENCES

1. Barr, D. R. (1966). On testing the equality of uniform and related distributions, J. Amer. Statist. Assoc., 61,856-864.
2. Bell, C . B. (1964). Automatic Distribution Free Statistical Signal Detection, Research report, U.S. Navy, Electronics Laboratory.
3. Choi, Y. J. (1981). Kolmogorov-Smirnov Test with Nuisance Parameters in the Uniform Case, M.S. Thesis, University of Washington, Washington.
4. Durbin, J. (1961). Some methods of constructing exact tests, Biometrika, 48, 41-55.
5. Durbin, J. (1973). Distribution Theory for Tests Based on the Sample Distribution Function, SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia.
6. Khatri, C. G. (1960). On testing the equality of parameters in k rectangular populations, J. Amer. Statist. Assoc., 55, 144-147.
7. Lehmann, E. L. (1959). Testing Statistical Hypotheses, Wiley, New York.
8. Lieberman, G. J. and Resnikoff, G. J. (1955). Sampling plans for inspection by Variables, J. Amer. Statist. Assoc., 50, 457-516.
9. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown, J. Amer. Statist. Assoc., 62, 399-402.
10. Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown, J. Amer. Statist. Assoc., 64, 387-389.

11. Murty, V. N. (1955). The distribution of the quotient of maximum values in samples from a rectangular distribution, J. Amer. Statist. Assoc., 50, 1136-1141.

12. Rider, P. R. (1951). The distribution of the quotient of ranges in samples from a rectangular population, J. Amer. Statist. Assoc., 46, 502-507.

13. Roy, S. N. (1953). On a heuristic method of test Construction and its use in multivariate analysis, Ann. Math. Statist., 24, 220-238.

14. Srinivasan, R. (1970). An approach to testing the goodness-of-fit of incompletely specified distributions, Biometrika, 57, 605-611.

APPENDIX

Tables

Graphs

Numerical Examples

[This appendix was prepared with the aid of S-M Lee and A. Mason]

APPENDIX

TABLE 1. Critical Values of $\hat{D}_m$ in the uniform case Parameter θ is unknown, level of significance α of $\hat{D}_m = \sup_x |F_m(x) - \hat{F}_\theta(x)|$

m \ α	.01	.05	.10	.15	.20	.25	.30
2	0.9909	0.9507	0.8996	0.8501	0.8001	0.7509	0.7018
3	0.8963	0.7746	0.6846	0.6472	0.6222	0.5967	0.5718
4	0.7878	0.6688	0.6071	0.5613	0.5241	0.4950	0.4783
5	0.7120	0.5935	0.5407	0.5030	0.4723	0.4465	0.4229
6	0.6482	0.5463	0.4922	0.4569	0.4304	0.4065	0.3871
7	0.5991	0.5032	0.4541	0.4216	0.3975	0.3779	0.3608
8	0.5603	0.4681	0.4224	0.3917	0.3685	0.3494	0.3341
9	0.5324	0.4457	0.3984	0.3699	0.3480	0.3290	0.3144
10	0.5025	0.4218	0.3789	0.3519	0.3313	0.3145	0.2993
11	0.4765	0.4009	0.3612	0.3342	0.3153	0.2987	0.2854
12	0.4618	0.3839	0.3473	0.3235	0.3033	0.2874	0.2742
13	0.4436	0.3724	0.3332	0.3084	0.2906	0.2761	0.2631
14	0.4263	0.3562	0.3203	0.2971	0.2788	0.2644	0.2523
15	0.4103	0.3419	0.3086	0.2853	0.2698	0.2555	0.2437
16	0.4012	0.3347	0.3019	0.2799	0.2636	0.2490	0.2370
17	0.3862	0.3245	0.2917	0.2713	0.2555	0.2421	0.2301
18	0.3813	0.3159	0.2823	0.2630	0.2476	0.2350	0.2245
19	0.3669	0.3057	0.2752	0.2555	0.2403	0.2276	0.2169
20	0.3553	0.2968	0.2672	0.2473	0.2332	0.2214	0.2113
21	0.3493	0.2918	0.2636	0.2445	0.2303	0.2179	0.2075
22	0.3401	0.2835	0.2567	0.2386	0.2245	0.2128	0.2030
23	0.3338	0.2789	0.2498	0.2317	0.2179	0.2069	0.1974
24	0.3255	0.2714	0.2435	0.2266	0.2136	0.2059	0.1933
25	0.3199	0.2652	0.2397	0.2226	0.2090	0.1984	0.1889
26	0.3146	0.2616	0.2350	0.2184	0.2060	0.1954	0.1864
27	0.3060	0.2564	0.2313	0.2156	0.2031	0.1925	0.1838
28	0.3029	0.2539	0.2286	0.2118	0.1996	0.1898	0.1811
29	0.2989	0.2481	0.2233	0.2073	0.1952	0.1849	0.1763
30	0.2914	0.2427	0.2187	0.2027	0.1908	0.1813	0.1732

*** UNIFORM TABLE BY MONTE CARLO 20000 TIMES REPEATED

Illustrative Examples

Example 1: Detection procedures for Model II: PN_2

Given a set of historical data: $X_1, \ldots, X_m$ we want to test the following PN situation:

$$PN_2: \mathcal{L} \in \Omega(URP)$$
$$N + S: \mathcal{L} \notin \Omega(URP)$$

Three statistics are used for two sets of simulated data:

(1) Kolmogorov-Smirnov

Decide $N + S$ iff $D_m^* > d_{m-1,\alpha}$ (from K-S table)

$$D_m^* = \sup_{0 \le u \le 1} \left| \frac{1}{m-1} \sum_{j=1}^{m-1} \varepsilon\,(u - U(j)) - u \right|$$

$$= \max_{1 \le j \le m-1} \left[\max \left\{ \frac{j}{m-1} - U(j), U(j) - \frac{j-1}{m-1} \right\} \right],$$

where $U(j) = \frac{X(j)}{X(m)}$.

