



3 4456 0363138 1

*OK*

# ornl

ORNL/TM-12167

**OAK RIDGE  
NATIONAL  
LABORATORY**

**MARTIN MARIETTA**

## Overdispersion: Notes on Discrete Distributions

K. O. Bowman  
L. R. Shenton  
M. A. Kastenbaum  
Karl Broman

OAK RIDGE NATIONAL LABORATORY

CENTRAL RESEARCH LIBRARY

CIRCULATION SECTION

6300N. ROOM 175

**LIBRARY LOAN COPY**

DO NOT TRANSFER TO ANOTHER PERSON

If you wish someone else to see this  
report, send in name with report and  
the library will arrange a loan.

NON-PROFIT

MANAGED BY  
MARTIN MARIETTA ENERGY SYSTEMS, INC.  
FOR THE UNITED STATES  
DEPARTMENT OF ENERGY

This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401, FTS 626-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/TM-12167

Engineering Physics and Mathematics Division

405  
22

Mathematical Sciences Section

**OVERDISPERSION:  
NOTES ON DISCRETE DISTRIBUTIONS**

K.O. Bowman †  
L.R. Shenton †  
M.A. Kastenbaum \*  
Karl Broman \*

- † Mathematical Sciences Section  
Oak Ridge National Laboratory  
Oak Ridge, Tennessee 37831-6367
- ‡ University of Georgia  
Athens, Georgia 30602
- \* 16933 Timberlakes Dr, S.W.  
Ft. Myers, Florida 33908
- University of Wisconsin  
Milwaukee, Wisconsin 53211

Date Published: September, 1992

Prepared by the  
Oak Ridge National Laboratory  
Oak Ridge, Tennessee 37831  
managed by  
Martin Marietta Energy Systems, Inc.  
for the  
U.S. DEPARTMENT OF ENERGY  
under Contract No. DE-AC05-84OR21400



3 4456 0363138 1



## Contents

1	Introduction . . . . .	1
2	Beta Distribution . . . . .	1
	2.1 Beta Density for $p$ . . . . .	1
	2.2 Moment Estimators . . . . .	2
	2.3 Properties . . . . .	2
3	Generalized Mixtures of Binomials . . . . .	5
	3.1 Density on $(0, 1)$ . . . . .	5
	3.2 Transformed Beta Density . . . . .	5
	3.3 Pearson Type IV Density (Student's $t$ included) . . . . .	9
	3.4 Applications . . . . .	11
4	Truncated Densities for $p$ . . . . .	13
5	Further Binomial Mixtures . . . . .	15
	5.1 Exponential Density . . . . .	15
	5.2 Modified Exponential . . . . .	17
	5.3 Another Example . . . . .	17
	5.4 Gamma (Chi-squared) . . . . .	17
6	Miscellaneous Generalizations . . . . .	18
	6.1 Binomial Gamma-Gamma Mixtures . . . . .	18
	6.2 Mixtures and Multivariate Densities . . . . .	19
7	Mixtures of Logarithmic Distributions . . . . .	21
	7.1 General Forms . . . . .	21
	7.2 Gamma Density . . . . .	22
	7.3 Riemann Zeta Function $\zeta(s)$ Density . . . . .	24
	7.4 Uniform Density . . . . .	25
	7.5 Triangular Density . . . . .	25
	7.6 Trigonometric Density . . . . .	25
	7.7 Beta Density . . . . .	25
8	Concluding Remarks . . . . .	26
9	References . . . . .	26
A	Graphical Analysis . . . . .	28
	A.1 Model 1 . . . . .	28
	A.2 Model 2 . . . . .	28
	A.3 Model 3 . . . . .	29
	A.4 Model 4 . . . . .	29
	A.5 Model 5 . . . . .	30
	A.6 Model 6 . . . . .	30
	A.7 Model 7 . . . . .	30
	A.8 Model 8 . . . . .	31
	A.9 Model 9 . . . . .	31
	A.10 Model 10 . . . . .	32

A.11 Model 12	32
A.12 Model 13	33
A.13 Model 14	33
A.14 Model 15	34
A.15 Model 16	34
A.16 Model 17	35
A.17 Model 18	36
A.18 Model 19	37
A.19 Model 20	38
A.20 Model 22	38
A.21 Model 24	39
A.22 Model 25	40
A.23 Model 26	40
A.24 Model 27	41
B Three Dimensional Graphs	42
C Two Dimensional Graphs	72

### **Acknowledgements**

A research supported under contract ERD-89-828 with the Center for Indoor Air Research by Martin Marietta Energy Systems, Inc., under contract DE-AC05-84OR21400 with the U.S. Department of Energy. Karl Broman's work was supported by National Science Foundation, Research Experiences for Undergraduates program at University of Tennessee, Knoxville, Tennessee.



**OVERDISPERSION:  
NOTES ON DISCRETE DISTRIBUTIONS**

K.O. Bowman  
L.R. Shenton  
M.A. Kastenbaum  
Karl Broman

**Abstract**

We introduce mixtures of binomial distributions derived by assuming that the probability parameter  $p$  varies according to some law. We use the transformation  $p = \exp(-t)$  and consider various appropriate densities for the transformed variables. In the process, the Laplace transform becomes the fundamental entity. Large numbers of new binomial mixtures are generated in this way. Some transformations may involve several variates that lead to "multivariate" binomial mixtures. An extension of this to the logarithmic distribution, with parameter  $p$ , is possible. Frullani integrals and Laplace transforms are encountered.

Graphical representations of some of the more significant distributions are given. These include probability functions, regions of validity, and three dimensional representations of probability functions showing the response to variation of parameters when two parameters are involved.



## 1. Introduction

Here we describe new mixtures of binomial distributions in which the probability parameter, transformed exponentially, is taken to be a random variable with a known probability structure. The derived binomial mixtures may depend on several parameters, which, in general, can be estimated by using the first few sample moments. Once the parameters are estimated, the corresponding probabilities are readily computed by the use of finite differences. Such methods are applied to the data of Bender, *et al.* [1] on spontaneous chromosomal aberrations in human blood lymphocytes. These authors observe that the Poisson model works well in some cases, but that the Poisson characteristic, mean equals variance, does not always hold, and that, in general, the variances exceed the means. The new binomial mixtures presented here include this type of “overdispersed” model, as well as models for “underdispersed” data.

Various generalizations are mentioned including (a) representation for the parameter  $p$  as  $t^2/(a^2 + t^2)$ ,  $-\infty < t < \infty$ ; (b) truncated densities for  $p$ ; (c) multivariate densities for  $p$ . A link is given to deviates derived from other densities by varying the parameters. One typical case is the variation of the Poisson parameter as a  $\chi^2$  variate. Mixtures relating to the logarithmic distribution are studied.

Historically, Skellam [11] studied a beta modification of the binomial, deriving the expression for factorial moments, as well as an iterative scheme for maximum likelihood (m.l.) estimation of the two beta density parameters. He applied his results to data on the secondary association of chromosomes in *Brassica oleracea* (Catchside [3]). Shenton [10] studied the asymptotic efficiency of the method of moments, deriving the first terms in a rather complicated series. He pointed out that the Skellam modified-binomial may be regarded as a negative hypergeometric distribution which arises also in the relation between the binomial and negative binomial distributions. In a related study, Bowman, Kastenbaum and Shenton [2] show that the efficiency of the method of moments is fairly high over a segment of the parameter space, and does not contradict Shenton's [10] assertion that the method of moments rarely has low efficiency. Indeed the study shows a joint efficiency of 99% when the asymmetry of the basic distribution is low. The authors also present new expressions for the likelihood equations and a brief study of extended series for moments of the moment estimators.

## 2. Beta Distribution

### 2.1. Beta Density for $p$

The binomial probability function being

$$B(x; k, p) = \binom{k}{x} p^x q^{k-x}, \quad (x = 0, 1, \dots, k; 0 \leq p \leq 1, p + q = 1; k = 1, 2, \dots)$$

the beta modified form is

$$S(x; k, \alpha, \beta) = \int_0^1 \binom{k}{x} p^x (1-p)^{k-x} p^{\alpha-1} (1-p)^{\beta-1} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} dp$$

$$= \binom{k}{x} \frac{\Gamma(\alpha + x)}{\Gamma(\alpha)} \frac{\Gamma(\beta + k - x)}{\Gamma(\beta)} \frac{\Gamma(\alpha + \beta)}{\Gamma(k + \alpha + \beta)}.$$

$$(x = 0, 1, \dots, k; \alpha > 0, \beta > 0)$$

Let

$$\begin{aligned} P_0 = S(0; k, \alpha, \beta) &= \frac{\Gamma(\beta + k)}{\Gamma(\beta)} \frac{\Gamma(\alpha + \beta)}{\Gamma(k + \alpha + \beta)} \\ &= \frac{\{(k + \beta - 1) \cdots \beta\}}{(k + \alpha + \beta - 1) \cdots (\alpha + \beta)}. \end{aligned} \quad (1)$$

Then

$$\frac{P_{x+1}}{P_x} = \frac{k - x}{x + 1} \cdot \frac{\alpha + x}{k + \beta - x - 1}, \quad (x = 0, 1, \dots, k - 1) \quad (2)$$

is a simple recursive relationship, with  $\sum_x P_x = 1$ . It may be readily shown that

$$\alpha = \frac{-\mu'_1(\mu'_2 - k\mu'_1)}{k(\mu_2 - \mu'_1) + \mu_1'^2}, \quad \beta = \left(\frac{k}{\mu'_1} - 1\right) \alpha \quad (3)$$

where  $\mu'_s$  and  $\mu_s$  are, respectively, the  $s$ th crude and central moments of distribution.

## 2.2. Moment Estimators

For the moment estimators of  $\alpha$ ,  $\beta$  use equation (3), with sample moments  $m'_1$ ,  $m'_2$  replacing  $\mu'_1$ ,  $\mu'_2$ . Thus

$$\begin{cases} \alpha^* = m'_1(-m'_2 + km'_1)/\{k(m_2 - m'_1) + m_1'^2\} \\ \beta^* = (k/m'_1 - 1)\alpha^* \quad (m'_s = \sum_{i=1}^n x_i^s/n, m_2 = m'_2 - m_1'^2) \end{cases} \quad (4)$$

## 2.3. Properties

Table 1 demonstrates how  $\alpha$  and  $\beta$  values affect the modified distributions, and how the results compare with the basic binomial distribution. Values of  $P_x$ , for  $k = 5$ , ( $x = 0, \dots, k$ ), are derived from equations (1) and (2). Similar calculations will yield values of  $P_x$  for  $k = 20$ . These were omitted from Table 1 in the interest of conserving space. The binomial parameter  $p = \alpha/(\alpha + \beta)$ ,  $\mu'_1 = k\alpha/(\alpha + \beta)$ ,  $\sigma$  is the standard deviation, and  $\sqrt{\beta_1}$  and  $\beta_2$  are skewness and kurtosis parameters.

We note, from Table 1, that with increases in  $\alpha$ , the beta distribution approaches the binomial with  $p = \alpha/(\alpha + \beta)$ . For small  $\alpha$ , the beta distributions are skewed and platikurtic, and have larger standard deviations. In other words, these distributions are overdispersed in comparison with the binomial, in general. Illustrations of the densities for  $k = 20$  are given in Figures 1.

Table 1: Moments of the Beta Distribution

$k = 5$						
	$\beta = \alpha$			$\beta = 3\alpha/4$		
$\alpha$	4.0000	8.0000	Binomial	4.0000	8.0000	Binomial
$\sigma$	1.3437	1.2426	1.1180	1.3553	1.2454	1.1066
$\sqrt{\beta_1}$	0.0000	0.0000	0.0000	-0.1991	-0.1721	-0.1291
$\beta_2$	2.2531	2.4115	2.5528	2.2630	2.4234	2.5758
$x$	$P_x$					
0	0.0707	0.0511	0.0312	0.0455	0.0294	0.0145
1	0.1768	0.1703	0.1562	0.1299	0.1176	0.0964
2	0.2525	0.2786	0.3125	0.2165	0.2353	0.2570
3	0.2525	0.2786	0.3125	0.2597	0.2941	0.3427
4	0.1768	0.1703	0.1562	0.2273	0.2311	0.2285
5	0.0707	0.0511	0.0312	0.1212	0.0924	0.0609
	$\beta = \alpha/2$			$\beta = \alpha/4$		
$\alpha$	4.0000	8.0000	Binomial	4.0000	8.0000	Binomial
$\sigma$	1.3214	1.2054	1.0541	1.1547	1.0445	0.8944
$\sqrt{\beta_1}$	-0.5045	-0.4346	-0.3162	-1.1135	-0.9574	-0.6708
$\beta_2$	2.4818	2.5891	2.6838	3.6429	3.4551	3.0447
$x$	$P_x$					
0	0.0238	0.0128	0.0041	0.0079	0.0030	0.0003
1	0.0794	0.0641	0.0412	0.0317	0.0200	0.0064
2	0.1587	0.1648	0.1646	0.0794	0.0719	0.0512
3	0.2381	0.2747	0.3292	0.1587	0.1798	0.2048
4	0.2778	0.3022	0.3292	0.2778	0.3297	0.4096
5	0.2222	0.1813	0.1317	0.4444	0.3956	0.3277
$k = 20$						
	$\beta = \alpha$			$\beta = \alpha/4$		
$\alpha$	4.0000	8.0000	Binomial	4.0000	8.0000	Binomial
$\sigma$	3.9441	3.2540	2.2361	4.0658	3.3320	2.2131
$\sqrt{\beta_1}$	0.0000	0.0000	0.0000	-0.1835	-0.1447	-0.0645
$\beta_2$	2.4312	2.6444	2.7764	2.4243	2.6395	2.7879
	$\beta = \alpha/2$			$\beta = 3\alpha/4$		
$\alpha$	4.0000	8.0000	Binomial	4.000	8.000	Binomial
$\sigma$	4.0630	3.3076	2.108	3.6515	2.9542	1.7889
$\sqrt{\beta_1}$	-0.4717	-0.3743	-0.1581	-1.0563	-0.8463	-0.3354
$\beta_2$	2.6099	2.7596	2.8419	3.6911	3.4860	3.0224

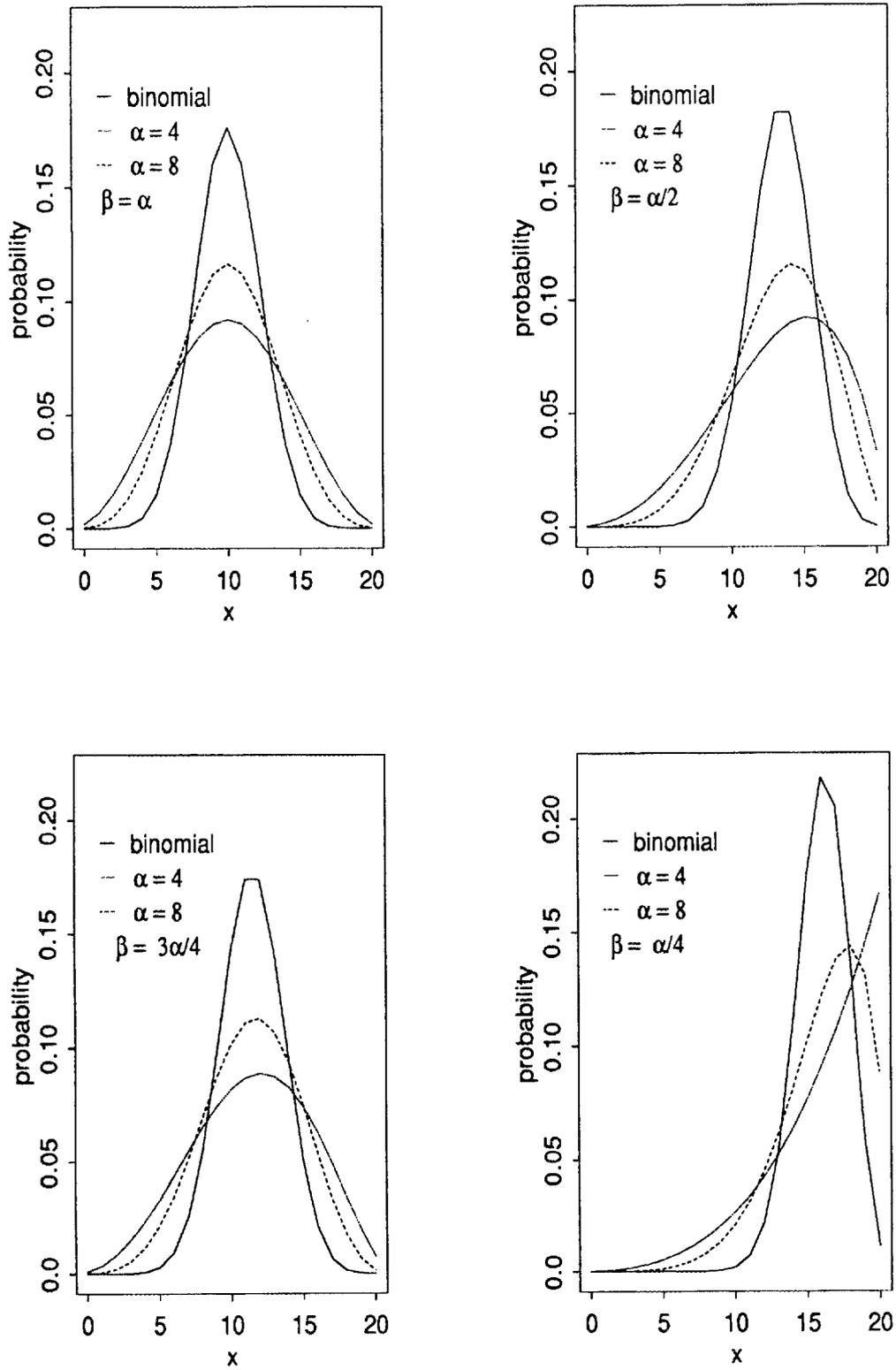


Figure 1:  $k = 20$

### 3. Generalized Mixtures of Binomials

#### 3.1. Density on (0, 1)

The Skellam [11] distribution treats the probability parameter  $p$  as a beta random variable with density

$$D(p) = p^{\alpha-1}(1-p)^{\beta-1}\Gamma(\alpha+\beta)/\{\Gamma(\alpha)\Gamma(\beta)\} \quad (\alpha, \beta > 0).$$

Families of binomial distributions exist that are generalizations of the beta-binomial. A more general binomial mixture is

$$B(x; n) = \binom{n}{x} \int_0^1 p^x(1-p)^{n-x} f(p) dp \quad (x = 0, \dots, n), \quad (5)$$

where  $f(p)$  is a valid density on interval (0, 1). The  $r$ th factorial moment,

$$\mu'_{[r]} = \int_0^1 \sum_{x=0}^n \binom{n}{x} x^{(r)} p^x(1-p)^{n-x} f(p) dp = n^{(r)} \int_0^1 p^r f(p) dp,$$

where  $x^{(r)} = x(x-1)\cdots(x-r+1)$ , is remarkable for its simplicity; note that  $\mu'_{[1]}$  represents the first non-central moment. For a probability parameter  $p$ , with distribution function  $\sigma(p)$ , the factorial moment generating function is

$$\int_0^1 (1+p\alpha)^n d\sigma(p).$$

From this expression, both non-central and central moments may be set up. However, the central moments may be computed more directly from  $\mu_r = E(x - \mu'_{[1]})^r$ . Hald [5] gives asymptotic results for forms like (5) with  $f(p)$  a density on (0, 1). In particular he gives numerical error values when  $f(\cdot)$  is a beta density.

#### 3.2. Transformed Beta Density

Since  $p$  satisfies  $0 \leq p \leq 1$ , we may consider the transformation,  $p = e^{-t}$ , ( $0 \leq t < \infty$ ) and the resulting, generalized binomial distribution

$$B_g(x; n, \phi) = \binom{n}{x} \int_0^\infty e^{-xt}(1-e^{-t})^{n-x} \phi(t; \underline{\alpha}) dt \quad (6)$$

where  $\phi(t; \underline{\alpha})$  is a valid density on  $(0, \infty)$ , with parameters  $\alpha_1, \dots, \alpha_m$ . More generally, we may use the Stieltjes form

$$B_g(x; n, \sigma) = \binom{n}{x} \int_0^\infty e^{-xt}(1-e^{-t})^{n-x} d\sigma(t).$$

For  $\phi(t; \underline{\alpha}) = 0$  (except when  $t = \log(1/p)$ ), (6) reduces to the binomial. Another possible transformation is  $p = t^2/(1+t^2)$ , with  $-\infty < t < \infty$ .

The probability function in (6) is expressible in terms of the Laplace transform of  $\phi(\cdot)$  or  $\sigma(\cdot)$ . Thus if

$$\Phi(s; \phi) = \int_0^{\infty} e^{-st} \phi(t; \underline{\alpha}) dt \quad (s > 0),$$

is assumed to exist, then

$$B_g(x; n, \phi) = \binom{n}{x} \sum_{r=0}^{n-x} (-1)^r \binom{n-x}{r} \Phi(x+r; \phi), \quad (7)$$

with a similar expression for the case of a discontinuous density. The summatory term in (7) will be recognized as an advancing difference.

Factorial moments readily result from

$$\mu'_{[r]} = n^{(r)} \int_0^{\infty} e^{-rt} \phi(t; \alpha) dt = n^{(r)} \Phi(r; \phi).$$

A list of possible forms for these generalizations is given in Table 2 and Table 3. It seems reasonable to describe these forms as mixtures which, to our knowledge, have not been described in the literature.

One case that does not arise obviously from the variation of parameters is model 29 for which

$$\phi_g(t) = k(Ae^{-at} + Be^{-bt} + \dots)/t^m; \quad (0 < t < \infty; m = 1, 2, \dots)$$

we assume positivity and integrability. In particular  $\int_0^{\infty} \phi(t) dt = 1$ . Then for the transform, (Hardy [6], p126), take

$$\Phi(s) = \frac{k(-1)^m}{(m-1)!} \{Aa^{m-1} \ln(a+s) + Bb^{m-1} \ln(b+s) + \dots\} = \frac{\sum Aa^{m-1} \ln(a+s)}{\sum Aa^{m-1} \ln a}.$$

For example, for the case

$$\phi(t) = k(Ae^{-at} + Be^{-bt} + Ce^{-ct} - e^{-dt})/t^3,$$

we may take as a valid density

$$\begin{aligned} A &= (d-b)(c-d)/\{(a+b)(c-a)\}, \\ B &= (a-d)(d-c)/\{(a-b)(b-c)\}, \\ C &= (b-d)(d-a)/\{(b-c)(c-a)\}, \end{aligned}$$

and similar confluent expressions, when there are equivalences. In particular, for the two-parameter case,

$$\phi(t) = ke^{-at}(1 - e^{-bt})^3/t^3, \quad (a, b > 0),$$

we have

Table 2: Generalized Binomial Distributions (1 - 15)

Model j	$\phi(t)$	$\Phi(s) = \int_0^\infty e^{-st}\phi(t)dt$	limits
1	$\begin{cases} \phi(t) = 0 \\ \phi(\ln \frac{1}{b}) = 1 \end{cases}$	$p^s$	$0 < p < 1$
2	$\frac{e^{-bt}t^{\rho-1}b^\rho}{\Gamma(\rho)}$	$\left(\frac{b}{b+s}\right)^\rho$	$b, \rho > 0$
3	$\frac{e^{-bt} - e^{-at}}{t \ln(a/b)}$	$\ln\left(\frac{a+s}{b+s}\right) / \ln\left(\frac{a}{b}\right)$	$a > b > 0$
4	$\frac{k}{2\sqrt{\pi t^3}} e^{-\frac{k^2}{4t}}$	$e^{-k\sqrt{s}}$	$k > 0$
5	$\frac{E_1(t/k)}{k}$	$\frac{\ln(1+ks)}{ks}$	$k > 0$
6	$\frac{e^{-bt} - e^{-at}}{(\sqrt{a}-\sqrt{b})2\sqrt{\pi t^3}}$	$\frac{\sqrt{s+a}-\sqrt{s+b}}{\sqrt{a}-\sqrt{b}}$	$b > a > 0$
7	$2aK_0(2\sqrt{at})$	$\int_0^\infty \frac{e^{-\theta}d\theta}{1+\theta s/a}$	$a > 0$
8	$\frac{e^{-at}(1+at) - e^{-bt}(1+bt)}{t^2(b-a)}$	$1 - \frac{s}{b-a} \ln\left(\frac{b+s}{a+s}\right)$	$b > a > 0$
9	$\frac{\int_0^\infty \left\{ \frac{e^{-s(t+B^{-1})}}{t+B^{-1}} - \frac{e^{-s(t+A^{-1})}}{t+A^{-1}} \right\} dt}{(B-A)}$	$\frac{\ln\{(Bs+1)/(As+1)\}}{(B-A)s}$	$B > A > 0$
10	$\frac{e^{-at} \left(1+at + \frac{a^2t^2}{2!}\right) - e^{-bt} \left(1+bt + \frac{b^2t^2}{2!}\right)}{t^2(b-a)/2}$	$\frac{\frac{b^2+2bs}{2(b+s)} - \frac{a^2+2as}{2(a+s)} - s \ln\left(\frac{b+s}{a+s}\right)}{(b-a)/2}$	$b > a > 0$
11	$\left(\frac{e^{-at}-e^{-bt}}{t}\right)^m e^{ct} / K(m, a, b; c)$	$\frac{K(m, a, b; c+s)}{K(m, a, b; c)}$	
12	$\frac{e^{-t/B} - e^{-t/A}}{B-A}$	$\frac{1}{(sA+1)(sB+1)}$	$B > A > 0$
13	$\frac{\left(\frac{t}{B}+1\right)e^{-\frac{t}{B}} - \left(\frac{t}{A}+1\right)e^{-\frac{t}{A}}}{2(B-A)}$	$\frac{2+s(A+B)}{2(sA+1)^2(sB+1)^2}$	$B > A > 0$
14	$\frac{2t^{(R-1)/2} K_{R-1}\{\sqrt{(t/A)}\}}{A^{(R+1)/2}\Gamma(R)}$	$\frac{1}{\Gamma(R)} \int_0^\infty \frac{e^{-u}u^{R-1}du}{1+Asu}$	$R, A > 0$
15	$e^{-at} \sin(t\sqrt{u}) \frac{a^2+u}{\sqrt{u}}$	$\frac{a^2+u}{(a+s)^2+u}$	$a, u > 0$

Functions in 5, 7, 11, and 14 are classical functions defined, for example, in *Handbook of Mathematical function*, AMS 55, National Bureau of Standard.

Table 3: Generalized Binomial Distributions (16 - 29)

Model j	$\phi(t)$	$\Phi(s) = \int_0^\infty e^{-st} \phi(t) dt$	limits
16	$\frac{e^{-at} \cos(t\sqrt{u})(u+a^2)}{a}$	$\left(\frac{a+s}{a}\right) \left(\frac{u+a^2}{u+(a+s)^2}\right)$	$a, u > 0$
17	$\frac{a^2+b^2}{a} \tanh\left(\frac{b\pi}{2a}\right) e^{-bt}  \sin(at) $	$\frac{a^2+b^2}{a^2+(s+b)^2} \frac{\coth\left(\frac{(b+s)\pi}{2a}\right)}{\coth\left(\frac{b\pi}{2a}\right)}$	$b > 0$
18	$\frac{b(b^2+4a^2)}{2a^2} e^{-bt} \sin^2(at)$	$\frac{b(b^2+4a^2)}{(b+s)\{(b+s)^2+4a^2\}}$	$b > 0$
19	$\frac{(b^3+4a^2b)e^{-bt} \cos^2(at)}{2a^2+b^2}$	$\frac{\{2a^2+(s+b)^2\}(b^3+4a^2b)}{(2a^2+b^2)\{(b+s)^3+4a^2(b+s)\}}$	$b > 0$
20	$\frac{4e^{-bt} \sin^2(at)}{t \ln(1+4a^2/b^2)}$	$\ln \frac{1+4a^2/(b+s)^2}{\ln(1+4a^2/b^2)}$	$b > 0$
21	$(2k/t)e^{-bt} \sinh(at)$	$k\{Ei(-a-B) - Ei(a-B)\}$	$t > 0, b > a$ $B = b + s$
22	$\frac{e^{-bt} \cosh\{2\sqrt{at}\}\sqrt{b}}{e^{a/b}\sqrt{\pi t}}$	$\frac{e^{a/(s+b)}\sqrt{b}}{e^{a/b}\sqrt{s+b}}$	$b, a > 0$
23	$\frac{k \sinh^2(\sqrt{bt})}{\sqrt{t}} e^{-bt}$	$\frac{k}{2} \sqrt{\frac{\pi}{b+s}} \left\{ e^{\frac{a}{b+s}} - 1 \right\}$	$(b > 0)$
24	$ce^{-at-k^2/(4t)}/(2\sqrt{\pi t^3})$	$e^{k\{\sqrt{a}-\sqrt{a+s}\}}$	$a > 0, k > 0$
25	$\frac{c}{\sqrt{k\pi}} e^{-at} \sinh\{2\sqrt{kt}\}$	$e^{\frac{k}{a+s} - \frac{k}{a}} / (1+s/a)^{3/2}$	$k > 0, a > 0$
26	$ce^{-at} \operatorname{erfc}\left(\frac{k}{2\sqrt{t}}\right)$	$\frac{e^{k\{\sqrt{a}-\sqrt{a+s}\}}}{1+s/a}$	$a > 0, k > 0$
27	$c\{e^{-at}e^{-k^2/(4t)}\}/\sqrt{\pi t}$	$\frac{e^{k\{\sqrt{a}-\sqrt{a+s}\}}}{\sqrt{1+s/a}}$	$k > 0, a > 0$
28	(see paragraph 3.3)	$\frac{\int_{-\infty}^{\infty} \frac{i^{2x} e^{-k \tan^{-1}(t) dt}}{(1+t^2)^n (1+t^2)^m}}{\int_{-\infty}^{\infty} \frac{e^{-k \tan^{-1}(t) dt}}{(1+t^2)^m}}$	$m > 1/2$
29	$\frac{k(Ae^{-at} + Be^{-bt} + \dots)}{t^m}$		$0 < t < \infty;$ $m = 1, 2, \dots$

$$\begin{aligned}\int_0^\infty \phi(t)dt &= \int_0^\infty k\{e^{-at} - 3e^{-(a+b)t} + 3e^{-(a+2b)t} - e^{-(a+3b)t}\}/t^3 dt \\ &= \frac{k}{2!}\{a^2 \ln a - 3(a+b)^2 \ln(a+b) + 3(a+2b)^2 b(a+2b) - (a+3b)^2 b(a+3b)\}.\end{aligned}$$

If

$$\begin{aligned}\Psi(s) &= (a+b)^2 \ln(a+s) - 3(a+b+s)^2 \ln(a+b+a) + 3(a+2b+s)^2 \ln(a+2b+s) \\ &\quad - (a+3b+s) \ln(a+3b+s),\end{aligned}$$

then

$$\Phi(s) = \Psi(s)/\Psi(0).$$

Clearly  $\Phi(s)$  is a valid density over  $s > 0$ .

