Introduction to Bayesian Statistics - 4

PhD Physics course (XXVIII ciclo)
Università di Trieste
Edoardo Milotti

A few simple applications of Bayesian techniques

1. straight-line fit
2. weighted mean
3. systematic errors
4. a two-dimensional location problem
5. search for weak signals in spectra

1. Straight-line fit

$$y_i = ax_i + b + \varepsilon_i$$

y_i measured value

x_i independent variable ("exactly" known)

a, b fit parametes: eventually we expect to find pdf's for these parameters

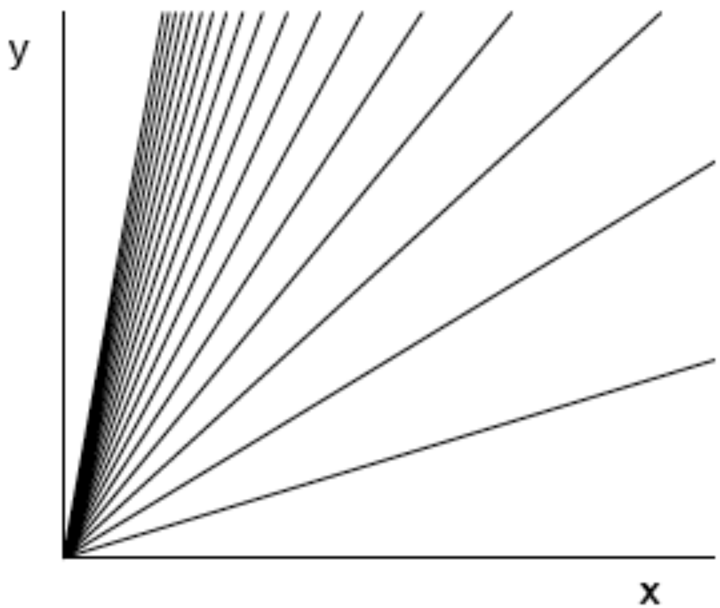
ε_i statistical error

$$\langle \varepsilon_i \rangle = 0; \quad \langle \varepsilon_i^2 \rangle = \sigma^2 \quad \Rightarrow$$ the statistical measurement error has a Gaussian distribution

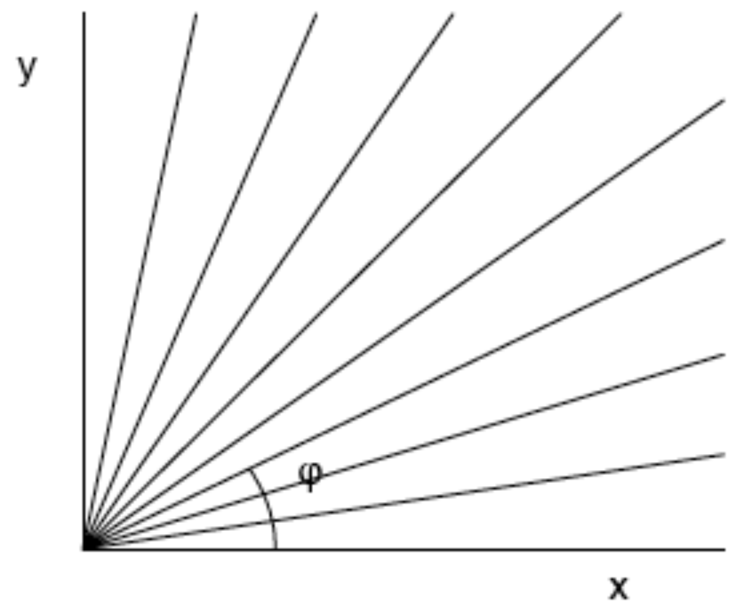
likelihood

$$p(\mathbf{y} \mid a, b, \mathbf{x}, \sigma) = \left(2\pi\sigma^2\right)^{-N/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i - a x_i - b\right)^2\right]$$

prior angular distribution



uniform *a*



uniform angle

The uniform distribution of a introduces an angular bias. The least informative choice corresponds to a uniform angular distribution

$$p_\varphi(\varphi) = \frac{1}{\pi}; \quad -\frac{\pi}{2} \le \varphi < \frac{\pi}{2}$$

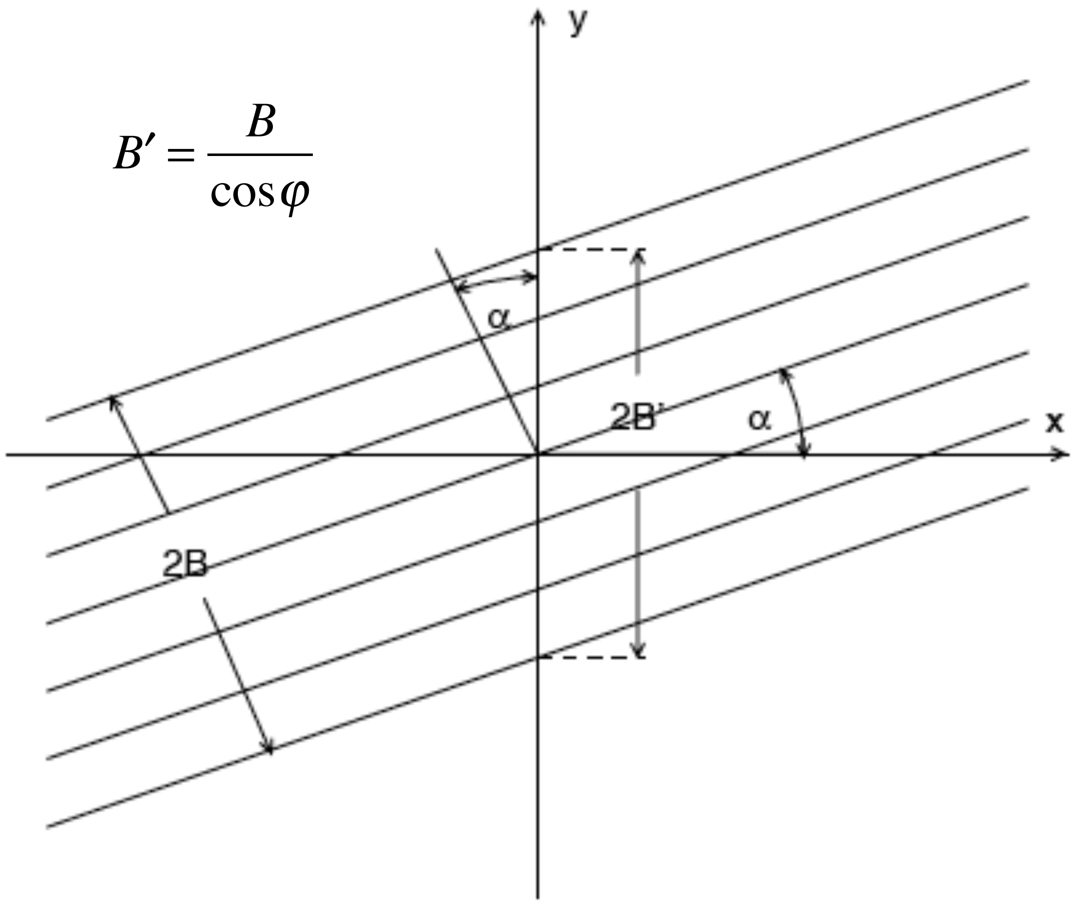
and we obtain the distribution of *a* with the transformation method:

$$a = \tan\varphi$$

$$\Rightarrow \quad p_\varphi(\varphi)d\varphi = p_a(a)da = p_a(a)d(\tan\varphi) = p_a(a)\sec^2\varphi d\varphi$$

$$\Rightarrow \quad p_a(a) = \frac{1}{\pi\sec^2\varphi} = \frac{1}{\pi(1+\tan^2\varphi)} = \frac{1}{\pi(1+a^2)}$$

prior distribution of *b*: improper uniform distribution, related to the distribution of *a*



