International Journal of Theoretical Physics, Vol. 38, No. 10, 1999

Spatial Statistics and Information Geometry for Parametric Statistical Models of Galaxy Clustering

C. T. J. Dodson[1]

Received March 2, 1999

Poisson spatial processes of points and of extended objects representing smoothed clusters of galaxies are considered; some results are obtained for planar representations of random filaments, which may help interpret the findings of the Las Campanas Redshift Survey. Based on a model for the void probability function, a family of gamma-related distributions is investigated as a three-dimensional model for the clustering of galaxies. The unclustered models in this family correspond to the random case and to maximum information-theoretic entropy. The Riemannian information metric and Gaussian curvature are derived for the parameter space of the family of models, which provides a background on which to write dynamics for cluster evolution.

1. INTRODUCTION

For a general account of large-scale structures in the universe, see, for example, Peebles [18] and Fairall [9], the latter providing a comprehensive atlas. See also Cappi *et al.* [1], Coles [2], Labini *et al.* [13, 14], Vogeley *et al.* [21], and van der Weygaert [22] for further recent discussion of large structures. The Las Campanas Redshift Survey is currently the most comprehensive deep survey, providing some 26,000 data points in a slice out to 500 h^{-1} Mpc. Doroshkevich *et al.* [7] (cf. also Fairall [9 §5.4] and his Figure 5.5) revealed a rich texture of filaments, clusters and voids and suggested that it resembled a composite of three apparently Poisson processes:

1. *Superlarge-scale sheets*: 60% of galaxies, characteristic separation about 77 h^{-1} Mpc.

[1] Department of Mathematics, University of Manchester Institute of Science and Technology, Manchester M60 1QD, U.K.; e-mail: dodson@umist.ac.uk.

2. *Rich filaments*: 20% of galaxies, characteristic separation about 30 h^{-1} Mpc.

3. *Sparse filaments*: 20% of galaxies, characteristic separation about 13 h^{-1} Mpc.

There is a body of theory that provides the means to calculate the variance of density in planar Poisson processes of arbitrary rectangular elements, using arbitrary finite cells of inspection [3]. We provide details of this method. In principle, it may be used to interpret the survey data by finding a best fit for filament and sheet sizes—and perhaps their size distributions—and detecting departures from Poisson processes. For analyses using 'counts in cells' for other surveys, see Efstathiou [8] and Szapudi *et al.* [19]. A hierachy of *N*-point correlation functions needed to represent clustering of galaxies in a complete sense was provided by White [23] and he provided explicit formulas, including their continuous limit.

2. SPATIAL STOCHASTIC PROCESSES

The basic random model for spatial stochastic processes representing the distribution of galaxies in space is that arising from a Poisson process of mean density n galaxies per unit volume in a large box—the region covered by the catalogue data to be studied. Then, the probability of finding exactly m galaxies in a *given sample region of volume v* is

$$P_m = \frac{(nv)^m}{m!} e^{-nv} \qquad \text{for} \quad m = 0, 1, 2, \ldots \tag{1}$$

The Poisson probability distribution (1) has mean equal to its variance, $\langle m \rangle = \mathrm{Var}(m) = nv$, and this is used as a reference case for comparison with observational data. Complete sampling of the available space using cells of volume v will reveal clustering if the variance of local density over the cells exceeds n. Moreover, the covariance of density between cells encodes correlation information about the spatial process being observed. The correlation function [18] is the ratio of the covariance of density of galaxies in cells separated by distance r, divided by the variance of density for the chosen cells,

$$\xi(r) = \frac{\mathrm{Cov}(r)}{\mathrm{Cov}(0)} = \frac{\langle m(r_0)m(r_0 + r)\rangle}{\langle m \rangle^2} - 1 \tag{2}$$

In the absence of correlation, we expect $\xi(r)$ to be zero or at least decay rapidly to zero with the separation distance r. In practice, we find that $\xi(r)$ resembles an exponential decay for r not too large.

