On Gauss' charaterization of the normal distribution

Adelchi Azzalini

Università di Padova, Italia

joint work with Marc G. Genton
to appear in *Bernoulli* . . . sometime

The essence of the problem

- consider a simple random sample $x_1, \ldots, x_n$ from a certain population
- the following fact is well known:
  assume the parent population is $N(\mu, \sigma^2)$
  $\Longrightarrow$ MLE of μ is the sample mean
- now reverse the components:
  assume that MLE of μ is the sample mean
  $\stackrel{?}{\Longrightarrow}$ the parent population is $N(\mu, \sigma^2)$
- hence a characterization problem

A bit more in detail

- original formulation/solution in Gauss (1809)
- subsequent work of Teicher (1961)
- some conditions needed, specifically
  - the mean must be MLE for all samples of size $n = 2, 3$
  - some (weak) regularity conditions must hold
- extension to multivariate case by Marshall & Olkin (1993), and Stadje (1993)
- alternative approach by Hürlimann (1998)

Univariate location family

Notation & assumptions:

- consider location family with density $f(x-\mu)$
- a simple random sample of size n, $n \geq 3$, is drawn
- assume f differentiable and f' continuous at least at one point
- for each set of sample values $x_1, \ldots, x_n$, the sample mean $\bar{x} = \sum_i x_i/n$ is a solution of the likelihood equation

Statement:
$f(x) = \phi(x; \sigma^2)$, the $N(0, \sigma^2)$ density, for some $\sigma^2 > 0$

Univariate location family. Proof

Proof.

Denote $g(x) = \frac{\mathrm{d}}{\mathrm{d}x} \log f(x)$. Lik.eqn. is $\sum_{i=1}^{n} g(x_i - \mu) = 0$.

- Consider a sample with $(u, u, \ldots, u)$ for some u; $\bar{x} = u$. Since the solution of lik.eqn.

$$n\, g(u - \mu) = 0$$

  is $\hat{\mu} = u$, then $g(0) = 0$.

- Consider the sample $(2u, 0, u, \ldots, u)$; hence $\bar{x} = u$. Since the solution of lik.eqn.

$$g(u) + g(-u) + (n-2)\, g(0) = 0$$

  is $\hat{\mu} = u$, hence $g(-u) = -g(u)$.

Univariate location family. Proof (cntd)

Proof (ctd)

- Consider next a sample $(u, v, -(u+v), 0, \ldots, 0)$, such that $\hat{\mu} = 0$, which requires

$$g(u) + g(v) = g(u+v)$$

  i.e. Cauchy functional equation, leading to

$$g(x) = -c\,x$$

- integration of g gives

$$f(x) = d - \tfrac{1}{2}\,c\,x^2$$

  for some d; here $c > 0$ to ensure $\int f = 1$
- hence $f(x) = \phi(x; c^{-1})$. QED

Univariate location family. A counter-example

If $n = 2$, the statement fails.

- consider density

$$f(x) = \text{constant} \times e^{-x^2/2+w(x)}$$

  where w is an even function (hence w' is odd)
- then likelihood equation is

$$\sum_{i=1}^{n}(x_i - \mu) = \sum_{i=1}^{n} w'(x_i - \mu)$$

- if $n = 2$, then $\bar{x}$ is always a solution,
  even e. g. for $w(x) = \cos(x)$ and $f \neq \phi$

Multivariate location family

- for p-dimensional distribution, the argument is similar
- use appropriate generalisation of Cauchy equation, leading to

$$g(x) = -C\,x$$

  for $p \times p$ matrix C
- $\int g = \text{constant} - \frac{1}{2} x^\top C x$
- must be $C \geq 0$ for integrability $\Longrightarrow f(x)$ is $\phi_p(x; C^-)$
- however if C not positive definite, then continuous differentiability would not hold
- hence $f(x)$ is $\phi_p(x; C^{-1})$. QED

Multivariate location family with extra parameter

- Consider p-dimensional case with extra parameter θ, $\theta \in \mathbb{R}^q$
- assume that, for any fixed θ, $f(x;\theta)$ admits partial derivatives and the gradient is continuous at least at one point
- for any fixed θ, we can argue as before and conclude that $f(x;\theta)$ is of type $N_p(0, \Sigma(\theta))$

References

Gauss, C. F. (1809), *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, Perthes et Besser, Hamburg. English translation by C. H. Davis, reprinted by Dover, New York (1963).

Hald, A. (1998), *A History of Mathematical Statistics from 1750 to 1930*, New York: J. Wiley & Sons.

Hürlimann, W. (1998), On the Characterization of Maximum Likelihood Estimators for Location-Scale Families, *Communications in Statistics, Theory and Methods*, 27, 495-508.

Marshall, A. W., and Olkin, I. (1993), Maximum Likelihood Characterizations of Distributions, *Statistica Sinica*, 3, 157-171.

Stadje, W. (1993), ML Characterization of the Multivariate Normal Distribution, *Journal of Multivariate Analysis*, 46, 131-138.

Teicher, H. (1961), Maximum Likelihood Characterization of Distributions, *Annals of Mathematical Statistics*, 32, 1214-1222.