

# Aplicaciones de aprendizaje estadístico a estimación temporal con programas de intervalo fijo

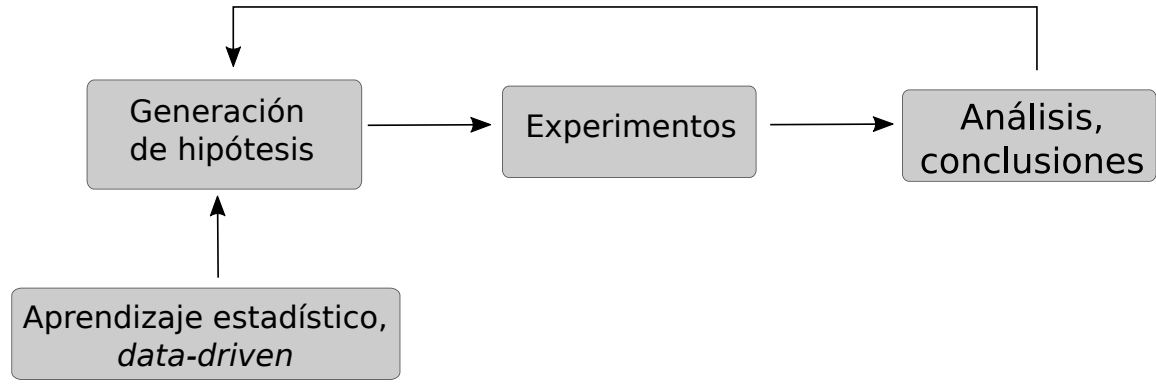
Emmanuel Alcalá & Jonathan Buriticá

Universidad de Guadalajara, CEIC

✉ jealcalat@gmail.com — 🌐 <https://github.com/jealcalat>

## Introducción

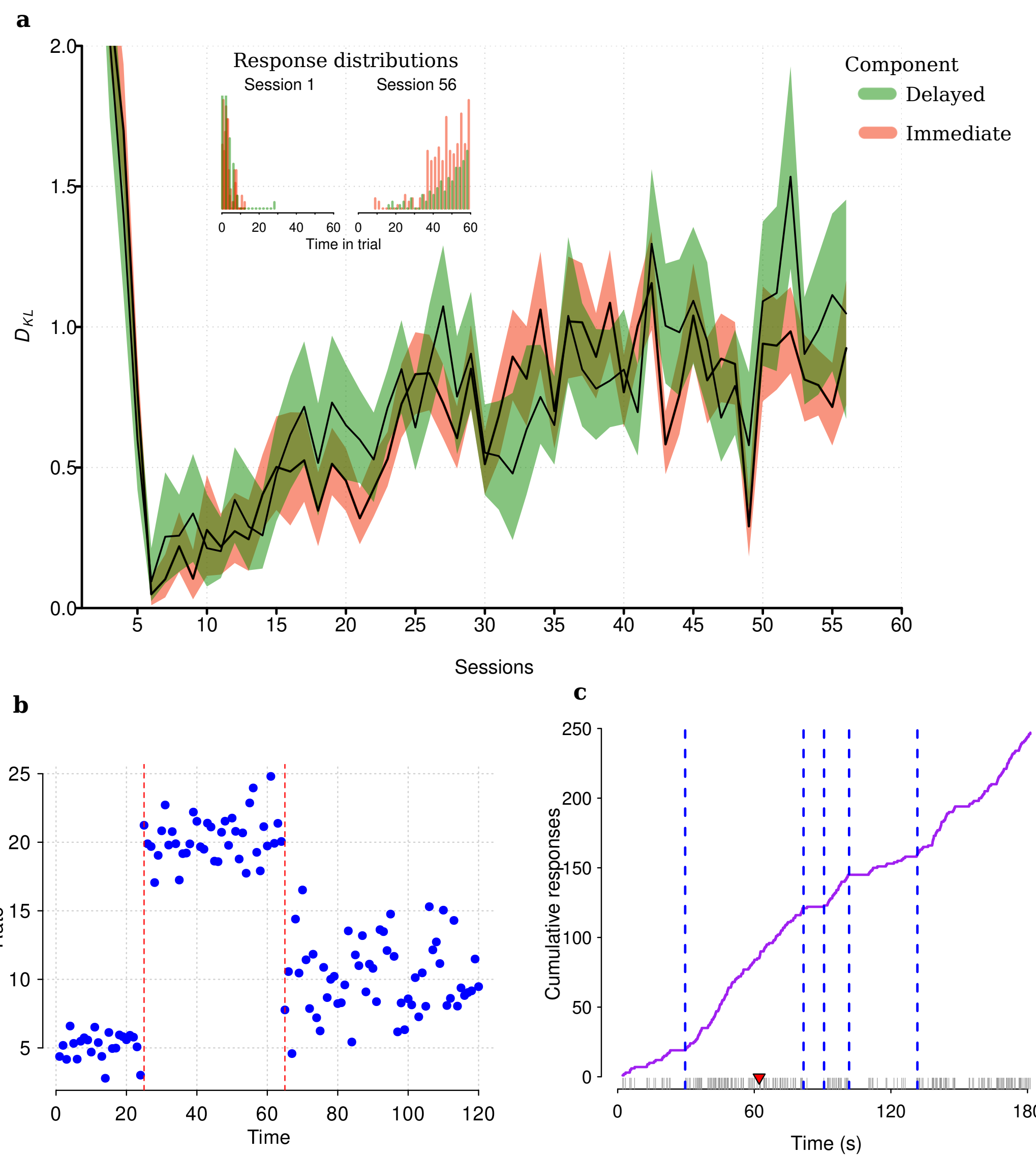
- El aprendizaje estadístico es un conjunto de métodos y algoritmos cuyo propósito es entender y/o predecir un conjunto dado de datos.
- Las técnicas de aprendizaje estadístico proporcionan una forma libre de hipótesis para analizar datos (Valleta et al., 2017).
- A menudo no se requiere de asumir una forma funcional de variables ni distribucionales de los datos.
- Suelen dar más peso a los datos que a las hipótesis (*data-driven (DD) vs hypothesis-driven (HD)*), permitiendo explorar características de los datos que probablemente no se hubiesen explorado con un enfoque dirigido a hipótesis (*hypothesis-free*).
- El enfoque *DD* es complementario al *HD* (ver diagrama).
- Se presentan algunas aplicaciones con datos de experimentos de estimación temporal: estimación de densidad por kernel, regresión no lineal con *bootstrap* no paramétrico y estimación de puntos de cambio usando árboles de regresión.



