

Outlier-prone Distributions and
the Maximum of a Poisson Process

Richard F. Green

Department of Statistics

CONTENTS

Introduction to Part 1	p. 1
Part 1: A Theorem about Outlier-prone Families of Distributions	p. 2
Introduction to Part 2	p. 4
Part 2: Outlier-prone and Outlier-resistant Distributions	p. 7
Introduction to Part 3	p. 24
Part 3: A Theorem on the Limiting Behavior of the Maximum of a Poisson Process	p. 27
Bibliography	p. 51

INTRODUCTION TO PART 3

The problem considered in the last part of this thesis was suggested in a conversation with Ted Cohn. When recordings are made in the retinal ganglion of cats whose eyes have been dark- or light-adapted seemingly random impulses are observed which may be considered as noise. Barlow and Levick (1969) have used the statistical properties of this noise to determine objective thresholds of visual perception. They look at the distribution of impulse numbers in fixed intervals and assume normality. They consider the threshold to be 2.88 standard deviations above the mean.

Since an animal is interested in a change of illumination at any time, not just during a fixed interval, I have considered the distribution of the maximum number of impulses in an interval of fixed length but which moves in time. I have assumed that noise is Poisson (Barlow and Levick found that this isn't true). My conclusion is that asymptotically the distribution of the maximum is the same as the distribution of one plus the maximum of $[\lambda T]$ independent and identically distributed Poisson variables with parameter λ , where λ is the intensity of the Poisson noise. Here T is the time during which my detector is operating. This result is a

precise version (asymptotically) of what Naus (1963) refers to as the Berg-Mack approximation (which was asserted only for $\lambda < .1$) which arises in a similar problem.

This other problem, considered by Naus in his thesis (1963) and by Silberstein (1945), Berg (1945) and Mack (1948) is to determine the probability of observing in an interval of length p at least n points out of N points distributed randomly on $(0, 1)$.

In his thesis Naus conjectures that for fixed Np as $p \rightarrow 0$ the average size of the maximum cluster gets very large and perhaps becomes infinite. That it does in fact become infinite is easy to see from my result.

Newell (1962) considers a problem equivalent to mine, namely the "distribution for the smallest distance between any pair of k^{th} nearest neighbor random points on a line." His approach is more general than mine since he considers general renewal processes while I confine myself to the Poisson process.

The main area for application of this idea is that already outlined by Naus, namely testing for clustering.

I conjecture that for points distributed randomly with intensity λ in an n -dimensional cube

with volume V the distribution of the maximum number of points in any unit cube with the same orientation as the larger cube is asymptotically equivalent to that of n plus the maximum of $\lfloor \lambda^{nV} \rfloor$ independent Poisson random variables with parameter λ .

One further point is not explicitly considered in the thesis. A generalization of the ballot theorem to three candidates is easy to obtain using the method used in proving the sublemma to Lemma 2 in this final part.

PART 3: A THEOREM ON THE LIMITING BEHAVIOR OF THE
MAXIMUM OF A POISSON PROCESS

Definition: For $\lambda > 0$, $t > 0$ let $N_\lambda(t)$ be the number of counts in a Poisson process during time $(0, t]$. That is, $N_\lambda(t + w) - N_\lambda(t)$ has poisson distribution with parameter $w\lambda$ for $w > 0$ and the numbers of counts over disjoint intervals are independent.

Definition: $Y_{\lambda, w}(t) = N_\lambda(t) - N_\lambda(t - w)$, $t \geq w$.

Definition: $Z_{\lambda, w}(T) = \max_{t \in [w, T]} Y_{\lambda, w}(t)$.

Notation: $Y_\lambda(t) = Y_{\lambda, 1}(t)$ and

$$Z_\lambda(T) = Z_{\lambda, 1}(T).$$

Definition: Let $U_i(\lambda)$ be i. i. d. $P(\lambda)$, $i = 1, 2, \dots$

That is, $P(U_i(\lambda) = k) = \frac{e^{-\lambda} \lambda^k}{k!}$.

Definition: $V_\lambda(T) = \max_{i=1, 2, \dots, [T]} U_i(\lambda)$.

Definition: $p_{k, \lambda}(T) = P(Z_\lambda(T) = k)$.

Definition: $q_{k, \lambda}(T) = P(V_\lambda(\lambda T) + 1 = k)$.

Theorem: $\sum_{k=0}^{\infty} |p_{k, \lambda}(T) - q_{k, \lambda}(T)| \xrightarrow{T \rightarrow \infty} 0$.

To prove this theorem I will use three lemmas.

Lemma 1: $p_{k, \lambda}(2) \sim \frac{e^{-\lambda} \lambda^k}{(k-1)!}$.

That is, $\frac{p_{k, \lambda}(2)}{\frac{e^{-\lambda} \lambda^k}{(k-1)!}} \xrightarrow{k \rightarrow \infty} 1$.

Lemma 2: Let $A_i^n = (\max_{t \in (i, i+1]} Y_\lambda(t) \geq n)$.

For any natural number i : $\frac{P(A_i^n A_{i+1}^n)}{P(A_i^n)} \xrightarrow{n \rightarrow \infty} 0$.

Lemma 3: If

1. $A_1^k, A_2^k, \dots, A_n^k$ are events such that event A_i^k is independent of any set of events not involving $A_{i-1}^k, A_i^k, A_{i+1}^k$ for all $i = 1, 2, \dots, n$.
2. $P(A_1^k) = P(A_2^k) = \dots = P(A_n^k) = \beta_k$.
3. $P(A_i^k A_{i+1}^k) = \alpha_k \beta_k$ for all $i = 1, 2, \dots, n-1$.

then

$$\left| P\left(\bigcup_{i=1}^n A_i^k\right) - (1 - (1 - P(A_1^k))^n) \right|$$

$$\leq e^{\beta_k n} \left(\exp\left(\frac{\alpha_k' + \beta_k}{1 - \beta_k} \beta_k n\right) - 1 \right)$$

where $\alpha_k' = \max(\alpha_k, \beta_k)$.

