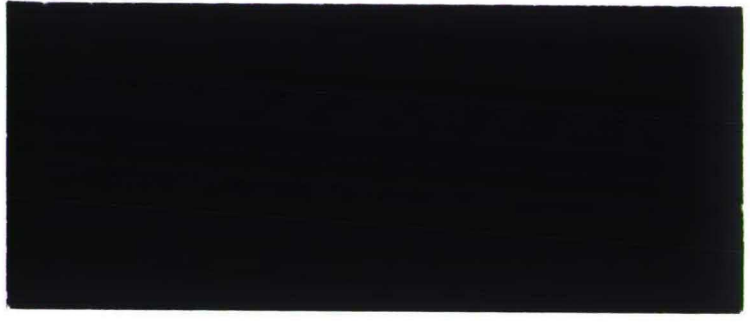
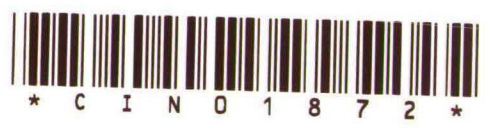


ECO  
CBM  
RR  
7626  
1992  
578

UNIVERSITY  
COLLEGE  
UNIVERSITEIT  
BRABANT

POSTBOX 90153  
5000 LE TILBURG  
THE NETHERLANDS



R41  
Statistical Distribution

DEPARTMENT OF ECONOMICS  
RESEARCH MEMORANDUM



**CHARACTERIZING DISTRIBUTIONS BY  
QUANTILE MEASURES**

R.Th.A. Wagemakers, J.J.A. Moors,  
M.J.B.T. Janssens

**FEW 578**

Communicated by Dr. R.M.J. Heuts

## CHARACTERIZING DISTRIBUTIONS BY QUANTILE MEASURES

R.Th.A. Wagemakers\*

J.J.A. Moors\*

M.J.B.T. Janssens\*

**Abstract.** Modelling an empirical distribution by means of a simple theoretical distribution is an interesting issue in applied statistics. A reasonable first step in this modelling process is to demand that measures for location, dispersion, skewness and kurtosis for the two distributions coincide. Up to now, the four measures used hereby were based on moments.

In this paper measures are considered which are based on quantiles. Of course the four values of these quantile measures do not uniquely determine the modelling distribution. They do, however, within specific systems of distributions, like Pearson's or Johnson's.

This opens the possibility of modelling - within a specific system - an empirical distribution by means of quantile measures. Since moment-based measures are sensitive for outliers, this approach may lead to a better fit.

\* Tilburg University, P.O. Box 90153, 5000 LE Tilburg, Netherlands.

### 1. A quantile measure for kurtosis

Consider a random variable  $\underline{x}$  with mean  $\mu = E(\underline{x})$  and central moments

$$\mu_i = E(\underline{x} - \mu)^i, \quad i = 2, 3, \dots$$

The (very familiar) moment-based measures for location, dispersion, skewness and kurtosis now are

- the mean  $\mu$
- the variance  $\mu_2$
- the third standardized moment  $\beta_1 = \mu_3 / \mu_2^{3/2}$
- the fourth standardized moment  $\beta_2 = \mu_4 / \mu_2^2$

They all exist provided  $E(\underline{x}^4) < \infty$ .

For the first three measures quantile-based alternatives are well-known. Defining quartiles  $Q_i$  by

$$P(\underline{x} < Q_i) \leq i/4, \quad P(\underline{x} > Q_i) \leq 1 - i/4$$

for  $i = 1, 2, 3$ , they are given by

- the median  $Q = Q_2$
- the half interquartile range  $R = (Q_3 - Q_1)/2$
- Bowley's skewness measure  $S = (Q_3 - 2Q_2 + Q_1)/(Q_3 - Q_1)$

provided that  $Q_3 \neq Q_1$ . Moors (1986, 1988) presented a new interpretation of kurtosis as well as a quantile-based alternative for  $\beta_2$ . Define octiles  $E_i$  by

$$P(\underline{x} < E_i) \leq i/8, \quad P(\underline{x} > E_i) \leq 1 - i/8$$

for  $i = 1, 2, \dots, 7$ . Then the quantile measure T for kurtosis reads

$$T = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2}$$

provided that  $E_6 \neq E_2$ . Note that T is much less sensitive for outliers than  $\beta_2$ ; it can be calculated by graphical means. Furthermore, T exists even for distributions without finite moments; e.g. T = 2 for the Cauchy distribution.

The quartet (Q,R,S,T) can be seen as an alternative to  $(\mu, \mu_2, \beta_1, \beta_2)$ . Like  $\beta_1$  and  $\beta_2$ , S and T remain unchanged under linear transformations: these four quantities are location-scale-invariant. This is the main reason why in the sequel attention is focussed on the pair (S,T).

## 2. The Pearson system of distributions

The Pearson system of distributions is based on the following differential equation:

$$\frac{d \log f(x)}{dx} = \frac{x}{B_0 + B_1 x + B_2 x^2}$$

Solutions f are densities within the Pearson system. These solutions depend on the zeros of the denominator or - more specifically - on the quantity

$$K = B_1^2 / (4B_0 B_2)$$

For  $K < 0$ ,  $0 < K < 1$ ,  $K > 1$  three main types of distributions arise; the limiting cases  $K = 0$ ,  $K = 1$  or  $K \rightarrow \infty$  lead to transition types. Table 1 shows the details.

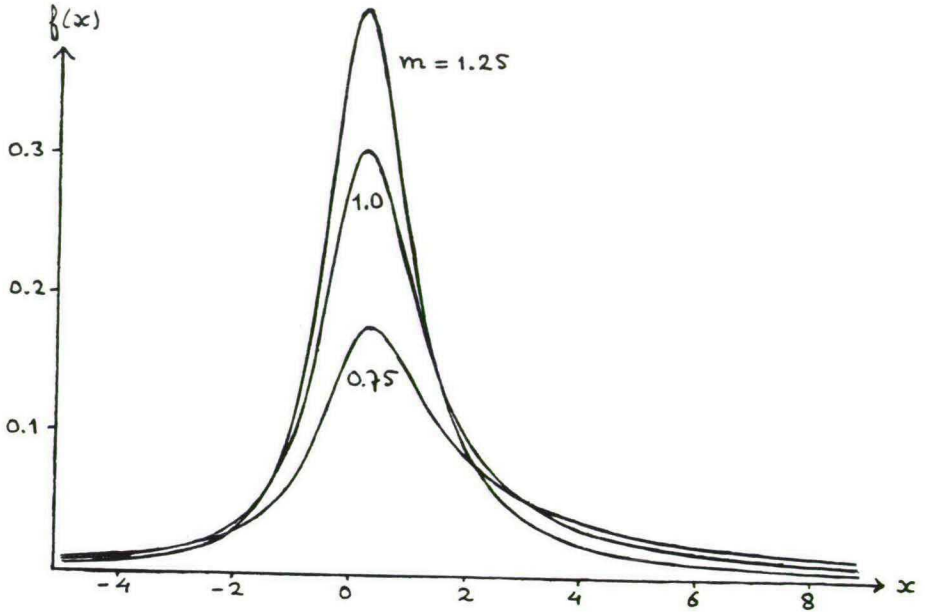
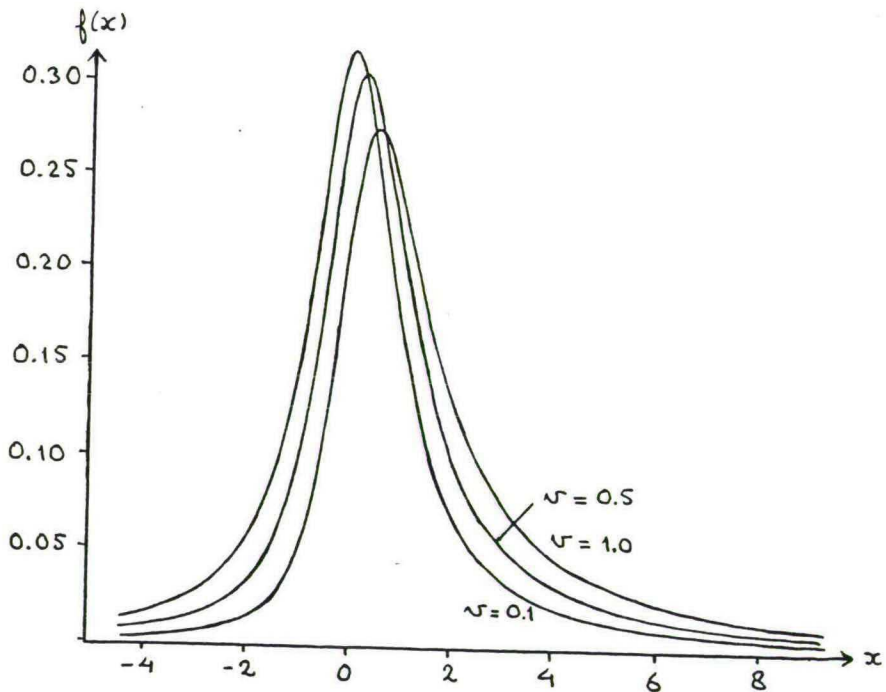
Table 1. Outline of the Pearson system.

