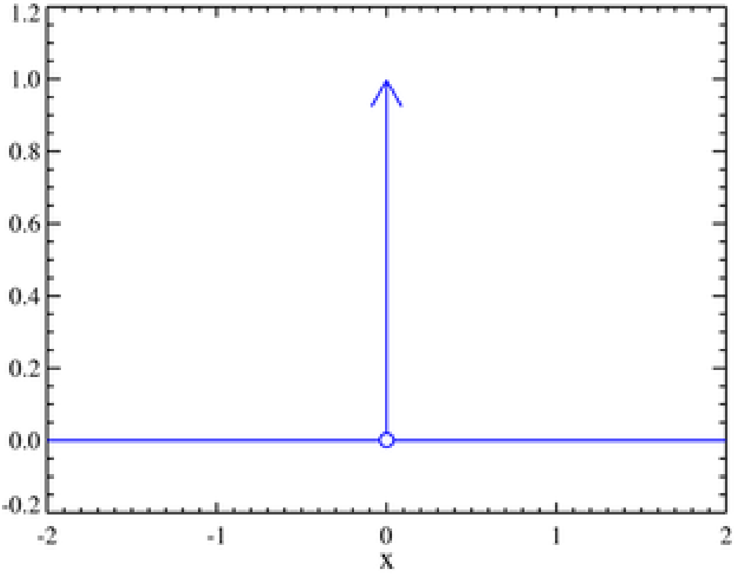
Recent Math Problems Arising in Statistics involving Generalized Functions, Group Representations, Inequalities, and the Riemann Zeta Function

Arthur Berg

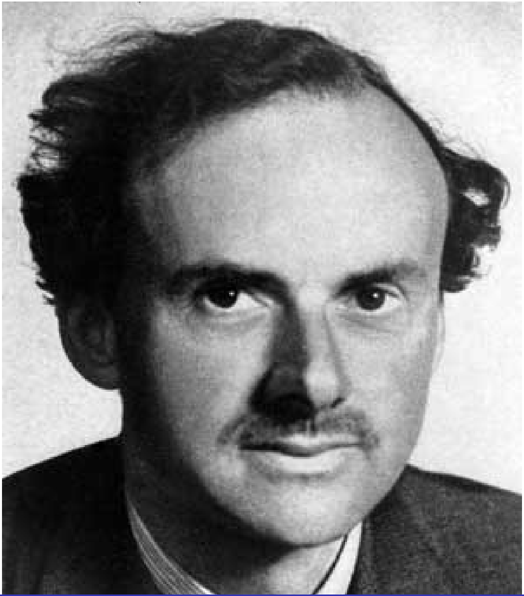
Department of Statistics, University of Florida

July 2, 2008

Part I: Generalized Functions



Paul Dirac (1902-1984)



Sergi Sobolov (1908-1989) and Laurent Schwartz (1915-2002)

Schwartz Distribution

Let

$$f : \mathbb{R} \to \mathbb{R}$$

be a locally integrable function and let

$$\varphi : \mathbb{R} \to \mathbb{R}$$

be infinitely differentiable with compact support.
Set

$$\langle f, \varphi \rangle = \int_{\mathbb{R}} f(x)\phi(x)\, dx$$

f can therefore be viewed as a continuous linear functional on test functions ϕ.
The derivative of the distribution S is naturally defined to be

$$\langle S', \varphi \rangle \triangleq -\langle S, \varphi' \rangle$$

Fourier Transform and Tempered Distributions

The Fourier transform can be defined on *tempered* distributions.
Test functions are from the Schwartz space — infinitely differentiable rapidly decreasing functions—$\forall n, m$

