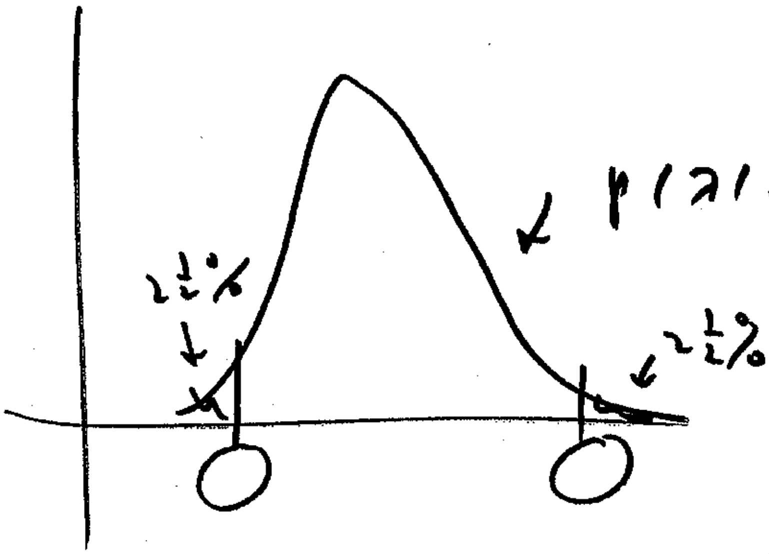
extra notes Feb 1 ①

part 2



$(y \mid \lambda) \sim$ Poisson(λ)

$E(y) = \lambda$

$$E(y_{n+1} \mid y) = \text{post. mean for } \lambda$$

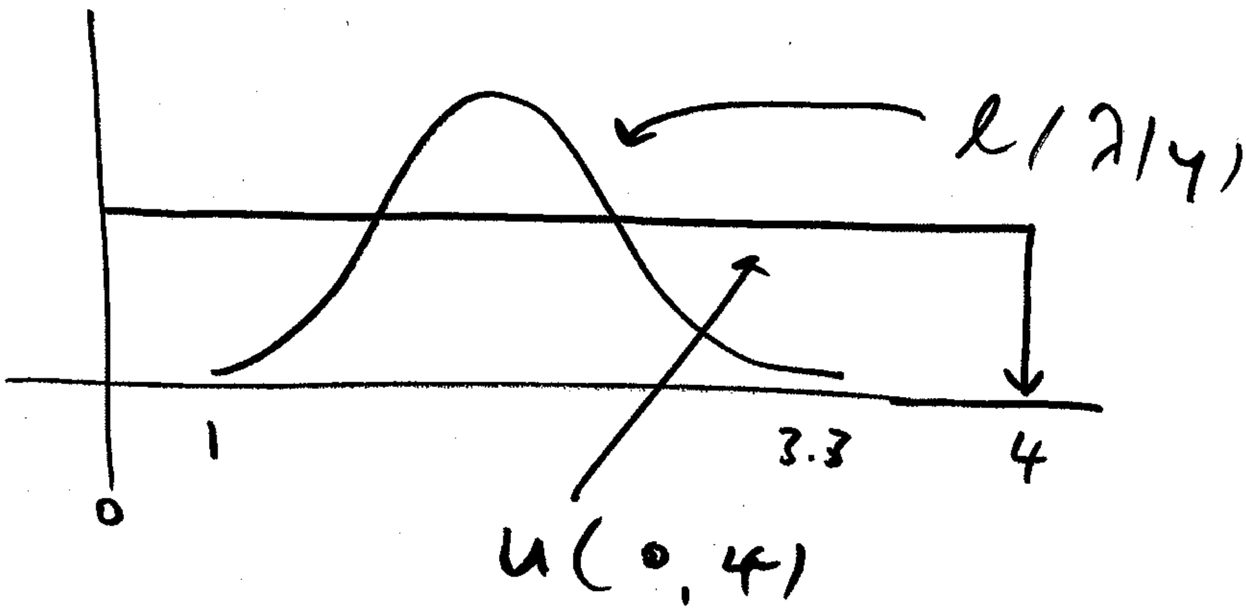
$$= \frac{\alpha^*}{\beta^*} = \frac{\alpha + s}{\beta + n}$$

