



CHICAGO JOURNALS



The University of Chicago

Statistical Comments on Cohen's Application of a Simple Stochastic Population Model to Natural Primate Troops

Author(s): Niels Keiding

Reviewed work(s):

Source: *The American Naturalist*, Vol. 111, No. 982 (Nov. - Dec., 1977), pp. 1211-1219

Published by: [The University of Chicago Press](#) for [The American Society of Naturalists](#)

Stable URL: <http://www.jstor.org/stable/2460266>

Accessed: 19/10/2012 14:36

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press, The American Society of Naturalists, The University of Chicago are collaborating with JSTOR to digitize, preserve and extend access to *The American Naturalist*.

<http://www.jstor.org>

of selfings. However, there may be cases where, under the present model, high selfing would compensate a genotype for having low pollen fertility. In cases where $w_3 < w_2$, it will be more rational to accredit the pollen parent with all self offspring. The same would hold for the selfing component of hermaphrodites in gynodioecious populations; for example, under Ross and Weir's (1975) "limited-fitness model," pollen availability determines the extent to which ovules are utilized in selfings.

LITERATURE CITED

- Ross, M. D. 1977. Frequency-dependent fitness and differential outcrossing in hermaphrodite populations. *Amer. Natur.* 110:200-202.
- Ross, M. D., and B. S. Weir. 1975. Maintenance of male sterility in plant populations. III. Mixed selfing and random mating. *Heredity* 35:21-29.

AUGUSTE HOROVITZ

DEPARTMENT OF PLANT GENETICS
WEIZMANN INSTITUTE OF SCIENCE
REHOVOT, ISRAEL
June 3, 1977

STATISTICAL COMMENTS ON COHEN'S APPLICATION OF A
SIMPLE STOCHASTIC POPULATION MODEL TO NATURAL
PRIMATE TROOPS

Cohen (1969) gave some examples of application of a simple stochastic birth-immigration-death-emigration (BIDE) model to describe the natural dynamics of primate troops. While incorporating stochastic features into the description, such a model does not take into account any aspects of age structure, density dependence, or other social structures within the troops. The aim of this paper is not, however, to discuss whether the model is a biologically uninteresting oversimplification, but rather to reconsider the statistical aspects of Cohen's paper and to present some statistical methods which might also be useful in other similar studies.

Cohen compared the theory to two types of data: (1) size distributions of troops assumed to be in an equilibrium state and (2) a set of observations (Altmann and Altmann 1970) of all vital events in a baboon troop over a year. Distributions of troop sizes for howler monkeys, gibbons, colobus and langurs, and baboons were analyzed. For the gibbons I propose here an alternative and possibly more natural interpretation of the BIDE model which yields a considerably better fit of the equilibrium distribution.

The statistical analysis of the actual dynamics over 1 yr of the baboon troop

may benefit from the advances in statistical theory since the time of Cohen's paper (1969). I present here a direct (maximum-likelihood) attack on the inference problem as an alternative to Cohen's "jackknife" approach. (It may, in fact, be shown that the two approaches are approximately equivalent.) A reiteration of the jackknife calculations shows that some numerical slips must have occurred in Cohen's computations, and I present the corrected jackknife analysis as well. Also, the problem of testing the fit of the BIDE model is considered in some detail.

Most of the analyses suggest that the dynamics are reasonably well described by the BIDE model, and that the stationary distribution corresponding to the dynamical data is compatible with that of the other troops observed by Altmann and Altmann (this is opposite to Cohen's conclusion). On the other hand, I also point out that the pattern of vital events over the year of study might suggest some seasonality (which is not assumed in the BIDE model), and that the dynamically studied troop has a stationary expectation well below its actual size. This may suggest that the dynamics of the troop concern a "demographically active subgroup," or that some other form of heterogeneity within the troop has to be included in the description. A similar conclusion was reached by Cohen (1975) in a recent study of the social structure of orangutans. (The BIDE model assume equal birth, death, immigration, and emigration rates for all individuals in the population, and "averaging" of different individual rates over the troop such as suggested by Cohen [1969, p. 458] will not lead to a BIDE process for the whole troop.)

