M1 MAS - Statistics

Exercices Chap 5 Bayesian statistics

(1)

Exercice 1 . Let $\alpha, \beta > 0$. $\Theta \sim \beta eta(\alpha, \beta)$

- Θ takes values in $[0,1]$, with density function such that

$$\pi_{\Theta}(\theta) = \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} \text{ and } B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

- $\mathbb{E}[\Theta] = \frac{\alpha}{\alpha+\beta}$ and $Var(\Theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

- $\beta eta(1,1) \sim \mathcal{U}([0,1])$.
- Consider n realizations of an iid random sample X following Bernoulli:

$X_{1:n} \sim \mathcal{B}(\theta)^{\otimes n}$, we are interested in $\theta \in [0,1]$.

Let $S_n = X_1 + \dots + X_n$.

(1) - Let $\hat{\theta}_n = \frac{1}{n} S_n = \frac{1}{n}\sum_{i=1}^n x_i$ with $\sum_{i=1}^n X_i \sim \mathcal{N}(n\theta, n\theta(1-\theta))$.

Thus $\hat{\theta}_n \sim \mathcal{N}\left(\theta, \frac{\theta(1-\theta)}{n}\right)$. $\Rightarrow \frac{(\hat{\theta}_n - \theta)}{\sqrt{\frac{\theta(1-\theta)}{n}}} \sim \mathcal{N}(0,1)$

$$\mathbb{P}\left(-1{,}96 \leq \sqrt{n}\frac{\hat{\theta}_n-\theta}{\sqrt{\theta(1-\theta)}} \leq 1{,}96\right) = 0{,}95, \text{ or } \mathbb{P}\left(-q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)} \leq \sqrt{n}\frac{\hat{\theta}_n-\theta}{\sqrt{\theta(1-\theta)}} \leq q_{1-\frac{\alpha}{2}}^{\mathcal{N}(0,1)}\right) = 1-\alpha$$

Hence $\mathbb{P}\left(-\frac{\sqrt{\theta(1-\theta)}}{\sqrt{n}} q_{1-\frac{\alpha}{2}} - \hat{\theta}_n \cancel{\leq} -\theta \cancel{\leq} \sqrt{\frac{\theta(1-\theta)}{n}} q_{1-\frac{\alpha}{2}} - \hat{\theta}_n\right) = 1-\alpha$

$$\mathbb{P}\left(\hat{\theta}_n - \sqrt{\frac{\hat{\theta}_n(1-\hat{\theta}_n)}{n}} q_{1-\frac{\alpha}{2}} \leq \theta \leq \hat{\theta}_n + \sqrt{\frac{\hat{\theta}_n(1-\hat{\theta}_n)}{n}} q_{1+\frac{\alpha}{2}}\right) = 1-\alpha$$

Finally $IC_{1-\alpha}(\theta) = \left[\hat{\theta}_n - q_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\theta}_n(1-\hat{\theta}_n)}{n}} ; \hat{\theta}_n + q_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\theta}_n(1-\hat{\theta}_n)}{n}}\right]$.

→ With $x_{obs} = (1, 1, \dots, 1)$, the realized bounds of the interval would be:

$$\hat{\theta}_n^{obs} = \frac{1}{n} S_n = \frac{1}{n} \times n = 1. \Rightarrow IC^{obs} = [1, 1] \Rightarrow \text{ponctual!}$$

→ With $x_{obs} = (0, 0, \dots, 0, 1)$, the realized bounds of the interval would be:

$$\hat{\theta}_n^{obs} = \frac{1}{n} S_n = \frac{1}{n} \Rightarrow IC^{obs} = \left[\frac{1}{n} - q_{1-\frac{\alpha}{2}}\sqrt{\frac{\frac{1}{n}(1-\frac{1}{n})}{n}} ; \frac{1}{n} + \dots \right]$$

soit $IC^{obs} = \left[\frac{1}{n} - q_{1-\frac{\alpha}{2}}\sqrt{\frac{n-1}{n^3}} ; \frac{1}{n} + q_{1-\frac{\alpha}{2}}\sqrt{\frac{n-1}{n^3}}\right] \underset{n\to+\infty}{\sim} \left[\frac{1}{n} - q_{1-\frac{\alpha}{2}}\frac{1}{n} ; \frac{1}{n} + \frac{q_{1-\frac{\alpha}{2}}}{n}\right]$

$$\simeq \left[\frac{1-q_{1-\frac{\alpha}{2}}}{n} ; \frac{1+q_{1-\frac{\alpha}{2}}}{n}\right]. \quad \underline{ex}: \alpha = 5\% \Rightarrow IC \simeq \left[-\frac{1}{n}, \frac{3}{n}\right]$$

but θ cannot be < 0! since $\Theta \subseteq [0, 1]$.

