数理解析研究所講究録
749 巻 1991 年 147-161

Exponential Dispersion Model と 巾変換

広島大学　総合科学部
西井 龍映 (Ryuei NISHII)

SUMMARY

It is known that a normalizing transformation of an exponential family is determined by the differential equation. We shall prove that this correspondence between the families and the transformations is one-to-one within the class of all exponential families. This result yields that the Box-Cox transformations are normalizing transformations of the exponential dispersion models with power variance functions. The correspondence is plotted graphically. Using this relation we propose a new parameter estimation procedure for the exponential dispersion models with power variance functions, and the procedure is examined by the numerical example. Variance-stabilizing transformations are also discussed.

1. Introduction

Box and Cox (1964) proposed a family of power transformations such that the transformed data has clearly defined properties, e.g., constancy of variance and/or normality. It is well known that familiar distributions are closely related to the Box-Cox transformations. For example, the square transformation is a variance-stabilizing transformation of a Poisson distribution (Anscombe, 1948), and the cube transformation is a normalizing transformation of a gamma distribution (Wilson & Hilferty, 1931). We try to unify the results obtained in the literature.

In general it is known (Konishi, 1981) that a normalizing (variance-stabilizing) transformation of an exponential family is determined by the differential equation. We shall prove that this correspondence between the families and the transformations is one-to-one within the class of all exponential families. This assertion also holds within the class of all exponential dispersion models. The exponential dispersion model is a generalization of an exponential family, and was reviewed by Jørgensen (1987). This result yields that the Box-Cox transformations are normalizing (variance-stabilizing) transformations of exponential dispersion models with power variance functions.

In Section 2 we review exponential dispersion models with power variance functions. Section 3 gives an one-to-one correspondence between exponential families and the normalizing (variance-stabilizing) transformations. This theorem is applied to the Box-Cox transformations in Section 4, and the correspondence is plotted out in FIGUREs 1 and 2. Using this correspondence we propose a new parameter estimation procedure for the exponential dispersion model with power variance function in Section 5. The procedure is examined by the numerical example.

2. Exponential dispersion models

Suppose a d-variate random vector $\underset{\sim}{X}$ comes from a distribution with density function of the form

$$dF(\underset{\sim}{x};\lambda,\underset{\sim}{\theta})/d\nu(\underset{\sim}{x}) \equiv f(\underset{\sim}{x};\lambda,\underset{\sim}{\theta}) = a(\underset{\sim}{x};\lambda)e^{\lambda\{\underset{\sim}{\theta}^T\underset{\sim}{x} - \kappa(\underset{\sim}{\theta})\}} \tag{1}$$

where a and κ are given functions, and $\underset{\sim}{\theta}$ varies in Θ of a subset of $\mathbb{R}^d$ and λ varies in Λ of a subset of $\mathbb{R}_+$. Here ν is a σ-finite measure on $\mathbb{R}^d$. The parameter $1/\lambda$ stands for a measure of dispersion. This family is called an *exponential dispersion model* and denoted by $ED(\lambda,\underset{\sim}{\theta})$.

The exponential dispersion model was introduced by Nelder and Wedderburn (1972) with d = 1, and by Jørgensen (1987) in multivariate case. Such an original idea may be found in Tweedie (1947). The exponential dispersion model is used for error distributions of generalized linear models. Allowing that $\lambda \equiv 1$, we can regard that the model is an generalization of the exponential family.

Another exponential dispersion models are used for describing discrete distributions. Discrete dispersion models, written as $ED^*(\lambda,\underset{\sim}{\theta})$, consist of densities of the form

$$dG(\underset{\sim}{y};\lambda,\underset{\sim}{\theta})/d\nu(\underset{\sim}{y}) \equiv g(\underset{\sim}{y};\lambda,\underset{\sim}{\theta}) = a(\underset{\sim}{y};\lambda)e^{\underset{\sim}{\theta}^T\underset{\sim}{y} - \lambda\kappa(\underset{\sim}{\theta})}. \tag{2}$$