### 3.3. Pearson Type IV Density (Student's $t$ included)

A different kind of mixture, Model 28, (Table 3), is based on Karl Pearson's Type IV density. In this case

$$\Phi(s) = \int_0^\infty e^{-st} \phi(t) dt = \int_{-\infty}^\infty \frac{t^{2x} e^{-k \tan^{-1}(t)} dt}{(1+t^2)^n (1+t^2)^m} / \int_{-\infty}^\infty \frac{e^{-k \tan^{-1}(t)} dt}{(1+t^2)^m} \quad (m > 1/2).$$

We use

$$\phi(t; a, k) = y_0 e^{-k \tan^{-1}(t/a)} / (t^2/a^2 + 1)^m \quad (-\infty < t < \infty; m \geq 1/2),$$

and

$$B_g(n; x, \phi) = \binom{n}{x} \int_{-\infty}^\infty \left( \frac{t^2}{t^2 + a^2} \right)^x \left( \frac{a^2}{t^2 + a^2} \right)^{n-x} \phi(t; a, k) dt,$$

which is independent of  $a$ , so that

$$B_g(n; x, \phi) = \binom{n}{x} \int_{-\infty}^\infty \frac{t^{2x}}{(1+t^2)^n} \phi(t; 1, k) dt.$$

Defining

$$Q(x; m+n, k) = y_0 \int_{-\infty}^\infty \frac{t^x e^{-k \tan^{-1} t} dt}{(1+t^2)^{m+n}} = Q_x, \quad \text{where } y_0 = \left\{ \int_{-\infty}^\infty \frac{e^{-k \tan^{-1} t} dt}{(t^2+1)^m} \right\}^{-1},$$

we can either integrate by parts and derive the recurrence

$$Q_x = \{(x-1)Q_{x-2} - kQ_{x-1}\} / (2N - x - 1) \quad (N = m+n; x = 1, \dots), \quad (8)$$

with  $Q_0$  to be determined, or make the transformation  $t = \tan \theta$  and consider integrals such as

$$G(M, k) = \int_{-\pi/2}^{\pi/2} \cos^M \theta e^{-k\theta} d\theta = \frac{M(M-1)}{M^2 + k^2} G(M-2, k), \quad (\text{see K. Pearson}[8]).$$

$Q_0$ , in (8), is given by

$$\begin{aligned} Q_0 &= \int_{-\pi/2}^{\pi/2} \cos^{2n+2m-2} \theta e^{-k\theta} d\theta / \int_{-\pi/2}^{\pi/2} \cos^{2m-2} \theta e^{-k\theta} d\theta \\ &= G(2n+2m-2, k) / G(2m-2, k), \\ &= \frac{(2n+2m-2)^{(2n)}}{\prod_{r=1}^n \{k^2 + (2n+2m-2r)^2\}} \quad (n+m \geq 1/2, n=0, \dots), \end{aligned}$$

and  $Q_1 = -kQ_0/(2n+2m-2)$ . Similarly for the  $r$ th factorial moment of the binomial distribution,

$$\mu'_{[r]} = y_0 n^{(r)} \int_{-\pi/2}^{\pi/2} \sin^{2r} \theta \cos^{2m-2} \theta e^{-k\theta} d\theta = n^{(r)} \sum_{s=0}^r (-1)^s \binom{r}{s} \frac{G(2m+2s-2, k)}{G(2m-2, k)}.$$

In particular,

$$\mu'_{[1]} = \frac{n(k^2 + 2m)}{(k^2 + 4m^2)}, \quad \mu'_{[2]} = n^{(2)} \frac{\{k^4 + (12m+4)k^2 + 12m(m+1)\}}{(k^2 + 4m^2)\{k^2 + 4(m+1)^2\}}.$$

Moment estimators ( $k^*$ ,  $m^*$ ), based on these equations, are found from the relationships

$$Am^{*3} + Bm^{*2} + Cm^* + D = 0, \quad k^{*2} = 2m^*(2m^*\nu_1 - 1)/(1 - \nu_1),$$

where

$$\nu_1 = m'_{[1]}/n, \quad \nu_2 = n^{(2)}/m'_{[2]},$$

$$A = \frac{4\nu_1^2(1-\nu_2)}{(1-\nu_1)^2} + \frac{8\nu_1}{1-\nu_1} + 4,$$

$$B = -\frac{4\nu_1(1-\nu_2)}{(1-\nu_1)^2} + \frac{4\nu_1(2-3\nu_2)}{1-\nu_1} + \frac{4-8\nu_1}{1-\nu_1} - 3\nu_2,$$

$$C = \frac{1-\nu_2}{(1-\nu_1)^2} + \frac{4\nu_1(1-\nu_2)}{1-\nu_1} - \frac{2(2-3\nu_2)}{1-\nu_1} + 4 - 3\nu_2,$$

$$D = -2(1-\nu_2)/(1-\nu_1).$$

The validity of these equations is conditioned by the restraints,  $m^* \geq 1/2$  and  $2m^*\nu_1 \geq 1$ .

Properties and figures, related to the generalized binomial distributions may be found in the appendix.

### 3.4. Applications

We now apply the moment approach to Catcheside's [3] data on the secondary association of chromosomes in *Brassica*, using Models 2, 15, and 28. (Note: There may be no solutions to the moment equations. In fact, solutions may exist that do not yield valid probabilities.) Thus for Model 2

$$m'_1 = 3 \left( \frac{b^*}{b^* + 1} \right)^{\rho^*} = 1.741840, \quad m'_{[2]} = 6 \left( \frac{b^*}{b^* + 2} \right)^{\rho^*} = 2.148368,$$

and  $b^* = 7.537223$ ,  $\rho^* = 4.363957$ , with  $\phi(t) = \phi_0 e^{-7.537223t} t^{3.363957}$ . For Model 15 the density is not always positive. However, the moments

$$\mu'_1 = \frac{k(a^2 + u)}{(a + 1)^2 + u}, \quad \mu'_{[2]} = \frac{k^{(2)}(a^2 + u)}{(a + 2)^2 + u},$$

yield solutions

$$a = -(\nu_1 - 4\nu_2 + 3\nu_1\nu_2)/(2\nu_1 - 4\nu_2 + 2\nu_1\nu_2),$$

and

$$u = \{(a + 1)^2\nu_1 - a^2\}/(1 - \nu_1),$$

where

$$\nu_1 = \mu'_{[1]}/k, \quad \nu_2 = \mu'_{[2]}/k^{(2)}, \quad k^{(2)} = k(k - 1).$$

Estimates of the Model 15 parameters are

$$a^* = 1.574527, \quad u^* = 3.264956.$$

For Model 28,

$$Pr(X = x) = \binom{k}{x} \int_{-\infty}^{\infty} \frac{t^{2x} e^{-w \tan^{-1}(t)} dt}{(1 + t^2)^k (1 + t^2)^m} / \int_{-\infty}^{\infty} \frac{e^{-w \tan^{-1}(t)} dt}{(1 + t^2)^m}. \quad (m > 1/2)$$

Moment estimators are found from

$$m'_{[1]} = k(w^2 + 2m)/(w^2 + 4m^2),$$

$$m'_{[2]} = k^{(2)} \frac{\{w^4 + (12m + 4)w^2 + 12m(m + 1)\}}{(w^2 + 4m^2)\{w^2 + 4(m + 1)^2\}}.$$

In Table 4, the fits of three generalized binomials (Model 2, 15 and 28) are compared with Skellam's modified binomial. Also the results of fitting Models 2 and 15 to a second set of data presented by Skellam appear in Table 4.

Here the computed moments are  $m'_1 = 0.465224$ , and  $m'_{[2]} = 0.44204$ , and the corresponding parameter estimates are

$$\left\{ \begin{array}{ll} \text{Model 2} & b^* = 1.1183, \quad \rho^* = 3.7171, \\ \text{Model 15} & a^* = -0.1062, \quad u^* = 0.0695. \end{array} \right.$$

Table 4: Skellam's Example

Secondary Association of Chromosomes in Brassicca					
Number of Associations	Observed	Skellam	Model 2	Model 15	Model 28
0	32	33.92	33.90	31.98	30.37
1	103	97.24	97.28	103.06	99.59
2	122	127.76	127.73	121.94	133.97
3	80	78.08	78.09	80.02	73.20
Degrees of Freedom		1	1	1	1
$\chi^2$		0.76	0.75	0.00007	1.91
Accident Data					
Number of Associations	Observed	Skellam	Model 2	Model 15	
0	447	448	448.32	442.07	
1	132	129	128.05	142.99	
2	42	47	46.77	38.45	
3	21	17	17.20	15.03	
4	3	5	5.49	6.28	
5 or more	2	1.4	1.17	2.18	
Degrees of Freedom		3	3	3	
$\chi^2$		2.60	3.18	5.32	

In this application we note that the Model 2 parallels the Skellam very closely, but that the Model 15 cannot be excluded as a reasonable fit.

The data presented by Bender, *et al.* [1] provide additional applications. These authors caution that their data, in many instances, depart from the Poisson, with variances generally larger than means. The results of applying Models 4, 24, and 15 to one of these data sets are presented in Table 5. For Model 4,

$$m'_1 = 3e^{-w^*} = 0.0028, \quad m'_{[2]} = 6e^{-w^*\sqrt{2}} = 0.00026,$$

and  $w^* = 6.9819$ , with  $\phi(t) = \phi_0 e^{12.19/t/t^{3/2}}$ .

For Model 24,

$$m'_1 = 3e^{w^*\{\sqrt{a^*} - \sqrt{a^*+1}\}} = 0.0028, \quad m'_{[2]} = 6e^{w^*\{\sqrt{a^*} - \sqrt{a^*+2}\}} = 0.00026.$$

and  $w^* = 7.4431$ ,  $a^* = 0.0041$  with  $\phi(t) = \phi_0 e^{-0.0041t - 13.85/t/t^{3/2}}$

For Model 15,

$$a^* = -0.4494, \quad u^* = -0.2018.$$

The parameters estimated for the Model 2 fit to the second set of data are  $b^* = 1.7718$ ,  $\rho^* = 5.2926$ .

Model 4 and Model 24 are variations of the Poisson distribution with the variance greater than the mean. As the  $\chi^2$  values show, they provide a satisfactory fit to the

Table 5: Bender's Data

Number of Aberrations	Observed	Poisson	Model 4	Model 24	Model 15
0	85571	85561.33	85573.77	85571.60	85570.72
1	220	238.34	213.93	218.11	220.84
2	8	0.33	12.30	9.95	7.16
3	1	-	-	0.34	1.28
Degrees of Freedom		1	1	1	1
$\chi^2$		229.20	1.06	1.68	0.16

Number of Aberrations	Observed	Poisson	Model 2
0	267	244.61	271.14
1	110	137.42	100.92
2	33	38.60	37.80
3	14	7.23	13.54
4	3	} 1.14	4.33
5	1		} 1.27
6	1		
Degrees of Freedom		3	3
$\chi^2$		27.73	2.33

data. Their meaning, however, in a biological context, remains to be explained.

#### 4. Truncated Densities for $p$

Finally we note that the generalized binomial mentioned in Section 1 belongs to the family

$$B(n; \phi) = \binom{n}{x} \int_0^1 p^x (1-p)^{n-x} \phi(p) dp,$$

$\phi(p)$  being a valid density on  $(0, 1)$ . If  $\phi(\cdot)$  is a point mass distribution, then the classical binomial is returned. A truncated form, namely,

$$B(n; \phi; u, v) = \binom{n}{x} \int_u^v p^x (1-p)^{n-x} \psi(p) dp \quad (0 \leq u < v \leq 1),$$

where  $\psi(\cdot)$ , a valid density on  $(u, v)$ , suggests itself. Thus the beta density,

$$B(n; \phi; u, v) = \binom{n}{x} \int_u^v p^x (1-p)^{n-x} \frac{(p-u)^{\alpha-1} (v-p)^{\beta-1} \Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)(v-u)^{\alpha+\beta-1}} dp, \quad (9)$$

involves four parameters  $\alpha$ ,  $\beta$ ,  $u$ , and  $v$ , where  $u, v$  relate to the range of  $p$ . A moment solution using the first two moments and the trial values of  $u, v$  may be optimized for a minimum  $\chi^2$  value.

Expression (9) may be written as

$$B(n; \phi; u, v) = \binom{n}{x} \int_0^1 \{u + (v - u)y\}^x \{1 - u - (v - u)y\}^{n-x} \frac{y^{\alpha-1}(1-y)^{\beta-1}\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} dy$$

involving four parameters  $u$ ,  $v$ ,  $\alpha$ , and  $\beta$ , the first two of which define the range of the probability parameter  $p$  in (9). The “binomial” probabilities  $B(n; \phi; u, v)$  may be evaluated by quadrature or by expanding the two binomial components of the integrand.

Given frequency data on the integers  $0, \dots, n$  where  $n \geq 6$  it is possible to use the first four sample moments  $m'_{[s]}$  ( $s = 1 \sim 4$ ) to determine a solution, if solutions exist. However,  $m'_{[3]}$  and  $m'_{[4]}$  may involve high sampling variation in general. An alternative is to select values of  $u$ ,  $v$ , and then use the first two moments  $m'_{[1]}$  and  $m'_{[2]}$  to determine  $\alpha^*$  and  $\beta^*$ . If  $\alpha^*$  and  $\beta^*$  are positive, then a unique solution is available. Thus, there are two criteria for deciding on the optimum choice of  $u$  and  $v$ :

- Minimize the goodness of fit  $\chi^2$  value;
- Compute the skewness and kurtosis, and look for the closest approximation to the normal values, (0, 3).

Define  $\nu_1 = m'_{[1]}/n$ ,  $\nu_2 = m'_{[2]}/n^{(2)}$ . Then

$$\alpha^* = \frac{(\nu_1 - u)\{(u + v)\nu_1 - uv - \nu_2\}}{(v - u)(\nu_2 - \nu_1^2)}, \quad \beta^* = \alpha^* \frac{(v - \nu_1)}{(\nu_1 - u)},$$

are solutions, provided they are positive.

It is also useful to guess at values of  $\alpha^*$ ,  $\beta^*$  in order to arrive at estimates of  $u$  and  $v$ . In this case

$$u^* = \nu_1 - \sqrt{\{(\nu_2 - \nu_1^2)(\alpha^* + \beta^* + 1)\alpha^*/\beta^*\}}, \quad v^* = u^* + (\nu_1 - u^*)(\alpha^* + \beta^*)/\alpha^*.$$

For an example, David and Johnson [4] gave data on the distribution of the number of defective teeth in 11 year old boys. These data were used by Rao and Chakravarti [9] for the test for a Poisson distribution. Here  $n = 12$ ,  $N = 265$ ,  $m'_{[1]} = 2.573585$ ,  $m'_{[2]} = 9.841509$ , with variance  $m_2 = 5.791755$ , greater than the mean.

For the model of equation (9) we calculate a Pearson  $\chi^2$ -value of 6.81, with five degrees of freedom, when  $u = 0.001$ ,  $v = 0.85$ ,  $\alpha^* = 0.942848$  and  $\beta^* = 2.80707$ . This fit (Table 6) appears to be satisfactory. By way of comparison, when  $u = 0$ , and  $v = 1$ , we found  $\alpha^* = 1.05056$ ,  $\beta^* = 3.84794$ , and  $\chi^2 = 6.78$  with seven degrees of freedom.

Another alternative is to use the model of Model 1 (Table 2) transformed onto (0, 1). We calculate a Pearson  $\chi^2$ -value of 7.01, with five degrees of freedom when  $u=0.001$ ,  $v = 0.800$ ,  $b^* = 1.2898$  and  $\rho^* = 2.2996$ . By way of comparison, when  $u = 0$  and  $v = 1$ , we found  $b^* = 1.7506$ ,  $\rho^* = 3.4073$ , and  $\chi^2 = 7.30$  with seven degrees of freedom.

In a number of cases which we have studied, the introduction of the interval parameters ( $u, v$ ) does not always provide improvement.

Table 6: David and Johnson Data

Defected teeth	Number of boys	Poisson	Equation (9)	Model 1
0	61	20.21	61.81	61.58
1	47	52.01	50.10	50.60
2	43	66.93	40.47	40.48
3	35	57.41	32.36	32.13
4	28	36.94	25.34	25.14
5	15	19.01	19.26	19.19
6	20	8.16	14.06	14.12
7	5	3.00	9.69	9.82
8	5	0.96	6.16	6.27
9	2			
10	1			
11	2	0.37	5.74	5.69
12	1			
Degrees of Freedom		8	5	5
$\chi^2$		224.30	6.81	7.00

## 5. Further Binomial Mixtures

### 5.1. Exponential Density

In Section 3, we introduced binomial mixtures defined as

$$\Pr(X = x) = \int_0^{\infty} e^{-xt}(1 - e^{-t})^{k-x} \phi(t) dt \quad (x = 0, \dots, n; k = 1, \dots)$$

where  $\phi(t)$  is a valid probability density on  $(0, \infty)$ . For the most part we do not use the purely discontinuous case. Here we show a link with distributions arising as probability densities subject to random variable parameters; for example, the Skellam [11] distribution, which assumes that the binomial probability parameter  $p$  has a beta distribution.

In addition, we list further examples of binomial mixtures including multivariate densities which may have interest or applicability. Diagrams are presented to give some insight into the question of modality of mixtures, because we have encountered serious problems with approaches that are purely mathematical.

Let

$$\phi_n(t, \alpha) = \alpha e^{-\alpha t} \quad (t > 0; \alpha > 0)$$

be the basic distribution, where  $\alpha$  has the distribution

$$\phi_p(\alpha) = k/\alpha, \quad (0 < A \leq \alpha < B; k = 1/\ln(B/A)). \quad (10)$$

Then the (variated) basic distribution becomes

$$\begin{aligned}\phi_v(t) &= \{e^{-At} - e^{-Bt}\}/\{t \ln(B/A)\}, & (0 < t < \infty) \\ &= 0, & (t < 0).\end{aligned}\tag{11}$$

The corresponding binomial mixture is

$$\Pr(X = x) = \int_0^\infty e^{-xt}(1 - e^{-t})^{k-x}\phi_v(t)dt.\tag{12}$$

Upon expanding the second binomial factor in the integrand, we require the fundamental entities

$$\Phi(s) = \int_0^\infty e^{-st}\phi_v(t)dt,$$

which, from (12), yields  $(A \rightarrow A + s, B \rightarrow B + s)$

$$\Phi(s) = \ln\{(B + s)/(A + s)\}/\ln(B/A).$$

We notice the role played by Laplace transforms in (12), and the appearance of Frullani integrals (see Hardy [6]) from (12). We have, for simplicity, sometimes omitted the parameter in probabilities such as  $\phi_v(t)$ , which doubtless should be written  $\phi_v(t; A, B)$ .

There are several generalizations. First, take the cases:

$$\begin{aligned}\phi_b(x) &= \alpha^2 x e^{-\alpha x} & (x > 0) \\ &= 0 & (x < 0)\end{aligned}$$

with  $\alpha$  varying as (10). Then

$$\begin{aligned}\phi_v(t) &= \{(1 + At)e^{-At} - (1 + Bt)e^{-Bt}\}/\{t \ln(B/A)\}, & (t > 0) \\ &= 0, & (t < 0)\end{aligned}$$

and

$$\Phi(s) = \left\{ \ln \left( \frac{B + s}{A + s} \right) + \left( \frac{A}{A + s} - \frac{B}{B + s} \right) \right\} / \ln(B/A),$$

and

$$\Pr(X = x) = \binom{k}{x} \left\{ \Phi(x) - \binom{k-x}{1} \Phi(x+1) + \binom{k-x}{2} \Phi(x+2) \dots \right\}.$$

Take another case:

$$\phi_b(t) = \alpha^{m+1} t^m e^{-\alpha t} / m!, \quad (t > 0; \alpha > 0, m = 0, \dots).$$

Then if  $\alpha$  varies as in (10)

$$\phi_v(t) = \{H_m(t, A) - H_m(t, B)\}/\{t \ln(B/A)\}, \quad (t > 0)$$

where

$$H_m(t, k) = e^{-kt} \sum_{\lambda=0}^m \frac{k^\lambda t^\lambda}{s!}, \quad (13)$$

and

$$\Phi(s) = \left\{ \ln \left( \frac{B+s}{A+s} \right) + \sum_{\lambda=1}^m \frac{1}{\lambda} \left( \frac{A^\lambda}{(A+s)^\lambda} - \frac{B^\lambda}{(B+s)^\lambda} \right) \right\} / \ln(B/A).$$

### 5.2. Modified Exponential

$$\phi_b(t) = e^{-t/\alpha} / \alpha, \quad (0 < t < \infty; \alpha > 0)$$

with  $\alpha$  varying as in (10). Then

$$\phi_v(t) = E_1(B^{-1}) - E_1(A^{-1})$$

where  $E_1(x)$  is the exponential integral. Thus

$$\phi_v(t) = \int_0^\infty \left( \frac{e^{-t(\theta+B^{-1})}}{\theta+B^{-1}} - \frac{e^{-t(\theta+A^{-1})}}{\theta+A^{-1}} \right) \frac{d\theta}{(B-A)}$$

and for the binomial mixture

$$\Phi(s) = \frac{1}{(B-A)s} \ln \left( \frac{Bs+1}{As+1} \right).$$

(Note that  $\Phi(0) = \lim_{s \rightarrow 0} \Phi(s) = 1$ ).

### 5.3. Another Example

$$\phi_b(t) = \alpha^3 t^2 e^{-\alpha t} / 2, \quad (t > 0; \alpha > 0)$$

where

$$\phi_p(\alpha) = 1 / \{\alpha \ln(B/A)\}. \quad (A < t < B; B > A > 0)$$

Then

$$\phi_v(t) = \{H_2(t, A) - H_2(t, B)\} / \{t \ln(B/A)\}, \quad (t > 0)$$

(see (13)) and

$$\Phi(s) = \left\{ \ln \left( \frac{B+s}{A+s} \right) + \frac{A}{A+s} - \frac{B}{B+s} + \frac{1}{2} \left( \frac{A^2}{(A+s)^2} - \frac{B^2}{(B+s)^2} \right) \right\} / \{\ln(B/A)\}.$$

### 5.4. Gamma (Chi-squared)

$$\phi_b(t) = \frac{t^{\rho-1} e^{-t/\phi}}{\phi^\rho \Gamma(\rho)}, \quad (0 < t < \infty; \rho > 0, \phi > 0)$$

where  $\phi$  has the density

$$\phi_p(t) = \frac{t^{r-1}e^{-t/A}}{A^r\Gamma(r)}. \quad (r > 0; A > 0, t > 0)$$

Then

$$\phi_v(t) = \int_0^\infty \frac{t^{\rho-1}e^{-t/\phi}}{\phi^\rho\Gamma(\rho)} \frac{\phi^{r-1}e^{-\phi/A}}{A^r\Gamma(r)} d\phi = \frac{t^{\rho-1}}{A^r\Gamma(\rho)\Gamma(r)} \int_0^\infty \phi^{r-\rho-1} e^{-\frac{t}{\phi} - \frac{\phi}{A}} d\phi$$

which in terms of a Bessel function, gives

$$\phi_v(t) = \frac{2}{A^{(r+1)/2}\Gamma(r)} t^{(r-1)/2} K_{r-1}\{2\sqrt{t/A}\}.$$

For the Laplace transform function, after simplification,

$$\Phi(s) = \frac{1}{\Gamma(r)} \int_0^\infty \frac{e^{-u}u^{r-1}}{1+Aus} du.$$

## 6. Miscellaneous Generalizations

### 6.1. Binomial Gamma-Gamma Mixtures

Let

$$g(t; a, r) = e^{-at}t^r a^{r+1}/r! \quad (a > 0, r > -1; 0 < t < \infty)$$

and

$$h(t; r) = e^{-t}t^r/r!,$$

so that  $\max(h)$  occurs when  $t = r$ . It follows that

$$H(t; r) = e^{-t}t^r/(e^{-r}r^r) \geq 1 \quad \text{for } r > 0.$$

The binomial gamma-gamma mixture is

$$\Pr(X = x) = \binom{k}{x} \int_0^\infty H^x(t; r_0) \{1 - H(t; r_0)\}^{k-x} g(t; a, r) dt, \quad (r_0 > 0, a > 0, r > -1)$$

and

$$\begin{aligned} \Phi(s; k, r_0; a, r) &= \int_0^\infty H^s(t; r_0) g(t; a, r) dt = \int_0^\infty \frac{e^{-st}t^{sr_0}}{e^{-sr_0}r^{sr_0}} \cdot \frac{e^{-at}t^r a^{r+1}}{r!} dt \\ &= \frac{e^{sr_0} a^{r+1} (sr_0 + r)!}{r_0^{sr_0} r! (a + s)^{sr_0 + r + 1}} \end{aligned}$$

with

$$\Phi(0; k, r_0; a, r) = 1.$$

The binomial gamma-gamma mixture is now defined as

$$\Pr(X = x) = \binom{k}{x} \sum_{\lambda=0}^{k-x} (-1)^\lambda \binom{k-x}{\lambda} \Phi(x + \lambda, k, r_0; a, r),$$

depending on the parameters  $k, r_0, a$ , and  $r$ .

In particular with  $r_0 = 1$

$$\Pr(X = x) = \binom{k}{x} \int_0^\infty (te^{1-t})^x (1 - te^{1-t})^{k-x} \frac{a^{r+1} e^{-at} t^r}{r!} dt, \quad (r > -1)$$

and

$$\Phi(s; r) = a^{r+1} e^s (s+r)! / \{(a+s)^{r+s+1} r!\},$$

with

$$\mu'_1 = ka^{r+1}(r+1)e/(a+1)^{r+2}.$$

## 6.2. Mixtures and Multivariate Densities

Consider the mixture

$$\binom{k}{x} \int_0^1 \int_0^1 \cdots \int_0^1 q_m^x (1 - q_m)^{k-x} \prod_{r=1}^m \phi_r(p_r) dp_r, \quad (q_m = \prod_{r=1}^m p_r)$$

where the density of  $p_1$  is  $\phi_1(p_1)$  for  $(0 \leq p_1 \leq 1)$ , and 0 elsewhere, and similarly for the remaining variates  $p_2, p_3, \dots, p_m$ , which we assume to be independent. Then by expansion

$$\Pr(X = x) = \binom{k}{x} \left\{ \Phi_m(x) - \binom{k-x}{1} \Phi_m(x+1) + \cdots + (-1)^k \binom{k-x}{k-x} \Phi_m(k) \right\}$$

where

$$\Phi_m(s) = \Psi_1(s) \Psi_2(s) \cdots \Psi_m(s),$$

and

$$\Psi_1(s) = \int_0^1 p_1^s \phi_1(p_1) dp_1.$$

Similarly for  $\Psi_2(s), \dots, \Psi_m(s)$ . In the case of identically and independently distributed (i.i.d.) variates,

$$\Pr(X = x) = \binom{k}{x} \left\{ \Psi_m(x) - \binom{k-x}{1} \Psi_m(x+1) + \cdots \right\}, \quad (x = 0, 1, \dots, k).$$

Examples

- Let the common density be the uniform  $U(0, 1)$ . Then

$$\Pr(X = x) = \binom{k}{x} \left\{ \frac{1}{(x+1)^m} - \binom{k-x}{1} \frac{1}{(x+2)^m} + \binom{k-x}{2} \frac{1}{(x+3)^m} \cdots \right\},$$

$$(m = 1, 2, \dots; x = 0, 1, \dots, k).$$

If  $m = 1$ , the binomial mixture becomes the discrete uniform with probabilities  $1/(k+1)$ .

The factorial moments are,

$$\mu'_{[r]} = k^{(r)}/(r+1)^m.$$

and clearly as  $m \rightarrow \infty$ , the binomial mixture tends to a point mass at  $x = 0$ . *In general the multivariate case tends to concentrate the binomial probabilities near the origin.*

- If the variate follows a beta density with parameters  $\alpha$  and  $\beta$ , then, in the multivariate case (i.i.d.)

$$\Pr(X = x) = \binom{k}{x} \int_0^1 \int_0^1 \cdots \int_0^1 (1 - \prod_1^m p_r)^{k-x} \prod_{r=1}^m \left\{ \frac{p_r^{\alpha-1} (1-p_r)^{\beta-1} \Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} dp_r \right\},$$

with

$$\Phi_m(s; \underline{p}; \alpha, \beta) = \left\{ \frac{\Gamma(\alpha+s)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(s+\alpha+\beta)} \right\}^m, \quad (m = 1, 2, \dots).$$

Alternatively, one may use

$$\Pr(X = x) = \binom{k}{x} \int \frac{e^{-tx} (1 - e^{-t})^{k-x} dt}{(B - A)},$$

the integration being from  $A$  to  $B$  ( $B > A > 0$ ). Here

$$\Phi_m(s; A, B) = \left\{ \frac{e^{-sA} - e^{-sB}}{s(B - A)} \right\}^m.$$

For example, in Skellam's example ( $k = 3$ ) we could try  $A = 0, B > 0$  with

$$\mu'_1 = k \left( \frac{1 - e^{-B}}{B} \right)^m.$$

- For the mixture use

$$\Pr(X = x) = \binom{k}{x} \int_0^\infty \int_0^\infty \cdots \int_0^\infty e^{-xT} (1 - e^{-T})^{k-x} \prod_{\theta=1}^m a_\theta e^{-a_\theta t_\theta} dt_\theta$$

where  $T = t_1 + t_2 + \cdots + t_m$ , for which

$$\Phi_m(s) = \prod_{\theta=1}^m \frac{a_\theta}{(s + a_\theta)}, \quad (a_\theta > 0, \theta = 1, 2, \dots, m).$$

depending on  $m$  parameters. If the range of each variate is  $0 \leq A \leq t < B$ , then

$$\Phi_m(s) = \left\{ \frac{e^{-sA} - e^{-sB}}{s(B - A)} \right\}^m \quad (s = 1, 2, \dots)$$

- For another example, take

$$\Pr(X = x) = \binom{k}{x} \int_0^A e^{-xt} (1 - e^{-t})^{k-x} dt / A,$$

$$\Phi(s) = \left( \frac{1 - e^{-sA}}{sA} \right) \quad (A > 0).$$

### Multivariate

$$\Phi^*(s) = \{\Phi(s)\}^m \quad (m = 1, 2, \dots).$$

Guess  $m$  (positive integer), and solve  $\mu'_1 = k\{\Phi(1)\}^m$  for  $A$ . Search for  $\min \chi^2$ .

- For yet another example, take

$$\Pr(X = x) = \binom{k}{x} \int_0^A e^{-xt} (1 - e^{-t})^{k-x} \frac{2t dt}{A^2} \quad (A > 0),$$

$$\Phi(s) = 2 \left\{ \frac{1 - (As + 1)e^{-As}}{A^2 s^2} \right\}.$$

For the multivariate case (i.i.d.), we have  $\Phi_m(s)$  is  $\{\Phi(s)\}^m$ .

