Discussion Paper No. 24

POISSON PROCESS AND DISTRIBUTION-FREE STATISTICS, I

by

Meyer Dwass

November 20, 1972

1. INTRODUCTION

Suppose that $Y(t)$, $0 \leq t < \infty$, is a Poisson process with stationary increments and parameter λ. Let T equal the time t at which $Y(t) = t$ for the last time. That is

$$T = \sup\,(t\colon Y(t) = t)$$

The random variable T assumes the values $0,1,\ldots,\infty$. We have the following basic facts which are easily verified:

$$P(Y(t) < t,\ \text{all } t > 0) = P(T = 0) = \begin{cases} 1-\lambda, & \text{if } \lambda < 1 \\ 0, & \text{otherwise} \end{cases} \tag{1.1}$$

(See Appendix A)

$$P(T < \infty) = \begin{cases} 1 & \text{if } \lambda < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$P(T = k) = \frac{(\lambda k)^k e^{-\lambda k}}{k!}(1-\lambda),\ k=0,1,\ldots \quad \text{if } \lambda < 1 \tag{1.2}$$

Let U be a function of Y( ) with the following property:

The value of U is completely determined by the function $Y(t)$ for $0 \leq t \leq T$. (1.3)

Denote

$$G_n(t) = \frac{Y(tn)}{n}, \quad 0 \leq t \leq 1.$$

Subject to the condition that $T = n$, the behavior of $G_n(t)$ is exactly that to see that N_n is "distribution free". That is the distribution of N_n does not depend on the form of the cdf F, where $F(u) = P(X_i < u)$, as long as F is continuous. We assume here that $F(u) = u,\ 0 \leq u \leq 1$.)

It is obvious that N is geometrically distributed when $\lambda < 1$. Specifically,

$$P(N \geq k) = \lambda^k, \quad k = 0,1,\ldots \tag{1.7}$$

The counterpart of equation (1.4) is

$$P(N \geq k) = \sum_{n=0}^{\infty} P(N_n \geq k) \frac{n^n}{n!} (\lambda e^{-\lambda})^n (1 - \lambda)$$

It will be seen in Appendix C that

$$\frac{\lambda^k}{(1-\lambda)} = \sum_{n=0}^{\infty} \frac{n^{n-k}}{(n-k)!} (\lambda\ e^{-\lambda})^n , \quad k = 0,1,\ldots$$

(This follows from equation (C.1) with u = 0.)

Hence the counterpart of the relationship (1.4) is

$$P(N_n \geq k) = \frac{n^{n-1}}{(n-k)!} \cdot \frac{n!}{n^n} = \frac{n!}{(n-k)!\,n^k}$$

By a simple application of Stirling's formula, it follows that

$$\lim_{n\to\infty} P\left(\frac{N_n}{\sqrt{n}} \geq t\right) = e^{-t^2/2} \tag{1.8}$$

a result which was first proved by Smirnov [8].

We close this introduction by considering a second example.

Example 2 Let

L = number of ladder points of $Y(t) - t,\ t \geq 0$. That is L equals the number of times that $Y(t) - t$ achieves positive maxima which exceed all preceding maxima. (See Figure 1.)

Similarly we let L_n denote the number of ladder points in the empirical cdf F_n.[1/] The random variable L is also geometrically distributed with

$$P(L \geq k) = \lambda^k$$

Hence it follows that

$$P(N_n \geq k) = P(L_n \geq k) = \frac{n!}{(n-k)!n^k}$$

and

$$\lim_{n\to\infty} P\left(\frac{L_n}{\sqrt{n}} \geq t\right) = e^{-t^2/2}$$

2. ONE-SIDED MAXIMA. FIRST APPROACH

We define

$$M = \sup_{t>0} Y(t) - t$$

which is the maximum exceedance of the Poisson process Y(t) above the straight line t. Suppose that $\lambda < 1$. If $M \geq u$, (where u is an arbitrary positive number not necessarily an integer) this means that Y( ) intersects the straight line t + u, necessarily only a finite number of times. The probability that such an intersection last occurs at height n is

$$\frac{[\lambda(n-u)]^n}{n!} e^{-\lambda(n-u)} (1-\lambda).$$

Hence

$$P(M \geq u) = \sum_{n>u} \frac{(n-u)^n}{n!} (\lambda e^{-\lambda})^n e^{\lambda u}(1-\lambda) \qquad (2.1)$$

