CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

On Testing for a non-homogeneous
Poisson process.

A Thesis submitted

by

Aline Chouinard

to

The School of Graduate Studies of

The University of Ottawa

in partial fulfillment of the requirements

for the degree of
Master of Science
in the subject of
Mathematics

November, 1982

CONTENTS

Acknowledgement

I would like to thank Dr. D. McDonald for suggesting the topic of this thesis. His assistance and guidance were essential to the preparation of this thesis and I wish to express my sincere appreciation of his effort.

I also wish to extend my thanks to the National Research Council of Canada and the Department of Mathematics at the University of Ottawa for their financial support during the preparation of this thesis. The rapid and accurate work done by Mrs. Michelle Lukaszczyk in the typing of this paper was also much appreciated.

Section 1

Introduction

Officials at the Museum of Man in Ottawa presented us with the following problem. Their security service has recorded the number of persons entering the museum every hour and also the number of persons leaving it every hour, over several years. They would like to know the distribution of the viewing time of visitors to the museum.

As Mr. Fortin of the Museum of Man pointed out it's easy to estimate the average daily sojourn time. We just count the total number of viewing hours in a day and divide by the number of visitors that day. At the beginning of each hour it is not difficult to calculate the number of visitors who are in the museum (for the next hour). We merely take the total number of arrivals until that time and subtract from it the number of departures. For each hour of the day we have the number of viewers. Adding up over all the hours we have the total number of viewing hours. The table below gives an example.

	came in	actually in	left	stay over
1st hour	5	5	2	3
2nd hour	3	6	5	1
3rd hour	8	9	9	0

Here 16 persons came to the museum and they (alltogether) spent 20 hours in it.

This problem can of course be expressed in more general terms as: find the sojourn time distribution of certain objects in a system for which we record the number of objects coming in and out of the system for each time unit.

To find out further information about the sojourn time distribution a probabilistic model is necessary.

Section 2

A probabilistic model

We postulate the following hypothesis. The stochastic process by which visitors arrive at the museum is a non-homogeneous Poisson process and there exists a sojourn time independent of the arrival time of the people to the museum.

Let us first define a non-homogeneous Poisson process.

<u>Definition</u> (stated in Ross p. 24): $[N(t), t \geq 0]$ is said to be a non-homogeneous Poisson process with intensity function $\lambda(t)$ if

i) $N(0) = 0$

ii) $[N(t), t \geq 0]$ has independent increments

iii) $P[2 \text{ or more events in } (t, t+h)] = o(h)$

where $o(h)$ means $\lim_{h \to 0} \frac{f(h)}{h} = 0$

iv) $P[\text{exactly 1 event in } (t, t+h)] = \lambda(t)h + o(h)$

Moreover, note that we can show

$$P[N(t) = n] = e^{-m(t)} \frac{[m(t)]^n}{n!} \qquad n \geq 0$$

where $m(t) = \int_0^t \lambda(s)ds$. (This is done in Ross for the homogeneous case on p. 14-15).

In a later section we will develop at test to verify this hypothesis,

Section 3

The departure process

In this section, we will show that the process by which people leave the museum is also non-homogeneous Poisson.

For arrivals let

$X_{(i)}$: the time between the $(i-1)^{th}$ and i^{th} event.

$S_{(i)}$: the waiting time until the i^{th} event occurs i.e. $S_{(n)} = \sum_{i=1}^{n} X_{(i)}$

$N(t)$: the number of events that have occurred by time t.

The $S_{(i)}$ defined above are clearly ordered in increasing size. Consider a random permutation Π of $1,2,\ldots,n$ (each permutation has an equal probability). Define $S_i = S_{(\pi(i))}$. The random variables S_i are the unordered waiting times.

For departures, let

$Y(t)$ = the number of departures from the museum that have occurred by time t

$$W_i = \begin{cases} 1 & \text{if the visitor to enter the museum at time } S_i \text{ (unordered) leaves before time } t \\ 0 & \text{if he stays beyond time } t. \end{cases}$$

By definition, $Y(t) = \sum_{i=1}^{N(t)} W_i$.

Let us also denote by $G(x)$ the distribution of the sojourn time. (i.e. $G(x) = P\{\text{sojourn time} \le x\}$).

Theorem 1:

Let $\{N(t),\ t\geq 0\}$ be a non-homogeneous Poisson process with mean value function $m(t)$. Given $N(t) = n$, the ordered set of arrival times $\{S_{(i)},\ i=1,\ldots,n\}$ has the same distribution as the order statistics of n independent and identically distributed random variables having density function $f(s) = \frac{\lambda(s)}{m(t)}$ for $s \in (0,t)$, where $m(t) = \int_0^t \lambda(s)ds$.

Proof:

The conditional joint density function of $S_{(1)},\ldots,S_{(n)}$ given $N(t)$ is:

$$f_{S_{(1)},\ldots,S_{(n)}|N(t)}(t_1,\ldots,t_n|n) \equiv \lim_{\substack{h_1\downarrow 0\\ h_2\downarrow 0\\ \vdots\\ h_n\downarrow 0}} \frac{P\{S_{(1)}\in(t_1,t_1+h_1],\ldots,S_{(n)}\in(t_n,t_n+h_n],\ N(t) = n\}}{P\{N(t)=n\}\quad h_1\ h_2\ldots h_n}$$

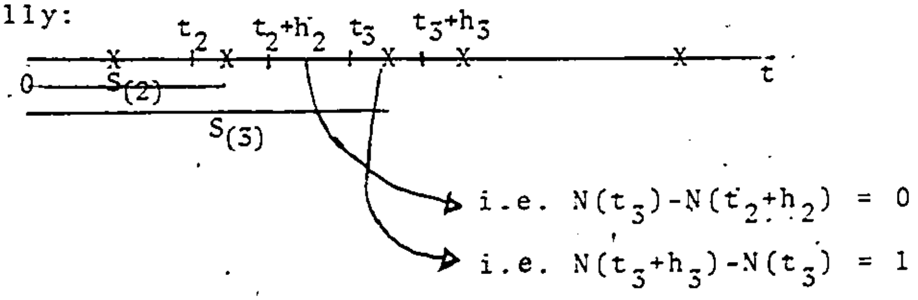
where h_i is such that $t_i+h_i < t_{i+1}$ and $0 < t_1 < t_2 < \ldots < t_n < t$

$$\equiv \lim P\{N(t_1)=0,\ N(t_1+h_1)-N(t_1) = 1,\ N(t_2)-N(t_1+h_1) = 0,\ldots$$

$$N(t_n)-N(t_{n-1}+h_{n-1}) = 0,\ N(t_n+h_n)-N(t_n) = 1,$$

$$N(t)-N(t_n+h_n)=0\} \div P\{N(t)=n\}\ h_1\ h_2\ldots h_n$$

Graphically:



i.e. $N(t_3)-N(t_2+h_2) = 0$

i.e. $N(t_3+h_3)-N(t_3) = 1$

$$= \lim \frac{e^{-m(t_1)}[\lambda(t_1)h_1+o(h_1)]e^{-(m(t_2)-m(t_1+h_1))}\cdots[\lambda(t_n)h_n+o(h_n)]e^{-(m(t)-m(t_n+h_n))}}{e^{-m(t)}\frac{(m(t))^n}{n!}\quad h_1\cdots h_n}$$

because $N(t)$ has independent increments and

$$P\{N(t_i+h_i) - N(t_i) = 1\} = \lambda(t_i)h_i + o(h_i) \quad \text{and}$$

$$P\{N(t_{i+1}) - N(t_i+h_i) = 0\} = e^{-\int_{t_i+h_i}^{t_{i+1}} \lambda(s)ds}$$

$$= e^{-(m(t_{i+1}) - m(t_i+h_i))}.$$

$$= \lim \frac{n!}{(m(t))^n}\left[\frac{\lambda(t_1)h_1+o(h_1)}{h_1}\right]\cdots\left[\frac{\lambda(t_n)h_n+o(h_n)}{h_n}\right]\cdot e^{-(m(t_1)-m(t_1+h_1))} \cdots e^{-[m(t_n)-m(t_n+h_n)]} \qquad (1)$$

Note that

$-m(t_i) + m(t_i+h_i) = \int_{t_i}^{t_i+h_i} \lambda(s)ds$ but $[t_i, t_i+h_i)$ has measure zero as $h_i \to 0$.

So,

$$f_{S_{(1)},\ldots,S_{(n)}|N(t)}(t_1,\ldots,t_n|n) =$$

$$= \frac{n!}{(m(t))^n}\ \lambda(t_1)\ldots\lambda(t_n)$$

$$= n!\ \frac{\lambda(t_1)}{m(t)} \cdots \frac{\lambda(t_n)}{m(t)} = n!\ \ f(t_1)\ldots f(t_n) \qquad \square$$

Now we are ready to show that indeed $Y(t)$ is also a non-homogeneous Poisson process. Let us find the characteristic function of $Y(t)$ (denoted ϕ_Y)

$$\phi_Y(u) = E[e^{iuY(t)}] \qquad u \in R$$

$$= \sum_{n=0}^{\infty} E[e^{iuY(t)}|N(t)=n]\ P\ [N(t)=n].$$

But $E[e^{iuY(t)}\,|N(t)=n] \equiv E[e^{iu\sum_{j=1}^{N(t)} W_j}|N(t)=n]$

$\equiv E[e^{iu\sum_{j=1}^{n} W_j}] = \qquad = \prod_{j=1}^{n} \phi_{W_j}(u)$

because the characteristic function of a sum of independent random variables is equal to the product of the individual characteristic functions, also $\prod_{j=1}^{n} \phi_{W_j}(u) = [\phi_{W_j}(u)]^n$ since the W_j are independent, identically distributed.

Observe that,

$P(W_j=1)$ = P (the visitor who came at time S_j leaves before time t)

$= \int_0^t$ P (the visitor who came at time S_j leaves before time t / $S_j = s$)

$P(S_j \epsilon ds)$

$= \int_0^t P(\text{sojourn time} \le t-s)\ P(S_j \epsilon\, ds)$

$= \int_0^t G(t-s) \frac{\lambda(s)}{m(t)}\, ds$ by theorem 1

and $P(W_j=0) = 1 - \int_0^t G(t-s) \frac{\lambda(s)}{m(t)}\, ds$.

So $E[e^{iu\,W_j}] = \sum_{w=0}^{1} e^{iu\,w}\, P(W_j=w)$

$= 1 - \int_0^t G(t-s) \frac{\lambda(s)}{m(t)}\, ds + e^{iu} \int_0^t G(t-s) \frac{\lambda(s)}{m(t)}\, ds$

$= \int_0^t [1-G(t-s) + e^{iu}\, G(t-s)] \frac{\lambda(s)}{m(t)}\, ds = I$

So that $E[e^{iuY(t)}] = \sum_{n=0}^{\infty} I^n \cdot P[N(t)=n] = \sum_{n=0}^{\infty} I^n\, e^{-m(t)} \frac{[m(t)]^n}{n!}$

$= e^{-m(t)} \sum_{n=0}^{\infty} \{\frac{1}{m(t)} \int_0^t [1-G(t-s)+e^{iu}G(t-s)]\lambda(s)ds\}^n \frac{(m(t))^n}{n!}$

$$= e^{-m(t)} \sum_{n=0}^{\infty} \{\int_0^t [1+(e^{iu}-1)\ G(t-s)]\lambda(s)ds\}^n \frac{1}{n!}$$

$$= e^{-m(t)}\ \exp\{\int_0^t [1+(e^{iu}-1)G(t-s)]\lambda(s)ds\}$$

$$= e^{-m(t)}\ \exp\{m(t)\ +\ \int_0^t (e^{iu}-1)G(t-s)\lambda(s)ds\}$$

$$= \exp\ \{(e^{iu}-1)\ \int_0^t\ G(t-s)\lambda(s)ds\}$$

which is the characteristic function of a Poisson distribution with mean function given by

$r(t) = \int_0^t G(t-s)\lambda(s)ds.$

Section 4

The sojourn time distribution

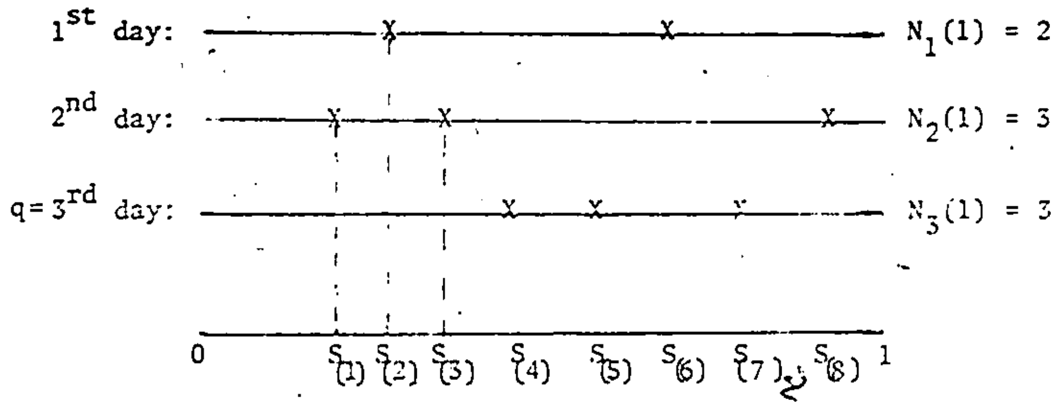
Now we shall find an expression for the sojourn time distribution. As seen at the end of the last section, we have $r(t) = \int_0^t G(t-s)\lambda(s)ds = \int_0^t G(t-s)dm(s)$ since $m(s) = \int_0^s \lambda(x)dx$. Thus, r is the convolution of G and m. Although we do not know m or r we may estimate them since they are mean value functions. We shall take a sample of q days and we redefine:

$N_j(x)$ = the number of arrivals counted by time x on the j^{th} day.

$M_j(x)$ = the number of departures counted by time x on the j^{th} day.

$S^j_{(i)}$ = the waiting time until the i^{th} arrival on the j^{th} day.

Now we may order the arrival times among q days (or among the $n_1 + \ldots + n_q = n$ observations) and call them $S_{(i)}$. Similarly we will call $T_{(i)}$ the ordered departure times among q days. Let us clarify by an example: the (X) represent the occurence of an event on the time interval (0,1).



$S_{(1)} = S^2_{(1)}, \; S_{(2)} = S^1_{(1)}, \; S_{(3)} = S^2_{(2)}, \; S_{(4)} = S^3_{(1)}, \; S_{(5)} = S^3_{(2)}, \; S_{(6)} = S^1_{(2)}, \; S_{(7)} = S^3_{(3)},$
$S_{(8)} = S^2_{(3)}$.

Explicitely we estimate m(x) by $\hat{m}(x) = \sum_{j=1}^{q} \frac{N_j(x)}{q}$ where q is the number of days and $N_j(x)$ is the number of arrivals counted by time x on the j^{th} day. Similarly, we estimate r(x) by $\hat{r}(x) = \sum_{j=1}^{q} \frac{M_j(x)}{q}$ where $M_j(x)$ is the number of departures counted by time x on the j^{th} day.