$$\sup_{x\in\mathbb{R}} |x^n \varphi^{(m)}(x)| < \infty$$

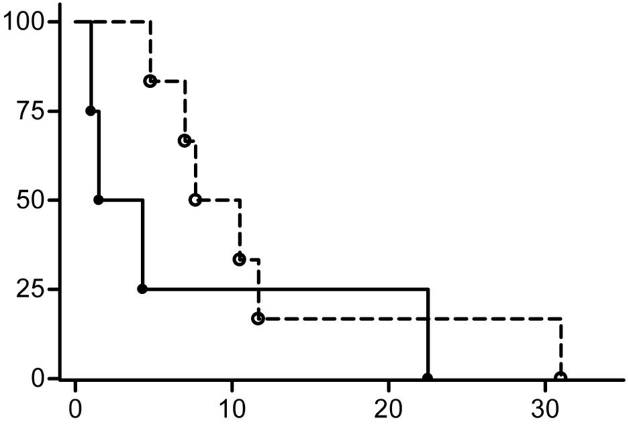
$$\check{\varphi}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi(s) e^{-isx}\, ds$$

$$\langle \check{S}, \varphi \rangle = \langle S, \check{\varphi} \rangle$$

Empirical Distribution Function and Survival Function

Let $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} F(x)$. The survival function, $S(x) = 1 - F(x)$, has estimate

$$\hat{S}(t) = 1 - \hat{F}(t) = 1 - \frac{1}{n}\sum_{j=1}^{n} I(X_i \leq t)$$



Bias of Smoothed EDF

$$\hat{F}_h(t) = \int_{-\infty}^{t} \hat{f}_h(x)\,dx = \frac{1}{n}\sum_{j=1}^{n} \bar{K}\left(\frac{t-X_j}{h}\right)$$

where $\bar{K}(t) = \int_{-\infty}^{t} K(x)\,dx$.

$$\mathrm{E}\left[\hat{F}_h(t)\right] = \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}\left[\bar{K}\left(\frac{t-X_i}{h}\right)\right],$$

$$\begin{aligned}
\mathrm{E}\left[\bar{K}\left(\frac{t-X_i}{h}\right)\right] &= \int_{-\infty}^{\infty} \bar{K}\left(\frac{t-x}{h}\right) f(x)\,dx \\
&= \int_{-\infty}^{\infty} \bar{K}\left(\frac{t-x}{h}\right)\,dF(x) \\
&= \underbrace{\bar{K}\left(\frac{t-x}{h}\right) F(x)\Big|_{x=-\infty}^{x=\infty}}_{=0} + \frac{1}{h}\int F(x) K\left(\frac{t-x}{h}\right)\,dx \\
&= F \star K_h(t)
\end{aligned}$$

where $K_h(t) = \frac{1}{h} K\left(\frac{t}{h}\right)$.

Fourier Transform of a CDF

Note that

$$F(t) = \int_{-\infty}^{t} f(x)\,dx = \int_{-\infty}^{\infty} f(x)H(t-x)\,dx = f \star H(t)$$

where $H(x)$ is the Heaviside step function given by $H(x) = 1(x > 0)$. Therefore

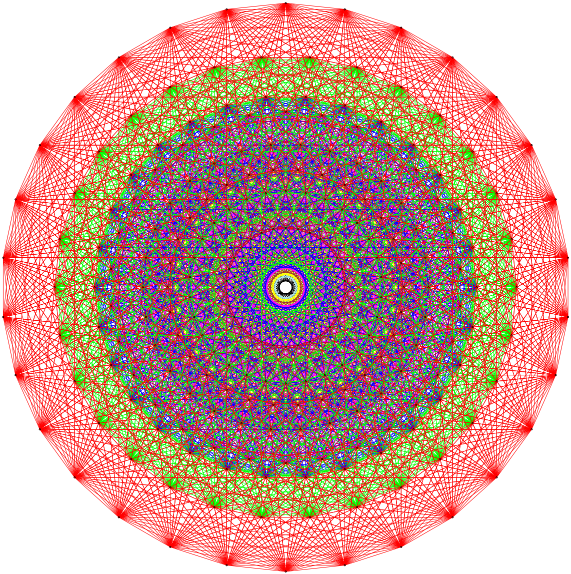
$$\begin{aligned} \mathcal{F}\left(F(t)\right) &= \phi(s)\left(\pi\delta(s) + \frac{1}{is}\right) \\ &= \pi\phi(0)\delta(s) + \frac{\phi(s)}{is} \\ &= \pi\delta(s) + \frac{\phi(s)}{is}. \end{aligned}$$

where $\phi(s)$ is the characteristic function (the Fourier transform of f).