Let $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ be random vectors with density functions of the form (1) and (2) respectively. Then cumulant generating functions are given by

$$\log E_F[\exp(\underset{\sim}{t}^T\underset{\sim}{X})] = \lambda\{\kappa(\underset{\sim}{\theta} + \underset{\sim}{t}/\lambda) - \kappa(\underset{\sim}{\theta})\} \quad \text{and} \tag{3}$$

$$\log E_G[\exp(\underset{\sim}{t}^T\underset{\sim}{Y})] = \lambda\{\kappa(\underset{\sim}{\theta} + \underset{\sim}{t}) - \kappa(\underset{\sim}{\theta})\}. \tag{4}$$

Hence $\kappa(\underset{\sim}{\theta})$ is called a *cumulant generator*. Immediately it holds that $E[\underset{\sim}{X}] = \partial\kappa(\underset{\sim}{\theta})/\partial\underset{\sim}{\theta}$, $Var[\underset{\sim}{X}] = (1/\lambda)\partial^2\kappa(\underset{\sim}{\theta})/\partial\underset{\sim}{\theta}\partial\underset{\sim}{\theta}^T$, $E[\underset{\sim}{Y}] = \lambda\cdot\partial\kappa(\underset{\sim}{\theta})/\partial\underset{\sim}{\theta}$ and $Var[\underset{\sim}{Y}] = \lambda\cdot\partial^2\kappa(\underset{\sim}{\theta})/\partial\underset{\sim}{\theta}\partial\underset{\sim}{\theta}^T$. Barndorff-Nielsen (1978) showed that the mapping $\underset{\sim}{\theta} \longrightarrow \partial\kappa(\underset{\sim}{\theta})/\partial\underset{\sim}{\theta}$ ($\equiv \underset{\sim}{\mu}$, say) is one-to-one. We define the inverse mapping by $\underset{\sim}{\theta} = \xi(\underset{\sim}{\mu})$. The d×d matrix $\partial^2\kappa(\underset{\sim}{\theta})/\partial\underset{\sim}{\theta}\partial\underset{\sim}{\theta}^T\big|_{\underset{\sim}{\theta}=\xi(\underset{\sim}{\mu})}$ ($\equiv V(\underset{\sim}{\mu})$, say) is called a *variance function*, and the variance function characterizes the distribution.

Theorem 1 (Jørgensen, 1987). An exponential dispersion model is characterized within the class of all exponential dispersion models by its variance functions.

In the sequel we consider univariate (d = 1) exponential dispersion models. Especially the exponential dispersion model with power variance function

$$V(\mu) = \mu^p$$

is very attractive. Letting $\alpha = (p-2)/(p-1)$, we denote this model by $ED^{(\alpha)} = ED^{(\alpha)}(\lambda,\theta)$, and its cumulant generator by $\kappa_\alpha(\theta)$. Giving various values to α, we get important families.

$ED^{(2)}$: the normal family with $p = 0$, $\kappa_2(\theta) = \theta^2$,
$ED^{(-\infty)}$: the Poisson family with $p = 1$, $\kappa_{-\infty}(\theta) = e^\theta$,
$ED^{(0)}$: the gamma family with $p = 2$, $\kappa_0(\theta) = -\log(-\theta)$,
$ED^{(1/2)}$: the inverse gaussian family with $p = 3$, $\kappa_{1/2}(\theta) = \sqrt{-2\theta}$.

In general Jørgensen (1987) proved that there exists an exponential dispersion model with $\kappa_\alpha(\theta) = (\alpha-1)\alpha^{-1}\{\theta/(\alpha-1)\}^\alpha$ when $p \le 0$, $1 < p < 2$ or $p > 2$ (equivalently $\alpha \le 2$), and in other case there exists no exponential dispersion model. Further he showed that a distribution in $ED^{(\alpha)}$ is continuous on $\mathbb{R}_+$ with an atom at $x = 0$ if $\alpha < 0$, is continuous on $\mathbb{R}_+$ if $0 \le \alpha < 1$, and is continuous on $\mathbb{R}$ if $1 < \alpha \le 2$. The exact density in $ED^{(\alpha)}(\lambda,\theta)$ for $0 < \alpha < 1$, $1 < \alpha < 2$ is given by