Using the Laplace transform (denoted $\mathcal{L}$), we get

$$\mathcal{L}(r(t)) = \mathcal{L}(m(t)) \cdot \mathcal{L}(G(t)) \qquad (1)$$

so that using the inverse transform (denoted $\mathcal{L}^{-1}$) we now get,

$$G(t) = \mathcal{L}^{-1}\left\{\frac{\mathcal{L}(r(t))}{\mathcal{L}(m(t))}\right\} \qquad (2)$$

But $\mathcal{L}(m(s)) = \int_0^\infty e^{-st}\, dm(t)$, let us replace m(t) by $\hat{m}(t)$

$$\mathcal{L}(m(s)) = \int_0^\infty e^{-st}\, d\hat{m}(t) = \sum_{i=1}^{N} e^{-sS(i)} \frac{1}{q}$$

Similarly, we have an estimate of $\mathcal{L}(r(s))$

$$\mathcal{L}(r(s)) = \sum_{i=1}^{N} e^{-sT(i)} \frac{1}{q}$$

Readily we can obtain an estimate of G(t) via (2).

Observe that $\mathcal{L}(G(s)) = E(e^{-sG})$. So that $(\mathcal{L}G)^{(1)}(0) = -\mu_G$ (1^{st} moment of a random variable having distribution G), and $(\mathcal{L}G)^{(2)}(0) = m_2^G$ (2^{nd} moment of a random variable having distribution G).

$$\text{Here } \mathcal{L}(G(s)) = \frac{\mathcal{L}(r(s))}{\mathcal{L}(m(s))} \simeq \frac{\sum_{i=1}^{N} e^{-sT(i)}}{\sum_{i=1}^{N} e^{-sS(i)}} .$$

Let us compute the estimates for μ_G and σ_G^2.

$$\hat{\mu}_G = -(\hat{\mathcal{L}} G)^{(1)}(0) = -\frac{1}{N}\left(\sum_{i=1}^{N} S_{(i)} - \sum_{i=1}^{N} T_{(i)}\right) = \frac{1}{N}\sum_{i=1}^{N}(T_{(i)} - S_{(i)}).$$

$$\hat{m}_2^G = \frac{1}{N}\left\{\sum_{i=1}^{N} T_{(i)}^2 - \sum_{i=1}^{N} S_{(i)}^2\right\} + \frac{2}{N^2}\left\{\left(\sum_{i=1}^{N} S_{(i)}\right)^2 - \left(\sum_{i=1}^{N} S_{(i)}\right)\left(\sum_{i=1}^{N} T_{(i)}\right)\right\}$$

$$\hat{\sigma}_G^2 = \frac{1}{N}\left\{\sum_{i=1}^{N} T_{(i)}^2 - \sum_{i=1}^{N} S_{(i)}^2\right\} + \frac{1}{N^2}\left\{\left(\sum_{i=1}^{N} S_{(i)}\right)^2 - \left(\sum_{i=1}^{N} T_{(i)}\right)^2\right\}$$

The mean value of G is computed as the average length of time spent by the N persons in the museum. The variance of G can be expressed in the following way:

$$\hat{\sigma}_G^2 = \frac{1}{N}\sum_{i=1}^{N}\left(T_{(i)} - \frac{\sum_{i=1}^{N} T_{(i)}}{N}\right)^2 - \frac{1}{N}\sum_{i=1}^{N}\left(S_{(i)} - \frac{\sum_{i=1}^{N} S_{(i)}}{N}\right)^2 . \qquad (3)$$

Consider T the distribution of the departure time and S the distribution of the arrival time. Then for any individual, his time of departure is his time of arrival plus his sojourn time (i.e. T=S+G). As the sojourn time and arrival time are independent, then $\sigma_G^2 = \sigma_T^2 - \sigma_S^2$ and this is given by (3).

This is not really a solution to the real life problem. In the case of the museum the departure process is censored when the museum closes so we cannot observe $\hat{f}(x)$. The average viewing time per visitor suggested by Mr Fortin is clearly less than the mean sojourn time. This problem deserves further study.

Section 5

Time transformation of a nonhomogeneous Poisson process

We can reduce a nonhomogeneous Poisson process to a homogeneous one with a transformation. For each time value s we define a new time value (denoted $\tilde{s}$) as follows:

$$s \to \tilde{s} = \frac{m(s)}{m(1)} \in [0,1]$$

note: $s = m^{-1}(\tilde{s}\ m(1))$.

For the process to be unaltered we set $\tilde{N}(\tilde{s}) = N(s)$. Let us compute $P[\tilde{N}(\tilde{s})=n]$ and show that $\tilde{N}(\tilde{s})$ is homogeneous.

$$P[\tilde{N}(\tilde{s})=n] = P[N(s)=n] = P[N(m^{-1}[\tilde{s}\ m(1)])=n]$$

$$= e^{-m(m^{-1}[\tilde{s}\ m(1)])}[m(m^{-1}[\tilde{s}\ m(1)])]^n/n! = e^{-\tilde{s}\ m(1)}[\tilde{s}\ m(1)]^n/n!$$

Hence $\tilde{N}(\tilde{s})$ is a homogeneous Poisson process with rate $m(1)$. Observe that to each S_i there corresponds a $\tilde{S}_i = \frac{m(S_i)}{m(1)}$.

An analogue to theorem 1 in the homogeneous case exists and its proof is a corollary to the proof of theorem 1 and is done in Ross p. 18.

<u>Theorem 2</u>: Given $N(1) = n$ ($N(1)$, Poisson process) the n ordered arrival times $\tilde{S}_{(1)},\dots,\tilde{S}_{(n)}$ have the same distribution as the order statistics corresponding to n independent random variables uniformly distributed on the interval $(0,1)$.

Thus given that $N(1) = \tilde{N}(\tilde{1}) = n$, $\tilde{S}_1,\ldots,\tilde{S}_n$ are distributed independently and uniformly in the interval $(0,1)$. So we may test if $\tilde{N}(\tilde{x})$ is Poisson by testing that $\tilde{S}_1,\ldots,\tilde{S}_n$ come from a uniform $(0,1)$ population. This however is not possible since we can only estimate $\tilde{S}_i = \frac{m(S_i)}{m(1)}$ by

$$\frac{\hat{m}(S_i)}{\hat{m}(1)} = \frac{\sum_{j=1}^{q} N_j(S_i)}{\sum_{j=1}^{q} N_j(1)} = \frac{k}{n} \quad \text{where} \quad S_i = S_{(k)} .$$

This estimate is not random.

Let us recapitulate. We have $N_i(1)$, $(i=1,\ldots,q)$ which are independent identically distributed random variables. Suppose $N_i(1) = n_i$, we have a sequence $S^i_{(1)},\ldots S^i_{(n_i)}$ which represent the waiting time until the $1^{st},\ldots,n_i^{th}$ event and the upper index refers to the day.

Given that $N_1(1) = n_1,\ldots,N_q(1) = n_q$, the points $S^1_{(1)},\ldots,S^1_{(n_1)},\ldots,S^q_{(1)},\ldots,S^q_{(n_q)}$ are ordered as

$$\widetilde{S^1_{(1)}} = \frac{m(S^1_{(1)})}{m(1)},\ldots,\widetilde{S^1_{(n_1)}} = \frac{m(S^1_{(n_1)})}{m(1)},\ldots,\widetilde{S^q_{(1)}} = \frac{m(S^q_{(1)})}{m(1)},\ldots,\widetilde{S^q_{(n_q)}} = \frac{m(S^q_{(n_q)})}{m(1)}$$

which (since m & $\hat{m}$ are increasing functions) in turn are ordered as their estimate,

$$\frac{\hat{m}(S^1_{(1)})}{\hat{m}(1)}, \dots, \frac{\hat{m}(S^1_{(n_1)})}{\hat{m}(1)}, \dots, \frac{\hat{m}(S^q_{(1)})}{\hat{m}(1)}, \dots, \frac{\hat{m}(S^q_{(n_q)})}{\hat{m}(1)}$$

which take their value in the set $\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$. We know that given $N_k(1) = n_k$, $\frac{m(S^k_{(1)})}{m(1)}, \dots, \frac{m(S^k_{(n_k)})}{m(1)}$ are distributed as n_k order statistics of independent uniformly distributed random variables. Hence given $N_1(1) = n_1, \dots, N_q(1) = n_q$, $\frac{\hat{m}(S^1_{(1)})}{\hat{m}(1)}, \dots, \frac{\hat{m}(S^q_{(n_q)})}{\hat{m}(1)}$ are ordered as the order statistics of n uniformly distributed random variables on (0,1).

We would like to know the probability of a specific arrangement of the events among the q days. From the set $\{\frac{1}{n}, \dots, \frac{n}{n}\}$, n_i values will be estimates for $S^{\tilde{i}}_1, \dots, S^{\tilde{i}}_{n_i}$. This is like drawing n_1 values from an urn containing balls numbered from $\frac{1}{n}$ to $\frac{n}{n}$ and assigning them to day 1, then drawing n_2 balls from the same urn and assigning them to day 2, etc... Hence there are $\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_{q-1}}{n_q} = \frac{n!}{n_1!n_2!\cdots n_q!}$ possible arrangements of the events among q days.

Section 6

The test for randomness

From the previous sections it is clear that if a simple point process is a non-homogeneous Poisson process we may test both

a) $\{N_i(1)\}_{i=1}^{q}$ are iid Poisson random variables

b) The ranks of the arrival times among the q days are "randomly distributed." That is, given $\{N_i(1)=n_i\}_{i=1}^{q}$ the distribution of the ranked arrival times is the same as the distribution of the ranks $\{1,\dots n\}$ drawn randomly into a groups of sizes $n_1, n_2, \dots n_q$. This is our null hypothesis H_0.

For b) let us rank the observations by assigning rank 1 to the observation which occured at time $S_{(1)}$, rank 2 to the one that occurred at time $S_{(2)}$, etc... where $S_{(i)}$ are the ordered arrival time among q days. Then denote R_{ij} the rank of the j^th^ observation of the i^th^ day such that $R_{i1}<R_{i2}<\dots<R_{in_i}$.

Now $E(\frac{R_{ij}}{n+1}) = \frac{j}{n_i+1}$ (shown in appendix section 1, proposition 2), so under the null hypothesis, we expect the values $\frac{R_{ij}}{n+1}$ to be close to $\frac{j}{n_i+1}$. This leads to the conclusion that the sum of the squared differences, $(\frac{R_{ij}}{n+1} - \frac{j}{n_i+1})^2$ should be small under H_0. Hence, reject

$$H_0 \text{ if } C = \sum_{i=1}^{q} \sum_{j=1}^{n_i} \left(\frac{R_{ij}}{n+1} - \frac{j}{n_i+1}\right)^2 > k \qquad (1)$$

We wish now to determine the critical value k that corresponds to a given significance level α. Chebyshev's inequality says:

$$P\{|C-\mu_C|<k\} \geq 1 - \frac{\sigma_C^2}{k^2} .$$

We have $\alpha = \frac{\sigma_C^2}{k^2}$, i.e. $k^2 = \frac{\sigma_C^2}{\alpha}$ and the inequality becomes:

$$P\{|C-\mu_C| < \frac{\sigma_C}{\sqrt{\alpha}}\} \geq 1-\alpha$$

In section 2 of the appendix, we compute μ_C and σ_C^2. This was under the condition that $N_1(1) = n_1, \ldots, N_q(1) = n_q$. For a) we will use a goodness of fit test to check if the $N_i(1)$'s are distributed like Poisson variables. We define a statistic $f(N_1(1), N_2(1), \ldots, N_q(1))$ which, under the null hypothesis that the $N_i(1)$'s are Poisson, satisfies

$$P(f(N_1(1), \ldots, N_q(1)) > k_1 | H_0) = \alpha_1. \qquad (2)$$

A bootstrap procedure will suffice. First estimate the intensity by $\Lambda = \frac{1}{q} \sum_{i=1}^{q} N_i(1)$. Next define classes

$I_i = \{n_i, n_i+1, \ldots n_{i+1}-1\}$ for $i = 1, \ldots \ell-1$, $I_\ell = \{n_\ell, n_{\ell+1} \ldots\}$

where $n_1 = 0$ and $\{n_i\}_{i=1}^{\ell}$ is a strictly increasing sequence of integers. Pick these integers (i) such that the $p_i = \sum_{k \in I_i} e^{-\Lambda} \frac{\Lambda^k}{k!}$ are approximately equal and not "too" small ($q \times p_i \geq 5$). Finally define

$$f(n_1, n_2, \ldots n_q) = \sum_{i=1}^{\ell} \frac{[\theta_i - q \cdot p_i]^2}{q\, p_i}$$

where θ_i designates the number of the n_k's in class I_i. $f(N_1(1), \ldots N_q(1))$ has, under H_0, a limiting χ^2 distribution with $\ell-2$ degrees of freedom; k_1 (in 2) is therefore determined.

Thus we reject H_o if (1) or (2) is satisfied. We perform the joint test using Chebyshev's inequality.

$$P\{|C-\mu_C| < k \text{ and } f(N_1(1),\ldots,N_q(1)) < k_1 | H_0\}$$

$$= E\{P[|C-\mu_C| < k | N_i(1) = n_i, i=1,\ldots,q] \; I_{f(N_1(1),\ldots,N_q(1))<k_1} | H_0\}$$

$$\geq (1-\alpha) \; E\{I_{f(N_1(1),\ldots,N_q(1))<k_1} | H_0\}$$

$$\geq (1-\alpha) \; (1-\alpha_1).$$

Section 7

Conclusion

The statistic C was used to test that a specific arrangement of the events among q days has probability $\frac{1}{\binom{n}{n_1 \ldots n_q}}$. When we want to test that there is no difference among q treatments and we have n_i observations available for each treatment. We may rank the $n = n_1 + \ldots + n_q$ observations at the end of the experiment. And if in fact there is no difference any arrangement of the ranks has probability $\frac{1}{\binom{n}{n_1 \ldots n_q}}$. This is usually done using Kruskal Wallis test. And $P(K \geq c) \to \chi^2_{q-1}(c)$ as $n_1, \ldots, n_q \to \infty$, K is the Kruskal Wallis statistic, χ^2_{q-1} designated the chi-square distribution with q-1 degrees of freedom. For our problem a Kruskal Wallis test is not appropriate because we want the n_i's to be small with q increasing instead.

There are many further problems to investigate. For instance, what is a good approximation of $P(C \geq k)$ when $q \to \infty$? Or as in the Kruskal Wallis case, what is a good approximation of $P(C \geq k)$ when $n_1, \ldots, n_q \to \infty$?

Bibliography

Lehman, E.L. (1975): Nonparametrics, statistical methods based on ranks, McGraw-Hill, Toronto.