|                        | Name          | Type | Density*                         | Range          | Parameters                       |
|------------------------|---------------|------|----------------------------------|----------------|----------------------------------|
| $K < 0$                | Beta 1        | I    | $x^{p-1}(1-x)^{q-1}$             | $[0,1]$        | $p, q > 0$                       |
| $K = 0$                | Student       | VII  | $(1+x^2/n)^{-(n+1)/2}$           | $\mathbb{R}$   | $n > 0$                          |
| $0 < K < 1$            | Arctan        | IV   | $(1+x^2)^{-m} \exp[v \arctan x]$ | $\mathbb{R}$   | $m > 1/2,$<br>$v \in \mathbb{R}$ |
| $K = 1$                | Inverse gamma | V    | $x^{-(p+1)} e^{-1/x}$            | $\mathbb{R}^+$ | $p > 0$                          |
| $K > 1$                | Beta 2        | VI   | $x^{p-1}/(x+1)^{p+q}$            | $\mathbb{R}^+$ | $p, q > 0$                       |
| $K \rightarrow \infty$ | Gamma         | III  | $x^{p-1} e^{-x}$                 | $\mathbb{R}^+$ | $p > 0$                          |

\* up to normalizing constant.

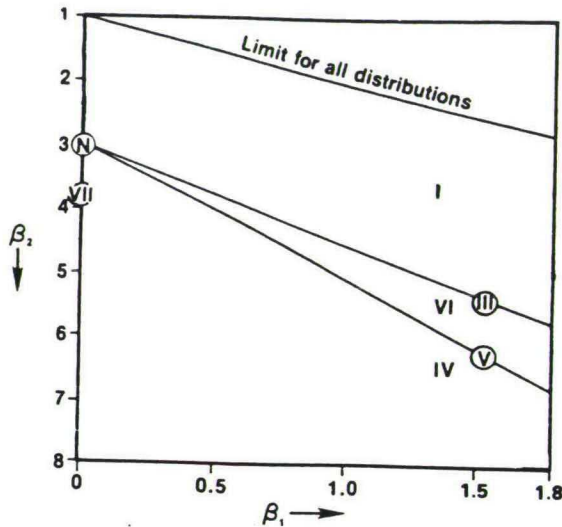
The column 'Type' contains the Roman numbers originally used by Pearson to indicate the different classes of distributions. (The missing type II consists of the symmetrical Beta 1 distributions.) Location-scale parameters have been deleted from the densities, as well as normalizing constants. See for details about all this Stuart & Ord (1987), p. 210 ff.

Since the Arctan distributions are relatively unfamiliar, Figures 1 and 2 show some densities for type IV.

Figure 1. Densities of Pearson type IV;  $v = 0.5$ .Figure 2. Densities of Pearson type IV;  $m = 1$ .

For our purposes, the main property of the Pearson system is that any (location-scale-free) distribution has a unique pair of values for the measures  $\beta_1$  and  $\beta_2$ . In other words, there is a one-one relation between the distributions in Table 1 and points in the  $(\beta_1, \beta_2)$ -plane. Figure 3 shows this relation; the (symmetric) half-plane with  $\beta_1 < 0$  is omitted. Compare Stuart & Ord (1987), p. 211.

Figure 3. The  $(\beta_1, \beta_2)$ -plane for the Pearson system.



The main types appear to occupy separate parts of the plane. Transition type III corresponds to the straight line  $2\beta_2 - 3\beta_1 = 6$ ; the set of type V distributions is slightly curved. The Pearson system leaves unoccupied the upper righthand corner above the line  $\beta_2 - \beta_1 = 1$ .

In summary: all distributions in the Pearson system can be characterized by the quartet  $(\mu_1, \mu_2, \beta_1, \beta_2)$ . Hence, the empirical counterpart  $(\bar{x}, s^2, b_1, b_2)$  of this quartet corresponds with exactly one distribution within the Pearson system. This distribution can be taken as a simple model for the empirical distribution, based on moment measures. Note that from the quartet the parameters of the corresponding Pearson distribution can be found analytically.



### 3. Characterizing the Pearson system by quantile measures

In this section the behaviour is investigated of the pair (S,T) for distributions in the Pearson system. First of all, convergence of distributions implies convergence of the pair (S,T). In particular, the (S,T)-values of a transition type arise as limits of the (S,T)-values of main type distributions. This statement will be proved here for one limiting case only: I  $\rightarrow$  III.

Starting point is the following limiting property of the gamma function  $\Gamma$ :

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n+\rho)}{n^\rho \Gamma(n)} = 1$$

which can be proved by means of Stirling's formula. Let the distribution of a random variable  $\underline{x}$  be denoted by  $\mathcal{L}(\underline{x})$ .

**Lemma 1.** If  $\underline{x}_n \sim \text{Be}(\rho, n)$ ,

$$\mathcal{L}(n\underline{x}_n) \rightarrow \Gamma(1, \rho)$$

holds for  $n \rightarrow \infty$ .

**Proof.** Let  $q_n$  be the density of  $n\underline{x}_n$  and  $p$  that of  $\Gamma(1, \rho)$ ; then it is sufficient to show that  $q_n \rightarrow p$  pointwise if  $n \rightarrow \infty$ . Now  $\underline{x}_n$  has density

$$B(\rho, n)^{-1} x_n^{\rho-1} (1-x_n)^{n-1},$$

where  $B(\rho, n) = \Gamma(\rho)\Gamma(n)/\Gamma(\rho+n)$ . For  $y = n\underline{x}_n$  it follows:

$$\begin{aligned} q_n(y) &= \frac{1}{nB(\rho, n)} x_n^{\rho-1} (1-x_n)^{n-1}, \quad 0 < x_n < 1, \quad \text{with } x_n = y/n, \\ &= \frac{1}{nB(\rho, n)} \left(\frac{y}{n}\right)^{\rho-1} \left(1-\frac{y}{n}\right)^{n-1}, \quad 0 < y < n, \\ &= \frac{1}{n^\rho B(\rho, n)} y^{\rho-1} \left(1-\frac{y}{n}\right)^{n-1}. \end{aligned}$$

Now the limits

$$\lim_{n \rightarrow \infty} (1-y/n)^n = e^{-y}, \quad \lim_{n \rightarrow \infty} \frac{1}{n^{\rho} B(\rho, n)} = \frac{1}{\Gamma(\rho)}$$

imply

$$\lim_{n \rightarrow \infty} q_n(y) = \frac{1}{\Gamma(\rho)} y^{\rho-1} e^{-y} = p(y)$$

which proves the lemma.  $\square$

**Theorem 1.** Let  $(S_n, T_n)$  and  $(S_0, T_0)$  denote the quantile measures of skewness and kurtosis for  $Be(\rho, n)$  and  $\Gamma(1, \rho)$ , respectively. Then

$$\lim_{n \rightarrow \infty} S_n = S_0, \quad \lim_{n \rightarrow \infty} T_n = T_0$$

**Proof.** Since  $S$  and  $T$  are invariant under linear transformations, they are identical for  $\underline{x}_n$  and  $n\underline{x}_n$ . Now, the theorem is an immediate result of the Lemma.  $\square$

So the conclusion is, that smooth transitions between the various types exist in the  $(S, T)$ -plane - just as in the  $(\beta_1, \beta_2)$ -plane.