## 7. Mixtures of Logarithmic Distributions

### 7.1. General Forms

For the logarithmic distribution

$$\Pr(X = x) = p^x / \{x \ln(1 - p)^{-1}\}, \quad (x = 1, 2, \dots; 0 < p < 1).$$

Following the ideas in Section 3, we consider the mixture

$$\Pr(X = x) = \frac{1}{x} \int_0^\infty \frac{(1 - e^{-t})^x}{t} \phi(t; \beta) dt \quad (14)$$

for a valid density (parameters  $\beta_1, \beta_2, \dots$ ) on  $(0, \infty)$ . The probability generating function (p.g.f.) is

$$P(\omega) = - \int_0^\infty \frac{\phi(t; \beta)}{t} \ln(\omega e^{-t} + 1 - \omega) dt = \sum_1^\infty \omega^x \Pr(X = x), \quad (|\omega| \leq 1) \quad (15)$$

with factorial moments, if they exist,

$$\begin{aligned} \mu'_{[1]} &= \int_0^\infty \frac{e^t(1-e^{-t})}{t} \phi(t; \beta) dt, \\ \mu'_{[2]} &= (s-1)! \int_0^\infty \frac{e^{st}(1-e^{-t})^s}{t} \phi(t; \beta) dt, \quad (s = 2, 3, \dots). \end{aligned}$$

## 7.2. Gamma Density

Here

$$\phi(t; b, \rho) = b^\rho e^{-bt} t^{\rho-1} / \Gamma(\rho), \quad (b, \rho, t > 0)$$

and

$$\Pr(X = x; b, \rho) = \frac{1}{x} \int_0^\infty \frac{(1 - e^{-t})^x}{t} \frac{b^\rho e^{-bt} t^{\rho-1}}{\Gamma(\rho)} dt.$$

Factorial moments of order  $s$  or less exist, if  $b > s$ . In fact

$$\mu'_{[s]} = s! \left( \frac{b}{b-s} \right)^\rho \Pr(X = s; b-s, \rho) \quad (b > s)$$

There are three cases to consider:

- $\rho > 1$ . By expansion of  $(1 - e^{-t})^x$ ,

$$\Pr(X = x; b, \rho) = \frac{b}{(\rho-1)x} \frac{1}{x} \sum_{r=0}^x (-1)^r \binom{x}{r} \left( \frac{b}{b+r} \right)^{\rho-1}.$$

Note if  $\rho = 2$ , then

$$\begin{aligned} \Pr(X = x; b, 2) &= \frac{b^2}{x} \int_0^\infty (1 - e^{-t})^x e^{-bt} dt = \frac{b^2}{x} \int_0^1 (1-p)^x p^{b-1} dp \\ &= \frac{b^2}{x} \frac{\Gamma(x+1)\Gamma(b)}{\Gamma(x+b+1)} = \frac{b^2(x-1)!\Gamma(b)}{\Gamma(x+b+1)}. \end{aligned}$$

For example,

$$\Pr(X = 1; b, 2) = b/(b + 1)$$

$$\Pr(X = 2; b, 2) = 1!b/\{(b + 2)(b + 1)\}$$

$$\Pr(X = 3; b, 2) = 2!b/\{(b + 3)(b + 2)(b + 1)\} \text{ etc.}$$

This is a factorial-type series (see Johnson and Kotz [7]). The mean exists if  $b > 1$ , the variance if  $b > 2$  and so on. The distribution will be long-tailed if  $b(> 0)$  is small. The probability generating function is

$$\sum \omega^x \Pr(X = x) = -b^2 \int_0^\infty e^{-bt} \ln(\omega e^{-t} + 1 - \omega) dt,$$

with  $s$ th factorial moment

$$\mu'_{[s]} = b^2(s - 1)! \int_0^\infty e^{(b-s)t} (1 - e^{-t})^s dt \quad (b > s)$$

- $\rho = 1$ . Here

$$\Pr(X = x; b, 1) = \frac{b}{x} \int_0^\infty \frac{(1 - e^{-t})^x}{t} e^{-bt} dt = \frac{b}{x} \sum_{r=0}^x (-1)^{r+1} \binom{x}{r} \ln(b + r),$$

this being a Frullani integral (Hardy [6]).

- $0 < \rho < 1$ . To allow for the apparent singularity at  $t = 0$  in (14), we consider

$$\Pr(X = x; b, \rho) = \lim_{\epsilon \rightarrow +0} \frac{1}{x} \int_\epsilon^\infty \frac{(1 - e^{-t})^x}{t} \frac{b^\rho t^{\rho-1} e^{-bt}}{\Gamma(\rho)} dt,$$

for which the integral of a finite sum is that of the sum. A typical term is now

$$\frac{1}{x} (-1)^r \binom{x}{r} \int_\epsilon^\infty \frac{b^\rho}{\Gamma(\rho)} t^{\rho-2} e^{-(b+r)t} dt,$$

and the integral component is

$$\begin{aligned} \frac{1}{\rho - 1} \int_\epsilon^\infty b^\rho e^{-(b+r)t} d(t^{\rho-1}) &= \left| \frac{b^\rho e^{-(b+r)t} t^{\rho-1}}{\rho - 1} \right|_\epsilon^\infty - \frac{b^\rho}{(1 - \rho)} \int_\epsilon^\infty t^{\rho-1} (b + r) e^{-(b+r)t} dt, \\ &= -\frac{b^\rho \epsilon^{\rho-1} e^{-(b+r)\epsilon}}{\rho - 1} - \frac{b^\rho (b + r)}{\Gamma(\rho)(1 - \rho)} \int_\epsilon^\infty e^{-(b+r)t} t^{\rho-1} dt. \end{aligned}$$

The first term in the sum is dominated ( $\epsilon \rightarrow 0$ ) by

$$\epsilon^{\rho-1} \sum_{r=0}^x \binom{x}{r} \{1 - (b + r)\epsilon + \dots\} = -\epsilon^{\rho-1} \cdot \epsilon \sum_{r=0}^x (b + r) \binom{x}{r}, \quad (\epsilon \rightarrow 0)$$

Table 7: Example

$x$	$\Pr(X = x; 1, 1/4)$	$\Pr(X = x; 5, 1/4)$
1	0.909057	0.976876
2	0.056052	0.020036
3	0.015682	0.002360
4	0.006712	0.000494
5	0.003551	0.000284
6	0.002135	0.000050
$\leq 7$	0.006811	0.000021

and will approach 0, since  $\rho > 0$ . Finally for  $0 < p < 1$ .

$$(1 - \rho) \Pr(X = x; b, \rho) = \frac{1}{x} \sum_{r=0}^x (-1)^{r+1} \binom{x}{r} (b+r) \left(\frac{b}{b+r}\right)^\rho.$$

In particular

$$\Pr(X = 1; b, \rho) = -b^\rho \{b^{1-\rho} - (b+1)^{1-\rho}\} / (1 - \rho),$$

$$2 \Pr(X = 2; b, \rho) = -b^\rho \{b^{1-\rho} - 2(b+1)^{1-\rho} + (b+2)^{1-\rho}\} / (1 - \rho),$$

$$3 \Pr(X = 3; b, \rho) = -b^\rho \{b^{1-\rho} - 3(b+1)^{1-\rho} + 3(b+2)^{1-\rho} - (b+3)^{1-\rho}\} / (1 - \rho).$$

values for  $b = 1, \rho = 1/4$  are shown in Table 7. From (15), moments do not exist.

By contrast, for  $b = 5, \rho = 1/4$ , and mean approximately equal to 1.027, derived values are shown in the Table 7. Expression (15) expands as

$$1 + \frac{\rho}{2!b} + \frac{\rho(\rho+1)}{3!b^2} + \frac{\rho(\rho+1)(\rho+2)}{4!b^3} + \dots \sim 1.0273$$

for  $b = 5, \rho = 1/4$ .

### 7.3. Riemann Zeta Function $\zeta(s)$ Density

Take

$$\phi(t) = \left( \frac{t^{s-1}}{e^{-t} - 1} \right) / \{\zeta(s)\Gamma(s)\}. \quad (s > 1)$$

Then

$$\begin{aligned} \Pr(X = x) &= \frac{1}{x} \int_0^\infty \frac{(1 - e^{-t})^x}{t} \frac{t^{s-1} e^{-t}}{(1 - e^{-t}) \Gamma(s) \zeta(s)} dt \\ &= \frac{1}{x \Gamma(s) \zeta(s)} \int_0^\infty (1 - e^{-t})^{x-1} e^{-t} t^{s-2} dt \end{aligned}$$

$$= \frac{1}{x(x-1)\zeta(s)} \left\{ 1 - \binom{x-1}{1} \frac{1}{2^{s-1}} + \binom{x-1}{2} \frac{1}{3^{s-1}} - \dots \right\}.$$

In particular for  $s = 2$ ,  $\Pr(X = x) = 6/(\pi^2 x^2)$  ( $x = 1, 2, \dots$ ).

#### 7.4. Uniform Density

$$\Pr(X = x) = \frac{1}{x} \int_0^A \frac{(1 - e^{-t})^x}{t} \frac{dt}{A}, \quad (A > 0).$$

#### 7.5. Triangular Density

Here

$$\begin{aligned} \phi(t) &= 2t/A^2, & (0 < t < A) \\ &= 0, & \text{otherwise} \end{aligned}$$

and

$$\Pr(X = x) = \frac{1}{x} \int_0^A (1 - e^{-t})^x \frac{2}{A^2} dt = \frac{2}{Ax} \left\{ 1 - \binom{x}{1} \frac{(1 - e^{-A})}{A} + \binom{x}{2} \frac{(1 - e^{-2A})}{2A} - \dots \right\}.$$

#### 7.6. Trigonometric Density

$$\phi(t) = y_0 e^{-bt} \sin^2(at), \quad (t \geq 0, b > 0, a \text{ real})$$

and

$$\int_0^\infty e^{-st} \frac{\sin^2(at)}{t} dt = \frac{1}{4} \ln \left( 1 + \frac{4a^2}{s^2} \right), \quad (s > 0).$$

Thus

$$\begin{aligned} \Pr(X = x) &= \frac{1}{x} \frac{(b^2 + 4a^2)b}{8a^2} \\ &\times \left[ \ln \left( 1 + \frac{4a^2}{b^2} \right) - \binom{x}{1} \ln \left( 1 + \frac{4a^2}{(b+1)^2} \right) + \binom{x}{2} \ln \left( 1 + \frac{4a^2}{(b+2)^2} \right) - \dots \right]. \end{aligned}$$

#### 7.7. Beta Density

$$\begin{aligned} \Pr(X = x) &= \frac{1}{x} \int_0^1 \frac{p^x}{\ln\{1/(1-p)\}} \frac{p^{\alpha-1}(1-p)^{\beta-1} \Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} dp \quad (\alpha, \beta > 0) \\ &= \frac{1}{x} \int_0^\infty \frac{(1 - e^{-t})^{x+\alpha-1}}{t} \{e^{-(\beta-1)t} e^{-t}\} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} dt \\ &= \frac{1}{x} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \int_0^\infty \frac{(1 - e^{-t})^{x+\alpha-1}}{t} e^{-\beta t} dt. \end{aligned}$$

This reduces to a Frullani integral if  $\alpha$  is a positive integer. The  $s$ th factorial moment is

$$\mu'_{[s]} = \frac{(s-1)!\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \int_0^\infty \frac{(1-e^{-t})^{\alpha+s-1}}{t} e^{-(\beta-s)t} dt \quad (\beta > s, s = 1, 2, \dots; \alpha > 0).$$

## 8. Concluding Remarks

Using a truncated form of the beta density, or any other appropriate form of the distribution of the binomial parameter  $p$ , it may be possible to gain some insight into the variation of  $p$  among sets of observations. The search technique for truncation parameters suggested here becomes feasible with the aid of a computer. Suitable examples to test models seem to be rare in the literature, perhaps because poor fits of the binomial were naturally withheld from publication especially in the early part of the century. Nevertheless, with the accumulation of massive data sets in biology and other areas of research, one should expect to find numerous applications in which the probability parameter  $p$  is confined to a subinterval of  $(0, 1)$ .

## 9. References

- [1] M.A. Bender, et al. On the distribution of spontaneous chromosomal aberrations in human peripheral blood lymphocytes in culture. *Mutation Research*, 244:215–220, 1990.
- [2] K.O. Bowman, M.A. Kastenbaum, and L.R. Shenton. The negative hypergeometric distribution and estimation by moments. *Commun. statist. Simula-Computa.*, 1991.
- [3] D.G. Catcheside. Secondary pairing in *Brassica oleracea*. *Cytologia, Fujii Jub. Vol.*, pages 366–378, 1937.
- [4] F.N. David and N.L. Johnson. Truncated Poisson. *Biometrics*, 10:275–285, 1952.
- [5] A. Hald. The mixed binomial distribution and the posterior distribution of  $p$  for a continuous prior distribution. *JRSS Series B*, 30:359–367, 1968.
- [6] G.H. Hardy. On the Frullani integral  $\int_0^\infty (\phi(ax^m) - \phi(bx^m)) \{\ln(x)\}^p / x dx$ . *Quarterly Journal of Pure and Applied Mathematics*, 33:113–144, 1901.
- [7] N.L. Johnson and S. Kotz. *Discrete Distributions*. Houghton Mifflin Company, Boston, 1969.
- [8] K. Pearson. System of frequency curves. *Phil. Trans. A*, 186:343–414, 1885.
- [9] C.R. Rao and I.M. Chakravarti. Some small sample tests of significance for a Poisson distribution. *Biometrics*, 12:264–282, 1956.
- [10] L.R. Shenton. Maximum likelihood and the efficiency of the method of moments. *Biometrika*, 37:111–116, 1950.

- [11] J.G. Skellam. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. *J. Roy Statist. Soc., Series B*, 10:257-261, 1948.

APPENDIX

**A. Graphical Analysis**

The following is a graphical analysis of a set of generalized mixtures of binomial distributions.

In addition to the graphical analysis, the skewness and kurtosis of the distributions were used in order to compare them with the Pearson-type frequency curves. Models 7 and 14 were not compared in this fashion, because of the difficulty in obtaining data.

In what follows,  $n$  is taken to be 20. "sd" denotes the standard deviation, "skew" denotes the skewness, and "kurt" denotes the kurtosis.

**A.1. Model 1**

The model 1 distribution is simply the binomial distribution. A short description will be helpful in comparing it with the other mixtures.

For the model 1 distribution,  $f(x; p = p_0) = f(n - x; p = 1 - p_0)$ , for  $0 < p_0 < 1$ . The model 1 distributions are in the form of bell-shaped curves. When  $p = 0.5$ , the distribution has skewness=0, the kurtosis is at a minimum and the standard deviation is at a maximum. Thus as  $p$  deviates from 0.5, the kurtosis increases and the skewness decreases. When  $p$  is very close to 0 or 1, the distribution becomes J-shaped. When  $p < 0.5$ , the distribution is skewed to the left (skewness > 0); when  $p > 0.5$ , the distribution is skewed to the right (skewness < 0).

For  $0.03 \leq p \leq 0.97$ , the model 1 distribution could be classified as a Pearson type I distribution. For  $p \leq 0.29$  or  $p \geq 0.98$ , it could be classified as a Pearson type I(J) distribution.

$p$ :	0.1000	0.3000	0.5000	0.7000	0.9000
sd:	1.3416	2.0494	2.2361	2.0494	1.3416
skew:	0.5963	0.1952	0.0000	-0.1952	-0.5963
kurt:	3.2556	2.9381	2.9000	2.9381	3.2556

**A.2. Model 2**

The model 2 distributions include J-shaped and bell-shaped curves. For small  $b$ , increasing  $p$  changes the distribution from a U-shaped curve to a J-shaped curve with skewness < 0 to a bell-shaped curve with skewness > 0. For larger  $b$ , increasing  $p$  changes the distribution from a J-shaped curve with skewness < 0 to a bell-shaped curve with skewness < 0 to a bell-shaped curve with skewness > 0.

In general, for  $p < 1$ , the model 2 distribution could be classified as Pearson type I(U) when  $b < 0.5$  and Pearson type I(J) when  $b > 1.0$ . For  $p > 1$ , it could be classified as Pearson type I(J) for  $b < 1$  and Pearson type I for  $b > 2$  or  $b > 3$ .

Note that  $f(x; b = 1, p = 1)$  is a discrete uniform distribution.

$p$		$b = 0.50$	$b = 1.00$	$b = 2.0$	$b = 4.0$	$b = 8.0$
$b/2$	sd:	6.1950	5.7910	5.0553	4.2248	3.4980
	skew:	-1.1629	-0.8181	-0.5671	-0.3927	-0.2770
	kurt:	3.0240	2.5053	2.3922	2.4870	2.6379
$b$	sd:	6.9394	6.0553	4.9876	4.0181	3.2845
	skew:	-0.3025	0.0000	0.1820	0.2488	0.2401
	kurt:	1.6560	1.7945	2.1348	2.4692	2.6916
$3b/2$	sd:	6.7409	5.4956	4.2789	3.3340	2.7016
	skew:	0.2208	0.5255	0.6712	0.6614	0.5672
	kurt:	1.6616	2.2163	2.7866	3.0881	3.1342
$2b$	sd:	6.1824	4.7140	3.4775	2.6312	2.1224
	skew:	0.6393	0.9625	1.0776	0.9948	0.8297
	kurt:	2.1391	3.1221	3.8578	3.9680	3.7118

### A.3. Model 3

The model 3 distribution is in general J-shaped. Fixing  $b$  and increasing  $a$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a relatively uniform curve to a J-shaped curve with skewness  $< 0$ . Fixing  $a$  and increasing  $b$  results in similar changes.

For  $b < 0.4$  and  $a < 1$ , the model 3 distribution could be classified as Pearson type I(U). For larger values of  $a$  and  $b$ , it could be classified as Pearson type I(J).

$b$		$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/5$	sd:	6.2675	6.3548	5.6090	4.3073
	skew:	0.6926	0.0524	-0.5424	-1.1064
	kurt:	2.1743	1.7219	2.2354	3.6662
$2a/5$	sd:	6.3021	5.8948	4.7321	3.3308
	skew:	0.3960	-0.2194	-0.7810	-1.2739
	kurt:	1.8724	1.8925	2.8178	4.4619
$3a/5$	sd:	6.2338	5.5389	4.2219	2.8597
	skew:	0.2209	-0.3739	-0.9030	-1.3443
	kurt:	1.7920	2.0766	3.1838	4.8197
$4a/5$	sd:	6.1448	5.2673	3.8879	2.5799
	skew:	0.0966	-0.4823	-0.9876	-1.3976
	kurt:	1.7792	2.2422	3.4576	5.0770

### A.4. Model 4

The model 4 distribution is generally bell-shaped. For very small  $k$ , the distribution is J-shaped with skewness  $< 0$ . Increasing  $k$  causes the distribution to become bell-shaped,

decreasing the skewness and the kurtosis. When the skewness becomes positive, the kurtosis begins to increase. For large  $k$ , the distribution is again bell-shaped.

For  $0.03 \leq k \leq 0.51$ , the model 4 distribution could be classified as a Pearson type I distribution. For  $k \leq 0.02$  or  $k \geq 0.52$ , it could be classified as Pearson type I(J).

$k$ :	0.1000	0.4000	1.0000	2.0000	4.0000
sd:	1.3123	2.1023	2.1566	1.5298	0.5997
skew:	-0.6170	-0.1620	0.1225	0.4767	1.6065
kurt:	3.2807	2.9263	2.9150	3.1273	5.4808

#### A.5. Model 5

The model 5 distribution is either J-shaped or U-shaped. Increasing  $k$  changes the curve from J-shaped with skewness  $< 0$  to U-shaped to J-shaped with skewness  $> 0$ .

For  $k \leq 1.55$ , the model 5 distribution could be classified as Pearson type I(J). For  $k \geq 1.56$ , it could be classified as Pearson type I(U).

$k$ :	1.0000	2.0000	4.0000	8.0000
sd:	5.5153	6.5717	6.9035	6.5275
skew:	-0.7655	-0.2162	0.3390	0.9235a
kurt:	2.5280	1.7058	1.6734	2.4338

#### A.6. Model 6

The model 6 distribution is either J-shaped or U-shaped. Fixing  $a$  and increasing  $b$  changes the distribution from a U-shaped curve with skewness  $> 0$  to a J-shaped curve with skewness  $< 0$ . Fixing  $b$  and increasing  $a$  causes a similar change.

For  $a < 0.1$  or  $b < 1$ , the model 6 distribution could be classified as Pearson type I(U). For larger values of  $a$  and  $b$ , it could be classified as Pearson type I(J).

$b$		$a = 0.25$	$a = 0.5$	$a = 1.0$	$a = 2.0$
$3a/2$	sd:	7.3742	6.6608	5.3703	3.8488
	skew:	0.0410	-0.4574	-0.9812	-1.5095
	kurt:	1.4726	1.8338	2.9257	4.8784
$3a$	sd:	7.0902	6.0679	4.6375	3.1870
	skew:	-0.2446	-0.7621	-1.3123	-1.8648
	kurt:	1.5946	2.3543	3.9950	6.6561
$9a/2$	sd:	6.8379	5.6778	4.2301	2.8578
	skew:	-0.4325	-0.9680	-1.5447	-2.1251
	kurt:	1.7779	2.8347	4.9160	8.1789

#### A.7. Model 7

The model 7 distribution is U-shaped. Increasing  $a$  changes the skewness from positive to negative. For small and large values of  $a$ , the kurtosis is larger and the standard deviation is smaller.

The model 7 distribution has not been classified in regard to the Pearson distributions.

<i>a</i> :	0.1000	0.5000	1.0000	2.0000	5.0000
sd:	6.1092	7.1723	6.7103	5.6720	3.8671
skew:	1.3962	0.0879	-0.4377	-0.9935	-1.8017
kurt:	3.5853	1.5177	1.8184	2.9097	6.1301

### A.8. Model 8

The model 8 distribution is generally J-shaped. Fixing *b* and increasing *a* changes the distribution from a J-shaped curve with skewness > 0 to a relatively uniform curve to a J-shaped curve with skewness < 0. Fixing *a* and increasing *b* is accompanied by similar changes.

For large values of *a*, the model 8 distribution could be classified as Pearson type I(J). For smaller values of *a*, it could be classified as Pearson type I(U) for small values of *b* and Pearson type I(J) for larger values of *b*.

<i>b</i>		<i>a</i> = 0.5	<i>a</i> = 1.0	<i>a</i> = 2.0	<i>a</i> = 4.0	<i>a</i> = 8.0
3 <i>a</i> /2	sd:	6.2552	5.8417	4.6582	3.2532	2.0967
	skew:	0.4272	-0.1884	-0.7376	-1.2057	-1.5739
	kurt:	1.9092	1.8885	2.7486	4.2386	5.9437
2 <i>a</i>	sd:	6.2640	5.6448	4.3647	2.9873	1.9082
	skew:	0.2633	-0.3379	-0.8791	-1.3390	-1.6969
	kurt:	1.8030	2.0249	3.1028	4.7806	6.6019
3 <i>a</i>	sd:	6.1926	5.3074	3.9468	2.6387	1.6697
	skew:	0.0155	-0.5728	-1.1122	-1.5696	-1.9184
	kurt:	1.7633	2.3535	3.8066	5.8436	7.9062

### A.9. Model 9

The model 9 distribution is, in general, J-shaped. Fixing *B* and increasing *A* changes the distribution from a J-shaped curve with skewness < 0 to a relatively uniform curve to a J-shaped curve with skewness > 0.

For *A* > 1, the model 9 distribution can be classified as Pearson type I(U). For *A* < 1, it can be classified as Pearson type I(J) for small values of *B* and Pearson type I(U) for larger values of *B*.

$B$		$A = 0.3$	$A = 0.6$	$A = 1.0$	$A = 2.0$
5A/4	sd:	4.2728	5.5775	6.1474	6.1179
	skew:	-0.8601	-0.3331	0.1002	0.7537
	kurt:	3.0730	2.0336	1.7791	2.3057
3A/2	sd:	4.4996	5.7391	6.2110	6.0481
	skew:	-0.7946	-0.2535	0.1877	0.8539
	kurt:	2.8902	1.9447	1.7855	2.4763
2A	sd:	4.8900	5.9821	6.2774	5.9037
	skew:	-0.6806	-0.1172	0.3367	1.0250
	kurt:	2.6014	1.8311	1.8394	2.8211

### A.10. Model 10

The model 10 distribution is either J-shaped or bell-shaped. Fixing  $b$  and increasing  $a$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a flatter, more centered, bell-shaped curve to a bell-shaped curve with skewness  $< 0$  and higher kurtosis. Fixing  $a$  and increasing  $b$  causes similar changes.

For  $a < 1.3$ , the model 10 distribution could be classified as Pearson Type I(J) for small or large values of  $b$  and Pearson Type I for intermediate values of  $b$ . For  $a > 1.4$ , it could be classified as Pearson Type I for small values of  $b$  and Pearson Type I(J) for large values of  $b$ .

$b$		$a = 0.5$	$a = 1.0$	$a = 2.0$	$a = 5.0$
3a/2	sd:	3.9785	4.9485	4.8753	3.5205
	skew:	1.6844	0.6834	-0.0245	-0.6946
	kurt:	5.4265	2.5750	2.1392	3.0663
5a/2	sd:	4.5993	5.1678	4.6626	3.1303
	skew:	1.1523	0.3136	-0.3227	-0.9472
	kurt:	3.5594	2.1403	2.3302	3.7333
5a	sd:	5.3055	5.1975	4.2377	2.6376
	skew:	0.4842	-0.2136	-0.8004	-1.3998
	kurt:	2.2141	2.1141	3.1157	5.3675

### A.11. Model 12

The model 12 distribution is bell-shaped or J-shaped. Fixing  $B$  and increasing  $A$  changes the distribution from a bell-shaped curve with skewness  $< 0$  to a J-shaped curve with skewness  $> 0$ . Fixing  $A$  and increasing  $B$  causes similar changes.

For  $A \leq 0.73$ , the model 12 distribution could be classified as Pearson type I for small values of  $B$  and Pearson type I(J) for larger values of  $B$ . For  $A \geq 0.74$ , it could be classified as Pearson type I(J).

$B$		$A = 0.3$	$A = 0.6$	$A = 1.0$	$A = 2.0$
$5A/4$	sd:	4.6598	5.0213	4.5697	3.3081
	skew:	-0.1692	0.4846	1.1192	2.3596
	kurt:	2.2305	2.3225	3.5089	8.7901
$3A/2$	sd:	4.7996	4.9985	4.4347	3.1219
	skew:	-0.0917	0.5926	1.2602	2.5831
	kurt:	2.1701	2.4435	3.9063	10.1458
$2A$	sd:	5.0114	4.9194	4.1899	2.8284
	skew:	0.0457	0.7829	1.5094	2.9814
	kurt:	2.0996	2.7252	4.7249	12.8617

### A.12. Model 13

The model 13 distribution is bell-shaped or J-shaped. Fixing  $B$  and increasing  $A$  changes the distribution from a bell-shaped curve to a J-shaped curve with skewness  $> 0$ . Fixing  $A$  and increasing  $B$  causes a similar change.

For  $A \leq 0.6$ , the model 13 distribution could be classified as Pearson type I for small values of  $B$  and Pearson type I(J) for larger values of  $B$ . For  $A \geq 0.625$ , it could be classified as Pearson type I(J).

$B$		$A = 0.2$	$A = 0.4$	$A = 0.6$	$A = 0.8$
$3A/2$	sd:	4.2827	4.3238	3.8053	3.2608
	skew:	-0.0390	0.6550	1.2192	1.7366
	kurt:	2.3256	2.7438	4.0718	6.0489
$2A$	sd:	4.4825	4.2209	3.5606	2.9678
	skew:	0.0920	0.8586	1.4882	2.0747
	kurt:	2.2667	3.0903	4.9704	7.6843
$3A$	sd:	4.6912	3.9714	3.1625	2.5454
	skew:	0.3294	1.2100	1.9508	2.6570
	kurt:	2.2787	3.9303	6.9322	11.1696

### A.13. Model 14

The model 14 distribution is U-shaped or J-shaped. For  $R = 1$ , it is U-shaped and for  $R \geq 2$  it is J-shaped. Fixing  $R$  and increasing  $A$  increases the skewness and the kurtosis.

The model 14 distribution has not been classified in regard to the Pearson distributions.

$R$		$A = 1.0$	$A = 2.0$	$A = 3.0$
1	sd:	6.7066	7.1663	7.1396
	skew:	-0.4334	0.0911	0.3967
	kurt:	1.8170	1.5186	1.6494
2	sd:	6.7248	6.2878	5.7890
	skew:	0.3106	0.9307	1.3355
	kurt:	1.6978	2.5102	3.5200
3	sd:	6.3103	5.4690	4.8417
	skew:	0.7811	1.5073	2.0080
	kurt:	2.2748	4.1088	6.0751

#### A.14. Model 15

For many values of  $a$  and  $u$ , the model 15 distribution assumes negative values for extreme values of the variate (this phenomenon is noted and in sequel, we refer to it a “bad” case). For  $a = a_0$ , there is some value  $u_0$  (depending on  $a$ ) for which  $f(x; a_0, u)$  is positive (a “good” case) for  $u < u_0$  but “bad” for  $u > u_0$ . For  $a < 20$ , the value  $u_0(a)$  increases at a not quite linear rate. For  $20 < a < 30$ , the values fluctuate rather intricately, finally dropping down to  $u_0 = 0$  for  $a = 35$ , meaning that the model 15 distribution yields negative values for  $a > 35$ .

The model 15 distribution is J-shaped or bell-shaped. Fixing  $u$  and increasing  $a$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a bell-shaped curve with skewness  $< 0$ .

The model 15 distribution could be classified as Pearson type I(J) for  $a < 0.75$  and Pearson type I for  $a > 1.25$ . For  $0.75 < a < 1.25$ , it could be classified as Pearson type I(J) for small values of  $u$  and Pearson type I for larger values of  $u$ .

$u$		$a = 0.5$	$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/5$	sd:	3.8795	4.7567	4.8653	4.1105	2.9906
	skew:	1.6909	0.8047	0.1286	-0.4034	-0.8120
	kurt:	5.5454	2.8344	2.1504	2.5161	3.3935a
$2a/5$	sd:	4.0854	4.7396	4.7336	3.9989	2.9357
	skew:	1.4215	0.6904	0.0900	-0.4090	-0.8076
	kurt:	4.5719	2.6676	2.1675	2.5339	3.3842
$3a/5$	sd:	4.1969	4.6794	4.5956	3.8899	2.8824
	skew:	1.2389	0.6111	0.0660	-0.4093	-0.8017
	kurt:	4.0136	2.5683	2.1778	2.5385	3.3688

#### A.15. Model 16

For many values of  $a$  and  $u$ , the model 16 distribution assumes negative values (and is thus “bad”). Just as for the model 15 distribution, for  $a = a_0$ , there is some value

$u_0$  (depending on  $a$ ) for which  $f(x; a_0, u)$  is “good” for  $u < u_0$  but “bad” for  $u > u_0$ . Similarly to the model 15 distribution, the value of  $u_0(a)$  increases at a not quite linear rate for  $a < 15$ , and then fluctuates intricately for  $15 < a < 25$ , at which point it quickly drops to 0. Thus the model 16 distribution yields negative values for  $a > 25$ .

The model 16 distribution is J-shaped or bell-shaped. Fixing  $u$  and increasing  $a$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a bell-shaped curve to a J-shaped curve with skewness  $> 0$ .

For  $a < 1$ , the model 16 distribution could be classified as Pearson I(U) for very small values of  $u$  and type I(J) for larger values of  $u$ . For  $1 < a < 3$ , it could be classified as Pearson type I(J) for small values of  $u$  and type I for larger values of  $u$ . For  $a > 3$ , it could be classified as Pearson type I(J).

$u$		$a = 0.5$	$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/20$	sd:	6.1577	5.9238	4.9357	3.5866	2.3680
	skew:	0.5276	-0.0477	-0.5812	-1.0544	-1.4417
	kurt:	2.0257	1.8260	2.4310	3.6923	5.2896
$a/10$	sd:	6.0873	5.7762	4.8127	3.5213	2.3425
	skew:	0.4334	-0.0866	-0.5898	-1.0497	-1.4349
	kurt:	1.9558	1.8574	2.4571	3.6815	5.2576
$3a/20$	sd:	5.9726	5.6125	4.6859	3.4556	2.3171
	skew:	0.3581	-0.1141	-0.5918	-1.0420	-1.4271
	kurt:	1.9156	1.8810	2.4652	3.6565	5.2191

#### A.16. Model 17

The model 17 distribution is either J-shaped or bell-shaped, and may have more than one “peak”. For small  $a$ , increasing  $b$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a bell-shaped curve with decreasing skewness. For larger  $a$ , the curve changes from a curve that has a J-shape on the left and a bell-shape on the right (a sideways S-curve) to a simple bell-shaped curve. Similar changes occur when  $b$  is fixed and  $a$  is increased.

For  $a < 3$  (approximately), the model 17 distribution could be classified as Pearson type I(J) for small values of  $b$  and type I for larger values of  $b$ . For  $a > 3$ , it could be classified as Pearson type I(U) for small values of  $b$ , type I(J) for intermediate values of  $b$ , and I for larger values of  $b$ .

Note that  $f(x; a, b) = f(x; -a, b)$ .

$b$		$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/10$	sd:	3.2092	4.6997	5.8236	6.1156
	skew:	2.5979	1.5092	0.7924	0.1673
	kurt:	9.8685	4.2197	2.3187	1.7550a
$a/2$	sd:	4.6644	5.1610	4.6341	3.4854
	skew:	0.8535	-0.0068	-0.6317	-1.0998
	kurt:	2.9835	2.0971	2.6624	3.9494
$a$	sd:	4.7792	4.3977	3.3883	2.3340
	skew:	0.4622	-0.2537	-0.8165	-1.1969
	kurt:	2.4265	2.4996	3.6198	4.9412
$2a$	sd:	4.6931	3.7346	2.6038	1.7212
	skew:	0.0653	-0.4426	-0.8172	-1.1068
	kurt:	2.1828	2.6401	3.5028	4.4115

### A.17. Model 18

The model 18 distribution is either J-shaped or bell-shaped, but it may have several “peaks”, forming the sideways S-curve as in model 17. When  $a$  is small, increasing  $b$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a bell-shaped curve with skewness  $< 0$ . For larger  $a$ , increasing  $b$  changes the distribution from a sideways S-curve (J-shaped on the left and bell-shaped on the right) to a simple bell-shaped curve and finally back to a J-shaped curve with skewness  $< 0$ . Similar changes are observed when  $b$  is fixed and  $a$  increases.

For  $a < 2$ , the model 18 distribution could be classified as Pearson type I(J) for small values of  $b$  and type I for larger values of  $b$ . For  $a > 2$ , it could be classified as Pearson type I(U) for small values of  $b$ , I(J) for intermediate values of  $b$ , and I for larger values of  $b$ . For  $a > 5$ , there is a range of values of  $b$  for which the model 18 distribution could be classified as Pearson type III.

As for the model 17 distribution,  $f(x; a, b) = f(x; -a, b)$ .

$b$		$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/10$	sd:	2.8406	4.5006	5.7369	6.0805
	skew:	2.4291	1.4155	0.7577	0.1542
	kurt:	9.0316	3.8612	2.2276	1.7385
$a/2$	sd:	3.9363	4.7311	4.4341	3.4066
	skew:	0.8320	-0.0903	-0.7014	-1.1384
	kurt:	3.2064	2.2089	2.8114	4.0962
$a$	sd:	3.9882	3.9546	3.1947	2.2742
	skew:	0.5917	-0.2025	-0.7869	-1.1369
	kurt:	2.8848	2.7403	3.8311	4.9959
$2a$	sd:	4.1086	3.5270	2.5838	1.7734
	skew:	0.3320	-0.2064	-0.5828	-0.8658
	kurt:	2.5035	2.5776	3.1687	3.8290

#### A.18. Model 19

The model 19 distribution is generally J-shaped, but it may have several “peaks”, forming sideways-S curves or W-shaped curves. Fixing  $a$  and increasing  $b$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a W-shaped curve (J-shaped at both ends with a bell-shaped in the center) to a backwards sideways-S curve (J-shaped on the right and bell-shaped on the left) and finally to a J-shaped curve with skewness  $< 0$ . Similar changes occur when  $b$  is fixed and  $a$  increases.

The classification of the model 19 distribution with regard to the Pearson frequency curves is quite complicated. For small values of  $b$ , the distribution could be classified as Pearson type I(U). For larger values of  $b$ , it could be classified as type I(J), I or III. For large values of  $b$ , the distribution is type I(J).

As for models 17 and 18,  $f(x; a, b) = f(x; -a, b)$ .

$b$		$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/10$	sd:	4.9755	5.7537	6.3096	6.3166
	skew:	2.2988	1.6784	0.9289	0.2320
	kurt:	6.9581	4.5010	2.6102	1.8253
$a/5$	sd:	7.1262	6.7517	5.4217	3.7985
	skew:	0.1141	-0.2719	-0.6751	-1.1125
	kurt:	1.4685	1.5430	2.2349	3.5731
$a$	sd:	5.9061	4.9260	3.5369	2.2816
	skew:	-0.7175	-1.2778	-1.6837	-2.0170
	kurt:	2.4774	3.7455	5.3564	7.3280
$2a$	sd:	4.0938	2.8546	1.8659	1.1871
	skew:	-0.9441	-1.6183	-2.1196	-2.3953
	kurt:	3.6488	6.7665	9.9997	11.9942

#### A.19. Model 20

The model 20 distribution is very similar to the model 18 distribution, forming J-shaped curves, sideways-S curves, and bell-shaped curves, and undergoing similar changes. As with models 17, 18, and 19,  $f(x; a, b) = f(x; -a, b)$ .

The classification of model 20 with regard to the Pearson type frequency curves is nearly identical to that of the model 18 distribution.

$b$		$a = 1$	$a = 2$	$a = 4$	$a = 8$
$a/10$	sd:	4.6975	5.6233	5.6753	4.8714
	skew:	0.8362	-0.0340	-0.6567	-1.2128
	kurt:	2.9182	1.8871	2.2967	3.6729
$a/5$	sd:	4.6422	4.3944	3.5068	2.4749
	skew:	0.4706	-0.3396	-1.0111	-1.4703
	kurt:	2.5289	2.7006	4.2488	6.2015
$a$	sd:	4.6090	3.8950	2.8326	1.9136
	skew:	0.3006	-0.3054	-0.7842	-1.1211
	kurt:	2.3369	2.6325	3.7361	4.8805
$2a$	sd:	4.5127	3.4821	2.3929	1.5810
	skew:	-0.0053	-0.4737	-0.8173	-1.0984
	kurt:	2.2203	2.7025	3.4827	4.3223

#### A.20. Model 22

The model 22 distribution is either J-shaped or U-shaped. Fixing  $a$  and increasing  $b$  changes the distribution from a J-shaped curve with skewness  $> 0$  to a U-shaped curve

to a J-shaped curve with skewness  $< 0$ . Fixing  $b$  and increasing  $a$  is accompanied by similar changes.

For  $a < 2$ , the model 22 distribution could be classified as Pearson type I(U) for small values of  $b$  and type I(J) for large values of  $b$ . For  $2 < a < 7$ , it could be classified as Pearson type I(J) for small and large values of  $b$  and type I(U) for intermediate values of  $b$ . For  $a > 7$ , it could be classified as Pearson type I(J) for all values of  $b$ .

$b$		$a = 0.5$	$a = 1$	$a = 2$	$a = 4$
$a/2$	sd:	6.3199	6.8660	6.9023	6.1453
	skew:	1.4557	0.8816	0.3191	-0.2277
	kurt:	3.6704	2.3017	1.6790	1.8032
$a$	sd:	7.3303	6.8491	5.6579	4.1023
	skew:	0.3403	-0.1646	-0.6755	-1.1667
	kurt:	1.5829	1.6108	2.3088	3.7024
$2a$	sd:	6.5140	5.1404	3.5957	2.3032
	skew:	-0.4623	-0.9759	-1.4709	-1.9026
	kurt:	1.8669	2.9632	4.7981	7.1018

#### A.21. Model 24

The model 24 distribution is either J-shaped or bell-shaped. Fixing  $a$  and increasing  $k$  changes the distribution from a J-shaped curve with skewness  $< 0$  to a bell-shaped curve with increasing skewness. Fixing  $b$  and increasing  $a$  is accompanied by similar changes.

For  $a < 0.5$ , the model 24 distribution could be classified as Pearson type I(U) for small values of  $k$  and type I(J) for larger values of  $k$ . For  $a > 0.5$ , it could be classified as Pearson type I(J) for small and large values of  $k$  and type I for intermediate values of  $k$ .

$k$		$a = 1$	$a = 2$	$a = 4$	$a = 8$
$a/4$	sd:	3.0508	3.0808	2.9579	2.7770
	skew:	-2.5650	-1.6889	-0.9948	-0.4989
	kurt:	10.6347	6.3018	4.0322	3.1037
$a$	sd:	4.6123	4.0205	3.1944	2.3809
	skew:	-0.7615	-0.2466	0.1356	0.3841
	kurt:	2.8878	2.4704	2.6325	2.9658
$3a/2$	sd:	4.6937	3.7214	2.6324	1.7194
	skew:	-0.3234	0.1427	0.4754	0.7050
	kurt:	2.2917	2.4432	2.9622	3.4601

### A.22. Model 25

The model 25 distribution is either J-shaped or bell-shaped. Fixing  $a$  and increasing  $k$  changes the distribution from a bell-shaped curve to a J-shaped curve with skewness  $> 0$ . Fixing  $k$  and increasing  $a$  is accompanied by the opposite change: from a J-shaped curve with skewness  $> 0$  to a bell-shaped curve with decreasing skewness.

For small values of  $a$  or large values of  $k$ , the model 25 distribution could be classified as Pearson type I(J). For large values of  $a$ , it could be classified as Pearson type I.

$k$		$a = 1$	$a = 2$	$a = 4$	$a = 8$
$a/2$	sd:	5.2518	5.3782	4.5127	3.2398
	skew:	0.8814	0.1464	-0.4286	-0.8718
	kurt:	2.7856	2.0058	2.4149	3.4187
$a$	sd:	4.8416	5.3328	4.7589	3.5585
	skew:	1.2362	0.4071	-0.2174	-0.6923
	kurt:	3.6963	2.1679	2.2050	2.9958
$2a$	sd:	3.8874	4.8758	4.8753	3.9442
	skew:	1.9691	0.8783	0.1317	-0.4065
	kurt:	6.6518	2.9536	2.1646	2.5539

### A.23. Model 26

For many values of  $a$  and  $k$ , the model 26 distribution assumes negative values (and is thus “bad”). Similarly to model 15 and 16, for a given  $k = k_0$ , there is some value  $a_0$  (depending on  $k$ ) for which  $f(x, a, k_0)$  is “good” for  $a > a_0$  but “bad” for  $a < a_0$ . It appears that this value is  $a(k) = 6.5$  (approximately) for all values of  $k$ . [A more exact value is  $a(k) = 6.497596$ ]

The model 26 distribution is a sideways, backwards S-curve (J-shaped on the right and bell-shaped on the left. It is the point at  $x = 19$  that becomes negative, while  $x = 20$  is very large (giving the curve the J-shape) and for  $x < 19$  the curve is bell-shaped. Fixing  $a$  and increasing  $b$  results in a decrease in the value at  $x = 20$  and changes the bell-shaped part of the curve from skewness  $< 0$  to skewness  $> 0$ .

For small values of  $k$ , the model 26 distribution could be classified as Pearson type I(J). For larger values of  $k$ , it could be classified as Pearson type I.

$a$		$k = 0.5$	$k = 1.0$	$k = 2.0$	$k = 4.0$	$k = 8.0$
6.5	sd:	3.0450	4.0008	4.8858	5.1607	4.1030
	skew:	-2.0870	-1.2271	-0.5058	0.1849	0.9914
	kurt:	6.9579	3.6768	2.2840	2.1720	3.7200
8	sd:	2.7583	3.6414	4.4890	4.8317	3.9880
	skew:	-2.0819	-1.2440	-0.5495	0.1002	0.8266
	kurt:	7.1020	3.8508	2.4405	2.2266	3.3824
10	sd:	2.4709	3.2793	4.0857	4.4915	3.866
	skew:	-2.0776	-1.2632	-0.5970	0.0116	0.663
	kurt:	7.2544	4.0377	2.6103	2.2970	3.1086
15	sd:	2.0117	2.6959	3.4245	3.9128	3.6352
	skew:	-2.0706	-1.2972	-0.6787	-0.1360	0.4085
	kurt:	7.5058	4.3592	2.9110	2.4522	2.8128

#### A.24. Model 27

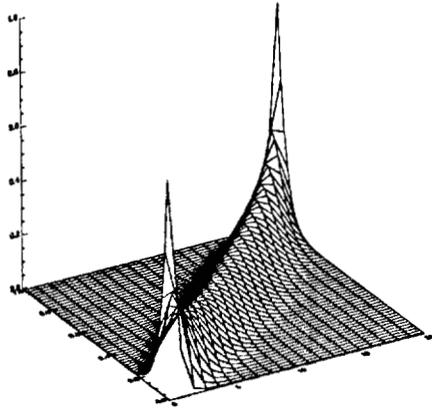
The model 27 distribution takes the form of a bell-shaped or J-shaped curve. Fixing  $k$  and increasing  $a$  changes the distribution from a sideways-S curve (J-shaped on the left and bell-shaped on the right) to a simple bell-shaped curve to a J-shaped curve with skewness  $> 0$ . Fixing  $a$  and increasing  $k$  is accompanied by similar changes.

For  $k < 1$ , the model 27 distribution could be classified as Pearson type I(U) for small values of  $a$  and type I(J) for larger values of  $a$ . For  $k > 1$ , it could be classified as Pearson type I(J) for small values of  $a$  and type I for larger values of  $a$ .

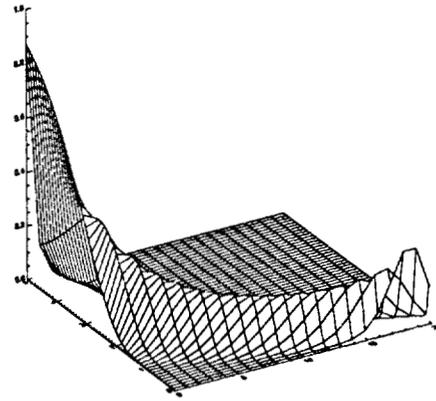
$k$		$a = 1.0$	$a = 2.0$	$a = 4.0$	$a = 8.0$
$a/2$	sd:	5.6099	4.4989	3.6004	2.9637
	skew:	-0.3550	-0.4107	-0.2433	-0.0118
	kurt:	2.0268	2.4298	2.6093	2.7212
$a$	sd:	5.2450	4.1383	3.1522	2.3362
	skew:	-0.0252	0.0485	0.2388	0.4194
	kurt:	1.9736	2.3117	2.6674	3.0030
$2a$	sd:	4.3326	3.0671	1.9734	1.1711
	skew:	0.4651	0.6326	0.8214	1.0560
	kurt:	2.3883	2.9859	3.6304	4.2393
$5a$	sd:	2.0613	0.9926	0.4154	0.1410
	skew:	1.5266	1.8655	2.8110	7.2022
	kurt:	5.5795	7.4999	11.9636	55.6491

### B. Three Dimensional Graphs

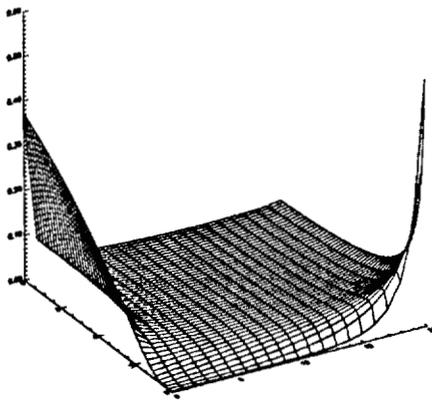
Three axes refer to probability, a parameter and the variable  $x$ ;  $n$  is taken to be 20.



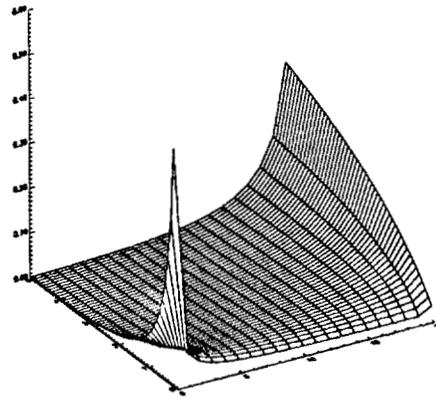
Model 1



Model 4



Model 5



Model 7

Figure 2: Models 1, 4, 5, and 7

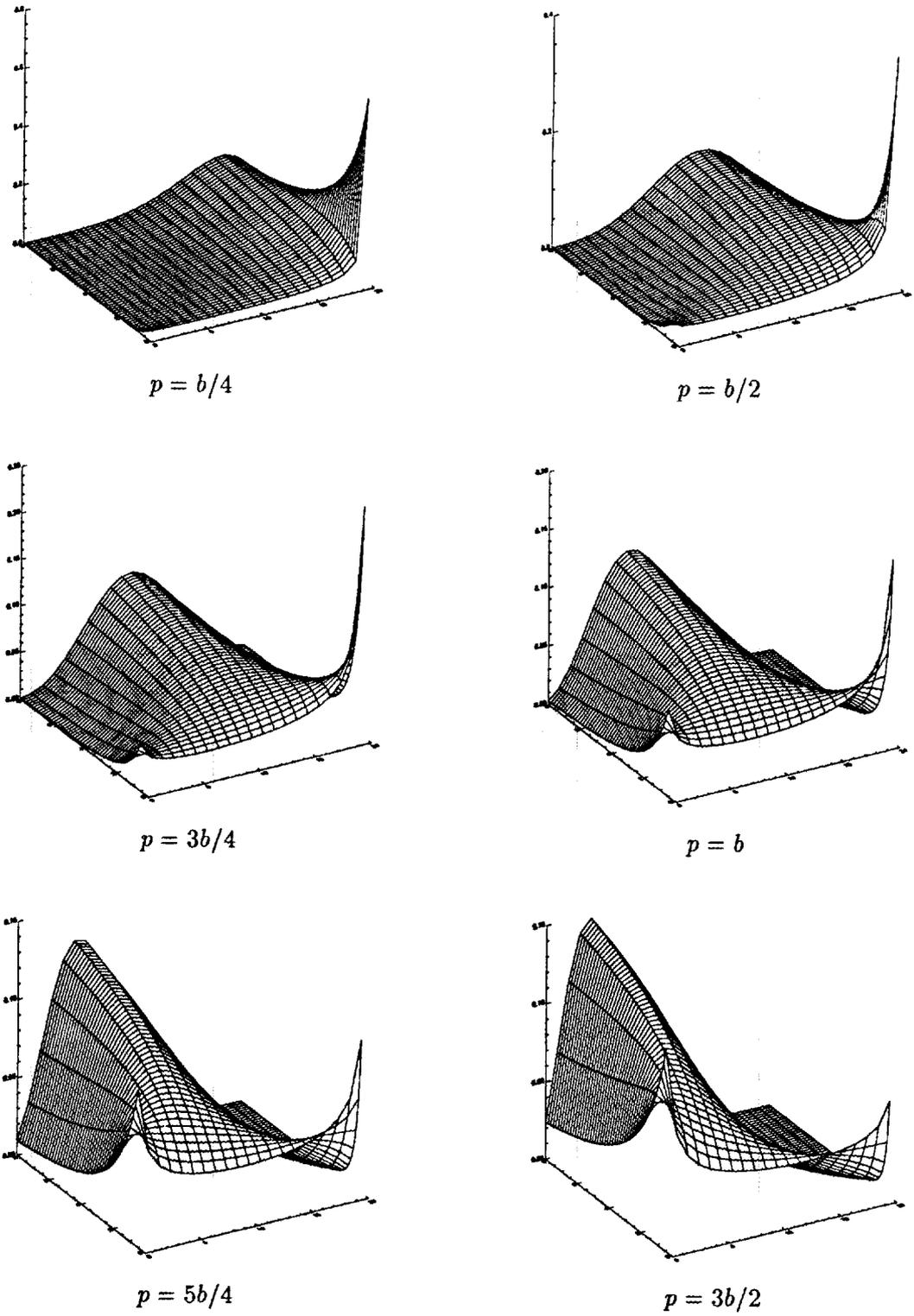
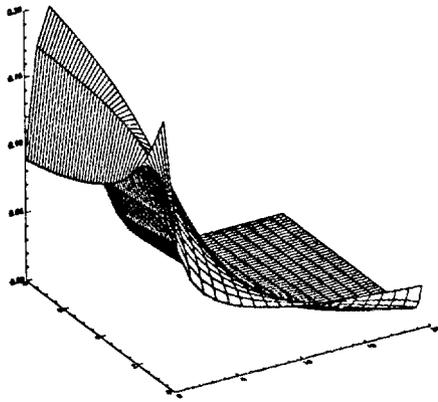
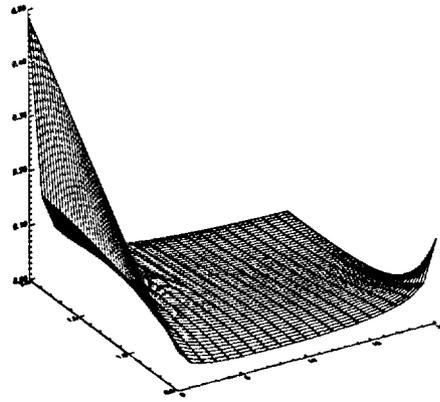


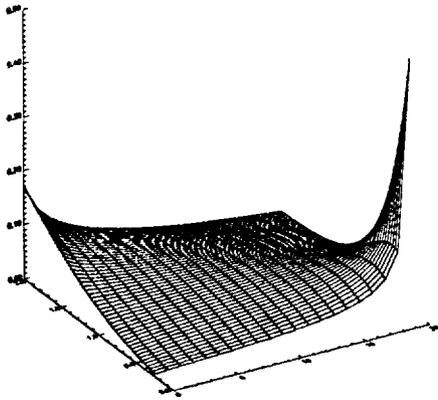
Figure 3: Model 2



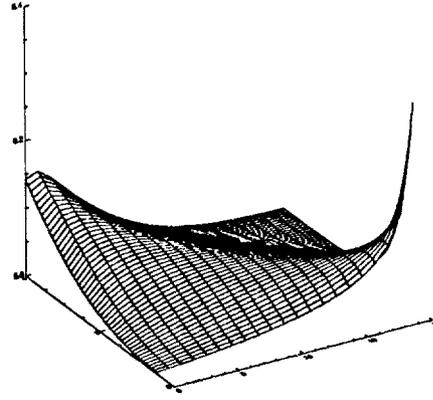
$p = 2$



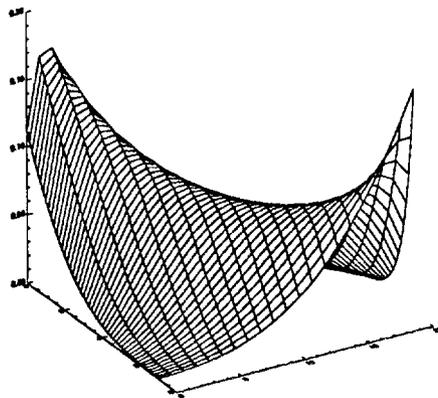
$b = 0.5$



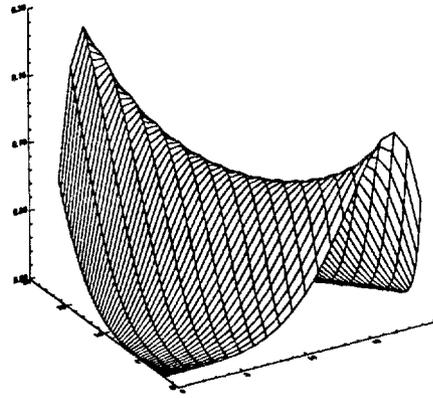
$b = 1$



$b = 2$



$b = 4$



$b = 8$

Figure 4: Model 2

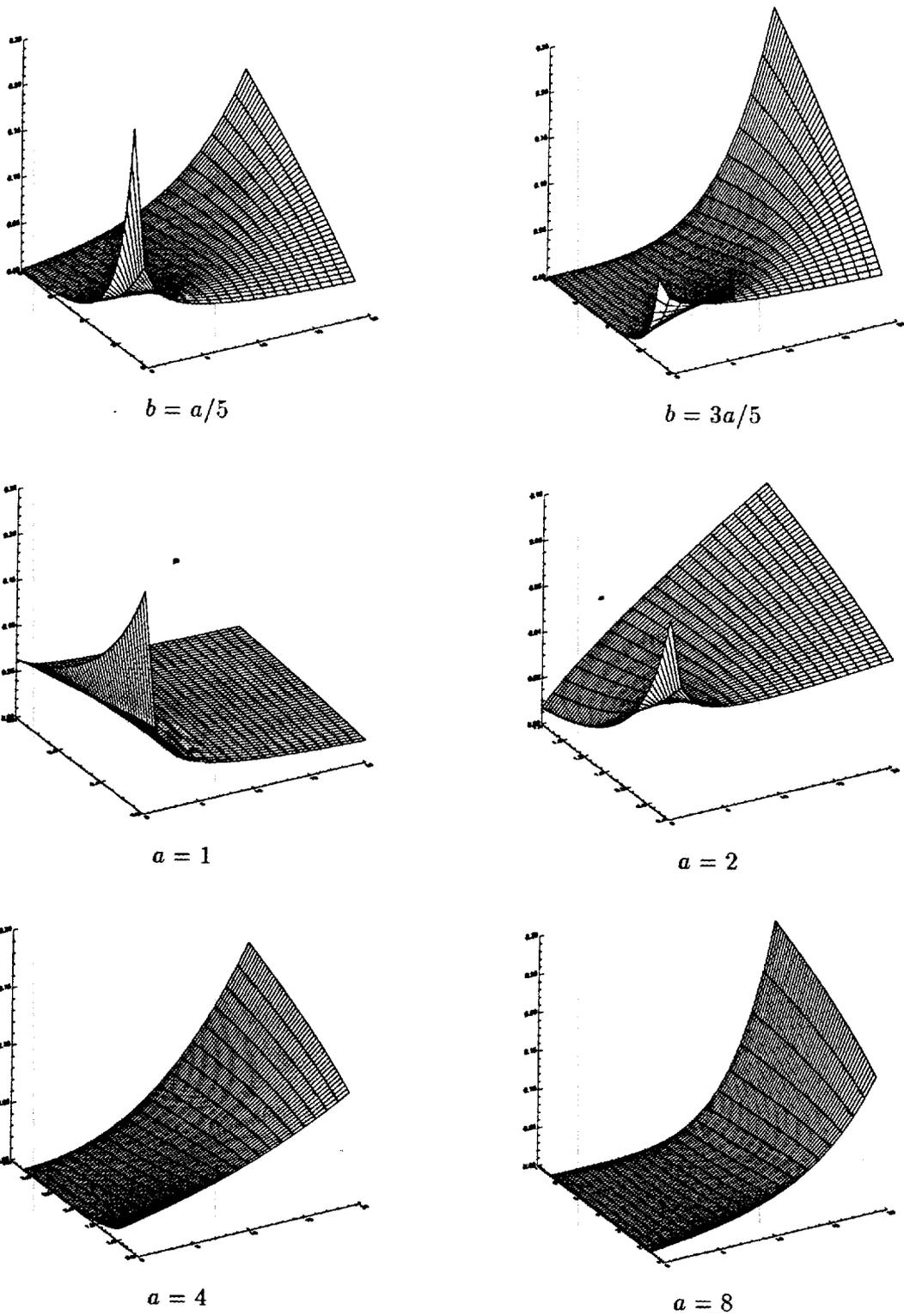
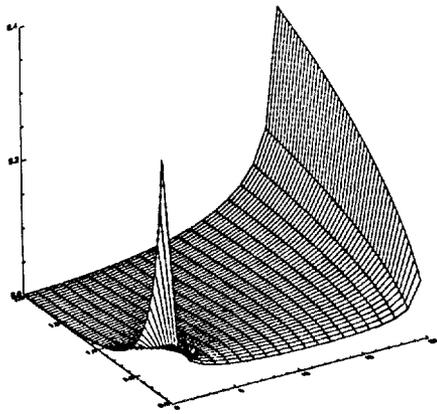
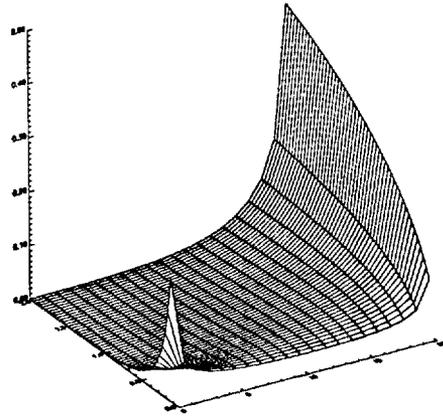