$$f_\alpha(x;\lambda,\theta) = c\cdot S_\alpha(cx)\exp\left(\lambda\{\theta x - \frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^\alpha\} \right),$$

where $c = c(\lambda) = \alpha^{1/\alpha}(1-\alpha)^{1-1/\alpha}\lambda^{1-1/\alpha}$. Here $S_\alpha(x)$ is a probability density of a stable distribution with index α, see, e.g., p.583 of Feller (1971).

When $0 < \alpha < 1$,

$$S_\alpha(x) = -\frac{1}{\pi x}\sum_{k=1}^{\infty}\frac{\Gamma(\alpha k+1)}{k!}(-x^{-\alpha})^k \sin(\alpha\pi k) \text{ for } x > 0 \text{ ; } = 0 \text{ for } x \le 0,$$

and when $1 < \alpha < 2$,

$$S_\alpha(x) = -\frac{1}{\pi|x|}\sum_{k=1}^{\infty}\frac{\Gamma(1+k/\alpha)}{k!}(-|x|)^k \sin(\pi k/2) \text{ for } x \ne 0.$$

The convergence of $S_\alpha(x)$ is very slow. Hence it is hard to obtain the maximum likelihood estimator of (α,λ,θ). Stable distributions were extensively studied by Zolotarev (1986).

When $\alpha < 0$, the distribution in $ED^{(\alpha)}(\lambda,\theta)$ is given by the Poisson mixture of a gamma distribution. More precisely, the mean of the Poisson distribution is given by $\lambda(\alpha-1)\alpha^{-1}\{\theta/(\alpha-1)\}^\alpha$, and parameters of the gamma distribution $ED^{(0)}(\lambda',\theta')$ are given by $\lambda' = -\alpha$ and $\theta' = \theta\lambda$. See, e.g., Siegel (1985).

3. Normalizing and variance-stabilizing transformations

Let $\bar{X}$ be a mean of an n-random sample from a univariate exponential dispersion model $ED(\lambda,\theta)$. Then by calculating the moment generating function of $\bar{X}$ by (2), we hold that $\bar{X}$ follows $ED(n\lambda,\theta)$. On the other hand by the central limit theorem a distribution of $\sqrt{n\lambda}\{\bar{X} - \kappa'(\theta)\}$ converges to a normal distribution $N(0, \kappa''(\theta))$ as n tends to infinity. Substituting $n\lambda$ to λ, and $\bar{X}$ to X ($\sim ED(\lambda,\theta)$), we know that a distribution of $\sqrt{\lambda}\{X - \kappa'(\theta)\}$ converges to a normal distribution $N(0, \kappa''(\theta))$ as λ tends to infinity. This implies that for a smooth function h(x), it holds that

$$\sqrt{\lambda}\{h(X) - h(\kappa'(\theta))\} \xrightarrow{L} N[0, \kappa''(\theta)\{h'(\kappa'(\theta))\}^2] \ (\lambda \longrightarrow \infty). \tag{5}$$

In this article we treat the asymptotics that λ tends to infinity when the underlying distribution F is a member of the exponential dispersion model. If F is a member of an exponential family, substituting X with $\bar{X}$ and λ with n in (5), we consider the asymptotics that n tends to infinity.

Let h(x) be a twice continuously differentiable monotone function with $h'(x) > 0$. We transform X ($\sim ED(\lambda,\theta)$) into h(X). We regard that h(x) and $c_1h(x) + c_2$ ($c_1 > 0$) provide the same transformation. Now standardizing h(X) by the asymptotic bias and the asymptotic variance, we put

$$Z_h(X) = \frac{\sqrt{\lambda}}{h'\{\kappa'(\theta)\}\sqrt{\kappa''(\theta)}}\left(h(X) - h\{\kappa'(\theta)\} - \frac{\kappa''(\theta)h''\{\kappa'(\theta)\}}{2\lambda h'\{\kappa'(\theta)\}}\right). \tag{6}$$