$$V(y_{n+1} \mid y) = \frac{\alpha^*}{\beta^*}\left(1 + \frac{1}{\beta^*}\right) = O(1) + O\left(\frac{1}{n}\right)$$

$$= \frac{\alpha + s}{\beta + n}\left(1 + \frac{1}{\beta + n}\right) \doteq \bar{y}\left(1 + \frac{1}{n}\right) = \bar{y} + \frac{\bar{y}}{n}$$

$$V(\lambda \mid y) = \frac{\alpha^*}{(\beta^*)^2} = \frac{\alpha + s}{(\beta + n)^2} \doteq \frac{s}{n^2} = \frac{\bar{y}}{n} \quad (\alpha = \beta = 0)$$

$$= \underline{\underline{O}}\left(\frac{1}{n}\right)$$



binary coherence

$$\{Y_i\} \text{ exch} \xrightarrow{\text{coherence}} \left\{ \begin{array}{l} \theta \sim p(\theta) \\ (Y_i \mid \theta) \stackrel{iid}{\sim} \text{Bernoulli}(\theta) \end{array} \right\}$$

↓ have to think about

deF (1925) ← (hierarchical model)

θ = "underlying death rate" = $P(Y_i = 1 \mid \theta)$

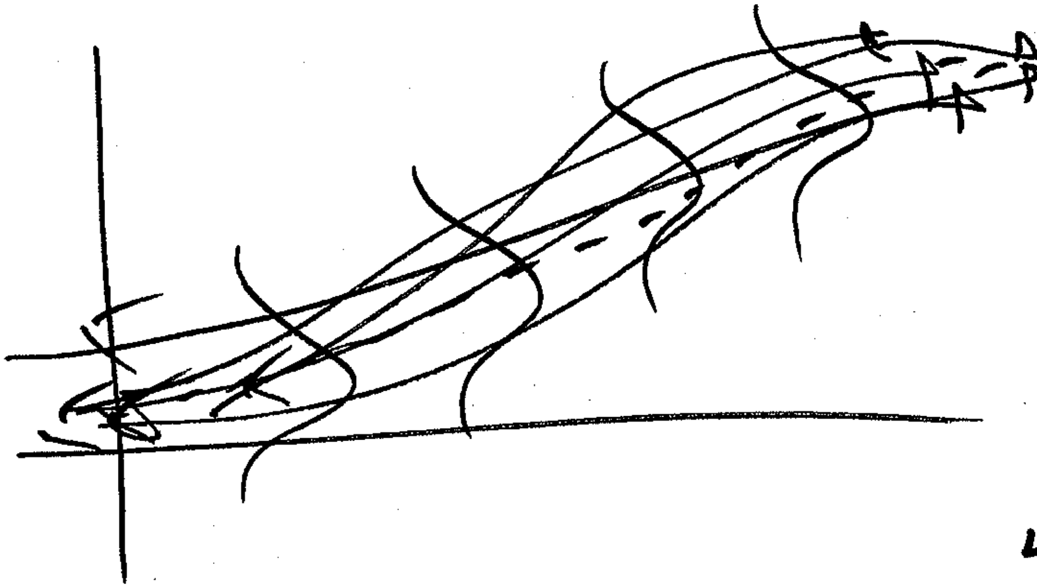
continuous

$$\{Y_i\} \text{ exch} \xrightarrow{\text{coh.}} \left\{ \begin{array}{l} F \sim p(F) \\ (Y_i \mid F) \stackrel{iid}{\sim} F \end{array} \right\}$$