Ross, S.M. (1970): Applied probability models with optimization applications, Holden-Day, San Francisco

Appendix

Section 1

Proposition 1: $\sum_{k=j}^{n-(n_i-j)} \binom{k}{j}\binom{n-k}{n_i-j} = \binom{n+1}{n_i+1}$

Proof:

The binomial theorem says:

$$(1+y)^m = 1 + my + \frac{m(m-1)}{2}y^2 + \dots + y^m$$

Thus

$$(1-x)^{-(j+1)} = 1 + (j+1)x + \frac{(j+1)\ (j+2)}{2}x^2 + \dots$$

$$= \sum_{k=j}^{\infty} \binom{k}{j} x^{k-j}, \qquad (1)$$

and $$(1-x)^{-(n_i-j+1)} = \sum_{\ell=n_i-j}^{\infty} \binom{\ell}{n_i-j} x^{\ell-(n_i-j)}. \qquad (2)$$

We notice that taking $\ell = n-k$ in r.h.s. of (2) and if we multiply it with the r.h.s. of (1), we get

$$\sum_{k=j}^{\infty} \binom{k}{j} \binom{n-k}{n_i-j} x^{n-n_i} = \sum_{k=j}^{n-(n_i-j)} \binom{k}{j} \binom{n-k}{n_i-j} x^{n-n_i}$$

Thus we are looking for the coefficient of x^{n-n_i} in the product of (1) with (2) i.e. in $(1-x)^{-(n_i+2)}$.

But $$(1-x)^{-(n_i+2)} = \sum_{m=n_i+1}^{\infty} \binom{m}{n_i+1} x^{m-(n_i+1)}.$$

We will find the coefficient of x^{n-n_i} if we let $m = n + 1$ and so the required coefficient is $\binom{n+1}{n_i+1}$.

Prop. 2: $E\,R_{i\,j} = \frac{j(n+1)}{(n_i+1)}$.

Proof: $R_{i\,j}$ takes on the values: $j, j+1, \ldots, n-(n_i-j)$ with corresponding probabilities:

$$\frac{\binom{j-1}{j-1}\binom{n-j}{n_i-j}}{\binom{n}{n_i}}, \frac{\binom{j}{j-1}\binom{n-(j+1)}{n_i-j}}{\binom{n}{n_i}}, \ldots, \frac{\binom{n-(n_i-j)-1}{j-1}\binom{n-[n-(n_i-j)]}{n_i-j}}{\binom{n}{n_i}}.$$

Then

$$E\,R_{i\,j} = \sum_{k=j}^{n-(n_i-j)} k\,\frac{\binom{k-1}{j-1}\binom{n-k}{n_i-j}}{\binom{n}{n_i}}$$

$$= \sum_{k=j}^{n-(n_i-j)} \frac{j\binom{k}{j}\binom{n-k}{n_i-j}}{\binom{n}{n_i}} \quad \text{because } k\binom{k-1}{j-1} = j\binom{k}{j}$$

$$= \frac{j}{\binom{n}{n_i}}\binom{n+1}{n_i+1} \qquad \text{by prop. 1.}$$

$$= j\frac{(n+1)}{(n_i+1)}.$$

Note 1: Let $X_i = \sum_{j=1}^{n_i} \frac{j\,R_{ij}}{(n_i+1)(n+1)}$

$$E\,X_i = \sum_{j=1}^{n_i} \frac{j^2}{(n_i+1)^2} = \frac{n_i(2n_i+1)}{6(n_i+1)}$$

Proposition 3: $\sum_{k=j}^{n-(n_i-j)} k\binom{k}{j}\binom{n-k}{n_i-j} = j\binom{n+1}{n_i+1} + (j+1)\binom{n+1}{n_i+2}$

Proof:

We know that,

$$x^j(1-x)^{-(j+1)} = \sum_{k=j}^{\infty}\binom{k}{j}x^k$$

Taking the derivative

$$j\,x^{j-1}(1-x)^{-(j+1)} + (j+1)x^j(1-x)^{-(j+2)} = \sum_{k=j}^{\infty} k\binom{k}{j}x^{k-1}. \qquad (1)$$

Also, $$(1-x)^{-(n_i-j+1)} = \sum_{\ell=n_i-j}^{\infty}\binom{\ell}{n_i-j}x^{\ell-(n_i-j)}. \qquad (2)$$

Again as in Prop. 1, taking $\ell = n - k$ in (2) and multiplying it with (1), we get:

$$\sum_{k=j}^{\infty} k\binom{k}{j}\binom{n-k}{n_i-j}\, x^{n-n_i+j-1} = \sum_{k=j}^{n-(n_i-j)} k\binom{k}{j}\binom{n-k}{n_i-j}x^{n-n_i+j-1}$$

thus we are looking for the coefficient of x^{n-n_i+j-1} in the product of (1) and (2) which is:

$$j\,x^{j-1}(1-x)^{-(n_i+2)} + (j+1)x^j(1-x)^{-(n_i+3)}$$

$$= j\sum_{m=n_i+1}^{\infty}\binom{m}{n_i+1}x^{m-n_i+j-2} + (j+1)\sum_{m=n_i+2}^{\infty}\binom{m}{n_i+2}x^{m-n_i-2+j}$$

to find the coefficient of x^{n-n_i+j-1} let $m = n + 1$, this leads to:

$$j\binom{n+1}{n_i+1} + (j+1)\binom{n+1}{n_i+2}.$$

Note 2: $$ER^2_{i\,j} = \sum_{k=j}^{n-(n_i-j)} \frac{k^2\binom{k-1}{j-1}\binom{n-k}{n_i-j}}{\binom{n}{n_i}} \quad \text{as in Prop. 2.}$$

$$= \frac{j}{\binom{n}{n_i}} \sum_{k=j}^{n-(n_i-j)} k\binom{k}{j}\binom{n-k}{n_i-j} \quad \text{because} \quad k^2\binom{k-1}{j-1} = jk\binom{k}{j}$$

$$= \frac{j}{\binom{n}{n_i}} \left[j\binom{n+1}{n_i+1} + (j+1)\binom{n+1}{n_i+2}\right] \text{ by Prop. 3.}$$

$$= j^2\,\frac{(n+1)}{(n_i+1)} + j(j+1)\frac{(n+1)(n-n_i)}{(n_i+1)(n_i+2)}$$

$$= j\,\frac{(n+1)}{(n_i+1)(n_i+2)}[j(n_i+2) + (j+1)(n-n_i)]$$

$$= j\,\frac{(n+1)}{(n_i+1)(n_i+2)}\,[n-n_i + (n+2)j].$$

Proposition 4: $$\sum_{k=j}^{n-(n_i-j)} \binom{k}{j} \cdot \sum_{\lambda=k+\gamma-j}^{n-(n_i-\gamma)} \lambda\binom{\lambda-k-1}{\gamma-j-1}\binom{n-\lambda}{n_i-\gamma}$$

$$= \binom{n+1}{n_i+1}\left[\frac{(n+2)\gamma + n-n_i}{n_i+2}\right].$$

Proof:

i) First, we want to evaluate: $$\sum_{\lambda=k+\gamma-j}^{n-(n_i-\gamma)} \lambda\binom{\lambda-k-1}{\gamma-j-1}\binom{n-\lambda}{n_i-\gamma} \qquad (1)$$

Observe:

$$(1-x)^{-(n_i-\gamma+1)} = \sum_{p=n_i-\gamma}^{\infty} \binom{p}{n_i-\gamma} x^{p-(n_i-\gamma)} \tag{2}$$

and $$(1-x)^{-(\gamma-j)}\, x^{\gamma+k-j} = \sum_{m=\gamma-j-1}^{\infty} \binom{m}{\gamma-j-1} x^{m+k+1} \tag{3}$$

Let us differentiate (3):

$$(\gamma-j)\,(1-x)^{-(\gamma-j+1)} x^{\gamma+k-j} + (1-x)^{-(\gamma-j)} (\gamma+k-j) x^{\gamma+k-j-1}$$

$$= \sum_{m=\gamma-j-1}^{\infty} \binom{m}{\gamma-j-1} (m+k+1)\, x^{m+k} \tag{4}$$

Taking $p = n-\lambda$ in (2) and replacing m by $\lambda-k-1$ in (4) and then taking their product, we get:

$$\sum_{\lambda=\gamma-j+k}^{\infty} \lambda \binom{\lambda-k-1}{\gamma-j-1} \binom{n-\lambda}{n_i-\gamma}\, x^{n-n_i+\gamma-1}$$

Thus we are looking for the coefficient of $x^{n-n_i+\gamma-1}$ in the product of (2) with (4), which is:

$$(\gamma-j)x^{\gamma+k-j}\,(1-x)^{-(n_i-j+2)} + (\gamma+k-j)x^{\gamma+k-j-1}(1-x)^{-(n_i-j+1)}$$

$$= (\gamma-j) \sum_{q=n_i-j+1}^{\infty} \binom{q}{n_i-j+1} x^{q-n_i+\gamma-1+k} + (\gamma+k-j) \sum_{q=n_i-j}^{\infty} \binom{q}{n_i-j} x^{q-n_i-1+\gamma+k}$$

to find the coefficient of $x^{n-n_i+\gamma-1}$ let $q = n-k$, and so

$$\sum_{\lambda=\gamma-j+k}^{n-(n_i-\gamma)} \lambda \binom{\lambda-k-1}{\gamma-j-1} \binom{n-\lambda}{n_i-\gamma} = (\gamma-j)\binom{n-k}{n_i-j+1} + (\gamma+k-j)\binom{n-k}{n_i-j}.$$

ii) Now

$$\sum_{k=j}^{n-n_i-j} \cdot \binom{k}{j} \sum_{\lambda=k+\gamma-j}^{n-(n_i-\gamma)} \lambda \binom{\lambda-k-1}{\gamma-j-1}\binom{n-\lambda}{n_i-\gamma}$$

$$= \sum_{k=j}^{n-(n_i-j)} (\gamma-j)\binom{k}{j}\left[\binom{n-k}{n_i-j+1} + \binom{n-k}{n_i-j}\right] + \sum_{k=j}^{n-(n_i-j)} k\binom{k}{j}\binom{n-k}{n_i-j}.$$

But repeating Prop. 1 we find that;

$$\sum_{k=j}^{n-(n_i-j)} \binom{k}{j}\binom{n-k}{n_i-j+1} = \binom{n+1}{n_i+2}$$

Thus by Prop. 1' and 3 we get

$$(\gamma-j)\left[\binom{n+1}{n_i+2} + \binom{n+1}{n_i+1}\right] + j\binom{n+1}{n_i+1} + (j+1)\binom{n+1}{n_i+2}$$

$$= \gamma\binom{n+1}{n_i+1} + (\gamma+1)\binom{n+1}{n_i+2} = \binom{n+1}{n_i+1}\left[\gamma+(\gamma+1)\frac{(n-n_i)}{(n_i+2)}\right]$$

$$= \binom{n+1}{n_i+1}\left[\frac{(n+2)\gamma + n-n_i}{n_i+2}\right].$$

Note 3: To compute $P(R_{i\ j}=k,\ R_{i\ \gamma}=\lambda,\ (j<\gamma))$ we remark that we may choose among $(k-1)$ values to fill the $(j-1)$ first places, among $(\lambda-1-k)$ values to fill the $(\gamma-1-j)$ places between k and λ and among $(n-\lambda)$ values to fill the last $(n_i-\gamma)$ places. So we have, for $j < \gamma$

$$E(R_{i\ j}\ R_{i\ \gamma}) = \frac{1}{\binom{n}{n_i}} \sum_{k=j}^{n-(n_i-j)} \sum_{\lambda=k+\gamma-j}^{n-(n_i-\gamma)} k\lambda\binom{k-1}{j-1}\binom{\lambda-1-k}{\gamma-1-j}\binom{n-\lambda}{n_i-\gamma}$$

$$= \frac{1}{\binom{n}{n_i}} \sum_{k=j}^{n-(n_i-j)} j\binom{k}{j} \sum_{\lambda=k+\gamma-j}^{n-(n_i-\gamma)} \lambda \binom{\lambda-1-k}{\gamma-1-j}\binom{n-\lambda}{n_i-\gamma}.$$

$$= \frac{j}{\binom{n}{n_i}} \binom{n+1}{n_i+1}\left[\frac{(n+2)\gamma+n-n_i}{n_i+2}\right] \text{ by prop. 4}$$

$$= j\,\frac{(n+1)}{(n_i+1)(n_i+2)}[(n+2)\gamma+n-n_i].$$

Also

$$E(X_i^2) = \frac{1}{(n+1)^2} \quad E\Big(\sum_{j=1}^{n_i} \frac{j\,R_{i\,j}}{n_i+1}\Big)^2$$

where

$$E\Big(\sum_{j=1}^{n_i} \frac{j\,R_{i\,j}}{n_i+1}\Big)^2 = E\Big\{\sum_{j=1}^{n_i} \frac{j^2\,R_{i\,j}^2}{(n_i+1)^2} + \sum_{j<\gamma} j\,\gamma\,\frac{R_{i\,j}\,R_{i\,\gamma}}{(n_i+1)^2}$$

$$+ \sum_{\gamma<j} j\,\gamma\,\frac{R_{i\,j}\,R_{i\,\gamma}}{(n_i+1)^2}\Big\}$$

$$= \sum_{j=1}^{n_i} \frac{j^3}{(n_i+1)^3}\,\frac{n+1}{n_i+2}\,[n-n_i+j(n+2)]$$

$$+ \sum_{j<\gamma} \frac{j^2\gamma}{(n_i+1)^3}\,\frac{n+1}{n_i+2}\,[n-n_i+\gamma(n+2)] + \sum_{\gamma<j} \frac{j\,\gamma^2}{(n_i+1)^3}\cdot\frac{n+1}{n_i+2}[n-n_i+j(n+2)]$$

$$= \frac{n+1}{(n_i+1)^3(n_i+2)} \begin{array}{l} \{n-n_i[\sum_{j=1}^{n_i} j^3 + \sum_{j<\gamma} j^2\gamma + \sum_{\gamma<j} j\,\gamma^2] \\ +(n+2)[\sum_{j=1}^{n_i} j^4 + \sum_{j<\gamma}(j\gamma)^2 + \sum_{\gamma<j}(j\,\gamma)^2]\} \end{array}$$

