Probabilistic models
-
Bayesian Methods
-
Discussion

Marylou Gabrié

CDS, NYU

March 16, 2021

Bayesian decision for absolute loss is median

posterior $p(\theta|\mathcal{D})$ $\quad \theta \in \mathbb{R} = \Theta$

$\ell(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$

$$r(\hat{\theta}) = \mathbb{E}\left[|\theta - \hat{\theta}| \,|\, \mathcal{D}\right] = \int_{-\infty}^{+\infty} |\theta - \hat{\theta}| \, p(\theta|\mathcal{D}) \, d\theta$$

Bayesian risk ?

$$\mathbb{1}_{\theta > \hat{\theta}}(\theta) = \begin{cases} 0 & \text{if } \theta < \hat{\theta} \\ 1 & \text{if } \theta > \hat{\theta} \end{cases}$$

$\mathbb{1}$

$$\frac{dr}{d\hat{\theta}} = \int_{-\infty}^{+\infty} \nabla_{\hat{\theta}} |\theta - \hat{\theta}| \, p(\theta|\mathcal{D}) \, d\theta = \int_{-\infty}^{+\infty} \nabla_{\hat{\theta}} |\theta - \hat{\theta}| \left(\mathbb{1}_{\theta > \hat{\theta}} + \mathbb{1}_{\theta < \hat{\theta}}\right) p(\theta|\mathcal{D})$$

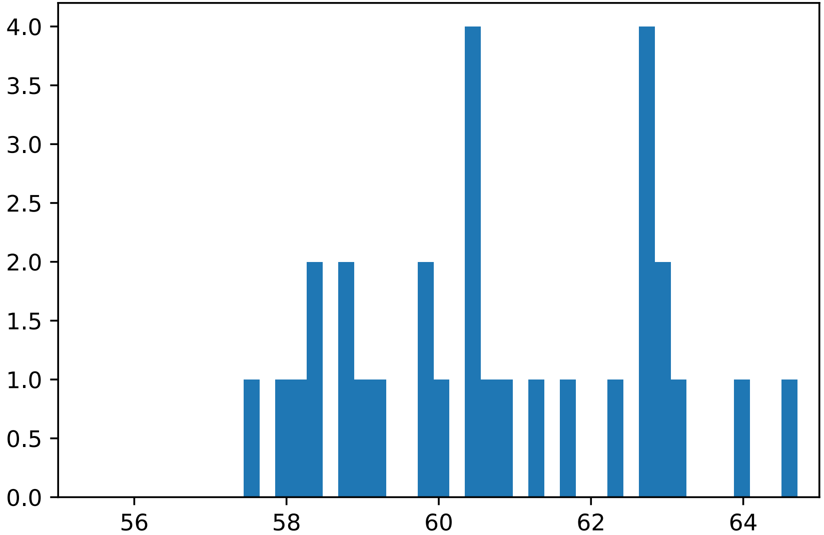
$$= \int_{-\infty}^{+\infty} (-1) \mathbb{1}_{\theta > \hat{\theta}} \, p(\theta|\mathcal{D}) \, d\theta + \int_{-\infty}^{+\infty} (1) \, \mathbb{1}_{\theta > \hat{\theta}} \, p(\theta|\mathcal{D}) \, d\theta = 0$$

$$\Rightarrow \int_{-\infty}^{\hat{\theta}} d\theta \, p(\theta|\mathcal{D}) = \int_{\hat{\theta}}^{+\infty} d\theta \, p(\theta|\mathcal{D}) \quad \Rightarrow \quad \hat{\theta} = \text{median of } p(\theta|\mathcal{D})$$