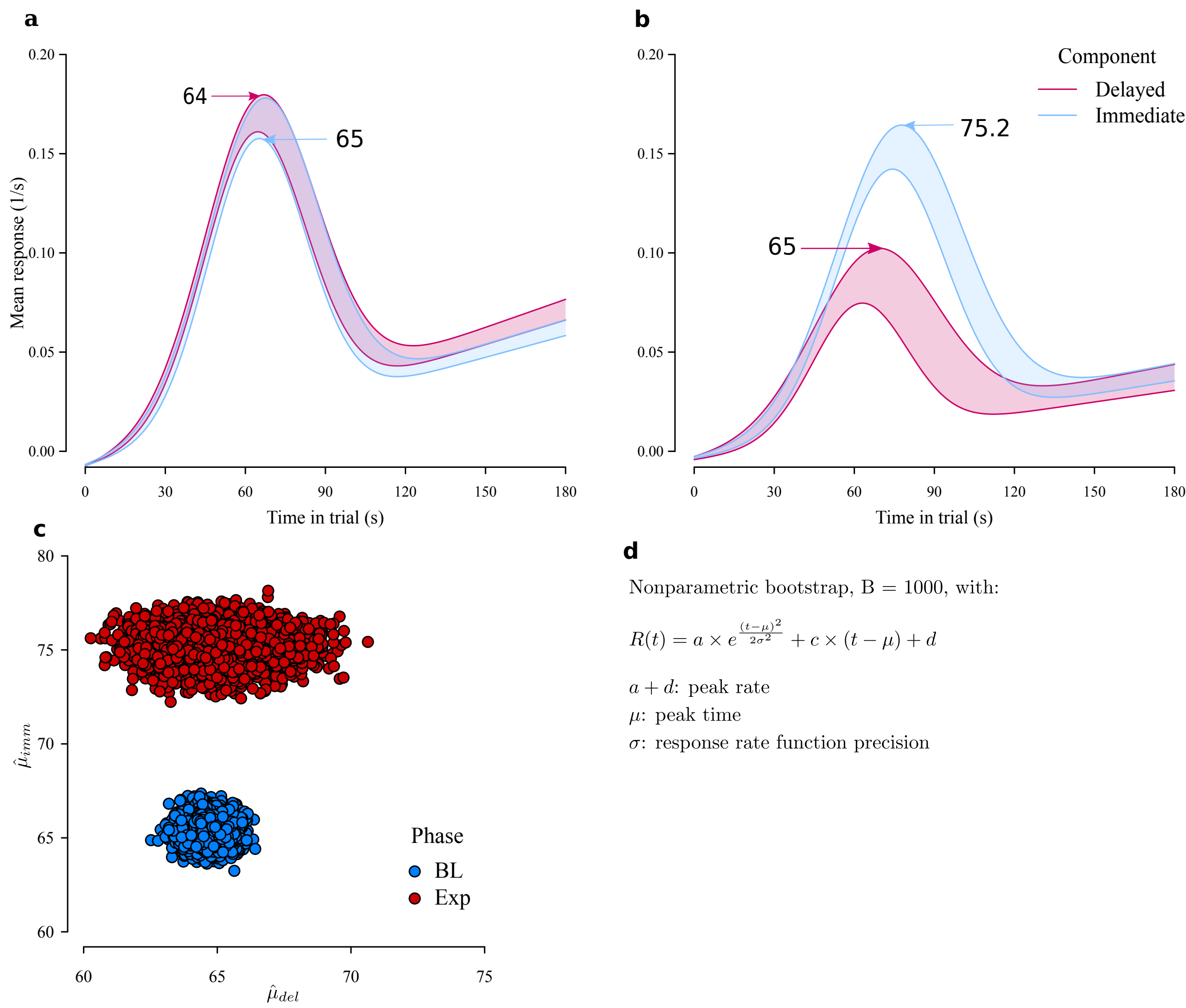
## Método

- Datos ejemplares de un experimento de procedimiento de pico con un programa múltiple IF 60 IF 60 (Buriticá & Alcalá, 2019). En la primera fase se entrenaron IFs de forma separada, con 1 ensayo de pico por cada 10 componentes. En la segunda fase, un componente (Inmediato) se acopló al otro (Demorado), donde el componente demorado se convirtió en *tandem* IF 54 s TF 6s.
- **Análisis:**
  - Como métrica de discriminación de ensayos IF, se estimó  $D_{KL}(P||Q) = \int_0^{60} P(x) \log_2 \frac{P(x)}{Q(x)} dx$ , donde  $Q(x) \sim U(0, 60)$  y  $P(x)$  se obtuvo usando estimación de densidad por kernel. La distribución  $U(\cdot)$  dado que se asume que el animal responde a tasa constante. La integral se computó de forma numérica. Esto se hizo con los  $t$  de cada respuesta por ensayo por componente.  
 $Q(x) = 0 \iff P(x) = 0$
  - *Bootstrap* no paramétrico de datos de ensayos de pico, usando una función gaussiana con rampa (ver Figura 2). Se usaron 1000 remuestras, por cada una de ellas se obtuvieron los coeficientes de la función para construir intervalos de confianza al 95 %.
  - Estimación de puntos de cambio en ensayos individuales de pico, usando árboles de regresión (algoritmo CART).

## Resultados



**Figura 1:** **a:**  $D_{KL}(P||Q)$  de ensayos IF. En las primeras sesiones los sujetos responden más al inicio del ensayo; aproximadamente en la sesión 6, su distribución es indistinguible de una uniforme, a partir de lo cual comienza a diverger (ver insertos). **b:** Ejemplo de puntos de cambio identificados con árboles de regresión, usando datos simulados de distribuciones normales. **c:** registro acumulativo (línea morada) y estimación de puntos de cambio de un ensayo de pico de un sujeto típico, usando el algoritmo CART. La partición recursiva identificó más de 2 puntos de cambio.



**Figura 2:** **a** y **b:** funciones de respuesta obtenidas con los parámetros de la función gaussiana, usando los cuantiles 0.025 y 0.975. Las flechas muestran el tiempo de pico promedio por función. **c:** promedios estimados del tiempo de pico. Se muestran los valores estimados del componente inmediato como función del demorado. Nótese que el valor promedio estimado parece cambiar poco para el componente demorado entre fases, pero su variabilidad es mucho mayor durante la fase experimental (rojo). **d:** Función gaussiana con rampa usada para la regresión no lineal.

## Conclusiones

1. La divergencia  $D_{KL}$  muestra sesión a sesión cuánto se aleja la distribución de respuestas con respecto de una distribución uniforme (que asume que los sujetos responden a tasa constante durante el ensayo IF).