Observe that

$$\sum_{j=1}^{n_i} j^4 + \sum_{j<\gamma}(j\,\gamma)^2 + \sum_{\gamma<j}(j\,\gamma)^2 = (\sum_{j=1}^{n_i} j^2)^2 \quad \text{and,}$$

$$\sum_{j<\gamma} j^2\gamma = \sum_{j=1}^{n_i-1} j^2 \cdot \sum_{\gamma=j+1}^{n_i} \gamma = \sum_{j=1}^{n_i-1} j^2\{\frac{n_i(n_i+1)}{2} - \frac{j(j+1)}{2}\}$$

$$= \frac{n_i(n_i+1)}{2} \frac{(n_i-1)n_i(2n_i-1)}{6} - \frac{1}{2}\sum_{j=1}^{n_i-1} j^4 - \frac{1}{2}\sum_{j=1}^{n_i-1} j^3 .$$

$$= \frac{(n_i-1)n_i^2(n_i+1)(2n_i-1)}{12} - \frac{1}{2} \frac{(n_i-1)n_i(2n_i-1)(3n_i^2-3n_i-1)}{30} - \frac{1}{2} \frac{(n_i-1)^2 n_i^2}{4}$$

$$= \frac{1}{120}(n_i-1)n_i[10n_i(n_i+1)(2n_i-1)-2(2n_i-1)(3n_i^2-3n_i-1)-15(n_i-1)n_i]$$

$$= \frac{1}{120}(n_i-1)n_i[10(2n_i^3+n_i^2-n_i)-2(6n_i^3-9n^2+n_i+1)-15(n_i^2-n_i)]$$

$$= \frac{1}{120}n_i(n_i-1)[8n_i^3+13n_i^2+3n_i-2]$$

$$E(\sum_{j=1}^{n_i} \frac{j\,R_{ij}}{n_i+1})^2 = \frac{n+1}{(n_i+1)^3(n_i+2)} \{(n-n_i)[\frac{n_i^2(n_i+1)^2}{4}$$

$$+ \frac{1}{60}n_i(n_i-1)(8n_i^3+13n_i^2+3n_1-2)] + (n+2)[\frac{n_i^2(n_i+1)^2\;(2n_i+1)^2}{36}]\}$$

$$= \frac{n+1}{(n_i+1)^3(n_i+2)} \{(n-n_i)[\frac{n_i^2(n_i+1)^2}{4} + \frac{1}{60} n_i(n_i-1)(n_i+1)(8n_i^2+5n_i-2)]$$

$$+ (n+2)[\frac{n_i^2(n_i+1)^2(2n_i+1)^2}{36}]\}$$

$$= \frac{(n+1)n_i}{(n_i+1)^2(n_i+2)} \{(n-n_i)[\frac{n_i(n_i+1)}{4} + \frac{1}{60}(n_i-1)(8n_i^2+5n_i-2)]$$

$$+ \frac{(n+2)n_i(n_i+1)(2n_i+1)^2}{36}\} .$$

Furthermore
Var X_i

$$= \frac{1}{(n+1)^2} \frac{(n+1)\ n_i}{(n_i+1)^2(n_i+2)} \{(n-n_i)[\frac{n_i(n_i+1)}{4} + \frac{1}{60}(n_i-1)(8n_i^2+5n_i-2)]$$

$$+ \frac{(n+2)n_i(n_i+1)(2n_i+1)^2}{36}\}$$

$$- \frac{n_i^2(2n_i+1)^2}{36(n_i+1)^2}$$

$$= \frac{1}{(n+1)} \frac{n_i}{(n_i+1)^2(n_i+2)} \{\frac{(n-n_i)}{60}[15n_i(n_i+1) + (n_i-1)(8n_i^2+5n_i-2)]\}.$$

$$+ \frac{n_i^2(2n_i+1)^2}{36(n_i+1)^2} [\frac{(n+2)(n_i+1)}{(n_i+2)(n+1)} - 1]$$

$$= \frac{1}{60} \frac{n_i(n-n_i)}{(n+1)(n_i+1)^2(n_i+2)} [15n_i^2+15n_i+8n_i^3+5n_i^2-2n_i-8n_i^2-5n_i+2]$$

$$+ \frac{n_i^2\ (2n_i+1)^2}{36(n_i+1)^2} [\frac{(n\ n_i+n+2n_i+2)-(n\ n_i+2n+n_i+2)}{(n_i+2)(n+1)}]$$

$$= \frac{1}{30}\,\frac{n_i(n-n_i)}{(n+1)(n_i+1)^2(n_i+2)}\,[4n_i^3+6n_i^2+4n_i+1]$$

$$-\;\frac{n_i^2(2n_i+1)^2}{36\;(n_i+1)^2}\;\frac{(n-n_i)}{(n_i+2)(n+1)}$$

$$= \frac{1}{30}\;\frac{n_i(n-n_i)}{(n+1)(n_i+1)^2(n_i+2)}\;(2n_i+1)(2n_i^2+2n_i+1)$$

$$-\;\frac{n_i^2(2n_i+1)^2\;(n-n_i)}{36\;(n+1)(n_i+1)^2(n_i+2)}$$

$$= \frac{n_i(n-n_i)(2n_i+1)}{(n+1)(n_i+1)^2(n_i+2)}\;\left[\frac{(2n_i^2+2n_i+1)}{30}\;-\;\frac{n_i(2n_i+1)}{36}\;\right]$$

$$= \frac{n_i(n-n_i)(2n_i+1)}{(n+1)(n_i+1)^2(n_i+2)}\;\frac{1}{180}\;[(12n_i^2+12n_i+6)-(10n_i^2+5n_i)]$$

$$= \frac{n_i(n-n_i)(2n_i+1)}{180(n+1)(n_i+1)^2(n_i+2)}\;[2n_i^2+7n_i+6]$$

$$= \frac{n_i(n-n_i)(2n_i+1)}{180(n+1)(n_i+1)^2(n_i+2)}\;(n_i+2)(2n_i+3)$$

$$= \frac{n_i(n-n_i)(2n_i+1)(2n_i+3)}{180\;(n+1)(n_i+1)^2}$$

Proposition 5:

a) $$\sum_{j=0}^{k}\sum_{\ell=0}^{k-j} j^2\binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$= k\binom{n-1}{n_i-1}\binom{n-n_i}{n_\beta} + k(k-1)\binom{n-2}{n_i-2}\binom{n-n_i}{n_\beta}$$

b) $$\sum_{j=0}^{k}\sum_{\ell=0}^{k-j} j^2\ell\binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$= k(k-1)\binom{n-2}{n_i-1}\binom{n-n_i-1}{n_\beta-1} + k(k-1)(k-2)\binom{n-3}{n_i-2}\binom{n-n_i-1}{n_\beta-1}$$

c) $$\sum_{j=0}^{k}\sum_{\ell=0}^{k-j} j^2\ell^2\binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$= k(k-1)\binom{n-2}{n_i-1}\binom{n-n_i-1}{n_\beta-1}$$

$$+ k(k-1)(k-2)\left\{\binom{n-3}{n_i-2}\binom{n-n_i-1}{n_\beta-1}+\binom{n-3}{n_i-1}\binom{n-n_i-2}{n_\beta-2}\right\}$$

$$+ k(k-1)(k-2)(k-3)\binom{n-4}{n_i-2}\binom{n-n_i-2}{n_\beta-2}.$$

Proof:

We know that

$$(1+x+y)^k = \sum_{j=0}^{k}\sum_{\ell=0}^{k-j}\binom{k}{j}\binom{k-j}{\ell}x^j y^\ell \qquad (1)$$

also $$(1+x+y)^{n-k} = \sum_{\theta=0}^{n-k} \sum_{\psi=0}^{n-k-\theta} \binom{n-k}{\theta}\binom{n-k-\theta}{\psi} x^{\theta} y^{\psi} \qquad (2)$$

taking the derivative of (1) w.r. to x and multiplying by x:

$$kx(1+x+y)^{k-1} = \sum_{j=0}^{k} \sum_{\ell=0}^{k-j} j\binom{k}{j}\binom{k-j}{\ell} x^{j} y^{\ell} \text{ .}$$

differentiate again w.r. to x:

$$k(1+x+y)^{k-1} + k(k-1)x(1+x+y)^{k-2} = \sum_{j=0}^{k} \sum_{\ell=0}^{k-j} j^{2}\binom{k}{j}\binom{k-j}{\ell} x^{j-1} y^{\ell} \qquad (3)$$

differentiate now w.r. to y and multiply by y:

$$k(k-1)y(1+x+y)^{k-2} + k(k-1)(k-2)xy(1+x+y)^{k-3}$$

$$= \sum_{j=0}^{k} \sum_{\ell=0}^{k-j} \ell\, j^{2} \binom{k}{j}\binom{k-j}{\ell} x^{j-1} y^{\ell} \qquad (4)$$

differentiate again w.r. to y:

$$k(k-1)(1+x+y)^{k-2} + k(k-1)(k-2)(x+y)(1+x+y)^{k-3}$$

$$+ k(k-1)(k-2)(k-3)\; xy(1+x+y)^{k-4} = \sum_{j=0}^{k} \sum_{\ell=0}^{k-j} \ell^{2} j^{2} \binom{k}{j}\binom{k-j}{\ell} x^{j-1} y^{\ell-1} \qquad (5)$$

taking in (2) $\psi = n_{\beta}-\ell$ and $\theta = n_{i}-j$ and multiplying it by (3), (4) and (5) we get:

$$\sum_{j=0}^{k} \sum_{\ell=0}^{k-j} j^{2}\binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_{i}-j}\binom{n-k-n_{i}+j}{n_{\beta}-\ell} x^{n_{i}-1} y^{n_{\beta}}$$

$$= \sum_{j=0}^{k} \sum_{\ell=0}^{k-j} j^{2}\binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_{\beta}-\ell}\binom{n-k-n_{\beta}+\ell}{n_{i}-j} x^{n_{i}-1} y^{n_{\beta}} \qquad (6)$$

$$\sum_{j=0}^{k}\sum_{\ell=0}^{k-j} \ell j^2 \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_i-j}\binom{n-n-n_i+j}{n_\beta-\ell} x^{n_i-1} y^{n_\beta}$$

$$= \sum_{j=0}^{k}\sum_{\ell=0}^{k-j} \ell j^2 \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-k-n_\beta+\ell}{n_i-j} x^{n_i-1} y^{n_\beta} \qquad (7)$$

and

$$\sum_{j=0}^{k}\sum_{\ell=0}^{k-j} \ell^2 j^2 \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_i-j}\binom{n-k-n_i+j}{n_\beta-\ell} x^{n_i-1} y^{n_\beta-1}$$

$$= \sum_{j=0}^{k}\sum_{\ell=0}^{k-j} \ell^2 j^2 \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-k-n_\beta+\ell}{n_i-j} x^{n_i-1} y^{n_\beta-1} \qquad (8)$$

Thus for a) we want the coefficient of $x^{n_i-1} y^{n_\beta}$ in the product of (2) with (3) which is:

$$k(1+x+y)^{n-1} + k(k-1)x(1+x+y)^{n-2}$$

$$= k \sum_{\rho=0}^{n-1}\sum_{\sigma=0}^{n-1-\rho} \binom{n-1}{\rho}\binom{n-1-\rho}{\sigma} x^\rho y^\sigma + k(k-1)\sum_{\tau=0}^{n-2}\sum_{\sigma=0}^{n-2-\tau} \binom{n-2}{\tau}\binom{n-2-\tau}{\sigma} x^{\tau+1} y^\sigma$$

to find the coefficient of $x^{n_i-1} y^{n_\beta}$ let $\rho = n_i-1$, $\sigma = n_\beta$ and $\tau = n_i-2$ so we obtain,

$$k\binom{n-1}{n_i-1}\binom{n-n_i}{n_\beta} + k(k-1)\binom{n-2}{n_i-2}\binom{n-n_i}{n_\beta}.$$

Similarly for b) we want the coefficient of $x^{n_i-1} y^{n_\beta}$ in the product of (2) with (4) which is:

$$k(k-1)y(1+x+y)^{n-2} + k(k-1)(k-2)xy(1+x+y)^{n-3}$$

$$= k(k-1)\sum_{\rho=0}^{n-2}\sum_{\sigma=0}^{n-2-\rho}\binom{n-2}{\rho}\binom{n-2-\rho}{\sigma}x^{\rho}y^{\sigma+1} + k(k-1)(k-2)\sum_{\tau=0}^{n-3}\sum_{\sigma=0}^{n-3-\tau}\binom{n-3}{\tau}\binom{n-3-\tau}{\sigma}x^{\tau+1}y^{\sigma+1}.$$

To find the coefficient of $x^{n_i-1}\ y^{n_\beta}$ let $\rho = n_i-1$, $\sigma = n_\beta-1$ and $\tau = n_i-2$; so we obtain:

$$k(k-1)\binom{n-2}{n_i-1}\binom{n-1-n_i}{n_\beta-1} + k(k-1)(k-2)\binom{n-3}{n_i-2}\binom{n-n_i-1}{n_\beta-1}$$

In the same way again for c) we want the coefficient of $x^{n_i-1}\ y^{n_\beta-1}$ in the product of (2) with (5) which is:

$$k(k-1)(1+x+y)^{n-2} + k(k-1)(k-2)(x+y)(1+x+y)^{n-3}$$
$$+ k(k-1)(k-2)(k-3)xy(1+x+y)^{n-4} = k(k-1)\sum_{\rho=0}^{n-2}\sum_{\sigma=0}^{n-2-\rho}\binom{n-2}{\rho}\binom{n-2-\rho}{\sigma}x^{\rho}y^{\sigma}$$

$$+ k(k-1)(k-2)\{\sum_{\tau=0}^{n-3}\sum_{\sigma=0}^{n-3-\tau}\binom{n-3}{\tau}\binom{n-3-\tau}{\sigma}x^{\tau+1}y^{\sigma}$$

$$+ \sum_{\rho=0}^{n-3}\sum_{\psi=0}^{n-3-\rho}\binom{n-3}{\rho}\binom{n-3-\rho}{\psi}x^{\rho}\ {}^{\psi+1}\}$$

$$+ k(k-1)(k-2)(k-3)\sum_{\tau=0}^{n-4}\sum_{\psi=0}^{n-4-\tau}\binom{n-4}{\tau}\binom{n-4-\tau}{\psi}x^{\tau+1}\ y^{\psi+1}$$

to find the coefficient of $x^{n_i-1}\ y^{n_\beta-1}$ let $\rho = n_i-1$, $\sigma = n_\beta-1$, $\tau = n_i-2$ and $\psi = n_\beta-2$, hence we get:

$$k(k-1)\binom{n-2}{n_i-1}\binom{n-n_i-1}{n_\beta-1}$$

$$+ k(k-1)(k-2)\{\binom{n-3}{n_i-2}\binom{n-n_i-1}{n_\beta-1} + \binom{n-3}{n_i-1}\binom{n-n_i-2}{n_\beta-2}\}$$

$$+ k(k-1)(k-2)(k-3)\binom{n-4}{n_i-2}\binom{n-n_i-2}{n_\beta-2}.$$

Note 4:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=1}^{n} (k-1)k = \frac{(n-1)(n)(n+1)}{3}$$

$$\sum_{k=1}^{n} (k-2)(k-1)k = \frac{(n-2)(n-1)(n)(n+1)}{4}$$

$$\sum_{k=1}^{n} (k-3)(k-2)(k-1)k = \frac{(n-3)(n-2)(n-1)n(n+1)}{5}$$

$$\sum_{k=1}^{n} (k-4)(k-3)(k-2)(k-1)k = \frac{(n-4)(n-3)(n-2)(n-1)n(n+1)}{6}$$

Proposition 6:

$$\sum_{j=1}^{n_i} \sum_{\gamma=1}^{n_\beta} \frac{j\gamma}{(n+1)^2(n_i+1)(n_\beta+1)}$$

$$\Bigg\{ \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{\gamma-1} \sum_{\lambda=k+\gamma-\ell}^{n-(n_\beta-\gamma)} k\lambda \binom{k-1}{j-1} \binom{k-j}{\ell} \binom{\lambda-1-k}{\gamma-1-\ell} \sum_{m=0}^{n_i-j} \binom{\lambda-k-(\gamma-\ell)}{m}$$

$$\binom{n-\lambda}{n_\beta-\gamma} \binom{n-\lambda-(n_\beta-\gamma)}{n_i-(j+m)}$$

$$+ \sum_{\lambda=\gamma}^{n-(n_\beta-\gamma)} \sum_{\ell=0}^{j-1} \sum_{k=\lambda+j-\ell}^{n-(n_i-j)} k\lambda \binom{\lambda-1}{\gamma-1} \binom{\lambda-\gamma}{\ell} \binom{k-1-\lambda}{j-1-\ell} \sum_{m=0}^{n_i-j} \binom{k-\lambda-(j-\ell)}{m}$$

$$\binom{n-k}{n_i-j} \binom{n-k-(n_i-j)}{n_\beta-(\gamma+m)} \Bigg\} \qquad \text{where } \beta \neq i$$

$$= \frac{1}{6} \frac{n_i n_\beta}{(n_i+1)(n_\beta+1)} \frac{1}{(n+1)^2} \frac{n!}{n_i! n_\beta! (n-n_i-n_\beta)!}$$

$$\{\frac{n^2}{6}(4n_i n_\beta+2n_i+2n_\beta+1) + \frac{n}{20}(24n_i n_\beta+11n_i+11n_\beta+4)$$

$$+ \frac{1}{60}(32n_i n_\beta+13n_i+13n_\beta+2)\}.$$

<u>Proof</u>: Let us consider the part of the expression where $k < \lambda$.