1/ From now on throughout this paper it will be assumed that F_n is the empirical cdf of n independent random variables, each uniformly distributed on [0,1].

$$P(M_n < u) = \sum_{\substack{i+j=n \\ i \le u}} u \binom{n}{i} \frac{(i-u)^i \; (j+u)^{j-i}}{n^n} \tag{2.5}$$

Equations (2.4), (2.5) are well-known. See [1], [2], [3], [4]. It is easy to give a direct proof of (2.4) which is analogous to that of (2.1) using the relation (B.1) in Appendix B. The argument goes as follows: $(M_n \geq u)$ means that $F_n(t)$ crosses the line $t + \frac{u}{n}$ somewhere. The probability of a last crossing at height i/n is the binomial probability

$$\binom{n}{i} \left(\frac{i}{n} - \frac{u}{n}\right)^i \left(1 - \frac{i}{n} + \frac{u}{n}\right)^{n-i}$$

times the conditional probability that the part of $F_n(t)$ traversing between $(\frac{i}{n} - \frac{u}{n}, \frac{i}{n})$ and $(1,1)$ stays below $t + \frac{u}{n}$. By (B.1), Appendix B this equals

$$\frac{1 - (1 - \frac{u}{n})}{1 - (\frac{i}{n} - \frac{u}{n})} = \frac{u/n}{(n-i+u)/n}$$

Hence

$$P(M_n \geq u) = \sum_{i>w} \binom{n}{i} \left(\frac{i}{n} - \frac{u}{n}\right)^i \left(1 - \frac{i}{n} + \frac{u}{n}\right)^{n-i-1} \frac{u}{n}$$

which is the same as (2.4). In Section 3 we will give a more intuitive explanation of (2.1) and (2.4).

3. ONE SIDED MAXIMA. SECOND APPROACH

We first need a technical preliminary about the Poisson process which is interesting in its own right. Define

$$J = \begin{cases} \text{value of } Y(t) - t \text{ at the first instant that } Y(t) > t \text{ if this takes place,} \\ 0, \text{ otherwise.} \end{cases}$$

