38

Problem Section

It was not possible to publish solutions to problems in this issue. It is not our intention to let this happen again. For the moment I can only apologize. I have also to apologize for a small, but annoying printing error in the formulation of problem 285 (W. ALBERS). In one of the formulas the letter "s" occurs, it should read as "x". Below the corrected version is reprinted together with a couple of new problems.

New Problems

*Problem 285*** (W. ALBERS)

Let Φ denote the standard normal distribution function and ϕ the density function. Define

$$k(x)=\phi(x)/(1-\Phi(x)).$$

Determine the smallest $c \geqslant 1$ such that

$$k(x) > xx \pm \frac{\quad}{c+xk(x)} \text{ for all } x \geqslant 0.$$

Problem 290 (T.J. WANSBEEK)

Let $X_1,\ldots,X_n$ be independently distributed as $N_p(0,\Sigma)$, where $n>p$ and Σ is nonsingular. Define the matrix $X=(X_1,\ldots,X_n)^T$. Calculate the expectation of $X(X^TX)^{-1}X^T$.

Problem 291 (J.A. WELLNER, Seattle, USA)

Suppose that $\{N_t : t \geqslant 0\}$ is a standard Poisson process with rate 1. For $i=1,2,\ldots$ let $X_i=N_i-N_{i-1}$. So the X_i's are i.i.d. having the Poisson (1) distribution. Additionally, let $Y_i=T_i-T_{i-1}$, where $T_i=\inf\{t>0: N_t=i\}$. So T_i is the time of the i-th event, Y_i is the i-th interarrival time, and the Y_i's are i.i.d. Exponential (1) distributed.

A. Show that $\overline{XY}_n=n^{-1}\sum_{i=1}^{n} X_iY_i \xrightarrow{p} 1$.

B.** Show that $\sqrt{n}(\overline{XY}_n-1)$ converges in distribution and find the limiting distribution.

Arranging the possible outcomes in increasing order, δ^* can be given as follows

$$\begin{aligned}
\delta^*(a,a,a) &= a,\\
\delta^*(a,a,a+1) &= a + p/(1-p),\\
\delta^*(a,a,a+2) &= a+1,\\
\delta^*(a,a+1,a+1) &= a + (1-2p)/(1-p),\\
\delta^*(a,a+1,a+2) &= a+1,\\
\delta^*(a,a+2,a+2) &= a+1.
\end{aligned}$$

This estimator is unbiased, and its variance is $2p^2(3-11p+11p^2)/(1-p)$, which is smaller than the variances of the mean and the median for fixed $p \in [0, \frac{1}{2}]$.

Problem 290 (T. J. Wansbeek)

Let $X_1, ..., X_n$ be independently distributed as $N_p(0, \Sigma)$, where $n > p$ and Σ is non-singular. Define the matrix $X = (X_1, ..., X_n)'$. Calculate the expectation of $X(X'X)^{-1}X'$.

The problem was solved by R. M. de Jong, E. A. van der Meulen, H. Neudecker, and A. G. M. Steerneman. Neudecker applies a result of Mardia (viz. Theorem 2.8.3 from K. V. Mardia, J. T. Kent, and J. M. Bibby (1979), Multivariate analysis, Academic Press). The other solutions are obtained using invariance.

Solution by E. A. van der Meulen

Set $Y' = \Sigma^{-1/2}X' = (Y_1, ..., Y_n)$. Hence $Y_1, ..., Y_n$ is an independent random sample from the $N_p(0, I_p)$ distribution. As

$$X(X'X)^{-1}X' = Y(Y'Y)^{-1}Y'$$

it follows that without loss of generality we may assume $\Sigma = I_p$. Note that vec $Y \sim N_{p \times n}(0, I_p \otimes I_n)$ and that for any $n \times n$ orthogonal matrix O, vec $OY = (I_p \otimes O)$ vec $Y \sim N_{p \times n}(0, I_p \otimes I_n)$. Hence $\mathscr{L}\{OY\} = \mathscr{L}\{Y\}$ for all orthogonal O. This easily leads to

$$E\{Y(Y'Y)^{-1}Y'\} = OE\{Y(Y'Y)^{-1}Y'\}O',$$

for all orthogonal matrices O. Hence $E\{Y(Y'Y)^{-1}Y'\} = cI_n$. It remains to establish the constant c. To this end, note that

$$cn = E\{\text{trace } Y(Y'Y)^{-1}Y'\} = E\{\text{trace } (Y'Y)^{-1}Y'Y\} = p.$$

Hence $c = p/n$.

Problem 291 (J. A. Wellner, Seattle, USA)

Suppose that $\{N_t : t \geq 0\}$ is a standard Poisson process with rate 1. For $i = 1, 2, ...$ let $X_i = N_i - N_{i-1}$. So the X_i's are i.i.d. having the Poisson (1) distribution. Additionally, let $Y_i = T_i - T_{i-1}$, where $T_i = \inf\{t > 0 : N_t = i\}$. So T_i is the time of the ith event, Y_i is the ith interarrival time, and the Y_i's are i.i.d. Exponential (1) distributed.

A. Show that $\overline{XY}_n = n^{-1} \sum_{i=1}^{n} X_i Y_i \xrightarrow{p} 1$.

B.** Show that $\sqrt{n}(\overline{XY}_n - 1)$ converges in distribution and find the limiting distribution.

No solutions were submitted. The problem will be kept open until further notice.

New problems

*Problem 304*** (W. REY)

The problem we consider goes on *balanced randomization of experiments*. It is first introduced and then the question comes.

