$$\begin{matrix} p_J(\theta) \\ t \downarrow \uparrow t^{-1} \\ p_J(\phi) \end{matrix}$$

$$\phi = t(\theta) = \frac{\theta}{1-\theta} \longleftrightarrow \theta = t^{-1}(\phi) = \frac{\phi}{1+\phi}$$

$$p_J(\theta) = \text{Beta}\left(\tfrac{1}{2}, \tfrac{1}{2}\right) = \frac{1}{B(\frac{1}{2},\frac{1}{2})} \theta^{-\frac{1}{2}} (1-\theta)^{-\frac{1}{2}}$$

drop abs value

$$p_J(\phi) = p_J(t^{-1}(\phi)) \left| \frac{d}{d\phi}\left[\frac{\phi}{1+\phi}\right] \right| = \frac{1}{B(\frac{1}{2},\frac{1}{2})} \left(\frac{\phi}{1+\phi}\right)^{-\frac{1}{2}} \left(\frac{1}{1+\phi}\right)^{-\frac{1}{2}} \left|\frac{1}{(1+\phi)^2}\right|$$

$$= \underbrace{\frac{1}{B(\frac{1}{2},\frac{1}{2})}}_{= \boxed{\frac{1}{\pi}}} \frac{\phi^{-\frac{1}{2}}}{1+\phi} = \underbrace{\frac{1}{\pi} \phi^{-\frac{1}{2}} (1+\phi)^{-1}}_{\text{want this.}} = F_{1,1} \text{ distribution.}$$

$$p(X|\phi) = \binom{n}{x}\left(\frac{\phi}{1+\phi}\right)^x \underbrace{\left(1 - \frac{\phi}{1+\phi}\right)^{n-x}}_{= \left(\frac{1}{1+\phi}\right)^{n-x}} = \binom{n}{x} \frac{\phi^x}{(1+\phi)^n} \Bigg\} = \mathcal{L}(\phi; x)$$

PMF of Binomial by odds.

JMF always same as likelihood.

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

$$\ell(\phi; x) = \ln\left(\binom{n}{x}\right) + x\ln(\phi) - n\ln(1+\phi) \quad \text{take } \frac{d}{d\phi}$$

$$\ell'(\phi; x) = \frac{x}{\phi} - \frac{n}{1+\phi}, \quad \ell''(\phi; x) = -\frac{x}{\phi^2} + \frac{n}{(1+\phi)^2} \quad \text{hence}$$

for Bin $E[X] = n\theta$

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

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

$$p_J(\phi) \propto \sqrt{\frac{n}{\phi(1+\phi)^2}} \propto \frac{1}{\sqrt{\phi}} \frac{1}{1+\phi} = \phi^{-\frac{1}{2}}(1+\phi)^{-1} \propto F_{1,1} = \frac{1}{\pi}\phi^{-\frac{1}{2}}(1+\phi)^{-1}$$

This verifies that Jeffrey's procedure works for Binomial model and odds reparameterization. Lets prove for all....

Given $p(X|\theta)$ and t $\Rightarrow$ $P(X|\phi)$, assume $P_J(\theta) \propto \sqrt{I(\theta)}$, prove $P(\phi) \propto \sqrt{I(\phi)}$

$\boxed{|x| = \sqrt{x^2}}$

Proof: $P_\phi(\phi) = P_\theta(\theta)\left|\frac{d}{d\phi}[\theta]\right| \propto \sqrt{I(\theta)}\left|\frac{d\theta}{d\phi}\right| = \sqrt{I(\theta)\frac{d\theta}{d\phi}\frac{d\theta}{d\phi}} =$

Fact: ⊛ $I(\theta) := VAR_X[\ell'(\theta;X)] = \ldots\ldots = E_X[-\ell''(\theta;x)] = \ldots = E_X[\ell'(\theta;x)^2]$

⊛ use this here

(Proof continued)

$$\sqrt{E_X[\ell'(\theta;x)^2]\frac{d\theta}{d\phi}\frac{d\theta}{d\phi}} = \sqrt{E_X\left[\frac{\partial \ell}{\partial\theta}\frac{\partial \ell}{\partial\theta}\right]\frac{\partial\theta}{\partial\phi}\frac{\partial\theta}{\partial\phi}} = \sqrt{E_X\left[\frac{\partial \ell}{\partial\theta}\frac{\partial \ell}{\partial\theta}\frac{\partial\theta}{\partial\phi}\frac{\partial\theta}{\partial\phi}\right]}$$

from FACT ABOVE again ⊛

$$= \sqrt{E_X\left[\frac{\partial \ell}{\partial\phi}\frac{\partial \ell}{\partial\phi}\right]} = \sqrt{E_X[\ell'(\phi;x)^2]} \overset{\downarrow}{=} \sqrt{I(\phi)}$$

