Scaling:
The Constant Method

from Fechner through Thurstone to

Bock & Jones, 1968

Compare each of several objects to a "constant," and judge

$$X_j > X_c$$

or not. A model for this (out of Fechner (1860), through Thurstone (1927a and 1927b) has "discriminal processes"

$$\nu_j = \mu_j + \epsilon_j \qquad \nu_c = \mu_c + \epsilon_c$$

distributed normally:

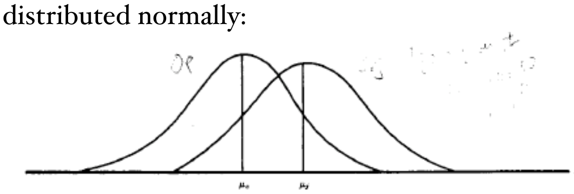


FIG. 2.1. Schematic marginal distributions of v_j and v_c.

Then the difference

$$\nu_{jc} = \nu_j - \nu_c = (\mu_j - \mu_c) + (\epsilon_j - \epsilon_c) = \mu_{jc} + \epsilon_{jc}$$

is normally distributed with

$$\mu_{jc} = \mu_j - \mu_c$$

$$\sigma_{jc}^2 = \sigma_j^2 + \sigma_c^2 - 2\rho_{jc}\sigma_j\sigma_c$$

and

$$P_{jc} = P(X_j > X_c) = \frac{1}{\sqrt{2\pi}\sigma_{jc}} \int_0^\infty \exp\left[-\frac{1}{2}\left(\frac{y - \mu_{jc}}{\sigma_{jc}}\right)^2\right] dy$$

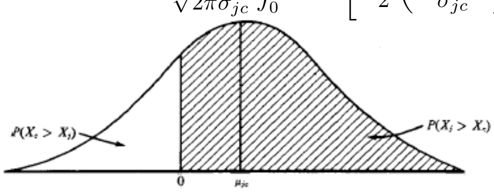


FIG. 2.2. Schematic distribution of v_{jc}.

As Bock & Jones say (p. 19), "it is convenient" to set

$$\sigma_{jc}^2 = 1$$

re-arrange the normal integral to use the standard normal, so

$$P_{jc} = \frac{1}{\sqrt{2\pi}} \int_{-\mu_{jc}}^\infty \exp\left(-\frac{1}{2}z^2\right) dz = \Phi(\mu_{jc})$$

For psychophysical problems with known values of *x*, we then make a linear model

$$\mu_{jc} = \alpha + \beta x_{jc}$$

and estimate the parameters by maximum likelihood, which is our goal here.

What we're about:

- 2-parameter maximum likelihood, using
- (minimally) multivariate Newton-Raphson

What this was about historically:

- Fechner offered this model "beginning scientific psychology"
- Thurstone generalized this model to cases with no physical x, especially using paired comparisons and "successive categories" (aka rating scales)
- IRT combines this model with latent variables

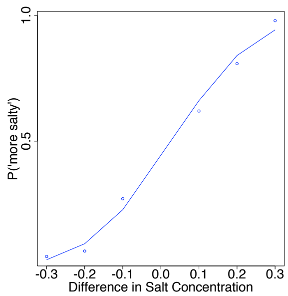
The data for the example come from Table 2.1 of Bock & Jones:
The numbers of judges who rate one (of six) solutions "saltier" than a standard:

TABLE 2.1. Graphical Solution for the Constant Method*

Concentration, %, x_j	N_j	r_{jc}	p_j	P_{jL}*	P_{jU}*	$y_j = \hat{Y}_{jc}$	Y_{jL}	Y_{jU}
1.3	49	48	.980	.906	.999	2.05	1.32	3.09
1.2	47	38	.809	.688	.897	.87	.49	1.26
1.1	50	31	.620	.493	.737	.31	.02	.63
.9	48	13	.271	.168	.398	− .61	− .96	− .26
.8	48	3	.062	.017	.155	−1.54	−2.12	−1.02
.7	49	2	.041	.007	.124	−1.74	−2.46	−1.16

* Confidence bounds on the P_j are based upon individual confidence coefficients, $1 - \gamma = .90$; P_{jL} and P_{jU} are determined from (2.18) and (2.19), respectively.