$$B' = \frac{B}{\cos\varphi}$$

$$p(b \mid a=0) = \frac{1}{2B}; \quad p(b \mid a) = \frac{1}{2B'} = \frac{\cos\varphi}{2B} = \frac{1}{2B}\cdot\frac{1}{\sqrt{1+a^2}}$$

we obtain the posterior from Bayes' theorem

$$p(a,b \mid \mathbf{y},\mathbf{x},\sigma) = \frac{p(\mathbf{y} \mid a,b,\mathbf{x},\sigma)}{\int_{-\infty}^{+\infty} da \int_{-B/\cos\varphi}^{B/\cos\varphi} db\; p(\mathbf{y} \mid a,b,\mathbf{x},\sigma)\cdot p(a,b)} \cdot p(a,b)$$

where the prior is

$$p(a,b) = p(b \mid a)\cdot p(a) = \left(\frac{1}{2B}\cdot\frac{1}{\sqrt{1+a^2}}\right)\left(\frac{1}{\pi\left(1+a^2\right)}\right)$$

$$\propto \frac{1}{\left(1+a^2\right)^{3/2}}$$

finally we find

$$p\left(a,b \mid \mathbf{y},\mathbf{x},\sigma\right)=\frac{\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i-ax_i-b\right)^2\right]}{\left\{\int_{-\infty}^{+\infty}da\int_{-B/\cos\varphi}^{B/\cos\varphi}db\,\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i-ax_i-b\right)^2\right]\cdot\frac{1}{\left(1+a^2\right)^{3/2}}\right\}}\cdot\frac{1}{\left(1+a^2\right)^{3/2}}$$

$$\approx\frac{\frac{1}{\left(1+a^2\right)^{3/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i-ax_i-b\right)^2\right]}{\left\{\int_{-\infty}^{+\infty}\frac{da}{\left(1+a^2\right)^{3/2}}\int_{-\infty}^{+\infty}db\,\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i-ax_i-b\right)^2\right]\right\}}$$

This expression has a partly Gaussian structure, and now we rearrange the quadratic expression in the exponential

$$\sum_{i=1}^{N}(y_i - ax_i - b)^2 = \sum_{i=1}^{N}\left[(y_i - ax_i)^2 - 2b(y_i - ax_i) + b^2\right]$$

$$= \sum_{i=1}^{N}(y_i - ax_i)^2 - 2b\sum_{i=1}^{N}(y_i - ax_i) + Nb^2$$

$$= N\left\{\left[b^2 - 2b\frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i) + \left(\frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2\right] + \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)^2 - \left(\frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2\right\}$$

$$= N\left\{\left(b - \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2 + \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)^2 - \left(\frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2\right\}$$

$$= N\left(b - \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2 + N\left(\frac{1}{N}\sum_{i=1}^{N}y_i^2 - 2a\frac{1}{N}\sum_{i=1}^{N}x_i y_i + a^2\frac{1}{N}\sum_{i=1}^{N}x_i^2\right) - N\left(\frac{1}{N}\sum_{i=1}^{N}y_i - a\frac{1}{N}\sum_{i=1}^{N}x_i\right)^2$$

$$= N\left(b - \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2 + N\left(\operatorname{var} y - 2a\operatorname{cov}(x,y) + a^2\operatorname{var} x\right)$$

therefore the normalization integral becomes

$$\int_{-\infty}^{+\infty}\frac{da}{(1+a^2)^{3/2}}\exp\left[-\frac{N}{2\sigma^2}\left(\operatorname{var} y - 2a\operatorname{cov}(x,y) + a^2\operatorname{var} x\right)\right]\int_{-\infty}^{+\infty}db\,\exp\left[-\frac{N}{2\sigma^2}\left(b - \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)\right)^2\right]$$

$$= \sqrt{\frac{2\pi\sigma^2}{N}}\int_{-\infty}^{+\infty}\frac{da}{(1+a^2)^{3/2}}\exp\left[-\frac{N}{2\sigma^2}\left(\operatorname{var} y - 2a\operatorname{cov}(x,y) + a^2\operatorname{var} x\right)\right]$$

Approximate integration of the remaining integral

$$\int_{-\infty}^{+\infty} \frac{da}{\left(1+a^2\right)^{3/2}} \exp\left[-\frac{N}{2\sigma^2}\left(\operatorname{var} y - 2a\operatorname{cov}(x,y) + a^2 \operatorname{var} x\right)\right]$$

We evaluate this integral by integrating about the peak of the integrand, assuming that the peak is narrow.

We start with the logarithm of the integrand, we find its maximum and we Taylor expand about the maximum

$$\Phi(a) = -\frac{3}{2}\ln\left(1+a^2\right) - \frac{N}{2\sigma^2}\left(\operatorname{var} y - 2a\operatorname{cov}(x,y) + a^2 \operatorname{var} x\right)$$

$$\Phi(a) = -\frac{3}{2}\ln\left(1+a^2\right) - \frac{N}{2\sigma^2}\left(\mathrm{var}\, y - 2a\,\mathrm{cov}(x,y) + a^2\,\mathrm{var}\, x\right)$$

$$\frac{d\Phi}{da} = -\frac{3a}{1+a^2} + \frac{N}{\sigma^2}\left(\mathrm{cov}(x,y) - a\,\mathrm{var}\, x\right) = 0$$

we find *a* from this cubic equation

note that when N>>1 the peak is at position $a_0 \approx \frac{\mathrm{cov}(x,y)}{\mathrm{var}\, x}$

We use the Newton-Raphson method for the solution of the cubic equation:

$$f(a_0) = -\frac{3a_0}{1+a_0^2}$$

$$f'(a_0) = -3\frac{1-a_0^2}{(1+a_0^2)^2} - \frac{N}{\sigma^2}\mathrm{var}\, x \approx -\frac{N}{\sigma^2}\mathrm{var}\, x$$

then

$$\delta a_1 = -\frac{3a_0}{1+a_0^2}\frac{\sigma^2}{N\,\mathrm{var}\,x} \qquad\qquad a_1 = a_0 - \frac{3a_0}{1+a_0^2}\frac{\sigma^2}{N\,\mathrm{var}\,x} \qquad (1)$$

Now, to complete the expansion, we must evaluate the second derivative at a_1:

$$\left.\frac{d^2\Phi}{da^2}\right|_{a=a_1} = -3\frac{1-a_1^2}{(1+a_1^2)^2} - \frac{N}{\sigma^2}\mathrm{var}\,x = -\frac{1}{2\sigma_1^2} \qquad (2)$$

$$\Phi(a) \approx \Phi(a_1) + \frac{1}{2}\left.\frac{d^2\Phi}{da^2}\right|_{a_1}(a-a_1)^2 = \Phi(a_1) - \frac{(a-a_1)^2}{2\sigma_1^2}$$

we find this by using equations (1) and (2)

Now we complete the evaluation of the integral

$$\int_{-\infty}^{+\infty} \frac{da}{\left(1+a^2\right)^{3/2}} \exp\left[-\frac{N}{2\sigma^2}\left(\operatorname{var} y - 2a\operatorname{cov}(x,y) + a^2 \operatorname{var} x\right)\right]$$

$$= \int_{-\infty}^{+\infty} \exp\left[\Phi(a)\right] da$$

$$\approx \int_{-\infty}^{+\infty} \exp\left[\Phi(a_1) - \frac{\left(a-a_1\right)^2}{2\sigma_1^2}\right] da = \sqrt{2\pi\sigma_1^2} \exp\left[\Phi(a_1)\right]$$

and finally we find the posterior distribution.

