Question: Should we report posterior mean, median or mode?
which is best?
what does it mean for an estimator to be "best"?

Answer: Depends on loss function.

We can view Bayes estimators as "risk minimizers"

Setup

parameters $\underline{\theta}$
Data $\underline{y}$
Loss function $L(\underline{\theta}, a)$ – determines the cost of deciding that the parameter has value a when it is in fact equal to $\underline{\theta}$.

The Bayes estimator, $\hat{\underline{\theta}}_B$ minimizes posterior expected loss.

Example: $L(\theta, a) = (\theta - a)^2$

Posterior Risk : $R(a \| \underline{y}) = \underbrace{\int (\theta - a)^2 p(\theta | \underline{y})\, d\theta}_{\mathbb{E}[L(\theta, a) | \underline{y}]}$

$$\hat{\theta}_B := \underset{a}{\operatorname{argmin}}\ \mathbb{E}\left[(\theta - a)^2 | \underline{y}\right]$$, expand out:

$$= \underset{a}{\operatorname{argmin}}\ \mathbb{E}[\theta^2 | \underline{y}] - 2\mathbb{E}[\theta \cdot a | \underline{y}] + \mathbb{E}[a^2 | \underline{y}]$$

* expectation taken wrt $\theta | \underline{y}$. =>

$$= \underset{a}{\operatorname{argmin}}\ \underbrace{\mathbb{E}[\theta^2 | y]}_{\text{constant in } a} - 2a\,\mathbb{E}[\theta | \underline{y}] + a^2$$

* differentiate w.r.t. a ; set $= 0$:

$$-2\,\mathbb{E}[\theta | y] = -2a \Rightarrow \boxed{a = \mathbb{E}[\theta | y]}$$

Ex Absolute error loss

$$L(\theta, a) = |\theta - a|$$

$$\text{Risk}(a|y) = \mathbb{E}\left[|\theta - a| \,|\, y\right] = \int_{-\infty}^{\infty} |\theta - a|\, p(\theta|y)\, d\theta$$

* split integral at a:

$$= \int_{-\infty}^{a} (a - \theta)\, p(\theta|y)\, d\theta + \int_{a}^{\infty} (\theta - a)\, p(\theta|y)\, d\theta$$

* Differentiate w.r.t. a:

$$\frac{d}{da} R(a|y) = \int_{-\infty}^{a} p(\theta|y)\, d\theta - \int_{a}^{\infty} p(\theta|y)\, d\theta$$

* set $= 0$:

$$\int_{-\infty}^{a} p(\theta|y)\, d\theta = \int_{a}^{\infty} p(\theta|y)\, d\theta$$

$$\Rightarrow P(\theta \leq a|y) = \frac{1}{2}$$

$$\Rightarrow \boxed{\hat{\theta}_B := \text{posterior median}}$$

Contrast w/ frequentist estimators:

Frequentist risk: $R(\theta, \delta) = \mathbb{E}_Y\left[L(\overset{\text{fixed}}{\theta}, \delta(Y))\right]$

$$= \int L(\theta, \delta(Y)) \cdot p(Y|\theta)\, dY$$