Proof of Lemma 1: I will find an expression for $p_{k,\lambda}(2) = P(\max_{t \in [1,2]} Y_\lambda(t) = k)$ and look at what happens to it as k becomes large. To find $p_{k,\lambda}(2)$ I will condition successively on $Y_{\lambda,2}(2)$ and the pair $(Y_\lambda(1), Y_\lambda(2))$. $Y_{\lambda,2}(2)$ has a Poisson distribution with parameter 2λ . $Y_\lambda(1)$ given $Y_{\lambda,2}(2)$ will have a binomial distribution with $p = \frac{1}{2}$ and $n = Y_{\lambda,2}(2)$.

Using the reflection principle for random walks it

follows that $P(\max_{t \in [1,2]} Y_\lambda(t) = k \mid Y_{\lambda,2}(2) = n, Y_\lambda(1) = m)$

$$= \frac{\binom{n}{k} - \binom{n}{k+1}}{\binom{n}{m}} \text{ for } n \in (k, k+1, \dots, 2k),$$

$$m \in (n-k, n-k+1, \dots, k).$$

$$p_{k,\lambda}(2) = \sum_{n=k}^{2k} \sum_{m=n-k}^k \frac{e^{-2\lambda} (2\lambda)^n}{n!} \binom{n}{m} \left(\frac{1}{2}\right)^n \frac{\binom{n}{k} - \binom{n}{k+1}}{\binom{n}{m}}$$

$$= \sum_{n=k}^{2k} \frac{e^{-2\lambda} \lambda^n}{n!} \sum_{m=n-k}^k \left(\binom{n}{k} - \binom{n}{k+1} \right).$$

But $\sum_{m=n-k}^k \left(\binom{n}{k} - \binom{n}{k+1} \right) = \frac{(2k-n+1)^2 n!}{(k+1)!(n-k)!}$, so

$$p_{k,\lambda}(2) = \sum_{n=k}^{2k} \frac{e^{-2\lambda} \lambda^n}{(k+1)!} \frac{(2k-n+1)^2}{(n-k)!}$$

$$= \frac{e^{-\lambda} \lambda^k}{(k-1)!} \frac{e^{-\lambda}}{k(k+1)} \sum_{i=0}^k \frac{(k+1-i)^2 \lambda^i}{i!}.$$

But $\frac{e^{-\lambda}}{k(k+1)} \sum_{i=0}^k \frac{(k+1-i)^2 \lambda^i}{i!} \xrightarrow{k \rightarrow \infty} 1$.

Therefore $p_{k,\lambda}(2) \sim \frac{e^{-\lambda} \lambda^k}{(k-1)!}$

This completes the proof of Lemma 1.

Proof of Lemma 2: First I will prove the following

Sublemma: $P(A_1^n A_2^n \mid Y_\lambda(1) = i, Y_\lambda(2) = j, Y_\lambda(3) = k)$

$$\leq P(A_1^n \mid Y_\lambda(1) = i, Y_\lambda(2) = j) P(A_2^n \mid Y_\lambda(2) = j, Y_\lambda(3) = k).$$

Write this as:

$$P(A_1^n A_2^n \mid i, j, k) \leq P(A_1^n \mid i, j) P(A_2^n \mid j, k).$$

Proof of Sublemma:

$$\text{For } i \geq n, P(A_1^n A_2^n \mid i, j, k) = P(A_2^n \mid j, k).$$

$$\text{For } k \geq n, P(A_1^n A_2^n \mid i, j, k) = P(A_1^n \mid i, j).$$

$$\text{For } j \geq n, P(A_1^n \mid i, j) = P(A_2^n \mid j, k) =$$

$$= P(A_1^n A_2^n \mid i, j, k) = 1.$$

$$\text{For } i+j < n \text{ or } j+k < n, P(A_1^n A_2^n \mid i, j, k) = 0.$$

Therefore we need only consider

$$i, j, k < n, i+j \geq n, j+k \geq n.$$

The conditions

$$Y_\lambda(1) = i, Y_\lambda(2) = j, Y_\lambda(3) = k$$

mean that i points are distributed uniformly and independently in $(0,1]$, j points are distributed uniformly and independently in $(1,2]$, and k points are distributed uniformly and independently in $(2,3]$ respectively.

I propose to use the following trick. Label each of the i points in $(0,1]$ with a letter L (for "left"), label each of the j points in $(1,2]$ with a letter C (for "center"), and the k points in $(2,3]$ with a letter R (for "right"). Now shift each of the L labelled points one unit to the right and each of the R labelled points one unit to the left. Now

we have $i+j+k$ points in the interval $(1,2]$, i labelled L, j labelled C, and k labelled R. All $i+j+k$ points are uniformly and independently distributed.

Now we consider starting at the point 1 and moving toward 2, noticing the label on each successive point encountered. This procedure will produce a random sequence of i L's, j C's and k R's. If, reading this sequence from the left, at some point we have read at least $n-j$ more R's than C's then event A_2^n will have occurred. If at some point $n-i$ more C's than L's have been read then A_1^n will have occurred.

We know from the reflection principle for a one-dimensional random walk that

$$P(A_1^n \mid i, j) = \frac{\binom{i+j}{n}}{\binom{i+j}{i}}; \quad P(A_2^n \mid j, k) = \frac{\binom{j+k}{n}}{\binom{j+k}{j}}.$$

Now I will represent the random sequence of L's, C's and R's as a walk where I make the following identification:

- L: move one unit in (compass) direction 0° ,
- R: move one unit in direction 120° , and
- C: move one unit in direction 240° .

The event A_2^n corresponds to this walk hitting the vertical line r passing through a point $n-j$ steps toward R from the origin. The event A_1^n corresponds to this walk hitting the 120° - 300° line l passing

through a point $n-i$ steps toward C from the origin.
(See diagram, next page.)