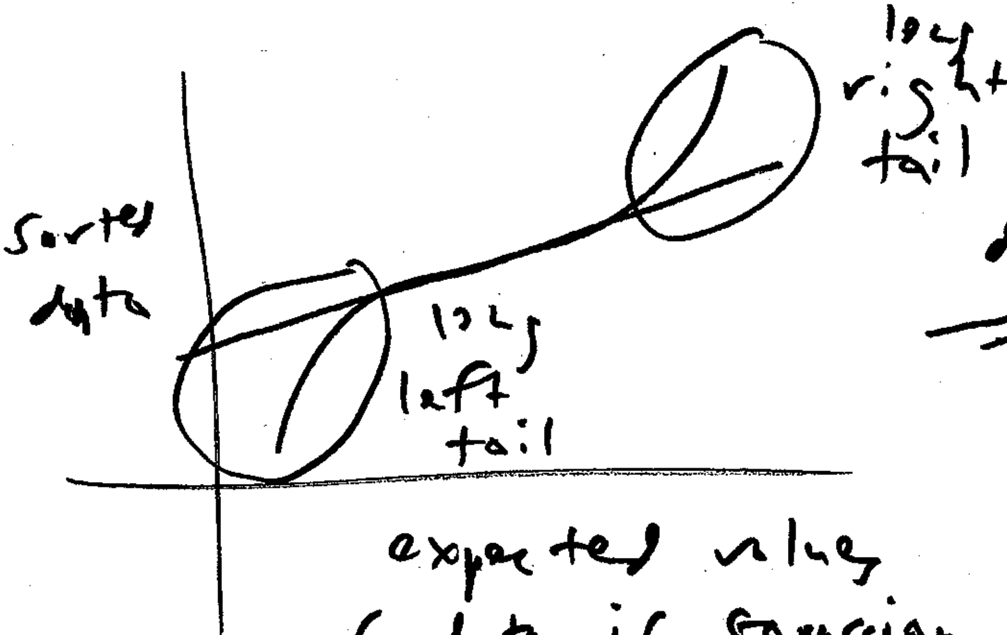
deF (1935)

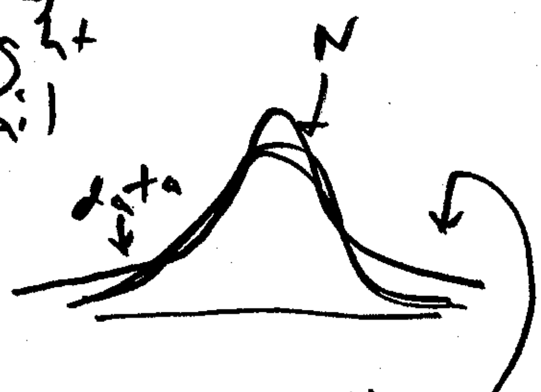
↓

have to think about F = "underlying CDF"



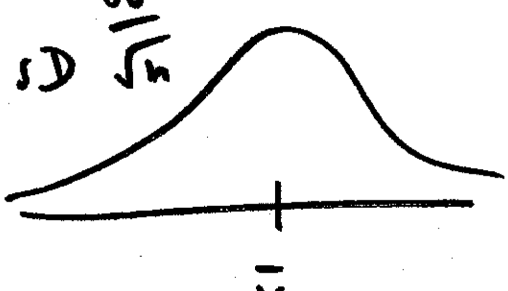
putting dist. on F → Bayesian nonparametrics





lik: $t_\nu(\mu, \sigma^2)$

shape loc scale



$\ell(\mu \mid y)$ in $N(\mu, \sigma_0^2)$ known

$$\left\{\begin{array}{l} \mu \sim N(\mu_0, \sigma_\mu^2) \\ (y_i \mid \mu) \stackrel{iid}{\sim} N(\mu, \underset{\text{known}}{\sigma_0^2}) \\ i = 1, \dots, n \end{array}\right\} \longrightarrow (\mu \mid y) = (\mu \mid \bar{y}) = N(\mu_*, \sigma_*^2)$$

$$\mu_* = \frac{\left(\frac{1}{\sigma_\mu^2}\right)\mu_0 + \left(\frac{n}{\sigma_0^2}\right)\bar{y}}{\frac{1}{\sigma_\mu^2} + \frac{n}{\sigma_0^2}}$$ = weighted ave of prior mean μ_0 & data mean $\bar{y}$ weighted by precisions

with weights given by $\boxed{\left(\frac{1}{\sigma_\mu^2}\right)}$ & $\boxed{\left(\frac{n}{\sigma_0^2}\right)}$, respectively; and