A given design of experiment is composed of $b = 8$ blocks and each block consists of a sequence of $k = 4$ units out of $t = 5$ possible treatments. It is known that the results may depend on the order in which the units are performed and it is desired to get rid of this dependence. That can be done poorly by randomizing and much better by arranging the four units of each block according to the following design:

1	2	3	4
2	5	1	3
3	1	4	5
4	3	5	2
5	4	2	1
1	5	3	2
2	4	1	5
3	1	2	4

Each treatment occurs either $[k/t]$ or $\{k/t\}$ times in each block; the number of treatment occurrences per column is either $\{b/t\}$ of $[b/t]$. Considering the sequences of two successive units, there are $t(t-1)$ possible different pairs and there are $b(k-1)$ pair occurrences in the design; thus in this example, 16 pairs should occur once and 4 pairs should occur twice. This is precisely achieved.

The design is (modestly) known as an "universal optimal repeated measurements design" or an $\mathbf{RM}(t, b, k)$, in the notation of HEDAYAT, A. and K. AFSARINEJAD (1978), Repeated measurements designs II, *Annals of Statistics 6,* 619–628. The theoretical knowledge is limited and mainly concentrated on the Latin squares $\mathbf{RM}(t, t, t)$.

The question is: How can you construct a $\mathbf{RM}(t, b, k)$?

Problem 305 (B. C. HOMBAS, Athens, Greece)

In proving the Central Limit Theorem the following inequality was applied

$$\left|\operatorname{sgn} h - \frac{2}{\pi}\int_0^T \frac{\operatorname{sgn} ht}{t}\,dt\right| \leq \min\left(1, \frac{2}{|h|T}\right), \quad \text{for} \quad T>0 \quad \text{and} \quad h \neq 0.$$

Prove this inequality.

Problem Section

Solutions

Problem 291 (J.A. WELLNER, Seattle, USA)

Suppose that $\{N_t : t \geq 0\}$ is a standard Poisson process with rate 1. For $i = 1, 2, \dots$ let $X_i = N_i - N_{i-1}$. So the X_i's are i.i.d. having the Poisson(1) distribution. Additionally, let $Y_i = T_i - T_{i-1}$, where $T_i = \inf\{t > 0 : N_t = i\}$. So T_i is the time of the i-th event, Y_i is the i-th interarrival time, and the Y_i's are i.i.d. Exponential(1) distributed.

A. Show that $\overline{XY}_n = n^{-1} \sum_{i=1}^{n} X_i Y_i \xrightarrow{p} 1$.

B.** Show that $\sqrt{n}(\overline{XY}_n - 1)$ converges in distribution and find the limiting distribution.

Part B of this problem has not been solved yet. Part A was solved by S. Foss, G. Hooghiemstra, and M. Keane. Their solution will be given below.

Solution by S. FOSS, G. HOOGHIEMSTRA, *and* M. KEANE

It will be shown that for $n \to \infty$,

$$\mathrm{E}\sum_{k=1}^{n} X_k Y_k = n - (\frac{2}{\pi})^{\frac{1}{2}}.n^{\frac{1}{2}} + O(1),$$

and that

$$\mathrm{Var}(\sum_{k=1}^{n} X_k Y_k) = (2 - \frac{2}{\pi})n + o(n).$$

Together these two asymptotic results show that

$$\sum_{k=1}^{n} X_k Y_k \to 1,$$

in probability, as $n \to \infty$.

We constantly use well known identities for the Poisson, uniform, and binomial distribution. For an overview of these identities we refer to the appendix at the end of this solution.

Lemma 1. For all integers i and j, satisfying $i \geq 0$ and $j > 0$,

$$\mathrm{E}N_i Y_j = i - P(N_i \geq j).$$

Proof.

$$
\begin{aligned}
EN_iY_j &= E\{N_iE(Y_j|N_i)\} = E\{N_i1(N_i \le j-1)\} + \\
&\quad +E\{\frac{iN_i}{(N_i+1)}1(N_i \ge j-1)\} \\
&= iP(N_i < j-1) + iP(N_i \ge j-1) - P(N_i \ge j). \quad \square
\end{aligned}
$$

From the above lemma we obtain for $i = 1, 2, \ldots$,

$$
\begin{aligned}
(1) \quad EX_iY_i &= EN_iY_i - EN_{i-1}Y_i = i - P(N_i \ge i) - \{i - 1 - P(N_{i-1} \ge i)\} \\
&= 1 + P(N_{i-1} \ge i) - P(N_i \ge i).
\end{aligned}
$$

Summing from $i = 1$ to n gives

$$
(2) \quad E(\sum_{i=1}^{n} X_iY_i) = n - \sum_{i=1}^{n-1} P(N_i = i) - P(N_n \ge n).
$$

It is not difficult to see that

$$
(3) \quad E(\sum_{i=1}^{n} X_iY_i) = n - (\frac{2}{\pi})^{\frac{1}{2}}.n^{\frac{1}{2}} + O(1), \quad n \to \infty.
$$

There is no advantage in calculating $EX_i^2Y_i^2$ through the calculation of $EN_i^2Y_j^2$, therefore we follow a direct approach.

