Bayes Days 2000 at LANL

Three-Day Minicourse on Bayesian Analysis in Physics

Lectures presented by

Prof. Volker Dose
Max Planck Institute for Plasma Physics

Sponsored by
Enhanced Surveillance Program, Los Alamos National Laboratory

For more information, look on the web:
http://public.lanl.gov/kmh/course/BD2000.html

Organized by Ken Hanson, DX-3, 505-667-1402, kmh@lanl.gov

Outlier tolerant Parameter Estimation

V. Dose and W. von der Linden

Los Alamos Nat. Lab
April 3 – 5, 2000

Data Pooling

		Failures	Successes	% Success
Exp. A	old	16519	4343	20.8 ± 0.3
	new	742	122	14.1 ± 1.1
Exp. B	old	3876	14488	78.9 ± 0.3
	new	1233	3907	76.0 ± 0.6
Pooled	old	20395	18831	48.0 ± 0.3
	new	1975	4029	67.1 ± 0.6

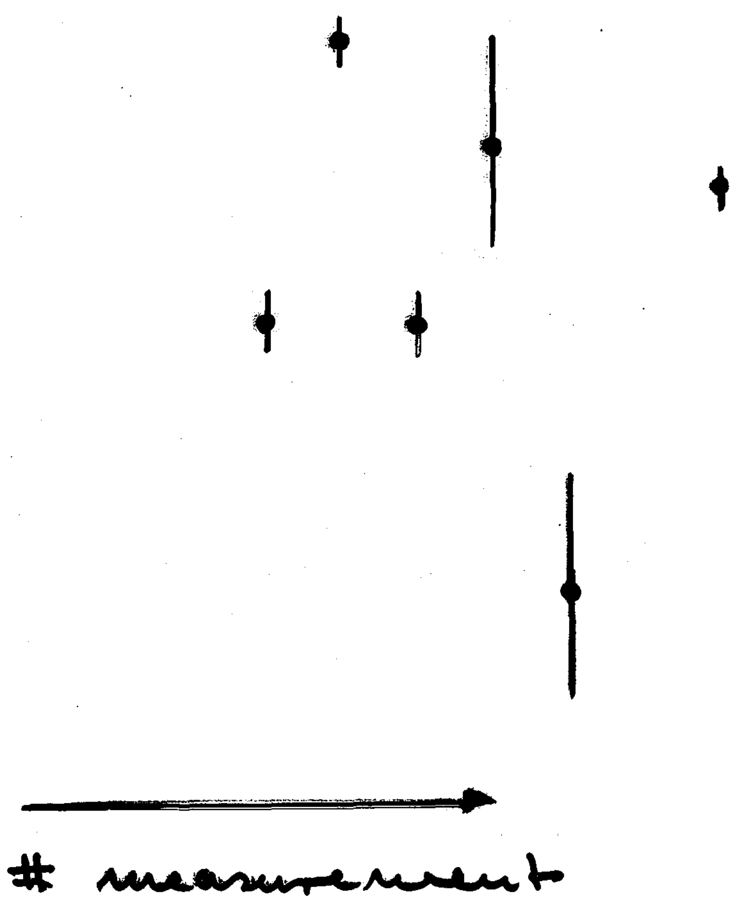
NUCLEAR FUSION, Vol.25, No.1 (1985)

GLOBAL ENERGY CONFINEMENT SCALING FOR NEUTRAL-BEAM-HEATED TOKAMAKS

S.M. KAYE, R.J. GOLDSTON
Plasma Physics Laboratory,
Princeton University,
Princeton, New Jersey,
United States of America

ABSTRACT. A total of 677 representative discharges from seven neutral-beam-heated tokamaks have been used to study the parametric scaling of global energy confinement time. Contributions to this data base were from Asdex, DITE, D-III, ISX-B, PDX, PLT and TFR, and were taken from results of gettered, L-mode type discharges. Assuming a power law dependence of τ_E on the discharge parameters κ, I_p, B_t, $\bar{n}_e$ P_{tot}, a and R, standard multiple linear regression techniques were used in two steps to determine the scaling. The results indicate that the discharges used in the study are well described by the scaling

$$\tau_E \propto \kappa^{0.28} B_T^{-0.09} I_p^{1.24} \bar{n}_e^{0.26} P_{tot}^{-0.58} a^{-0.49} R^{1.65} .$$



After a thorough analysis using a number of least-squares algorithms, the initial group of 38 items of stochastic input data was reduced to 22 items by deleting those that were either highly inconsistent with the remaining data or had assigned uncertainties so large that they carried negligible weight.

The 1986 CODATA Recommended Values of the Fundamental Physical Constants
(http://physics.nist.gov/cuu)

The large change in K_V (*K_V relates the SI unit of volt to a calibration standard*) and hence in many other quantities between 1973 and 1986 would have been avoided if two determinations of F (*the Faraday number*) which seemed to be discrepant with the remaining data had not been deleted in the 1973 adjustment.