② - Posterior distribution: we look for the distribution $\Theta | X$:

$$f_{\Theta|X=x}(\theta) = \frac{\pi_{(\Theta,X)}(\theta, x)}{f_X(x)} = \frac{f_{X|\Theta=\theta}(x)\, \pi_\Theta(\theta)}{\int_\Theta f_{X|\Theta=\theta}(x)\, \pi_\Theta(\theta)\, d\theta}$$ with X a random vector $(X_1, X_2, \dots, X_n)$

X_i iid
$$= \frac{\prod_{i=1}^n f_{X_i|\Theta=\theta}(x_i)\, \pi_\Theta(\theta)}{\int_\Theta \prod_{i=1}^n f_{X_i|\Theta=\theta}(x_i)\, \pi(\theta)\, d\theta} = \frac{\prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} \times \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1 \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} \times \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta}$$

$$= \frac{\theta^{\sum_{i=1}^n x_i + \alpha - 1}(1-\theta)^{n-\sum_{i=1}^n x_i + \beta - 1}\cancel{\frac{1}{B(\alpha,\beta)}}}{\int_0^1 \theta^{\sum x_i + \alpha - 1}(1-\theta)^{n-\sum x_i + \beta - 1}\cancel{\frac{1}{B(\alpha,\beta)}}\, d\theta} = \dots$$

→ The denominator is independent from θ;

→ We recognize at the numerator the density function of the Beta distribution.

with new parameters $\begin{cases} \alpha' = \alpha + \sum_{i=1}^{n} x_i \\ \beta' = \beta + n - \sum_{i=1}^{n} x_i \end{cases}$

3. - $\mathbb{E}[\Theta | X = x_{obs}] = \frac{\alpha'}{\alpha'+\beta'}$ and $Var(\Theta | X = x_{obs}) = \frac{\alpha'\beta'}{(\alpha'+\beta')^2(\alpha'+\beta'+1)}$

4. - $\mathbb{E}[\Theta|X] = \frac{\alpha + \sum_{i=1}^{n} X_i}{\alpha + \cancel{\sum_{i=1}^{n} X_i} + \beta + n - \cancel{\sum_{i=1}^{n} X_i}} = \frac{\alpha + \sum_{i=1}^{n} X_i}{\alpha + \beta + n}$

Is $\hat{\theta} = \mathbb{E}[\Theta|X]$ a consistent estimator of θ?

- $\mathbb{E}[\hat{\theta}] = \mathbb{E}[\mathbb{E}[\Theta|X]] = \mathbb{E}\left[\frac{\alpha + \sum X_i}{\alpha+\beta+n}\right] = \frac{1}{\alpha+\beta+n}(\alpha + n \times \theta)$

$$= \frac{\alpha + n\theta}{\alpha+\beta+n} = \frac{\alpha}{\alpha+\beta+n} + \frac{n\theta}{\alpha+\beta+n} \xrightarrow[n\to+\infty]{} 0 + \frac{\cancel{n}\theta}{\cancel{n}} = \theta.$$

It is "asymptotically" unbiased.

- $Var(\hat{\theta}) = Var(\mathbb{E}[\Theta|X]) = Var\left(\frac{\alpha + \sum X_i}{\alpha+\beta+n}\right) = \frac{1}{(\alpha+\beta+n)^2} Var\left(\alpha + \sum_{i=1}^{n} X_i\right)$

$$= \frac{1}{(\alpha+\beta+n)^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{n\theta(1-\theta)}{(\alpha+\beta+n)^2} = \frac{n\theta}{(\alpha+\beta+n)^2} - \frac{n\theta^2}{(\alpha+\beta+n)^2} \xrightarrow[n\to+\infty]{} 0$$

⇒ The quadratic risk of $\hat{\theta} = \mathbb{E}[\Theta|X]$ tends to 0 when $n \to +\infty$.

⇒ $\hat{\theta} = \mathbb{E}[\Theta|X] \xrightarrow[n\to+\infty]{MQ} \theta \Rightarrow \hat{\theta} \xrightarrow[n\to+\infty]{\mathbb{P}} \theta$

⇒ $\hat{\theta} = \mathbb{E}[\Theta|X]$ is a consistent estimator of θ.