Another way to detect clustering is to use an increasing sequence v_1, $v_2, \ldots$ of sampling cell volumes; in the absence of correlation we expect that the variance of numbers found using these cells will be the average numbers of galaxies in them, $nv_1, nv_2, \ldots$, respectively. Suppose that a sampling cell of volume v_1 contains exactly k sampling cells of volume v_2; then Var_1, the variance of density of galaxies found using v_1, is expressible as

$$\mathrm{Var}_1 = \frac{1}{k}\mathrm{Var}_2 + \frac{k-1}{k}\mathrm{Cov}_{1,2} \tag{3}$$

where Var_2 is the variance found using the smaller cells and $\mathrm{Cov}_{1,2}$ is the average covariance among the smaller cells in the larger cells. As $k \to \infty$, so $(1/k)\,\mathrm{Var}_2 \to 0$ and Var_1 tends to the mean covariance among *points* inside the v_1 cells. Now, the mean covariance among points inside v_1 cells is the expectation of the covariance between pairs of points separated by distance r, taken over all possible values for r inside a v_1 cell. Explicitly

$$\mathrm{Var}_1 = \int_0^D \mathrm{Cov}(r)\, b(r)\, dr \tag{4}$$

where b is the probability density function for the distance r between pairs of points chosen independently and at random in a v_1 cell and D is the diameter or maximum dimension of such a cell.

Ghosh [10] gave some examples of different functions b and some analytic results are known for covariance functions arising from spatial point processes—by representing the clusters as 'smoothed out' lumps of matter—see ref. 3 for the case of arbitrary rectangles in planar situations. It is convenient to normalize Eq. (4) by division through by $\mathrm{Cov}(0) = \mathrm{Var}(0)$, which is known for a Poisson process; this gives the 'between-cell' variance for complete sampling using v_1 cells. Then we obtain

$$\mathrm{Var}_1 = \mathrm{Var}(0)\int_0^D \alpha(r) b(r)\, dr \tag{5}$$

where α is the point autocorrelation function for the particular type of lumps of matter being used to represent a cluster of galaxies; typically, $\alpha(r) \approx e^{-r/d}$ for 'small' r and d is of the order of the smallest dimension of a cluster. Since it involves finite cells, Var_1 is in principle measurable, so (5) can be compared with observational data once the type of sampling cell and representative extended matter object are chosen. We return to this in the sequel and provide examples for a two-dimensional model. From Labini *et al.* [14] we note that experimentally for clusters of galaxies

$$\xi(r) \approx \left(\frac{25}{r}\right)^{1.7} \qquad \text{with } r \text{ in } \mathrm{h}^{-1}\ \mathrm{Mpc} \tag{6}$$

which for $2 < r < 10$ resembles $e^{-r/d}$ for suitable d near 1.8.

3. GALACTIC CLUSTER SPATIAL PROCESSES

From the atlases shown in Fairall [9] and surveys discussed by Labini *et al.* [14], one may estimate in a planar slice a representative galactic 'wall' filament thickness of about 5 h^{-1} Mpc and a wall 'thickness-to-length' aspect ratio A in the range $10 < A < 50$. Then, in order to represent galactic clustering as a Poisson process of wall filaments of length λ and width ω, we need the point autocorrelation function α for such filaments. In two dimensions it was shown in ref. 3 that the function α is given in three parts for rectangles of length λ and width ω by the following.

For $0 < r \leq \omega$

$$\alpha_1(r) = 1 - \frac{2}{\pi}\left(\frac{r}{\lambda} + \frac{r}{\omega} - \frac{r^2}{2\omega\lambda}\right) \tag{7}$$

For $\omega < r \leq \lambda$

$$\alpha_2(r) = \frac{2}{\pi}\left(\arcsin\frac{\omega}{r} - \frac{\omega}{2\lambda} - \frac{r}{\omega} + \sqrt{\left(\frac{r^2}{\omega^2} - 1\right)}\right) \tag{8}$$

For $\lambda < r \leq \sqrt{(\lambda^2 + \omega^2)}$

$$\alpha_3(r) = \frac{2}{\pi}\left(\arcsin\frac{\omega}{r} - \arccos\frac{\lambda}{r} - \frac{\omega}{2\lambda} - \frac{\lambda}{2\omega} - \frac{r^2}{2\lambda\omega} + \sqrt{\left(\frac{r^2}{\lambda^2} - 1\right)} + \sqrt{\left(\frac{r^2}{\omega^2} - 1\right)}\right) \tag{9}$$

For small r, as expected even in three dimensions, $\alpha(r) \approx e^{-2r/\pi\omega}$.