Quantiles for the type I, III, V and VI can be found directly by means of the statistical computer package SAS, while for VII Smirnov (1961) was used. For type IV a special program was written which uses numerical integration. This led to outcomes that differed slightly from the values in Johnson et al (1963). Hence, another program was written, which confirmed our previous results. Table 2 is a brief abstract from the extensive results in Wagemakers (1991).

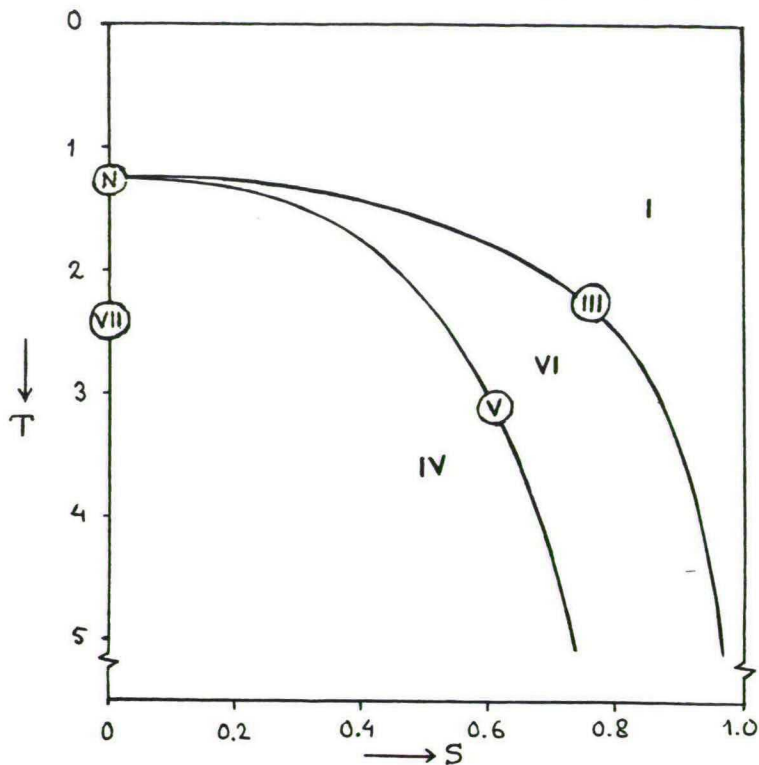
Table 2. Octiles and (S,T)-values for the Pearson system.

| Type | $p_1$ | $p_2$ | $E_1$  | $E_2$  | $E_3$  | $E_4$ | $E_5$ | $E_6$ | $E_7$ | S     | T     |
|------|-------|-------|--------|--------|--------|-------|-------|-------|-------|-------|-------|
| I    | 0.30  | 0.3   | 0.007  | 0.068  | 0.236  | 0.500 | 0.764 | 0.932 | 0.993 | 0.000 | 0.529 |
| I    | 0.30  | 1.2   | 0.001  | 0.008  | 0.030  | 0.079 | 0.168 | 0.317 | 0.557 | 0.541 | 1.349 |
| I    | 0.30  | 2.4   | 0.000  | 0.003  | 0.013  | 0.035 | 0.076 | 0.153 | 0.300 | 0.579 | 1.581 |
| I    | 1     | 1     | 0.125  | 0.250  | 0.375  | 0.500 | 0.625 | 0.750 | 0.875 | 0.000 | 1.000 |
| I    | 1     | 4     | 0.033  | 0.069  | 0.111  | 0.159 | 0.218 | 0.293 | 0.405 | 0.197 | 1.190 |
| I    | 1     | 8     | 0.017  | 0.035  | 0.057  | 0.083 | 0.115 | 0.159 | 0.229 | 0.230 | 1.244 |
| III  | 0.30  |       | 0.001  | 0.007  | 0.027  | 0.073 | 0.165 | 0.343 | 0.740 | 0.606 | 2.004 |
| III  | 0.75  |       | 0.058  | 0.153  | 0.283  | 0.454 | 0.688 | 1.034 | 1.650 | 0.317 | 1.347 |
| III  | 1     |       | 0.134  | 0.288  | 0.470  | 0.693 | 0.981 | 1.386 | 2.079 | 0.262 | 1.306 |
| III  | 5     |       | 2.617  | 3.369  | 4.020  | 4.671 | 5.390 | 6.274 | 7.599 | 0.104 | 1.243 |
| IV   | 0.70  | 0.1   | -12.11 | -2.070 | -0.479 | 0.347 | 1.444 | 4.542 | 26.37 | 0.269 | 5.530 |
| IV   | 0.70  | 0.5   | -1.862 | -0.014 | 0.869  | 2.150 | 5.020 | 14.62 | 84.46 | 0.704 | 5.617 |
| IV   | 0.70  | 1.0   | 0.222  | 1.298  | 2.761  | 5.544 | 12.26 | 35.12 | 202.1 | 0.749 | 5.687 |
| IV   | 1     | 0.1   | -2.065 | -0.822 | -0.282 | 0.124 | 0.553 | 1.193 | 2.802 | 0.061 | 2.001 |
| IV   | 1     | 0.5   | -1.013 | -0.237 | 0.207  | 0.630 | 1.175 | 2.117 | 4.726 | 0.263 | 2.027 |
| IV   | 1     | 1.0   | -0.249 | 0.316  | 0.775  | 1.313 | 2.100 | 3.566 | 7.808 | 0.387 | 2.072 |
| V    | 0.30  |       | 1.352  | 2.916  | 6.074  | 13.67 | 36.94 | 144.9 | 1467  | 0.849 | 10.11 |
| V    | 0.75  |       | 0.606  | 0.967  | 1.453  | 2.202 | 3.538 | 6.519 | 17.33 | 0.555 | 2.637 |
| V    | 1     |       | 0.481  | 0.721  | 1.020  | 1.443 | 2.128 | 3.476 | 7.489 | 0.476 | 2.142 |
| V    | 5     |       | 0.132  | 0.159  | 0.186  | 0.214 | 0.249 | 0.297 | 0.382 | 0.204 | 1.362 |
| VI   | 0.30  | 0.3   | 0.007  | 0.073  | 0.308  | 1.000 | 3.244 | 13.79 | 143.1 | 0.865 | 10.22 |
| VI   | 0.30  | 1.2   | 0.001  | 0.008  | 0.031  | 0.086 | 0.202 | 0.465 | 1.255 | 0.660 | 2.370 |
| VI   | 0.30  | 2.4   | 0.000  | 0.003  | 0.013  | 0.036 | 0.083 | 0.180 | 0.428 | 0.630 | 2.026 |
| VI   | 1     | 1     | 0.143  | 0.333  | 0.600  | 1.000 | 1.667 | 3.000 | 7.000 | 0.500 | 2.171 |
| VI   | 1     | 4     | 0.034  | 0.075  | 0.125  | 0.189 | 0.278 | 0.414 | 0.682 | 0.325 | 1.456 |
| VI   | 1     | 8     | 0.017  | 0.037  | 0.061  | 0.091 | 0.130 | 0.189 | 0.297 | 0.294 | 1.377 |
| VII  | 0.30  |       | -1.486 | -0.923 | -0.439 | 0.000 | 0.439 | 0.923 | 1.486 | 0.000 | 1.135 |
| VII  | 0.75  |       | -9.277 | -2.197 | -0.733 | 0.000 | 0.733 | 2.197 | 9.277 | 0.000 | 3.888 |
| VII  | 1     |       | -2.414 | -1.000 | -0.414 | 0.000 | 0.414 | 1.000 | 2.414 | 0.000 | 2.000 |
| VII  | 1.5   |       | -1.134 | -0.577 | -0.258 | 0.000 | 0.258 | 0.577 | 1.134 | 0.000 | 1.517 |