⑤- Credibility intervals:

We remind that for a given prior $\pi(\theta)$ on Θ, an interval $IC_{1-\alpha}^{x^{obs}}(\theta)$ is a credibility interval with security $(1-\alpha)$ if

$$\underset{\text{distribution of } \Theta}{\mathbb{P}_{\pi}}\left(\theta \in IC_{1-\alpha}^{x_{obs}}(\theta) \mid X = x^{obs}\right) = 1-\alpha.$$

With $\mathbb{P}_{\pi}\left(\theta \in IC_{1-\alpha}^{x_{obs}}(\theta) \mid X = x^{obs}\right) = \int_{\Theta} \mathbb{1}_{\{\theta \in IC_{1-\alpha}^{x_{obs}}\}} \underbrace{\pi(\theta \mid X=x)}_{\sim \text{Beta}(\alpha', \beta')} d\theta$:

We know that $\Theta \mid X=x \sim \text{Beta}(\alpha'(x), \beta'(x))$ with $\begin{cases} \alpha'(x) = \alpha + \sum_{i=1}^{n} x_i^{obs} \\ \beta'(x) = \beta + n - \sum_{i=1}^{n} x_i^{obs} \end{cases}$

Taking $\hat{\theta} = \mathbb{E}[\Theta \mid X]$, we obtain

$$IC_{1-\alpha}^{x_{obs}}(\theta) = \left[q_{\alpha/2}^{\text{Beta}(\alpha',\beta')}, q_{1-\frac{\alpha}{2}}^{\text{Beta}(\alpha',\beta')}\right]$$

$\Rightarrow$ The credibility interval is always an interval, and $\theta \in [0,1]$!

Exercice 2 We consider the bayesian model: $X_{1:n} \mid \Theta = \theta \sim \mathcal{G}(\theta)^{\otimes n}$

$\Theta \sim \text{Beta}(\alpha, \beta)$.

①- The posterior distribution is given by:

$$\pi(\theta \mid x) = \frac{\pi(\theta)\, L(\theta \mid x)}{\mathbb{P}(X=x)}$$

, where $\mathcal{L}(\theta \mid x)$ denotes the likelihood of the sample x considering the statistical model.

$$\propto \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)} \times \prod_{i=1}^{n}(1-\theta)^{x_i-1}\theta$$

Indeed, we have

$$L(\theta \mid x) = f_{(X_1,\dots,X_n)}((x_1,\dots,x_n) \mid \theta)$$
$$\overset{X_i \text{ iid}}{=} \prod_{i=1}^{n} f_{X_i}(x_i \mid \theta)$$
$$= \prod_{i=1}^{n} \mathbb{P}(X_i = x_i \mid \theta)$$
$$= \prod_{i=1}^{n}(1-\theta)^{x_i-1}\theta$$

$$\propto \underbrace{\theta^{\alpha-1+n}(1-\theta)^{\beta-1+\sum_{i=1}^{n} x_i - n}}_{} \times K$$

one recognize the density of a Beta distribution!

Then $\Theta \mid X=x \sim \text{Beta}(\alpha', \beta')$, with

$$\begin{cases} \alpha' = \alpha + n \\ \beta' = \beta + \sum x_i - n \end{cases}$$

The posterior expectation immediatly follows:

$$\mathbb{E}[\Theta | X = x] = \frac{\alpha'}{\alpha'+\beta'} = \frac{\alpha+n}{\alpha+\cancel{n}+\beta+\sum x_i-\cancel{n}} = \frac{\alpha+n}{\alpha+\beta+\sum_{i=1}^{n} x_i}.$$

2 - Denote by $\hat{\theta} = \mathbb{E}[\Theta | X] = \hat{\theta}(X)$

We have $\hat{\theta} = \frac{\alpha+n}{\alpha+\beta+\sum_{i=1}^{n} X_i}$. Is $\hat{\theta}$ a consistent estimator of θ?

- Bias: $\mathbb{E}[\hat{\theta}] = \mathbb{E}\left[\frac{\alpha+n}{\alpha+\beta+\sum X_i}\right] = (\alpha+n)\, \mathbb{E}\left[\frac{1}{\alpha+\beta+\sum X_i}\right] \underset{\text{transfer theorem}}{=} \cdots$
- Variance of $\hat{\theta}$: ...

However, there is much simpler: indeed, when $n \to +\infty$, the impact of the hyperparameters α and β of the prior becomes negligible... that means that $\hat{\theta}(X) \xrightarrow[n\to+\infty]{} \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\overline{X}_n}$.

By the SLNN (Strong Law of Large Numbers), we know that $\overline{X}_n \xrightarrow[n\to+\infty]{a.s.} \mathbb{E}[X_i]$

If g is a regular function (continuous), then

$$g(\overline{X}_n) \xrightarrow[n\to+\infty]{a.s.} g(\mathbb{E}[X_i])$$. Take $g(x) = \frac{1}{x}$ here, we thus have:

$$\frac{1}{\overline{X}_n} \xrightarrow[n\to+\infty]{a.s} \frac{1}{\mathbb{E}[X_i]} = \frac{1}{1/\theta} = \theta$$. Since almost sure convergence implies also convergence in probability, we have $\hat{\theta}(X) \xrightarrow[n\to+\infty]{\mathbb{P}} \theta$.