$$\boxed{\sigma_*^2 = \frac{1}{\frac{1}{\sigma_\mu^2} + \frac{n}{\sigma_0^2}}}$$

σ_μ^2 is prior variance of μ

$$\text{lik} = N\left(\bar{y}, \boxed{\frac{\sigma_0^2}{n}}\right)$$

$\frac{1}{\sigma_\mu^2}$ = prior precision

$\frac{n}{\sigma_0^2}$ = likelihood precision

$$\frac{1}{\sigma_*^2} = \frac{1}{\sigma_\mu^2} + \frac{n}{\sigma_0^2}$$

post prec. = prior prec + lik. prec

$$\mu_* = \frac{\left(\frac{\sigma_0^2}{\sigma_\mu^2}\right)\mu_0 + n\bar{y}}{\left(\frac{\sigma_0^2}{\sigma_\mu^2}\right) + n}$$

so prior sample size is $n_0 = \frac{\sigma_0^2}{\sigma_\mu^2}$

---

joint

$p(\theta_1, \theta_2, \ldots, \theta_{50} \mid y)$ hard to visualize

easier $p(\theta_1 \mid y), \; p(\theta_2 \mid y), \ldots,$ marginals

($k = 50$ of these)

$$p(\theta_1 \mid y) = \underbrace{\int \cdots \int}_{\leftarrow 49 \rightarrow} p(\theta_1, \ldots, \theta_{50} \mid y) \, \partial\theta_2 \ldots \partial\theta_{50}$$

#3 (PS 3)

$$\left\{ \begin{array}{l} \sigma^2 \sim p(\sigma^2) \\ (d_i \mid \sigma^2) \overset{iid}{\sim} N(0, \sigma^2) \end{array} \right\}$$

$\underline{\text{lik}}$: $p(d_1, \ldots, d_n \mid \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} d_i^2}$

$$\ell(\sigma^2 \mid y) = c\,(\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} d_i^2\right]$$

so $\sum_{i=1}^{n} d_i$ or equiv. $\frac{1}{n}\sum_{i=1}^{n} d_i^2$ is suff.
for σ^2 in this model

$$\ell\ell(\sigma^2 \mid y) = c - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} d_i^2$$

η

MLE of σ^2

$$\frac{\partial \ell\ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n} d_i^2 = 0$$

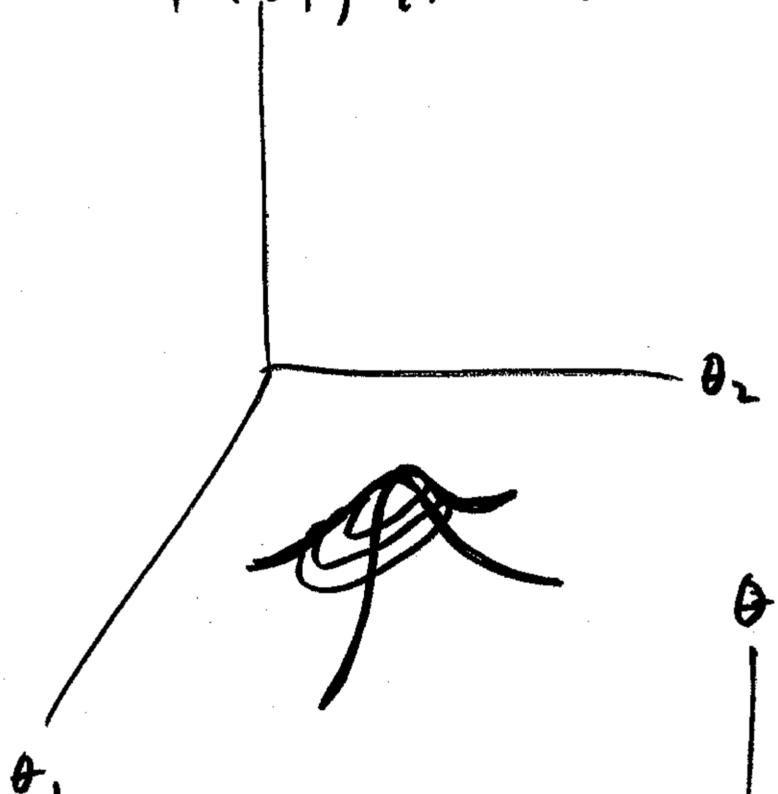
iff $\frac{n}{\sigma^2} = \frac{\sum_{i=1}^{n} d_i^2}{(\sigma^2)^2} \rightarrow \hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n} d_i^2$