Note that for random *squares* of side length s, $\omega = \lambda = s$ and we have only two cases:

For $0 < r \leq s$

$$\alpha_1(r) = 1 - \frac{2}{\pi}\left(\frac{2r}{s} - \frac{r^2}{2s^2}\right) \tag{10}$$

For $s < r \leq \sqrt{(2s^2)}$

$$\alpha_3(r) = \frac{2}{\pi}\left(\arcsin\frac{s}{r} - \arccos\frac{s}{r} - 1 - \frac{r^2}{2s^2} + 2\sqrt{\left(\frac{r^2}{s^2} - 1\right)}\right) \tag{11}$$

This case may be used to represent in two dimensions clusters of galaxies as a Poisson process of smoothed-out squares of matter—the sheetlike elements of Doroshkevich *et al.* [7].

Next we need b, the probability density function for the distance r between pairs of points chosen independently and at random in a suitable inspection cell. From ref. 10 for square inspection cells of side length x, for $0 \leq r \leq x$

$$b(r, x) = \frac{4r}{x^4}\left(\frac{\pi x^2}{2} - 2rx + \frac{r^2}{2}\right) \tag{12}$$

and for $x \leq r \leq D = \sqrt{2}x$

$$b(r, x) = \frac{4r}{x^4}\left(x^2\left(\arcsin\left(\frac{x}{r}\right) - \arccos\left(\frac{x}{r}\right)\right) + 2x\sqrt{(r^2 - x^2)} - \frac{1}{2}(r^2 + 2x^2)\right) \tag{13}$$

A plot of this function is given in Fig. 1.

Ghosh [10] gave also the form of b for other types of cells; for arbitrary rectangular cells those expressions can be found in ref. 3. It is of interest to note that for small values of r, so $r \ll D$, the formulas for plane convex cells of area A and perimeter P all reduce to

$$b(r, A, P) = \frac{2\pi r}{A} - \frac{2Pr^2}{A^2}$$

which would be appropriate to use when the filaments are short compared

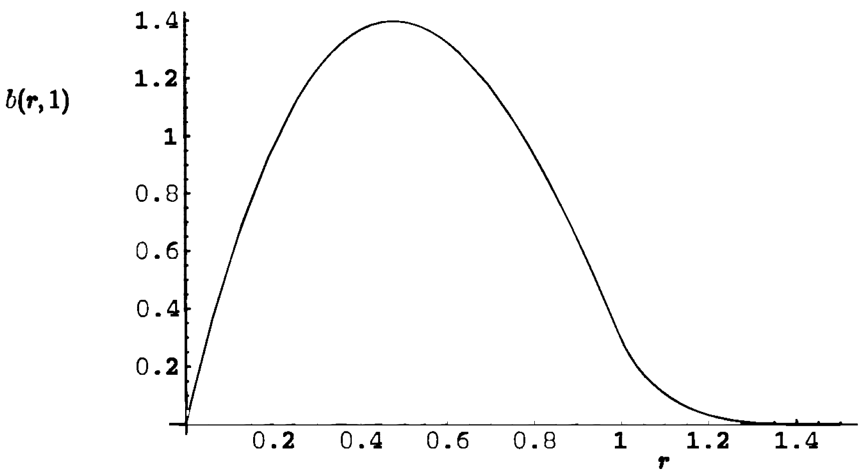

Fig. 1. Probability density function $b(r, 1)$ for the distance r between two points chosen independently and at random in a unit square.

with the dimensions of the cell. The filaments are supposed to be placed at random and independently in the plane and hence their variance contributions can be summed in a cell to give the variance for zonal averages—that is, the between-cell variance for complete sampling schemes. So, the variance between cells is the expectation of the covariance function, taken over all possible pairs of points in the cell, as given in (5). We rewrite this for square cells of side length x as

$$\mathrm{Var}(x) = \mathrm{Var}(0) \int_0^{\sqrt{2}x} \alpha(r)\, b(r, x)\, dr \tag{14}$$

Using this equation, in Fig. 2 we plot Var(x)/Var(0) against inspection cell size x h^{-1} Mpc for the case of filaments with width $\omega = 5$ h^{-1} Mpc and length $\lambda = 100$ h^{-1} Mpc. Note that Var(x) is expressible also as an integral of the (point) power spectrum over wavelength interval $[x, \infty)$ and that Landy *et al.* [15] detected evidence of a strong peak at 100 h^{-1} Mpc in the power spectrum of the Las Campanas Redshift Survey (cf. also Lin *et al.* [17]).