The 1986 CODATA Recommended Values of the Fundamental Physical Constants (http://physics.nist.gov/cuu)

Also: Physics Today, August 1999, BG7

The generalized arithmetic mean

$$p(\vec{x}|\mu,\vec{\sigma},I) = \left\{\prod_i \frac{1}{\sigma_i\sqrt{2\pi}}\right\} \cdot \exp\left\{-\frac{1}{2}\sum_i \frac{(x_i-\mu)^2}{\sigma_i^2}\right\}$$

(A) True errors $\vec{\sigma}$ assumed to be known

Bayes theorem

$$p(\mu|\vec{x},\vec{\sigma},I) = \frac{p(\mu|I)p(\vec{x}|\mu,\vec{\sigma},I)}{p(\vec{x}|\vec{\sigma},I)}$$

Flat prior on μ

$$p(\mu|\vec{x},\vec{\sigma},I) \sim \exp\left\{-\frac{1}{2}\sum_i \frac{(x_i-\mu)^2}{\sigma_i^2}\right\}$$

$$\begin{aligned} \langle\mu\rangle &= \sigma_\mu^2 \cdot \sum_i x_i/\sigma_i^2 \\ \sigma_\mu^2 &= \left\{\sum_i 1/\sigma_i^2\right\}^{-1} \end{aligned}$$

σ_μ^2 is independent of $\{x_i\}$

(B) True errors unknown

Experimental data $\{x_i, s_i\}$

$$\begin{aligned} p(\vec{x}|\mu,\vec{s},I) &= \int p(\vec{x},\vec{\sigma}|\mu,\vec{s},I)d^N\sigma \\ &= \int p(\vec{\sigma}|\mu,\vec{s},I)p(\vec{x}|\mu,\vec{\sigma},\vec{s},I)d^N\sigma \end{aligned}$$

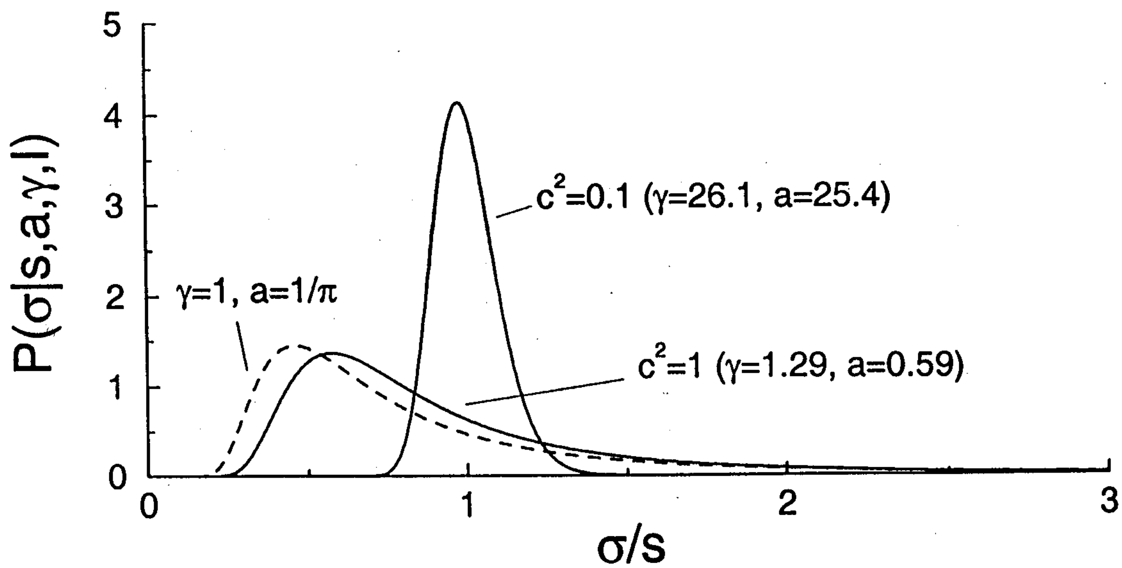
$$p(\sigma|s,a,\gamma,I) = 2\ \frac{a^\gamma}{\Gamma(\gamma)} \cdot \left(\frac{s}{\sigma}\right)^{2\gamma+1} e^{-a\frac{s^2}{\sigma^2}}\ \frac{d\sigma}{s}$$

