

EST-46115: Modelación Bayesiana

Profesor: Alfredo Garbuno Iñigo — Primavera, 2021.

Objetivo. Repasar notación que utilizaremos a lo largo del curso. Y a la vez, establecer la motivación de los temas que trataremos en la materia.

Lecturas recomendadas: Notas del [curso de fundamentos](#) (2022) y sección 1 de [2].

1. NOTACIÓN

Usaremos la convención usual en probabilidad.

1.1. Variables aleatorias

Una variable aleatoria X está definida a través de un **espacio de probabilidad** $(\mathcal{X}, \mathcal{F}, \pi)$. La función $\pi : \mathcal{F} \rightarrow [0, 1]$ se llama la **función de distribución** de la variable aleatoria X . Escribimos $X \sim \pi$.

1.2. Distribución paramétrica

Decimos que una función de distribución es **paramétrica** si se puede identificar completamente la distribución con respecto a un **vector de parámetros** $\theta \in \mathbb{R}^p$. Esto lo denotamos de la siguiente manera

$$\pi_{\theta}(x) \quad \text{ó} \quad \pi(x; \theta), \quad (1)$$

y si $\theta \neq \theta'$ entonces $\pi_{\theta}(x) \neq \pi_{\theta'}(x)$ para cualquier x en el **soporte**.

1.3. Valores esperados

El **valor esperado** de una variable aleatoria $X \sim \pi$ se define como

$$\mathbb{E}[X] = \int_{\mathcal{X}} x \pi(x) dx. \quad (2)$$

1.3.1. Abuso de notación La **función de densidad** está definida como $d\pi/dx$. En la definición de valor esperado deberíamos de haber escrito $d\pi(x)$ o bien $\pi(dx)$ (integrales de Lebesgue). Pero para no ofuscar notación, y dado que no veremos cosas exóticas en el curso, lo obviaremos...

1.4. Estadísticas de interés

La definición se puede extender con $f : \mathcal{X} \rightarrow \mathbb{R}$ y se calcula como

$$\mathbb{E}[f(X)] = \int_{\mathcal{X}} f(x) \pi(x) dx. \quad (3)$$

Denotaremos de la siguiente manera

$$\pi(f) := \mathbb{E}[f(X)]. \quad (4)$$

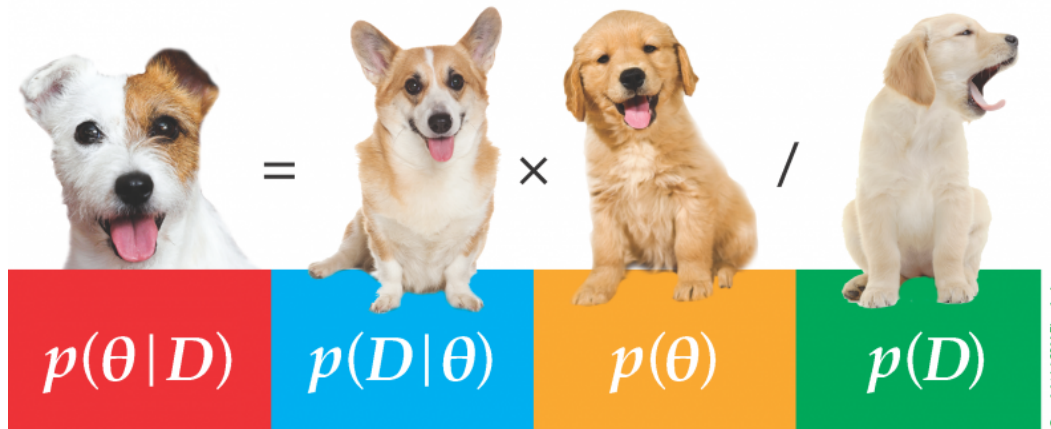


FIGURA 1. Tomado de [4] .