These spatial statistical models may be used in two distinct ways. If observational data are available for Var(x) for a range of x values, for example, by digitizing catalogue data on 2-dimensional slices, then attempts may be made to find the best fit for λ and ω. That would give statistical estimates of filament sizes on the presumption that the underlying process of filaments is Poisson. On the other hand, given observed $\mathrm{Var}^{\mathrm{obs}}(x)$ for a range of x, the variance ratio of this to (14)

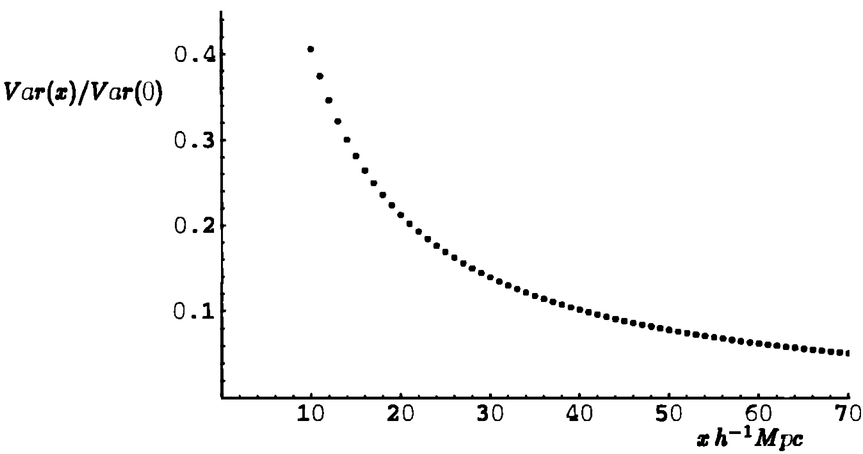

Fig. 2. Plot of the relative between-cell variance for a planar Poisson process of filaments with width $\omega = 5$ h^{-1} Mpc and length $\lambda = 100$ h^{-1} Mpc sampled using square cells of side length x h^{-1} Mpc from Eq. (14).

$$VR(x) = \frac{\text{Var}^{\text{obs}}(x)}{\text{Var}(x)} \tag{15}$$

will be an increasing function of x if there is a tendency of the filaments to cluster.

According to the Las Campanas Redshift Survey, some 40% of galaxies out to 500 h^{-1} Mpc are contained in filaments and the remainder in 'sheets,' which we may interpret perhaps as squares, both apparently following a Poisson process. Such a composite spatial structure may be represented easily with our model if the individual Poisson processes are independent; then the net variance for any choice of inspection cells is the weighted sum of the variances for the individual processes. So the between-cell variance (14) becomes a weighted sum of integrals, using the appropriate α functions for the constituent representative lumps of matter—perhaps squares for sheets and two kinds of filaments, dense and light.

4. GAMMA MODELS FOR INTERGALACTIC VOIDS

A family of parametric statistical models was developed in ref. 5 for the probability density function for intergalactic void volume V, including the Poisson process for galaxies as a special case. There are of course many choices for such families, but we chose one based on gamma distributions that had been successful in modeling void size distributions in terrestrial stochastic porous media [6]. The family of gamma distributions has event space $\Omega = \mathbb{R}^+$, parameters $\mu, \beta \in \mathbb{R}^+$, and probability density functions given by

$$f(V; \mu, \beta) = \left(\frac{\beta}{\mu}\right)^{\beta} \frac{V^{\beta-1}}{\Gamma(\beta)} e^{-V\beta/\mu} \tag{16}$$

Then $\bar{V} = \mu$ and Var $(V) = \mu^2/\beta$ and we see that μ controls the mean of the distribution while the spread and shape are controlled by $1/\beta$, the square of the coefficient of variation.