Moments

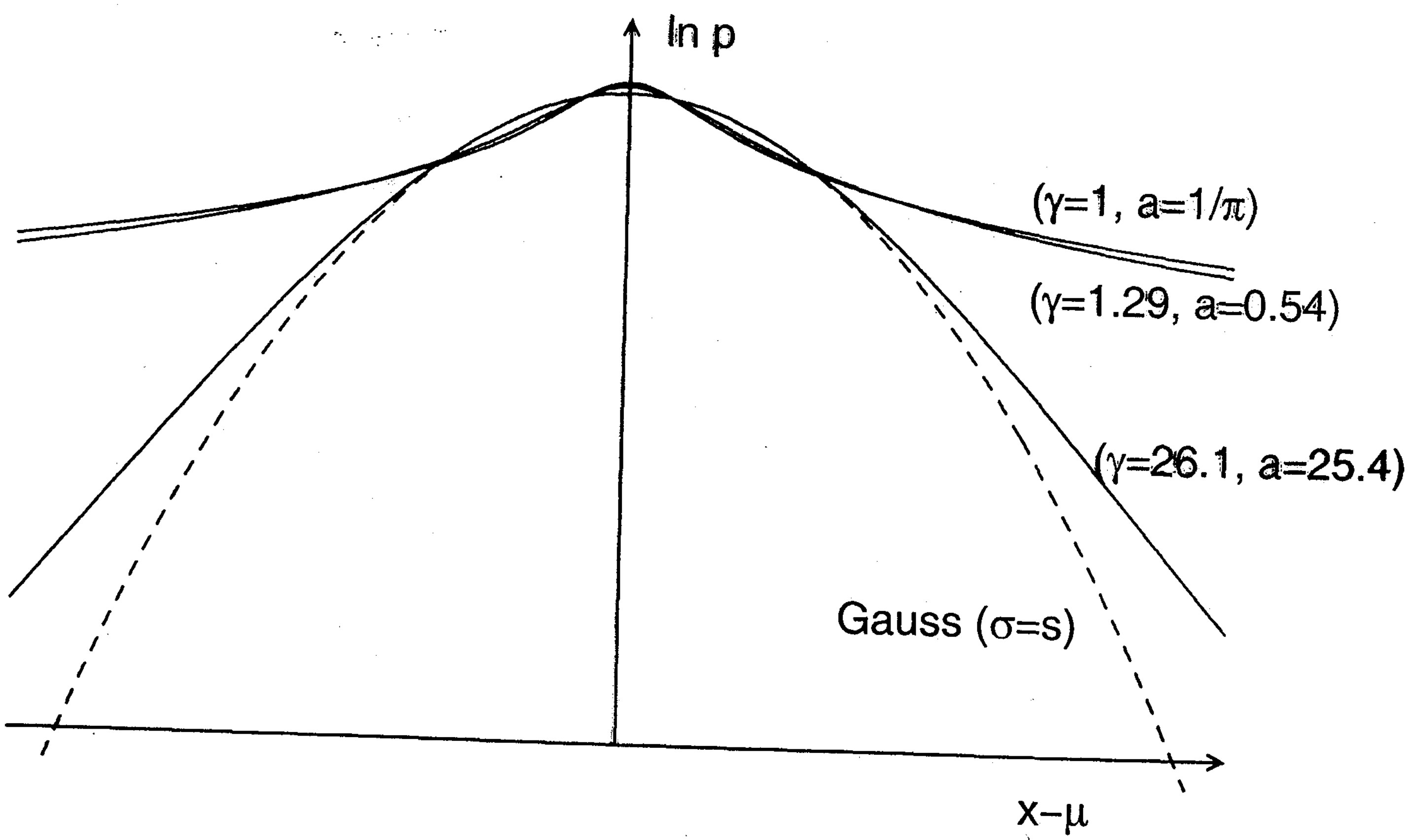
$$\left\langle \frac{\sigma}{s} \right\rangle = \sqrt{a}\ \Gamma(\gamma - 1/2)/\Gamma(\gamma) \stackrel{!}{=} 1$$

$$\left\langle \frac{\sigma^2}{s^2} \right\rangle = a\ \Gamma(\gamma - 1)/\Gamma(\gamma) = a/(\gamma - 1)$$

$$\left\langle \Delta\frac{\sigma^2}{s^2} \right\rangle = a/(\gamma - 1) - 1 \stackrel{!}{=} c^2$$



$$
\begin{aligned}
p(x|\mu, s, I) &= \frac{2}{\sqrt{2\pi}}\frac{a^\gamma}{\Gamma(\gamma)}\frac{1}{s}\int \frac{d\sigma}{\sigma}\left(\frac{s}{\sigma}\right)^{2\gamma+1} e^{-a\frac{s^2}{\sigma^2}-\frac{1}{2\sigma^2}(x-\mu)^2} \\
&= \frac{1}{s\sqrt{2\pi}}\,\frac{a^\gamma}{\Gamma(\gamma)}\,\frac{\Gamma(\gamma+1/2)}{\left\{a+(x-\mu)^2/2s^2\right\}^{\gamma+1/2}}
\end{aligned}
$$



Change of properties

$$p(\vec{x}|\mu,\vec{\sigma},I) = \left\{\prod_i \frac{1}{\sigma_i\sqrt{2\pi}}\right\} \prod_i \exp\left\{-\frac{1}{2\sigma_i^2}(x_i-\mu)^2\right\}$$

$$= k(\vec{\sigma}) \prod_i f\left(\frac{(x_i-\mu)^2}{\sigma_i^2}\right) = k(\vec{\sigma}) f\left(\sum_i \frac{(x_i-\mu)^2}{\sigma_i^2}\right)$$

$$p(\vec{x}|\mu,\vec{s},I) = \left\{\prod_i \frac{1}{s_i\sqrt{2\pi}}\right\} \prod_i \left\{\frac{a^\gamma}{\Gamma(\gamma)} \frac{\Gamma(\gamma+1/2)}{\left\{a+\frac{(x_i-\mu)^2}{2\sigma^2}\right\}^{\gamma+1/2}}\right\}$$

in general multimodal!