The value $P(A_1^n A_2^n | i, j, k)$ is the ratio of the number of ways of "walking" from the origin to the finishing point and hitting both r and l en route to the total number of ways of going from the origin to the finishing point. This total number is the multinomial coefficient:

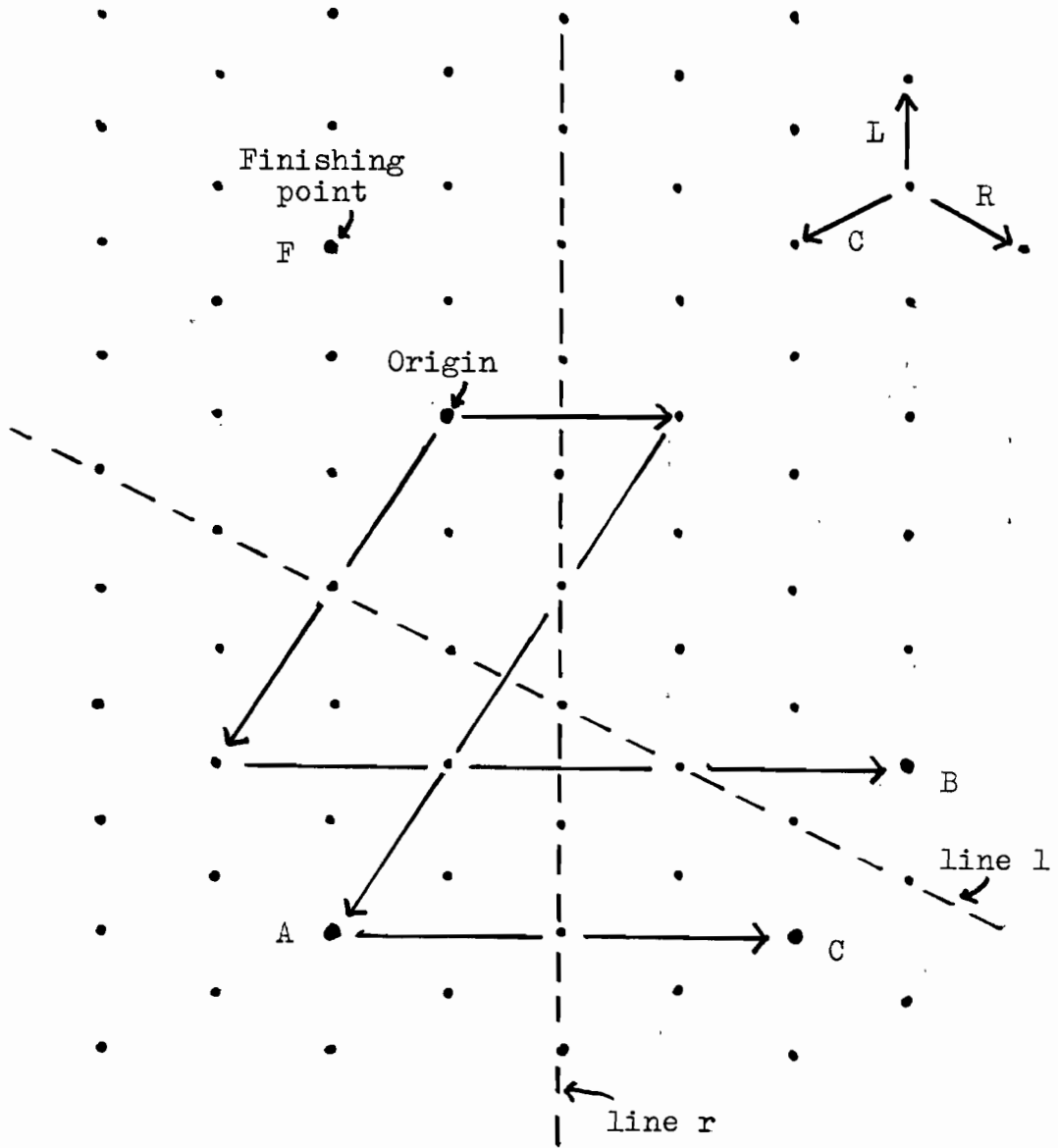
$$\binom{i+j+k}{i, j, k}.$$

Symmetry considerations show that the number of ways of getting from the origin to the finishing point while hitting r and l en route equals the number of ways of getting from the point obtained by reflecting the origin about r and then reflecting this point about l to the finishing point plus the number of ways of getting from the point obtained by reflecting the origin about l and then reflecting this point about r to the finishing point minus the number of ways of getting from the point obtained by reflecting the origin about r , then reflecting this point about l , and then reflecting this point about r to the finishing point (all these new walks consist of $i+j+k$ steps). Using multinomial coefficients this number is

$$\binom{i+j+k}{2n-j, i+j-n, j+k-n} + \binom{i+j+k}{n, n, i+j+k-2n} - \binom{i+j+k}{2n-j, j, i+j+k-2n}.$$

Diagram.

The desired number N_n is the sum of the number of ways of getting from A to F and the number of ways of getting from B to F minus the number of ways of getting from C to F. Each of these walks will consist of $i+j+k$ steps.



Call this number N_n . The sublemma asserts that

$$* \quad \frac{N_n}{\binom{i+j+k}{i, j, k}} \leq \frac{\binom{i+j}{n}}{\binom{i+j}{i}} \frac{\binom{j+k}{n}}{\binom{j+k}{j}}.$$

It may happen that $i+j+k-2n < 0$. In this case

$$N_n = \binom{i+j+k}{2n-j, i+j-n, j+k-n}$$

and we have

$$\begin{aligned} \frac{N_n}{\binom{i+j+k}{i, j, k}} &= \frac{i!j!k!}{(2n-j)!(i+j-n)!(j+k-n)!} \\ &\leq \frac{i!j!j!k!}{n!n!(i+j-n)!(j+k-n)!} = \frac{\binom{i+j}{n}}{\binom{i+j}{i}} \frac{\binom{j+k}{n}}{\binom{j+k}{j}} \end{aligned}$$

since $n!n! \leq j!(2n-j)!$

Thus we may assume $i+j+k-2n \geq 0$. Assertion * is equivalent to

$$\begin{aligned} &\frac{i!j!k!}{(2n-j)!(i+j-n)!(j+k-n)!} + \frac{i!j!k!}{n!n!(i+j+k-2n)!} - \\ &- \frac{i!j!k!}{(2n-j)!j!(i+j+k-2n)!} \leq \frac{i!j!j!k!}{n!n!(i+j-n)!(j+k-n)!}. \end{aligned}$$