Moreover

$$p(a,b\,|\,\mathbf{y},\mathbf{x},\sigma) \propto \frac{1}{\left(1+a^2\right)^{3/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(y_i - a x_i - b\right)^2\right]$$

$$\approx \exp\left[-\Phi(a_1) - \frac{(a-a_1)^2}{2\sigma_1^2}\right] \exp\left[-\frac{N}{2\sigma^2}\left(b - \frac{1}{N}\sum_{i=1}^{N}\left(y_i - a_1 x_i\right)\right)^2\right]$$

and thus we see that:

$$\langle a \rangle = a_1; \quad \operatorname{var} a = \sigma_1^2;$$

$$\langle b \rangle = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - a_1 x_i\right); \quad \operatorname{var} b = \frac{\sigma^2}{N}$$

2. Weighted mean

We consider known Gaussian errors

The likelihood function is

$$P(\mathbf{d} \mid \mu, \boldsymbol{\sigma}) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left[-\frac{(d_k-\mu)^2}{2\sigma_k^2}\right]$$

$$= \frac{1}{(2\pi)^{N/2}} \left(\prod_{k=1}^{N} \frac{1}{\sigma_k}\right) \exp\left[-\frac{1}{2}\sum_{k=1}^{N} \frac{(d_k-\mu)^2}{\sigma_k^2}\right]$$

Using an improper uniform prior we find

$$P(\mu \mid \mathbf{d},\boldsymbol{\sigma}) = \frac{P(\mathbf{d} \mid \boldsymbol{\sigma},\mu)}{\int_\mu P(\mathbf{d} \mid \boldsymbol{\sigma},\mu)\cdot P(\mu)\,d\mu}\cdot P(\mu) \rightarrow \frac{P(\mathbf{d} \mid \boldsymbol{\sigma},\mu)}{\int_\mu P(\mathbf{d} \mid \boldsymbol{\sigma},\mu)\,d\mu}$$

and then

$$P(\mu \mid \mathbf{d},\boldsymbol{\sigma}) = \frac{\exp\left[-\frac{1}{2}\sum_{k=1}^{N}\frac{(d_k-\mu)^2}{\sigma_k^2}\right]}{\int_{-\infty}^{+\infty}\exp\left[-\frac{1}{2}\sum_{k=1}^{N}\frac{(d_k-\mu)^2}{\sigma_k^2}\right]d\mu}$$

the exponent can be rearranged as usual

$$\sum_{k=1}^{N}\frac{(d_k-\mu)^2}{\sigma_k^2}=\sum_{k=1}^{N}\frac{d_k^2-2\mu d_k+\mu^2}{\sigma_k^2}=\mu^2\sum_{k=1}^{N}\frac{1}{\sigma_k^2}-2\mu\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}+\sum_{k=1}^{N}\frac{d_k^2}{\sigma_k^2}$$

$$=\left(\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\right)\left[\mu^2-2\mu\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\right)+\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\right)^2\right]+\left[\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\right)^2\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2}-\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\right)^2\right]$$

$$=\left(\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\right)\left[\mu-\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\right)\right]^2+\left[\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\right)^2\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2}-\left(\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\right)^2\right]$$

thus we see that the mean has a Gaussian posterior distribution and that

$$\bar{\mu}=\frac{\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}}{\sum_{k=1}^{N}\frac{1}{\sigma_k^2}};\qquad \sigma_\mu^2=\left(\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\right)^{-1}$$

3. Miscalibrated Gaussian measurement errors

Now we consider a case where we must find the mean value with given measurement errors, and where the errors are Gaussian and they are systematically multiplied by an unknown scale factor.

the likelihood is

$$P(\mathbf{d} \mid \mu, \boldsymbol{\sigma}, \alpha) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\alpha^2\sigma_k^2}} \exp\left[-\frac{(d_k - \mu)^2}{2\alpha^2\sigma_k^2} \right]$$

$$= \frac{1}{(2\pi)^{N/2}\alpha^N} \left(\prod_{k=1}^{N} \frac{1}{\sigma_k} \right) \exp\left[-\frac{1}{2\alpha^2} \sum_{k=1}^{N} \frac{(d_k - \mu)^2}{\sigma_k^2} \right]$$

we must rearrange the exponent as usual ...

$$\sum_{k=1}^{N}\frac{(d_k-\mu)^2}{\sigma_k^2}=\sum_{k=1}^{N}\frac{d_k^2}{\sigma_k^2}-2\mu\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}+\mu^2\sum_{k=1}^{N}\frac{1}{\sigma_k^2}=\frac{ND}{\sigma_M^2}-2\mu\frac{NM}{\sigma_M^2}+\mu^2\frac{N}{\sigma_M^2}$$

$$=\frac{N}{\sigma_M^2}\left(D-2\mu M+\mu^2\right)$$

dove $\frac{1}{\sigma_M^2}=\frac{1}{N}\sum_{k=1}^{N}\frac{1}{\sigma_k^2};\quad M=\sum_{k=1}^{N}\frac{d_k}{\sigma_k^2}\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2};\quad D=\sum_{k=1}^{N}\frac{d_k^2}{\sigma_k^2}\Big/\sum_{k=1}^{N}\frac{1}{\sigma_k^2}$

therefore the likelihood is

$$P(\mathbf{d}\mid\mu,\boldsymbol{\sigma},\alpha)=\frac{1}{(2\pi)^{N/2}\alpha^N}\left(\prod_{k=1}^{N}\frac{1}{\sigma_k}\right)\exp\left[-\frac{N}{2\alpha^2\sigma_M^2}\left(D-2\mu M+\mu^2\right)\right]$$

Now we estimate the scale factor from Bayes' theorem

$$P(\alpha \mid \mathbf{d},\boldsymbol{\sigma}) = \frac{P(\mathbf{d} \mid \boldsymbol{\sigma},\alpha)}{\int_\alpha P(\mathbf{d} \mid \boldsymbol{\sigma},\alpha)\cdot P(\alpha)\,d\alpha}\cdot P(\alpha)$$

however we need first to marginalize the likelihood with respect to the mean, which in this case is a nuisance parameter

we take a uniform prior for the mean

$$P(\mathbf{d} \mid \boldsymbol{\sigma},\alpha) = \int_\mu P(\mathbf{d} \mid \mu,\boldsymbol{\sigma},\alpha) P(\mu \mid \boldsymbol{\sigma},\alpha)\,d\mu$$

$$= \frac{1}{W}\int_{\mu_{\min}}^{\mu_{\max}} P(\mathbf{d} \mid \mu,\boldsymbol{\sigma},\alpha)\,d\mu$$

$$\approx \frac{1}{W}\frac{1}{(2\pi)^{N/2}\alpha^N}\left(\prod_{k=1}^{N}\frac{1}{\sigma_k}\right)\int_{-\infty}^{+\infty}\exp\left[-\frac{N}{2\alpha^2\sigma_M^2}\left(D-2\mu M+\mu^2\right)\right]d\mu$$

$$\left(W = \mu_{\max} - \mu_{\min}\right)$$

as usual ...

$$D - 2\mu M + \mu^2 = \mu^2 - 2\mu M + M^2 + D - M^2$$
$$= (\mu - M)^2 + D - M^2$$