Finally, $\hat{\theta}(X)$ is therefore a consistent estimator of θ.

Exercice 3 We consider an n-sized sample drawn from a Gaussian distribution such that $X = X_{1:n} \sim \mathcal{N}(\mu, \sigma^2)^{\otimes n}$, with known σ^2.

We observe $x = (x_1, \dots, x_n)$, and want to infer μ.

① Let us consider two estimators X_1 and $\bar{X}_n$ of μ.

$\rightarrow X_1$: • $\mathbb{E}[X_1 - \mu] = \mathbb{E}[X_1] - \mu = \mu - \mu = 0$
• $Var(X_1) = \sigma^2$

$\rightarrow \bar{X}_n$: • $\mathbb{E}[\bar{X}_n] = \mathbb{E}[X_i] = \mu$
• $Var(\bar{X}_n) = \frac{\sigma^2}{n}$

Both estimators are unbiased. However, $\bar{X}_n$ has a lower variance, and thus a lower MSE. (MSE = bias^2 + variance).

② We assume the following Bayesian model:

$$X_i \mid \mu = \mu \sim \mathcal{N}(\mu, \sigma^2)$$
$$\mu \sim \mathcal{N}(m, \tau^2)$$

The posterior distribution is given by:

$$\pi(\mu \mid x) = \frac{\pi(\mu) L(\mu \mid x)}{f_X(x)} = \frac{\pi(\mu) f_X(x \mid \mu)}{f_X(x)} = \frac{1}{f_X(x)} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\mu-m)^2}{2\tau^2}} \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\propto \frac{1}{\sqrt{2\pi\tau^2}\left(\sqrt{2\pi\sigma^2}\right)^n} e^{-\frac{(\mu-m)^2}{2\tau^2}} e^{-\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}} \propto e^{-\frac{(\mu-m)^2}{2\tau^2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i^2 - 2\mu x_i + \mu^2)}$$

$$\propto e^{-\frac{(\mu^2 - 2m\mu + m^2)}{2\tau^2} - \frac{1}{2\sigma^2}\left(\sum x_i^2 - 2\mu \sum x_i + n\mu^2\right)}$$

---- leads to complex computations.

It is better to separate each term of the computation: denote by $s^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2$

We have $\sum_i (x_i - \mu)^2 = \sum_i \left[(x_i - \bar{x}) - (\mu - \bar{x})\right]^2 = \sum_i (x_i - \bar{x})^2 + \sum_i (\bar{x} - \mu)^2 + 2\sum_i (x_i - \bar{x})(\mu - \bar{x})$

$$= ns^2 + n(\bar{x} - \mu)^2$$

since $\sum_i (x_i - \bar{x})(\mu - \bar{x}) = (\mu - \bar{x})\left[\left(\sum x_i\right) - n\bar{x}\right] = 0$

Hence $L(\mu|x) = \frac{1}{(2\pi)^{n/2}} \frac{1}{\sigma^n} e^{-\frac{1}{2\sigma^2}\left[ns^2 + n(\bar{x}-\mu)^2\right]}$

$$\propto \left(\frac{1}{\sigma^2}\right)^{n/2} e^{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2} e^{-\frac{ns^2}{2\sigma^2}} \overset{\sigma^2 \text{ known}}{\propto} \mathcal{N}\left(\bar{x}, \frac{\sigma^2}{n}\right)$$

Now remind that the prior follows $\pi(\mu) \propto \mathcal{N}(\mu_0, \sigma_0^2) \propto e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2}$

We can get the posterior distribution:

$$\pi(\mu|x) \propto \pi(\mu) L(\mu|x) \propto e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2} e^{-\frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2}$$

$$= e^{-\frac{1}{2\sigma^2}\sum_i (x_i^2+\mu^2-2x_i\mu) - \frac{1}{2\sigma_0^2}(\mu^2+\mu_0^2-2\mu_0\mu)}$$

$$\propto e^{-\frac{\mu^2}{2}\left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}\right) + \mu\left(\frac{\mu_0}{\sigma_0^2}+\frac{\sum_i x_i}{\sigma^2}\right) - \left(\frac{\mu_0^2}{2\sigma_0^2}+\frac{\sum_i x_i^2}{2\sigma^2}\right)}$$

$$:= e^{-\frac{1}{2\sigma_n^2}\left(\mu^2 - 2\mu\mu_n + \mu_n^2\right)} = e^{-\frac{1}{2\sigma_n^2}(\mu-\mu_n)^2} \quad \text{with :}$$

matching coefficients of μ^2, we find σ_n^2 is given by

$$-\frac{\mu^2}{2\sigma_n^2} = -\frac{\mu^2}{2}\left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}\right) \iff \boxed{\sigma_n^2 = \frac{\sigma^2\sigma_0^2}{n\sigma_0^2+\sigma^2} = \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}}}$$