Let K be the third cumulant of $Z_h(X)$:

$$K = \frac{1}{3}\left(\frac{\kappa'''(\theta)}{\{\kappa''(\theta)\}^{3/2}} + 3\sqrt{\kappa''(\theta)}\,\frac{h''\{\kappa'(\theta)\}}{h'\{\kappa'(\theta)\}}\right). \tag{7}$$

Then we get the asymptotic expansion of the distribution function as

$$\Pr[Z_h(X) \le z] = \Phi(z) - \frac{1}{2\sqrt{\lambda}}\cdot K(z^2 - 1)\phi(z) + O(\tfrac{1}{\lambda}), \tag{8}$$

where $\phi(z)$ and $\Phi(z)$ are the density and the distribution function of the standard normal distribution respectively.

When Y has a discrete distribution of the form (2), we consider the transformation from Y to $h(Y/\lambda)$. Then

$$Z_h(Y/\lambda) = \frac{\sqrt{\lambda}}{h'\{\kappa'(\theta)\}\sqrt{\kappa''(\theta)}}\left(h(Y/\lambda) - h\{\kappa'(\theta)\} - \frac{\kappa''(\theta)h''\{\kappa'(\theta)\}}{2\lambda h'\{\kappa'(\theta)\}}\right)$$

converges a normal distribution $N(0, \kappa''(\theta))$ in law as λ tends to infinity. For a discrete distribution, (8) is not valid because the finite correction term of order $O(1/\sqrt{\lambda})$ should be added. See e.g. Siotani and Fujikoshi (1984). However the following approximation for the density

$$\Pr[Z_h(Y/\lambda) = z] = \phi(z) + \frac{1}{2\sqrt{\lambda}}\cdot K\cdot(z^3 - 3z)\phi(z) + O(\tfrac{1}{\lambda})$$

is valid when $Z_h(Y/\lambda)$ takes the value z with positive probability, where K is of the form as (7).

A function h(x) is called a *variance-stabilizing transformation of an exponential dispersion model* (or *of an exponential family*) when the asymptotic variance of $Z_h(X)$ of (6) is independent of θ, i.e.,

$$h'\{\kappa'(\theta)\}\sqrt{\kappa''(\theta)} \equiv 1 \quad \text{or equivalently} \quad h\{\kappa'(\theta)\} = \int\sqrt{\kappa''(\theta)}d\theta. \qquad (9)$$

The most important transformation may be a *normalizing transformation* which vanishes the coefficient of $O(1/\sqrt{\lambda})$ in the expansion (8). Such a characterization was established by Konishi (1981, 1987). In our setup, a normalizing transformation is given by a solution of the differential equation that K of (7) equals zero. This equation has the unique solution

$$h\{\kappa'(\theta)\} = \int\kappa''(\theta)^{2/3}d\theta. \qquad (10)$$

Both transformations (9) and (10) are given by the solutions of the differential equations. Conversely for a given h(x), (9) and (10) determine $\kappa(\theta)$. However, a family with the cumulant generator $\kappa(\theta)$ may not exist. Thus we have

Theorem 2. Let h(x) be a given twice continuously differentiable monotone function. Then

(1) h(x) is a variance-stabilizing transformation of an exponential

dispersion model (of an exponential family) whose cumulant generator $\kappa(\theta)$ satisfies $\kappa''(\theta) = [h'\{\kappa'(\theta)\}]^{-2}$, and

(2) $h(x)$ is a normalizing transformation of an exponential dispersion model (of an exponential family) whose cumulant generator $\kappa(\theta)$ satisfies $\kappa''(\theta) = [h'\{\kappa'(\theta)\}]^{-3}$.

4. The Box-Cox transformations

As a special case of Theorem 2, consider the signed Box-Cox transformations

$$h_q(x) = \mathrm{sign}(x)\cdot|x|^q \quad (q \neq 0); \quad h_0(x) = \log x,$$

and the exponential dispersion model $ED^{(\alpha)}$ with power variance function.