The special case $\beta = 1$ corresponds to the situation when V represents the random or Poisson process (1); then the distribution of void volumes is exponential with $\mu = 1/n$. The family of gamma distributions (16) can model a range of stochastic processes corresponding to nonindependent 'clumped' events for $\beta < 1$ and dispersed events for $\beta > 1$ as well as the random case [4, 6]. Figure 3 shows a family of gamma distributions, all of unit mean, with $\beta = 0.5, 1, 2, 5$.

Shannon's information-theoretic 'entropy' or 'uncertainty' for such stochastic processes [e.g., 12] is given, up to a factor, by the negative of the expectation of the logarithm of the probability density function (16), that is,

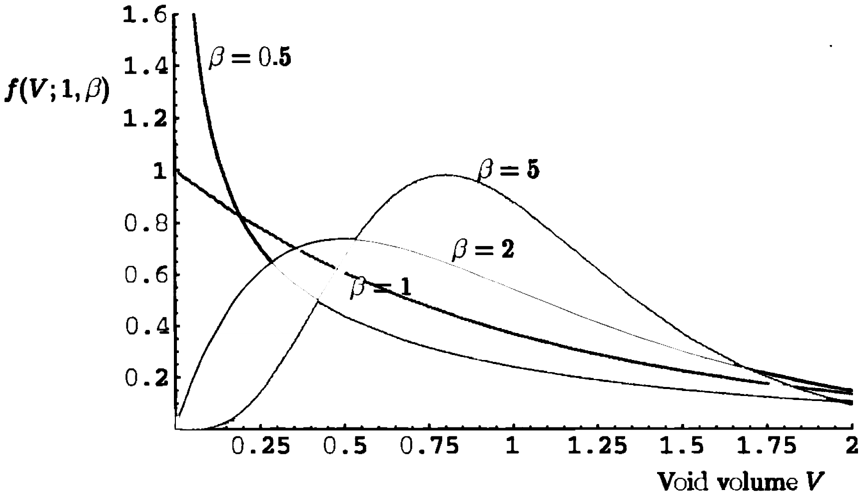

Fig. 3. Probability density functions $f(V; \mu, \beta)$ from (16) for gamma distributions representing the intergalactic void volume V with unit mean $\mu = 1$, and $\beta = 0.5, 1, 2, 5$. The case $\beta = 1$ corresponds to a 'random' distribution from an underlying Poisson process of galaxies, $\beta < 1$ corresponds to clustering, and $\beta > 1$ corresponds to dispersion.

$$S_f(\mu, \beta) = -\int_0^{\infty} \log(f(V; \mu, \beta) f(V; \mu, \beta)\, dx \tag{17}$$

$$= \beta + (1 - \beta)\frac{\Gamma'(\beta)}{\Gamma(\beta)} + \log\frac{\mu\Gamma(\beta)}{\beta} \tag{18}$$

In particular, at unit mean, the maximum entropy (or maximum uncertainty) occurs at $\beta = 1$, which is the random case, and then $S_f(\mu, 1) = 1 + \log \mu$.

The 'maximum likelihood' estimates $\hat{\mu}, \hat{\beta}$ of μ, β can be expressed in terms of the mean and mean logarithm of a set of independent observations $X = \{X_1, X_2, \ldots, X_n\}$. These estimates are obtained in terms of the properties of X by maximizing the 'log-likelihood' function [4, 5] with the following result

$$\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{19}$$

$$\log\hat{\beta} - \frac{\Gamma'(\hat{\beta})}{\Gamma(\hat{\beta})} = \overline{\log X} - \log\bar{X} \tag{20}$$

where $\overline{\log X} = (1/n)\,\Sigma_{i=1}^{n} \log X_i$.