Then matching coefficients of μ, we get:

$$-\frac{2\mu\mu_n}{-2\sigma_n^2} = \mu\left(\frac{\sum_{i=1}^n x_i}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right) \iff \frac{\mu_n}{\sigma_n^2} = \frac{\sigma_0^2 n\bar{x} + \sigma^2\mu_0}{\sigma^2\sigma_0^2}$$

Therefore $\mu_n = \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0 + \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\bar{x} = \boxed{\sigma_n^2\left(\frac{\mu_0}{\sigma_0^2}+\frac{n\bar{x}}{\sigma^2}\right) = \mu_n}$

Finally, we thus have $\pi(\mu|x) \propto \mathcal{N}(\mu_n, \sigma_n^2)$

③ - • What is the bayesian estimator for the quadratic error?

This is $\hat{\mu}^B = \mathbb{E}[\mu | X] = \mathbb{E}[\mu | X = (X_1, \dots, X_n)]$ where we know that the posterior is also Gaussian with parameters μ_n and σ_n^2.

Hence, $\hat{\mu}^B = \mu_n = \sigma_n^2 \left(\frac{\mu_0}{\sigma_0^2} + \frac{n\bar{x}}{\sigma^2} \right) = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}} \left(\frac{\mu_0}{\sigma_0^2} + \frac{n \bar{X}_n}{\sigma^2} \right)$

• What is the risk of $\hat{\mu}^B$?

$\rightarrow \mathbb{E}[\hat{\mu}^B] = \mathbb{E}\left[\frac{1}{n/\sigma^2 + 1/\sigma_0^2} \left(\frac{\mu_0}{\sigma_0^2} + \frac{n\bar{X}_n}{\sigma^2} \right) \right] = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}} \left[\frac{\mu_0}{\sigma_0^2} + \frac{1}{\sigma^2} n \underbrace{\mathbb{E}[\bar{X}_n]}_{= \mathbb{E}[X_i] = \mu} \right]$

$= \frac{\sigma^2 \cancel{\sigma_0^2}}{n\sigma_0^2 + \sigma^2} \times \frac{\mu_0}{\cancel{\sigma_0^2}} + \frac{\cancel{\sigma^2} \sigma_0^2}{n\sigma_0^2 + \sigma^2} \times \frac{n}{\cancel{\sigma^2}} \mu = \frac{\sigma^2 \mu_0}{n\sigma_0^2 + \sigma^2} + \frac{n\sigma_0^2 \mu}{n\sigma_0^2 + \sigma^2}$

$= \frac{\sigma^2 \mu_0 + n\sigma_0^2 \mu}{n\sigma_0^2 + \sigma^2} = \frac{\mu \left(n\sigma_0^2 + \frac{\sigma^2 \mu_0}{\mu} \right)}{n\sigma_0^2 + \sigma^2} = \mu \left(\frac{n\sigma_0^2 + \sigma^2 + \frac{\sigma^2 \mu_0}{\mu} - \sigma^2}{n\sigma_0^2 + \sigma^2} \right)$

$= \mu \left(1 + \frac{\sigma^2 \left(\frac{\mu_0}{\mu} - 1 \right)}{n\sigma_0^2 + \sigma^2} \right) = \mu + \frac{\sigma^2 (\mu_0 - \mu)}{n\sigma_0^2 + \sigma^2} \neq \mu.$

$\rightarrow \mathrm{Var}(\hat{\mu}^B) = \dots$

Exercice 5

1 - Denoting by $\pi(.|X)$ the density of posterior distribution, we have:

$$\forall \theta \in \mathbb{R}^{+*}, \quad \pi(\theta|X) \propto \theta^{p-1} e^{-\lambda\theta} \prod_{i=1}^{n} e^{-\theta} \frac{\theta^{X_i}}{X_i!}$$
$$\propto \theta^{p+n\bar{X}_n-1} e^{-(n+\lambda)\theta} \mathbb{1}_{\theta>0}.$$

Hence, $\Theta|X \sim \text{Gamma}(p+n\bar{X}_n, n+\lambda)$.

2 - We have $T^*(X) = \mathbb{E}[\Theta|X] = \dfrac{p+n\bar{X}_n}{\lambda+n}$. $T^*(X)$ is the bayesian estimator related to the quadratic loss function.