... therefore the likelihood is:

$$P(\mathbf{d} \mid \boldsymbol{\sigma}, \alpha) \approx \frac{1}{W} \frac{1}{(2\pi)^{N/2} \alpha^N} \left(\prod_{k=1}^{N} \frac{1}{\sigma_k} \right) \int_{-\infty}^{+\infty} \exp\left\{ -\frac{N}{2\alpha^2 \sigma_M^2} \left[(\mu - M)^2 + D - M^2 \right] \right\} d\mu$$
$$= \frac{1}{W} \frac{1}{(2\pi)^{N/2} \alpha^N} \left(\prod_{k=1}^{N} \frac{1}{\sigma_k} \right) \exp\left(-\frac{N\left(D - M^2\right)}{2\alpha^2 \sigma_M^2} \right) \sqrt{\frac{2\pi\alpha^2\sigma_M^2}{N}}$$

$$P(\alpha \mid \mathbf{d},\boldsymbol{\sigma}) = \frac{P(\mathbf{d} \mid \boldsymbol{\sigma},\alpha)}{\int_\alpha P(\mathbf{d} \mid \boldsymbol{\sigma},\alpha)\cdot P(\alpha)}\cdot P(\alpha)$$

$$= \frac{\frac{1}{\alpha^{N-1}}\exp\left(-\frac{N\left(D-M^2\right)}{2\alpha^2\boldsymbol{\sigma}_M^2}\right)}{\int_\alpha \frac{1}{\alpha^{N-1}}\exp\left(-\frac{N\left(D-M^2\right)}{2\alpha^2\boldsymbol{\sigma}_M^2}\right)\cdot P(\alpha)\,d\alpha}\cdot P(\alpha)$$

$$P(\alpha) \propto \frac{1}{\alpha}$$

the prior should be scale-independent and therefore we take Jeffrey's prior

$$P\left(\alpha \mid \mathbf{d},\boldsymbol{\sigma}\right)=\frac{\dfrac{1}{\alpha^{N-1}}\exp\left(-\dfrac{A^2}{\alpha^2}\right)\cdot\dfrac{1}{\alpha}}{\int_\alpha \dfrac{1}{\alpha^{N-1}}\exp\left(-\dfrac{A^2}{\alpha^2}\right)\cdot\dfrac{1}{\alpha}\,d\alpha}; \qquad A^2=\frac{N\left(D-M^2\right)}{2\sigma_M^2}$$

$$P\left(\alpha \mid \mathbf{d},\boldsymbol{\sigma}\right)\to\frac{\dfrac{1}{\alpha^{N-1}}\exp\left(-\dfrac{A^2}{\alpha^2}\right)\cdot\dfrac{1}{\alpha}}{\int_0^\infty \dfrac{1}{\alpha^{N-1}}\exp\left(-\dfrac{A^2}{\alpha^2}\right)\cdot\dfrac{1}{\alpha}\,d\alpha}$$

$$\int_0^\infty \frac{1}{\alpha^{N-1}} \exp\left(-\frac{A^2}{\alpha^2}\right) \cdot \frac{1}{\alpha} d\alpha = \int_0^\infty \frac{1}{\alpha^N} \exp\left(-\frac{A^2}{\alpha^2}\right) d\alpha$$

$$\frac{A^2}{\alpha^2} = x; \quad \alpha = \frac{A}{\sqrt{x}}; \quad d\alpha = -\frac{A}{2x^{3/2}} dx$$

$$\int_0^\infty \frac{x^{N/2}}{A^N} \exp(-x) \frac{A}{2x^{3/2}} dx = \frac{1}{2A^{N-1}} \int_0^\infty x^{\frac{N-1}{2}-1} \exp(-x) dx = \frac{1}{2A^{N-1}} \cdot \Gamma\left(\frac{N-1}{2}\right)$$

$$P(\alpha \mid \mathbf{d}, \boldsymbol{\sigma}) \to \frac{\dfrac{2A^{N-1}}{\alpha^N} \exp\left(-\dfrac{A^2}{\alpha^2}\right)}{\Gamma\left(\dfrac{N-1}{2}\right)}$$

we take the MAP estimate the scale parameter from the pdf

$$P(\alpha|\mathbf{d},\boldsymbol{\sigma}) \to \frac{\dfrac{2A^{N-1}}{\alpha^N}\exp\left(-\dfrac{A^2}{\alpha^2}\right)}{\Gamma\left(\dfrac{N-1}{2}\right)}$$

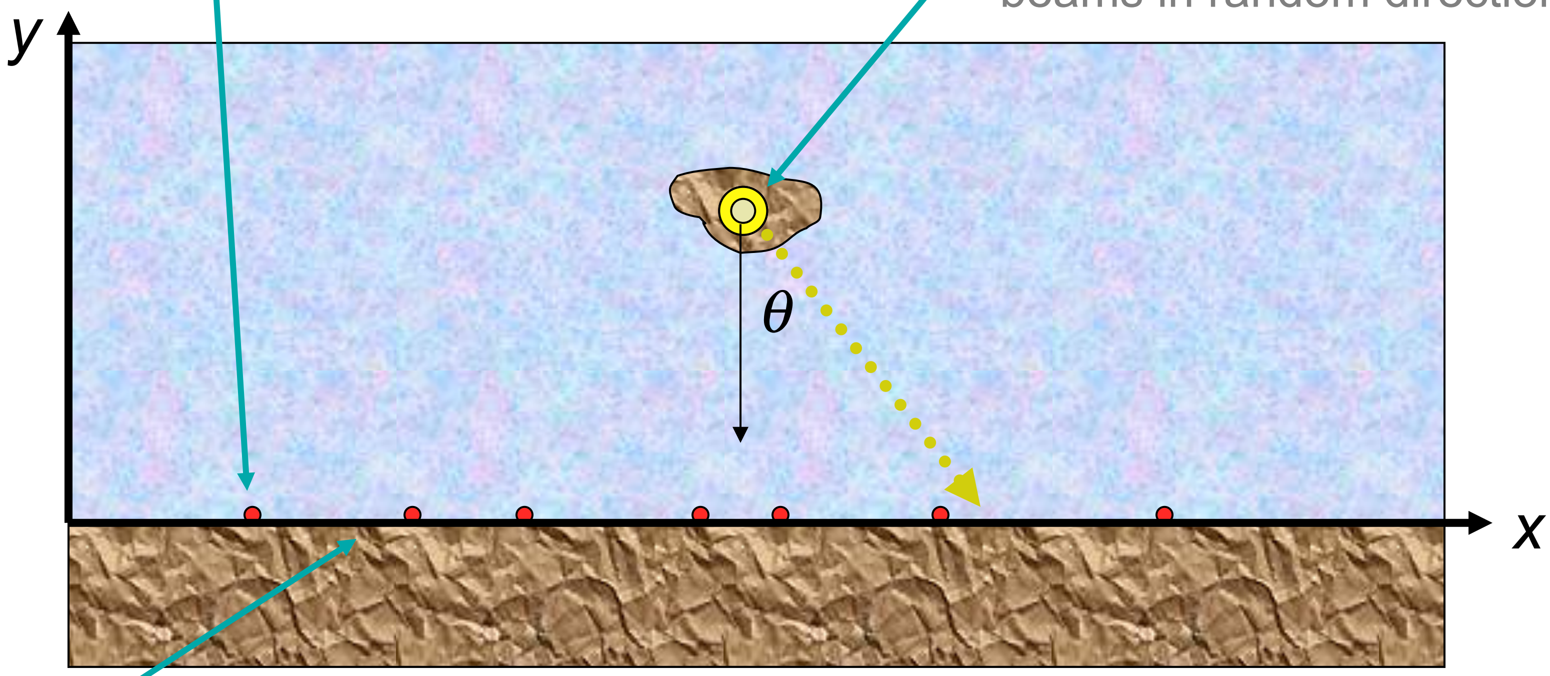
$$\frac{d}{d\alpha}P(\alpha|\mathbf{d},\boldsymbol{\sigma}) \propto -\frac{N}{\alpha^{N+1}}\exp\left(-\frac{A^2}{\alpha^2}\right)+\frac{2A^2}{\alpha^{N+3}}\exp\left(-\frac{A^2}{\alpha^2}\right)=0$$

⇒ $N\alpha^2 = 2A^2$ ⇒ $\alpha_{MAP} = \sqrt{\dfrac{2}{N}}A = \sqrt{\dfrac{\left(D-M^2\right)}{\sigma_M^2}}$

4. A two-dimensional problem: location of a lighthouse (S. Gull)

omnidirectional detectors at position (x_i, y_i) ($i=1,\dots,N$)

lighthouse at (x_0, y_0), which sends highly collimated beams in random directions



rectilinear coastline

$$x - x_0 = y_0 \tan\theta$$

$$\Delta x = y_0 \sec^2\theta \, \Delta\theta$$

individual events correspond to the detection of a certain emission angle, or rather to a certain position on the coastline: this position depends on parameters x_0 and y_0

