341 Lec11

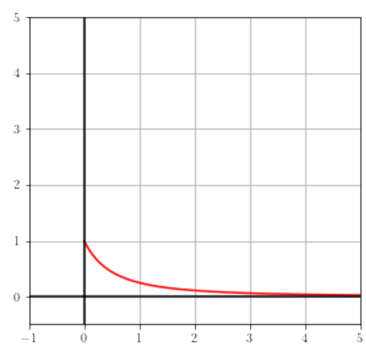
Andrew Claros

“Odds”

- Consider $X \sim \text{Bernoulli}\,(\theta)\,, \theta = \mathbb{P}\,(x=1)\,, \theta \in \Theta = (0,1)$
- We can “*parametrize*” the Bernoulli.
- $\phi = t(\theta) = \frac{\theta}{1-\theta}, \phi \in (0,\infty), \Rightarrow \phi - \theta\phi = \theta$
- $\Rightarrow \phi = \theta + \theta\phi$
- $\Rightarrow \phi = \theta(1+\phi)$
- $\theta = t^{-1}\theta = \frac{\phi}{1+\phi}$

Odds as prior

- Laplace: $\mathbb{P}\,(\theta) = \text{U}\,(0,\,1)$ Therefore we want: $\mathbb{P}\,(\phi) = \text{Uniform}$ but this is impossible because we cannot have a uniform prior of $\text{U}\,(0,\,\infty)$
- $x \sim \text{Bernoulli}\,(\phi) = (\frac{\phi}{1+\phi})^x(\frac{1}{1+\theta})^{1-x} = \frac{\theta^x}{1+\phi}$
- $\mathcal{P}_\phi = \mathcal{P}_\theta(t^{-1}(\theta))|\frac{d}{d\theta[t^{-1}(\theta)]}| = \underbrace{\mathcal{P}_\theta(\frac{\phi}{1+\phi})}_{1}|\frac{d}{d\theta}[\frac{\phi}{1+\phi}]|$
- $= |\frac{d}{d\phi}[\frac{\phi}{1+\phi}] = |\frac{(1+\theta)(1)-\phi(1)}{(1+\phi)^2}|\frac{1}{(1+\phi)^2} \Rightarrow$ Fisher-Snedecor distribution



- This is a valid density function: $\int_0^\infty \frac{1}{(1+\phi)}d\phi = [\frac{\phi}{1+\phi}]_0^\infty = 1-0$

Conclusion: By being indifferent on the probability scale, we <u>cannot</u> be indifferent on the odds scale

- Fisher used this conclusion to discredit Laplace’s prior and Bayesian statistics.
- changing parametrization changes inference

Fixing the problem:

- Let θ be a parameter of $\mathcal{F}$ and $t(\theta) = \phi$, a 1:1 reparametrization.
- $F := \mathbb{P}\left(x|\theta\right) \overset{procedure}{\longrightarrow} \mathbb{P}\left(\theta\right)$
- $\mathbb{P}\left(x|\theta\right) \leftrightarrows_{t}^{t^{-1}} \mathbb{P}\left(x|\phi\right)$
- $\mathbb{P}\left(x|\phi\right) \overset{procedure}{\longrightarrow} \mathbb{P}\left(\phi\right)$

Jeffrey's Priors

Kernals

- $f(x;\theta) \propto k(x;\theta) \Rightarrow \exists c \in \mathbb{R}, f(x;\theta) \quad ck(x;\theta)$ where k,f are 1:1
- meaning: f, k differ by c
- $\int_{\text{Supp}[X]} f(x:\theta)dx = 1 \Rightarrow \int ck(x;\theta)dx = 1 \Rightarrow \int k(x;\theta)d\theta = \frac{1}{c} = \frac{1}{\int_{\text{Supp}[X]} k(x;\theta)dx}$

Apply to $\text{Beta}\left(\alpha,\, \beta\right)$

- $Y \sim \text{Beta}\left(\alpha,\, \beta\right) = \frac{1}{\beta(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1} \propto y^{\alpha-1}(1-y)^{\beta-1} = k(y;\alpha,b)$

Apply to $\mathcal{F}$: $\text{Binomial}\left(n,\, \theta\right)$, n fixed, $\mathbb{P}\left(\theta\right) = \text{Beta}\left(\alpha,\, \beta\right)$

- $\mathbb{P}\left(\theta|x\right) = \frac{\mathbb{P}(x|\theta)\mathbb{P}(\theta)}{\mathbb{P}(x)} \propto \mathbb{P}\left(x|\theta\right)\mathbb{P}\left(\theta\right) = \underbrace{\binom{n}{x}\theta^x(1-\theta)^{n-x}}_{\mathbb{P}(x|\theta)} \underbrace{\frac{1}{\beta(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}_{\mathbb{P}(\theta)}$

$\propto \underbrace{\theta^x(1-\theta)^{n-x}\theta^{\alpha-1}}_{\mathbb{P}(x|\theta)} \underbrace{(1-\theta)^{\beta-1}}_{\mathbb{P}(\theta)} = \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1} \propto \text{Beta}\left(x+\alpha,\, n-x+\beta-1\right)$