EQUILIBRIUM DISTRIBUTIONS OF GIBBON TROOPS

Cohen (1969, p. 464) suggested the following mechanism for gibbon troop formation: Each troop consists of a father and a mother and a variable number of children. Therefore, the births to a troop happen with an intensity independent of troop size and may therefore be described as "immigrations." In other words, the birth parameter λ in the BIDE process should be set equal to zero, and the equilibrium distribution is then Poisson. Cohen interpreted the size of the troop, including the parents, as governed by the BIDE process and accordingly fitted a 0,1-truncated Poisson distribution to the observed data. Another possibility would be to describe the number of children by a BIDE process and thus fit a complete Poisson distribution to the number of children, that is, to troop size minus two. This gives a much nicer fit (see table 1), although this fact should be considered in the light of the extra flexibility that the shifting of the size distribution allows.

DYNAMICS OF BABOON TROOPS

Cohen analyzed a set of observations on a baboon troop obtained in the Amboseli reserve, Kenya, by Altmann and Altmann (1970). Notice that baboon troops, in contrast to the gibbon troops discussed above, consist of several parents and children. The data (from Altmann and Altmann 1970) are given

TABLE 1

OBSERVED AND FITTED DISTRIBUTIONS OF NUMBERS OF CHILDREN IN GIBBON TROOPS

No. CHILDREN	<i>Hylobates lar</i>			TOTAL		
	Obs.*	Cohen's Fit	Poisson	Obs.†	Cohen's Fit	Poisson
0	8	10.9	8.4	12	14.4	11.4
1	15	12.5	14.9	17	15.3	18.3
2	12	10.7	13.0	14	12.2	14.8
3	9	7.3	7.6	9	7.8	8.0
≥ 4‡	5	7.6	5.0	5	7.2	4.6

* Mean number of children equals 1.756.

† Mean number of children equals 1.614.

‡ Interpreted as 4 when computing mean but as ≥ 4 in computing the expected values.

in table 2. I want to describe this by a BIDE process with intensities λ (birth), μ (death + emigration), and ν (immigration). Let us first estimate these parameters assuming the model to hold true and later develop some tests of the adequacy of the model.

The maximum-likelihood estimators (MLE), which were already given by Cohen, and estimates of their variances are derived in Appendix A. The numerical results are presented in table 3 along with the jackknife estimates computed according to the description by Cohen (1969) (the latter are different from Cohen's as regards μ and p).

It may be shown for occurrence/exposure rates such as the MLE of λ and μ that, if the number of subdivisions on which the jackknife method is based is large, the jackknife estimator will converge toward the occurrence/exposure rate itself, and the usual jackknife variance estimate will converge to an obvious variance estimate based on ML theory. In accordance with this result the ML and jackknife estimates of λ and μ are seen to be identical to two significant digits, and this trivially holds true for ν , $\hat{\nu}$ being a simple average. For p and r the difference between MLE and jackknife is still easily within the standard error.

With the MLE estimates, or the revised jackknife, there is no significant difference between the present estimates of $r = \nu/\lambda$ and $p = 1 - \lambda/\mu$ and those ($r = 2.99$, $p = 0.06$) obtained by Cohen from fitting the stationary negative binomial distribution to the sizes of 51 baboon troops observed by Altmann and Altmann. This conclusion is opposite to Cohen's.

TEST FOR FIT OF THE BIDE MODEL

I approach the problem of evaluating the fit of the model from three different viewpoints. First, the time intervals between events are compared to their expected exponential distribution; second, the pattern of events (births, deaths, and immigrations) is studied; and finally, the expected stationary distribution is computed.

TABLE 2
 BIRTH (B), IMMIGRATION (I), DEATH (D), AND EMIGRATION (E) IN
 ALTMANN'S MAIN STUDY TROOP

Row	After This Many Days	At This Troop Size	This Event Occurred
1.....	41	40	B
2.....	5	41	B
3.....	22	42	B
4.....	2	43	D
5.....	17	42	D
6.....	26	41	I
7.....	0	42	I
8.....	55	43	B
9.....	35	44	I
10.....	20	45	E
11.....	5	44	D
12.....	6	43	E
13.....	32	42	D
14.....	4	41	D
15.....	0	40	D
16.....	22	39	D
17.....	10	38	B
18.....	0	39	B
19.....	7	40	D
20.....	4	39	B
21.....	17	40	D
22.....	11	39	E
23.....	3	38	B
24.....	4	39	D
25.....	8	38	D
26.....	2	37	D
27.....	5	36	B
28.....	10	37	B
	Final	38	...

