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

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

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

$$= \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}\right)} \frac{\phi^{-\frac{1}{2}}}{1+\phi} = \underbrace{\frac{1}{\pi} \phi^{-\frac{1}{2}} (1+\phi)^{-1}}_{} = F_{1,1} \text{ distr.}$$

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

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

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

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

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

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

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

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

This verifies that Jeffrey's procedure works for the Binomial model and the odds reparameterization. Let's now prove it for all models and all reparameterizations:

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

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

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

$$I(\theta) := \text{Var}_X[\ell'(\theta; x)] = \ldots = E_X[-\ell''(\theta; x)] = \ldots = E_X[\ell'(\theta; x)^2]$$

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

We have three principled non-informative priors AKA "objective"
(a) Laplace / uniform
(b) Haldane
(c) Jeffreys

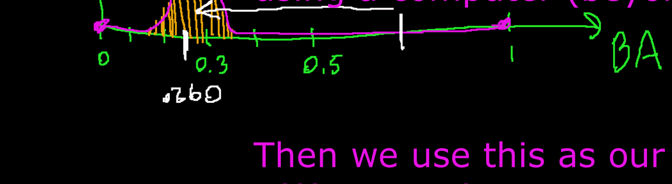
Informative priors i.e. subjective priors! Imagine you are trying to infer a *new* baseball player's batting ability θ, the probability he gets a hit during an at bat. The batting ability is usually inferred by the "batting average", BA := x / n = thetahatmle where x is the # of hits and n is the # of relevant at bats.

The problem is the MLE is a poor estimate if n is low. For example n = 3, x = 2 => BA = 2/3 = .667. But this batting ability is impossible. In fact the highest BA ever recorded in baseball history is .366 by Ty Cobb.

Will Bayes estimates with uninformative priors help you here? NO Consider Laplace uniform prior => thethatmmse = 3/5 = 0.600 which is also absurd.

We can solve this by using an uninformative prior that provides an "empirical Bayes" estimate i.e. uses historical data. Here's how...

Look at previous data. Let's subset on all players that have at least 500 at bats (arbitrary cutoff I know, but we have to start somewhere). If you plot the BA's, you get something like this:



Fit a beta distribution using MLEs and find that alphaMLE = 78.7 and betaMLE = 224.8 using a computer (beyond scope of class).

shrink hard to this

Then we use this as our prior!
P(θ) = Beta(78.7, 224.8) => E[θ] = .260,
n_0 = 303.5

Let's use this prior to estimate θ for our new batter:

$$\hat{\theta}_{MMSE} = (1-\rho)\hat{\theta}_{MLE} + \rho E[\theta]$$

$$= \frac{3}{303.5+3} \cdot .667 + \frac{303.5}{303.5+3} \cdot .260$$

$$= 1\% \cdot .667 + 99\% \cdot .260 = .263$$

The use case for informative priors is when you believe the new rv behaves like historical rv's behaved. Then you can use old data to fit an empirical Bayes prior which will be informative with high shrinkage. Then use that to do your inference.

$$\mathcal{F}: \text{Bin}(n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \qquad \theta = \frac{\lambda}{n}$$

Imagine n --> infinity and θ --> 0 but nθ = λ > 0 but not too big. What is an approximate PMF for this binomial?

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

$$= \frac{\lambda^x}{x!} \lim \frac{\overset{n(n-1)\cdot\ldots\cdot(n-x+1)}{\cancel{n!}}}{\cancel{(n-x)!}\, n^x} \lim \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \lim \frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{x!} \lim \frac{\overbrace{n \cdot (n-1) \cdot \ldots \cdot (n-x+1)}^{x \text{ terms}}}{\underbrace{n \cdot n \cdot \ldots \cdot n}_{x \text{ terms}}} \left(e^{-\lambda}\right)(1) = \frac{\lambda^x e^{-\lambda}}{x!} = \text{Poisson}(\lambda)$$

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