1.5. Probabilidad condicional

La probabilidad condicional de A dado el evento B se denota $\pi(A|B)$ y está definida como

$$\pi(A|B) = \frac{\pi(A \cap B)}{\pi(B)} \quad (5)$$

2. REPASO INFERENCIA

Repaso de inferencia bajo un enfoque frecuentista.

2.1. Regla de Bayes

La regla de Bayes utiliza la definición de probabilidad condicional para hacer inferencia a través de

$$\pi(A|B) = \frac{\pi(B|A)\pi(A)}{\pi(B)}. \quad (6)$$

2.2. Ejemplos

- Verosimilitud: $x|\theta \sim \text{Binomial}(n, \theta)$ + Previa: $\theta \sim \text{Beta}(\alpha, \beta)$ = Posterior: ?
- Verosimilitud: $x|\theta \sim \text{Uniforme}(0, \theta)$ + Previa: $\theta \sim \text{Pareto}(\theta_0)$ = Posterior: ?

3. REPASO INFERENCIA

Repaso de inferencia bajo un enfoque bayesiano.

3.1. Ejemplo

Este ejemplo fue tomado de [3].

3.2. Diferentes previas, diferentes posteriores

```
1 modelo_beta <- function(params, n = 5000){
2   rbeta(n, params$alpha, params$beta)
3 }
```

```

1 escenarios <-
2   tibble(analista = fct_inorder(c("Ignorante", "Indiferente",
3                                 "Feminista", "Ingenuo")),
4         alpha = c(1, .5, 5, 14),
5         beta = c(1, .5, 11, 1)) |>
6   nest(params.previa = c(alpha, beta)) |>
7   mutate(muestras.previa = map(params.previa, modelo_beta))

```

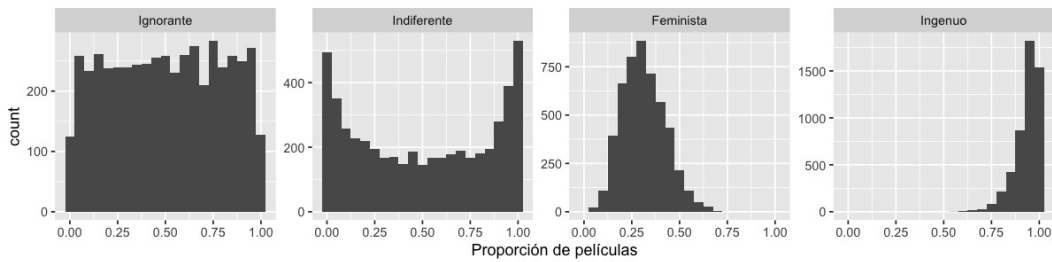
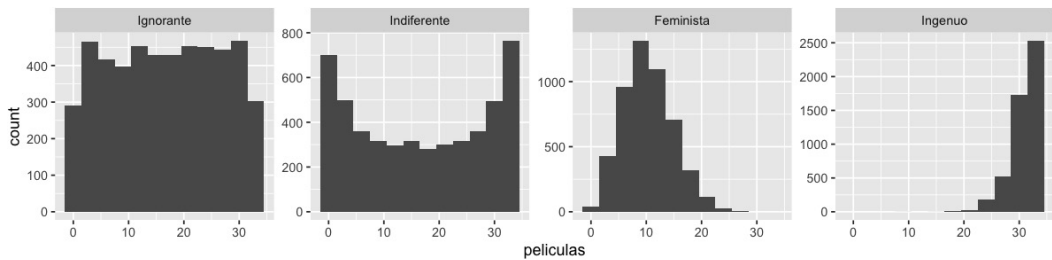
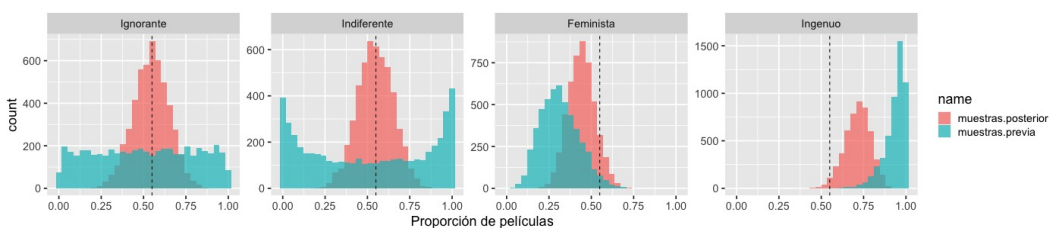
FIGURA 2. Muestras de $\theta \sim$ Previa .

