$n=100,\ x=43,\ P(\theta)=\text{Beta}(1,1) \Rightarrow P(\theta|x)=\text{Beta}(44,58).\quad \alpha_0=5\%,\ \delta=1\%.$

$H_0: \theta \in [.49, .51]$

$$p_{val} = P(H_0|x) = \int_{.49}^{.51} d\theta = \text{pbeta}(.51,44,58) - \text{pbeta}(.49,44,58) = 0.06 \Rightarrow \text{Retain } H_0.$$

$$\| \qquad F_{\theta|x}(\theta_0+\delta) - F_{\theta|x}(\theta_0-\delta) \qquad \cancel{\alpha_0} = 5\%$$

There is insufficient evidence to prove this coin is unfair.

Last topic before the midterm

$\mathcal{F}: \text{Bin}(n,\theta)$ with n fixed, $P(\theta)=\text{Beta}(\alpha,\beta) \Rightarrow P(\theta|x) = \text{Beta}(\alpha+x, \beta+n-x)$

Laplace $P(\theta) = \text{Beta}(\underset{\alpha}{1}, \underset{\beta}{1}) \Rightarrow n_0 = 2$ pseudotrials, $x_0 = 1$ pseudosuccesses

Laplace's uniform prior is "flat" in an effort to be "objective" i.e. let the data speak for itself and not to be "subjective" i.e. allow your personal biases to be part of your inferential conclusion.

Can we be more objective? Can we create a prior that has no part in the inferential conclusion? This would mean n_0 = 0. How about α = β = 0?

$$P(\theta) = \text{Beta}(0,0) = \frac{1}{B(0,0)} \theta^{-1}(1-\theta)^{-1} \Big\}\ \text{not a PDF}$$

There's a problem with this. The parameter space for the beta is α > 0 and β > 0. If α = β = 0, this is not a PDF since its integral over the support diverges. This makes it an "improper prior" since it is not a true random variable.

But do we care? Churning through the math, we get the posterior:

$$P(\theta|x) = \text{Beta}(x, n-x)$$

This posterior is proper as long as x < n and x > 0 which means you need to have at least one success and at least one failure in your data. If it's proper, you have full bayesian inference: point estimates, CR's, p-vals... However, you always have thetahatmmse:

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

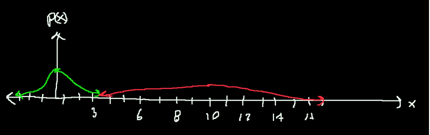
Also, ρ = 0 (no shrinkage). I believe this prior was first introduced
by Haldane in 1932 so we'll call it the "Haldane prior".

Midterm I ↑

Midterm I ↓

Back to probability class... We will introduce mixture / compound
distributions e.g.

$$X \sim \begin{cases} \underbrace{N(0,1^2)}_{\text{model}} & \underbrace{\text{wp}}_{\text{mixture}}\ \frac{1}{2} \\ N(10,2^2) & \text{wp}\ \frac{1}{2} \end{cases}$$



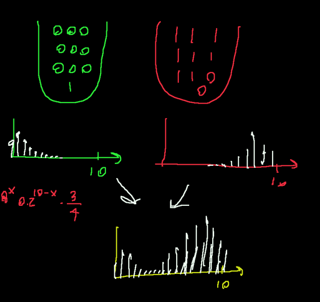
$$P(X) = \int_{\Theta} P(X, \vec{\theta})\, d\vec{\theta} \text{ if } \vec{\theta} \text{ is continuous}$$

$$= \sum_{\vec{\theta} \in \Theta} P(X, \vec{\theta}) \text{ if } \vec{\theta} \text{ is discrete}$$

$$= \sum_{\vec{\theta} \in \Theta} P(X|\vec{\theta}) P(\vec{\theta})$$

$$= \frac{1}{\sqrt{2\pi 1^2}} e^{-\frac{1}{2 \cdot 1^2}(x-0)^2}(0.5) + \frac{1}{\sqrt{2\pi 2^2}} e^{-\frac{1}{2 \cdot 2^2}(x-10)^2}(0.5)$$

$$X \sim \begin{cases} \text{Bin}(10, 0.1) & \text{wp}\ \frac{1}{4} \\ \text{Bin}(10, 0.8) & \text{wp}\ \frac{3}{4} \end{cases}$$



$$P(X) = \binom{10}{x} 0.1^x\, 0.9^{10-x} \cdot \frac{1}{4} + \binom{10}{x} 0.8^x\, 0.2^{10-x} \cdot \frac{3}{4}$$

Have we seen P(X) before that's the result of a margining making P(X) a mixture/compound distribution? Yes...