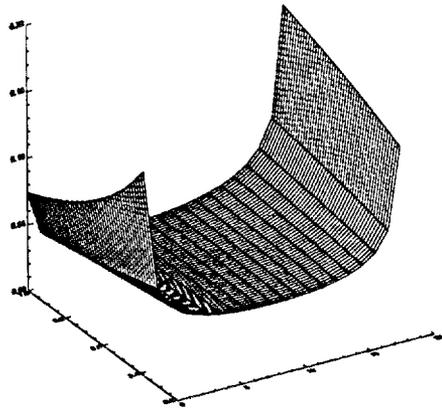
Figure 5: Model 3



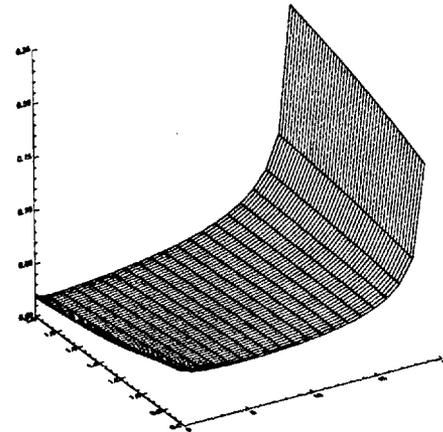
$b = 3a/2$



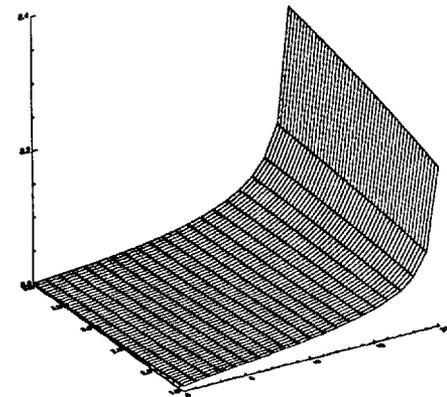
$b = 4a$



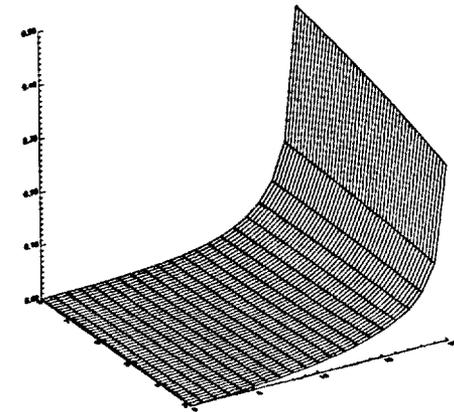
$a = 0.25$



$a = 0.5$



$a = 1$



$a = 2$

Figure 6: Model 6

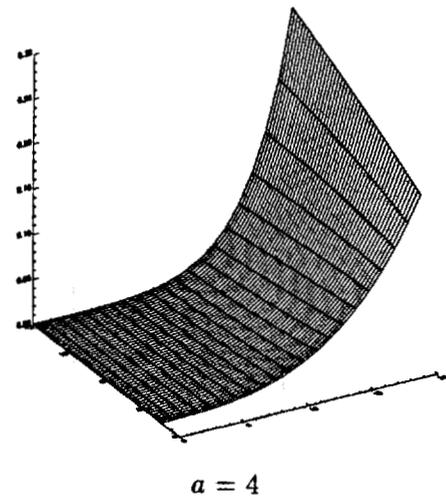
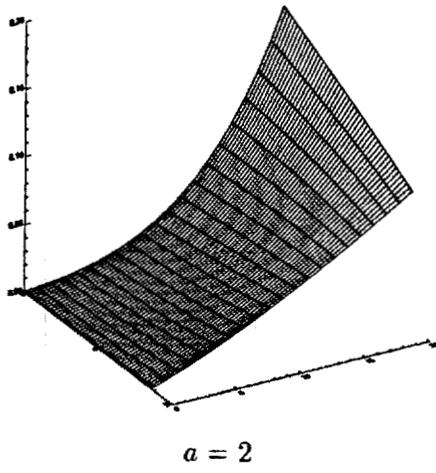
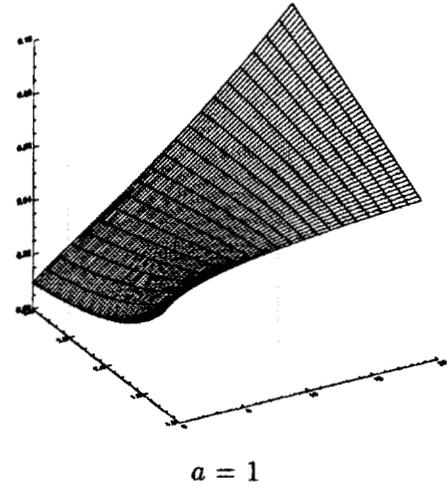
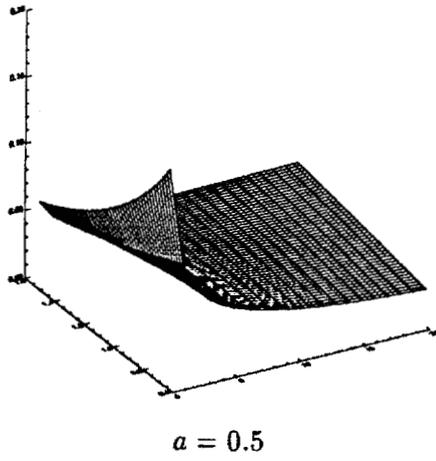
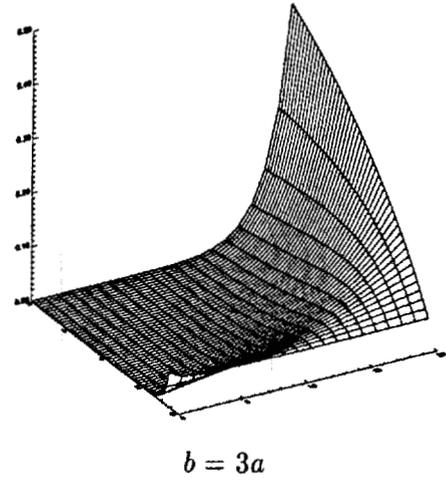
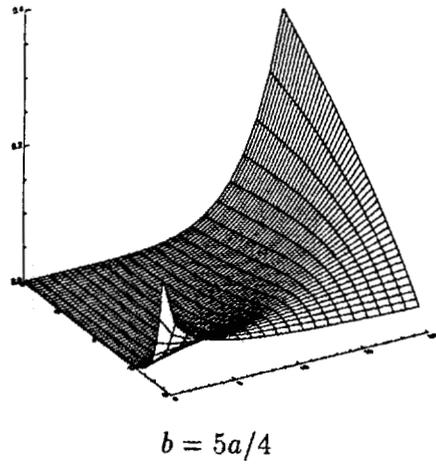
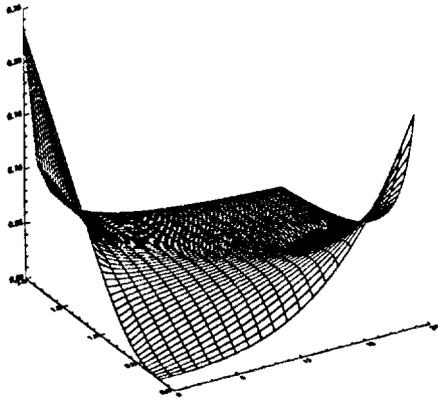
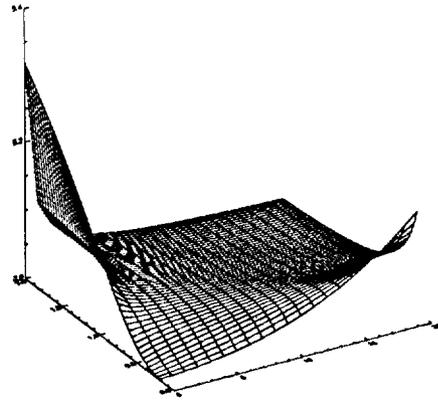


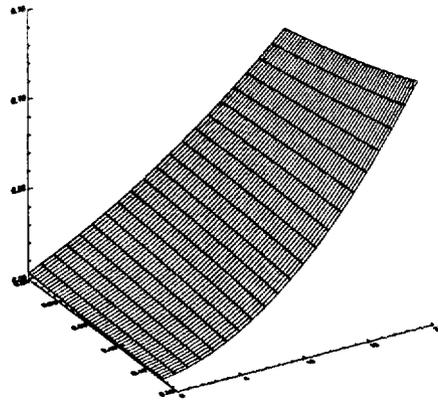
Figure 7: Model 8



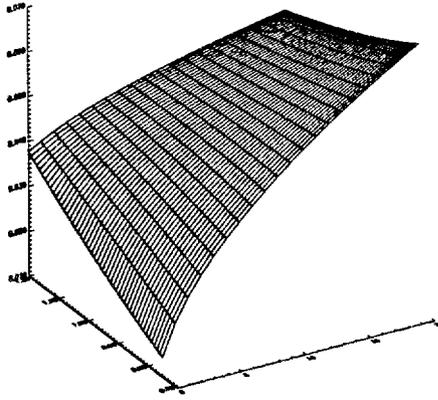
$B = 5A/4$



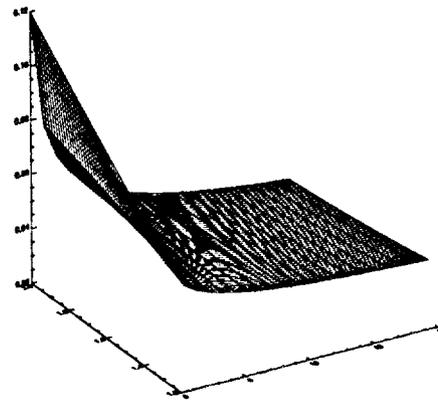
$B = 2A$



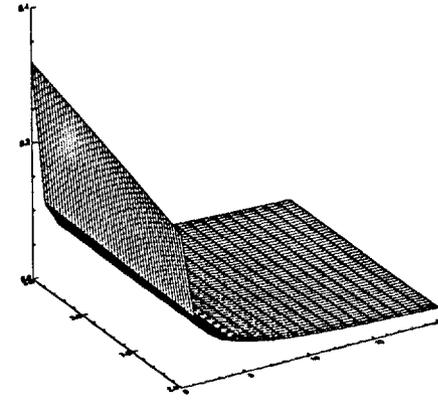
$A = 0.3$



$A = 0.6$

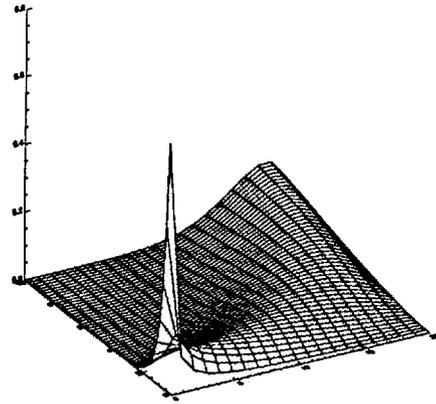


$A = 1$

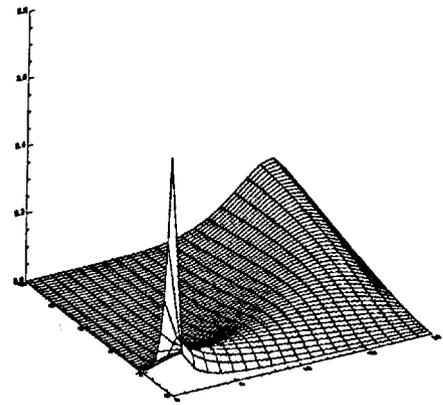


$A = 2$

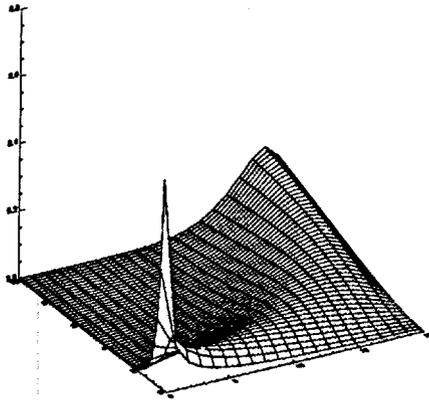
Figure 8: Model 9



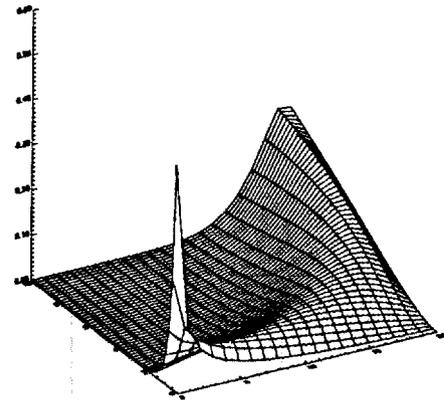
$$b = 5a/4$$



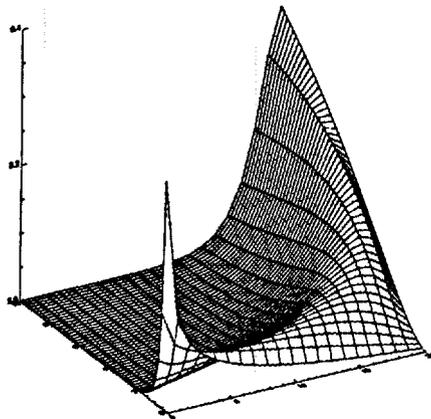
$$b = 3a/2$$



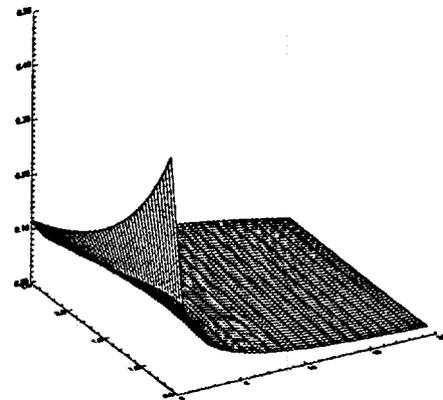
$$b = 2a$$



$$b = 3a$$



$$b = 5a$$



$$a = 0.5$$

Figure 9: Model 10

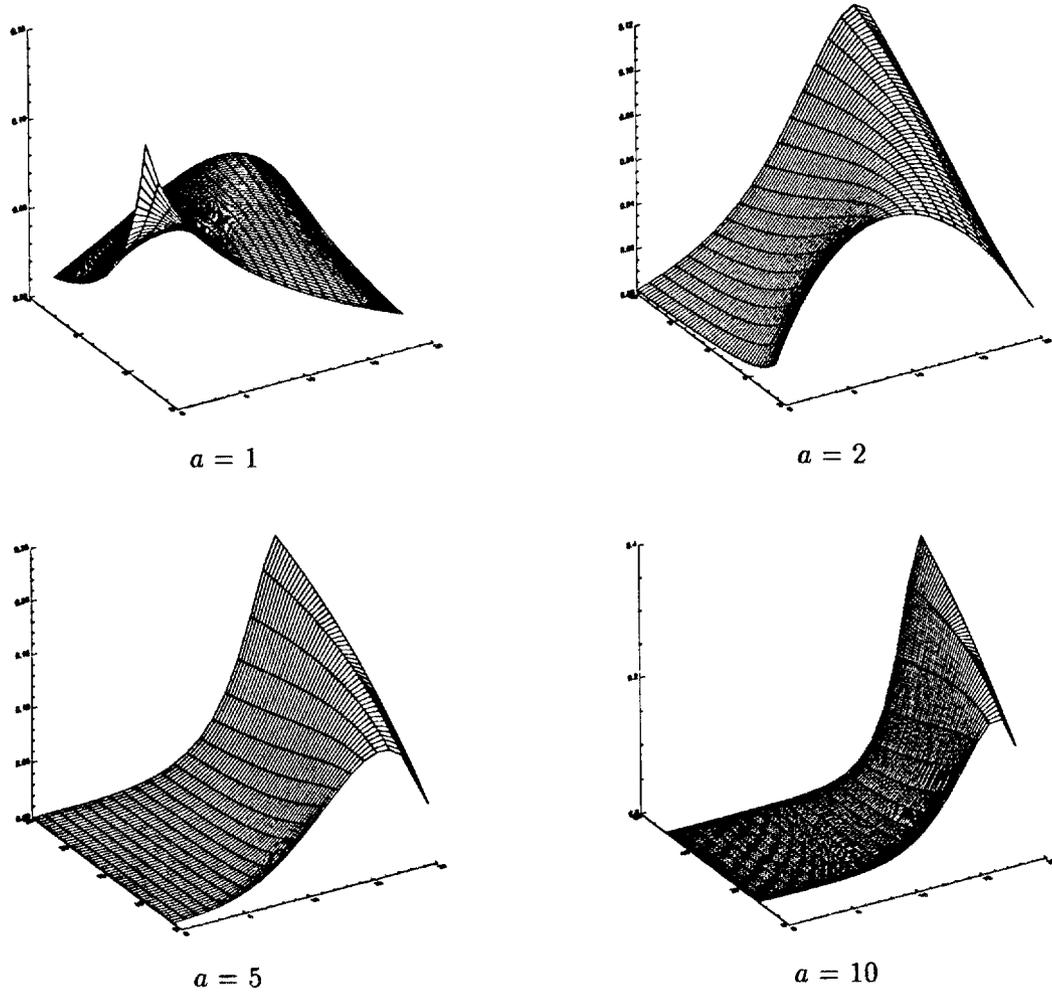
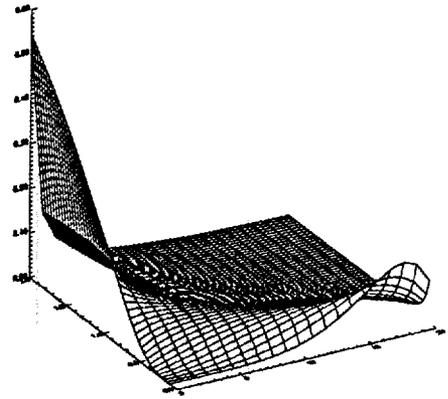
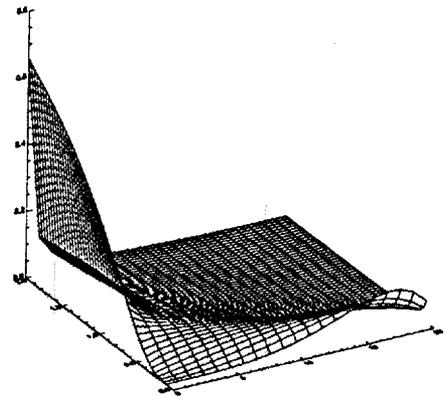