Theorem 3.

(1a) The variance-stabilizing transformation of $ED^{(\alpha)}$ ($\alpha \leq 2$, $\alpha \neq 1$) is given by $h_q(x)$ with $q = \alpha/\{2(\alpha-1)\}$,

(1b) the variance-stabilizing transformation of the gamma family $ED^{(0)}$ is given by $h_0(x) = \log x$,

(1c) the variance-stabilizing transformation of the Poisson family $ED^{(-\infty)}$ is given by $h_{1/2}(x)$ (Anscombe, 1948),

and each converse of (1a)-(1c) is also valid within the class of all exponential dispersion models or within the class of all exponential families.

Note that $q = \alpha/\{2(\alpha-1)\}$ takes the values $q < 1/2$ or $q \geq 1$ as α varies in $\alpha \leq 2$, $\alpha \neq 1$. The inverse function $\alpha = 2q/(2q-1)$ is plotted in FIGURE 1. Taking $\alpha = 1/2$ in (1a) we know that the inverse gaussian family $ED^{(1/2)}$ is corresponding to $h_{-1/2}(x) = 1/\sqrt{x}$.

Theorem 4.

(2a) The normalizing transformation of $ED^{(\alpha)}$ ($\alpha \le 2$, $\alpha \ne 1/2, 1$) is given by $h_r(x)$ with $r = (2\alpha-1)/\{3(\alpha-1)\}$,

(2b) the normalizing transformation of the inverse gaussian family $ED^{(1/2)}$ is given by $h_0(x)$ (Whitmore-Yalovsky, 1978),

(2c) the normalizing transformation of the gamma family $ED^{(0)}$ is given by $h_{1/3}(x)$ (Wilson-Hilferty, 1931),

(2d) the normalizing transformation of the Poisson family $ED^{(-\infty)}$ is given by $h_{2/3}(x)$ (Blom, 1954),

and each converse of (2a)-(2d) is also valid within the class of all exponential dispersion models or within the class of all exponential families.

Note that $r = (2\alpha-1)/\{3(\alpha-1)\}$ takes the values $r < 2/3$ or $r \ge 1$ as α varies $\alpha \le 2$, $\alpha \ne 1/2, 1$. The inverse function $\alpha = (3r-1)/(3r-2)$ is illustrated in FIGURE 2. Recall the distribution of $ED^{(\alpha)}$ with $\alpha < 0$ takes zero with positive probability. However in Theorems 3 and 4, this causes no problem because $q = \alpha/\{2(\alpha-1)\}$ of (1a) and $r = (2\alpha-1)/\{3(\alpha-1)\}$ of (2a) are positive when $\alpha < 0$.

Corollary 5. Suppose a function $h(x)$ is a variance-stabilizing and normalizing transformation of an exponential family. Then $h(x)$ should be of the form $c_1x + c_2$ with constants $c_1 \ne 0$ and c_2, and the family should be normal.

Proof. Let $h(x)$ and $\kappa(\theta)$ satisfy the relations (9) and (10). From (9) we have $h'\{\kappa'(\theta)\} = \kappa''(\theta)^{-1/2}$ and $h''\{\kappa'(\theta)\} = -1/2\ \kappa''(\theta)^{-5/2}$.

Hence we have $-1/2\cdot\kappa'''(\theta)/\kappa''(\theta)^{3/2} \equiv 0$ or $\kappa'''(\theta) \equiv 0$. Thus $\kappa(\theta) = c_1\theta^2 + c_2\theta + c_3$ and $h'(2c_1\theta + c_2) = (2c_1)^{-1/2}$ for some constants $c_1 > 0$, c_2 and c_3. This completes the proof.

The z-transformation of a sample correlation coefficient is a variance-stabilizing and normalizing transformation. We remark that a distribution of a sample correlation coefficient is not a member of an exponential family.