Posterior estimate of $\langle\sigma/s\rangle$

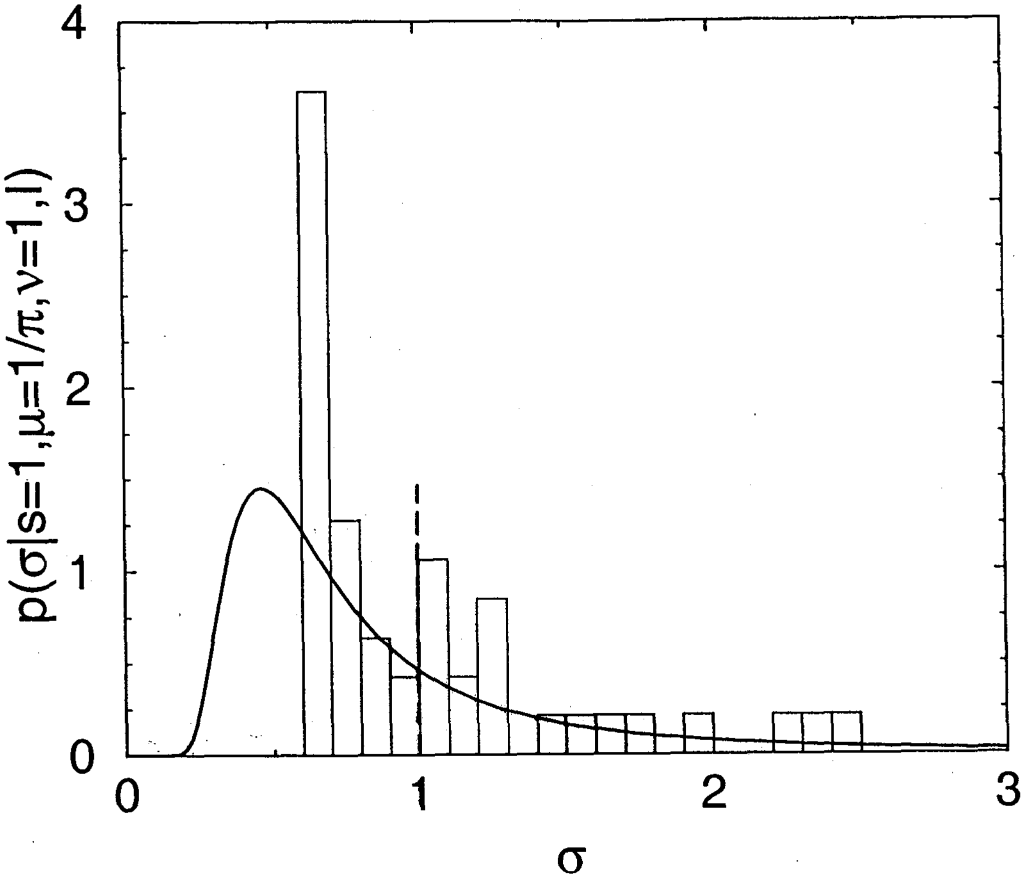
$$p(\vec{\sigma}|\vec{x},\vec{s},a,\gamma,I) = p(\vec{\sigma}|\vec{s},a,\gamma,I)p(\vec{x}|\vec{\sigma},\vec{s},a,\gamma,I)/p(\vec{x}|\vec{s},a,\gamma,I)$$

$$p(\vec{x}|\vec{\sigma},I) = \int p(\vec{x},\mu|\sigma,I)d\mu = \int p(\mu|I)p(\vec{x}|\mu,\vec{\sigma},I)d\mu$$

$$\boxed{\langle\sigma_k\rangle = s_k \frac{\Gamma(\gamma)}{\Gamma(\gamma+1/2)} \int \left\{a+\frac{(x_k-\mu)^2}{2s_k^2}\right\}^{1/2} \rho(\mu,\vec{x})d\mu/Z}$$

$$\rho(\mu,\vec{x}) = \prod_i \left\{a+\frac{(x_i-\mu)^2}{2s_i^2}\right\}^{-(\gamma+1/2)} p(\mu|I)$$