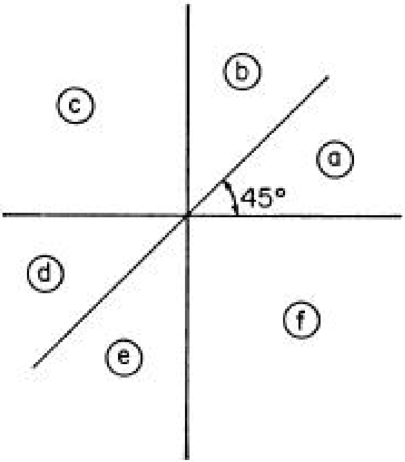
Bias Calculation Continued

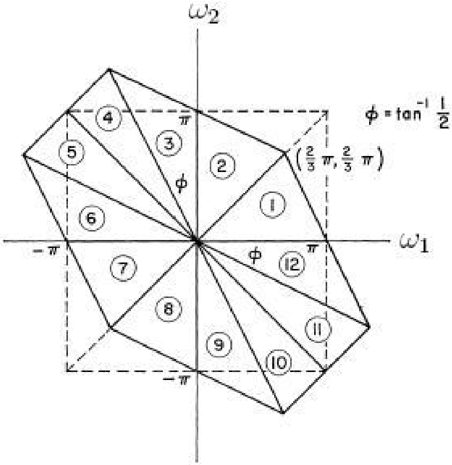
$$
\begin{aligned}
\text{bias}\left(\hat{F}_h(t)\right) &= K_h \star F(t) - F(t) \\
&= \mathcal{F}\left(\mathcal{F}^{-1}\left(K_h \star F(t) - F(t)\right)\right) \\
&= \mathcal{F}\left(\mathcal{F}^{-1}\left(K_h\right) \cdot \mathcal{F}^{-1}\left(F\right) - \mathcal{F}^{-1}\left(F\right)\right) \\
&= \mathcal{F}\left(\left(\mathcal{F}^{-1}\left(K_h\right) - 1\right) \mathcal{F}^{-1}\left(F\right)\right) \\
&= \mathcal{F}\left(\left(\kappa(sh) - 1\right)\left(\pi\delta(s) + \frac{\phi(s)}{is}\right)\right) \\
&= \mathcal{F}\left(\left(\kappa(sh) - 1\right)\frac{\phi(s)}{is}\right) - \pi\mathcal{F}\left(\left(\kappa(sh) - 1\right)\delta(s)\right) \\
&= \mathcal{F}\left(\left(\kappa(sh) - 1\right)\frac{\phi(s)}{is}\right) - \underbrace{\pi\mathcal{F}\left(\left(\kappa(sh) - 1\right)\Big|_{s=0}\right)}_{=0} \\
&= \frac{1}{2\pi}\int_{|s|>1/h}\left(\kappa(sh) - 1\right)\frac{\phi(s)}{is}\,ds.
\end{aligned}
$$

Part II: Group Representations



Symmetries of the Bivariate ACF and Bispectrum





Symmetries of ACF

One variable: $C(x) = C(-x)$

Two variables:

$$\begin{aligned} C(x,y) &\overset{1}{=} C(x,y) \\ &\overset{2}{=} C(-x, y-x) \\ &\overset{3}{=} C(y,x) \\ &\overset{4}{=} C(x-y, -y) \\ &\overset{5}{=} C(-y, x-y) \\ &\overset{6}{=} C(y-x, -x) \end{aligned}$$

Eq. 2 + Eq. 3 $\Rightarrow$ Eq. 6

$$C(x,y) = C(y,x) = C(-x, y-x) \quad \text{suffices}$$

$\star$ We wish 1) understand the symmetries and 2) symmetrize a non-symmetric function.