i) First, we prove: $$\sum_{m=0}^{n_i-j} \binom{\lambda-k-(\gamma-\ell)}{m} \binom{n-\lambda-(n_\beta-\gamma)}{n_i-(j+m)} = \binom{n+\ell-n_\beta-k}{n_i-j} \quad (1)$$

We know,

$$(1+x)^{\lambda-k-(\gamma-\ell)} = \sum_{m=0}^{\lambda-k-(\gamma-\ell)} \binom{\lambda-k-(\gamma-\ell)}{m} x^m \quad \text{and} \quad (2)$$

$$(1+x)^{n-\lambda-(n_\beta-\gamma)} = \sum_{\rho=0}^{n-\lambda-(n_\beta-\gamma)} \binom{n-\lambda-(n_\beta-\gamma)}{\rho} x^\rho \quad (3)$$

As before, taking $\rho = n_i - (j+m)$ in (3) and multiplying with (2) we get;

$$\sum_{m=0}^{\lambda-k-(\gamma-\ell)} \binom{\lambda-k-(\gamma-\ell)}{m} \binom{n-\lambda-(n_\beta-\gamma)}{n_i-(j+m)} x^{n_i-j}$$

Thus we are looking for the coefficient of x^{n_i-j} in the product of (2) and (3) which is:

$$(1+x)^{n-n_\beta-k+\ell} = \sum_{q=0}^{n-n_\beta-k+\ell} \binom{n-n_\beta-k+\ell}{q} x^q$$

to find coefficient of x^{n_i-j} let $q = n_i-j$ and so the required coefficient is: $\binom{n-n_\beta-k+\ell}{n_i-j}$.

ii) Now, let us sum over $\lambda\,(\lambda>k)$.

$$\sum_{\lambda=k+\gamma-\ell}^{n-(n_\beta-\gamma)} \lambda\binom{\lambda-1-k}{\gamma-1-\ell}\binom{n-\lambda}{n_\beta-\gamma} = (\gamma-\ell)\binom{n-k}{n_\beta-\ell+1} + (\gamma+k-\ell)\binom{n-k}{n_\beta-\ell}$$

by part i) of the proof of proposition 4 with j and n_i replaced by ℓ and n_β respectively.

But

$$(\gamma-\ell)\binom{n-k}{n_\beta-\ell+1} + (\gamma-\ell+k)\binom{n-k}{n_\beta-\ell} =$$

$$(\gamma-\ell)\left[\binom{n-k}{n_\beta-\ell+1} + \binom{n-k}{n_\beta-\ell}\right] + k\binom{n-k}{n_\beta-\ell} =$$

$$(\gamma-\ell)\binom{n-k}{n_\beta-\ell}\left[\frac{n-k-n_\beta+\ell}{n_\beta-\ell+1} + 1\right] + k\binom{n-k}{n_\beta-\ell} =$$

$$\binom{n-k}{n_\beta-\ell}\left[\frac{(\gamma-\ell)(n-k+1)+k(n_\beta-\ell+1)}{n_\beta-\ell+1}\right] =$$

$$\binom{n-k}{n_\beta-\ell}\left[\frac{\gamma(n-k+1)-\ell(n+1)+k(n_\beta+1)}{n_\beta-\ell+1}\right]$$

iii) So we rewrite the part of the expression where $k<\lambda$ as:

$$\sum_{j=1}^{n_i}\sum_{\gamma=1}^{n_\beta}\frac{j\gamma}{(n+1)^2(n_i+1)(n_\beta+1)}\left\{\sum_{k=j}^{n-(n_i-j)}\sum_{\ell=0}^{\gamma-1} k\binom{k-1}{j-1}\binom{k-j}{\ell}\binom{n-n_\beta-k+\ell}{n_i-j}\right.$$

$$\left.\times\binom{n-k}{n_\beta-\ell}\left[\frac{\gamma(n-k+1)-\ell(n+1)+k(n_\beta+1)}{n_\beta-\ell+1}\right]\right\}.$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \{ \sum_{\gamma=\ell+1}^{n_\beta} [\frac{\gamma^2(n-k+1)+\gamma[k(n_\beta+1)-\ell(n+1)]}{n_\beta-\ell+1}] \frac{1}{(n_\beta+1)} \}$$

$$\times \frac{j}{(n+1)^2(n_i+1)} j \binom{k}{j} \binom{k-j}{\ell} \binom{n-n_\beta-k+\ell}{n_i-j} \binom{n-k}{n_\beta-\ell}$$

Note that $(\gamma > \ell)$.

We may sum over γ

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2}{(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j} \binom{k-j}{\ell} \binom{n-k}{n_\beta-\ell} \binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \{(n-k+1)[\frac{n_\beta(n_\beta+1)(2n_\beta+1)-\ell(\ell+1)(2\ell+1)}{6}]$$

$$+ (k(n_\beta+1)-\ell(n+1))[\frac{n_\beta(n_\beta+1)-\ell(\ell+1)}{2}]\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2}{(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j} \binom{k-j}{\ell} \binom{n-k}{n_\beta-\ell} \binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \frac{1}{6}\{(n-k+1)(n_\beta-\ell)[2n_\beta^2+n_\beta(3+2\ell)+1+2\ell^2+3\ell]$$

$$+ 3[k(n_\beta+1)-\ell(n+1)](n_\beta-\ell)(n_\beta+\ell+1)\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2 \quad (n_\beta-\ell)}{(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j} \binom{k-j}{\ell} \binom{n-k}{n_\beta-\ell} \binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \frac{1}{6}\{(n-k+1)[2n_\beta^2+n_\beta(3+2\ell)+1+2\ell^2+3\ell] + 3(n_\beta+\ell+1)[k(n_\beta+1)-\ell(n+1)]\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)}{6(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times\{(n-k+1)[2(n_\beta+1)^2 + n_\beta(2\ell-1)-1+2\ell^2+3\ell]$$

$$+ 3((n_\beta+1)+\ell)[k(n_\beta+1)-\ell(n+1)]\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)}{6(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \{(n-k+1)[2(n_\beta+1)^2+(n_\beta+1)(2\ell-1)+2\ell^2+\ell]$$

$$+ 3((n_\beta+1)+\ell)\ [k(n_\beta+1)-\ell(n+1)]\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)}{6(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\{(n_\beta+1)^2[2(n-k+1)+3k] + (n_\beta+1)[(2\ell-1)(n-k+1)-3\ell(n+1)+3\ell k] +$$

$$(2\ell^2+\ell)(n-k+1)-3\ell^2(n+1)\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)}{6(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}$$

$$\times \binom{n-n_\beta-k+\ell}{n_i-j} \begin{Bmatrix} (n_\beta+1)^2(2n+k+2) \\ +(n_\beta+1)(n-k+1)(-\ell-1) \\ +(n+1)(-\ell^2+\ell)-k(2\ell^2+\ell) \end{Bmatrix}$$

Factor $(n_\beta-\ell+1)$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)(n_\beta-\ell+1)}{6(n_\beta-\ell+1)(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}$$

$$\times \binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}\{(n_\beta+1)(2n+k+2) + (n+1)(\ell-1) + k(2\ell+1)\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)}{6(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \{\ell(n+1+2k)-(n+1)+k+n_\beta(2n+k+2)+2n+k+2\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2(n_\beta-\ell)}{6(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\{\ell(n+1+2k) + n_\beta[2n+2+k] + n+1+2k\}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2}{6(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \begin{Bmatrix} -\ell^2(n+1+2k) + \ell n_\beta(n+1+2k) - \ell n_\beta(2n+2+k) \\ +n_\beta^2(2n+2+k) - \ell(n+1+2k) + n_\beta(n+1+2k) \end{Bmatrix}$$

$$= \sum_{j=1}^{n_i} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{n_\beta-1} \frac{j^2}{6(n_\beta+1)(n+1)^2(n_i+1)} \binom{k}{j}\binom{k-j}{\ell}\binom{n-k}{n_\beta-\ell}\binom{n-n_\beta-k+\ell}{n_i-j}$$

$$\times \{-\ell^2(n+1+2k) + \ell[n_\beta(-n-1+k)-(n+1+2k)] + n_\beta[n_\beta(2n+2+k) + (n+1+2k)]\}$$

$$= \sum_{k=1}^{n-1} \frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)}$$

$$\times \{-(n+1+2k) \left[\begin{array}{l} k(k-1)\binom{n-2}{n_i-1}\binom{n-n_i-1}{n_\beta-1} + \\ + k(k-1)(k-2)\left[\binom{n-3}{n_i-2}\binom{n-n_i-1}{n_\beta-1} + \binom{n-3}{n_i-1}\binom{n-n_i-2}{n_\beta-2}\right] \\ + k(k-1)(k-2)(k-3)\binom{n-4}{n_i-2}\binom{n-n_i-2}{n_\beta-2} \end{array} \right]$$

$$- (n_\beta(n+1-k)+n+1+2k) \left[\begin{array}{l} k(k-1)\binom{n-2}{n_i-1}\binom{n-n_i-1}{n_\beta-1} \\ +k(k-1)(k-2)\binom{n-3}{n_i-2}\binom{n-n_i-1}{n_\beta-1} \end{array} \right]$$

$$+ n_\beta[n_\beta(2(n+1)+k)+n+1+2k] \left[\begin{array}{l} k\binom{n-1}{n_i-1}\binom{n-n_i}{n_\beta} \\ +k(k-1)\binom{n-2}{n_i-2}\binom{n-n_i}{n_\beta} \end{array} \right] \}$$

by proposition 5.

$$= \sum_{k=1}^{n-1} \frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)} \binom{n}{n_i}\binom{n-n_i}{n_\beta}$$

$$\times \{-(n+1+2k) \left[\begin{array}{l} k(k-1)\dfrac{n_i n_\beta}{n(n-1)} + \dfrac{k(k-1)(k-2)}{n(n-1)(n-2)}[(n_i-1)n_i n_\beta + n_i n_\beta(n_\beta-1)] \\ + \dfrac{k(k-1)(k-2)(k-3)}{n(n-1)(n-2)(n-3)} n_i(n_i-1)n_\beta(n_\beta-1) \end{array} \right]$$

$$-(n_\beta(n+1-k)+n+1+2k)[\frac{k(k-1)}{n(n-1)}\, n_i n_\beta + \frac{k(k-1)(k-2)}{n(n-1)(n-2)}(n_i-1)n_i n_\beta]$$

$$+n_\beta[n_\beta(2(n+1)+k)+n+1+2k][\frac{k}{n}\, n_i + \frac{k(k-1)}{n(n-1)}\, n_i(n_i-1)]\}.$$

$$=\sum_{k=1}^{n-1} \frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)} \binom{n}{n_i}\binom{n-n_i}{n_\beta}$$

$$\times \{\frac{k}{n}\, n_i n_\beta[n_\beta(2(n+1)+k)+n+1+2k]$$

$$+ \frac{k(k-1)}{n(n-1)}\, n_i n_\beta[-(n+1+2k)-(n_\beta(n+1-k)+n+1+2k)+(n_i-1)(n_\beta(2n+2+k)+n+1+2k)]$$

$$+ \frac{k(k-1)(k-2)}{n(n-1)(n-2)}\, n_i n_\beta[-(n+1+2k)(n_i-1+n_\beta-1)-(n_i-1)(n_\beta(n+1-k)+n+1+2k)]$$

$$- \frac{k(k-1)(k-2)(k-3)}{n(n-1)(n-2)}\, n_i n_\beta(n_i-1)(n_\beta-1)(n+1+2k)\}.$$

$$=\sum_{k=1}^{n-1} \frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)} \binom{n}{n_i}\binom{n-n_i}{n_\beta}$$

$$\times \{\frac{k}{n}\, n_i n_\beta[(n+1)(2n_\beta+1)+k(n_\beta+2)]$$

$$+ \frac{k(k-1)}{n(n-1)}\, n_i n_\beta\left[\begin{matrix} k(-2+n_\beta-2+ (n_i-1)(n_\beta+2)) \\ -(n+1)-n_\beta(n+1)-(n+1)+(n_i-1)(n+1)(2n_\beta+1) \end{matrix}\right]$$

$$+\frac{k(k-1)(k-2)}{n(n-1)(n-2)}\quad n_i n_\beta \left[\begin{matrix} k[-2(n_i+n_\beta-2))-(n_i-1)(-n_\beta+2)] \\ -(n+1)(n_i+n_\beta-2)-(n_i-1)(n+1)(n_\beta+1)\end{matrix}\right]$$

$$-\frac{k(k-1)(k-2)(k-3)}{n(n-1)(n-2)(n-3)}\quad n_i n_\beta (n_i-1)(n_\beta-1)(n+1+2k)\}.$$

$$=\sum_{k=1}^{n-1}\frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)}\binom{n}{n_i}\binom{n-n_i}{n_\beta}$$

$$\times\ \{\frac{k}{n}\,n_i n_\beta (n+1)(2n_\beta+1) + \frac{k^2}{n}\,n_i n_\beta (n_\beta+2) + \frac{k^2(k-1)}{n(n-1)}\,n_i n_\beta (n_i n_\beta+2n_i-6)$$

$$+\frac{k(k-1)}{n(n-1)}\,n_i n_\beta (n+1)(2n_i n_\beta-3n_\beta+n_i-3) + \frac{k^2(k-1)(k-2)}{n\ (n-1)(n-2)}\,n_i n_\beta (n_i n_\beta-4n_i-3n_\beta+6)$$

$$-\frac{k(k-1)(k-2)}{n(n-1)(n-2)}\,n_i n_\beta (n+1)(n_i+n_\beta-2+(n_i-1)(n_\beta+1))$$

$$-\frac{2k^2(k-1)(k-2)(k-3)}{n(n-1)(n-2)(n-3)}\quad n_i n_\beta (n_i-1)(n_\beta-1)$$

$$-\frac{k(k-1)(k-2)(k-3)}{n(n-1)(n-2)(n-3)}\quad n_i n_\beta (n_i-1)(n_\beta-1)(n+1)\}$$

$$=\sum_{k=1}^{n-1}\frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)}\binom{n}{n_i}\binom{n-n_i}{n_\beta}\frac{n_i n_\beta}{n}$$

$$\times\ \{k\ (n+1)(2n_\beta+1) + k^2(n_\beta+2)$$

$$+ [(k+1)k(k-1)-k(k-1)] \frac{(n_i n_\beta+2n_i-6)}{n-1}$$

$$+ k(k-1)\frac{(n+1)}{(n-1)} (2n_i n_\beta-3n_\beta+n_i-3)$$

$$+ [(k+1)k(k-1)(k-2)-k(k-1)(k-2)]\frac{(n_i n_\beta-4n_i-3n_\beta+6)}{(n-1)(n-2)}$$

$$- k(k-1)(k-2) \frac{(n+1)}{(n-1)(n-2)} (n_i n_\beta+2n_i-3)$$

$$- 2[(k+1)k(k-1)(k-2)(k-3)-k(k-1)(k-2)(k-3)] \frac{(n_i-1)(n_\beta-1)}{(n-1)(n-2)(n-3)}$$

$$- k(k-1)(k-2)(k-3)\frac{(n_i-1)(n_\beta-1)(n+1)}{(n-1)(n-2)(n-3)} \}.$$

We sum now over k using note 4.