(2) Srinivasan-type

Decide $N + S$ iff $\tilde{D}_m > \tilde{d} = (\frac{m-1}{m}) d_{m-1,\alpha}$

$$\tilde{D}_m = \sup_{0 \le u \le 1} (\frac{m-1}{m}) \left| \frac{1}{m-1} \sum_{j=1}^{m-1} \varepsilon\,(u - U(j)) - u \right| = (\frac{m-1}{m}) D_m^*$$

(3) Lilliefors-type

Decide $N + S$ iff $\hat{D}_m > d^*$ (from Table 1)

$$\hat{D}_m = \sup_{0 \le u \le 1} \left| \frac{1}{m} \sum_{j=1}^{m-1} \varepsilon\,(u - U(j)) - u \right|$$

$$= \max_{1 \le j \le m-1} \left[\max\left\{ \frac{j}{m} - U(j),\ U(j) - \frac{j-1}{m} \right\} \right]$$

The data sets are given in Table 2. The graphs showing inter-arrivals, waiting times and counting process for both processes are shown in diagrams 1 to 6. The computed values for the above three statistics are listed at the bottom of the Table 2.

<u>Conclusion</u>. For the both sets of data all detectors, at PFA $\alpha = .01$ decide in favor of PN_2.

<u>Note</u>. For m greater than 30 the exact Kolmogorov-Smirnov statistic asymptotic values are adequate to approximate $\hat{D}_m$ when $m > 30$. For example.

PFA α:	.01	.05	.10	.20
$\hat{D}_{m>30} \approx D_m$:	$\frac{1.63}{\sqrt{m}}$	$\frac{1.36}{\sqrt{m}}$	$\frac{1.22}{\sqrt{m}}$	$\frac{1.07}{\sqrt{m}}$

(2-sided test)

TABLE 2: Simulated Data for URP(θ) X_j = jth interarrival; $W_j = \sum_{i=1}^{j} X_i$ = jth waiting time.

m	URP(2.5)x		URP(10)x	
	(X_m)	(W_m)	(X_m)	(W_m)
1	0.723	0.723	6.239	6.239
2	1.462	2.186	8.044	14.282
3	2.258	4.443	1.373	15.656
4	2.217	6.660	5.163	20.818
5	1.242	7.902	6.336	27.154
6	0.991	8.893	7.403	34.557
7	2.180	11.072	7.020	41.577
8	2.328	13.400	5.071	46.548
9	1.489	14.889	7.485	54.134
10	0.524	15.413	5.620	59.754
11	2.220	17.633	3.081	62.835
12	1.160	18.794	4.854	67.690
13	1.570	20.364	8.760	76.449
14	1.117	21.481	8.006	84.455
15	1.145	22.625	0.426	84.881
16	0.587	23.212	8.975	93.856
17	1.909	25.121	8.545	102.401
18	1.478	26.599	8.542	110.943
19	1.110	27.709	6.887	117.830
20	1.238	28.947	5.015	122.845
21	2.305	31.251	2.012	124.857
22	0.525	31.776	1.844	126.701
23	2.080	33.856	2.978	129.679
24	0.382	34.238	8.582	138.261
25	2.116	36.354	1.215	139.476
26	1.695	38.049	2.189	141.665
27	2.315	40.364	8.128	149.794
28	0.904	41.268	9.536	159.330
29	0.632	41.900	5.358	164.688
30	0.703	42.604	5.770	170.458

Detection Statistic ($m = 30$)	Computed Value		Critical Value
	URP ($\theta = 2.5$)	URP ($\theta = 10$)	PFA: $\alpha = .01$
Kolmogorov-Smirnov: D_m	0.191	0.233	0.290
Srinivasan-type: $\tilde{D}_m$	0.184	0.225	0.2852 $[\tilde{D}_m = \frac{(m-1)}{m} D_m]$
Lilliefors-type: $\hat{D}_m$	0.194	0.242	0.2914

Example 2. Detection procedures for Model III: PN_3

Given two sets of data:

$$X_1, X_2, \ldots, X_m \text{ i.i.d. } U(0, \theta_1)$$

$$Y_1, Y_2, \ldots, Y_n \text{ i.i.d. } U(0, \theta_2).$$

Consider the following Model III-type situations:

Case 1.

$$PN_3': \ \mathcal{X}_1 = \mathcal{X}_2 \in \Omega(URP) \quad (\theta_1 = \theta_2)$$

$$N + S: \ \mathcal{X}_1 \text{ and } \mathcal{X}_2 \in \Omega(URP) \quad (\theta_1 < \theta_2)$$

Decide $N + S$ iff $X(m) \leq \alpha^{1/m}(\frac{N}{n})^{1/m} \max\{X(m), Y(n)\}$

Case 2.

$$PN_3^*: \ \mathcal{X}_1 = \mathcal{X}_2 \in \Omega(URP) \quad (\theta_1 = \theta_2)$$

$$N + S: \ \mathcal{X}_1 \text{ and } \mathcal{X}_2 \in \Omega(URP) \quad (\theta_1 > \theta_2)$$

Decide $N + S$ iff $Y(n) \leq \alpha^{1/n}(\frac{N}{m})^{1/n} \max\{X(m), Y(n)\}$

Case 3.

$$PN_3: \ \mathcal{X}_1 = \mathcal{X}_2 \in \Omega(URP) \quad (\theta_1 = \theta_2)$$

$$N + S \ \ \mathcal{X}_1 \neq \mathcal{X}_2 \ [\mathcal{X}_i \in \Omega(URP)] \quad (\theta_1 \neq \theta_2)$$

