

Lista 5 - Power Prior

Enunciado

- Dados históricos: $Y_{0,i} \sim \mathcal{N}(\theta_0, \sigma_0^2)$, $i = 1, \dots, n_0$
- Dados atuais: $Y_i \sim \mathcal{N}(\theta, \sigma^2)$, $i = 1, \dots, n$
- Priori inicial para θ : $\pi(\theta) \propto 1$
- As variâncias são conhecidas
- **Questão 1:** Considerando $\theta = \theta_0$, encontre a Power Prior e a distribuição a posteriori para θ .
- **Questão 2:** Estudo de simulação: Sugiro verificar via simulação se este procedimento bayesiano apresenta probabilidade de cobertura de 95% para diferentes valores de a_0 . É necessário assumir que $\sigma^2 = \sigma_0^2$?

Questão 1

A Power Prior é definida por:

$$\pi(\theta \mid \mathbf{y}_0, a_0) \propto [\mathcal{L}(\theta_0 \mid \mathbf{y}_0, \sigma_0^2)]^{a_0} \pi(\theta) \quad (1)$$

onde $\mathcal{L}(\theta_0 \mid \mathbf{y}_0, \sigma_0^2)$ é a função de verossimilhança dos dados históricos e a_0 é um hiperparâmetro que controla as caudas da priori.

Então distribuição a posteriori para θ fica:

$$\pi(\theta \mid \mathbf{y}, \mathbf{y}_0, a_0) \propto \mathcal{L}(\theta \mid \mathbf{y}, \sigma^2) [\mathcal{L}(\theta_0 \mid \mathbf{y}_0, \sigma_0^2)]^{a_0} \pi(\theta) \quad (2)$$

As funções de verossimilhança dos dados atuais e históricos são dadas por:

$$\mathcal{L}(\theta \mid \mathbf{y}, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \theta)^2 \right\} \quad (3a)$$

$$\mathcal{L}(\theta_0 \mid \mathbf{y}_0, \sigma_0^2) = \left(\frac{1}{\sqrt{2\pi\sigma_0^2}} \right)^{n_0} \exp \left\{ -\frac{1}{2\sigma_0^2} \sum_{i=1}^{n_0} (y_{0,i} - \theta_0)^2 \right\} \quad (3b)$$

Para conseguirmos melhor interpretar a Power Prior, vamos re-escrever as verossimilhanças em função das estatísticas suficientes:

$$\mathcal{L}(\theta \mid \mathbf{y}, \sigma^2) \propto \exp \left\{ -\frac{n}{2\sigma^2} (\theta - \bar{y})^2 \right\} \quad (4a)$$

$$\mathcal{L}(\theta_0 \mid \mathbf{y}_0, \sigma_0^2) \propto \exp \left\{ -\frac{n_0}{2\sigma_0^2} (\theta_0 - \bar{y}_0)^2 \right\} \quad (4b)$$

onde \bar{y} e \bar{y}_0 são as médias das observações atuais e passadas, respectivamente. Podemos então re-escrever a posteriori como:

$$\pi(\theta \mid \mathbf{y}, \mathbf{y}_0, a_0) \propto \exp \left\{ -\frac{n}{2\sigma^2} (\theta - \bar{y})^2 \right\} \exp \left\{ -\frac{a_0 n_0}{2\sigma_0^2} (\theta_0 - \bar{y}_0)^2 \right\} \pi(\theta) \quad (5)$$

Agora, assumindo que $\theta = \theta_0$ e $\pi(\theta) \propto 1$, com um pouco de algebrismo (semelhante ao caso da inferência Bayesiana com o modelo normal conjugado) é possível combinar os dois termos exponenciais da eq. (5) em um único, obtendo então o resultado de que a posteriori possui distribuição normal:

$$\pi(\theta \mid \mathbf{y}, \mathbf{y}_0, a_0) \sim \mathcal{N}(\theta \mid \mu_*, \sigma_*^2) \quad (6)$$

com

$$\mu_* = \frac{\frac{n}{\sigma^2} \bar{y} + a_0 \frac{n_0}{\sigma_0^2} \bar{y}_0}{\frac{n}{\sigma^2} + a_0 \frac{n_0}{\sigma_0^2}} = \frac{\tau \bar{y} + a_0 \tau_0 \bar{y}_0}{\tau + a_0 \tau_0} \quad (7a)$$

$$\sigma_*^2 = \frac{1}{\frac{n}{\sigma^2} + a_0 \frac{n_0}{\sigma_0^2}} = \frac{1}{\tau + a_0 \tau_0} \quad (7b)$$

Interpretação

Para este caso em que adotamos a priori imprópria de Bayes-Laplace para $\pi(\theta)$ e assumimos que $\theta = \theta_0$, a aplicação da Power Prior com o modelo normal conduz a uma posteriori normal com parâmetros influenciados pelos dados atuais, pelos dados históricos e também pelo hiperparâmetro a_0 da Power Prior.