The usual Riemannian information metric on the parameter space $\mathcal{S} = \{(\mu, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+\}$ is given by

$$ds^2_{\mathcal{S}} = \frac{\beta}{\mu^2}\,d\mu^2 + \left(\psi'(\beta) - \frac{1}{\beta}\right)d\beta^2 \qquad \text{for} \quad \mu, \beta \in \mathbb{R}^+ \tag{21}$$

where $\psi(\beta) = \Gamma'(\beta)/\Gamma(\beta)$ is the digamma function, the logarithmic derivative of the gamma function. For more details see refs. 16 and 4. The one-dimensional subspace parametrized by $\beta = 1$ corresponds to the available 'random' processes and the 'length' of any path in $\mathcal{S}$ is given via (21); locally, minimal paths in $\mathcal{S}$ are given by the geodesics defined by (21). The Gaussian curvature of the surface $\mathcal{S}$ is

$$K_{\mathcal{S}}(\mu, \beta) = \frac{3 - 4\,\beta\psi'(\beta) - \beta^2\psi''(\beta)}{2\,\beta(-1 + \beta\psi'(\beta))^2} \qquad \text{for} \quad \mu, \beta \in \mathbb{R}^+ \tag{22}$$

$$K_{\mathcal{S}}(\mu, \beta) \to -1 \qquad \text{as} \quad \beta \to 0 \tag{23}$$

$$K_{\mathcal{S}}(\mu, \beta) \to -2 \qquad \text{as} \quad \beta \to \infty \tag{24}$$

Some examples of geodesic sprays in the vicinities of the points

$$(\mu, \beta) = (1, 0.5), (1, 1), (1, 2)$$

are shown in Fig. 4.

The potential benefit of geometrizing the parameter space of our statistical models lies in its provision of standard Riemannian metric structure for applying variational methods and flows to represent evolutionary dynamics. Then statistical physics is translated into geometrical language with all of the normal vector calculus, using a metric that is derived from well-tried, information-theoretic, maximum likelihood principles—which is appropriate for fitting observational data.

5. GALAXY CLUSTER STATISTICS

We make an attempt here to model the cluster statistics by making use of the parametric models for the intergalactic void statistics described above.

Observationally, in a region where there are found large voids we would expect to find a lower local density of galaxies and vice versa. Then the two random variables, local void volume V and local density of galaxies n, are presumably inversely related. Many choices are possible; we take a simple functional form using an exponential and normalize the local density of galaxies to be bounded above by 1. Denoting the random variable representing this normalized local density by ν, we put

$$\nu(V) = e^{-v} \tag{25}$$

This assumption provides through (16) a probability density function for ν parametrized by μ, β and given by

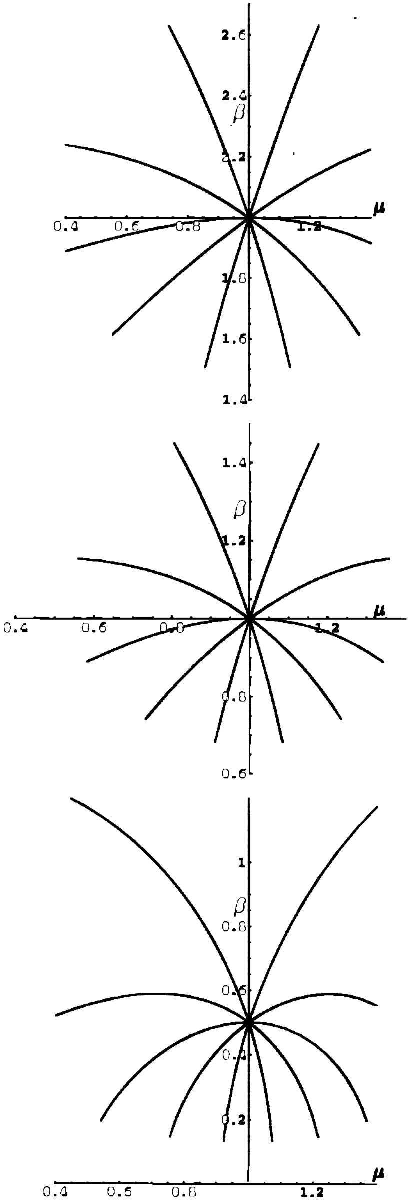

Fig. 4. Geodesic sprays in the gamma manifold, radiating from the points with unit mean $\mu = 1$, and $\beta = 0.5, 1, 2$ increasing vertically. The case $\beta = 1$ corresponds to an exponential distribution from an underlying Poisson process of galaxies; $\beta < 1$ corresponds to galactic clustering and β increasing aboved 1 corresponds to the opposite process, dispersion leading to greater uniformity. The Riemannian metric is given by $ds_S^2 = \frac{\beta}{\mu^2}\, d\mu^2 + \left(\psi'(\beta) - \frac{1}{\beta}\right) d\beta^2$.