Decide $N + S$ iff

$$X(m) \leq \alpha^{1/m}(\frac{N}{n})^{1/m} \max\{X(m), Y(n)\}$$

or

$$Y(n) \leq \alpha^{1/n}\left(\frac{N}{m}\right)^{1/n} \max\{X(m),\ Y(n)\}$$

Note that if $n = m$

$$\alpha^{1/m}\left(\frac{N}{n}\right)^{1/m} \max\{X(m),\ Y(n)\} = \alpha^{1/n}\left(\frac{N}{m}\right)^{1/n} \max\{X(m),\ Y(n)\}$$

$$= \alpha^{1/n}\,(2)^{1/n} \max\{X(m),\ Y(n)\}$$

$$= (2\alpha)^{1/n} \max\{X(m),\ Y(n)\},$$

so we decide $N + S$ iff

$$\min\{X(m),\ Y(n)\} \leq (2\alpha)^{1/n} \max\{X(m),\ Y(n)\}$$

ie

$$\frac{\max\{X(m),\ Y(n)\}}{\min\{X(m),\ Y(n)\}} \geq (2\alpha)^{-1/n}\ .$$

Furthermore, notice that all the detection statistics developed for this problem namely T, T_1, T_2, T_3 are equivalent and SDF under the pure noise situation. One could have used any of the above detection procedures. In this example detection procedures T and T_3 are used. The generated data and the application of the above detection procedures is given below.

TABLE 3. Simulated Data for URP(θ)

Observation	$\theta_1 = 1$ (Set 1)	$\theta_2 = 2.5$ (Set 2)	$\theta_3 = 5$ (Set 3)	$\theta_4 = 10$ (Set 4)
1	0.580	0.723	1.892	6.239
2	0.951	1.462	0.467	8.044
3	0.786	2.258	1.563	1.373
4	0.298	2.217	3.330	5.163
5	0.454	1.242	3.252	6.336
6	0.006	0.991	2.551	7.403
7	0.276	2.180	0.919	7.020
8	0.306	2.328	1.932	5.071
9	0.689	1.489	0.581	7.485
10	0.383	0.524	4.259	5.620
11	0.133	2.220	3.551	3.081
12	0.832	1.160	3.611	4.854
13	0.583	1.570	3.207	8.760
14	0.099	1.117	0.079	8.006
15	0.277	1.145	0.358	0.426
16	0.620	0.587	4.770	8.975
17	0.084	1.909	2.115	8.545
18	0.990	1.478	1.937	8.542
19	0.979	1.110	3.021	6.887
20	0.694	1.238	4.485	5.015
21	0.934	2.305	0.382	2.012
22	0.212	0.525	2.880	1.844
23	0.131	2.080	4.540	2.978
24	0.863	0.382	2.860	8.582
25	0.819	2.116	0.608	1.215
26	0.541	1.695	3.922	2.189
27	0.019	2.314	2.945	8.128
28	0.314	0.904	1.395	9.536
29	0.765	0.632	2.263	5.358
30	0.942	0.703	1.166	5.770

Let $X_j(30)$ for $j = 1, 2, 3, 4$ denote the maximum value in the set j. Then $X_1(30) = 0.990$, $X_2(30) = 2.315$, $X_3(30) = 4.770$, and $X_4(30) = 0.536$. Choose PFA, $\alpha = .01$.

Decision Rules

Case 1.

For PN_3' versus $N + S$ decide $N + S$, if and only if,

$$X_1(30) \leq \alpha^{1/30} \left(\frac{60}{30}\right)^{1/30} \max\{X_1(30), X_2(30)\} =$$

$$= (2\alpha)^{1/30} (2.315) = 2.03;$$

so one decides in favor of $N + S$.

Case 2.

To detect PN_3' against $N + S$ in this situation decide $N + S$, if and only if,

$$X_2(30) \leq \alpha^{1/30} \left(\frac{60}{30}\right)^{1/30} \max\{X_3(30), X_2(30)\} =$$

$$= (2\alpha)^{1/30} (4.770) = 4.19$$

Thus, we conclude for $N + S$.

Case 3.

For detecting PN_3 against $S + N$ in this case one concludes $N + S$, if and only if,

$$T = \frac{\max\{X_3(30), X_4(30)\}}{\text{Min}\{X_3(30), X_4(30)\}} \geq (2\alpha)^{-1/30} = 1.14$$

Since $T = \frac{9.536}{4.770} = 1.99$, one concludes in favor of $N + S$.

Example 3. Epileptic Seizures and URP.

The situation below is very similar to a signal detection problem. First we need two definitions.

(a) Epilepsy: Chronic disorder of the central nevous (CNR) system characterized by recurrent (multiple) seizures which do not occur: (i) only during hospitalization for an acute systematic illness; or (ii) only in association with fever; or (iii) as a result of development or degenerative diseases of the central nervous system of CNR infection.

(b) Seizure: The clinical manifestation of abnormal paroxysmal discharges of neurons in the brain producing convulsive movements and/or sensory, vegetative or psychic dysfunction with or without loss of consiousness.

Epileptic seizures were measured on an epileptic female patient, whose age was 12 years, from 7:02 a.m. to 7:02 p.m. The beginning times and the duration of each seizure were recorded. The total number of seizures was 20 for 12 hours. We wish to decide whether the inter-arrival times have a uniform distribution. The data obtained were as follows:

TABLE 4. Epileptic Seizure Data

Start 7:02 a.m. Stop 7:02 p.m.

Seizure	Time	Duration	Inter-arrival time(min) (X_n)
1	9:16	0.2	134
2	9:16	0.5	0
3	10:06	0.3	50
4	10:16	0.1	10
5	10:52	0.2	36
6	10:54	0.2	2
7	10:58	0.2	4
8	11:04	0.2	6
9	11:14	0.2	10
10	11:20	0.2	6
11	11:46	0.2	26
12	13:06	0.1	90
13	13:15	0.5	9
14	13:16	0.2	1
15	13:46	0.1	30
16	17:54	0.2	248
17	18:20	0.2	26
18	18:25	0.2	5
19	18:46	1.0	21
20	18:48	0.2	2

Detection of the Model PN versus S + N

$PN(H_0)$: The inter-arrival time distribution is $U(0,\theta)$ for some $\theta > 0$, where $X_1, X_2, \ldots, X_{20}$ are the inter-arrival times of a renewal process and are i.i.d. r. vs.

