3/7/17

$\underline{S}(\theta; x) = \ell'(\theta; x)$

$I(\theta) = \mathrm{Var}[S(\theta; x)] \ldots = E(S(\theta; x)^2]$

$= \ldots = E(-\ell''(\theta; x))$

Fisher information

$X \sim \text{Binom}(n, \theta)$

$L(\theta; x) = P(X) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$

$\ell(\theta; x) = \ln \binom{n}{x} + x \ln(\theta) + (n-x)\ln(1-\theta)$

$S(\theta; x) = \frac{x}{\theta} - \frac{n-x}{1-\theta}$

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

$I(\theta) = E[-\ell''(\theta; x)] = E\left[\frac{x}{\theta^2} + \frac{n-x}{(1-\theta)^2}\right]$

$= \frac{E(X)}{\theta^2} + \frac{n - E(X)}{(1-\theta)^2}$

$= n\left(\frac{\theta}{\theta^2} + \frac{(1-\theta)}{(1-\theta)^3}\right) = \frac{n\theta}{\theta^2} + \frac{n - n\theta}{(1-\theta)^2} = n\left(\frac{1}{\theta(1-\theta)}\right)$

$= n\left(\frac{1}{\theta} + \frac{1}{1-\theta}\right)$

$X_1 \sim \text{Binom}(1, .5) \Rightarrow I(\theta) = 4$

$X_2 \sim \text{Binom}(1, .01) \Rightarrow I(\theta) = 101.01$

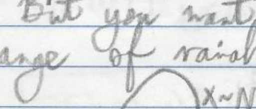
Given $F = P(X|\theta)$, pick $p(\theta)$.

$\phi = f(\theta)$ s.t. f is 1:1 smooth

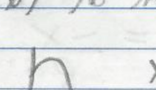
$p(x|\theta) \xrightarrow{\text{pick}} p(\theta)$

$p(x|\phi) \xrightarrow{\text{pick}} p(\phi)$

But you want $p(\theta), p(\phi)$ to be related via change of variables



$X \sim N(0,1)$



$X^2 \sim \chi^2_1$

$P(\theta) \propto \sqrt{I(\theta)}$ Jeffreys Prior

Pseudocounts

$X \sim \text{Binom}(n, \theta)$

$\Rightarrow p(\theta) \propto \sqrt{n\left(\frac{1}{\theta(1-\theta)}\right)} \propto \sqrt{\frac{1}{\theta(1-\theta)}} = \theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$

$\text{Beta}(0,0)$

$\Updownarrow$

$\propto \text{Beta}\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{\underbrace{B\left(\frac{1}{2}, \frac{1}{2}\right)}_{\pi}} \theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$

$\Updownarrow$

$\text{Beta}(1,1)$

$= \frac{1}{\pi\sqrt{\theta(1-\theta)}}$ Arcsine distribution

$p(\theta) = \frac{1}{\pi\sqrt{\theta(1-\theta)}}$

$R = f(\theta) = \frac{\theta}{1-\theta}$, $\theta = f^{-1}(R) = \frac{R}{R+1}$

$X \sim \text{Binom}(n, R)$

$P(X|R) = \binom{n}{x}\left(\frac{R}{R+1}\right)^x \Big(\underbrace{1 - \frac{R}{R+1}}_{\frac{1}{R+1}}\Big)^{n-x}$

$= \binom{n}{x} \frac{R^x}{(R+1)^n}$

$\ell(x; R) = \ln\binom{n}{x} + x\ln(R) - n\ln(R+1)$

$\ell'(x; R) = \frac{x}{R} - \frac{n}{R+1}$

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

$I(R) = E(-\ell''(R; x))$

$= E\left[\frac{x}{R^2} - \frac{n}{(R+1)^2}\right] = \frac{E(X)}{R^2} - \frac{n}{(R+1)^2}$