Proposition 1. For $i \ge 2$,

$$
\begin{aligned}
(4) \quad EX_i^2Y_i^2 &= 4 - 10P(N_i = i) + (\frac{8}{e} - 4)P(N_{i-1} = i) - 2P(N_i = i-3) + \\
&\quad -\frac{4}{i}P(N_i = i-2),
\end{aligned}
$$

for $i = 1$,

$$
EX_1^2Y_1^2 = 4 - \frac{10}{e}.
$$

Proof. For $i \ge 2$ we can write

$$
\begin{aligned}
EX_i^2Y_i^2 &= E\{X_i^2Y_i^21(N_{i-1} \ge i)\} + E\{X_i^2Y_i^21(N_i < i-1)\} + \\
&\quad +E\{X_i^2Y_i^21(N_{i-1} \le i-1 \le N_i)\} \\
&= F_1 + F_2 + F_3,
\end{aligned}
$$

which can be evaluated as follows:

$$
\begin{aligned}
F_1 &= 2E\{E(Y_i^2|N_{i-1})1(N_{i-1} \ge i)\} \\
&= 2E\{\frac{2(i-1)^2}{(N_{i-1}+1)(N_{i-1}+2)}1(N_{i-1} \ge i)\}
\end{aligned}
$$

$$
\begin{aligned}
&= 4(i-1)^2\mathrm{E}\{\frac{1(N_{i-1}\geq i)}{(N_{i-1}+1)(N_{i-1}+2)}\} \\
&= 4P(N_{i-1}\geq i+2), \\
F_2 &= \mathrm{E}\{X_i^2Y_i^2 1(N_i < i-1)\} = 2\mathrm{E}\{X_i^2 1(N_i < i-1)\} \\
&= 2\mathrm{E}\{\mathrm{E}(X_i^2|N_i)1(N_i < i-1)\} = 2\mathrm{E}\{(\frac{i-1}{i^2}N_i + \frac{1}{i^2}N_i^2)1(N_i < i-1)\} \\
&= 2\frac{i-1}{i^2}iP(N_i < i-2) + \frac{2}{i^2}i^2P(N_i < i-3) + \frac{2}{i^2}iP(N_i < i-2) \\
&= 2\{P(N_i \leq i-3) + P(N_i \leq i-4)\}, \\
F_3 &= \mathrm{E}X_i^2Y_i^2 1(N_{i-1}\leq i-1\leq N_i)\} \\
&= \mathrm{E}\{X_i^2\mathrm{E}(Y_i^2 N_{i-1}, N_i)1(N_{i-1}\leq i-1\leq N_i)\} \\
&= \mathrm{E}\{X_i^2 1(X_i\geq 1)[\frac{2}{(X_i+1)(X_i+2)} + \frac{2(i-1)^2}{(N_{i-1}+1)(N_{i-1}+2)}1(N_{i-1}=i-1) + \\
&\quad +2.1(N_i = i-1)].1(N_{i-1}\leq i-1\leq N_i)\} \\
&= \mathrm{E}\{\frac{2X_i^2}{(X_i+1)(X_i+2)}1(N_{i-1}\leq i-1\leq N_i)\} + \mathrm{E}\{\frac{2(i-1)^2}{i(i+1)}X_i^2 1(N_{i-1}=i-1\leq N_i)\} + \\
&\quad +\mathrm{E}\{2X_i^2 1(N_{i-1}\leq i-1 = N_i)\} = J_1 + J_2 + J_3,
\end{aligned}
$$

where

$$
\begin{aligned}
J_3 &= 2\mathrm{E}\{X_i^2 1(N_i = i-1)\} = 2\mathrm{E}\{\mathrm{E}(X_i^2|N_i)1(N_i = i-1)\} \\
&= 2\mathrm{E}\{[(i-1)\frac{1}{i}\frac{i-1}{i} + \frac{(i-1)^2}{i^2}]1(N_i = i-1)\} \\
&= 4\frac{(i-1)^2}{i^2}P(N_i = i-1) \\
&= 4\frac{(i-1)}{i}P(N_i = i-2), \\
J_2 &= \frac{2(i-1)^2}{i(i+1)}\mathrm{E}\{X_i^2 1(N_{i-1} = i-1)\} \\
&= \frac{2(i-1)^2}{i(i+1)}\mathrm{E}\{X_i^2\}P(N_{i-1} = i-1) \\
&= 4\frac{(i-1)^2}{i(i+1)}P(N_{i-1} = i-1) \\
&= 4P(N_{i-1} = i+1), \\
J_1 &= 2\mathrm{E}\{\frac{X_i^2}{(X_i+1)(X_i+2)}1(N_{i-1}\leq i-1\leq N_i)\} \\
&= 2\mathrm{E}\{1(N_{i-1}\leq i-1)\frac{X_i^2}{(X_i+1)(X_i+2)}1(X_i \geq i-1-N_{i-1})\} \\
&= 2\mathrm{E}\{1(N_{i-1}\leq i-1)\mathrm{E}\{(1 - \frac{3}{X_i+1} + \frac{4}{(X_i+1)(X_i+2)}).
\end{aligned}
$$

$$
\begin{aligned}
& .1(X_i \geq i-1-N_{i-1})|N_{i-1})\} \\
= \; & 2\mathrm{E}\{1(N_{i-1} \leq i-1)(\phi(i-1-N_{i-1}) - 3\phi(i-N_{i-1}) + 4\phi(i+1-N_{i-1}))\},
\end{aligned}
$$

where $\phi(j)$ is defined in formula (13) of the Appendix. Hence

$$
\begin{aligned}
J_1 &= 2\sum_{k=0}^{i-1} P(N_{i-1}=k)\left[\frac{e^{-1}}{(i-1-k)!} - \frac{2e^{-1}}{(i-k)!} + 2\phi(i+1-k)\right] \\
&= 2P(N_i = i-1) - 4P(N_i = i) + \frac{4}{e}P(N_{i-1}=i) + 4P(N_{i-1} \leq i-1) + \\
&\quad -4P(N_i \leq i) + \frac{4}{e}P(N_{i-1}=i).
\end{aligned}
$$