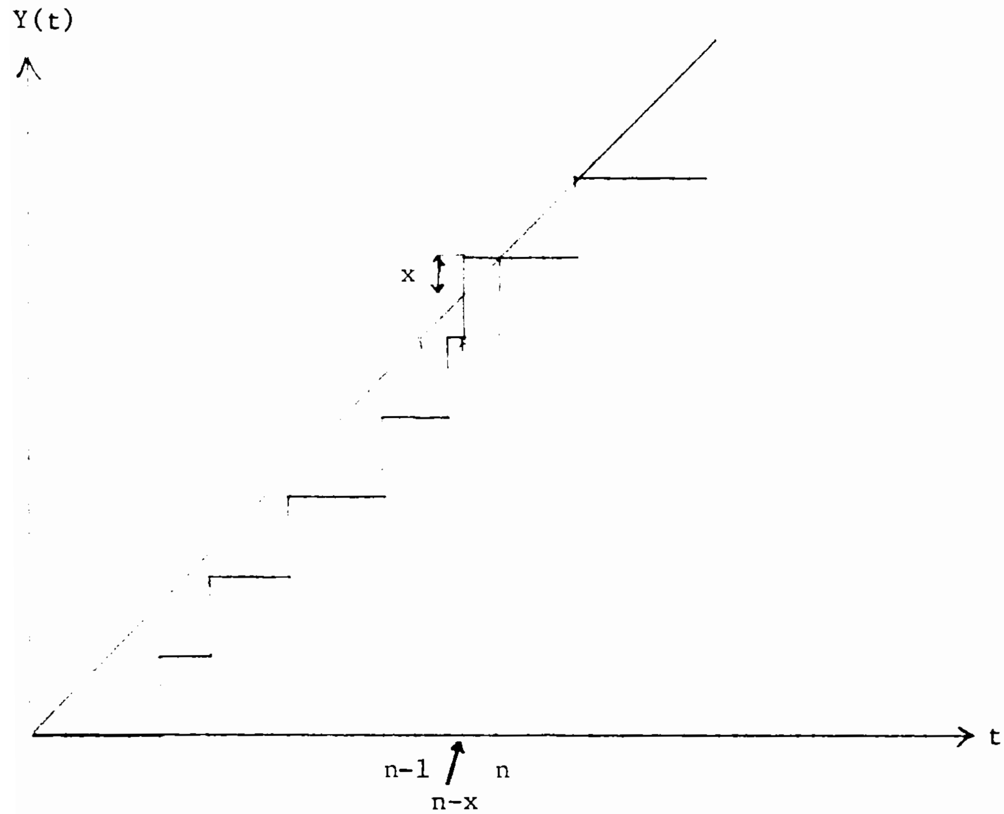

Figure 2

In other words, M is a sum of a geometric number of uniformly distributed random variables.

It is well-known that if $U_1, U_2, \ldots$ are independent random variables, each uniformly distributed on $[0,1]$, then

$$P\left(\sum_{1}^{n} U_i < u\right) = \sum_{i=0}^{n} \binom{n}{i}(-1)^i [c(u-i)]^n / n! \tag{3.3}$$

where

$$c(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x > 0 \end{cases}$$

Hence we can derive the distribution of M once again, using Theorem 3.3, by the calculation

$$P(M<u) = \sum_{n=0}^{\infty} \left\{ \sum_{i=0}^{n} \binom{n}{i} (i)^i \, [c(u-i)]^n / n! \right\} \lambda^n (1-\lambda)$$

$$\sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \sum_{n \geq i} \frac{[c(u-i)]^n}{(n-i)!} \lambda^n (1-\lambda)$$

$$= \sum_{i=0} \frac{(-1)^i}{c!} [(\lambda c(u-i)] e^{\lambda c(u-1)} (1-\lambda)$$

$$= \sum_{0 \leq i < u} \frac{(i-u)^i}{i!} (\lambda e^{-\lambda})^i e^{\lambda u} (1-\lambda)$$

which agrees with (2.2).

The Laplace transform of M is

$$Ee^{-\theta M} = \sum_{r=0}^{\infty} \left(\frac{1-e^{-\theta}}{\theta}\right)^r \lambda^r (1-\lambda)$$

$$= \frac{\theta(1-\lambda)}{\theta - \lambda(1-e^{-\theta})}$$

conditionally distributed as n independent random variables, each uniformly distributed on [0,1].

(b) M_n is distributed as

$$U_1 + \dots + U_{L_n}$$

where $U_1, U_2, \dots$ is a sequence of independent, uniform [0,1] random variables, independent of L_n.

It is now easy to determine the asymptotic distribution of M_n.

Theorem 3.5

$$\lim P\left(\frac{M_n}{\sqrt{n}} \geq t\right) = e^{-2t^2}$$

Proof:

$\frac{M_n}{\sqrt{n}}$ is distributed as

$$\frac{U_1 + \dots + U_{L_n}}{\sqrt{n}} = \frac{U_1 + \dots + U_{L_n}}{L_n} \frac{L_n}{\sqrt{n}}$$

Since L_n converges in probability to ∞ as $n \to \infty$ (Example 2, Introduction) it follows that $(U_1 + \dots + U_{L_n})/L_n$ converges in probability to $\frac{1}{2}$. Hence (by Example, 2, Introduction)

$$\lim P\left(\frac{M_n}{\sqrt{n}} \geq t\right) = \lim P\left(\frac{1}{2} \frac{L_n}{\sqrt{n}} \geq t\right) = e^{-(2t^2)/2}$$

which completes the proof. The above result was first proved by Kolmogorov [5].