HW

- $Y \sim \mathcal{N}\left(\theta,\, \sigma^2\right) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\theta)^2}}_{f(y;\theta,\sigma^2)} \propto \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right) = \exp\left(-\frac{1}{2\sigma^2}(y^2 - 2y\theta + \theta^2)\right)$

$\exp\left(-\frac{y^2}{2\sigma^2}\right)\exp\left(\frac{y\theta}{\sigma^2}\right)\exp\left(-\frac{\theta^2}{2\sigma^2}\right) \propto \exp\left(-\frac{y^2}{2\sigma^2}\right)\exp\left(\frac{y\theta}{\sigma^2}\right) = k(y;\theta,\sigma^2)$

Going backwards: $c = \frac{1}{2\pi\sigma^2} e^{-\frac{\theta^2}{2\sigma^2}}$

Fisher Information $x =< x_1, \ldots, x_n >$

- Recall: $\mathcal{L}(\theta;x) = \mathbb{P}\left(x;\theta\right)$
- $\ell\theta; x := \ln(\mathcal{L}(\theta;x))$
- Score function: $\mathcal{S}(\theta;x) := \frac{d}{d\theta}[\ell(\theta;x)]$. $\mathcal{S}(\theta;x) \overset{\text{set}}{=} 0$, solve for θ
- Fisher Information: $\mathcal{I}(\theta) := \text{Var}_X[\mathcal{S}(\theta;x)] \ldots = \mathbb{E}_X[-\ell''(\theta;x)]$

Fisher information of one $X \sim \text{Bernoulli}\,(\theta)$

- $\mathcal{L}(\theta;x) = \theta^x(1-\theta) \Rightarrow \ell(\theta;x) = x\ln(\theta) + (1-x)\ln(1-\theta)$
  $\ell'(\theta;x) = \frac{x}{\theta} - \frac{1-x}{1-\theta} \Rightarrow \ell''(\theta;x) = -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2}$

- $\mathcal{L}(\theta) = \mathbb{E}_x[\frac{x}{\theta^2} + \frac{1-x}{(1-\theta)^2}] = \frac{1}{\theta^2}\mathbb{E}\,[X] + \frac{1}{(1-\theta)^2}(1-\mathbb{E}\,[X])$
  $= \frac{1}{\theta^2}\theta + \frac{1}{(1-\theta)^2}(1-\theta) = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}$

Application

Theorem Jeffrey's Prior: $\mathbb{P}_J(\theta) \propto \sqrt{\mathcal{I}(\theta)}$

- $\mathcal{F}$: Binomial $(n,\, \theta) \Rightarrow \mathcal{L}(\theta;x) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \Rightarrow \mathcal{L}(\theta;x) = \ln(\binom{n}{x}) + x\ln(\theta) + (n-x)\ln(1-\theta)$

- $\ell'(\theta;x) = \frac{x}{\theta} - \frac{n-x}{1-\theta} \Rightarrow \ell'' = -\frac{x}{\theta^2} - \frac{n-x}{(1-x)^2}$

- $\mathcal{I} = \mathbb{E}\,[-\ell''] = \mathbb{E}\left[\frac{X}{\theta^2} + \frac{n-X}{(1-\theta)^2}\right] = \frac{1}{\theta^2}\mathbb{E}\,[X] + \frac{1}{(1-\theta)^2}(n-\mathbb{E}\,[X])$
  $= \frac{1}{\theta^2}n\theta + \frac{1}{(1-\theta)^2}(n-n\theta) = n(\frac{1}{\theta} + \frac{1}{1-\theta}) = \frac{n}{\theta(1-\theta)}$

Answer: $\mathbb{P}_J(\theta) \propto \sqrt{\frac{n}{\theta(1-\theta)}} \propto \sqrt{\frac{1}{\theta(1-\theta)}} = \theta^{\frac{1}{2}}(1-\theta)^{\frac{1}{2}} \propto \text{Beta}\left(\frac{1}{2},\, \frac{1}{2}\right)$ Jeffrey's Prior is Beta $\left(\frac{1}{2},\right.$

- $\mathbb{P}\,(X|\theta) \overset{\text{Jeffrey's Procedure}}{\longrightarrow} \mathbb{J}\theta = \text{Beta}\left(\frac{1}{2},\, \frac{1}{2}\right)$

- Through $\mathbb{P}\,(X|\theta) \leftrightarrows_t^{t^{-1}} \mathbb{P}\,(X|\phi)$

- $\mathbb{P}\,(X|\phi) \overset{\text{Jeffrey's procedure}}{\longrightarrow} \mathbb{J}\phi =?$

- $\mathbb{P}_J(\theta) \leftrightarrows_t^{t^{-1}} \mathbb{P}_J(\phi)$

- Later will verify using $\phi = t(\theta) = \frac{\theta}{(1-\theta)}$ ("*Odds*").