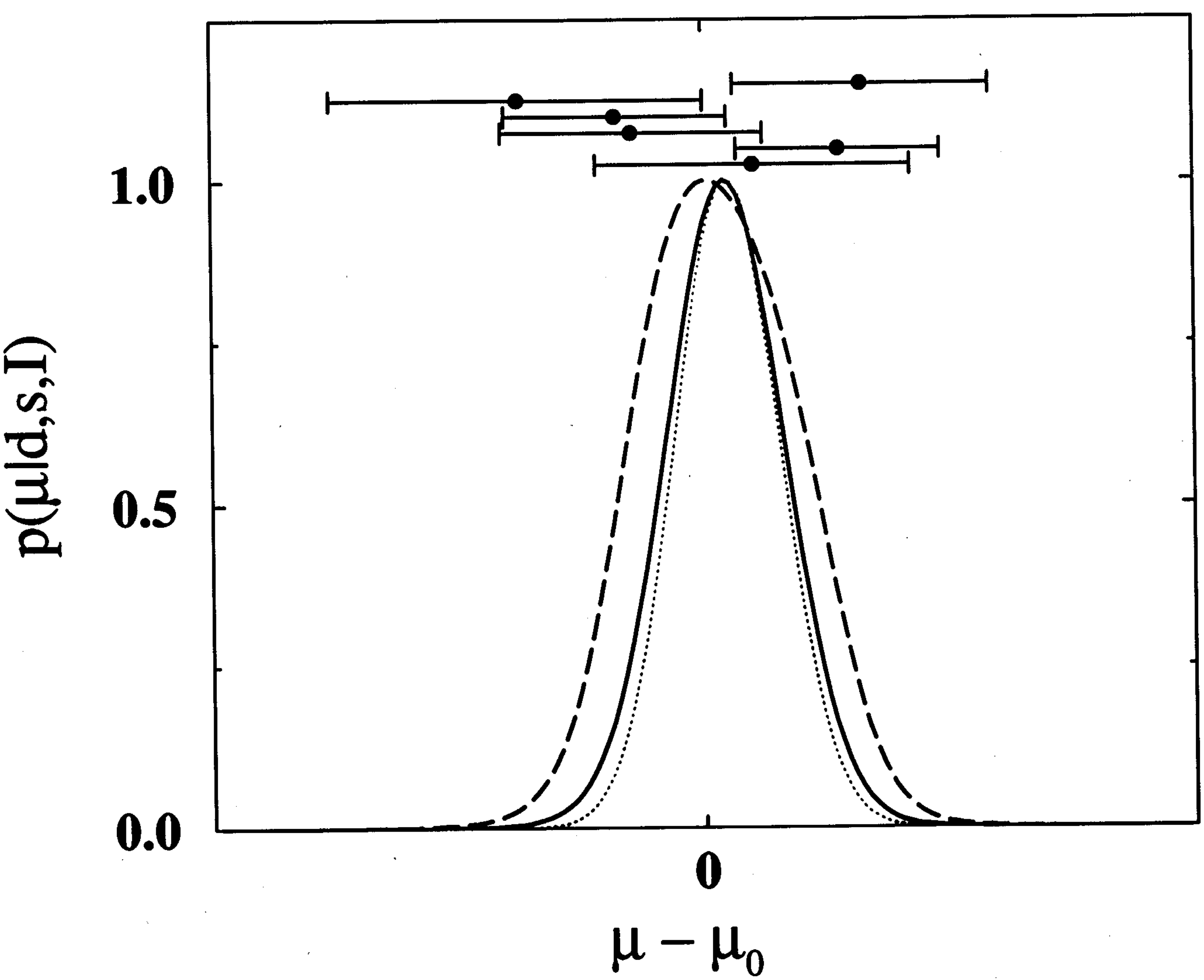
$$Z = \int d\mu\ \rho(\mu,\vec{x})$$



p(σ|s=1,μ=1/π,ν=1,I)
σ
0
1
2
3
4

Quantum Hall Effect

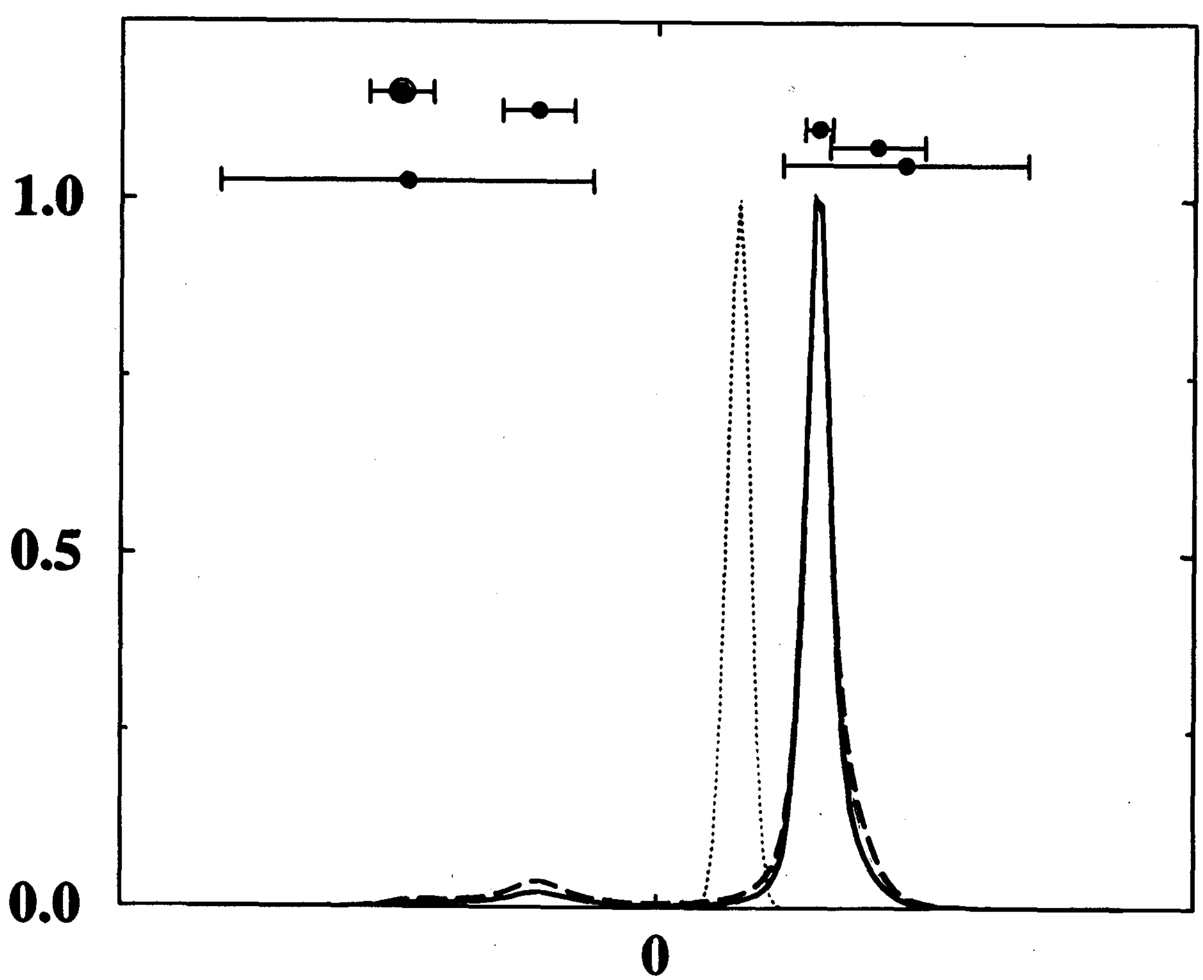
$h/e^2 = 25812.8056(12)\ \Omega$



Proton gyromagnetic ratio

$$\gamma_p = \mu_p / I_p = (\mu_p/\mu_n)\,(F/M_p)$$

$$\gamma_p = 26752.2128(81)\ [10^4 s^{-1} T^{-1}]$$



Of all the gyromagnetic ratio data, the most glaring discord comes from the NPL low-field value. The measurements of the proton resonance frequency were completed in December, 1975 after which the coil dimensions were measured, but no verification was made (by repeating the frequency measurements), that the measurement process did not affect the coils. Because this result is so discrepant, and because the measurements were forced to terminate prematurely, we consider it to be an incomplete effort which should not be included in the final adjustment.

Codata Bulletin 63 (1986) 8

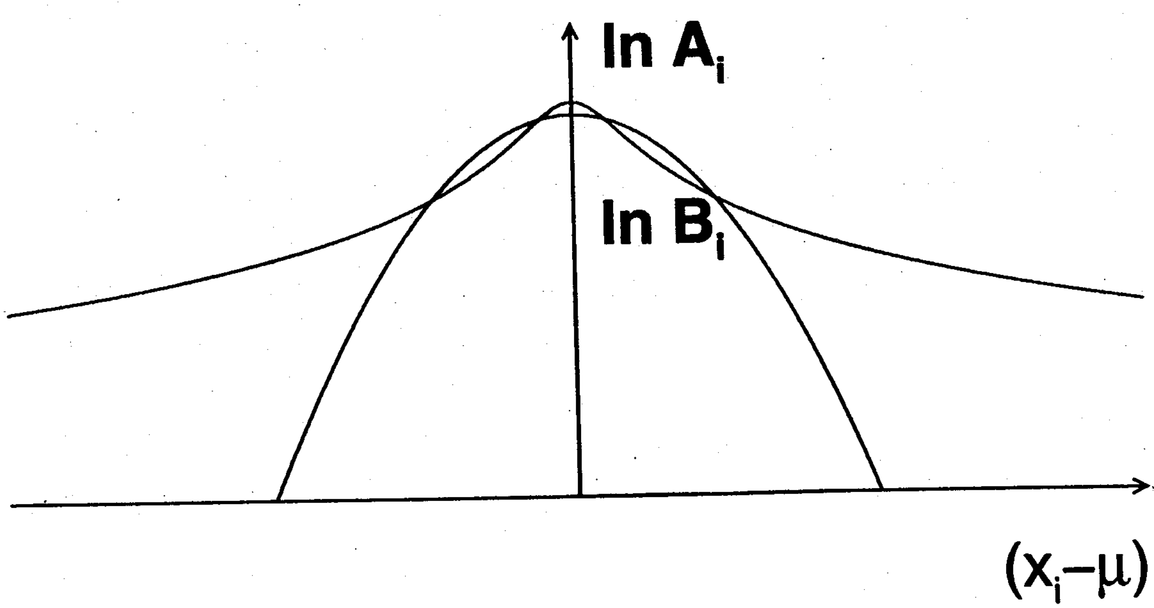
regular data and outlier

$$p(\sigma_i|s_i,a,\gamma,\beta,I) = \beta p(\sigma_i|s_i,a,\gamma,I) + (1-\beta)\delta(\sigma_i - s_i)$$

