minimum mean absolute error i.e. $\hat{\theta}_{MMAE} = \underset{\hat{\theta} \in \Theta}{\text{argmin}} \left\{ E\left[|\theta - \hat{\theta}| \mid x \right] \right\}$

$$\hat{\theta}_{MMAE} := Med[\theta | x] = a \quad s.t. \int_{-\infty}^{a} P(\theta|x)\, d\theta = \frac{1}{2}$$

Using our model: iid bern(θ) and data x = <0,1,1>, we can compute the MMAE Bayesian point estimate:

$$\int_0^a 12\theta^2(1-\theta)\, d\theta = 12\left[\frac{\theta^3}{3} - \frac{\theta^4}{4}\right]_0^a = 12\left(\frac{a^3}{3} - \frac{a^4}{4}\right) \overset{set}{=} \frac{1}{2}$$

$$\Rightarrow -\frac{1}{4}a^4 + \frac{1}{3}a^3 + 0a^2 + 0a - \frac{1}{24} = 0 \Rightarrow a \approx 0.614$$

This is a "quartic equation" and has a formulaic solution. You can look it up. The answer is

These are the three bayesian point estimates we will use for the rest of the class i.e. $\hat{\theta}_{MMSE}, \hat{\theta}_{MMAE}, \hat{\theta}_{MAP}$

The data x = <0,1,1> was a specific case. We will now solve this generally for any dataset x = <x_1, ..., x_n>. Also using Laplace's prior of indifference, θ ~ U(0, 1).

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)} = \frac{P(X|\theta)\cancel{P(\theta)}^{1}}{\int_0^1 P(X|\theta)\cancel{P(\theta)}^{1}\, d\theta} = \frac{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i}}{\int_0^1 \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}\, d\theta}$$

This integral in the denominator is a special integral and is known as the "beta function": The beta function has no closed form solution but can be calculated to arbitrary precision using a scientific calculator.

$$B(\alpha,\beta) := \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

$$= \frac{1}{B(\sum x_i + 1, n - \sum x_i + 1)}\, \theta^{\sum x_i + 1 - 1}(1-\theta)^{n - \sum x_i + 1 - 1} = Beta\left(\sum x_i + 1, n - \sum x_i + 1\right)$$

We just derived that the posterior for the iid bernoulli likelihood is a beta distribution. Let's go back to probability class and examine the beta distribution...

$$Y \sim Beta(\alpha,\beta) \overset{PDF}{:=} \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1} = p(y)$$

$$Supp[Y] = (0,1). \quad \int_0^1 \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1}\, dy = \frac{1}{\cancel{B(\alpha,\beta)}} \underbrace{\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\, dy}_{\cancel{B(\alpha,\beta)}} = 1 \checkmark$$

$\alpha \in ?, \beta \in ?$ $\quad \alpha > 0, \beta > 0.$

$$\alpha = 0, \beta = 1 \Rightarrow \int_0^1 \frac{1}{y}\, dy = \infty$$

$$E[Y] = \int_0^1 y\, p(y)\, dy = \int_0^1 y \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1}\, dy = \frac{1}{B(\alpha,\beta)} \int_0^1 y^{\alpha+1-1}(1-y)^{\beta-1}\, dy = \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)}$$

To simplify this, we need the gamma function:

$$\Gamma(\alpha) := \int_0^\infty t^{\alpha-1} e^{-t}\, dt, \quad \alpha > 0$$

Facts: ① $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, ② $B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

$$= \frac{\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}}{B(\alpha,\beta)} = \frac{\frac{\alpha\,\cancel{\Gamma(\alpha)\Gamma(\beta)}}{(\alpha+\beta)\cancel{\Gamma(\alpha+\beta)}}}{\cancel{B(\alpha,\beta)}} = \frac{\alpha}{\alpha+\beta}$$

$Var[Y] =$ on HW

$$Mode[Y] = \underset{y \in (0,1)}{\text{argmax}} \left\{ \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1} \right\} = \text{argmax}\left\{ (\alpha-1)\ln(y) + (\beta-1)\ln(1-y) \right\}$$

deriv.

$$\Rightarrow \frac{\alpha-1}{y} - \frac{\beta-1}{1-y} \overset{set}{=} 0 \Rightarrow y_* = \frac{\alpha-1}{\alpha+\beta-2}$$

If we take the second derivative to check if it's negative, we find it's only negative if both alpha and beta are greater than or = 1.

$Med[Y]$ has no closed form expression and thus must be done with a computer. We will denote the answer to this using notation from the R programming language: qbeta(0.5, alpha, beta).

Let's take a look at the shapes of the beta distribution: