Let $X|\theta \sim \text{Binom}(n,\theta)$, $\theta \sim \text{Beta}(\alpha,\beta)$ and $\theta|X \sim \text{Beta}(\overbrace{\alpha+x}^{\alpha'}, \overbrace{\beta+n-x}^{\beta'})$. Then

$$X^*|X \sim \text{BetaBinom}(n^*,\alpha',\beta') = \binom{n^*}{x^*}\frac{B(\overbrace{\alpha+x}^{\alpha'}+x^*, \overbrace{\beta+n-x}^{\beta'}+n^*-x^*)}{B(\underbrace{\alpha+x}_{\alpha'}, \underbrace{\beta+n-x}_{\beta'})}$$

Posterior Predictive Distribution: $\mathbb{P}(X^* \mid X) = \int_\Theta \mathbb{P}(X^* \mid \theta)\,\mathbb{P}(\theta \mid X)\,d\theta$ (the distribution of function X^* given data x)
$\mathbb{P}(X)$ is the distribution of data observed $= \int_\Theta \mathbb{P}(X \mid \theta)\,\mathbb{P}(\theta)\,d\theta$
Prior Predictive Distribution: $\mathbb{P}(X \mid \{\}) = \int \mathbb{P}(X \mid \theta)\,\mathbb{P}(\theta \mid \{\})\,d\theta$

Let $X \sim \text{BetaBinom}(n,\alpha,\beta)$. If $\theta \sim U(0,1) = \text{Beta}(1,1)$, this is an uninformative prior, as well as a indifference or Laplace prior. It says there is one success and one failure. The most uninformative prior is $\theta \sim \text{Beta}(0,0)$. However, this is "illegal" because α and β are not in the parameter space and thus do not form a true PDF. This prior is called an improper prior, as well as Haldane prior.

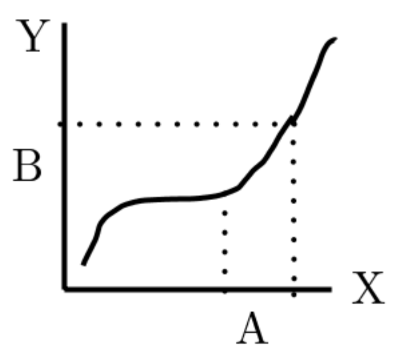
Let's say we go along with $\theta \sim \text{Beta}(0,0)$. Then $\theta|X \sim \text{Beta}(x, n-x)$. From this,

$$\hat{\theta}_{\text{MMSE}} = \frac{x}{n} = \hat{\theta}_{\text{MLE}}$$

This posterior could be improper if $x = 0$ (no successes) or if $x = n$ (no failures). Therefore, be careful when using "improper" priors as your posterior could also be improper.
Note: Beta(0, 0) and Beta(1, 1) are both uninformative but only Beta(1, 1) is indifferent.

Reparameterization: $R = \text{Odds}(\theta) = \frac{\theta}{1-\theta}$. For example, $R = \text{Odds}(0.9) = \frac{0.9}{1-0.9} = 9$. Note that $\theta = (0,1)$ and $R = (0,\infty)$.

Let X and Y be two random variables related by a 1-1 inverse transform. This means $Y = t(X)$ and $X = t^{-1}(Y)$. We know $f_X(x)$, the PDF of X. We want the PDF of Y, $f_Y(y)$.



Since $\mathbb{P}(X \in A) \approx f_X(x)A$ and $\mathbb{P}(Y \in B) \approx f_Y(y)B$

$$f_X(x)|dx| = f_Y(y)|dy| \to f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

By the above equations, we can substitute for X:

$$f_Y(y) = f_X(t^{-1}(y))\left|\frac{d}{dy}[t'(x)]\right|$$

Since $R = t(\theta) = \frac{\theta}{1-\theta}$, then $\theta = t^{-1}(R) = \frac{R}{R+1}$ Therefore

$$f_R(r) = f_\theta(t^{-1}(r))\left|\frac{d}{dr}[t^{-1}(r)]\right| = f_Y(\frac{r}{r+1})\left|\frac{d}{dr}p\frac{r}{r+1}\right| = (1)\left|-\frac{1}{(r+1)^2}\right| = \frac{1}{(r+1)^2}$$