SOURCE.—Altman and Altmann 1970.

TABLE 3
 MAXIMUM-LIKELIHOOD AND JACKKNIFE ESTIMATES

	ML	SD	Jackknife	SD
λ	$6.5 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	$6.5 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$
μ	$9.7 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$	$9.7 \cdot 10^{-4}$	$3.3 \cdot 10^{-4}$
v	$8.0 \cdot 10^{-3}$	$4.6 \cdot 10^{-3}$	$8.0 \cdot 10^{-3}$	$5.7 \cdot 10^{-3}$
$r = v/\lambda$	12.4	8.2	7.8	12.5
$p = 1 - \lambda/\mu$33	.27	.40	.37
Correlation (r, p)37

First, as explained in Appendix B, the normalized waiting times

$$(\tau_{n+1} - \tau_n) \left(X_{\tau_n} + \frac{v}{\lambda + \mu} \right) \approx (\tau_{n+1} - \tau_n)(X_{\tau_n} + 5) = Y_n$$

should be exponentially distributed with expectation $(\lambda + \mu)^{-1}$ and independent. Figure 1 shows, on a logarithmic ordinate, one minus the observed and

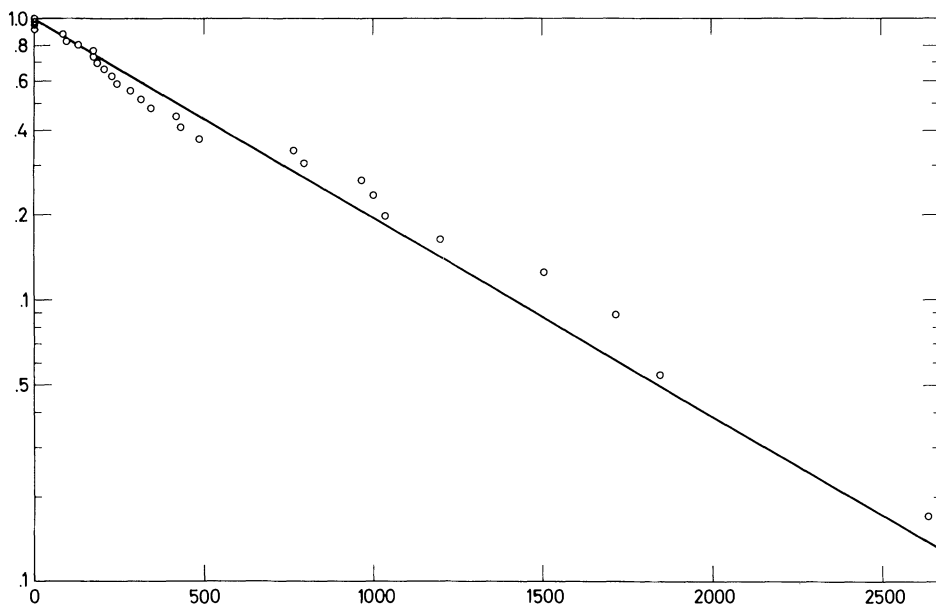


FIG. 1.—Observed (dots) and expected (line) distribution of normalized inter-event times.

expected distribution functions of the Y_n 's, indicating a fair fit. Also, under this hypothesis, $2 \sum Y_n/s$, where s is the standard deviation of the Y_n 's, should be approximately χ^2 distributed with 56 df. The observed value is 52.3, also indicating a good fit.

A useful way of obtaining an overall impression of the constancy of the intensities is to plot Aalen's (1975) nonparametric estimates of the integrated intensity. The plots are explained in Appendix C and shown in figure 2 for each of the three types of events: immigrations, births, and deaths + emigrations. Although the fit to the exponential interevent times was satisfactory above, a definite seasonality in the vital events is obvious, in particular concerning deaths and emigrations of which many more take place in the latter half of the year.