We have 3 principled non-informative (AKA objective): (do not influence your result too much)

(a) Laplace / uniform
(b) Haldane
(c) Jeffreys

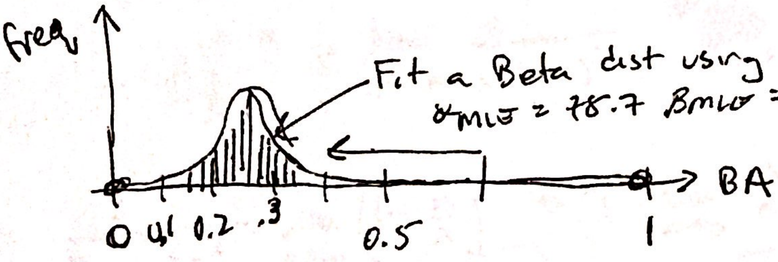
} All have low shrinkage.

Now informative priors i.e. subjective priors. Going to have a lot of weight. Imagine you are trying to infer a new ball players batting ability θ, the probability he gets a hit during an at bat. The batting ability is usually inferred by batting average. $\boxed{BA := \frac{x}{n}\ ;\ x = \text{hits},\ n = \text{relevant at bats} = \hat{\theta}_{MLE} = \bar{x}}$

The problem is the MLE a poor estimate if n is low. Ex: $n=3$, $x=2$ $\Rightarrow BA = \frac{2}{3} = .667$. But this batting ability is impossible. In fact the highest BA ever recorded is .366 Ty Cobb. Will Bayes estimates with uninformative priors help you here?

Take laplace prior $\Rightarrow \hat{\theta}_{MMSE} = 3/5 = .600$ also absurd. Not helpful.

We can solve this by using an uninformative prior that ~~implements~~ provides an "empirical Bayes" estimate. (uses historical data). Here's how...
Look at previous data. Lets subset on all players with at least 500 at bats (arbitrary cut off). If you plot batting averages you get something like this:



Fit a Beta dist using MLEs and find that $\alpha_{MLE} = 78.7$ $\beta_{MLE} = 224.8$ Then we use this as our prior

avg BA = .260

hence $P(\theta) = \text{Beta}(78.7, 224.8) \Rightarrow E[\theta] = .260$, $n_0 = 303.5$, ~~[illegible]~~

ghost ~~[illegible]~~ AT BATS

shrink towards this hard

Lets use this prior to estimate θ for our new batter!

$$\hat{\theta}_{MMSE} = (1-\rho)\hat{\theta}_{MLE} + \rho E[\theta] = \frac{3}{303.5+3} \cdot (.667) + \frac{303.5}{(303.5)+3} \cdot (.260)$$

$$= 1\% \, (.667) + 99\% \, (.260) = .263$$

The use case for informative priors is when you believe the new r.v. behaves like historical r.v.'s behaved. Then you use old data to fit an empirical Baye's prior which will be informative, high shrinkage. Then you use this to do your inference.

Done with Beta-Binomial. Next model Poisson.

$X : \text{Bin}(n, \theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$, now imagine $n \to \infty$ and $\theta \to 0$. but $n\theta = \lambda > 0$, $\theta = \frac{\lambda}{n}$ but not too big. What is ~~the PMF of the number~~ an approximate PMF for this Binomial.

$$\lim_{n\to\infty} \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} = \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \frac{\lambda^x}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{n!}{(n-x)!\,n^x} \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n \left(1-\lim_{n\to\infty}\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{\overbrace{n\cdot(n-1)\cdot\ldots\cdot(n-x+1)}^{x \text{ terms}}}{\underbrace{n\cdot n\cdot\ldots\cdot n}_{x \text{ terms}}} \cdot e^{-\lambda} \cdot (1)$$

goes to 1

$$= \frac{\lambda^x e^{-\lambda}}{x!} = \boxed{\text{Poisson}(\lambda)}$$

Poisson is an approximation of a Binomial if n is large and θ is small.