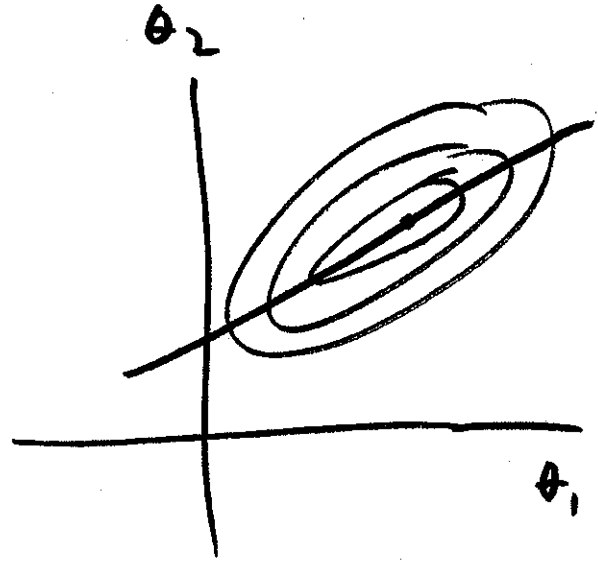
$= 191.8$

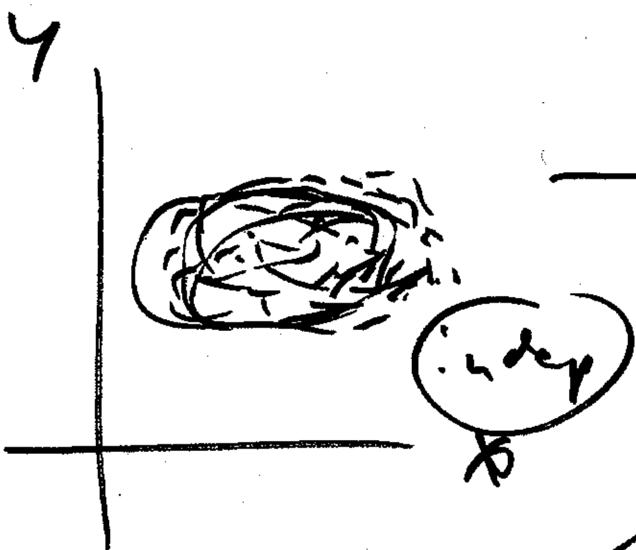
$$\ell(\eta \mid y) = c\, \eta^{-\frac{n}{2}} \exp\left(-\frac{ssq}{2\eta}\right)$$