Nas equações (7a) e (7b), note que os termos $\frac{n}{\sigma^2}$ e $\frac{n_0}{\sigma_0^2}$ podem ser interpretados como a precisão dos dados presentes (τ) e passados (τ_0), respectivamente).

Para o caso em que $a_0 = 1$, recuperamos o cenário de inferência bayesiana com o modelo normal conjugado, ou seja: a interpretação é que temos uma priori normal $\mathcal{N}(\bar{y}_0, \sigma_0^2/n_0)$, construída com os dados passados, e então esta priori é atualizada com informações pelos dados presentes (modelo normal), resultando assim em uma posteriori (também normal) cuja média μ_* é média das médias dos dados atuais (\bar{y}) e passados ponderada (\bar{y}_0) pelas respectivas precisões atual (τ) e passada (τ_0); e cuja variância σ_*^2 também um balanço entre a variância da amostra atual e da passada.

Agora, quando $0 \leq a_0 < 1$, o que este hiperparâmetro faz é controlar o peso (influência) dos dados passados: quanto maior a_0 , mais os dados passados influenciam na posteriori. No caso trivial em que $a_0 = 0$, os dados passados são ignorados e os estimadores obtidos são os de máxima verossimilhança.

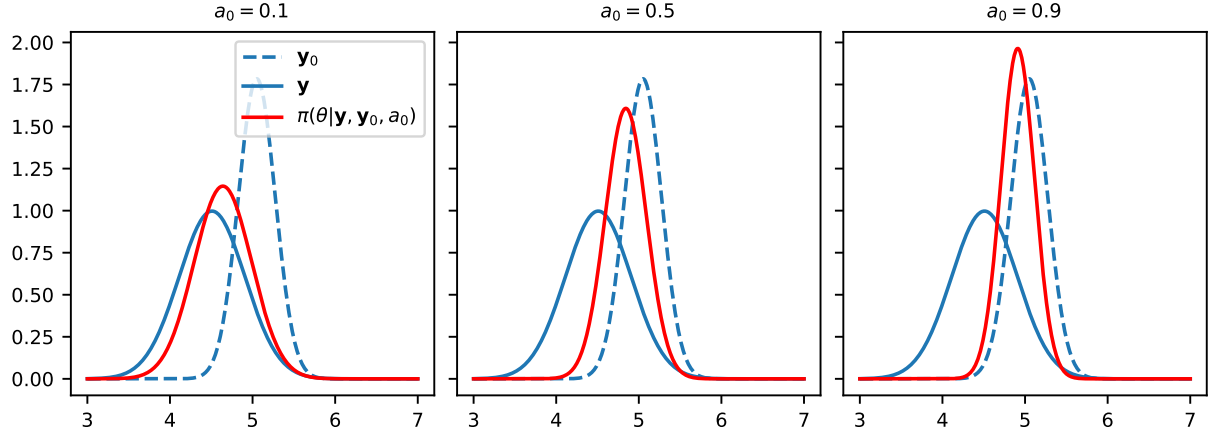


Figura 1: Exemplo simulado para a questão 1. Observe que, para $a_0 = 0.1$, a posteriori sofre pouca influência dos dados históricos \mathbf{y}_0 . A situação oposta ocorre como $a_0 = 0.9$. Com $a_0 = 0.5$, ocorre um cenário de meio-termo.

Por fim, ressaltamos que neste desenvolvimento não precisamos assumir $\sigma^2 = \sigma_0^2$, apenas $\theta = \theta_0$.

Para ilustrarmos, no gráfico da fig. 1 consideramos $\theta = \theta_0 = 5$, $\sigma_0^2 = 4$, $\sigma^2 = 16$, $n_0 = 80$, $n = 100$ e $a_0 = \{0.1, 0.5, 0.9\}$. Observe que, para $a_0 = 0.1$, a posteriori sofre pouca influência dos dados históricos \mathbf{y}_0 . A situação oposta ocorre com $a_0 = 0.9$. Por fim, para $a_0 = 0.5$, ocorre uma situação de equilíbrio: a posteriori é influenciada com o mesmo peso pelos dados históricos e pelos dados atuais.

