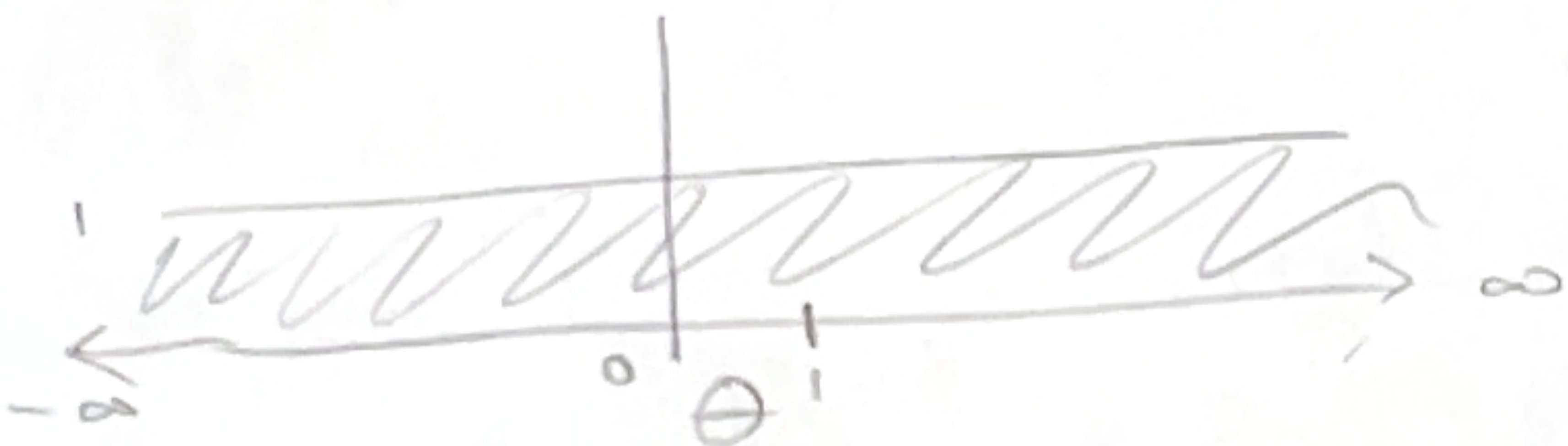
$$\int_0^{\sigma^2}\int_{-\infty}^{\infty} p(\theta,\sigma^2)\,d\theta\,d\sigma^2$$

$$= \int_0^{\sigma^2}\frac{1}{\sigma^2}\,d\sigma^2 \cdot \underbrace{\int_{-\infty}^{\infty} 1\cdot d\theta}_{\theta\big|_{-\infty}^{\infty}}$$



constant

$\mathrm{var}(Ax)$

r.v.

$A\,\mathrm{var}(x)\,A^T$

$$\hat\beta = \underset{p\times n\ \ n\times p}{(X^TX)^{-1}}\ \underset{p\times n\ \ n\times 1}{X^T y}$$

$$\mathrm{var}(\hat\beta) = (X^TX)^{-1} X^T \underset{\uparrow\ \sigma^2 I}{\mathrm{var}(y)}\, X (X^TX)^{-1}$$

variance $= \sigma^2 \cancel{(X^TX)^{-1} X^TX} (X^TX)^{-1}$

unit info.

Precision $= \frac{1}{\sigma^2}(X^TX)\cdot\frac{1}{n}$

$\det(\underset{n \times n}{cA}) = c^n \cdot \det(A)$

MLE σ^2 in MVN

$$p(Y \mid \beta, \sigma^2 I) = \left(2\pi \left| I\sigma^2 \right|\right)^{-1/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}(Y - X\hat{\beta})^T I (Y - X\hat{\beta})\right\}$$

$$= c\,(\sigma^2)^{-n/2} \cdot \exp\left\{-\frac{1}{2\sigma^2} SSR\right\}$$

rewrite as $\lambda = 1/\sigma^2$

$$L(\lambda) = p(Y \mid \beta, \lambda) = c\,\lambda^{n/2} \exp\left\{-\frac{\lambda}{2} \cdot SSR\right\}$$