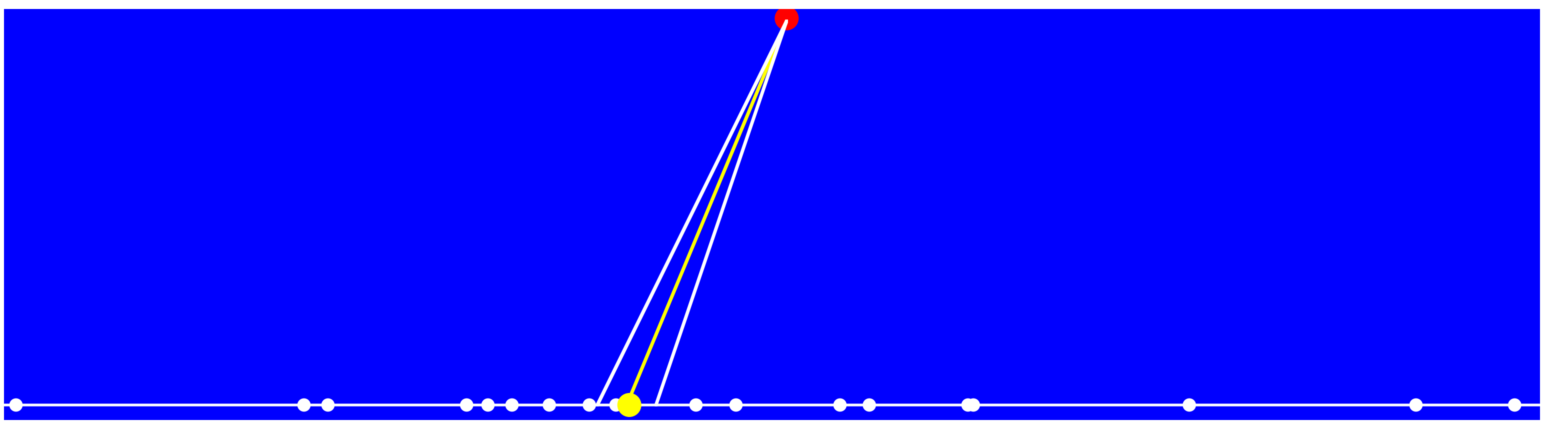
$$p_\theta(\theta \mid I) = \frac{1}{\pi}$$ uniform prior pdf for the emission angle

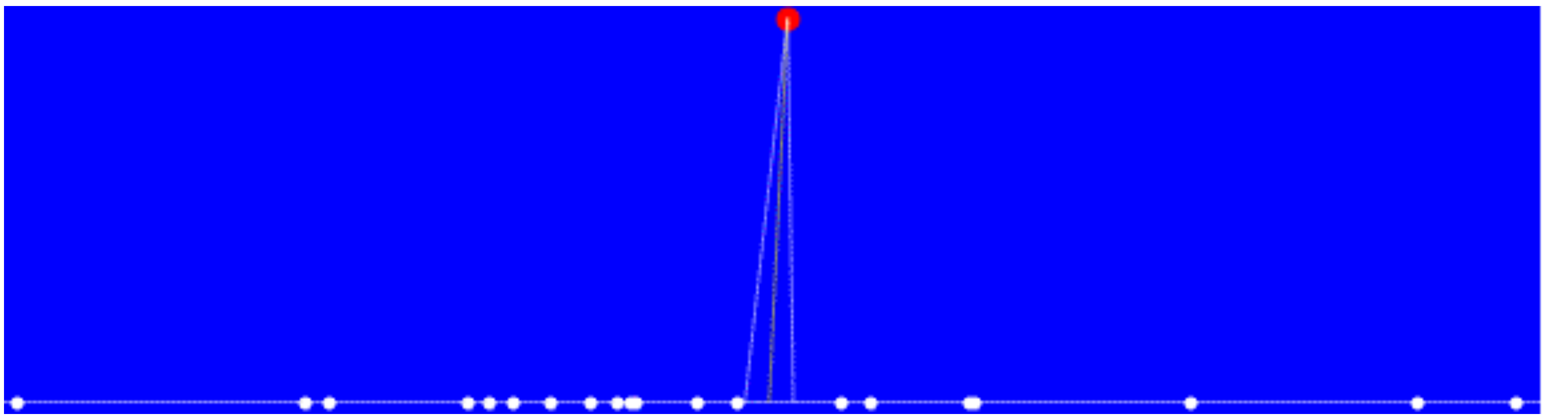
$$p_\theta(\theta \mid I)\, d\theta = p_x(x \mid x_0, y_0, I)\, dx = p_x(x \mid x_0, y_0, I)\, y_0 \sec^2\theta\, d\theta$$

$$= p_x(x \mid x_0, y_0, I) \frac{y_0^2 + (x - x_0)^2}{y_0} d\theta$$

therefore the prior pdf of x is

$$p_x(x \mid x_0, y_0, I) = \frac{1}{\pi} \cdot \frac{y_0}{y_0^2 + (x - x_0)^2}$$





likelihood

$$p\left(\mathbf{x} \mid x_0, y_0, I\right) = \prod_{k=1}^{N} p\left(x_k \mid x_0, y_0, I\right)$$

$$= \frac{y_0^N}{\pi^N} \prod_{k=1}^{N} \frac{1}{y_0^2 + \left(x_k - x_0\right)^2}$$

posterior pdf from Bayes' theorem

$$p\left(x_0, y_0 \mid \mathbf{x}, I\right) = \frac{p\left(\mathbf{x} \mid x_0, y_0, I\right)}{\int_{x,y} p\left(\mathbf{x} \mid x_0, y_0, I\right) \cdot p\left(x_0, y_0 \mid I\right) dx_0 dy_0} \cdot p\left(x_0, y_0 \mid I\right)$$

case 1: y_0 is not interesting, therefore we marginalize the likelihood with respect to y_0 (we assume a uniform prior for y)

$$\int_y p(\mathbf{x} \mid x_0, y_0, I) \cdot p(y_0 \mid x_0, I)\, dy_0 = \frac{1}{Y} \int_{y_{\min}}^{y_{\max}} \frac{y_0^N}{\pi^N} \prod_{k=1}^{N} \frac{1}{y_0^2 + (x_k - x_0)^2}\, dy_0$$

$$Y = y_{\max} - y_{\min}$$

posterior pdf from Bayes' theorem, after marginalization

$$p(x_0 \mid \mathbf{x}, I) = \frac{p(\mathbf{x} \mid x_0, I)}{\int_x p(\mathbf{x} \mid x_0, I) \cdot p(x_0 \mid I)\, dx_0} \cdot p(x_0 \mid I)$$

important subcase: a simple marginalization (y_0 = 1, we know a priori the distance of the lighthouse from the coastline)