One can solve the problem graphically, and Bock & Jones do so:



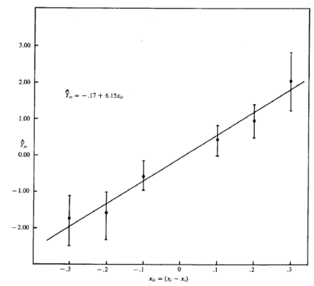
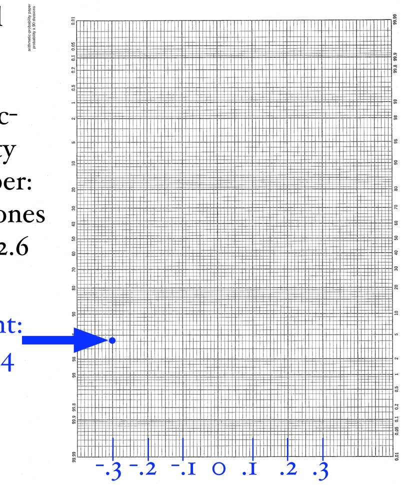
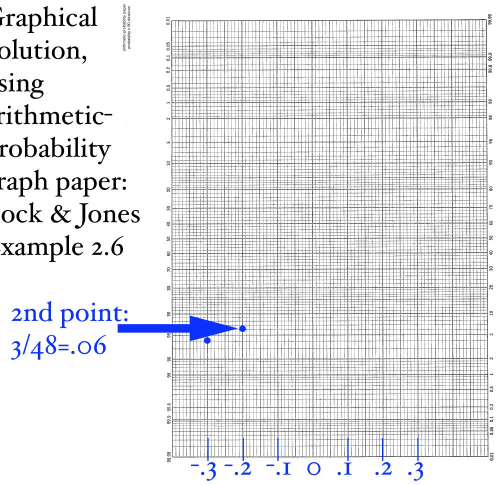


FIG. 2.5. Plot of graphical solution for the constant method.

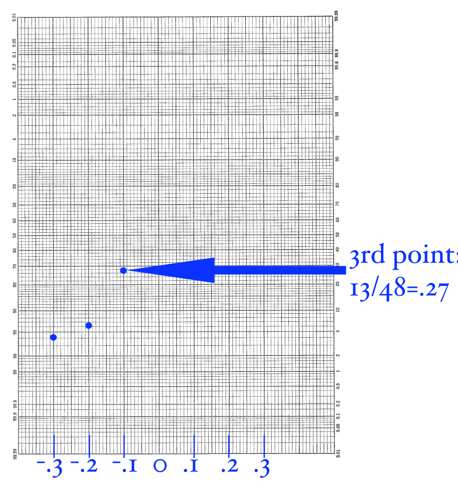
Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6



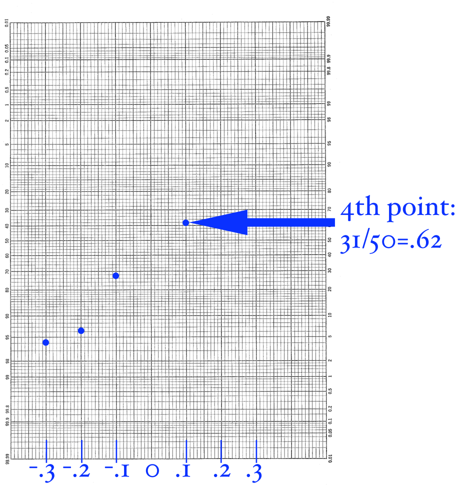
Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6



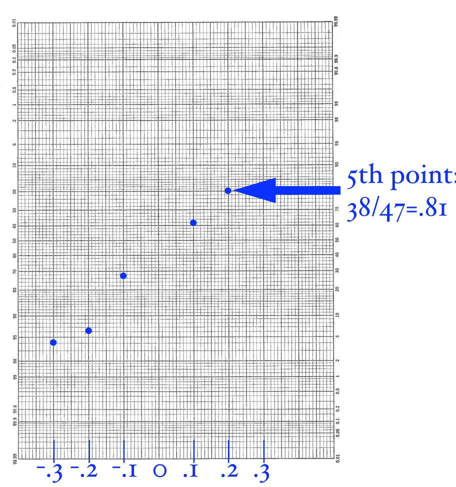
Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6