$$= \frac{1}{6(n+1)^2(n_\beta+1)(n_i+1)} \binom{n}{n_i}\binom{n-n_i}{n_\beta} \frac{n_i\, n_\beta}{n}$$

$$\times \{\frac{(n+1)(2n_\beta+1)(n-1)n}{2} + \frac{(n_\beta+2)(n-1)n(2n-1)}{6}$$

$$+ \frac{(n_i n_\beta+2n_i-6)}{(n-1)} [\frac{(n+1)n(n-1)(n-2)}{4} - \frac{n(n-1)(n-2)}{3}]$$

$$+\ \frac{(n+1)(2n_i n_\beta - 3n_\beta + n_i - 3)}{(n-1)} \quad \frac{n(n-1)(n-2)}{3}$$

$$+\ \frac{(n_i n_\beta - 4n_i - 3n_\beta + 6)}{(n-1)(n-2)} \left[\frac{(n+1)n(n-1)(n-2)(n-3)}{5} - \frac{n(n-1)(n-2)(n-3)}{4}\right]$$

$$-\ \frac{(n+1)(n_i n_\beta + 2n_i - 3)}{(n-1)(n-2)} \quad \frac{n(n-1)(n-2)(n-3)}{4}$$

$$-\ \frac{2(n_i-1)(n_\beta-1)}{(n-1)(n-2)(n-3)} \left[\frac{(n+1)n(n-1)(n-2)(n-3)(n-4)}{6} - \frac{n(n-1)(n-2)(n-3)(n-4)}{5}\right]$$

$$-\ \frac{(n_i-1)(n_\beta-1)(n+1)}{(n-1)(n-2)(n-3)} \left[\frac{n(n-1)(n-2)(n-3)(n-4)}{5}\right] \Big\}.$$

$$= \frac{1}{6}\,\frac{n_i n_\beta}{(n+1)^2(n_\beta+1)(n_i+1)} \quad \frac{n!}{n_i!\,n_\beta!\,(n-n_i-n_\beta)!}$$

$$\times \Big\{\frac{(n-1)(n+1)(2n_\beta+1)}{2} + \frac{(n-1)(2n-1)(n_\beta+2)}{6} + (n_i n_\beta + 2n_i - 6)\frac{(n-2)}{12}(3n-1)$$

$$+\ (2n_i n_\beta - 3n_\beta + n_i - 3)\frac{(n+1)(n-2)}{3}$$

$$+\ (n_i n_\beta - 4n_i - 3n_\beta + 6)\frac{(n-3)}{20}(4n-1)$$

$$-\ (n_i n_\beta + 2n_i - 3)\frac{(n+1)(n-3)}{4}$$

$$-\ (n_i-1)(n_\beta-1)\frac{(n-4)}{15}(5n-1)$$

$$-\ (n_i-1)(n_\beta-1)\frac{(n+1)(n-4)}{5}\Big\}.$$

iv) To get the part of the expression where $\lambda < k$, just interchange n_i and n_β.

So adding these, we obtain:

$$\frac{1}{6}\ \frac{n_i n_\beta}{(n+1)^2(n_\beta+1)(n_i+1)}\ \frac{n!}{n_i!\,n_\beta!\,(n-n_i-n_\beta)!}$$

$$\times \{\frac{(n-1)(n+1)}{2}(2n_\beta+2n_i+2) + \frac{(n-1)(2n-1)(n_\beta+n_i+4)}{6}$$

$$+ \frac{(n-2)(3n-1)}{12}(2n_in_\beta+2n_i+2n_\beta-12) + \frac{(n-2)(n+1)}{3}(4n_in_\beta-2n_\beta-2n_i-6)$$

$$+ \frac{(n-3)(4n-1)}{20}(2n_in_\beta-7n_i-7n_\beta+12) - \frac{(n+1)(n-3)}{4}(2n_in_\beta+2n_i+2n_\beta-6)$$

$$- 2\,\frac{(n-4)(5n-1)}{15}(n_i-1)(n_\beta-1) - 2\,\frac{(n+1)(n-4)}{5}(n_i-1)(n_\beta-1)\}$$

$$= \frac{1}{6}\ \frac{n_i n_\beta}{(n_i+1)(n_\beta+1)(n+1)^2}\ \frac{n!}{n_i!\ n_\beta!\,(n-n_i-n_\beta)!}$$

$$\times \{(n^2-1)((n_i+n_\beta)+1) + (2n^2-3n+1)\frac{((n_i+n_\beta)+4)}{6}$$

$$+ (3n^2-7n+2)\frac{(n_in_\beta+(n_i+n_\beta)-6)}{6} + \frac{2}{3}(n^2-n-2)(2n_in_\beta-(n_\beta+n_i)-3)$$

$$+ (4n^2-13n+3)\,\frac{(2n_in_\beta-7(n_i+n_\beta)+12)}{20} - (n^2-2n-3)\frac{(n_in_\beta+(n_i+n_\beta)-3)}{2}$$

$$- [\frac{2(5n^2-21n+4)+6(n^2-3n-4)}{15}](n_in_\beta-(n_i+n_\beta)+1)\}.$$

Regrouping the terms:

$$= \frac{1}{6}\;\frac{n_in_\beta}{(n_i+1)(n_\beta+1)(n+1)^2}\;\frac{n!}{n_i!n_\beta!(n-n_i-n_\beta)!}$$

$$\times \{n^2[n_in\;(\frac{1}{2}+\frac{4}{3}+\frac{2}{5}-\frac{1}{2}-\frac{16}{15}) + (n_i+n_\beta)(1+\frac{1}{3}+\frac{1}{2}-\frac{2}{3}-\frac{7}{5}-\frac{1}{2}+\frac{16}{15})$$

$$+ (1+\frac{4}{3}-3-2+\frac{12}{5}+\frac{3}{2}-\frac{16}{15})]$$

$$+ n[n_in_\beta(-\frac{7}{6}-\frac{4}{3}-\frac{13}{10}+1+4) + (n_i+n_\beta)(-\frac{1}{2}-\frac{7}{6}+\frac{2}{3}+\frac{91}{20}+1-4)$$

$$+ (-2+7+2-\frac{39}{5}-3+4)]$$

$$+n_in_\beta(\frac{1}{3}-\frac{8}{3}+\frac{3}{10}+\frac{3}{2}+\frac{16}{15}) + (n_i+n_\beta)(-1+\frac{1}{6}+\frac{1}{3}+\frac{1}{4}-\frac{21}{20}+\frac{3}{2}-\frac{16}{15})$$

$$+ (-1+\frac{2}{3}-2+4+\frac{9}{5}-\frac{9}{2}+\frac{16}{15})\}.$$

$$= \frac{1}{6}\;\frac{n_in_\beta\qquad n!}{(n_i+1)(n_\beta+1)(n+1)^2\;n_i!n_\beta!(n-n_i-n_\beta)!}$$

$$\times \{n^2[\tfrac{2}{3} n_i n_\beta + \frac{(n_i+n_\beta)}{3} + \tfrac{1}{6}] + n[n_i n_\beta(\tfrac{6}{5}) + (n_i+n_\beta)\tfrac{11}{20} + \tfrac{1}{5}]$$

$$+ n_i n_\beta(\tfrac{8}{15}) + (n_i+n_\beta)(\tfrac{13}{60}) + \tfrac{1}{30}\}$$

$$= \frac{1}{6} \; \frac{n_i n_\beta}{(n_i+1)(n_\beta+1)(n+1)^2} \; \frac{n!}{n_i!\,n_\beta!\,(n-n_i-n_\beta)!}$$

$$\times \{\frac{n^2}{6}(4n_i n_\beta+2(n_i+n_\beta)+1) + \frac{n}{20}(24n_i n_\beta+11(n_i+n_\beta)+4)$$

$$+ \frac{1}{60}(32n_i n_\beta+13(n_i+n_\beta)+2)\}.$$

Note 5:

$$EX_iX_\beta = E\{\sum_{j=1}^{n_i} \frac{j\,R_{i\,j}}{(n+1)(n_i+1)} \cdot \sum_{\gamma=1}^{n_\beta} \frac{\gamma\,R_{\beta\,\gamma}}{(n+1)(n_\beta+1)}\}$$

$$= \frac{1}{(n+1)^2(n_i+1)(n_\beta+1)} \sum_{j=1}^{n_i} \sum_{\gamma=1}^{n_\beta} j\,\gamma\,E\,R_{i\,j}\,R_{\beta\,\gamma} .$$

For the part of the expression where $k < \lambda$; to compute $P(R_{i\,j} = k,\ R_{\beta\,\gamma} = \lambda$ for $i \neq \beta$ and $k < \lambda)$, remark to fill the $(j-1)$ places before the j^{th} place we may choose among $(k-1)$ values, there are now $(k-j)$ values smaller than k from which we may choose ℓ values to fill the first ℓ places before the γ^{th} place. Then remains $(\lambda-1-k)$ values smaller than λ but larger than k to fill the $(\lambda-1-\ell)$ remaining places before λ. Now there remains $(\lambda-k-(\gamma-\ell))$ values smaller than λ from which we take m values to fill m places immediately after k. There is now $(n-\lambda)$ values left from which we choose successively $(n_\beta-\gamma)$ and $(n_i-(j+m))$ values to fill the places after λ and after the $(j+m)^{th}$ place respectively.

Hence, for $k < \lambda$

$$ER_{i\,j}R_{\beta\,\gamma} = \frac{1}{\binom{n}{n_i}\binom{n-n_i}{n_\beta}} \sum_{k=j}^{n-(n_i-j)} \sum_{\ell=0}^{\gamma-1} \sum_{\lambda=k+\gamma-\ell}^{n-(n_\beta-\gamma)} k\lambda \binom{k-1}{j-1}\binom{k-j}{\ell}\binom{\lambda-1-k}{\gamma-1-\ell} \sum_{m=0}^{n_i-j} \binom{\lambda-k-(\gamma-\ell)}{m}$$

$$\binom{n-\lambda}{n_\beta-\gamma}\binom{n-\lambda-(n_\beta-\gamma)}{n_i-(j+m)}$$

and

$$EX_iX_\beta = \frac{1}{\binom{n}{n_i}\binom{n-n_i}{n_\beta}} \quad \frac{1}{6} \quad \frac{n_in_\beta}{(n_i+1)(n_\beta+1)} \quad \frac{1}{(n+1)^2} \quad \frac{n!}{n_i!n_\beta!(n-n_i-n_\beta)!}$$

$$\{\frac{n^2}{6}(4n_in_\beta+2n_i+2n_\beta+1) + \frac{n}{20}(24n_in_\beta+11n_i+11n_\beta+4) + \frac{1}{60}(32n_in_\beta+13n_i+13n_\beta+2)\}$$

by prop. 6

Furthermore

$$Cov(X_i,X_\beta) = EX_iX_\beta - EX_iEX_\beta$$

$$= \frac{1}{6}\frac{n_in_\beta}{(n_i+1)(n_\beta+1)(n+1)^2}\{\frac{n^2}{6}(4n_in_\beta+2n_i+2n_\beta+1)$$

$$+ \frac{n}{20}\,(24n_in_\beta+11n_i+11n_\beta+4) \;+\; \frac{1}{60}\,(32n_in_\beta+13n_i+13n_\beta+2)\}$$

$$- \frac{n_i(2n_i+1)}{6(n_i+1)}\;\frac{n_\beta(2n_\beta+1)}{6(n_\beta+1)}$$

$$= \frac{n_in_\beta}{(n_i+1)(n_\beta+1)(n+1)^2}\,\{\frac{n^2}{36}\,(2n_i+1)(2n_\beta+1) \;+\; \frac{n}{120}(24n_in_\beta+11n_i+11n_\beta+4)$$

$$+ \frac{1}{360}\,(32n_in \;+13n_i+13n \;+2) \;-\; \frac{(2n_i+1)(2n_\beta+1)(n+1)^2}{36}\;\}$$

$$= \frac{n_in_\beta}{360(n_i+1)(n_\beta+1)(n+1)^2}\,\{3n(24n_in_\beta+11n_i+11n_\beta+4)$$

$$+ \;(32n_in_\beta+13n_i+13n_\beta+2) \;-10(2n+1)(4n_in_\beta+2n_i+2n_\beta+1)\}$$

$$= \frac{n_in_\beta}{360(n_i+1)(n_\beta+1)(n+1)^2}\,\{n[n_in_\beta(72-80)+(n_i+n_\beta)(33-40) \;+\; (12-20)]$$

$$+ \;n_in_\beta(32-40)+(n_i+n_\beta)(13-20) \;+\; (2-10)\}$$

$$= \frac{n_in_\beta}{360(n_i+1)(n_\beta+1)(n+1)^2}\,\left\{\begin{array}{l} n[-8n_in_\beta-7(n_i+n_\beta)-8]+ \\ \;[-8n_in_\beta-7(n_i+n_\beta)-8] \end{array}\right\}$$

$$= \frac{n_in_\beta}{360(n_i+1)(n_\beta+1)(n+1)}\;(-8n_in_\beta-7(n_i+n_\beta)-8).$$

Appendix

Section 2

We want to compute the mean and variance of the statistic C of Section 6.