In fact, Aalen (personal communication) has applied the "cumulative total time on test statistic" by Barlow et al. (1972, p. 268) to the sequence of deaths and emigrations. This statistic tests constancy of the intensity against the alternative of increasing intensity. Aalen finds the value

$$T = \left(\sum_{i=1}^{n-1} \sum_{j=1}^i E_j \right) / \left(\sum_{i=1}^n E_i \right) = 10.23,$$

where

$$E_i = \int_{d_{i-1}}^{d_i} X_u du$$

is the exposure time of the process between the time d_i of the i th death or

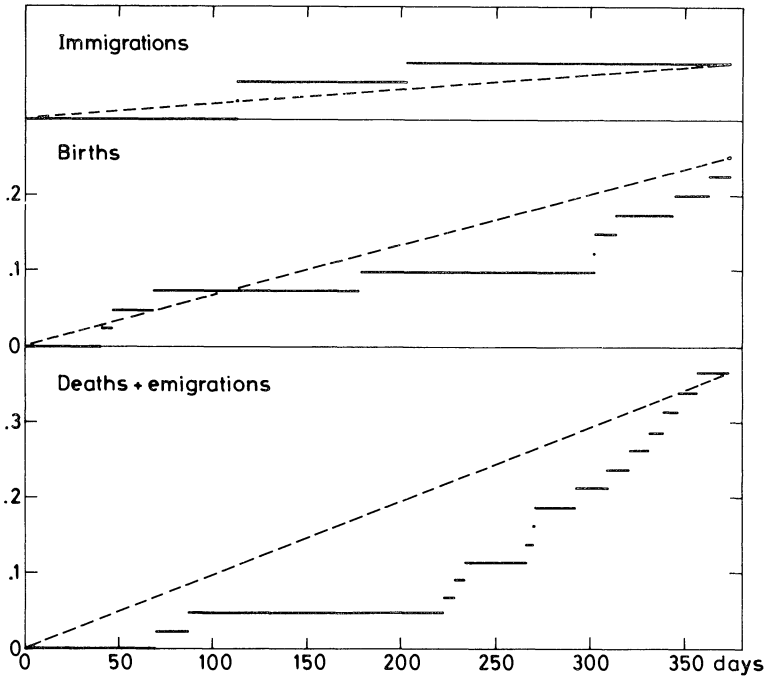


FIG. 2.—Aalen's (1975) nonparametric estimator of the integrated intensities compared with the expected straight lines.

emigration and d_{i+1} (set $d_0 = 0$). From Barlow et al. (1972, table 6.1), one can see that this will lead to rejection of the hypothesis of constant intensity at the .5% level.

I next propose a check on the sequence of vital events, without regard to calendar time. It will be very difficult to make any checks concerning the immigrations since only three of those were observed. If immigrations are disregarded, there results a linear birth-and-death process which may be considered as made up of waiting times and a random walk of births (B) and losses (D and E) (cf. Moran 1951, 1953). This imbedded random walk is here given by

BBB B BB B B BB
DD EDEDDDD D DE DDD

I tabulate numbers of adjacent pairs of events (D meaning D or E)

B	B	B	D	as	4	5	9
D	B	D	D		5	10	15
					9	15	24

The χ^2 test of independence gives $\chi^2 = 0.296$, $f = 1$, showing no deviation from the random-walk hypothesis.

Finally, much of Cohen's discussion assumes that the process is in a stationary regime, and it would therefore be interesting to compare the actual sample function with the corresponding stationary negative binomial distribution. The estimated expectation and standard deviation are 24.8 and 8.6. Although the standard error of the stationary expectation may be estimated as 14.3 (as explained in Appendix A), it is still striking that the process revolves slowly around the observed time average of 41.3; that is, under the model the process spends the time of observation in a rather extreme part of its stationary distribution. An exact use of this fact as counterevidence to the model would, however, need a study of the transient behavior of the BIDE process beyond the scope of the present paper. Another explanation would be that there is some heterogeneity in the troop such that the vital events governing the BIDE process correspond to what is going on in a "demographically active" subtroop (compare the fact that the stationary expectation is lower than the time average).