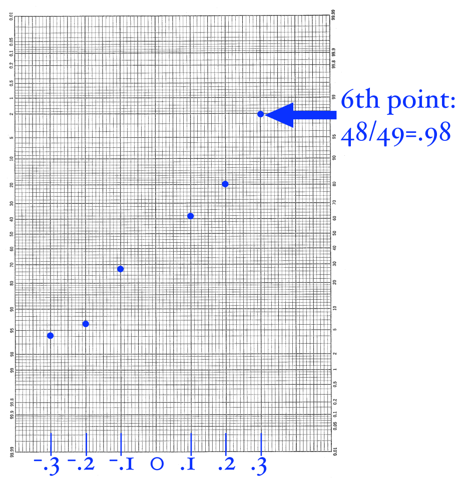
Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6



Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6

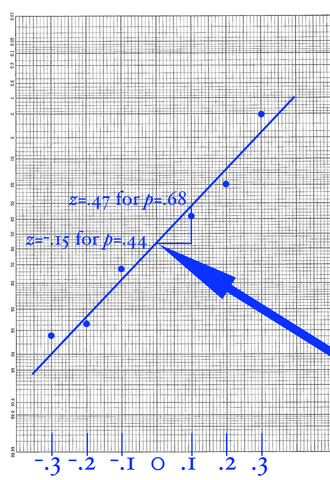


Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6



Graphical Solution, using arithmetic-probability graph paper: Bock & Jones Example 2.6

[Bock & Jones reported an intercept of -.17, and slope 6.15.]



Draw a line "by eye" & read slope & intercept from the graph.

Here, intercept is z=-.15 for p=.44, and slope is 10*[.47-(-.15)] = 6.2.

There are many other ways to solve the estimation problem.

Bock & Jones go on about Urban's "minimum normit chi-square" solution (which we will not).

Among "advanced solutions for the constant method" we have maximum likelihood.

The likelihood for the observations is

$$L = \prod_{j=1}^{n} \frac{N_{jc}!}{r_{jc}!(N_{jc} - r_{jc})!} P_{jc}^{r_{jc}} Q_{jc}^{(N_{jc}-r_{jc})}$$

so the loglikelihood is

$$\ell = \sum_{j=1}^{n} \log \frac{N_{jc}!}{r_{jc}!(N_{jc} - r_{jc})!} + r_{jc} \log P_{jc} + (N_{jc} - r_{jc}) \log Q_{jc}$$

where

$$P_{jc} = \Phi(\alpha + \beta x_{jc})$$

Bock & Jones (pp. 53-56) do the derivatives of the loglikelihood w.r.t. alpha and beta in great detail.

Newton-Raphson here involves (Bock & Jones, p. 56):

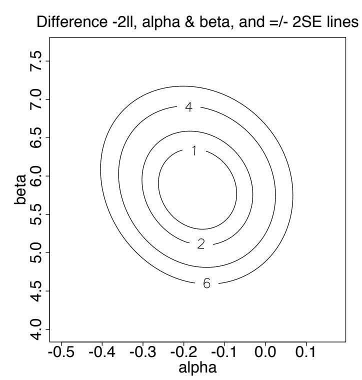
$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{k+1} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{k} - \begin{bmatrix} \frac{\partial^2 \ell}{\partial \alpha^2} & \frac{\partial^2 \ell}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ell}{\partial \beta \partial \alpha} & \frac{\partial^2 \ell}{\partial \beta^2} \end{bmatrix}_{k}^{-1} \begin{bmatrix} \frac{\partial \ell}{\partial \alpha} \\ \frac{\partial \ell}{\partial \beta} \end{bmatrix}_{k}$$

Bock's IRT Chapter 2 (p. 70) describes the Fisher-scoring version of multivariate Newton-Raphson as:

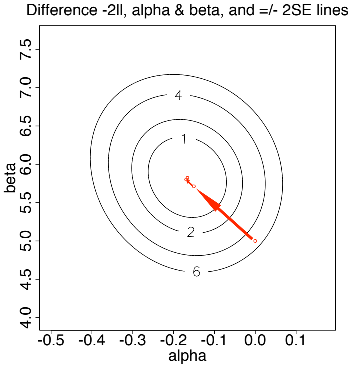
$$\hat{\theta}_{i+1} = \hat{\theta}_i + I^{-1}(\hat{\theta}_i) G(\hat{\theta}_i)$$

(The sign difference is due to the fact that the information matrix is the negative (expected value of the) matrix of second derivatives.