Let us first reduce the expression for C.

$$C = \sum_{i=1}^{q} \sum_{j=1}^{n_i} \left(\frac{R_{ij}}{n+1} - \frac{j}{n_i+1}\right)^2$$

$$= \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{R_{ij}^2}{(n+1)^2} - 2 \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j\ R_{ij}}{(n+1)(n_i+1)} + \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j^2}{(n_i+1)^2}$$

$$= \frac{1}{(n+1)^2} \sum_{k=1}^{n} k^2 - 2 \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j\ R_{ij}}{(n+1)(n_i+1)} + \sum_{i=1}^{q} \frac{1}{(n_i+1)^2} \frac{n_i(n_i+1)(2n_i+1)}{6}$$

$$= \frac{n(2n+1)}{6(n+1)} - 2 \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j\ R_{ij}}{(n+1)(n_i+1)} + \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

Let us compute μ_C.

$$E(C) = \frac{n(2n+1)}{6(n+1)} - 2 \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j\ E(R_{ij})}{(n+1)(n_i+1)} + \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{n_i+1}$$

$$= \frac{n(2n+1)}{6(n+1)} - 2 \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j^2\ (n+1)}{(n+1)(n_i+1)^2} + \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

by Prop. 2 of section 1.

$$= \frac{n(2n+1)}{6(n+1)} - 2\sum_{i=1}^{q} \frac{1}{(n_i+1)^2} \cdot \frac{n_i(n_i+1)(2n_i+1)}{6} + \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

$$= \frac{n(2n+1)}{6(n+1)} - \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

Let us compute σ_c^2 .

Var C = $E(C^2) - E^2(C)$

But

$$E(C^2) = E\{[\frac{n(2n+1)}{6(n+1)} - 2\sum_{i=1}^{q}\sum_{j=1}^{n_i} \frac{j\, R_{ij}}{(n_i+1)(n+1)} + \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}]^2\}$$

$$= E\{[\frac{n(2n+1)}{6(n+1)}]^2 + \frac{4}{(n+1)^2}[\sum_{i=1}^{q}\sum_{j=1}^{q} \frac{j\, R_{ij}}{n_i+1}]^2 + [\sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}]^2$$

$$- \frac{4}{6}\,\frac{n(2n+1)}{(n+1)} \sum_{i=1}^{q}\sum_{j=1}^{n_i} \frac{j\, R_{ij}}{(n_i+1)(n+1)} + \frac{2\, n(2n+1)}{6(n+1)} \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

$$- 4 \cdot \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)} \cdot \sum_{i=1}^{q}\sum_{j=1}^{n_i} \frac{j\, R_{ij}}{(n_i+1)(n+1)}\}$$

$$= [\frac{n(2n+1)}{6(n+1)}]^2 + \frac{4}{(n+1)^2} E[(\sum_{i=1}^{q}\sum_{j=1}^{n_i} \frac{j\, R_{ij}}{n_i+1})^2] + [\sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}]^2$$

$$- \frac{2}{3}\,\frac{n(2n+1)}{(n+1)} \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)} + \frac{n(2n+1)}{3(n+1)} \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

$$- 4\,(\sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)})^2 .$$

because

$$E(\sum_{i=1}^{q}\sum_{j=1}^{n_i} \frac{j\, R_{ij}}{(n_i+1)(n+1)}) = E(\sum_{i=1}^{q} X_i) = \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}$$

by note_1 of section 1.

$$E(C^2) = \{\frac{n(2n+1)}{6(n+1)} - 2\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j^2}{(n_i+1)^2} + \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}\}^2$$

$$+ \frac{4}{(n+1)^2}\, E[(\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{n_i+1})^2] - 4[\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j^2}{(n_i+1)^2}]^2$$

Note that

$$\mathrm{Var}[\frac{1}{n+1}\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{n_i+1}] = \frac{1}{(n+1)^2}\, E[(\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{(n_i+1)})^2] - [\sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}]^2$$

$$= \frac{1}{(n+1)^2}\, E[(\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{n_i+1})^2] - [\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j^2}{(n_i+1)^2}]^2.$$

Thus

$$E(C^2) = \{\frac{n(2n+1)}{6(n+1)} - 2\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j^2}{(n_i+1)^2} + \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}\}^2$$

$$+ 4\,\mathrm{Var}[\frac{1}{n+1}\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{n_i+1}].$$

$$= \{\frac{n(2n+1)}{6(n+1)} - \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}\}^2 + 4\,\mathrm{Var}[\frac{1}{n+1}\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{n_i+1}].$$

and so,

$$\mathrm{Var}\,C = \{\frac{n(2n+1)}{6(n+1)} - \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}\}^2 + 4\,\mathrm{Var}[\frac{1}{n+1}\sum_{i=1}^{q}\sum_{j=1}^{n_i}\frac{j\,R_{ij}}{n_i+1}]$$

$$- \left\{\frac{n(2n+1)}{6(n+1)} - \sum_{i=1}^{q} \frac{n_i(2n_i+1)}{6(n_i+1)}\right\}^2$$

$$= 4\ \mathrm{Var}\left[\frac{1}{n+1} \sum_{i=1}^{q} \sum_{j=1}^{n_i} \frac{j\ R_{i\ j}}{n_i+1}\right].$$

$$= 4\ \mathrm{Var}\left[\sum_{i=1}^{q} X_i\right]$$

$$= 4 \sum_{i=1}^{q} \mathrm{Var}\ X_i + 4 \sum_{i \neq \beta} \mathrm{Cov}(X_i,\ X_\beta).$$

$$\mathrm{Var}(C) = 4 \sum_{i=1}^{q} \frac{n_i(n-n_i)(2n_i+1)(2n_i+3)}{180\ (n+1)(n_i+1)^2}$$

$$+\ 4 \sum_{i \neq \beta} \frac{n_i n_\beta\ (-8n_i n_\beta - 7(n_i+n_\beta) - 8)}{360(n_i+1)(n_\beta+1)(n+1)}$$

Observe that

$$\sum_{i \neq \beta} \frac{n_i n_\beta\ (-8n_i n_\beta - 7(n_i+n_\beta) - 8)}{360(n_i+1)(n_\beta+1)(n+1)}$$

$$= \sum_{i=1}^{q} \sum_{\beta=1}^{n_i} \frac{n_i n_\beta\ (-8n_i n_\beta - 7(n_i+n_\beta) - 8)}{360(n_i+1)(n_\beta+1)(n+1)} - \sum_{i=1}^{q} \frac{n_i^2(-8n_i^2 - 14n_i - 8)}{360(n_i+1)^2(n+1)}$$

$$\frac{\mathrm{Var}\ (C)}{4} = \sum_{i=1}^{q} \frac{n_i(n-n_i)(2n_i+1)(2n_i+3)}{180(n+1)(n_i+1)^2} - \sum_{i=1}^{q} \frac{2n_i^2(-4n_i^2 - 7n_i - 4)}{360(n_i+1)^2(n+1)}$$

$$+\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i n_\beta\;(-8n_i n_\beta-7(n_i+n_\beta)-8)}{360(n_i+1)(n_\beta+1)(n+1)}$$

$$=\sum_{i=1}^{q}\frac{n_i}{180\;(n_i+1)^2(n+1)}[(n-n_i)(2n_i+1)(2n_i+3)+n_i(4n_i^2+7n_i+4)]$$

$$+\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i n_\beta\;(-8n_i n_\beta-7(n_i+n_\beta)-8)}{360(n_i+1)(n_\beta+1)(n+1)}$$

$$=\sum_{i=1}^{q}\frac{n_i}{180(n_i+1)^2(n+1)}[n(2n_i+1)(2n_i+3)-n_i(4n_i^2+8n_i+3)+n_i(4n_i^2+7n_i+4)]$$

$$+\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i n_\beta\;(-8n_i n_\beta-7(n_i+n_\beta)-8)}{360(n_i+1)(n_\beta+1)(n+1)}$$

$$=\sum_{i=1}^{q}\frac{n_i}{180(n_i+1)^2(n+1)}[n(2n_i+1)(2n_i+3)-n_i(n_i-1)]$$

$$+\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i n_\beta(-8n_i n_\beta-7n_i-7n_\beta-8)}{360(n_i+1)(n_\beta+1)(n+1)}$$

but

$$\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i n_\beta\;(-8n_i n_\beta-7n_i-7n_\beta-8)}{360(n_i+1)(n_\beta+1)(n+1)}=\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i n_\beta}{360(n_i+1)(n+1)}(-8n_i-7+\frac{n_i-1}{n_\beta+1})$$

$$=\sum_{i=1}^{q}\frac{n_i(-8n_i-7)n}{360(n_i+1)(n+1)}+\sum_{i=1}^{q}\sum_{\beta=1}^{q}\frac{n_i(n_i-1)}{360(n_i+1)(n+1)}\cdot\frac{n_\beta}{n_\beta+1};$$

$$= \sum_{i=1}^{q} \frac{n_i n(-8n_i-7)}{360(n_i+1)(n+1)} + \frac{1}{360(n+1)} \sum_{i=1}^{q} \frac{n_i(n_i-1)}{(n_i+1)} \sum_{\beta=1}^{q} \frac{n_\beta}{n_\beta+1}$$

$$= \sum_{i=1}^{q} \frac{n n_i(-8n_i-7)}{360(n+1)(n_i+1)} + \frac{1}{360(n+1)}\left(\sum_{i=1}^{q} \frac{n_i(n_i+1)-2n_i}{n_i+1}\right)\left(\sum_{\beta=1}^{q} \frac{n_\beta+1-1}{n_\beta+1}\right)$$

$$= \sum_{i=1}^{q} \frac{n\, n_i(-8n_i-7)}{360(n+1)(n_i+1)} + \frac{1}{360(n+1)}\left(n-2\sum_{i=1}^{q} \frac{n_i}{n_i+1}\right)\left(\sum_{\beta=1}^{q} \left[1-\frac{1}{n_\beta+1}\right]\right)$$

$$= \frac{n}{360(n+1)} \sum_{i=1}^{q} \frac{n_i(-8n_i-7)}{(n_i+1)} + \frac{1}{360(n+1)}\left(n-2\sum_{i=1}^{q} \frac{n_i+1-1}{n_i+1}\right)\left(q-\sum_{\beta=1}^{q} \frac{1}{n_\beta+1}\right)$$

$$= \frac{n}{360(n+1)} \sum_{i=1}^{q} \frac{n_i(-8n_i-7)}{n_i+1} + \frac{1}{360(n+1)}\left(n-2q+2\sum_{i=1}^{q} \frac{1}{n_i+1}\right)\left(q-\sum_{\beta=1}^{q} \frac{1}{n_\beta+1}\right)$$

$$\frac{\text{Var }(C)}{4} = \sum_{i=1}^{q} \frac{2n_i}{360(n+1)(n_i+1)^2}\left[n(2n_i+1)(2n_i+3)-n_i(n_i-1)\right]$$

$$+ \frac{1}{360(n+1)}\left\{\sum_{i=1}^{q} \frac{n\, n_i(-8n_i-7)}{n_i+1} + \left(n-2q+2\sum_{i=1}^{q} \frac{1}{n_i+1}\right)\left(q-\sum_{\beta=1}^{q} \frac{1}{n_\beta+1}\right)\right\}$$

$$= \frac{1}{360(n+1)}\left\{\sum_{i=1}^{q} \frac{n_i n}{n_i+1}\left[\frac{2(2n_i+1)(2n_i+3)}{n_i+1} + (-8n_i-7)\right]\right.$$

$$- \sum_{i=1}^{q} \frac{2n_i^2(n_i-1)}{(n_i+1)^2} + nq-2q^2+2q \sum_{i=1}^{q} \frac{1}{n_i+1}$$

$$- n \sum_{\beta=1}^{q} \frac{1}{n_\beta+1} + 2q \sum_{\beta=1}^{q} \frac{1}{n_\beta+1} - 2\Big(\sum_{i=1}^{q} \frac{1}{n_i+1}\Big)\Big(\sum_{\beta=1}^{q} \frac{1}{n_\beta+1}\Big)\}$$

but

i) $\sum_{i=1}^{q} \frac{1}{n_i+1} = \sum_{\beta=1}^{q} \frac{1}{n_\beta+1}$

ii) $$\sum_{i=1}^{q} \frac{2n_i^2(n_i-1)}{(n_i+1)^2} = \sum_{i=1}^{q} \Big[\frac{2(n_i+1)^2(n_i-1)}{(n_i+1)^2} - \frac{2(2n_i+1)(n_i-1)}{(n_i+1)^2}\Big]$$

$$= \sum_{i=1}^{q} 2(n_i-1) - 2 \sum_{i=1}^{q} \Big[\frac{(2n_i+1)(n_i+1)}{(n_i+1)^2} \quad \frac{-2(2n_i+1)}{(n_i+1)^2}\Big]$$

$$= 2n-2q-2 \sum_{i=1}^{q} \Big[\frac{2n_i+1}{n_i+1} \quad \frac{-2(2n_i+1)}{(n_i+1)^2}\Big]$$

$$= 2n-2q-2 \ \Big\{ \sum_{i=1}^{q} \frac{2(n_i+1)-1}{n_i+1} - 2 \sum_{i=1}^{q} \frac{2(n_i+1)-1}{(n_i+1)^2} \Big\}$$

$$= 2n-2q-2 \ \Big\{2q - \sum_{i=1}^{q} \frac{1}{n_i+1} - 2 \sum_{i=1}^{q} \frac{2}{n_i+1} + 2 \sum_{i=1}^{q} \frac{1}{(n_i+1)^2} \Big\}$$

$$= 2n-6q + 10 \sum_{i=1}^{q} \frac{1}{n_i+1} - 4 \sum_{i=1}^{q} \frac{1}{(n_i+1)^2}$$

$$\frac{\text{Var (C)}}{4} = \frac{1}{360(n+1)} \{n \sum_{i=1}^{q} \frac{n_i}{n_i+1}[\frac{2(2n_i+1)(2n_i+3)}{n_i+1} - (8n_i+7)]$$

$$- 2n+6q-10 \sum_{i=1}^{q} \frac{1}{n_i+1} + 4 \sum_{i=1}^{q} \frac{1}{(n_i+1)^2} + n_q-2q^2$$

$$+ 4q \sum_{i=1}^{q} \frac{1}{n_i+1} - n \sum_{\beta=1}^{q} \frac{1}{n_\beta+1} - 2(\sum_{i=1}^{q} \frac{1}{n_i+1})^2\}.$$

also

$$\frac{2(2n_i+1)(2n_i+3)}{n_i+1} - (8n_i+7) = \frac{2(4n_i^2+8n_i+3) - (8n_i+7)(n_i+1)}{n_i+1}$$