Hours of sun during January in Vienna

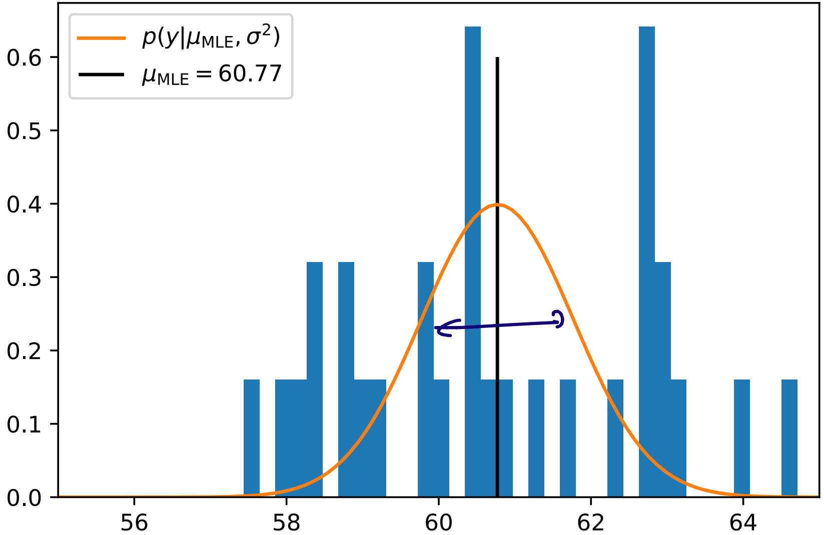
- $N = 30$ measurements y_i
- What could you choose as a parametric family? Gaussian $p(y;\mu,\sigma^2) = \frac{e^{-(y-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$
- What is the MLE?

Assume we know σ^2. μ_{MLE}? $\mu_{MLE} = \frac{1}{N}\sum_{i=1}^{N} y_i$



Hours of sun during January in Vienna

- $N = 30$ measurements y_i
- What could you choose as a parametric family?
- What is the MLE?



Hours of sun during January in Vienna

- $N = 30$ measurements y_i $\quad \mathcal{D} = \{y_i\}_{i=1}^{N}$
- Now meteorologists tells you they have a prior: $p(\mu) = \mathcal{N}(\mu; \mu_0, 1) = \frac{1}{\sqrt{2\pi}} e^{-(\mu-\mu_0)^2/2}$
- What is the posterior?

$$p(\mu \mid \mathcal{D}, \sigma^2) \propto \text{likelihood} \times \text{prior} \propto \prod_{i=1}^{N} e^{-(y_i-\mu)^2/2\sigma^2} \, e^{-(\mu-\mu_0)^2/2}$$

$$\propto e^{-\sum_{i=1}^{N} \frac{(y_i-\mu)^2}{2\sigma^2} - (\mu-\mu_0)^2/2}$$

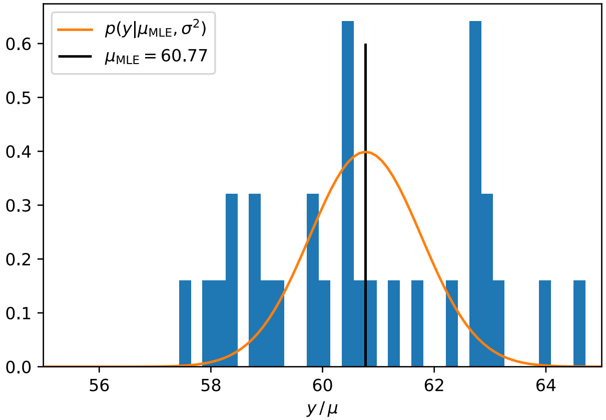
$$\propto e^{-\frac{(\mu-\mu')^2}{2\sigma'^2}}$$