The result (4) now follows from the equality

$$\mathrm{E}X_i^2Y_i^2 = F_1 + F_2 + J_1 + J_2 + J_3.$$

Finally for $i=1$, it is straightforward that

$$\mathrm{E}X_1^2Y_1^2 = 4 - \frac{10}{e}. \quad \square$$

Now we consider the variance:

$$(5) \quad \mathrm{Var}(\sum_{i=1}^{n} X_iY_i) = \sum_{i=1}^{n}\mathrm{Var}(X_iY_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{Cov}(X_iY_i, X_jY_j).$$

From the two preceeding subsections we see that

$$(6) \quad \mathrm{Var}(X_iY_i) = 3 + \frac{1}{(2\pi i)^{\frac{1}{2}}}(\frac{8}{e} - 14) + O(\frac{1}{i}), \quad i \to \infty.$$

Hence

$$(7) \quad \sum_{i=1}^{n}\mathrm{Var}(X_iY_i) = 3n + O(n^{\frac{1}{2}}), \quad n \to \infty.$$

In order to calculate the term $\mathrm{Cov}(X_iY_i, X_jY_j)$ we introduce the notation:

$$N(i,j) := N_j - N_i, \quad i \leq j,$$

the number of Poisson points contained in $(i,j]$. With this notation one can write, for $1 \leq i < j \leq n$,

$$
\begin{aligned}
(8) \quad \mathrm{Cov}(X_iY_i, X_jY_j) \;=\; & \mathrm{Cov}(N_iY_i, N(i,j)Y_j) \\
& -\mathrm{Cov}(N_iY_i, N(i,j-1)Y_j) \\
& -\mathrm{Cov}(N_{i-1}Y_i, N(i-1,j)Y_j) \\
& +\mathrm{Cov}(N_{i-1}Y_i, N(i-1,j-1)Y_j).
\end{aligned}
$$

In order to calculate $\mathrm{E}N_iY_iN(i,j)Y_j$ we condition on the σ-field $\mathcal{F} := \sigma\{N_i, N(i,j)\}$:

$$\mathrm{E}N_iY_iN(i,j)Y_j = \mathrm{E}\{N_iN(i,j)\mathrm{E}(Y_iY_j|\mathcal{F})\}.$$

The conditional expectation $\mathrm{E}(Y_iY_j|\mathcal{F})$ is easy to calculate, but consists of 6 terms (originally 13, but that number can be reduced to 6). We obtain for $i < j$,

(9) $\mathrm{E}N_iY_iN(i,j)Y_j =$

$$\begin{aligned}
= \; & \mathrm{E}N_iN(i,j)1(N_i < N_j \le i-1) \\
& + \mathrm{E}N_iN(i,j)\frac{i}{(N_i+1)}1(i-1 \le N_i < N_j \le j-1) \\
& + \mathrm{E}N_iN(i,j)\frac{j-i}{(N(i,j)+1)}1(N_i \le i-1 \le N_j \le j-1) \\
& + \mathrm{E}N_iN(i,j)\frac{(j-i)^2}{(N(i,j)+1)(N(i,j)+2)}1(N_i \le i-1, N_j \ge j-1) \\
& + \mathrm{E}N_iN(i,j)\frac{i^2}{(N_i+1)(N_i+2)}1(j-1 \le N_i, N_j \ge j) \\
& + \mathrm{E}N_iN(i,j)\frac{i(j-i)}{(N_i+1)(N(i,j)+1)}1(i-1 \le N_i \le j-1 \le N_j).
\end{aligned}$$

Similarly the three other expectations follow:
for $i < j-1$,

(10) $\mathrm{E}N_iY_iN(i,j-1)Y_j =$

$$\begin{aligned}
= \; & \mathrm{E}N_iN(i,j-1)1(N_i < N_{j-1} \le i-1) \\
& + \mathrm{E}N_iN(i,j-1)\frac{i}{(N_i+1)}1(i-1 \le N_i < N_{j-1} \le j-1) \\
& + \mathrm{E}N_iN(i,j-1)\frac{j-i-1}{(N(i,j-1)+1)}1(N_i \le i-1 \le N_{j-1} \le j-1) \\
& + \mathrm{E}N_iN(i,j-1)\frac{(j-i-1)^2 1(N_i \le i-1, N_{j-1} \ge j-1)}{(N(i,j-1)+1)(N(i,j-1)+2)} \\
& + \mathrm{E}N_iN(i,j-1)\frac{i^2}{(N_i+1)(N_i+2)}1(j-1 \le N_i, N_{j-1} \ge j) \\
& + \mathrm{E}N_iN(i,j-1)\frac{i(j-i-1)1(i-1 \le N_i \le j-1 \le N_{j-1})}{(N_i+1)(N(i,j-1)+1)},
\end{aligned}$$

for $i < j$,

(11) $\mathrm{E}N_{i-1}Y_iN(i-1,j)Y_j =$

$$\begin{aligned}
&= \mathrm{E}N_{i-1}N(i-1,j)1(N_{i-1} < N_j \le i-1)\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j)\frac{i-1}{(N_{i-1}+1)}1(i-1 \le N_{i-1} < N_j \le j-1)\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j)\frac{j-i+1}{(N(i-1,j)+1)}1(N_{i-1} \le i-1 \le N_j \le j-1)\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j)\frac{(j-i+1)^2 1(N_{i-1} \le i-1, N_j \ge j-1)}{(N(i-1,j)+1)(N(i-1,j)+2)}\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j)\frac{(i-1)^2}{(N_{i-1}+1)(N_{i-1}+2)}1(j-1 \le N_{i-1}, N_j \ge j)\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j)\frac{(i-1)(j-i+1)1(i-1 \le N_{i-1} \le j-1 \le N_j)}{(N_{i-1}+1)(N(i-1,j)+1)},
\end{aligned}$$