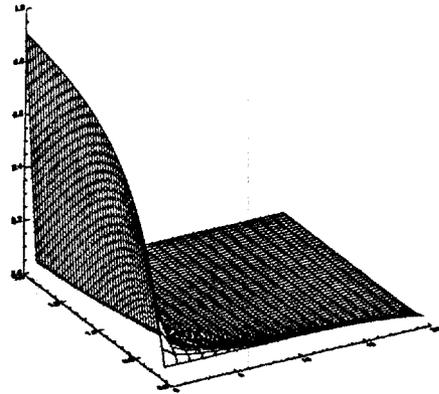
Figure 10: Model 10



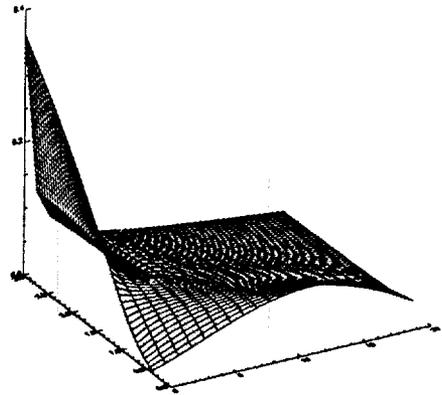
$$B = 5A/4$$



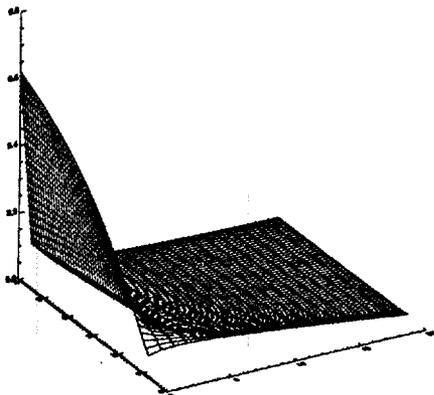
$$B = 2A$$



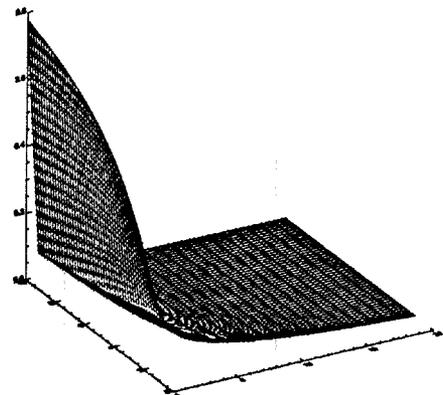
$$B = 10A$$



$$A = 0.3$$

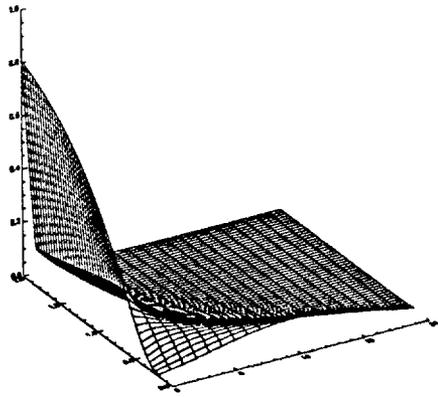


$$A = 0.6$$

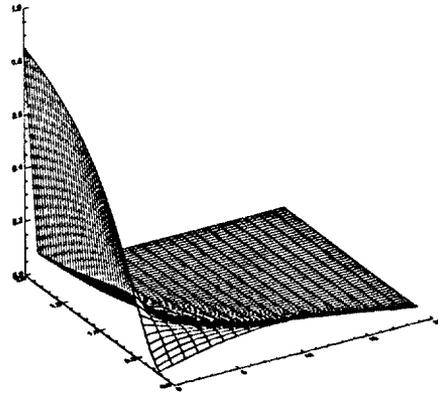


$$A = 1$$

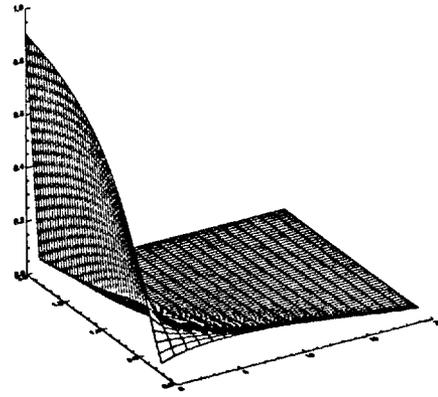
Figure 11: Model 12



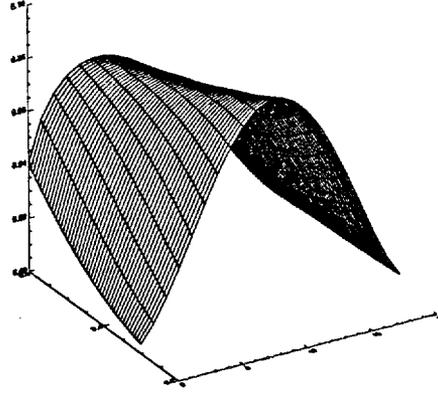
$$B = 3A/2$$



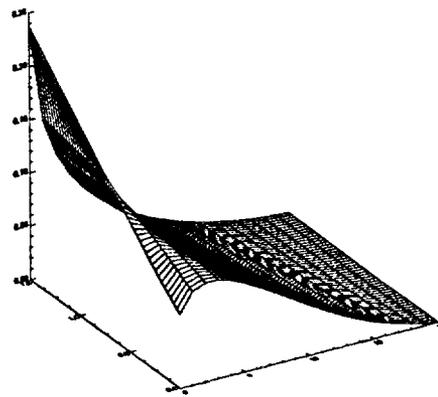
$$B = 2A$$



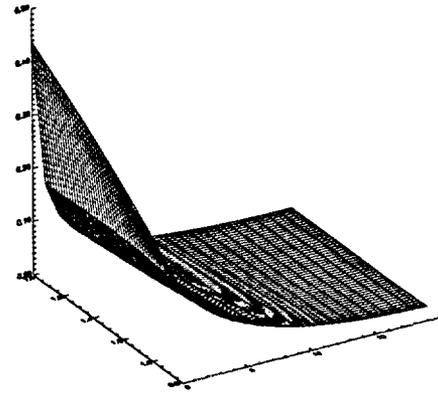
$$B = 3A$$



$$A = 0.2$$

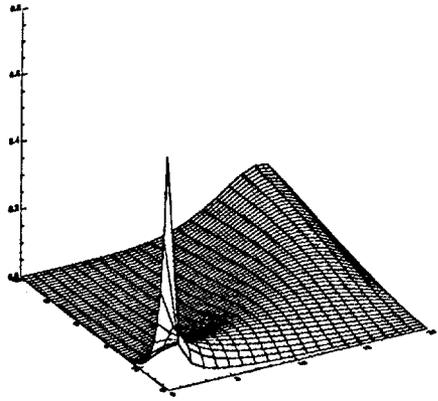


$$A = 0.4$$

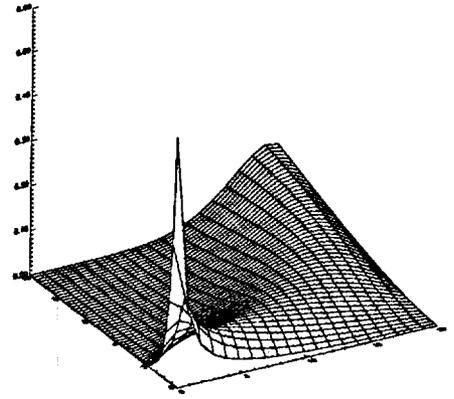


$$A = 0.6$$

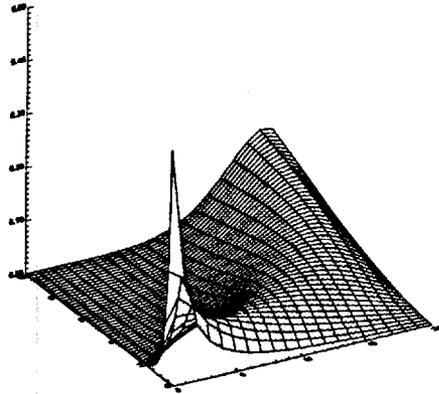
Figure 12: Model 13



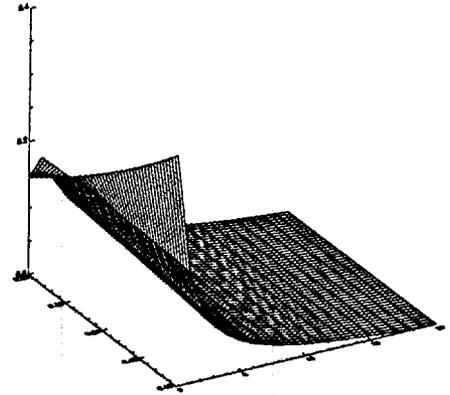
$u = a/5$



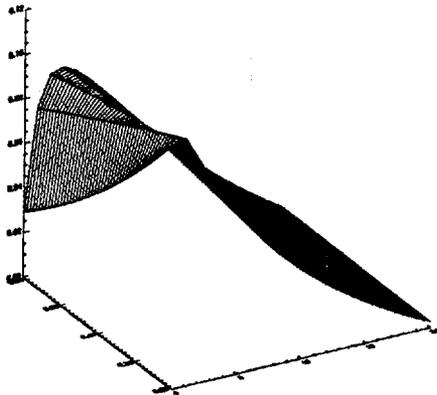
$u = 3a/5$



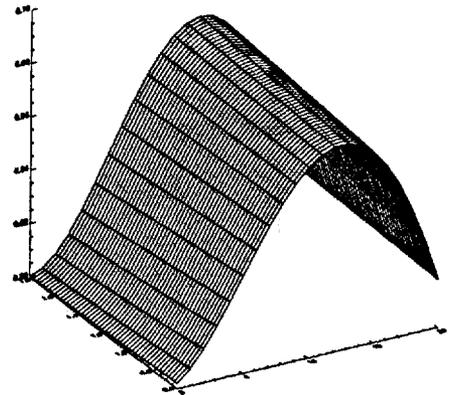
$a = 0.5$



$a = 1$

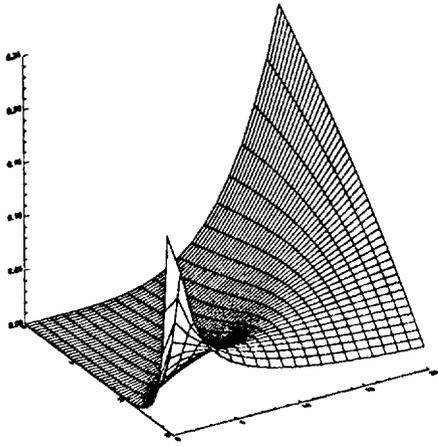


$a = 3$

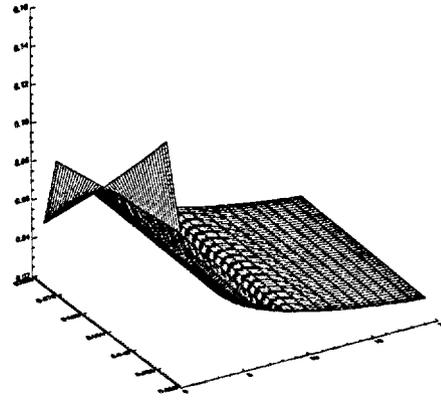


$a = 6$

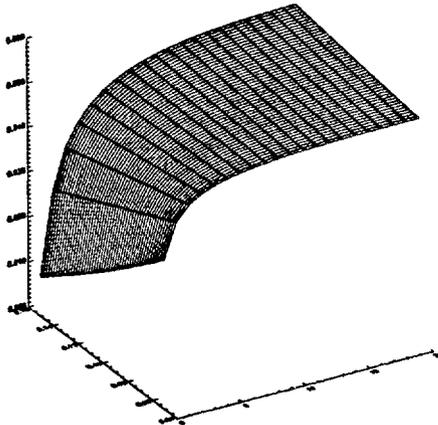
Figure 13: Model 15



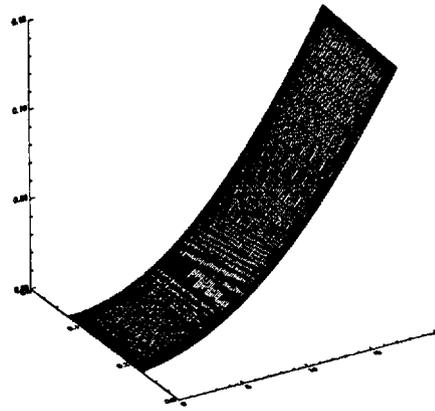
$$u = 3a/20$$



$$a = 0.5$$

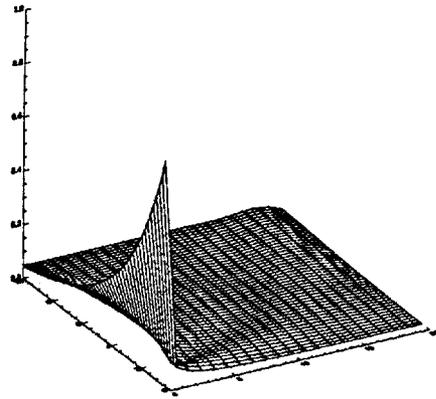


$$a = 1$$

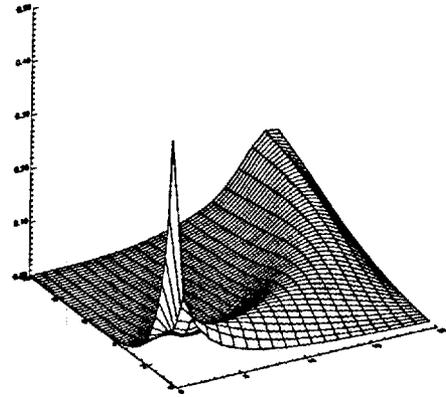


$$a = 3$$

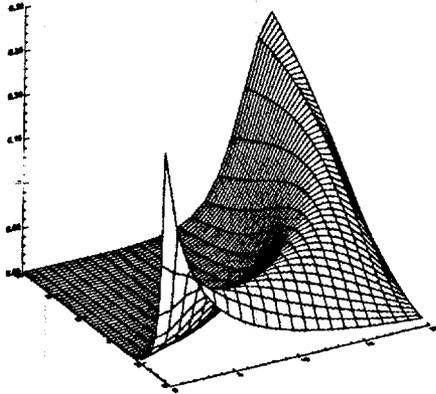
Figure 14: Model 16



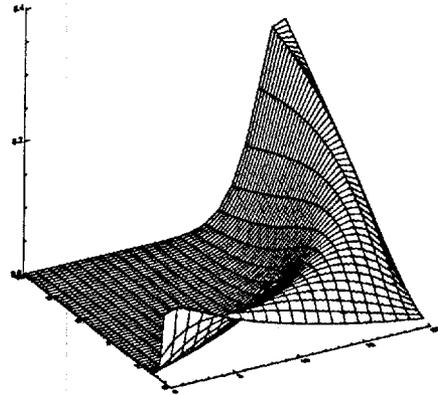
$$b = a/10$$



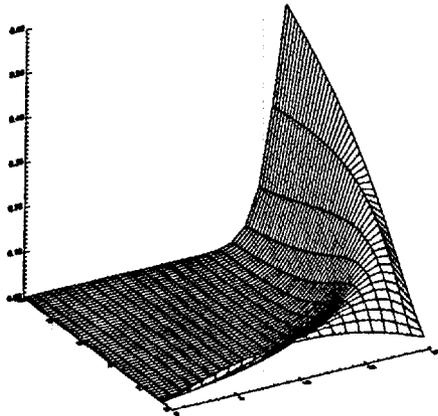
$$b = a/2$$



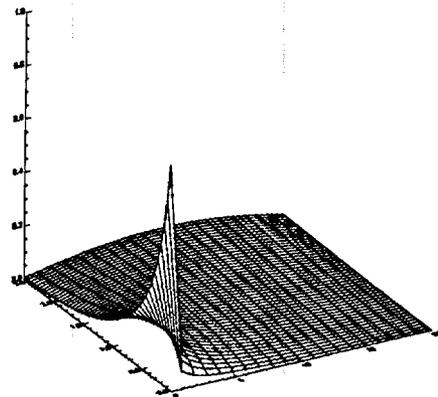
$$b = a$$



$$b = 2a$$

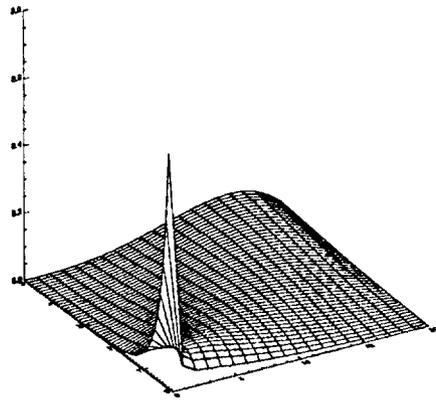


$$b = 5a$$

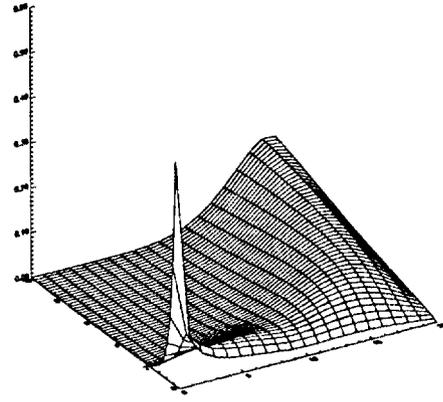


$$a = 0.5$$

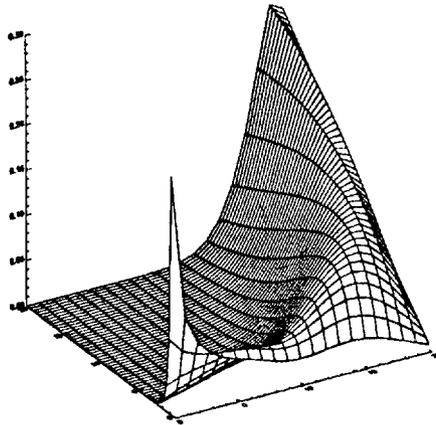
Figure 15: Model 17



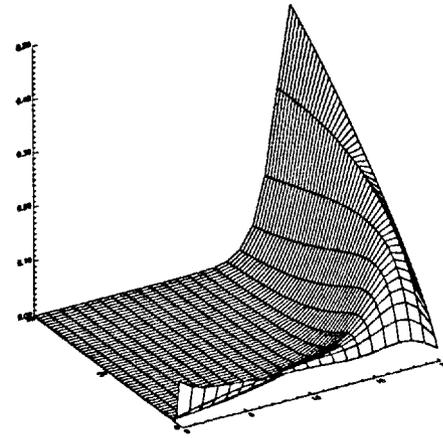
$a = 1$



$a = 2$

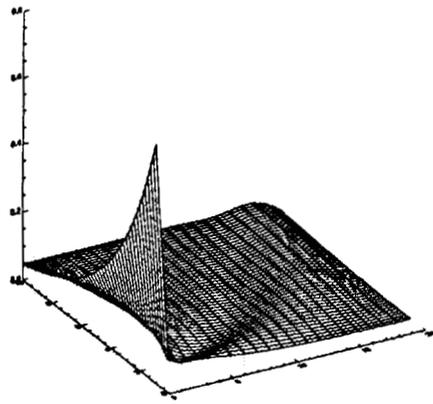


$a = 4$

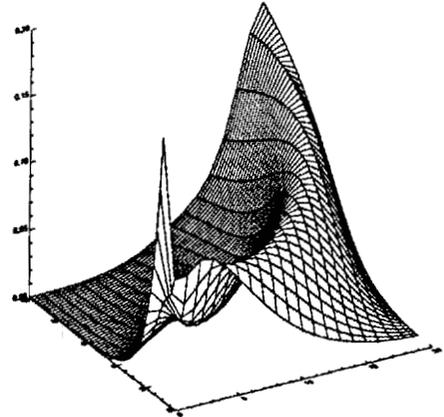


$a = 8$

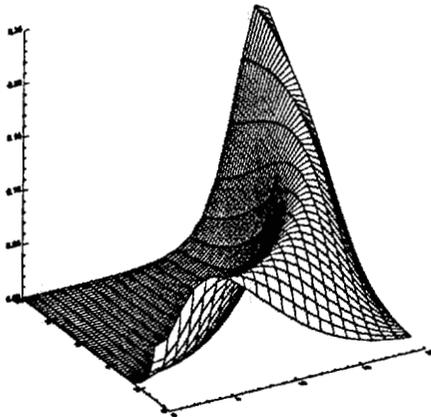
Figure 16: Model 17



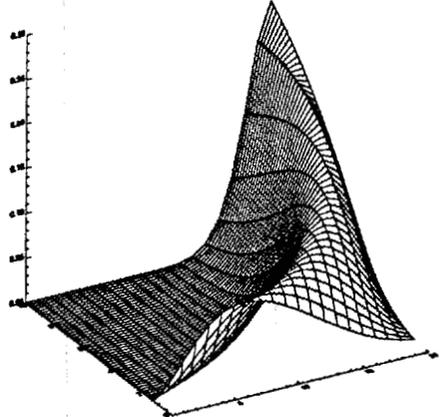
$$b = a/10$$



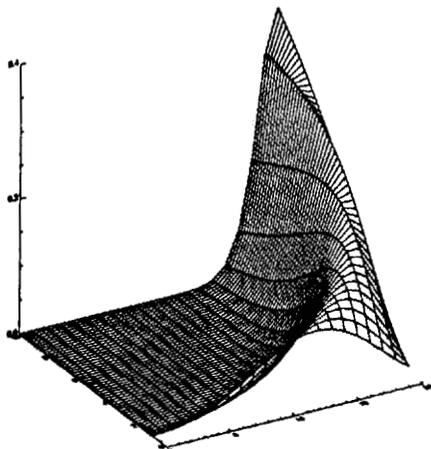
$$b = a/2$$



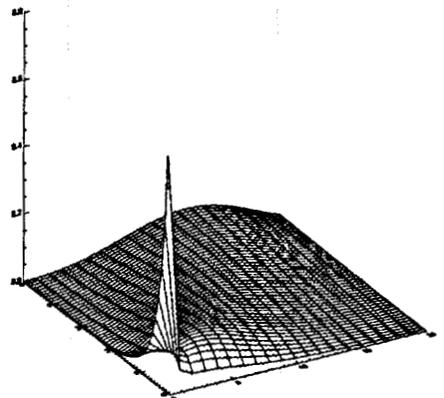
$$b = a$$



$$b = 2a$$

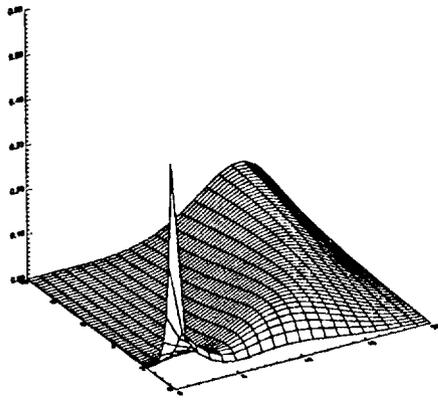


$$b = 5a$$

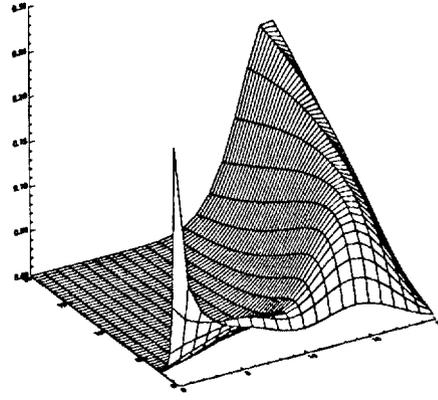


$$a = 1$$

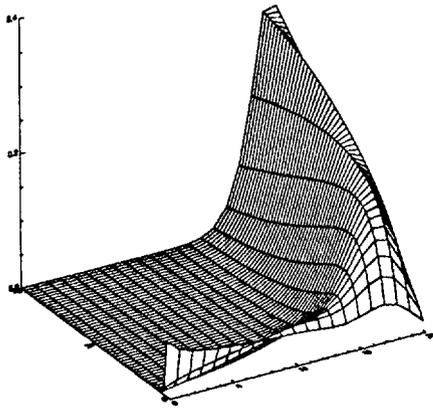
Figure 17: Model 18



$a = 2$

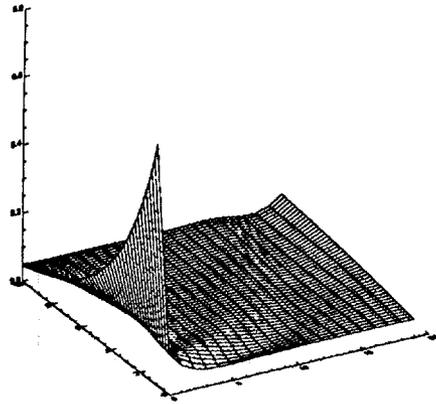


$a = 4$

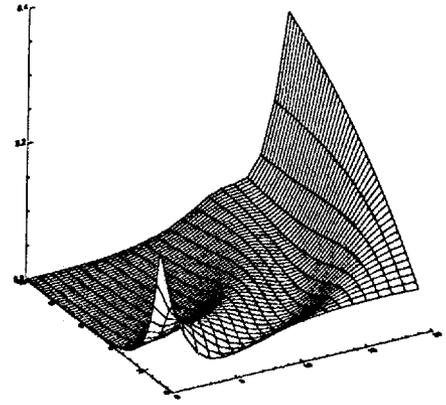


$a = 8$

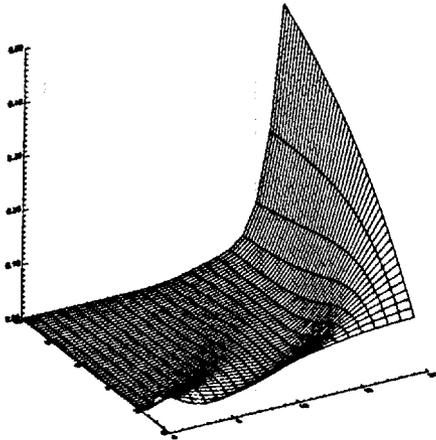
Figure 18: Model 18



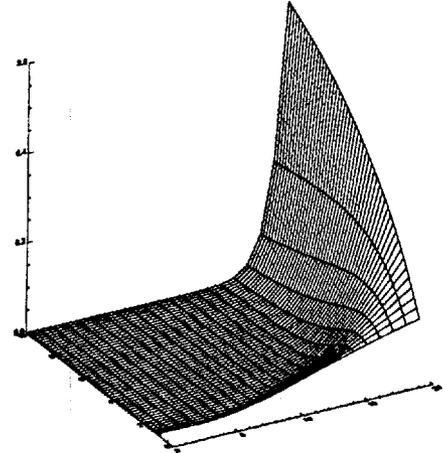
$b = a/10$



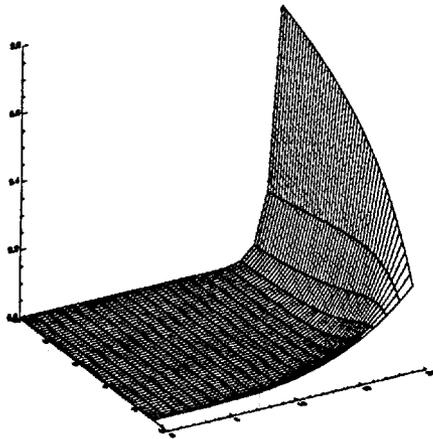
$b = a/2$



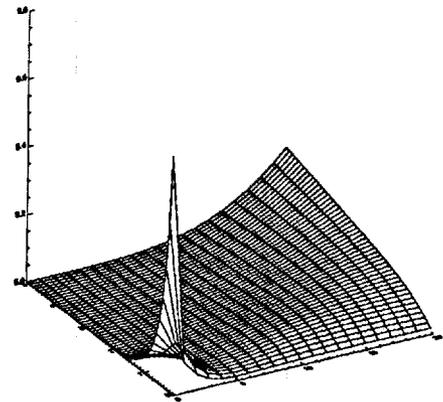
$b = a$



$b = 2a$



$b = 5a$



$a = 1$

Figure 19: Model 19

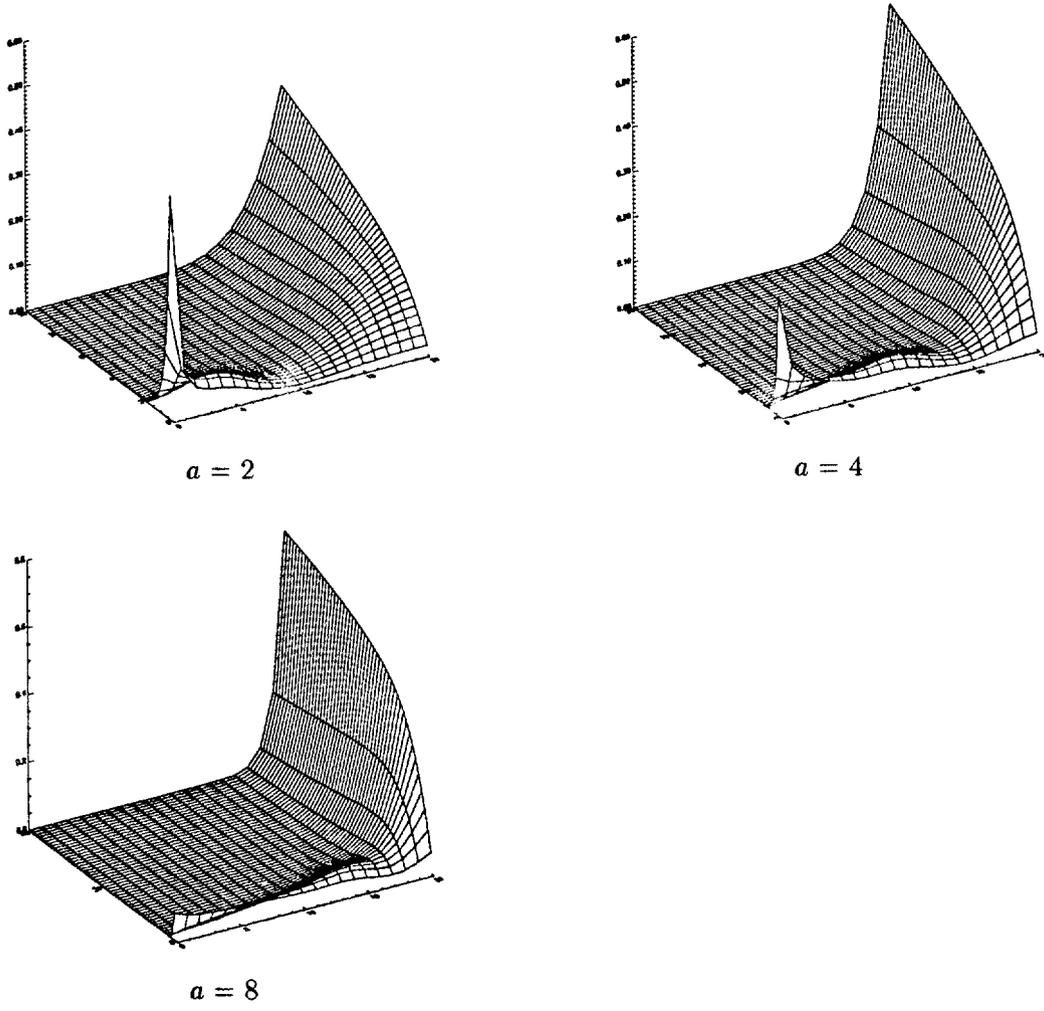


Figure 20: Model 19

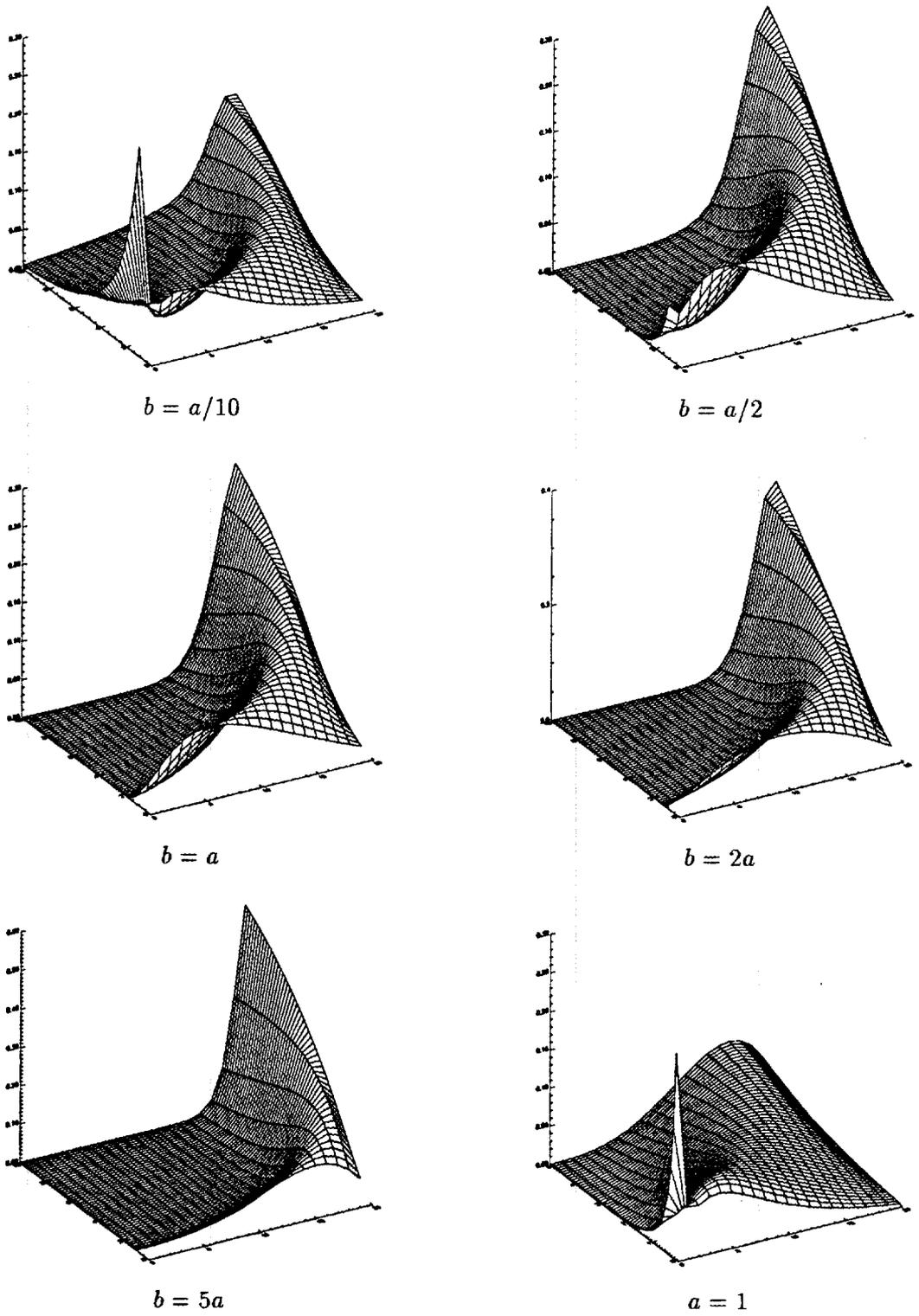
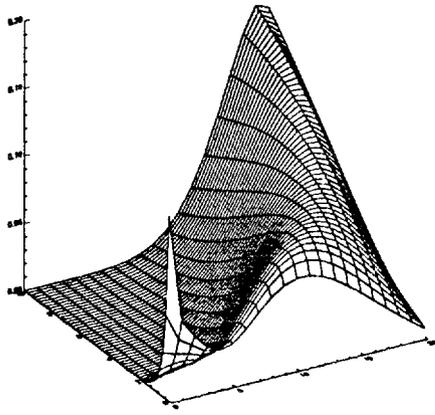
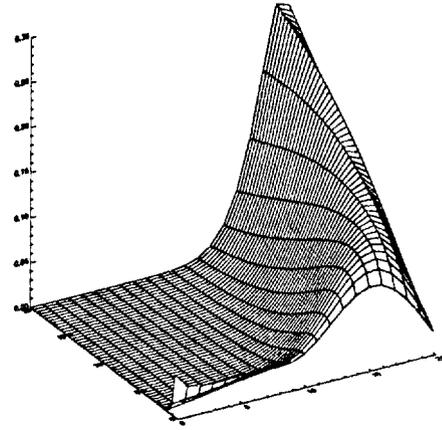