ACKNOWLEDGMENTS

This paper grew out of a visit by J. E. Cohen to the University of Copenhagen supported by Danish Natural Sciences Research Council. I am grateful to Joel Cohen for useful comments and for providing additional references. Substantial suggestions on an earlier draft were received from Odd Aalen, Niels Becker, Jan M. Hoem, Steffen Lauritzen, and Mats Rudemo.

APPENDIX A

ESTIMATION IN THE BIDE PROCESS

I refer to standard probabilistic treatments like that of Bailey (1964, pp. 91-101). Maximum-likelihood estimation in an immigration-death process was studied by Beneš (1957) and in a linear birth-and-death process by Keiding (1975). With notation as above, the likelihood function based on continuous observation of a BIDE process (X_u) in a time interval (0, t) is

$$v^I t \lambda^{B_t} \mu^{D_t} e^{-(\lambda + \mu) S_t - vt},$$

where I_t = number of immigrations, B_t = number of births, D_t = number of losses, and

$$S_t = \int_0^t X_u du = \text{time at risk.}$$

This gives (exactly as Cohen [1969, pp. 470-471]) the *ML* estimators $\hat{v} = I_t/t$, $\hat{\lambda} = B_t/S_t$, $\hat{\mu} = D_t/S_t$, $\hat{\rho} = \hat{v}/\hat{\lambda} = (I_t S_t)/t B_t$, $\hat{q} = \hat{\lambda}/\hat{\mu} = B_t/D_t$. Since we have a population size of about 40 and λt and μt around .3, the small-sample study of Beyer et al. (1976) will indicate that a large population approximation will be suitable for deriving asymptotic distributions. In fact, if both population size and

vt are large, $(\hat{\lambda}, \hat{\mu}, \hat{v})$ is asymptotically normal with asymptotically independent components (since the likelihood factorizes).

Now $\hat{v} = I_t/t$ is Poisson with expectation $E(\hat{v}) = v$ and variance $\sigma_v^2 = v/t$. For $\hat{\lambda}$ and $\hat{\mu}$, we get $-D_{\lambda}^2 \log L = B_t/\lambda^2$, $-D_{\mu}^2 \log L = D_t/\mu^2$, and an approximation to the variance σ_{λ}^2 of $\hat{\lambda}$ is then given by $(E[-D_{\lambda}^2 \log L])^{-1}$.

It is possible to derive $E(B_t)$, but since this will have to be estimated anyway, one might as well estimate σ_{λ}^2 by $\hat{\lambda}^2/B_t = B_t/S_t^2$ and σ_{μ}^2 by D_t/S_t^2 . By the usual theorem on differentiable transformations we get, approximately,

$$\begin{aligned} \text{Var}(\hat{r}) &= \frac{v^2}{\lambda^4} \sigma_{\lambda}^2 + \frac{1}{\lambda^2} \sigma_v^2, \\ \text{Var}(\hat{q}) &= \text{Var}(\hat{p}) = \frac{1}{\mu^2} \sigma_{\lambda}^2 + \frac{\lambda^2}{\mu^4} \sigma_{\mu}^2, \\ \text{Cov}(\hat{r}, \hat{p}) &= -\text{Cov}(\hat{r}, \hat{q}) = -\frac{v}{\mu\lambda^2} \sigma_{\lambda}^2. \end{aligned}$$

The numerical values of all these quantities in the present case are quoted in table 3. Finally, one may similarly derive

$$\text{Var}\left(\frac{\hat{r}\hat{q}}{\hat{p}}\right) = \text{Var}\left(\frac{\hat{v}}{\hat{\mu} - \hat{\lambda}}\right) = \frac{1}{(\mu - \lambda)^2} \sigma_v^2 + \left(\frac{v}{\mu - \lambda}\right)^2 (\sigma_{\mu}^2 + \sigma_{\lambda}^2),$$

where rq/p is the stationary expectation.