for $i < j$,

(12) $\mathrm{E}N_{i-1}Y_iN(i-1,j-1)Y_j$

$$\begin{aligned}
&= \mathrm{E}N_{i-1}N(i-1,j-1)1(N_{i-1} < N_{j-1} \le i-1)\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j-1)\frac{i-1}{(N_{i-1}+1)}1(i-1 \le N_{i-1} < N_{j-1} \le j-1)\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j-1)\frac{(j-i)1(N_{i-1} \le i-1 \le N_{j-1} \le j-1)}{(N(i-1,j-1)+1)}\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j-1)\frac{(j-i)^2 1(N_{i-1} \le i-1, N_{j-1} \ge j-1)}{(N(i-1,j-1)+1)(N(i-1,j-1)+2)}\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j-1)\frac{(i-1)^2 1(j-1 \le N_{i-1}, N_{j-1} \ge j)}{(N_{i-1}+1)(N_{i-1}+2)}\\
&\quad +\mathrm{E}N_{i-1}N(i-1,j-1)\frac{(i-1)(j-i)1(i-1 \le N_{i-1} \le j-1 \le N_{j-1})}{(N_{i-1}+1)(N(i-1,j-1)+1)}.
\end{aligned}$$

According to formula (5) the above expressions have to be summed over i and j. In order to see the effect of the cancelation of terms we introduce for $i \le j$,

$$\begin{aligned}
(13)\quad \phi(i,j) &:= \mathrm{E}N_iY_iN(i,j)Y_j\\
\tilde{\phi}(i,j) &:= \mathrm{E}N_iY_iN(i,j)Y_{j+1}\\
\psi(i,j) &:= \mathrm{E}N_iY_{i+1}N(i,j)Y_j\\
\tilde{\psi}(i,j) &:= \mathrm{E}N_iY_{i+1}N(i,j)Y_{j+1},
\end{aligned}$$

and the four associated products of the expectations:

$$\begin{aligned}
(14)\quad a(i,j) &:= \mathrm{E}N_iY_i\mathrm{E}N(i,j)Y_j\\
\tilde{a}(i,j) &:= \mathrm{E}N_iY_i\mathrm{E}N(i,j)Y_{j+1}\\
b(i,j) &:= \mathrm{E}N_iY_{i+1}\mathrm{E}N(i,j)Y_j\\
\tilde{b}(i,j) &:= \mathrm{E}N_iY_{i+1}\mathrm{E}N(i,j)Y_{j+1}.
\end{aligned}$$

Lemma 2.

$$(15)\quad 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{Cov}(X_iY_i, X_jY_j)$$

$$\begin{aligned}
= \quad & \{2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n}[\phi(i,j)-\tilde{\phi}(i,j)-\psi(i,j)+\tilde{\psi}(i,j)] \\
& \quad -2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n}[a(i,j)-\tilde{a}(i,j)-b(i,j)+\tilde{b}(i,j)]\} \\
& +\{2\sum_{i=1}^{n-2}(\tilde{\phi}(i,n)-\tilde{\psi}(i,n))-2\sum_{i=1}^{n-2}(\tilde{a}(i,n)-\tilde{b}(i,n))\} \\
& +\{2\sum_{i=1}^{n-2}(\phi(i,i+1)-\tilde{\phi}(i,i+1)+\tilde{\psi}(i,i+1)) \\
& \quad -2\sum_{i=1}^{n-2}(a(i,i+1)-\tilde{a}(i,i+1)+\tilde{b}(i,i+1))\} \\
& +\{2\phi(n-1,n)-2a(n-1,n)\}.
\end{aligned}$$

Proof. From (8) and the definitions (13) and (14) we obtain

$$2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{Cov}(X_iY_i, X_jY_j) =$$

$$\begin{aligned}
= \quad & 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\{\phi(i,j)-\tilde{\phi}(i,j-1)-\psi(i-1,j)+\tilde{\psi}(i-1,j-1)\}+ \\
& -2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\{a(i,j)-\tilde{a}(i,j-1)-b(i-1,j)+\tilde{b}(i-1,j-1)\}.
\end{aligned}$$

Now observe that

$$(16)\quad 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}[\phi(i,j)-\tilde{\phi}(i,j-1)-\psi(i-1,j)+\tilde{\psi}(i-1,j-1)]+$$

$$-2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n}[\phi(i,j)-\tilde{\phi}(i,j)-\psi(i,j)+\tilde{\psi}(i,j)]$$

$$\begin{aligned}
&= 2\sum_{i=1}^{n-1}\phi(i,i+1) - 2\sum_{i=1}^{n-1}\tilde{\phi}(i,i) - 2\sum_{i=1}^{n-2}\tilde{\phi}(i,i+1) + \\
&\quad +2\sum_{i=1}^{n-2}\tilde{\phi}(i,n) - 2\sum_{j=2}^{n}\psi(0,j) - 2\sum_{i=1}^{n-2}\tilde{\psi}(i,n) \\
&\quad +2\sum_{i=1}^{n-2}\tilde{\psi}(i,i+1) + 2\sum_{j=1}^{n-1}\tilde{\psi}(0,j) \\
&= 2\sum_{i=1}^{n-2}[\tilde{\phi}(i,n) - \tilde{\psi}(i,n)] \\
&\quad +2\sum_{i=1}^{n-2}[\phi(i,i+1) - \tilde{\phi}(i,i+1) + \tilde{\psi}(i,i+1)] \\
&\quad +2\phi(n-1,n),
\end{aligned}$$

where the second equality follows since $\tilde{\phi}(i,i) = \psi(0,i) = \tilde{\psi}(0,i) = 0$, $\forall i$. The same equality holds with $\phi, \tilde{\phi}, \psi$, and $\tilde{\psi}$ replaced by $a, \tilde{a}, b$, and $\tilde{b}$, respectively. Together these equalities imply (15). □

The four terms between brackets on the right-hand side of (15) will be evaluated below.