Or, equivalently

$$\begin{aligned} &\frac{1}{(2n-j)!(i+j-n)!(j+k-n)!} - \frac{1}{(2n-j)!j!(i+j+k-2n)!} + \\ &+ \frac{1}{n!n!(i+j+k-2n)!} - \frac{j!}{n!n!(i+j-n)!(j+k-n)!} \leq 0, \end{aligned}$$

$$\begin{aligned} \text{or } &\frac{1}{(2n-j)!} \left(\frac{1}{(i+j-n)!(j+k-n)!} - \frac{1}{j!(i+j+k-2n)!} \right) - \\ &- \frac{j!}{n!n!} \left(\frac{1}{(i+j-n)!(j+k-n)!} - \frac{1}{j!(i+j+k-2n)!} \right) \leq 0, \end{aligned}$$

$$\text{or } \left(\frac{1}{(2n-j)!} - \frac{j!}{n!n!} \right) \left(\frac{1}{(i+j-n)!(j+k-n)!} - \frac{1}{j!(i+j+k-2n)!} \right) \leq 0,$$

$$\text{or } \frac{n!n! - j!(2n-j)!}{n!n!(2n-j)!} \left(\frac{j!(i+j+k-2n)! - (i+j-n)!(j+k-n)!}{(i+j-n)!(j+k-n)!j!(i+j+k-2n)!} \right) \leq 0.$$

But $n!n! \leq j!(2n-j)!$ and

$$j!(i+j+k-2n)! \geq (i+j-n)!(j+k-n)! \text{ since}$$

$$2n-i-k \geq |i-k| \text{ with our restrictions on } i, k \text{ and } n.$$

Thus we have verified * and our Sublemma is

proved.

Now I return to the proof of Lemma 2.

$$P(A_1^n A_2^n) \leq \sum_{\substack{i+j \geq n \\ j+k \geq n \\ 0 \leq i, j, k \leq n}} \frac{e^{-3\lambda} \lambda^{i+j+k}}{i!j!k!} P(A_1^n | i, j) P(A_2^n | j, k) \\ + P(i > n) + P(j > n) + P(k > n).$$

$$P(i > n) = P(j > n) = P(k > n) = \sum_{k=n+1}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!}.$$

$$\text{But } \sum_{k=n+1}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \text{ is } o(P(A_1^n)).$$

Thus we need only sum up to $i, j, k = n$.

$$P(A_1^n A_2^n) \leq \sum_{\substack{i+j \geq n \\ j+k \geq n \\ 0 \leq i, j, k \leq n}} \frac{e^{-3\lambda} \lambda^{i+j+k}}{i!j!k!} \frac{\binom{i+j}{n}}{\binom{i+j}{i}} \frac{\binom{j+k}{n}}{\binom{j+k}{j}} + o(P(A_1^n)).$$

The first term on the right above equals

$$\sum_{\substack{i+j \geq n \\ j+k \geq n \\ 0 \leq i, j, k \leq n}} \frac{e^{-3\lambda} \lambda^{i+j+k}}{n!n!} \frac{j!}{(i+j-n)!(j+k-n)!},$$

fix j, k and sum over $i = n-j, n-j+1, \dots, n$ using

$$r = i - (n-j),$$

$$= \sum_{\substack{j+k \geq n \\ 0 \leq j, k \leq n}} \frac{e^{-3\lambda} \lambda^{n+k} j!}{n! n! (j+k-n)!} \sum_{r=0}^j \frac{\lambda^r}{r!},$$

fix j , sum over $k = n-j, n-j+1, \dots, n$ using
 $s = k - (n-j)$,

$$\begin{aligned} &= \sum_{0 \leq j \leq n} \frac{e^{-3\lambda} \lambda^{2n-j} j!}{n! n!} \sum_{s=0}^j \frac{\lambda^s}{s!} \sum_{r=0}^j \frac{\lambda^r}{r!} \\ &\leq \sum_{0 \leq j \leq n} \frac{e^{-\lambda} \lambda^{2n-j} j!}{n! n!} \\ &= \frac{e^{-\lambda}}{n!} \left(\frac{\lambda^n n!}{n!} + \frac{\lambda^{n+1} (n-1)!}{n!} + \dots + \frac{\lambda^{2n} 0!}{n!} \right) \\ &= \frac{e^{-\lambda} \lambda^n}{n!} \left(1 + \frac{\lambda}{n} + \frac{\lambda^2}{(n)_2} + \dots + \frac{\lambda^n}{n!} \right) \\ &= o(P(A_1^n)) \text{ since by Lemma 1 } P(A_1^n) \sim \frac{e^{-\lambda} \lambda^n}{(n-1)!}. \end{aligned}$$

Thus we have

$$P(A_1^n A_2^n) \text{ is } o(P(A_1^n)) \text{ or}$$

$$\frac{P(A_1^n A_2^n)}{P(A_1^n)} \xrightarrow{n \rightarrow \infty} 0.$$

But by definition of $N_\lambda(t)$ we have

$$\frac{P(A_i^n A_{i+1}^n)}{P(A_i^n)} = \frac{P(A_1^n A_2^n)}{P(A_1^n)} \text{ for } i = 1, 2, \dots, n-1.$$

This completes the proof of Lemma 2.

Proof of Lemma 3: I will expand $P(\bigcup_{i=1}^n A_i^k)$

and $1 - (1 - P(A_1^k))^n$ and compare the terms of corresponding order.

$$* \quad P\left(\bigcup_{i=1}^n A_i^k\right) = \sum_{i=1}^n P(A_i^k) - \sum_{i < j} P(A_i^k A_j^k) + \\ + \sum_{i < j < l} P(A_i^k A_j^k A_l^k) - \dots$$

$$** \quad 1 - (1 - P(A_1^k))^n = nP(A_1^k) - \binom{n}{2}P(A_1^k)^2 \\ + \binom{n}{3}P(A_1^k)^3 - \dots$$

I think of the event $A_{i_1}^k A_{i_2}^k \dots A_{i_r}^k$ as consisting of isolated singles, doubles, triples, etc. For instance, the event $A_2^k A_5^k A_7^k A_8^k A_{10}^k A_{11}^k A_{12}^k A_{14}^k$ consists of the three singles A_2^k , A_5^k and A_{14}^k , the one double $A_7^k A_8^k$ and the one triple $A_{10}^k A_{11}^k A_{12}^k$. The probability of any particular single is β_k in both * and **. The probability of any particular double is $\alpha_k \beta_k$ in * and (corresponds to the probability) β_k^2 in **. The probability of any particular triple is $\leq \beta_k^2$ in * and $= \beta_k^3$ in **. In general a 2r-tuple has probability $\leq \alpha_k \beta_k^r$ in * and $= \beta_k^{2r}$ in **. A 2r+1-tuple has probability $\leq \beta_k^{r+1}$ in * and $= \beta_k^{2r+1}$ in **. Thus the difference between the probabilities assigned to any particular 2r-tuple by the expressions * and ** is $\leq \alpha_k \beta_k^r$. The difference for any particular 2r+1-tuple is $\leq \beta_k^{r+1}$.