N + S: Not as above.

Data: 134, 0, 50, 10, 36, 2, 4, 6, 10, 6, 26, 80, 9, 1, 30, 248, 26, 5, 21, 2 ($m = 20$). Choose PFA, $\alpha = .01$.

(1) The Lilliefors-type Detection procedure.

We reject PN, if and only if, $\hat{D}_m = \sup_x |F_m(x) - \hat{F}_\theta(x)| > 0.3553$, where "0.3553" is the .99th percentile of the $\hat{D}_m$ distribution in the uniform case for $m = 20$, and is obtained from Table 1.

In this case $\hat{\theta} = \max_{1 \le j \le 20} \{X_j\} = X_{(20)} = 248$; $\hat{F}_\theta(x) = x|\hat{\theta} = x(248)^{-1}$

where $0 < x < X_{(20)}$. The computations for $\hat{D}_{20}$ are given in the Table 5 below. Since $\hat{D}_{20} = 0.6548 > 0.3553$, we reject PN situation for a PFA of 0.01 that the data is uniform.

TABLE 5: Computations for the Lilliefors-type Detection Procedure, $\hat{D}_m$; Epileptic Seizure Data

i	$i/m = F_m(x)$	$X_{(i)}$	$\hat{F}_\theta(x_{(i)}) = \frac{x_{(i)}}{X_{(m)}}$	$\hat{F}_\theta(x_{(i)}) - \frac{i-1}{m}$	$\frac{i}{m} - \hat{F}_\theta(x_{(i)})$
1	.0500	0	.0000	.0000	.0500
2	.1000	1	.0040	-.0460	.0960
3	.1500	2	.0081	-.0460	.1419
4	.2000	2	.0081	-.1919	.1919
5	.2500	4	.0081	-.1419	.2339
6	.3000	5	.0161	-.1839	.2798
7	.3500	6	.0202	-.2298	.3258
8	.4000	6	.0242	-.2758	.3758
9	.4500	9	.0363	-.3258	.4137
10	.5000	10	.0403	-.3637	.4597
11	.5500	10	.0403	-.4597	.5097
12	.6000	21	.0847	-.4653	.5153
13	.6500	26	.1048	-.4952	.5452
14	.7000	26	.1048	-.5424	.5952
15	.7500	30	.1376	-.5624	.6124
16	.8000	36	.1452	-.6048	.6548
17	.8500	50	.2016	-.5984	.6484
18	.9000	80	.3226	-.5274	.5774
19	.9500	134	.5403	-.3597	.4097
20	1.0000	248	1.000	0.0500	0.0000

(2) <u>The Detection based on a M-S-N.</u>

In this case we reject PN, if and only if, $D_m^* > 0.3014$, where "0.3014" is the .95th percentile of the D_{m-1} distribution for $m = 20$ obtained from the standard Kolmogorov-Smirnov table. Here, a M-S-N is given by $\underset{\sim}{V}^*$, which is defined as

$$\underset{\sim}{V}^* = \left\{ \frac{X_{(1)}}{X_{(m)}}, \ldots, \frac{X_{(m-1)}}{X_{(m)}} \right\} \overset{d}{=} \{U_{(1)}, \ldots, U_{(m-1)}\},$$

where $U_1, \ldots, U_{m-1}$ are i.i.d. U(0,1) r. vs. For the Epileptic Seizure data with $m = 20$, one gets

$$\begin{aligned}\underset{\sim}{V}^* = \{&0, 0.0040, 0.0081, 0.0081, 0.0161, 0.0202,\\ &0.0242, 0.0242, 0.0363, 0.0403, 0.0403, 0.0847,\\ &0.1048, 0.1048, 0.1210, 0.1452, 0.2016, 0.3226,\\ &0.5403\}.\end{aligned}$$

The computations for D_{20}^* are given in Table 6 below, where $Y_i = U_{(i)}$, $i = 1, 2, \ldots, 19$.

TABLE 6. Computations for a M-S-N Detection Statistic D_m^*; Epileptic Seizure Data.

i	i/19	$F_\theta(y_i)=y_i$	$F_\theta(y_i)-\frac{i-1}{19}$	$\frac{i}{19}-F_\theta(y_i)$
1	.0526	0.0000	0.0000	0.0526
2	.1053	.0040	-.0486	.1013
3	.1579	.0081	-.0972	.1498
4	.2105	.0081	-.1498	.2024
5	.2632	.0161	-.1944	.2471
6	.3158	.0202	-.2430	.2956
7	.3684	.0242	-.2916	.3442
8	.4211	.0242	-.3442	.3969
9	.4737	.0363	-.3848	.4374
10	.5263	.0403	-.4334	.4860
11	.5789	.0403	-.4860	.5386
12	.6316	.0847	-.4960	.5469
13	.6842	.1048	-.5268	.5794
14	.7368	.1048	-.5768	.6320
15	.7895	.1210	-.6158	.6685
16	.8421	.1452	-.6443	.6969
17	.8947	.2016	-.6405	.6931
18	.9474	.3226	-.5721	.6248
19	1.0000	.5403	-.4071	.4597

Since $D_{20}^* = 0.6969 > 0.352$,

We reject the null hypothesis for a PFA level of 0.01 that the data is uniform.

Discussion of the Detection Results

The PN situation is rejected by all procedures by a "large margin". This seems to indicate that the inter-arrivals are decidedly non-uniform.