Constructing the Symmetries from Permutations

⋆Each equation corresponds to a permutation.

For simplicity, assume $\mathrm{E}[X_t] = 0$.

The equation corresponding to the permutation $\sigma = (12)$ is

$$\begin{aligned} C(x,y) &= \mathrm{E}[X_t X_{t+x} X_{t+y}] \\ &\overset{(12)}{=} \mathrm{E}[X_{t+x} X_t X_{t+y}] \\ &= \mathrm{E}[X_t X_{t-x} X_{t-x+y}] \\ &= C(-x, y-x) \end{aligned}$$

A Group Representation

Define $\psi : \mathbb{R}^3 \to \mathbb{R}^2$ given by $\psi(a,b,c) \mapsto (b-a, c-a)$. Take $\sigma = (12)$, then

$$(x,y) \longrightarrow (0,x,y) \xrightarrow{\sigma} (x,0,y) \xrightarrow{\psi} (-x, y-x) \longrightarrow \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}$$

$$e \longleftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \longleftrightarrow C(x,y) \qquad (13) \longleftrightarrow \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} \longleftrightarrow C(x-y,-y)$$

$$(12) \longleftrightarrow \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \longleftrightarrow C(-x,y-x) \qquad (123) \longleftrightarrow \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \longleftrightarrow C(-y,x-y)$$

$$(23) \longleftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \longleftrightarrow C(y,x) \qquad (132) \longleftrightarrow \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \longleftrightarrow C(y-x,-x)$$

Suppose $\sigma = (12)$ and $\tau = (13)$, then $\gamma = (132) = \sigma\tau$ and

$$\rho(\gamma) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} = \rho(\sigma)\rho(\tau)$$

Implications of the Representation

Theorem (Berg 2007)

The mapping $\rho : S_n \to \mathrm{GL}_{n-1}(\mathbb{R})$, described above, is a faithful group representation.

1. Symmetrizing lag-windows – generalizing current constructions

   Current construction: $\tilde{f}(x,y) = f(x)f(y)f(x-y)$

   $$\tilde{f}(x,y) = h(f(x,y), f(-x,y-x), f(y,x), f(x-y,-y), f(-y,x-y), f(y-x,-x))$$

   where h is any symmetric function of its six arguments.

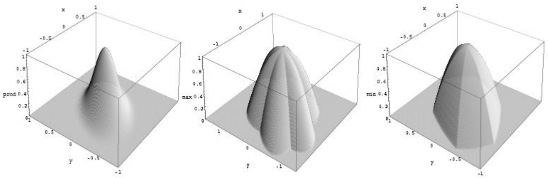


Figure: $\tilde{f}$ with $h = \prod x_i$, $h = \max(x_i)$, $h = \min(x_i)$, and $f(x,y) = (1 - x^2 - y^2)^+$.

Implications of the Representation II

2. Generalization of the Gabr-Rao optimal window

$$\Lambda_{\text{opt}}(\boldsymbol{\omega}) = \alpha \left(1 - \beta \left(\sum_{i=1}^{k-1} \omega_i^2 + \sum_{i<j} \omega_i \omega_j\right)\right)^+$$