I will characterize any event $A_{i_1}^k A_{i_2}^k \dots A_{i_r}^k$

by a non-increasing sequence of natural numbers representing its singles, doubles, triples, etc. For instance, (3, 2, 1, 1, 1) will represent an event having one triple, one double and three singles. Such an event will have probabilities under * and ** differing by no more than

$$\beta_k^2 \alpha_k' \beta_k \beta_k \beta_k \beta_k = \alpha_k' \beta_k^6.$$

for the triple) for the three singles
for the double

There will be no more than $\frac{n^5}{3!}$ ($> (1, 1, 3)$) distinct events characterized by this sequence. Thus the difference between * and ** contributed by such (3, 2, 1, 1, 1) events will be less than

$$\frac{\alpha_k' \beta_k^{n^5}}{3!}.$$

Now I will obtain the desired bound on the difference between * and ** by summing the (absolute) differences for each sequence of natural numbers. First sum over all possible numbers of singles. (Clearly the sequence shouldn't start with a 1 because this would mean the sequence was all 1's and the indicated events would have the same probabilities under * and **.) For instance, consider sequences of the form

$$(5, 3, 3, 2, \underbrace{1, 1, \dots, 1})$$

i 1's where $i = 0, 1, 2, \dots$

For the 5, 3, 3, 2 we have an upper bound on the difference = $\alpha'_k \beta_k^8 n^4 / 2$. Now, when we consider the singles we must multiply by the factor by which the probability of an individual event is reduced (β_k) and the number of ways the single can be located.

For i (# of 1's) = 0 multiply by ≤ 1

$$= 1 \quad " \quad " \leq \beta_k n$$

$$= 2 \quad " \quad " \leq \beta_k^2 n^2 / 2$$

$$= 3 \quad " \quad " \leq \beta_k^3 n^3 / 3!$$

etc.

Sum and get an upper bound. In the above example:

$$\alpha'_k \beta_k^8 \frac{n^4}{2} \sum_{i=0}^{\infty} \frac{(\beta_k n)^i}{i!} = \alpha'_k \beta_k^8 \frac{n^4}{2} e^{\beta_k n}.$$

Now sum over the doubles: consider two cases:

(1) No triples or higher -tuples,

(2) Higher -tuples.

Case (1):

For i (# of 2's) = 1 multiply by $\leq \alpha'_k \beta_k n$

$$= 2 \quad " \quad " \leq (\alpha'_k \beta_k n)^2 / 2$$

$$= 3 \quad " \quad " \leq (\alpha'_k \beta_k n)^3 / 3!$$

etc.

$$\sum_{i=1}^{\infty} \frac{(\alpha'_k \beta_k n)^i}{i!} = e^{\alpha'_k \beta_k n} - 1.$$

Case (2):

For i (# of 2's) = 0 multiply by ≤ 1

$$= 1 \quad " \quad " \leq \alpha'_k \beta_{k^n}$$

$$= 2 \quad " \quad " \leq (\alpha'_k \beta_{k^n})^2 / 2$$

etc.

$$\sum_{i=0}^{\infty} \frac{(\alpha'_k \beta_{k^n})^i}{i!} = e^{\alpha'_k \beta_{k^n}}.$$

Similarly, sum over the triples: consider two cases:

(1) No 4-tuples or higher -tuples,

(2) Higher -tuples.

$$\text{Case (1): } \sum_{i=1}^{\infty} \frac{(\beta_{k^n}^2)^i}{i!} = e^{\beta_{k^n}^2} - 1.$$

$$\text{Case (2): } \sum_{i=0}^{\infty} \frac{(\beta_{k^n}^2)^i}{i!} = e^{\beta_{k^n}^2}.$$

Similarly, sum over the 4-tuples: consider two cases:

(1) No 5-tuples or higher -tuples,

(2) Higher -tuples.

$$\text{Case (1): } \sum_{i=1}^{\infty} \frac{(\alpha'_k \beta_{k^n}^2)^i}{i!} = e^{\alpha'_k \beta_{k^n}^2} - 1$$

$$\text{Case (2): } \sum_{i=0}^{\infty} \frac{(\alpha'_k \beta_{k^n}^2)^i}{i!} = e^{\alpha'_k \beta_{k^n}^2}.$$

Do the same for 5-tuples and beyond. Finally obtain

$$\left| P\left(\bigcup_{i=1}^n A_i^k\right) - (1 - (1 - P(A_1^k))^n) \right| \leq$$

$$\begin{aligned}
&\leq (e^{\alpha'_k \beta_k^n} - 1)e^{\beta_k^n} \\
&\quad + (e^{\beta_k^{2n}} - 1)e^{\alpha'_k \beta_k^n} e^{\beta_k^n} \\
&\quad + (e^{\alpha'_k \beta_k^{2n}} - 1)e^{\beta_k^{2n}} e^{\alpha'_k \beta_k^n} e^{\beta_k^n} \\
&\quad + \dots
\end{aligned}$$

Collapse and get

$$\begin{aligned}
&= -e^{\beta_k^n} + \\
&\quad + e^{\beta_k^n(1+\beta_k+\beta_k^2+\dots)} + \alpha'_k \beta_k^n(1+\beta_k+\beta_k^2+\dots) \\
&= e^{\frac{\beta_k^n}{1-\beta_k}} + \frac{\alpha'_k \beta_k^n}{1-\beta_k} - e^{\beta_k^n} \\
&= e^{\beta_k^n} \left(e^{\frac{\alpha'_k + \beta_k}{1-\beta_k} \beta_k^n} - 1 \right).
\end{aligned}$$

This completes the proof of the lemmas. In addition to these lemmas I will need several facts which I will state without proof.

Unless otherwise specified I will assume a and b are constants $\in (0,1)$.