$$P(\theta|x) = \frac{P(X|\theta)P(\theta)}{P(X)} = \text{Beta}(\alpha+x, \beta+n-x)$$

$$P(X) = \int_{\Theta} P(X|\theta) P(\theta)\, d\theta$$

$\mathcal{F}: \text{Bin}(n,\theta)$ n fixed, $P(\theta) = \text{Beta}(\alpha,\beta)$

$$P(X) = \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} d\theta = \frac{\binom{n}{x}}{B(\alpha,\beta)} \underbrace{\int_0^1 \theta^{\alpha+x-1}(1-\theta)^{\beta+n-x-1} d\theta}_{\text{beta function}}$$

$$= \frac{\binom{n}{x}}{B(\alpha,\beta)} B(\alpha+x, \beta+n-x) = \text{BetaBinomial}(n, \alpha, \beta)$$

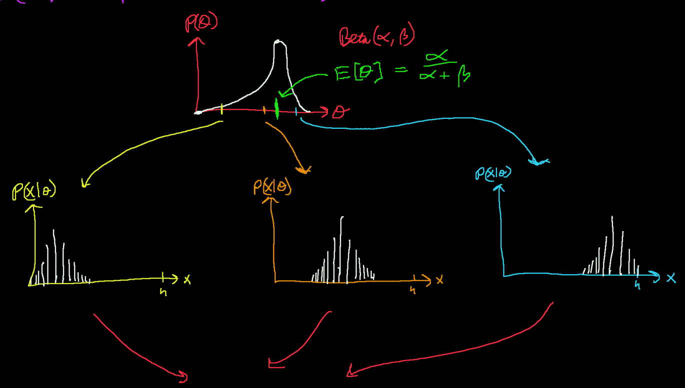
$Y \sim \text{BetaBinomial}(n,\alpha,\beta)$, $\text{Supp}[Y] = \{0, 1, \ldots, n\}$, $n \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$

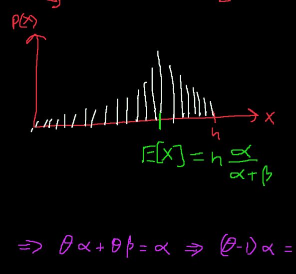
$$E[Y] = \ldots = n\frac{\alpha}{\alpha+\beta}, \quad \text{Var}[Y] = \ldots = n\frac{\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

since the beta function is not available in closed form, the PMF/CDF are not available in closed form. To compute, you need a computer. Here's the notation we'll use in this class (the R notation):

$$P(Y=y) = \text{dbetabinom}(y, n, \alpha, \beta)$$

$$P(Y \le y) = \text{pbetabinom}(y, n, \alpha, \beta)$$





$$E[X] = n\frac{\alpha}{\alpha+\beta}$$

which is why the betabinomial rv is sometimes called the "overdispersed binomial"

$$\text{let } \theta := \frac{\alpha}{\alpha+\beta} \Rightarrow \theta\alpha + \theta\beta = \alpha \Rightarrow (\theta-1)\alpha = -\theta\beta \Rightarrow \beta = \alpha\frac{1-\theta}{\theta}$$

$\Downarrow$

$E[X] = n\theta$ an intuitive formula for the betabinomial expectation since it is the same as binomial expectation.

let $\alpha \to \infty$ but keep $\theta = \frac{\alpha}{\alpha+\beta}$ constant $\Rightarrow 1-\theta = \frac{\beta}{\alpha+\beta}$

$$\lim_{\alpha\to\infty} \text{Var}[X] = \lim_{\alpha\to\infty} n \frac{\cancel{\alpha}\left(\alpha\frac{1-\theta}{\theta}\right)\left(\alpha + \alpha\frac{1-\theta}{\theta} + n\right)}{\left(\cancel{\alpha} + \cancel{\alpha}\frac{1-\theta}{\theta}\right)^2\left(\alpha + \alpha\frac{1-\theta}{\theta} + 1\right)} = n \lim_{\alpha\to\infty} \frac{\frac{1-\theta}{\theta}}{\left(1 + \frac{1-\theta}{\theta}\right)^2} \cancel{\lim_{\alpha\to\infty} \frac{\alpha + \alpha\frac{1-\theta}{\theta} + n}{\alpha + \alpha\frac{1-\theta}{\theta} + 1}}$$

$$= n \lim_{\alpha\to\infty} \frac{\theta(1-\theta)}{(\cancel{\theta} + 1 - \cancel{\theta})^2} = n\theta(1-\theta)$$

which is the same variance as the rv Bin(n, θ)

$$\text{Var}[X] = n\frac{\alpha\beta}{(\alpha+\beta)^2} \frac{\alpha+\beta+n}{\alpha+\beta+1} = \underbrace{n\theta(1-\theta)}_{\text{Var Binomial}} \underbrace{\frac{\alpha+\beta+n}{\alpha+\beta+1}}_{\text{overdispersion}}$$

$$\text{overdispersion} \in (1, n)$$

Thus, the betabinomial is a much more flexible model.