|       | I   | III    | IV  | V      | VI  | VII |
|-------|-----|--------|-----|--------|-----|-----|
| $p_1$ | $p$ | $\rho$ | $m$ | $\rho$ | $p$ | $n$ |
| $p_2$ | $q$ |        | $v$ |        | $q$ |     |

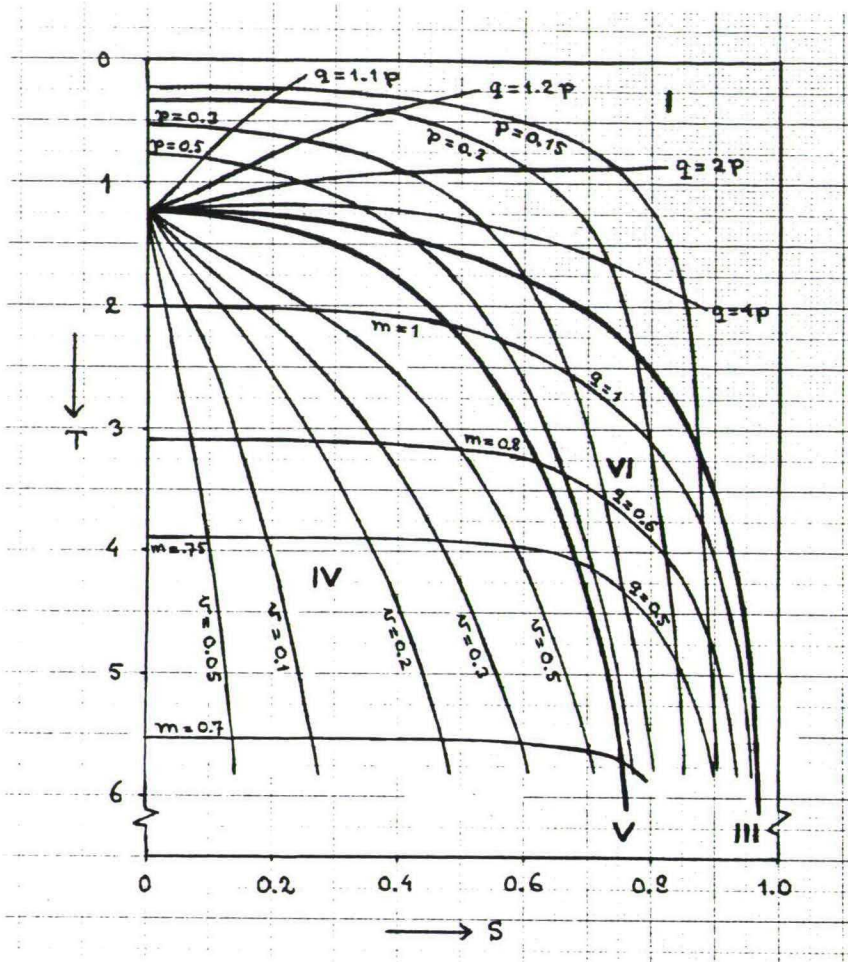
There appears to be one-one relation between the (location-scale free) Pearson distributions and pairs of (S,T)-values. So, just like the  $(\beta_1, \beta_2)$ -plane, the (S,T)-plane is subdivided into separate sets corresponding to the main types; the demarcation lines are given by the transition types. See Figure 4; compare it to Figure 3. The half-plane with  $S < 0$  has been omitted.

Figure 4. The (S,T)-plane for the Pearson system.



Since within a given Pearson type the location-scale parameter is uniquely determined by  $(Q,R)$ , all distributions in the Pearson system can be characterized by the quartet  $(Q,R,S,T)$ . Hence, the empirical counterpart  $(q,r,s,t)$  of this quartet determines exactly one distribution within the Pearson system. Again, this gives a simple model for the empirical distribution, now based on quantile measures. However, the parameters of the corresponding Pearson model have to be found numerically. This can be done by trial-and-error, using the programs mentioned above. Another possibility is to develop a nomogram from which for given  $(S,T)$ -values the corresponding Pearson distribution can be read. In Figure 5 such a nomogram is sketched; of course, to attain numerical accuracy, a much more detailed nomogram is necessary.

Figure 5. Sketch of nomogram for the Pearson system.



In principle, there now are two ways to find a model within the Pearson system for a given frequency distribution. An interesting question is which model fits best; this question is discussed in some more detail in Section 6.

#### 4. The Johnson system

Another subdivision of the  $(\beta_1, \beta_2)$ -plane was obtained by Johnson (1949). His system of distributions consists of three different types of transformations of a standard normal variable  $z$ . Using

$$\underline{x} = (z - \gamma) / \delta$$

for given constants  $\gamma$  and  $\delta$ , these transformations are

$$y = \varphi_L(\underline{x}) = \exp(\underline{x})$$

$$y = \varphi_B(\underline{x}) = \exp(\underline{x}) / [1 + \exp(\underline{x})]$$

$$y = \varphi_U(\underline{x}) = [\exp(\underline{x}) - \exp(-\underline{x})] / 2$$

Details of the resulting distributions are shown in Table 3.

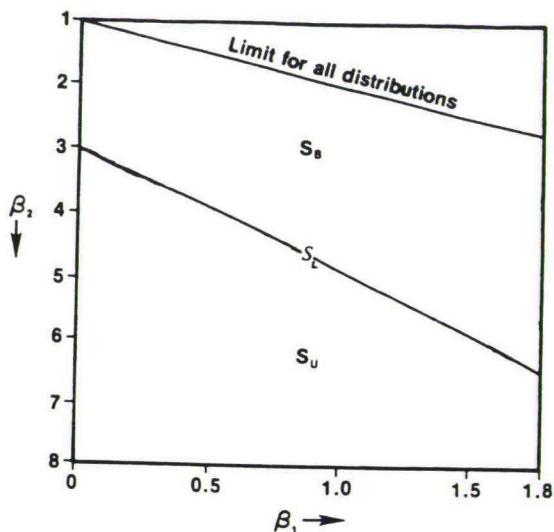
Table 3. Outline of the Johnson system.

| Name            | Type Density  | Range          | Parameters  |
|-----------------|---|----------------|---|
| Lognormal       | $S_L \frac{\delta}{\sqrt{2\pi}} \frac{1}{y} \exp[-(\delta \log y)^2 / 2]$                                       | $\mathbb{R}^+$ | $\delta \in \mathbb{R}^+$                             |
| Bounded range   | $S_B \frac{\delta}{\sqrt{2\pi}} \frac{1}{y(1-y)} \exp[-\{\gamma + \delta \log(\frac{y}{1-y})\}^2 / 2]$          | $[0, 1]$       | $\gamma \in \mathbb{R},$<br>$\delta \in \mathbb{R}^+$ |
| Unbounded range | $S_U \frac{\delta}{\sqrt{2\pi}} \frac{1}{\sqrt{1+y^2}} \exp[-\{\gamma + \delta \log(y + \sqrt{1+y^2})\}^2 / 2]$ | $\mathbb{R}$   | $\gamma \in \mathbb{R},$<br>$\delta \in \mathbb{R}^+$ |

For the lognormal distributions the - location - parameter  $\gamma$  has been deleted.