The idea is to locate the maximum of the loglikelihood surface



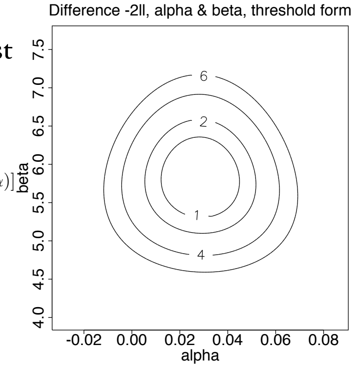
For example, starting at (0, 5) Newton iterations may look like this:



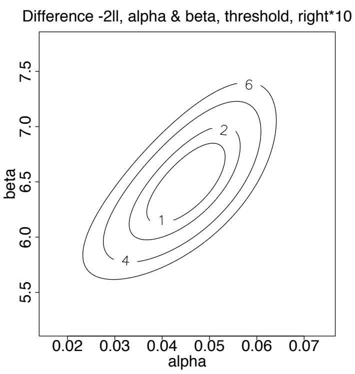
Likelihoods are not always so Gaussian-looking:

This is the same model and data cast in slope-threshold form:

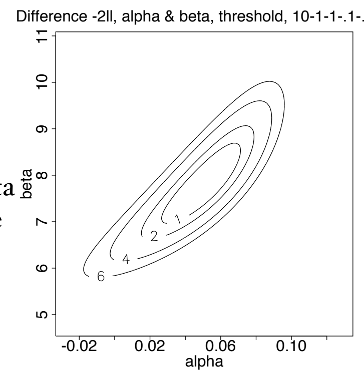
$P_{jc} = \Phi[\beta(x_{jc} - \alpha)]$



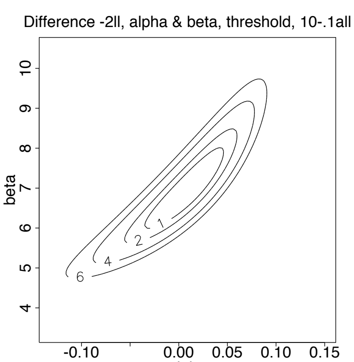
Slope-threshold form, and ten times the data on the right:



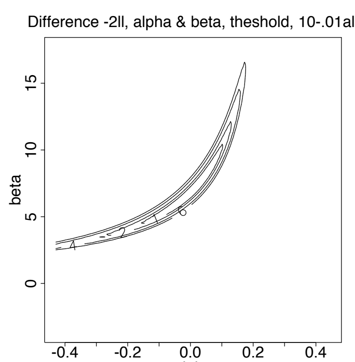
Slope-threshold form, and ten times the data at .3, but only .1 of the data for negative *x*.



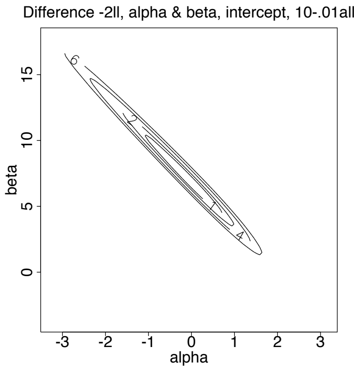
Slope-threshold form, and ten times the data at .3, but only .1 of the data for all other *x*.



Slope-threshold form, and ten times the data at .3, but only .01 of the data for all other *x*.



Slope-intercept form removes the curve, even with ten times the data at .3, but only .01 of the data for all other x.



On the programming side:

- R's glm function
- Using R's nlm (nonlinear minimizer; nlminb in Splus), both
    - without derivatives and
    - with derivatives
- C++, using Davies' NEWMAT maximizer

Next: Commentary on likelihood and MCMC