APPENDIX B

THE DISTRIBUTION OF THE INTEREVENT TIMES OF A BIDE PROCESS

Let $0 = \tau_0 < \tau_1 < \dots < \tau_n = t$ be the times at which events happen in the BIDE process (X_n) . It is then well known and follows in particular directly from the minimal construction of the process (Feller 1971, pp. 326 ff.) that given the development of the process up to the random time τ_i , the distribution of the interevent time $\tau_{i+1} - \tau_i$ is exponential with expectation

$$(\lambda + \mu)X_{\tau_i} + v.$$

It follows that the random variables $(\tau_{i+1} - \tau_i)[X_{\tau_i} + v/(\lambda + \mu)]$, $i = 0, \dots, n - 1$, are independently identically exponentially distributed with expectation $(\lambda + \mu)^{-1}$. References to this fact for $v = 0$, that is, for the linear birth-and-death process, and extensive applications to inference problems were given by Athreya and Keiding (1977). It is therefore reasonable to assume that since $v/(\lambda + \mu)$ is estimated by

$$\hat{v}/(\hat{\lambda} + \hat{\mu}) = 4.96 \approx 5, \quad Y_i = (\tau_{i+1} - \tau_i)(X_{\tau_i} + 5)$$

will be approximately independent identically distributed exponential with expectation $(\lambda + \mu)^{-1}$. This yields a way of checking the interevent time structure of the BIDE process. I do this by plotting the distribution of (Y_i) and by computing the sum $\sum Y_i$ and standard deviation s of the Y_i 's. Each Y_i is approximately χ^2 distributed with 2 df and scale parameter $[2(\lambda + \mu)]^{-1}$, and s is a consistent estimator of $(\lambda + \mu)^{-1}$ so that

$$2 \sum Y_i/s$$

is approximately χ^2 with $2n = 56$ df.

APPENDIX C

AALEN'S NONPARAMETRIC ESTIMATOR OF THE INTEGRATED INTENSITIES

Recently Aalen (1975) proposed a nonparametric estimator of the integrated intensity of a counting process (N_t) generalizing the well-known Kaplan-Meier nonparametric method of estimating a distribution function from censored data. Aalen's method assumes that the intensity has the multiplicative form $Y_t \alpha_t$, and in that case

$$\int_0^t \frac{1}{Y_s} dN_s$$

is an unbiased estimator of

$$\int_0^t \alpha_s ds.$$

(The interpretation of the estimator is as follows: Each time, t , the counting process $[N_s]$ makes a jump, an amount Y_t^{-1} is added to the estimator.) In the particular case of a constant intensity α , a plot of Aalen's estimator will, therefore, by comparison with the straight line αt , supply a check on the model.

LITERATURE CITED

- Aalen, O. O. 1975. Statistical inference for a family of counting processes. Ph.D. diss. University of California, Berkeley.
- Altmann, Stewart A., and Jeanne Altmann. 1970. Baboon ecology: African field research. Karger, Basel.
- Athreya, K. B., and N. Keiding. 1977. Estimation in continuous-time branching processes. In *Sankhyā*, Ser. A (in press).
- Bailey, N. T. J. 1964. The elements of stochastic processes with applications to the natural sciences. Wiley, New York.
- Barlow, R. E., D. J. Bartholomew, J. M. Bremner, and H. D. Brunk. 1972. Statistical inference under order restrictions. Wiley, London.
- Beneš, V. E. 1957. A sufficient set of statistics for a simple telephone exchange model. *Bell Syst. Tech. J.* 36:939-964.
- Beyer, J. E., N. Keiding, and W. Simonsen. 1976. The exact behavior of the maximum likelihood estimator in the pure birth process and the pure death process. *Scand. J. Statist.* 3:61-72.
- Cohen, J. E. 1969. Natural primate troops and a stochastic population model. *Amer. Natur.* 103:455-477.
- . 1975. The size and demographic composition of social groups of wild orang-utans. *Anim. Behav.* 23:543-550.
- Feller, W. 1971. An introduction to probability theory and its applications. Vol. 2. 2d ed. Wiley, New York.
- Keiding, N. 1975. Maximum likelihood estimation in the birth-and-death process. *Ann. Statist.* 3:363-372.
- Moran, P. A. P. 1951. Estimation methods for evolutive processes. *J. Roy. Statist. Soc. Ser. B*, 13:141-146.
- . 1953. The estimation of parameters of a birth and death process. *J. Roy. Statist. Soc., Ser. B*, 15:241-245.

NIELS KEIDING

INSTITUTE OF MATHEMATICAL STATISTICS
 UNIVERSITY OF COPENHAGEN
 COPENHAGEN, DENMARK
 August 26, 1976