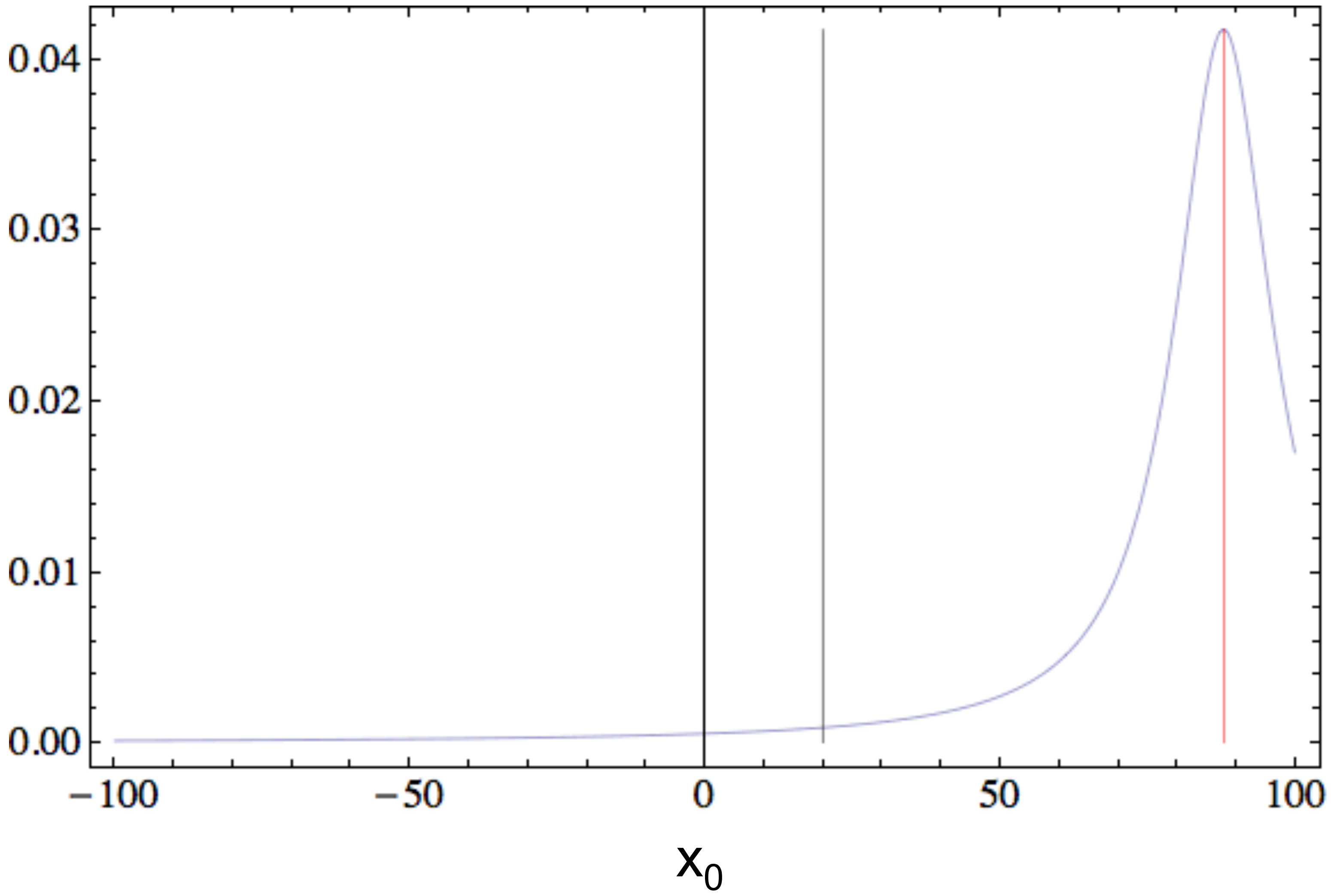
$$\int_y p\left(\mathbf{x} \mid x_0, y_0, \theta, I\right) \cdot \delta\left(y_0 - 1\right) dy_0 = \frac{1}{\pi^N} \prod_{k=1}^{N} \frac{1}{1 + \left(x_k - x_0\right)^2}$$

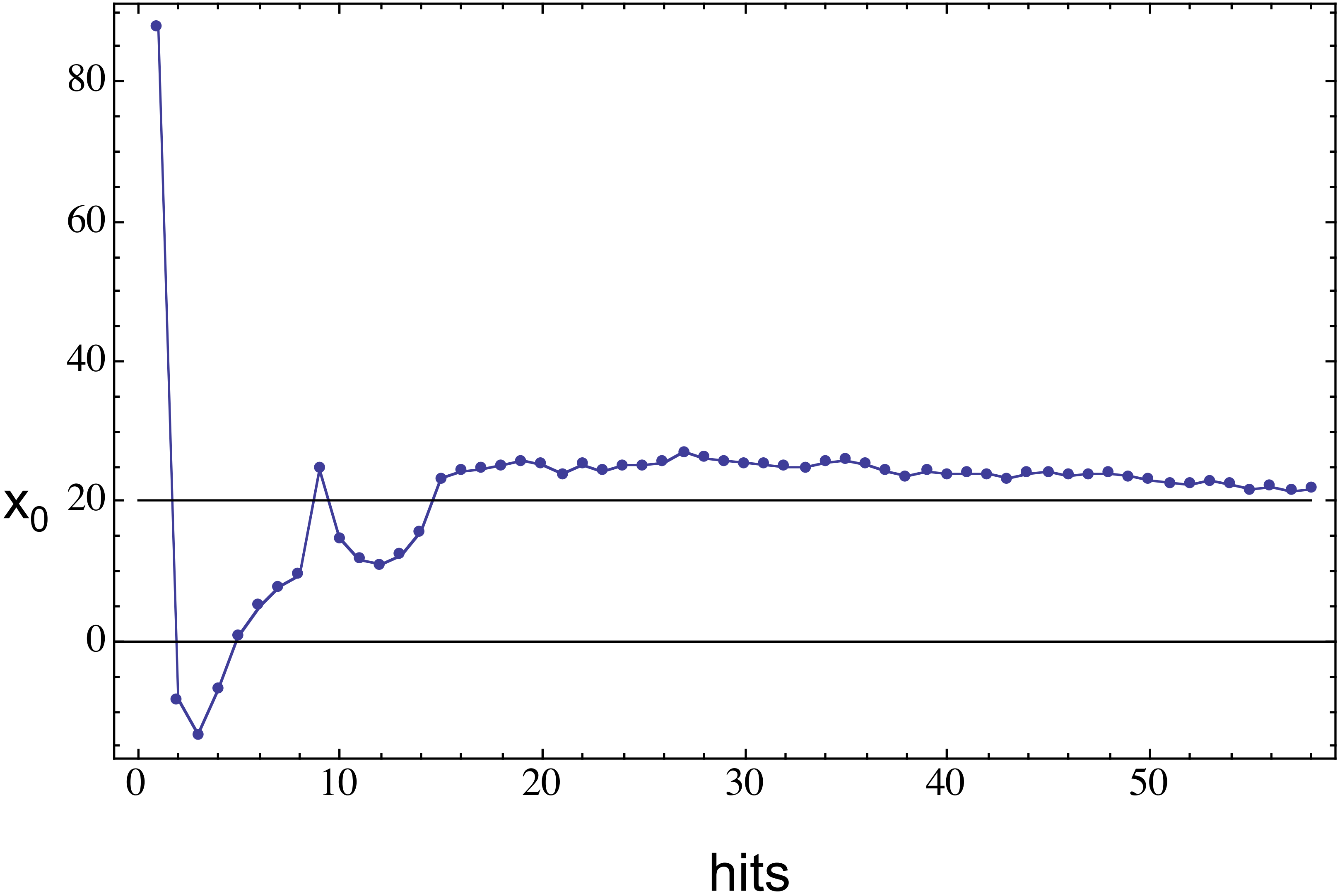
pdf localized at y_0 = 1

posterior pdf from Bayes' theorem, after marginalization (improper uniform prior for *x*)