Proposition 2.

$$\text{(17)}\qquad 2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n}[\phi(i,j) - \tilde{\phi}(i,j) - \psi(i,j) + \tilde{\psi}(i,j)] = 2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n} P(N_i = i, N_j = j+1),$$

$$\text{(18)}\qquad 2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n}[a(i,j) - \tilde{a}(i,j) - b(i,j) + \tilde{b}(i,j)] = 2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n} P(N_i = i)\{P(N_j = j) - P(N_i = j)\}.$$

Proof. The proof of (18) is immediate from Lemma 1 and is therefore omitted. For the proof of (17) we use (9)–(12), from which we obtain for $j \geq i+2$,

$$\phi(i,j) - \tilde{\phi}(i,j) - \psi(i,j) + \tilde{\psi}(i,j)$$

$$
\begin{aligned}
= \quad & -\mathrm{E}\{N_i N(i,j)\frac{i}{N_i+1}1(N_i = i-1, N_j = j)\} \\
& +\mathrm{E}\{N_i N(i,j)\frac{j-i}{N(i,j)+1}1(N_i = i, N_j = j)\} \\
& -\mathrm{E}\{N_i N(i,j)\frac{(j-i)^2}{(N(i,j)+1)(N(i,j)+2)}1(N_i = i, N_j = j-1)\} \\
& +\mathrm{E}\{N_i N(i,j)\frac{i(j-i)}{(N_i+1)(N(i,j)+1)}1(N_i = i-1, N_j = j-1)\}.
\end{aligned}
$$

Note that from the 6 terms in each of (9)–(12) two terms vanish. From the above expression we obtain after some calculus

$$
\phi(i,j) - \tilde{\phi}(i,j) - \psi(i,j) + \tilde{\psi}(i,j) = P(N_i = i)P(N(i,j) = j-i+1). \quad \square
$$

Corollary 1. For $n \to \infty$,

$$
\begin{aligned}
(19) \qquad & 2\sum_{i=1}^{n-2}\sum_{j=i+2}^{n}\{[\phi(i,j) - \tilde{\phi}(i,j) - \psi(i,j) + \tilde{\psi}(i,j)] + \\
& \quad -[a(i,j) - \tilde{a}(i,j) - b(i,j) + \tilde{b}(i,j)]\} \\
& \quad = (1 - \frac{2}{\pi})n + o(n).
\end{aligned}
$$

Proposition 3. For $n \to \infty$,

$$
\begin{aligned}
(20) \qquad & 2\sum_{i=1}^{n-2}(\tilde{\phi}(i,n) - \tilde{\psi}(i,n)) \\
& \quad = 2\sum_{i=1}^{n-2}\{iP(N_i = i) - (n-i)P(N_i = i)\} - n + O(n^{\frac{1}{2}}),
\end{aligned}
$$

$$
(21) \qquad 2\sum_{i=1}^{n-2}(\tilde{a}(i,n) - \tilde{b}(i,n)) = -2\sum_{i=1}^{n-2}(n-i)P(N_i = i) + O(n^{\frac{1}{2}}).
$$

Proof. The proof of (21) is straightforward and therefore omitted; the proof of (20) follows below.
Combining (13), (10), and (12) we obtain for $i < n$,

$$
\begin{aligned}
& \tilde{\phi}(i,n) - \tilde{\psi}(i,n) = \\
& \quad [-\mathrm{E}N_i N(i,n)1(N_n = i) + \mathrm{E}N_i N(i,n)\frac{i}{N_i+1}1(N_i = i-1 < N_n \leq n) \\
& \quad +\mathrm{E}N_i N(i,n)\frac{n-i}{N(i,n)+1}\{1(N_i \leq i-1 = N_n) - 1(N_i = i \leq N_n \leq n)\}
\end{aligned}
$$

$$-\mathrm{E}N_iN(i,n)\frac{(n-i)^2}{(N(i,n)+1)(N(i,n)+2)}1(N_i=i,N_n\geq n)$$
$$+\mathrm{E}N_iN(i,n)\frac{i(n-i)}{(N_i+1)(N(i,n)+1)}1(N_i=i-1,N_n\geq n)].$$

After sommation from 1 to $n-2$, and evaluation of a part of the expectations this results in

$$2\sum_1^{n-2}(\tilde{\phi}(i,n)-\tilde{\psi}(i,n))$$
$$=2\sum_1^{n-2}\{-i(m-i)P(N_n=i-2)+(n-i)\mathrm{E}\frac{N_iN(i,n)}{N(i,n)+1}1(N_n=i-1)$$
$$+iP(N_i=i)+P(N_i=i)P(N(i,n)\geq n-i+2)-(n-i)P(N_i=i)$$
$$-iP(N_i=i)e^{-(n-i)}\}.$$