$$= n\frac{\frac{R}{R+1}}{R^2} - \frac{n}{(R+1)^2} \qquad = p(\phi)$$

$$= n\left(\frac{\cancel{R+1}\,R}{R(R+1)} - \frac{R}{(R+1)^2}\right)$$

$$= n\,\frac{1}{R(R+1)^2}$$

$$\Rightarrow p(R) \propto \sqrt{\frac{n}{R(R+1)^2}} \propto \frac{1}{\sqrt{R}}\frac{1}{R+1} \propto \frac{1}{\pi}\frac{1}{\sqrt{R}}\frac{1}{\sqrt{R+1}}$$

$$p_R(R) = p_\theta(f^{-1}(R))\left|\frac{d}{dR}\left[f^{-1}(R)\right]\right|$$

$$= \frac{1}{\pi}\left(\frac{R}{R+1}\right)^{-1/2}\left(\frac{1}{R+1}\right)^{-1/2}\left(\frac{1}{(R+1)^2}\right)$$

$$= \frac{1}{\pi}R^{-1/2}(R+1)^{1}\frac{1}{(R+1)^2} = \frac{1}{\pi}\frac{1}{\sqrt{R}}\frac{1}{(R+1)}$$

$P(X|\theta),\ P(X|\phi),$

$p(\theta) \propto \sqrt{I(\theta)}$

$p(\phi) \propto \sqrt{I(\phi)}$

$$p(\phi) = p_\theta(f^{-1}(\phi))\left|\frac{d}{d\phi}\left[f^{-1}(\phi)\right]\right| \propto \sqrt{I(\phi)}$$

w/ $\theta = f^{-1}(\phi)$

$$= p_\theta(\theta)\left|\frac{d\theta}{d\phi}\right| \propto \sqrt{I(\theta)}\left|\frac{d\theta}{d\phi}\right| = \sqrt{I(\theta)\frac{d\theta}{d\phi}\frac{d\theta}{d\phi}}$$

$$= \sqrt{E[S(\theta;x)]\frac{d\theta}{d\phi}\frac{d\theta}{d\phi}} \qquad S = \ell'$$

$$= \sqrt{E\left[S(\theta;x)^2\frac{d\theta}{d\phi}\frac{d\theta}{d\phi}\right]}$$

$$E\left(\left(\frac{dl}{d\theta}\right)^2\right) \to s$$

$$= \sqrt{E\left[\frac{dl}{d\phi}\frac{dl}{d\theta}\frac{d\theta}{d\phi}\frac{d\theta}{d\phi}\right]}$$

$$= \sqrt{E[S(\phi, x)^2]} = \sqrt{I(\phi)}$$

$X \mid \theta \sim \text{Binom}(n, \theta)$

$\to \theta \sim \text{Beta}(\alpha, \beta)$

$\theta \mid x \sim \text{Beta}(\alpha + x, \beta + n - x) \Rightarrow \hat{\theta}_{MMSE} = \dfrac{\overbrace{\alpha + x}^{\alpha'}}{\underbrace{n + \alpha + \beta}_{\beta'}}$

batting average (BA), θ.

$\hat{\theta} := BA = \dfrac{\#\text{hits}}{\#\text{at bats}} = \dfrac{x}{n}$

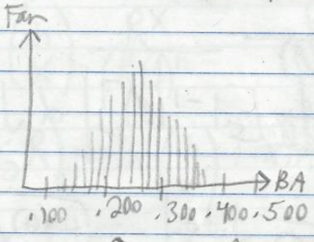
$\#\text{hits} \sim \text{Binom}(\#\text{at bats}, \theta)$

$n = 2$

$x = 0 \Rightarrow BA = 0$

$x = 1 \Rightarrow BA = 0.5$

$x = 2 \Rightarrow BA = 1$



Fix a $\theta \sim \text{Beta}(\alpha, \beta)$

$\hat{\alpha}_{MLE} = 78.7$

$\hat{\beta}_{MLE} = 224.8$

$\hat{\alpha}_{MLE} + \hat{\beta}_{MLE} = 303.5 \Rightarrow \hat{\theta}$

$\Rightarrow \hat{\theta}_{MMSE} = \dfrac{x + 78.7}{n + 303.5}$

Empirical Bayes

① Get all data

② Fit prior to all data
③ Use this fit's hyperparameters for inference

Finished Beta-Binomial

$F = \text{Geometric} \Rightarrow X|\theta \sim (1-\theta)^x \theta$

X represents # of "failures"