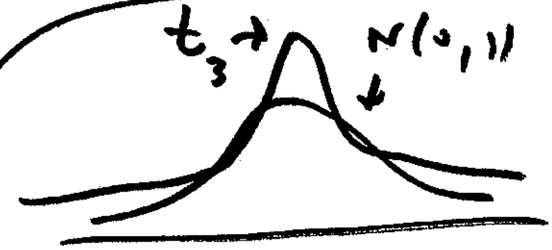
$$= \cancel{\Gamma}\ \chi^{-2}(\ ,\ )$$

$f(\theta_1, \theta_2)$ = bivariate normal







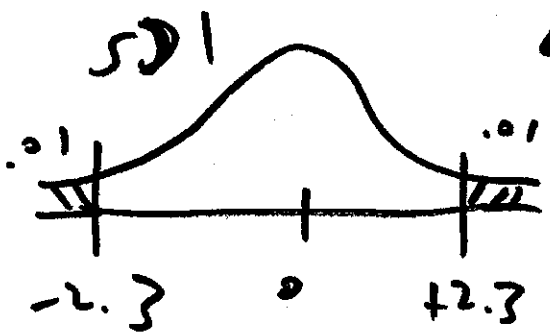


$$\frac{\bar{y}-c}{s/\sqrt{n}} \underset{H_0}{\sim} t_{n-1}$$

standard t dist

suggest you get winBUGS working
before next Tue if not already up & running

RS



dist. of z if null true

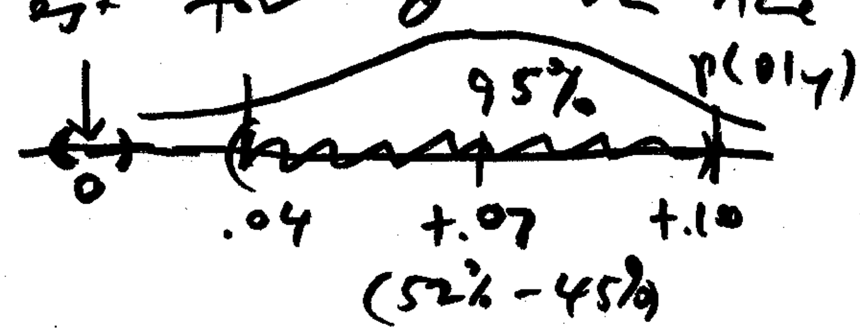
$$\boxed{\begin{array}{c} p = 0.02 \\ \updownarrow \\ z = \pm 2.3 \end{array}}$$

$$z = \frac{\text{signal}}{\text{noise}} = \frac{(\overline{\text{new}} - \overline{\text{old}})}{\widehat{SE}(\overline{\text{new}} - \overline{\text{old}})}$$

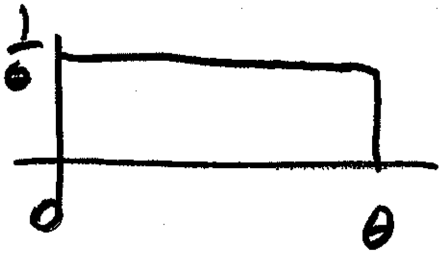
if need both $(\overline{\text{new}} - \overline{\text{old}})$ & $\widehat{SE}(\overline{\text{new}} - \overline{\text{old}})$ (and you do), why even bother with the fraction?

$$\boxed{(\overline{\text{new}} - \overline{\text{old}})} \pm 2 \boxed{\widehat{SE}(\overline{\text{new}} - \overline{\text{old}})}$$

= 95% int. est for θ on the right scale



$(Y_i \mid \theta) \overset{iid}{\sim} U(0, \theta)$



$$p(y_i \mid \theta) = \left\{ \begin{array}{ll} \frac{1}{\theta} & \text{if } 0 \le y_i \le \theta \\ 0 & \text{else} \end{array} \right\}$$

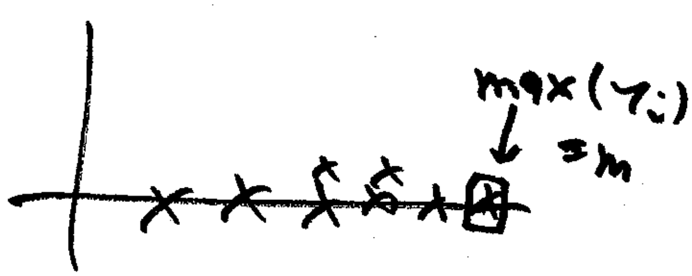
$$= \frac{1}{\theta} I(0 \le y_i \le \theta)$$

($I(A) = 1$ if A true, 0 else)