3 - We are now interested in the risk of this estimator:

$\forall \theta > 0$, we have

$$R(\theta, T^*) = \mathbb{E}_X\left[\left(\frac{p+n\bar{X}_n}{\lambda+n} - \theta\right)^2\right] = \mathbb{E}\left[\left(\frac{n(\bar{X}_n-\theta)}{\lambda+n} + \frac{p-\lambda\theta}{\lambda+n}\right)^2\right]$$

$$= \frac{n^2}{(\lambda+n)^2} \underbrace{\mathbb{E}\left[(\bar{X}_n-\theta)^2\right]}_{=(\text{bias}(\bar{X}_n))^2 + Var(\bar{X}_n)} + \frac{(p-\lambda\theta)^2}{(n+\lambda)^2} + \frac{2n(p-\lambda\theta)}{(n+\lambda)^2}\mathbb{E}[\bar{X}_n-\theta], \text{ with } \begin{cases} \mathbb{E}[\bar{X}_n-\theta]=0 \\ Var(\bar{X}_n) = \frac{\theta}{n} \end{cases}$$

$$= \frac{n^{\cancel{2}}}{(\lambda+n)^2} \times \frac{\theta}{\cancel{n}} + \frac{(p-\lambda\theta)^2}{(n+\lambda)^2} = \frac{n\theta + \lambda^2\left(\frac{p}{\lambda}-\theta\right)^2}{(n+\lambda)^2} \Rightarrow \text{Still depends on } \theta, \text{ which is random!}$$

4 - If we want to get the Bayesian risk, we thus have to consider the expectation over the distribution of Θ (prior distribution):

$$R^B = \mathbb{E}_\Theta\left[R(\theta, T^*)\right] = \mathbb{E}\left[\frac{n\Theta + \lambda^2\left(\frac{p}{\lambda}-\Theta\right)^2}{(n+\lambda)^2}\right]$$

$$= \frac{1}{(n+\lambda)^2}\left[n\,\mathbb{E}[\Theta] + \lambda^2\,\mathbb{E}\left[\left(\frac{p}{\lambda}-\Theta\right)^2\right]\right] = \frac{1}{(n+\lambda)^2}\left[n\frac{p}{\lambda} + \lambda^2 Var(\Theta)\right]$$

$$= \frac{1}{(n+\lambda)^2}\left[n\frac{p}{\lambda} + \cancel{\lambda^2}\frac{p}{\cancel{\lambda^2}}\right] = \frac{1}{\lambda(n+\lambda)^2}(np+\lambda p) = \frac{p}{\lambda(n+\lambda)} \quad \text{(is now a constant!)}$$

⑤ - We seek the minimizer of the posterior risk; where the posterior risk of some estimator T for the loss function ℓ follows:

$$\rho_\pi^\ell(T|X) = \int_\Theta \theta^3 e^{-2\theta} (\theta - T)^2 \, \pi(\theta|X) \, d\theta .$$

Furthermore, we know that $\Theta|X \sim \text{Gamma}(p + n\bar{X}_n, \lambda + n)$; thus:

$$\rho_\pi(T|X) = \frac{(\lambda+n)^{p+n\bar{X}_n}}{\Gamma(p+n\bar{X}_n)} \int_0^{+\infty} (\theta - T)^2 \theta^3 e^{-2\theta} \theta^{p+n\bar{X}_n - 1} e^{-(\lambda+n)\theta} d\theta$$

$$= \frac{(\lambda+n)^{p+n\bar{X}_n}}{\Gamma(p+n\bar{X}_n)} \int_0^{+\infty} (\theta - T)^2 \theta^{p+n\bar{X}_n+2} e^{-\theta(\lambda+n+2)} d\theta$$

We notice that $\rho_\pi^\ell(T|X)$ is a function of T, proportional to $\mathbb{E}[(\tilde{\theta} - T)^2 | X]$ where the distribution of $\tilde{\theta}|X$ follows a Gamma such that $\tilde{\theta}|X \sim \text{Gamma}(n\bar{X}_n + p + 3, \lambda + n + 2)$. We also know that for all Y random variable with finite second moment, the function $a \mapsto \mathbb{E}[(Y-a)^2|X]$ takes its minimum value at $a = \mathbb{E}[Y|X]$. This leads to the minimizer:

$$T_\ell(X) = \frac{n\bar{X}_n + p + 3}{\lambda + n + 2} .$$

Exercice 6 : Un portefeuille d'assurance automobile est composé de 35% de bons conducteurs, 40% de conducteurs moyens et 25% de mauvais conducteurs. L'actuaire a estimé que les bons conducteurs ont en moyenne un accident tous les 10 ans, les moyens 2 accidents, et les mauvais 6 accidents. L'actuaire suppose de plus que la fréquence des accidents a une distribution de Poisson. Pour simplifier, les sinistres coûtent tous 1€. (6)

a) What is the probability to have 1 claim for a policyholder taken at random?