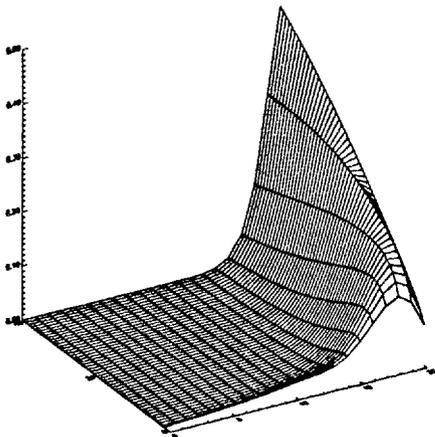
Figure 21: Model 20



$a = 2$

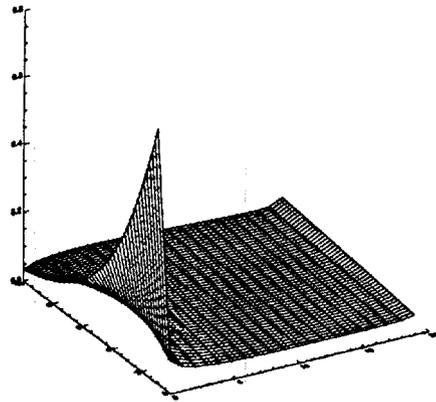


$a = 4$

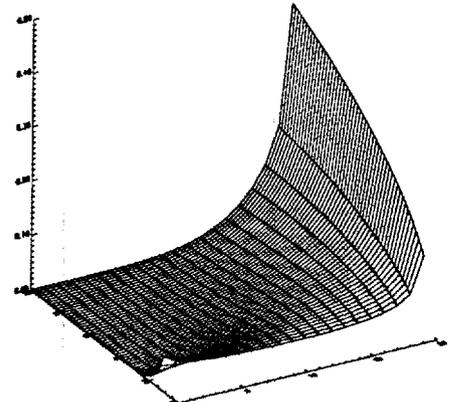


$a = 8$

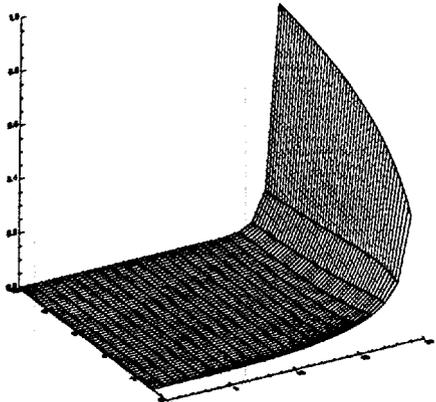
Figure 22: Model 20



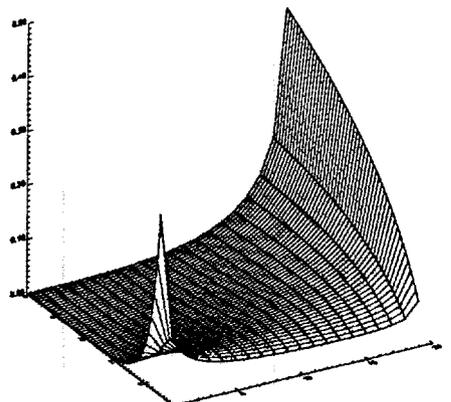
$b = a/2$



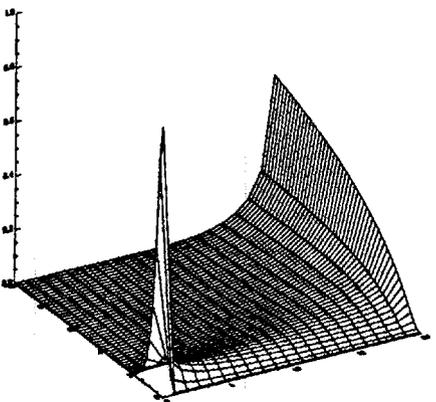
$b = a$



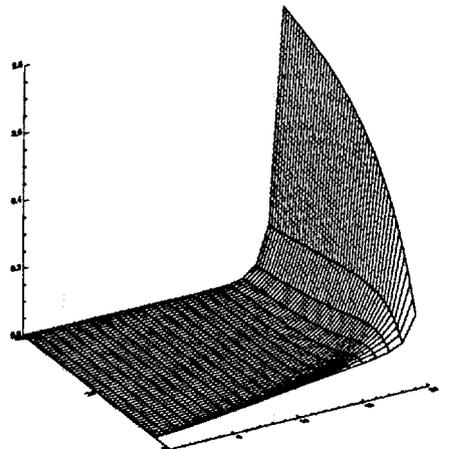
$b = 5a$



$a = 1$

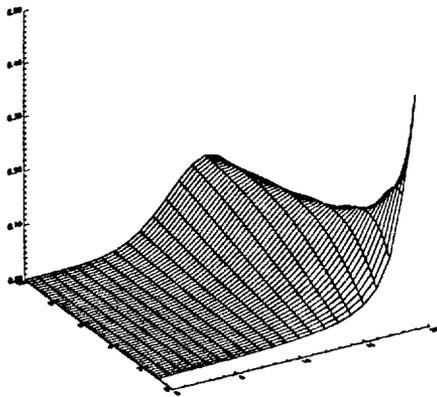


$a = 2$

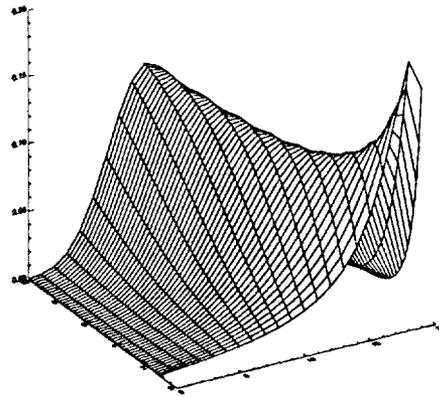


$a = 8$

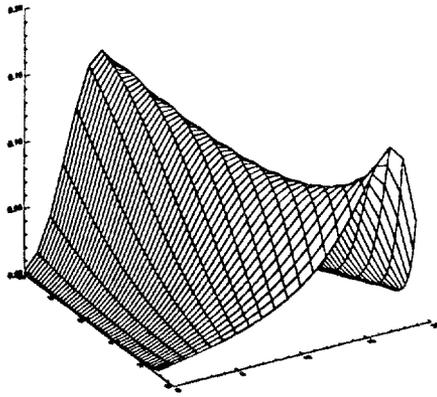
Figure 23: Model 22



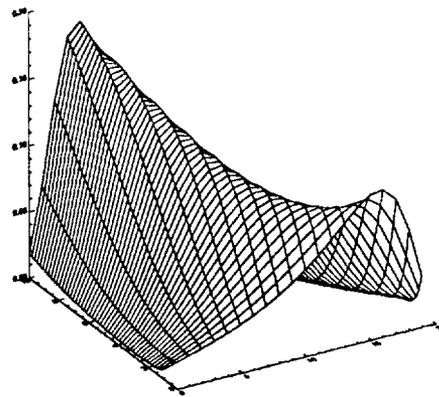
$$k = a/4$$



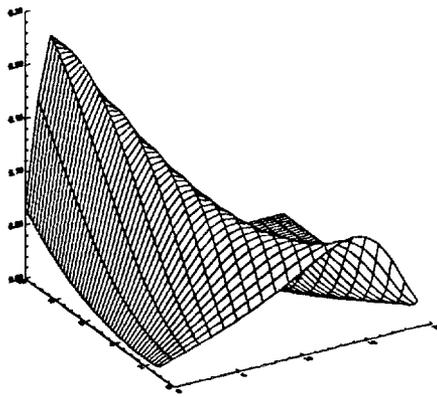
$$k = a/2$$



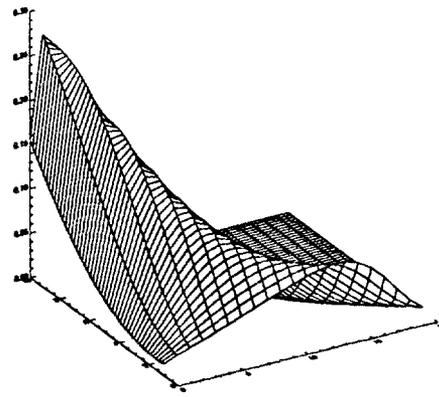
$$k = 3a/4$$



$$k = a$$

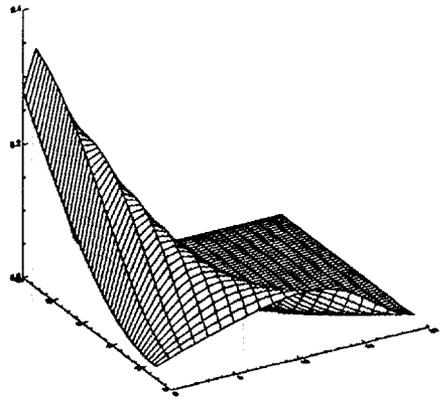


$$k = 5a/4$$

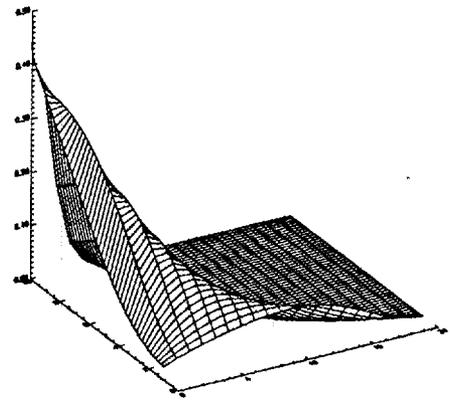


$$k = 3a/2$$

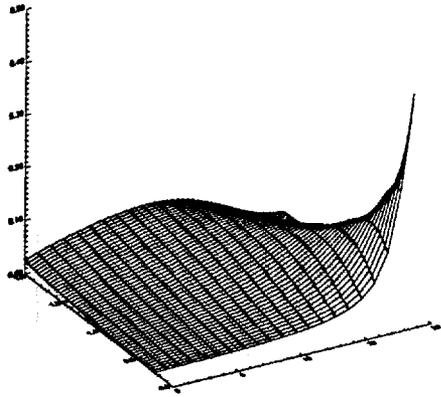
Figure 24: Model 24



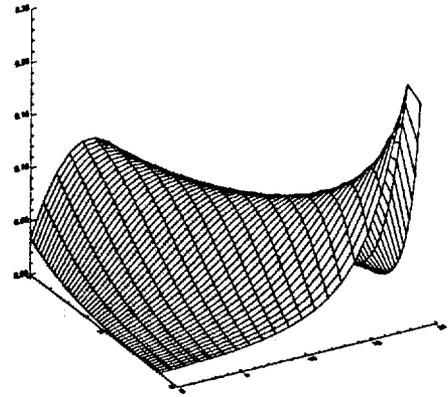
$$k = 7a/4$$



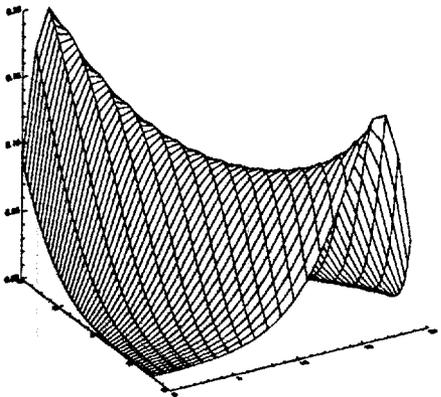
$$k = 2a$$



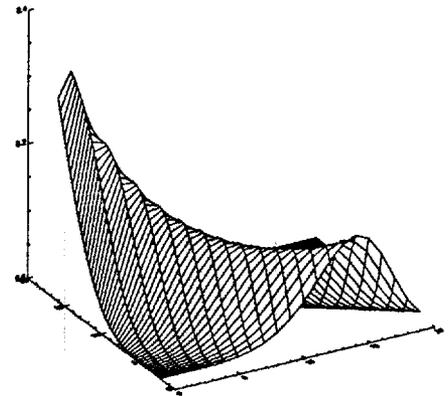
$$a = 1$$



$$a = 2$$

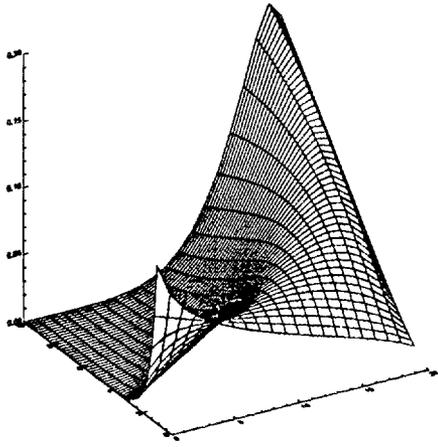


$$a = 4$$

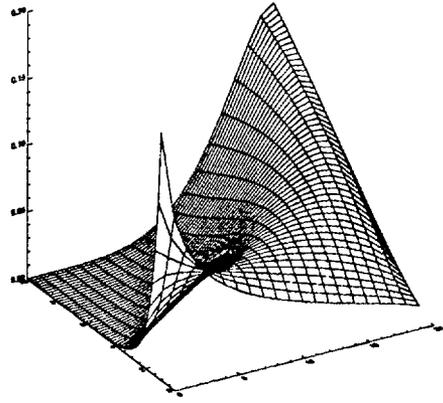


$$a = 8$$

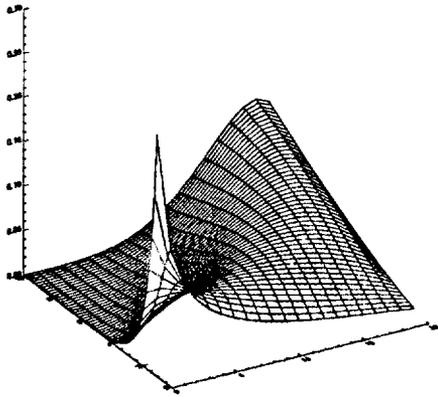
Figure 25: Model 24



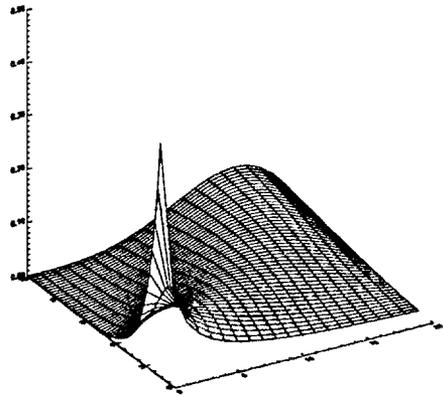
$$k = a/10$$



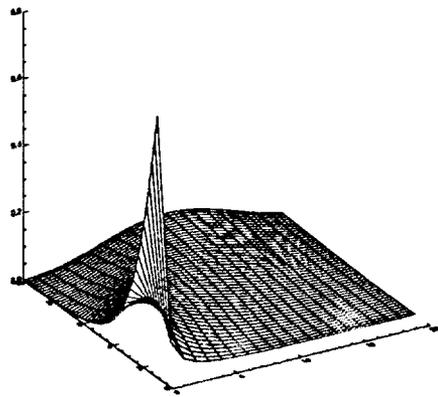
$$k = a/2$$



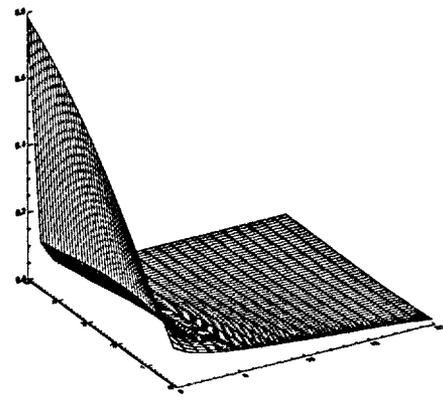
$$k = a$$



$$k = 2a$$



$$k = 5a$$



$$a = 1$$

Figure 26: Model 25

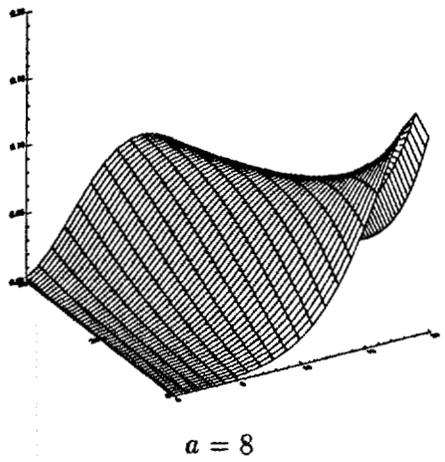
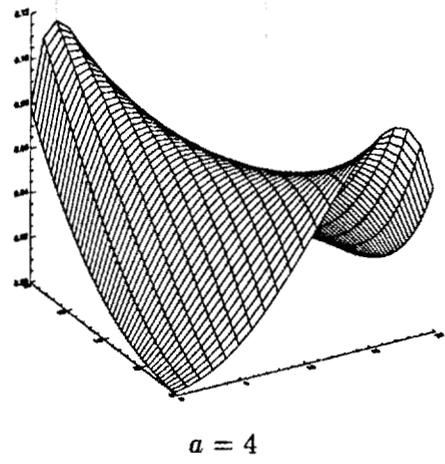
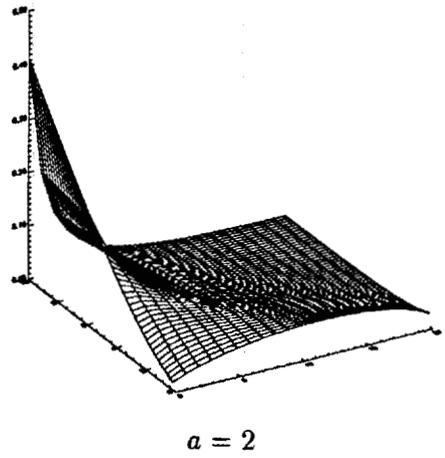
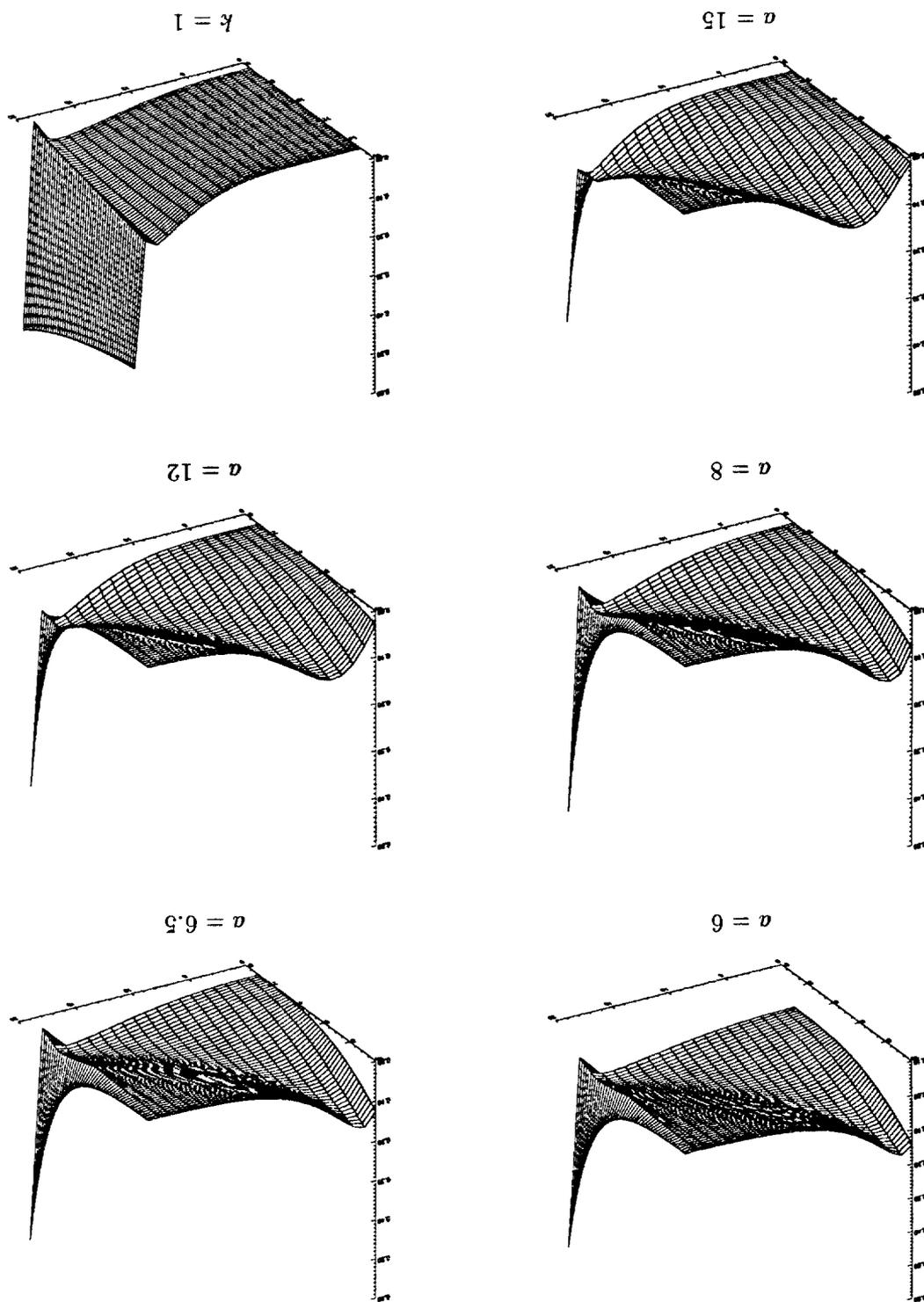
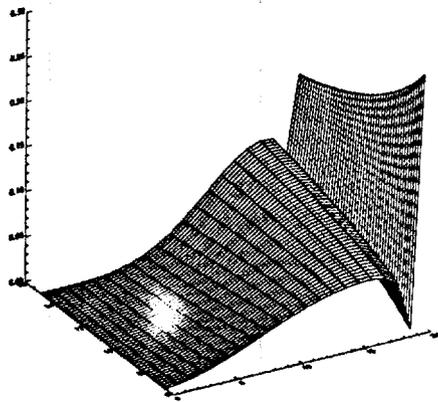


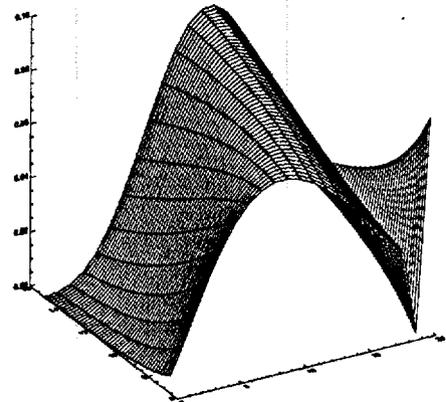
Figure 27: Model 25

Figure 28: Model 26

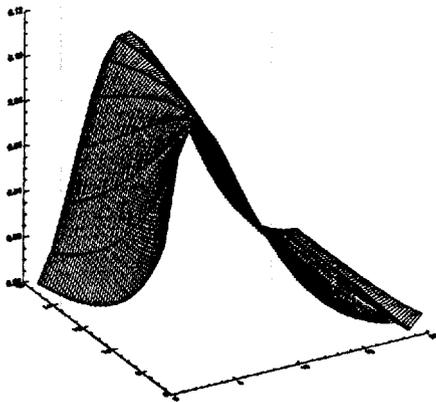




$k = 2$

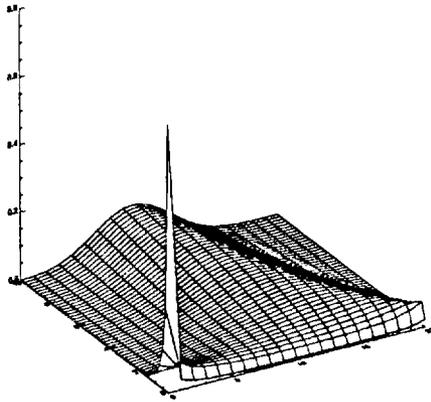


$k = 4$

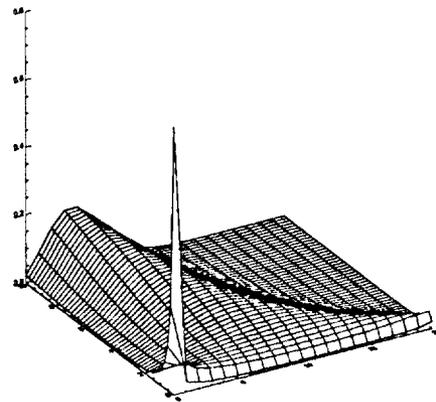


$k = 8$

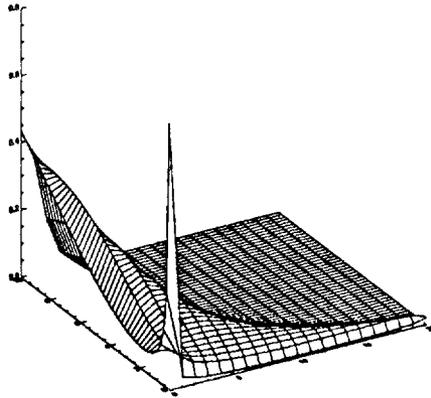
Figure 29: Model 26



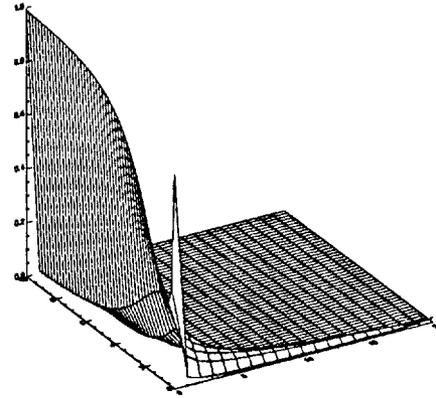
$k = a/2$



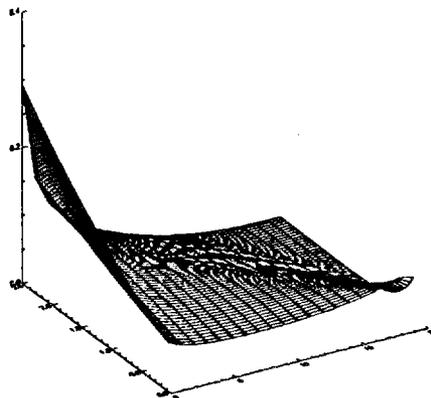
$k = a$



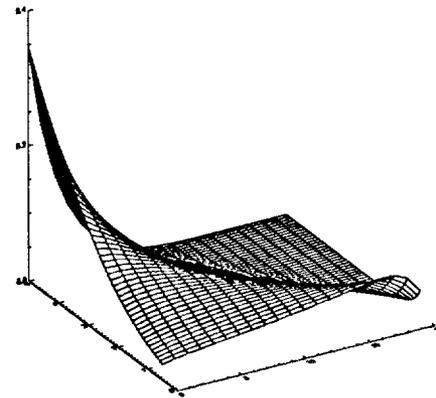
$k = 2a$



$k = 5a$



$a = 0.5$



$a = 1$

Figure 30: Model 27

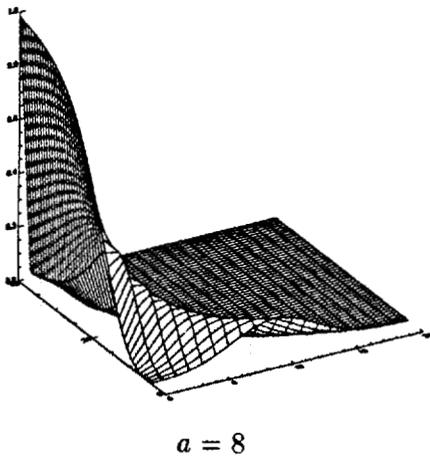
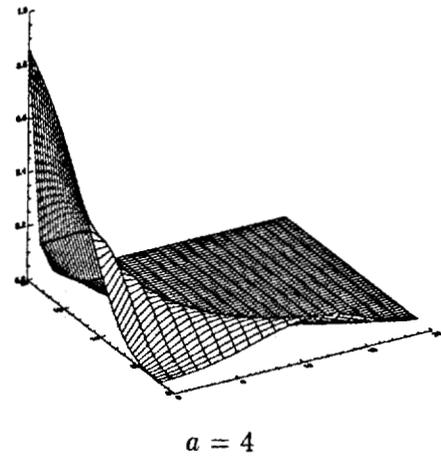
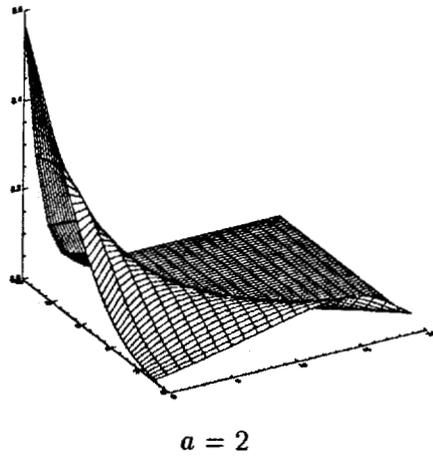


Figure 31: Model 27

### C. Two Dimensional Graphs

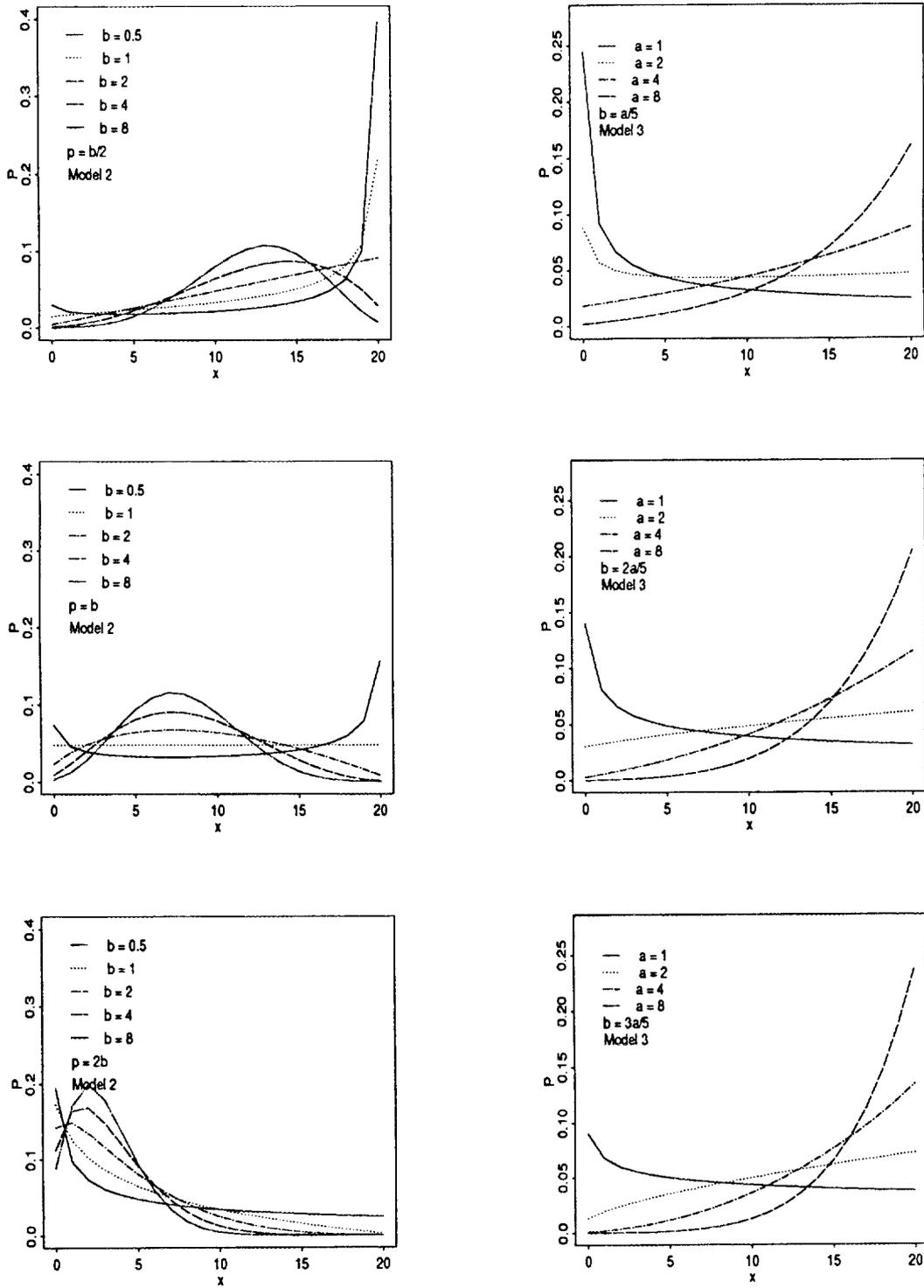


Figure 32: Models 2 and 3

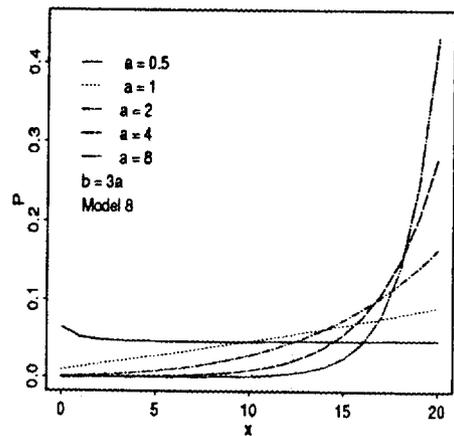
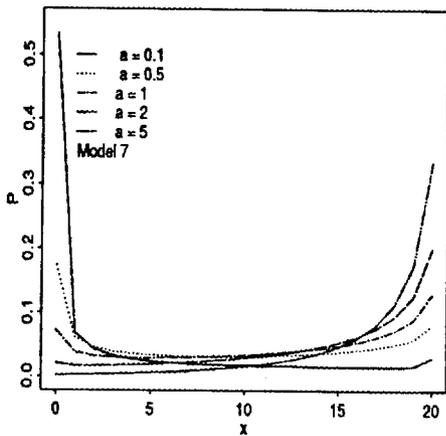
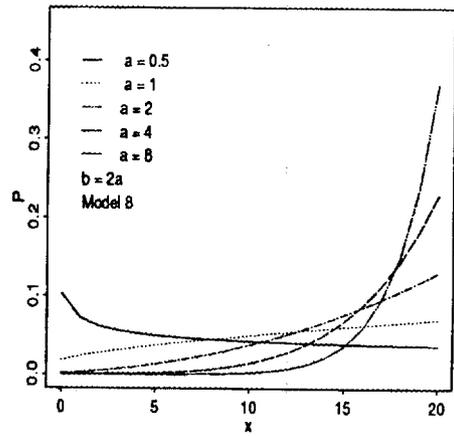
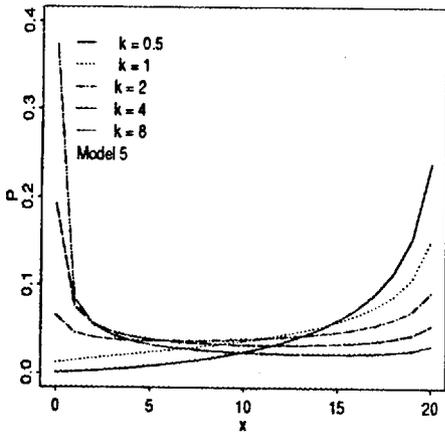
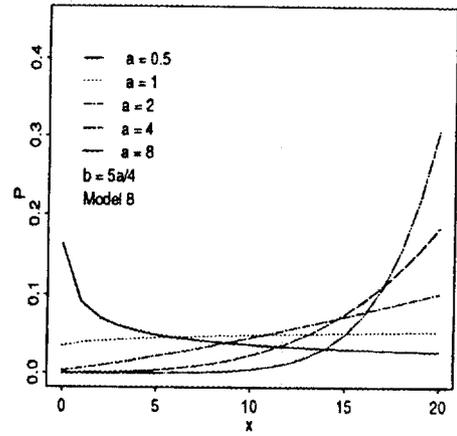
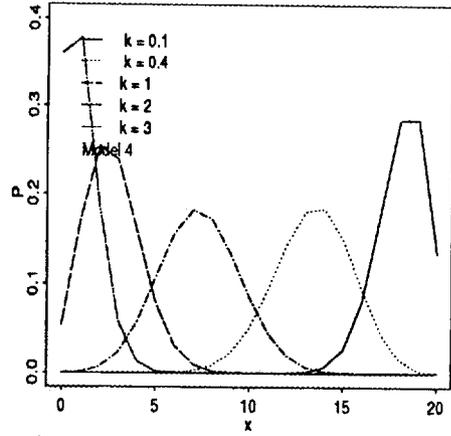


Figure 33: Models 4, 5, 7, and 8

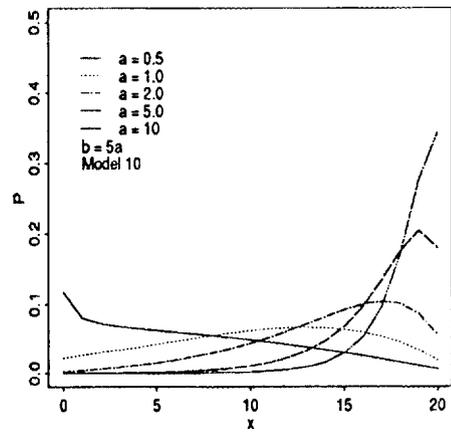
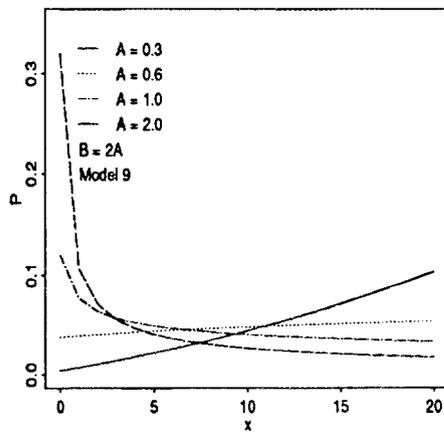
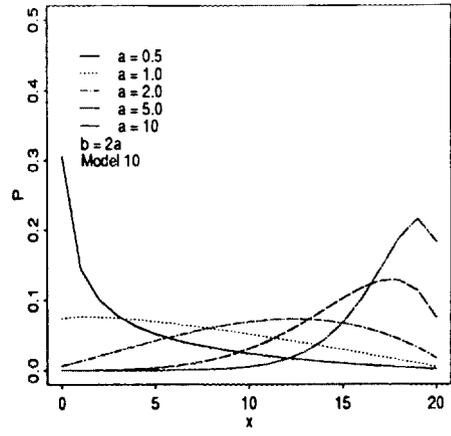
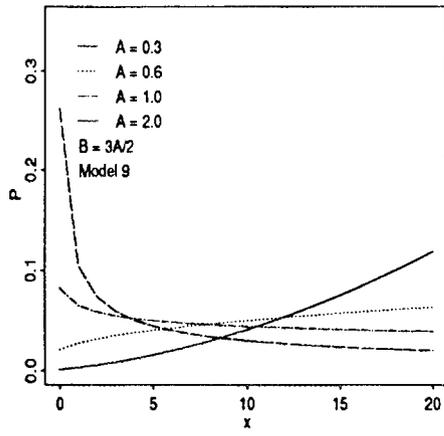
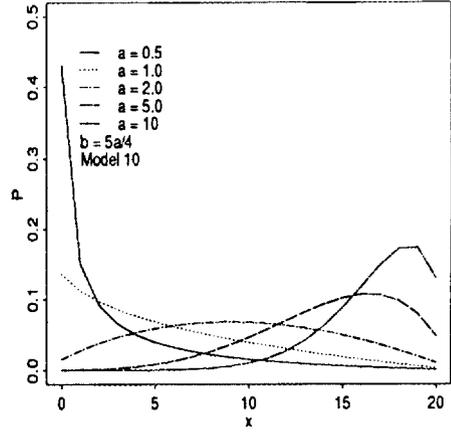
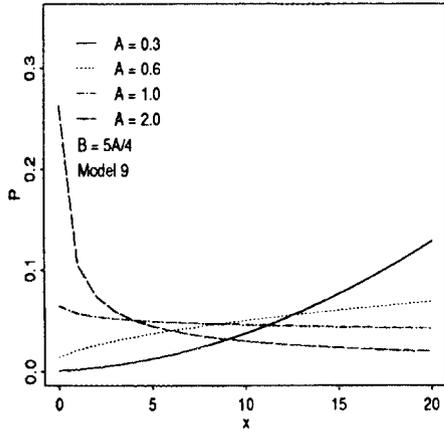