Diagram 1: URP(2.5); Interarrivals

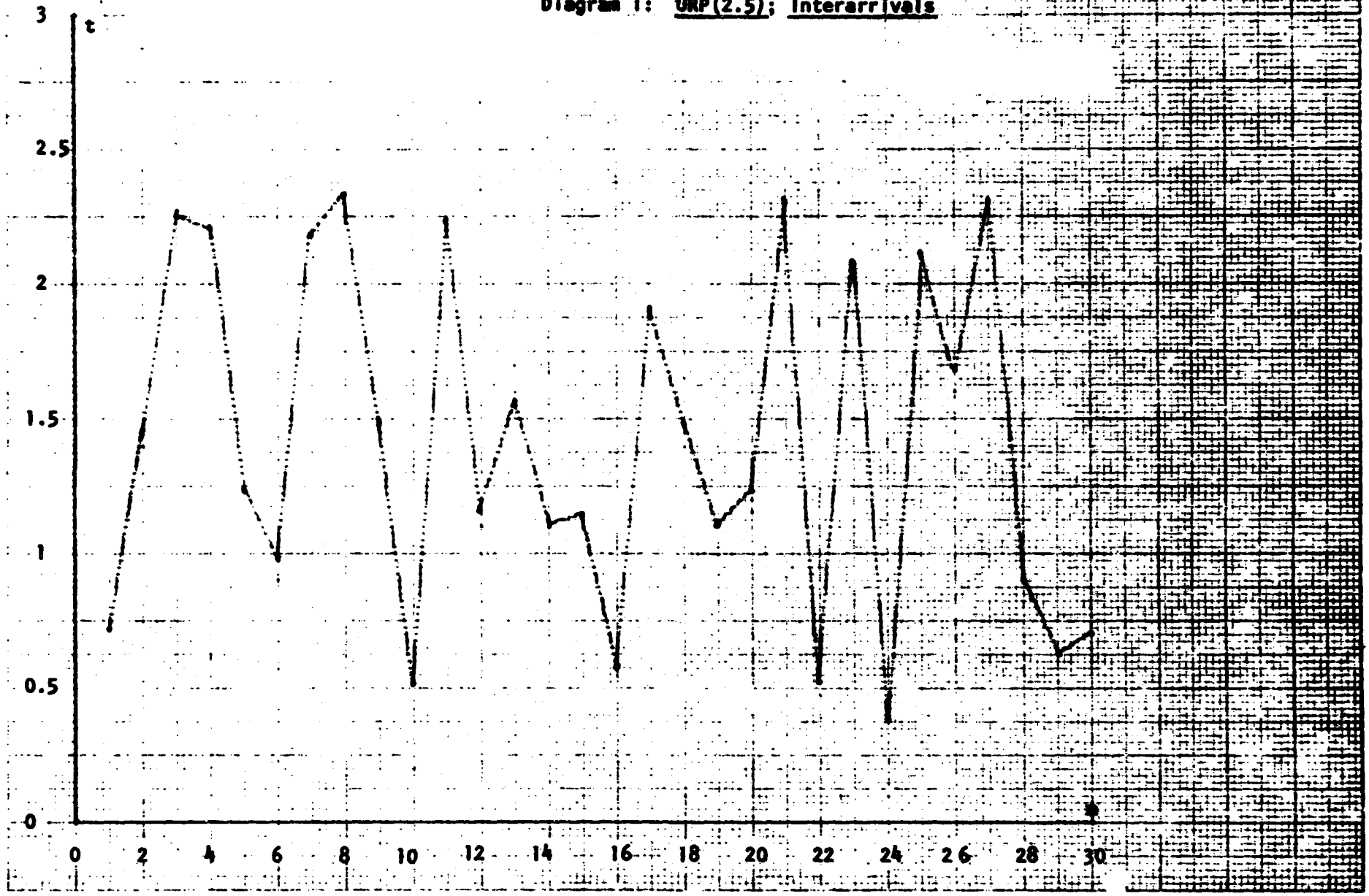
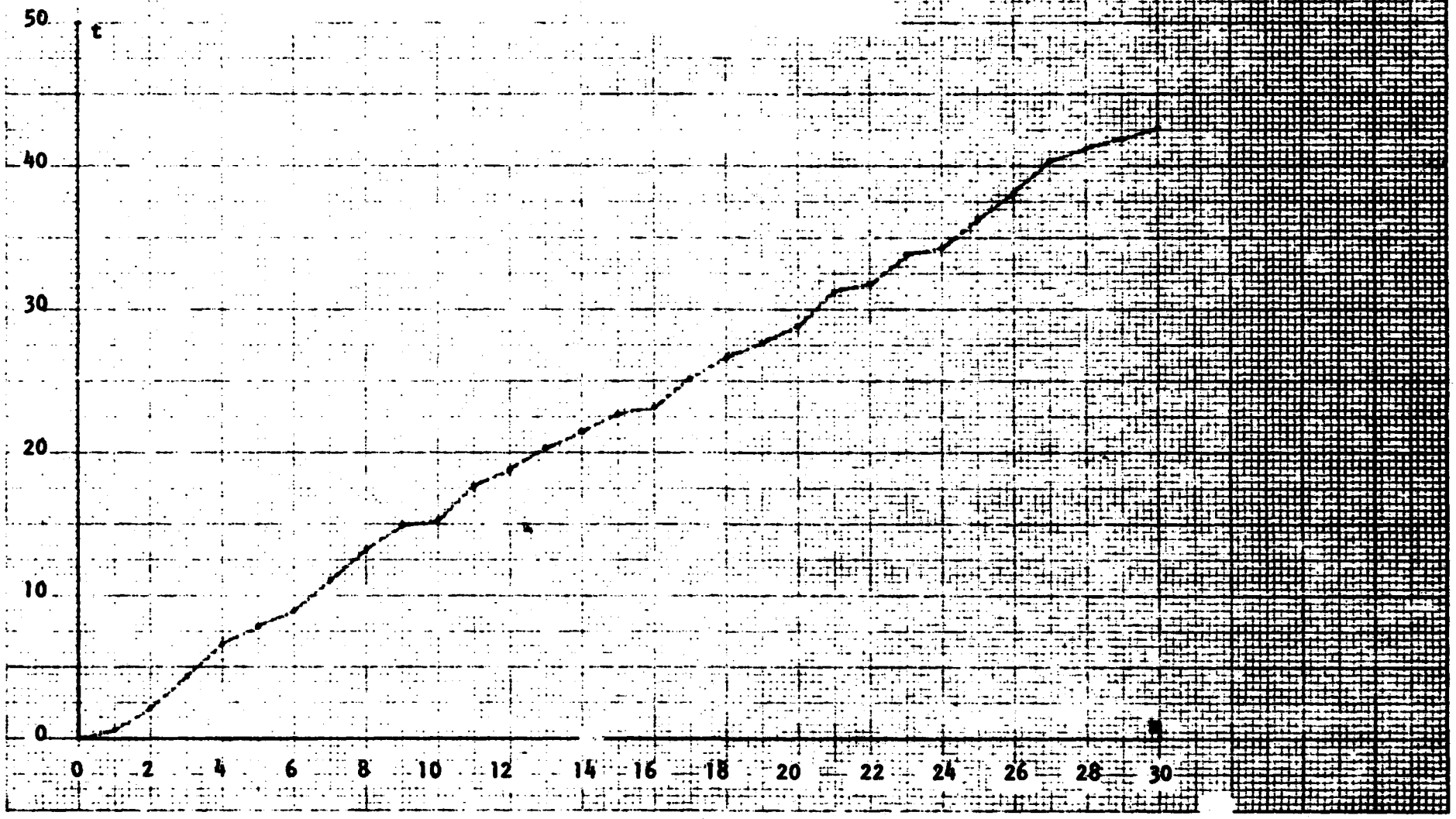