Let us consider now the random variables $I_1, I_2, \dots$ defined earlier. This is the succession of excesses of $Y(t)$ over the line t. The counterparts for empirical cdf's is $I_1^{(n)}, I_2^{(n)}, \dots$ the succession of excesses of $n(F_n(t) - t)$

$$N(1) = N,\ L(1) = L,\ N_n(1) = N_n,\ L_n(1) = L_n.$$

It should be emphasized, however, that the random variable T still refers to last crossing of the line t (c.f. Section 1.). Equivalent to (1.1) is the relationship

$$P(Y(t) < \rho t \text{ all } t > o) = \begin{cases} 1 - \frac{\lambda}{\rho} \text{ if } \lambda < \rho \\ 0 \text{ otherwise.} \end{cases}$$

Hence $N(\rho)$ is geometrically distributed with

$$P(N(\rho) \geq k) = \left(\frac{\lambda}{\rho}\right)^k, \quad k = 0,1,\ldots$$

Since

$$P(N(\rho) \geq k) = P(N \geq k)/\rho^k$$

it follows that for $\rho > 1$,

$$P(N_n(\rho) \geq k) = P(N_n \geq k)/\rho^k = \frac{1}{\rho^k} \quad \frac{n!}{(n-k)!\ n^k} \quad k = 0,1,2,\ldots \tag{4.1}$$

Before going further, let us point out one interesting consequence of (4.1); namely,

$$P(N_n(\rho) \geq 1) = \frac{1}{\rho} \tag{4.2}$$

But this is just

$$1 - P(F_n(t) < \rho t, \text{ all } 0 < t \leq 1)$$

which we know equals $1/\rho$ by (B.1),Appendix B. Thus (4.1) generalizes that assertion.

We now see an asymptotic result for $N_n(\rho)$ by letting ρ vary with n, as follows.

To obtain a version of Theorem 3.4 for $M_n(\rho)$ we consider the random variables

$$I_1, I_2, \ldots, \qquad J_1, J_2, \ldots$$

$$I_1^{(n)}, I_2^{(n)}, \ldots, \quad J_1^{(n)}, J_2^{(n)}, \ldots$$

where the role of the line t is now played by ρt. In other words, $I_1, I_2, \ldots$ is the succession of excesses of $Y(t)$ over the line ρt; $J_1, J_2, \ldots$ is the succession of increases of $Y(t) - \rho t$ at the succession of ladder points of $Y(t) - \rho t$. (Strictly speaking the notation for the I_i's and J_i's should indicate their dependence on ρ but we suppress the ρ in the interest of typography.) Since when $\rho = 1$, the conditional distributions of the jumps as described in Theorems 3.2 and 3.4 do not depend on the actual value of λ as long as $\lambda < 1$, it follows similarly that the following is true:

<u>Theorem 4.1</u>

$$P((I_1, \ldots, I_k) \in A \mid N(\rho) \geq k) \qquad \text{(a)}$$

$$= P((J_1, \ldots, J_k) \in A \mid L(\rho) \geq k) \qquad \text{(b)}$$

$$= P((I_1^{(n)}, \ldots I_k^{(n)} \in A \mid N_n(\rho) \geq k) \qquad \text{(c)}$$

$$= P((J_1^{(n)}, \ldots, J_k^{(n)} \in A \mid L_n(\rho) \geq k) \qquad \text{(d)}$$

$$= P((U_1, \ldots, U_k) \in A)$$

where $U_1, \ldots, U_k$ are independent random variables each uniformly distributed on $[0,1]$. For (a) and (b) we need to assume that $\rho \geq \lambda$. For (c) and (d) we need $\rho \geq 1$.

It follows now, as in Theorem 3.4, that

<u>Lemma 5.1</u> Suppose $\lambda < 1$.

$$P((-r < Y(t) - t < s;\ 0 \leq t \leq T) \cup (T = 0))$$

$$= \frac{P(M < s)\ P\ (M < r)}{P(M < r + s)}$$

Proof: Define

$$R_1(r,s) = P(Y(t) \text{ hits } t\text{-}r \text{ and avoids } t + s \text{ en route})$$

$$R_2(r,s) = P(Y(t) \text{ hits } t + s \text{ and avoids } t\text{-}r \text{ en route})$$