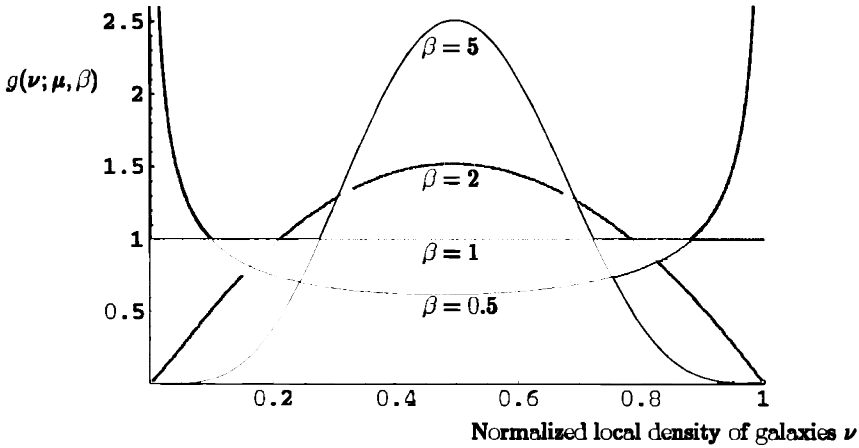

Fig. 5. Probability density function $g(\nu;\ \mu,\ \beta)$ from (26) for distributions representing the normalized local density of galaxies, $\nu \in [0, 1]$, with central mean $\langle\nu\rangle = 0.5$ and $\beta = 0.5$, 1, 2, 5. The case $\beta < 1$ corresponds to clustering in the underlying spatial process of galaxies; conversely, $\beta > 1$ corresponds to dispersion.

$$g(\nu;\ \mu,\ \beta) = \left(\frac{\beta}{\mu}\right)^{\beta} \frac{\nu^{\beta/\mu - 1}}{\Gamma(\beta)} \left|\log \nu\right|^{1-\beta} \tag{26}$$

This distribution for local galactic number density has mean $\langle\nu\rangle$ and variance $\mathrm{Var}(\nu)$ given by

$$<\nu> = \left(\frac{\beta}{\beta + \mu}\right)^{\beta} \tag{27}$$

$$\mathrm{Var}\ (\nu) = \left(\frac{\beta}{\beta + 2\mu}\right)^{\beta} - \left(\frac{\beta}{\beta + \mu}\right)^{2\beta} \tag{28}$$

Figure 5 shows the distribution (26) for mean normalized density $\langle\nu\rangle = 0.5$ and $\beta = 0.5$, 1, 2, 5. Note that as $\beta \to 1$, the distribution (26) tends to the uniform distribution. For $\beta < 1$ we have clustering in the underlying process, with the result that the population has high- and low-density peaks. Other choices of functional relationship between local void volume and local density of galaxies would lead to different distributions; for example, $\nu(V) = \exp(-V^k)$ for $k = 2, 3, \ldots$, would serve. However, the persisting qualitative observational feature that would discriminate among the parameters is the prominence of a central modal value—indicating a smoothed or dispersed structure, or the prominence of high- and low-density peaks—indicating clustering.