Questão 2 - Simulação

Nesta questão, à partir da posteriori, determinamos o intervalo de credibilidade 95% para θ e então verificamos a influência do hiperparâmetro a_0 e das variâncias na probabilidade de cobertura simulada, mantendo fixo os demais parâmetros. Adotamos as seguintes estratégias e valores para a simulação:

- Médias populacionais: $\theta = \theta_0 = 5$;
- Variâncias populacionais:
 - Cenário 1: $\sigma_0^2 = \sigma^2 = 4$;
 - Cenário 2: $\sigma_0^2 = 4, \sigma^2 = 16$;
 - Cenário 3: $\sigma_0^2 = 16, \sigma^2 = 4$;
- Tamanho das amostras: $n_0 = 80, n = 100$;
- Valores de a_0 discretizados: $(0.01, 0.02, \dots, 1.00)$;
- Número de simulações para cada a_0 : 1000.

Os resultados são mostrados no gráfico da fig. 2.

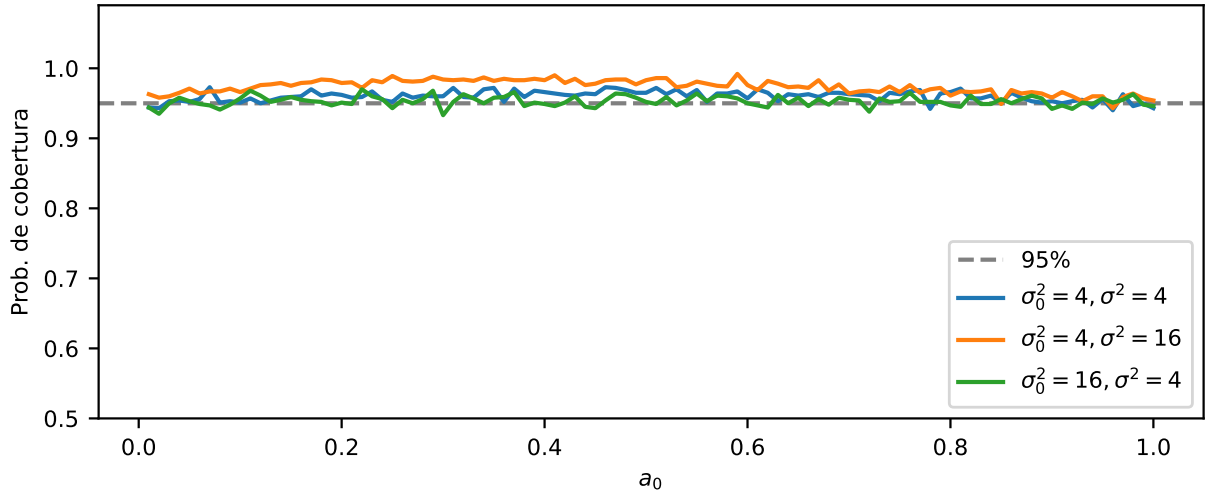


Figura 2: Simulação da probabilidade de cobertura 95% para a inferência com a Power Prior com diferentes valores de a_0 e diferentes valores de variâncias para os dados passado e presente. Observe que, para os 3 cenários, a probabilidade de cobertura ficaram próximas ao valor esperado, exceto talvez no caso em que $\sigma^2 = 16$ e $\sigma_0^2 = 4$, o qual apresentou probabilidade de cobertura ligeiramente acima do esperado para valores de a_0 na região central (próximos a $a=0.5$)

Observe que, para os três cenários, as probabilidades de cobertura simuladas ficaram próximas ao valor esperado de 95% para toda a faixa de valores de a_0 . Exceto talvez para o caso em que $\sigma^2 = 16$ e $\sigma_0^2 = 4$. Neste caso, para valores de a_0 próximos 0.5, aparentemente a probabilidade de cobertura ficou ligeiramente superior à esperada. Porém, não conseguimos avaliar se trata-se de um efeito realmente sistemático (decorrente do fato de estarmos considerando uma variância dos dados atuais bastante superior aos dados históricos), ou é algum efeito aleatório de nossa simulação. Em todo caso, **não há necessidade de assumirmos $\sigma^2 = \sigma_0^2$** .

Implementação

A simulação foi implementada em Python, o código-fonte está disponível no repositório <https://github.com/cleitonmoya/ib/>.