Let N denote the claim frequency: $N \sim \mathcal{P}(\theta)$

the r.v. Θ represents the risk profile of the driver : $\Theta \in \{\theta^m ; \theta^M ; \theta^b\}$ (bad, standard, good)

Moreover, we know :
$$\begin{cases} N \mid \Theta = \theta^m \sim \mathcal{P}(\theta^m = 0{,}6) \\ N \mid \Theta = \theta^M \sim \mathcal{P}(\theta^M = 0{,}2) \\ N \mid \Theta = \theta^B \sim \mathcal{P}(\theta^B = 0{,}1) \end{cases}$$

$$\mathbb{P}(N=1) = \mathbb{P}(N=1 \mid \Theta = \theta^m)\,\mathbb{P}(\Theta = \theta^m) + \mathbb{P}(N=1 \mid \Theta = \theta^M)\,\mathbb{P}(\Theta = \theta^M) + \mathbb{P}(N=1 \mid \Theta = \theta^B)\,\mathbb{P}(\Theta = \theta^B)$$

$$= \left(e^{-0{,}6}\frac{0{,}6^1}{1!}\right) \times 0{,}25 + \left(e^{-0{,}2}\frac{0{,}2^1}{1!}\right) \times 0{,}4 + \left(e^{-0{,}1}\frac{0{,}1^1}{1!}\right) \times 0{,}35$$

$$= 0{,}1795$$

b) Compute the premium for each of the 3 types of drivers, knowing that the risk premium is $\mathbb{E}[S \mid \Theta = \theta]$ (indeed, claim cost is 1€).

- $\pi(\theta^m) = \mathbb{E}[S \mid \Theta = \theta^m] = \mathbb{E}[N \mid \Theta = \theta^m] = 0{,}6$
- $\pi(\theta^M) = \mathbb{E}[S \mid \Theta = \theta^M] = 0{,}2$
- $\pi(\theta^B) = \mathbb{E}[S \mid \Theta = \theta^B] = 0{,}1$.

c) Calculate the collective premium:

$$\pi^{coll} = \mathbb{E}_{\Theta}[\pi(\Theta)] = \pi(\theta^m)\,\mathbb{P}(\Theta = \theta^m) + \pi(\theta^M)\,\mathbb{P}(\Theta = \theta^M) + \pi(\theta^B)\,\mathbb{P}(\Theta = \theta^B)$$

$$= 0{,}6 \times 0{,}25 + 0{,}2 \times 0{,}4 + 0{,}1 \times 0{,}35$$

$$= 0{,}265.$$

d) Calculate the bayesian premium for year 6:

We look for $\mathbb{E}[S \mid s = (s_1, s_2, s_3, s_4, s_5)]$.

We have
$$\mathbb{E}[S \mid s] = \mathbb{E}[S \mid \Theta = \theta^m]\,\mathbb{P}(\Theta = \theta^m \mid s) + \mathbb{E}[S \mid \Theta = \theta^M]\,\mathbb{P}(\Theta = \theta^M \mid s) + \mathbb{E}[S \mid \Theta = \theta^B]\,\mathbb{P}(\Theta = \theta^B \mid s)$$

$\Rightarrow$ We need the posterior distribution:

- $$\mathbb{P}(\Theta = \theta^m \mid s) = \frac{\mathbb{P}(s \mid \Theta = \theta^m)\,\mathbb{P}(\Theta = \theta^m)}{\sum_\theta \mathbb{P}(S = s \mid \Theta = \theta)\,\mathbb{P}(\Theta = \theta)} = \frac{\prod_{i=1}^{5} \mathbb{P}(S = s_i \mid \Theta = \theta^m)\,\mathbb{P}(\Theta = \theta^m)}{\sum_\theta \mathbb{P}(S = s \mid \Theta = \theta)\,\mathbb{P}(\Theta = \theta)}$$

$$= \frac{0{,}25\left[e^{-0{,}6} \times 0{,}6^1/1! \times e^{-0{,}6} \times 0{,}6^0/0! \times e^{-0{,}6} \times 0{,}6^1/1! \times e^{-0{,}6} \times 0{,}6^1/1! \times e^{-0{,}6}\, 0{,}6^0/0!\right]}{\underbrace{0{,}25\left[(0{,}6\, e^{-0{,}6})^3 (e^{-0{,}6})^2\right]}_{= 0{,}00\cancel{\;}26885} + \underbrace{0{,}4\left[(0{,}2\, e^{-0{,}2})^3 (e^{-0{,}2})^2\right]}_{0{,}001177214} + \underbrace{0{,}35\left[(0{,}1\, e^{-0{,}1})^3 (e^{-0{,}1})^2\right]}_{0{,}000212285}}$$

$$= 0{,}6598$$

- $\mathbb{P}(\Theta = \theta^M \mid s) = 0{,}2886742$
- $\mathbb{P}(\Theta = \theta^B \mid s) = 0{,}0520563$

$\Rightarrow$ Check that $\sum_\theta \mathbb{P}(\Theta = \theta \mid s) = 1$.