REFERENCES

1. A. Cappi, S. Maurogordato, and M. Lachiéze-Rey, A scaling law in the distribution of galaxy clusters, *Astron. Astrophys.* **243** (1991) 28–32.
2. P. Coles, Understanding recent observations of the large-scale structure of the universe, *Nature* **346** (1990) 446–447.
3. C. T. J. Dodson, Spatial variability and the theory of sampling in random fibrous networks, *J. R. Stat. Soc.* **33** (1971) 88–94.
4. C. T. J. Dodson, Gamma manifolds and stochastic geometry, in Proceedings of the Workshop on Recent Topics in Differential Geometry, Santiago de Compostela 16-19 July 1997, *Public. Dept. Geom. Topol.* **89** (1998) 85–92.
5. C. T. J. Dodson, A geometrical representation for departures from randomness of the intergalactic void probability function, Presented at NATO Advanced Study Institute, Workshop on Statistics of Cosmological Data Sets, Isaac Newton Institute, Cambridge 8–13 August 1999.
6. C. T. J. Dodson and W. W. Sampson, Modeling a class of stochastic porous media, *Appl. Math. Lett.* **10** (1997) 87–89.
7. A. G. Doroshkevich, D. L. Tucker, A. Oemler, R. P. Kirshner, H. Lin, S. A. Shectman, S. D. Landy, and R. Fong, Large- and superlarge-scale structure in the Las Campanas Redshift Survey, *Mon. Not. R. Astr. Soc.* **283** (1996) 1281–1310.
8. G. Efstathiou, Counts-in-cells comparisons of redshift surveys, *Mon. Not. R. Astr. Soc.* **276** (1995) 1425–1434.
9. A. P. Fairall, *Large-Scale Structure in the Universe*, Wiley-Praxis, Chichester, England (1998).
10. B. Ghosh, Random distances within a rectangle and between two rectangles, *Calcutta Math. Soc.* **43** (1951) 17–24.
11. S. Ghigna, S. Borgani, M. Tucci, S. A. Bonometto, A. Klypin, and J. R. Primack, Statistical tests for CHDM and lambda CDM cosmologies, *Astrophys. J.* **479** (1997) 580–591.
12. E. T. Jaynes, Information theory and statistical inference, *Phys. Rev.* **106** (1957) 620–630; **108** (1957) 171–190.
13. F. Sylos Labini, A. Gabrielli, M. Montuori, and L. Pietronero, Finite size effects on the galaxy number counts: Evidence for fractal behavior up to the deepest scale, *Physica A* **226** (1996) 195–242.
14. F. Sylos Labini, M. Montuori, and L. Pietronero, Scale invariance of galaxy clustering, *Phys. Rep.* **293** (1998) 61–226.
15. S. D. Landy, S. A. Shectman, H. Lin, R. P. Kirshner, A. A. Oemler, and D. Tucker, Two-dimensional power spectrum of the Las Campanas redshift survey: Detection of excess power on 100 h^{-1} Mpc scales, *Astrophys. J.* **456** (1996) L1–7.
16. S. L. Lauritzen, Statistical manifolds, in *Differential Geometry in Statistical Inference*, Institute of Mathematical Statistics, Berkeley (1987), pp 163–218.
17. H. Lin, R. P. Kirshner, S. A. Schectman, S. D. Landy, A. Oemler, D. L. Tucker, and P. L. Schechter, The power spectrum of galaxy clustering in the Las Campanas Redshift Survey, *Astrophys. J.* **471**, (1996) 617–635.
18. P. J. E. Peebles, *Large Scale Structure of the Universe*, Princeton University Press, Princeton, New Jersey (1980).
19. I. Szapudi, A. Meiksin, and R. C. Nichol, Higher order statistics from the Edinburgh Durham Southern Galaxy Catalogue Survey. 1. Counts in cells, *Astrophys. J.* **473** (1996) 15–2l.
20. M. Tribus, R. Evans, and G. Crellin, The use of entropy in hypothesis testing, in *Proceedings Tenth National Symposium on Reliability and Quality Control 7-9 January 1964*; cf. M. Tribus, *Thermostatics and Thermodynamics*, Van Nostrand, Princeton, New Jersey (1961).

21. M. S. Vogeley, M. J. Geller, C. Park, and J. P. Huchra, Voids and constraints on nonlinear clustering of galaxies, *Astron. J.* **108** (1994) 745–758.
22. R. van der Weygaert, Quasi-periodicity in deep redshift surveys, *Mon. Not. R. Astr. Soc.* **249** (1991) 159–163.
23. S. D. M. White, The hierarchy of correlation functions and its relation to other measures of galaxy clustering, *Mon. Not. R. Astr. Soc.* **186** (1979) 145–154.