$$\begin{aligned} p(x_i|s_i,\mu,I) &= \int p(x_i,\sigma_i|\mu,I)d\sigma_i \\ &= \int p(\sigma_i|s_i,a,\gamma,\beta,I)p(x_i|\mu,\sigma_i,I)d\sigma_i \\ &= \beta A_i(\mu) + (1-\beta)B_i(\mu) \end{aligned}$$

$$A_i(\mu) = \frac{1}{s_i\sqrt{2\pi}}\frac{a^\gamma}{\Gamma(\gamma)}\frac{\Gamma(\gamma+1/2)}{\left\{a+(x_i-\mu)^2/2s_i^2\right\}^{\gamma+1/2}}$$

$$B_i(\mu) = \frac{1}{s_i\sqrt{2\pi}}\exp\left\{-\frac{1}{2s_i^2}(x_i-\mu)^2\right\}$$



β marginalization

$$p(\vec{x}|\mu, \vec{s}, a, \gamma, I) \sim \int d\beta p(\beta|I) \prod_{i=1}^{N} \{\beta A_i(\mu) + (1-\beta)B_i(\mu)\}$$

$$\prod_{i=1}^{N} \{\beta A_i + (1-\beta)B_i\} = \sum_{i=1}^{N} \binom{N}{i} \beta^i (1-\beta)^{N-i} C_i^{(N)}(\mu) = P_N$$

choose $p(\beta|I)$ flat in $0 \leq \beta \leq 1 \Rightarrow p(\beta|I) = 1$

$$\boxed{p(\vec{x}|\mu, \vec{s}, a, \gamma, I) \sim \frac{1}{N+1} \sum_{i=0}^{N} C_i^{(N)}(\mu)}$$

determination of $C_i^{(N)}$

$$P_N = P_{N-1}(A_N \beta + B_N(1-\beta))$$

$$C_i^N = \frac{i}{N} A_N C_{i-1}^{(N-1)} + \left(\frac{N-1}{N}\right) B_N C_i^{(N-1)}$$