Let $\theta \sim U(0,1)$ or $\theta \sim \text{Beta}(0,0)$ (uninformative). If under a reparameterization $\phi = t(\theta)$, what if I had a protocol which allows us to pick a priors given $\mathcal{F}$:

$$\mathbb{P}\left(X \mid \theta\right) \overset{\text{pick}}{\to} \mathbb{P}\left(\theta\right) \text{ and } \mathbb{P}\left(X \mid \phi\right) \overset{\text{pick}}{\to} \mathbb{P}\left(\phi\right)$$

such that we have $P(\phi) = p(t^{-1}(\phi))\left|\frac{d}{dt}t^{-1}(\phi)\right|$ (Jeffrey's prior).

$\mathbb{P}\left(\theta \mid X\right) = \frac{\mathbb{P}(X \mid \theta)\mathbb{P}(\theta)}{\mathbb{P}(X)} \propto \mathbb{P}\left(X \mid \theta\right)\mathbb{P}\left(\theta\right)$
in fact, $f(x;\theta) \propto g(x;\theta)$ where g is a kernel of f. This means $f(x;\theta) = \frac{1}{c}g(x;\theta)$.

$$\int f(x)\,dx = 1 \to \int g(x)\,dx = \int cf(x)\,dx = c\underbrace{\int f(x)\,dx}_{1} \to c = \int g(x)\,dx$$

Note: f and g are 1-1.

Let $X|\theta \sim \text{Binom}(n,\theta)$ and $\theta \sim \text{Beta}(\alpha,\beta)$.

$$\begin{aligned}
\mathbb{P}\left(\theta \mid X\right) \propto \mathbb{P}\left(X \mid \theta\right)\mathbb{P}\left(\theta\right) &= \binom{n}{x}\theta^x(1-\theta)^{n-x}\frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\
&\propto \theta^x(1-\theta)^{n-x}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\
&= \theta^{\overbrace{x+\alpha-1}^{a}}(1-\theta)^{\overbrace{n-x+\beta-1}^{b}} \\
&= \text{Beta}(x+\alpha, n-x+\beta)
\end{aligned}$$

$$\theta \sim \text{Beta}(\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \propto \underbrace{\theta^a(1-\theta)^b}_{\text{kernel of the beta}}$$

$X|\theta \sim \text{Binom}(n,\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x} = (\frac{n!}{x!(n-x)!})\theta^x(1-\theta)^n(1-\theta)^{-x} \propto \frac{1}{x!(n-x)!}(\frac{\theta}{1-\theta})^x$

Likelihood: $\mathcal{L}(\theta;x) = \mathbb{P}\left(x;\theta\right)$
Log-Likelihood: $l(\theta;x) = \ln(\mathcal{L}(\theta;x))$
Score Function: $s(\theta;x) = l'(\theta;x)$
Fisher Information: $I(\theta) = \text{Var}_x[s(\theta;x)] = \cdots = \text{E}_x[s(\theta;x)^2] = \cdots = \text{E}_x[-l''(\theta;x)]$

The Fisher Information measures the information in X about θ.

Let $X \sim \text{Binom}(n;\theta)$ Then

$$\begin{aligned}
X \sim \text{Binom}(n;\theta) &= \binom{n}{x}\theta^x(1-\theta)^{n-x} \\
l(\theta;x) &= ln\frac{n}{x} + x\ln\theta + (n-x)\ln(1-\theta) \\
l'(\theta;x) &= \frac{x}{\theta} - \frac{n-x}{1-\theta} \\
l''(\theta;x) &= \frac{-x}{\theta^2} - \frac{n-x}{(1-\theta)^2} \\
I(\theta) &= \text{E}_x[-l''(\theta;x)] \\
&= \text{E}[\frac{x}{\theta^2} + \frac{n-x}{(1-\theta)^2}] \\
&= \frac{\text{E}[X]}{\theta^2} + \frac{n-\text{E}[X]}{(1-\theta)^2} \\
&= \frac{n\theta}{\theta^2} + \frac{n-n\theta}{(1-\theta)^2} \\
&= n(\frac{1}{\theta} + \frac{1}{1-\theta}) \\
&= n\frac{1}{\theta(1-\theta)}
\end{aligned}$$

The Fisher information for the Binomial distribution is $n\frac{1}{\theta(1-\theta)}$.