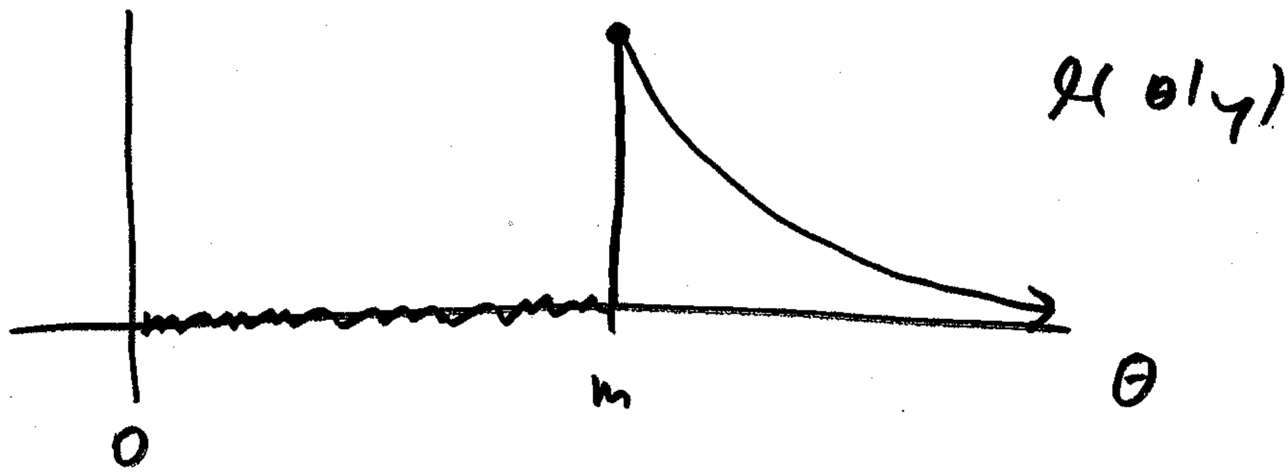
$$\ell(\theta \mid y) = \prod_{i=1}^{n} \frac{1}{\theta} I(0 \le y_i \le \theta)$$

$$= \theta^{-n} I(0 \le y_1 \le \theta \text{ and } 0 \le y_2 \le \theta \text{ and } \dots \text{ and } 0 \le y_n \le \theta)$$

$$\left(\begin{array}{c} \text{all of } y_i \ge 0 \\ \text{and } \le \theta \end{array} \right) \longleftrightarrow \left\{ 0 \le \max(y_i) \le \theta \right\}$$



so $\ell(\theta|y) = c\,\theta^{-n}\, I[\theta \ge \max(y_i)]$



by inspection $\hat{\theta}_{MLE} = \max(y_i)$

$\theta \sim \text{pareto}(\alpha, \beta) \to p(\theta) = \begin{cases} \alpha \beta^{\alpha} \theta^{-(\alpha+1)} & \text{if } \theta \ge \beta \\ 0 & \text{else} \end{cases}$

$\ell(\theta|y) = c\,\text{Pareto}(n-1, m)$ $\qquad \theta^{-n}$

conj. prior $\theta \sim \text{Pareto}(\alpha, \beta)$

$(\theta|y) \sim \text{Pareto}[\underline{\alpha+n}, \max(\beta, m)]$

data: $m = 5.1$, $n = 11$

$p(\theta|y) = \text{Pareto}(10, \underline{5.1})$

prior: Pareto (2.5, 4)

---

post: $p(\theta|y) \sim$ Pareto (13.5, 5.1)

$$p(\theta|y) = \text{Pareto}(\alpha+n, \max(\beta, m))$$

$$V(\theta) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)}$$

$(y_i|\theta) \sim$ iid $U(0,\theta)$

as $n \to \infty$
$m \to \theta$

$$V(\theta|y) = \frac{(\alpha+n)\left[\max(\beta,m)\right]^2}{(\alpha+n-1)^2(\alpha+n-2)}$$