5. Applications for parameter estimations

As we have already seen, the maximum likelihood estimators of the parameters of $ED^{(\alpha)}(\lambda,\theta)$ are hard to obtain. The moment method for parameter estimation based on a random sample from $ED^{(\alpha)}(\lambda,\theta)$ is :

$$\hat{\alpha} = (2\hat{\sigma}^4 - \hat{\kappa}_3\bar{x})/(\hat{\sigma}^4 - \hat{\kappa}_3\bar{x}), \qquad \hat{\theta} = (\hat{\alpha} - 1)\bar{x}^{(\hat{\alpha}-1)^{-1}}, \qquad \text{and} \quad \hat{\lambda} = \bar{x}^{(\hat{\alpha}-2)/(\hat{\alpha}-1)}/\hat{\sigma}^2, \tag{12}$$

where $\hat{\kappa}_3$ is an estimate of the third cumulant (see Hougaard, 1986). On the other hand, Theorem 4 may provide an alternative procedure for estimating the index α. The proposed procedure is as follows.

Step 1: Fit a normal distribution after the signed Box-Cox transformations from x_i to $\mathrm{sign}(x_i)|x_i|^r$ and find the power, say $\tilde{r}$, which maximizes the likelihood.

Step 2: If $2/3 < \tilde{r} < 1$, then we stop to fit $ED^{(\alpha)}$ for the data.

Step 3: If $\tilde{r} = 1/3$, a gamma distribution is fitted,
if $\tilde{r} = 2/3$, a Poisson distribution is fitted,
and in other case parameters are estimated by

$$\tilde{\alpha} = (3\tilde{r}-1)/(3\tilde{r}-2),\ \tilde{\theta} = (\tilde{\alpha}-1)\bar{x}^{(\tilde{\alpha}-1)^{-1}},\ \tilde{\lambda} = \bar{x}^{(\tilde{\alpha}-2)/(\tilde{\alpha}-1)}/\hat{\sigma}^2. \quad (13)$$

Step 4: Obtain likelihoods of $ED^{(\alpha)}$ with $\tilde{\alpha}$ of (13), $\hat{\alpha}$ of (12), $\alpha = 0$ (gamma), $\alpha = 1/2$ (inverse gaussian) and $\alpha = 2$ (normal). Then choose the best model by using an information criterion such as AIC.

Concerning on Step 3, $\tilde{\theta}$ is a conditional maximum likelihood estimate given $\tilde{\alpha}$, and $\tilde{\lambda}$ is obtained by the moment method.

This procedure is examined for the weight data of 98 newly-enrolled male students of a university of Hiroshima. The estimated mean and variance are $\hat{\mu}$ = 61.02(Kg) and $\hat{\sigma}^2$ = 38.79. The following table summarizes the performance of the fitted models.

TABLE Weights (Kg) of 98 freshemen

data set : (†45) 51, 6*52, 3*53, 4*54, 7*55, 2*56, 9*57, 7*58, 4*59, 6*60, 4*61, 5*62, 8*63, 5*64, 9*65, 5*66, 2*67, 68, 70, 72, 3*73, 74, 76, 79, 82 (†94)

(n*w implies that w (Kg) is repeated n times.
The numerals were rounded off at the first decimal.
The original sample size was 100. The minimum and maximum values †45 and †94 of the original data are treated as outliers, and omitted.)

Fitted model	-2·logarithm of likelihood	# of parameters	
Normal	636.62	2	(α = 2)
Inverse gaussian	630.398	2	(α = 1/2)
Log normal	630.397	2	
Normal after Box-Cox transformation	626.10	3	(estimated power $\tilde{r}$ = -1.835)
$ED^{(\alpha)}$ (moment method)	625.96	3	($\hat{\alpha}$ = .85088)
$ED^{(\alpha)}$ (from Box-Cox)	625.73*	3	($\tilde{\alpha} = (3\tilde{r}-1)/(3\tilde{r}-2)$ = .86676)

$ED^{(\alpha)}$ (MLE of α) 625.67 3 ($\hat{\alpha}_{ML}$ = .878)

(* : proposed method)

Hence our procedure works well because its likelihood is very close to that based on the maximum likelihood principle. This data may follow a distribution of an exponential dispersion model with power variance function. We also note that the likelihoods of an inverse gaussian distribution and a log-normal distribution are nearly equal. This phenomena is interpreted by (2b) of Theorem 4. Actually the estimated inverse gaussian distribution is $ED^{(1/2)}(6147, -0.01343)$, and the estimated λ = 6147 is fairly large.