Fact 1: For any $\epsilon > 0$ there exists an integer N_0 such that

$$|e^{-aN} - (1-a)^N| < \epsilon \text{ for } N \geq N_0.$$

Fact 2: For any $\epsilon > 0$ there exists a $\delta > 0$ such that for all positive integers N ,

$$|e^{-aN} - e^{-bN}| < \epsilon \text{ for } \left| \frac{a}{b} - 1 \right| < \delta.$$

Fact 3: For any constants $c > 0$ and $\epsilon > 0$ there exists an integer N_0 such that for all $a \in (0,1)$

$$e^{-a(N-c)} - e^{a(N+c)} < \varepsilon \text{ for } N \geq N_0.$$

Fact 4: Let $b_k = \frac{e^{-\lambda} \lambda^k}{k!}$, $\lambda > 0$.

$$\frac{b_{k+1}}{b_k} = \frac{\lambda}{k+1} \xrightarrow{k \rightarrow \infty} 0.$$

Fact 5: $b_k \sim \sum_{i=k}^{\infty} b_i$.

Fact 6: $\frac{\sum_{i=k}^{\infty} b_i}{\sum_{i=k}^{\infty} b_i} \xrightarrow{k \rightarrow \infty} 0$.

Now return to the proof of the Theorem.

To prove: $\sum_{k=0}^{\infty} |p_{k,\lambda}(T) - q_{k,\lambda}(T)| \xrightarrow{T \rightarrow \infty} 0,$

(I will suppress the λ .)

or for any $\varepsilon > 0$, $\lambda > 0$ there exists an integer T_0

(≥ 2 , say) such that

$$\sum_{k=0}^{\infty} |p_k(T) - q_k(T)| < \varepsilon \text{ for } T \geq T_0.$$

Define: $P_k(T) = \sum_{i=k}^{\infty} p_i(T),$

$$Q_k(T) = \sum_{i=k}^{\infty} q_i(T),$$

$$R_k(T) = 1 - (1 - P(A_1^k))^{[T]}$$

where A_1^k is the event defined in Lemma 2 on p. 28.

The proof of the theorem will consist of cleverly choosing k_1 and using the following fact:

for any non-negative integer k_1 we have

$$\begin{aligned} \sum_{k=0}^{\infty} |p_k(T) - q_k(T)| &\leq \sum_{i=0}^{k_1} p_i(T) + \sum_{i=0}^{k_1} q_i(T) \\ &+ |p_{k_1+1}(T) - q_{k_1+1}(T)| \\ &+ |p_{k_1+2}(T) - q_{k_1+2}(T)| \\ &+ \sum_{i=k_1+3}^{\infty} p_i(T) + \sum_{i=k_1+3}^{\infty} q_i(T) \end{aligned}$$

and from this we get

$$\begin{aligned} \sum_{k=0}^{\infty} |p_k(T) - q_k(T)| &\leq 2 - P_{k_1+1}(T) - Q_{k_1+1}(T) \\ &+ |P_{k_1+1}(T) - R_{k_1+1}(T)| + |R_{k_1+1}(T) - Q_{k_1+1}(T)| \\ &+ 2|P_{k_1+2}(T) - R_{k_1+2}(T)| + 2|R_{k_1+2}(T) - Q_{k_1+2}(T)| \\ &+ |P_{k_1+3}(T) - R_{k_1+3}(T)| + |R_{k_1+3}(T) - Q_{k_1+3}(T)| \\ &+ P_{k_1+3}(T) + Q_{k_1+3}(T). \end{aligned}$$

Given any $\varepsilon > 0$, $\lambda > 0$ find k_0 and T_0 . Then

for any $T \geq T_0$ find the largest integer k such that

$$e^{-P(A_1^{k+1})(T/2)} < \varepsilon_a \quad (\varepsilon_a \text{ suitably chosen}).$$

Call this integer k_1 .

How large must k_0 be? Choose k_0 large enough

so that

$$1. \left| e^{-\sum_k b_i T} - e^{-P(A_1^{k+1})T} \right| < \epsilon_3, \text{ for } k \geq k_0.$$

Such a k_0 exists by Lemma 1 and Fact 2 for any $\epsilon_3 > 0$. (I will specify all my ϵ_i 's at the end of the proof.) Here I use

$$b_i = \frac{e^{-\lambda} \lambda^i}{i!}.$$

$$2. \left(\left(\frac{1}{\epsilon_a} \right)^3 \frac{\alpha'_{k+2} + \beta_{k+2}}{1 - \beta_{k+2}} - 1 \right) < \epsilon_a^3 \epsilon_b \text{ for } k \geq k_0$$

where $\alpha_k = P(A_1^k A_2^k) / P(A_1^k)$

$$\beta_k = P(A_1^k)$$

and $\alpha'_k = \max(\alpha_k, \beta_k)$.

Such a k_0 exists by Lemmas 1 and 2.

$$3. \left| e^{-\sum_{k+1} b_i [T]} - e^{-P(A_1^{k+2})[T]} \right| < \frac{1}{2} \epsilon_a^3$$

for $k \geq k_0$. Such a k_0 exists by Lemma 1 and Fact 2.

$$4. \frac{\sum_{k+2}^{\infty} b_i}{\sum_{k+1}^{\infty} b_i} \frac{\log(1 - \epsilon_4)}{\log \frac{\epsilon_a}{2}} \text{ for } k \geq k_0.$$

Such a k_0 exists by Fact 6.

How large must T_0 be? Choose T_0 large enough so that for any $a \in (0, 1)$

$$1. e^{-P(A_1^{k_0+1})[T_0/2]} < \epsilon_a.$$

$$2. \left| e^{-a[\lambda T]} - (1-a)^{[\lambda T]} \right| < \epsilon_1 \text{ for } T \geq T_0$$

$$|e^{-a[T-1]} - (1-a)^{[T-1]}| < \epsilon_1 \quad \text{for } T \geq T_0.$$

Such T_0 exists by Fact 1.

$$\left. \begin{aligned} 3. \quad & |e^{-a\lambda[T]} - e^{-a[\lambda T]}| < \epsilon_2 \\ & |e^{-a[\lambda T]} - e^{-a\lambda T}| < \epsilon_2 \\ & |e^{-a[T]} - e^{-aT}| < \epsilon_2 \\ & |e^{-a[T]} - e^{-a[T-1]}| < \epsilon_2 \end{aligned} \right\} \quad \text{for } T \geq T_0.$$

Such T_0 exists by Fact 3.