Start:

$$P_1 = \beta A_1 + (1-\beta)B_1 = C_0^{(1)}(1-\beta) + C_1^{(1)}\beta$$

$$\boxed{C_0^{(1)} = B_1(\mu), \quad C_1^{(1)} = A_1(\mu)}$$

β distribution

Bayes theorem

$$p(\beta|\vec{x},\vec{s},a,\gamma,I) \sim p(\beta|I) \ \underbrace{p(\vec{x}|\vec{s},\beta,a,\gamma,I)}_{\int d\mu\ p(\mu,\vec{x}|\vec{s},\beta,a,\gamma,I)}$$

$$\boxed{p(\beta|\vec{x},\vec{s},a,\gamma,I) \sim \int d\mu \sum_{i=0}^{N} \binom{N}{i} C_i^N(\mu)\beta^i(1-\beta)^{N-i}}$$

Outlier identification

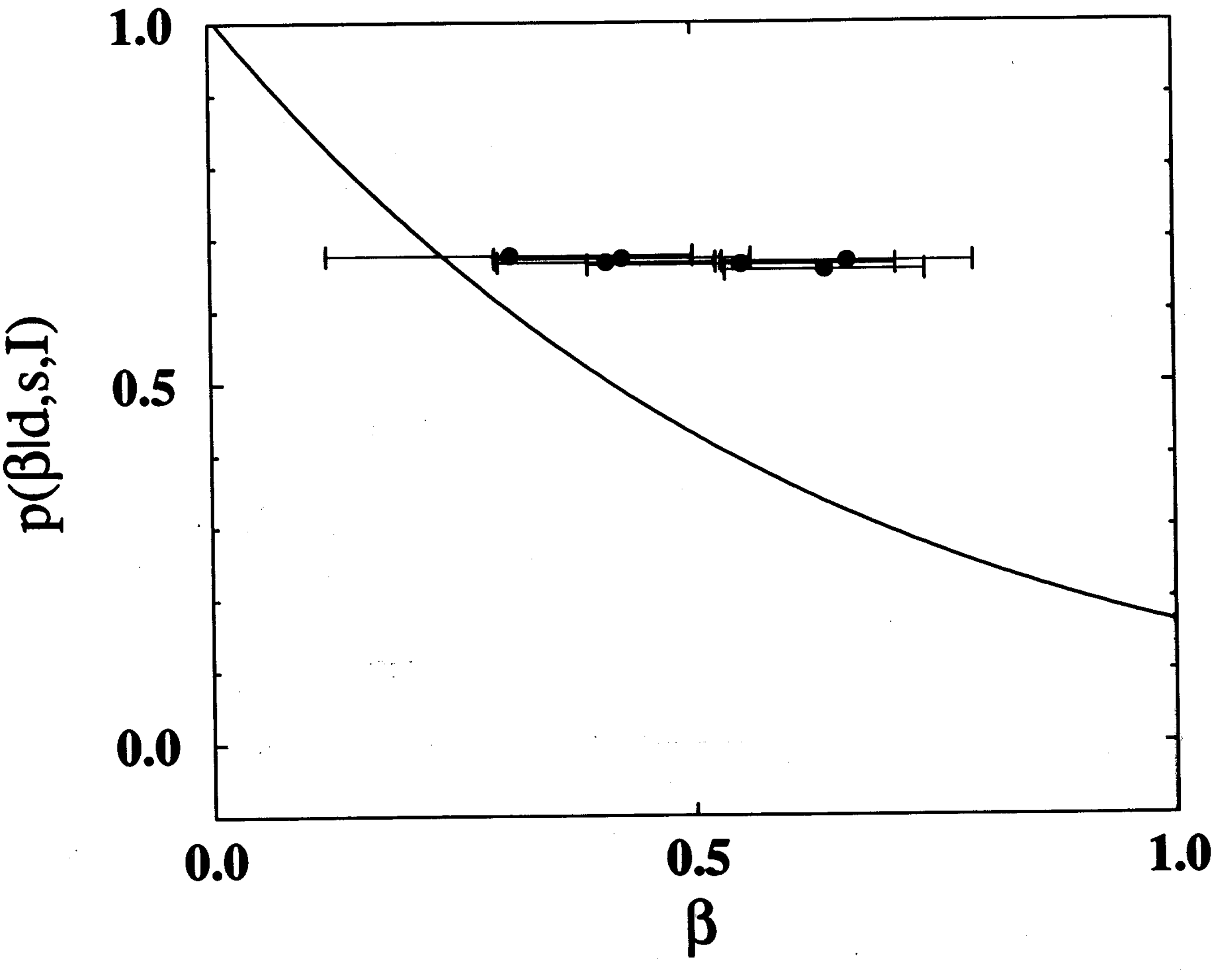
$$p(Q_k|\vec{x},\vec{s},a,\gamma,I) = \int p(\beta|I)\cdot p(Q_k|\vec{x},\vec{s},a,\gamma,\beta,I)\ d\beta$$

$$p(Q_k|\vec{x},\vec{s},a,\gamma,\beta,I) = \frac{p(Q_k|\vec{s},a,\gamma,\beta,I)\cdot p(\vec{x}|Q_k,\vec{s},a,\gamma,\beta,I)}{p(\vec{x}|\vec{s},a,\gamma,\beta,I)}$$