ACKNOWLEDGMENT

I would like to express my hearty thanks to Prof. S. Konishi, Institute of Statistical Mathematics, for drawing my attention to this study. Thanks are also due to Prof. I. Kubo, Hiroshima University, for introducing me the book by Zolotarev.

REFERENCES

Anscombe, F.J. (1948) The transformation of Poisson, binomial, and negative binomial data. *Biometrika* 35, 246-54.

Barndorff-Nielsen, O.E. (1978) *Information and Exponential Families in Statistical Theory.* Wiley

Blom, G. (1954) Transformation of the binomial, negative binomial, Poisson and χ^2 distributions. *Biometrika* 41, 302-16.

Box, G.E.P. & Cox, D.R. (1964) An analysis of transformations. *J. R. Statist. Soc. B* 26, 211-52.

Feller, W. (1971) *An Introduction to Probability Theory and Its*

Applications (Second Edition). Wiley.

Hougaard, P. (1986) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387-96.

Jørgensen, B. (1987) Exponential dispersion models. *J. R. Statist. Soc. B* 49, 127-62.

Konishi, S. (1981) Normalizing transformations of some statistics in multivariate analysis. *Biometrika* 68, 647-51.

Konishi, S. (1987) Transformations of statistics in multivariate analysis. *Advances in Multivariate Statistical Analysis*, 213-31, A.K.Gupta (ed.), D. Reidel Publishing Co.

Nelder, J.A. & Wedderburn, R.W.M. (1972) Generalized linear models. *J. R. Statist. Soc. A* 135, 370-84.

Siegel, A.F. (1985) Modelling data containing exact zeros using zero degrees of freedom. *J. R. Statist. Soc. B* 47, 267-71.

Siotani, M. and Fujikoshi, Y. (1984) Asymptotic approximations for the distributions of multinomial goodness-of-fit statistics. *Hiroshima Math. J.* 14, 115-24.

Tweedie, M.C.K. (1947) Functions of a statistical variate with given means, with special reference to Laplacian distributions. *Proc. Cambridge Phil. Soc.* 49, 41-9.

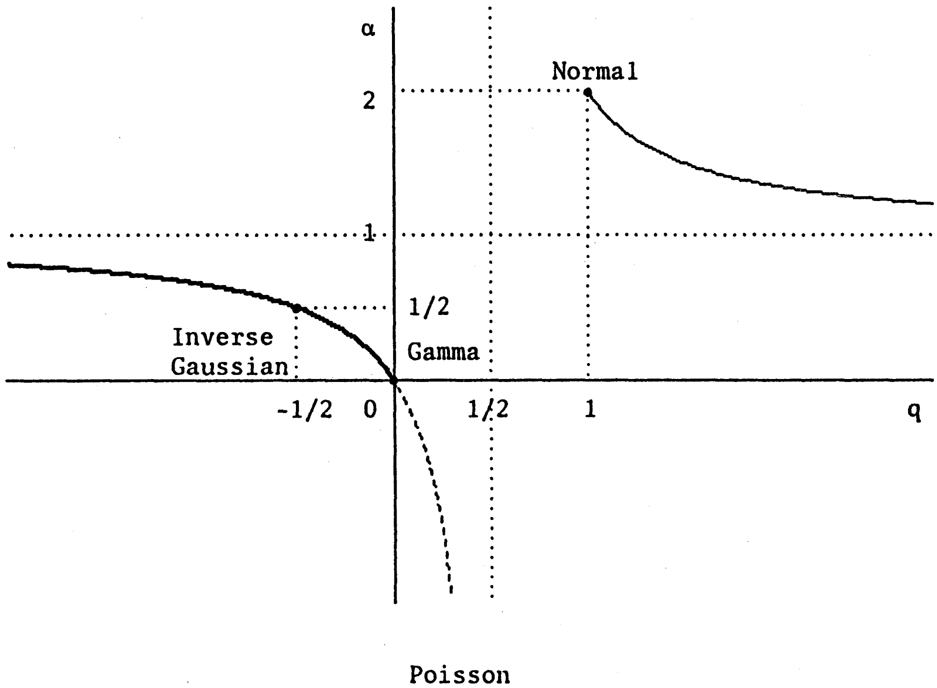
Whitmore, G.A. and Yalovsky, M. (1978) A normalizing logarithmic transformation for inverse Gaussian random variables. *Technometrics* 20, 207-8.