FIGURA 3. Distribución predictiva previa

```

1 update_rule <- function(params){
2   tibble(alpha = params$alpha + data$PASS,
3         beta = params$beta + data$FAIL)
4 }
5 escenarios <- escenarios |>
6   mutate(params.posterior = map(params.previa, update_rule),
7         muestras.posterior = map(params.posterior, modelo_beta))

```



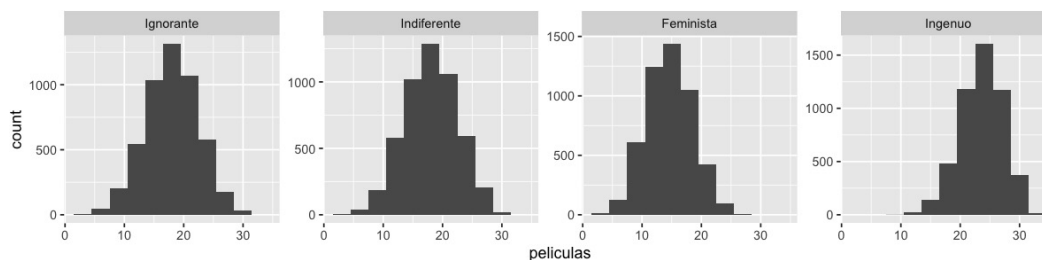
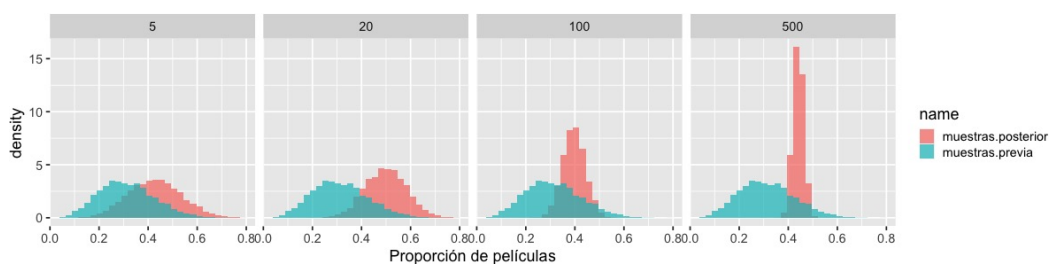


FIGURA 4. Predictiva posterior.

3.3. Diferentes datos, diferentes posteriores



3.4. Análisis secuencial

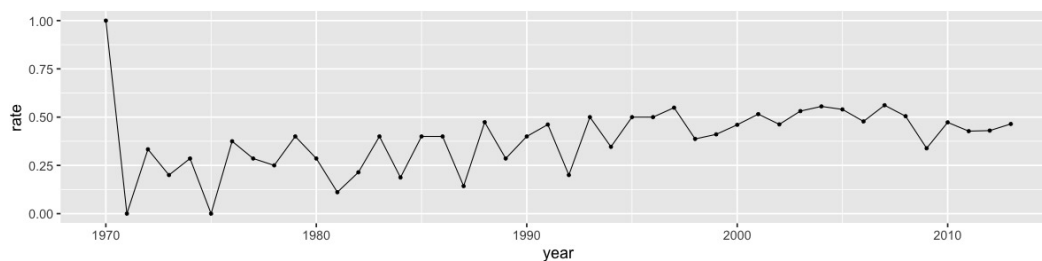


FIGURA 5. Histórico de la proporción de películas que pasan la prueba de Bechdel por año.