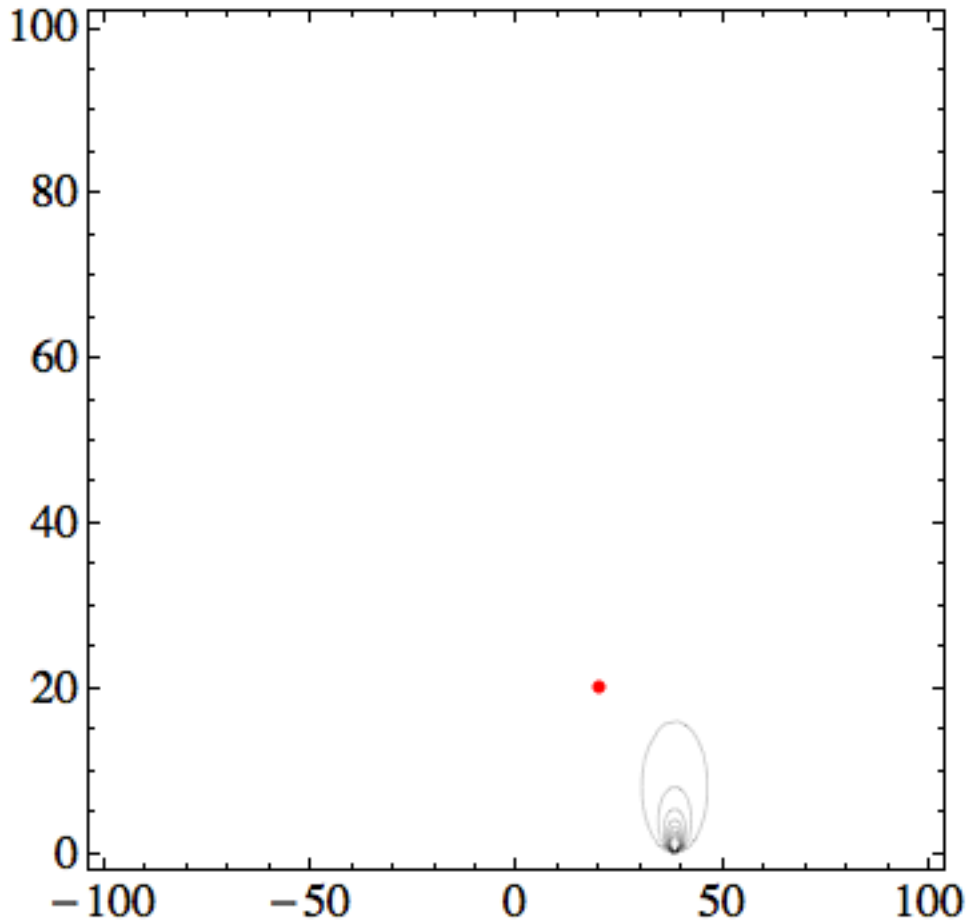
$$p\left(x_0 \mid \mathbf{x}, I\right) = \frac{p\left(\mathbf{x} \mid x_0, I\right)}{\int_x p\left(\mathbf{x} \mid x_0, I\right) \cdot p\left(x_0 \mid I\right) dx_0} \cdot p\left(x_0 \mid I\right) = \frac{\prod_{k=1}^{N} \frac{1}{1 + \left(x_k - x_0\right)^2}}{\int_{x_{\min}}^{x_{\max}} \prod_{k=1}^{N} \frac{1}{1 + \left(x_k - x_0\right)^2} dx_0}$$

numerical experiments




80
60
40
20
0
x_0
0
10
20
30
40
50
hits

2D posterior pdf (without marginalization of y_0)



5. Search of signals in spectra (Caldwell and Kröninger, PRD 74 (2006) 092003)

Consider the search for sparse signals in a spectrum where

- The spectrum is confined to a certain region of interest.
- The spectral shape of a possible signal is known.
- The spectral shape of the background is known.
- The spectrum is divided into bins and the event numbers in the bins follow Poisson distributions.

The competing hypotheses are

- H_{bkg} = background only
- H_{bs} = background + signal

then

$$p\left(H_{bkg} \middle| spectrum, I\right) = \frac{p\left(spectrum \middle| H_{bkg}, I\right)}{p\left(spectrum \middle| I\right)} p\left(H_{bkg} \middle| I\right)$$

$$p\left(H_{bs} \middle| spectrum, I\right) = \frac{p\left(spectrum \middle| H_{bs}, I\right)}{p\left(spectrum \middle| I\right)} p\left(H_{bs} \middle| I\right)$$

$$p\left(spectrum \middle| I\right) = p\left(spectrum \middle| H_{bkg}, I\right) p\left(H_{bkg} \middle| I\right) + p\left(spectrum \middle| H_{bs}, I\right) p\left(H_{bs} \middle| I\right)$$

A specific spectral shape depends on the average number of background (B) and signal (S) events, and we can write

prior distribution for the average B

$$p\left(spectrum \middle| H_{bkg}, I\right) = \int_B p\left(spectrum \middle| B, I\right) p_B\left(B\right) dB$$

$$p\left(spectrum \middle| H_{bs}, I\right) = \int_B p\left(spectrum \middle| B, S, I\right) p_B\left(B\right) p_S\left(S\right) dB\, dS$$

the possible spectra are the results of many possibile choices of the background and of the signal rates, and therefore of the average number of background and signal events; here we marginalize over these dependencies

prior distribution for the average S

We introduce the normalized spectral shapes of background and signal

$$f_B(E); \quad f_S(E);$$

Then we find the average number of events in each bin

$$\nu_i(B,S) = \nu_i(E_i, \Delta E_i, B, S) = B \int_{E_i}^{E_i+\Delta E_i} f_B(E)\,dE + S \int_{E_i}^{E_i+\Delta E_i} f_S(E)\,dE$$

An observed spectrum is defined by the numbers of counts in each bin: $\{n_i\}_{i=1,n}$ and since we assume a Poisson statistics in each bin, we find the following likelihoods for a given spectral observation

$$p(spectrum|B,I) = \prod_{i=1}^{N} \frac{[\nu_i(B,0)]^{n_i}}{n_i!} \exp[-\nu_i(B,0)]$$

$$p(spectrum|B,S,I) = \prod_{i=1}^{N} \frac{[\nu_i(B,S)]^{n_i}}{n_i!} \exp[-\nu_i(B,S)]$$

Finally we find:

$$p\left(spectrum \middle| H_{bkg}, I\right) = \int_B p\left(spectrum \middle| B, I\right) p_B(B)\, dB$$

$$= \int_B \prod_{i=1}^{N} \frac{\left[\nu_i(B,0)\right]^{n_i}}{n_i!} \exp\left[-\nu_i(B,0)\right] p_B(B)\, dB$$

$$p\left(spectrum \middle| H_{bs}, I\right) = \int_B p\left(spectrum \middle| B, S, I\right) p_B(B)\, p_S(S)\, dB$$

$$= \int_B \prod_{i=1}^{N} \frac{\left[\nu_i(B,S)\right]^{n_i}}{n_i!} \exp\left[-\nu_i(B,S)\right] p_B(B)\, p_S(S)\, dB$$

The final, complete expressions are:

$$p\left(H_{bkg}\middle|spectrum,I\right)=\frac{p\left(spectrum\middle|H_{bkg},I\right)}{p\left(spectrum\middle|I\right)}p\left(H_{bkg}\middle|I\right)$$

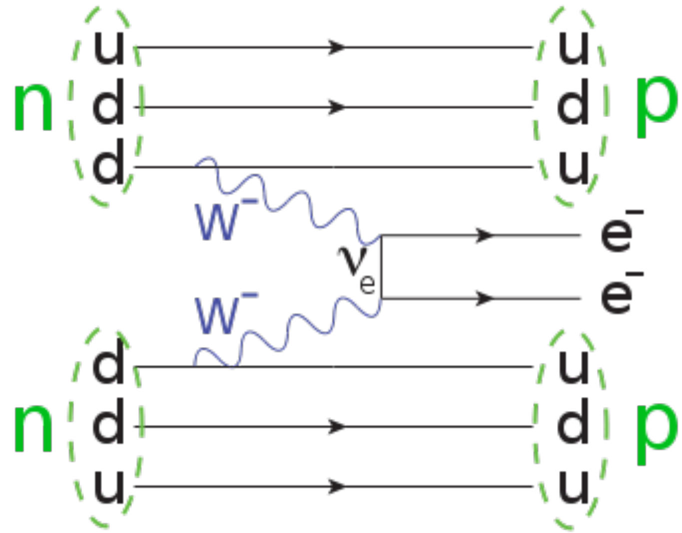
$$=\frac{\int_B\prod_{i=1}^{N}\frac{\left[\nu_i(B,0)\right]^{n_i}}{n_i!}\exp\left[-\nu_i(B,0)\right]p_B(B)dB}{\int_B\prod_{i=1}^{N}\frac{\left[\nu_i(B,0)\right]^{n_i}}{n_i!}\exp\left[-\nu_i(B,0)\right]p_B(B)dB\;p\left(H_{bkg}\middle|I\right)+\int_B\prod_{i=1}^{N}\frac{\left[\nu_i(B,S)\right]^{n_i}}{n_i!}\exp\left[-\nu_i(B,S)\right]p_B(B)p_S(S)dB\;p\left(H_{bs}\middle|I\right)}p\left(H_{bkg}\middle|I\right)$$