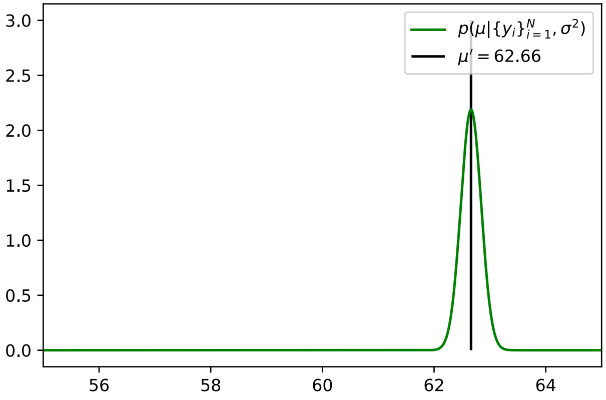
$$\begin{cases} \sigma'^2 = \left(1 - \frac{N}{\sigma^2}\right)^{-1} \\ \mu' = \sigma'^2 \left(\mu_0 + \frac{\sum_{i=1}^{N} y_i / N}{\sigma^2}\right) \end{cases}$$

μ_{MLE}

Hours of sun during January in Vienna

- $N = 30$ measurements y_i
- Now meteorologists tells you they have a prior: $p(\mu) = \mathcal{N}(\mu; \mu_0, 1)$ $\mu_0 = 63$
- What is the posterior?

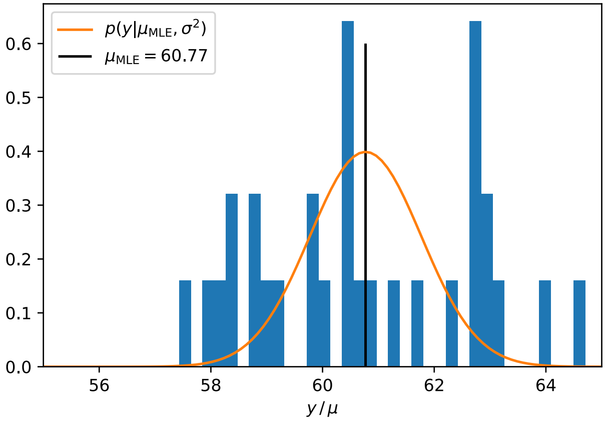


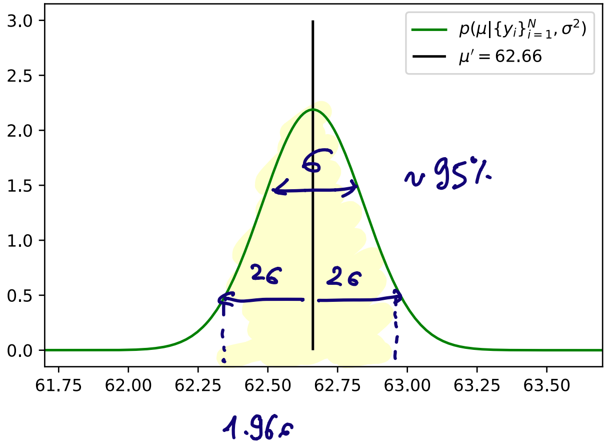


Hours of sun during January in Vienna

- $N = 30$ measurements y_i
- Now meteorologists tells you they have a prior: $p(\mu) = \mathcal{N}(\mu; \mu_0, 1)$
- What is the posterior? What is the credible set?

$p(\mu \in [62.3, 62.9]) \geq 0.95$





Hours of sun during January in Vienna

- $N = 30$ measurements y_i
- Now meteorologists tells you they have a prior: $p(\mu) = \mathcal{N}(\mu; \mu_0, 1)$
- The posterior is $p(\mu|\{y_i\}_{i=1}^N, \sigma^2) = \mathcal{N}(\mu; \mu', \sigma'^2)$.
- What are the point estimates of μ minimizing squared loss, absolute loss and 0-1 loss?

$$\hat{\mu}_{MSE} = \min_{\hat{\mu}} \int d\mu \; p(\mu | \mathcal{D}, \sigma^2) (\mu - \hat{\mu})^2 = \int d\mu \; \mu \; p(\mu | \mathcal{D}, \sigma^2) = \mu'$$

$$\hat{\mu}_{AL} = \mu'$$

$$\hat{\mu}_{MAP} = \arg\max_{\hat{\mu}} p(\hat{\mu} | \mathcal{D}, \sigma^2) = \mu'$$