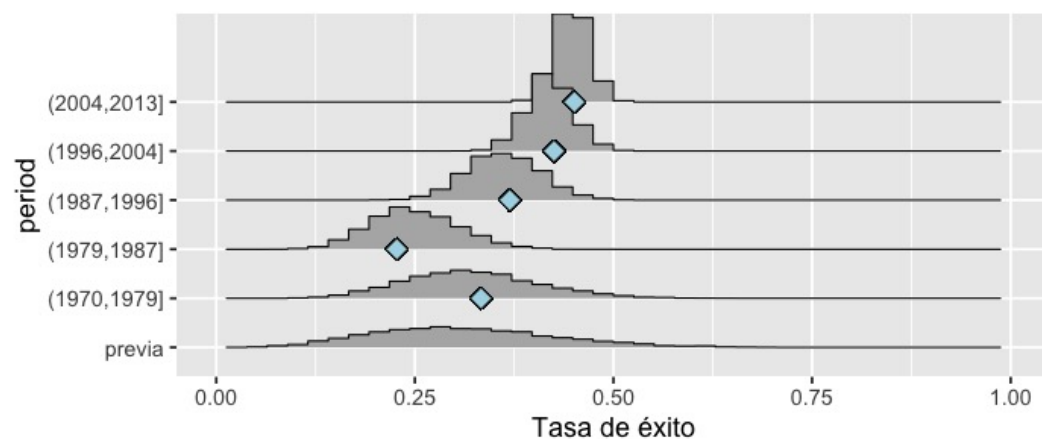


FIGURA 6. La posterior de hoy puede ser la previa de mañana.

3.5. Tarea

Echenle un ojo a la sección 5.2 de [Bayes rules!](#) donde se expone a detalle un modelo más del análisis conjugado. ¿Puedes identificar la distribución predictiva?

4. MOTIVACIÓN

Por medio de metodología Bayesiana podemos cuantificar incertidumbre en:

- Observaciones.
- Parámetros.
- Estructura.

Es fácil especificar y ajustar modelos. Pero hay preguntas cuyas respuestas no han quedado claras:

1. Construcción.
2. Evaluación.
3. Uso.

Programación probabilística.

Los aspectos del flujo de trabajo Bayesiano consideran ([2]):

1. Construcción iterativa de modelos.
2. Validación de modelo (computacional).
3. Entendimiento de modelo.
4. Evaluación de modelo.

4.1. Distinción importante

Inferencia no es lo mismo que análisis de datos o que un flujo de trabajo.

Inferencia (en el contexto bayesiano) es formular y calcular con probabilidades condicionales.

4.2. ¿Por qué necesitamos un flujo de trabajo?

- El cómputo puede ser complejo.
- Expandir nuestro entendimiento en aplicaciones.
- Entender la relación entre modelos.
- Distintos modelos pueden llegar a distintas conclusiones.

4.3. Proceso iterativo

- La gente de ML sabe que el proceso de construcción de un modelo es iterativo, ¿por qué no utilizarlo?

Una posible explicación puede encontrarse en [1]. El argumento es formal en cuanto a actualizar nuestras creencias como bayesianos. Sin embargo, con cuidado y un procedimiento científico puede resolver el asunto.

REFERENCIAS

- [1] A. Gelman and Y. Yao. Holes in Bayesian statistics. *Journal of Physics G: Nuclear and Particle Physics*, 48(1):014002, jan 2021. ISSN 0954-3899, 1361-6471. . 5