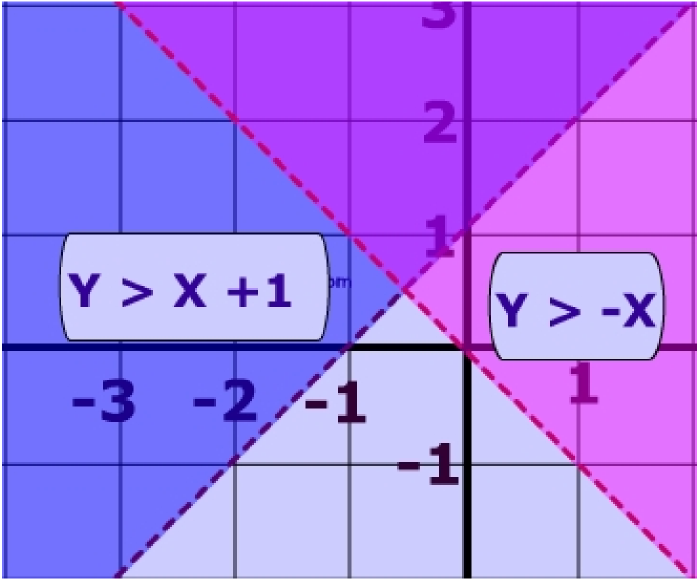
Theorem (Berg 2007)

Let $\Lambda(\boldsymbol{\omega})$ be any nonnegative kernel that integrates to one and satisfies all the necessary symmetries. Also assume

$$\int_{\mathbb{R}^{k-1}} \omega_j^2 \Lambda(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \int_{\mathbb{R}^{k-1}} \omega_j^2 \Lambda_{\text{opt}}(\boldsymbol{\omega})\, d\boldsymbol{\omega}$$

for $j = 1, \ldots, n-1$. Then $\|\Lambda\|_{L_2} \geq \|\Lambda_{\text{opt}}\|_{L_2}$.

Part III: Inequalities



Agresti Example

	Lung Cancer	Heat Disease
Nonsmokers	p_1=.0001	p_3=.00413
Smokers	p_2=.0014	p_4=.00669

	Lung Cancer	Heat Disease
Relative Risk	14	1.62
Odds Ratio	14.02	1.62
Risk Difference	.00130	.00256

Formulas

$$\mathrm{RR} = \frac{p_2}{p_1} \qquad \mathrm{OR} = \frac{p_2/(1-p_2)}{p_1/(1-p_1)} \qquad \mathrm{RD} = p_2 - p_1$$

Inequality Problem

A General Result

Suppose $p_1, p_2, p_3, p_4 \in (0,1)$ satisfy $p_1 + p_4 \leq 1$ or $p_2 + p_3 \leq 1$. Show that

$$\frac{p_2}{p_1} < \frac{p_4}{p_3} \qquad (i)$$

and

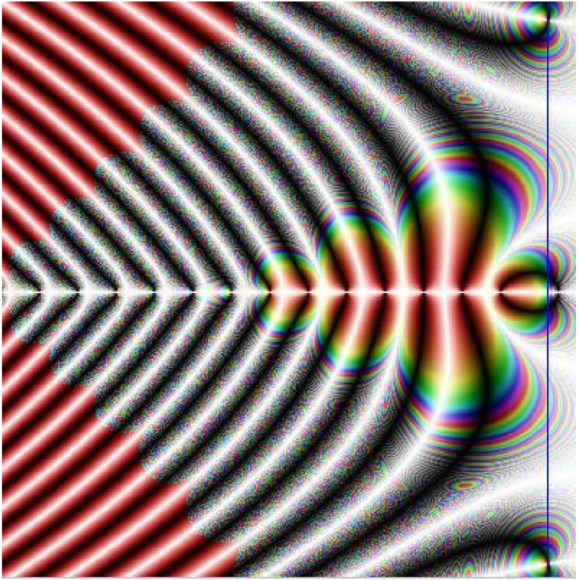
$$p_2 - p_1 < p_4 - p_3 \qquad (ii)$$

imply

$$\frac{p_2/(1-p_2)}{p_1/(1-p_1)} < \frac{p_4/(1-p_4)}{p_3/(1-p_3)} \qquad (iii)$$

And also (i) and (ii) imply (iii) when $<$ is replaced with $>$ in each of the inequalities.