Like for the Pearson system, any distribution of the Johnson system has a unique pair of values for the measures  $\beta_1$  and  $\beta_2$ . Figure 6 shows how the  $(\beta_1, \beta_2)$ -plane is split by the curve  $S_L$  in separate parts  $S_B$  and  $S_U$ . See for details Stuart & Ord (1987), p. 234 ff. The half-plane with  $\beta_1 < 0$  has been omitted.

Figure 6. The  $(\beta_1, \beta_2)$ -plane for the Johnson system.



The quartet  $(\bar{x}, s^2, b_1, b_2)$  of empirical measures determines a unique distribution within the Johnson system, as was the case for the Pearson system.

##### 5. Characterization of the Johnson system by quantile measures

The  $(S, T)$ -values of the transition type  $S_L$  arise as limits of the  $(S, T)$ -values of either  $S_U$  or  $S_B$  type distributions. E.g. for the latter type this is implied by the following property: if  $\underline{x}_n \sim S_B(n, \delta)$ ,  $\mathcal{L}(\underline{x}_n) \rightarrow S_L(\delta)$  for  $n \rightarrow \infty$ . So, in the  $(S, T)$ -plane the set defined by  $S_L$  is a smooth transition between the sets corresponding to  $S_U$  and  $S_B$ .

Since Johnson distributions are transformations of the standard normal, octiles are easily calculated. Table 4 gives a brief summary of the extensive tables in Wagemakers (1991).

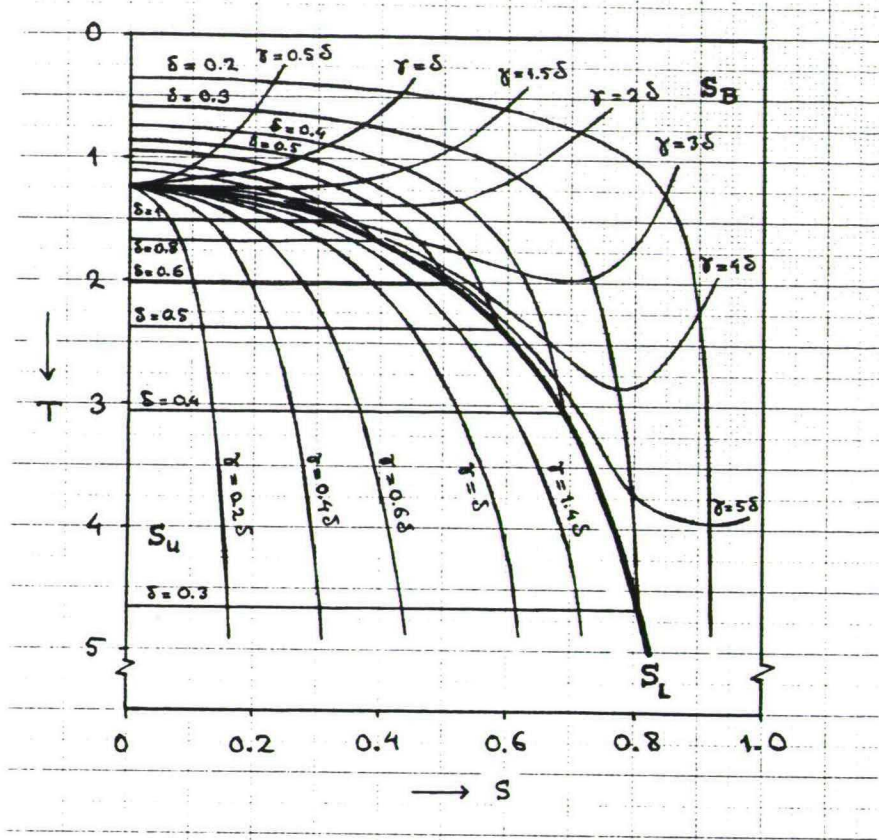
Table 4. Octiles and (S,T)-values for the Johnson system.

| Type  | $\delta$ | $\gamma$ | $E_1$  | $E_2$  | $E_3$  | $E_4$ | $E_5$ | $E_6$ | $E_7$ | S     | T     |
|-------|----------|----------|--------|--------|--------|-------|-------|-------|-------|-------|-------|
| $S_L$ | 0.30     |          | 0.022  | 0.106  | 0.346  | 1.000 | 2.893 | 9.472 | 46.27 | 0.809 | 4.666 |
| $S_L$ | 0.75     |          | 0.216  | 0.407  | 0.654  | 1.000 | 1.529 | 2.458 | 4.636 | 0.422 | 1.728 |
| $S_L$ | 1        |          | 0.317  | 0.509  | 0.727  | 1.000 | 1.375 | 1.963 | 3.159 | 0.328 | 1.510 |
| $S_L$ | 5        |          | 0.795  | 0.874  | 0.938  | 1.000 | 1.066 | 1.144 | 1.259 | 0.067 | 1.244 |
| $S_B$ | 0.30     | 0        | 0.021  | 0.096  | 0.257  | 0.500 | 0.743 | 0.905 | 0.979 | 0.000 | 0.583 |
| $S_B$ | 0.30     | 0.30     | 0.008  | 0.037  | 0.113  | 0.269 | 0.516 | 0.777 | 0.945 | 0.374 | 0.722 |
| $S_B$ | 0.30     | 1.20     | 0.000  | 0.002  | 0.006  | 0.018 | 0.050 | 0.148 | 0.459 | 0.780 | 2.840 |
| $S_B$ | 0.30     | 2.10     | 0.000  | 0.000  | 0.000  | 0.001 | 0.003 | 0.009 | 0.041 | 0.808 | 4.506 |
| $S_B$ | 1        | 0        | 0.240  | 0.338  | 0.421  | 0.500 | 0.579 | 0.663 | 0.760 | 0.000 | 1.111 |
| $S_B$ | 1        | 1        | 0.104  | 0.158  | 0.211  | 0.269 | 0.336 | 0.419 | 0.538 | 0.150 | 1.179 |
| $S_B$ | 1        | 5        | 0.002  | 0.003  | 0.005  | 0.007 | 0.009 | 0.013 | 0.021 | 0.321 | 1.496 |
| $S_B$ | 1        | 7        | 0.000  | 0.001  | 0.001  | 0.001 | 0.001 | 0.002 | 0.003 | 0.324 | 1.508 |
| $S_U$ | 0.30     | 0.00     | -23.12 | -4.683 | -1.273 | 0.000 | 1.273 | 4.683 | 23.12 | 0.000 | 4.666 |
| $S_U$ | 0.30     | 0.12     | -15.49 | -3.096 | -0.712 | 0.411 | 2.042 | 7.030 | 34.51 | 0.307 | 4.666 |
| $S_U$ | 0.30     | 0.30     | -8.482 | -1.599 | -0.062 | 1.175 | 3.868 | 12.85 | 62.88 | 0.616 | 4.666 |
| $S_U$ | 0.30     | 0.60     | -3.051 | -0.251 | 1.082  | 3.627 | 10.66 | 34.99 | 171.0 | 0.780 | 4.666 |
| $S_U$ | 1        | 0.0      | -1.421 | -0.727 | -0.324 | 0.000 | 0.324 | 0.727 | 1.421 | 0.000 | 1.510 |
| $S_U$ | 1        | -0.5     | -0.697 | -0.175 | 0.182  | 0.521 | 0.913 | 1.464 | 2.508 | 0.150 | 1.510 |
| $S_U$ | 1        | -1.0     | -0.151 | 0.231  | 0.735  | 1.175 | 1.735 | 2.574 | 4.236 | 0.248 | 1.510 |
| $S_U$ | 1        | -2.0     | 0.956  | 1.749  | 2.593  | 3.627 | 5.032 | 7.218 | 11.65 | 0.313 | 1.510 |

Again, from the empirical measures (q,r,s,t) a model can be found within the Johnson system. This may be done numerically or graphically, by means of a nomogram. Figure 7 sketches such a nomogram.