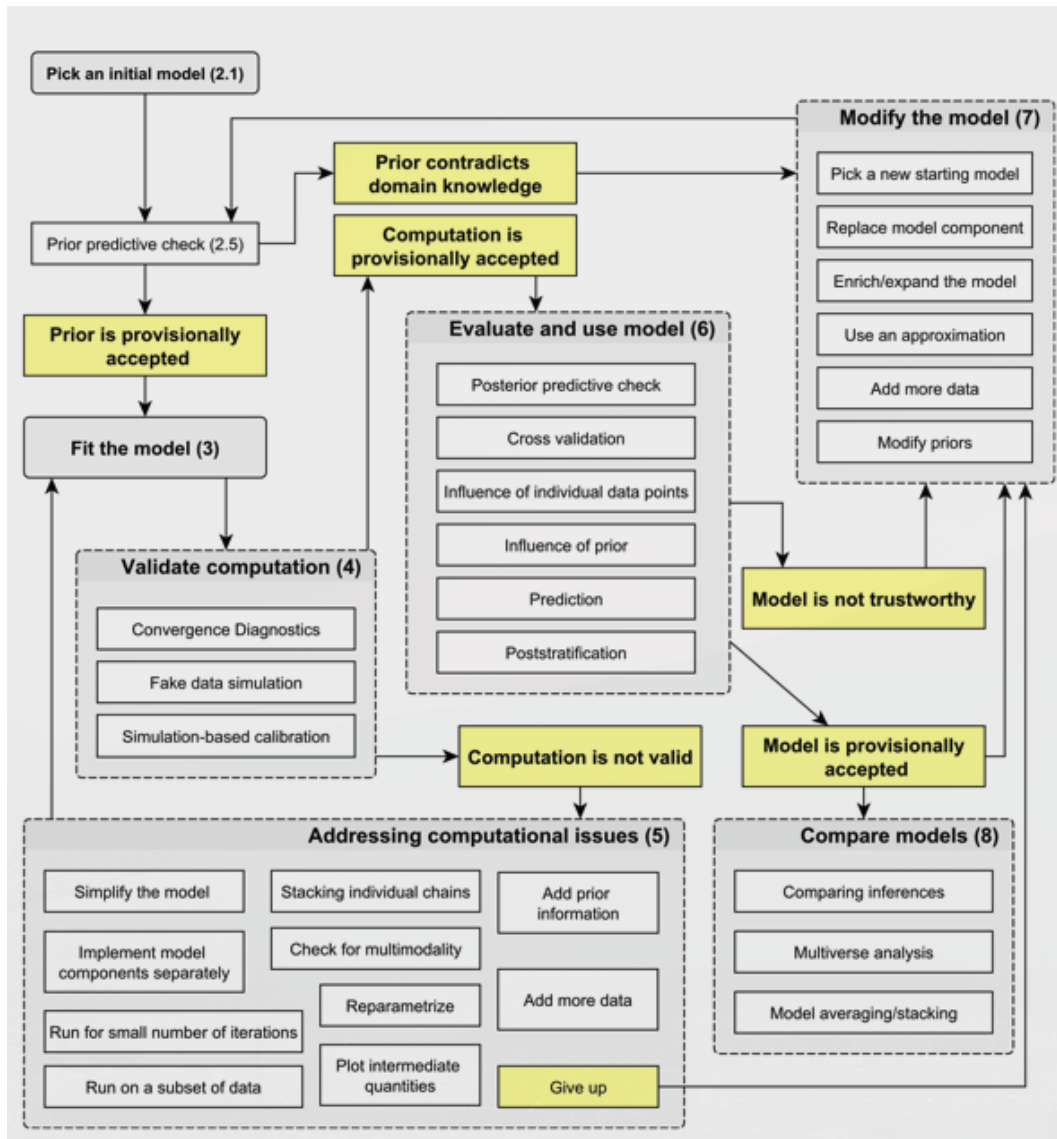


FIGURA 7. Tomado de [2].

- [2] A. Gelman, A. Vehtari, D. Simpson, C. C. Margossian, B. Carpenter, Y. Yao, L. Kennedy, J. Gabry, P.-C. Bürkner, and M. Modrák. Bayesian workflow. *arXiv preprint arXiv:2011.01808*, 2020. 1, 5, 6
- [3] A. Johnson, M. Ott, and M. Dogucu. *Bayes Rules! An Introduction to Applied Bayesian Modeling*. 2021. 2
- [4] J. Kruschke. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press, 2014. 2