$$p\left(H_{bs}\middle|spectrum,I\right)=\frac{p\left(spectrum\middle|H_{bs},I\right)}{p\left(spectrum\middle|I\right)}p\left(H_{bs}\middle|I\right)$$

$$=\frac{\int_B\prod_{i=1}^{N}\frac{\left[\nu_i(B,S)\right]^{n_i}}{n_i!}\exp\left[-\nu_i(B,S)\right]p_B(B)p_S(S)dB}{\int_B\prod_{i=1}^{N}\frac{\left[\nu_i(B,0)\right]^{n_i}}{n_i!}\exp\left[-\nu_i(B,0)\right]p_B(B)dB\;p\left(H_{bkg}\middle|I\right)+\int_B\prod_{i=1}^{N}\frac{\left[\nu_i(B,S)\right]^{n_i}}{n_i!}\exp\left[-\nu_i(B,S)\right]p_B(B)p_S(S)dB\;p\left(H_{bs}\middle|I\right)}p\left(H_{bs}\middle|I\right)$$

We can use these expressions to test hypotheses (by means of Bayes factors), and find values for B and S.

The work of Caldwell and Kröninger has been carried out in the context of GERDA (GERmanium Detector Array), an experiment that aims to detect weak signals from neutrinoless beta decay in germanium detectors kept in a very low background environment.



Feynman diagram of neutrinoless double-beta decay, with two neutrons decaying to two protons. The only emitted products in this process are two electrons, which can occur if the neutrino and antineutrino are the same particle (i.e. Majorana neutrinos) so the same neutrino can be emitted and absorbed within the nucleus.
In conventional double-beta decay, two antineutrinos - one arising from each W vertex - are emitted from the nucleus, in addition to the two electrons. The detection of neutrinoless double-beta decay is thus a sensitive test of whether neutrinos are Majorana particles. (from Wikipedia)

Artist's view of GERDA
http://www.mpi-hd.mpg.de/gerda/

The 2007 Europhysics Conference on High Energy Physics IOP Publishing
Journal of Physics: Conference Series **110** (2008) 082010 doi:10.1088/1742-6596/110/8/082010

Table 1. Estimate of the background level expected in the GERDA experiment for a simplified Phase II setup at the present level of R&D.

Detector part	Contribution [10^{-4} counts/(kg·keV·y)]
Germanium detector (cosmogenic ^{68}Ge)	10.8
Germanium detector (cosmogenic ^{60}Co)	0.3
Germanium detector (bulk)	3.0
Germanium detector (surface)	3.5
Cabling	7.6
Copper holder	3.4
Electronics	3.5
Cryogenic liquid	0.1
Infrastructure	2.9
Muons and neutrons	2.0
Total	37.1

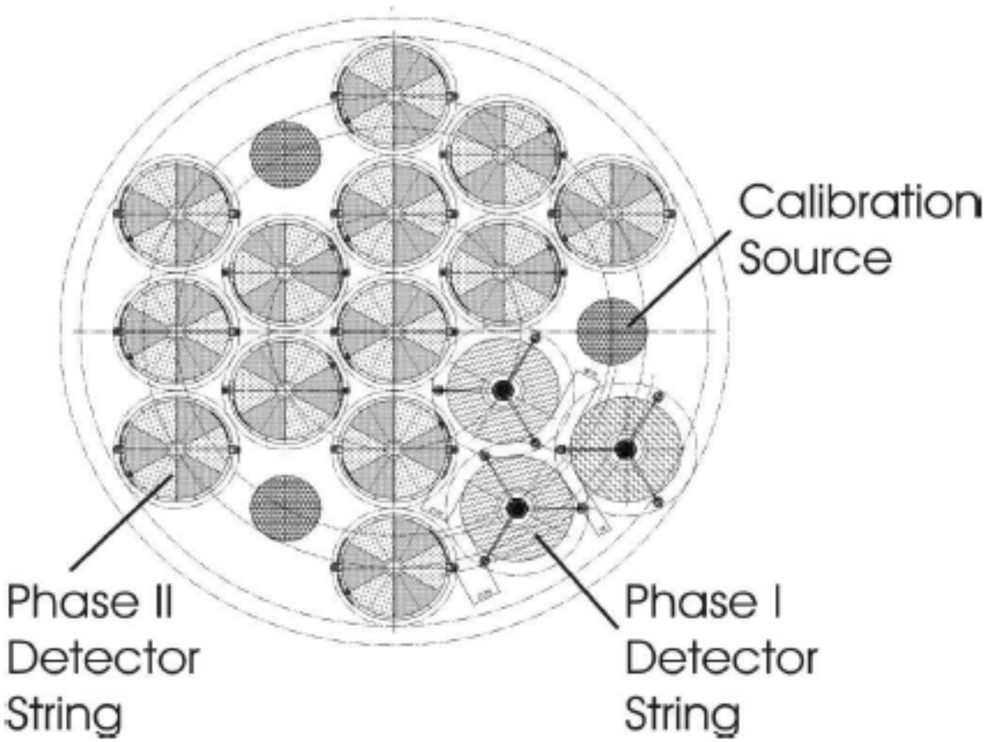


Figure 1. Detector array indicating the possible positions for the Phase I and Phase II detectors as well as for calibration sources.

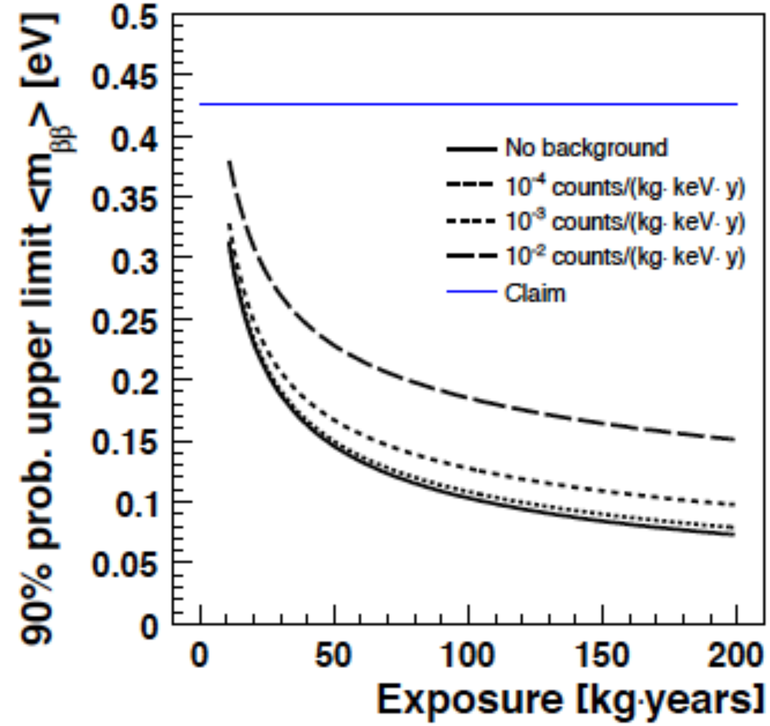


Figure 2. Expected 90% prob. upper limit on the effective Majorana neutrino mass, $\langle m_{\beta\beta} \rangle$, using the nuclear matrix elements presented in [14] with $\langle M^{0\nu} \rangle = 3.92$.

References:

• V. Dose: “Bayes in five days”, lecture notes, Max-Planck Research School on bounded plasmas, Greifswald, may 14-18 2002

• S. F. Gull: “Bayesian Inductive Inference and Maximum Entropy”, in G.J.Erickson and C. R. Smith (eds.): “Maximum Entropy and Bayesian Methods in Science and Engineering”, Kluwer, 1988