Diagram 2: URP(2,5): Waiting times



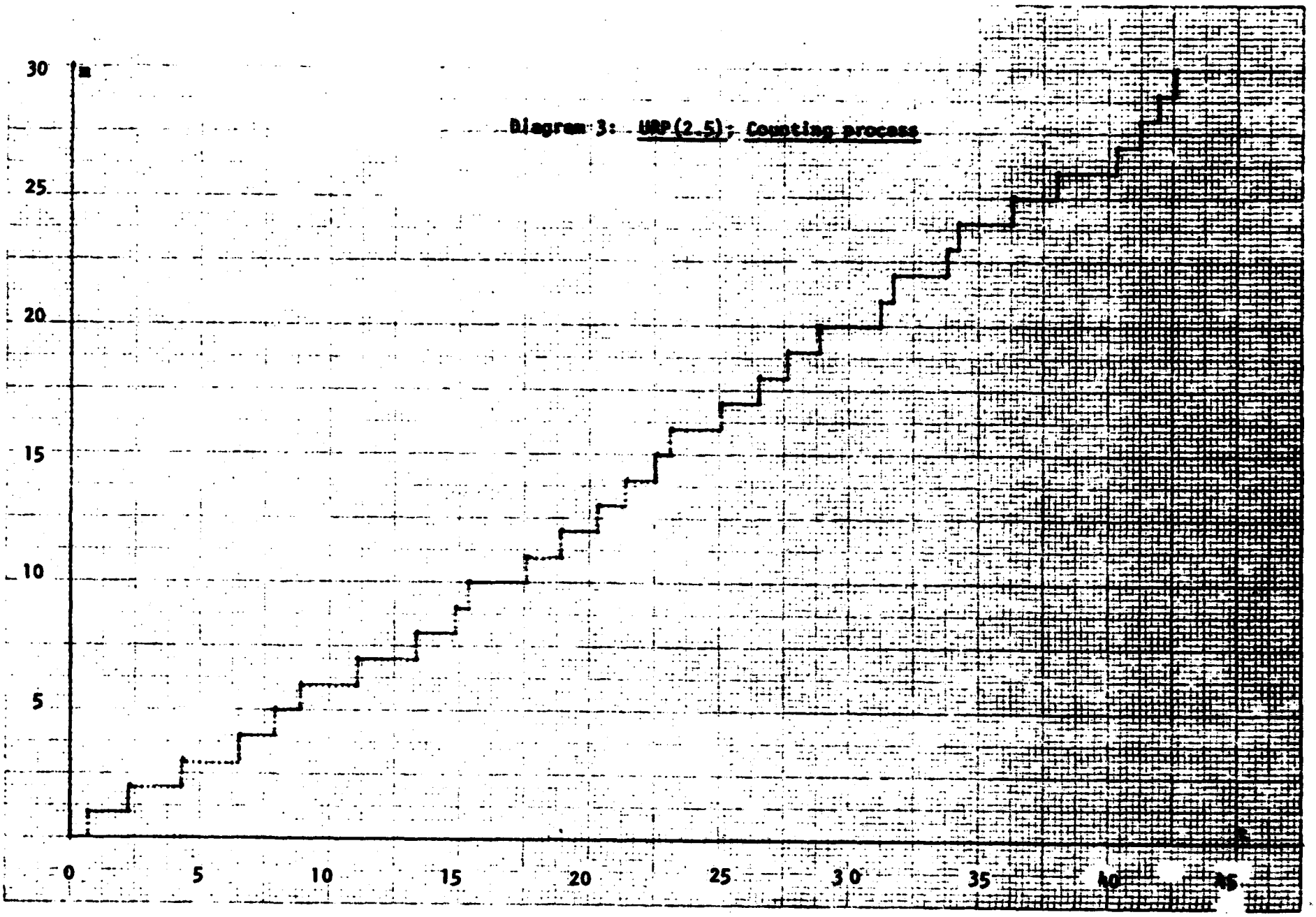


Diagram 3: HRP(2,5); Counting process