Finally,

$$\mathbb{E}[S|s] = \underbrace{\mathbb{E}[N|\Theta=\theta^m]}_{\sim \mathcal{P}(0,6)} \mathbb{P}(\Theta=\theta^m|s) + \underbrace{\mathbb{E}[N|\Theta=\theta^M]}_{\sim \mathcal{P}(0,2)} \mathbb{P}(\Theta=\theta^M|s) + \underbrace{\mathbb{E}[N|\Theta=\theta^B]}_{\sim \mathcal{P}(0,1)} \mathbb{P}(\Theta=\theta^B|s)$$

$$= 0,6 \times 0,6592 + 0,2 \times 0,2886742 + 0,1 \times 0,0520563$$

$$= 0,4585.$$

Exercise 7) Les montants de sinistres d'un contrat furent $\begin{cases} S_1 = 7 \\ S_2 = 13 \\ S_3 = 1 \\ S_4 = 4 \end{cases}$, au cours de 4 premières années d'assurance.

Votre expérience antérieure avec ce type de contrat vous permet de postuler le modèle suivant pour les montants de sinistres:

$$\begin{cases} \mathbb{P}(S=x \mid \Theta=\theta) = \binom{x+4}{4} \theta^5 (1-\theta)^x, & x = 0,1,\dots \\ u(\theta) = 504\, \theta^5 (1-\theta)^3, & 0 < \theta < 1. \end{cases}$$

Calculate the bayesian premium of 5th year:

We look for $\mathbb{E}[S \mid s = (s_1, s_2, s_3, s_4)]$. As previously,

$$\mathbb{E}[S|s] = \int_\Theta \mathbb{E}[S|\Theta=\theta]\, f_{\Theta|s}(\theta)\, d\theta$$

$\Rightarrow$ One needs the distribution of posterior: $\Theta | S=s$:

$$f_{\Theta|S=s}(\theta) = \frac{\mathbb{P}(S=s|\Theta=\theta)\, u(\theta)}{\int_\Theta \mathbb{P}(S=s|\Theta=\theta)\, u(\theta)\, d\theta} = \frac{\overbrace{\prod_{i=1}^{4} \mathbb{P}(S=s_i|\Theta=\theta)\, u(\theta)}^{A}}{\underbrace{\int_0^1 \mathbb{P}(S=s|\Theta=\theta)\, u(\theta)\, d\theta}_{B}}$$

$$A = 504\, \theta^5 (1-\theta)^3 \Big[\underbrace{\binom{7+4}{4}}_{330} \theta^5 (1-\theta)^7 \times \underbrace{\binom{13+4}{4}}_{2380} \theta^5 (1-\theta)^{13} \times \underbrace{\binom{1+4}{4}}_{5} \theta^5 (1-\theta)^1 \times \underbrace{\binom{4+4}{4}}_{70} \theta^5 (1-\theta)^4 \Big]$$

$=$

Ainsi, $A = 504\,\theta^5(1-\theta)^3\left[\theta^{20}(1-\theta)^{25} \times 330 \times 2380 \times 5 \times 70\right]$

D'autre part, $=$

$$B = \int_0^1 \mathbb{P}(S=s \mid \Theta=\theta)\, u(\theta)\, d\theta$$

$$= \int_0^1 A\, u(\theta)\, d\theta = \int_0^1 504\,\theta^5(1-\theta)^3\left[\theta^{20}(1-\theta)^{25} \times 330 \times 2380 \times 5 \times 70\right] u(\theta)\, d\theta$$

$$= 504 \times 330 \times 2380 \times 5 \times 70 \times \int_0^1 \theta^{25}(1-\theta)^{28} \times 504\,\theta^5(1-\theta)^3\, d\theta$$

$$= (504)^2 \times 330 \times 2380 \times 5 \times 70 \times \int_0^1 \theta^{30}(1-\theta)^{31}\, d\theta$$

Ainsi $$f_{\Theta \mid S=s}(\theta) = \frac{\cancel{504} \times \cancel{330} \times \cancel{2380} \times \cancel{5} \times \cancel{70} \times \theta^{25}(1-\theta)^{28}}{(504)^{\cancel{2}} \times \cancel{330} \times \cancel{2380} \times \cancel{5} \times \cancel{70} \times \int_0^1 \theta^{30}(1-\theta)^{31}\, d\theta}$$

$$= \frac{\theta^{25}(1-\theta)^{28}}{504 \underbrace{\int_0^1 \theta^{30}(1-\theta)^{31}\, d\theta}_{= 6{,}93 \times 10^{-20}}}$$ avec $\int_0^1 \theta^{30}(1-\theta)^{31}\, d\theta = \int_0^1 \left(\theta(1-\theta)\right)^{30}(1-\theta)\, d\theta$