MATH 341 - 2/22 lec 06

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

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

Quantile which gives you the 50% mark.

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

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

← set equal to $\frac{1}{2}$ to find median point.

$-\infty$ → 0 since parameter space was only (0,1)

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

this is a quartic equation, has a formula for soln. The answer is $a \approx .614$.

~~These~~ These are the 3 bayesian point estimates:

$\hat{\theta}_{MAP}, \hat{\theta}_{MMSE}, \hat{\theta}_{MMAE}$ ← They are functions of the posterior dist.

The data $x = (0,1,1)$ was a specific case. We will now solve generally for any data set $x = (x_1, x_2, \dots, x_n)$. Also using laplace prior $\boxed{\theta \sim U(0,1)}$.

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)} = \frac{P(x|\theta)P(\theta)}{\int_0^1 p(x|\theta)p(\theta)\,d\theta}$$

(says this is 1 for both)

prior predictive distribution margin out θ

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

$$\int_0^1 \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}\,d\theta$$ ← this is a special integral, known as the "beta function".

$$B(\alpha,\beta) := \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,dt.$$ The beta function has no closed form sol, but can be computed with sci-calc.

Continue with calculation:

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

→ this is known as the Beta PDF.

$$\text{Beta}\left(\sum x_i+1,\, n-\sum x_i+1\right)$$

we just derived that the posterior for the iid Bern likelihood is a beta distribution. Let's examine beta distribution

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

$\text{Supp}[Y] = (0,1)$ Verify this is a pdf. does this integrate to 1?

$$\int_0^1 \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1}\,dy = \frac{1}{B(\alpha,\beta)}\int_0^1 \overbrace{y^{\alpha-1}(1-y)^{\beta-1}\,dy}^{B(\alpha,\beta)\text{ function}} = \int_0^1 \frac{B(\alpha,\beta)}{B(\alpha,\beta)} = 1$$

$\alpha \in ?\ \beta \in ?$ $\alpha > 0,\ \beta > 0$ thats it.

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

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

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 y^{\alpha+1-1} (1-y)^{\beta-1} dy = \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} \quad (*)$$

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 $B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

use this guy. ⟶

In the numerator.

$$(*) = \frac{\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}}{B(\alpha,\beta)} \overset{\text{using ①}}{=} \frac{\frac{\alpha\,\boxed{\Gamma(\alpha)\Gamma(\beta)}}{(\alpha+\beta)\,\boxed{\Gamma(\alpha+\beta)}}}{B(\alpha,\beta)} = \frac{\alpha}{\alpha+\beta} = E[Y]$$

VAR[Y] = on HW.

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

argmax is immune to monotonic transformations. get rid of constant and take natural log.

$= \text{argmax} \{ (\alpha-1)\ln(y) + \beta-1 \ln(1-y)$. TAKE derivative, set to 0.

derivative $= \dfrac{\alpha-1}{y} - \dfrac{\beta-1}{1-y} \overset{\text{set}}{=} 0 \implies \underset{y^*}{\text{mode}} = \dfrac{\alpha-1}{\alpha+\beta-2}$

Take derivative again and see if its negative, we find its only negative if both α and β are greater than or equal to 1.

Med[Y] has no closed form expression, and thus must be done with a computer. We will denote the answer using notation from R prog language: qbeta$(0.5, \alpha, \beta)$. $= \hat{\theta}_{MMAE}$

Lets take a look at some common shapes of Beta distribution.