Figure 7. Sketch of nomogram for the Johnson system.



Again, there are two ways to find a suitable model within the Johnson system; an important question is whether the moment-based or the quantile-based approach is better.

## 6. Discussion and further research

In this paper an alternative method was developed to find a suitable model for an empirical frequency distribution within a given system of theoretical distributions. For this class of potential models both Pearson's and Johnson's system of distributions was considered. Our method is based on the four quantile measures

(Q, R, S, T)

for location, dispersion, skewness and kurtosis; all of them can be calculated from the seven octiles.

Attention was concentrated on the behaviour of S and T; our main result is that both in Pearson's and in Johnson's system there is a one-one correspondence between the (location-scale free) distributions and the values of the pair (S,T).

An interesting next question is of course whether this quantile-based method gives a better fit than the classical approach, which is based on the moments

$(\mu, \mu_2, \beta_1, \beta_2)$

As a first step in answering this question, the limit distributions of the empirical measures  $(\underline{s}, \underline{t})$  and  $(\underline{b}_1, \underline{b}_2)$  are being investigated. For the standard normal distribution we obtained the following results:

$$\sqrt{n} \begin{bmatrix} \underline{s} - S \\ \underline{t} - T \end{bmatrix} \xrightarrow{\mathcal{L}} N_2 \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.839 & 0 \\ 0 & 3.153 \end{bmatrix} \right)$$

$$\sqrt{n} \begin{bmatrix} \underline{b}_1 - \beta_1 \\ \underline{b}_2 - \beta_2 \end{bmatrix} \xrightarrow{\mathcal{L}} N_2 \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 & 0 \\ 0 & 24 \end{bmatrix} \right)$$

where  $\xrightarrow{\mathcal{L}}$  denotes convergence in distribution for  $n \rightarrow \infty$ . Note that for  $N(0,1)$

$$(S, T) = (0, 1.233), \quad (\beta_1, \beta_2) = (0, 3)$$

holds. To check the (co)variances a simulation study was made. From  $N(0,1)$  200 random sample of size  $n$  were drawn and for all of them the statistics

$$(s, t), \quad (b_1, b_2)$$

were calculated. From the 200 replicated values the estimated variances and the covariance of each pair was found. Tables 5 and 6 present the results.

**Table 5.** Simulated means and (co)variance of  $(\underline{s}, \underline{t})$ ;  
200 replicated samples from  $N(0,1)$ .

| n        | Simulated value of         |                                  |                     |                     |                                      |
|----------|----------------------------|----------------------------------|---------------------|---------------------|--------------------------------------|
|          | $\sqrt{n}E(\underline{s})$ | $\sqrt{n}E(\underline{t}-1.233)$ | $nV(\underline{s})$ | $nV(\underline{t})$ | $nCov(\underline{s}, \underline{t})$ |
| 50       | -0.224                     | -0.149                           | 1.887               | 3.940               | 0.256                                |
| 100      | 0.103                      | 0.395                            | 1.746               | 3.378               | -0.034                               |
| 200      | -0.039                     | 0.395                            | 1.931               | 3.487               | -0.026                               |
| 2000     | -0.048                     | 0.654                            | 1.524               | 3.815               | -0.283                               |
| $\infty$ | 0                          | 0                                | 1.839               | 3.153               | 0                                    |

**Table 6.** Simulated means and (co)variance of  $(\underline{b}_1, \underline{b}_2)$ ;  
200 replicated samples from  $N(0,1)$ .

| n        | Simulated value of           |                                |                       |                       |  |
|----------|------------------------------|--------------------------------|-----------------------|-----------------------|--|
|          | $\sqrt{n}E(\underline{b}_1)$ | $\sqrt{n}E(\underline{b}_2-3)$ | $nV(\underline{b}_1)$ | $nV(\underline{b}_2)$ | $nCov(\underline{b}_1, \underline{b}_2)$ |
| 50       | -0.240                       | -0.893                         | 5.068                 | 15.099                | -0.068                                   |
| 100      | 0.095                        | -0.840                         | 4.683                 | 17.781                | -0.035                                   |
| 200      | -0.013                       | -0.461                         | 5.453                 | 21.181                | 1.185                                    |
| 2000     | 0.296                        | 0.570                          | 5.180                 | 23.137                | 0.283                                    |
| $\infty$ | 0                            | 0                              | 6                     | 24                    | 0  |

A full report, with much more general results, is in preparation.

The simulation results appear to be in agreement with the theoretical values in the last lines of the two tables. Note that  $(S, T)$  can be estimated with a greater accuracy than  $(\beta_1, \beta_2)$ . Of course, this does not imply that the quantile-based approach is to be preferred. To admit such a conclusion, some measure of fit will have to be chosen and compared for

both methods of modelling. We plan to make such a comparison in due course.

Apart from the Pearson and Johnson systems of distributions, other systems may be taken as the class of potential models. Interesting candidates are the Schmeiser-Deutsch (1978) system of distributions and Burr's system, cf. Stuart & Ord (1987), p. 242. A final question is how to select a suitable system to start with.

#### **Acknowledgement**

We are very grateful to an unknown referee who - by slashing the first draft of this paper - forced us to improve it considerably. Tables 5 and 6 are based on work by Victor Coenen.

## References

- Johnson, N.L. (1949), *Systems of frequency curves generated by methods of translations*, Biometrika 36, 149-176.
- Johnson, N.L., E. Nixon and D.E. Amos (1963), *Table of percentage points of Pearson curves, for given  $\sqrt{\beta_1}$  and  $\beta_2$ , expressed in standard measure*, Biometrika 50, p. 459-498.
- Moors, J.J.A. (1986), *The meaning of kurtosis: Darlington reexamined*, The American Statistician 40, 283-284.
- Moors, J.J.A. (1988), *A quantile alternative for kurtosis*, The Statistician 37, 25-32.
- Schmeiser, B.W. and S.J. Deutsch (1978), *A versatile four parameter family of probability distributions suitable for simulation*, AIEE Transactions 9, 176-181.
- Smirnov, N.V. (1961), *Tables for the distributions and density functions of t-distribution ('Students'-distribution)*, Mathematical Tables Series 16, Pergamon Press.
- Stuart, A and J.K. Ord (1987), *Kendall's advanced theory of statistics*, Griffin London.
- Wagemakers, R.Th.A. (1991), *Fitting frequency distribution to theoretical distributions by means of quantiles*, Tilburg University (in Dutch).

## IN 1991 REEDS VERSCHENEN

- 466 Prof.Dr. Th.C.M.J. van de Klundert - Prof.Dr. A.B.T.M. van Schaik  
Economische groei in Nederland in een internationaal perspectief
- 467 Dr. Sylvester C.W. Eijffinger  
The convergence of monetary policy - Germany and France as an example
- 468 E. Nijssen  
Strategisch gedrag, planning en prestatie. Een inductieve studie binnen de computerbranche
- 469 Anne van den Nouweland, Peter Borm, Guillermo Owen and Stef Tijs  
Cost allocation and communication
- 470 Drs. J. Grazell en Drs. C.H. Veld  
Motieven voor de uitgifte van converteerbare obligatieleningen en warrant-obligatieleningen: een agency-theoretische benadering
- 471 P.C. van Batenburg, J. Kriens, W.M. Lammerts van Bueren and R.H. Veenstra  
Audit Assurance Model and Bayesian Discovery Sampling
- 472 Marcel Kerkhofs  
Identification and Estimation of Household Production Models
- 473 Robert P. Gilles, Guillermo Owen, René van den Brink  
Games with Permission Structures: The Conjunctive Approach
- 474 Jack P.C. Kleijnen  
Sensitivity Analysis of Simulation Experiments: Tutorial on Regression Analysis and Statistical Design
- 475 C.P.M. van Hoesel  
An  $O(n \log n)$  algorithm for the two-machine flow shop problem with controllable machine speeds
- 476 Stephan G. Vanneste  
A Markov Model for Opportunity Maintenance
- 477 F.A. van der Duyn Schouten, M.J.G. van Eijs, R.M.J. Heuts  
Coordinated replenishment systems with discount opportunities
- 478 A. van den Nouweland, J. Potters, S. Tijs and J. Zarzuelo  
Cores and related solution concepts for multi-choice games
- 479 Drs. C.H. Veld  
Warrant pricing: a review of theoretical and empirical research
- 480 E. Nijssen  
De Miles and Snow-typologie: Een exploratieve studie in de meubelbranche
- 481 Harry G. Barkema  
Are managers indeed motivated by their bonuses?

- 482 Jacob C. Engwerda, André C.M. Ran, Arie L. Rijkeboer  
Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation  $X + A^T X^{-1} A = I$
- 483 Peter M. Kort  
A dynamic model of the firm with uncertain earnings and adjustment costs
- 484 Raymond H.J.M. Gradus, Peter M. Kort  
Optimal taxation on profit and pollution within a macroeconomic framework
- 485 René van den Brink, Robert P. Gilles  
Axiomatizations of the Conjunctive Permission Value for Games with Permission Structures
- 486 A.E. Brouwer & W.H. Haemers  
The Gewirtz graph - an exercise in the theory of graph spectra
- 487 Pim Adang, Bertrand Melenberg  
Intratemporal uncertainty in the multi-good life cycle consumption model: motivation and application
- 488 J.H.J. Roemen  
The long term elasticity of the milk supply with respect to the milk price in the Netherlands in the period 1969-1984
- 489 Herbert Hamers  
The Shapley-Entrance Game
- 490 Rezaul Kabir and Theo Vermaelen  
Insider trading restrictions and the stock market
- 491 Piet A. Verheyen  
The economic explanation of the jump of the co-state variable
- 492 Drs. F.L.J.W. Manders en Dr. J.A.C. de Haan  
De organisatorische aspecten bij systeemontwikkeling een beschouwing op besturing en verandering
- 493 Paul C. van Batenburg and J. Kriens  
Applications of statistical methods and techniques to auditing and accounting
- 494 Ruud T. Frambach  
The diffusion of innovations: the influence of supply-side factors
- 495 J.H.J. Roemen  
A decision rule for the (des)investments in the dairy cow stock
- 496 Hans Kremers and Dolf Talman  
An SLSPP-algorithm to compute an equilibrium in an economy with linear production technologies

- 497 L.W.G. Strijbosch and R.M.J. Heuts  
Investigating several alternatives for estimating the compound lead time demand in an (s,Q) inventory model
- 498 Bert Bettonvil and Jack P.C. Kleijnen  
Identifying the important factors in simulation models with many factors
- 499 Drs. H.C.A. Roest, Drs. F.L. Tijssen  
Beheersing van het kwaliteitsperceptieproces bij diensten door middel van keurmerken
- 500 B.B. van der Genugten  
Density of the F-statistic in the linear model with arbitrarily normal distributed errors
- 501 Harry Barkema and Sytse Douma  
The direction, mode and location of corporate expansions
- 502 Gert Nieuwenhuis  
Bridging the gap between a stationary point process and its Palm distribution
- 503 Chris Veld  
Motives for the use of equity-warrants by Dutch companies
- 504 Pieter K. Jagersma  
Een etiologie van horizontale internationale ondernemingsexpansie
- 505 B. Kaper  
On M-functions and their application to input-output models
- 506 A.B.T.M. van Schaik  
Productiviteit en Arbeidsparticipatie
- 507 Peter Borm, Anne van den Nouweland and Stef Tijs  
Cooperation and communication restrictions: a survey
- 508 Willy Spanjers, Robert P. Gilles, Pieter H.M. Ruys  
Hierarchical trade and downstream information
- 509 Martijn P. Tummers  
The Effect of Systematic Misperception of Income on the Subjective Poverty Line
- 510 A.G. de Kok  
Basics of Inventory Management: Part 1  
Renewal theoretic background
- 511 J.P.C. Blanc, F.A. van der Duyn Schouten, B. Pourbabai  
Optimizing flow rates in a queueing network with side constraints
- 512 R. Peeters  
On Coloring  $j$ -Unit Sphere Graphs



- 513 Drs. J. Dagevos, Drs. L. Oerlemans, Dr. F. Boekema  
Regional economic policy, economic technological innovation and networks
- 514 Erwin van der Krabben  
Het functioneren van stedelijke onroerend-goed-markten in Nederland - een theoretisch kader
- 515 Drs. E. Schaling  
European central bank independence and inflation persistence
- 516 Peter M. Kort  
Optimal abatement policies within a stochastic dynamic model of the firm
- 517 Pim Adang  
Expenditure versus consumption in the multi-good life cycle consumption model
- 518 Pim Adang  
Large, infrequent consumption in the multi-good life cycle consumption model
- 519 Raymond Gradus, Sjak Smulders  
Pollution and Endogenous Growth
- 520 Raymond Gradus en Hugo Keuzenkamp  
Arbeidsongeschiktheid, subjectief ziektegevoel en collectief belang
- 521 A.G. de Kok  
Basics of inventory management: Part 2  
The (R,S)-model
- 522 A.G. de Kok  
Basics of inventory management: Part 3  
The (b,Q)-model
- 523 A.G. de Kok  
Basics of inventory management: Part 4  
The (s,S)-model
- 524 A.G. de Kok  
Basics of inventory management: Part 5  
The (R,b,Q)-model
- 525 A.G. de Kok  
Basics of inventory management: Part 6  
The (R,s,S)-model
- 526 Rob de Groof and Martin van Tuijl  
Financial integration and fiscal policy in interdependent two-sector economies with real and nominal wage rigidity

- 527 A.G.M. van Eijs, M.J.G. van Eijs, R.M.J. Heuts  
Gecoördineerde bestelsystemen  
een management-georiënteerde benadering
- 528 M.J.G. van Eijs  
Multi-item inventory systems with joint ordering and transportation  
decisions
- 529 Stephan G. Vanneste  
Maintenance optimization of a production system with buffercapacity
- 530 Michel R.R. van Bremen, Jeroen C.G. Zijlstra  
Het stochastische variantie optiewaarderingsmodel
- 531 Willy Spanjers  
Arbitrage and Walrasian Equilibrium in Economies with Limited Infor-  
mation