$$\log L(\lambda) = \frac{n}{2}\log\lambda - \frac{\lambda}{2} SSR + c$$

$$\frac{d}{d\lambda}\log L(\lambda) = \frac{n}{\cancel{2}\lambda} - \frac{SSR}{\cancel{2}}$$

set equal to 0

$$\lambda = \frac{n}{SSR} \Rightarrow \hat{\sigma}^2 = \frac{SSR}{n}$$

Exercise : Unit information

$$\frac{1}{n} \log \left(\theta^{\Sigma y_i} (1-\theta)^{n-\Sigma y_i}\right)$$

$$= \frac{1}{n}\left[\Sigma y_i \log\theta + (n-\Sigma y_i)\log(1-\theta)\right]$$

$$= \frac{1}{\cancel{n}}\left[\cancel{n}\bar{y}\log\theta + \cancel{n}(1-\bar{y})\log(1-\theta)\right] \quad A$$

Set $\log p(\theta) =$ to A $+ c$

so when I exponentiate,

$$p(\theta) = c \cdot \theta^{\bar{y}}(1-\theta)^{1-\bar{y}}$$

$$\theta \sim \text{beta}(\bar{y}+1, 2-\bar{y})$$

Exercise Jeffreys prior

$$p(Y|\theta) = \frac{\theta^y e^{-\theta}}{y!}$$

$$\ell(\theta) = \log p(Y|\theta) = y \log\theta - \theta - \log(y!)$$

$$\frac{d}{d\theta}\ell(\theta) = \frac{y}{\theta} - 1$$

$$\frac{d^2}{d\theta^2}\ell(\theta) = \frac{-y}{\theta^2}$$

$$-\mathbb{E}\, \frac{-y}{\theta^2}\Big|\theta = \int \frac{y}{\theta^2} p(y|\theta)\, dy$$

$$= \frac{1}{\theta^2} = \underbrace{\int y p(y|\theta)\, dy}_{\theta}$$

$$I(\theta) = \frac{1}{\theta}$$

We know $p(\theta) \propto \sqrt{I(\theta)} = \theta^{-1/2}$

$$p(\theta|\vec{y}) \propto \underbrace{\theta^{\Sigma y_i - 1/2} e^{-n\theta}}_{\text{gamma}(\Sigma y_i + 1/2,\ n)}$$

Let $\phi = \log\theta \Rightarrow \theta = e^{\phi}$

$$p(\phi) = p(\theta)\left|\frac{d\theta}{d\phi}\right| = e^{-\phi/2} \cdot e^{\phi} = e^{\phi/2}$$

$$\begin{array}{ccc} p(Y|\theta) & & p(Y|\phi = g(\theta)) \\ \downarrow & & \downarrow \\ J(\theta) & \longleftrightarrow & J(\phi) \end{array}$$

Reference priors.

Goal: <u>max</u> KL divergence btwn $p(\theta)$
Kullback-Leibler
& $p(\theta|\vec{y})$ (averaged over possible data).

<u>KL</u>-divergence:

$$D_{KL}(p(\theta|y) \| p(\theta)) = \int p(\theta|y) \log \frac{p(\theta|y)}{p(\theta)} d\theta$$

Avg over ys:

$$\mathbb{E}_y\left[D_{KL}(p(\theta|y) \| p(\theta))\right] = \int p(y) \int p(\theta|y) \log \frac{p(\theta|y)}{p(\theta)} d\theta$$

$$= \iint p(y,\theta) \log \frac{p(\theta,y)}{p(\theta)p(y)} d\theta\, dy$$

$$= I(\theta, y) \quad \text{"mutual info"}$$

To choose a reference prior, one must find

$$p(\theta) = \underset{p(\theta)}{\operatorname{argmax}}\; I(\theta, y)$$