Figure 34: Models 9 and 10

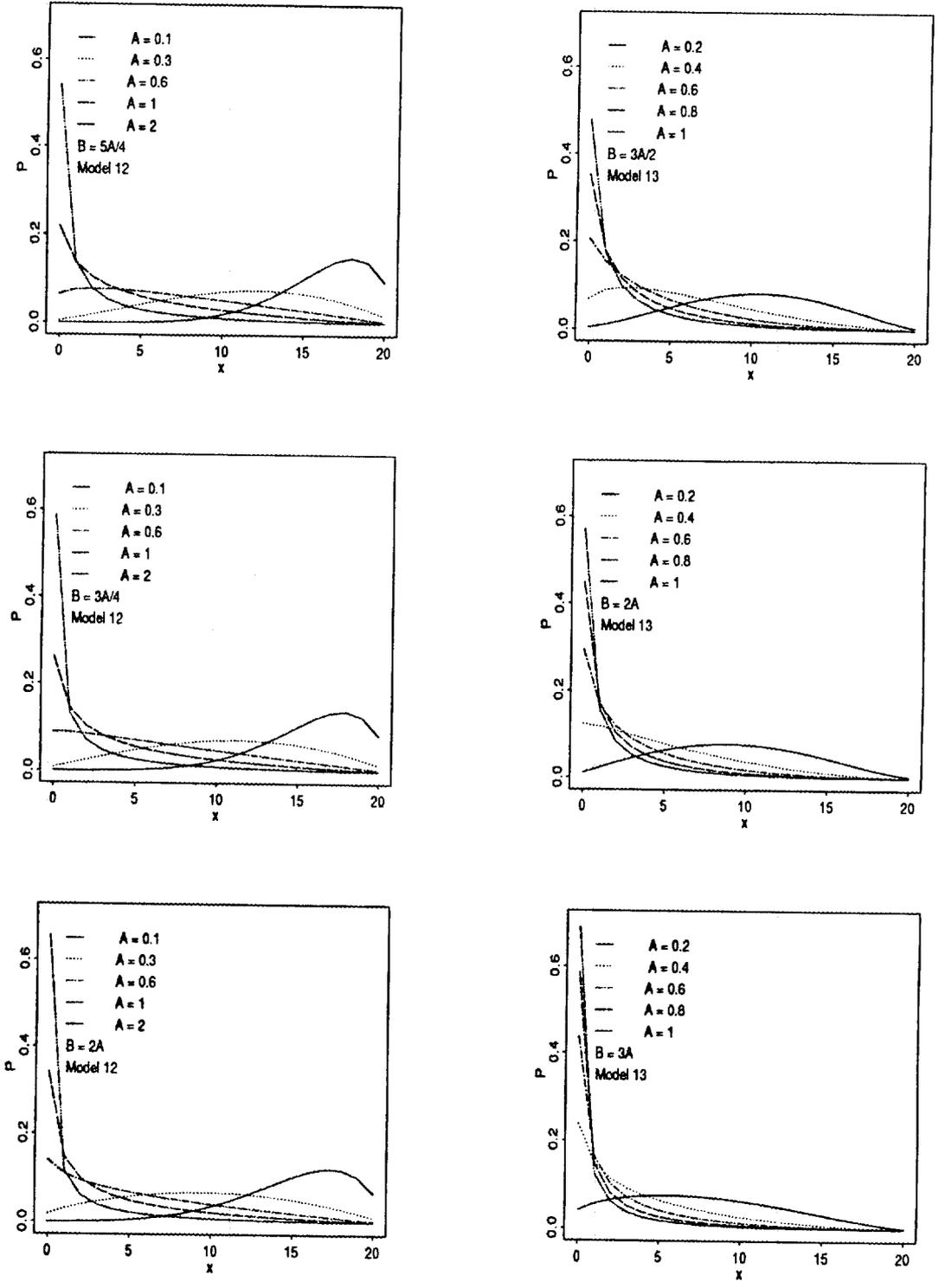


Figure 35: Models 12 and 13

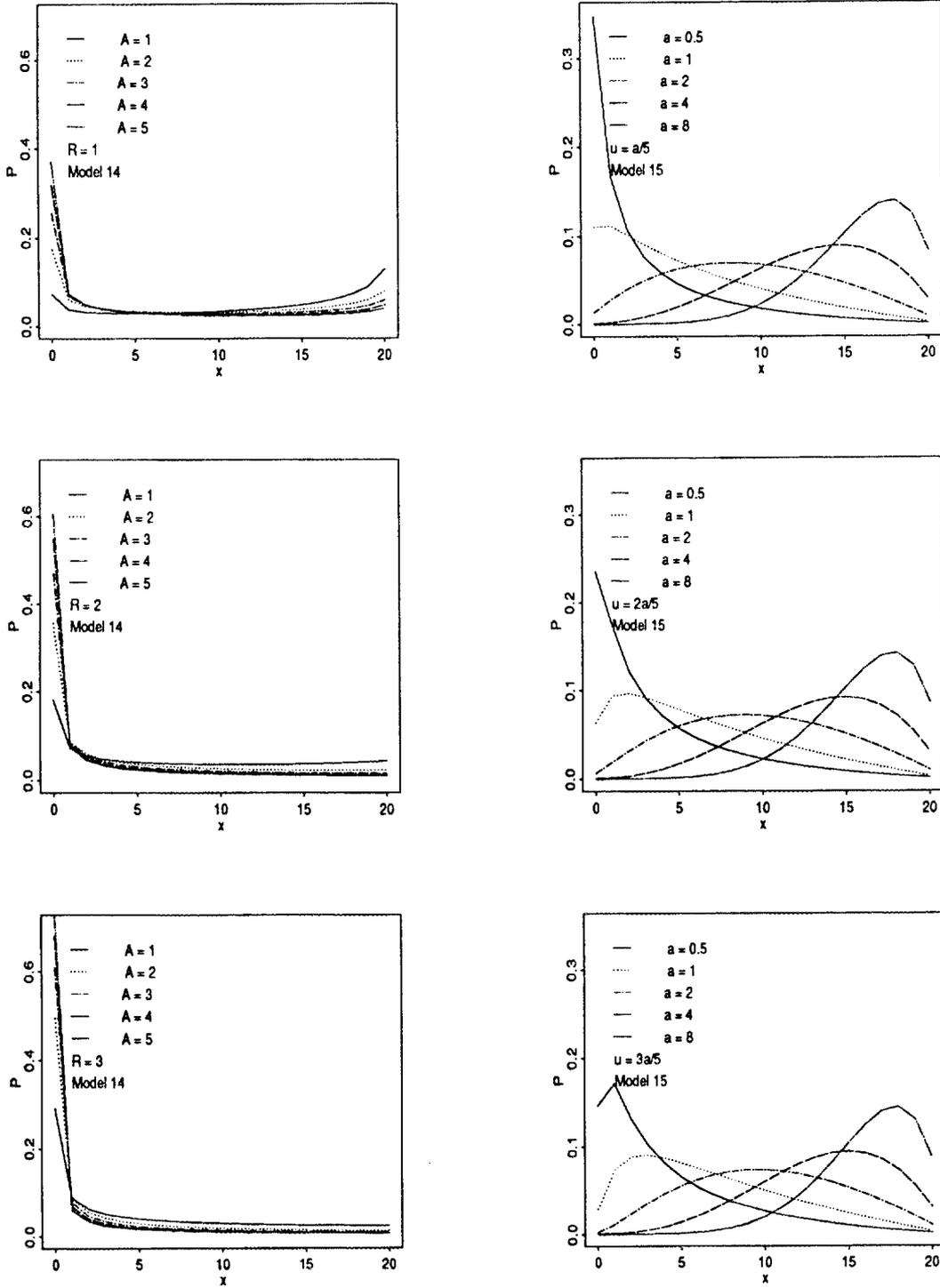


Figure 36: Models 14 and 15

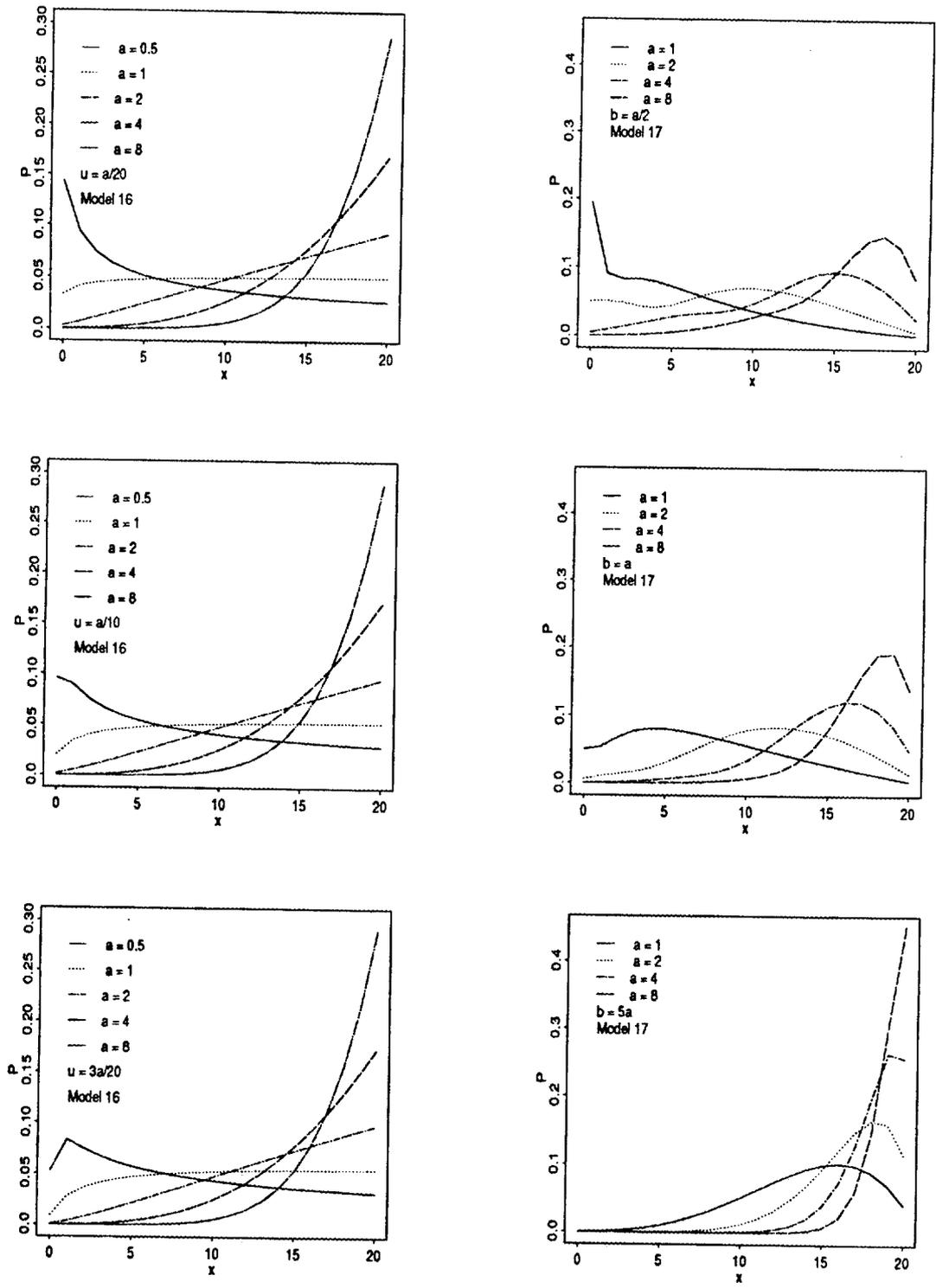


Figure 37: Models 16 and 17

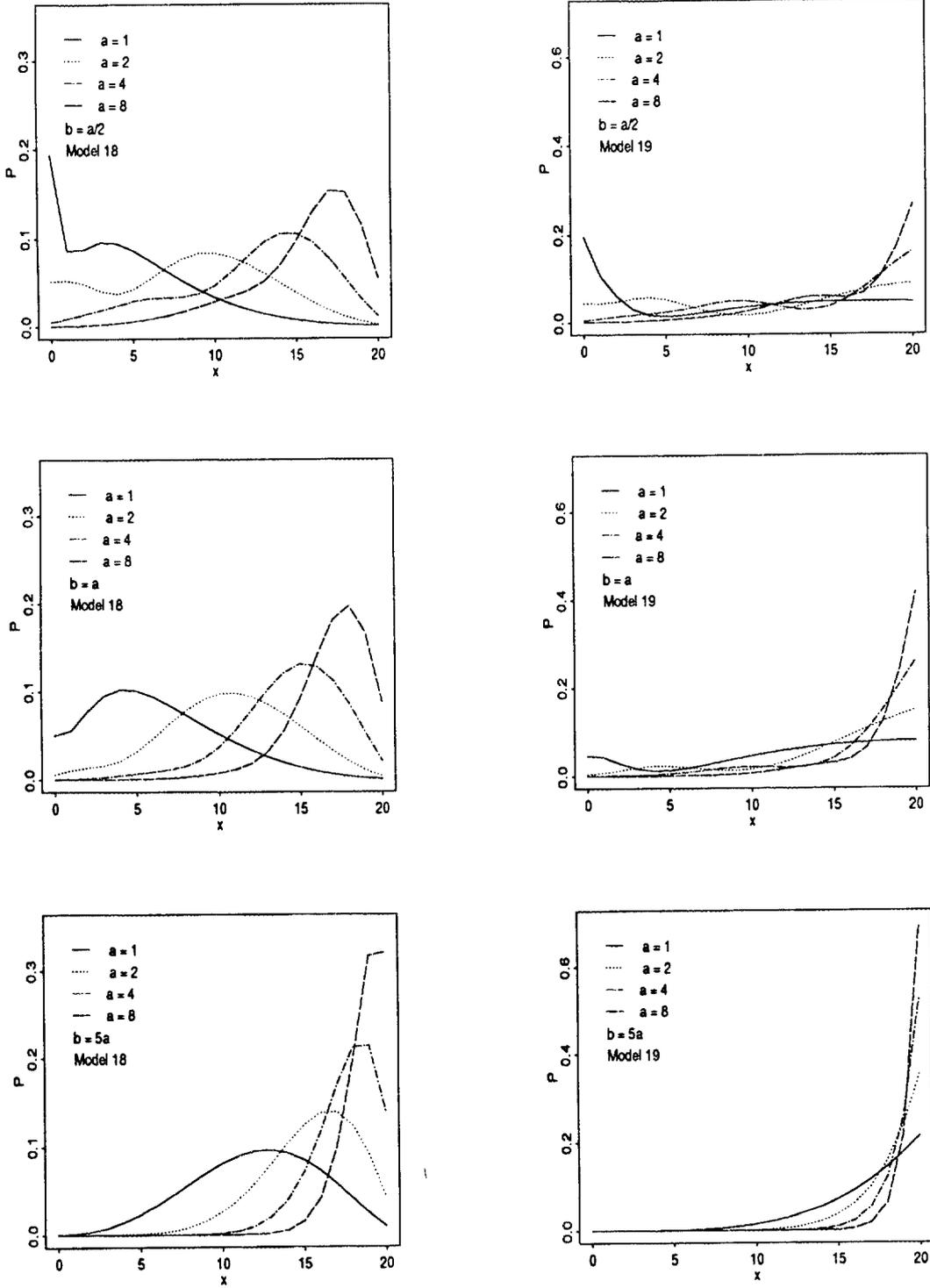


Figure 38: Models 18 and 19

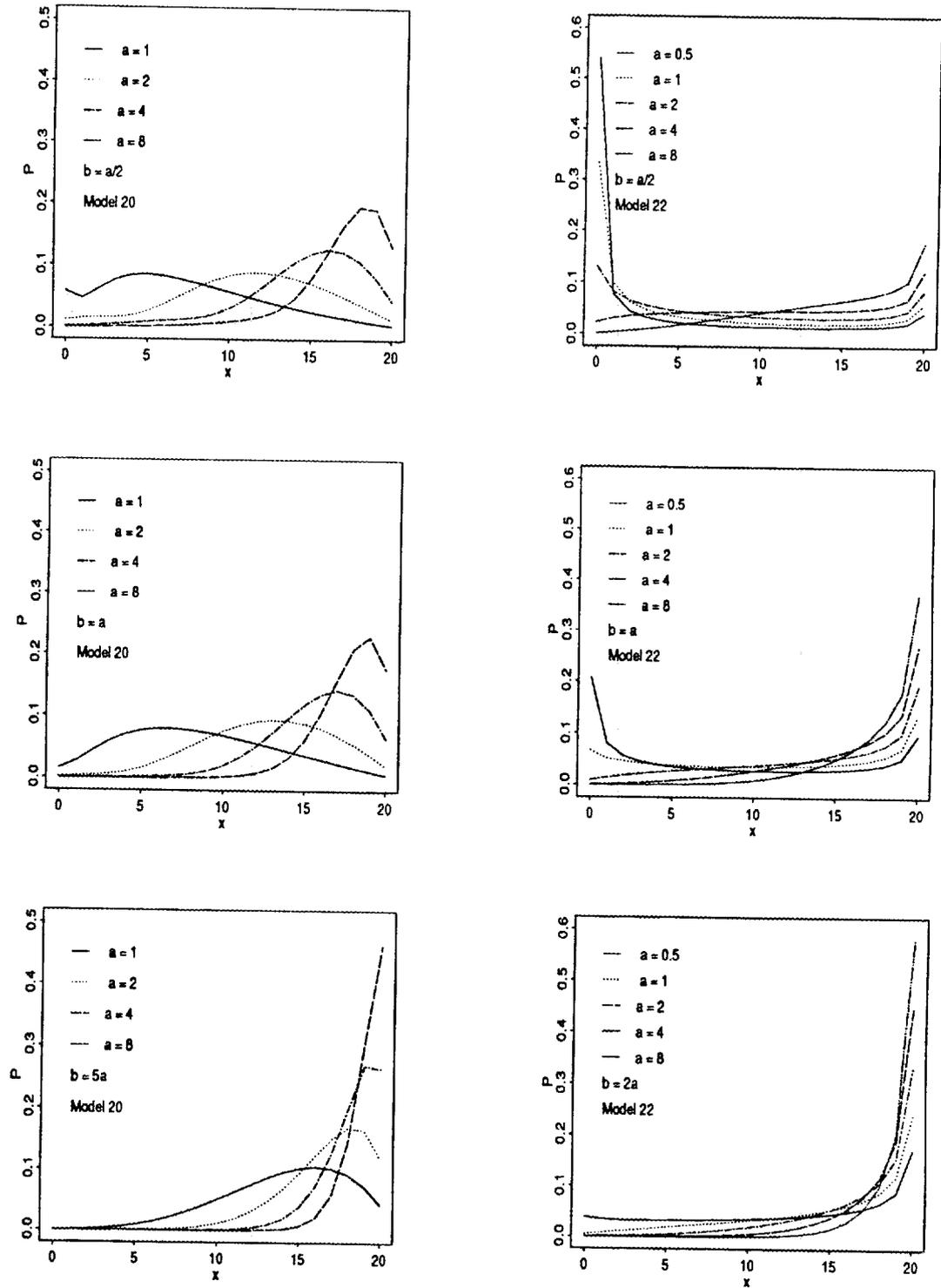


Figure 39: Models 20 and 22

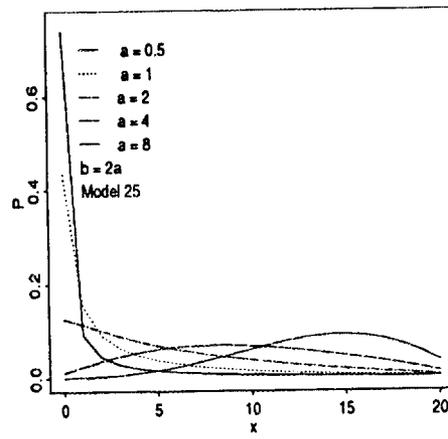
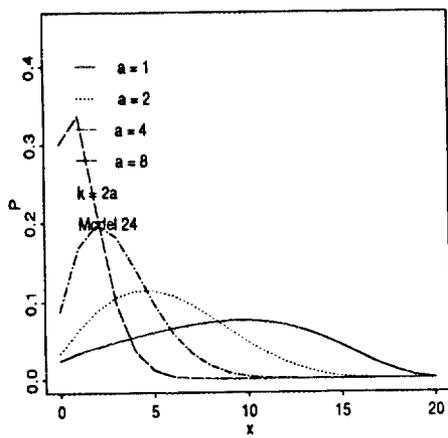
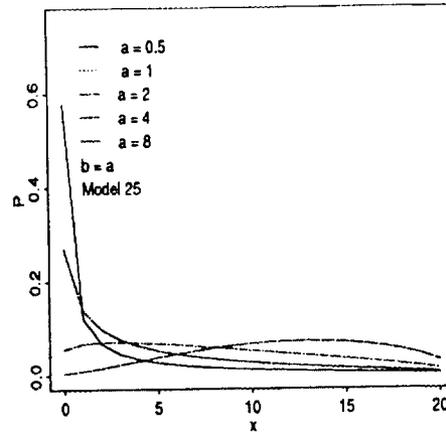
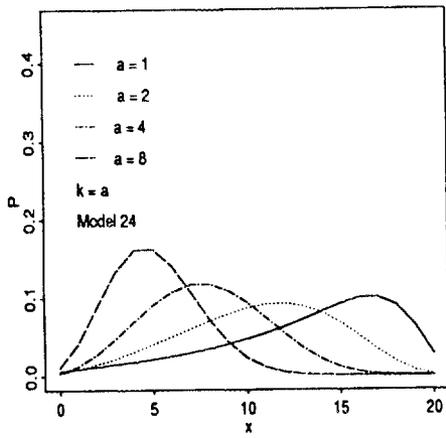
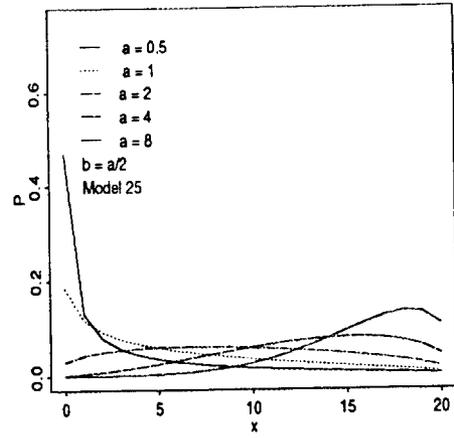
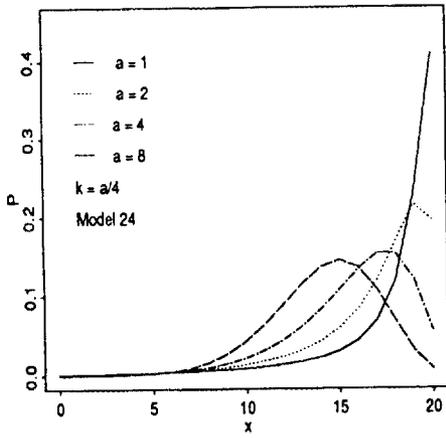


Figure 40: Models 24 and 25

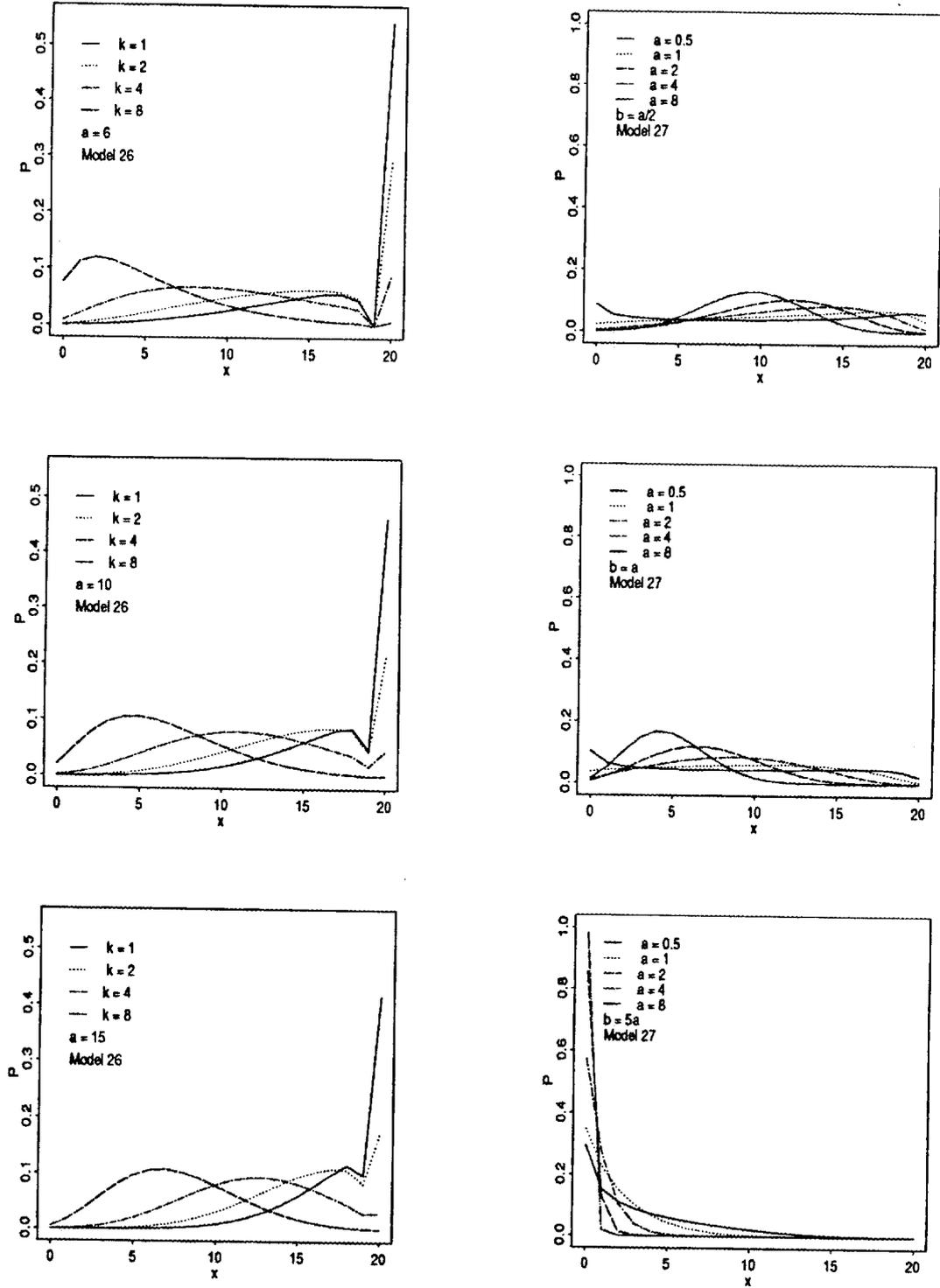


Figure 41: Models 26 and 27



**INTERNAL DISTRIBUTION**

- |                      |                                      |
|----------------------|--------------------------------------|
| 1. B. R. Appleton    | 20. G. Ostrouhov                     |
| 2. C. K. Bayne       | 21-25. S. A. Raby                    |
| 3. J. J. Beauchamp   | 26. R. L. Schmoyer                   |
| 4-8. K. O. Bowman    | 27-31. R. F. Sincovec                |
| 9. C. C. Brandt      | 32-36. R. C. Ward                    |
| 10-11. T. S. Darland | 37. D. A. Wolf                       |
| 12. D. M. Flanagan   | 38. T. Wright                        |
| 13. E. L. Frome      | 39. Central Research Library         |
| 14. T. L. Hebble     | 40. ORNL Patent Office               |
| 15. E. Leach         | 41. K-25 Plant Library               |
| 16. S. Lenhart       | 42. Y-12 Technical Library           |
| 17. Shaw-Pin Miaou   | 43. Laboratory Records - RC          |
| 18. T. J. Mitchell   | 44-45. Laboratory Records Department |
| 19. M. D. Morris     |                                      |

**EXTERNAL DISTRIBUTION**

46. Shunichi Amari, Faculty of Engineering, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan
47. George A. Barnard, Mill House, 54 Hurst Green, Brightlingsea, Colchester, Essex C07 0EH, United Kingdom
48. Roger W. Brockett, Division of Applied Sciences, Harvard University, Cambridge, MA 02138
- 49-53. Karl Broman, 734 Alcatraz Ave., No.2, Oakland, CA 94609
54. Robert H. Byers Jr, Statistics and Data Branch, Centers for Disease Control, HIV/AIDS Division, 1509 Sheridan Walk NE, Atlanta, GA 30324
55. John Cavallini, Acting Director, Scientific Computing Staff, Applied Mathematical Sciences, Office of Energy Research, U.S. Department of Energy, Washington, DC 20585
56. Gary A. Churchill, Plant Bleeding and Biometry Department, Cornell University, 337 Warren Hall, Ithaca, NY 14853
57. P. C. Consul, Mathematics and Statistics Department, University of Calgary, Calgary, T2N 1N4, Canada
58. Edwin L. Crow, U.S. Department of Commerce, Boulder Laboratories ITS.N3, 325 Broadway, Boulder, CO 80303-3328
59. John J. Dorning, Department of Nuclear Engineering Physics, Thornton Hall, McCormick Road, University of Virginia, Charlottesville, VA 22901
60. James B. Douglas, Statistics Department, University of New South Wales, Box 1, Kensington, NSW 2033, Australia

61. Max Eisenberg, Center for Indoor Air Research, 1099 Winterson Road, Suite 280, Linthicum, MD 21090
62. Jerome Friedman, Department of Statistics, Sequoia Hall, Stanford University, Stanford, CA 94305
63. John Gurland, Statistics Department, University of Wisconsin, 12210 Dayton St, Madison, WI 53706
64. Kermit Hutcheson, Statistics Department, University of Georgia, Athens, GA 30602
65. N. L. Johnson, Statistics Department, University of North Carolina, Chapel Hill, NC 27599-3260
- 66-70. M. A. Kastenbaum, 16933 Timberlakes Dr., S.W., Ft. Myers, FL 33908
71. A. D. Kemp, Statistics Department, University of St Andrews North Haugh, St Andrew KY16 9SS, Scotland
72. Samuel Kotz, Management Sciences and Statistics, University of Maryland, College Park, MD 20742
73. James E. Leiss, Rt. 2, Box 142C, Broadway, VA 22815
74. Neville Moray, Department of Mechanical and Industrial Engineering, University of Illinois, 1206 West Green Street, Urbana, IL 61801
75. T. Okuno, 17-8 Minami-cho 4, Tanashi, Tokyo, 188A, Japan
76. J. K. Ord, Management Science Department, Pennsylvania State University, 303 Bean Business Administration Bldg., University Park, PA 16802
77. Jerome Sacks, National Institute of Statistical Sciences, P.O.Box 14162, Research Triangle Park, NC 27709-4162
- 78-82. L. R. Shenton, University of Georgia, Athens, GA 30602
83. Daniel L. Solomon, Department of Statistics, North Carolina State University, P.O.Box 5457, Raleigh, NC 27650
84. Alan Stuart, 82 Somerset Road, Wimbledon, London SW19 5JX, United Kingdom
85. Werner Stuetzle, Department of Statistics, GN-22, University of Washington, Seattle, WA 98195
86. Ray A. Waller, S-1, Statistics, Los Alamos National Laboratory, P.O.Box 1663, Los Alamos, NM 87545
87. Mary F. Wheeler, Rice University, Dept. of Mathematical Sciences, P.O.Box 1892, Houston, TX 77001
88. Alice Zeiger, Center for Indoor Air Research, 1099 Winterson Road, Suite 280, Linthicum, MD 21090
89. Office of Assistant Manager for Energy Research and Development, U.S. Department of Energy, Oak Ridge Operations Office, P.O.Box 2001, Oak Ridge, TN 37831-8600
- 90-99. Office of Scientific and Technical Information, P.O.Box 62, Oak Ridge, TN 37830