$$= \frac{1}{n_i+1} [(8n_i^2+16n_i+6) - (8n_i^2+15n_i+7)] = \frac{1}{n_i+1} (n_i-1)$$

$$\frac{\text{Var (C)}}{4} = \frac{1}{360} \frac{1}{n+1} \{n \sum_{i=1}^{q} \frac{n_i(n_i-1)}{(n_i+1)^2} - 2n+6q+nq-2q^2 +$$

$$(-10+4q-n) \sum_{i=1}^{q} \frac{1}{n_i+1} + 4 \sum_{i=1}^{q} \frac{1}{(n_i+1)^2} - 2(\sum_{i=1}^{q} \frac{1}{n_i+1})^2\}$$

but

$$\sum_{i=1}^{q} \frac{n_i(n_i-1)}{(n_i+1)^2} = \sum_{i=1}^{q} [\frac{(n_i+1)^2 - 3n_i-1}{(n_i+1)^2}] = \sum_{i=1}^{q} [1- \frac{3n_i+1}{(n_i+1)^2}]$$

$$= q - \sum_{i=1}^{q} \frac{3(n_i+1)-2}{(n_i+1)^2} = q - \sum_{i=1}^{q} [\frac{3}{n_i+1} - \frac{2}{(n_i+1)^2}]$$

$$\frac{\text{Var (C)}}{4} = \frac{1}{360(n+1)} \{n\ q-3n\sum_{i=1}^{q} \frac{1}{n_i+1} + 2n\sum_{i=1}^{q} \frac{1}{(n_i+1)^2} - 2n+6q$$

$$+ nq-2q^2 + (-10+4q-n)\sum_{i=1}^{q} \frac{1}{n_i+1} + 4\sum_{i=1}^{q} \frac{1}{(n_i+1)^2} - 2(\sum_{i=1}^{q} \frac{1}{n_i+1})^2\}$$

$$= \frac{1}{360(n+1)}\{(\sum_{i=1}^{q} \frac{1}{n_i+1})(-10+4q-4n) + (\sum_{i=1}^{q} \frac{1}{(n_i+1)^2})(2n+4)$$

$$-2(\sum_{i=1}^{q} \frac{1}{n_i+1})^2 -2n+6q+2nq-2q^2\}$$

$$= \frac{1}{360(n+1)}\{2(2q-2n-5)(\sum_{i=1}^{q} \frac{1}{n_i+1}) + 2(n+2)(\sum_{i=1}^{q} \frac{1}{(n_i+1)^2})$$

$$-2(\sum_{i=1}^{q} \frac{1}{n_i+1})^2 + 2(nq-n+3q-q^2)\}$$

$$= \frac{1}{180}(n+1)\{(2q-2n-5)(\sum_{i=1}^{q} \frac{1}{n_i+1}) + (n+2)(\sum_{i=1}^{q} \frac{1}{(n_i+1)^2})$$

$$-(\sum_{i=1}^{q} \frac{1}{n_i+1})^2 + n(q-1) + q(-q+3)\}.$$

TABLES

The following program first makes the list of all distinct permutations given n and q. For example if n = 3(N) and q = 2 (i,Q) it produces the following matrices (MA):

$$\begin{pmatrix}1 & 3 & 5\\ 2 & 4 & 6\end{pmatrix}, \quad \begin{pmatrix}1 & 3 & 6\\ 2 & 4 & 5\end{pmatrix}, \quad \begin{pmatrix}1 & 3 & 4\\ 2 & 5 & 6\end{pmatrix},$$

$$\begin{pmatrix}1 & 4 & 5\\ 2 & 3 & 6\end{pmatrix}, \quad \begin{pmatrix}1 & 4 & 6\\ 2 & 3 & 5\end{pmatrix}, \quad \begin{pmatrix}1 & 5 & 6\\ 2 & 3 & 4\end{pmatrix},$$

$$\begin{pmatrix}1 & 2 & 5\\ 3 & 4 & 6\end{pmatrix}, \quad \begin{pmatrix}1 & 2 & 6\\ 3 & 4 & 5\end{pmatrix}, \quad \begin{pmatrix}1 & 2 & 4\\ 3 & 5 & 6\end{pmatrix},$$

$\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}$. The idea of how to do this is due to Dr. S. Bainbridge. After each matrix is created a subprogram "Classe" computes the standardized value of the statistic C(STAT) then in position L of the array NTAB, "classe" stores the number of times this value stat has occurred. Going back to the main program, it then finds the 10%, 5%, 1% (NTEN, NFIV, NONE) largest value.

```
C
      READ(5,999)N,IQ
  999 FORMAT(2I5)
      NO=N*IQ
      DIMENSION MA(15,15),NTAB(1000)
      NU=0
      DOUBLE PRECISION IB1,IB2,IFAC
      IB1=IFAC(NO)
      IB2=IFAC(IQ)*IFAC(N)**IQ
      NP=IB1/IB2
      WRITE(6,9)NP
    9 FORMAT(1X, I12)
      NTEN=NP/10
      NFIV=NP*5/100
      NONE=NP/100
      DO 55 I=1,1000
   55 NTAB(I)=0
      X0=FLOAT(IQ)
      X1=FLOAT(NO)
      X2=FLOAT(N)
C VARIANCE OF C
       A=(X1+2.-X0)/(X2+1.)+2.*(X0-X1)-5.
       B=A*X0/(X2+1.)+X1*(X0-1.)-X0*(X0-3.)
       VARC=B/( 45.*(X1+1.))
      WRITE(6,88)VARC
   88 FORMAT(1X,F10.4)
C
C THE FIRST COLUMN
C
      MA(1,1)=1
      I=2
   10 M=I
   20 IF (M.GT.((I-1)*N+1))  GOTO40
   30 IF(M.GT.MA(I-1,1))  GOTO32
   31 M=M+1
      GOTO 20
   32 MA(I,1)=M
   40 I=I+1
      IF(I.LE.IQ)GOTO10
   99 J=2
C THE OTHERS COLUMNS
C
  100 I=1
  200 M=J
  300 IF(M.GT.(NO-N+J))GOTO 1
  310 L=1
  320 K=1
  400 IF(M.NE.MA(K,L))GOTO420
  410 M=M+1
      GOTO 300
  420 K=K+1
      IF(K.LE.IQ)GOTO400
  430 L=L+1
      IF(L.LT.J)GOTO 320
```

```
  440 KO=1                                                              CHS00550
  450 IF(KO.GE.I)GOTO600                                                CHS00560
  500 IF(M.EQ.MA(KO,J))GOTO410                                          CHS00570
  510 KO=KO+1                                                           CHS00580
      GOTO450                                                           CHS00590
  600 IF(M.LE.MA(I,J-1))GOTO410                                         CHS00600
  610 MA(I,J)=M                                                         CHS00610
  700 I=I+1                                                             CHS00620
      IF(I.LE.IQ)GOTO200                                                CHS00630
  710 J=J+1                                                             CHS00640
      IF(J.LE.N)GOTO100                                                 CHS00650
  800 IO=1                                                              CHS00660
      JO=1                                                              CHS00670
C                                                                       CHS00680
      NU=NU+1                                                           CHS00690
C                                                                       CHS00700
      CALL CLASSE(IQ,N,NO,NTAB,MA,VARC)                                 CHS00710
      IF(NU.GE.NP)GOTO 4                                                CHS00720
    1 IO=IO+1                                                           CHS00730
C                                                                       CHS00740
C TO PRODUCE THE NEXT MATRIX                                            CHS00750
C                                                                       CHS00760
      IF(IO.LE.IQ)GOTO3                                                 CHS00770
    2 IO=1                                                              CHS00780
      JO=JO+1                                                           CHS00790
      IF(JO.GT.N)GOTO4                                                  CHS00800
    3 J=N-JO+1                                                          CHS00810
      I=IQ-IO+1                                                         CHS00820
      IF(J.NE.1)GOTO5                                                   CHS00830
    6 IF(MA(I,1).GE.((I-1)*N+1))GOTO1                                   CHS00840
    7 M=MA(I,1)+1                                                       CHS00850
      GOTO 20                                                           CHS00860
    5 IF (MA(I,J).GE.(NO-N+J))GOTO1                                     CHS00870
    8 M=MA(I,J)+1                                                       CHS00880
      GOTO 300                                                          CHS00890
C                                                                       CHS00900
    4 NSUM=NTAB(1000)                                                   CHS00910
      IJ=1000                                                           CHS00920
      GOTO66                                                            CHS00930
   11 IJ=IJ-1                                                           CHS00940
      IF(IJ.LT.1)GOTO44                                                 CHS00950
      NSUM=NSUM+NTAB(IJ)                                                CHS00960
   66 IF(NSUM.LT.NONE)GOTO11                                            CHS00970
      A00=FLOAT(IJ)                                                     CHS00980
      A01=A00/100.                                                      CHS00990
   22 IJ=IJ-1                                                           CHS01000
      IF(IJ.LT.1)GOTO44                                                 CHS01010
      NSUM=NSUM+NTAB(IJ)                                                CHS01020
      IF(NSUM.LT.NFIV)GOTO22                                            CHS01030
      B00=FLOAT(IJ)                                                     CHS01040
      B05=B00/100.                                                      CHS01050
   33 IJ=IJ-1                                                           CHS01060
      IF(IJ.LT.1)GOTO44                                                 CHS01070
      NSUM=NSUM+NTAB(IJ)                                                CHS01080
```

```
      IF(NSUM.LT.NTEN)GOTO33                                            CHS01090
      COO=FLOAT(IJ)                                                     CHS01100
      C10=COO/100.                                                      CHS01110
   44 WRITE(6,1000)N,IQ,A01,B05,C10                                     CHS01120
 1000 FORMAT(1X,2I5/1X,3F8.4)                                           CHS01130
      STOP                                                              CHS01140
      END                                                               CHS01150
                                                                        CHS01160
      DOUBLE PRECISION FUNCTION IFAC(KL)                                CHS01170
      IFAC=1                                                            CHS01180
      IF (KL.LE.1)RETURN                                                CHS01190
      DO 7 JU=2,KL                                                      CHS01200
      XX=FLOAT(JU)                                                      CHS01210
    7 IFAC=IFAC*XX                                                      CHS01220
      RETURN                                                            CHS01230
      END                                                               CHS01240
C                                                                       CHS01250
      SUBROUTINE CLASSE(IQ,N,NO,NTAB,MA,VARC)                           CHS01260
C                                                                       CHS01270
      DIMENSION MA(15,15),NTAB(1000)                                    CHS01280
C TO CALCULATE STAT                                                     CHS01290
      C=0.                                                              CHS01300
      DO 2 J=1,N                                                        CHS01310
      X1=FLOAT(NO)                                                      CHS01320
      X2=FLOAT(N)                                                       CHS01330
      CS=0.                                                             CHS01340
       DO 1 I=1,IQ                                                      CHS01350
       MAIJ=MA(I,J)                                                     CHS01360
       XO=FLOAT(IQ)                                                     CHS01370
       X4=FLOAT(MAIJ)                                                   CHS01380
       X3=FLOAT(J)                                                      CHS01390
C                                                                       CHS01400
C                                                                       CHS01410
       F=X2+1.                                                          CHS01420
       D=(X2*(2.*X2+1.)/6.)*XO                                          CHS01430
       E=X1+1.                                                          CHS01440
    1  CS=CS+X4                                                         CHS01450
    2 C=C+(CS*X3)                                                       CHS01460
C C-EC=ST                                                              CHS01470
      ST=2.*(D-C/E)/F                                                   CHS01480
      STAT=ST/SQRT(VARC)                                                CHS01490
      STO=STAT*100.                                                     CHS01500
      L=INT(STO)                                                        CHS01510
      IF(L.LT.1)RETURN                                                  CHS01520
      IF(L.GT.1000)L=1000                                               CHS01530
      NTAB(L)=NTAB(L)+1                                                 CHS01540
      RETURN                                                            CHS01550
      END                                                               CHS01560
C                                                                       CHS01570
      END                                                               CHS01580
```

Documentation

VARIABLES:

N : number of observations per day

IQ: number of days

NO: total number of observations

MA: matrices we will generate, to find all possibilities of the distribution of the ranks.

NTAB: this vector has as index the value of the statistic (C * 100) and

NTAB(ij): number of times the value (ij/100) was computed from the different matrices MA.

NU: counts the number of matrices MA generated

NP: total number of matrices MA to be generated

VARC: the variance of C

M: the value we will possibly give to MA(i,j)

NSUM: add up the values of NTAB

IO: used to find in which line of MA we will change an element

JO: used to find in which column of MA we will change an element

A01: the value such that P(C ≥ A01) ≥ .01

B05: the value such that P(C ≥ B05) ≥ .05

C10: the value such that P(C ≥ C10) ≥ .1

NTEN: the number of values of C greater than C10

NFIV: the number of values of C greater than B05

NONE: the number of values of C greater than A01

Subroutines called

ifac: computes the factorial of its argument
classe

algorithm:

-initialisation

-we compute the variance

- to generate the different matrices MA:
  - for the 1st column:
    - we check that M is smaller or equal to (i-1)n+1
    - and that M is larger than MA(i-1,1)
  - for the other columns:
    - we check that M is smaller or equal to N-n+j
    - and that M is different of all MA(i,j)'s we have already determined
    - also that M is greater than MA(i,j-1)
- after having called classe we may generate the next matrix MA:
  - choose an element of MA to modify.

    1st MA(iq,n) i.e. IO = 1, JO = 1

    2nd MA(iq-1,n) IO = 2, JO = 1

    :

    iQth MA(i,n) IO = iQ, JO = 1

    iQ+1th MA(iq,n-1) IO = 1, JO = 2

    etc...

    but restarting this process after generating a new matrix MA.
  - if the element we selected say MA(i,j) did not reach its maximum we add 1 to its present value and use the appropriate column treatment
- we count the number of values of the statistics C are greater or equal to (ij/100) until NSUM is greater or equal to NONE then we have found A01
- we continue in the same way until we find B05 and C10.

Subroutine classe:

algorithm:

- computes the statistic C for each matrix MA generated and calls it STAT
- makes of STAT an index in NTAB(i.e.L)
- update NTAB.

P(STAT ≤ k) = p (in parentheses we give the value of n and q)

p.	90%	95%	99%
k	(2,3) 2.07	2.07	2.07
k	(2,4) 1.29	1.93	2.58
k	(2,5) 1.20	1.65	2.56
k	(2,6) 1.35	1.69	2.37
k	(3,2) 2.54	2.54	2.54
k	(3,3) 1.63	2.03	2.58
k	(3,4) 1.51	1.76	2.78
k	(4,2) 2.28	3.65	3.65
k	(4,3) 1.53	1.97	3.29
k	(5,2) 1.90	1.76	4.77
k	(6,2) 1.57	2.35	4.90
k	(7,2) 1.42	2.14	4.14
k	(8,2) 1.26	2.13	3.99
k	(9,2) 1.27	2.04	3.83