Next we use formula (12) of the Appendix to compute

$$\mathrm{E}\{\frac{N_iN(i,n)}{N(i,n)+1}1(N_n=i-1)\}.$$

Putting things together and using that

$$\sum_1^{n-2}\{-i(n-i)P(N_m=i-2)+(n-i)\mathrm{E}\frac{N_iN(i,n)1(N_n=i-1)}{N(i,n)+1}\}$$
$$=\sum_1^{n-2}P(N_n=i-1)\{-i+n(\frac{i}{n})^i\}=-\frac{n}{2}+O(n^{\frac{1}{2}});$$
$$\sum_1^{n-2}P(N_i=i)P(N(i,n)\geq n-i+2)=O(n^{\frac{1}{2}});$$
$$\sum_1^{n-2}iP(N_i=i)\exp\{-(n-i)\}=O(n^{\frac{1}{2}}),$$

we finally obtain (20). □

Proposition 4. For $n\to\infty$,

$$(22)\quad 2\sum_{i=1}^{n-2}\{\phi(i,i+1)-\tilde{\phi}(i,i+1)+\tilde{\psi}(i,i+1)\}$$
$$=2\sum_{i=1}^{n-2}\{2iP(N_{i+1}\leq i)-(i+1)P(N_{i+1}=i-1)+iP(N_i\geq i+1)+$$

$$-iP(N_i \le i-1)\} - 2n + O(n^{\frac{1}{2}}),$$

(23) $$2\sum_{i=1}^{n-2}\{a(i,i+1) - \tilde{a}(i,i+1) + \tilde{b}(i,i+1)\}$$

$$= 2\sum_{i=1}^{n-2}\{iP(N_{i+1} \le i) + iP(N_i \ge i+1)\} - n + O(n^{\frac{1}{2}}).$$

Proof. The proof of (23) is straightforward and therefore omitted. We focus on the proof of (22). In the proof of Proposition 2 (last line) it was seen that for $j \ge i+2$,

$$\phi(i,j) - \tilde{\phi}(i,j) - \psi(i,j) + \tilde{\psi}(i,j) = P(N_i = i, N_j = j+1).$$

This relation is not valid for $j = i+1$, because according to (13),

$$\psi(i,i+1) = \mathrm{E}N_iY_{i+1}N(i,i+1)Y_{i+1} = \mathrm{E}N_iX_{i+1}y_{i+1}^2,$$

and hence we *cannot* apply (11) to calculate $\psi(i,i+1)$. Instead we can conclude that

(24) $$\phi(i,i+1) - \tilde{\phi}(i,i+1) + \tilde{\psi}(i,i+1) = P(N_i = i, N_{i+1} = i+2) + \chi(i),$$

where $\chi(i), i \ge 1$, is defined by (compare with the right-hand side of (11)),

$$\begin{aligned}\chi(i) :=\ & \mathrm{E}X_{i+1}N_i1(N_i < N_{i+1} \le i)\\ &+ i\mathrm{E}\frac{X_{i+1}N_i}{(N_i+1)}1(i \le N_i < N_{i+1} \le i)\\ &+ \mathrm{E}\frac{X_{i+1}N_i}{(X_{i+1}+1)}1(N_i \le i \le N_{i+1} \le i)\\ &+ \mathrm{E}\frac{X_{i+1}N_i}{(X_{i+1}+1)(X_{i+1}+2)}1(N_i \le i \le N_{i+1})\\ &+ i^2\mathrm{E}\frac{X_{i+1}N_i}{(N_i+1)(N_i+2)}1(i \le N_i, N_{i+1} \ge i+1)\\ &+ i\mathrm{E}\frac{X_{i+1}N_i}{(N_i+1)(X_{i+1}+1)}1(i \le N_i \le i \le N_{i+1})\\ =\ & A_1 + A_2 + \ldots + A_6.\end{aligned}$$

The calculation of $A_1, \ldots, A_6$ (except perhaps that of A_4) is straightforward,

$$\begin{aligned}A_1 &= iP(N_{i+1} \le i-2),\\ A_2 &= 0,\\ A_3 &= \frac{i}{e}P(N_i = i) - P(N_{i+1} = i-1),\end{aligned}$$

$$
\begin{aligned}
A_4 &= \mathrm{E}(N_i 1(N_i \le i)\{\phi(i+1-N_i) - 2\phi(i+2-N_i)\}) \\
&= 2iP(N_{i+1} = i) - \frac{2i}{e}P(N_i = i) + iP(N_{i+1} \le i-1) - iP(N_i \le i-1), \\
A_5 &= iP(N_i \ge i+1) - 2P(N_i \ge i+2), \\
A_6 &= \frac{i}{e}P(N_i = i+1).
\end{aligned}
$$

Next we use that

$$
\begin{aligned}
&2\sum_{i=1}^{n-2}\{\frac{i}{e}P(N_i = i+1) - \frac{i}{e}P(N_i = i)\} = O(n^{\frac{1}{2}}), \\
&2\sum_{i=1}^{n-2}\{-2P(N_i \ge i+2)\} = -2n + O(n^{\frac{1}{2}}),
\end{aligned}
$$

to conclude that

$$
\begin{aligned}
2\sum_{i=1}^{n-2}\chi(i) = 2\sum_{i=1}^{n-2}\{&2iP(N_{i+1} \le i) - (i+1)P(N_{i+1} = i-1) \\
&+iP(N_i \ge i+1) - iP(N_i \le i-1)\} - 2n + O(n^{\frac{1}{2}}).
\end{aligned}
$$

This proves (22), since

$$
2\sum_{i=1}^{n-2}P(N_i = i, N_{i+1} = i+1) = \frac{2}{e}\sum_{i=1}^{n-2}P(N_i = i) = O(n^{\frac{1}{2}}). \quad \square
$$