Part IV: Riemann Zeta Function



Power Law Distribution

Let $X_1, \ldots, X_n$ from a power law distribution

$$f(x) = f_\beta(x) = \frac{c}{x^\beta}; \qquad x = 1, 2, \ldots$$

for some $\beta > 1$; note that $c = c_\beta = \frac{1}{\zeta(\beta)}$ where ζ is the Riemann zeta function.

Consider the problem of estimating β.

Define

$$H_n(x) = n\hat{f}(x) = \sum_{i=1}^{n} 1[X_i = x].$$

Natural estimate of β given from a linear regression on

$$\{(\log x, \log H_n(x)) : x = 1, 2, \ldots\},$$

Optimal Least Squares

An optimal estimate of β is a particular linear combination of $\hat{c}$ and $\hat{\beta}$. Supposing the regression routine produced the covariance matrix

$$\begin{pmatrix} \hat{\tau}_{11} & \hat{\tau}_{12} \\ \hat{\tau}_{21} & \hat{\tau}_{22} \end{pmatrix}$$

for the covariance of

$$\begin{pmatrix} \hat{\beta} \\ \hat{c} \end{pmatrix}$$

Note that

$$\beta = \beta(c) = \zeta^{-1}\left(1/c\right) \quad \text{and} \quad \beta'(c) = \frac{-1}{c^2 \zeta'\left(\zeta^{-1}\left(1/c\right)\right)}$$

Basing an estimate of β on $\tilde{\beta} = \zeta^{-1}(1/\hat{c})$, we have a correlated pair $\begin{pmatrix} \hat{\beta} \\ \tilde{\beta} \end{pmatrix}$ of asymptotically unbiased estimates of β with covariance efficiently estimated by

$$\begin{pmatrix} \hat{\tau}_{11} & \tilde{\tau}_{12} \\ \tilde{\tau}_{21} & \tilde{\tau}_{22} \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ 0 & \beta'(\hat{c}) \end{pmatrix} \begin{pmatrix} \hat{\tau}_{11} & \hat{\tau}_{12} \\ \hat{\tau}_{21} & \hat{\tau}_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \beta'(\hat{c}) \end{pmatrix}$$

Optimal Least Squares II

There is an optimal convex combination $\hat{\beta}^* = a\hat{\beta} + (1-a)\hat{\beta}$, with a given by

$$a = \frac{\tilde{\tau}_{22} - \tilde{\tau}_{12}}{\hat{\tau}_{11} - 2\tilde{\tau}_{12} + \tilde{\tau}_{22}},$$

Main Theorem

Theorem (Abramson, Berg, Meyers 2008)

Let $\beta > 1$ *be fixed and* $X_1, X_2, \ldots \overset{iid}{\sim} f(x) = cx^{-\beta}(x = 1, 2, \ldots)$*. Define* $H_n(x) = \sum_{i=1}^{n} 1\,[X_i = x]$ *and* $M_{nk} = M_n = \min\{x : H_n(x) \leq k\}$ *where* k *is any nonnegative integer. Provided a sequence* y_n *satisfies*

$$y_n e^{-ncy_n^{-\beta}} \longrightarrow 0 \qquad (\star)$$

as $n \to \infty$*, it follows that* $\Pr[M_n > y_n] \to 1$ *as* $n \to \infty$*.*

In estimating rates for y_n, we would like to approximate the difference between

$$\zeta(s) \qquad \text{and} \qquad \int_1^\infty x^{-s}\,dx$$

And show

$$\zeta(2) - 1 = \frac{\pi^2}{6} - 1 = \sup_{s \in (1,2]} \left[\zeta(s) - \int_1^\infty x^{-s}\,dx\right] \qquad (\star\star)$$