Since $\lambda < 1$,

$$P(Y(t) \text{ hits } t - r) = 1 = R_1(r,s) + R_2(r,s)$$

By the regenerative nature of the Poisson process,

$$P((Y(t) \text{ hits either } t - r \text{ or } t + s \quad 0 \leq t \leq T) \cap (T > 0))$$

$$= R_2(r,s) + R_1(r,s)\ (1 - P(M < r)).$$

We also have

$$P(M < s) = R_1(r,s)\ P(M < r + s)$$

Hence,

$$P((\ r < Y(t) - t < s,\ 0 \leq t \leq T) \cup (T = 0))$$

$$= 1 - P((Y(t) \text{ hits either } t - r \text{ or } t + s,\ 0 \leq t \leq T) \cap (T > 0))$$

$$= 1 - R_2\ (r,s) + R_1(r,t)(1 - P(M < r))$$

$$= R_2(r,s)\ P\ (M < r)$$

$$= \frac{P(M < r)\ P\ (M < s)}{P(M < r + s)}$$

which completes the proof.

Substituting in the right side of (5.3) and cancelling out the $(1-\lambda)$'s gives (5.1) with $\lambda e^{-\lambda} = e^{-1}u$. Since $\lambda e^{-\lambda}/e^{-1}$ defines a 1-1 map of [0,1] onto itself this gives (5.1). Similarly, if we use (2.2) and substitute in (5.3) this gives (5.2).

From (5.2) it follows that the $Q_n(r,s)$ terms satisfy the following difference equations:

$$\sum_{\substack{i+j=n \\ j\le r+s}} Q_i(r,s)\,\frac{i^i}{i!}\,\frac{(j-r-s)^j}{j!} = \begin{cases} \displaystyle\sum_{\substack{i+j=n \\ i\le r \\ j\le s}} \frac{(i-r)^i}{i!}\,\frac{(j-s)^i}{j!} & \text{if } n \le r+s \\ 0 & \text{if } n > r+s \end{cases}$$

These can be used for iterative numerical computation of the $Q_n(r,s)$. For $r = s$ the above difference equations are similar to ones developed by other methods by Massey [6].

APPENDIX B

Let $F_n(t)$, $0 \leq t \leq 1$ be an empirical cdf of n independent random variables, each uniformly distributed over [0,1]. Then if $\gamma > 1$

$$P(F_n(t) < \gamma t \quad \text{all } t \text{ in } [0,1]) = 1 - \frac{1}{\gamma} \tag{B.1}$$

This is easily proved by induction on n as follows. Let U be the largest of the n points selected in [0,1]. Then U has the integrating density nu^{n-1} in [0,1]. By the induction hypothesis

$$P(F_n(t) < \gamma t \text{ all } t \mid U = u) = \begin{cases} [u - \frac{(n-1)}{n\gamma}]/u & \text{if } u > 1/\gamma \\ 0 & \text{otherwise} \end{cases}$$

Hence

$$P(F_n(t) < \gamma t \text{ all } t) = \int_0^1 (u - \frac{(n-1)}{n\gamma})\ (1/u)\ nu^{n-1}\, du = 1 - \frac{1}{\gamma}$$

Since (B.1) clearly holds for n = 1, this completes the proof. Another proof of (B.1) appears in Section 4 (See 4.1).

$$(n + u)^{n-k} = (n + u)^{n-k-1} (n - k + k + u)$$

Hence, from (C.1)

$$\lambda^k e^{\lambda u}/(1-\lambda) = \sum_{n\geq k+1} \frac{(n+u)^{n-k-1}}{(n-k-1)!} (\lambda e^{-\lambda})^n + (k + u) \sum_{n\geq k} \frac{(n+u)^{n-k-1}}{(n-k)!}(\lambda e^{-\lambda})^n$$

The first expression on the right can be evaluated by (C.1) to equal

$$(\lambda e^{-\lambda}) \lambda^k e^{\lambda(u+1)}/(1 - \lambda)$$

Hence considering the second expression on the right,

$$(k+u) \sum_{n\geq k} \frac{(n+u)^{n-k-1}}{(n-k)!}(\lambda e^{-\lambda})^n = (\lambda^k e^{\lambda u} - (\lambda e^{-\lambda}) \lambda^k e^{\lambda(u+1)})/(1-\lambda)$$

$$= \lambda^k e^{\lambda u}$$

which proves (C.2) also in the case that $u \geq 0$. It follows that the left sides of (C.1) and (C.2) have absolutely convergent power series expansions in u for all u. Hence from considerations of analyticity both sides of (C.1) and (C.2) are equal for all u.