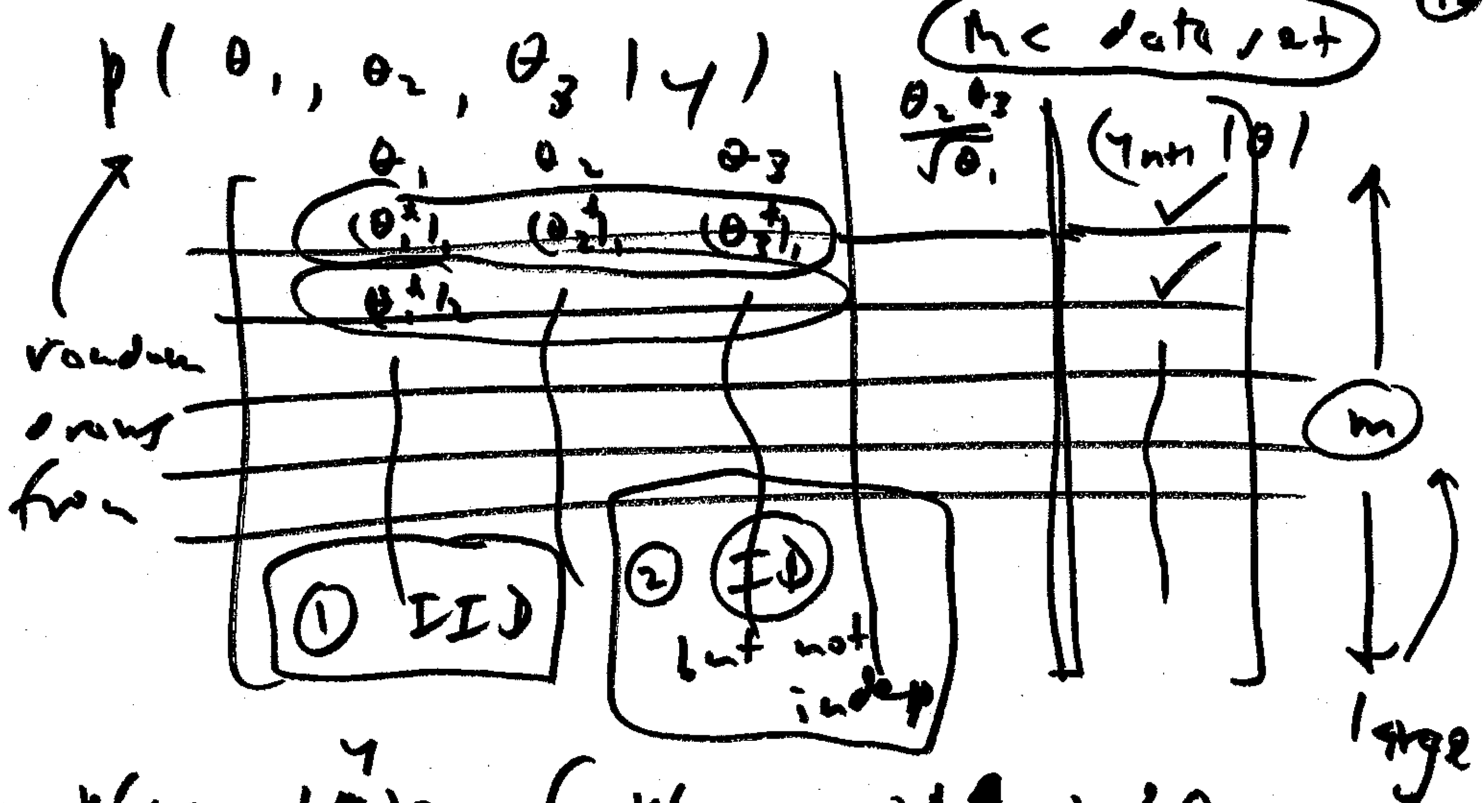
usually $V(\theta|y) = \downarrow 0$ as $n \uparrow$

& moreover $V(\theta|y) = \underline{\underline{O}}\left(\frac{1}{n}\right)$
$SD(\theta|y) \doteq c/\sqrt{n}$

but here $V(\theta|y) \downarrow 0$ as $n \uparrow$

but $V(\theta|y) = \underline{\underline{O}}\left(\frac{1}{n^2}\right) \leftarrow$
$SD(\theta|y) = c/n$

$$p(\theta_1, \theta_2, \theta_3 \mid y)$$



$$p(y_{n+1} \mid y) = \int p(y_{n+1}, \theta \mid y) \, d\theta$$

$$= \int p(y_{n+1} \mid \theta, \cancel{y}) \, p(\theta \mid y) \, d\theta$$

$$= \int p(y_{n+1} \mid \theta) \, p(\theta \mid y) \, d\theta$$

$= $ mixture of $p(y_{n+1} \mid \theta)$ weighted by $p(\theta \mid y)$

$$\boxed{(y_{n+1} \mid y) \sim \begin{array}{l} (\theta \mid y) \\ (y_{n+1} \mid \theta, \cancel{y}) \end{array}}$$

$$p(\underset{\sim}{\theta} \mid \underset{\sim}{y}) = c \, p(\underset{\sim}{\theta}) \, \ell(\underset{\sim}{\theta} \mid \underset{\sim}{y})$$