It follows from (5), (7), (15), Corollary 1, and the Propositions 3 and 4 that

$$
\begin{aligned}
&\mathrm{Var}(\sum_{i=1}^{n}(X_iY_i - 1)) = \mathrm{Var}(\sum_{i=1}^{n}X_iY_i) = \\
&= 3n + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{Cov}(X_iY_i, X_jY_j) + O(n^{\frac{1}{2}}) \\
&= 3n + (1 - \frac{2}{\pi})n + 2\sum_{i=1}^{n-2}iP(N_i = i) - n + \\
&\quad +2\sum_{i=1}^{n-2}\{iP(N_{i+1} \le i) - (i+1)P(N_{i+1} = i-1) \\
&\quad -iP(N_i \le i-1)\} - n \\
&\quad +2\phi(n-1,n) - 2a(n-1,n) + o(n).
\end{aligned}
$$

The proof is finished by noting that:

$$2\sum_{i=1}^{n-2}\{iP(N_i = i) - (i+1)P(N_{i+1} = i-1)\} = O(n^{\frac{1}{2}}),$$

$$2\sum_{i=1}^{n-2}\{iP(N_{i+1} \le i) - iP(N_i \le i-1)\}$$

$$= 2\{(n-1)P(N_{n-1} \le n-2) - P(N_1 = 0) - \sum_{i=1}^{n-2} P(N_{i+1} \le i)\} = O(n^{\frac{1}{2}}),$$

$$2\phi(n-1,n) - 2a(n-1,n) = o(n).$$

Appendix

In the formulæbelow the random variable ξ is Poisson distributed with parameter λ; $Y_1, \ldots, Y_n$ is a sample from the uniform distribution on the interval $(0, a)$; $Y_{(1)}, \ldots, Y_{(n)}$, the ordered sample, and $U_1, \ldots, U_{n+1}$ are defined by:

$$U_1 := Y_{(1)}, \quad U_i := Y_{(i)} - Y_{(i-1)}, \, 2 \le i \le n, \quad U_{n+1} := a - Y_{(n)}.$$

Further we denote by β a random variable wich is binomial distributed with parameters n and $p \in (0,1)$; $q = 1 - p$.

The relations below are well known and/or easily derived.

(1) $$\mathrm{E}(\xi 1(\xi < k)) = \lambda P(\xi < k-1),$$

(2) $$\mathrm{E}\{\frac{1}{\xi+1} 1(\xi \ge k)\} = \frac{1}{\lambda} P(\xi \ge k+1),$$

(3) $$\mathrm{E}\{\frac{1(\xi \ge k)}{(\xi+1)(\xi+2)}\} = \frac{1}{\lambda^2} P(\xi \ge k+2),$$

(4) $$\mathrm{E}\{\xi^2 1(\xi < k)\} = \lambda^2 P(\xi < k-2) + \lambda P(\xi < k-1),$$

(5) $$\mathrm{E}\frac{\xi^2}{(\xi+1)(\xi+2)} = 1 - \frac{1}{\lambda}(3 + e^{-\lambda}) + \frac{4}{\lambda^2}(1 - e^{-\lambda}),$$

(6) $$\mathrm{E}U_j = \frac{a}{n+1},$$

(7) $$\mathrm{E}U_j^2 = \frac{2a^2}{(n+1)(n+2)},$$

(8) $$\mathrm{E}U_i U_j = \frac{a^2}{(n+1)(n+2)}, \quad i \ne j,$$

(9) $$\mathrm{E}\beta^2 = npq + n^2p^2,$$

(10) $$\mathrm{E}\{\frac{1}{1+\beta}\} = (1 - q^{n+1})\frac{1}{(n+1)p},$$

$$\text{(11)} \quad \mathrm{E}\{\frac{1}{(1+\beta)(2+\beta)}\} = (1-q^{n+2}-(n+2)pq^{n+1})\frac{1}{(n+1)(n+2)p^2},$$

$$\text{(12)} \quad \mathrm{E}\{\frac{\beta(n-\beta)}{1+\beta}\} = nq - \frac{q}{p}(1-q^n).$$

We define for $j = 1, 2, \ldots$,

$$\text{(13)} \quad \phi(j) := \sum_{i=j}^{\infty} \frac{e^{-1}}{i!}.$$

New problems

Problem 311 (F. W. STEUTEL AND J. G. F. THIEMANN)

In "Probability" by G. Grimmett and D. Welsh the following problem is discussed. At time zero I amoebas are present. They split into two amoebas each after independent exponentially distributed times with expectation $1/\lambda$, and the "new-born" amoebas behave in the same fashion independent of all others. The number of amoebas present at time t is denoted by $M_I(t)$, and the result to be proved reads (Theorem 11D, p. 179),

$$P(M_I(t)=k) = \binom{k-1}{I-1} e^{-I\lambda t}(1-e^{-\lambda t})^{k-I}, \qquad (k = I, I+1, \ldots).$$

In the book the problem is attacked by use of the probability generating function of M_I, for which a partial differential equation is derived, whose solution is "a skill not required of the reader". Prove the result in a few lines.

Problem 312 (B. C. HOMBAS)

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the distribution for which the probability density function is

$$f(x,\theta) = 1 + \frac{2x-1}{6}\cos\theta + \frac{3x^2-1}{6}\sin\theta \quad x \in [0,1], \quad \theta \in [0, 100\pi].$$

a) Show that the distribution satisfies all regularity conditions for the existence of a consistent solution to the ML equation.
b) Show that for all n and all samples in general there are fifty solutions and only one is consistent.

Problems marked with * are nonelementary, of problems marked with ** no solution is known to the editor; unmarked problems are not necessarily simple, Solutions of the problems in this issue should arrive before September 30. 1994 or, in case the issue appears late, four weeks after its appearance. Problems (preferably with solutions) and solutions (type-written on separate sheets bearing the name and the address of the solver) are welcomed by the column editor.