In the above inequalities the absolute value signs aren't always necessary. The square brackets about T , λT and $T-1$ indicate the greatest integer function.

Now consider the expression in the middle of page 43 term by term. k_1 will be as specified on page 43, namely, the largest integer such that

$$e^{-P(A_1^{k_1+1})[T/2]} < \epsilon_a.$$

For $T \geq T_0$ we have $k_1 \geq k_0$.

$$Q_{k_1+1}(T) = 1 - (1 - \sum_{k_1}^B b_i)^{[\lambda T]}$$

$$R_{k_1+1}(T) = 1 - (1 - P(A_1^{k_1+1}))^{[T]}.$$

First we consider the term

$$|R_{k_1+1}(T) - Q_{k_1+1}(T)|.$$

$$\begin{aligned}
& \left| R_{k_1+1}(T) - Q_{k_1+1}(T) \right| \\
&= \left| \left(1 - \sum_{k_1}^{\infty} b_i\right) [\lambda T] - \left(1 - P(A_1^{k_1+1})\right) [T] \right| \\
&\leq \left| \left(1 - \sum_{k_1}^{\infty} b_i\right) [\lambda T] - e^{-\sum_{k_1}^{\infty} b_i [\lambda T]} \right| \\
&\quad + \left| e^{-\sum_{k_1}^{\infty} b_i [\lambda T]} - e^{-\sum_{k_1}^{\infty} b_i \lambda T} \right| \\
&\quad + \left| e^{-\lambda \sum_{k_1}^{\infty} b_i T} - e^{-P(A_1^{k_1+1}) T} \right| \\
&\quad + \left| e^{-P(A_1^{k_1+1}) T} - e^{-P(A_1^{k_1+1}) [T]} \right| \\
&\quad + \left| e^{-P(A_1^{k_1+1}) [T]} - \left(1 - P(A_1^{k_1+1})\right) [T] \right| \\
&\leq \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_2 + \epsilon_1 = \underline{2\epsilon_1 + 2\epsilon_2 + \epsilon_3}.
\end{aligned}$$

Similarly:

$$2 \left| R_{k_1+2}(T) - Q_{k_1+2}(T) \right| \leq \underline{4\epsilon_1 + 4\epsilon_2 + 2\epsilon_3}$$

$$\text{and } \left| R_{k_1+3}(T) - Q_{k_1+3}(T) \right| \leq \underline{2\epsilon_1 + 2\epsilon_2 + \epsilon_3}.$$

Now consider the term

$$2 - P_{k_1+1}(T) - Q_{k_1+1}(T).$$

$$P_{k_1+1}(T) = 1 - \sum_{i=0}^{k_1} p_i(T).$$

$$\text{But } \sum_{i=0}^{k_1} p_i(T) \leq \left(1 - P(A_1^{k_1+1})\right) [T/2] \leq e^{-P(A_1^{k_1+1}) [T/2]} \leq \epsilon_a.$$

$$\text{Therefore } P_{k_1+1}(T) > 1 - \epsilon_a.$$

The first inequality above follows from the idea

used in the proof of Lemma 3.

$$R_{k_1+1}(T) = 1 - (1 - P(A_1^{k_1+1}))^{[T]}$$

$$> 1 - e^{-P(A_1^{k_1+1})[T]} > 1 - \epsilon_a.$$

$$Q_{k_1+1}(T) \geq R_{k_1+1}(T) - |R_{k_1+1}(T) - Q_{k_1+1}(T)|$$

$$> 1 - \epsilon_a - 2\epsilon_1 - 2\epsilon_2 - \epsilon_3.$$

$$\text{So } 2 - P_{k_1+1}(T) - Q_{k_1+1}(T) < \underline{2\epsilon_a + 2\epsilon_1 + 2\epsilon_2 + \epsilon_3}.$$

$$\text{The term } |P_{k_1+1}(T) - R_{k_1+1}(T)| < \underline{\epsilon_a}$$

$$\text{since } P_{k_1+1}(T) > 1 - \epsilon_a$$

$$\text{and } R_{k_1+1}(T) > 1 - \epsilon_a.$$

Now consider the term

$$2|P_{k_1+2}(T) - R_{k_1+2}(T)|.$$

$$P\left(\bigcup_{i=1}^{[T]-1} A_i^{k_1+2}\right) \leq P_{k_1+2}(T) \leq P\left(\bigcup_{i=1}^{[T]} A_i^{k_1+2}\right).$$

I look at two cases, corresponding to the bounds on $P_{k_1+2}(T)$ shown above.

Case 1:

$$\left| P\left(\bigcup_{i=1}^{[T]} A_i^{k_1+2}\right) - (1 - (1 - P(A_1^{k_1+2}))^{[T]}) \right|$$

$$< e^{\beta_{k_1+2}[T]} \left(e^{\frac{\alpha_{k_1+2} + \beta_{k_1+2}}{1 - \beta_{k_1+2}} \beta_{k_1+2}[T]} - 1 \right)$$

$$< \epsilon_b.$$

Since $e^{-P(A_1^{k_1+2})[T/2]} \geq \epsilon_a$ by choice of k_1 we have

$$e^{-P(A_1^{k_1+2})[T]} \geq \epsilon_a^3$$

But $e^{P(A_1^{k_1+2})[T]} = e^{\beta_{k_1+2}[T]}$ so

$$e^{\beta_{k_1+2}[T]} \leq \frac{1}{\epsilon_a^3} \text{ and}$$

$$\dots ((e^{\beta_{k_1+2}[T]})^{\frac{\alpha'_{k_1+2} + \beta_{k_1+2}}{1 - \beta_{k_1+2}}} - 1) \leq \epsilon_a^3 \epsilon_b.$$