Diagram 4: UNP(10); Interarrivals

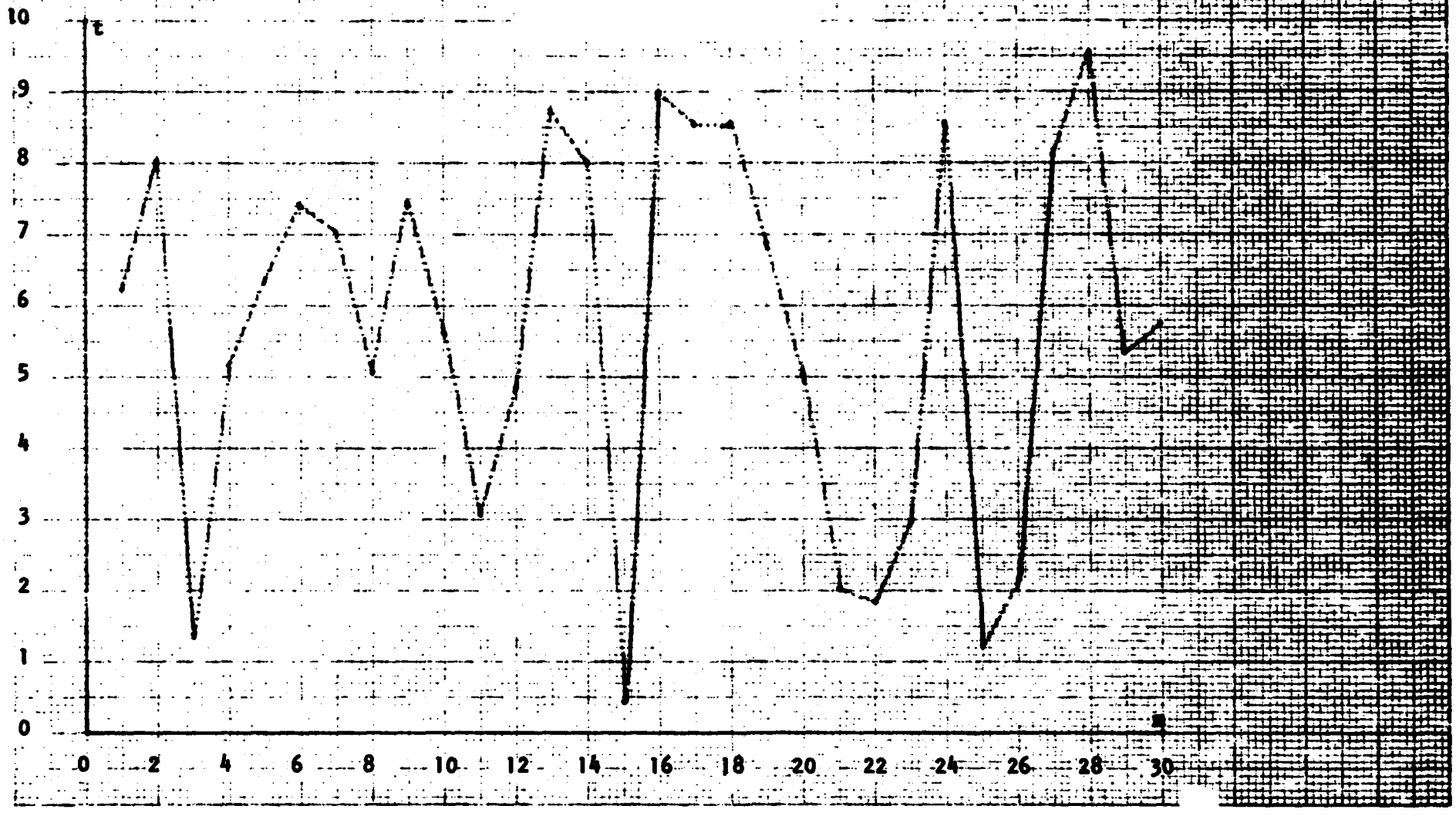
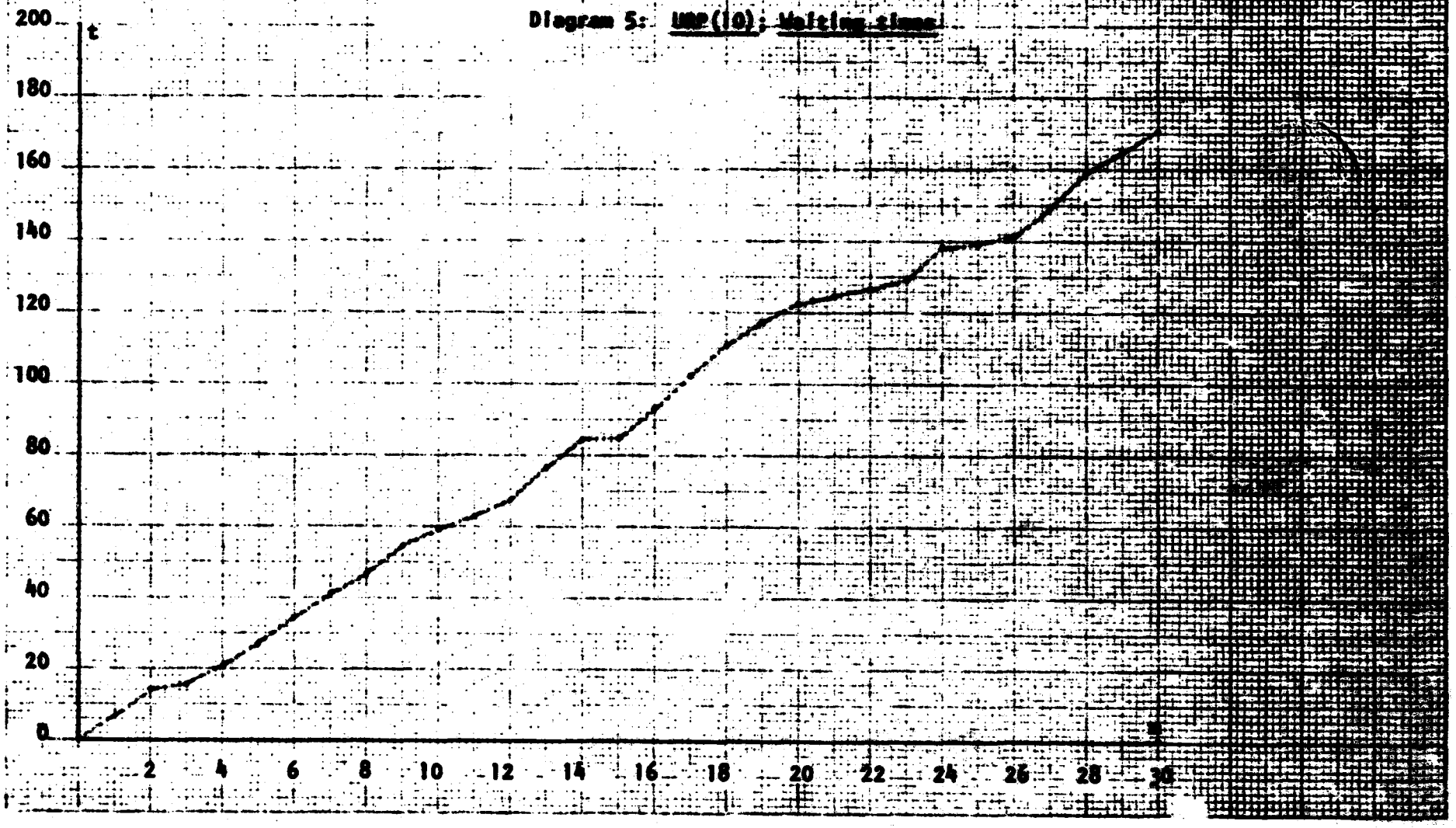


