

tarea 1

Gabriel Bermudez, Matias Bajac

2025-06-12

1)

$$f(x_1, x_2, \dots, x_n | \lambda) = f(\vec{x} | \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} f(\vec{x} / \lambda) = \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} \dots \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} =$$

$$\frac{[\lambda^{x_1} \lambda^{x_2} \dots \lambda^{x_n}] [e^{-\lambda} e^{-\lambda} \dots e^{-\lambda}]}{x_1! x_2! \dots x_n!} = \frac{\lambda^{\sum_{i=1}^n x_i} e^{-\lambda n}}{\prod_{i=1}^n x_i!}$$

2) Compute the Maximum Likelihood Estimator (MLE) $\hat{\lambda}_{MLE}$ of λ given \underline{X} .

$$\ell(\lambda | \underline{X}) = \sum X_i \ln(\lambda) - \lambda n$$

$$\frac{\partial \ell(\lambda | \underline{X})}{\partial \lambda} = \frac{\sum X_i}{\lambda} - n = 0 \iff \lambda = \frac{\sum X_i}{n}$$

$$\implies \boxed{\hat{\lambda}_{MLE} = \bar{X}}$$

3) Choose (and justify your choice) the functional form of the prior distribution on λ considering one of the distributions presented in class, i.e Gaussian (Normal), Gamma, Exponential, Beta, Uniform

Debido a que el soporte de la distribución Poisson es $\lambda > 0$, es razonable pensar que una priori conjugada para λ podría ser una priori $\lambda \sim \text{Gamma}(\alpha, \beta)$ con soporte $\alpha > 0$ y $\beta > 0$.

4) Compute the posterior distribution of λ and a Bayesian estimator $\hat{\lambda}_{PO}$

$$f(\lambda | x_1, x_2, \dots, x_n) \propto f(\lambda) f(\vec{x} | \lambda) \propto \lambda^{\alpha-1} e^{-\beta\lambda} \lambda^{\sum x_i} e^{-\lambda n} \propto$$

$$e^{-\lambda n - \beta\lambda} \lambda^{\alpha-1 + \sum_{i=1}^n x_i} \propto e^{\lambda(n+\beta)} \lambda^{\alpha-1 + \sum_{i=1}^n x_i}$$

$$\lambda | x_1, \dots, x_n \sim \text{Gamma}(\alpha, \sum_{i=1}^n x_i + \beta)$$

Bajo pérdida cuadrática, el estimador bayesiano es la media a posteriori

$$\hat{\lambda}_{PO} = E(\lambda | \vec{X}) = \frac{\alpha + \sum_{i=1}^n x_i}{\beta + n}$$

5) (Optional) Comment about what happens too $\hat{\lambda}_{PO}$ when (like done in class)

- letting the prior variance going to 0
- letting the prior variance going to ∞

- letting the sample size going to ∞

$$\text{media a priori } \hat{\lambda} = \frac{\alpha k + \sum_{i=1}^n x_i}{\beta_k + n}$$

$$k \rightarrow 0 \text{ varianza a priori} \rightarrow \infty \rightarrow \hat{\lambda} \rightarrow \sum_{i=1}^n \frac{x_i}{n} = \bar{x} \sim \text{EMV}$$

$$k \rightarrow \infty \text{ varianza a priori} \rightarrow 0 \rightarrow \hat{\lambda} \rightarrow \frac{\alpha}{\beta} \rightarrow \text{media a priori}$$

$$n \rightarrow \infty \rightarrow \hat{\lambda} \sim \frac{\alpha + \mathcal{K}\bar{x}}{\beta + \mathcal{K}} \rightarrow \bar{x} \sim \text{EMV}$$

- 6) Compute the posterior predictive distribution for a next observation X_{n+1} given the sample X

$$f_{X_{n+1}}(X_{n+1}/X) = \int_0^\infty \underbrace{\frac{\lambda^{x_{n+1}} e^{-\lambda}}{x_{n+1}!}}_{f(x_{n+1}/\lambda)} \underbrace{\frac{(n+\beta)}{\Gamma\alpha + \sum x_i} \lambda^{\alpha + \sum x_i - 1} e^{-\lambda(n+\beta)}}_{f(\lambda/X=x)} d\lambda =$$

$$\frac{(n+\beta)^{\alpha + \sum x_i}}{\Gamma\alpha + \sum x_i} \frac{1}{x_{n+1}!} \int_0^\infty \lambda^{x_{n+1}-1 + \alpha + \sum x_i} e^{-\lambda(1+\beta+n)} d\lambda$$

- 7) Describe, in your words, why MLE and posterior mean are sound choices as estimators of a parameter

El estimador máximo verosímil da la media muestral el cual parece ser una buena elección en inferencia clásica ya que usa únicamente la información de los datos, dando como resultado un estadístico de resumen mientras tanto en un enfoque bayesiano, la media a posteriori (pérdida cuadrática) incorpora la información a priori además de la verosimilitud de los datos.

- 9) (Optional) You will do the exercise considering i.i.d. Poisson distributions but you should comment about such choice, i.e. on the assumptions that we are making considering Poisson and considering i.i.d.. Are they completely justified or not? If not, what are the critical aspects?

Considering that the Poisson Distribution is only defined for positive and discrete values, this aligns with the data we have: because of course there can't be negative deaths or a fraction of them. Also, the Poisson Distribution is skewed to the right, more notoriously when the mean is *small*, meaning there is a bigger probability mass at the lower values of *number of deaths*, which makes sense with the situation we are modelling as well.

- 10) Compute the MLE of λ .

En el ejercicio número 2 se llegó a la conclusión que si los datos siguen una distribución de Poisson, entonces el MLE de λ es igual a la **media muestral**. En este caso:

$$\hat{\lambda}_{MLE} = \bar{X} = 10$$

11. Based on your experience in your home country, assign the parameters to the prior distribution on λ . As in class, it might be useful to think what is the "physical meaning" of λ and then think of its possible value, to be considered as a mean, and a trust on such assessment, which could be considered as a variance.

Teniendo en cuenta que: - En 2024 se registraron casi 15 muertes en accidentes automovilísticos por 100,000 habitantes en Uruguay, y - Que esperamos que este año sea mejor en este sentido,

Vamos a querer que la prior mean de $\pi(\lambda)$ sea 14. Recordando que elegimos como distribución a priori una Gamma ($\mathcal{G}(\alpha, \beta)$), podemos plantear la siguiente igualdad:

$$\mathbb{E}[\pi(\lambda)] = \frac{\alpha}{\beta} = 14$$

$$\alpha = 14\beta$$

$$\implies (\alpha, \beta) = (14, 1)$$

Lo que nos deja con una confianza (Trust) en nuestra estimación de:

$$\mathbb{V}(\pi(\lambda)) = \frac{\alpha}{\beta^2} = 14 \implies \text{sd}(\text{'Trust'}) = \sqrt{14} \simeq 3.74$$