Wilson, E.B. & Hilferty, M.M. (1931) The distribution function of chi-squares. *Proceedings of the National Academy of Sciences* 17, 684-8.

Zolotarev, V.M. (1986) *One-dimensional Stable Distributions*. Translations Math. Mono. 65, A.M.S.

FIGURE 1. Variance-Stabilizing transformation, Plot of $\alpha = 2q/(2q-1)$

The signed Box-Cox transformation $h_q(x) = \text{sign}(x)|x|^q$ ($q < 1/2$, $1 \le q$) ; $h_0(x) = \log(x)$ ($q = 0$) is a variance-stabilizing transformation of the exponential dispersion model $ED^{(\alpha)}$ with power variance function of $\alpha = 2q/(2q-1)$ ($\alpha \le 2$, $\alpha \ne 1$).



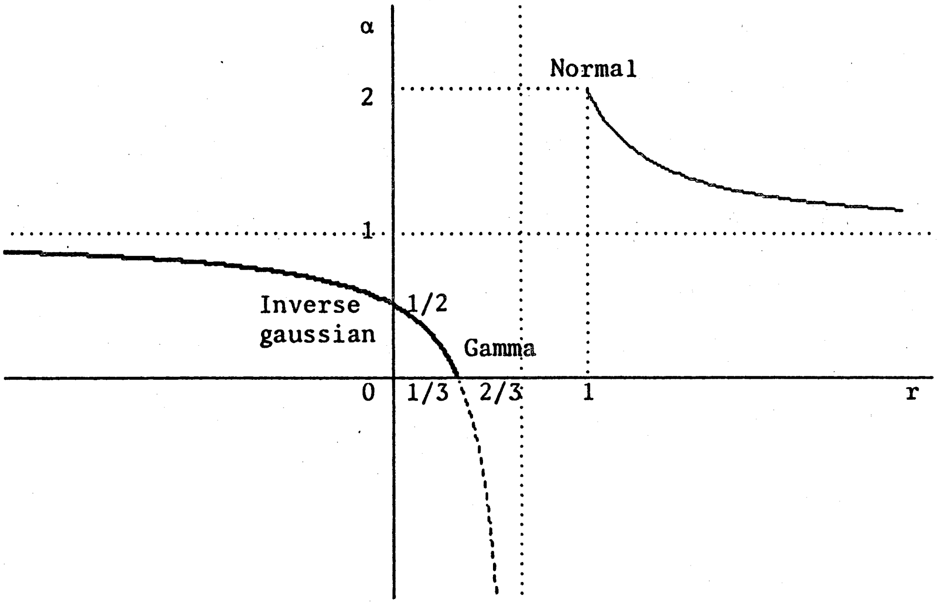
---- Poisson mixtures of gamma distributions, support $\{0\} \cup (0,\infty)$

—— Generated by extreme stable distributions, support $(-\infty,\infty)$

▬▬ Generated by positive stable distributions, support $(0,\infty)$

FIGURE 2. Normalizing transformation, Plot of $\alpha = (3r-1)/(3r-2)$

The signed Box-Cox transformation $h_r(x)$ ($r < 2/3$, $1 \le r$) is a normalizing transformation of $ED^{(\alpha)}$ with $\alpha = (3r-1)/(3r-2)$ ($\alpha \le 2$, $\alpha \ne 1$).



Poisson