2. En la Figura 2 **c** se obtienen más de 2 puntos de cambio, lo que sugiere que la asunción de tasa alta - baja - alta podría resultar de forzar los datos a ajustarse a esta asunción. Es común que los ensayos que no se adecúan a esta asunción, bajo algún estadístico de error, se descarten (e.g., Balci et al., 2009; Church et al., 1994).
3. El aprendizaje estadístico ofrece diversos métodos para explorar exhaustivamente un conjunto de datos, no solo algunas características.
4. Si bien el enfoque basado en modelos (HD) ayuda a evaluar teorías (e.g., hipótesis bajo un modelo matemático plausible), el enfoque dirigido por datos (DD) permite explorar estructuras ocultas o características de los datos que no son aparentes bajo el enfoque HD. Esto, a su vez, podría ayudar a generar nuevas hipótesis basadas en dichas características (e.g., más de dos puntos de cambio en ensayos de pico).
5. La potencia de las computadoras modernas también permite evaluar modelos o hipótesis sin la necesidad explícita de asunciones distribucionales (e.g., *bootstrap* para construir intervalos de confianza no basados en teorías asintóticas).

## Referencias

Balci, F., Gallistel, C. R., Allen, B. D., Frank, K. M., Gibson, J. M., & Brunner, D. (2009). Acquisition of peak responding: what is learned?. *Behavioural Processes*, 80(1), 67-75.

Buriticá, J., & Alcalá, E. (2019). Increased generalization in a peak procedure after delayed reinforcement. *Behavioural Processes*, 169, 103978. doi: 10.1016/j.beproc.2019.103978

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, 20(2), 135.

Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124, 203-220. doi: 10.1016/j.anbehav.2016.12.005

Vu, M. A. T., Adahi, T., Ba, D., Buzsáki, G., Carlson, D., Heller, K., ... & Mayberg, H. S. (2018). A shared vision for machine learning in neuroscience. *Journal of Neuroscience*, 38(7), 1601-1607. doi: 10.1523/JNEUROSCI.0508-17.2018.

## Agradecimientos

A los generosos e increíbles gurús de Stackoverflow.