$\text{Supp}(X) = \{0, 1, \ldots\} = \mathbb{N}_0$

$\Theta = (0, 1)$

$E(X) = \frac{1}{\theta} - 1$

$\theta \uparrow \Rightarrow X \downarrow$

$\theta \downarrow \Rightarrow X \uparrow$

$X_1|\theta, \ldots, X_n|\theta \overset{iid}{\sim} \text{Geo}(\theta)$

$P(\theta|X) \propto P(X|\theta)P(\theta) = \overbrace{(1-\theta)^{\Sigma x_i}\theta^n}^{\text{kernel for beta}} P(\theta)$

$P(X|\theta) = \prod_{i=1}^{n} (1-\theta)^{x_i}\theta = (1-\theta)^{\Sigma x_i}\theta^n$

Let $P(\theta) = \text{Beta}(\alpha, \beta)$

$\propto \theta^n (1-\theta)^{\Sigma x_i} \theta^{\alpha-1}(1-\theta)^{\beta-1}$

$= \theta^{n+\alpha-1}(1-\theta)^{\Sigma x_i + \beta - 1}$

$\propto \text{Beta}(n+\alpha, \Sigma_i x_i + \beta)$

$\Rightarrow$ Beta is also the conjugate prior for the geometric r.v.

$X_1|\theta, \ldots, X_n|\theta \overset{iid}{\sim} \text{Geom}(\theta)$

$\theta \sim \text{Beta}(\alpha, \beta)$ hyperparameters

$\theta|x_1, \ldots, x_n \sim \text{Beta}(\alpha + n, \beta + \Sigma x_i)$

$$\hat{\theta}_{MMSE} = \frac{\alpha+n}{\alpha+n+\beta+\sum x_i}$$

$$\hat{\theta}_{MAE} = qbeta(0.5, \alpha+n, \beta+\sum x_i)$$

$$\hat{\theta}_{MAP} = \frac{\alpha+n-1}{\alpha+n+\beta+\sum x_i-2}$$ where appropriate

CR: qbeta

Hypothesis: p-values

α: pseudo # of trials

β: in the α pseudotrials, this is the sum total of failures

Haldane $\theta \sim Beta(0,0)$ $\alpha=0$
complete ignorance $\beta=0$

Laplace $\theta \sim U(0,1) = Beta(1,1) \Rightarrow \alpha=1$
$\beta=1$

Jeffry's $\theta \sim$

$$\mathcal{L}(\theta;x) = (1-\theta)^{\sum x_i}\theta^n$$

$$\ell(\theta;x) = \sum x_i \ln(1-\theta) + n\ln(\theta)$$

$$\ell'(\theta;x) = -\frac{\sum x_i}{1-\theta} + \frac{n}{\theta}$$

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

$$I(\theta) = E(-\ell''(\theta;x))$$

$$= E\left(\frac{\sum x_i}{(1-\theta)^2} + \frac{n}{\theta^2}\right) = \frac{E(\sum X_i)}{(1-\theta)^2} + \frac{n}{\theta^2}$$

$$= \frac{nE(x)}{(1-\theta)^2} + \frac{n}{\theta^2} = n\left(\frac{\frac{1}{\theta}-1}{(1-\theta)^2} + \frac{1}{\theta^2}\right)$$

$$= n\left(\frac{\frac{1-\theta}{\theta}}{(1-\theta)^2} + \frac{1}{\theta^2}\right) = n\left(\frac{1}{\theta(1-\theta)} + \frac{1}{\theta^2}\right)$$

$$= n\left(\frac{1}{\theta^2(1-\theta)}\right) \propto Beta(0, \tfrac{1}{2})$$

$$P(\theta) \propto \sqrt{I(\theta)} = \sqrt{n\frac{1}{\theta^2(1-\theta)}} \propto \theta^{-1}(1-\theta)^{-\frac{1}{2}}$$

Jeffrey's $\theta \sim Beta(0, \tfrac{1}{2}) \Rightarrow \alpha = 0$, $\beta = \tfrac{1}{2}$

(improper)

Similar to Wilson estimate