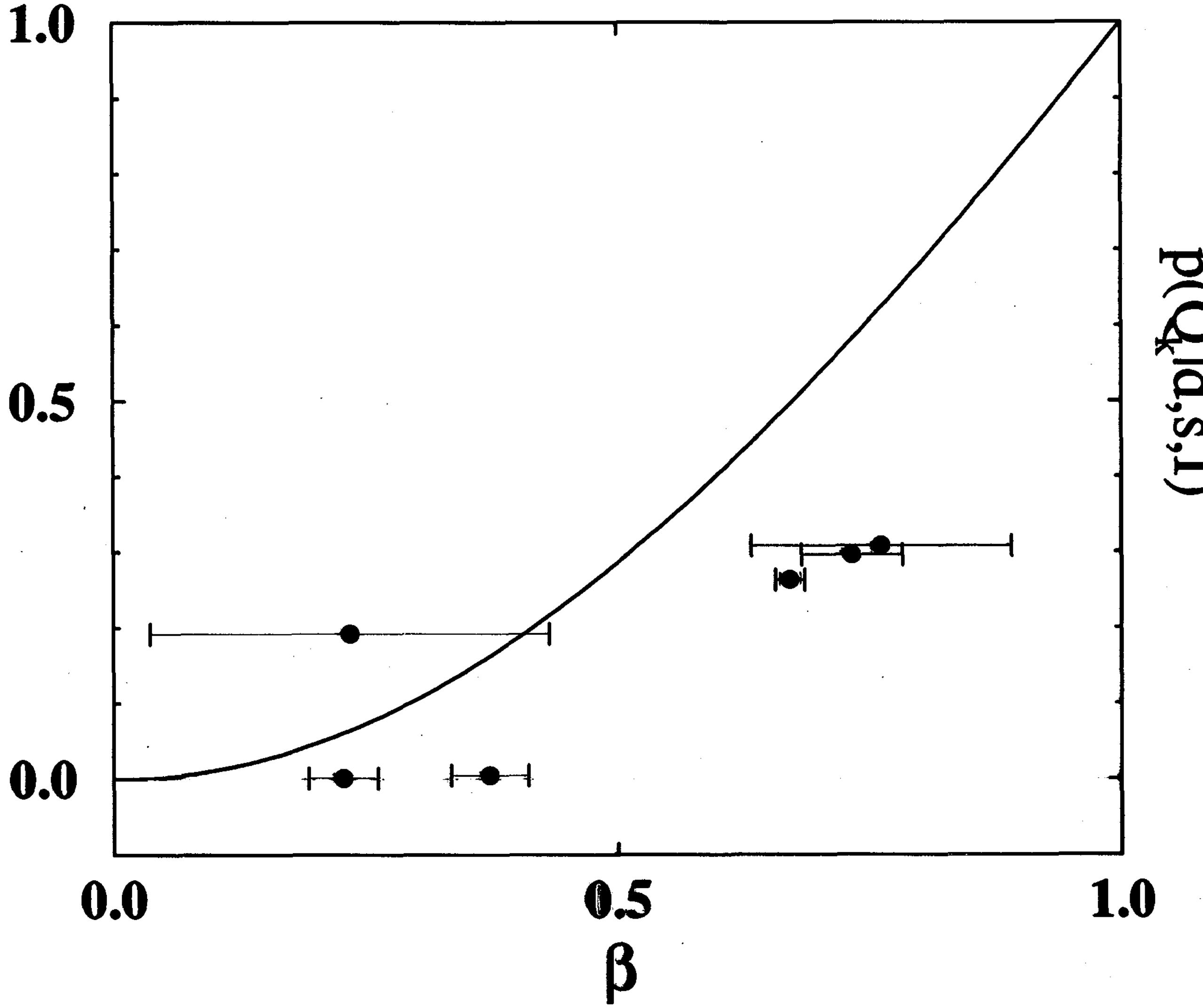
$$p(\vec{x}|Q_k,\vec{s},a,\gamma,\beta,I) = \int p(\mu|I)\cdot p(\vec{x}|Q_k,\mu,\vec{s},a,\gamma,\beta,I)\ d\mu$$

$$p(\vec{x}|Q_k,\mu,\vec{s},a,\gamma,\beta,I) \sim A_k \prod_{i\neq k} \{\beta A_i(\mu) + (1-\beta)B_i(\mu)\}$$

Quantum Hall Effect



Proton gyromagnetic ratio



NUCLEAR FUSION, Vol.25, No.1 (1985)

GLOBAL ENERGY CONFINEMENT SCALING FOR NEUTRAL-BEAM-HEATED TOKAMAKS

S.M. KAYE, R.J. GOLDSTON
Plasma Physics Laboratory,
Princeton University,
Princeton, New Jersey,
United States of America

ABSTRACT. A total of 677 representative discharges from seven neutral-beam-heated tokamaks have been used to study the parametric scaling of global energy confinement time. Contributions to this data base were from Asdex, DITE, D-III, ISX-B, PDX, PLT and TFR, and were taken from results of gettered, L-mode type discharges. Assuming a power law dependence of τ_E on the discharge parameters κ, I_p, B_t, $\bar{n}_e$ P_{tot}, a and R, standard multiple linear regression techniques were used in two steps to determine the scaling. The results indicate that the discharges used in the study are well described by the scaling $\tau_E \propto \kappa^{0.28} B_T^{-0.09} I_p^{1.24} \bar{n}_e^{0.26} P_{tot}^{-0.58} a^{-0.49} R^{1.65}$.