## IN 1992 REEDS VERSCHENEN

- 532 F.G. van den Heuvel en M.R.M. Turlings  
Privatisering van arbeidsongeschiktheidsregelingen  
Refereed by Prof.Dr. H. Verbon
- 533 J.C. Engwerda, L.G. van Willigenburg  
LQ-control of sampled continuous-time systems  
Refereed by Prof.dr. J.M. Schumacher
- 534 J.C. Engwerda, A.C.M. Ran & A.L. Rijkeboer  
Necessary and sufficient conditions for the existence of a positive  
definite solution of the matrix equation  $X + A^*X^{-1}A = Q$ .  
Refereed by Prof.dr. J.M. Schumacher
- 535 Jacob C. Engwerda  
The indefinite LQ-problem: the finite planning horizon case  
Refereed by Prof.dr. J.M. Schumacher
- 536 Gert-Jan Otten, Peter Borm, Ton Storcken, Stef Tijs  
Effectivity functions and associated claim game correspondences  
Refereed by Prof.dr. P.H.M. Ruys
- 537 Jack P.C. Kleijnen, Gustav A. Alink  
Validation of simulation models: mine-hunting case-study  
Refereed by Prof.dr.ir. C.A.T. Takkenberg
- 538 V. Feltkamp and A. van den Nouweland  
Controlled Communication Networks  
Refereed by Prof.dr. S.H. Tijs
- 539 A. van Schaik  
Productivity, Labour Force Participation and the Solow Growth Model  
Refereed by Prof.dr. Th.C.M.J. van de Klundert
- 540 J.J.G. Lemmen and S.C.W. Eijffinger  
The Degree of Financial Integration in the European Community  
Refereed by Prof.dr. A.B.T.M. van Schaik
- 541 J. Bell, P.K. Jagersma  
Internationale Joint Ventures  
Refereed by Prof.dr. H.G. Barkema
- 542 Jack P.C. Kleijnen  
Verification and validation of simulation models  
Refereed by Prof.dr.ir. C.A.T. Takkenberg
- 543 Gert Nieuwenhuis  
Uniform Approximations of the Stationary and Palm Distributions  
of Marked Point Processes  
Refereed by Prof.dr. B.B. van der Genugten

- 544 R. Heuts, P. Nederstigt, W. Roebroek, W. Selen  
Multi-Product Cycling with Packaging in the Process Industry  
Refereed by Prof.dr. F.A. van der Duyn Schouten
- 545 J.C. Engwerda  
Calculation of an approximate solution of the infinite time-varying  
LQ-problem  
Refereed by Prof.dr. J.M. Schumacher
- 546 Raymond H.J.M. Gradus and Peter M. Kort  
On time-inconsistency and pollution control: a macroeconomic approach  
Refereed by Prof.dr. A.J. de Zeeuw
- 547 Drs. Dolph Cantrijn en Dr. Rezaul Kabir  
De Invloed van de Invoering van Preferente Beschermingsaandelen op  
Aandelenkoersen van Nederlandse Beursgenoteerde Ondernemingen  
Refereed by Prof.dr. P.W. Moerland
- 548 Sylvester Eijffinger and Eric Schaling  
Central bank independence: criteria and indices  
Refereed by Prof.dr. J.J. Sijben
- 549 Drs. A. Schmeits  
Geïntegreerde investerings- en financieringsbeslissingen; Implicaties  
voor Capital Budgeting  
Refereed by Prof.dr. P.W. Moerland
- 550 Peter M. Kort  
Standards versus standards: the effects of different pollution  
restrictions on the firm's dynamic investment policy  
Refereed by Prof.dr. F.A. van der Duyn Schouten
- 551 Niels G. Noorderhaven, Bart Nooteboom and Johannes Berger  
Temporal, cognitive and behavioral dimensions of transaction costs;  
to an understanding of hybrid vertical inter-firm relations  
Refereed by Prof.dr. S.W. Douma
- 552 Ton Storcken and Harrie de Swart  
Towards an axiomatization of orderings  
Refereed by Prof.dr. P.H.M. Ruys
- 553 J.H.J. Roemen  
The derivation of a long term milk supply model from an optimization  
model  
Refereed by Prof.dr. F.A. van der Duyn Schouten
- 554 Geert J. Almekinders and Sylvester C.W. Eijffinger  
Daily Bundesbank and Federal Reserve Intervention and the Conditional  
Variance Tale in DM/\$-Returns  
Refereed by Prof.dr. A.B.T.M. van Schaik
- 555 Dr. M. Hetebrij, Drs. B.F.L. Jonker, Prof.dr. W.H.J. de Freytas  
"Tussen achterstand en voorsprong" de scholings- en personeelsvoor-  
zieningsproblematiek van bedrijven in de procesindustrie  
Refereed by Prof.dr. Th.M.M. Verhallen

- 556 Ton Geerts  
Regularity and singularity in linear-quadratic control subject to implicit continuous-time systems  
Communicated by Prof.dr. J. Schumacher
- 557 Ton Geerts  
Invariant subspaces and invertibility properties for singular systems: the general case  
Communicated by Prof.dr. J. Schumacher
- 558 Ton Geerts  
Solvability conditions, consistency and weak consistency for linear differential-algebraic equations and time-invariant singular systems: the general case  
Communicated by Prof.dr. J. Schumacher
- 559 C. Fricker and M.R. Jaïbi  
Monotonicity and stability of periodic polling models  
Communicated by Prof.dr.ir. O.J. Boxma
- 560 Ton Geerts  
Free end-point linear-quadratic control subject to implicit continuous-time systems: necessary and sufficient conditions for solvability  
Communicated by Prof.dr. J. Schumacher
- 561 Paul G.H. Mulder and Anton L. Hempenius  
Expected Utility of Life Time in the Presence of a Chronic Noncommunicable Disease State  
Communicated by Prof.dr. B.B. van der Genugten
- 562 Jan van der Leeuw  
The covariance matrix of ARMA-errors in closed form  
Communicated by Dr. H.H. Tigelaar
- 563 J.P.C. Blanc and R.D. van der Mei  
Optimization of polling systems with Bernoulli schedules  
Communicated by Prof.dr.ir. O.J. Boxma
- 564 B.B. van der Genugten  
Density of the least squares estimator in the multivariate linear model with arbitrarily normal variables  
Communicated by Prof.dr. M.H.C. Paardekooper
- 565 René van den Brink, Robert P. Gilles  
Measuring Domination in Directed Graphs  
Communicated by Prof.dr. P.H.M. Ruys
- 566 Harry G. Barkema  
The significance of work incentives from bonuses: some new evidence  
Communicated by Dr. Th.E. Nijman

- 567 Rob de Groof and Martin van Tuijl  
Commercial integration and fiscal policy in interdependent, financially integrated two-sector economies with real and nominal wage rigidity.  
Communicated by Prof.dr. A.L. Bovenberg
- 568 F.A. van der Duyn Schouten, M.J.G. van Eijs, R.M.J. Heuts  
The value of information in a fixed order quantity inventory system  
Communicated by Prof.dr. A.J.J. Talman
- 569 E.N. Kertzman  
Begrotingsnormering en EMU  
Communicated by Prof.dr. J.W. van der Dussen
- 570 A. van den Elzen, D. Talman  
Finding a Nash-equilibrium in noncooperative N-person games by solving a sequence of linear stationary point problems  
Communicated by Prof.dr. S.H. Tijs
- 571 Jack P.C. Kleijnen  
Verification and validation of models  
Communicated by Prof.dr. F.A. van der Duyn Schouten
- 572 Jack P.C. Kleijnen and Willem van Groenendaal  
Two-stage versus sequential sample-size determination in regression analysis of simulation experiments
- 573 Pieter K. Jagersma  
Het management van multinationale ondernemingen: de concernstructuur
- 574 A.L. Hempenius  
Explaining Changes in External Funds. Part One: Theory  
Communicated by Prof.Dr.Ir. A. Kapteyn
- 575 J.P.C. Blanc, R.D. van der Mei  
Optimization of Polling Systems by Means of Gradient Methods and the Power-Series Algorithm  
Communicated by Prof.dr.ir. O.J. Boxma
- 576 Herbert Hamers  
A silent duel over a cake  
Communicated by Prof.dr. S.H. Tijs
- 577 Gerard van der Laan, Dolf Talman, Hans Kremers  
On the existence and computation of an equilibrium in an economy with constant returns to scale production  
Communicated by Prof.dr. P.H.M. Ruys

**Bibliotheek K. U. Brabant**



**17 000 01109978 6**