Diagram 5: UMP(10); Waiting times



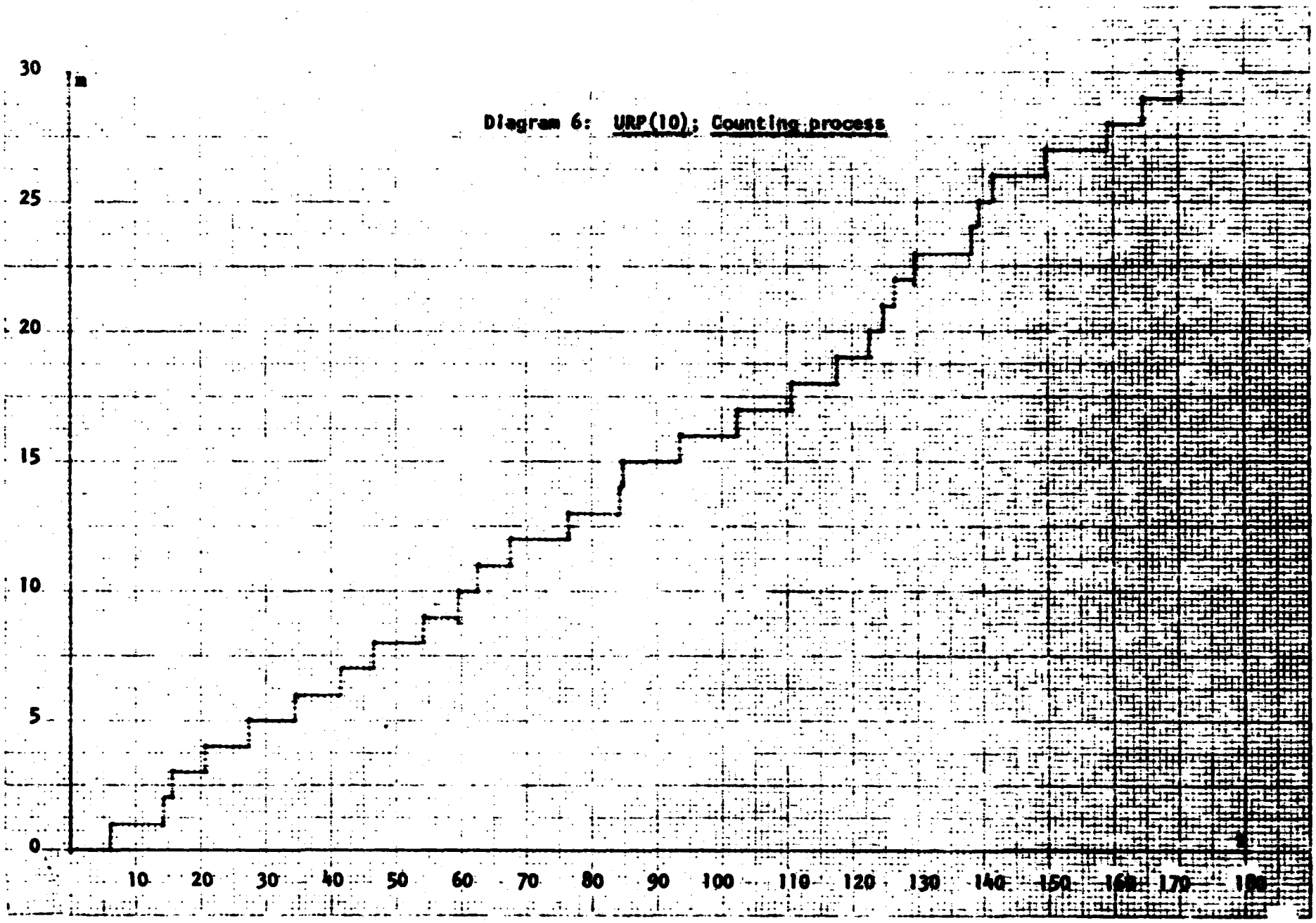


Diagram 6: URP(10); Counting process
Im
30
25
20
15
10
5
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180