Case 2:

$$\begin{aligned} & \left| P\left(\bigcup_{i=1}^{[T]-1} A_i^{k_1+2}\right) - (1 - (1 - P(A_1^{k_1+2}))^{[T]}) \right| \\ & \leq \left| P\left(\bigcup_{i=1}^{[T]-1} A_i^{k_1+2}\right) - (1 - (1 - P(A_1^{k_1+2}))^{[T]-1}) \right| \\ & \quad + \left| (1 - (1 - P(A_1^{k_1+2}))^{[T]-1}) - (1 - (1 - P(A_1^{k_1+2}))^{[T]}) \right| \\ & \leq \epsilon_b + \left| (1 - P(A_1^{k_1+2}))^{[T]} - e^{-P(A_1^{k_1+2})[T]} \right| \\ & \quad + \left| e^{-P(A_1^{k_1+2})[T]} - e^{-P(A_1^{k_1+2})([T]-1)} \right| \\ & \quad + \left| e^{-P(A_1^{k_1+2})([T]-1)} - (1 - P(A_1^{k_1+2}))^{[T]-1} \right| \\ & < \epsilon_b + 2\epsilon_1 + \epsilon_2. \end{aligned}$$

So, in either case

$$\left| P_{k_1+2}(T) - R_{k_1+2}(T) \right| < \epsilon_b + 2\epsilon_1 + \epsilon_2$$

$$\text{and } 2 \left| P_{k_1+2}(T) - R_{k_1+2}(T) \right| < \underline{2\epsilon_b + 4\epsilon_1 + 2\epsilon_2}.$$

Similarly

$$\left| P_{k_1+3}(T) - R_{k_1+3}(T) \right| < \underline{\varepsilon_b + 2\varepsilon_1 + \varepsilon_2}.$$

Finally, look at $P_{k_1+3}(T)$ and $Q_{k_1+3}(T)$. We already

know that

$$e^{-P(A_1^{k_1+2})[T]} \geq \varepsilon_a^3.$$

$$e^{-\lambda \sum_{k_1+1} b_i [T]} \geq \frac{1}{2} \varepsilon_a^3 \text{ by choice of } k_0,$$

$$e^{-\lambda \sum_{k_1+2} b_i [T]} \geq 1 - \varepsilon_4 \text{ by choice of } k_0.$$

$$\left| e^{-\lambda \sum_{k_1+2} b_i [T]} - e^{-\sum_{k_1+2} b_i [\lambda T]} \right| < \varepsilon_2,$$

$$\left| e^{-\sum_{k_1+2} b_i [\lambda T]} - (1 - \sum_{k_1+2} b_i) [\lambda T] \right| < \varepsilon_1.$$

$$Q_{k_1+3}(T) = (1 - \sum_{k_1+2} b_i) [\lambda T] < \underline{\varepsilon_1 + \varepsilon_2 + \varepsilon_4}.$$

$$P_{k_1+3}(T) \leq Q_{k_1+3}(T) + \left| P_{k_1+3}(T) - R_{k_1+3}(T) \right| + \left| R_{k_1+3}(T) - Q_{k_1+3}(T) \right|$$

$$< (\varepsilon_1 + \varepsilon_2 + \varepsilon_4) + (\varepsilon_b + 2\varepsilon_1 + \varepsilon_2) + (2\varepsilon_1 + 2\varepsilon_2 + \varepsilon_3).$$

$$P_{k_1+3}(T) < \underline{5\varepsilon_1 + 4\varepsilon_2 + \varepsilon_3 + \varepsilon_4 + \varepsilon_b}.$$

We have, then, that

$$\sum_{k=0}^{\infty} \left| p_k(T) - q_k(T) \right| < 22\varepsilon_1 + 18\varepsilon_2 + 6\varepsilon_3 + 2\varepsilon_4 + 3\varepsilon_a + 4\varepsilon_b.$$

When \mathbf{E} is chosen, let

$$\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = \varepsilon_a = \varepsilon_b = \frac{\varepsilon}{55}.$$

Thus, for $T \geq T_0$

$$\sum_{k=0}^{\infty} |p_k(T) - q_k(T)| < \varepsilon.$$

This completes the proof of the theorem.

BIBLIOGRAPHY

- Anscombe, F. J. Rejection of outliers. Technometrics, 2 (1960), 123-147.
- Barlow, H. B., and Levick, W. R. Three factors limiting the reliable detection of light by retinal ganglion cells of the cat. J. Physiol., 200 (1969), 1-24.
- Barlow, H. B., and Levick, W. R. Changes in the maintained discharge with adaptation level in the cat retina. J. Physiol., 202 (1969), 699-718.
- Berg, W. F. Aggregates in one- and two-dimensional random distributions. (Developability of silver specks of known dimensions and the size of photographic sensitivity specks.) The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science ("Phil. Mag."), 36: 7th Series (1945), 337-346.
- Dixon, W. J. Analysis of extreme values. Ann. Math. Statist., 21 (1950), 488-506.
- Dixon, W. J. Ratios involving extreme values. Ann. Math. Statist., 22 (1951), 68-78.
- Dixon, W. J. Processing data for outliers. Biometrics, 9 (1953), 74-89.

- Geffroy, J. Contributions à la théorie des valeurs extrêmes. Publ. Inst. Statist. Univ. Paris, 7 and 8 (1958 and 1959), 37-185.
- Gnedenko, B. Sur la distribution limite du terme maximum d'une série aléatoire. Annals of Math., 44 (1943), 423-453.
- Grubbs, F. E. Procedures for detecting outlying observations in samples. Technometrics, 11 (1969), 1-21.
- Gumbel, E. J. Statistics of Extremes. New York: Columbia University Press, (1958).
- Mack, C. An exact formula for $Q_k(n)$, the probable number of k-aggregates in a random distribution of n points. Phil. Mag., 39: 7th Series (1948), 778-790.
- Mises, R. von. Über die Variationsbreite einer Beobachtungsreihe. Sitzungsber. d. Berliner Math. Ges., 22 (1923), 3-8.
- Naus, J. Clustering of Random Points in Line and Plane. Thesis, Harvard University (1963).
- Newell, G. F. Distribution for the smallest distance between any pair of kth nearest neighbor random points on a line. Proceedings of Symposium on Time Series Analysis (Brown University, 1962),

89-103. New York: Wiley.

Neyman, J. and Scott, E. L. Outlier proneness of phenomena and of related distributions.

Optimizing Methods in Statistics. New York: Academic Press, (1971).

Silberstein, L. The probable number of aggregates in random distributions of points. Phil. Mag., 36: 7th Series (1945), 319-336.